



T.C.
NECMETTİN ERBAKAN ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ



SOSYAL MEDYA MESAJLARINDA VERİ
MADENCİLİĞİ İLE BİLGİ KEŞFİ ve GÖRSEL
ANALİTİK ORTAMDA SUNULMASI:
COVID-19 TWEET VERİ SETİ ÖRNEĞİ

Burak ÇAĞLAR

DOKTORA TEZİ

Harita Mühendisliği Anabilim Dalı

Aralık-2023
KONYA
Her Hakkı Saklıdır

TEZ KABUL VE ONAYI

Burak ÇAĞLAR tarafından hazırlanan “Sosyal medya mesajlarında veri madenciliği ile bilgi keşfi ve görsel analitik ortamda sunulması: Covid-19 tweet veri seti örneği” adlı tez çalışması 28/12/2023 tarihinde aşağıdaki jüri tarafından oy birliği ile Necmettin Erbakan Üniversitesi Fen Bilimleri Enstitüsü Harita Mühendisliği Anabilim Dalı’nda DOKTORA TEZİ olarak kabul edilmiştir.

Jüri Üyeleri

İmza

Başkan

Prof. Dr. İ. Öztuğ BİLDİRİCİ

.....

Danışman

Doç. Dr. Hüseyin Zahit SELVİ

.....

Üye

Prof. Dr. İbrahim YILMAZ

.....

Üye

Prof. Dr. S. Savaş DURDURAN

.....

Üye

Doç. Dr. İlkay BUĞDAYCI

.....

Fen Bilimleri Enstitüsü Yönetim Kurulu’nun/.../20.. gün ve sayılı kararıyla onaylanmıştır.

Prof. Dr. Şerife Yurdağül KUMCU
FBE Müdürü

TEZ BİLDİRİMİ

Bu tezdeki bütün bilgilerin etik davranış ve akademik kurallar çerçevesinde elde edildiğini ve tez yazım kurallarına uygun olarak hazırlanan bu çalışmada bana ait olmayan her türlü ifade ve bilginin kaynağına eksiksiz atıf yapıldığını bildiririm.

DECLARATION PAGE

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

İmza

Burak ÇAĞLAR

Tarih: Aralık / 2023

ÖZET

DOKTORA TEZİ

SOSYAL MEDYA MESAJLARINDA VERİ MADENCİLİĞİ İLE BİLGİ KEŞFİ ve GÖRSEL ANALİTİK ORTAMDA SUNULMASI: COVID-19 TWEET VERİ SETİ ÖRNEĞİ

Burak ÇAĞLAR

Necmettin Erbakan Üniversitesi Fen Bilimleri Enstitüsü
Harita Mühendisliği Anabilim Dalı

Danışman: Doç. Dr. Hüseyin Zahit SELVİ

2023, 269 Sayfa

Jüri

Prof. Dr. İ. Öztuğ BİLDİRİCİ

Prof. Dr. İbrahim YILMAZ

Prof. Dr. S. Savaş DURDURAN

Doç. Dr. Hüseyin Zahit SELVİ

Doç. Dr. İlkey BUĞDAYCI

Günümüzde, sosyal medya iletişim ve kişisel ilgi alanlarını oluşturmak için en etkili araçtır. Akıllı cihazların ve sosyal platformların yaygın kullanımıyla, sosyal medyanın önemi artmış ve birçok kişi tarafından tercih edilir hale gelmiştir. Sosyal medya, insanlar arasındaki etkileşimi destekleyerek, yapılandırılmamış, ayrıntılı ve büyük ölçekli dijital verilerin kaynağı haline gelmiştir.

COVID-19 pandemisi, çağımızın en önemli küresel krizi olarak kabul edilmektedir. Bu pandemi, başlangıçta bir sağlık krizi olarak başlamış olsa da hızla tüm dünyaya yayılan benzersiz bir sosyo-ekonomik ve çevresel krize dönüşmüştür. Pandemi yönetiminin büyük bir kısmını mekânsal gelişimin takibi oluşturmaktadır. Salgınların kontrol altına alınması, temas takibi ve yayılımın önlenmesi için etkili bir toplum müdahalesinin sağlanabilmesi adına mekânsal bilgilere olan ihtiyaç oldukça büyüktür. Kriz yönetimi için kaliteli veri ve istatistiksel sonuçlar büyük önem taşımaktadır. Mekânsal görsel analitik, kullanıcıların mekânsal verileri kullanarak örüntüleri tespit etmeleri ve gelecekteki sonuçları tahmin etmeleri için kartografya, veri yönetimini, veri madenciliğini, arayüz tasarımı ve bilişsel bilim vb. farklı disiplinlerden yararlanmaktadır.

Bu tez çalışmasının temel amacı, bilgi biliminin ve kartografyanın etkileşimli yapısını yansıtan Web tabanlı mekânsal görsel analitik uygulama örneği geliştirilmesidir. Bu doğrultuda, eski adıyla Twitter yeni adıyla X sosyal medya platformunda COVID-19 pandemi döneminde aşılarda hakkında oluşturulan tweet veri seti kullanılmıştır. Tweet verilerinin mekânsal gelişiminin izlenmesi amacıyla konum bilgisi çıkarımı gerçekleştirilmiştir. Daha sonra, tweet veri setinde yer alan saklı örüntülerin keşfi gerçekleştirilmiştir. Veri setinde keşfi yapılan yeni bilgiler kullanılarak, K-ortalamlar ve Ward yöntemiyle kümeleme analizi gerçekleştirilmiştir. Mesajların oluşturulduğu ülkelerin benzerlikleri ve farklılıkları ortaya konmuştur. Tweet veri setinde yer alan mesajların, veri setinden keşfedilen yeni bilgilerin, kümeleme analizi sonucu elde edilen bilgiler ile üretilen tematik haritaların birbirleriyle etkileşim içinde olduğu, araştırmacının belirlediği sorgulamalar kapsamında filtrele araçlarının yer aldığı, veri setindeki mesajların mekânsal - zamansal gelişiminin aynı ekran üzerinden takip edildiği ve araştırmaya ait sonuçların farklı grafiksel araçlarla kullanıcıya aktarıldığı Web tabanlı interaktif bir uygulama geliştirilmiştir.

Anahtar Kelimeler: Büyük veri, Kartografya, Kümeleme Analizi, Mekânsal Görsel Analitik, Sosyal Medya, Tematik Harita

ABSTRACT

Ph.D THESIS

INFORMATION DISCOVERY WITH DATA MINING IN SOCIAL MEDIA MESSAGES and PRESENTATION IN VISUAL ANALYTICS: COVID-19 TWEET DATASET EXAMPLE

Burak AĐLAR

THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCE OF
NECMETTİN ERBAKAN UNIVERSITY
DOCTOR OF PHILOSOPHY
IN GEOMATIC ENGINEERING

Advisor: Assoc. Prof. Hüseyin Zahit SELVİ

2023, 269 Pages

Jury

Prof. Dr. İ. ÖztuĐ BİLDİRİCİ
Prof. Dr. İbrahim YILMAZ
Prof. Dr. S. SavaĐ DURDURAN
Assoc. Prof. Hüseyin Zahit SELVİ
Assoc. Prof. İlkey BUĐDAYCI

Nowadays, social media is the most effective tool for communication and building personal interests. With the widespread use of smart devices and social platforms, the importance of social media has increased and has become preferred by many people. Social media has become a source of Big data, unstructured, granular and large-scale digital data, supporting interaction between people.

The COVID-19 pandemic is recognized as the most significant global crisis of our time. Although this pandemic initially started as a health crisis, it has rapidly evolved into a unique socio-economic and environmental crisis that has spread across the globe. Tracking spatial development constitutes a large part of pandemic management. There is a great need for spatial information to ensure an effective community response to control outbreaks, contact tracing and prevention of spread. Quality data and statistical results are essential for crisis management. Geovisual analytics utilizes different disciplines such as cartography, data management, data mining, interface design and cognitive science to help users detect patterns and predict future outcomes using spatial data.

The main objective of this thesis is to develop a Web-based spatial visual analytics application that reflects the interactive nature of information science and cartography. In this context, the tweet data set created on the social media platform, formerly Twitter and now X, about vaccines during the COVID-19 pandemic period was used. Accordingly, the tweet data series created on the social media platform Twitter about vaccines during the COVID-19 pandemic period was used. In order to track the spatial evolution of the tweet data, location information extraction was performed. After that, the hidden patterns in the tweet data set were discovered. Using the new information discovered in the dataset, clustering analysis was performed using K-means and Ward methods. The similarities and differences of the countries where the messages were created were revealed. A Web-based interactive application has been developed in which the messages in the Tweet dataset, the new information discovered from the dataset, the information obtained as a result of the clustering analysis and the thematic maps produced interact with each other, filtering tools are included with in the scope of the queries determined by the researcher, the spatial - temporal development of the messages in the dataset is followed on the same screen, and the results of the research are transferred to the user with different graphical tools.

Keywords: Big Data, Cartography, Cluster Analysis, Geovisual Analytics, Social Media, Thematic Map

ÖNSÖZ

Lisans öğrenimimden itibaren, akademik yolculukta her daim rehberim ve mesleki kazanımlarımda üzerimde çok fazla emeği olan tez danışmanım Doç. Dr. Hüseyin Zahit SELVİ'ye teşekkür ederim. Yine lisans öğreniminden bugüne kadar bilgi birikimleri ve tecrübeleriyle yol gösteren Prof. Dr. İ. Öztuğ BİLDİRİCİ ve Doç. Dr. İlkay BUĞDAYCI'ya teşekkür ederim.

Doktora çalışmam sırasında benden desteğini, zamanını ve sabrını esirgemeyen, beni anlayışla karşılayan sevgili eşim Nuriye ÇAĞLAR'a, sevgileriyle bana güç veren kızlarım Feyza ve İpek Ravza'ya teşekkürlerimi sunarım.

Burak ÇAĞLAR
KONYA-2023



İÇİNDEKİLER

ÖZET	iv
ABSTRACT	v
ÖNSÖZ	vi
İÇİNDEKİLER	vii
SİMGELER VE KISALTMALAR	x
ŞEKİL LİSTESİ	xi
ÇİZELGE LİSTESİ	xv
1. GİRİŞ	1
1.1. Tezin Amacı	3
1.2. Tezin Organizasyonu	4
2. KAYNAK ARAŞTIRMASI	5
2.1. Büyük Veri	5
2.2. Mekânsal Büyük Veri	7
2.3. Veri Madenciliği	11
2.3.1. Tanımlayıcı analiz yöntemleri	11
2.3.1.1. Kümeleme analizi.....	12
2.3.1.2. Birliktelik kuralları ve ilişki analizi	12
2.3.2. Tahmin edici analiz yöntemleri	13
2.3.2.1. Sınıflandırma ve regresyon yöntemi	13
2.4. Veri Tabanlarında Bilgi Keşfi	13
2.5. Görselleştirme.....	14
2.6. Coğrafi Görselleştirme	18
2.7. Görsel Analitik	19
2.7.1. Görsel analitik alanındaki araştırmalar	20
2.7.2. Görsel analitik uygulama alanları	21
2.7.3. Görsel analitik süreci	24
2.8. Mekânsal Görsel Analitik	26
2.8.1. Analitik akıl yürütme	26
2.8.2. İnteraktif coğrafi görselleştirme.....	27
2.8.3. Hesaplamalı yöntemler	28
2.8.4. Teknik gelişmeler.....	28
2.8.5. Mekânsal görsel analitikte kullanıcı merkezli tasarım.....	29
2.9. Sosyal Medya	29
2.10. Eski Adıyla Twitter, Yeni Adıyla X.....	33
2.10.1. Mekânsal özellikler ve göstergeler	35
2.11. Araştırma Konusu Kapsamında Daha Önceden Gerçekleştirilen Çalışmalar ..	37
3. MATERYAL ve YÖNTEM	43

3.1. Kümeleme Analizi	43
3.1.1. Ward yöntemi.....	43
3.1.2. K-ortalamlar (k-means) yöntemi.....	44
3.1.3. Küme geçerliliği teknikleri	45
3.1.3.1. Kofenetik (cophonetic) korelasyon katsayısı	45
3.1.3.2. Silhouette katsayısı.....	45
3.1.4. KNIME	46
3.1.5. RStudio	46
3.2. Ayırık Küresel Grid Sistemleri.....	46
3.2.1. Uber H3.....	49
3.2.1.1. H3 indeksleme yapısı	54
3.2.1.2. H3 indekslemesinde kullanılan fonksiyonlar	57
3.2.1.3. Uber H3'ün mekânsal analizlerde kullanılması	59
3.3. Web Tabanlı İnteraktif CBS Uygulaması Geliştirmede Kullanılacak Yazılımlar	62
3.3.1. Arcgis World Geocoding	63
3.3.2. Arcgis Web Appbuilder	64
3.3.3. Arcgis Dashboards	65
3.3.4. Arcgis Hub	66
3.3.5. Ülkelere ait idari sınır verileri.....	66
4. UYGULAMA.....	67
4.1. Veri Setinin Elde Edilmesi	67
4.2. Veri Setinin İncelenmesi	68
4.3. Veri Setinden Coğrafi Bilgi Çıkarımı ve Coğrafi İşaretlemenin Gerçekleştirilmesi.....	70
4.4. Veri Setindeki Gizli Verilerin Keşfi.....	75
4.4.1. COVID-19 aşılıları hakkındaki tweetler hangi ülkelerden oluşturulmuştur?	76
4.4.2. COVID-19 aşılıları hakkındaki tweetler hangi kullanıcılar tarafından oluşturulmuştur?	77
4.4.3. Ülkelerin COVID-19 aşılıları hakkındaki tweet atan kullanıcı sayıları kaçtır?	78
4.4.4. Pfizer Biontech aşısı hakkında oluşturulan tweetler hangileridir?	80
4.4.5. Sinopharm aşısı hakkında oluşturulan tweetler hangileridir?.....	80
4.4.6. Sinovac aşısı hakkındaki tweetler hangileridir?	81
4.4.7. Moderna aşısı hakkındaki tweetler hangileridir?.....	81
4.4.8. Oxford AstraZeneca aşısı hakkındaki tweetler hangileridir?.....	81
4.4.9. Covaxin aşısı hakkındaki tweetler hangileridir?.....	82
4.4.10. Sputnik V aşısı hakkındaki tweetler hangileridir?	82
4.4.11. Johnson&Johnson aşısı hakkındaki tweetler hangileridir?	82
4.4.12. Covid kelimesi hakkındaki tweetler hangileridir?	83
4.4.13. Vaccine (aşı) kelimesi hakkındaki tweetler hangileridir?.....	83
4.5. Veri setinde Keşfi Yapılan Bilgilerin Tematik Haritalar ile Sunumu	86
4.6. Nokta Geometrisinden Alan Geometrisine Geçiş	87
4.6.1. Convex hull (dışbükey kabuk).....	87
4.6.2. Uber H3: Hiyerarşik altıgen hücrelerle mekânsal indeksleme	90
4.7. Dünya Geneline Uygulanmakta Olan COVID-19 Aşılılarının İncelenmesi ..	103
4.8. Kümeleme Analizi	107

4.8.1. Ülkelerin X kullanıcı sayıları ve tweet sayılarına göre ward yöntemiyle hiyerarşik kümeleneşmesi	108
4.8.2. Ülkelerin X kullanıcı sayıları ve tweet sayılarına göre k-ortalama (k-means) yöntemiyle hiyerarşik olmayan kümeleme	109
4.8.3. Ülkelerin oluşturulan toplam tweet sayısı, covid kelimesi içeren tweet sayısı ve vaccine kelimesi içeren tweet sayıları göre ward yöntemiyle kümeleneşmesi	112
4.8.4. Ülkeleri oluşturulan toplam tweet sayısı, covid kelimesi içeren tweet sayısı ve vaccine kelimesi içeren tweet sayıları göre k-ortalama (k-means) yöntemiyle hiyerarşik olmayan kümeleme	114
4.9. Web Tabanlı İnteraktif Uygulama Geliştirilmesi	116
4.9.1. Coğrafi verilerin WFS servisi kullanılarak bulut sunucu üzerinde paylaşılması.....	117
4.9.2. ArcGIS Online Map Viewer Classic üzerinde Web haritalarının oluşturulması.....	119
4.9.3. ArcGIS Web AppBuilder ile interaktif veri editörü uygulaması geliştirme	126
4.9.4. Covid-19 Aşılılarıyla ilgili tweetleri gösteren bilgi panosunun oluşturulması	131
4.9.5. Ülkelerin X kullanıcı sayıları ve tweet sayılarına göre kümeleme sonuçlarını gösteren bilgi panosunun oluşturulması.....	141
4.9.6. Ülkelerin tweet sayısı, “covid” kelimesi içeren tweet sayısı ve “vaccine” kelimesi içeren tweet sayılarına göre kümeleme sonuçlarını gösteren bilgi panosunun oluşturulması	146
4.9.7. İnteraktif veri editörünün ve bilgi panolarının birleştirilerek web tabanlı olarak kullanıcıların erişimine açılması	150
5. ARAŞTIRMA BULGULARI ve DEĞERLENDİRME	156
6. SONUÇLAR ve ÖNERİLER	171
KAYNAKLAR	177
EKLER	188
ÖZGEÇMİŞ	254

SİMGELER VE KISALTMALAR

Kısaltmalar

API	Application Programming Interface
CCC	Cophenetic Correlation Coefficient
CGV	Coordinated Graph Visualization
D3	Data Drive Documents
DGGS	Discrete Global Grid Systems
ESS	Emergency Support Services
GAV	GeoAnalytics Visualization
GDB	Geodatabase
GeoPKKD	Geographic Privacy-aware Knowledge Discovery and Delivery
GISCO	Geographical information system of the Commission
GPS	Global Positioning System
HTML	HyperText Markup Language
IEEE	Institute of Electrical and Electronics Engineers
JSON	JavaScript Object Notation
KDD	Knowledge Discovery in Databases
KML	Keyhole Markup Language
NIH	National Institute Of Health
NSF	National Science Foundation
NVAC	National Visualization and Analytics Center
OASIS	Open Advanced System for disaster and emergency management
OGC	Open Geospatial Consortium
SIGKDD	Special Interest Group on Knowledge Discovery and Data Mining
VGI	Volunteered Geographic Information
VGTC	Visualization and Graphics Technical Community
WCS	Web Coverage Service
WFS	Web Feature Service
WHO	World Health Organization
WMS	Web Map Service
URL	Uniform Resource Locator

ŞEKİL LİSTESİ

Şekil 2.1	Büyük verinin özellikleri (Demchenko ve ark., 2013; Narasimhan ve ark., 2014).	6
Şekil 2.2	Hindistan'da mobil internet trafiği (Meeker, 2012)	8
Şekil 2.3	İnsanoğlu'nun karşılaşacağı 15 küresel zorluk (Shekhar, 2012).....	9
Şekil 2.4	Google Maps tarafından sunulan konum tabanlı hizmetler	10
Şekil 2.5	Veri tabanlarında bilgi keşfinin adımları (Han ve Kamber, 2006)	12
Şekil 2.6	Bilgi keşfi sürecinin aşamaları (Fayyad ve ark., 1996)	13
Şekil 2.7	Görselleştirme için referans modeli (Card ve ark., 1999).....	15
Şekil 2.8	İnternet tabanlı görselleştirme adımları (Upton ve ark., 1989; Haber ve McNabb, 1990)	16
Şekil 2.9	Yayıncı bir görsel oluşturur, görüntüleyici statik bir görüntüyü analiz eder (Wood ve ark., 1996)	17
Şekil 2.10	Yayıncı modeli oluşturur, görüntüleyici görüntü üretimini kontrol eder (Wood ve ark., 1996)	17
Şekil 2.11	Görüntüleyici görseli oluşturur (Wood ve ark., 1996).....	17
Şekil 2.12	Yayıncı, görüntüleyicinin kontrolü altında geometri modeli oluşturur (Wood ve ark., 1996)	18
Şekil 2.13	Kartografya küpü (MacEachren, 1994)	19
Şekil 2.14	Coordinated Graph Visualization (CGV) ile gerçekleştirilen iklim modelleme simülasyonu (Tominski ve ark., 2009).....	22
Şekil 2.15	Bir üniversite ağının SSH servisine dağılmış ağ saldırısının NFlowVis ile analizi (Mansmann ve ark., 2009).....	23
Şekil 2.16	Görsel analitik süreci (Keim ve ark., 2010).....	24
Şekil 2.17	GAV (GeoAnalytics visualization) araç seti (Robinson, 2017).....	27
Şekil 2.18	Sosyal medya kullanımına genel bakış (URL-3)	31
Şekil 2.19	Dünya'da en çok kullanılan sosyal medya platformları (URL-3)	31
Şekil 2.20	Sosyal medya kullanımının ana sebepleri (URL-3).....	32
Şekil 2.21	X (Eski adıyla Twitter) kullanım istatistikleri (URL-3)	34
Şekil 2.22	X kullanıcısının konum göstergeleri (Ajao ve ark., 2015).....	35
Şekil 2.23	Fortaleza'daki cinayetlerin haritası ile WikiCrimes ana arayüzüne genel bakış, nokta haritaları ve ana alanlarını açıklayan metinler (Furtado ve ark., 2010)	37
Şekil 2.24	SensePlace2 mekânsal analitik uygulaması kullanıcı ara yüzü (MacEachren ve ark., 2011)	38
Şekil 2.25	Etkileşimli tahmin edici görsel analiz ortamı arayüzü (Maciejewski ve ark., 2011)	38
Şekil 2.26	GCAM isimli Su ile ilgili iklim değişikliği senaryolarını araştıran Web tabanlı bir görselleştirme aracı (Luo ve ark., 2015).....	39
Şekil 2.27	Nijer Nehri havzasındaki su kıtlığını keşfetmek için oluşturulan interaktif görsel analiz aracı (Lei ve ark., 2015).....	39
Şekil 2.28	STempo mekânsal görsel analitik uygulaması kullanıcı ara yüzü (Robinson ve ark., 2016)	40
Şekil 2.29	Birleşik Krallık'taki hava ulaşım trafiğinin mekânsal ve zamansal görselleştirilmesi (Andrienko ve ark., 2017)	41
Şekil 3.1	k-Ortalama algoritması ile kümeleme (Han ve Kamber, 2006).....	44
Şekil 3.2	Gerçekte var olmayan temel katılar ve küresel gösterimleri: Soldan sağa tetrahedron, küp, oktahedron, dodecahedron, icosahedron (Lei ve ark., 2020).	47

Şekil 3.3	En yaygın uygulanan mozaikleme şemaları (Kmoch ve ark., 2022)	48
Şekil 3.4	H3'ün harita projeksiyonu için kullandığı yirmiyüzlü yüzey ve yirmiyüzlü yüzeye odaklanan gnomonik projeksiyon (URL-8).....	50
Şekil 3.5	Bir üçgenin komşularına, bir karenin komşularına ve bir altıgenin komşularına uzaklıkları (URL-8).....	50
Şekil 3.6	Tek bir yirmiyüzlü yüzünde grid oluşturmak için H3 kullanılmaktadır (URL-8).....	51
Şekil 3.7	R. Buckminster Fuller'in yirmiyüzlü yerleşimine ait dört görünüm (URL-17)	51
Şekil 3.8	H3, kullanıcının alanları daha küçük parçalara ve daha küçük altıgenlere ayırmasını sağlamaktadır (URL-8)	52
Şekil 3.9	Kaliforniya şehrinin yoğunlaştırılmış ve seyrekleştirilmiş hücre kümeleriyle temsil edilmesi (URL-11)	53
Şekil 3.10	"85283473ffffff" H3 indeksine sahip H3 altıgen hücresi ve bilgileri (URL-16)	54
Şekil 3.11	Hücreler için H indeksinin bit düzeni (URL-13)	55
Şekil 3.12	6 Farklı H3 yönlendirilmiş kenar indekslerine sahip H3 altıgen hücresi (URL-14).....	55
Şekil 3.13	Yönlendirilmiş kenarlar için H3İndeksinin bit düzeni (URL-14).....	56
Şekil 3.14	6 Farklı H3 köşe noktası indeksine sahip H3 altıgen hücresi (URL-15)	56
Şekil 3.15	Köşe noktaları için H3 indeksinin bit düzeni (URL-15).....	57
Şekil 3.16	Bir şehirdeki arabaların harita üzerinde gösterimi, harita üzerinde H3 altıgen hücreleriyle birlikte gösterimi ve H3 altıgen hücreleri içerisinde kalan arabaların toplam sayılarını renkler ve renk tonlarıyla gösterimi (URL-8).....	59
Şekil 3.17	New York şehrinin (Manhattan) posta kodu haritası (URL-8).....	60
Şekil 3.18	San Francisco şehri kaplayan rastgele oluşturulmuş altıgen kümeler (URL-8).....	61
Şekil 3.19	Bir veri setinin noktalar, ilçe sınırları ve H3 altıgenleriyle gösterimi (URL-18)	62
Şekil 3.20	Metinsel bir ifadenin coğrafi kodlaması	63
Şekil 3.21	Konumu bilinen bir noktanın ters coğrafi kodlaması	64
Şekil 3.22	ArcGIS Web AppBuilder kullanıcı arayüzü	64
Şekil 3.23	ArcGIS Dashboards kullanıcı arayüzü.....	65
Şekil 3.24	ArcGIS Hub Web sitesi oluşturma kullanıcı arayüzü.....	66
Şekil 4.1	COVID-19 All Vaccines Tweets” veri seti "user_location” öznitelik alanı veri detayı.....	74
Şekil 4.2	"user_location" öznitelik bilgisi "Turkey-Israel" olan verisi örneği.....	75
Şekil 4.3	Ülkeler ve “kullanıcı bazlı atılan tweet sayıları” veri tabloları arasında birden çoka (1:N) ilişki	78
Şekil 4.4	Nokta kümesi (solda), dışbükey kabuk alanı (ortada) ve dışbükey kabuk (sağda) (URL-30).....	88
Şekil 4.5	Hindistan ve ABD’de oluşturulan tweetlerin dışbükey kabukları	88
Şekil 4.6	Farklı ülkelerden oluşturulan tweetlerin dışbükey kabukları.....	89
Şekil 4.7	Fransa ve bağlı topraklarından oluşturulan tweetlerin dışbükey kabukları .	89
Şekil 4.8	Çözünürlük = 0 seviyesinde oluşturulan 122 adet H3 altıgen hücreleri	91
Şekil 4.9	Türkiye'nin sekiz farklı seviyedeki H3 altıgen hücrelerle indekslenmesi...	92
Şekil 4.10	“tweetcount” ve “usercount” değişkenlerinin dendrogram üzerinde gösterimi	108
Şekil 4.11	Knime platformunda k-ortalama kümeleme analizi	110

Şekil 4.12	"Line Plot" aracıyla oluşturulan Silhouette katsayı grafiği.....	111
Şekil 4.13	Serpilme grafiği üzerinde kümelerin ve küme elemanlarının gösterimi....	112
Şekil 4.14	"tweetcount", "covidcount" ve "vaccinecount" değişkenlerinin dendrogram üzerinde gösterimi.....	113
Şekil 4.15	Toplam tweet sayısı ve Covid kelimesi içeren tweet sayısına göre kümeleri gösteren serpilme grafiği.....	114
Şekil 4.16	Toplam tweet sayısı ve vaccine kelimesi içeren tweet sayısına göre kümeleri gösteren serpilme grafiği.....	115
Şekil 4.17	Covid kelimesi içeren tweet sayısı ve vaccine kelimesi içeren tweet sayısına göre kümeleri gösteren serpilme grafiği	115
Şekil 4.18	Feature (Detay) servisi oluşturma süreci (URL-32)	117
Şekil 4.19	Coğrafi verilerin ArcGIS Pro üzerinden detay servisi olarak yayınlanması	117
Şekil 4.20	ArcGIS Online üzerinden coğrafi verilere erişimin sağlanması	118
Şekil 4.21	ArcGIS Online genel bakış sekmesi	118
Şekil 4.22	Map Viewer Classic arayüzüne coğrafi veri katmanlarının entegre edilmesi	120
Şekil 4.23	Coğrafi veri katmanlarının Map Viewer arayüzünde harita üzerinde görünümü	121
Şekil 4.24	Coğrafi veri katmanlarının görünür ölçek aralıklarının ayarlanması.....	121
Şekil 4.25	Coğrafi verilerin işaretleştirme ve renk değişikliklerinin gerçekleştirilmesi	122
Şekil 4.26	Coğrafi veriler için etiket oluşturulması	124
Şekil 4.27	Coğrafi veriler için açılır pencere (pop-up) oluşturma	125
Şekil 4.28	Web haritası üzerinde yer işareti oluşturma.....	125
Şekil 4.29	ArcGIS Online üzerinden Web AppBuilder uygulamasına giriş.....	127
Şekil 4.30	Web AppBuilder uygulama geliştirme platformu.....	127
Şekil 4.31	Web AppBuilder uygulaması gereçlerinin bulunduğu arayüz.....	128
Şekil 4.32	"Covid-19 Vaccines Data Editor" Web uygulaması arayüzü	130
Şekil 4.33	ArcGIS Online üzerinden Dashboards uygulamasına giriş	132
Şekil 4.34	Dashboards uygulama geliştirme platformu	132
Şekil 4.35	Dashboards uygulamasında Web haritası ayarlarının gerçekleştirildiği arayüz.....	133
Şekil 4.36	Dashboards uygulamasında öğelerin eklenileceği alanlar gösteren arayüz.....	133
Şekil 4.37	COVID-19 aşılılarıyla ilgili tweetleri gösteren bilgi panosu taslak şeması.	135
Şekil 4.38	Dashboards uygulamasıyla geliştirilen COVID-19 aşılılarıyla ilgili tweetleri gösteren bilgi panosu	136
Şekil 4.39	Bilgi paneli uygulamasında bulunan sorgu aracı örnekleri.....	138
Şekil 4.40	Tweet özneliklerinin gösterildiği bilgi penceresi	138
Şekil 4.41	Yatay sütun grafik üzerinde tweet sayıları.....	139
Şekil 4.42	Tweet verilerinin zamansal (1 ay aralıklı) dağılımlarının dikey sütun grafik üzerinde gösterimi.....	139
Şekil 4.43	Pfizer Biontech aşısı hakkındaki bilgi penceresi	140
Şekil 4.44	Sinopharm aşısı hakkındaki bilgi penceresi.....	140
Şekil 4.45	Sinovac aşısı hakkındaki bilgi penceresi	140
Şekil 4.46	Moderna aşısı hakkındaki bilgi penceresi	140
Şekil 4.47	Oxford AstraZeneca aşısı hakkındaki bilgi penceresi	140
Şekil 4.48	Covaxin aşısı hakkındaki bilgi penceresi.....	140
Şekil 4.49	"Covid" kelimesi içeren tweetler hakkındaki bilgi penceresi	141
Şekil 4.50	"Vaccine" kelimesi içeren tweetler hakkındaki bilgi penceresi.....	141

Şekil 4.51	Ülkelerin X kullanıcı sayıları ve tweet sayılarına göre kümeleme sonuçlarını gösteren bilgi panosu taslak şeması	142
Şekil 4.52	Ülkelerin X kullanıcı sayıları ve tweet sayılarına göre kümeleme sonuçlarını gösteren bilgi panosu	143
Şekil 4.53	K-ortalamlar yöntemiyle elde edilen kümeleme sonuçlarının bilgi paneli üzerinden görünümü	144
Şekil 4.54	Ülke bazlı K-ortalamlar yöntemiyle elde edilen kümeleme sonuçlarının bilgi paneli üzerinden görünümü	144
Şekil 4.55	Ward yöntemiyle elde edilen kümeleme sonuçlarının bilgi paneli üzerinden görünümü	145
Şekil 4.56	Ülke bazlı Ward yöntemiyle elde edilen kümeleme sonuçlarının bilgi paneli üzerinden görünümü	146
Şekil 4.57	Ülkelerin toplam tweet sayıları, "Covid" kelimesi içeren tweet sayıları ve "Vaccine" kelimesi içeren tweet sayıları göre kümeleme sonuçlarını gösteren bilgi panosu	148
Şekil 4.58	Dikey grafik üzerinden küme seçimi gerçekleştirildiğinde ülkeler listesi, Web haritası, çizgisel grafik ve gösterge sayaçlarının etkileşimi	149
Şekil 4.59	Web haritası üzerinden bir ülke seçimi gerçekleştirildiğinde ülkeler listesi, Web haritası, çizgisel grafik ve gösterge sayaçlarının etkileşimi	149
Şekil 4.60	Web haritası üzerinden birden fazla ülke seçimi gerçekleştirildiğinde ülkeler listesi, Web haritası, çizgisel grafik ve gösterge sayaçlarının etkileşimi... ..	150
Şekil 4.61	Ülkelerin tweet sayıları, "Covid" kelimesi içeren tweet sayıları ve "Vaccine" kelimesi içeren tweet sayılarının ward yöntemiyle kümeleme analizi sonuçlarını gösteren bilgi panosu	151
Şekil 4.62	Covid-19 aşılı veri editörü Web sayfası görünümü	152
Şekil 4.63	Covid-19 aşılı bilgi panosu Web sayfası görünümü	153
Şekil 4.64	Ülkelerin kümeleme sonuçlarının bilgi panosu üzerinde gösterimi-I.....	154
Şekil 4.65	Ülkelerin kümeleme sonuçlarının bilgi panosu üzerinde gösterimi-II	155
Şekil 5.1	COVID-19 aşısı hakkında tweet oluşturulan ilk 15 ülke	156
Şekil 5.2	COVID-19 aşısı hakkında tweet oluşturan kullanıcı sayıları bakımından ilk 15 ülke	157
Şekil 5.3	COVID-19 aşısı hakkında oluşturulan tweetlerin zamansal dağılımları ...	157
Şekil 5.4	Pfizer Biontech aşısı hakkında oluşturulan tweetlerin zamansal dağılımları	158
Şekil 5.5	Sinopharm aşısı hakkında oluşturulan tweetlerin zamansal dağılımları....	158
Şekil 5.6	Sinovac aşısı hakkında oluşturulan tweetlerin zamansal dağılımları	159
Şekil 5.7	Moderna aşısı hakkında oluşturulan tweetlerin zamansal dağılımları.....	159
Şekil 5.8	Oxford AstraZeneca aşısı hakkında oluşturulan tweetlerin zamansal dağılımları	160
Şekil 5.9	Covaxin aşısı hakkında oluşturulan tweetlerin zamansal dağılımları.....	160
Şekil 5.10	Sputnik V aşısı hakkında oluşturulan tweetlerin zamansal dağılımları	161
Şekil 5.11	Johnson&Johnson aşısı hakkında oluşturulan tweetlerin zamansal dağılımları	161
Şekil 5.12	"Covid" hakkında oluşturulan tweetlerin zamansal dağılımları	162
Şekil 5.13	"Vaccine" hakkında oluşturulan tweetlerin zamansal dağılımları	162
Şekil 5.14	COVID-19 aşısı hakkında tweetlerin oluşturulmasında kullanılan ilk 15 uygulama platformu	163

ÇİZELGE LİSTESİ

Çizelge 3.1	Farklı H3 çözünürlük seviyeleri için hücre bilgileri tablosu (URL-20)..	52
Çizelge 4.1	Ülkeler bazında COVID-19 aşıları hakkında atılan tweet sayıları.....	76
Çizelge 4.2	"user_name" ve "user_created" öznitelik alanlarıyla yeni öznitelik alanı oluşturulması	78
Çizelge 4.3	Tweet sayılarının kullanıcılara göre dağılımı	78
Çizelge 4.4	COVID-19 aşısı hakkında tweet oluşturan kullanıcı sayısının ülkelere göre dağılımı	79
Çizelge 4.5	DETECTLANGUAGE aracının örnek kullanımı	83
Çizelge 4.6	GOOGLETRANSLATE aracı ile "aşı" kelimesinin farklı dillere dönüştürülmesi	84
Çizelge 4.7	İngilizceden farklı dillerde oluşturulan tweet sayılarının ülkelere göre dağılımı	84
Çizelge 4.8	Türkiye'de İngilizceden farklı dillerde kelime içeren tweetler	85
Çizelge 4.9	Sekiz farklı aşı markası ile "Covid" ve "Vaccine" kelimelerini içeren tweetlerin dağılımları	86
Çizelge 4.10	Türkiye'nin içinde bulunduğu sekiz farklı seviyedeki H3 altıgen hücre bilgileri	90
Çizelge 4.11	Coğrafi işaretlemesi gerçekleştirilen tweetlerin üç farklı seviyedeki H3 altıgen hücre bilgileri	93
Çizelge 4.12	H3 hiyerarşik altıgen hücrelerinin kategorisel dağılımı.....	93
Çizelge 4.13	Ülkelerin X kullanıcı sayıları ve tweet sayılarına göre Ward yöntemiyle oluşturulan kümelere ait istatistikî bilgiler.....	109
Çizelge 4.14	Ülkelerin X kullanıcı sayıları ve tweet sayılarına göre k-ortalamlar yöntemiyle oluşturulan kümelere ait istatistikî bilgiler.....	112
Çizelge 4.15	Ülkelerin tweet sayıları, covid kelimesi içeren tweet sayıları ve vaccine kelimesi içeren tweet sayılarına göre Ward yöntemiyle oluşturulan kümelere ait istatistikî bilgiler.....	114
Çizelge 4.16	Ülkelerin tweet sayıları, covid kelimesi içeren tweet sayıları ve vaccine kelimesi içeren tweet sayılarına göre k-ortalamlar yöntemiyle oluşturulan kümelere ait istatistikî bilgiler.....	116

1. GİRİŞ

Bugünün dünyasında, sosyal medya en güçlü iletişim ve kişisel ilgi alanlarını oluşturma aracı olarak öne çıkmaktadır. Akıllı cihazların ve sosyal platformların geniş kullanımıyla, sosyal medyanın önemi artmakta ve birçok insanın dikkatini çekmektedir. Sosyal medya aracılığıyla insanlar arasındaki etkileşim, yapılandırılmamış, ayrıntılı ve büyük ölçekli dijital verilerin potansiyel bir kaynağına dönüşmüştür. Bu günlerde, büyük veri kavramı oldukça yaygın hale gelmiştir. Büyük veri, büyük veri kümelerini, sosyal medya analitiğini, yeni nesil veri yönetimi yeteneklerini, anlık veri akışını ve daha fazlasını içermektedir. Uzun yıllardır, dünya genelinde depolanan ve işlenmeyi bekleyen veri miktarı sürekli olarak artmaktadır. İnsanoğlu'nun yaşadığı bu dönem, verinin hammadde olarak kabul edildiği bilgi çağı olarak nitelendirilmektedir.

Günlük hayatın her dakikasında, Google kullanıcıları ortalama 5.9 milyon arama sorgusu işlemekte, Elektronik posta hizmeti kullanıcıları ortalama 231.4 milyon posta göndermekte, Facebook kullanıcıları ortalama 1.7 milyon içerik paylaşmakta, Instagram kullanıcıları ortalama 66 bin fotoğraf paylaşmakta, Youtube kullanıcıları ortalama 500 saatlik video yüklemekte, Snapchat kullanıcıları ortalama 2.43 milyon snap göndermekte ve X (Eski adıyla Twitter) kullanıcıları ortalama 347.2 bin tweet atmaktadır (URL-1).

Van Wijk (2005), veri artışındaki yaşanan ciddi büyümeyi veri patlatması olarak nitelendirmekte ve verilerdeki bu artışın daha karmaşık analiz yöntemlerine neden olacağını belirtmektedir. Ayrıca, insanların veri analizi yeteneği neredeyse sabit kalmakta; bu da mevcut verilerle yapılan analizler arasındaki boşluğun daha da genişlemesine yol açmaktadır (Maletic ve Marcus, 2010).

Bu büyük veri kümelerinden bilgi çıkarmak için iki farklı yaklaşım bulunmaktadır. **Veri madenciliği** yöntemleri, büyük veri miktarları içerisinde önceden bilinmeyen örüntüleri araştırırken, **görselleştirme** ise örüntü tanıma için son derece gelişmiş insan algısını kullanarak veri kümesinin temsilini oluşturmaktadır.

Mevcut verilere daha özelleştirilmiş, daha hızlı ve daha doğru öngörüler elde etmek amacıyla otomatik ve görsel yöntemlerin birleşimiyle ortaya çıkarak ilk dönemlerinde “etkileşimli insan-makine ara yüzleri tarafından kolaylaştırılan analitik akıl yürütme bilimi” (Wong ve Thomas, 2004) şeklinde tanımlanan **görsel analitik** disiplini geliştirilmiştir.

Görsel analitik alanının geliştirilmesiyle ortaya çıkan **mekânsal görsel analitik**, mekânsal bilgilerin etkileşimli görsel ara yüzlerle ele alındığı analitik akıl yürütme bilimi olup; coğrafi mekân ve bu mekândaki nesnelere, olaylar, olgular ve işlemleri içeren problemleri ele almaktadır (Thomas ve Cook, 2005; Andrienko ve ark., 2007; Andrienko ve ark., 2011). Mekânsal görsel analitiğin esas hedefi, incelenen olgulardaki değişikliklerin etkisini öne çıkarmak amacıyla çoklu koordinat görüntüleri ve dinamik sorgu prensiplerini kullanarak etkileşimli bir ortam oluşturmaktır.

COVID-19 pandemisi, çağımızın en belirgin küresel krizi olarak kabul edilmektedir. Bu pandemi, günlük yaşamlarımıza ve dünya görüşümüze derinlemesine etkilerde bulunan bir durum olarak öne çıkmaktadır. Başlangıçta bir sağlık krizi olarak ortaya çıkan COVID-19, kısa bir süre içerisinde hızla yayılarak tüm ülkelere ve kıtalara yayılan benzersiz bir sosyo-ekonomik ve çevresel kriz haline gelmiştir. Sonuç olarak, dünya genelinde her ölçekte (yerel, bölgesel ve küresel düzeyde) endüstriler ve sektörler, eşi görülmemiş sosyal ve ekonomik değişiklikler ve güçlüklerle karşılaşmıştır (Rajabifard ve ark., 2021).

COVID-19 pandemi döneminde, insanlar yeni normların uygulandığı bir yaşam biçimini benimsemiş ve dünyanın dört bir yanındaki bireyler, sadece sağlıklarını değil aynı zamanda işlerini, geçim kaynaklarını, sosyal ve aile ilişkilerini de kaybetme riskiyle karşı karşıya kalmıştır. Hayatın normal akışına ne zaman döneceğine dair gerçek anlayışların olmadığı bir dönem yaşanmıştır. İnsanoğlu'nun yakın zaman geçirmiş olduğu COVID-19 pandemisi gibi afet durumlarında kriz yönetimi için, kaliteli veri ve istatistiksel sonuçlar, özellikle konum bilgileri ve sonuç ürünleri, gelecekteki dayanıklılığımızı ve veri altyapımızın sağlamlığını güçlendirmek adına kritik bir rol oynamaktadır. Bu veriler, genel kamu güvenliği stratejilerimize geniş bir bakış açısı sunarak hayati önem taşımaktadırlar. Küresel bir krizde, dünya liderleri ve karar verici mekanizmalar, acil ve etkili yanıtlar sunabilmek, doğru seçimleri yapmak, karmaşıklıkla başa çıkmak ve bu şekilde dünyayı daha dirençli ve sürdürülebilir bir geleceğe yönlendirmek için erişilebilir ve güvenilir bilgilere hızla erişme ihtiyacı duymaktadırlar (Rajabifard ve ark., 2021).

Pandemi yönetiminin büyük bir kısmı, doğası gereği coğrafi ve mekânsal açıdan ele alınması gereken bir durumdur. Salgınların kontrol altına alınması, temas takibi ve yayılımın önlenmesi için etkili bir toplum müdahalesinin sağlanabilmesi adına mekânsal

bilgilere olan ihtiyaç oldukça büyüktür ve konum bilgisi bu süreçte ayrı bir önem taşımaktadır.

COVID-19 salgını, mevcut krizi hem yerel hem de küresel ölçekte daha iyi anlamamıza ve gelecekteki dünya çapındaki pandemilere daha etkin bir şekilde hazırlanmamıza yardımcı olabilecek mekânsal ve görsel altyapının önemini vurgulamaktadır. Mekânsal verileri gerçek zamanlı haritalar ve konum bilgileriyle izleme ve paylaşma ihtiyacı sürekli olarak artmaktadır. Bu nedenle, veri analitiği araçları ve çözümleriyle birlikte mekânsal ve görsel özelliklere sahip bir platform, pandemi gibi kriz dönemlerinin başlangıcında destek sağlamak, toplumu bilgilendirmek ve hazırlıklı tutmak için kritik bir rol oynamaktadır (Rajabifard ve ark., 2021).

Konum bilgisi, temel verilerin toplanması, mevcut krizin haritalandırılması, müdahale değişkenlerinin modellenerek sonuçlarının görselleştirilmesi, temas takibi ve riskli bölgelerin tespiti, yüksek riskli konumların idaresi ve acil yardımın en fazla gereksinim duyulan bölgelere dağıtımı gibi işlevleri sağlayarak, krizin yönetilmesinde önemli bir rol üstlenmektedir.

Sosyal medya kullanımı, birçok bireyin günlük yaşamının ayrılmaz bir parçası haline gelmiştir. Bu platformların acil durum senaryolarında yaygın olarak kullanılması son derece doğaldır. Sosyal ağlar, belirsizliği azaltan, hızlı ve güvenilir bilgi akışı sağlayarak kriz anlarında vatandaşların özgüvenini etkileyebilmektedir (Jurgens ve Helsloot, 2018). Özellikle, afet durumları bağlamında X (Eski adıyla Twitter) kullanımı üzerine yapılan çalışmalar, retweetlerin ve özel mesajların acil durum bilgi yayılımını (Sutton ve ark., 2014) ve afetler sırasında X (Eski adıyla Twitter)'de bilgi paylaşım davranışlarını (Toriumi ve ark., 2013; Mendoza ve ark., 2010) ortaya koymaktadır.

1.1. Tezin Amacı

Mekânsal görsel analitik sistemleri farklı disiplinlerden bilim ve teknolojinin birleşmesiyle oluşmaktadır. Görsel analitik sistemlerinin merkezinde görselleştirme yer almaktadır. Mekânsal görsel analitik, kullanıcıların mekânsal verileri kullanarak örüntüleri tespit etmeleri ve gelecekteki sonuçları tahmin etmeleri için kartografya, veri yönetimi, veri madenciliği, arayüz tasarımı ve bilişsel bilim vb. farklı disiplinlerden yararlanmaktadır.

Bu tez çalışmasının temel amacı, bilgi biliminin ve kartografyanın etkileşimli yapısını yansıtan Web tabanlı mekânsal görsel analitik uygulama örneği

geliştirilmesidir. Bu çalışma kapsamında ilk olarak, COVID-19 pandemisi gibi kriz durumlarında sosyal medya platformlarında ortaya çıkacak metinsel mesaj veri setlerinden, hedeflenen konuya ilişkin içeriklerin mekânsal yayılımını görselleştirmek amacıyla konumsal bilgi çıkarımı gerçekleştirilecektir. İkinci aşamada, araştırmacının mevcut veri seti içerisinde keşfetmeyi hedeflediği konularla ilgili saklı örüntülerin tespiti gerçekleştirilecektir. Üçüncü aşamada, veri setinden keşfedilen yeni bilgiler kullanılarak, veri madenciliği yöntemlerinden biri olan kümeleme analizi yöntemiyle, tweetlerin oluşturulduğu ülkelerin benzerliklerin ve farklılıkların belirlenmesi amacıyla analiz işlemi gerçekleştirilecektir. Çalışmanın son aşamasında, veri setinde yer alan mesajların, veri setinden keşfedilen yeni bilgilerin, kümeleme analizi sonucu elde edilen bilgiler ile üretilen tematik haritaların birbirleriyle etkileşim içinde olduğu, araştırmacının belirlediği sorgulamalar kapsamında filtrele araçlarının yer aldığı, veri setindeki mesajların mekânsal - zamansal gelişiminin aynı ekran üzerinden takip edildiği ve araştırmaya ait sonuçların farklı grafiksel araçlarla kullanıcıya aktarıldığı Web tabanlı interaktif bir uygulama geliştirilecektir.

Bu tez kapsamında gerçekleştirilecek bu araştırmanın, kriz durumları gibi yerel ya da küresel ölçekte ortaya çıkabilecek mekâna ilişkin problemlerin tahmininde, analizinde ve çözümünde önemli rol oynayacağı düşünülmektedir.

1.2. Tezin Organizasyonu

Bu tez çalışması altı ana başlık altında toplanmıştır. Birinci bölümde, tezin konusu kapsamında giriş yapılarak, tezin amacı ve kapsamı hakkında bilgiler verilmektedir. İkinci bölümde, büyük veri, mekânsal büyük veri, veri madenciliği, veri tabanlarında bilgi keşfi, görselleştirme, görsel analitik, mekânsal görsel analitik, sosyal medya ve X (Eski adıyla Twitter) konularına genel bir bakış sunulacak ve tez konusu kapsamında gerçekleştirilen daha önceki çalışmalar hakkında bilgiler verilmektedir. Üçüncü bölümde, bu çalışma kapsamında kullanılan materyal ve yöntemler hakkında bilgiler verilmektedir. Dördüncü bölümde, X (Eski adıyla Twitter) mesajlarının yer aldığı veri seti kullanılarak, tweet mesajlarının konumsal işaretlenmesi, veri setinde yer alan gizli verilerin keşfi, kümeleme analizi ve Web tabanlı mekânsal görsel analitik uygulama örneği gerçekleştirilmektedir. Beşinci bölümde, araştırma bulgularına ilişkin değerlendirmeler yapılmaktadır. Altıncı bölümde ise, tez çalışmasında elde edilen sonuç ve öneriler açıklanmaktadır.

2. KAYNAK ARAŞTIRMASI

2.1. Büyük Veri

Bilgi ve iletişim teknolojilerinin ilerlemesiyle birlikte, internet üzerindeki Web sayfaları, mobil cihazlar, sosyal medya platformları ve sensörler gibi pek çok kaynaktan günlük hayatımızla ilgili veriler toplanmaktadır. Bu durum, giderek artan bir veri yığınının varlığına işaret etmektedir. Bu verilerin değeri anlaşıldıkça, düzenli bir yapıda toplanmaları, saklanmaları, analiz edilmeleri ve sonuç olarak bilgiye dönüştürülerek kullanıcıların erişimine sunulmaları önem kazanmaktadır. Hem akademik çevrelerde hem de kamu kurumlarında, verilerin etkili bir şekilde yönetilmesi birincil hedef haline gelmektedir.

Büyük veri kavramı, ilk defa 1990'ların ortalarında bilimsel çevrelerde görülmüş ve son on yılda akademik alanlardan sanayi kuruluşlarına, kamu kurumlarından diğer organizasyonlara kadar geniş bir yelpazede dikkat çeken küresel bir odak noktası haline gelmiştir (Li ve ark., 2016).

Büyük veri terimi hala kesin bir tanıma kavuşmamış olsa da akademik, teknolojik ve endüstriyel çevreler, bu terimi çeşitli şekillerde açıklamışlardır (Chen ve ark., 2014). Büyük veri, hesaplama ve iletişim yeteneklerimiz karşısından büyük bir hızla akan veridir (Mutukrishnan, 2005). Büyük veri, tipik ve klasik veri tabanı yazılım araçlarının kabiliyetlerinin ötesinde verileri alan, depolayan, yöneten, analiz eden ve erişime sunan araçları tanımlamak için kullanılmıştır (Manyika ve ark., 2011). Büyük veri, karmaşık ve gelişmiş veri depolama, yönetme, analiz ve görselleştirme teknolojileri gerektiren veri kümeleri ve analitik teknikleridir (Chen ve ark., 2012).

Büyük veri, geleneksel donanım, yazılım ve veri tabanı teknolojileriyle kolaylıkla yakalanamayan, depolanamayan, yönetilemeyen, analiz edilemeyen ve sunulamayan büyük veri hacimlerine sahip yapılandırılmış ve yapılandırılmamış veri kümeleri olarak kabul edilmektedir (Li ve ark., 2016).

Laney 2001 yılında ilk defa büyük verinin özelliklerini "Volume (Hacim) - Velocity (Hız) - Variety (Çeşitlilik)" 3V modeli olarak açıklamıştır (Laney, 2001). Büyük veriyi tanımlamak için ilk olarak 3V modeli kullanılırken, büyük veri hakkında yapılan çalışmalar sonucu Value (Değer) ve Veracity (Doğruluk) özellikleri de eklenerek Şekil 2.1'de görülmekte olan 5V modeline ulaşılmıştır (Demchenko ve ark., 2013; Narasimhan ve ark., 2014).



Şekil 2.1 Büyük verinin özellikleri (Demchenko ve ark., 2013; Narasimhan ve ark., 2014)

Hacim (volume), büyük verinin en iyi bilinen özelliğidir. Verinin boyutunu ifade etmektedir. Uzaktan algılama görüntü verileri için petabayt boyutunda veri arşivleri gerekmektedir. Gerçek zamanlı sensör verileri, konum tabanlı sosyal medya verileri ve sürekli artan gönüllü coğrafi verileri yalnızca veri depolama konusunu değil, aynı zamanda büyük analiz konusunu da gündeme getirmektedir (Dasgupta, 2013).

Sensörler, sosyal medya, Web sayfaları, mobil cihazlar vb. veri üreten kaynakları hızla artması beraberinde veri üretim **hızı (velocity)** da artmaktadır. Çok sık üretilen yüksek çözünürlüklü uydu görüntüsü verileri, sensör gözlemlerinin sürekli akışı, gerçek zamanlı GNSS verileri, sosyal medya verilerinin tamamında üretme hızını ve talebi karşılamak için yüksek hızlı veri işlemeyi gerektirir (Dasgupta, 2013).

Çeşitlilik (variety), bir veri kümesindeki yapısal heterojenliği ifade etmektedir. Teknolojik gelişmeler, firmaların çeşitli yapılandırılmış, yarı yapılandırılmış ve yapılandırılmamış verileri kullanmasına izin vermektedir (Gandomi ve Haider, 2015). Mevcuttaki tüm verilerden sadece %5'ini yapılandırılmış veriler (Elektronik tablolarda veya ilişkisel veri tabanlarında tutulan veriler) oluşturmaktadır (Cukier, 2010). Harita verileri, görüntü verileri, coğrafi etiketli metin verileri, yapılandırılmış ve yapılandırılmamış veriler, raster ve vektör verileri, çoğu karmaşık yapıya sahip tüm bu farklı türdeki veriler daha verimli modeller, yapılar, indeksler ve veri yönetimi stratejileri ve teknolojileri gerektirmektedir. Örneğin: NoSQL kullanımı (Li ve ark., 2016).

Oracle'ın tanımına göre, büyük veriler genellikle nispeten “düşük değer yoğunluğu” ile tanımlanmaktadır. Yani, orijinal formda alınan veriler genellikle hacmine göre düşük bir **değere (value)** sahiptir. Bununla birlikte, bu tür büyük hacimli veriler analiz edilerek yüksek bir değer elde edilebilmektedir (Gandomi ve Haider, 2015).

Doğruluk (veracity), büyük veri için iki farklı şekilde ele alınmaktadır. Bunlardan birincisi büyük veri kaynaklarının güvenilirliğini, ikincisi ise verinin hedef kitleye uygunluğunu göstermektedir. Büyük verinin doğruluğu, farklı kaynaklardan gelmesi ve veriyi toplayan kurumlar tarafından kontrolünün sağlanmaması nedeniyle en zor bileşen olmaktadır (Sütcü ve Aytakin, 2013; Bayrakçı, 2015). Mekânsal büyük verilerin çoğu, düşük veya bilinmeyen doğruluğa sahip doğrulanmamış kaynaklardan gelmektedir. Doğruluk seviyesi veri kaynaklarına bağlı olarak değişmektedir. Kaynak verilerinin kalite değerlendirmesi ve analiz sonuçlarının “istatistiksel” olarak nasıl artırılacağı gündeme gelmektedir (Li ve ark., 2016).

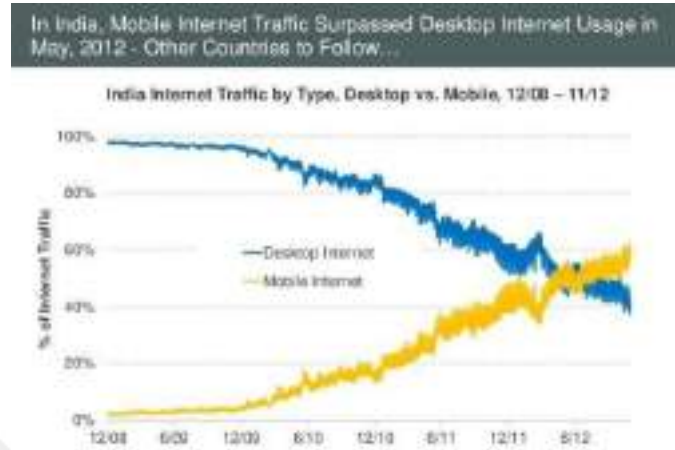
2.2. Mekânsal Büyük Veri

Yeryüzünde veya yakınında belirli bir anlama sahip olan doğal (nehir, orman vb.) ve yapay (yol, bina vb.) nitelikteki coğrafi verileri belirli bir referans sistemine göre yerini ve biçimini belirten vektör ve raster verilere **mekânsal veri** denilmektedir (Çağlar, 2018).

CBS verileriyle çalışanların en aşına tanımlardan birisi “Verilerin %80'i coğrafidir” ifadesidir (Morais, 2012). Bu ifade ve “Büyük Veri” tanımlamaları dikkate alındığında mekânsal veriler içinde “**Mekânsal Büyük Veri**” tanımlaması yapılabilmektedir.

Mekânsal büyük veri, genel olarak mevcut hesaplama sistemlerinin kapasitesini aşan mekânsal veri setlerini ifade etmektedir (Lee ve Kang, 2015). McKinsey Global Institute, 2009 yılında kişisel konum verileri havuzunun 1Petabayt düzeyinde olduğunu ve yılda %20 oranında büyüdüğünü söylemektedir (Dasgupta, 2013). Birleşmiş Milletler Küresel Coğrafi Mekânsal Bilgi Yönetimi Girişimi tahminine göre, her gün 2.5 kuintilyon bayt (2.5 milyar Gb) veri üretilmekte ve verilerin büyük bir kısmı konumla ilişkilidir. Ayrıca Google'da günde yaklaşık 25 PB veri üretilmektedir ve verilerin önemli bir kısmı mekânsal zamansal veriler alanına girmektedir (Vatsavai ve ark., 2012). Mekânsal büyük verideki bu eğilim, mobil cihazların kullanımının artmasıyla

doğrusal bir şekilde büyümeye devam etmektedir. Şekil 2.2’de görüldüğü üzere; Hindistan’da, mobil cihazlardan gelen internet trafiği, masaüstü bilgisayarlardan gelen internet trafiğini aşmaktadır (Meeker, 2012).



Şekil 2.2 Hindistan'da mobil internet trafiği (Meeker, 2012)

Mekânsal büyük verilerin artışıyla birlikte, mekânsal olarak zorunlu kılınan içeriklerin modellenmesi ve simülasyonu için yüksek performanslı bilgi işlem becerisine daha fazla ihtiyaç duyulmaktadır. Ancak, sınırlı işleme gücü nedeniyle, birçok uygulamada yüksek hacimli veya yüksek hızlı coğrafi mekânsal veri toplamasından tam olarak yararlanmak zordur. Hadoop, Hive ve MongoDB gibi büyük veri platformları kullanıcılar için dağıtık ve paralel hesaplamalar yapan büyük veri analizi yazılımları geliştirmektedir (Lee ve Kang, 2015).

Mekânsal büyük veri, insanoğlu için bir fırsattır (Eagle ve Greene, 2014). Milenyum Projesiyle, insan türünün karşı karşıya kalacağı tespit edilen 15 küresel zorluk Şekil 2.3’te gösterilmektedir. Shashi Shekhar, “Sürdürülebilir Kalkınma ve İklim Değişikliği, Temiz Su, Nüfus ve Kaynaklar, Sağlık Sorunları, Barış ve Çatışma, Uluslararası Organize Suçlar ve Enerji” başlıkları altında toplanan yedi zorluğun mekânsal büyük verilerle ilgili olduğunu söylemektedir (Shekhar, 2012).

Enerji alanında, Ford firmasının mekânsal büyük veriler kullanarak geliştirdiği “Ekonomik Rota” uygulaması kullanılarak enerji tasarrufu sağlanabilen iyi bir örnektir. Bu teknoloji, seyahat etme süresi ya da seyahat mesafesi yerine yakıt tüketimini en aza indirmektedir. “En Hızlı Rota” seçeneğine kıyasla, “Ekonomik Rota” seçeneği kullanılarak bazı araçlarda %15’e varan oranlarda yakıt tüketimi sağlanmaktadır (Valdes-Dapena, 2011).

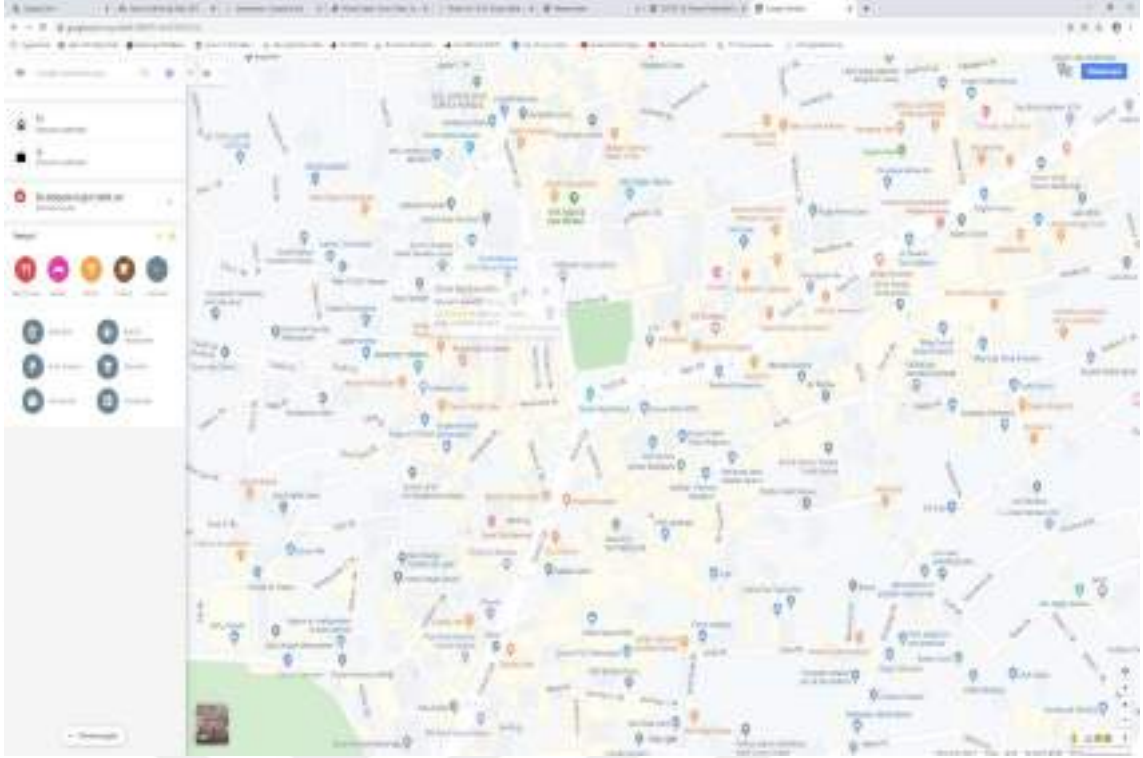


Şekil 2.3 İnsanoğlu'nun karşılaştığı 15 küresel zorluk (Shekhar, 2012)

Mekânsal büyük verinin barındırdığı konum özelliği tartışılmaz bir güçtür. WPP grup CEO'su Sir Martin Sorrell, “Konum belirleme pazarlamacılar için kutsal bir kâse” demektedir (Sorrell, 2011). Büyük veri analizi, konumun gücünü artırmanın etkili bir yoludur (Percivall, 2013).

İnsan hareketliliği analizinden yararlanabilecek diğer bir örnek doğrudan pazarlamadır. Bu alan Tobler’ın ilk coğrafya yasası “Her şey diğer her şeyle ilgilidir, fakat yakın şeyler uzak şeylerden daha fazla ilişkilidir.” tarafından desteklenmektedir (Tobler, 1970). Yani, bir kullanıcının geçerli konumuna yakın olan hizmet veya mağazaların önerilmesi, uzaktaki kişilere önermekten daha etkili olmaktadır. Konum hedeflemenin, mobil reklamcılık performansının çok yüksek olduğu bilinmektedir. Günümüzde kullanıcılar konum tabanlı hizmetlere bağlandığında (Google Maps, Foursquare), kullanıcıya yakın olan birçok ürün/hizmet sağlayıcının konumu bildirim olarak ekranda görülmektedir (Şekil 2.4).

Mekânsal büyük veriler; raster veriler ve vektör veriler olarak kategorize edilmektedir. Raster veriler; uydulardan, hava araçlarından, güvenlik kamera görüntülerinden ve son zamanların popüler teknolojisi olan insansız hava araçlarından (drone) elde edilen coğrafi görüntüleri içermektedir. Vektör veriler; noktalar, çizgiler ve alanlardan oluşmaktadır. Klasik anlamda harita verileri vektör yapısına aittir.



Şekil 2.4 Google Maps tarafından sunulan konum tabanlı hizmetler

Gündelik hayatımızın bir parçası olan konum tabanlı hizmet sağlayıcılarında (Google Maps, Foursquare) bulunduğumuz konum ve hedef noktalarımız *nokta*, takip edeceğimiz güzergahlar *çizgi* ve çevremizdeki binalar, parklar vb. alanlar *alan* vektör yapısıyla kullanımımıza sunulmaktadır.

Yakın zamana kadar mekânsal veri toplama, teknik olarak zor, ekonomik olarak pahalı ve kullanım olarak karmaşık cihazlar ile gerçekleştiriliyordu. Gelişen teknoloji sayesinde mekânsal veri, hepimizin kullandığı akıllı telefonlar, tabletler, saatler gibi cihazlarla farkında olarak ya da olmayarak topladığımız bir ürün haline gelmiştir. Sosyal ağlar üzerinden toplanan ve Gönüllü Coğrafi Bilgi (VGI) adı verilen veriler, coğrafi sensör ağlarından gelen veriler, konumsal bilgilerin kullanılabilirliğinin artmasına neden olmuştur (Sester ve ark., 2014).

İnsanlar “sensör” olarak düşünüldüğünde Gönüllü Coğrafi Bilgi (VGI) adı verilen verilerin gündelik hayatta kolaylık sağlayacak mekânsal verilerin toplanmasında yardımcı olmaktadır. Gönüllü Coğrafi Bilgi edinimi, katılımcı ve fırsatçı olarak ayırt edilmektedir. Katılımcı veri toplama, bilinçli bir süreçte, nesnelere ve özelliklerini açıkça seçen ve bu bilgilere katkıda bulunan bir kullanıcı tarafından gerçekleştirilmektedir (Örneğin, OpenStreetMap). Fırsatçı veri toplama bilinçsiz bir şekilde, çoğunlukla belirli bir amaç olmadan gerçekleşmektedir. Öne çıkan bir örnek,

trafik sıkışıklığı gibi trafik bilgilerini belirlemek için cep telefonu verilerinin kullanılmasıdır (Li ve ark., 2016). Gönüllü Coğrafi Bilgi adı verilen özel verilere bir afet (felaket) anında olduğu gibi ihtiyaç duyulduğunda önemli bir veri kaynağı olduğunu kanıtlamıştır (Goodchild ve Glennon, 2010). Sosyal medya ve hizmetler, erken müdahale ve kriz yönetimi için yeni bilgi kaynakları olarak görülmektedir. Fuchs ve ark. (2013), Almanya'daki büyük çaplı taşkın olaylarını tespit etmek için X (Eski adıyla Twitter) akışlarını değerlendirmiştir. 8 aylık bir sürede yaklaşık 6 milyon tweet kaydedilmiştir. Konumsal-zamansal kümelemeyle birlikte belirli anahtar kelimelerin dahil edilmesiyle, bazı olayların tespiti gerçekleştirilmiştir (Fuchs ve ark., 2013).

2.3. Veri Madenciliği

Veri madenciliği terimi, veri tabanı teknolojileri, makine öğrenimi, istatistik, örüntü tanıma ve yapay zekâ alanlarına dayanan disiplinler arası bir araştırma alanını tanımlamaktadır (Han ve Kamber, 2006).

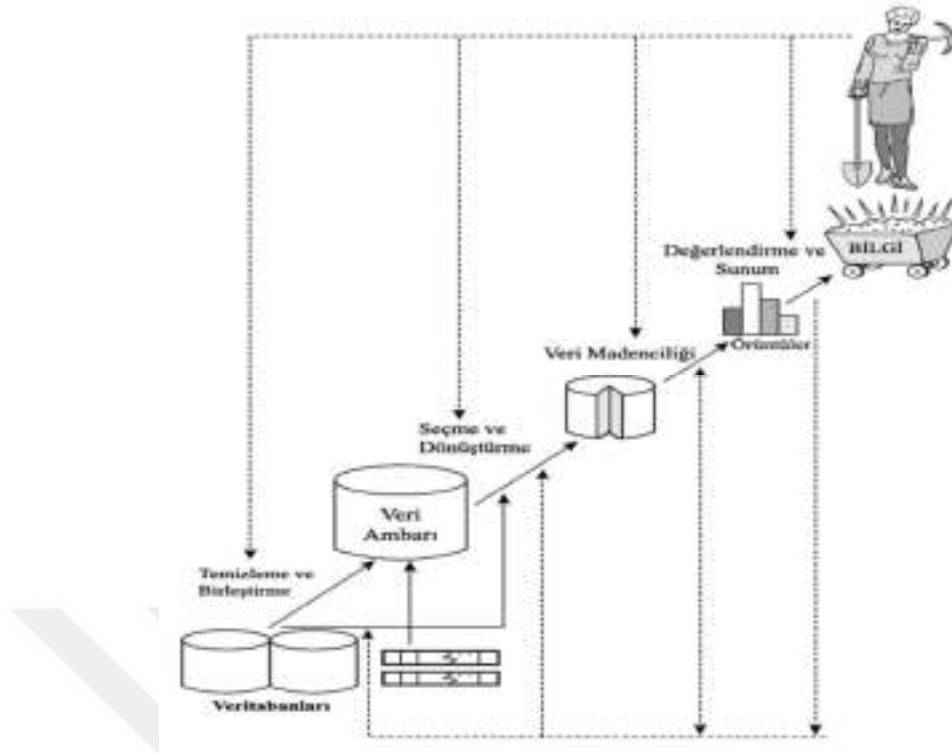
Veri madenciliği, verilerden örüntüler çıkarmak için özel algoritmaların uygulanmasıdır (Fayyad ve ark., 1996). Veri madenciliği, veri tabanlarında, veri ambarlarında veya diğer bilgi havuzlarında depolanan büyük miktardaki verilerden ilginç bilgileri keşfetme sürecidir (Han ve Kamber, 2006). Veri madenciliği, ilginç örüntüler ve ilişkiler için büyük veri tabanlarını eleme sürecini tanımlamak için kullanılan bir terimdir (Maimon ve Rokach, 2010). Bu tanımların ortak noktası, veri madenciliğinin temelinde, içinde saklı olan örüntüleri ve düzenlilikleri ortaya çıkarmak amacıyla büyük miktarda veriyle ilgilenmesidir. Ayrıca, veri madenciliği terimi genellikle otomatik veya yarı otomatik bir süreç olarak adlandırılmaktadır (Olafsson ve ark., 2008).

Literatürde veri madenciliği genellikle bilgi keşfi süreciyle eş anlamlı olarak kullanılmaktadır, ancak diğer bir yandan da Şekil 2.5'te görüldüğü gibi bu süreç içerisinde yalnızca tek bir adımı olarak görülmektedir (Han ve Kamber, 2006).

Veri madenciliği tanımlayıcı ve tahmin edici olarak ayrılmaktadır (Han ve Kamber, 2006).

2.3.1. Tanımlayıcı analiz yöntemleri

Tanımlayıcı veri madenciliği yöntemleri, veri setinin mevcut durumu hakkında açıklamalar yapmaktadır. Bunlar aşağıdaki sınıflara ayrılabilir.



Şekil 2.5 Veri tabanlarında bilgi keşfinin adımları (Han ve Kamber, 2006)

2.3.1.1. Kümeleme analizi

Kümeleme, benzer verileri/nesnelere tanımlamayı ve bunları kümeler halinde gruplamayı amaçlamaktadır. Kümelerin hangi özelliklere göre oluşturulacağı başlangıçta sabitlenmez ve prosedüre bırakılmaktadır. Burada, bir küme içindeki öğeler mümkün olduğunca benzer, kümeler ise mümkün olduğunca birbirinden farklıdır (Han ve Kamber, 2006).

2.3.1.2. Birliktelik kuralları ve ilişki analizi

Birliktelik Kuralları ve İlişkilendirme yöntemi, veriler arasındaki bağımlılıkları bulmak için kullanılmaktadır. Bu yöntemin en sık kullanılan örneği, sıklıkla birlikte satın alınan ürünleri tanımlayan alışveriş sepeti analizidir (Han ve Kamber, 2006). Birliktelik kuralının hesaplanmasında Destek ve Güven ölçütü önemlidir.

Bir birliktelik kuralının desteği, birliktelik kuralının tüm öğelerinin bir işlemde (bir tuple) meydana gelme sıklığını tanımlamaktadır. Bir birliktelik kuralı (A -> B) için yüzde 10 veya 0,1'lik bir destek, tüm işlemlerin yüzde 10'unda hem A hem de B'nin bulunduğu anlamına gelmektedir. Ayrıca, destek mutlak olarak (meydana gelme sayısı) veya göreceli olarak (toplam işlem sayısına göre) hesaplanabilmektedir (Han ve Kamber, 2006).

Bir birliktelik kuralının güveni, kuralın doğru olduğu zamanların yüzdesini tanımlamaktadır. Yüzde 70 veya 0,7 güvene sahip bir birliktelik kuralı (A -> B), A'nın gerçekleştiği durumların yüzde 70'inde B'nin de gerçekleştiği anlamına gelmektedir (Han ve Kamber, 2006).

2.3.2. Tahmin edici analiz yöntemleri

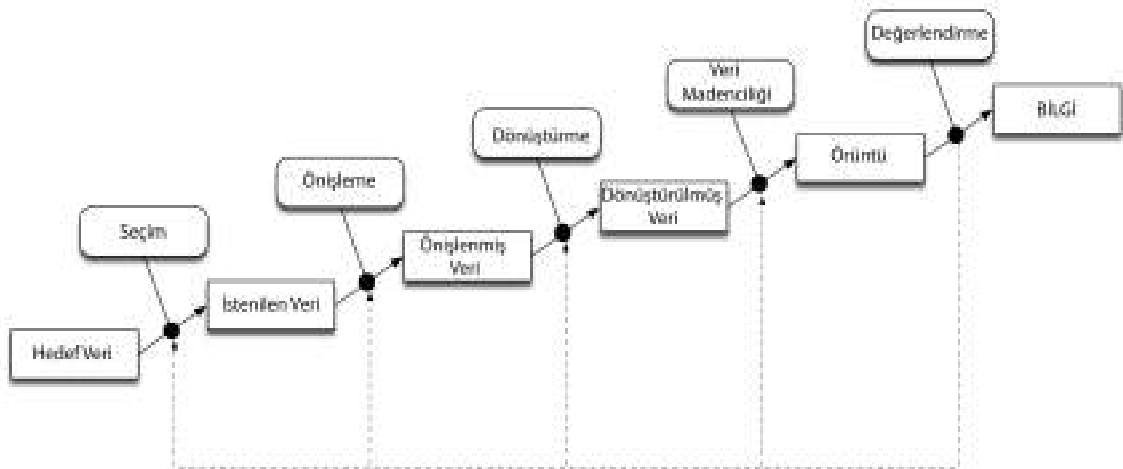
Tahmin Edici veri madenciliği yöntemleri henüz mevcut olmayan veya gelecekteki durumları tahmin etmeye çalışmaktadır.

2.3.2.1. Sınıflandırma ve regresyon yöntemi

Sınıflandırma yöntemi, bir eğitim veri seti temelinde belirli özelliklere sahip farklı sınıflar oluşturmaktadır. Yeni ögeler daha sonra bu sınıflara ayrılmaktadır. Bunun bir örneği, yaş ve gelir gibi kriterlere dayalı olarak insanların kredi notu skorlarının belirlenmesidir. Regresyon yöntemi, halihazırda bilinen değerlere dayalı olarak yeni değerlere beklenen bir değer atayan bir fonksiyon tanımlamaya çalışmaktadır. Bunun önceki verilerde gerçekleşmesi gerekmemektedir (Han ve Kamber, 2006).

2.4. Veri Tabanlarında Bilgi Keşfi

Veri tabanlarında Bilgi Keşfi (KDD), bazı kaynaklarda Veriden Bilgi Keşfi (Han ve Kamber, 2006) bazı kaynaklarda da Bilgi Keşfi ve Veri Madenciliği (Puolamäki ve ark., 2010) olarak da adlandırılmaktadır. Veri tabanlarında Bilgi Keşfi, ham veriden kullanılabilir ve değerli bilgiye kadar olan değer zinciri şeklinde tanımlanmaktadır (Puolamäki ve ark., 2010). Bilgi Keşfi süreci Şekil 2.6'da şematik olarak gösterilmektedir (Fayyad ve ark., 1996).



Şekil 2.6 Bilgi keşfi sürecinin aşamaları (Fayyad ve ark., 1996)

Seçim aşamasında, Mevcut bir veri kümesinden (farklı veri kaynakları veya bir veri kaynağının alt kümesi), sürecin yürütüleceği tüm veriler seçilmektedir. Gerekli tüm veriler bütünleştirilmezse, süreç başarısız olabilir. Veri alt kümesi hedef veri olarak adlandırılmaktadır (Fayyad ve ark., 1996).

Önişleme aşamasında, mümkün olan en iyi veri kalitesini sağlamak için veriler daha ileri analizler için hazırlanmaktadır. Veriler daha sonra ön işlemden geçirilmiş veriler olarak adlandırılmaktadır (Fayyad ve ark., 1996).

Dönüştürme aşamasında, veriler veri madenciliği adımı için hazırlanmaktadır. Bu aşamada, veri seti içerisinde öznitelik seçimi, verilere üzerinde ayrıştırma ve genelleştirme, verilerin normalleştirilmesi ve yeni öznitelik oluşturma işlemleri uygulanabilmektedir. Bu adım genellikle başarı için belirleyici aşamadır (Maimon ve Rokach, 2010).

Veri madenciliği aşamasında, bağlı olarak bir veri madenciliği algoritması seçilmekte ve hazırlanan veri kümesinde örüntüler aranmaktadır (Fayyad ve ark., 1996; Maimon ve Rokach, 2010).

Değerlendirme aşamasında, veri madenciliği yöntemleriyle bulunan örüntüler kullanıcıya sunulmakta ve kullanıcı tarafından yorumlanmaktadır (Fayyad ve ark., 1996; Maimon ve Rokach, 2010).

Tüm bilgi keşif süreci yinelemeli bir süreçten oluşmaktadır. Yani her adım istenildiği sıklıkta tekrarlanabilir veya gerekirse bir önceki adıma geri dönmek ve kazanılan öngörüyle o noktadan devam edilebilmektedir (Han ve Kamber, 2006; Maimon ve Rokach, 2010).

2.5. Görselleştirme

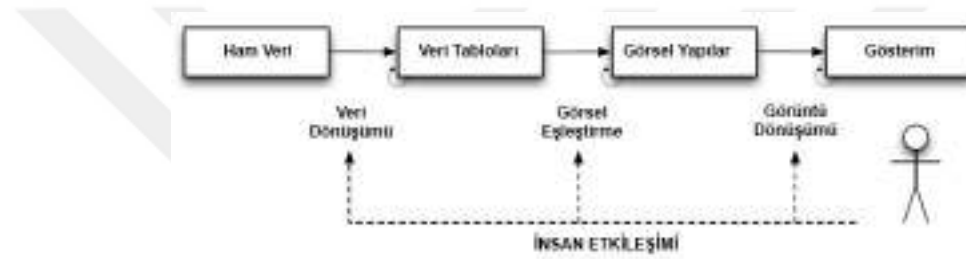
Görselleştirme terimi literatür araştırmalarında farklı şekillerde tanımlanmaktadır. MacEachren (1994) görselleştirmeyi, “düşünsel simgeleri geliştirmeye yarayan insani bir kabiliyet ve bu kabiliyeti arttıran ve kolaylaştıran araçların kullanımı” şeklinde tanımlamaktadır.

Card ve ark. (1999) görselleştirmeyi, “bilgisayar destekli, etkileşimli, görsel veri temsillerinin bilişi güçlendirmek için kullanılması” olarak tanımlamaktadır. Bu tanım görselleştirme için temel bir tanım olup farklı alanlara ait veriler;

Bilimsel alana ait veriler için görselleştirme terimi, "Bilimsel verilerin bilgisayar destekli, etkileşimli, görsel temsillerinin, tipik olarak fiziksel temelli, bilişi güçlendirmek için kullanılması." şeklinde tanımlanırken,

Bilgi görselleştirilmesi, finansal verilerin görselleştirilmesi veya soyut verilerin görselleştirilmesi alanlarında kullanılan görselleştirme terimi, "Bilişi güçlendirmek için soyut, fiziksel temelli olmayan verilerin bilgisayar destekli, etkileşimli, görsel temsillerinin kullanılması" olarak tanımlanmaktadır (Card ve ark., 1999).

Card ve ark. (1999) görselleştirme için referans model sunmaktadır (Şekil 2.7). Bu referans model verinin ham halinden başlayarak verinin görsel temsiline kadar olan süreci tanımlamaktadır.



Şekil 2.7 Görselleştirme için referans modeli (Card ve ark., 1999)

Veri Dönüşümü: Veriler başlangıçta ham veri olarak adlandırılan herhangi bir formatta bulunmaktadır. Ham verilerin öncelikle ilişkisel açıklamalara, veri tablolarına dönüştürülmesi gerekmektedir. Ham veriler bu adımda yapılandırılmakta ve sonraki adımlarda daha kolay işlenmektedir. Bu aşamada, ham verilerde sıklıkla bulunan yanlış veya eksik değerler düzeltilebilir ya da belirlenebilir. Ayrıca, ek bilgiler hesaplanabilir ve verilere eklenebilir (Card ve ark.,1999).

Görsel Eşleştirme: Bu adımda veri tabloları, çoğu durumda çeşitli olasılıkların mevcut olduğu görsel yapılarla eşleştirilmektedir. Anlamlı bir eşleme bulmak, iki ön koşulu yerine getirmesi gereken önemli bir görevdir. Eşleştirme 'anlamlı' olmalıdır, yani görsel ilişkisi bulunan tüm veriler eşleştirilmelidir. Ayrıca görselleştirme 'etkili' olmalıdır, yani mümkün olduğunca çabuk yorumlanmalıdır (Card ve ark.,1999).

İnsan algısı görselleştirmenin verimliliğinde belirleyici bir rol oynamaktadır. Görsel algıya öncelik verilmelidir. Görsel algı kavramı, insanların istemsiz olarak farklı görsel özellikleri tanıma yeteneğini tanımlamaktadır. Bu şekilde, kullanıcı karelerden oluşan bir küme içerisinde bir üçgenin bulunup bulunmadığını saniyenin çok küçük bir bölümünde değerlendirebilmektedir (Healey ve ark., 1996).

Görüntü Dönüşümü: Görselleştirme sürecinin son adımında, görüntülerin görünümü, statik bir görüntüden daha fazla bilgi elde etmek için etkileşimli olarak dönüştürülebilmektedir. Burada üç temel kavram tanımlanmaktadır:

Konum belirleyiciler, daha ayrıntılı bilgi görüntülemek için bir öğeyi işaretleyerek ek bilgi sağlamaktadır. Alternatif olarak, bir ayrıntı penceresi yerine, çevreleyen bölge farklı bir şekilde, örneğin büyük ölçüde büyütülmüş olarak görüntülenebilmektedir.

Bakış Açısı Kontrolörleri, verilere bakış şeklimizi değiştirmenize olanak tanımaktadır. Yakınlaştırma (bir alanı büyütme) ve kaydırma (ekran bölümünü hareket ettirme) en etkin kontrol araçlarıdır.

Distorsiyon, mevcut görsel yapı içerisine detayların dahil edilmesini sağlamaktadır. Burada verilerin bir alanının büyütülmüş bir şekilde vurgulanmasından bahsedilmektedir. Bu teknik kullanılarak, bağlam korunurken kısmi bir alan değerlendirilmektedir (Card ve ark., 1999).

Web tabanlı gerçekleştirilen modern bilimsel görselleştirmede, görselleştirme sürecinin adımları Şekil 2.8’de görüldüğü gibi **Filtreleme** (veri seçimi ve hazırlanması), **Eşleştirme** (geometrik olmayan verilerin geometrik nesnelere ve renklere eşleştirilmesi) ve **Renderlama** (görüntü oluşturma) şeklinde tanımlanmaktadır (Upson ve ark., 1989; Haber ve McNabb, 1990).



Şekil 2.8 İnternet tabanlı görselleştirme adımları (Upson ve ark., 1989; Haber ve McNabb, 1990)

Bu model İnternet üzerinden görselleştirmeyi anlamamıza yardımcı olmaktadır. Veri açıkça yayıncı tarafında yer almaktadır. Görsel ise Görüntüleyici tarafında bulunmaktadır (Wood ve ark., 1996).

Senaryo I: Yayıncı bir görsel oluşturur

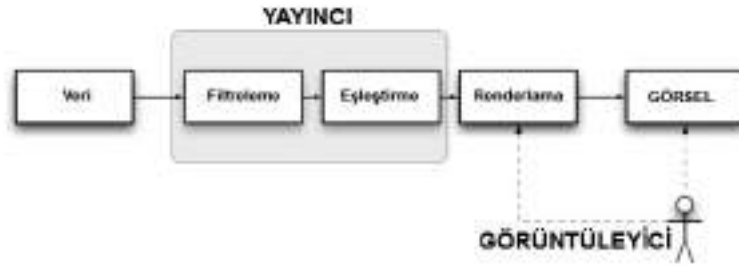
En basit durumda (Şekil 2.9), arka plan bilgisine sahip yayıncı, daha sonra bir görüntüleyici tarafından yorumlanabilecek bir görsel oluşturur. Ancak, görüntüleyicinin bu görsel ile etkileşime girme imkânı yoktur (Wood ve ark., 1996).



Şekil 2.9 Yayınıcı bir görsel oluşturur, görüntüleyici statik bir görüntüyü analiz eder (Wood ve ark., 1996)

Senaryo II: Yayınıcı bir geometri modeli oluşturur

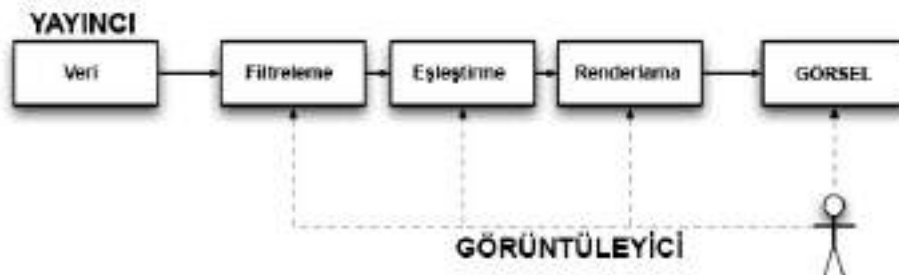
Bu senaryoda (Şekil 2.10), görüntüleyiciye görselleştirme üzerinde kontrol verilmektedir. Böylece, yayınıcı yalnızca temel geometri modelini oluşturduğundan, etkileşim görselleştirme içinde gerçekleşebilmektedir. Ancak görüntüleyici, etkili bir görselleştirme için çok önemli bir adım olan eşleştirmeye müdahale edememektedir (Wood ve ark., 1996).



Şekil 2.10 Yayınıcı modeli oluşturur, görüntüleyici görüntü üretimini kontrol eder (Wood ve ark., 1996)

Senaryo III: Görüntüleyici görselleştirmeyi oluşturur

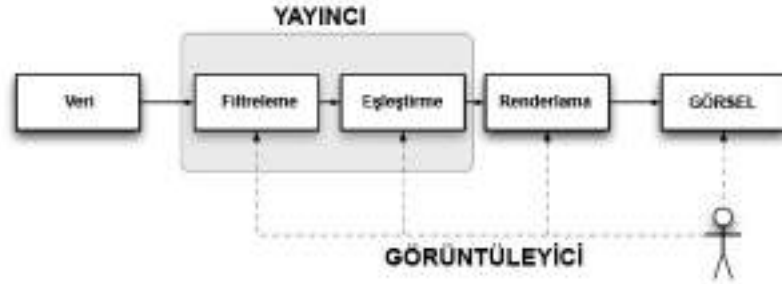
Üçüncü senaryoda (Şekil 2.11), görüntüleyici görselleştirme sürecinin münferit adımları üzerinde tam kontrole sahiptir. Yayınıcı sadece ham verileri sağlamaktadır. Bu durumda, görüntüleyiciye mümkün olan en büyük analiz özgürlüğünü verilmektedir. Ancak görüntüleyicinin bilgi ve hesaplama kapasitesine sahip olması gerekmektedir (Wood ve ark., 1996).



Şekil 2.11 Görüntüleyici görseli oluşturur (Wood ve ark., 1996)

Senaryo IV: Yayıncı, görüntüleyicinin kontrolü altında bir geometri modeli oluşturur

Son senaryo (Şekil 2.12), senaryo II ve senaryo III'ün dezavantajlarını ele almaktadır. Prosedür senaryo II'deki ile aynıdır, ancak görüntüleyici arayüzler aracılığıyla filtreleme ve eşleme için parametreler belirleme imkanına sahip olmaktadır (Wood ve ark., 1996).

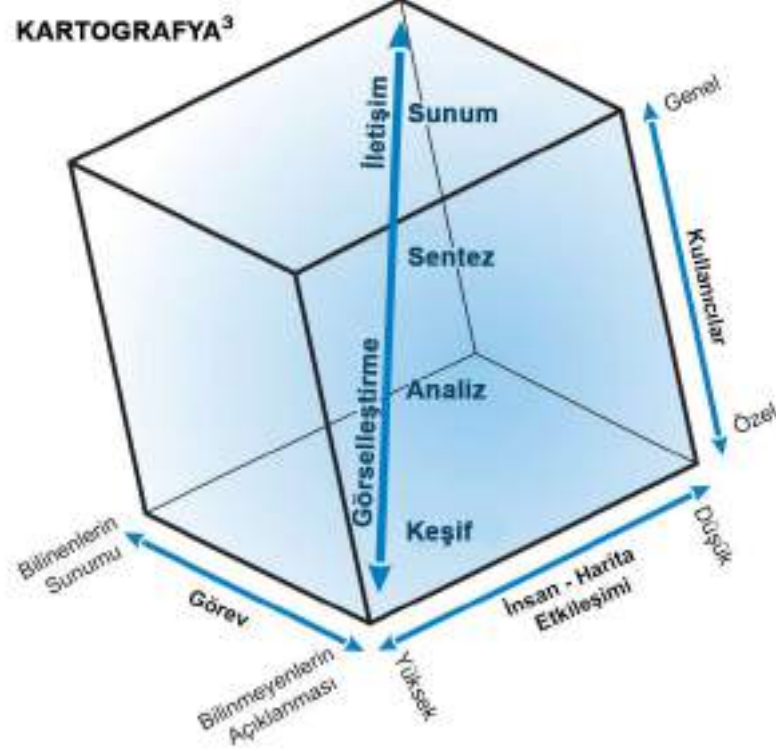


Şekil 2.12 Yayıncı, görüntüleyicinin kontrolü altında geometri modeli oluşturur (Wood ve ark., 1996)

2.6. Coğrafi Görselleştirme

Coğrafi görselleştirme, kâğıt ortamında, bilgisayar ekranında ya da diğer ortamlarda (görme ile ilişkili en güçlü insan bilgi işleme yeteneklerini kullanacak şekilde) mekânsal içerikleri ve problemleri görünür hale getiren somut görsel gösterimlerin kullanımı şeklinde tanımlanmaktadır (MacEachren vd., 1992; Slocum vd., 2008). MacEachren (1994) görselleştirmenin kapsamının anlaşılabilmesi için kullanışlı bir kavramsal araç olan kartografya küpünü sunmakta (Şekil 2.13) ve görselleştirmenin ayırt edici özelliklerini kapsamak için üç eksen kullanmaktadır. İlk eksen, coğrafi görselleştirme için kullanıcı kitlesinin kapsamını ifade etmektedir. Bu eksen aralığı sadece yapımcıları tarafından kullanılan özel haritalardan başlayarak, kamusal kullanım için yayınlanan genel haritalara kadar uzanmaktadır.

İkinci eksen haritanın kullanımda 'etkileşim' sunma derecesini göstermektedir. Üçüncü eksen ise bilinmeyen örüntülerin ortaya çıkarılmasından bilinen gerçeklerin sunulmasına kadar uzanan bir süreklilik gösteren veri ilişkilerini göstermektedir. Geleneksel anlamda tasarlanan haritaların çoğu, kartografya küpünün (genel, statik ve bilinen bilgileri sunan) köşesinde olduğu kabul edilmektedir. Görsel analitik araçlar ile tasarlanan haritalar ise genellikle karşı köşede bulunmaktadır. Bu haritalar yüksek seviyede insan-harita etkileşimi ile karakterize edilmekte ve genellikle özel verilerle ya da belirli alanlardaki uzmanlık verilerine göre tasarlanmaktadır (MacEachren, 1994).



Şekil 2.13 Kartografya küpü (MacEachren, 1994)

2.7. Görsel Analitik

Görsel analitiğin, çoklu süreçleri ve çeşitli uygulama alanlarını içeren çok disiplinli doğası nedeniyle tanımlanması kolay değildir. İlk dönemlerde “etkileşimli insan-makine ara yüzleri tarafından kolaylaştırılan analitik akıl yürütme bilimi” olarak tanımlanmaktadır. (Wong ve Thomas, 2004).

ABD Ulusal bilim Vakfı (NSF) ve Ulusal Sağlık Enstitüsü (NIH) tarafından hazırlanan “Illuminating the Path: the R&D Agenda for visual analytics” raporunda,

“Görsel analitik, etkileşimli görsel ara yüzlerin kolaylaştırdığı bir analitik akıl yürütme bilimidir. İnsanlar bilgileri sentezlemek ve büyük, dinamik, belirsiz ve çoğu zaman çelişen verilerden öngörü elde etmek için görsel analitik araçlarını ve tekniklerini kullanır; böylelikle beklenenleri tespit eder ve beklenmeyenleri keşfeder; zamanında, savunulabilir ve anlaşılır değerlendirmeler yapar ve eylem için değerlendirmeyi etkili biçimde iletir.” şeklinde tanımlanmaktadır. (Thomas ve Cook, 2005).

Görsel analitik, çok büyük ve karmaşık veri setleri temelinde etkin bir anlayış muhakeme ve karar verme için otomatik analiz tekniklerini etkileşimli görselleştirmeyle kombine edilmektedir. Bu bağlamda görsel analitik,

- Büyük, dinamik, belirsiz ve genellikle uyumsuzluk içindeki verilerden bilgileri sentezleyen ve öngörü kazandıran
- Beklenenleri tespit eden, beklenmeyenleri keşfeden
- Doğru zamanda, savunulabilir ve anlaşılabilir değerlendirmeler yapan
- Bu değerlendirmeleri eylem için etkin biçimde iletmeye,

sağlayan araçların ve tekniklerin oluşturulması şeklinde tanımlanmaktadır (Keim ve ark., 2010).

2.7.1. Görsel analitik alanındaki araştırmalar

Görsel analitik konusunun uluslararası alandaki önemi, IEEE Görselleştirme, IEEE Bilgi Görselleştirme Sempozyumu ve Eurographics / IEEE-VGTC Görselleştirme Sempozyumu gibi önde gelen uluslararası konferanslarda görsel analizin güçlü varlığıyla yansıtılmaktadır. Ekim 2006'da ilk IEEE Görsel Analiz Bilimi ve Teknolojisi Sempozyumu düzenlenmiştir. Görsel analitik de bu alanda büyük dergilerde yerini bulmuştur. Görsel analitikle ilgili özel konular IEEE Görselleştirme İşlemleri ve Bilgisayar Grafikleri, Bilgisayar ve Grafik, Coğrafi Bilgi Bilimi Uluslararası Dergisi ve SIGKDD Keşiflerinde yayımlanmıştır (Järvinen ve ark., 2009). Amerika Birleşik Devletleri'nde Ulusal Sağlık Enstitüsü ve Ulusal Bilim Vakfı "Görselleştirme Araştırma Zorlukları" raporunu yayınlamıştır (Johnson ve ark., 2006). ABD İç Güvenlik Bakanlığı (DHS), iç güvenlik için görsel analitik üzerine araştırma girişimi başlatmıştır. 2004 yılında kurulan "Ulusal Görselleştirme ve Analiz Merkezi" (NVAC) bu araştırmaları koordine etmektedir (Järvinen ve ark., 2009). ABD'de sınır güvenliği gibi güvenlik uygulamalarına odaklanan görsel analitik araştırma zorluklarını "Illuminating the Path" kitabında ortaya konmuştur (Thomas ve Cook 2005). Kanada'da, Görsel Analitik ile ilgili önemli araştırma girişimleri sürmekte ve Kanada hükümeti tarafından desteklenen ek görsel analiz programları Simon Fraser Üniversitesi'nden Brian Fisher tarafından hazırlanmaktadır. Avustralya'da, Sydney Üniversitesi'nden Peter Eades, ulusal bir görsel analitik araştırma girişimi başlatmak için girişimde bulunmuştur. Bu uluslararası girişimlerin tümü, görsel analitik konusunda Avrupa Koordinasyon Eylemine desteklerini belirtmektedir (Järvinen ve ark., 2009). Avrupa'da, görsel analitik ile ilgili birkaç ulusal girişim başlatılmıştır. Almanya'da, Alman Araştırma Vakfı DFG, Konstanz Üniversitesi'nden Daniel Keim tarafından yönetilen çalışma grupları konsorsiyumu tarafından Nisan 2007'de önerilen Ölçeklenebilir Görsel Analizler ile

ilgili stratejik bir araştırma girişimini kabul etmiştir. Program, bir dizi bireysel araştırma projesini destekleyerek ulusal görsel analiz araştırmasını ilerletmeyi hedeflemektedir. İsviçre'de, bilimsel görselleştirme ve görsel analitik için ulusal bir merkez kurma çabaları devam etmektedir. Avusturya'da son zamanlarda, Avusturya FIT-IT programı içerisinde özel olarak belirtilen görsel analitik ilişkileri olan Görsel Hesaplama üzerine bir araştırma girişimi başlatılmıştır (Järvinen ve ark., 2009).

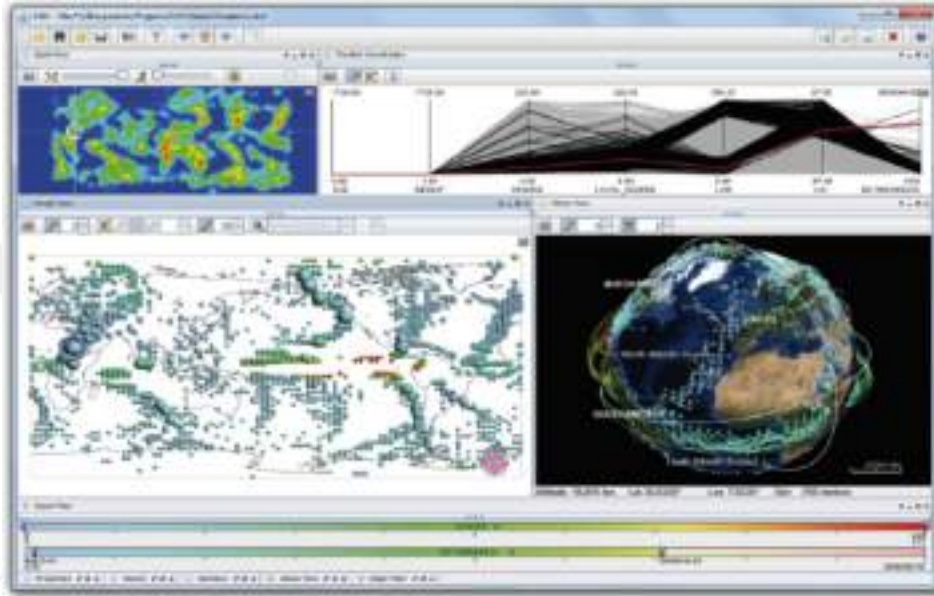
Veri madenciliği ve makine öğrenimi araştırmaları KNet, MLNet NoEs ve KDubiQ CA'nın odak noktasındadır. Coğrafyaya atıfta bulunulan hareketliliğin görselleştirilmesi, analizi ve yönetimi FET-Open projesi GeoPKDD'un odağını oluşturmaktadır. Afet ve acil durum yönetimi uygulamaları, OASIS ve ESS projelerinin odak noktasındadır. Bilgi alma ile ilgili görselleştirme problemleri, bu proje kapsamında yapılan kullanıcı ara yüzü çalışmasının bir parçası olarak DELOS NoE'de ele alınmıştır (Järvinen ve ark., 2009).

VisMaster Koordinasyon Eylem Projesi 2008'de başlamıştır. VisMaster'ın amacı, bir Avrupa görsel analitik girişimi için beklentileri kapsayacak şekilde kritik bir disiplinler arası Avrupalı araştırmacı kitlesini bir araya getirmektir (Keim, 2008).

2.7.2. Görsel analitik uygulama alanları

Görsel analitik, büyük bilgi alanlarının işlenmesi ve analiz edilmesi gereken uygulama alanlarında önemlidir. Başlıca uygulama alanları fizik ve astronomidir. Yaygın veri analizi teknikleriyle, gökbilimciler ilgili verileri gürültüden ayırabilir, benzerlikleri veya karmaşık kalıpları analiz edebilir ve evren hakkında faydalı bilgiler edinmektedirler. Görsel analitik yaklaşımı, standart algoritmik araçlarla bulunamayacak olan büyük ve dinamik veri akışlarında beklenmeyen olayları belirleme sürecini önemli ölçüde desteklemektedir. İklimin ve havanın izlenmesi, tüm dünya çapında aynı anda algılayıcılar tarafından uydulardan kısa zaman aralıklarında çok fazla miktarda verinin toplandığı bir alandır. Görsel bir yaklaşım, bu büyük miktardaki verilerin yorumlanmasına ve iklim faktörlerinin bağımlılıklarına ve böylelikle kolayca tespit edilemeyecek iklim değişikliği senaryolarına ilişkin ön görüş kazanmalarına yardımcı olmaktadır. Hava tahminlerinin yanı sıra, mevcut uygulamalar küresel ısınmayı, kutupların erimesini, stratosferik ozon azalmasının yanı sıra kasırga ve tsunami uyarılarını görselleştirmektedir (Keim ve ark., 2010).

Şekil 2.14'te çok etkileşimli bir grafik görselleştirme sistemi olan Coordinated Graph Visualization (CGV) tarafından sağlanan iklim modellerinin simülasyonu için görsel destek görülmektedir. Farklı görselleştirme görevlerini desteklemek için esnek bir yerleştirme çerçevesi yardımıyla görünüm toplulukları dinamik olarak oluşturulabilmektedir. CGV, sonsuz grid ve radar görünümü ile artırılmış navigasyonun yanı sıra gelişmiş dinamik filtreleme, grafik lensler, kenar tabanlı navigasyon içermektedir (Tominski ve ark., 2009).



Şekil 2.14 Coordinated Graph Visualization (CGV) ile gerçekleştirilen iklim modelleme simülasyonu (Tominski ve ark., 2009)

Acil durum yönetimi alanında, görsel analitik ile acil durumun ilerlemesi belirlenebilir ve olası hasarları sınırlamak için alınması gereken önlemleri (örneğin, fiziksel karşı önlemlerin veya nüfusun tahliyesi) belirlenebilir. Bu tür senaryolar, taşkın veya tsunami, volkanlar, fırtına, yangın veya hastalık salgınlarının büyümesi (örneğin N1H1 virüsü) gibi doğal veya meteorolojik felaketleri, ayrıca endüstriyel kazaları, nakliye kazaları veya kirliliği gibi insan kaynaklı teknolojik felaketleri içerebilir. Güvenlik ve coğrafi grafikler için görsel analitik önemli bir araştırma konusudur. Terörizm bilişiminden sınır korumaya, yol tespitinden ağ güvenliğine kadar bu sektördeki uygulama alanı geniştir. Görsel analitik, çok büyük veri kümelerinde benzerlik ve anormalliklerin araştırılmasını ve tespitini destekler. Örneğin, dünya çapında, günde 210 milyardan fazla e-posta, 4 milyar SMS, 90 milyon tweet ve IP veri paketi sayısı 9000 milyarı aşmış durumdadır. Avrupa düzeyinde bir belge işleme örneği olarak, Europe Media Monitor medya portalları, resmî Web siteleri ve haber ajansları

olmak üzere 2.500 haber kaynağından haber belgeleri toplamaktadır ve günlük olarak 80.000-100.000 makaleyi 43 dilde işlemektedir (Keim ve ark., 2010).

Şekil 2.15'te NFlowVis kullanılarak bir üniversite ağının SSH servisine dağılmış bir ağ saldırısının analizini gösterilmektedir. Arka plandaki TreeMap, iç ağ yapısını ve bu ağda yer alan ana bilgisayarları dikdörtgenler şeklinde göstermektedir. TreeMap dışında yer alan renkli noktalar ise dış ağda yer alan ana bilgisayarları temsil etmektedir. Hiyerarşik kenar demetleri, üst kısımdaki ana bilgisayarlardan gelen dağıtılmış saldırı gibi iletişim modellerini ortaya koymaktadır (Mansmann ve ark., 2009).



Şekil 2.15 Bir üniversite ağının SSH servisine dağılmış ağ saldırısının NFlowVis ile analizi (Mansmann ve ark., 2009)

Biyoloji ve tıpta, 3 boyutlu dijital rekonstrüksiyon ve görselleştirme için bilgisayarlı tomografi ve ultrason görüntüleme, gigabaytlar boyutunda tıbbi veri üretilmektedir. Biyo-bilişimin uygulama alanı, büyük miktarlarda biyolojik verileri analiz etmek için görsel analitik tekniklerini kullanmaktadır (Keim ve ark, 2010).

Görsel analitik için diğer bir önemli uygulama alanı, iş zekasıdır. Yüz binlerce varlığa sahip olan finansal piyasa, günlük olarak büyük miktarlarda veri üretmekte ve bu da yıllar içinde oldukça yüksek veri hacimleriyle sonuçlanmaktadır. Örneğin, günde 300 milyondan fazla kredi kartı işlemi olduğu tahmin edilmektedir. Bu alandaki asıl zorluk, verilerin tarihsel ve güncel durumlarını anlamak için çeşitli bakış açıları ve varsayımlar altında analiz etmek ve ardından eğilimleri tahmin etmek veya tekrar eden durumları belirlemek için piyasayı izlemektir. Bu alandaki diğer kilit uygulamalar

dolandırıcılık tespiti, tüketici verilerinin analizi, sosyal veriler ve sağlık hizmetleriyle ilgili verilerden oluşmaktadır.

2.7.3. Görsel analitik süreci

Görsel Analitik süreci, verilerden bilgi edinmek için otomatik ve görsel analiz yöntemlerini insan etkileşimi yoluyla sıkı bir bağlantıyla birleştirmektedir. Şekil 2.16'da görsel analiz sürecinde farklı aşamalara (ovalerle temsil edilir) ve bunların geçişlerine (oklar) soyut bir genel bakış gösterilmektedir (Keim ve ark., 2010).



Şekil 2.16 Görsel analitik süreci (Keim ve ark., 2010)

Birçok uygulama senaryosunda, görsel veya otomatik analiz yöntemlerinin uygulanabilmesi için heterojen veri kaynaklarının entegre edilmesi gerekmektedir. Bu nedenle, ilk adım genellikle daha ileri araştırmalar için farklı sunumlar elde etmek üzere verileri önceden işlemek ve dönüştürmektir. Diğer tipik ön işleme görevleri arasında veri temizleme, normalleştirme, gruplama veya heterojen veri kaynaklarının entegrasyonu yer almaktadır (Keim ve ark., 2010).

Dönüşümden sonra, analist görsel veya otomatik analiz yöntemlerini uygulamayı seçebilmektedir. Önce otomatik bir analiz kullanılırsa, orijinal verilerin modellerini oluşturmak için veri madenciliği yöntemleri uygulanmaktadır. Bir model oluşturulduktan sonra analist, modeli verilerle etkileşime sokarak yapılabilecek modeli en iyi şekilde değerlendirmeli ve geliştirmelidir. Görselleştirme, analistlerin parametreleri değiştirerek veya başka analiz algoritmaları seçerek otomatik yöntemlerle etkileşime girmesini sağlamaktadır. Model görselleştirme, üretilen modellerin

bulgularını değerlendirmek için kullanılmaktadır. Görsel ve otomatik yöntemler arasında geçiş yapmak, görsel analiz süreci için karakteristiktir ve ön sonuçların sürekli olarak iyileştirilmesine ve doğrulanmasına yol açmaktadır. Yanıltıcı sonuçlar, ara bir işlemde erken bir aşamada keşfedildiği takdirde daha iyi sonuçlar ve daha yüksek bir güven sağlanmaktadır. Öncelikle görsel veri keşfi yapılırsa, kullanıcı oluşturulan hipotezleri otomatik bir analizle doğrulaması gerekmektedir. Anlayışlı bilgileri ortaya çıkarmak için görselleştirme ile kullanıcı etkileşimi (örneğin, farklı veri alanlarını yakınlaştırarak veya veriler üzerinde farklı görsel görünümüler göz önünde bulundurarak) gerekmektedir. Görselleştirmelerdeki bulgular, otomatik analizde model oluşturmayı yönlendirmek için kullanılabilir (Keim ve ark., 2010).

Görsel analiz süreci, otomatik analiz yöntemlerini ve etkileşimli görsel sunumları bir araya getirmeyi amaçlamaktadır (Keim ve ark., 2010). Shneiderman verilerin ekranda nasıl gösterilmesi gerektiğini

*“Öncelikle geneline bak,
yakınlaştır ve filtrele,
talep üzerine detaylandır.”*

şeklinde açıklamaktadır (Shneiderman, 1996). Ancak, eldeki büyük veri kümeleriyle, ilginç örüntüleri kaybetmeden genel bir görselleştirme oluşturmak zordur. Bu işlemle kullanıcılara daha fazla nelerin inceleneceği hakkında çok az bilgi verildiğinden, yakınlaştırma ve filtreleme tekniklerini etkin bir şekilde gereksiz kılmaktadır. Görsel analizler bağlamındaki bu rehber

*“Öncelikle analiz et,
önemli olanı göster,
yakınlaştır / filtrele,
daha fazla analiz et,
talep üzerine detaylandır.”*

olarak genişletilebilir. Görselleştirmede yalnızca veriyi almanın ve görüntülemenin yeterli olmadığı; bunun yerine verilerinin kullanıcının ilgisine göre analiz edilmesi, verilerin en alakalı yönlerinin gösterilmesi ve aynı zamanda kullanıcının talebi üzerine verilerin detaylarının sağlandığı etkileşim modellerinin sunulması gerekmektedir (Keim ve ark., 2006).

2.8. Mekânsal Görsel Analitik

Mekânsal görsel analitik, mekânsal bilgilerin etkileşimli görsel ara yüzlerle ele alındığı analitik akıl yürütme bilimidir (Andrienko ve ark., 2007). Mekânsal Görsel Analitik, 2000'lerin sonunda görsel analitik alanının gelişmesiyle ortaya çıkmıştır (Thomas ve Cook, 2005). Mekânsal görsel analitik, coğrafi mekân ve bu mekândaki nesnelere, olaylar, olgular ve işlemleri içeren problemleri ele alır (Andrienko ve ark., 2011).

Mekânsal görsel analitik, kullanıcıların mekânsal verileri kullanarak örüntüleri tespit etmeleri ve gelecekteki sonuçları tahmin etmeleri için kartografya, hesaplamalı yöntemler, ara yüz tasarımı, biliş bilimi vb. alanların her birini etkileşimli haritalama sisteminde birleştirmektedir. Mekânsal görsel analitik, büyük veri ile ortaya çıkan hacim, hız, çeşitlilik, doğruluk ve diğer yönlerden kaynaklanan zorluklardan büyük ölçüde etkilenmektedir (Robinson, 2017). Mekânsal görsel analitikte özellikle mekânsal-zamansal problemlerin analizi ile ilgili zorluklara dikkat edilmelidir (Andrienko ve ark., 2011).

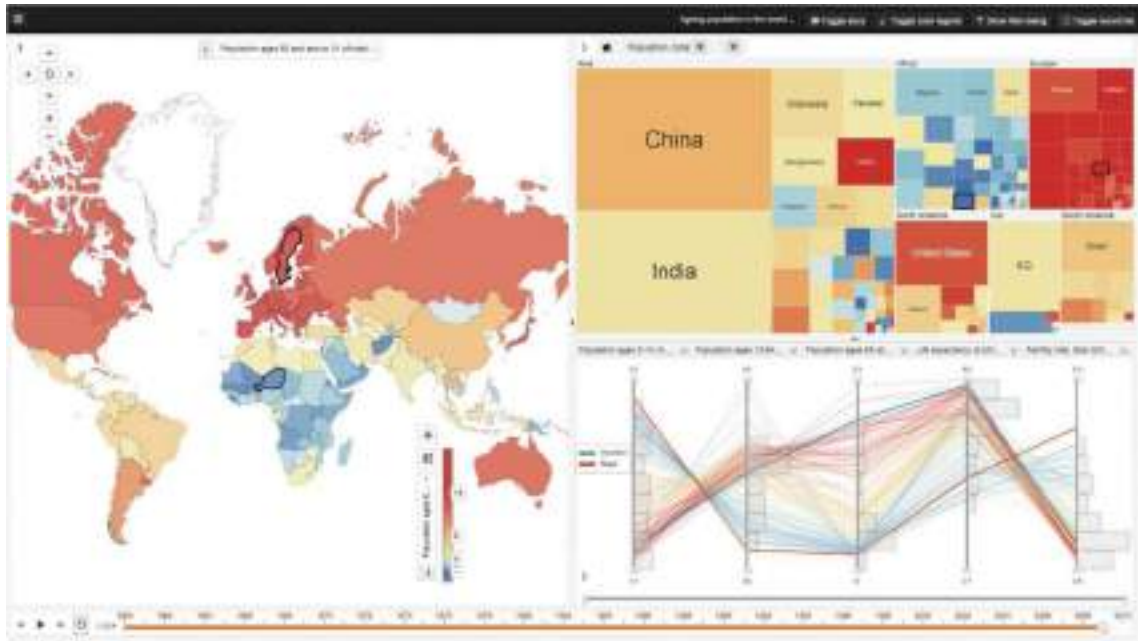
2.8.1. Analitik akıl yürütme

Mekânsal Görsel Analitik biliminin en önemli farklılaşma noktası, bilginin içindeki örüntüleri bulmak için bilgi inceleme süreci olan analitik akıl yürütmeyi temel almasıdır. Mekânsal görsel analitik, öncelikle görselleştirmeye odaklanmak yerine, insanların görsel ara yüzleri kullanarak coğrafi bilgileri analiz etmelerini ve anlamlandırmalarını destekleme konusunda bilimsel gelişmeleri hedef almaktadır. Bu nedenle yeni yaklaşımlar sadece coğrafi bilgilere görsel ara yüzler sağlamak için değil, coğrafi bilgi ile akıl yürütme ve algılama mekanizmaları içinde geçerlidir. Mekânsal görsel analitikte analitik akıl yürütme desteği genellikle örüntüleri tespit etmek ve/veya gelecekteki sonuçları tahmin etmek için kullanılan hesaplama yöntemleri şeklinde gelmektedir. Analitik akıl yürütme, analitik süreçleri destekleyecek şekilde uyarlanmış coğrafi bilgilere ilişkin ara yüzlerde ve analistlerin işlerini yürütmek için kullandıkları iş akışlarında da görülebilir. Bu nedenle mekânsal görsel analitik araştırmaları, örüntü tespiti için yeni nicel yöntemlerin geliştirilmesinden (Shekhar ve ark., 2011) analist algılama süreçlerinin kalitatif çalışmalarına (Pirolli ve Card, 2005) kadar analitik spektrumun her iki tarafını da içermektedir (Robinson, 2017).

2.8.2. İnteraktif coğrafi görselleştirme

Mekânsal görsel analitik araçlarında bulunan görselleştirmeler tematik haritalar, dağılım grafikleri, paralel koordinat grafikleri, zaman çizelgeleri, uzay-zaman küpleri ve çok çeşitli diğer gösterim tekniklerini içermektedir.

Şekil 2.17'deki GAV (GeoAnalytics visualization) araç seti, çok değişkenli mekânsal zamansal verileri keşfetmek için koordineli bir şekilde kullanılabilen çok çeşitli görselleştirme bileşenleri içeren mekânsal görsel analitik sistemidir. Bu örnekte, ülkelere göre küresel nüfus eğilimleri, koroplet harita, ağaç haritası (treemap) ve paralel koordinat grafiği kullanılarak karşılaştırılmıştır (Robinson, 2017).



Şekil 2.17 GAV (GeoAnalytics visualization) araç seti (Robinson, 2017)

Etkileşim, genellikle bir görünümdeki değişikliklerin eş zamanlı olarak diğer kullanılabilir görünümde yayınlanmasına izin vererek desteklenmektedir. Bu tür sistemler genel olarak **koordineli görünüm görselleştirmeleri** olarak tanımlanmaktadır (Robinson, 2017).

Mekânsal Görsel Analitik sistemleri genellikle yüksek seviyede etkileşimi desteklemektedir (Roth, 2013). Coğrafi verilere uygulanan gösterim yöntemleri ve karmaşık veri kümelerinden örüntü çıkarmada kullanılan hesaplama teknikleri için arayüz kontrolleri sağlanmaktadır. Mekânsal görsel analitikte etkileşime odaklanmak, kullanıcıların gördüklerini değerlendirmek ve düzeltmek için yinelemeli katılımı desteklemeyi, böylelikle tümdengimsel ve tümevarımsal muhakeme biçimlerini mümkün kılmayı amaçlamaktadır (Gahegan, 2005). Etkileşim, mekânsal

görselleştirmenin yükselmesinin ardında merkezi bir faktör ve günümüzde Mekânsal Görsel Analitik alanının selefidir (Robinson, 2017). Mekânsal Görsel Analitikte etkileşim, görsel anlatımı değiştirme yeteneğinin yanı sıra, temel anlatım yöntemlerini ve hikâyeleştirmeyi desteklemeye yönelik araçları da içerebilmektedir (Van Ho ve ark., 2012).

2.8.3. Hesaplamalı yöntemler

Mekânsal görsel analitik sistemlerinde, coğrafi mekânsal verilerdeki örüntü ve kümeleri tanımlamak için **veri madenciliği, keşifsel veri analizi, hesaplamalı hareket analizi** vb. teknikler kullanılmaktadır. Ayrıca, analistlerin mekânsal görsel analitik sistemlerini daha önce bilinen hedef örüntüye dayalı örnekleri aramak için örüntü eşleştirme araçlarından faydalanması mümkündür. Bu yaklaşımların her biri, etkileşimli bir görselleştirme ortamına dahil edilebilecek mekânsal analiz yöntemlerinin geliştirilmesine dayanmaktadır. Bu gereklilikten çeşitli bilimsel ve teknolojik sorunlar ortaya çıkmaktadır. Örneğin, hesaplama yöntemlerinin verileri gerçek zamanlı etkileşimini ve analitik parametrelerin manipülasyonunu sağlayacak şekilde hızlıca işlemesi önem arz etmektedir. (Robinson, 2017).

Mekânsal görsel analitikte hesaplama yöntemleri, küme tespiti, coğrafi ağırlıklı regresyon, kendi kendini düzenleyen haritalar, adlandırılmış varlık tanıma ve diğer birçok yaklaşımı içermektedir. Bu yöntemlerin uygulanmasındaki temel hedef, analitik akıl yürütmeyi mekânsal bilgilerle desteklemektir (Robinson, 2017).

2.8.4. Teknik gelişmeler

Mekânsal Görsel Analitik sistemlerin gelişimi, teknolojik gelişmenin birkaç aşamasından geçmiştir. İlk olarak Masaüstü bilgisayar ortamları için tasarlanmıştır. Daha sonra genellikle Java ve ActionScript dilleri kullanılarak yazılmış, Web tabanlı görselleştirme uygulamaları ortaya çıkmıştır. Bugün, geliştiricilerin seçebileceği birçok uygulama yazılımları bulunmaktadır. Örneğin açık kaynak kodlu ve özel Coğrafi Bilgi Sistemi platformlarına dayalı özel uygulamalar oluşturmak ve sunmak için özel uygulama programı arabirimlerini (API) kullanılmaktadır. Ayrıca hesaplama adımlarını ve veri depolama için bulut tabanlı hizmetler yaygın olarak kullanılmaktadır. Günümüzde Web üzerinden görselleştirme tasarımında, PostGIS gibi mekânsal veri tabanlarıyla birlikte Data Drive Documents (D3) kütüphanesi ve tamamen işlevsel bir mekânsal görsel analitik ortamı formüle etmek için R gibi istatistiksel analiz

kütüphaneleri çoğunlukla kullanılmaktadır. Tableau ve Insights for ArcGIS gibi ticari araçlar, ara yüz, veri tabanı ve hesaplama ihtiyaçlarını tek bir bütünleşmiş platformda ele alan birleştirilmiş sistemler sağlamaktadır (Robinson, 2017).

2.8.5. Mekânsal görsel analitikte kullanıcı merkezli tasarım

Mekânsal görsel analitikte temel düşünce, araçların yüksek derecede fayda ve kullanılabilirliği desteklemek için kullanıcı merkezli tasarım ilkelerini yansıtmasını ve görüntüleyicinin değerlendirmeleri ile yinelemeli olarak tasarlanmasını sağlamaktır. Mekânsal görsel analitik sistemlerinin **faydası** ve **kullanılabilirliği** arasında önemli ayırım yapılmaktadır. **Kullanılabilirlik**, belirli bir sisteme ilişkin genel kullanım kolaylığını ifade ederken, **fayda** bir aracın belirli sorunları çözme yeteneğini ifade etmektedir. Araçların yüksek derecede faydalı ve düşük seviyede kullanılabilir olması ya da bunun tam tersi mümkündür. Analitik akıl yürütmeyi desteklemek için, kullanıcıların gerçekleştirmek istedikleri görevler belirlenmeli ve bu görevleri kolay ve verimli bir şekilde destekleyen ara yüzler tasarlanmalıdır. Mekânsal görsel analitikte kullanıcı merkezli yaklaşım, temel kullanıcı gereksinimlerini tanımlamak için **ihtiyaç değerlendirmesinden** başlayarak, kullanıcı ihtiyaçlarını karşılayan ve etkin bir şekilde çalışan son bir tasarıma ulaşılan kadar döngüsel olarak prototip değerlendirmesiyle devam ettirilmelidir (Robinson ve ark., 2005).

2.9. Sosyal Medya

Web 2.0 teknolojisinin gelişimiyle ortaya çıkan sosyal medya platformları, iletişimin karşılıklı ve etkileşimli bir şekilde sağlandığı kullanıcı odaklı içerik üretiminin gerçekleştiği alanlardır. Sosyal medyanın insanoğlu'nun hayatına entegrasyonu birlikte, insanların iletişim kurma biçimleri, paylaşma alışkanlıkları ve yaşam tarzları ciddi şekilde değişiklik göstermektedir. Teknolojik gelişmelerin ürünü olan sosyal medya platformları, iletişimi desteklemenin yanında, haberler erişim, alışveriş, eğlence, arkadaşlık ağları oluşturma vb. amaçlara da hizmet vermektedir.

Sosyal medya, “Tüm dünyadaki bilgisayar ağlarını ve kurumsal bilgisayar sistemlerini internet teknolojisiyle bir araya getirerek, kullanıcılara farkı sanal ortamlarda iletişim kurmalarını sağlayan internet tabanlı hizmetler” şeklinde tanımlanmaktadır. (Kuşay, 2010).

Mayfield (2008) sosyal medya'nın beş temel özelliğini şu şekilde açıklamıştır;

Katılım: Her bireyin içeriğe katkıda bulunabilme fırsatını sağlayan bir platform sunma niteliğini ifade etmektedir. Kullanıcıların birbirleriyle etkileşimde bulunması ve kendilerinde katkıda bulunmaları sosyal medya içeriğini daha zenginleştirmektedir.

Açıklık: Sosyal medya, temelde açık katılım prensibi üzerine kurulmuştur. Ana hedef, etkileşimin artırılması için yorum yapmak ve içerik paylaşmaktır. Bu nedenle, içeriğe erişimi engellemek, kısıtlamak veya şifrelemek gibi kısıtlayıcı yaklaşımlar genellikle tercih edilmemektedir.

Sohbet: Sosyal medya etkileşimli bir iletişime hizmet etmesi nedeniyle tek yönlü iletişim sağlayan geleneksel medyadan ayrılmaktadır.

Topluluk: Sosyal medya, toplulukların hızlı ve etkili bir şekilde bir araya gelmesine imkân tanımaktadır. Bu sayede topluluklar, tercih ettikleri fotoğrafları, siyasi görüşlerini, en sevdikleri televizyon programlarını ve kişisel zevklerini serbestçe paylaşabilmektedirler.

Bağlantısallık: Sosyal medya, çeşitli linkler aracılığıyla farklı internet sitelerine erişim imkânı sunmaktadır. Bu linkler, çeşitli topluluklar veya kaynaklara kolayca erişimi sağlayarak sosyal medyanın içeriğinin daha zengin hale gelmesini sağlayan önemli öğelerden birini oluşturmaktadır (Mayfield, 2008).

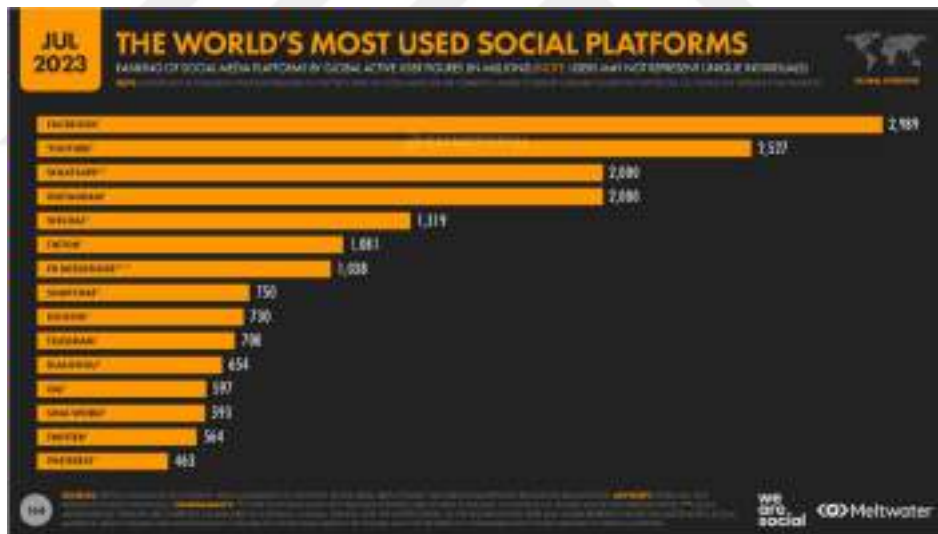
İnternet kullanımı ve sosyal medya kullanımına ilişkin dijital raporlar her yıl hazırlanarak kullanıcıları erişimine açılmaktadır. Bu raporlar, ülkesel ve dünya genelinde insanları sosyal medya kullanımlarına ilişkin bilgileri, sosyal medya platformlarının kullanım oranlarını, sosyal medyayı kullanımlarının nedenleri vb. konular hakkında istatistiki bilgiler yer almaktadır.

2023 yılına ait dijital rapora (URL-3) göre, sosyal medya kullanıcılarının dünya genelindeki sayısı 4.88 milyar olarak belirlenmiştir. 2012 yılından itibaren 173 milyon yeni kullanıcının sosyal medya kullanmaya başladığı görülmektedir. Bir sosyal medya kullanıcısının her ay ortalama 6,7 farklı sosyal medya platformunu kullanarak, günde ortalama 2 saat 26 dakikasını sosyal medyada kullandığı görülmektedir. Sosyal medya kullanıcılarının %53,6'lık oranının erkek, %46,4'lük oranının kadın kullanıcılardan oluştuğu da raporda görülmektedir (Şekil 2.18).



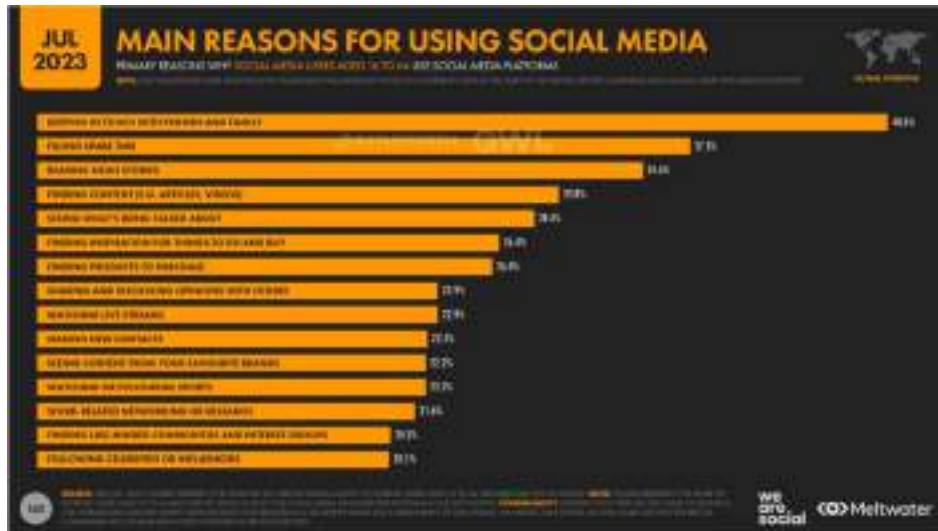
Şekil 2.18 Sosyal medya kullanımına genel bakış (URL-3)

2023 yılına ait dijital rapora (URL-3) göre, Facebook'un aylık 2.989 milyar, Whatsapp'ın aylık 2 milyar, Instagram'ın aylık 2 milyar aktif kullanıcılarının olduğu görülmektedir. Youtube'un reklam erişim potansiyeli'nin 2.527 milyar olduğu da aynı raporda (URL-3) yer almaktadır (Şekil 2.19).



Şekil 2.19 Dünya'da en çok kullanılan sosyal medya platformları (URL-3)

14 – 64 yaş aralığındaki insanların sosyal medya'yı kullanmalarındaki ana sebepler içerisinde; %48,6'lık oranla aileleri ve arkadaşları ile iletişim sağlamak olurken, %37 gibi azımsanamayacak bir oranın boş vakit geçirmek için sosyal medyayı kullandıkları aynı raporda (URL-3) görülmektedir (Şekil 2.20)



Şekil 2.20 Sosyal medya kullanımının ana sebepleri (URL-3)

Sosyal medyanın temel bileşenleri, kullanıcılar ve bu kullanıcıların sergilediği davranışlardır. Sosyal medya kullanıcıları, tarih damgaları, metin, medya ve muhtemelen coğrafi bilgiler gibi unsurları içeren iletleri oluşturabilir ya da başkaları tarafından gönderilen iletleri kendi hesaplarında paylaşabilmektedirler. Bu eylemler, sosyal medya platformlarında bilginin yayılmasını sağlamaktadır. Kullanıcılar tarafından üretilen içerik çevrimiçi sosyal iletişim aracılığıyla yayılmakta ve tüm sosyal medya verileri kaydedilmektedir. Devasa bir hacme kavuşarak Büyük Veri formuna dönüşen Sosyal medya verileri içerisinde ilk bakışta fark edilemeyen birçok örüntü barındırmaktadır. Bu nedenle birçok araştırmacının temel hedef noktasında sosyal medya verileri yer almaktadır (Chen ve ark., 2017).

Sosyal medya verilerinin en önemli özelliklerinden biri mekânsal ve zamansal bilgiyi içermesidir. Sosyal medya platformlarında, kullanıcılar tarafından paylaşılan iki tür mekânsal bilgi bulunmaktadır. Bunlardan ilki, kullanıcıların genellikle yaşam yerlerini ifade eden kaba coğrafi özellikli veridir. İkincisi ise kesin koordinat bilgilerine sahip coğrafi etiketlenmiş iletlerdir. Araştırmacılar mekânsal sosyal medya verilerinde yer alan mekânsal bilgileri kullanarak mesajların farklı şehirler, bölgeler ya da ülkeler arasında yayılımlarını ve tartışılma durumlarını analiz edebilmektedirler (Cao ve ark., 2012; Zhang ve ark., 2013). Sosyal medya platformlarının, COVID-19 pandemisi gibi afet durumlarında etkin yaygın olarak kullanılması şaşırtıcı olmamaktadır. Sosyal medya platformları, belirsizliği azaltıcı, hızlı ve güvenilir bilgi akışı sağlayarak kriz anlarında vatandaşların özgüvenini etkileyebilmektedir (Jurgens ve Helsloot, 2018).

2.10. Eski Adıyla Twitter, Yeni Adıyla X

Jack Dorsey ve ark. tarafından ilk olarak 2006 yılında mikro blog uygulaması şeklinde ortaya çıkarılan Twitter platformu, kullanıcılarına gündelik hayatı ile ilgili düşüncelerini, yaşadığı toplumu ilgilendiren siyasal, ekonomik, askeri, sağlık, eğitim vb. olaylarla ilgili düşüncelerini 280 karakterden oluşan “tweet” adı verilen mesajlarla özgürce ifade edebildikleri ve birbirleriyle etkileşim halinde oldukları sosyal ağ platformudur. Twitter uygulaması, kullanıcılara hikâye ve fotoğraf paylaşımıyla birlikte diğer sosyal uygulama içeriklerinin de platform içerisinde kullanılması ve paylaşılmasına olanak tanımaktadır (URL-2; Öztürk, 2022).

Tesla ve SpaceX'in CEO'su Elon Musk, Twitter'ı 44 Milyar dolara satın alarak 23 Temmuz 2023'te attığı tweet ile, *“Yakında Twitter markasına ve yavaş yavaş tüm kuşlara veda ediyoruz”* ifadelerini kullanarak Mavi kuş logosunun “Siyah X” logosu ile değiştirildiğini ve Twitter isminin X olarak değiştirildiğini duyurmuştur. Elon Musk, 21 Şubat 2023 tarihinde attığı *“Algoritmamız gelecek hafta açık kaynak haline geldiğinde ilk başta hayal kırıklığına uğramaya hazır olun, ancak hızla gelişecek!”* tweeti ile platformun kaynak kodlarını açık kaynak haline getireceğini duyurmuş ve kaynak kodların bir kısmını Github'da yayımlamıştır. Bu sayede, isteyenler şirketin tescilli yazılımını inceleyebilecek, Twitter'ın kodunu değiştirmek için geliştiricilere fikir sunabilecek veya algoritmayı kendi uygulamalarında kullanabileceklerdir (URL-33; URL-34; URL-35).

Günümüzde “kişisel medya” aracı olarak tanımlanan X platformu, kullanıcılarının öz denetiminin ön planda olduğu, kendisini takip edecek kullanıcı hesaplarını belirleyebildiği, kendini takip eden hesaplar hakkında bilgi alabildiği, ne kadar takipçi hesabı olduğunu görebildiği sosyal platformdur (Lüders, 2008; Schmidt, 2016).

Gelişen teknolojiyle birlikte mobil cihazların hayatımızın bir parçası olmasıyla birlikte kullanıcıların X platformuna erişimleri ve kullanımları aynı doğrultuda yükseliş göstermektedir. 2023 yılına ait dijital raporda (URL-3), X'de (Eski adıyla Twitter) 372,9 milyon kullanıcıya ulaşabileceği görülmektedir. Tüm dünyadaki insanların %4,6'sının X'i (Eski adıyla Twitter) kullandığı, X (Eski adıyla Twitter) kullanıcılarının %35,7'sinin kadın kullanıcı ve %64,3'ünün erkek kullanıcı olduğu, 13 yaş üzeri nüfusun %5,9'unun X (Eski adıyla Twitter)'ı kullandığı görülmektedir (Şekil 2.21).



Şekil 2.21 X (Eski adıyla Twitter) kullanım istatistikleri (URL-3)

X platformunun doğal afetleri ve hastalık salgınlarını tespit etmek için iyi bir platform olduğu bilinmektedir. Doğal afet ya da hastalık salgınından etkilenen kullanıcıların konumlarını doğru bir şekilde çıkarabilme yeteneği hayat kurtarabilmekte ve kriz yönetimine yardımcı olabilmektedir (Ajao ve ark., 2015).

X platformu, kullanıcılarına birincil adresleri olan yaşadıkları şehir ismini X kullanıcı hesabı kaydı sırasında metinsel bilgi olarak girme imkânı sağlamaktadır. Ancak gerçekte kullanıcıların %14'ünden daha azı doğru ya da eksiksiz olarak tamamlamaktadır (Ajao ve ark., 2015). Hecht ve ark. (2011), X kullanıcılarının %34'ünün bu alan yanlış veya hayali yer adları verdiklerini keşfetmişlerdir. Bu alan metinsel içerikli alan olduğundan, X kullanıcılarının bu alana girebilecekleri metinsel ifadeleri düzenleyememekte ve çeşitli kurallarla zorlayamamaktadır (Hecht ve ark., 2011).

X, kullanıcılarının deneyimini geliştirmek için konum koordinatlarının tweet'lere meta veri olarak eklenmesine izin vermektedir. Kullanıcının gönderdiği tweetlere o anki konumu ekleyebilmekte, bu işleme de coğrafi etiketleme (geotagging) denilmektedir. Coğrafi etiketli mesajlar, kullanıcının mevcut konumu hakkında ya da bir tweet'in kaynağı hakkında belirli bir mesafeye kadar doğru bir tahmin verebilmektedir (Ajao ve ark., 2015). Benzer şekilde, günümüzde üretilen akıllı cihazların neredeyse tamamı entegre GPS ile gelse de X kullanıcılarından %0,5'inden daha azı gizlilik, siber zorbalık ve takip etme konularında duydukları endişe nedeniyle akıllı cihazlarının konum tabanlı hizmetlerini açmaktadırlar. (Li ve ark., 2012). Diğer kullanıcılar ise, güç tasarrufu

yapmak ve pillerinin hızla bitmesini önlemek için konum tabanlı hizmetlerini kapatmaktadırlar (Lin ve ark., 2010).

2.10.1. Mekânsal özellikler ve göstergeler

X kullanıcılarının konumunu çıkarmaya yardımcı olan çeşitli göstergeler (Şekil 2.22) farklı araştırmalarda kullanılmaktadır (Ajao ve ark., 2015).



Şekil 2.22 X kullanıcılarının konum göstergeleri (Ajao ve ark., 2015).

Tweet Metnindeki Konum Belirteci: Tweet mesaj metnini işlenmesindeki zorluğun sebebi, tweet'lerin dilbilgisi kuralları ve anlamsal bütünlük ile yazılan makalelerin aksine, anlamlandırılmış yapıda ve herhangi bir dilbilgisi kuralına uymadan oluşturulmalarından kaynaklanmaktadır. Tweet'lerin içerisinde belirli kurallar uyan kısaltmalar ve standart olmayan kısaltmalar da bulunmaktadır. Tweet'lerin çoğu mobil cihazlar üzerinden gönderildiği için kullanıcılarının yazım hataları ve kısaltmalar konusunda geniş bir hareket alanı bulunmaktadır. Örneğin, Birleşik Krallık'ın kısaltması 'UK', 'GB', 'GBR' ya da 'GR8 Britain' şeklinde yazılmaktadır. İyi bir içerik analizi yaklaşımı, mesaj içinde ifade edilen konum varlıklarının tüm olası durumlarını dikkate alacaktır. Mesajların içinde yer alan konumlar belirlense bile, bunun otomatik olarak kullanıcı konumunu ya da tweet'in konumunu ifade etmeyeceğini bilinmesi gerekmektedir. Örneğin, bir tweet'in 'Ankara' adını içermesi bu tweeti atanın Ankara'da yaşadığı ya da bu tweeti Ankara'da attığı anlamına gelmemektedir. (Ajao ve ark., 2015; Ikawa ve ark., 2013).

Arkadaş Ağı: Bir X kullanıcısının takipçilerinden aynı lokasyonda bulunanların, o kullanıcının konumu için iyi bir mekânsal gösterge olduğu varsayılmaktadır. İki X kullanıcısının aynı şehirde ikamet etmeleri halinde sık sık iletişim kurmalarının muhtemel olduğunu düşünülmektedir (Jurgens, 2013; McGee ve ark., 2011).

Kullanıcı Profil Bilgileri: Bir X kullanıcı hesabının oluşturulması aşamasında verilen konum bilgileri, reklamcılarının müşterilerini doğru bir şekilde hedefleyerek, onların konumlarıyla alakalı reklam yönlendirmeleri yapmalarını sağlayabilmektedir.

Acil servislerin ve ilk müdahale ekiplerinin bir kriz veya afet konumunu kolaylıkla tespit etmelerine yardımcı olmaktadır. Genellikle konum alanı, kullanıcıların şehir isimlerini manuel olarak yazmalarına olanak tanıyan metinsel formatta bulunmaktadır. Konum ifadeleri genellikle İstanbul, Türkiye gibi şehir ve ülke düzeyinde olmaktadır. Bununla birlikte, 'Ev', 'İş', 'Deniz' ve 'Güneş' gibi anlamsız ifadeler, doğal dil işleme teknolojilerini ve makine öğrenimi algoritmalarının konum varlıklarını etkili bir şekilde çıkarmasını zorlaştırmakta, hatta yanıltıcı sonuçlar vermesine sebep olmaktadır (Ajao ve ark., 2015).

Coğrafi Etiketler: Günümüzde akıllı telefonların çoğu bütünleşik GPS teknolojisiyle donatılmıştır. Entegre GPS sistemi sayesinde mobil kullanıcının coğrafi konumu (enlem/boylam koordinatları) doğru bir şekilde belirlenebilmektedir. Coğrafi etiket (geotag) özelliği genellikle kullanıcıların gizlilik endişeleri duymaları sebebiyle isteğe bağlı etkinleştirilen bir özelliktir. Kullanıcıların %0,5'ten daha azının konum hizmetlerini açtığı düşünülmekte, bu da coğrafi etiketleme işlemini zorlaştırmaktadır. Bu gösterge, kullanıcının mobil olduğu ve konum bilgisini sık sık güncellediği durumlarda çok yararlı olmaktadır (Ajao ve ark., 2015; Li ve ark., 2012).

URL Linkleri: Bazı çalışmalarda, metin gövdesi içindeki URL bağlantıları kullanıcıların konumunu çıkarmak için mekânsal gösterge olarak kullanılmıştır. Ücretsiz bir çevrimiçi sorgu hizmeti olan InfoDB, adresleri ve alan adlarını coğrafi konum ile eşleştirmektedir. InfoDB veri tabanına IP adresleri girilerek ülke düzeyinde konum sorgusu gerçekleştirilmektedir (Schulz ve ark., 2013).

Üçüncü Parti Yazılımlar: Konum tabanlı sosyal medya sitelerinin popülerliği, Coğrafi sosyal ağ olarak da adlandırılan etkileşim araçlarını mümkün kılmaktadır. Foursquare ve Yelp uygulamaları, şirketlere, küçük işletmelere ve restoranlara coğrafi bir veri tabanına dahil olarak gerçekleştirilecek sorgulamalarda listelenmelerini sağlayan başarılı uygulamalardır. Çevrimiçi kullanıcılar, ilgilendikleri bir yerin, örneğin İstanbul'daki bir lokantanın konumunu yalnızca çevrimiçi dizinlerinde arama yaparak bulabilmektedir. Foursquare, kullanıcılarının X hesaplarını genellikle coğrafi etiketli olan Foursquare gönderilerine bağlamalarına izin vererek, X'de konum tabanlı hizmet servislerine izin vermemelerine rağmen Foursquare mesajı gönderisinden konumlarını çıkarmalarına olanak tanımaktadır (Li ve Sun, 2014; Ajao ve ark., 2015).

Zaman Dilimleri: Tweetlerdeki meta veriler genellikle mesajın zaman damgasını ve X API tarafından yakalanan saat dilimini içermektedir. Zaman dilimleri, konumun en azından ülke düzeyinde ayrıntıya kadar çıkarılmasına olanak tanımaktadır. Mesaj metninin içinde konum bilgisinin kısıtlı bulunduğu durumlarda faydalı olmaktadır (Mahmud ve ark., 2014).

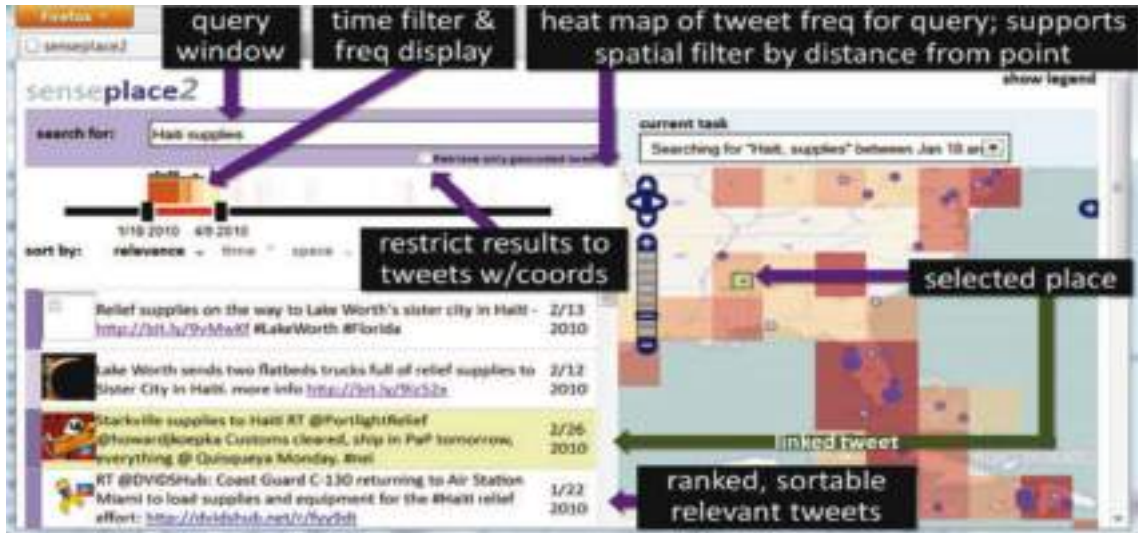
2.11. Araştırma Konusu Kapsamında Daha Önceden Gerçekleştirilen Çalışmalar

Furtado ve ark. (2010), Şekil 2.23'te suç olaylarını kaydetmek ve araştırmak için haritaların kullanımına ve doğrudan manipülasyonuna dayalı işbirlikçi ortam sunan Wiki Crimes isimli Web 2.0 uygulaması geliştirmişlerdir. Wiki Crimes uygulaması ile suç bilgilerine daha fazla şeffaflık ve aleniyet kazandırmak, vatandaşlara önleme araçları sağlamak ve eksik raporlama olgusunu azaltmak amaçlanmıştır (Furtado ve ark., 2010).



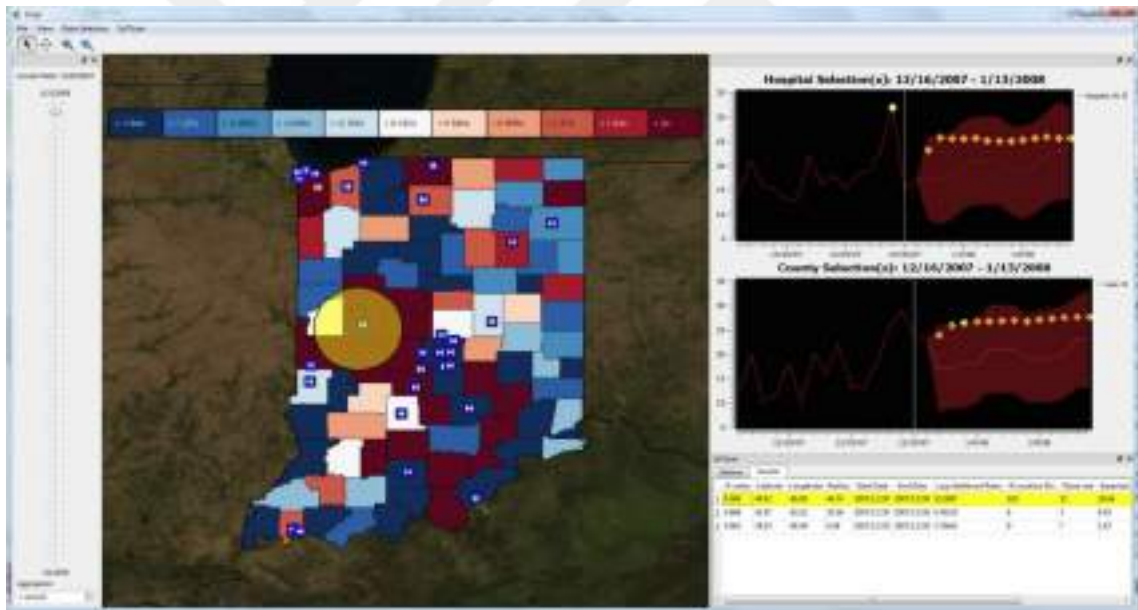
Şekil 2.23 Fortaleza'daki cinayetlerin haritası ile WikiCrimes ana arayüzüne genel bakış, nokta haritaları ve ana alanlarını açıklayan metinler (Furtado ve ark., 2010)

Mekânsal görsel analitik'in çeşitli alanlarda uygulamaları bulunmaktadır. MacEachren ve ark. (2011), Şekil 2.24'te geliştirdikleri SensePlace2 adlı uygulamada kriz yönetimi alanında X mesajlarını kullanarak coğrafi temelli durumsal farkındalık ile ilgili mekânsal görsel analitik yaklaşımı sunmuşlardır (MacEachren ve ark., 2011).



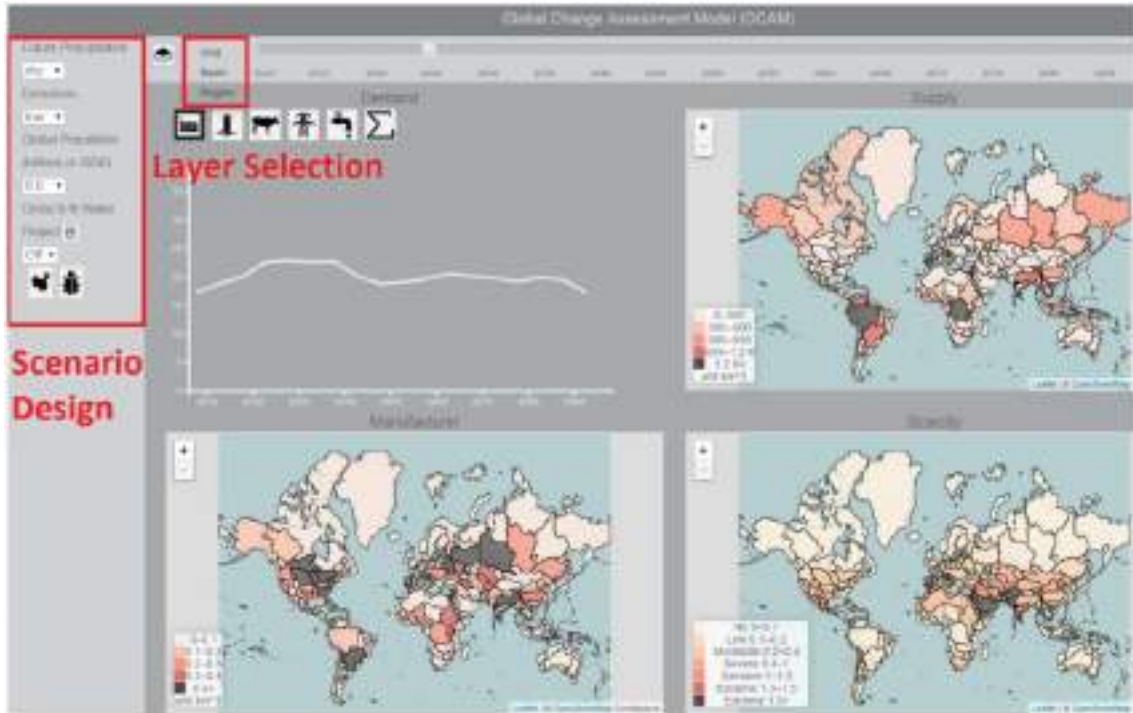
Şekil 2.24 SensePlace2 mekânsal analitik uygulaması kullanıcı ara yüzü (MacEachren ve ark., 2011)

Maciejewski ve ark. (2011), Şekil 2.25'te epidemiyoloji alanında hotspot (sıcak nokta) analizi hakkında akıllı analiz yaklaşımını sunmuşlardır (Maciejewski ve ark., 2011).



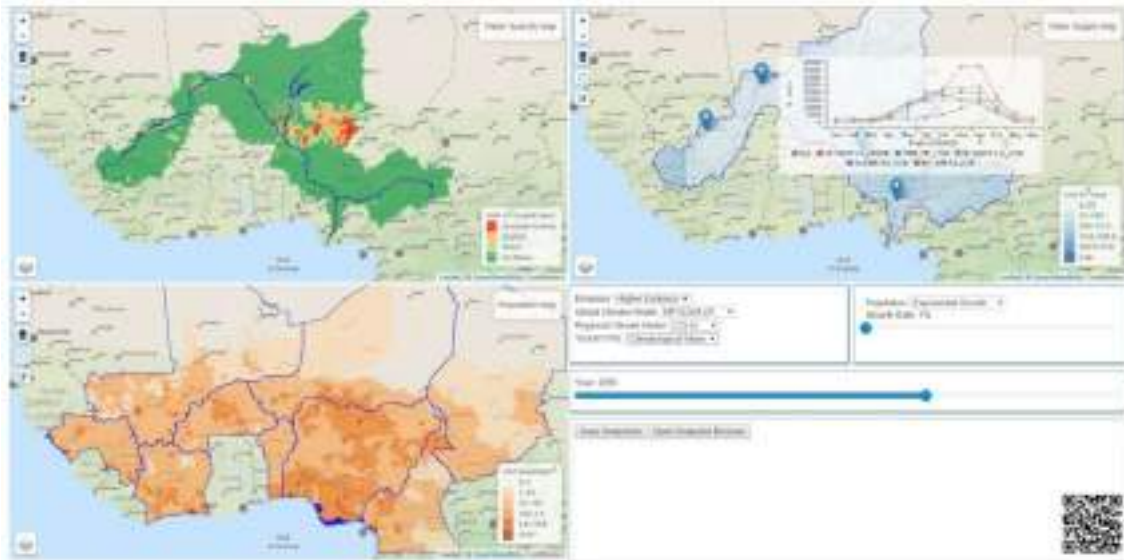
Şekil 2.25 Etkileşimli tahmin edici görsel analiz ortamı arayüzü (Maciejewski ve ark., 2011)

Luo ve ark. (2015), Şekil 2.26'da küresel değişim değerlendirme modelini kullanarak, iklim senaryolarını keşfetmeyi amaçlayan bir çalışma yürütmüşlerdir. Bu çalışma ile kullanıcıların su kıtlığını zamansal değişim, mekânsal farklılıklar ve gelecekteki çeşitli iklim politikası senaryolarıyla karşılaştırarak incelemelerine olanak tanıyan bir Web tabanlı mekânsal görsel analitik yaklaşımı sunmuşlardır (Luo ve ark., 2015).



Şekil 2.26 GCAM isimli Su ile ilgili iklim değişikliği senaryolarını araştıran Web tabanlı bir görselleştirme aracı (Luo ve ark., 2015)

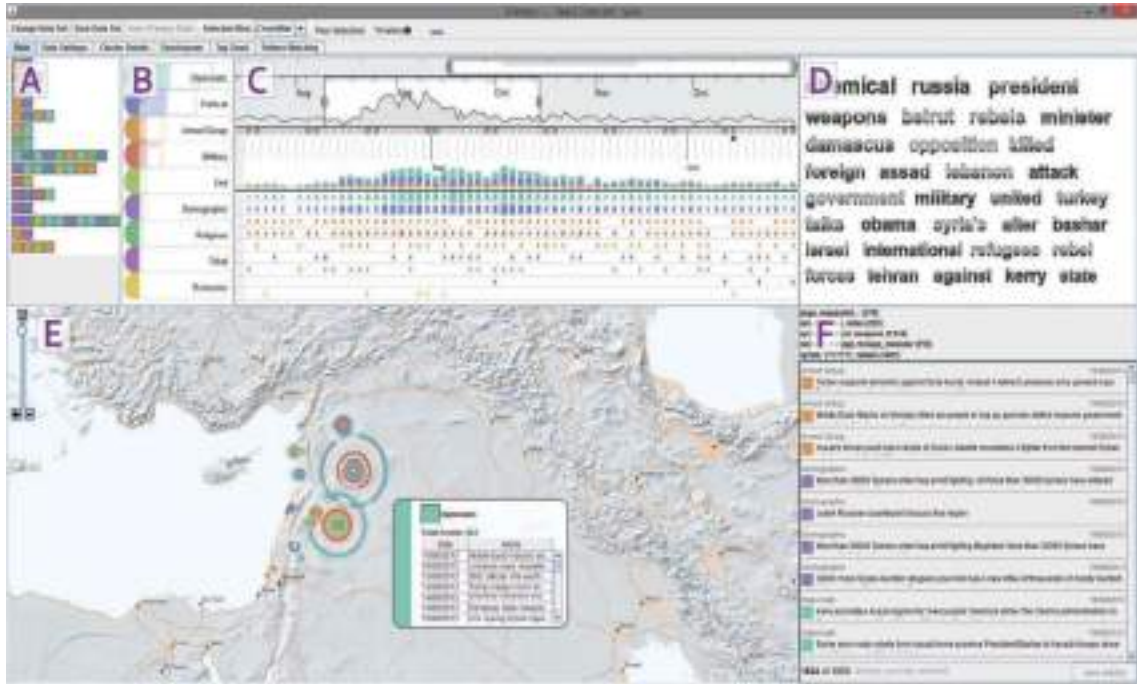
Lei ve ark. (2015), Şekil 2.27’de Nijer Nehri Havzası’ndaki su talebi, arzı ve kıtlığını farklı senaryolar altında incelemek amacıyla, iklim modelleri, iklim politikaları ve gelecekteki nüfus artışı kombinasyonları altında simülasyon sonuçlarını keşfetmek için bir mekânsal görsel analiz aracı geliştirmişlerdir (Lei ve ark., 2015).



Şekil 2.27 Nijer Nehri havzasındaki su kıtlığını keşfetmek için oluşturulan interaktif görsel analiz aracı (Lei ve ark., 2015)

Robinson ve ark. (2016), Şekil 2.28’te gerçekleştirdikleri çalışmada Suriye’deki siyasi, sosyal, ekonomik ve askeri olayların örüntülerini anlamak için mekânsal-zamansal verilerin değerlendirilmesinde STempo mekânsal görsel analitik sistem

tasarımı yaklaşımını sunmuşlardır. STempo araç seti, haber makalelerinden toplanan mekânsal-zamansal olaylarda örüntü keşfini ve analizini desteklemek için tasarlanmış bir mekânsal görsel analiz sistemidir (Robinson ve ark., 2016).

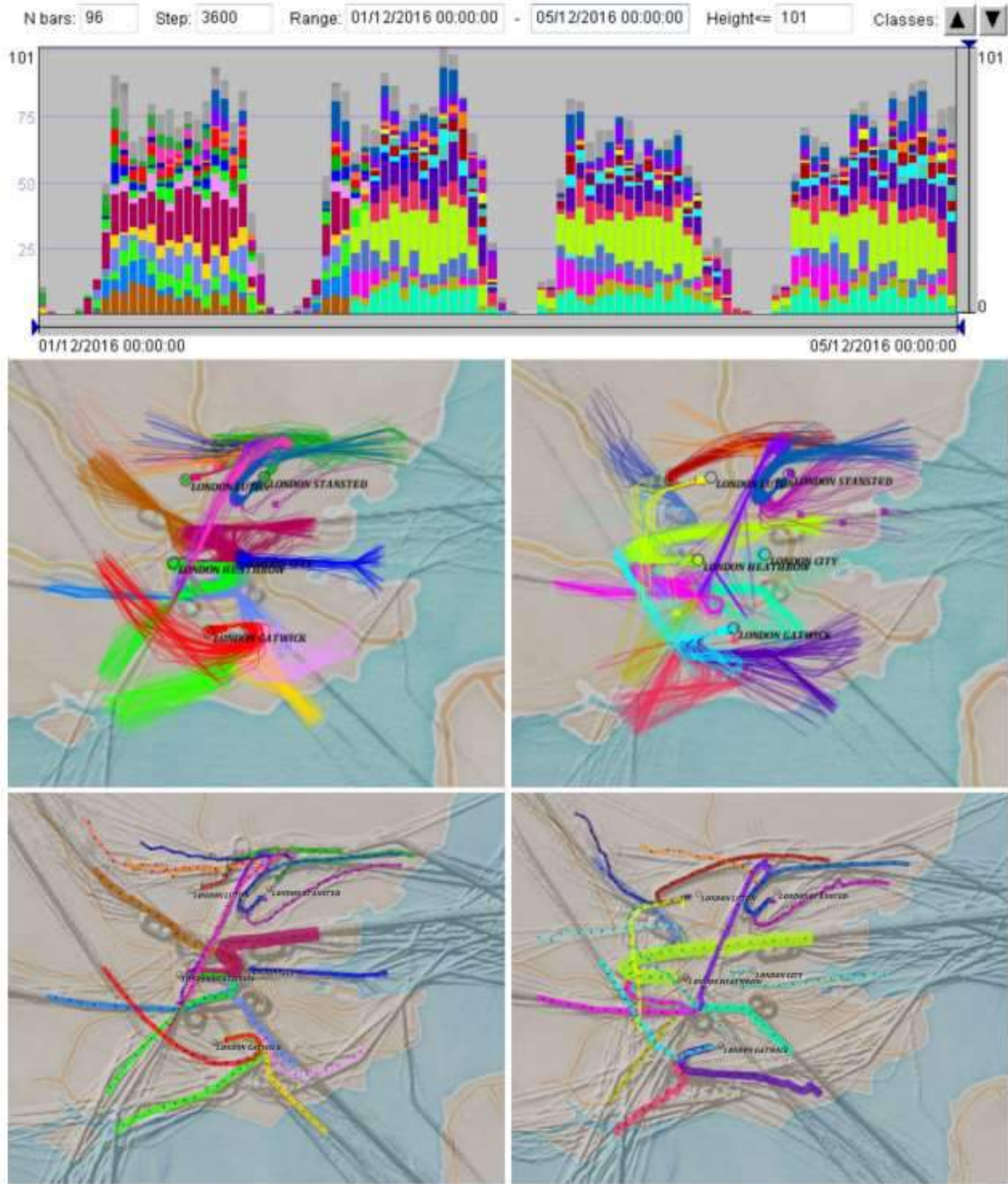


Şekil 2.28 STempo mekânsal görsel analitik uygulaması kullanıcı ara yüzü (Robinson ve ark., 2016)

Andrienko ve ark. (2017), Şekil 2.29’da hareket ve ulaşım sistemlerinin görsel analitik uygulanarak akıllı ulaşım sistemlerinin sağlanması ve geliştirilmesi yaklaşımında bulunmuşlardır (Andrienko ve ark., 2017). Ulaşım analizine yönelik bu mekânsal görsel analiz örneğinde, görüntünün üst kısmında yer alan zamansal histogramdaki çubuklar, Birleşik Krallık'taki havalimanlarına saatlik aralıklarla gelen uçuş sayısını göstermektedir. Histogram çubuk segmentleri, bu uçuşların ait olduğu rota tabanlı kümelerine göre renklendirilmiştir. Birinci gün, ikinci ve dördüncü günlere kıyasla daha farklı bir karışıma sahip gibi görünmektedir. Görüntünün ortasında, her bir uçuşun yörüngesinin son bölümleri rota tabanlı kümelerine göre renklendirilmiştir. Görüntünün alt kısmında, genel veri kümesindeki kalıpları genelleştirmek için her bir kümenin merkezi yörüngesi kullanılarak birinci ve üçüncü günler gösterilmektedir. (Andrienko ve ark., 2017).

Murthy ve Longwell (2012), 2010 yılı Pakistan sel felaketi sırasında sosyal iletişim platformu X'in kullanımını inceleyerek, kullanıcıların geleneksel ve sosyal medya bağlantılarını tweetleme/retweetleme eğiliminde olup olmadıklarını, bu kullanıcıların hangi ülkelerden tweet attıklarını, atılan tweetler ile Pakistan sel

felaketinin gerçekleştiği konum arasında bir korelasyon olup olmadığını araştırmışlardır (Murthy ve Longwell, 2012).



Şekil 2.29 Birleşik Krallık'taki hava ulaşım trafiğinin mekânsal ve zamansal görselleştirilmesi (Andrienko ve ark., 2017)

Nair ve ark. (2017), afet yönetimi alanında sosyal medyanın kullanımını incelemek için 2015 yılı Chennai taşkını hakkında atılan X verileri üzerinde, atılan tweetlerin taşkın ile alakalı olup olmadığını ve taşkın hakkındaki atılan tweetlerin en çok hangi kullanıcılar tarafından gerçekleştirildiğinin tespitine yönelik araştırma gerçekleştirmişlerdir (Nair ve ark., 2017).

Lamsal (2020), COVID-19 döneminde atılan X verilerini 40 farklı hashtag'e göre analizini gerçekleştirmiştir. Yaklaşık 310 milyondan fazla tweet verisinden yaklaşık 141 bin tweet verisinin coğrafi etiketli olduğunu tespit etmiştir.

Eligüzel (2021), Nepal Depremi hakkındaki X verilerini kullanarak metin sınıflandırması ve konum verilerine yönelik yeni bir yöntem geliştirmiştir. Olası bir deprem senaryosunda kapsamlı deprem yönetimi için X platformundan faydalı bilgiler elde etmeye odaklanmıştır (Eligüzel, 2021).

Preda (URL-4), Tüm dünyada yaygın olarak kullanılan Pfizer/BioNTech, Sinopharm, Sinovac, Moderna, Oxford/AstraZeneca, Covaxin ve Sputnik V aşılı hakkında atılan tweetleri toplayarak, Kaggle platformu üzerinden araştırmacıların kullanımına açmıştır (URL-4).

3. MATERYAL ve YÖNTEM

3.1. Kümeleme Analizi

Kümeleme, benzer verileri/nesneleri tanımlamayı ve bunları kümeler halinde gruplamayı amaçlamaktadır. Kümelerin hangi özelliklere göre oluşturulacağı başlangıçta sabitlenmez ve prosedüre bırakılmaktadır. Burada, bir küme içindeki öğeler mümkün olduğunca benzer, kümeler ise mümkün olabildiğince birbirinden farklıdır (Han ve Kamber, 2006). Kümeleme analizi yöntemleri, temelde hiyerarşik ve hiyerarşik olmayan kümeleme yöntemleri olmak üzere ayrılmaktadır. Hiyerarşik kümeleme algoritmaları, birleştirici (agglomerative) ve ayrıştırıcı (divisive) yöntemleri olarak ayrılırken; hiyerarşik olmayan kümeleme algoritmaları, bölümlenmeli (partitioning), yoğunluk temelli (density based), grid tabanlı ve model tabanlı yöntemler şeklinde ayrılmaktadır (Silahtaroglu, 2013). Bu tez çalışması kapsamında hiyerarşik birleştirici kümeleme algoritmalarından Ward yöntemi ve hiyerarşik olmayan bölümlenmeli algoritmalarından k-ortalama yöntemini kullanılacaktır.

3.1.1. Ward yöntemi

1963 yılında J.H. Ward tarafından geliştirilmiştir. Hiyerarşik kümeleme analiz yöntemlerinden olan Ward yöntemi birleştirici kümeleme yöntemleri arasında hata kareler toplamı kriterini esas alarak, grup içi dağılımı minimize ederek kümelerin oluşmasını sağlayan ve en iyi sonuç veren yöntem olarak kabul edilmektedir. (Hands ve Everitt, 1987; Ferreira ve Hitchcock, 2009; Murtagh ve Legendre, 2014). Ward tekniğinin amacı, kümelemenin her aşamasında küme içi hata kareler toplamındaki artışı en aza indirmektir. k kümesinde yer alan n_k noktanın k kümesinin ortalama vektörüne olan öklid uzaklıkları toplamı hata kareler toplamıdır ve W_k olarak ifade edilir;

$$W_k = \sum_{i=1}^p \sum_{j=1}^{n_k} (x_{ijk} - \bar{x}_{ik})^2 = \sum_{i=1}^p \sum_{j=1}^{n_k} x_{ijk}^2 - n_k \sum_{i=1}^p \bar{x}_{ik}^2 \quad (3.1)$$

şeklinde hesaplanmaktadır. Burada W_k değeri $k=1, 2, \dots, n$ kümelerde hesaplanarak küme içi hata kareler toplamı,

$$W = \sum_{i=1}^n W_k \quad (3.2)$$

şeklinde hesaplanmaktadır. W 'de en küçük artışa sahip olan p ve q kümeleri birleştirilerek t kümesi elde edilir. W 'deki bu artış,

$$DW_{pq} = W_t - W_p - W_q \quad (3.3)$$

eşitliği ile hesaplanır. Böylece n birim $(n-1)$ kümeye ayrılır. Bu şekilde küme sayısı $k=1$ oluncaya kadar W artış değerleri bulunarak birimlerin aşamalı biçimde birbirlerine bağlanmaları sağlanır.

Analiz sonucunda Ward Yöntemi kullanıldığında, birimler değişik seviyelerde başarılı bir şekilde bir araya geldikleri “dendrogram” adlı şemada gösterilir. Ward yöntemi aykırı noktalara duyarlı bir yöntemdir (Dibb, 1998).

3.1.2. K-ortalamlar (k-means) yöntemi

İlk olarak 1967 yılında Mac Queen tarafından önerilen ve hiyerarşik olmayan k-Ortalama algoritması, sürekli olarak kümelerin yinelendiği ve en uygun çözüme ulaşıncaya kadar devam eden döngüsel bir algoritmadır. Bu algoritmanın genel prensibi, n adet veri ögesinden oluşan bir veri kümesini, araştırmacının ön bilgisi ve deneyimine dayalı olarak belirlenen k adet kümeye bölmektir. Ana hedef, bölümlenme işlemi sonucunda elde edilen kümeler arasındaki benzerliği minimize etmek ve aynı zamanda kümeler içindeki benzerliği en üst düzeye çıkarmaktır. Algoritmanın temel kod yapısı aşağıdaki gibi sunulabilir (Silahtaroglu, 2013).

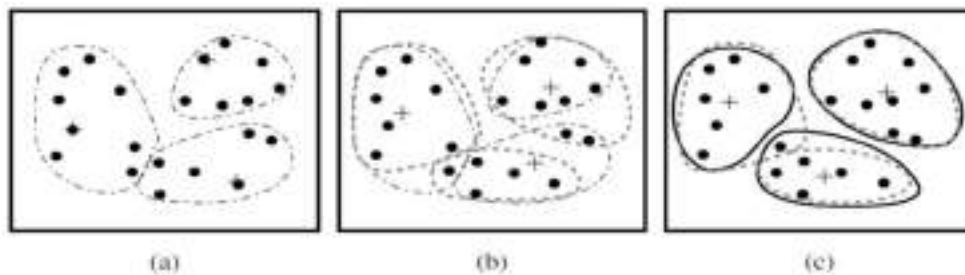
Girdiler: $D = \{t_1, t_2, \dots, t_n\}$ // eldeki veri tabanı

k // verilen küme sayısı

Adımlar:

1. Keyfi olarak m_1, m_2, \dots, m_k ortalama belirle (Eldeki veri tabanından rastgele)
2. Her bir t_i yi en yakın olduğu m_i nin kümesine ata
3. Kümelere ait m_1, m_2, \dots, m_k değerlerini yeniden hesapla
4. Küme elemanlarında herhangi bir değişiklik yoksa dur
5. İlk adımdan itibaren tekrar et.

Çıktı: k adet küme



Şekil 3.1 k-Ortalama algoritması ile kümeleme (Han ve Kamber, 2006)

3.1.3. Küme geçerliliği teknikleri

3.1.3.1. Kofenetik (cophonetic) korelasyon katsayısı

Hiyerarşik kümeleme yöntemlerinde, kümeleme sonuçlarının değerlendirilmesinde kofenetik korelasyon katsayısı kullanılmaktadır. Bir veri seti için en doğru kümelemenin gerçekleştiğini, kofenetik korelasyon katsayısının (*CCC*) yüksekliği göstermektedir (Saraçlı ve Akşit, 2022).

$$CCC = \frac{\sum_{i<j}(x(i,j)-x)(t(i,j)-t)}{\sqrt{\sum_{i<j}[(x(i,j)-x)^2] \sum_{i<j}[(t(i,j)-t)^2]}} \quad (3.4)$$

$x(i,j)=|X_i-X_j|$ = Öklid uzaklığı

$t(i,j)=|T_i-T_j|$ = Dendrogram uzaklığı

CCC katsayısı [-1,1] aralığında değer almaktadır. *CCC* değerinin 1 olması, veri kümesi için hesaplanan benzerlik matrisi ile hiyerarşik kümeleme sonucunda elde edilen ağaç diyagramında aynı kümede gruplanan veri çiftlerinin yakınlık seviyeleri değerlerinden oluşan kofenetik matrisin yüksek derecede benzer olduğunu göstermektedir. Kofenetik değer 0,80 veya daha büyük değer olduğunda kullanılan değişkenlerin birbiriyle ilişkili olduğu kabul edilmektedir (Romesburg, 1984).

3.1.3.2. Silhouette katsayısı

Hiyerarşik olmayan kümeleme yöntemlerinde Silhouette katsayısı, kümeleme tekniğinin kalitesini hesaplamak için kullanılan ölçüttür. Silhouette indeksi, her birimin atandığı kümeye uygunluğunun belirlenmesi amacıyla Rousseeuw tarafından önerilmiştir (Rousseeuw, 1987). Silhouette indeksi,

$$sil(i) = \frac{b(i)-a(i)}{\max(a(i),b(i))} \quad (3.5)$$

formülüyle hesaplanmaktadır. Burada, $a(i)$ 'nci birimin kendi kümesindeki tüm noktalara olan ortalama uzaklıklarını (benzerliğini) ve $b(i)$ i'nci birimin diğer kümelerdeki tüm noktalara olan ortalama uzaklıkların minimumunu ifade etmektedir. $sil(i)$ değeri 1'e yaklaşırsa i'nci birimin atandığı kümeye daha uyduğu, $sil(i)$ değeri 0'a yaklaşırsa veya negatif olursa i'nci birimin atandığı kümeye uygun olmadığı şeklinde yorumlanmaktadır. Tüm kümelemenin kalitesi,

$$sil(C) = \frac{1}{n} \sum_{s_i \in S} sil(i) \quad (3.6)$$

şeklinde tanımlanmaktadır. Bu formülere göre, maksimum ortalama Silhouette değerine ulaşan küme sayısı uygun küme sayısı olarak alınmaktadır. Silhouette katsayısı -1 ve +1 aralığında değer almaktadır. Silhouette katsayısının 1 değerine yaklaşması oluşan kümelerin birbirinden oldukça uzak olduğu ve açıkça ayırt edildiği, -1 değerine yaklaşması ise oluşan kümelerin yanlış şekilde oluşturulduğu anlamına gelmektedir (Rousseeuw, 1987).

3.1.4. KNIME

KNIME (The Konstanz Information Miner), 2004 yılında Konstanz Üniversitesi'nde geliştirilmeye başlayan açık kaynak kodlu bir veri analizi platformudur. Bu platform veri analizi süreçlerinin görsel olarak tasarlanmasını ve etkileşimli şekilde yürütülmesini sağlayan modüler bir ortamdır. Yeni algoritmaların, veri yönetim araçlarının veya görselleştirme yöntemlerinin yeni modüller veya düğümler (node) şeklinde kolayca entegre edilmesini sağlayan araştırma, öğretim ve iş birliği platformu olarak tasarlanmıştır. Knime, açık kaynaklı bir veri analizi platformu olarak öne çıkmakta ve kullanıcılarına birçok avantaj sunmaktadır. Büyük veri işleme süreçlerini görsel şekilde tasarlama yeteneği, bu platformu özellikle büyük veri analizi konusunda tercih edilen bir araç haline getirmektedir (URL-5).

3.1.5. RStudio

RStudio veri yönetimini, analiz işlemleri, istatistiki hesaplamalar ve veri görselleştirmeleri için kullanılan bütünleşik geliştirme ortamıdır (IDE). RStudio, R programlama dili odaklı çalışmaktadır. R dilinin kolay kullanımını sağlamak ve sonuçlar elde etmek için kullanıcı dostu etkileşimli bir arayüz sunmaktadır. RStudio, geniş bir R kullanıcı topluluğuna sahiptir ve kullanıcılar için forumlar, belgeler ve eğitim materyalleri sunmaktadır. RStudio'nun açık kaynak topluluğu tarafından geliştirilen ücretsiz RStudio Desktop ve ticari kullanıma yönelik RStudio Server Pro gibi farklı sürümleri bulunmaktadır (URL-6).

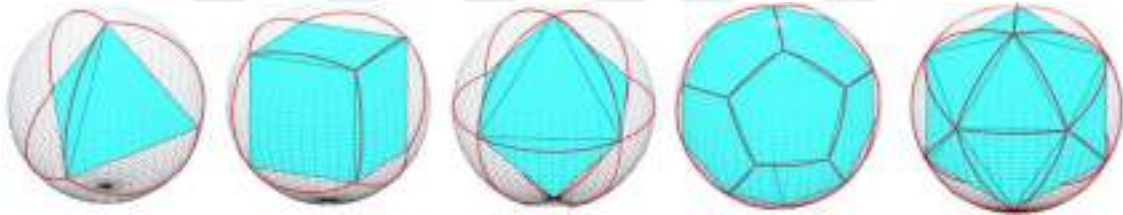
3.2. Ayrık Küresel Grid Sistemleri

Dijital Dünya'nın oluşturulmasında, veri küpleri ve Ayrık Küresel Grid Sistemleri (DGGS) ve olmak üzere iki ana yaklaşım üzerinde durulmaktadır (Purss ve ark., 2019; Alderson ve ark., 2020).

Veri küpleri, yapısı gereği mekânsal ya da mekânsal olmayan çeşitli öznitelik/koordinat eksenlerine göre sıralanan, sorgulamaya hazır mekânsal-zamansal verileri depolamak için kullanılan n boyutlu dizilerdir (Alderson ve ark., 2020).

Ayrık Küresel Grid Sistemleri'nde, Dünya'nın yüzeyi düzenli hiyerarşik hücrelerle kaplanmaktadır. Bu hücreler veri atamak ve almak amacıyla bir indeksleme mekanizması kullanılarak adreslenmektedir. DGGS teorisi, bilim topluluklarında yıllardır tartışılmaktadır. Ancak OGC Konsorsiyumu tarafından 2017 yılında hazırlanan özet şartnamesinde de gösterilmesi ve pratik uygulamaları ortaya çıkmasıyla birlikte özellikle son yıllarda araştırmacıların ilgisi daha da artmaktadır (Kmocho ve ark., 2022).

DGGS genel olarak, uyumlu hiyerarşik veri küpleri ve mekânsal veriler için küresel olarak uygulanabilen referans sistemi ve mekânsal veri modeli olarak kabul edilmektedir (Goodchild, 2018; Purss ve ark., 2019). Kürenin (veya elipsoidin) ayrıklaştırılması, gerçekte var olmayan bir temel katının yüzeylerinden türetilmektedir (Şekil 3.2). En yaygın olarak uygulanan temel katılar küp ve yirmiyüzlüdür (Kmocho ve ark., 2022).



Şekil 3.2 Gerçekte var olmayan temel katılar ve küresel gösterimleri: Soldan sağa tetrahedron, küp, oktahedron, dodekahedron, icosahedron (Lei ve ark., 2020)

DGGS uygulamalarının hücre bölünmeleri; temel katı seçimi, hiyerarşik bölümlenme şeması ve hücreler için coğrafi referanslama veya projeksiyon seçimi nedenleriyle farklı geometrik özelliklere sahiptir (Kmocho ve ark., 2022).

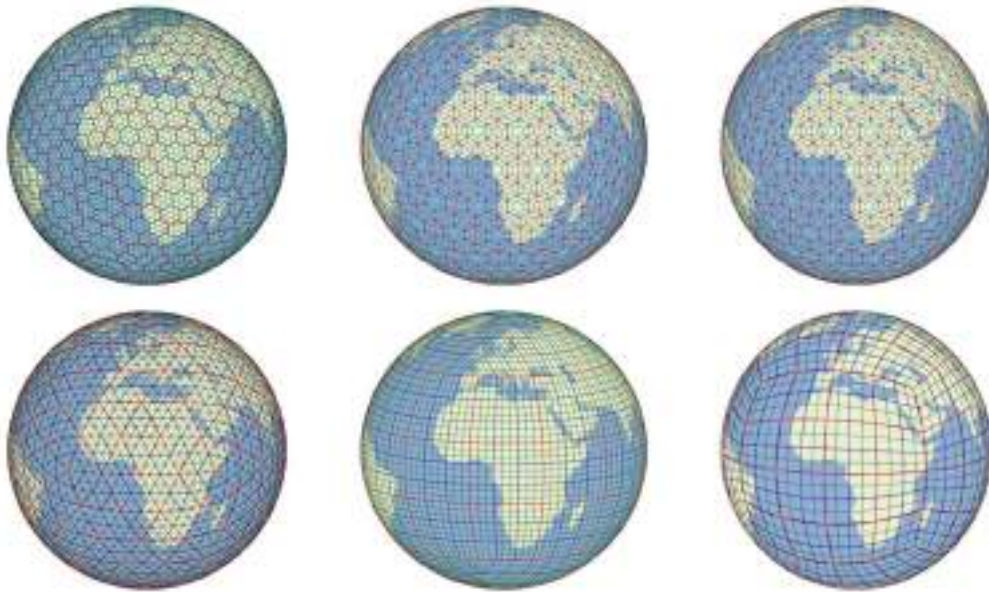
Bir DGGS uygulaması,

- Bölme yüzeyi olarak temel katı seçimi,
- Temel katının Dünya'ya göre yönünün belirlenmesi,
- Temel katının yüzeylerinde hiyerarşik mekânsal bölümlenme yönteminin tanımlanması,
- Düzlemsel bölme yüzeyinin küresel ya da elipsoidal yüzeye izdüşürülmesi,
- Grid hücrelerinin indekslenmesi ve adreslenmesi için bir yöntemin belirlenmesi olmak üzere bir dizi temel adımı gerektirmektedir (Kmocho ve ark., 2022).

Literatürde ve mevcut yazılımlarda kullanılan mozaikleme yöntemlerinin çoğu,

- Dikdörtgen uyumlu iyileştirmelerle temel katı olarak küp,
- Dörtlü üçgen ağlar,
- Altıgen sistemlerden oluşmaktadır (Kmocho ve ark., 2022).

Bir mozaikleme sisteminin açıklığı, o sistemdeki bir ana hücrenin bir sonraki daha üst çözünürlükteki alt hücreleriyle olan ilişkisini tanımlamaktadır. Örneğin; bir kare (ya da herhangi bir dörtgen) veya bir üçgen 4 eşit hücreye bölebildiği için açıklıkları 4 olmaktadır. Altıgenler için 3, 4 ve 7 olmak üzere farklı açıklıkları bulunmaktadır. Ancak sadece 7 açıklıklı altıgen sistemlerinde neredeyse tam uyumlu bir ebeveyn-çocuk ilişkisinden bahsedilmektedir. Aşağıdaki Şekil 3.3'te üst sırada soldan sağa 7 Açıklıklı Altıgen (7H), 3 Açıklıklı Altıgen (3H), 4 Açıklıklı Altıgen (4H); alt sırada soldan sağa 4 Açıklıklı Üçgen (4T), 9 Açıklıklı rHEALPix, 4 Açıklıklı Kareler olmak üzere en yaygın olarak kullanılan yüzey mozaiklerini gösterilmektedir (Kmocho ve ark., 2022).



Şekil 3.3 En yaygın uygulanan mozaikleme şemaları (Kmocho ve ark., 2022)

Suç analizi (Jendryke ve McClure, 2019), orman yangını modellemesi (Robertson ve ark., 2020), kıyı ortamlarının karakterizasyonu (Bousquin, 2021), deniz trafiği için risk analizi (Rawson ve ark., 2021) veya taşkın haritalaması (Chaudhuri ve ark., 2021) gibi birçok alanda Ayırık Küresel Grid Sistemleri (DGGS) tabanlı Coğrafi Bilgi sistemleri (CBS) analizleri gerçekleştirilmektedir.

Araştırma ve uygulama alanında, açık kaynak kodlu ve ücretsiz olarak kullanıma sunulan Uber H3 (URL-7), rHEALPix (Gibb ve ark.,2016), DGGRID (Sahr, 2023) ve S2 (Veach ve ark., 2017) olmak üzere farklı DGGS yazılımları bulunmaktadır.

3.2.1. Uber H3

Grid sistemleri, büyük mekânsal veri kümelerini analiz etmek için yerküreyi tanımlanabilir grid hücrelerine ayırmada büyük öneme sahiptir (URL-8).

UBER ekibi, ulaşım alanında sevkiyat ve yolculuk ücretlerini etkili bir şekilde optimize etmek, mekânsal verileri keşfetmek ve görselleştirmek için H3'ü geliştirmiştir. H3, dinamik fiyatlar belirlemek ve şehir çapında başka kararlar almak için coğrafi bilgileri analiz edilmesini sağlamaktadır. Market alanlarının analizi ve optimizasyonu için grid sistemi olarak H3 kullanılmaktadır. H3 sistemi, açık kaynaklı olarak Github platformunda diğer kullanıcıların kullanımına sunulmaktadır (URL-8).

Uber ekibi, altıgen bir grid sistemin faydalarını hiyerarşik bir indeksleme sistemi ile birleştirmek için H3'ü oluşturmuşlardır. Küresel bir grid sistemi genellikle bir harita projeksiyonu ve harita üzerine yerleştirilmiş grid olmak üzere en az iki unsuru gerektirmektedir. Dünya üzerindeki üç boyutlu bir konumdan harita üzerinde iki boyutlu bir noktaya gitmek için bir harita projeksiyonu gerekmektedir. Daha sonra harita üzerine grid yerleştirilerek küresel bir grid sistemi oluşturulmaktadır (URL-8).

Bu süreç, farklı harita projeksiyonlarını ve gridleri birleştirerek gerçekleştirilmektedir. Mercator projeksiyonu ve kare grid en yaygın örneğidir. Bu basit yöntem işe yarayacak olsa da birtakım dezavantajları bulunmaktadır. En başta Mercator projeksiyonunun da önemli bir boyut bozulması gerçekleşmektedir. Bu neden bazı hücreler çok farklı alanlara sahip olacaktır. Kare gridlerin ayrıca, analiz için kullanıldığında birden fazla katsayı seti gerektiren dezavantajları bulunmaktadır. Bu dezavantaj, karelerin dört ana yönde ortak kenar paylaşımları ve dört çapraz yönde bir tepe noktasını paylaşımları bakımından iki farklı türdeki komşuluk ilişkisinin sonucu olarak ortaya çıkmaktadır (URL-8).

Uber tarafından geliştirilen H3 mekânsal indeksleme sistemi, yeryüzü küresini yüksek hassasiyetli altıgen bölümlere ayıran ayırık küresel grid sistemidir (Sahr, 2023; URL-10; URL-18).

Uber ekibi, harita projeksiyonu için, yirmiyüzlü yüzeyle odaklanan gnomonik projeksiyonları kullanmaktadır. Yirmiyüzlü tabanlı bir harita projeksiyonu, tek bir

düzlem yerine yirmi ayrı iki boyutlu düzlemle sonuçlanmaktadır. Yirmiyüzlü, her seferinde iki boyutlu bir harita oluşturarak birçok şekilde açılabilir. Bununla birlikte H3, grid sistemini oluşturmak için yirmiyüzlüyü açmaz ve bunun yerine gridleri yirmiyüzlü yüzeylerin üzerine yerleştirerek jeodezik ayrık bir küresel grid sistemi (DGGS) oluşturmaktadır. Uber ekibi, Şekil 3.4'te görüldüğü üzere H3'ün harita projeksiyonu için yirmiyüzlü yüzeylere (solda) odaklanan gnomonik projeksiyonları kullanmakta ve Dünya'yı küresel bir yirmiyüzlü olarak yansıtmaktadır (sağda) (URL-8).



Şekil 3.4 H3'ün harita projeksiyonu için kullandığı yirmiyüzlü yüzey ve yirmiyüzlü yüzeye odaklanan gnomonik projeksiyon (URL-8)

Hücre şekli olarak bir altıgen kullanmak H3 için çok önemlidir. Aşağıdaki Şekil 3.5'te gösterildiği gibi, komşuluk ilişkileri bakımından kare hücreler iki farklı mesafeye, üçgen hücreler üç farklı mesafeye ve altıgen hücreler altıgen merkez noktası ile komşu altıgenler arasında yalnızca tek bir mesafeye sahiptir. Bu özellik analiz yapmayı ve gradyanları yumuşatmayı büyük ölçüde basitleştirmektedir (URL-8; URL-9).



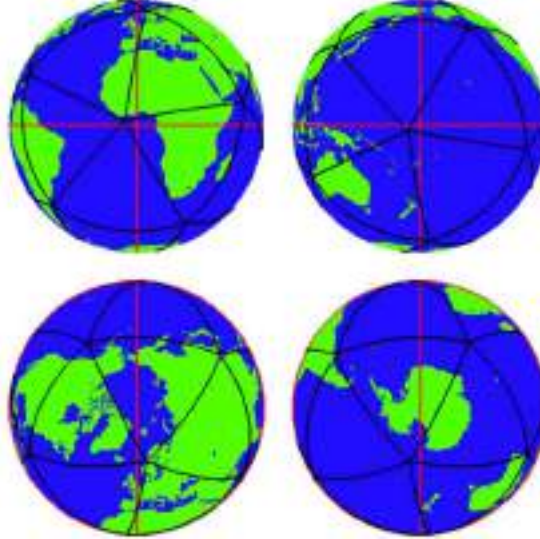
Şekil 3.5 Bir üçgenin komşularına, bir karenin komşularına ve bir altıgenin komşularına uzaklıkları (URL-8)

H3 gridleri, Şekil 3.6'da görüldüğü gibi yüzey başına on hücre olacak şekilde Dünya üzerine 122 temel hücre yerleştirilerek inşa edilmektedir. Bazı hücreler birden fazla yüze sahiptir. Yirmiyüzlüyü sadece altıgenlerle döşemek mümkün olmadığından, her yirmiyüzlünün köşesine birer tane olmak üzere on iki beşgen hücre eklenmektedir.



Şekil 3.6 Tek bir yirmiyüzlü yüzünde grid oluşturmak için H3 kullanılmaktadır (URL-8)

Yirmiyüzlü, R. Buckminster Fuller'in Dymaxion oryantasyonu kullanılarak küreye sabitlenmektedir (İpbüker, 2005). R. Buckminster Fuller'ın Dymaxion oryantasyonu, bir yirmiyüzlünün tepe noktasını 5.24539° batı boylamında, 2.300882° Kuzey enleminde ve bitişik bir tepe noktasını ilk tepe noktasından 7.46658° azimut açısıyla yerleştirmektedir. Bu yerleşimde, Şekil 3.7'de görüldüğü gibi yirmiyüzlülerin tüm köşeleri okyanus yüzeylerinde yer almaktadır. Bu şekilde yüzeye çıkan beşgenler önlenmektedir (URL-8; URL-17).



Şekil 3.7 R. Buckminster Fuller'in yirmiyüzlü yerleşimine ait dört görünüm (URL-17)

Hiyerarşik indeksleme gridi, her altıgenin alt birim altıgenlere bölünebileceği anlamına gelmektedir. Bu nedenle H3, ortalama $4.357.449,41 \text{ km}^2$ 'lik bir alandan (çözünürlük 0) $0,895 \text{ m}^2$ 'ye (çözünürlük 15) kadar dünyadaki herhangi bir konumu on altı farklı seviyede indeksleyebilmektedir. H3 mekânsal indeksleme sisteminde, sıfır çözünürlük seviyesinden onbeşinci çözünürlük seviyesinde altıgenlerin kapladıkları

ortalama alanları, altıgenlerin ortalama kenar uzunluklarını ve her seviyedeki benzersiz hücreleri Çizelge 3.1’de gösterilmektedir (URL-20).

Çizelge 3.1 Farklı H3 çözünürlük seviyeleri için hücre bilgileri tablosu (URL-20)

H3 Çözünürlük Seviyesi	Ortalama Altıgen Alanı (km ²)	Ortalama Altıgen Kenar Uzunluğu (km)	Benzersiz Hücre Sayısı
0	4,357,449.416078381	1281.256011	122
1	609,788.441794133	483.0568391	842
2	86,801.780398997	182.5129565	5,882
3	12,393.434655088	68.97922179	41,162
4	1,770.347654491	26.07175968	288,122
5	252.903858182	9.854090990	2,016,842
6	36.129062164	3.724532667	14,117,882
7	5.161293360	1.406475763	98,825,162
8	0.737327598	0.531414010	691,776,122
9	0.105332513	0.200786148	4,842,432,842
10	0.015047502	0.075863783	33,897,029,882
11	0.002149643	0.028663897	237,279,209,162
12	0.000307092	0.010830188	1,660,954,464,122
13	0.000043870	0.004092010	11,626,681,248,842
14	0.000006267	0.001546100	81,386,768,741,882
15	0.000000895	0.000584169	569,707,381,193,162

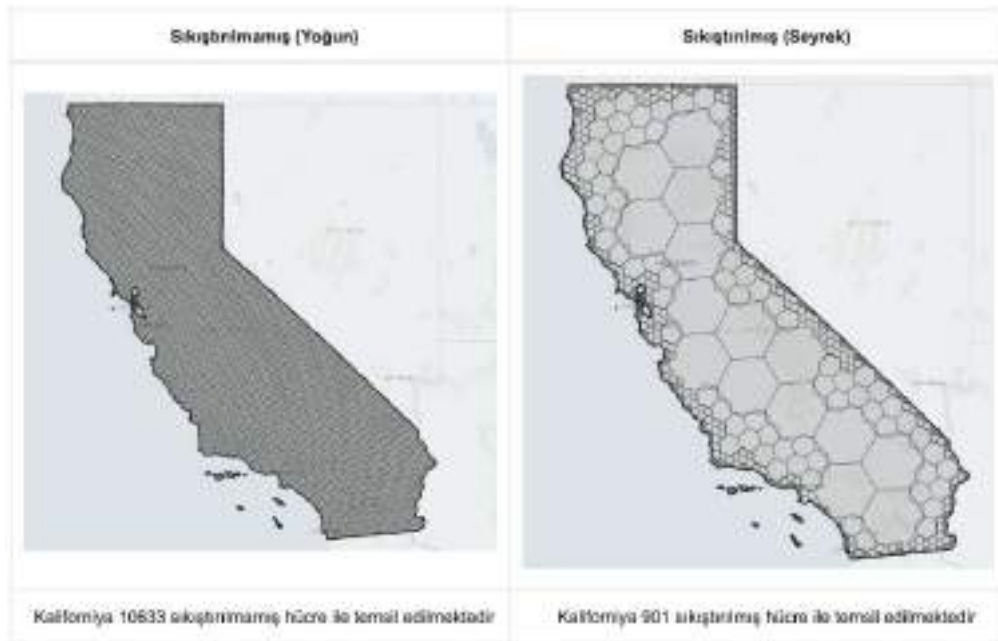
İlk H3 çözünürlüğü (çözünürlük 0), 110 altıgen ve 12 yirmiyüzlü tepe merkezli beşgen olmak üzere toplam 122 temel hücrelerden oluşmaktadır. Bu temel hücrelere, merkez noktalarının enlem değerlerine göre 0 – 121 aralığında değerler atanmaktadır. 0 numaralı temel hücre en kuzeydeki merkez noktaya sahip hücreyi belirtirken, 121 numaralı temel hücre en güneydeki merkez noktaya sahip hücreyi göstermektedir. Her üst seviyedeki çözünürlük, bir alt seviyedeki çözünürlükteki alanın yedide bir oranındaki alandan oluşan hücrelerden oluşmaktadır. Altıgen hücreler mükemmel bir şekilde yedi altıgene bölünemediklerinden, her alt seviyedeki hücreler yalnızca Şekil 3.8’de görüldüğü gibi yaklaşık olarak bir ana hücre içinde bulunmaktadır (URL-8; URL-18).



Şekil 3.8 H3, kullanıcının alanları daha küçük parçalara ve daha küçük altıgenlere ayırmasını sağlamaktadır (URL-8)

Alt hücrelerin tanımlayıcıları, ata hücrelerini daha alt bir çözünürlükte bulmak için kolayca kesilebilmektedir. Bu işlem verimli bir indeksleme sağlamaktadır. Alt hücreler sadece yaklaşık olarak barındırıldığı için, kesme işlemi sabit oranda şekil bozulması gerçekleştirmektedir. Bu bozulma sadece bir hücre tanımlayıcısının kesilmesi sırasında ortaya çıkmaktadır. Konumları belirli bir çözünürlükte indekslerken hücre sınırlarında bozulma ortaya çıkmamaktadır (URL-8). İndeksleme işlemindeki coğrafi sınırlama yaklaşık olsa da mantıksal sınırlama kesindir. Bu sayede, H3 gridleri ile farklı çözünürlük seviyelerindeki indekslenmiş veri setlerinin verimli bir şekilde ilişkilendirilmesine izin verilmektedir. H3 indeksinin yapısı, coğrafi olarak yakın konumların sayısal olarak da yakın indekslere sahip olacağı anlamına gelmektedir (URL-11).

Bir konum için kesinlik veya belirsizliğin mekânsal indekste kodlanması gerektiğinde hiyerarşik yapı analizlerde de kullanılabilir. Örneğin; bir GPS alıcısında alınan bir nokta, sinyal hassasiyetinin daha düşük olduğu durumlarda daha alt bir çözünürlükte indekslenebilir ya da daha üst çözünürlükteki hücrelerin her birinde çok az veri noktasının bulunduğu durumlarda bazı hücreler bir ana hücrede toplanabilmektedir (URL-11). Hiyerarşik kümeleme, bitişik hücre kümelerini yoğunlaştırmaya ve seyrekletirmeye izin vermektedir. Şekil 3.9'da Kaliforniya şehrinin yoğunlaştırılmış ve seyrekleştirilmiş hücre kümeleriyle temsili görülmektedir (URL-11).



Şekil 3.9 Kaliforniya şehrinin yoğunlaştırılmış ve seyrekleştirilmiş hücre kümeleriyle temsili edilmesi (URL-11)

3.2.1.1. H3 indekisleme yapısı

H3 sistemi, hücrelerin merkezine, her hücrenin tüm tepe noktalarına ve her hücrenin tüm kenarlarına sırasıyla kök hücre sine ve sahip hücre sine göre hiyerarşik bir dizin atamaktadır (URL-12).

Hücre Modu: Herhangi bir çözünürlük seviyesindeki r hücre sinin indeksi, r hücre sinin çözünürlük 0 seviyesindeki temel hücre numaralı ile başlamaktadır. Bu işlemi, 0-6 arası r basamaklı bir dizi izlemektedir. Buradaki her i 'nci d_i basamağı, d_i ile d_{i-1} arasında basamaklandırılan 7 hücre den birini belirtmektedir. 0 çözünürlüğündeki temel hücrelerin her birine bir yerel altıgen koordinat sistemi atanır ve bu koordinat sistemi o temel hücre nin tüm hiyerarşik indekisleme alt hücrelerini yönlendirmek için kullanılmaktadır. Her çözünürlük seviyesinde 0-6 rakamlarının atanması bir Merkezi Konum İndeksleme düzenlemesini kullanmaktadır (Sahr, 2023). Çocuk altıgenler doğrusal olarak ana altıgenlerden daha küçüktür (URL-13). Şekil 3.10'da 5'inci çözünürlük seviyesinde H3 indeksine sahip H3 altıgen hücre si ve hücre bilgileri görülmektedir.



Şekil 3.10 "85283473fffff" H3 indeksine sahip H3 altıgen hücre si ve bilgileri (URL-16)

Bir H3 Hücre indeksi (mod 1), H3 grid sistemindeki belirli bir çözünürlükte bir hücre yi (altıgen veya beşgen) temsil etmektedir. H3 Hücre indeksinin bileşenleri aşağıdaki gibi en yüksek bit önce olacak şekilde 64 bitlik bir tamsayı halinde paketlenmektedir;

- 1 bit ayrılmıştır ve 0'a ayarlanmıştır,
- 4 bit H3 Hücre indeks modunu belirtmek için,
- 3 bit ayrılmıştır ve 0'a ayarlanmıştır,
- 4 bit 0-15 hücre çözünürlüğünü belirtmek için,
- 7 bit 0-121 temel hücre sini belirtmek için,

- 3 bit Çözünürlük 1'den hücrenin çözünürlüğüne kadar sonraki her 0-6 hanesini belirtmek için (toplam 45 bit) çözünürlük 1-15 için ayrılmıştır.

Kullanılmayan her basamak için üç bit 7'ye ayarlanmaktadır.

Hücreler için bir H3 İndeksinin bit düzeni Şekil 3.11'de gösterilmektedir. "Mod-Bağımlı" alanının yorumu, indeks moduna bağlı olarak farklılık göstermektedir (URL-13).

	0x0F	0x0E	0x0D	0x0C	0x0B	0x0A	0x09	0x08	0x07	0x06	0x05	0x04	0x03	0x02	0x01	0x00
0x50	Reserved (R)	Mode (M)			Reserved (R)			Resolution			Base cell					
0x20	Base cell		Digit 1		Digit 2		Digit 3		Digit 4		Digit 5		Digit 6		Digit 7	
0x10	Digit 8		Digit 9		Digit 10		Digit 11		Digit 12		Digit 13		Digit 14		Digit 15	
0x00	Digit 16		Digit 17		Digit 18		Digit 19		Digit 20		Digit 21		Digit 22		Digit 23	

Şekil 3.11 Hücreler için H indeksinin bit düzeni (URL-13)

Yönlendirilmiş Kenar Modu: H3 Yönlendirilmiş Kenar indeksi (mod 2), ("kaynak" hücre ve komşu "hedef") hücre olmak üzere iki hücre arasındaki tek yönlü bir kenarı temsil etmektedir (URL-14). Şekil 3.12'de farklı H3 yönlendirilmiş kenar indekslerine sahip H3 altıgen hücresi görülmektedir.



Şekil 3.12 6 Farklı H3 yönlendirilmiş kenar indekslerine sahip H3 altıgen hücresi (URL-14)

H3 Yönlendirilmiş Kenar indeksinin bileşenleri, aşağıdaki gibi en yüksek bit önce olacak şekilde 64 bitlik bir tamsayı halinde paketlenmektedir;

- 1 bit ayrılmıştır ve 0'a ayarlanmıştır,
- 4 bit H3 Tek Yönlü Kenar indeks modunu belirtmek için,
- 3 bit Kaynak hücrenin kenarını (1-6) belirtmek için,

- Kök hücrenin dizin bitleriyle eşleşen sonraki bitler.

Yönlendirilmiş kenarlar için bir H3 İndeksinin bit düzeni Şekil 3.13'te gösterilmektedir. "Mod-Bağımlı" alanının yorumu, indeks moduna bağlı olarak farklılık göstermektedir (URL-14).

0x0F	0x0E	0x0D	0x0C	0x0B	0x0A	0x09	0x08	0x07	0x06	0x05	0x04	0x03	0x02	0x01	0x00
0x50	Resonans 10	Mode 01		Edge				Resolution					Base cell		
0x20	Base cell		Digit 1		Digit 2		Digit 3		Digit 4		Digit 5				
0x10	Digit 1		Digit 6		Digit 7		Digit 8		Digit 9		Digit 10				
0x00	Digit 10	Digit 11		Digit 12		Digit 13		Digit 14		Digit 15					

Şekil 3.13 Yönlendirilmiş kenarlar için H3İndeksinin bit düzeni (URL-14)

Köşe Nokta Modu: H3 Köşe nokta indeksi (mod 4), H3 grid sisteminde üç hücre tarafından paylaşılan tek bir topolojik köşe noktasını temsil etmektedir. Bir H3 köşe noktası, rastgele üç komşu hücreden birine "owner" olarak atanmakta ve köşe noktasının coğrafi koordinatları ile standart indeksi hesaplamakta kullanılmaktadır (URL-15). Şekil 3.14'te farklı H3 köşe noktası indeksine sahip H3 altıgen hücresi görülmektedir.



Şekil 3.14 6 Farklı H3 köşe noktası indeksine sahip H3 altıgen hücresi (URL-15)

H3 Köşe noktası indeksinin bileşenleri, aşağıdaki gibi en yüksek bit önce olacak şekilde 64 bitlik bir tamsayı halinde paketlenmektedir.

- 1 bit ayrılmıştır ve 0'a ayarlanmıştır,
- 4 bit H3 Köşe noktası indeks modunu belirtmek için,
- 3 bit sahip hücrenin köşe noktasının köşe numarasını (0-5) belirtmek için,


```
h3.latLng_to_cell(lat, lng, resolution)
```

Grid uzayında hücrenin merkezini bulmak için,

```
h3.cell_to_latlng(cell)
```

Hücrenin sınırlarını bulmak için,

```
h3.cell_to_boundary(cell, geo_json=False)
```

ifadeleri kullanılmaktadır (URL-19).

Fonksiyonların Java Dilindeki Kullanımı

Konum bilgisi içeren hücrenin indeksinin döndürerek, konumu belirtilen çözünürlükte indekslemek için,

```
long latLngToCell(double lat, double lng, int res);
String latLngToCellAddress(double lat, double lng, int
                           res);
```

Grid uzayında hücrenin merkezini bulmak için,

```
LatLng cellToLatLng(long cell);
LatLng cellToLatLng(String cellAddress);
```

Hücrenin sınırlarını bulmak için,

```
List<LatLng> cellToBoundary(long cell);
List<LatLng> cellToBoundary(String cellAddress);
```

ifadeleri kullanılmaktadır (URL-19).

Fonksiyonların JavaScript (Live) Dilindeki Kullanımı

Konum bilgisi içeren hücrenin indeksinin döndürerek, konumu belirtilen çözünürlükte indekslemek için,

```
h3.latLngToCell(lat, lng, res)
```

Grid uzayında hücrenin merkezini bulmak için,

```
h3.cellToLatLng(cell)
```

Hücrenin sınırlarını bulmak ve GeoJson formatında almak için,

```
h3.cellToBoundary(cell, [formatAsGeoJson])
```

ifadeleri kullanılmaktadır (URL-19).

3.2.1.3. Uber H3'ün mekânsal analizlerde kullanılması

Uber pazarında her gün milyonlarca olay meydana gelmektedir. Platformdaki diğer işlemlerin yanı sıra her dakika, sürücüler sürüş talep etmekte, sürücü ortakları yolculuk başlatmakta ve aç kullanıcılar yemek istemektedir. Her olay belirli bir konumda gerçekleşmektedir. Örneğin, bir müşteri evinden bir araçla yolculuk talebinde bulunmakta ve bir sürücü bu isteği sadece birkaç kilometre ötedeki arabasında kabul etmektedir (URL-8).

Bu etkinlikler, Uber'in hizmetlerindeki kullanıcıları açısından pazarı daha iyi anlamasına ve optimize etmesine olanak tanımaktadır. Örneğin, bu olaylar şehrin belirli bir bölümünde arzdan daha fazla talep olduğunu söyleyebilmekte ve buna göre fiyatlandırmayı ayarlama veya platforma uberPOOL'da belirli bir sürücüye yakın mesafede iki yolculuk talebi olduğunu bildirebilmektedir (URL-8).

Uber pazarındaki verilerden bilgi ve öngörü elde etmek, verilerin tüm şehir genelinde analiz edilmesini gerektirmektedir. Şehirler coğrafi olarak çeşitli olduğundan, bu analizin ince bir ayrıntı düzeyinde yapılması gerekir. Bir olayın meydana geldiği tam konum olan en ince ayrıntı düzeyinde analiz çok zor ve pahalıdır. Bir şehir içindeki mahalleler gibi alanlarda analiz yapmak çok daha pratiktir (URL-8).

Şekil 3.16'da görülen birinci harita şehirdeki arabaları, ikinci harita H3 altıgen hücreleri içerisinde kalan arabaları, üçüncü harita H3 altıgen hücreleri içerisinde kalan arabaların toplam sayılarını renkler ve tonlarıyla ifade edilmektedir (URL-8).



Şekil 3.16 Bir şehirdeki arabaların harita üzerinde gösterimi, harita üzerinde H3 altıgen hücreleriyle birlikte gösterimi ve H3 altıgen hücreleri içerisinde kalan arabaların toplam sayılarını renkler ve renk tonlarıyla gösterimi (URL-8)

Uber bu tür olguları altıgen alanlara (hücelere) ayırmak için grid sistemi kullanmaktadır. Veri noktaları altıgen hücreler halinde toplanarak kullanılmaktadır. Örneğin; Uber hizmet verdiği her şehirde arz ve talebi altıgen hücreler kapsamında

ölçerek dalgalanma fiyatlandırmasını hesaplamaktadır. Bu altıgen hücreler, Uber pazarına ilişkin analizlerin temelini oluşturmaktadır (URL-8).

Şehirdeki insanlar genellikle hareket halindedir. Kullanıcıların bir şehirde hareket ederken ortaya çıkan nicemleme (kuantalama) hatasını en aza indirme açısından Altıgen hücreler önemli bir seçim olmaktadır. Altıgen hücreler, ayrıca Elasticsearch sorgu motorunu kullanarak yarıçapları kolayca tahmin etmeye imkân sağlamaktadır (URL-8).

Olayları alanlara bölmek için, alanları etrafındaki çokgen bölgelere ayırmak gibi kullanılacak farklı yöntemlerde bulunmaktadır. Bu alanlar Şekil 3.17’de görüldüğü gibi bir şehrin posta kodu alanları gibi alanlar olabilmektedir. Ancak bu tür alanlar, analiz için kullanışlı olmayan sıra dışı şekillerle ve boyutlara sahiptirler. Bölgeler, Uber ekipleri tarafından o şehir hakkındaki bilgilerine göre çizilebilir. Ancak bu tür bölgelerin şehirler değiştikçe sık sık güncellenmesi gerekecek ve alan sınırları genellikle keyfi olarak tanımlanacaktır (URL-8).



Şekil 3.17 New York şehrinin (Manhattan) posta kodu haritası (URL-8)

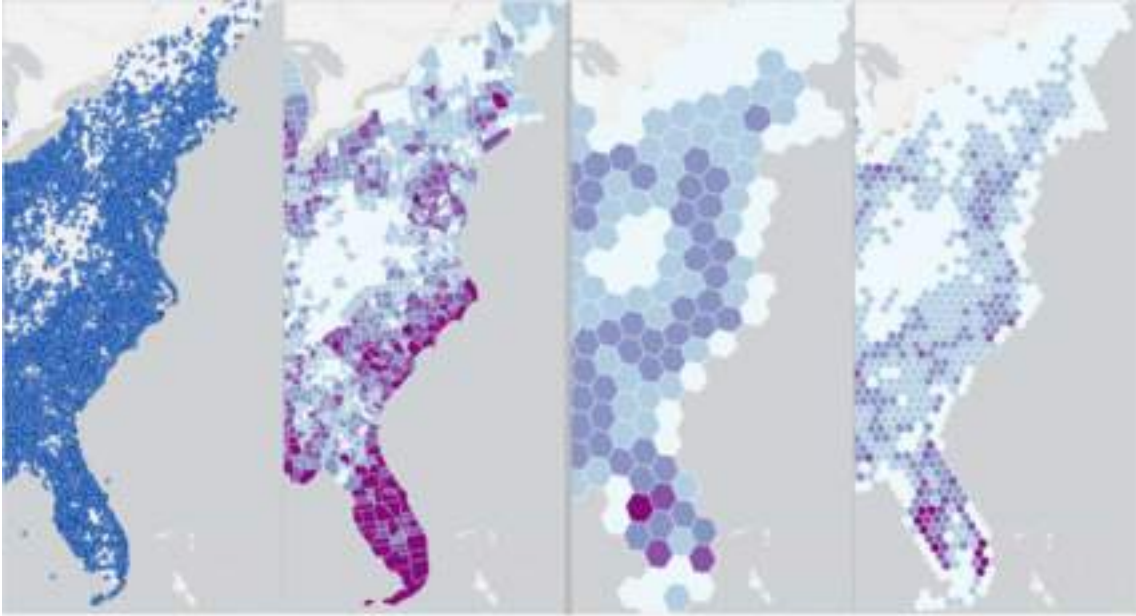
Grid sistemleri, Uber’in faaliyet gösterdiği şehirlerin genelinde benzer şekillere ve boyutlara sahip olmakta ve keyfi değişikliklere tabi değildirler. Grid sistemleri şehirlerdeki sokaklara ve mahalle sınırlarına tam olarak oturmasa da grid hücrelerini kümeleyerek mahalleri etkin bir şekilde temsil etmede kullanılmaktadır. Şekil 3.18’de San Francisco şehrini kaplayan rastgele oluşturulmuş altıgen kümeler görülmektedir. Amaç fonksiyonları kullanılarak kümeleme yapılabilen ve analiz için çok daha kullanışlı şekiller üretilebilmektedir (URL-8).



Şekil 3.18 San Francisco şehrini kaplayan rastgele oluşturulmuş altıgen kümeler (URL-8)

Mekânsal gruplama ya da birleştirme, büyük nokta veri setlerindeki örüntüleri analiz etmede kullanılan en iyi yöntemlerdir. Belirli bir bölge, eyalet, ülke gibi idari sınırlara konu olan verileri birleştirme ya da özetleme işlemlerinde altıgen grid yapısını kullanmanın çeşitli avantajları bulunmaktadır.

Tek tip bir altıgen grid kullanılarak sayım istatistikleri kolaylıkla görüntülenebilmektedir. İdari sınır verileri, boyut ve şekil bakımından düzensiz verilerdir. Bu bakımdan, daha büyük belirgin olarak görüldüğü ve sayısal nitelikler bakımından normalleştirme gereken durumlarda boyut yanlılığına neden olabilmektedir. Ayrıca idari sınırların aksine, altıgen gridleri zamansal değişime uğramadığı için farklı zaman dilimlerine gerçekleşen değişimlerin tespiti altıgen grid sayesinde kolaylıkla yapılabilmektedir. Özellikle, bir Dünya modeli üzerine inşa edilen H3 altıgen gridleri konumları her çözünürlükte kararlı kaldıkları için dikkat çekici niteliktedir. Bu nedenlerle, H3 altıgenleri çok ölçekli analiz ve projede standart grid olarak kullanılmaktadır. Şekil 3.19'daki ilk haritada ABD'de doğu bölgesinde konumlandırılan veri seti nokta kümeleri halinde, ikinci haritada nokta kümelerinin ilçe sınırları kullanılarak gruplanması, üçüncü haritada noktaların çözünürlük 3 seviyesindeki H3 altıgenleriyle kümelenmesi ve dördüncü haritada noktaların çözünürlük 4 seviyesindeki H3 altıgenleriyle kümelenmesi gösterilmektedir (URL-18).



Şekil 3.19 Bir veri setinin noktalar, ilçe sınırları ve H3 altıgenleriyle gösterimi (URL-18)

3.3. Web Tabanlı İnteraktif CBS Uygulaması Geliştirmede Kullanılacak Yazılımlar

CBS veri setlerinin ve uygulamalarının, yerel ağ üzerinde veya Web ortamında paylaşılması ve düzenlenmesi, özel yazılımlar ve üst seviye yazılım bilgisi gerektirdiğinden kullanıcılar/araştırmacılar açısından zorlu bir süreç olmaktadır. ESRI firması geliştirmiş olduğu ArcGIS Enterprise ürünleriyle, güçlü bir CBS sunucusunun yanı sıra çalışmalarınızı bulut tabanlı ve yerel ağ üzerinde düzenlemek ve paylaşmak için özel Web tabanlı CBS altyapısı içeren tam özellikli bir haritalama ve analiz platformunu ticari olarak sunmaktadır. ESRI firması akademik araştırmalar yapan öğrenciler için profesyonel anlamda haritalar oluşturmak ve Web tabanlı uygulamalar geliştirmek için birçok ArcGIS uygulamasını, ArcGIS Pro modüllerini, ArcGIS bulut tabanlı hizmetleri kullanım kredisini ve ESRI eğitim sitesine sınırsız erişimi yıllık 100\$ gibi bir ücretle bireysel kullanım lisansıyla öğrencilere vermektedir.

Bu çalışma kapsamında, X mesaj metinlerinden elde edilecek konumsal verilerin coğrafi kodlama (geocoding) işleminde ArcGIS World Geocoding hizmeti kullanılacaktır. Masaüstü platformda verilerin haritalar üzerine işlenmesinde, öznitelik tablolarının oluşturulmasında, mekânsal analiz işlemlerinde ve tematik haritaların oluşturulmasında ArcGIS Pro yazılımı kullanılacaktır. Masaüstü platformda hazırlanan mekânsal verilerin, öznitelik tablolarının, tematik haritaların Web ortamında hazırlanması ve sunumu için ArcGIS Online isimli bulut platformundan

yararlanılacaktır. ArcGIS Online platformu üzerinde servis edilecek konumsal veriler, öznetelik tabloları ve tematik haritalar ile interaktif tabanlı Web uygulaması geliřtirmede ArcGIS Web AppBuilder, ArcGIS Dashboards ve ArcGIS Hub yazılımları kullanılacaktır.

3.3.1. Arcgis World Geocoding

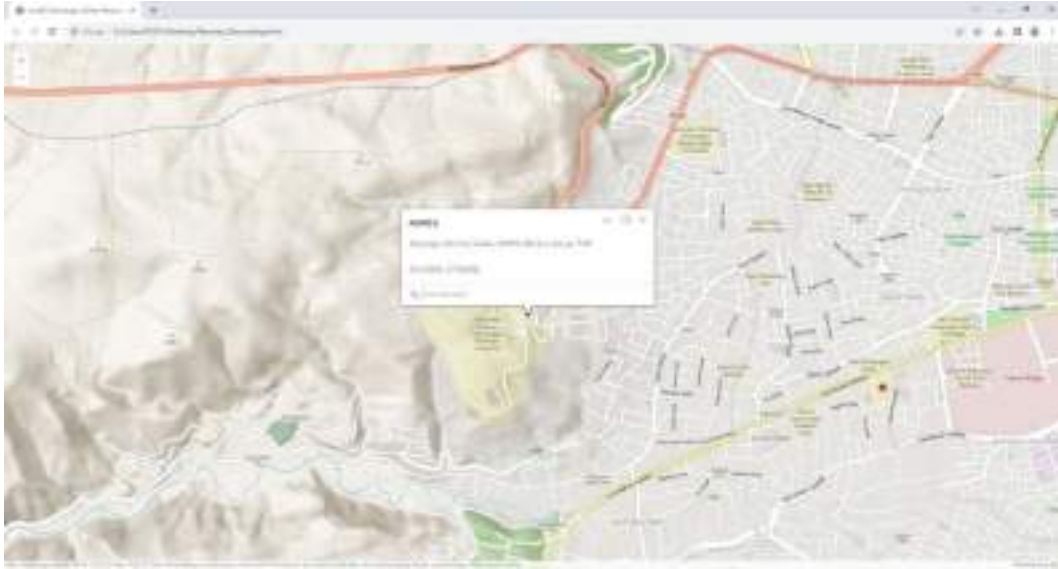
Coğrafi kodlama, metinsel ifadeleri bir adrese veya bir konuma dönüřtürme işlemi olarak tanımlanmaktadır. ArcGIS World Geocoding hizmeti sayesinde, coğrafi kodlama (metinden adres bulma) ve ters coğrafi kodlama (bir konumun yakınındaki adres metnini bulma) işlemleri gerçekleştirilmektedir (URL-21).

Metinsel bir ifadeye ait coğrafi kodlama işlemini gerçekleřtiren HTML kod bloğuna Ek-1.1’de verilmektedir. Bu örnekte, coğrafi kodlaması gerçekleřtirecek metin olarak “Köyceğiz Mah. Demeç Sok. No:44/1 Meram/ KONYA” ifadesi tanımlanmaktadır. HTML kod bloğuna ait Web sayfası görüntüsü Şekil 3.20’de görülmektedir.



Şekil 3.20 Metinsel bir ifadenin coğrafi kodlaması

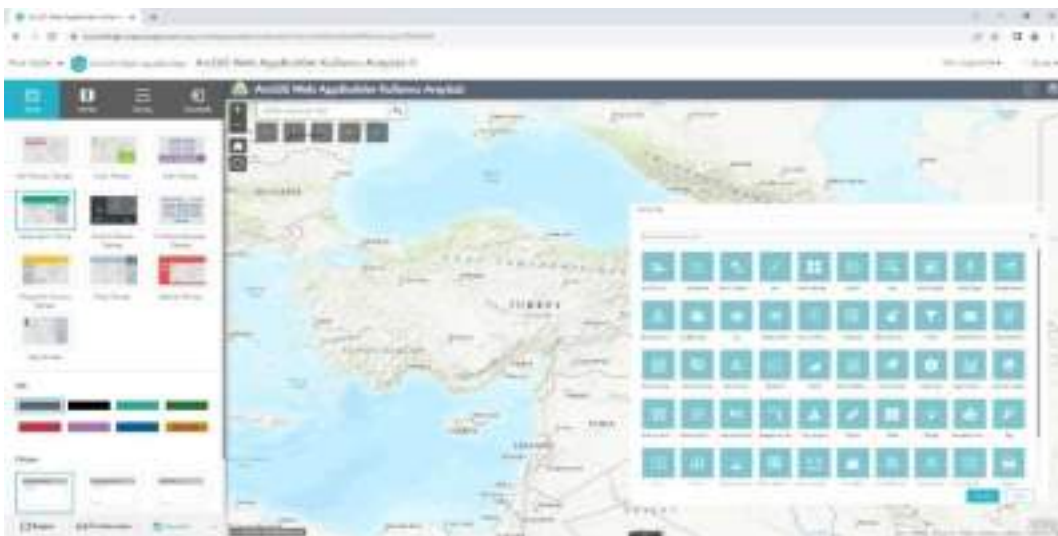
Konumu bilinen bir noktanın adresinin metinsel olarak belirlenmesi (ters coğrafi kodlama) işlemine gerçekleřtiren HTML kod bloğuna Ek-1.2’de verilmektedir. HTML kod bloğuna ait Web sayfası görüntüsü Şekil 3.21’de görülmektedir.



Şekil 3.21 Konumu bilinen bir noktanın ters coğrafi kodlaması

3.3.2. Arcgis Web Appbuilder

ArcGIS Web AppBuilder uygulaması ile masaüstü tarayıcılarda, akıllı telefonlarda ve tabletlerde çalışabilecek HTML/JavaScript tabanlı ileri düzey CBS uygulamaları herhangi bir kodlama bilgisine ihtiyaç duymadan oluşturulabilmektedir. ArcGIS Web AppBuilder uygulaması sorgu, coğrafi işlem ve baskı gibi kullanıma hazır araçlar ile kullanıcının ihtiyacına yönelik uygulamalar oluşturmaya imkân tanımaktadır. İleri düzey kullanıcılar için kendi oluşturdukları araçları uygulamaya ekleme ve uygulama içerisinde yer alan iş akış sekmelerine erişim sağlayarak kişiselleştirilmiş uygulamalar geliştirmeye imkân tanımaktadır. Şekil 3.22’de ArcGIS Web AppBuilder kullanıcı arayüzü ve uygulamada kullanıcıya sunulan araçlar görülmektedir (URL-22).



Şekil 3.22 ArcGIS Web AppBuilder kullanıcı arayüzü

3.3.3. Arcgis Dashboards

Gösterge panosu, belirli bir iş süreci veya uygulamayla ilişkili temel performans göstergelerini, ölçümleri ve önemli veri noktalarını incelemek ve görselleştirmek amacıyla kullanılan bir tür grafik arayüzüdür. Web tabanlı gösterge panosu oluşturmak karmaşık ve zorlu bir süreçtir. Gösterge panosu oluşturmak için iyi düzeyde programlama bilgisi gerekmektedir. Programcı olmayan biri için bir gösterge tablosu oluşturmak, harita, grafik, tablo vb. öğeleri tek bir Web sayfasında bir araya getirmek ve birbirleriyle etkileşime girmelerini sağlamak gerçekten çok zordur.

ArcGIS Dashboards, programcı olmayan bir kişinin bile harita, grafik, öznetelik tablosu vb. öğelerle zengin içeriklere sahip gösterge panosu oluşturmasına imkân sağlamaktadır. Araçlar gösterge panolarının yapı taşıdır. ArcGIS Dashboards, yakınlaştırma/uzaklaştırma, gösterge, pasta grafik, sütun grafik, harita, tablo vb. liste, ayarlı araçları kullanıcıya sunmaktadır. Kullanıcı sadece aracı eklemek, yapılandırmasını ayarlamak ve Web sayfası üzerindeki yerleşimleri ayarlamak zorundadır. Kullanıcı deneyimine dayanarak bir aracın farklı araçlar ya da öğeler ile iletişimini de yapılandırabilmektedir. ArcGIS Dashboards kullanılarak geliştirilen uygulamalar, duyarlı tasarım ilkelerini takip etmekte ve farklı ekran boyutlarına sahip masaüstü ya da mobil cihazlarda düzgü şekilde görüntülenebilmektedir. Şekil 3.23'te ArcGIS Dashboards kullanıcı arayüzü görülmektedir (URL-23).



Şekil 3.23 ArcGIS Dashboards kullanıcı arayüzü

3.3.4. Arcgis Hub

ArcGIS Hub, bilgi odaklı girişimler aracılığıyla verileri, insanları ve araçları koordinasyonun sağlayan, kolay yapılandırılabilen topluluk katılım platformudur. Devlet kurumları, gönüllü hizmet veren kuruluşlar ve akademik topluluklar vb. her düzeydeki organizasyon, ArcGIS Hub platformunu kullanarak iletişimi, iş birliğini, veri paylaşımını ve katılımı en üst düzeye çıkarmaktadırlar. ArcGIS Hub ile kuruluşlar, kendi ürettikleri ve yetkili oldukları verileri toplulukların daha iyi kararlar alabilmesi için açık veri olarak paylaşabilmekte, herhangi bir kodlama bilgisi olmaksızın bileşenleri sürükleyip bırakarak sınırsız sayıda Web sitesi ve Web sayfası oluşturulabilmektedirler. Şekil 3.24'te ArcGIS Hub platformunda Web sitesi oluşturmada kullanılan arayüz ve araçlar görülmektedir (URL-24).



Şekil 3.24 ArcGIS Hub Web sitesi oluşturma kullanıcı arayüzü

3.3.5. Ülkelere ait idari sınır verileri

Bu çalışmada, Dünya genelinde sosyal medya platformlarında yer alan mesaj metinlerinden elde edilecek konumsal verilerin haritalar üzerine işlenmesinde, mekânsal analiz işlemlerinde ve tematik haritaların oluşturulmasında idari sınır verisi olarak Eurostat bünyesindeki Coğrafi Bilgi Sistemi Komisyonu (GISCO) tarafından hazırlanan 03/04/2020 tarihli ve 1:1 milyon ölçekli Ülke İdari Sınır verileri kullanılmıştır (URL-26).

4. UYGULAMA

Dünya Sağlık Örgütü'nün (WHO), 11 Mart 2020'de yeni tip koronavirüs (COVID-19) pandemisini ilan etmesiyle birlikte, insanların alışlagelmiş hayat tarzları tamamen değişime uğramıştır. Dünya genelinde hükümet yetkilileri, pandemiye kontrol etmek ve gelişimini önlemek adına sokağa çıkma yasağı uygulamışlardır. İnsanlar ailelerinden, arkadaşlarından, çalışma ortamlarından, okullarından, alışveriş ve eğlence merkezlerinden uzaklaşarak kendilerini evlerine kapatmışlardır. Böylelikle, sosyal medya platformları insanların aile ve arkadaşlarıyla iletişim kurmalarında, profesyonel anlamda işlerini gerçekleştirmelerinde, alışverişlerinde, dış ortamdan haber edinmelerinde, duygu ve düşüncelerini ifade etmelerinde birincil kaynak haline dönüşmüştür. COVID-19 pandemi dönemine ilişkin küresel istatistiklere bakıldığında 31 Ekim 2022 tarihinde vaka sayılarının 627 milyonu geçtiği, 6.5 milyon üzerinde ölüm vakasının gerçekleştiği, Dünya Sağlık Örgütü (WHO) tarafından acil kullanım onayı (EUA) verilen Pfizer/BioNTech, Sinopharm, Sinovac, Moderna, Oxford/AstraZeneca, Covaxin, Sputnik V vb. aşılarla 12.8 milyar doz üzerinde aşılama yapıldığı görülmektedir (Dayan 2021; URL-25).

Bu tez çalışmasında, COVID-19 döneminde sosyal medya platformlarının trend konusu olan COVID-19 aşıları hakkında atılan tweetlerin Dünya genelindeki konumsal çıkarımları gerçekleştirilmiş, tweet trafiğinin mekânsal ve zamansal gelişimi analiz edilmiş, kümeleme analizi yöntemiyle farklı öznitelik bilgileri kullanılarak tweet atılan konumların, bölgelerin ve ülkelerin benzerlik ve farklılıkları ortaya konmaya çalışılmıştır. Tweet verilerinin mekânsal-zamansal gelişiminin görselleştirilmesi ve analiz sonuçlarının göstergeler, grafikler, haritalar üzerinde etkileşimli bir şekilde görüntüleyiciye aktarılması için Web tabanlı görsel analitik uygulama örneği oluşturulmuştur.

4.1. Veri Setinin Elde Edilmesi

Bu tez çalışması kapsamında, sosyal medya platformu X'de COVID-19 aşıları hakkında atılan tweet mesajlarının analizi hedeflenmiştir. X şirketi, X uygulama programlama arayüzü (API) aracılığıyla şirketlerin, uygulama geliştiricilerin, araştırmacıların programa dayalı olarak X'e ve verilerine erişim olanağı sağlamaktadır. Ancak kullanıcıların X (Twitter) API hesabı açtırabilmeleri için <https://developer.twitter.com> adresinden X'in geliştirici platformuna üyelik başvurularını yapmaları ve başvurularının X tarafından kabul edilmesi gerekmektedir.

Özellikle COVID-19 pandemisi döneminde, araştırmacıların X verilerine artan yoğun ilgisi karşısında X firması üyelik başvuru taleplerini *“Başvurunuz incelendi. İlginiz için Teşekkürler. Maalesef başvurunuzu onaylayamıyoruz.”* şeklinde hiçbir sebep göstermeksizin onaylamamaktadır.

Bu çalışmada X veri seti olarak, Gabriel Preda tarafından dünya genelinde Pfizer/BioNTech, Sinopharm, Sinovac, Moderna, Oxford/AstraZeneca, Covaxin ve Sputnik V aşılı hakkında toplanan ve hidratlanmış şekilde Kaggle platformunda “CC0: Public Domain” lisansı ile açık veri olarak sunduğu CSV formatındaki “COVID-19 All Vaccines Tweets” veri seti kullanılmıştır (URL-4).

4.2. Veri Setinin İncelenmesi

“COVID-19 All Vaccines Tweets” veri seti incelendiğinde, veri setinin 16 öznitelik başlığı altında 218.412 tweet verisinden oluştuğu görülmektedir. Veri setinde her bir tweet verisi için oluşturulan 16 farklı öznitelik alanı aşağıdaki şekilde incelenmektedir.

id: X’deki nesnelere ait benzersiz kimliğini göstermektedir. 64 bitlik tamsayılardan oluşmakta olup, 20 basamaklı bir tamsayıdır.

user_name: X kullanıcısının kendisini tanımladığı X ekran adı, tanıtıcı veya takma isimdir. Genellikle maksimum 15 karakterden oluşan karakter tipindeki alandır.

user_location: Kullanıcı profilinde bir konum belirtmişse bu alanı ifade etmektedir. Kullanıcı X uygulamasına konum tabanlı servis erişimine izin vermiş ve tweet atarken konum hizmetini açmış ise tweet’in atıldığı coğrafi konumu gösteren karakter tipindeki alandır.

user_description: Kullanıcının profil bilgisinin (biyografisini) tanımlandığı karakter tipindeki alandır.

user_created: Kullanıcının hesabının oluşturulduğu tarih ve saati (UTC) gösteren tarih (ISO 8601) tipindeki alandır.

user_followers: Kullanıcının takipçi sayısı gösteren alandır

user_friends: Kullanıcının takip ettiği kullanıcı sayısını gösteren alandır.

user_favourite: Kullanıcının favori sayısını gösteren alandır.

user_verified: Kullanıcının doğrulanmış hesap olup olmadığını gösteren boolean (doğru-yanlış) tipindeki alandır.

date: Tweet'in oluşturulma zamanını gösteren tarih (ISO 8601) tipindeki alandır.

text: Tweet metninin UTF-8 formatında yer aldığı alandır.

hashtags: Kullanıcının oluşturmuş olduğu tweet hakkındaki bilgileri (sözcüklerin başına “#” işareti ekleyerek) grupladığı alandır.

source: Kullanıcının tweet oluşturduğu uygulamanın ismini gösteren karakter tipindeki alandır. X kullanıcısının Web üzerinden, mobil cihazdan ya da başka bir uygulama üzerinden paylaşım yaptığını göstermektedir.

retweets: Kullanıcının oluşturduğu tweet'in başka kullanıcılar tarafından tweet'lenme sayısını göstermektedir.

favorites: Kullanıcının oluşturduğu tweet'in başka kullanıcılar tarafından favorilenme sayısını göstermektedir.

is_retweet: Kullanıcının oluşturduğu tweet'in başka kullanıcı tarafından tekrar tweetlenme durumunu “True” ve “False” (doğru-yanlış) şeklinde gösterildiği alandır.

Veri setinde yer alan “id” öznitelik alanı incelendiğinde, veri setinde 218.412 adet benzersiz değer olduğu yani tweet oluşturulduğu görülmüştür.

Veri setinde yer alan “user_name” alanı istatistiksel olarak incelendiğinde, 82.806 adet farklı kullanıcı adı tanımlaması olduğu, 1 adet kullanıcı adı verisinin tanımlanmadığı (Null) ve en çok tekrar eden kullanıcı adının “CoWIN Blore 18-44” olduğu tespit edilmiştir.

Veri setinde yer alan “user_location” alanı istatistiksel olarak incelendiğinde, 24.592 adet farklı konumsal ifadenin bulunduğu, 64.348 adet kullanıcı konum verisinin tanımlanmadığı (Null) ve en çok tekrar eden kullanıcı konum verisinin “Bengaluru, India” olduğu tespit edilmiştir.

Veri setinde yer alan “user_created” alanı istatistiksel olarak incelendiğinde, kullanıcı hesaplarından en eskisinin 15 Temmuz 2006 21:06:00 tarihinde oluşturulduğu, en yeni kullanıcı hesabının 5 Kasım 2021 16:28:50 tarihinde oluşturulduğu ve veri

setindenki en fazla tweet'in (12.210 adet) 29 Haziran 2021 09:33:13 tarihinde açılan kullanıcı hesabı tarafından oluşturulduğu tespit edilmiştir.

Veri setinde yer alan “user_verified” alanı istatistiksel olarak incelendiğinde, 18.501 adet tweet'in doğrulanmış kullanıcı hesaplarından oluşturulduğu ve 199.911 adet tweet'in doğrulanmamış kullanıcı hesaplarından oluşturulduğu tespit edilmiştir.

Veri setinde yer alan “date” alanı istatistiksel olarak incelendiğinde, COVID-19 aşılıları hakkındaki ilk tweet'in 12 Aralık 2020 tarihinde oluşturulduğu, en son tweet'in 6 Kasım 2021 10:50:32 tarihinde oluşturulduğu ve en fazla tweet'in (151 adet) 3 Kasım 2021 tarihinde oluşturulduğu tespit edilmiştir.

Veri setinde yer alan “text” alanı istatistiksel olarak incelendiğinde, tüm tweet verilerinin metin alanının bir ifade ile tanımlandığı, en kısa tweet metin uzunluğunun 13 karakter ve en uzun tweet metninin 227 karakter uzunluğuna sahip olduğu tespit edilmiştir.

Veri setinde yer alan “hashtags” alanı istatistiksel olarak incelendiğinde, 57.777 adet farklı hashtags tanımlaması yapıldığı, 45.949 adet tweet'in hashtags verisi girilmediği, en çok tekrar eden (11.113 adet) “[‘COVAXIN', 'BBMP']” hashtags verisi olduğu, en uzun hashtags verisinin 16 adet gruplandırma sözcüğü içeren “[‘jandj', 'bbcworld', 'cnn', 'nyt', 'wapo', 'dw', 'un', 'eu', 'boris', 'london', 'uk', 'astrazeneca', 'pfizer', 'moderna', 'novovax', 'Corona']” hashtags verisi olduğu tespit edilmiştir.

Veri setinde yer alan “source” alanı istatistiksel olarak incelendiğinde, tweet verilerinin 368 farklı uygulama platformundan oluşturulduğu, 115 adet tweet'in oluşturulduğu platform bilgisinin bulunmadığı (Null), en fazla tweet verisinin (56.496 adet) “Twitter for Android” uygulama platformu ile oluşturulduğu tespit edilmiştir.

4.3. Veri Setinden Coğrafi Bilgi Çıkarımı ve Coğrafi İşaretlemenin Gerçekleştirilmesi

“COVID-19 All Vaccines Tweets” veri setinde yer alan tweet'lerin harita ekranı üzerinde konumlandırılması için “user_location” öznitelik verileri kullanılmıştır. X kullanıcılarına “user_location” öznitelik alanına istedikleri metinsel ifadeyi girmelerini kısıtlamadığı için yeryüzünde bir konum ifade eden tanımlamaların bulunması ve “user_location” öznitelik alanı boş bırakılan mesajların belirlenmesi işlemi gerçekleştirilmiştir. Bu işlemler için sql sorgu dilinden yararlanılmıştır.

Kullanıcı profil kısmında konum bilgisini belirten alana herhangi bir ifade yazmamış (boş bırakmış) ya da X uygulamasında konum tabanlı hizmetlerden yararlanılmasını engellemiş ise bu tweetler;

```
user_location IS NULL
```

sql sözdizimi ile tespit edilmiştir.

Kullanıcı profil kısmında konum bilgisini belirten alana konumunu koordinat çiftleri ile manuel olarak tanımlamış ya da X uygulamasının konum tabanlı hizmetlerden yararlanmasına izin vererek kullanıcının tweet attığı sırada konumunu açarak attığı tweetler;

```
user_location LIKE '%°%' And
user_location LIKE '%N%' And
user_location LIKE '%S%' And
user_location LIKE '%E%' And
user_location LIKE '%W%' And
user_location LIKE '%K%' And
user_location LIKE '%G%' And
user_location LIKE '%D%' And
user_location LIKE '%B%'
```

sql sözdizimleri ile belirlenmiştir.

Kullanıcı profil kısmında konum bilgisi alanına ülke ismini ana dilinde kullandığı yazım karakterleri ile ifade etmişse bu tweetler;

```
user_location LIKE '%الإمارات العربية المتحدة%' Or
user_location LIKE '%Босна И Херцеговина%' Or
user_location LIKE '%Deutschland%' Or
user_location LIKE '%France%' Or
user_location LIKE '%საქართველო%' Or
user_location LIKE '%조선민주주의인민공화국%' Or
```

```

user_location LIKE '%Северна Македонија%' Or
user_location LIKE '%Portugal%' Or
user_location LIKE '%ประเทศไทย%' Or
user_location LIKE '%Türkiye%' Or

```

ve diğer ülkeler ait tanımlardan oluşan sql sözdizimleri ile belirlenmiştir.

Kullanıcı profil kısmında konum bilgisi alanına ülke ismini İngilizce karşılığı ile ifade etmişse bu tweetler;

```

user_location LIKE '%The United Arab Emirates%' Or
user_location LIKE '%Bosnia and Herzegovina%' Or
user_location LIKE '%Germany%' Or
user_location LIKE '%France%' Or
user_location LIKE '%Georgia%' Or
user_location LIKE '%North Korea%' Or
user_location LIKE '%North Macedonia%' Or
user_location LIKE '%Portugal%' Or
user_location LIKE '%Thailand %' Or
user_location LIKE '%Turkey%' Or

```

ve diğer ülkeler ait İngilizce tanımlardan oluşan sql sözdizimleri ile belirlenmiştir.

Kullanıcı profil kısmında konum bilgisi alanına ülkesini id/kod bilgileri ile ifade etmişse ile ifade etmişse bu tweetler;

```

user_location LIKE '%AE%' Or
user_location LIKE '%BA%' Or
user_location LIKE '%DE%' Or
user_location LIKE '%FR%' Or
user_location LIKE '%GE%' Or
user_location LIKE '%KP%' Or
user_location LIKE '%MK%' Or

```

```

user_location LIKE '%PT%' Or
user_location LIKE '%TH%' Or
user_location LIKE '%TR%' Or

```

ve diğer ülkeler ait id/kod tanımlarından oluşan sql sözdizimleri ile belirlenmiştir.

Kullanıcı profil kısmında konum bilgisi alanına ülkesini ISO3 ülke kodu bilgileri ile ifade etmişse bu tweetler;

```

user_location LIKE '%ARE%' Or
user_location LIKE '%BIH%' Or
user_location LIKE '%DEU%' Or
user_location LIKE '%FRA%' Or
user_location LIKE '%GEO%' Or
user_location LIKE '%PRK%' Or
user_location LIKE '%MKD%' Or
user_location LIKE '%PRT%' Or
user_location LIKE '%THA%' Or
user_location LIKE '%TUR%' Or

```

ve diğer ülkeler ait ISO3 ülke kodu tanımlarından oluşan sql sözdizimleri ile belirlenmiştir.

Kullanıcı profil kısmında konum bilgisi alanına bir şehir ismi tanımlamışsa bu kullanıcıların konumlarının tespitinde simplemaps firması tarafından hazırlanan veri setinden yararlanılmıştır. Bu veri setinde Dünya'da öne çıkan şehirlere (büyükşehirler, başkentler vb.) ait şehir isimleri, şehirlerin coğrafi konum bilgileri, ait olduğu ülke bilgileri yer almaktadır. Örneğin kullanıcı profil kısmında konum bilgisi alanına bu veri setinde yer alan Hindistan'ın öne çıkan şehirlerinden olup olmadığının sorgusu;

```

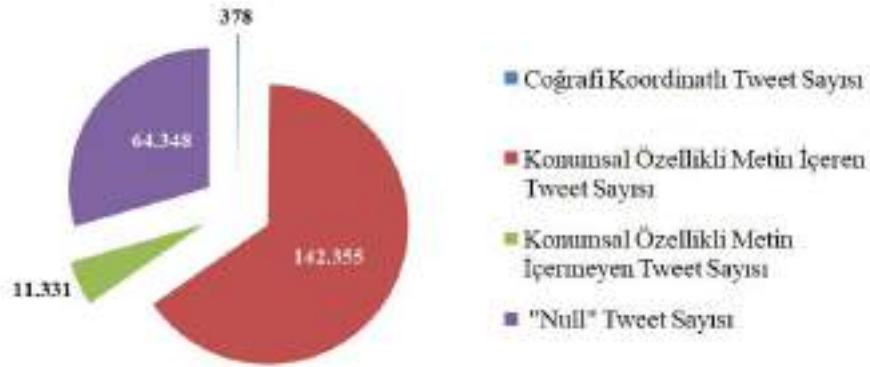
user_location LIKE '%Delhi%' Or
user_location LIKE '%Mumbai%' Or
user_location LIKE '%Bangalore%' Or

```

ve veri setinde yer alan diğer şehir isimlerinden oluşan sql sözdizimleri ile belirlenmiştir.

“COVID-19 All Vaccines Tweets” veri setindeki, “user_location” özneliği alanında coğrafi koordinat çiftleri, kullanıcının anadilindeki ülke isimleri, İngilizce ülke isimleri, ülke kodları, ISO3 ülke kodları ve şehir isimleri sorgulamalarıyla eşleşen ve “Null” değere sahip olmayan tweetlerin ArcGIS World Geocoding servisi kullanılarak Dünya haritası üzerinde konumlandırma işlemi gerçekleştirilmiştir.

“COVID-19 All Vaccines Tweets” veri setinde yer alan 218.412 adet tweetten 142.733 adedinin Dünya haritası üzerinde konumlandırması gerçekleştirilmiştir. 142.733 adet tweet verisinden sadece 378 adet tweetin coğrafi koordinat içerdiği tespit edilmiştir. Dünya haritası üzerinde konumlandırılması sağlanamayan 75.679 adet tweet verisinden 64.348 tweet verisinin “user_location” değerinin “Null” olduğu ve geriye kalan 11.331 adet verinin konum ifade etmeyen, “A Place of Hope”, “all around the world”, “always somewhere...”, “Anywhere But Here”, “Around the World”, “At Home”, “At Moon”, “At Sea”, “At your Heart”, “At Home”, “Behind You”, “Church”, “City In The Sea”, “Cloud”, “Could Be Anywhere”, “Deep Web”, “Does it matter?”, “Earth”, “Living in my dreams”, “Mars”, “Metro”, “Not from your country!” vb. kelime yada cümleler içerdiği tespit edilmiştir (Şekil 4.1).



Şekil 4.1 COVID-19 All Vaccines Tweets” veri seti "user_location” öznelik alanı veri detayı

X uygulamasında, “user_location” öznelik bilgisinde yer alan konum bilgilerinin kullanıcı tarafından metinsel girildiği (konum tabanlı hizmetler kullanılarak atılan hariç) göz önüne alındığında, kullanıcının konum bilgisi olarak birden fazla coğrafi konumu ifade eden tanımlamalar yapabileceği muhtemeldir. Bu gibi durumlarda Geocoding servislerinin, konumsal tanımlamaları ifade eden kelime dizisi içinde son kelimeyi birincil konum, önceki ifadeleri ise alt seviye konum olarak algıladığı ve işlem yaptığı bilinmektedir.

Örneğin; kullanıcı “user_location” öznitelik bilgisini “Turkey – Israel” kelime dizisi ile tanımlamışsa Geocoding servisi konumsal ifadenin ana belirteci olarak “Israel” kelimesi kullanacak, alt seviye konum olarak da “Turkey” kelimesini kullanacaktır. Geocoding işlemi sonucunda bu konumsal ifadenin İsrail devletinde bir konum olarak işlendiği görülmektedir (Şekil 4.2).



Şekil 4.2 "user_location" öznitelik bilgisi "Turkey-Israel" olan verisi örneği

X kullanıcısının, tweet attığı anda birden fazla lokasyonda bulunma ihtimalinin bulunamayacağı gerçeği ve “user_location” özniteliğinde belirtilen konum bilgilerinin yüzde yüz doğru bir konum belirten tanımlamalar olmadığı gerçeği göz önüne alınarak Geocoding servisinin bu şekildeki konumlandırma işlemleri kabul edilmiştir.

4.4. Veri Setindeki Gizli Verilerin Keşfi

Bu çalışmadaki hedeflerden biri, büyük veri içerisinde yer alan gizli örüntülerin keşfedilmesi, keşfedilen örüntülerle analiz çalışması gerçekleştirilerek sonuç ürün olan anlamlandırılmış verinin yani bilginin kullanıcıya sunulmasıdır.

“COVID-19 All Vaccines Tweets” veri seti incelendiğinde, toplamda 218.412 adet tweet oluşmaktadır. Tweet’lerin 142.733 adedinin coğrafi işaretlemesi Geocoding servisi kullanılarak gerçekleştirilmiştir. Ancak, “142.733 adet tweet hangi ülkelerden oluşturulmuştur? ya da “Ülkelerdeki COVID-19 hakkında tweet atan kullanıcı sayısı nedir?”. Çalışma kapsamında veri setinde yer alan ancak doğrudan erişilemeyen aşağıdaki soruların cevapları, veri setinde gerçekleştirilen sql sorgulamaları ve mekânsal analiz teknikleriyle elde edilerek yeni öznitelik alanları oluşturulmuştur.

4.4.1. COVID-19 aşıları hakkındaki tweetler hangi ülkelerden oluşturulmuştur?

Bu sorunun tespitinde mekânsal analiz yöntemlerinde biri olan Bindirme/Çakıştırma (Overlay) analizinden yararlanılmıştır. Dünya haritası üzerine konumlandırılan tweet veri noktaları ile ülkelere ait idari sınır verileri çakıştırma analizi uygulanarak COVID-19 aşıları hakkındaki atılan tweetlerin hangi ülkelerden oluşturuldukları ve ülkelere ait oluşturulan tweet sayılarının bilgileri elde edilmiştir (Çizelge 4.1).

Çizelge 4.1 Ülkeler bazında COVID-19 aşıları hakkında atılan tweet sayıları

Ülke Adı	Tweet Sayısı	Ülke Adı	Tweet Sayısı	Ülke Adı	Tweet Sayısı
Afganistan	30	Almanya	1195	Umman	51
Arnavutluk	43	Gana	75	Pakistan	2677
Cezayir	19	Cebelitarık	5	Filistin	18
Andorra	7	Yunanistan	118	Panama	27
Angola	45	Grönland	11	Papua Yeni Gine	18
Antarktika	118	Guam	9	Paraguay	22
Antigua ve Barbuda	3	Guatemala	25	Peru	88
Arjantin	878	Guernsey	2	Filipinler	2919
Ermenistan	69	Gine	4	Polonya	96
Arunaçal Pradeş	8	Guyana	17	Portekiz	71
Avustralya	1715	Haiti	8	Porto Riko	72
Avusturya	193	Honduras	14	Katar	398
Azerbaycan	90	Hong Kong	554	Romanya	17
Bahamalar	9	Macaristan	136	Rusya	1436
Bahreyn	48	İzlanda	26	Ruanda	21
Bangladeş	666	Hindistan	56080	Saint Helena	6
Barbados	17	Endonezya	379	Saint Kitts and Nevis	8
Belarus	23	Iran	90	Saint Lucia	1
Belgium	679	Iraq	22	Saint Vincent ve Grenadinler	1
Benin	3	İrlanda	901	Samoa	1
Bermuda	10	Man Adası	18	San Marino	1
Bhutan	13	İsrael	190	Sao Tome ve Prinsipe	1
Bolivya	24	İtalya	766	Suudi Arabistan	173
Bosna Hersek	16	Jamaika	86	Senegal	6
Botsvana	20	Cemmu ve Keşmir	597	Senkaku Adaları	1
Brezilya	284	Japonya	564	Sırbistan	134
British Virjin Adaları	8	Jersey	21	Seyşeller	10
Bruney	21	Ürdün	37	Sierra Leone	17
Bulgaristan	57	Kazakistan	10	Singapur	481
Burkina Faso	13	Kenya	636	Sint-Maarten	1
Burundi	3	Kuveyt	78	Slovakya	107
Kamboçya	146	Kırgızistan	4	Slovenya	20
Kamerun	24	Laos	8	Solomon Adaları	5
Kanada	9543	Letonya	4	Somali	34
Cayman Adaları	3	Lübnan	184	Güney Afrika	1079
Orta Afrika Cumhuriyeti	64	Lesotho	3	Güney Kore	108
Çad	3	Liberya	4	Güney Sudan	2
Şili	121	Libya	12	İspanya	370
Çin	2668	Lihtenştayn	4	Sri Lanka	2333

Kolombiya	123	Litvanya	18	Sudan	6
Komorlar	1	Lüksemburg	21	Surinam	2
Kongo	9	Makao	8	İsveç	159
Cook Adaları	2	Madagaskar	2	İsviçre	469
Costa Rica	5	Malavi	6	Suriye	25
Kotdivuar	6	Malezya	1664	Tacikistan	7
Hırvatistan	24	Maldivler	49	Tayland	937
Küba	27	Mali	10	Doğu Timor	8
Kıbrıs	117	Malta	53	Togo	4
Çekya	144	Marşal Adaları	3	Tonga	2
Kongo Demokratik Cumhuriyeti	5	Morityus	57	Trinidad and Tobago	184
Danimarka	233	Meksika	228	Tunus	81
Cibuti	1	Moldova	4	Türkiye	517
Dominika	1	Monako	5	Türkmenistan	1
Dominik Cumhuriyeti	22	Moğolistan	21	Uganda	100
Ekvator	36	Karadağ	30	Ukrayna	108
Mısır	113	Fas	64	Birleşik Arap Emirlikleri	1190
El Salvador	11	Mozambik	10	Birleşik Krallık	8072
Ekvator Ginesi	1	Myanmar/Burma	124	Tanzanya	22
Estonya	13	Namibya	46	Amerika	30922
Esvatini	3	Nepal	118	Uruguay	7
Etiyopya	24	Hollanda	590	US Virgin Adaları	40
Falkland Adaları	5	Yeni Kaledonya	26	Özbekistan	11
Faroe Adaları	1	Yeni Zelanda	81	Vanuatu	3
Fiji	39	Nikaragua	6	Vatikan	1
Finlandiya	148	Nijer	3	Venezuela	106
Fransa	1021	Nijerya	260	Vietnam	80
Gabon	5	Kuzey Kore	8	Yemen	3
Gambiya	14	Kuzey Makedonya	28	Zambiya	4
Gürcistan	24	Norveç	148	Zimbabve	361

4.4.2. COVID-19 aşılı hakkında tweetler hangi kullanıcılar tarafından oluşturulmuştur?

Veri seti incelendiğinde, tweet verilerinin öncelikli anahtarı 20 basamaktan oluşan “tweet_id” alanıdır. X kullanıcısının adını belirten “user_name” ve tweet’in oluşturulduğu tarihi belirten “date” öznitelik alanları incelendiğinde; aynı isme sahip birden fazla kullanıcı olduğu ve bu kullanıcıların aynı zamanlarda tweet oluşturduğu görülmektedir. Bu nedenle, bu sorunun tespitinde veri setindeki 16 öznitelik alanında doğrudan bir sorgulama ya da filtreleme işlemi gerçekleştirilememektedir.

Aynı isme sahip birden fazla kullanıcının birbirinden ayırma işleminde, Dünya’nın farklı bölgelerinde aynı zaman diliminde aynı kullanıcı adı ile bir X hesabı oluşturulamayacağı öngörülmüştür. Bu nedenle; “user_name” ve “user_created” öznitelik alanları birleştirilerek yeni bir öznitelik alanı oluşturulmuştur (Çizelge 4.2).

Çizelge 4.2 "user_name" ve "user_created" öznitelik alanlarıyla yeni öznitelik alanı oluşturulması

user_name	user_created	name_date
A**** K*** G*****	3.11.2015 13:40:58	A**** K*** G*****@3.11.2015 13:40:58

Çizelge 4.2’de görüldüğü üzere yeni oluşturulan öznitelik alanı incelendiğinde, bir isim ve zaman dilimini saniye esaslı alan tarihin birleştirilmesiyle oluşturulan yeni veriyi içeren tweetlerin, Dünya’nın farklı bölgelerinde oluşturulma ihtimali gerçekten çok düşüktür. Elde edilen yeni “name_date” öznitelik alanında gerçekleştirilen gruplama ve filtreleme işlemleriyle kullanıcı bazlı tweet sayıları bilgisi elde edilmiştir. “COVID-19 All Vaccines Tweets” veri setinde yer alan 218.412 adet tweet verisi 86.558 farklı kullanıcı tarafından oluşturulmuştur. Çizelge 4.3’te COVID-19 aşılı hakkında en çok tweet atan 10 kullanıcı görülmektedir.

Çizelge 4.3 Tweet sayılarının kullanıcılara göre dağılımı

Name_date	User_name	Toplam Tweet Sayısı
CoWIN Blore 18-44@29.06.2021 09:33:13	CoWIN Blore 18-44	12210
CowinBangalore@17.05.2021 18:56:45	CowinBangalore	11495
VaxBLR@21.06.2021 08:44:34	VaxBLR	5710
Owl 🦉 @26.05.2021 19:13:22	Owl 🦉	4851
Workout Solutions@20.09.2010 17:01:08	Workout Solutions	2251
Owl 🦉 @22.05.2021 15:14:51	Owl 🦉	1851
Owl 🦉 @23.05.2021 16:18:16	Owl 🦉	1719
eLéPhunk@16.05.2014 04:25:35	eLéPhunk	836
Xukki 🌍 @21.05.2020 15:54:09	Xukki 🌍	672
Sputnik V@11.08.2020 09:12:38	Sputnik V	507

4.4.3. Ülkelerin COVID-19 aşılı hakkındaki tweet atan kullanıcı sayıları kaçtır?

Bu sorunun çözümünde çalışma kapsamında kullanılan “Ülkeler” veri tablosu ve “Kullanıcı Bazlı Atılan Tweet Sayıları” veri tablosu arasında birden çoka (1:N) veri tabanı ilişkisi kurulmuştur (Şekil 4.3).

**Şekil 4.3** Ülkeler ve “kullanıcı bazlı atılan tweet sayıları” veri tabloları arasında birden çoka (1:N) ilişki

Şekil 4.3’te görüldüğü gibi “Ülkeler” ve “Kullanıcı Bazlı Atılan Tweet Sayıları” veri tabloları arasında ilişki kurulduktan sonra ülke bazlı yapılan gruplama ve filtreleme yöntemleriyle ülke bazlı kullanıcı bilgileri elde edilmiştir (Çizelge 4.4).

Çizelge 4.4 COVID-19 aşısı hakkında tweet oluşturan kullanıcı sayısının ülkelere göre dağılımı

Ülke Adı	User Count	Ülke Adı	User Count	Ülke Adı	User Count
Afganistan	22	Almanya	654	Umman	27
Arnavutluk	19	Gana	52	Pakistan	1627
Cezayir	14	Cebelitarık	4	Filistin	12
Andorra	5	Yunanistan	73	Panama	13
Angola	14	Grönland	3	Papua Yeni Gine	5
Antarktika	22	Guam	7	Paraguay	14
Antigua ve Barbuda	3	Guatemala	18	Peru	68
Arjantin	212	Guernsey	2	Filipinler	2115
Ermenistan	17	Gine	4	Polonya	65
Arunaçhal Pradeş	5	Guyana	9	Portekiz	45
Avustralya	845	Haiti	8	Porto Riko	27
Avusturya	90	Honduras	9	Katar	72
Azerbaycan	56	Hong Kong	148	Romanya	13
Bahamalar	9	Macaristan	50	Rusya	234
Bahreyn	25	İzlanda	16	Ruanda	16
Bangladeş	249	Hindistan	16153	Saint Helena	1
Barbados	14	Endonezya	202	Saint Kitts and Nevis	5
Belarus	14	İran	56	Saint Lucia	1
Belgium	255	İraq	12	Saint Vincent ve Grenadinler	1
Benin	3	İrlanda	445	Samoa	1
Bermuda	3	Man Adası	9	San Marino	1
Bhutan	10	İsrael	57	Sao Tome ve Prinsipe	1
Bolivya	17	İtalya	323	Suudi Arabistan	93
Bosna Hersek	15	Jamaika	33	Senegal	3
Botsvana	18	Cemmu ve Keşmir	238	Senkaku Adaları	1
Brezilya	115	Japonya	245	Sırbistan	41
British Virjin Adaları	3	Jersey	17	Sejšeller	8
Bruney	17	Ürdün	26	Sierra Leone	7
Bulgaristan	13	Kazakistan	10	Singapur	312
Burkina Faso	8	Kenya	342	Sint-Maarten	1
Burundi	3	Kuveyt	36	Slovakya	30
Kamboçya	46	Kırgızistan	3	Slovenya	16
Kamerun	15	Laos	7	Solomon Adaları	5
Kanada	3835	Letonya	4	Somali	27
Cayman Adaları	2	Lübnan	77	Güney Afrika	422
Orta Afrika Cumhuriyeti	27	Lesotho	3	Güney Kore	56
Çad	2	Liberya	3	Güney Sudan	2
Şili	84	Libya	9	İspanya	206
Çin	260	Lihtenştayn	2	Sri Lanka	460
Kolombiya	82	Litvanya	11	Sudan	5
Komorlar	1	Lüksemburg	16	Surinam	2
Kongo	5	Makao	6	İsveç	89
Cook Adaları	2	Madagaskar	2	İsviçre	258
Kosta Rika	5	Malavi	6	Suriye	16
Kotdivuar	6	Malezya	744	Tacikistan	5
Hırvatistan	16	Maldivler	31	Tayland	296
Küba	14	Mali	7	Doğu Timor	6
Kıbrıs	33	Malta	16	Togo	3
Çekya	52	Marşal Adaları	3	Tonga	2
Kongo Demokratik Cumhuriyeti	4	Morityus	18	Trinidad and Tobago	82
Danimarka	75	Meksika	164	Tunus	42
Cibuti	1	Moldova	4	Türkiye	127
Dominika	1	Monako	5	Türkmenistan	1
Dominik Cumhuriyeti	12	Moğolistan	18	Uganda	66

Ekvator	20	Karadağ	7	Ukrayna	30
Mısır	53	Fas	34	Birleşik Arap Emirlikleri	405
El Salvador	10	Mozambik	7	Birleşik Krallık	4661
Ekvator Ginesi	1	Myanmar/Burma	97	Tanzanya	18
Estonya	8	Namibya	25	Amerika	18326
Esvatini	3	Nepal	82	Uruguay	7
Etiyopya	15	Hollanda	368	US Virgin Adaları	11
Falkland Adaları	4	Yeni Kaledonya	14	Özbekistan	7
Faroe Adaları	1	Yeni Zelanda	58	Vanuatu	2
Fiji	23	Nikaragua	5	Vatikan	1
Finlandiya	72	Nijer	3	Venezuela	30
Fransa	426	Nijerya	148	Vietnam	43
Gabon	5	Kuzey Kore	8	Yemen	3
Gambiya	11	Kuzey Makedonya	7	Zambiya	4
Gürcistan	15	Norveç	91	Zimbabve	155

4.4.4. Pfizer Biontech aşısı hakkında oluşturulan tweetler hangileridir?

Pfizer Biontech aşısı hakkında oluşturulan tweetlerin tespitinde, “COVID-19 All Vaccines Tweets” veri setindeki “text” ve “hashtags” öznitelik alanları kullanılmıştır. “text” ve “hashtags” öznitelik alanlarında gerçekleştirilen

```
text LIKE '%pfiz%' Or text LIKE '%bion%' Or text LIKE '%Pfiz%' Or text LIKE '%Bion%' Or
```

```
hashtags LIKE '%pfiz%' Or hashtags LIKE '%bion%' Or
hashtags LIKE '%Pfiz%' Or hashtags LIKE '%Bion%' Or text
LIKE '%BFI%' Or hashtags LIKE '%BFI%' Or text LIKE '%BION%'
Or hashtags LIKE '%BION%'
```

şeklinde oluşturulan sql sorguları ile Pfizer Biontech aşısı hakkındaki oluşturulan tweetler belirlenmiştir.

4.4.5. Sinopharm aşısı hakkında oluşturulan tweetler hangileridir?

Sinopharm aşısı hakkında oluşturulan tweetleri tespitinde, “COVID-19 All Vaccines Tweets” veri setindeki “text” ve “hashtags” öznitelik alanları kullanılmıştır. “text” ve “hashtags” öznitelik alanlarında gerçekleştirilen

```
text LIKE '%Sinopharm%' Or text LIKE '%sinopharm%' Or
text LIKE '%SINOPHARM%' Or
```

```
hashtags LIKE '%Sinopharm%' Or hashtags LIKE
'%sinopharm%' Or hashtags LIKE '%SINOPHARM%'
```

şeklinde oluşturulan sql sorguları ile Sinopharm aşısı hakkında oluşturulan tweetler belirlenmiştir.

4.4.6. Sinovac aşısı hakkındaki tweetler hangileridir?

Sinovac aşısı hakkında oluşturulan tweetleri tespitinde, “COVID-19 All Vaccines Tweets” veri setindeki “text” ve “hashtags” öznitelik alanları kullanılmıştır. “text” ve “hashtags” öznitelik alanlarında gerçekleştirilen

```
text LIKE '%Sinovac%' Or text LIKE '%sinovac%' Or text
LIKE '%SINOVAC%' Or
```

```
hashtags LIKE '%Sinovac%' Or hashtags LIKE '%sinovac%'
Or hashtags LIKE '%SINOVAC%'
```

şeklinde oluşan sql sorguları ile Sinovac aşısı hakkında oluşturulan tweetler belirlenmiştir.

4.4.7. Moderna aşısı hakkındaki tweetler hangileridir?

Moderna aşısı hakkında oluşturulan tweetleri tespitinde, “COVID-19 All Vaccines Tweets” veri setindeki “text” ve “hashtags” öznitelik alanları kullanılmıştır. “text” ve “hashtags” öznitelik alanlarında gerçekleştirilen

```
text LIKE '%Moderna%' Or text LIKE '%moderna%' Or text
LIKE '%MODERNA%' Or
```

```
hashtags LIKE '%Moderna%' Or hashtags LIKE '%
moderna%' Or hashtags LIKE '% MODERNA%'
```

şeklinde oluşan sql sorguları ile Moderna aşısı hakkında oluşturulan tweetler belirlenmiştir.

4.4.8. Oxford AstraZeneca aşısı hakkındaki tweetler hangileridir?

Oxford AstraZeneca aşısı hakkında oluşturulan tweetleri tespitinde, “COVID-19 All Vaccines Tweets” veri setindeki “text” ve “hashtags” öznitelik alanları kullanılmıştır. “text” ve “hashtags” öznitelik alanlarında gerçekleştirilen

```
text LIKE '%Astra%' Or text LIKE '%Zeneca%' Or text
LIKE '%astra%' Or text LIKE '%zeneca%' Or text LIKE
'%oxford%' Or text LIKE '%Oxford%' Or text LIKE '%ASTRA%'
Or text LIKE '%ZENECA%' Or
```

```
hashtags LIKE '%Astra%' Or hashtags LIKE '%Zeneca%' Or
hashtags LIKE '%astra%' Or hashtags LIKE '%zeneca%' Or
```

```
hashtags LIKE '%oxford%' Or hashtags LIKE '%Oxford%' Or
hashtags LIKE '%ASTRA%' Or hashtags LIKE '%ZENECA%'
```

şeklinde oluşan sql sorguları ile Oxford AstraZeneca aşısı hakkında oluşturulan tweetler belirlenmiştir.

4.4.9. Covaxin aşısı hakkındaki tweetler hangileridir?

Covaxin aşısı hakkında oluşturulan tweetleri tespitinde, “COVID-19 All Vaccines Tweets” veri setindeki “text” ve “hashtags” öznitelik alanları kullanılmıştır. “text” ve “hashtags” öznitelik alanlarında gerçekleştirilen

```
text LIKE '%Covaxin%' Or text LIKE '%covaxin%' Or text
LIKE '%COVAX%' Or
```

```
hashtags LIKE '%Covaxin%' Or hashtags LIKE '%covaxin%'
Or hashtags LIKE '%COVAX%'
```

şeklinde oluşan sql sorguları ile Covaxin aşısı hakkında oluşturulan tweetler belirlenmiştir.

4.4.10. Sputnik V aşısı hakkındaki tweetler hangileridir?

Sputnik V aşısı hakkında oluşturulan tweetleri tespitinde, “COVID-19 All Vaccines Tweets” veri setindeki “text” ve “hashtags” öznitelik alanları kullanılmıştır. “text” ve “hashtags” öznitelik alanlarında gerçekleştirilen

```
text LIKE '%Sputnik%' Or text LIKE '%sputnik%' Or text
LIKE '%SPUTN%' Or
```

```
hashtags LIKE '%Sputnik%' Or hashtags LIKE
'%sputnik%' Or hashtags LIKE '%SPUTN%'
```

şeklinde oluşan sql sorguları ile Sputnik V aşısı hakkında oluşturulan tweetler belirlenmiştir.

4.4.11. Johnson&Johnson aşısı hakkındaki tweetler hangileridir?

Johnson&Johnson aşısı hakkında oluşturulan tweetleri tespitinde, “COVID-19 All Vaccines Tweets” veri setindeki “text” ve “hashtags” öznitelik alanları kullanılmıştır. “text” ve “hashtags” öznitelik alanlarında gerçekleştirilen

```
text LIKE '%Johnson%' Or text LIKE '%johnson%' Or text
LIKE '%JOHNSON%' Or
```

```
hashtags LIKE '%Johnson%' Or hashtags LIKE '%johnson%'
Or hashtags LIKE '%JOHNSON%'
```

şeklinde oluşan sql sorguları ile Johnson&Johnson aşısı hakkında oluşturulan tweetler belirlenmiştir.

4.4.12. Covid kelimesi hakkındaki tweetler hangileridir?

Covid kelimesi hakkında oluşturulan tweetleri tespitinde, “COVID-19 All Vaccines Tweets” veri setindeki “text” ve “hashtags” öznitelik alanları kullanılmıştır. “text” ve “hashtags” öznitelik alanlarında gerçekleştirilen

```
text LIKE '%Covid%' Or text LIKE '%covid%' Or text
LIKE '%COVID%' Or
```

```
hashtags LIKE '%Covid%' Or hashtags LIKE '%covid%' Or
hashtags LIKE '%COVID%'
```

şeklinde oluşan sql sorguları ile Covid kelimesi hakkında oluşturulan tweetler belirlenmiştir.

4.4.13. Vaccine (aşı) kelimesi hakkındaki tweetler hangileridir?

“COVID-19 All Vaccines Tweets” veri setinde yer alan tweetler Dünya genelinde toplanan veriler olduğu için çoklu dil içerebileceği, bu yüzden “vaccine” ya da “aşı” kelimesinin sorgulamasının yapılmadan çoklu dil karşılaştırmasının yapılması ve dil farklılığı olması durumunda farklı dillerden İngilizceye çevirinin yapılması gerektiği belirlenmiştir.

Veri setinde yer alan tweetlerin “text” ve “hashtags” alanları Google tarafından sunulan **DETECTLANGUAGE** aracı ile kullanılarak hangi dilde yazıldığının tespiti gerçekleştirilmiştir. **DETECTLANGUAGE** aracı kullanılırken sözdizimi şu şekildedir;

DETECTLANGUAGE (text_or_range)

text_or_range- Metin veya değerlendirilecek metin içeren hücrelere referans.

Çizelge 4.5 DETECTLANGUAGE aracının örnek kullanımı

Text	Formula	Result
I'm analyzing tweets about Covid19.	=DETECTLANGUAGE ()	en

Ich analysiere Tweets über Covid19.	de
Estoy analizando tweets sobre Covid19.	es
أقوم بتحليل التغريدات حول Covid19.	ar

Veri setinde yer alan tweetlerin “text” ve “hashtags” alanlarının **DETECTLANGUAGE** aracı ile kullanılarak hangi dilde yazıldığının tespitinden sonra farklı dillerde yazılan tweetler **GOOGLETRANSLATE** aracı kullanılarak İngilizce’ye dönüştürülmüştür. **GOOGLETRANSLATE** aracı kullanılırken sözdizimi şu şekildedir;

GOOGLETRANSLATE (text, [source_language, target_language])

text- Çevrilecek metin.

source_language - [İSTEĞE BAĞLI -varsayılan olarak]- Kaynak dilin iki harfli dil kodu girilmektedir.

target_language - [İSTEĞE BAĞLI - varsayılan olarak sistem dili]- Hedef dilin iki harfli dil kodu girilmektedir.

Çizelge 4.6 GOOGLETRANSLATE aracı ile "aşı" kelimesinin farklı dillere dönüştürülmesi

Dil	Formül	Sonuç
Almanca	GOOGLETRANSLATE ("aşı"; "tr"; "de")	Impfung
İngilizce	GOOGLETRANSLATE ("aşı"; "tr"; "en")	vaccine
İspanyolca	GOOGLETRANSLATE ("aşı"; "tr"; "es")	vacuna
Fransızca	GOOGLETRANSLATE ("aşı"; "tr"; "fr")	vaccin
Arapça	GOOGLETRANSLATE ("aşı"; "tr"; "ar")	مصل

Veri setinde yapılan dil incelemesinde; toplam 218.412 adet tweet verisinden 1.623 adetinin dilinin İngilizce olmayıp diğer dillerde yazıldığı tespit edilmiştir. Ayrıca farklı dilde yazılan 1.623 adet tweet verisinden sadece 591 adedinin harita üzerinde bir konumla eşleştiği görülmüştür (Çizelge 4.7).



Çizelge 4.7 İngilizceden farklı dillerde oluşturulan tweet sayılarının ülkelere göre dağılımı

Ülke Adı	Tweet Sayısı	Ülke Adı	Tweet Sayısı	Ülke Adı	Tweet Sayısı
Hindistan	137	İsviçre	5	Danimarka	1
Brezilya	99	Avusturya	4	Mısır	1
Tayland	40	Bangladeş	4	Finlandiya	1
Amerika	37	Kolombiya	4	Yunanistan	1
Filipinler	24	Cemmu ve Keşmir	4	Macaristan	1
Arjantin	19	Kenya	4	Lübnan	1
Rusya	19	Peru	4	Meksika	1
Almanya	18	Birleşik Krallık	4	Myanmar/Burma	1
İtalya	16	Venezuela	4	Nepal	1
Pakistan	13	Antarktika	3	Nijerya	1
Fransa	12	Belgium	3	Norveç	1
Sri Lanka	11	Endonezya	3	Portekiz	1
Türkiye	11	Birleşik Arap	3	Porto Riko	1

		Emirlikleri			
Hollanda	10	Şili	2	Romanya	1
Malezya	9	Ürdün	2	Suudi Arabistan	1
Hong Kong	6	Kuveyt	2	Singapur	1
Japonya	6	Fas	2	Tunus	1
Paraguay	6	Polonya	2	Ukrayna	1
İspanya	6	Bolivya	1	Uruguay	1
Kanada	5	Kamboçya	1	Zimbabve	1
Çin	5	Kosta Rika	1	Danimarka	1

Örneğin, Türkiye’de farklı kullanıcılar tarafından oluşturulan atılan ve içerisinde İngilizceden farklı dillerde kelime bulunan 11 adet tweet Çizelge 4.8’de görülmektedir.

Çizelge 4.8 Türkiye’de İngilizceden farklı dillerde kelime içeren tweetler

Tweet Id	User Name	Text
1405637758094356484	A**** K*** G*****	#biontec #Sinovac #asıyahayır #TedirginOlmaAsıOl #DenekOlmaTurkiye #Gazze halk https://t.co/uWx5wYdPjs
1422998727854329857	 B**** K***** A****®	Sence? #Limak #cnzburak #TermikSantrali #BirUEniversite #YurticiSarkoyYalancidir #TamamUlanTamam #Sinovac... https://t.co/Hgza0GcqZZ
1435175624914636803	D**** Y***** G**	 6 ay sonra loading 2 doz #Sinovac - 1 doz #Biontek %51,4 2 doz #Biontek %39,44 https://t.co/xokNIP2kpO
1363560121725558784	G**** D****	#DerbyMilano #Marasli #MenajerimiAra #Hercai #HanKer #MileyCyrus #PlasticHearts #RedVelvet #SputnikV #MucizeDoktor... https://t.co/vdZldWwJsg
1428958237081096194	S*****	I.Doiz 142.785 %72 II.Doiz 118.744 %60 #Karabuk @karabukism @KarabukValiligi @fuatgurel76 @KarabukEAH... https://t.co/8B3IwCTEsn
1429140963952365568	S*****	I.Doiz 143.155 %72 II.Doiz 119.119 %60 #Karabuk @karabukism @KarabukValiligi @fuatgurel76 @KarabukEAH... https://t.co/T5Z66694jr
1380455182484520965	L**** Y**	ExPM and Slovak Finance Minister Matovič to meet @Viktor_Orban & #PéterSzijjártó in Budapest today, to discuss "othe... https://t.co/Hs2AHYbf7L
1381960984332406785	a***** k*****	Covid19 #aşı #Biontech #Sinovac #meb #millieğitimbakanlığı #sağlıkbakanlığı #sağlıkdiliyorum https://t.co/AqzIUX85YQ
1399449698767888388	I****	Turkey reports 6,493 new coronavirus cases, 122 deaths #CoronavirusPandemic #yasaklar #Turkey #kabinetoplantısı... https://t.co/sflfLj7lZe
1399499933296709635	I****	Turkey reports 6,493 new coronavirus cases, 122 deaths #CoronavirusPandemic #yasaklar #Turkey #kabinetoplantısı... https://t.co/CQXs20HAfz
1399995107831648257	N**** T*****	#DünyaSağlıkÖrgütü'nden #Sinovac 'a onay https://t.co/jfZdbIETfc

Veri setindeki “text” ve “hashtags” öznitelik alanlarında farklı dillerin İngilizce’ye dönüştürülmesinin ardından “vaccine” kelimesi hakkında oluşturulan tweetleri tespitinde,

```
text LIKE '%vaccine%' Or text LIKE '%Vaccine%' Or text
LIKE '%VACCINA%' Or text LIKE '%vaccina%' Or text LIKE
'%vac%' Or text LIKE '%VAC%' Or
```

```
hashtags LIKE '%vaccine%' Or hashtags LIKE '%Vaccine%'
Or hashtags LIKE '%VACCINA%' Or hashtags LIKE '%vaccina%'
Or hashtags LIKE '%vac%' Or hashtags LIKE '%VAC%'
```

şeklinde oluşan sql sorguları ile vaccine (Aşı) kelimesi hakkında oluşturulan tweetler belirlenmiştir.

COVID-19 All Vaccines Tweets” veri seti incelendiğinde gerçekleştirilen veri keşfi sonrasında keşfi gerçekleştirilen 8 farklı aşı markası ile “Covid” ve “Vaccine” kelimelerine ait tweetlerin dağılımları Çizelge 4.9’da görülmektedir.

Çizelge 4.9 Sekiz farklı aşı markası ile "Covid" ve "Vaccine" kelimelerini içeren tweetlerin dağılımları

Konu	Konumlandırılmış Tweet Sayısı	Konumlandırılmamış Tweet Sayısı	Toplam Tweet Sayısı
Pfizer Biontech	14.232	7.252	21.484
Sinopharm	6.588	2.279	8.867
Sinovac	7.433	3.527	10.960
Moderna	31.711	14.296	46.007
Oxford AstraZeneca	5.712	2.613	8.325
Covaxin	37.700	32.144	69.844
Sputnik V	12.713	4.990	17.703
Johnson&Johnson	1.383	764	2.147
Covid	41.530	13.699	55.229
Vaccine (Aşı)	69.608	26.691	96.299

4.5. Veri setinde Keşfi Yapılan Bilgilerin Tematik Haritalar ile Sunumu

Bu çalışmada kullanılan X veri setinde farklı sorgu ve analiz teknikleriyle elde edilen COVID-19 aşı markaları ile “covid” ve “vaccine” kelimeleri hakkında tweetler Dünya’nın farklı ülkelerinden oluşturulmuştur. Bu nedenle elde edilen bilgilerin harita üzerinde gösteriminde altlık olarak Dünya haritası kullanılmıştır. Dünya haritasının projeksiyon seçiminde, Arthur Robinson tarafından yeryüzünün tamamını göstermek üzere geliştirilen Robinson projeksiyonu tercih edilmiştir. Robinson projeksiyonu,

- Ana karalar parçalanmaksızın yeryüzünün tamamının kesintisiz gösterebilme,

- Ana karalarda en az şekil bozulması ile mümkün olduğunca alan oranlarının korunması,
- Her yaştan kullanıcıların anlayabileceği basit bir coğrafi ağ,
- Tüm şeklin 1:1.4 oranında olup kâğıt boyutlarına uygun olması,

özelliklerine sahiptir (Bildirici, 2018). Sekiz farklı COVID-19 aşısı markaları ile “covid” ve “vaccine” kelimeleri hakkında oluşturulan tweetlerin Dünya haritası üzerindeki yayılımlarını gösteren nokta işaretli haritalar (dot-map) üretilmiştir (Ek-2.1 – Ek-2.10).

4.6. Nokta Geometrisinden Alan Geometrisine Geçiş

Mekânsal verilere bakıldığında, temel vektör mekânsal veri elemanlarının nokta, çizgi ve alan olduğu bilinmektedir. Kartografyada, bir su kuyusunu nokta geometrisi ile temsil ederken, bu su kuyusundan çıkan suyun taşınmasında kullanılan su hattını çizgi geometrisiyle ve sulanan bölge ise alan geometrisi ile temsil edilmektedir.

Mekânsal analiz işlemleri açısından bakıldığında veriler genellikle haritalar ve öznitelik tabloları üzerinde bir araya getirilerek kullanılmaktadır. Örneğin, bir bölgede yapılacak hafriyat çalışmasındaki kübaj hesapları için, yüksekliği temsil eden arazi detayları noktasal veri olarak toplanmaktadır. Toplanan noktasal veriler arasında farklı enterpolasyon yöntemleriyle nokta sıklaştırması yapılmakta ve elde edilen tüm nokta verilerinin birleştirilmesiyle verinin alansal olarak ifade edildiği bir yüzey alanı oluşturulmaktadır.

Bu çalışma kapsamında, Dünya'nın farkı noktalarından oluşturulan tweetlerin analiz işlemleri için, yani tweet atılan noktalar yerine tweet atılan bölgelerin belirlenmesi için nokta geometrisinden alan geometrisine geçiş amaçlanmıştır. Bu amaç doğrultusunda öncelikli olarak noktasal tweet verisinin Convex Hull'u (Dışbükey Kabuk) belirlenmiştir.

4.6.1. Convex hull (dışbükey kabuk)

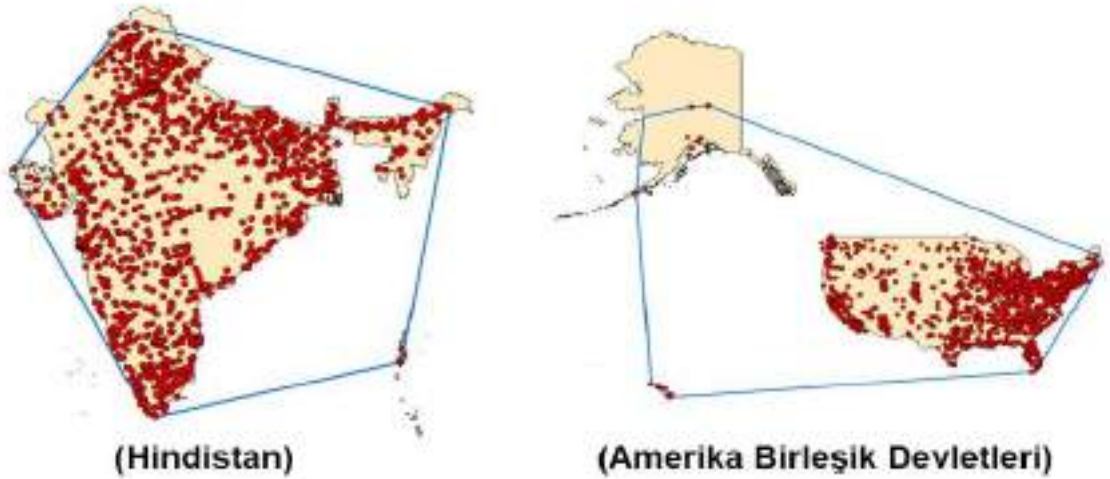
Bir nokta kümesinin dışbükey kabuğu, tüm noktaları içeren en küçük dışbükey kümedir. Matematiksel olarak ifade edilirse, herhangi bir S kümesinin dışbükey kabuğu, S 'yi içeren tüm dışbükey kümelerin kesişimidir. Şekil 4.4'te, solda gösterilen düzlemdeki noktalar kümesi S olarak kabul edilirse, S 'nin dışbükey kabuğu ortada şekilde gösterilen dolu kümedir. Bu örnekte sonlu bir nokta kümesi ele alındığından, dışbükey kabuk her zaman köşeleri başlangıç kümesi t 'nin bir alt kümesi olan bir

çokgen olacaktır. Bu nedenle aşağıda şekilde sağda gösterilen çokgen sınırı dışbükey kabuk (convex hull) olarak kabul edilmektedir (URL-30).



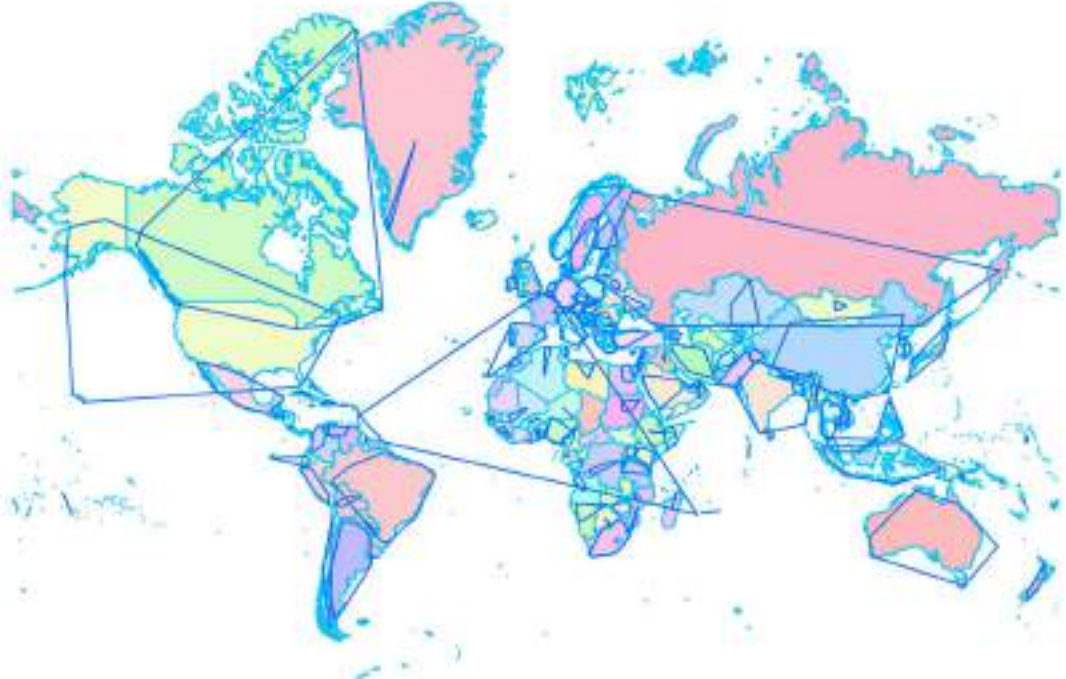
Şekil 4.4 Nokta kümesi (solda), dışbükey kabuk alanı (ortada) ve dışbükey kabuk (sağda) (URL-30)

Nokta geometrisi şeklinde bulunan tweet verilerinin oluşturulduğu ülke isimleri özniteliğine göre gruplandırılarak tweetlerin oluşturulmuş oldukları alanların dışbükey kabukları (convex hull) belirlenmiştir. En çok tweet oluşturulan ülkeler olan Hindistan ve Amerika Birleşik Devletleri'nde oluşturulan tweetlerin dışbükey kabukları Şekil 4.5'te mavi çoklu doğru ile gösterilmektedir.



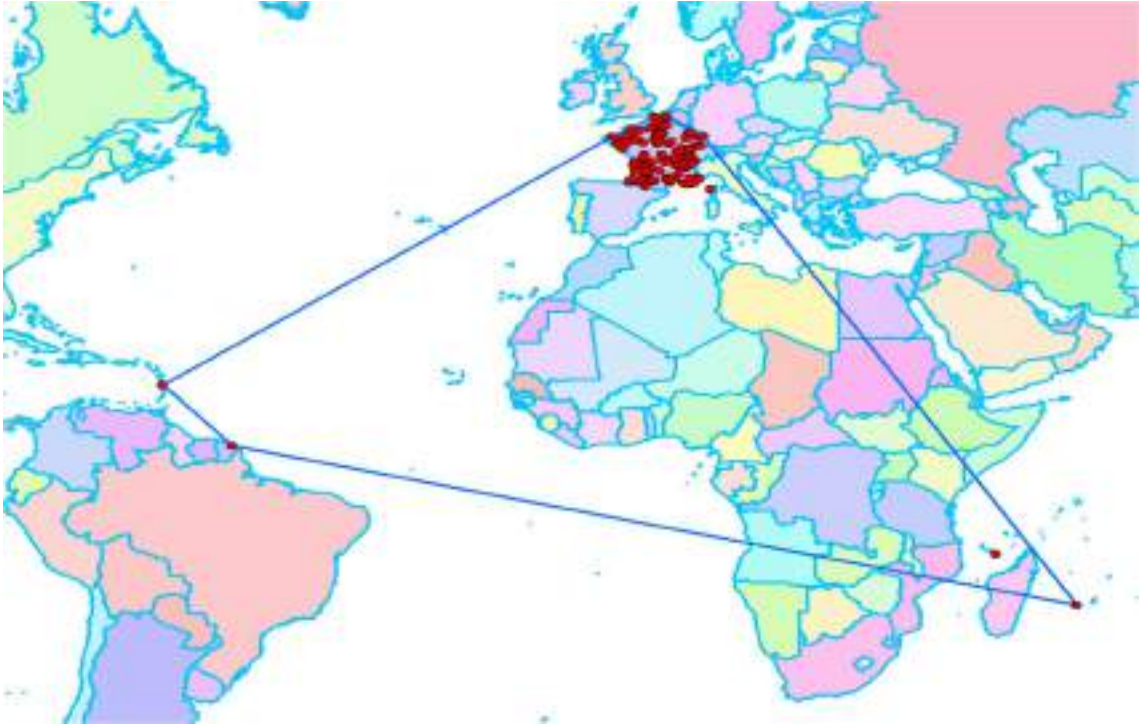
Şekil 4.5 Hindistan ve ABD'de oluşturulan tweetlerin dışbükey kabukları

Dışbükey kabuklar ülkeler bazında tek başına incelendiklerinde belirli bir alanı ifade etmelerine karşın, ülkelerin bir arada toplu gösterimlerinde oluşturulan Dışbükey kabukların birbirleriyle kesişmekte olduğu ya da diğer ülkelere ait dışbükey kabukları örttüğü görülmektedir (Şekil 4.6).



Şekil 4.6 Farklı ülkelerden oluşturulan tweetlerin dışbükey kabukları

Ayrıca, Dünya’da bilinen coğrafi konumları haricinde farklı bölgelerde de bağlı toprak parçaları bulunan ülkelerden (Örneğin Fransa’nın deniz aşırı bölgelerde toprak parçası bulunmaktadır) oluşturulan tweetlerin dışbükey kabukları başka ülkelere atılan tweetlerin dışbükey kabuklarını örtmektedir (Şekil 4.7).



Şekil 4.7 Fransa ve bağlı topraklarından oluşturulan tweetlerin dışbükey kabukları

Mevcut veri seti için dışbükey kabuk yöntemi değerlendirildiğinde, elde edilen sonuçların çalışmanın niteliği açısından uygun olmadığı görülmektedir. Tabii ki bu durum dışbükey kabuk yönteminden kaynaklanmadığı bilinmekte olup, dışbükey kabuk oluşturmada kullanılan düşük hassasiyetli (ülke adı) öznitelik verisi yerine yüksek hassasiyetli veri (eyalet, şehir, mahalle vb.) kullanılırsa elde edilen sonuçların daha başarılı olacağı düşünülmektedir.

4.6.2. Uber H3: Hiyerarşik altıgen hücrelerle mekânsal indeksleme

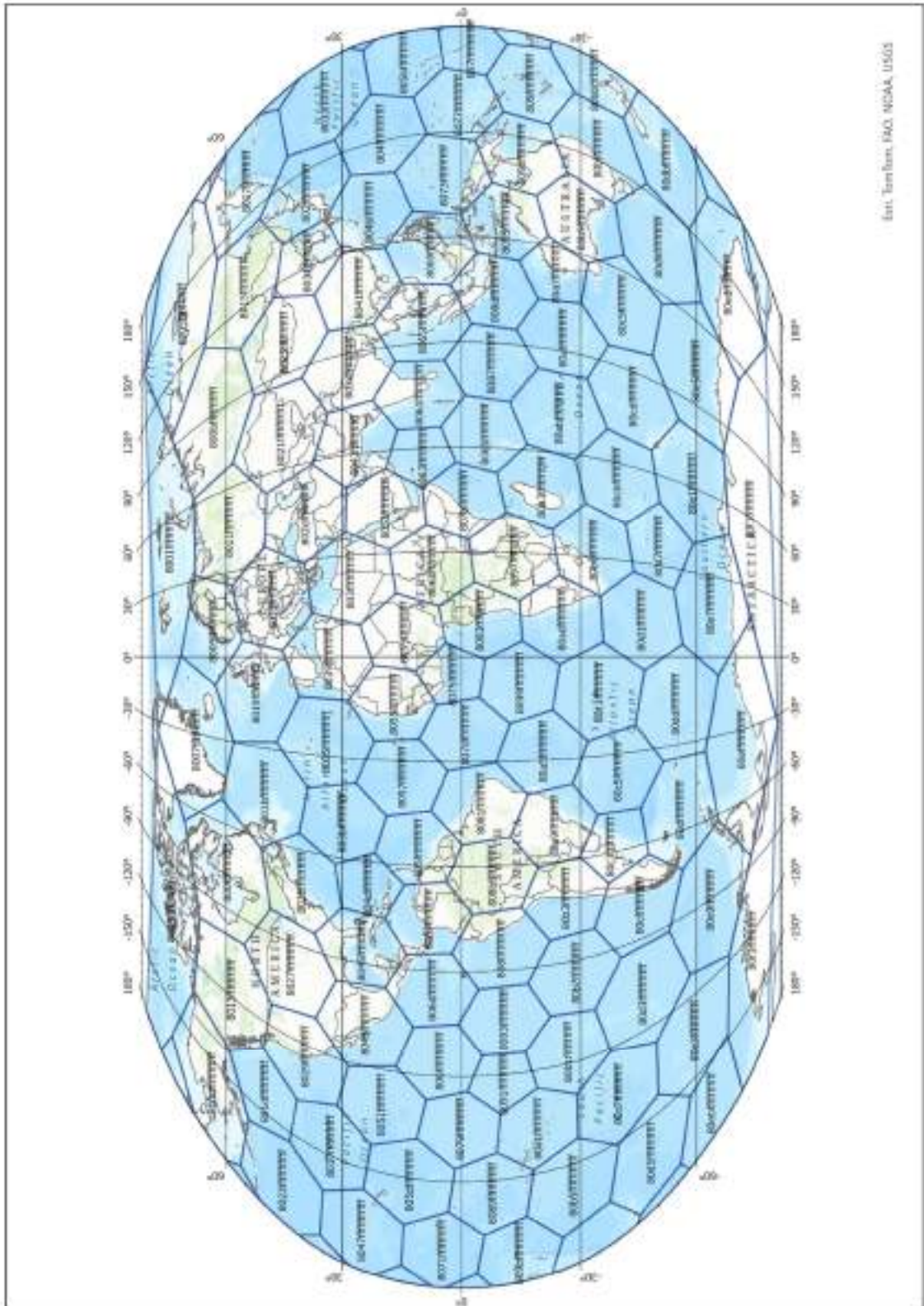
Bölüm 3.2 ve alt başlıklarında açıklandığı üzere, bu yöntemde yerküre mekânsal veri kümelerini analiz etmek amacıyla tanımlanabilir grid hücrelerine ayrılmaktadır. H3 hiyerarşik algoritması ile yerküre 16 farklı çözünürlük seviyesinde modellenmektedir. İlk çözünürlük seviyesi olan 0'ncı seviye de yer küre 122 benzersiz hücre ile temsil edilirken, maksimum çözünürlük seviyesi olan 15'inci seviyede yer küre 569.707.381.193.162 adet benzersiz hücre ile temsil edilmektedir. Şekil 4.8'de Dünya haritası üzerinde 0'ncı çözünürlük seviyesinde üretilen 122 adet H3 altıgen hücreler görülmektedir.

Düzenli altıgen grid hücrelerin en önemli avantajı; obje sayıları, alanları, uzunlukları, yoğunlukları, sosyal istatistikî veriler, sıcaklıklar, yükseklikler, doğal ve ekonomik göstergeler vb. yapıdaki heterojen verilerin hücre tabanlı depolanabilmesidir. Düzenli bir grid hücre yapısında toplanan veriler farklı analiz yöntemleriyle değerlendirilerek, yerkürenin farklı bir bölgesindeki düzenli grid hücre ile kolaylıkla ve şeffaflıkla karşılaştırılmaktadır.

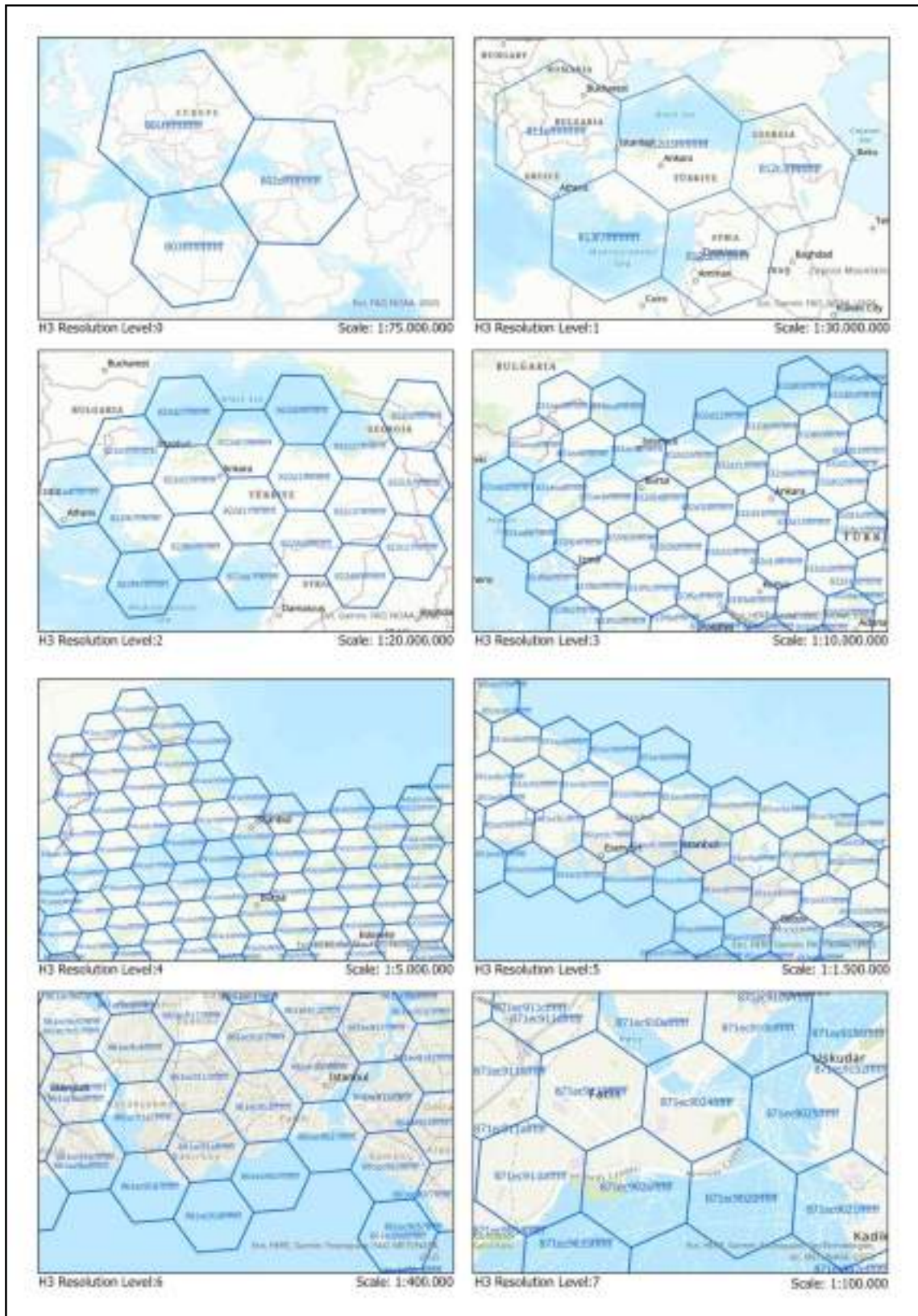
Yerküre üzerinde Türkiye'nin bulunduğu bölgeyi H3 hiyerarşik altıgen hücreler ile modellediğimizde sekiz farklı seviye için oluşturulan hücre sayıları Çizelge 4.10'da ve harita gösterimleri Şekil 4.9'da görülmektedir.

Çizelge 4.10 Türkiye'nin içinde bulunduğu sekiz farklı seviyedeki H3 altıgen hücre bilgileri

H3 Çözünürlük Seviyesi	Benzersiz Hücre Sayısı
0	3
1	5
2	19
3	88
4	487
5	3.102
6	20.703
7	141.988



Şekil 4.8 Çözünürlük = 0 seviyesinde oluşturulan 122 adet H3 altgen hücreleri



Şekil 4.9 Türkiye'nin sekiz farklı seviyedeki H3 altıgen hücrelerle indekslenmesi

Dünya haritası üzerinde coğrafi işaretlenmesi gerçekleştirilen sekiz farklı aşık markasıyla, “covid” ve “vaccine” kelimeleri hakkında oluşturulan tweetlerin H3 hiyerarşik altıgen hücreleriyle mekânsal indekslemesi gerçekleştirilmiştir. Veri setinde

yer alan tweet verilerin coğrafi işaretlemesinde kullanılan adres bilgisi hassasiyetinin düşük olması (şehir seviyesinde) olması nedeniyle altıgen hücrelerin ortalama kenar uzunluğunun yaklaşık 182 km olduğu 2'nci çözünürlük seviyesine kadar H3 altıgen hücreleri oluşturularak mekânsal indekslemeler gerçekleştirilmiştir. Üç farklı çözünürlük seviyesinde elde edilen toplam hücre sayıları Çizelge 4.11'de gösterilmekte olup, oluşturulan H3 hiyerarşik altıgen hücre sayılarının sekiz farklı aşı markasına, “covid” ve “vaccine” kelimelerine göre dağılımları Çizelge 4.12'de gösterilmektedir.

Çizelge 4.11 Coğrafi işaretlemesi gerçekleştirilen tweetlerin üç farklı seviyedeki H3 altıgen hücre bilgileri

H3 Çözünürlük Seviyesi	Benzersiz Hücre Sayısı
0	78
1	303
2	934

Çizelge 4.12 H3 hiyerarşik altıgen hücrelerinin kategorisel dağılımı

Tweet Konusu	H3 Hiyerarşik Altıgen Hücre Sayıları		
	Çözünürlük:0	Çözünürlük:1	Çözünürlük:2
Pfizer Biontech	69	222	546
Sinopharm	53	167	341
Sinovac	61	196	426
Moderna	72	242	648
Oxford AstraZeneca	60	172	386
Covaxin	63	185	401
Sputnik V	64	197	449
Johnson&Johnson	51	119	241
Covid	73	255	677
Vaccine	75	284	796

Coğrafi işaretlemesi gerçekleştirilen tweetlerin H3 hiyerarşik altıgen hücreleriyle mekânsal indeksleme sonucu elde edilen Çizelge 4.11 ve Çizelge 4.12 karşılaştırıldığında; tüm çözünürlük seviyelerinde sekiz farklı aşı markası içerisinde en fazla hücre sayısının Moderna aşısı hakkında olduğu, en az hücre sayısının ise Johnson&Johnson aşısı hakkında olduğu görülmektedir. Çözünürlük seviyesi arttıkça hücresel yayılımların genelleştirmeden uzaklaşarak daha hassaslaştığı görülmektedir. Örneğin; 0'nci çözünürlük seviyesinde “covid” kelimesi hakkında tweetleri içeren hücre oranının yaklaşık %94 iken 2'nci çözünürlük seviyesinde yaklaşık %72 olduğu görülmektedir. Sekiz farklı aşı markasıyla, “covid” ve “vaccine” kelimeleri hakkında oluşturulan üç farklı çözünürlük seviyesindeki H3 hiyerarşik altıgen hücreleri beş farklı sınıfa ayrılarak koroplet haritalar hazırlanmıştır (Ek-3.1 – Ek-3.30).

Pfizer Biontech aşısı hakkındaki tweetlerin H3 hiyerarşik altıgen hücreleriyle çözünürlük 0 seviyesinde oluşturulan Ek-3.1’de yer alan koroplet harita incelendiğinde, en yüksek tweet oluşumlarının Amerika Birleşik Devletleri’ni ve Kanada’nın güney bölgesini kapsayan 3 hücre, Birleşik Krallık ve Avrupa Kıtası’nı kapsayan 2 hücre ve Hindistan’ın kuzey bölgesi, Bangladeş ve Nepal bölgelerini kapsayan 1 hücre olmak üzere, toplam 6 farklı mekânsal indekse sahip hücrelerde olduğu görülmektedir.

H3 hiyerarşik altıgen hücrelerinin çözünürlük 1 seviyesinde oluşturulduğu Ek-3.2’de yer alan koroplet harita incelendiğinde, Pfizer Biontech aşısı hakkında oluşturulan tweetlerin konumlarının bir seviyede daha hassaslaştığı ve tweetlerin daha doğru bir mekânsal indekslemeye gerçekleştiği görülmektedir. Bu seviyede Pfizer Biontech aşısı hakkında oluşturulan tweetlerin en yoğun olduğu hücrelerin Birleşik Krallık ve Hollanda’yı kapsayan 1 hücre olduğu, ikinci derece yoğunlukların ise Amerika Birleşik Devletleri’nin doğu bölgesini kapsayan 3 hücre ve orta kesiminde yer alan 1 hücre, Almanya – Fransa – İsviçre – İtalya bölgesini kapsayan 1 hücre ve Hindistan’ın kuzey bölgesini kapsayan 1 hücre olmak üzere, toplam 6 farklı mekânsal indeksli hücreler olduğu görülmektedir.

Çözünürlük 2 seviyesindeki H3 hiyerarşik altıgen hücrelerinden oluşan Ek-3.3’te yer alan koroplet harita incelendiğinde, Pfizer Biontech aşısı hakkında oluşturulan tweetlerin en yoğun bölgelerin Amerika Birleşik Devletleri’nin Philadelphia, New York, Hartford şehirlerini kapsayan 1 hücre ve Kansas’ın batı bölgesini kapsayan 1 hücre, Kanada’nın Toronto şehrini kapsayan 1 hücre, Birleşik Krallık’ın güney bölgesi ile Fransa’nın Lille şehrini kapsayan 2 hücre, Hindistan’ın New Delhi, Jaipur, Jabalpur şehirlerinin kapsayan 2 hücre olmak üzere toplam 7 farklı mekânsal indeksli hücreler olduğu görülmektedir.

Pfizer Biontech aşısı hakkında üç farklı çözünürlük seviyesi için hazırlanan Ek-3.1 – Ek-3.2’de yer alan koroplet haritalar ve Ek-2.1’de yer alan nokta işaretli harita karşılaştırıldığında, Pfizer Biontech aşısı hakkında oluşturulan tweetlerin Amerika Birleşik Devletleri, Kanada, Avrupa kıtası ve Hindistan merkezli olduğu görülmektedir. H3 hiyerarşik altıgen hücrelerin, Pfizer Biontech aşısı hakkında oluşturulan tweetlerin konumları ve sayıları hakkında bilgiyi daha anlaşılabilir ve analiz işlemlerinde daha kolay kullanılabilir yapıda sunduğu görülmektedir.

Sinopharm aşısı hakkındaki tweetlerin H3 hiyerarşik altıgen hücreleriyle çözünürlük 0 seviyesinde oluşturulan Ek-3.4’te yer alan koroplet harita incelendiğinde,

en yüksek tweet oluşumunun Hindistan'ın güneyini kapsayan 1 hücre olduğu görülmektedir. İkinci derece yoğunlukların ise Afganistan, Birleşik Arap Emirlikleri, Umman, İran'ın doğusu ve Hindistan'ın kuzeydoğusunu kapsayan 1 hücre, Çin'in doğusu, Güney Kore ve Kuzey Kore'yi kapsayan 1 hücre olmak üzere toplam 2 hücre olduğu görülmektedir.

Genelleştirme kavramında uzaklaşarak H3 hiyerarşik altıgen hücrelerinin çözünürlük 1 seviyesinde oluşturulduğu Ek-3.5'te yer alan koroplet harita incelendiğinde, Sinopharm aşısı hakkında oluşturulan tweetlerin konumlarının bir seviyede daha hassaslaştığı ve tweetlerin daha doğru bir mekânsal indekslemeye gerçekleştiği görülmektedir. Bu seviyede Sinopharm aşısı hakkında oluşturulan tweetlerin en yoğun olduğu hücrelerin Sri Lanka'yı kapsayan 1 hücre olduğu görülmektedir. İkinci derece yoğunlukların ise Pakistan'ın doğusunu kapsayan 1 hücre, Çin'in Pekin bölgesini kapsayan 1 hücre olmak üzere, toplam 2 farklı mekânsal indeksli hücreler olduğu görülmektedir.

Çözünürlük 2 seviyesindeki H3 hiyerarşik altıgen hücrelerinden oluşan Ek-3.6'da yer alan koroplet harita incelendiğinde, Sinopharm aşısı hakkında oluşturulan tweetlerin en yoğun bölgelerin Sri Lanka'nın Colombo, Kandy, Kalmunai şehirlerini kapsayan 2 hücre ve Çin'in Pekin şehrini kapsayan 1 hücre olmak üzere, toplam 3 farklı mekânsal indeksli hücreler olduğu görülmektedir.

Sinopharm aşısı hakkında üç farklı çözünürlük seviyesi için hazırlanan Ek-3.4 – Ek-3.6'da yer alan koroplet haritalar ve Ek-2.2'de yer alan nokta işaretli harita karşılaştırıldığında, Sinopharm aşısı hakkında oluşturulan tweetlerin Çin, Sri Lanka, Pakistan, Bangladeş ve Birleşik Arap Emirlikleri merkezli oldukları görülmektedir. H3 hiyerarşik altıgen hücrelerin, Sinopharm aşısı hakkında oluşturulan tweetlerin konumları ve sayıları hakkında bilgiyi daha anlaşılabilir ve analiz işlemlerinde daha kolay kullanılabilir yapıda sunduğu görülmektedir.

Sinovac aşısı hakkındaki tweetlerin H3 hiyerarşik altıgen hücreleriyle çözünürlük 0 seviyesinde oluşturulan Ek-3.7'de yer alan koroplet harita incelendiğinde, en yüksek tweet oluşumunun Endonezya'yı, Filipinler'i ve Vietnam'ı kapsayan 1 hücre olduğu görülmektedir. İkinci derece yoğunlukların ise Afganistan, Birleşik Arap Emirlikleri, Umman, İran'ın doğusu ve Hindistan'ın kuzey doğusunu kapsayan 1 hücre, Çin'in doğusu, Güney Kore ve Kuzey Kore'yi kapsayan 1 hücre ve Kamboçya,

Malezya, Myanmar, Tayland'ı kapsayan 1 hücre olmak üzere toplam 3 hücre olduğu görülmektedir.

H3 hiyerarşik altıgen hücrelerinin çözünürlük 1 seviyesinde oluşturulduğu Ek-3.8'de yer alan koroplet harita incelendiğinde, Sinopharm aşısı hakkında oluşturulan tweetlerin konumlarının bir seviyede daha hassaslaştığı ve tweetlerin daha doğru bir mekânsal indekslemeye gerçekleştiği görülmektedir. Bu seviyede Sinopharm aşısı hakkında oluşturulan tweetlerin en yoğun olduğu hücrelerin Filipinler'i kapsayan 1 hücre ve Çin'in Pekin bölgesinin kapsayan 1 hücre olduğu görülmektedir. İkinci derece yoğunlukların ise, Pakistan'ın doğu ve Hindistan'ın kuzeydoğu bölgesinin kapsayan 1 hücre, Kamboçya, Tayland ve Vietnam'ı kapsayan 1 hücre, Malezya'yı kapsayan 1 hücre, Endonezya'yı kapsayan 1 hücre, Filipinler kapsayan 1 hücre ve Hong Kong'u kapsayan 1 hücre olmak üzere, toplam 6 farklı mekânsal indeksli hücreler olduğu görülmektedir.

Çözünürlük 2 seviyesindeki H3 hiyerarşik altıgen hücrelerinden oluşan Ek-3.9'da yer alan koroplet harita incelendiğinde, Sinopharm aşısı hakkında oluşturulan tweetlerin en yoğun bölgelerin Çin'in Pekin şehrinin kapsayan 1 hücre ve Filipinler'in Manila şehrini kapsayan 1 hücre olmak üzere, toplam 2 farklı mekânsal indeksli hücreler olduğu görülmektedir.

Sinovac aşısı hakkında üç farklı çözünürlük seviyesi için hazırlanan Ek-3.7 – Ek-3.9'da yer alan koroplet haritalar ve Ek-2.3'te yer alan nokta işaretli harita karşılaştırıldığında, Sinovac aşısı hakkında oluşturulan tweetlerin Çin, Filipinler, Malezya, Tayland merkezli oldukları görülmektedir. H3 hiyerarşik altıgen hücrelerin, Sinovac aşısı hakkında oluşturulan tweetlerin konumları ve sayıları hakkında bilgiyi daha anlaşılabilir ve analiz işlemlerinde daha kolay kullanılabilir yapıda sunduğu görülmektedir.

Moderna aşısı hakkındaki tweetlerin H3 hiyerarşik altıgen hücreleriyle çözünürlük 0 seviyesinde oluşturulan Ek-3.10'da yer alan koroplet harita incelendiğinde, en yüksek tweet oluşumlarının Amerika Birleşik Devletleri'nin ve Kanada'nın güney bölgesini kapsayan 3 hücre olduğu görülmektedir. İkinci seviye oluşumların ise Amerika Birleşik Devletleri'nin güney sahillerini kapsayan 1 hücre, Birleşik Krallık ve Avrupa Kıtası'nı kapsayan 2 hücre, Hindistan, Bangladeş, Nepal ve Pakistan'ın kuzey doğu bölgesinin kapsayan 1 hücre olmak üzere toplam 3 hücre olduğu görülmektedir.

H3 hiyerarşik altıgen hücrelerinin çözünürlük 1 seviyesinde oluşturulduğu Ek-3.11’de yer alan koroplet harita incelendiğinde, Moderna aşısı hakkında oluşturulan tweetlerin en yoğun olduğu hücrelerin Amerika Birleşik Devletleri’nin doğu bölgesini kapsayan 3 hücrenin, orta bölgesini kapsayan 1 hücrenin ve batı bölgesini kapsayan 1 hücrenin, Birleşik Krallık, Belçika, Hollanda’yı kapsayan 1 hücre olmak üzere, toplam 6 farklı mekânsal indeksli hücreler olduğu görülmektedir.

Çözünürlük 2 seviyesindeki H3 hiyerarşik altıgen hücrelerinden oluşan Ek-3.12’de yer alan koroplet harita incelendiğinde, Moderna aşısı hakkında oluşturulan tweetlerin en yoğun bölgelerin Amerika Birleşik Devletleri’nin Philadelphia ve New York şehirlerini kapsayan 1 hücre, Garden City ve Liberal şehirlerini kapsayan 1 hücre, Los Angeles ve San Diego şehirlerini kapsayan 1 hücre ve Kanada’nın Toronto, Hamilton ve Barrie şehirlerinin kapsayan 1 hücre olmak üzere, toplam 4 farklı mekânsal indeksli hücreler olduğu görülmektedir.

Moderna aşısı hakkında üç farklı çözünürlük seviyesi için Ek-3.10 – Ek-3.12’de yer alan koroplet haritalar ve Ek-2.4’te yer alan nokta işaretli harita karşılaştırıldığında, Moderna aşısı hakkında oluşturulan tweetlerin Amerika Birleşik Devletleri, Kanada, Birleşik Krallık ve Hindistan merkezli oldukları görülmektedir. H3 hiyerarşik altıgen hücrelerin, Moderna aşısı hakkında oluşturulan tweetlerin konumları ve sayıları hakkında bilgiyi daha anlaşılabilir ve analiz işlemlerinde daha kolay kullanılabilir yapıda sunduğu görülmektedir.

Oxford AstraZeneca aşısı hakkındaki tweetlerin H3 hiyerarşik altıgen hücreleriyle çözünürlük 0 seviyesinde oluşturulan Ek-3.13’te yer alan koroplet harita incelendiğinde, en yüksek tweet oluşumlarının Birleşik Krallık bölgesini kapsayan 1 hücre olduğu görülmektedir. İkinci seviye oluşumların ise Avrupa Kıtası’nı kapsayan 1 hücre, Hindistan, Bangladeş, Nepal ve Pakistan’ın kuzey doğu bölgesinin kapsayan 1 hücre olmak üzere toplam 2 hücre olduğu görülmektedir.

H3 hiyerarşik altıgen hücrelerinin çözünürlük 1 seviyesinde oluşturulduğu Ek-3.14’te yer alan koroplet harita incelendiğinde Oxford AstraZeneca aşısı hakkında oluşturulan tweetlerin en yoğun olduğu hücrelerin Birleşik Krallık, Belçika, Hollanda’yı kapsayan 1 hücre olduğu görülmektedir. İkinci seviye oluşumların ise, Kanada’nın Toronto, Ottawa ve Montreal şehirlerinin kapsayan 1 hücre ve Hindistan’ın New Delhi, Lucknow, Jabalpur, Indore şehirlerini kapsayan 1 hücre olmak üzere, toplam 2 farklı mekânsal indeksli hücreler olduğu görülmektedir.

Çözünürlük 2 seviyesindeki H3 hiyerarşik altıgen hücrelerinden oluşan Ek-3.15'te yer alan koroplet harita incelendiğinde, Oxford AstraZeneca aşısı hakkında oluşturulan tweetlerin en yoğun bölgelerin Birleşik Krallık'ın Londra ve Fransa'nın Lille şehirlerinin kapsayan 1 hücre, Birleşik Krallık'ın Birmingham ve Cardiff şehirlerini kapsayan 1 hücre, Birleşik Krallık'ın Nottingham ve Leeds şehirlerinin kapsayan 1 hücre olmak üzere, toplam 3 farklı mekânsal indeksli hücreler olduğu görülmektedir.

Oxford AstraZeneca aşısı hakkında üç farklı çözünürlük seviyesi için hazırlanan Ek-3.13 – Ek-3.15'te yer alan koroplet haritalar ve Ek-2.5'te yer alan nokta işaretli harita karşılaştırıldığında, Oxford AstraZeneca aşısı hakkında oluşturulan tweetlerin Birleşik Krallık, Kanada ve Hindistan merkezli oldukları görülmektedir. H3 hiyerarşik altıgen hücrelerin, Oxford AstraZeneca aşısı hakkında oluşturulan tweetlerin konumları ve sayıları hakkında bilgiyi daha anlaşılabilir ve analiz işlemlerinde daha kolay kullanılabilir yapıda sunduğu görülmektedir.

Covaxin aşısı hakkındaki tweetlerin H3 hiyerarşik altıgen hücreleriyle çözünürlük 0 seviyesinde oluşturulan Ek-3.16'da yer alan koroplet harita incelendiğinde, en yüksek tweet oluşumunun Hindistan'ın güneyini ve Sri Lanka'yı kapsayan 1 hücre ve Hindistan'ın kuzey doğusunu, Nepal ve Bangladeş'i kapsayan 1 hücre olmak üzere, toplam 2 farklı mekânsal indeksli hücreler olduğu görülmektedir. İkinci derece yoğunlukların ise, Amerika Birleşik Devletleri'ni ve Kanada'nın güney bölgesini kapsayan 3 hücre, Birleşik Krallık'ı kapsayan 1 hücre ve Afganistan, Birleşik Arap Emirlikleri, Umman, İran'ın doğusu ve Hindistan'ın kuzey doğusunu kapsayan 1 hücre olmak üzere, toplam 5 farklı mekânsal indeksli hücreler olduğu görülmektedir.

Çözünürlük 1 seviyesindeki H3 hiyerarşik altıgen hücrelerinden oluşan Ek-3.17'de yer alan koroplet harita incelendiğinde, Covaxin aşısı hakkında oluşturulan tweetlerin en yoğun bölgelerin Hindistan'ın New Delhi, Lucknow, Jabalpur, Indore şehirlerini kapsayana 1 hücre ve Hindistan'ın Bengaluru ve Kochi şehirlerini kapsayan 1 hücre olmak üzere, toplam 2 farklı mekânsal indeksli hücreler olduğu görülmektedir. İkinci derece yoğunlukların ise, Hindistan'ın Haydarabad, Kolhapur ve Nagpur şehirlerini kapsayan 1 mekânsal indeksli hücre olduğu görülmektedir.

Çözünürlük 2 seviyesindeki H3 hiyerarşik altıgen hücrelerinden oluşan Ek-3.18'de yer alan koroplet harita incelendiğinde, Covaxin aşısı hakkında oluşturulan

tweetlerin en yoğun bölgenin Hindistan'ın Bengaluru ve Mysure şehirlerini kapsayan 1 mekânsal indeksli hücre olduğu görülmektedir.

Covaxin aşısı hakkında üç farklı çözünürlük seviyesi için hazırlanan Ek-3.16 – Ek-3.18'de yer alan koroplet haritalar ve Ek-2.6'da yer alan nokta işaretli harita karşılaştırıldığında, Covaxin aşısı hakkında oluşturulan tweetlerin yüksek yoğunluklar Hindistan, daha sonra Amerika Birleşik Devletleri ve Birleşik Krallık merkezli oldukları görülmektedir. H3 hiyerarşik altıgen hücrelerin, Covaxin aşısı hakkında oluşturulan tweetlerin konumları ve sayıları hakkında bilgiyi daha anlaşılabilir ve analiz işlemlerinde daha kolay kullanılabilir yapıda sunduğu görülmektedir.

Sputnik V aşısı hakkındaki tweetlerin H3 hiyerarşik altıgen hücreleriyle çözünürlük 0 seviyesinde oluşturulan Ek-3.19'da yer alan koroplet harita incelendiğinde, en yüksek tweet oluşumunun Hindistan'ın güneyini ve Sri Lanka'yı kapsayan 1 hücre ve Hindistan'ın kuzey doğusunu, Nepal ve Bangladeş'i kapsayan 1 hücre olmak üzere, toplam 2 hücre olduğu görülmektedir. İkinci derece yoğunlukların ise, Birleşik Krallık'ı kapsayan 1 hücre, Avrupa Kıtası'nı kapsayan 1 hücre ve Rusya'yı kapsayan 1 hücre olmak üzere, toplam 5 farklı mekânsal indeksli hücreler olduğu görülmektedir.

Çözünürlük 1 seviyesindeki H3 hiyerarşik altıgen hücrelerinden oluşan Ek-3.20'de yer alan koroplet harita incelendiğinde, Sputnik V aşısı hakkında oluşturulan tweetlerin en yoğun bölgenin Hindistan'ın New Delhi, Lucknow, Jabalpur, Indore şehirlerini kapsayana 1 hücre olduğu görülmektedir. İkinci derece yoğunlukların ise, Hindistan'ın Haydarabad, Kolhapur ve Nagpur şehirlerini kapsayan 1 hücre, Hindistan'ın Bengaluru ve Kochi şehirlerini kapsayan 1 hücre ve Rusya'nın Moskova, Smolensk ve Kursk şehirlerini kapsayan 1 hücre olmak üzere, toplam 3 farklı mekânsal indeksli hücreler olduğu görülmektedir.

Çözünürlük 2 seviyesindeki H3 hiyerarşik altıgen hücrelerinden oluşan Ek-3.21'de yer alan koroplet harita incelendiğinde, Sputnik V aşısı hakkında oluşturulan tweetlerin en yoğun bölgelerin Hindistan'ın New Delhi, Jaipur ve Churu şehirlerini kapsayan 1 hücre, Hindistan'ın Jabalpur ve Begamganj şehirlerini kapsayan 1 hücre ve Hindistan'ın Bengaluru, Mysure ve Shivamogga şehirlerini kapsayan 1 hücre olmak üzere, toplam 3 farklı mekânsal indeksli hücreler olduğu görülmektedir. İkinci derece yoğunluğun ise Hindistan'ın Pune, Mumbai ve Ahmadnagar şehirlerini kapsayan 1

hücre ve Rusya'nın Moskova ve Tver şehirlerini kapsayan 1 hücre olmak üzere, toplam 3 farklı mekânsal indeksli hücreler olduğu görülmektedir.

Sputnik V aşısı hakkında üç farklı çözünürlük seviyesi için hazırlanan Ek-3.19 – Ek-3.21'de yer alan koroplet haritalar ve Ek-2.7'de yer alan nokta işaretli harita karşılaştırıldığında, Sputnik V aşısı hakkında oluşturulan tweetlerin yüksek yoğunluklar Hindistan, daha sonra ise Rusya merkezli oldukları görülmektedir. H3 hiyerarşik altıgen hücrelerin, Sputnik V aşısı hakkında oluşturulan tweetlerin konumları ve sayıları hakkında bilgiyi daha anlaşılabilir ve analiz işlemlerinde daha kolay kullanılabilir yapıda sunduğu görülmektedir.

Johnson&Johnson aşısı hakkındaki tweetlerin H3 hiyerarşik altıgen hücreleriyle çözünürlük 0 seviyesinde oluşturulan Ek-3.22'de yer alan koroplet harita incelendiğinde, en yüksek tweet oluşumunun ABD'ni ve Kanada'nın güney bölgesini kapsayan 3 hücre ve Birleşik Krallık'ı kapsayan 1 hücre olmak üzere, toplam 4 farklı mekânsal indeksli hücreler olduğu görülmektedir. İkinci derece yoğunlukların ise, ABD'nin güney doğu bölgesini kapsayan 1 hücre ve Meksika'nın bir kısmını kapsayan 1 hücre, Avrupa Kıtası'nı kapsayan 1 hücre ve Hindistan'ın kuzey doğusunu, Nepal ve Bangladeş'i kapsayan 1 hücre olmak üzere, toplam 4 farklı mekânsal indeksli hücreler olduğu görülmektedir.

Çözünürlük 1 seviyesindeki H3 hiyerarşik altıgen hücrelerinden oluşan Ek-3.23'te yer alan koroplet harita incelendiğinde, Johnson&Johnson aşısı hakkında oluşturulan tweetlerin en yoğun bölgelerin ABD'nin Philadelphia ve New York şehirlerini kapsayan 1 hücre, Detroit ve Cincinnati şehirlerini kapsayan 1 hücre, Dallas ve Oklahoma City şehirlerini kapsayan 1 hücre, Los Angeles, Las Vegas ve Carson City şehirlerini kapsayan 1 hücre ve Birleşik Krallık'ın Londra, Cardiff, Glasgow şehirleri ile Hollanda'nın Amsterdam ve Fransa'nın Lille şehirlerini kapsayan 1 hücre olmak üzere, toplam 3 farklı mekânsal indeksli hücreler olduğu görülmektedir.

Çözünürlük 2 seviyesindeki H3 hiyerarşik altıgen hücrelerinden oluşan Ek-3.24'te yer alan koroplet harita incelendiğinde, Johnson&Johnson aşısı hakkında oluşturulan tweetlerin en yoğun bölgenin Amerika Birleşik Devletleri'nin Philadelphia ve Hartford şehirlerini kapsayan tek hücre olduğu görülmektedir. İkinci derece yoğunluğun ise, ABD'nin Los Angeles ve San Diego şehirlerini kapsayan 1 hücre, San Francisco şehrini kapsayan 1 hücre, Garden City ve Dodge City'i kapsayan 1 hücre, Birleşik Krallık'ın Darlington ve Norwich şehirlerini kapsayan 1 hücre, Birleşik

Krallık'ın Londra, Fransa'nın Lille ve Belçika'nın Brugge şehirlerinin kapsayan 1 hücre olmak üzere, toplam 5 farklı mekânsal indeksli hücreler olduğu görülmektedir.

Johnson&Johnson aşısı hakkında üç farklı çözünürlük seviyesi için hazırlanan Ek-3.22 – Ek-3.24'te yer alan koroplet haritalar ve Ek-2.8'de yer alan nokta işaretli harita karşılaştırıldığında, Johnson&Johnson aşısı hakkında oluşturulan tweetlerin yüksek yoğunlukla ABD, daha sonra ise İngiltere merkezli oldukları görülmektedir. H3 hiyerarşik altıgen hücrelerin, Sputnik V aşısı hakkında oluşturulan tweetlerin konumları ve sayıları hakkında bilgiyi daha anlaşılabilir ve analiz işlemlerinde daha kolay kullanılabilir yapıda sunduğu görülmektedir.

Covid kelimesi hakkındaki tweetlerin H3 hiyerarşik altıgen hücreleriyle çözünürlük 0 seviyesinde oluşturulan Ek-3.25'te yer alan koroplet harita incelendiğinde, en yüksek tweet oluşumunun Hindistan'ın güneyi ve Sri Lanka'yı kapsayan 1 hücre olduğu görülmektedir. İkinci derece yoğunlukların ise, Amerika Birleşik Devletleri'nin kuzeydoğu bölgesini kapsayan 1 hücre ve Hindistan'ın kuzeydoğusunu, Nepal ve Bangladeş'i kapsayan 1 hücre olmak üzere, toplam 2 farklı mekânsal indeksli hücreler olduğu görülmektedir.

Çözünürlük 1 seviyesindeki H3 hiyerarşik altıgen hücrelerinden oluşan Ek-3.26'da yer alan koroplet harita incelendiğinde, Covid kelimesi hakkında oluşturulan tweetlerin en yoğun bölgelerin Hindistan'ın New Delhi, Lucknow, Jabalpur, Indore şehirlerini kapsayan 1 hücre, Hindistan'ın Bengaluru ve Kochi şehirlerini kapsayan 1 hücre olmak üzere, toplam 2 farklı mekânsal indeksli hücreler olduğu görülmektedir. İkinci derece yoğunluğun ise, Birleşik Krallık'ın Londra, Cardiff, Glasgow şehirleri ile Hollanda'nın Amsterdam ve Fransa'nın Lille şehirlerini kapsayan 1 hücre, Kanada'nın Toronto ve Ottawa şehirlerini kapsayan 1 hücre, Amerika Birleşik Devletleri'nin Philadelphia ve Boston şehirlerini kapsayan 1 hücre, Dallas ve Oklahoma City'i kapsayan 1 hücre, Washington ve Cincinnati şehirlerini kapsayan 1 hücre, Hindistan'ın Haydarabad, Kolhapur ve Nagpur şehirlerini kapsayan 1 hücre ve Çin'in Pekin şehrini kapsayan 1 hücre olmak üzere, toplam 7 farklı mekânsal indeksli hücreler olduğu görülmektedir.

Çözünürlük 2 seviyesindeki H3 hiyerarşik altıgen hücrelerinden oluşan Ek-3.27'de yer alan koroplet harita incelendiğinde, Covid kelimesi hakkında oluşturulan tweetlerin en yoğun bölgenin Hindistan'ın Bengaluru ve Mysure şehirlerini kapsayan tek hücre olduğu görülmektedir. İkinci derece yoğunlukların ise, Hindistan'ın Jabalpur

şehrini kapsayan 1 hücre, New Delhi ve Jalpur şehirlerini kapsayan 1 hücre, Pune, Mumbai ve Ahmadnagar şehirlerini kapsayan 1 hücre, Çin'in Pekin şehrini kapsayan 1 hücre, ABD'nin Philadelphia ve New York şehirlerini kapsayan 1 hücre, Kanada'nın Toronto şehrini kapsayan 1 hücre, ABD'nin Garden City şehrini kapsayan 1 hücre olmak üzere, toplam 7 farklı mekânsal indeksli hücreler olduğu görülmektedir.

Covid kelimesi hakkındaki tweetlerin üç farklı çözünürlük seviyesi için hazırlanan Ek-3.25 – Ek-3.27'de yer alan koroplet haritalar ve Ek-2.9'da yer alan nokta işaretli harita karşılaştırıldığında, tweetlerin yüksek yoğunlukla Hindistan, Amerika Birleşik Devletleri ve Çin merkezli oldukları görülmektedir.

Vaccine kelimesi hakkındaki tweetlerin H3 hiyerarşik altıgen hücreleriyle çözünürlük 0 seviyesinde oluşturulan Ek-3.28'de yer alan koroplet harita incelendiğinde, en yüksek tweet oluşumunun Hindistan'ın güneyini ve Sri Lanka'yı kapsayan 1 hücre ve Hindistan'ın kuzeydoğusunu, Nepal ve Bangladeş'i kapsayan 1 hücre olmak üzere, toplam 2 farklı mekânsal indeksli hücrelerde olduğu görülmektedir. İkinci derece yoğunluğun ise Amerika Birleşik Devletleri'nin kuzeydoğu ve Kanada'nın güney bölgesini kapsayan 2 hücre ve Birleşik Krallık'ı kapsayan 1 hücre olmak üzere, toplam 3 farklı mekânsal indeksli hücreler olduğu görülmektedir.

Çözünürlük 1 seviyesindeki H3 hiyerarşik altıgen hücrelerinden oluşan Ek-3.29'da yer alan koroplet harita incelendiğinde, Vaccine kelimesi hakkında oluşturulan tweetlerin en yoğun bölgelerin Hindistan'ın New Delhi, Jabalpur, Indore şehirlerini kapsayan 1 hücre, Hindistan'ın Bengaluru ve Kochi şehirlerini kapsayan 1 hücre olmak üzere, toplam 2 farklı mekânsal indeksli hücreler olduğu görülmektedir. İkinci derece yoğunluğun ise, Birleşik Krallık'ın Londra, Glasgow şehirleri ile Hollanda'nın Amsterdam ve Fransa'nın Lille şehirlerini kapsayan 1 hücre, Hindistan'ın Haydarabad ve Nagpur şehirlerini kapsayan 1 hücre, Kanada'nın Toronto ve Ottawa şehirlerini kapsayan 1 hücre, Amerika Birleşik Devletleri'nin Philadelphia ve Boston şehirlerini kapsayan 1 hücre, Dallas ve Oklahoma City'yi kapsayan 1 hücre, Washington ve Cincinnati şehirlerini kapsayan 1 hücre olmak üzere, toplam 6 farklı mekânsal indeksli hücreler olduğu görülmektedir.

Çözünürlük 2 seviyesindeki H3 hiyerarşik altıgen hücrelerinden oluşan Ek-3.30'da yer alan koroplet harita incelendiğinde, Covid kelimesi hakkında oluşturulan tweetlerin en yoğun bölgelerin Hindistan'ın Bengaluru ve Mysure şehirlerini kapsayan 1 hücre, Hindistan'ın Jabalpur ve Sagar şehirlerini kapsayan 1 hücre, New Delhi ve

Jalpur şehirlerini kapsayan 1 hücre olmak üzere, toplam 3 farklı mekânsal indeksli hücreler olduğu görülmektedir. İkinci derece yoğunlukların ise, Hindistan'ın Pune ve Mumbai şehirlerini kapsayan 1 hücre, ABD'nin Philadelphia ve New York şehirlerini kapsayan 1 hücre, ABD'nin Garden City şehrini kapsayan 1 hücre, Kanada'nın Toronto şehrini kapsayan 1 hücre, Çin'in Pekin şehrini kapsayan 1 hücre, Birleşik Krallık'ın Londra, Fransa'nın Lille ve Belçika'nın Brugge şehirlerinin kapsayan 1 hücre olmak üzere toplam 6 farklı mekânsal indeksli hücreler olduğu görülmektedir.

Vaccine kelimesi hakkındaki tweetlerin üç farklı çözünürlük seviyesi için hazırlanan Ek-3.28 – Ek-3.30'da yer alan koroplet haritalar ve Ek-2.10'da yer alan nokta işaretli harita karşılaştırıldığında, tweetlerin yüksek yoğunlukla Hindistan, Amerika Birleşik Devletleri, İngiltere ve Çin merkezli oldukları görülmektedir.

Noktasal dağılımdaki mekânsal verileri gruplama ya da birleştirme, veri setlerini analiz etmede kullanılan en iyi yöntemlerdendir. Uber H3 yöntemiyle Covid-19 aşılıları hakkındaki tweetler ile, “covid” ve “vaccine” kelimesi hakkındaki noktasal dağılımdaki tweetlerin hiyerarşik altıgen hücreler ile mekânsal gruplandırması gerçekleştirilmiştir. Oluşturulan hiyerarşik altıgen hücreler kullanılarak sayım istatistikleri kolaylıkla hesaplanmış, elde edilen bilgiler tematik haritalar ile görselleştirilmiştir.

4.7. Dünya Geneline Uygulanmakta Olan COVID-19 Aşılarının İncelenmesi

Bu çalışma kapsamında kullanılan COVID-19 aşılıları hakkındaki tweet verilerinin coğrafi işaretlemeleri ve Uber H3 yöntemiyle mekânsal indekslemeleri sonucu elde edilen sonuçların karşılaştırılması amacıyla, ülkelerin COVID-19 pandemi döneminde salgınla mücadele kapsamında kullandıkları aşı türleri, aşı türüne bağlı aşılana birey sayıları ve aşı türüne bağlı uygulanan toplam aşı miktarlarına ilişkin araştırma gerçekleştirilmiştir. Araştırma sonucunda, ülkelerin kullanmış oldukları aşı türleri, aşılama tarihleri, toplam aşı miktarları, toplam aşılana birey sayılarına ilişkin veri setine erişim sağlanmıştır (URL-31). Elde edilen veri seti incelendiğinde, veri setinin “location”, “date”, “vaccine”, “source_url”, “total_vaccinations”, “people_vaccinated”, “people_fully_vaccinated” ve “total_booster” olmak üzere 8 farklı öznitelik alanından oluşmaktadır.

location: COVID-19 aşısının kullanıldığı ülkeyi,

date: COVID-19 aşılıları hakkındaki istatistiki verileri tarihini,

vaccine: Ülkelerin kullanmış oldukları COVID-19 aşılılarını,

source_url: COVID-19 aşıları hakkındaki istatistiki verilerin kaynağını,

total_vaccinations: Ülkede gerçekleştirilen toplam aşı miktarını,

people_vaccinated: Ülkede aşılanan birey sayısını,

people_fully_vaccinated: Kullanılan aşı türüne bağlı olarak tüm dozları aşılanan birey sayısını,

total_boosters: Ülkede takviye (ek) olarak gerçekleştirilen aşılama miktarını

ifade etmektedir.

Ülkelerin uygulamış oldukları aşıların tarihsel bazlı grup halinde verildiği görülmektedir. Veriler, ülkelerin uygulamış oldukları tüm aşılarla ait toplam istatistiki sonuçları göstermektedir. Bu verileri kullanarak her aşı türü için uygulanan aşı dozunu, aşılanan birey sayısını tespit etmek imkânsızdır. Bu veri setinden çıkarılabilecek en anlamlı bilgi hangi ülkenin hangi aşığı kullandığına ilişkin bilgidir.

Bu amaçla veri seti üzerinde ön işleme gerçekleştirilmiştir. Veri setindeki veriler farklı ülkelerin kaynaklarından edinildiği için öncelikle tarih formatındaki farklılıklar giderilerek tek yapıda filtreleme yapılabilen tarih verileri elde edilmiştir. Veriler, Covid-19 Tweet veri setinin kapsadığı 12/12/2020-10/10/2021 tarihleri arasını kapsayacak şekilde filtelenmiştir. Daha sonra, “vaccine” öznitelik alanında düzenlemeler yapılarak her ülkenin kullanmış oldukları aşı markaları ayrı ayrı tespit edilmiştir. COVID-19 tweet veri setinde yer alan 8 farklı aşı markası için veri setinde eşleştirmeler yapılarak, 8 farklı aşı markasının hangi ülkelerde kullanıldığının tespiti yapılmıştır. Elde edilen sonuçlar ile tematik haritalar oluşturulmuştur (Ek 4.1 – Ek-4.8).

Pfizer Biontech aşısının ülkelere göre kullanım durumunu gösteren Ek-4.1’de yer alan koroplet harita incelendiğinde; aşının Afrika Kıtasının bazı ülkeleri ve Asya Kıtasının Çin, Rusya, Kazakistan, Hindistan, Türki Cumhuriyetler, İran ve Kamboçya vd. ülkeler haricindeki ülkelerde kullanıldığı görülmektedir. Kullanıcıların attığı tweetleri gösteren Ek-3.1 – Ek-3.3’te yer alan koroplet haritalar ve Ek-2.1’de yer alan nokta işaretli harita incelendiğinde de Pfizer Biontech aşısı hakkında oluşturulan tweetlerin Amerika Birleşik Devletleri, Kanada, Avrupa kıtası merkezli oldukları tespit edilmiştir. Hindistan’da Pfizer Biontech aşısının kullanımını olmamasına rağmen, Pfizer Biontech aşısı hakkında atılan tweetlerin olduğu görülmektedir. Hindistan haricinde elde edilen haritaların birbirleriyle tutarlı olduğu görülmektedir.

Sinopharm aşısının ülkelere göre kullanım durumunu gösteren Ek-4.2’de yer alan koroplet harita incelendiğinde; aşının Asya Kıtası’nın güney kısmında yer alan ülkelerde (Hindistan harici), Afrika Kıtası ülkelerinin genelinde ve Güney Amerika Kıtası ülkelerinde (Kolombiya, Brezilya ve Uruguay harici) kullanıldığı görülmektedir. Kullanıcıların attığı tweetleri gösteren Ek-3.4 – Ek-3.6’da yer alan koroplet haritalar ve Ek-2.2’de yer alan nokta işaretli harita incelendiğinde, Sinopharm aşısı hakkında oluşturulan tweetlerin Çin, Sri Lanka, Pakistan, Bangladeş ve Birleşik Arap Emirlikleri merkezli oldukları tespit edilmiştir. Elde edilen bilgiler ve üretilen haritaların birbirleriyle tutarlı olduğu görülmektedir.

Sinovac aşısının ülkelere göre kullanım durumunu gösteren Ek-4.3’te yer alan koroplet harita incelendiğinde; aşının Asya Kıtası’nın Çin, Pakistan, Malezya, Singapur vd. ülkeleri, Afrika Kıtası’nın Libya, Mısır, Sudan vd. ülkeleri, Meksika, Kolombiya, Şili, Uruguay, Türkiye vd. ülkelerde kullanıldığı görülmektedir. Kullanıcıların attığı tweetleri gösteren Ek-3.7 – Ek-3.9’da yer alan koroplet haritalar ve Ek-2.3’te yer alan nokta işaretli harita incelendiğinde, Sinovac aşısı hakkında oluşturulan tweetlerin Çin, Filipinler, Malezya, Tayland merkezli oldukları tespit edilmiştir. Elde edilen bilgiler ve üretilen haritaların birbirleriyle tutarlı olduğu görülmektedir.

Moderna aşısının ülkelere göre kullanım durumunu gösteren Ek-4.4’te yer alan koroplet harita incelendiğinde; aşının Kuzey Amerika Kıtası’nda, Güney Amerika Kıtası’nda Arjantin, Paraguay ve Kolombiya’da, Avrupa Kıtasının tamamına yakınında, Afrika Kıtası’nın bazı ülkelerinde, Asya Kıtası’nın Tacikistan, Özbekistan, Güney Kore, Bangladeş, Endonezya, Tayland, Filipinler ve Avustralya’da kullanıldığı görülmektedir. Kullanıcıların attığı tweetleri gösteren Ek-3.10 – Ek-3.12’de yer alan koroplet haritalar ve Ek-2.4’te yer alan nokta işaretli harita incelendiğinde, Moderna aşısı hakkında oluşturulan tweetlerin Amerika Birleşik Devletleri, Kanada ve Birleşik Krallık merkezli oldukları tespit edilmiştir. Hindistan’da Moderna aşısının kullanımı olmamasına rağmen, Pfizer Biontech aşısı hakkında atılan tweetlerin olduğu görülmektedir. Hindistan haricinde elde edilen haritaların birbirleriyle tutarlı olduğu görülmektedir.

Oxford AstraZeneca aşısının ülkelere göre kullanım durumunu gösteren Ek-4.5’te yer alan koroplet harita incelendiğinde; aşının Amerika Birleşik Devletleri, Grönland, Venezuela, Türkiye, Kazakistan, Rusya, Kırgızistan, Çin ve bazı Afrika Kıtası ülkeler hariç olmak üzere nerdeyse tüm ülkelerde kullanıldığı görülmektedir. Kullanıcıların attığı tweetleri gösteren Ek-3.13 – Ek-3.15’te yer alan koroplet haritalar

ve Ek-2.5'te yer alan nokta işaretli harita incelendiğinde, Oxford AstraZeneca aşısı hakkında oluşturulan tweetlerin Birleşik Krallık, Kanada ve Hindistan merkezli oldukları tespit edilmiştir. Elde edilen bilgiler ve üretilen haritaların birbirleriyle tutarlı olduğu görülmektedir.

Covaxin aşısının ülkelere göre kullanım durumunu gösteren Ek-4.6'da yer alan koroplet harita incelendiğinde; aşının Hindistan, İran, Botsvana, Orta Afrika Cumhuriyeti, Komorlar, Morityus ve Paraguay ülkelerinde kullanıldığı görülmektedir. Kullanıcıların attığı tweetleri gösteren Ek-3.16 – Ek-3.18'de yer alan koroplet haritalar ve Ek-2.6'da yer alan nokta işaretli harita karşılaştırıldığında, Covaxin aşısı hakkında oluşturulan tweetlerin yüksek yoğunlukla Hindistan merkezli olduğu tespit edilmiştir. Amerika Birleşik Devletleri ve Birleşik Krallık'ta Covaxin aşısının kullanımı olmamasına rağmen, Covaxin aşısı hakkında atılan tweetlerin olduğu görülmektedir. Amerika Birleşik Devletleri ve Birleşik Krallık haricinde elde edilen haritaların birbirleriyle tutarlı olduğu görülmektedir.

Sputnik V aşısının ülkelere göre kullanım durumunu gösteren Ek-4.7'de yer alan koroplet harita incelendiğinde; aşının Asya Kıtası'nın Rusya, Türki Cumhuriyetler, İran, Irak, Birleşik Arap Emirlikleri, Umman, Hindistan, Vietnam, Filipinler, Sri Lanka vd. ülkelerinde, Afrika Kıtası'nın Fas, Tunus, Libya, Mısır vd. ülkelerinde, Avrupa Kıtası'nda Bosna Hersek, Macaristan, Sırbistan, Moldova vd. ülkelerinde, Meksika, Venezuela, Arjantin, Paraguay vd. ülkelerde kullanıldığı görülmektedir. Kullanıcıların attığı tweetleri gösteren Ek-3.19 – Ek-3.21'de yer alan koroplet haritalar ve Ek-2.7'de yer alan nokta işaretli harita incelendiğinde, Sputnik V aşısı hakkında oluşturulan tweetlerin yüksek yoğunlukla Hindistan, daha sonra ise Rusya merkezli oldukları tespit edilmiştir. Elde edilen bilgiler ve üretilen haritaların birbirleriyle tutarlı olduğu görülmektedir.

Johnson&Johnson aşısının ülkelere göre kullanım durumunu gösteren Ek-4.8'de yer alan koroplet harita incelendiğinde; aşının Kuzey Amerika Kıtası ülkelerinde, Güney Amerika Kıtası'nın Brezilya, Kolombiya vd. ülkelerinde, Avrupa Kıtası ülkelerinde, Afrika Kıtası'nın doğu bölgesindeki ülkelerde, Asya Kıtası'nın Afganistan, Nepal, Kamboçya, Güney Kore, Endonezya, Filipinler vd. ülkelerinde kullanıldığı görülmektedir. Kullanıcıların attığı tweetleri gösteren Ek-3.22 – Ek-3.24'te yer alan koroplet haritaları ve Ek-2.8'de yer alan nokta işaretli harita incelendiğinde, Johnson&Johnson aşısı hakkında oluşturulan tweetlerin yüksek yoğunlukla Amerika

Birleşik Devletleri, daha sonra ise İngiltere merkezli oldukları tespit edilmiştir. Elde edilen bilgiler ve üretilen haritaların birbirleriyle tutarlı olduğu görülmektedir.

4.8. Kümeleme Analizi

Bu çalışmada, COVID-19 aşılıları hakkındaki veri setinden Bölüm 4.4'te keşfedilen yeni bilgiler kullanılarak, kümeleme analizi yöntemiyle ülkelerin birbirlerine olan benzerlikleri ve farklılıklarının tespit edilmesi amaçlanmıştır. Bu kapsamda, “Ülkelerin X kullanıcı sayıları ve oluşturulan tweet sayılarına göre kümelenmesi” ve “Ülkelerin oluşturulan toplam tweet sayısı, covid kelimesi içeren tweet sayısı ve vaccine kelimesi içeren tweet sayılarına göre kümelenmesi” olmak üzere iki farklı kümeleme analizi işlemi hedeflenmiştir.

Kümeleme analizi yöntemleri, aynı veri seti içerisinde yer alan verileri birbirleriyle ilişki durumunu dikkate almadan kümelemektedir. Ancak bu durum kullanıcının kümeleme yönteminden beklediği sonuçlar için olumsuz sonuçlar çıkarabilmektedir (Selvi ve Çağlar, 2017). Bu nedenle, kümeleme analizinden önce verilerini birbirleriyle ilişkisi kontrol edilerek, verilerin ilişki düzeyi belirlenmelidir. Pearson korelasyon testi, veriler arasında bir ilişki olup olmadığının belirlenmesinde kullanılmaktadır. Pearson korelasyon katsayısı (4.1) eşitliğiyle hesaplanmaktadır.

$$r = \frac{n \sum XY - \sum X \sum Y}{\sqrt{[n \sum X^2 - (\sum X)^2][n \sum Y^2 - (\sum Y)^2]}} \quad (4.1)$$

Pearson korelasyon (r) katsayısı, [-1,1] aralığında değerler almaktadır. Korelasyon katsayısının (1) değerine yaklaşması korelasyon testine giren değişkenler arasında yüksek düzeyde olumlu ilişkinin olduğu, (-1) değerine yaklaşması ise değişkenler arasında yüksek düzeyde olumsuz bir ilişkinin bulunduğunu göstermektedir.

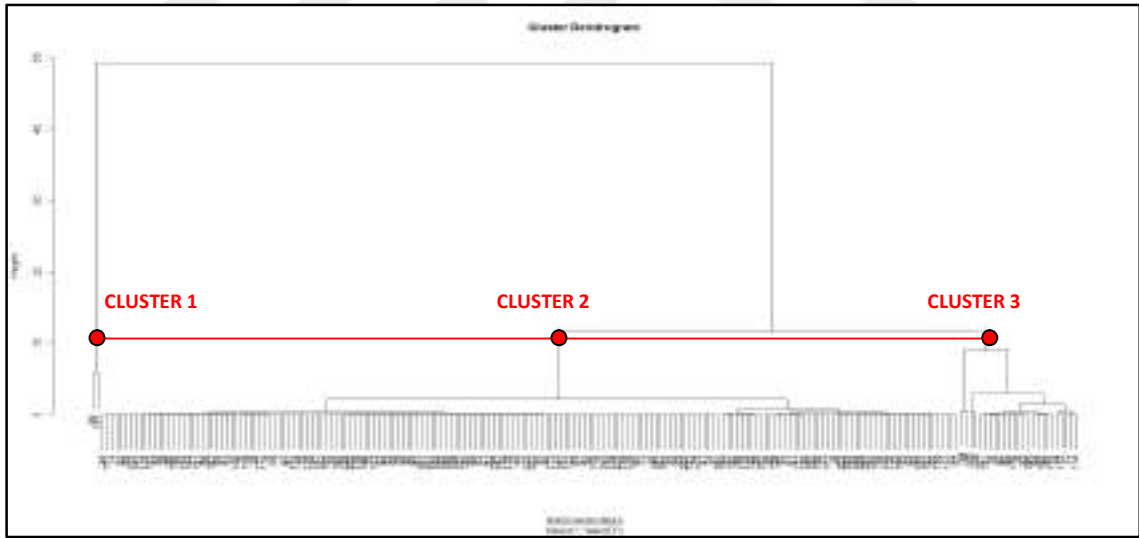
“Ülkelerin X kullanıcı sayıları ve tweet sayılarına göre kümelenmesi” ve “Ülkelerin tweet sayısı, covid kelimesi içeren tweet sayısı ve vaccine kelimesi içeren tweet sayılarına göre kümelenmesi” olmak üzere iki farklı kümeleme analiz çalışmasında, Ward yöntemiyle hiyerarşik kümeleme ve K-Ortalamalar (K-Means) hiyerarşik olmayan kümeleme yöntemleri ayrı ayrı kullanılmıştır.

4.8.1. Ülkelerin X kullanıcı sayıları ve tweet sayılarına göre ward yöntemiyle hiyerarşik kümelmesi

Ward yöntemiyle hiyerarşik kümeleme analizi işlemleri RStudio platformunda R dili kullanılarak gerçekleştirilmiştir. RStudio platformunda Ward yöntemiyle hiyerarşik kümeleme işleminde Ek-5.1’te yer alan kodlamalar kullanılmıştır.

Kümeleme analizi işleminde kullanılacak değişkenler arasındaki ilişki düzeyinin belirlenmesi amacıyla Pearson korelasyon testi gerçekleştirilmiştir. “tweetcount” ve “usercount” değişkenleri arasındaki Pearson korelasyon katsayısının 0,9395855 çıkması, iki değişken arasında olumlu anlamda yüksek seviyede ilişki olduğunu göstermektedir.

“tweetcount” ve “usercount” değişkenleri kullanılarak, Ward yöntemiyle hiyerarşik kümeleme işleminde küme sayısının ve elemanlarının belirlenmesi amacıyla dendrogram çizelgesi oluşturulmuştur (Şekil 4.10). Dendrogram incelendiğinde, kırmızı renkteki referans çizgi ile ağaç yapısının kesiştikleri düğüm noktaları birer kümeyi ifade etmektedir.



Şekil 4.10 “tweetcount” ve “usercount” değişkenlerinin dendrogram üzerinde gösterimi

Birden fazla değişken kullanılarak gerçekleştirilen hiyerarşik kümeleme analizinde oluşan kümeler yorumlanırken dikkat edilmelidir. Hiyerarşik kümeleme yönteminin uygunluğunun belirlenmesinde Kofenetik korelasyon katsayısından yararlanılmaktadır. Ward yöntemiyle hiyerarşik kümeleme işleminde kullanılan veriler arası kofenetik korelasyon katsayısı 0,844 olarak hesaplanmıştır. Romesburg (1984),

kofenetik değerin 0,80 veya daha büyük değerin olmasının değişkenlerin birbirleriyle yüksek ilişkili ve kümelemenin kabul edilebilir olduğunu gösterdiğini belirtmiştir.

Ülkelerin X kullanıcı sayıları ve tweet sayılarına göre Ward yöntemiyle hiyerarşik kümeleme işlemine ait ülke bazlı kümeleme sonuçları Ek-5.2’de verilmiştir. Ward yöntemiyle oluşturulan kümelere ait istatistikî bilgiler Çizelge 4.13’te gösterilmektedir.

Çizelge 4.13 Ülkelerin X kullanıcı sayıları ve tweet sayılarına göre Ward yöntemiyle oluşturulan kümelere ait istatistikî bilgiler

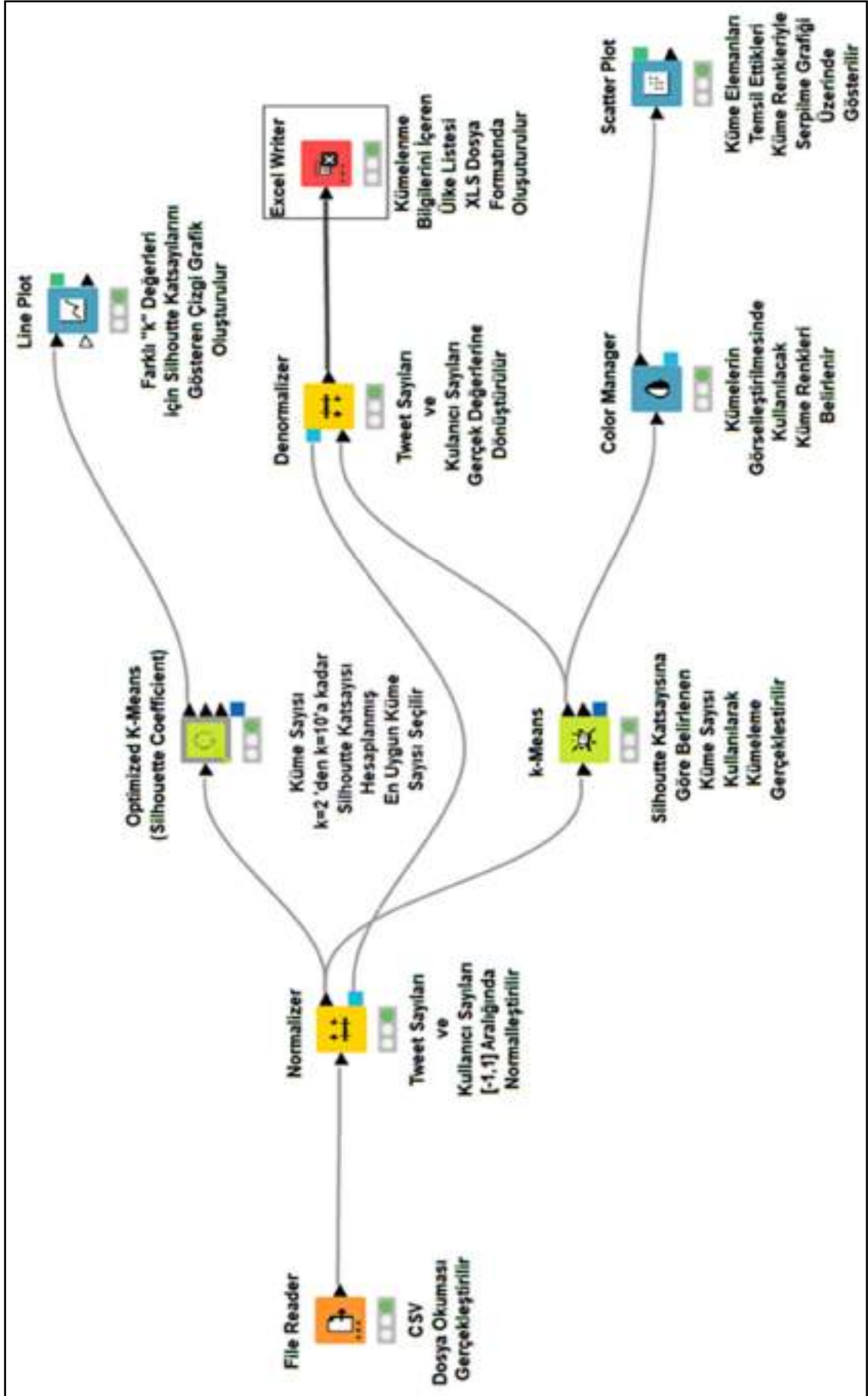
Küme No	Küme Üyesi Ülke Sayısı	Tweet Sayıları	Kullanıcı Sayıları
1	2	56080 - 30922	18326 - 16153
2	25	9543 - 469	4661 - 212
3	180	554 - 1	206 - 1

4.8.2. Ülkelerin X kullanıcı sayıları ve tweet sayılarına göre k-ortalama (k-means) yöntemiyle hiyerarşik olmayan kümeleme

K-ortalama yöntemiyle hiyerarşik olmayan kümeleme analizi işlemleri KNIME platformunda gerçekleştirilmiştir. Knime platformunda, veri yönetim araçları, analiz araçları, istatistik ve görselleştirme araçları “Node” adı verilen yapılarda bulunmaktadır. Knime platformunda gerçekleştirilen kümeleme analizi tamamen görsel arayüz üstünde “Node” adı verilen araçlar kullanılarak gerçekleştirilmiştir (Şekil 4.11).

Verilerin Knime platformuna Şekil 4.11’te yer alan “File Reader” aracı kullanılarak entegre edilmesi sağlanmıştır. Kümeleme analizinde kullanılacak veriler arasında sayısal değerlerde çok büyük fark olması, birim uyumsuzluğu vb. bazı durumlarda veri normalizasyon işlemi gerekmektedir. Veri normalizasyon işlemi için “Normalizer” aracı kullanılmıştır.

K-ortalama yöntemine göre gerçekleştirilecek kümeleme işleminde küme sayısının belirlenmesi kullanıcının deneyimiyle belirlenmektedir. Kullanıcının küme sayısını belirlemesi yardımcı olan yöntemler bulunmaktadır. Silhouette katsayısı kullanıcının kümeleme işleminde kullanacağı k katsayısını belirlemeye yardımcı olmaktadır. Knime platformunda, Silhouette katsayısının belirlemede “Optimized K-Means (Silhouette Coefficient)” aracı kullanılmıştır. Optimized K-Means aracı kullanılarak $k=2$ değerinden $k=9$ değerine kadar küme değerleri Silhouette katsayı değerleri hesaplanmıştır.



Şekil 4.11 Knime platformunda k-ortalama kümeleme analizi

Optimized K-Means (Silhouette Coefficient) aracından elde edilen Silhouette katsayılarına ait çizgisel grafiğin oluşturulması “Line Plot” aracı kullanılmıştır. “Line Plot” aracı ile üretilen Silhouette katsayılarına ait çizgisel grafik Şekil 4.12’de gösterilmektedir.



Şekil 4.12 "Line Plot" aracıyla oluşturulan Silhouette katsayı grafiği

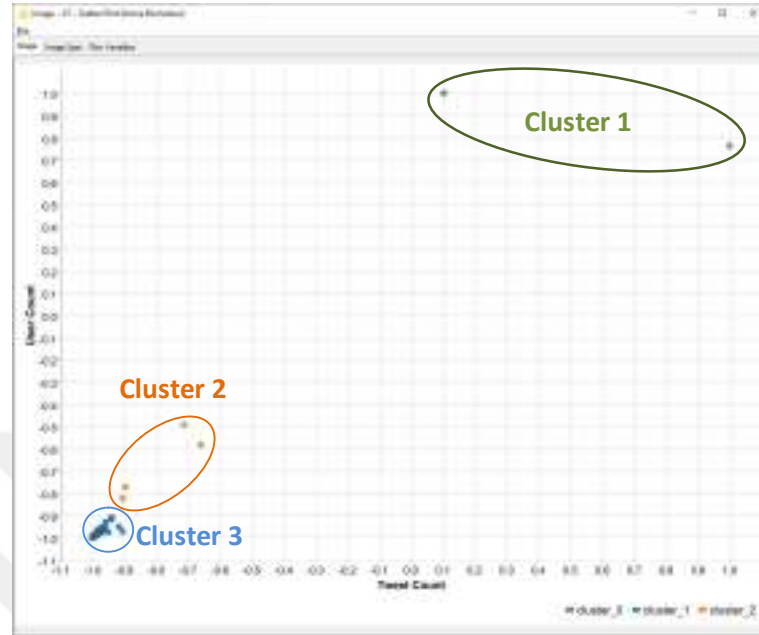
Şekil 4.12’de görüleceği üzere K-ortalamalar yöntemiyle yapılacak kümeleme analizinde en uygun k değerinin 3 olduğu görülmektedir. Kümeleme analizi işlemi için, Silhouette yöntemiyle k katsayısının belirlenmesinin ardından “k-Means” aracı ile kümeleme analizi işlemi gerçekleştirilmiştir.

“k-Means” aracı ile kümeleme analizi işleminin gerçekleştirilmesinden sonra “Color Manager” aracı kullanılarak k-ortalama yöntemiyle oluşturulan kümelerin görselleştirmede temsil edileceği renkler belirlenmiştir.

“Color Manager” aracı ile kümeleme analizi işlemi sonucu oluşturulan kümelerin görselleştirmede kullanılacak renklerinin belirlenmesinin ardından “Scatter Plot” aracı kullanılarak kümeleme analizi sonucu oluşan kümelerin ve elemanlarının serpilme grafiği (Şekil 4.13) oluşturulmuştur.

Kümeleme analizi işlemi sonucu oluşturulan kümelerin oluşturulmasında normalizasyonu gerçekleştirilen öznitelik değerleri kullanılmıştır. Bu öznitelik değerlerinin gerçek değerlerine dönüştürülmesinde “Denormalizer” aracı kullanılmıştır.

Knime platformunda k-ortalamlar yöntemiyle kümeleme analizi işlemi sonucu oluşturulan ülkeler bazındaki kümeleme bilgileri “Excel Writer” aracı kullanılarak Excel dosya formatında platform dışına çıkarılmıştır.



Şekil 4.13 Serpilme grafiği üzerinde kümelerin ve küme elemanlarının gösterimi

Ülkelerin X kullanıcı sayıları ve tweet sayılarına göre k-ortalamlar (k-Means) yöntemiyle hiyerarşik kümelendirme işlemine ait ülke bazlı kümelendirme sonuçları Ek-5.2’de verilmiştir. k-ortalamlar yöntemiyle oluşturulan kümelere ait istatistikî bilgiler Çizelge 4.14’te gösterilmektedir.

Çizelge 4.14 Ülkelerin X kullanıcı sayıları ve tweet sayılarına göre k-ortalamlar yöntemiyle oluşturulan kümelere ait istatistikî bilgiler

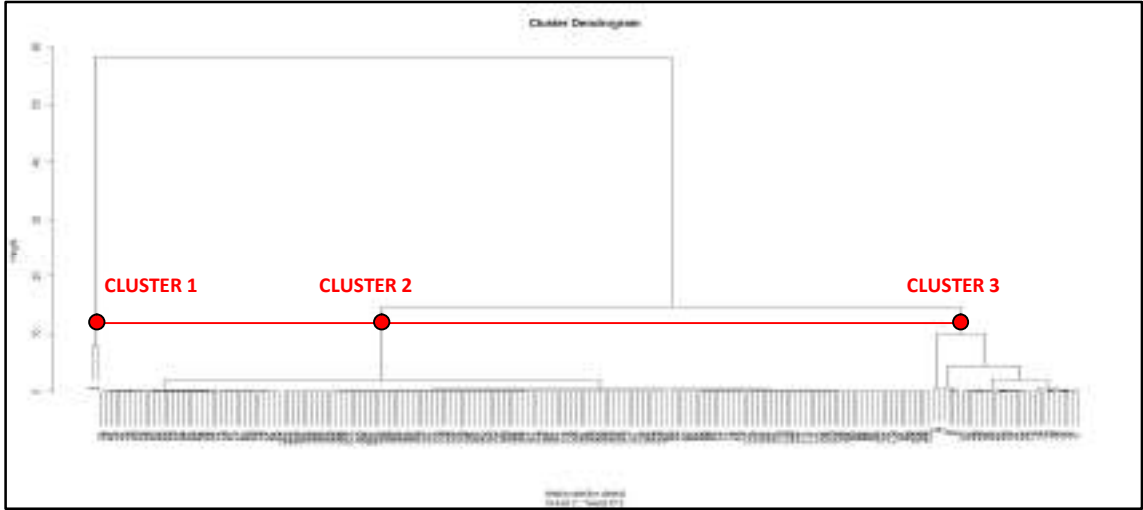
Küme No	Küme Üyesi Ülke Sayısı	Tweet Sayıları	Kullanıcı Sayıları
1	2	56080 - 30922	18326 - 16153
2	4	9543 - 2677	4661 - 1627
3	180	2668 - 1	845-1

4.8.3. Ülkelerin oluşturulan toplam tweet sayısı, covid kelimesi içeren tweet sayısı ve vaccine kelimesi içeren tweet sayıları göre ward yöntemiyle kümelendirme

Bu bölümde gerçekleştirilen işlemler Bölüm 4.8.1’de gerçekleştirilen işlemlerle aynı nitelikteki işlemler olduğu için detaylı açıklama yapılmamıştır. Kümeleme işlemi kullanılan öznitelik değişkenleri ve öznitelik değişken sayısındaki farklılıklar ilgili işlem adımlarında açıklanmıştır. Kümeleme analizinin gerçekleştirilmesinde Ek-5.3’te verilen r kod bloğu kullanılmıştır.

Kümeleme analizi işlemi için kullanılacak üç farklı değişken arasındaki ilişki düzeyinin belirlenmesi amacıyla 3 farklı Pearson korelasyon katsayısı hesaplanmıştır. Değişken çiftleri arasındaki Pearson korelasyon katsayılarının 0.9976956, 0.996624 ve 0.9920709 çıkması, bu değişken çiftlerinin yüksek seviyede ilişkili olduğunu göstermektedir.

“tweetcount”, “covidcount” ve “vaccinecount” değişkenleri kullanılarak, Ward yöntemiyle hiyerarşik kümeleme işlemi küme sayısının ve elemanlarının belirlenmesi amacıyla dendrogram çizelgesi oluşturulmuştur (Şekil 4.14). Dendrogram incelendiğinde, kırmızı renkteki referans çizgi ile ağaç yapısının kesiştiği düğüm noktaları birer kümeyi ifade etmektedir.



Şekil 4.14 “tweetcount”, “covidcount” ve “vaccinecount” değişkenlerinin dendrogram üzerindeki gösterimi

Ward yöntemiyle hiyerarşik kümeleme işlemi için kullanılan veriler arasındaki kofenetik korelasyon katsayısı 0,795 olarak hesaplanmıştır. Bulunan kofenetik korelasyon katsayısı değeri Romesburg (1984) tarafından açıklanan 0,80 değerine yaklaşık eşit kabul edilerek kümelemenin kabul edilebilir olduğu değerlendirilmiştir.

Ülkelerin oluşturulan toplam tweet sayısı, covid kelimesi içeren tweet sayısı ve vaccine kelimesi içeren tweet sayıları kullanıcı sayılarına göre ward yöntemiyle hiyerarşik kümeleme işlemine ait ülke bazlı kümeleme sonuçları Ek-5.4’te verilmiştir. Ward yöntemiyle oluşturulan kümelere ait istatistikî bilgiler Çizelge 4.15’te gösterilmektedir.

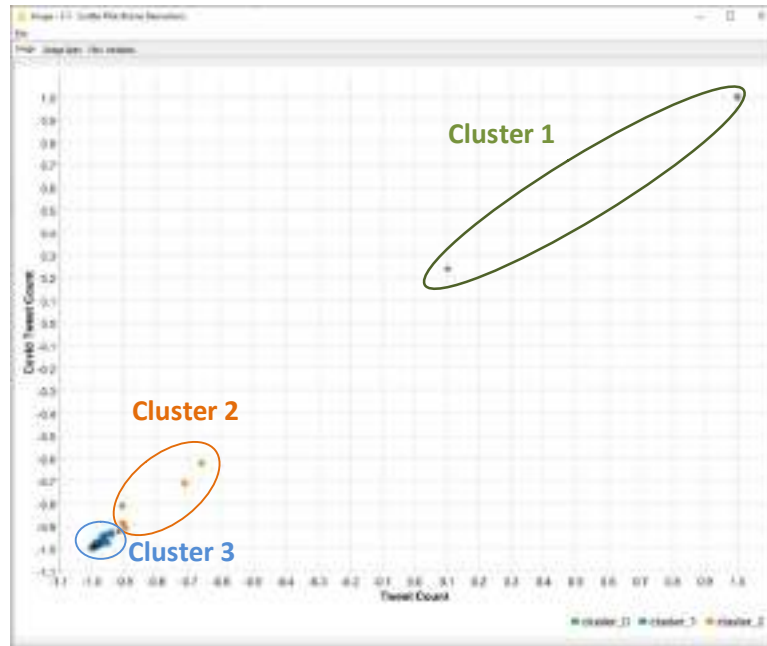
Çizelge 4.15 Ülkelerin tweet sayıları, covid kelimesi içeren tweet sayıları ve vaccine kelimesi içeren tweet sayılarına göre Ward yöntemiyle oluşturulan kümelere ait istatistikî bilgiler

Küme No	Küme Üyesi Ülke Sayısı	Tweet Sayıları	Covid İçeren Tweet Sayısı	Vaccine İçeren Tweet Sayısı
1	2	56080 - 30922	15443 - 9563	22913 - 16383
2	31	9543 - 361	2911 - 100	4358 - 193
3	180	284 - 1	102 - 0	143 - 0

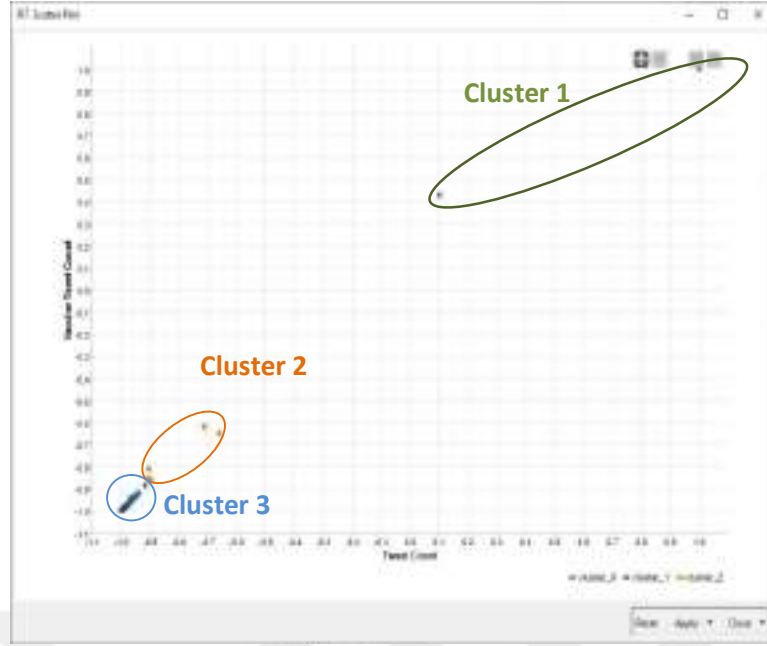
4.8.4. Ülkeleri oluşturulan toplam tweet sayısı, covid kelimesi içeren tweet sayısı ve vaccine kelimesi içeren tweet sayıları göre k-ortalama (k-means) yöntemiyle hiyerarşik olmayan kümeleme

Bu bölümde de gerçekleştirilen işlemler Bölüm 4.8.2'de gerçekleştirilen işlemlerle aynı nitelikteki işlemler olduğu için detaylı açıklama yapılmamıştır. KNIME platformunda Şekil 4.11'de yer alan iş akış diyagramı gerçekleştirilerek k-ortalamlar yöntemiyle kümeleme işlemi gerçekleştirilmiştir.

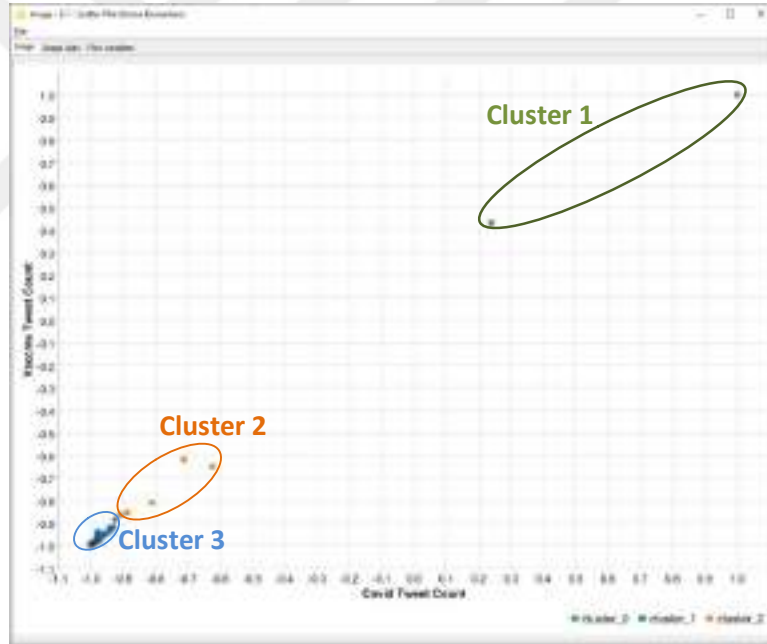
K-ortalamlar yöntemiyle gerçekleştirilecek kümeleme işleminde küme sayısının belirlenmesi Silhouette katsayısı kullanılmıştır. $k=2$ değerinden $k=9$ değerine kadar küme değerleri için Silhouette katsayı değerleri hesaplanmıştır. Kümeleme analizi işlemi sonucu elde edilen kümeleme sonuçları Şekil 4.15, Şekil 4.16 ve Şekil 4.17'de yer alan serpilme grafikleri üzerinde ayrı ayrı gösterilmiştir.



Şekil 4.15 Toplam tweet sayısı ve Covid kelimesi içeren tweet sayısına göre kümeleri gösteren serpilme grafiği



Şekil 4.16 Toplam tweet sayısı ve vaccine kelimesi içeren tweet sayısına göre kümeleri gösteren serpilme grafiği



Şekil 4.17 Covid kelimesi içeren tweet sayısı ve vaccine kelimesi içeren tweet sayısına göre kümeleri gösteren serpilme grafiği

Ülkelerin oluşturulan toplam tweet sayısı, Covid kelimesi içeren tweet sayısı ve Vaccine kelimesi içeren tweet sayıları kullanıcı sayılarına göre k-Ortalamlar (k-Means) yöntemiyle hiyerarşik kümelendirme işlemine ait ülke bazlı kümelendirme sonuçları Ek-5.4'te verilmiştir. k-Ortalamlar (k-Means) yöntemiyle oluşturulan kümelere ait istatistikî bilgiler Çizelge 4.16'da gösterilmektedir.

Çizelge 4.16 Ülkelerin tweet sayıları, covid kelimesi içeren tweet sayıları ve vaccine kelimesi içeren tweet sayılarına göre k-ortalamlar yöntemiyle oluşturulan kümelere ait istatistikî bilgiler

Küme No	Küme Üyesi Ülke Sayısı	Tweet Sayıları	Covid İçeren Tweet Sayısı	Vaccine İçeren Tweet Sayısı
1	2	56080 - 30922	15443 - 9563	22913 - 16383
2	5	9543 - 2668	2911 - 706	4358 - 1553
3	200	2333 - 1	585 - 0	1327 - 0

4.9. Web tabanlı interaktif uygulama geliştirilmesi

ArcGIS Pro, Netcad ve QGIS gibi coğrafi bilgi sistemi yazılımlarında hazırlanan verilerin ve haritaların Web üzerinden paylaşılması ve sunulması için Open Geospatial Consortium (OGC) bazı teknik standartlar tanımlamıştır.

Web Map Service (WMS), masaüstü yazılım platformlarında oluşturulan haritaların Web üzerinde JPG, TIF ve PNG gibi raster format yapılarında paylaşmak ve harita tabanlı uygulama oluşturmak amacıyla kullanılan servistir. WMS yapısıyla yayınlanan veri raster formatta sunulduğu için, WMS servisi kullanılarak oluşturulan Web tabanlı haritaların işlevselliği harita yakınlaştırma, uzaklaştırma kaydırma gibi sınırlı özelliklere sahiptir. WMS servisi türünden veri akışı tek taraflıdır. İstemci ilgilendiği veri talebini sunucuya iletir, sunucuda istemciden gelen veri talebini tek taraflı olarak istemciye aktarmaktadır (URL-27).

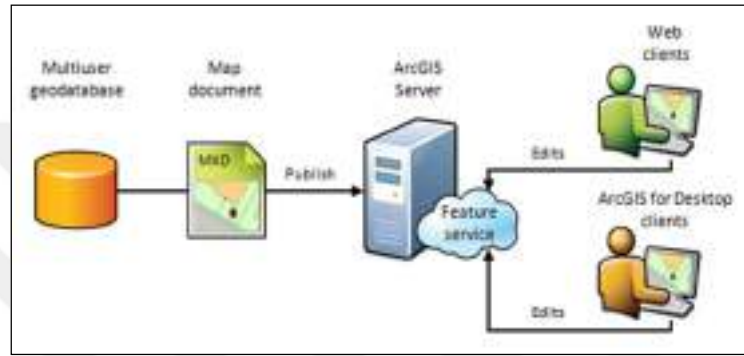
Web Feature Service (WFS), masaüstü yazılım platformlarında harita oluşturma sürecinde kullanılan nokta, çizgi ve poligon yapısındaki vektör verileri Web üzerinde paylaşmak ve harita tabanlı uygulama oluşturmak amacıyla kullanılan servistir. Web tabanlı harita uygulamalarında harita verilerinin düzenlenmesi, sorgulanması, güncellenmesi ya da silme işlemlerinin gerçekleştirilmesi isteniyorsa mutlaka WFS servisi türü kullanılmalıdır. WFS servisi türünde veri akışı çift taraflı gerçekleşmektedir. Yani veri akışı sadece sunucudan istemciye doğru değildir. İstemci Web üzerinde yaptığı veri değişikliklerini ya da oluşturduğu yeni verileri sunucuya ileterek veri kaynağı üzerinde değişiklikler sağlayabilmektedir (URL-28).

Web Coverage Service (WCS), masaüstü yazılım platformlarında oluşturulan raster veri setlerinin Web üzerinde paylaşılmasında kullanılan servistir. WCS servisinin raster veri formatı paylaşması yönüyle WMS servisine benzemesinin yanında, WMS servisinden ayıran en büyük fark kullanıcılara istemci üzerinde WCS servisiyle

yayınlanan verilerle eğitim, bakı, kabartma vb. raster tabanlı analiz işlemlerine imkân sunmasıdır (URL-29).

4.9.1. Coğrafi verilerin WFS servisi kullanılarak bulut sunucu üzerinde paylaşılması

Web üzerinden coğrafi verilerin sorgulanabildiği, yeniden düzenlenebildiği, yeni verilerin oluşturulabildiği ve gerektiği durumlarda silinebildiği işlev yeteneklerine sahip harita tabanlı etkileşimli uygulamalar geliştirmek için coğrafi verilerin detay servisi olarak sunucu üzerine yayınlanması gerekmektedir (Şekil 4.18).



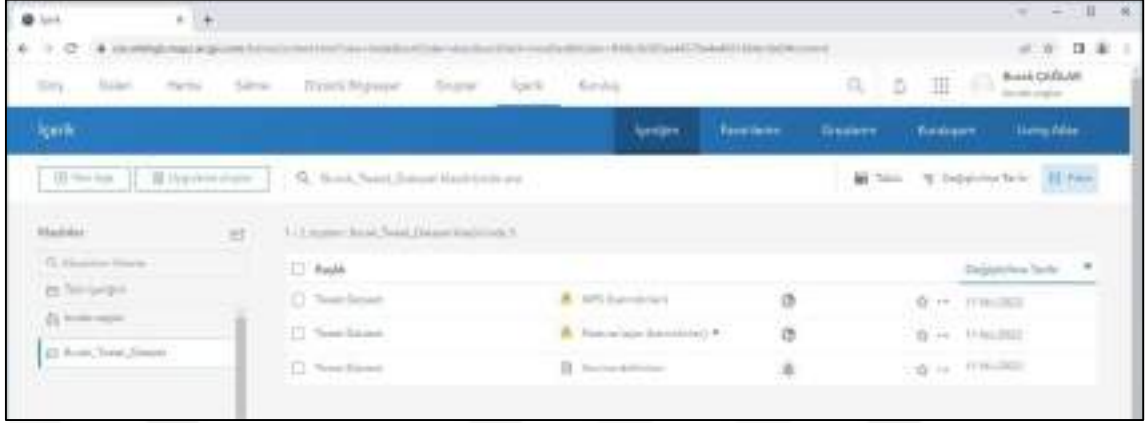
Şekil 4.18 Feature (Detay) servisi oluşturma süreci (URL-32)

ArcGIS Pro masaüstü yazılımında, harita üretiminde kullanılan tweet noktaları, ülke idari sınırları ve H3 altıgen hücreleri ArcGIS Online üzerinde Web tabanlı harita uygulaması oluşturma amacıyla Şekil 4.19 görüldüğü gibi ArcGIS sunucuları üzerinde yayınlanmıştır.



Şekil 4.19 Coğrafi verilerin ArcGIS Pro üzerinden detay servisi olarak yayınlanması

ArcGIS Pro yazılımından ArcGIS sunucularına yüklenen tweet noktaları, ülke idari sınırları ve H3 altıgen hücrelerine ait coğrafi veri katmanları ve WFS servislerine ArcGIS Online üzerinden Şekil 4.20’de görüldüğü gibi erişim sağlanmıştır.



Şekil 4.20 ArcGIS Online üzerinden coğrafi verilere erişimin sağlanması

ArcGIS Online üzerinden, ArcGIS sunucuları üzerinde yayınlanan coğrafi verilere ilişkin yönetimsel hizmetler gerçekleştirilmektedir. Şekil 4.21’de görülen Genel Bakış sekmesi altından coğrafi veri setinin barındırdığı katmanlar görüntülenebilmekte, katmanlar Map Viewer sekmesiyle çevrimiçi harita oluşturma ekranında açılmaktadır. ArcGIS sunucuları üzerinde bulunan veri seti ArcGIS Desktop ya da ArcGIS Pro yazılımlarında açılabilir. Coğrafi veriler Shapefile, CSV, KML, Excel, GDB ve GeoJSON dosya formatlarında sunucu dışına çıkarılabilmektedir.



Şekil 4.21 ArcGIS Online genel bakış sekmesi

ArcGIS Online Veri sekmesi üzerinde coğrafi veri katmanlarına ait öznetelik tabloları görüntülenebilmekte, yeni öznetelik alanları eklenebilmekte, mevcut öznetelik alanlarına ait tanımlamalar değiştirilebilmekte ya da istenildiği takdirde mevcut öznetelik alanları silinebilmektedir.

ArcGIS Online Görselleştirme sekmesinden coğrafi veri katmanları çevrimiçi Dünya Haritası üzerinde gösterimi yapılmak istenen katman seçilerek gösterilmekte, coğrafi verinin sembol ve renk seçimleri, istenilen özniteliğe bağlı olarak gösterim şeklinin belirlenmesi vb. seçeneklerle kullanıcı dostu arayüzü ile kolaylıkla coğrafi verileri yönetme imkânı sunmaktadır.

4.9.2. ArcGIS Online Map Viewer Classic üzerinde Web haritalarının oluşturulması

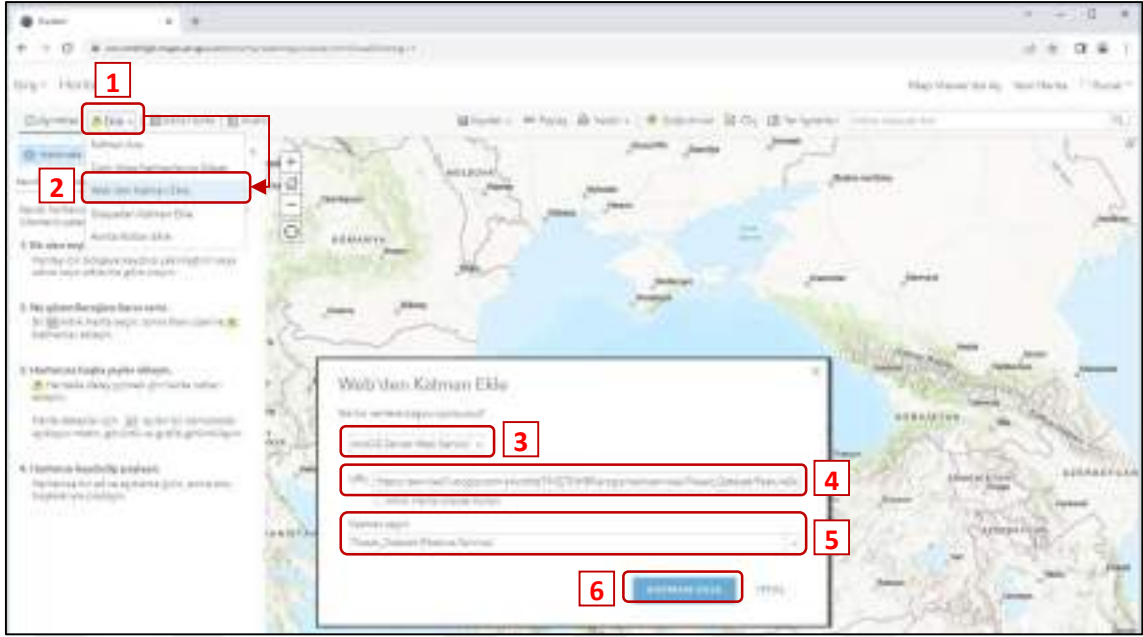
Web haritaları coğrafi veri katmanları, altlık haritalar ve yer işaretlerini içermektedir. Web haritasında kullanılacak coğrafi veriler için işaretleştirme, etiket, görünürlük ölçeği ayarları, katman seçim ayarları, açılır bilgi penceresi ayarlamaları ArcGIS Online Map Viewer Classic üzerinden gerçekleştirilerek Web haritaları oluşturulmaktadır. Web tabanlı çalışan ArcGIS WebApp uygulaması ya da ArcGIS Dashboard paneli geliştirmeden önce mutlaka Web haritalarının oluşturulması gerekmektedir. Web haritalarının projeksiyon sistemi olarak, çevrimiçi harita servisleri tarafından kullanılan ve yeryüzünün $\pm 85.05129^\circ$ enlemleri arasında kalan bölümünün gösterildiği Web Mercator projeksiyonu kullanılmıştır (Bildirici, 2015; Battersby vd., 2014; Bildirici ve Kırtıloğlu, 2016). ArcGIS Online Map Viewer Classic platformu üzerinde Web haritası oluşturma işlemleri adım adım aşağıdaki şekilde uygulanmıştır.

Map Viewer Classic harita oluşturma arayüzüne ArcGIS sunucuları üzerine servis edilen coğrafi veri katmanları ve tablolar Şekil 4.22’de görülen işlem adımlarıyla entegre edilmiştir.

(1) Map Viewer Classic uygulama arayüzünde “Ekle sekmesi” seçildikten sonra, verinin eklenecek katmanın hangi kaynaktan ekleneceği seçilmektedir.

(2) Bu çalışmada coğrafi veri katmanları Web ortamında bulunduğu için “Webden Katman Ekle” yöntemi seçilmiştir.

(3) “Webden Katman Ekle” seçeneğiyle karşımıza gelen yeni pencerede, kullanıcıya eklenmek istenen coğrafi veri yapısının formatı hakkında bilgi istenilmektedir. Katman olarak eklenebilecek veri formatları arasında ArcGIS Server Web Servisi, WMS OGC Web Servisi, WMTS OGC Web Servisi, WFS OGC Web Servisi, Döşeme (Raster Katmanı), KML, GeoRSS ve CSV Dosya formatları bulunmaktadır. Coğrafi verilerimiz ArcGIS sunucuları üzerinde barındırıldığı için ArcGIS Server Web Servisi seçeneği seçilmiştir.



Şekil 4.22 Map Viewer Classic arayüzüne coğrafi veri katmanlarının entegre edilmesi

(1) Map Viewer Classic uygulama arayüzünde “Ekle sekmesi” seçildikten sonra, verinin eklenecek katmanın hangi kaynaktan ekleneceği seçilmektedir.

(2) Bu çalışmada coğrafi veri katmanları Web ortamında bulunduğu için “Webden Katman Ekle” yöntemi seçilmiştir.

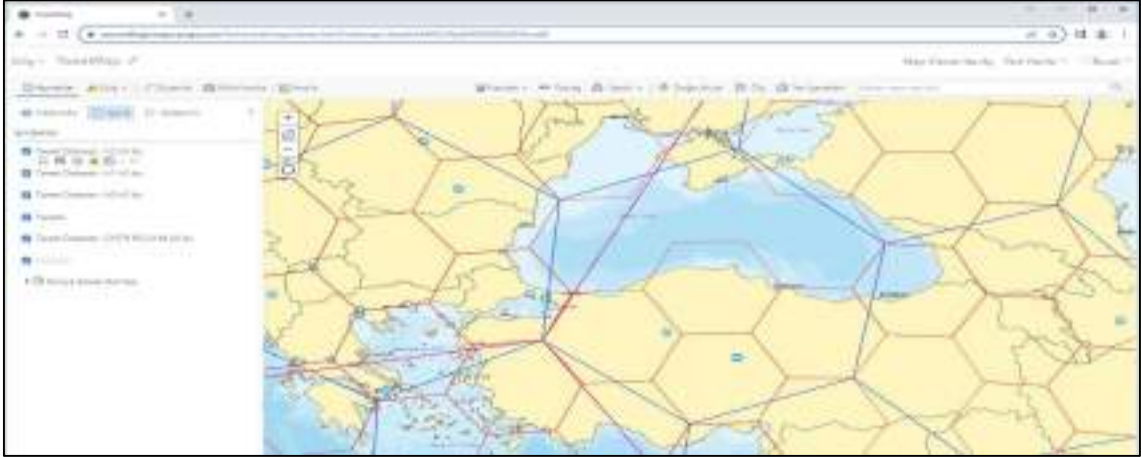
(3) “Webden Katman Ekle” seçeneğiyle karşımıza gelen yeni pencerede, kullanıcıya eklenmek istenen coğrafi veri yapısının formatı hakkında bilgi istenilmektedir. Katman olarak eklenebilecek veri formatları arasında ArcGIS Server Web Servisi, WMS OGC Web Servisi, WMTS OGC Web Servisi, WFS OGC Web Servisi, Döşeme (Raster Katmanı), KML, GeoRSS ve CSV Dosya formatları bulunmaktadır. Coğrafi verilerimiz ArcGIS sunucuları üzerinde barındırıldığı için ArcGIS Server Web Servisi seçeneği seçilmiştir.

(4) ArcGIS Server sunucuları üzerinde yer alan coğrafi veri katmanlarımızın URL tanımlaması girilmiştir. Coğrafi veri katmanlarımızın barındırıldığı URL adresi “https://services7.arcgis.com/****/arcgis/rest/services/Tweet_Dataset/FeatureServer” şeklinde tanımlanmıştır.

(5) URL tanımlaması yapılan ArcGIS Feature Server sunucu üzerinde çalışmak istenilen katman/katmanlar seçilerek, (6) katmanların Map Viewer Classic uygulama arayüzüne eklenilmesi sağlanmıştır.

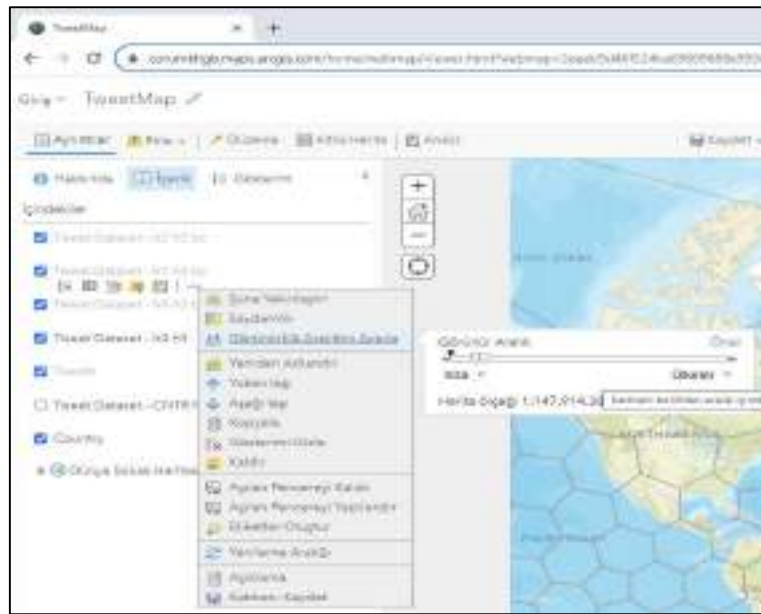
Coğrafi veri katmanları Map Viewer Classic arayüzüne eklendiğinde tüm katmanlar ekran üzerinde renk, işaret, görünürlük ölçeği, etiket düzenlemeleri, açılır

bilgi panelleri, harita yer işaretleri ayarlanmaksızın rastgele gösterilmektedir (Şekil 4.23). Kullanıcı dostu arayüze sahip bir Web haritası oluşturmak için coğrafi veri katmanlarına ait işaretlerin, renk, görünür ölçek aralıklarının, veri seçildiğinde açılır bilgi panellerinin, harita üzerinde belirli bölümleri referanslayacak yer işaretlerinin ayarlanması gerekmektedir.



Şekil 4.23 Coğrafi veri katmanlarının Map Viewer arayüzünde harita üzerinde görünümü

Coğrafi veri katmanlarının hem sunucu üzerinden veri istenmesi aşamasında veri trafiğinde yığılmayı önlemek hem de oluşturulacak Web haritalarında farklı çözünürlük seviyelerinde veri genelleştirmesi sağlayarak coğrafi verinin daha anlaşılabilir olmasını sağlamak için ölçek değişkenli görünür ölçek aralıkları tanımlanmıştır. Coğrafi veri katmanlarının görünür ölçek aralıkları Şekil 4.24'te görülen işlem adımlarıyla tanımlanmıştır.

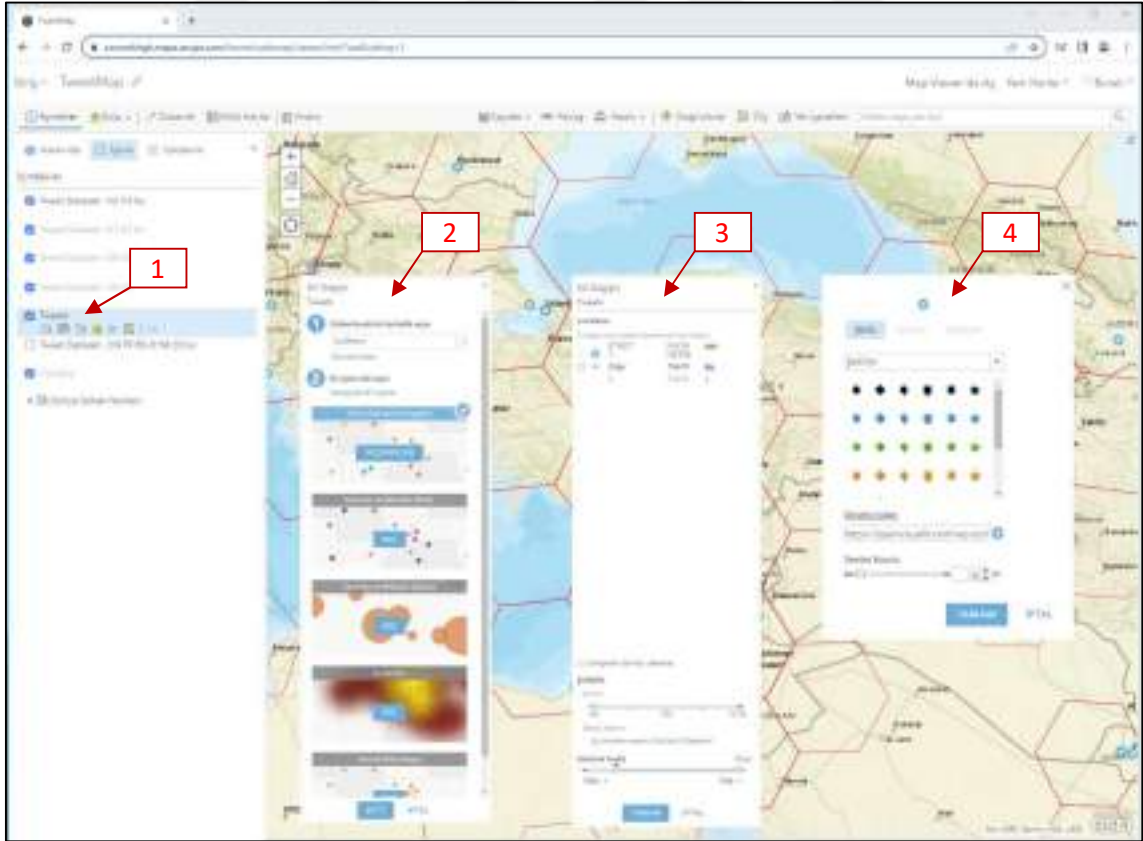


Şekil 4.24 Coğrafi veri katmanlarının görünür ölçek aralıklarının ayarlanması

Coğrafi veriler ile analog haritalar oluşturmada bilginin genelleştirilmesi, bilginin anlamsal olarak işaretlenmesi (semiyoji) ve estetiksel üretimine dikkat ediliyorsa, Web haritalarında da aynı temel prensiplere uyulması kullanıcının Web haritasıyla olan etkileşiminin sağlanması açısından önem arz etmektedir. Bu bakımdan farklı coğrafi katmanlarda yer alan verilerin birbirleriyle net bir şekilde ayırt edilmesini sağlamak ve kullanıcının görsel algısını artırmak için veri yapılarına göre işaret ve renk tanımlaması Şekil 4.25’te görülen işlem adımlarıyla gerçekleştirilmiştir.

(1) Map Viewer Classic uygulama arayüzünde coğrafi veri katmanı seçildikten sonra “Stil Değiştir” sekmesi seçilerek coğrafi verinin hangi öznelik bilgisine ve gösterim stiliyle görselleştirilmesinin belirlendiği (2) numaralı arayüze geçilmektedir.

(2) Görselleştirme stillerinin belirlendiği arayüzde, veri yapısının temsiline uygun olarak Türler (Benzersiz simgeler), Sayımlar ve Miktarlar (Renk), Sayımlar ve Miktarlar (Boyut), Isı Haritası ve Konum (Tek Simge) olmak üzere farklı gösterim stilleri sunulmaktadır. Kullanıcı kendi tasarımına uygun gösterim stilini seçtikten sonra stil değişikliklerinin ve tanımlamalarının yapıldığı (3) numaralı arayüze geçilmektedir.



Şekil 4.25 Coğrafi verilerin işaretleştirme ve renk değişikliklerinin gerçekleştirilmesi

(3) Stil deęişikliklerinin ve tanımlamaların yapıldığı bu arayüzde, Web haritası üzerinde gösterilecek coęrafi veriler belirlenmekte (kullanıcı öznitelik deęişkenlerine baęlı olarak göstermek istemedięi verileri pasif hale getirebilir), coęrafi verinin harita lejantında görülen etiketleri tanımlanmakta, coęrafi verinin şeffaflık seviyesi ayarlamaları yapılmakta ve coęrafi verinin görünürlük aralıkları belirlenmektedir. İşaretleştirmenin yapılabilmesi için ilgili katmanın sol tarafında yer alan “Simge Deęiştir” sekmesi seçilerek (4) numaralı arayüze geçilmektedir.

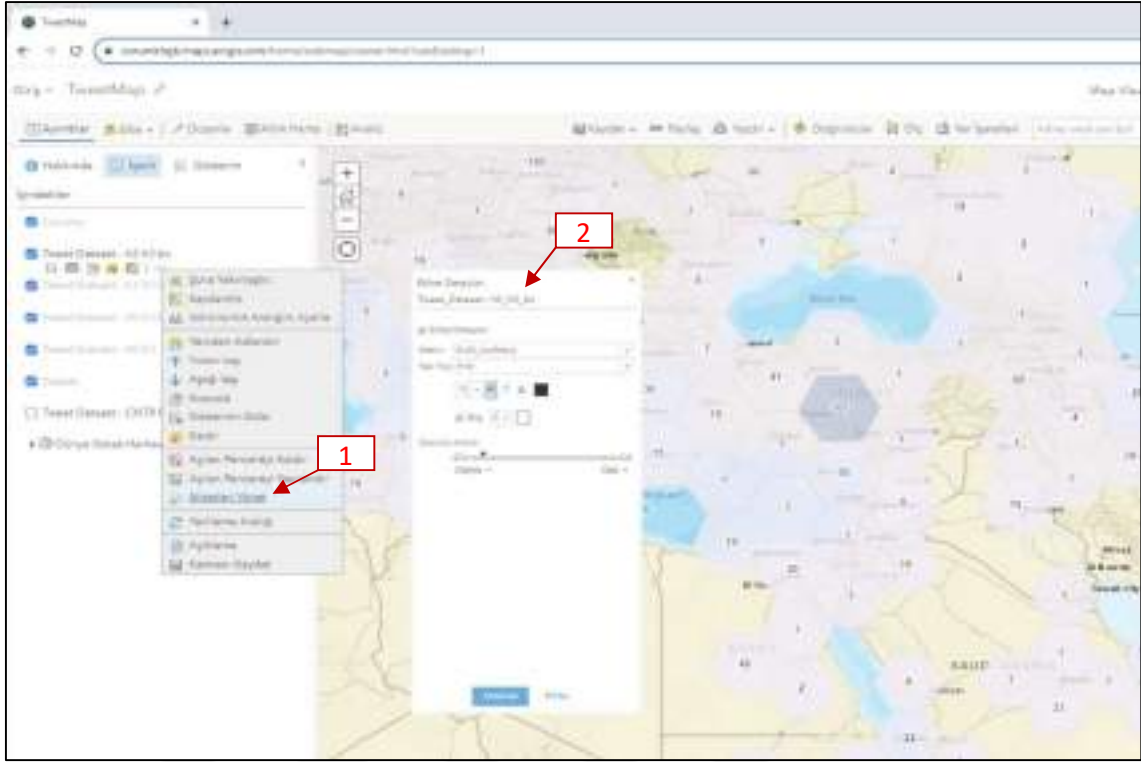
(4) Bu arayüzde coęrafi verinin temsilinde kullanılacak işaretler seçilmekte ve işaret boyutları piksel ölçütlü olarak ayarlanmaktadır. Bu çalışmada, tweet verilerinin temsilinde X’in önceki logosu olan mavi arka planlı kuş logosu kullanılmıştır. Bu şekildeki kullanıcının kendine özel tasarımları kullanabilmesi için bu arayüzde “Görüntü kullan” sekmesi seçilerek, görüntü dosyasının bulunduğu URL adresinden eklenmesi gerekmektedir.

Web haritası üzerinde coęrafi verilerin yapısına göre barındırdıkları öznitelik bilgilerine ait (detay kodları, isimlendirmeleri, miktarları vb.) özellikler etiket şeklinde Web haritaları üzerine yerleştirilebilmektedir. Bu işlem için ilgili coęrafi veri katmanı seçilerek Şekil 4.26’da görülen işlem adımları gerçekleştirilmektedir.

(1) Map Viewer Classic uygulama arayüzünde coęrafi veri katmanı seçildikten sonra “Etiketleri Yönet” sekmesi seçilerek coęrafi veri için oluşturulacak etiket detaylarının bulunduğu (2) numaralı arayüze geçilmektedir.

(2) Bu arayüzde coęrafi veriye ait oluşacak etiketin kullanılacağı öznitelik bilgisi, etiket oluşturmada kullanılacak yazı tipi ve yazı tipi özellikleri ve oluşturulacak etiketlerin görünürlük aralığı seçilmektedir. Şekil 4.26’da görülen ekran görüntüsünde çözünürlük 2 seviyesinde oluşturulan Uber H3 altıgen hücrelerinin kapladığı alandaki tweet sayıları için etiket oluşturulmuştur.

İnteraktif Web harita kullanıcısının harita ekranındaki coęrafi verinin barındırdığı; öznitelik bilgilerine, farklı formatlardaki (PDF, JPG, DOC, XLS vb.) dosya eklerine, harici bir Web sayfasına yöneltme linklerine erişimi Web haritası üzerinde coęrafi veri işaretlendiğinde aktif olan Açılır Pencere (Pop-up) üzerindeki tanımlanan alanlar ile sağlanmaktadır.

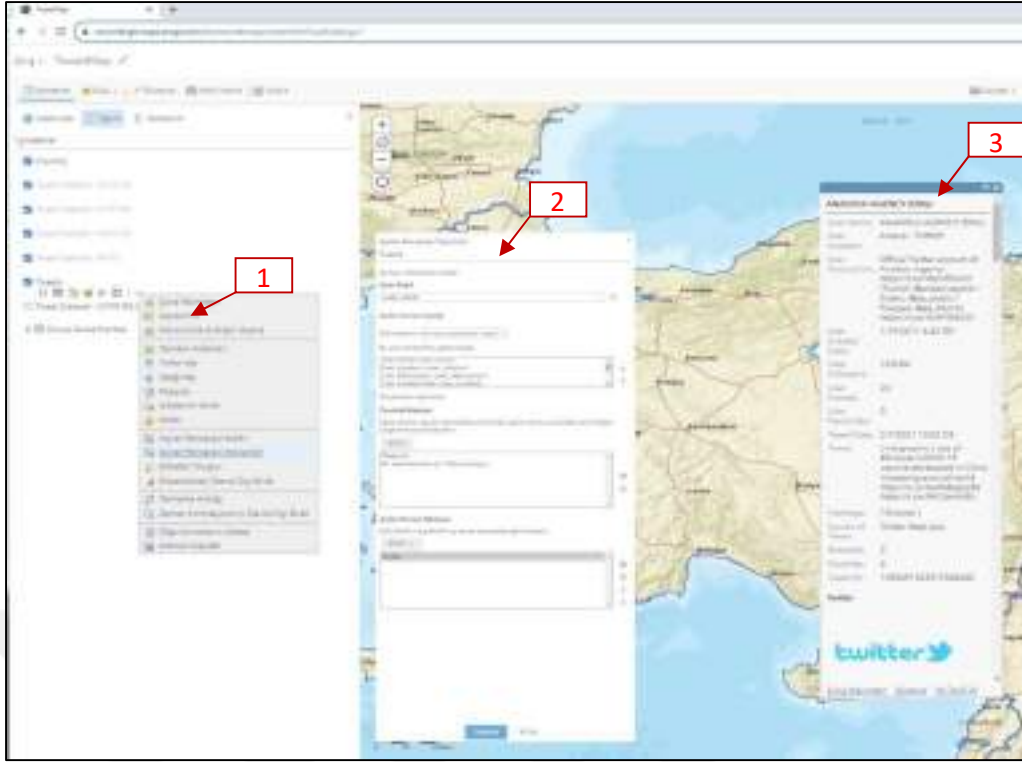


Şekil 4.26 Coğrafi veriler için etiket oluşturulması

Map Viewer Classic uygulaması üzerinde bir coğrafi veri katmanı için “Açılır Pencere (Pop-up)” oluşturulma adımları Şekil 4.27’de gösterilmektedir.

(1) Map Viewer Classic uygulama arayüzünde coğrafi veri katmanı seçildikten sonra “Açılan Pencereyi Yapılandır” sekmesi seçilerek coğrafi veri için oluşturulacak açılır pencerede yer alacak detaylarının bulunduğu (2) numaralı arayüze geçilmektedir.

(2) Bu arayüzde coğrafi veriye ait açılır pencerelerde yer alacak açılır pencere başlığı, açılır pencerede yer alacak öznitelik bilgileri ve görüntülenme etiketleri, açılır pencerede görüntülenecek coğrafi veri ekleri, grafikler, URL linkleri tanımlamaları gerçekleştirilmektedir. Bu arayüzde yapılacak doğru tanımlamalar sonrasında, harita ekranı üzerinde coğrafi veri seçildiğinde harita ekranı üzerinde (3) numara pencere görüldüğü şekilde açılır pencere aktif olmaktadır. Bu açılır pencere üzerinde harita ekranı üzerinde seçilen veriye ait öznitelik bilgileri sunulmakta, bu örnekte görüldüğü gibi en altta yer alan Twitter logosu seçildiğinde ise harita ekranı üzerinde seçilen tweet verisinin yayınlandığı gerçek X sayfasına erişim sağlanmaktadır.



Şekil 4.27 Coğrafi veriler için açılır pencere (pop-up) oluşturma

Web haritası üzerinde oluşturulacak “Yer İşaretleri” sayesinde, kullanıcı Web haritası üzerinde önceden tanımlanan bölgelere (kıta, ülke, şehir vb.) ya da adreslere kolaylıkla erişim sağlayabilmektedir. Web haritası üzerinde “Yer İşaretleri” oluşturma işlemi için öncelikle harita ekranı üzerinde “Yer İşareti” tanımlanacak alana yaklaşılmalı, ardından Şekil 4.28’de görüldüğü gibi Map Viewer Classic uygulama arayüzünde yer alan “Yer İşaretleri” sekmesi işaretlenerek açılan pencerede “Yer İşareti Ekle” alanına tanımlanacak alanın tanımlaması yapılarak yer işareti oluşturulmalıdır.



Şekil 4.28 Web haritası üzerinde yer işareti oluşturma

Bu bölümde bir Web haritası oluşturulma adımlara tüm yönleriyle açıklanmaya çalışılmıştır. Bu işlem adımlarında oluşturulan TweetMap olarak isimlendirilen bu Web haritası tez çalışması kapsamında üretilen haritalardan sadece bir tanesidir. Tez çalışması kapsamında toplamda 3 farklı Web haritası oluşturulmuştur.

Ülkelerin X kullanıcı sayıları ve tweet sayılarına göre kümelenmesini gösteren iki değişkenli koroplet Web haritası ve ülkelerin oluşturulan toplam tweet sayısı, Covid kelimesi içeren tweet sayısı ve vaccine kelimesi içeren tweet sayılarına göre kümelenmesini gösteren üç değişkenli koroplet Web haritası hazırlanmıştır. Çok değişkenli koroplet haritaların hazırlanmasında, ülkelerin idari sınır verileri, Ward yöntemiyle elde edilen kümeleme analizi sonuçları ve k-ortalamalar (k-Means) yöntemiyle elde edilen kümeleme analizi sonuçları kullanılmıştır.

Web haritaları oluşturma işlemlerinin tamamlanmasının ardından, Web haritalarının interaktif Web tabanlı kullanılabilmesi için ArcGIS Web AppBuilder ile interaktif veri editörü uygulaması; ArcGIS Dashboard ile tweet verilerinden elde edilen bilgilerin ve kümeleme analizi sonuçlarının kullanıcıya interaktif haritalar, gösterge sayaçları, filtreleme araçları ve dinamik grafikler ile sunulduğu dashboards uygulaması oluşturmaya geçilmiştir.

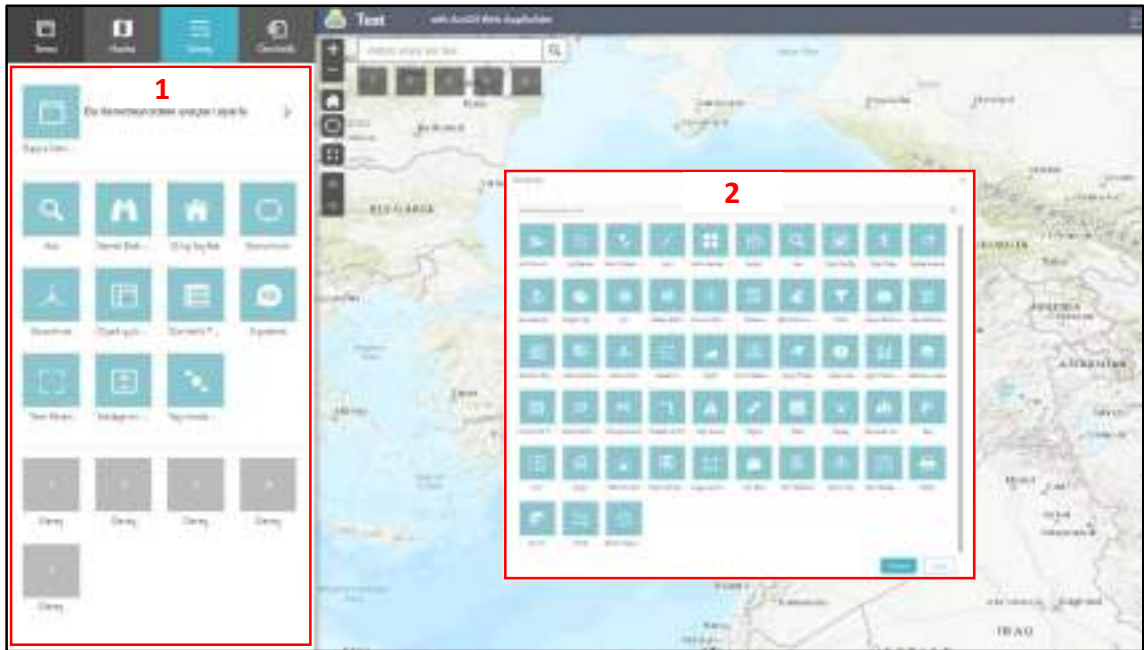
4.9.3. ArcGIS Web AppBuilder ile interaktif veri editörü uygulaması geliştirme

Tweet verileriyle Web haritası oluşturulduktan sonra kullanıcıların, tweet verilerine erişmesini, geometrik olarak düzenleme yapmasını, zamansal olarak filtrelemesini, ekrandan seçtiği detaylar hakkında istatistiki bilgiler edinmesini, tweet verilerinin zamansal gelişimini izlemesini sağlayan Web tabanlı uygulama geliştirilmiştir. ArcGIS Web AppBuilder üzerinde temel anlamda uygulama geliştirmede kullanılan işlem adımları aşağıdaki şekilde sıralanmaktadır.

ArcGIS Online platformu üzerinde Şekil 4.29’da görüldüğü gibi “Uygulama Oluştur” sekmesi üzerinden Web AppBuilder sekmesi seçilerek, bir sonraki pencerede uygulamaya ait başlık, uygulamanın tanınmasına yardımcı etiketler, uygulama hakkında özet bilgi ve uygulamanın kaydedileceği klasör bilgileri seçilerek uygulama başlatılmaktadır.

“Stil” sekmesinin altında yer alan “Düzen” bölümünde, harita ekranı üzerinde yer alan harita kontrol ve uygulama araçlarının arayüz üzerinde yerleşim düzenine ait şablon yapısı seçilmektedir. “Tema” sekmesinin hemen sağında yer alan “Harita” sekmesinde, geliştirilecek Web uygulamasında kullanılacak Web haritası ve katmanları uygulama arayüzüne eklenmektedir. “Harita” sekmesinin sağında yer alan “Gereç” sekmesi seçildiğinde Şekil 4.31’de görülmekte olan 1 numaralı uygulama arayüzü gelmektedir. Bu arayüz üzerinde turkuaz renkte görülmekte olan araçlar Web AppBuilder uygulamasının, geliştirilecek uygulama için ön tanımlı olarak uygulama oluşturucusuna sunduğu araçlardır. Bu araçların her birinin sağ üst köşesinde yer alan göz simgesi aktif/pasif hale getirilerek araçların uygulama arayüzünü eklenmesi ve kaldırılması sağlanmaktadır.

Şekil 4.31’de görülmekte olan 1 numaralı uygulama arayüzünün alt kısmında yer alan gri renkli 5 gereç ise Web AppBuilder tarafından uygulama geliştiricisi tarafından eklenebilecek araçların seçildiği alandır. Bu sekmeler seçildiğinde Şekil 4.31’de 2 numaralı uygulama arayüzünde görülmekte olan 53 adet farklı uygulama araçlarından eklenilmek istenen araç seçilerek eklenebilmektedir. Bu tez çalışmasında geliştirilen “Covid-19 Vaccines Data Editör” isimli Web uygulaması Şekil 4.32’de yer almaktadır.



Şekil 4.31 Web AppBuilder uygulaması gereçlerinin bulunduğu arayüz

Uygulama arayüzün sol üst bölümünde ve sağ üst bölümünde, harita ekranını yönetmekte kullanılan yakınlaştırma / uzaklaştırma, varsayılan ekran yayılımı, önceki /

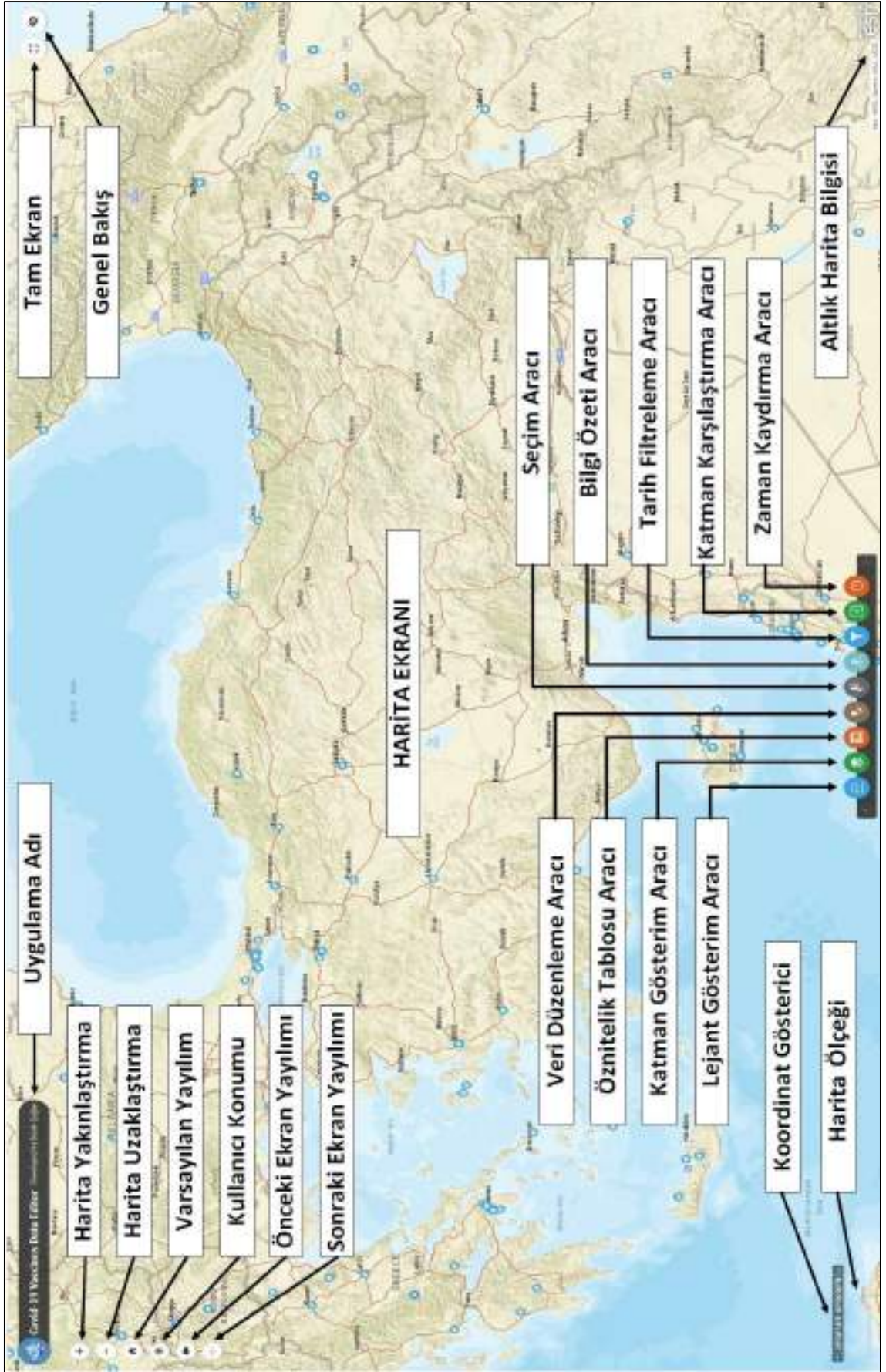
sonraki ekran yayılımları, tam ekran ve genel bakış araçları bulunmaktadır. Uygulama arayüzünün sol alt bölgesinde, harita ekranı üzerinde imlecin gezdiği koordinatları gösteren ya da harita ekranı üzerinde bir konum işaretlendiğinde koordinat bilgisini kullanıcıya gösteren araç ve Web haritasına ait ölçek yer almaktadır. Uygulama arayüzünün alt-orta bölgesinde coğrafi verilere ait gösterim, düzenleme, seçim, bilgi özeti, tarih filtreleme, katman karşılaştırma ve zaman kaydırma araçları yer almaktadır.

Lejant gösterim aracı, dinamik yapıda olup ekran yayılımında yer alan coğrafi verilere ait işaretleri ve açıklamalarını göstermektedir.

Katman gösterim aracı, uygulama kullanılan Web haritasının içerdiği coğrafi veri katmanlarının gösterildiği alandır. Dinamik yapıda olup, haritanın görünürlük ölçeğine bağlı olarak ekranda görülen coğrafi katmanların aktif ve pasif durumunu göstermektedir.

Öznetelik Tablosu aracı, Web uygulaması üzerinde coğrafi verilere ait öznetelik tablosunu aktif / pasif hale getiren araçtır. Öznetelik tablosu dinamik yapıda olup, harita ekranı yayılımında bulunan coğrafi verilere ait detayları listelemektedir.

Veri düzenleme aracı, coğrafi verilere ilişkin geometrik ve özneteliksel bilgilerin düzenlenmesini sağlayan araçtır.



Şekil 4.32 "Covid-19 Vaccines Data Editor" Web uygulaması arayüzü

Seçim aracı, coğrafi verilerin noktasal, çizgisel, dikdörtgen, daire, çoklu doğruyla, kement yöntemiyle seçim yapılmasını sağlayan araçtır.

Bilgi özeti aracı, harita ekranında yer alan tweet sayısı ve tweet içeriklerinin gösterildiği pencereyi aktif/pasif duruma getiren araçtır. Dinamik yapıda olup, harita yayılım değiştiğinde bilgi özeti penceresinde yer alan bilgilerde değişmektedir.

Tarih filtreleme aracı, harita ekranı üzerinde görülmekte olan tweet verilerinin kullanıcı tarafından belirtilen tarihler arasında filtrelenmesini sağlayan araçtır.

Katman karşılaştırma aracı, harita ekranı üzerinde bulunan coğrafi verileri, harita ekranını dikey olarak bölerek karşılaştırmalarını sağlayan araçtır.

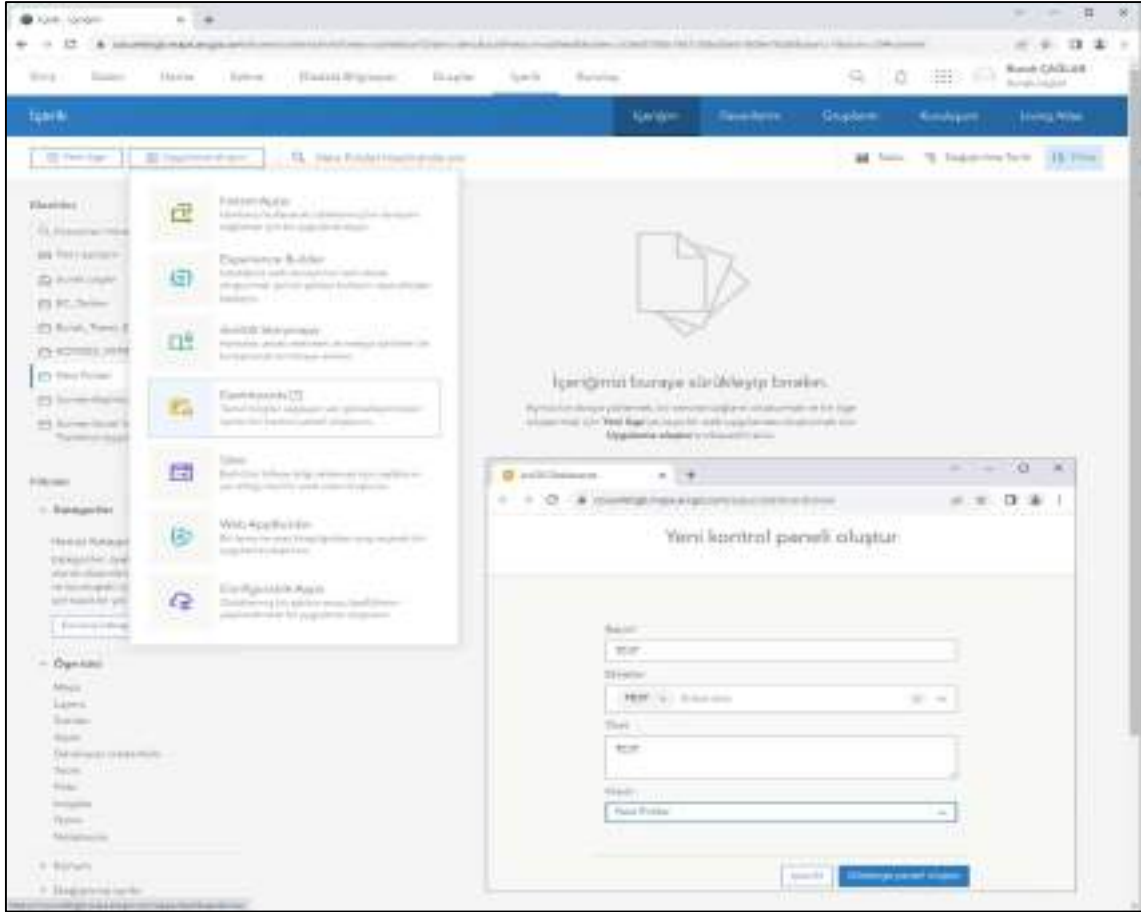
Zaman kaydırma aracı, harita ekranı üzerinde görülen tweet verilerini uygulama geliştiricisinin belirlemiş olduğu zaman periyoduna göre zamansal gelişimini animasyon halinde gösterilmesini sağlayan araçtır.

4.9.4. Covid-19 aşılılarıyla ilgili tweetleri gösteren bilgi panosunun oluşturulması

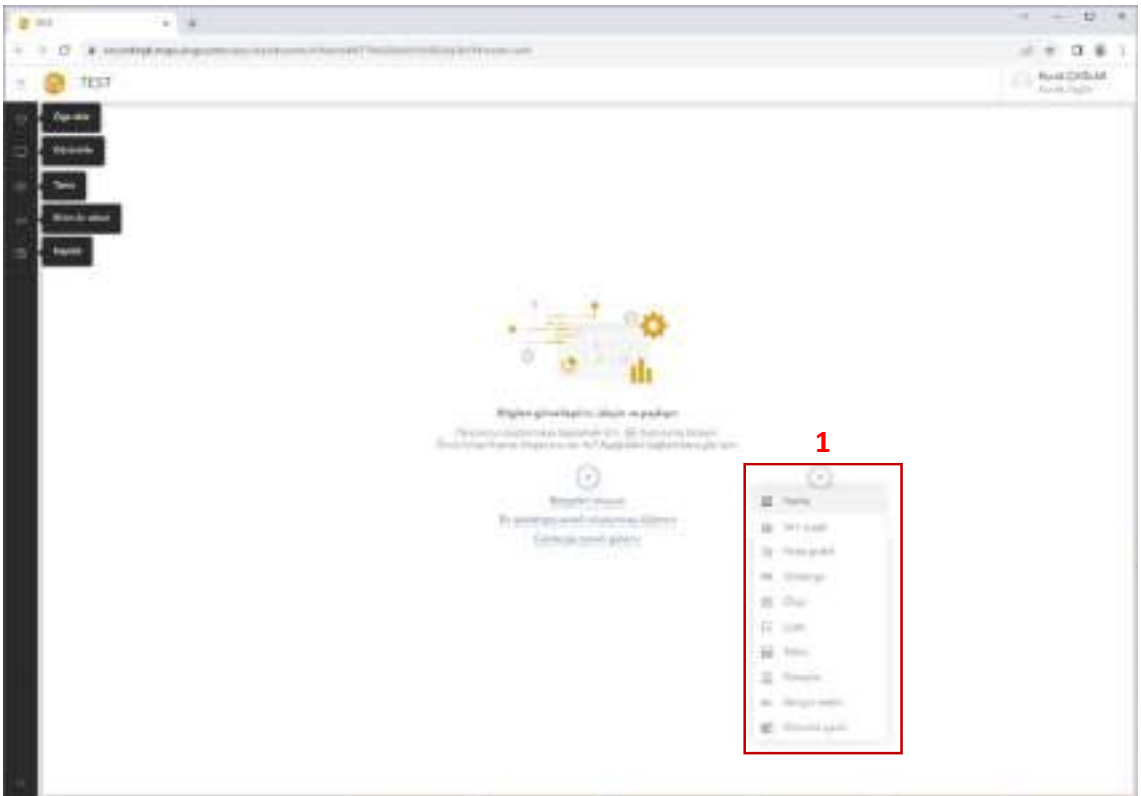
Covid-19 aşılıları hakkındaki tweet verilerinden elde edilen bilgilerin kullanıcıya interaktif harita, gösterge sayaçları, filtreleme araçları ve dinamik grafikler ile sunum amacıyla ArcGIS Dashboards uygulaması ile bilgi panosu oluşturulması amaçlanmıştır.

ArcGIS Online platformu üzerinde Şekil 4.33'te görüldüğü gibi "Uygulama Oluştur" sekmesi üzerinden Dashboards sekmesi seçilerek, bir sonraki pencerede uygulamaya ait başlık, uygulamanın tanınmasına yardımcı etiketler, uygulama hakkında özet bilgi ve uygulamanın kaydedileceği klasör bilgileri seçilerek uygulama başlatılmaktadır.

Web AppBuilder uygulamasına başlatıldığında Şekil 4.34'te görülen uygulama geliştirme arayüzüne erişim sağlanmaktadır. Dashboards uygulama geliştirme platformu ekranında, sol üstte yer alan "Öğe ekle" butonuna ya da arayüz ekranının ortasında yer alan (+) sembolü seçildiğinde, uygulama arayüzünde Şekil 4.34'te görülen (1) numaralı pencere açılmaktadır. Bu pencere yardımıyla uygulama arayüzüne, Web haritaları, seri çizelge, pasta grafik, gösterge, ölçü, liste, tablo, detaylar, zengin metin ve gömülü içerik ekleme araçları aktif hale gelmektedir. Bir konu hakkında gösterge paneli oluşturmak için, öncelikli olarak o konuya ilişkin Web haritası oluşturmalı ve dashboards arayüzüne ilk olarak Web haritasının eklenmesi gerekmektedir.



Şekil 4.33 ArcGIS Online üzerinden Dashboards uygulamasına giriş



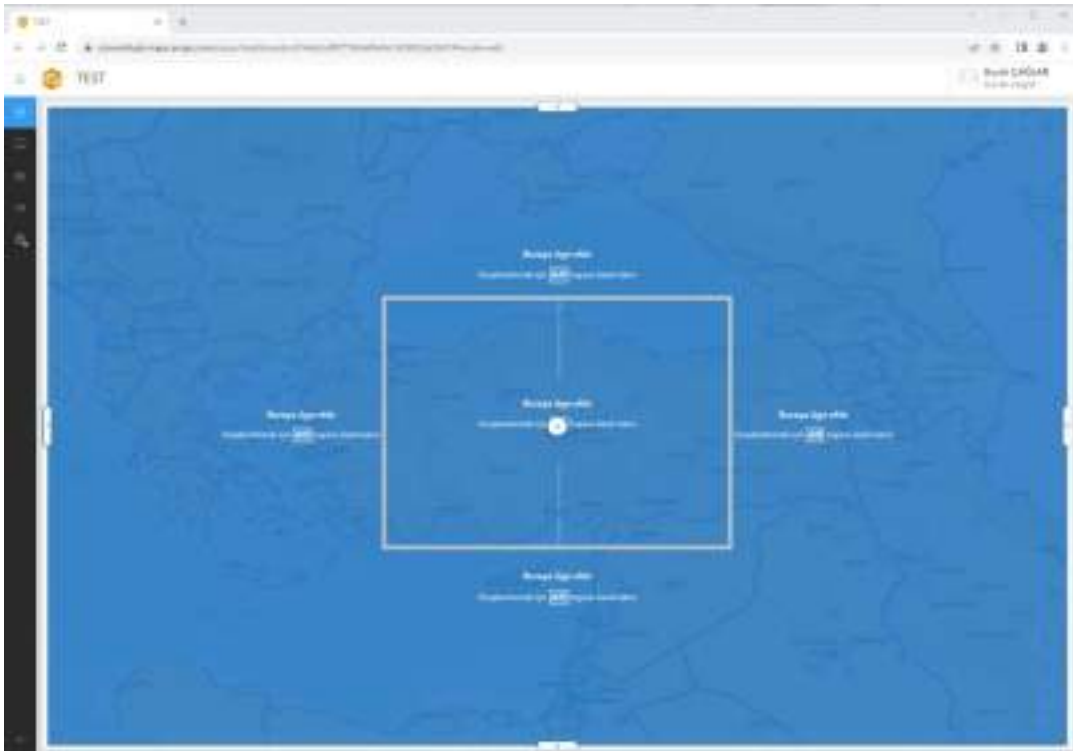
Şekil 4.34 Dashboards uygulama geliştirme platformu

Eklenilmek istenilen harita seçildikten sonra Web haritasına ait ayarların gerçekleştirildiği Şekil 4.35’te görülen pencere gelmektedir.



Şekil 4.35 Dashboards uygulamasında Web haritası ayarlarının gerçekleştirildiği arayüz

Web haritasının Dashboards uygulama platformuna eklenilmesinin ardından, bilgi panelini oluşturacak liste, grafik vb. araçların uygulama platformuna eklenmesi gerekmektedir. Şekil 4.34’te görülmek olan uygulama arayüzünün sol üst köşesinde yer alan “Öğe Ekle” butonu seçildiğinde, uygulama arayüzünün üst, alt, sağ, sol ve orta kısmında (+) işareti ortaya çıkmaktadır. Bu (+) işaretlerine seçim için yaklaşıldığında Şekil 4.36’da görülen “Buraya Öğe Ekle” pencereleri aktif hale gelmektedir.



Şekil 4.36 Dashboards uygulamasında öğelerin eklenileceği alanlar gösteren arayüz

Kullanıcı tarafından oluşturulması planlanan Dashboards uygulamasına ait seri çizelge, pasta grafik, gösterge, ölçü, liste, tablo, detaylar, zengin metin ve gömülü içerik öğeleri genel parametre ayarları yapılarak ve Şekil 4.36’da belirtilen alanlara eklenerek Dashboards uygulaması oluşturulmaktadır.

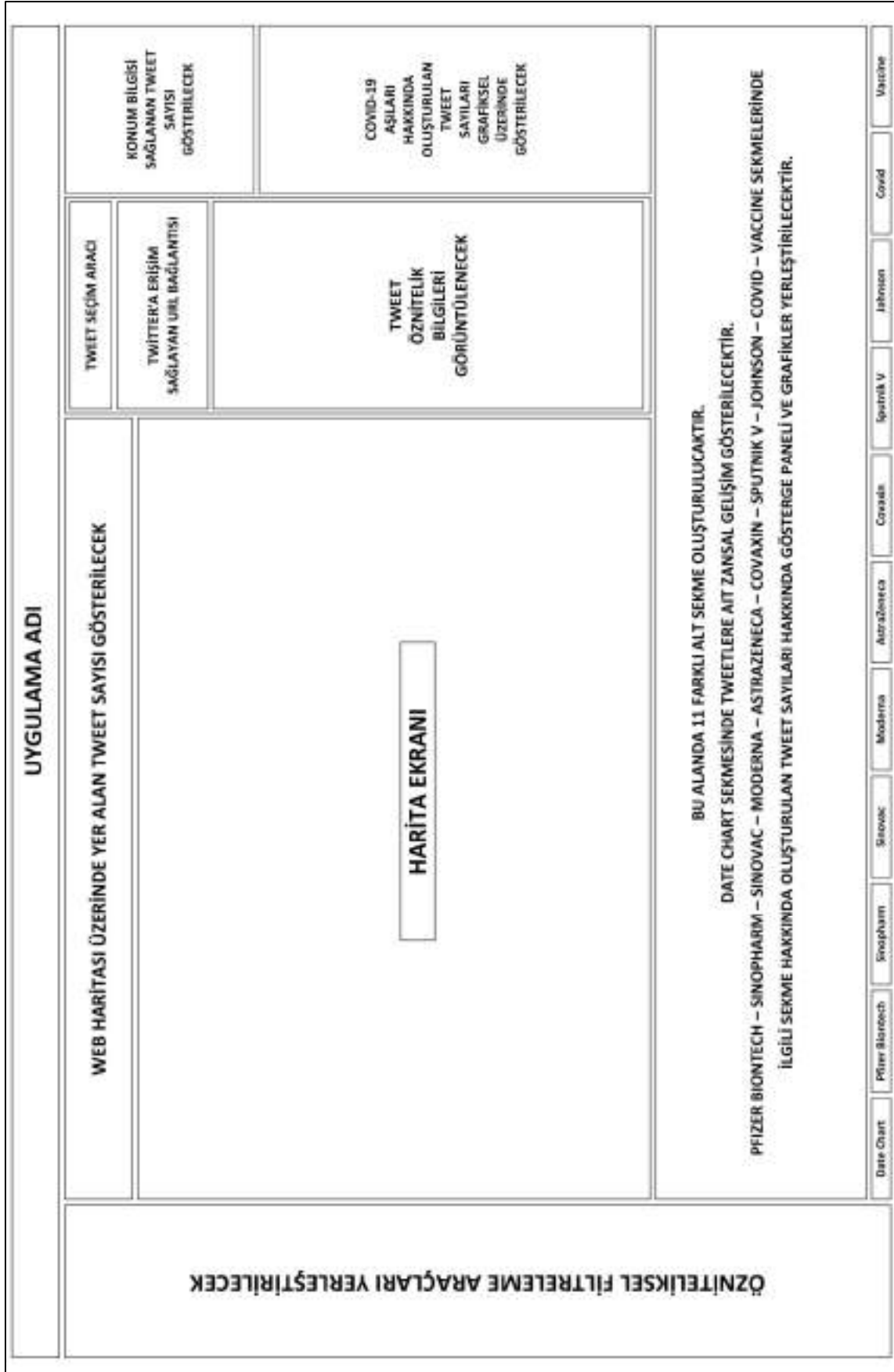
Dashboards uygulamasıyla hazırlanan bilgi panellerinin amaçlanan konu hakkında detaylı harita, bilgi ve grafikleri yansıtmalarının yanı sıra kullanıcı dostu arayüze sahip olması çok önemlidir. Bilgi panelleri kullanacak insanlar, tek ekran üzerinde gösterilen birden fazla öğe arasında algı karışıklığı yaşamamalıdır. Kullanıcıya aktarılabilecek bilgilerin olabildiğince sade ve anlaşılır öğelerle görüntüleyiciye aktarılması gereklidir. Bu nedenle öncelikle bilgi panelinin taslak şeması oluşturulmalı daha sonra uygulama oluşturma aşamasına geçilmelidir. Covid-19 aşılılarıyla ilgili tweetleri gösteren bilgi panosuna ait taslak şema Şekil 4.37’de görülmektedir.

Şekil 4.37’de yer alan taslak şema doğrultusunda, Dashboards platformunda geliştirilen bilgi paneli uygulaması Şekil 4.38’de görülmektedir. Şekil 4.38’de yer alan bölümleri adım adım aşağıdaki şekilde açıklanmaktadır.

(1) numaralı bölüm, tweet coğrafi veri setinden oluşturulan Web haritasının eklendiği bölümdür. Harita ekranı üzerinde, tweet verilerinin mekânsal takibi gerçekleştirilmektedir. Bir konumdaki tweet verisini ya da bir bölgedeki tweet verilerini harita ekranının sınırlandırdığı yayılım alanında görüntülendiğinde, bilgi paneli göstergeleri harita yayılım alanıyla etkileşimli çalışarak harita ekranı üzerinde görülmekte olan tweet verilerinden üretilen bilgileri kullanıcıya göstermektedir.

(2) numaralı bölüm, geliştirilen bilgi paneli uygulamasının isminin bulunduğu alandır.

(3) numaralı bölüm, uygulama geliştiricisinin tanımlamış olduğu özneliksel sorgulamaların kullanıcıya sunulduğu alandır. Bilgi paneli uygulama kullanıcısı, bu alandan yer alan ön tanımlı sorgulamaları kullanarak, yaptığı sorgulama sonuçlarını etkileşimli olarak harita ekranı üzerinde mekânsal takibini ve diğer grafik göstergeler üzerinde ise istatistiksel takibini gerçekleştirmektedir. Bu alanda yer alan sorgu araçları tek başına ya da çoklu olarak kullanılabilir. Sorgulama araçları dinamik yapıda çalışmakta olup, harita ekranı yayılımındaki yer alan coğrafi verilere göre sorgulama araçları içeriğini güncellemektedir.



Şekil 4.37 COVID-19 aşılarıyla ilgili tweetleri gösteren bilgi panosu taslak şeması

Bu bölümde,

Tweet verilerinin atıldığı ülkelere göre sorgulanması,

Tweet verilerinin zamansal (tarih) aralığına göre sorgulanması,

Tweet verilerinin oluşturulduğu uygulama platformuna göre sorgulanması,

Pfizer Biontech aşısı hakkında oluşturulan tweetlerin sorgulanması,

Pfizer Biontech aşısı hakkındaki tweetleri oluşturan kullanıcıların sorgulanması,

Sinopharm aşısı hakkında oluşturulan tweetlerin sorgulanması,

Sinopharm aşısı hakkındaki tweetleri oluşturan kullanıcıların sorgulanması,

Sinovac aşısı hakkında oluşturulan tweetlerin sorgulanması,

Moderna aşısı hakkında oluşturulan tweetlerin sorgulanması,

Moderna aşısı hakkındaki tweetleri oluşturan kullanıcıların sorgulanması,

Oxford AstraZeneca aşısı hakkında oluşturulan tweetlerin sorgulanması,

Oxford AstraZeneca aşısı hakkındaki tweetleri oluşturan kullanıcıların sorgulanması,

Covaxin aşısı hakkında oluşturulan tweetlerin sorgulanması,

Covaxin aşısı hakkındaki tweetleri oluşturan kullanıcıların sorgulanması,

Sputnik V aşısı hakkında oluşturulan tweetlerin sorgulanması,

Sputnik V aşısı hakkındaki tweetleri oluşturan kullanıcıların sorgulanması,

Johnson&Johnson aşısı hakkında oluşturulan tweetlerin sorgulanması,

Johnson&Johnson aşısı hakkındaki tweetleri oluşturan kullanıcıların sorgulanması,

“Covid” kelimesi hakkında oluşturulan tweetlerin sorgulanması,

“Covid” kelimesi hakkındaki tweetleri oluşturan kullanıcıların sorgulanması,

“Vaccine” kelimesi hakkında oluşturulan tweetlerin sorgulanması,

“Vaccine” kelimesi hakkındaki tweetleri oluşturan kullanıcıların sorgulanması işlemlerine ait sorgu araçları bulunmaktadır. Şekil 4.39’da sorgu aracı örnekleri görülmektedir.



Şekil 4.39 Bilgi paneli uygulamasında bulunan sorgu aracı örnekleri

(4) numaralı bölüm, harita ekranı yayılımında yer alan toplam tweet sayısının gösterildiği alandır.

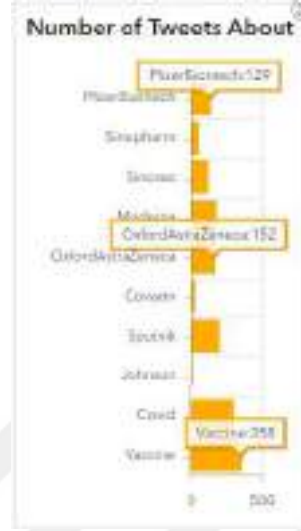
(5) numaralı bölüm, harita ekranı yayılımında yer alan tweetlere ait öznitelik bilgilerinin, görüntülediği alandır. Bu bölümün en üstünde yer alan oklar yardımıyla tweetler arasında geçiş sağlanmaktadır. Twitter logosuna tıklandığında ise seçili tweet'in yayınlandığı X sayfasına erişim sağlanmaktadır. Bu bölüm dinamik yapıda olup, harita ekranı, sorgu araçları ve diğer grafiklerle etkileşimli çalışmaktadır (Şekil 4.40).



Şekil 4.40 Tweet özniteliklerinin gösterildiği bilgi penceresi

(6) Bu bölüm, toplam tweet sayısının ve coğrafi olarak konumlandırılmış tweet sayısının at nalı formundaki grafik üzerinde gösterildiği alandır. Grafik statik yapıdadır.

(7) Bu bölüm, Covid-19 aşılıları hakkındaki tweetler ile “covid” ve “vaccine” kelimesi hakkındaki tweet sayılarının yatay sütun grafik üzerinde gösterildiği alandır. Bu alandaki grafik dinamik yapıda olup, harita ekranı yayılımı ve sorgu araçlarıyla etkileşimli olarak çalışmaktadır (Şekil 4.41).



Şekil 4.41 Yatay sütun grafik üzerinde tweet sayıları

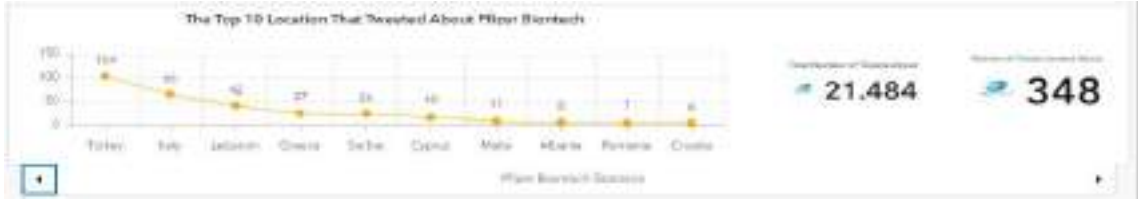
(8) Bu bölüm, 11 farklı alt sekmeye göre değişen dinamik yapıdaki bilgi penceresidir. Date Chart sekmesinde tweetlerin zamansal dağılımları dikey sütun grafik üzerinde gösterilmektedir. Harita ekran yayılımı ve sorgu araçlarıyla etkileşimli çalışmaktadır (Şekil 4.42).



Şekil 4.42 Tweet verilerinin zamansal (1 ay aralıklı) dağılımlarının dikey sütun grafik üzerinde gösterimi

Covid-19 aşılılarına ait alt sekmelerde ise Şekil 4.43'te görülen çizgisel grafik ve bilgi pencereleri yer almaktadır. Çizgisel grafik, harita ekranı yayılımında olan ülkelerden ilgili Covid-19 aşısı hakkında (Bu örnekte Pfizer Biontech) en çok tweet atılan 10 ülkeyi göstermektedir. Çizgisel grafik dinamik yapıda olup, harita ekran yayılımı ve sorgu araçlarıyla etkileşimli çalışmaktadır. Çizgisel grafik, hemen sağında yer alan gösterge tweet veri setindeki Pfizer Biontech aşılıları hakkında konumlandırılan ve konumlandırılmayan toplam tweet sayısını (21.484) göstermektedir. Bu bilgi penceresi statik yapıdadır. Bu bilgi penceresinin hemen sağında bulunan bilgi penceresi ise harita ekranında yer alan konumlandırılmış toplam tweet sayısını göstermektedir. Bu

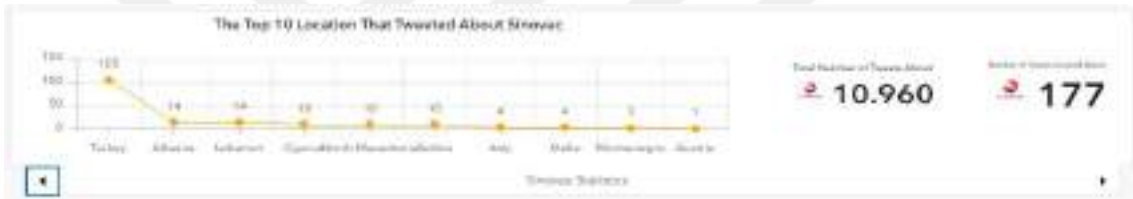
bilgi penceresi de dinamik yapıda olup, harita ekran yayılımı ve sorgu araçlarıyla etkileşimli çalışmaktadır. Diğer Covid-19 aşılı, “covid” ve “vaccine” kelimesini içeren tweetlere ait istatistikler Şekil 4.43 – 4.50’de gösterilmektedir.



Şekil 4.43 Pfizer Biontech aşısı hakkındaki bilgi penceresi



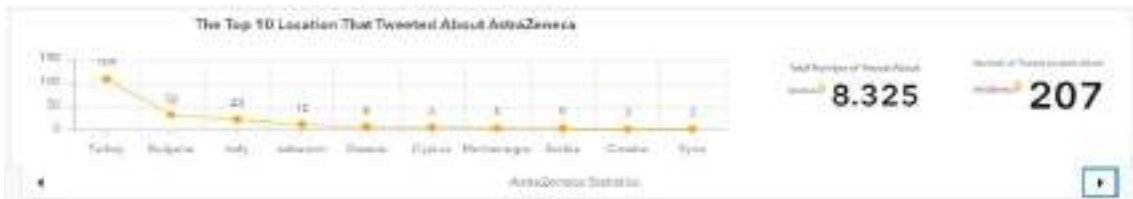
Şekil 4.44 Sinopharm aşısı hakkındaki bilgi penceresi



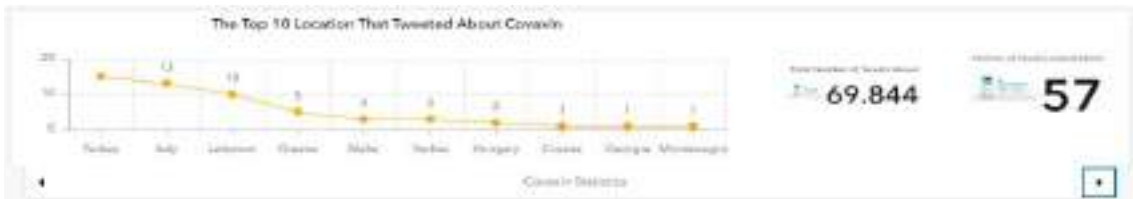
Şekil 4.45 Sinovac aşısı hakkındaki bilgi penceresi



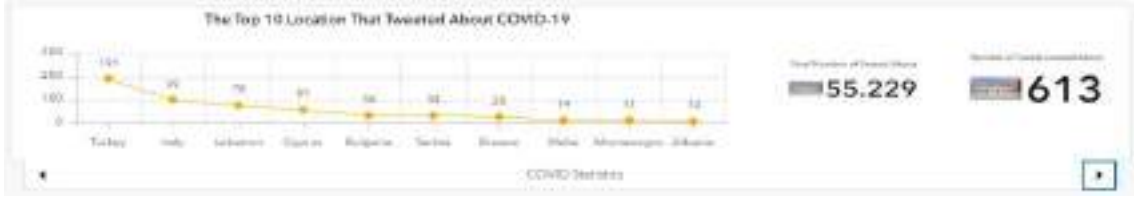
Şekil 4.46 Moderna aşısı hakkındaki bilgi penceresi



Şekil 4.47 Oxford AstraZeneca aşısı hakkındaki bilgi penceresi



Şekil 4.48 Covaxin aşısı hakkındaki bilgi penceresi



Şekil 4.49 "Covid" kelimesi içeren tweetler hakkındaki bilgi penceresi



Şekil 4.50 "Vaccine" kelimesi içeren tweetler hakkındaki bilgi penceresi

4.9.5. Ülkelerin X kullanıcı sayıları ve tweet sayılarına göre kümeleme sonuçlarını gösteren bilgi panosunun oluşturulması

Covid-19 aşılıları hakkındaki tweet verilerinden elde edilen X kullanıcı sayısı ve oluşturulan tweet sayıları bilgileri kullanılarak, ward yöntemi ve k-ortalamlar yöntemi ile gerçekleştirilen kümeleme analizlerine ait sonuçların interaktif Web haritaları ve gösterge sayaçlarıyla kullanıcıya sunumu amacıyla ArcGIS Dashboards uygulaması ile bilgi panosu oluşturulması amaçlanmıştır. Ülkelerin X kullanıcı sayıları ve tweet sayılarına göre kümeleme sonuçlarını gösteren bilgi panosuna ait taslak şema Şekil 4.51'de görülmektedir.

Şekil 4.51'de yer alan taslak şema doğrultusunda, dashboards platformunda geliştirilen bilgi paneli uygulaması Şekil 4.52'de görülmektedir. Şekil 4.52'te yer alan bölümler aşağıda açıklanmaktadır.

(1) numaralı bölüm, geliştirilen bilgi paneli uygulamasının isminin bulunduğu alandır.

(2) numaralı bölüm, ülke liste isimlerinin liste formunda yer aldığı alandır. Dinamik yapıda olup, harita ekran yayılımları ve gösterge sayaçlarıyla etkileşimli çalışmaktadır.

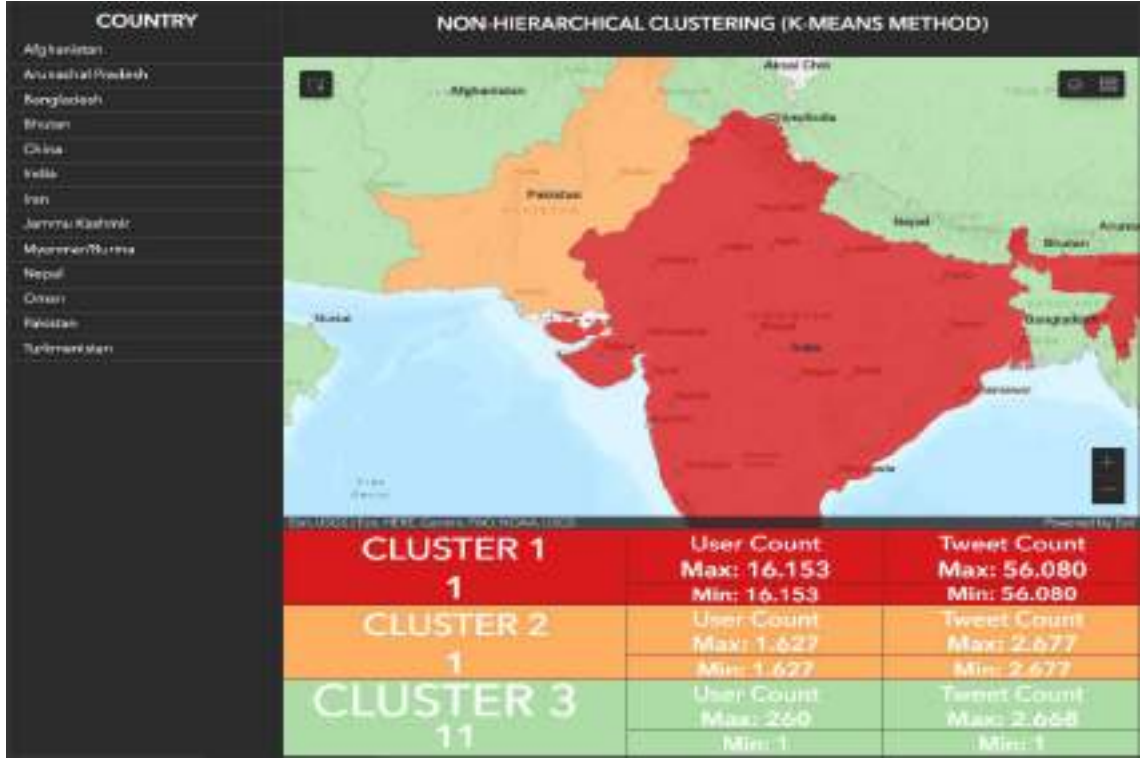
(3) numaralı bölüm, ülkelerin X kullanıcı sayıları ve tweet sayılarına göre k-ortalamlar yöntemiyle kümeleme analizinden elde edilen kümeleme sonuçlarıyla iki değişkenli koroplet Web haritasının bulunduğu alandır. Web haritası, (2) numaralı bölümde yer alan ülkeler listesi ve (5) numaralı bölümdeki k-ortalamlar yöntemine ait kümeleme sonuçlarının gösterildiği alan ile etkileşimli olarak çalışmaktadır.

UYGULAMA ADI	
ÜLKELER LİSTESİ	KÜMELEME YÖNTEMİ (K-MEANS)
	KÜMELEME YÖNTEMİ (WARD)
	HİYERARŞİK OLMAYAN KÜMELEME K-MEANS YÖNTEMİ HARİTA EKRANI
	HİYERARŞİK KÜMELEME WARD YÖNTEMİ HARİTA EKRANI
	HİYERARŞİK OLMAYAN K-MEANS YÖNTEMİYLE KÜMELEME ANALİZİ SONUÇLARININ GÖSTERİLECEĞİ GÖSTERGE SAYAÇLARI
	HİYERARŞİK KÜMELEME WARD YÖNTEMİYLE KÜMELEME ANALİZİ SONUÇLARININ GÖSTERİLECEĞİ GÖSTERGE SAYAÇLARI

Şekil 4.51 Ülkelerin X kullanıcı sayıları ve tweet sayılarına göre kümeleme sonuçlarını gösteren bilgi panosu taslak şeması

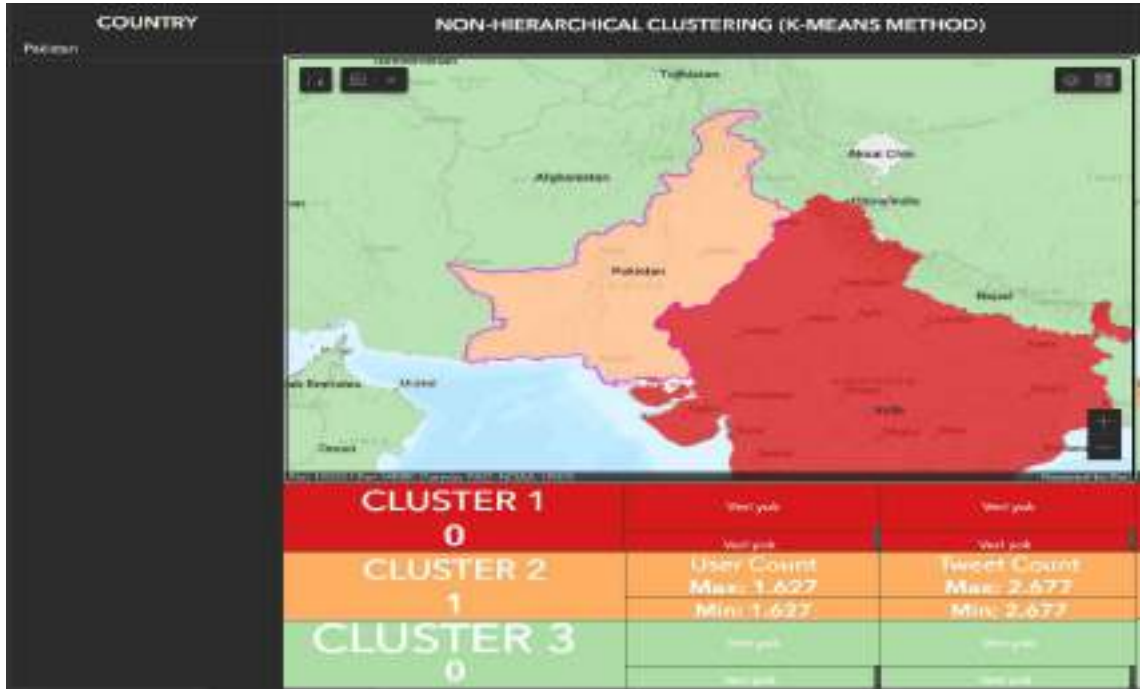
(4) numaralı bölüm, ülkelerin X kullanıcı sayıları ve tweet sayılarına göre ward yöntemiyle kümeleme analizinden elde edilen kümeleme sonuçlarıyla iki değişkenli koroplet Web haritasının bulunduğu alandır. Web haritası, (2) numaralı bölümde yer alan ülkeler listesi ve (6) numaralı bölümdeki ward yöntemine ait kümeleme sonuçlarının gösterildiği alan ile etkileşimli olarak çalışmaktadır.

(5) numaralı bölüm, ülkelerin X kullanıcı sayıları ve tweet sayılarına göre k-ortalamlar yöntemiyle kümeleme analizine ait sonuçların gösterildiği alandır. Bölüm esasen 15 adet farklı gösterge sayaçlarından oluşmaktadır. Gösterge sayaçlarının hepsi dinamik yapıda olup, (2) numaralı bölümde yer alan ülkeler listesi ve (3) numaralı bölümde yer alan harita ekran yayılımıyla etkileşimli olarak çalışmaktadır. Şekil 4.53'te harita ekran yayılım değiştirildiğinde, ülkeler listesinin ve kümeleme bilgilerinin gösteren göstergelerinde değiştiği görülmektedir.



Şekil 4.53 K-ortalamalar yöntemiyle elde edilen kümeleme sonuçlarının bilgi paneli üzerinden görünümü

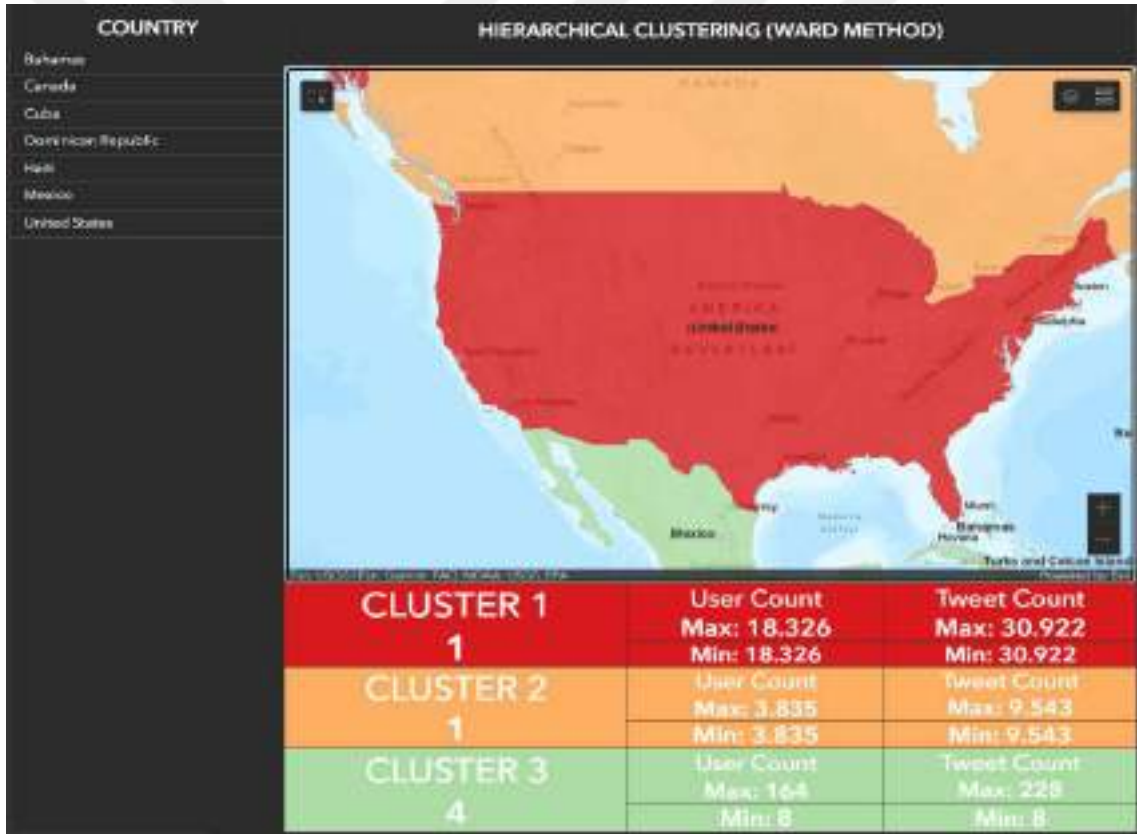
Ayrıca, (2) numaralı bölümde yer alan ülkeler listesi ya da (3) numaralı harita ekranı bölümünde yer alan çok değişkenli Web haritası üzerinden bir ülke seçildiğinde, seçilen ülkenin hangi kümede yer aldığı ve ülkenin kümelemede kullanılan öznitelik değerleri gösterge sayaçları üzerinde görüntülenmektedir (Şekil 4.54).



Şekil 4.54 Ülke bazlı K-ortalamalar yöntemiyle elde edilen kümeleme sonuçlarının bilgi paneli üzerinden görünümü

(6) numaralı bölüm, ülkelerin X kullanıcı sayıları ve tweet sayılarına göre ward yöntemiyle kümeleme analizine ait sonuçların gösterildiği alandır. Bu bölüm (5) numaralı bölü gibi 15 adet farklı gösterge sayaçlarından oluşmaktadır. Gösterge sayaçlarının hepsi dinamik yapıda olup, (2) numaralı bölümde yer alan ülkeler listesi ve (4) numaralı bölümde yer alan harita ekran yayılımıyla etkileşimli olarak çalışmaktadır. Şekil 4.55'te harita ekran yayılım değiştirildiğinde, ülkeler listesinin ve kümeleme bilgilerinin gösteren göstergelerinde değiştiği görülmektedir.

Ayrıca, (2) numaralı bölümde yer alan ülkeler listesi ya da (4) numaralı harita ekranı bölümünde yer alan çok değişkenli Web haritası üzerinden bir ülke seçildiğinde, seçilen ülkenin hangi kümede yer aldığı ve ülkenin kümelemede kullanılan öznitelik değerleri gösterge sayaçları üzerinde görüntülenmektedir (Şekil 4.56).



Şekil 4.55 Ward yöntemiyle elde edilen kümeleme sonuçlarının bilgi paneli üzerinden görünümü



Şekil 4.56 Ülke bazlı Ward yöntemiyle elde edilen kümeleme sonuçlarının bilgi paneli üzerinden görünümü

4.9.6. Ülkelerin tweet sayısı, “covid” kelimesi içeren tweet sayısı ve “vaccine” kelimesi içeren tweet sayılarına göre kümeleme sonuçlarını gösteren bilgi panosunun oluşturulması

Covid-19 aşılıları hakkındaki tweet verilerinden elde edilen toplam tweet sayıları, “covid” kelimesi içeren tweet sayıları ve “vaccine” kelimesi içeren tweet sayıları bilgileri kullanılarak, ward yöntemi ve k-ortalamlar yöntemi ile gerçekleştirilen kümeleme analizlerine ait sonuçların interaktif Web haritaları, gösterge sayaçları ve grafiklerle kullanıcıya sunumu amacıyla ArcGIS Dashboards uygulaması ile bilgi panosu oluşturulması amaçlanmıştır. Ülkelerin tweet sayıları, “covid” kelimesi içeren tweet sayıları ve “vaccine” kelimesi içeren tweet sayılarına göre kümeleme sonuçlarını gösteren bilgi panosunu tasarımında da Şekil 4.51’de görülmekte olan taslak şema kullanılmıştır. Taslak şema doğrultusunda, dashboards platformunda geliştirilen bilgi paneli uygulaması Şekil 4.57’de görülmektedir. Şekil 4.57’de yer alan bölümler aşağıda açıklanmaktadır.

(1) numaralı bölüm, geliştirilen bilgi paneli uygulamasının isminin bulunduğu alandır.

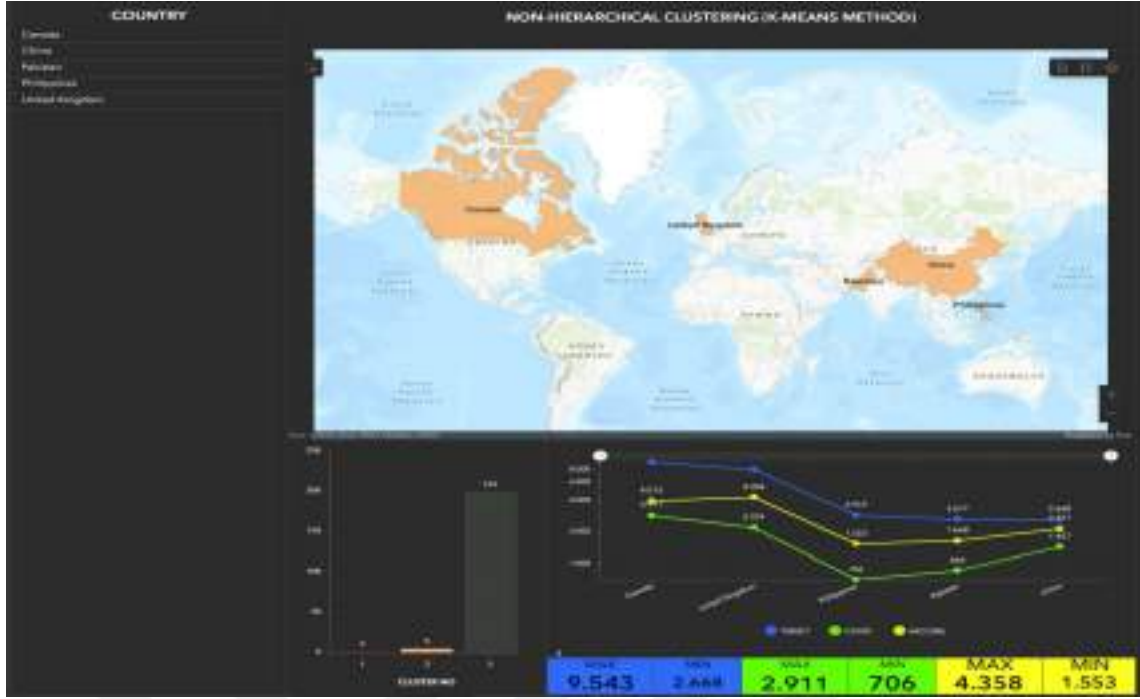
(2) numaralı bölüm, ülke liste isimlerinin liste formunda yer aldığı alandır. Dinamik yapıda olup, harita ekran yayılımları ve gösterge sayaçlarıyla etkileşimli çalışmaktadır.

(3) numaralı bölüm, ülkelerin oluşturulan toplam tweet sayısı, “covid” kelimesi içeren tweet sayısı ve “vaccine” kelimesi içeren tweet sayılarına göre k-ortalamlar yöntemiyle kümeleme analizinden elde edilen kümeleme sonuçlarıyla üç değişkenli koroplet Web haritasının bulunduğu alandır. Web haritası, (2) numaralı bölümde yer alan ülkeler listesi ve (5) numaralı bölümdeki k-ortalamlar yöntemine ait kümeleme sonuçlarının gösterildiği alan ile etkileşimli olarak çalışmaktadır.

(4) numaralı bölüm, ülkelerin oluşturulan toplam tweet sayısı, “covid” kelimesi içeren tweet sayısı ve “vaccine” kelimesi içeren tweet sayılarına göre ward yöntemiyle kümeleme analizinden elde edilen kümeleme sonuçlarıyla üç değişkenli koroplet Web haritasının bulunduğu alandır. Web haritası, (2) numaralı bölümde yer alan ülkeler listesi ve (6) numaralı bölümdeki ward yöntemine ait kümeleme sonuçlarının gösterildiği alan ile etkileşimli olarak çalışmaktadır.

(5) numaralı bölüm, ülkelerin oluşturulan toplam tweet sayısı, “covid” kelimesi içeren tweet sayısı ve “vaccine” kelimesi içeren tweet sayılarına göre k-ortalamlar yöntemiyle kümeleme analizine ait sonuçların gösterildiği alandır. Bölüm esasen 1 adet dikey sütun grafik, 1 adet çizgisel grafik ve 6 farklı gösterge sayaçlarından oluşmaktadır. Dikey sütun grafik, çizgisel grafik ve 6 farklı gösterge sayaçlarının tamamı dinamik yapıdadır. Grafikler ve gösterge sayaçları birbirleriyle, (2) numaralı bölümde yer alan ülkeler listesi ve (3) numaralı bölümde yer alan harita ekran yayılımıyla etkileşimli olarak çalışmaktadır (Şekil 4.58).

Dikey grafik sütun grafik kümeleme analizi işlemi sonucu oluşan kümeleri ve küme elemanı ülke sayısını göstermektedir. Dikey grafik üzerinde bir küme seçildiği zaman, harita ekranımı yayılımında yer alan diğer kümelere ait ülkelerin gösterimi kapatılmaktadır. Aynı zamanda ülkeler listesi seçili küme elemanı ülkeleri göstermekte, çizgisel grafik seçili küme elemanı ülkelere ait ülkelerin üç farklı öznitelik değişkenlerini göstermekte ve 6 farklı gösterge sayacı da seçili kümenin üç farklı öznitelik değişkenine ait maksimum ve minimum değerlerini göstermektedir.



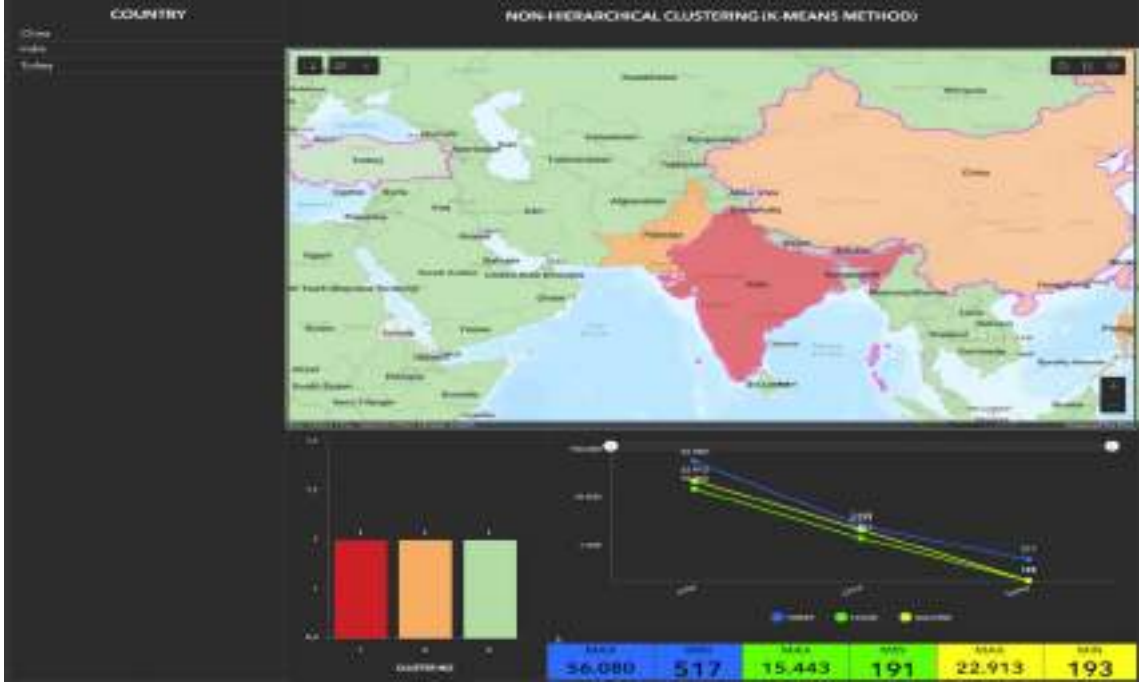
Şekil 4.58 Dikey grafik üzerinden küme seçimi gerçekleştirildiğinde ülkeler listesi, Web haritası, çizgisel grafik ve gösterge sayaçlarının etkileşimi

Çok değişkenli Web haritası üzerinde bir ülke seçildiğinde, ülkeler listesinde sadece seçilen ülkenin ismi, kümelerin gösteriminin gerçekleştirildiği dikey sütun grafik üzerinde sadece seçili ülkenin bulunduğu küme, çizgisel grafik üzerinde ve gösterge sayaçları üzerinde sadece seçili ülkeye ait üç farklı öznitelik gösterilmektedir (Şekil 4.59).



Şekil 4.59 Web haritası üzerinden bir ülke seçimi gerçekleştirildiğinde ülkeler listesi, Web haritası, çizgisel grafik ve gösterge sayaçlarının etkileşimi

Çok değişkenli Web haritası üzerinde birden fazla ülke seçimi gerçekleştirildiğinde ise; ülkeler listesinde seçilen ülkelerin ismi, kümelerin gösteriminin gerçekleştirildiği dikey sütun grafik üzerinde sadece seçili ülkenin bulunduğu kümeler ayrı ayrı, çizgisel grafik üzerinde seçili ülkelere ait üç farklı öznitelik değişkeni, gösterge sayaçları üzerinde seçili ülkelerin üç farklı öznitelik değişkenlerinin maksimum ve minimum değer aralıkları gösterilmektedir (Şekil 4.60).



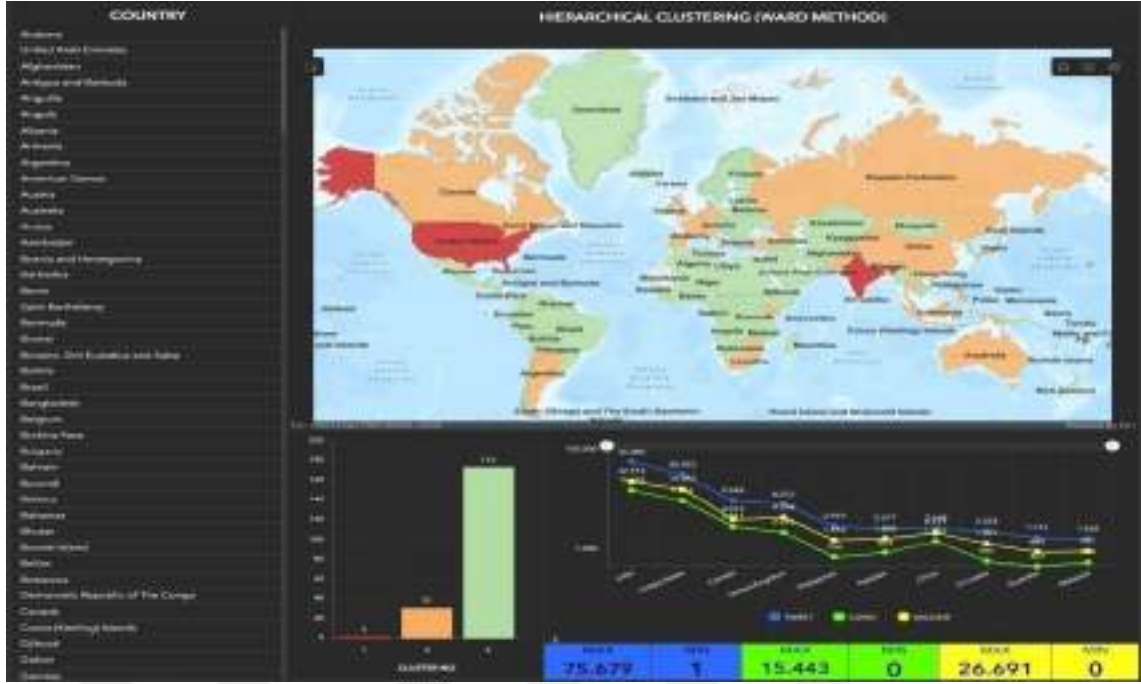
Şekil 4.60 Web haritası üzerinden birden fazla ülke seçimi gerçekleştirildiğinde ülkeler listesi, Web haritası, çizgisel grafik ve gösterge sayaçlarının etkileşimi

(6) numaralı bölüm, ülkelerin oluşturulan toplam tweet sayısı, “covid” kelimesi içeren tweet sayısı ve “vaccine” kelimesi içeren tweet sayılarına göre ward yöntemiyle kümeleme analizine ait sonuçların gösterildiği alandır. (5) numaralı içi hazırlanan grafikler ve göstergeler bu bölümde de bulunmaktadır ve çalışma prensipleri aynı şekildedir (Şekil 4.61).

4.9.7. İnteraktif veri editörünün ve bilgi panolarının birleştirilerek Web tabanlı olarak kullanıcıların erişimine açılması

Bu tez çalışması kapsamında geliştirilmesi tamamlanan;

- Covid-19 aşılı veri editörü,
- Covid-19 aşılı bilgi panosu,
- Ülkelerin X kullanıcı sayıları ve tweet sayılarına göre kümeleme sonuçlarını gösteren bilgi panosu,

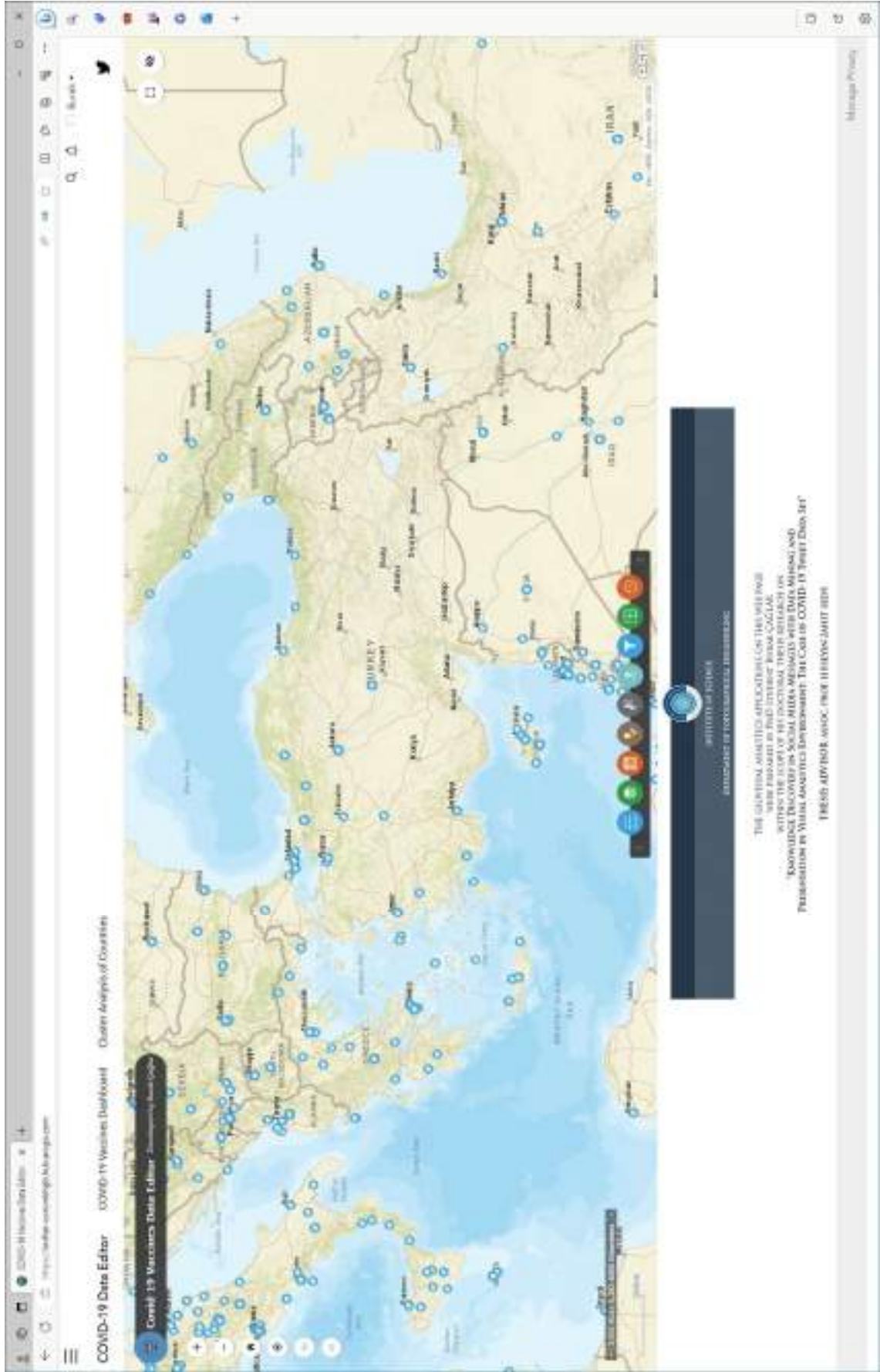


Şekil 4.61 Ülkelerin tweet sayıları, “Covid” kelimesi içeren tweet sayıları ve “Vaccine” kelimesi içeren tweet sayılarının ward yöntemiyle kümeleme analizi sonuçlarını gösteren bilgi panosu

- Ülkelerin tweet sayısı, “Covid” kelimesi içeren tweet sayısı ve “Vaccine” kelimesi içeren tweet sayılarına göre kümeleme sonuçlarını gösteren bilgi panosu,

uygulamaları ArcGIS Hub platformu kullanılarak tek bir platform altında birleştirilmiştir.

“COVID-19 All Vaccines Tweets” setinde yer alan mesajların, veri setinden keşfedilen yeni bilgilerin, kümeleme analizi sonucu elde edilen bilgiler ile üretilen çok değişkenli haritaların birbirleriyle etkileşim içinde olduğu, ön tanımlı sorgulamalar kapsamında filtreleme araçlarının yer aldığı, veri setindeki mesajların mekânsal-zamansal gelişiminin aynı ekran üzerinden takip edildiği ve sonuçların farklı grafiksel araçlarla kullanıcıya aktarıldığı Web tabanlı interaktif bir uygulama <https://twitter-corumkhgb.hub.arcgis.com/> adresi üzerinden Web tabanlı olarak erişime açılmıştır. İlgili Web sayfası üzerinden erişime açılan uygulamanın farklı sayfalarına ait görüntüler Şekil 4.62 – Şekil 4.65’te yer almaktadır.



Şekil 4.62 Covid-19 aşları veri editörü Web sayfası görünümü



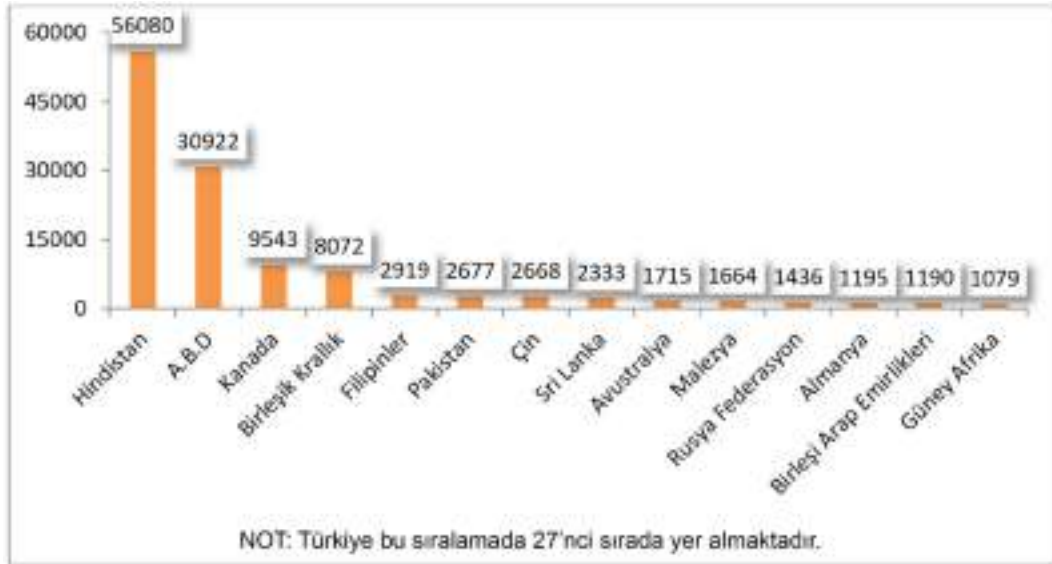
Şekil 4.63 Covid-19 aşılıları bilgi panosu Web sayfası görünümü

5. ARAŞTIRMA BULGULARI ve DEĞERLENDİRME

Bu tez çalışmasında, 2020 yılının aralık ayı itibarıyla Acil Kullanım Onayı (EUA) olarak Dünya'nın birçok ülkesinde uygulanmaya başlayan COVID-19 aşılı hakkında X sosyal medya platformunda oluşturulan tweet veri seti kullanılmıştır. COVID-19 tweet veri setinde, COVID-19 aşılı hakkında oluşturulan tweetlerin mekânsal bazlı incelenmesi, mekânsal olguların zamansal gelişiminin incelenmesi, mekânsal olguların benzer ve farklılıkların ortaya konulması amacıyla farklı konularda veri keşfi sağlanmıştır.

Veri setinde yer alan 218.412 adet tweetten 142.733 adedinin Dünya haritası üzerinde konumlandırması gerçekleştirilmiştir. 142.733 adet tweet verisinden sadece 378 adet tweetin coğrafi koordinat içerdiği tespit edilmiştir. Tüm veri seti içerisinde, konum servisi kullanılarak oluşturulan tweet sayısının 378 adet olması (%0,17) olması sosyal medya platformu kullanıcılarının konumlarını paylaşmada duydukları güvensizliği ortaya koymaktadır.

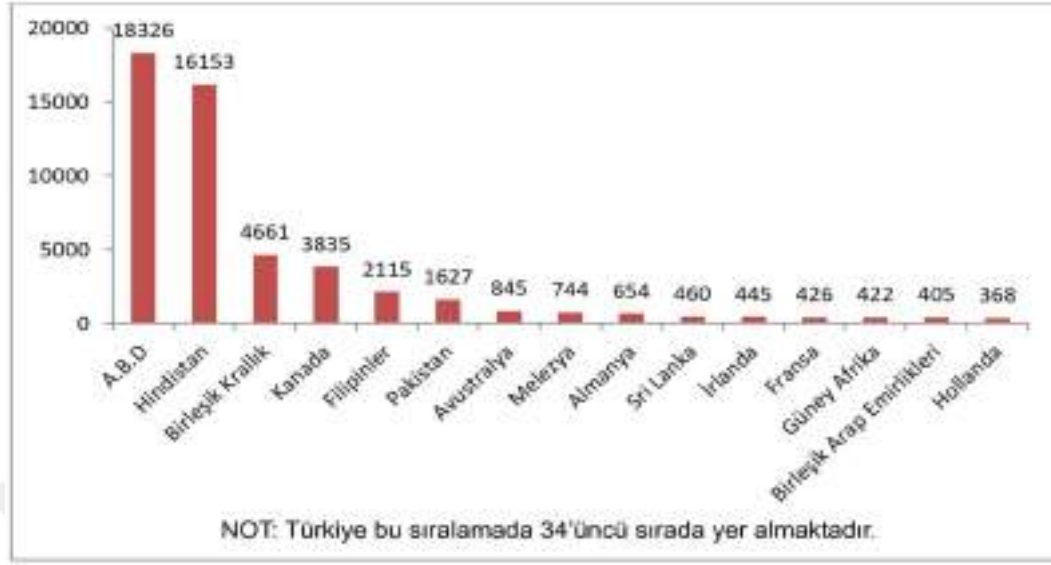
COVID-19 aşılı hakkında oluşturulan tweet sayıları ülke bazlı incelendiğinde, ilk sırada 56.080 tweet ile Hindistan, ikinci sırada 30.922 tweet ile Amerika Birleşik Devletleri ve üçüncü sırada 9.543 adet tweet ile Kanada'nın yer aldığı tespit edilmiştir. COVID-19 Aşılı hakkında tweet oluşturulan ilk 15 ülke Şekil 5.1'de gösterilmektedir.



Şekil 5.1 COVID-19 aşılı hakkında tweet oluşturulan ilk 15 ülke

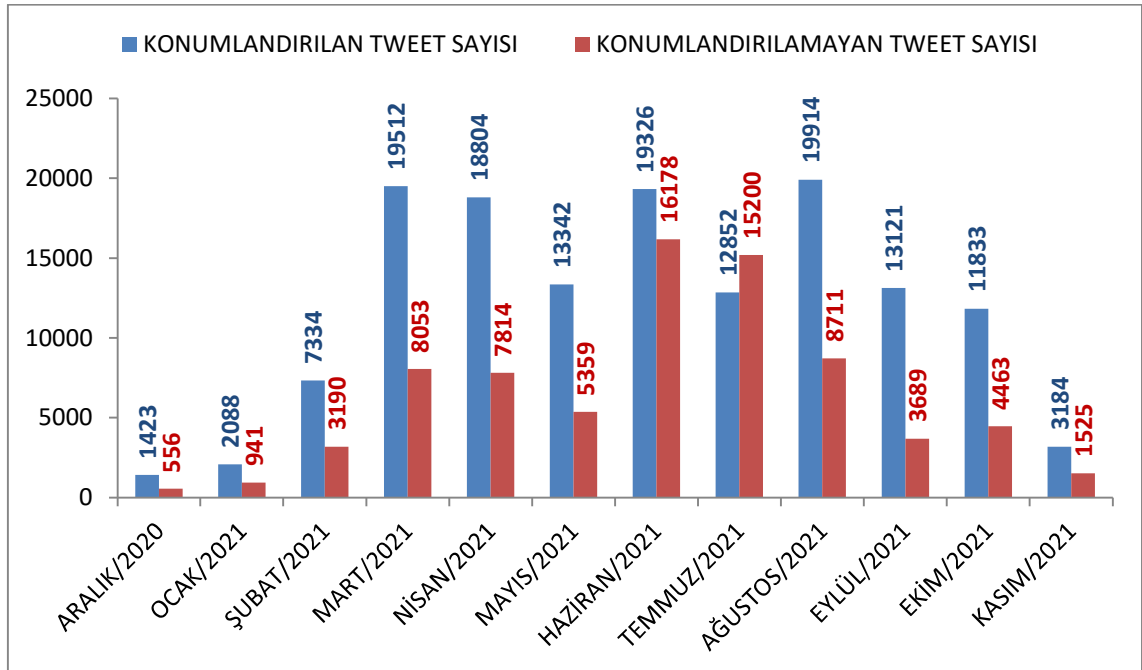
COVID-19 aşılı hakkında tweet oluşturan kullanıcı sayıları ülke bazlı incelendiğinde, ilk sırada 18.326 adet X kullanıcısı ile Amerika Birleşik Devletleri, ikinci sırada 16.153 adet X kullanıcısı ile Hindistan ve üçüncü sırada 4.661 adet X

kullanıcısı ile Birleşik Krallık yer almaktadır. COVID-19 Aşıları hakkında tweet oluşturan kullanıcılar dikkate alındığında ilk 15 ülke Şekil 5.2’de gösterilmektedir.



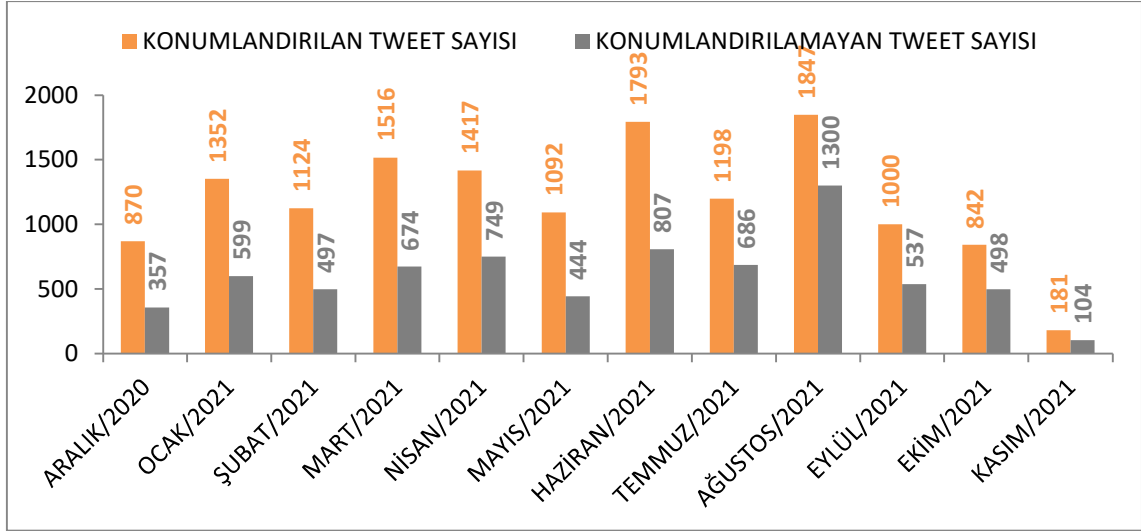
Şekil 5.2 COVID-19 aşısı hakkında tweet oluşturan kullanıcı sayıları bakımından ilk 15 ülke

COVID-19 aşısı hakkında oluşturulan tweet sayılarının zamansal dağılımları incelendiğinde, en fazla tweet'in (35.504 adet) 2021 yılı haziran ayında oluşturulduğu, en az tweet'in (1.979 adet) ise 2020 yılı aralık ayında oluşturulduğu tespit edilmiştir. COVID-19 aşısı hakkında oluşturulan tweetlerin zamansal dağılımları Şekil 5.3'te gösterilmektedir.



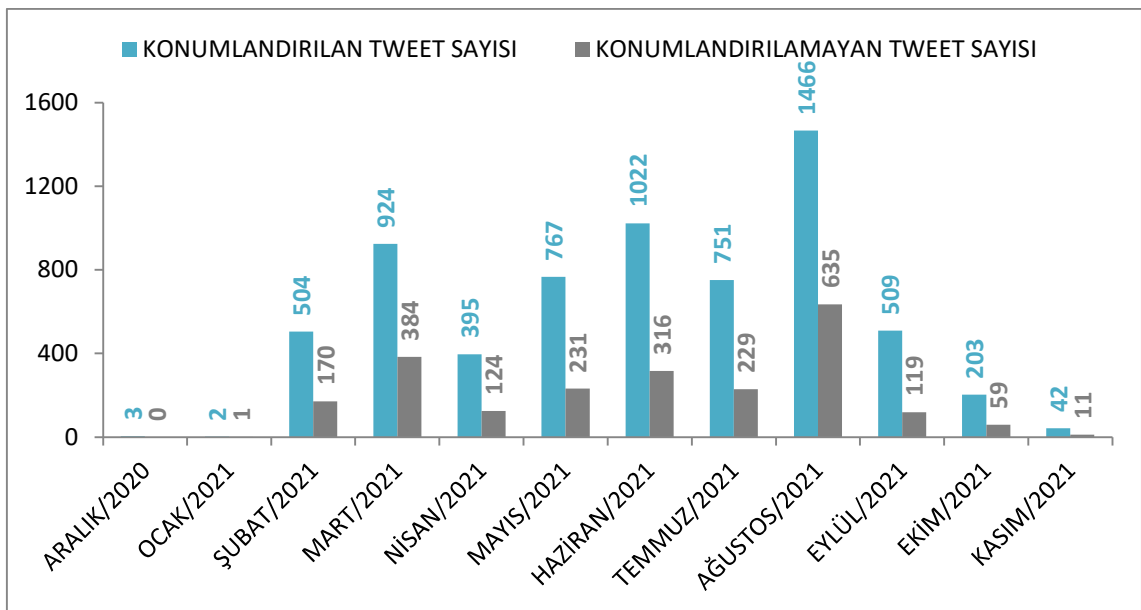
Şekil 5.3 COVID-19 aşısı hakkında oluşturulan tweetlerin zamansal dağılımları

Pfizer Biontech aşısı hakkında oluşturulan tweet sayılarının zamansal dağılımları incelendiğinde, en fazla tweet'in (3.147 adet) 2021 yılı ağustos ayında oluşturulduğu, en az tweet'in (285 adet) ise 2021 yılı kasım ayında oluşturulduğu tespit edilmiştir. Pfizer Biontech aşısı hakkında oluşturulan tweetlerin zamansal dağılımları Şekil 5.4'te gösterilmektedir.



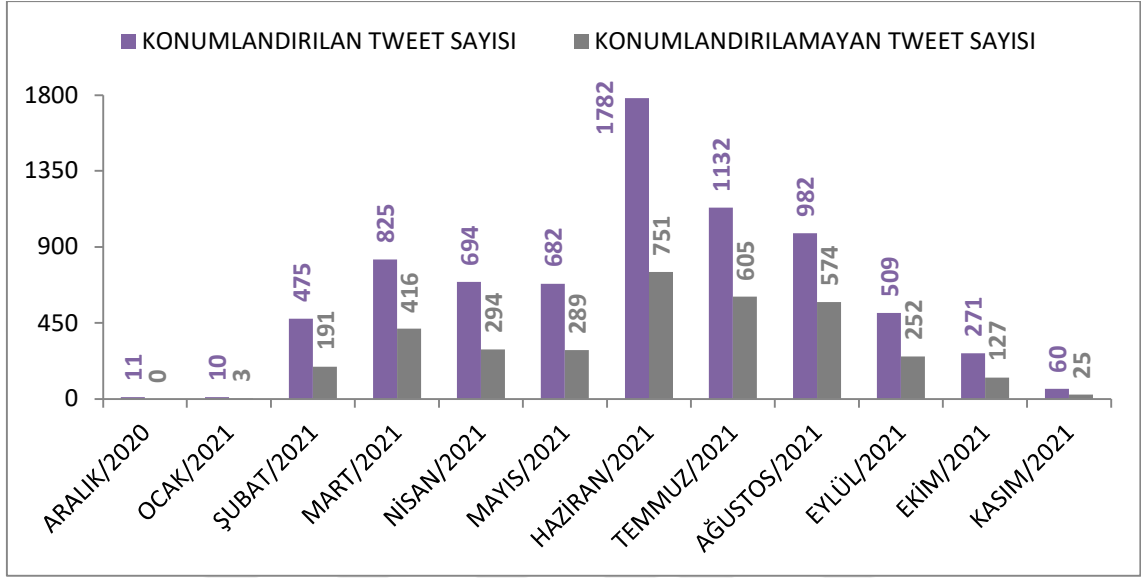
Şekil 5.4 Pfizer Biontech aşısı hakkında oluşturulan tweetlerin zamansal dağılımları

Sinopharm aşısı hakkında oluşturulan tweet sayılarının zamansal dağılımları incelendiğinde, en fazla tweet'in (2.101 adet) 2021 yılı ağustos ayında oluşturulduğu, en az tweet'in (3 adet) ise 2020 yılı aralık ve 2021 yılı ocak aylarında oluşturulduğu tespit edilmiştir. Sinopharm aşısı hakkında oluşturulan tweetlerin zamansal dağılımları Şekil 5.5'te gösterilmektedir.



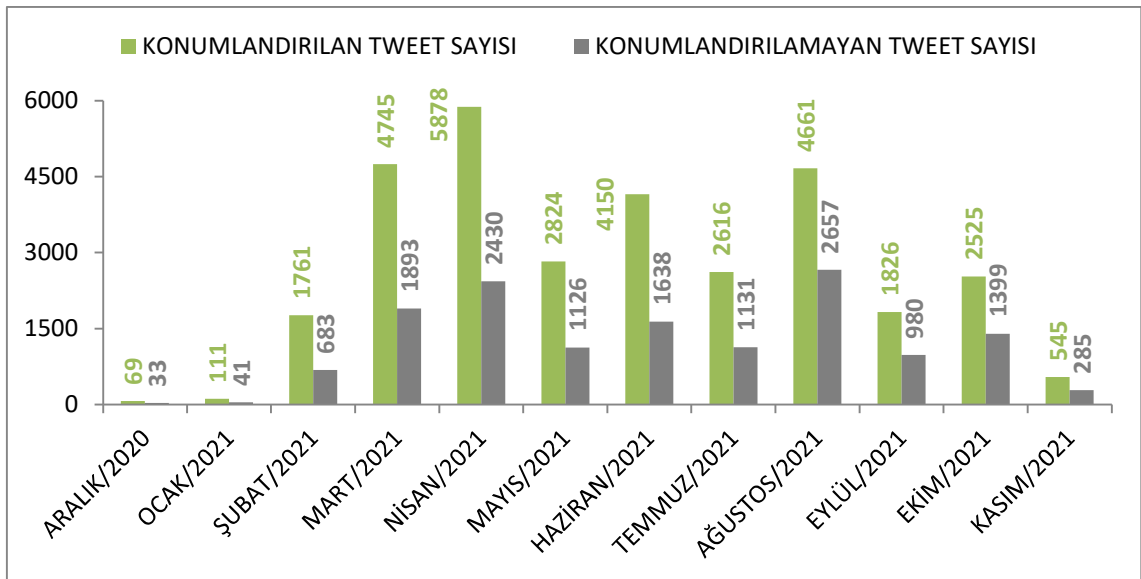
Şekil 5.5 Sinopharm aşısı hakkında oluşturulan tweetlerin zamansal dağılımları

Sinovac aşısı hakkında oluşturulan tweet sayılarının zamansal dağılımları incelendiğinde, en fazla tweet'in (2.533 adet) 2021 yılı haziran ayında oluşturulduğu, en az tweet'in (11 adet) ise 2020 yılı aralık ayında oluşturulduğu tespit edilmiştir. Sinovac aşısı hakkında oluşturulan tweetlerin zamansal dağılımları Şekil 5.6'da gösterilmektedir.



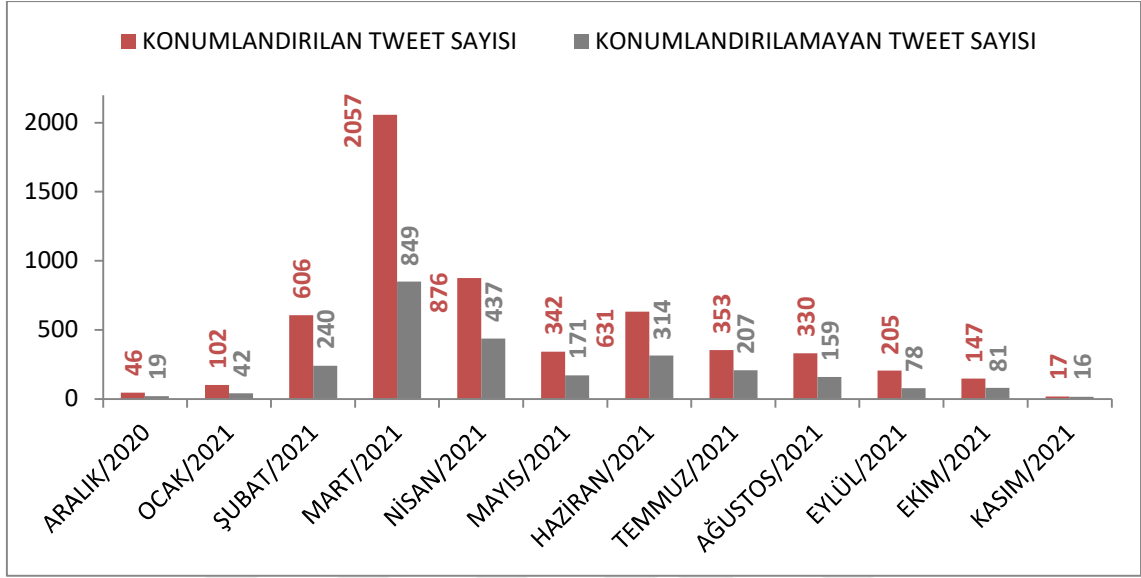
Şekil 5.6 Sinovac aşısı hakkında oluşturulan tweetlerin zamansal dağılımları

Moderna aşısı hakkında oluşturulan tweet sayılarının zamansal dağılımları incelendiğinde, en fazla tweet'in (8.308adet) 2021 yılı nisan ayında oluşturulduğu, en az tweet'in (102 adet) ise 2020 yılı aralık ayında oluşturulduğu tespit edilmiştir. Moderna aşısı hakkında oluşturulan tweetlerin zamansal dağılımları Şekil 5.7'de gösterilmektedir.



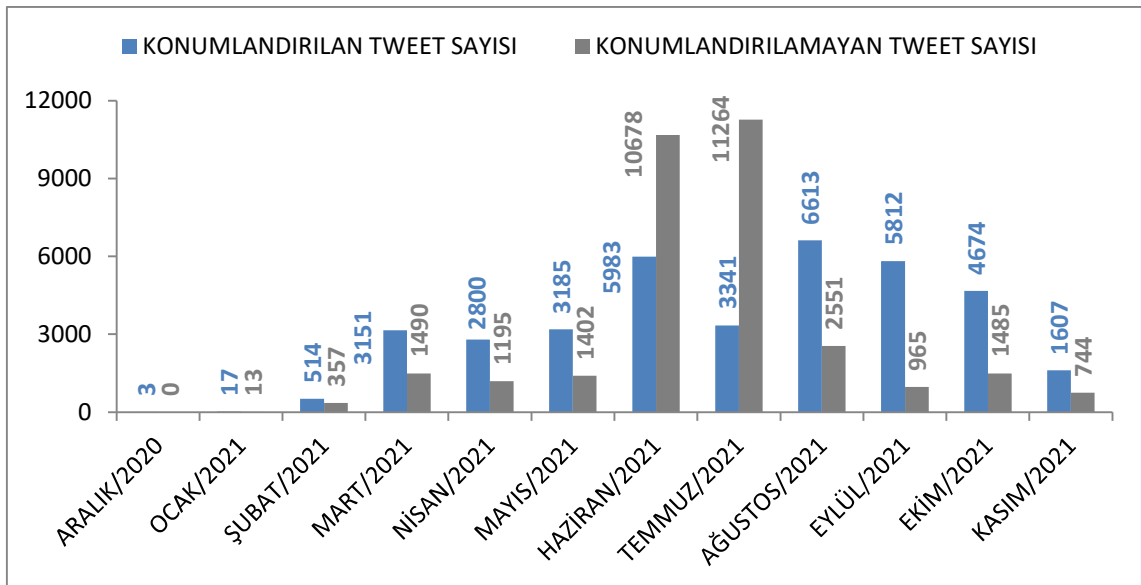
Şekil 5.7 Moderna aşısı hakkında oluşturulan tweetlerin zamansal dağılımları

Oxford AstraZeneca aşısı hakkında oluşturulan tweet sayılarının zamansal dağılımları incelendiğinde, en fazla tweet'in (2.906 adet) 2021 yılı mart ayında oluşturulduğu, en az tweet'in (33 adet) ise 2021 yılı kasım ayında oluşturulduğu tespit edilmiştir. Oxford AstraZeneca aşısı hakkında oluşturulan tweetlerin zamansal dağılımları Şekil 5.8'de gösterilmektedir.



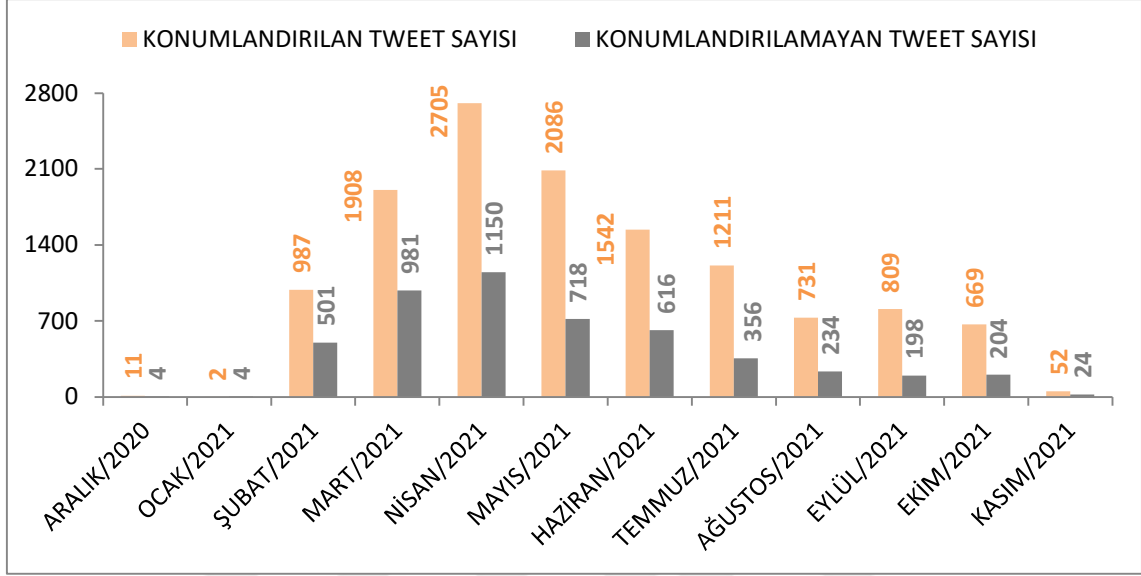
Şekil 5.8 Oxford AstraZeneca aşısı hakkında oluşturulan tweetlerin zamansal dağılımları

Covaxin aşısı hakkında oluşturulan tweet sayılarının zamansal dağılımları incelendiğinde, en fazla tweet'in (16.601 adet) 2021 yılı haziran ayında oluşturulduğu, en az tweet'in (3 adet) ise 2020 yılı aralık ayında oluşturulduğu tespit edilmiştir. Covaxin aşısı hakkında oluşturulan tweetlerin zamansal dağılımları Şekil 5.9'da gösterilmektedir.



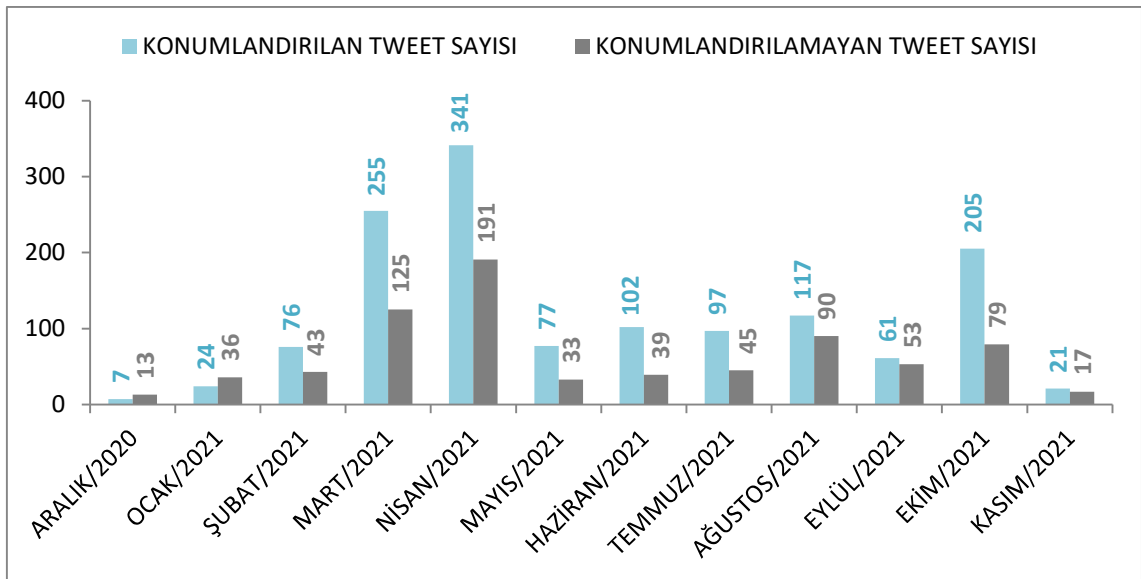
Şekil 5.9 Covaxin aşısı hakkında oluşturulan tweetlerin zamansal dağılımları

Sputnik V aşısı hakkında oluşturulan tweet sayılarının zamansal dağılımları incelendiğinde, en fazla tweet'in (3.855 adet) 2021 yılı nisan ayında oluşturulduğu, en az tweet'in (6 adet) ise 2021 yılı ocak ayında oluşturulduğu tespit edilmiştir. Sputnik V aşısı hakkında oluşturulan tweetlerin zamansal dağılımları Şekil 5.10'da gösterilmektedir.



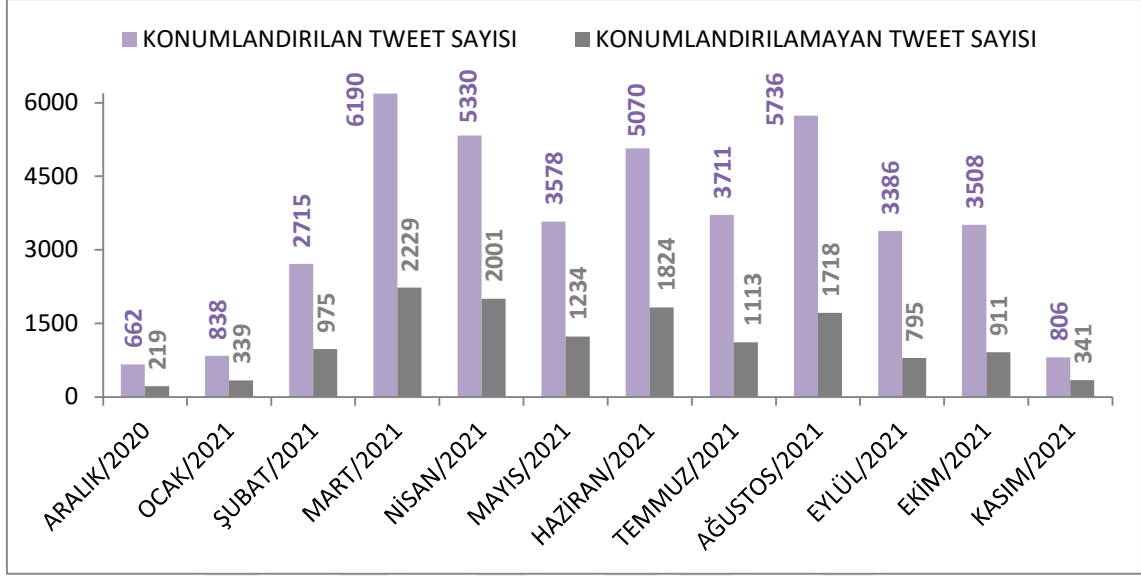
Şekil 5.10 Sputnik V aşısı hakkında oluşturulan tweetlerin zamansal dağılımları

Johnson&Johnson aşısı hakkında oluşturulan tweet sayılarının zamansal dağılımları incelendiğinde, en fazla tweet'in (532 adet) 2021 yılı nisan ayında oluşturulduğu, en az tweet'in (20 adet) ise 2020 yılı aralık ayında oluşturulduğu tespit edilmiştir. Johnson&Johnson aşısı hakkında oluşturulan tweetlerin zamansal dağılımları Şekil 5.11'de gösterilmektedir.



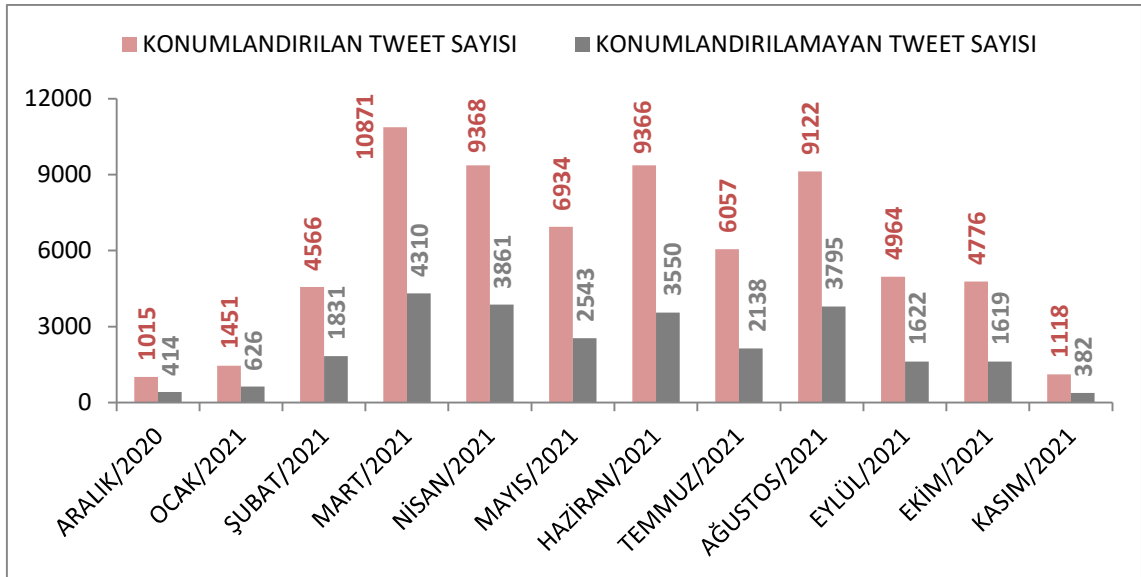
Şekil 5.11 Johnson&Johnson aşısı hakkında oluşturulan tweetlerin zamansal dağılımları

“Covid” kelimesi hakkında oluşturulan tweet sayılarının zamansal dağılımları incelendiğinde, en fazla tweet’in (8.419 adet) 2021 yılı mart ayında oluşturulduğu, en az tweet’in (881 adet) ise 2020 yılı aralık ayında oluşturulduğu tespit edilmiştir. “Covid” kelimesi hakkında oluşturulan tweetlerin zamansal dağılımları Şekil 5.12’de gösterilmektedir.



Şekil 5.12 "Covid" hakkında oluşturulan tweetlerin zamansal dağılımları

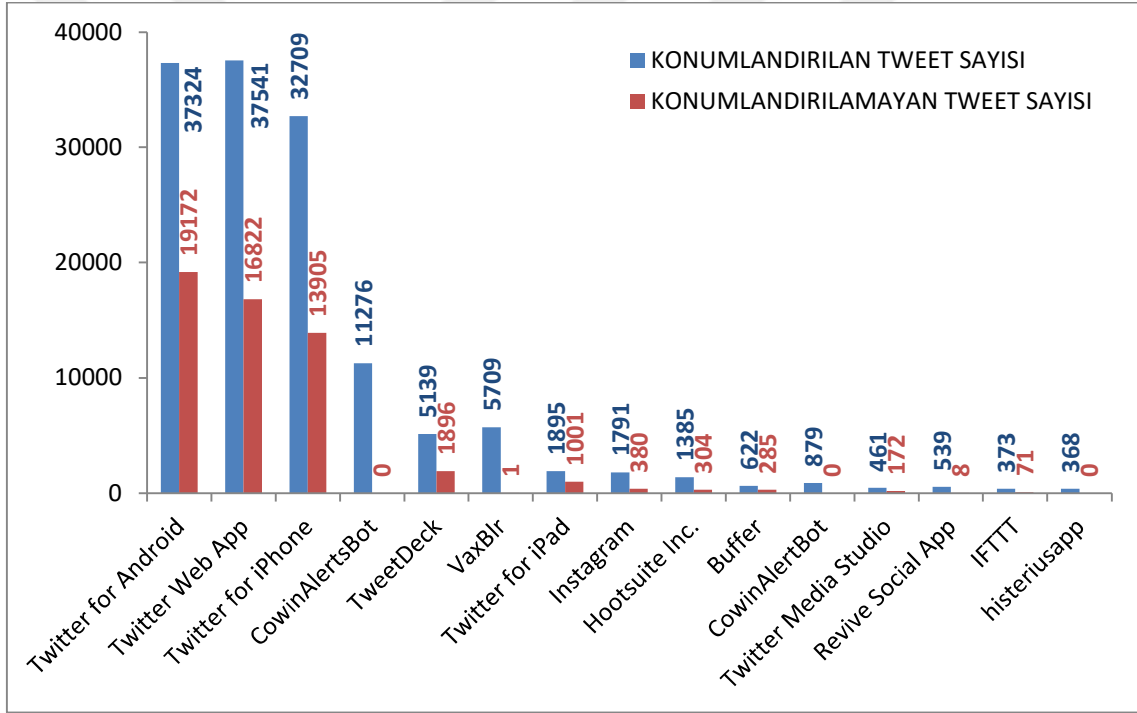
“Vaccine” kelimesi hakkında oluşturulan tweet sayılarının zamansal dağılımları incelendiğinde, en fazla tweet’in (15.181 adet) 2021 yılı mart ayında oluşturulduğu, en az tweet’in (1.429 adet) ise 2020 yılı aralık ayında oluşturulduğu tespit edilmiştir. “Vaccine” kelimesi hakkında oluşturulan tweetlerin zamansal dağılımları Şekil 5.13’te gösterilmektedir.



Şekil 5.13 "Vaccine" hakkında oluşturulan tweetlerin zamansal dağılımları

COVID-19 aşıları hakkında oluşturulan tweetlerin zamansal dağılımlarını gösteren Şekil 5.3 – Şekil 5.13’te yer alan grafikler incelendiğinde, X platformunda 2021 yılı şubat ayı itibarıyla COVID-19 aşıları hakkında oluşturulan tweetlerin artmaya başladığı, nisan – mayıs – haziran – temmuz – ağustos ayları boyunca yüksek sayılarda devam ettiği ve eylül – ekim ayları itibarıyla düşüş eğilimine geçtiği tespit edilmiştir.

COVID-19 aşıları hakkındaki tweetlerin kullanıcılara tarafından hangi uygulama platformu üzerinde oluşturulduğu incelendiğinde, ilk sırada 56.496 adet tweetle “Twitter for Android” uygulama platformu, ikinci sırada 54.363 adet tweetle “Twitter Web App” uygulama platformu, üçüncü sırada 46.614 adet tweetle “Twitter for iPhone” uygulama platformu yer almaktadır. COVID-19 aşıları hakkında tweetlerin oluşturulmasında kullanılan ilk 15 uygulama platformu Şekil 5.14’te gösterilmektedir.



Şekil 5.14 COVID-19 aşısı hakkında tweetlerin oluşturulmasında kullanılan ilk 15 uygulama platformu

Pfizer Biontech aşısı hakkında oluşturulan tweetlerin mekânsal dağılımı gösteren tek değişkenli nokta işaretli harita incelendiğinde (Ek-2.1), Pfizer Biontech aşısı hakkında oluşturulan tweetlerin ağırlıklı olarak ABD’nin doğu kısmında, Kuzeybatı Avrupa’da, Hindistan’ın kuzeydoğu bölgesinde olduğu görülmüştür.

Sinopharm aşısı hakkında oluşturulan tweetlerin mekânsal dağılımı gösteren tek değişkenli nokta işaretli harita incelendiğinde (Ek-2.2), Sinopharm aşısı hakkında oluşturulan tweetlerin ağırlıklı olarak Hindistan’ın kuzeybatı bölümünde, Sri Lanka’da, Kuzeybatı Avrupa’da ve ABD’nin doğu bölgesinde olduğu görülmüştür.

Sinovac aşısı hakkında oluşturulan tweetlerin mekânsal dağılımı gösteren tek değişkenli nokta işaretli harita incelendiğinde (Ek-2.3), Sinovac aşısı hakkında oluşturulan tweetlerin ağırlıklı olarak Uzak Doğu Asya bölgesinde, Hindistan'ın genelinde, Birleşik Krallık'ta, Kuzeybatı Avrupa'da ve ABD'nin doğu bölgesinde olduğu görülmüştür.

Moderna aşısı hakkında oluşturulan tweetlerin mekânsal dağılımı gösteren tek değişkenli nokta işaretli harita incelendiğinde (Ek-2.4), Moderna aşısı hakkında oluşturulan tweetlerin ağırlıklı olarak ABD, Avrupa'da, Hindistan'da ve Uzak Doğu Asya bölgesinde olduğu görülmüştür.

Oxford AstraZeneca aşısı hakkında oluşturulan tweetlerin mekânsal dağılımı gösteren tek değişkenli nokta işaretli harita incelendiğinde (Ek-2.5), Oxford AstraZeneca aşısı hakkında oluşturulan tweetlerin ağırlıklı olarak Avrupa'da ve ABD'nin doğu bölgesinde olduğu görülmüştür.

Covaxin aşısı hakkında oluşturulan tweetlerin mekânsal dağılımı gösteren tek değişkenli nokta işaretli harita incelendiğinde (Ek-2.6), Covaxin aşısı hakkında oluşturulan tweetlerin ağırlıklı olarak Hindistan'ın geneli, Kuzeybatı Avrupa'da ve ABD'nin doğu bölgesinde olduğu görülmüştür.

Sputnik V aşısı hakkında oluşturulan tweetlerin mekânsal dağılımı gösteren tek değişkenli nokta işaretli harita incelendiğinde (Ek-2.7), Sputnik V aşısı hakkında oluşturulan tweetlerin ağırlıklı olarak Hindistan'ın geneli, Avrupa ve ABD'nin doğu bölgesinde olduğu görülmüştür.

Johnson&Johnson aşısı hakkında oluşturulan tweetlerin mekânsal dağılımı gösteren tek değişkenli nokta işaretli harita incelendiğinde (Ek-2.8), Johnson&Johnson aşısı hakkında oluşturulan tweetlerin ağırlıklı olarak ABD'nin doğu bölgesinde ve Kuzeybatı Avrupa'da olduğu görülmüştür.

“Covid” kelimesi oluşturulan tweetlerin mekânsal dağılımı gösteren tek değişkenli nokta işaretli harita incelendiğinde (Ek-2.9), “Covid” kelimesi hakkında oluşturulan tweetlerin ağırlıklı olarak ABD, Avrupa, Hindistan ve Uzak Doğu Asya bölgesinde olduğu görülmüştür.

“Vaccine” kelimesi oluşturulan tweetlerin mekânsal dağılımı gösteren tek değişkenli nokta işaretli harita incelendiğinde (Ek-2.10), “Vaccine” kelimesi hakkında

oluşturulan tweetlerin ağırlıklı olarak Hindistan, ABD, Avrupa, Uzak Doğu Asya, Afrika kıtasının güney ve batı bölgelerinde olduğu görülmüştür.

Uber H3 hiyerarşik altıgen hücrelerle mekânsal indeksleme yöntemi kullanılarak, yerkürenin farklı konumlarında oluşturulan tweetlerin farklı çözünürlük seviyelerindeki mekânsal indeksli hücrelerle kümelendirilmesi gerçekleştirilmiştir. Üç farklı çözünürlük seviyesinde oluşturulan hiyerarşik hücreler kullanılarak, COVID-19 aşıları hakkında oluşturulan tweetlerin bulunduğu mekânsal hücreler, toplam tweet sayısı özniteliğinin doğal kırımlı değerleri esas alınarak 5 farklı sınıfta görselleştirildiği tek değişkenli koroplet haritalar hazırlanmıştır (Ek-3.1 – Ek-3.30).

Pfizer Biontech aşısı hakkında üç farklı çözünürlük seviyesi için hazırlanan Ek-3.1 – Ek-3.3'te yer alan koroplet haritalar incelendiğinde, Pfizer Biontech aşısı hakkında oluşturulan tweetlerin Amerika Birleşik Devletleri, Kanada, Avrupa kıtası ve Hindistan merkezli olduğu görülmüştür.

Sinopharm aşısı hakkında üç farklı çözünürlük seviyesi için hazırlanan Ek-3.4 – Ek-3.6'da yer alan koroplet haritaları incelendiğinde, Sinopharm aşısı hakkında oluşturulan tweetlerin Çin, Sri Lanka, Pakistan, Bangladeş ve Birleşik Arap Emirlikleri merkezli oldukları görülmüştür.

Sinovac aşısı hakkında üç farklı çözünürlük seviyesi için hazırlanan Ek-3.7 – Ek-3.9'da yer alan koroplet incelendiğinde, Sinovac aşısı hakkında oluşturulan tweetlerin Çin, Filipinler, Malezya, Tayland merkezli oldukları görülmüştür.

Moderna aşısı hakkında üç farklı çözünürlük seviyesi için hazırlanan Ek-3.10 – Ek-3.12'de yer alan koroplet incelendiğinde, Moderna aşısı hakkında oluşturulan tweetlerin Amerika Birleşik Devletleri, Kanada, Birleşik Krallık ve Hindistan merkezli oldukları görülmüştür.

Oxford AstraZeneca aşısı hakkında üç farklı çözünürlük seviyesi için hazırlanan Ek-3.13 – Ek-3.15'te yer alan koroplet haritalar incelendiğinde, Oxford AstraZeneca aşısı hakkında oluşturulan tweetlerin Birleşik Krallık, Kanada ve Hindistan merkezli oldukları görülmüştür.

Covaxin aşısı hakkında üç farklı çözünürlük seviyesi için hazırlanan Ek-3.16 – Ek-3.18'de yer alan koroplet haritalar incelendiğinde, Covaxin aşısı hakkında oluşturulan tweetlerin yüksek yoğunluklar Hindistan, daha sonra Amerika Birleşik Devletleri ve Birleşik Krallık merkezli oldukları görülmüştür.

Sputnik V aşısı hakkında üç farklı çözünürlük seviyesi için hazırlanan Ek-3.19 – Ek-3.21’de yer alan koroplet haritalar incelendiğinde, Sputnik V aşısı hakkında oluşturulan tweetlerin yüksek yoğunluklar Hindistan, daha sonra ise Rusya merkezli oldukları görülmüştür.

Johnson&Johnson aşısı hakkında üç farklı çözünürlük seviyesi için hazırlanan Ek-3.22 – Ek-3.24’te yer alan koroplet haritalar incelendiğinde, Johnson&Johnson aşısı hakkında oluşturulan tweetlerin yüksek yoğunlukla Amerika Birleşik Devletleri, daha sonra ise İngiltere merkezli oldukları görülmüştür.

“Covid” kelimesi hakkındaki tweetlerin üç farklı çözünürlük seviyesi için hazırlanan Ek-3.25 – Ek-3.27’de yer alan koroplet haritalar incelendiğinde, tweetlerin yüksek yoğunlukla Hindistan, Amerika Birleşik Devletleri ve Çin merkezli oldukları görülmüştür.

“Vaccine” kelimesi hakkındaki tweetlerin üç farklı çözünürlük seviyesi için hazırlanan Ek-3.28 – Ek-3.30’da yer alan koroplet haritalar incelendiğinde, tweetlerin yüksek yoğunlukla Hindistan, Amerika Birleşik Devletleri, İngiltere ve Çin merkezli oldukları görülmüştür.

Tweet verileri ile üretilen Ek-2.1 – Ek-2.10’da yer alan tek değişkenli nokta işaretli haritalar ile Uber H3 altıgen hücreleriyle mekânsal indeksleme sonucu oluşturulan tek değişkenli koroplet haritalar karşılaştırıldığında; bu haritalar arasındaki en büyük farkın verinin temsil biçimindeki (nokta – alan) farklılık olduğu görülmektedir. Her iki yöntemle oluşturulan haritalarda aşılardan kullanılan bölgeler hakkındaki bilgi edinilmektedir. Ancak, Uber H3 algoritması sayesinde coğrafi verilerin temsil biçiminde nokta geometrisinden alan geometrisine geçiş sağlanarak; nokta işaretli haritalarda görülmekte verilerin üst üste çakışması sonucu ortaya çıkan görsel karmaşıklık ortadan kaldırılmış, alan geometrisindeki H3 altıgen hücreleriyle verilerin en temel de sayım (miktar) özelliği olan istatistiksel analizlerle temsili ve farklı çözünürlük seviyeleri ya da aynı çözünürlük seviyesindeki farklı hücrelerle istatistiksel karşılaştırma imkânı elde edilmiştir.

Covid-19 aşılardan oluşturulan tweetlerin mekânsal dağılımını gösteren nokta işaretli haritalar ve Uber H3 yöntemiyle hiyerarşik mekânsal indeksleme sonucu elde edilen koroplet haritalardan elde edilen sonuçların desteklenmesi ve karşılaştırılması amacıyla; ülkelerin COVID-19 pandemi döneminde salgınla mücadele

kapsamında kullandıkları aşılama bilgilerini gösteren veri seti kullanılarak, sekiz farklı aşı markasının hangi ülkeler tarafından kullanıldığını gösteren koroplet haritalar hazırlanmıştır (Ek-4.1 – Ek-4.8).

Pfizer Biontech aşısının ülkelere göre kullanım durumunu gösteren Ek-4.1’de gösterilen koroplet harita, Ek-2.1’de yer alan nokta işaretli harita ve Ek-3.1 – Ek-3.3’te yer alan koroplet haritalarla karşılaştırıldığında; Pfizer Biontech aşısı hakkında oluşturulan tweetlerin Amerika Birleşik Devletleri, Kanada, Avrupa kıtası merkezli oldukları görülmüştür. Hindistan’da Pfizer Biontech aşısının kullanımı olmamasına rağmen, Pfizer Biontech aşısı hakkında oluşturulan tweetlerin olduğu tespit edilmiştir. Hindistan haricinde elde edilen haritaların birbirleriyle tutarlı olduğu belirlenmiştir.

Sinopharm aşısının ülkelere göre kullanım durumunu gösteren Ek-4.2’de gösterilen koroplet harita, Ek-2.2’de yer alan nokta işaretli harita ve Ek-3.4 – Ek-3.6’da yer alan koroplet haritalarla karşılaştırıldığında; Sinopharm aşısı hakkında oluşturulan tweetlerin Çin, Sri Lanka, Pakistan, Bangladeş ve Birleşik Arap Emirlikleri merkezli oldukları tespit edilmiştir. Elde edilen bilgiler ve üretilen haritaların birbirleriyle tutarlı olduğu belirlenmiştir.

Sinovac aşısının ülkelere göre kullanım durumunu gösteren Ek-4.3’te gösterilen koroplet harita, Ek-2.3’te yer alan nokta işaretli harita ve Ek-3.7 – Ek-3.9’da yer alan koroplet haritalarla karşılaştırıldığında; Sinovac aşısı hakkında oluşturulan tweetlerin Çin, Filipinler, Malezya, Tayland merkezli oldukları tespit edilmiştir. Elde edilen bilgiler ve üretilen haritaların birbirleriyle tutarlı olduğu belirlenmiştir.

Moderna aşısının ülkelere göre kullanım durumunu gösteren Ek-4.4’te gösterilen koroplet harita, Ek-2.4’te yer alan nokta işaretli harita ve Ek-3.10 – Ek-3.12’de yer alan koroplet haritalarla karşılaştırıldığında; Moderna aşısı hakkında oluşturulan tweetlerin Amerika Birleşik Devletleri, Kanada ve Birleşik Krallık merkezli oldukları tespit edilmiştir. Hindistan’da Moderna aşısının kullanımı olmamasına rağmen, Pfizer Biontech aşısı hakkında atılan tweetlerin olduğu tespit edilmiştir. Hindistan haricinde elde edilen haritaların birbirleriyle tutarlı olduğu belirlenmiştir.

Oxford AstraZeneca aşısının ülkelere göre kullanım durumunu gösteren Ek-4.5’te gösterilen koroplet harita, Ek-2.5’te yer alan nokta işaretli harita ve Ek-3.13 – Ek-3.15’te yer alan koroplet haritalarla karşılaştırıldığında; Oxford AstraZeneca aşısı hakkında oluşturulan tweetlerin Birleşik Krallık, Kanada ve Hindistan merkezli

oldukları tespit edilmiştir. Elde edilen bilgiler ve üretilen haritaların birbirleriyle tutarlı olduğu belirlenmiştir.

Covaxin aşısının ülkelere göre kullanım durumunu gösteren Ek-4.6'da gösterilen koroplet harita, Ek-2.6'da yer alan nokta işaretli harita ve Ek-3.16 – Ek-3.18'de yer alan koroplet haritalarla karşılaştırıldığında; Covaxin aşısı hakkında oluşturulan tweetlerin yüksek yoğunlukla Hindistan merkezli olduğu tespit edilmiştir. ABD ve Birleşik Krallık'ta Covaxin aşısının kullanımı olmamasına rağmen, Covaxin aşısı hakkında atılan tweetlerin olduğu görülmüştür. ABD ve Birleşik Krallık haricinde elde edilen haritaların birbirleriyle tutarlı olduğu belirlenmiştir.

Sputnik V aşısının ülkelere göre kullanım durumunu gösteren Ek-4.7'de gösterilen koroplet harita, Ek-2.7'de yer alan nokta işaretli harita ve Ek-3.19 – Ek-3.21'de yer alan koroplet haritalarla karşılaştırıldığında; Sputnik V aşısı hakkında oluşturulan tweetlerin yüksek yoğunlukla Hindistan, daha sonra ise Rusya merkezli oldukları tespit edilmiştir. Elde edilen bilgiler ve üretilen haritaların birbirleriyle tutarlı olduğu belirlenmiştir.

Johnson&Johnson aşısının ülkelere göre kullanım durumunu gösteren Ek-4.8'de gösterilen koroplet harita, Ek-2.8'de yer alan nokta işaretli harita ve Ek-3.22 – Ek-3.24'te yer alan koroplet haritalarla karşılaştırıldığında; Johnson&Johnson aşısı hakkında oluşturulan tweetlerin yüksek yoğunlukla Amerika Birleşik Devletleri, daha sonra ise İngiltere merkezli oldukları tespit edilmiştir. Elde edilen bilgiler ve üretilen haritaların birbirleriyle tutarlı olduğu belirlenmiştir.

Ülkelerin k-ortalamalar yöntemiyle tweet sayıları ve kullanıcı sayıları değişkenlerine göre kümeleme işlemi sonucunda 3 farklı küme oluşmuştur. ABD ve Hindistan (2 ülke) yüksek seviyede benzerlik göstererek 1 nolu kümeyi, Birleşik Krallık – Filipinler – Kanada – Pakistan (4 ülke) orta seviyede benzerlik göstererek 2 nolu kümeyi ve diğer ülkeler (201 ülke) düşük seviyede benzerlik göstererek 3 nolu kümeyi oluşturmuşlardır.

Ülkelerin ward yöntemiyle tweet sayıları ve kullanıcı sayıları değişkenlerine göre kümeleme işlemi sonucunda 3 farklı küme oluşmuştur. ABD ve Hindistan (2 ülke) yüksek seviyede benzerlik göstererek 1 nolu kümeyi, Almanya - Arjantin - Avustralya - Bangladeş - Belçika - Birleşik Arap Emirlikleri - Birleşik Krallık - Çin - Filipinler - Fransa - Güney Africa - Hollanda - İrlanda - İsviçre - İtalya - Jammu Keşmir - Japonya -

Kanada - Kenya - Malezya - Pakistan - Rusya - Singapur - Sri Lanka – Tayland (25 ülke) orta seviyede benzerlik göstererek 2 nolu kümeyi ve diğer ülkeler (180 ülke) düşük seviyede benzerlik göstererek 3 nolu kümeyi oluşturmuşlardır.

Ülkelerin k-ortalamalar yöntemiyle tweet sayıları, “covid” kelimesi hakkındaki tweet sayıları ve “vaccine” kelimesi hakkındaki tweet sayıları değişkenlerine göre kümeleme işlemi sonucunda 3 farklı küme oluşmuştur. ABD ve Hindistan (2 ülke) yüksek seviyede benzerlik göstererek 1 nolu kümeyi, Birleşik Krallık – Çin – Filipinler – Kanada – Pakistan (5 ülke) orta seviyede benzerlik göstererek 2 nolu kümeyi ve diğer ülkeler (200 ülke) düşük seviyede benzerlik göstererek 3 nolu kümeyi oluşturmuşlardır.

Ülkelerin ward yöntemiyle tweet sayıları, “covid” kelimesi hakkındaki tweet sayıları ve “vaccine” kelimesi hakkındaki tweet sayıları değişkenlerine göre kümeleme işlemi sonucunda 3 farklı küme oluşmuştur. ABD ve Hindistan (2 ülke) yüksek seviyede benzerlik göstererek 1 nolu kümeyi, Almanya - Arjantin - Avustralya - Bangladeş - Belçika - Birleşik Arap Emirlikleri - Birleşik Krallık - Çin - Endonezya - Filipinler - Fransa - Güney Afrika - Hollanda - Hong Kong - İrlanda - İspanya - İsviçre - İtalya - Jammu Keşmir - Japonya - Kanada - Katar - Kenya - Malezya - Pakistan - Rusya - Singapur - Sri Lanka - Tayland - Türkiye - Zimbabve (31 ülke) orta seviyede benzerlik göstererek 2 nolu kümeyi ve diğer ülkeler (174 ülke) düşük seviyede benzerlik göstererek 3 nolu kümeyi oluşturmuşlardır.

COVID-19 aşılıları hakkındaki tweet veri setinde gerçekleştirilen coğrafi kodlama, veri setinden bulunan gizli verilerin keşfi, mekânsal ve özneliksel analizler, tweet verilerinin Uber H3 yöntemiyle mekânsal indekslenmesi, k-ortalamalar ve ward yöntemleri kullanılarak kümeleme analizi işlemlerinden elde edilen bilgiler kullanılarak; görüntüleyicinin bir görüntüleme ekranına sahip masaüstü ya da mobil cihaz (tablet ve akıllı telefon) üzerinden herhangi bir işletim sistemi kullanarak (Windows, Android, iOS) erişim sağladığı Web tabanlı bir mekânsal görsel analitik uygulama örneği oluşturulmuştur.

Bu tez çalışması sonucu oluşturulan mekânsal görsel analitik uygulamasını klasik haritalama ile karşılaştırarak değerlendirildiğinde;

Klasik haritacılık yöntemleriyle tweet verilerinin coğrafi dağılımlarını açık ve anlaşılır temsil eden statik haritalar ile son kullanıcıya sunmak yerine; geliştirilen mekânsal görsel analitik uygulaması ile, tweet verilerinin basitçe temsil edilmesinin

ötesine geçerek, verilerdeki saklı örüntülerin ve ilişkilerin keşfi sağlanarak, sonuçlar kullanıcıya dinamik ve etkileşimli haritalarla sunulmuştur.

Veri boyutu ve karmaşıklığı açısından, klasik haritalama daha az sayıdaki coğrafi verileri ve mekânsal ilişkileri görselleştirme için uygundur. Mekânsal görsel analitik uygulaması ile, büyük veri niteliğinde olan yaklaşık 218 bin tweet işlenmek ve analiz edilmek üzere mekânsal, zamansal ve öznitelik bilgileriyle sisteme entegre edilerek son kullanıcıya sunulmuştur.

Son kullanıcının etkileşimi açısından, klasik haritalama uygulamaları genellikle statiktir ve harita oluşturucusunun belirlediği özellik dahilinde etkileşime girebilirler. Mekânsal görsel analitik uygulamasında ise kullanıcının etkileşimi dinamik bir özelliktedir. Son kullanıcılar verilerle etkileşime girebilmekte, öznitelikleri değiştirebilmekte, interaktif haritaları yakınlaştırıp uzaklaştırabilmekte verilere ait istatistiksel bilgileri grafikler ve gösterge sayaçları üzerinden görüntüleyebilmekte, istenilen katmanları görünürlüğünü kontrol edebilmekte, bilgileri özniteliksel, mekânsal ve zamansal olarak filtreleyebilmektedirler.

6. SONUÇLAR ve ÖNERİLER

Günümüzde sosyal medya platformları, sahip oldukları milyarlarca kullanıcı tarafından hayatlarının her anına ilişkin paylaşılan mesajlar, resimler, videolar, dokümanlar vb. içeriklerden oluşan çok büyük miktardaki veriyi kendi veri tabanlarında barındıkları için, büyük veri alanında önemli bir kaynak olarak kabul edilmektedirler. Pazar araştırma analistleri, işletmeler, kurum ve kuruluşlar Dünya genelinde geniş ve çeşitli kullanıcı kitlesine sahip olan sosyal medya platformlarında paylaşılan verileri kullanarak hedef kitleleri hakkında bilgi edinmektedirler. Sosyal medya platformlarındaki verilerin gerçek zamanlı olması, olayların çok daha hızlı takip edilmesini sağlamaktadır. Vaka analizi, kriz yönetimi gibi hızlı tepki gerektiren durumlarda gerçek zamanlı veriler büyük avantaj sağlamaktadır.

Bu tez çalışmasında, X (Eski adıyla Twitter) isimli sosyal medya platformunda COVID-19 pandemi döneminde salgınla mücadelede kullanılan aşılarda hakkında oluşturulan tweet veri setinden; coğrafi kodlama, mekânsal ve özniteliksiz analizler, kümeleme analizi yöntemleri kullanılarak bilgi çıkarımları gerçekleştirilmiştir. Tweet verilerinin ve analiz sonuçlarının, mekânsal-zamansal gelişiminin aynı ekran üzerinde etkileşimli ve dinamik olarak çalışan Web haritaları, sorgu ve filtreleme araçları, grafikler ve gösterge sayaçları üzerinde Web tabanlı olarak görüntüleyiciye aktarıldığı interaktif mekânsal görsel analitik uygulaması geliştirilmiştir.

Öncelikli olarak bir araştırma kapsamında X sosyal medya platformunda oluşturulan tweetlerin kullanılması planlanıyor ise, mutlaka öncelikli olarak X Developer (Geliştirici) hesabı edinilerek, developer hesabına özgü tanımlanan “*API key*”, “*API secret key*”, “*Access Token*” ve “*Access Token Secret*” anahtarların edinilmesi gerekmektedir. X firmasının özellikle COVID-19 pandemi döneminde araştırmacıların tweet verileri üzerinde gerçekleştirmek istedikleri yoğun analiz isteği karşısında, X Developer hesap başvurularını “*Your Twitter developer account application was not approved. Thank you for your interest in developer access. We are unable to serve your use case at this time. Thank you for your interest in building on Twitter.*” şeklinde hiçbir sebep göstermeksizin reddetmesi ve tekrar hesap başvurusuna kapatması bu tez çalışmasında da karşılaşılan en büyük sorun olmuştur. Ancak Gabriel Preda tarafından Kaggle platformunda “CC0:Public Domain” lisansı ile açık veri olarak sunduğu “COVID-19 All Vaccines Tweets” veri seti, bu tez çalışması kapsamında gerçekleştirilecek araştırma için temel veri seti olarak kullanılmıştır.

Tweet mesajlarının mekânsal takibinin yapılabilmesi için, tweetlerin Bölüm 2.10.1’de açıklanan mekânsal göstergeleri kullanarak konum çıkarımlarının gerçekleştirilmesi gerekmektedir. Bu tez çalışmasında, tweetlerin konum çıkarımında mekânsal gösterge olarak “user_location” öznitelik alanı kullanılmıştır. Tweet veri setinde yer alan 218.412 adet tweet verisinden, sadece 378 adet tweetin (yaklaşık yüzde 0,17) coğrafi koordinat içerdiği tespit edilmiştir. Bu sonuç, X kullanıcılarının mobil cihazlarında tweet oluştururken konum tabanlı hizmetler servisine izin vermediğinin göstergesidir. Tweet veri setindeki 142.733 adet (yaklaşık yüzde 65) tweet verisinden konum çıkarımı gerçekleştirilerek harita üzerinden işaretlenmesi gerçekleştirilmiştir. Ancak bu işlemde karşılaşılan diğer bir sorun ise, kullanıcıların konum bilgilerini tanımladıkları metinsel alana birden fazla konum belirteci tanımlamış olmaları sebebiyle hangi konumun kullanıcının gerçek konumu olduğunun tahmin edilememesidir. X firmasının, kendi platformunda paylaşılan verilerin gerçek zamanlı veriler olmasından dolayı özellikle vaka analizi ve kriz yönetimi gibi önemli alanlarda temel veri kaynaklarından biri olabileceğini dikkate alması gerekmektedir. Bu kapsamda, kullanıcı profil tanımlama kısmında yer alan konum bilgisini metinsel tanımlamaya kapatarak, en düşük hassasiyetli konum verisinin kullanıcının sabit/mobil cihazı üzerinden elde ettiği IP sunucu adresinin ya da bu konuma yakın kesin adres tanımlamalarının açılır liste üzerinden kullanıcıya sunulması seçiminin sağlanması faydalı olacaktır. Ya da X uygulamasının kullanımında, kullanıcıya konum tabanlı hizmetler servisinin zorunlu kullanım izninin sağlanması gerektiği değerlendirilmektedir.

Tweet verileri harita üzerinde işaretlendiğinde, konumsal verilerinin birçoğunun birbirine çok yakın ya da aynı konum bilgisini içeren veriler olduğu gözlemlenmiştir. Mekânsal analizlere bakıldığında, verilerin genellikle haritalar ve öznitelik tabloları üzerinde bir araya getirilerek kullanıldığı görülmektedir. Bu çalışmada, verileri bir araya getirerek mekânsal analizler gerçekleştirmek amacıyla, nokta geometrisinden alan geometrisine dönüşüm sağlamak üzere Convex Hull (Dışbükey Kabuk) algoritması kullanılmıştır. Tweet verilerinin Convex Hull algoritmasıyla nokta geometrisinden alan geometrisine dönüşümünde tüm verilerin için ortak parametre “Ülke adı” verisi kullanılmıştır. Ancak “Ülke adı” parametresinin düşük hassasiyetli bir konum belirteci olması nedeniyle Convex Hull algoritmasında beklenen başarı elde edilememiştir. Bunun üzerine tweet verilerini nokta geometrisinden alan geometrisine dönüştürmek

için Uber H3 algoritması kullanılmıştır. Uber H3 algoritması kullanılarak, tweet verilerinin bulunduğu bölgelerde üç farklı çözünürlük seviyesinde mekânsal indeks özellikli altıgen hücreler üretilmiştir. Üç farklı çözünürlük seviyesinde üretilen altıgen hücreler içerisinde yer alan tweet verileri özniteliksel olarak birleştirilerek H3 altıgen hücreleriyle ilişkilendirilmiştir. Böylelikle tweet verilerinin temsil ettiği özniteliksel özellikleri bir araya getirilerek farklı çözünürlük seviyesindeki hiyerarşik altıgen hücrelerle temsili sağlanmıştır. Uber H3 algoritmasının, mekânsal verileri maksimum çözünürlük seviyesi olan 15'inci seviyede kenar uzunluğu ortalama 0,60 m ve ortalama alanı 0,90 m² olan altıgen hücrelerle indeksleme özelliği gerçekten göz ardı edilemeyecek niteliktedir. Özellikle hücreler arasında olan aile-çocuk (parent-child) ilişki modeli mekânsal analizlere sağlayacağı avantajlar açısından büyük önem taşımaktadır.

Tweet verilerinin noktasal temsilinden Uber H3 altıgen hücreleri kullanılarak alansal olarak temsiline geçişte mekânsal analizler açısından önemli kazanımlar elde edilmektedir. Mekânsal ilişki yönünden bakıldığında, nokta verileri genellikle tek başına mekânsal örüntüler ve ilişkiler hakkında kapsamlı bir anlayış sağlamamaktadır. Noktaların alanlar halinde toplanması daha geniş bir bakış açısı sağlamakta ve indekslenmiş bölgelerdeki mekânsal dağılımların anlaşılmasına yardımcı olmaktadır. Uber H3 altıgen alanlarındaki tek tiplilik, hesaplamaları basitleştirdiği ve mekânsal ilişkileri tutarlı bir şekilde temsil etmesi açısından analiz ve görselleştirme yöntemleri açısından kolaylık sağlamaktadır. Verilerin alanlara toplanması, alansal birimler için daha uygun olan istatistiksel yöntemlerin uygulanmasına olanak tanımaktadır. Bu sayede, kümeleme analizi ve mekânsal regresyon gibi veri madenciliği araçlarının kullanılmasını sağlamaktadır. Alan verileri genellikle nokta verilerine göre görsel olarak daha sezgiseldir. Alan verileriyle oluşturulan haritalar, mekânsal örüntülerin daha geniş bir kitleye daha net bir şekilde iletilmesini sağlayarak daha iyi anlama ve karar vermeyi kolaylaştırmaktadır. Birçok mekânsal veri kümesi doğal olarak alansal düzeyde düzenlenmektedir. Nokta verilerinin alan verilerine dönüştürülmesi, diğer ilgili veri setleriyle entegrasyonu kolaylaştırarak mekânsal ilişkilerin ve örüntülerin daha kapsamlı bir şekilde analiz edilmesini sağlamaktadır. Alan düzeyinde analiz, karar vericilere eyleme geçirilebilir öngörüler sağlamaktadır. İster demografik örüntüleri ister suç noktalarını veya hastalık yaygınlığını anlamak olsun, alansal veriler etkili karar destek sistemlerine katkıda bulunmaktadır.

Kartografik sınıflandırma ve kümeleme analizi yöntemleri mekânsal analiz ve veri temsilde kullanılan yöntemlerdir. Genellikle önceden belirlenmiş sınıflara dayanan kartografik sınıflandırmanın aksine, kümeleme algoritmaları hemen görülemeyen gizli yapıları ve gruplamaları ortaya çıkarabilmektedir. Bu bakımdan kümeleme yöntemleri, amaç mekânsal veriler içindeki doğal örüntüleri ve ilişkileri keşfetmek olduğunda özellikle avantajlıdır. Kümeleme yöntemleri, denetimsiz bir şekilde çalışmakta, yani önceden etiketlenmiş eğitim verileri gerektirmemektedir. Bu esneklik, sınıfların veya kategorilerin sayısının bilinmediği, geliştiği veya kolayca tanımlanamadığı durumlarda etkili olmakta ve veri analizine daha iyi bir yaklaşım sağlamaktadır. Kümeleme, zaman içinde mekânsal verilerdeki dinamik değişikliklere uyum sağlamaktadır. Buna karşılık, kartografik sınıflandırma, gelişen veri kümeleriyle karşılaşıldığında manuel ayarlamalar veya yeniden sınıflandırma gerektirebilmekte, bu da kümelemeyi sürekli ve uyarlanabilir analizler için daha uygun bir seçim haline getirmektedir. Özellikle bu çalışmada olduğu gibi veri sayısının artarak büyük veri formatına dönüştüğü, veri kontrolünün zorlaştığı ve birden fazla değişkene bağlı olarak üretilen çok değişkenli haritalarda kümeleme yöntemlerinin daha etkili olduğu görülmektedir.

Tez çalışması kapsamında, k-ortalamlar (k-Means) ve Ward yöntemi olmak üzere iki farklı kümeleme algoritması kullanılmıştır. Tweet veri setini kümeleme işlemlerinde, Ward yönteminin k-ortalamlar yöntemi göre daha hızlı kümeleme işlemi gerçekleştirdiği görülmüştür. Ancak k-ortalamlar yöntemiyle oluşturulan kümelerin öznelik değişkenlerine bakıldığında maksimum ve minimum değerlerinin diğer kümelerden daha iyi ayrıştığı görülmektedir. Bu tez çalışmasında kullanılan veri seti için k-ortalamlar yöntemiyle kümeleme analizi sonuçlarının daha doğru kümeleme sonuçları verdiği gözlemlenmiştir.

Kümeleme analizinde kullanılan RStudio ve KNIME yazılımlarına baktığımızda, her iki yazılımın da veri madenciliği ve veri analizi gibi veri odaklı işlemlerde kullanılan açık kaynak kodlu yazılımlar olduğu görülmektedir. RStudio yazılımı, R programlama dili odaklı olup, R kodlarının oluşturulması ve çalıştırılması için tasarlanmıştır. KNIME yazılımı ise, görsel bir iş akışı (workflow) tasarım aracı mantığıyla çalışmakta olup, R, Python, Java vb. birden fazla programlama dili ile çalışabilmektedir. RStudio temelde bir kod geliştirme ortamı olup, daha çok R kodu yazarak veri analizi gerçekleştiren kullanıcılar için uygundur. KNIME ise, görsel iş akışı

araçlarıyla kod yazma ihtiyacını en aza indirmektedir. Bu nedenle kodlama bilgisine sahip olmayan kullanıcılar tarafından daha çok tercih edilmektedir.

Tez çalışmasının son aşamasında geliştirilen mekânsal görsel analitik uygulaması, temel olarak COVID-19 ile ilgili tweetlere yönetici düzeyinde erişim sağlayabilen, tweet verilerinin mekânsal analizleri, özniteliksiz analizleri ve kümeleme analizi sonuçlarını entegre eden bir Web tabanlı uygulama örneğidir. Bu tezin temel amacı, bilgi bilimi ve kartografinin etkileşimli yapısını kullanarak Web tabanlı bir mekânsal görsel analitik uygulama örneği oluşturmaktır. Uygulama, Web haritaları, sorgulama ve filtreleme araçları, grafik araçlar ve gösterge sayaçları gibi etkileşimli öğeleri içerir, bu sayede analiz sonuçları kullanıcıya anlamlı bir şekilde iletmektedir.

Sonuç olarak mekânsal görsel analitik, geleneksel haritalama yöntemlerinden dinamik bir ayrılış olarak, mekânsal veri analizine daha etkileşimli, çok boyutlu ve zamansal olarak farkında bir yaklaşım sunmaktadır. Teknoloji ilerlemeye devam ettikçe ve dünyamız giderek daha fazla veri odaklı ve karmaşık hale geldikçe, mekânsal bilgileri gerçek zamanlı olarak etkileşimli bir şekilde keşfetme ve analiz etme yeteneği çok önemli hale gelecek ve mekânsal görsel analitik muhtemelen giderek daha hayati bir rol oynayacaktır. Klasik haritalama coğrafi bilgilerin aktarılması için değerli bir araç olmaya devam ederken, mekânsal analitik dönüştürücü bir yaklaşım olarak ortaya çıkmakta ve kullanıcıların mekânsal modeller hakkında daha derin öngörüler elde etmelerini ve çeşitli sektörlerde daha bilinçli kararlar almalarını sağlamaktadır.

Bu tezin her aşamasına araştırmacı ve uygulayıcı bir perspektiften bakıldığında, uygulamaların yarıya yakın kısmının masaüstü yazılımlarla ve diğer yarıya yakın kısmının Web tabanlı yazılımlar üzerinde gerçekleştirilmiştir. Tasarımcı tarafından belirlenen kullanıcı arayüzleri, Web haritaları, grafikler ve gösterge sayaçları gibi öğeler, görüntüleyiciye başarılı bir şekilde sunulmuştur.

Bu tez araştırması kapsamında edinilen bilgi ve deneyimler doğrultusunda; bu çalışmanın bir sonraki aşaması, görüntüleyici pozisyonunda bulunan son kullanıcının analizini hedeflediği konuya ilişkin tweetleri X API'siyle otomatik olarak internet tabanlı veri tabanlarına çekebildiği, kullanıcıya sunulacak Web tabanlı uygulama arayüzleriyle veri görüntüleme, filtreleme, düzeltme işlemlerinin gerçekleştirildiği, yine Web tabanlı olarak kümeleme analizi yöntemlerinin uygulandığı, elde edilen bilgilerle Web tabanlı olarak haritaların yapılabildiği, grafiklerin oluşturulabildiği ve kullanıcının kendisine ait görsel tasarımla sonuçlandırdığı bir uygulama oluşturma platformunun

geliştirilmesine odaklanmak olacaktır. Tabii ki bahsedilen bu işlemlerin, mutlak suretle bir kişinin tek başına planlaması, gerçekleştirilmesi, uygulaması, yönetmesi ve kullanıcı tarafından gelecek geri bildirimler doğrultusunda düzeltmeler gerçekleştirilmesi gereken işlem adımlarıyla düşünüldüğünde çok zorlu bir süreç olarak görülmektedir. Ancak konularında deneyimli bir araştırma ekibi tarafından gerçekleştirilmesi imkânsız olmayan bir süreç olduğu değerlendirilmektedir.



KAYNAKLAR

- Ajao, O., Hong, J., Liu, W., 2015, A survey of location inference techniques on Twitter. *Journal of Information Science*, s.1–10.
- Alderson, T., Purss, M., Du, X., Mahdavi-Amiri, A., ve Samavati, F., 2020, Digital Earth Platforms, in: *Manual of Digital Earth*, edited by Guo, H., Goodchild, M. F., and Annoni, A., s.25–55, Springer, Singapore, 1 edn., <https://doi.org/10.1007/978-981-32-9915-3>
- Andrienko, G., Andrienko, N., Jankowski, P., Keim, D., Kraak, M. J., MacEachren, A. M. ve Wrobel, S., 2007, Geovisual analytics for spatial decision support: setting the research agenda. *International Journal of Geographical Information Science*, 21(8), s.839-857.
- Andrienko, G., Andrienko, N., Keim, D., MacEachren, A. ve Wrobel, S., 2011, Challenging problems of geospatial visual analytics, *Journal of Visual Languages and Computing* 22 (4), s.251-256.
- Andrienko, G., Andrienko, N., Fuchs, G., & Garcia, J. M. C., 2017, Clustering Trajectories by Relevant Parts for Air Traffic Analysis. *IEEE Transactions on Visualization and Computer Graphics*, 24(1), Forthcoming.
- Bayrakçı, S., 2015, Sosyal Bilimlerdeki Akademik Çalışmalarda Büyük Veri Kullanımı. (Yüksek Lisans Tezi). İstanbul: Marmara Üniversitesi Sosyal Bilimler Enstitüsü
- Battersby, S. E., Finn, M. P., Usery, E. L. ve Yamamoto, K. H., 2014, Implications of web Mercator and its use in online mapping. *Cartographica: The International Journal for Geographic Information and Geovisualization*, 49(2), 85-101.
- Bildirici, İ. Ö., 2015, The web Mercator projection: A cartographic analysis. *Proceedings of the 1st ICA European Symposium on Cartography*. Vienna, Avusturya, 221-231.
- Bildirici, İ. Ö. ve Kırtıloğlu, O. S., 2016, Web merkator projeksiyonu, *Harita Teknolojileri Elektronik Dergisi*, 8 (2): 99-109
- Bildirici, İ. Ö., 2018, Kartografya: Harita Tasarımı ve Kullanımı için Gerekli Bilim, Sanat ve Teknik. Atlas Akademi Yayınevi, Konya
- Bloch, M., Cox, A., McGinty, J.C., Quealy, K., 2010, A peek into Netflix queues, <http://www.nytimes.com/interactive/2010/01/10/nyregion/20100110-netflix-map.html>, The New York Times.
- Bousquin, J., 2021, Discrete Global Grid Systems as scalable geospatial frameworks for characterizing coastal environments, *Environmental Modelling & Software*, 146, s.105–210, <https://doi.org/10.1016/j.envsoft.2021.105210>
- Cao, N., Lin, Y., Sun, X., Lazer, D., Liu, S., Qu, H., 2012, Whisper: Tracing the spatiotemporal process of information diffusion in real time. *IEEE Transactions on Visualization and Computer Graphics* 18, s.2649–2658.

- Card, S.K., Mackinlay, J.D., Shneiderman, B., 1999, Readings in Information Visualization: Using Vision To Think., Kapitel 1, s. 1–34
- Chaudhuri, C., Gray, A., ve Robertson, C., 2021, InundatEdv1.0: a height above nearest drainage (HAND)-based flood risk modeling system using a discrete global grid system, *Geoscientific Model Development*, 14, s.3295–3315, <https://doi.org/10.5194/gmd-14-3295-2021>
- Chen, H., Chiang, R. H., Storey, V., 2012, “Business intelligence and analytics: from big data to big impact”. *MIS quarterly*, s.1165-1188.
- Chen, M., Mao, S., Liu, Y., 2014. Big data: a survey. *Mobile Networks Appl.* 19 (2), s.171–209.
- Chen, S., Lin, L. ve Yuan, X., 2017, Social Media Visual Analytics. *Computer Graphics Forum*, 36: s.563-587. doi:10.1111/cgf.13211
- Cukier, K., 2010, Data, data everywhere; A special report on managing information, *The Economist*
- Çağlar, B., 2018, Mekânsal Verilerin Kümeleme Analizi ile Değerlendirilmesi (Yüksek Lisans Tezi), Necmettin Erbakan Üniversitesi Fen Bilimleri Enstitüsü
- Dasgupta, A., 2013. Big data: the future is in analytics. *Geospatial World*.
- Dayan, S., 2021, COVID-19 ve Aşı, *Dicle Tıp Dergisi / Dicle Medical Journal*, 48 (Özel Sayı / Special Issue): s.98-113
- Demchenko, Y., Ngo, C. ve Membre, P., 2013, Architecture Framework and Components for the Big Data Ecosystem Draft Version 0.2 (SNE Technical Report SNE-UVA-2013-02). System and Network Engineering Group & Universiteit Van Amsterdam.
- Dibb, S., 1998, Market Segmentation: Strategies for success, *Marketing Intelligence & Planning*, 16/7, s. 394–406.
- Eagle, N. ve Greene, K., 2014, Reality Mining: Using Big Data to Engineer a Better World, 1st edition, The MIT Press.
- Eligüzel, N., 2021, Using Twitter for situational awareness after an earthquake: The roles of text categorization and location information, Doktora Tezi, Gaziantep Üniversitesi Fen Bilimleri Enstitüsü, Gaziantep.
- Fayyad, U. M., Piatetsky-Shapiro, G. Smyth, P., 1996, From Data Mining to Knowledge Discovery in Databases. *AI Magazine*, 17(3):37.
- Fayyad, U. M., Piatetsky-Shapiro, G. Smyth, P., 1996, From Data Mining to Knowledge Discovery: An Overview. In U. M. Fayyad, G. Piatetsky-Shapiro, P. Smyth, Uthurusamy, R. *Advances in Knowledge Discovery and Data Mining*, Kapitel 1, s. 1–34.

- Ferreira, L. ve Hitchcock, D. B., 2009, A comparison of hierarchical methods for cluster functional data, *Communications in Statistics-Simulation and Computation*, 38(9), s.1925-1949.
- Fuchs, G., Andrienko, N., Andrienko, G., Bothe, S., Stange, H., 2013. Tracing the German centennial flood in the stream of tweets: first lessons learned. In: *Proceedings of the Second ACM SIGSPATIAL International Workshop on Crowdsourced and Volunteered Geographic Information*, s. 31–38.
- Furtado, V., Ayres, L., De Oliveira, M., Vasconcelos, E., Caminha, C., D’Orleans, J., Belchior, M., 2010, Collective intelligence in law enforcement—the WikiCrimes system. *Inf Sci* 180(1), s.4–17.
- Gahegan, M., 2005, Beyond tools: visual support for the entire process of GIScience. In J. Dykes, A. M. MacEachren, & M. J. Kraak (Eds.), *Exploring Geovisualization*, London, UK: Elsevier s. 83-99.
- Gandomi, A. ve Haider, M., 2015, Beyond the Hype: Big Data Concepts, Methods, and Analytics. *International Journal of Information Management*, 35(2), s.137-144.
- Gibb, R., Raichev, A., ve Speth, M., 2016, The rHEALPix DGGS preprint, https://raichev.net/files/rhealpix_dggs_preprint.pdf, Erişim Tarihi: 20 Nisan 2022
- Goodchild, M.F., Glennon, J.A., 2010, Crowdsourcing geographic information for disaster response: a research frontier, *International Journal of Digital Earth*, 3:3, s.231-241.
- Haber, R.B., ve McNabb, D.A., 1990, *Visualization in Scientific Computing*, IEEE Computer Society Press, s. 74-93.
- Han, J. ve Kamber, M., 2006, *Data mining: concepts and techniques*. Morgan Kaufmann.
- Hands, S. ve Everitt, B., 1987, A Monte Carlo study of the recovery of cluster structure in binary data by hierarchical cluster techniques. *Multivar. Behav. Res.* 22, s.235-243.
- Healey, C.G., Booth, K.S., Enns, J.T., 1996, High-Speed Visual Estimation Using Preattentive Processing. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 3(2), s.107–135.
- Hecht, B., Hong, L., Suh, B. ve Chi, E.H., 2011, Tweets from Justin Bieber's heart: the dynamics of the location field in user profiles, *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems ACM*, s.237-246.
- Ikawa, Y., Vukovic, M., Rogstadius, J. ve Murakami, A., 2013, Location-based insights from the social Web, *Proceedings of the 22nd international conference on World Wide Web companion International World Wide Web Conferences Steering Committee*, s.1013-1016.
- İpbüker, C., 2005, Çok Yüzlüler Üzerine İzdüşüm Uygulamaları: Fuller Projeksiyonu, *Jeodezi, Jeoinformasyon ve Arazi Yönetimi Dergisi*, 92, s.38-43.

- Järvinen, P., Puolamäki, K., Siltanen, P. ve Ylikerälä, M., 2009, Visual Analytics Final Report, VTT Technical Research Centre of Finland
- Jendryke, M. ve McClure, S. C., 2019, Mapping crime – Hate crimes and hate groups in the USA: A spatial analysis with gridded data, *Applied Geography*, 111, 102-117, <https://doi.org/10.1016/j.apgeog.2019.102072>
- Jurgens, D., 2013, That's What Friends Are For: Inferring Location in Online Social Media Platforms Based on Social Relationships, ICWSM.
- Jurgens, M. ve Helsloot, I., 2018, The effect of social media on the Dynamics of (self) resilience during disasters: A literature review. *Journal of Contingencies and Crisis Management*, 26 (1), s.79-88.
- Keim, D. A., 2008, Visual Analytics in Europe. NVAC Views.
- Keim, D.A., Kohlhammer, J., Mansmann, F., May, T., Wanner, F., 2010. Visual analytic". In: Keim, D.A., Kohlhammer, J., Ellis, G. ve Mannsmann, F. (eds.) *Mastering the Information Age: Solving Problems with Visual Analytics*, Goslar: Eurographics Association, s.7-18.
- Keim, D.A., Mansmann, F., Schneidewind, J. ve Ziegler, H., 2006, Challenges in visual data analysis. In *Information Visualization (IV 2006)*, Invited Paper, July 5-7, London, United Kingdom. IEEE Press.
- Kmoch, A., Matsibora, O., Vasilyev, I., ve Uuemaa, E., 2022, Applied open-source Discrete Global Grid Systems, *AGILE GIScience Ser.*, 3, 41, <https://doi.org/10.5194/agile-giss-3-41-2022>
- Kuşay, Y., 2010, "Sosyal Medyanın Gücü ve Uygulama Örnekleri, İkinci Medya Çağında İnternet, Alfa Yayınevi, İstanbul.
- Laney, D., 2001. *3D Data Management: Controlling Data Volume, Velocity, and Variety. Application Delivery Strategies.*
- Lamsal, R., 2020, Design and analysis of a large-scale COVID-19 tweets dataset, *Applied Intelligence*, Volume, 51, s.2790–2804
- Lee, J. ve Kang, M., 2015, *Geospatial Big Data: Challenges and Opportunities*, Big Data Research (Elsevier), s.74–81.
- Lei, T. L., Liang, X., Mascaro, G., Luo, W., White, D., Westerhoff, P. and Maciejewski, R., 2015, An Interactive Web-Based Geovisual Analytics Tool to Explore Water Scarcity in Niger River Basin, Conference: *Visualization in Environmental Sciences (EnvirVis 2015)*, Cagliari, Italy
- Lei, K., Dongxu, Q., ve Xiaolin, T., 2020, "A New Coordinate System for Constructing Spherical Grid Systems" *Applied Sciences* 10, no.2: 655. <https://doi.org/10.3390/app10020655>

- Li, C. ve Sun, A., 2014, Fine-grained location extraction from tweets with temporal awareness, Proceedings of the 37th international ACM SIGIR conference on Research & development in information retrieval, ACM, s.43-52.
- Li, R., Wang, S. ve Chang, K.C., 2012, Multiple location profiling for users and relationships from social network and content, Proceedings of the VLDB Endowment, vol. 5(11), s.1603-1614.
- Li, S., Dragicevic, S., Castro, F.A., Sester, M., Winter, S., Çöltekin, A., Pettit, C., Jiang, B., Haworth, J., Stein, A., Cheng, T., 2016, Geospatial big data handling theory and methods: A review and research challenges, ISPRS Journal of Photogrammetry and Remote Sensing, s.119-133.
- Lin, K., Kansal, A., Lymberopoulos, D. ve Zhao, F., 2010, Energy-accuracy trade-off for continuous mobile device location, Proceedings of the 8th international conference on Mobile systems, applications, and services ACM, s.285-298.
- Luo, W., Chang, Z., Kong, L., Link, R., Hejazi, M., Clarke, ., Maciejewski, R., 2015, Web-Based Visualization of the Global Change Assessment Model, Workshop on Visualisation in Environmental Sciences (EnvirVis).
- Lüders, M., 2008, Conceptualizing Personal Media. *New Media & Society*, 10 (5).
- MacEachren, A. M., Buttenfield, B. P., Campbell, J. B., DiBiase, D. W., ve Monmonier, M., 1992, Visualization. *Geography's inner worlds: Pervasive themes in contemporary American geography*, s. 99–137.
- MacEachren, A., 1994, Visualization in Modern Cartography: Setting the Agenda, *Visualization in Modern Cartography*, Vol.2, Pergamon Press, s.1-12.
- MacEachren, A. M., Jaiswal, A., Robinson, A. C., Pezanowski, S., Savelyev, A., Mitra, P., Zhang, X., ve Blanford, J. ,2011, SensePlace2: Geotwitter Analytics Support for Situation Awareness. 2nd IEEE Conference on Visual Analytics Science and Technology 2011, VAST 2011- Providence, RI, United States, s. 181-190.
- Maciejewski, R., Hafen, R., Rudolph, S., Larew, S. G., Mitchell, M. A., Cleveland, W. S., & Ebert, D. S., 2011, Forecasting Hotspots—A Predictive Analytics Approach. *Visualization and Computer Graphics*, IEEE Transactions on, 17(4), s. 440-453. doi:10.1109/tvcg.2010.82
- Mahmud, J., Nichols, J. ve Drews, C., 2014, Home Location Identification of Twitter Users, *ACM Transactions on Intelligent Systems and Technology (TIST)*, vol.5(3)
- Maimon, O. ve Rokach, L., 2010, *Data Mining and Knowledge Discovery Handbook*, Kapitel 1, S. 1–15. doi:10.1007/978-0-387-09823-4.
- Maletic, J.I. ve Marcus, A., 2010, Data Cleansing: A Prelude to Knowledge Discovery, *Data Mining and Knowledge Discovery Handbook*, Kapitel 2, s. 19–32. doi: 10.1007/978-0-387-09823-4.
- Mansmann, F., Fischer, F., North, S. C. ve Keim, D. A., 2009, Visual support for analyzing network trac and intrusion detection events using treemap and graph

- representations. In CHiMiT '09: Proceedings of the Symposium on Computer Human Interaction for the Management of Information Technology, s.19–28.
- Mayfield, A., 2008, What is Social Media, http://crmexchange.com/uploadedFiles/White_Papers/PDF/What_is_Social_Media_iCrossing_ebook.pdf, Erişim Tarihi: 02.04.2023.
- McGee, J., Caverlee, J.A. ve Cheng, Z., 2011, A geographic study of tie strength in social media, Proceedings of the 20th ACM international conference on Information and knowledge management, ACM, s.2333-2336.
- Meeker, M., 2012, KPCB internet trends year-end update, <http://www.slideshare.net/kleinerperkins/2012-kpcb-internet-trends-year-end-update>.
- Mendoza, M., Poblete, B., Castillo, C., 2010, Twitter Under Crisis: Can we trust what we RT? SOMA 2010- Proceedings of the 1st Workshop on Social Media Analytics. doi: 10.1145/1964858.1964869.
- Morais, C.D., 2012, Where is the Phrase “80% of Data is Geographic” From?. <<http://www.gislounge.com/80-percent-data-is-geographic/>> (Erişim Tarihi: 10 Nisan 2022).
- Murtagh, F. ve Legendre, P., 2014, Ward’s hierarchical agglomerative cluster method: Which algorithms implement ward’s criterion? Journal of Classification, 31(3), s.274-295.
- Murthy, B. ve Longwell, S.A., 2012, Twitter And Disasters, Information Communication&Society, 16(6), s.1-19.
- Mutukrishnan, S., 2005, Data Stream: Algorithms and Applications. Foundations and Trends® in Theoretical Computer Science: Vol. 1: No.2: s.117-236.
- Nair, M.R., Ramya, G.R., Sivakumar, P.B., 2017, Usage and analysis of Twitter during 2015 Chennai flood towards, disaster management, Procedia Computer Science, 115 Os.,350–358.
- Narasimhan, R. ve Bhuvaneshwari, T., 2014, Big Data- A Brief Study. International Journal of Scientific & Engineering Research, 5(9), s.350-353.
- Olafsson, S., Li, X., Wu, S., 2008, Operations research and data mining, European Journal of Operational Research, 187 (3), s. 1429-1448, doi:10.1016/j.ejor.2006.09.023.
- Öztürk, F., E., 2022, Ağ Erkeklikleri: Hegemonik Erkekliğin Dijital Dönüşümü, Doktora Tezi, Marmara Üniversitesi Sosyal Bilimler Enstitüsü, İstanbul.
- Percivall, G., 2013, The power of location, <http://www.opengeospatial.org/blog/1817>, Open Geospatial Consortium.

- Pirolli, P., ve Card, S., 2005, The sensemaking process and leverage points for analyst technology as identified through cognitive task analysis. Paper presented at the International Conference on Intelligence Analysis, McLean, VA.
- Puolamäki, K., Bertone, A., Therón, R., Huisman, O., Johansson, J., Miksch, S., Papapetrou, P., Rinzivillo, S., 2010, Data Mining, Mastering the Information Age - Solving Problems with Visual Analytics, Kapitel 2, S. 39–56. Eurographics Association, Goslar.
- Purss, M. B. J., Peterson, P. R., Strobl, P., Dow, C., Sabeur, Z. A., Gibb, R. G. and Ben, J., 2019, Datacubes: A Discrete Global Grid Systems Perspective, *Cartographica*, 54, s.63–71, <https://doi.org/10.3138/cart.54.1.2018-0017>
- Rajabifard, A., Paez, D., Foliente, G., 2021, COVID-19 Pandemic, Geospatial Information, and Community Resilience, CRC Press, doi:10.1201/9781003181590
- Rawson, A., Sabeur, Z. A., ve Brito, M., 2021, Intelligent geospatial maritime risk analytics using the Discrete Global Grid System, *Big Earth Data*, s. 1–29, <https://doi.org/10.1080/20964471.2021.1965370>
- Robertson, C., Chaudhuri, C., Hojati, M., ve Roberts, S. A., 2020, An integrated environmental analytics system (IDEAS) based on a DGGS, *ISPRS Journal of Photogrammetry and Remote Sensing*, 162, 214–228, <https://doi.org/10.1016/j.isprsjprs.2020.02.009>
- Robinson, A.C., 2017, Geovisual Analytics, *The Geographic Information Science & Technology Body of Knowledge (3rd Quarter 2017 Edition)*, (ed.). doi: 10.22224/gistbok/2017.3.6.
- Robinson, A.C., Chen, J., Lengerich, G., Meyer, H., ve MacEachren, A. M., 2005, Combining usability techniques to design geovisualization tools for epidemiology. *Cartography and Geographic Information Science*, 32(4), 243-255. doi:10.1559/152304005775194700
- Robinson, A.C., Peuquet, D. J., Pezanowski, S., Hardisty, F. A., ve Swedberg, B., 2016, Design and evaluation of a geovisual analytics system for uncovering patterns in spatio-temporal event data. *Cartography and Geographic*.
- Romesburg H.C., 1984, *Cluster Analysis for Researchers*, Belmont, CA: Life time Learning Publications
- Rousseeuw P.J., 1987, Silhouettes: A Graphical Aid to the Interpretation and Validation of Cluster Analysis, *Journal of Computational and Applied Mathematics*, 20, s.53-65.
- Roth, R.E., 2013, Interactive Maps: What we know and what we need to know. *The Journal of Spatial Information Science*, 6 s.59-115.
- Sahr, K., 2023, DGGRID version 7.8, <https://www.discreteglobalgrids.org/software/>, Erişim tarihi: 1 Mayıs 2023

- Saraçlı, S. ve Akşit, M., 2022, Comparison of Hierarhic Clustering Methods with Cophenetic Correlation Coefficient in Big Data, Afyon Kocatepe Üniversitesi Fen ve Mühendislik Bilimleri Dergisi, s.552-559
- Schmidt, J.H., 2016, Twitter Kişisel Kamuların Doğuşu, Twitter ve Toplum, İstanbul: Kafka, s.47-62.
- Schulz, A., Hadjakos, A., Paulheim, H., Nachtwey, J. ve Mühlhäuser, M., 2013, A Multi-Indicator Approach for Geolocalization of Tweets, ICWSM
- Selvi, H.Z. ve Çağlar, B., 2017, Çok Değişkenli Haritalama İçin Kümeleme Yöntemlerinin Kullanılması, Niğde Ömer Halisdemir Üniversitesi Mühendislik Bilimleri Dergisi, 6(2), s.415-429.
- Sester, M., Arsanjani, J.J., Klammer, R., Burghardt, D., Haurert, J.-H., 2014, Integrating and Generalising Volunteered Geographic Information. Abstracting Geographic Information in a Data Rich World – Methodologies and Applications of Map Generalisation. Springer, Heidelberg, s. 119–155.
- Shekhar, S., 2012, Spatial big data challenges, in: Keynote at ARO/NSF Workshop on Big Data at Large: Applications and Algorithms, Durham, NC.
- Shekhar, S., Evans, M. R., Kang, J. M., ve Mohan, P., 2011, Identifying patterns in spatial information: A survey of methods. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 1(3), s.193-214.
- Shneiderman, B., 1996, The eyes have it: A task by data type taxonomy for information visualizations. In IEEE Symposium on Visual Languages, s.336–343.
- Silahtaroglu, G., 2013, Veri Madenciliği (Kavram ve Algoritmaları), Papatya Yayıncılık, İstanbul.
- Slocum, T., McMaster, R., Kessler, F., ve Howard, H., 2008, Thematic Cartography and Geovisualization. Pearson Prentice Hall, Upper Saddle River, New Jersey, 3. baskı.
- Sorrell, S.M., 2011, The power of apps, in: The 2011 GSMA Mobile World Congress, <http://www.youtube.com/watch?v=5gfTQUq0mHw>.
- Sütcü, C. S., ve AYTEKİN, Ç., 2013, Elektronik Ticaretten Sosyal Ticarete Dönüşüm Süresinde Ölçümleme. İstanbul: Derin Yayınevi.
- Thomas, J. ve Cook, K., 2005, Illuminating the Path: Research and Development Agenda for Visual Analytics, IEEE Press. s.194.
- Tobler, W., 1970, A computer movie simulating urban growth in the Detroit region, Econ. Geogr. 46(2), s.234–240.
- Tominski, C., Abello, J. ve Schumann, H., 2009, CGV–An interactive graph visualization system. Computers & Graphics, 33(6):s.660–678.

- Toriumi, F., Sakaki, T., Shinoda, K., Kazama, K., Kurihara, S., Noda, I., 2013. Information sharing on Twitter during the 2011 catastrophic earthquake. 1025-1028. doi:10.1145/2487788.2488110.
- Upton, C., Faulhaber, T., Kamins, D., Laidlaw, D.H., Schlegel, D., Vroom, J., Gurwitz, R. ve Van Dam, A., 1989, The Application Visualization System: A Computational Environment for Scientific Visualization, IEEE Comput. Graph. Appl. 9, 4, s.30–42, doi:10.1109/38.31462
- Valdes-Dapena, P., 2011, GPS systems that save gas, http://money.cnn.com/2011/03/03/autos/navigation_gps_fuel_economy/ /CNN Money
- Van Ho, Q., Lundblad, P., Astrom, T., ve Jern, M., 2012, A Web-enabled visualization toolkit for geovisual analytics, Information Visualization, 11(1), s.22-42.
- Van Wijk, J.J., 2005, "The value of visualization," VIS 05. IEEE Visualization, 2005., Minneapolis, MN, USA, s. 79-86, doi: 10.1109/VISUAL.2005.1532781.
- Vatsavai, R.R., Ganguly, A., Chandola, V., Stefanidis, A., Klasky, S., Shekhar, S., 2012, Spatiotemporal data mining in the era of big spatial data: algorithms and applications, in: Proceedings of 2nd ACM SIGSPATIAL International Workshop on Analytics for Big Geospatial Data, Redondo Beach, CA, s.1–10.
- Veach, E., Rosenstock, J., Engle, E., Snedegar, R., Basch, J., and Manshreck, T., 2017, S2 Geometry Library, <https://s2geometry.io>, Erişim Tarihi: 20 Nisan 2023
- Wong, P.C ve Thomas, J., 2004, Visual analytics. IEEE Computer Graphics and Applications, 24(5), s.20–21.
- Wood, J., Brodlie, K. ve Wright, H., 1996, Visualization over the World Wide Web and its application to environmental data. In Proceedings of the 7th conference on Visualization '96 (VIS '96). IEEE Computer Society Press, Washington, DC, USA.
- Zhang, C., Liu, Y., Wang, C., 2013, Time-space varying visual analysis of micro-blog sentiment. S.64-71.
- URL-1: <https://www.domo.com/data-never-sleeps#data> [Erişim Tarihi: 10 Şubat 2023]
- URL-2: <https://tr.wikipedia.org/wiki/Twitter> [Erişim Tarihi: 8 Ocak 2022]
- URL-3: <https://wearesocial.com/uk/blog/2023/01/digital-2023/> [Erişim Tarihi: 10 Ağustos 2023]
- URL-4: <https://www.kaggle.com/datasets/gpreda/all-covid19-vaccines-tweets> [Erişim Tarihi: 6 Kasım 2021]
- URL-5: <https://www.knime.com/software-overview> [Erişim Tarihi: 10 Mart 2022]
- URL-6: <https://posit.co/products/open-source/rstudio/> [Erişim Tarihi: 11 Mart 2022]

- URL-7: <https://h3geo.org/> [Erişim Tarihi: 5 Aralık 2022]
- URL-8: <https://www.uber.com/en-IN/blog/h3/> [Erişim Tarihi: 5 Aralık 2022]
- URL-9: <https://h3geo.org/docs/highlights/aggregation> [Erişim Tarihi: 5 Aralık 2022]
- URL-10: <https://h3geo.org/docs/> [Erişim Tarihi: 5 Aralık 2022]
- URL-11: <https://h3geo.org/docs/highlights/indexing> [Erişim Tarihi: 7 Aralık 2022]
- URL-12: <https://h3geo.org/docs/core-library/h3Indexing> [Erişim Tarihi: 7 Aralık 2022]
- URL-13: <https://h3geo.org/docs/library/index/cell> [Erişim Tarihi: 7 Aralık 2022]
- URL-14: <https://h3geo.org/docs/library/index/directededge> [Erişim Tarihi: 7 Aralık 2022]
- URL-15: <https://h3geo.org/docs/library/index/vertex> [Erişim Tarihi: 8 Aralık 2022]
- URL-16: <https://observablehq.com/@nrabinowitz/h3-index-inspector#85283473ffffff> [Erişim Tarihi: 7 Aralık 2022]
- URL-17: <http://webpages.sou.edu/~sahrk/dgg/orientation/dymorient.html> [Erişim Tarihi: 5 Aralık 2022]
- URL-18: <https://www.esri.com/arcgis-blog/products/arcgis-pro/analytics/use-h3-to-create-multiresolution-hexagon-grids-in-arcgis-pro-3-1/> [Erişim Tarihi: 20 Nisan 2023]
- URL-19: <https://h3geo.org/docs/api/indexing> [Erişim Tarihi: 9 Aralık 2022]
- URL 20: <https://h3geo.org/docs/core-library/restable/> [Erişim Tarihi: 5 Aralık 2022]
- URL-21: <https://developers.arcgis.com/documentation/mapping-apis-and-services/geocoding/> [Erişim Tarihi: 10 Ocak 2023]
- URL-22: <https://developers.arcgis.com/web-appbuilder/guide/xt-welcome.htm> [Erişim Tarihi: 10 Ocak 2023]
- URL-23: <https://www.esri.com/en-us/arcgis/products/arcgis-dashboards/overview> [Erişim Tarihi: 13 Ocak 2023]
- URL-24: <https://www.esri.com/en-us/arcgis/products/arcgis-hub/overview> [Erişim Tarihi: 13 Ocak 2023]
- URL-25: <https://covid19.who.int/> [Erişim Tarihi: 31 Ekim 2022]
- URL-26: <https://ec.europa.eu/eurostat/web/gisco/geodata/reference-data/administrative-units-statistical-units/countries> [Erişim Tarihi: 14 Ocak 2022]
- URL-27: <https://www.ogc.org/standard/wms> [Erişim Tarihi: 16 Ocak 2022]
- URL-28: <https://www.ogc.org/standard/wfs> [Erişim Tarihi: 16 Ocak 2022]
- URL-29: <https://www.ogc.org/standard/wcs> [Erişim Tarihi: 16 Ocak 2022]

- URL-30: <https://www.jgibson.id.au/articles/convex-hull/> [Eriřim Tarihi: 30.07.2022]
- URL-31: https://github.com/owid/covid-19-data/tree/master/public/data/vaccinations/country_data]Eriřim Tarihi: 10 Őubat 2022]
- URL-32: <https://enterprise.arcgis.com/en/server/10.5/publish-services/windows/what-is-a-feature-service-.htm> [Eriřim Tarihi: 05 Őubat 2022]
- URL-33: <https://tr.euronews.com/2023/07/23/elon-musk-twitterin-mavi-kus-logosunu-degistirecegini-duyurdu> [Eriřim Tarihi: 10 Kasım 2023]
- URL-34: <https://webrazzi.com/2023/02/22/elon-musk-twitterin-algoritmasini-onumuzdeki-hafta-acik-kaynak-haline-getirecegini-soyledi/> [Eriřim Tarihi: 10 Kasım 2023]
- URL-35: <https://www.indyturk.com/node/620776/yařam/twitterin-algoritması-herkese-açıldı-iřte-öğrenilenler> [Eriřim Tarihi: 10 Kasım 2023]

EKLER

Ek-1.1 Metinsel bir ifadeye ait coğrafi kodlama işlemini gerçekleştiren HTML kod bloğu

```

<html>
<head>
<meta charset="utf-8">
<meta name="viewport" content="initial-scale=1, maximum-
scale=1, user-scalable=no">
<title>ArcGIS Developer Guide: Forward geocoding</title>
<!-- ArcGIS Mapping APIs and Location Services Developer
Guide
    Learn more:
https://developers.arcgis.com/documentation/mapping-apis-
and-services/search/
-->
<style>
    html, body, #viewDiv {
        padding: 0;
        margin: 0;
        height: 100%;
        width: 100%;
    }
</style>

<link rel="stylesheet"
href="https://js.arcgis.com/4.27/esri/themes/light/main.css
">
<script src="https://js.arcgis.com/4.27"></script>

<script>
    require([
        "esri/config",
        "esri/Map",
        "esri/views/MapView",
        "esri/Graphic",
        "esri/rest/locator"
    ], (esriConfig, Map, MapView, Graphic, locator) => {

        esriConfig.apiKey = "YOUR_API_KEY";

        const map = new Map({
            basemap: "arcgis-imagery" //Basemap styles service
        });

```

```

const view = new MapView({
  container: "viewDiv",
  map: map,
  constraints: {
    snapToZoom: false
  }
});

view.popup.actions = [];
view.when(()=>{

const geocodingServiceUrl = "https://geocode-
api.arcgis.com/arcgis/rest/services/World/GeocodeServer";

  const params = {
    address: {
      "address": "Köyceğiz Mah. Demeç Sok. No:44/1
Meram/ KONYA"
    }
  }

  locator.addressToLocations(geocodingServiceUrl,
params).then((results) => {
    showResult(results);
  });

function showResult(results) {
  if (results.length) {
    const result = results[0];
    console.log(result.address)
    view.graphics.add(new Graphic({
      symbol: {
        type: "simple-marker",
        color: "#000000",
        size: "8px",
        outline: {
          color: "#ffffff",
          width: "1px"
        }
      },
      geometry: result.location,
      attributes: {
        title: "NECMETTİN ERBAKAN ÜNİVERSİTESİ
MÜHENDİSLİK FAKÜLTESİ",

```

```

        address: result.address,
        score: result.score
    },
    popupTemplate: {
        title: "{title}",
        content: result.address + "<br><br>" +
result.location.longitude.toFixed(5) + "," +
result.location.latitude.toFixed(5)
    }
    });
    if (results.length) {
        const g = view.graphics.getItemAt(0);
        view.openPopup({
            features: [g],
            location: g.geometry
        });
    }
    view.goTo({
        target: result.location,
        zoom: 13
    });
}
});
});
</script>
</head>
<body>
<div id="viewDiv"></div>
</body>
</html>

```

Ek-1.2 Konumu Bilinen Bir Noktanın Adresinin Metinsel Olarak Belirlenmesi (Ters Coğrafi Kodlama) İşlemine Gerçekleştiren HTML Kod Bloğu

```

<html>
<head>
<meta charset="utf-8">
<meta name="viewport" content="initial-scale=1, maximum-
scale=1, user-scalable=no">
<title>ArcGIS Developer Guide: Reverse geocoding</title>
<!-- ArcGIS Mapping APIs and Location Services Developer
Guide
    Learn more:
https://developers.arcgis.com/documentation/mapping-apis-
and-services/search/
-->
<style>
    html, body, #viewDiv {
        padding: 0;
        margin: 0;
        height: 100%;
        width: 100%;
    }
</style>

<link rel="stylesheet"
href="https://js.arcgis.com/4.27/esri/themes/light/main.css
">
<script src="https://js.arcgis.com/4.27"></script>

<script>
    require([
        "esri/config",
        "esri/Map",
        "esri/views/MapView",
        "esri/Graphic",
        "esri/rest/locator"
    ],(esriConfig, Map, MapView, Graphic, locator)=> {
        esriConfig.apiKey = "YOUR_API_KEY";
        const map = new Map({
            basemap: "osm-standard-relief", //Basemap styles
service
        });
        const view = new MapView({
            container: "viewDiv",
            map: map,

```

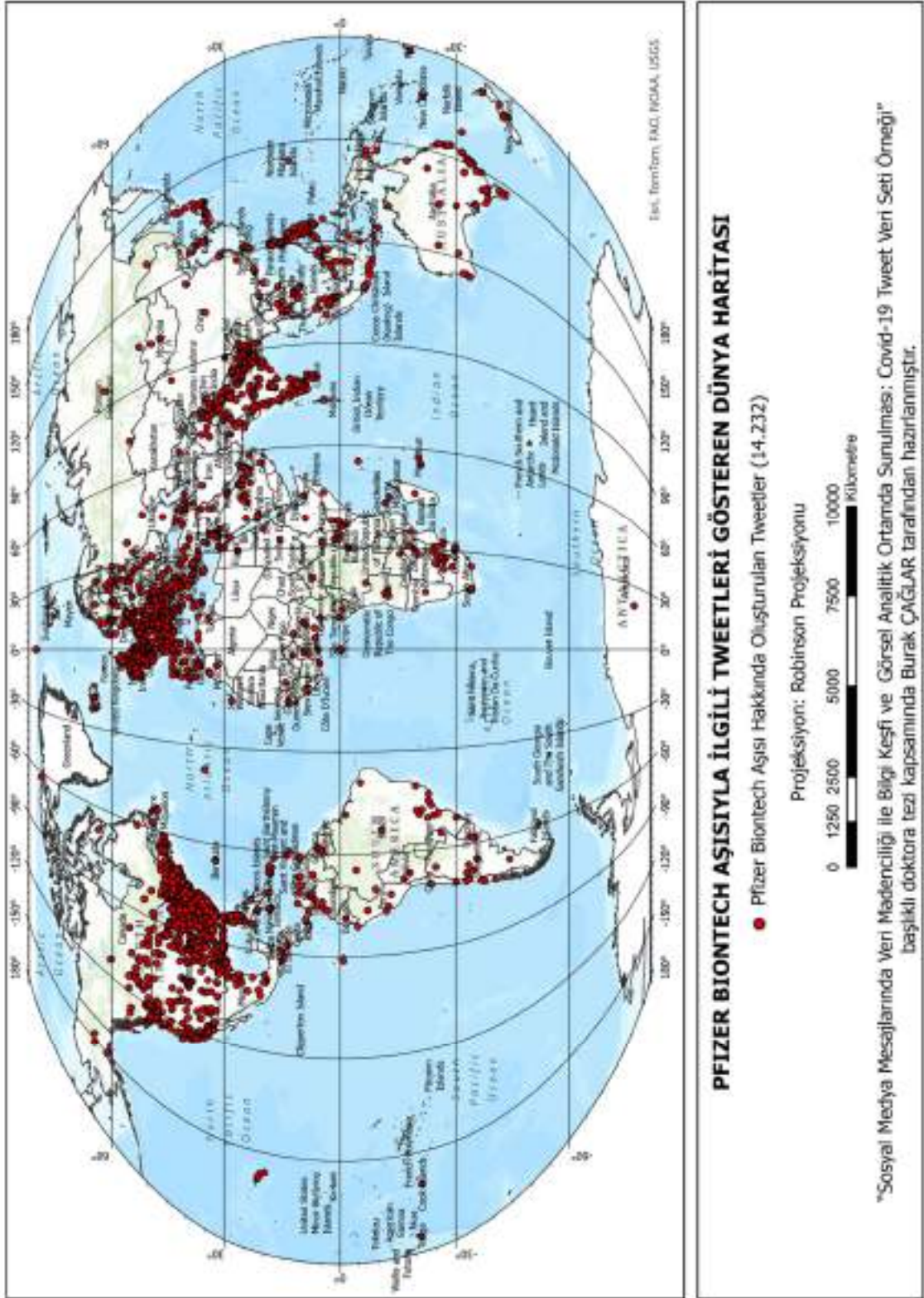
```

    center: [32.419, 37.866], //???
    zoom: 15,
    constraints: {
      snapToZoom: false
    }
  });
  view.popup.actions = [];
  view.when(() => {
    reverseGeocode(view.center);
  });
  view.on("click", (event) => {
    reverseGeocode(event.mapPoint);
  });
function reverseGeocode(pt) {
  const geocodingServiceUrl = "https://geocode-
api.arcgis.com/arcgis/rest/services/World/GeocodeServer" ;
  const params = {
    location: pt
  };
  locator.locationToAddress(geocodingServiceUrl,
params).then(
  (response) => {
    if (response) {
      showPopup(response);
    }
  },
  (err) => {
    showPopup("No address found.", pt);
  }
);
}
//Information shown when map is clicked
function showPopup(response) {
  view.openPopup({
    title: response.attributes.PlaceName || "ADRES",
    content:
      response.attributes.LongLabel +
      "<br><br>" +
      response.location.longitude.toFixed(5) +
      ", " +
      response.location.latitude.toFixed(5),
    location: response.location,
  });
  addGraphic(response);
}

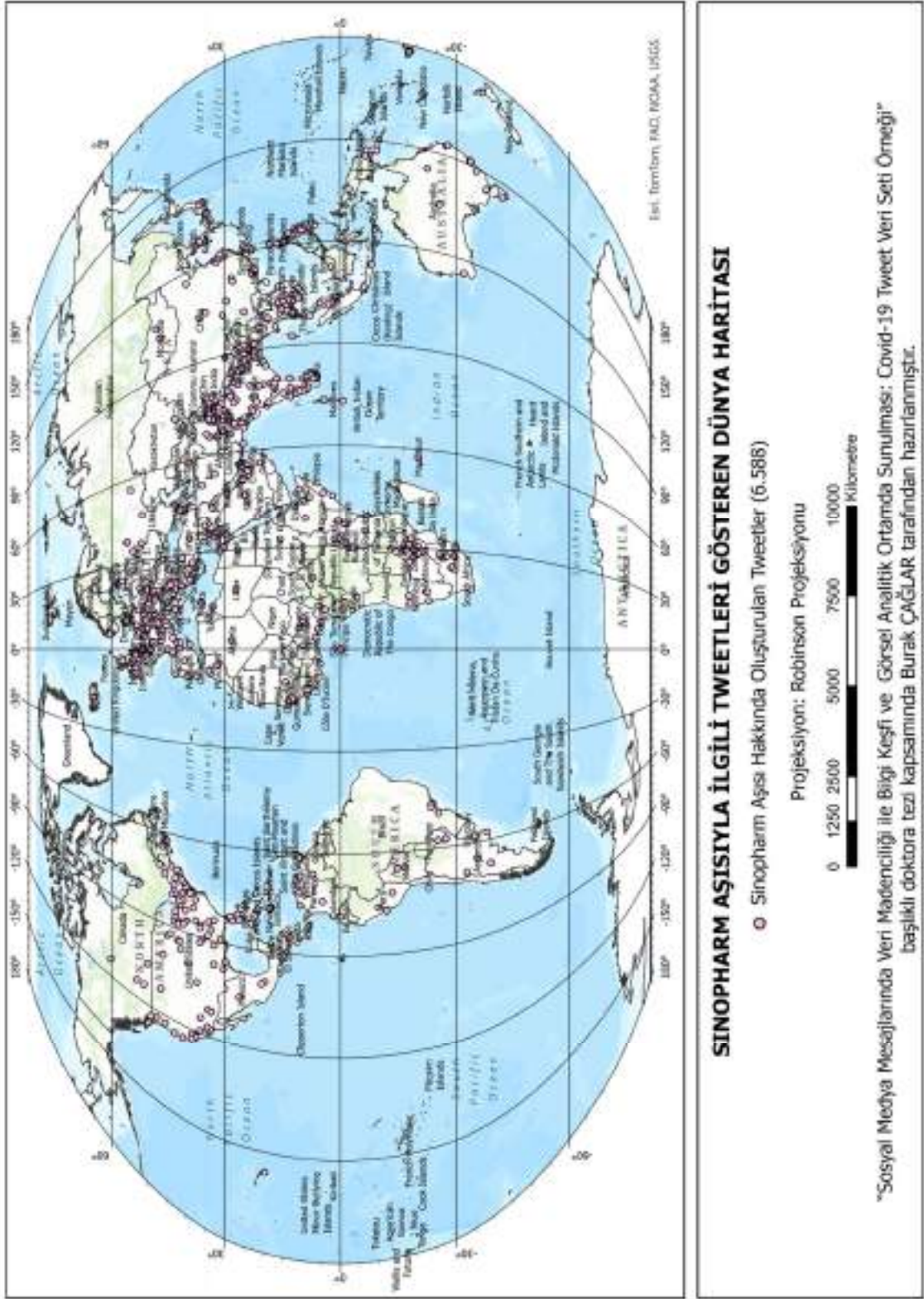
```

```
    }  
  
    function addGraphic(response) {  
      if (!response.location) {  
        return;  
      }  
      view.graphics.removeAll();  
      view.graphics.add(  
        new Graphic({  
          symbol: {  
            type: "simple-marker",  
            outline: {  
              color: "white",  
              width: 1.5,  
            },  
            color: "black",  
            size: 6,  
          },  
          geometry: response.location,  
        })  
      );  
    }  
  });  
</script>  
</head>  
<body>  
<div id="viewDiv"></div>  
</body>  
</html>
```

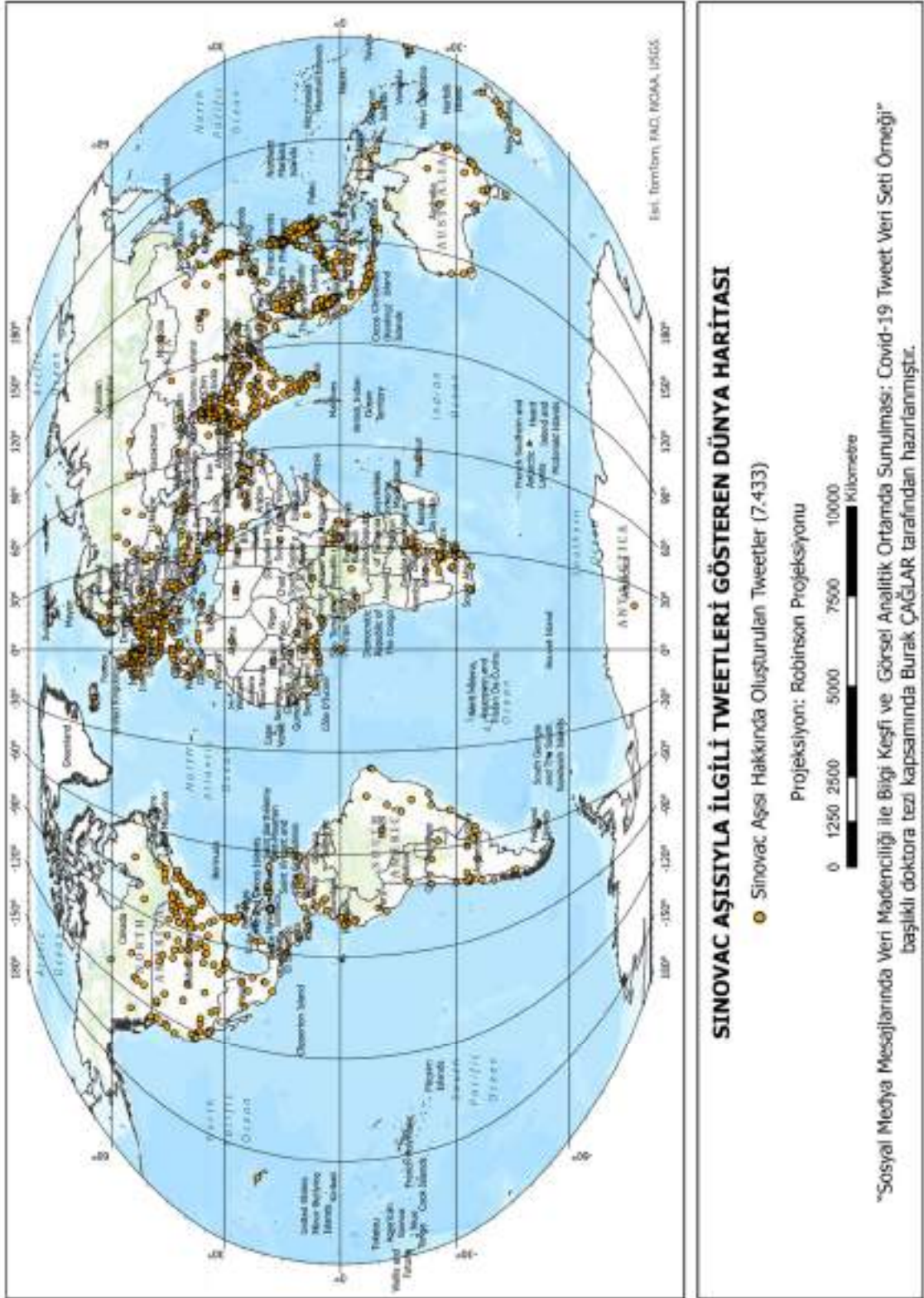
Ek-2.1 Pfizer Biontech Aşısı Hakkında Oluşturulan Tweetlerin Dünya Haritası Üzerinde Yayılımı



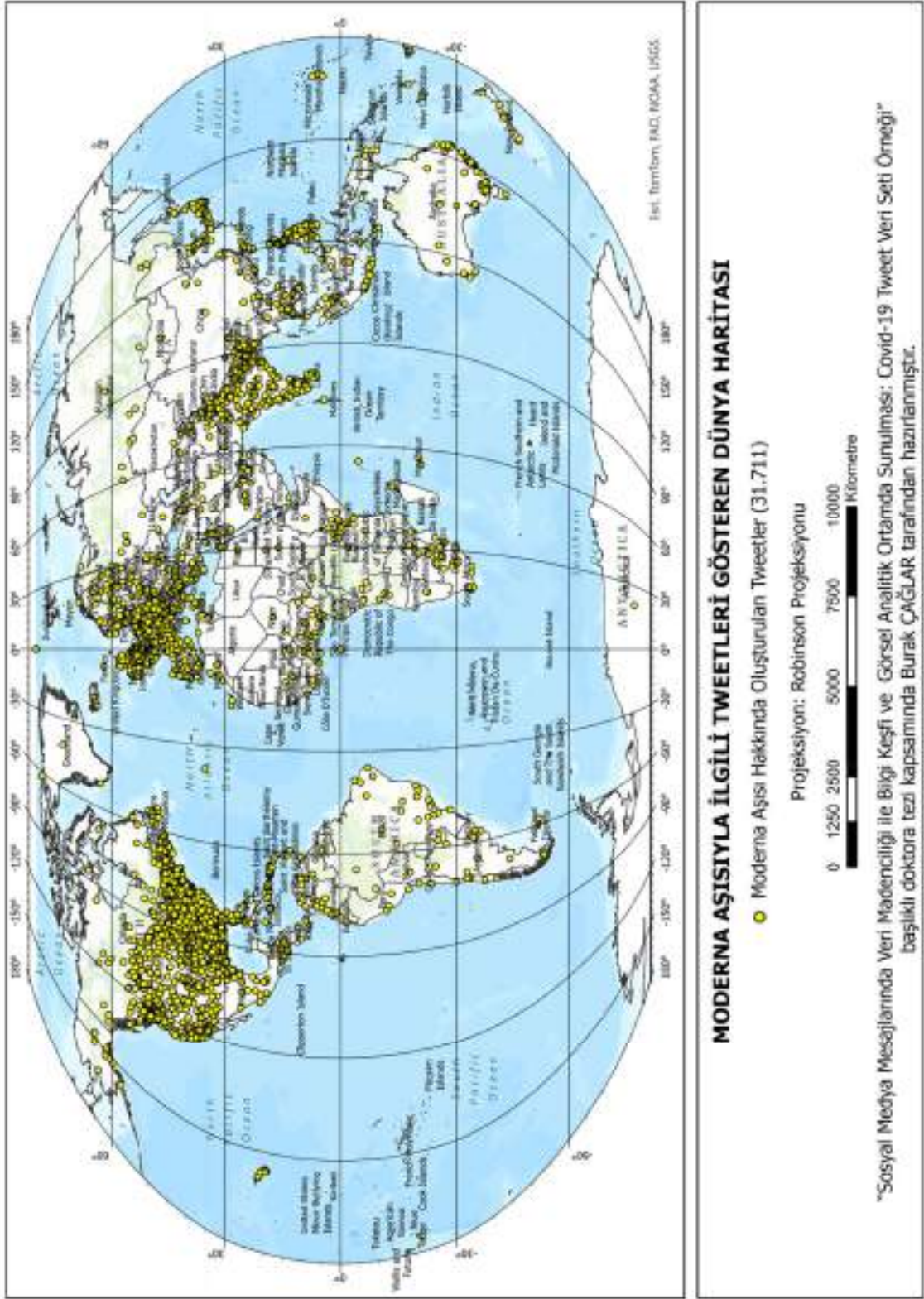
Ek-2.2 Sinopharm Aşısı Hakkında Oluşturulan Tweetlerin Dünya Haritası Üzerinde Yayılımı



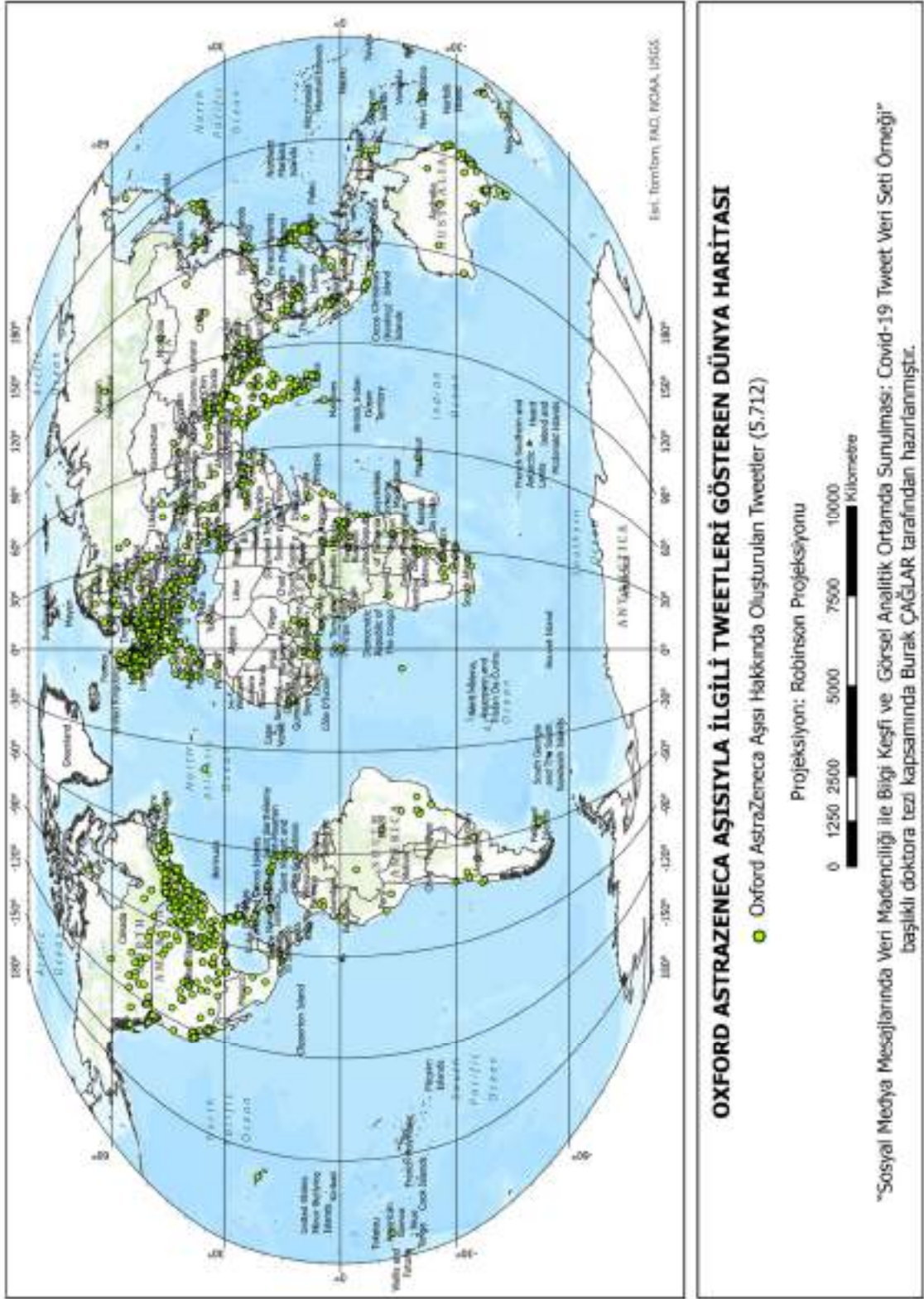
Ek-2.3 Sinovac Aşısı Hakkında Oluşturulan Tweetlerin Dünya Haritası Üzerinde Yayılımı



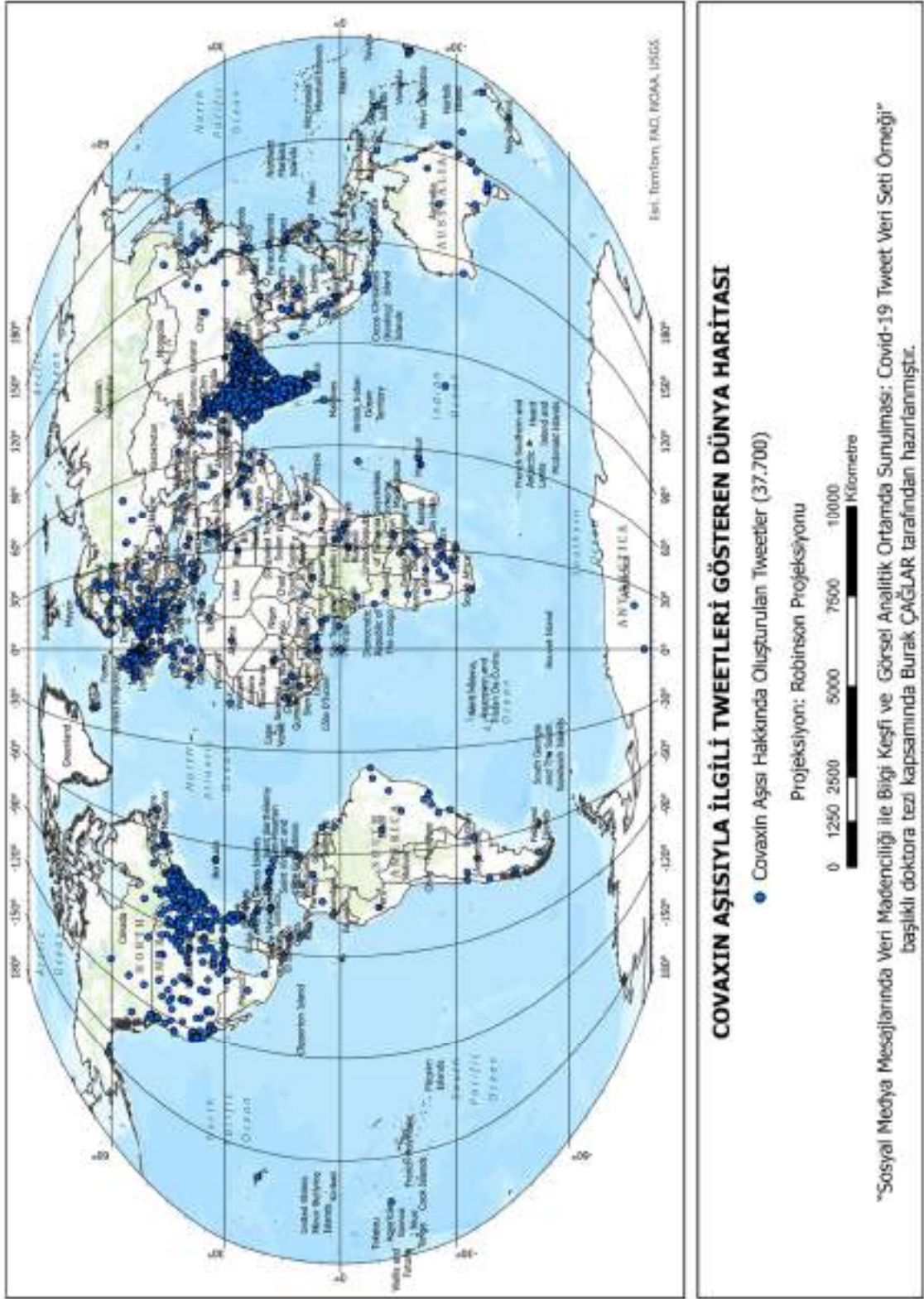
Ek-2.4 Moderna Aşısı Hakkında Oluşturulan Tweetlerin Dünya Haritası Üzerinde Yayılımı



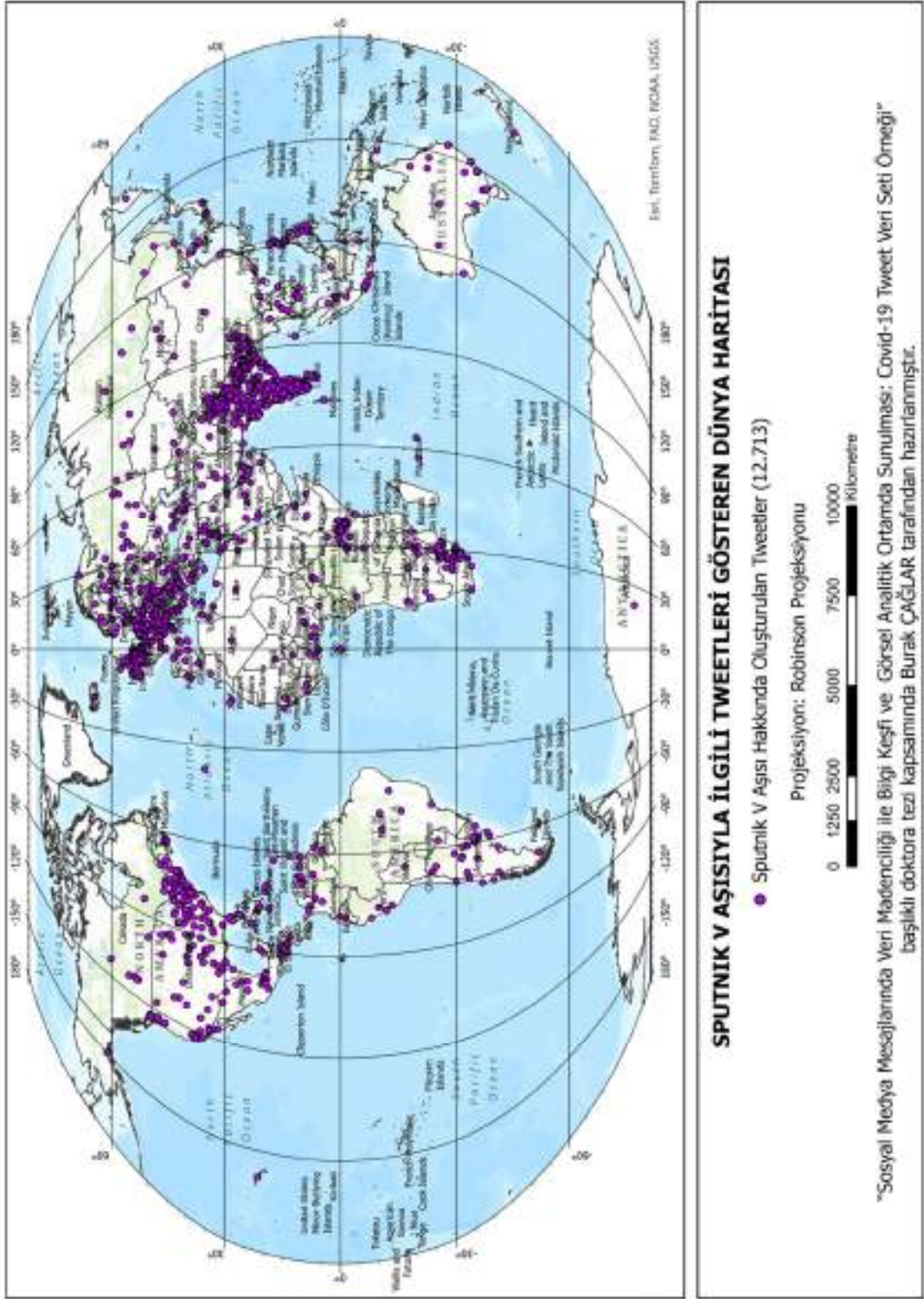
Ek-2.5 Oxford AstraZeneca Aşısı Hakkında Oluşturulan Tweetlerin Dünya Haritası Üzerinde Yayılımı



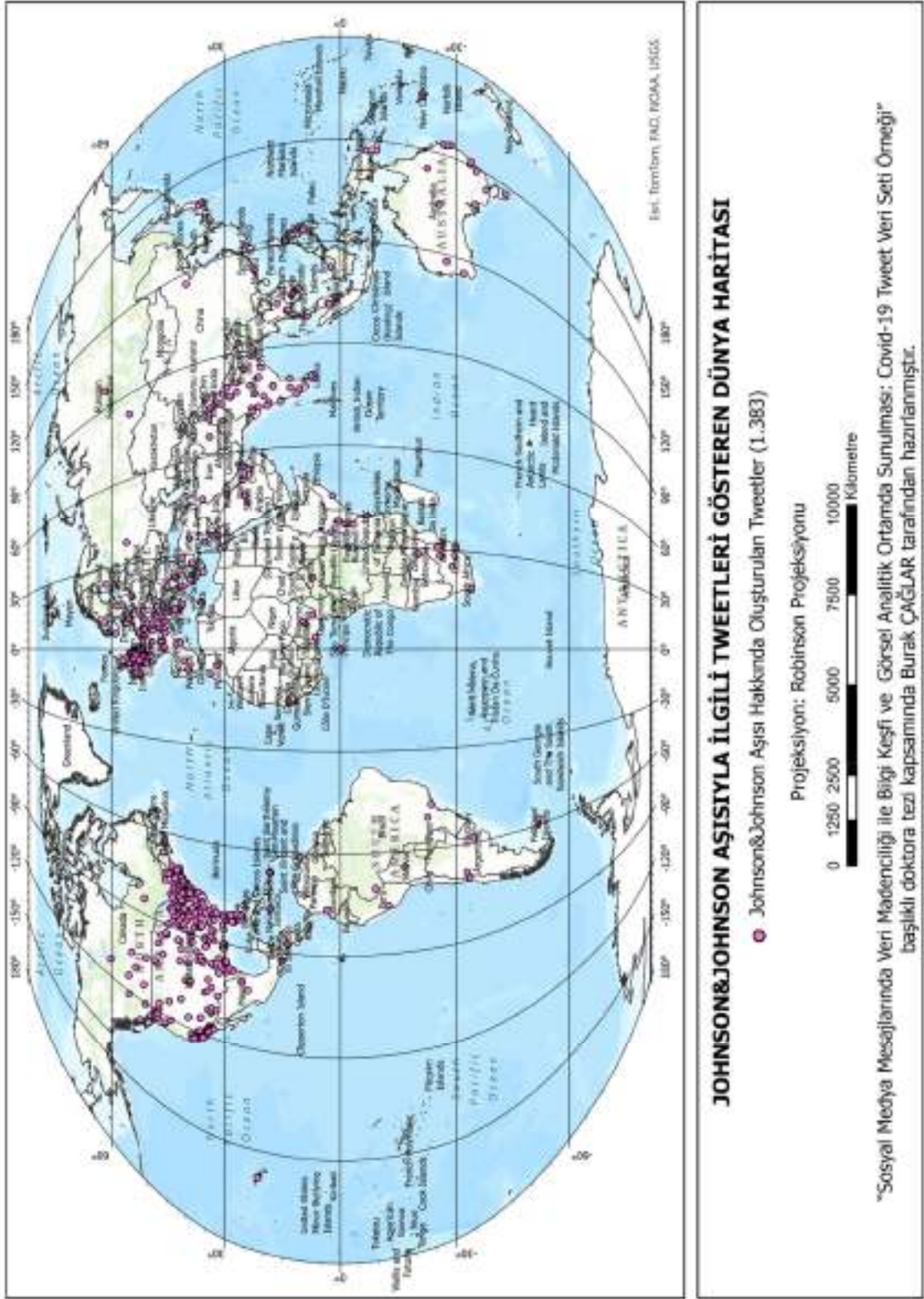
Ek-2.6 Covaxin Aşısı Hakkında Oluşturulan Tweetlerin Dünya Haritası Üzerinde Yayılımı



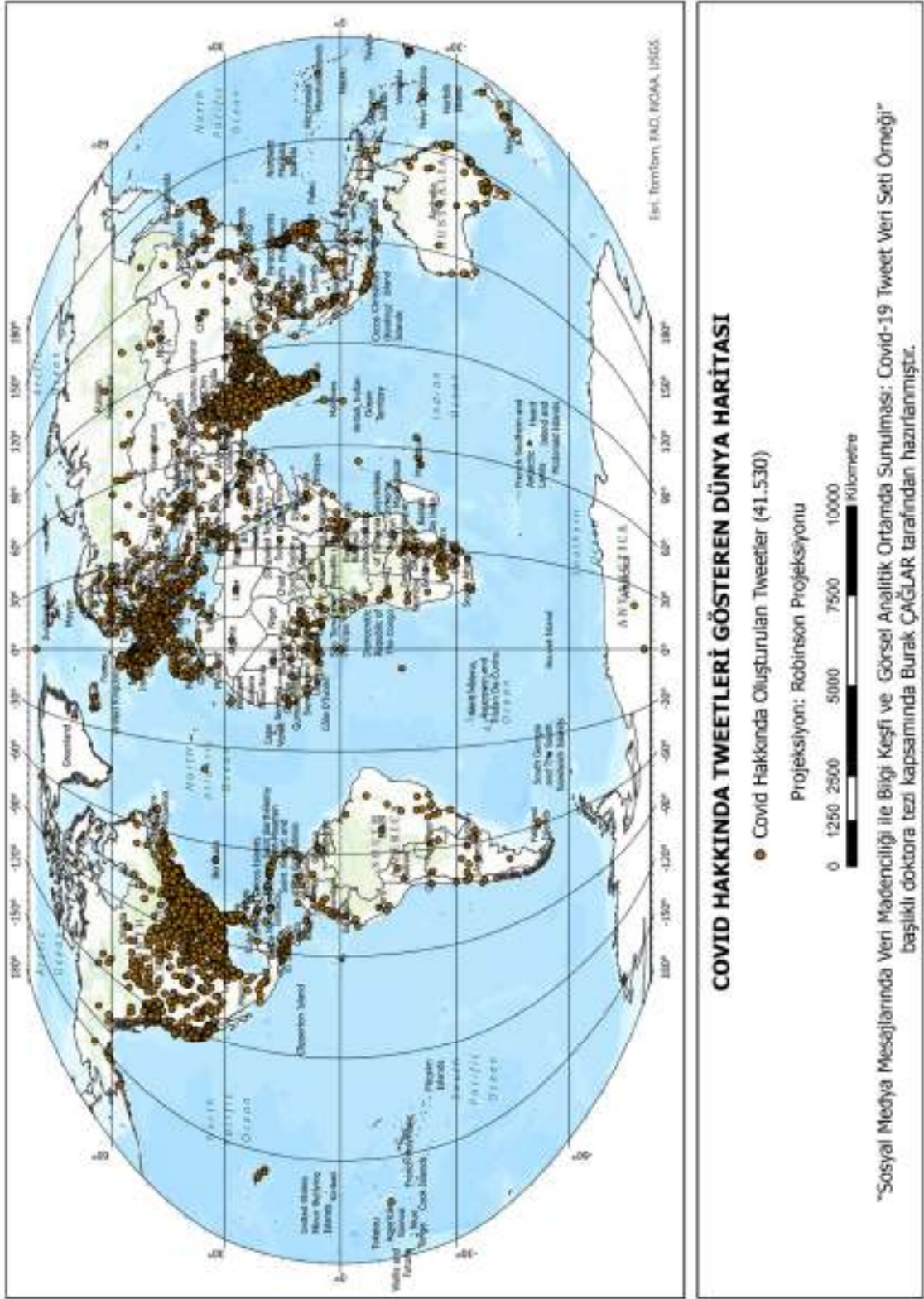
Ek-2.7 Sputnik V Aşısı Hakkında Oluşturulan Tweetlerin Dünya Haritası Üzerinde Yayılımı



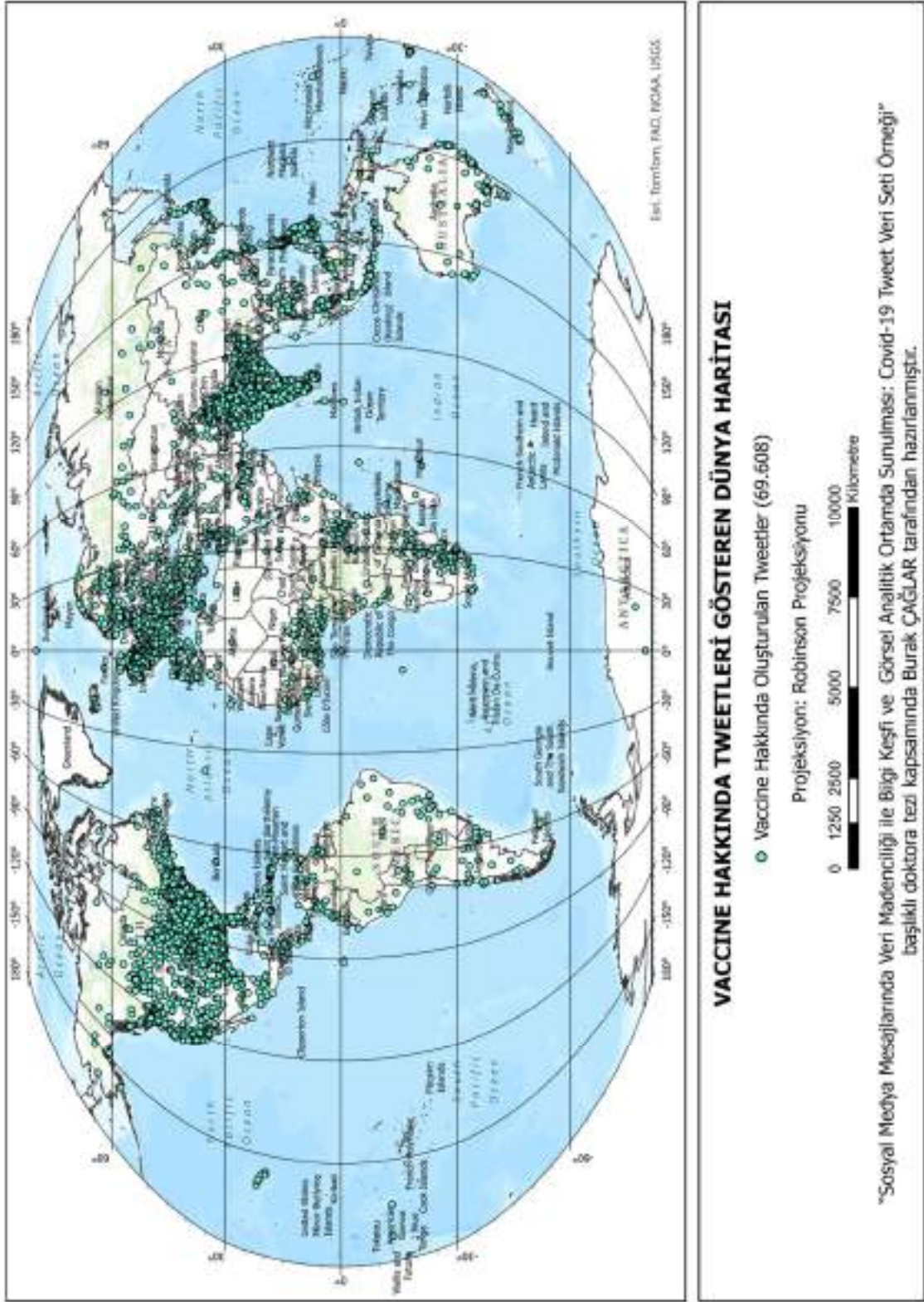
Ek-2.8 Johnson&Johnson Aşısı Hakkında Oluşturulan Tweetlerin Dünya Haritası Üzerinde Yayılımı



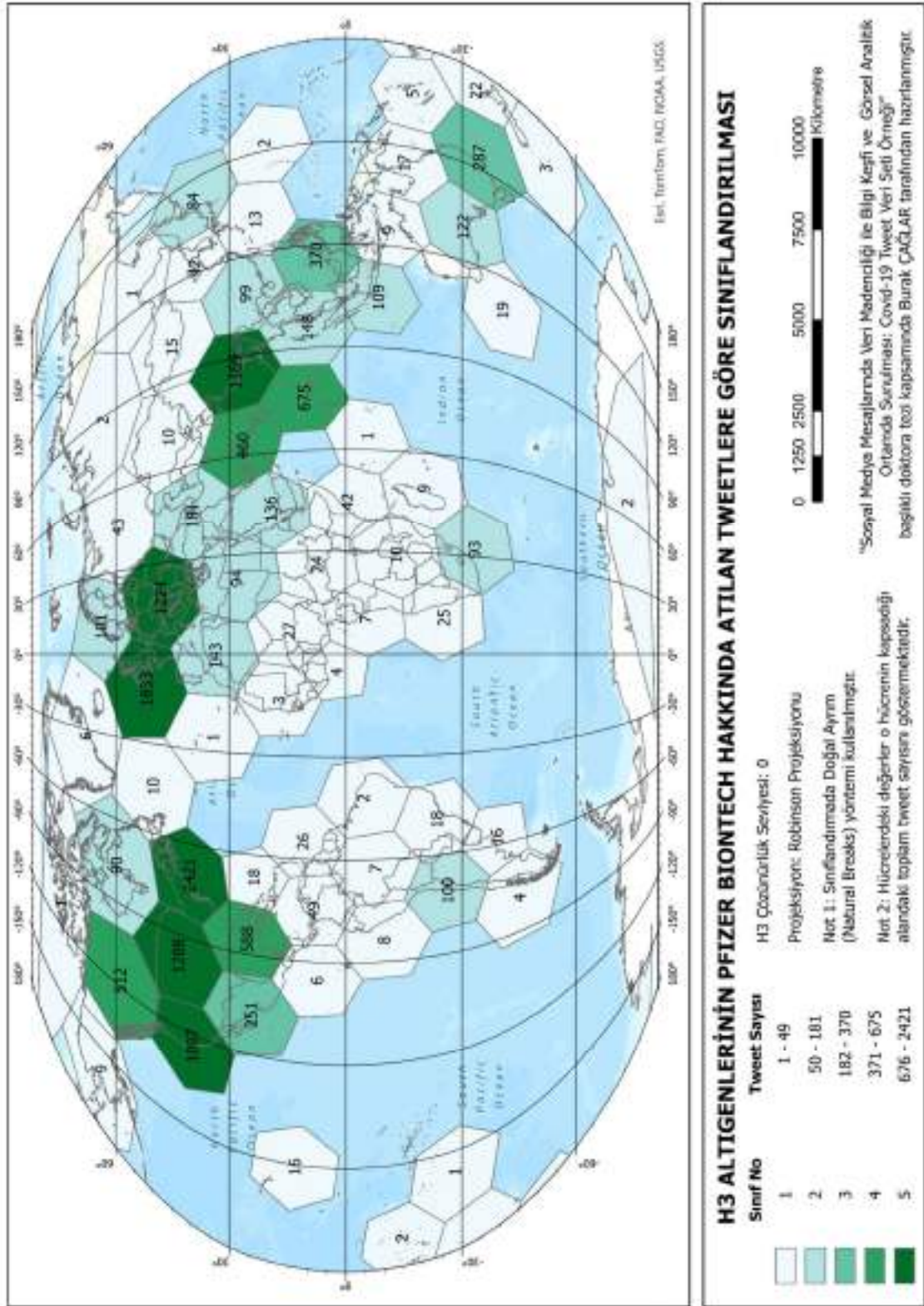
Ek-2.9 Covid Kelimesi Hakkında Oluşturulan Tweetlerin Dünya Haritası Üzerinde Yayılımı



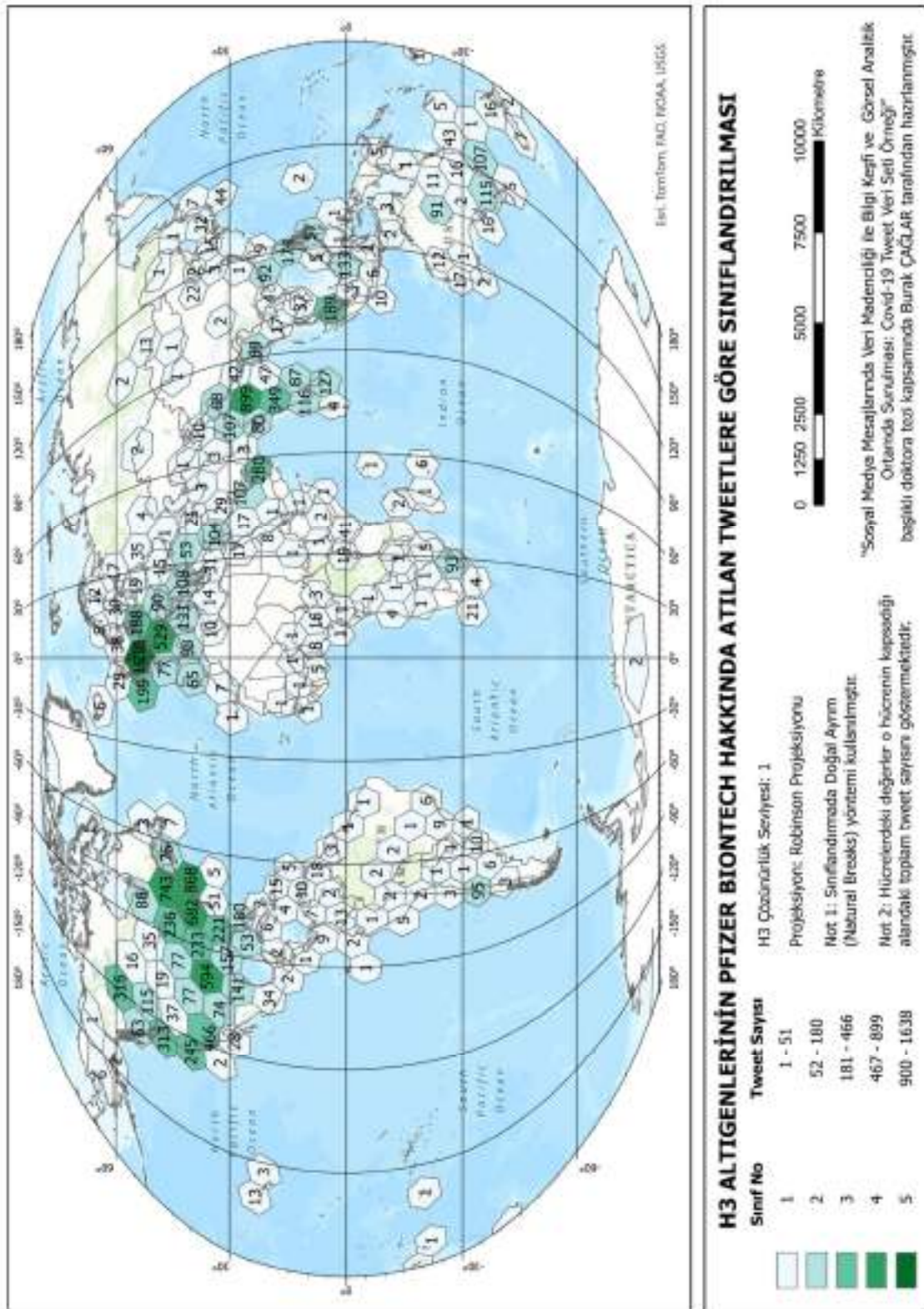
Ek-2.10 Vaccine Kelimesi Hakkında Oluşturulan Tweetlerin Dünya Haritası Üzerinde Yayılımı



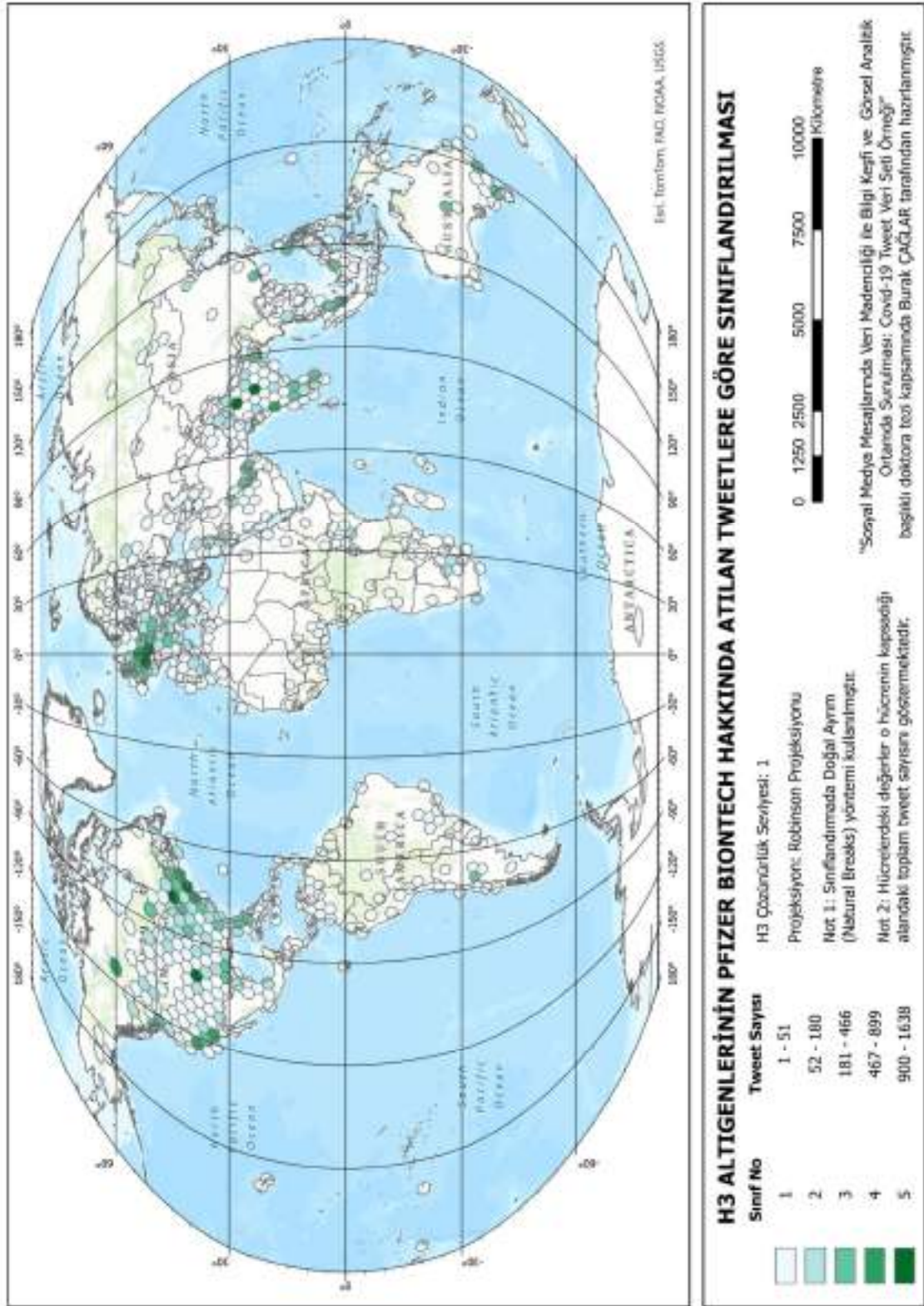
Ek-3.1 Pfizer Biontech Aşısı Hakkındaki Tweetleri Kapsayan Çözünürlük=0 Seviyesindeki H3 Hiyerarşik Altıgen Hücreleri



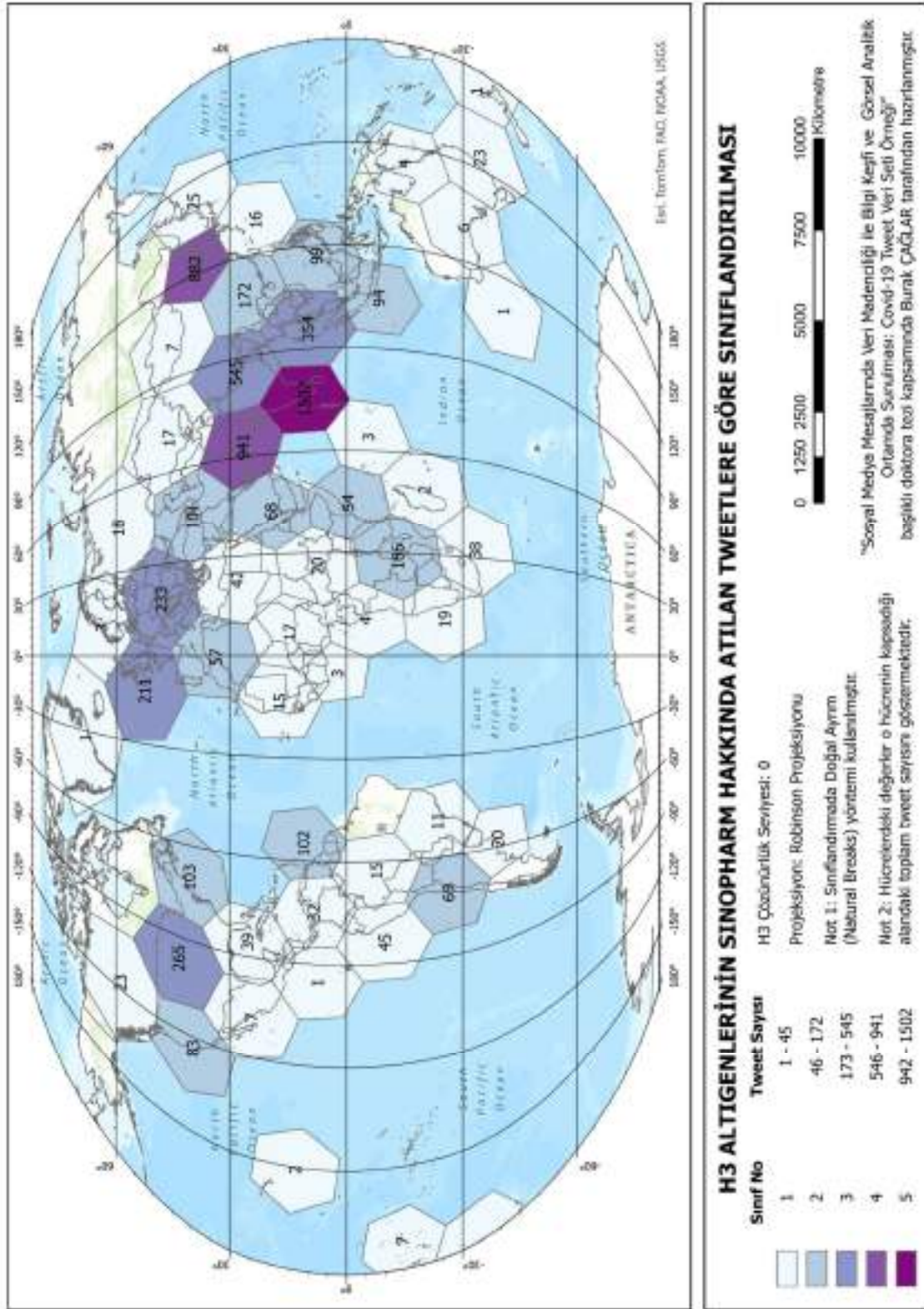
Ek-3.2 Pfizer Biontech Aşısı Hakkındaki Tweetleri Kapsayan Çözünürlük=1 Seviyesindeki H3 Hiyerarşik Altıgen Hücreleri



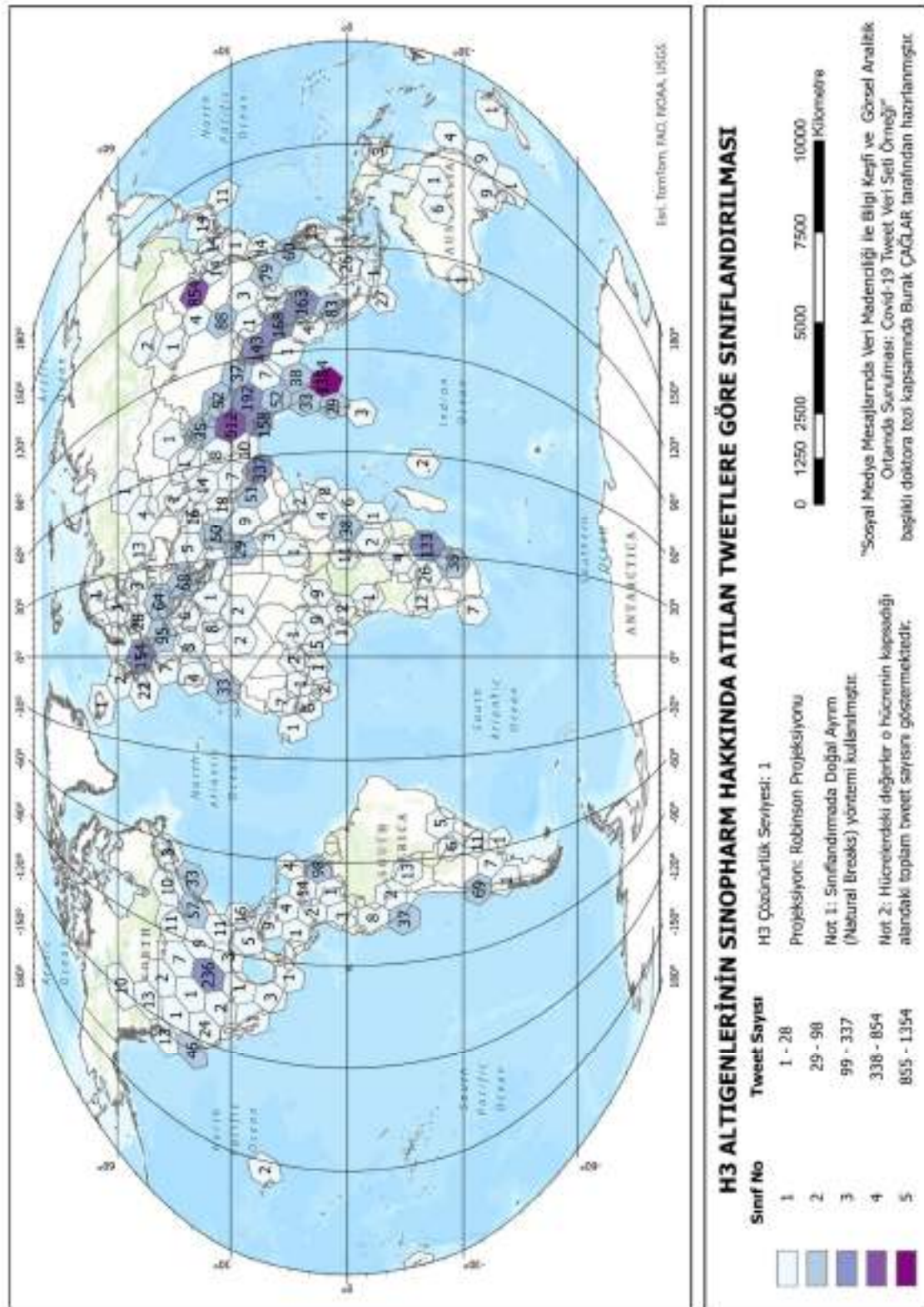
Ek-3.3 Pfizer Biontech Aşısı Hakkındaki Tweetleri Kapsayan Çözünürlük=2 Seviyesindeki H3 Hiyerarşik Altıgen Hücreleri



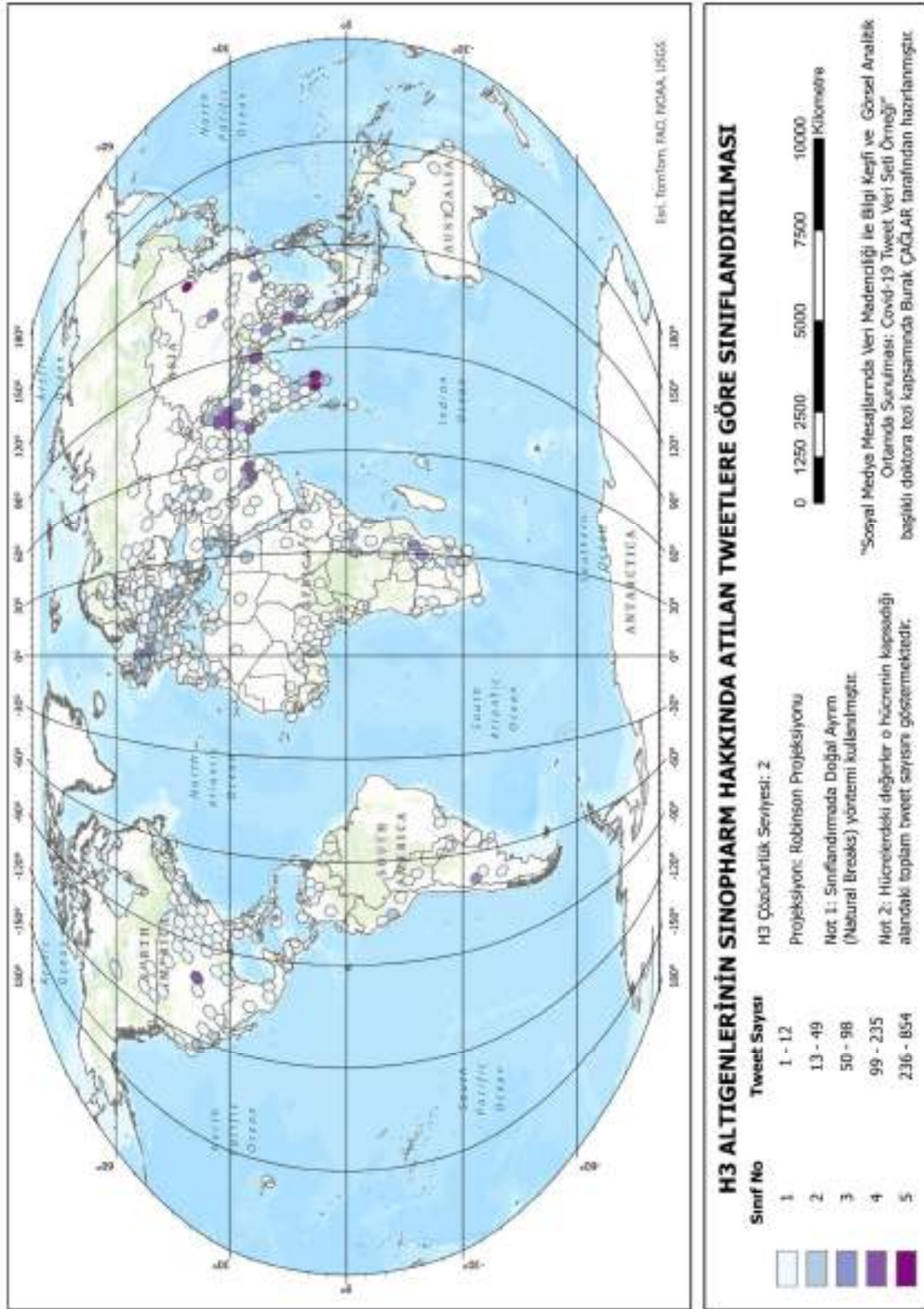
Ek-3.4 Sinopharm Aşısı Hakkındaki Tweetleri Kapsayan Çözünürlük=0 Seviyesindeki H3 Hiyerarşik Altıgen Hücreleri



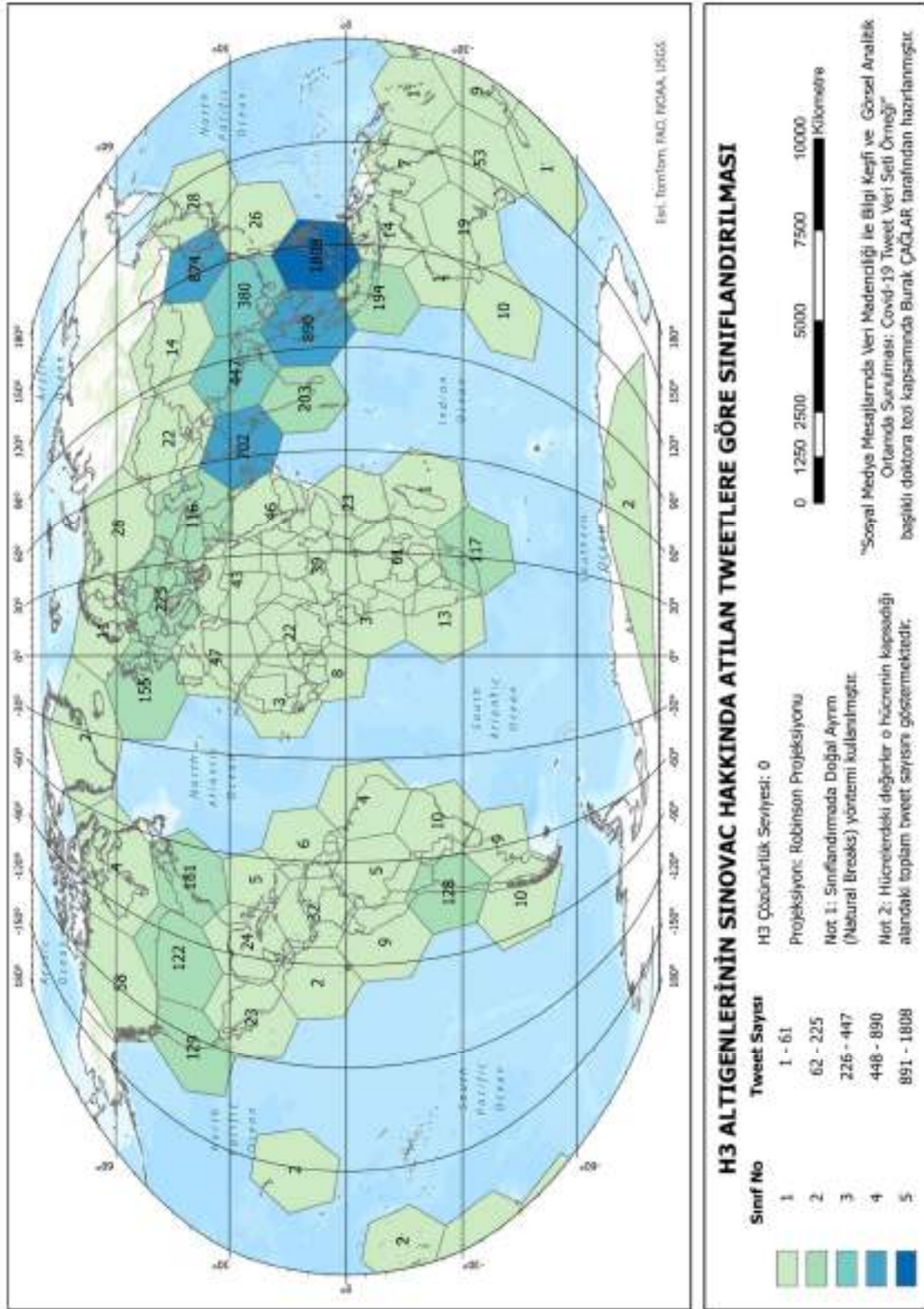
Ek-3.5 Sinopharm Aşısı Hakkındaki Tweetleri Kapsayan Çözünürlük=1 Seviyesindeki H3 Hiyerarşik Altıgen Hücreleri



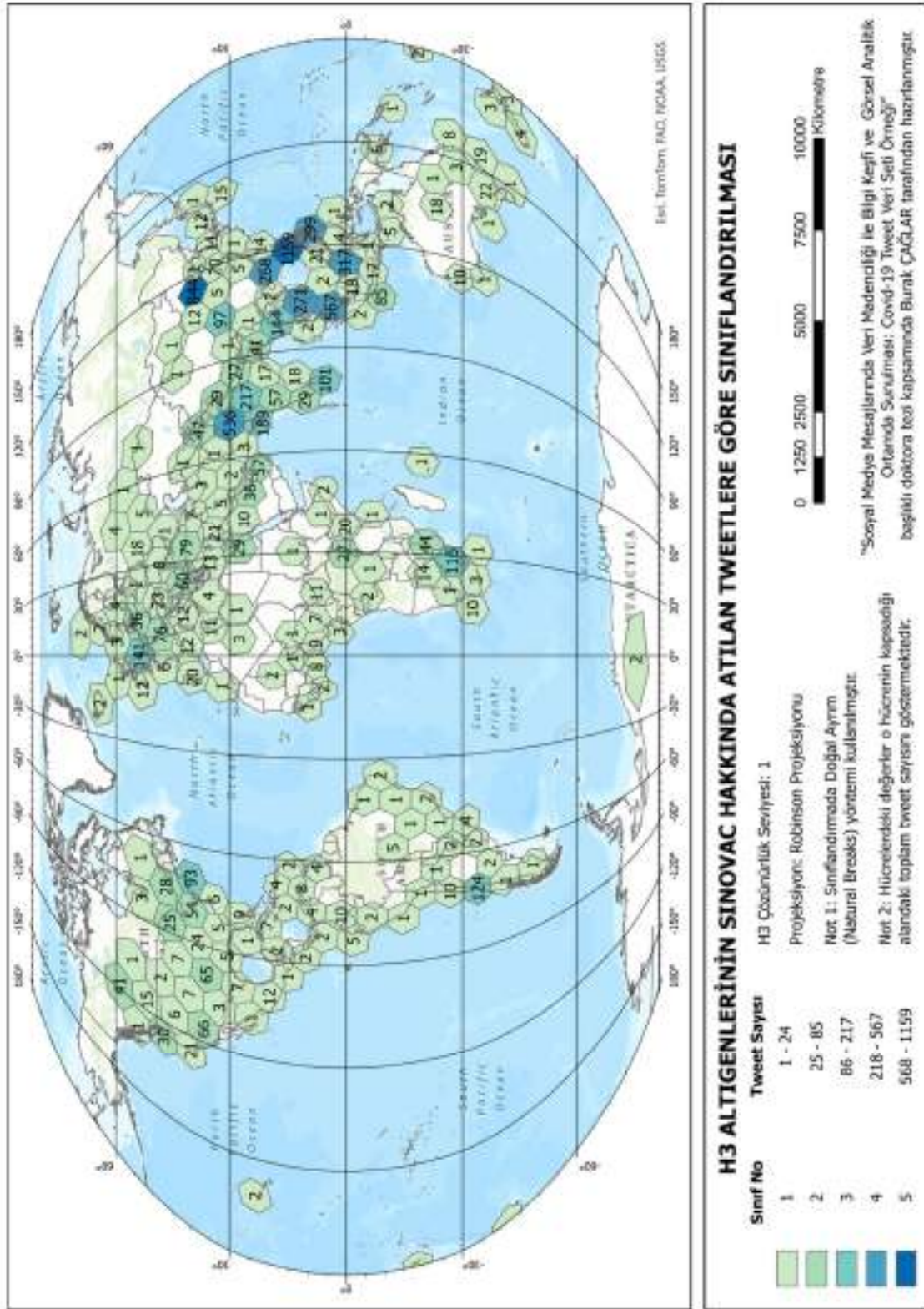
Ek-3.6 Sinopharm Aşısı Hakkındaki Tweetleri Kapsayan Çözünürlük=2 Seviyesindeki H3 Hiyerarşik Altıgen Hücreleri



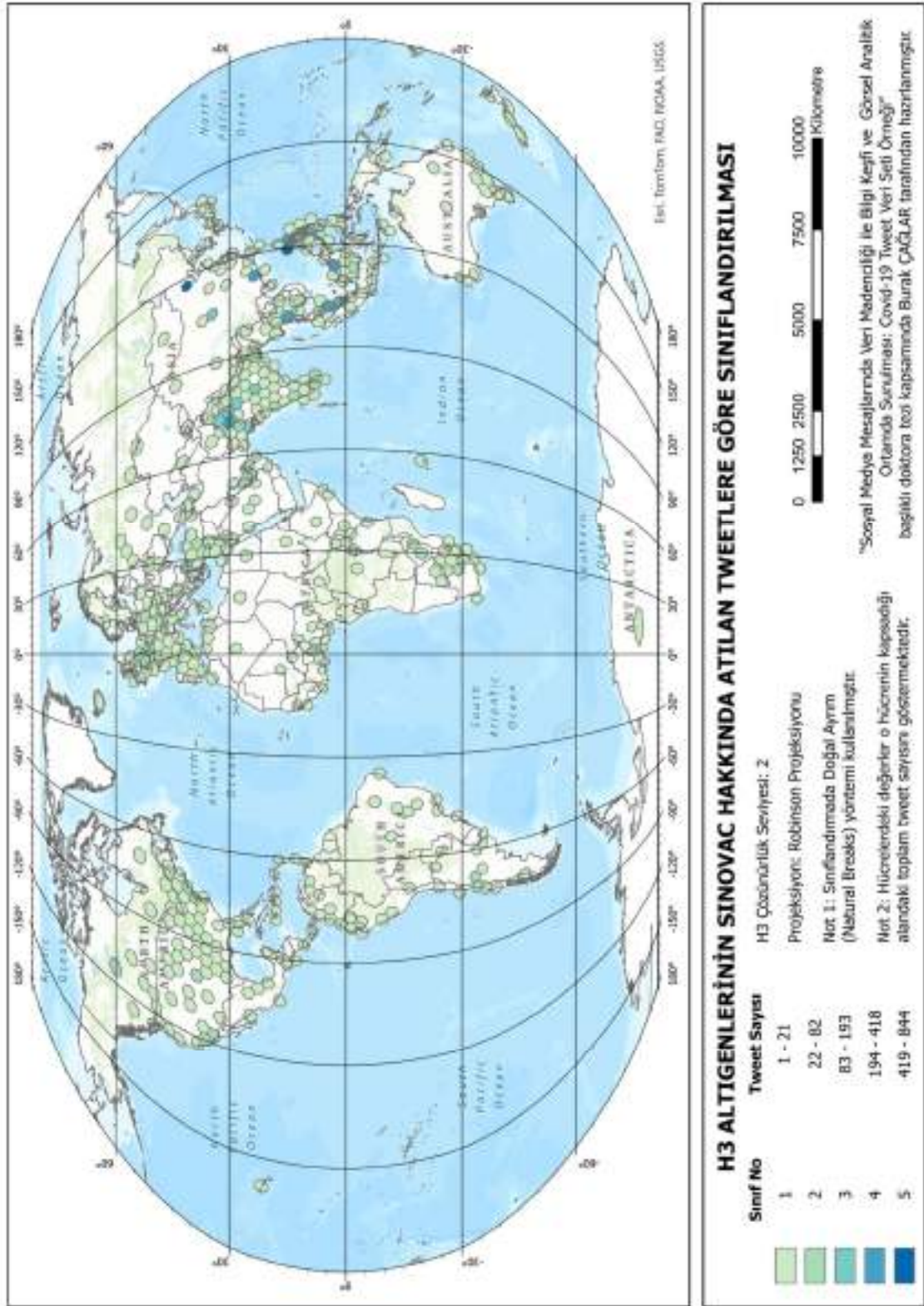
Ek-3.7 Sinovac Aşısı Hakkındaki Tweetleri Kapsayan Çözünürlük=0 Seviyesindeki H3 Hiyerarşik Altıgen Hücreleri



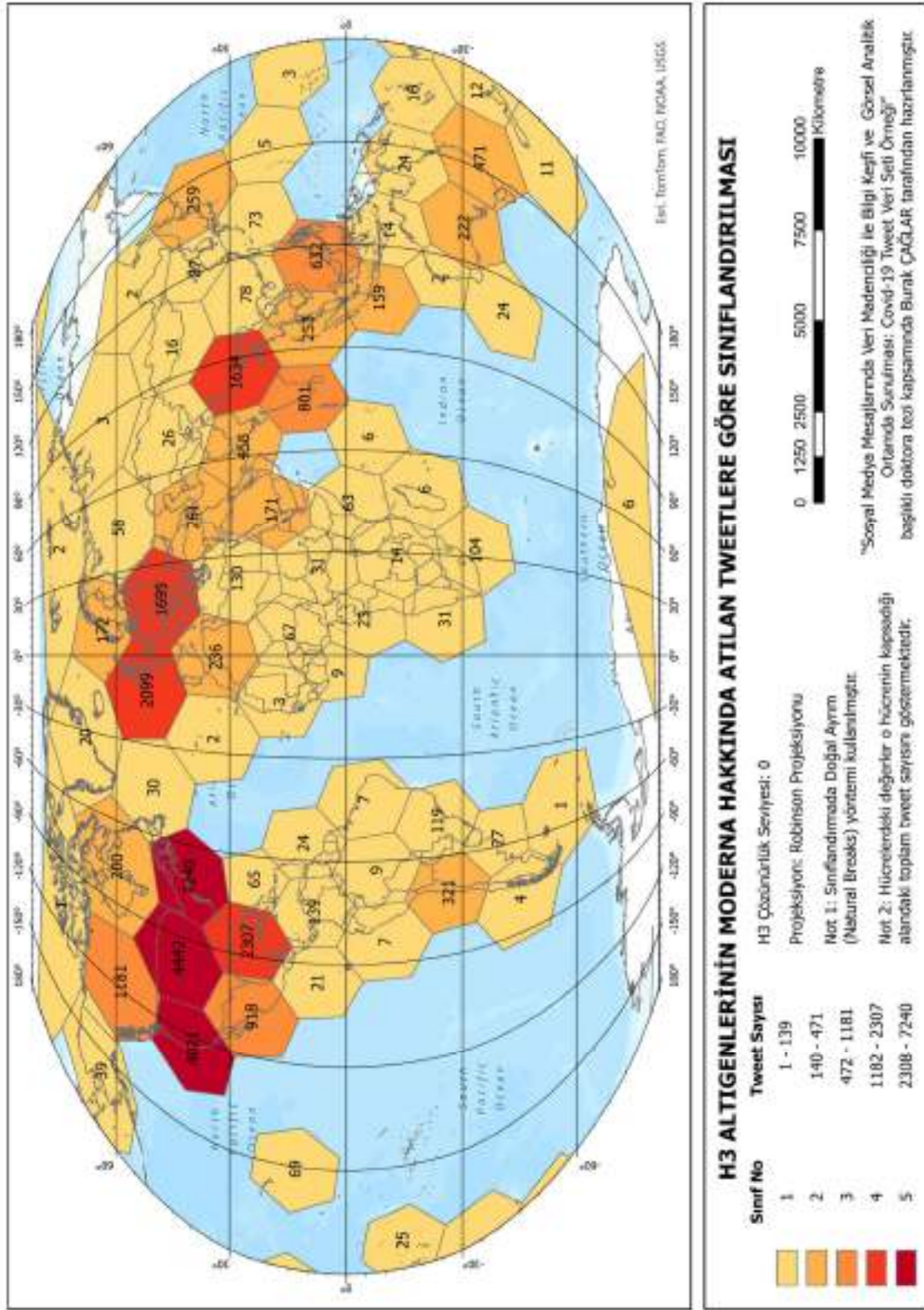
Ek-3.8 Sinovac Aşısı Hakkındaki Tweetleri Kapsayan Çözünürlük=1 Seviyesindeki H3 Hiyerarşik Altıgen Hücreleri



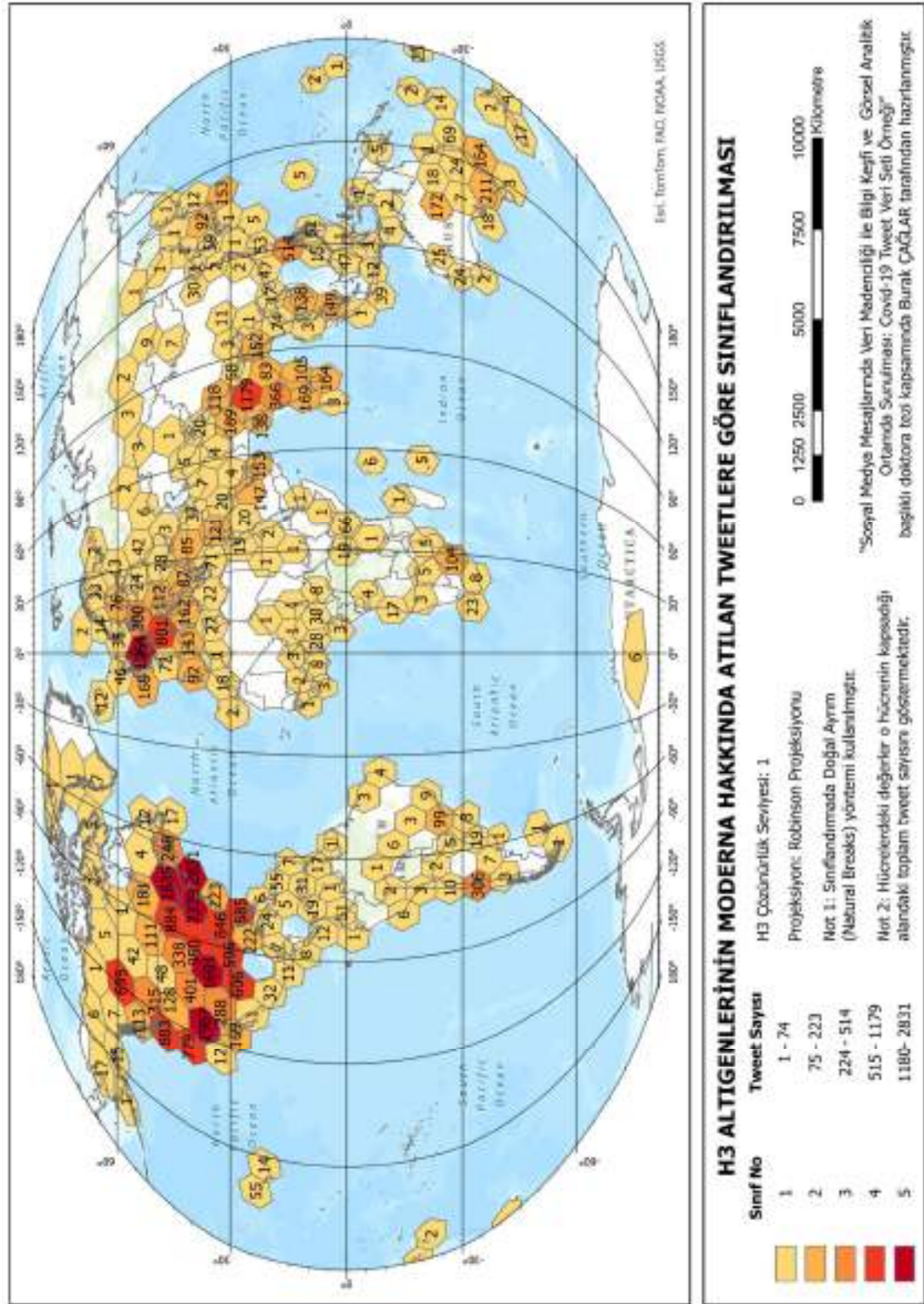
Ek-3.9 Sinovac Aşısı Hakkındaki Tweetleri Kapsayan Çözünürlük=2 Seviyesindeki H3 Hiyerarşik Altıgen Hücreleri



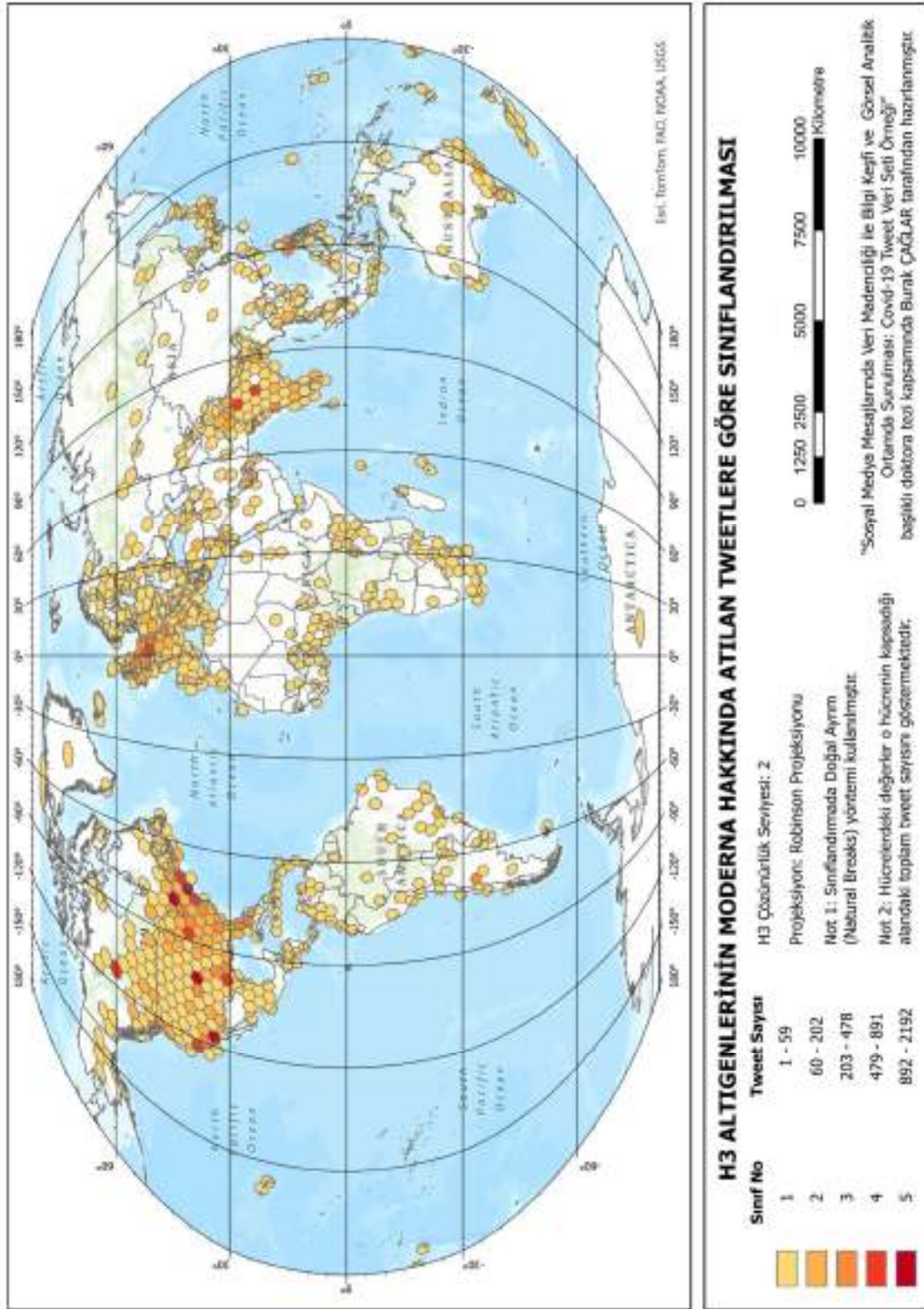
Ek-3.10 Moderna Aşısı Hakkındaki Tweetleri Kapsayan Çözünürlük=0 Seviyesindeki H3 Hiyerarşik Altıgen Hücreleri



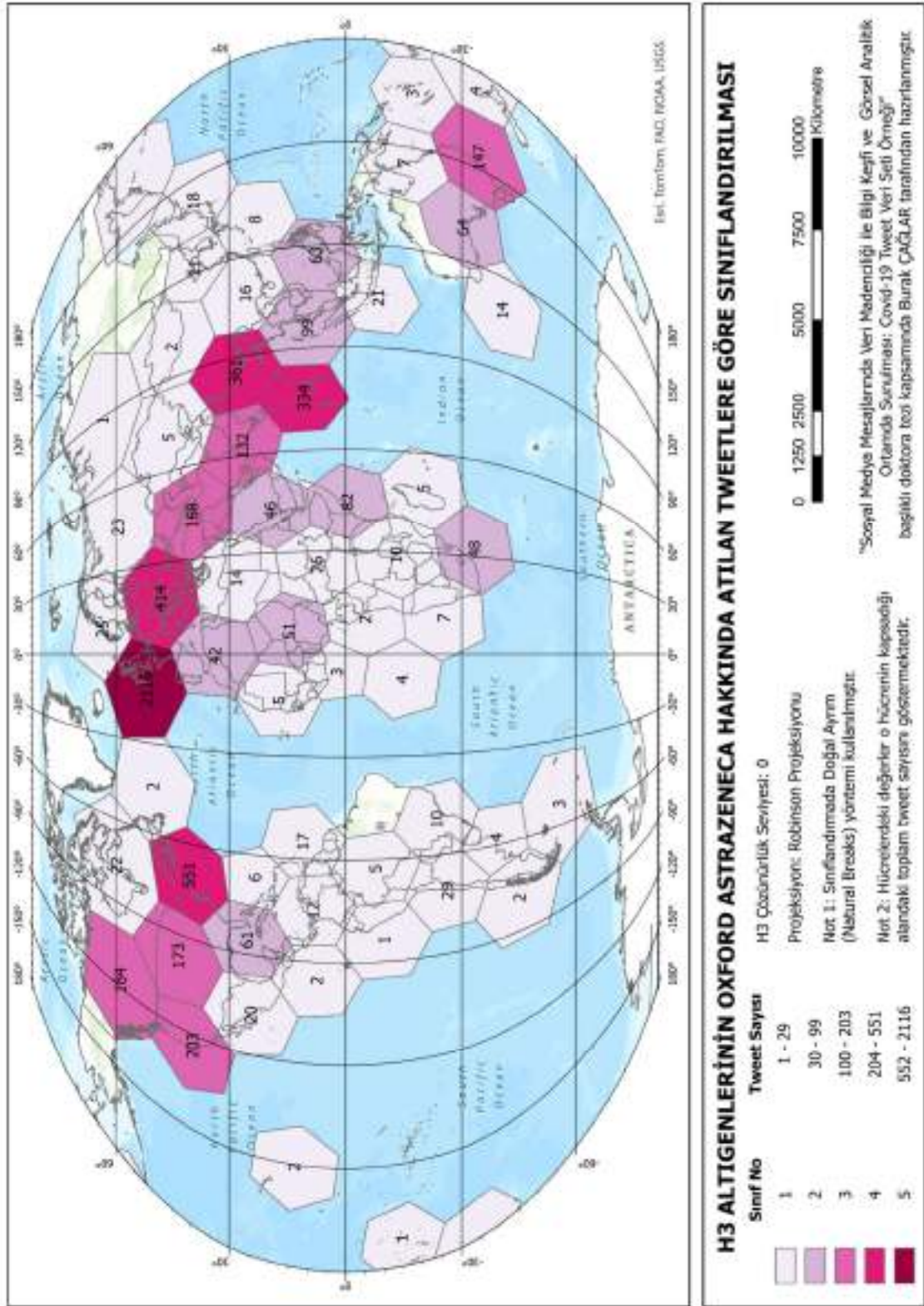
Ek-3.11 Moderna Aşısı Hakkındaki Tweetleri Kapsayan Çözünürlük=1 Seviyesindeki H3 Hiyerarşik Altıgen Hücreleri



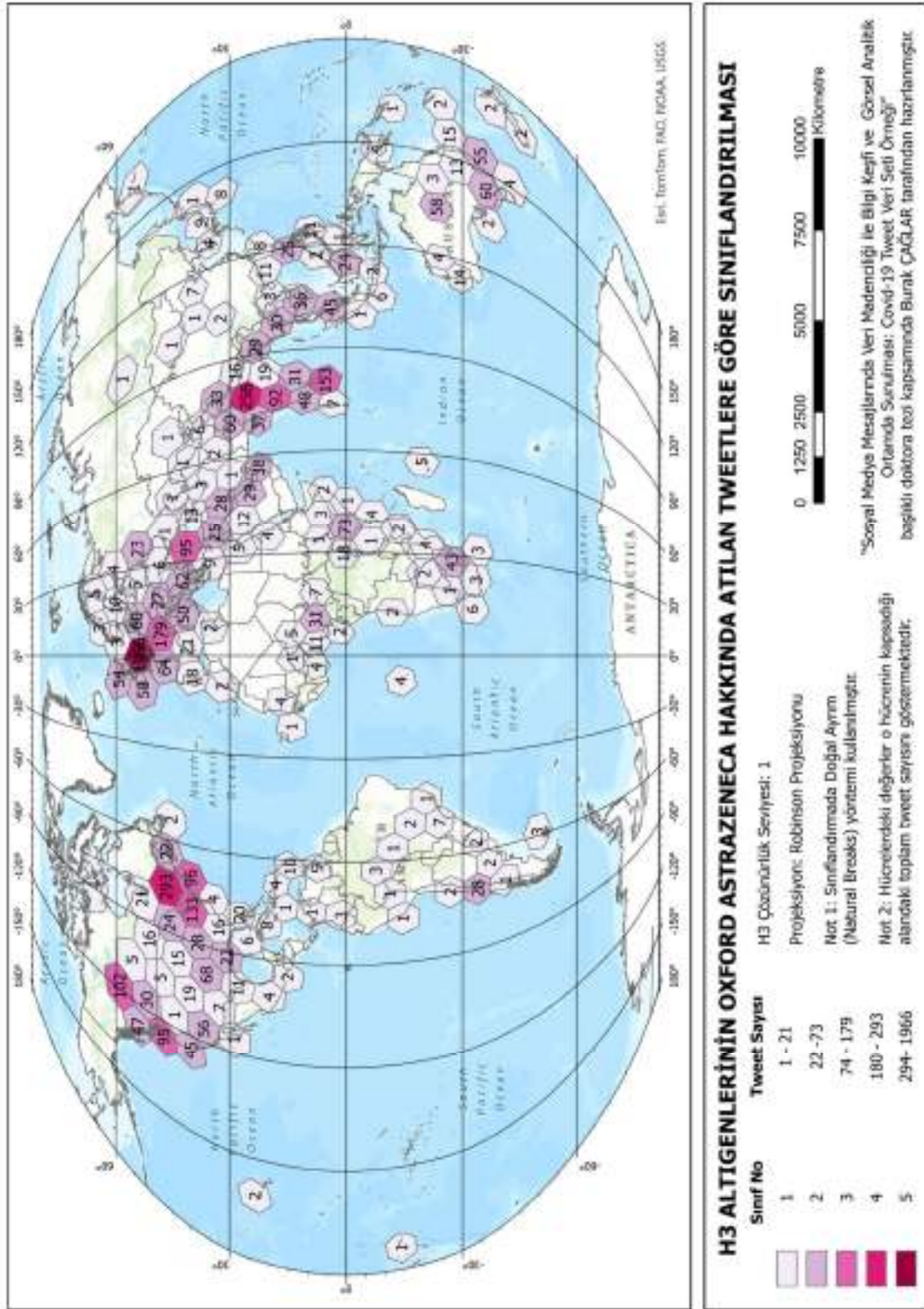
Ek-3.12 Moderna Aşısı Hakkındaki Tweetleri Kapsayan Çözünürlük=2 Seviyesindeki H3 Hiyerarşik Altıgen Hücreleri



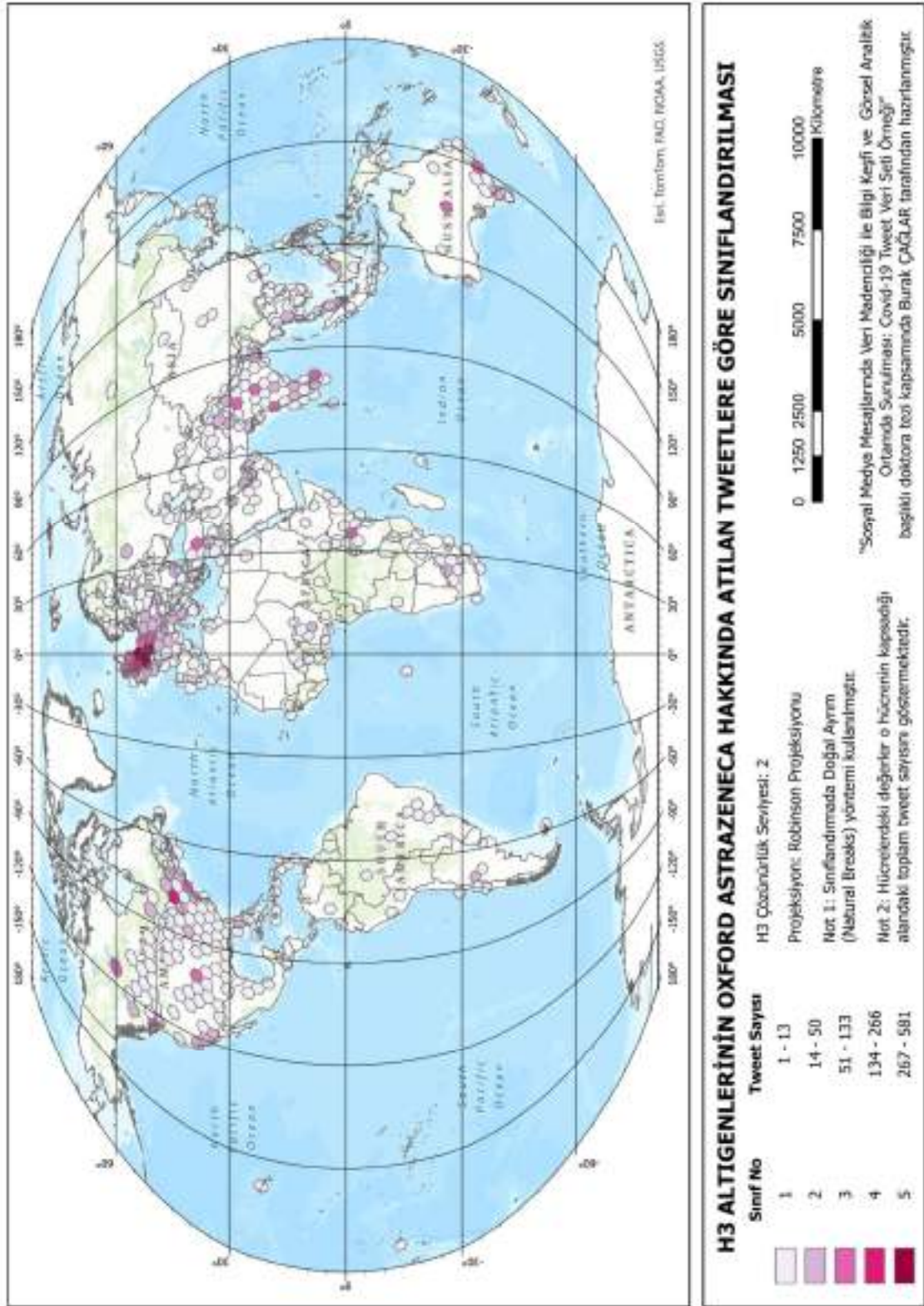
Ek-3.13 Oxford AstraZeneca Aşısı Hakkındaki Tweetleri Kapsayan Çözünürlük=0 Seviyesindeki H3 Hiyerarşik Altıgen Hücreleri



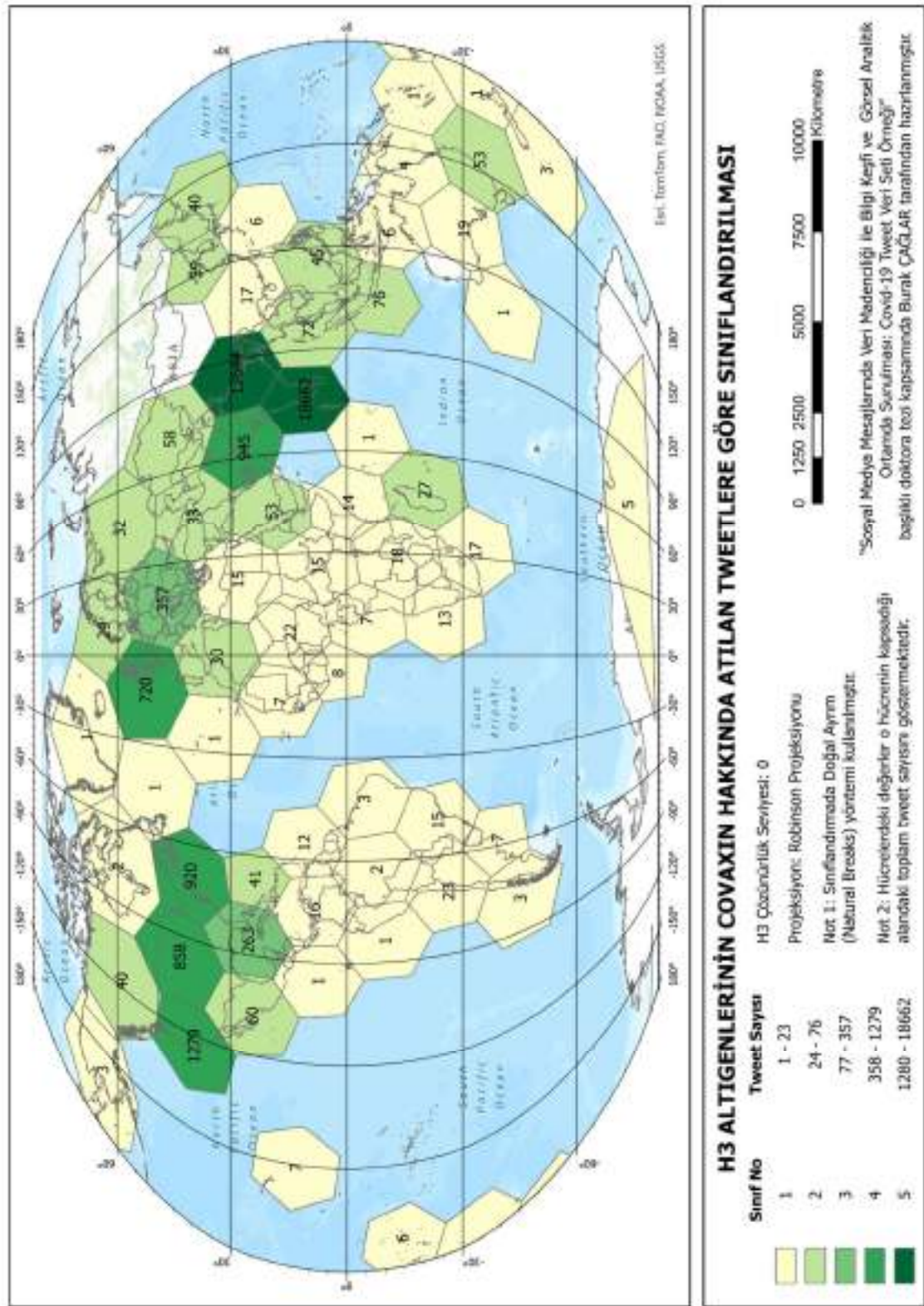
Ek-3.14 Oxford AstraZeneca Aşısı Hakkındaki Tweetleri Kapsayan Çözünürlük=1 Seviyesindeki H3 Hiyerarşik Altıgen Hücreleri



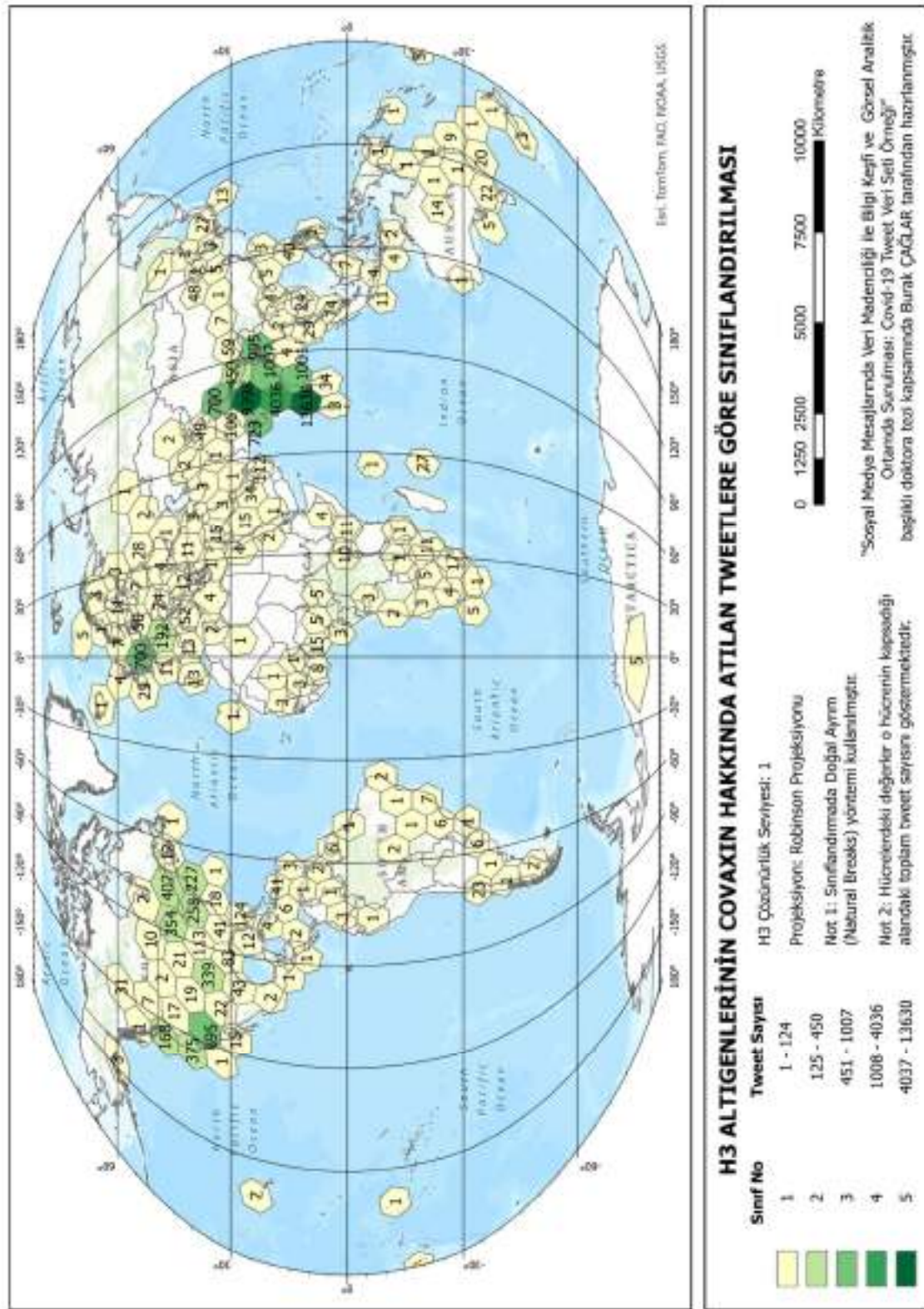
Ek-3.15 Oxford AstraZeneca Aşısı Hakkındaki Tweetleri Kapsayan Çözünürlük=2 Seviyesindeki H3 Hiyerarşik Altıgen Hücreleri



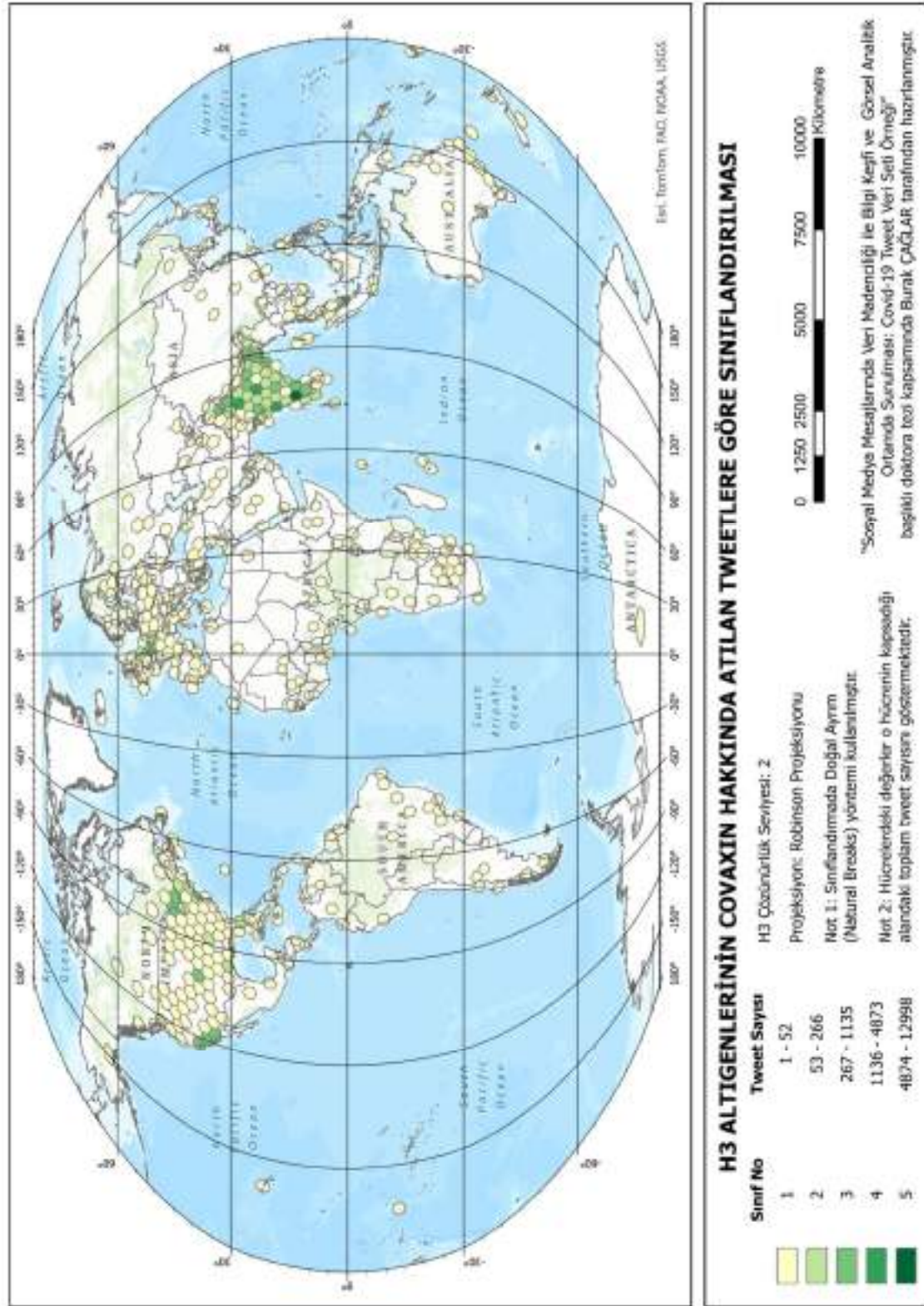
Ek-3.16 Covaxin Aşısı Hakkındaki Tweetleri Kapsayan Çözünürlük=0 Seviyesindeki H3 Hiyerarşik Altıgen Hücreleri



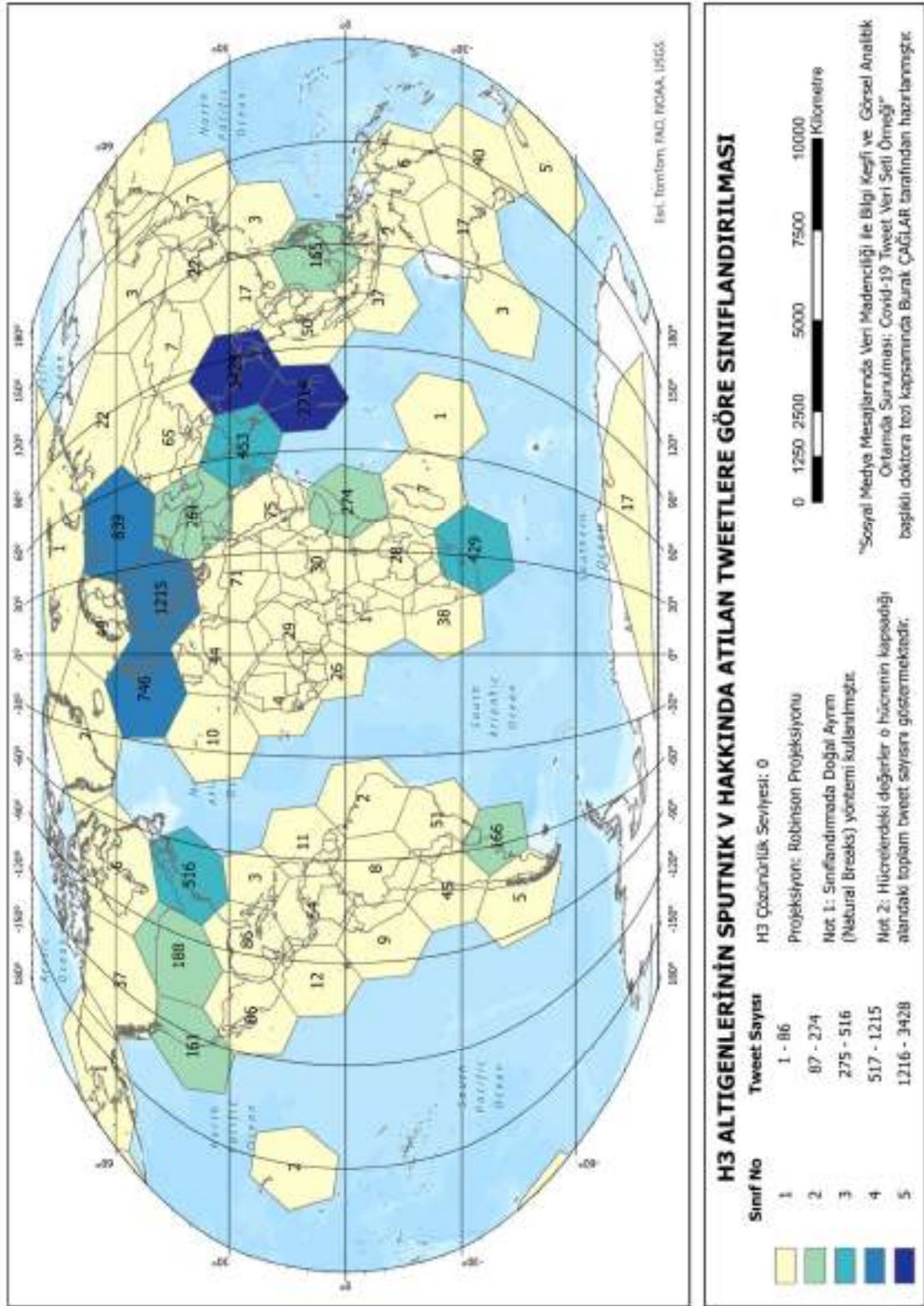
Ek-3.17 Covaxin Aşısı Hakkındaki Tweetleri Kapsayan Çözünürlük=1 Seviyesindeki H3 Hiyerarşik Altıgen Hücreleri



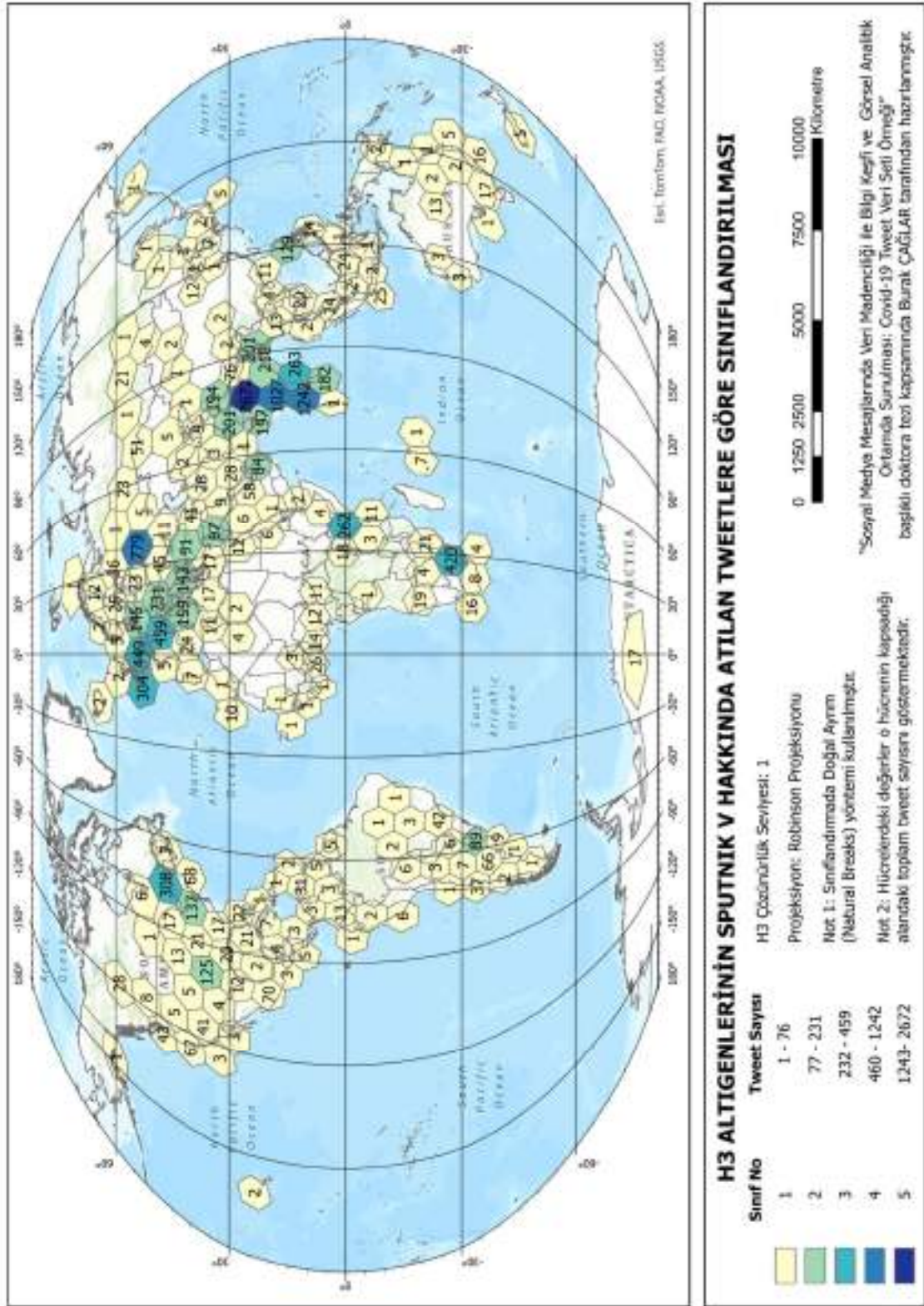
Ek-3.18 Covaxin Aşısı Hakkındaki Tweetleri Kapsayan Çözünürlük=2 Seviyesindeki H3 Hiyerarşik Altıgen Hücreleri



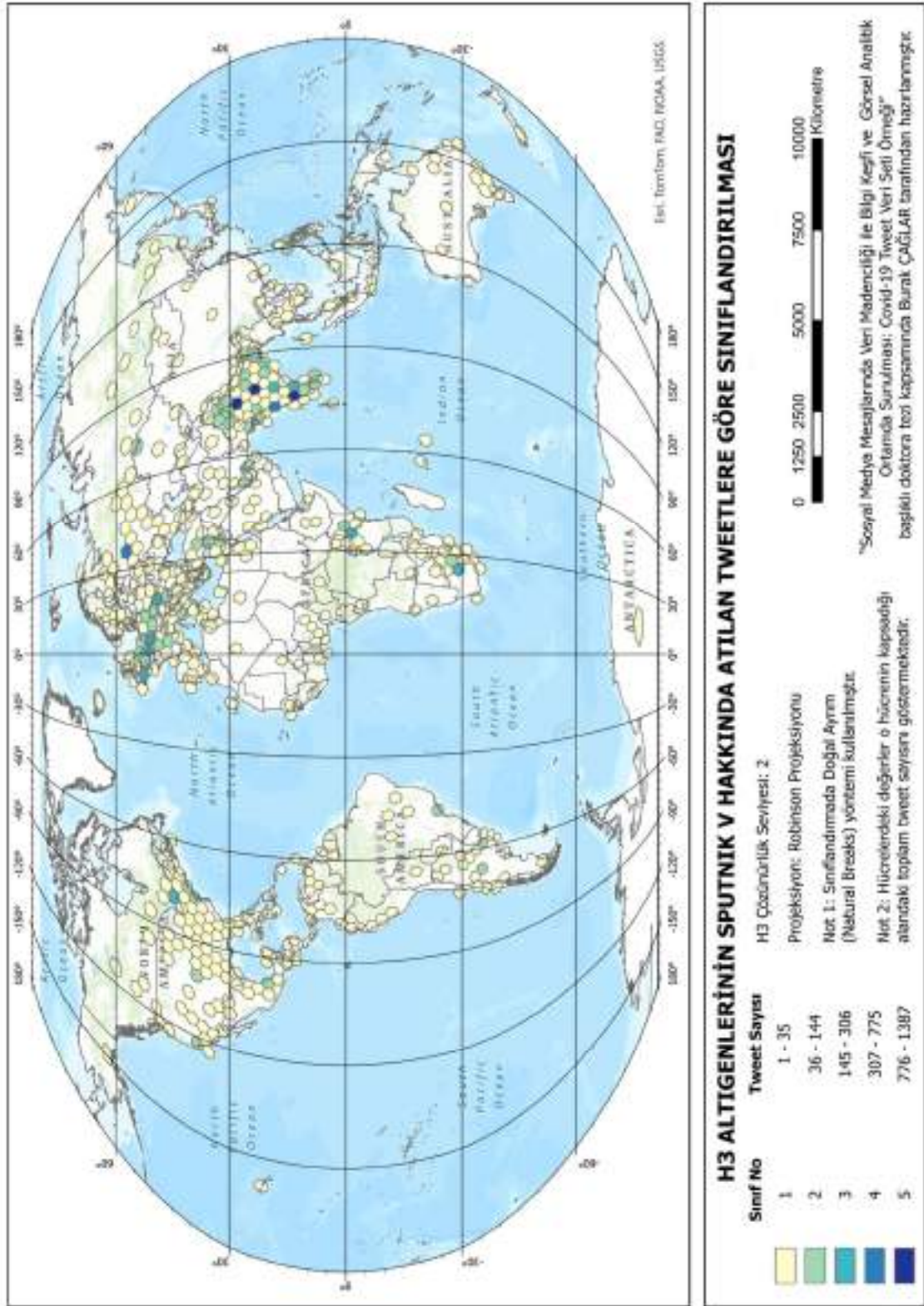
Ek-3.19 Sputnik V Aşısı Hakkındaki Tweetleri Kapsayan Çözünürlük=0 Seviyesindeki H3 Hiyerarşik Altıgen Hücreleri



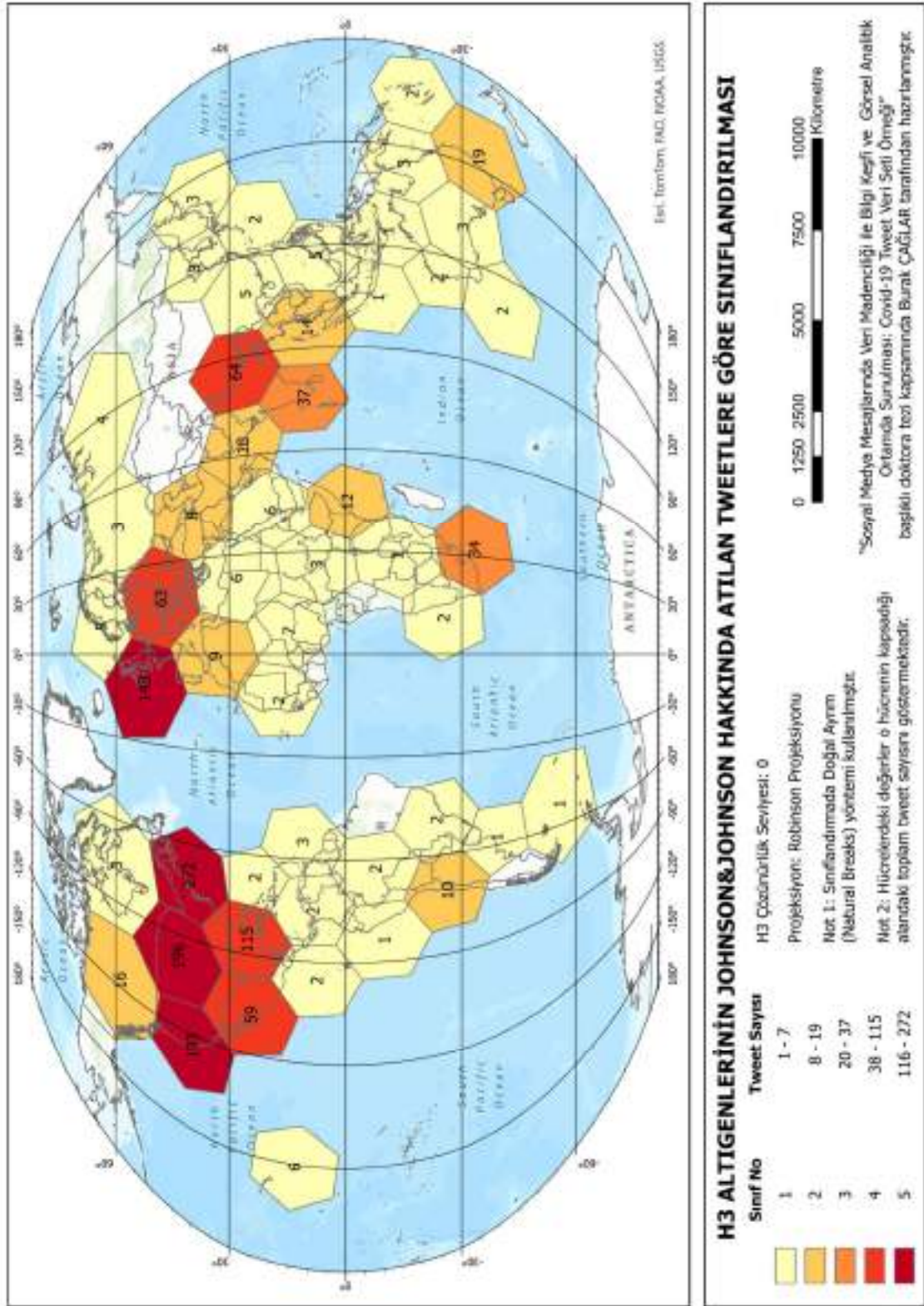
Ek-3.20 Sputnik V Aşısı Hakkındaki Tweetleri Kapsayan Çözünürlük=1 Seviyesindeki H3 Hiyerarşik Altıgen Hücreleri



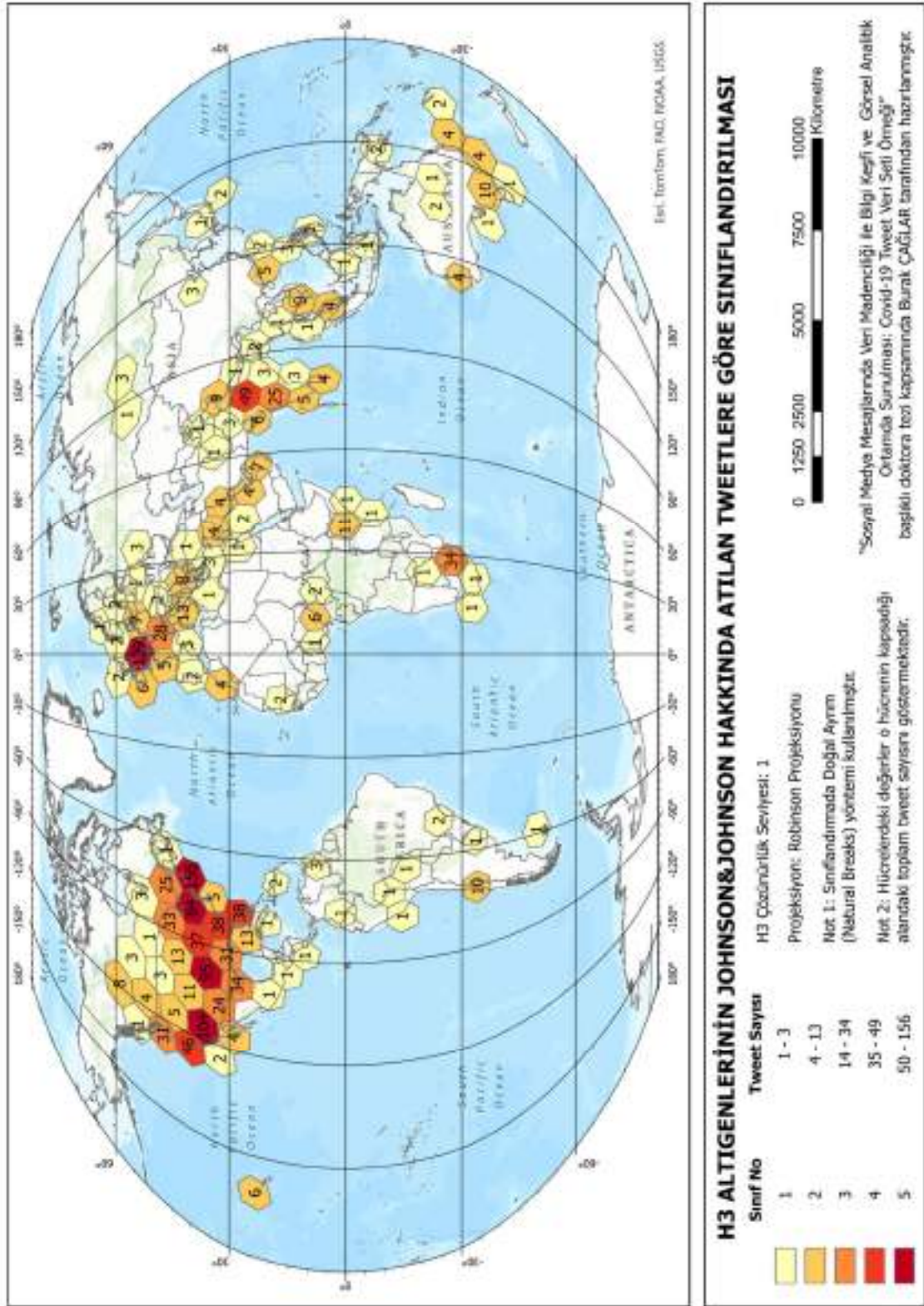
Ek-3.21 Sputnik V Aşısı Hakkındaki Tweetleri Kapsayan Çözünürlük=2 Seviyesindeki H3 Hiyerarşik Altıgen Hücreleri



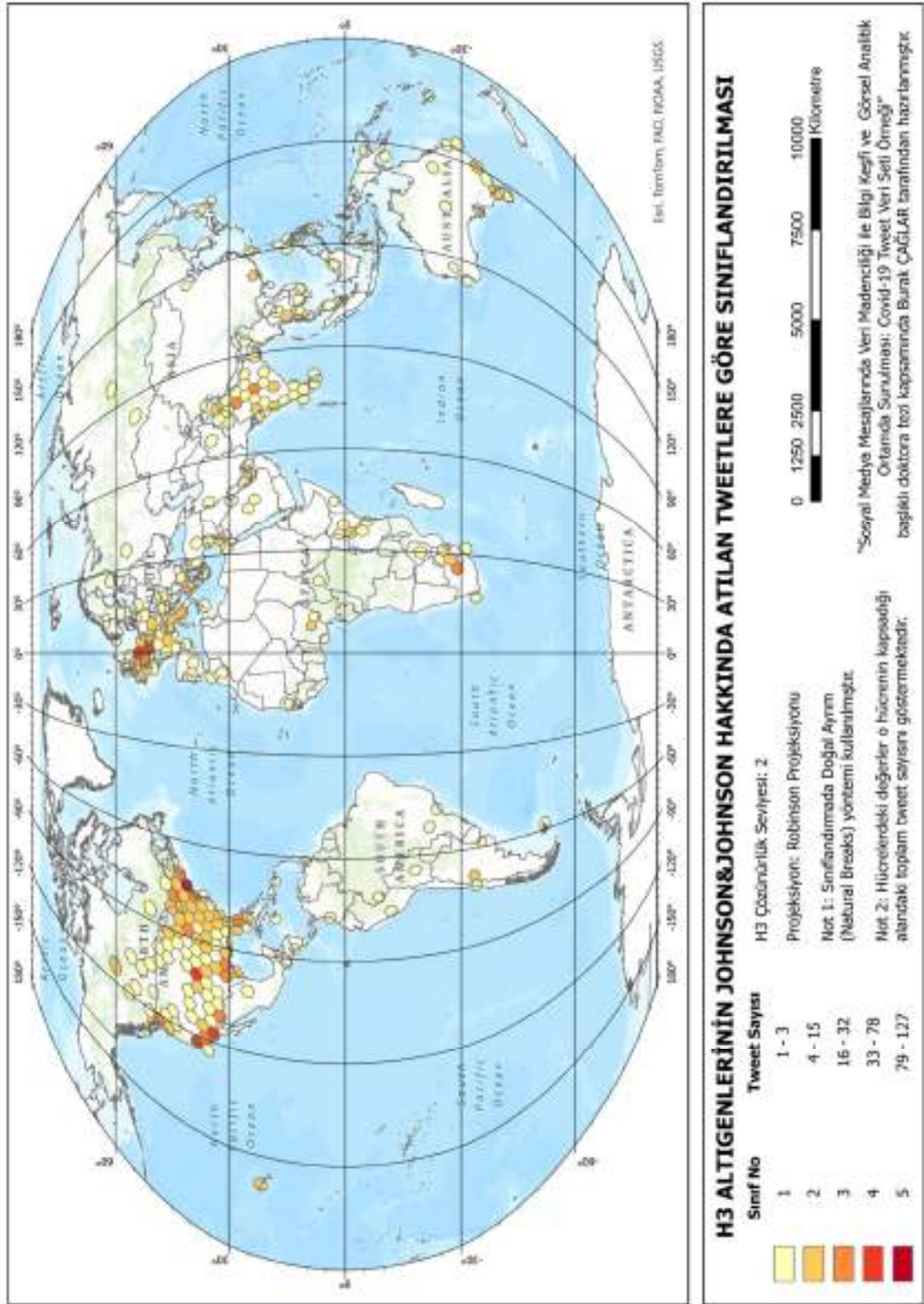
Ek-3.22 Johnson&Johnson Aşısı Hakkındaki Tweetleri Kapsayan Çözünürlük=0 Seviyesindeki H3 Hiyerarşik Altıgen Hücreleri



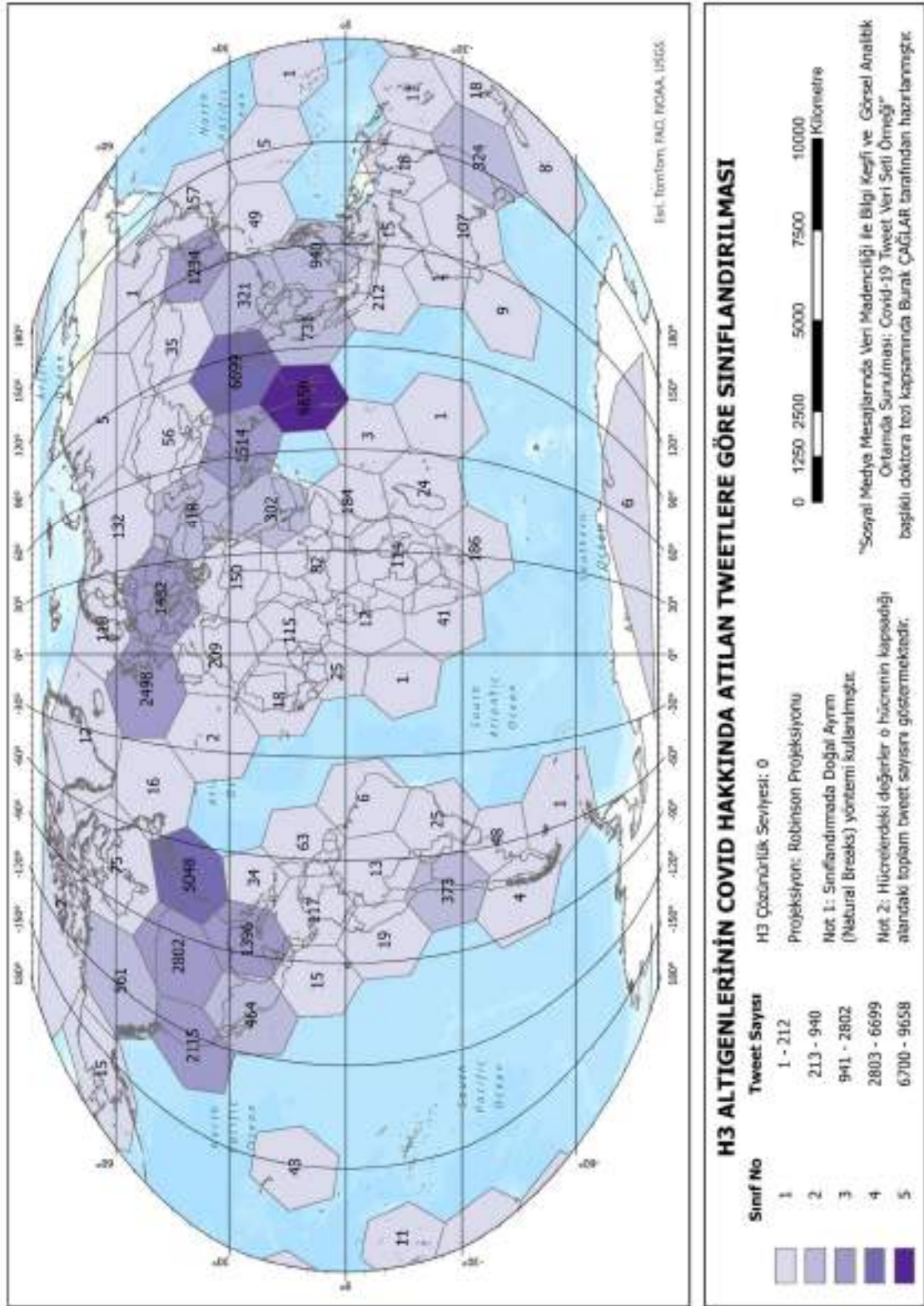
Ek-3.23 Johnson&Johnson Aşısı Hakkındaki Tweetleri Kapsayan Çözünürlük=1 Seviyesindeki H3 Hiyerarşik Altıgen Hücreleri



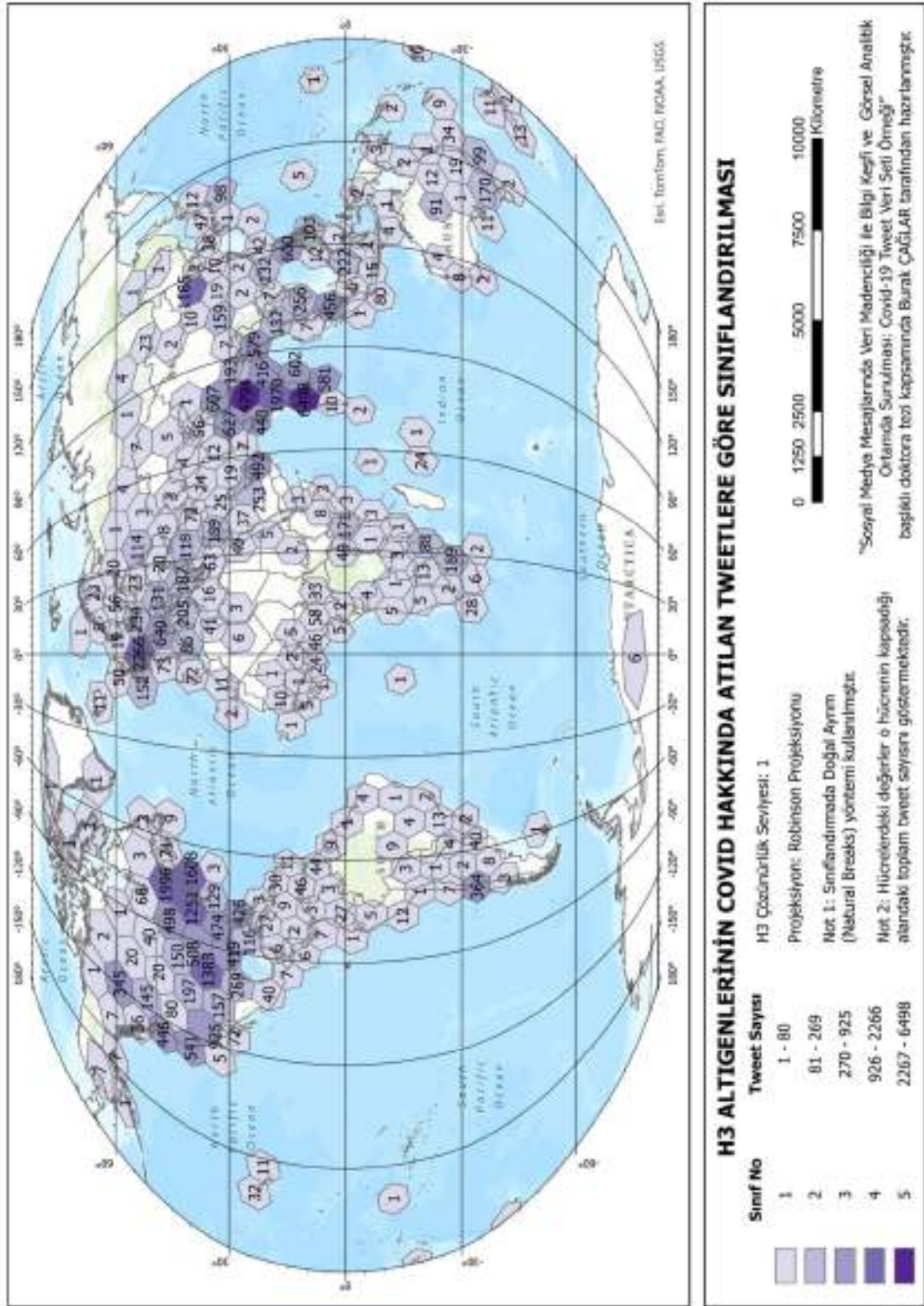
Ek-3.24 Johnson&Johnson Aşısı Hakkındaki Tweetleri Kapsayan Çözünürlük=2 Seviyesindeki H3 Hiyerarşik Altıgen Hücreleri



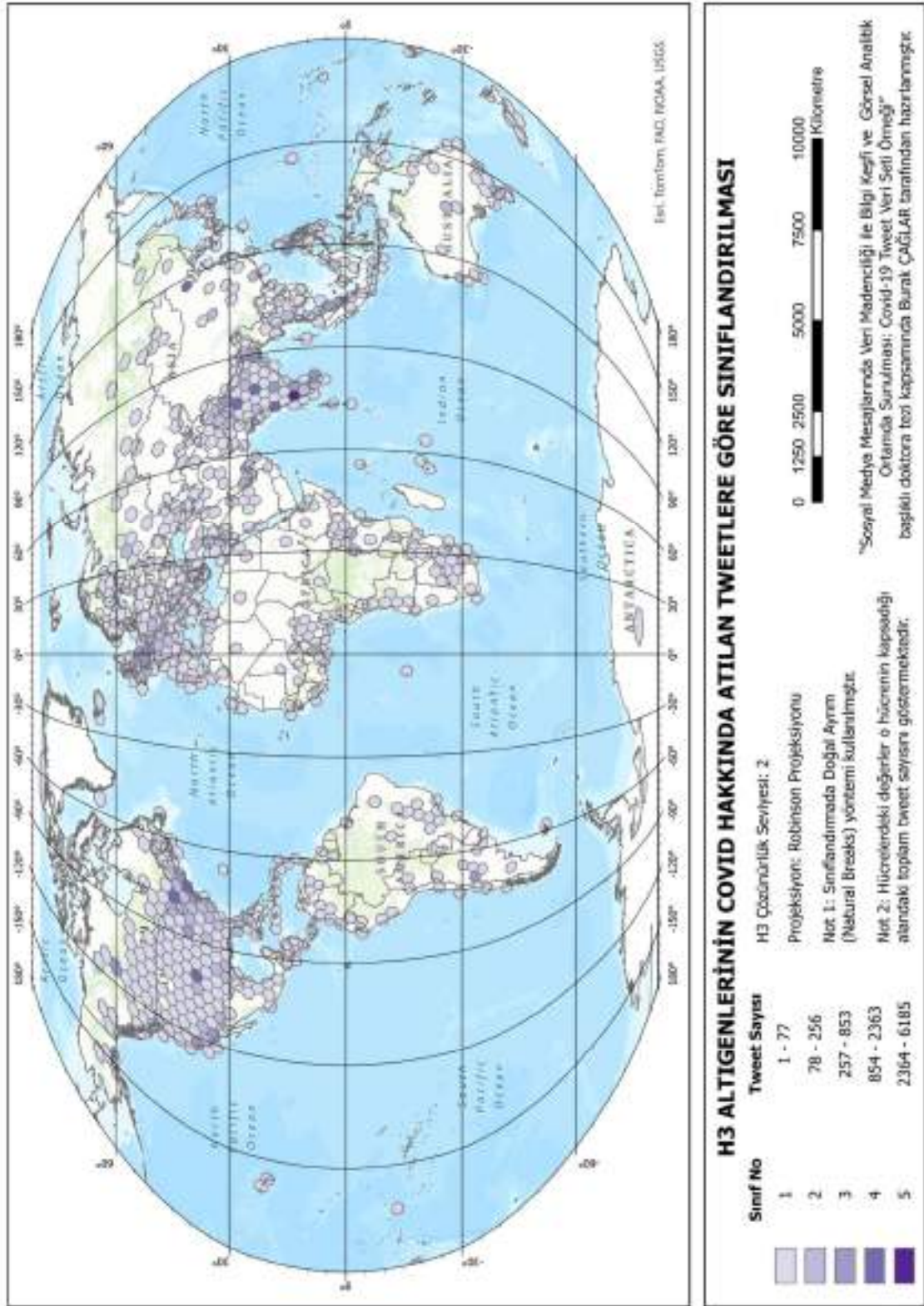
Ek-3.25 Covid Kelimesi Hakkındaki Tweetleri Kapsayan Çözünürlük=0 Seviyesindeki H3 Hiyerarşik Altıgen Hücreleri



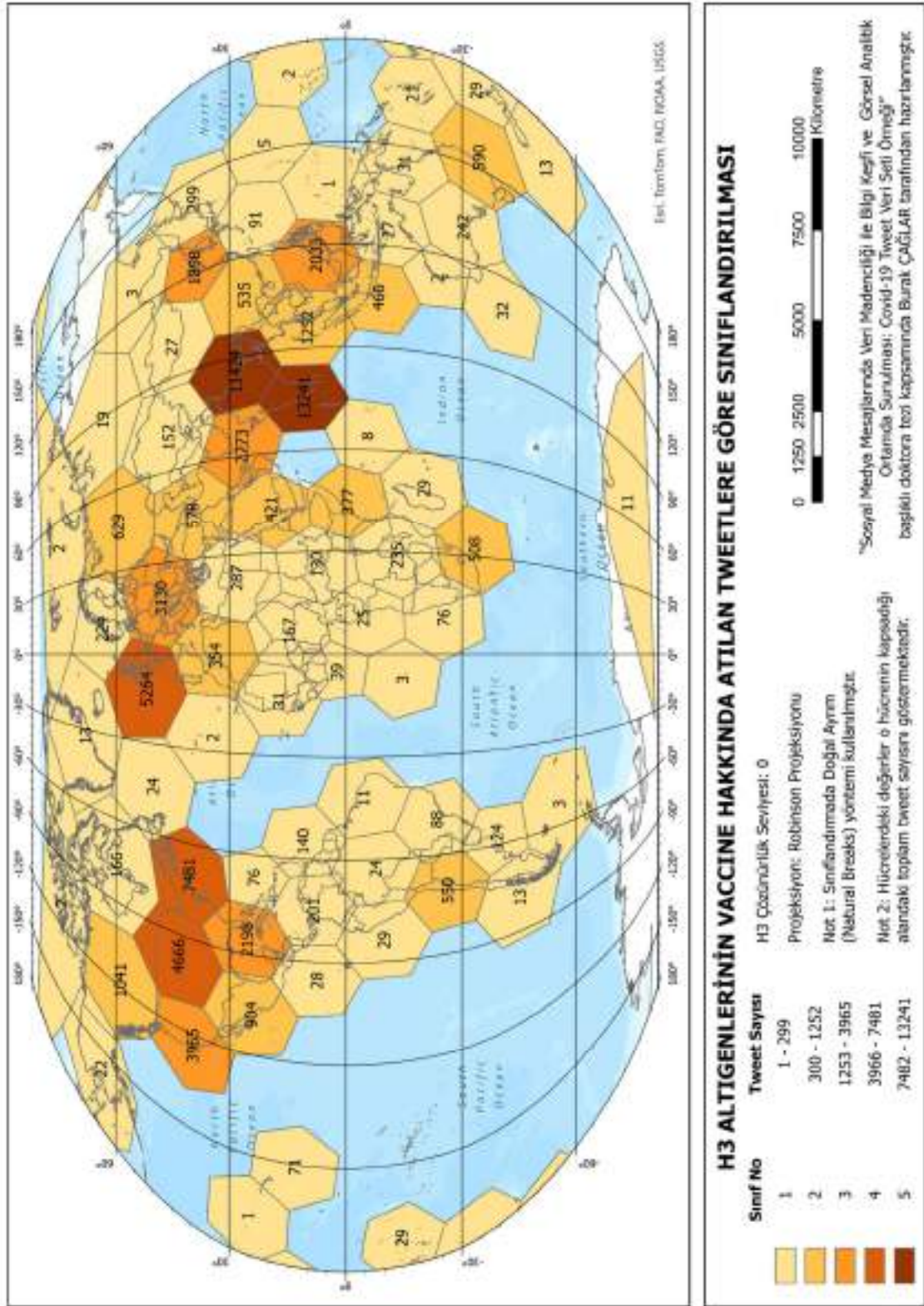
Ek-3.26 Covid Kelimesi Hakkındaki Tweetleri Kapsayan Çözünürlük=1 Seviyesindeki H3 Hiyerarşik Altıgen Hücreleri



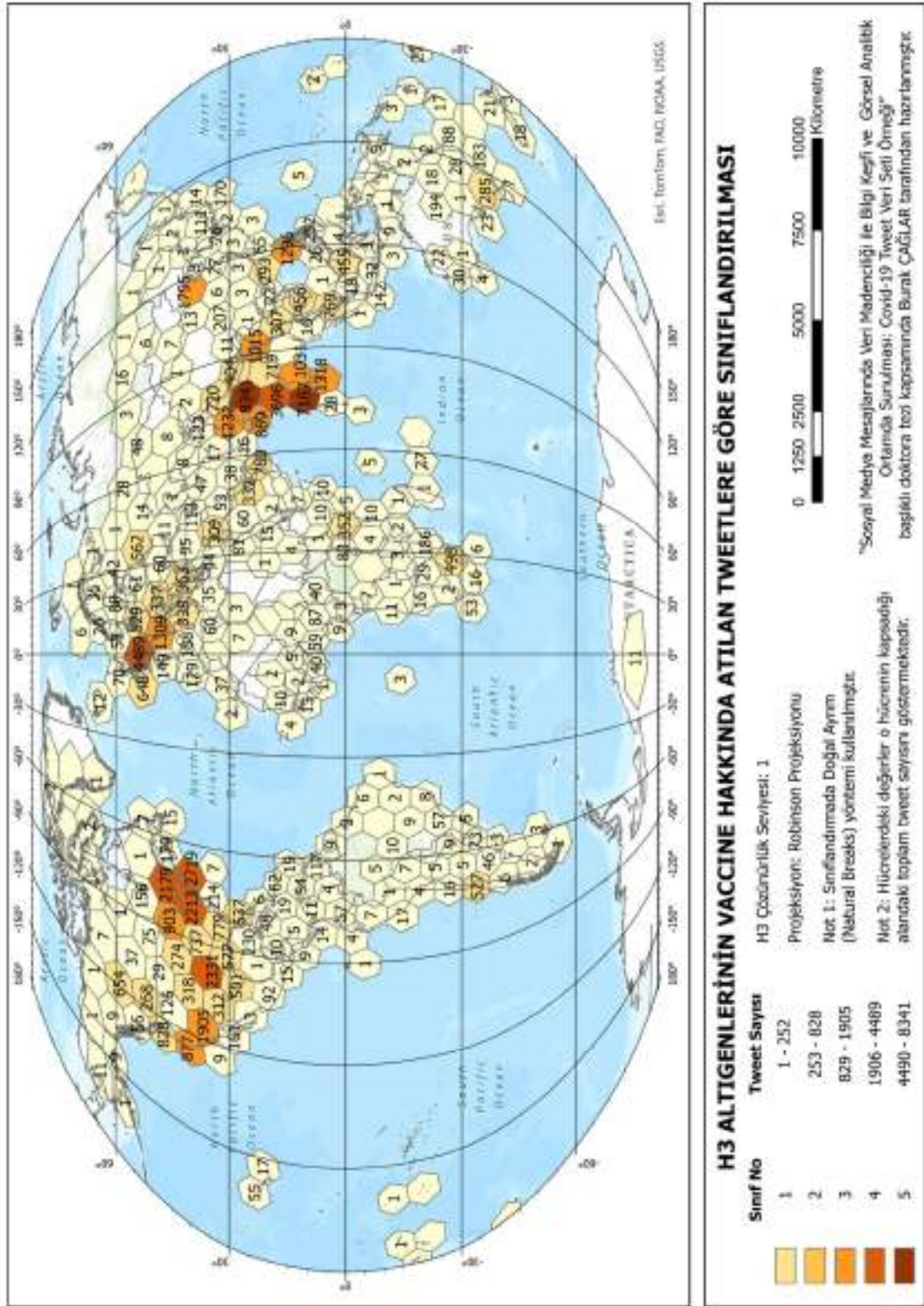
Ek-3.27 Covid Kelimesi Hakkındaki Tweetleri Kapsayan Çözünürlük=2 Seviyesindeki H3 Hiyerarşik Altıgen Hücreleri



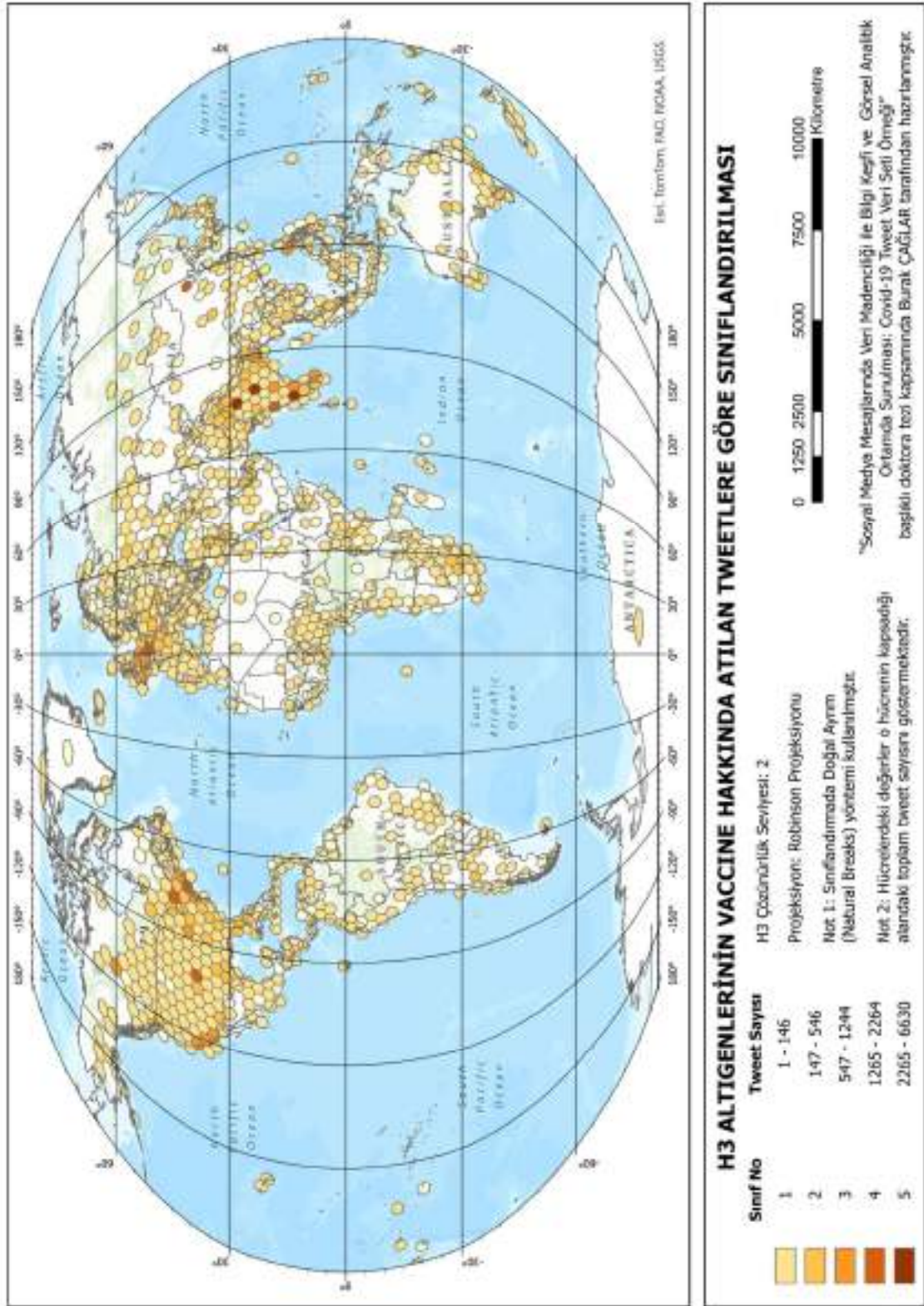
Ek-3.28 Vaccine Kelimesi Hakkındaki Tweetleri Kapsayan Çözünürlük=0 Seviyesindeki H3 Hiyerarşik Altıgen Hücreleri



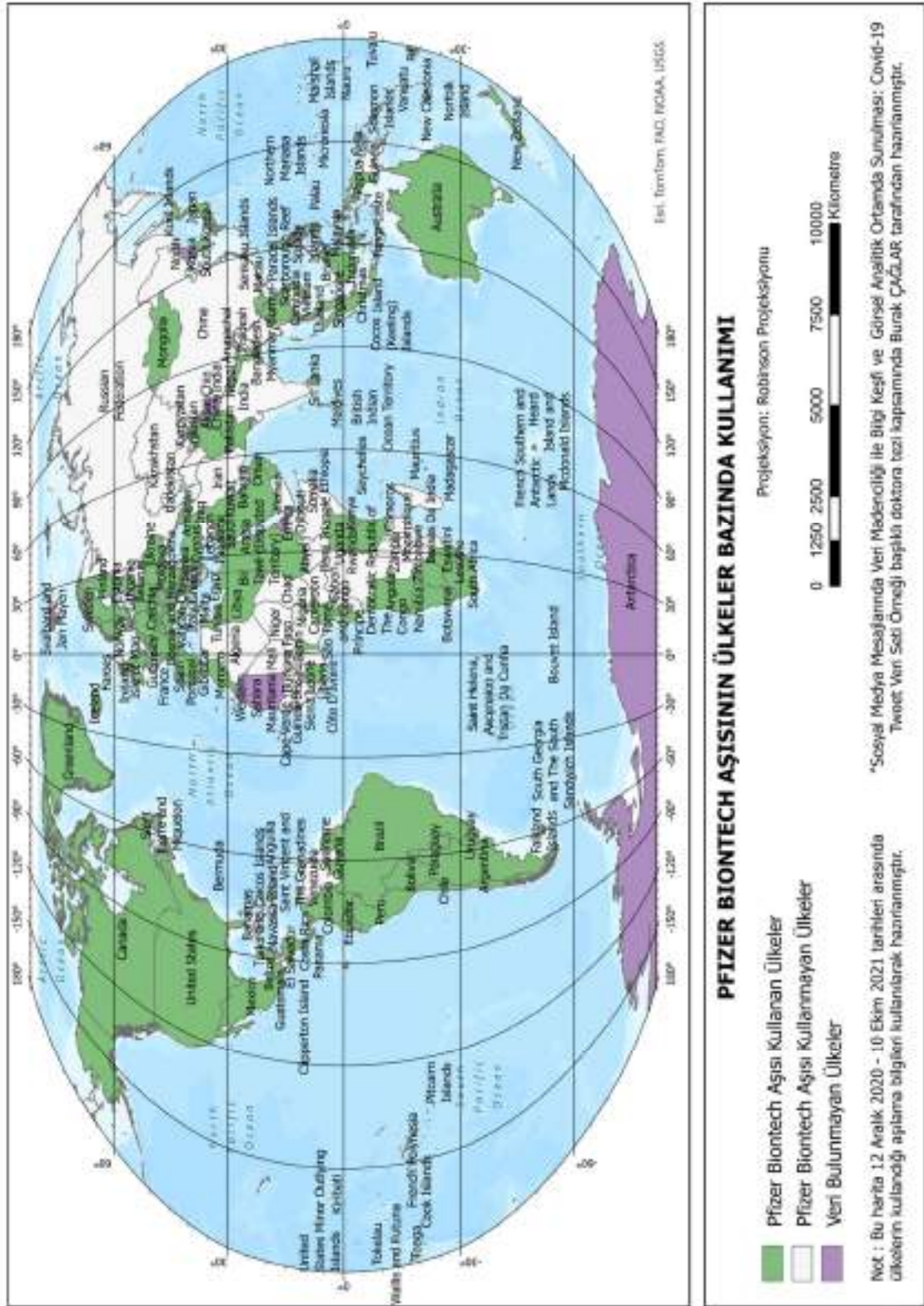
Ek-3.29 Vaccine Kelimesi Hakkındaki Tweetleri Kapsayan Çözünürlük=1 Seviyesindeki H3 Hiyerarşik Altıgen Hücreleri



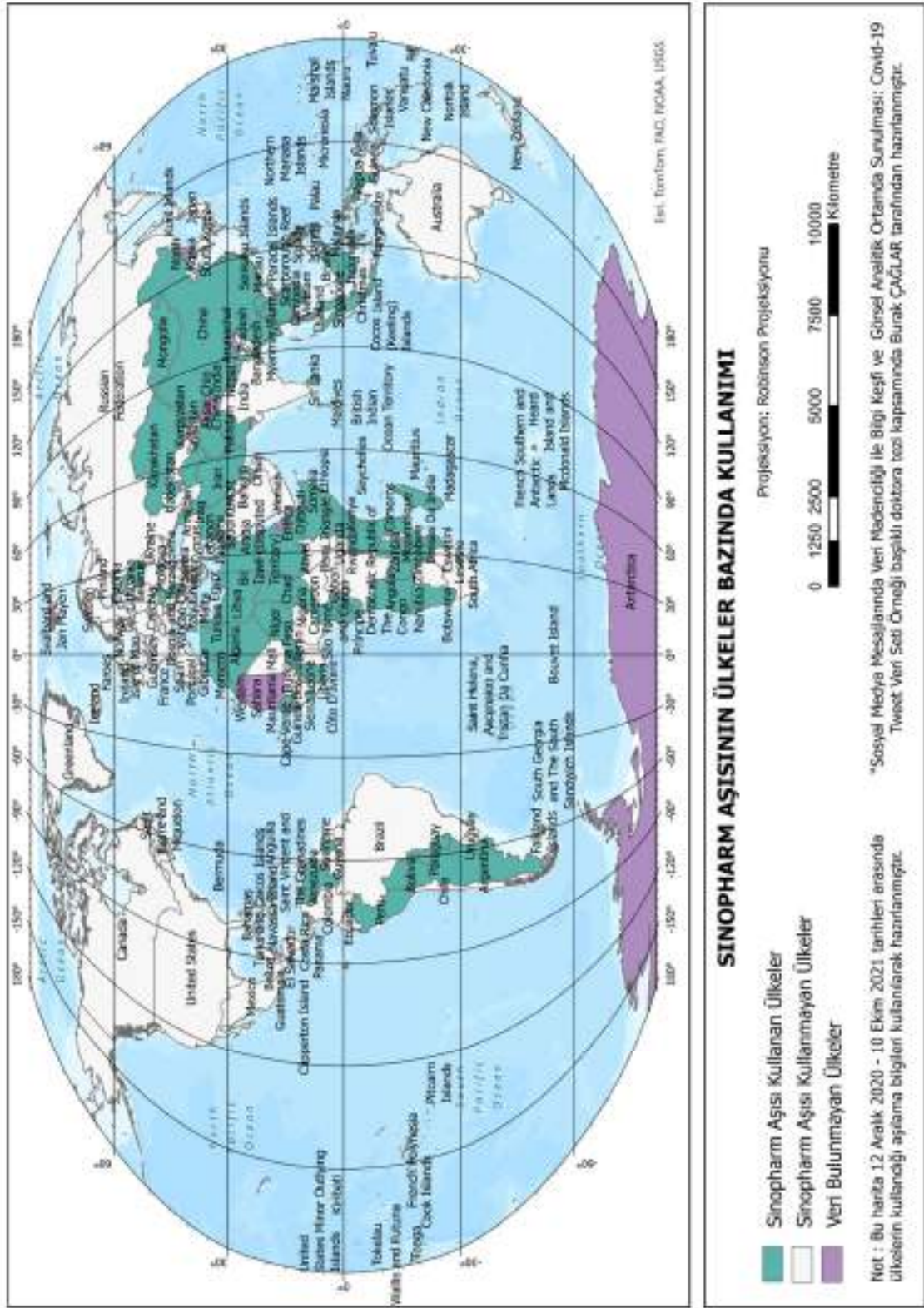
Ek-3.30 Vaccine Kelimesi Hakkındaki Tweetleri Kapsayan Çözünürlük=2 Seviyesindeki H3 Hiyerarşik Altıgen Hücreleri



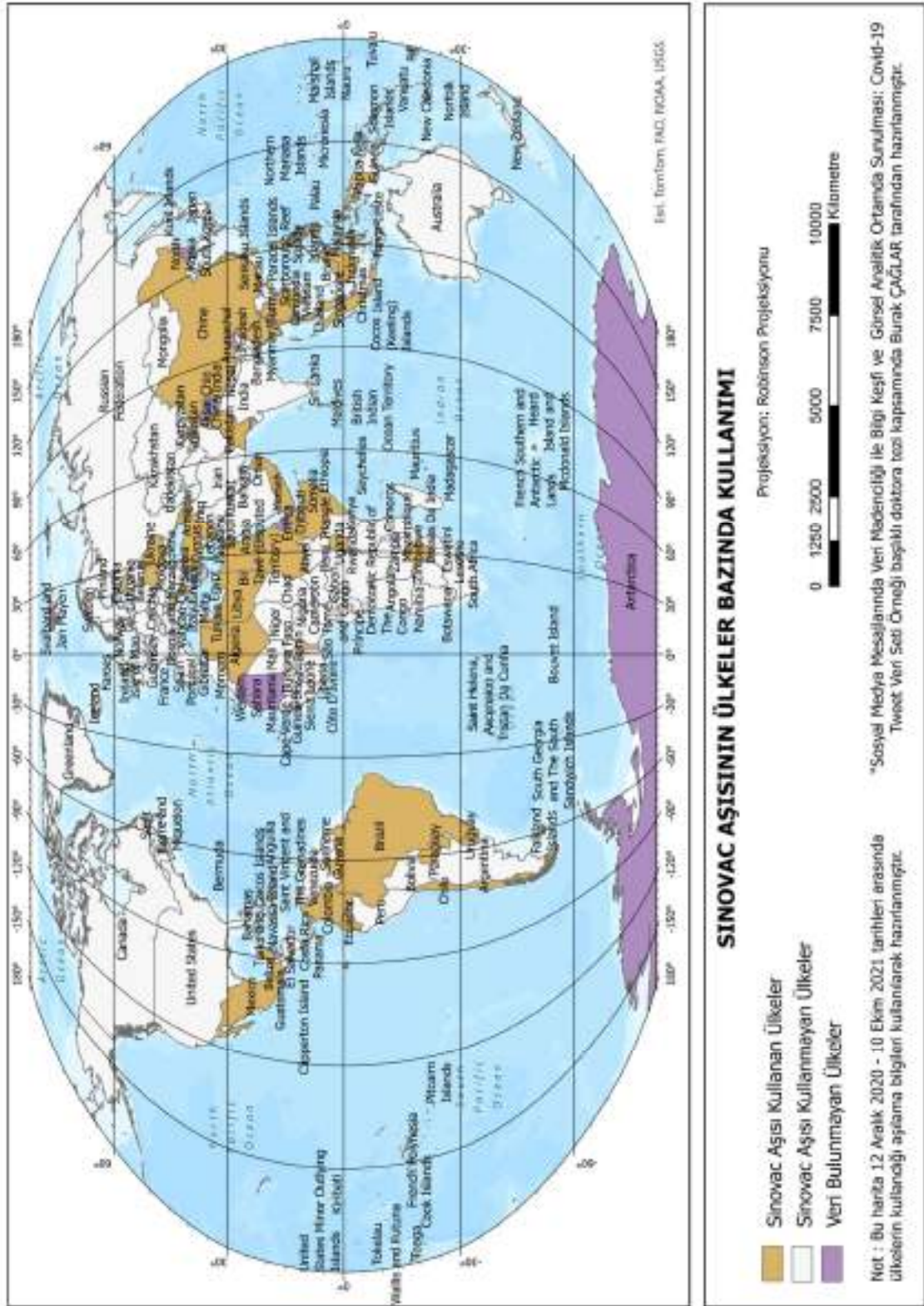
Ek-4.1 Pfizer Biontech Aşısının Ülkelere Göre Kullanımı



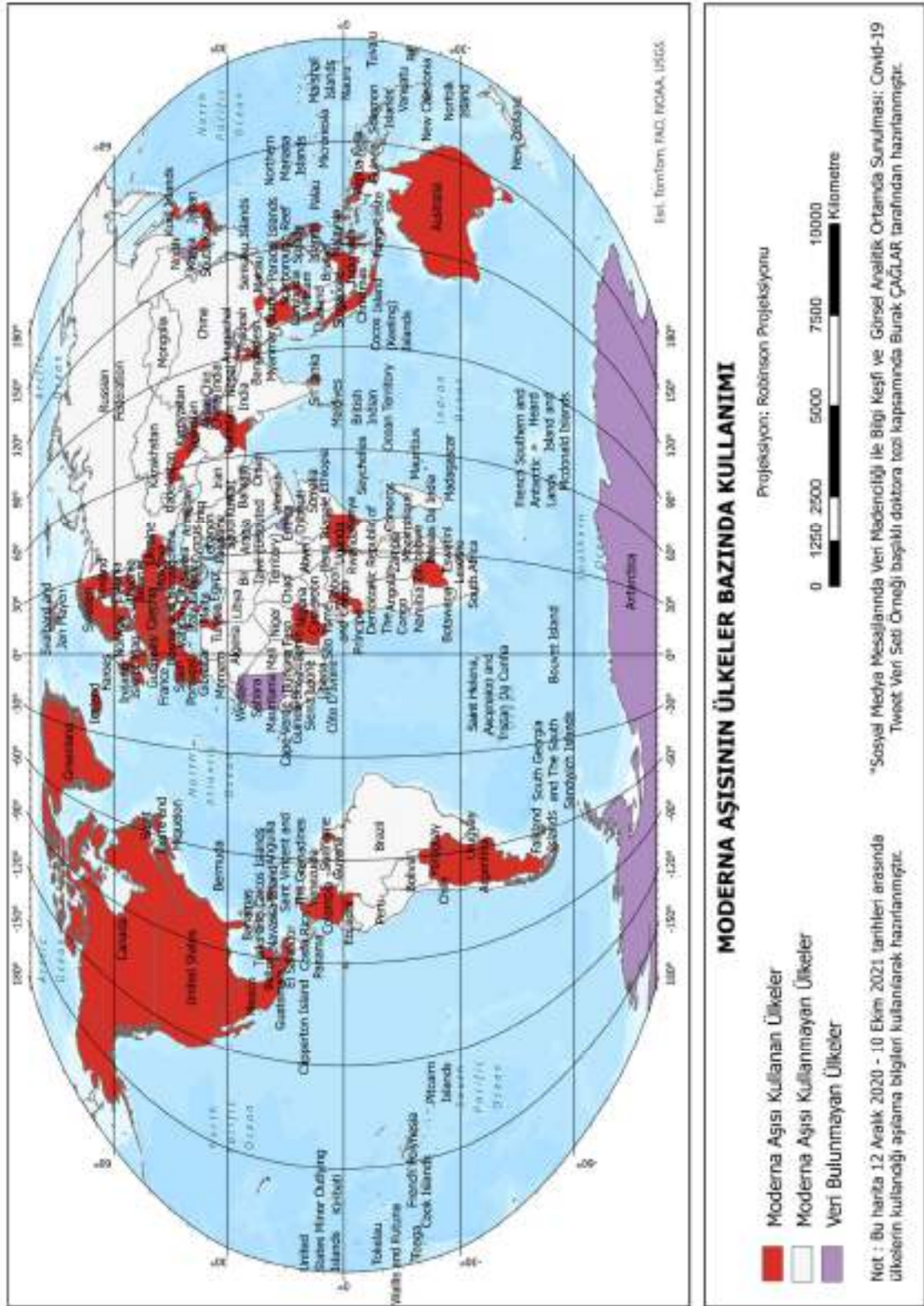
Ek-4.2 Sinopharm Aşısının Ülkelere Göre Kullanımı



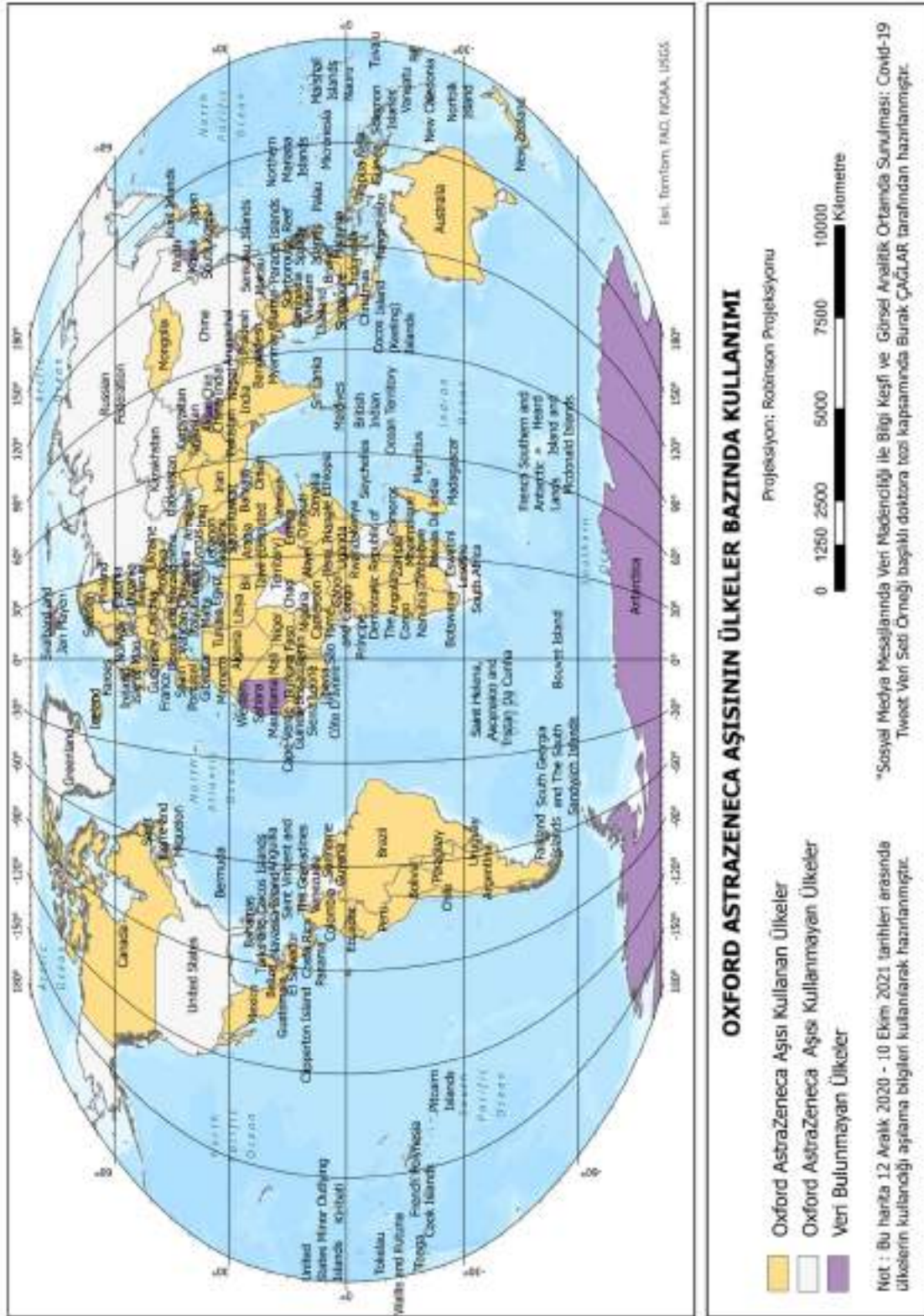
Ek-4.3 Sinovac Aşısının Ülkelere Göre Kullanımı



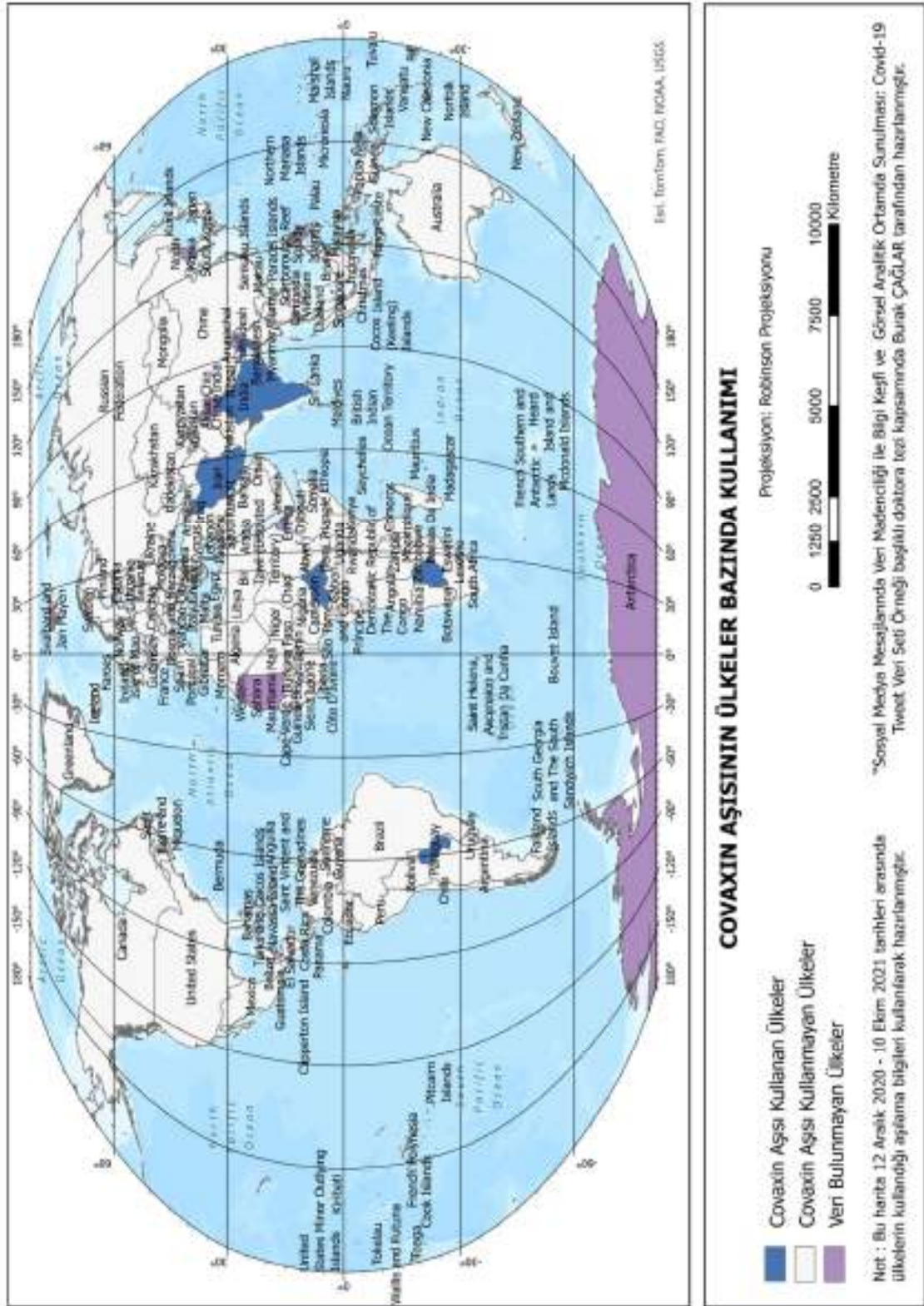
Ek-4.4 Moderna Aşısının Ülkelere Göre Kullanımı



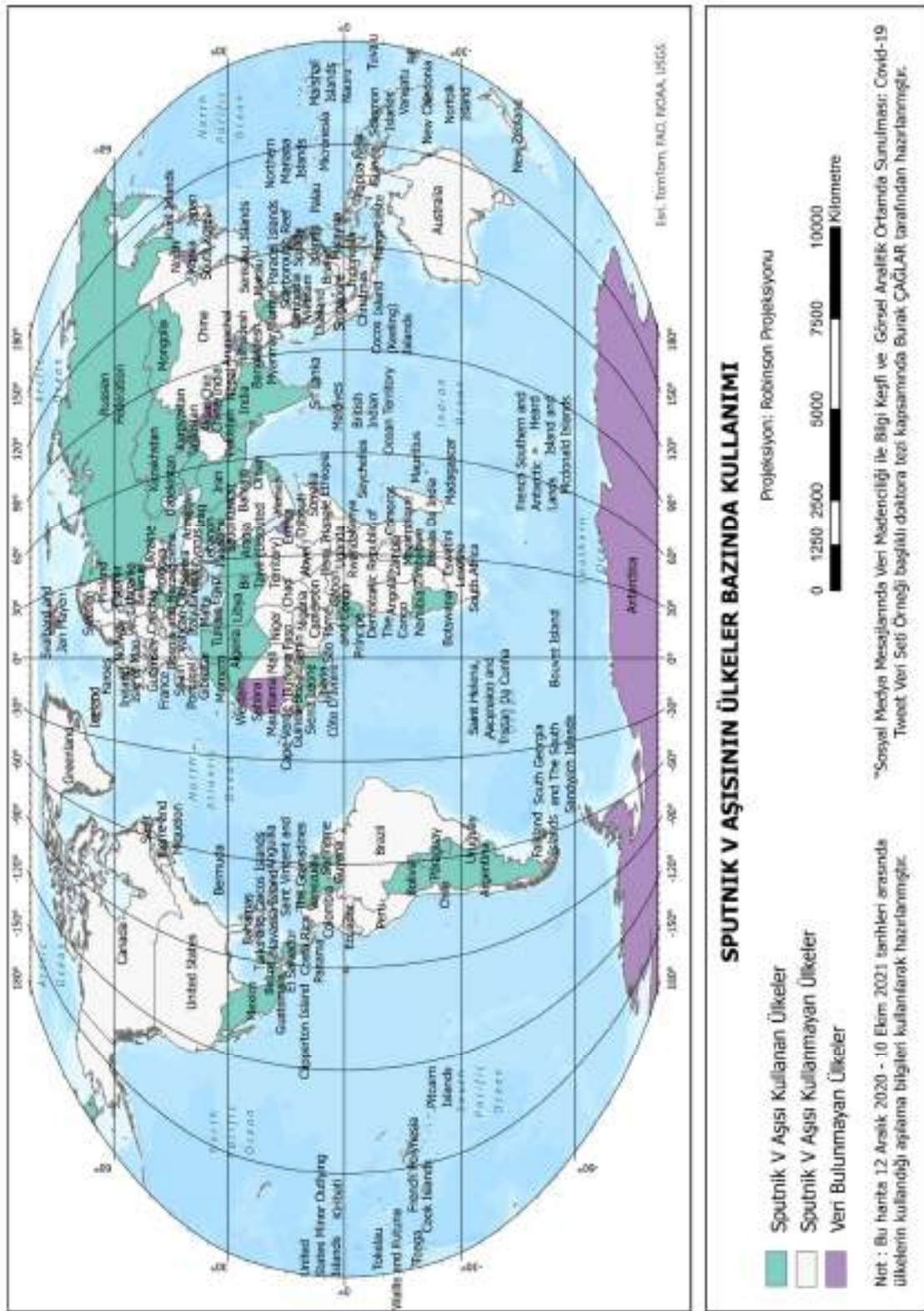
Ek-4.5 Oxford AstraZeneca Aşısının Ülkelere Göre Kullanımı



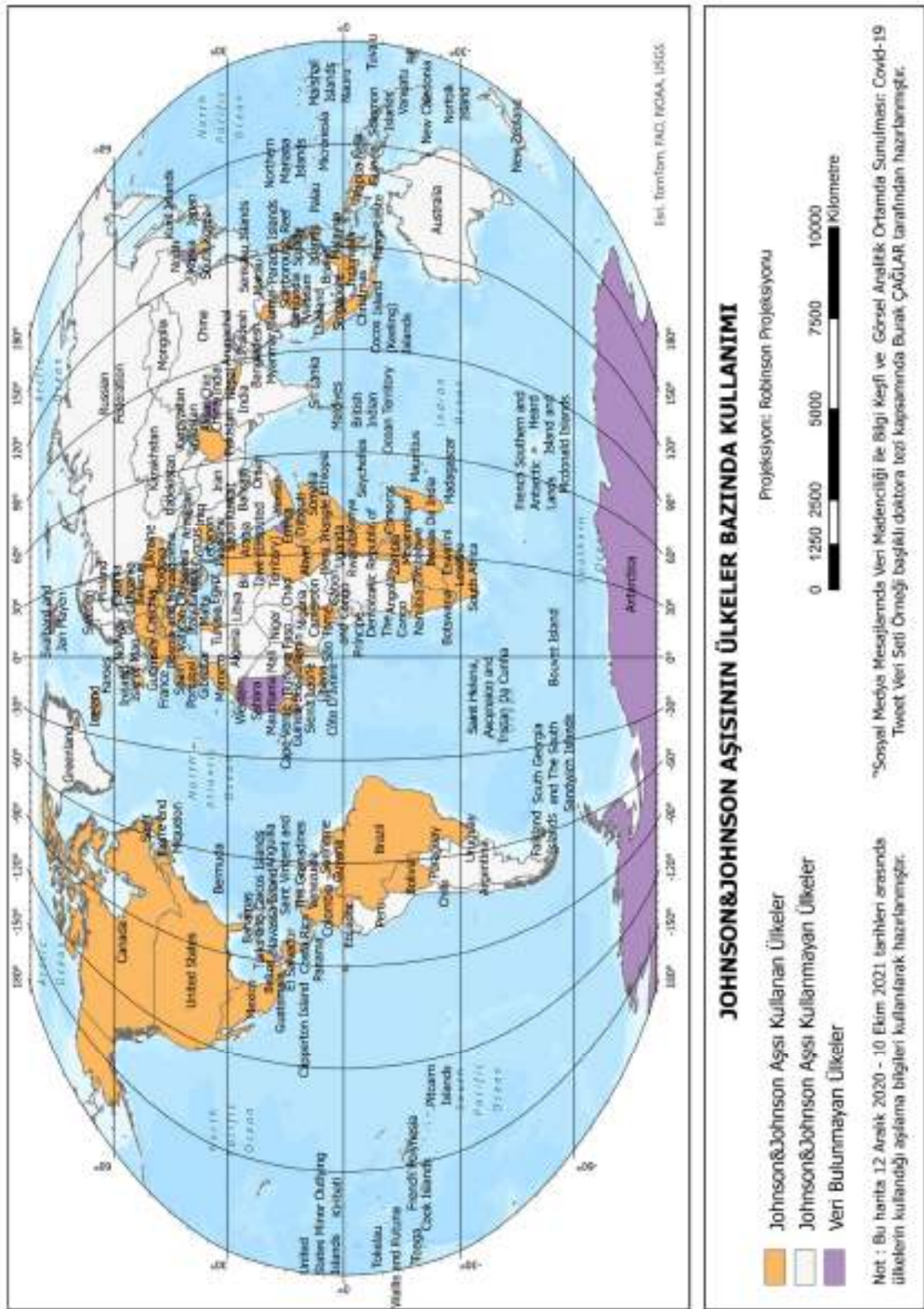
Ek-4.6 Covaxin Aşısının Ülkelere Göre Kullanımı



Ek-4.7 Sputnik V Aşısının Ülkelere Göre Kullanımı



Ek-4.8 Johnson&Johnson Aşısının Ülkelere Göre Kullanımı



Ek-5.1 RStudio Platformunda Ülkelerin X Kullanıcı Sayıları ve Tweet Sayılarına Göre Ward Yöntemiyle Hiyerarşik Kümeleme İşlemine ait Kod Bloğu

```
>library(readxl)
>Tweet_User <- read_excel("C:/***.xlsx")
>View(Tweet_User)
>summary(Tweet_User [, -1])
>tweetcount<-Tweet_User$`Tweet Count`
>usercount<-Tweet_User$`User Count`
>cor.test(tweetcount,usercount, method = c("pearson"))
>tu.data=Tweet_User[,2:3]
>wardsmethod<-hclust(dist(scale(tu.data)), method="ward.D")
>plot(wardsmethod)
>rect.hclust(wardsmethod,3)
> euclismtrx<-dist(tu.data, method = "euclidean",diag =
FALSE, upper = FALSE, p=2)
> cophoneticmtrx<-cophenetic(wardsmethod)
> cor(euclismtrx,cophoneticmtrx)
```

Ek-5.2 Ülkelerin X Kullanıcı Sayıları, Tweet Sayıları ve Kümeleme Sonuçları

Iso3 Code	Ülke Adı	Tweet Sayısı	Kullanıcı Sayısı	K-Means Küme No	Ward Küme No
IND	Hindistan	56080	16153	Cluster 1	Cluster 1
USA	Amerika	30922	18326	Cluster 1	Cluster 1
CAN	Kanada	9543	3835	Cluster 2	Cluster 2
GBR	Birleşik Krallık	8072	4661	Cluster 2	Cluster 2
PHL	Filipinler	2919	2115	Cluster 2	Cluster 2
PAK	Pakistan	2677	1627	Cluster 2	Cluster 2
CHN	Çin	2668	260	Cluster 3	Cluster 2
LKA	Sri Lanka	2333	460	Cluster 3	Cluster 2
AUS	Avustralya	1715	845	Cluster 3	Cluster 2
MYS	Malezya	1664	744	Cluster 3	Cluster 2
RUS	Rusya	1436	234	Cluster 3	Cluster 2
DEU	Almanya	1195	654	Cluster 3	Cluster 2
ARE	Birleşik Arap Emirlikleri	1190	405	Cluster 3	Cluster 2
ZAF	Güney Afrika	1079	422	Cluster 3	Cluster 2
FRA	Fransa	1021	426	Cluster 3	Cluster 2
THA	Tayland	937	296	Cluster 3	Cluster 2
IRL	İrlanda	901	445	Cluster 3	Cluster 2
ARG	Arjantin	878	212	Cluster 3	Cluster 2
ITA	İtalya	766	323	Cluster 3	Cluster 2
BEL	Belgium	679	255	Cluster 3	Cluster 2
BGD	Bangladeş	666	249	Cluster 3	Cluster 2
KEN	Kenya	636	342	Cluster 3	Cluster 2
XH	Cemmu ve Keşmir	597	238	Cluster 3	Cluster 2
NLD	Hollanda	590	368	Cluster 3	Cluster 2
JPN	Japonya	564	245	Cluster 3	Cluster 2
HKG	Hong Kong	554	148	Cluster 3	Cluster 3
TUR	Türkiye	517	127	Cluster 3	Cluster 3
SGP	Singapur	481	312	Cluster 3	Cluster 2
CHE	İsviçre	469	258	Cluster 3	Cluster 2
QAT	Katar	398	72	Cluster 3	Cluster 3
IDN	Endonezya	379	202	Cluster 3	Cluster 3
ESP	İspanya	370	206	Cluster 3	Cluster 3
ZWE	Zimbabve	361	155	Cluster 3	Cluster 3
BRA	Brezilya	284	115	Cluster 3	Cluster 3
NGA	Nijerya	260	148	Cluster 3	Cluster 3
DNK	Danimarka	233	75	Cluster 3	Cluster 3
MEX	Meksika	228	164	Cluster 3	Cluster 3
AUT	Avusturya	193	90	Cluster 3	Cluster 3
ISR	İsrael	190	57	Cluster 3	Cluster 3
LBN	Lübnan	184	77	Cluster 3	Cluster 3
TTO	Trinidad and Tobago	184	82	Cluster 3	Cluster 3

SAU	Suudi Arabistan	173	93	Cluster 3	Cluster 3
SWE	İsveç	159	89	Cluster 3	Cluster 3
FIN	Finlandiya	148	72	Cluster 3	Cluster 3
NOR	Norveç	148	91	Cluster 3	Cluster 3
KHM	Kamboçya	146	46	Cluster 3	Cluster 3
CZE	Çekya	144	52	Cluster 3	Cluster 3
HUN	Macaristan	136	50	Cluster 3	Cluster 3
SRB	Sırbistan	134	41	Cluster 3	Cluster 3
MMR	Myanmar/Burma	124	97	Cluster 3	Cluster 3
COL	Kolombiya	123	82	Cluster 3	Cluster 3
CHL	Şili	121	84	Cluster 3	Cluster 3
ATA	Antarktika	118	22	Cluster 3	Cluster 3
GRC	Yunanistan	118	73	Cluster 3	Cluster 3
NPL	Nepal	118	82	Cluster 3	Cluster 3
CYP	Kıbrıs	117	33	Cluster 3	Cluster 3
EGY	Mısır	113	53	Cluster 3	Cluster 3
KOR	Güney Kore	108	56	Cluster 3	Cluster 3
UKR	Ukrayna	108	30	Cluster 3	Cluster 3
SVK	Slovakya	107	30	Cluster 3	Cluster 3
VEN	Venezuela	106	30	Cluster 3	Cluster 3
UGA	Uganda	100	66	Cluster 3	Cluster 3
POL	Polonya	96	65	Cluster 3	Cluster 3
AZE	Azerbaycan	90	56	Cluster 3	Cluster 3
IRN	İran	90	56	Cluster 3	Cluster 3
PER	Peru	88	68	Cluster 3	Cluster 3
JAM	Jamaika	86	33	Cluster 3	Cluster 3
NZL	Yeni Zelanda	81	58	Cluster 3	Cluster 3
TUN	Tunus	81	42	Cluster 3	Cluster 3
VNM	Vietnam	80	43	Cluster 3	Cluster 3
KWT	Kuveyt	78	36	Cluster 3	Cluster 3
GHA	Gana	75	52	Cluster 3	Cluster 3
PRI	Porto Riko	72	27	Cluster 3	Cluster 3
PRT	Portekiz	71	45	Cluster 3	Cluster 3
ARM	Ermenistan	69	17	Cluster 3	Cluster 3
CAF	Orta Afrika Cumhuriyeti	64	27	Cluster 3	Cluster 3
MAR	Fas	64	34	Cluster 3	Cluster 3
BGR	Bulgaristan	57	13	Cluster 3	Cluster 3
MUS	Morityus	57	18	Cluster 3	Cluster 3
MLT	Malta	53	16	Cluster 3	Cluster 3
OMN	Umman	51	27	Cluster 3	Cluster 3
MDV	Maldivler	49	31	Cluster 3	Cluster 3
BHR	Bahreyn	48	25	Cluster 3	Cluster 3
NAM	Namibya	46	25	Cluster 3	Cluster 3
AGO	Angola	45	14	Cluster 3	Cluster 3
ALB	Arnavutluk	43	19	Cluster 3	Cluster 3

VIR	US Virgin Adaları	40	11	Cluster 3	Cluster 3
FJI	Fiji	39	23	Cluster 3	Cluster 3
JOR	Ürdün	37	26	Cluster 3	Cluster 3
ECU	Ekvator	36	20	Cluster 3	Cluster 3
SOM	Somali	34	27	Cluster 3	Cluster 3
AFG	Afganistan	30	22	Cluster 3	Cluster 3
MNE	Karadağ	30	7	Cluster 3	Cluster 3
MKD	Kuzey Makedonya	28	7	Cluster 3	Cluster 3
CUB	Küba	27	14	Cluster 3	Cluster 3
PAN	Panama	27	13	Cluster 3	Cluster 3
ISL	İzlanda	26	16	Cluster 3	Cluster 3
NCL	Yeni Kaledonya	26	14	Cluster 3	Cluster 3
GTM	Guatemala	25	18	Cluster 3	Cluster 3
SYR	Suriye	25	16	Cluster 3	Cluster 3
BOL	Bolivya	24	17	Cluster 3	Cluster 3
CMR	Kamerun	24	15	Cluster 3	Cluster 3
HRV	Hırvatistan	24	16	Cluster 3	Cluster 3
ETH	Etiyopya	24	15	Cluster 3	Cluster 3
GEO	Gürcistan	24	15	Cluster 3	Cluster 3
BLR	Belarus	23	14	Cluster 3	Cluster 3
DOM	Dominik Cumhuriyeti	22	12	Cluster 3	Cluster 3
IRQ	İraq	22	12	Cluster 3	Cluster 3
PRY	Paraguay	22	14	Cluster 3	Cluster 3
TZA	Tanzanya	22	18	Cluster 3	Cluster 3
BRN	Brunei	21	17	Cluster 3	Cluster 3
JEY	Jersey	21	17	Cluster 3	Cluster 3
LUX	Lüksemburg	21	16	Cluster 3	Cluster 3
MNG	Moğolistan	21	18	Cluster 3	Cluster 3
RWA	Ruanda	21	16	Cluster 3	Cluster 3
BWA	Botsvana	20	18	Cluster 3	Cluster 3
SVN	Slovenya	20	16	Cluster 3	Cluster 3
DZA	Cezayir	19	14	Cluster 3	Cluster 3
IMN	Man Adası	18	9	Cluster 3	Cluster 3
LTU	Litvanya	18	11	Cluster 3	Cluster 3
PSE	Filistin	18	12	Cluster 3	Cluster 3
PNG	Papua Yeni Gine	18	5	Cluster 3	Cluster 3
BRB	Barbados	17	14	Cluster 3	Cluster 3
GUY	Guyana	17	9	Cluster 3	Cluster 3
ROU	Romanya	17	13	Cluster 3	Cluster 3
SLE	Sierra Leone	17	7	Cluster 3	Cluster 3
BIH	Bosna Hersek	16	15	Cluster 3	Cluster 3
GMB	Gambiya	14	11	Cluster 3	Cluster 3
HND	Honduras	14	9	Cluster 3	Cluster 3
BTN	Bhutan	13	10	Cluster 3	Cluster 3
BFA	Burkina Faso	13	8	Cluster 3	Cluster 3

EST	Estonya	13	8	Cluster 3	Cluster 3
LBY	Libya	12	9	Cluster 3	Cluster 3
SLV	El Salvador	11	10	Cluster 3	Cluster 3
GRL	Grönland	11	3	Cluster 3	Cluster 3
UZB	Özbekistan	11	7	Cluster 3	Cluster 3
BMU	Bermuda	10	3	Cluster 3	Cluster 3
KAZ	Kazakistan	10	10	Cluster 3	Cluster 3
MLI	Mali	10	7	Cluster 3	Cluster 3
MOZ	Mozambik	10	7	Cluster 3	Cluster 3
SYC	Seyşeller	10	8	Cluster 3	Cluster 3
BHS	Bahamalar	9	9	Cluster 3	Cluster 3
COG	Kongo	9	5	Cluster 3	Cluster 3
GUM	Guam	9	7	Cluster 3	Cluster 3
XD	Arunaçhal Pradeş	8	5	Cluster 3	Cluster 3
VGB	British Virjin Adaları	8	3	Cluster 3	Cluster 3
HTI	Haiti	8	8	Cluster 3	Cluster 3
LAO	Laos	8	7	Cluster 3	Cluster 3
MAC	Makao	8	6	Cluster 3	Cluster 3
PRK	Kuzey Kore	8	8	Cluster 3	Cluster 3
KNA	Saint Kitts and Nevis	8	5	Cluster 3	Cluster 3
TLS	Doğu Timor	8	6	Cluster 3	Cluster 3
AND	Andorra	7	5	Cluster 3	Cluster 3
TJK	Tacikistan	7	5	Cluster 3	Cluster 3
URY	Uruguay	7	7	Cluster 3	Cluster 3
CIV	Kotdivuar	6	6	Cluster 3	Cluster 3
MWI	Malavi	6	6	Cluster 3	Cluster 3
NIC	Nikaragua	6	5	Cluster 3	Cluster 3
SHN	Saint Helena	6	1	Cluster 3	Cluster 3
SEN	Senegal	6	3	Cluster 3	Cluster 3
SDN	Sudan	6	5	Cluster 3	Cluster 3
CRI	Kosta Rika	5	5	Cluster 3	Cluster 3
COD	Kongo Demokratik Cumhuriyeti	5	4	Cluster 3	Cluster 3
FLK	Falkland Adaları	5	4	Cluster 3	Cluster 3
GAB	Gabon	5	5	Cluster 3	Cluster 3
GIB	Cebelitarık	5	4	Cluster 3	Cluster 3
MCO	Monako	5	5	Cluster 3	Cluster 3
SLB	Solomon Adaları	5	5	Cluster 3	Cluster 3
GIN	Gine	4	4	Cluster 3	Cluster 3
KGZ	Kırgızistan	4	3	Cluster 3	Cluster 3
LVA	Letonya	4	4	Cluster 3	Cluster 3
LBR	Liberya	4	3	Cluster 3	Cluster 3
LIE	Lihtenştayn	4	2	Cluster 3	Cluster 3
MDA	Moldova	4	4	Cluster 3	Cluster 3
TGO	Togo	4	3	Cluster 3	Cluster 3
ZMB	Zambiya	4	4	Cluster 3	Cluster 3

ATG	Antigua ve Barbuda	3	3	Cluster 3	Cluster 3
BEN	Benin	3	3	Cluster 3	Cluster 3
BDI	Burundi	3	3	Cluster 3	Cluster 3
CYM	Cayman Adaları	3	2	Cluster 3	Cluster 3
TCD	Çad	3	2	Cluster 3	Cluster 3
SWZ	Esvatini	3	3	Cluster 3	Cluster 3
LSO	Lesotho	3	3	Cluster 3	Cluster 3
MHL	Marşal Adaları	3	3	Cluster 3	Cluster 3
NER	Nijer	3	3	Cluster 3	Cluster 3
VUT	Vanuatu	3	2	Cluster 3	Cluster 3
YEM	Yemen	3	3	Cluster 3	Cluster 3
COK	Cook Adaları	2	2	Cluster 3	Cluster 3
GGY	Guernsey	2	2	Cluster 3	Cluster 3
MDG	Madagaskar	2	2	Cluster 3	Cluster 3
SSD	Güney Sudan	2	2	Cluster 3	Cluster 3
SUR	Surinam	2	2	Cluster 3	Cluster 3
TON	Tonga	2	2	Cluster 3	Cluster 3
COM	Komorlar	1	1	Cluster 3	Cluster 3
DJI	Cibuti	1	1	Cluster 3	Cluster 3
DMA	Dominika	1	1	Cluster 3	Cluster 3
GNQ	Ekvator Ginesi	1	1	Cluster 3	Cluster 3
FRO	Faroe Adaları	1	1	Cluster 3	Cluster 3
LCA	Saint Lucia	1	1	Cluster 3	Cluster 3
VCT	Saint Vincent ve Grenadinler	1	1	Cluster 3	Cluster 3
WSM	Samoa	1	1	Cluster 3	Cluster 3
SMR	San Marino	1	1	Cluster 3	Cluster 3
STP	Sao Tome ve Prinsipe	1	1	Cluster 3	Cluster 3
XN	Senkaku Adaları	1	1	Cluster 3	Cluster 3
SXM	Sint-Maarten	1	1	Cluster 3	Cluster 3
TKM	Türkmenistan	1	1	Cluster 3	Cluster 3
VAT	Vatikan	1	1	Cluster 3	Cluster 3

Ek-5.3 RStudio Platformunda Ülkelerin Oluşturulan Toplam Tweet Sayısı, Covid Kelimesi İçeren Tweet Sayısı ve Vaccine Kelimesi İçeren Tweet Sayıları Göre Ward Yöntemiyle Kümelenmesi İşlemine ait Kod Bloğu

```
>library(readxl)
>Tweet_Covid_Vaccine<-
>read_excel("C:/Users/GIS01/Desktop/.../Tweet_Covid_Vaccine.xlsx")
>View(Tweet_Covid_Vaccine)
>summary(Tweet_Covid_Vaccine[, -1])
>tweetcount<-Tweet_Covid_Vaccine$`Tweet Count`
>covidcount<-Tweet_Covid_Vaccine$`Covid Tweet Count`
>vaccinecount<-Tweet_Covid_Vaccine$`Vaccine Tweet Count`
>cor.test(tweetcount,covidcount, method = c("pearson"))
>cor.test(covidcount,vaccinecount, method = c("pearson"))
>cor.test(tweetcount,vaccinecount, method = c("pearson"))
>tcv.data=Tweet_Covid_Vaccine[,3:5]
>wardsmethod<-hclust(dist(scale(tcv.data)),method="ward.D")
>plot(wardsmethod)
>rect.hclust(wardsmethod,3)
>euclidmtrx<-dist(tcv.data,method = "euclidean",diag=FALSE,
upper = FALSE, p=2)
>cophoneticmtrx<-cophenetic(wardsmethod)
>cor(euclidmtrx,cophoneticmtrx)
>cluster<-cutree(wardsmethod,3)
>table=data.frame(Tweet_Covid_Vaccine, cluster)
>View(table)
>write.csv(table,"cluster_result.csv")
```

Ek-5.4 Ülkelerin Tweet Sayıları, Covid Kelimesi Hakkında Tweet Sayıları, Vaccine Kelimesi Hakkında Tweet Sayıları ve Kümeleme Sonuçları

ISO3 Code	Ülke Adı	Tweet Sayısı	Covid Hakkında Tweet Sayısı	Vaccine Hakkında Tweet Sayısı	K-Means Küme No	Ward Küme No
IND	Hindistan	56080	15443	22913	Cluster 1	Cluster 1
USA	Amerika	30922	9563	16383	Cluster 1	Cluster 1
CAN	Kanada	9543	2911	4013	Cluster 2	Cluster 2
GBR	Birleşik Krallık	8072	2224	4358	Cluster 2	Cluster 2
PHL	Filipinler	2919	706	1553	Cluster 2	Cluster 2
PAK	Pakistan	2677	864	1668	Cluster 2	Cluster 2
CHN	Çin	2668	1457	2171	Cluster 2	Cluster 2
LKA	Sri Lanka	2333	585	1327	Cluster 3	Cluster 2
AUS	Avustralya	1715	457	889	Cluster 3	Cluster 2
MYS	Malezya	1664	554	927	Cluster 3	Cluster 2
RUS	Rusya	1436	172	713	Cluster 3	Cluster 2
DEU	Almanya	1195	317	636	Cluster 3	Cluster 2
ARE	Birleşik Arap Emirlikleri	1190	477	753	Cluster 3	Cluster 2
ZAF	Güney Afrika	1079	224	560	Cluster 3	Cluster 2
FRA	Fransa	1021	234	502	Cluster 3	Cluster 2
THA	Tayland	937	304	522	Cluster 3	Cluster 2
IRL	İrlanda	901	142	628	Cluster 3	Cluster 2
ARG	Arjantin	878	390	613	Cluster 3	Cluster 2
ITA	İtalya	766	239	387	Cluster 3	Cluster 2
BEL	Belgium	679	140	383	Cluster 3	Cluster 2
BGD	Bangladeş	666	209	347	Cluster 3	Cluster 2
KEN	Kenya	636	173	353	Cluster 3	Cluster 2
XH	Cemmu ve Keşmir	597	216	277	Cluster 3	Cluster 2
NLD	Hollanda	590	122	270	Cluster 3	Cluster 2
JPN	Japonya	564	161	309	Cluster 3	Cluster 2
HKG	Hong Kong	554	122	267	Cluster 3	Cluster 2
TUR	Türkiye	517	191	193	Cluster 3	Cluster 2
SGP	Singapur	481	117	281	Cluster 3	Cluster 2
CHE	İsviçre	469	144	245	Cluster 3	Cluster 2
QAT	Katar	398	222	273	Cluster 3	Cluster 2
IDN	Endonezya	379	120	234	Cluster 3	Cluster 2
ESP	İspanya	370	116	193	Cluster 3	Cluster 2
ZWE	Zimbabve	361	100	213	Cluster 3	Cluster 2
BRA	Brezilya	284	36	105	Cluster 3	Cluster 3
NGA	Nijerya	260	102	143	Cluster 3	Cluster 3
DNK	Danimarka	233	53	137	Cluster 3	Cluster 3
MEX	Meksika	228	47	119	Cluster 3	Cluster 3
AUT	Avusturya	193	47	96	Cluster 3	Cluster 3

ISR	Israel	190	79	122	Cluster 3	Cluster 3
LBN	Lübnan	184	78	128	Cluster 3	Cluster 3
TTO	Trinidad and Tobago	184	44	115	Cluster 3	Cluster 3
SAU	Suudi Arabistan	173	61	105	Cluster 3	Cluster 3
SWE	İsveç	159	46	76	Cluster 3	Cluster 3
FIN	Finlandiya	148	47	69	Cluster 3	Cluster 3
NOR	Norveç	148	28	79	Cluster 3	Cluster 3
KHM	Kamboçya	146	62	113	Cluster 3	Cluster 3
CZE	Çekya	144	26	67	Cluster 3	Cluster 3
HUN	Macaristan	136	21	84	Cluster 3	Cluster 3
SRB	Sırbistan	134	34	93	Cluster 3	Cluster 3
MMR	Myanmar/Burma	124	9	112	Cluster 3	Cluster 3
COL	Kolombiya	123	27	52	Cluster 3	Cluster 3
CHL	Şili	121	33	68	Cluster 3	Cluster 3
ATA	Antarktika	118	6	11	Cluster 3	Cluster 3
GRC	Yunanistan	118	28	71	Cluster 3	Cluster 3
NPL	Nepal	118	30	80	Cluster 3	Cluster 3
CYP	Kıbrıs	117	57	77	Cluster 3	Cluster 3
EGY	Mısır	113	40	83	Cluster 3	Cluster 3
KOR	Güney Kore	108	37	70	Cluster 3	Cluster 3
UKR	Ukrayna	108	19	57	Cluster 3	Cluster 3
SVK	Slovakya	107	18	54	Cluster 3	Cluster 3
VEN	Venezuela	106	49	59	Cluster 3	Cluster 3
UGA	Uganda	100	36	68	Cluster 3	Cluster 3
POL	Polonya	96	18	51	Cluster 3	Cluster 3
AZE	Azerbaycan	90	47	41	Cluster 3	Cluster 3
IRN	İran	90	36	68	Cluster 3	Cluster 3
PER	Peru	88	17	26	Cluster 3	Cluster 3
JAM	Jamaika	86	28	53	Cluster 3	Cluster 3
NZL	Yeni Zelanda	81	26	42	Cluster 3	Cluster 3
TUN	Tunus	81	39	52	Cluster 3	Cluster 3
VNM	Vietnam	80	23	51	Cluster 3	Cluster 3
KWT	Kuveyt	78	19	41	Cluster 3	Cluster 3
GHA	Gana	75	21	37	Cluster 3	Cluster 3
PRI	Porto Riko	72	22	33	Cluster 3	Cluster 3
PRT	Portekiz	71	22	40	Cluster 3	Cluster 3
ARM	Ermenistan	69	21	62	Cluster 3	Cluster 3
CAF	Orta Afrika Cumhuriyeti	64	33	39	Cluster 3	Cluster 3
MAR	Fas	64	11	37	Cluster 3	Cluster 3
BGR	Bulgaristan	57	34	44	Cluster 3	Cluster 3
MUS	Morityus	57	24	26	Cluster 3	Cluster 3
MLT	Malta	53	14	30	Cluster 3	Cluster 3
OMN	Umman	51	16	29	Cluster 3	Cluster 3
MDV	Maldivler	49	12	31	Cluster 3	Cluster 3
BHR	Bahreyn	48	12	30	Cluster 3	Cluster 3

NAM	Namibya	46	5	16	Cluster 3	Cluster 3
AGO	Angola	45	8	13	Cluster 3	Cluster 3
ALB	Arnavutluk	43	10	30	Cluster 3	Cluster 3
VIR	US Virgin Adaları	40	6	12	Cluster 3	Cluster 3
FJI	Fiji	39	10	27	Cluster 3	Cluster 3
JOR	Ürdün	37	18	26	Cluster 3	Cluster 3
ECU	Ekvator	36	4	20	Cluster 3	Cluster 3
SOM	Somali	34	9	21	Cluster 3	Cluster 3
AFG	Afganistan	30	17	22	Cluster 3	Cluster 3
MNE	Karadağ	30	11	22	Cluster 3	Cluster 3
MKD	Kuzey Makedonya	28	4	21	Cluster 3	Cluster 3
CUB	Küba	27	12	15	Cluster 3	Cluster 3
PAN	Panama	27	7	14	Cluster 3	Cluster 3
ISL	İzlanda	26	11	12	Cluster 3	Cluster 3
NCL	Yeni Kaledonya	26	9	17	Cluster 3	Cluster 3
GTM	Guatemala	25	8	19	Cluster 3	Cluster 3
SYR	Suriye	25	8	17	Cluster 3	Cluster 3
BOL	Bolivya	24	4	11	Cluster 3	Cluster 3
CMR	Kamerun	24	11	16	Cluster 3	Cluster 3
HRV	Hırvatistan	24	3	10	Cluster 3	Cluster 3
ETH	Etiyopya	24	8	10	Cluster 3	Cluster 3
GEO	Gürcistan	24	7	20	Cluster 3	Cluster 3
BLR	Belarus	23	4	13	Cluster 3	Cluster 3
DOM	Dominik Cumhuriyeti	22	1	13	Cluster 3	Cluster 3
IRQ	Iraq	22	9	17	Cluster 3	Cluster 3
PRY	Paraguay	22	4	9	Cluster 3	Cluster 3
TZA	Tanzanya	22	3	10	Cluster 3	Cluster 3
BRN	Bruney	21	6	10	Cluster 3	Cluster 3
JEY	Jersey	21	5	15	Cluster 3	Cluster 3
LUX	Lüksemburg	21	6	11	Cluster 3	Cluster 3
MNG	Moğolistan	21	6	11	Cluster 3	Cluster 3
RWA	Ruanda	21	11	14	Cluster 3	Cluster 3
BWA	Botsvana	20	3	10	Cluster 3	Cluster 3
SVN	Slovenya	20	6	8	Cluster 3	Cluster 3
DZA	Cezayir	19	12	14	Cluster 3	Cluster 3
IMN	Man Adası	18	2	7	Cluster 3	Cluster 3
LTU	Litvanya	18	9	14	Cluster 3	Cluster 3
PSE	Filistin	18	4	12	Cluster 3	Cluster 3
PNG	Papua Yeni Gine	18	3	9	Cluster 3	Cluster 3
BRB	Barbados	17	7	7	Cluster 3	Cluster 3
GUY	Guyana	17	9	10	Cluster 3	Cluster 3
ROU	Romanya	17	7	14	Cluster 3	Cluster 3
SLE	Sierra Leone	17	5	12	Cluster 3	Cluster 3
BIH	Bosna Hersek	16	4	9	Cluster 3	Cluster 3
GMB	Gambiya	14	9	11	Cluster 3	Cluster 3

HND	Honduras	14	4	7	Cluster 3	Cluster 3
BTN	Bhutan	13	4	10	Cluster 3	Cluster 3
BFA	Burkina Faso	13	2	5	Cluster 3	Cluster 3
EST	Estonya	13	2	8	Cluster 3	Cluster 3
LBY	Libya	12	3	8	Cluster 3	Cluster 3
SLV	El Salvador	11	3	3	Cluster 3	Cluster 3
GRL	Grönland	11	1	1	Cluster 3	Cluster 3
UZB	Özbekistan	11	4	7	Cluster 3	Cluster 3
BMU	Bermuda	10	3	7	Cluster 3	Cluster 3
KAZ	Kazakistan	10	3	8	Cluster 3	Cluster 3
MLI	Mali	10	2	4	Cluster 3	Cluster 3
MOZ	Mozambik	10	3	6	Cluster 3	Cluster 3
SYC	Sejšeller	10	1	5	Cluster 3	Cluster 3
BHS	Bahamalar	9	3	6	Cluster 3	Cluster 3
COG	Kongo	9	2	5	Cluster 3	Cluster 3
GUM	Guam	9	5	5	Cluster 3	Cluster 3
XD	Arunaçhal Pradeş	8	4	4	Cluster 3	Cluster 3
VGB	British Virjin Adaları	8	1	4	Cluster 3	Cluster 3
HTI	Haiti	8	0	4	Cluster 3	Cluster 3
LAO	Laos	8	0	2	Cluster 3	Cluster 3
MAC	Makao	8	3	7	Cluster 3	Cluster 3
PRK	Kuzey Kore	8	1	3	Cluster 3	Cluster 3
KNA	Saint Kitts and Nevis	8	2	7	Cluster 3	Cluster 3
TLS	Doğu Timor	8	2	7	Cluster 3	Cluster 3
AND	Andorra	7	2	2	Cluster 3	Cluster 3
TJK	Tacikistan	7	0	4	Cluster 3	Cluster 3
URY	Uruguay	7	2	3	Cluster 3	Cluster 3
CIV	Kotdivuar	6	3	3	Cluster 3	Cluster 3
MWI	Malavi	6	1	2	Cluster 3	Cluster 3
NIC	Nikaragua	6	2	5	Cluster 3	Cluster 3
SHN	Saint Helena	6	1	3	Cluster 3	Cluster 3
SEN	Senegal	6	2	3	Cluster 3	Cluster 3
SDN	Sudan	6	2	6	Cluster 3	Cluster 3
CRI	Kosta Rika	5	2	4	Cluster 3	Cluster 3
COD	Kongo Demokratik Cumhuriyeti	5	0	3	Cluster 3	Cluster 3
FLK	Falkland Adaları	5	1	3	Cluster 3	Cluster 3
GAB	Gabon	5	1	3	Cluster 3	Cluster 3
GIB	Cebelitarık	5	1	4	Cluster 3	Cluster 3
MCO	Monako	5	0	3	Cluster 3	Cluster 3
SLB	Solomon Adaları	5	2	3	Cluster 3	Cluster 3
GIN	Gine	4	0	1	Cluster 3	Cluster 3
KGZ	Kırgızistan	4	3	2	Cluster 3	Cluster 3
LVA	Letonya	4	1	0	Cluster 3	Cluster 3
LBR	Liberya	4	1	1	Cluster 3	Cluster 3
LIE	Lihtenştayn	4	0	3	Cluster 3	Cluster 3

MDA	Moldova	4	0	3	Cluster 3	Cluster 3
TGO	Togo	4	2	4	Cluster 3	Cluster 3
ZMB	Zambiya	4	1	2	Cluster 3	Cluster 3
ATG	Antigua ve Barbuda	3	1	3	Cluster 3	Cluster 3
BEN	Benin	3	0	1	Cluster 3	Cluster 3
BDI	Burundi	3	1	2	Cluster 3	Cluster 3
CYM	Cayman Adaları	3	1	1	Cluster 3	Cluster 3
TCD	Çad	3	0	1	Cluster 3	Cluster 3
SWZ	Esvatini	3	0	2	Cluster 3	Cluster 3
LSO	Lesotho	3	0	0	Cluster 3	Cluster 3
MHL	Marşal Adaları	3	1	2	Cluster 3	Cluster 3
NER	Nijer	3	0	0	Cluster 3	Cluster 3
VUT	Vanuatu	3	0	1	Cluster 3	Cluster 3
YEM	Yemen	3	0	1	Cluster 3	Cluster 3
COK	Cook Adaları	2	1	1	Cluster 3	Cluster 3
GGY	Guernsey	2	0	0	Cluster 3	Cluster 3
MDG	Madagaskar	2	0	1	Cluster 3	Cluster 3
SSD	Güney Sudan	2	0	1	Cluster 3	Cluster 3
SUR	Surinam	2	0	0	Cluster 3	Cluster 3
TON	Tonga	2	0	0	Cluster 3	Cluster 3
COM	Komorlar	1	0	1	Cluster 3	Cluster 3
DJI	Cibuti	1	0	0	Cluster 3	Cluster 3
DMA	Dominika	1	0	0	Cluster 3	Cluster 3
GNQ	Ekvator Ginesi	1	0	0	Cluster 3	Cluster 3
FRO	Faroe Adaları	1	0	1	Cluster 3	Cluster 3
LCA	Saint Lucia	1	0	1	Cluster 3	Cluster 3
VCT	Saint Vincent ve Grenadinler	1	0	0	Cluster 3	Cluster 3
WSM	Samoa	1	0	1	Cluster 3	Cluster 3
SMR	San Marino	1	0	1	Cluster 3	Cluster 3
STP	Sao Tome ve Prinsipe	1	0	0	Cluster 3	Cluster 3
XN	Senkaku Adaları	1	0	1	Cluster 3	Cluster 3
SXM	Sint-Maarten	1	0	0	Cluster 3	Cluster 3
TKM	Türkmenistan	1	1	1	Cluster 3	Cluster 3
VAT	Vatikan	1	0	0	Cluster 3	Cluster 3