

KEŞİFSEL VERİ ANALİZİ ve GÖRSELLEŞTİRME / COVID-19

Simay UĞUR

07 06 2020

Contents

1 Giriş	3
2 Verilerin Okutulması	3
3 Verilerin Düzenlenmesi	3
3.1 Onaylanan, Vefat ve İyileşen Veri Setlerinin Birleştirilmesi	4
4 Dünya Çapında Durum	6
5 Veri Setine Yeni Değişkenler Eklenmesi ve Düzenlenmesi	6
5.1 Aktif Vaka Sayısının Belirlenmesi	6
5.2 Veri Setinin <i>ülke</i> ve <i>tarih</i> Değişkenlerine Göre Düzenlenmesi	6
5.3 Hatalı ve Tutarsız Değerlerin Düzenlenmesi	7
5.4 Elde Edilen Değişkenlere Göre Ölüm Oranlarının Hesaplanması	8
6 Harita Nesneleri ile Çalışmak	10
7 Verilerin Görselleştirilmesi	11
7.1 Verilerin Uzun Formata Dönüştürülmesi	11
7.2 Salgın Grafikleri Neden Logaritmiktir?	11
8 Dünya Çapında Vakaların Görselleştirilmesi	12
8.1 Türkiye’de Vakaların Durumunu Görselleştirme	13
9 Dünya Çapında, Aktif ve Günlük Onaylanan Vakalarının Görselleştirilmesi	14
9.1 Vefat Eden ve İyileşen Vakaları Görselleştirme	15
10 Ölüm Oranları Grafikleri	17
10.1 Dünyada’daki Ölüm Oranları Grafikleri	17
10.2 Türkiye’deki Ölüm Oranları Grafikleri	18
10.3 Türkiye, Amerika, Brezilya ve Çin’deki Yeni Onaylanan Vakalar	19
11 Onaylanan Vaka Sayısı En Yüksek İlk 15 Ülke	22
11.1 Vefat Eden, İyileşen ve Aktif Vakalar	26
12 SONUC	28
13 KAYNAKÇA	29

Kütüphaneler

```
library(magrittr) # pipe operations
library(lubridate) # date operations
library(tidyverse) # ggplot2, tidyr, dplyr...
library(gridExtra) # multiple grid-based plots on a page
library(ggforce) # accelerating ggplot2
library(leaflet) # map
library(tidycovid19)
library(remotes)
library(ggplot2)
library(RColorBrewer)
library(coronavirus)#
library(tidyverse)
library(coronavirus)
library(zoo)
library(tidycovid19)
library(yaml)
library(scales)
library(tinytex)
library(dplyr)
```

1 Giriş

Koronavirusler (CoV), soğuk algınlığından Orta Doğu Solunum Sendromu (MERS-CoV) ve Şiddetli Akut Solunum Sendromu (SARS-CoV) gibi daha ciddi hastalıklara kadar çeşitli hastalıklara neden olan büyük bir virüs ailesidir.

İlk kez Aralık ayında görülen ve “2019-nCoV” olarak adlandırılan virüs, Koronavirüs ya da Corona Virüsü adıyla biliniyor. Dünya Sağlık Örgütü (WHO) 11 Şubat’ta koronavirüs kaynaklı hastalığa Covid-19 adını verdi. Daha sonra küresel salgın anlamına gelen “pandemi” olarak ilan edildi.

WHO Başkanı Tedros Adhanom Ghebreyesus, Covid-19’un “korona”nın co’su, “virüs”ün vi’si ve hastalık kelimesinin İngilizcesi “disease” sözcüğünün d’sinden türetildiğini söyledi.

Türkiye’de Corona Virüsü 11 Mart 2020 tarihinden itibaren resmi olarak kayıtlara geçmiştir.

17 Haziran 2020 itibarıyla Dünya çapında SARS-CoV-2 virüsü nedeniyle 8.299.389 vaka, 446.929 ölüm ve 4.347.514 iyileşme gözlenmiştir.[1]

2 Verilerin Okutulması

```
data1.confirmed <-read.csv("https://raw.githubusercontent.com/CSSEGISandData/COVID-19/master/csse_covid_19_data/csse_covid_19_data/deaths <-read.csv("https://raw.githubusercontent.com/CSSEGISandData/COVID-19/master/csse_covid_19_data/csse_covid_19_data/recovered <-read.csv("https://raw.githubusercontent.com/CSSEGISandData/COVID-19/master/csse_covid_19_data/csse_covid_19_data/
```

Aşağıda verilerin hangi zaman aralığından itibaren analiz edildiği bilgisi görülmektedir. Onaylanan vakaların 1 Şubat 2020 den itibaren başlayan bugünü kapsayan veriler görülmektedir.

```
n.col <- ncol(data1.confirmed)
## get dates from column names
dates <- names(data1.confirmed)[5:n.col] %>% substr(2,8) %>% lubridate::mdy()
range(dates)
```

```
## [1] "2020-01-22" "2020-06-23"
```

Veri setinin başlangıç tarihi 22 Ocak 2020 ve son güncelleme tarihi 15 Haziran 2020 dir.

```
min.date <- min(dates)
max.date <- max(dates)
min.date.txt <- min.date %>% format('%d %b %Y')
max.date.txt <- max.date %>% format('%d %b %Y') %>% paste('UTC')
```

3 Verilerin Düzenlenmesi

Aşağıda Üç dosyada ayrı ayrı verilen veriler ülke ve tarih bilgisi tutarlı olacak şekilde tek bir veri nesnesinde birleştiren R kodları yazılmıştır. İşlemler sonucunda tüm veriler tek nesnede ulke, tarih, onaylanan_sayi, vefat_sayi, iyilesen_sayi gibi değişkenler ile yer almalıdır.

Veri seti için gerekli olmayan değişkenler (Province.State, Lat, Long) çıkartılmıştır.

Country.Region değişkeninin adı ülke olarak değiştirilmiştir. Tarihler uzun formattan geniş formata çevrilmiştir. Bu durum her ülke için günlük sayıları vermektedir.

Tarih değişkeninin başındaki X’ler kaldırılmış ve tarih formatına çevrilmiştir.

Bu işlemler 3 veri seti içinde gerçekleştirilmiştir.

- **Onaylanan:** Onaylanmış vakalar 14 Nisan’dan itibaren CDC yönergelerine uygun olarak olası **pozitif vakaları** ve **olası vakaları** içerir.

```
veriler.onaylanan <- data1.confirmed %>% select(-c(Province.State, Lat, Long))
veriler.onaylanan%<>% rename(ulke=Country.Region)
veriler.onaylanan %<>% gather(key=tarih, value=sayi, -ulke)
veriler.onaylanan %<>% mutate(tarih = tarih %>% substr(2,8) %>% mdy())
str(veriler.onaylanan)
```

```
## 'data.frame': 40964 obs. of 3 variables:
## $ ulke : Factor w/ 188 levels "Afghanistan",...: 1 2 3 4 5 6 7 8 9 9 ...
## $ tarih: Date, format: "2020-01-22" "2020-01-22" ...
## $ sayi : int 0 0 0 0 0 0 0 0 0 0 ...
```

- Vefat:

```
veriler.vefat <- data1.deaths %>% select(-c(Province.State, Lat, Long))
veriler.vefat %<>% rename(ulke=Country.Region)
veriler.vefat %<>% gather(key=tarih, value=sayi, -ulke)
veriler.vefat %<>% mutate(tarih = tarih %>% substr(2,8) %>% mdy())
str(veriler.vefat)
```

```
## 'data.frame': 40964 obs. of 3 variables:
## $ ulke : Factor w/ 188 levels "Afghanistan",...: 1 2 3 4 5 6 7 8 9 9 ...
## $ tarih: Date, format: "2020-01-22" "2020-01-22" ...
## $ sayi : int 0 0 0 0 0 0 0 0 0 0 ...
```

- İyileşen:

```
veriler.iyilesen <- data1.recovered %>% select(-c(Province.State, Lat, Long))
veriler.iyilesen %<>% rename(ulke=Country.Region)
veriler.iyilesen %<>% gather(key=tarih, value=sayi, -ulke)
veriler.iyilesen %<>% mutate(tarih = tarih %>% substr(2,8) %>% mdy())
str(veriler.iyilesen)
```

```
## 'data.frame': 38962 obs. of 3 variables:
## $ ulke : Factor w/ 188 levels "Afghanistan",...: 1 2 3 4 5 6 7 8 9 9 ...
## $ tarih: Date, format: "2020-01-22" "2020-01-22" ...
## $ sayi : int 0 0 0 0 0 0 0 0 0 0 ...
```

3.1 Onaylanan, Vefat ve İyileşen Veri Setlerinin Birleştirilmesi

Onaylanan, Vefat, İyileşen veri setlerinin tarih ve ülkelerinin tutarlı olması için ülke ve tarihe göre grupta yapararak sayı değerleri toplanmıştır.

Her veri setinin içindeki sayı değişkenleri yeni isim (olgusayısı, vefatsayısı, iyilesensayısı) atanmıştır.

İlk çıktı Onaylanan vakaların veri çerçevesini göstermektedir.İkinci çıktı Vefat edenlerin veri çerçevesinin gösterir. Üçüncü çıktı İyileşen vakaların veri setini bilgisini göstermektedir.

Belirtilen üç veri seti *Veriler* dataframe olarak birleştirilmiştir.

Veriler dataframe'inde 27072 gözlemden oluşmaktadır. 5 değişken bulunmaktadır bunlar: Ülkeler, tarih, onaylanan vakaların sayısı, vefat eden vakaların sayısı, iyileşen vakaların sayısıdır.

22 Ocak 2020 UTC'den beri, COVID-19 vakası doğrulanmış 188 ülke vardır.

```
veriler.onaylanan %<>% group_by(ulke, tarih) %>%
  summarise(sayi=sum(sayi, na.rm=T)) %>% as.data.frame()
veriler.vefat %<>% group_by(ulke, tarih) %>%
  summarise(sayi=sum(sayi, na.rm=T)) %>% as.data.frame()
veriler.iyilesen %<>% group_by(ulke, tarih) %>%
```

```

summarise(sayi=sum(sayi, na.rm=T)) %>% as.data.frame()

veriler.onaylanan %<>% rename(olgu_sayi=sayi)
veriler.vefat %<>% rename(vefat_sayi=sayi)
veriler.iyilesen %<>% rename(iyilesen_sayi=sayi)
str(veriler.onaylanan)

## 'data.frame': 28952 obs. of 3 variables:
## $ ulke : Factor w/ 188 levels "Afghanistan",...: 1 1 1 1 1 1 1 1 1 1 ...
## $ tarih : Date, format: "2020-01-22" "2020-01-23" ...
## $ olgu_sayi: int 0 0 0 0 0 0 0 0 0 0 ...

str(veriler.vefat)

## 'data.frame': 28952 obs. of 3 variables:
## $ ulke : Factor w/ 188 levels "Afghanistan",...: 1 1 1 1 1 1 1 1 1 1 ...
## $ tarih : Date, format: "2020-01-22" "2020-01-23" ...
## $ vefat_sayi: int 0 0 0 0 0 0 0 0 0 0 ...

str(veriler.iyilesen)

## 'data.frame': 28952 obs. of 3 variables:
## $ ulke : Factor w/ 188 levels "Afghanistan",...: 1 1 1 1 1 1 1 1 1 1 ...
## $ tarih : Date, format: "2020-01-22" "2020-01-23" ...
## $ iyilesen_sayi: int 0 0 0 0 0 0 0 0 0 0 ...

veriler <- veriler.onaylanan%>% merge(veriler.vefat, all=T) %>% merge(veriler.iyilesen, all=T)
str(veriler)

## 'data.frame': 28952 obs. of 5 variables:
## $ ulke : Factor w/ 188 levels "Afghanistan",...: 1 1 1 1 1 1 1 1 1 1 ...
## $ tarih : Date, format: "2020-01-22" "2020-01-23" ...
## $ olgu_sayi : int 0 0 0 0 0 0 0 0 0 0 ...
## $ vefat_sayi : int 0 0 0 0 0 0 0 0 0 0 ...
## $ iyilesen_sayi: int 0 0 0 0 0 0 0 0 0 0 ...

```

4 Dünya Çapında Durum

Aşağıda tüm dünya için günlük olgu.sayi, vefat.sayi, iyilesen.sayi değerlerini hesaplanmıştır. Dünya çapında hergün için toplam olgu sayısı, toplam vefat edenlerin sayısı, toplam iyileşenlerin sayısı görülmektedir.

```
data.dunya <- veriler %>% group_by(tarih) %>%
  summarise(ulke='dunya',
            olgu_sayi= sum(olgu_sayi, na.rm=T),
            vefat_sayi = sum(vefat_sayi, na.rm=T),
            iyilesen_sayi = sum(iyilesen_sayi, na.rm=T))
head(data.dunya)
```

```
## # A tibble: 6 x 5
##   tarih      ulke  olgu_sayi vefat_sayi iyilesen_sayi
##   <date>    <chr>    <int>    <int>    <int>
## 1 2020-01-22 dunya      555        17        28
## 2 2020-01-23 dunya      654        18        30
## 3 2020-01-24 dunya      941        26        36
## 4 2020-01-25 dunya     1434        42        39
## 5 2020-01-26 dunya     2118        56        52
## 6 2020-01-27 dunya     2927        82        61
```

Bulunan değerleri **veriler** dataframe'ne eklenmiştir.

```
veriler<-veriler%<>% rbind(data.dunya)
```

5 Veri Setine Yeni Değişkenler Eklenmesi ve Düzenlenmesi

5.1 Aktif Vaka Sayısının Belirlenmesi

Verilere aktif vaka sayısını gösteren (aktif.olgu) yeni bir değişken ekleyen R kodları aşağıdaki gibidir.

```
veriler %<>% mutate(aktif.olgu = olgu_sayi - vefat_sayi - iyilesen_sayi)
head(veriler)
```

```
##           ulke      tarih olgu_sayi vefat_sayi iyilesen_sayi aktif.olgu
## 1 Afghanistan 2020-01-22         0         0         0         0
## 2 Afghanistan 2020-01-23         0         0         0         0
## 3 Afghanistan 2020-01-24         0         0         0         0
## 4 Afghanistan 2020-01-25         0         0         0         0
## 5 Afghanistan 2020-01-26         0         0         0         0
## 6 Afghanistan 2020-01-27         0         0         0         0
```

5.2 Veri Setinin *ülke* ve *tarih* Değişkenlerine Göre Düzenlenmesi

Verilerin ülke ve tarihe göre sıralı olmasını sağlayan R kodları aşağıdaki gibidir. Veriler Tarih ve ülkeye göre düzenledi.

```
veriler %<>% arrange(ulke,tarih)
head(veriler,7)
```

```
##           ulke      tarih olgu_sayi vefat_sayi iyilesen_sayi aktif.olgu
## 1 Afghanistan 2020-01-22         0         0         0         0
## 2 Afghanistan 2020-01-23         0         0         0         0
## 3 Afghanistan 2020-01-24         0         0         0         0
## 4 Afghanistan 2020-01-25         0         0         0         0
## 5 Afghanistan 2020-01-26         0         0         0         0
```

```
## 6 Afghanistan 2020-01-27      0      0      0      0
## 7 Afghanistan 2020-01-28      0      0      0      0
```

5.3 Hatalı ve Tutarsız Değerlerin Düzenlenmesi

Her bir gün için *yeni.olgu*, *yeni.vefat* ve *yeni.iyilesen* sayılarını hesaplayan ve verilere ekleyen R kodlarına yer verilmiştir. (İlk gün sayılan “2020-01-22” için ‘NA’ değeri atanmıştır..) Değişkenler ‘lag’ fonksiyonu kullanarak 1 gecikme ile elde edilecektir.

```
n <- nrow(veriler)
day1 <- min(veriler$tarih)
veriler %<>%
  mutate(yeni.olgu = ifelse(tarih == day1, NA, olgu_sayi - lag(olgu_sayi,n=1)),
         yeni.vefat = ifelse(tarih == day1, NA, vefat_sayi - lag(vefat_sayi, n=1)),
         yeni.iyilesen = ifelse(tarih == day1, NA, iyilesen_sayi - lag(iyilesen_sayi, n=1)))
head(veriler)
```

```
##      ulke      tarih olgu_sayi vefat_sayi iyilesen_sayi aktif.olgu
## 1 Afghanistan 2020-01-22      0      0      0      0
## 2 Afghanistan 2020-01-23      0      0      0      0
## 3 Afghanistan 2020-01-24      0      0      0      0
## 4 Afghanistan 2020-01-25      0      0      0      0
## 5 Afghanistan 2020-01-26      0      0      0      0
## 6 Afghanistan 2020-01-27      0      0      0      0
##      yeni.olgu yeni.vefat yeni.iyilesen
## 1      NA      NA      NA
## 2      0      0      0
## 3      0      0      0
## 4      0      0      0
## 5      0      0      0
## 6      0      0      0
```

Hatalı veri girişi veya sonradan yapılan düzeltmeler nedeniyle günlük verilerde tutarsızlıklar olabilir.

yeni.olgu, *yeni.vefat* ve *yeni.iyilesen* değişkenlerini inceleyerek bu tutarsızlıkları belirleyen ve gideren R kodları aşağıdaki gibidir.

```
veriler%<>% mutate(yeni.olgu = ifelse(yeni.olgu < 0, 0, yeni.olgu),
                  yeni.vefat = ifelse(yeni.vefat < 0, 0, yeni.vefat),
                  yeni.iyilesen = ifelse(yeni.iyilesen < 0, 0, yeni.iyilesen))
head(veriler)
```

```
##      ulke      tarih olgu_sayi vefat_sayi iyilesen_sayi aktif.olgu
## 1 Afghanistan 2020-01-22      0      0      0      0
## 2 Afghanistan 2020-01-23      0      0      0      0
## 3 Afghanistan 2020-01-24      0      0      0      0
## 4 Afghanistan 2020-01-25      0      0      0      0
## 5 Afghanistan 2020-01-26      0      0      0      0
## 6 Afghanistan 2020-01-27      0      0      0      0
##      yeni.olgu yeni.vefat yeni.iyilesen
## 1      NA      NA      NA
## 2      0      0      0
## 3      0      0      0
## 4      0      0      0
## 5      0      0      0
## 6      0      0      0
```

5.4 Elde Edilen Değişkenlere Göre Ölüm Oranlarının Hesaplanması

Toplam ölümlere ve iyileşmiş vakalara dayalı olan *ölüm oranı* **ust.mortalite.oran** değişkeni olarak değiştirilmiştir ve veriler nesnesine eklenmiştir. Ölüm oranının üst sınırı ve nedenleri:

- 1) Koronavirüs patlak verdiğinde iyileşmiş vakalardan çok daha fazla ölüm oldu bunun nedeni virüsün kuluçka süresi 2-14 gün arası olmasından kaynaklanmaktadır.
 - 2) Zaman geçtikçe günlük ölüm sayısı azalacak ve iyileşen vaka sayısında artış olacaktır.
- **Üst Mortalite:** Üst mortalite orana bakılarak hangi ülkenin daha etkili önlemler ve tedaviler kullanılıp kullanılmadığı anlaşılabilir.

```
$ ust.mortalite.oran=100*vefat.sayi/(vefat.sayi+iyilesen.sayi) $
```

```
veriler %<>% mutate(ust.mortalite.oran = (100 * vefat_sayi / (vefat_sayi + iyilesen_sayi)) %>%  
round(1))
```

- **Alt Mortalite:** toplam teyit edilen vakalara bağlı ölüm oranı, alt.mortalite.oran değişkeni olarak tanımlanmıştır. Onaylanan vakalarda yeni ölümler vardır ve olmaya devam edecektir.

```
$ alt_sınır=100*vefat.sayi/olgu.sayi $
```

```
veriler %<>% mutate(alt.mortalite.oran = (100 * vefat_sayi / olgu_sayi) %>% round(1))
```

- **Günlük Mortalite:** Günlük gerçekleşen ölüm sayısı ve iyileşen vakaların sayısına bağlı olarak *günlük ölüm oranı*, günlük.oran değişkeni olarak tanımlanıp verilere eklenmiştir. Günlük oran yukarıdaki ikisinden daha değişkendir, oran daha çabuk değişir.

Bununla birlikte, bize mevcut durum hakkında bir ipucu verebilir: çok ciddi olup olmadığı veya daha iyiye giden bir süreç olduğu hakkında bilgi verebilir.

```
gunluk.oran = 100 * yeni.vefat / (yeni.vefat + yeni.iyilesen)
```

```
veriler %<>% mutate(gunluk.oran = (100 * yeni.vefat / (yeni.vefat + yeni.iyilesen )) %>%  
round(1))
```

```
head(veriler,2)
```

```
##          ulke      tarih olgu_sayi vefat_sayi iyilesen_sayi aktif.olgu  
## 1 Afghanistan 2020-01-22         0         0         0         0  
## 2 Afghanistan 2020-01-23         0         0         0         0  
##   yeni.olgu yeni.vefat yeni.iyilesen ust.mortalite.oran alt.mortalite.oran  
## 1         NA         NA         NA         NaN         NaN  
## 2         0         0         0         NaN         NaN  
##   gunluk.oran  
## 1         NA  
## 2        NaN
```

Covid-19 salgını her ülkede farklı tarihlerde başladığı için ölüm oranları olgu sayıları, iyileşen sayıları her ülke için farklılık gösterir ve oranladığımız değerler bazı ülkeler için 'NaN' değeri getirir. Aşağıdaki kod NaN olarak görülen yerlere 0 (sıfır) değeri atanmasını sağlar.

```
veriler$gunluk.oran=ifelse(veriler$gunluk.oran=="NaN",0,veriler$gunluk.oran)  
veriler$ust.mortalite.oran=ifelse(veriler$ust.mortalite.oran=="NaN",0,veriler$ust.mortalite.oran)  
veriler$alt.mortalite.oran=ifelse(veriler$alt.mortalite.oran=="NaN",0,veriler$alt.mortalite.oran)  
head(veriler,3)
```

```
##          ulke      tarih olgu_sayi vefat_sayi iyilesen_sayi aktif.olgu  
## 1 Afghanistan 2020-01-22         0         0         0         0  
## 2 Afghanistan 2020-01-23         0         0         0         0  
## 3 Afghanistan 2020-01-24         0         0         0         0
```


	yeni.olgu	yeni.vefat	yeni.iyilesen	ust.mortalite.oran	alt.mortalite.oran
## 1	NA	NA	NA	0	0
## 2	0	0	0	0	0
## 3	0	0	0	0	0

	gunluk.oran
## 1	NA
## 2	0
## 3	0

6 Harita Nesneleri ile Çalışmak

En son güne ait sayıları ayrı bir değişkene alındı. Gereken değişkenleri seçip, ülke ve güncel sayı değişkenini metin olarak birleştiriyoruz.

Güncel durum onaylanan covid-19 vakaları (en son güne ait sayılar) haritada gösterilmiştir.

Kırmızı noktaların büyüklüğü vaka sayıları ile doğru orantılıdır. Yeni tip Covid-19'dan en çok etkilenen kıta Amerika kıtasıdır.

Güncel onaylanan vakaların Türkiye'deki sayısı 178239 dir. İtalya'da güncel vakaların sayısı 236989, Yunanistan'da onaylanmış vakaların sayısı 3121 dir. Amerikada Güncel vaka sayısı 2.163.290 dir. En yüksek vaka Amerika'da ikinci en yüksek vaka Brezilya'dadır.

```
x <- data1.confirmed
x$confirmed <- x[, ncol(x)]
x %<>% select(c(Country.Region, Province.State, Lat, Long, confirmed)) %>%
  mutate(txt=paste0(Country.Region, ' - ', Province.State, ': ', confirmed))
m <- leaflet(width=1500, height=900) %>% addTiles()
m %<>% addCircleMarkers(x$Long, x$Lat,
  radius=0.03*sqrt(x$confirmed),
  stroke=F,color='red', fillOpacity=0.3,popup=x$txt)
m
```

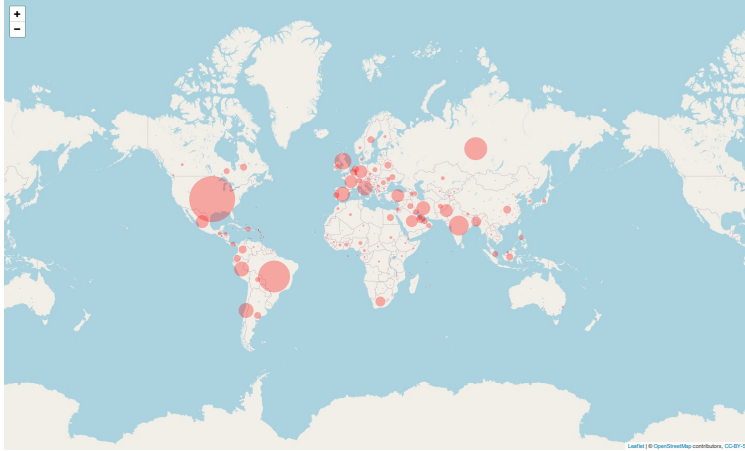


Figure 1: Dünya Verisi için Onaylanan Vaka Sayısı

7 Verilerin Görselleştirilmesi

Veriler nesnesindeki sayı değişkenleri ile oran değişkenlerini ayırıp iki ayrı veri nesnesi oluşturuldu. Veriler nesnesinde ki sayı değişkenleri ve oran değişkenleri farklı anlamlar ifade ettikleri için 2 ayrı veri çerçevesine atandı. Bu durumda Görselleştirmelerde daha etkili ve uygun sonuçlar alınacaktır.

```
data.long <- veriler %>%
  select(c(ulke, tarih, olgu_sayi, aktif.olgu, iyilesen_sayi, vefat_sayi))
rates.long <- veriler %>%
  select(c(ulke, tarih, ust.mortalite.oran, alt.mortalite.oran, gunluk.oran ))
```

7.1 Verilerin Uzun Formata Dönüştürülmesi

Günlük onaylanan, iyileşen ve vefat eden vakaların sayılarını her ülke için belirtebilmek için veri seti geniş formattan uzun formata dönüştürülmüştür.

```
data.long <- data.long %>%
  gather(key=type, value=count, -c(ulke, tarih))
rates.long <- rates.long %>%
  # mutate(country=factor(country, levels=top.countries)) %>%
  gather(key=type, value=count, -c(ulke, tarih))
head(rates.long, 3)
```

```
##           ulke      tarih      type count
## 1 Afghanistan 2020-01-22 ust.mortalite.oran    0
## 2 Afghanistan 2020-01-23 ust.mortalite.oran    0
## 3 Afghanistan 2020-01-24 ust.mortalite.oran    0
```

```
head(data.long, 2)
```

```
##           ulke      tarih      type count
## 1 Afghanistan 2020-01-22 olgu_sayi    0
## 2 Afghanistan 2020-01-23 olgu_sayi    0
```

data.long nesnesindeki tanım değişkeninin değerleri, olgu, aktif, iyilesen, vefat olarak tekrar düzenleyen ve faktör türüne dönüştüren R kodları aşağıda yer almaktadır.

rates.long nesnesindeki tanım değişkeninin değerlerini, gunluk mortalite, alt mortalite, ust mortalite olarak tekrar düzenlendi ve factor türüne dönüştürülmüştür.

```
data.long %<>% mutate(type=recode_factor(type, olgu_sayi='Olgu',
                                           aktif.olgu='Aktif',
                                           iyilesen_sayi='iyilesen' , vefat_sayi='Vefat'))

rates.long %<>% mutate(type=recode_factor(type, gunluk.oran='gunluk.mortalite', alt.mortalite.oran='alt.mortalite',
                                           ust.mortalite.oran='ust.mortalite'))
library("leaflet")
head(data.long, 2)
```

```
##           ulke      tarih type count
## 1 Afghanistan 2020-01-22 Olgu    0
## 2 Afghanistan 2020-01-23 Olgu    0
```

7.2 Salgın Grafikleri Neden Logaritmiktir?

Matematikçi Dr. Salih Durhan, bunun sebebinin doğrusal grafiklerin ilk günlerdeki vaka sayılarını görmek ve zaman içerisindeki artış hızını doğru yorumlayabilmek için kullanışlı olmamasından kaynaklı olduğunu söylüyor:

- “İlk günlerde bir günde yaşanan ufak artışlar, son günlerdeki binlerce artışla karşılaştırıldığında çok küçük ve önemsiz gelebilir. Ancak artış hızlarını karşılaştırdığımızda ilk günlerdeki vaka artış hızının son günlere göre çok daha yüksek olduğunu görebiliriz. Artış hızının nasıl değiştiğini kolayca görmenin yolu logaritmik tablolara bakmaktır. Böylelikle çok büyük sayılarla küçük sayıları aynı grafik üzerinde rahatça görebiliriz.”

8 Dünya Çapında Vakaların Görselleştirilmesi

Dünya çapındaki durumu tarihe göre grafikler ile göstermek için gereken R kodları yazılmıştır. Sayıları ve logaritmik ölçeğe dönüştürülmüş sayıları kullanarak iki ayrı grafik elde edilmiştir.

Grafiklerde iyileşen, vefat, aktif ve toplam sayılarını birlikte gösteren çizgi ve alan grafiği olarak gösterilmiştir.

Alan grafiğinde aktif vaka sayısı Şubat ayından Mart ayının sonuna kadar yavaş bir artış göstermektedir. Pandeminin dünyaya yayılmasıyla birlikte Nisan ayının başından itibaren Haziran ayının ortasına kadar çok hızlı bir şekilde vaka sayıları artmıştır.

İyileşen vaka sayısı Mart ayından itibaren başlamaktadır. Nisan ayından itibaren Haziran ayının ortasına kadar iyileşen vaka sayısı artış göstermektedir.

Vefat eden vaka sayısı Haziran ayına kadar artmış ve Haziran ayından sonra sabit hale gelmiştir.

Çizgi grafiğinde kullandığımız **log ölçeğinde** her şey doğrusal hale gelir ve ülkeler arasındaki F(x) karşılaştırmasını gözlemlemeyi kolaylaştırır. Sağdaki grafikte vaka sayılarının 10'un kuvvetleri şeklinde arttığını görüyoruz.

Çizgi grafiğine bakıldığında aktif vaka sayısı ve olgu sayısı Şubat ayının ortasına kadar aynı desende ilerlemiştir.

İyileşen ve vefat eden vaka sayısı Şubat ayına kadar aynı sayıda artmıştır. İyileşen vaka sayısı Şubat ayından sonra hızlı bir şekilde artmıştır. Mart ayının son haftasında iyileşen vaka sayısı

Nisan ayının başına kadar sabit ilerlemiştir. Nisan ayından itibaren Haziran ayının başına kadar iyileşen vaka sayısı artmış ve Haziran ayından itibaren Aktif vaka sayısını geçmiştir.

Nisan ayının ortalarından itibaren Haziran ayının başına kadar vefat eden vaka sayısı küçük artışlarla devam etmiştir.

```
world.long <- data.long %>% filter(ulke == 'dunya')

dunya.alan.graph <- world.long %>% filter(type != 'Olgu') %>%
  ggplot(aes(x=tarih, y=count)) +
  geom_area(aes(fill=type), alpha=0.5) +
  labs(title=paste0('Dünya Capında Vaka Sayısı (günlük ölçeği) - ',max.date.txt)) +
  scale_fill_manual(values=c( 'green', 'black','blue')) +
  theme(legend.title=element_blank(), legend.position='bottom',
        plot.title = element_text(size=8),
        axis.title.x=element_blank(),
        axis.title.y=element_blank(),
        legend.key.size=unit(0.2, 'cm'),
        legend.text=element_text(size=9),
        axis.text=element_text(size=9),
        axis.text.x=element_text(angle=45, hjust=1))

dunya.cizgi.graph <- world.long %>%
  ggplot(aes(x=tarih, y=count)) +
  geom_line(aes(color= type),size=1) +
  labs(title=paste0('Dünya Capında Vaka Sayilari (Log Olcek)', max.date.txt)) +
```

```

scale_color_manual(values=c('purple', 'red', 'green', 'black'))+
theme(legend.title=element_blank(), legend.position='bottom',
      plot.title = element_text(size=9),
      axis.title.x=element_blank(),
      axis.title.y=element_blank(),
      legend.key.size=unit(0.2, 'cm'),
      legend.text=element_text(size=9),
      axis.text=element_text(size=9),
      axis.text.x=element_text(angle=45, hjust=1)) +
geom_hline(aes(yintercept = max(world.long$count)),
           size=1,color="red")+
scale_y_continuous(trans='log10',
                   labels = trans_format("log10", math_format(10^.x)),
                   breaks = c(1000,10000,100000,1e05))

grid.arrange(dunya.alan.graph, dunya.cizgi.graph, ncol=2)

```

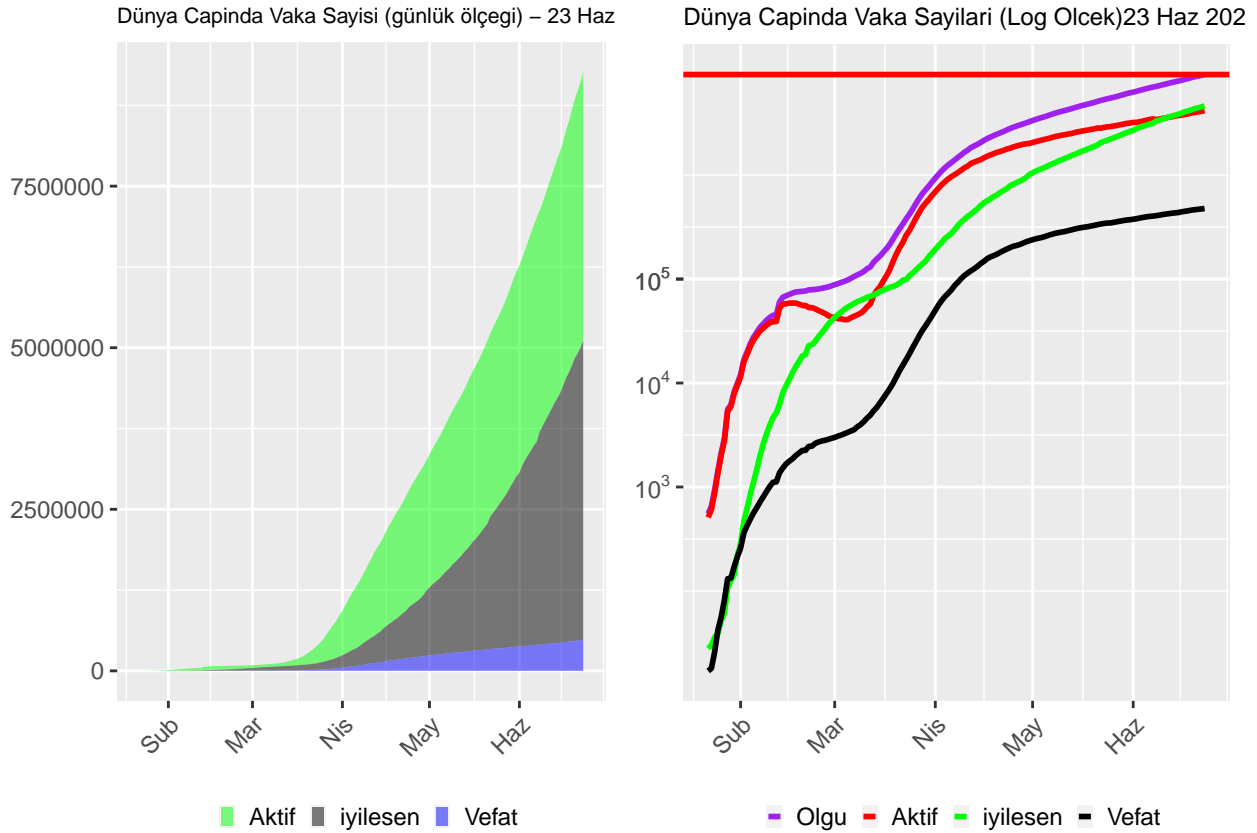


Figure 2: Dünya Çapında Vakaların Alan ve Çizgi Grafiği

8.1 Türkiye’de Vakaların Durumunu Görselleştirme

11 Mart Türkiye’de ilk Covid-19 vakası görülmesinden itibaren grafiğe göre aktif vaka sayısı görülmeye devam etmektedir. İlk iyileşen vakalar Nisan ayının ilk haftalarında.Dünyadaki aktif ve iyileşen vakaların alan grafiği ile Türkiye’deki aktif vaka ve iyileşen vakaların alan grafiğini karşılaştıracak olursak Türkiye’de iyileşen vakaların sayısı aktif vakalardan dünyaya göre daha fazladır.Bunun nedeni Türkiye’nin aldığı sınır kapılarını

kapama ,karantina dönemleri,Günlük test sayılarının fazlalığı , Hastalığın evrelerine göre verilen ilaçlar gibi etkenler söz konusu olabilir.

İyileşenlerin artış hızı Olguların artış hızına benzer şekilde artmıştır.Hastalığın kuluçka döneminden dolayı ortaya çıkış aralığı farklıdır.Vefat edenler Mart ayının yaklaşık ikinci haftasından itibaren hızlı yükselmiş ,karantina süreci,alınan önlemler ve mevsimsel etkilerden dolayı Mayıs ayının ilk haftalarından itibaren ölümlerin artış hızı azalan şekilde artmıştır.Aktif vakalar ile Onaylanan vakaların artışı Nisan ayının ortasına kadar aynı şekilde artarak devam etmiştir. Nisan ayının son haftalarından itibaren alınan önlemler karantina vb sonra aktif vaka sayısı artış hızı azalmaya başlamıştır.

```
son.tarih <- max(world.long$tarih)
turkiye.long <- veriler %>% filter(ulke == 'Turkey', tarih>"2020-03-10")
```

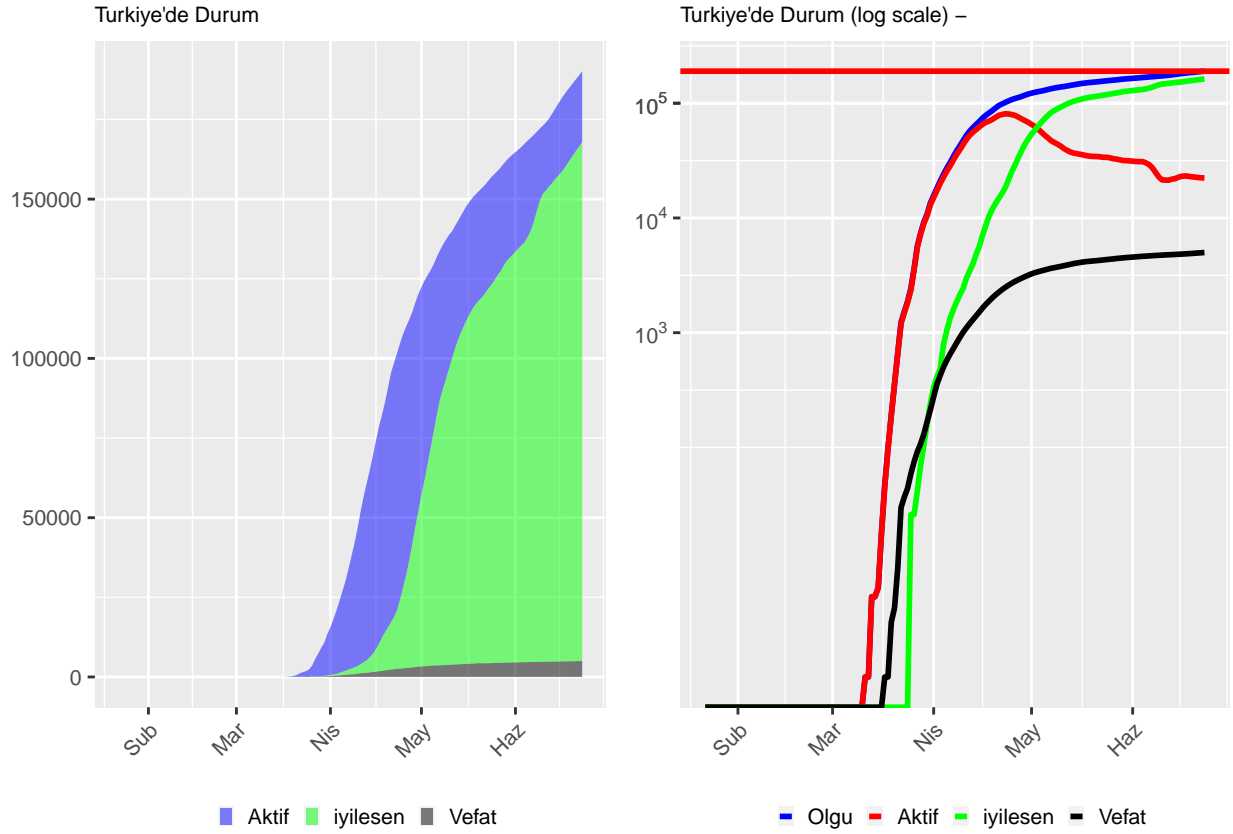


Figure 3: Türkiye

9 Dünya Capında, Aktif ve Günlük Onaylanan Vakalarının Görselleştirilmesi

Dünya çapındaki aktif olgular (aktif.olgu) ve günlük yeni (yeni.olgu) olgular grafiksel olarak iki ayrı şekilde gösterilmiştir.

```
data.dunya <- veriler %>% filter(ulke == 'dunya')
n <- nrow(data.dunya)
```

Mevcut covid-19 vakaları (soldaki tabloya bakınız) ve yeni teşhis konulan günlük covid-19 (sağdaki tabloya bakınız) artışı gösterilmektedir.Mavi çizgiler düzeltilmiş koşullu ortalamalardır ve etraflarındaki gri bant%

95 güven aralığını gösterir.

Dünyada aktif covid-19 vakası 4 milyonu geçmiştir. Salgın şiddetli bir şekilde devam etmektedir. Özellikle şuan kış mevsiminde olan Güney Amerika ülkeleri ,nüfusu oldukça fazla olan Amerika ve Avrupada pandemi etkisini önemli ölçüde göstermektedir.

```
aktif.olgu.graph <- ggplot(data.dunya, aes(x=tarih, y=aktif.olgu)) +  
  geom_smooth( method = 'loess' )+  
  geom_point() +  
  xlab('') + ylab('Sayı') + labs(title="Dunya'da Aktif Olgu") +  
  theme(axis.text.x=element_text(angle=45, hjust=1))+  
  scale_y_continuous()  
  
gunluk.oran.graph <- ggplot(data.dunya, aes(x=tarih, y=yeni.olgu)) +  
  geom_point() + geom_smooth() +  
  xlab('') + ylab('Sayı') + labs(title="Dunya'da Gunluk Yeni Olgu") +  
  theme(axis.text.x=element_text(angle=45, hjust=1))  
  
grid.arrange(aktif.olgu.graph,gunluk.oran.graph, ncol=2)
```

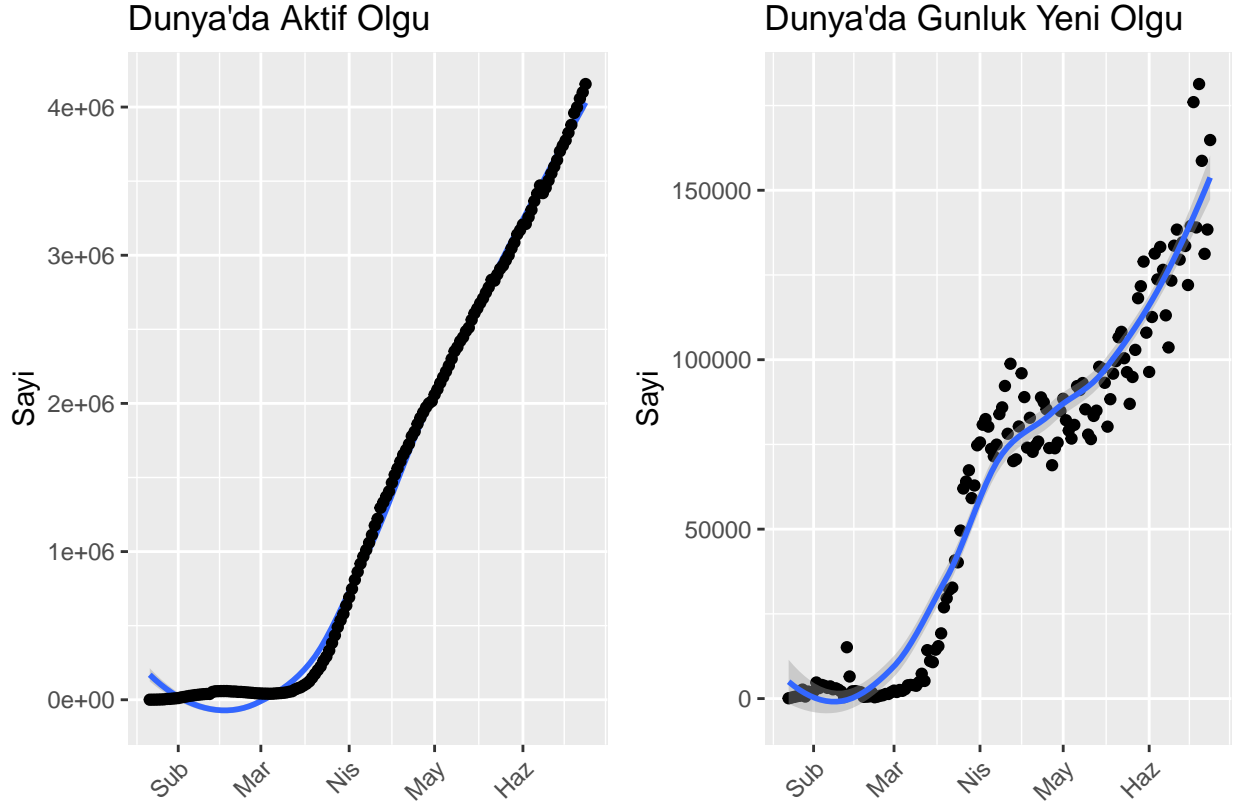


Figure 4: Aktif Olgu ve Günlük Yeni Olgu Grafiği

9.1 Vefat Eden ve İyileşen Vakaları Görselleştirme

Vefat edenlerin grafiği Nisadan bu yana sabit hızla artmaya devam etmektedir.

İyileşen vakaların grafiğine baktığımızda iyileşme logaitmik yani üssel şekilde gerçekleşmiştir. Yani iyileşen

sayısı katlanarak artmıştır.

Yeni iyileşen vakaların grafiği toplamda iyileşen vakaların grafiğiyle örtüşmektedir.

Şubattan mayıs başına kadar artış gösteren yeni vefat edenlerin grafiği mayıs başından itibaren azalışa geçmiştir.

```
olum.graph <- ggplot(data.dunya, aes(x=tarih, y=vefat_sayi)) +  
  geom_point() + geom_smooth() +  
  xlab('') + ylab('Sayı') + labs(title='Vefat Eden Vaka Grafiği') +  
  theme(axis.text.x=element_text(angle=45, hjust=1))  
  
iyilesen.graph <- ggplot(data.dunya, aes(x=tarih, y=iyilesen_sayi)) +  
  geom_point() + geom_smooth() +  
  xlab('') + ylab('Sayı') + labs(title='İyileşen Vaka Grafiği') +  
  theme(axis.text.x=element_text(angle=45, hjust=1))  
  
yeni.olum.graph <- ggplot(data.dunya, aes(x=tarih, y=yeni.vefat)) +  
  geom_point() + geom_smooth() +  
  xlab('') + ylab('Sayı') + labs(title='Yeni Vefat Eden Vaka Grafiği') +  
  theme(axis.text.x=element_text(angle=45, hjust=1))  
  
yeni.iyilesen.graph <- ggplot(data.dunya, aes(x=tarih, y=yeni.iyilesen)) +  
  geom_point() + geom_smooth() +  
  xlab('') + ylab('Sayı') + labs(title='Yeni İyileşen Vaka Grafiği') +  
  theme(axis.text.x=element_text(angle=45, hjust=1))  
  
grid.arrange(olum.graph, iyilesen.graph, yeni.olum.graph, yeni.iyilesen.graph, nrow=2)
```

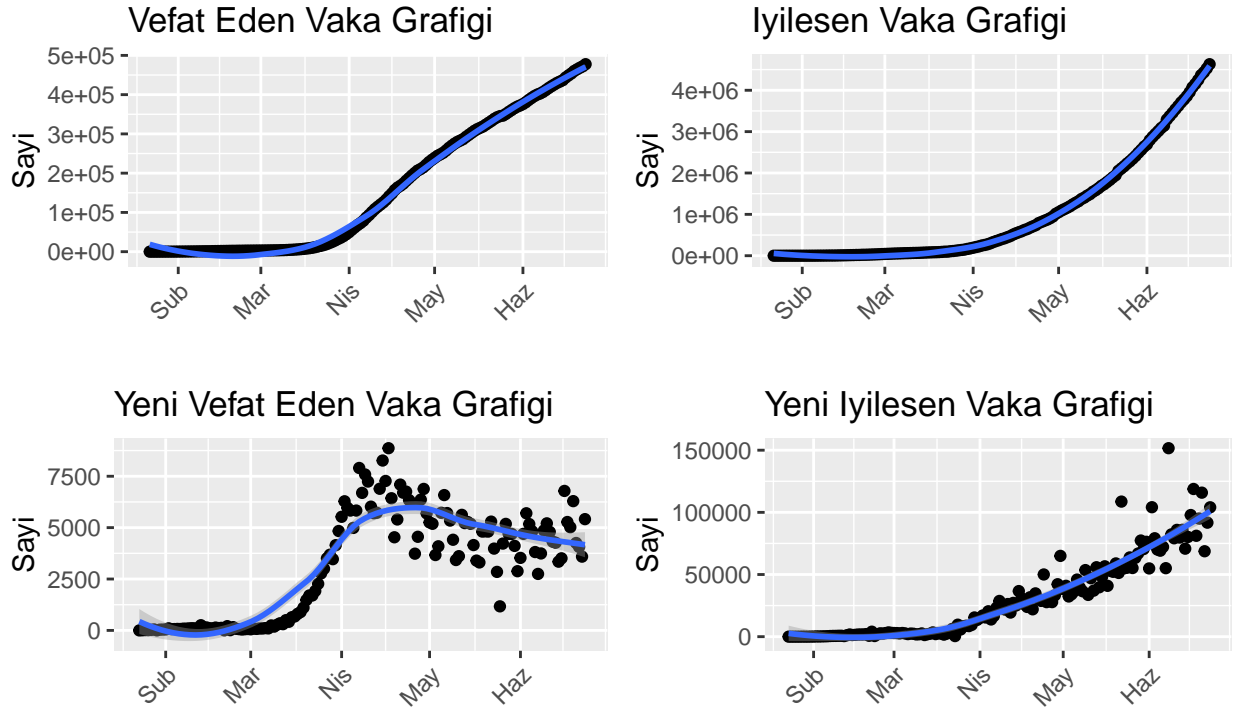


Figure 5: Vefat Eden - İyileşen Vaka Sayılarının Grafikleri

10 Olum Oranları Grafikleri

Ölüm oranları üç farklı şekilde hesaplanmıştır:Alt Mortalite Oran,Üst Mortalite Oran,Günlük Oran.

10.1 Dünyada'daki Ölüm Oranları Grafikleri

Dünya genelinde üst mortalite sınırı azalmıştır,bu bize İyileşen vaka sayılarının gerçekleşen vefatlara göre daha çok arttığını gösterir.

Mart ayı sonlarında günlük ölüm oranındaki (yeşil) artış, durumun önemli ölçüde değiştiğini gösteriyor (aslında kötüleşiyor) ve üst / alt sınırların üstünde kısa bir süre artmaya devam etmiştir.Muhtemel bir nedeni bu dalgalanma İran, Avrupa ve ABD'deki koronavirüs salgının etkisinin en çok hissedildiği yerler olmasıdır.Günlük ölüm oranı pandemi sürecinin başında ve Mart ayından Nisan ayının ortasına kadar Üst ölüm değerlerinin de üstünde kalmıştır.

Dünya için ölüm oranlarının grafiğine bakıldığında Şubat ayının sonundan Mart ayının başına kadar ve Mayıs ayının sonundan belirtilen son güne kadar günlük ölüm oranı alt ölüm oranının altında kaldığı görülmektedir.Bu bize günlük iyileşen vakaların vefat eden hastalara göre sayılarının fazla olduğunu gösterir.Bu durumda günlük ölüm oranı pandemi başında ve Mart ayından Nisan ayına kadar oldukça yüksek ve Haziran ayından itibaren de düşük oranlarla devam etmektedir.

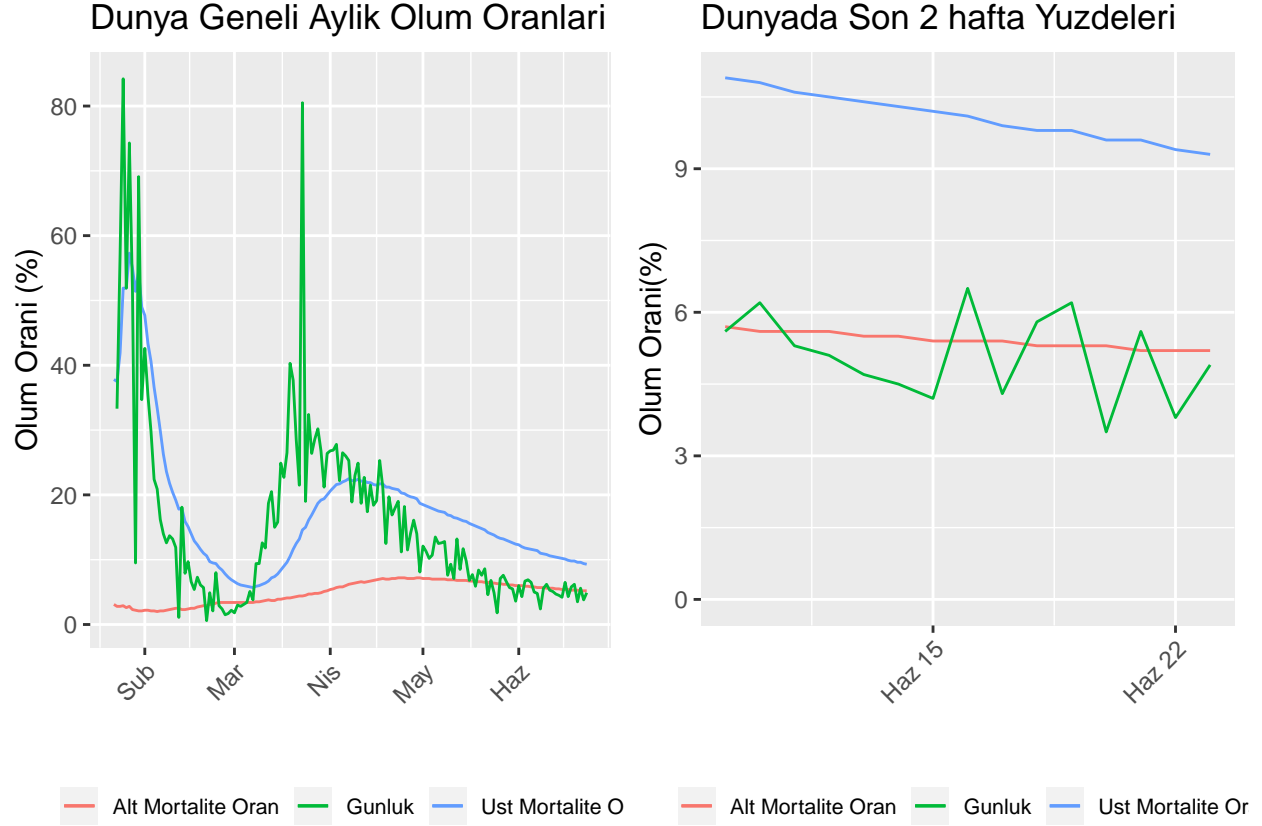
Mart ayının başından Mayıs ayının ilk haftasına kadar alt mortalite oran artmış,mayıstan günümüze kadar sabit ilerlemiştir.Bu bize hastalığın öldürücülüğü hakkında oransal değerler vermemizi sağlar.

Son 2 hafta ölüm oranına bakacak olursa üst mortalite oran azalmış,Alt mortalite oran sabit ve günlük ölüm oranı alt mortalite sınırının çevresinde artan azalan şekilde değişkenlik gösterir.

```
plot8 <- ggplot(data.dunya, aes(x=tarih)) +  
  geom_line(aes(y=ust.mortalite.oran, colour='Ust Mortalite Oran')) +  
  geom_line(aes(y=alt.mortalite.oran, colour='Alt Mortalite Oran')) +  
  geom_line(aes(y=gunluk.oran, colour='Gunluk')) +  
  xlab('') + ylab('Olum Orani (%)') + labs(title='Dunya Geneli Aylık Olum Oranlari') +  
  theme(legend.position='bottom', legend.title=element_blank(),  
        legend.text=element_text(size=8),  
        legend.key.size=unit(0.5, 'cm'),  
        axis.text.x=element_text(angle=45, hjust=1))
```

```
y.max <- data.dunya[n-(14:0), ] %>% select(ust.mortalite.oran, alt.mortalite.oran, gunluk.oran) %>% max
```

```
plot9 <- ggplot(data.dunya[n-(14:0),], aes(x=tarih)) +  
  geom_line(aes(y=ust.mortalite.oran, colour='Ust Mortalite Oran')) +  
  geom_line(aes(y=alt.mortalite.oran, colour='Alt Mortalite Oran')) +  
  geom_line(aes(y=gunluk.oran, colour='Gunluk')) +  
  xlab('') + ylab('Olum Orani(%)') + labs(title='Dunyada Son 2 hafta Yuzdeleri') +  
  theme(legend.position='bottom', legend.title=element_blank(),  
        legend.text=element_text(size=8),  
        legend.key.size=unit(0.5, 'cm'),  
        axis.text.x=element_text(angle=45, hjust=1)) +  
  ylim(c(0, y.max))  
grid.arrange(plot8, plot9, ncol=2)
```



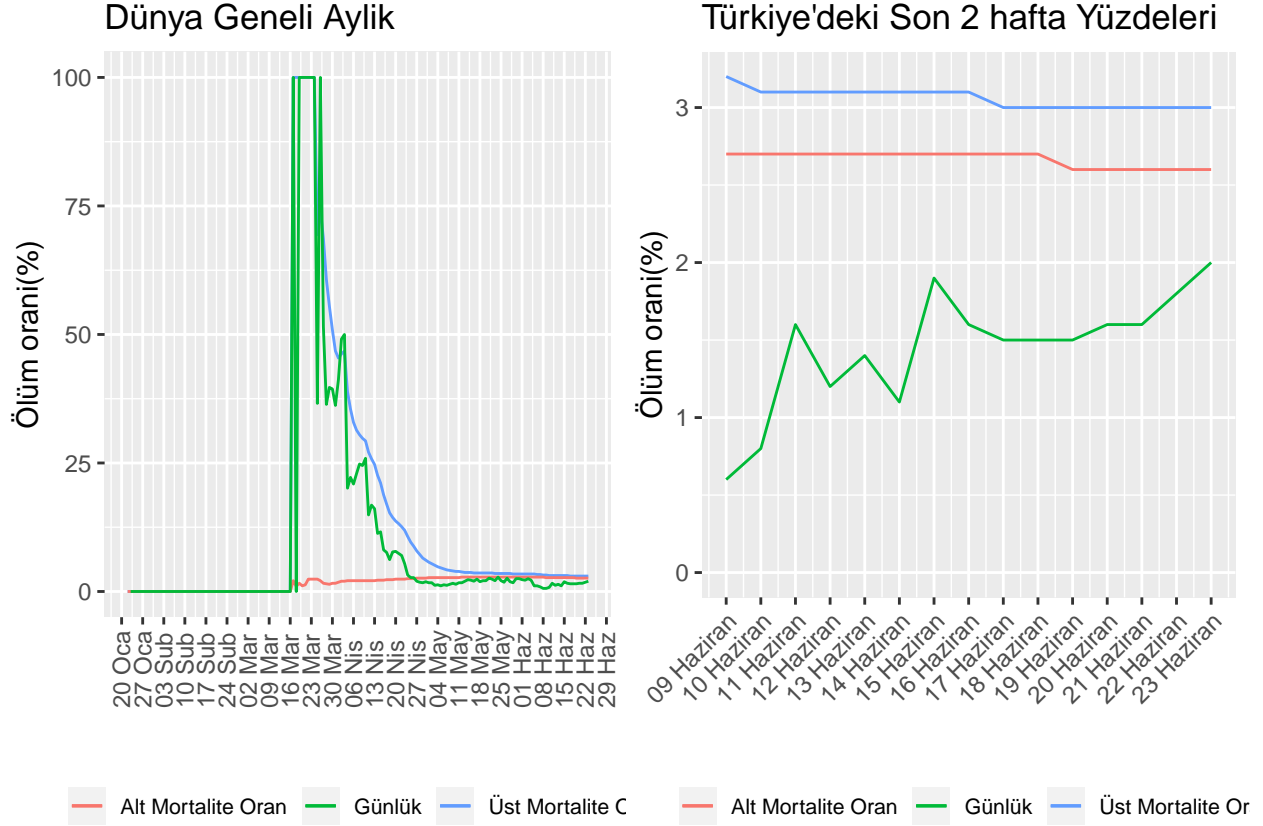
10.2 Türkiye'deki Olum Oranları Grafikleri

Üstte oluşturulan grafikler Türkiye verilerini kullanarak tekrar oluşturulmuştur.

Günlük ölüm oranı en yüksek 16 Mart 23 Mart haftalarında. Bunun nedeni covid-19 hastalığının yeni ülkemizde görülmeye başlaması ,hastalığın ilk ortaya çıkma ve virüsten tamamen kurtulma süresin 15-20 gün gibi uzun süre olması ve dolayısıyla iyileşen vakaların o haftalarda görülememesinden kaynaklanmaktadır. Günlük ölümler Mart ayının 2. ve 3. haftalarından itibaren iyileşen vakaların olmasıyla düşmeye başlamıştır.

Nisanın ilk haftalarından itibaren Üst Mortalite oran giderek azalmıştır ,bu iyileşen vakarın sayısının vefat edenlere göre fazla olduğunun göstergesidir. Türkiyede karantina uygulamaları başlamış ,sağlık seötörünün süreci kontrol altına alması kullanılan tedavi yöntemleri,yapılan test sayılarının arttırılması virüsün yayılmasının kontrol altına alınmaya çalışılması iyileşen sayılarında artışa yeni onaylanan vakaların sayısının azalmasına neden olmuştur.

Grafiğe yakından bakarsak son 2 hafta yüzdelere göre alt mortalite oran 19 hazirandan itibaren çok az azalmıştır,onaylanan vaka sayılarının arttığını gösterir,buna normalleşme süreci başlamasından .Üst mortalite oran 17 hazirandan itibaren %3 çizgisinde devam etmektedir.



Dağılım Grafikleri

10.3 Türkiye, Amerika, Brezilya ve Çin'deki Yeni Onaylanan Vakalar

Türkiyede Yeni vakaların aylara göre histogram grafiği aşağıda gösterilmiştir.Yeni onaylanan vakalar en çok Nisan ayı boyunca görülmüştür Nisan ayının yaklaşık ikinci haftasından itibaren yani salgının Türkiye'deki 5. haftasında vaka sayıları en yüksek noktayı görmüş ve azalmaya başlamıştır.Alınan karantina önlemleri,maske takma zorunluluğu,filyasyon yöntemi yeni vaka sayılarının Nisan ayının son haftalarında azalmaya başlamasının sonucudur.1 Haziran normalleşme süreci başlamıştır ve vakaların sayısında tekrar artış olduğu grafikte görülmektedir.

Amerika'daki Yeni vakaların aylara göre histogram grafiğini inceleyecek olursak dağılım sola çarpıktır. Mart ayının son haftalarından itibaren hızlı bir şekilde artan vaka sayıları Nisan ayında 30000'i aşmıştır.Nisanın son haftasında Haziran aya kadar vaka sayıları 20000 civarına gerilemiştir ,karantinanın belirli eyaletlerde uygulanması,vaka sayılarının çokluğu ve virüsün bulaşıcılığının fazla olması vaka sayılarının hızlı azalmasına engel olmuştur.Haziran ayının ikinci haftasından itibaren vaka sayılarında artış başlamıştır.

Brezilya Yeni vakaların aylara göre histogram grafiğinde dağılım sola çarpıktır,Vaka sayıları Nisan ayından itibaren artmaya başlamıştır.Haziranın ilk haftasından itibaren vaka sayıları yirmi bini ((20000) geçmiştir,Haziran ayının üçüncü haftası yeni vaka sayıları en çok olan haftadır 40000'i aşkın yeni vakalar vardır.Nedenleri iklim kaynaklı olması,vaka sayılarının çok fazla olması sağlık sisteminde ki vaka fazlalığından dolayı yük ve virüsün bulaşımın fazla olması nedenler arasında gösterilebilir.

Çin Yeni vakaların aylara göre histogram grafiğine baktığımızda veriler sağa çarpık dağılım göstermektedir.Çinde covid-19 yeni vakaları Şubattan öncesine dayanmaktadır ve salgın kontrol altına alınmış gözükmemektedir ve yeni vakaların sayısı oldukça azdır.Çinin bu süreçte aldığı uzun karantina önlemleri,filyasyon yöntemi uygulamaları salgını önemli ölçüde durdurmada başarılı olduklarını göstermişti,yeni vakaların dağılım grafiği de bunu gösterir.

```

df <- download_merged_data(cached = TRUE, silent = TRUE)

g<-df %>%
  filter(iso3c == "TUR") %>%
  rename(tarih=date)%>%
  mutate(
    yeni_vakalar = confirmed - lag(confirmed),
    ave_new_cases = rollmean(yeni_vakalar, 7, na.pad=TRUE, align="right")
  ) %>%
  filter(!is.na(yeni_vakalar), !is.na(ave_new_cases)) %>%
  ggplot(aes(x = tarih)) +
  geom_bar(aes(y = yeni_vakalar), stat = "identity", fill = "mistyrose3") +
  geom_line(aes(y = ave_new_cases), color = "mediumorchid4",size=1) +
  labs(title="Türkiye") +
  theme_minimal()

g1<-df %>%
  filter(iso3c == "USA") %>%
  rename(tarih=date)%>%
  mutate(
    yeni_vakalar = confirmed - lag(confirmed),
    ave_new_cases = rollmean(yeni_vakalar, 7, na.pad=TRUE, align="right")
  ) %>%
  filter(!is.na(yeni_vakalar), !is.na(ave_new_cases)) %>%
  ggplot(aes(x = tarih)) +
  geom_bar(aes(y = yeni_vakalar), stat = "identity", fill = "lightsteelblue") +
  geom_line(aes(y = ave_new_cases), color = "mediumorchid4",size=1) +
  labs(title="Amerika") +
  theme_minimal()

g2<-df %>%
  filter(iso3c == "BRA") %>%
  rename(tarih=date)%>%
  mutate(
    yeni_vakalar = confirmed - lag(confirmed),
    ave_new_cases = rollmean(yeni_vakalar, 7, na.pad=TRUE, align="right")
  ) %>%
  filter(!is.na(yeni_vakalar), !is.na(ave_new_cases)) %>%
  ggplot(aes(x = tarih)) +
  geom_bar(aes(y = yeni_vakalar), stat = "identity", fill = "palevioletred") +
  geom_line(aes(y = ave_new_cases), color = "mediumorchid4",size=1) +
  labs(title="Brezilya") +
  theme_minimal()

g3<-df %>%
  filter(iso3c == "CHN") %>%
  rename(tarih=date)%>%
  mutate(
    yeni_vakalar = confirmed - lag(confirmed),
    ave_new_cases = rollmean(yeni_vakalar, 7, na.pad=TRUE, align="right")
  ) %>%

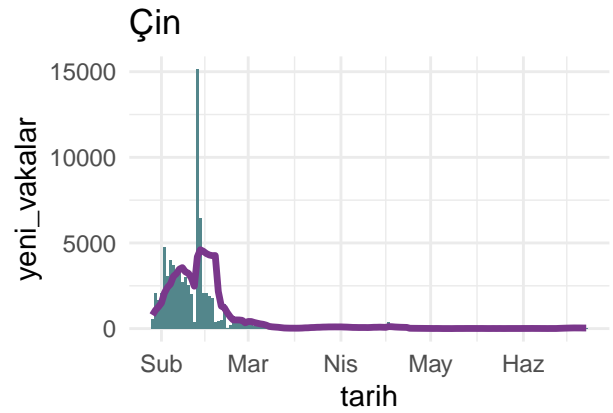
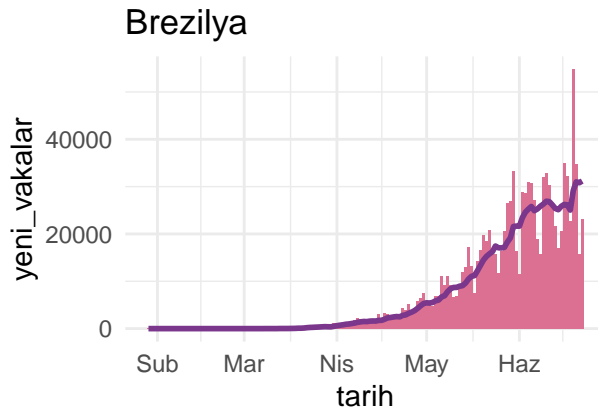
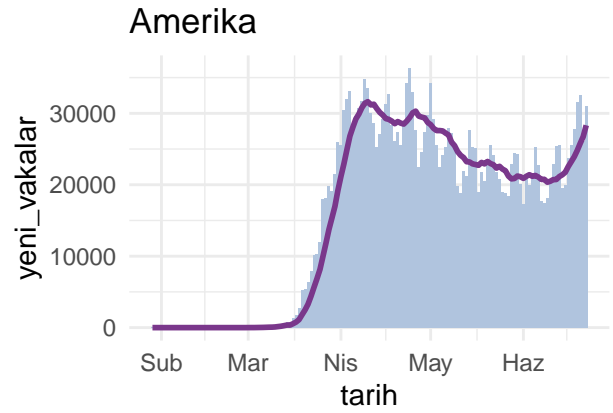
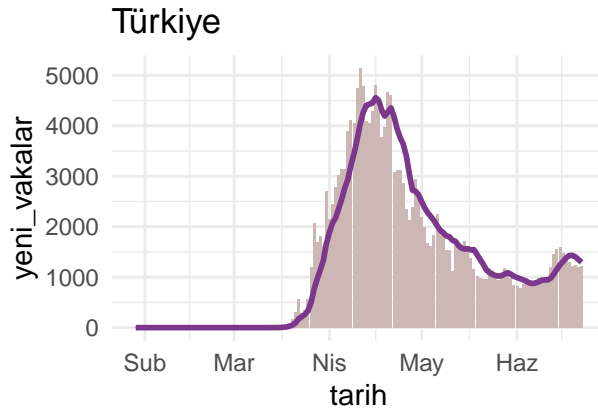
```

```

filter(!is.na(yeni_vakalar), !is.na(ave_new_cases)) %>%
  ggplot(aes(x = tarih)) +
  geom_bar(aes(y = yeni_vakalar), stat = "identity", fill = "cadetblue4") +
  geom_line(aes(y = ave_new_cases), color = "mediumorchid4", size=1.2) +
  labs(title="Çin") +
  theme_minimal()

grid.arrange(g,g1,g2,g3,ncol=2)

```



11 Onaylanan Vaka Sayısı En Yüksek İlk 15 Ülke

Aşağıda son güne ait covid-19 pandemisinin ülkelere göre verileri bulunmaktadır. 23 Haziran 2020 tarihli Türkiye verisine göre bin iki yüz altmış sekiz (1268) yeni vaka, 22312 aktif olgu yirmi yedi (27) vefat eden vardır.

##	ulke	tarih	olgu_sayi	yeni.olgu	aktif.olgu	iyilesen_sayi
## 1	Afghanistan	2020-06-23	29481	324	19603	9260
## 2	Albania	2020-06-23	2047	52	807	1195
## 3	Algeria	2020-06-23	12076	156	2541	8674
## 4	Andorra	2020-06-23	855	0	6	797
## 5	Angola	2020-06-23	189	3	102	77
## 6	Antigua and Barbuda	2020-06-23	26	0	1	22
##	vefat_sayi	yeni.vefat	olum.orani	ranking		
## 1	618	20	2.1	42		
## 2	45	1	2.2	103		
## 3	861	9	7.1	64		
## 4	52	0	6.1	133		
## 5	10	0	5.3	160		
## 6	3	0	11.5	175		

Dünyada vakaların en çok görüldüğü ilk 15 ülke aşağıdaki gibidir. Türkiye’de bu ilk 15 ülkenin içinde yer almaktadır.

En çok onaylanan Vakaların üçte biri (1/3) Avrupa’dan dır .Yaklaşık üçte birini de Güney Amerika ülkeri oluşturmaktadır.

Güney Amerika ülkeleri şuan kış dönemini yaşadıkları için salgının etkisini daha büyük ölçüde yaşamaktadırlar.

İlk etapta “toplumsal bağışıklık kazanılması amacıyla virüsün kontrollü yayılması” stratejisini izleyen İngiliz hükümeti artan ölüm oranlarıyla bu stratejisinden vazgeçti ancak alınan karantina önlemleri geç başlamasından dolayı salgın hala etkisini İngilterede göstermektedir.

```
k <- 15
```

```
top.countries <- data.latest.all %>% filter(ranking <= k + 1) %>%  
arrange(ranking) %>% pull(ulke) %>% as.character()  
top.countries %>% setdiff('dunya') %>% print()
```

```
## [1] "US" "Brazil" "Russia" "India"  
## [5] "United Kingdom" "Peru" "Chile" "Spain"  
## [9] "Italy" "Iran" "France" "Germany"  
## [13] "Mexico" "Turkey" "Pakistan"
```

Aşağıda Vaka Sayısı en çok olan ilk 15 ülkenin toplam ölüm oranı ,toplam vefat edenlerin sayısı ,toplam iyileşenlerin sayısı,toplam aktif olgu sayıları gösterilmiştir.

Dünyada Toplam Vakaların sayısı 9 milyonu geçmiştir.Onalayanan vakaların iki milyonu geçkini Amerikadandır.1 milyonu geçen vaka sayısı ile Breziya ikinci pandeminin en çok yaşadığı ülkedir.

Bu tabloda Ölüm oranı en yüksek olan 1. ülke %14.5 ile İtalya 2.ülke %14.0 ile İngilteredir.

Tabloya göre,Ölüm oranı en düşük olan ülke %1.4 ile Rusya, Ölüm oranı en düşük olan ülke 2. ülke %1.8 ile Türkiyedir.

```
data.latest <- data.latest.all %>% filter(!is.na(ulke)) %>%  
mutate(ulke=ifelse(ranking <= k + 1, as.character(ulke), 'Others')) %>%  
mutate(ulke=ulke %>% factor(levels=c(top.countries, 'Others')))  
data.latest %<>% group_by(ulke) %>%  
summarise(olgu_sayi=sum(olgu_sayi),
```

```

      yeni.olgu=sum(yeni.olgu),
      aktif.olgu=sum(aktif.olgu),
      iyilesen_sayi=sum(iyilesen_sayi),
      vefat_sayi=sum(vefat_sayi),
      yeni.vefat=sum(yeni.vefat)) %>%
mutate(olum.orani=(100 * vefat_sayi/olgu_sayi) %>% round(1))
data.latest %<>% select(c(ulke,olgu_sayi,vefat_sayi,olum.orani,
                        yeni.olgu,yeni.vefat,aktif.olgu ))
head(data.latest,15)

```

```

## # A tibble: 15 x 7
##   ulke          olgu_sayi vefat_sayi olum.orani yeni.olgu yeni.vefat aktif.olgu
##   <fct>          <int>      <int>      <dbl>      <dbl>      <dbl>      <int>
## 1 dünya          9263466    477584        5.2    164823      5413    4155491
## 2 US              2347022    121228        5.2     34720       826    1578246
## 3 Brazil          1145906     52645        4.6     39436      1374    465298
## 4 Russia           598878      8349        1.4       7413       153    234682
## 5 India            456183     14476        3.2     15968       465    183022
## 6 United Kingd~   307682     43011        14        921       280    263341
## 7 Peru            260810      8404        3.2     3363       181    103969
## 8 Chile           250767      4505        1.8     3804        3     35692
## 9 Spain           246752     28325       11.5       248        1     68051
## 10 Italy            238833     34675       14.5       113        18     19573
## 11 Iran            209970      9863        4.7     2445       121     30947
## 12 France          197804     29723        15        423        57     93086
## 13 Germany         192480      8914        4.6       712        15      7741
## 14 Mexico          191410     23377       12.2     6288       793     24387
## 15 Turkey          190165      5001        2.6     1268        27     22316

```

Aşağıdaki bar grafiğinde Vaka sayısı en çok olan ilk 15 ülkenin ; Toplam Onaylanan Vakaların sayısı, Total Ölüm Oranı (%) ,Yeni Onaylanan Vakaların sayısı,Günlük Ölümünün sayısı ,Güncel Covid-19 Vakaların sayısı görülmektedir.

- Vaka sayısı en çok olan ilk 15 ülkenin Toplam onaylanan bar grafiğine bakıldığında Amerika dünyadaki diğer ülkeler grubunun vaka sayılarında üstündedir.Şuan Salgında dünya geneli birinci ülke amerika ikinci ülke Brezilyadır.
- Vaka sayısı en çok olan ilk 15 ülkenin Total Ölüm Oranı Bar grafiğine göre Ölüm oranı en yüksek olan 1. ülke %14.5 ile İtalya 2.ülke %14.0 ile İngilteredir ve üçüncü ölüm oranı en yüksek ülke %12.2 ile Mexico dur.Tabloya göre,Ölüm oranı en düşük olan ülke %1.4 ile Rusya, Ölüm oranı en düşük olan ülke 2. ülke %1.8 ile Chile.
- Yeni Onaylanan Vakaların bar grafına baktığımızda diğer dünya ülkelerinde total sayı en fazladır.Tek bir ülke olarak grafikte en yüksek günlük vaka sayısı Brezilyaya aittir onu takip eden ikinci ülke Amerikadır.
- Günlük ölümler grafiği incelendiğinde en yüksek günlük ölüm sayısına sahip ilk üç ülke sırasıyla Brezilya(1374) ,Amerika(826) ve Meksika'dır(793).Dünyadaki diğer ülkelerin günlük vefat edenlerin sayısı 1039dur.

-Aktif Vakalar grafiği incelendiğinde Amerika en yüksek aktif covid-19 vakasına sahip ülkedir ,aktif vakaların sayısı yaklaşık 1.5 milyondur.

-İyileşenlerin bar grafiğine bakıldığında grafikte olan diğer dünya ülkelerindeki iyileşen sayı en fazla sayıdadır.Bunun nedeni Çin ,Güney Kore,Japonya gibi ülkelerin salgını kontrol altına alıp hatta durdurmasıdır daha önceki vakalar iyileşmiştir.

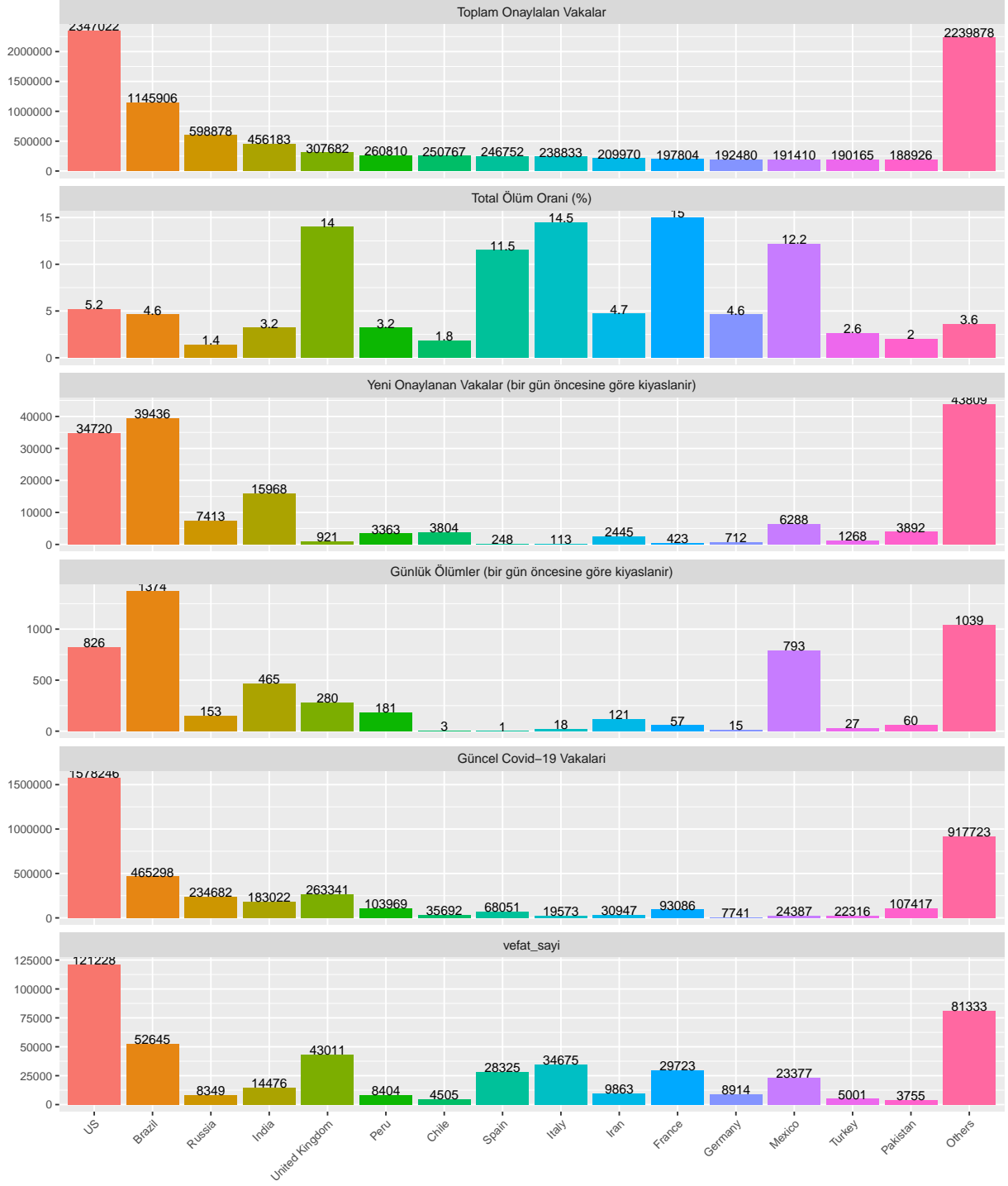
- Toplamda Vefat edenlerin bar grafiği incelendiğinde en yüksek sayıda ölüm Amerika ülkesine aittir yaklaşık 121 bin kişi pandeminin başlangıcından buyana hayatlarını kaybetmişlerdir.Grafiklerde göze

çarpan ülkeler (sayıca fazla ölüme sahip): Amerika,Brezilya,İspanya ,İngiltere ve İtalyadır.

```
## alan çizgileri için geniş den uzun formata dönüştürme
data.latest.long <- data.latest %>% filter(ulke!='dunya') %>%
  gather(key=type, value=count, -ulke)
## uygun metinlerle ve istenen sırada göstermek için faktör düzeylerini #ayarlama
data.latest.long1 <- data.latest.long %<>% mutate(type=recode_factor(type,
  olgu_sayi='Toplam Onaylanan Vakalar',
  olum_orani='Total Ölüm Oranı (%)',
  yeni_olgu='Yeni Onaylanan Vakalar (bir gün öncesine göre kıyaslanır)',
  yeni_vefat='Günlük Ölümler (bir gün öncesine göre kıyaslanır)',
  aktif_olgu='Güncel Covid-19 Vakaları',
))

## bar chart
data.latest.long1 %>% ggplot(aes(x=ulke, y=count, fill=ulke, group=ulke)) +
  geom_bar(stat='identity') +
  geom_text(aes(label=count, y=count), size=3, vjust=0) +
  xlab('') + ylab('') +
  labs(title=paste0('En çok Vakaların Görüldüğü 15 Ülke - ', max.date.txt)) +
  scale_fill_discrete(name='Ulke', labels=aes(count)) +
  theme(legend.title=element_blank(),
    legend.position='none',
    plot.title=element_text(size=8),
    axis.text=element_text(size=8),
    axis.text.x=element_text(angle=45, hjust=1)) +
  facet_wrap(~type, ncol=1, scales='free_y')
```


En çok Vakaların Görüldüğü 15 Ülke – 23 Haz 2020 UTC



11.1 Vefat Eden, İyileşen ve Aktif Vakalar

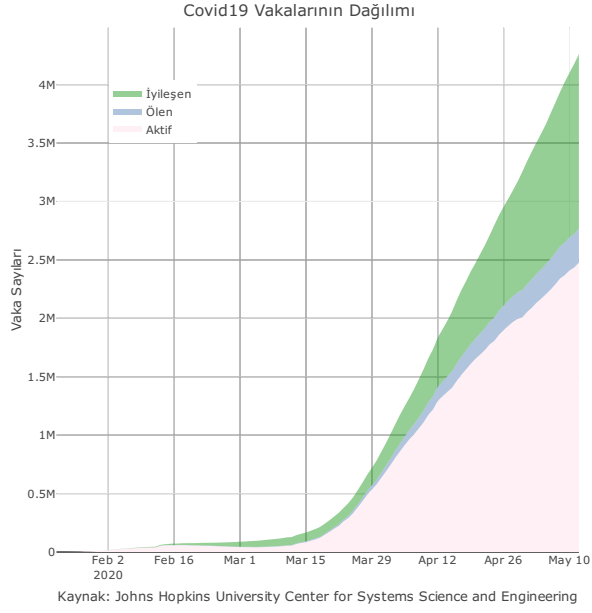
Aşağıdaki analizde Ölenlerin, İyileşenlerin ve Aktif Vakaların interaktif alan grafiği gösterilmiştir.

Aktif vakaların sayısı İyileşen ve ölen vakalara göre çok fazladır. Bu bize pandeminin hala etkili bir şekilde sürdüğünü göstermektedir.

12 Mayıs'ta aktif vakaların dünya genelinde sayıları 2.476.391 dir. 12 Mayıs'ta dünya genelinde İyileşen vakaların sayısı 1.493.414 dir ve vefat edenlerin sayısı toplamda 286.333 tür.

```
library(plotly)

coronavirus %>%
  group_by(type, date) %>%
  summarise(total_cases = sum(cases)) %>%
  pivot_wider(names_from = type, values_from = total_cases) %>%
  arrange(date) %>%
  mutate(aktif = confirmed - death - recovered) %>%
  mutate(aktif_total = cumsum(aktif),
         recovered_total = cumsum(recovered),
         death_total = cumsum(death)) %>%
  plot_ly(x = ~ date,
          y = ~ aktif_total,
          name = 'Aktif ',
          fillcolor = '#FFF0F5',
          type = 'scatter',
          mode = 'none',
          stackgroup = 'one') %>%
  add_trace(y = ~ death_total,
            name = "Ölen ",
            fillcolor = '#B0C4DE') %>%
  add_trace(y = ~recovered_total,
            name = 'İyileşen',
            fillcolor = 'mistyrose3') %>%
  layout(title = "Covid19 Vakalarının Dağılımı",
         legend = list(x = 0.1, y = 0.9),
         yaxis = list(title = "Vaka Sayıları"),
         xaxis = list(title = "Kaynak: Johns Hopkins University Center for Systems Science and Engineer
```



12 SONUC

Analizi yapılan veriler John Hopkins Üniversitesi tarafından sağlanan covid19 pandemisine ait ülkelerin verileri <https://github.com/CSSEGISandData/COVID-19> sitesinden indirilmiştir. Veri seti Eyalet, ülke, enlem, boylam ve tarihler olarak alınmıştır. İleri görselleştirme için 'tidyverse' ve 'dplyr' paketleri ile veriler düzenlenmiştir. 'ggplot2' paketi ile düzenlenen verilerin görselleştirilmesi yapılmıştır. 'dplyr' ve 'tidyverse' paketinde olmayan işlevler için farklı kütüphaneler kullanılmıştır. 'coronavirus' ve 'tidycovid19' paketi kullanılarak interactive görseller eklenmiştir.

Analiz bize ülkelerin covid-19 süreçleri hakkında bilgi vermektedir, covid-19 pandemisini daha iyi anlamamızı sağlamaktadır. Karantina uygulanan ülkelerin, fiblasyon yöntemi ile salgını kontrol altına çalışan ülkelerin iyileşen sayılarında ki artış vefat sayılarını geçmiş ve günlük vaka sayıları azaldığı görülmüştür. 1 Haziran'da bir çok Avrupa ve dünya ülkeleri normalleşme sürecine geçmeye başlamıştır. Grafiklerdeki Haziran sonrası vaka artışlarının bir nedenide budur.

13 KAYNAKÇA

- 1) BBC (2020). Koronavirüs nedir: Covid-19'a karşı hangi önlemler alınmalı, virüsün özellikleri neler?
URL <https://www.bbc.com/turkce/haberler-dunya-51177538>
- 2) WorldMeter (2020) Covid-19 Koronavirüs Salgını URL <https://www.worldometers.info/coronavirus/>
- 3) Shiny Tutorial URL <https://shiny.rstudio.com/tutorial/written-tutorial/lesson1/>