

Değişkenler Arındaki İlişkilerin ve Zaman Serilerinin Görselleştirilmesi

Rümeysa Kurt

30.11.2022

ÖZET

Bu raporda 3 farklı veri seti incelenmiştir. İlk veri seti Amazon'un 2009-2019 yılları arasında en çok satan 50 kitabını içermektedir. İlk kısımda kitap fiyatları ve okuyucu yorumları arasındaki ilişki görselleştirilmiştir, Diğer kısımda ise kitap fiyatları ve okuyucu yorumları arasındaki ilişki kitap türüne göre görselleştirilmiştir. İkinci veri seti Game of Thrones dizisini içermektedir. İlk kısımda IMDB oy sayıları ve bölüm süreleri arasındaki ilişki görselleştirilmiş olup diğer kısımda ise IMDB oy sayıları ve bölüm süreleri arasındaki ilişki yardımcı oyuncunun (Star2) Nikolaj Coster-Waldau olup olmaması durumuna göre görselleştirilmiştir. Üçüncü veri seti döviz kurlarıdır, priceR paketi kullanılarak ilk kısımda Türk lirasının, Euro, Amerikan doları ve Pound karşılıklarının değişimi görselleştirilmiştir. İkinci kısımda ise seçtiğiniz 3 farklı döviz kurunun Amerikan doları karşılıkları araştırılıp görselleştirilmiştir. Görselleştirme araçlarını kullanmak ve çeşitli işlemleri yapmak için `install.packages("ggplot2")`, `install.packages("dplyr")`, `install.packages("tidyr")` paketleri indirilmiştir. Son olarak oluşan grafikler yorumlanmıştır.

Gerekli Paketlerin İndirilmesi

```
install.packages("dplyr")
install.packages("tidyr")
install.packages("ggplot2")
install.packages("priceR")
library(dplyr)
library(tidyr)
library(ggplot2)
library(priceR)
```

1.2009-2019 Yılları Arasında Amazon'da En Çok Satan 50 Kitabı İçeren Veri Setinin İncelenip İlgili Konular Dahilinde Görselleştirilmesi

Veri seti 2009'dan 2019'a kadar Amazon'un en çok satan 50 kitabını içermektedir. Kitaplar kurgu ve kurgu olmayan olarak iki sınıfa ayrılmıştır. Veri setinde 7 değişken ve 550 gözlem olduğu görülmüştür. Veri setinde yer alan değişkenler şunlardır: Kitap adı, Yazar, Kullanıcı reytingi, Yorumlar, Fiyat, Yıl, Tür.

```
library(readr)
bestsellers_with_categories <- read_csv("bestsellers with categories.csv")
```

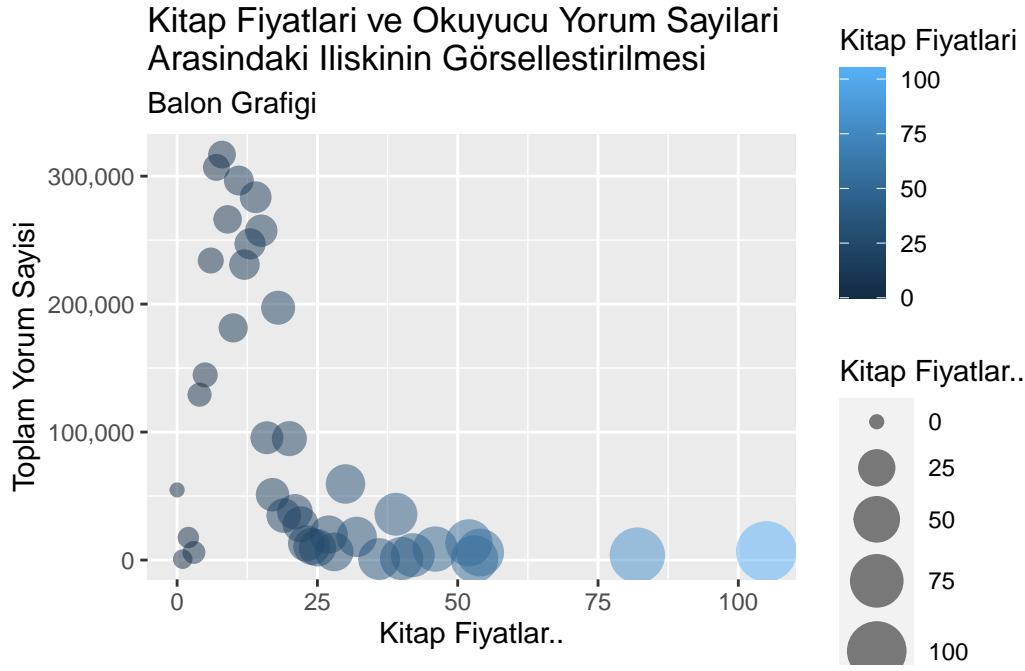
1.1 Kitap fiyatları ve okuyucu yorum sayıları arasındaki ilişkiyi veri görselleştirme yöntemleriyle araştırınız. Kitap fiyatlarındaki artışın okuyucu yorum sayıları üzerinde nasıl bir etkisi vardır?

```
okuyucuyorumsayısı <- bestsellers_with_categories %>%
group_by(Price,Reviews) %>%
summarise(sayı = n())
```

```
toplamyorumsayısı <- okuyucuyorumsayısı %>%
group_by(Price) %>%
summarise(Toplam_yorum_sayısı = sum(Reviews))
toplamyorumsayısı
```

```
# A tibble: 40 x 2
  Price Toplam_yorum_sayısı
  <dbl>          <dbl>
1     0          54787
2     1           720
3     2         17609
4     3          5983
5     4        129219
6     5        144662
7     6        234065
8     7        306840
9     8        316865
10    9        266221
# ... with 30 more rows
```

```
ggplot(toplamyorumsayısı, aes(x=Price, y=Toplam_yorum_sayısı, size=Price, color=Price)) +
  scale_y_continuous(labels = scales::comma)+
  geom_point(alpha=0.5) +
  scale_size(range=c(2, 10), name='Kitap Fiyatları') +
  labs(x="Kitap Fiyatları", y="Toplam Yorum Sayısı",
  title = "Kitap Fiyatları ve Okuyucu Yorum Sayıları
  Arasındaki İlişkinin Görselleştirilmesi",
  color = "Kitap Fiyatları",
  subtitle = "Balon Grafiği")
```



Yorum

Kitap fiyatları ve okuyucu yorum sayıları arasındaki ilişki balon grafiği ile görselleştirilmiş olup x ekseninde kitap fiyatları, y ekseninde ise toplam yorum sayıları yer almaktadır. Balonların büyüklüğü ve rengin açılması kitap fiyatlarına bağlı değişmektedir. Fiyatlar arttıkça balonun genişliği artmaktadır ve mavi rengin tonu açılmaktadır. Grafiğe bakıldığında fiyatlar arttıkça okuyucu yorumlarının düştüğünü görüyoruz. En çok yorum alan kitapların fiyatlarının 0 ile 12 arasında olduğunu grafiğe bakarak söyleyebiliriz.

1.2 Kitap fiyatları ve okuyucu yorum sayıları arasındaki ilişkiyi, kitap türüne göre veri görselleştirme yöntemleriyle araştırınız. Kitap fiyatlarındaki artışın okuyucu yorum sayıları üzerinde etkisi kitap türüne göre bir değişiklik göstermekte midir?

```
kitap <- bestsellers_with_categories %>%  
  group_by(Price,Reviews,Genre) %>%  
  summarise(sayı=n())
```

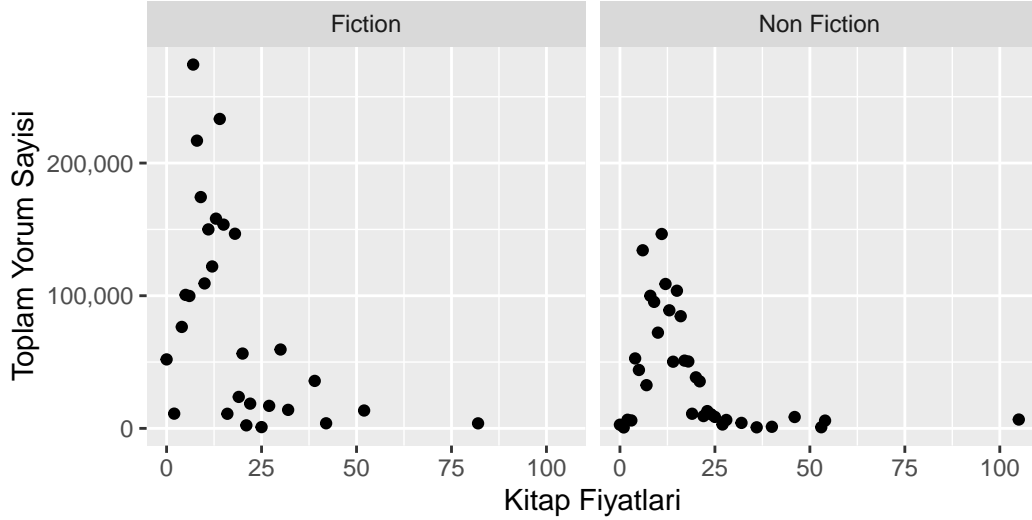
```
kitap_2 <- kitap %>%  
  group_by(Price, Genre) %>%  
  summarise(Toplamyorumsayısı=sum(Reviews))  
kitap_2
```

```
# A tibble: 63 x 3  
# Groups:   Price [40]  
  Price Genre      Toplamyorumsayısı  
  <dbl> <chr>          <dbl>  
1     0 Fiction          52013  
2     0 Non Fiction       2774  
3     1 Non Fiction        720  
4     2 Fiction          11107  
5     2 Non Fiction        6502  
6     3 Non Fiction        5983  
7     4 Fiction          76491  
8     4 Non Fiction       52728  
9     5 Fiction         100628  
10    5 Non Fiction       44034  
# ... with 53 more rows
```

```
ggplot(kitap_2, aes(x = Price, y = Toplamyorumsayısı)) +  
  geom_point() +  
  labs(title = "Kitap Fiyatları ve Okuyucu Yorum Sayıları Arasındaki  
İliskinin Kitap Türüne Göre Görselleştirilmesi",  
        x = "Kitap Fiyatları",  
        y = "Toplam Yorum Sayısı",  
        subtitle = "Scilim Grafiği") +  
  scale_y_continuous(labels = scales::comma)+  
  facet_grid(~Genre)
```

Kitap Fiyatları ve Okuyucu Yorum Sayıları Arasındaki İlişkinin Kitap Türüne Göre Görselleştirilmesi

Saçılım Grafiği



Yorum

Kitap fiyatları ve okuyucu yorum sayıları arasındaki ilişki kitap türlerine göre saçılım grafikleri kullanılarak görselleştirilmiştir. İki farklı kitap türü olduğundan daha iyi yorum yapabilmek için yan yana saçılım grafiği tercih edilmiş olup x eksenine kitap fiyatları, y eksenine toplam yorum sayısı konumlandırılmıştır. Fiyatı 0-50 arasındaki kitapların okuyucu yorumlarına baktığımızda kurgusal kitapların kurgusal olmayanlara göre daha fazla yorum aldığını söyleyebiliriz. Her iki kitap türü içinde bakmak gerekirse kitap fiyatları arttıkça yorum sayıları düşmektedir yorumunu yapabiliriz.

2.Game of Thrones Dizisi İle İlgili Verileri İçeren Veri Setinin İncelenip İlgili Konular Dahilinde Görselleştirilmesi

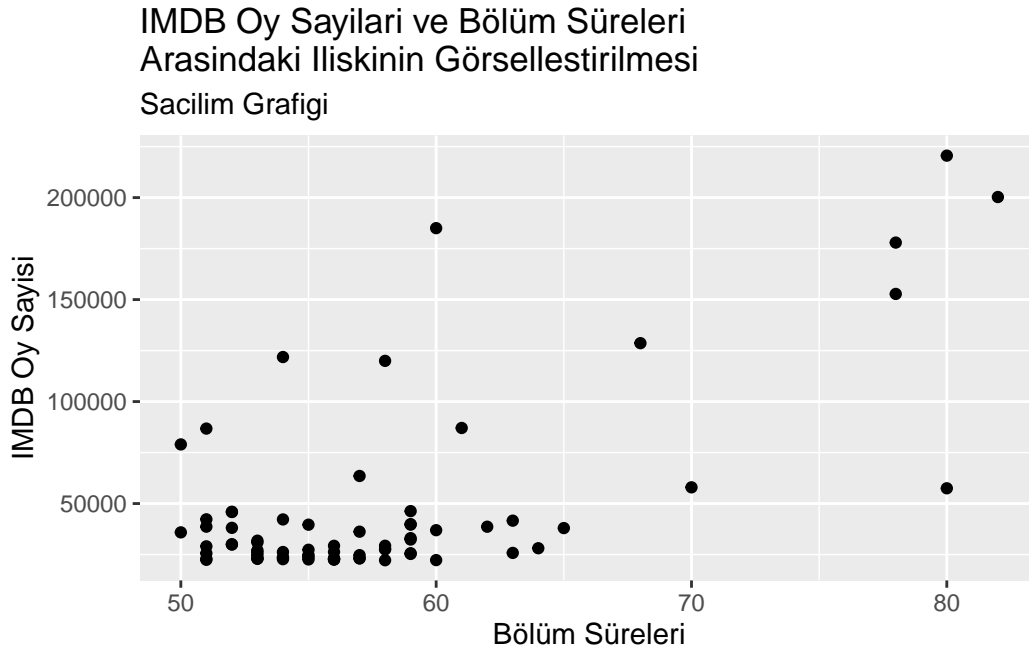
Bu veri seti Game of Thrones dizi ile ilgili verileri içermektedir. Game of Thrones veri setinde 18 değişken ve 73 gözlem yer almaktadır. Veri setinde bulunan değişkenler şunlardır:Sezon, Bölüm, Başlık, Yayın Tarihi, Rating, Oy, Özet, Yazar1, Yazar2, Başrol1, Başrol2, Başrol3, Kullanıcı yorumları, Eleştirmen yorumları, US yorumları, Süre, Yönetmen, Bütçe tahmini.

```
library(readr)
GOT_episodes_v4 <- read_csv("GOT_episodes_v4.csv")
```

2.1 IMDB oy sayıları ve bölüm süreleri arasındaki ilişkiyi veri görselleştirme yöntemleriyle araştırınız. Bölüm süresinin artması oy sayıları üzerinde nasıl bir etki göstermiştir?

```
got_veri <- GOT_episodes_v4 %>%
  group_by(Votes, Duration) %>%
  summarise(sayı = n())

ggplot(got_veri, aes(x = Duration, y = Votes)) +
  geom_point() +
  labs(x="Bölüm Süreleri", y="IMDB Oy Sayısı",
       title = "IMDB Oy Sayıları ve Bölüm Süreleri
Arasındaki İlişkinin Görselleştirilmesi",
       subtitle = "Saçılım Grafiği")
```



2.1 IMDB oy sayıları ve bölüm süreleri arasındaki ilişkiyi, yardımcı oyuncunun (Star2) Nikolaj Coster-Waldau olup olmaması durumuna göre veri görselleştirme yöntemleriyle araştırınız. Nikolaj Coster-Waldau'nun yardımcı oyuncu olarak oynayıp oynamamasının, oy sayıları ve bölüm süreleri arasındaki ilişki üzerinde bir etkisi var mıdır?

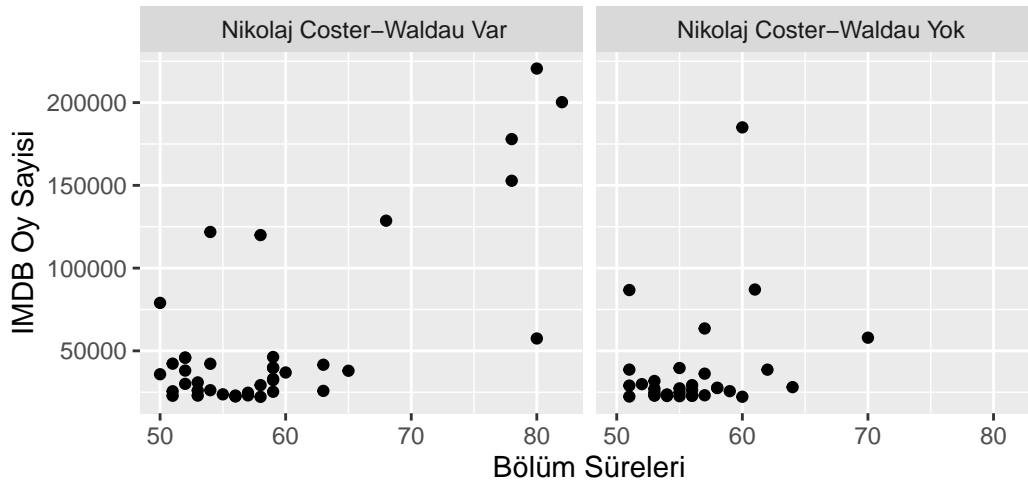
```
got_veri_oyuncu <- GOT_episodes_v4 %>%  
group_by(Star_2,Duration,Votes) %>%  
summarise(sayı = n())
```

```
oyuncu <- got_veri_oyuncu %>%  
mutate(yardımcı_oyuncu = ifelse(Star_2 == "Nikolaj Coster-Waldau",  
                                "Nikolaj Coster-Waldau Var",  
                                "Nikolaj Coster-Waldau Yok"))
```

```
ggplot(oyuncu, aes(x = Duration, y = Votes)) +  
  geom_point() +  
  labs(title = "IMDB Oy Sayıları ve Bölüm Süreleri Arasındaki İlişkinin  
Nikolaj Coster-Waldau'nun Olup Olmaması Durumuna  
Göre Görselleştirilmesi",  
x="Bölüm Süreleri", y="IMDB Oy Sayısı",  
subtitle = "Sacilim Grafiği") +  
  facet_grid(~yardımcı_oyuncu)
```

IMDB Oy Sayıları ve Bölüm Süreleri Arasındaki İlişkinin
Nikolaj Coster-Waldau'nun Olup Olmaması Durumuna
Göre Görselleştirilmesi

Sacilim Grafiği



Yorum

IMDB oy sayıları ve bölüm süreleri arasındaki ilişkiyi, yardımcı oyuncunun (Star2) Nikolaï Coster-Waldau olup olmaması durumuna göre saçılım grafikleri kullanılarak görselleştirilmiştir. Oyuncunun olup olmama durumuna göre görselleştirilmesi istendiği için yan yana serpilme diyagramı kullanılmış olup x eksenine bölüm süreleri, y eksenine IMDB oy sayıları konumlandırılmıştır. Nikolaï Coster-Waldau'nun yer aldığı bölümlerde bölüm süresi arttığında alınan oy sayılarında arttığı görülmektedir. Nikolaï Coster-Waldau'nun olmadığı bölümlerde ise en yüksek oy sayısına bölüm süresi 60 dakika olduğunda rastlanmıştır.

3.Döviz Kurları

priceR paketi kullanılarak elde edilmiştir. Ülkelerin geçmiş döviz kurlarını içerir.

3.1 Son 10 yılda, Türk lirasının, Euro, Amerikan doları ve Pound karşılıklarının değişimini veri görselleştirme yöntemleriyle araştırınız.

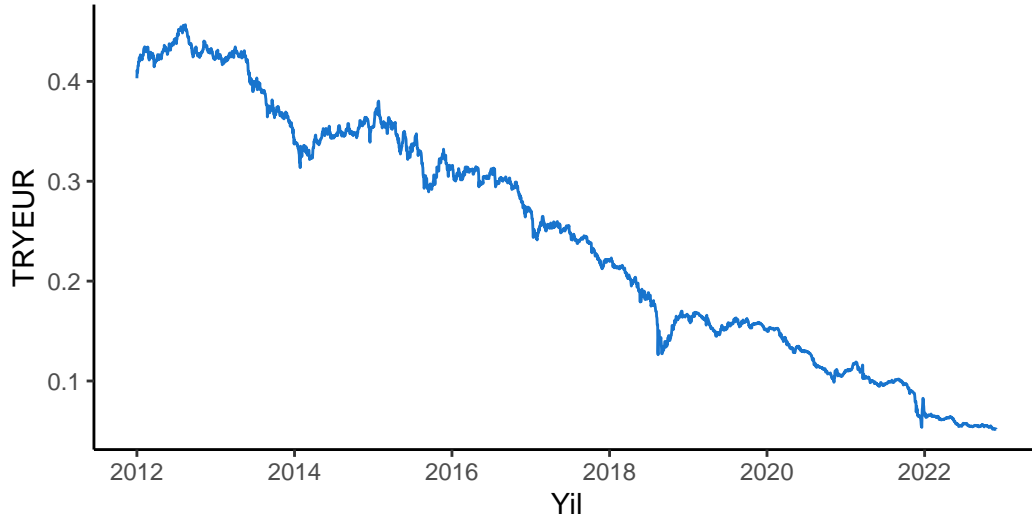
TRYEUR Grafiği

```
w <- historical_exchange_rates(from = "TRY", to = "EUR",
start_date = "2012-01-01", end_date = "2022-11-30")

ggplot(w, aes(x = date, y = w$one_TRY_equivalent_to_x_EUR)) +
  geom_line(color = "dodgerblue3") +
  labs(x = "Yıl",
       y = "TRYEUR",
       subtitle = "Çizgi Grafiği",
       title = "Son 10 Yılda Türk Lirasının Euro
Karsısındaki Değişiminin Grafiği") +
  theme_classic()
```


Son 10 Yilda Türk Lirasinin Euro Karsisindaki Degisiminin Grafigi

Çizgi Grafigi



Yorum

Son 10 yılda Türk lirasının Euro karşısındaki değişimini görselleştirmek için çizgi grafiği kullanılmıştır. 1 Türk lirasının Euro karşısında en değerli olduğu yıl ise 2013 yılıdır. Türk lirası yıllar içinde Euro karşısındaki değerini kaybetmiştir. Grafikte yıllar içinde dalgalanmalar mevcuttur. Türk lirası 2018 yılının ortalarına doğru o zamana kadar geçen sürede Euro karşısında en büyük kaybı yaşamıştır. 2022 yılından itibaren ise Türk lirasında ciddi bir düşüş gözlenmiştir.

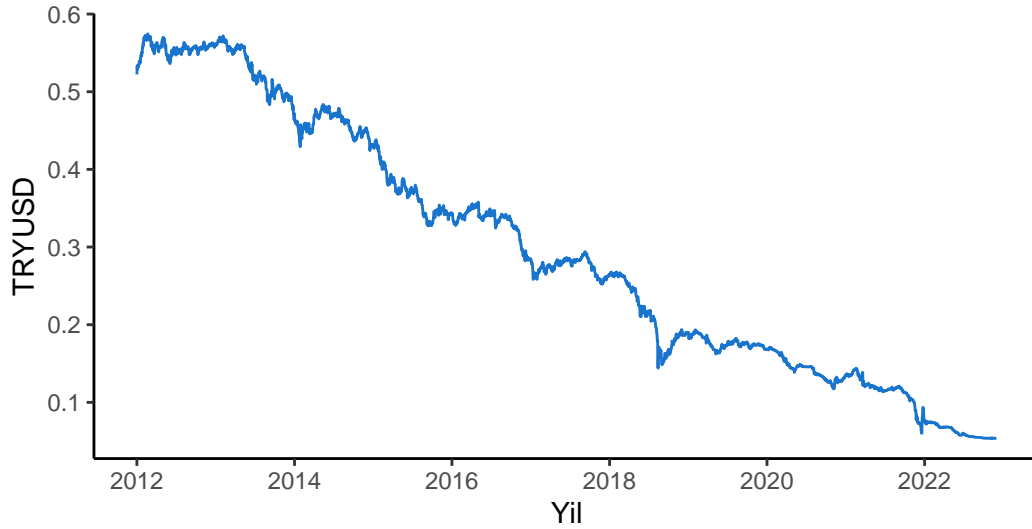
TRYUSD Grafiği

```
v <- historical_exchange_rates(from = "TRY", to = "USD",
start_date = "2012-01-01", end_date = "2022-11-30")

ggplot(v, aes(x = date, y = v$one_TRY_equivalent_to_x_USD)) +
  geom_line(color = "dodgerblue3") +
  labs(x = "Yil",
       y = "TRYUSD",
       subtitle = "Çizgi Grafiği",
       title = "Son 10 Yilda Türk Lirasinin Amerikan Dolari
Karsisindaki Degisiminin Grafigi") +
  theme_classic()
```

Son 10 Yilda Türk Lirasinin Amerikan Doları Karşısındaki Değişiminin Grafiği

Çizgi Grafiği



Yorum

Son 10 yılda Türk lirasının Amerikan doları karşısındaki değişimini görselleştirmek için çizgi grafiği kullanılmıştır. 1 Türk lirasının Amerikan doları karşısında en değerli olduğu yıl ise 2012 yılıdır. Grafiğe bakıldığında son 10 yıl içinde Türk lirasının Amerikan doları karşısında değer kaybettiği görülmektedir. Grafikte dalgalanmalar mevcuttur. 2022 yılının başlarında bir artık gözlenmiş olup ondan sonra günümüze kadar olan sürede ise Türk lirası Amerikan doları karşısında en düşük seviyeleri görmüştür.

TRYGBP Grafiği

```
z <- historical_exchange_rates(from = "TRY", to = "GBP",
start_date = "2012-01-01", end_date = "2022-11-30")

ggplot(v, aes(x = date, y = z$one_TRY_equivalent_to_x_GBP)) +
  geom_line(color = "dodgerblue3") +
  labs(x = "Yıl",
       y = "TRYGBP",
       subtitle = "Çizgi Grafiği",
       title = "Son 10 Yilda Türk Lirasinin Pound
Karsisindaki Degisiminin Grafiği") +
  theme_classic()
```

Son 10 Yilda Türk Lirasinin Pound Karsisindaki Degisiminin Grafigi

Çizgi Grafigi



Yorum

Son 10 yılda Türk lirasının pound karşısındaki değişimini görselleştirmek için çizgi grafiği kullanılmıştır. Grafikte dalgalanmalar mevcuttur. Geçen 10 yıl içinde Türk lirası Pound karşısında değer kaybetmiştir. 1 Türk lirasının pound karşısında en değerli olduğu yıl 2013 iken en değersiz olduğu yıl ise 2022'dir

3.2 Seçtiğiniz 3 farklı döviz kurunun Amerikan doları karşılıklarını veri görselleştirme yöntemleriyle araştırınız.

MXNUSD Grafiği

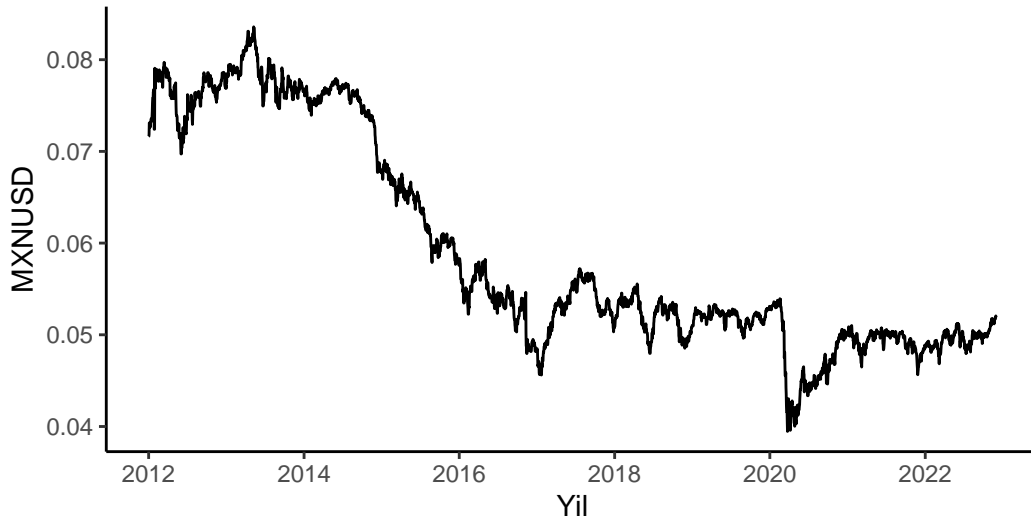
```
a <- historical_exchange_rates(from = "MXN", to = "USD",
start_date = "2012-01-01", end_date = "2022-11-30")

ggplot(a, aes(x = date, y = a$one_MXN_equivalent_to_x_USD)) +
  geom_line() +
  labs(x = "Yil",
       y = "MXNUSD",
       subtitle = "Çizgi Grafiği",
       title = "Son 10 Yilda Meksika Pesosu'nun Amerikan Dolari
```

```
Karsisindaki Degisiminin Grafigi") +  
  theme_classic()
```

Son 10 Yilda Meksika Pesosu'nun Amerikan Dolari Karsisindaki Degisiminin Grafigi

Çizgi Grafigi



Yorum

Son 10 yılda Meksika Pesosu'nun Amerikan doları karşısındaki değişimini görselleştirmek için çizgi grafiği kullanılmıştır. Grafiğe bakıldığında dalgalanmalar mevcuttur. 1 Meksika Pesosu'nun Amerikan doları karşısında en değerli olduğu yıl 2013 olurken en çok değer kaybettiği yıl ise 2020 olmuştur. 2020 yılı içerisinde artış azalışlar gözlenmektedir ve Meksika Pesosu'nun son aylarda değer kazandığı görülmektedir.

JPYUSD Grafiği

```
b <- historical_exchange_rates(from = "JPY", to = "USD",  
  start_date = "2012-01-01", end_date = "2022-11-30")  
  
ggplot(a, aes(x = date, y = b$one_JPY_equivalent_to_x_USD)) +  
  geom_line() +  
  labs(x = "Yil",  
    y = "JPYUSD",  
    subtitle = "Çizgi Grafigi",
```

```

    title = "Son 10 Yilda Japon Yeni'nin Amerikan Dolari
Karsisindaki Degisiminin Grafigi") +
    theme_classic()

```



Yorum

Son 10 yılda Japon Yeni'nin Amerikan doları karşısındaki değişimini görselleştirmek için çizgi grafiği kullanılmıştır. Grafikte dalgalanmalar olduğu görülmektedir. 1 Japon Yeni'nin Amerikan doları karşısında en değerli olduğu yıl 2012 olurken en çok değer kaybettiği yıl ise 2022 olmuştur.

SEKUSD Grafiği

```

c <- historical_exchange_rates(from = "SEK", to = "USD",
start_date = "2012-01-01", end_date = "2022-11-30")

ggplot(c, aes(x = date, y = c$one_SEK_equivalent_to_x_USD)) +
  geom_line() +
  labs(x = "Yil",
       y = "SEKUSD",
       subtitle = "Çizgi Grafigi",

```

```
title = "Son 10 Yilda İsveç Kronu'nun Amerikan Dolari  
Karsisindaki Degisiminin Grafigi") +  
theme_classic()
```



Yorum

Son 10 yılda İsveç Kronu'nun Amerikan doları karşısındaki değişimini görselleştirmek için çizgi grafiği kullanılmıştır. Grafikte dalgalanmalar olduğu görülmektedir. 1 İsveç Kronu'nun Amerikan doları karşısında en değerli olduğu yıl 2013 olurken en çok değer kaybettiği yıl ise 2022 olmuştur. 2014-2015 yılları arasında İsveç Kronu Amerikan doları karşısında çok büyük bir düşüş yaşamış o yıllardan sonra bir daha eski değerini kazanamamıştır.