

# 3 veri setinin incelenmesi ve göselleştirilmesi

Ahmet Batuhan Özdoğan

3.12.2022

## Özet

Bu ödevde Kaggle sitesinden alınmış 3 farklı veri setinin görselleştirme ve yorumlama çalışmaları bulunmaktadır. Birinci veri seti En Çok Satan Kitaplar ,ikinci veri seti Döviz Kurları ve üçüncü veri seti ise Marvel filmleri olarak kullanılmıştır

## Gerekli Kütüphanlerin yüklenmesi

```
install.packages("ggplot2")
install.packages("quarto")
install.packages("tidyverse")
install.packages("dplyr")
install.packages("ggridges")
install.packages("readxl")
install.packages("priceR")
install.packages("ggpubr")
library(ggpubr)
library(priceR)
library(ggridges)
library(ggplot2)
library(tidyverse)
library(dplyr)
```

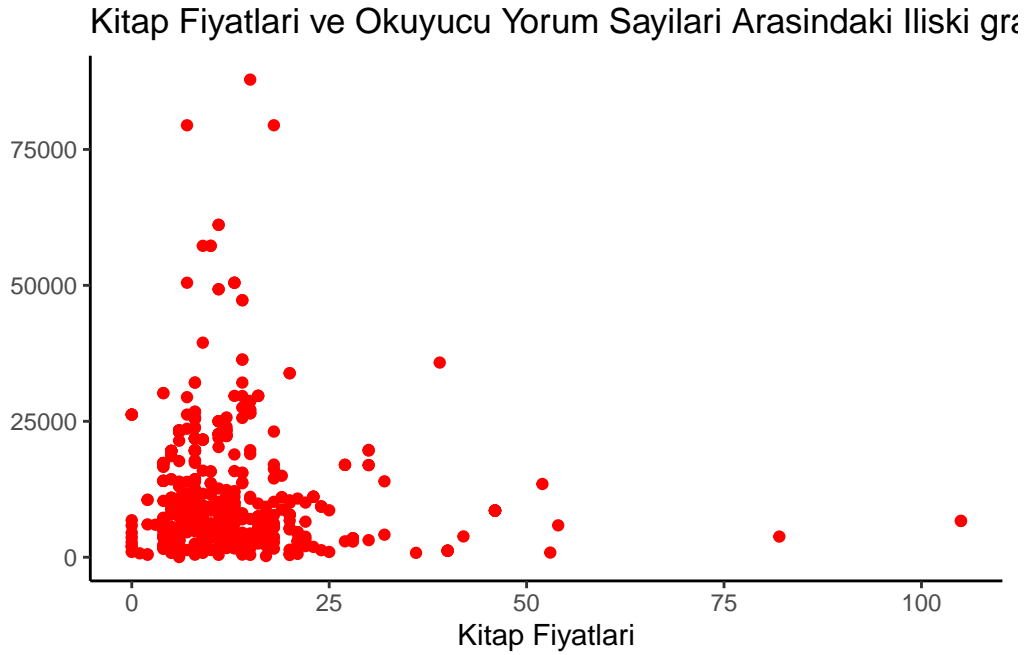
## 1. veri

### Veri Setinin Hakkında

Bu veri setimizde En çok satan kitapların incelemesi ve görselleştirilmesi bulunmaktadır

### Kitap fiyatları ve okuyucu yorum sayıları arasındaki ilişki

```
library(readr)
kitap <- read_csv("bestsellers with categories.csv")
kitap <- kitap%>%
add_column(tür = if_else(kitap$Genre == "Fiction", "Kurgusal", "Kurgusal Olmayan"))
ggplot(kitap, aes(Price, Reviews))+
geom_point(color = "red" )+
labs(x = "Kitap Fiyatlari", y = "" ,
title = "Kitap Fiyatlari ve Okuyucu Yorum Sayilari Arasindaki Iliski grafiği")+
theme_classic()
```

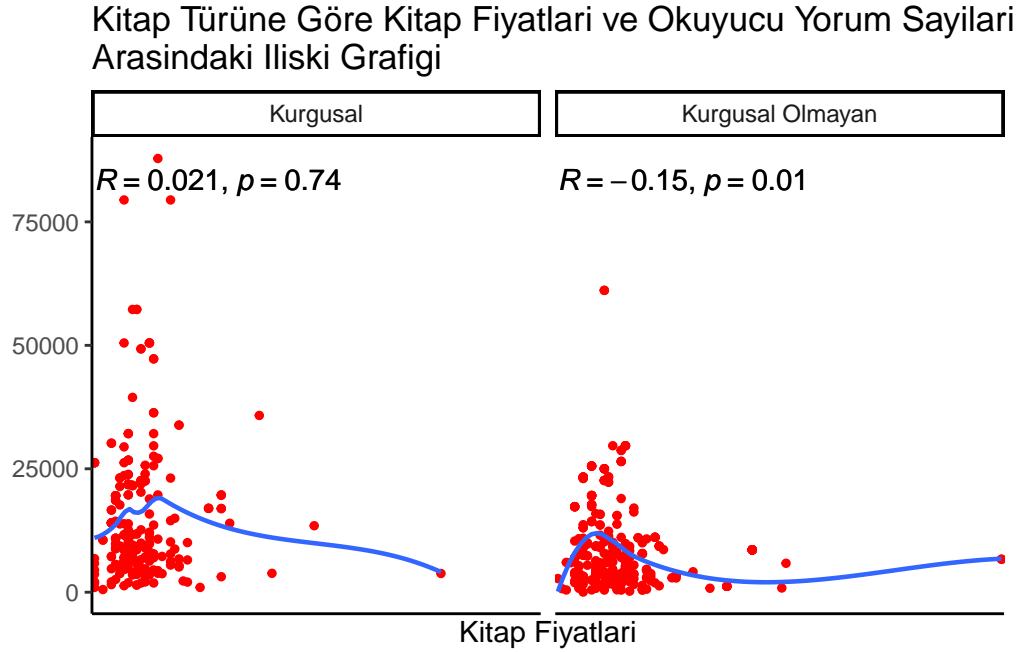


## Rapor

Grafiğe Baktığımızda kitap fiyatları arttıkça okuyucu yorum sayısının azaldığı kitap fiyatları azaldıkça okuyucu yorum sayısının çoğaldığı gözükmektedir

### Kitap Türüne Göre Kitap Fiyatları ve Okuyucu Yorum Sayıları Arasındaki İlişki

```
ggplot(kitap, aes(x=Price,y=Reviews))+  
geom_point(size = 1 , color = "red" )+  
geom_smooth(se = FALSE, size = 0.8)+  
ggpubr::stat_cor()+  
facet_wrap(~kitap$tür) +  
ggpubr::stat_cor() +  
labs(x = "Kitap Fiyatları", y = "",  
title = "Kitap Türüne Göre Kitap Fiyatları ve Okuyucu Yorum Sayıları  
Arasındaki İlişki Grafiği",  
fill = "Kitap Türü") +  
scale_x_discrete(labels = c("Kurgusal","Kurgusal olmayan")) +  
theme_classic()
```



## Rapor

Grafikteki sonuçlara göre, kurgusal olan ve kurgusal olmayan kitaplarımızda grafiğimize göre bariz bir biçimde fiyat arttıkça yorum sayısı düşmektedir yani yorum sayısında kitapların cinsine göre bir ayırım söz konusu değildir

## İkinci veri

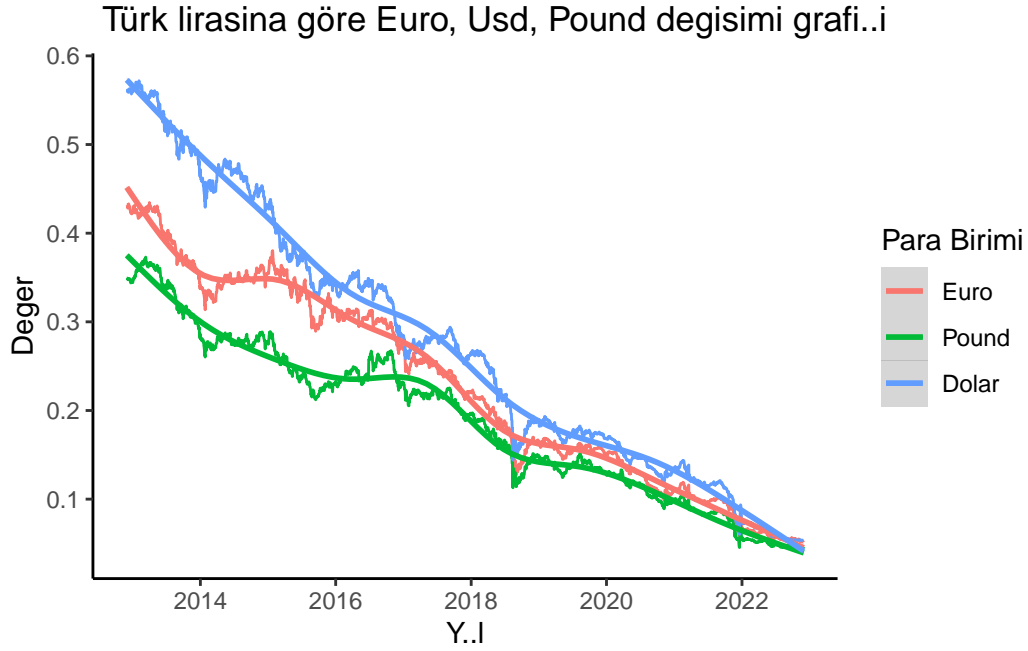
### Veri Setinin Hakkında

Bu veri seti son 10 yıldaki döviz değişimlerini içermektedir

### Türk lirasına göre Euro, Usd, Pound degisimi grafiği

```
veri1 <- historical_exchange_rates(from = "TRY" , to ="EUR" ,
start_date = "2012-11-30" , end_date = "2022-11-30" )
veri2 <- historical_exchange_rates(from = "TRY" , to ="USD" ,
start_date = "2012-11-30" , end_date = "2022-11-30" )
veri3 <- historical_exchange_rates(from = "TRY" , to ="GBP" ,
start_date = "2012-11-30" , end_date = "2022-11-30" )

doviz <- veri1 %>% left_join(veri2, by = "date" )
doviz2 <- doviz %>% left_join(veri3, by = "date" )
dovizlast <- doviz2 %>%
pivot_longer(cols= starts_with("one"))
ggplot(dovizlast , aes(x = dovizlast$date , y= dovizlast$value , color = dovizlast$name))
geom_line() +
geom_smooth(se=TRUE) +
scale_color_discrete(labels = c("Euro", "Pound","Dolar")) +
labs (x ="Yıl" , y = "Deger" ,
title = " Türk lirasına göre Euro, Usd, Pound degisimi grafiği " ,
color ="Para Birimi") +
theme_classic()
```



## Rapor

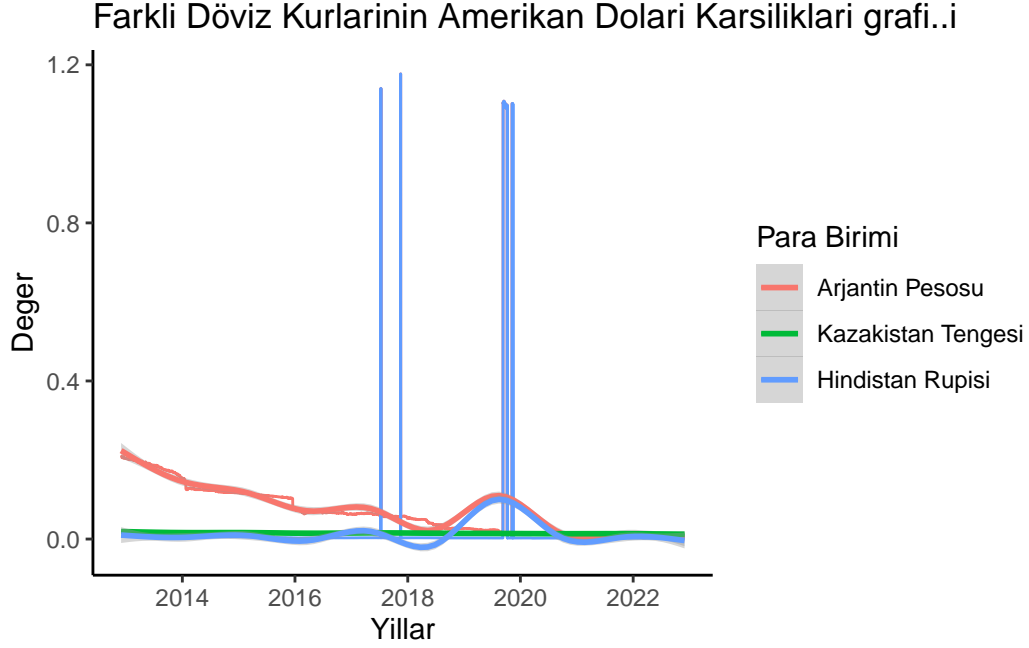
Grafikteki sonuçlara göre, son 10 yılda Türk lirasının bütün para birimlerine göre değer kaybettiği gözlemlenmektedir

## Farklı Döviz Kurlarının Amerikan Doları Karşılıkları grafiği

```
price1 <- historical_exchange_rates(from = "ARS" , to ="USD" ,
start_date = "2012-11-30" , end_date = "2022-11-30" )
price2 <- historical_exchange_rates(from = "KZT" , to ="USD" ,
start_date = "2012-11-30" , end_date = "2022-11-30" )
price3 <- historical_exchange_rates(from = "INR" , to ="USD" ,
start_date = "2012-11-30" , end_date = "2022-11-30" )

doviz3 <- price1 %>% left_join(price2, by = "date" )
doviz4 <- doviz3 %>% left_join(price3, by = "date" )
dovizson <- doviz4 %>%
pivot_longer(cols= starts_with("one"))
ggplot(dovizson , aes(x = dovizson$date , y= dovizson$value , color = dovizson$name)) +
geom_line() +
```

```
geom_smooth(se=TRUE) +  
scale_color_discrete(labels = c("Arjantin Pesosu" , "Kazakistan Tengeşi" , "Hindistan Rupisi"  
labs (x ="Yıllar" , y = "Değer" ,  
title = "Farklı Döviz Kurlarının Amerikan Doları Karşılıkları grafiği " ,  
color ="Para Birimi") +  
theme_classic()
```



## Rapor

Grafikteki sonuçlarımıza göre, dolara göre incelediğimiz para birimlerimiz genel olarak hep aynı oranlarda kaldığı gözlemlenmektedir. Arjantin Pesosu'nun yıllar içinde değer kaybettiğini, Hindistan Rupisi'nde 2018 ve 2020 yıllarında garip bir biçimde yükseldiğini sonrasında ise yeniden düştüğünü gözlemlemek mümkündür.

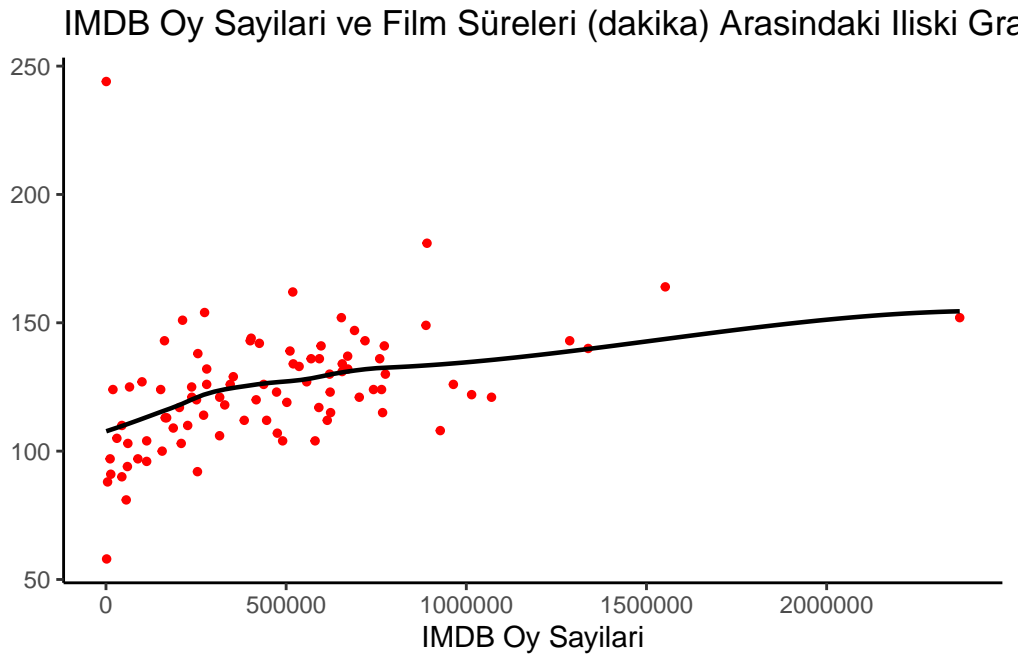
## Üçüncü veri

### Veri Setinin Hakkında

Bu veri seti Marvel Filmleri hakkında bilgiler içermektedir.

## IMDB Oy Sayilari ve Film Süreleri (dakika) Arasindaki Iliski

```
library(readr)
marvel <- read_csv("mdc.csv")
ggplot(marvel, aes(x = imdb_votes, y = runtime))+
  geom_point(size = 1, color = "red")+
  geom_smooth(se = FALSE, size = 0.8, color = "black")+
  labs(x="IMDB Oy Sayilari", y= "",
  title = "IMDB Oy Sayilari ve Film Süreleri (dakika) Arasindaki Iliski Grafigi") +
  theme_classic()
```

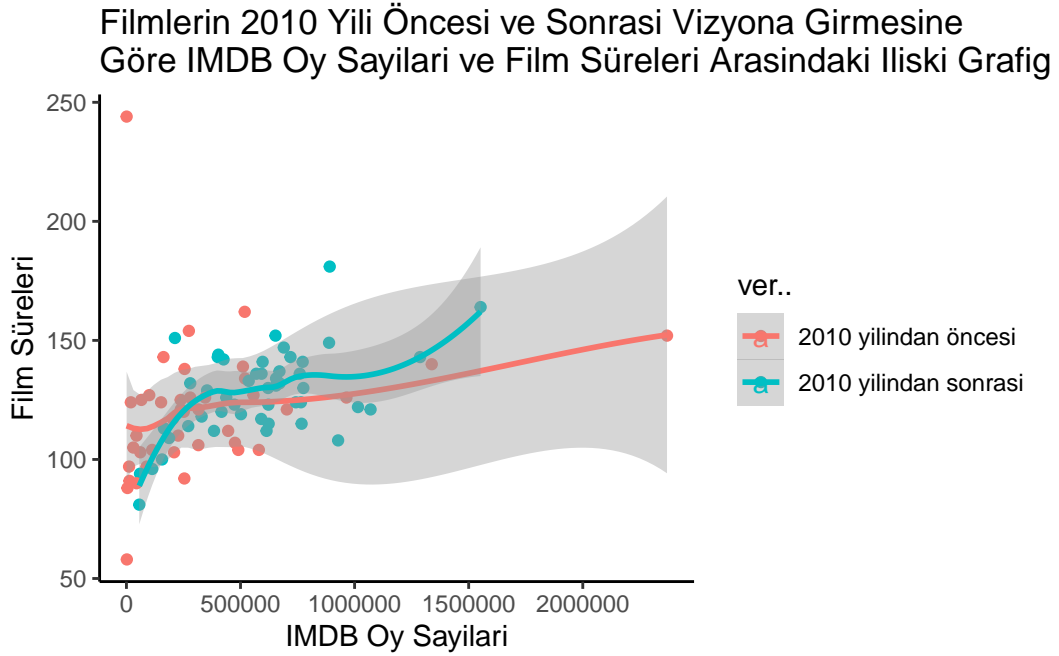


## Rapor

Grafikteki sonuçlarımıza göre Film süreleri genel olarak 75 ile 150 dakika aralığında çıkmaktadır ve en çok oyun ise film süreleri 100 dk ile 150 dakika arasında olan filmlerde alındığı gözlenmektedir

IMDB oy sayıları ve film süreleri arasındaki ilişkiyi, filmin 2010 yılı öncesi ve sonrası vizyona girmesi durumuna göre ilişkisi

```
marvel <- marvel %>%  
add_column(veri = if_else(marvel$year < 2010, "2010 yilindan öncesi", "2010 yilindan sonra"))  
ggplot(marvel, aes(x = imdb_votes, y = runtime, color= veri))+  
geom_point()+  
geom_smooth(se = TRUE)+  
ggpubr::stat_cor(method = "loess") +  
labs(x="IMDB Oy Sayilari",  
      y= "Film Süreleri",  
      title = "Filmlerin 2010 Yili Öncesi ve Sonrasi Vizyona Girmesine  
Göre IMDB Oy Sayilari ve Film Süreleri Arasindaki Iliski Grafigi") +  
theme_classic()
```



## Rapor

Grafikteki sonuçlarımıza bakacak olursak film sürelerinin artmasıyla kullanılan oy sayılarının arttığını söyleyebiliriz.