

# **En Çok Satan Kitaplar, Harry Potter Karakterleri ve Marvel Filmleri Grafik ve Yorumları**

Zeynep Afra Sezer

## **Özet**

Bu raporda 2009-2019 yılları arasında Amazonun en çok satan kitapları, Harry Potter karakterleri ve Marvel filmleri veri setleri incelenip yorumlandırılmıştır. En çok satan kitaplar veri seti; türlerine ve aldıkları okuyucu puanlarına göre oranlandırılıp grafiklendirilmiştir. Harry Potter veri seti; karakterlerin cinsiyetlerine, karakterlerin evlerine ve karakterlerin Muggle olmaları durumlarına göre oranlandırılıp grafiklendirilmiştir. Marvel filmleri veri seti ise; filmlerin türlerine ve uzunluklarına göre oranlandırılıp grafiklendirilmiştir.

Bu rapor için kullanılan kütüphaneler;

```
install.packages("readr")
install.packages("dplyr")
install.packages("ggplot2")
install.packages("MetBrewer")
install.packages("tidyverse")
library(readr)
library(dplyr)
library(ggplot2)
library(MetBrewer)
library(tidyverse)
```

## EN ÇOK SATAN KİTAPLAR

Bu veri setinde 2009-2019 yılları arasında Amazon'un en çok satan 50 kitabının veri kümesi verilmiştir. Veriler kurgusal ve kurgusal olmayan olarak kategorize edilmiştir.

Veri Tanımı;

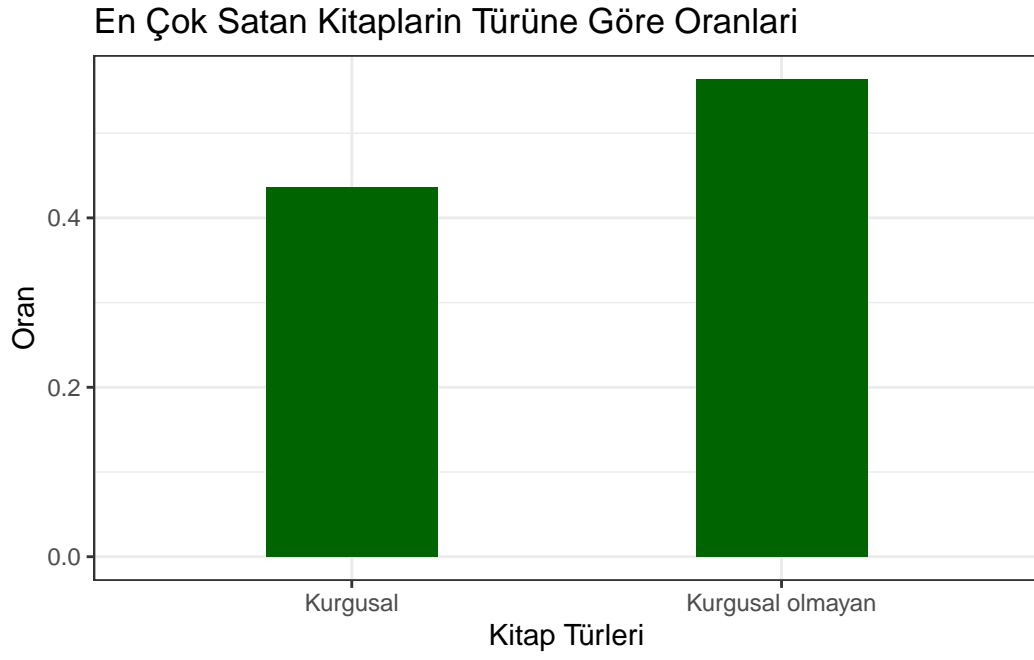
NAME: Kitabın adı, AUTHOR: Kitabın yazarı, USER RATING: Amazon kullanıcı değerlendirmesi, REVIEWS: Amazon'da yazılı yorum sayısı, PRICE: Kitabın fiyatı, YEAR: En çok satanlar listesinde yer aldığı yıllar, GENRE: Kurgusal- Kurgusal olmayan

### 1.1 En Çok Satan Kitapların Türlerine Göre Oranları

```
books <- read_csv("books.csv")

books <- books %>%
  group_by(Genre) %>%
  summarise(sayi = n()) %>%
  mutate(oran = sayi / sum(sayi))

ggplot(books, aes(x = Genre, y = oran)) +
  geom_bar(stat = "identity", width = 0.4,
          fill = "darkgreen") +
  labs(x = "Kitap Türleri", y = "Oran",
       title = "En Çok Satan Kitapların Türüne Göre Oranları") +
  scale_x_discrete(labels = c("Kurgusal", "Kurgusal olmayan")) +
  theme_bw()
```



### Yorum

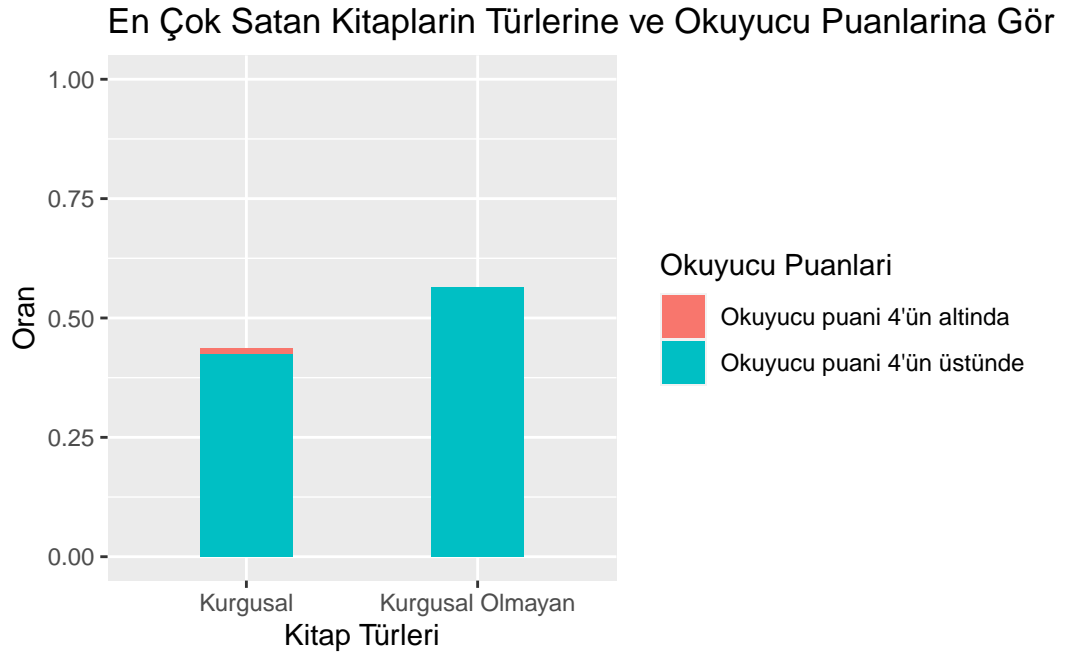
Kurgusal olmayan kitapların okunma oranı, kurgusal olan kitapların okunma oranından daha fazladır. Okuyucuların kurgusal olmayan kitap türünü daha çok tercih ettiğini söyleyebiliriz.

## 1.2 En Çok Satan Kitapların Türlerine ve Okuyucu Puanlarına Göre Oranları

```
books <- read_csv("books.csv")
books <- books %>%
  add_column(ur = if_else(books$`User Rating` < 4, "Okuyucu puanı 4'ün altında",
                        "Okuyucu puanı 4'ün üstünde"))

books1 <- books %>%
  group_by(Genre, ur) %>%
  summarise(sayi = n()) %>%
  mutate(oran = sayi / nrow(books))

books1 %>%
  ggplot()+
  geom_bar(aes(x = Genre, y = oran, fill = ur),
           stat = "identity",
           width = 0.4) +
  labs(x = "Kitap Türleri",
       y = "Oran",
       title = "En Çok Satan Kitapların Türlerine ve Okuyucu Puanlarına Göre Oranları",
       fill = "Okuyucu Puanları")+
  scale_x_discrete(labels = c("Kurgusal", "Kurgusal Olmayan")) +
  ylim(0, 1)
```



## Yorum

Okuyucuların kurgusal kitap türüne büyük oranda 4'ün üstünde puan verdiğini söyleyebiliriz. Kurgusal olmayan kitaplara ise 4'ün altında puan verilmemiştir. Yine, kurgusal olan kitapların kurgusal olmayan kitaplara göre daha az tercih edildiğini söyleyebiliriz.

## HARRY POTTER KARAKTERLERİ

Bu veri setinde Harry Potter filmlerinin karakterleri ile ilgili veriler verilmiştir.

Veri Tanımı;

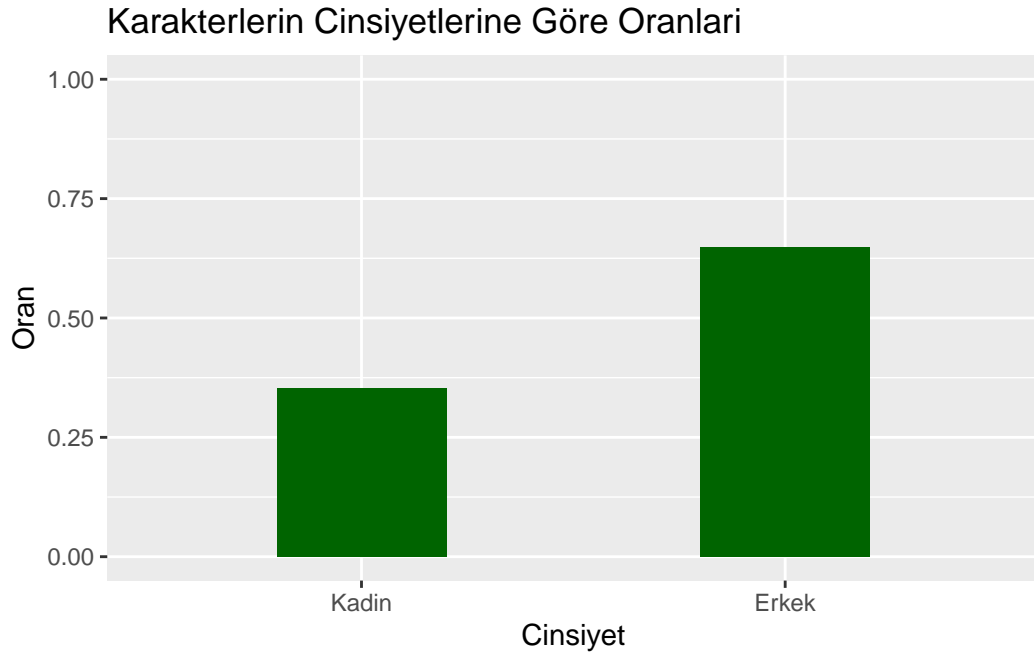
NAME: Karakterlerin isimleri, GENDER: Karakterlerin cinsiyetleri, JOB: Karakterlerin işleri, HOUSE: karakterlerin evleri, WAND: karakterlerin sahip olduğu asalar, PATRONUS: karakterlerin Patronus'ları, SPECIES: Türler, BLOOD STATUS: Karakterler saf kan mı Muggle doğumlu mu?, HAIR COLOUR: Karakterlerin saç renkleri

### 2.1 Karakterlerin Cinsiyetlerine Göre Oranı

```
characters <- read_delim("Characters.csv",
                        delim = ";", escape_double = FALSE,
                        trim_ws = TRUE)

harry1 <- characters %>%
  group_by(Gender) %>%
  drop_na(Gender)%>%
  summarise(sayi = n()) %>%
  mutate(oran = sayi / sum(sayi))

ggplot(harry1, aes(x = Gender, y = oran))+
  labs(x = "Cinsiyet",
       y = "Oran",
       title = "Karakterlerin Cinsiyetlerine Göre Oranları")+
  scale_x_discrete(labels = c("Kadin", "Erkek"))+
  geom_bar(stat = "identity", fill = "darkgreen",
          width = 0.4)+
  ylim(0, 1)
```



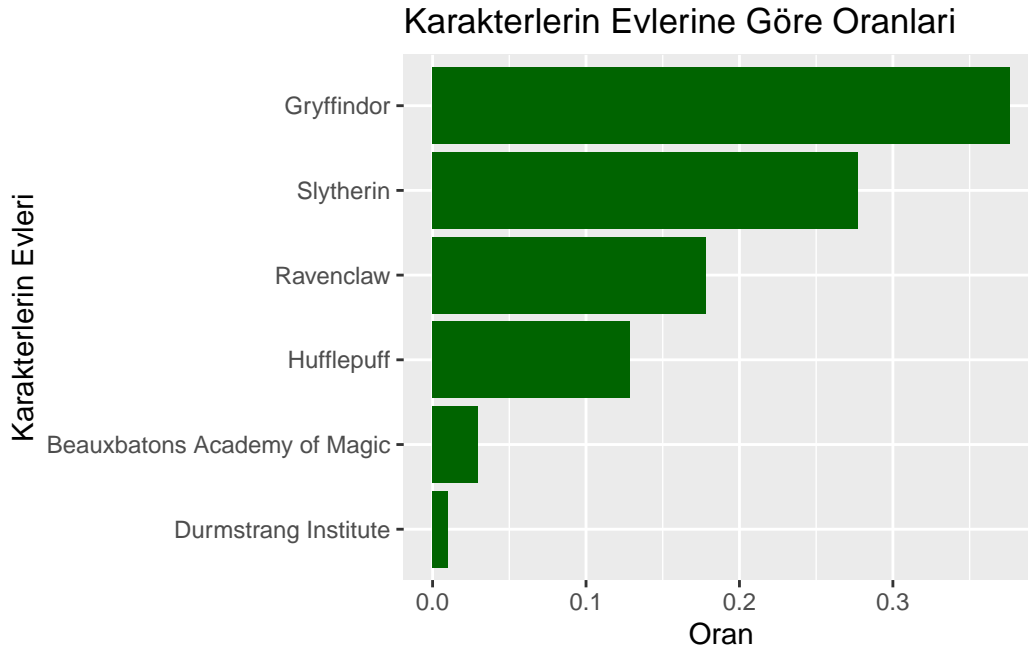
### Yorum

Harry Potter filmlerinde kadın karakterlerden çok erkek karakterlere yer verilmiştir. Kadın karakterlerin oranı %30 civarındayken, erkek karakterlerin oranı %70 civarındadır.



## 2.2 Karakterlerin Evlerine Göre Oranları

```
harry2 <- characters %>%  
  drop_na(House)%>%  
  group_by(House) %>%  
  summarise(sayi2 = n()) %>%  
  mutate(oran = sayi2 / sum(sayi2))  
  
ggplot(harry2, aes(x = oran, y = reorder(House, +oran)))+  
  labs(x = "Oran",  
       y = "Karakterlerin Evleri",  
       title = "Karakterlerin Evlerine Göre Oranlari") +  
  geom_bar(stat = "identity", fill = "darkgreen")
```



### Yorum

Harry Potter filmlerinde karakterlerin yaklaşık olarak %40'ı Gryffindor'dur. En az oranla üye sahibi olan ev ise Durmstrang Insitute'dır.

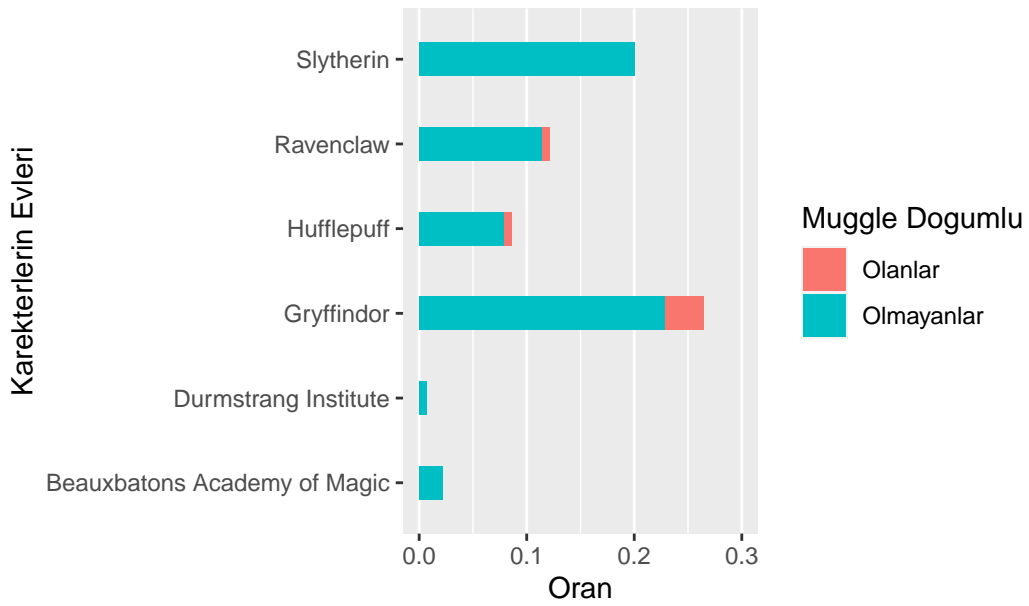
## 2.3 Karakterlerin Evlerine ve Muggle Olması Durumlarına Göre Oranları

```
characters <- characters%>%
  add_column(Muggle= if_else(characters$`Blood status`== "Muggle-born", "Olanlar",
                             "Olmayanlar"))

harry3 <- characters %>%
  group_by(House, Muggle) %>%
  drop_na(House, Muggle)%>%
  summarise(sayi = n())%>%
  mutate(oran = sayi / nrow(characters))

harry3%>%
  ggplot()+
  geom_bar(aes(x = House, y = oran, fill = Muggle),
           stat = "identity",
           width = 0.4) +
  labs(x = "Karakterlerin Evleri",
       y = "Oran",
       title = "Karakterlerin Evlerine ve Muggle Olması Durumlarına Göre Oranları",
       fill = "Muggle Doğumlu")+
  coord_flip() +
  ylim(0, 0.3)+
  theme(panel.grid.major.y = element_blank(),
        panel.grid.minor.y = element_blank())
```

## Karakterlerin Evlerine ve Muggle Olması D



## Yorum

Slytherin, Durmstrang Institute ve Beauxbatons Academy of Magic; Muggle doğumlu olmayan büyücülere ev sahipliği yapmaktadır bir diğer deyişle, saf-kan büyücülerden oluşmaktadırlar. Ravenclaw ve Hufflepuff evleri ise büyük oranda Muggle doğumlu olmayan büyücülere sahip olsa da az bir oranla Muggle doğumlu olan büyücülere de ev sahipliği yapmaktadır. Gryffindor ise diğer beş eve göre en fazla oranla Muggle doğumlu büyücülere ev sahipliği yapmaktadır.

## MARVEL FİLMLERİ

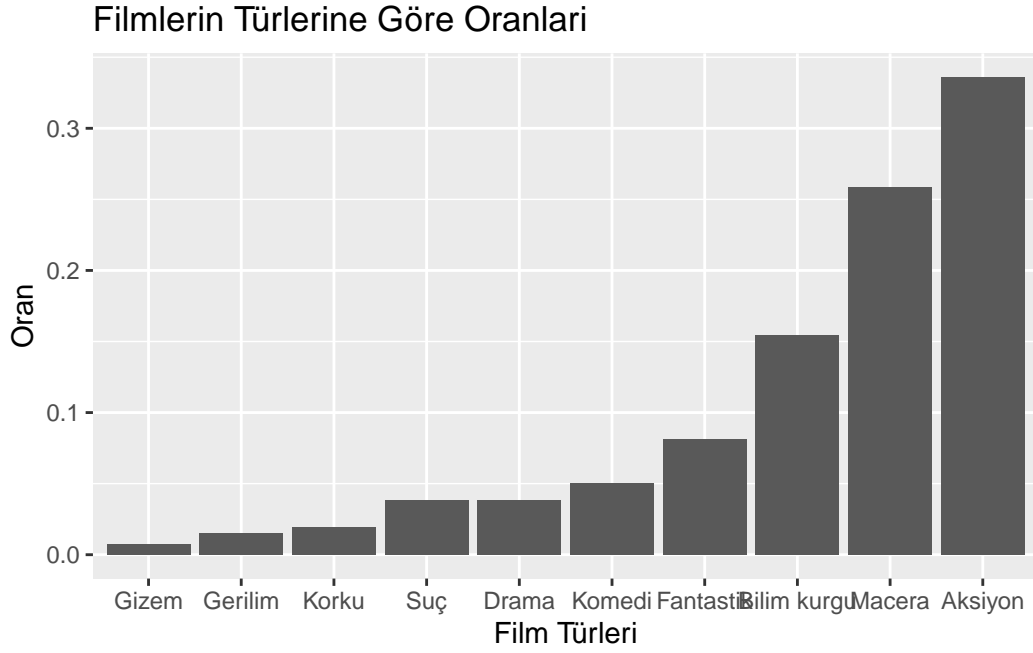
Bu veri setinde 1944 - 2020 yılları arasındaki tüm Marvel filmlerinin verileri verilmiştir.

Veri Tanımı;

TITLE: Filmin adı, Yıl: Çıkış tarihi, GENRE: Filmin türü, RUNTIME: dakika cinsinden film uzunluğu, MPAA\_RATING: Sinema Filmleri Birliği film derecelendirme sistemi, IMDB\_RATING: Yüzde 10 ile 1-10 ölçeğinde sayısal Derecelendirmesi, IMDB\_VOTES izleici puamı, GROSS: brüt kar, DIRECTOR: Film yönetmeni, STARS: en iyi 4 oyuncu, DESCRIPTION: filmin karakter açıklaması, CRIT\_CONSENSUS: Eleştirmenlerin film hakkındaki fikirleri (Profesyonel)

### 3.1 Filmlerin Türlerine Göre Oranları

```
marvel = read_csv("mdc.csv")
marvel <- marvel %>%
  add_column(süre = if_else(marvel$runtime < 120, "120'nin altında", "120'nin üstünde"))
marvel1 <- marvel %>%
  tidyr::separate_rows(genre, sep = ", ") %>%
  group_by(genre) %>%
  summarise(sayi = n()) %>%
  mutate(oran = sayi / sum(sayi))
marvel1%>%
  ggplot()+
  geom_col(aes(x = reorder(genre, +oran), y = oran)) +
  labs(x="Film Türleri", y= "Oran",
       title = "Filmlerin Türlerine Göre Oranları") +
  scale_x_discrete(labels = c("Gizem", "Gerilim", "Korku", "Suç", "Drama", "Komedi",
                              "Fantastik", "Bilim kurgu", "Macera", "Aksiyon"))
```

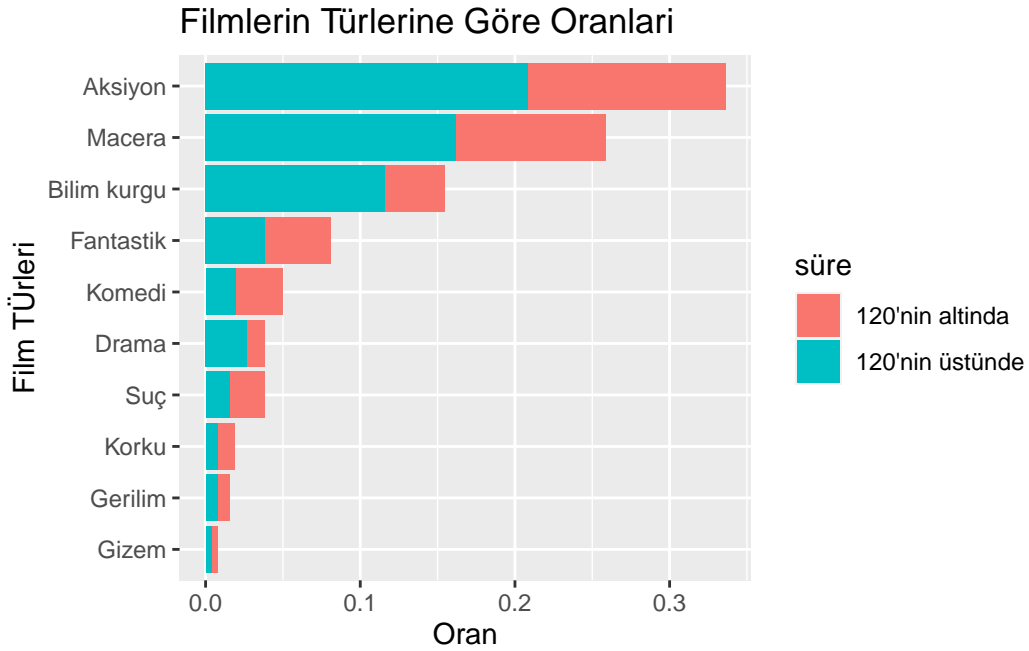


## Yorum

Marvel filmleri büyük oranla aksiyon türünden oluşmaktadır. Ardından macera ve bilim kurgu türleri en fazla orana sahiptir. Filmler en az oranla gizem türündedir. Suç ve drama türleri eşit orandadır.

## 3.2 Filmlerin Türlerine ve Uzunluklarına Göre Oranları

```
marvel2 <- marvel%>%
  tidyr::separate_rows(genre, sep = ", ") %>%
  group_by(genre, süre) %>%
  summarise(sayi = n()) %>%
  mutate(oran = sayi / sum(marvel1$sayi))
marvel2 %>%
  ggplot()+
  geom_bar(aes(x = reorder(genre, +oran), y = oran , fill = süre),
    stat = "identity")+
  coord_flip()+
  labs(x="Film Türleri", y= "Oran",
    title = "Filmlerin Türlerine Göre Oranları")+
  scale_x_discrete(labels = c("Gizem", "Gerilim", "Korku", "Suç", "Drama", "Komedi",
    "Fantastik", "Bilim kurgu", "Macera", "Aksiyon"))
```



## Yorum

Aksiyon, macera, bilim kurgu, drama türlerinde filmlerin süreleri büyük oranda 120 dakikanın üstündedir. Fantastik, gerilim ve gizem türlerinde ise süre dağılım oranlarının eşit olduğunu

söyleyebiliriz. Komedi, suç ve korku türlerinde de filmlerin süresi büyük oranda 120 dakikanın altındadır.