Anderson Amorim

https://www.kaggle.com/datasets/octopusteam/imdb-top-1000-movies/data

```
a <- ggplot(filmes, aes(x = averageRating)) +
  geom_density(fill = "lightblue", alpha = 0.5, adjust = 1.5) +
  labs(x = "Avaliação Média", y = "Densidade") +
  theme_tufte() +
  theme(axis.text.x = element_text(angle = 0, vjust = 0.5, hjust=0.5)) +
  geom_vline(aes(xintercept = mean(averageRating)),
             color="gray", linetype="dashed", size=1)

b <- ggplot(filtered_filmes, aes(x = numVotes, y = averageRating)) +
  geom_point(alpha = 0.6, color = "lightblue") +
  geom_segment(aes(x = 32000, y = 9.4, xend = 32000, yend = 14.4),
               linetype = 3, color = "gray", size = 1) +
  geom_segment(aes(x = 770000, y = 9.4, xend = 770000, yend = 14.4),
               linetype = 3, color = "gray", size = 1) +
  geom_segment(aes(x = 32000, y = 9.4, xend = 770000, yend = 9.4),
               linetype = 3, color = "gray", size = 1) +
  geom_segment(aes(x = 32000, y = 14.4, xend = 770000, yend = 14.4),
               linetype = 3, color = "gray", size = 1) +
  labs(x = "Número de Votos",
       y = "Avaliação Média") +
  scale_x_log10(labels = scales::comma) +
  theme_tufte()

c <- filmes %>%
  group_by(releaseYear) %>%
  summarise(numMovies = n()) %>%
  ggplot(aes(x = releaseYear, y = numMovies)) +
  geom_line(color = "lightblue") +
  labs(x = "Ano de Lançamento", y = "Número de Filmes") +
  theme_tufte()

d <- filmes %>%
  separate_rows(genres, sep =",") %>%
  group_by(genres) %>%
  summarise(numMovies=n(), avgRating=mean(averageRating)) %>%
  ggplot(aes(x=reorder(genres,-numMovies), y=numMovies)) +
    geom_bar(stat="identity", fill="lightblue") +
    coord_flip() +
    labs(x="Gênero", y="Número de Filmes") +
    theme_tufte()
```

1000 filmes mais bem avaliados - IMDB