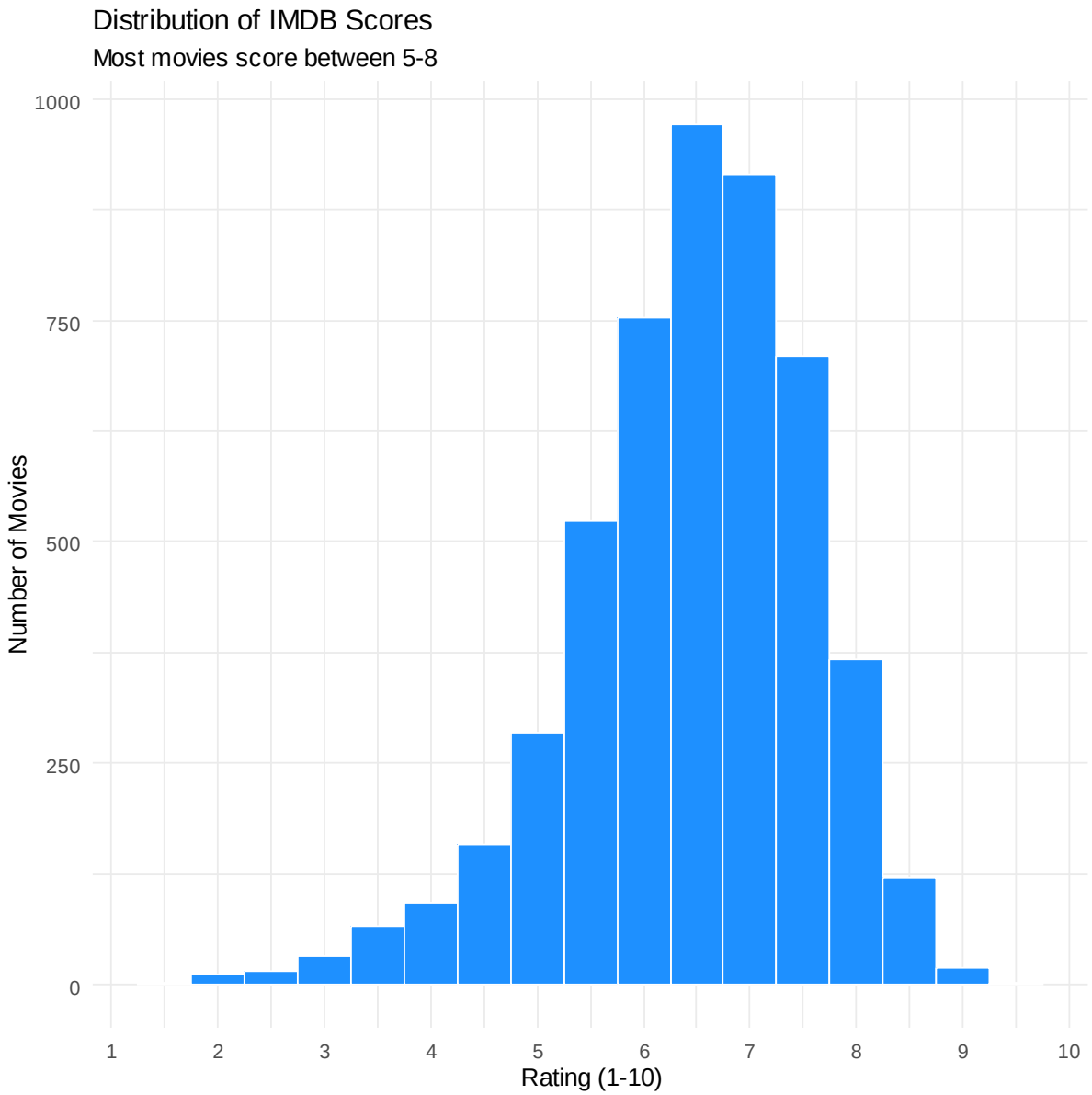
AIM: **Movie Ratings Analysis.**

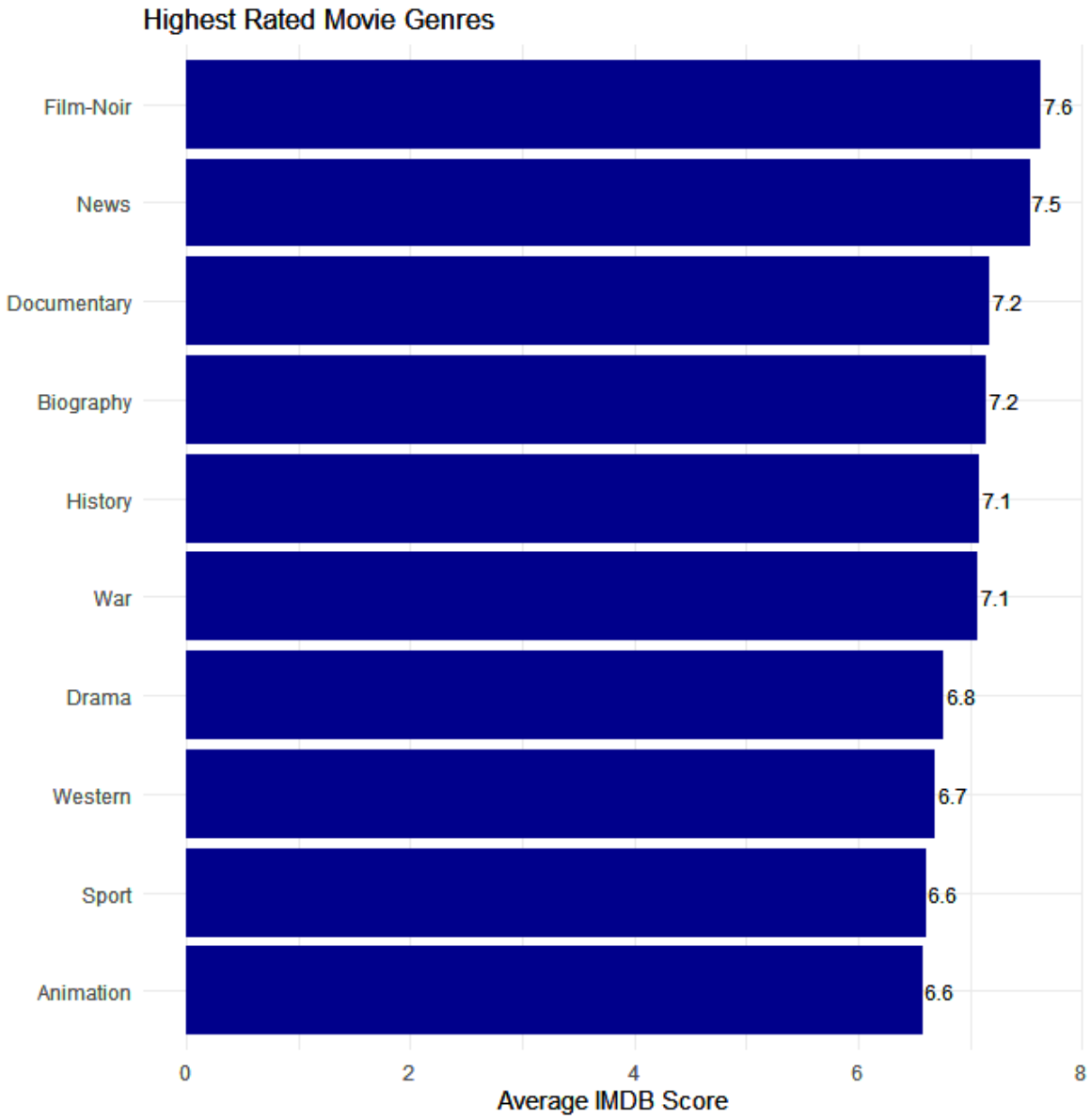
Code:

```
install.packages("tidyverse")
library(tidyverse)
movies <- read_csv("movie_metadata.csv")
ggplot(movies, aes(x = imdb_score)) +
 geom_histogram(binwidth = 0.5, fill = "dodgerblue", color = "white") +
 labs(title = "Distribution of IMDB Scores",
      subtitle = "Most movies score between 5-8",
      x = "Rating (1-10)",
      y = "Number of Movies") +
  scale_x_continuous(breaks = 1:10) +
  theme_minimal()
movies %>%
  separate_rows(genres, sep = "\\|") %>%
  group_by(genres) %>%
  summarise(avg_rating = mean(imdb_score, na.rm = TRUE)) %>%
  arrange(desc(avg_rating)) %>%
  head(10) %>%
  ggplot(aes(x = reorder(genres, avg_rating), y = avg_rating)) +
  geom_col(fill = "darkblue") +
  coord_flip() +
  labs(title = "Highest Rated Movie Genres",
      x = "",
      y = "Average IMDB Score") +
  geom_text(aes(label = round(avg_rating, 1)),
          hjust = -0.1, size = 3) +
  theme_minimal()
```

dataset: https://www.kaggle.com/datasets/carolzhangdc/imdb-5000-movie-dataset

OUTPUT:




Highest Rated Movie Genres
Film-Noir
News
Documentary
Biography
History
War
Drama
Western
Sport
Animation
7.6
7.5
7.2
7.2
7.1
7.1
6.8
6.7
6.6
6.6
0
2
4
6
8
Average IMDB Score