

# Microsoft Movie Analysis

---

Calvin Ruto

May 22nd, 2022

# Summary

Analysis of TheMovieDB dataset for films the head of Microsoft movies studio can publish reveals that movies whose genres are **superhero, fantasy and science fiction** are the most popular as well as best average rated movies genres in the current market. On the other hand, analysis of the Rotten Tomatoes dataset show that the most popular movie genres are **Action and Adventure, Comedy|Drama|Romance, Action and Adventure|Art House and International, Art House and International|Documentary, Action and Adventure|Mystery and suspense.**

# Outline

- Business Problem
- Data
- Methods
- Results
- Conclusions

# Business Problem

- Recommend the genre of best-rated and popular movies currently trending in the market through analysis of the various movies datasets.

# Data

- The dataset contains more than 20000 different movies from 1930 to date.
- The datasets sources I used are:
  - Rotten tomatoes.
  - TheMovieDB.
- TheMovieDB includes properties like movie popularity, average\_vote rating and movie titles both original and current.
- Rotten tomatoes dataset includes properties like genres, movie rating and movie titles.

# Methods

- **Data preparation**

Cleaned the data by removing duplicates and checking and removing missing values where necessary.

- **Analysis**

Performed analysis for the most popular and best rated movies as well as movie genres. Went further and analyzed the movies that were produced after the year 2010 so as to get the current movie trends in the market.

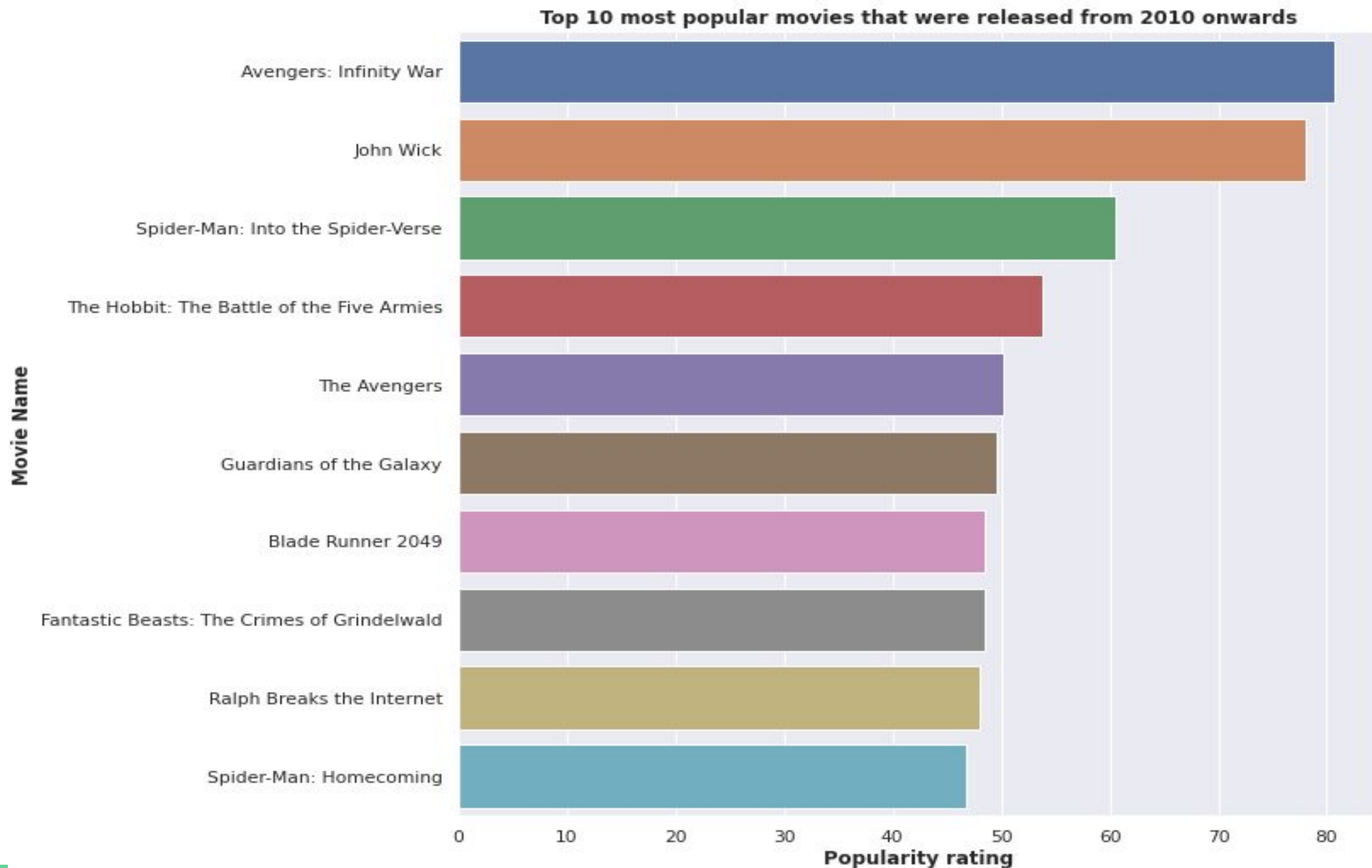
- **Visualization**

Drew visualizations on some of the analysis done.

# Results - TheMovieDB

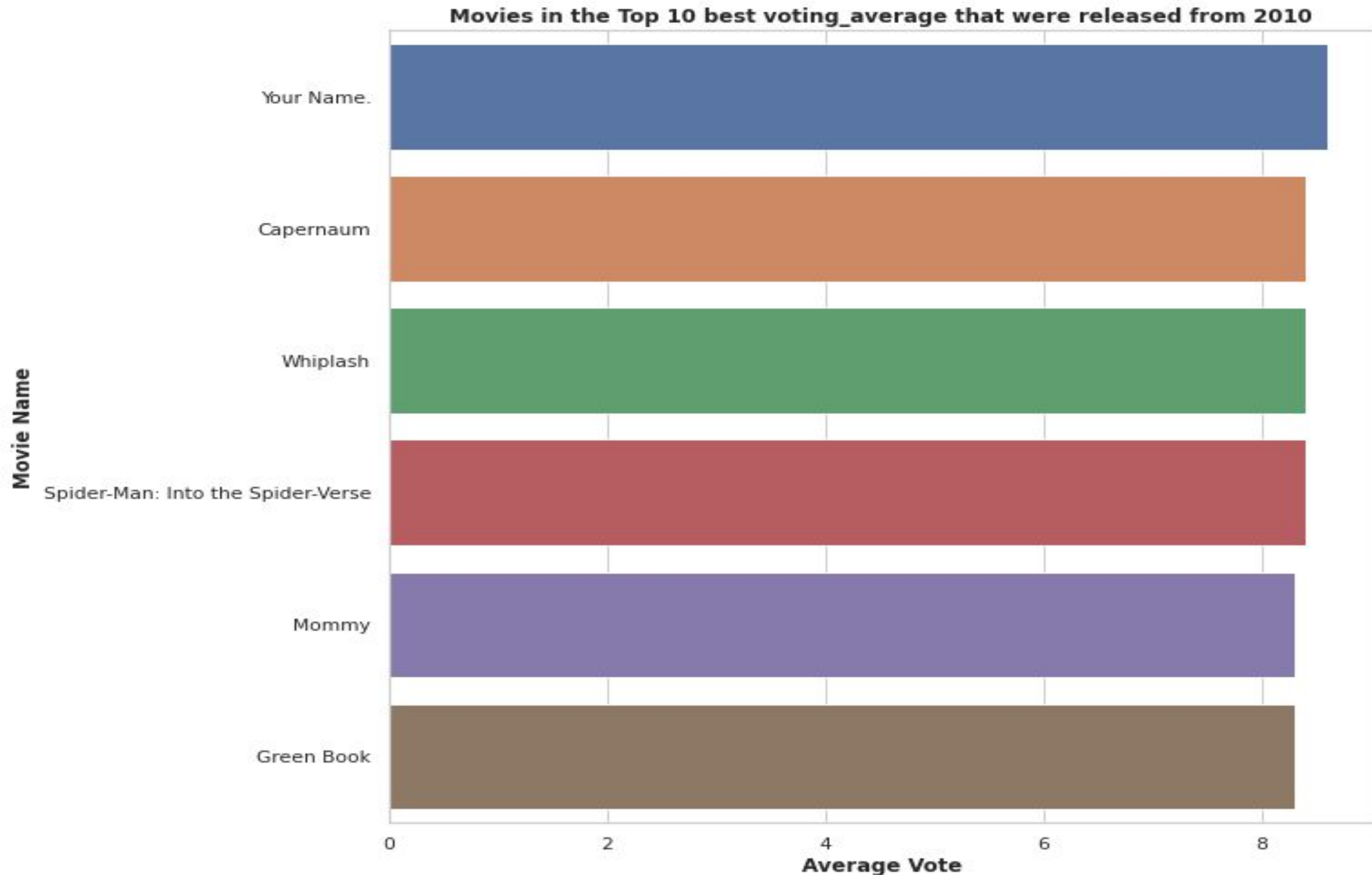
The plot shows the movies from 2010 that were most popular.

Avengers: Infinity war was the most popular with a score of 80.773



# Results - TheMovieDB

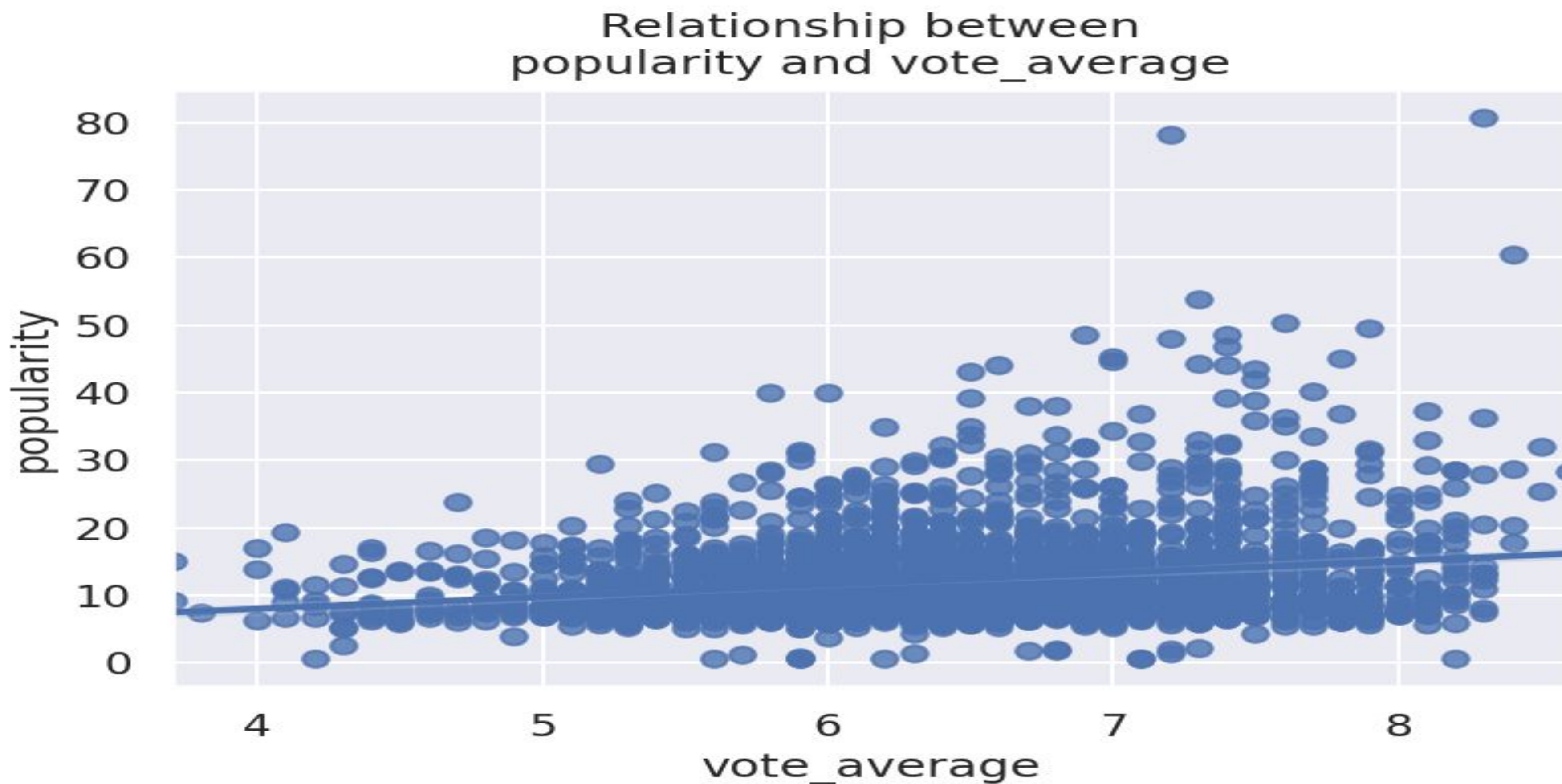
The plot shows the top 10 movies from 2010 that had the highest average votes . Your name is the highest with a score of 8.6





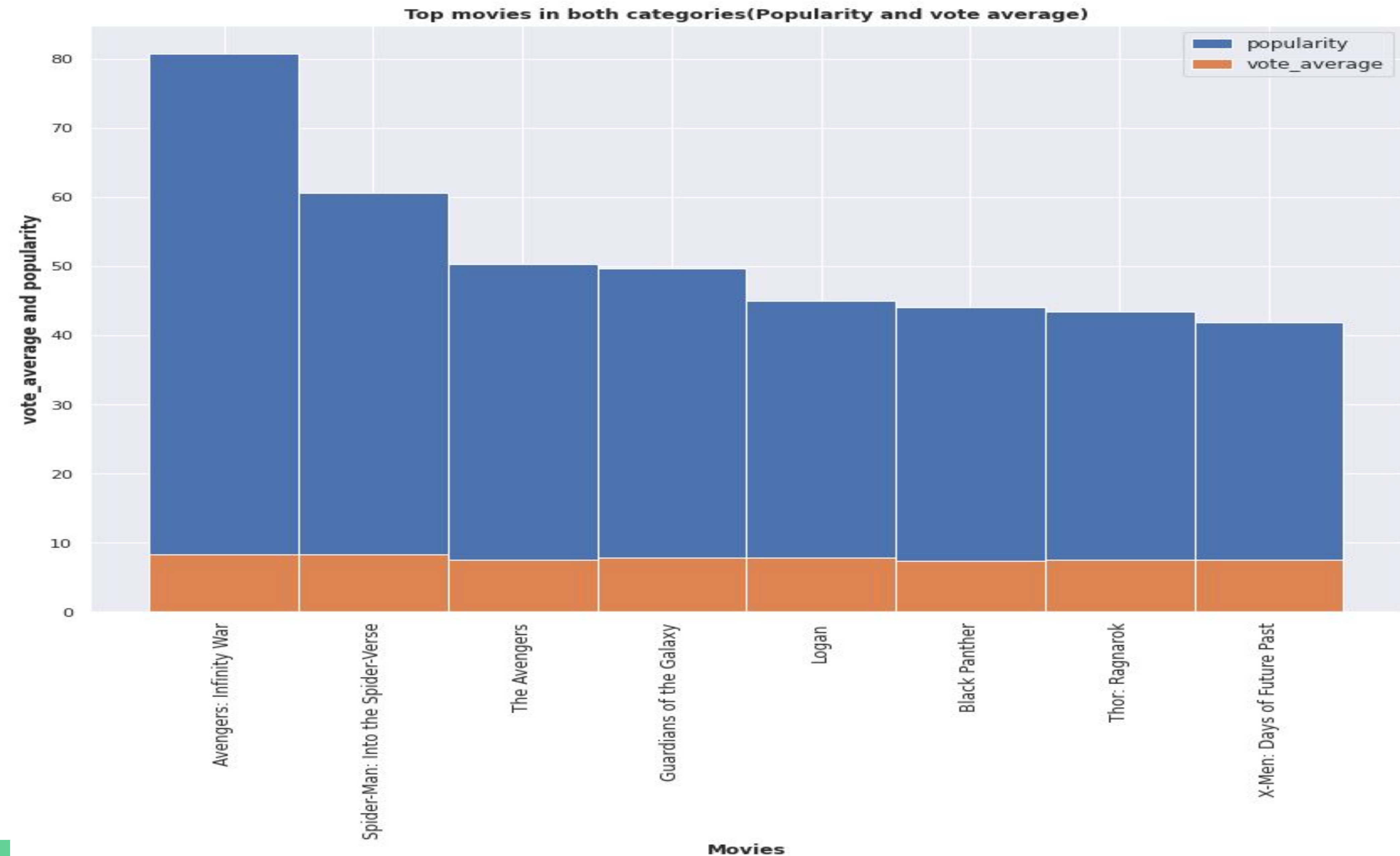
# Results - TheMovieDB

The plot shows the correlation between popularity and vote\_average. With an increase in movie popularity, the average-rating also increases though the relationship is weak meaning such occurrences are not very common



# Results - TheMovieDB

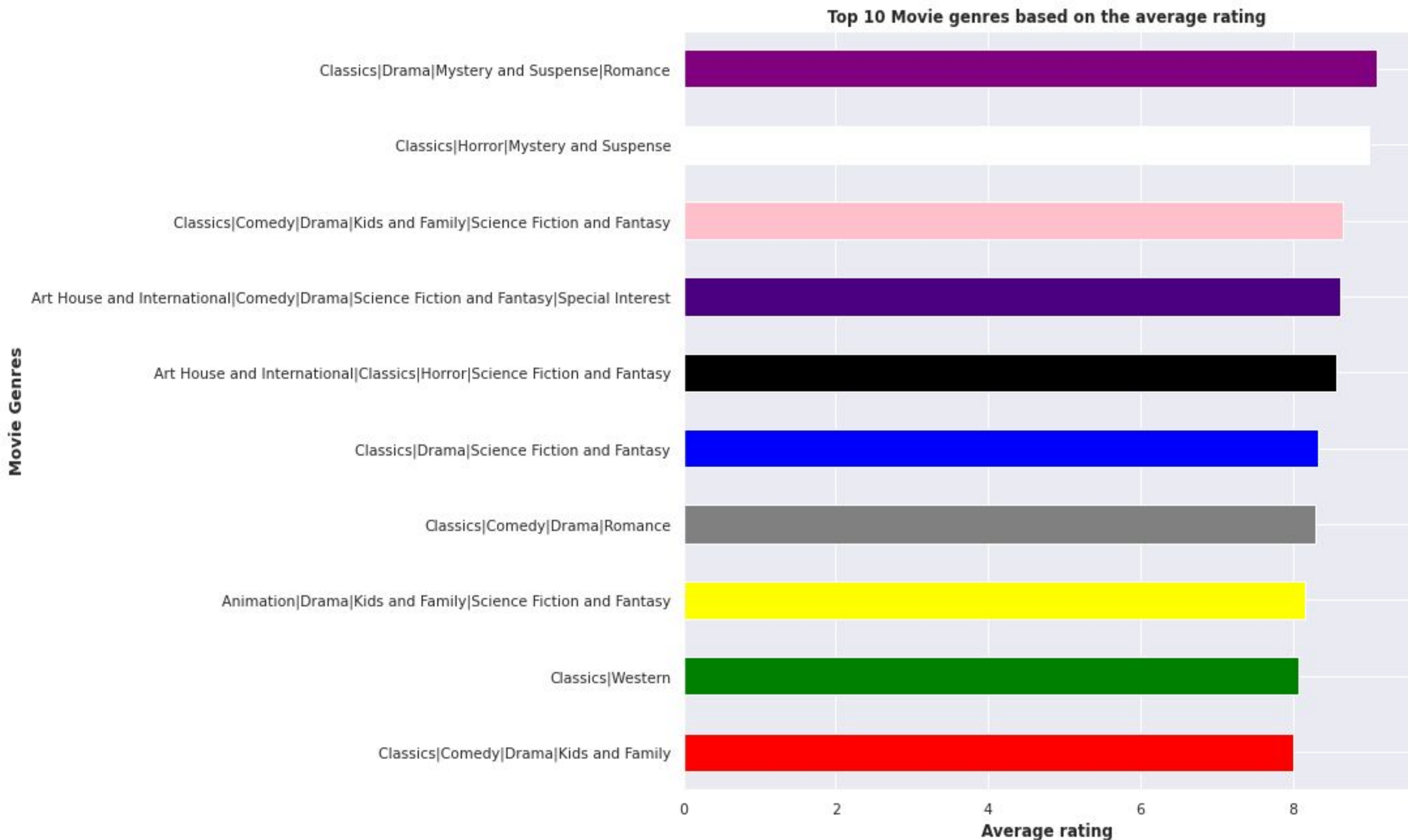
Comparing the top movies in both categories ,i.e., popularity and average vote score. This helps to get the top movies based on popularity as well as vote average.



# Results - Rotten Tomatoes

The plot shows the top movie genres based on average rating.

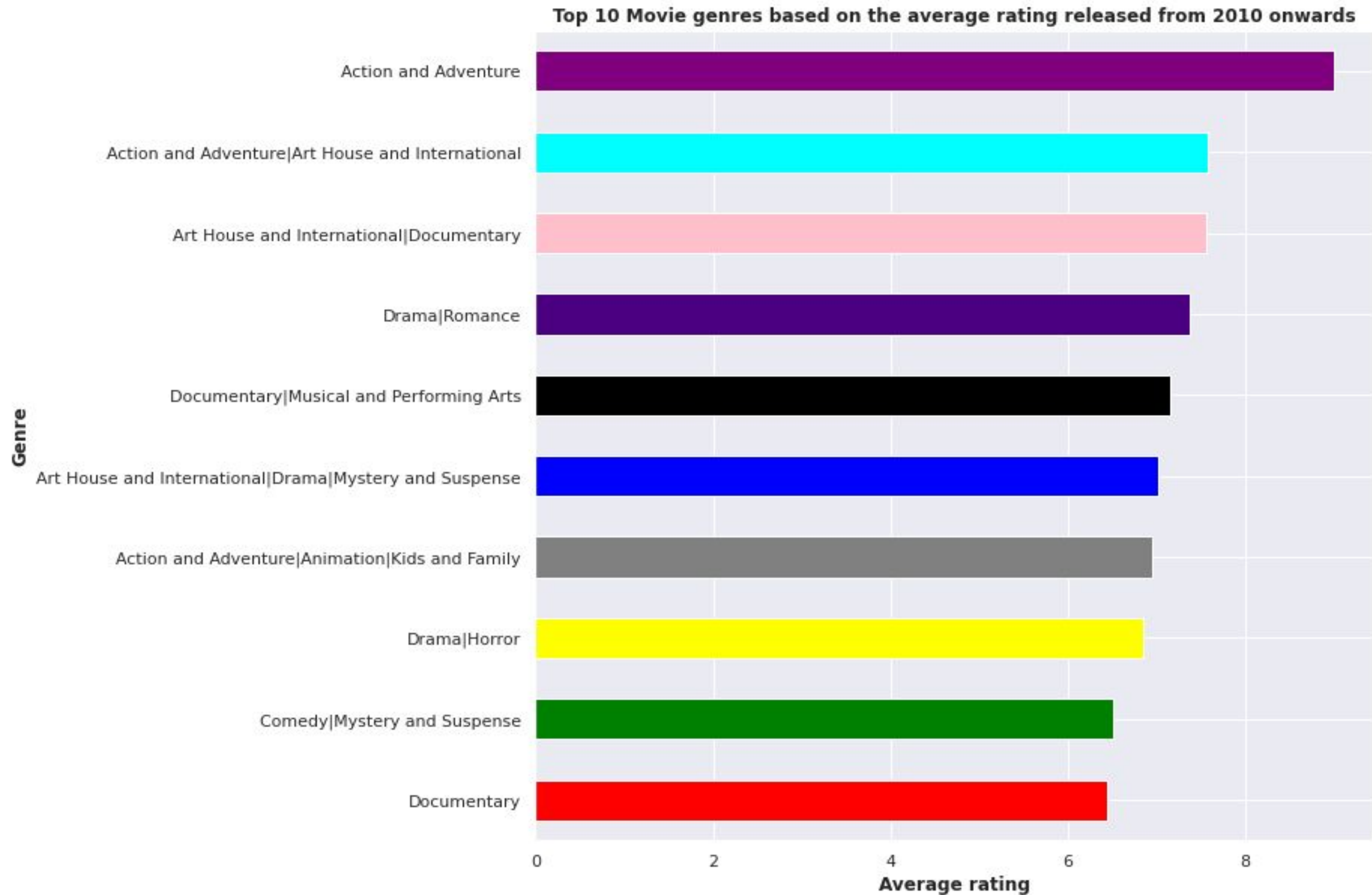
Classics|Drama|Mystery and Suspense|Romance is the top rated movie genre.





# Results - Rotten Tomatoes

The plot shows the top movie genres based on average rating released from 2010 onwards. Action and Adventure is the top rated movie genre.



# Conclusions

After the analysis that was done on the TheMovieDB dataset, most movies that are currently trending in the market belong to **Superhero, Science fiction and Fantasy** genres.

On the other hand, Rotten tomatoes dataset shows that the most popular movie genres are **Action and Adventure, Comedy|Drama|Romance, Action and Adventure|Art House and International, Art House and International Documentary, Action and Adventure|Mystery and suspense.**

However, TheMovieDB dataset was limited in that the movies were not be classified into the genre names they belong to. Therefore, the analysis was limited based on this classification and I had to make my own conclusions on the genre the top movies belong to.

# Recommendations

Based on the analysis and conclusions mentioned above, I would recommend to the head of microsoft movie studio to make films that fall into the genres outlined below. The genres include:

- I. Action and Adventure.
- II. Comedy|Drama|Romance.
- III. Action and Adventure|Art House and International.
- IV. Art House and International | Documentary.
- V. Action and Adventure|Mystery and suspense.

This list shows the genres to consider when making films with the first genre having the highest priority while the fifth one the least.

Thank You!

**Email:** calvin.ruto@student.moringaschool.[com](mailto:calvin.ruto@student.moringaschool.com)

**GitHub:** @clain6

**LinkedIn:** [linkedin.com/in/username/](https://www.linkedin.com/in/username/)