

2023

SAGE HOEPFL

Data Science Project: Disney Success Data

Problem Statement

WHERE DOES THE SOURCE OF SUCCESS RESIDE?

The problem at hand is to identify the key factors and values that have contributed to the tremendous success of Disney as a company.

LET'S GET STARTED



Benefits of Analysis



Continued Success

This analysis can provide valuable insights into how other organizations can replicate Disney's success and improve their own performance in the industry.



Director Success

This analysis offers a perspective on the correlation between a director's contribution and the success of a movie, as measured by the total gross income attributed to that director.



Movie Success

This analysis offers valuable insights into how the genre and title of a movie contribute to the success of Disney films.

Methodology

1.

To compare values between the movies' total gross dataset and Disney directors dataset, **I merged the two datasets.**

2.

I **utilized groupby functions** to establish the relationship between directors, genre, rating, top movies, release date, and total gross income.

3.

Visualizations were generated to illustrate the relationship between these distinct values. A machine learning model was then created to predict the director based off the movie title.

Challenges

1.

To deal with the challenge of having the total gross income format in a string that included symbols like '\$' and ',', I resolved the issue by utilizing the replace function and converting the string into an integer format.

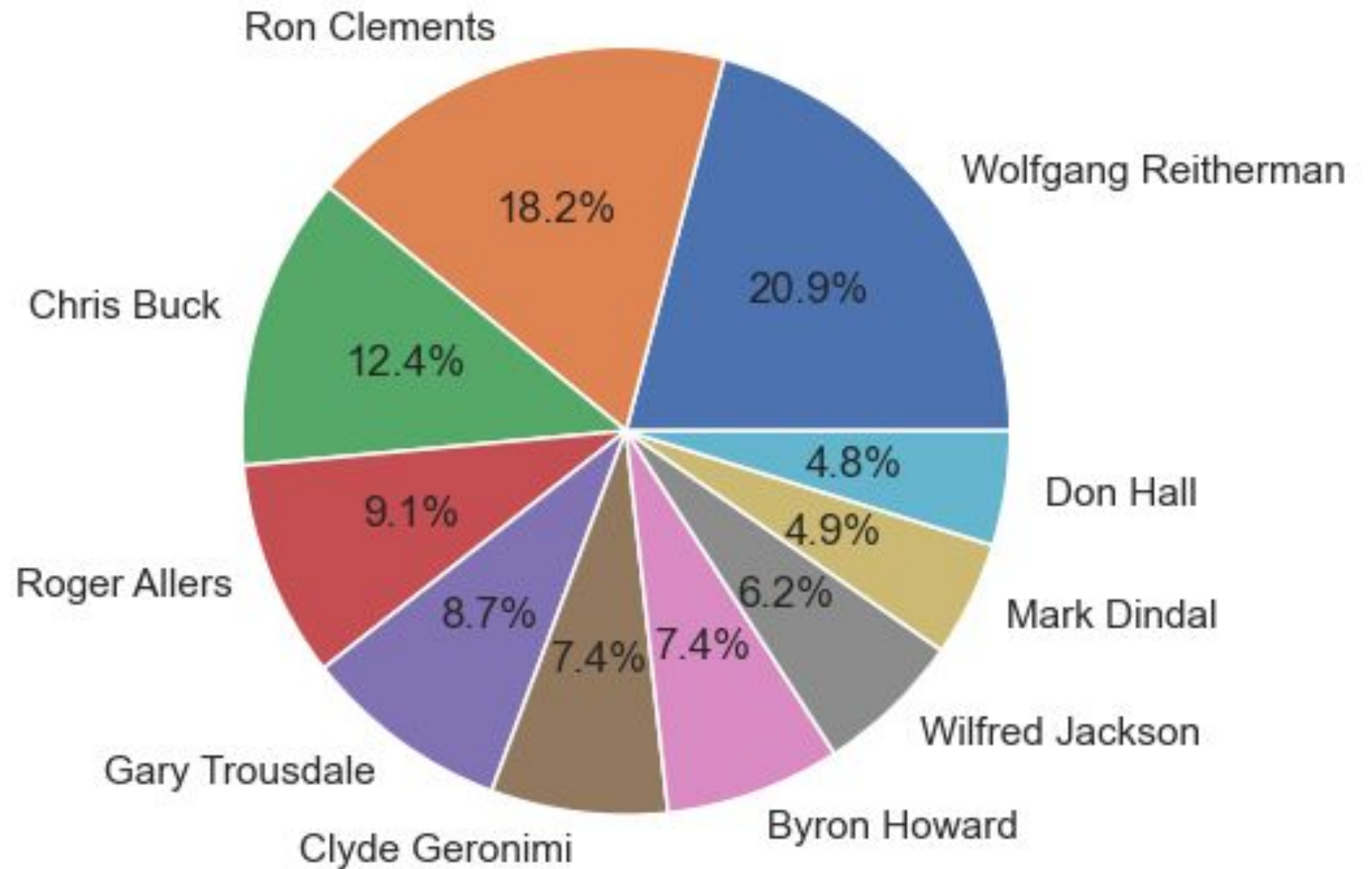
2.

While developing a scatter plot visualization for top-grossing movies, I encountered an issue where the highest values were displayed at the bottom of the chart. To resolve this, I sorted the total gross values in ascending order.

Which top ten directors performed the best?

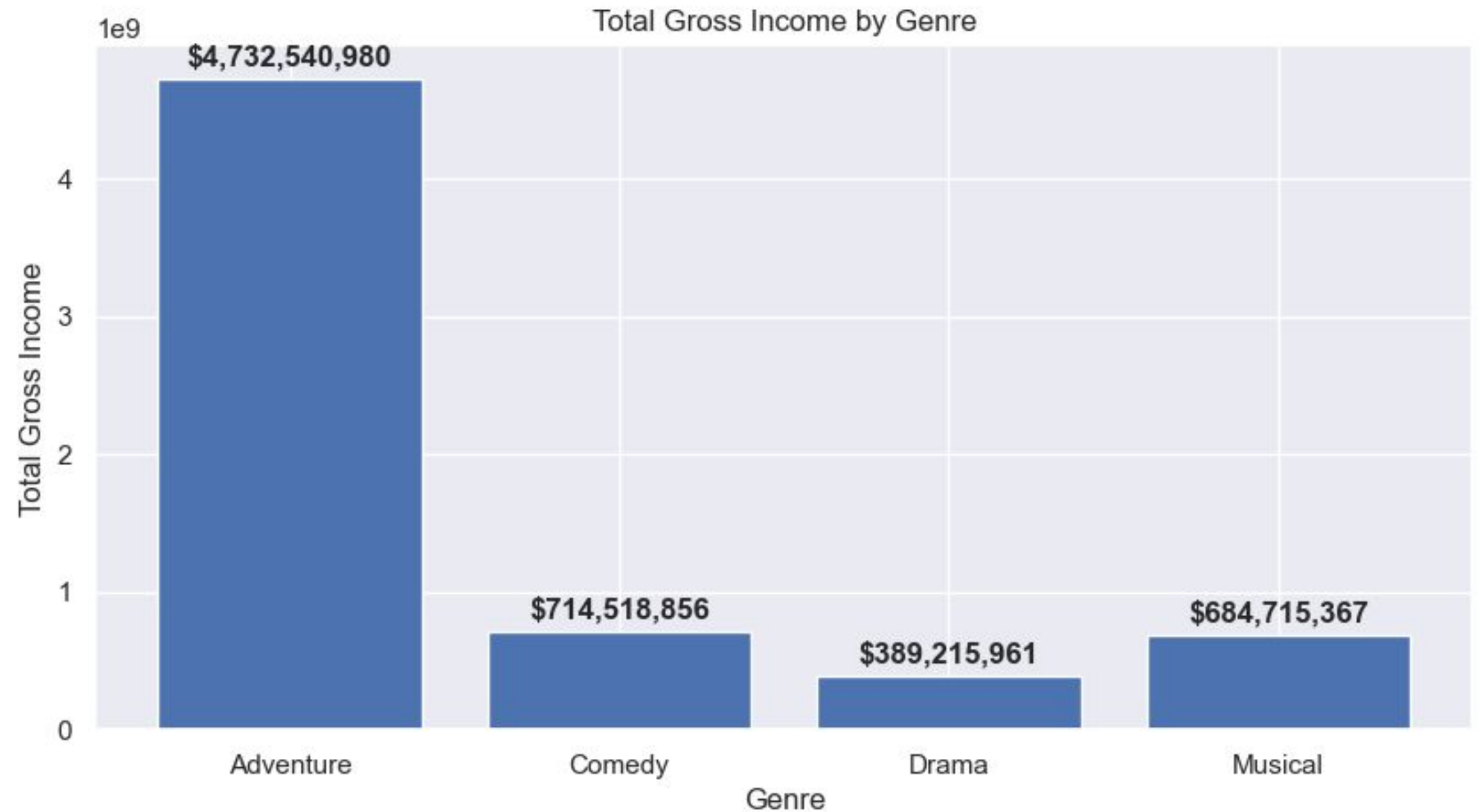
Top 10 Directors by Total Gross

Wolfgang
Reitherman is the
top director,
then Ron
Clements, then
Chris Buck.



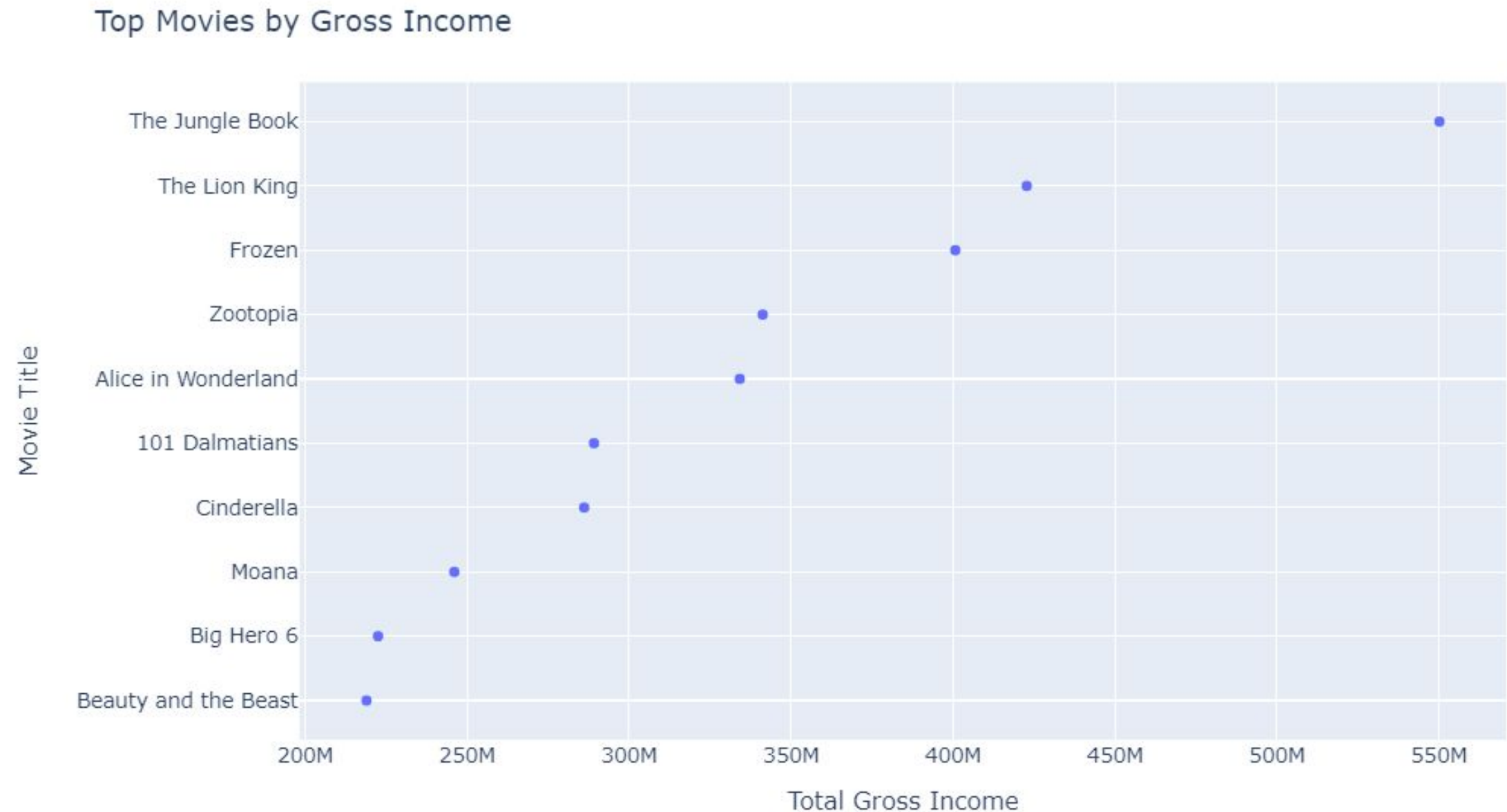
Which genre performed the best?

Adventure performed the best regarding total gross income.



Which movie performed the best?

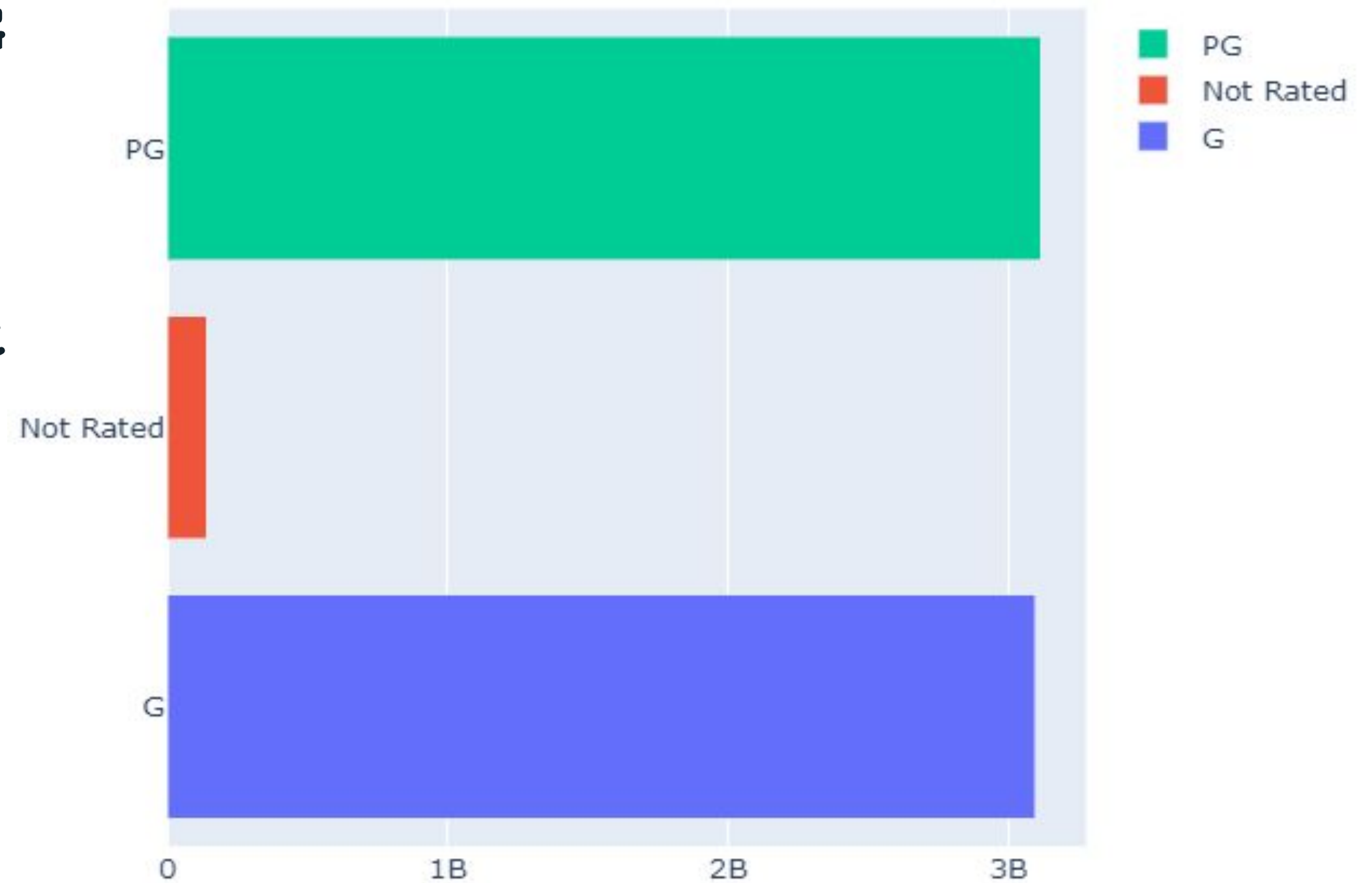
Jungle Book
performed the best
due to the 3
Jungle Book movies
being recorded in
the dataset



Does the rating affect the gross income?

The rating between PG and G movies in regards to total gross income does not show a large difference, as not rated provides less income.

Total Gross Income by MPAA Rating



Machine Learning Model

I trained a machine learning model to accurately predict the director of a movie based on its title input. The model performed well and produced accurate director predictions.





Recommendations

DIRECTORS

To maintain success in movies, it is advisable to either retain the same directors for future projects or select directors who possess a similar directing style.

MOVIES

Generating sequels for successful movies or those with a comparable style is recommended.

GENRE

It's important to produce a larger number of adventure-themed movies while creating films since they generally perform well.

RATING

Keep ratings of movies to either PG or G as Not Rated does not provide as much income as the other ratings.

Future Work

Hire Diverse and emerging directors



Disney could expand its pool of movie directors by hiring more diverse and emerging talents. This could bring fresh perspectives and storytelling techniques to their movies.

Focus on family-friendly movies



Disney has built a strong brand around family-friendly movies, and they could continue this trend by producing more movies that families can enjoy together.

Experiment with different genres



While Disney is well-known for its animated and fantasy movies, they could also explore other genres like action and science fiction to appeal to a wider range of audiences.

Expand to different ratings



Disney could also explore producing movies with higher ratings like PG-13 or even R-rated movies. This could appeal to older audiences and allow Disney to explore more mature themes and stories.



SAGE HOEPFL

Thank you!

Questions?