

THE MOVIE DATASET

BUILD A DATASTORY MIDTERM PROJECT

By Oluwasegun Alabi

Problem Statement

- What are the key features that are commonly found in the top 5% of highly-rated movies?

Analysis Overview

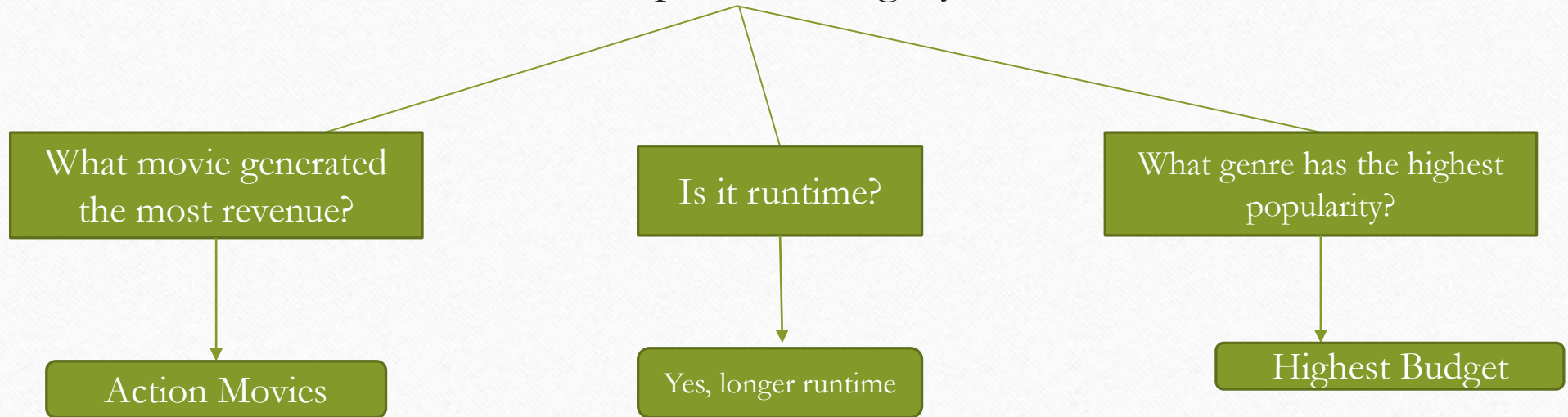
- Movie genre by profit.
- Movie genre by revenue.
- Movie genre by popularity.

Executive summary

- The analysis will be examined based on several factors that contributed to the high rating of top 5% of movies, including the movie's runtime, genre, and budget.
- Predicting the success of future movies can be improved based on the top 5% of movie genres that have the highest popularity, profitability, and revenue.

Issue Tree

What are the top 5% of highly rate movies?



Genre By Profits

- The profits will be determined by deducting the budget from the revenue.
- A bar chart plot of the movie genre against the profits.

Genre By Revenue

- An histogram plot will be used to display the movie genre plotted against revenue.
- This would reflect the highest revenue generated and what movie genre.

Genre By Popularity

- We would take a look at the most popular distribution against the movie titles and genre.
- The movie genre will be plotted against the movie popularity.
- An horizontal bar chart will be use to display the results.

Movie Runtime

- A chart or a tree map will be use to show the highest movie runtime.
- This will allows us to detect highly rated movie drawn from a conclusion of movie runtime.

Limitations

- Data Collection
- Data processing
- Data Cleaning
- Data insights

Conclusion/Next steps

- After a clean data, we identify sources for missing budget and revenue values.
- Visualization would be created after cleansing and the insights would drive recommendations for the business.