



Microsoft Studio (EDA)

JIMCOLLINS WAMAE

12TH MARCH 2023

Overview

- ▶ Microsoft is looking at starting up a movie studio. This presentation will bring forth the various insights achieved through exploratory data analysis of the provided movie datasets so as to generate business recommendations that will enable the business in selecting the choice of films to produce so as to achieve success in the new business venture.



Business Understanding

Microsoft would like to venture into the movie industry by opening a movie studio. The success of the business will depend on how they will approach the market through delivery of quality products. Therefore there need be insights on how to achieve the needed quality so as to identify the type of films to create.

We will conduct an exploratory data analysis so as to explore what types of films are currently doing the best at the box office. We will then translate our findings into actionable insights that the head of Microsoft's new movie studio can use to help decide what type of films to create.

Data understanding.

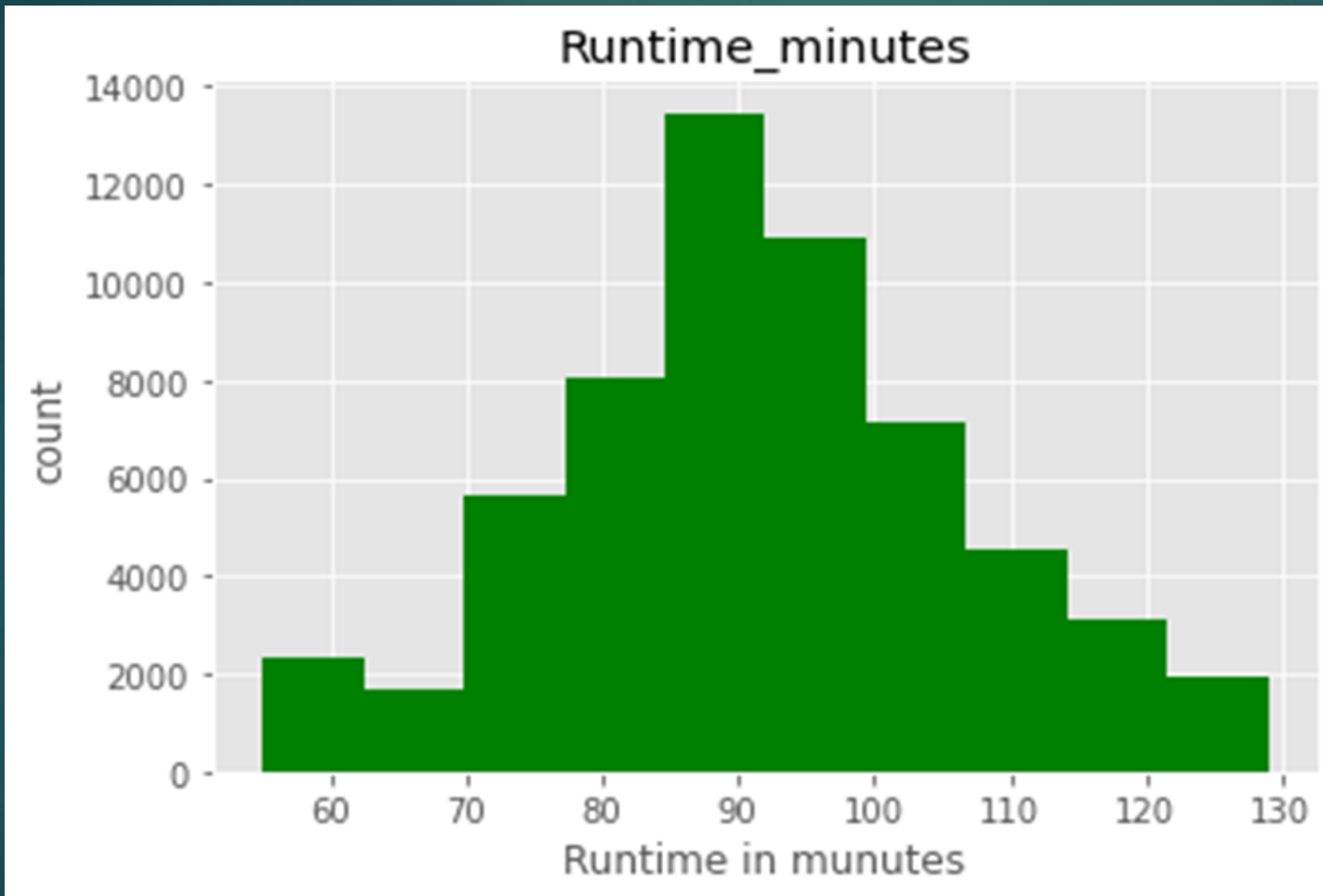
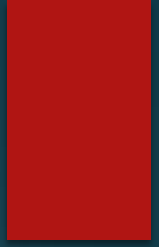
We have been provided with various data sets for the study. This will give us an overview of the data used to generate our insights.

1. The first data set we will use is the im.db database . It will give us data that relates to the movie genres and ratings from the tables movie basics and movie ratings. We will use the persons table and the directors to generate insight about director performance as well as writer performances.
2. The second data set is the movie budget csv file that will help us visualize revenue generated both from the local market as well as the global market. We used it to also analyze the budget allocated for the movies. This will help us analyze the total profits generated by the movies.

Objectives

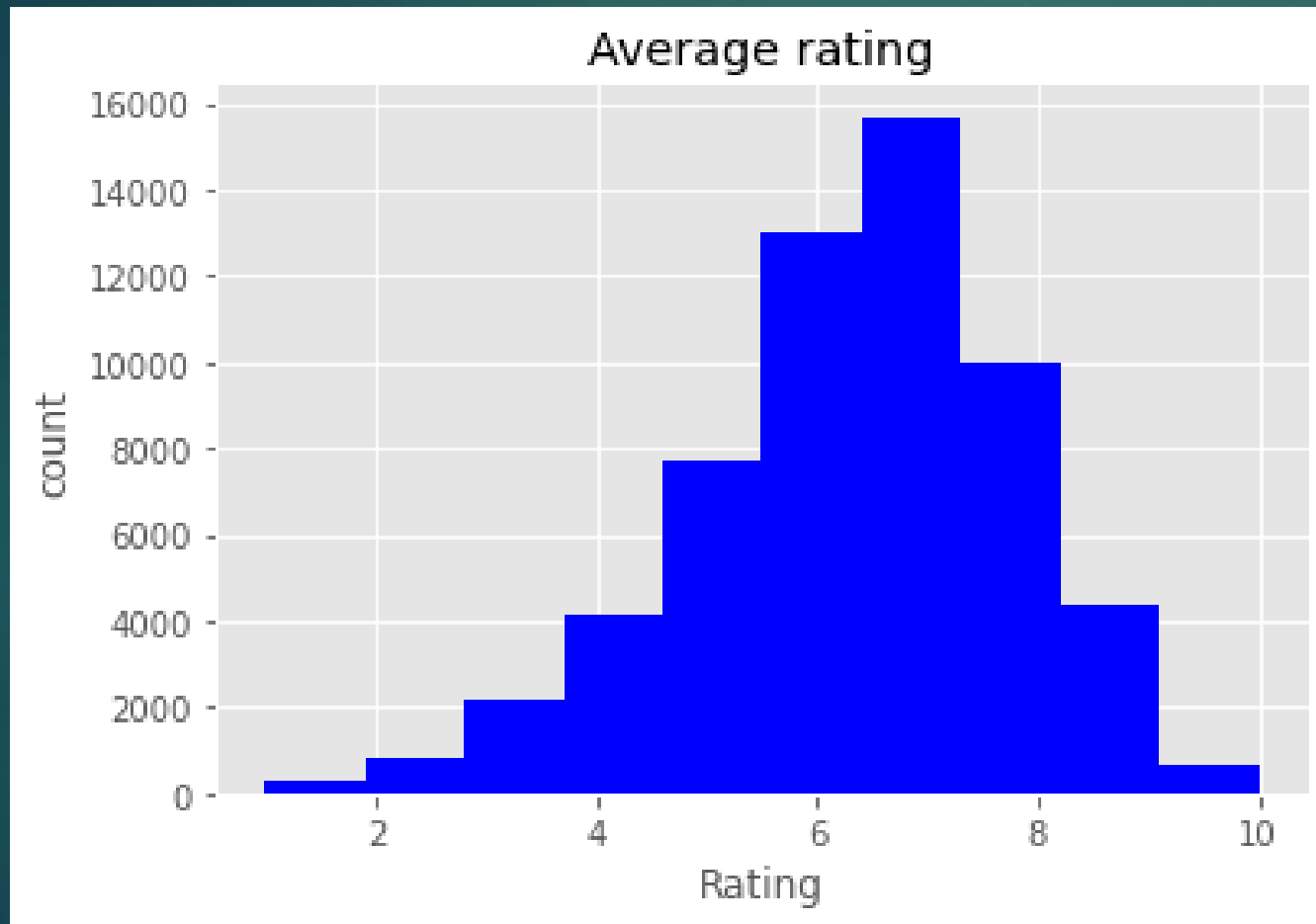
- ▶ Identifying the effect of runtime minutes on movie average rating.
- ▶ What are the top performing movie genres according to genres average ratings?
- ▶ What is the effect of directors and writers selection towards movie success.
- ▶ What is the correlation between budget and revenue generated for a given movie?

Run time minutes analysis



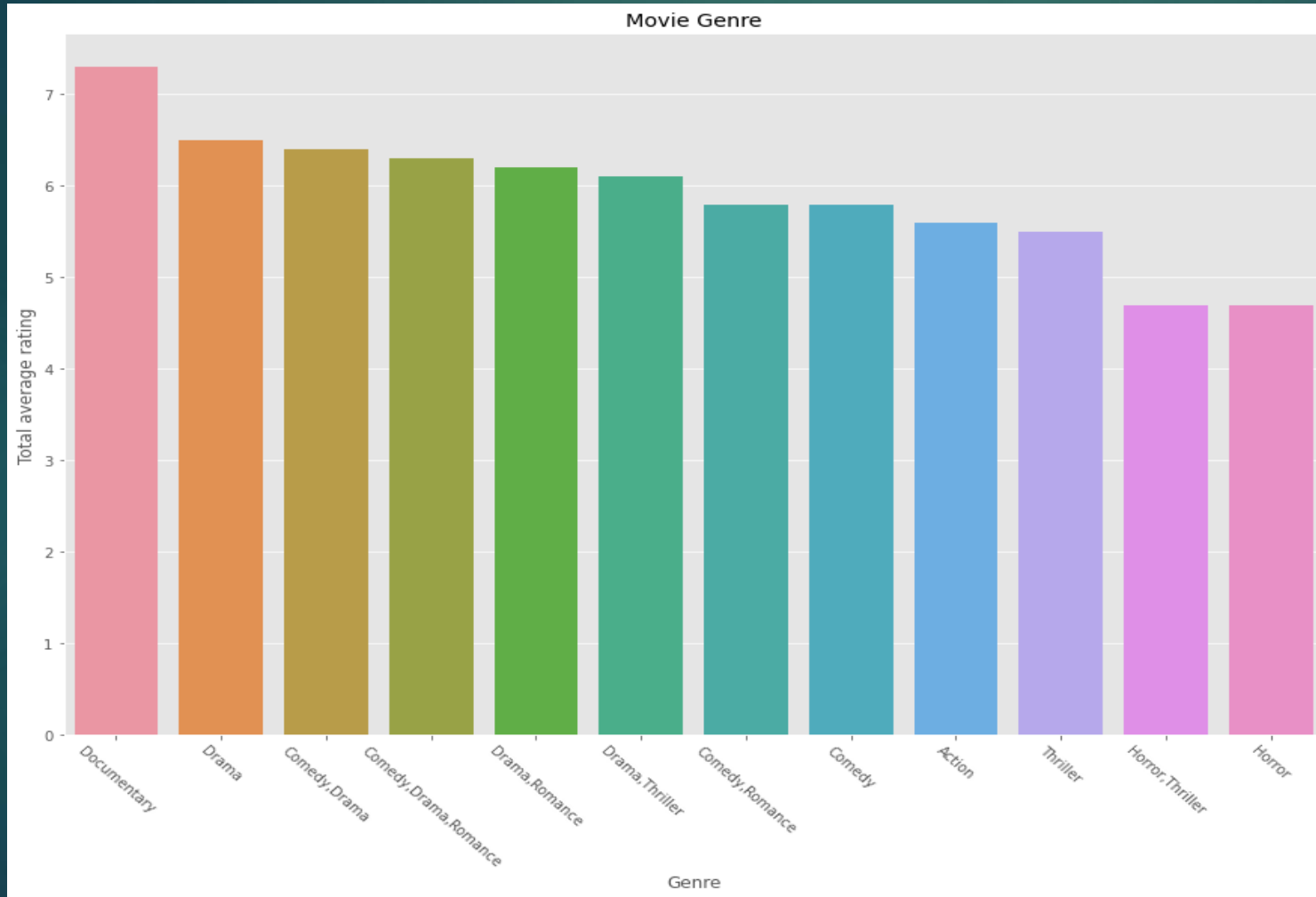
- This is a visualization of the number of movies produced within a given runtime. Most movies are produced within 80 – 100 minutes

Average rating analysis



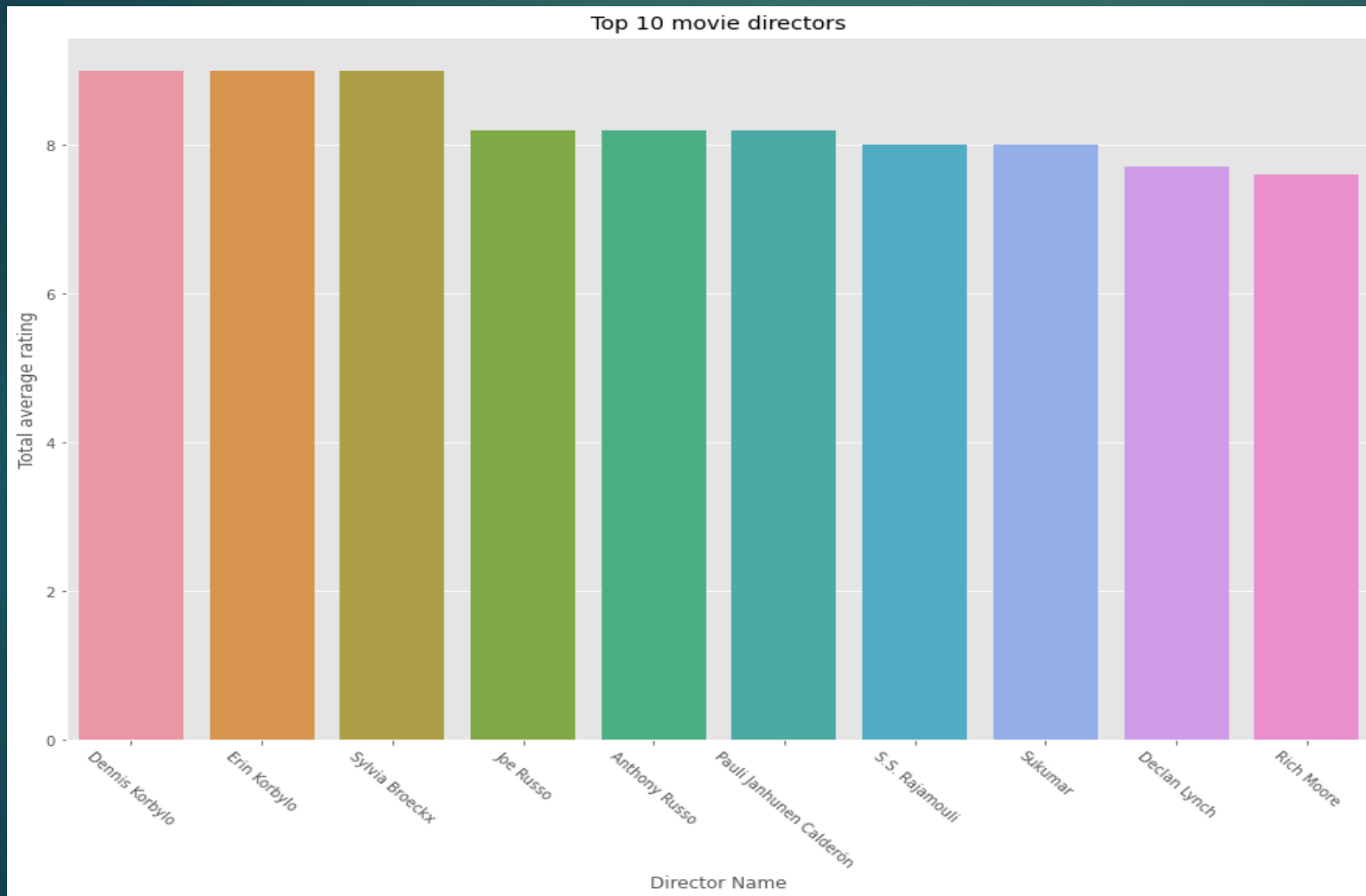
- ▶ According to the data provides most movies get a rating of 5-8.

Top performing movie Genres



- ▶ The top performing genres according to average rating. The best genres to produce happen to be having a total average rating of 4.5 and above.

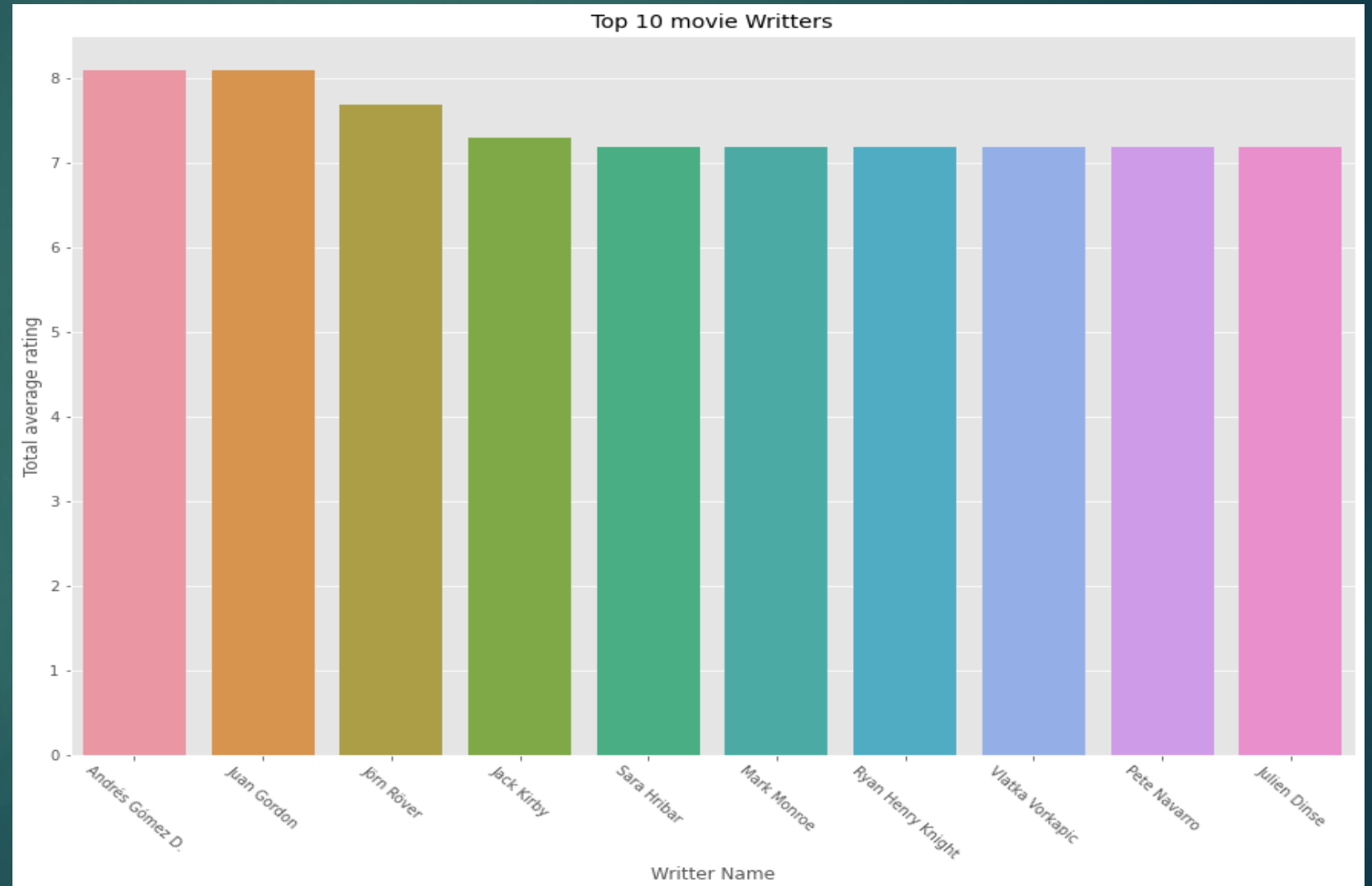
Movie directors analysis.



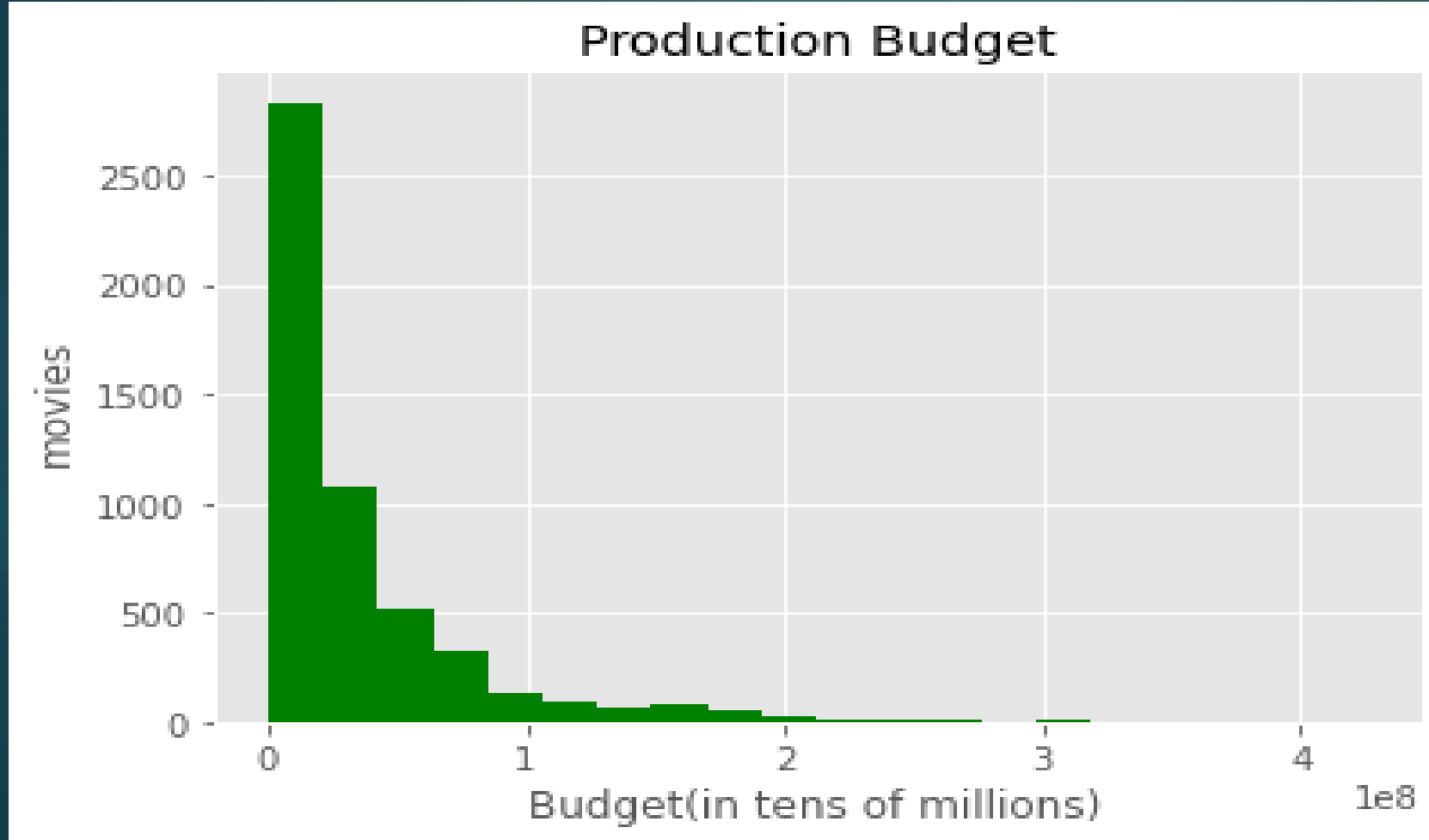
- ▶ This is a visualization of the best 10 directors together with their average ratings in the industry having directed a minimum of 20 films in their careers.

Movie writers analysis

- ▶ This visualization identifies the 10 best writers in the industry with a career of more than 20 successful films.
- ▶ The best movie writers in the movie industry have a cumulative average rating of 7 and above.

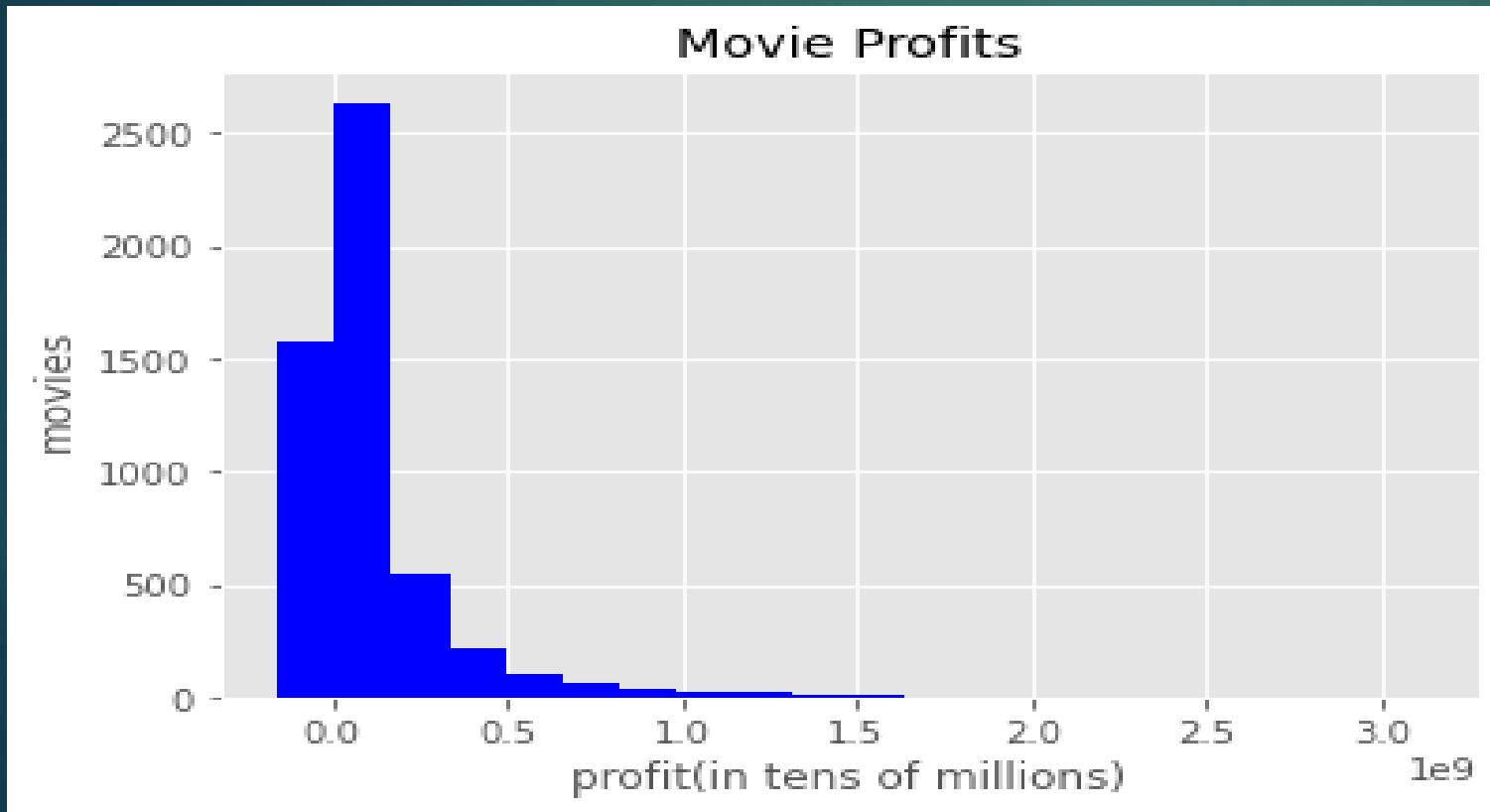


Average Movie budgets analysis.



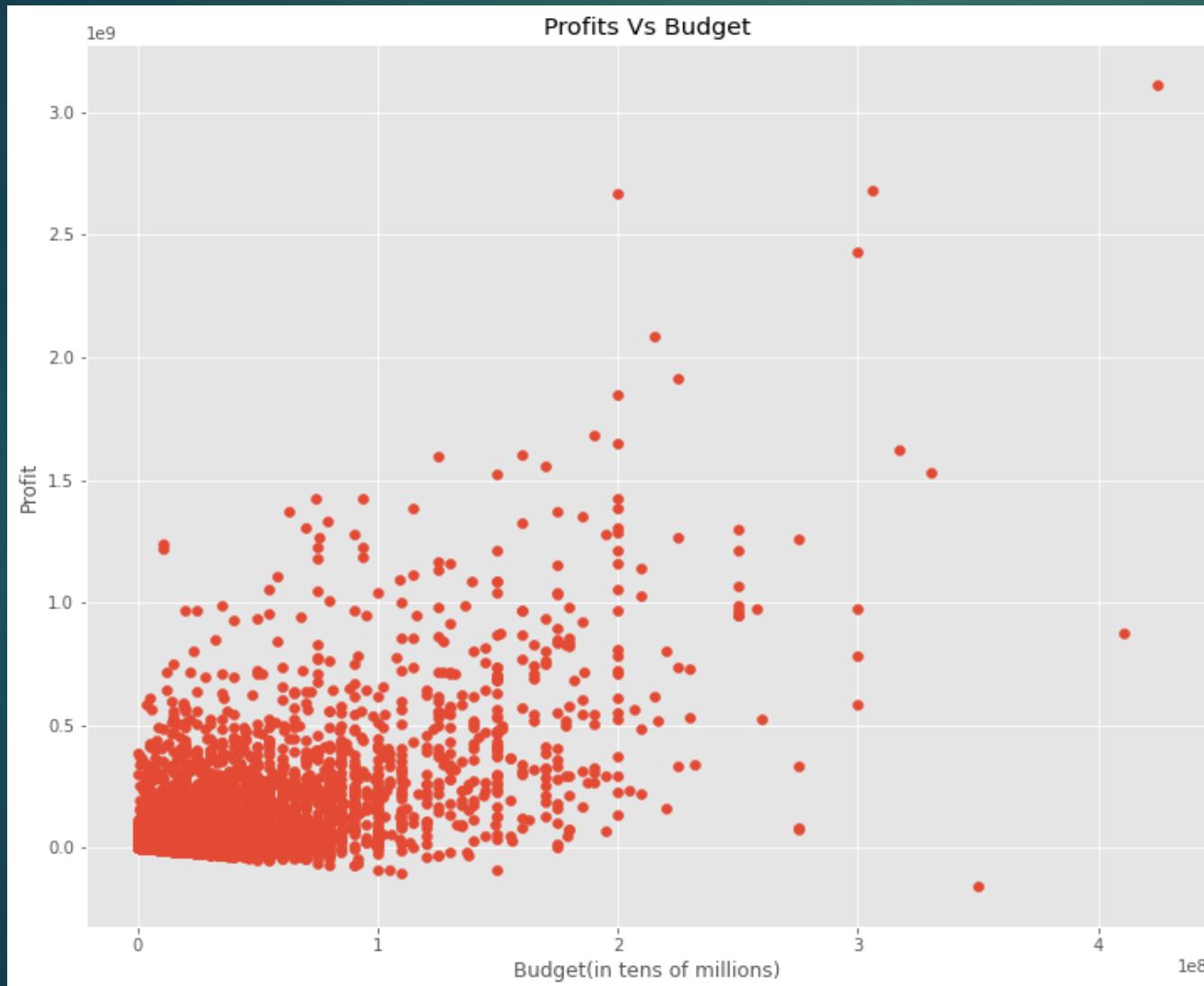
- This visualization shows that most movies are produced under a relatively low budget.

Average movie profits analysis.



This visualization shows that the profits generated are also relatively low. This is expected because of the low budgets as well.

Correlation between profits and budget



- ▶ This scatter plot identifies a strong positive correlation between profit and budget signifying that the higher the budget the higher the profits.
- ▶ Therefore allocating maximum resources generate maximum profits.

Findings

- ▶ Runtime is inversely related to movie average ratings. Long movies have lower ratings compared to short movies.
- ▶ Different genres perform differently in the movie industry with regard to the ratings.
 - ▶ Documentary(7.3)
 - ▶ Drama(6.5)
 - ▶ Comedy,Drama(6.4)
 - ▶ Comedy,Drama,Romance(6.3)
 - ▶ Drama,Romance(6.2)
 - ▶ Drama,Thriller(6.1)
 - ▶ Comedy,Romance(5.8)
 - ▶ Comedy (5.8)
 - ▶ Action(5.6)
 - ▶ Thriller(5.5)
 - ▶ Horror,Thriller(4.7)
 - ▶ Horror(4.7)
- ▶ Different writers and directors perform differently based on factors such as experience and expertise in the industry by looking at their average rating movies performance.
- ▶ Movie budget and revenue generation are directly related where high budget generates more revenue.

Recommendations

1. The studio should aim at producing movies with relatively short runtime as it is inversely related to average rating. This is because the movies with short runtimes ended up having more average ratings. This will entail avoiding unnecessary scenes and being brief and on point with the shooting.
2. The studio should aim at producing the top preferred genres that have a high average rating which means they will generate high profits compared to the relatively unwatched genres.
3. The studio should aim at using the top directors and writers according to experience in the industry according to their average ratings as they will enable the studio to produce top tier movie quality that will in turn generate better returns for the business.
4. The studio should aim at assigning the required budget and not under funding their movie projects since from the data movies with high budgets generated higher revenue. This entails using the best technology and tools possible to get the best quality.



NEXT STEPS

1. Remove outliers.
2. Focus on the most important aspects.
3. Use more visualisations.

THANK YOU

Questions

- ▶
- ▶
- ▶
- ▶

CONTACT INFO

- ▶ JIMCOLLINS WAMAE
- ▶ Email - jimwamaef12@gmail.com
- ▶ LinkedIn – jim wamae
- ▶ Phone - +254-722614538.