DATA ANALYSIS PHASE 1 PRESENTATION SLIDES

Student Name: Frank Oyugi

Introduction

- In this project 5 datasets containing various information about films are analyzed,explored,cleaned and conclusions are drawn so as to make accurate recommendations to microsoft,a company trying to get into the film-making industry.

Data Overview

Some of the information contained in these datasets includes:

- ➢ Movie title
- ➢ Rating
- ➢ Genre
- ➢ Runtime
- ➢ Production budget
- ➢ Domestic Gross

Methodology

The following analytical methods and techniques were used to analyze this data:

➢ Pandas:A python library used for data manipulation.

➢ Numpy:A python library used for numerical operations.

➢ Matplotlib:A plotting library in python.

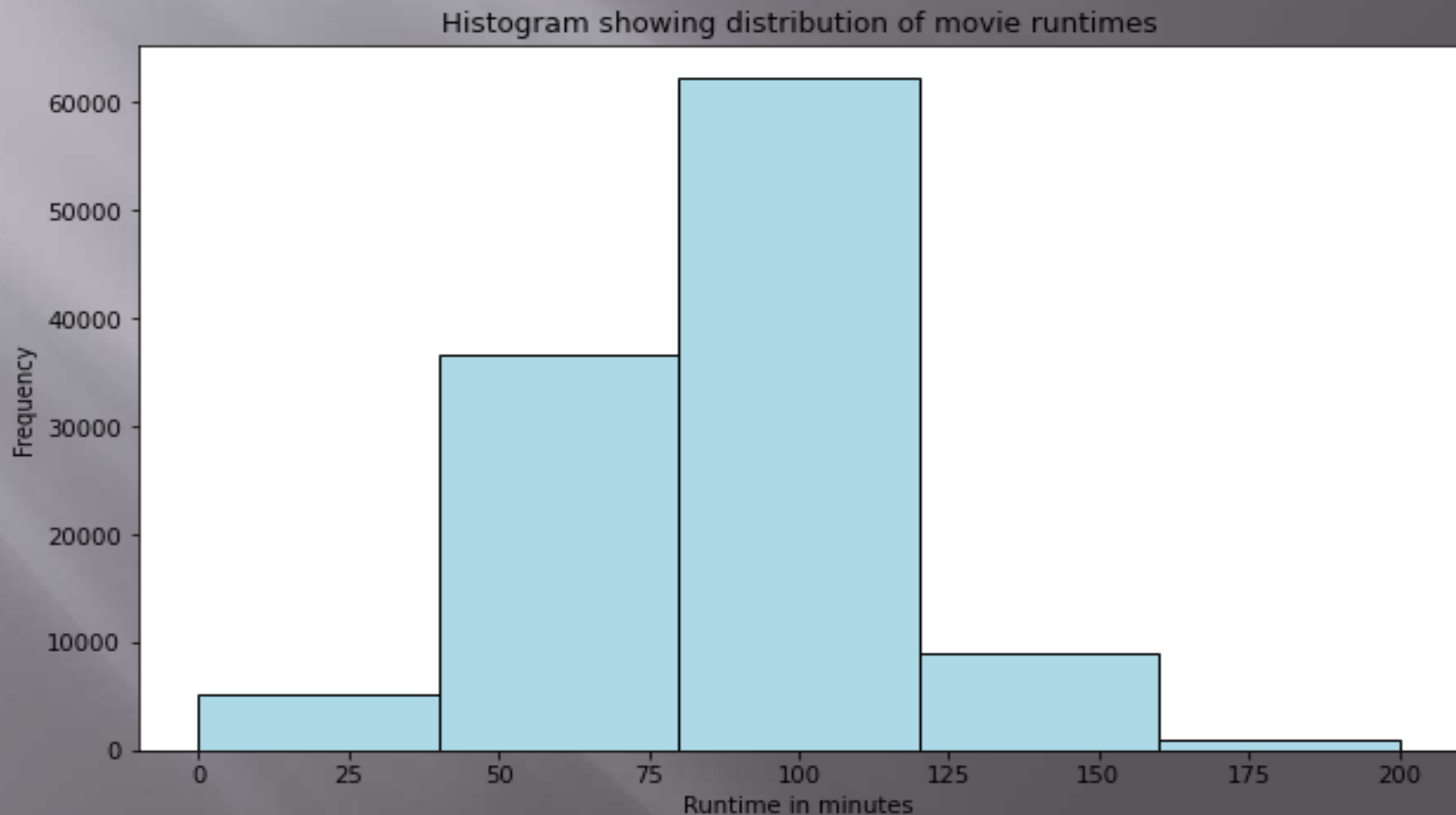➢ Seaborn:A data visualization library based on matplotlib.

Visualizations

In this project some of the visualizations used include:

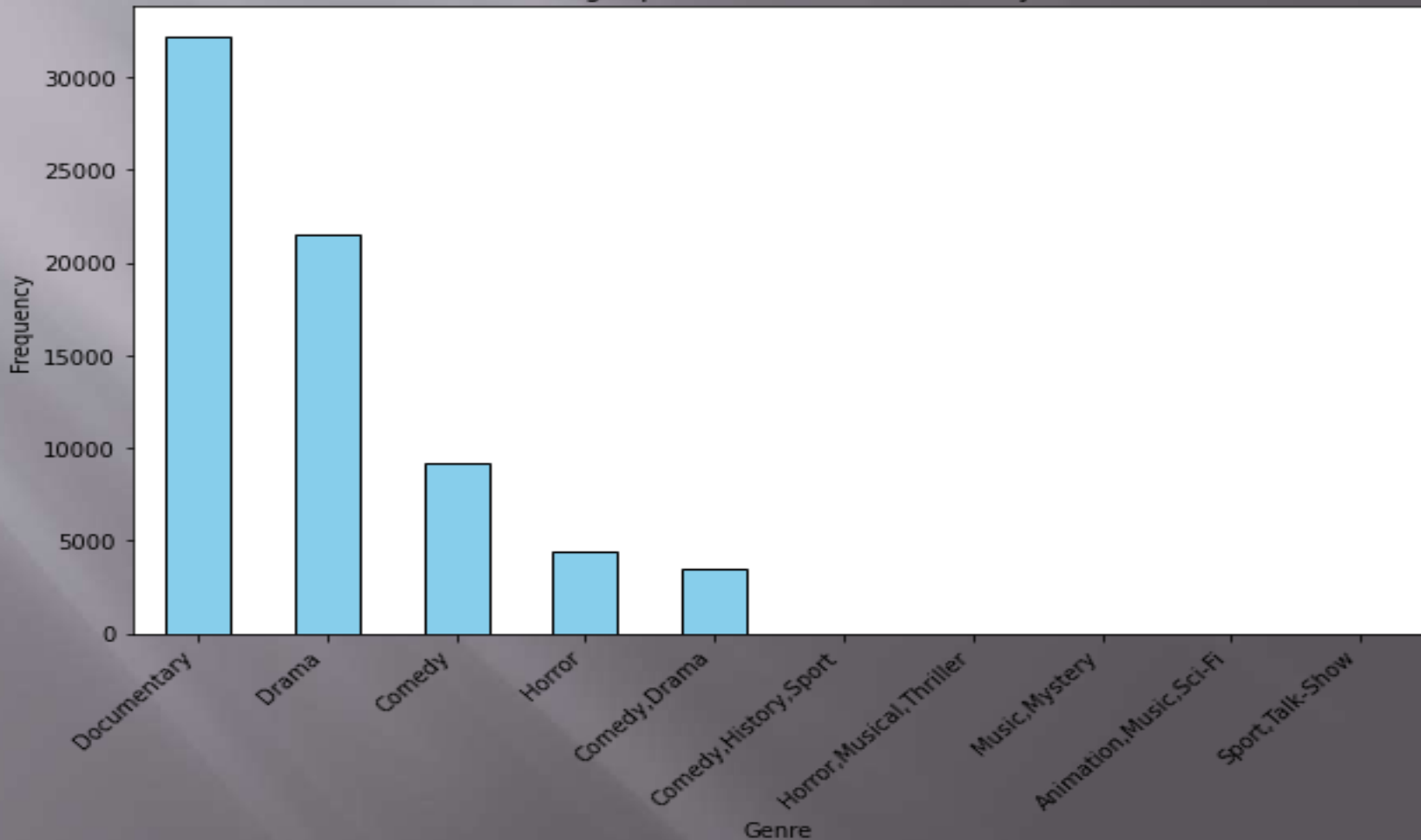➤ Heatmaps:Show correlation between two variables.

Correlation Heatmap: Domestic Gross vs. Worldwide Gross
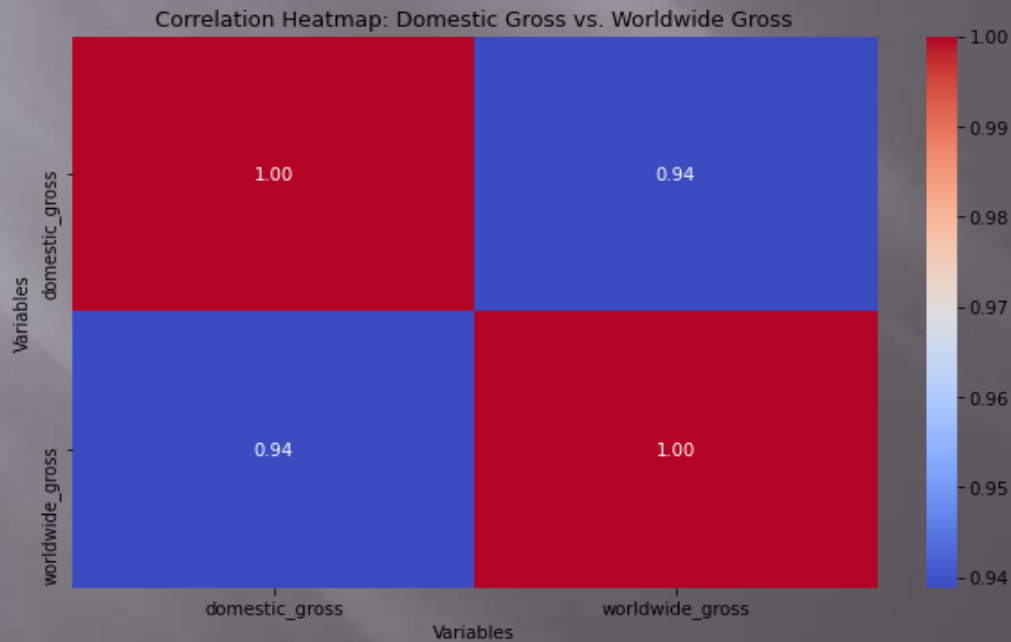
Visualizations

➢ Histograms:Show the frequenc distribution of a numerical variable.

Histogram showing distribution of movie runtimes

Visualizations

➢ Bar Chart:Used to compare two or more categorical variables.

Bar Chart Showing Top 5 and Bottom 5 Genres by Occurrences

Recommendations

The recommendations made from the project are as follows:

- Invest heavily in the production budget.The heat maps show that their is a positive correlation between production budget and domestic gross.

- Focus on making documentaries and dramas.The histogram above shows that these two genres are the most popular ones.

Recommendations

- Hire 'Omar Pasha' as your writer and director.He has contributed to the highest number of films in this dataset.

- Avoid hiring 'Raphaelle Ayach' as your writer and director.He has contributed to the lowest number of films in this dataset.

- make movies which are between 90 and 110 minutes long.The average runtime is 100 minutes.

Recommendations

- Do not make musicals and fantasies.They are the least popular genres.
- Target movie ratings of 6.5 or higher.6.5 is the average movie rating.