MOVIE DATA ANALYSIS

Analysis using IMDB dataset

Summary

This project will use Exploratory Data Analysis (EDA) to investigate the trends in the video production sector and help Microsoft decide on what films to invest on..

Outline

- Business Problem
- Data
- Methods
- Results
- Conclusions

Business Problem

Microsoft wants to invest in the movies and film industry but does not have data informed decisions on what movie genres to invest in.

Data

-Data was from IMDB movie data set

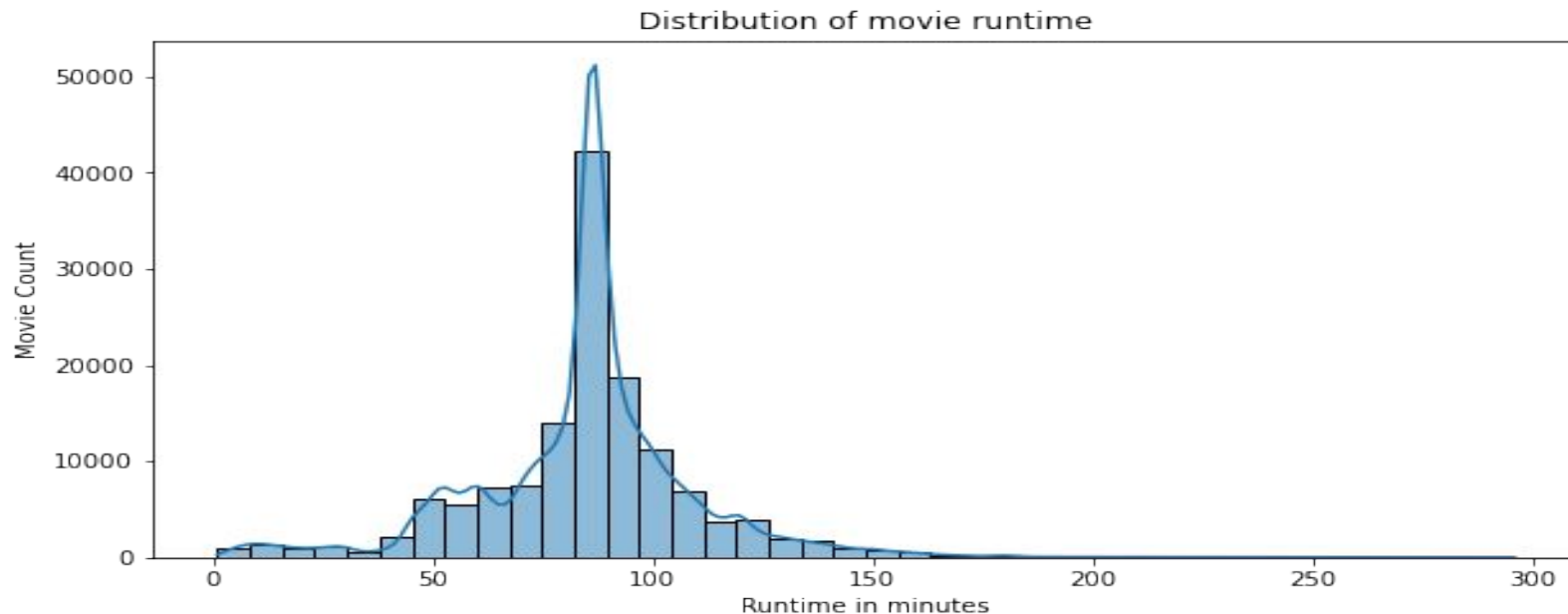-The dataset had 140736 rows and 6 columns

Methods

1.Data cleaning: Remove missing values and outliers detection

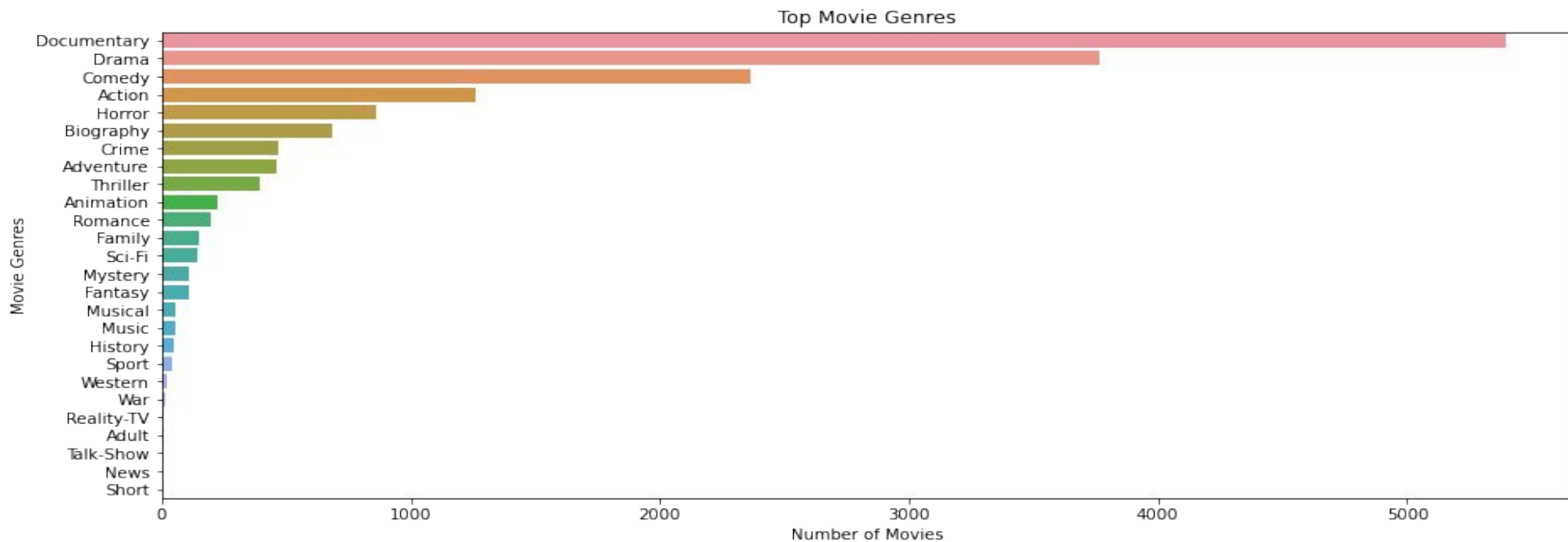2.Data Analysis: Grouped columns to see their relation , Trend analysis , visualization

Results

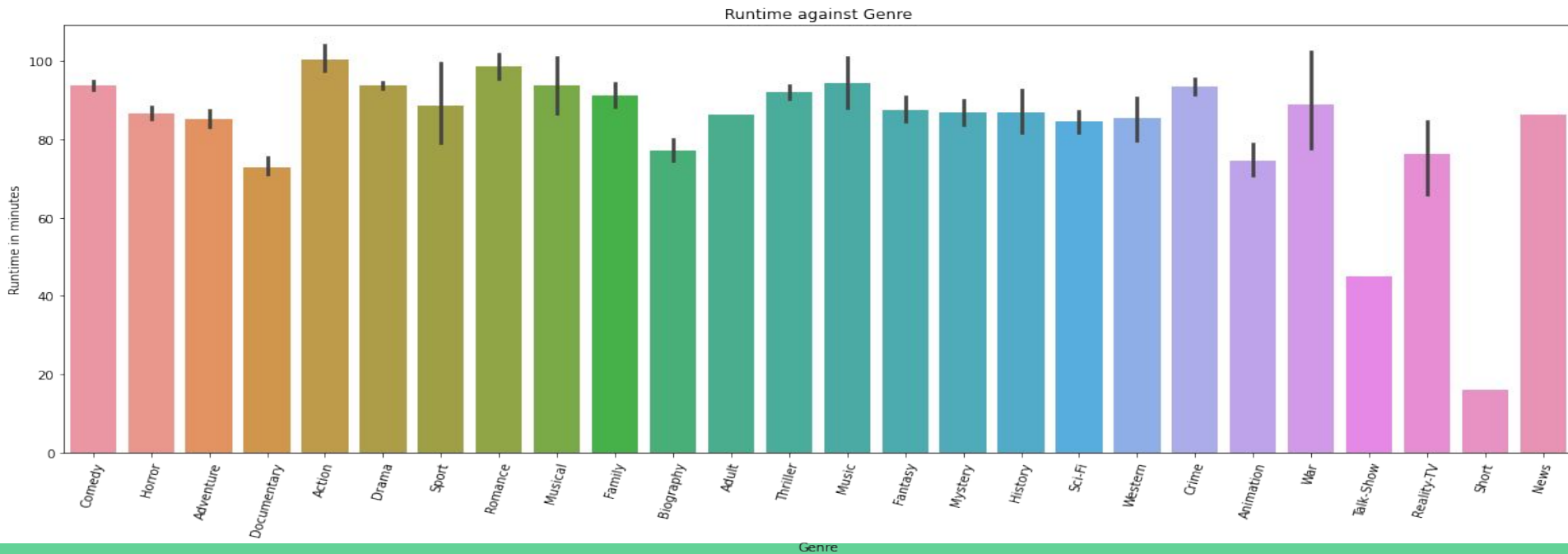**Key finding:** Films that are of 80 to 90 minutes range tend to be more successful

Distribution of movie runtime

Results

-Most common Genres in the industry are : Documentary, Drama, Comedy and action. So this are the genres to invest in.

Top Movie Genres

Result

- Documentary films tend to have a shorter runtime but are the most popular genre so investing in them might be profitable.

Runtime against Genre

Conclusions

I would recommend Microsoft to invest in Documentaries more, Drama, Comedy and action movies.

The company should also focus on producing movies with the lengths of 80 to 90 minutes long.

Thank You!

Email: jomosamson12@gmail.com
GitHub: @sam-trades
LinkedIn: linkedin.com/in/username/