PHASE 1 PROJECT

PRESENTED BY BONFACE MANGOKA

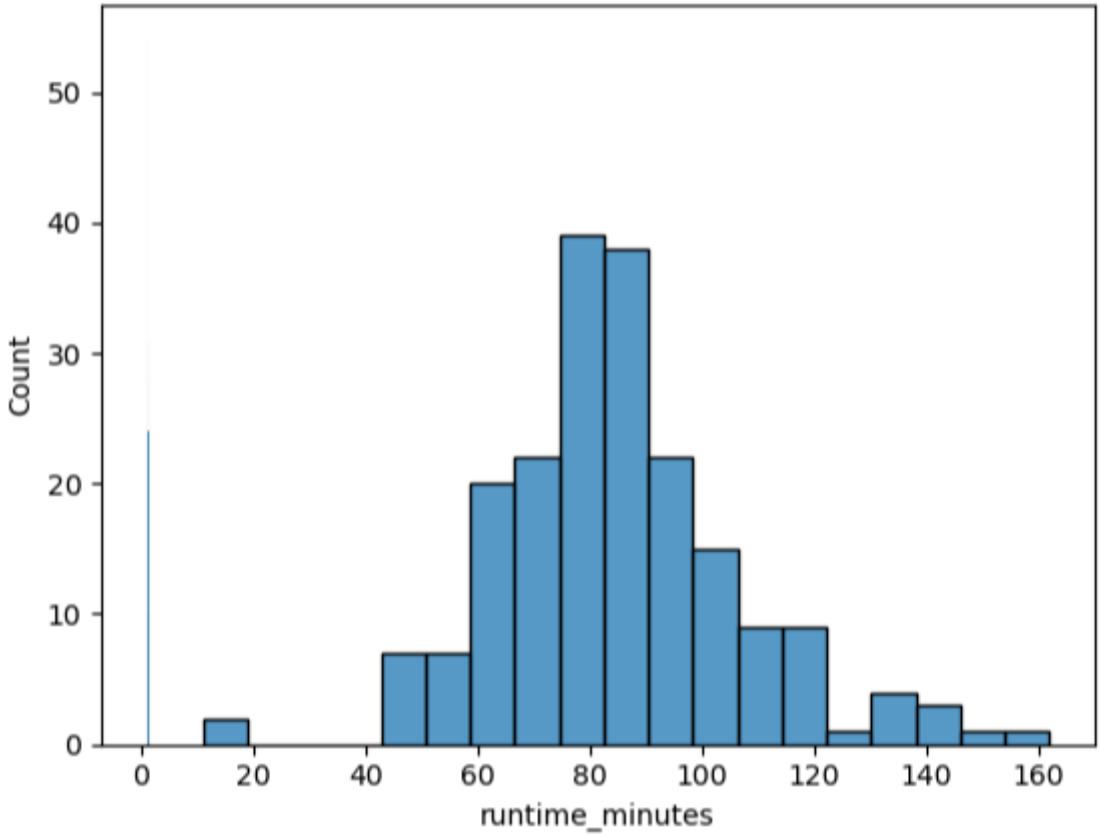
DATA ANALYSIS USING PANDAS LIBRARIES

Business Problem

- Microsoft sees all the big companies creating original video content and they want to get in on the fun. They have decided to create a new movie studio, but they don’t know anything about creating movies. i am charged with exploring what types of films are currently doing the best at the box office then translate those findings into actionable insights that the head of Microsoft's new movie studio can use to help decide what type of films to create.

The Data

- The data provided was pulled from several sources
- Box Office MojoLinks to an external site. IMDBLinks to an external site. Rotten TomatoesLinks to an external site. TheMovieDBLinks to an external site. The NumbersLinks to an external site.
- I used the below data for the project.
- imdb.title.basics imdb.title.ratings bom.movie_gross


Count
50
40
30
20
10
0
0
20
40
60
80
100
120
140
160
runtime_minutes

Average runtime has highest rating

original_title	start_year	runtime_minu
Freeing Bernie Baran	2010	1
Revolution Food	2015	
Hercule contre Hermès	2012	
Requiem voor een Boom	2016	
Fly High: Story of the Disc Dog	2019	
Dedicated Life: Phoebe Brand Beyond the Group	2015	
I Was Born Yesterday!	2015	
Exteriores: Mulheres Brasileiras na Diplomacia	2018	
The Dark Knight: The Ballad of the N Word	2018	1
Atlas Mountain: Barbary Macaques - Childcaring...	2015	

Documentaries have high rating

Conclusion

- Data visualized was sample data

Recomendations

From the data analyzed I would recommend;

- Movies with average minutes because it has the highest ratings.
- Documentary movies since they have the highest ratings.
- Drama movies mostly have poor ratings.