Halah Salem Banuqaytah
Project 1 (Proposal)

Recommendations for Cinematic Production Companies

❖ Background and Questions:

The movie industry that includes showing the movies only in cinema is facing a tough time, due to the massive impact caused by the new way of delivering movies to the public by streaming services. Streaming platforms such as Netflix and apple tv are starting to dominate the field with movies from their production and provide the audience with many features. That put the big cinematic production companies in a difficult situation for losing their audience to the new direction of movies streaming, which will eventually lead to losing more profits.

This project aims to provide the production companies with recommendations to produce movies to restore the audience and make more profit.

❖ We are going to answer these questions:

Q1 - Can the actor or the director's names affect the movie revenue?

Q2- Do action movies make more profits than any other genre?

Q3- Does a high budget mean more profit?

Q4- Predicting the movie profit depending by the Budget?

❖ Data description:

This dataset includes more than 7500 movies from 1980 to 2020. The dataset has 15 feature which was extracted from IMDb website, and it is available in Kaggle under the name (Movie Industry) For each movie we have the following features:

1- Name (movie official name)
2- Rating (The rating of the audience)
3- Genre (The type of the movie – only the main genre for each movie)
4- Year (year that movie was released)
5- Released (full releasing history-MM-DD-YYYY)
6- Score (movie score in IMDb)
7- Votes (votes for the score in IMDb)
8- Director (The movie director)
9- Writer (the movie writer)
10- Star (The main actor or actress of the movie)
11- Country (The country of the movie)

12- Budget (The movie's budget)
13- Gross (The profits of the movie)
14- Company (Production company of the movie)
15- Runtime (The length of the movie)

❖ Tools:

To answer the previous questions, we will use:

- ✓ Numpay
- ✓ Pandas
- ✓ Matplot
- ✓ Seaborn
- ✓ Sklearn
- ✓ LinearRegression Model

The tools will be more as we continue farther with the analysis.

❖ Methodology:

1- I will use the EDA part to answer some of the questions that don't need a model to solve

2- I am going to use the LinearRegression Model to predict the profit of a movie depending on the budget since I found that there's a high correlation between them. For now, I just create the training model and get some errors that I will start to work on to solve them.