



Data Visualisation with Matplotlib Project

Project Deliverable

- Your deliverable will be a notebook with your solution.

Instructions

Background Information

Mima is a startup film company that aims to enter the film industry.

As a data scientist consultant working for the startup, you're required to perform an analysis and provide recommendations on the kinds of movies that the startup should create in order to have a profitable business.

After sourcing for an existing movie dataset and using your existing knowledge of the Matplotlib library, you start work on your analysis which entails answering the following questions:

- Do movies with a higher budget end up being popular?
- Does the length of the movie affect the vote count and popularity?
- Does higher popularity mean higher profits?
- What features are associated with the top 10 revenue movies?
- Which genres are most popular from year to year?

You can use the following guiding notebook [[Link](#)] to get started working on your analysis.

Dataset

- Dataset CSV (URL): <http://bit.ly/MoviesDS>
- This dataset contains information about 10,000 movies collected from The Movie Database (TMDb).
- Columns like 'cast' and 'genres', contain multiple values separated by pipe (|) characters.

- You can leave out the characters in the 'cast' column.
- The final two columns ending with “_adj” show the budget and revenue of the associated movie in terms of 2010 dollars, accounting for inflation over time.

Source

<https://www.themoviedb.org/>