

MOVIE STUDIO PRODUCTION PROJECT

1. BUSINESS UNDERSTANDING

Link to Jupyter Notebook :

https://flatiron.illumidesk.com/user/johnmarkchege/notebooks/dsc-csv/Untitled.ipynb?kernel_name=python3

Link to Github : <https://github.com/JohnmarkChege>

1.1 UNDERSTANDING THE PROBLEM

I have assumed the role of a data scientist under the Microsoft Corporation. The task is to explore the type of films doing best at the box office using the exploratory data analysis. Databases related to other top movie studios such as IMDB and Box Office Mojo are to be analysed with attention to key details such as Movie Profits, Movie Genre, Movie Actors and lastly Movie Release Dates. I will then translate these findings into actionable insights that the head of Microsoft's new movie studio can use to help decide what type of films to create.

2. DATA UNDERSTANDING

DATA COLLECTION

I have sourced two datasets from GitHub that are from both **Box Office Mojo** and **IMDB**.