

PROJECT REPORT

Dataset used for the investigation: TMDb movie data (consisting of information about more than 10k movies)

Questions proposed: Varied questions are proposed, accordingly to the different aspects/ information given in the dataset. These are the questions that are being analyzed for the investigation. These include: finding the maximum, and minimum profits of a movie, and what characteristics are important to make a profitable movie? This is just a glimpse of what I've done in the project.

Description about Investigation: Starting from the questioning phase, what I basically have done is simply followed the data analysis process. I went through all the steps, perhaps, not in the same order. I did it as per I needed. For example, even towards the end, I needed to clean some data, so I cleaned it there and then. I cleaned the data, did some data wrangling to make the data consistent, and then did the exploratory data analysis, leading to making conclusions.

Documentation of any data wrangling you did, Summary statistics and plots communicating your final results :

All the phases have been done in much detail in the jupyter notebook.