

PHASE 1 PROJECT SUBMISSION

Student name : Robert Agwingi

Student pace : Part-time

Review date : 9.11.2021

Instructor: Asha Deen

Blog post URL : <https://github.com/RobertAgwingi/Phase-1-Project--Moringa.git>

The imdb.title.basics.csv file in the zipped data folder contains movie titles with their start year, genre and runtime minutes.

From the findings, there are 146144 titles. 31739 have a runtime of 0. Therefore a percentage of 78.28% have watched the shows. The most watched shows are documentaries in genre. The mean runtime minutes is at 86.187247. A movie title(Logistics- a documentary) started in the year 2018 is the most watched with a record of 51420.

The bom.movie_gross.csv file in the zipped data folder contains movie titles with their studios, domestic gross, foreign gross and the year.

From the findings , there are 3387 entries. Only 5 of them don't have a studio. The domestic gross mean is at 2.8745. The highest domestic gross was 9.3670 in the year 2015.

The imdb.title.ratings.csv file in the zipped data folder contains indices with the average rating and number of votes. There were 73856 entries with no non-null count. The highest average rating is 10.0 and highest number of votes of 1841066. Index 63498 had the highest number of votes with an average rating of 8.8. The average rating mean is at 6.3327 and the mean number of votes is at 3.5237.

From the graphical representation the average rating was high between rating 6-8.