**Dataset Analyzed**

IMDb Movie Database

**Data Cleaning Performed**

1. On running summary statistics, observations were some of the columns would be useful in analysis, while some of them will not be contributing in data exploration. Hence, the columns - *imdb\_id, original\_title, cast, homepage, director, tagline, genres, keywords, overview* and *production\_companies* were dropped from the dataset
2. Columns *budget*, *revenue* and *runtime* reported lot of zero values. A movie cannot report budget, revenue or runtime as zero. The zero values will kill the visualizations in the histogram, hence the zero values are replaced with null values

**Questions Answered**