

Test

Daniel Caley

9/18/2021

- Boxplot Rating
 - Breaking IMDB score by rating shows that PG, PG-13, and R all have around the same average and are distributed similarly.
 - The tails and outliers are different for PG, PG-13, and R, but not too different.
 - When looking at the Rating score for G the average and median is different than the remaining cohort of data.
- Boxplot Duration Bins
 - When looking at the boxplot the duration bin between 1.5 and 2.0 hours appear to have somewhat normally distributed data.
 - The remaining duration bins either become smaller and/or are not as evenly distributed.
- Histogram International
 - Breaking the data out as International or domestic, there is quite more data in the domestic category.
 - Both have a sizeable amount of data and are left skewed.
 - Domestic here means that the film exclusively released in the United States and no other country.
- Overall
 - When breaking data into multiple dimensions the information can be less actionable due to not having enough data.
 - The data here, as different dimensions are applied, shows that in most cases the data is not unevenly distributed.
 - A caution will be as all of these dimension added into the model will play in the ability to have predictive capabilities.