Intercom: Product Analyst Screener 2018

1) Products

If you worked at Google, which of their products would you like to work on most? And why?

2) Data, Data, Data

If you had access to any dataset from any company or organization in the world…

1. What would the company / dataset be?
2. What would you like to explore within the dataset?
3. Why this dataset / this exploration?

3) Technical Analytics Task

We would like you to explore and analyze an IMDB Movie Dataset with almost 60,000 movie titles, genres and ratings, answering the questions outlined below. We'd like you to submit your answers in a word or PDF file, along with your workings (e.g. Excel file, iPython Notebook, R Script, etc).

Short-Answer Questions

1. What are the 3 most common ratings (1-10) for movies in the list? Round each rating up or down to the nearest whole number.
2. Are there more R-Rated movies or PG-13 movies in this list?
3. Are there more Drama, Comedy, or Romance films in this list?
4. How many movies are described as both Action and Comedy (but no other genre)?
5. What is the average, median, 25th percentile and 75th percentile of ratings?
6. Which of the following has the strongest correlation coefficient (r)? Rating vs. Votes, Length vs. Rating, or Year vs. Rating. What is that correlation?
7. If you plot Length vs. Rating, and you look for the most obvious outlier, what is the name of that movie?

Open-Ended Questions

Please limit your answer to 2 pages including charts for each question below.

1. Imagine you are an executive producer today and you are tasked with picking the genre of Startup Studios' next movie. Using the data in this file, which genre would you recommend? Why?
2. Please tell us something interesting about this data. You can present it in any way you like.

Data Summary

Attached is one CSV file with the following fields:

- Title: Title of the movie
- Year: Year of release
- Budget: Total budget (if known) in US Dollars
- Length: Length in minutes
- Rating: Average IMDB user rating
- Votes: Number of IMDB users who rated this movie
- r1-r10: Distribution of votes for each rating, to midpoint of nearest decile: 0 = no votes, 4.5 = 1-9% votes, 14.5 = 11-19% of votes, etc. Due to rounding errors these may not sum to 100.
- Mpaa: MPAA rating
- Action, animation, comedy, drama, documentary, romance, short: Binary variables representing if movie was classified as belonging to that genre