Caroline Sudhakar

3/8/23

Crowdfunding Report

Question 1. Given the provided data, what are three conclusions that we can draw about crowdfunding campaigns?

Conclusion 1: Looking at all countries for the outcomes of each parent category, we can see that journalism is technically the most successful, simply because all of the projects were successful. The worst category for a project is games.

Conclusion 2: This conclusion is similar to conclusion 1, except instead of each parent category, we are looking at the sub categories. Additionally, we used country and parent category as the main filters. Audio and world music are the most successful projects, with a success rate of 100%. The worst one is science fiction.

Conclusion 3: This conclusion is a bit different, we have the filters as parent category and years. Once again, we are looking at the outcomes, but this time we use the months to see the success rate. The month with the highest success rate is June, with a success rate of 63.9%. The month that failed the most is August.

Question 2. What are some limitations of this dataset?

The biggest limitation to this dataset is the countries column. There are 195 countries around the globe, but the dataset only has 7 countries in it. Having more countries could show a difference in the data. Additionally, there are only 1000 entries in the data. This could be classified as a limitation since there could be more, but for some people it would be enough. When it comes to the concept of total entries, it’s up to the user to decide how many there should be.

Question 3. What are some other possible tables and/or graphs that we could create, and what additional value would they provide?

A table that can be created could be which country seems to have the most backers. Having this table could indicate that location also plays a factor into whether a project would be successful or not.

(Below questions are a part of the Statistical Analysis)

Question 4. Use your data to determine whether the mean or the median better summarizes the data.

The mean better summarizes the data due to the fact that we can determine how many backers could be required for a successful project. The median wouldn’t make any sense to use since it’s the middle value of being successful.

Question 5. Use your data to determine if there is more variability with successful or unsuccessful campaigns. Does this make sense? Why or why not?

There is more variability with the successful campaigns. This does make sense since the successful campaigns have more data than the unsuccessful one.