Homework 1 Kickstarter analysis

Questions:

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

2. What are some limitations of this dataset?

3. What are some other possible tables and/or graphs that we could create?

Answers:

1. Three conclusions can be drawn from the data analysis of the Kickstarter data. The first conclusion the most popular campaign category to be created is theater. The theater category had most successful campaigns but also had the most failed campaigns. The least commonly created category was journalism. All of the campaigns in journalism were cancelled. The most common sub-category created was plays. This correlated to the most successful campaigns being of sub-category plays, and also the most failed campaigns also being sub-category plays.
2. Some of the limitations I see in this data is there is no timeline on how the funding of each campaign progressed. I would like to see how much money was raised over time in each campaign as it progressed to see if the there is a steady rise in funding over time or if it comes in at certain points like the beginning and end of the campaign. This data will be useful to determine if a campaign looks like it is going to progress to a success or remain a failure. Another limitation in the data is that you only see the final amount raised as a lump sum and it is not broken down by individual backers or groups of backers. This data would be useful to see if campaigns do better with a large number of small backers or if large contributions help to push a campaign towards success. Another limitation of the data is there is no reward value shown for each of the campaigns. Generally, a campaign would have some way to reward its backers for their contributions. The value of these contributions would most likely play a role in if someone was willing to donate to specific campaign.
3. The data could also be put into a table and graph to separate out by country to determine which type of campaigns were most and least successful in the different countries Kickstarter campaigns are created in. The next analysis would be to compare campaign status with either staff and/or spotlight to see the trend in data. This would indicate if the campaign has a greater chance of success when it is either a spotlight or staff pick. This would be useful to know if you created a campaign on the chances if it will be successful or not.

Bonus:

For this data the mean is more useful in summarizing the data because of the large range between the campaign minimum and maximums. The median number is less than a majority of the campaign numbers.

Based on the data analysis the variance is greater with the successful campaigns compared to the unsuccessful campaigns. This makes sense because in the data analysis completed there were more successful campaigns compared to unsuccessful campaigns thus more opportunities for a variance. Also, when a campaign is successful it is generally going to have more backers supporting in order to raise the needed funds to achieve success. While a failed campaign will generally lack enough supporters to raise the fund necessary. Therefore, campaigns that are successful will have more backers and thus a chance of more variance.