Matthew Groh

Module 1 Report

**Crowdfunding Goal Analysis**

Given the provided data, one conclusion that we can draw is that crowdfunding campaigns are more successful than not in every parent category except for games. A second conclusion is that there are far more campaigns in the sub-category of plays than in any other category, and because of this, the number of both successful and unsuccessful play campaigns is more than any other sub-category. Finally, a third conclusion we can draw is that most campaigns have a funding goal either between 5,000 and 10,000 or greater than 50,000, and the percentage of successful versus failed campaigns is much more even compared to nearly every other goal bracket.

One limitation of this dataset is that it does not have a currency standard. While most of the data is shown in USD, there is also data in CAD, DKK, GBP, and likely more. Since many currencies are represented without any conversion, some metrics can be misrepresented, such as the funding goals primarily featured in the fifth sheet.

A new graph that we could create is a comparison of the count of each outcome versus the country the campaign was in, with additional filters based on parent or sub-categories. This could provide data on where crowdfunding campaigns are most common, which categories are most prevalent in each country, and the likelihood of a successful campaign by country based on their outcome percentages.

**Statistical Analysis**

In this case, the median better summarizes the data. While there are not heavy outliers in either data set, there is not a particularly even distribution of data across all points. In both the successful and failed campaigns, there is a higher concentration of campaigns with less than the mean number of backers compared to those with more than the mean, and since the maximum of each category is much greater than their respective means, using the mean would skew the data more toward the maximum, which would not best represent the data.

According to the data, there is more variability in successful campaigns. This does make sense because between the two data sets, successful campaigns have a larger mean and maximum than failed campaigns. Additionally, the minimum number of backers in successful campaigns is very close to that of failed campaigns, and there are many more successful campaigns with backer counts closer to their maximum than failed campaigns do. All these factors contribute to successful campaigns having more variability.