Crowdfunding Report

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

A striking conclusion that I would take away from this data is that there seems to be an overwhelming amount of support for the theater industry, specifically plays. Our top 3 categories for funding came from Theatre, Film/Video, and Rock music. The entertainment platforms did the best job of acquiring funding. Another conclusion I would make is that during the summer months, we saw a large increase in our ability to acquire funds and towards august and september we see a very significant decrease in acquiring funds. There are a lot of conclusions that make sense in this data set given the time ranges. Something interesting that I noticed is that not only in 2020, but in 2019 I see that in the summer time there was not a lot of success in crowdfunding overall, and that in the spring of 2019, and the fall of 2019 there is a significant number of more successful campaigns.

2. What are some limitations of this dataset?

I think one of the largest limitations of this data set is the inability to acquire the quality of campaigns, events or fundraisers for crowdfunding. It would be interesting to see which events were attaining more funders, and other details such as what types of food, and what kinds of technology. Are the music events strictly concerts? Are there community events attached to the large portions of crowdfunding? I would also like to see which cities/states within the countries had the most success so that we could have a better understanding of the strengths/opportunities for crowdfunding in each area. Something I was thinking about when we are comparing other countries is inputting something involving a country's GOP so that we could get a better understanding of how the industries affect the economies within each country, and how the events truly compare to other countries. For example,It’s a little challenging to be able to understand that when we are comparing the dollar amount to EU, and within these data sets we don’t have a standard of what things these countries rely on the most as far as economy.

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

I think having individualized pie charts for successful, failed, and unsuccessful in order to compare all of the categories would be a bit more appealing to the eye. I would even go as far as creating pie charts or even bar charts and taking only the successful campaigns. This would be able to separate the unsuccessful. If I were helping a business in crowdfunding I could then show them the types of events that have been successful and the types that have not. I would even like to create separate tables for individual countries so that I could better cater to the specific country for an individual business.

**Statistical Summary:**

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

Considering the large variance with both data sets, the median calculations are a better representation of the data as a whole. In addition, the standard deviation for both data sets is quite large meaning there is an overwhelming number of data sets outside the average number of successful and unsuccessful campaigns.

2. 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 in the successful campaigns. This does make sense. When looking at the data in the graphs, you can see that the very successful campaigns are only in a few different categories. The most successful campaigns were in the theater category, and those campaigns were dramatically more successful than any other campaigns. When compared to the unsuccessful campaigns, it is visibly clear by looking at the graphs and numerical data, that they are much closer in numbers. Additionally, there is a significant number of more successful campaigns than unsuccessful, meaning it is more likely that the variance would be larger given that it is a larger set of data.