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

Conclusion 1: In the Category Pivot Table, the Category theater has the highest number of campaigns but you’re likely to have a successful campaign in music than in theater. According to the data There is a 33 percent chance that you’ll have a successful campaign in technology. Also, there is a low number of food campaigns being successful. So I don’t recommend starting a food Kickstarter.

Conclusion 2: In the Sub-Category Pivot Table, you’ll likely have a successful campaign in these sub-categories: classical music, documentary, electronic, hardware, metal, nonfiction, podcasts, pop, rock, shorts, table-top games, and television. You’ll likely have a failed campaign in jazz, drama, animation, children’s books, fiction, gadgets, mobile games, nature, people, places, restaurants, and video games.

Conclusion 3: There is a higher rate of successful

Kickstarter campaigns than failed campaigns throughout the years of 2009-2017. Successful campaigns peak in the month of May and it slowly decreases until September. I highly suggest starting a campaign from February through May, you’ll have a better chance having a successful Kickstarter. Successful campaigns drastically decrease in the month of December due to the assumption that backers may be doing Christmas shopping at that time.

1. What are some limitations of this dataset?

Some limitations that I can see from this dataset are the use of currency and the lack of information of backer donations. From the data, it is not clear if money from other countries is converted to US dollars and if so, it would be a good addition to add a table to see what the original price was in that specific campaign and country. Also, there is a lack of information of backer donations. The table doesn’t show how much each backer specifically donated as well as the time when backers donated in the duration of the campaign being live. This information will be helpful to further analyze this data.

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

A few other possible graphs that we could add to this dataset would be a stacked Histogram and a scatter line plot. A stacked Histogram would be helpful to see what campaigns failed, succeeded, and were live over a period of time. Another possible graph that could be useful would be a scatter line plot. This graph would be beneficial to see what the campaigns goal were on the y-axis, measured by when the campaign ended (year’s timeline). Also, the dots on the scatter plot could be color coordinated to see if the campaigns failed or not.

Bonus Statistical Analysis

1. In my data, I determined that the median summarized the data more meaningfully. Most of the number of backers in each campaign were close together rather than spread out. Also, there were outliers in the data, which impacted the calculation of the mean.
2. Finding whether successful campaigns or unsuccessful campaigns have more variability doesn’t make sense to me in this data because I can’t determine which category is more spread out due to the fact that both categories have a different amount of campaigns.