**Kickstarter Projects Analysis**

**Will Cutrone**

**1) Conclusions**

1) More than half of Kickstarter campaigns are successful, indicating that they are generally fairly successful ventures. From the data, 53.1% of them succeeded, 37.2% failed, 8.5% were canceled, and 1.2% are still live. Out of the live, ongoing campaigns only 16% have met a successful outcome (meaning the amount raised is greater than or equal to the goal). And only 20% of the live campaigns have reached 50% or more of their fundraising goal. Therefore, the prognosis for the current live campaigns is not looking very promising, as it seems that many of them will fall short of the mark. Given this, it seems that the number of successful Kickstarter campaigns in the long run will not be too much more than the 53.1% success figure we currently have.

2) In terms of category, Kickstarter campaigns related to the performing arts seem to be most successful. The categories of theater, music and film/video are the three most successful categories, respectively. The category of theater is particularly successful. Theater has the highest percentage of successful campaigns per category, as 60.2% of theater campaigns end up being successful. Breaking down the theater category, 76.5% of these campaigns are of the sub-category: plays. Plays are very successful as well in their own right, with a 65.1% success rate. Bottom line: if you are going to start a Kickstarter campaign for theater, there’s a pretty good chance it will succeed. This is the category I would feel most comfortable starting one in. The performing arts in general are not a bad idea either. Campaigns having to do with theater, music, and film/video account for 75% of all successful Kickstarter campaigns. In addition to theater campaigns having the highest success rate, they also have the highest average donation per backer out of any of the categories.

3) Although theater campaigns had the most success in terms of number of projects that met their goals, it was actually the technology campaigns that I believe were the most effective at raising money, as these ventures in the long run made more total money across all projects than any other categories. Out of all categories, technology had the highest surplus in terms of funding. These successful technology campaigns totaled about 13 million dollars in surplus funding. The surplus value of funding for all successful theater campaigns meanwhile was around 6.6 million dollars. Technology also had 90 campaign’s funding exceed 200% of the goal, whereas for theater that number was 33. Furthermore, technology also had the second highest average donation per backer after theater. Therefore, although theater campaigns had better numbers in terms of amount of successful projects, technology campaigns generated larger amounts of money.

**2) Limitations of the Data**

One limitation I noticed was that it would have been nice to know more specifics about the backers and what each of them paid. This information would tell us more about whether the backers are generally paying similar amounts, or if one or a select few backers are paying out very large amounts of money to fund the campaign. If you had the case where only one or very few backers are paying the vast majority of the needed money, than the value for average donation is not really an accurate representation of what each backer is most likely to be paying, as it will be skewed by the outliers who are paying large sums of money. It could also indicate that the key to having a successful Kickstarter campaign is more about securing certain high-rolling backers, rather than having a high average amount of donation per backer.

**3) Other Possible Tables/Graphs to look at**

One other value I would be interested in exploring would be the date that the fundraising goal was met for the successful projects. This would indicate how much time it would typically take for successful campaigns to meet their goal. This information would also be helpful in knowing if certain categories of projects achieve success quicker than others.

**Bonus Statistical Analysis**

**-** When considering the relationship between outcome and number of backers, it would be more meaningful to use the median as a measure of central tendency. This is because we have both a large variance and large standard deviation for this data. Basically, both values indicate that the data is overall very spread out from the mean. The mean in this data set is also being skewed by some pretty large outliers in both data sets (successful and unsuccessful campaigns), making the mean an inaccurate representation of true central tendency. Therefore, the median would be a better measure.

- There is significantly more variability with successful campaigns. I think that this generally makes sense. This is because you can have a wide amount of combinations that can lead to a successful project. You can have projects where a small amount of people give a lot of money to meet the goal, and you can have projects where a large amount of people give a small amount of money to achieve the goal, and you can have many scenarios in between. Therefore, there are many possibilities with regards to number of backers that could make a campaign successful. With unsuccessful campaigns, I wouldn’t expect to see nearly as many possibilities. After all, the main reason why campaigns fail is most likely because there aren’t enough backers, and the backers that do back the project don’t give very much. Therefore, it is more likely that you will see a general pattern of a lower number of backers, which will lead to a smaller variance. The huge difference in median and maximum number of backers is also going to lead to smaller variance as well. Successful campaigns have a median of 62 backers, while unsuccessful campaigns have a median of 4. Far fewer people are generally backing unsuccessful campaigns, leading to less possibilities and less variance. In addition, the maximum number of backers for successful campaigns is 26,457, while the maximum for unsuccessful campaigns is 1,293. Again, this huge difference is going to result in less potential possibilities concerning backing for unsuccessful campaigns due to general lack of people backing, which will lead to less variance. Therefore, it would make sense to see a smaller variance with unsuccessful campaigns.