Report On Crowdfunding

**Question#1**: Given the provided data, what are three conclusions that we can draw about crowdfunding campaigns?





CONCLUSION #1:

Visually, by looking at the graph it is very clear that the theatre category got the most traction. Statistically, the theatre category has a 54% success rate with backers and 38.3% of the total backers failed in their campaign. The ratio between failures and successes is similar to other categories, however, the popularity is outstanding, taking up 34% of the whole in terms of categories alone. Our findings in terms of “success” and “fail” based on categories are similar. Film & video and music are almost identical in terms of backers’ participation. The ratio between success and failure for these two categories is similar to others, and they both take up approximately17% of campaign votes. They are 2nd and 3rd but half her percentage of the theatre. In conclusion, it seems like that arts are a lot more important to backers seeing as the 1st, 2nd and 3rd most popular categories (Theatre, Music and Film & Video) are all part of arts.

Some of the obvious less popular categories are technology, publishing, and journalism which are opposite from the most popular categories. Evidently, the type of category (arts, maths, sciences, etc.) matters by looking at the visualized and numerical data.

CONCLUSION #2 :

In this graph of subcategory, we can see that “plays” got the most popularity.

Between this category and others, there is a drastic difference. This takes 34% of all sub-categories. With the most success, failures, etc. This is the most successful, with all other categories' traction being very underwhelming. This is very evident because this is the only category with more than 5% of the total traction. Documentary, food trucks, rock, video games, animation, and web, although not doing great in the grand scheme of things, have at least some traction. On the other hand, audio, metal, radio, and podcast did significantly worse than the already low bar, with less than 5% total of traction among all categories.

CONCLUSION #3:

Looking at this line graph depicting the relationship between a campaigning launched years and backers’ participation, we can see that the relationship between successful and failed companies was positive until 2019. After 2019, we can also see that the trend for successes is upwards and the trend for failures proceeded to go downwards. Here, we need to investigate what happened from 2019 to 2020.

A reasonable conclusion as to why the company flourished (the rate of failed backers went down) during 2019-2020 is because that is the peak of Covid-19, meaning this company might have been online or had a large social media presence.

There is little to no progress on live or cancelled backers for this company, therefore there is no conclusion.

LIMITATIONS:

**Question #2** : What are some limitations of this dataset?

A common limitation for the possibilities of this data is a timeline. This was not given to us. I think that having a set of numbers depicting the number of years or months to track progress from the launch to end date would be good. The only set of data with this information is the line graph because it is best fit for data over time.

Doing this would help investors make a judgement of the success rate of the company over time proving its value better.

# SUGGESTIONS

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

I would suggest another graph based on outcomes in different countries. Consideration regarding geographical locations such as currency in terms of making financial decisions would be good for investors and companies. Comparing countries would be more interesting for the backers.

STATISTICAL ANALYSIS

**Question #4:** Use your data to determine whether the mean or the median better summarizes the data.

For “Successful” campaigns there is a huge difference between the mean (851) and the median (201). We can observe that the data is right skewed, meaning that some campaigns’ high backers number increased the value of mean. Therefore, the median represents better measure for the central tendency because it’s less biased by the extreme values.

For “failed” campaigns The mean (585.62) is also higher than the median (114.50) but the difference is less than the successful campaign. There is some bias in the mean so median describes the data better.

**Question #5 :** Use your data to determine if there is more variability with successful or unsuccessful campaigns. Does this make sense? Why or why not?

In order to determine variability with successful or failed campaigns, we can look at the variance and the standard deviation. Successful campaigns have a variance of 1606216.59 which is higher than the variance for failed campaigns. Higher standard deviance means the data points are more spread out. In the same way, standard deviation for successful campaigns (1266.24) is higher than the standard deviation of failed campaigns(959.99). The higher the standard deviation gets, the higher the data points, on average further from the mean.

We can see that successful campaigns show more variability on both variance and standard deviation in the number of backers than failed candidates. This is because successful candidates can vary widely in their level of success. We can see that some candidates barely met their goals while others exceeded them significantly which attracted more candidates. Since the range of outcomes is wide, therefore we can see a higher spread of candidates. On the other hand, failed campaigns fail to attract a large number of candidates and thus tend to cluster around lower numbers of candidates. The number of candidates are low, therefore the variability is low.

In conclusion, for both successful and failed campaigns, the median better summarizes the data because it is less affected by the extreme values and provides a more accurate picture of the typical number of backers. In terms of variability the backers’ number for successful campaigns is higher than failed campaigns which makes sense because of the wide range of success levels that can occur in such campaigns and vice versa.

