Matthew Yurich

Excel Challenge

4/30/2021

When looking at the provided data of various kickstarter campaigns, there are many conclusions that can be drawn about the overall successfulness of these campaigns. Based on the data, it can be seen clearly that the success of kickstarter campaigns can vary based on the category of each campaign. As shown on the bar chart of success, failure and cancellations of each category, Theater has the greatest number of kickstarters created and the greatest number of campaigns successful. Correspondingly, the categories of journalism and food had low numbers of campaigns and only a small percentage of those were successful. Another conclusion that can be made is that the number of successful campaigns can vary widely between each sub-category as well. For example, the music category had a high percentage of campaigns that were successful. After taking a look at the sub-category bar chart, it can be seen how much it varies depending on the genre of music. Rock music has a 100% success rate of kickstarting the campaign, but Jazz music had a 0% success rate. Last, it can be concluded that the highest number of campaigns that are successful happen between January and June in any given year. Cancellations stay consistent throughout the year and failed campaigns stay consistent. Successful campaigns also have the sharpest decline during a given year. There is a steady drop from June to September and Shar decline in the month of December.

There are several limitations to this data. The data we gathered focused mainly on the number of successful kickstarters based on category, sub-category or year. We also, created new data such as the Average Donation and Percent Funded. More data focused on the Percent Funded of each campaign as compared to the number of successful campaigns could be helpful using a line chart. This information would help show how much funding is required for a campaign to be successful and to see if a goal needs to even be reached. Other limitations in the data that was created is the length of time between the launch of a kickstarter campaign and the end date of collecting donations. Using a line graph, this would show any patterns between the number of successful campaigns and how long they collected donations. Last, the dataset may not contain all the results of kickstarter campaigns. This assignment looks at 4,000 campaigns and only one-third of kickstarters are successful. Based on my findings in this dataset about half of all campaigns were successful. This could skew the results of finding patterns without the full dataset.

In the summary of statistics between the number of successful and unsuccessful campaigns and backers count, the median summarizes the data more meaningfully. This is because the majority of values are lower with several outliers that are significantly higher than most of the data. This inflates the mean and skews the data. The medians of 62 for successful campaigns and 3 for unsuccessful campaigns shows that most data values are low. There is more variability between successful campaigns than in unsuccessful campaigns. This makes sense because there is a greater spread of data in successful campaigns and successful campaigns had more backers for each kickstarter. The high values of the dataset in successful campaigns drives up the mean. The variance is higher because more values from the data set are farther from the mean than in unsuccessful campaigns which had a low median of 3.