Moriah Taylor

5 March 2021

Excel Homework 1

**Background**

Kickstarter is an online crowdfunding service which has been used to raise over $2 billion for over 300,000 projects. This dataset includes key information such as the initial goal, number of backers, and outcome of approximately 4000 projects.

Chart, bar chart

Description automatically generated**Conclusions**

Given the provided data, one conclusion we can draw about Kickstarter campaigns is that the most common sub-category of campaign by far is “plays” (see Figure 1b) and subsequently, the most common category of campaign is “theater” (see Figure 1a). This is purely speculative, but one possible explanation for this could be that the theater community is very tight-knit and supportive of each people’s creative endeavors.

Figure 1a: Category vs. Outcome (Stacked Column Chart)

Chart

Description automatically generated

Figure 1b: Sub-Category vs. Outcome (Stacked Column Chart)

Chart, line chart

Description automatically generatedChart, line chart

Description automatically generatedAnother conclusion we can draw is that there is a correlation between the goal amount of a campaign and its outcome. Shown on a line graph (see Figure 2), as the goal amount increases, the percentage of failed campaigns increases. This makes intuitive sense because the higher the goal, the more backers and the more money is needed for the campaign to be considered a success. A third observation that is worth noting is that there does not seem to be much of a relationship between the time of year that a campaign launches and the propensity of it to being canceled, successful, or failed; however, there are three notable pearks in the number of successful campaigns in the months of February, June, and October when compared the months surrounding them (see Figure 3). This could indicate that there is something worth looking into there.

Figure 2: Goal Amount vs. Outcome (Line Graph)

**Limitations**

One of the limitations of this data is that there is not a way to determine whether the idea for the Kickstarter is original or well-thought-out. People tend to back ideas that are innovative – they are not going to put money into something that already exists in some comparable form. In addition, it is hard to garner support for a campaign that does not seem to be planned or organized well. If a campaign does not state exactly what it intends to do with the money raised and/or the idea itself does not seem fully fledged, it is less likely someone would invest their money into it. Chart, pie chart

Description automatically generatedThis quality is not captured in the dataset, nor does there seem to exist a metric that could evaluate this. Another limitation of this data is that the overwhelming majority of campaigns originate in the United States (see Figure 4) so it would be difficult to draw meaningful conclusions about Kickstarter campaigns outside of the United States due to such small sample sizes. Lastly, there are unknown factors which could cause campaigns to fluctuate in both successfulness and in quantity overall, such as a country’s economic state. One can imagine that if a country’s economy is not doing well, there may be an increase in the number of Kickstarter campaigns, a decrease in the number of successful campaigns and/or a decline in the number of backers and size of donations.

Figure 3: Month of Campaign vs. Outcome (Line Graph)

Figure 4: Where Kickstarter Campaigns Originate (Pie Chart)

**Future Methods**

For further exploration of this dataset, we could compare campaign outcomes to whether the campaign was a staff pick, was a spotlight campaign, or had neither of these endorsements using a PivotTable or a clustered bar graph. It is possible these endorsements could lead to more successful campaigns, or the difference could be negligible. Another visualization that could be created would be a line graph which related how long a campaign ran (difference between the launch time and deadline) and outcome. Naturally, one would be inclined to think that the longer a campaign runs, the higher chance it has to be successful.

**Other Statistical Analyses – Number of Backers per Campaign**

When the number of backers per campaign is represented with a histogram, it is clear the data is incredibly right skewed for both the successful campaigns as well as the failed campaigns. In this case, the median is a more appropriate measure of center because the large outliers pull the mean to a higher value whereas the median is unaffected. For successful campaigns, the median number of backers is 62 (SD=844.30; Var=712840.99). For failed campaigns, the median number of backers is 4 (SD=61.43; Var=3773.22). There is more variability with the number of backers for successful campaigns which makes sense because for a campaign to be successful, it requires more money and therefore numerous backers, whereas a failed campaign would be more likely to have less or zero backers and less likely to have thousands of backers.