Sally Anderson

UCI Data Analytics Boot Camp

Oct. 5, 2023

Module 1 Challenge:

**Excel Crowdfunding**

Based on the analyzed data, there are several conclusions that we can draw, but only three will be highlighted here. First, the “theater” category generated the most amount of successful crowdfunding campaigns, which only had one sub-category of “plays,” with 344 successful campaigns. Second and thirdly, the categories that were the most successful were “music” and “film & video”. The sub-category that was the most successful for the “music” was by far “rock,” with 85 successful campaigns. The sub-category that was the most successful in the “film & video” was “documentary,” with 60 successful campaigns.

Stating the limitations within the dataset is important because it can help create a situation where further investigation can be done to have a greater understanding, and possible improvements, to draw more detailed conclusions for informed decision making. One limitation of this dataset is that it is uncertain about why certain campaigns were canceled. Having this information could help prevent campaigns from being canceled prematurely. Furthermore, another limitation is where, and how, the backers were recruited. If there are several events, or online engagements, this could be valuable information about how to target backers to influence a successful outcome.

However, we could still draw more conclusions within the dataset that has already been provided. We could add analysis and tables to analyze in greater depth the length of each campaign, and if it affects the outcome. The duration can possibly cause the success rate of the campaign. And we could add if the length of the campaign is different between the different categories and sub-categories.

**Statistical Analysis**

An additional analysis was done on the data of the successful and failed campaigns: the calculation of normal distribution, which helps visualize the data so statistical analysis can be easier to interpret. The bell curves in Figure 1. and Figure 2. show a skewed distribution. Because the data is not symmetrical, the median is better to use in summarizing the data than the mean. Both graphs show a right-skewed distribution which indicates the frequency of fewer number of backers is more common than frequency of larger number of backers.

Furthermore, there is more variability within the successful campaigns and less variability within the failed campaigns. This could make sense because there would be a greater likelihood of success with the greater the tries through the addition of more backers. The lesser variability could cause failure earlier in the process of the crowdfunding campaign.

**Figure 1**. Bell Curve of the Backer Count for Successful Campaigns.

**Figure 2**. Bell Curve of the Backer Count for Failed Campaigns.

**References**

1. Teacher's Tech. (2022, Nov. 21). How to Create a Bell Curve in Microsoft Excel. [Video file]. URL https://www.youtube.com/watch?v=K1xQK6pkGGg
2. Khan Academy. (2017, July 7). Median, mean and skew from density curves | AP Statistics | Khan Academy. [Video file]. URL https://www.youtube.com/watch?v=JFesFhraX2M
3. Steven Bradburn. (2021, June 2). Variance - Clearly Explained (How to Calculate Variance). [Video file]. URL https://www.youtube.com/watch?v=x0rmUXWtSS8
4. Daniel Storage. (2019, June 18). Measures of Variability (Range, Standard Deviation, Variance). [Video file]. URL https://www.youtube.com/watch?v=s7WTQ0H0Acc