

1. Given the provided data, what are three conclusions that we can draw about crowdfunding campaigns?

- A. Though not proven to be statistically significant at this stage of the analysis, the data set provided concluded there is a 56.5% chance of a successful outcome. Therefore, based on the data provided there is more than half the chance of a successful outcome with crowdfunding campaigns.
- B. The top 20 in dollar amount pledges were successful and accomplished at least 100% or more of the pledge goal. However, considering all of the campaigns, not all that reached 100% or more of the pledge goal were successful.
- C. The most common crowdfunding campaign category is for theater. The sub-category within theater that had the most campaigns was plays. 54% of campaigns related to plays were successful, while 38% were unsuccessful. All plays that were 100% funded were successful.

2. What are some limitations of this dataset?

Without running statistical tests, understanding standard deviation, and limited graphs/visuals, the interpretation of the graphs asked to configure for the exercise could provide a false guidance if used for future trends. There is a limited understanding of how significant the data is without plotting additional graphs and performing regression tests, t-tests, and correlation tests.

Additionally, omitting statistical calculations from the analysis such as the standard deviation, variance, min/max, z-score, mean and median, illustrates inconclusive results. The lack of understanding statistical information such as outliers could ultimately misinform decision makers when using this data to determine trends and/or building business cases.

Limitation Highlights

- T-test calculation to reveal if the data is statistically significant
- Understanding the variance in the data – the spread away from the mean
- Standard deviation – how far is the spread away from the mean to allow us to tell if there are any outliers as outliers may be skewing the conclusion

3. What are some other possible tables and/or graphs that we could create, and what additional value would they provide?
 - A. Box and whisker graph – visually shows the spread of the data to understand data for analysis
 - B. Scatter plot graph – visualization comparing two variables to show a relationship and any patterns in the data (predict trends via trend lines)