**Alex Ivanoff**

**Report on Crowdfunding Campaigns**

1. **Three Conclusions Based on the Dataset**
2. **Success Rate and Overfunding:** Campaigns can exceed their funding goals significantly, as seen in the case of "Odom Inc," which raised 1040% of its goal. Successful campaigns often gather substantial overfunding, suggesting that exceeding the initial goal is a trend of well-planned campaigns.
3. **Country and Category Influence:** Most successful campaigns in the dataset come from the US and fall into categories like technology and music. This implies that certain regions and sectors may have higher success rates, potentially due to larger audience bases or higher demand. Certain subcategories, like "plays" and "music", dominate the crowdfunding landscape in volume. For instance: "Plays" had the highest count of campaigns with 187 successful and 132 failed campaigns. "Music" also had a high number of both successes and failures.
4. **Seasonal Trends in Campaign Outcomes**: Successful campaigns appear relatively stable across months (e.g., ~40-60 successes monthly) with peaks during warmer month. Failed campaigns fluctuate more significantly, especially in February, June, and December.
5. **Limitations of the Dataset**
6. **Temporal Bias:** The dataset extents different years and currencies without adjusting for economic inflation or currency fluctuations, which may distort comparisons.
7. **Suggestions for Additional Tables/Graphs**
8. **Pledge Distribution:** A histogram or boxplot of pledged amounts could reveal patterns such as the most common funding ranges or outliers.
9. **Timeline Analysis:** A line graph showing the frequency and success rate of campaigns over time would help ascertain trends or seasonal patterns.
10. **Correlation Analysis:** possibly a heatmap showing correlations between variables, (for example: goal, number of backers, percent funded) could identify the most critical factors influencing success of the campaign.
11. **Compare Mean and Median**

In our case, the mean and median are not close, with the mean being much higher than the median, which indicates that the data is right-skewed.

Box plots confirmed that the dataset contains a significant number of high outliers, which can distort the mean and reduce its reliability. The standard deviation is also high, indicating greater variability in the data, further supporting the presence of skewness. Since the mean is sensitive to extreme values, while the median is hardier against outliers, the median is a better choice for summarizing the number of backers in this case. Successful campaigns have higher standard deviation, their results are more variable.