Lucinda Hodgson

Oct 5 2023

Crowdfunding Questions:

**Given the provided data the following conclusions can be drawn:**

1. In this dataset the most successful crowdfunding events relate to entertainment venues (Theater, Music, Film&Video).
2. Overall, the most successful time of year for crowdfunding, in this dataset, is during the summer months. This might be because people are looking beyond their own circumstances and feeling freer during summer. Notable exceptions are technology and publishing who peak earlier in the calendar year.
3. The higher the goal amount the less chance of success. Higher starting goals may dissuade people from participating due to perceived futility of reaching the higher amounts.

**Limitations:**

1. This is a small dataset.
2. There may be other variables that would negate trends seen here, such as median household income, type of request or location (rural/urban). It is possible that with further examination of donor demographics important information could be gained in comparison to category and amount that would help target the appropriate audience for each category.
3. Comparing countries is problematic. Each county on the list has cultural factors in place that cannot be seen in the data making it impossible to look at the set as a whole. Although this data can be looked at by individual country, it is hard to draw a conclusion about the entire data set. It is also difficult to draw equivalence between currencies. Exchange rates can be volatile and the amounts in question could vary greatly depending on exchange rates.

**Options for further tables/graphs:**

1. Creating a box/whisker plot to look at outliers.
2. Identifying what the best money goal is to have the greatest chance of success.
3. Looking at each country individually and seeing if there is a marked difference.
4. calculate z-score, create summary by mean and median.

**Goal Analysis**

1. The median is the better metric to analyze the data as it is not affected by outliers and the variation seen in this data set. The median is less affected by skewed data. There is no pretty bell curve here.
2. There is more variability with successful campaigns. The standard deviation is higher for successful vs failed- 1276 v 961. This may be because successful campaigns have a wide range of outcomes that translate to success, but with failed campaigns, the base is just failure- goal not met.