Data Analysis Report

Crowdfunding Campaigns

CONCLUSIONS - General:

* Based on the Crowdfunding data, we can conclude the following:
  1. Theater and plays during the summer months will most likely generate successful campaigns.
  2. Crowdfunding campaigns have more successes than failures – and the number of successful campaigns is greater than failed and canceled combined.
  3. The US is the biggest generator of crowdfunding campaigns.
* Given the Crowdfunding data, we can identify the following limitations:
  1. The data lacks detailed geographical data, (e.g., specific states and counties within each country); perhaps this would only be beneficial for the US, however, as they are the biggest generator of campaigns.
  2. The definition of “pledged” is not explained – does this mean promised but not yet reflected on the books?
  3. The currency/goal/pledged columns lack clarity – would be beneficial if these amounts were in a standard measure to more accurately compare totals by country.
* In addition to the pivot tables and graphs employed – the following would also aid in greater analysis of the data:
  1. A table listing the top number or percentage of backers by successful outcomes; this information could be used to reach out to backers when creating new campaigns.
  2. A table breaking out countries into more detailed geographical locations; this data would be helpful if campaigns are city/county/etc. specific.
  3. A line graph showing the timing of donations to help target outreach efforts; for example, knowing when donations tend to come in could help determine the best time to contact backers, send informational brochures, etc.

CONCLUSIONS – Crowdfunding Goal & Statistical Analysis:

* Based on the data, I believe the median is better in summarizing the data. Both data sets have a large spread between the lowest and highest value. This creates a wide distribution with multiple outliers. Using the median would be best to summarize the data since it is less skewed by those lower and higher numbers. But I think using both would give the most accurate picture of the data; it’s clear to see that successful campaigns tend to have more backers when looking at both the mean and median.
* Based on the data, successful campaigns have more variability due to the greater range of backers and donations: The more variable, the more successful. Failed campaigns have lower variability due to the greater number of zeroes in backers; these zeroes lower the variance value. Perhaps further analysis could be conducted to determine if there’s a threshold of number of backers to determine the chances of success?