Jillian Walter

9/29/2024

Crowdfunding Analysis – Week 1 Challenge

**CROWDFUNDING GOAL ANALYSIS**

* Given the provided data, what are three conclusions that we can draw about crowdfunding campaigns?
  1. Technology, Photography and Journalism are spaces that offer slightly less competition at a lower historical risk
     + When looking at Parent Category, crowdfunding projects in the Theater, Film & Video and Music are the most popular, with >100 projects to date.
     + Technology, Photography and Publishing are slightly less risky, with a success rate of >60%; Crowdfunding for Games and Food have a higher chance of failing or being cancelled than other categories
     + Though small at only 4 projects, Journalism has 100% success rates, indicating white space in the market and a potentially less risky endeavor.
  2. Starting a project in June will set up the crowdfunding campaign for the most success
     + June-July have historically driven the highest success rates at >50% success with lower rates of failure/cancellation
     + Failure and Cancellation shares spike in August, and the average campaign duration is about 15 days, an early start in June will back up the campaign for the most success in the June-July timeframe
  3. The crowdfunding goal should fall between $10k-$35k, given that the 37 historical projects in this range have a 92% success rate
* What are some limitations of this dataset?
  1. There is limited background as to what the blurb next to the company means and the value it provides. Overall some column names could provide more detail or explanation, such as staff-pick and spotlight
  2. There is minimal distinction between cancelled and failed projects, and the set is missing explanations as to why a project may have been cancelled or failed

* What are some other possible tables and/or graphs that we could create, and what additional value would they provide?
  1. Scatterplots could help portray the average donation x number of campaigns by parent category, such that potential investors or project starters could see the most profitable business areas in terms of donations, or identify any white space in the market where there are few campaigns.
  2. Bar charts showing the average donation by country by parent category could help identify where to seek out donors for different project types based on the highest average donation.
  3. Box and Whisker plots could help identify any outliers in the data, as well as the 25-75% range of pledged donation and average donation amounts such that project starters can forecast the number of backers needed for future projects.

**STATISTICAL ANALYSIS**

* Use your data to determine whether the mean or the median better summarizes the data.
  + The data provided can be better summarized by the median. This data has a large range and high variance, with few highly-backed campaigns skewing the mean upwards. Only 163 out of the 586 campaigns are above the mean of 812 for successful campaigns, while 201 represents the true “middle” of the data
* Use your data to determine if there is more variability with successful or unsuccessful campaigns. Does this make sense? Why or why not?
  + There is more variability with backers for successful campaigns, as this data has a higher standard deviation/variance. This makes sense as successful campaigns may have longer lifecycles and therefore accumulate backers along the way, while failed campaigns tend to have less funding and therefore fewer backers, making the number of backers naturally closer to 0.