June 16, 2022

Module 1 Challenge

By Vedrana Basimamovic

**Kickstarter Analysis** Georgia Tech Data Boot camp

**1 Background**

Crowdfunding platforms like Kickstarter and Indiegogo have been growing in success and popularity since the late 2000s. People are using crowdfunding to launch new products and generate buzz, but not every project has found success. In order to receive funding, the project must meet or exceed an initial goal, so many organizations spend ample resources analyzing old projects in an attempt to discover “the trick” to finding success.

**2 Objective**

For this week's Challenge, we have organized and analyzed a database of 1,000 sample projects to uncover any hidden trends.

**3 Conclusions**

Based on the provided dataset, we can conclude that majority of the campaigns had good success (out of 1000 total, we had 565 successful vs. 364 failed). However, I found it interesting that if a campaign is even at the 99% of their goal, it has been deemed as failed, i.e. Huff LLC.



It would be beneficial to know how the failed criteria was set and what drove the decision to only count campaigns above 100% as successful. Journalism category had the 100% success rate, but only 4 campaigns, while Theater category had the largest number of campaigns of 344, of which 54% were successful. Games category has the lowest success rate of 44%.

In addition, Plays Subcategory outnumbers all the other subcategories by over 7-fold, with total of 344 campaigns, but with success rate of only 54%, compared to Audio Subcategory, which had 4 campaigns at 100% success rate. Plays and Audio are the only subcategories within their assigned Categories, Journalism and Theater, respectably. Where other Categories have multiple subs, i.e., Film & Video have 6, as does Music. Having subcategories help aids in having more granular data but can easily create outliers on our analysis that are easily overlooked.

Graphical user interface

Description automatically generated with medium confidence

And lastly, we can conclude that the most successful season for crowdfunding is summertime, or rather the months of June and July. Shoulder months and Winter tend to still be more successful than not, however, Winter tends to have the greatest number of cancelled campaigns, in addition to August.

Chart, line chart

Description automatically generated

**3.2**

Even though we have plethora of data, there are some limitations to this data set. Size of the dataset does not appear to be even, from campaign to campaign, and we do not know the crowdfunding source, making it hard to level the data baseline and have 1:1 comparison. Knowing what strategies were used in crowdfunding (in person, via phone, social media, where free give-a-ways handed out like play tickets, etc.), would give us one more set of data to more accurately find a possible link between outcomes.

**3.3**

By creating additional graphs and tables, we would be able to answer more detailed questions behind the successes and failures of the campaigns. For example, we should be looking at the longevity of a campaign within Categories and their respected Subcategories. Displaying a timeline (4 days vs 2 weeks) with outcomes would give us further insight on outliers and if longevity has had any impact of campaign success. If we knew, we could chart the backer’s touchpoints or as I mentioned earlier, a strategy used in a campaign. We do have geographical location by county, but not any external factors like demographics, however we can still use this data to graph the outcome by country, based on a category.