Kickstarter Data Analysis

Jessica Fernandez

August 20, 2020

**What are the trends seen in Kickstarter campaigns?**

* + 1. MUSIC has a great success rate as a Parent Category
       1. 100% Success rate for Sub-Categories CLASSICAL, ELECTRONIC, METAL, POP, and ROCK. The highest of which was ROCK at 260 successful campaigns
    2. The most campaigns attempted are in Sub-Category PLAYS, a total of 1066 campaigns. These campaigns only had a 65% success rate.
    3. In all years, January and May have the highest successful campaigns.
    4. The smaller the fundraising goal, the bigger the difference between Successful vs Failed attempts. The bigger the goal, the more even the success-to-failed ratio.

**What are the limitations seen in this dataset?**

* + 1. There may be limitations based on seasonal ability to host certain campaigns
    2. Data collected from different countries may not be accurate

**What other tables and graphs can we create?**

* + 1. Number of Backers Vs Time LINE GRAPH
    2. Number of Backers vs Category BAR GRAPH
    3. Number of Backers vs Subcategory BAR GRAPH
    4. Number of Backers Vs State BAR GRAPH
    5. Number of Backers Vs Average Donation

**Measures of Central Tendency**

Since one of the best ways to measure the success of a campaign is to look at the number of campaign backers, it is important to note the difference between the average number of doners in a successful campaign versus an unsuccessful campaign. It is clear that on average there are far more backers in successful campaigns than failed ones. The mean and median are meaningful to summarize the data, and to show how backers are an essential part of a successful campaign. However, it is also important to check the variance of the dataset.

We notice there is a larger variance for the successful campaigns than the unsuccessful. This just means there are extremes in the data and we can see that this makes sense if we compare the range in the MIN and MAX values of both columns. For successful campaign there is a bigger gap between MIN-MAX.