Excel Challenge Report

Angela Angulo

7/24/2021

1. **Given the provided data, what are three conclusions we can draw about Kickstarter campaigns?**

Three conclusions that we can draw about Kickstarter campaigns based on this data are:

* Out of all the parent categories, only the film & video, music, and theater categories had a greater number of successful campaigns than failed or canceled campaigns. The most successful sub-categories from these groups were: documentary, shorts, and television (all 100% success) for film & video; classical, electronic, metal, pop (all 100% success), and indie rock (87% success) for music; and plays (65% success) for theater.
* Campaigns in the food category had the lowest success rate (17%), with all of the campaigns under the group’s largest sub-category, food trucks, being failed or canceled. Small batch campaigns, however, seem to do well, with all of the concluded campaigns ending successfully.
* May and June seem to be the best months to launch a Kickstarter campaign, while September and December seem to be the worst.

1. **What are some limitations of this dataset?**

Some of the limitations of this dataset include:

* About three-quarters of the campaigns included in the dataset are from the US alone.
* The websites for the campaigns are not included in the data, so we can’t tell if the appearance of the campaign websites (i.e.: difference in appearance, having photos and/or videos to showcase the product, etc.) potentially affected rates of success.
* The data does not include descriptions of the inclusion/exclusion criteria for the categories and sub-categories, so it is difficult to tell if the way campaign types were aggregated had any effect on the categories’ and sub-categories’ overall rates of success.
* The data does not indicate the reasons why a campaign was canceled – they assumption that they were canceled due to insufficient backers and funding may not be accurate for all canceled campaigns.
* From the data included, there is no way to tell whether or not a successful campaign was primarily backed by a small percentage of the campaigns total number of backers (i.e: 3-4 large donors out of a total of 100 backers).

1. **What are some other possible tables and/or graphs that we could create?**

Some other tables and graphs that could be created based on the data set include:

* A table or bar graph to compare the number of successful, failed, and canceled campaigns based on the amount of time a campaign was live (could be filtered by category and/or sub-category).
* A table to compare goal amounts, total number of backers, and average donation size to see if there are any patterns based on these variables in the number of successful, failed, and canceled campaigns (goal amounts and average donation would ideally take into account conversion rates for currency used).

BONUS

1. **Use your data to determine whether the mean or the median summarizes the data more meaningfully.**

Since the distribution of the number of backers seem to be heavily skewed to the right for both successful and failed campaigns, the median number of backers would be a more meaningful summary of the data because, compared to the mean, it is less affected by outliers.

1. **Use your data to determine if there is more variability with successful or unsuccessful campaigns. Does this make sense? Why or why not?**

Based on the variance and standard deviations for the two sets of data, there is more variability in the number of backers for successful campaigns than for unsuccessful campaigns. This makes sense because the data set for successful campaigns includes those that reached over 100% - some up to 200% - of their goal amount, while the failed campaigns all fell short of their goal amounts. This can be seen translated in the ranges in the number of backers for both sets: successful campaigns had 1 to 26,457 backers (median = 62), while unsuccessful campaigns had only 0 - 1,293 backers (median = 4).