Lipika Rana

2/1/2021

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

2. What are some limitations of this dataset?

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

1. Three conclusion we can draw about Kickstarter Campaigns are
   1. Having a launch (going live) in the beginning of the year (January, February, March) are the most popular months to start going live of a product.
   2. Theater is the most popular category of the people creating a campaign.
   3. Around 53% of campaigns are successful when a campaign is started on Kickstarter.
2. The limitations of this dataset are that it is not a dataset of a specific country. Each country has a different economic status and status of people, so the data set conclusions cannot be applied for every country. For example, AU has very minimal datasets so making conclusions based off the whole data set for AU is inaccurate. We would need more data points for more of the countries or make this dataset for one specific country. Also, data is limited just to kickstart, not other sources.
3. Other possible tables and/or graphs we could create are make a more filtered pivot table with category and subcategory and then filter by stage of project (success, cancel, live) and then counts of each category to know when category and subcategory is the most successful or another is which category got the most donations (a graph displaying that).

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

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

BONUS

1. The median summarizes the data more meaningfully. This is because the data is skewed. There is a high standard deviation. If you look at the median it is drastically lower, then median. The median means more then half the data has been covered by that point. The max number acts as an outlier which skews the average.
2. There is more variability with successful campaigns than unsuccessful campaigns. This makes sense because there is such a high range of backers needed to have a successful campaign. There is no set range, there can be as low as 1 backer to have a successful campaign. The standard deviation and variance are also a larger number than the unsuccessful campaigns.