Tricia Harris

12/12/2020

Excel Homework

DU-Data Analytics Bootcamp

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

Three conclusions that can be drawn from the provided data include:

1. Kickstarter campaigns appear to be more popular in later years compared to earlier years. This means the competition to get funding has likely increased.
2. There appears to be a spike in successful campaigns that launched in the springtime compared to launching in the fall. This could indicate that a successful Kickstarter campaign is more likely if it is launched in the Spring.
3. Campaigns with a lower goal appear to have a higher chance of success. The data becomes a little more sporadic in the upper limits of the goal range, but generally, the lower the goal the more successful the campaign.
4. What are some limitations of this dataset?

Some limitations of this dataset include:

1. This data is very broad. Having broad data makes it difficult to draw specific conclusions as to what exactly might produce a successful campaign. To identify strategic moves relevant to building a new Kickstarter campaign, one would need to drill down into specific criteria relevant to the project, such as country, category/sub-category, etc.
2. Some of the conclusions are challenging to draw because some countries have many data points where others don’t. This limits the ability to draw strong conclusions on the full data set.
3. What are some other possible tables and/or graphs that we could create?

Other possible data points to consider include:

* 1. Number of Backers compared to campaign state by category. This could be shown as a stacked bar graph. This could help show whether there is a trend in the number of backers needed to be successful and how/if this varies depending on category type. For example, some categories may show they are successful with less backers, but that could be due to their being one or two large backers that fund the project, etc.
  2. Evaluate each campaign state by length of time open compared to # of backers. This would be shown as a line graph (one per campaign state). This could show time related results of Kickstarter campaigns and be useful in determining how long a campaign should stay open or how long it takes to become successful.

Statistical Bonus:

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

The median is more useful for summarizing the data for successfully campaign backers. The dataset is greatly dispersed, as shown by the high standard deviation and the delta between the max and min. The standard deviation for the failed campaign backers is not as dramatic as it is for the successful campaign backers, however, it still shows the data as being relatively dispersed, therefore, the median is more useful than the mean for assessing the failed data as well.

\* 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 successful campaigns. This does make sense because there are a few outliers that are significantly skewing the data for successful campaigns.