Maxim Izotov

15th August, 2020

1. Excel Homework
2. **Given the provided data, what are three conclusions we can draw about Kickstarter campaigns?**
   1. Theater is the most abundant and most successful category of Kickstarter campaigns.
   2. Plays is the most abundant and most successful sub-category of Kickstarter campaigns in the theater category.
   3. Kickstarter saw the largest amount of new campaigns in 2015 and 2016 compared to other years.
3. **What are some limitations of this dataset?**
   1. There is no date for Q1 2017 and beyond. This lack of data makes our conclusions inaccurate in the present day.
   2. The dataset may not have all the categories of campaigns that Kickstarter offers. This lack of data limits our ability to analyze Kickstarter campaigns in more detail in reference to each other.
   3. The dataset does not include thorough data on donations, such as outliers who may have contributed the majority of the pledged amount. This data would allow us to analyze how large donations affect success for Kickstarter campaigns.
   4. This dataset does not include a True/False value for whether the Kickstarter campaign was used as the prim ry source of funding. This lack of data prevents us from analyzing how making Kickstarter a primary source of fundraising affects a campaign’s success.
4. **What are some other possible tables and/or graphs that we could create?**
   1. A graph of how many campaigns were successful that received spotlight vs. those that did not.
   2. A graph of how many campaigns succeed based on the pledged amount (did campaigns that asked for less money see more success?).
   3. A graph of how the length of the campaign affected its success.

**Bonus Statistical Analysis**

1. **Use your data to determine whether the mean or the median summarizes the data more meaningfully.** 
   1. I think the median summarizes the data more meaningfully. The mean of 17 for the failed campaign seems a little skewed by high outliers considering the median is 4. I think the same is true for successful campaigns.
   2. There is more variability in the successful campaign because the standard deviation is higher. It makes sense that the data for successful campaigns is more dispersed relative to its mean because there is a larger number of potential backers. For the failed campaigns, we will generally see a smaller number of backers because the project did not generate enough support and can only have so many backers until it succeeds.