Mathew Miller

Data Analysis Boot Camp: Homework 1

Instructor: Heather Bree

12-19-2020

Kickstarter Report

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

This data shows that successful funding of Kickstarter projects is highly biased towards the categories of film & video, music, and theater which will all be considered entertainment. This information may be misleading however because when looking deeper at the subcategories, most successful Kickstarter projects look to be mostly plays. One conclusion could be that looking at category data alone is not sufficient in identifying successful campaigns.

Since we have determined that plays tend to have a much greater rate of success in the Kickstarter world, we might presume that this would be a good avenue in which to “kickstart” like ideas. I think it would be safe to conclude that this would be a good area to do further analysis. One could analyze the genre of plays, venue sizes and cost for admissions to gain more insight. This information might give greater clarity on which types of plays tend to be more successful and which Kickstarter goals could confidently be set.

Among some of the other conclusions that can be drawn from this information, one of the most telling is how much the success rate of Kickstarter goals drops as the amount of the goal increases. From the range of $1 to $50,000 and above, the success rate drops over 50%. This is more than double the amount of the failed and canceled goals which both changed about 15% over the same range. While categories and subcategories show us where success might be, my conclusion is that this information alone is not sufficient in identifying Kickstarter success; categories should be looked at with other data sets to better understand successful goal trends.

1. What are some limitations of this dataset?

Some of the limitations of this data set is that it is overly broad. It does not categorize the pledgers themselves and the donation amounts are averages. Since we are not able further analyze this and other data, we are not able to remove outliers to better predict and recommend projects.

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

Other possible tables or graphs we can create are pie charts which make weight easy to see of one data set. Also, scatter plots are good for two data sets where we want to see the correlation of the distribution, the clustering, and outliers. Histograms are good for identifying outliers.