D Vickers

Homework 1

1/19/19

**What are three conclusions we can make about Kickstarter campaign given the provided data?**

1. The Kickstarter crowdfunding service has raised over 2 million dollars and the funding data provides an abundance of numerical information that can be grouped, graphed and analyzed in many ways. Three conclusions we can make from looking at the data are:
2. Out of the 2,185 successfully funded projects, almost 40% (839) of them were in the theater category followed by music with 25% (509). Journalism had 0 successfully funded projects.
3. Within the sub-categories, plays had the greatest amount of successfully funded projects although they also had the greatest number of failed outcomes within the sub-categories. Perhaps, the greater the attempts, the greater the probability of increasing your chances at funding.

**What are some of the limitations of the data set?**

Between the years of 2009 and 2017, approximately 54% of campaigns were successfully funded using Kickstarter. However, the data has some limitations as it does not answer other important factors that can tell the whole story. Here are a few questions that come to mind when looking at what was given:

* What correlation is there, if any, between the months with the greatest number of successful outcomes Feb/May/Oct/Nov. Are the spikes in these months due to the following?
  + Feb – tax returns, recovering from the holidays?
  + May-Mothers’ Day and donations given as gifts, increase in leisure time, increase in temperature, the real estate market?
  + Oct/Nov- elections, holiday season and family time, football.
* Overall, what was happening with our stock market, politics, the economy, oil, real estate during 2009-2017? How did these factors affect our data?
* The data also does not tell us how wide the Kickstarter “net” is, who their population and demographic include, etc.

**What are some other possible tables/graphs that we could create?**

While certain types of data lend themselves to different types of representation, this set of information could also be represented through circle graphs to display the percentage in category and subcategories that were funded, scatterplots to show relationships/correlations between data points and hierarchy charts to compare the different categories.