Taylor Greenhut

01 – Excel Homework

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

* Arts categories (such as theater, music, and film/video) have the best/most popular success on Kickstarter.
* Food & Journalism/Publishing categories have the lowest success in funding on Kickstarter.
* The US accounts for by far the most Kickstart campaigns, accounting for 74% of total campaigns (2nd highest is Great Britain with 15%).

**What are some limitations of this dataset?**

* This data set does not tell us if any failed/canceled campaigns ran a second campaign later on. Thus, a failed/canceled campaign could have later became successful after a relaunch.
* This data set also does not tell us if any campaigns changed their goals (to be higher or lower) during the campaign. This means that a campaign could have originally set an unachievable goal, but later lower it to become successful.

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

* Campaign success per goal amount
  + This table would let us get an idea of what the average successful and unsuccessful campaign’s goal amount was.
* Money raised per category & sub-Category
  + This would tell us which categories bring in the most/least total money raised.
* Backer count per category & sub-category
  + This would tell us which categories bring in the most/least donors.
* Average donation per category & sub-category
  + This would help see which categories raise the highest/lowest donations per donor

Bonus Section:

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

* I believe that the mean of the data summarizes the data more meaningfully. Being the average of the data set, I believe that it gives us a number the I closer to the majority, rather than selecting the middle number of the data set. For instance, for failed campaigns it would be unfair to look at the median of 4 and believe that this number represents most of failed campaigns. Many campaigns raised well over the number 4, having the mean (17.71) a better representation. This also applies to successful campaigns (mean = 194.43 while the media = 62).

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

* Both successful and failed campaigns have high variability, but successful has much more variability. Successful campaigns have a variance of almost 190x that of failed, and a higher standard deviation by almost 14x. This does make sense to me because looking at the mean, min, and max of both data sets, we can see that successful campaigns have a much wider variety of data points without even calculating the variance or standard deviation.