**University of Arizona Data Analytics Bootcamp**

**Excel Homework: Kickstart My Chart**

**Julia Richard**

**Analysis**

The first conclusion we can draw, from the provided data, is that more than half of Kickstarter campaigns meet their funding goals. This should make us stop and wonder about our data set, because we were provided background information that informed us “Of the more than 300,000 projects launched on Kickstarter, *only a third* have made it through the funding process with a positive outcome”.

The second conclusion we can draw is that approximately one third of Kickstarter campaigns are related to theater initiatives. This surprised me a little because my personal experience (bias) with Kickstarter has been through Facebook campaigns, seeking funding to manufacture some kind of gadget (like a picnic cooler that has a built-in blender). Based on our dataset, less than 0.5% of campaigns are related to “gadgets”.

A third conclusion we can draw from our dataset is that there is not a significant amount of campaign launch variability from month to month. However, there is an initial surge in activity at the beginning of the year (Jan, Feb), another uptick in the summer months (May, Jun, Jul) and a drop off in activity at year end (Dec). This may be of interest to a campaign that has flexibility in choosing a launch date. They could choose to launch in December and potentially have less competition for available crowd funds.

**Limitations**

A couple of limitations with our data set include (1) the data only represents campaigns through 2017 and (2) the data only represents 4,000 from more than 300,000 campaigns. As a result, we are not capturing the most recent (and potentially most relevant) data trends. Also, our data set only represents 1.3% of all Kickstarter campaigns. Trends in the remaining 98.7% of campaigns may tell a very different story.

**Other Tables or Graphs?**

One additional table or graph we could create and analyze could include the datapoints in the “Spotlight” column. I don’t know exactly what “Spotlight” is, but I imagine a campaign receives some kind of promotional help from Kickstarter, maybe for a fee or percentage of funds raised. Regardless of what “Spotlight” is, there seems to be a visually evident correlation between “true” and “successful” campaigns and “false” and “failed or cancelled” campaigns.

Another analysis we could run and illustrate is the difference between “pledged” and “goal”. Let’s call this difference “over subscription”. I am interested in knowing what kinds of things drive funding above the stated goal. This information could be used to develop a strategy or rule-of-thumb for establishing an initial funding goal. Analysis of the “Bonus” exercise indicates the highest percentage of “successful” campaigns set a goal of $5,000 or less. The highest percentage of “failed” or “cancelled” campaigns set a goal of $45,000 or more. It would be fun to take a deeper dive into these figures.

Lastly, it would also be fun to augment this dataset with some outside data that reflects social media outreach. We could then try to predict the optimum number of followers (or outreach) that correlates with reaching a funding goal.

**Bonus Statistical Analysis**

For our analysis of the number of backers for “successful” compared to “failed” campaigns, the median is a more meaningful summary of central tendency. For both “successful” and “failed” campaigns, there are a few campaigns with atypically high backer participation. These few data points are pulling the mean up for both data sets. The median is more accurate and, in both cases, illustrates that there are many campaigns that have a small amount of backer participation.

The “successful” campaigns have more variability in the number of backers than “failed” campaigns (a standard deviation of 844 for “successful” campaigns, compared to a standard deviation of 61 for “failed” campaigns). Numerically, this makes sense. The range in the number of backers for a “successful” campaign is more than 20 times greater than the range in the number of “failed” campaigns. Intuitively, this also makes sense. A “successful” campaign could have 1 backer, who funds the entire campaign, or have many thousands of interested backers, as with the 3D Doodle Pen. In contrast, a “failed” campaign could have 1 backer, but there are no “failed” campaigns with tens of thousands of backers.