**Module 1 Challenge**

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

Based on the data, we can draw the following three conclusions:

1. People working in art production-based industries (film and video, music and theater) are more likely to launch crowdfunding campaigns than people in other industries, like technology. Perhaps people working in artistic industries are more optimistic or have more collectivistic cultural views than those in other industries. Or perhaps the crowdfunding sites advertised more heavily to artistic communities thus drawing in more arts-based campaigns. More data are needed to explore the differences in campaign categories.
2. The United States had the highest number of total campaigns at 754. The other countries averaged just 39 total campaigns with Italy having the highest at 48 and Switzerland having the lowest at 22 total campaigns. Perhaps advertising was concentrated in the United States and had a smaller footprint in other countries. Perhaps other countries offer more grant-based programs for personal startups than the United States. More data is needed to explore the difference in the number of campaigns per country.
3. Campaigns with mid-ranged goals between 10,000 to 30,000 had the highest rate of success when compared to campaigns with goals below 10,000 or above 30,000. In addition, campaigns with a goal of only 100 had a 100% failure rate.

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

One limitation of the data is that we are only provided with the average donation amount, not the amount of each individual donation. When outliers are present, the mean is not always a good representation of the data spread.

**What are some other possible tables and/or graphs that we could create, and what additional value would they provide?**

We could have also created a graph that shows the relationship, if any, between campaigns that were listed as “Staff Pick” or “Spotlight” and the rate of success. This data would be a starting point for building a list of factors contributing to successful campaigns. This would be good marketing data for prospective crowdfunding clients.

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

The median in this data set summarizes the data better than the mean because the range and high standard deviation in both successful and failed campaigns indicate outliers.

A screenshot of a document

Description automatically generatedA table with numbers and text

Description automatically generated

There are 565 successful campaigns with a backer range of 7279 and a standard deviation of 1266. Likewise, there are 364 failed campaigns with a backer range from 6080 and a standard deviation of 960. As such, the median in both of these cases better represent the data.

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