Crowdfunding Project

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

* 1. From the summary tables, we can say that most of the campaigns were successful across all Parent Categories, except for Games, for which there are still some live campaigns.
  2. We can also say that most of the campaigns were successful across all Sub-Categories, except for Film & Video – Science fiction, which most failed.
  3. There seems to be a lot more crowdfunding campaigns for art related projects: theater plays, music and film & video.

# What are some limitations of this dataset?

* 1. There are big differences in the sample size for the different categories, some seem to be too small to be actually representative of those categories of campaigns’ rate of success.
  2. There are also big differences in the goals for the campaigns, so comparing the success of a campaign with a goal of, for example $1,700 vs. the failure of a $128,000 might not give us the most accurate result as to the impact of the type of campaign.
  3. We also see variations in the length of time the campaign was live, this might have a big impact on the ability to reach the goal. Maybe with longer live times, the failed campaigns could have been successful.
  4. This data set contains campaigns that are still live, which could change the percentage of campaigns that succeeded vs. failed based on the outcome, it’d be good to filter those out.
  5. We should also filter out the cancelled campaigns since we don’t know or can’t compare the reason for the cancellation.

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

* 1. I’d probably start by dividing the campaigns by goal size, probably by quartiles, and then do percentage comparisons instead of total counts.
  2. Depending on what we want to know about these campaigns we could make different comparisons.
     1. Filter the campaigns by country to see if different countries support different kinds of campaigns.
     2. Create a table that compares the % of successful campaigns created on each month, so we can see if there’s a trend regarding the month the campaign starts, the more likely it is to succeed, and we could filter that by country and/or year.
     3. We could make some comparisons also based on the staff picks and spotlight columns, with column charts, we could see if there’s a trend for the different kinds of campaigns to be more successful if they are marked as staff pick and/or spotlight.

# Statistical Analysis – Backer’s count

For the number of backers for successful and failed campaigns, the median represents the data better since the mean is being pulled towards higher numbers by the few outlier numbers.

The variability for both sets of data is close, they are both skewed to the right, towards higher number of backers, by some very big outliers. The successful campaigns have a slightly bigger variance. Having a bigger variance and slightly higher range makes sense since the higher the number of backers, the more likely a campaign is to succeed. It’s probably easier to get a lot of people to donate a small amount of money, than to have a few donating large amounts of money.