## Crowdfunding Goal Analysis

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

When considering successful or failed campaign outcomes, there were three factors that stood out in the data: campaign category, number of campaign entries, and the amount of money the backers pledged.

* 1. The top three categories are in theater, and practically tying for second place, film & video and music.
  2. All three categories have a higher number of campaigns, and consequently higher success rates, as well as failure rates.
  3. Additionally, those three categories, with outliers in technology, garnered the highest percentage funding, often surpassing their goal amount.

Perhaps as when playing the lottery, success in crowdfunding is a game of chance. The odds of winning or having a successful campaign are determined by the campaign category, the number of entries, and how much people are willing to contribute.

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

1. The start and end dates are in the 1980s. This is peculiar as online crowdfunding wasn’t around during that time.
2. There is limited information about the canceled campaigns.

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

1. Campaign outcomes by goal amount – a visual of this information could help determine if there is a range of goal money that correlates with successfully funded campaigns.
2. Campaign outcomes by average donation in US dollars – a visual of this information could help determine if there is a correlation between a certain amount of money donated to a campaign and its success outcome.
3. Campaign outcomes by blurb – a visual of this information could show if there is a specific phrase or key words that show up more often with successful campaign outcomes.

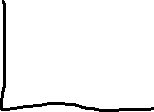
## Statistical Analysis

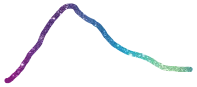
**Use your data to determine whether the mean or the median better summarizes the data.**

The data would be best summarized using the median because of the outliers in the data set. These outliers affect the mean, causing the data to be right-skewed.

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

The data shows that there is more variability with successful campaigns than unsuccessful campaigns. For example, when comparing the variance in the number of backers for both outcomes, the variance in successful campaigns is larger than the variance in unsuccessful campaigns. This makes sense because there are more data entries being analyzed with the successful campaigns which has led to a larger spread. Additionally, if looking at a sketch of a bell curve for both campaigns’ number of backers (see below), then visually the successful campaign’s distribution is wider (more variable) than the unsuccessful campaign’s distribution.







Successful campaigns

Unsuccessful campaigns

