## Data Analysis ##

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

1. **Categories Matter** – The theater category is highly competitive, however, lead to a higher percentage of failures than a category like film & video.
2. **Size of Project Goal** – The sweet spot for project goal amount is $1000-$4999 with a success rate of 83% given the large dataset of 231 projects. There are other projects that have a higher success percentage but have a small dataset.
3. **Project Start Date is important** – A project that starts in July is more like to see success than a project started in September.

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

Limitations of the dataset include:

* Incomplete data – we don’t have information regarding marketing strategies or project descriptions. This could help us determine what descriptions and marketing strategies led to successful or unsuccessful projects.
* Information about project users – It appears that some of the names on the projects are individuals and some are organizations. This data would help us to better determine if individuals or organizations are more likely to succeed.

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

* **Outcomes of Projects by Country** to see if there is an emerging market
* **Outcomes vs. Staff Pick or Spotlight** to see if the inclusion of being featured by the company helps improve outcomes.

## Statistical Analysis ##

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

The median best summarizes the data because it is less sensitive to outliers. The data is skewed right which tells me there are more small donor projects than there are large donor projects. The mean (851.15) gives you a false sense that most projects have close to 1000 donors per project. Sharing that story with customers could provide disappointment when a project doesn’t reach that level. Using the mean to say our projects see 200+ donors per project would be a truer statement.

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

There is more variability with successful projects. This makes sense because they have a wider range of outcomes. The outliers show that some projects perform very well while a majority just meet the success threshold. Unsuccessful projects are clustered tightly.

**A graph with a line and a line

Description automatically generated with medium confidenceA graph with blue and white lines

Description automatically generated**