Kickstarter Challenge- Analyzing donations in crowdfunding

Given a set of data regarding crowdfunding projects, we will be analyzing data to try and notate different trends. Specifically in this analysis, we are focused on Theatre Outcomes based on launch date, and Play outcomes compared to their original goals.

The purpose of this analysis is to see if we can find hidden trends based on launch dates. This can offer some assistance in getting donations based on dates that doners tend to be more generous. We are also comparing original goals with actual pledges to ensure expectations are realistic and obtainable.

The data set contained information organized into a category called “Category and Subcategory.” We separated this information out into two separate columns to be able to sort independently. We also converted the “Deadline’ and “Launched at” date stamps into a readable format so we could use this in our analysis as well.

**Outcomes** **based on Launch Date**: This data shows that the most successful theatre launch dates were in May followed by a slow decline over the months of June, July and August. Failed theatre outcomes showed consistent trends over these same months, peaking in October. Cancelled Theater launches were basically strait lined at a low number. January, June, and October had around the same number of failed campaigns. Theatre launches should be attempted in May through August and Goals should remain modest in the less than $1,000 to $4,999 range.

**Outcomes based on Goals:** The percentage of successful fundraising occurred in the less than $1,000 bucket. This makes sense since the average person would be able to donate within this range. The next most successful range was in the 1,000 to $4,000 range with only a 3% difference in success rate from the smaller bucket. As the ranges increased, the percentage of success dropped. There were only two successful collections in the greater than $50,000 range, and zero cancellations.

**Challenges:** One of the most challenging aspects of this data, was the different parent category and subcategory distinction and analytics. Trying to give meaningful data required careful segregation of each section.

**Limitations:** Some of the data limitations that I would notate would be in the area of “Failed and cancelled” reasons. If we could categorize why something is failing or reasons why cancellations are occurring, we could remove them from the data set if they are not representative of trends.

**Other thoughts:** In addition to reviewing launch dates, and goal setting, we could also look for trends regarding the subcategory and “Successes, failures, and cancellations” based off certain types of theatrical productions. We have the available data for “Animation, documentary, drama, science fiction, etc.” So, I would be interested to see if there was a certain type of genre that had more successes vs. others.