**Module 1 Challenge**

* **Given the provided data, what are three conclusions that we can draw about crowdfunding campaigns?**
  + July -> August sees a huge drop in success rate and an increase in fail and cancelled rate. This infers that there’s something about the shift to August from July that heavily affects the success rate of these crowdfunded campaigns.
  + Food & Games are the two lowest performing parent categories.
    - Food: 48% success rate (2nd lowest) & 43% fail rate (2nd highest)
    - Games: 44% success rate (lowest) & 48% fail rate (highest)
      * \*We ignored journalism as a parent category because of the small sample size.
  + Investors seem to favor the arts and entertainment spaces. The three parent categories with the most successful campaigns are theater, music, and film/video.
    - Theater: 187 successful campaigns
    - Film & Video: 102 successful campaigns
    - Music: 99 successful campaigns
  + The higher your goal amount the more likely you are to fail. Starting at $3500 and moving upwards the chance of failure starts to increase and goals greater than 5k have the second overall highest chances to fail and the lowest chance to succeed.
* **What are some limitations of this dataset?**
  + No stats or data on the performance of the companies after the crowdfund event.
    - It would be valuable to see the post-launch success of the higher pledged start-ups and determine if higher pledged amounts lead to eventual success.
  + No info on individual investments. Only total data when it comes to backers count and pledged amount. Data surrounding the individual investments would allow us to determine top investors and most attractive companies by showing the variance in investment amounts.
    - For example, some companies might have a lot of investors and a high pledged amount, but 99% of those pledges could be very small amounts and the 1% could be an enormous investment. This would skew the data and trick us into believing there’s strong confidence coming from a higher number of investors in a certain company.
  + The different currencies and no conversion rates.
    - A higher pledge amount in USD could be extremely higher or lower when converted to one of the many other currencies in the dataset.
* **What are some other possible tables and/or graphs that we could create, and what additional value would they provide?**
  + Pivot Table: Backers Count Grouped By Subcategory Filtered by Parent Category
    - This would show us which categories investors are flocking to and what, if any, trends there seem to be.
      * For ex, we could see if there’s been a huge recent uptick in investments in the mobile game category or if investors are no longer investing in shorts.
      * This would help entrepreneurs who may be considering which industry or category to enter.
  + Pivot Table and Line Chart: Backers Count Grouped Month
    - This would show us which months of the year are the most investors making investments into startups.
      * For ex, is there a specific season where investors have more money to invest? Do they run out of money any specific time of the year?
    - This would help start-up companies determine when to begin their crowdfunding campaign.
  + Bar Chart: Comparing the Number of Investments by Country
    - This would show us which countries are harvesting the most start-up companies and can show us recent shifts.

**Statistical Analysis**

* **Use your data to determine whether the mean or the median better summarizes the data.**
  + I believe the median would be the better number to summarize both datasets because the median lies in a better part of both histograms. The first category on both histograms represents the largest sample in the datasets and the median in both lies right in the middle of this column. The mean on the other hand lies outside of this first column.
* **Use your data to determine if there is more variability with successful or unsuccessful campaigns. Does this make sense? Why or why not?**
  + Both charts are strongly left skewed. However, successful campaigns have a higher upper bound than unsuccessful campaigns, which leads me to believe it has more variability.