* **Given the provided data, what are three conclusions we can draw about Kickstarter campaigns?**
  + Based on the outcomes by category statistics shown in (Figure 1), ***Music, Theater, Film & Video shows the highest success rates, Music of 77.14%, Theater 60.23%, Film & Video 57.69% in order not including live or canceled concerts***. These categories showed high success rates and low failure rates within most countries. ***Theater had the had the highest count of campaigns of 1393 with the lowest cancelation rates making it a more consistent category than the rest in most countries compared to the other categories having most of its success rate with the plays sub-categories.***

**Figure 1: Outcome Based on Category**

* + ***Looking more in depth to the subcategories within each category shown (Figure 2), plays had the highest success rates putting in the first place within the theater category and all other categories***. Rock witnessed a high success rate with no failure rate nor cancelation rates making it a more consistent subcategory than others with the theater category and all others. Finally, documentary, hardware and indie rock showed consisted rates with low to no failure rates or cancelation rates which all fall within the top three categories. ***Looking into both categories and sub-category analysis even though music had the highest success rates but was less funded than theater.***

**Figure 2: Outcome Based on Sub-Category**

* + ***Overall, between 2009-2017, Success rates dropped specially in the last few months of each year compared to failure rates. Figure 3 shows an increase in failure rates between***

***April-July and drop-in success rates within the same period. Overall cancelations stayed consistent most months.***

**Figure 3: Outcomes Based on Launch Date**

* **What are some limitations of this dataset?**
  + During analyzing the data, it was challenging to determine where the highest success rates occurred within each category, sub-category, and country. The data could be setup better with more metrics that better describing the success/failure/ cancelation rates within each category, country, sub-category. Its not clear how each category is being funded from which country.
  + there’s not enough evidence within the dataset to support the differences between each category and why each category performed better than the other. What is the interest per country and why based on their cultural background for example?
  + It is not clear from the data what each campaign goals to a success and what is the hopes behind each accomplishment within each category.
  + There is no explanation within the dataset on why success rates showed a decline compared to failure rates overall. More metrics and data to support can help analyze this case study.
* **What are some other possible tables and/or graphs that we could create?**
  + Line Chart analyzing successful/failed/ live/canceled by year and not month only between 2009-2017 to determine the driving factor for each rate to act the way it acted based on Figure 3.
  + Bar Char for each sub-category in a category separately from the others to support our conclusions on why music for example such success rates had even though it had higher cancelation rates than theater. Why plays within theater had the highest success rates?
  + Box Plots to determine your outliers, your center of data, variance, and standard deviation. This can help determine the type of data you are dealing wells and can lead to support your conclusions with better reasoning.
  + Within the pivot tables, displaying the timeline within each graph, sub-category to each category within the same pivot table, the goal, and average donations to help better understand the driving factor for the success rate/failure rate for each category within each country.

**Bonus:**

* **Use your data to determine whether the mean or the median summarizes the data more meaningfully**.
  + Based on the resulting statistics of the data, the median represents the data better than the mean. There are several reasons for this:
    - Mean is three times the median which indicates the data is very spread apart indicating that taking the middle is of the data set is more accurate than averaging it for better resulting analysis.
    - The difference between the maximum and minimum value is 26457 indicating how spread out the data and indicators for outliers
    - Variance and Standard deviation explain how spread each data point from the middle of the dataset being the median so averaging the data will lead to inaccurate results
  + This explanation applies to both successful and failed outcomes
* **Use your data to determine if there is more variability with successful or unsuccessful campaigns. Does this make sense? Why or why not?**
  + The resulting standard deviations explains the higher variability in successful campaigns compared to the unsuccessful campaigns. Successful campaigns had more data points but showed wider spread in data than the successful data. The higher variability in both resulting data makes difficult to approach the data to make sense out of it and indicate higher volume of outliers within the dataset. Overall using median over mean can help direct us to understand the conclusions more than the mean.