**Kickstarter-Analysis Report**

**From the analysis using Pivot Tables:**

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

1. Overall, more than 50% of the Kickstart projects went successful. Out of all the categories, projects under **THEATER**, **MUSIC** and **FILM & VIDEO** ranked top 3 number of successful projects. These also show that people prefer the **entertainment-related** projects more.
2. **Theater** projects went super successful in **Asian** countries; **technology** projects drew most attention from **European** countries, but they did not have high successful rates. As for **American** countries**, theater**, **music** and **film & video** attracted more attention. However, although **theater** projects have the greatest number of successful outcomes, it is more likely to pledge money for **music** projects because the percentage of successful outcomes hits the most under the music category. American people (especially the **US** people) are more likely to spend for **rock** (and **indie rock**) music.
3. **Theater** project went super successful (with zero failure and other) in 2013 and prior. **Music** projects were generally successful in all year. The rest of the categories had closed numbers of successful and failed outcomes. In addition, theater's successful projects peaked during **summer time**, which shows that **weather** impacts people's willingness to visit theaters. The rest of the successful projects have relatively flatter curve across the years, but data shows that people were not likely to spend money in **December**.

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

The sources of data come mostly from the US, which consist **3038** data entries out of **4114** total entries. This shows the bias of the raw data and therefore can impact the accuracy of the conclusions drawn from the data analysis. For example, one of the conclusions that I drew above is that the music projects that focus on rock music tend to have higher chance on successful results. However, since most of the data were collected from **US**, the conclusion is more likely showing that people from the **US** (not necessary the people from worldwide) prefer rock music.

**What are some other possible tables and/or graphs that we could create?**

1. **Scatter Plot graph** - we can create a scatter plot graph with the y-axis being the number of successful/failed/live/canceled projects and the x-axis being the goal amounts or number of backers, in order to determine if there are any positive or negative correlations/trends between them.
2. **Summary table or in a Pivot Table** - Firstly, use the project ended dates minus the project created dates to calculate the durations of all projects. Next, filter the projects by their categories and then by the project states (i.e., successful or failed), and calculate the means and standard deviations of the duration of the projects in each category and state. As a result, we should have a summary table to show how much time on average would a successful project need in certain category.

**From the analysis using Pivot Tables:**

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

Median summarizes the data more meaningfully. Based on our data, both successful and failed campaigns have their means greater than the medians, which indicates that both the distributions are skewed. Also, based on the calculated Z-Scores, extreme values exist and have impact on the means. Thus, median is more meaningful in this case as it is used to measure data when distribution is skewed or when extreme values take place.

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

Successful campaigns have more variability than unsuccessful campaigns because of the much larger variance and standard deviation. This makes sense because in the common world, successful campaigns usually have more variance in turns of their results -- for example, a super successful campaign can be so popular when it has 50 times more backers than the other successful ones. On the other hand, failed projects most likely have few, sometimes zero, backers. So, in the failed case, there is likely less variance to it. Additionally, the successful projects have no upper bound (there is no limit to the number of backers) whereas the failed ones have a clear lower bound (zero backer). Hence, it makes sense that successful campaigns have more variability than the unsuccessful ones.