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

1a. One major conclusion we can draw about Kickstarter campaigns is that the Category of **“Theater”** seems to be very successful with about **839 successful** campaigns as opposed to the lowest category of **“food”** with an abysmal **34 successful** campaigns. Furthermore, it seems that the majority of successful **“Theater”** campaigns have come from the US with a total of **525 successful campaigns**. However, it should be noted that most Kickstarter campaigns came from the US (regardless of state), with a Grand total of **3038** out of **4114** total campaigns worldwide.

1b. Since we already looked at the parent category of “Theater” we can take a closer look at the subcategories. Amongst the **sub-categories** of **musical, plays, and spaces**, **Plays** were the most successful at **694 successful** campaigns, followed by **85 for spaces** and **60 for musicals.**

1c. Surprisingly, across all years, categories and countries, there did seem to be a trend in terms of which month of the year had the most successfully completed campaigns. The month of **May** was the highest with a rate of **234** successful campaigns and **December** was the lowest with a rate of **111** successful campaigns.

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

One major limitation of this dataset is that it does not tell us which State (i.e. California, Texas etc.) that these campaigns were run out of. With the United States being the Country with the most **campaigns (3038/4114)** it would be nice to know which of the States had the most successful campaigns vs. the most failures etc. This could provide some very important information for such a big Country.

Another possible limitation is the fact that we may be limited in terms of where the majority of the backers are from. This dataset could have included the country of origin from where the backers are from initially, and POSSIBLY more demographic information on the backers themselves. This can include age, gender, ethnicity, religion etc. This can give us invaluable information in terms of who is backing which project and see if there are any specific trends as to why that may be.

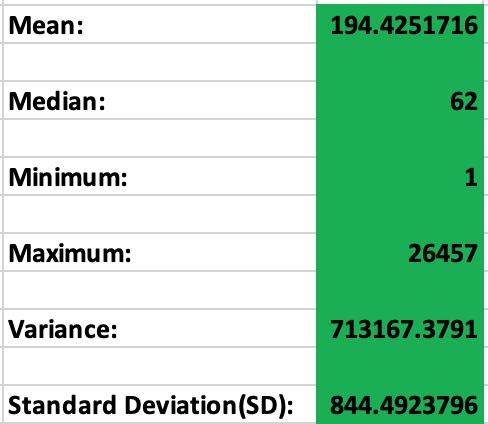
1. **What are some other possible tables and/or graphs that we could create?**

One interesting column of data that we were not asked to look at was the “Staff Pick” Column. Through this we could possibly see (by using staff pick as a filter) whether or not what the staff thought was a good idea actually came into fruition and became a successful campaign. Of course, we could also look at this by the category as well. That would be interesting to me.

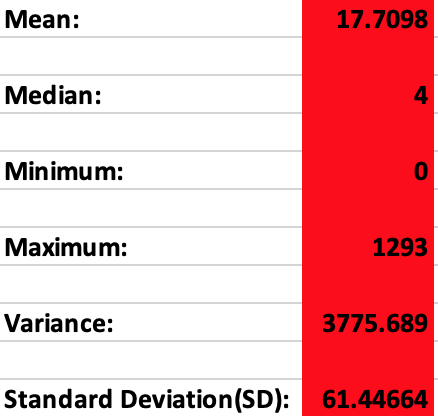
**Bonus Statistical Analysis Question Answers (See Below)**

**Bonus Statistical Analysis Questions:**

Successful Outcome Descriptive Stats



Failed Outcome Descriptive Stats



* **Use your data to determine whether the mean or the median summarizes the data more meaningfully.**
  + Personally, I believe that the **median** summarizes the data more meaningfully simply due to the fact that there seem to be so many outliers when it comes to the number of backers. This is especially true when it comes to the **successful** **outcome backers** because if you look at the **range**, it ranges from a **minimum** of only **1** backer to an outstanding **maximum of 26,457 backers**. These numbers are clearly outliers and will thus skew the mean which is why the **median would be better**. This same logic applies to the Failed outcome backers, however, to a less severe degree.
* **Use your data to determine if there is more variability with successful or unsuccessful campaigns. Does this make sense? Why or why not?**
  + There is more variability in **successful** campaigns. This makes sense because of the fact that both the variance and standard deviation are significantly higher than that of the unsuccessful campaigns
  + In general, when we look at the variance it describes how much variation exists in the data and the number, we get from the successful campaign section is pretty enormous.
  + Furthermore, we can already tell from the minimum and maximum values that there theoretically will be more variance in the successful campaigns
  + Also, it makes sense that the “failed” campaigns will have a number of backers that are closer together, mainly because we are probably dealing with a lower number of backers in general close to 0.