**Excel Challenge Report**

**2/27/2023**

**Meardreed Vilméus**

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

3 conclusions we can draw after analyzing the crowdfunding data are listed below.

* The top categories by number of crowdfunding seekers are theater, film, and music. Although journalism has a very low number of projects, they show a 100% success rate.
* July and January are the month with the most launched projects.
* Projects with a goal of more than $50000 are less likely to be successful with a 37% chance of success. Whereas projects with a goal ranging $15000 to $25000 and $30000 to $34999 tend to have a 100% success rate. Most projects have a goal of $1000 to $10000. Although this would be more accurate after universalizing the currencies.
  + **What are some limitations of this dataset?**

The first limitation is the currency. We would have to convert all currencies to US dollars, assuming we are in the US, to properly evaluate the funding by goal to see if the brackets proposed in the outcome by goal are truly accurate.

The length of the project from launch to deadline would be helpful in determining if projects with a lengthier span tend to raise more money. Or if the length of fundraising is proportionate to the goal.

* + **What are some other possible tables and/or graphs that we could create, and what additional value would they provide?**

To dwell on the length of project limitation, we can add a column with the fundraising length to the Crowdfunding table. Import the length data, the goal and parent category to the new table with the outcome by backers. And see how the length and goals impact the number of backers for a project.

We can also look at the goals by project category to check which type of projects attracts the most investors with a pivot table.

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

The data ranges from 16 to 7295 for the successful backers, the mean is 851 and the median is 201. We can observe in the box and whiskers that the mean is a better representative of the data than the median. However, both datasets are skewed to the right per scatter chart and the means are heavily affected by the outliers.

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

In this particular case, both scatter graphs show similar trendline and the R2 values are close. However, the variability of the successful campaign is greater and that makes sense because the size of the data is larger which open doors for more outliers.