**Class**: USC-LA-DATA-PT-08-2020-U-C

**Student Name**: Lisa Man

Excel Homework Assignment – Excel Challenge

For each requirements/asks in the homework assign, see comments in **blue** to locate the answers.

**1. Kickstarter Excel**

1.1 Requirements - Use conditional formatting to fill each cell in the `state` column with a different color, depending on whether the associated campaign was successful, failed, or canceled, or is currently live.

See tab 1 Starter Data, column F

1.2 Requirements - Create a new column O called `Percent Funded` that uses a formula to uncover how much money a campaign made to reach its initial goal.

See tab 1 Starter Data, column O

1.3 Requirements - Use conditional formatting to fill each cell in the `Percent Funded` column using a three-color scale. The scale should start at 0 and be a dark shade of red, transitioning to green at 100, and blue at 200.

See tab 1 Starter Data, column O

1.4 Requirements - Create a new column P called `Average Donation` that uses a formula to uncover how much each backer for the project paid on average.

See tab 1 Starter Data, column P

1.5 Requirements - Create two new columns, one called `Category` at Q and another called `Sub-Category` at R, which use formulas to split the `Category and Sub-Category` column into two parts.

See tab 1 Starter Data, column T for Category and column U for Sub-Category.

1.6 Requirements - Create a new sheet with a pivot table that will analyze your initial worksheet to count how many campaigns were successful, failed, canceled, or are currently live per \*\*category\*\*.

See tab 2 Pivot category state

1.7 Requirements - Create a stacked column pivot chart that can be filtered by country based on the table you have created.

See tab 2 Pivot category state

1.8 Requirements - Create a new sheet with a pivot table that will analyze your initial sheet to count how many campaigns were successful, failed, or canceled, or are currently live per \*\*sub-category\*\*.

See tab 3 pivot Subcategory

1.9 Requirements - Create a stacked column pivot chart that can be filtered by country and parent-category based on the table you have created.

See tab 3 pivot Subcategory

\* The dates stored within the `deadline` and `launched\_at` columns use Unix timestamps. Fortunately for us, [there is a formula](https://www.extendoffice.com/documents/excel/2473-excel-timestamp-to-date.html) that can be used to convert these timestamps to a normal date.

1.10 Requirements - Create a new column named `Date Created Conversion` that will use [this formula](https://www.extendoffice.com/documents/excel/2473-excel-timestamp-to-date.html) to convert the data contained within `launched\_at` into Excel's date format.

See tab 1 Starter Data column Q

1.11 Requirements - Create a new column named `Date Ended Conversion` that will use [this formula](https://www.extendoffice.com/documents/excel/2473-excel-timestamp-to-date.html) to convert the data contained within `deadline` into Excel's date format.

See tab 1 Starter Data column R

1.12 Requirements - Create a new sheet with a pivot table with a column of `state`, rows of `Date Created Conversion`, values based on the count of `state`, and filters based on `parent category` and `Years`.

See tab 4 Pivot Date Created\_State

1.13 Requirements - Now create a pivot chart line graph that visualizes this new table.

See tab 4 Pivot Date Created\_State

1.14 Create a report in Microsoft Word and answer the following questions.

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

2. What are some limitations of this dataset?

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

See Word file title Report for KickStarter Analysis

**2 . Bonus**

2.1 Requirements Create a new sheet with 8 columns:

\* `Goal`

\* `Number Successful`

\* `Number Failed`

\* `Number Canceled`

\* `Total Projects`

\* `Percentage Successful`

\* `Percentage Failed`

\* `Percentage Canceled`

See tab 5 Bonus line chart

2.2 Requirements - In the `Goal` column, create 12 rows with the following headers:

\* Less than 1000

\* 1000 to 4999

\* 5000 to 9999

\* 10000 to 14999

\* 15000 to 19999

\* 20000 to 24999

\* 25000 to 29999

\* 30000 to 34999

\* 35000 to 39999

\* 40000 to 44999

\* 45000 to 49999

\* Greater than or equal to 50000

See tab 5 Bonus line chart

2.3 Requirements - Using the `COUNTIFS()` formula, count how many successful, failed, and canceled projects were created with goals within the ranges listed above. Populate the `Number Successful`, `Number Failed`, and `Number Canceled` columns with this data.

See tab 5 Bonus line chart, column B, C, and D

2.4 Requirements - Add up each of the values in the `Number Successful`, `Number Failed`, and `Number Canceled` columns to populate the `Total Projects` column. Then, using a mathematical formula, find the percentage of projects that were successful, failed, or canceled per goal range.

See tab 5 Bonus line chart, column E, F, G, and H

2.5 Requirements - Create a line chart that graphs the relationship between a goal's amount and its chances at success, failure, or cancellation.

See tab 5 Bonus line chart

**3. Bonus Statistical Analysis**

3.1 Requirements - Create a new worksheet in your workbook, and create a column each for the number of backers of successful campaigns and unsuccessful campaigns.

See tab 6 Bonus\_Statistic Analysis

3.2 Requirements - Use Excel to evaluate the following for successful campaigns, and then for unsuccessful campaigns:

\* The mean number of backers.

\* The median number of backers.

\* The minimum number of backers.

\* The maximum number of backers.

\* The variance of the number of backers.

\* The standard deviation of the number of backers.

See tab 6 Bonus\_Statistic Analysis Table A

3.3 Requirements - Use your data to determine whether the mean or the median summarizes the data more meaningfully.

Answer - For both dataset (successful and failed states), using the median to determine the number of backer count is more meaningful. This is because the number of campaigns with backer counts below and under the median is about the same. Using the mean is less meaningful for these two datasets because the percentage of campaign below the mean is significantly high, approximately 80% of the campaigns are below the mean. Therefore, using the median is more predicative.

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

Answer – There is more variability in successful campaign than in the failed campaigns. The standard deviation and the interquartile range found in the successful campaign are higher than the failed campaigns. The dataset makes sense. From the review of the successful campaign dataset, there were 392 successful campaigns that had reached at least 100% funding goals. Also, there were 244 (11% over the total successful population) successful campaigns with backers count over 309 (upper 3rd boundary quartile), and the highest backers count was 26,457. This indicates that the successful campaigns had more variances.

From reviewing the dataset over the failed projects, none of the failed projects were able to meet the funding goals (less than 100%). The median of backer counts for failed campaigns was only 4. The lower of the funds contributed to reach the project goals, the lesser of the backers count. It is highly resulted in an unsuccessful campaign.