**Shadee Tabasi**

**Excel Homework: Kickstart My Chart + Bonus**

**November 10, 2020**

Table of Contents

[Primary Excel Homework 1](#_Toc55998589)

[Given the provided data, what are three conclusions we can draw about Kickstarter campaigns? 1](#_Toc55998590)

[What are some limitations of this dataset? 1](#_Toc55998591)

[What are some other possible tables and/or graphs that we should create? 2](#_Toc55998592)

[Bonus Statistical Analysis 3](#_Toc55998593)

[In the bonus statistical analysis exercise, determine whether the mean or median summarizes the data more meaningfully. 3](#_Toc55998594)

[Use your data to determine if there is more variability with successful or unsuccessful campaigns. Does this make sense? Why or why not? 4](#_Toc55998595)

# Primary Excel Homework

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

Based on the provided data, we can make the following assumptions:

1. The highest volume of Kickstarter campaigns appear to be theater-related – both globally and in the US – signaling a crowded space and market. Plays in particular are responsible for the greatest portion of those campaigns.
2. Journalism Kickstarters have had the least amount of success as all campaigns within this date range appeared to have been canceled. These all fall under the sub-category of “audio” and include radio shows and podcasts.
3. It appears that the number of successful campaigns decrease over the course of a year – with the most successful months happening in H1 (January 1 – June 30) and a decreasing volume of successful campaigns in H2 (July 1 – December 31).

## What are some limitations of this dataset?

1. One limitation or irregularity within this data set is that some of the years (2017, 2009) are missing data for a few months of the year. This may be skewing data – especially in the last line graph where we are trying to show Kickstarters over time.
2. For storytelling purposes, not knowing the reason for canceled Kickstarter campaigns, may also limit us in how we explain and interpret the data.
3. In comparison to the other “states”, we have a very limited amount of live campaign data to analyze.
4. In this particular exercise, we hadn’t converted currency to one standard (e.g. AUS to USD) and therefore, we are not comparing apples to apples when talking about campaign investments.

## What are some other possible tables and/or graphs that we should create?

A couple of other tables and/or graphs that we can create to tell an interesting story with this data set:

1. A table and bar chart that show the **average donation per backer by Category** could tell us what campaigns people are most likely to invest the greatest amount in. In the example below, we see that technology campaigns received the highest average donation per backer. We can even filter it by country to see how that may fluctuate globally.

Chart, waterfall chart

Description automatically generated

1. A line graph showing the **total amount pledged to Kickstarter campaigns over the years** will show us trends in backer spending for a specific category or overall. For example, we see an increasing amount spent overall over the years, but a decreasing amount for specific categories (e.g. journalism). See filters at top which can help us view category-specific data or country-specific data.

Graphical user interface, chart, line chart

Description automatically generated

# Bonus Statistical Analysis

## In the bonus statistical analysis exercise, determine whether the mean or median summarizes the data more meaningfully.

Through an analysis of the data, median should be utilized as the measure to summarize the data most meaningfully. When breaking down the data into quartiles, solving for the interquartile range (IQR) and parameters for outliers (Q1-1.5\*IQR AND Q3-1.5\*IQR), we find that there are many outliers in the data set (e.g. MAX # of backers) and the mean is likely to be skewed.

Table

Description automatically generated

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

There appears to be more variability within successful campaigns. This may make sense because successful campaigns may require varying levels of investment and number of backers, while failed/unsuccessful campaign are likely to have a much lower volume of backers across the board.