Steven Rojano / September 20, 2020 / Assignment # 1 / Data Analytics Bootcamp – Arizona

**What are three conclusions we can make about Kickstarter campaigns given the provided data?**

1. Theater is the front runner throughout all countries in the parent category, with well over half of its campaigns having a successful outcome. The most popular class of theatre is enjoyed from Plays, which account for roughly 14% of the success within the group.

A screenshot of a cell phone

Description automatically generated

**A picture containing sitting, pencil, person, table

Description automatically generated**

The first half of the year resulted in the highest production of Kickstarter campaigns for both Theatre and Non-Theatre Category’s, with a peak for theatre from April to July. When the Theatre group is removed, the months from January – May had the highest number of Non-Theatre Kickstarters.

A close up of a map

Description automatically generated

A close up of a map

Description automatically generated

1. Notably, the peak year of Kickstarter Crowdsourcing didn’t see a significant increase until 2013 – 2016. This could be that Kickstarter was still fairly new (founded in 2009) or may be partially due to the financial crisis.

A close up of a map

Description automatically generated

1. Furthermore, there is a trend of those startups that set a goal to be funded less than $1000 and resulted in over 70% having a successful project. The highest range listed at more than $50,000 resulted in nearly 60% of failed project.

A screenshot of a cell phone

Description automatically generated

A close up of a map

Description automatically generated

3. Lastly, the following summary statistic table represents the number of successful and failed backers that each startup had in this dataset. The mean should be used to summarize the data more meaningfully, because it is the most central measurement of the data set for each group. Standard Deviation differs from each group with the successful being more spread out and the failed being clustered together.

A screenshot of a cell phone

Description automatically generated

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

First, the data set does not represent the true number of Kickstarter campaigns. We are given information that only 1/3 of the 300,000 Kickstarter campaigns experience success, and the data only contains 4100 of those campaigns.

Additionally, there should be more information on each Kickstarter mission. The blurbs. The attributes to the class descriptions in the column with the header blurb are not sufficient information that clarifies the reason why this Kickstarter program became successful. Having a metric that measures the longevity of the company’s success would be a key factor in determining what industry/parent category has experienced true success.

Finally, with that being said, I do believe that you can take similar approaches when applying independent research in other industries. The data set is a good foundation on what you can do to manipulate data. This helped me develop a good approach to discovering trends and creating several different data visualization techniques when presenting information to any group of people. Though the data did not represent the true amount of campaigns that occur, I was successful in telling a story and analyzing the data set given.