Insights form the Kick-Start Data Set

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

* If you want to put your money into Kickstart projects, the “Theater” is a good category to take a bet. From the category chart, theater, has the highest count of successful campaigns (839 projects are successful) compared to the other categories and “plays” has the highest count of successful campaigns compared to other subcategories.
* If you want to launch a project in Kickstart, the possibility of success is higher than failed or canceled. From the data set, the number of successful projects is 2185 which is higher then failed projects which its number is 1530. The canceled project is only 8% (“canceled” is 349, total number is 4064) of the total of projects.
* If the projects’ goal is less than $4999, then the percentage of successful projects will be higher than 50%. If the projects’ goal is higher $44999, then the failed percentage will higher than 40%. So, the conclusion is the smaller goal’s projects will have higher possibilities to success.

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

* Should add more location information, such cities or state
* More detailed and precise category definition to fit the project types
* The “live” state is an uncertain feature when we want to predict the projects whether is “successful” or “failed” because it is still on-going.

1. What are some other tables/graphs that we could create?

* We can plot the following tables or graphs
* Project success rates over time by category
* How much money was successfully raised over the certain year
* Project Success rates by countries
* Project Duration vs Successful rate
* How much is pledged in percentage of the goal depending on the duration of projects