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

1. There are more successful campaigns in Theatre category than in any other category.
2. In Theatre category, plays get the most successful campaign. 82% of the total successful campaigns are in Theatre category.
3. Kickstart campaigns gained more popularity in the years 2013-2016

What are some limitations of this dataset?

1. The currency of the campaigns is based on the country it was campaigned. However, the data will need to be converted into uniform currency if an overall comparison is required.
2. Further breakdown of states or territories within each country is missing. States and territories breakdown will give us more insight into preferences of categories and sub categories by geography, metro vs regional breakdown etcetera.

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

1. A line or bar graph based on percent funded could give us more insight into what categories and sub-categories people tend fund more.
2. Scatter plot graph can be created to show relationship between the project being in spotlight or not and the chances of it being successful
3. Scatter plot graph can be created to show relationship between the project being a staff pick and the chances of it being successful

 Use your data to determine whether the mean or the median summarizes the data more meaningfully.

When looked at the mean and median of the dataset, the average backers for successful campaign is 194 and 18 for failed. However, the maximum of the same range is 26,457 and 1,293. Therefore it can be concluded that there are few outliers.

Further analysis of the outliers needed to be done in order to remove the outliers.

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 is more variability in successful campaigns in comparison to unsuccessful campaigns. However, the variance and standard deviation will only provide limited insight for the comparison as the definition of a successful campaigns is based on the dollar value raised and number of backers for each campaign. The variability of the dataset will not provide the reasons why a campaign was successful or not.