**Challenge module 1**

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

As a result, obtained from the pivot table, the sub-category "plays" had the highest success rate among other sub-categories. We can also infer from the results obtained in the pivot table for parent categories that theatre projects were both the most successful and the most failed projects simultaneously. Based on the calculated data, we can conclude that there are more successful backers (46,160) than failed backers (15,599).

1. What are some limitations of this dataset?

Some datasets, such as "staff\_pick" and "spotlight," were not necessary but are still present in the dataset, creating a large but unused dataset. Another limitation is the absence of actual factors like economic conditions, politics, and other environmental factors that can influence behavior and potentially impact crowdfunding success.

1. What are some other possible tables and/or graphs that we could create, and what additional value would they provide?

Another possible table we can create is the duration of the project, from the launch date to the data conversion endpoint. This table will provide additional information on the time period of both successful and failed projects. Additionally, we can filter the dataset by currency. We can create a table with datasets indicating the percentage of live projects currently running. Furthermore, we can generate a table based on the success, failure, cancellation, or current live status of campaigns per sub-category, providing additional information based on each sub-category.

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

As there are more variance in successful and unsuccessful campaigns, as with the result of standard deviations, there are other factors that affect the for campaigns successful and unsuccessful.

1. A brief and compelling justification of whether the mean or median better summarises the data.

The large difference between the mean and median for both successful and failed backers,

Mean for Successful Backers: 733, Mean for Failed Backers: 503

Median for Successful Backers: 198, Median for Failed Backers: 88 and the presence of relatively high maximum values, the median may be a more robust summary measure in this context. It provides a value that is less influenced by extreme observations, offering a better representation of the typical number of backers in both successful and failed campaigns.