Module 1 Challenge

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

* The subcategory of crowdfunding campaigns showed a lack of influence on the outcome. The sub-category plays had the greatest number of successful and failed campaigns.
* Campaigns with successful outcomes had lower average goal amounts compared to campaigns with failed outcomes.
* Campaigns with successful outcomes had greater backercounts on average than campaigns with failed outcomes.

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

* Currency. Multiple countries and currencies were represented in the dataset which could cause misrepresentation in data when comparing goal and pledged amounts if we do not create a base currency to compare.
* The sample size is a limitation of the dataset. It is difficult to prove whether or not a crowdfunding campaign is more likely to be successful based on the category or sub-category, since the sample included a limited number of campaigns in categories such as food, games, and journalism.
* Since each campaign started and ended during different years and months, the data may not provide an accurate insight into success by month.
* Outliers in the dataset can distort the averages and misrepresent the data.

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

* Recreate the dataset with a base currency to standardize the goal and pledged amounts in order to accurately compare average donations, goal, and pledged amounts by outcomes.
* A bar chart showing the average pledged amount by sub-category could provide insight to which sub-categories attract more funding from donors on average.

**Use your data to determine whether the mean or the median better summarizes the data.**

* The medium best summarizes the data. Due to outliers in data, the mean is positively skewed for both the successful and failed backers count as it is greater than the medium. The medium represents the data more accurately without being skewed by the outliers.

**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 campaigns with successful outcomes. This makes sense when analyzing campaign outcome based on backers count since there is a greater difference between the minimum and maximum of the successful backers count to the minimum and maximum of failed successful backers count.