1. Out of the 1000 crowdfunding projects, 565 of them were considered successful. Looking at the “Percentage Funded” column that we created, many of these “successful” projects were pledged more than two or three times the initial goal. While the pivot charts we created provided many insights into what makes a crowdfunding project likely to be a success, three stood out to me.
2. Theater and Film & Video were the two parent categories with the most crowdfunding projects and most of them were successful. If I were to do a more detailed exploration of this data, I would look to see how the goals of theater projects compared to the other categories. Instead, when I looked at the sub-categories, Theater only had “Play” as a sub-category which leads me to believe that these projects were largely for community theater programs and may not provide as much relevant information to organizations as Film & Video would. The Film & Video and Music categories were very similar. However, having these categories as the Top 3 indicates that crowdfunding is most common and most successful in areas where there is already a strong internet community.
3. The number of crowdfunding projects by year remains relatively stable between 2010 and 2019. The data collected for this dataset stops in 2020. Economic, public health, and cultural considerations aside (I’m aware that this is a serious limitation), crowdfunding should be expected to remain an accepted and effective means to allocate funds in the United States specifically.
4. Across the top three categories, most successful crowdfunding projects were launched in July and very few were started in the spring whereas most failed projects were launched in May or January. The time of year is likely related to how many backers will be reached but more statistical analysis is necessary to confidently draw this conclusion.
5. This dataset provides a lot of good information about categories and how much of a project was funded, if at all. I would say that the data is lacking specificity and diversity when it comes to the sub-categories. For example, crowdfunding technology would be an interesting area to analyze but it only includes “web” and “wearables” as sub-categories. There is no way of knowing what exactly what these categories encompass but technology is such a large area of production, there should also be more than two sub-categories.

In addition, the dataset does not provide any information about the backers such as where they are from, their profession or income, etc. Information like this would be incredibly useful to organizations looking to reach more people and fund their projects quicker.

Finally, the dataset is outdated. For the purpose of this module challenge I doubt this is a relevant issue but outside of a classroom setting, the most recent data would be the most useful and most reliable. It would also have the most context so there would be more conclusions to draw from it.

1. If I were to do a deeper analysis of this data, I would create a pivot table and chart for “Date Ended Conversion” so I could see which projects were funded the quickest. I would also create a bar chart comparing how expensive a project was to how many backers it took to fund it, and what the average donation was. Finally, I would create a chart that showed how much money was being collected by category. For example, Theater had the most successful projects, but could that be because the goal for these projects was so little compared to areas like technology? More information about the backers and funds would be useful.

Bonus

1. In this case, the median number of backers provides more insight than the mean. The average number of backers of successful projects is 851 which is misleading given the maximum number of backers is 7295 and the minimum is 16. The median is only 201 which means that more than half of the projects had 201 backers or less. The large maximum value skews the data and makes the average misleading.
2. There is more variability in the successful projects. This is likely because there were more of these projects and therefore more opportunity for outliers as seen in the difference between the mean and median for successful outcomes.