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

The first conclusion is that theater/plays is the most popular area that people are looking to be crowdfunded. Further research into the cause is needed, but a possible outcome is because of the rise of technology it has reduced the popularity and therefore funding of this area of entertainment where its double that of film or music. Second conclusion is that crowdfunding is not the most reliable source of funding as shown by the success rate. For the three most popular parent categories, the success rate sits around the 56% success rate (film and video 57%, music 57%, and theater 54%), which can be quite risky for both the investors and people trying to raise the money. A third conclusion that we can draw is that even though the number of campaigns goals ranging between $10,000 and $50,000 is small, it has a high success rate of 79%. Again, more investigation and additional sample size for those target goals would be helpful to gain confidence in those results, but it seems that investors are a lot more willing to invest into projects of the middle ranges.

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

One limitation, which was mentioned in the previous answer, is that sample size for the middle ranges of funding goals ($10,000 to $50,000). The success rate is looking positive for those areas, however a larger sample size is needed to have confidence in the conclusion. Even though the current data population shows the analyst the categories in which crowdfunding is most popular, if there was a client wanting to decide if they wanted to invest into an area outside of film and video; music, and theater, the data does not have a very big sample population to give the investor confidence that their money will lead to a successful outcome.

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

Another graph that could be valuable would be comparing the number of campaigns to the year they were created. The analyst could then try to link the relationship between the number of campaigns for each year and the significant events and market performance of that year. It could also be broken down into number of successful or failed campaigns related to its respective year. For example, it may show that in 2016 the rate of success was low as perhaps the value of the dollar was low or interest rates were high and so investors were a lot more cautious spending money to ensure they had savings for cost of living and not wanting to risk it. An area that the data was not used was the average donation amount. If we compared this to the category, we could draw conclusions as to if there are any trends related to the industry perhaps showing that investors are sympathetic to certain areas of campaigns and are willing to invest greater amounts. Additionally, we could compare this to outcome of the campaign to show if the there is a relationship there. For example, it may show that the larger the average donation size the likelihood that the campaign is successful increases, whereas if the donation size is small it is a riskier investment,

From the bonus section, we can see that an additional source of information would be creating a box and whisker plot using the table created related to the number of backers. This chart would show that the quartiles are heavily weighted towards the lower end of investors, showing that crowdfunding is perhaps not a popular source of investment, and high number of backers are quite rare.

**BONUS QUESTIONS**

**Does the mean or median better summarises the data?**

If we create a box and whisker plot, we can see that the mean is more centered to the inter-quartile range, whereas the median is heavily weighted to the first quartile. It can also be seen that the third quartile is large and so the mean better represents this portion of data.

**Is there more variability with successful or unsuccessful campaigns? Does this make sense? Why or why not?**

There is greater variability in the unsuccessful campaigns. This does make sense as minimum number of backers is largely skewed by the low counts of backers, as at the end of the day you need numerous backers to reach a goal and so this lowers the first quartile thus increasing the variance. Comparatively, for successful campaigns the first quartile is higher as for a campaign to be successful you need numerous investors thus raising the bar.