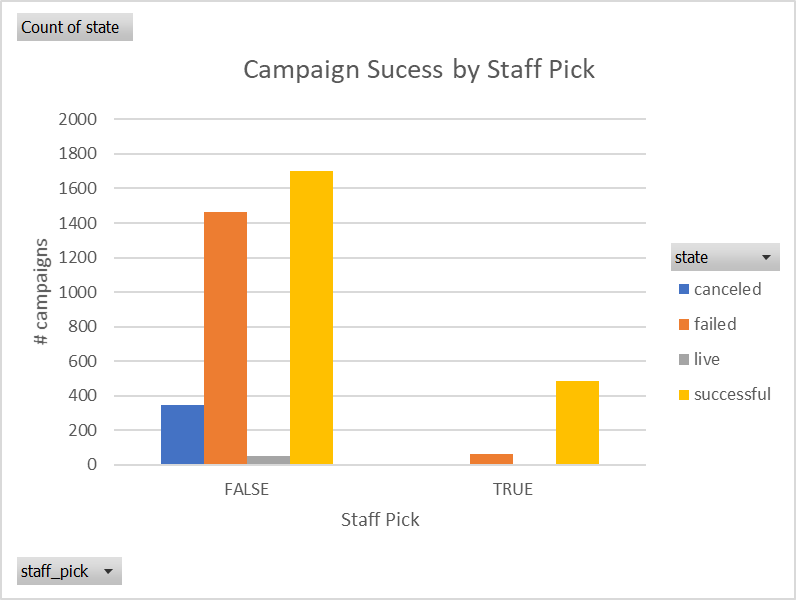
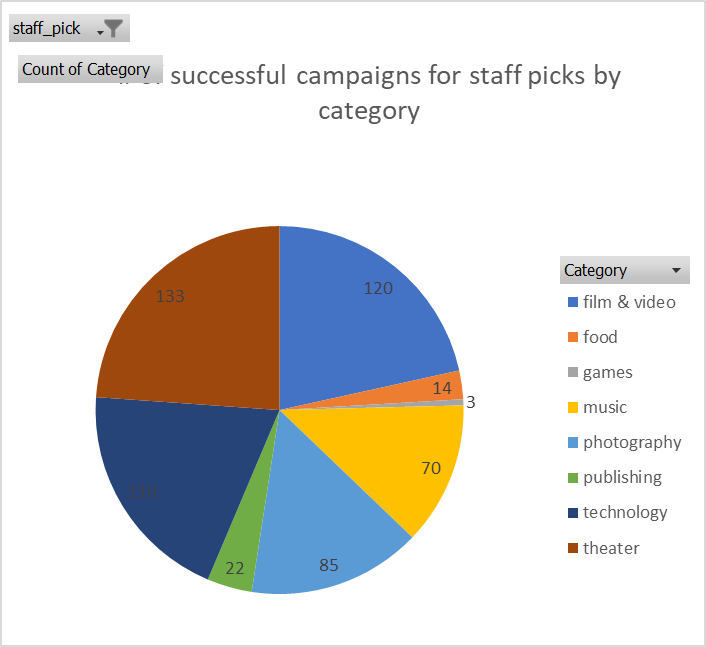
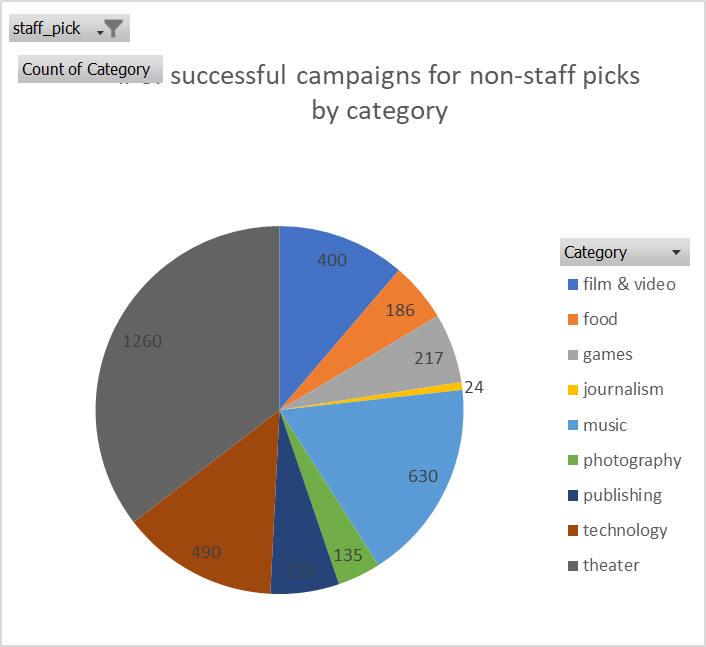
Excel Homework – Diane Tiblin

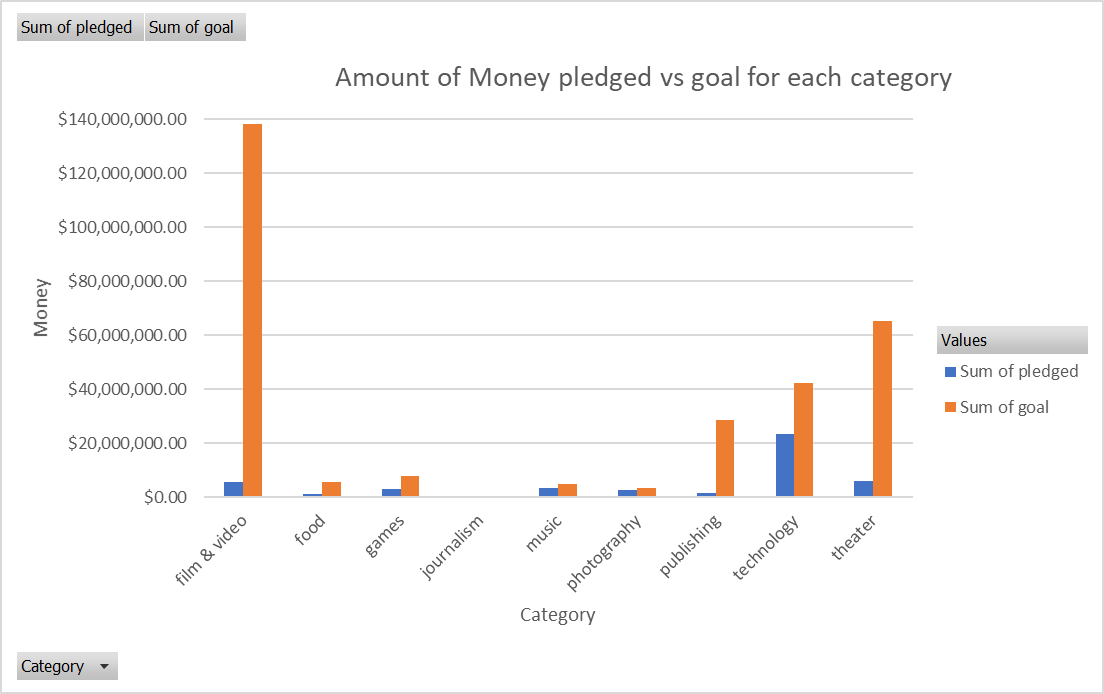
1. **Given the provided data, what are three conclusions we can draw about Kickstarter campaigns?**
   1. There are an abundance of projects in the arts field that are lacking in funding from traditional sources and have turned to Kickstarter, with theater being the largest category in terms of both # of campaigns & # of successful campaigns
   2. The month with the fewest # of successful campaigns is December, and the only month in which there were more failed campaigns than successful ones. We can conclude that December is probably not a good month to begin a kickstarter campaign.
   3. In general, as your original goal amount increases, the likelihood of completing the campaign successfully decreases
2. **What are some limitations of this dataset?**
   1. We cannot fully infer based on the data what exactly caused people to choose to help fund or not fund the projects which was the ultimate goal. Did the funded projects just have better marketing/outreach to potential donors or just a bigger network of potential donors? Were they just better projects? Is there some benefit provided for helping those projects get funded? There are so many factors that could have influenced successful campaign completion that are not reflected in this data set or able to be reliably determined as causal vs correlated.
   2. This sample data set appears like it may not be representative of the full population…the background indicates that only 1/3 of Kickstarter projects are funded successfully but the sample we have here indicates a success rate of over 50% so we may not be getting a fully accurate picture of what’s happening with the unsuccessful projects or an accurate comparison between successful vs unsuccessful.
   3. There is quite a bit of variability in the dataset in terms of the goals so it can be hard to make accurate deductions when analyzing it as a whole. It may make more sense to break the data into smaller sets for comparison such as in the bonus activity.
3. **What are some other possible tables and/or graphs that we could create?**
   1. One interesting chart/table would be the relationship between staff pick & campaign success or percentage funded. I would hypothesize that the attention garnered from being a staff pick (and presumably featured on the website) would make your project more likely to reach success and to receive more additional funding beyond the original goal. Without running any stats, you can see in the chart I created that the staff picks seem to have a higher success rate.

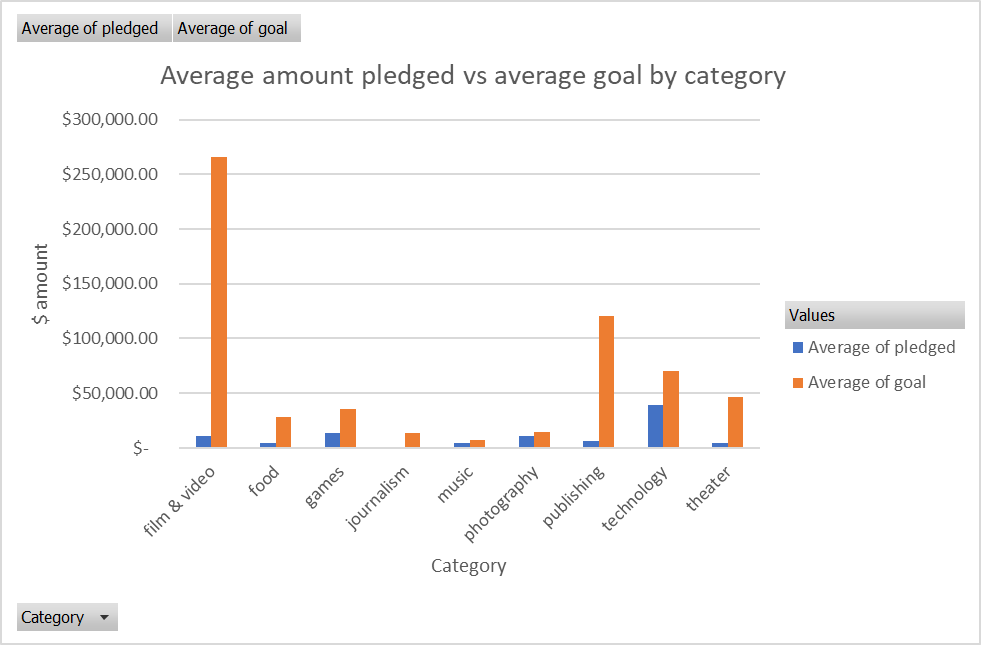


Then I was thinking about how the staff picks might influence a particular category or if they are reflective of the relative # of projects on the site. Maybe the staff is biased toward a particular category. So I created pie charts to illustrate how their picks shake out:

* 1. Another interesting chart is the amount of money pledged per category. First you would want to convert to a standardized currency (which I did not at this point in time), but at first glance it is still interesting to see that the category with by far the most # successful campaigns (theatre) does not appear to have anywhere near the most total money pledged (technology). Similarly interesting is a comparison of the amount of money pledged per category vs the goals. You might assume that the category with the most successful campaigns would have the closest ratio of amount pledged vs the goal but this is not the case. Most likely there are several outliers in the theater category throwing off the average I would guess.





Bonus Report:

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

The median is a better summary of the central tendency in this case because the data contains a lot of outliers skewing the mean to the higher end of the data set.

 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 with the successful campaigns as evidenced by the variance and std deviation. I think this makes sense since the successful campaigns would be more likely to have reached a larger # of backers.