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

1. The overwhelming majority of the sample of crowdfunded projects are theater plays.
2. The most reliably successful group of campaigns – according to the data, and with an emphasis on “reliable” – are the $1000-$4999 goal campaigns, regardless of category or time frame.
3. Campaigns tend to do better over the summer – possibly due to school vacations, with more free time for a family to attend the aforementioned theater plays.

* What are some limitations of this dataset?

1. The sample size is rather small to draw any sort of real conclusion for most of the analyzed data – The “most successful” group of each metric tends to be so because there are too few samples.
2. There are very few non-art-related projects, so this data is almost useless to anyone outside of an art field.
3. The data on each project is very limited – for example, it doesn’t take into account marketing/advertisement, and I suspect that the greatest indicator of success is the number of people you can communicate your campaign to.

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

1. Outcome vs Staff Picks and Success vs Spotlights, to determine whether or not the platforms, themselves, are good at picking out whether or not a project will succeed, and whether or not to go through the extra effort to get your project picked.
2. Outcome vs Average Donation, to determine the best supporter pricing tiers to set – since a funder can only pledge to a project once, you want to set tiers low enough that people will donate, but high enough that you get an amount of money that will allow you to succeed.
3. Outcome vs Campaign Duration – we analyzed start time, but not how long the campaign runs. Do longer campaigns have a higher chance of succeeding?
4. Outcome vs Country – Are those from specific countries more likely to succeed? And is this due to cultural phenomena, or just plain old marketing?

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

The medians better summarize the data for both the successful and failed campaigns: there are no outliers on the lower end for each case, but many outliers on the higher end for both cases, which heavily skews the mean upward. The median, in this case, allows us to ignore those 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 successful campaigns. This makes sense from a pure numbers perspective, simply because a failed campaign has a limited area in which it can fail – 0 <= x < goal – where a successful campaign has a theoretically infinite area in which it can succeed – goal <= x.