**Excel Homework: Charting Crowdfunding:**

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

* 1. The fund-raising goal bucket of $1000-$4999 included the highest number of successful campaigns. With more than two thirds of successful campaigns having a goal of $9999 or less.
* 2. Theater is the most popular category for seeking crowdfunding while journalism is the least popular.
* 3. Two categories – food and games – have the smallest success rate with more failed and canceled campaigns combined than successful ones.

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

* While other categories are broken down into further sub-categories, Theater has only one sub-category of Plays which includes all the data for that parent category. Other similar categories (I.e. music & film/video) have sub-categories that describe specific genres. This doesn’t really help to show if there are specific plays that raise money more successfully than others.
* The data covers more than 10 years of campaigns and with a sample size of 1000, that equates to approximately 100 campaigns per year which is not a very large sample size. More trends can be identified with more data – specifically more campaigns.

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

* One additional graph that would be helpful is comparing the fundraising goal with the category or sub-category. This would help answer the question, “Do certain categories require more funding than others to get ideas off the ground?”
* Another graph that would be helpful is looking at the length of the campaign compared to the success or failure of it. This would be helpful to answer the questions, “Does the length of a campaign impact the success or failure?”

**Bonus Statistical Analysis:**

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

* Based on the standard deviation and variance, there is a large spread in the number of backers for both the successful and failed campaigns. Since the spread is large, the median summarizes the data in a more meaningful way. The mean is skewed by higher values in the dataset

1. **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 successful campaigns, but there are also more successful campaigns than failed campaigns by about 20% which allows for a greater spread.