HOMEWORK 1 - EXCEL

Gustavo Maldonado

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

* There are 3 categories where the campaigns are more successful: Music (77%), Theater (60%) & Film&Video (58%).
* There are 3 months where the campaigns showed to be more successful: May (61%), April (60%) & February (57%).
* Looks like there are 23 events/subcategories where the effectiveness of the campaign is less than 50%, should help as statement to focus on those which their success is clearer than others.

1. What are some limitations of this dataset?

* The targets of each audience, would be interesting to have range of ages, ethnicity, location, political preferences, salary, education,… to identify if there is some variables that can bring more information about the profile of people/places where the campaign was successful.
* More information about the brackers as name of the association, company, amount per person, location.

1. What are some other possible tables and/or graphs that we could create?

* Pie graph to have a full picture by category according to its weights in % of total.
* State by location, to identify where the state of the campaigns.

BONUS - STATISCAL ANALYSIS

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

* Both are important, but in this case as the values min/max are hugely disperse, I would choose the median because the queuing in this data creates a distortion of the mean.

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 in the Successful campaign, this is easy to see in the variance & standard deviation. Meanwhile the parameters of the failed campaigns are less.
* For the Successful campaign the dispersion between the min/max is huge and the σ explains the complexity of being predictable. For my understanding, is better to have a small σ, because it means that the dispersion of data is closer at the mean value. In this case there is 23% chances to have ± lose or win according to the average for successful campaigns, besides failed where the percentage is 28%.