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

Answer: Taking a glance at the data visualization, Campaigns organized during the month of June and July were more successful and during the same period not much campaigns were canceled. At the same time, number of failed campaigns tends to increase but the highest number of total campaigns organized was in also July.

Secondly, number of successful campaigns were lowest in august which could’ve been influenced by the increasing number of canceled or failed campaigns as those values were one of the higher values in their corresponding datasets.

Furthermore, failed campaigns were almost the same during beginning and ending of the year, While the difference between those periods increase in failed and successful campaigns, with successful campaigns having the most variance between the start and end of the year. While, Live campaigns were almost within the same range and not showing much variability.

* What are some limitations of this dataset?

Answer: Dataset were missing a lot of values and those inconsistencies can be seen when we filter the data based of category and year, that would eventually affect the decision-making part of the analysis.

As well as, the data we have is just for 10 years and not for a bigger time range. Hence, we can apply descriptive analysis but cannot apply predictive analysis based on what happened in the past.

Moreover, we didn’t have much information regarding why it was happening and for that reason we cannot perform diagnostic analysis.

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

Answer: We could use clustered column representation, that will provide the comparison in between successful, failed, successful and live within the month. Whist applying filters, lines in line graph tends to intersect due to fluctuating values, hence, doesn’t provide proper visualization. In which case, clustered column will be more efficient. Also, radar can also be considered to visualize this kind of data as it excels at showing relative comparison, highlighting outliers, data simplification etc. (Jaspersoft, n.d.)

**Statistical Analysis Justification:**

As we check the median value of backers for successful campaigns which was 201 while the maximum value was 7265, it is evident that more values were lying between 16, i.e. is minimum value, than to the values between median and maximum value as median is much closer to minimum value rather than maximum value. Hence, most of the values were less than 201, which is supported the mean value, i.e. 851.1469 being closer to minimum value rather than maximum value.

And same thing is true in case of the failed campaigns as median, i.e. 114.5, and mean, i.e. 585.6154, were closer to minimum value than maximum value.

Conclusion being that a greater number of values lies from 16 to 201 and from 0 to 114.5 for successful and failed campaigns respectively.