Bonus Material for Homework 1

Matthew Berry

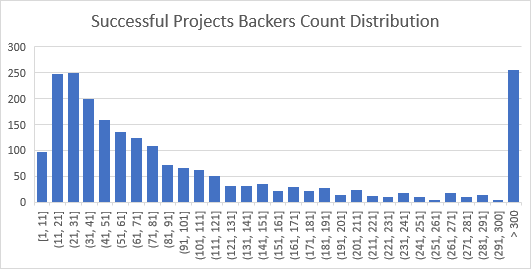
University of Denver University College

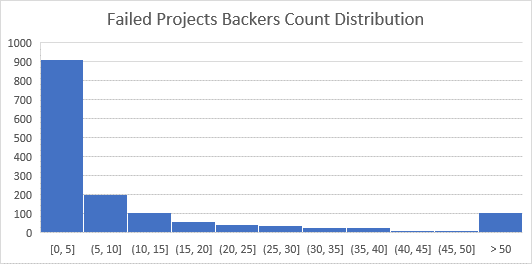
**Bonus Statistical Analysis**

1. **See Workbook for the requested Summary Statistics Table.**
2. **Use your data to determine whether the mean or the median summarizes the data more meaningfully.**

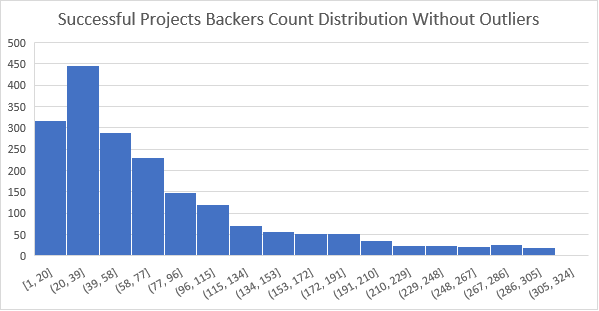
Using the mean to summarize the data is most appropriate when the data is normally (or symmetrically) distributed (or close to normally distributed as “normal distribution” is theoretical) and outliers are not skewing the mean.

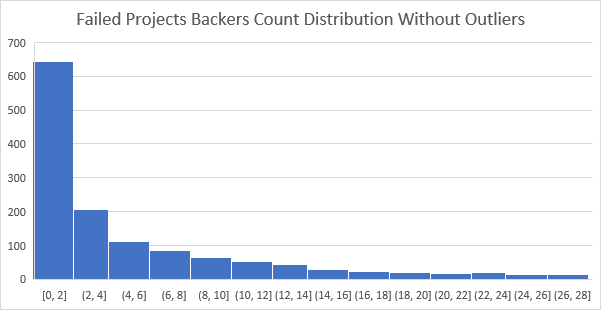
According to the histograms below neither of the datasets are symmetrically distributed.





Generally, if the distribution of data is skewed to the left, the mean is less than the median. If the distribution of data is skewed to the right, the median is less than the mean. In both datasets the median is less than the mean and the histograms confirm the datasets are both skewed right and are not symmetrically distributed. Even if we remove outliers from both datasets the distribution is still skewed right, as illustrated by the histograms below and the fact the median is still less than the mean in both datasets.





Therefore, because neither dataset is symmetrically distributed, and both contain outliers that skew the mean, using the median to summarize the data in both datasets is more “meaningful.”

1. **Use your data to determine if there is more variability with successful or unsuccessful campaigns. Does this make sense? Why or why not?**

The variance of the successful dataset is greater than the variance of the unsuccessful dataset. Also, the standard deviation for the successful greater than the standard deviation for the unsuccessful projects. The successful projects dataset has more variability than the unsuccessful dataset.

This does not make sense to me. It seems to me the variability should be similar.