Ticket Metrics and Probablity

Chris Mulvihill

5/2/2021

# Hypothesis 1:

## Outliers Impact The Average

I believe that the ticket metrics are being weighted in either direction by *extreme outliers*. This would cause the average (or mean) to be swayed away from a true representation of the majority of tickets.

For example: if you take a series of numbers like this one:

## [1] 4 5 5 6 3 4 100

The average of these numbers is

## [1] 18.14286

However, the majority of these numbers are much lower than the average. Only one number, 100, is higher than the average. This shows that a single value that is well above or well below the rest of the data will sway the average and misrepresent the data as a whole.

In this case, it would be better to use the median (the middle number) to find where the majority of the data is.

## [1] 5

The median shows us that the middle value is 5. Since this is relatively far from the mean, we can assume that there are extreme outliers that are impacting the data.

In this summary, we see that the majority of the values are between 4 and 5.5. since the mean is beyond the 3rd quartile, we have further evidence of our suspicions that outliers are impacting our analysis.

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 3.00 4.00 5.00 18.14 5.50 100.00