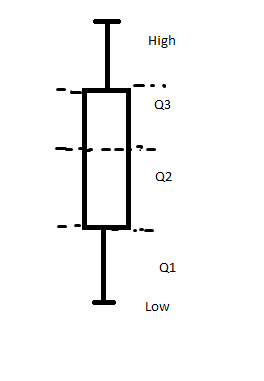
1. Difference between outlier and Anomaly

Outlier is an extreme data point.

1. What is relation between IQR and Boxplot

In exploratory data analysis, a boxplot is often used for descriptive statistics. For finding outliers one of the best visualisation methods used is called Box Plot sometimes also called Whisker Plots. In this method, we can clearly see the outliers that lie beyond the low and high limits of the whisker plot. Univariate analysis is the most efficient way of determining outliers in a dataset. IQR plots can visualize the first quartile, second quartile and third quartile of the data distribution making it easy for outlier visualization.



Using these quartile points is used Q3 is the 75th Percentile of the data. Q2 is the 50th percentile of the data. Also, Q1 is the 25th percentile of the data. Similarly using this IQR is found by taking the difference between Q3 and Q1. In IQR it is always assumed that data is distributed normally.

IQR = Q3-Q1

Low = Q1 – 1.5 \* IQR

High = Q3 – 1.5 \* IQR

So any point that comes below low or above high as per the above statistical equation will be considered as an outlier or extreme point.

1. Code attached as ipynb file.
2. Can I noisy Dataset mislead the model generating Algorithm to generate an Inaccurate model? Explain How?