1. Write the Gaussian Distribution empirical formula?

Ans: , where

1. What is Z-score, and why it is important?

Ans: Z score is used to standardize the data on a common reference. Z score allows us to calculate the probability of the score occurring within the normal distribution and enables us to compare the multiple scores that are from different normal distribution.

1. What is an outlier, exactly?

Ans: An outlier is an observation that lies at an abnormal distance from other values in a random sample from a population. It is up to the analyst to decide what will be considered as an outlier. Normally, different tests like Z-test, t-test, IQR etc are used to find the outliers

1. What are our options for dealing with outliers in our dataset?

Ans: Following are the ways to deal with outlier:

* Remove the outliers (trim the data)
* Replace the values of outliers or reducing the weight of outliers through weight adjustment

1. Write the sample and population variances equations and explain Bessel correction?

Ans: Population variance:

Sample variance:

In statistics, Bessel’s correction is the use of (n-1) instead of n in the formula of simple variance of a sample from the population. Based on various iterations, scientist found that (n-1) gives better estimation towards the population variance than n.