

## WORKSHEET 4 STATISTICS

## ASSIGNMENT-6

**ANS. NO.1 (D)**

**ANS.NO.2 (A)**

**ANS.NO.3 (A)**

**ANS.NO.4 (C)**

**ANS.NO.5 (B)**

**ANS.NO.6 (A)**

**ANS.NO.7 (C)**

**ANS.NO.8 (B)**

**ANS.NO.9 (B)**

**ANS.NO.10**

Histograms are a special kind of bar graph that shows a bar for a range of data values instead of a single value.

A box plot is a data display that draws a box over a number line to show the interquartile range of the data. The 'whiskers' of a box plot show the least and greatest values in the data set.

**ANS.NO.11**

The key point is to choose metrics that clearly indicate where you are now in relation to your goals. Good metrics can be improved. Good metrics measure progress, which means there, needs to be room for improvement.

**ANS.NO.12**

To assess statistical significance, use hypothesis testing. The null hypothesis and alternate hypothesis would be stated first. Second, calculate the p-value, which is the likelihood of getting the test's observed findings if the null hypothesis is true. Finally, select the threshold of significance (alpha) and reject the null hypothesis if the p-value is smaller than the alpha.

— In other words, the result is statistically significant.

**ANS.NO.13**

Exponential distributions do not have a log-normal distribution or a Gaussian distribution. In fact, any type of data that is categorical will not have these distributions as well.

Example: Duration of a phone car, time until the next earthquake, etc.

**ANS.NO.14****Income is the best example:**

Income is the classic example of when to use the median instead of the mean because its distribution tends to be skewed. The median indicates that half of all incomes fall below and half are above it. For these data, the mean overestimates where most household incomes fall.

**ANS.NO.15**

The likelihood is the probability that a particular outcome is observed when the true value of the parameter is , equivalent to the probability mass on ; it is not a probability density over the parameter . The likelihood, should not be confused with, which is the posterior probability of given the data