

STATISTICS WORKSHEET-1

1) a

2) a

3) b

4) d

5) c

6) b

7) d

8) a

9) c

10) The normal distribution is the probability distribution in which most of the data points cluster towards the middle of the curve, this point is called the mean. The mean, median and mode are all the same for this distribution.

11) In small datasets we can handle missing data by dropping rows/columns but in large datasets we can't drop the row/columns as it leads to loss of information and hence wrong estimations. We can substitute the missing value with the mean /median value of a particular column if that column is numerical and with mode (most frequently used value) if the column is categorical.

12) A/B testing is a statistical split testing where two variants of a webpage or app are compared against each other to see which performs better. Both of them are shown to users and data is gathered from them.

13) Yes It is acceptable as the sample size of the dataset remains the same and the estimate of the mean remains unbiased.

14) It is used to predict the value of a variable based on the value of other variable/variables. The variable whose value we predict is called dependent variable and the variable by which we predict the dependent variable is called independent variable.

15) two branches of statistics are descriptive and inferential.