Statistics Assignment

1) A)

2) D)

3) C)

4) D)

5) B)

6) B)

7) A)

8) A)

9) C)

10) A normal distribution is an arrangement of a data set in which most values cluster in the middle of the range and the rest taper off symmetrically toward either extreme. A graphical representation of a normal distribution is sometimes called a bell curve because of its flared shape. The precise shape can vary according to the distribution of the population but the peak is always in the middle and the curve is always symmetrical. In a normal distribution, the mean, mode and median are all the same.

11) There are three main approaches to handle missing data: (1) *Imputation*—where values are filled in the place of missing data, (2) *omission*—where samples with invalid data are discarded from further analysis and (3) *analysis*—by directly applying methods unaffected by the missing values.

12) A/B Testing is the testing in which we compare data of two variation of same element between A and B and concluding which of the variant is more concluding.

13) True, imputing the mean preserves the mean of the observed data. So if the data are missing completely at random, the estimate of the mean remains unbiased. Since most research studies are interested in the relationship among variables, mean imputation is not a good solution

14) Linear regression is an attempt to model the relationship between two variables by fitting a linear equation to observed data, where one variable is considered to be an explanatory variable and the other as a dependent variable.

15) There are three real branches of statistics: data collection, descriptive statistics and inferential statistics.

Descriptive statistics deals with the presentation and collection of data. This is usually the first part of a statistical analysis. It is usually not as simple as it sounds, and the statistician needs to be aware of designing experiments, choosing the right focus group and avoid biases that are so easy to creep into the experiment.

Inferential statistics, as the name suggests, involves drawing the right conclusions from the statistical analysis that has been performed using descriptive statistics. In the end, it is the inferences that make studies important and this aspect is dealt with in inferential statistics.