Statistics Worksheet 1

Mcq

1. a)

2.c)

3.b)

4.d)

5.c)

6.b)

7.b)

8.a)

9.c)

Subjective type questions:

10. It is the probability distribution which is around the mean, if we draw the shape on the graph, it will be like bell shape curve. It has zero skewness.

11.) In case of numerical data, we take the mean method. In case of categorical data, we take median as well average of the column and can find and fill the missing data.

12) An AB test is an example of statistical hypothesis testing, a process whereby a hypothesis is made about the relationship between two data sets and those data sets are then compared against each other to determine if there is a statistically significant relationship or not.

13) Yes .As in general, in continuous data , we see that there are lot of cases occur where there is missing data and we can use the mean of the numbers as it gives the overall average number. It would not affect the overall data much.

14) It is generally used for predicting the future results or outcome. It's like one company is investing some amount of money in TV, newspaper advertisements and how much money will the company will earn).it shows if we invest x amount of money in adds then how much company will earn in future. We should have the continuous data and we can train the model and it predicts the future.it basically involves inputs (independent) and produces the label (dependent variable).it is used in predicting the sales, weather forecasting etc.

15) There are two branches:

1. Inferential

2. Descriptive

1. Inferential: Here we have a large amount of data. We take the small samples that represent the data and conclude the results from it .Suppose we have the data on 1000 bulbs. And we want to know the

proportion of defective bulbs .We take a small sample from it and test it .If the sample does not contain any defective bulb then we will conclude that there are no defective bulbs in the entire lot. Here we don't test the entire lot but take small samples and analyze it to conclude results.

2. Descriptive statistics: It is the basically the data which anyone can analyses easily and can explain it conveniently. Suppose we consider an example of marks of students in a particular school. Then we can know how many students are exactly there who have scored more than 80 marks, less than 50 marks or how many students have scored 0, etc. We can know the exact number of such students. Hence here we are able to conclude reliable results (in proper numbers).