Python Assignment:

1. C
2. A
3. C
4. A
5. D
6. C
7. A
8. A
9. A,B,C
10. A,B

Statistics Assignment:

1. A
2. A
3. C
4. D
5. C
6. B
7. B
8. A
9. C
10. Normal distribution is a probability distribution which states 68% of the data is within 1 standard deviation of the mean, 95% is within 2 standard deviations from mean , and 99.7% of data is within 3 standard deviations from the mean.

11. There are 2 ways in which missing data can be handled. 1. Data deletion - In this method we delete the data lines where the data is missing. This method is suitable for cases where the % of missing lines is very less compared to overall data. Second method is data imputation, here the missing data is replaced with another value which can be mean of remaining data or any other value depending on the nature of data.

12. A/B testing is where 2 versions of the same variables are studies in a control enviroment to find out which version performs better.

13. Yes data imputation is a acceptable practice, it is used to bring harmony in the data also in cases where the missing data is high in % compared to overall data imputation is required.

14. Linear Regression is a common type of predictive analysis. There are 2 major objectives of Linear Regression 1. To predict the outcome 2. To find the variables which contribute the maximum in output. It is used to explain the relationship between dependent and independent variables on the outcome.

15. There are 4 major branches of statistics:

> Mathematical or theoretical Statistics

> Statistical Methods or functions

> Descriptive Statistics

> Inferential Statistics

Machine Learning Assignment:

1. A
2. A
3. C
4. B
5. C
6. B
7. D
8. D
9. A
10. A
11. B
12. A,B,C
13. Regularisation is a method used to solve under fitting and over fitting problem of regression model. In case of under fitting where the model is not able to find the trend line of the data. Regularisation helps to find the closest possible trend line to predict the more accurate test data or unseen data, the major problem in underfitting data is high bias. Incase of Over fitting where the model tries to make a trend line with maximum inputs the problem of high variance occurs, here the model works good on train data but performs drastically worse on test data.
14. Ridge and Lasso Regression
15. Error in linear regression is the difference between expected output and actual output in the machine learning model. For example the predicted price of a stock using linear regression model was 100 but the actual price of the stock is 80.