PYTHON – WORKSHEET 1

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

STATISTICS WORKSHEET-1

1. A
2. A
3. B
4. D
5. C
6. B
7. B
8. A
9. C
10. Normal distribution, also known as the Gaussian distribution, is a probability distribution that is symmetric about the mean, showing that data near the mean are more frequent in occurrence than data far from the mean. In graph form, normal distribution will appear as a bell curve.G
11. We can handle missing data by deletion methods, regression analysis to systematically eliminate data, imputation techniques.
    Imputation Techniques to handle missing data –

    1. Mean, Median, Mode
    2. Last observation carried forward
    3. Next observation carried forward
    4. Linear interpolation
    5. Common point imputation
    6. Random sampling imputation

12. A/B testing is one of the most popular controlled experiments used to optimize web marketing strategies. It allows decision makers to choose the best design for a website by looking at the analytics results obtained with two possible alternatives A and B.To understand what A/B testing is about, let's consider two alternative designs: A and B. Visitors of a website are randomly served with one of the two. Then, data about their activity is collected by web analytics. Given this data, one can apply statistical tests to determine whether one of the two designs has better efficacy. Now, different kinds of metrics can be used to measure a website efficacy. With discrete metrics, also called binomial metrics, only the two values 0 and 1 are possible.

13. No.

14. Linear regression analysis is used to predict the value of a variable based on the value of another variable. The variable you want to predict is called the dependent variable. The variable you are using to predict the other variable's value is called the independent variable.

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

MACHINE LEARNING WORKSHEET-1

1. A
2. A
3. B
4. B
5. C
6. B
7. D
8. B
9. A
10. B
11. B
12. A, B, C
13. Regularize means to make things regular or acceptable. This is exactly why we use it for. Regularizations are techniques used to reduce the error by fitting a function appropriately on the given training set and avoid overfitting.
14. Lasso, Ridge and Elastic -net regression.
15. An error term represents the margin of error within a statistical model; it refers to the sum of the deviations within the regression line, which provides an explanation for the difference between the theoretical value of the model and the actual observed results. The regression line is used as a point of analysis when attempting to determine the correlation between one independent variable and one dependent variable.