

Statistics Solution

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

10. A normal distribution is a distribution that is symmetric around the, this means there is more occurrence of data around the mean and the occurrence gradually goes down as you move away from the mean. In graphical presentation, it always appears a 'bell shaped' curve.

11. Handling missing data can be done primarily in two ways, imputation and removal of data. Before removing or imputing data, Data Scientists have to analyze carefully the reason why the data is missing. Mode, Median and Mean imputations are the most commonly used, however, in the case where we have a lot of missing data, Time Series Imputation will be more beneficial.

12. A/B testing is a way to compare the two versions of a variable to find out which performs better in a controlled environment.

13. Mean imputation of data is an acceptable practice in the case where we only have a few missing data points.

14. Linear regression is a linear approach to explain the relationship between a scalar response and one or more variables.

15. There exist two main branches of statistics which are Descriptive and Inferential Statistics.