WORKSHEET

STATISTICS WORKSHEET-1

1.(a)

2.(a)

3.(b)

4.(d)

5.(c)

6.(b)

7.(b)

8.(a)

9.(d)

10. A normal distribution is a type of continuous probability distribution in which most data points cluster toward the middle of the range, while the rest taper off symmetrically toward either extreme.

11.1. Missing data can be dealt with in a variety of ways. I believe the most common reaction is to ignore it.

2. Another common strategy among those who pay attention is imputation.

3. Mean imputation.

4. Substitution.

5. Hot deck imputation.

6. Cold deck imputation.

7. Regression imputation.

8. Stochastic regression imputation.

12. A/B testing is a methodology for comparing two versions of a webpage or app against each other to determine which one performs better. A/B testing is essentially an experiment where two or more variants of a page are shown to users at random, and statistical analysis is used to determine which variation performs better for a given conversion goal.

13. Mean imputation is typically considered terrible practice since it ignores feature correlation.

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.