WORKSHEET 1 STATISTICS

Q1 to Q9

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

Q10 to Q15

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. The middle of the range is also known as the mean of the distribution.

11. Missing data can be dealt with in a variety of ways. I believe the most common reaction is to ignore it. Choosing to make no decision, on the other hand, indicates that your statistical programme will make the decision for you. Your application will remove things in a listwise sequence most of the time. Depending on why and how much data is gone, listwise deletion may or may not be a good idea.

    Mean imputation

    Calculate the mean of the observed values for that variable for all non-missing people. It has the advantage of maintaining the same mean and sample size, but it also has a slew of drawbacks. Almost all of the methods described below are superior to mean imputation.

12. A/B testing is a type of experiment in which you split your web traffic or user base into two groups, and show two different versions

of a web page, app, email, and so on, with the goal of comparing the results to find the more successful version.

13. The process of replacing null values in a data collection with the data’s mean is known as mean imputation. mean imputation decreases the variance of our data while increasing bias. As a result of the reduced variance, the model is less accurate and the confidence interval is narrower.

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