STATISTICS WORKSHEET-1

ANS-1-A

ANS-2-A

ANS-3-B

ANS-4-D

ANS-5-C

ANS-6-B

ANS-7-B

ANS-8-A

ANS-9-C

ANS-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 graphical form, the normal distribution appears as a “bell curve”.

ANS-11-Mean, median and mode are the most common methods of imputing values when dealing with missing data. In cases where there are small number of missing observations, we can calculate the mean or median of existing observations. I recommend the best imputation method is mice-cart.

ANS-12-A/B testing, also called split testing, is a comparison test between two or more versions of the same page. It will divide the

traffic into all versions to determine which page design delivers the better results using metrics like the number of page views, session time or conversion rate.

ANS-13-Mean imputation does not preserve the relationships among variables. True, imputing the mean preserves the mean of the observed data. So if the data are missing completely at random, the estimate of the mean remains unbiased.

ANS-14-In statistics, linear regression is a linear approach for modeling the relationship between a scalar response and one or more explanatory variables. The case of one explanatory variable is called simple linear regression; for more than one, the process is called multiple linear regression.

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