STATISTICS WORKSHEET-3

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

10. Bayes' theorem describes the probability of occurrence of an event related to any condition. It is also considered for the case of conditional probability. Bayes theorem is also known as the formula for the probability of "causes". Let E1, E2…, En be a set of events associated with a sample space S, where all the events E1, E2…, En have a nonzero probability of occurrence and they form a partition of S. Let A be any event associated with S, then according to Bayes theorem, $P(Ei|A) = P(Ei)P(A|Ei)\Sigma k=1nP(Ek)P(A|Ek)$ for any k = 1, 2, 3, …., n

11. A z-score (also called a standard score) gives you an idea of how far from the mean a data point is. But more technically it's a measure of how many standard deviations below or above the population mean a raw score is. A z-score can be placed on a normal distribution curve. Z-scores range from -3 standard deviations (which would fall to the far left of the normal distribution curve) up to +3 standard deviations (which would fall to the far right of the normal distribution curve). In order to use a z-score, we need to know the mean μ and also the population standard deviation σ.

    The basic z score formula for a sample is: $z = (x – \mu) / \sigma$

12. A t-test is a statistical test that is used to compare the means of two groups. It is often used in hypothesis testing to determine whether a process or treatment actually has an effect on the population of interest, or whether two groups are different from one another. o The t-test is a test used for hypothesis testing in statistics. o Calculating a t-test requires

three fundamental data values including the o difference between the mean values from each data set, the standard deviation o of each group, and the number of data values. o T-tests can be dependent or independent.

13. A percentile (or a centile) is a measure used in statistics indicating the value below which a given percentage of observations in a group of observations fall. For example, the 20th percentile is the value (or score) below which 20% of the observations may be found. The term percentile and the related term percentile rank are often used in the reporting of scores from norm-referenced tests. Every score is in the 100th percentile. The 25th percentile is also known as the first quartile (Q1), the 50th percentile as the median or second quartile (Q2), and the 75th percentile as the third quartile (Q3). In general, percentiles and quartiles are specific types of quantiles.

14. ANOVA, which stands for Analysis of Variance, is a statistical test used to analyze the difference between the means of more than two groups. A one-way ANOVA uses one independent variable, while a two-way ANOVA uses two independent variables. An ANOVA test is a way to find out if survey or experiment results are significant. In other words, they help you to figure out if you need to reject the null hypothesis or accept the alternate hypothesis.

15. ANOVA is helpful for testing three or more variables. It is similar to multiple twosample t-tests. However, it results in fewer type I errors and is appropriate for a range of issues. ANOVA groups difference by comparing the means of each group and includes spreading out the variance into diverse sources. The one-way ANOVA can help you know whether or not there are significant differences between the means of your independent variables (such as the first example: age, sex, income). When you understand how each independent variable's mean is different from the others, you can begin to understand which of them has a connection to your dependent variable (landing page clicks), and begin to learn what is driving that behavior