Insert Title Here

Insert Your Name Here

Insert University Here

**Data Analysis: Descriptive Statistics and Assumption Testing**

**Correlation: Descriptive Statistics and Assumption Testing**

**Frequency Distribution Table:**



**Histogram:**



**Descriptive Statistics Table:**



**Measurement Scale:**

Ratio scale is used as a measurement scale here, as the:

**Measure of Central Tendency:**

The mean of the average number of days an employee remains sick in the company is measured at 7.1262. Standard deviation or the average variation per person is 0.1864. The median and mode, both remain at 7.

**Evaluation:** On an average, an employee takes approximately seven leaves due to illness as per the given data. In the histogram, we can visualize that the frequency on the 6 to 8 bracket is relatively high. In terms of normality, the curve is slightly skewed towards the left, the kurtosis is close to zero i.e. distribution is normally distributed. In the given dataset, one can conclude that the assumptions for parametric statistical testing were met.

**Simple Regression: Descriptive Statistics and Assumption Testing**

**Frequency Distribution Table:**



**Histogram.**



**Descriptive Statistics Table:**



**Measurement scale:** Since there is a true zero present here, we can conclude that ratio scales are used as measurement scales in the representation of this data.

**Measure of Central Tendency:**

Central tendency of this data is measured by its mean, standard deviation and skewness. Skewness and standard deviation indicate the direction in which most of the data lies.

**Evaluation.** The data gives us a normal distribution curve thus, it is normally distributed. The assumption regarding the independent variables being linearly correlated does not apply as there is only one independent variable. The variable in the first column is contract number.

**Multiple Regression: Descriptive Statistics and Assumption Testing**

**Frequency Distribution Table:**



**Histogram.**



**Descriptive Statistics Table:**



**Measurement Scale:** The measurement scale used in this data set is known as the ratio scale, as decibels can attain an actual value of zero.

**Measure of Central Tendency:** The mean value of the dependent variable i.e. decibels, is equal to 124.8359 and the standard deviation suggests that most of the values are concentrated near the mean value. The value of kurtosis and the shape of the histogram suggests that the frequency distribution is normal as well.

**Evaluation:** The initial assumption of normality is met in the data set as seen in the histogram curve. The assumption of multi-collinearity and multi-variate normality is also being fulfilled. Thus, we can conclude that the assumptions for parametric statistical testing were met.

**Independent Samples *t* Test: Descriptive Statistics and Assumption Testing**

**Frequency Distribution Table:**



Figure 1-Group A freq. distribution



Figure 2-Group B freq. distribution.

**Histogram**





**Descriptive Statistics Table:**



Figure 3-Statistical summary of group A data



Figure 4-statistical summary of group-B data

**Measurement Scale:** The measurement scale used in the representation of this dataset is the ratio scale.

**Measure of Central Tendency:** The central tendency for such a dataset is generally used as median because it is more robust in tackling the problem of outliers.

**Evaluation:** The assumptions of normality of both the variables are met as seen in the histogram and the value of kurtosis for both variables but the assumption of homogeneity of variances is not being met so we can conclude that the assumptions for parametric statistical testing were not met in the given data.

**Dependent Samples (Paired-Samples) *t* Test: Descriptive Statistics and Assumption Testing**

**Frequency Distribution Table:**



Figure 5-freq distribution table for per-exposure data



Figure 6-freq. distribution table for post-exposure data

**Histogram:**



Figure 7-Histogram for pre-exposure data



Figure 8-Histogram for post-exposure data

**Descriptive Statistics Table:**



Figure 9-Pre exposure statistics



Figure 10-Post exposure statistics

**Measurement Scale:** The ratio scale is used as a measurement scale in this dataset.

**Measure of Central Tendency:** The most appropriate measure for central tendency in a sampled t-test is median as it is the one which is least effected by outliers.

**Evaluation:** The assumption of normality for both the variables are met as seen in the histogram. In the given data, the assumption that both the variables need to be independent of each other is not being met. Therefore, we can conclude that the assumptions for parametric statistical testing is not being met.

**ANOVA: Descriptive Statistics and Assumption Testing**

**Frequency Distribution Table:**



Figure 11-Frequency dist. for D



Figure 12-freq. dist. table for A



Figure 13-freq. dist. table for C



Figure 14- freq. dist. table for B

**Histogram.**



Figure 15-Histogram for D



Figure 16- Freq. dist. table for C



Figure 17-freq. dist. table for B



Figure 18-freq. dist. table for A

**Descriptive statistics table.**



Figure 19-Statistics for D



Figure 20-Statistics for C



Figure 21-Statistics for B



Figure 22-Statistics for C

**Measurement scale.** Ratio scale is being used for measurement here.

**Measure of central tendency.** Mean is the most appropriate measurement measure of central tendency here.

**Evaluation.** The histogram for the dependent variable is normally distributed. The sample variance for all the given variables is not the same so we can conclude that the assumptions for parametric statistical testing were not met.