Statistics

[Enter your Name here]

[Enter name of Institution here]

Researchers perform their work with an awareness that the data at their disposal will break some assumptions underlying the statistical tests of significance. Analyzing such data may create problems for the researchers. Data is transformed in such a manner that assumptions are not broken or there may be some other tests used by the person. The non-parametric tests’ options have opened more ways for researchers to adopt (Ghasemi & Zahidiasi, 2012). The breaking of assumptions is coupled with changing characteristics of the data e.g. a researcher who classifies certain scores in groups above or below the median scores, will ignore the differences between different categories. The tests with no assumption regarding distribution of sample data will tend to accept the null hypothesis when it is actually false which means that these tests will lack power. This lack of power is the major handicap that stops researchers from using these non-parametric tests and an occasional significant result is necessary to maintain the response of research. A researcher does not take up the tedious work of transforming the data and simply goes ahead with using t-tests or analysis of variance. A significant result is either due to difference in means or difference in variance (Boneau, 1960).

 All forms of statistical analysis assume sound measurement which does not include any coding errors. Running descriptive statistics on data is a good practice because researchers then expect output in form of mean and standard deviation. Descriptive statistics also confirm that there are no outliers in the data that can distract the data. It is not a good practice to change the limits of data in an artificial way because in that case, correlation is disturbed leading to wrong estimation of effect size for that particular variable (Nimon, 2012). If there are some common factors affect both variable A and variable B, the correlation between these two variables is disturbed. This is especially true when same indicators affect the independent and dependent variables at the same time. Proper model specification is another way to conduct proper statistical research. Specification refers to including all relevant variables in the analysis and not omitting any variable that has some significant relationship with the dependent variable. It is also important that the direction of all relationships is correctly shown in the diagram. A good literature review is much helpful in identifying the variables properly. A lower value of coefficient of determination in case of regression model will show that important variables have been omitted. In studies with a single model, specification problem will be more serious. As far as chi-square test is concerned, there should be adequate cell count. A factor space is the set of all cells that are generated by crosstabulation of categorical dependent with all categorical factors. A large sample size may not guarantee adequate cell size. Most statistical procedures require an interval or ratio level of measurement. In social sciences, researchers use dichotomous or ordinal data. Ordinal data is only used in statistical research if there is normality in the error term. Violation of data level assumptions will mean that actual standard error will be much higher than computed value of standard error and significance is over estimated. Centering is not an assumption of any given statistical technique but it is strongly recommended to be attained. Without centering, there will be no real world meaning to the coefficients calculated (Garson, 2012).

**Analysis**

H0: There is no significant difference between the two samples

H1: There is a significant difference between two samples

Level of significance has been set at 0.05

Z = 

Putting values

(4.5-5.6)/$\sqrt{}$2.25/100+4/100

=-1.1/0.25

=-4.4

The tabulated value for z @ 0.05 level of significance is 1.96 and z calculated is greater than z tabulated so we reject H0 and state that there is a significant difference between the two groups.

# **References**

Boneau, C. (1960). THE EFFECTS OF VIOLATIONS OF ASSUMPTIONS UNDERLYING THE t TEST. *Psychological Bulletin*.

Garson, D. (2012). *Testing Statistical Assumptions.* Statistical Associates Publishing.

Ghasemi, A., & Zahidiasi, S. (2012). Normality Tests for Statistical Analysis: A Guide for Non-Statisticians. *International Journal of Endocrinology Metabolism*, 486-489.

Nimon, K. F. (2012). Statistical assumptions of substantive analyses across the general linear model: a mini-review. *Frontiers of Psychology*.