Unit 5 MBA5652 Assignment

[Author Name(s), First M. Last, Omit Titles and Degrees]

[Institutional Affiliation(s)]

Author Note

Unit 5 MBA5652 Assignment

This section provides the correlation, linear regression, and multiple regression test hypotheses, test results, findings, and analysis. The provided sun coast dataset has been employed to run the test in Excel Toolpak. The outputs from excel have been imported and analyzed below.

# Correlation: Hypothesis Testing

The correlation hypothesis testing involves testing the relationship between two variables A and B (Field, 2000). In this test, the relationship of health of employees and particulate matter (PM) will be tested using chi-square correlation test, and simple regression will show the percentage of relationship between the variables. The test results will predict if the higher level of air pollution measured in microns of particulate matter's size is associated with a higher number of leaves from the employees in different sites.

## Hypotheses

Null hypothesis Ho1: No statistically substantial relationship exists among particulate matter and number of sick leaves.

Experimental hypothesis Ha1:A statistically substantial relationship exists among particulate matter and number of sick leaves.

## Correlation output and interpretation



Fig. 1 Correlation output

It is evident from the table that the R-square is equal to 0.513, Pearson’s coefficient r is ‘-0.715’. Hence, the two variables are 71.5% related to each other. The negative sign indicates an inverse relationship between the two. Lower the particle size, higher the number of sick leaves. It is important to note that the results of correlation analysis only show the relationship, which means it does not predict cause and effect. For example, we cannot say that particle size is causing ill health. However, there is surely a strong relationship observed among the two variables.

 Therefore, we reject the null hypothesis and accept the experimental or alternative hypothesis, which states that a relationship exists between particulate matter size and a number of sick leaves.

# Simple Regression: Hypothesis Testing

The regression analysis in addition to proving the relationship between the variables, determines the percentage of relationship between them (Field, 2000). The simple regression in this case will predict to what extent the safety training is correlated to lost time hours. The training has been measured by amount of money spent on safety training.

## Hypotheses

Null hypothesis Ho1: Training has proven to be effective in reducing the lost-time hours.

Experimental hypothesis Ha1:Training has proven to be effective in reducing the lost-time hours.

## Simple/linear regression output and interpretation

Linear regression outputs from the data provided have been presented below.



Fig.2 (a) Scatterplot showing correlation training expenses and lost time hours

The scatterplot above shows inverse relationship between the two variables. The rise in money spent on training is accompanied by the reduction of a number of lost hours. The plot shows the Pearson’s coefficient (r) 0.1434. Hence, the relationship between the two variables was 14.34%. It is important to note that the outlier exists, which was removed for better interpretation of results.



Fig.2 (b) Scatterplot showing correlation training expenses and lost time hours after removal of outliers

The scatterplot after removal of outlier is presented above. The relationship after removing outliers is 14.37% which is not statistically different from that obtained before removal of the outlier.



Fig.3 (a) Simple regression test result

It is evident from the table above that r square value is 0.883. Hence, the results show that the variables are 88% related to each other. The multiple R is 0.939, R square 0.883, ANOVA F value 1851.86 and alpha a value is 273.45.



Fig.3 (b) Simple regression test result

Y = bx + a

Where b is safety training expenditure and a is intercept. Hence,

Y = -0.143 \* x + 273.45

The regression equation shows that lost time hours can be obtained by multiplying x to -0.143 and adding 273.45. Considering the above results, we do reject the null hypothesis and accept the experimental hypothesis that safety training is effective in reducing the lost time hours.

# Multiple Regression: Hypothesis Testing

The multiple regression is used to study a number of factors linked to a variable Creswell & Creswell, 2017). In this case, it is tested if the variables like angle, velocity, chord length, frequency and displacement are contributing to noise measured in decibels.

## Hypotheses

Null hypothesis Ho1: The variables frequency, angle, chord length, velocity and displacement do not contribute to noise (dB).

Experimental hypothesis Ha1:The variables frequency, angle, chord length, velocity and displacement contribute to noise (dB).

## Multiple regression output and interpretation

The excel outputs of multiple regression analyses have been presented below.

****

Fig.4 (a) Multiple regression output

 The Multiple R value observed is 0.00, R square is 0.36 and ANOVA F value 170.03.

****

Fig.4 (b) Multiple regression output

As evident from the table above, the regression outputs display non-significant outputs for angle and chord length and significant values for frequency, velocity, and displacement. Therefore, the velocity, frequency, and displacement contribute to noise at workplace. The amount of noise can be predicted by following equation if the frequency, velocity and displacement are known.

Dependent variable (Noise) = 126.82 – 0.00 (frequency) + 0.08(velocity) – 240(displacement), where 126.82 is constant alpha value.

# References

Creswell, J. W., & Creswell, J. D. (2017). *Research design: Qualitative, quantitative, and mixed methods approaches*. Sage publications.

Field, A. (2000). *Discovering statistics using SPSS:(and sex, drugs and rock'n'roll)* (Vol. 497). Sage.