Correlation and Regression Analysis Using Sun Coast Data Set

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

[Institutional Affiliation(s)]

Author Note

[Include any grant/funding information and a complete correspondence address.]

Contents

[Sun Coast Remediation Course Project Guidance 4](#_Toc21265889)

[Introduction 4](#_Toc21265890)

[Statement of the Problems 4](#_Toc21265891)

[Particulate Matter (PM) 4](#_Toc21265892)

[Safety Training Effectiveness 5](#_Toc21265893)

[Sound-Level Exposure 5](#_Toc21265894)

[New Employee Training 6](#_Toc21265895)

[Lead Exposure 6](#_Toc21265896)

[Return on Investment 6](#_Toc21265897)

[Research Objectives 7](#_Toc21265898)

[Research Questions and Hypotheses 7](#_Toc21265899)

[Statistical test based on data provided 8](#_Toc21265900)

[Data Analysis: Descriptive Statistics and Assumption Testing 8](#_Toc21265901)

[Correlation: Descriptive Statistics and Assumption Testing 8](#_Toc21265902)

[Simple Regression: Descriptive Statistics and Assumption Testing 10](#_Toc21265903)

[Multiple Regression: Descriptive Statistics and Assumption Testing 12](#_Toc21265904)

[Findings 15](#_Toc21265905)

[Recommendations 16](#_Toc21265906)

[References 17](#_Toc21265907)

# Sun Coast Remediation Course Project Guidance

Sun Coast provides remediation services to business and governmental organizations. Most of their contracts involve working within contamination sites where they remove toxic substances from soil and water. In addition to the toxicity of the air, water, and soil their employees come into contact with, the work environment is physically demanding and potentially contributory to injuries involving musculoskeletal systems, vision, and hearing. Sun Coast genuinely cares about the health, safety, and well-being of their 5,500 employees, but they are also concerned about worker compensation costs and potential long-term litigation from injuries and illness related to employment.

# Introduction

Senior leadership at Sun Coast has identified several areas for concern that they believe could be solved using business research methods. The previous director was tasked with conducting research to help provide information to make decisions about these issues. Although data were collected, the project was never completed. Senior leadership is interested in seeing the project through to fruition. The following is the completion of that project and includes the statement of the problems, literature review, research objectives, research questions and hypotheses, research methodology, design, and methods, data analysis, findings, and recommendations.

# Statement of the Problems

Six business problems were identified:

## Particulate Matter (PM)

There is a concern that job-site particle pollution is adversely impacting employee health. Although respirators are required in certain environments, PM varies in size depending on the project and job site. PM that is between 10 and 2.5 microns can float in the air for minutes to hours (e.g., asbestos, mold spores, pollen, cement dust, fly ash), while PM that is less than 2.5 microns can float in the air for hours to weeks (e.g. bacteria, viruses, oil smoke, smog, soot). Due to the smaller size of PM that is less than 2.5 microns, it is potentially more harmful than PM that is between 10 and 2.5 since the conditions are more suitable for inhalation. PM that is less than 2.5 is also able to be inhaled into the deeper regions of the lungs, potentially causing more deleterious health effects. It would be helpful to understand if there is a relationship between PM size and employee health. PM air quality data have been collected from 103 job sites, which is recorded in microns. Data are also available for average annual sick days per employee per job-site.

## Safety Training Effectiveness

Health and safety training is conducted for each new contract that is awarded to Sun Coast. Data for training expenditures and lost-time hours were collected from 223 contracts. It would be valuable to know if training has been successful in reducing lost-time hours and, if so, how to predict lost-time hours from training expenditures.

## Sound-Level Exposure

Sun Coast’s contracts generally involve work in noisy environments due to a variety of heavy equipment being used for both remediation and the clients’ ongoing operations on the job sites. Standard ear-plugs are adequate to protect employee hearing if the decibel levels are less than 120 decibels (dB). For environments with noise levels exceeding 120 dB, more advanced and expensive hearing protection is required, such as earmuffs. Historical data have been collected from 1,503 contracts for several variables that are believed to contribute to excessive dB levels. It would be important if these data could be used to predict the dB levels of work environments before placing employees on-site for future contracts. This would help the safety department plan for procurement of appropriate ear protection for employees.

## New Employee Training

All new Sun Coast employees participate in general health and safety training. The training program was revamped and implemented six months ago. Upon completion of the training programs, the employees are tested on their knowledge. Test data are available for two groups: Group A employees who participated in the prior training program and Group B employees who participated in the revised training program. It is necessary to know if the revised training program is more effective than the prior training program.

## Lead Exposure

Employees working on job sites to remediate lead must be monitored. Lead levels in blood are measured as micrograms of lead per deciliter of blood (μg/dL). A baseline blood test is taken pre-exposure and postexposure at the conclusion of the remediation. Data are available for 49 employees who recently concluded a 2-year lead remediation project. It is necessary to determine if blood lead levels have increased.

## Return on Investment

Sun Coast offers four lines of service to their customers, including air monitoring, soil remediation, water reclamation, and health and safety training. Sun Coast would like to know if each line of service offers the same return on investment. Return on investment data are available for air monitoring, soil remediation, water reclamation, and health and safety training projects. If return on investment is not the same for all lines of service, it would be helpful to know where differences exist.

# Research Objectives

Sun coast initiated different researches to improve their services. This section covers the three areas including efforts to reduce air pollution, noise pollution and measuring effectiveness of training sessions. Research objectives are listed below:

RO1: Determine if there is a relationship between PM (particulate matter) size and employee health.

RO2: Determine if training has been successful in reducing lost-time hours

RO3: Determine if frequency, angel, chord length, velocity and displacement contributed to noise (dB)

# Research Questions and Hypotheses

Based on the objectives above, the research questions and hypotheses are given below:

RQ1: Is there any relationship between PM (particulate matter) size and employee health.?

H01:There is no statistically significant relationship between PM (particulate matter) size and employee health.

HA1: There is a statistically significant relationship between PM (particulate matter) size and employee health.

RQ2: Have the training been successful in reducing lost-time hours?

H02: Training has not been successful in reducing lost-time hours

HA2: Training has been successful in reducing lost-time hours

RQ3: Does the frequency, angel, chord length, velocity and displacement contribute to noise (dB)?

H03: The frequency, angel, chord length, velocity and displacement does not contribute to noise (dB).

HA3: The frequency, angel, chord length, velocity and displacement contribute to noise (dB).

# Statistical test based on data provided

The chosen test for testing hypothesis 1, 2 and three are chi-square correlation test, simple regression and multiple regression respectively (Creswell & Creswell, 2017). The most suitable designs have been chosen for each question (Field, 2000).

# Data Analysis: Descriptive Statistics and Assumption Testing

This section provides the descriptive and specific test outputs from excel.

## Correlation: Descriptive Statistics and Assumption Testing



Fig. 1 Descriptive statistics of correlation test

It can be seen form the fig. 1 the total number of sites is 103, that the average size of PM is 5.66 microns and average number of annual sick days is 7 per site with same median and mode. The maximum particle size recorded is 10 and minimum 0.2.



Fig. 2 Histograms of correlation tests

The histograms show that the data follows a normal distribution pattern with a hump shaped curve.



Fig. 3(a) Correlation test result



Fig. 3(b) Correlation test result

The scatterplot shows a straight-line, proving correlation of two variables. (Fig. 3(a). Fig. 3(b)).

R-square value observed = 0.513

Pearson’s coefficient r is found to be -0.715

Hence the two variables, particle size and number of sick leaves show a strong relationship of 71.5% related to each.

## Simple Regression: Descriptive Statistics and Assumption Testing

The descriptive table below show the average expense on safety training to be 595.37 and lost hours to be 187.57. The maximum amount of money spent was 2271.86 and least was 20.45. the highest number of lost hours was 360 and minimum was 10.



Fig. 4 Descriptive statistics; safety training expenses and lost time hours



Fig. 5(a) Scatterplot showing correlation training expenses and lost time hours

The scatterplot of two variables show inverse relationship which means increase in money reduced the number of lost hours. The Pearson’s coefficient (r) before removing the outlier is 0.1434. hence, the relationship between the two variables was 14.34% before removal of outliers.



Fig. 5(b) Scatterplot showing correlation training expenses and lost time hours after removal of outliers

The relationship after removing outliers is 14.37%.



Fig. 6(a) Simple regression test result

The results of regression analysis show the variables are 88% related to each other.



Fig. 6(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

## Multiple Regression: Descriptive Statistics and Assumption Testing

Students should include this information here. Include frequency table, histogram, and descriptive statistics table. Evaluate and discuss the descriptive statistics and make an explicit statement about whether the assumptions for parametric statistical testing were met or not met. Delete these statements before you begin.

****

Fig. 7 (a) Descriptive statistics of frequency, angle and chord length

****

Fig. 7 (b) Descriptive statistics of velocity, displacement and noise in decibels

From fig 7(a) and fig 7(b) the mean frequency observed is 2886.38, mean angle 6.78, chord length 0.116, displacement 0.011 where mean noise value in decibels is 124.84

****

Fig. 8 Scatterplots of variables affecting decibels

****

Fig. 9 (a) Multiple regression output

****

Fig. (b) Multiple regression output

The regression outputs show non-significant values for angle and chord length and a significant value for frequency, velocity and displacement.

Dependent variable = 126.82 – 0.00 (frequency) + 0.08(velocity) – 240(displacement)

# Findings

By looking at the research objectives, it can be said that we have met all of them.

RO1: Determine if there is a relationship between PM (particulate matter) size and employee health.

RO2: Determine if training has been successful in reducing lost-time hours

RO3: Determine if frequency, angel, chord length, velocity and displacement contributed to noise (dB)

Hence, we accept our experimental hypotheses that there is statistically significant relationship between particulate matter size and employee health, Training has been successful in reducing lost-time hours. However, the third research proves half of the factors affect the decibels of noise.

# Recommendations

It is recommended to increase amount of money spent on training as it helps to improve performance, and efforts must be made to reduce lead exposure. Additionally, displacement, frequency and velocity of sounds must be addressed to reduce noise pollution.

# 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.