

Linear Regression Project

In this project, you will select 2 quantitative variables that you believe may be related in your selected data set. You will then investigate the relationship by constructing a scatter plot, computing correlation, and conducting a linear regression analysis.

Project Tasks

Below is the overall sequence of tasks you will follow to complete the linear regression project.

1. Choose a research question that can be addressed using the given data set.
2. Carry out your research:
 - a. Conduct linear regression analysis
 - b. Write your results in a report, using the outline given below
3. Upload your written report in a WORD file.

Project Report Outline

Below is a detailed outline of the content that should be included in your project report. The components are listed in outline form so that they can be used as a checklist. However, your project report is expected to be a formal paper (not an outline). Your results should be stated in complete sentences, and your paper should be written in paragraph form. Although you may choose to use headings, you should not number your paragraphs.

1. **Introduction.** State the topic of your study as a research question and as a specific hypothesis to be tested. Your hypothesis should indicate what type of correlation you expected to see (positive or negative) and how strong you expected the correlation to be (weak, moderate, strong). Your hypothesis should describe a specific result that you expected to find **and** the *practical reason* that you expected this result (your rationale).
2. **Define Population(s).** Define clearly the population(s) that you intend for your study to represent.
3. **Define Variable(s).** Define clearly the variable(s) that are analyzing (e.g. age, salary, price, miles per gallon, etc.) This must be specific: "time spent watching TV" is too vague; "number of hours spent watching TV in the last 3 days" would be specific enough. If your variable is a measurement (e.g. height) give units (e.g. inches). If your variable is score (e.g. on an achievement test), give the range of possible scores.
4. **Study Design.** Identify statistical procedures you used to analyze your data. Give relevant design details (e.g., which variable was selected as the explanatory variable, and which the response variable? Why? What type of correlation did you expect? And so on.)
5. **Results. Descriptive Statistics.** Give descriptive statistics **for each of your two quantitative variables.** Note that you will be reporting summary statistics for both your explanatory variable and your response variable. Report each set of descriptive statistics using both a table and a chart as described below. All tables and charts should be placed directly in your report.
 - a. **Table:** Give sample size, mean, standard deviation, and a 5-number summary for each variable.
 - b. **Chart:** Show boxplots that illustrates the distribution of each variable (using Boxplot Grapher)

6. **Results. Statistical Analysis.** Report the results of your analysis; include items below.
 - a. Scatter plot with a graph of the regression line
 - b. Value of the correlation coefficient r and interpretation of its meaning
 - c. Equation of the regression line
 - d. An example of a prediction using the regression equation
 - e. Value of r^2 for the regression model and interpretation of its meaning

7. **Findings.** Interpret the results of your statistical analysis in the context of your original research question. What possible lurking variables could have caused your results? Do your analyses support your expected findings? Explain.

8. **Discussion.** What conclusions, if any, do you believe you can draw as a result of your study? If the results were not what you expected, what factors might explain your results? What did you learn from your project about the population(s) you studied? What did you learn about the research variables? What did you learn about the specific statistical analysis you conducted?