

# Linear Regression Project

---

In this project, you will perform regression analysis on a set of data to develop a mathematical model that relates two variables. You will then use this model to make predictions.

## Objectives

Research Data

Produce a scatter plot using technology

Perform a regression analysis to find the equation of the line that best fits the data.

Display the results, plotted data and the regression equation together for visual comparisons.

Use the model to make predictions

Make conclusions about the data

## Guidelines

### 1. **Research and select your data.**

You are responsible for finding data and coming up with a question of interest.

Some sites to check for possible data include the following:

**Music Charts** - <https://www.billboard.com/charts>

**NFL Stats** – [www.nfl.com/stats](http://www.nfl.com/stats)

**Varied Statistics** - <https://www.usa.gov/statistics>

Source: \_\_\_\_\_

Question of Interest: \_\_\_\_\_

---

You will need at least **20 points**. You must define the explanatory and response variables. Clearly explain if it is appropriate to assign an explanatory and response variable OR if it is just a relationship between two variables. Check with your teacher to be sure the data you have data selected that would be a good fit for the project. *You may not use "year" or "season number" as a variable.*

Explanatory: \_\_\_\_\_ Response: \_\_\_\_\_

EXPLAIN:

### 2. **Create a scatter plot.**

Make a scatter plot using Microsoft Excel with these two variables. Include the appropriate labels, line of best fit and r-squared.

| Instructions using Excel   |  |   |
|--|--|---|
| <b>Step 1:</b> Input data into excel.<br>Column A = Explanatory data<br>Column B = Response data | <b>Step 2:</b> Select both columns by clicking the top left box and dragging your cursor to the bottom left box. | <b>Step 3:</b><br>Click <b>Insert</b> → <b>Recommended Charts</b> → <b>Scatter</b><br><br>Copy and paste scatterplot into word document |

**Instructions using Excel (CONT)**

**To display regression equation and R-Squared**

- 1) Left click on any data point.
- 2) Choose **"Add Trendline"** (towards the bottom of the list)
- 3) Choose **"Linear"**
- 4) Check **"Display Equation"** and **"Display R-Squared"**

**3. Regression analysis.**

Include the following:

- Regression Equation: \_\_\_\_\_
  - a. Interpret slope and y-intercept in context

Slope in context:

Y-Intercept in context:

- Correlation coefficient and description in context

Correlation Coefficient: \_\_\_\_\_ Strength and Direction: \_\_\_\_\_

Description in context:

- Predicted values and residuals – Identify any potential outliers – an example of a residual table is given below.

| x | Equation | Actual Y | Predicted Y | Residual |
|---|----------|----------|-------------|----------|
|   |          |          |             |          |

Potential Outliers:

What does a **positive** residual mean in context? **Negative?**

- R-squared with explanation

R<sup>2</sup>: \_\_\_\_\_ Explanation:

- Possible Lurking Variables

Percent due to lurking variables: \_\_\_\_\_ Name at least two lurking variables: \_\_\_\_\_

---

**4. Make conclusions**

Answer the following:

1. Is the regression line a good fit for your data? Why or Why not? **You must use linear regression vocabulary and calculations.**

2. What did you learn about your question of interest? Were your results surprising? Why or Why not?

Your report will be **at least 2 pages** not including the residual table. **The residual table must be included on a separate piece of paper.** Your report will be **typed, 12pt, times new roman font.**