# **Correlation and Simple Linear Regression V2**

## **Project Description:**

The present study shows data for prices per gallon of paint and the demand for each brand of paint. The sales manager at a home improvement store is interested in knowing if there is a significant relationship between the price a gallon of paint and the demand (number of gallons sold). A sample of 28 different paint brands is collected and analyzed.
We will use Excel formula to find the values of the covariance and correlation coefficient and analyze what the values mean. We will use the Excel Data Analysis Add-in to build a regression model and find the values of the errors in the model. We will identify the value of the coefficient of determination and analyze its meaning. We will use the regression model to make predictions. We will test the claim that there is no relationship between the two variables. We will compare the observed level of significance (p-value) with alpha and decide if we can conclude that there is relationship between the two variable.

## **Steps to Perform:**

| **Step** | **Instructions** | **Points Possible** |
| --- | --- | --- |
| 1 | Use a cell reference or a single formula where appropriate in order to receive full credit. Do not copy and paste values or type values, as you will not receive full credit for your answers.  Start Excel. Download and open the workbook named: **Correlation\_and\_Simple\_Linear\_Regression\_Start** | 0 |
| 2 | On the Data sheet, go to the **Data Tab** and click on **Data Analysis**. Double click on "Regression." For **Input Y Range**, select C1:C29 and for **Input X Range**, select B1:B29. Check the **Labels** box and select **Output Range** under Output Option and type E1. | 4 |
| 3 | In cell C4, find the covariance between the demand and price for the sample on the Data sheet.  Hint: use the COVARIANCE.S function. | 4 |
| 4 | In cell C5, find the correlation between the demand and price for the sample on the Data sheet.  Hint: use the CORREL function. | 4 |
| 5 | By assessing the values in cells C4 and C5, what can you say about the relationship between the Price of the paint and Demand? Choose your answer from the dropdown menu in cell C6. | 4 |
| 6 | Explain the answer you chose in cell C6. Choose your answer from the dropdown menu in cell C7. | 4 |
| 7 | In cell C8, find the Y-Intercept (B0) of the Regression Equation from the Data Analysis output table on the Data sheet. | 4 |
| 8 | In cell C9, find the Price (X variable) coefficient (B1) from the Data Analysis output table on the Data sheet. | 4 |
| 9 | In cell C10, find the Coefficient of Determination (R2) from the Data Analysis output table on the Data sheet. | 4 |
| 10 | In cell C11, find the Regression Sum of Squares (SSR) from the Data Analysis output table on the Data sheet. | 4 |
| 11 | In cell C12, find the Error Sum of Squares (SSE) from the Data Analysis output table on the Data sheet. | 4 |
| 12 | In cell C13, find the Total Sum of Squares (SST) from the Data Analysis output table on the Data sheet. | 4 |
| 13 | In cell C14, find the value of Coefficient of Determination (R2) using SSR and SST. | 5 |
| 14 | Explain what the value of R2 means. Choose your answer from the dropdown menu in cell C15. | 5 |
| 15 | In cell C16, find the lowest price for which we can predict demand. | 5 |
| 16 | In cell C17, find the highest price for which we can predict demand. | 5 |
| 17 | In cell C18, predict the demand for a paint that is priced at $50 per gallon. | 5 |
| 18 | Explain the relationship between the demand and price. Choose your answer from the dropdown menu in cell C19. | 5 |
| 19 | Is the correlation coefficient related to the Coefficient of Determination? Choose your answer from the dropdown menu in cell C20. | 5 |
| 20 | Test the hypothesis that there was no linear relationship between the paint price and demand.  In cell C21, find the observed level of significance (p-value) from the Data Analysis output table on the Data sheet. | 5 |
| 21 | Do you reject the claim that there is no linear relationship at alpha=0.05? Choose your answer from the dropdown menu in cell C22. | 5 |
| 22 | Explain your answer in cell C22. Choose your answer from the dropdown menu in cell C23. | 5 |
| 23 | In cell C24, insert a Scatter Plot that displays the relationship between the price and demand.  On the Data sheet, highlight the range B1:C29, click on the **Insert tab**, **Scatter charts**, then choose the first scatter plot. Right-click on the data point and select "**Add Trendline**". In the Format Trendline Window, check "**Display R-squared value on chart**". Select the Scatter Plot, cut and paste it in cell C24 and resize it to fit in the cell. | 6 |
| 24 | Save your file and submit for grading. | 0 |

|  |  |
| --- | --- |
| **Total Points** | **100** |