

CH 10: CORRELATION AND REGRESSION

The components and steps in Regression analysis consist of:

1. Create a **Scatterplot** to decide if there is a linear correlation
2. Calculate the correlation coefficient
3. Test the significance of this coefficient based on some α
4. Find the equation of the regression line
5. Predict values based on the regression line if the correlation is significant.

Minitab provides help in performing a number of the above tasks

For all of the discussion that follows, we assume the following data, already in a Minitab worksheet.

	C1	C2	
	Temp	Calls	
1	68	7	
2	74	4	
3	82	8	
4	88	10	
5	93	11	
6	99	9	
7	101	13	
8			

1. Create Scatterplot

- **Graph** → **Scatterplot**, then **Simple**
- Column C2 for **X-variable** and column C3 for **Y-variable**
- Click **Data View** and select **Symbols** for the data display.
- Click on **Labels** to set up the Title for the graph

2. Correlation Coefficient:

- **Stat** → **Basic Statistics** → **Correlation**
- Place the columns containing the x and y data into the **Variables** box.
- Check the box for **Display p-value**.
- Click OK. The output is placed in the session window.

Correlations: Temp, Calls

Pearson correlation of Temp and Calls = 0.811
P-Value = 0.027

3. To Test the significance of the correlation, use Hypothesis testing

4. Regression Line

- **Stat** → **Regression** → **Regression**
- Place the X variable (Temp) in the **Predictors** box
- Place the Y variable (Calls) in the **Response** box.
- Click OK
- The output is place in the session window

Regression Analysis: Calls versus Temp

The regression equation is
Calls = - 7.54 + 0.190 Temp

5. Predictions.

Assuming that the Hypothesis test above showed the correlation to be significant, you can use the above regression line to make predictions. For example, how many calls would be made if the temperature is 80 degrees?

$Calls = -7.54 + 0.19 \cdot (80) = 7.66$ or approximately 8 calls.