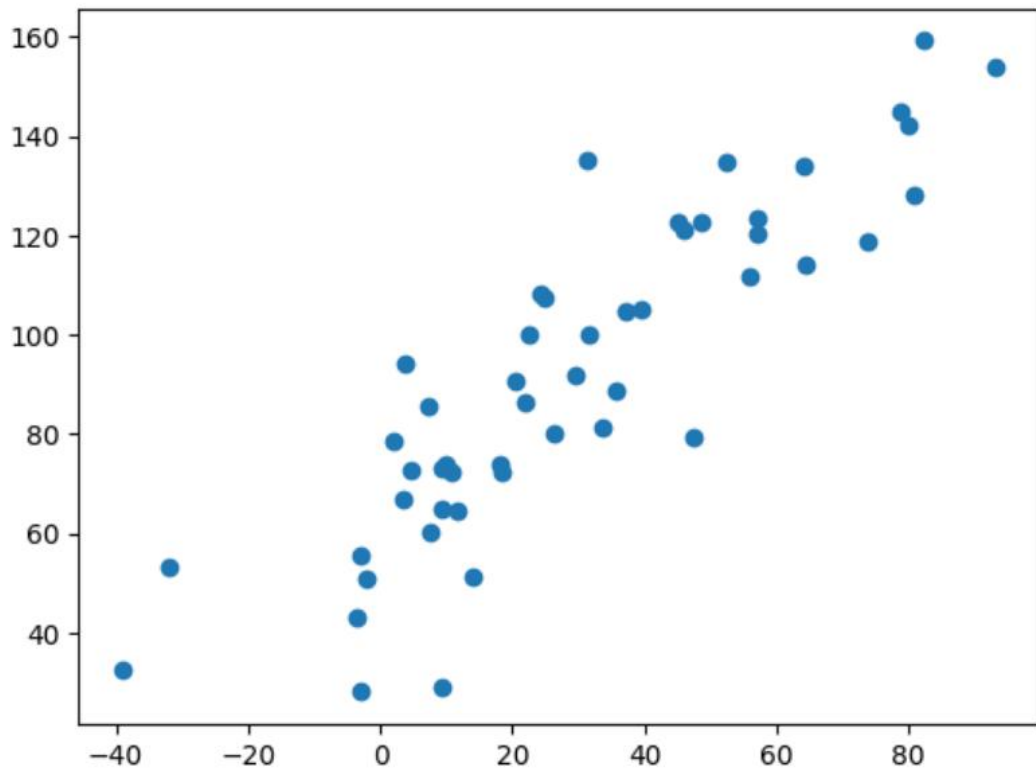


Activity 1

➤ Covariance Pearson correlation:

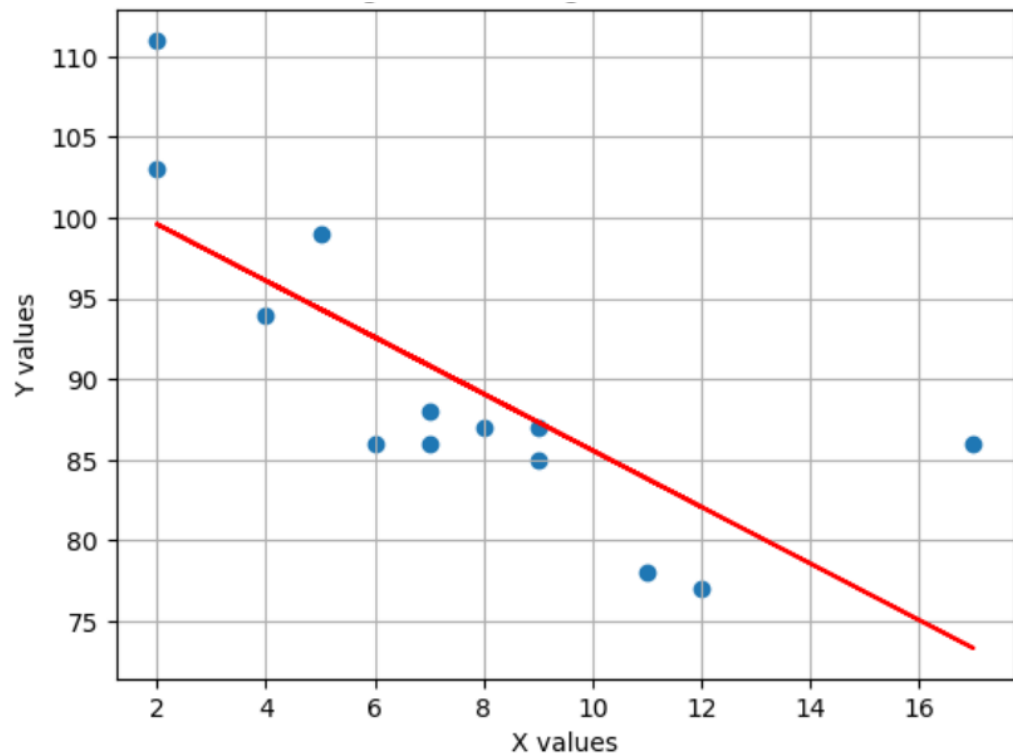


Data Summary:

- Data1: Mean = 29.235, Standard Deviation = 29.088
- Data2: Mean = 92.168, Standard Deviation = 32.717
- Covariance = 852.887
- Pearson's Correlation = 0.878

In this activity, we examined the relationship between two numerical variables using both covariance and Pearson's correlation. Covariance provided insight into the direction of the relationship indicating whether it was positive or negative, while Pearson's correlation offered a standardized measure of its strength and direction. Using Python's *scipy.stats* module, we computed both metrics and visualized the results through scatter plots, enhancing our understanding of the linear association between the two variables.

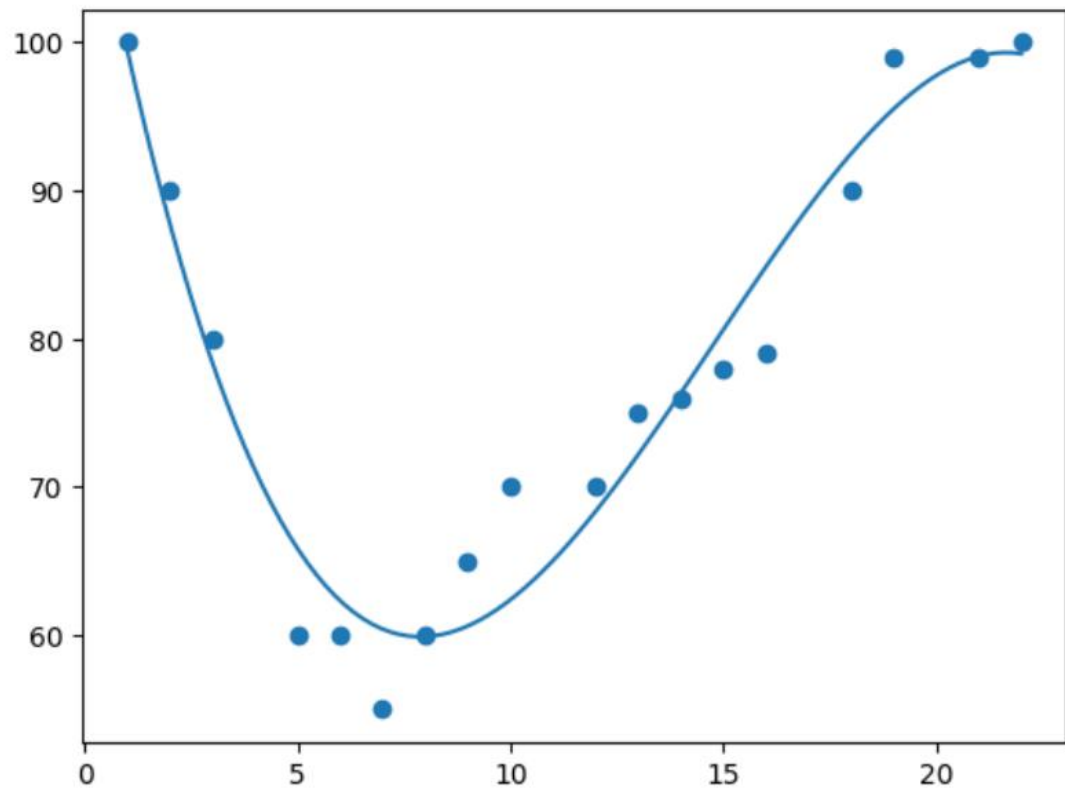
➤ **Linear Regression:**



- **Pearson's Correlation: -0.759**

In this activity, we used multiple Google Colab notebooks to explore common issues found in raw datasets and the methods used to address them. The resulting figure illustrates a linear regression model developed based on Pearson's correlation analysis, which revealed a moderate negative linear relationship between the variables

➤ Polynomial Regression



This activity involved applying polynomial regression to capture a non-linear trend in the dataset. Using NumPy's *polyfit* function with a third-degree polynomial, we modeled the curvature more precisely. A smooth regression curve was plotted with *linspace* to create detailed x-values, helping reveal complex patterns that linear models fail to detect in real-world data.