Title: Seattle Weather Prediction Using Linear Regression

**1. Detailed Description of Procedures**

In this project, we aim to analyze and predict the maximum temperature in Seattle based on historical weather data. To achieve this, we follow several steps, including data preprocessing, visualization, feature selection, model training, evaluation, and prediction.

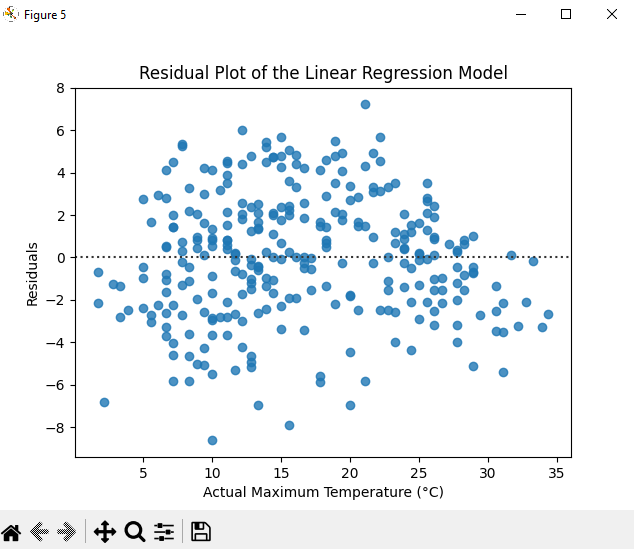
* Data Preprocessing: The dataset is loaded into a Pandas DataFrame and cleaned by removing missing or duplicate values. Outliers and inconsistencies in the data are handled by setting extreme temperature values to NaN. The 'date' column is converted to a datetime format, and additional features such as 'year', 'month', and 'day' are extracted.
* Visualization: Line plots, scatter plots, and histograms are created using Matplotlib and Seaborn libraries to visualize the relationships between variables and the distribution of maximum temperatures in the dataset.

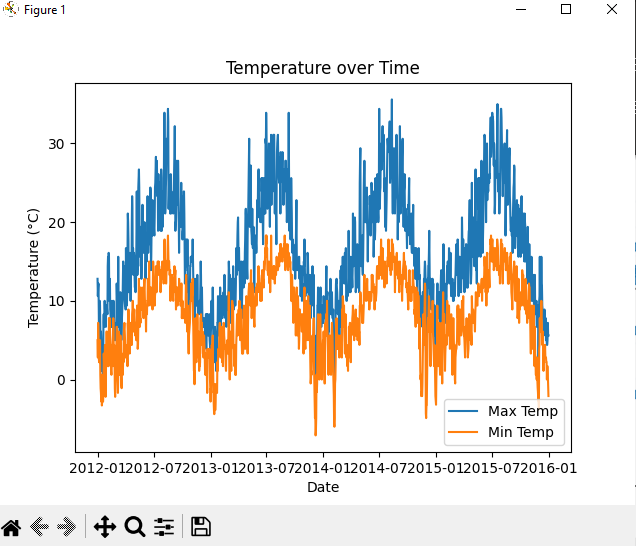
**2. Techniques and Algorithm**

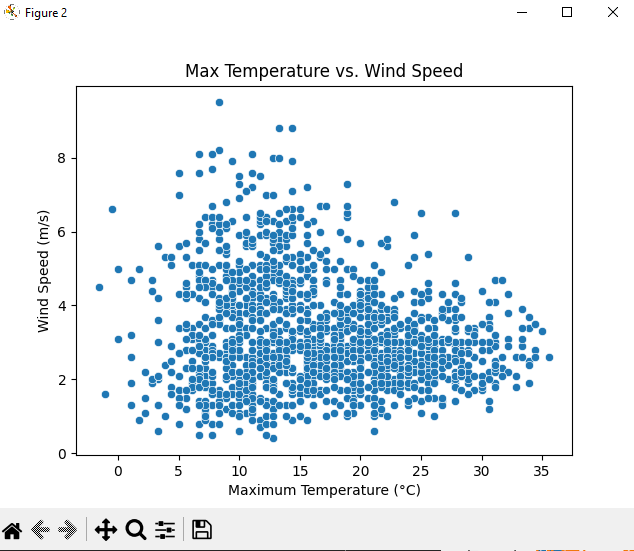
* Feature Selection: A list of relevant features is created, including 'year', 'month', 'day', 'temp\_min', 'wind', and 'precipitation'. The target variable is set as 'temp\_max'.
* Data Splitting: The dataset is split into training (80%) and testing (20%) sets using the train\_test\_split function from Scikit-learn.
* Model Training: A Linear Regression model is trained using the training set, with the selected features as input and the target variable as output.
* Model Evaluation: The trained model is evaluated using the test set to calculate the Mean Squared Error (MSE) and R2 score.
* Model Prediction: The model is used to predict the maximum temperature for a specific date with given feature values.
* Visualization of Model Performance: Scatter plots and residual plots are created to visually assess the model's performance.

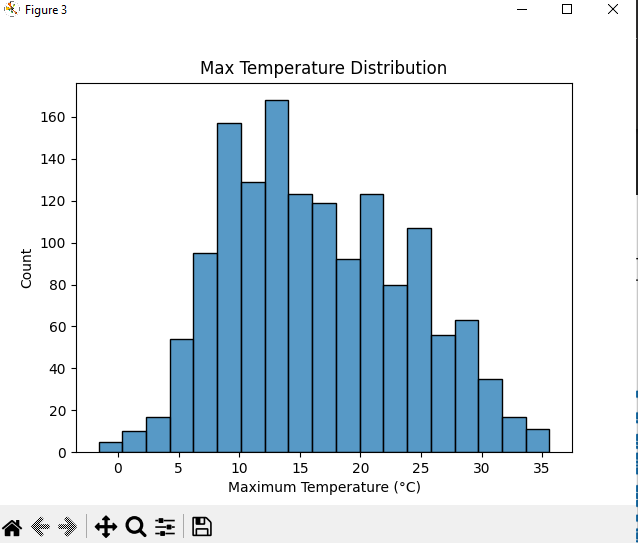
**3. Results**

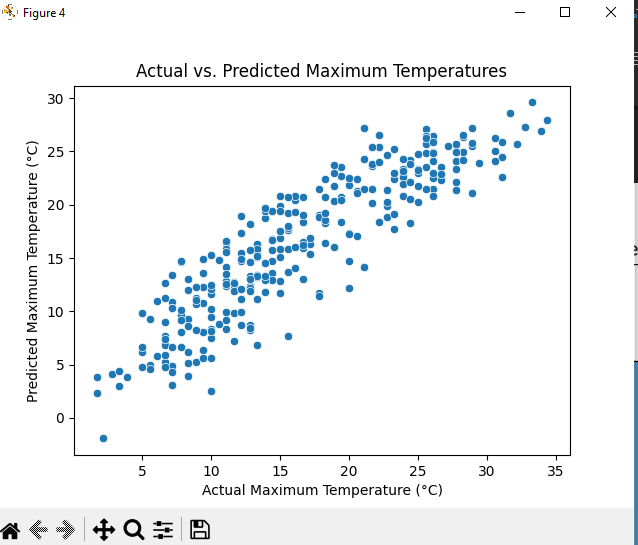
**Visualisation**



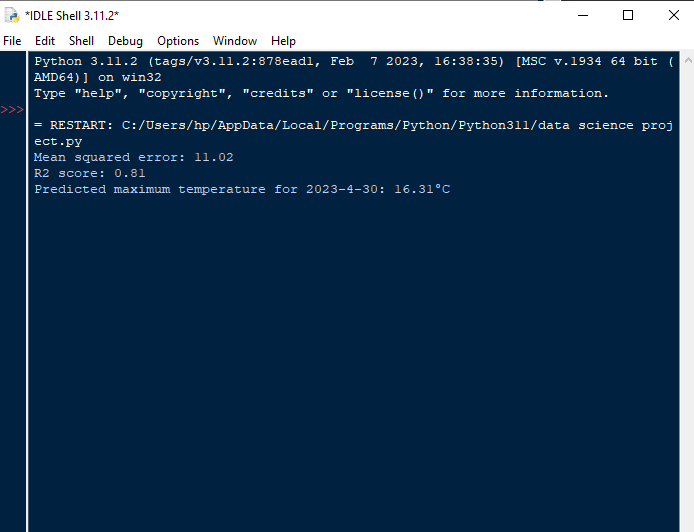








**Prediction**

****

**4. Conclusion**

The Linear Regression model has been successfully trained and evaluated on the Seattle weather dataset. The resulting R2 score indicates the proportion of variance in the dependent variable that is predictable from the independent variables. While the model provides a reasonable prediction of maximum temperatures, more complex algorithms, such as decision trees or neural networks, could potentially improve the model's performance.

In summary, this project demonstrates how to analyze and predict weather data using a Linear Regression model. Further improvements could be made by incorporating additional features or trying more advanced machine learning algorithms.