# ICP – 3

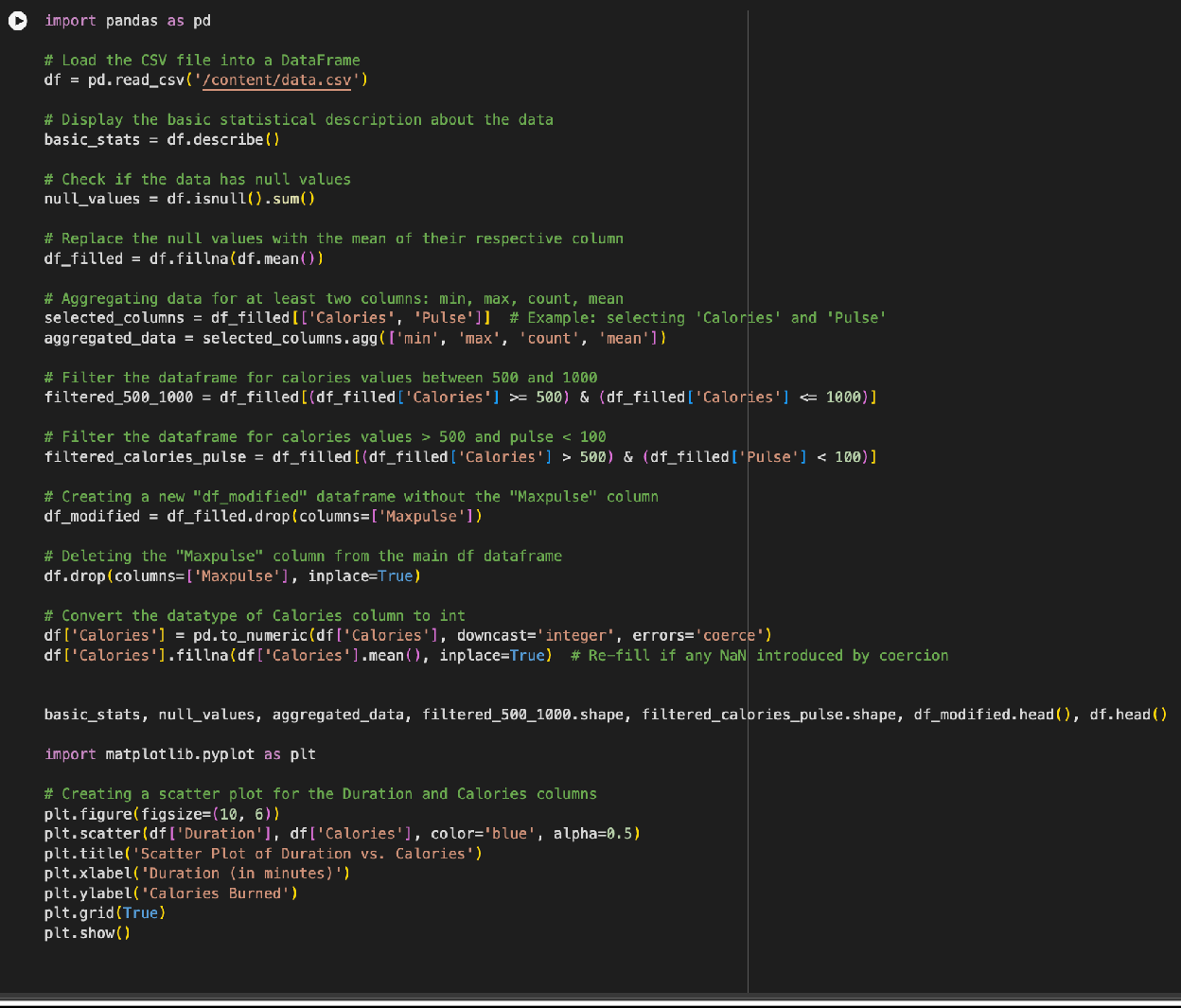
# Bhavana Parasa

**700761185**

**GitHub:**

# Q1:

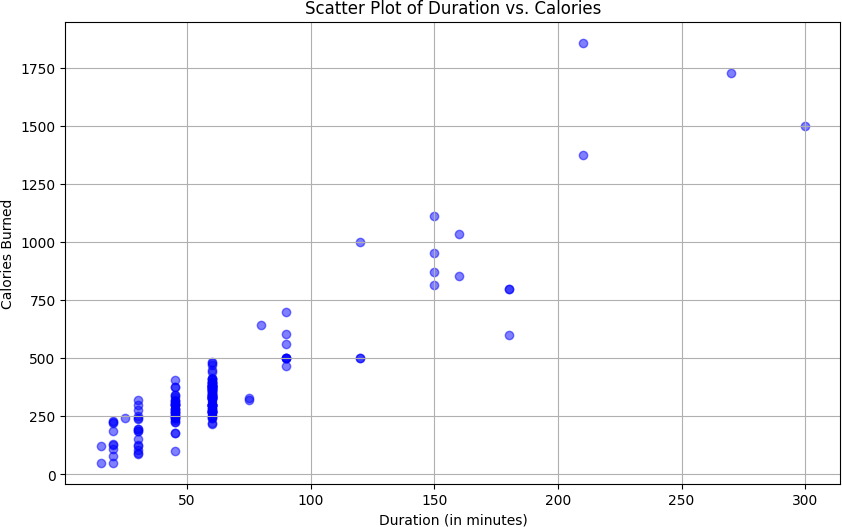
**Code:**



# Explanation:

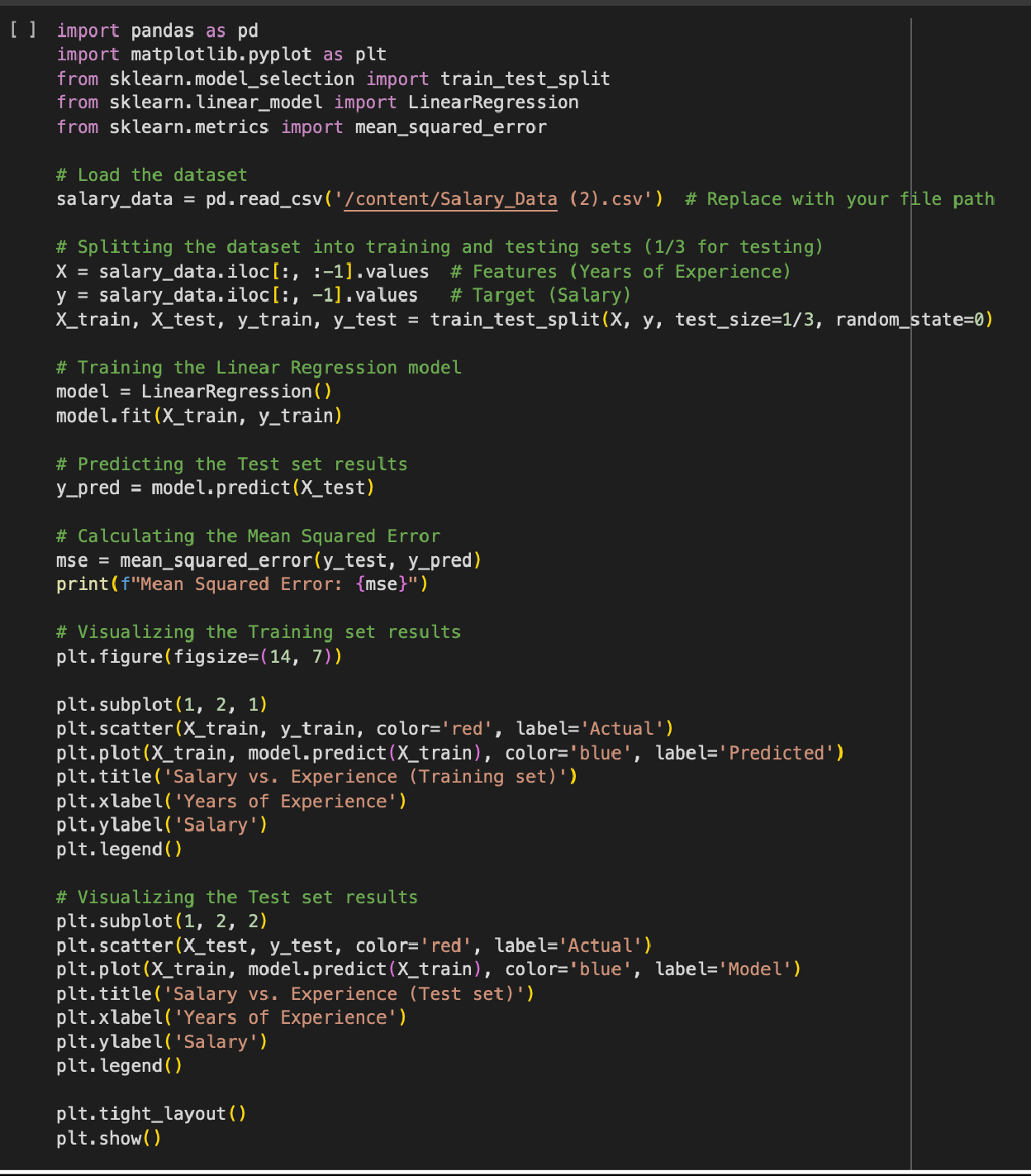
1. **Data Loading and Preparation**: Reads a dataset from a CSV file into a DataFrame and checks for null values.
2. **Statistical Analysis**: Generates basic statistical descriptions of the dataset, such as mean, standard deviation, minimum, and maximum values.
3. **Handling Missing Data**: Replaces any missing values in the dataset with the mean of their respective columns.
4. **Data Aggregation**: Aggregates data for 'Calories' and 'Pulse' columns to find their minimum, maximum, average, and count.
5. **Data Filtering**: Creates subsets of the data based on specific criteria related to 'Calories' and 'Pulse' values.
6. **DataFrame Modiﬁcation**: Removes a specific column ('Maxpulse') from the dataset and adjusts the 'Calories' column to integer data type.
7. **Data Visualization**: Constructs a scatter plot to visually represent the relationship between 'Duration' and 'Calories' in the data.

# Output:



**Q2:**

# Code:



**Explanation:**

1. **Imports Necessary Libraries**: Utilizes pandas for data handling, matplotlib for visualization, and sklearn for machine learning tasks.
2. **Loads and Prepares Data**: Reads a salary dataset into a DataFrame and splits it into features (years of experience) and target (salary).
3. **Splits Data into Train and Test Sets**: Divides the data into training and testing subsets, with one-third of the data reserved for testing.
4. **Trains a Linear Regression Model**: Fits a linear regression model to the training data to understand the relationship between experience and salary.
5. **Predicts Salaries and Evaluates Model**: Uses the model to predict salaries on the test set and calculates the mean squared error to assess the model's performance.
6. **Visualizes Results**: Creates scatter plots to visually compare actual vs. predicted salaries for both training and testing data.

# Output:

