BY : Ahmed Tarek Mohamed Sabry

1. I imported some libraries like "numpy , pandas , sklearn.svm "

2. using the pandas library to read a CSV (Comma-Separated Values)

3. head() returns the first 5 rows, but specifying 20 returns the first 20 rows.

4. **df.drop()**: This method is used to drop (i.e., remove) specified rows or columns from a DataFrame   axis=0 refers to rows, and axis=1 refers to columns.

5. Check nulls and drop it and check duplicated

6. clean up and fix the InvoiceDate column in a pandas. It specifically addresses issues where the time part of the InvoiceDate column contains invalid hours (e.g., hours greater than 23), which would make it impossible to parse the time correctly.

   Summary
   The function fix_hour cleans up time strings in the InvoiceDate column by fixing invalid hours.
   It applies the function to each entry in the InvoiceDate column.
   Finally, the cleaned column is converted to a valid datetime format, making it suitable for further processing.

7. convert the InvoiceDate column in the DataFrame df to datetime objects.

8. Day , month , years

9. Define fearures and target

10. Split data into training and testing sets

11. using the StandardScaler from scikit-learn to standardize (scale) the numerical features in the training and testing datasets

12. This code snippet is using the Support Vector Regression (SVR) model from the scikit-learn library to perform regression tasks.

13. Make predictions on the test data

14. Calculate MSE and R^2

15. visualizing the results of a regression model by creating various plots to compare actual values with predicted values.