**Task 1 – Pseudocode and Explanation**

# Pseudocode

BEGIN PROGRAM

IMPORT dataset "D598 Data Set.xlsx" INTO dataframe df

// Step 1: Identify duplicate rows

SET duplicates = rows in df that are duplicates

// Step 2: Group data by state and calculate descriptive statistics

SET state\_stats = group df by Business State

CALCULATE mean, median, min, max for numeric columns

// Step 3: Find companies with negative debt-to-equity ratio

SET negative\_DE = rows in df where Debt-to-Equity Ratio < 0

// Step 4: Compute Debt-to-Income Ratio for each business

FOR each row in df:

IF Revenue = 0:

IF Long-Term Debt = 0:

SET Debt-to-Income Ratio = 0

ELSE:

SET Debt-to-Income Ratio = 1

ELSE:

SET Debt-to-Income Ratio = Long-Term Debt ÷ Revenue

ADD Debt-to-Income Ratio column to df

// Step 5: Concatenate results

SET final\_df = df with Debt-to-Income Ratio column

OUTPUT duplicates, state\_stats, negative\_DE, final\_df

END PROGRAM

# Explanation of the Relationship Between Flowchart and Pseudocode

The pseudocode represents the **step-by-step logic** of the program in a textual format. It begins with importing the dataset, proceeds through identifying duplicates, summarizing statistics by state, filtering out companies with negative debt-to-equity ratios, computing a new debt-to-income ratio for each company, and finally concatenating the results into the final dataset. Each step is clearly written in plain language that mirrors the tasks shown in the flowchart.

The **flowchart** provides a **visual overview** of the program’s logic, while the pseudocode offers **structured written instructions**. Both align perfectly: for example, the “Compute Debt-to-Income Ratio” box in the flowchart corresponds to the loop and conditional logic in the pseudocode. Together, they demonstrate the planning phase of the project — the flowchart shows *what happens and in what order*, while the pseudocode specifies *how each step should be carried out*.