

## **PROBLEM STATEMENT**

Generate an **AI solution** that could help users modify their csv/excel data automatically, functions include: pick out the duplicate data, display numbers according to certain rules, add or delete rows. The program should support input the instructions/prompt and csv file, output the results. And the progress of modifying the excel should be recorded.

## **SOLUTION**

### **1. Data Loading and Initial Setup**

start by installing the openpyxl package and importing pandas to handle Excel files.

load an Excel file (dummy\_data.xlsx) into a DataFrame df and display its head to inspect the initial data.

### **2. Automated Data Modification Class**

define a class automated\_modification that encapsulates various data manipulation methods:

- pick\_duplicates: Identifies duplicate rows in the DataFrame.
- pick\_unique: Identifies unique rows in the DataFrame.
- remove\_duplicates: Removes duplicate rows, keeping the first occurrence.
- modification\_record: Maintains a history of modifications made to the DataFrame, including timestamps and the number of duplicate/unique rows.
- add\_row: Adds a new row to the DataFrame, accepting either a dictionary or a list/tuple.
- remove\_row: Removes a row by its index.
- filter\_rows: Filters rows based on specified conditions, supporting both exact matches and comparisons using operators.
- interactive\_modification: Provides an interactive console-based interface for users to modify the DataFrame through various operations

### **3. AI Interaction Class**

define another class AI Interaction that integrates AI capabilities to process user inputs:

- extract\_numbers: Extracts numbers from a given sentence using regular expressions.
- detect\_intent: Uses a rule-based approach to detect user intent from input strings, mapping them to specific actions like adding a row, removing a row, etc.
- process\_input: Maps detected intents to corresponding methods in the automated\_modification class to perform data operations.

Finally, I instantiate AI Interaction to process user inputs and perform data operations based on detected intents, with the understanding that changes are temporary unless saved.