BABU BANARASI DAS UNIVERSITY

Customer Insights Telecom Case Study

SUBMITTED TO:	SUBMITTED BY:
Mr. Ayushman	Name: Yash Tiwari
	Roll No: 1230258481
	Class: BCADS-36

Definition: The Unit of Analysis refers to the level at which data is grouped and analyzed. It is used to:

- Remove duplicate rows, ensuring each record is unique based on selected fields.
- Summarize multiple rows into one per group, thereby changing the unit of analysis.
- Create a binary (0/1) field based on specific conditions to label or filter data at a chosen level.
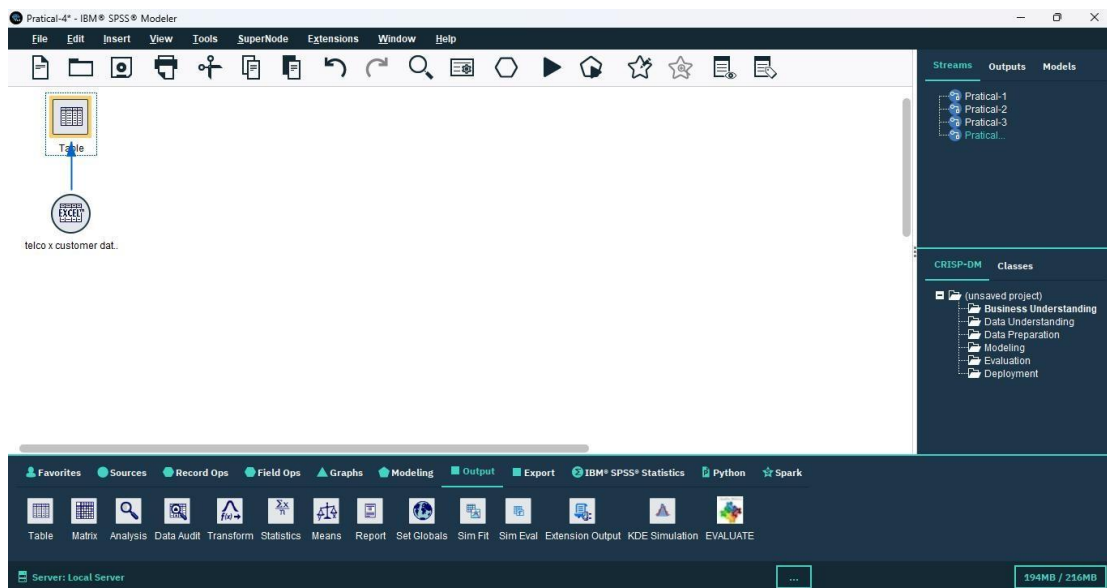
Outcomes/Learning: Learned to remove duplicate data, merge revenue using aggregation, and assign values using flag fields.

Required Tool: IBM SPSS Modeler tool.

Working: We used this data to identify and retain valid customer information, keeping only the known and relevant records while discarding the rest.
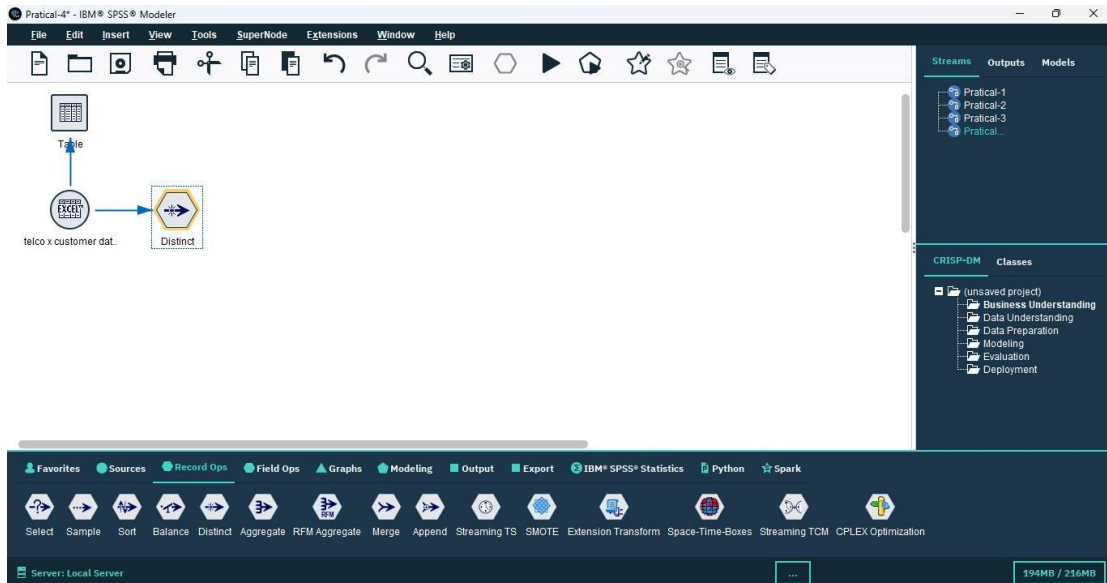
Step 1:

- Open **IBM SPSS Modeler**.
- Drag the **Excel Node** from the Source tab to the canvas.
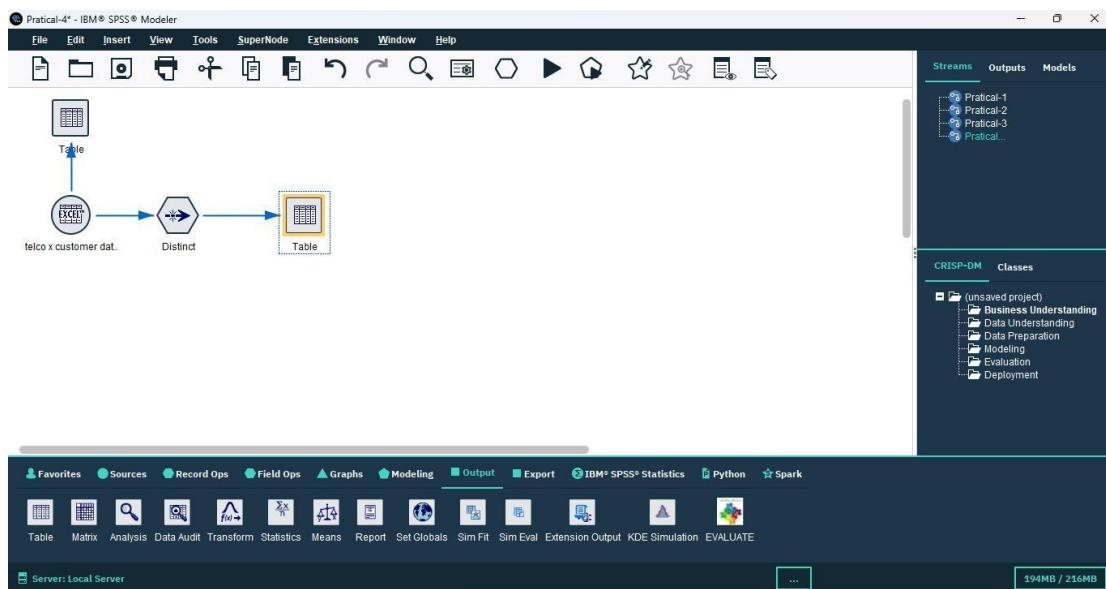- Import the Excel dataset into this node.

Step 2:

- Add a **Record Ops Node** and connect it to the Excel Node.
- Choose **Select Distinct** to remove duplicate records.
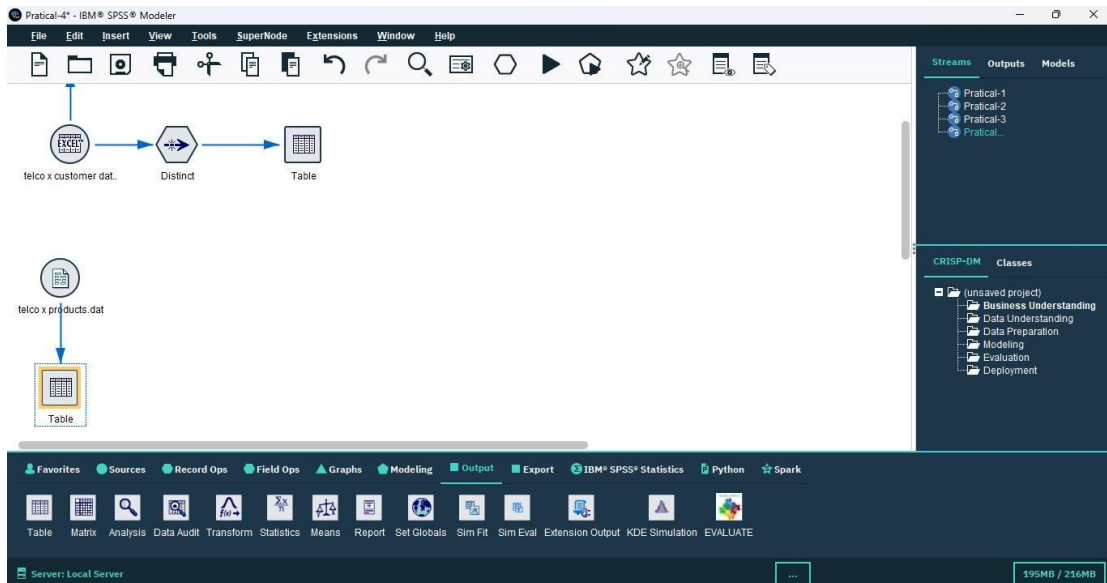- Set all fields as **key values** to keep only unique data.

Step 3:

- Connect a **Table Node** to the Record Ops Node.
- Run it to **verify duplicate removal**.
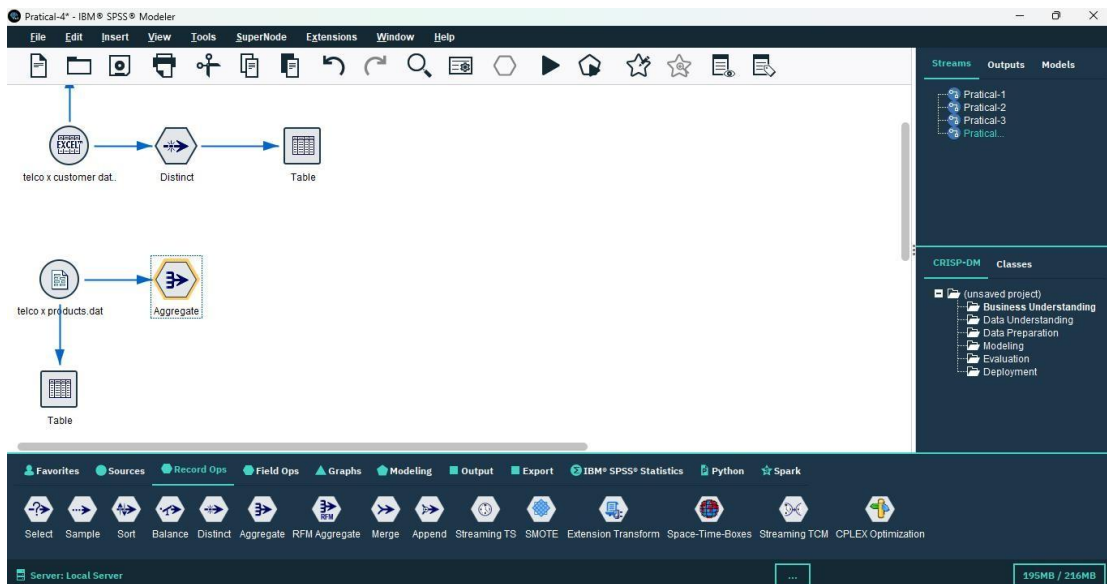- Check that only unique rows remain.

Step 4:

- Add a **Var File Node** to import a flat data file.
- Connect a **Table Node** to display the imported data.
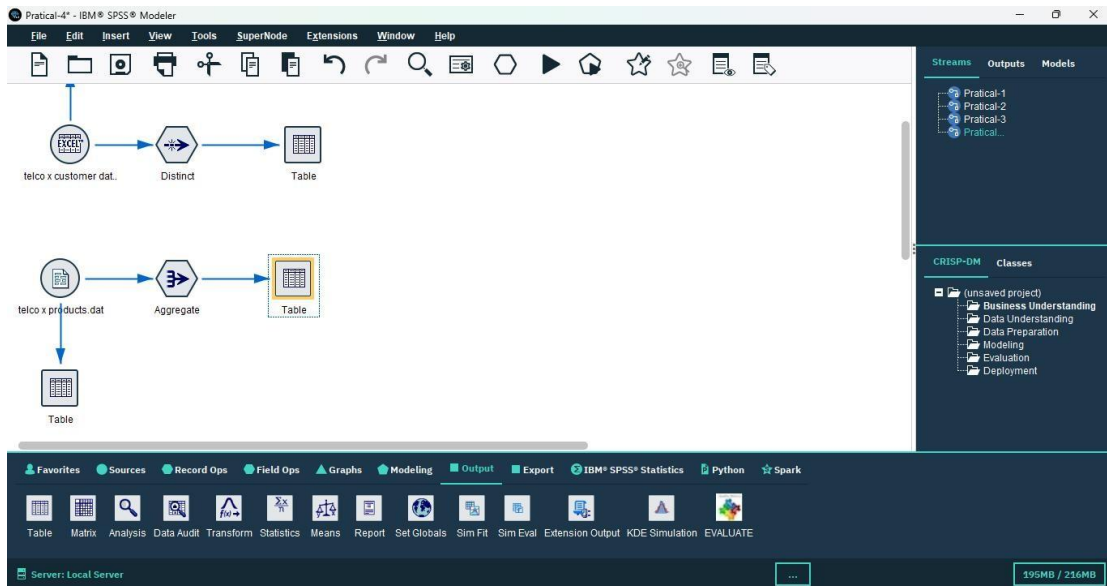- Ensure the data loads correctly.

Step 5:

- Connect a **Record Ops Node** to an **Aggregate Node**.
- Select **Revenue** as the key field.
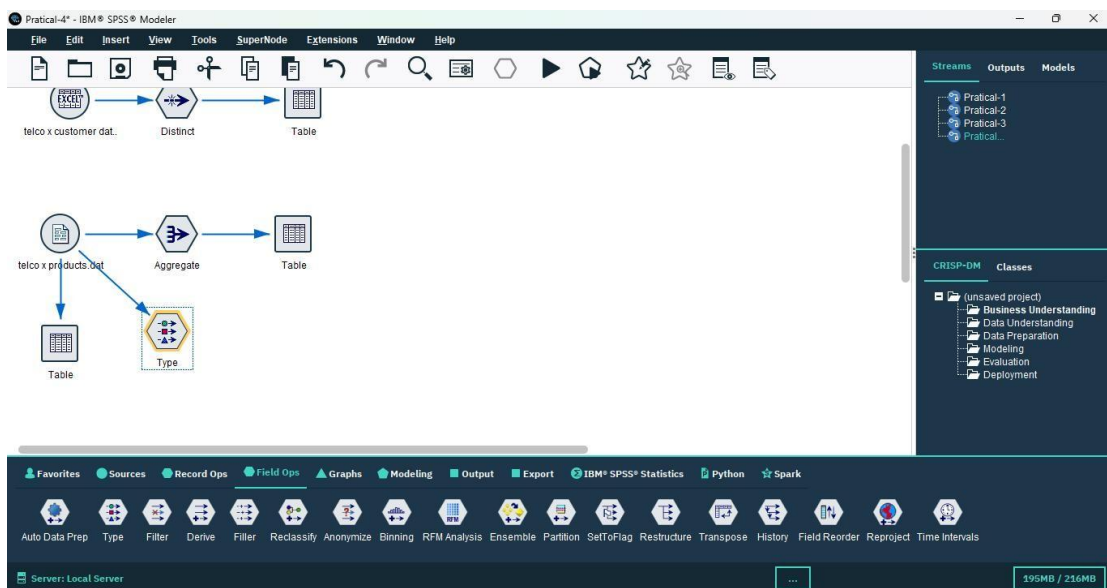- Apply **Sum** and **Mean** to calculate total and average revenue.

Step 6:

- Attach a **Table Node** to the Aggregate Node.
- Run it to view summarized output.
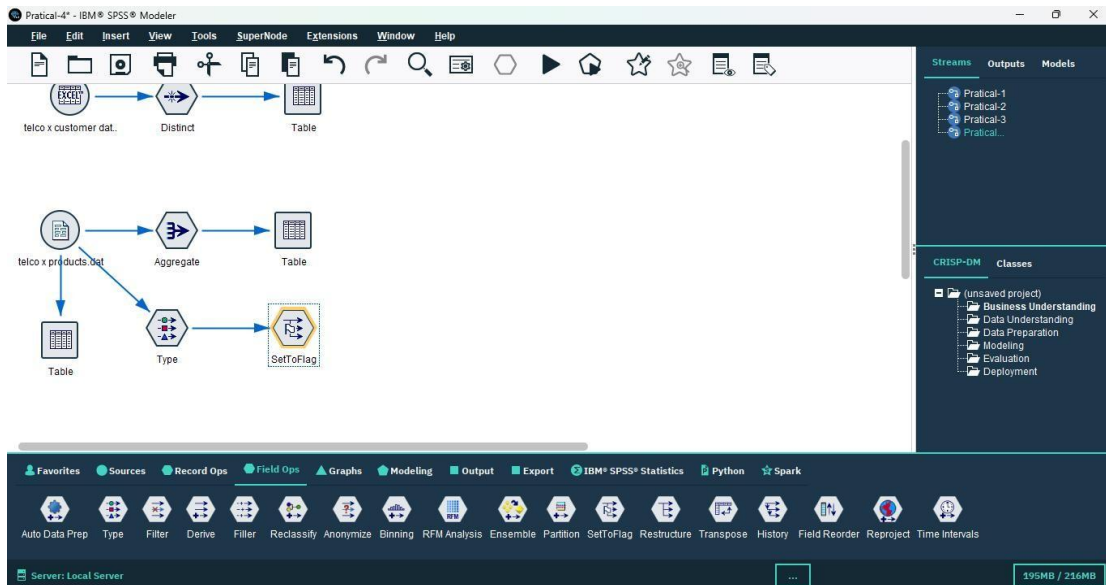- Confirm aggregated revenue values.

Step 7:

- Add a **Type Node** from Field Ops.
- Connect it to the previous node.
- It helps assign **correct field types**.

Step 8:

- Add a **Set to Flag Node** from Field Ops.
- Connect it to the Type Node.
- Convert product values into **Yes/No (0/1)** flags.

Step 9:

- Connect a final **Table Node** from Output category.
- Run the stream to view final data.
- Check that it shows **unique, aggregated, and flagged results**.