BABU BANARASI DAS UNIVERSITY
SCHOOL OF COMPUTER APPLICATION

Case Study

On

Identifying Relationships Affecting Customer Churn in the Telecom Industry

SUBMITTED TO:

Mr. Ayushman Bhadauria

SUBMITTED BY:

NAME: ANIKET KUMAR

CLASS: BCADS8-32

ROLL NO. :1230258076

Identifying Relationships Affecting Customer Churn in the Telecom Industry

Definition:
The agenda of this project is to identify key factors influencing customer churn in the telecom industry by analyzing relationships among variables like handset type, dropped calls, and number of gadgets. It defines how customer behavior and service quality impact retention and satisfaction.

Outcome:

1. This stream gives the outcome that

2. You will know **which customer behaviours or attributes** influence churn.

3. The telecom company can use this knowledge to **reduce churn** and improve **customer satisfaction**.

Required Tool:

IBM SPSS Modeler

Working:

The stream analyzes telecom customer data to identify relationships between variables such as handset type, number of gadgets, and dropped calls with customer churn. It helps determine which factors most influence customers' decisions to leave the network.

Steps to perform tasks:

Step 1: Import the dataset (telcoxdata.txt)

- Open IBM SPSS Modeler and create a new stream.

- From the Sources palette, drag the Var. File node (or Text File node) into the canvas.

- Browse and select the file telco1xdata.txt as your input data source.

- Connect it to a Table node and run it to preview the dataset.

- Verify that all variables (like handset, churn, dropped_calls, number_of_gadgets, etc.) are correctly imported.
This ensures the data is ready for further processing.
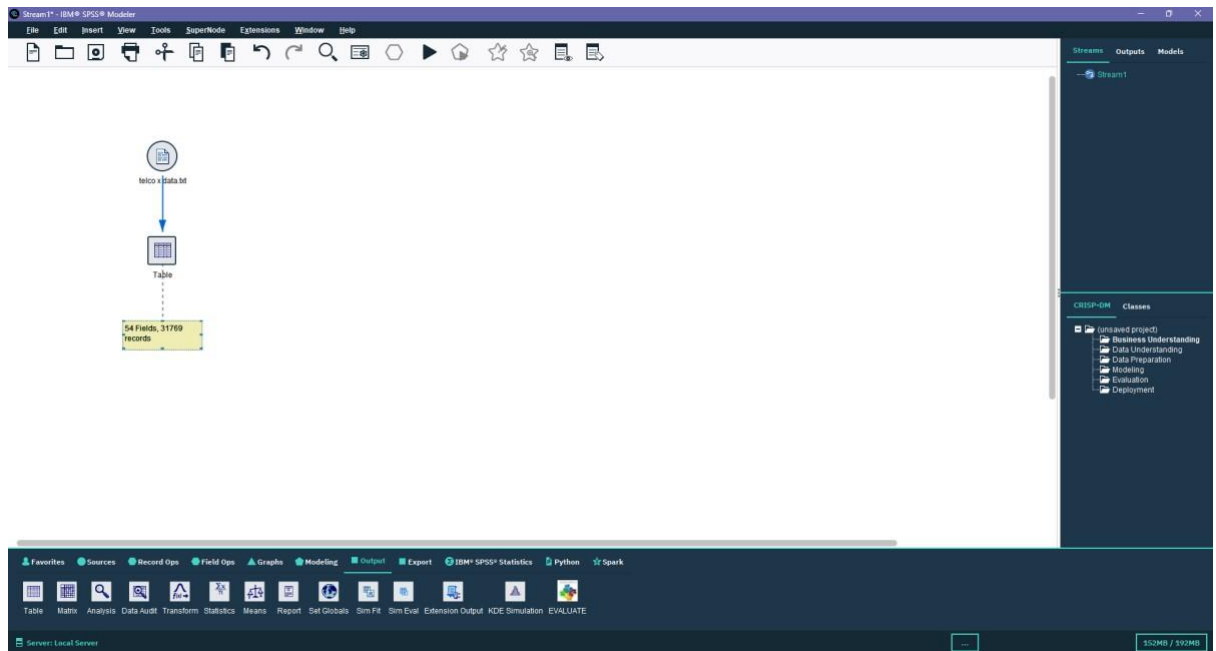
Step 2: Define variable types using the Type node

- From the Field Ops palette, drag and drop the Type node into the canvas and connect it to the input node.

- Open the Type node and define each field properly:

Assign appropriate measurement levels — *Nominal* for categorical data (like handset types), *Continuous* for numerical data (like dropped calls).

- This step helps SPSS Modeler understand the data type and purpose of each variable for accurate analysis
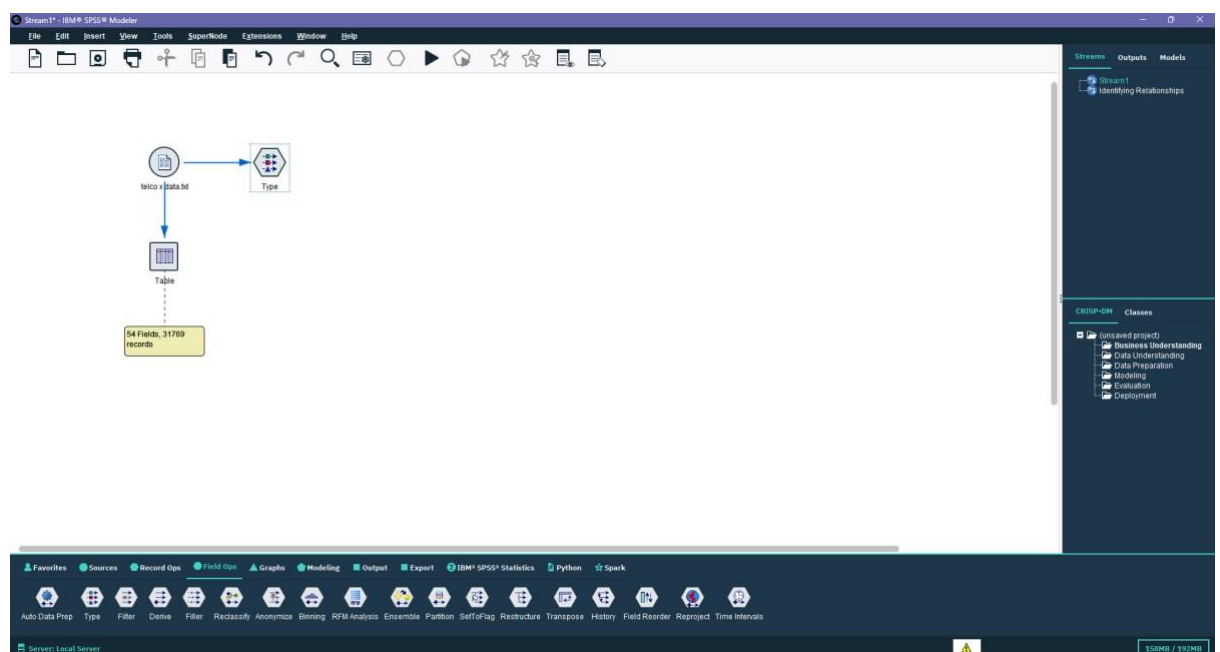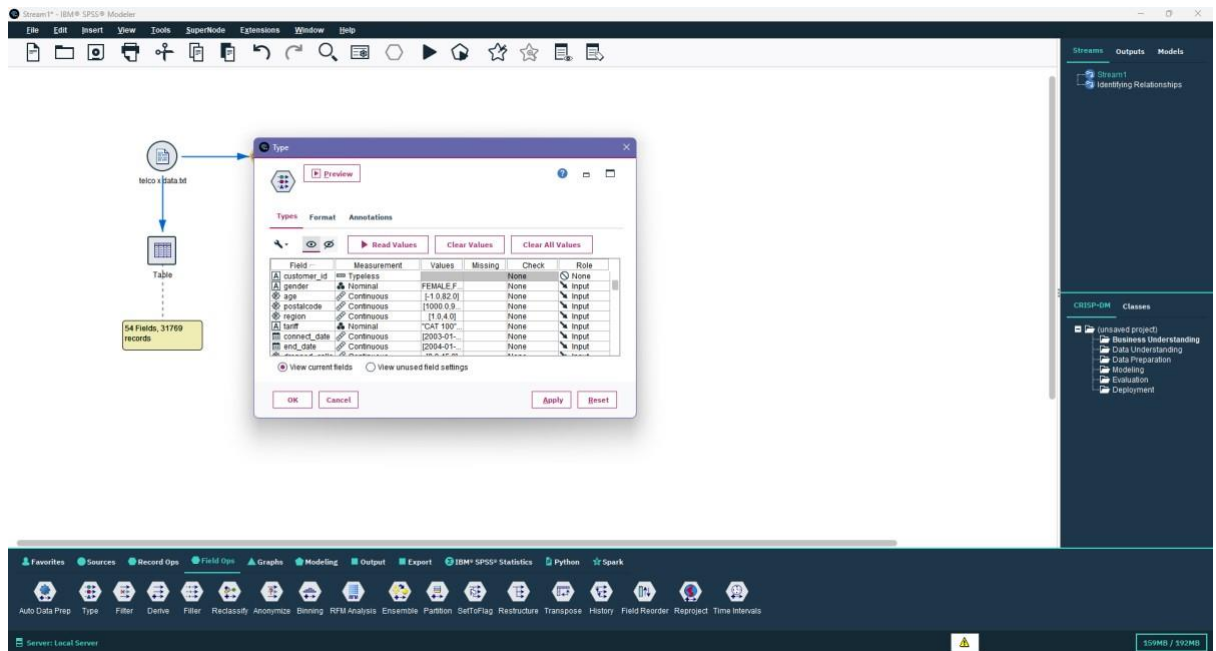
Step 3: Analyze handset and churn using the Matrix node

- From the Output palette, drag the Matrix node to the canvas and connect it to the Type node.

- Double-click the Matrix node and set handset as one variable and churn as the other.

- This cross-tabulation helps you identify the relationship between different handset types and customer churn rates — for example, which handset users are more likely to leave the network.
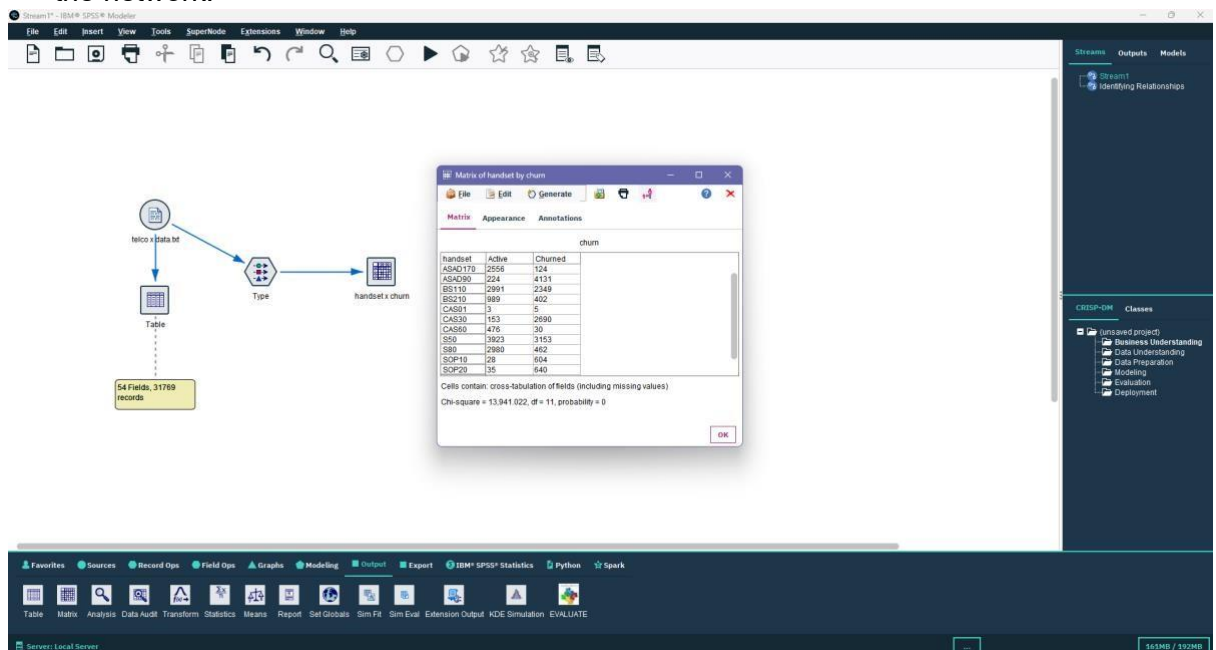
Step 4: Visualize handset distribution using the Distribution node

- From the Graphs palette, drag the Distribution node and connect it to the Type node.
- Configure the node to display the distribution of handset categories.
- This helps you see which handset models are most common among customers and compare them visually.
- You can interpret which handset categories have the highest user base or which ones dominate among churned customers.
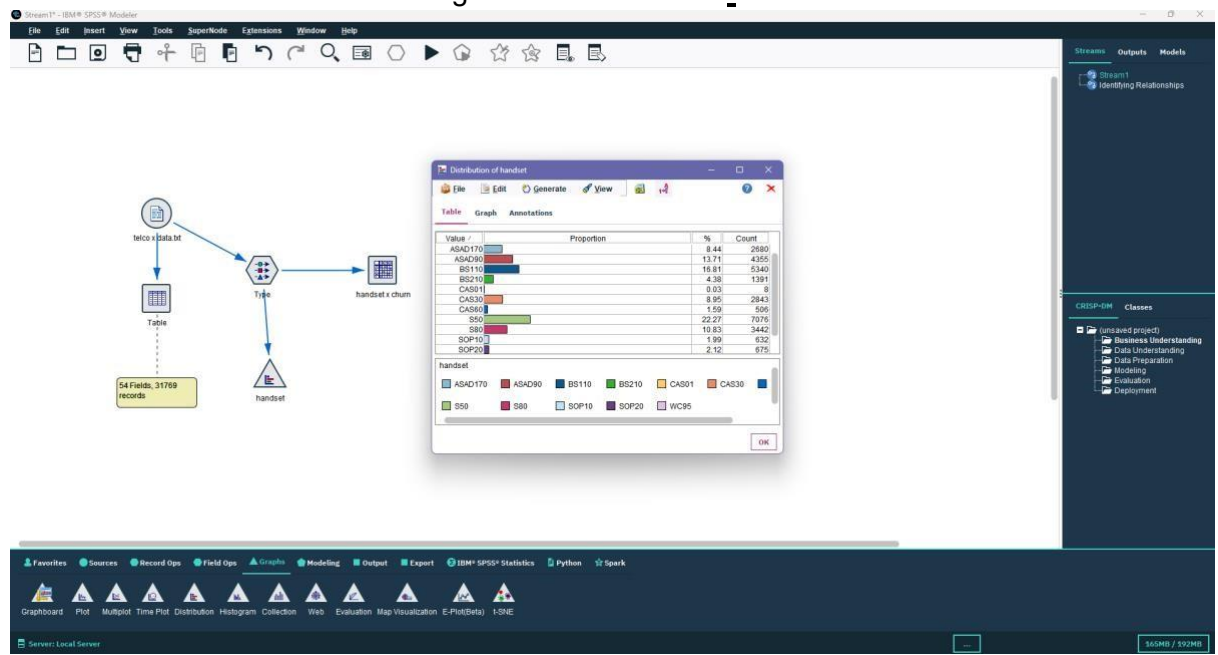
Step 5: Analyze dropped calls with the Means node

- Drag the Means node from the Output palette and connect it to the Type node.
- Choose dropped_calls as the field to analyze and churn as the grouping variable.
- This will calculate and compare the average number of dropped calls for customers who churned versus those who stayed.
- A higher mean value for churned customers may indicate poor network quality influencing churn.
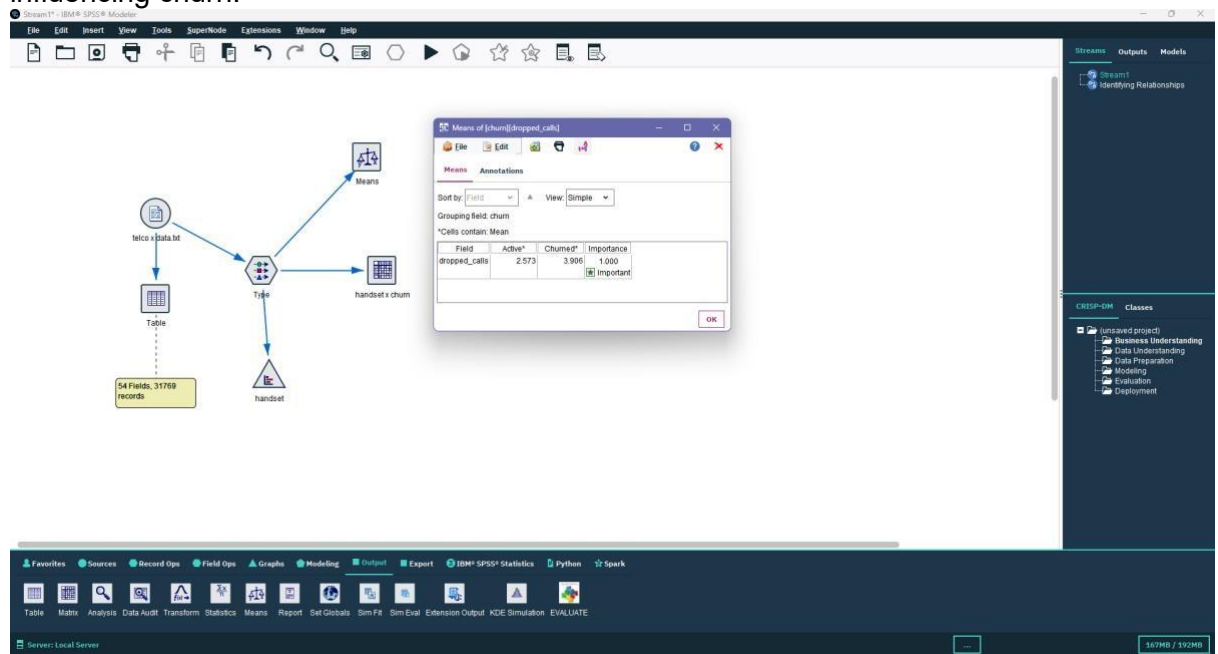
Step 6: Examine churn pattern using Histogram

- From the Graphs palette, drag the Histogram node and connect it to the Type node.
- Select dropped_calls or number_of_gadgets as the field to plot, and group them by churn.
- This visualization shows how churned and non-churned customers are distributed across these metrics.
- It helps in understanding whether higher dropped calls or device usage is associated with churn.
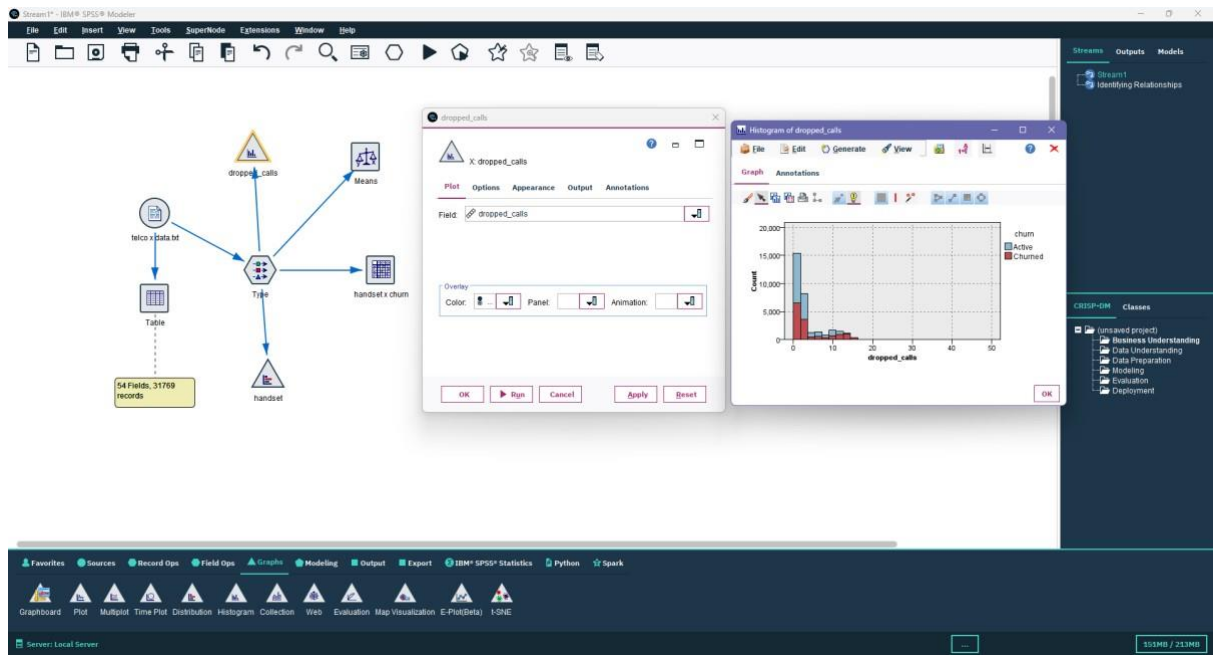
Step 7: Study revenue pattern using Plot node

- Drag the Plot node from the Graphs palette and connect it to the Type node.
- Select variables such as number_of_gadgets and revenue (if available).
- Use this node to visualize trends — for example, how customer spending or number of gadgets correlates with churn status.
- This can reveal patterns like "customers owning multiple gadgets tend to have higher loyalty or lower churn."
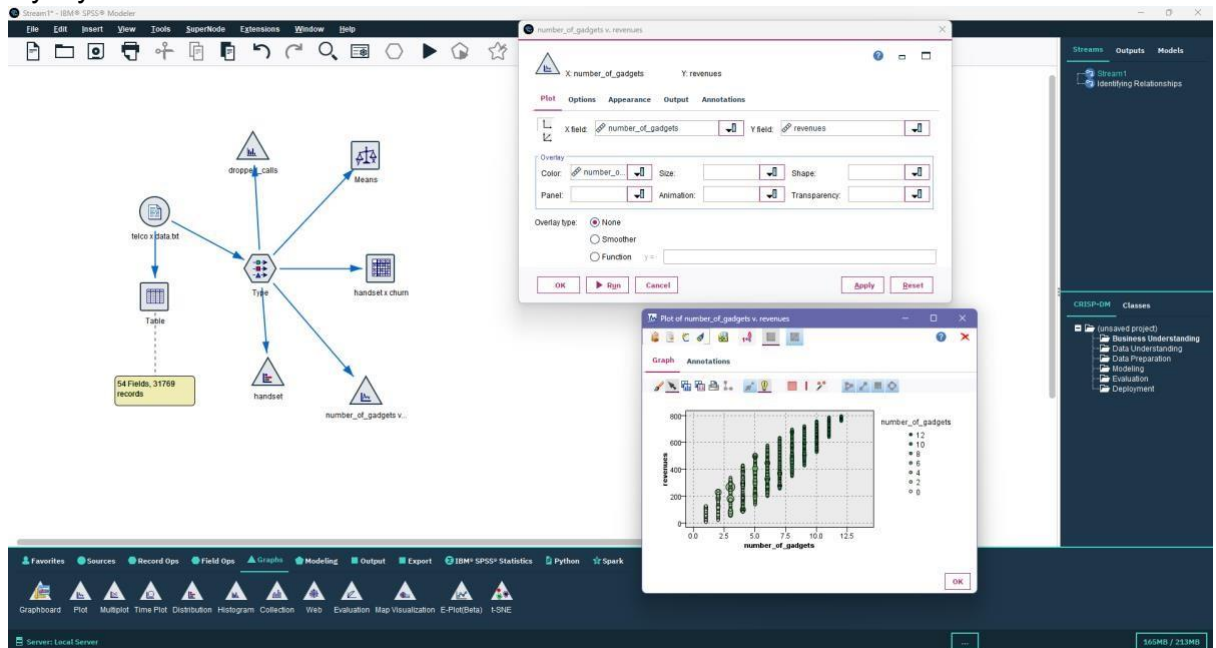
Step 8: Generate descriptive statistics using the Statistics node

- Drag the Statistics node from the Output palette and connect it to the Type node.
- Configure it to display descriptive measures such as:
  - ➢ Mean, Median, Mode
  - ➢ Minimum and Maximum
  - ➢ Standard Deviation and Variance
  - ➢ Count and Range
- These summary statistics help you understand the overall data distribution, variability, and key trends in customer behavior.