Babu Banarasi Das University

****

**Case Study on**

**Customer Churn Prediction**

**SUBMITTED TO: SUBMITTED BY:**

**Mr. Ayushman Bhardwaj Sir Name: Anand Kumar Gupta**

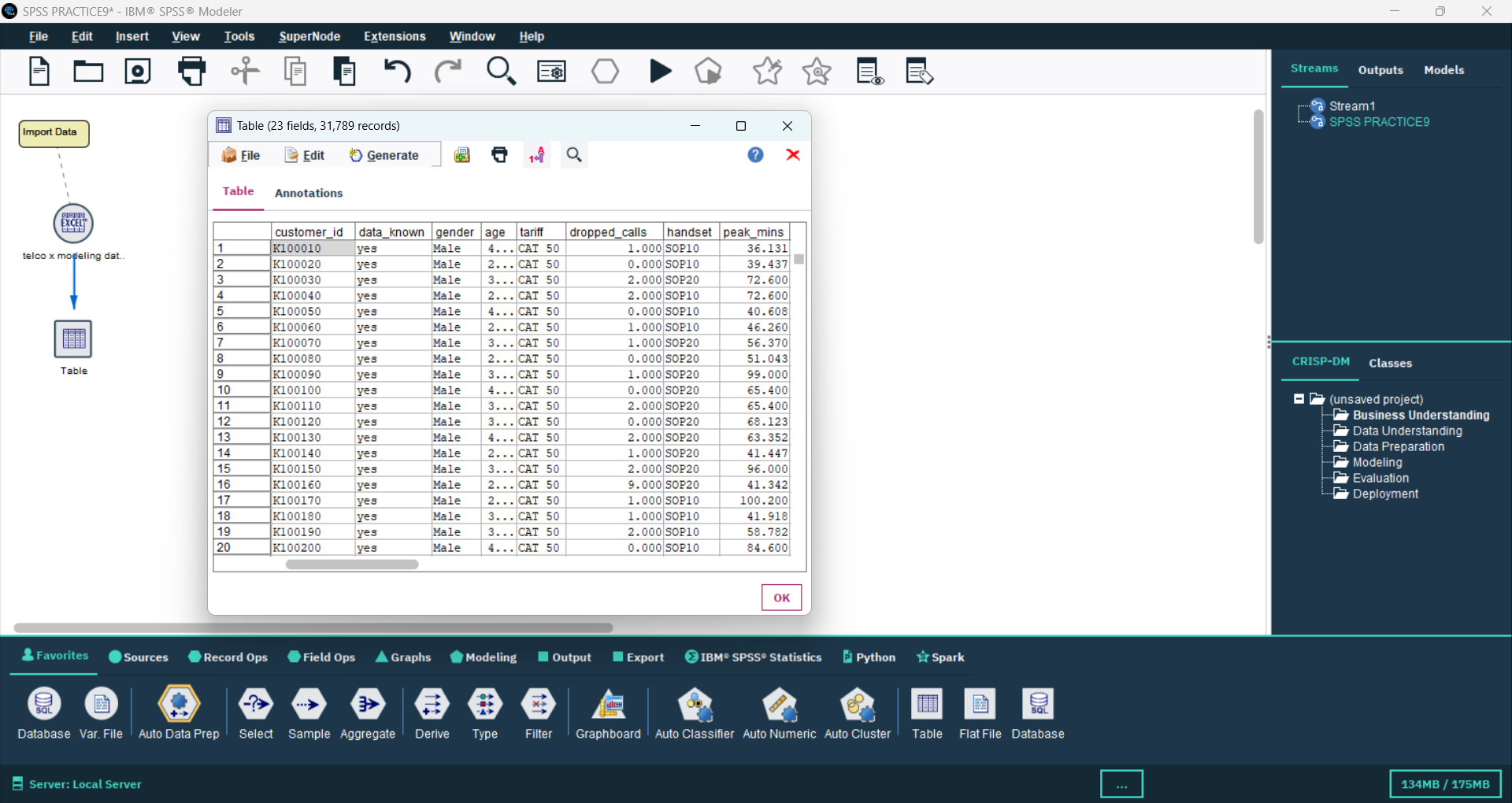
# Customer Churn Prediction – Practical Steps (SPSS Modeler)

This document describes all the steps followed in the SPSS Modeler stream for Customer Churn Prediction using the dataset 'telco x modeling data'. You can paste the corresponding screenshots below each step for your practical submission.

## 1. Excel Source Node (telco x modeling data)

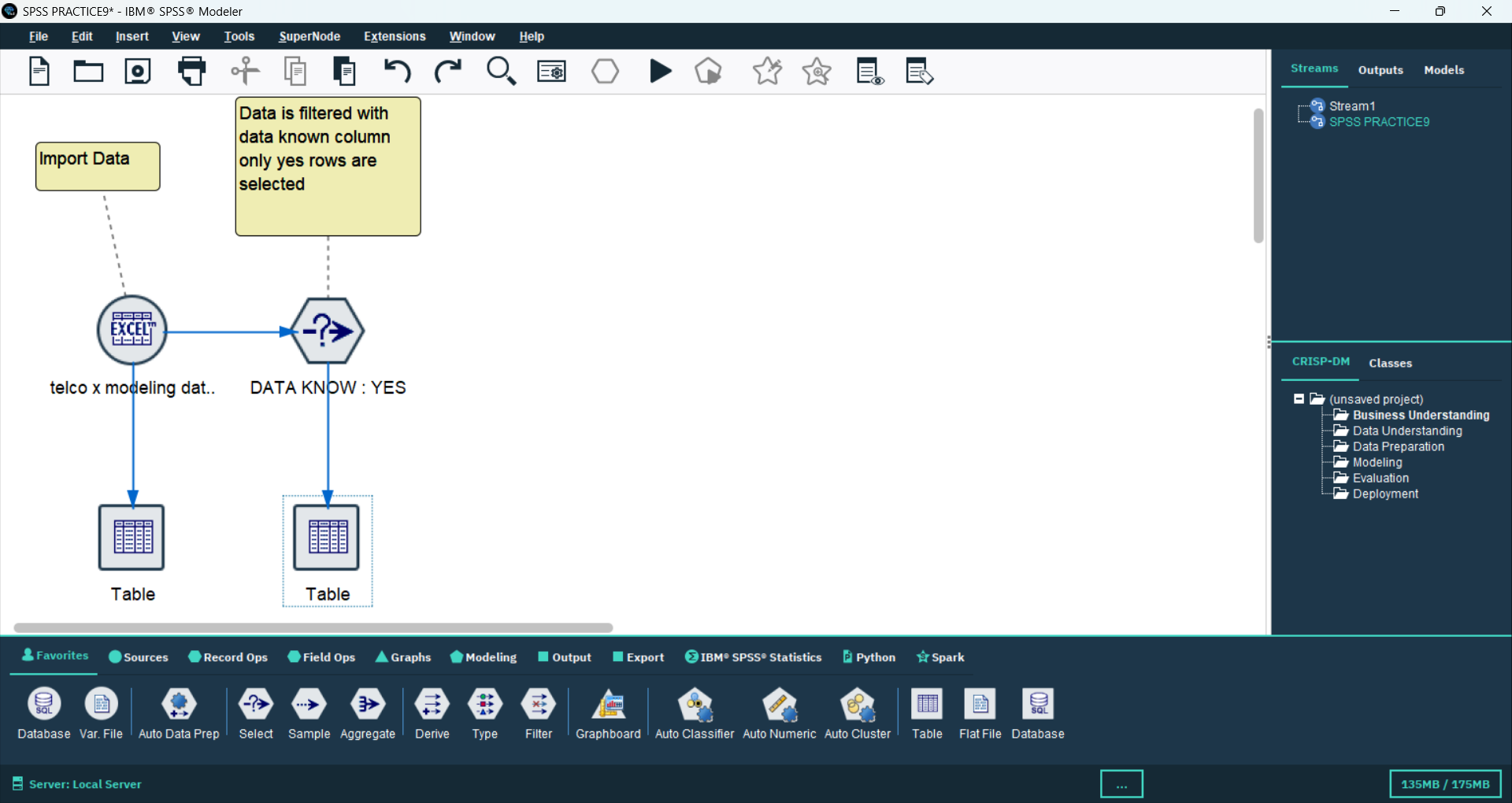
Imported the main dataset 'telco x modeling data' from an Excel file. This dataset contains customer details, service usage information, and churn status.

📸 Screenshot:



## 2. Filter Node (DATA KNOW: YES)

Filtered the dataset to include only rows where 'DATA KNOW' is 'YES' to ensure that only complete and valid records are analyzed.

📸 Screenshot:

## 3. Table Node (Preview 1)

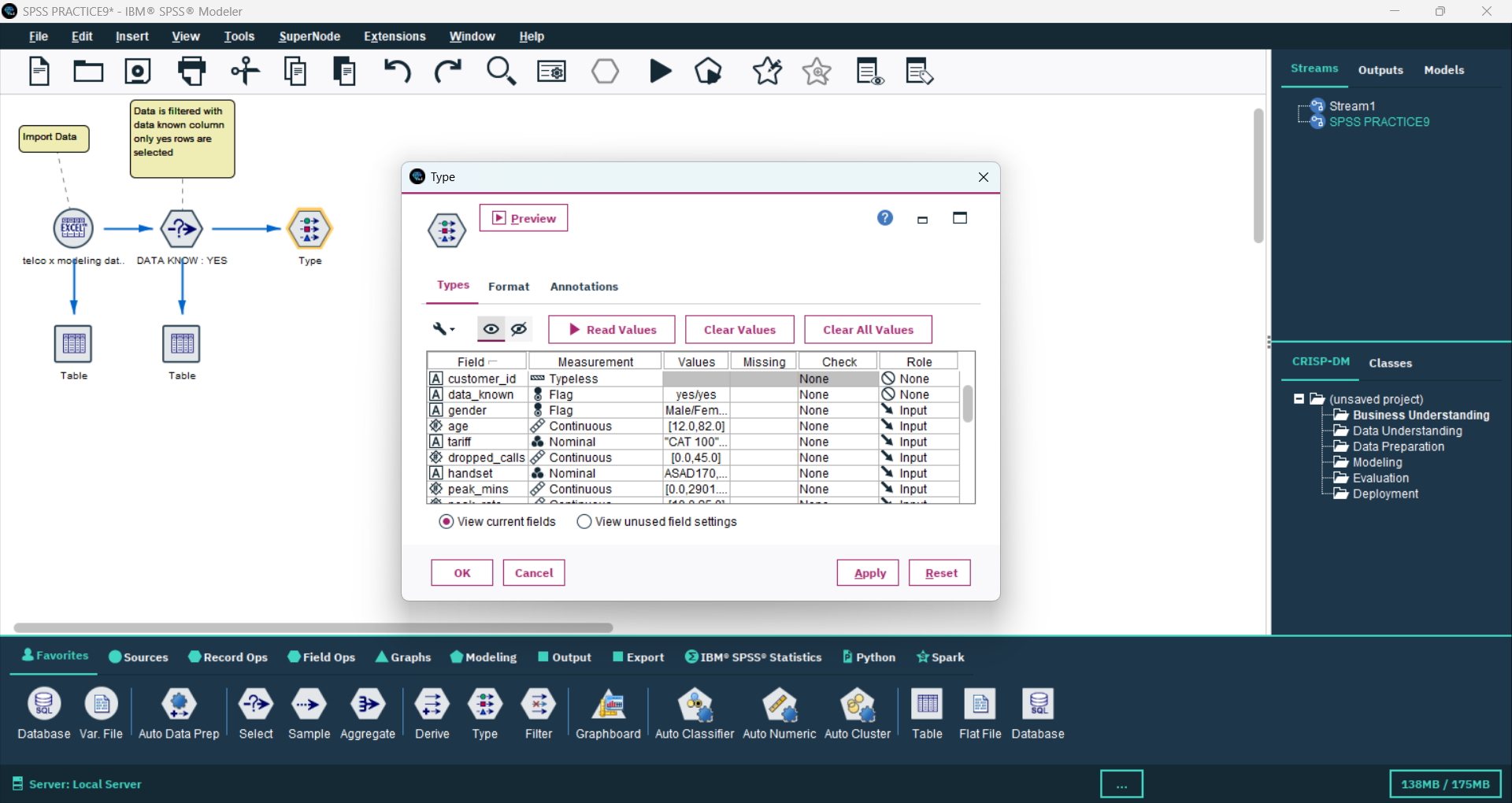
Used a Table node to preview the filtered data and confirm that the correct records were selected.

## 4. Table Node (Preview 2)

Another Table node used to view the dataset after filtering to verify the number of rows and columns retained.

## 5. Type Node

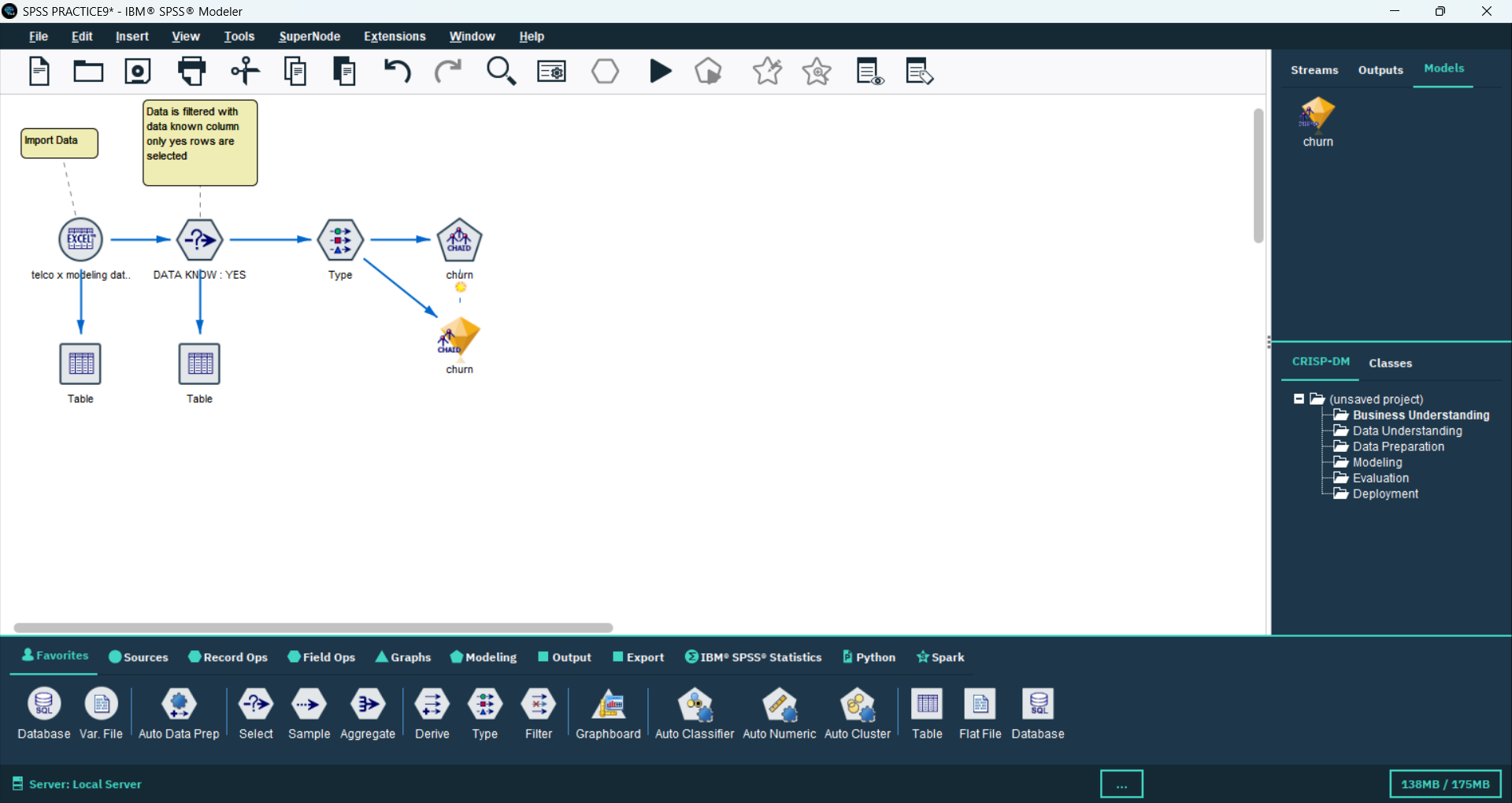
Defined variable roles and measurement levels. Set the target field as 'churn' and assigned input roles to the predictor variables.

📸 Screenshot:

## 6. CHAID Model Node (churn)

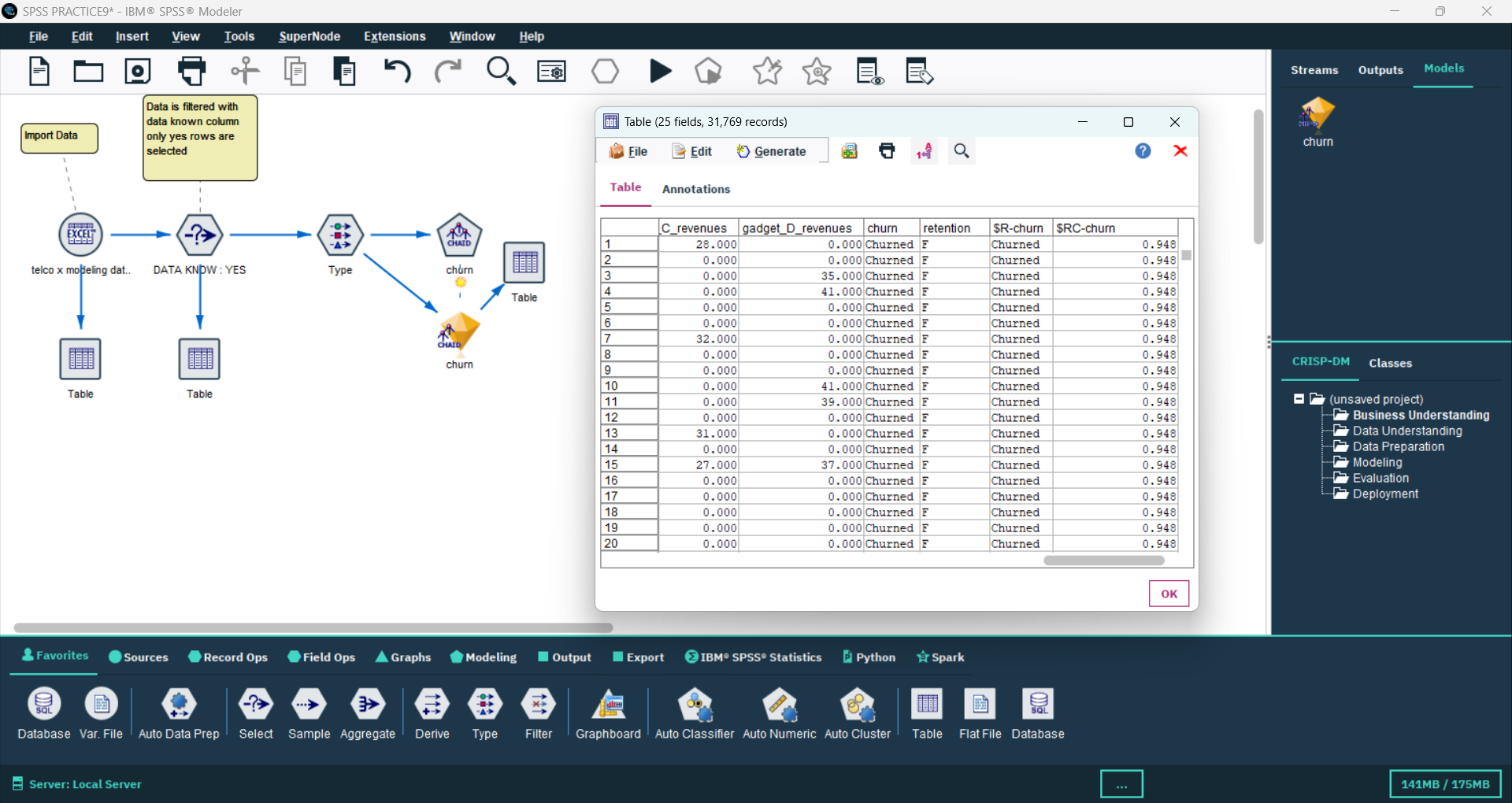
Built a CHAID Decision Tree model to predict customer churn. The target variable is 'churn', and other customer-related features are inputs.

📸 Screenshot:



## 7. Table Node (Model Output)

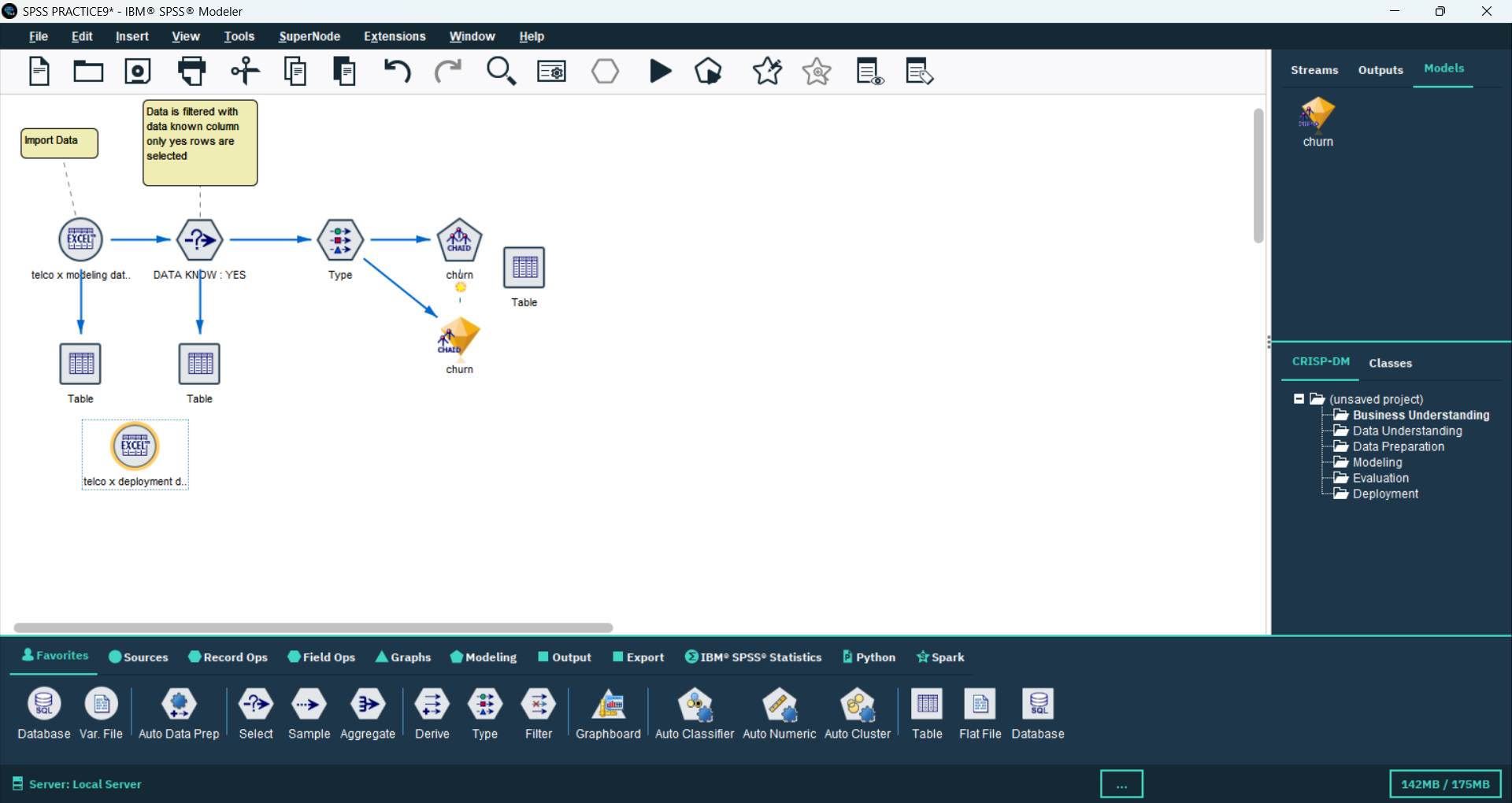
Displayed the results of the CHAID model, including predicted churn values and probabilities.

📸 Screenshot:

## 8. Excel Source Node (telco x deployment data)

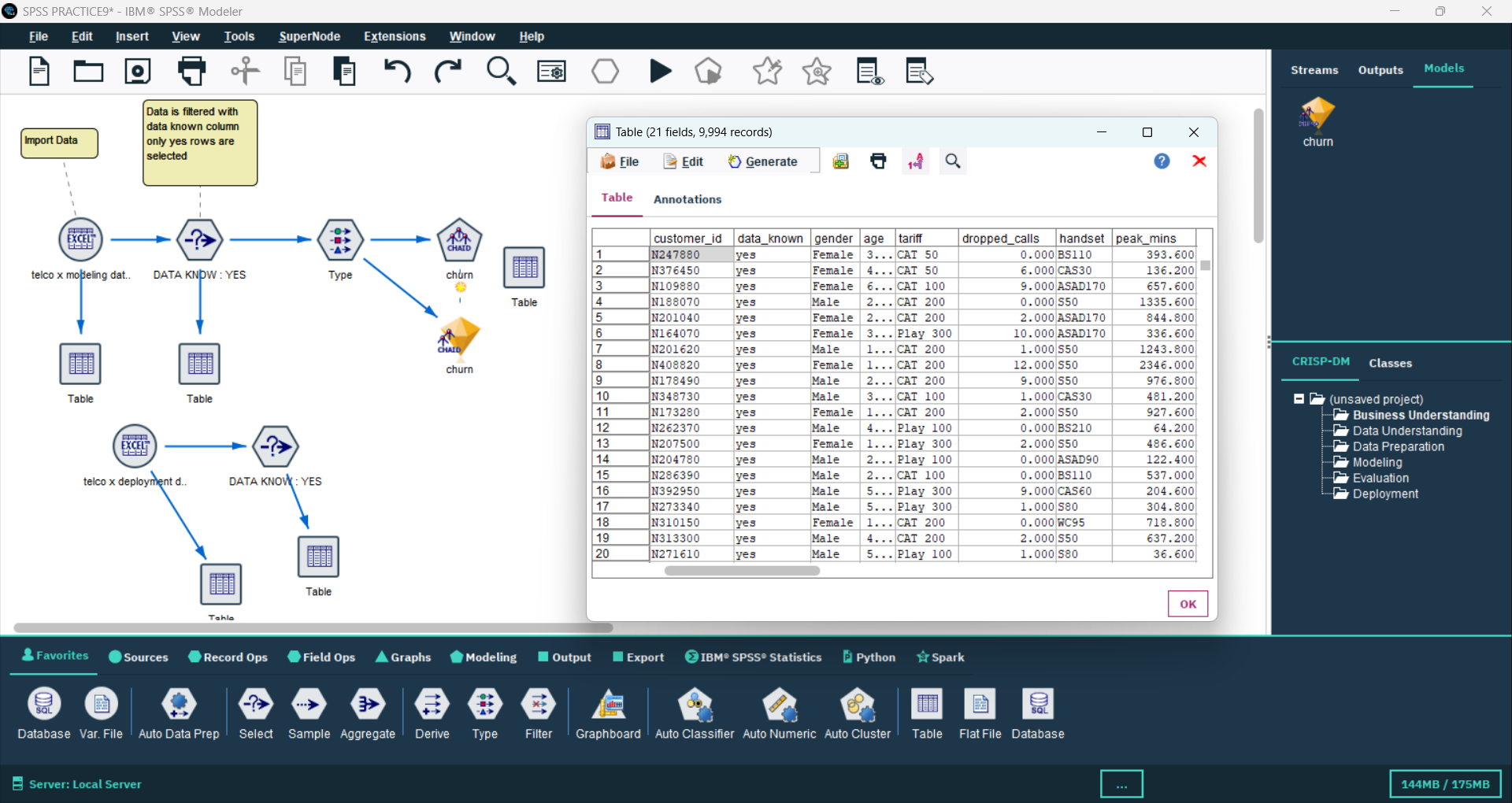
Imported another dataset 'telco x deployment data' used for applying (scoring) the trained churn model on unseen data.

📸 Screenshot:



## 9. Filter Node (DATA KNOW: YES - Deployment Data)

Filtered deployment data to include only rows where 'DATA KNOW' is 'YES' to maintain data consistency before applying the model.

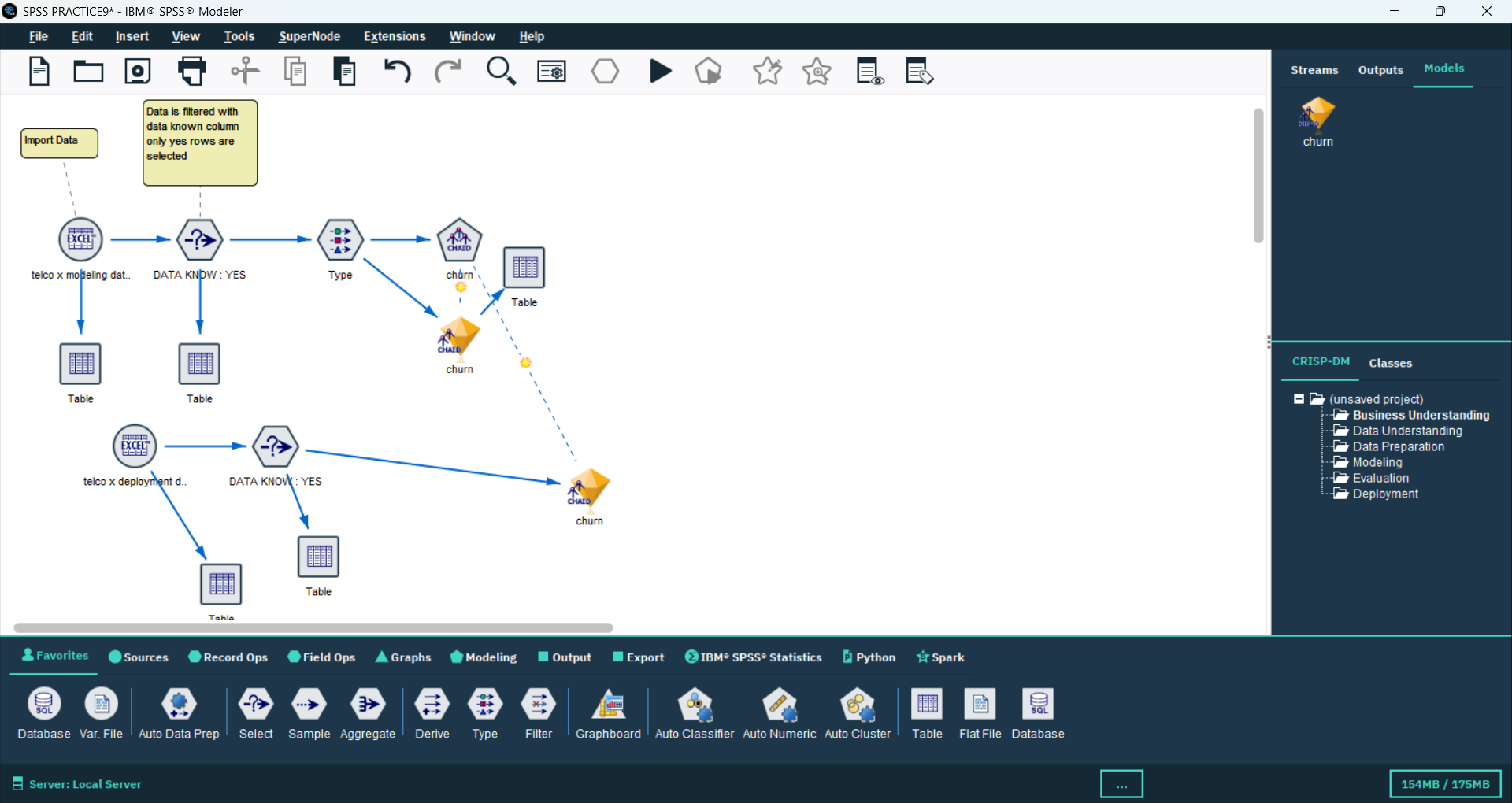
📸 Screenshot:

## 10. Table Node (Deployment Data Preview)

Viewed deployment dataset after filtering to confirm correct loading and structure.

## 11. CHAID Model Applied (churn)

Applied the previously trained CHAID churn model to the deployment dataset to predict churn probabilities for each customer.

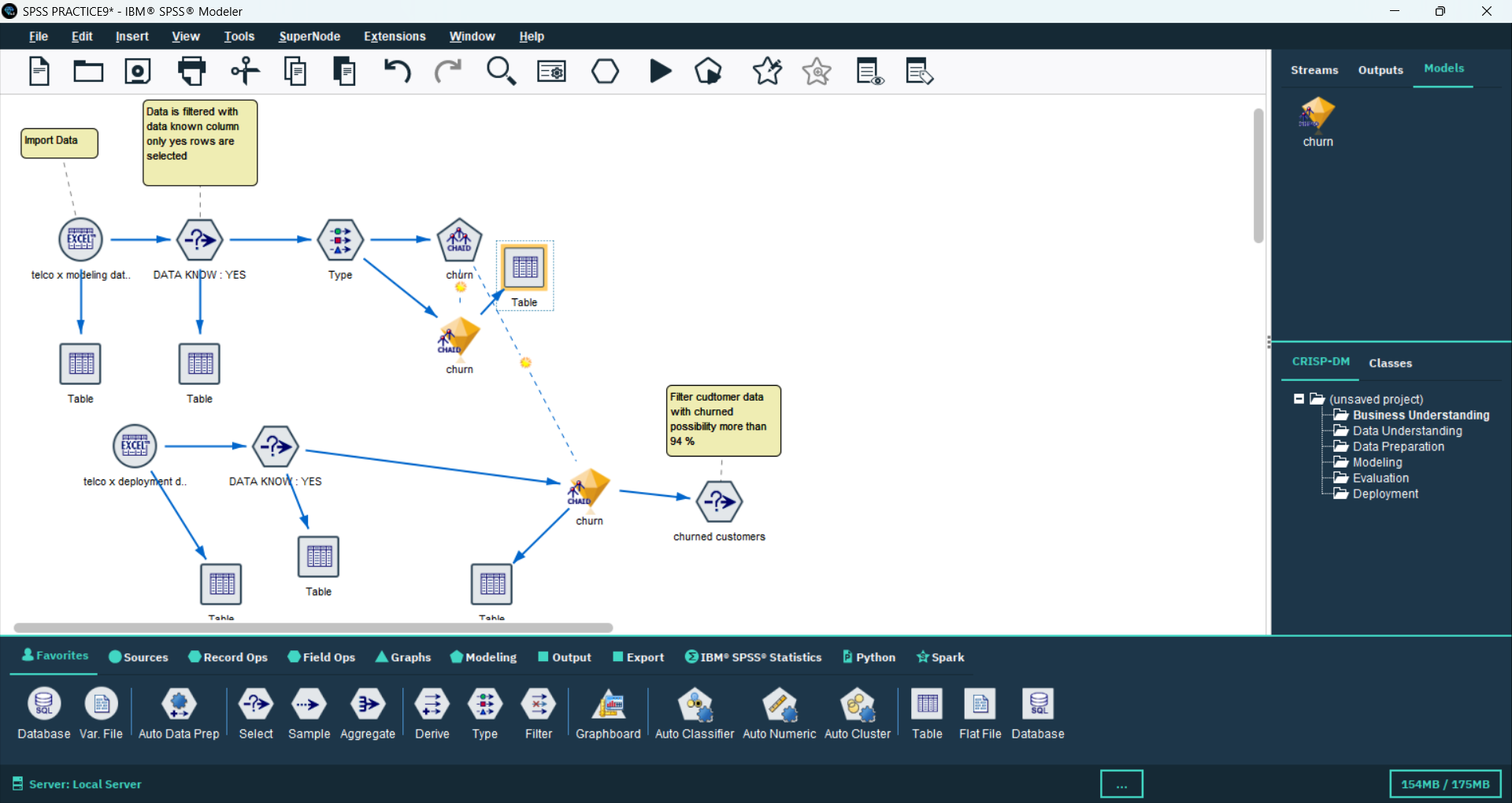
📸 Screenshot:

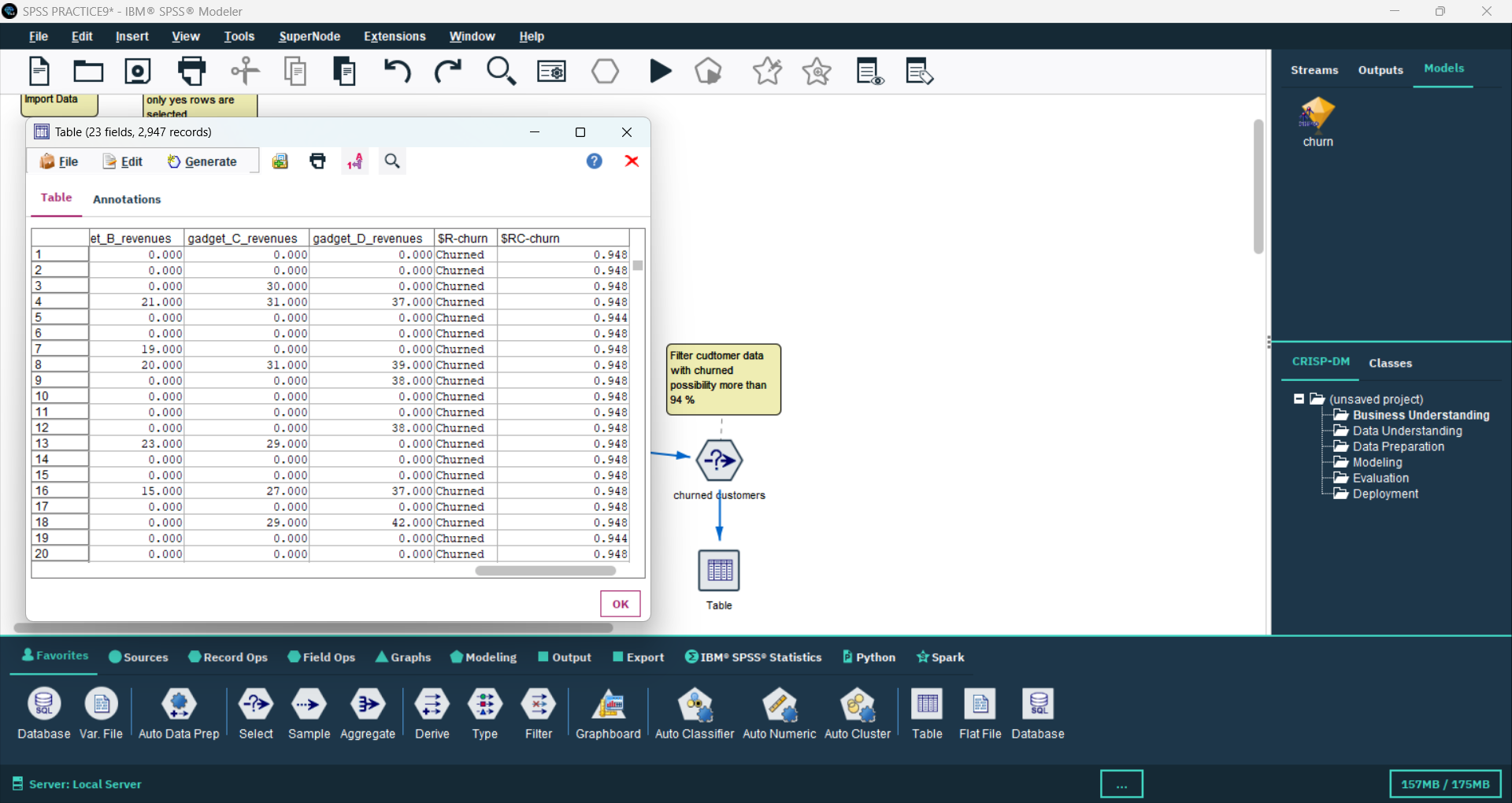


## 12. Filter Node (churned customers > 94%)

Filtered customer data to select only customers having churn probability greater than 94%, identifying the most likely churners.

📸 Screenshot:

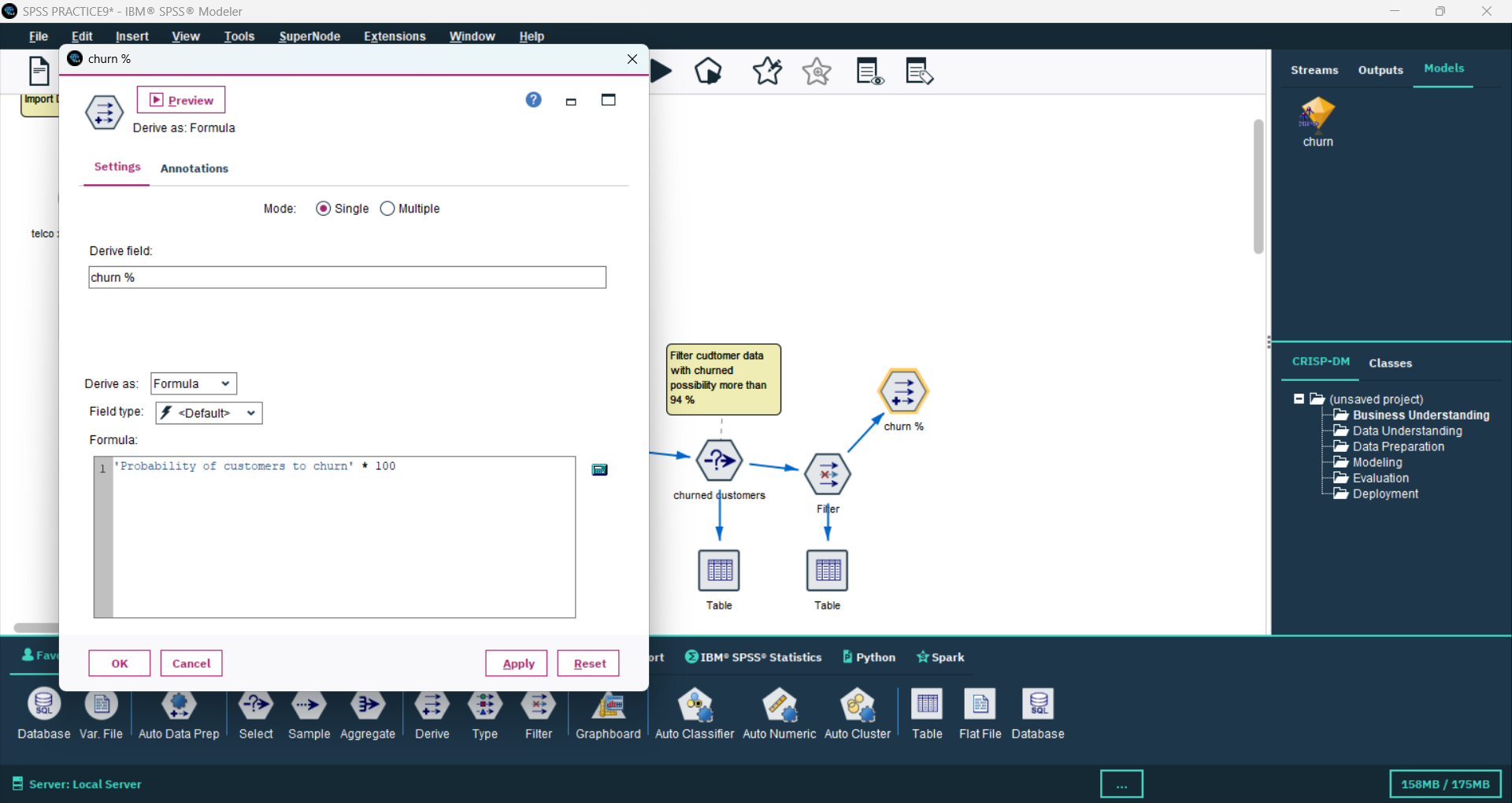




## 13. Filter Node (churn %)

Filtered and organized the output based on churn percentage for analysis and reporting.

📸 Screenshot:



## 14. Filter Node (churned customers)

Final filter applied to generate a refined list of churned customers ready for reporting.

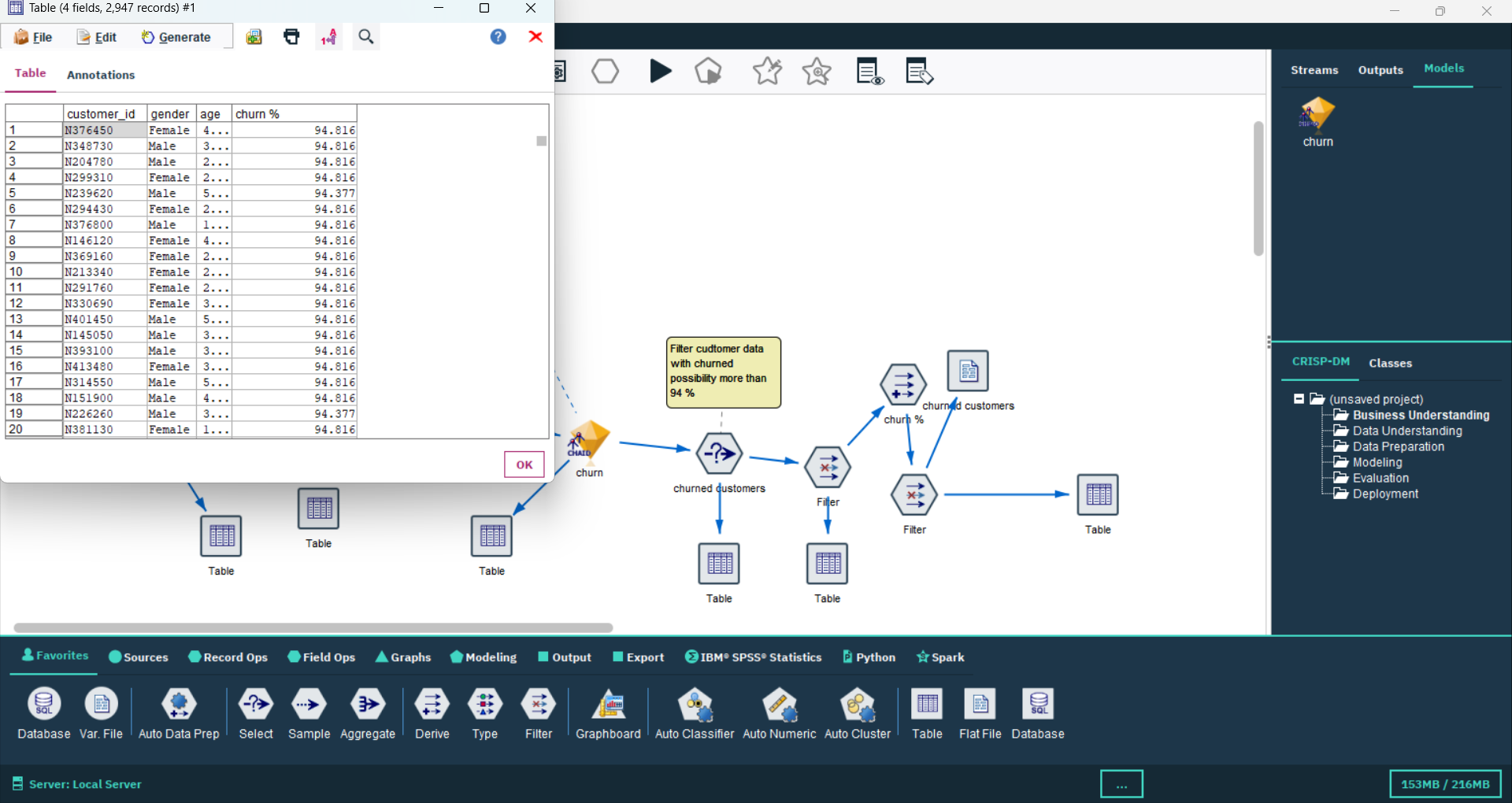
## 📸 Screenshot:

## 

## 15. Table Node (Final Output)

Displayed the final list of churned customers with churn probability and relevant customer information for business insight generation.

📸 Screenshot:



## Conclusion

The SPSS Modeler stream successfully implemented a complete Customer Churn Prediction pipeline using CHAID Decision Tree. It covered data import, cleaning, preparation, modeling, evaluation, and filtering of high churn probability customers. This workflow helps identify customers most at risk of leaving, allowing targeted retention strategies.