Babu Banarasi Das University

School Of Computer Application

Semester – 5

Case Study

On

Predicting Customer Churn in a Telecommunications Company

SUBMITTED TO:

Mr. Vikas Sir

SUBMITTED BY:

Name: SURAJ KUMAR YADAV

Roll NO:1230258438

Agenda/Definition: You work for a telecommunications firm and have to combine a number of datasets into a single dataset for analysis**.**

Outcomes/Learning: importance of nodes such as Churn, Type, Flat Files, Filter etc.

Required Tool: IBM SPSS Modular

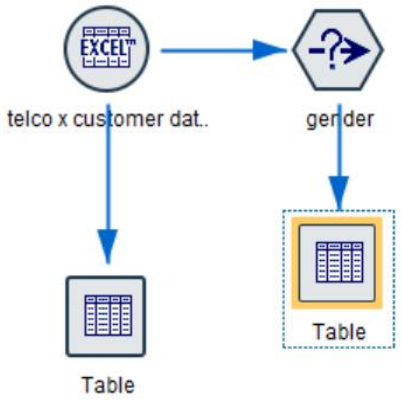
Working:

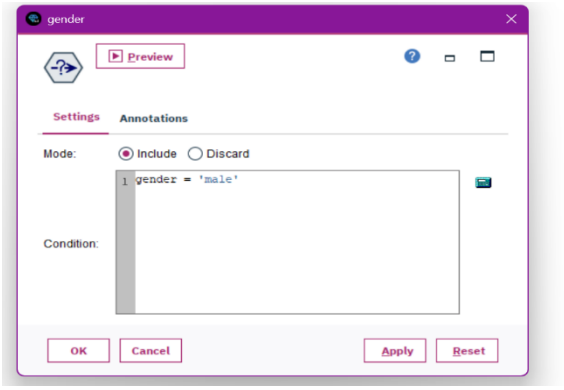
Step 1: Open IBM SPSS Modeler 18.6. We need to insert dataset now. So, we have a SOURCES Category from there we can choose the format of our dataset. Here. we will choose the Two STATISTICS NODE as the format of our dataset with the excel format

Step 2: Load the Excel data

Add an Excel (or File) source node and point it to your telco customer Excel file.

Run the node or open it to verify rows & column names.





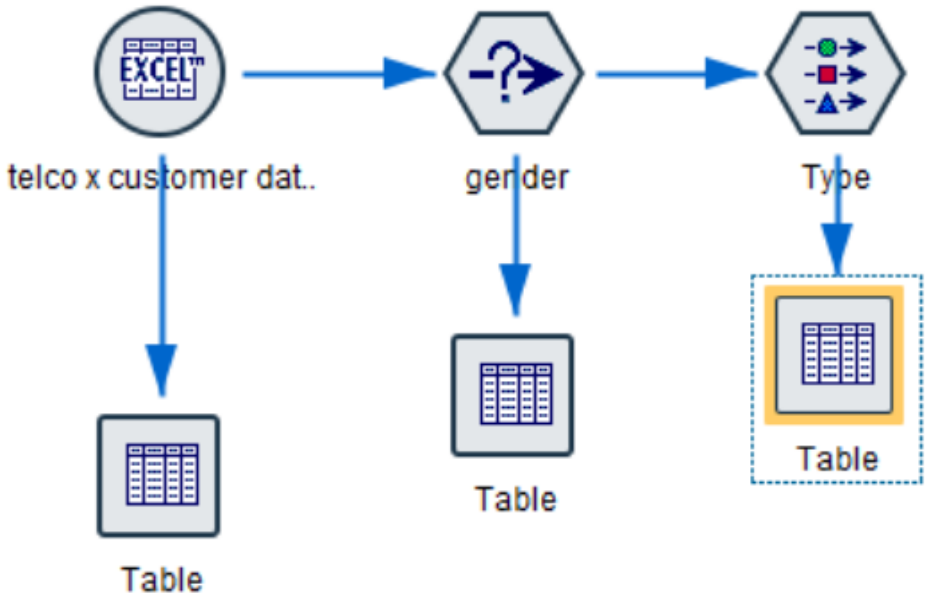
Step 3:

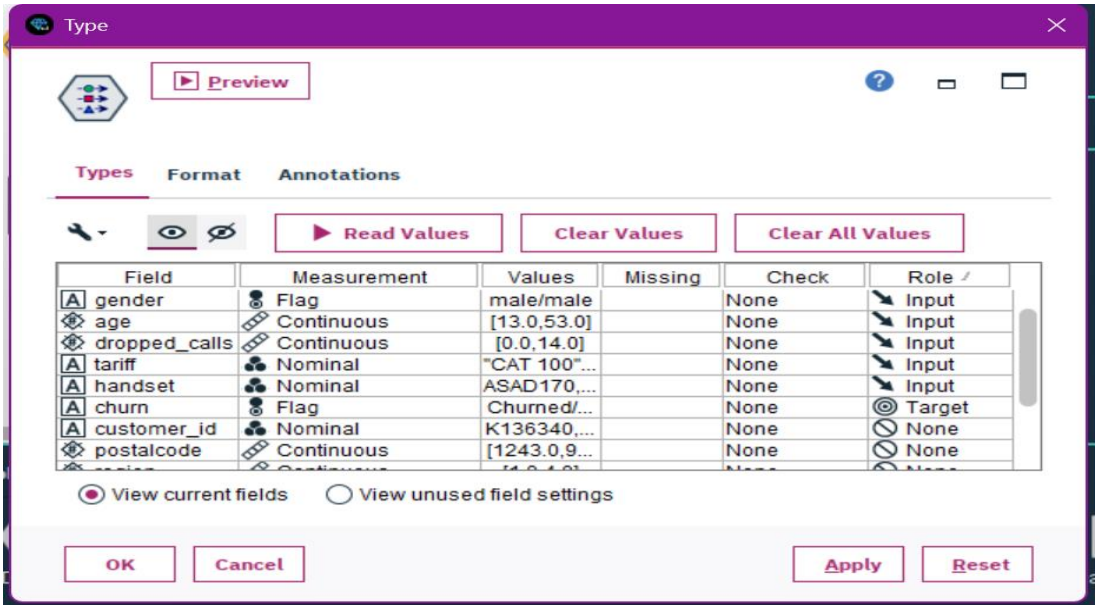
- **Add a Select after Excel.**
- **Set Mode = *Include* and Condition = gender = 'male' (or adjust as needed). This creates a branch with only male records.**
- **Run or preview the Select node to confirm filtered rows.**

Step 4:

A. **Add a Type node to control measurement levels for each field.**
B. **Click Read Values to let Modeler detect values, then set proper measurement for each field:**

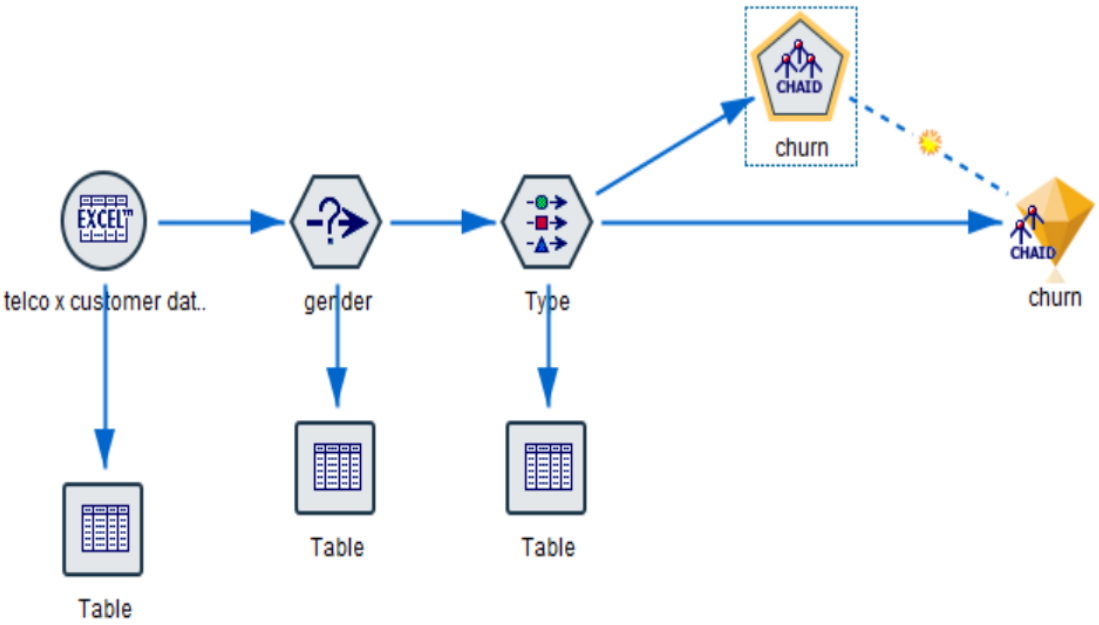
- **gender → Flag (or Nominal / Categorical)**
- **age → Continuous**
- **dropped_calls → Continuous**
- **tariff and handset → Nominal**
- **churn → Flag (target variable)**
- **Apply and run the Type node**

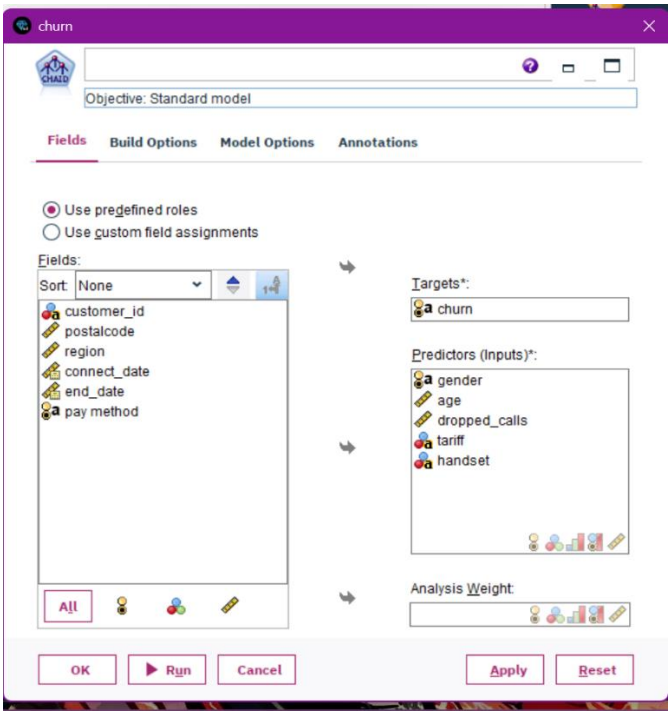




Step 5: Add the CHAID modeling node and connect it to the processed data output from the Type node. Double-click the CHAID node to open settings.

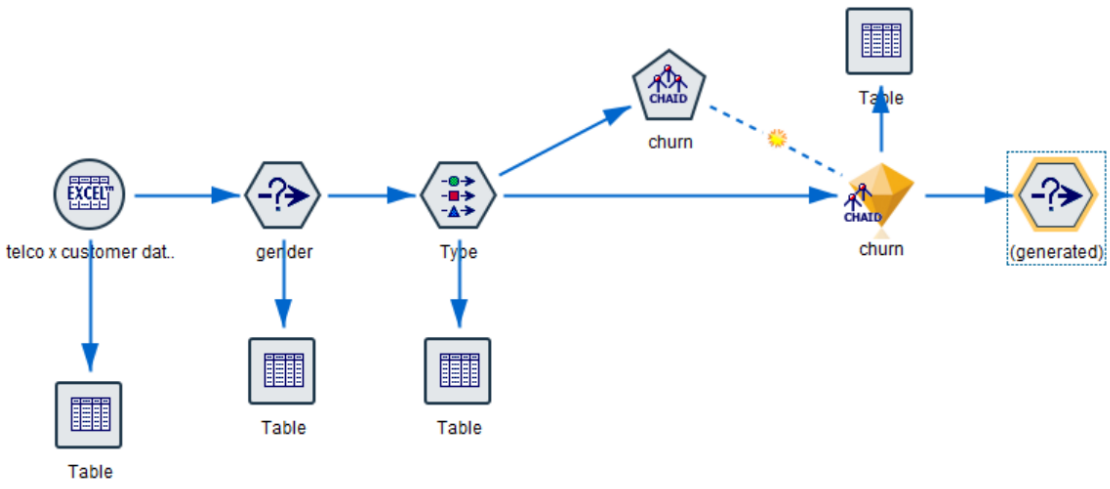
Confirm Target is churn. Move relevant Predictors (Inputs) such as gender, age, dropped_calls, tariff, handset into the predictor list.





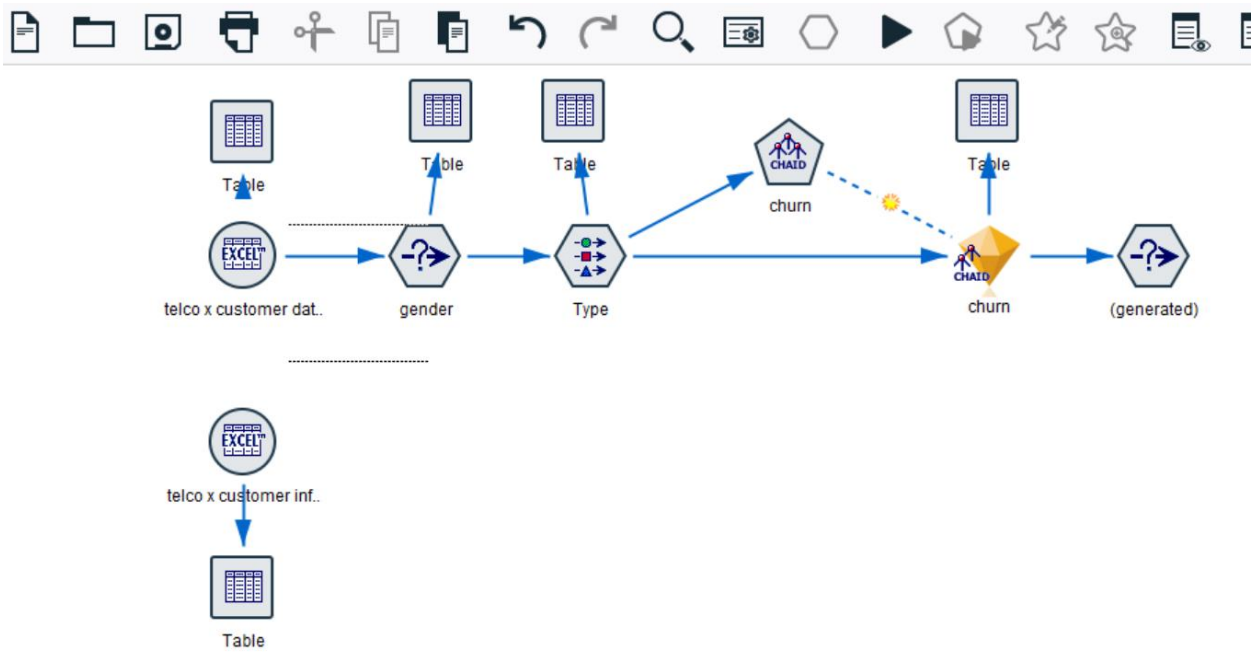
Step 6: In the CHAID node, optionally adjust Build Options (e.g., significance level, min node size) depending on your data size.

Click Run to build the model.



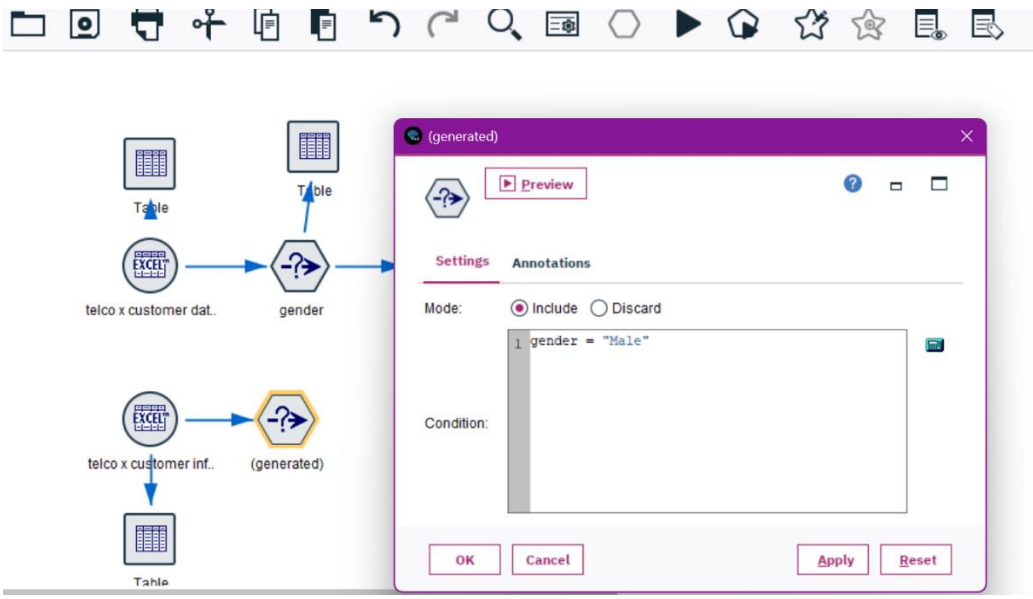
Step 7: After running, open the model node to view results, rules, and tree structure.

Connect the model output to a **Table** or output node to generate scored output for each row (predicted churn flag/probabilities).



Step 8: In this model ,we can take another data telco x customer info.xlse.

Now we can see the data through table.



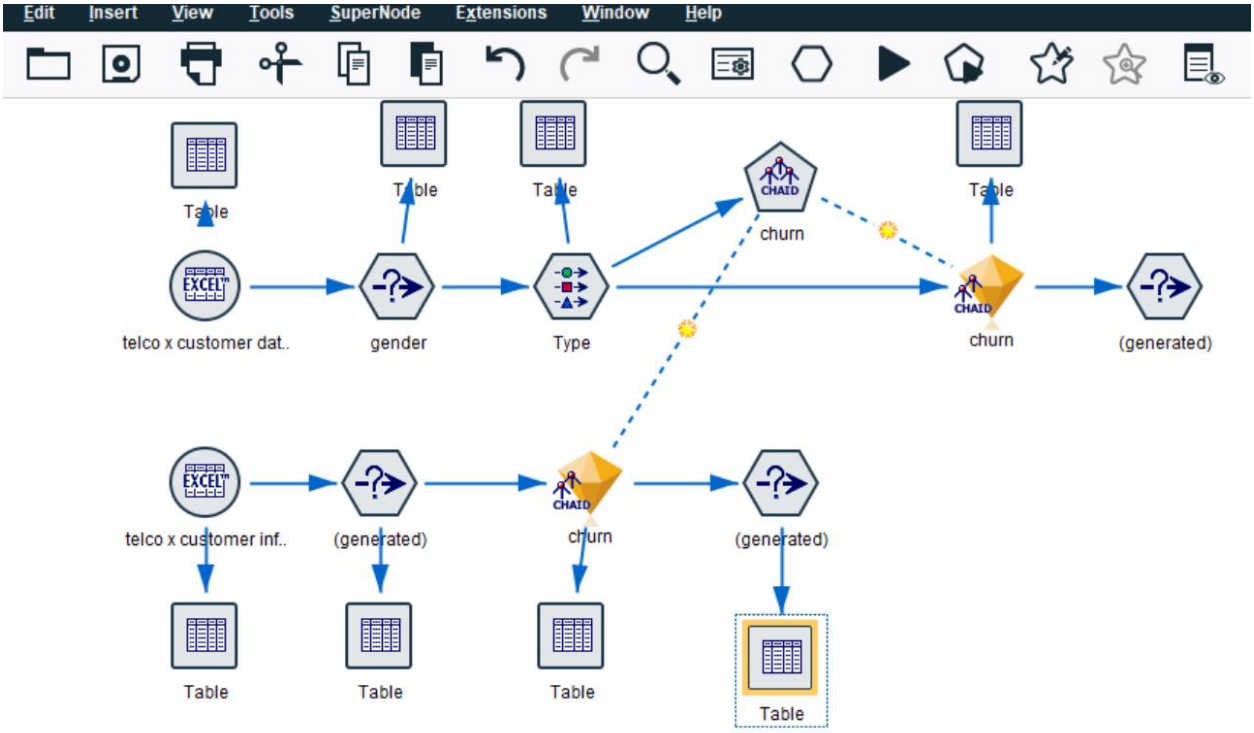
Step 9: Import a new Excel file (**telco x customer info.**) containing fresh customer records that do not yet have a churn flag.

Connect this dataset to the **Generated model node** (the gold diamond icon) created from the CHAID model.

This connection applies the trained churn prediction model to the new dataset — predicting churn for each customer based on the model's learned rules.

Add a **Table node** after the generated model to preview the results. You'll see the predicted churn flag and probability for each new customer record.

Optionally, export the scored dataset using a **File Output** node to Excel or CSV for reporting and business use.



Step 10: Connect the **Filter** node to the **Generated** node (the one created from applying the model to *telco x customer info*).

Double-click the **Filter** node.

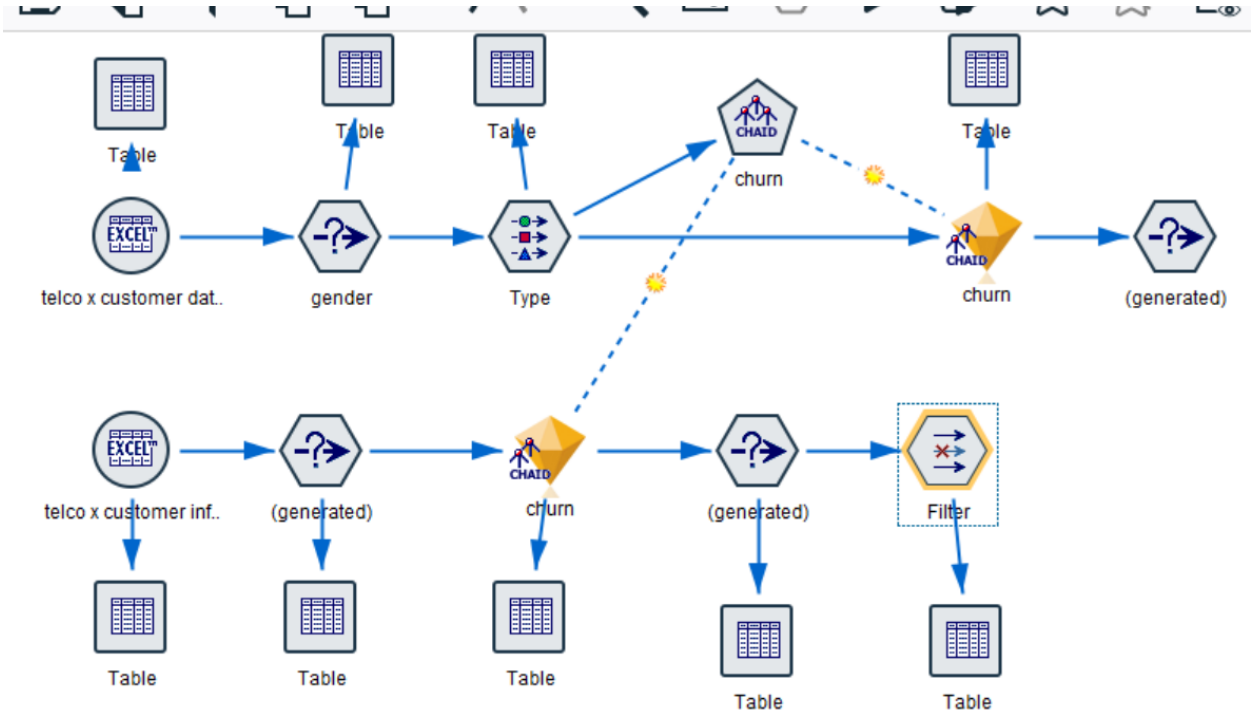
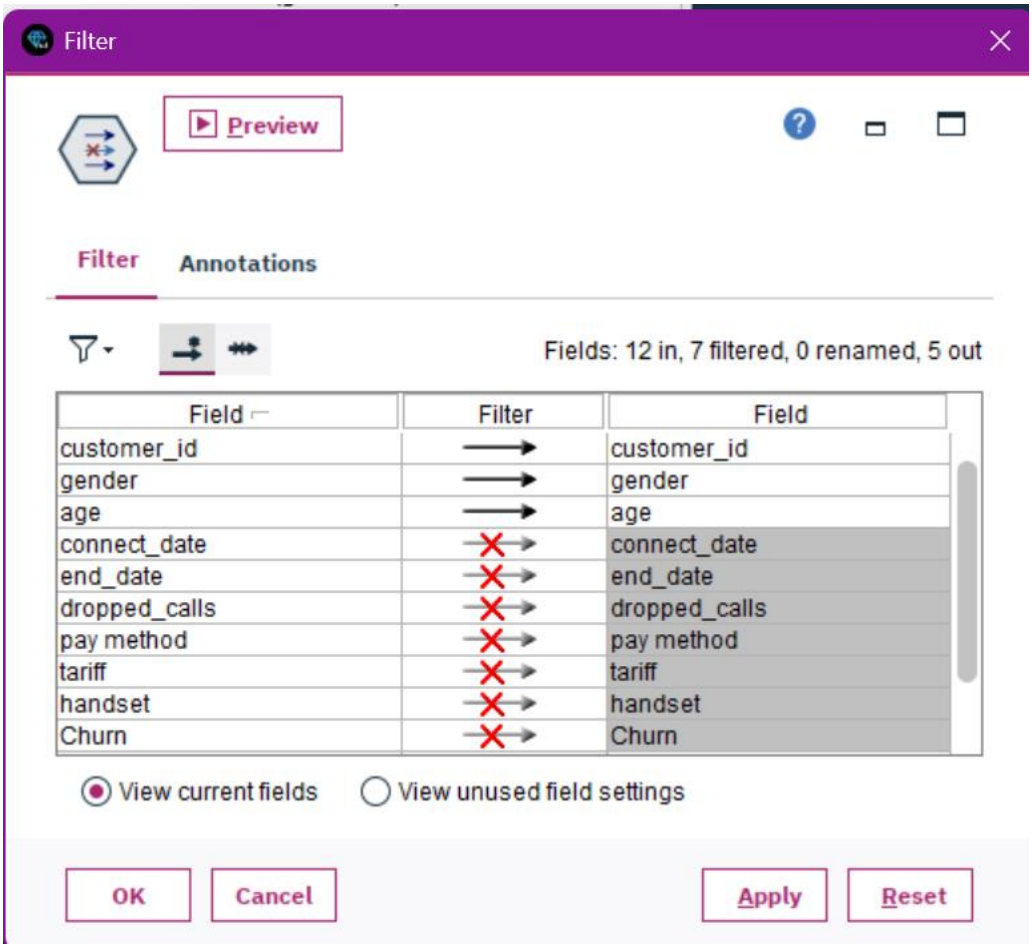
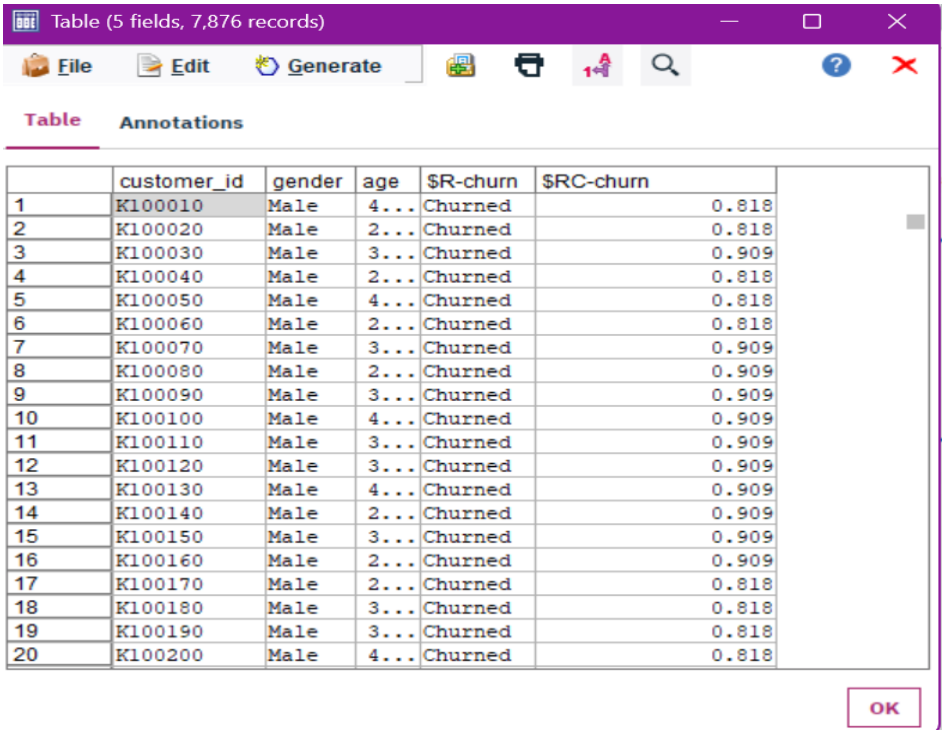


Table
Table
Table
CHAID
churn
Table
telco x customer dat..
gender
Type
CHAID
churn
(generated)
telco x customer inf..
(generated)
CHAID
churn
(generated)
Filter
Table
Table
Table
Table
Table


Filter
Preview
Filter
Annotations
Fields: 12 in, 7 filtered, 0 renamed, 5 out
Field
Filter
Field
customer_id
gender
age
connect_date
end_date
dropped_calls
pay method
tariff
handset
Churn
View current fields
View unused field settings
OK
Cancel
Apply
Reset



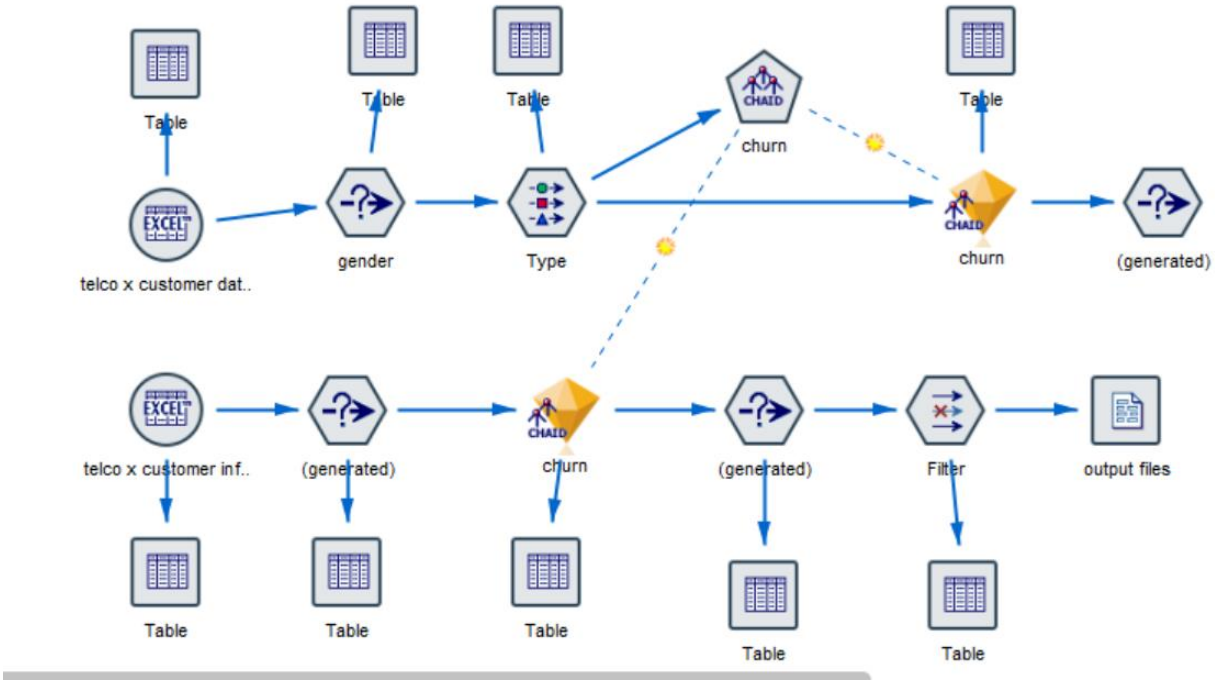
	customer_id	gender	age	$R-churn	$RC-churn
1	K100010	Male	4...	Churned	0.818
2	K100020	Male	2...	Churned	0.818
3	K100030	Male	3...	Churned	0.909
4	K100040	Male	2...	Churned	0.818
5	K100050	Male	4...	Churned	0.818
6	K100060	Male	2...	Churned	0.818
7	K100070	Male	3...	Churned	0.909
8	K100080	Male	2...	Churned	0.909
9	K100090	Male	3...	Churned	0.909
10	K100100	Male	4...	Churned	0.909
11	K100110	Male	3...	Churned	0.909
12	K100120	Male	3...	Churned	0.909
13	K100130	Male	4...	Churned	0.909
14	K100140	Male	2...	Churned	0.909
15	K100150	Male	3...	Churned	0.909
16	K100160	Male	2...	Churned	0.909
17	K100170	Male	2...	Churned	0.818
18	K100180	Male	3...	Churned	0.818
19	K100190	Male	3...	Churned	0.818
20	K100200	Male	4...	Churned	0.818

Step 11: Connect the **Output Files** node to the **Filter** node.

Double-click the **Output Files** node to configure it.

In the dialog box:

- Select **Excel (.xlsx)** as the output format (you can also choose CSV or SAV).
- Name the file something like **Predicted_Churn_Customers.xlsx**. Click **OK** and then **Run** the Output Files node.
- The filtered churn prediction results will now be saved as an Excel file containing only customers who are likely to churn.



This step allows you to **export actionable insights**, so the business team can take follow-up actions such as:

- Contacting customers at risk of churn
- Offering retention discounts or loyalty benefits