

**BABU BANARASI DAS  
UNIVERSITY**

**School of Computer Application (DS & AI)**

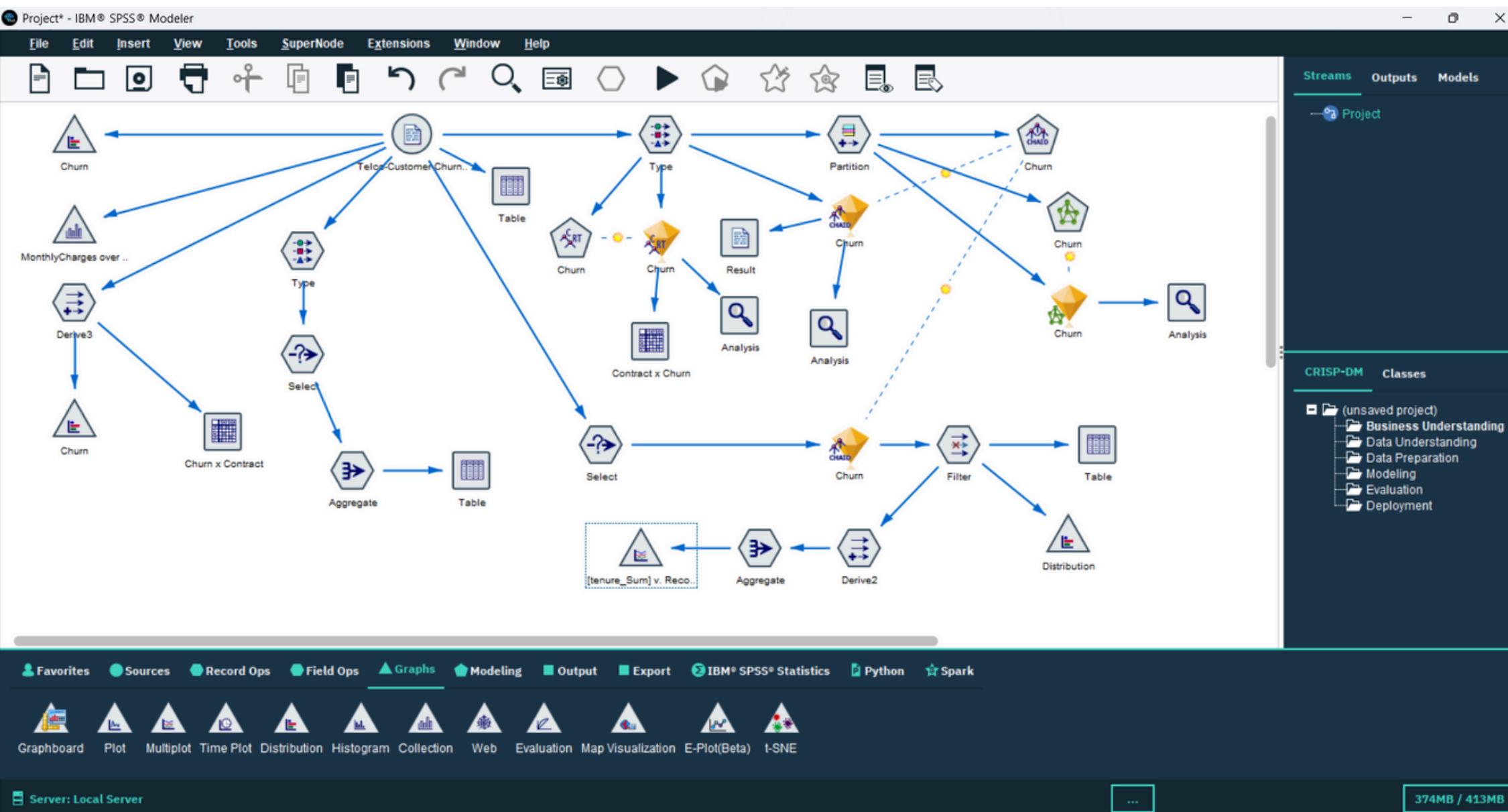
**Predictive Analysis**

**Aditya Yadav Amit Yadav**

**Submitted to: Mr Vikash (IBM)**

# **CHURN ANALYSIS OF TELECOM CUSTOMERS USING IBM SPSS MODELER ON AN OPEN SOURCE DATASET**

This is the Stream we made to analyze the telecom customer churn.



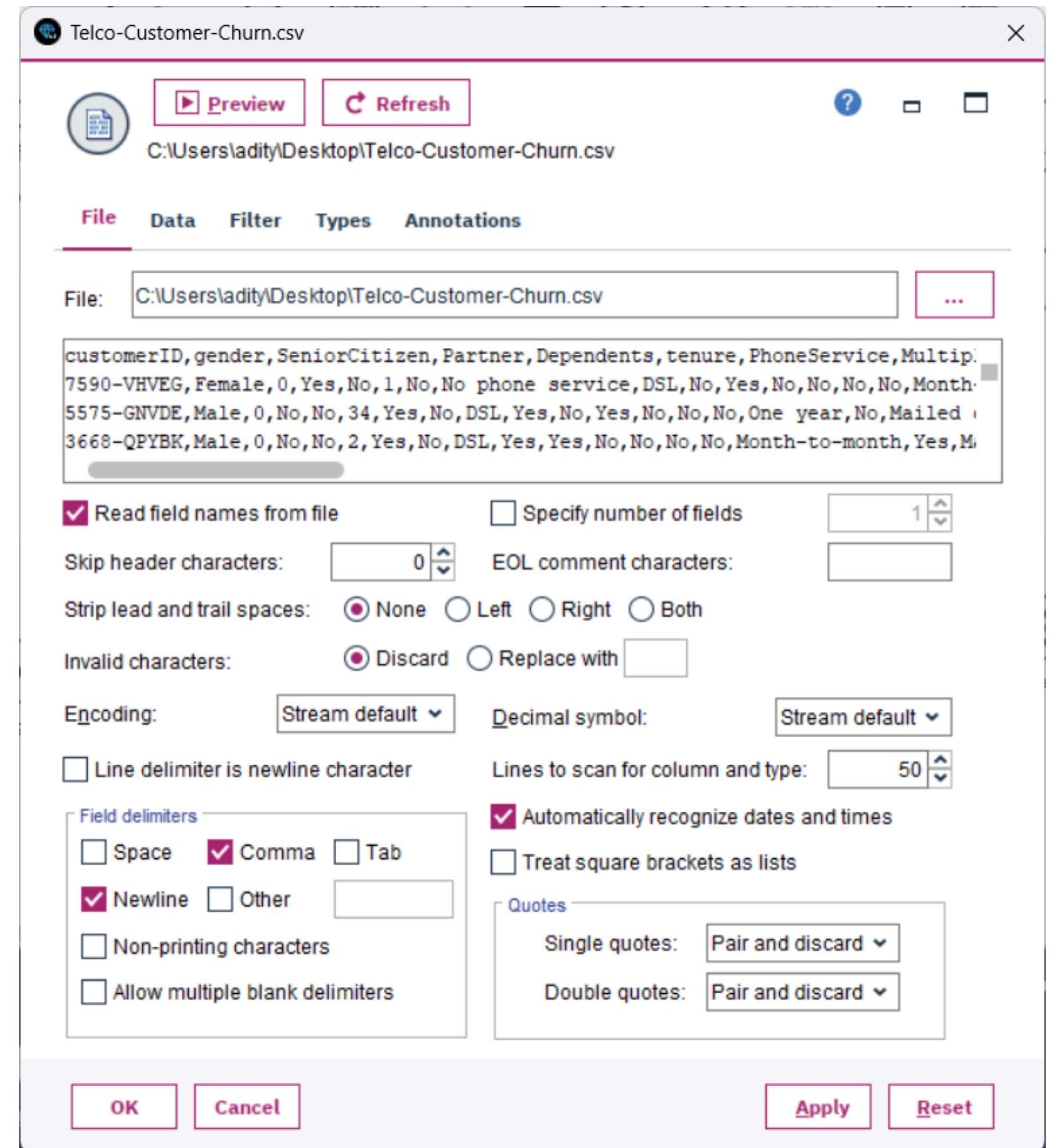
Now we will Understand the step-by-step procedure of creating this.

To Analyze the customer Churn we first need to have a proper dataset and upload it to IBM SPSS Modeler Using Var.file from Sourses.

The Data is in a CSV file format.

it contains 21 fields which are :-

customerID,gender,SeniorCitizen,Partner,  
Dependents,tenure,PhoneService,MultipleLines,  
InternetService,OnlineSecurity,OnlineBackup,  
DeviceProtection,TechSupport,StreamingTV,  
StreamingMovies,Contract,PaperlessBilling,  
PaymentMethod,MonthlyCharges>TotalCharges,  
Churn



Now we can use the Table view from the Output to view and review the data and data types to make changes before the analysis begins. This also gives us an overview of the data.

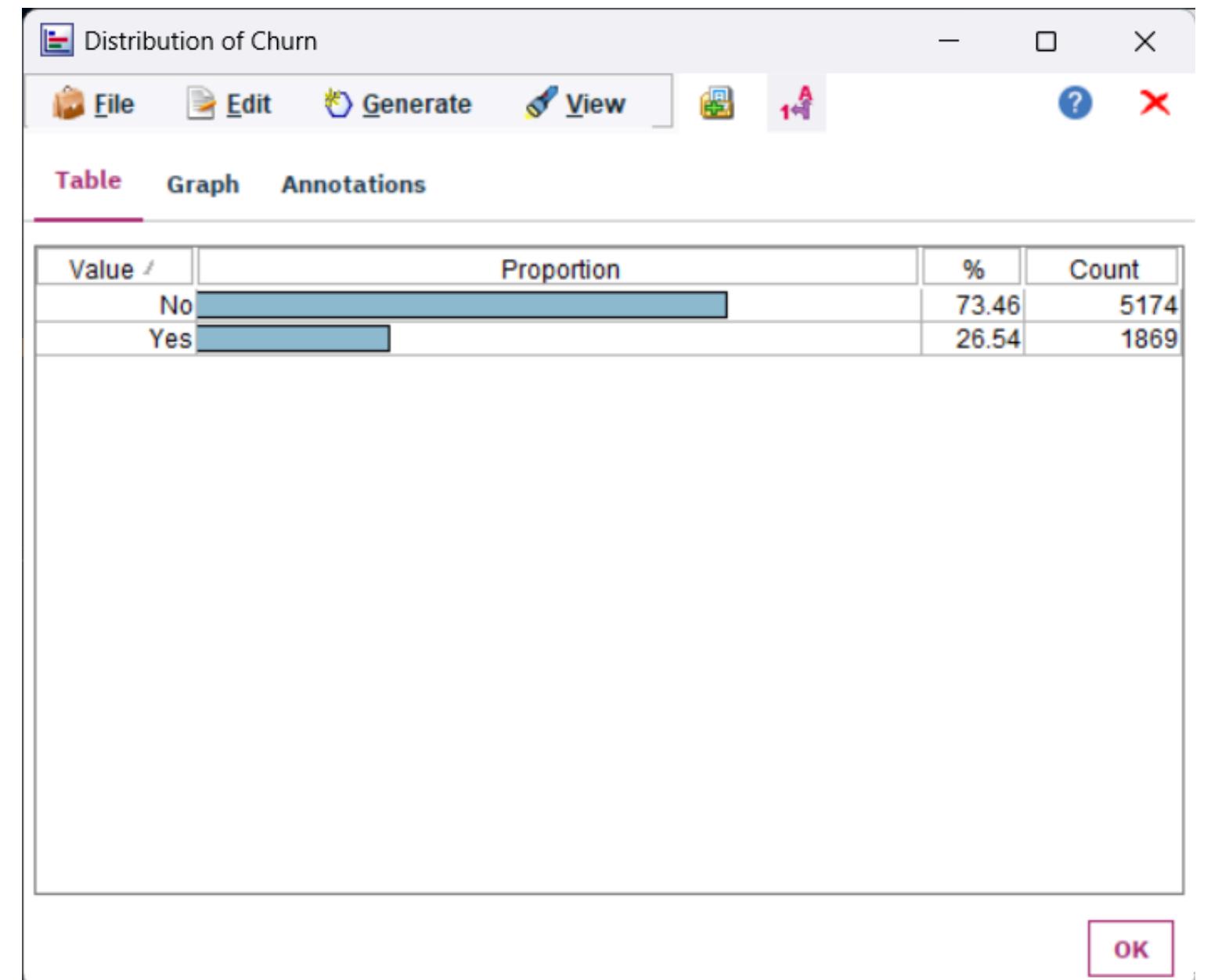
The screenshot shows a 'Table' view settings dialog with a tab bar at the top labeled 'Settings', 'Format' (which is selected), 'Output', and 'Annotations'. Below the tab bar is a table with columns for 'Field', 'Format', 'Justify', and 'Column Width'. The table lists various fields with their current settings. At the bottom of the dialog are two radio buttons: 'View current fields' (selected) and 'View unused field settings'. At the very bottom are five buttons: 'OK', 'Run', 'Cancel', 'Apply', and 'Reset'.

Field	Format	Justify	Column Width
A customerID	Auto	Auto	Auto
A gender	Auto	Auto	Auto
SeniorCitizen	####	Auto	Auto
A Partner	Auto	Auto	Auto
A Dependents	Auto	Auto	Auto
tenure	####	Auto	Auto
A PhoneService	Auto	Auto	Auto
A MultipleLines	Auto	Auto	Auto
A InternetService	Auto	Auto	Auto
A OnlineSecurity	Auto	Auto	Auto
A OnlineBackup	Auto	Auto	Auto
A DeviceProtection	Auto	Auto	Auto
A TechSupport	Auto	Auto	Auto
A StreamingTV	Auto	Auto	Auto
A StreamingMovies	Auto	Auto	Auto
A Contract	Auto	Auto	Auto
A PaperlessBilling	Auto	Auto	Auto
A PaymentMethod	Auto	Auto	Auto
MonthlyCharges	####.##	Auto	Auto
TotalCharges	####.##	Auto	Auto
A Churn	Auto	Auto	Auto

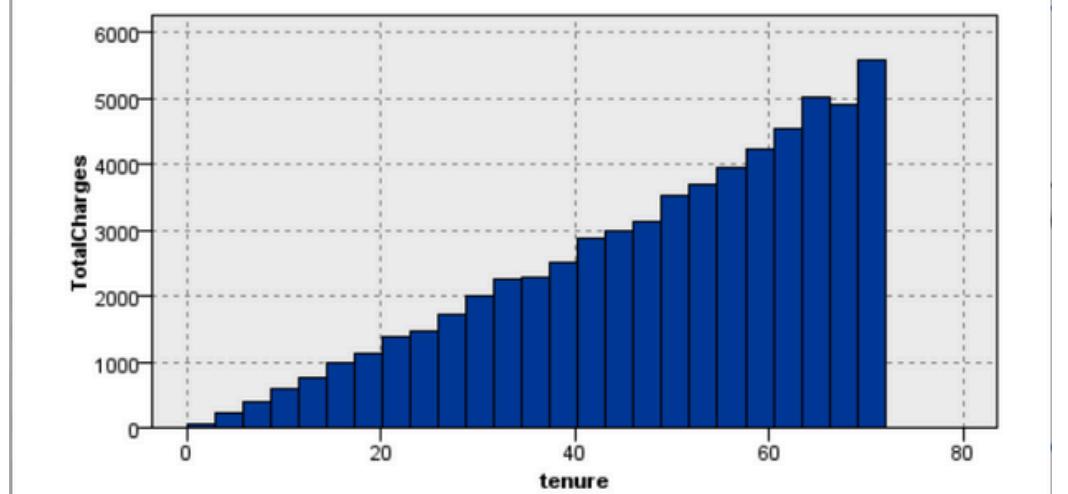
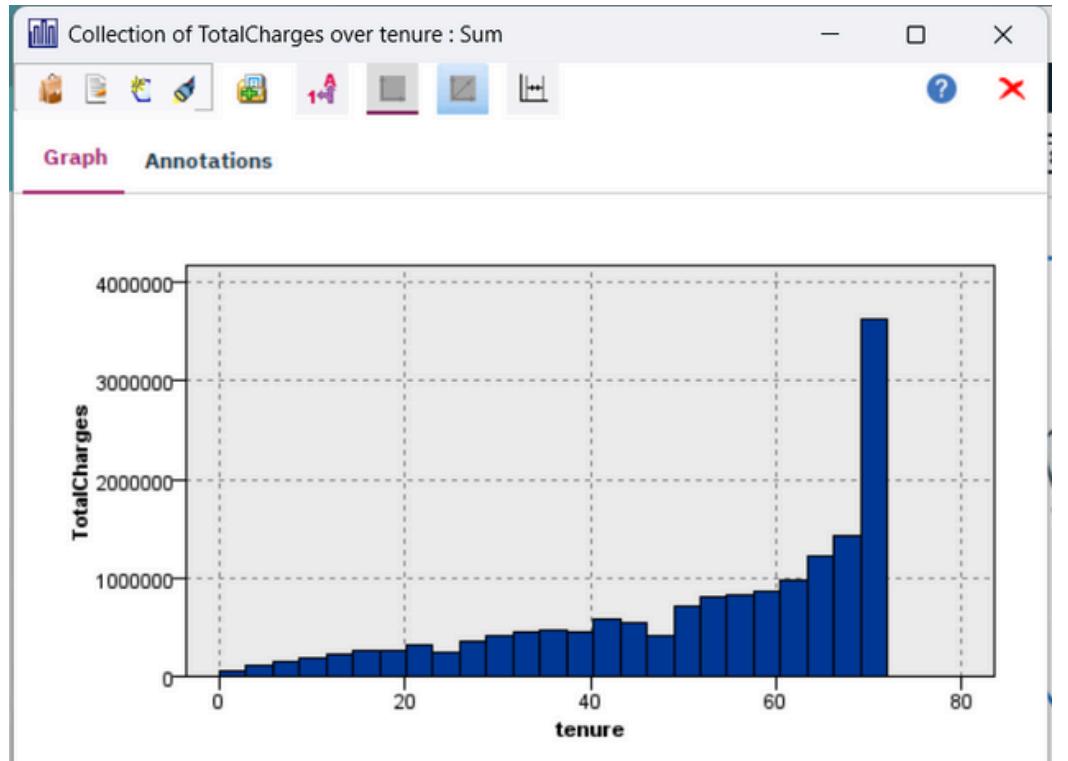
View current fields  View unused field settings

OK ▶ Run Cancel Apply Reset

We now use Graph view to analyse the rate of churn of customers during a particular time period.



We can analyse the data with various factors to try to understand the relationship between different fields of data, like the example shown in the graph.



Here we use a matrix from the output to show the more complex relationships between data in our dataset, which shows that the rate of churn decreases inversely as the tenure of the contract increases and is high among short-term customers.

Matrix of Churn by Contract

File Edit Generate

Matrix Appearance Annotations

Contract

Churn	Month-to...	One year	Two year
No	2220	1307	1647
Yes	1655	166	48

Cells contain: cross-tabulation of fields (including missing values)  
Chi-square = 1,184.597, df = 2, probability = 0

OK

Now we need to understand the reason we want to do a churn analysis, so we compare the total revenue and revenue lost due to churn. We assume that we have 10% revenue growth due to new customers, which shows that we still have a net revenue loss of 7.8%, and it will be difficult to sustain the operations for long.

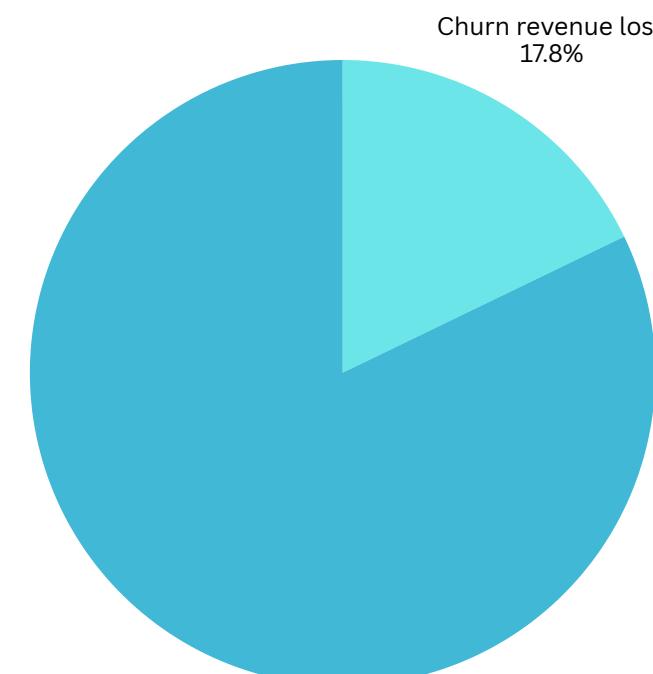


Table (3 fields, 1 records) #3

File Edit Generate

Annotations

	customerID_Count	TotalCharges_Sum	Record_Count
1	1869	2862926.900	1869

Type

Preview

?

Types Format Annotations

Read Values Clear Values Clear All Values

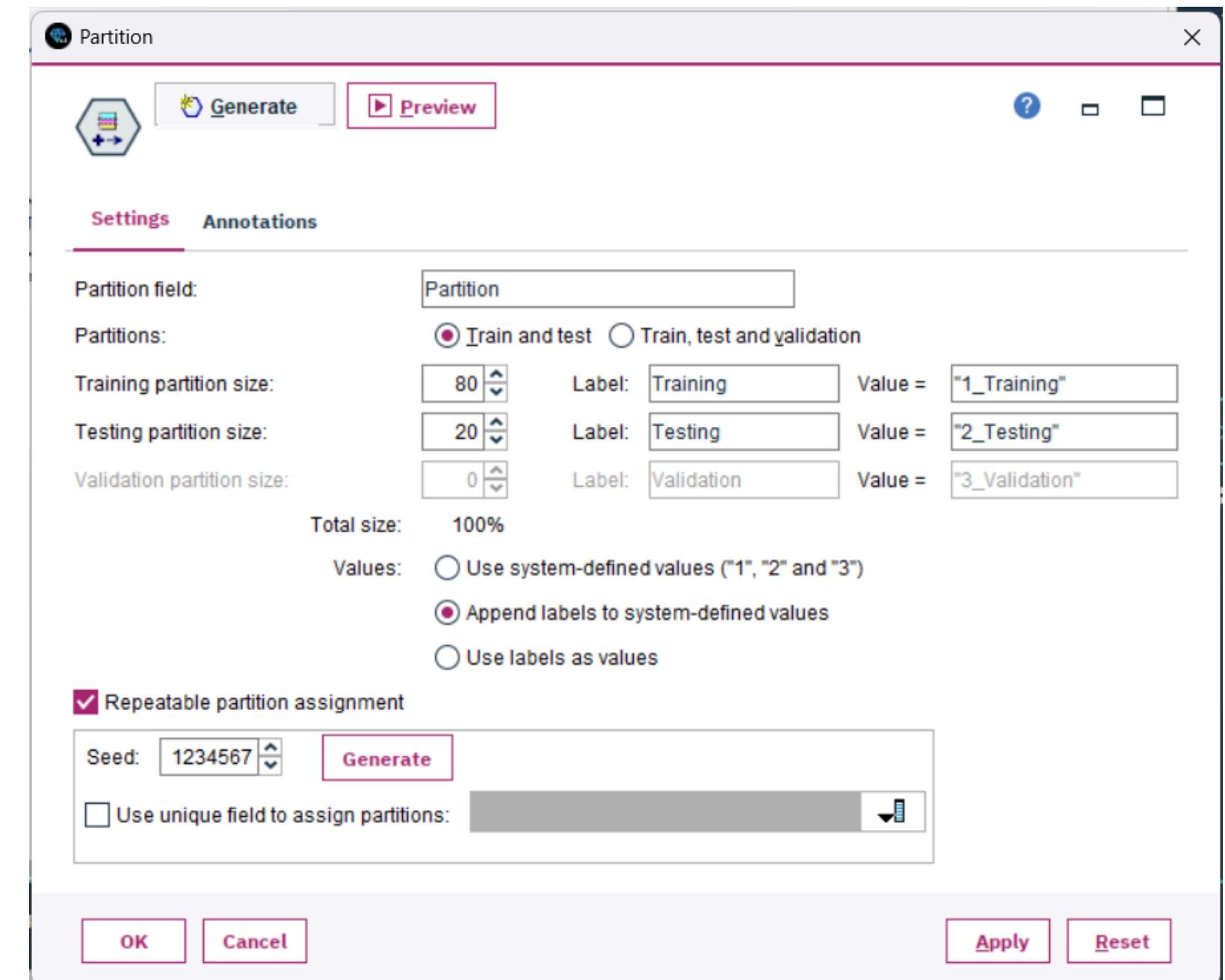
Field	Measurement	Values	Missing	Check	Role
A customerID	Typeless		None	None	<input checked="" type="checkbox"/> None
A gender	Flag	Male/Fem...	None	None	<input checked="" type="checkbox"/> Input
SeniorCitizen	Continuous	[0,1]	None	None	<input checked="" type="checkbox"/> Input
A Partner	Flag	Yes/No	None	None	<input checked="" type="checkbox"/> Input
A Dependents	Flag	Yes/No	None	None	<input checked="" type="checkbox"/> Input
tenure	Continuous	[0,72]	None	None	<input checked="" type="checkbox"/> Input
A PhoneService	Flag	Yes/No	None	None	<input checked="" type="checkbox"/> Input
A MultipleLines	Nominal	No,"No ph...	None	None	<input checked="" type="checkbox"/> Input
A InternetService	Nominal	DSL,"Fiber...	None	None	<input checked="" type="checkbox"/> Input
A OnlineSecurity	Nominal	No,"No int...	None	None	<input checked="" type="checkbox"/> Input
A OnlineBackup	Nominal	No,"No int...	None	None	<input checked="" type="checkbox"/> Input
A DeviceProtec...	Nominal	No,"No int...	None	None	<input checked="" type="checkbox"/> Input
A TechSupport	Nominal	No,"No int...	None	None	<input checked="" type="checkbox"/> Input
A StreamingTV	Nominal	No,"No int...	None	None	<input checked="" type="checkbox"/> Input
A StreamingMo...	Nominal	No,"No int...	None	None	<input checked="" type="checkbox"/> Input
A Contract	Nominal	Month-to...	None	None	<input checked="" type="checkbox"/> Input
A PaperlessBil...	Flag	Yes/No	None	None	<input checked="" type="checkbox"/> Input
A PaymentMet...	Nominal	"Bank tran...	None	None	<input checked="" type="checkbox"/> Input
MonthlyChar...	Continuous	[18.25,118...	None	None	<input checked="" type="checkbox"/> Input
TotalCharges	Continuous	[18.8,8684...	None	None	<input checked="" type="checkbox"/> Input
A Churn	Flag	Yes/No	None	None	<input checked="" type="checkbox"/> Target

View current fields  View unused field settings

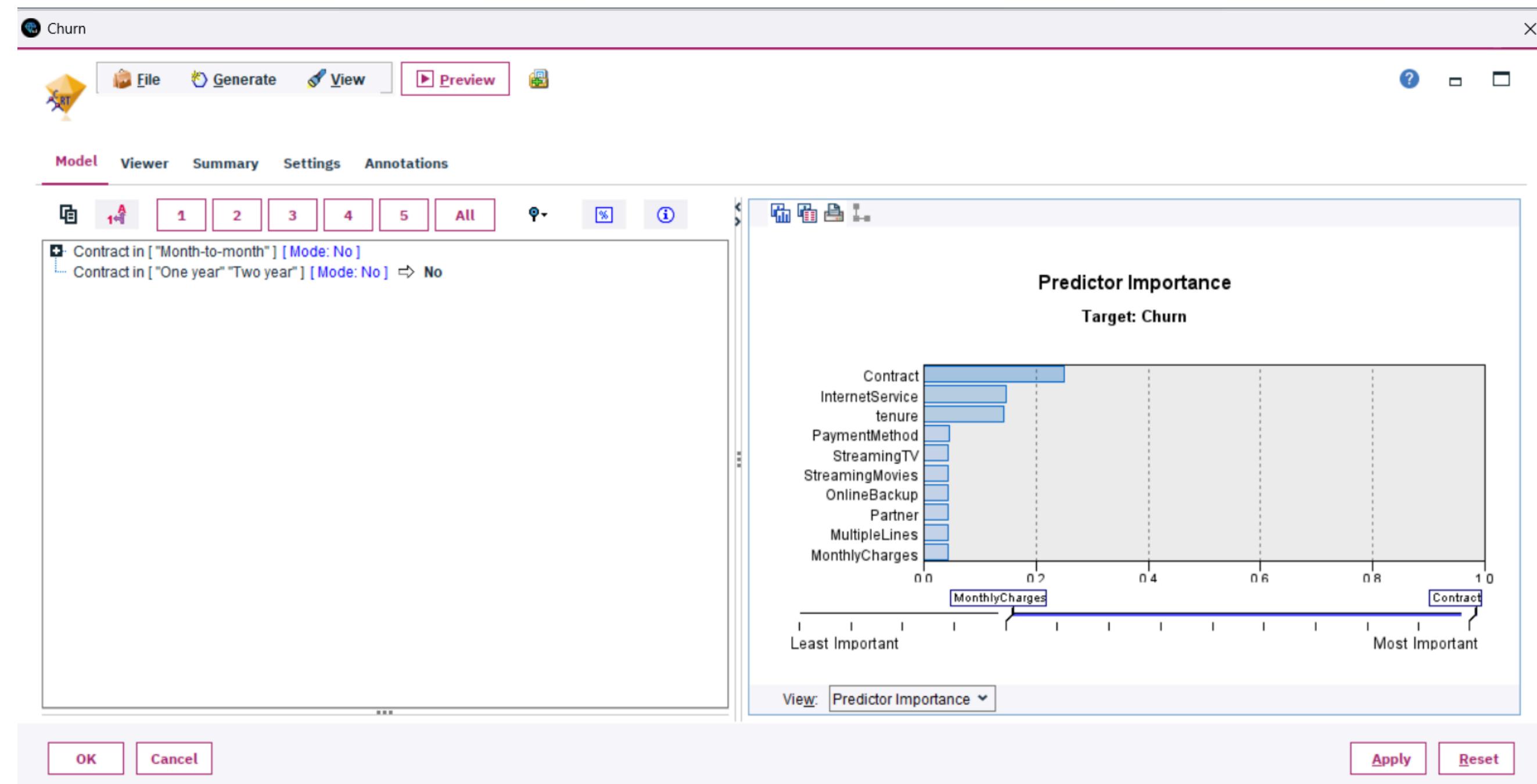
OK Cancel Apply Reset

We will now use the type field to specify the various types of data fields into various categories such as input, none, and target.

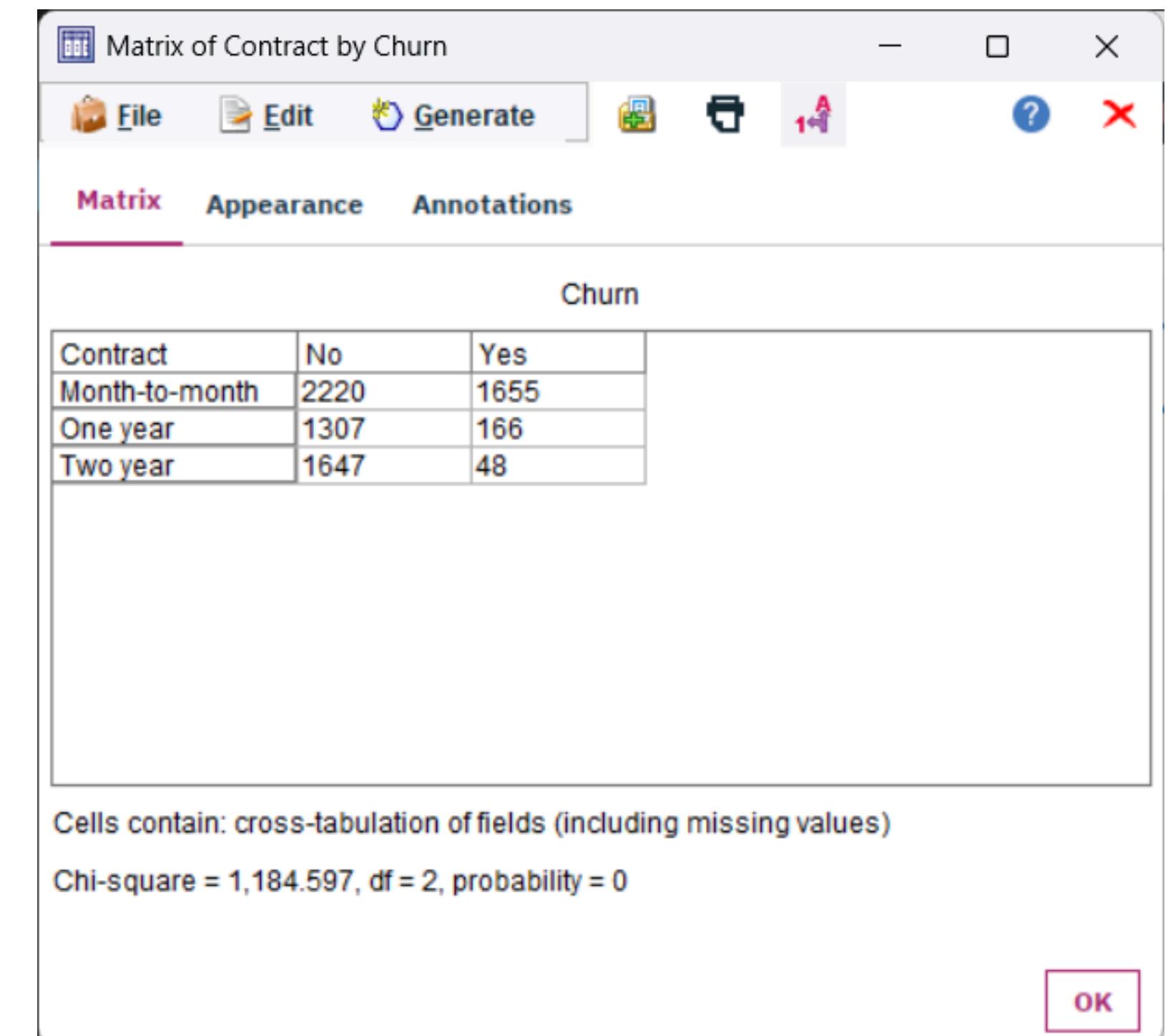
We can also use Partitioning to separate data into sets of training and testing data to test out the model accuracy.



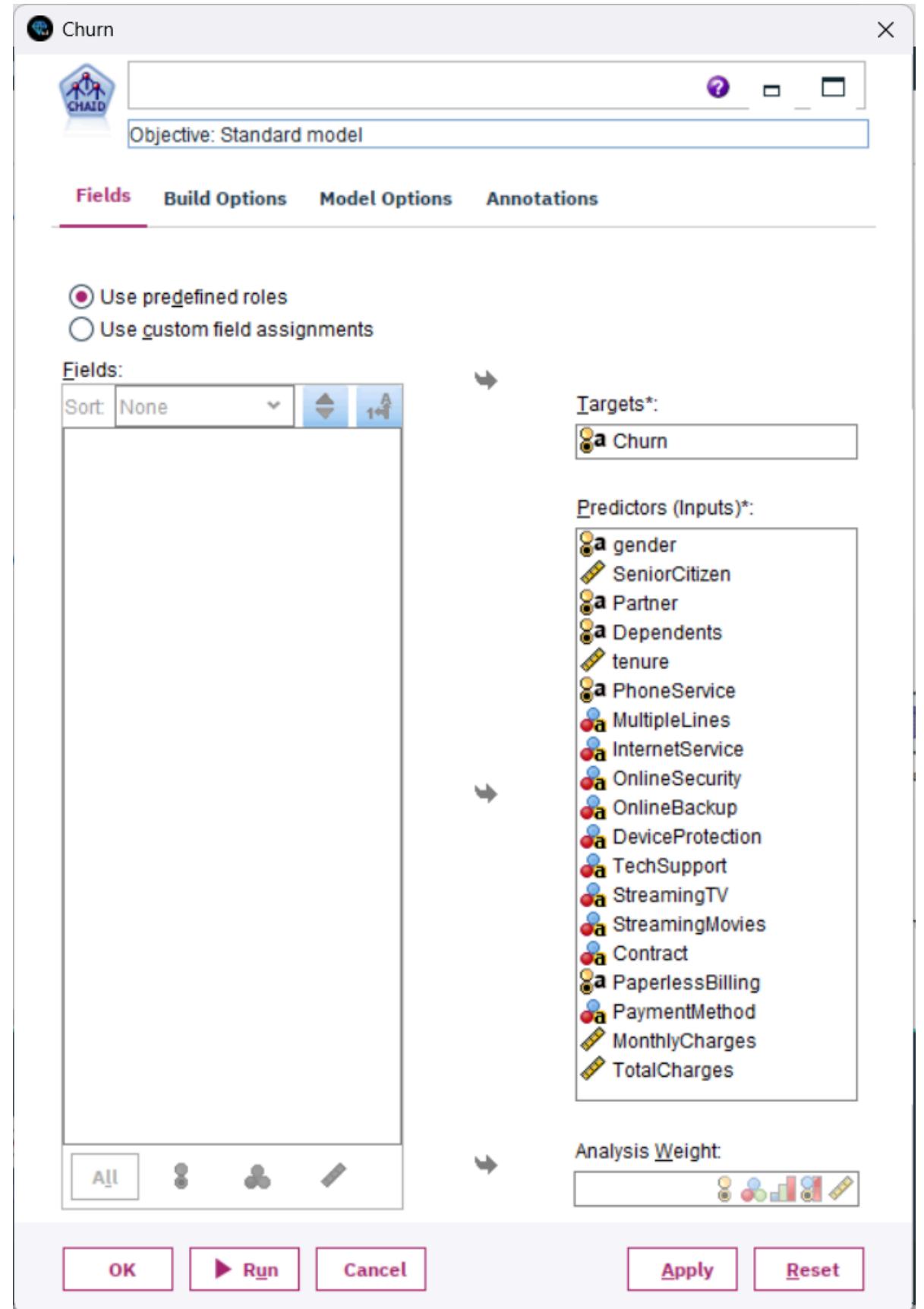
We use the CRT model to perform an analysis to find the factors affecting customer churn and make changes to improve these factors. For example, contract and internet service have a greater effect on customer churn than other fields such as monthly charges and payment method.



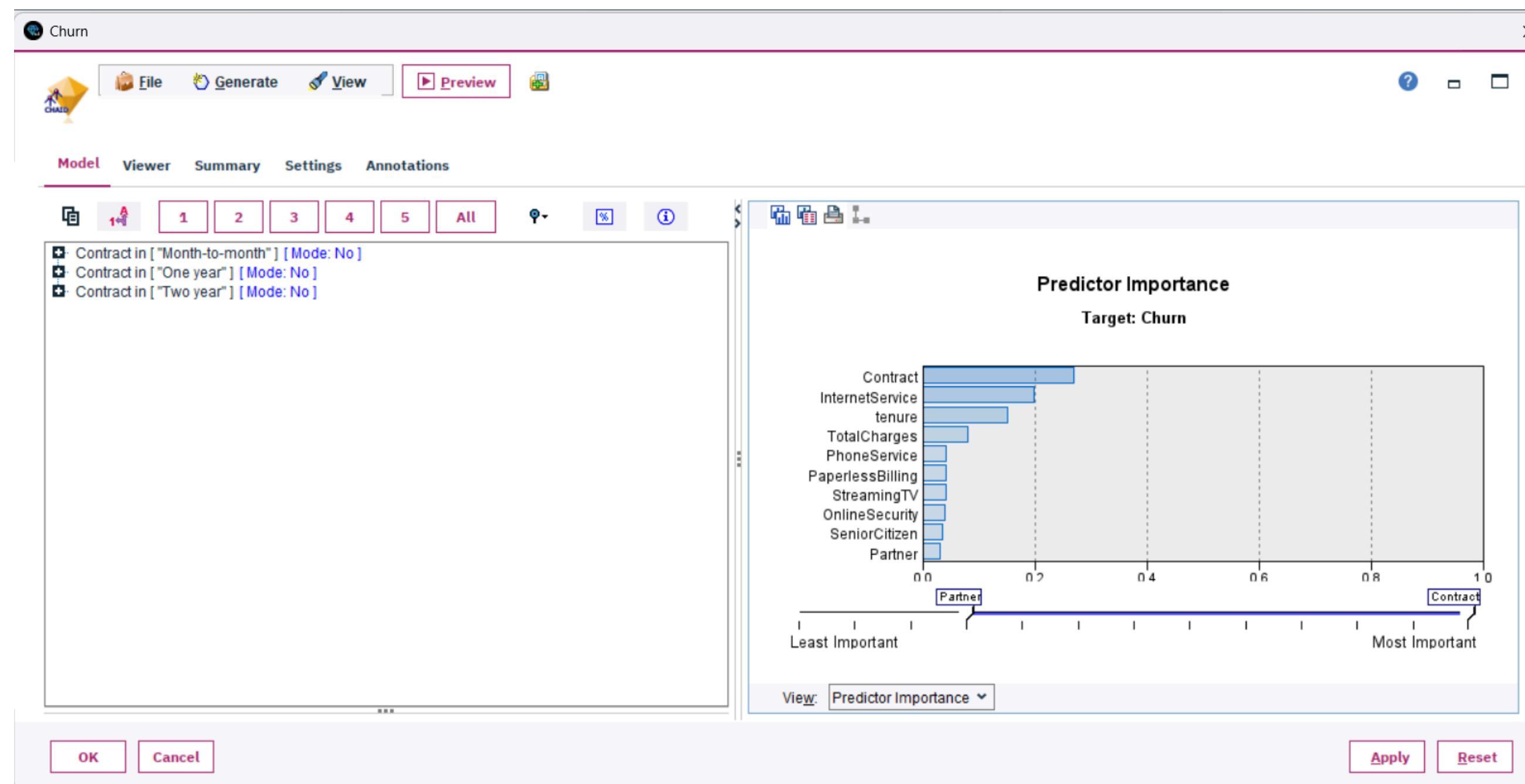
We now use an analysis node to analyse the output produced by the CRT model and analyse various fields.



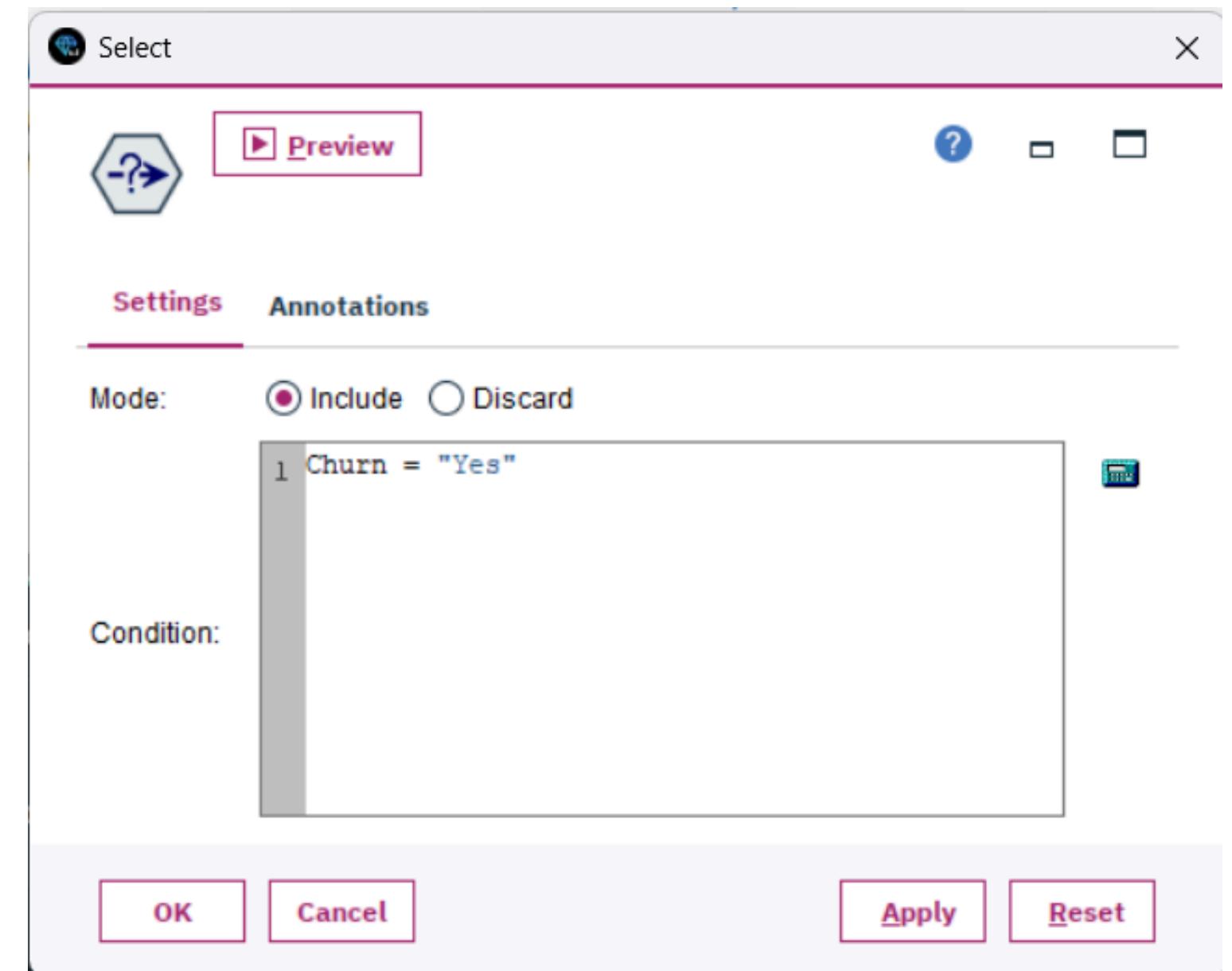
We will use the CHAID model to further improve our analysis, but first we have to specify the various parameters of the CHAID model in the setting of the model.



We use the CHAID model to perform an analysis to find the factors affecting customer churn and make changes to improve these factors. For example, contract and internet service have a greater effect on customer churn than other fields such as monthly charges and payment method.



Add a select node to the output of the CHAID model and add various conditions to use the model output for further analysis of the model output and visualisation.



Use a filter node to select the fields you want to specify and analyse through various nodes such as visualisation.

Filter

Preview

Filter Annotations

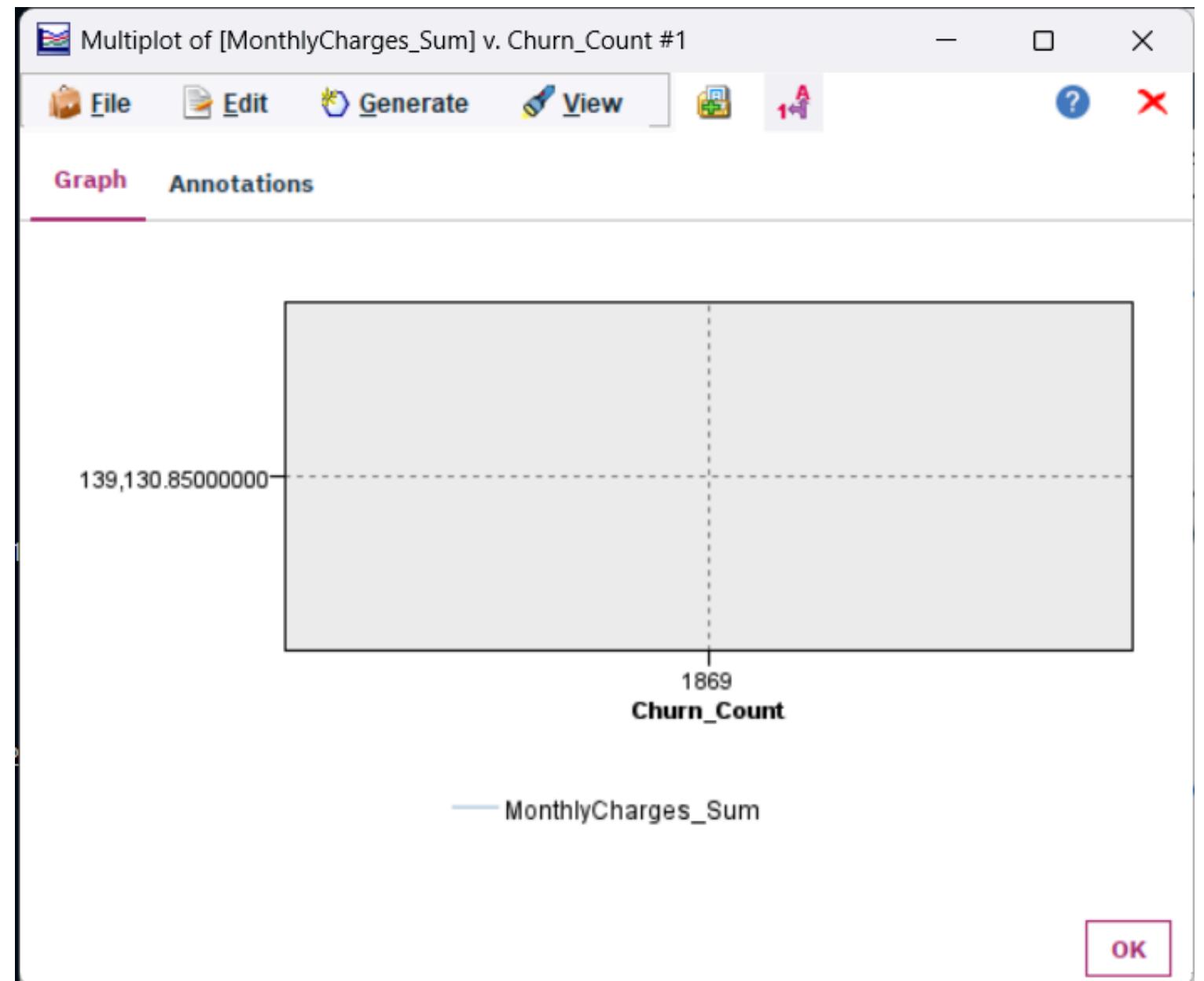
Fields: 23 in, 21 filtered, 0 renamed, 2 out

Field ↗	Filter	Field
customerID	✗ ➔	customerID
gender	✗ ➔	gender
SeniorCitizen	✗ ➔	SeniorCitizen
Partner	✗ ➔	Partner
Dependents	✗ ➔	Dependents
tenure	➡	tenure
PhoneService	✗ ➔	PhoneService
MultipleLines	✗ ➔	MultipleLines
InternetService	✗ ➔	InternetService
OnlineSecurity	✗ ➔	OnlineSecurity
OnlineBackup	✗ ➔	OnlineBackup
DeviceProtection	✗ ➔	DeviceProtection
TechSupport	✗ ➔	TechSupport
StreamingTV	✗ ➔	StreamingTV
StreamingMovies	✗ ➔	StreamingMovies
Contract	✗ ➔	Contract
PaperlessBilling	✗ ➔	PaperlessBilling
PaymentMethod	✗ ➔	PaymentMethod
MonthlyCharges	✗ ➔	MonthlyCharges
TotalCharges	✗ ➔	TotalCharges
Churn	➡	Churn
\$R-Churn	✗ ➔	\$R-Churn
\$RC-Churn	✗ ➔	\$RC-Churn

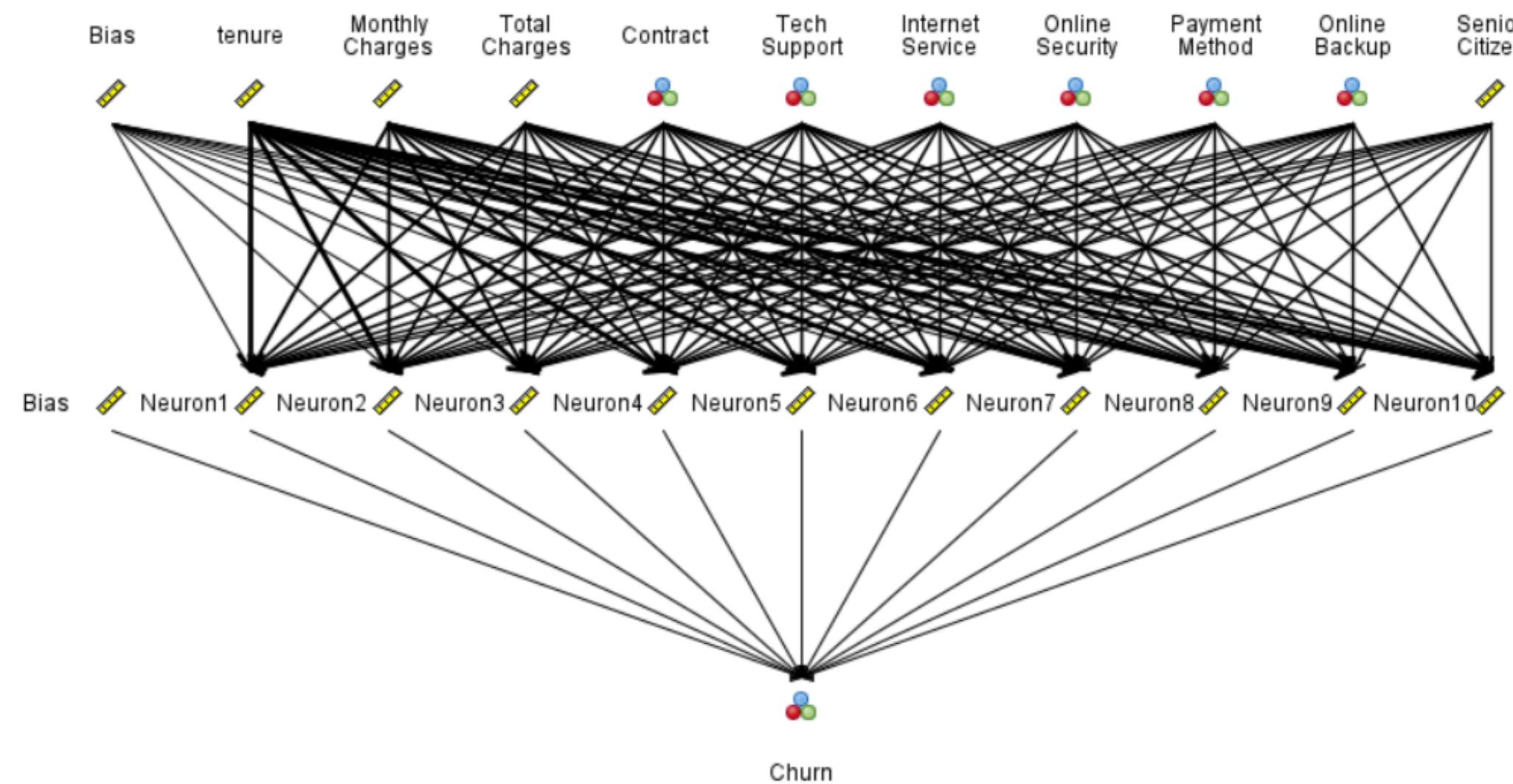
View current fields  View unused field settings

OK Cancel Apply Reset

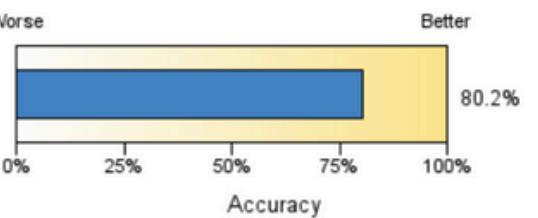
Here we analyse the total monthly loss of revenue due to customer churn. We can change the parameters in the previous two nodes to adjust the fields and visualise it more thoroughly.



For future analysis of customer churn probability, we can add a neural network node to create a model and add a database node to predict customer churn probability as we add new customers and prepare a report and take actions to reduce churn.



Model Summary	
Target	Churn
Model	Multilayer Perceptron
Stopping Rule Used	Error cannot be further decreased
Hidden Layer 1 Neurons	10

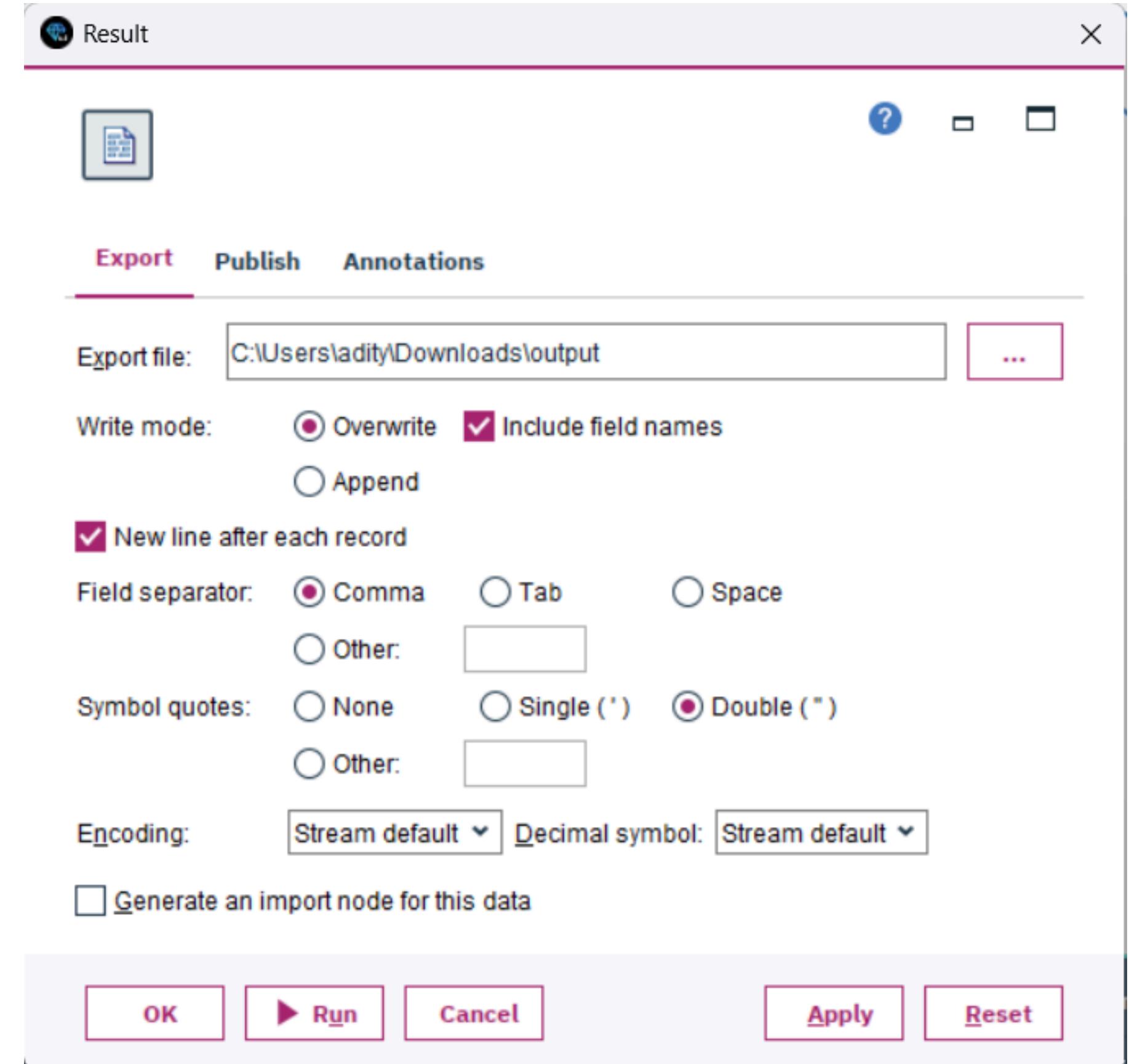


Here are the model specifications for 80% accuracy, which can be improved with changes in data fields and the dataset.

#### Classification for Churn

Overall Percent Correct = 80.3%

Observed	Predicted		Row Percent
	No	Yes	
No	90.2%	9.8%	
Yes	46.2%	53.8%	



Now we can produce an output of our churn prediction model analysis to our specified file.