

**BABU BANARASI DAS
UNIVERSITY**

School of Computer Applications (DS & AI)
Predictive Analysis

Aditya Yadav (1230258037)

Amit Yadav (1230258036)

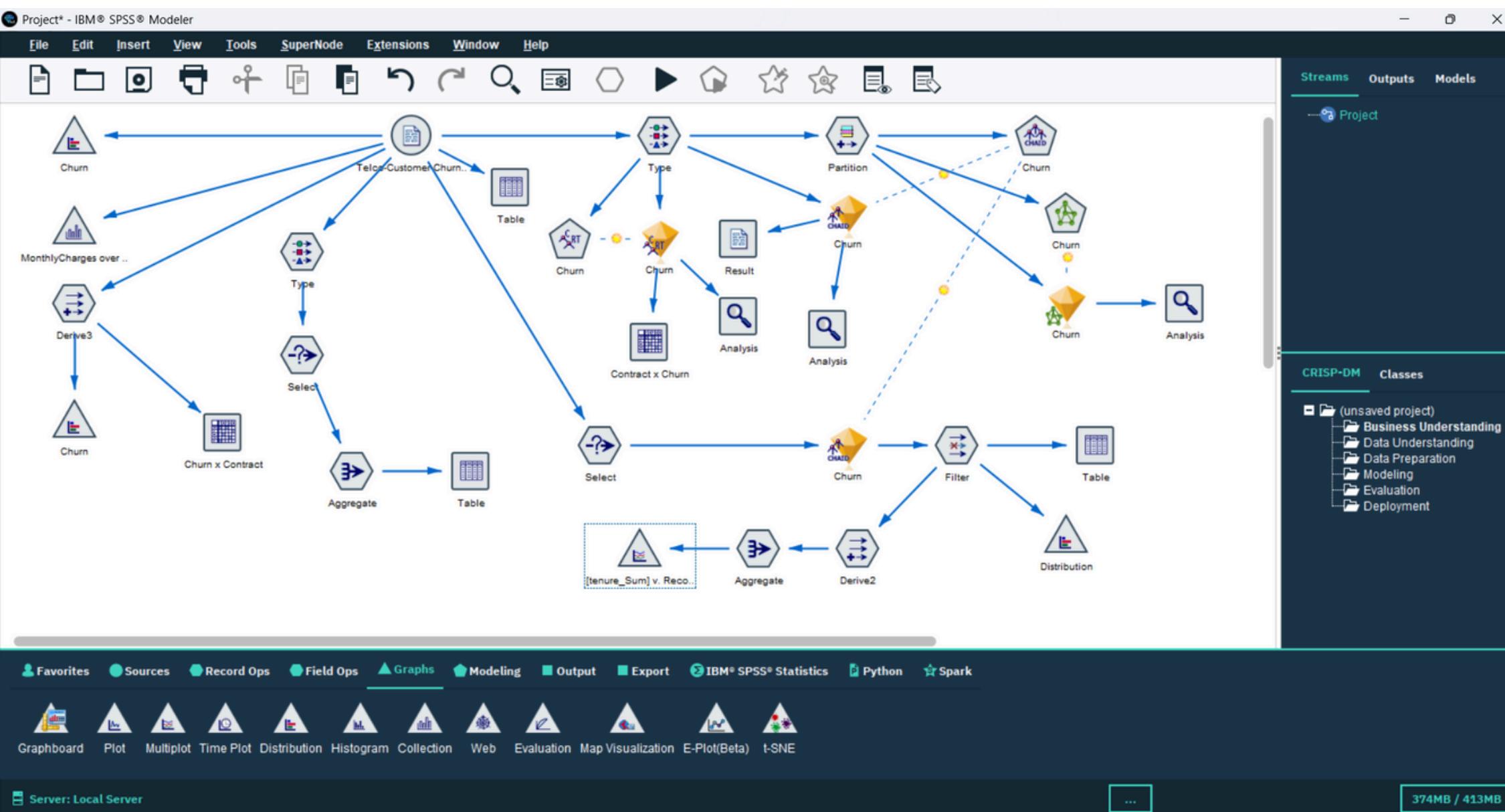
Aditya Srivastava (1230258065)

Aditi Singh (1230258025)

Submitted to: Mr. Vikash (IBM)

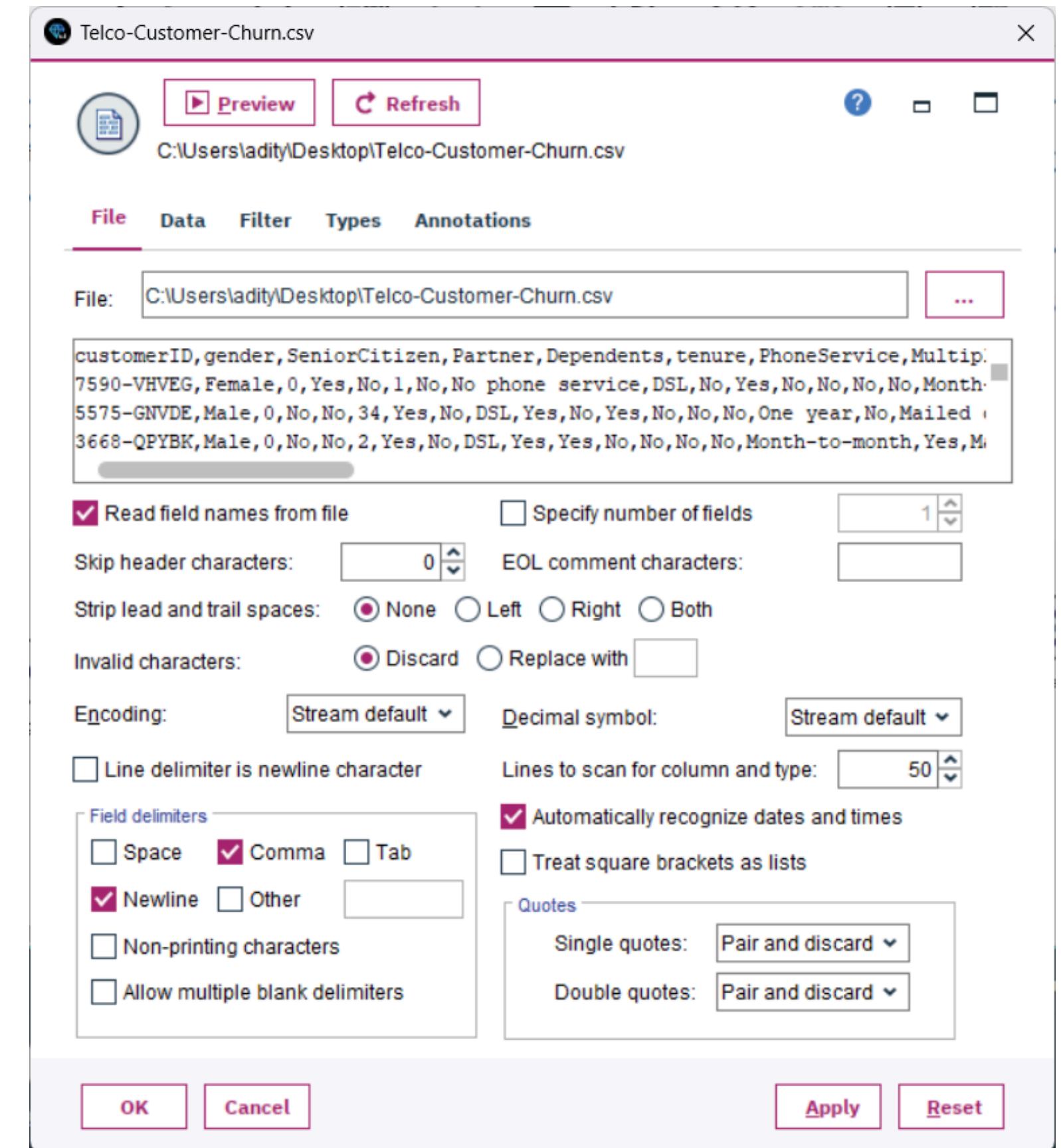
CHURN ANALYSIS OF TELECOM CUSTOMERS USING IBM SPSS MODELER ON AN OPEN SOURCE DATASET

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 analyse the customer churn, we first need to have a proper dataset and upload it to IBM SPSS Modeler using the Var.file from Sources. 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, and Churn.



Table

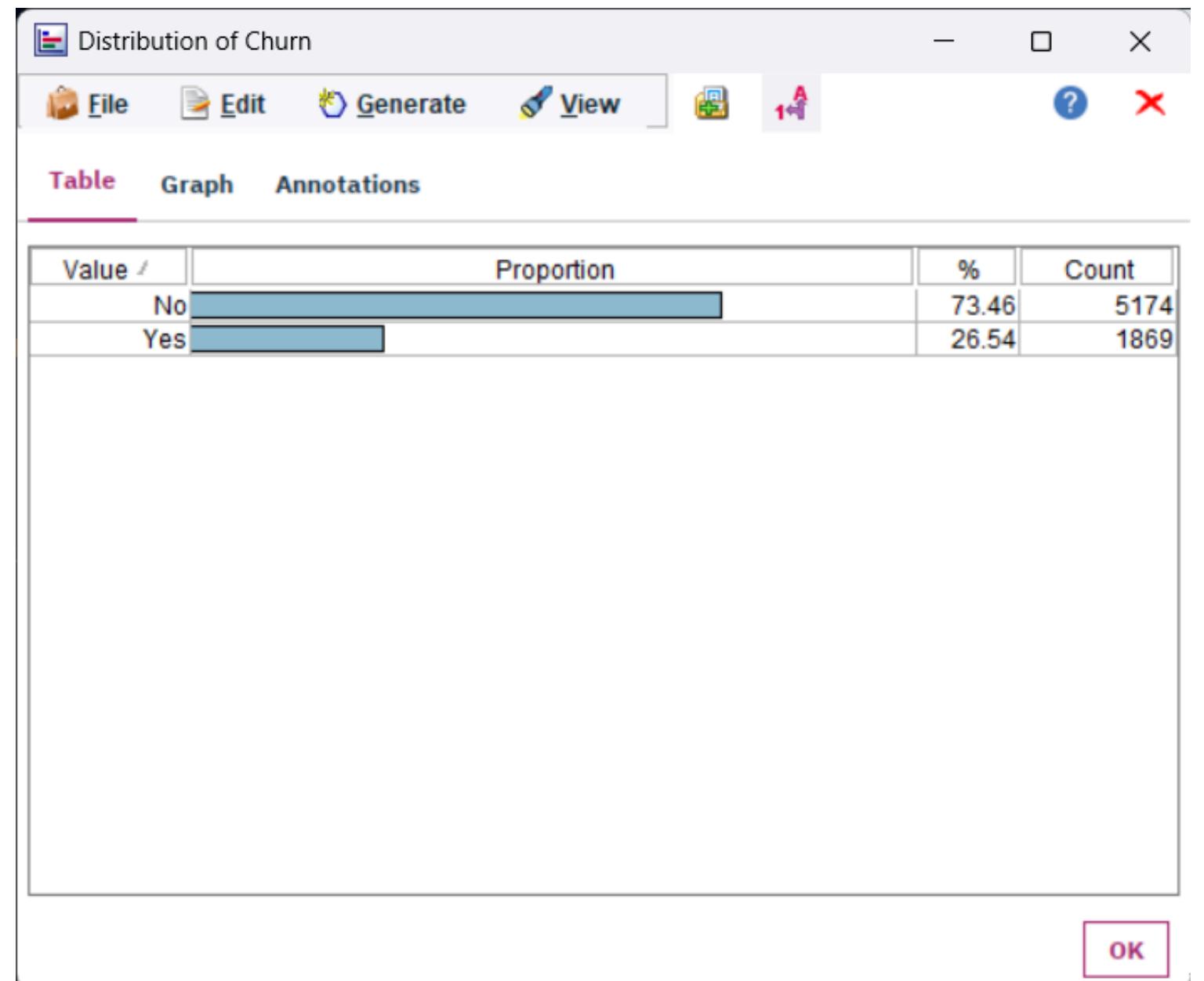
Field	Format	Justify	Column Width
A customerID		Auto	Auto
A gender		Auto	Auto
◇ SeniorCitizen	####	Auto	Auto
A Partner		Auto	Auto
A Dependents		Auto	Auto
◇ tenure	####	Auto	Auto
A PhoneService		Auto	Auto
A MultipleLines		Auto	Auto
A InternetService		Auto	Auto
A OnlineSecurity		Auto	Auto
A OnlineBackup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

View current fields
 View unused field settings

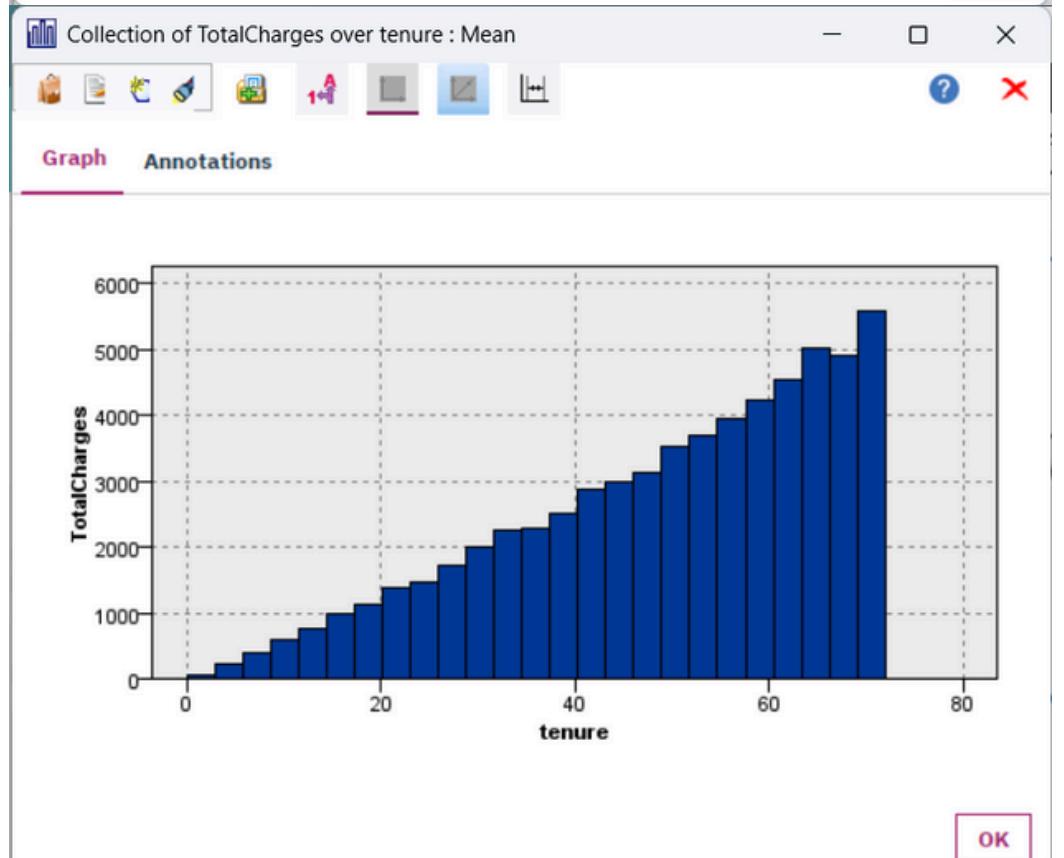
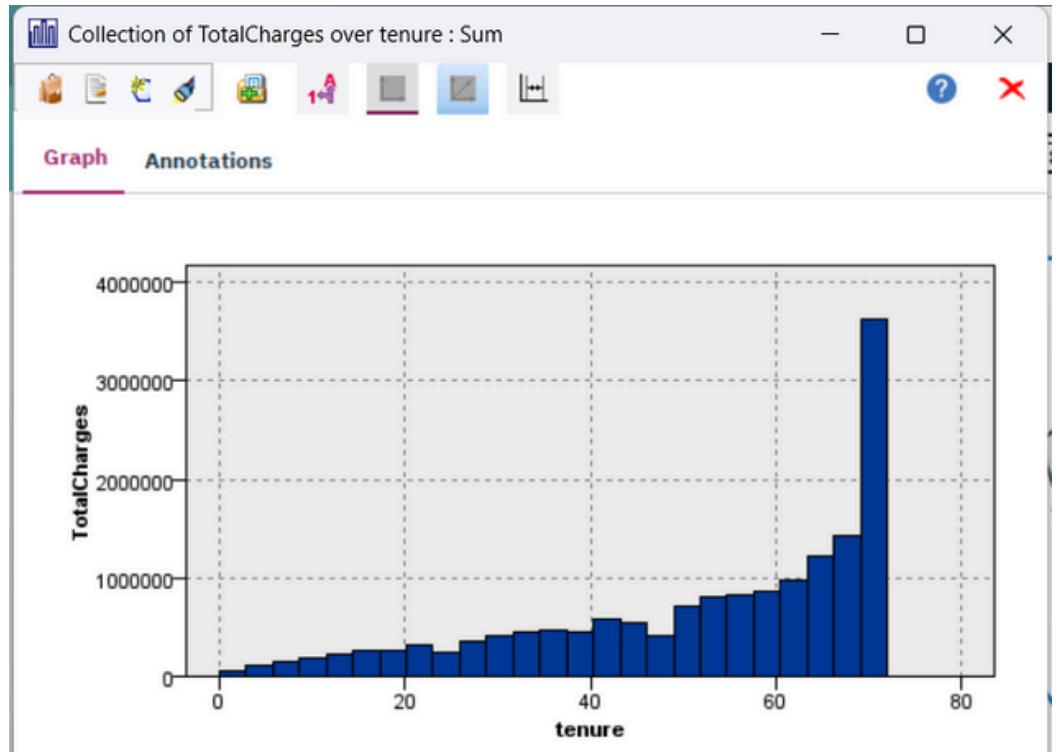
OK Run Cancel **Apply** **Reset**

Now, we can utilize the Table view from the Output to examine and assess the data along with its types, allowing for any necessary adjustments before the analysis starts. This also provides us with a comprehensive overview of the data.

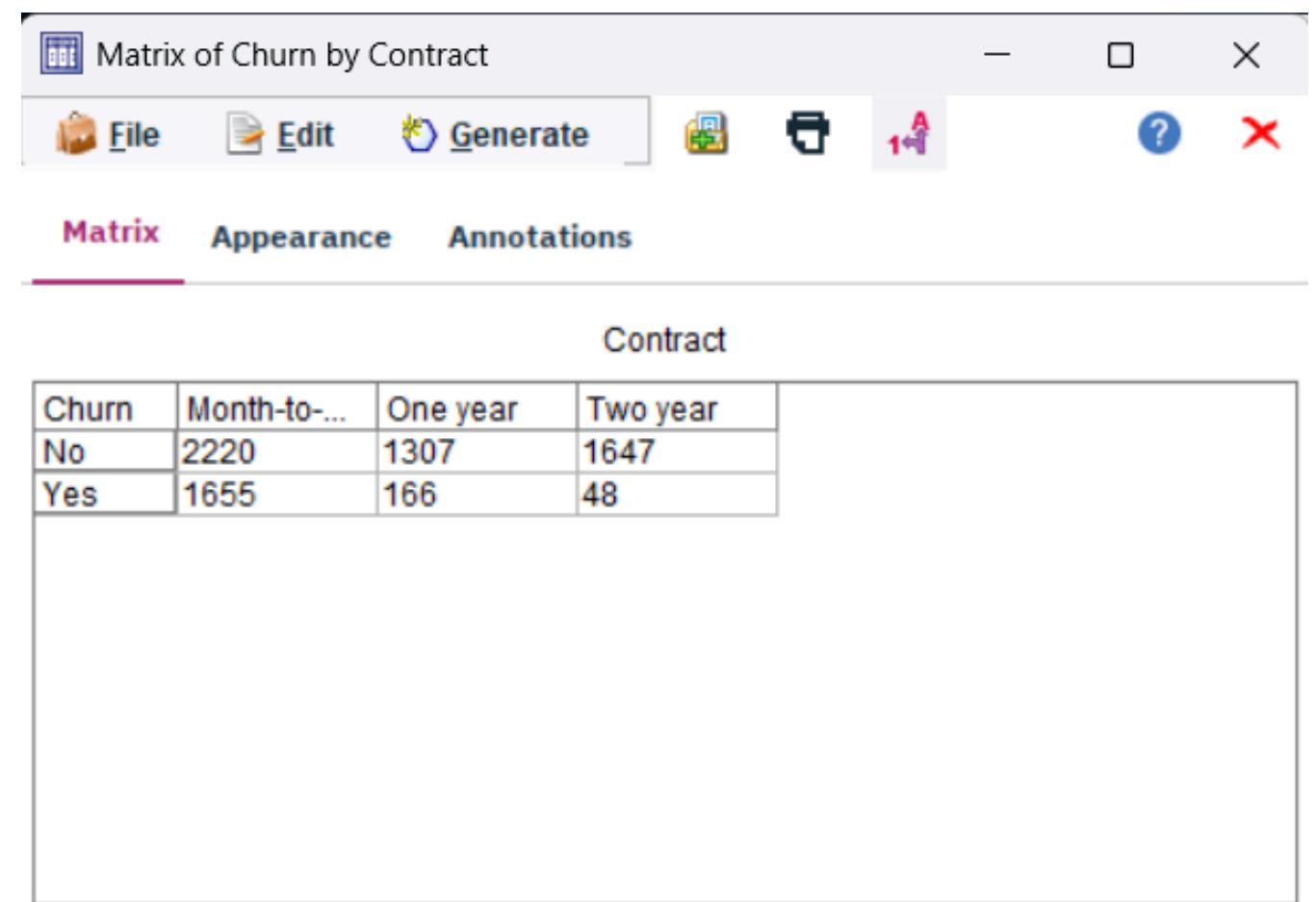
We now use Graph view to analyse the rate of churn of customers during a particular time period.



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 utilize a matrix derived from our output to illustrate the intricate relationships within our dataset. It reveals that the churn rate decreases inversely as contract tenure increases, with a notably high rate among short-term customers.



The screenshot shows a software window titled "Matrix of Churn by Contract". The menu bar includes "File", "Edit", "Generate", and other icons. Below the menu is a toolbar with "Matrix", "Appearance", and "Annotations" tabs, where "Matrix" is selected. The main area displays a table titled "Contract" with the following data:

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

To effectively conduct a churn analysis, it's essential to grasp the purpose behind it: comparing total revenue with the revenue lost due to churn. We estimate a 10% revenue growth from new customers, yet we still face a net revenue loss of 7.8%. This situation indicates that sustaining operations in the long run will be challenging.

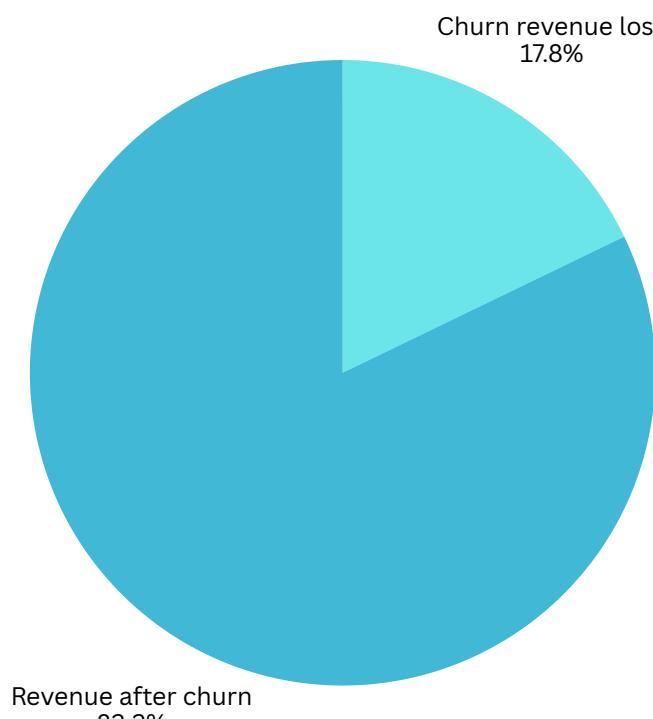


Table (3 fields, 1 records) #3

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

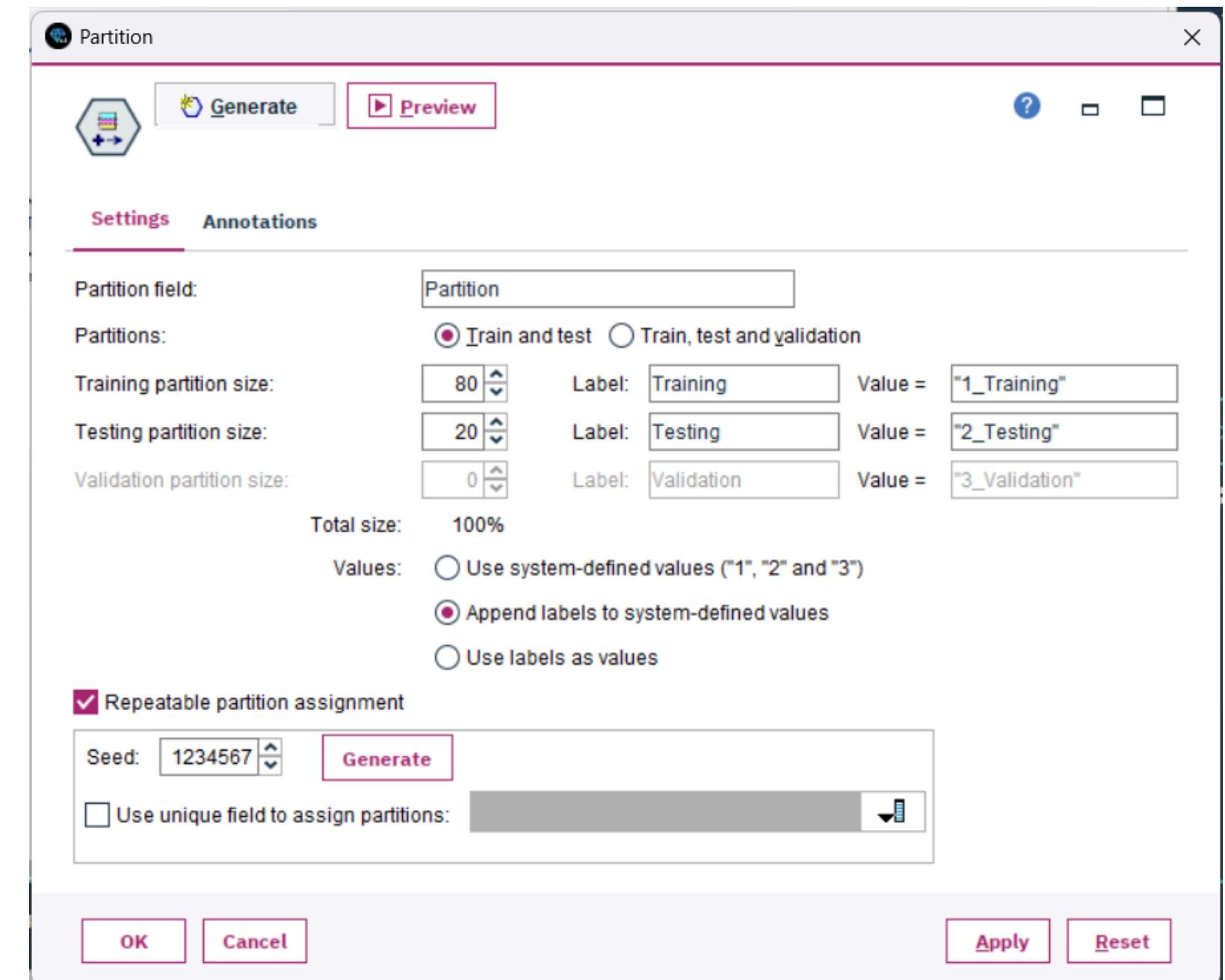
Field	Measurement	Values	Missing	Check	Role
A customerID	Typeless		None	None	<input checked="" type="checkbox"/> None
A gender	Flag	Male/Fem...	None	None	<input checked="" type="checkbox"/> Input
SeniorCitizen	Continuous	[0,1]	None	None	<input checked="" type="checkbox"/> Input
A Partner	Flag	Yes/No	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

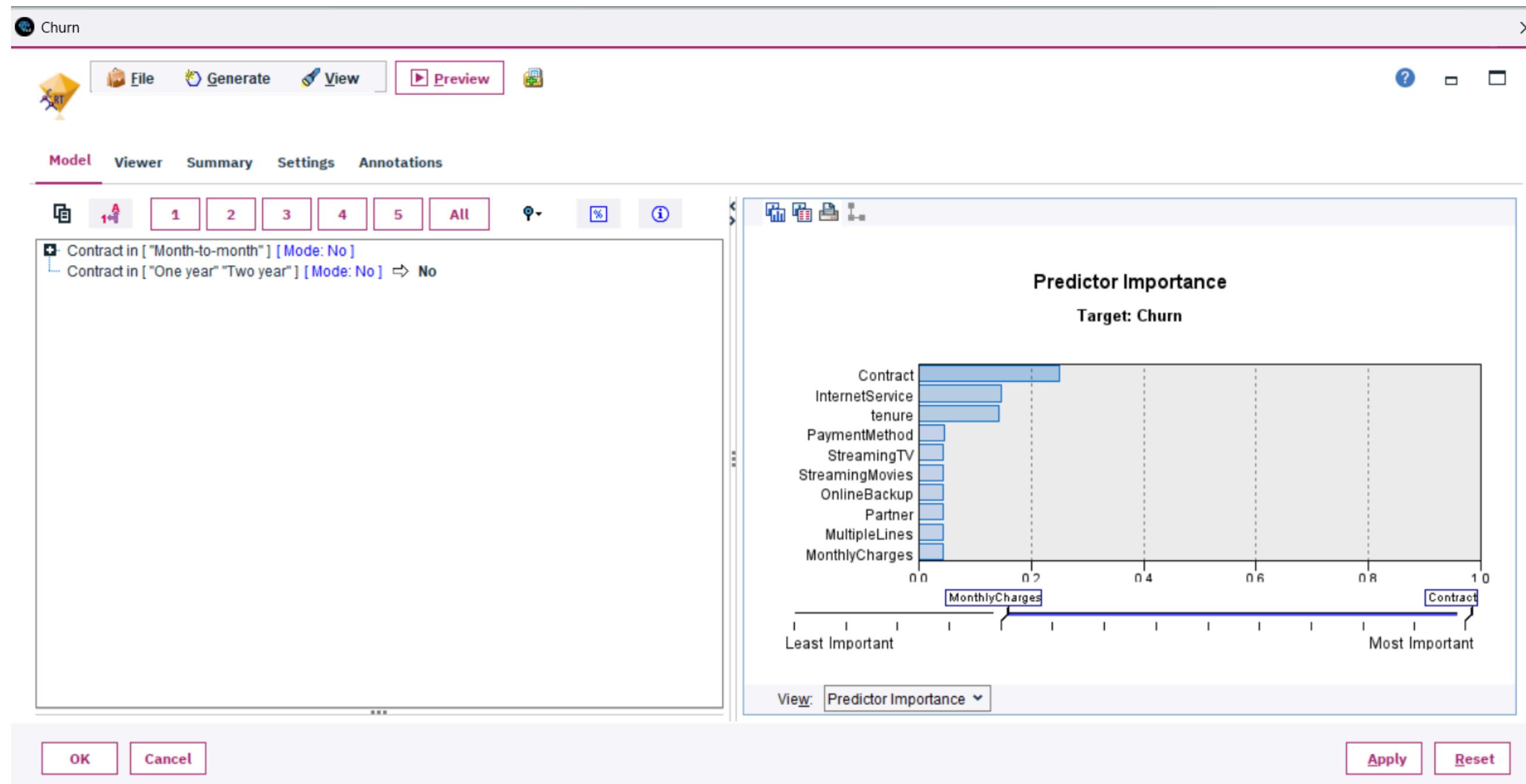
OK Cancel Apply Reset

We will now utilize the type field to categorize the different types of data fields into various classifications, including input, none, and target.

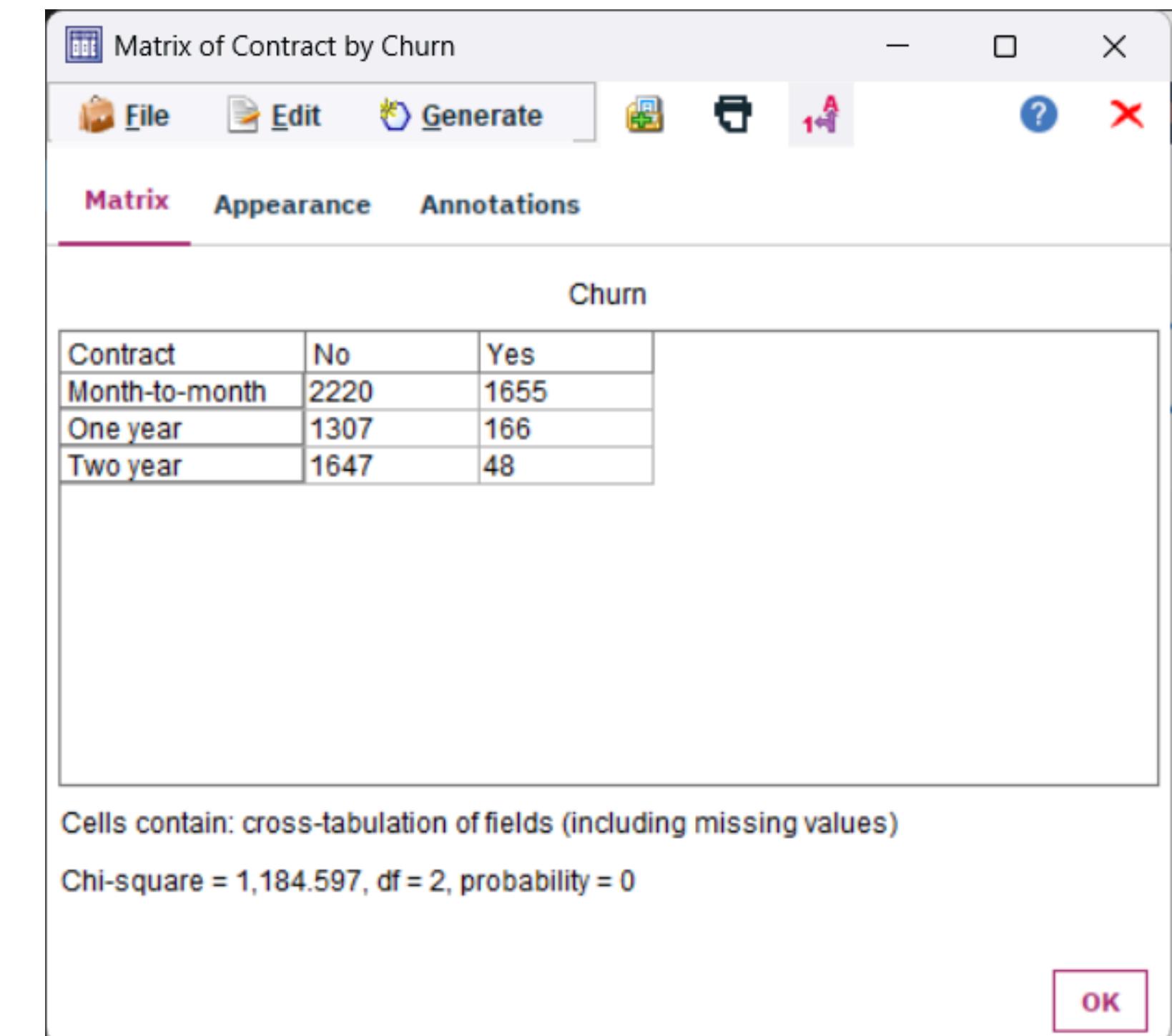
We can also use Partitioning to separate data into sets of training and testing data to test out the model accuracy.



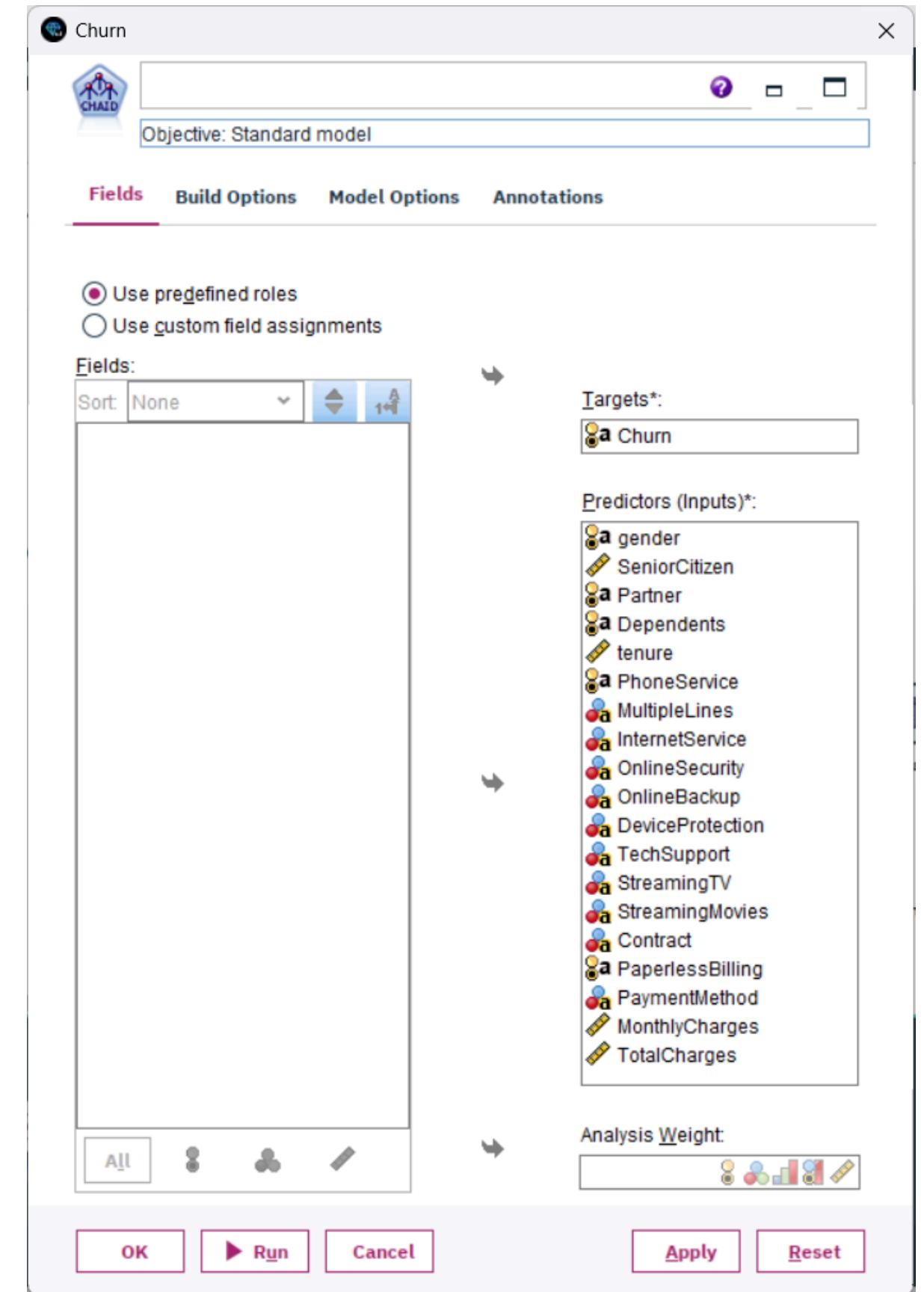
We utilize the CRT model to analyze the factors influencing customer churn and implement improvements accordingly. For instance, aspects like contract terms and internet service have a more significant impact on customer churn compared to other elements, such as monthly charges and payment methods.



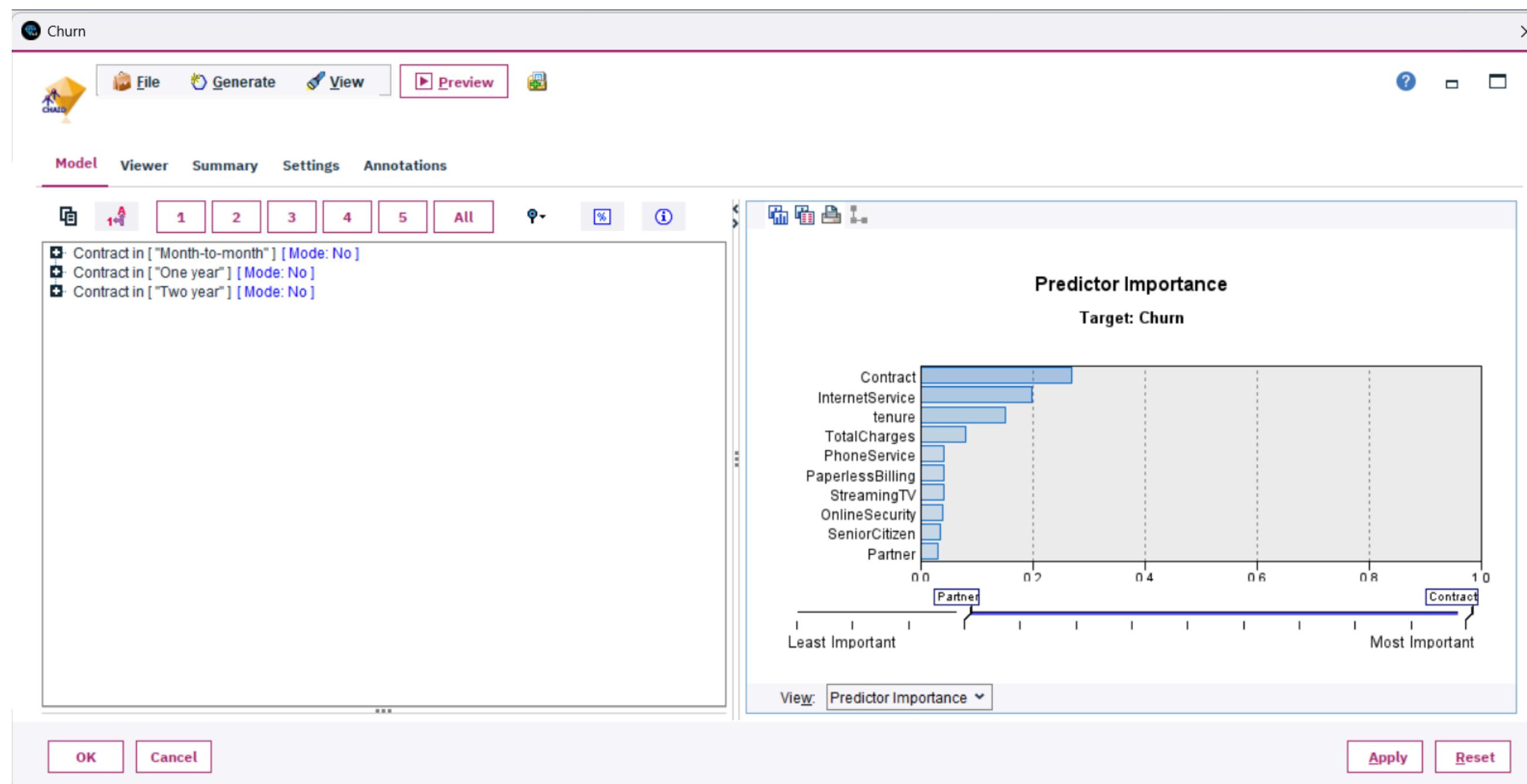
We now employ an analysis node to examine the output generated by the CRT model and evaluate various fields.



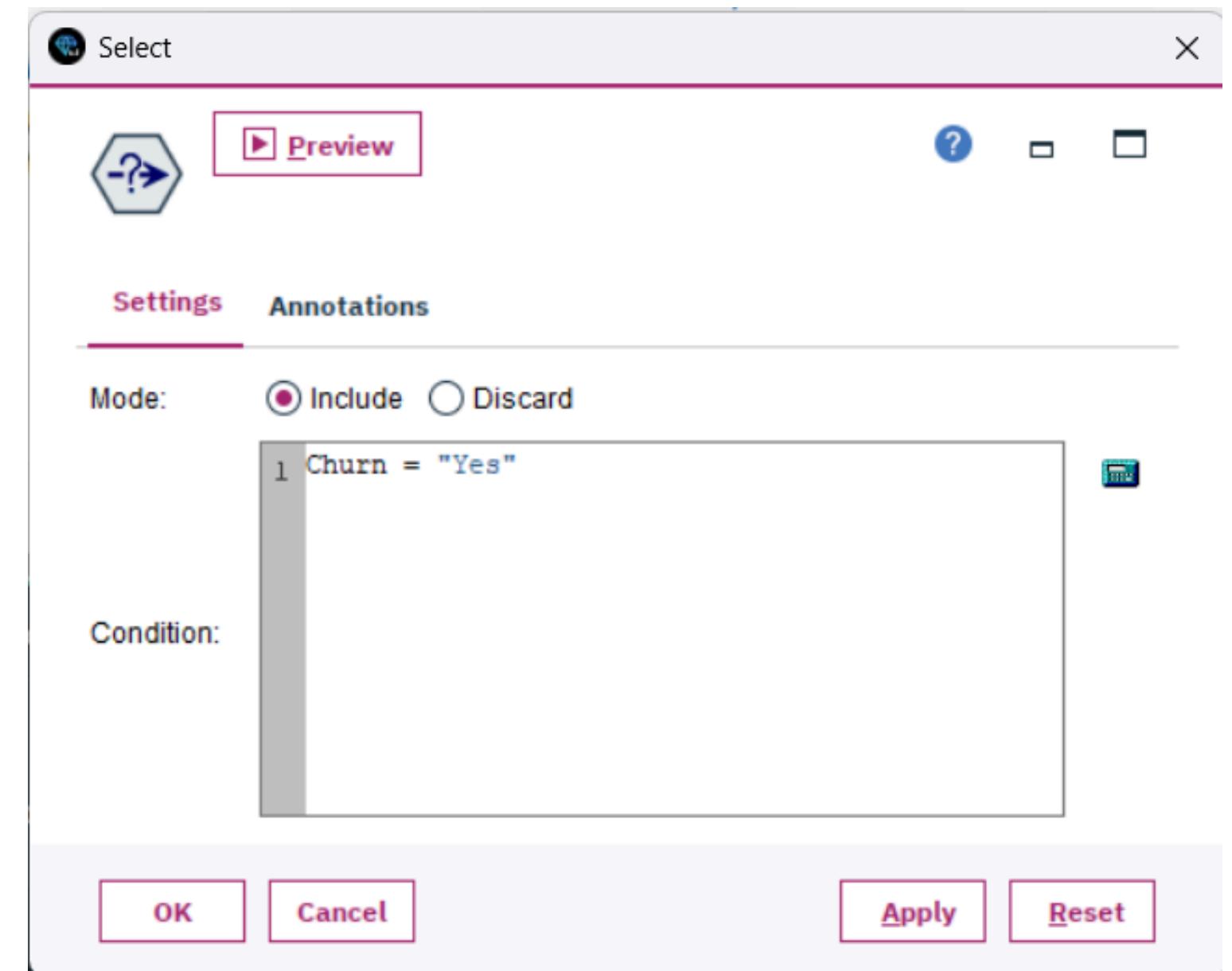
We will employ the CHAID model to enhance our analysis further; however, we first need to define the different parameters of the CHAID model within its context.



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

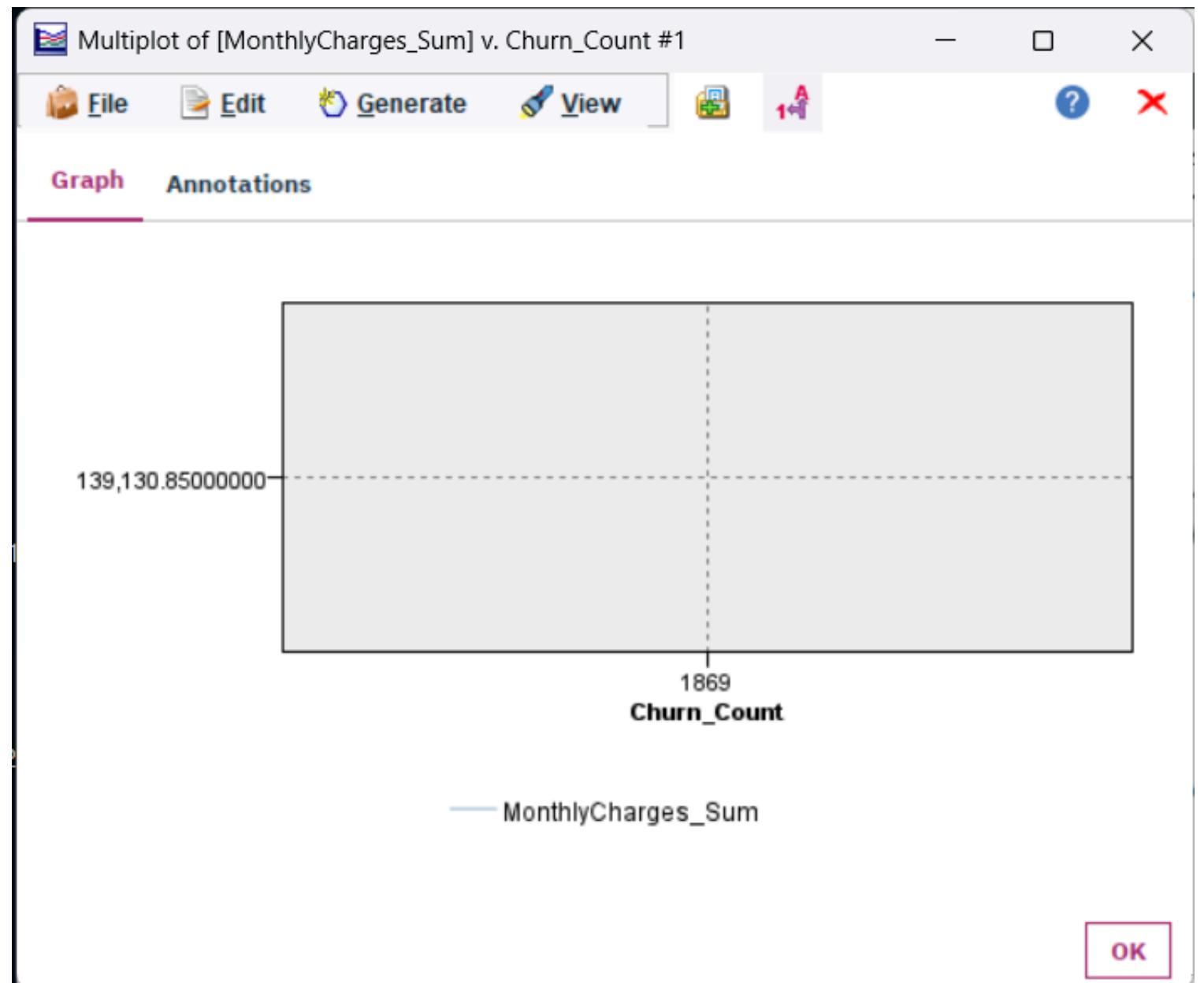
Fields: 23 in, 21 filtered, 0 renamed, 2 out

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

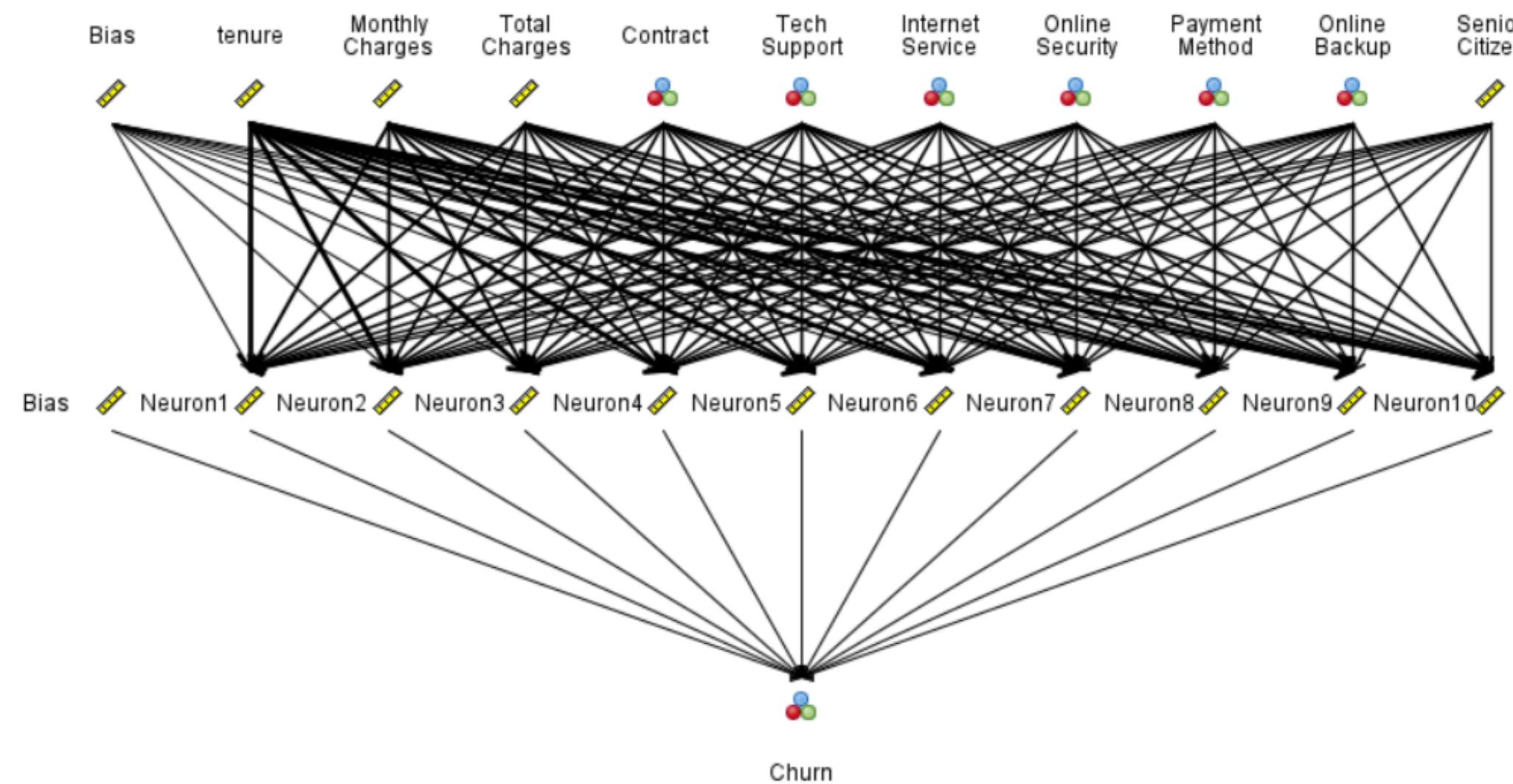
View current fields View unused field settings

OK Cancel Apply Reset

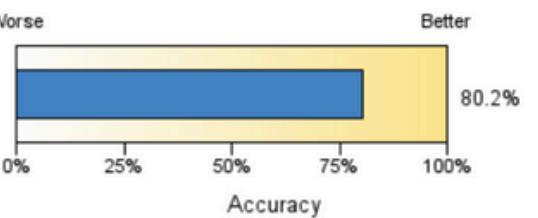
Here, we examine the overall monthly revenue loss resulting from customer churn. By adjusting the parameters in the previous two nodes, we can modify the fields and gain a more comprehensive visualization.



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 generate the output from our churn prediction model analysis to a designated file. This will enable us to review and share insights with stakeholders, providing a clearer understanding of customer behavior. The findings from this analysis are crucial for identifying patterns and trends that can refine our strategies to enhance customer retention. By analyzing the data, we can identify the factors leading to churn and create targeted interventions to boost customer satisfaction and loyalty. Additionally, this output lays the groundwork for future iterations of the model, facilitating continuous improvement and more precise predictions.

