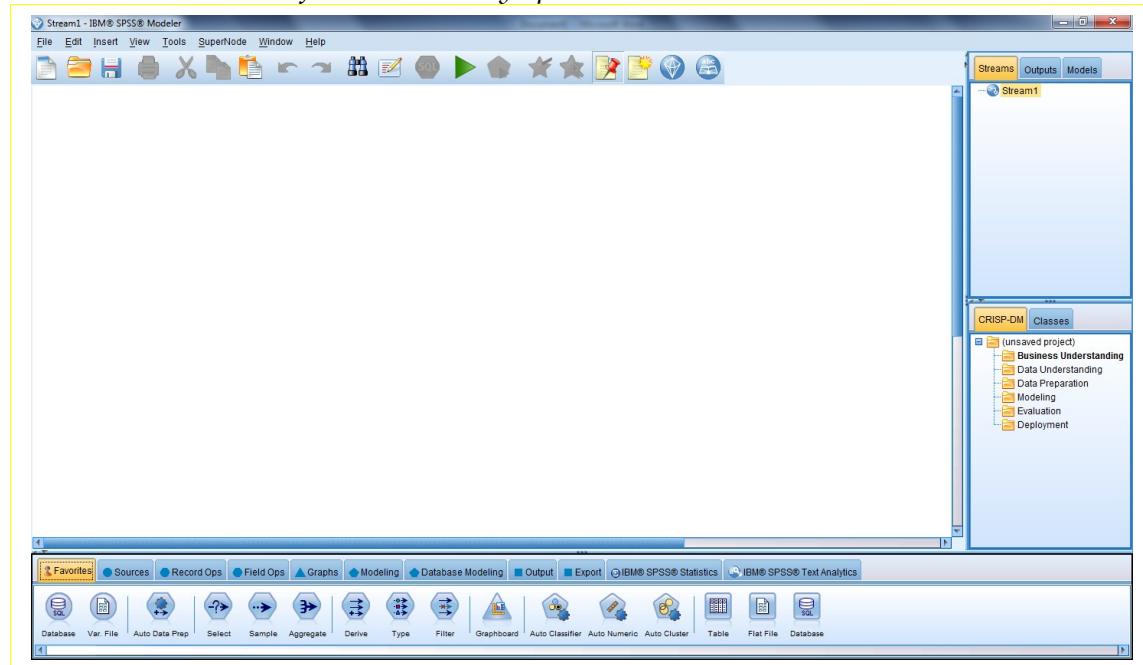


*SPSS MODELER 1 day Hands-on Labs*

# Lab 0 – Navigating IBM SPSS Modeler

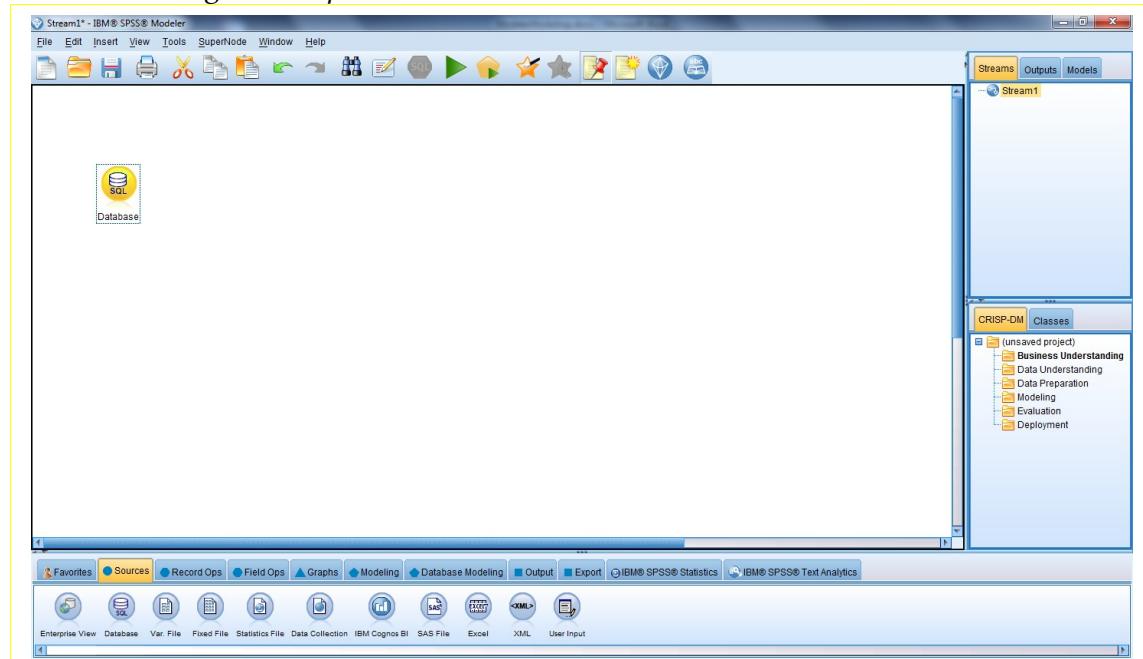
- 1) Start Modeler if it's not already open.



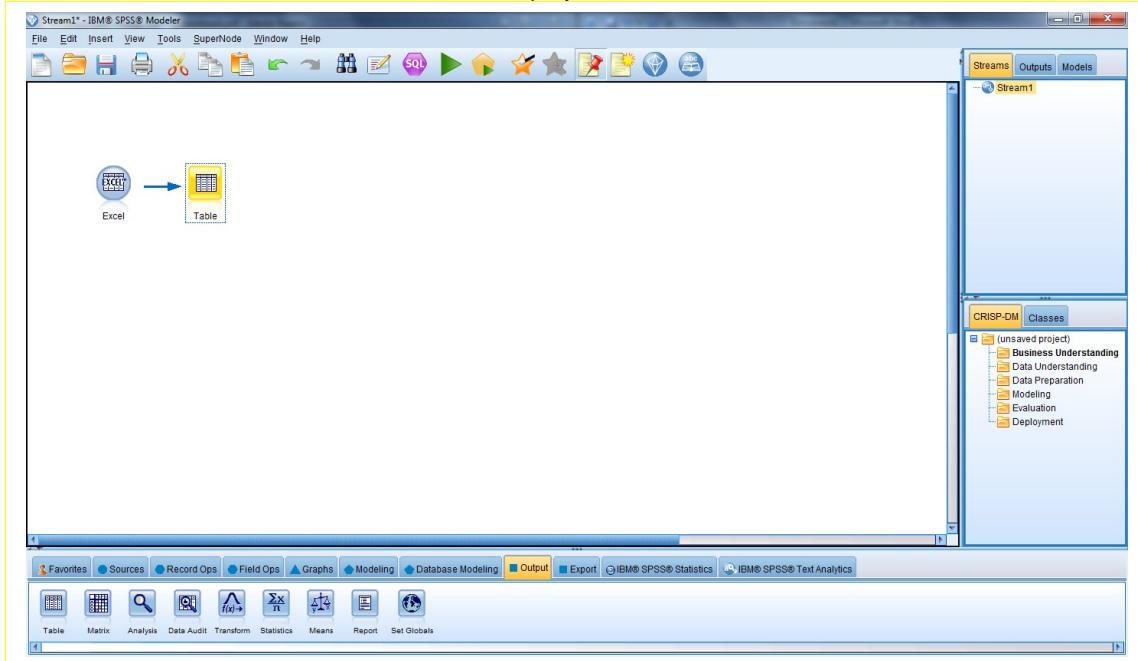
- 2) Review the various **palettes** near the bottom of the screen. The names of the **nodes** reflect the operation or function accomplished with that node.



- 3) From the **Sources** palette, select the **Database** node and add it to the canvas. This can be done by the following methods:
  1. Double-click the node on the palette.
  2. Drag and drop the node onto the canvas.



- 4) From the **Output** palette, select the **Table** node and add it to the canvas. Then join it to the Database node by the following methods:
1. Double-click the node to automatically add it to the stream and join it to the selected node.
  2. Drag and drop the node to the canvas. Select the first node, right click and select Connect from the context menu, and then left click on the second node.
  3. Click and hold the middle mouse button on the first node, move the cursor to the second node and release when the cursor is on top of the second node.

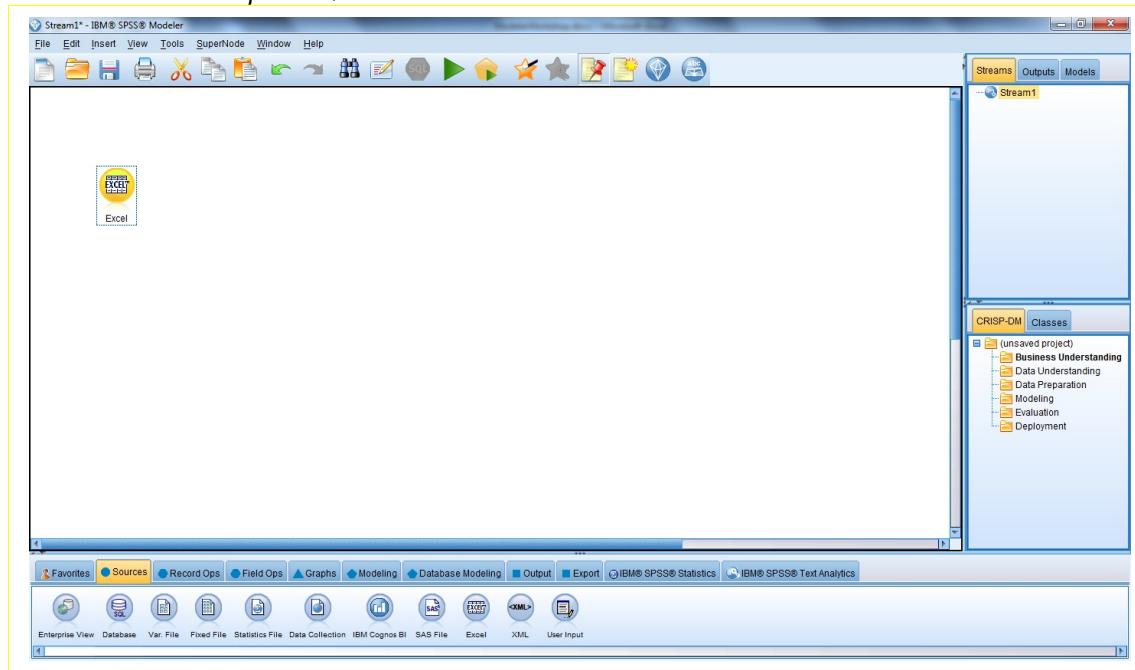


- 5) To delete the connection, right click on the joining arrow and select Delete Connection. The other methods listed in 4) can also be used to do this.  
Delete the connection.
- 6) To remove a node, select it and press the Delete button, or right click on the node and select Delete from the context menu.

## Lab 1 – Predictive In 30 Minutes

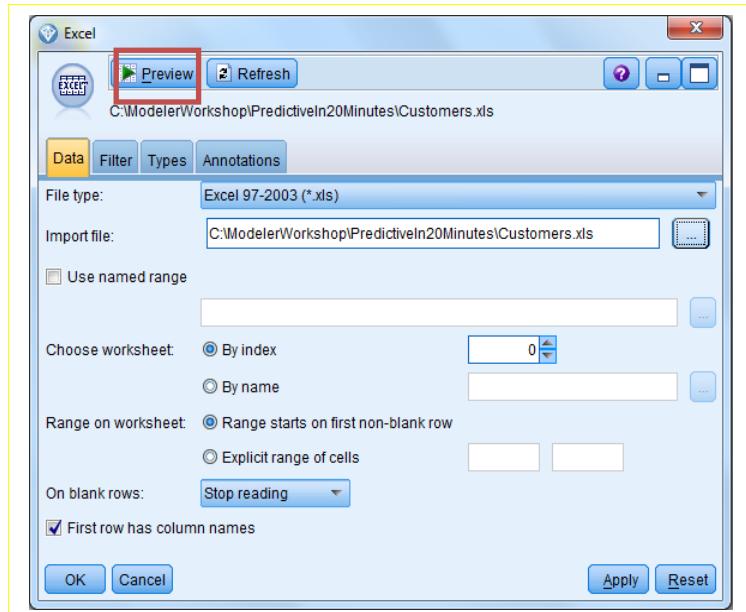
- 1) Start a new stream or remove all the nodes from the current stream.

From the **Sources** palette, add an Excel node.



- 2) Double-click the **Excel** node to open a dialog box. Use the Data tab to import the *Customers.xls* file from : C:\IDR Dwork\SPSS\SPSS Modeler1 Day\01 – PredictiveIn30Minutes\Customers.xls.

And click on the **Preview** button



- 3) After you clicked on the Preview button at the top of the dialog box to see the first 10 records in the file. This is an extraction of data from a telecommunications company's CRM system and the dataset includes customers' ID, gender, marital status and information about their services.

Preview from Customers.xls Node (17 fields, 10 records) #2

**File Edit Generate Table Annotations**

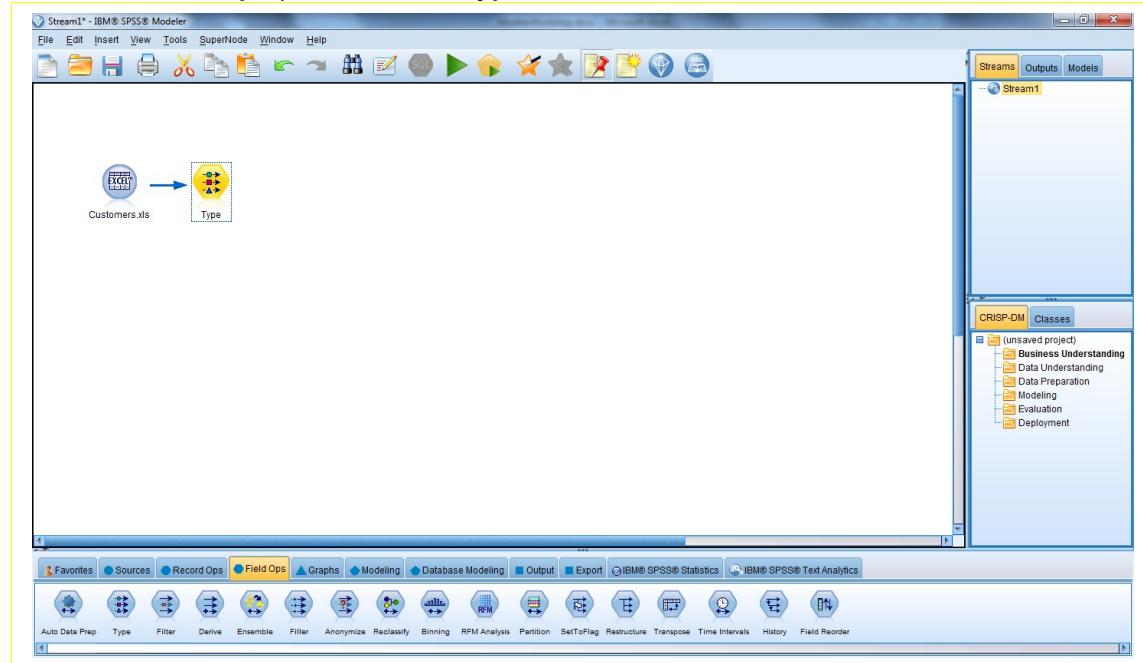
**Table**

**Annotations**

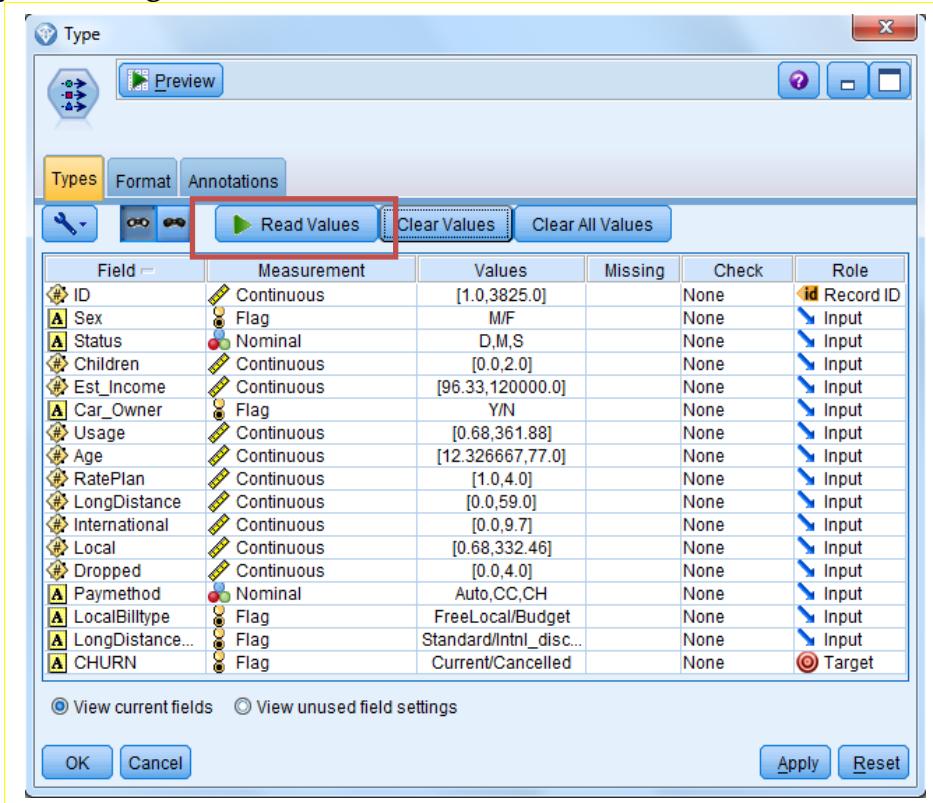
ID	Sex	Status	Children	Est_Income	Car_Owner	Usage	Age	RatePlan	LongDistance	International	Local	Dropped	Paymethod	LocalBillType	LongDistanceBillType	CHURN	
1	1.00	F	S	1.00	38000.00	N	229.64	24.39	3.00	23.56	0.00	206.08	0.00	CC	Budget	IntrL_discount	Cancelled
2	6.00	M		2.00	29516.00	N	75.29	49.43	2.00	29.78	0.00	45.50	0.00	CH	FreeLocal	Standard	Current
3	8.00	M		0.00	19732.80	N	47.25	50.67	3.00	24.41	0.00	22.44	0.00	CC	FreeLocal	Standard	Current
4	11...	M	S	2.00	96.33	N	59.01	56.47	1.00	26.13	0.00	32.88	1.00	CC	Budget	Standard	Current
5	14...	F	M	2.00	52004.80	N	28.14	25.14	1.00	5.03	0.00	23.11	0.00	CH	Budget	IntrL_discount	Cancelled
6	17...	M	M	2.00	53010.80	N	58.87	18.84	1.00	12.45	0.00	46.42	4.00	CC	FreeLocal	Standard	Current
7	18...	M	M	1.00	75004.50	N	58.72	64.80	1.00	26.52	0.00	32.19	0.00	CC	Budget	IntrL_discount	Current
8	21...	M	M	0.00	19749.30	N	34.17	60.37	3.00	20.22	0.00	13.94	0.00	CC	Budget	Standard	Current
9	22...	M	S	1.00	57626.90	Y	48.35	43.91	2.00	9.38	0.00	38.96	0.00	CC	Budget	Standard	Current
10	23...	M	M	2.00	20078.00	N	15.98	32.85	4.00	9.65	0.00	6.33	0.00	CC	Budget	IntrL_discount	Current

**OK**

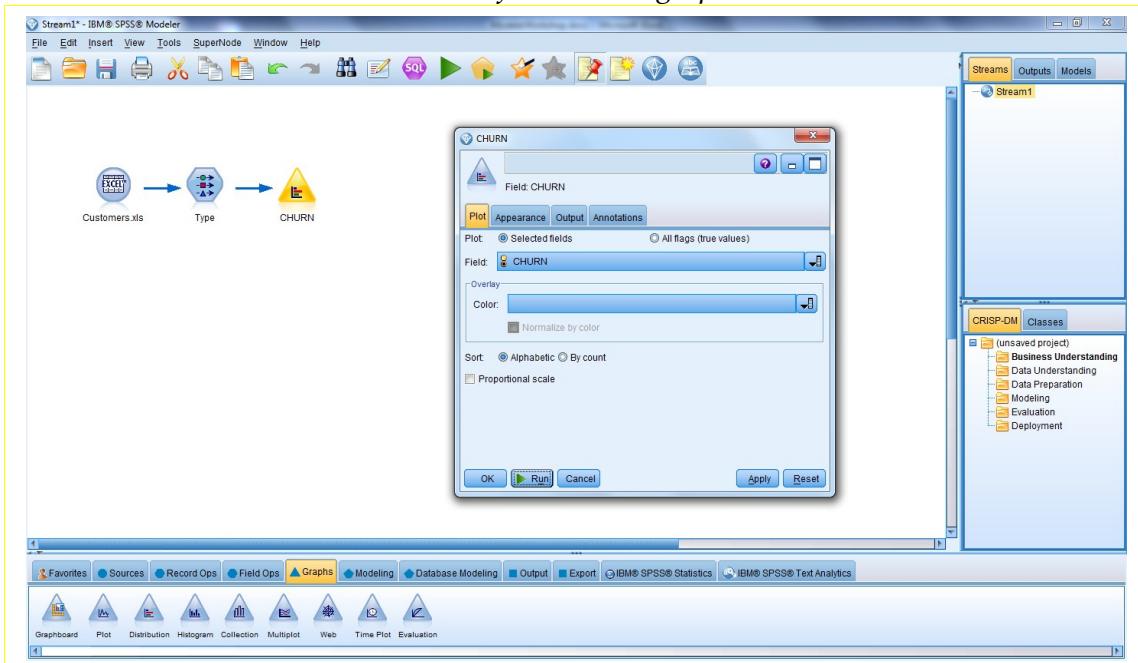
- 4) From the **Field Ops** palette, add a **Type** node to the canvas and connect it to the data source.



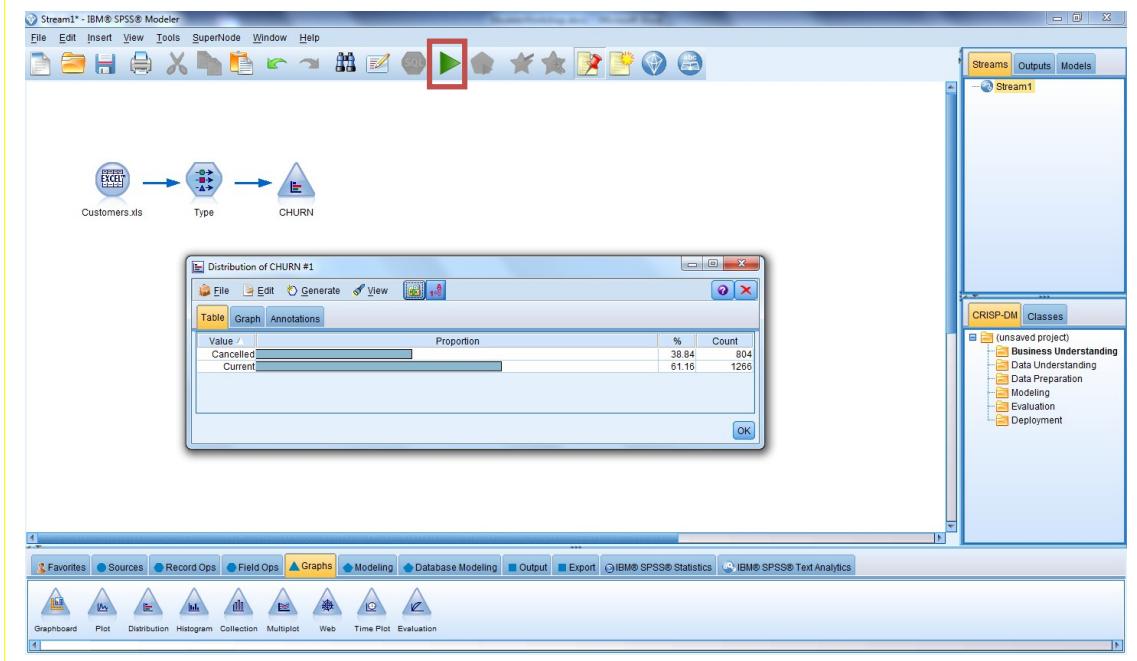
- 5) Open the **Type** node and click the **Read Values** button to scan the data as well as to display and update the range of values. Modify the **Role** of the **ID** field to **Record ID** and **CHURN** field to **Target**.



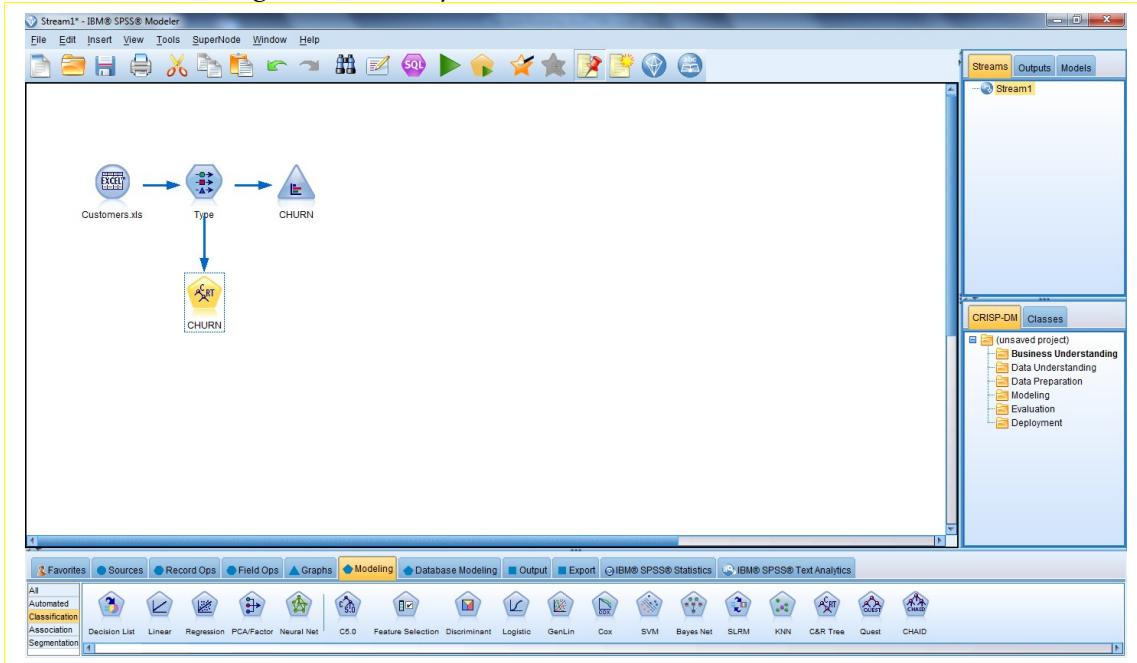
- 6) From the **Graphs** palette, add the **Distribution** node and connect it to the **Type** node. Edit the Distribution node so that the **CHURN** field will be graphed.



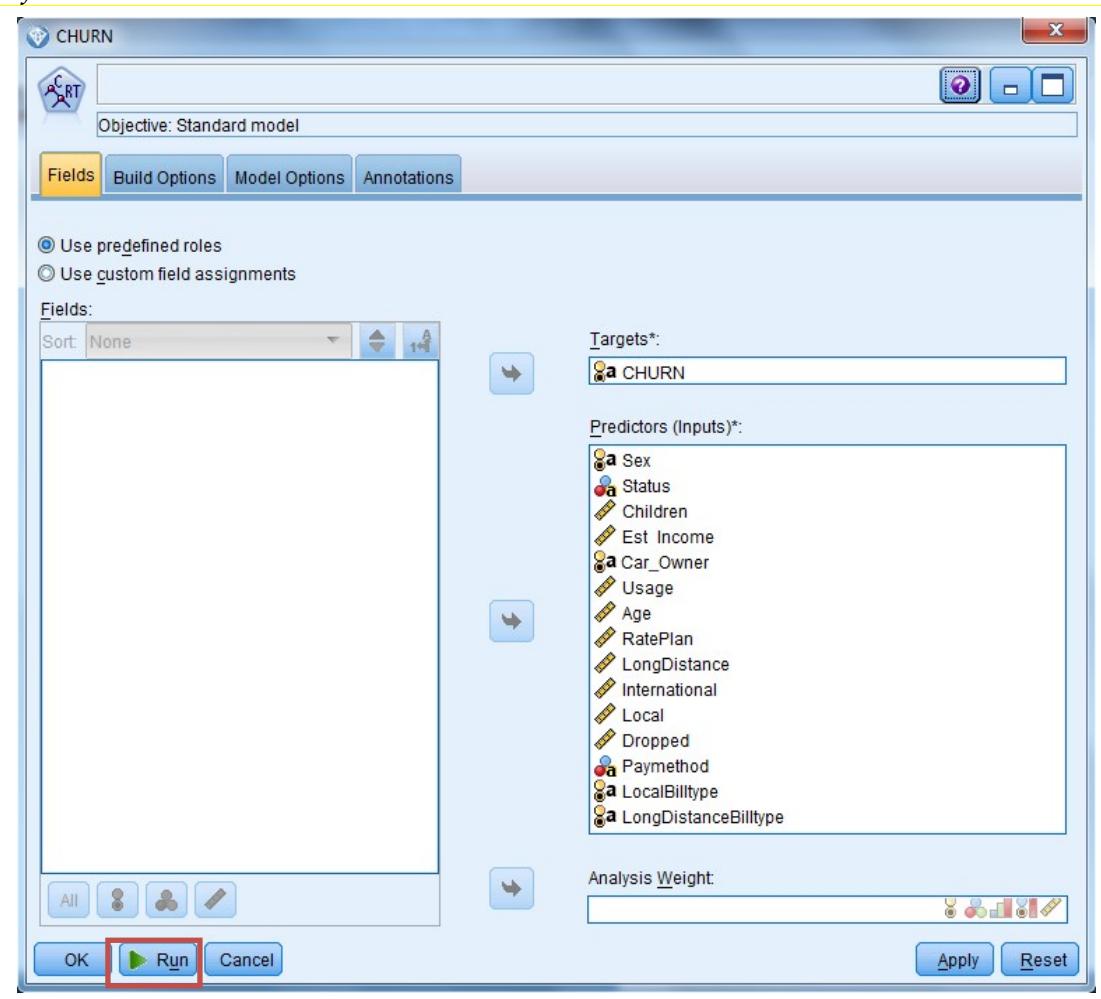
- 7) Run the stream to produce the graph. To do this, right click on the node and select Run from the context menu, or click the play button in the toolbar near the top of the screen.



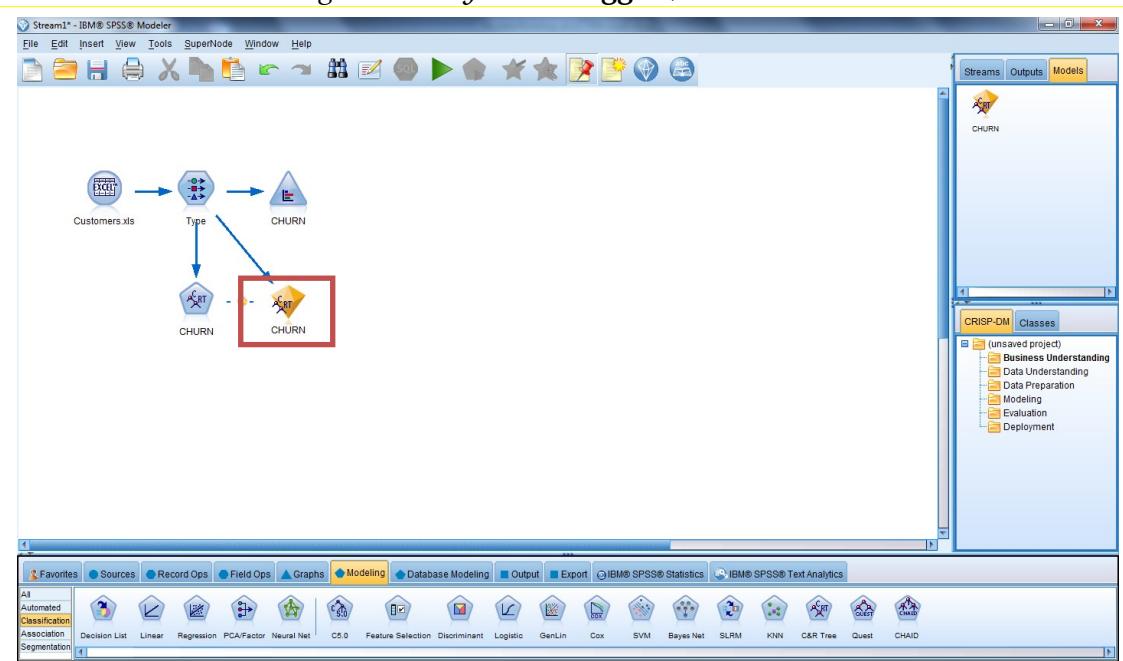
- 8) Of the 2,070 customers in this dataset, 38.84% have cancelled their contract and the remaining 61.16% are current. The task is to build a model to understand the relationships within the data that led to the canceled their contracts. A C&R Tree is needed to do this.
- 9) From the **Modeling** palette, add the **C&R** node to the Type node. Note that the C&R node name changes to CHURN when it is connected to the Type node. This is because we defined CHURN as the Target Role in Step 6.



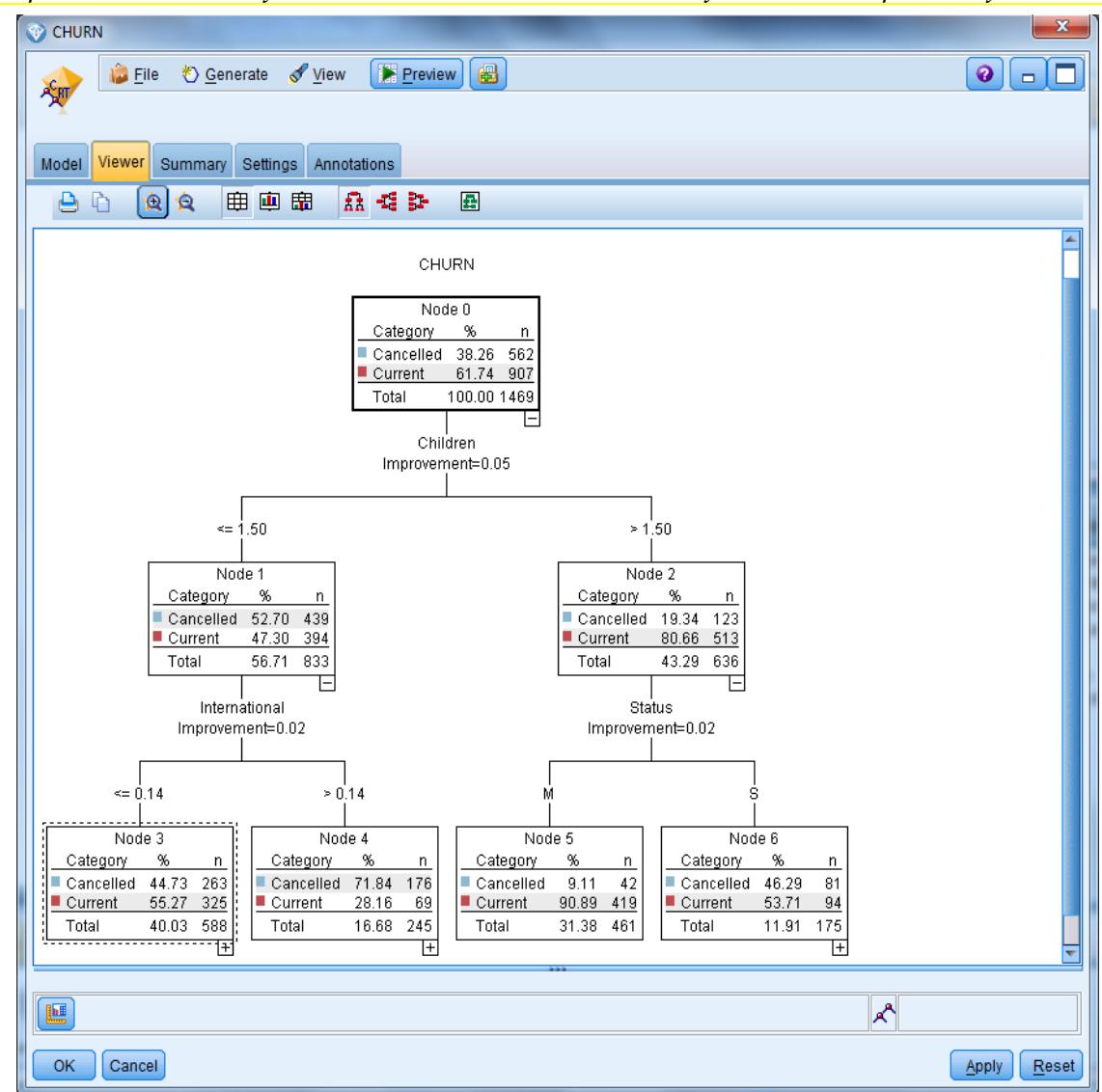
10) Open the C&R Tree node to view the settings before running the model. This example uses the defaults so click Run.



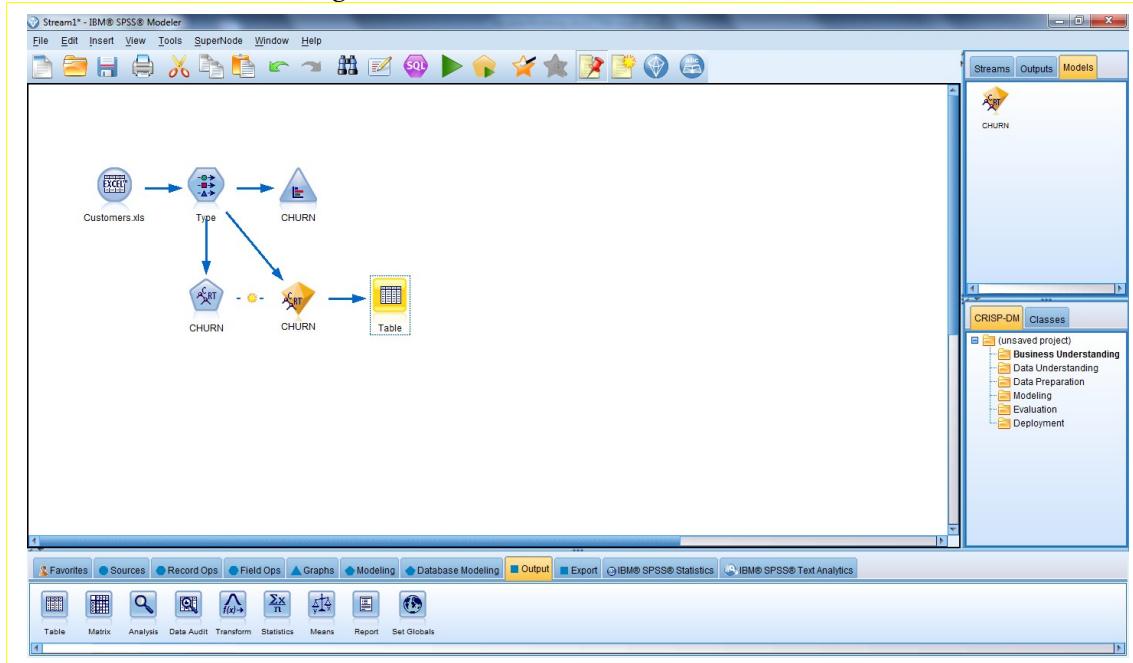
11) The C&R Tree model is generated (yellow nugget), added to the canvas and named CHURN.



12) Double-click on the generated model to see the outputs. The Model tab contains predictor importance and rules for the model. Click the Viewer tab for a visual depiction of the tree.



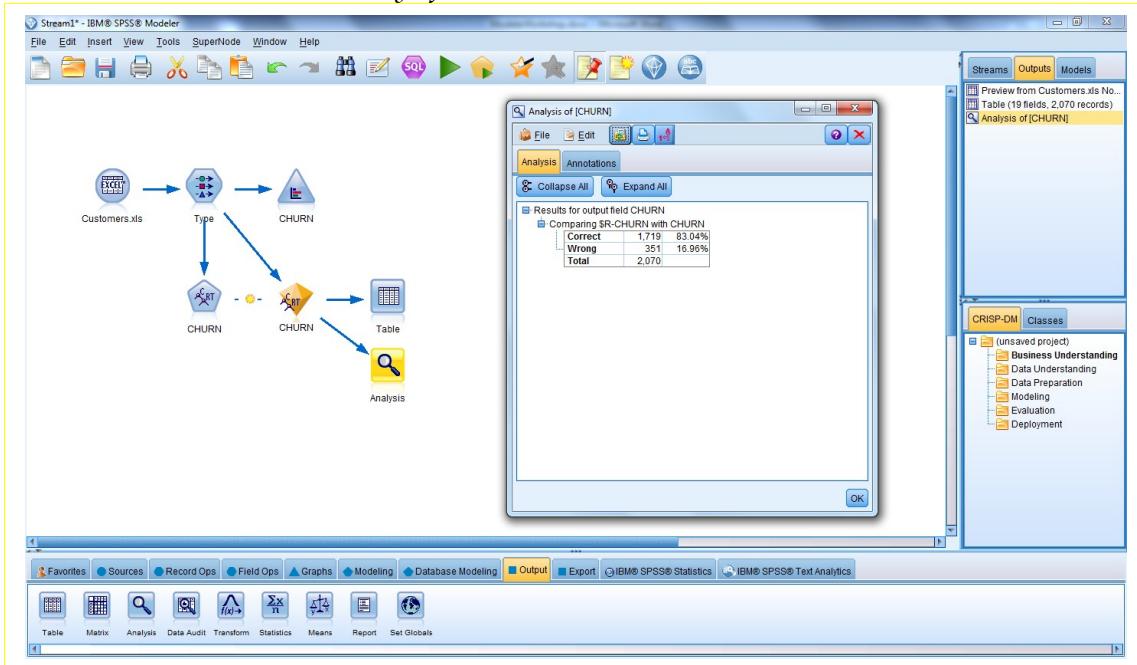
13) Add a **Table** node to the generated model and run it.



14) Look at the last three columns in the table. The third from last column is the actual outcome, whether the customer cancelled or is current; the second to last column is the prediction from the model; and the last column is the confidence in the prediction.

ID	Sex	Status	Children	Est_Income	Car_Owner	Usage	Age	RatePlan	LongDistance	International	Local	Dropped	Parmethod	LocalBillType	LongDistanceBillType	CHURN	\$R-CHURN	\$RC-CHURN	
1	1.00	F	S	1.00	38000.00	N	229.64	24.39	3.00	23.58	0.00	208.08	0.00	CC	Budget	Intl_discount	Cancelled	Cancelled	0.83
2	6.00	M	M	2.00	29816.00	N	75.29	49.43	2.00	29.78	0.00	45.50	0.00	CH	FreeLocal	Standard	Current	Current	0.91
3	8.00	M	M	0.00	19732.80	N	47.25	50.67	3.00	24.81	0.00	22.44	0.00	CC	FreeLocal	Standard	Current	Current	0.96
4	11..	M	S	2.00	96.33	N	59.01	56.47	1.00	26.13	0.00	32.88	1.00	CC	Budget	Standard	Current	Current	0.73
5	14..	F	M	2.00	52004.80	N	28.14	25.14	1.00	5.03	0.00	23.11	0.00	CH	Budget	Intl_discount	Cancelled	Current	0.91
6	17..	M	M	2.00	53010.80	N	58.87	18.84	1.00	12.45	0.00	46.42	4.00	CC	FreeLocal	Standard	Current	Current	0.91
7	18..	M	M	1.00	75004.50	N	58.72	64.80	1.00	26.52	0.00	32.19	0.00	CC	Budget	Intl_discount	Current	Cancelled	0.70
8	21..	M	M	0.00	19749.30	N	34.17	60.37	3.00	20.22	0.00	13.94	0.00	CC	Budget	Standard	Current	Current	0.96
9	22..	M	S	1.00	57626.90	Y	48.35	43.91	2.00	9.38	0.00	38.96	0.00	CC	Budget	Standard	Current	Cancelled	0.70
10	23..	M	M	2.00	20078.00	N	15.98	32.85	4.00	9.65	0.00	6.33	0.00	CC	Budget	Intl_discount	Current	Current	0.91
11	24..	M	M	2.00	47200.00	N	72.31	55.03	2.00	17.44	0.00	4.94	0.00	CH	FreeLocal	Standard	Current	Current	0.91
12	29..	M	M	1.00	7545.96	Y	206.75	16.75	3.00	22.39	0.00	178.36	0.00	CC	Budget	Standard	Cancelled	Cancelled	0.77
13	35..	F	S	0.00	78951.30	N	29.04	49.37	4.00	0.37	0.00	28.66	0.00	CC	FreeLocal	Standard	Current	Current	0.87
14	36..	F	S	1.00	17540.70	Y	36.20	62.79	4.00	22.17	0.57	13.45	0.00	Auto	Budget	Standard	Cancelled	Cancelled	0.82
15	37..	F	M	0.00	83891.90	Y	74.40	61.02	4.00	28.92	0.00	45.47	0.00	CH	Budget	Standard	Cancelled	Current	0.87
16	38..	F	M	2.00	28220.80	N	38.95	38.77	4.00	26.49	0.00	12.45	0.00	CC	FreeLocal	Standard	Cancelled	Current	0.91
17	40..	F	S	0.00	28589.10	N	100.28	15.60	4.00	13.19	0.00	87.09	0.00	CC	FreeLocal	Standard	Cancelled	Cancelled	0.83
18	42..	F	M	2.00	5237.63	N	78.02	48.75	2.00	13.32	0.05	56.64	0.00	CC	Budget	Standard	Current	Current	0.91
19	45..	M	S	2.00	89459.90	N	36.05	53.28	2.00	11.54	1.61	22.90	0.00	CC	FreeLocal	Standard	Cancelled	Cancelled	0.79
20	48..	F	S	1.00	13576.50	N	40.49	39.43	1.00	14.83	0.00	25.66	0.00	CC	Budget	Standard	Cancelled	Cancelled	0.83
21	52..	F	M	2.00	67388.00	N	96.33	53.12	3.00	4.79	0.50	91.04	1.00	CH	Budget	Standard	Current	Current	0.91
22	53..	F	M	1.00	57063.00	Y	98.10	52.33	4.00	16.79	0.00	81.30	0.00	CH	Budget	Standard	Current	Current	0.87
23	54..	F	M	2.00	8479.10	N	100.27	54.17	2.00	3.05	0.00	0.00	0.00	CH	FreeLocal	Standard	Cancelled	Current	0.91
24	69..	F	S	1.00	7379.41	N	61.80	42.67	4.00	11.65	0.00	50.84	0.00	CC	FreeLocal	Standard	Cancelled	Cancelled	0.93
25	61..	M	S	2.00	100020.00	N	314.62	50.00	4.00	21.37	0.00	293.24	0.00	CH	Budget	Standard	Current	Cancelled	0.55
26	62..	F	M	2.00	45287.60	Y	2.97	29.03	3.00	0.00	0.00	2.97	0.00	CC	Budget	Standard	Current	Current	0.91
27	63..	F	M	2.00	59613.10	N	175.78	34.32	1.00	12.37	0.00	163.41	0.00	CC	Budget	Intl_discount	Current	Current	0.91
28	65..	F	S	1.00	16326.70	Y	1.34	50.61	1.00	0.00	0.00	1.34	0.00	CC	FreeLocal	Intl_discount	Cancelled	Cancelled	0.83
29	68..	F	S	2.00	95597.8	N	77.91	51.38	3.00	18.31	0.00	59.59	0.00	CC	FreeLocal	Standard	Cancelled	Cancelled	0.79
30	71..	F	M	2.00	42031.80	N	59.78	50.75	4.00	20.45	0.00	39.33	0.00	CH	Budget	Intl_discount	Current	Current	0.91
31	73..	F	M	2.00	83284.10	Y	35.96	28.42	4.00	26.91	0.00	9.04	0.00	Auto	FreeLocal	Intl_discount	Current	Current	0.91
32	74..	M	M	1.00	39229.10	N	47.10	26.26	4.00	28.82	4.74	13.53	0.00	CC	Budget	Intl_discount	Current	Cancelled	0.81
33	75..	F	S	0.00	68427.40	N	73.81	42.39	3.00	23.76	0.00	50.05	0.00	Auto	FreeLocal	Standard	Current	Current	0.87
34	77..	M	S	0.00	3891.11	N	125.60	49.47	4.00	9.57	4.19	111.84	0.00	Auto	Budget	Intl_discount	Cancelled	Cancelled	0.82
35	78..	M	M	0.00	10242.60	N	300.00	37.65	4.00	18.00	0.00	204.30	0.00	CH	FreeLocal	Intl_discount	Cancelled	Cancelled	0.52
36	79..	F	S	2.00	86716.30	N	43.13	65.66	4.00	20.97	8.23	55.92	0.00	CC	FreeLocal	Standard	Cancelled	Cancelled	0.92
37	80..	F	S	2.00	66102.70	Y	58.07	62.48	2.00	11.70	0.00	47.17	0.00	Auto	Budget	Standard	Current	Current	0.73
38	81..	F	M	2.00	84799.30	N	33.47	50.43	4.00	14.22	8.80	10.45	0.00	CH	FreeLocal	Standard	Current	Current	0.91
39	83..	F	M	2.00	35976.50	Y	50.63	49.85	4.00	18.85	0.00	31.78	0.00	Auto	FreeLocal	Intl_discount	Current	Current	0.91
40	84..	F	M	0.00	32703.10	Y	72.89	64.92	4.00	16.80	0.00	56.09	0.00	CC	Budget	Intl_discount	Current	Current	0.96
41	85..	F	M	1.00	77129.10	Y	44.85	61.85	4.00	24.77	0.00	20.07	0.00	CH	FreeLocal	Intl_discount	Current	Current	0.87

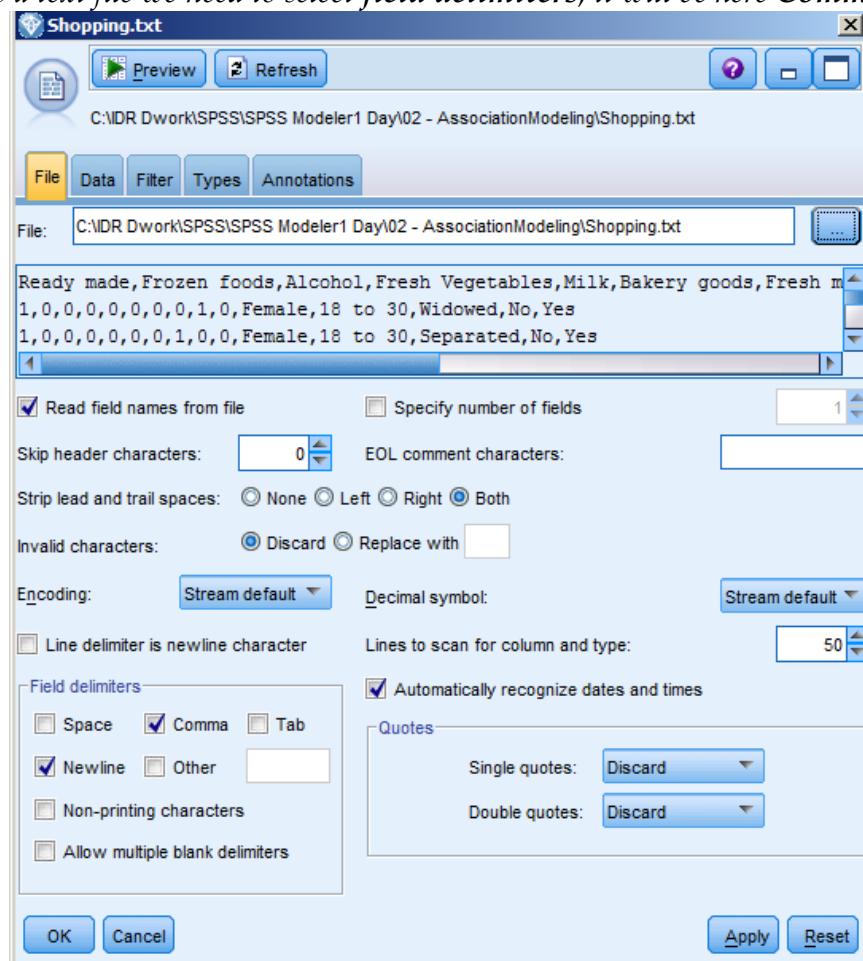
15) From the **Output** palette, add an **Analysis** node, join it to the generated model and run it. This shows the overall accuracy of the model.



16) Save the Stream to a local directory before closing it.

## Lab 2 – Association Modeling

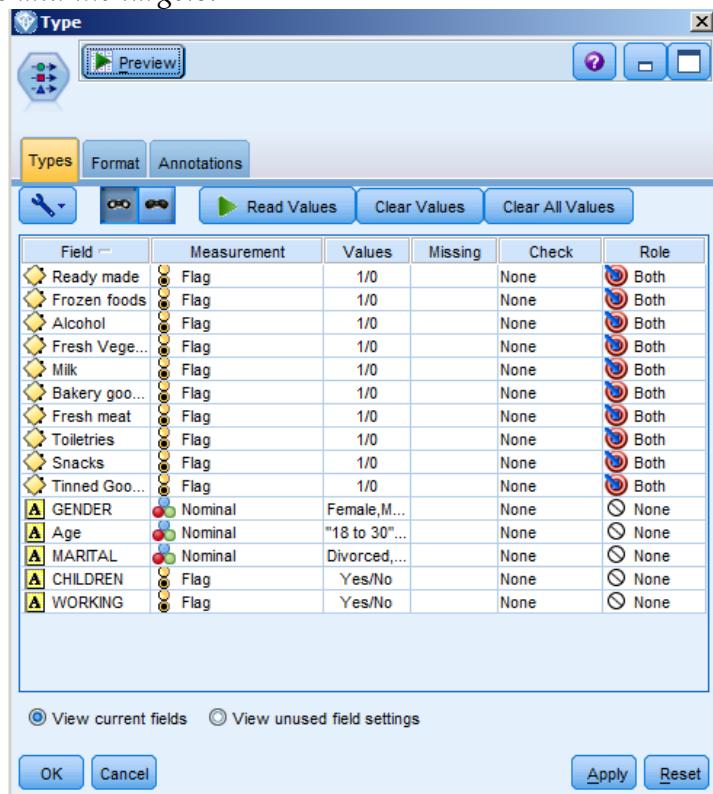
- 1) Start a new stream, From the Sources palette, add an **Var. File** node.
1. Double-click on the node to access the properties, in file click on the **3 dot** and look for the file **Shopping.txt** in **C:\IDR Dwork\SPSS\SPSS Modeler1 Day\02 – AssociationModeling** and click **open**
2. As it's a text file we need to select **field delimiters**, it will be here **Comma** and **Newline**



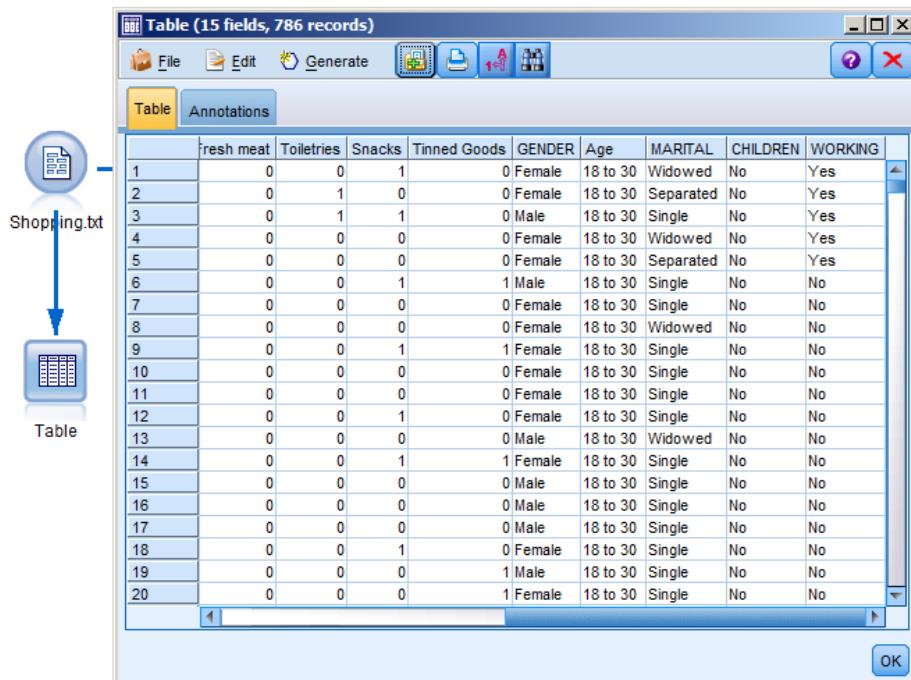
3. And click **preview** to make sure that fields delimiters are correct

2) Connect a **type** node to the **Var. file**, double click on it and click **read values**.

As we will use an Association model we will need to change roles to all the first 10 flags and set it to **both** because the task is to understand the relationship between the items being purchased, which are inputs and the targets.

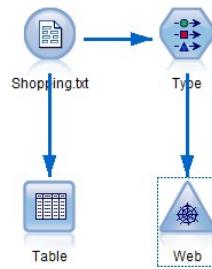


3) Now connect a **Table** node to the source file:



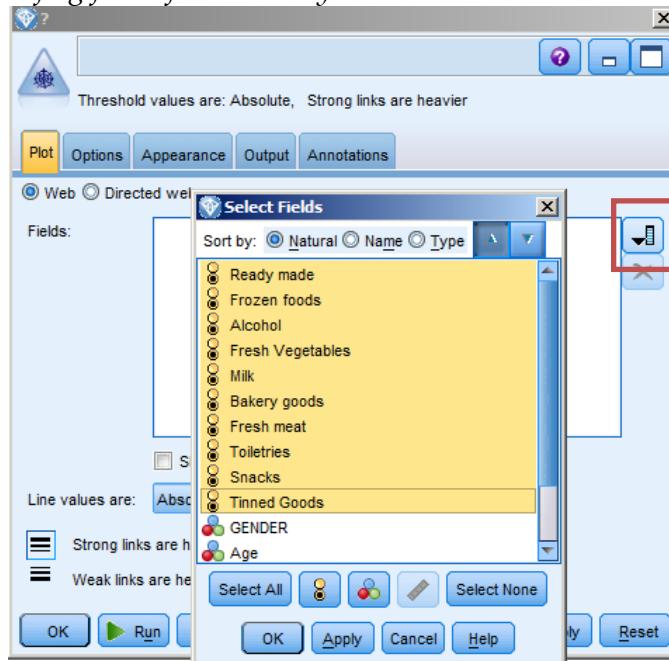
Run the Table node. In this example, the data is summarised transactional data with some demographics about the buyer. The first 10 fields are binary where 1 indicates the product was purchased and 0 indicates it was not purchased.

4) Now connect to the type node a **Web** node from the **Graphs** palette.

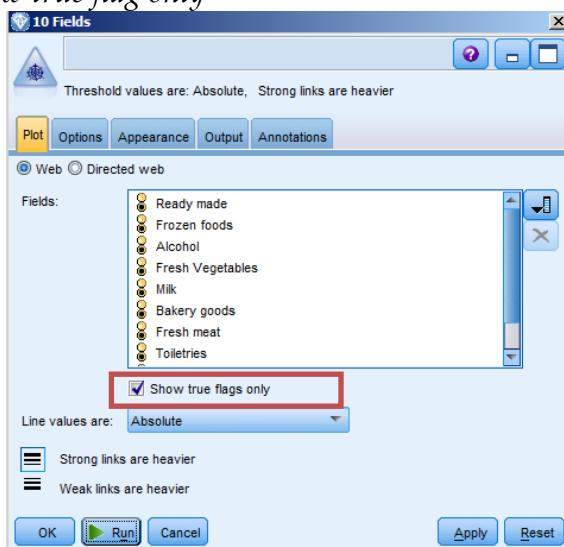


5) Double-click on it and edit the properties :

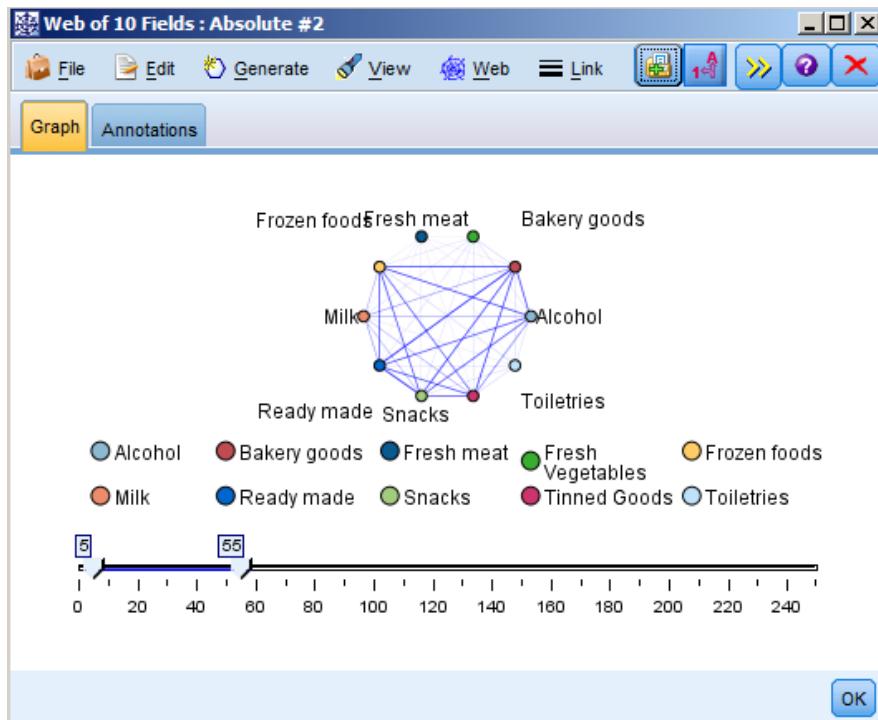
1. Add the 10 first flag fields from "Ready Made" to "Tinned Goods" and click OK



2. Tick the box "Show true flag only"

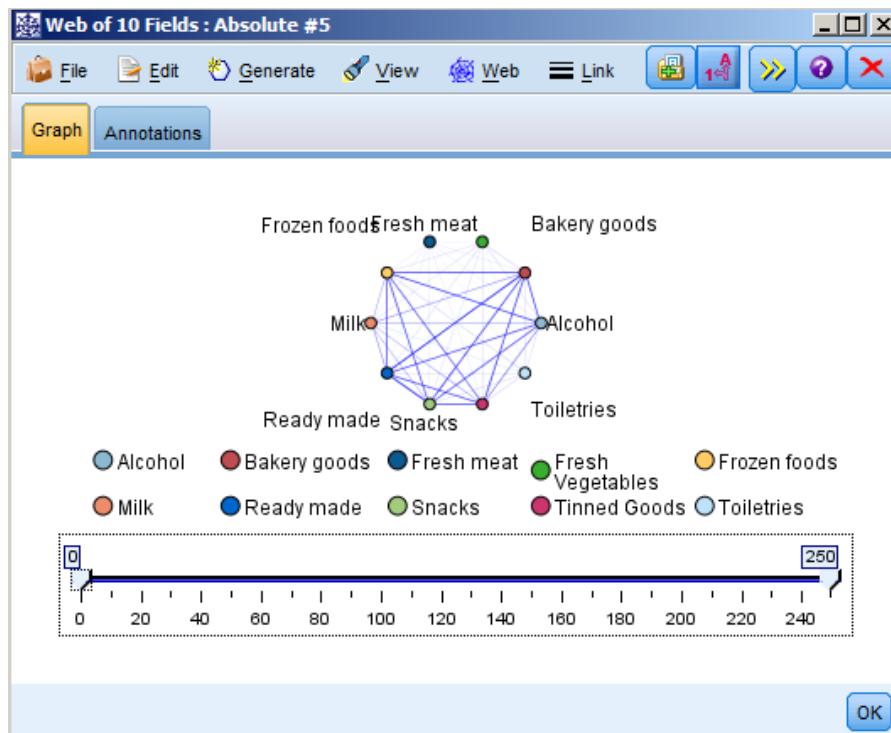


3. Click on **run**

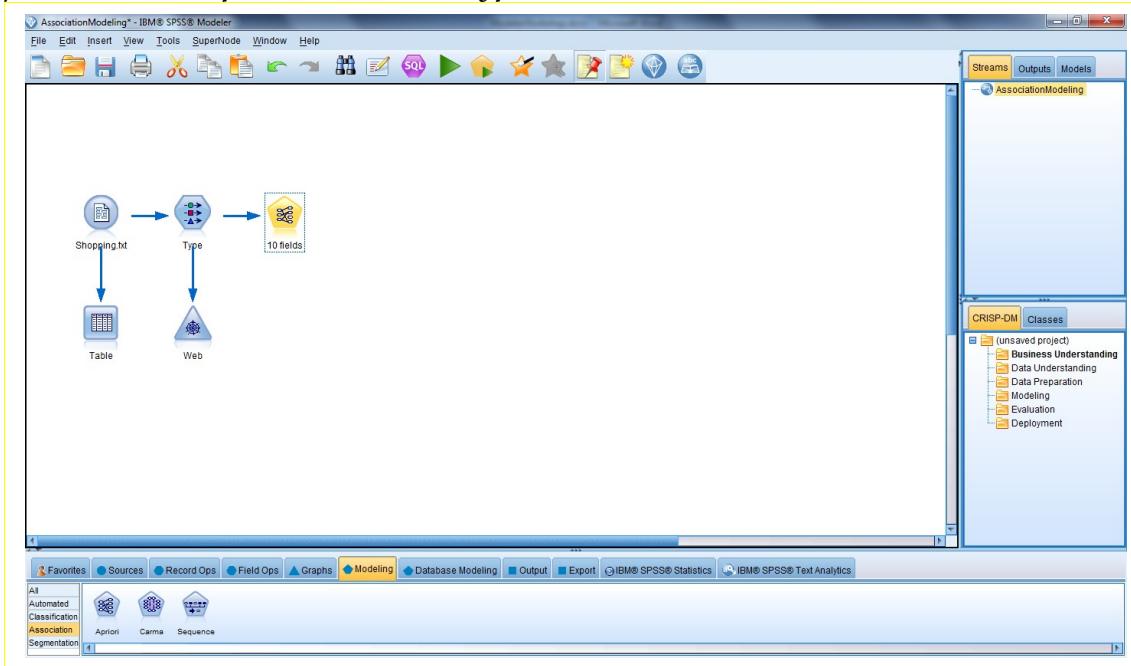


The output shows 10 items around the outside points of the graph. The lines connecting the points show the number of times the two items occurred in the same transaction.

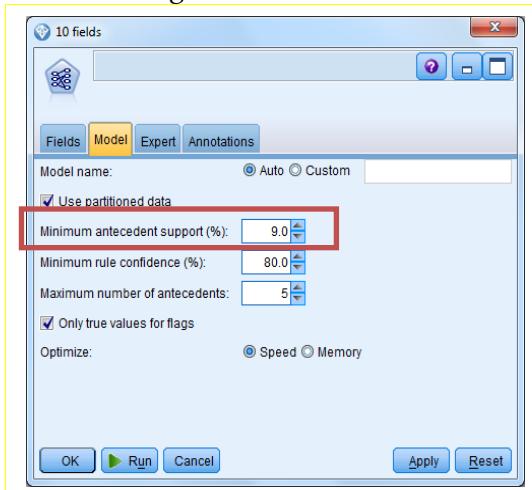
4. Use the **sliders** under the graph to limit the minimum and maximum transaction occurrences to shown in the graph.



- 6) After reviewing the visualisation, it is time to build a model to understand these relationships and to make predictions as to which item is the next likely purchase. From the **Modelling** palette, add an **Apriori** node to the **Type** node.



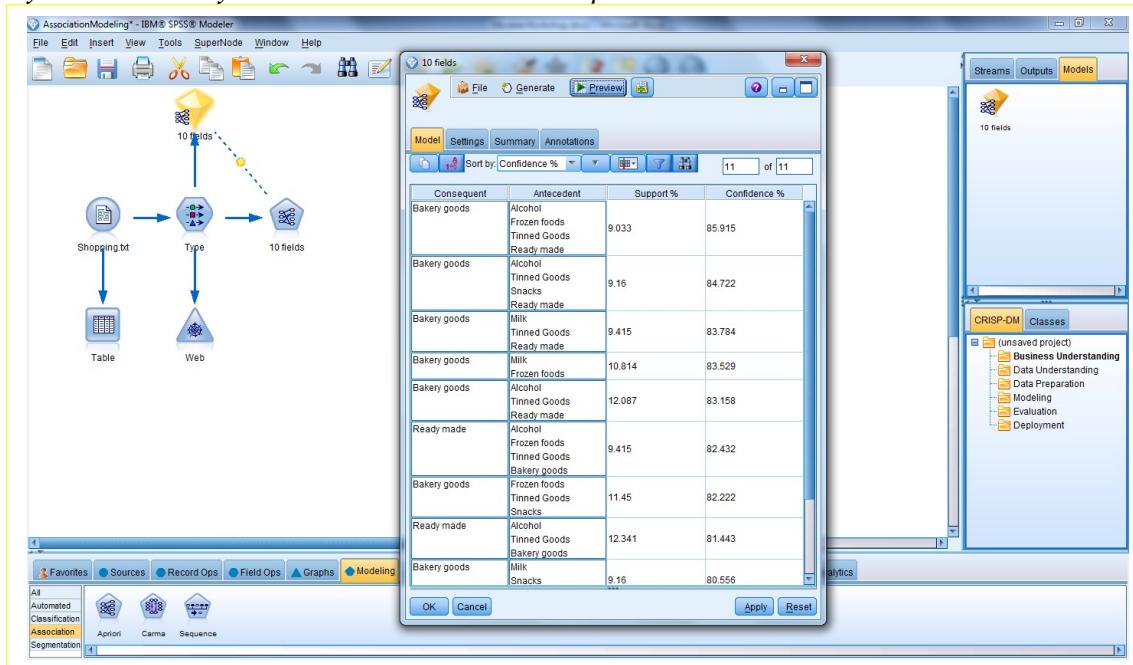
- 7) Edit the **Apriori** node settings in the Model tab, which is used to adjust the settings that control how rules are built.
1. The antecedents are the "If" parts of the rule, e.g., If Milk and Frozen Foods, then Bakery Goods. Change the Minimum antecedent support (%) to 9.0.



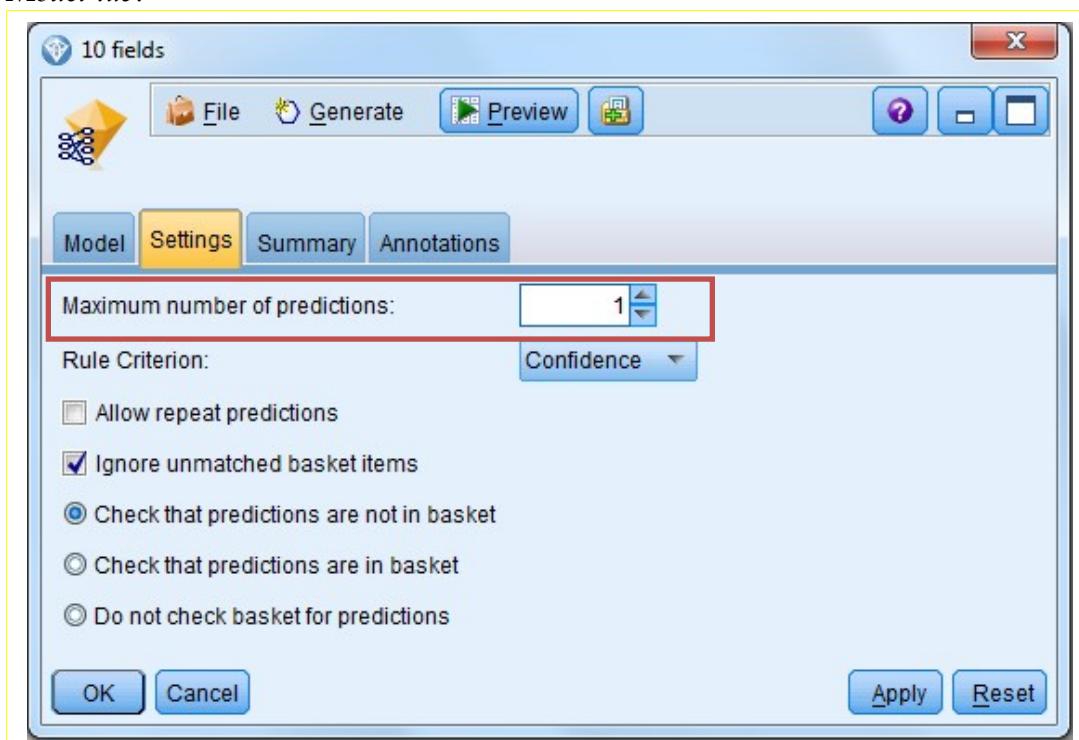
This means that only rules whose antecedents occur in at least 9% of the transactions will be displayed. (For more information, click the ? in the top right corner).

2. Click the **Run** button when finished.

- 8) Open the generated model (**yellow nugget**) to see the rules that have been found. The way to interpret this is that the antecedent occurred in the support % of the transactions. Of those, the confidence % also had the consequent.



- 9) Select the **Settings** tab and change the maximum number of predictions to 1 and click **OK**. This will now make one prediction for each transaction for the based upon the rules in the Model tab.



10) Add a **Table** node to the generated model and **run it**.

Three new columns were added to the right of the dataset.

The third to last column is the predicted recommendation, the second to last column is the confidence % based on other observed cases, and the last column shows the RuleID from the generated model that corresponds to that prediction.

	Ready made	Frozen foods	Alcohol	Fresh Vegetables	Milk	Bakery goods	Fresh meat	Toiletries	Snacks	Tinned Goods	GENDER	Age	MARITAL	CHILDREN	WORKIN	\$A-10 fields-1	\$A-10 fields-1	\$A-Rule_ID-1
1	1	0	0	0	0	0	0	0	0	1	0 Female	18 to 30	Widowed	No	Yes	Snull\$	Snull\$	
2	1	0	0	0	0	0	0	0	1	0	0 Female	18 to 30	Separated	No	Yes	Snull\$	Snull\$	
3	1	0	0	0	0	0	0	1	1	0	0 Male	18 to 30	Married	No	Yes	Snull\$	Snull\$	
4	1	0	0	0	0	0	1	1	0	0	0 Female	18 to 30	Widowed	No	Yes	Snull\$	Snull\$	
5	1	0	0	0	0	0	0	0	0	0	0 Female	18 to 30	Separated	No	Yes	Snull\$	Snull\$	
6	1	0	0	0	0	0	0	0	0	1	1 Male	18 to 30	Single	No	No	Snull\$	Snull\$	
7	1	0	0	0	0	0	1	0	0	0	0 Female	18 to 30	Single	No	No	Snull\$	Snull\$	
8	1	0	0	0	0	0	0	0	0	0	0 Female	18 to 30	Widowed	No	No	Snull\$	Snull\$	
9	1	0	0	0	0	0	1	0	0	1	1 Female	18 to 30	Single	No	No	Snull\$	Snull\$	
10	1	0	0	0	0	0	0	0	0	0	0 Female	18 to 30	Single	No	No	Bakery goods	0.838	4
11	1	0	0	0	0	0	0	0	0	0	0 Female	18 to 30	Single	No	No	Snull\$	Snull\$	
12	1	1	1	0	1	1	0	0	0	1	0 Female	18 to 30	Single	No	No	Snull\$	Snull\$	
13	1	0	0	0	0	0	0	0	0	0	0 Male	18 to 30	Widowed	No	No	Snull\$	Snull\$	
14	1	0	0	0	0	0	0	0	0	1	1 Female	18 to 30	Single	No	No	Snull\$	Snull\$	
15	1	0	0	0	0	0	0	0	0	0	0 Male	18 to 30	Single	No	No	Snull\$	Snull\$	
16	1	0	1	0	0	0	0	0	0	0	0 Male	18 to 30	Single	No	No	Snull\$	Snull\$	
17	1	0	0	0	0	0	0	0	0	0	0 Male	18 to 30	Single	No	No	Snull\$	Snull\$	
18	1	0	0	0	0	0	0	0	0	1	0 Female	18 to 30	Single	No	No	Snull\$	Snull\$	
19	1	0	0	0	0	0	0	1	0	0	1 Male	18 to 30	Single	No	No	Snull\$	Snull\$	
20	1	0	0	1	0	0	0	0	0	1	1 Female	18 to 30	Single	No	No	Snull\$	Snull\$	
21	1	0	0	0	0	0	0	0	0	0	0 Female	18 to 30	Single	No	No	Snull\$	Snull\$	
22	1	0	0	0	0	0	0	0	0	0	0 Male	18 to 30	Single	No	No	Snull\$	Snull\$	
23	1	0	0	0	0	0	0	0	0	0	0 Female	18 to 30	Widowed	No	No	Snull\$	Snull\$	
24	1	0	0	0	0	0	1	0	1	1	0 Female	18 to 30	Single	No	No	Snull\$	Snull\$	
25	1	0	0	0	0	0	0	0	0	1	0 Male	18 to 30	Married	No	Yes	Snull\$	Snull\$	
26	1	0	1	0	1	0	1	0	0	1	0 Male	18 to 30	Widowed	No	Yes	Bakery goods	0.806	5
27	1	0	0	0	0	0	0	0	0	0	0 Female	18 to 30	Widowed	Yes	Yes	Snull\$	Snull\$	

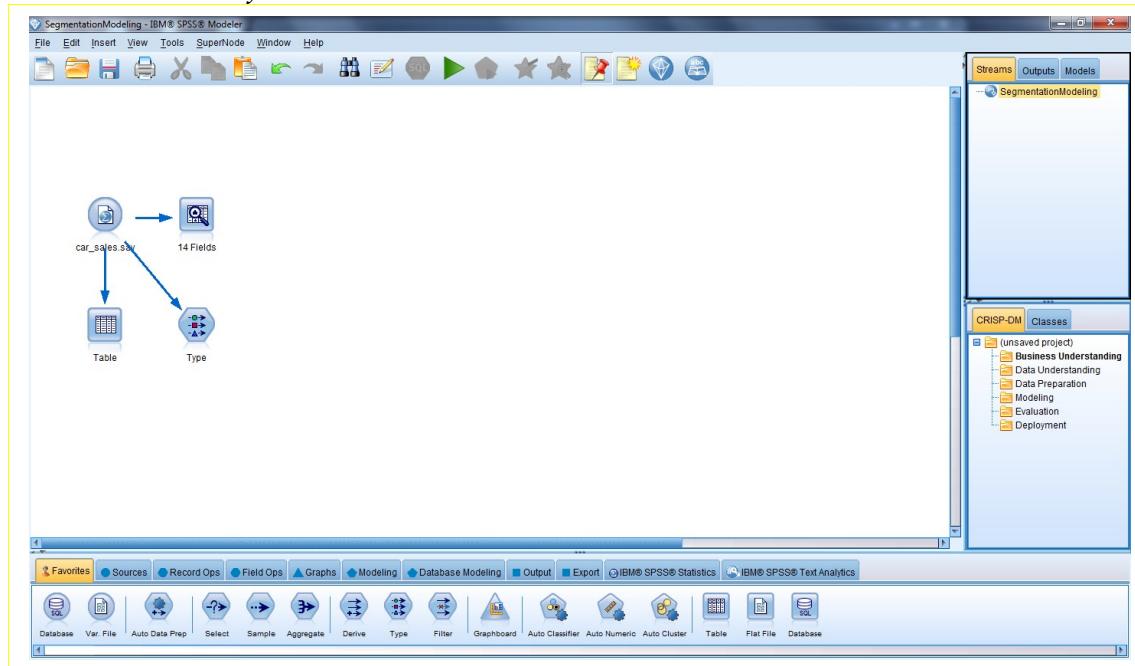
11) **Save the Stream to a local directory before closing it.**

## Lab 3 – Segmentation Modeling

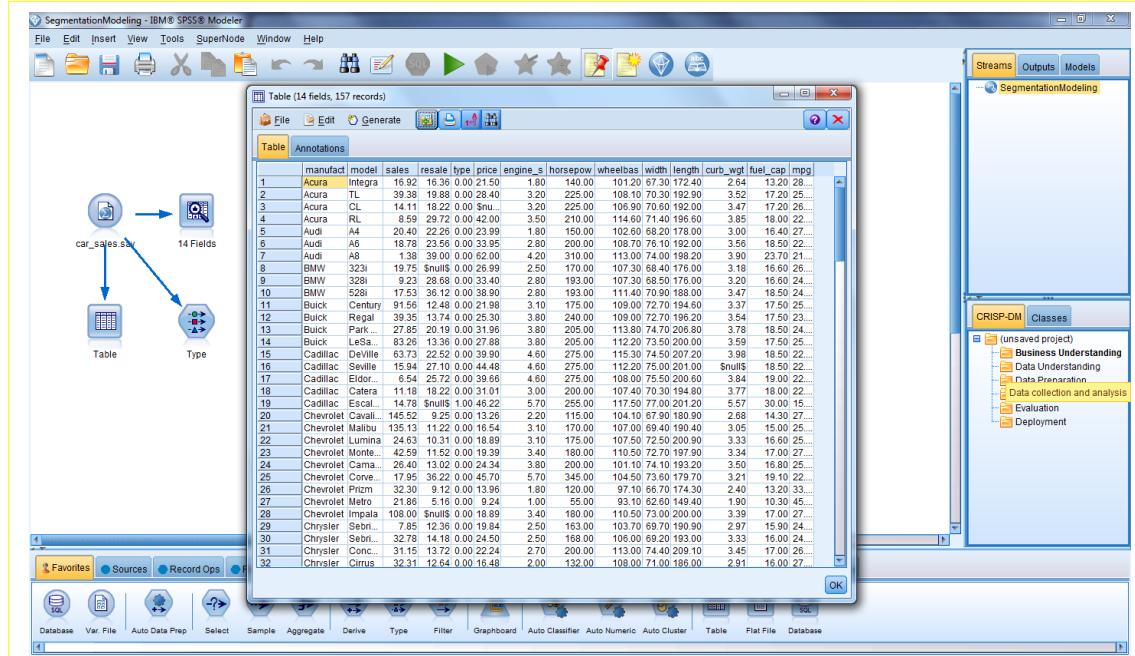
- 1) Open SegmentationModeling.str from the workshop directory.

Go to Menu, select File, select Open Stream and locate

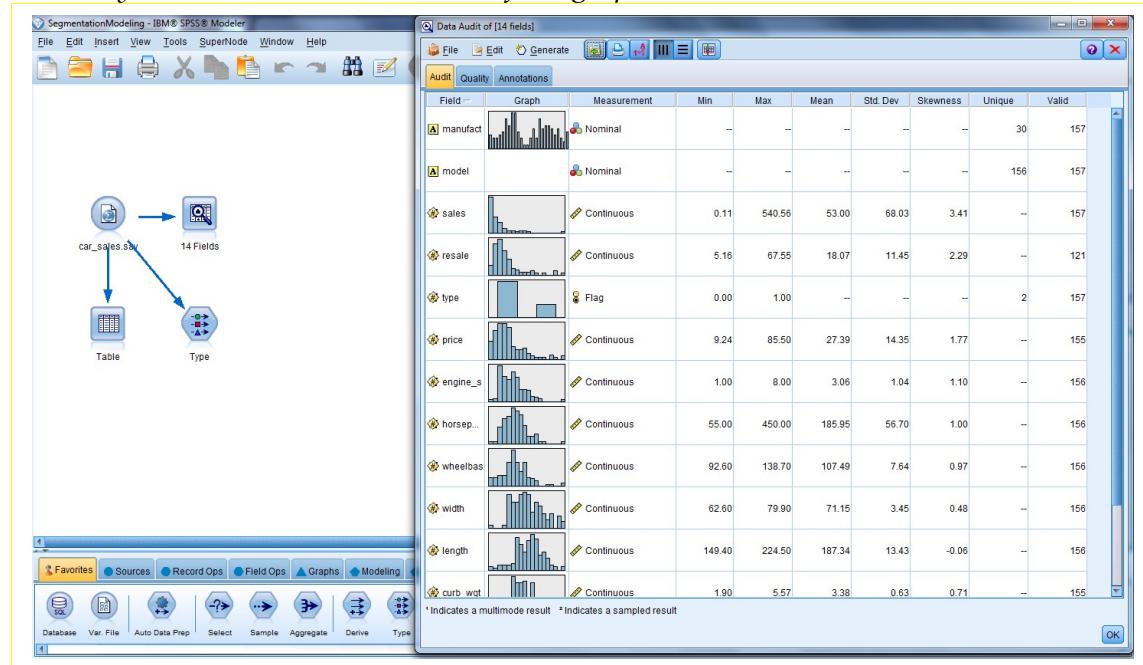
C:\IDR Dwork\SPSS\SPSS Modeler1 Day\03 - SegmentationModeling\SegmentationModeling.str or double-click the file.



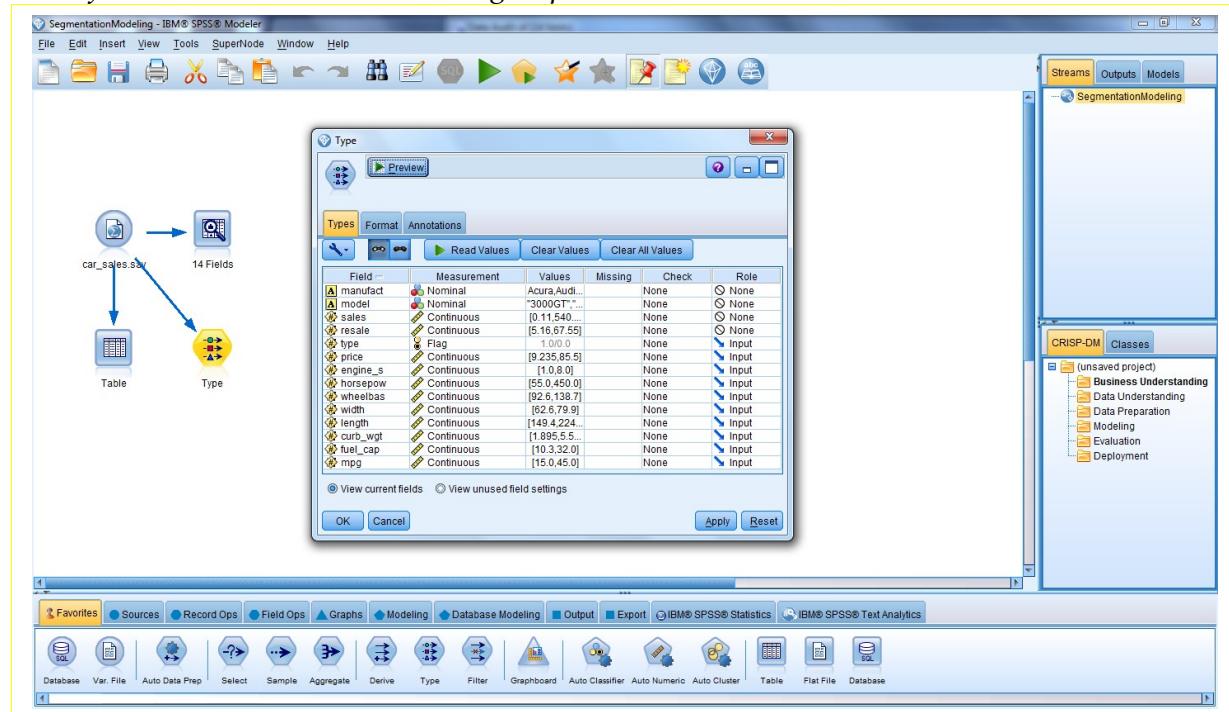
- 2) Run the Table node and view the data about the various makes and models of motor vehicles.



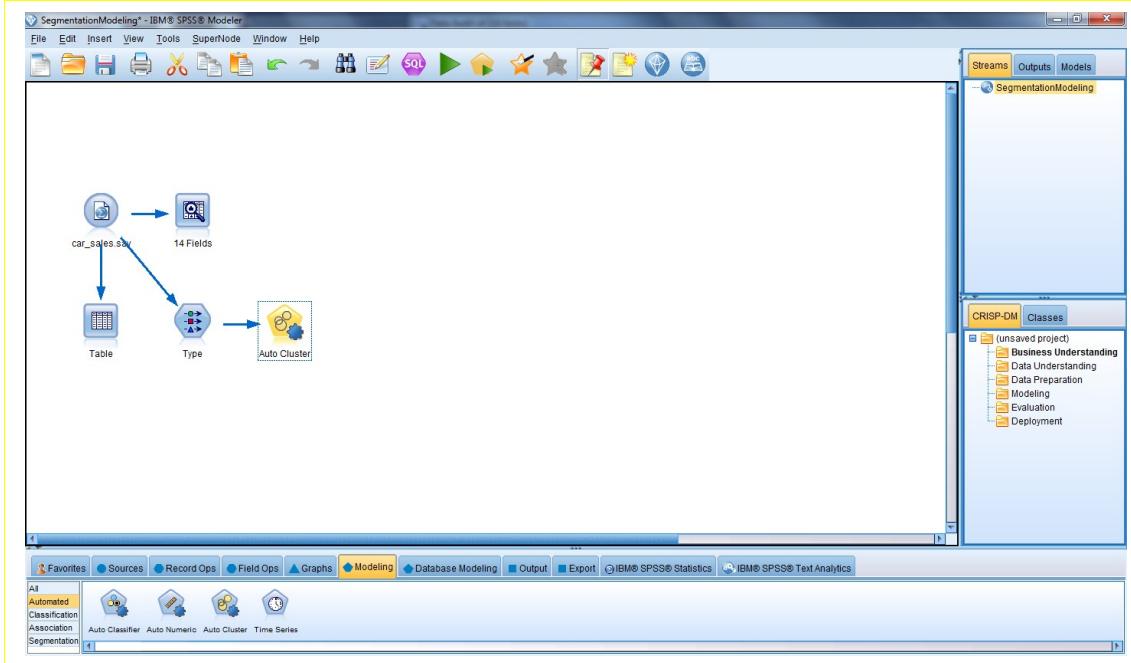
- 3) Run the **Data Audit** node, named 14 Fields, and view the distributions of the data fields and summary statistics. Double-click one of the graphs to see a more detailed version.



- 4) Open the **Type** node and confirm the **Roles** for the type and fields are set to **Input**. These fields will be used to cluster or group the vehicles based on their similarities.



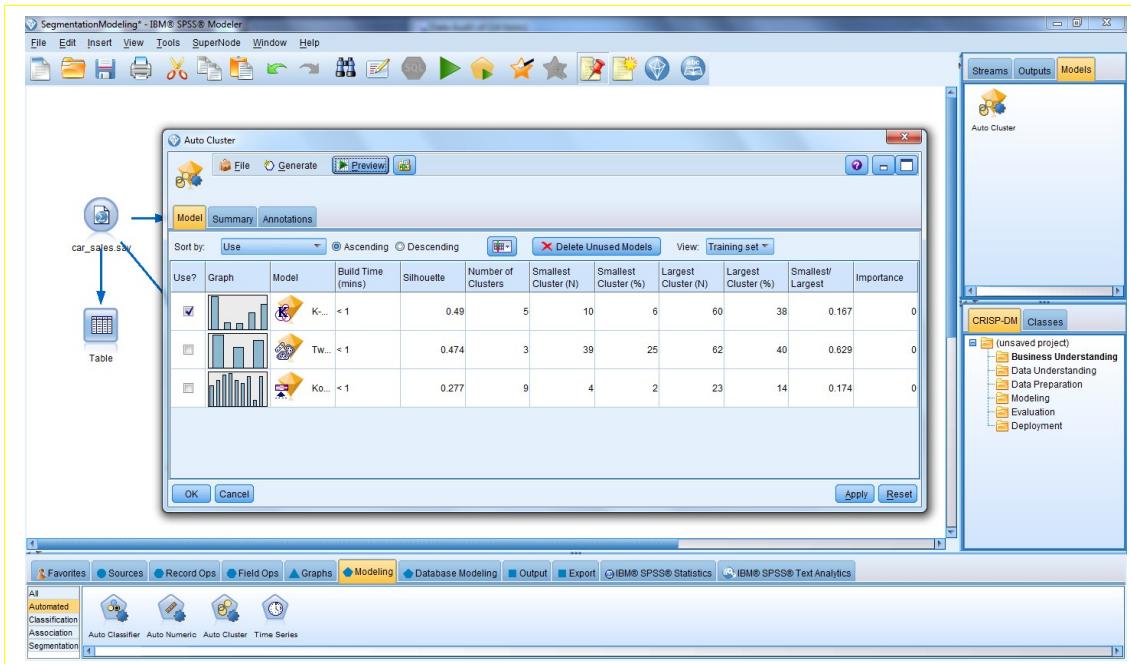
- 5) From the **Modelling** palette, add the **Auto Cluster** node to the **Type** node and run it.



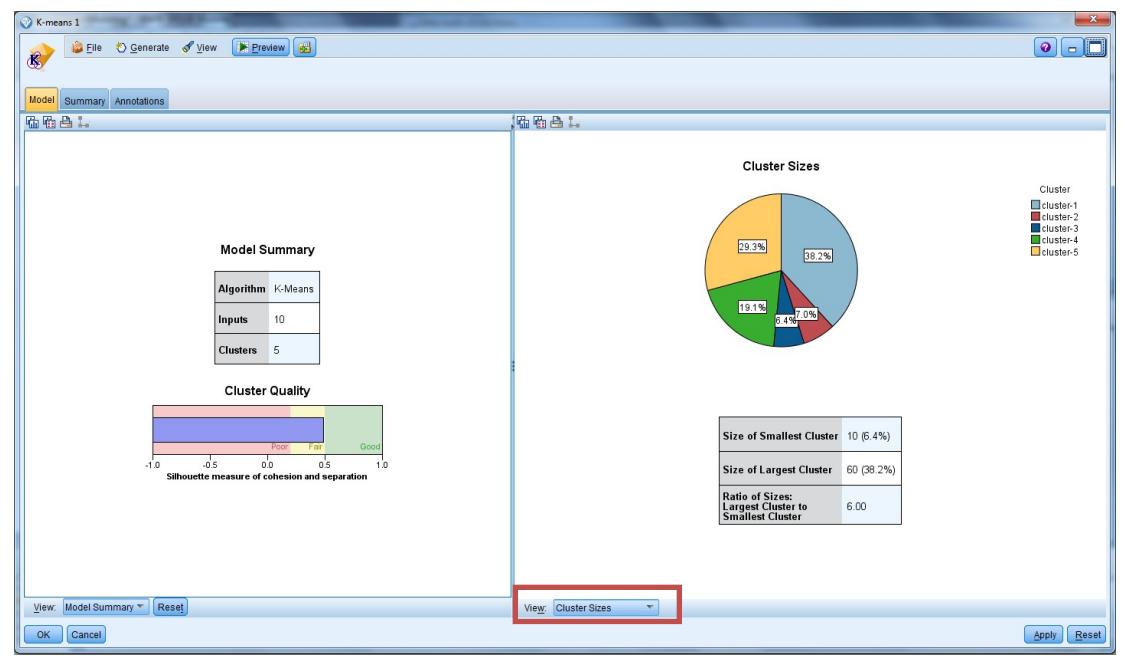
- 6) Double-click the generated model to view the results of the clustering methods.

This dialog box displays a graph illustrating the number of records in each group as well as the method and number of summary statistics used to evaluate the clustering method.

Double-click the **K-means** model to see more details.

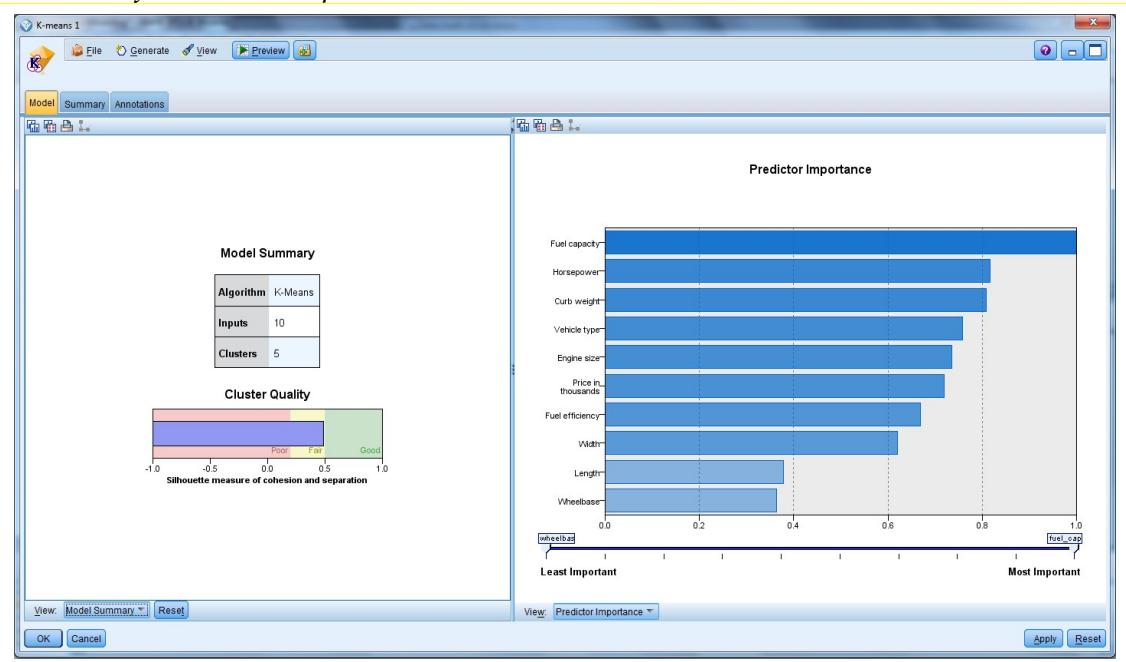


- 7) The output is divided into two viewers: on the left is a model summary; on the right, Cluster Sizes. From the right viewer's drop-down box, select **Predictor Importance**.



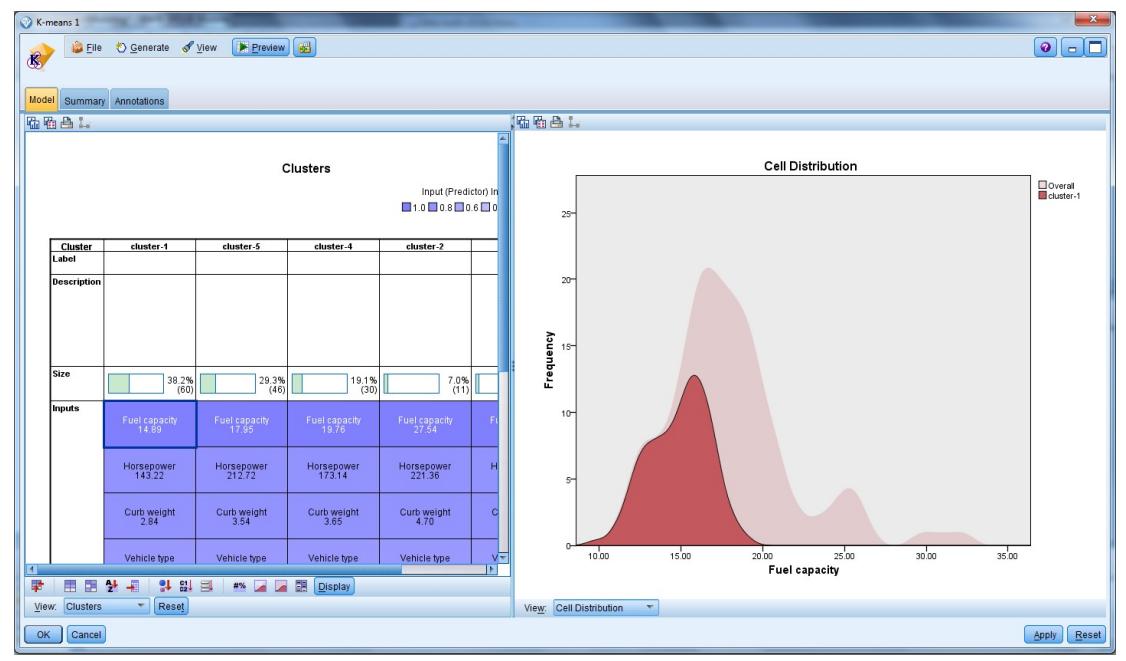
- 8) The right viewer displays a graph with the fields ranked in order of importance for cluster creation.

From the left viewer's drop-down box, select Clusters.

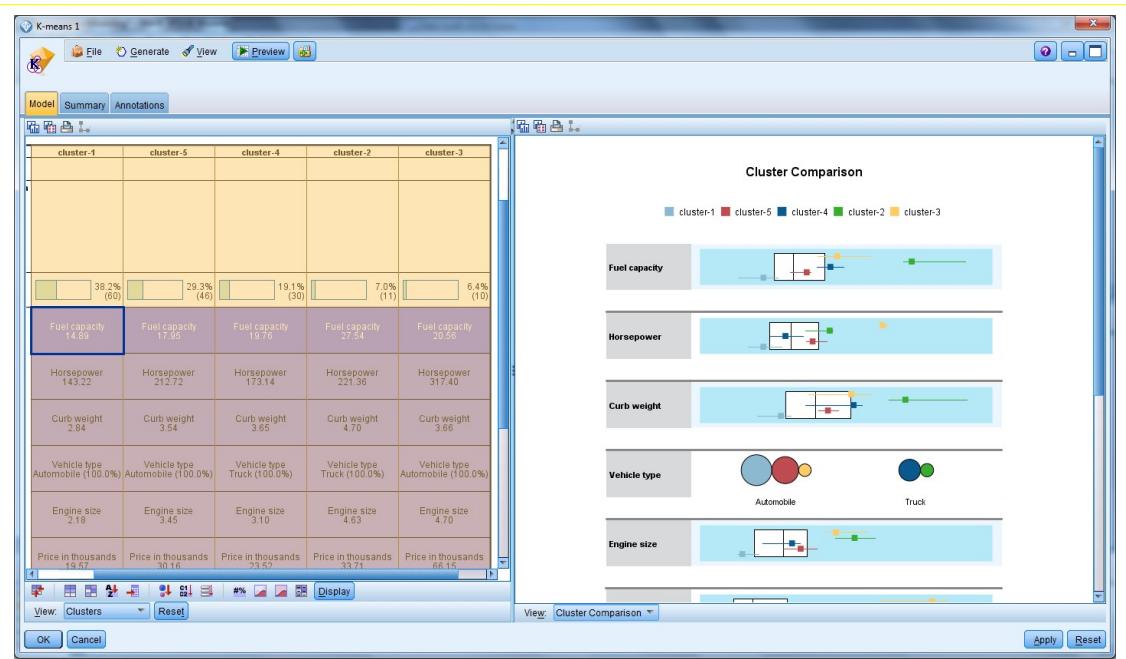


- 9) The left viewer displays a grid with the clusters listed from left to right in order of size – largest to smallest  
– and the fields listed in descending order based on importance.

*Click one of the cells to view how a cluster distribution compares to the remaining records in the right viewer.*



- 10) Select the heading for cluster 1, hold down the Shift button and click on cluster 3 to select the entire table.



*This will change the right viewer to the Cluster Comparison view, illustrating how the cluster are different based on the fields used in the models.*

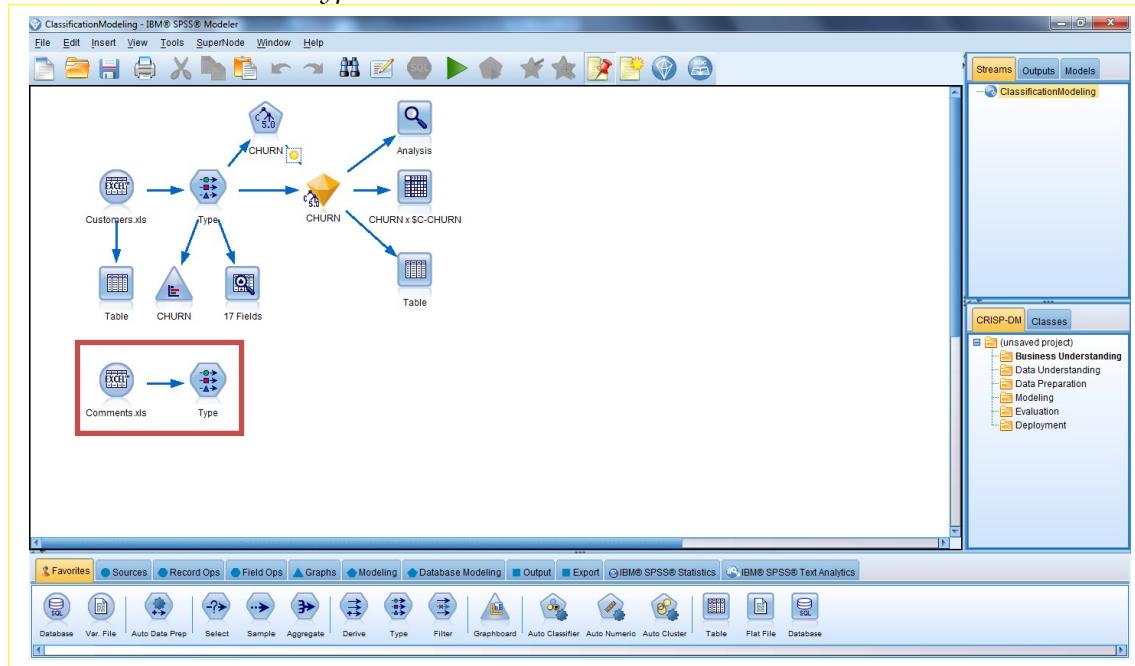
- 11) Add a **Table** node to the generated model to see the cluster membership variable added to the dataset.  
 12) Save the Stream to a local directory before closing it.

## Lab 4 – Classification Modeling

- 1) Open ClassificationModeling.str from the workshop directory:

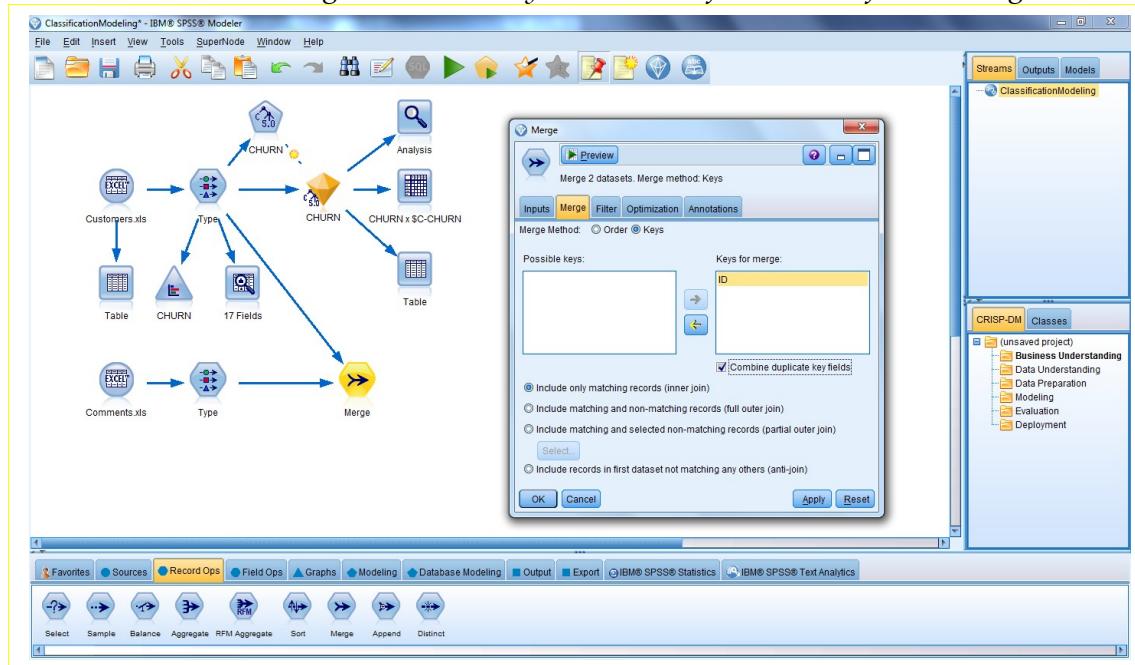
C:\IDR Dwork\SPSS\SPSS Modeler1 Day\04 - ClassificationModeling

Or, as an alternative, use the final stream from the Predictive in 30 Minutes exercise and add the Comments.xls and Type nodes.



- 2) From the Record Ops palette, add the Merge node.

Edit the node so the merge method is Keys and the ID field is used for the merge.

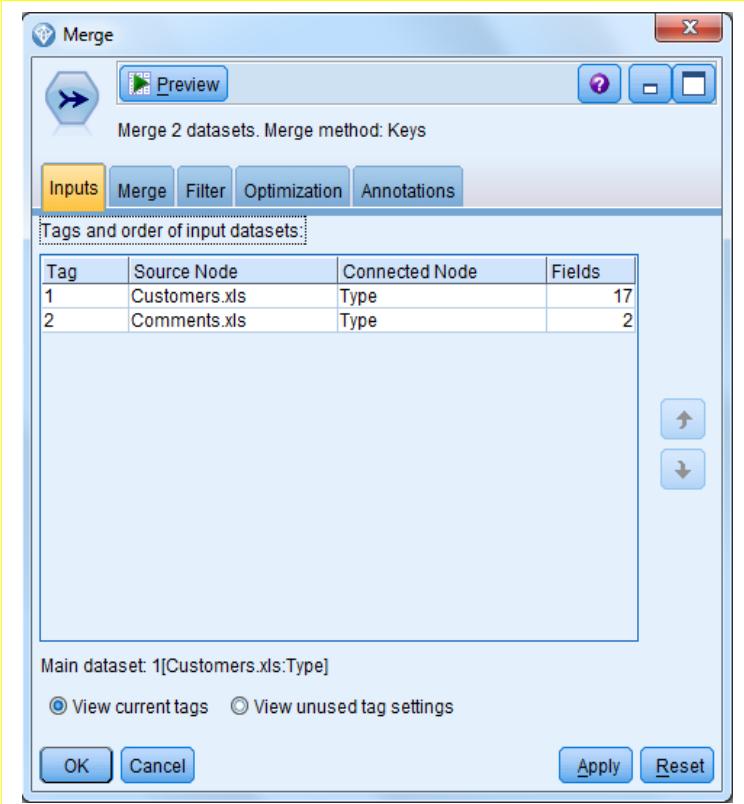


- 3) Check the Inputs tab to ensure the **Customers.xls** file is listed before **Comments.xls**.

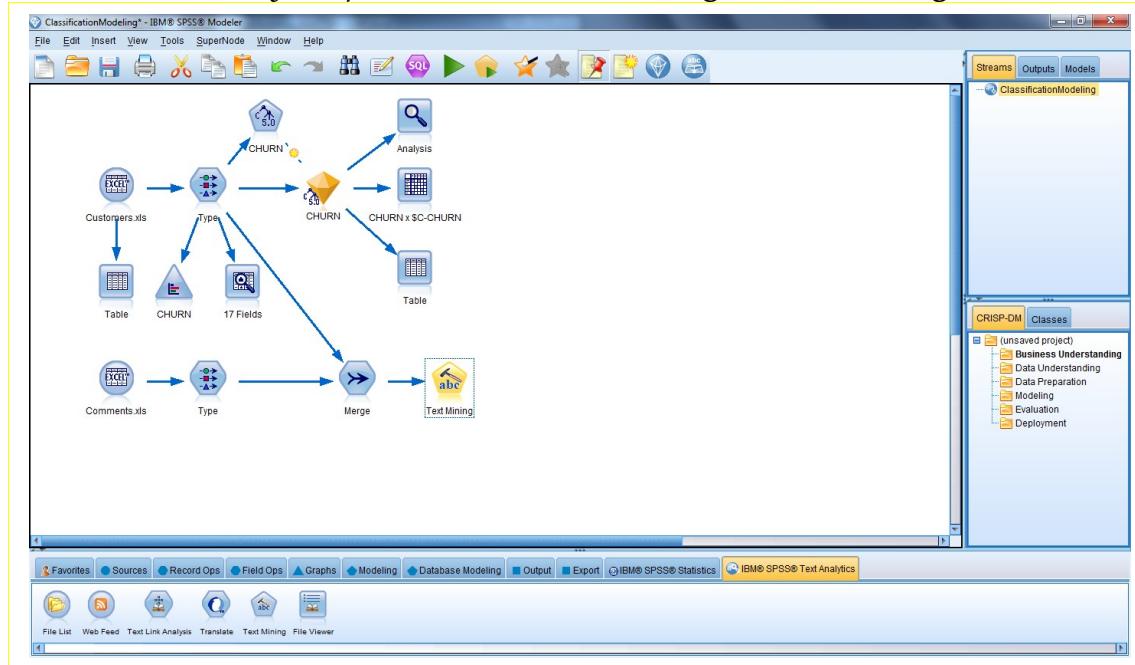
This order will **determine the order** the fields appear when they are merged.

Use the arrows on the right to adjust the order of the inputs, if necessary.

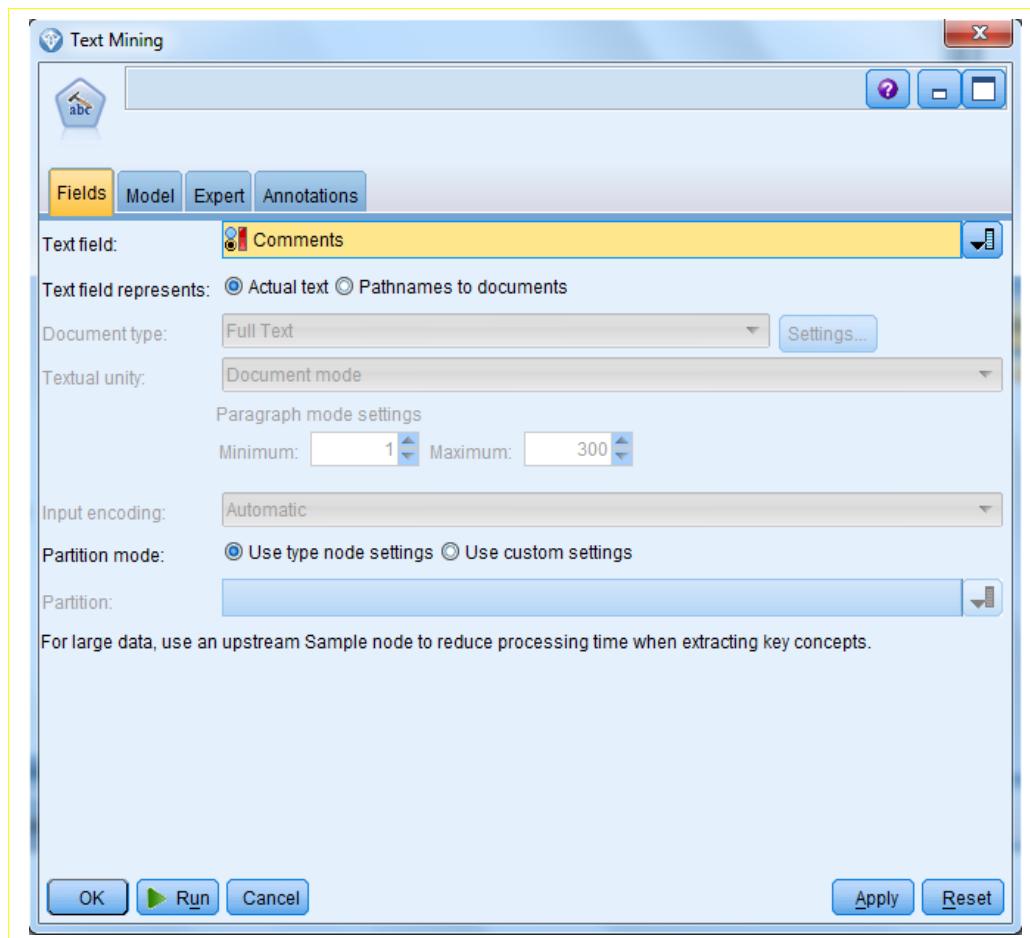
Click **OK** when finished.



- 4) From the **Text Analytics** palette, add the **Text Mining** node to the **Merge** node.

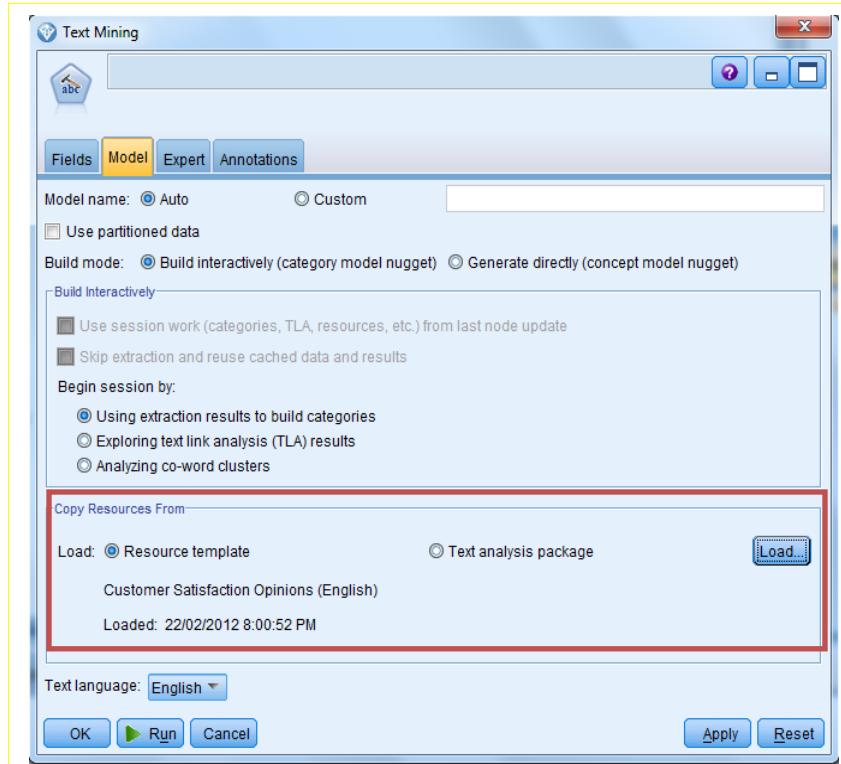


- 5) Edit the Text Mining node, using the Field tab to add **Comments** to the Text field.



- 6) Select the **Model** tab, scroll down to the **Copy Resources From** section and add **Customer Satisfaction Opinions Library**. This will load pre-built resources into the text mining process. Click RUN.

7)



- 8) Once the libraries and resources are loaded and the extraction process is complete, an **interactive workbench** is displayed.

Click **Build** and, if prompted, **Build Now**.

Use the defaults for this example.

The screenshot shows the Interactive Workbench - Comments application window. The top menu bar includes File, Edit, View, Generate, Categories, Tools, and Help. The toolbar contains icons for Build, Extend, Undo, Redo, Save, and Print. The main interface is divided into two panes:

- Categories pane (Top Left):** Shows a tree view of categories. The root node is "All Documents", which has children "Uncategorized" and "No concepts extracted". A status bar at the bottom right of this pane indicates: "To show bars or graphs after building categories, make a selection in another pane and click Display."
- Concepts pane (Bottom Left):** Shows a table of 240 concepts. The columns are Concept, In, Global, Docs, and Type. The data includes rows like "phone" (10 docs), "calls" (6 docs), "store" (5 docs), etc. A status bar at the bottom right of this pane indicates: "To populate data pane, make a selection in another pane and click Display to see corresponding results."

- 9) This populates the **categories** in the top left section. They can be expanded to view the concepts and descriptors that make them up.

The screenshot shows the Interactive Workbench - Comments application window with the Categories pane expanded. The tree view under "All Documents" shows the following structure and data:

- "All Documents" (Root)
  - "Uncategorized"
    - "No concepts extracted"
  - "phone" (10 docs)
  - "calls" (6 docs)
  - "store" (5 docs)
  - "family" (4 docs)
  - "sales" (4 docs)
  - "service" (4 docs)
  - "travels" (3 docs)
  - "battery" (3 docs)
  - "rates" (3 docs)
  - "minutes" (3 docs)
  - "charge" (3 docs)
  - "model" (2 docs)
  - "customer assistance" (2 docs)
  - "car" (2 docs)

- 10) Left click on the rates category, hold and drag it to the charge category.

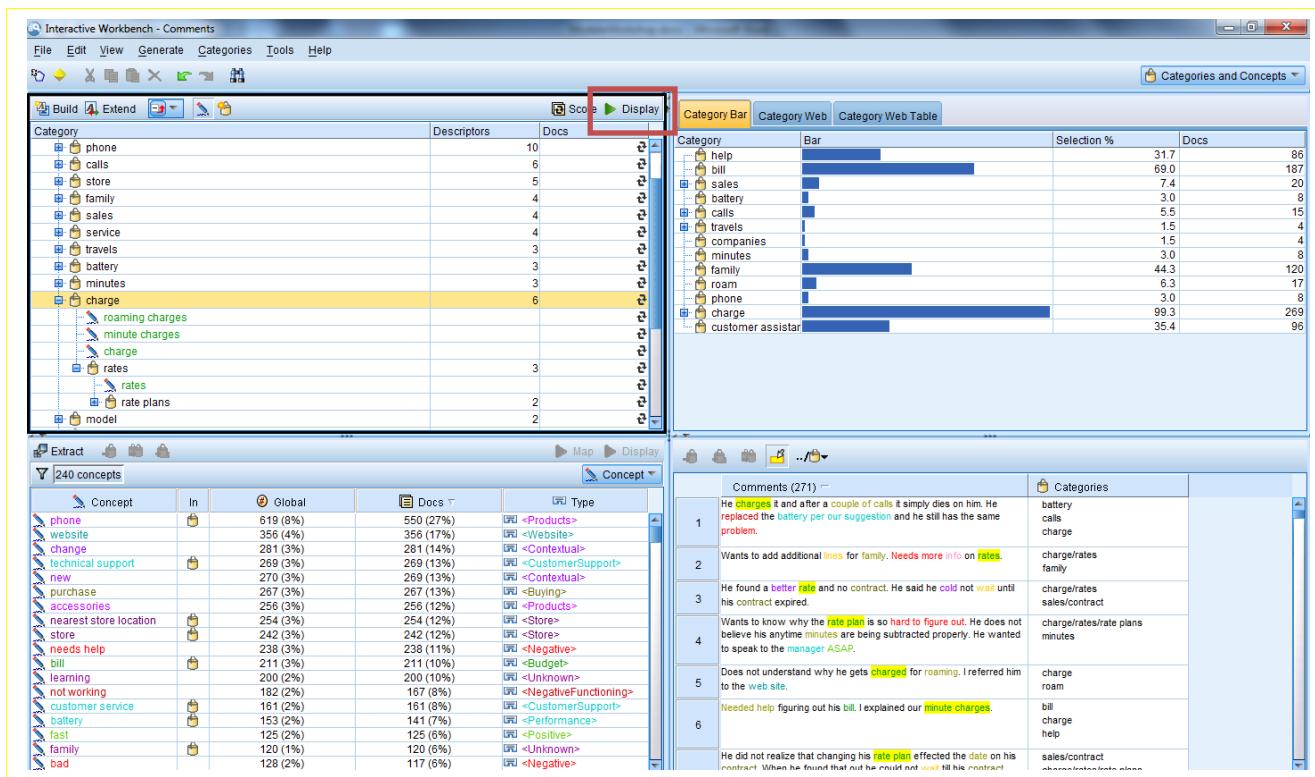
Release to combine the two when the icon, as seen below, appears.

Category	Descriptors	Docs
phone	10	2
calls	6	2
store	5	2
family	4	2
sales	4	2
service	4	2
travels	3	2
battery	3	2
<b>rates</b>	<b>3</b>	<b>2</b>
<b>rates</b>		
rate plans	2	2
minutes	3	2
<b>charge</b>	<b>3</b>	<b>2</b>
roaming charges		
minute charges		
charge		
model	2	2

- 11) The charge category is combined with the rates category and rate plans subcategory.

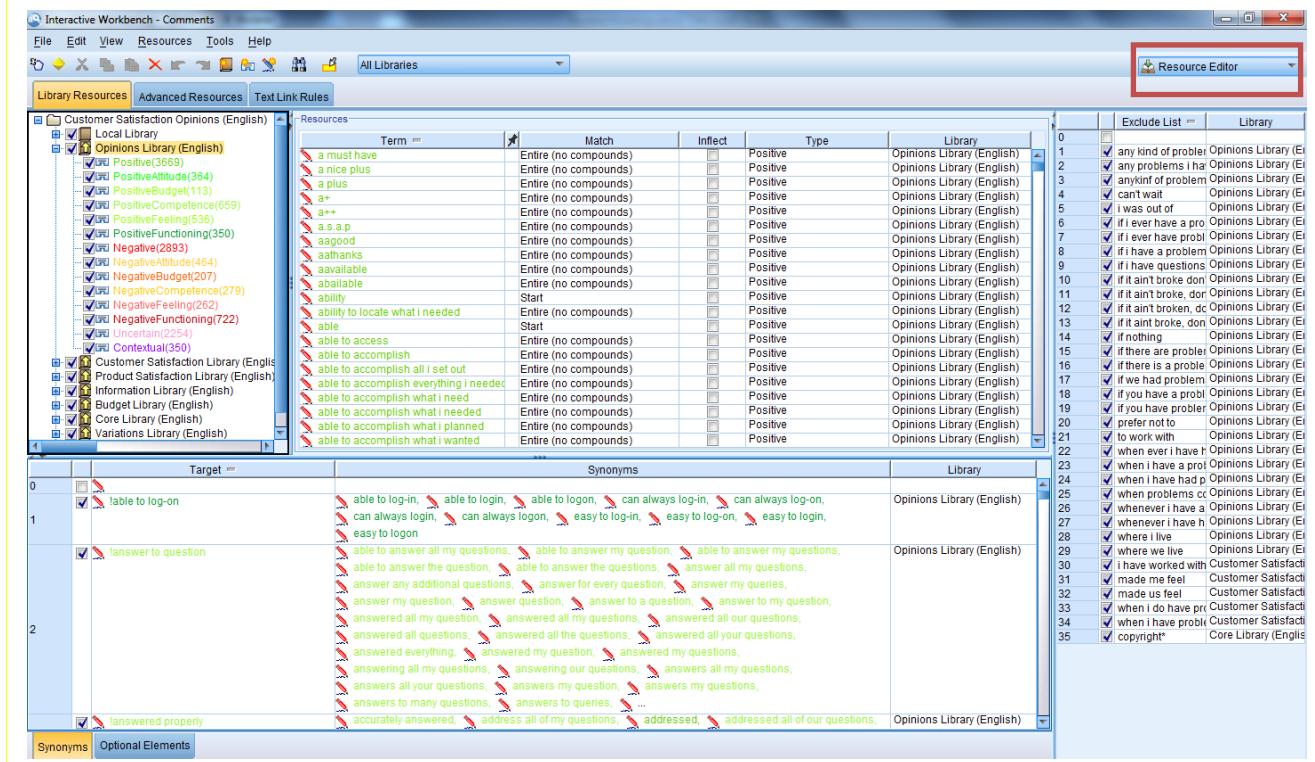
Category	Descriptors	Docs
phone	10	2
calls	6	2
store	5	2
family	4	2
sales	4	2
service	4	2
travels	3	2
battery	3	2
minutes	3	2
charge	6	2
roaming charges		
minute charges		
charge		
<b>rates</b>	<b>3</b>	<b>2</b>
<b>rates</b>		
rate plans	2	2
model	2	2

- 12) Select charge and click **Display** to view the individual comments with occurrences in the bottom right corner and co-occurrence other categories in the upper right corner.



13) Select **Resource Editor** from the drop-down menu in the top right corner.

*This is where the libraries are located and you can manipulate how concepts and sentiments are handled in the program.*



14) Move back to the **Categories and Concepts** view.

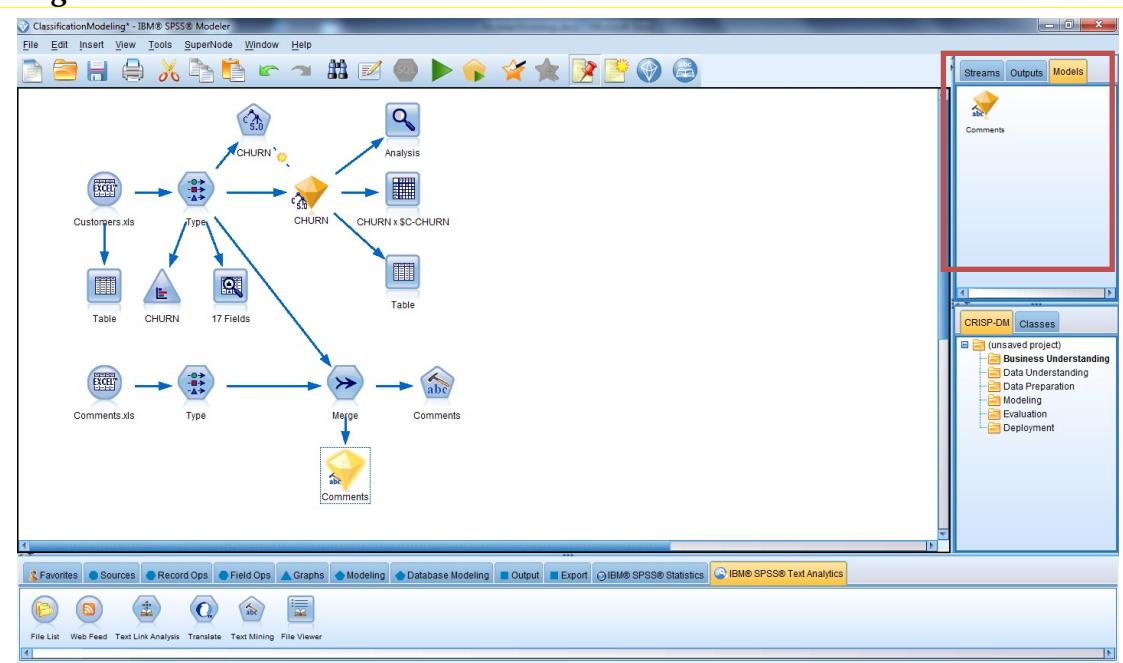
From the **Generate** menu, select **Generate Model**.

Close the window and, when prompted, select update then **OK**.

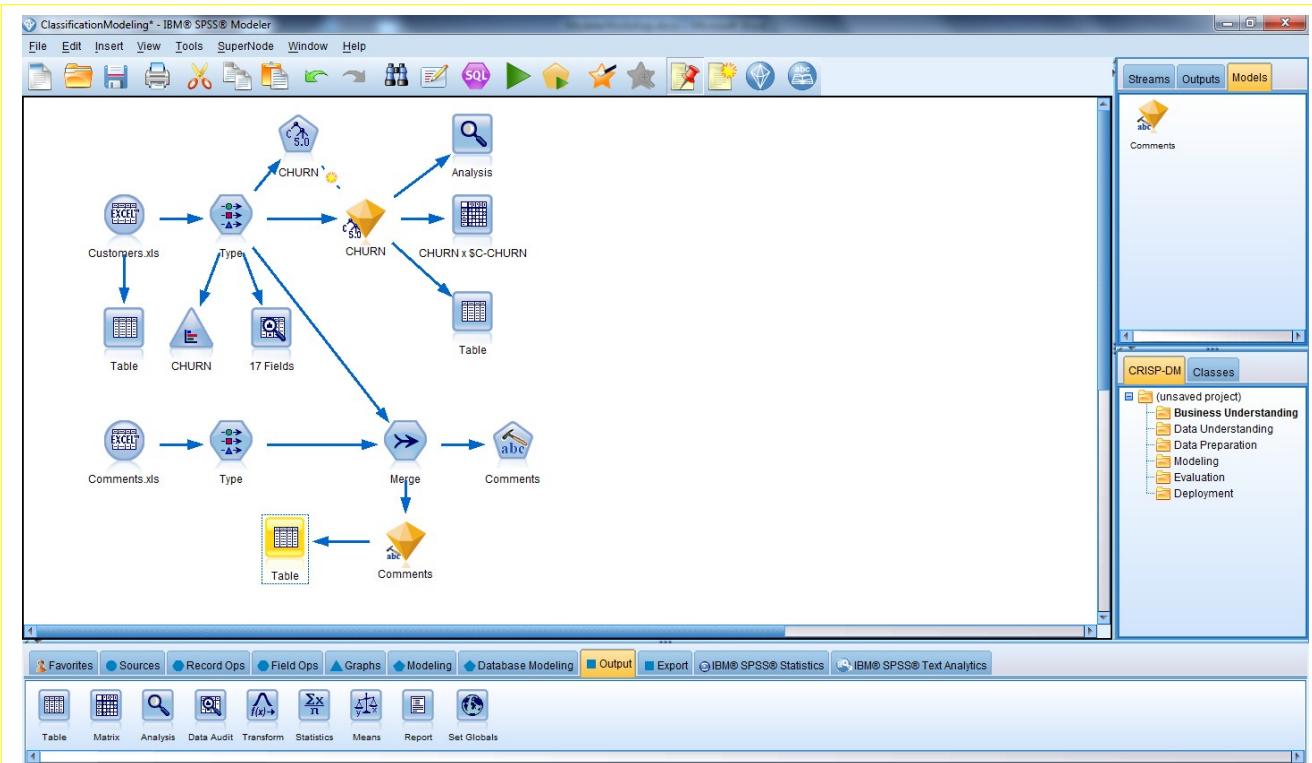
This returns you to the main screen.

The screenshot shows the Interactive Workbench - Concepts interface. The top menu bar has 'Generate' highlighted with a red box. The right-hand panel is titled 'Categories and Concepts' with a red box around it. Below it is a bar chart titled 'Category Bar' showing selection percentages for various categories like Bar, help, bill, sales, battery, etc. The bottom panel shows a table of '240 concepts' with columns for Concept, In, Global, Docs, and Type. A large text area on the right contains comments from users, such as 'He charged it and after a couple of calls it simply dies on him. He replaced the battery per our suggestion and he still has the same problem.' and 'Wants to add additional lines for family. Needs more info on rates.' The entire interface is enclosed in a yellow border.

15) From the **Models pane** in the top right, drag the generated model to the cases and add it to the **Merge** node.



16) From the **Output** palette, add a **Table** node to the **modeling** node and run the **Table** to see the comment categories.



17) Scroll right to view the new categories created.

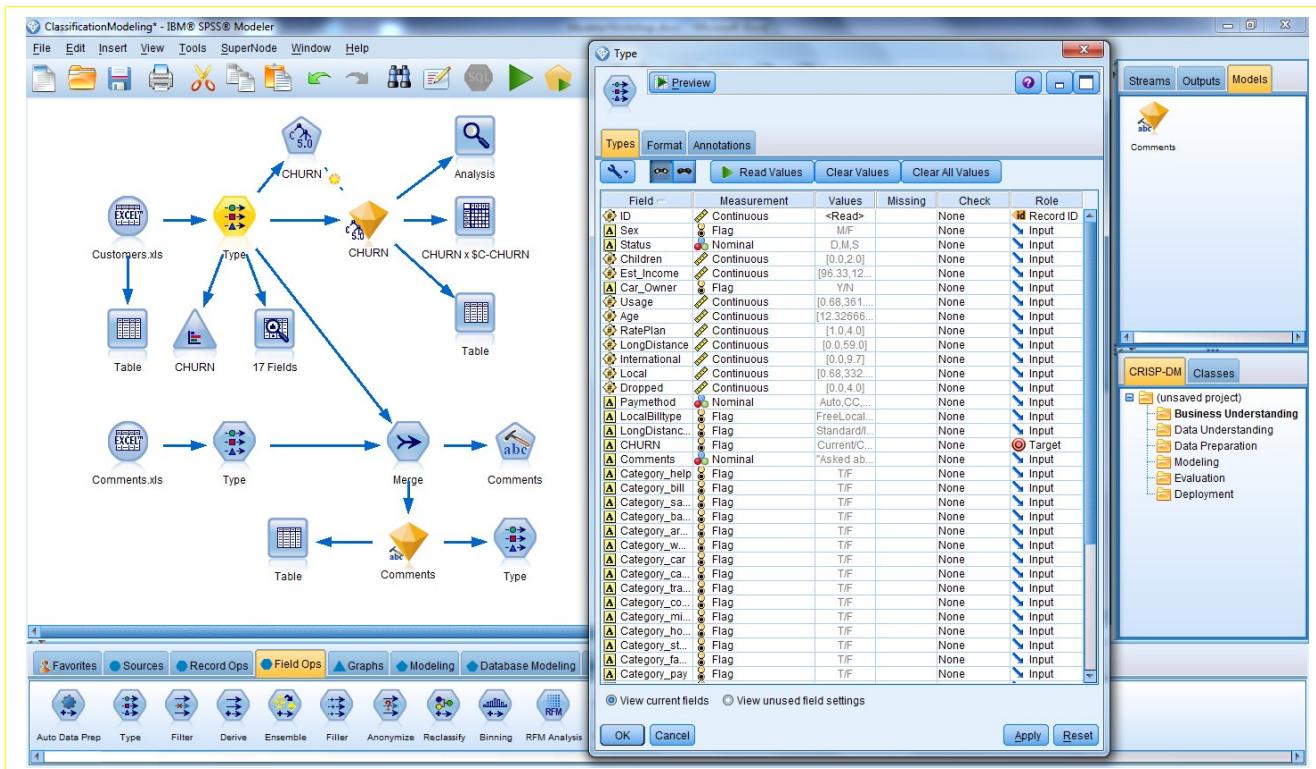
*There is a T if a comment was put into the category and an F if one wasn't.*

18) From the **Field Ops** palette, add a **Type** node to the model.

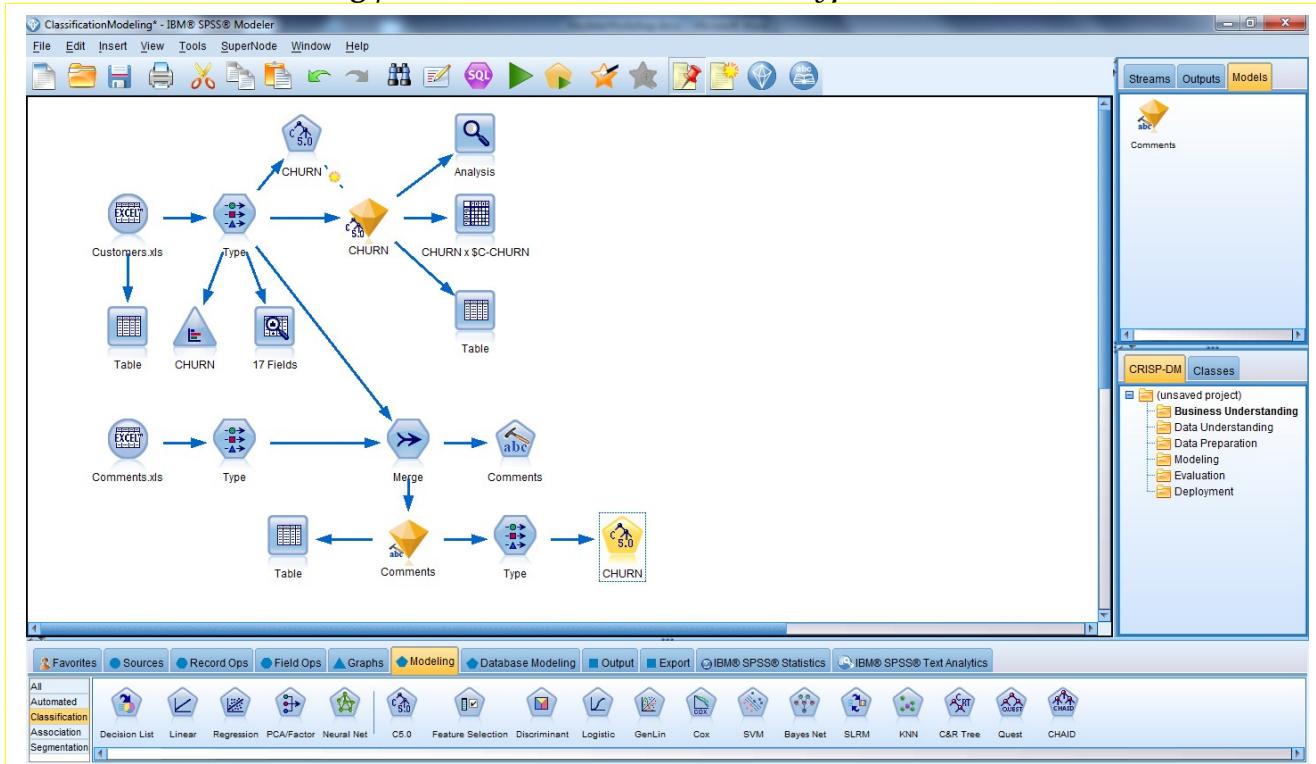
*Check that the Role for all fields except ID and CHURN, are set as Input.*

*ID should be Record ID and CHURN, to Target.*

*The additional fields will be used to improve the accuracy of the model.*

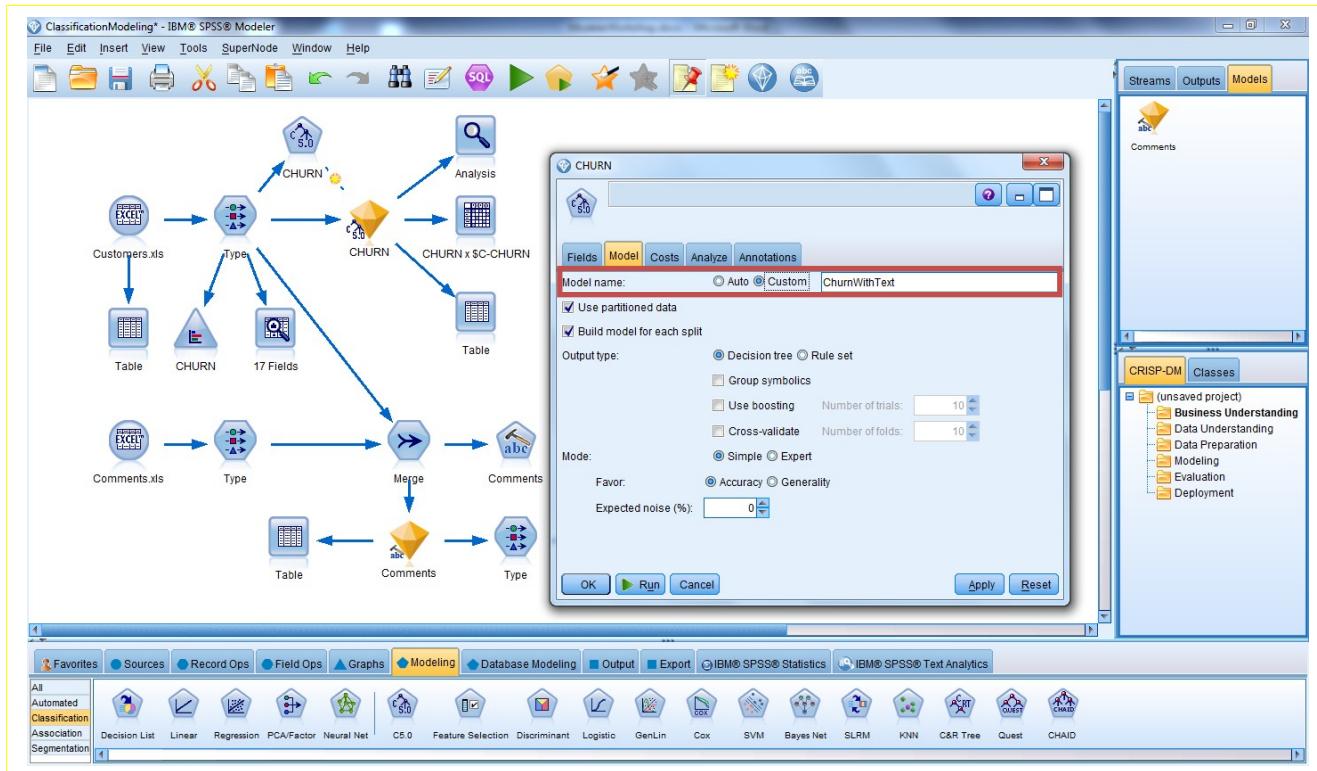


19) From the Modeling palette, add the C5.0 node to the Type node.

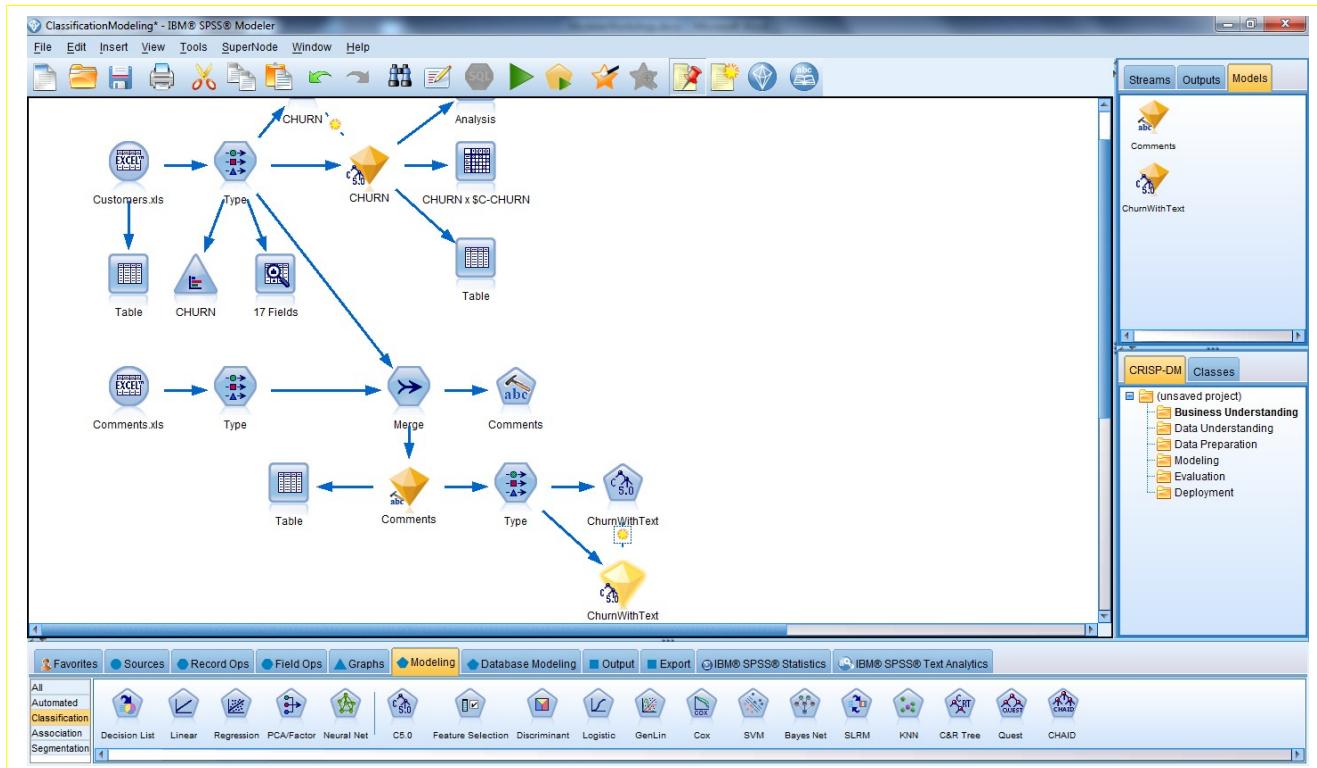


20) Edit the **modelling** node so the name of the generated model is ChurnWithText.

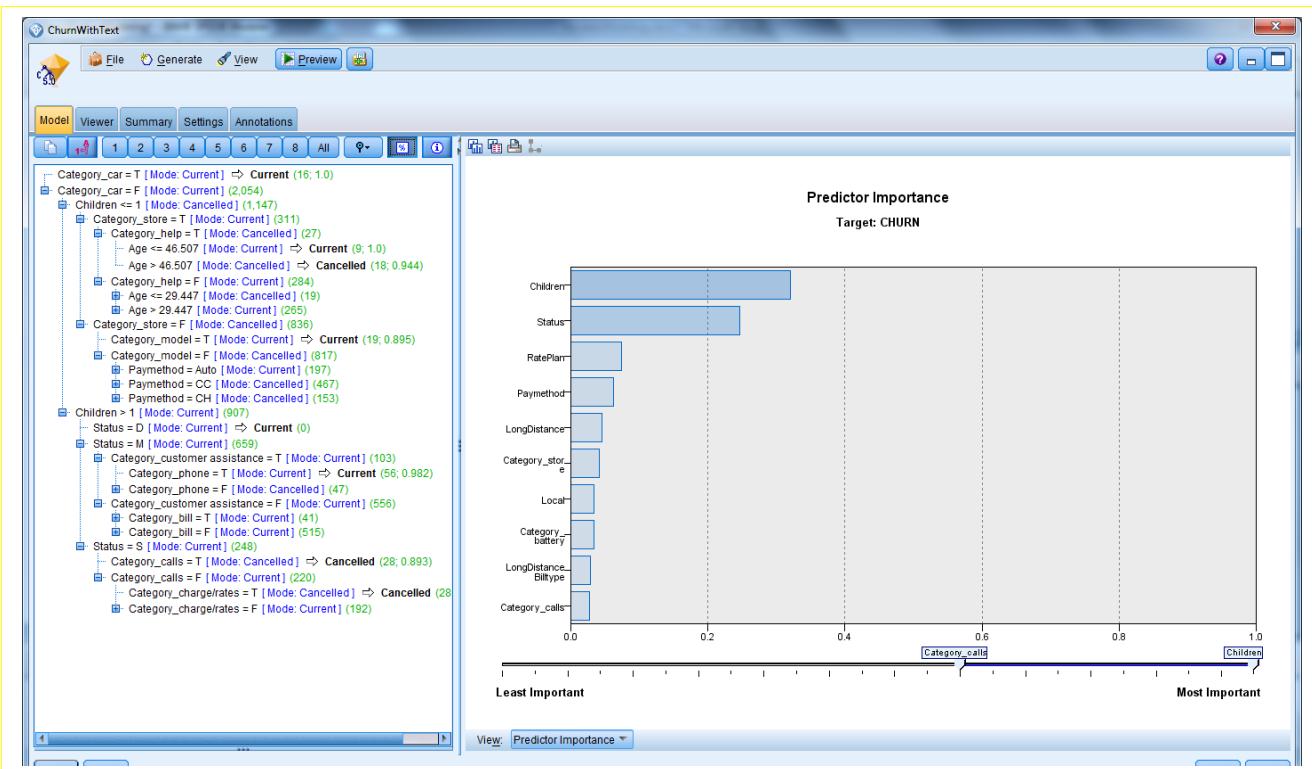
Run the model.



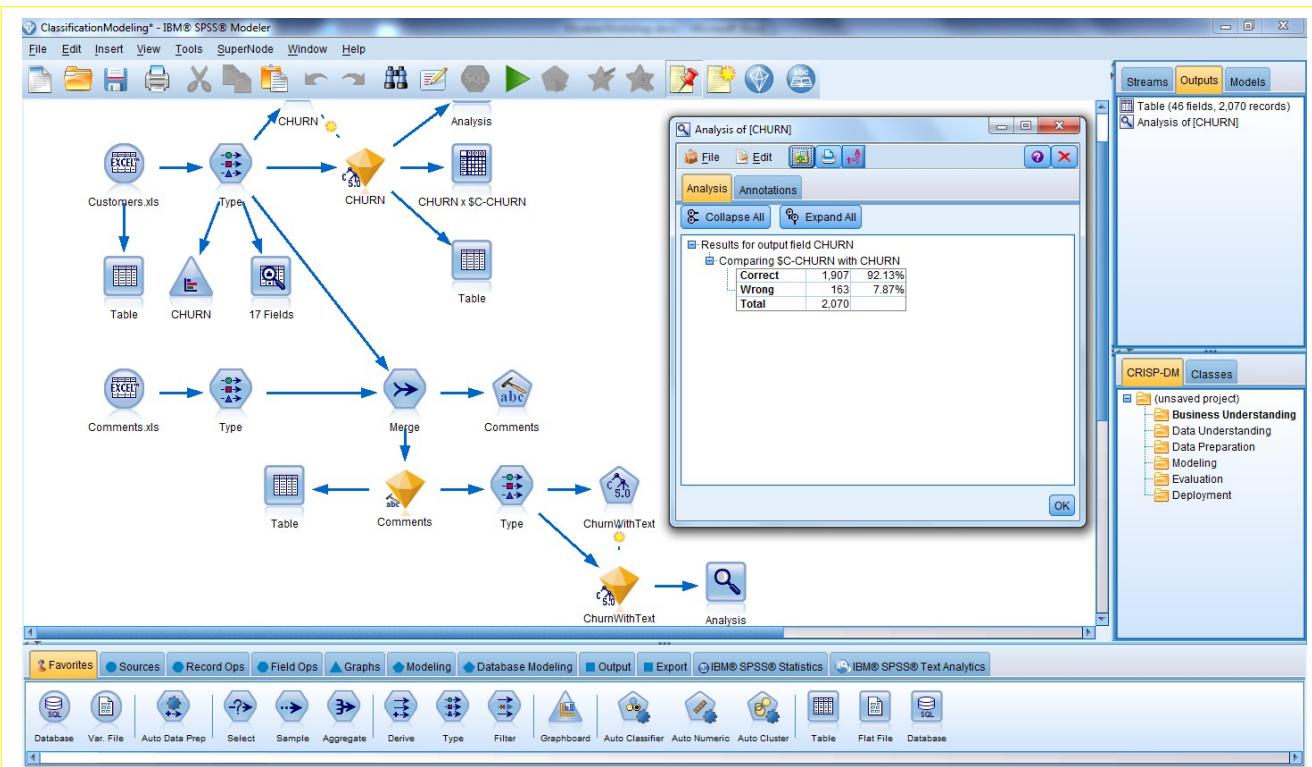
21) Double-click the generated model to view the **predictors** and influence the comment categories have on the model.



22) Expand the **model rules**. Select the **Viewer** tab to view the model and the relationship between the predictors and outcome.

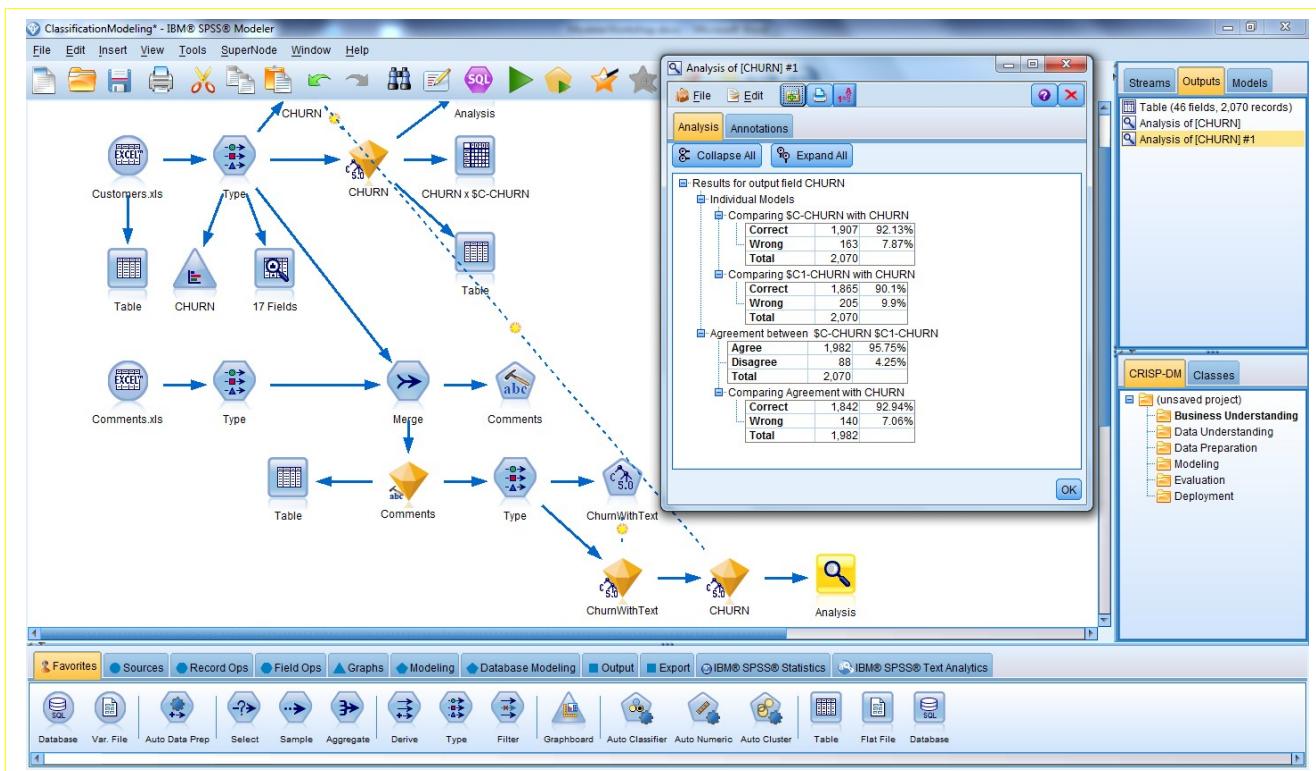


23) Add an Analysis node to the ChurnWithText model and run it to view the accuracy of the model.



24) To easily compare the two models, copy and paste the CHURN model and insert it between the ChurnWithText and Analysis nodes.

Run the Analysis node.

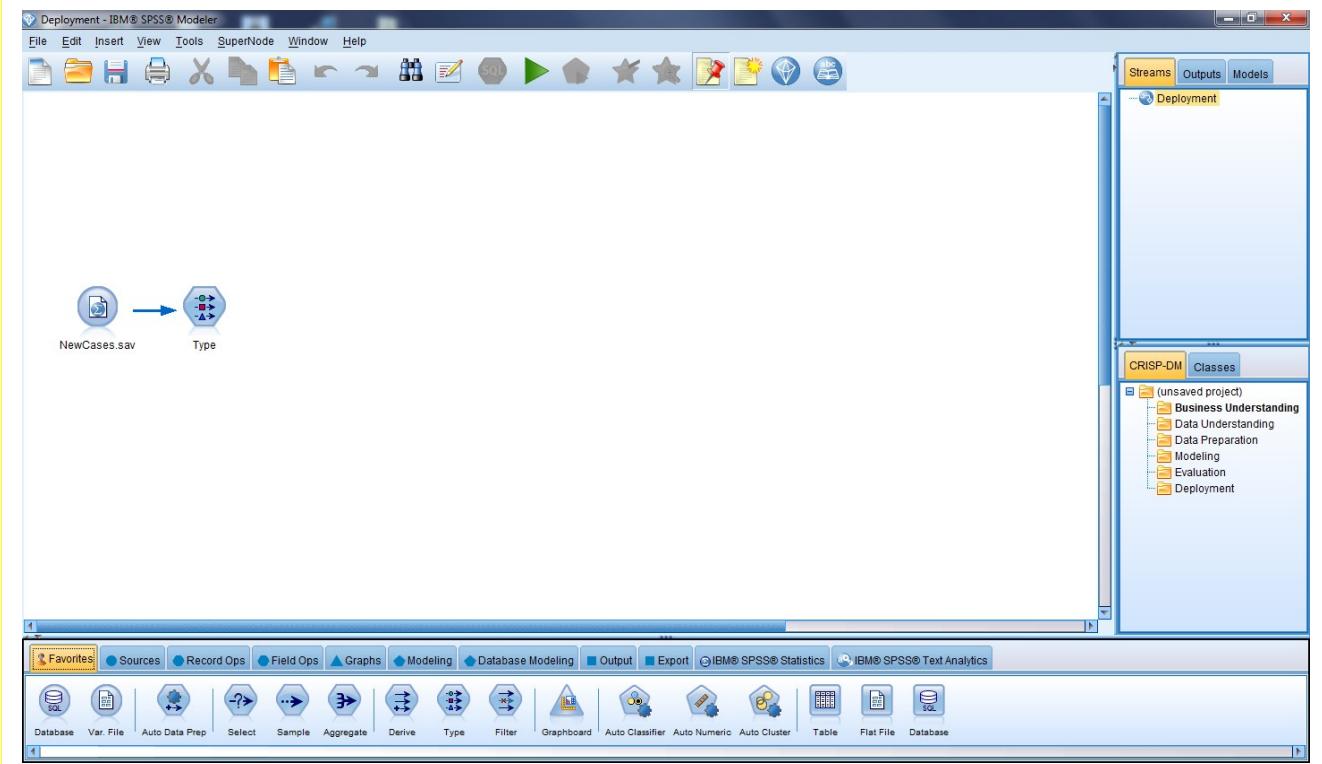


25) Save the Stream to a local directory before closing it.

## Lab 5 – Deployment

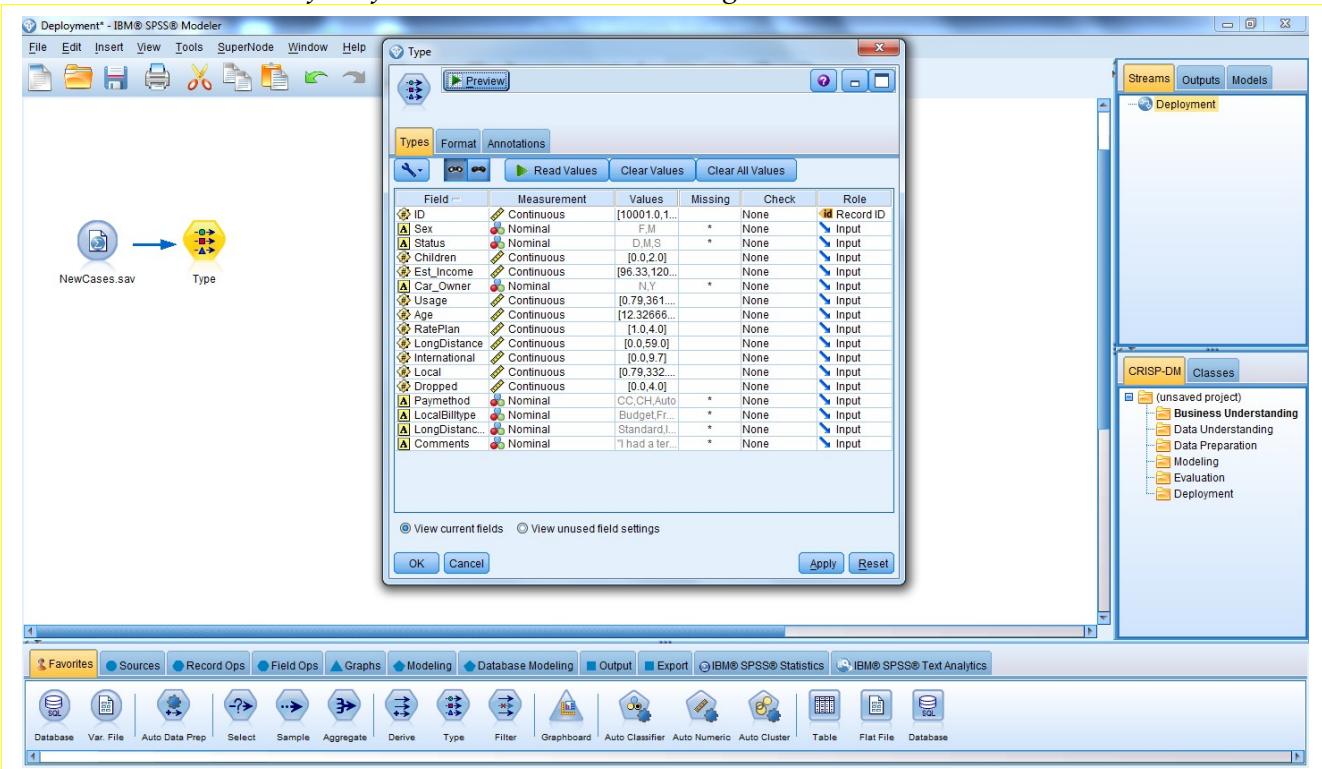
- 1) Open Deployment.str from the workshop directory:

C:\IDR Dwork\SPSS\SPSS Modeler1 Day\05 - Deployment

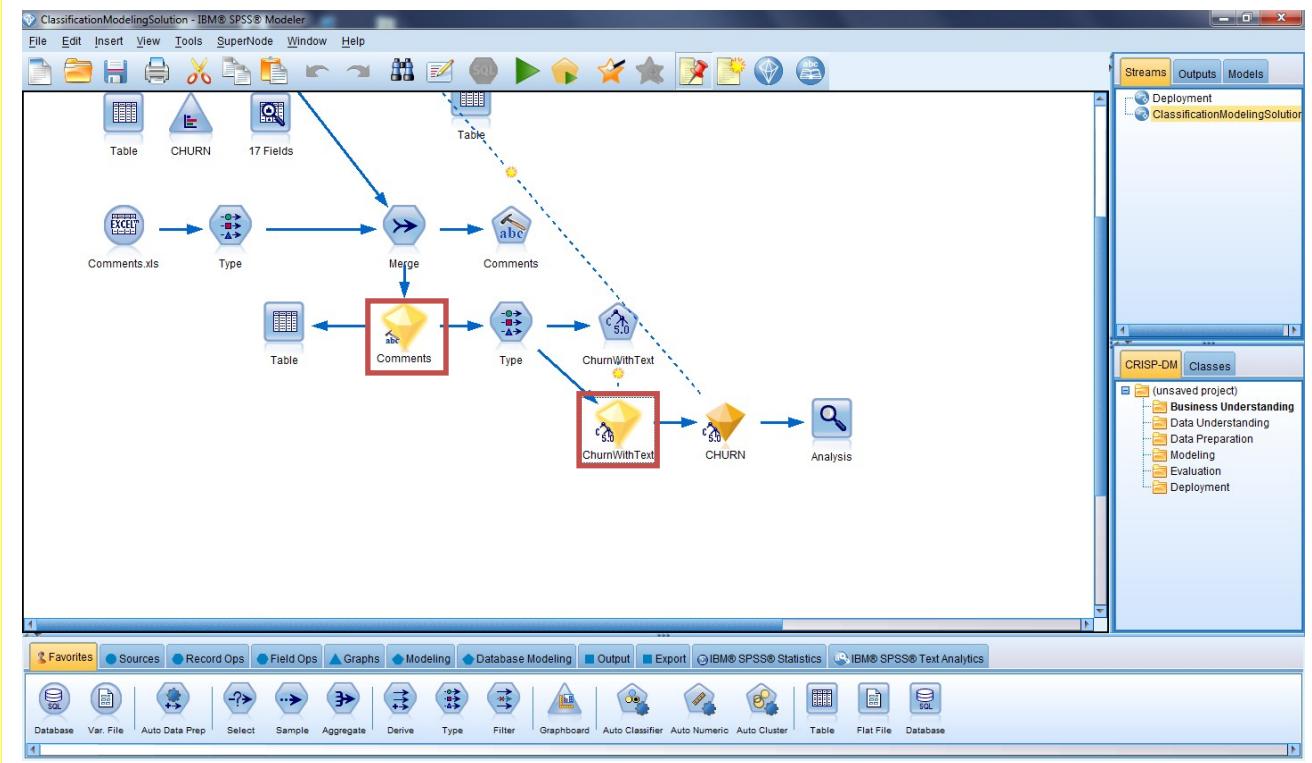


- 2) Check the Type node.

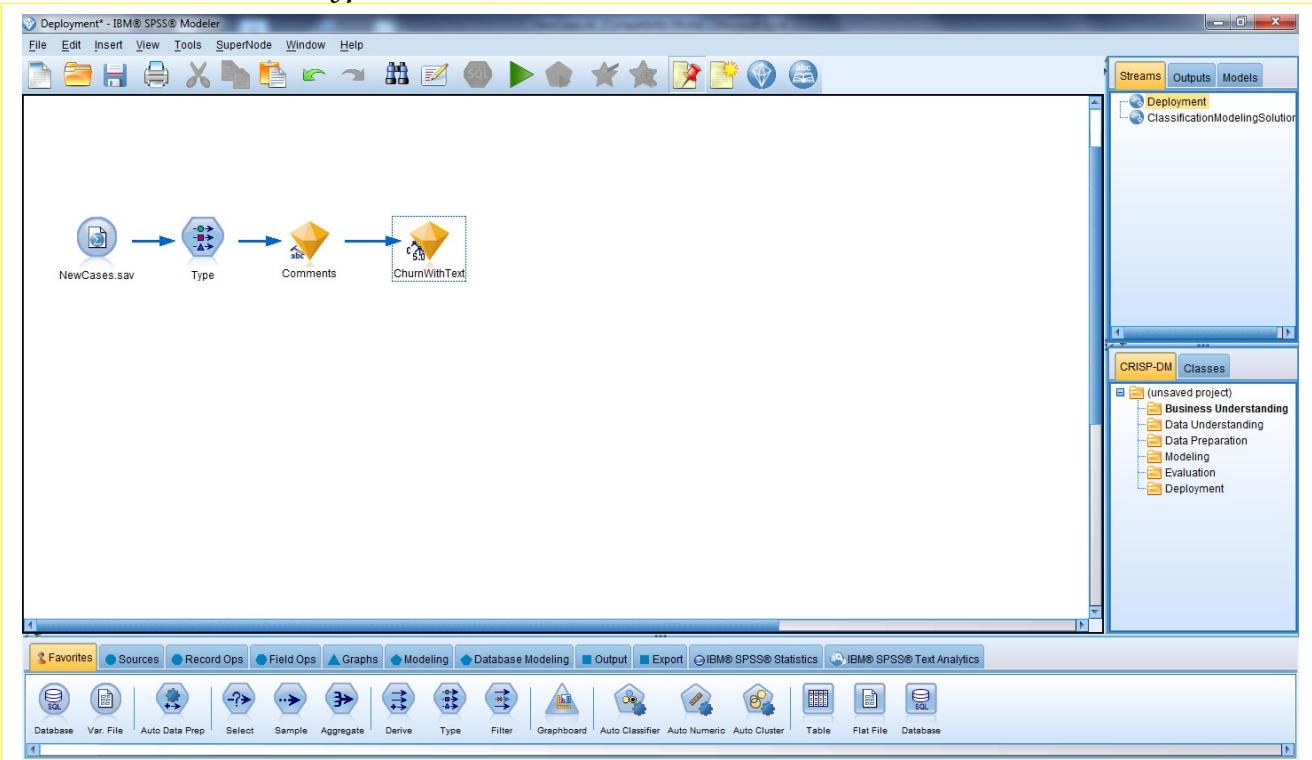
Note that none of the fields have a Role set to Target.



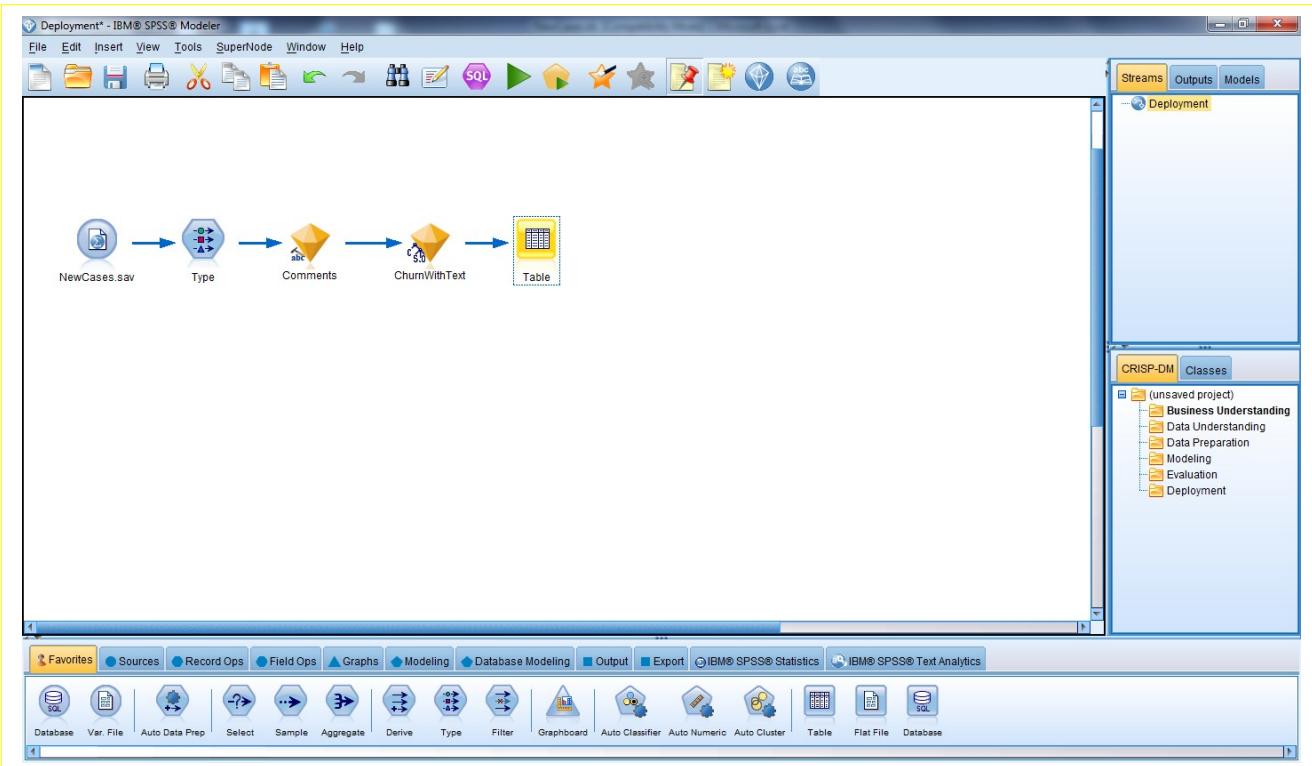
- 3) Copy the *Comments* and *ChurnWithText* models from the Classification stream.  
 Use *ClassificationSOLUTION.str* if needed.



- 4) Paste the models onto the canvas with the Deployment stream.  
 Add them to the *Type* node, as illustrated below.



- 5) Add a *Table* node to the *ChurnWithText* model.  
 Run the *Table*.



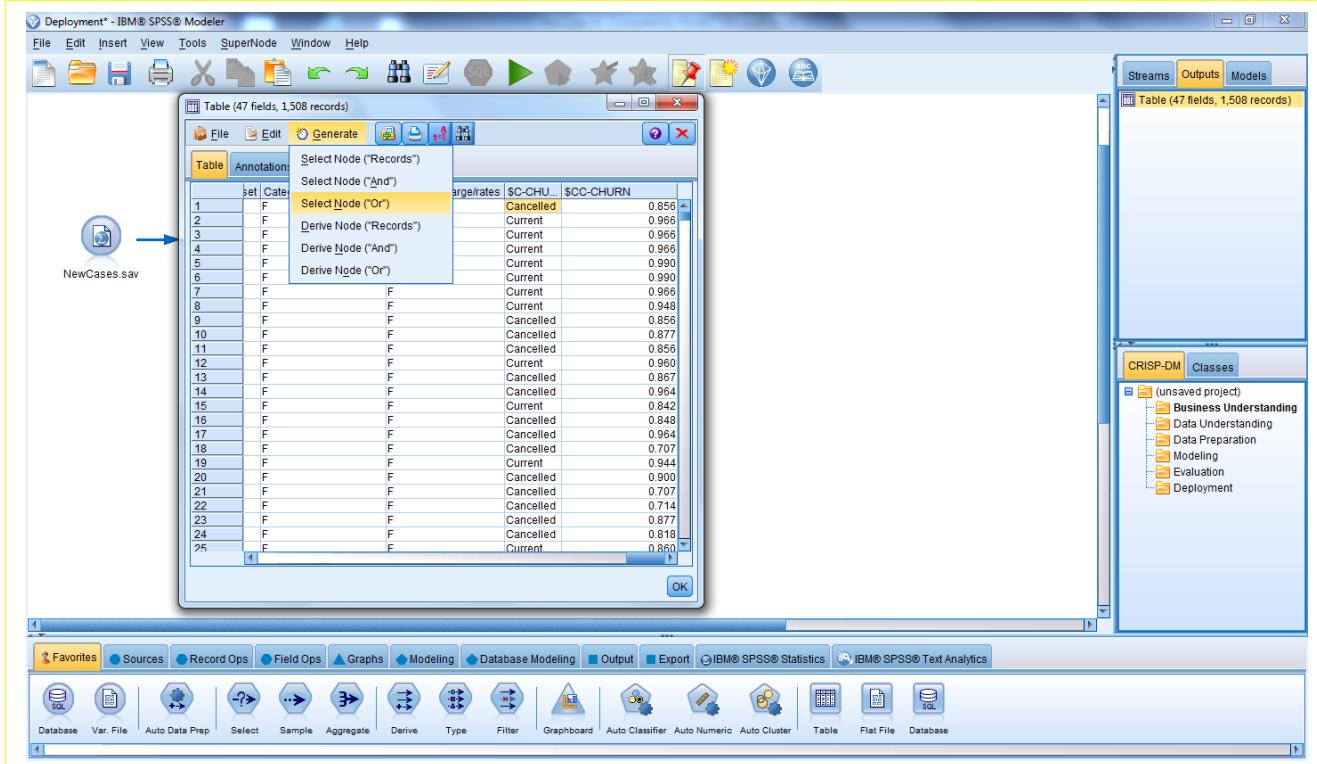
- 6) The new cases have been passed through the Comments model to extract concepts, which were then used as inputs in the C5.0 model to predict Churn.

The last two columns show the predicted outcome and calculated confidence.

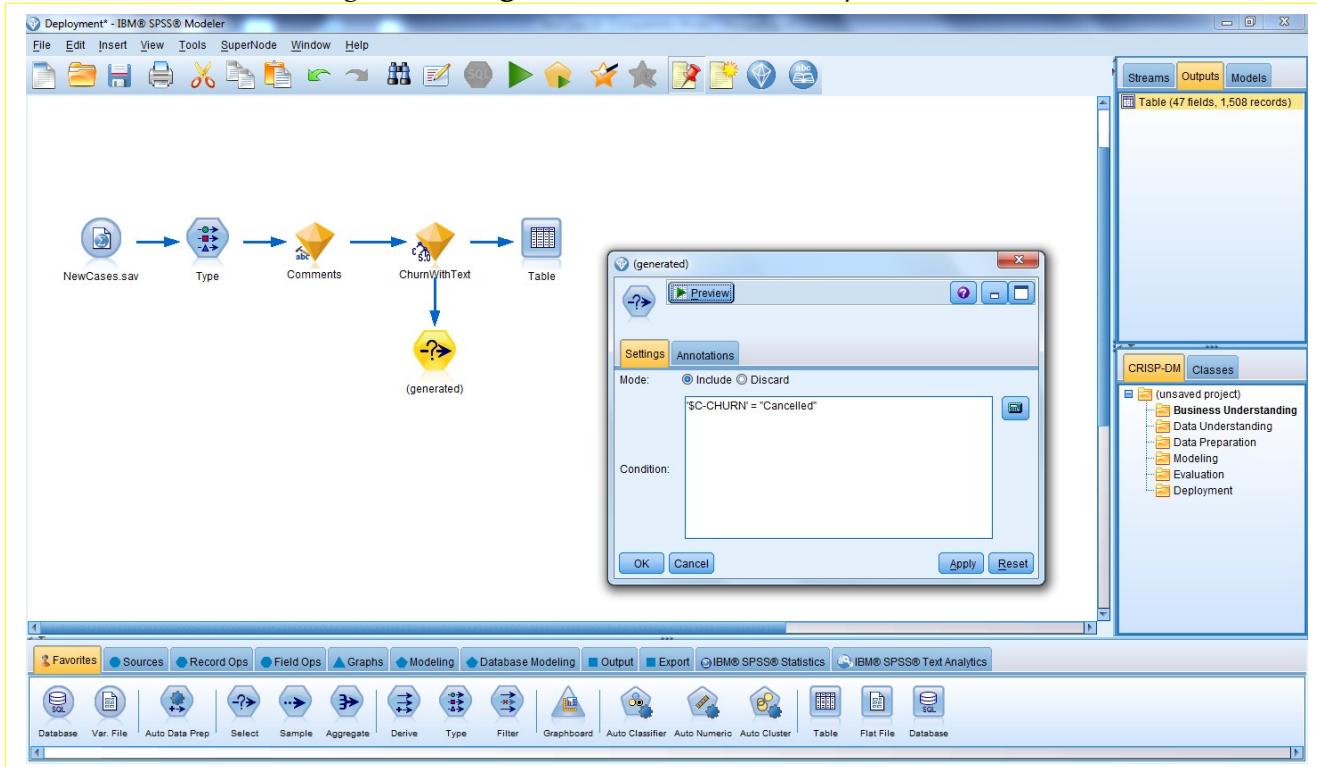
	Category A	Category B	Category C	Category D	Category E	Category F	Category G	Category H	Category I	Category J	Category K	Category L	Category M	Category N	Category O	Category P	Category Q	Category R	Category S	Category T	Category U	Category V	Category W	Category X	Category Y	SC-CHURN	SCC-CHURN
1	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	Cancelled	0.856	
2	F	F	F	F	T	F	F	F	F	T	F	F	T	F	F	F	F	F	F	F	F	F	F	F	Current	0.966	
3	F	F	F	F	T	F	F	F	F	T	F	F	T	F	F	F	F	F	F	F	F	F	F	F	Current	0.966	
4	F	F	F	F	T	F	F	F	F	T	F	F	T	F	F	F	F	F	F	F	F	F	F	F	Current	0.966	
5	F	F	F	F	T	F	F	F	F	T	F	F	T	F	F	F	F	F	F	F	F	F	F	F	Current	0.990	
6	F	F	F	F	T	F	F	F	F	T	F	F	T	F	F	F	F	F	F	F	F	F	F	F	Current	0.990	
7	F	F	F	F	T	F	F	F	F	T	F	F	T	F	F	F	F	F	F	F	F	F	F	F	Current	0.966	
8	F	F	F	F	T	F	F	F	F	T	F	F	T	F	F	F	F	F	F	F	F	F	F	F	Current	0.948	
9	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	Cancelled	0.856	
10	F	F	F	F	F	F	F	F	F	T	F	F	F	F	F	F	F	F	F	F	F	F	F	F	Cancelled	0.877	
11	F	F	F	F	F	F	F	F	F	T	F	F	F	F	F	F	F	F	F	F	F	F	F	F	Cancelled	0.856	
12	F	F	F	F	F	F	F	F	F	T	F	F	F	F	F	F	F	F	F	F	F	F	F	F	Current	0.960	
13	F	F	F	F	F	F	F	F	F	T	F	F	F	F	F	F	F	F	F	F	F	F	F	F	Cancelled	0.867	
14	F	F	F	F	F	F	F	F	F	T	F	F	T	F	F	F	F	F	F	F	F	F	F	F	Cancelled	0.964	
15	F	F	F	F	F	F	F	F	F	T	F	F	T	F	F	F	F	F	F	F	F	F	F	F	Current	0.842	
16	F	F	F	F	F	F	F	F	F	T	F	F	T	F	F	F	F	F	F	F	F	F	F	F	Cancelled	0.848	
17	F	F	F	F	F	F	F	F	F	F	F	F	T	F	F	F	F	F	F	F	F	F	F	F	Cancelled	0.964	
18	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	Cancelled	0.707	
19	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	Current	0.944	
20	F	F	F	F	F	T	F	F	F	T	F	F	F	F	F	F	F	F	F	F	F	F	F	F	Cancelled	0.900	
21	F	F	F	F	F	F	F	F	F	T	F	F	F	F	F	F	F	F	F	F	F	F	F	F	Cancelled	0.707	
22	F	F	F	F	F	F	F	F	F	T	F	F	F	F	F	F	F	F	F	F	F	F	F	F	Cancelled	0.714	
23	F	F	F	F	F	F	F	F	F	T	F	F	F	F	F	F	F	F	F	F	F	F	F	F	Cancelled	0.877	
24	F	F	F	F	F	F	F	F	F	T	F	F	F	F	F	F	F	F	F	F	F	F	F	F	Cancelled	0.819	
25	F	F	F	F	F	F	F	F	F	T	F	F	T	F	F	F	F	F	F	F	F	F	F	F	Current	0.860	
26	F	F	F	F	F	F	F	F	F	T	F	F	F	F	F	F	F	F	F	F	F	F	F	F	Current	0.769	
27	F	F	F	F	F	F	F	F	F	T	F	F	F	F	F	F	F	F	F	F	F	F	F	F	Current	0.842	
28	F	F	F	F	F	F	F	F	F	T	F	F	F	F	F	F	F	F	F	F	F	F	F	F	Current	0.769	
29	F	F	F	F	F	F	F	F	F	T	F	F	F	F	F	F	F	F	F	F	F	F	F	F	Cancelled	0.707	
30	F	F	F	F	F	F	F	F	F	T	F	F	F	F	F	F	F	F	F	F	F	F	F	F	Cancelled	0.964	
31	F	F	F	F	F	F	F	F	F	T	F	F	T	F	F	F	F	F	F	F	F	F	F	F	Cancelled	0.849	
32	F	F	F	F	F	F	F	F	F	T	F	F	T	F	F	F	F	F	F	F	F	F	F	F	Cancelled	0.964	
33	F	F	F	F	F	F	F	F	F	T	F	F	T	F	F	F	F	F	F	F	F	F	F	F	Current	0.966	
34	F	F	F	F	F	F	F	F	F	T	F	F	T	F	F	F	F	F	F	F	F	F	F	F	Cancelled	0.856	
35	F	F	F	F	F	F	F	F	F	T	F	F	T	F	F	F	F	F	F	F	F	F	F	F	Current	0.769	
36	F	F	F	F	F	F	F	F	F	T	F	F	F	F	F	F	F	F	F	F	F	F	F	F	Cancelled	0.867	
37	F	F	F	F	F	F	F	F	F	T	F	F	T	F	F	F	F	F	F	F	F	F	F	F	Cancelled	0.877	
38	F	F	F	F	F	F	F	F	F	T	F	F	F	F	F	F	F	F	F	F	F	F	F	F	Cancelled	0.964	
39	F	F	F	F	F	F	F	F	F	T	F	F	F	F	F	F	F	F	F	F	F	F	F	F	Current	0.796	
40	F	F	F	F	F	F	F	F	F	T	F	F	F	F	F	F	F	F	F	F	F	F	F	F	Cancelled	0.964	
41	F	F	F	F	F	T	F	F	F	T	F	F	T	F	F	F	F	F	F	F	F	F	F	F	Cancelled	0.700	
42	F	F	F	F	T	F	F	F	F	T	F	F	T	F	F	F	F	F	F	F	F	F	F	F	Current	0.990	
43	F	F	F	F	F	F	F	F	F	T	F	F	T	F	F	F	F	F	F	F	F	F	F	F	Cancelled	0.856	
44	F	F	F	F	T	F	F	F	F	T	F	F	T	F	F	F	F	F	F	F	F	F	F	F	Cancelled	0.707	

- 7) Select one of the cells with the value of Cancelled.

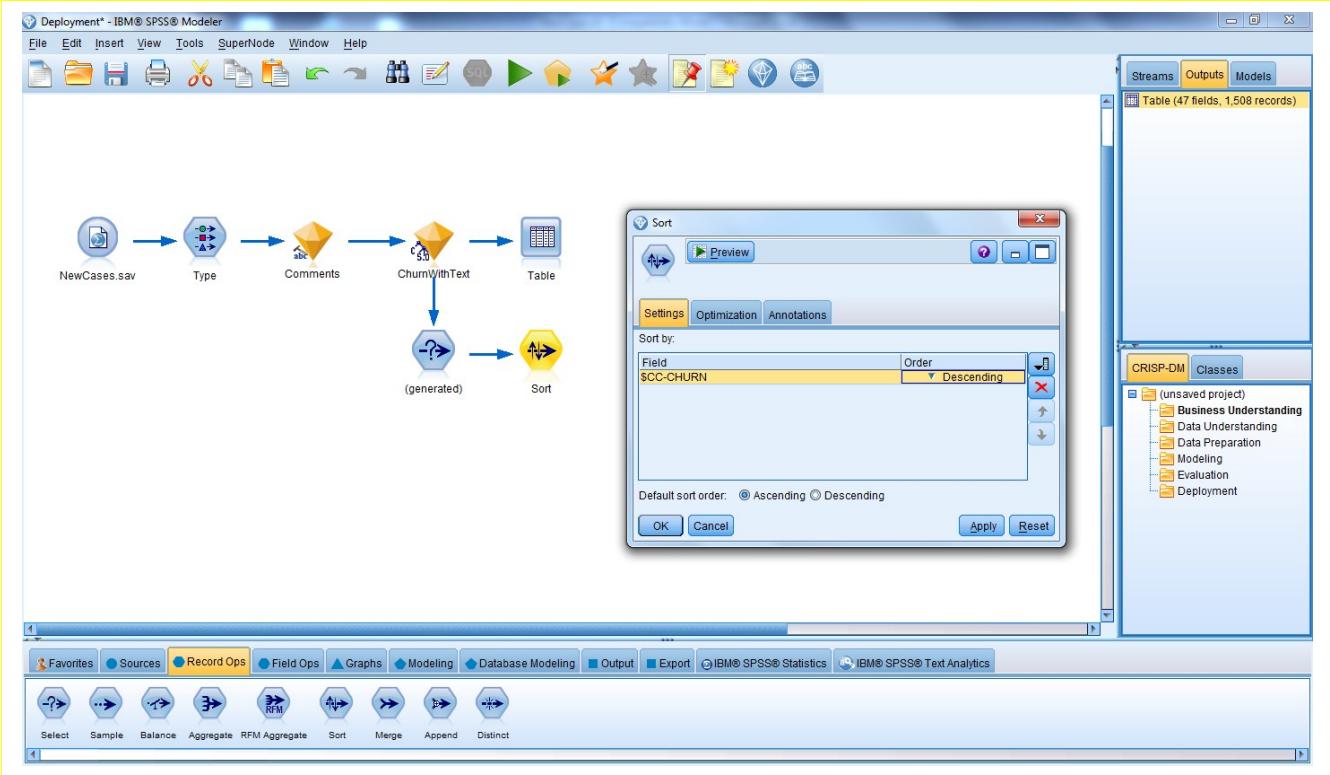
From the drop-down menu, select **Generate** and then select **Select Node ("Or")**.



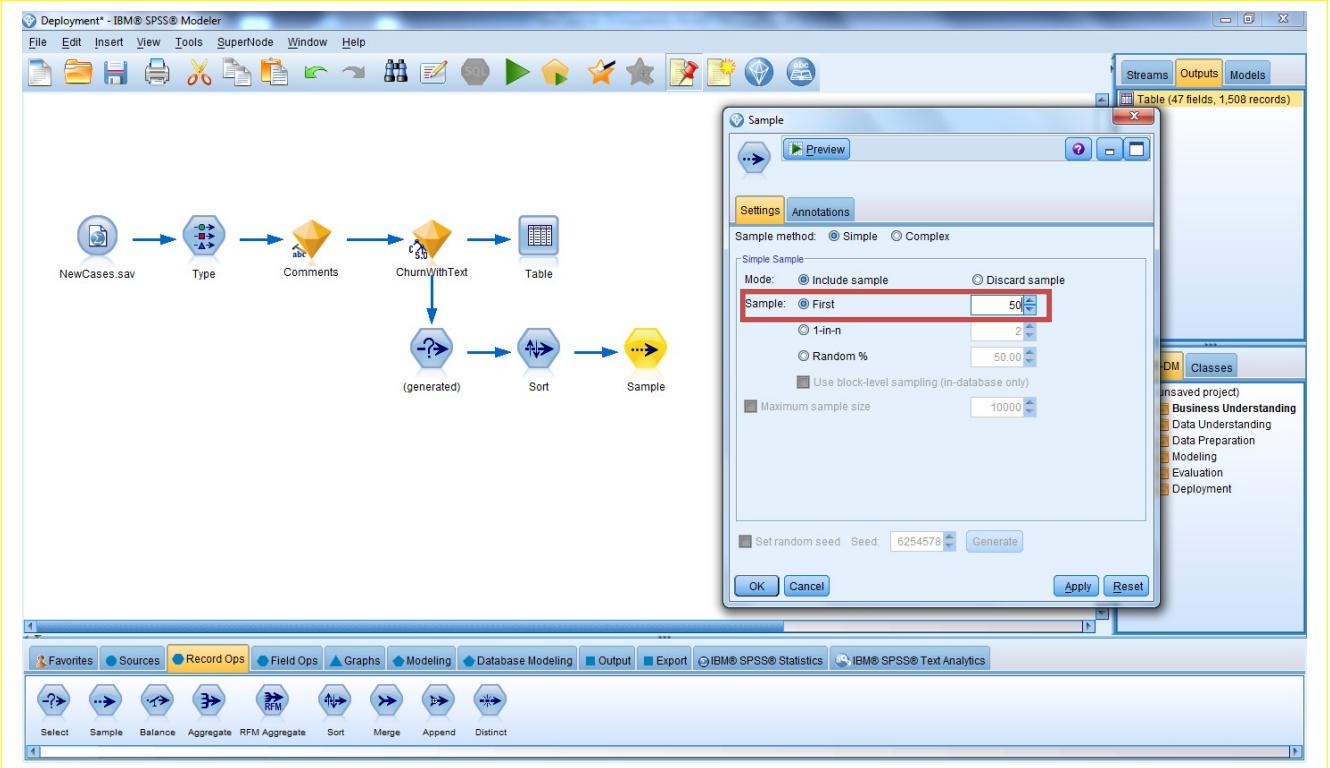
- 8) **Join the Generated node to the ChurnWithText model.**  
**Edit the note, using the Settings tab to select the records predicted to cancel.**



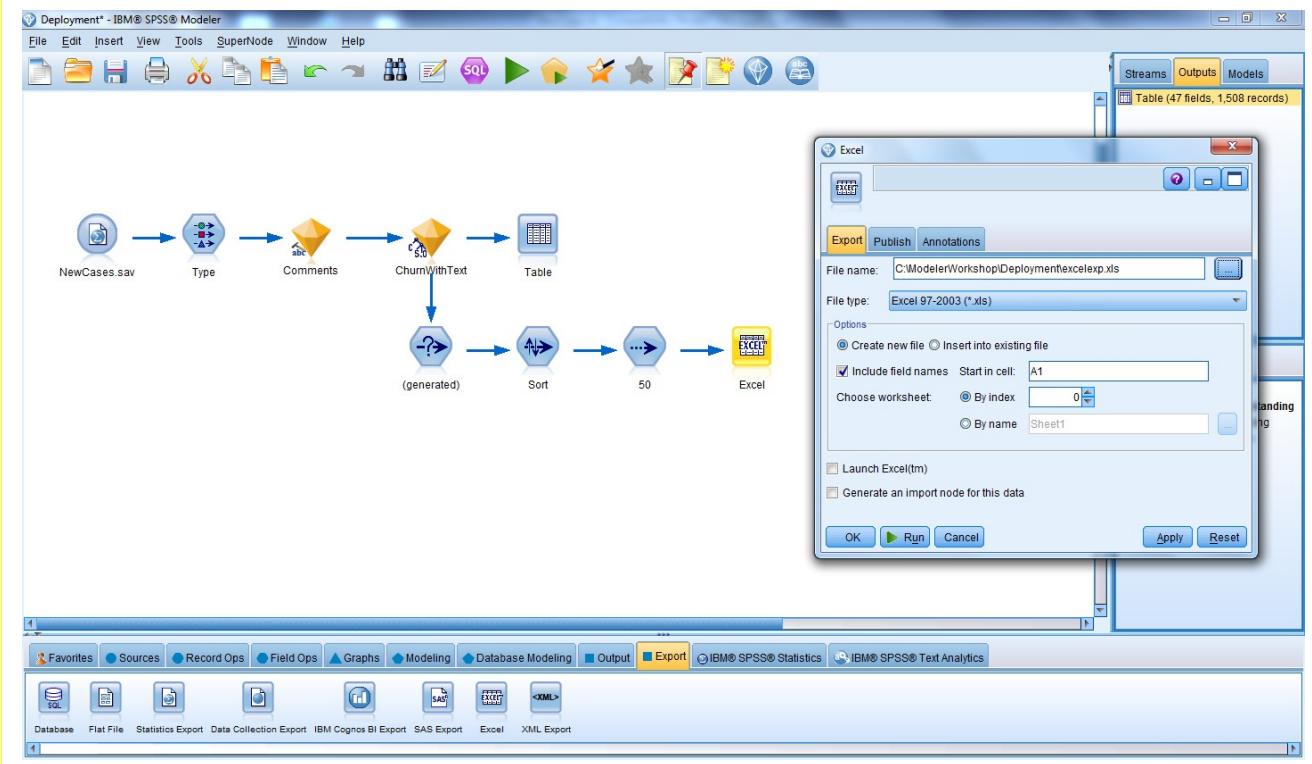
- 9) **Add a Sort node to the Generated node.**  
**Sort the records in descending order.**



10) Add a **Sample** node and edit the sample to be the first 50 records.



11) Finally add the **Excel export** node, use the **Export** tab to change the File name to a local directory and click Run.



12) Open the exported Excel file and confirm 50 records included (row 51 including the header).

The screenshot shows a Microsoft Excel spreadsheet titled 'excel.xls [Compatibility Mode] - Microsoft Excel'. The data consists of 50 rows of information, starting with a header row (row 1) and ending at row 51. The columns are labeled A through V. The data includes various demographic and usage information for individuals, such as ID, Sex, Status, Age, RatePlan, and various service metrics. The last few rows show a series of 'F' values across most columns, likely representing missing data or specific codes used in the dataset.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V
1	ID	Sex	Status	Children	Est_Incom	Car_Own	Usage	Age	RatePlan	LongDist	Internat	Local	Dropped	Paymethc	LocalBltt	LongDist	Comment	Category	Category	Category	Category	
17	13731	M	S	0	8942.7	N	16.91	53.69333	4	3.76	8.13	5.01	0	CC	Budget	Intrn_disco	Asked abo F	F	F	F	F	
18	12900	M	M	0	76161.4	Y	37.55	44.57333	1	11.96	0	25.58	0	CC	FreeLocal	Standard	Wants to ¢ F	F	F	F	F	
19	13723	M	M	1	84166.1	N	15.02	54.01333	1	3.28	0	11.74	1	CC	Budget	Standard	I expected F	F	F	F	F	
20	11332	F	S	0	5185.31	N	52.9	62.05333	3	16.39	5.99	30.51	0	CC	FreeLocal	Intrn_disco	He really tlf F	F	F	F	F	
21	13717	M	M	0	67388	N	96.33	53.12	4	4.79	0.5	91.04	1	CC	Budget	Standard	He expect F	F	F	F	F	
22	10304	M	M	1	75799.9	N	94.06	46.38	3	22.24	7.02	64.8	0	CC	FreeLocal	Intrn_disco	He really tlf F	F	F	F	F	
23	10910	M	S	1	94188.3	N	41.51	57.54667	4	20.86	0.86	19.78	0	CC	Budget	Standard	Wants to F	F	F	F	F	
24	12329	M	S	0	74460.7	Y	206.46	55.34	1	25.53	0.99	179.93	0	CC	FreeLocal	Standard	Cant get h F	F	F	F	F	
25	10266	F	S	0	2121.36	N	38.8	39.78667	3	2.02	0	36.78	1	CC	FreeLocal	Standard	He asked F	F	F	F	F	
26	12083	M	S	1	92647.5	N	32.13	56.04667	1	16.04	5.74	10.34	0	CC	Budget	Standard	Cant get h F	F	F	F	F	
27	10940	M	S	0	89142.7	N	16.91	53.69333	4	3.76	8.13	5.01	0	CC	Budget	Intrn_disco	Asked abo F	F	F	F	F	
28	13704	M	S	1	94188.3	N	41.51	57.54667	2	20.86	0.86	19.78	0	CC	Budget	Standard	Wants to F	F	F	F	F	
29	13703	M	M	0	67388	N	96.33	53.12	3	4.79	0.5	91.04	1	CC	Budget	Standard	He expect F	F	F	F	F	
30	13702	M	M	0	84166.1	N	15.02	54.01333	1	3.28	0	11.74	1	CC	Budget	Standard	I expected F	F	F	F	F	
31	13701	M	M	0	67388	N	96.33	53.12	4	4.79	0.5	91.04	1	CC	Budget	Standard	He expect F	F	F	F	F	
32	10950	M	M	1	66931.9	N	17.46	63.93333	3	11.72	0.39	5.35	0	CC	Budget	Standard	Wants to F	F	F	F	F	
33	10462	F	S	0	15889.3	N	29.17	61.69333	4	3.67	0	25.49	1	CC	Budget	Intrn_disco	Wants to ¢ F	F	F	F	F	
34	13689	M	M	0	84166.1	N	15.02	54.01333	1	3.28	0	11.74	1	CC	Budget	Standard	I expected F	F	F	F	F	
35	13688	M	M	1	67388	N	96.33	53.12	4	4.79	0.5	91.04	1	CC	Budget	Standard	He expect F	F	F	F	F	
36	10195	F	S	1	12525.3	N	40.59	53.05333	2	11.72	0	28.86	0	CC	FreeLocal	Intrn_disco	He wanted F	F	F	F	F	
37	11870	M	M	1	80000	Y	40.9	53.21333	3	15.27	0	25.62	0	CC	FreeLocal	Standard	transferred F	F	F	F	F	
38	13681	M	M	1	84166.1	N	15.02	54.01333	1	3.28	0	11.74	1	CC	Budget	Standard	I expected F	F	F	F	F	
39	13680	M	M	1	67388	N	96.33	53.12	3	4.79	0.5	91.04	1	CC	Budget	Standard	He expect F	F	F	F	F	
40	10191	M	M	0	92414.7	Y	40.06	45.51333	4	29.43	0	10.62	0	CC	FreeLocal	Standard	Hopes to ¢ F	F	F	F	F	
41	10974	F	S	0	15889.3	N	29.17	61.69333	3	3.67	0	25.49	1	CC	Budget	Intrn_disco	Wants to ¢ F	F	F	F	F	
42	10180	M	M	0	76161.4	Y	37.55	44.57333	4	11.96	0	25.58	0	CC	FreeLocal	Standard	Wants to ¢ F	F	F	F	F	
43	12315	F	S	1	13576.5	N	40.49	39.42667	1	14.83	0	25.66	0	CC	Budget	Standard	Asked how F	F	F	F	F	
44	11466	F	S	0	15889.3	N	29.17	61.69333	3	3.67	0	25.49	1	CC	Budget	Intrn_disco	Wants to ¢ F	F	F	F	F	
45	12327	M	S	1	92647.5	N	32.13	56.04667	4	16.04	5.74	10.34	0	CC	Budget	Standard	Cant get h F	F	F	F	F	
46	10174	M	M	1	63437.3	N	14.98	63.65333	4	0.68	3.35	10.94	0	CC	Budget	Intrn_disco	Asking for F	F	F	F	F	
47	11844	F	S	0	5185.31	N	52.9	62.05333	1	16.39	5.99	30.51	0	CC	FreeLocal	Intrn_disco	He really tlf F	F	F	F	F	
48	11645	M	S	1	92647.5	N	32.13	56.04667	4	16.04	5.74	10.34	0	CC	Budget	Standard	Cant get h F	F	F	F	F	
49	13641	M	M	0	92414.7	Y	40.06	45.51333	4	29.43	0	10.62	0	CC	FreeLocal	Standard	Hopes to ¢ F	F	F	F	F	
50	13632	M	M	1	76161.4	Y	37.55	44.57333	4	11.96	0	25.58	0	CC	FreeLocal	Standard	Wants to ¢ F	F	F	F	F	
51	12862	M	S	0	74460.7	Y	206.46	55.34	3	25.53	0.99	179.93	0	CC	FreeLocal	Standard	Cant get h F	F	F	F	F	
52																						
53																						

13) Save the Stream to a local directory before closing it.