Banking Services Case Study: Performing Association Analysis

By : Katragaddda Harika, Sonika Rajan

Market Basket Analysis

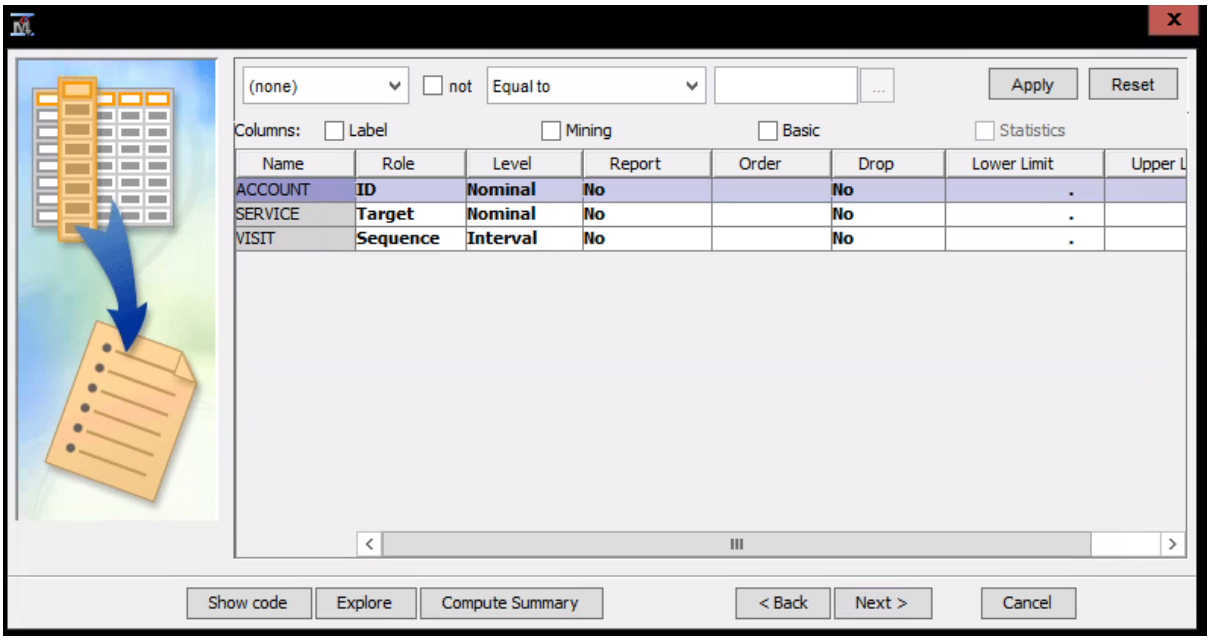
Created a new analysis diagram and data source for the **BANK** data set.

Property	Value
ID	EMWS3
Name	Association_analysis
Status	Open
Notes	
History	
Create Date	12/4/17 11:35 PM
Encoding	wlatin1 Western (Windows)
Data Representation	WINDOWS_64
Native OS	Yes

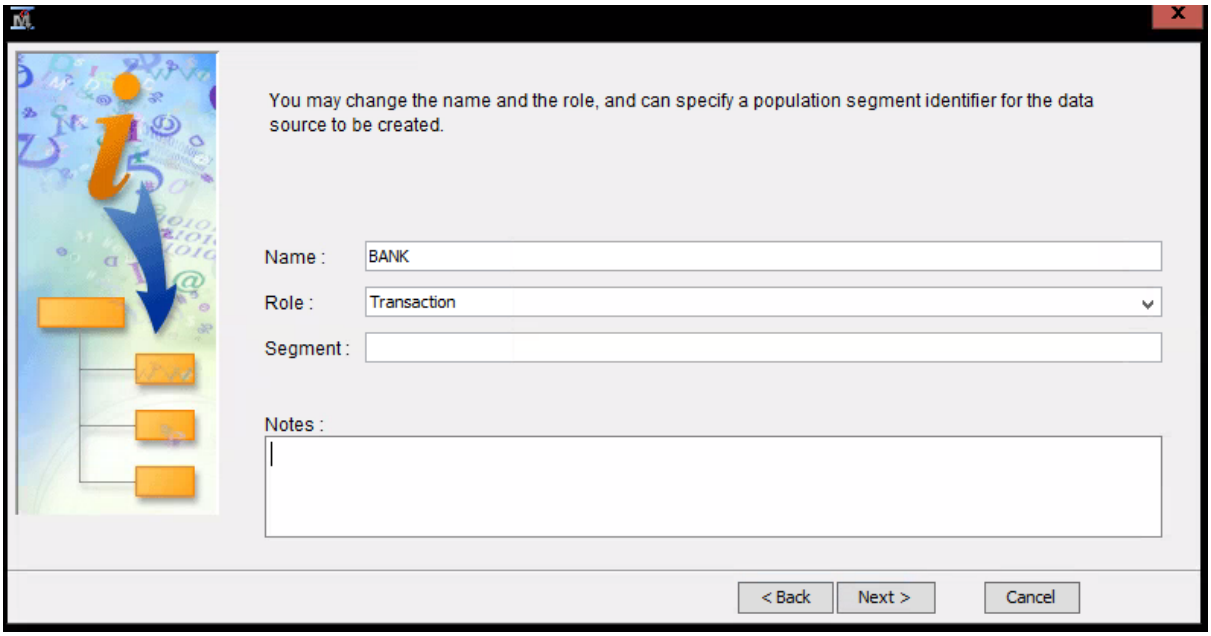
Roles for the variables are changed as: Account -> ID, SERVICE -> Target.

Also, as association analysis requires exactly one target variable and at least one ID variable we changed 'level of Account' and 'service' to Nominal. Both must have a nominal measurement level.

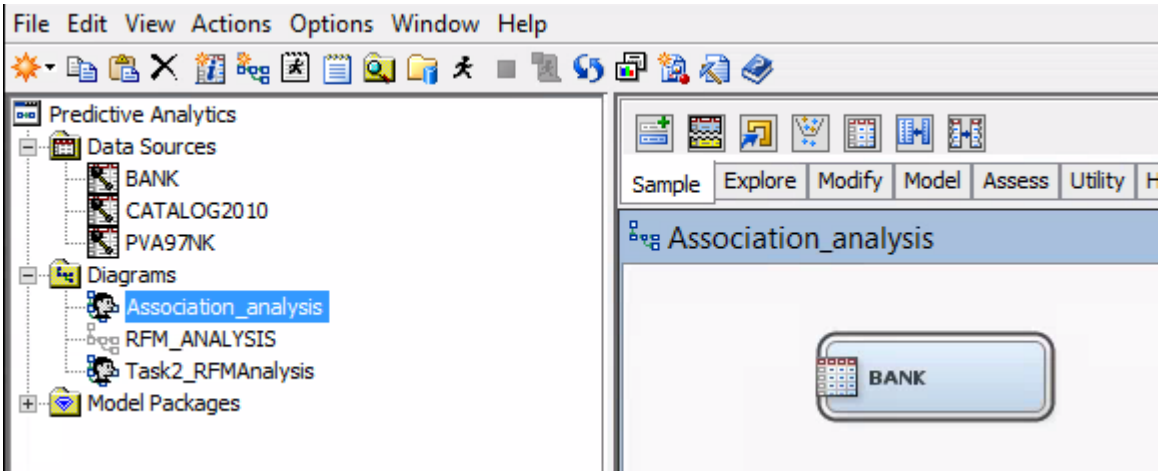
For sequence analysis we changed the role of VISIT -> Sequence



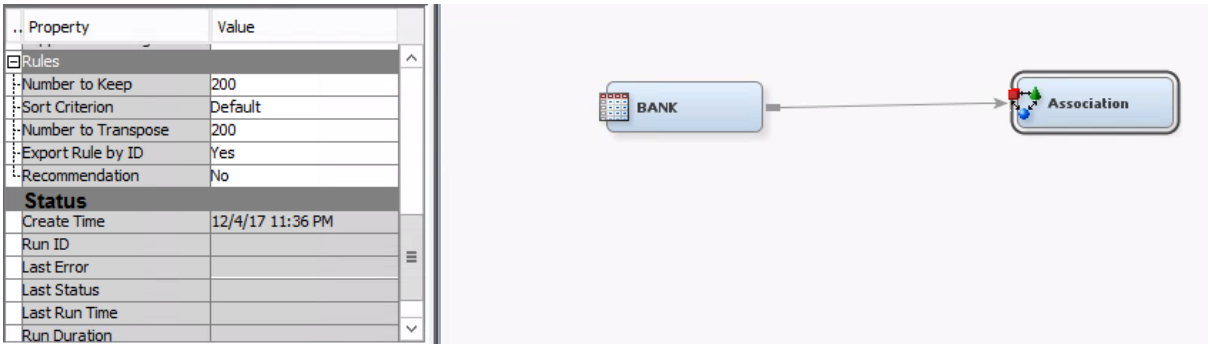
For an association analysis, the data source should have a role of Transaction so changed it to transaction.



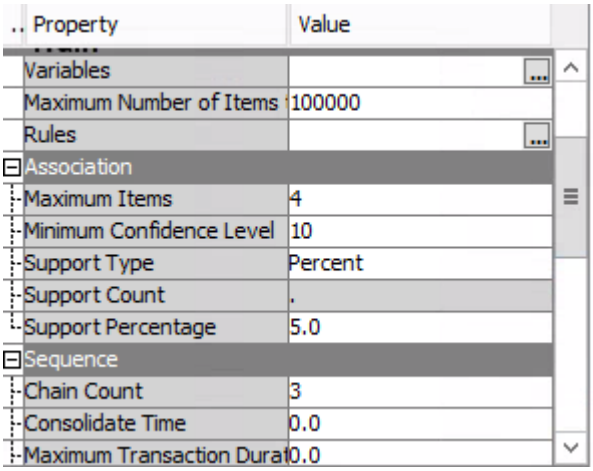
Created the Bank data set and it is dragged into the Association_analysis diagram:



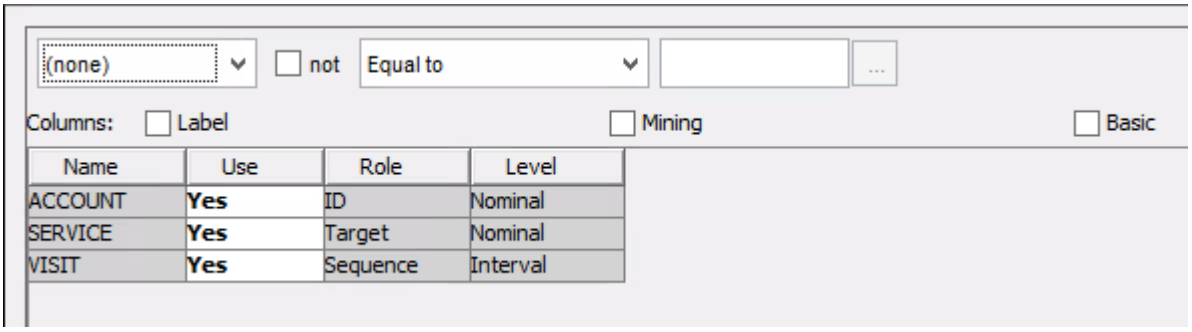
Association tool is dragged into the Associations Analysis diagram workspace. And the node is connected to the BANK database as we are carrying out association analysis.
The Export Rule by ID property determines whether the **Rule-by-ID** data is exported from the node and whether the Rule Description table is available for display in the Results window. Set the value for Export Rule by ID to **Yes**.



Default settings are used for the confidence, support and Maximum items.

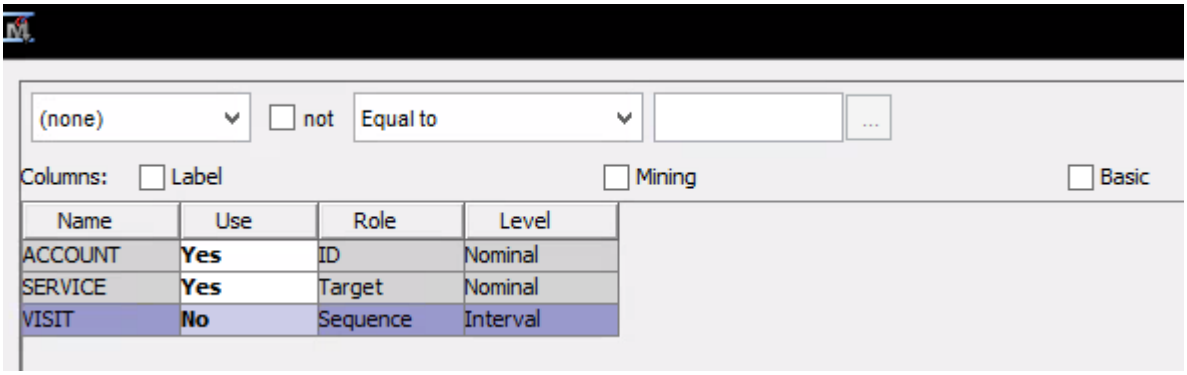


Variables can be seen in the dialog box for association node.

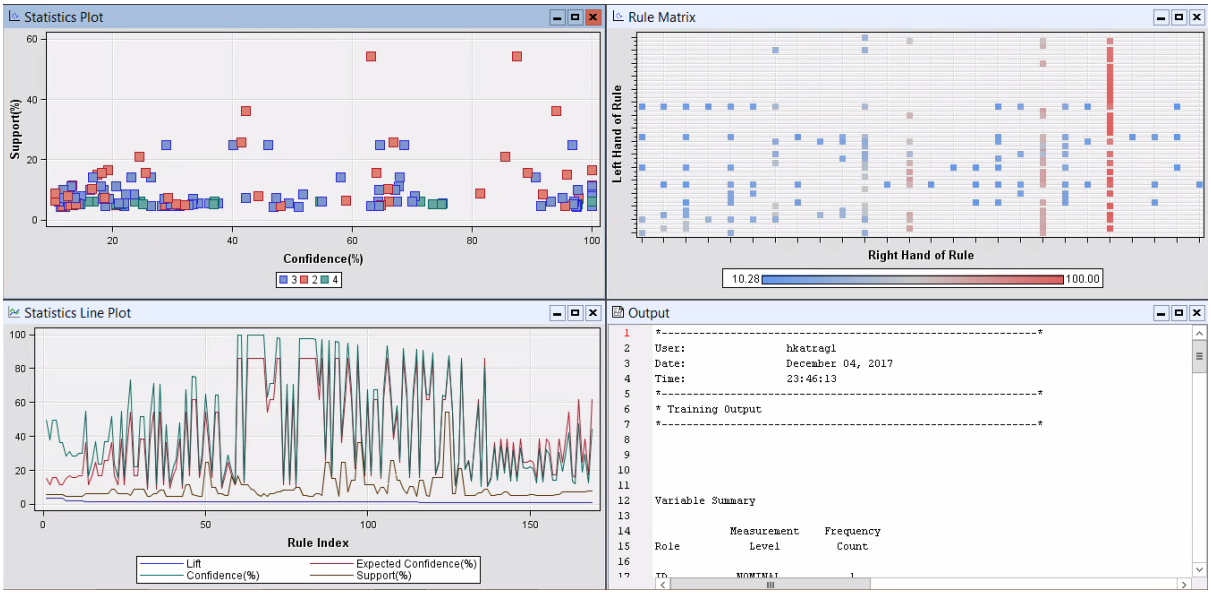


As we first want to perform a market basket analysis, we do not need the sequence variable.

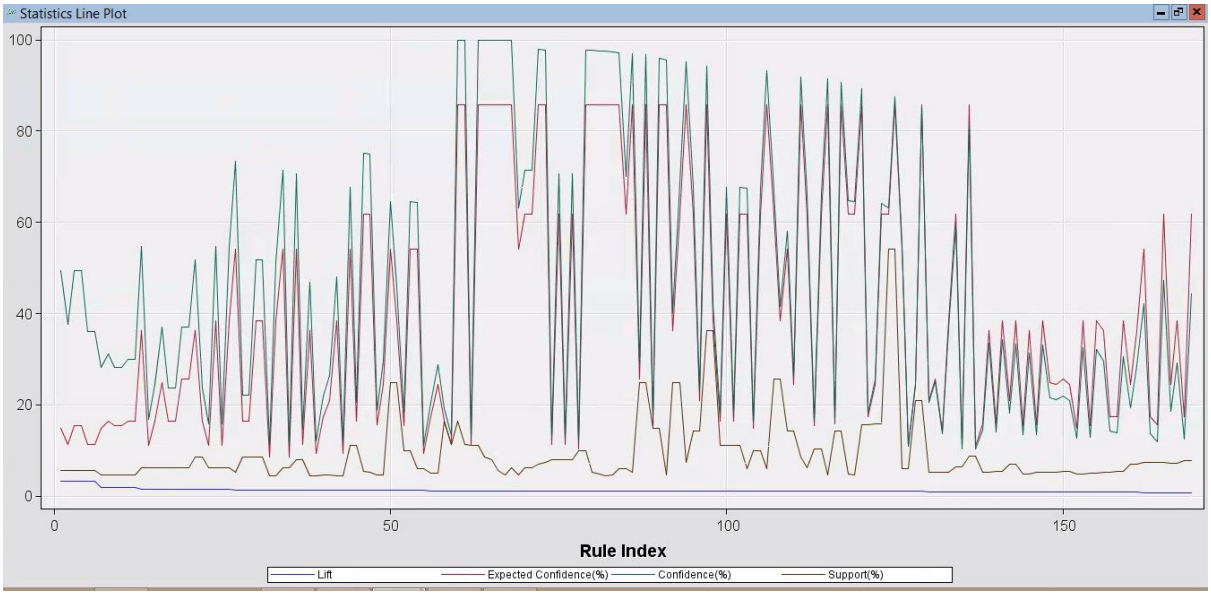
So we set USE = NO for the VISIT Variable.



Run the diagram for the Association node and view the results.
The Results -Node: Association Diagram window appears with the Statistics Plot, Statistics Line Plot, Rule Matrix, and Output windows visible.



The statistical line plot is generated as below:



Rule descriptions are viewed from View -> Rules -> Description. 169 Rules are generated.

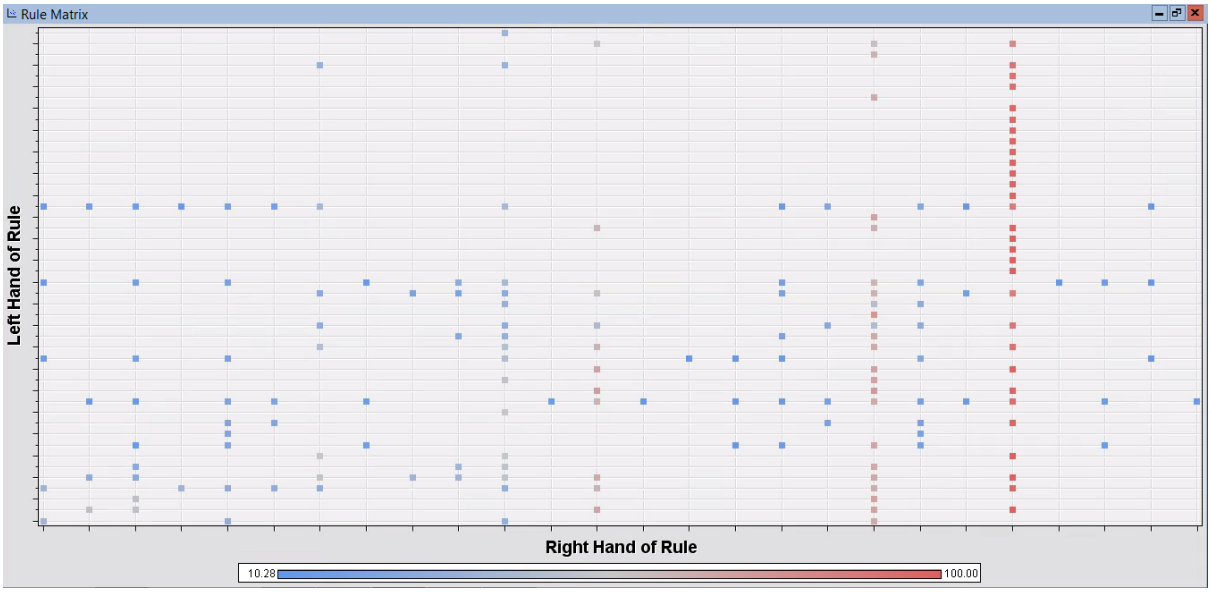
Rule Description

Map	Rule
RULE1	CKING & CCRD ==> CKCRD
RULE2	CKCRD ==> CKING & CCRD
RULE3	CKCRD ==> CCRD
RULE4	CKING & CKCRD ==> CCRD
RULE5	CCRD ==> CKCRD
RULE6	CCRD ==> CKING & CKCRD
RULE7	HMEQLC ==> CKING & CCRD
RULE8	CKING & CCRD ==> HMEQLC
RULE9	HMEQLC ==> CCRD
RULE10	HMEQLC & CKING ==> CCRD
RULE11	CCRD ==> HMEQLC
RULE12	CCRD ==> HMEQLC & CKING
RULE13	SVG & HMEQLC ==> CKING & ATM
RULE14	CKING & ATM ==> SVG & HMEQLC
RULE15	SVG & CKING & ATM ==> HMEQLC
RULE16	HMEQLC ==> SVG & CKING & ATM
RULE17	SVG & ATM ==> HMEQLC
RULE18	SVG & ATM ==> HMEQLC & CKING
RULE19	HMEQLC ==> SVG & ATM
RULE20	HMEQLC & CKING ==> SVG & ATM
RULE21	HMEQLC ==> CKING & ATM
RULE22	CKING & ATM ==> HMEQLC
RULE23	SVG & HMEQLC ==> ATM
RULE24	SVG & HMEQLC & CKING ==> ATM
RULE25	ATM ==> SVG & HMEQLC
RULE26	ATM ==> SVG & HMEQLC & CKING
RULE27	CD & ATM ==> SVG & CKING
RULE28	ATM ==> HMEQLC
RULE29	ATM ==> HMEQLC & CKING
RULE30	HMEQLC ==> ATM
RULE31	HMEQLC & CKING ==> ATM
RULE32	CKING & AUTO ==> ATM
RULE33	ATM ==> CKING & AUTO
RULE34	HMEQLC & ATM ==> SVG & CKING
RULE35	SVG & CKING ==> HMEQLC & ATM
RULE36	CKCRD ==> SVG & CKING
RULE37	SVG & CKING ==> CKCRD

Rule 1 has the highest lift i.e. checking, and credit card implies check card as rules are ordered in descending order of lift (also as seen in statistical line plot). This is not surprising given that many check cards include credit card logos. As lift is symmetric we can notice that there is symmetry in rules 1 and 2.

One of the higher lift rules is that a home equity line of credit (LOC) implies checking and check card (And vice versa) – Rule 7 as seen in the above list. Generally the customers with a home equity LOC, who do not already have a checking account, should be offered a checking account and check card with a special promotion.
The rule matrix plots the rules based on the items on the left side and right side of the rule. Based on the confidence of the rules, the points are colored.

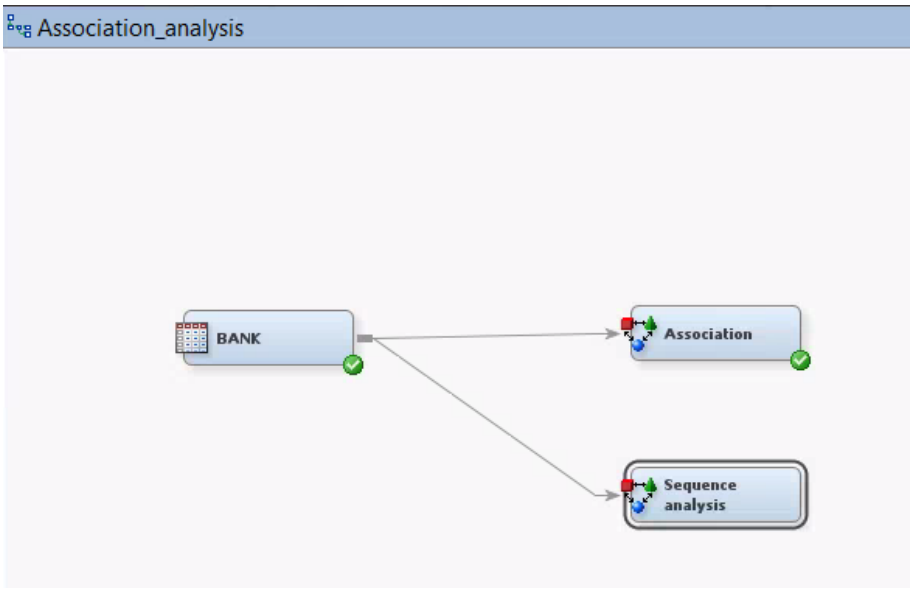
For example, the rules with the highest confidence are in the column in the picture above. Using the interactive feature of the graph, we can view that these rules are having checking on the right side of the rule.



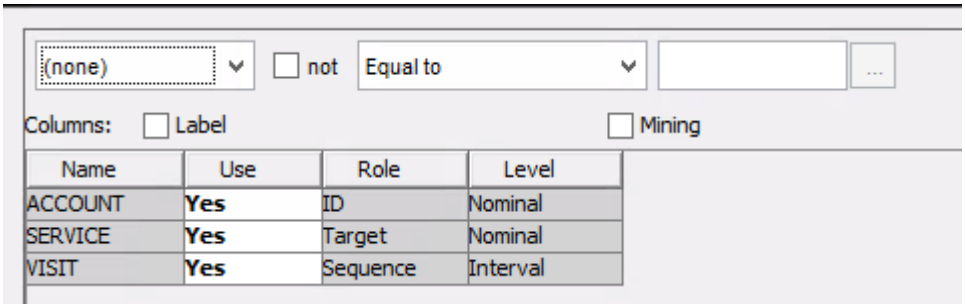
SEQUENCE ANALYSIS

In addition to the products owned by its customers, the bank is interested in examining the order in which the products are purchased to help with a best-next-offer (up-sell) campaign. The sequence variable in the data set i.e. VISIT enables you to conduct a sequence analysis.

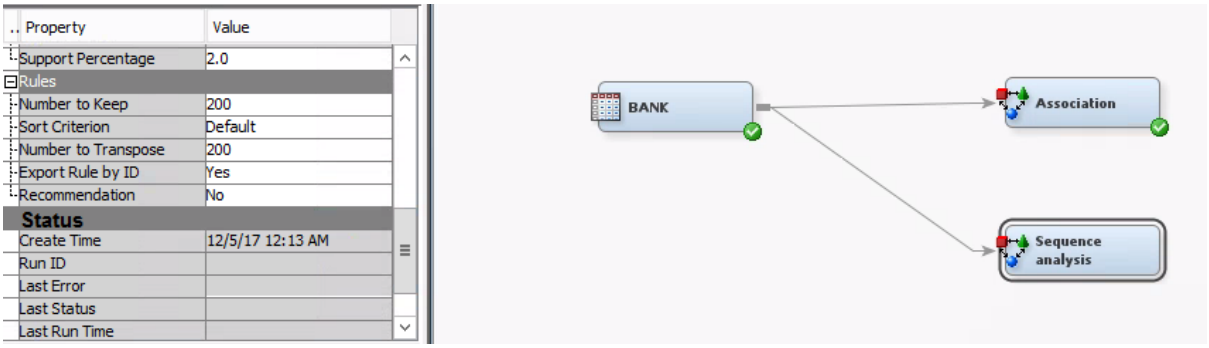
New association node created in the Associations Analysis diagram workspace and connected to BANK data source to conduct sequence analysis.



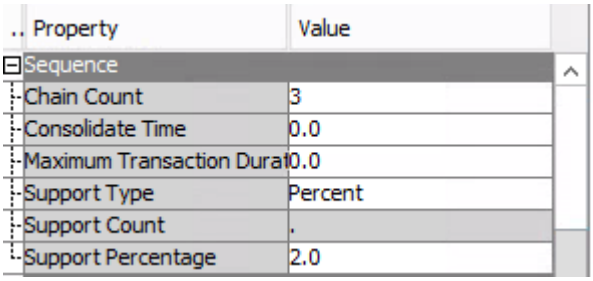
All variables have the USE value – Yes.



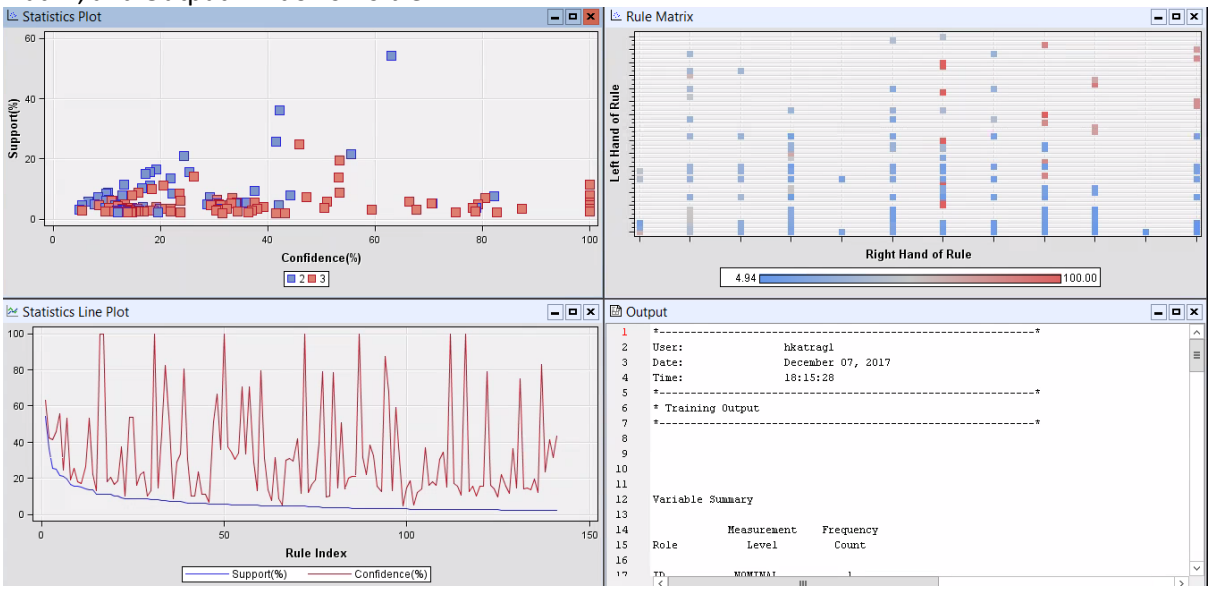
The diagram showing the final connections:



Sequence section in the Properties panel – Default properties are used for Chain Count, Consolidate Time, Maximum Transaction Duration, Support Type, Support Count and Support Percentage



Run node -> Association Diagram window appears with the Statistics Plot, Statistics Line Plot, Rule Matrix, and Output windows visible.

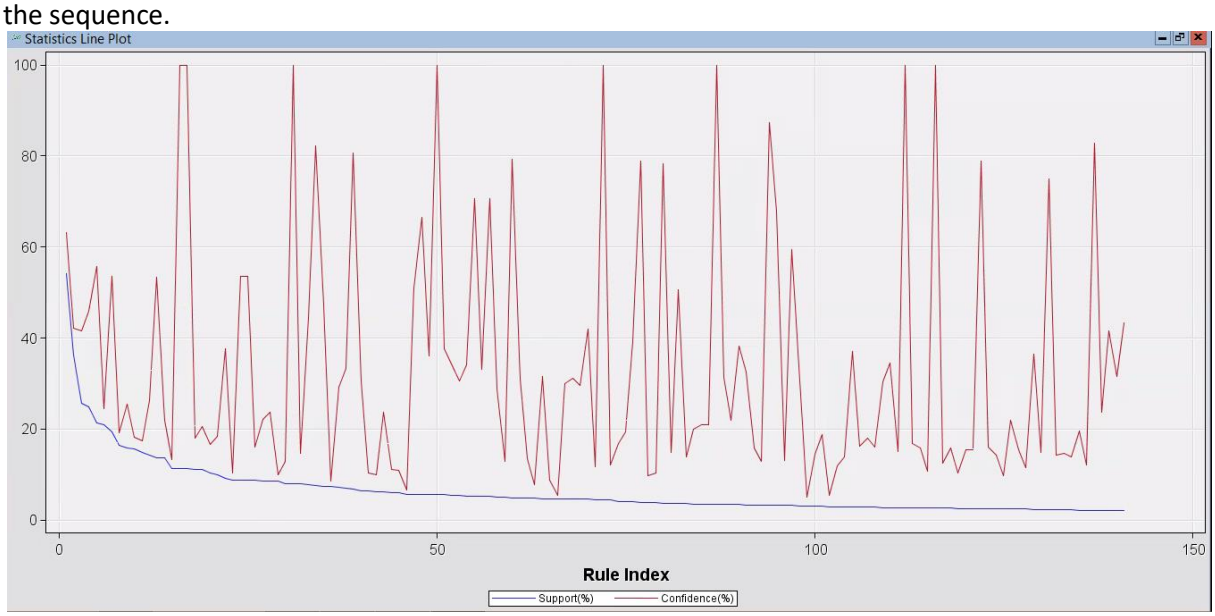


In case of sequence analysis as per statistical plot, the rules are ordered in the descending order of support percentage.

The statistics line plot graphs the support and confidence for each of the rules by rule index number.

The support percent is (transaction count / total number of customers), which would be the maximum transaction count.

The confidence percent is the transaction count divided by the transaction count for the left side of the sequence.



Rule Description

Output

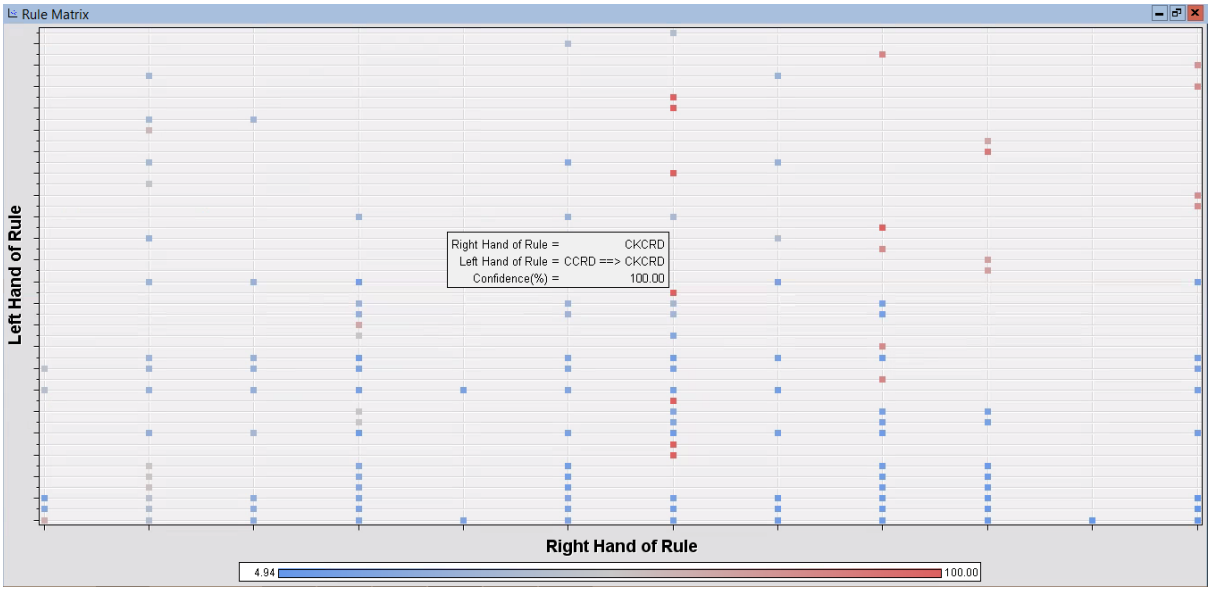
map ▲	RULE
RULE1	CKING ==> SVG
RULE10	CKING ==> MMDA
RULE100	CKING ==> CD ==> TRUST
RULE101	SVG ==> CD ==> CCRD
RULE102	CKING ==> SVG ==> TRUST
RULE103	CD ==> HMEQLC
RULE104	CKING ==> CD ==> HMEQLC
RULE105	MMDA ==> SVG ==> ATM
RULE106	MMDA ==> HMEQLC
RULE107	CKING ==> MMDA ==> HMEQLC
RULE108	MMDA ==> MMDA
RULE109	CKING ==> IRA ==> IRA
RULE11	CKING ==> CCRD
RULE110	MMDA ==> SVG ==> CD
RULE111	MMDA ==> CKCRD
RULE112	MMDA ==> CKCRD ==> CKCRD
RULE113	CKING ==> MMDA ==> CKCRD
RULE114	CD ==> CKCRD
RULE115	HMEQLC ==> MTG
RULE116	CKING ==> CD ==> CKCRD
RULE117	CD ==> CKCRD ==> CKCRD
RULE118	CKING ==> HMEQLC ==> MTG
RULE119	CD ==> AUTO
RULE12	CKING ==> SVG ==> CD
RULE120	HMEQLC ==> AUTO
RULE121	CKING ==> HMEQLC ==> AUTO
RULE122	CD ==> TRUST ==> TRUST
RULE123	SVG ==> CD ==> IRA
RULE124	MMDA ==> TRUST
RULE125	SVG ==> ATM ==> MTG
RULE126	SVG ==> HMEQLC ==> CKCRD
RULE127	CKING ==> CD ==> AUTO
RULE128	CKING ==> MMDA ==> TRUST
RULE129	SVG ==> IRA ==> ATM
RULE13	SVG ==> ATM ==> ATM
RULE130	SVG ==> CD ==> TRUST
RULE131	SVG ==> TRUST ==> TRUST

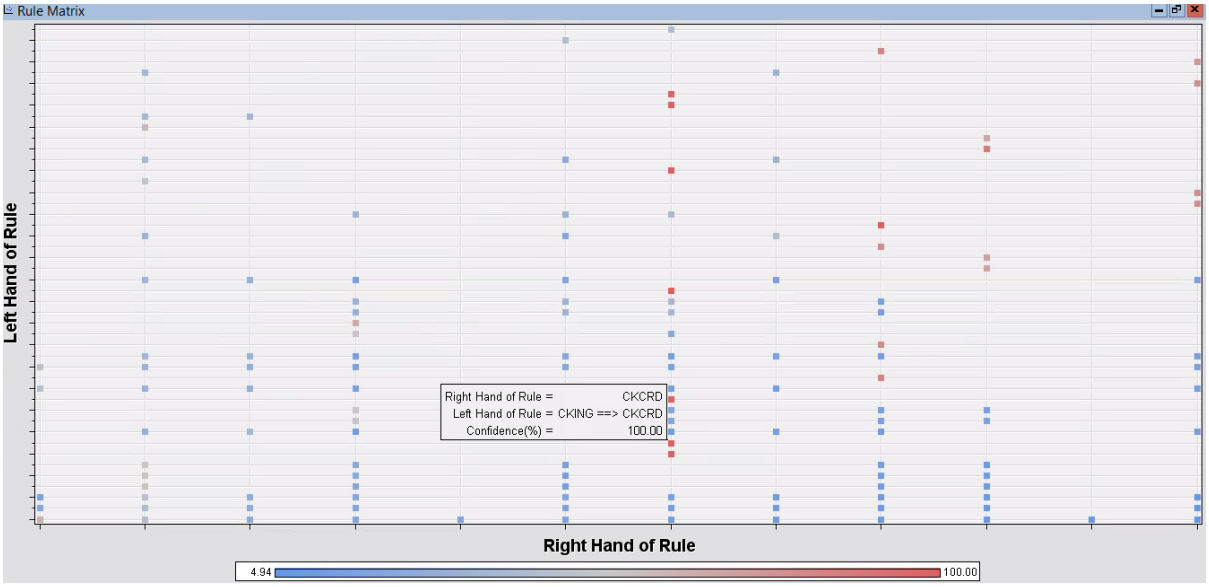
Rule 1 i.e. Checking -> Savings has the highest support percentage

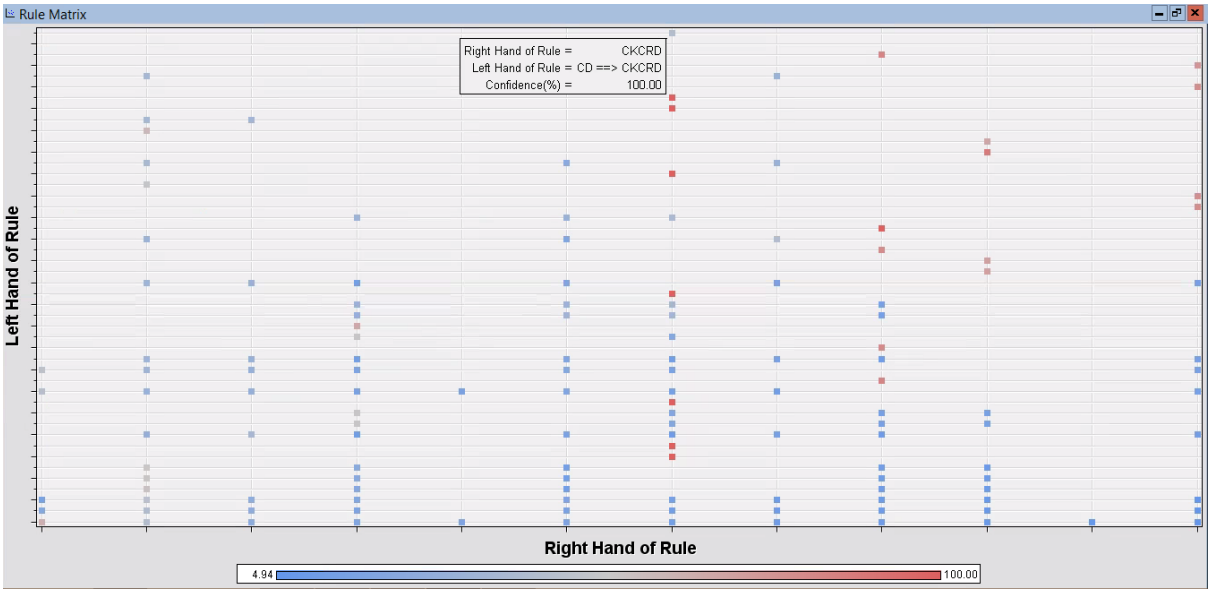
For most of the rules the confidence changes after the order of service acquisition is considered. For example, from the rule description above, if a customer already has checking and savings, they are likely to get an ATM card next.

The rule matrix plots the rules based on the items on the left side of the rule and the items on the right side of the rule. Based on the confidence of the rules, the points are colored.

In the rule matrix generated in sequence analysis, all the sequences with 100% Confidence are having the Check Card on the right hand of the rule.


Rule Matrix
Left Hand of Rule
Right Hand of Rule = CKCRD
Left Hand of Rule = CCRD ==> CKCRD
Confidence(%) = 100.00
Right Hand of Rule
4.94
100.00


Rule Matrix
Left Hand of Rule
Right Hand of Rule = CKCRD
Left Hand of Rule = CKING ==> CKCRD
Confidence(%) = 100.00
Right Hand of Rule
4.94
100.00


Rule Matrix
Left Hand of Rule
Right Hand of Rule = CKCRD
Left Hand of Rule = CD ==> CKCRD
Confidence(%) = 100.00
Right Hand of Rule
4.94
100.00