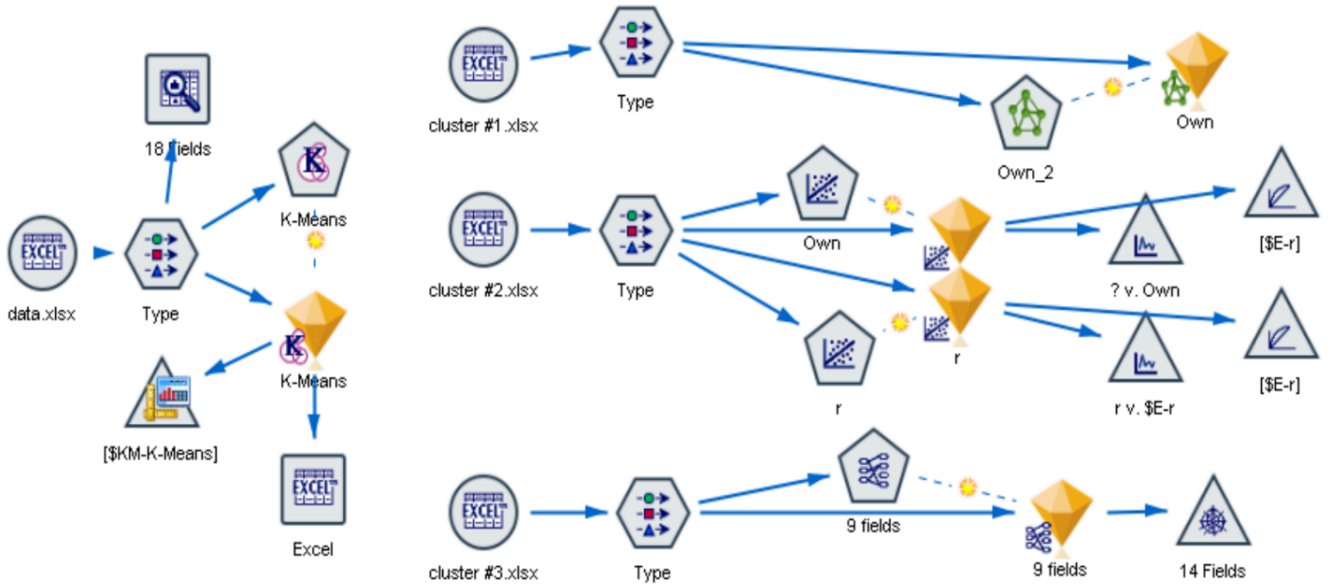
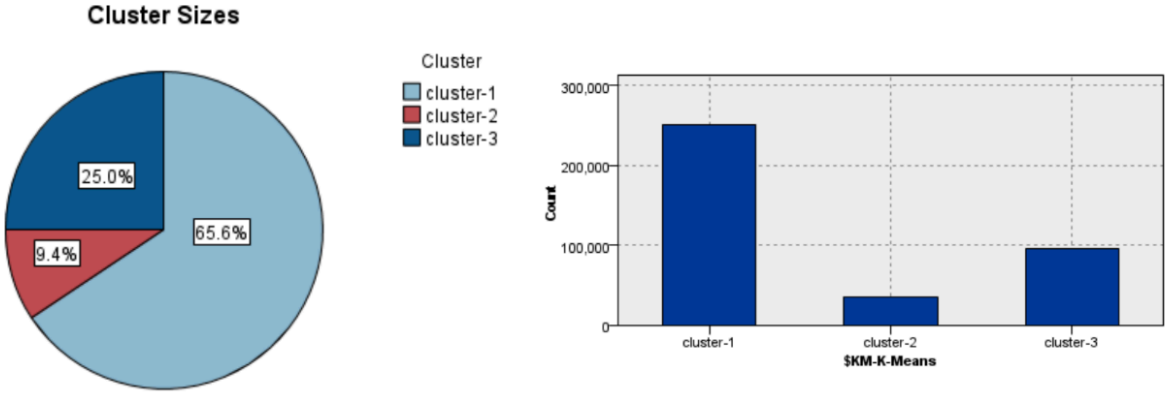


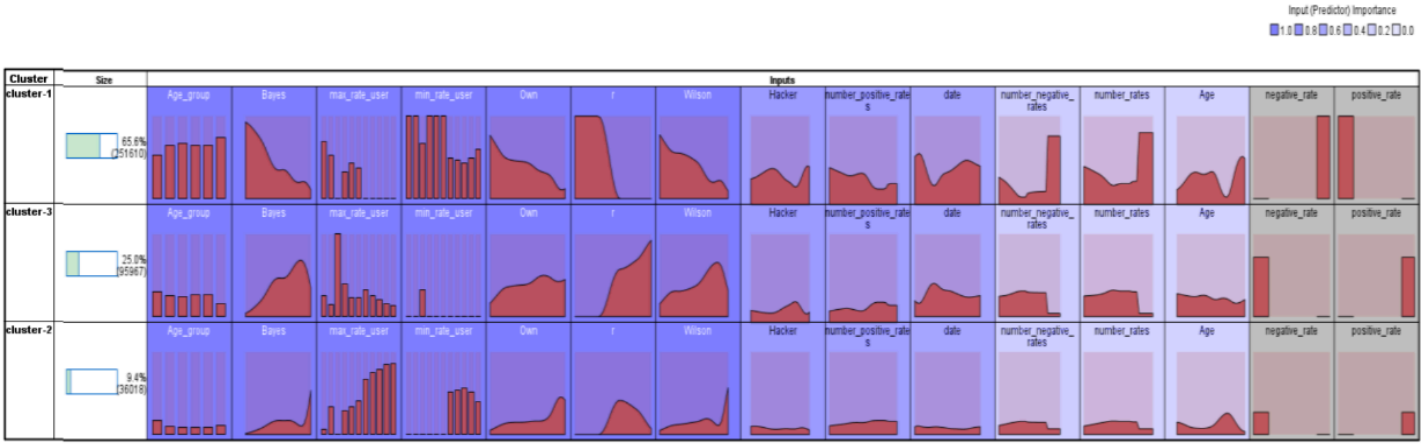
A flow of model building tools in IBM SPSS Modeler

Field	Sample Graph	Measurement	Min	Max	Mean	Std. Dev	Skewness	Unique	Valid
number_negative_rates		Continuous	0.000	2004.000	33.858	127.268	12.984	--	383595
number_rates		Continuous	1.000	2264.000	53.001	162.099	9.718	--	383595
date		Continuous	1901.000	2004.000	1995.310	7.424	-1.875	--	383595
Wilson		Continuous	0.000	0.716	0.126	0.113	0.835	--	383595
Bayes		Continuous	-0.056	1.230	0.313	0.158	0.501	--	383595
Hacker		Continuous	1.000	10.198	2.908	1.112	0.603	--	383595
Own		Continuous	0.001	9.831	1.234	0.989	1.785	--	383595
Country		Nominal	--	--	--	--	--	131	383595

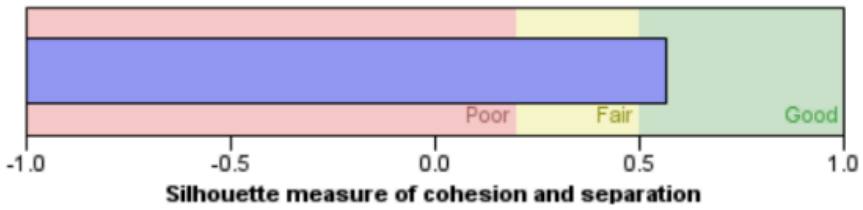
Features of variables using Data Audit

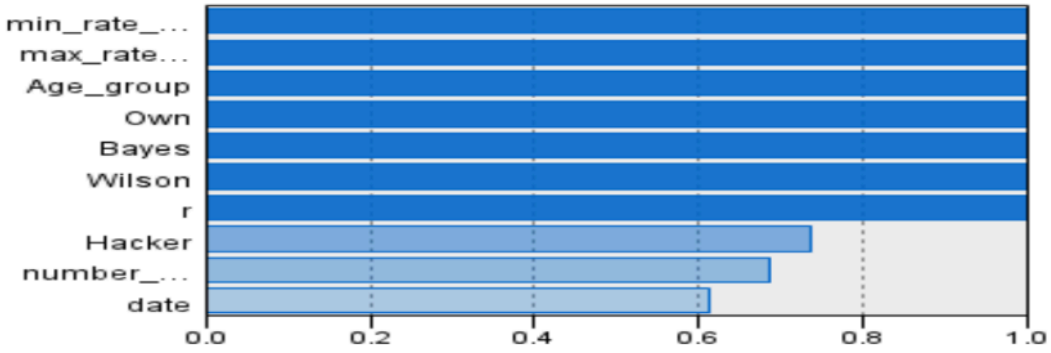


Quantitative distribution of clusters

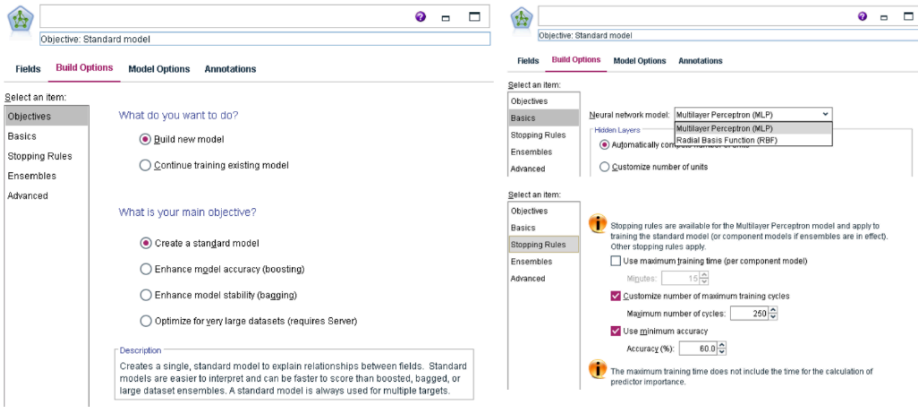


Characteristics of cluster indicators according to K-means





Importance of K-means predicates

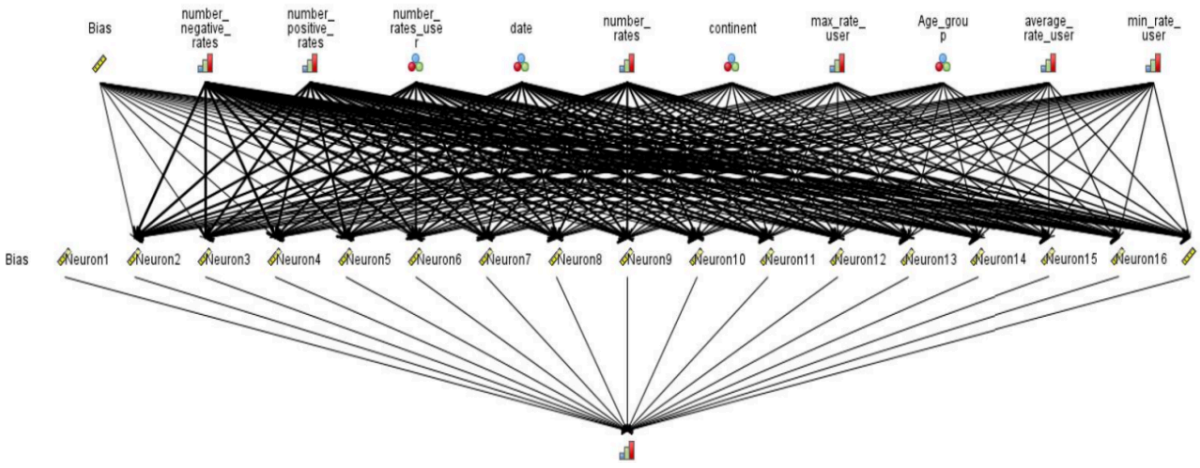


Neural network tab options

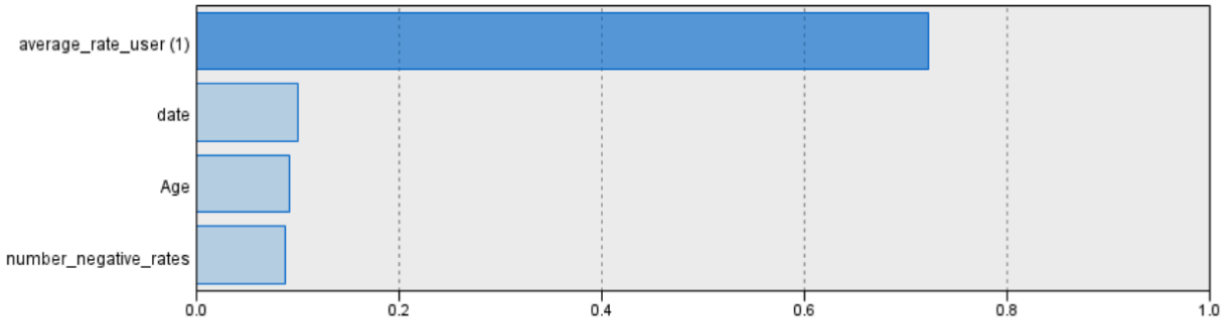
Observed	Predicted														
	0.000	1.000	2.000	3.000	4.000	5.000	6.000	7.000	8.000	9.000	10.000	11.000	12.000	13.000	15.000
0.000	0.0%	0.0%	8.7%	66.9%	24.4%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
1.000	0.0%	0.0%	11.5%	59.4%	28.8%	0.3%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
2.000	0.0%	0.0%	13.7%	54.7%	30.6%	0.9%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
3.000	0.0%	0.0%	7.5%	50.1%	37.0%	4.9%	0.4%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
4.000	0.0%	0.0%	3.7%	38.9%	47.3%	7.7%	2.5%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
5.000	0.0%	0.0%	1.9%	26.0%	40.2%	21.1%	10.7%	0.1%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
6.000	0.0%	0.0%	1.0%	13.7%	32.7%	23.8%	25.6%	2.7%	0.4%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
7.000	0.0%	0.0%	1.1%	7.2%	26.5%	17.7%	29.0%	16.0%	2.5%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
8.000	0.0%	0.0%	0.8%	4.4%	17.9%	5.0%	26.9%	27.5%	13.1%	0.0%	0.1%	0.0%	4.3%	0.0%	0.0%
9.000	0.0%	0.0%	1.1%	4.8%	19.0%	2.6%	24.2%	33.6%	14.7%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
10.000	0.0%	0.0%	0.1%	0.9%	8.5%	3.5%	2.5%	0.0%	0.0%	0.0%	84.4%	0.0%	0.0%	0.0%	0.0%
11.000	0.0%	0.0%	1.8%	9.6%	23.7%	0.0%	0.0%	0.0%	0.9%	0.0%	0.0%	0.0%	64.0%	0.0%	0.0%
12.000	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	2.5%	0.0%	0.0%	0.0%	97.5%	0.0%	0.0%
13.000	0.0%	0.0%	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
15.000	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	2.6%	2.6%	94.7%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%

Row Percent
100.00
80.00
60.00
40.00
20.00
0.00

Classification quality matrix



Neural network view



Importance of the 1st regression predicates

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.267[a]	.071	.071	1.245615

a. Predictors: (Constant), average_rate_user (1), number_negative_rates, Age, date

ANOVA

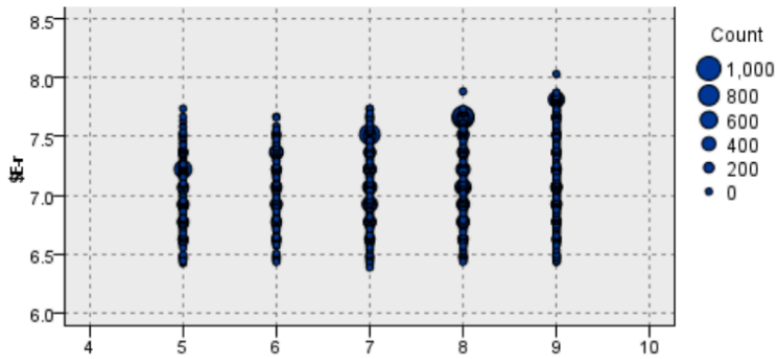
Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	4278.705	4	1069.676	689.421	.000[b]
	Residual	55750.562	35932	1.552		
	Total	60029.267	35936			

b. Predictors: (Constant), average_rate_user (1), number_negative_rates, Age, date

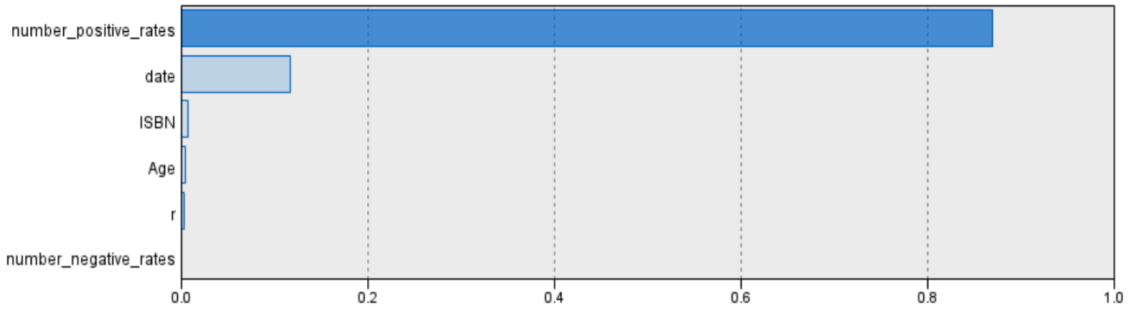
Coefficients

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	2.524	1.906		1.324	.185
	number_negative_rates	.000	.000	.010	1.948	.051
	date	.002	.001	.011	2.066	.039
	Age	6.891E-5	.000	.001	.140	.889
	average_rate_user (1)	.149	.003	.266	52.143	.000

Assessment of the quality of the 1st regression



Real and predicted product evaluation by the 1st regression



Importance of the 2nd regression predicates

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.708[a]	.502	.502	1.183348

a. Predictors: (Constant), ISBN, number_negative_rates, r, Age, date, number_positive_rates

ANOVA

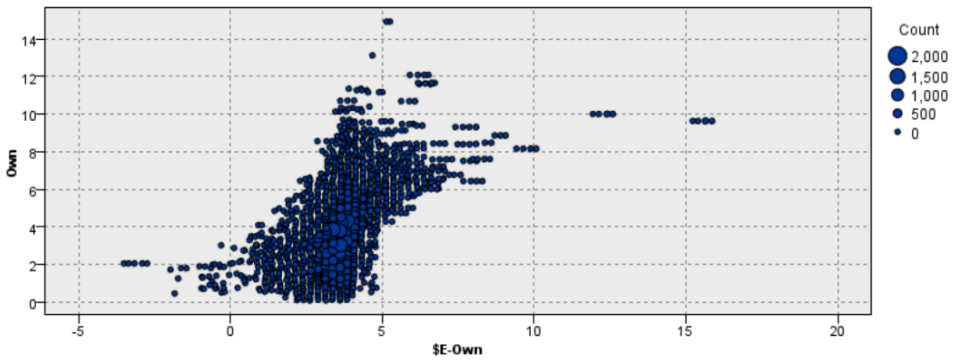
Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	354157.985	6	59026.331	42152.266	.000[b]
	Residual	351427.939	250964	1.400		
	Total	705585.924	250970			

b. Predictors: (Constant), ISBN, number_negative_rates, r, Age, date, number_positive_rates

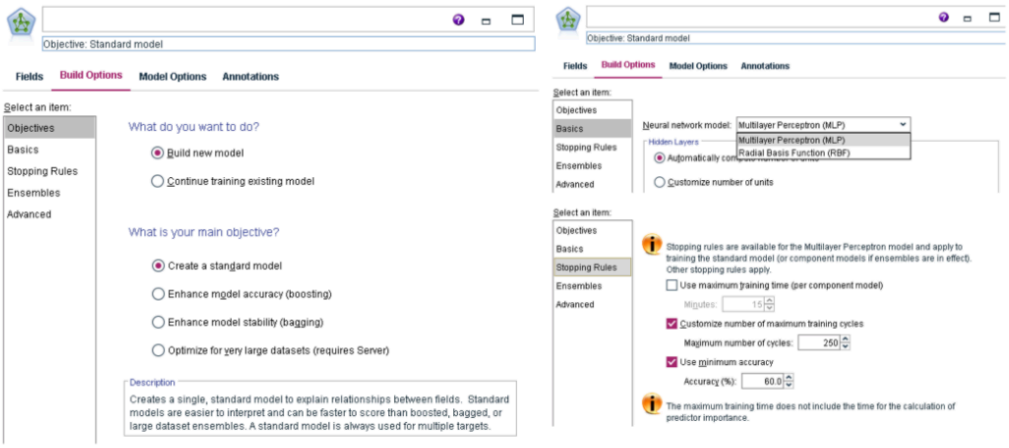
Coefficients

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-85.255	.647		-131.799	.000
	r	.144	.005	.041	29.315	<.001
	number_positive_rates	.031	.000	.835	444.742	.000
	number_negative_rates	-.004	.000	-.367	-196.794	.000
	date	.044	.000	.193	135.272	.000
	Age	-.005	.000	-.025	-17.834	<.001
	ISBN	-5.673E-14	.000	.000	-.025	.980

Assessment of the quality of the 2nd regression



Real and predicted product evaluation by the 2nd regression

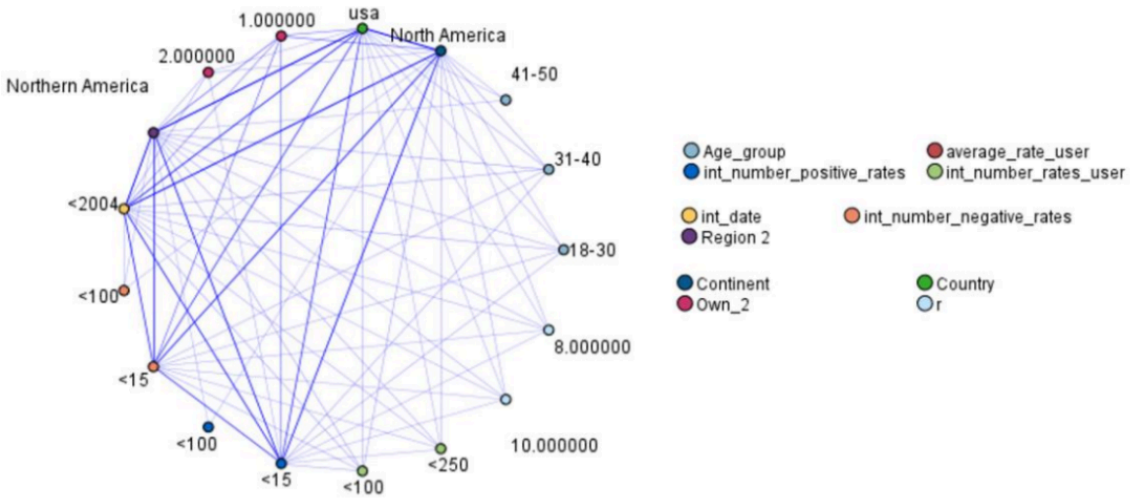


Apriori tab options

Консеквент	Антецедент	Екземпляри	Підтримка	Впевненість	Підтримка правила	Ліфт
Region 2 = Northern America	r = 10.0 and Own_2 = 1.0	6526	12,18	85,67	10,44	1,02
int_number_negative_rates = <15	r = 8.0 and Own_2 = 1.0	6843	12,78	84,39	10,78	1,13
int_date = <2004	average_rate_user = 5.0	5446	10,17	84,32	8,57	1,03
Continent = North America	r = 8.0 and Own_2 = 1.0	6843	12,78	84,26	10,76	1,00
int_date = <2004	average_rate_user = 2.0	9689	18,09	82,2	14,87	1,01
Continent = North America	r = 7.0	8386	15,66	80,97	12,68	0,96
Country = usa	average_rate_user = 1.0	7578	14,15	80,25	11,35	1,08
Region 2 = Northern America	average_rate_user = 3.0	9412	17,57	79,92	14,04	0,95
int_date = <2004	r = 10.0	13320	24,87	79,41	19,75	0,97
Continent = North America	average_rate_user = 5.0	5446	10,17	78,28	7,96	0,93
Region 2 = Northern America	average_rate_user = 4.0	6342	11,84	78,16	9,25	0,93
Region 2 = Northern America	average_rate_user = 5.0	5446	10,17	78,04	7,93	0,93
Country = usa	r = 10.0 and Own_2 = 1.0	6526	12,18	76,52	9,32	1,03
int_number_positive_rates = <15	r = 8.0 and Own_2 = 1.0	6843	12,78	74,73	9,55	1,00
int_date = <2004	r = 8.0 and Own_2 = 1.0	6843	12,78	72,82	9,30	0,89
int_date = <2004	Own_2 = 1.0	27308	50,98	72,18	36,8	0,88
int_number_negative_rates = <15	average_rate_user = 2.0	9689	18,09	72,09	13,04	0,96
int_number_negative_rates = <15	Own_2 = 2.0	13864	25,88	64,70	16,75	0,86

Джерело: сформовано автором

Quality assessment of the 2nd regression variables



Graph of the frequency of connections between variables