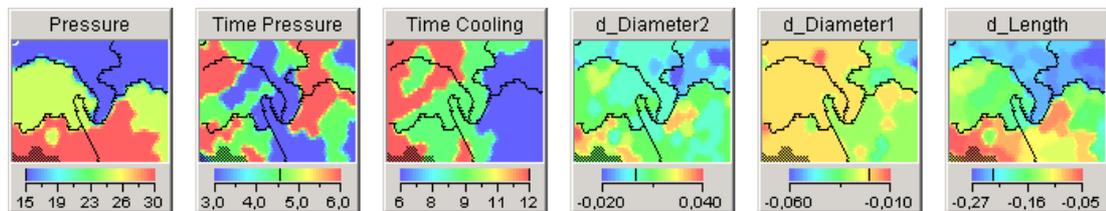


Viscovery SOMine 5 – New Features

Self-Organizing Maps powered by Viscovery®

With Viscovery SOMine version 5, the leading implementation of Self-Organizing Maps (SOM) combined with statistical methods is available to you. Following numerous customer requests, the latest SOMine generation has been re-developed based on the commercial Viscovery data mining suite.



The Viscovery SOMine visual analysis platform turns data mining into an intuitive experience and supports the realization of data mining projects in particular through the following comprehensive features:

- Project environment
- Workflow guidance
- Inline documentation

In addition to a multitude of new, powerful features — the most important of which are described in this paper — Viscovery SOMine version 5 supports diverse applications in the following fields:

- Explorative data mining
- Statistical analysis
- Profiling and segmentation
- Visual cluster analysis
- Classification

Viscovery SOMine 5 – New Features

Completely different, still the same

In Viscovery SOMine version 5 nothing is as it was except the intuitive understandability and the long-standing, proven interactivity of the Viscovery technology. Here are just a few of the highlights, which are described further in this document.

Viscovery supports you in completing complex data mining projects through a comprehensive project environment and workflows with well-defined, proven default settings. You can completely concentrate on your analysis — the Viscovery SOMine system handles all analysis steps, selected settings and data management.

Direct access to the data records underlying the SOM and their complete statistical contents is available at any time in both the analysis workflow and the visual analysis user interface. In addition to comprehensive descriptive statistics, you have access to cluster, principal component and correlation analyses, as well as comprehensive graphical evaluation using histograms, scatter plots and box plots in the context of each selection, down to the microcluster level.

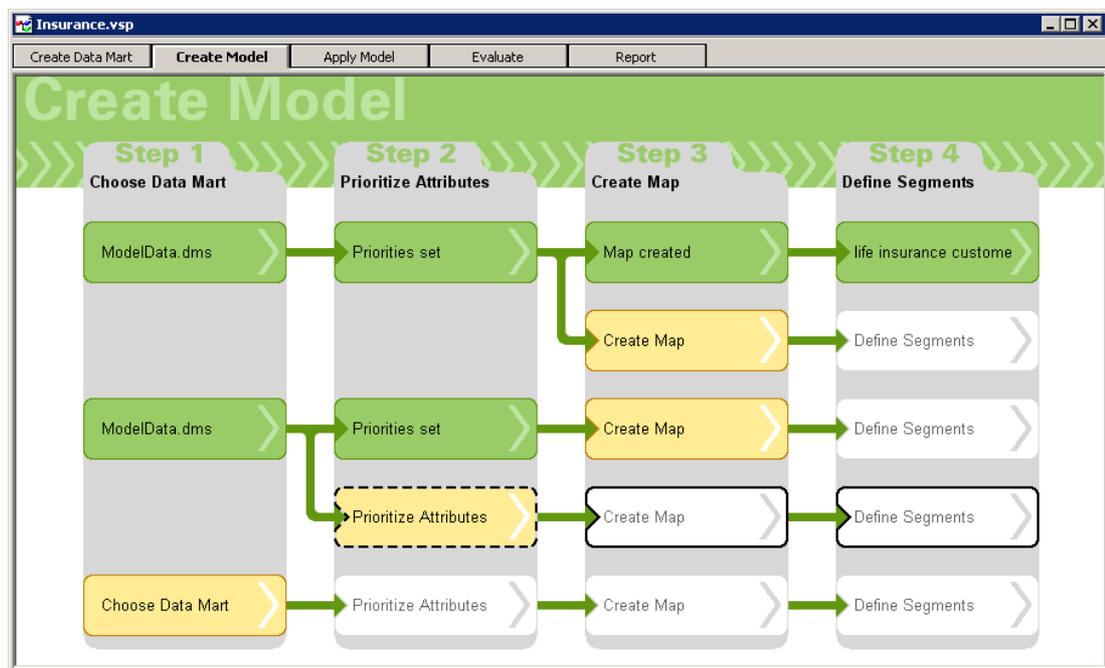
Viscovery SOMine version 5 is available in three editions: Basic, Expert and Enterprise. The editions vary in the availability of features and the number of processable data records and variables. The completely equipped "Enterprise Edition" contains additional special algorithms for high-performance processing of extremely high-dimensional data, for example, as needed in genetic analysis or text-mining applications.

Additionally, for SPSS® users, a special version of the Viscovery SOMine 5 system, which allows complementary use of the visual approach of the Viscovery system, is available.

Viscovery SOMine 5 – New Features

Analytical workflows safely guide you to your analysis goals

The analytical workflows, time-tested and proven as Viscovery business solutions, are now also available to you in the Viscovery SOMine system. In the workflows, you can set parameters for all steps in the preparation of analytical data marts, the generation and use of your data mining models, and their evaluation.



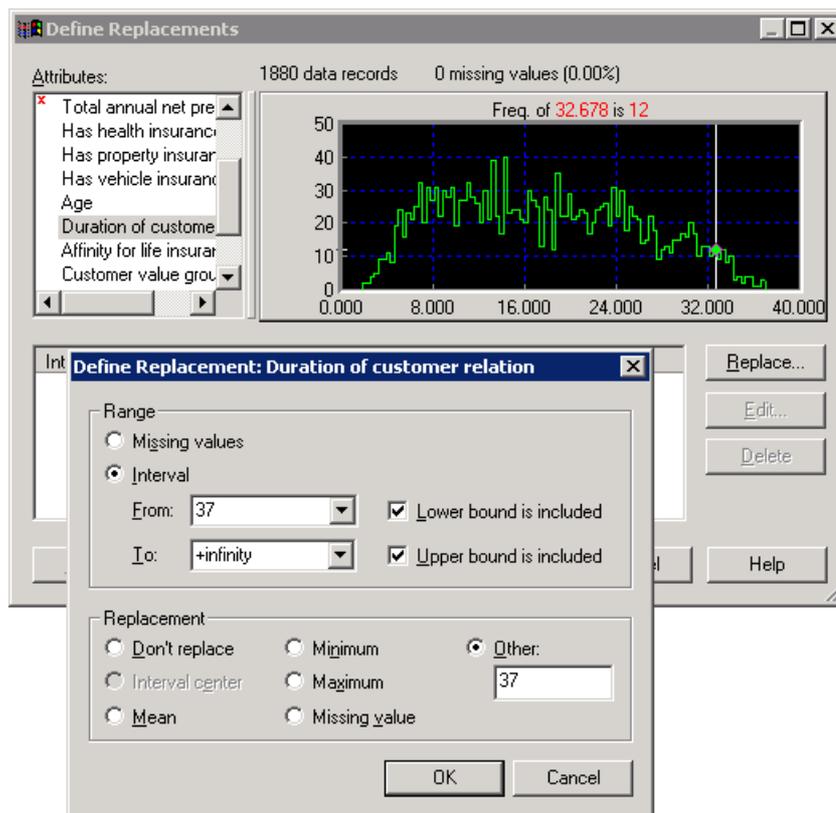
The individual analysis steps within each workflow have clearly defined semantics. You can create an alternative branch at any point in the workflow and vary the parameters to iterate your analysis, which is often necessary in data mining.

The SOMine system includes integrated reporting functionality, which provides a context-sensitive record of all performed steps — from importing data to the active workflow step — so you can always keep the overview. Additionally, you can document each analysis project and comment on the parameterization inline.

Viscovery SOMine 5 – New Features

Data structures and preprocessing

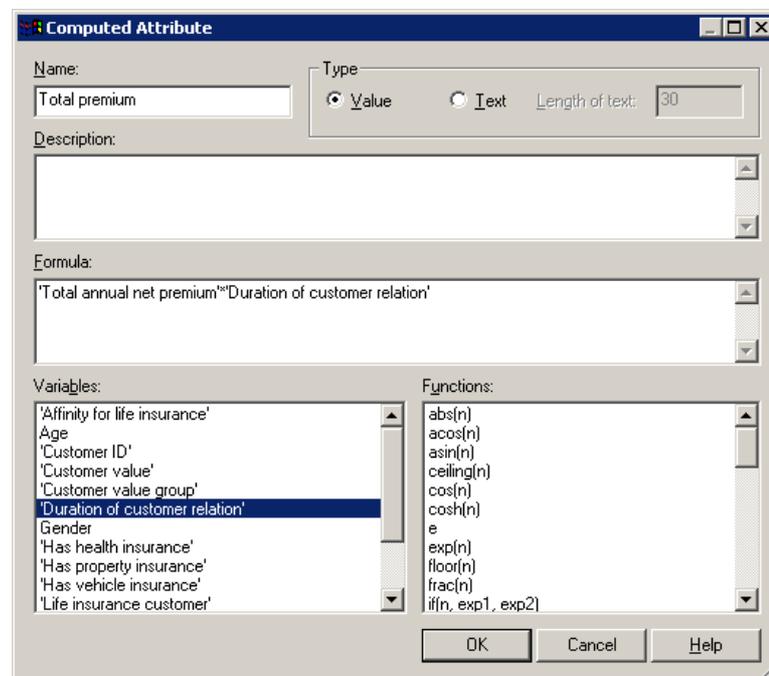
In addition to improved user ergonomics, SOMine version 5 includes numerous new data preprocessing functions that are known to be fundamental for the analysis process. For example, data records can be sampled or multiplied using user-defined formulas. Individual value ranges of variables can be replaced/transformed with a mouse click. High performance is always assured — also for large amounts of data.



Viscovery SOMine 5 – New Features

Definition of new variables

The derivation of new variables is now possible directly in the Viscovery SOMine system. Simply use formulas to define the desired quantities, which then become available to you as variables in the analysis process.

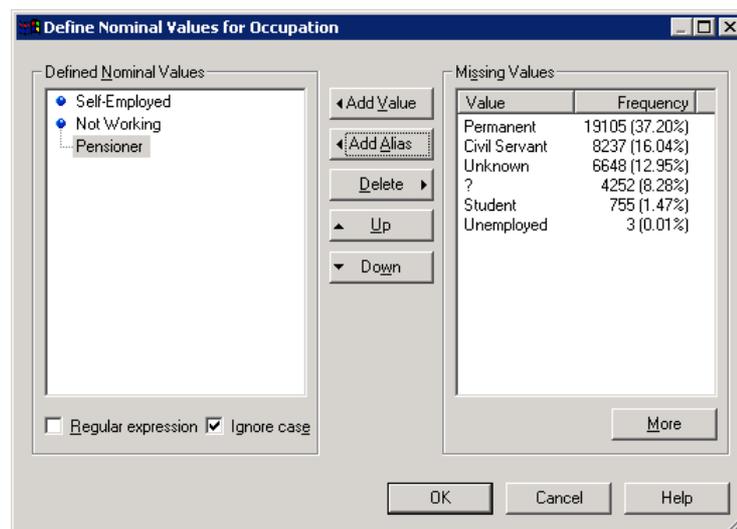


Managing nominal variables

You can now also use nominal variables in the Viscovery SOMine system (for example, the variable "Color" with values for red, blue, green or yellow). Simply use the mouse to define which (arrangements of) values you want to use in the analysis process; the Viscovery SOMine system takes care of the rest for you. Instead of having to painstakingly recode such non-metric variables by hand, the SOMine

Viscovery SOMine 5 – New Features

system takes care of the internal splitting into binary dummies and further on manages them automatically and transparently.



Automatic correlation compensation

Through correlations in the data, cluster analyses as well as the SOM data order may become biased. If you suspect correlations in your data, you can compensate for them by selecting one simple option. The Viscovery SOMine system automatically calculates and manages the optimal factors for internally rescaling correlated variables to establish a uniform data distribution.

Viscovery SOMine 5 – New Features

Prioritize Attributes

Attribute	Priority	Scaling	Minimum	Maximum	Mean	Std. Devi...	Missing V...	Transformation
Customer ID							no numeric...	
<input checked="" type="checkbox"/> Gender	1.00	Variance	0	1	0.425532	0.494555	0 (0.00%)	None
<input checked="" type="checkbox"/> Life insurance cus...	1.00	Variance	0	1	0.504255	0.500115	0 (0.00%)	None
<input checked="" type="checkbox"/> Number of policies	1.00	Variance	1	8	1.81277	1.29026	0 (0.00%)	None
<input checked="" type="checkbox"/> Total annual net p...	1.00	Variance	370	5990	1685.95	1347.44	0 (0.00%)	Sigmoid: offset 1685.9...
<input checked="" type="checkbox"/> Has health insuran...	1.00	Variance	0	1	0.517021	0.499843	0 (0.00%)	None
<input checked="" type="checkbox"/> Has property insur...	1.00	Variance	0	1	0.428723	0.495025	0 (0.00%)	None
<input checked="" type="checkbox"/> Has vehicle insura...	1.00	Variance	0	1	0.106383	0.308409	0 (0.00%)	None
<input checked="" type="checkbox"/> Age	1.00	Variance	24.4533	76.3215	50.106	13.2506	0 (0.00%)	None
<input checked="" type="checkbox"/> Duration of custo...	1.00	Variance	1.80696	37	17.552	8.32743	0 (0.00%)	None
<input checked="" type="checkbox"/> Affinity for life insur...	1.00	Variance	56.3046	124.121	98.2264	13.5938	0 (0.00%)	None
<input checked="" type="checkbox"/> Customer value gr...	1.00	Variance	0	3	1.58989	0.815891	0 (0.00%)	None
<input checked="" type="checkbox"/> Customer value	1.00	Variance	-4270.56	3171.25	120.612	1107.03	0 (0.00%)	None
<input checked="" type="checkbox"/> Occupation	1.00	Range	6 nomina...				289 (15.37%)	None
<input type="checkbox"/> Total premium	1.00	Variance	2281.87	210288	31432.5	34334.8	0 (0.00%)	None

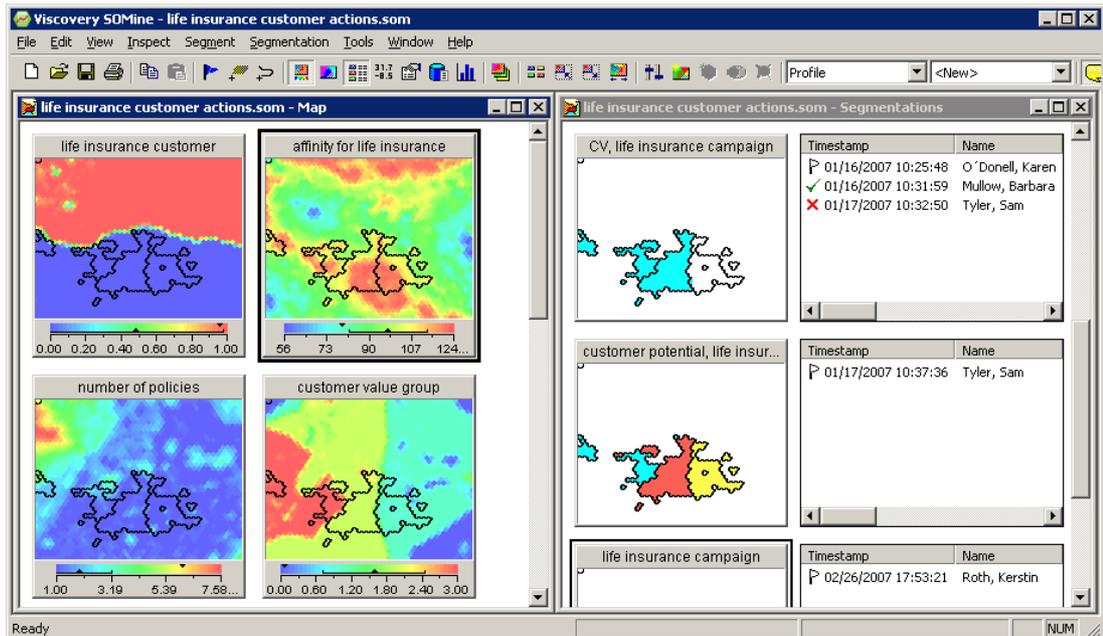
Compensate correlations Priority...

Automatic and manual definition of cluster boundaries

In addition to automatic clustering of the data record based on proven statistical techniques, it is now possible to manually override the cluster boundaries. Through the unique Viscovery cluster visualization, you can thus incorporate your domain-specific knowledge into the segmentation as needed.

You can create and save several segmentations in a model in parallel. Depending on the requirements or the application you can activate the required segmentation layer over the model with a mouse click.

Viscovery SOMine 5 – New Features

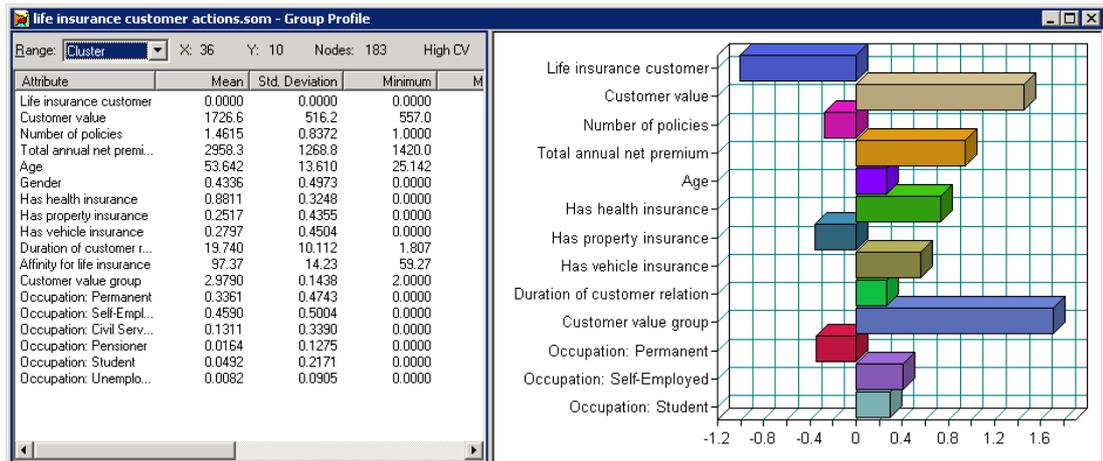


Statistical profiles

In addition to displaying descriptive analysis of the data records within any selection of the visual data representation, you can now display group profiles of the selected variables. These profiles show the deviations of the variable means within the selected group (in standard deviations) from the mean of the whole.

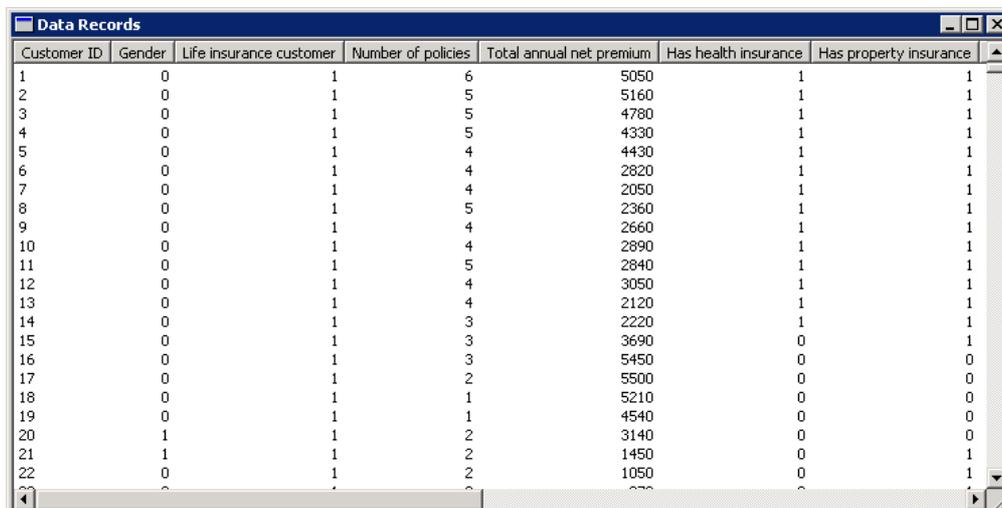
This presentation makes the description of the segments substantially easier in that it takes the preset levels of significance into account and expresses the statistical profile of the selected range at a glance.

Viscovery SOMine 5 – New Features



Access to the original data records

With Viscovery SOMine version 5 (available in the Expert Edition and Enterprise Edition) you can directly access the underlying data at any point in the workflow or for any area selections in the map. With this "Cluster Querying" functionality, you can carry out context-sensitive queries on the original data.



The screenshot shows the 'Data Records' window with the following data:

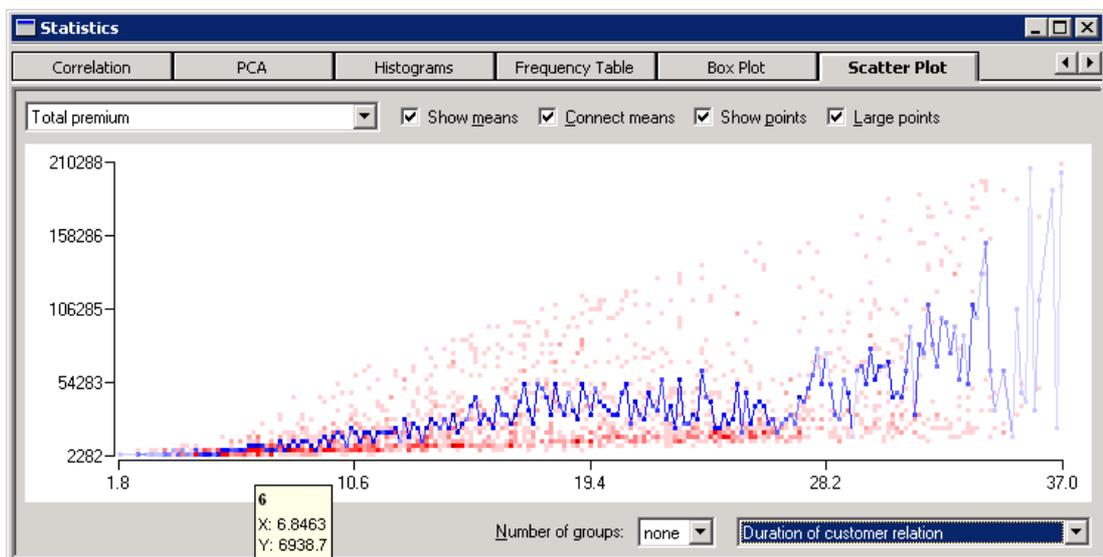
Customer ID	Gender	Life insurance customer	Number of policies	Total annual net premium	Has health insurance	Has property insurance
1	0	1	6	5050	1	1
2	0	1	5	5160	1	1
3	0	1	5	4780	1	1
4	0	1	5	4330	1	1
5	0	1	4	4430	1	1
6	0	1	4	2820	1	1
7	0	1	4	2050	1	1
8	0	1	5	2360	1	1
9	0	1	4	2660	1	1
10	0	1	4	2890	1	1
11	0	1	5	2840	1	1
12	0	1	4	3050	1	1
13	0	1	4	2120	1	1
14	0	1	3	2220	1	1
15	0	1	3	3690	0	1
16	0	1	3	5450	0	0
17	0	1	2	5500	0	0
18	0	1	1	5210	0	0
19	0	1	1	4540	0	0
20	1	1	2	3140	0	0
21	1	1	2	1450	0	1
22	0	1	2	1050	0	1

Viscovery SOMine 5 – New Features

In this way you can check, for example, analysis hypotheses by means of the original data records at the microcluster level. Alternatively, you can display all original data which fall into a segment or selection in real-time.

Statistical and graphical analysis

With the Viscovery SOMine 5 Expert Edition or Enterprise Edition, you can run comprehensive statistical analysis at any point in the workflow or over any selected area of the map. In addition to descriptive statistics, you have access to principal component and correlation analyses as well as histograms, scatter plots and box plots in any context, down to the microcluster level.



Viscovery SOMine 5 – New Features

About Viscovery Software GmbH

As one of the first data mining companies in Europe, Viscovery (formerly eudaptics software gmbh) is among the leading vendors of data mining solutions. Since September 2007 Viscovery has been part of the Biomax group.

The Viscovery suite comes with unique patented technology for exploratory analysis and statistical modeling of complex data. Viscovery business solutions enable high-performance, real-time integration of explorative models, classification models, non-linear predictions and scorings as well as workflow automation and embedded analytics.

For many years, the Viscovery software has been used by more than 300 customers in diversified areas ranging from banking, insurance, telecommunications, manufacturing, media and retail to research organizations and universities.

Free trial license of the Viscovery SOMine software

See the advantages of Viscovery SOMine version 5 for yourself! Get your free 30-day trial license of Viscovery SOMine from www.SOMine.info and let your analysis projects benefit from the performance and the manifold new features of the leading implementation of Self-Organizing Maps!

Viscovery is a registered trademark of Viscovery Software GmbH.
SPSS is a registered trademark of SPSS Inc.