

SALES STRATEGY PLANNING WITH THE NOVOSPARK@ VISUALIZER TOOL

OLGA PILIPCZUK – DMITRI EIDENZON

Abstract: In this paper we present a method for the analysis and visualization of multidimensional business data which was implemented in the NovoSpark@ Visualizer software system. An application example is provided, based on sales data. We describe the advantages and opportunities of visualization tool in sales strategy planning. We then discuss the results of a usability experiment. The results of this application indicate that it is possible to make quick decisions and to create the sales strategy by identifying anomalies and regularities using the NovoSpark@ Visualizer system.

Keywords: sales strategy planning, multidimensional data visualization, NovoSpark@ Visualizer

JEL Classification: M15, M31

1. INTRODUCTION

Ensuring the adequacy of the managerial decisions is possible only on condition of proper interpretation of data analysis results, which usually is impossible without visualization, especially when we try to process a large number of parameters simultaneously. In this case, the visual analysis becomes more complicated because of the increased volume of the input data. Additional graphics objects and verbal and numerical information describing these objects are emerged in the image and, therefore, increase number of efforts to interpret the results of the analysis. So, the need of multidimensional data visualization as a single image is appeared. This image allows for easy interpretation of results and increases the efficiency of tasks.

Sales strategy creation is a complicated process based on the planning of sales activities: methods of reaching clients, competitive differences and resources available. The information about the prices, demand, and concurrency are analyzed during this process. Usually the sales data has more than two dimensions. It should be analyzed with dynamic changes taking into account.

The main purpose of the article is to present the new approach for visualization of multidimensional sales data and to show the effects of its application. The example of decision making in the process of the sales strategy planning with NovoSpark@ Visualiser tool is provided.

2. TRADITIONAL METHODS OF MULTIDIMENSIONAL DATA VISUALIZATION

The traditional multidimensional visualization methods are: parallel coordinates, 3D parallel coordinates, worlds, line graphs, survey plot, scatter plots, variable binned scatter plots, scatter plot matrix and it's variations, star glyphs, treemaps, self organizing maps, dendrogram, radar charts, Voronoi diagrams, Chernoff's faces and also less popular parallel glyphs, Bertin's permutation matrices, words within worlds, table lens, VisDB, dynamic queries, attribute explorer. Very popular are the methods of dimensionality reduction [15, 24, 25]. The most widespread of them is the Principal Component Analysis [13].

However, most traditional methods and techniques of multidimensional data visualization have some disadvantages or weaknesses:

- a problem of organizing the plots on the screen in a meaningful manner,
- a lack of transparency in cases of a large number of features, due to which the analysis often becomes impossible,
- some of plots are unsuitable for the analysis of a non-linear data structure,
- some of methods do not provide the ability of dynamic data visualization,
- a lack of tolerance for sensory faults.

To eliminate the above problems many researches are still looking for the new methods or modernize or successfully modify the classical methods [1-4, 7-12, 14, 16-23]. The new method of multidimensional data visualization is also has been proposed by NovoSpark Corporation.

3. THE NOVOSPARK METHOD OF MULTIDIMENSIONAL DATA VISUALIZATION

The NovoSpark method is based on two isometric spaces, the objects from one data space are called the originals, while objects from the other space are called the images.

A point-observation \mathbf{A} in N -dimensional affine point-vector space R_N of the originals $A = (a_0, a_1 \dots a_{N-1})$ and form linear combination $f_A(t)$ of functions $\{P_i(t)\}_0^\infty$ by using the following formula [5, 6]:

$$f_A(t) = \sum_{i=1}^{N-1} a_i P_i(t) \quad (1)$$

where $P_i(t)$ - are orthogonal polynomials with weight 1 defined on the segment $t = [0, 1]$ (Legendre polynomials).

Image of a point in a multidimensional space can be "painted" in accordance with the function values. Such visualization of an image $f_A(t)$ as a surface is called a "spectrum" of the multidimensional point-observation. Applying a color palette emphasizes similarities and differences in images and allows viewing them in the

coordinate system $\{z, t\}$, where one can render an arbitrary image "width" along the z-axis.

Using NovoSpark method analysts can also visually identify abnormal observations. The image of a multidimensional interval is a two-dimensional region between the "minimum" and "maximum" images; this image is rendered in the coordinate system $\{f, t\}$ and is called a "cloud" of a multidimensional interval. Boundaries of the cloud are obtained from a linear combination of separate images for each parameter/variable from the coordinate space of the originals R_N . By fixing certain values of a multidimensional interval, for example, ± 3 standard deviations from the mean multivariate sample, one can visually identify abnormal observations. The images, even partially lying outside the interval "cloud" can be seen as abnormal with respect to the given interval.

NovoSpark method is especially helpful in next tasks of data analysis:

- Data preparation and study for analysis: detection of abnormal observations, identification of the most informative parameters / variables.
- Identification of data patterns: identification of static data clusters, identification of periodic segments for dynamic data.
- Selection of a reference as a single state, as a group state or as part of the process of periodically changing states.
- Comparison of references

4. THE NOVOSPARK@ VISUALIZER SOFTWARE

The NovoSpark@ Visualizer tool has been created for analysis and visualization of multidimensional data [26]. Visualizer consists of a set of pluggable executable modules for the Windows operating system.

The user interface provides the following functionality: synchronized and interactive display of data, images and results of data processing, flexible user interface, access to all functions through the user menu, context sensitive and dynamic help.

The main window shown on Figure 1 consists of configurable user interface elements. It allows synchronized interaction with the images and working with data sets. The

contents of the window including current settings and data state, is stored into the Visualizer project file. Multiple project windows can be opened simultaneously. Synchronized data and image rendering allows detailed data examination.

In addition to traditional help, Visualizer provides context-sensitive dynamic information about the current state of the image with the links to the relevant help topics that allow novice users to quickly get familiar with the visualization technology and capabilities of the system.

Visualizer provides variety of methods for data import and processing. A new data set can be created by a user, and if there are other data sets already loaded into the same project, completely copies the existing data structure, including the number and types of the parameters. The user is provided with a flexible and easily accessible tool for editing data structures: adding and deleting parameters, changing the order of parameters, modifying parameter types (numeric, textual or chronological), etc.

The data window shown on Figure 1 contains a table with all data values and a toolbar that provides access to the frequently used commands and configuration settings for this data set. The main window contains as many data windows, Figure 1 (2), as there are data sets loaded or created in the current project.

One of the important steps in each data analysis is data transformation. The traditional methods of data transformation used in the system are: normalization - to express the results in a single measurement system and standardization - to compare variables and/or data sets with different characteristics or distribution of units.

Visualizer creates two types of data images: "integral" images containing information about all parameters and "traditional" images. The integral images are: the "NovoSpark Curves" (see (1)), "Andrews Curves" and "Parallel Coordinates". An integral image is always rendered in the image window (Figure 1 (1)).

The available traditional images are: "Linear Plots", "Multiple Linear Plots", "Scatter Plot Matrix", "Polar Coordinates" and "Histograms". User can visualize traditional images that are displayed in the "Plot Window".

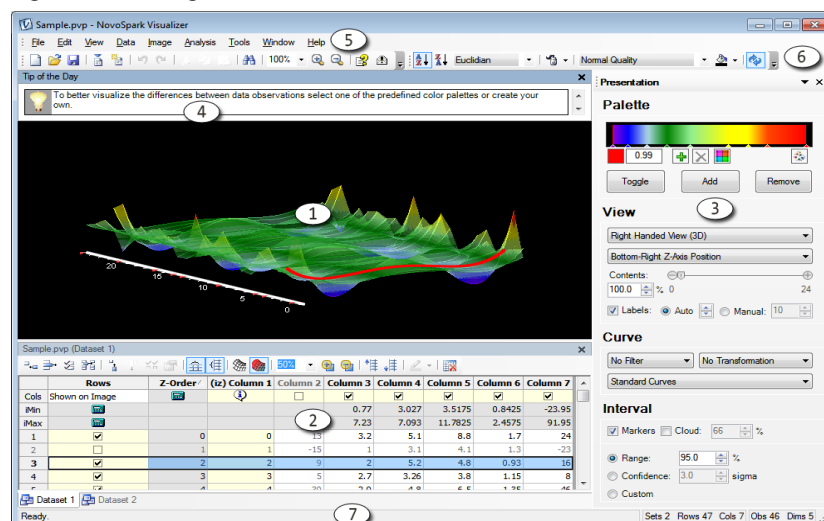


Figure 1 The NovoSpark@ Visualizer main window.

1) Image. 2) Data windows. 3) Task pane. 4) Tip of the day. 5) Main menu. 6) Tools panel. 7) Status bar.
Source: [26]

Table 1 The real estate company's sales in years 2010-2013 groped by quarters

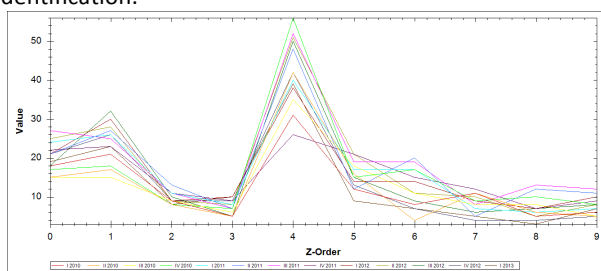
City	2010				2011				2012				2013
	I	II	III	IV	I	II	III	IV	I	II	III	IV	V
Poznan	18	15	15	17	24	21	27	22	21	25	18	21	19
Wroclaw	21	17	15	18	26	27	25	23	30	28	32	26	23
Gorzow	9	8	9	8	11	13	11	11	9	9	10	9	8
Sopot	5	5	6	7	8	7	7	9	10	8	5	9	10
Warsaw	31	42	35	56	40	26	52	48	38	51	50	32	39
Szczecin1	12	16	18	15	7	12	19	21	24	21	15	13	9
Szczecin2	8	4	11	17	17	20	19	15	14	11	9	7	7
Szczecin3	11	11	8	9	7	5	8	12	9	10	6	4	5
Stargard1	5	6	8	10	6	12	13	7	7	5	7	4	3
Stargard2	6	5	5	8	7	11	12	9	10	8	8	5	7

Visualizer supports traditional methods for multidimensional data analysis for classification of parameters, observations and individual data sets: descriptive statistics, discriminant analysis, factor analysis, cluster analysis, linear regression analysis, Kohonen self-organizing maps.

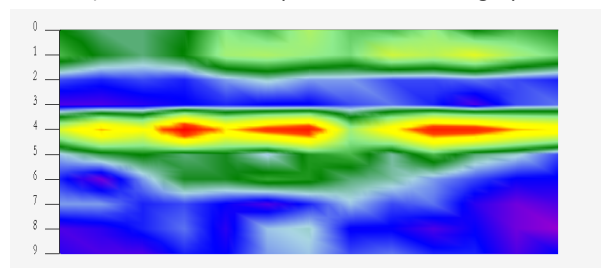
5. APPLICATION

In order to present the NovoSpark method a sales analysis has been made on the basis of real estate company data. During the research the number of soled apartments has been observed. Table 1 presents the number of sold apartments in years 2010-2013 groped by quarters. The knowledge about demand changes will be the bases for future sales planning. Firstly, the traditional visualization has been made. After that, authors provide the interpretation using color spectrum and compare the results. Figure 2a shows the data on company sales in 2010-2013. The x-axis shows the months, the y-axis - cities. In this figure all cities are labeled sequentially, i.e 0 presents Poznan, 1- Wroclaw etc.

Although the traditional graph shows quite clearly the sales level, the spectral interpretation (Fig. 2b) provides more information about the function value in the specified period of time. Both visualizations show that the most apartments have been sold in Warsaw, however the spectrum visualization provides the possibility of sales trend identification.



a) traditional interpretation with line graph



b) spectrum" interpretation

Figure 2 Visualization of the sales data
Source: own preparation

Figure 2b shows that the sales are growth tree times in analyzed period. It also allows observing and comparing the whole sales level. The blue color indicates that the whole real estate company sales level is low despite of three cities: Poznan, Wroclaw and Warsaw.

The next image provides the sales dynamics visualization. Every point of image can be analyzed using mouse inspect tool, which shows the parameters influence (Figure 3). Due to different colors mixing, it can be observed that the situation in market is unstable. The upward trend generally is observed in 2011.

Figure 4 provides the information about the abnormal observations. These observations are: III quarter of 2011 and III quarter of 2012.

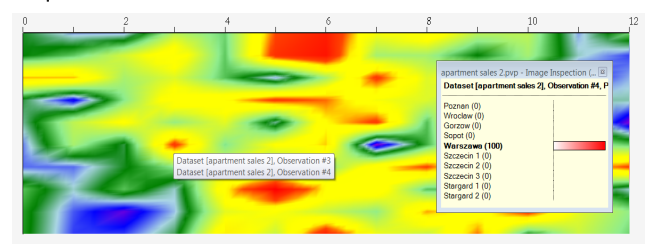


Figure 3 NovoSpark@Visualizer mouse inspect tool for dynamics analysis.

Source: own preparation

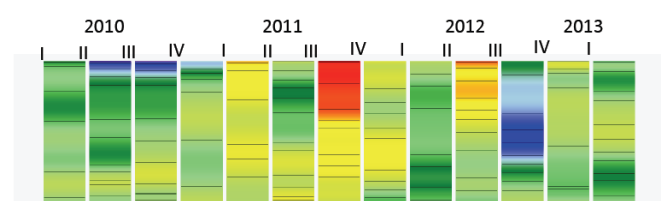


Figure 4 The NovoSpark curves for abnormal observations identification

Source: own preparation

The more detailed results could be obtained by using the "cloud" image. The observations: IV 2010, II, III 2011, III 2012 are identified as abnormal (Figure 5). The real estate company should take these periods under particular attention in sales analysis and look at them in details.

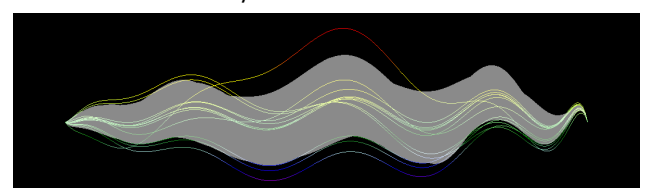


Figure 5 The "cloud" for abnormal observations identification

Source: own preparation

The presence of abnormal observations is the reason to make cause-effect analysis and to identify why the demand has changed in the observed period.

6. CONCLUSION

Presented in paper analysis is just a part of a whole sales strategy planning procedure. The visualization tool allows interpreting the data, identifying the most informative parameters and observing the anomalies and trends without complex calculations. The Novospark@ Visualization

software helps to avoid the disadvantages of many of traditional visualization methods, such as: a lack of transparency in cases of a large number of features, unsuitability for the analysis of a non-linear data structure, a lack of ability of dynamic data visualization, a lack of tolerance for sensory faults. The managerial decision making became less complicated by the integration of traditional analysis and new visualization method used in Novospark@ Visualizer.

REFERENCES

- [1] ALAHAKOON, D., HALGAMUGE, S., SRINIVASAN, B. *Dynamic self-organizing maps with controlled growth for knowledge discovery*. IEEE Transaction on Neural Networks. 2000, vol. 11, no. 3, pp. 601–614.
- [2] AMARASIRI, R., ALAHAKOON, D., SMITH, K., PREMARATNE, M. *HDGSOMr: a high dimensional growing self-organizing map using randomness for efficient web and text mining*. In: Proc. IEEE/WIC/ACM international conference on web intelligence. 2005, pp. 215–221.
- [3] AURENHAMMER, F., KLEIN, R., Voronoi, diagrams, in: *Handbook of Computational Geometry*. Edited by Sack, J. and Urrutia, J. Elsevier, 2000.
- [4] CHEN, C. *Information Visualization: Beyond the Horizon*, Springer, 2004.
- [5] EIDENZON D, VOLOVODENKO V, *Method for visualization of multidimensional data, Patent Application 20090252436*, USA, 2009.
- [6] EIDENZON D., SHAMRONI D., VOLOVODENKO V. *Method and System for Multidimensional Data Visualization*, Lambert Academic Publishing, Saarbrücken, 2013.
- [7] FANEA, E., CARPENDALE, S., INSELBERG, T. *An interactive 3D integration of parallel coordinates and star glyphs*, IEEE Symposium on Information Visualization, 2005, pp. 5-20.
- [8] FEW, S., *Multivariate Analysis Using Parallel Coordinates*, 2006, http://www.perceptualedge.com/articles/b-eye/parallel_coordinates.pdf
- [9] GEMIGNANI, Z. *Better Know a Visualization: Parallel Coordinates*, 2010.
- [10] HAO, M., DAYAL, U., SHARMA, R., KEIM, D., JANETZKO, H. *Variable Binned Scatter Plots*, *Information Visualization*, vol. 9 no. 3, 2010, pp. 194-203.
- [11] INSELBERG, A. *Parallel Coordinates: Visual Multidimensional Geometry and Its Applications*, Springer, 2009.
- [12] JOHANSSON, J., FORSELL, K., LIND, M., COOPER, M. *Perceiving Patterns in Parallel Coordinates: Determining Thresholds for Identification of Relationships*, *Information Visualization*, vol. 7 no. 2, 2008, pp. 152-162.
- [13] JOLLIFFE, I. *Principal component analysis*, New York: Springer, 2002.
- [14] KEIM, D., HAO, M., DAYAL, U., JANETZKO, H., BAK, P. *Generalized Scatter Plots*, *Information Visualization*, vol. 9 no. 4, 2010, pp. 301-311.
- [15] KIM, H., SCHULZE, J., CONE, A., SOSINSKY, G., MARTONE, M. *Dimensionality Reduction on Multi-Dimensional Transfer Functions for Multi-Channel Volume Data Sets*, *Information Visualization September 21*, vol. 9 no. 3, 2010, pp. 167-180.
- [16] LAGUS, K., KASKI, S., KOHONEN, T. *Missing massive document collections by the WEBSOM method*, *Information science*, 2004, vol. 163, no.1-3, pp. 135-156.
- [17] LARS, LINSEN, VAN LONG, TRAN, ROSENTHAL, PAUL. *Linking multidimensional feature space cluster visualization to multifield surface extraction*. IEEE Computer Graphics and Applications, 2009, vol. 29 no. 3, pp.85–89.
- [18] NATHAN, D., COORPRIDER, BURTON, ROBERT. *Extension of star coordinates into three dimensions*. In Proceedings of the SPIE, 2007, volume 6495.
- [19] REITSMA, R., TRUBIN, S., *Information Space Partitioning Using Adaptive Voronoi Diagrams*, *Information Visualization*, 2007, vol. 6 no. 2, pp. 123-138.
- [20] SCHRECK, T., BERNARD, J., VON LANDESBERGER, T., KOHLHAMMER J. *Visual Cluster Analysis of Trajectory Data with Interactive Kohonen Maps*, *Information Visualization*, 2009, vol. 8 no. 1, pp. 14-29.
- [21] SHNEIDERMAN, B., PLAISANT, C. *Treemaps for space-constrained visualization of hierarchies*, 2009.
- [22] SIPS, MIKE, NEUBERT, BORIS, LEWIS, JOHN, HANRAHAN, PAT. *Selecting good views of high dimensional data using class consistency*. Compute Graphics Forum, vol. 28, no. 3, 2009, pp. 831–838.
- [23] SUN,YANG, TANG, JIUYANG, TANG, DAQUAN, WEIDONG, XIAO. *Advanced star coordinates*. In Web-Age Information Management. The Ninth International conference WAIM, 2008, pp. 165–170.
- [24] VAN DER MAATEN, L. *An Introduction to Dimensionality Reduction Using Matlab*. Technical Report MICC-IKAT 07-07, Maastricht University, Maastricht, Netherlands, 2007.
- [25] VAN DER MAATEN, L., POSTMA, E., O., HERIK VAN DER, H. *Dimensionality Reduction: A Comparative Review*, *Journal of Machine Learning*, Tilburg University, Technical Report TICC-TR, 2009.
- [26] www.novospark.com

dr Olga PILIPCZUK

University of Szczecin, Mickiewicza 64, 71-101, Poland
e-mail: olga.pilipczuk@wneiz.pl

dr Dmitri EIDENZON

Novospark Corporation, Dansbury Drive 418
Waterloo, Canada
e-mail: dmitri.eidenzon@novospark.com