Visualization of Real-World Data: Supporting Insurance Brokers in the Process of Information Visualization

Patricia M Gouws, School of Computing, University of South Africa
PO Box 392, UNISA, 0003
gouwspm@unisa.ac.za

MR (Ruth) de Villiers, School of Computing, University of South Africa
PO Box 392, UNISA, 0003
dvillmr@unisa.ac.za

Abstract: This study considers the development and use of RoutePlanner, which is a model for the visualization process to transform raw data in a data warehouse into required insight, specifically in the domain of insurance brokerage. Each broker has a portfolio of clients, whose data is stored in a database that is accessed by a financial needs analysis system. This system does not provide an adequate overview of the portfolio, thus brokers are unable to obtain the holistic perspective required for the identification of client needs and marketing opportunities. This paper introduces RoutePlanner, a model of the visualization process, used to plan the route from raw data to required insight. Empirical research investigated whether the use of RoutePlanner led to increased awareness of visualization on the part of brokers and whether it resulted in successful visualizations of brokerage data by users who are not technological experts.

Key terms: Data warehouse, experiment, evaluation, guidelines, identification, selection, evaluation, visualization methods, visualization process.

Introduction

This study sets out to support visualization of data to addresses a real-world problem experienced by professional insurance brokers. Brokers in the South African target population under investigation are currently using a financial needs analysis system that, although comprehensive, does not provide holistic or historical overviews of the data in their client portfolios. Consequently it does not support users in the identification of particular client needs nor in determining appropriate marketing opportunities.

To address this problem, and to support brokers in acquiring enhanced insight into data, as they learn to explore the information space and to perceptively interpret the available information, we propose visualization of the data in their client portfolios. With this in mind, the first author developed RoutePlanner, a model of the visualization process. RoutePlanner is used to plan the 'route' of transformation from raw data to required insight.

Several kinds of experts are required to collaborate during the visualization process. The domain expert has domain knowledge and knows what insight is required. Domain experts are also the target audience and the end-users of the visualizations – in this case, the insurance brokers. The visualization expert has the knowledge, technological expertise, skill and experience to apply visualization methods. A data expert is familiar with the data in terms of its formats, location and availability, is responsible for the maintenance of the database, and may be a database administrator, designer or manager. An expert in the process may have more than one role.

To practically determine whether the use of this visualization process by domain experts (insurance brokers), in collaboration with data and visualization experts, leads to an increased awareness of visualization and results in the successful visualization of warehouse data, a contextualized experiment was conducted to test such a theory, along with an evaluation of RoutePlanner in use. The focus area of this study is, therefore, the process of visualization, while the application area is the domain of insurance brokerage.

To apply this study within its real-world context, a warehouse of data drawn from a stereotypical insurance broker was used during the experiment, consisting of a pre-study, an intervention using RoutePlanner, and a post-study. Furthermore, evaluation research was used to assess the utility of each stage of RoutePlanner. Quantitative and qualitative data was collected, analyzed and interpreted.

This paper introduces the components, stages and approach of RoutePlanner, a model for the visualization process; motivates the research design used in this study; discusses the research findings; and presents the conclusions and recommendations for future research.

RoutePlanner – a Model of the Visualization Process

A literature study was undertaken, which led to the development of a model of the visualization process. No single, absolute set of stages in the visualization route from raw data to required insight was encountered, so RoutePlanner was synthesized as a consolidated framework for this purpose. The literature study considered the following areas and extracted pertinent aspects, which subsequently became integral components and/or stages, of RoutePlanner:

- **Practical guidelines to be used in applying the visualization process:** Guidelines proposed by various authors comprise a body of knowledge that can support the experts, and also provide a valuable point of reference for the visualization experts. The guidelines of Brath (1999), environmental factors of Mann (1999) and Mann and Reiterer (2000), universal visualization issues of Spence (2001), and certain visualization trends synthesized by the first author, were included in the guideline component of RoutePlanner to assist and direct the experts prior to and throughout the stages of the visualization process, and hold potential to support subsequent independent use of RoutePlanner by domain experts.
- **Identification of which data aspects should be visualized and the associated user tasks:** RoutePlanner's identification stage is based on the technique of domain analysis of Espinosa, Hendrickson and Garret (1999) supplemented by the definition of abstract tasks by Shneiderman (1996), tasks for multi-dimensional data of OLIVE (1997), and the analysis cycle of Kimball, Reeves, Ross and Thornthwaite (1998). The discussions, using techniques suggested, results in a list of data issues to be visualized.
- **Selection of appropriate visualization methods for the tasks in hand:** Appropriate visualization methods (VMs) for given data must be selected from a vast collection of methods. Selection is considered an ad hoc process that requires experience, skill and intuition (de Oliveira and Levkowitz, 2003), and expertise may be acquired from a visualization expert, automated tools or from defined visualization method taxonomies. The taxonomies of Shneiderman (1996), OLIVE (1997), Keim (2002), De Oliveira and Levkowotz (2003) and Lengler and Eppler (2007) were included in RoutePlanner's selection stage to provide guidance for the selection of VMs. Public-domain VM systems were used for selection of suitable methods (VisTree (2008), XmdvTool (2008), InfoVis toolkit (2008) and TChart components).
- **Evaluation of selected visualizations to determine if they facilitate acquisition of the required insight:** The evaluation stage of RoutePlanner has three steps, namely: a general evaluation methodology of Wiss, Carr and Jonsson (1998) to compare suitability of two or more VMs, the insight-based methodology of Saraiya *et al* (2005) to identify insight gained from the visualizations, and a set of factors suggested by Kao and Ma (2000) to determine the success of the visualization.

Research Questions and Context of Study

The purpose of this empirical study is to investigate whether the use of RoutePlanner by the domain experts, in collaboration with a visualization expert and data expert, leads to an increased awareness of visualization and results in successful visualizations. The research questions posed in this paper are:

1. Does the use of RoutePlanner lead to an increased awareness of visualization on the part of insurance brokers?
2. Does the model of the visualization process result in successful visualizations of insurance brokerage data?

This section introduces the context of this study in its application domain of insurance brokerage. An insurance broker has a portfolio of clients, whose data is stored in a database accessed by a financial needs analysis program.

We propose the use of RoutePlanner by a collaboration of experts to gain enhanced insight into the visualized data, where snapshots of the database data are stored in a data warehouse.

Under the current situation, the insurance broker has a database containing the most recent data pertaining to each of the clients in his portfolio and uses a financial needs analysis program to access and manipulate this data. The broker selects a specific client from the portfolio and a particular unit of insurance analysis (e.g. Pension, Estate). The client's details for the selected unit are then presented to the broker who is, however, unable to gain a global overview of the data holistically or historically. Using this system, the insurance broker can view only the data of a single, selected client for the selected unit. For the broker to view the data of all clients in a portfolio, the system includes limited fixed queries, where results are presented as bland lists. Furthermore, the broker can view only the data relating to the current status of each client in the portfolio, as the database does not store historical data. There are, therefore, limitations on the insights a broker can obtain from the data in its present form and structure.

Under the system proposed by this paper, there would be a data warehouse instead of a database. For this research, a prototypical data warehouse was developed, into which regular snapshots of data from the insurance database of a stereotypical broker were loaded. Following the generation of the warehouse, visualization was used to explore the data and acquire deeper insights, so to overcome the limitations of the present system. Where the expertise is present to meaningfully interpret the visualization, a variety of useful insights may be acquired which can be used to support appropriate marketing by the broker and to contextually advise his clients. During this study this expertise resided in the visualization expert who guided the brokers in the intervention sessions, but a long term goal is to wean the brokers and train them towards independent usage of RoutePlanner.

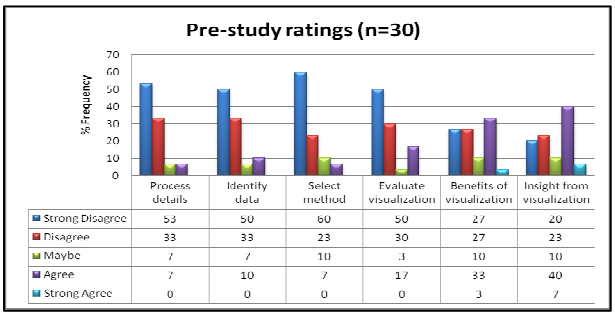
Research Methodology

The design of the study comprised a quasi-experiment and evaluation research as per Mouton (2003) and Olivier (2004). Domain expert participants were thirty insurance brokers (30), all of whom use the financial brokerage data. Some brokers had particular knowledge of the data in their brokerage databases and some were responsible for database maintenance and thus fulfilled the role of data expert. The first researcher served as the visualization expert. Several separate small-group sessions were held over a period of a month, attended by between one and three brokers, and lasted on average one and a half hours. At the beginning of a session, each participant was given two files, one containing colour-coded templates of RoutePlanner, and the other the associated questionnaires, similarly colour-coded. Participants completed the pre-study level-of-awareness questionnaire, the four stage-by-stage evaluation questionnaires, and then the post-study level-of-awareness questionnaire. The experiment was thus a single-group experiment in which the 'intervention' was the use of RoutePlanner, and the evaluation investigated each of the four stages/components of RoutePlanner. The session was facilitated by the researcher who used a PowerPoint presentation to structure the session and guide the processes.

The pre-study and post-study questionnaires were completed before and after the use of RoutePlanner, respectively. The evaluation-of-use questionnaire was completed to evaluate each stage of RoutePlanner. The learning experience was characterized by several of the 'C' features of contemporary learning theory, mentioned by De Villiers (2007). *Cognitive learning* was in evidence as participants reasoned their way through the identification of data items and mapped them to appropriate VMs. A *component-based* presentation style structured the sessions into manageable chunks, starting by referring to familiar visualizations such as bar charts, pie charts and scatter plots. This was not, however, at the expense of *constructivism*, in that although sessions were guided and advice was available, an open constructivist atmosphere prevailed as participants actively compared and selected VMs from the collection on hand, with a view to exploring data and gaining optimal insights. The team approach encouraged participative *collaborative learning* and lead to an open, discursive and communicative atmosphere. It appeared to be the best way for brokers to learn and to use RoutePlanner, namely, by consultative sessions with visualization experts. This approach helped brokers to overcome fear blocks relating to the unfamiliar. Finally, with regard to *creativity*, the set of processes contained in RoutePlanner form an innovative way of integrating existing guidelines, approaches in the body of visualization knowledge, and public-domain visualization resources into a consolidated supportive framework, with which even novices can rapidly gain confidence.

The quantitative and qualitative data was analysed and interpreted using descriptive statistics and thematic analysis respectively, to determine the utility of each stage in terms of positive and negative features, modifications suggested by participants, and the recognised relevance of each aspect to the work of the participant.

Results of data analysis and interpretation

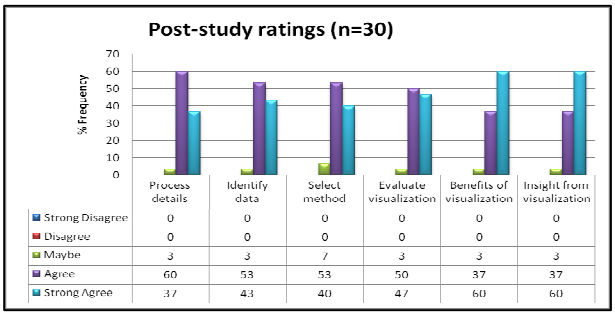


	Process details	Identify data	Select method	Evaluate visualization	Benefits of visualization	Insight from visualization
Strong Disagree	53	50	60	50	27	20
Disagree	33	33	23	30	27	23
Maybe	7	7	10	3	10	10
Agree	7	10	7	17	33	40
Strong Agree	0	0	0	0	3	7

Figure 1 Pre-study ratings

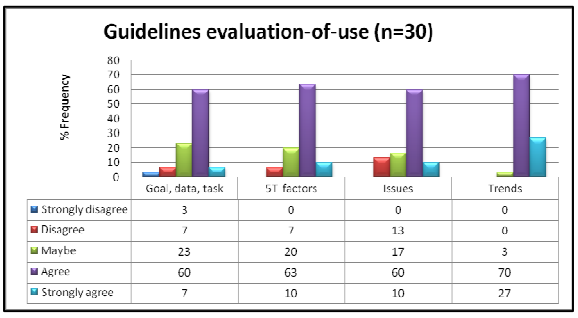
From the pre-study level-of-awareness data, it can be concluded that, although the participants were largely unaware of the details of the visualization process, as addressed by the first four statements in the questionnaire (average ratings 1.7, 1.8, 1.6 and 1.9), they were relatively aware of the benefits of the visualization and the insights to be gained from using it (average ratings of the last two statements: 2.6 and 2.9). Refer to Figure 1.

The post-study level-of-awareness data showed clearly, that awareness levels improved following the use of RoutePlanner. Most of the participants were acutely aware of the details of the visualization process (average ratings 4.3, 4.4, 4.3 and 4.4 for the first four statements) and had a greater awareness of its benefits, as well as the insights to be gained from the use of visualization in investigating their client data (average ratings 4.6 and 4.6 for the last two statements). Refer to Figure 2.



	Process details	Identify data	Select method	Evaluate visualization	Benefits of visualization	Insight from visualization
Strong Disagree	0	0	0	0	0	0
Disagree	0	0	0	0	0	0
Maybe	3	3	7	3	3	3
Agree	60	53	53	50	37	37
Strong Agree	37	43	40	47	60	60

Figure 2 Post-study ratings

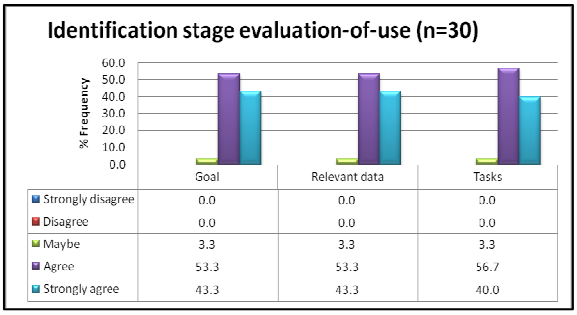


	Goal, data, task	5T factors	Issues	Trends
Strongly disagree	3	0	0	0
Disagree	7	7	13	0
Maybe	23	20	17	3
Agree	60	63	60	70
Strongly agree	7	10	10	27

Figure 3 Evaluation of guidelines

The Cochran-Armitage test was used to test for trend analysis of the experimental data. Item analysis was used in the calculation of a general awareness score for the pre-study and the post-study for each participant. In both analyses, the differences between the two sets of scores were proven significant. An evaluation study was conducted to assess the utility of each aspect of RoutePlanner (Figure 3 to 6).

RoutePlanner's guidelines: Each set of guidelines for visualization was evaluated: Brath's (1999) guidelines, the environmental factors of Mann (1999) and Mann & Reiterer (2000), the universal visualization issues of Spence (2001), and the trends in visualization. In each case, responses were positive with the majority of participants agreeing or strongly agreeing that they could use the guidelines. The percentage frequencies of 'Agree' and 'Strongly agree' varied from 67% to 97%. Refer to Figure 3.



	Goal	Relevant data	Tasks
Strongly disagree	0.0	0.0	0.0
Disagree	0.0	0.0	0.0
Maybe	3.3	3.3	3.3
Agree	53.3	53.3	56.7
Strongly agree	43.3	43.3	40.0

Figure 4 Evaluation of identification stage

Identification stage of RoutePlanner: Ninety seven percent (97%) of participants agreed or strongly agreed with the statements relating to the identification of data for visualization and the associated tasks. Refer to Figure 4.

Selection stage of RoutePlanner: Ninety seven percent (97%) of participants agreed or strongly agreed with the statements relating to the selection of appropriate VMs for the tasks. It was evident that they were learning to use the processes included in RoutePlanner. Refer to Figure 5.

Evaluation stage to determine whether the selected visualizations facilitate new insight: More than ninety percent (93%), agreed or strongly agreed that the processes for comparing visualizations and identifying the insight gained were effective, while 97% either agreed or strongly agreed that the factors to determine success of visualization or lack of success, can be applied. Refer to Figure 6.

From the qualitative data it could be concluded that the participants were orienting themselves accurately on the RoutePlanner journey and that they were able to use its stages. Furthermore, given the implementation of the evaluation stage, as undertaken during the small-group sessions, all the visualizations implemented during the intervention were deemed successful. A selection of these visualizations is presented in the associated poster.

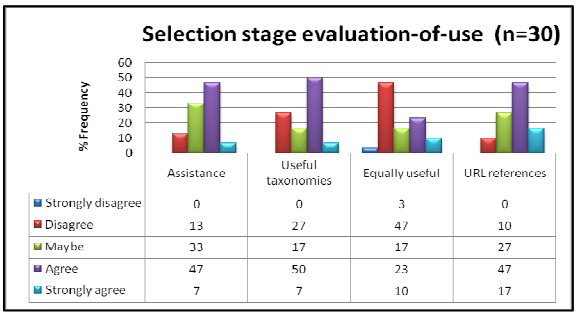


Figure 5 Evaluation of selection stage

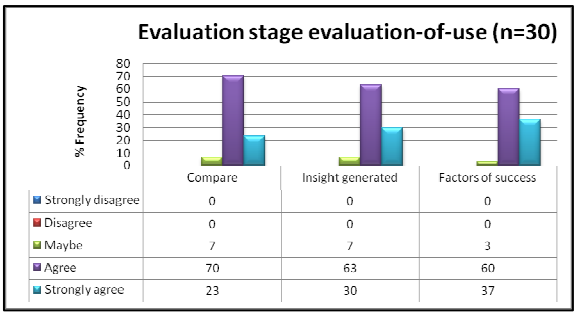


Figure 6 Evaluation of evaluation stage

Recommendations for Future Study

From the study, the following are recommended as future research:

- Revisit the participants to obtain metrics regarding the effect of the use of RoutePlanner on the insight gained into the data. Such a study would obtain quantifiable measurements indicating, not only the benefits of visualization, but also the extent of the learning acquired, i.e. the level of retention.
- Develop a visualization system comprising a set of visualization methods customized to the domain of insurance brokerage. Associated training should be developed to support the brokers in independent.
- Extend the use of RoutePlanner to other domains and undertake associated evaluation studies.

Conclusion

This study addresses the need for visualization in the real world domain of insurance brokerage. With this goal, RoutePlanner, a model of the visualization process, was developed. We now re-visit the research questions, namely:

- Does the use of RoutePlanner lead to an increased awareness of visualization on the part of brokers? RoutePlanner was used in an experiment to determine whether its use by domain experts, in collaboration with visualization- and a data expert, leads to an increased awareness of visualization and results in the successful visualization of data. In both Cochran-Armitage and item analysis, differences between pre- and post-study scores were significant, thus an increased awareness of visualization was achieved.
- Does the model of the visualization process result in successful visualizations of insurance brokerage data? From the evaluation-of-use findings, it was concluded that RoutePlanner has utility, all stages are relevant to the work of domain experts, and visualizations were successful in terms of the enhanced insight acquired. The collaboration between domain, data and visualization experts fostered the learning and application of RoutePlanner. Most of the participants (over 70%), who responded to the relevant question, preferred the supportive team approach rather than a solo visualization approach.

It is anticipated that, as technological expertise, exposure to visualization, training and orientation increase, and ultimately, when a visualization system is developed comprising VMs customized to the domain of insurance brokerage, brokers will acquire the expertise and skill to use the visualization process with decreasing collaboration.

This study makes a contribution to the focus area of the process of information visualization. RoutePlanner, as a model of the visualization process, is an innovative supportive framework that embodies existing guidelines, a variety of approaches in the body of visualization knowledge, and public-domain visualization resources. It thus serves to integrate multiple aspects of the visualization process for an audience that was originally uninformed in the area of visualization.

References

Brath, R. K. (1999). *Effective information visualization: Guidelines and metrics for 3D interactive representations of business data.* Unpublished MA. (Computer Science). Toronto: University of Toronto.

Card, S. K., Shneiderman, B., & Mackinlay, J. D. (1999). *Readings in information visualization: Using vision to think*. San Francisco, Calif.: Morgan Kaufmann Publishers.

de Oliveira, M. C. F., & Levkowitz, H. (2003). From visual data exploration to visual data mining: A survey. *IEEE Transactions on Visualization and Computer Graphics, 9*(3): 378-394.

De Villiers, M.R. (2007). The six C's framework for e-learning. In Buzzetto-More, N.A. (ed.) *Advanced Principles of Effective e-Learning* (pages 1-25). Santa Rosa, CA: Informing Science Press.

Espinosa, O. J., Hendrickson, C., & Garrett, J. H. (1999). Domain analysis: A technique to design a user-centered visualization framework. In: *IEEE Symposium on Information Visualization (Info Vis '99):* 44-52, 144. Washington, D.C, USA: IEEE Computer Society Press.

Gouws, P. M. (2009). *RoutePlanner – a model for the visualization of warehouse data.* Unpublished MSc. (Information Systems). Pretoria: University of South Africa.

InfoVis ToolKit (2008) *The InfoVis ToolKit [On-Line]* Available: http://ivtk.sourceforge.net/ . Accessed on 2008/12/19.

Kao, D. L., & Ma, K. L. (2000). The life cycle of a visualization case study. *IEEE Computer Graphics and Applications, 20*(5): 29-31.

Keim, D. A. (2002). Information visualization and visual data mining. *IEEE Transactions on Visualization and Computer Graphics, 8*(1): 1-8.

Lengler, R., & Eppler, M. J. (2007). Towards a periodic table of visualization methods for management. In: *Graphics and Visualization in Engineering (GVE2007):* 1. Calgary, AB: Acta Press.

Mann, T. M. (1999). Visualization of WWW search results. In: *Proceedings of the 10th International Workshop on Database and Expert Systems Applications (DEXA '99):* 264. Washington, DC, USA: IEEE Computer Society.

Mann, T. M., & Reiterer, H. (2000). Evaluation of different visualization of www search results. In: *Proceedings of the 11th International Workshop on Database and Expert Systems Applications:* 586. Washington, D.C., USA: IEEE Computer Society.

Mouton, J. (2003). *How to succeed in your master's and doctoral studies : A south african guide and resource book*. Pretoria: Van Schaik Publishers.

North, C. (2006). Toward measuring visualization insight. *IEEE Computer Graphics and Applications, 26*(3): 6-9.

OLIVE (1997) *On-line library of information visualization environments. [On-Line]* Available: http://otal.umd.edu/Olive/ Accessed on 2008/11/18.

Olivier, M. S. (2004). *Information technology research : A practical guide for computer science and informatics* (2nd ed.). Pretoria: Van Schaik Publishers.

Saraiya, P., North, C., & Duca, K. (2005). An insight-based methodology for evaluating bioinformatics visualizations. *IEEE Transactions on Visualization and Computer Graphics, 11*(4): 443-456.

Shneiderman, B. (1996). The eyes have it: A task by data type taxonomy for information visualizations. In: *Proceedings of the 1996 IEEE Symposium on Visual Languages (VL '96):* 336. Washington, DC, USA: IEEE Computer Society.

Spence, R. (2001). *Information visualization*. Essex, England: Addison-Wesley.

Wiss, U., Carr, D. A., & Jonsson, H. (1998). Evaluating three-dimensional information visualization designs: A case study of three designs. In: *Proceedings of the International Conference on Information Visualisation (IV '98):* 137. Washington, DC, USA: IEEE Computer Society.

VisTree (2008) *The Visualization Tree [On-Line].* Available: http://gbdi.icmc.usp.br/~junio/VisTree/Conteudo/Download.htm Accessed 2008/12/19.

XmdvTool (2008) *Interactive visual exploration of multivariate data [On-Line].* Available: http://davis.wpi.edu/~xmdv/ Accessed 2008/12/19.