TreemapBar: Visualizing Additional Dimensions of Data in Bar Chart

Mao Lin Huang[1], Tze-Haw Huang[1] and Jiawan Zhang[2]
[1] Faculty of Engineering and Information Technology, University of Technology, Sydney, Australia
[2] School of Computer Science and Technology, Tianjin University, China
{maolin@it.uts.edu.au, thhuang@it.uts.edu.au, jwzhang@tju.edu.cn}

Abstract

Bar chart is a very common and simple graph that is mainly used to visualize simple x, y plots of data for numerical comparisons by partitioning the categorical data values into bars and typically limited to operate on highly aggregated dataset. In today's growing complexity of business data with multi dimensional attributes using bar chart itself is not sufficient to deal with the representation of such business dataset and it also not utilizes the screen space efficiently. Nevertheless, bar chart is still useful because of its shape create strong visual attention to users at first glance than other visualization techniques. In this article, we present a treemap bar chart + tablelens interaction technique that combines the treemap and bar chart visualizations with a tablelens based zooming technique that allows users to view the detail of a particular bar when the density of bars increases. In our approach, the capability of the original bar chart and treemaps for representing complex business data is enhanced and the utilization of display space is also optimized.

Keywords--- **Information visualization, treemap, bar chart, tablelens, focus+context**

1. Introduction

The bar chart is the fundamental visualization method that was frequently employed in reporting analysis of mapping categorical data onto a visual display in simple *x* and *y* form for numerical comparisons. It operates on highly aggregated dataset and is only capable of interpreting the simple information. In business reporting, data analyst usually uses the color coding scheme and data labeling on bar chart to uncover additional dimension of business data and enhance the readability of graph to the users Nevertheless, bar chart in its nature is very basic (one dimension) visualization that merely meets the basic requirement of exponentially growing and complex dimensions of contemporary business data.

Generally, fundamental graphs can be classified into three major categories: *line, bar* and *pie* graphs in which bar is commonly used for visualizing volume data, count data and simple statistics. The popularity of bar chart can be attributed to its variable rectangle block with color coded bar since colored coding scheme created strong visual impact and therefore attracts more visual attention than a *point* or *line* on a display. For example, bar chart highlights the highest data value at a glance without forcing the reader to search for background information such as data label. However, if the dataset is multi attributes or dimensions then some information needs to be discarded in order to categorize the data into aggregated data values. This limitation is especially inconvenient if the decision maker is interested in exploring the relationship of multi attributed business dataset for the relationship between product sale volumes, quantities and types. The traditional approach usually uses separate bar charts to bring the information together to support the business decision and this chart separation creates excessive information overlap, eye movement of the user and does not support the discovery of subset linkage. Furthermore, E. Tufte [10] also pointed out that bar chart contains considerable redundancy, since data value is conveyed by several features. In order to overcome the shortfall of bar chart, we proposed a TreemapBar visualization technique that combines the treemap and bar chart that allows the redundant space inside the bars to be used for the display of another visualization to uncover more dimensions and attributes of data than the traditional bar chart.

The remainder of this paper is organized as follow; we discuss the related work in section 2 where the background information is given and in section 3 the detailed TreemapBar will be explained. In section 4 we will provide a case study about the application of TreemapBar and finally a conclusion.

2. Related Work

2.1 Bar Display

Bars are rectangular areas on a graph where their heights proportional to their data value. It has been mainly used to show highly aggregated data and often have a high degree of overlap [14], see Figure 1a.

978-0-7695-3733-7/09 $25.00 © 2009 IEEE
DOI 10.1109/IV.2009.22

2.2 Treemap

Treemap is a space filling visualization technique, first proposed by Shneiderman [1, 9] originally designed to fully utilize the screen space for recursively displaying the hierarchical structure in a rectangular area. The key ingredient of treemap is its layout algorithm [2, 6, 7] that works by recursively divide the available space into nested rectangles and also determines the aspect ratio of the division of rectangles. The optimal layout algorithm is to produce rectangles with aspect ratio closes to one as possible.

Slice and Dice Treemap [9] is the first and simplest treemap layout algorithm by recursively dividing a rectangle into sub-rectangles using parallel lines and sub-rectangles represent children to its parent rectangle.

Squarified Treemap [7] is a treemap layout algorithm that works by layout rectangles in horizontal and vertical rows. Either the rectangle is added to the current row or the current row is fixed and a new row is started n the remaining sub-rectangle based on the decision that as long as the worst aspect ratio of any rectangles in the current row keep improving.

Treemap appears to be the most desirable visualization technique in our scenario.

2.3 TableLens

Tablelens[12] is a focus+context [17] viewing technique that works effectively in browsing the tabular information by merging graphical and symbolic representation into an interactive view. An important feature of the TableLens is that the distortion of views in either horizontal or vertical order is independent from each other.

3. TreemapBar

TreemapBar is derived from both treemap and bar chart and the basic idea is to take advantage of treemap's space filling feature by embedding it within bars in order to fully utilize the display space inside bars for provision of additional dimensions of visualization. Each bar corresponds to a categorized data subset and the subset might contain hierarchical structures that will be encoded by the treemap. In Figures.1a and 1b, we compare the traditional bar chart with new TreemapBar that has equal width of bars respectively. Figure.1c shows the high density of bars and Figure.1d shows the application of a tablelens based focus+context viewing to allow users to see the detail of a particular bar in high dense bar charts. The TreemapBar extends the capability of the bar chart from the simple reporting to the visualization of hierarchical dataset.

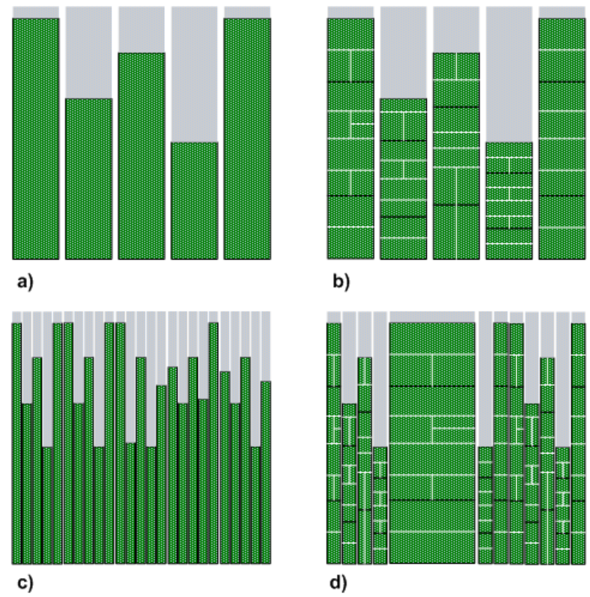
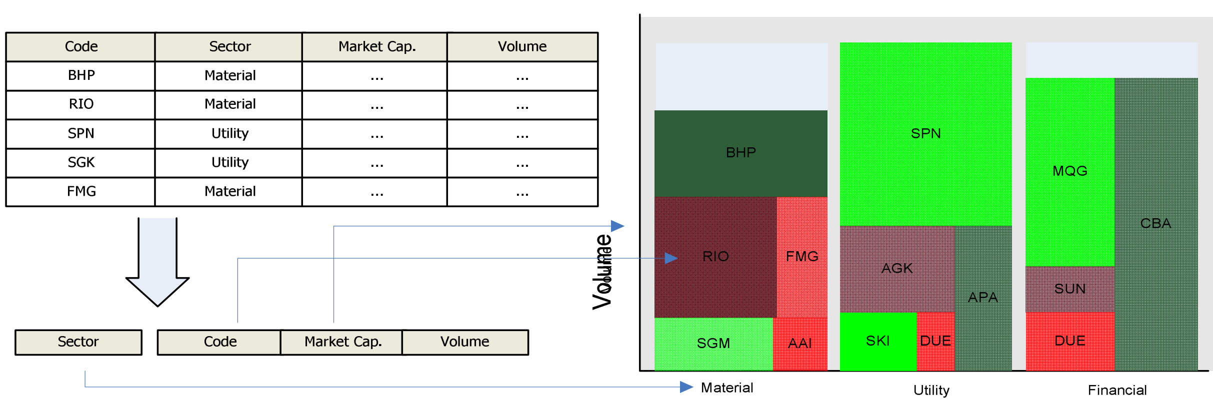


Figure.1 a) Traditional equal width bar chart b) TreemapBar c) Bar chart with high density d) TreemapBar chart with TableLens' focus.



Figur.2 Example of TreemapBar construction process starts from data partitioning.

3.1 TreemapBar Charts

In most cases, the categorical dataset naturally contains the hierarchical structures, for example, in stock market we can use treemap to present the business sector structures, stock industry classifications, and time hierarchies (by year, quarter, month and week) for time dimension. It is trivial for treemap bar chart to visualize the distribution of data values and also the hierarchical classifications of companies. Companies also have a natural size of area in treemaps namely their market capitalization. The price performance can be indicated by color coding. Essentially, the visualization process of TreemapBar starts with the partitioning of data clusters (e.g. business sectors) into bars of equal width and then draw the treemaps in each bar by using the Squarified Treemaps layout algorithm, where the import size of rectangular areas are real values (e.g. the capitalization of industry companies).

Figure.2 illustrated an example of constructing TreemapBar by mapping the financial stock market data into the visualization where:

1. "Code" C_g is the price change (capital gain/loss) indicated by colors. Thus, C_g is a function of the hexadecimal color value *rgb (x, y, z)*,
2. "Sector" is a group of industry companies who are running the same type of business,
3. "Market Cap" is the capitalization of a particular company, and
4. "Volume" is the total amount of money of daily transactions in stock exchange.

This example demonstrated that in comparison with the original bar chart, the proposed TreemapBar can show two additional dimensions of data attributes: 1) the business classification of companies, and 2) the market capitalizations of companies, in visualization. On the other hand, in comparison with the traditional treemaps application "SmartMoney" [21], our approach can show extra two types of information to the user; these are 1) the volume of the daily transactions, and 2) the comparison outcome of transaction volumes among different business sectors. In this research, we are not concerned with the information overloading because it involves with the human visual perception, intuition and preferential usage that greatly differs from one to another and it is beyond the scope of this work.

3.3 Focus+Context Viewing via Tablelens

The discernability issue is raised when the bar density increases and the hierarchical structure is deep. The discrepancy of treemap will be difficult to discern due to the increased child nodes in the hierarchical structure which requires the assistance of focus+context technique to highlight the current context information in order to enhance the readability and such issue is often addressed by the adoption of the local focus+context technique. The approach we use is similar to the TableLens in order to highlight the detail of a selected bar and shrink others as shown in Figure.3.

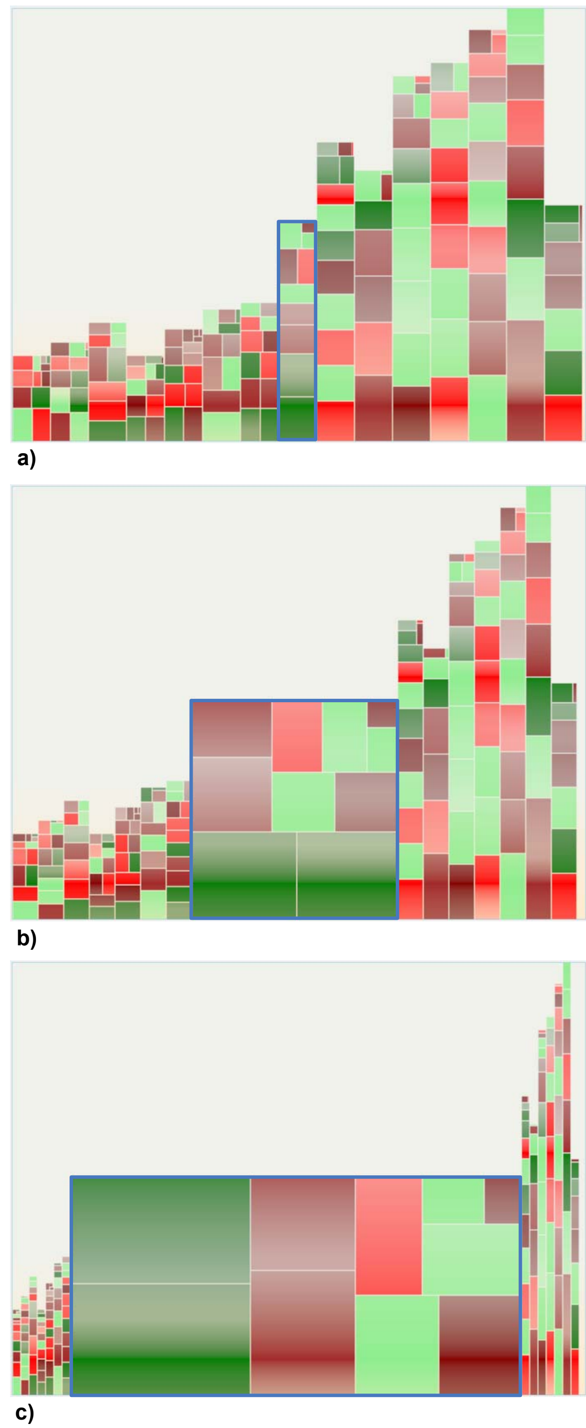


Figure.3 a) High density of treemap bar without tablelens, b) treemap bar + tablelens with moderate scale distortion c) treemap bar + tablelens with higher scale distortion.

The focus+context viewing scheme supports several human adjustments to control the display of focus area. The viewer can manually:

1. Adjust the amount of display space allocated to the focus bar without reducing the number of bars in the context view, see Figure 3b.
2. Adjust the amount of contents viewed within the focus bar without changing the size of the bar. The contents may be the enclosure structures and labels.
3. Adjust the location of the focus bar within the context, see Figure 3c.

3.4 Labeling

Effective visualization requires the appropriate labeling for the enhancement of human perception and cognition of the visualization context. It also assists the readers to clarify the implication of the graph objects for information they represent without referring to the actual dataset and the most common form of labeling is the textual label that attaches to an object.

In general, labeling is categorized into *static* and *dynamic* labeling [18] with the static labeling approach the placement of label is based on best effort method by aggregating the information to be labeled and the heuristic for the amount of information to be labeled is application dependent. For example, visualization might decide not to place a label on a graph object if such placement might degrade the visibility of the overall presentation.

In the dynamic approach, the labeling of an object can be turn on or off depends on the constraint i.e. textual label size is larger than the allowable space. The common condition to trigger the dynamic labeling is the cursor movement or mouse hover. According to Dogrusoz et al [19] that visualization with good label placement should exhibit the following basic rules:

- Elimination of ambiguity,
- Clarity, and
- Flexibility.

We use a static+dynamic labeling approach to place labels on the rectangles that is:

- The font size f_s of a label is set between the range 4 to 80 and it is a linear function of the size of rectangle in which the label is placed.
- If $4 \leq f_s \leq 80$, then we use the static labeling to place the label with the actual font size f_s into the rectangle.
- If the font size $f_s < 4$, then we assume that the font size is too small for viewers to perceive thus, a dynamic label placement approach is applied and we set $f_s = 4$.

The dynamic labeling is performed by cursor movement based on Excentric labeling [18] as illustrated in Figure.4.

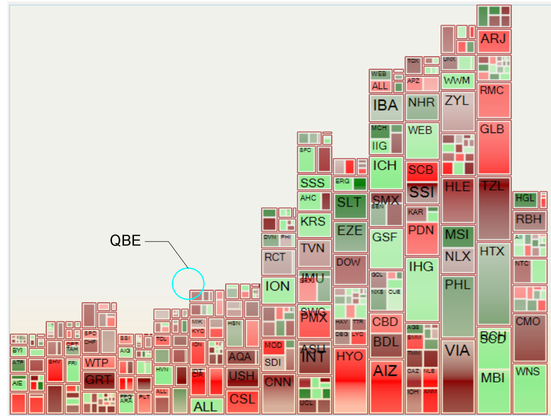


Figure.4 The static+dynamic labeling based on mouse hover.

4. Case Studies

4.1 Market Analysis via Sector Indices

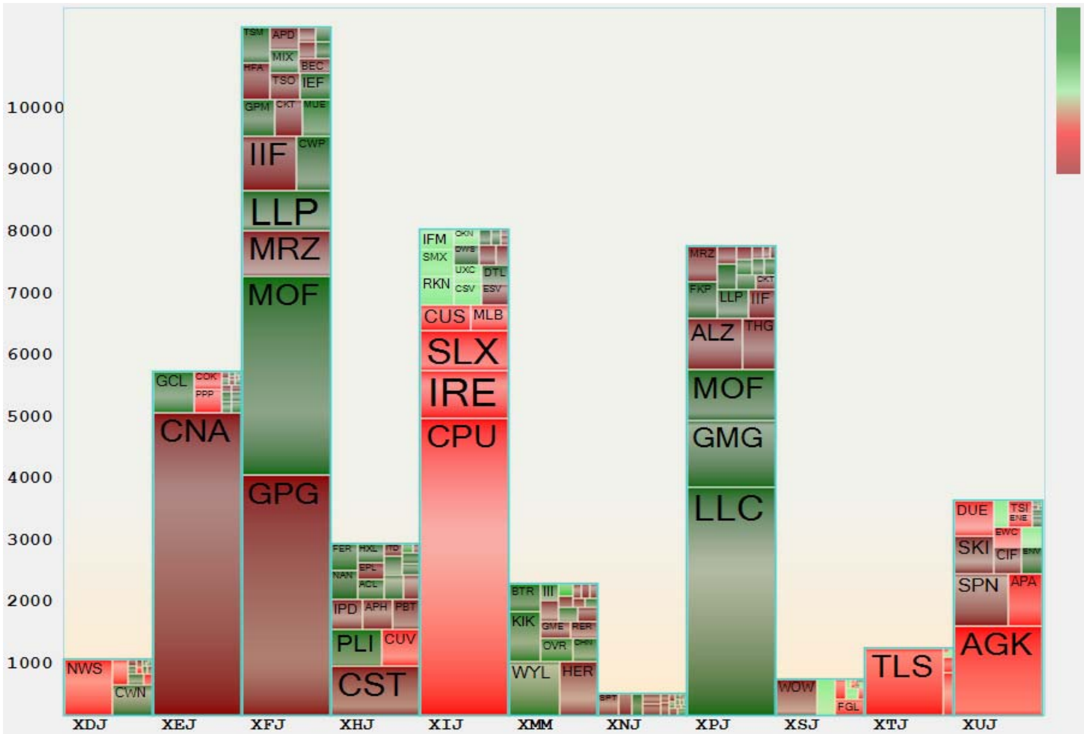
Treemap has been widely applied to various financial applications and the one of most famous implementation is the SmartMoney [21] which uses treemap to display the stock information. However, as discussed before the SmartMoney is unable to show the volume of the daily transactions through the *Y-axis*, as well as the comparison outcome of transaction volumes among different business sectors. Furthermore, the graphic format of bar chart is more user friendly that is commonly understood and acceptable than Treemaps which is relatively new to all kinds of users.

4.2 Stock Market Dataset

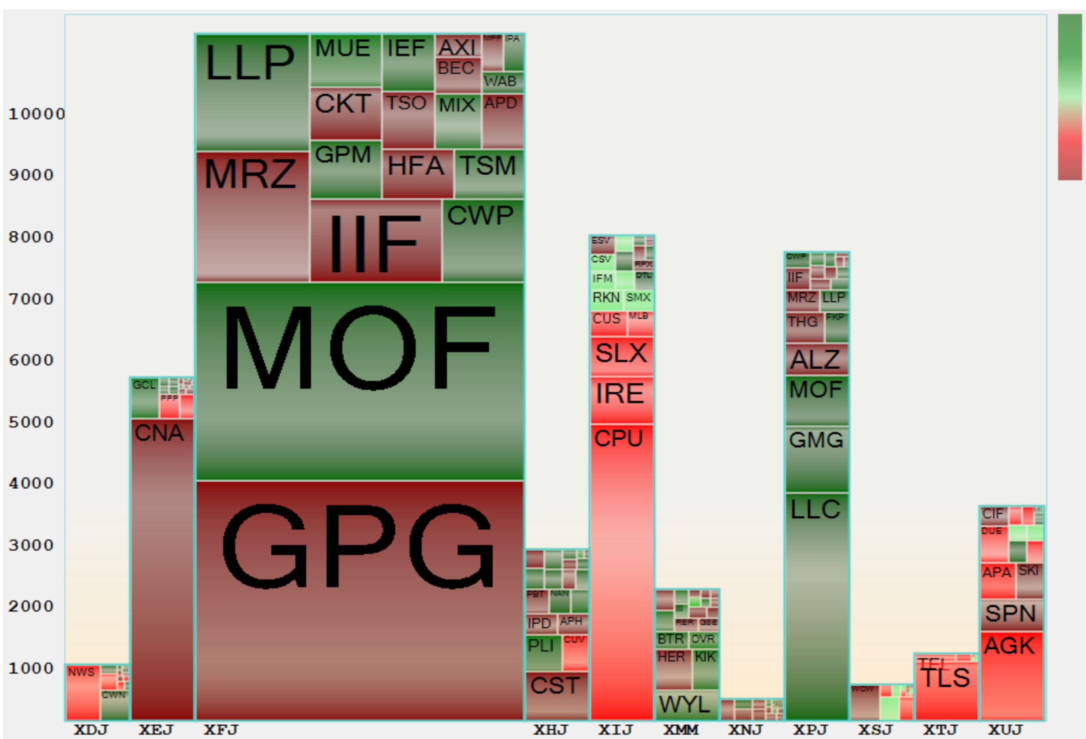
The example data used to generate the TreemapBar is based on Australia Stock Exchange (ASX) where companies are categorized into industrial sectors and each sector has its own sector index. For example, for financial and energy sector the market index is XFJ and XEJ respectively. The dataset is partitioned into categories according to sector indices and each sector contains number of rows (companies).

4.3 Stock Visualization via TreemapBar

In this case study, we selected top 10 gainers and 10 losers for each sector in ASX of a transaction day. The *X-axis* represents the sector category and *Y-axis* is the indices. Therefore, each bar consists of 20 rectangles. The size of rectangle determines by the market capitalization of the company. Figure.5a shows the visualization result of TreemapBar by applying it on ASX data and Figure.5b shows the application of focus+context viewing on the same TreemapBar.



a)



b)

Figure.5 a) Stock analysis with the normal view of TreemapBar, b) Stock analysis with a focused view of TreemapBar, where X axis represents the industry sectors and Y represents the index.

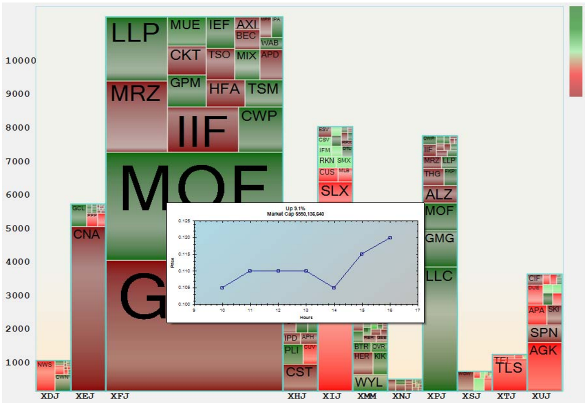


Figure.6 Exploring the change of share price of a trading day of a company though the TreemapBar visualization.

4.4 Data Exploration

Data exploration is one of important feature in our visualization because it can greatly enhance the user experience of TreemapBar. When the user clicks on a particular rectangle area (representing a company in an industry sector), a popup window appears that displays the daily share price trend in normal price chart manner as illustrated in Figure.6.

Conclusions

The primary aim of this research is attempting to extend the capability of original bar chart to visualize the dataset with multiple dimensions that meets the demand for visualizing growing complex business dataset and enhancing the user understanding of such multi dimensional datasets without adding additional graphs in the display. In conclusion, this paper proposed a new visualization technique TreemapBar with the TableLens viewing to increase the capability of traditional bar chart and treemaps in representing complex business data. Also we have applied our developed methodology to a financial stock analysis via sector indices.

References

[1] B. Johnson and B. Shneiderman, “Tree-Maps: A space-filling approach”, Proceeding of IEEE Conference on Visualization 1991. pp 284-291. San Diego, USA.

[2] M. Balzer and O. Deussen, “Voronoi Treemap”, IEEE Symposium on Information Visualization 2005. pp 49-56. Minneapolis, USA.

[3] S. Zhao, M. J. McGuffin and M. H. Chignell, “Elastic Hierarchies: Combining Treemaps and Node-link Diagrams”, IEEE Symposium on Information Visualization 2005. pp 57-64. Minneapolis, USA

[4] H. Lu and J. Fogarty, “Cascade Treemap: Examining the Visibility and Stability of Structure in Treemaps”, Graphics Interface 2008, pp 259-266.

[5] M. J. McGuffin, G. Davison, and R. Balakrishan, “Expand-Ahead: a Space filling strategy for browsing Trees”, IEEE Symposium on Information Visualization 2004. pp 119-126, 2004.

[6] B. Shneideman and M WattenBerg, “Ordered Treemap Layouts”, IEEE Symposium on Information Visualization 2001. pp 73-78.

[7] M. Bruls, K. Huizing and J. J. van Wijk, “Squarified Treemaps”, Proceeding of Joint Eurographics and IEEE TCVG Symposium on Visualization (TCVG) 2000, pp 33-42.

[8] A. Reese, “Bay! Bar Charts”, Significance. Volume 4, pp 41-44. March 2007.

[9] B. Shneiderman, “Tree Visualization with Tree-Maps: 2D Space-Filling Approach”, ACM Transactions on Graphics, pp92-99, 1992.

[10] E. R. Tufte, “The Visual Display of Quantitative Information”, Cheshire, Connecticut: Graphics Press 2001.

[11] Daniel A. Keim, Ming C. Hao and Umeshwar Dayal, “Hierarchical Pixel Bar Charts”, IEEE Transactions on Visualization and Computer Graphics. Volume 8, No. 3, July 2002.

[12] R. Rao, S. Card, “The Table Lens: Merging Graphical and Symbolic Representation in an Interactive Focus+Context Visualization for Tabular Information”, Proc. of SIGGHI Conference on Human Factors in Computing Systems, ACM, pp 318-322. April 1994.

[13] R. Vliegen, J.J. van Wijk, E.J. van der Linden, Visualizing Business Data with Generalized Treemaps”, IEEE Transaction on Visualization and Computer Graphics. Volume 12. No. 5. pp. 789-796, Sept. 2006.

[14] M. C Hao, U. Dayal and F. Casati, “Visual Mining Business Service using Pixel Bar Chart”, HP technical report. 2004.

[15] D. A Keim, M. C Hao and M. Lyons, “Value-Cell Bar Charts for Visualizing Large Transaction Data Sets”, IEEE Transaction on Visualization and Computer Graphics. Volume 13. Issue 4. pp 822-833. July 2007.

[16] D. A Keim, M. C. Hao, U. Dayal and M. Hsu, “Pixel Bar Charts: a Visualization Technique for Every Large Multi-Attribute Dataset without Aggregation”, IEEE InfoVis 2001. San Diego, USA.

[17] G. W. Furnas, “Generalized Fisheye Views”, ACM SIGCHI Conference on Human Factors in Computing System, pp 16-23. April 1986.

[18] J.D Fekete, C. Plaisant, “Excentric Labeling: Dynamic Neighborhood Labeling for Data Visualization”, CHI’99. pp 512-519. New York.

[19] U. Dogrusoz, K. G. Kakoulis, B. Madden and I. G. Tollis, “On Labeling in Graph Visualization”, Information Sciences, Volume 177. Issue 12. pp 2459-2472. 2007.

[20] S. G. Eick, “Visualization and Interaction Techniques”, in CHI’97 Tutorial notes on Information Visualization. ACM SIGCHI. March 1997.

[21] SmartMoney Software: http://www.smartmoney.com/ http://www.smartmoney.com/map-of-the-market/