DOI: 10.1111/cgf.12931
EuroVis 2016
R. Maciejewski, T. Ropinski, and A. Vilanova
(Guest Editors)

Volume 35 (2016), Number 3
STAR – State of The Art Report

A Survey on Visual Analysis Approaches for Financial Data

S. Ko[†], I. Cho[‡], S. Afzal[§], C. Yau[§], J. Chae[§], A. Malik[§], K. Beck[§], Y. Jang[¶], W. Ribarsky[‡], D. S. Ebert[§]

[†] Ulsan National Institute of Science and Technology (UNIST)
[‡] University of North Carolina at Charlotte
[§] Purdue University
[¶] Sejong University

Abstract

Market participants and businesses have made tremendous efforts to make the best decisions in a timely manner under varying economic and business circumstances. As such, decision-making processes based on financial data have been a popular topic in industries. However, analyzing financial data is a non-trivial task due to large volume, diversity and complexity, and this has led to rapid research and development of visualizations and visual analytics systems for financial data exploration. Often, the development of such systems requires researchers to collaborate with financial domain experts to better extract requirements and challenges in their tasks. Work to systematically study and gather the task requirements and to acquire an overview of existing visualizations and visual analytics systems that have been applied in financial domains with respect to real-world data sets has not been completed. To this end, we perform a comprehensive survey of visualizations and visual analytics. In this work, we categorize financial systems in terms of data sources, applied automated techniques, visualization techniques, interaction, and evaluation methods. For the categorization and characterization, we utilize existing taxonomies of visualization and interaction. In addition, we present task requirements extracted from interviews with domain experts in order to help researchers design better systems with detailed goals.

Categories and Subject Descriptors (according to ACM CCS): Document types [General and reference]: Surveys and overviews—

1. Introduction

The availability of different sources of financial data provides opportunities to market analysts and investors to extract new insights in order to make informed decisions. Such analyses guide them in the development of optimal investment and risk management strategies. To facilitate such analysis tasks, the analysts traditionally utilize visual analysis tools that are typically built atop standard statistical methods (e.g., moving averages, regression). Examples of these tools include line graphs and candlestick charts [EMB07] that are popular among financial market professionals for decision making tasks [SB13].

Despite the popularity of these techniques, these tools are often inadequate to handle the problems that arise from big data. With the advent of the digital age, an unprecedented amount of real-time financial data is generated in different sectors, including asset trading, news, and economic indicators. Thousands of stocks with multivariate attributes (e.g., price, time, volume) get traded every few seconds, and various reports on the stock market and economic indicators are continuously published for stakeholders. In the big data era [KKEM10, San13, TC05], the conventional approaches of analyzing these datasets using standard analysis techniques are often marred with limitations (e.g., scalability, overplotting issues, not enough inputs from domain experts). For example, line charts still remain popular among financial analysts as they incorporate different aggregation methods to reduce overplotting and occlusion issues [SB13]. However, such aggregations can cause a loss of information that may be undesirable in the analysis of financial data (e.g., comparison of sales and transactions). Analysts are often required to explore and consider discrete and fragmented pieces of information at the same time for making informed decisions.

The complexity and size of these financial datasets have attracted the attention of the information visualization and visual analytics communities. This has led to the rapid development of many visual analytic systems and interaction methods (e.g., bank wire transaction [CGK*07], corporate sales [KMJE12]). The development of such systems often requires visual analytics researchers to collaborate with financial experts to assess their requirements and challenges, and design appropriate visual analytics solutions to address them. This provides the impetus to systematically study and gather

[†] sako@unist.ac.kr
[‡] {icho1 | ribarsky}@uncc.edu
[§] {safzal | yauc | jchae | amalik | kaethe | ebertd}@purdue.edu
[¶] jangy@sejong.ac.kr, corresponding author

the needs and requirements of financial practitioners from different domains and to gain an overview of the existing visual analytics techniques that have been applied with respect to their datasets.

To address this need, researchers have conducted several surveys that provide an overview of the existing works in the financial domain. For example, Pryke [Pry10] discusses the role of visualization techniques in financial industry applications. Schwabish's survey [Sch14] provides an introduction and guidance to economists. More recently, Flood et al. [FLVW14] survey applications designed for monitoring financial stability. However, these surveys are limited in scope in terms of the number of systems surveyed and the financial domain targeted by the researchers. This encouraged us to conduct interviews with domain analysts (Section 4) that gave us new thoughts and insights. From the interviews, we found that there were various requirements, start and end points, and types of decisions made in financial analyses. For example, an analysis of financial anomalies usually required the analysts to form an interactive communication loop by two small groups with different task goals. In the analysis, one group started detecting and predicting anomalies as soon as data become available. Once this group derived initial results and moved to a second time analysis, another group started validating the initial results to generate reports and search for unknown and missing information. Then when the second finished their tasks, the conclusions from the group were transferred to the first group for further discussion. This may indicate that if researchers had considered real-world analytical patterns and requirements in the design phase, their final systems could have been be more beneficial.

To this end, we survey research papers in the information visualization and visual analytics domains that target the financial market sector. We categorize and characterize financial systems in terms of data sources, applied automated techniques and visualization techniques, interaction methods, and evaluation. In doing so, we utilize existing visualization [Kei02] and interaction [YaKSJ07] taxonomies to support effective understanding and navigation. Here, Keim's visualization taxonomy [Kei02] (Section 7) defines that existing data can be visualized with 2D, 3D visualization techniques, geometrically-transformed, iconic and stacked displays. Yi's interaction taxonomy [YaKSJ07] (Section 8) presents that existing interaction methods can be one of Select, Explore, Reconfigure, Encode, Abstract, Filter, and Connect.

The contributions of this work are two-fold. First, we identify and summarize general financial analysis task requirements by interviewing analysts with a background in financial analysis, risk management, and capital market investment management. Second, we go beyond the above surveys to perform a survey of recent developments across the information visualization and visual analytics domains categorizing previous approaches based on the input data, the automated techniques, visualizations, interaction, and evaluation methods. We concentrate mainly on research publications and also include visual analysis tools from the financial industry if enough information can be acquired from the research publications. Our work is intended to assist financial analysts and researchers in their common tasks, which include the following: 1) identify financial data sources used by stakeholders in their analysis and decision making, 2) provide an overview of data transformation and analysis techniques, and visualization techniques applied to produce desired results, 3) understand common user interaction approaches for analyzing financial data, and 4) provide evaluation strategies that are applied to visual analytic systems to show usability of the systems.

Financial Risk Management	Industrial Process Control
Operations Planning	Capital Markets Management
Military Strategic Planning	Network Monitoring
Marketing Analysis	Derivatives Trading
Fraud/Surveillance Analysis	Portfolio Management
Actuarial Modeling	Customer/Product Analysis
Budget Planning	Operations Management
Economic Analysis	Fleet/Shipping Admin

Table 1: *16 typical business application domains [Teg99].*

This paper is structured as follows. First, we present related work in surveying financial systems (Section 2) and the methodology used in our survey (Section 3). Then, we categorize previous work based on data type (Section 5), applied data transformation and analysis techniques (Section 6), visualization (Section 7), interaction techniques (Section 8), and evaluation methods (Section 9). After task requirements from analysts are described in Section 4, we discuss trends, findings and insights acquired from surveying papers in Section 10 and conclude our work in Section 11.

2. Previous Surveys

In this section, we discuss related surveys for visual financial data analysis. The survey conducted by Tegarden [Teg99] is the first survey where information visualization techniques in industry systems are highlighted. This work is important in that it defines 16 typical business application domains as shown in Table 1. There are two surveys from perspectives of economics [Pry10, Sch14]. With a goal to examine implications of visualizations on industry applications, Pryke [Pry10] conducted interviews with representatives of software companies and financial organizations and presents financial industry applications. After examination of applications based on interviews, Pryke concludes that visual approaches should be more highlighted for financial market analysis. However, little information is revealed on task requirements from the interviews due to a proprietary issue. The work conducted by Flood et al. [FLVW14] is different from others in that Flood's work mainly focuses on systems for monitoring financial stability and benefits of Visual Analytics [TC05] in performing the monitoring tasks. Dumas et al. [DML14, Fin] provide an excellent on-line visual overview with filtering options for financial visualization systems.

Compared to the surveys described above, there are three benefits of our work. We provide the most comprehensive and detailed survey on visualizations techniques and visual analytics approaches that have been applied to financial data. We classify previous systems for financial data analysis into several categories: data sources, automated techniques, visualizations, interactions, and evaluations, and derive patterns and insights for each category. In addition, we present task requirements based on interviews with

IV	International Conference on Information Visualization
InfoVis	IEEE Symposium on Information Visualization
VAST	IEEE Conference on Visual Analytics Science and Technology
EuroVis	Joint Eurographics-IEEE Symposium on Visualization
CHI	ACM Conference on Human Factors
VIS	IEEE Conference on Visualization
IVS	Information Visualization (published by SAGE)
JVIS	Journal of Visualization
SIGKDD	ACM Conference on Knowledge Discovery and Data Mining
CGA	Computer Graphics and Applications
HICSS	Hawaii International Conference on System Sciences
CG	ACM Conference on Computer Graphics

Table 2: *Searched venues for paper collection.*

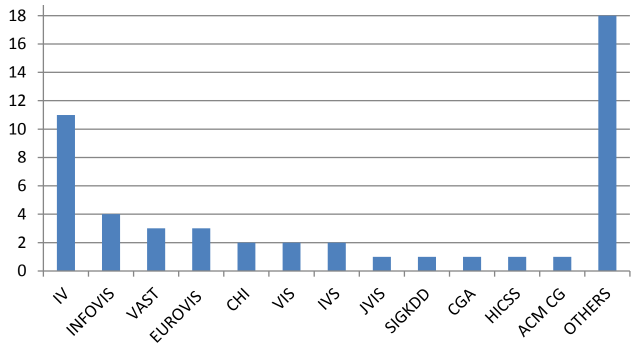

Figure 1: *Number of papers published from venues. The abbreviations of the venues are explained in Table 2. "Others" sums all venues other than those in Table 2.*

a number of analysts whose backgrounds are varied. We think that the presented requirements are able to complement low-level analysis operations in tasks (9 operations in [MSK*06] and 6 operations in [Mar07]). We believe that our work is able to support industry experts in finding the best suitable visualization techniques and visual analytics approaches for their data and to help researchers in the visualization and analytics communities understand what functionalities or visualizations are needed in the tasks for various business domains.

3. Methodology

In order to provide an overview of visualizations and visual analytics systems for financial data, we took several steps. The first step was collecting and reviewing papers. After this step, we found that there was little information about task requirements and a data exploration model directly derived from financial industry experts. Therefore, we decided to interview analysts from the financial domain to derive generalized domain requirements and a data exploration model. Lastly, with the extracted task requirements and a data exploration model in mind, we categorized the collected research papers. In the next paragraphs, we describe in detail the process of collecting papers, eliciting task requirements, and adapting a visual data exploration pipeline.

For the paper collection, we started from previous surveys described in Section 2 with a focus on papers that incorporate visualization techniques in the analysis loop. Then, we utilized search

ACM International Working Conference on Advanced Visual Interfaces
Innovative Applications of Artificial Intelligence Conference
Journal of Empirical Finance
European Conference on Information Systems
International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision
Asia-Pacific Symposium on Information Visualization
International Conference on Digital Government Research
New Frontiers in Applied Artificial Intelligence
Journal of Universal Computer Science
International Society for Optical Engineering
Procedia Computer Science
International Symposium on Visual Information Communication
Self-Organizing Maps, Lecture Notes in Computer Science
IEEE Symposium on Computational Intelligence for Financial Engineering and Economics
UKVAC Workshop on Visual Analytics
Tsinghua Science and Technology
Expert Systems with Applications
iConference

Table 3: *Conferences or journals in "Others" in Figure 1. We found only one paper from each conference or journal.*

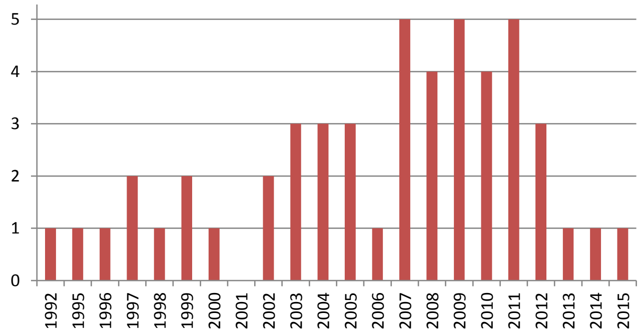

Figure 2: *Distribution of papers by publication years. Many papers were published between 2007 and 2011.*

tools, as well as the IEEE Xplore and ACM digital libraries with keywords combinations of "visualization", "financial", "economy", "exploration", and "analysis". We also individually explored prestigious venues for publications as shown in Table 2 and Figure 1 for related work in this domain. Here, we note that INFOVIS papers were published in IEEE Transactions on Computer Graphics (TVCG). Table 3 presents the papers in the "Others" category in Figure 1 and only one paper is found in each conference. In the end, we included 50 papers in total for our review.

While we were reviewing papers, we found that there was little information available about task requirements that were derived directly from financial analysts. We found one paper presenting types of action in tasks for a user study [MSK*06]. We believed such sparse information could limit researchers from acquiring detailed task requirements for their research. Therefore, we interviewed analysts from different financial sectors in order to systematically gather the general domain requirements for visual analytics researchers. The discussions, interviews, and task elicitation processes, and task requirements as presented in Section 4 contributed

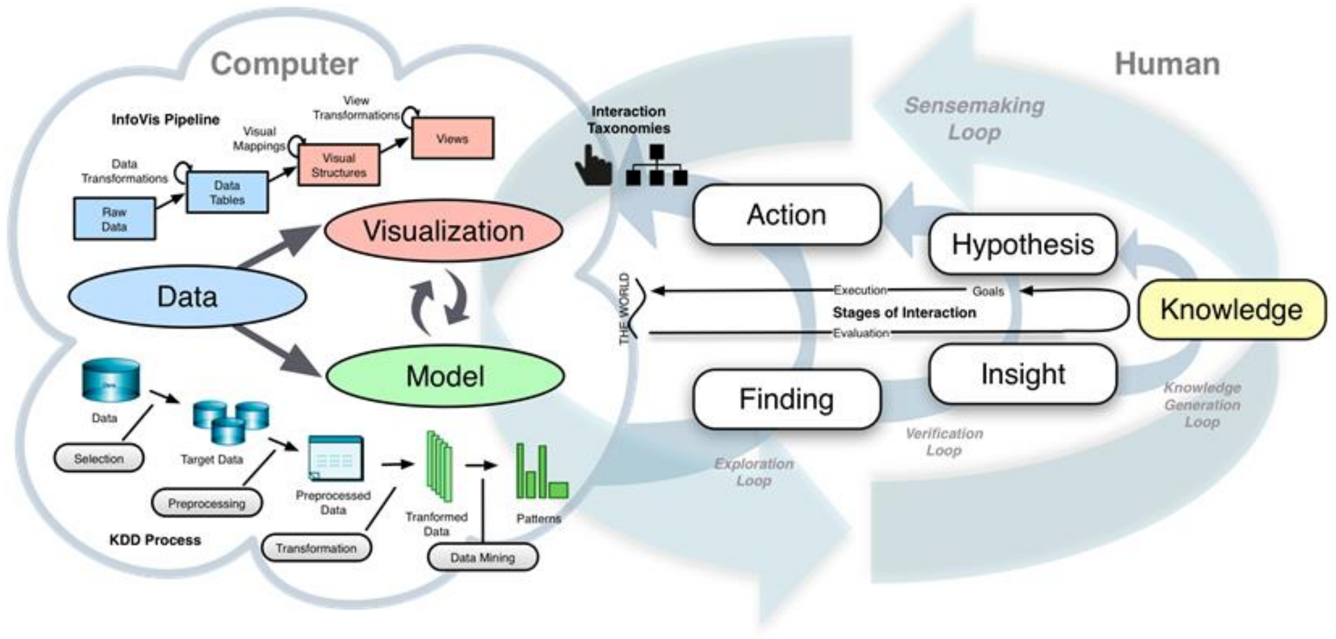

Figure 3: *Knowledge generation model for Visual Analytics [SSS*14]. As a feature noted from the interviews with analysts is that the model explicitly demonstrates two separate pipelines (InfoVis and KDD).*

to deriving a visual analysis pipeline in the financial domain. Next we discuss the financial visual analysis pipeline and knowledge generation model.

During reviews and discussions, we considered visual data exploration and analytics models [TC05, KKEM10] and the information foraging loop [PC05] to find out whether the models can be better used for describing the analysis process in the financial domain. At first, the analysis processes in the domain seemed well matched to the model proposed by Keim [KKEM10]. But, then we found that the preprocessing step (transformation) was underestimated in the model. This preprocessing is frequently required for certain tasks in the domain such as pattern detection or topic modeling (e.g., time-series stock data linked to news stories) to ingest the data into automated modules, followed by data visualization. This was also confirmed with the two analysts from a major nation-wide company (Section 4) who said they have extensively used and applied data mining techniques but did not expose detailed information due to propriety process information. Therefore, we think the knowledge generation model proposed by Sacha et al. [SSS*14] as shown in Figure 3 is more suitable to describe the visual analysis process for a financial data analysis process where preprocessing and data transformation steps are explicitly shown with the "KDD process." Another note is that sometimes analysts were dynamically teamed-up for different analytics goals and purposes. For example, in a team, one analyst focused on visual analysis with applying automated algorithms while another analyst concentrated on generating reports comparing the results generated from the visual analysis by another analyst (e.g., comparing anomalies to infer unknown facts in a report). This is interesting that the collaboration in the team is also represented in the knowledge generation model where the results from "Computer-side" and "Human" are iteratively communicated and discussed iteratively.

In this paper, summary information is provided with tables and figures. We provide a summary distribution of papers over their corresponding publication years in Figure 2. We summarize the results of our survey on financial data analysis systems in Tables 4-8. Here, we group papers by year and sort them alphabetically by first author name, placing each paper in the appropriate category. If at least one industry expert coauthored a paper (judged by affiliation), we color the paper author column dark blue to indicate collaborations between researchers and industry experts. We also provide the sum of the papers that have been published in collaboration with industry experts in the "From industry" column. In the next sections, we describe the task and system requirements and resulting recommendations derived from the interviews with analysts in the financial data analysis domain.

4. Survey of Task and System Requirements for Financial Data Analysis and Resulting Recommendations

Besides the categorization of financial visual analytics systems data, techniques, and evaluation, it is important to also examine the requirements of different groups of the financial analysts who will use these systems. We, therefore, interviewed domain experts from the financial risk management, economic analysis, and capital market management areas to obtain task requirements. Although the requirements summarized in this section were acquired from a subset of experts in the financial industry, we believe that they are representative.

We had five group discussion sessions where 2 to 9 analysts

	1992	1995	1996	1997		1998	1999		2000	2002		2003			2004		
	Jungmeister92 [JT92]	Strausfeld95 [Str95]	Ankerst96 [AKpK96]	Gross97 [GSF97]	Brodbeck97 [BCLC97]	Kirkland98 [KSH*98]	Gresh99 [GRTM99]	Wattenberg99 [Wat99]	Groenen00 [GF00]	Eklund02 [EBVV02]	Xiong02 [XPH02]	Csallner03 [CHLS03]	Simunic03 [Sim03]	Taskaya03 [TA03]	Dwyer04 [DG04]	Nesbitt04 [NB04]	Shen04 [SE04]
Stock	■		■					■	■				■	■	■		■
Fund		■									■	■					
Econ. Indi.				■													
Trans.					■	■	■									■	
Risk																	
Company Info										■							

	2005			2006	2007					2008				2009			
	Hao05 [HDKS05]	Lin05 [LCZ05]	Smeulders05 [SH05]	Schreck06 [SKM06]	Chang07 [CGK*07]	Mirel07 [MGBC07]	Schreck07 [STKF07]	Tekusova07 [TK07]	Ziegler07 [ZNK07]	Chiu08 [CHC08]	Dao08 [DBBS08]	Tekusova08 [TS08]	Ziegler08 [ZNK08]	Alsakran09 [AZZ09]	Huang09 [HLN09]	Rudolph09 [RSE09]	Sawant09 [Saw09]
Stock	■			■			■	■			■				■		■
Fund									■				■	■			
Econ. Indi.						■						■					
Trans.		■			■												
Risk																■	
Company Info	■		■							■							

		2010				2011					2012			2013		2014	2016		
	Schreck09 [SBVLK09]	Alsakran10 [AZZ10]	Lei10 [LZ10]	Wu10 [WP10]	Ziegler10 [ZJGK10]	Lei11 [LZ11]	Sarlin11a [Sar11]	Sarlin11b [SE11]	Savikhin11 [SLFE11]	Schaefer11 [SWK*11]	Brath12 [BM12]	Ko12 [KMJE12]	Wang12[WJCR12]	Sorenson13 [SB13]	Chen13[CRVC13]	Lemieux14 [LSL*14]	Wanner16 [WJS*16]	Sum	From Industry
Stock	■		■	■	■	■				■	■			■			■	24	6
Fund		■																7	2
Econ. Indi.							■											4	2
Trans.								■										7	5
Risk									■				■		■	■		5	0
Company Info												■						5	1

Table 4: *Categories for data sources. We derive 6 data sources. It is obvious that stock and fund data are the most popular data in both research communities and financial industries.*

participated. The analysts in the sessions had diverse backgrounds from economics and social science to quantitative analytics and statistics. Their current roles also varied including many who were interested in predictive analytics. Among the five sessions, the first two sessions were to discuss the work flow (task-level knowledge elicitation) and the other three sessions were to obtain general system requirements.

The first focus group was responsible for detecting anomalies from financial datasets (e.g., unusual variations in stock prices) and generating reports with evidence for further action. Analyzing such anomalies is itself a complex and time-consuming task. It requires analysts to look into events that have contextual evidence available through news data sources. It further requires analysts to understand historical patterns and causes of events and to distinguish whether events have regular or periodic properties before deciding if the events are considered to be anomalous. There is also a need to show statistical significance of the detected anomaly for guiding decision-making tasks. The second focus group consisted of analysts responsible for generating reports on future financial trends. While the initial procedure of detecting anomalies in trends is similar to that of the first, this group focuses on making a large number of comparisons among anomalies in the market data attributes to infer unknown facts.

Next, we interviewed three analysts from a large U.S. bank with backgrounds in computer science, economics, and mathematics who are tasked with finding patterns and trends from financial data and reporting results to decision makers. Their main duties include estimation of retail sales using customer transaction data combined with U.S. Census data. The analysts stated that due to the size and complexity of the recent datasets, various machine learning techniques are used, including k-means, principle component analysis, support vector machine, and linear and logistic regression to facilitate the analysis. In addition, several visual tools (e.g., Tableau, Spotfire) and statistical packages (e.g., SAS, R) are extensively used due to their effectiveness in exploring data. When asked which type of action was especially time-consuming, they stated that most of the difficulty during their tasks come from the need to spend time to format data to a digestible size, clean data, and automate their analysis results for presentation. A manager who leads a team of 6 analysts has additional tasks such as verifying data security, ensuring customer privacy, and removing unnecessary information encumbering the analytics processes.

Finally, we interviewed economists responsible for analyzing large scale financial trends (e.g., sector, nation-wide trends). The analysts' requirements were similar to those of the previous groups; however, these analysts stressed the importance of providing contextual evidence about anomalies that can help confirm certain hypotheses. These analysts were also interested in comparing the anomalies across different aggregation levels in support of anomaly explanations. Based on these discussions, we summarize detailed requirements for visual data analytics systems in the financial data exploration. In general, a visual analytics system for financial data should:

- R1 Provide ample information including historical and contextual data in order to help users find historical and recurring patterns, and causes of events,
- R2 Support users with automated techniques for detecting patterns, trends, and anomalies,
- R3 Allow interaction methods for drill-down and comparison analysis,
- R4 Present whether an anomaly (if detected automatically) is statistically significant and statistically significant ranges,
- R5 Forecast future trends based on user selected data ranges,
- R6 Provide functions for data cleaning, data shaping, and result customization for presentation,
- R7 Provide visualizations that avoid visual clutter and occlusion through data transformation (aggregation, clustering, filtering) and data cleaning (R6), but support details-on-demand (R3).

Here, by the contextual data in R1, we mean other auxiliary data that are often considered in conjunction with the main target data. For example, analysts stated in the interviews that they often combine news stories in analyzing stock data to gather contextual evidence to support or reject their hypotheses. In R3, various interaction methods exist based on different operations [Shn96] and tasks [AES05]. The analysts were especially interested in adjusting parameters of the applied data mining techniques (e.g., thresholds for anomaly judgment) and verifying the results in visualizations on-the-fly. In R7, when visualization techniques alone cannot resolve visual clutter and occlusion issues (e.g., visualizing hundreds of stocks), other approaches need to be considered including data cleaning, filtering, data transformation, aggregation, and clustering. But even with these approaches, there needs to be methods to support details-on-demand [Shn96] for certain tasks (e.g., comparison of sales or transactions). It should be noted that Kohavi et al. [KRS02] also discuss requirements in terms of time and performance constraints in business analytics systems (e.g., time for analysis cycle, integration of data, distribution of results). However, compared to Kohavi's work, the requirements discussed in this section provide task requirements and functionalities of systems at both an overview and detail level. In the next sections, we present our review on different aspects including data sources, visualization techniques, interaction methods, and evaluation.

5. Financial Data Sources

In this section, we summarize and categorize data sources used in financial analysis systems as follows: *stocks, funds, economic indicators, transactions, risk, and company information* as shown in Table 4. Note that we derived the six categories after analyzing the collected papers. Note that while stock and fund data are relevant to stocks, the others have relatively broader coverage. For example, the company information category covers all the information related to a business (e.g., profit, financial statements, sales, marketing data) and the transaction category considers all the data generated by transactions among different subjects (e.g., bank customers, companies, and countries). Visualizations for these data sources typically cover the domains of financial risk management, economic analysis, capital markets management, portfolio management, and marketing analysis (Table 1). Stock data containing time series of share prices of companies over a long time have been extensively used in financial data analyses, including trend, pattern, performance, and predictive analysis. It is notable that the number of stocks used in analyses ranges from a few to hundreds [Sim03, DG04, STKF07, DBBS08, SBVLK09, BM12, WJS*16] to several thousand stocks (5328 stocks in a visualization [AKpK96]). The stock data is often combined with news media data to extract additional knowledge for providing contextual information [WP10, TA03, WJS*16].

Analyzing groups of stocks is also popular (i.e., sector analysis and fund analysis). In sector analysis, stocks are usually aggregated based on industry-defined sectors (e.g., energy, healthcare, information technology) to be used in the analysis, while fund analysis uses groups of stocks gathered by investment consideration (e.g., mutual funds). Note that the data of groups of stocks compiled for industry sectors or funds possess hierarchies that are base data for hierarchy visualization techniques such as treemaps [Wat99].

Economic indicators are statistical information that affect financial markets as other factors (e.g., political news and corporate earning news, countries' or world's financial information from public data providers [oec, wor]) do. A basic approach in economic indicator analysis is to rely on a single indicator in the market such as consumer price indexes [GSF97] or inflation [TS08]. Compared to this approach, Sarlin [Sar11] uses 35 general economic variables for monitoring debt ratios, which is the largest number of variables used for indicator visualization.

Transactions are generated in many financial domains including bank, investment, and money exchange industries. Transaction analysis is important because it could reveal unknown transactional patterns, malfunctions in a business, and evidence or symptoms of security breaches. Many types of the transaction data have been analyzed with visualizations such as the transactions among customers (e.g., wire transfer [CGK*07], fixed income trading [BCLC97]), businesses (e.g., bids and asks [KSH*98, NB04] or option trading [GRTM99]), and countries (e.g., currency exchange rates among 23 countries [LCZ05, SE11]).

Risk analysis is an important part of business and investment, but the analysis of risk with visualizations is underutilized. Among the five risk analysis projects that we found, two projects [RSE09, SLFE11] show risk data generated from an economic model in user studies in order to find whether investors are able to recognize the risk in their investment and change their investment strategies according to the risk. The other three projects utilize credit data [WJCR12, LSL*14] and financial statements [CRVC13] for computing and presenting risks. Lemieux et al. [LSL*14] design a visual analytics system that visualizes multiple types of data including credit default swaps to help managers in banks understand systemic risk in the domain. In the system proposed by Wang et al. [WJCR12], consumer credit data generated from loans and credit products are used for assessing risk in banks' products. Risk of bankruptcy is directly computed and visualized from financial statements in the system proposed by Chen et al. [CRVC13].

Often, accessibility to data affects the focus of research. More than half of the papers use stock and fund data. This is expected due to easy access to financial data through on-line financial web sites. Therefore, stock data have been most utilized from early stages of financial visualizations. Compared to stock data, not much work has been done for visualizing economic indicators, although they are also accessible on-line.

It is notable that the visual analytics systems for analyzing transaction data were designed mainly when an author from an industry participated in the project. This is expected, because such data tend to be confidential and would normally not be accessible. However, we rarely find work performed with corporations for analyzing risks and company information. This came as a surprise initially because it is known that performance and risk evaluation of companies are very important areas for competitive intelligence [Kah98]. Additionally, marketing research provides insights on customers' fast-changing trends, demand, and buying patterns. We found only one paper [SH05] utilizing visualizations of customer data for marketing research. On the other hand, this can be understood since corporations tend not to expose projects they are interested in due to competition in the market and privacy issues existing in their data.

6. Applied Automated Techniques

There have been a number of automated techniques utilized for visual analysis of financial data as shown in Table 5. These tech-

	1997	1998	2000	2003	2004	2005	2007	2008	2009		2010	2010	2011			2012	2013	2016		
	Brodbeck97 [BCLC97]	Kirkland98 [KSH*98]	Groenen00 [GF00]	Simunic03 [Sim03]	Dwyer04 [DG04]	Lin05 [LCZ05]	Schreck07 [STKF07]	Ziegler08 [ZNK08]	Rudolph09 [RSE09]	Schreck09 [SBVLK09]	Lei10 [LZ10]	Ziegler10 [ZJGK10]	Savikhin11[SLFE11]	Sarlin11a [Sar11]	Sarlin11b [SE11]	Ko12 [KMJE12]	Chen13[CRVC13]	Wanner16 [WJS*16]	Sum	From Industry
K-means											■	■			■		■		4	1
SOM				■			■			■				■	■		■	■	7	3
MDS	■		■																2	1
PCA					■														1	0
Sampling						■											■		2	0
Moving Avg.																■		■	2	0
Regression																■			1	0
Decision Tree		■																	1	1
Specific Models								■	■				■						3	1

Table 5: *Categories for automated techniques. We categorize automated techniques used for financial data into 4 categories– clustering, dimensional reduction, statistical, and model- or theory-based techniques. For financial analysis, SOM (self-organizing maps) was used for both clustering and dimension-projection.*

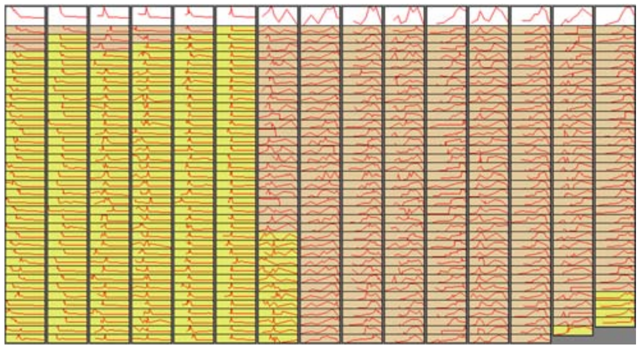

Figure 4: *Trajectory-based clusters are compared for analysis. Original data include 550 assets from sectors of software (yellow) and banks (brown) from 05/1994 to 06/2010 [ZJGK10].*

niques include clustering, dimensional reduction, trend and pattern analysis, and forecasting.

Clustering is a process of dividing data into groups of similar objects [Ber06]. Even though this may cause the loss of detailed information, clustering can effectively reduce the size of data for visualizations and also highlight other aspects of data. Efficacy to produce desired results of a clustering algorithm depends on the application. This leads to development of a variety of methods with different criteria. For financial data visualizations, k-means [Mac67] and Self-Organizing Map (SOM) [Koh90] are mainly utilized. In the k-means clustering, each of the clusters is presented by a mean (or weighted average) of data points in each cluster. Then, the discrepancies between the data points and the centroid are used to form clusters. With numerical data, this approach has an advantage in providing a good statistical and geometric sense in the results. For example, Ziegler et al. [ZJGK10] utilize the k-means algorithm for generating clusters with 550 assets due to its high speed computation, easy implementation, and ability to specify a desired amount of clusters as shown in Figure 4. In order to produce clusters for analyzing stock data based on trading patterns, Lei and Zhang [LZ10] utilize k-core, a variant of k-means, that allows pruning nodes and links whose degrees are less than k.

SOM [Koh90] is a neural network based on unsupervised learning technique that has gained popularity. As a clustering algorithm, the SOM produces topological clusters where similar clusters are neighboring. On the other hand, SOM projects multidimensional data into a two-dimensional grid as a projection technique. Therefore, SOM generates both clusters and topology-preserved mapping of prototype vectors on a grid structure in low dimensional space. Many visualization systems incorporating SOM tend to use the generated mapping data when the clusters are placed. For example, the approach shown in [Sim03, STKF07, STKF07, WJS*16] presents how SOM can be used for clustering with stock data and financial statement data [CRVC13], while the approach in [SE11] demonstrates how the mapping information can be utilized for placing clusters generated from multivariate economic indicators. There is a complementary work that provides a collection of SOM-based applications [DK98].

Sometimes the multidimensional analysis indicators can be better revealed if the multidimensional data can be placed in a lower dimensional space. To this end, two dimensional reduction techniques have mainly been used for financial data exploration: Multi Dimensional Scaling (MDS) [BG97] and Principle Component Analysis (PCA) [ED91]. Even though the notion of projecting multidimensional data onto a lower dimensional space is similar, there are differences between these two algorithms. MDS aims at presenting a proximity matrix based on Euclidean distances for a 2D or 3D space. On the other hand, PCA seeks to find the axes of greatest variance (i.e., principal component) and to capture the variance by projecting the data onto a plane. In addition, PCA can be more practical when scalability is an issue with large data [DE02]. When these projection techniques are used for transaction [BCLC97] and stock markets data [GF00, DG04], scatterplot-based visualizations are often utilized because the results of the algorithms need to be projected onto 2D [BCLC97, GF00] or 3D [DG04] grids.

There are techniques that are rarely used in financial research work including sampling, moving average, regression, decision tree, and risk/return models borrowed from economics. Sampling is a statistical process where a predetermined number of samples are taken from a large data collection. How many samples are taken from data depends on the type of analysis. The work by Lin et al. [LCZ05] is the only work using the sampling technique. In their work, the sampling technique is utilized in order to reduce data size in a stream of financial data. Applying the moving average technique basically helps reduce the impact of noise in the data and reveal interesting points [WJS*16]. More importantly for financial analysis, the moving average technique also can be used for trend analysis combined with regression techniques. The system proposed by Ko et al. [KMJE12] incorporates these two techniques for predicting future sales data for analyzing competitive advantages of corporates.

	Jungmeister92 [JT92]	Strausfeld95 [Str95]	Ankerst96 [AKpK96]	Gross97 [GSF97]	Brodbeck97 [BCLC97]	Kirkland98 [KSH*98]	Gresh99 [GRTM99]	Wattenberg99 [Wat99]	Groenen00 [GF00]	Eklund02 [EBVV02]	Xiong02 [XPH02]	Csallner03 [CHLS03]	Simunic03 [Sim03]	Taskaya03 [TA03]	Dwyer04 [DG04]	Nesbitt04 [NB04]	Shen04 [SE04]
Year	1992	1995	1996	1997		1998	1999		2000	2002		2003			2004		
Std (2D)											■		■	■			
Std (3D)		■		■		■	■				■				■	■	
GT					■				■	■							
Iconic																	■
Dense Pixel			■														
Stacked	■							■				■					

	Hao05 [HDKS05]	Lin05 [LCZ05]	Smeulders05 [SH05]	Schreck06 [SKM06]	Chang07 [CGK*07]	Mirel07 [MGBC07]	Schreck07 [STKF07]	Tekusova07 [TK07]	Ziegler07 [ZNK07]	Chiu08 [CHC08]	Dao08 [DBBS08]	Tekusova08 [TS08]	Ziegler08 [ZNK08]	Alsakran09 [AZZ09]	Huang09 [HLN09]	Rudolph09 [RSE09]	Sawant09 [Saw09]
Year	2005			2006	2007					2008				2009			
Std (2D)		■		■	■	■	■				■					■	
Std (3D)			■									■			■		
GT								■									■
Iconic										■							■
Dense Pixel									■				■	■			
Stacked	■														■		

	Schreck09 [SBVLK09]	Alsakran10 [AZZ10]	Lei10 [LZ10]	Wu10 [WP10]	Ziegler10 [ZJGK10]	Lei11 [LZ11]	Sarlin11a [Sar11]	Sarlin11b [SE11]	Savikhin11 [SLFE11]	Schaefer11 [SWK*11]	Brath12 [BM12]	Ko12 [KMJE12]	Wang12[WJCR12]	Sorenson13 [SB13]	Chen13[CRVC13]	Lemieux14 [LSL*14]	Wanner16 [WJS*16]	Sum	From Industry
Year	2009	2010				2011					2012			2013		2014	2016		
Std (2D)	■								■	■		■	■				■	16	3
Std (3D)											■							11	4
GT	■	■		■			■	■					■		■			12	6
Iconic			■			■								■				6	2
Dense Pixel			■		■					■		■	■				■	10	1
Stacked											■					■		7	3

Table 6: *Categories for visualization techniques based on the taxonomy in [Kei02]. As we expected standard charts have been a popular technique in analysis. Among advanced visualizations, dense and geometrically-transformed displays were used more than iconic and stacked.*

A decision tree is a technique utilizing a tree-like shape where each branch in the tree represents a consequence of a decision or event. Although a decision tree can be generated and used in machine learning, here we limited a decision tree as a visual representation for analyzing branches of selections. The decision tree technique is useful when a complex decision-making process proceeds with a series of selections among possible answers based on probability. The ADS (Advanced Detection System) [KSH*98] is a good example that shows the effectiveness of the decision-tree. The system has been used since 1997 to monitor regulatory breaks in stock trading, and it has detected over 7000 breaks. There are two research projects that utilize models originated from economics [RSE09, SLFE11] in order to visualize results of an investor's decisions. Both systems are used for user studies to find if investors can make better decisions with the support of a visual analytics system that visualizes the results of their investments.

Opportunities: Anomaly Detection and Modeling: We found two patterns in applying automated techniques to financial data. Often, some techniques have been used for making data more understandable by directly reducing data size (e.g., sampling) or grouping data into subsets of data (i.e., clustering and dimensional reduction). In many use case scenarios, it is observed that users pursue their exploration with interaction methods (e.g., Abstract and Filter in Section 8) after these techniques are applied. We also confirm that the number of utilized automated techniques is low when we consider other available dimensional reduction [Fod02] and clustering [Lia05, Ber02] algorithms in the data-mining research area. In addition, some trend analysis techniques (e.g., moving averages and regression) are not extensively used in financial visualization research. This is surprising because domain experts whom we have interviewed mentioned that they have extensively used automated statistical techniques in their work as described in Section 4. They are aware that accurate prediction of values and trends, and verification of changes in data or human activity could be very helpful in designing future strategies and could produce high returns with low risks in investments. Therefore, we suspect that when many businesses use a variety of techniques for a type of task (e.g., finding an anomaly), they try to avoid revealing their techniques in publications.

7. Visualization Techniques

As discussed in Section 5, our survey reveals that financial data mainly tend to be a set of time-series data with various types of attributes [AMST11]. In addition, we find that various visualization techniques have been applied to financial data, including parallel coordinate plots [Ins85], treemaps [Wat99], scatterplot matrices [EDF08], glyph based visualization techniques [Che73], stacked displays [HHN00], and dense displays [Kei00, WGK10]. In this section, we utilize Keim's [Kei02] classification to categorize the different approaches utilized for financial data: 2D/3D displays, geometrically transformed displays, iconic displays, dense pixel displays, and stacked displays (Table 6). When multiple visualization techniques are used in a paper, we categorize the paper based on the primary visualization technique for analysis. For each category, we add sub-categories if necessary in order to present differences within the same-category techniques.

7.1. Standard visualization techniques

The most common technique for visualizing financial data remains the family of standard charts (e.g., line, bar, pie charts, and their variations [Mur99]). These standard charts utilize 2D data whose basic idea is to present attribute values along the y-axis and time on x-axis. 3D visualization techniques can facilitate the analysis of data by adding an essential dimension to another axis. Although these charts help users in finding trends and key events (e.g., spikes), they suffer from issues of over-plotting, occlusion, and insufficient support for comparing multiple datasets [SB13]. In addition, these are not appropriate for analyzing higher dimensional data. In subsequent sections, we discuss these visualization techniques.

7.1.1. 2D Techniques

2D charts have been extensively used in the analysis of financial datasets. The basic use of these charts is an auxiliary visualization in multiple coordinated views (e.g., line graph view for stock prices, bar charts for credit analysis) [TA03, CGK*07, SWK*11, WJCR12, WJS*16]. Line graphs are also used in financial investment plan experiments [RSE09]. In the experiments, risk information is presented with line graphs to evaluate if investors are aided to understand investment risks with the line graphs. In order to alleviate the overplotting issue, Lin et al. [LCZ05] adapt the fish-eye technique [Fur86] in exploring currency exchange data streams divided by areas of interest.

Visualizations other than line graphs are also utilized. Dao et al. [DBBS08] utilize wedge charts as an extension of pie charts. In the wedge charts, each wedge is assigned to a stock, and the size and color of a wedge are mapped to velocity and force indicators in the market, respectively. 2D maps have been extensively used in information visualization, but only a few systems exist in the financial data analysis domain that project financial data onto geographical maps. Examples of such systems and visualizations include the visualizations for comparing regional sales competition [KMJE12], regional fund distributions of market values [XPH02], and changes in commodities markets [MGBC07].

Small multiples are a visualization technique where a series of basic graphics or charts on the same scale are aligned to allow easy comparisons [Tuf90]. There are two factors that need to be determined for using this technique: representation (e.g., line graphs) and an arrangement method (e.g., grid or radial). A basic approach utilized for analyzing stock data is to map data to representative charts (e.g., line chart) by using various layout algorithms (e.g., distance based mapping [Sim03], trajectory bundling [STKF07]). The charts are then distributed and organized in a topological order. There are projects ([Sim03, SBVLK09, STKF07]) that use a line chart as a basic representation combined with SOM for the chart arrangement for stock data analysis.

In summary, 2D displays are used for presenting time-series information in financial data, mainly with stock data. But the 2D displays in general are not equipped with other techniques for overcoming the overplotting issue except in the work by Lin et al. [LCZ05] that utilizes the fish-eye technique. We think that the overplotting issue partially leads to the adaptation of other visualizations techniques for financial analysis, leaving line graphs in auxiliary views.

7.1.2. 3D Techniques

Adding another dimension extends 2D visualizations to 3D visualizations. A great deal of research has demonstrated that using 3D visualizations could enhance understanding of data compared to 2D visualizations, and various 3D approaches have been adapted for financial data analysis. A standard way is to reuse known 2D visualizations. Bar (or column) charts have been popular for the extension of the standard charts (e.g., analysis for stock [Str95], trading transactions [KSH*98], funds [DG04], and customer data [SH05]). The basic approach used in this type of reuse is shown in Dwyer and Eades [DE02], whose visualization displays a navigable three-dimensional map with different sizes of columns presenting prices

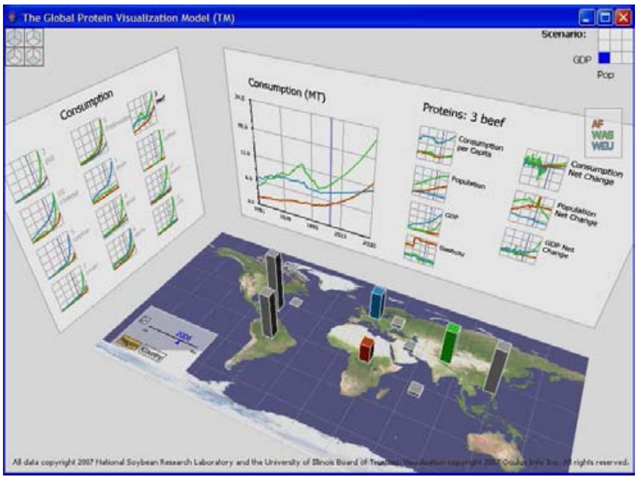

Figure 5: *An example of analysis of a commodity market in 3D. This workspace combines 2D line charts and 3D bar (or column) charts in a workspace [MGBC07].*

and volumes. Extending line graphs to 3D generates a surface chart [GRTM99, NB04] for analyzing stock prices. But the surfaces still may produce occlusion and cause difficulty in user understanding. For example, line graphs for presenting price-time information are extended in [NB04], but a large surface positioned at the front could hide other surfaces behind the front surface in the example. Mirel et al. [MGBC07] use a different approach in a way that they attach 2D visualizations to walls to form a 3D visualization space as shown in Figure 5.

Advanced visualizations have also been extended for 3D visualizations. Examples include 3D scatter plots, splat visualizations for analyzing fund composition [XPH02], and animating time-added 3D wedge charts for examining changes in inflation data through time [TS08]. Using spheres as a basis for organizing visualizations is a popular approach, because they support intuitive navigation and utilization of gestalt psychology [Kof13], and produce aesthetically appealing visualizations (e.g., [BM12, GSF97]). For example, in Brath and MacMurchy's work [BM12], SphereCorr spreads a dense graph for presenting stock data on a sphere based on weights of links, while SphereTree presents a hierarchy of stocks, consumer price index, and occupation and incomes with treemaps on a sphere. However, they also reported issues of depth perception and limited navigation methods.

We observe that the approach of using and extending the standard charts is still limited to presenting time-series data such as stock market data. Even though the approach has strength in certain tasks (e.g., rank and association [MSK*06]), it lacks a method to overcome visual clutter as the size of financial information becomes large. One effective approach is to use multiple coordinate views where the information with the standard charts can complement the information from other views. In our review, we find that only 14 visualizations incorporate multiple views to support financial data analysis.

7.2. Geometrically-transformed displays

Geometrically-transformed displays help users find interesting transformations of multidimensional datasets [Kei02]. In this category, scatterplot matrices, parallel coordinates, and cluster visualizations are included. It is notable that many visualization techniques in this category are often combined with automated algorithms. Conventionally, a scatterplot presents data as points in a 2D Cartesian space. A popular approach in using the scatterplot is to change color, size, shape, and orientation of the points with stock data [TK07, Saw09, WP10]. For example, stocks are placed based on monthly (x-axis) and yearly (y-axis) trading records, and each stock is assigned a color to present a degree of the price as shown in Figure 6. Often, the scatterplot is combined with automated algorithms. Two research projects [BCLC97, GF00] use Multi Dimensional Scaling in order to analyze trading data [BCLC97] and correlations of stock markets [GF00]. K-core (a family of k-means) is also utilized with a scatterplot to cluster stocks based on stock price variation in order to identify barriers and resilience in stock trading [LZ10]. Note that stock data are mainly utilized with scatterplots due to the fact that scatterplots assign properties (e.g., color, size) to individual elements and stock information is usually analyzed at the individual stock level.

On the other hand, parallel coordinate plots have not gained popularity for financial data analysis. The work of Alsakran et al. [AZZ10] is the only visualization utilizing the tile-based parallel coordinate plot for mutual fund data analysis. The system performs three processes in order to present a density of stock prices. The system first divides a visualization space by rectangular tiles. Then, the number of polylines (presenting stock prices) that passes each tile is used for computing density on each tile. By assigning different colors and opacity levels on the tiles according to the computed density, the visualization presents density of stock prices in a form of parallel coordinate plot. Note that parallel coordinates are rarely utilized for financial visualizations. This implies that multidimensional financial data have not been highlighted much yet, even though there exist several multidimensional financial data sets such as marketing data (e.g., customer data).

As discussed in Section 6, Self-Organizing Maps (SOM) generate both 2D projected data and clusters. We place cluster visualizations based on SOM in this section because SOM affects the geometry of the data and generates a new topology. There are systems using generated topological clusters that are laid out in a hexagonal grid in order to analyze companies' financial performances (e.g., operating margin) [EBVV02], debt indicators [Sar11], bankruptcy [CRVC13] and a currency crisis model [SE11]. SOM is one of the unsupervised algorithms. This implies that SOM may produce uninteresting results due to the lack of directions in learning. Relevant to this issue, Schreck et al. [SBVLK09] present an interactive visualization framework enabling combination of (unsupervised) automated algorithms and human supervision for clustering and visualizing trajectories of stock data.

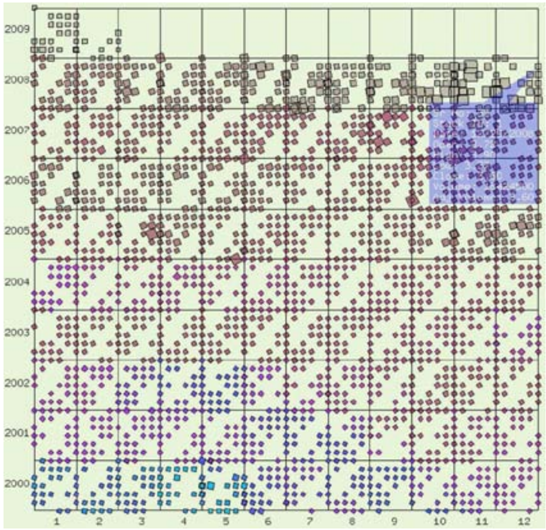

Figure 6: *Scatterplot example for Geometrically-transformed displays. Stock trading information (x-axis: month, y-axis: year) is presented by different colors, orientations, and sizes [Saw09].*

7.3. Iconic displays

Icon and glyph-based displays assign a dimension to a feature (e.g., shape, color) of the icon. Chiu et al.'s [CHC08] star-glyph follows this idea by placing each evaluation factor (e.g., size, return of assets) for each dimension. Their visualization allows efficient comparison of strengths and weaknesses among firms. Sawant [Saw09] also uses this idea in the visualization technique called "Psquares" (perceptual squares) when exploring stock price data where attributes of data are mapped to squares that vary in color, size, orientation, and transparency. These are then presented in a spiral layout or a scatter plot and help understanding stock trading patterns based on different time ranges (monthly or yearly). However, the large number of encodings mapped for the multiple attributes can make it difficult for analysts to discern patterns inherent in the data.

Lei and Zhang [LZ11] use the notion of visual signatures to map financial time series datasets. They create a galaxy spiral visualization to present the market index structure and another spiral visualization technique for displaying single stock data that allows investors to group stocks with similar short-term trading activities. Sorenson et al. [SB13] also utilize a visual mapping technique to describe discrete event data (e.g., news, earnings, announcements) to alphanumeric pictographics on line graphs, as opposed to using conventional abstract glyphs. However, there are usability issues in these mapping-based visualizations. The visual signature approach [LZ11] gains lower scores than RingMap [LZ10] in terms of usefulness, ease of use, and ease of learning. In the pictographics approach [SB13], there is a fixed number for mapping (36 glyphs) and occlusion issues. Shen and Eades [SE04] provide an ambient information visualization technique that maps financial data to an actual 3D tree where the trunk of the tree represents trade volumes and leaves show price. However, the focus of their work is to show the general changes in financial data in a visually-pleasing representation rather than to support actual analysis.

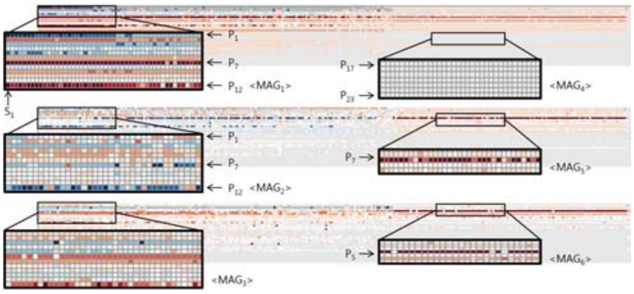

Figure 7: *An example of dense displays allowing analysis of sales, trend, and growth rates in 288 stores simultaneously [KMJE12]. Dense displays have an advantage in presenting large data.*

7.4. Dense pixel displays

Dense displays (or pixel-oriented/pixel-based displays) map an attribute value to a color-coded pixel in order to best utilize the screen space. The fundamental difference among the types of dense displays is the granularity that usually determines the layout of pixels. The standard layout strategy in this category is a grid layout as shown in the work by Ziegler et al. [ZNK07, ZNK08]. In the work, they compute performance and growth of different financial assets (bond and fund) for all possible time intervals and present them in a single dense view. Their dense matrix visualizations enable users to compare more than 12,000 assets simultaneously. In a similar manner, Schaefer et al. [SWK*11]'s visualization allows the analysis of 222 stocks over 5 years in the oil sector. Lastly, the Pixel Bar Charts technique [KHL*01] transforms a line graph into a bar graph densely filled with color-coded vertical lines. A series of these charts are utilized to analyze market stability and volatility [ZJGK10].

Dense displays have been well adapted for various sizes of pixels and layouts. Examples of systems adopting larger or varying-size of pixels include MarketAnalyzer [KMJE12], and Alsakran et al.'s [AZZ09] system. MarketAnalyer, as shown in Figure 7, maps different colors to the same size pixels for allowing competitive advantage analysis among 288 stores based on sales, trends in the sales, and growth data. Compared to MarketAnalyzer and Alsakran et al.'s [AZZ09] system generates larger pixels to help analysts detect suspicious activities in wire transfer transaction data [CGK*07] and find patterns in mutual fund based on various factors such as return against fund asset size (or cash holding, expense ratio). When the size of pixels becomes larger, the pixel-oriented visualizations tend to be more similar to heatmap visualizations [AvH04, WJCR12].

Dense displays are versatile because they are easily utilized in a number of layouts. One popular layout, other than grid for the pixel-oriented display, is the radial layout that is advantageous in finding cyclic patterns. "Circle Segments" [AKpK96] is the first work where dimensions of stock data are mapped to segments of a circle. This approach is sufficiently scalable compared to the line graphs for a large stock price data in that it allows analysis of 265,000 data items in 50 dimensions. In a similar way, Ringmaps [LZ10] are constructed with a varying size of multiple rings, each of which present multiple stocks within a close data range.

Figure 8: *Stacked displays have been popular in visualizing stock and fund data. Here, stocks included by multiple funds are presented by rectangles [CHLS03].*

7.5. Stacked displays

Stacked displays present data in a hierarchical manner. Treemaps [JS91] are a popular technique in this category whose goal is to visualize the hierarchy of multidimensional data. In treemaps, a dimension (e.g., a market sector) is connected to a rectangle whose properties (e.g., size, color, label, position) are determined by the connected dimension. This approach has been shown to have advantages for categorization and rank [MSK*06], but also has limitations such as lack of capability for presenting evolution of market or asset prices over longer time periods and a perceptual difficulty (e.g., hard to notice small size data in treemaps).

The strength in presenting hierarchies (or sectors) generates high demand in stock data analysis for providing overview stock markets (e.g., [JT92, Wat99, LSL*14, CHLS03]), where the volume, capital size, or risk of each sector determines the size of each rectangle as shown in Figure 8. Treemaps can also be used for computing layouts for placing financial data. Hao et al. [HDKS05] demonstrate how to place visualization areas (rectangles) to present sales data with bar charts. For 3D extensions, two approaches are used. The first method directly extends the 2D treemaps for industry sector analysis into 3D treemaps for visual financial surveillance by adding stock prices as height information [HLN09]. Another approach is to project treemaps onto a 3D sphere as Brath and MacMurchy [BM12] demonstrate for stock market analyses. However, both 3D approaches have a usability issue: occlusion and difficulty in data navigation. In the approach using additional data mapped to height, a cube with low height tends to be hidden by large surrounding cubes, and in the 3D sphere approach, users are frequently

	1992	1997	2003			2004	2005			2007		2008			2009	2010			2011		2012			2013	2014	2016		
	Jungmeister92 [JT92]	Brodbeck97 [BCLC97]	Csallner03 [CHLS03]	Simunic03 [Sim03]	Taskakya03 [TA03]	Dwyer04 [DG04]	Hao05 [HDKS05]	Lin05 [LCZ05]	Smeulders05 [SH05]	Chang07 [CGK*07]	Schreck07 [STKF07]	Dao08 [DBBS08]	Tekusova08a [TS08]	Ziegler08 [ZNK08]	Schreck09 [SBVLK09]	Alsakran10 [AZZ10]	Lei10 [LZ10]	Ziegler10 [ZJGK10]	Lei11 [LZ11]	Schaefer11 [SWK*11]	Brath12 [BM12]	Ko12 [KMJE12]	Wang12 [WJCR12]	Sorenson13 [SB13]	Lemieux14 [LSL*14]	Wanner16 [WJS*16]	**Sum**	**From Industry**
Select		■											■		■		■				■	■	■	■		■	9	4
Explore		■																									1	1
Reconfigure						■			■				■								■	■					5	1
Encode																			■								1	0
Abstract	■	■		■		■	■	■		■			■			■	■	■		■		■					13	4
Filter		■	■		■		■				■	■		■				■		■		■	■		■	■	13	2
Connect										■												■	■	■			4	2

Table 7: *Categories for interaction methods. Abstract, Filter and Connect have been the most popular interaction methods for visual financial analysis. The popularity of Connect implies wide adoption of multiple coordinated views.*

asked to perform rotation of the sphere to verify assets in both foreground and background.

7.6. Analysis of Visualization Techniques

We find that researchers have traditionally used standard charts (2D and 3D) for financial visualizations as shown in Table 6. We also observe that geometrically-transformed displays and dense displays are more preferred than iconic and stacked displays. On the other hand, we see different usage patterns over the years. The standard charts in 3D and stacked visualizations were used in early financial visualization systems followed by the standard 2D charts. Then, from 2007, dense displays including pixel-oriented visualizations gained popularity, followed by geometrically-transformed displays from 2009. We believe that a possible reason for this trend is the appearance of big data that is hard to visualize using conventional techniques and could be best approached with consideration of data mining techniques (for reduction of data size), geometrical-transformation, and layout algorithms.

Geometrically-transformed displays and standard charts are more preferred by industry. We believe that the popularity of using transformation techniques implies that financial datasets tend to be large, multidimensional, and hierarchical. Other than the standard charts and geometrically-transformed displays, the other categories have a similar number of publications involving industry experts.

8. Interaction Methods

Various interaction methods have been developed in order to support the different data types and analysis methods to enable users to effectively explore their data. In order to categorize the various interaction methods, a number of surveys are presented with a focus on low level interaction operations [Shn96], tasks [AES05], and benefits and user intentions [YaKSJ07]. In this work, we utilize Yi et al.'s [YaKSJ07] seven categories of interaction for the categorization of the interaction methods used in financial visualization systems. Note that not all the papers we reviewed described interaction methods in their visualizations or systems. Therefore, we exclude the papers from this list if enough information on the interaction methods cannot be discerned. Next, we briefly present the interaction taxonomy we use and examples that we considered. These interaction methods consist of: *select*, *explore*, *reconfigure*, *encode*, *abstract/elaborate*, *filter*, and *connect* as shown in Table 7.

8.1. Applied Interaction Methods

The "Select" interaction method enables users to *mark data in order to keep track* even when representations are changed. In order to make items visually distinctive, users click on individual items or draw an area that includes items within the area. The basic result of this interaction is to highlight specific items (e.g., stocks) and provide more information for the selected item on a (pop-up) tooltip (e.g., [BCLC97,LZ10,BM12,KMJE12,SB13,WJCR12,WJS*16]). This step is sometimes also used as a pre-step for another process (e.g., changing data ranges and hierarchy levels [TS08], allowing users to edit the stock trajectory for training an automated algorithm [SBVLK09]).

The "Explore" interaction method allows users to *examine different subsets of the data*. The most common example of this interaction type is panning that refers to the movement of a camera (e.g., panning on maps). In financial data analysis, this interaction method is rarely used because most visualizations present data according to a given screen space. The exception is Brodbeck et al.'s work [BCLC97] where data items are scattered beyond a given screen space.

The "Reconfigure" interaction method allows users to *view data from different perspectives*. The most common implementation of reconfigure is rotation in 3D visualizations to avoid occlusion (e.g., [TS08, DG04, BM12]). In particular, Brath and MacMurchy [BM12] project stock data onto a sphere and allow users to rotate the sphere to view the data that are at the back of the sphere. Sorting and other rearrangements of visual items are another example in this category. Two examples that explicitly rearrange the items include sorting the customer data for developing marketing strategies [SH05], and reordering of stores and products based on sales performance [KMJE12] that allows quick comparisons among data items.

The "Encode" interaction corresponds to *fundamental changes in visual representations*. Examples include changes in color, size, font, orientation, and shape of the different objects (e.g., from a pie chart to a histogram). Lei and Zhang [LZ11] present the only system that explicitly supports this interaction. In their visualization, the size of dots is allowed to be changed by users in order to encode different sizes of capitalization of industries.

The "Abstract/Elaborate" interaction allows users to change the abstraction level in the visual representation. More specifically, with support of this interaction, users are able to get both an overview and details of data items as needed. Zooming is an example interaction method in this category, and we find various usages of zooming in data exploration. When the data presented on the screen are too crowded, users are allowed to zoom-in to focus more on individual data items. Visualizations that adapt dense-displays for large data usually provide this interaction (e.g., [SWK*11, AZZ10, BCLC97, LZ10] for stock market data analysis). In the same context, this interaction is provided as a means to overcome visual clutter generated in the standard charts (e.g., standard charts for stock market data [Sim03, LCZ05], consumer price index [TS08], and fund [DG04] analysis). Additionally, this interaction also enables users to analyze data in different hierarchies that are generated by different types of aggregation. For example, stock market data are often aggregated by various hierarchical industry sector levels (e.g., Jungmeister and Turo [JT92] and Hao et al. [HDKS05] who utilize treemap visualizations to visualize hierarchical data). When data are clustered, a hierarchy is formed that can be utilized for stock pattern recognition as shown in [ZJGK10]. Data aggregation by date, time, or industry sectors also form a hierarchy. WireVis [CGK*07] supports analysis of data under different temporal hierarchies.

The "Filter" interaction helps users to *analyze data with specific conditions*. We find two approaches for specifying the conditions. The first approach is to select individual objects or attributes of the objects. This approach is realized by filtering using check boxes, or clicking on an object or attribute. The object-based filtering method is mostly used in stock analysis where individual stocks need to be filtered (e.g., [ZJGK10, DBBS08, HDKS05, STKF07, CHLS03, WJS*16]), while attribute-based options are used when auxiliary data are used together (e.g., stock analysis with new stories [TA03, WJS*16], income trading analysis with customer data [BCLC97]). The second approach for filtering is based on range. In general, this approach is realized by providing slider bars for range selection. For example, for stock analysis, slider bars have been utilized for selecting performance range [SWK*11] and also to allow modification of relevance functions to emphasize specific regions in the x and y-axes [ZNK08].

The "Connect" interaction means that data items are *highlighted and associated*. In general, systems with multiple coordinated views are assumed to provide this interaction, but we found only four papers in our review that utilize this interaction. For this interaction, users can click and select a region to select the data in one or more views. In MarketAnalyzer, a system designed for analyzing sales data [KMJE12], users can draw a rectangle in order to select stores and products. This action further aggregates sales for the selected stores and products in other linked views. Similarly, the WireVis system [CGK*07] designed for transaction analysis, allows brushing and linking using the mouse that are reflected in all other views including the main heatmap view. In the same context, RiskVA [WJCR12], a visual analytics system for consumer credit analysis, allows users to select a market to highlight all related cells in the heatmap that present consumer base. Finally, Sorenson and Brath [SB13] utilize the mouse hover functionality to help users determine the dates of important events using linked views.

8.2. Analysis of Interaction Methods

We notice from our survey that the Abstract and Filter interactions have been the most popular interaction methods in the visual analytics systems for financial data. We also observe that the Select interaction has gained popularity after 2008 and sometimes it is combined with Connect. This implies that the brushing and linking technique [Kei02] is gaining popularity. On the other hand, we find that the Explore, Reconfigure, Encode, and Connect interaction methods have not been utilized much for financial data analysis. One reason could be that we are strict in identifying when an interaction method is implemented in a system. Unless a clear description for an interaction method is provided, we do not acknowledge the system supports the interaction method. But, we think that it is also possible that the descriptions for the interaction methods might not be provided in the paper, even though the interactions are supported. For example, 3D visualization is a good candidate for Reconfigure. However, we found only three papers [DG04, SH05, BM12] that provide information on the interaction methods. Also, when we started our survey, we expected that many financial systems would provide sorting techniques (one implementation of Reconfigure) as shown in Merino et al.'s list of "type of action" [MSK*06]. However, the number of systems that support sorting is much less than our expectations. This also reduces the number of systems providing Reconfigure. Similarly, we find that only 3 systems provide Connect in a system, but actually there are 15 systems that incorporate multiple visualization views. The number of systems that provide the Connect interaction is also lower than our expectations (about half of the systems) because it is considered to be a common interaction method with multiple coordinated views [BWK00].

Furthermore, we note that the systems that incorporate spatial maps for data exploration usually support the panning interaction (Explore). But we have been unable to find many papers that adopt maps for financial data analysis. Ko et al. [KMJE12] provide a map in their system to show aggregated sales results, but do not allow panning. This indicates that spatial data visualizations have not been extensively utilized for financial data analysis, even though many business domains continuously generate spatial data (e.g., marketing analysis, customer analysis, fleet/shipping administration).

	1992	1995	1996	1997		1998	1999		2000	2002		2003			2004			2005			2006	2007					2008				2009					2010				2011					2012			2013		2014	2016		
	Jungmeister92 [JT92]	Strausfeld95 [Str95]	Ankerst96 [AKpK96]	Gross97 [GSF97]	Brodbeck97 [BCLC97]	Kirkland98 [KSH*98]	Gresh99 [GRTM99]	Wattenberg99 [Wat99]	Groenen00 [GF00]	Eklund02 [EBVV02]	Xiong02 [XPH02]	Csallner03 [CHLS03]	Simunic03 [Sim03]	Taskaya03 [TA03]	Dwyer04 [DG04]	Nesbitt04 [NB04]	Shen04 [SE04]	Hao05 [HDKS05]	Lin05 [LCZ05]	Smeulders05 [SH05]	Schreck06 [SKM06]	Chang07 [CGK*07]	Mirel07 [MGBC07]	Schreck07 [STKF07]	Tekusova07 [TK07]	Ziegler07 [ZNK07]	Chiu08 [CHC08]	Dao08 [DBBS08]	Tekusova08 [TS08]	Ziegler08 [ZNK08]	Alsakran09 [AZZ09]	Huang09 [HLN09]	Rudolph09 [RSE09]	Sawant09 [Saw09]	Schreck09 [SBVLK09]	Alsakran10 [AZZ10]	Lei10 [LZ10]	Wu10 [WP10]	Ziegler10 [ZJGK10]	Lei11 [LZ11]	Sarlin11a [Sar11]	Sarlin11b [SE11]	Savikhin11 [SLFE11]	Schaefer11 [SWK*11]	Brath12 [BM12]	Ko12 [KMJE12]	Wang12[WJCR12]	Sorenson13 [SB13]	Chen13[CRVC13]	Lemieux14 [LSL*14]	Wanner16 [WJS*16]	Sum	From Industry
Use Case	■	■	■	■	■	■		■	■		■	■	■		■		■		■	■			■	■	■	■	■	■	■			■		■			■	■	■					■	■			■	■	■	■	33	9
Case Study							■			■				■		■		■				■								■	■				■	■					■	■				■	■					14	7
User Feedback															■										■				■								■			■					■			■		■		8	2
User Study																■																	■				■						■									4	1
Performance																		■			■																												■			3	0

Table 8: *Categories for evaluation methods. Use cases were mostly preferred for evaluation of visualizations or systems followed by case studies. In contrast, user study and evaluation by performance were not popular.*

9. Evaluations

Evaluating the performance of proposed visualizations or systems is an important aspect that needs to be considered before actual deployment of a system. In order to evaluate visualization techniques, Merino et al. [MSK*06] perform empirical performance experiments to validate the suitability of visualizations for different type of actions (e.g., identify, locate, distinguish, categorize, cluster, rank, compare, associate, and correlate). Marghescu [Mar07], on the other hand, focuses on the evaluation of visualization techniques for multidimensional data with low level business and data mining questions, including outlier detection, clustering description, and comparison of items. Compared to the two previous approaches, Jeong et al. [JDL*08] select a different approach to evaluate a visual analytics system. They focus on the analysis of user interaction that was recorded from their user study with the Wirevis system [CGK*07]. Their efforts are unique because they design a new visualization system whose sole purpose is to analyze the recorded interaction data. With the new system, they clarify the relationship between user interaction and financial visual analysis.

During our survey, we identified different types of evaluation methods that were utilized in visual analytics systems for financial data and categorized them into five different categories. These are: use case, case study, user feedback, user study, and performance. A summary of our review has been provided in Table 8. When a paper presents a use case, it generally focuses on describing a hypothetical use case scenario using the proposed technique or system. On the other hand, when a case study is used for evaluation, our criterion is that there are domain experts who use the system with domain-specific data sets and present detailed information on the results produced during analysis. We count a case study in a paper if at least one author is from industry or the paper explicitly mentions that external experts are involved in the evaluation process. If not, a paper is categorized as a use case.

Compared to a use case and a case study, a user study for evaluation is different in that there are actual participants (with no previous background of the proposed visualization or system) who use the system to solve questions designed to evaluate features of the systems (e.g., user studies with the surface visualization designed for stock trading pattern analysis [NB04], RingMap [LZ10] designed for stock market analysis, RiskVA [WJCR12] designed for consumer credit analysis), or to verify human response based on academic models (e.g., observing human decision-making processes in investments based on economic risk models [RSE09, SLFE11]).

Sometimes qualitative feedbacks are provided from users in user studies or domain experts in case studies to discuss pros and cons and personal preference of a visualization or system. However, the personal preference alone in the feedback, especially from users without domain knowledge, might not be a reliable indicator on usefulness of a proposed visualization or system, because the preference is not equal to usefulness and also there are many factors (e.g., appearance of a visualization) that may cause biases on a visualization. Finally, the performance category means that authors run their algorithms and compare results with other known standard methods. We find two papers that use this evaluation technique [HDKS05,SKM06]. Both of these works present methods for placing stock data, but different factors are evaluated, e.g., space efficiency in [HDKS05] and normalized ordering error in [SKM06].

We notice from Table 8 that use cases are the most popular method for evaluation, followed by case studies. As expected, it is rare that qualitative feedback alone is provided in order to evaluate a system. Such feedback is provided mainly along with use cases. We also confirm that not many techniques are tested using the performance testing strategy. There are two papers published in 2005 [HDKS05] and 2006 [SKM06] that conducted their evaluation by measuring performance in order to show usefulness of proposed algorithms for real-world data sets. We note that performance testing may raise issues during analysis as memory and other performance issues will need to be factored in as well.

10. Discussion

Financial data analysis is important in that it directly impacts assets and investments. Although there have been many visualization systems for analyzing financial data, financial market participants still use conventional visualization techniques in their everyday analysis. Why state-of-the-art visualization techniques have not been disseminated could be explained based on sparse collaboration between researchers and market participants; researchers cannot access real-world financial data, whereas analysts and investors have a hard time finding more suitable visualization techniques than what have been used for their analysis. To mitigate this, our work focuses on two aspects; 1) we interviewed a number of analysts with varying backgrounds in industry in order to clarify what analysts think is important in visualizations and analysis systems and what functionalities are needed for their tasks, and 2) we surveyed the visualization and visual analytics systems that support the financial data exploration in order to enhance the understanding of researchers, analysts, and investors on visual financial data exploration systems. Here we summarize, discuss, and derive insights.

There are 16 confirmed financial domains (Table 1), but not many domains have benefited from visualizations and visual analytics systems. We think, based on our survey, the work from the visual analytics community can cover up to 9 business domains (e.g., financial risk management, fraud/surveillance, economic analysis, capital management, and portfolio management) in terms of data sources and visualizations. In addition, most research on visualizations and systems concentrates on stock and fund data. This can be explained in terms of easy access to real-world data and the number of stakeholders. In contrast, not much research and collaboration have been performed for analyzing companies' performance. This is expected because such analysis could be involved with confidential data (e.g., sales). In the same context, we think that there have been some gaps in data sharing in analyses due to proprietary and privacy issues. If some data from multiple disciplines can be discussed and visualized together, analysts may find other insights. Therefore, many business domains remain underutilized that otherwise could inspire new visualization techniques for visual analytics communities.

Broadly speaking, there are three categories that utilize automated techniques: data reduction, clustering, and forecasting. Considering the fact that not many automated techniques were applied to financial data in the past, we think that there are still many opportunities for research in this aspect. The analysts we interviewed actually stressed that many automated algorithms are applied to their data in everyday analysis that are not available in the current visual analysis tools including Tableau and Spotfire. In addition, we think that it is possible that this narrow use of the automated techniques might be caused by the use of a single type of data– time-series. Thus applying different automated techniques to other data types could inspire new insights from data and corresponding visualizations. The analysts we interviewed also stressed the importance of the role of automated anomaly detection with multiple data sources. In this context, one can combine financial data with text data, such as financial news articles and microblog posts, in order to provide rich information to analyze the financial data. For example, news articles may explain detected anomalies [WJS*16] and social media posts may predict the stock market [BM11, BMZ11, MG12]. However, there still remain many challenges in connecting textual data to financial data due to the complexity of analyzing and correlating heterogeneous data sources.

In terms of preference among automated techniques, it is noticeable that the techniques for reduction (data size and dimensionality) have been popular. We conjecture that the advent of big data boosts adoption of automated techniques in visual analysis systems in order to reduce data size and produce clusters. In the same context, visualizations that are able to present more (processed) data will gain popularity in use as geometrically transformed and dense pixel displays have demonstrated.

Many businesses use various automated algorithms based on our interviews with the analysts, but they tend not to expose their approaches in publications. We think that this has restricted researchers from acquiring detailed information for researching and evaluating the performance of the applied algorithms on various financial data sets. In this context, it is worth mentioning that Sarlin [Sar15] provides good experimental comparison results among dimensional and data reduction algorithms applied to financial data.

Concerning interaction methods, we see strong preferences for the Abstract and Filter methods. It is possible that this preference is caused by the data used or a type of task during analysis. We think this implies a general process in analysis (e.g., visual information seeking mantra [Shn96]). Users often start analyzing data with higher abstraction levels in visual representations. Then, in order to move to another point of analysis, users tend to use filtering or change abstraction levels. Once a new insight is found, the data is reprocessed with automated techniques if applicable. In fact, this pattern of data analysis leads us to the knowledge generation model as shown in Figure 3. On the other hand, Connect has not gain much popularity, which is different from our initial expectation.

Explore is defined as a way to allow examination of different subset of data and we conservatively follow the definition for Explore categorization. This means, without explicit description in the reviewed papers, we did not consider a system for Interaction. This may be the reason behind such a low number of systems that support Explore. In addition, the additional constraint in the definition of Explore can further introduce difficulties in classification– "Explore interactions do not necessarily make complete changes in the data being viewed [YaKSJ07]." For example, some new data items are frequently entered while others are removed in a view. If this happens by rotating a sphere, it is difficult to decide whether the interaction activity is Explore or Reconfigure. Currently, it is categorized as Reconfigure [YaKSJ07].

Conventionally, evaluations of visualizations and systems heavily rely on use cases. This may prevent industry experts from using proposed techniques or approaches with confidence for big data analysis. In order to overcome this, evaluation via performance experiments is needed to provide results such as memory consumption and computation time. In addition, not many papers utilize domain experts for evaluation. This may be caused by the fact that domain experts who can bring their data for use on a system are not always available. In this case, we think very detailed qualitative feedback by non-expert users should be pursued, not to mention user studies with questions designed for public users.

In the paper collection stage, we found that there are visual systems for analyzing shareholder (stockholder) networks [Chu06, Chu09, FFS15, VLGS09]. In general the analyses for the networks require a data scraping phase from the web (e.g., [Chu06]) to collect information of financial entities including shareholders, supplier-customer, and corporates. Because the number of entities could be in the millions, automated algorithms are frequently applied such as MDS or SOM to reduce the size of the collected data. Once this pre-processing stage is done, the network is visualized using graph or network visualization techniques [Ead84, KV06, HMM00]. Although they are not the main focus of this survey, we introduce them here because they can be used for auxiliary financial analysis steps. Chung [Chu06, Chu09] proposes a visual tool called Stakeholder Network Visualizer for analyzing the relationship between stakeholders in a business intelligence environment. In the work, similarities between stakeholders are measured based on the textual information and referencing links on their web pages. The similarities are utilized for presenting a node-link diagram. Landesberger et al. [VLGS09]'s work propose a similar work where SOM is used for clustering shareholders. The clusters are projected onto a graph for analysis. Tekusova and Kohlhammer [TK08] also introduce a similar work, but their system provides information of cash flow and control rights of the network. In terms of data sources, the work by Fujita et al. [FFS15] is somewhat different in that their system provides the information of an economic network model comprising million firms and the links of supplier-customer and ownership relations on graphs.

11. Conclusion

In this work, we have presented a survey of the research on visual analysis approaches for exploring financial data. The fundamental motivation of this work was that there was little information derived from analysts. This encouraged us to conduct many interviews with domain analysts in the financial domain. With the information from the interviews in mind, we derived useful task requirements in the financial data analysis domain and categorize existing financial analysis approaches. Our survey shows that there are trends and preferences in data sources, automated techniques, visualizations, interaction methods and evaluation. Conversely, this implies that there are still many under-examined business domains due to several reasons (e.g., proprietary and privacy issues) where research on new techniques and systems are needed as simple line charts are still used with help of aggregation. We believe that the lessons from our survey shed light on understanding state-of-the-art visual analytics approaches for financial data analysis.

Acknowledgements

This work was supported in part by the U.S. Department of Homeland Security's VACCINE Center under Award Number 2009-ST-061-CI0001. Ko's work was supported in part by the 2016 Research Fund (1.160038.01) of UNIST (Ulsan National Institute of Science and Technology). Jang's work was supported by Institute for Information & communications Technology Promotion(IITP) grant funded by the Korea government(MSIP) (No.R0190-15-2016). The authors would like to thank the analysts for their valuable feedback.

References

[AES05] AMAR R. A., EAGAN J., STASKO J. T.: Low-level components of analytic activity in information visualization. In *Proceedings of IEEE Symposium on Information Visualization* (2005), pp. 111–117. 5, 12

[AKpK96] ANKERST M., KEIM D. A., PETER KRIEGEL H.: Circle segments: A technique for visually exploring large multidimensional data sets. In *Proceedings of IEEE Conference on Visualization, Hot Topic Session* (1996). 6, 11

[AMST11] AIGNER W., MIKSCH S., SCHUMANN H., TOMINSKI C.: *Visualization of Time-Oriented Data*. Human-Computer Interaction. Springer, 2011. 8

[AvH04] ABELLO J., VAN HAM F.: Matrix zoom: A visual interface to semi-external graphs. In *Proceedings of IEEE Symposium on Information Visualization* (2004), pp. 183–190. 11

[AZZ09] ALSAKRAN J., ZHAO Y., ZHAO X.: Visual analysis of mutual fund performance. In *Proceedings of International Conference on Information Visualisation* (2009), pp. 252–259. 11

[AZZ10] ALSAKRAN J., ZHAO Y., ZHAO X.: Tile-based parallel coordinates and its application in financial visualization. *Proceedings of International Society for Optical Engineering 7530* (2010), 753003–753003–12. 10, 13

[BCLC97] BRODBECK D., CHALMERS M., LUNZER A., COTTURE P.: Domesticating bead: adapting an information visualization system to a financial institution. In *Proceedings of IEEE Symposium on Information Visualization* (Oct 1997), pp. 73–80. 6, 7, 10, 12, 13

[Ber02] BERKHIN P.: *Survey of clustering data mining techniques*. Technical report, Accrue Software, 2002. 8

[Ber06] BERKHIN P.: A survey of clustering data mining techniques. In *Grouping Multidimensional Data: Recent Advances in Clustering* (2006), Springer, pp. 25–71. 7

[BG97] BORG I., GROENEN P.: *Modern multidimensional scaling, theory and applications*. Springer, 1997. 7

[BM11] BOLLEN J., MAO H.: Twitter mood as a stock market predictor. *Computer 44*, 10 (2011), 91–94. 15

[BM12] BRATH R., MACMURCHY P.: Sphere-based information visualization: Challenges and benefits. In *Proceedings International Conference on Information Visualization* (2012), pp. 1–6. 6, 9, 11, 12, 13

[BMZ11] BOLLEN J., MAO H., ZENG X.: Twitter mood predicts the stock market. *Journal of Computational Science 2*, 1 (2011), 1–8. 15

[BWK00] BALDONADO M. Q. W., WOODRUFF A., KUCHINSKY A.: Guidelines for using multiple views in information visualization. In *Proceedings of ACM International Working Conference on Advanced Visual Interfaces* (2000), pp. 110–119. 13

[CGK*07] CHANG R., GHONIEM M., KOSARA R., RIBARSKY W., YANG J., SUMA E., ZIEMKIEWICZ C., KERN D., SUDJIANTO A.: Wirevis: Visualization of categorical, time-varying data from financial transactions. In *Proceedings of IEEE Conference on Visual Analytics Science and Technology* (2007), pp. 155–162. 1, 6, 9, 11, 13, 14

[CHC08] CHIU T.-F., HONG C.-F., CHIU Y.-T.: Visualization of financial trends using chance discovery methods. In *Proceedings of New Frontiers in Applied Artificial Intelligence* (2008), Springer, pp. 708–717. 10

[Che73] CHERNOFF H.: Using faces to represent points in k–dimensional space graphically. *Journal of the American Statistical Association 68*, 342 (1973), 361–368. 8

[CHLS03] CSALLNER C., HANDTE M., LEHMANN O., STASKO J.: Fundexplorer: supporting the diversification of mutual fund portfolios using context treemaps. In *Proceedings of IEEE Symposium on Information Visualization* (Oct 2003), pp. 203–208. 11, 13

[Chu06] CHUNG W.: Evaluating the use of a visual approach to business stakeholder analysis. *In Proceedings of SIGHCI, Paper 9* (2006). 15, 16

[Chu09] CHUNG W.: Enhancing business intelligence quality with visualization: An experiment on stakeholder network analysis. *Pacific Asia Journal of the Association for Information Systems 1*, 1 (2009), 3. 15, 16

[CRVC13] CHEN N., RIBEIRO B., VIEIRA A., CHEN A.: Clustering and visualization of bankruptcy trajectory using self-organizing map. *Expert Systems with Applications 40*, 1 (2013), 385–393. 6, 7, 10

[DBBS08] DAO H. T., BAZINET A. L., BERTHIER R., SHNEIDERMAN B.: Nasdaq velocity and forces: An interactive visualization of activity and change. *Journal of Universal Computer Science 14*, 9 (2008), 1391–1410. 6, 9, 13

[DE02] DWYER T., EADES P.: Visualising a fund manager flow graph with columns and worms. In *Proceedings International Conference on Information Visualisation* (2002), pp. 147–152. 7, 9

[DG04] DWYER T., GALLAGHER D. R.: Visualising changes in fund manager holdings in two and a half-dimensions. *Information Visualization 3*, 4 (Dec. 2004), 227–244. 6, 7, 9, 12, 13

[DK98] DEBOECK G., KOHONEN T.: *Visual Explorations in Finance with Self-Organizing Maps*. Springer, 1998. 7

[DML14] DUMAS M., MCGUFFIN M. J., LEMIEUX V. L.: Financevis.net - a visual survey of financial data visualizations. In *Poster Abstracts of IEEE Conference on Visualization* (2014). 2

[Ead84] EADES P.: A heuristic for graph drawing. *Congressus Numerantium 42*, 10 (Jun 1984), 149–160. 16

[EBVV02] EKLUND T., BACK B., VANHARANTA H., VISA A.: Assessing the feasibility of self organizing maps for data mining financial information. *European Conference on Information Systems* (2002), 528–537. 10

[ED91] EVERITT B. S., DUNN G.: *Applied Multivariate Data Analysis*. Arnold, 1991. 7

[EDF08] ELMQVIST N., DRAGICEVIC P., FEKETE J.-D.: Rolling the dice: Multidimensional visual exploration using scatterplot matrix navigation. *IEEE Transactions on Visualization and Computer Graphics 14*, 6 (2008), 1539–1148. 8

[EMB07] EDWARDS R., MAGEE J., BASSETTI W.: *Technical Analysis of Stock Trends*. AMACOM, 2007. 1

[FFS15] FUJITA Y., FUJIWARA Y., SOUMA W.: Visualizing large-scale structure of a million-firms economic network. In *Proceedings of SIGGRAPH Asia 2015 Visualization in High Performance Computing* (2015), ACM, pp. 13:1–13:4. 15, 16

[Fin] FinanceVisBrowser. http://financevis.net/. Accessed by 5 Aug 2015. 2

[FLVW14] FLOOD M., LEMIEUX V., VARGA M., WONG B.: The application of visual analytics to financial stability monitoring. *Social Science Research Network, SSRN Scholarly Paper ID 2438194* (May 2014). 2

[Fod02] FODOR I.: *A Survey of Dimension Reduction Techniques*. Tech. rep., Lawrence Livermore National Laboratory, 2002. 8

[Fur86] FURNAS G. W.: Generalized fisheye views. In *Proceedings of ACM CHI Conference on Human Factors in Computing Systems* (1986), pp. 16–23. 9

[GF00] GROENEN P. J., FRANSES P. H.: Visualizing time-varying correlations across stock markets. *Journal of Empirical Finance 7*, 2 (2000), 155–172. 7, 10

[GRTM99] GRESH D., ROGOWITZ B., TIGNOR M., MAYLAND E.: An interactive framework for visualizing foreign currency exchange options. In *Proceedings of IEEE Conference on Visualization* (1999), pp. 453–562. 9

[GSF97] GROSS M. H., SPRENGER T. C., FINGER J.: Visualizing information on a sphere. In *Proceedings of IEEE Symposium on Information Visualization* (1997), pp. 11–16. 6, 9

[HDKS05] HAO M. C., DAYAL U., KEIM D. A., SCHRECK T.: Importance-driven visualization layouts for large time series data. In *Proceedings of IEEE Symposium on Information Visualization* (2005), pp. 203–210. 11, 13, 14

[HHN00] HAVRE S., HETZLER B., NOWELL L.: Themeriver: visualizing theme changes over time. In *Proceedings of IEEE Symposium on Information Visualization* (Oct. 2000), IEEE, pp. 115–124. 8

[HLN09] HUANG M. L., LIANG J., NGUYEN Q. V.: A visualization approach for frauds detection in financial market. In *Proceedings International Conference on Information Visualization* (July 2009), pp. 197–202. 11

[HMM00] HERMAN I., MELANÇON G., MARSHALL M. S.: Graph visualization and navigation in information visualization: a survey. *IEEE Transactions on Visualization and Computer Graphics 6*, 1 (2000), 24–43. 16

[Ins85] INSELBERG A.: The plane with parallel coordinates. *The Visual Computer 1*, 2 (1985), 69–91. 8

[JDL*08] JEONG D. H., DOU W., LIPFORD H. R., STUKES F., CHANG R., RIBARSKY W.: Evaluating the relationship between user interaction and financial visual analysis. In *Proceedings of IEEE Symposium on Information Visualization* (2008), IEEE Computer Society, pp. 83–90. 14

[JS91] JOHNSON B., SHNEIDERMAN B.: Tree-maps: A space-filling approach to the visualization of hierarchical information structures. In *Proceedings of IEEE Conference on Visualization* (1991), pp. 284–291. 11

[JT92] JUNGMEISTER W.-A., TURO D.: *Adapting treemaps to stock portfolio visualization*. Tech. rep., University of Maryland, 1992. 11, 13

[Kah98] KAHANER L.: *Competitive Intelligence: How to Gather Analyze and Use Information to Move Your Business to the Top*. Touchstone Press, New York, U.S.A., 1998. 6

[Kei00] KEIM D. A.: Designing pixel-oriented visualization techniques: Theory and applications. *IEEE Transactions on Visualization and Computer Graphics 6*, 1 (2000), 59–78. 8

[Kei02] KEIM D.: Information visualization and visual data mining. *IEEE Transactions on Visualization and Computer Graphics 8*, 1 (2002), 1–8. 2, 8, 10, 13

[KHL*01] KEIM D. A., HAO M. C., LADISCH J., HSU M., DAYAL U.: Pixel bar charts: A new technique for visualizing large multi-attribute data sets without aggregation. In *Proceedings of IEEE Symposium on Information Visualization* (2001), pp. 113–120. 11

[KKEM10] KEIM D. A., KOHLHAMMER J., ELLIS G., MANSMANN F.: *Mastering The Information Age-Solving Problems with Visual Analytics*. Florian Mansmann, 2010. 1, 4

[KMJE12] KO S., MACIEJEWSKI R., JANG Y., EBERT D. S.: Marketanalyzer: An interactive visual analytics system for analyzing competitive advantage using point of sale data. *Computer Graphics Forum 31*, 3 (2012), 1245–1254. 1, 7, 9, 11, 12, 13

[Kof13] KOFFKA K.: *Principles of Gestalt Psychology*. Routledge, 2013. 9

[Koh90] KOHONEN T.: The self-organizing map. *Proceedings of IEEE 78*, 9 (1990), 1464–1480. 7

[KRS02] KOHAVI R., ROTHLEDER N., SIMOUDIS E.: Emerging trends in business analytics. *ACM Communications 45*, 8 (2002), 45–48. 6

[KSH*98] KIRKLAND J. D., SENATOR T. E., HAYDEN J. J., DYBALA T., GOLDBERG H. G., SHYR P.: The nasd regulation advanced detection system (ads). In *Proceedings of National/Tenth Conference on Artificial Intelligence/Innovative Applications of Artificial Intelligence* (1998), American Association for Artificial Intelligence, pp. 1055–1062. 6, 8, 9

[KV06] KENNINGS A. A., VORWERK K.: Force-directed methods for generic placement. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems 25*, 10 (2006), 2076–2087. 16

[LCZ05] LIN L., CAO L., ZHANG C.: The fish-eye visualization of foreign currency exchange data streams. In *Proceedings of Asia-Pacific Symposium on Information Visualization* (2005), Australian Computer Society, pp. 91–96. 6, 7, 9, 13

[Lia05] LIAO T. W.: Clustering of time series data: A survey. *Pattern Recognition 38*, 11 (2005), 1857–1874. 8

[LSL*14] LEMIEUX V. L., SHIEH B. W., LAU D., JUN S. H., DANG T., CHU J., TAM G.: Using visual analytics to enhance data exploration and knowledge discovery in financial systemic risk analysis: The multivariate density estimator. In *Proceedings of iConference* (2014), iSchools, pp. 649–653. 6, 11

[LZ10] LEI S., ZHANG K.: A visual analytics system for financial time-series data. In *Proceedings of International Symposium on Visual Information Communication* (2010), ACM, pp. 20:1–20:9. 7, 10, 11, 12, 13, 14

[LZ11] LEI S., ZHANG K.: Visual signatures for financial time series. In *Proceedings of International Symposium on Visual Information Communication* (2011), ACM, pp. 16:1–16:10. 10, 13

[Mac67] MACQUEEN J.: Some methods for classification and analysis of multivariate observations. In *Proceedings of the 5th Berkeley Symp. on Mathematics Statistics and Probability* (1967). 7

[Mar07] MARGHESCU D.: *Multidimensional Data Visualization Techniques for Financial Performance Data: A Review*. Tech. rep., Turku Centre for Computer Science, 2007. 3, 14

[MG12] MITTAL A., GOEL A.: Stock prediction using twitter sentiment analysis. *Standford University, CS229 (2011 http://cs229. stanford. edu/proj2011/GoelMittal-StockMarketPredictionUsingTwitterSentimentAnalysis. pdf)* (2012). 15

[MGBC07] MIREL B., GOLDSMITH P., BRATH R., CORT B.: Visual analytics for model-based policy analysis: Exploring rapid changes in commodities markets. In *Proceedings of International Conference on Digital Government Research: Bridging Disciplines & Domains* (2007), Digital Government Society of North America, pp. 312–313. 9

[MSK*06] MERINO C. S., SIPS M., KEIM D. A., PANSE C., SPENCE R.: Task-at-hand interface for change detection in stock market data. In *Proceedings of ACM Conference on Advanced Visual Interfaces* (2006), pp. 420–427. 3, 9, 11, 13, 14

[Mur99] MURPHY J.: *Technical Analysis of the Financial Markets: A Comprehensive Guide to Trading Methods and Applications*. New York Institute of Finance, 1999. 8

[NB04] NESBITT K., BARRASS S.: Finding trading patterns in stock market data. *IEEE Computer Graphics and Applications 24*, 5 (2004), 45–55. 6, 9, 14

[oec] Oecd data. `https://data.oecd.org/`. 6

[PC05] PIROLLI P., CARD S.: The sensemaking process and leverage points for analyst technology as identified through cognitive task analysis. In *Proceedings of International Conference on Intelligence Analysis* (2005), pp. 2–4. 4

[Pry10] PRYKE M.: Money's eyes: the visual preparation of financial markets. *Economy and Society 39*, 4 (2010), 427–459. 2

[RSE09] RUDOLPH S., SAVIKHIN A., EBERT D.: Finvis: Applied visual analytics for personal financial planning. In *Proceedings of IEEE Conference on Visual Analytics Science and Technology* (2009), pp. 195–202. 6, 8, 9, 14

[San13] SANTOVENA A. Z.: *Big data : evolution, components, challenges and opportunities*. Master's thesis, Massachusetts Institute of Technology, Sloan School of Management, 2013. 1

[Sar11] SARLIN P.: Sovereign debt monitor: A visual self-organizing maps approach. In *Proceedings of IEEE Symposium on Computational Intelligence for Financial Engineering and Economics* (2011), pp. 1–8. 6, 10

[Sar15] SARLIN P.: Data and dimension reduction for visual financial performance analysis. *Information Visualization 14*, 2 (2015). 15

[Saw09] SAWANT A.: Stockviz: Analyzing the trend of stocks in major auto, oil, consumer, and technology companies. In *Proceedings of International Conference on Modeling, Simulation & Visualization Methods* (2009), pp. 278–284. 10

[SB13] SORENSON E., BRATH R.: Financial visualization case study: Correlating financial timeseries and discrete events to support investment decisions. In *Proceedings of International Conference on Information Visualization* (2013), pp. 232–238. 1, 8, 10, 12, 13

[SBVLK09] SCHRECK T., BERNARD J., VON LANDESBERGER T., KOHLHAMMER J.: Visual cluster analysis of trajectory data with interactive kohonen maps. *Information Visualization 8*, 1 (Jan. 2009), 14–29. 6, 9, 10, 12

[Sch14] SCHWABISH J.: An economist's guide to visualizing data. *Journal of Economic Perspectives 28*, 1 (2014), 209–234. 2

[SE04] SHEN X., EADES P.: Using moneytree to represent financial data. In *Proceedings International Conference on Information Visualization* (2004), pp. 285–289. 10

[SE11] SARLIN P., EKLUND T.: Fuzzy clustering of the self-organizing map: Some applications on financial time series. In *Advances in Self-Organizing Maps*, vol. 6731 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, 2011, pp. 40–50. 6, 7, 10

[SH05] SMEULDERS R., HEIJS A.: Interactive visualization of high dimensional marketing data in the financial industry. In *Proceedings International Conference on Information Visualization* (2005), pp. 814–817. 6, 9, 13

[Shn96] SHNEIDERMAN B.: The eyes have it: A task by data type taxonomy for information visualizations. In *Proceedings of IEEE Symposium on Visual Languages* (1996), pp. 336–343. 5, 6, 12, 15

[Sim03] SIMUNIC K.: Visualization of stock market charts. In *Proceedings of International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision* (2003). 6, 7, 9, 13

[SKM06] SCHRECK T., KEIM D., MANSMANN F.: Regular treemap layouts for visual analysis of hierarchical data. In *Proceedings of ACM Conference on Computer Graphics* (2006), pp. 184–191. 14

[SLFE11] SAVIKHIN A., LAM H. C., FISHER B., EBERT D.: An experimental study of financial portfolio selection with visual analytics for decision support. In *Proceedings of Hawaii International Conference on System Sciences* (Jan 2011), pp. 1–10. 6, 8, 14

[SSS*14] SACHA D., STOFFEL A., STOFFEL F., KWON B. C., ELLIS G., KEIM D. A.: Knowledge generation model for visual analytics. *IEEE Transactions on Visualization and Computer Graphics 20*, 12 (Dec. 2014), 1604–1613. 4

[STKF07] SCHRECK T., TEKUŠOVÁ T., KOHLHAMMER J., FELLNER D.: Trajectory-based visual analysis of large financial time series data. *ACM SIGKDD Exploration Newsletter 9*, 2 (2007), 30–37. 6, 7, 9, 13

[Str95] STRAUSFELD L.: Financial viewpoints: Using point-of-view to enable understanding of information. In *Proceedings of ACM CHI Conference on Human Factors in Computing Systems* (1995), ACM, pp. 208–209. 9

[SWK*11] SCHAEFER M., WANNER F., KAHL R., ZHANG L., SCHRECK T., KEIM D. A.: A Novel Explorative Visualization Tool for Financial Time Series Data Analysis. In *Proceedings of UKVAC Workshop on Visual Analytics* (2011). 9, 11, 13

[TA03] TASKAYA T., AHMAD K.: Bimodal visualisation: A financial trading case study. In *Proceedings International Conference on Information Visualization* (2003), IEEE Computer Society, pp. 320–326. 6, 9, 13

[TC05] THOMAS J. J., COOK K. A. (Eds.): *Illuminating the Path: The R&D Agenda for Visual Analytics*. IEEE Press, 2005. 1, 2, 4

[Teg99] TEGARDEN D. P.: Business information visualization. *Communications of the Association for Information Systems 1*, 4 (1999), 197–204. 2

[TK07] TEKUSOVA T., KOHLHAMMER J.: Applying animation to the visual analysis of financial time-dependent data. In *Proceedings International Conference on Information Visualization* (2007), pp. 101–108. 10

[TK08] TEKUŠOVÁ T., KOHLHAMMER J.: Visual analysis and exploration of complex corporate shareholder networks. In *Proceedings of SPIE Conference on Visualization and Data Analysis* (2008), pp. 68090F–68090F. 16

[TS08] TEKUSOVA T., SCHRECK T.: Visualizing time-dependent data in multivariate hierarchic plots - design and evaluation of an economic application. In *Proceedings International Conference on Information Visualization* (July 2008), pp. 143–150. 6, 9, 12, 13

[Tuf90] TUFTE E. R.: *Envisioning Information*. Graphics Press, 1990. 9

[VLGS09] VON LANDESBERGER T., GÖRNER M., SCHRECK T.: Visual analysis of graphs with multiple connected components. In *Proceedings of IEEE Symposium on Visual Analytics Science and Technology* (2009), IEEE, pp. 155–162. 15, 16

[Wat99] WATTENBERG M.: Visualizing the stock market. In *Proceedings of ACM CHI Conference on Human Factors in Computing Systems, Extended Abstracts* (1999), pp. 188–189. 6, 8, 11

[WGK10] WARD M. O., GRINSTEIN G. G., KEIM D. A.: *Interactive Data Visualization - Foundations, Techniques, and Applications*. A K Peters, 2010. 8

[WJCR12] WANG X., JEONG D., CHANG R., RIBARSKY W.: Riskva: A visual analytics system for consumer credit risk analysis. *Tsinghua Science and Technology 17*, 4 (Aug 2012), 440–451. doi:10.1109/TST.2012.6297590. 6, 9, 11, 12, 13, 14

[WJS*16] WANNER F., JENTNER W., SCHRECK T., STOFFEL A., SHARALIEVA L., KEIM D. A.: Integrated visual analysis of patterns in time series and text data - workflow and application to financial data analysis. *Information Visualization 15*, 1 (2016), 75–90. 6, 7, 9, 12, 13, 15

[wor] Wolrd bank data. http://data.worldbank.org. 6

[WP10] WU E., PHILLIPS P.: Financial markets in motion: Visualising stock price and news interactions during the 2008 global financial crisis. *Procedia Computer Science 1*, 1 (2010), 1765–1773. 6, 10

[XPH02] XIONG F., PRAKASH E., HO K.: E-r modeling and visualization of large mutual fund data. *Journal of Visualization 5*, 2 (2002), 197–204. 9

[YaKSJ07] YI J. S., AH KANG Y., STASKO J. T., JACKO J. A.: Toward a deeper understanding of the role of interaction in information visualization. *IEEE Transactions on Visualization and Computer Graphics 13*, 6 (2007), 1224–1231. 2, 12, 15

[ZJGK10] ZIEGLER H., JENNY M., GRUSE T., KEIM D.: Visual market sector analysis for financial time series data. In *Proceedings of IEEE Conference on Visual Analytics Science and Technology* (2010), pp. 83–90. 7, 11, 13

[ZNK07] ZIEGLER H., NIETZSCHMANN T., KEIM D.: Visual exploration and discovery of atypical behavior in financial time series data using two-dimensional colormaps. In *Proceedings International Conference on Information Visualization* (July 2007), pp. 308–315. 11

[ZNK08] ZIEGLER H., NIETZSCHMANN T., KEIM D. A.: Visual analytics on the financial market: Pixel-based analysis and comparison of long-term investments. In *Proceedings International Conference on Information Visualization* (2008), pp. 287–295. 11, 13