Proceedings of the 2023 International Conference on Software Engineering and Machine Learning
DOI: 10.54254/2755-2721/27/20230065

Review on audit data visualization method based on R language

Weixuan Ding

School of International Education, Jinling Institute of Technology, NanJing, JiangSu, China, 211169

youngboy0810@163.com

Abstract. To be competitive in today's market, companies must constantly assess and improve their operations, and financial data analysis has emerged as a crucial tool for this purpose. Executives can utilize data analysis to gain a deeper understanding of the underlying facts in their data and make better informed decisions about their company and the market. Multiple data models in the language make it possible to have a fully functional language environment with tools for statistical analysis and visual visualization of data. With its ability to enhance the quality of work in statistical computations and graphical analysis, the R language is ideally suited to the industrial data analysis environment. In this paper, we use a literature review methodology to examine the domestic and international literature on R language big data visualization auditing, analyze the visualization application areas of R language big data auditing, discuss the benefits and challenges of big data auditing, demonstrate how R language can be adapted to real-world auditing scenarios, and offer reasonable recommendations and optimistic outlooks for the field's future.

Keywords: R language, data visualization, big data auditing.

1. Introduction

Humans are currently riding the big data wave, which is the result of the recent data boom. When dealing with audit projects in many sectors, the auditing sector encounters a similarly massive volume of data information. The adoption of big data will have a substantial impact on promoting national auditing services for national government and will result in profound shifts in the way auditors think about their work [1], the range of activities audits can do, and the weight audit findings carry. Big data finance is gaining traction as a result of the Internet's quick and robust growth in the financial sector. This has opened up new possibilities for the financial sector. Many financial institutions have actively explored and achieved significant progress in the application of big data, despite the fact that they must deal with a mountain of customer data, a mountain of transaction records, and a mountain of information sources. Simultaneously, as computer-aided audit matures, financial audit in the implementation of computer software technology places greater emphasis on the widespread adoption of computers, the Internet, automated services, and similar innovations. Therefore, a key research direction for computer-aided auditing in China is how to apply new technologies and methods in the big data environment to the practical work of financial auditing, how to push the boundaries in big

DOI: 10.54254/2755-2721/27/20230065

data collection, analysis, and mining, how to use big data technology to rank problem data, track the flow of funds, and determine audit priorities while avoiding the risks brought about by big data.

While traditional methods of data presentation have been phased out in favor of R language-based application statistical analysis tools, visualization technology has finally reached the stage where it may be put to practical use. When it comes to financial data analysis, integrating visualization technology with enterprise management has the potential to not only display the results of enterprise financial data analysis more intuitively for managers to make efficient decisions, but also to improve the economic efficiency of enterprises in a roundabout way and to introduce the business status of enterprises to potential investors in a way that is both accessible and informative. For this reason, advancements in the field of data visualization software are warranted. In light of the limitations of the standard financial statistical tools and methods, this paper provides a thorough and in-depth analysis of the financial data of various businesses by contrasting the use of the R language in a number of targeted audit cases and providing a graphical representation of the results of analyses of key financial metrics like solvency, operating capacity, and profitability within the R language's software environment. Business organizations can then use this information as a springboard for appropriate preventative action. The author provides a thorough and detailed explanation of the relevant concepts of R language language code, and applies the common functions and models in R language combined with financial data, with a focus on verifying the advantages of R language in the application of financial data visualization, allowing business managers to form a more reliable opinion of the financial state of their company thanks to the information provided by R language.

2. Analysis

Currently, the main methods for visual analysis of financial data based on R language are scatter diagram, face chart, scarter diagram, etc.

2.1. Scatter diagram

In the practical application of the R language, the relationship between two variables is typically represented graphically using a scatter diagram, as shown in figure 1, where each point represents a sample from the target data set and the more data it contains, the more accurate the comparison [3].

In the course of business operations, a great deal of information is generated, which in turn generates a great deal of financial data. In order to better compare with competitors in the same industry, corporate finance personnel must deal with more massive data. Due to the absence of reasonable judgments regarding the relationships between variables, financial reports are limited in a variety of ways for their users [4]. However, scatter diagrams do not present this issue since, on the one hand, each point represents independent data and, on the other hand, the relationship between the individual data can be adequately predicted by examining the entire picture. By studying the scatter distribution, it is possible to successfully address problems that cannot be resolved using conventional data display methods, and to make the internal logical relationship of data more obvious [5]. In addition, the financial data can be marked on the scatter diagram with the aid of a specific function model in the R language environment, allowing financial data organizers to perform preliminary data processing for financial data users when using the R language system to organize the data and assist financial data users in locating the key data.

DOI: 10.54254/2755-2721/27/20230065

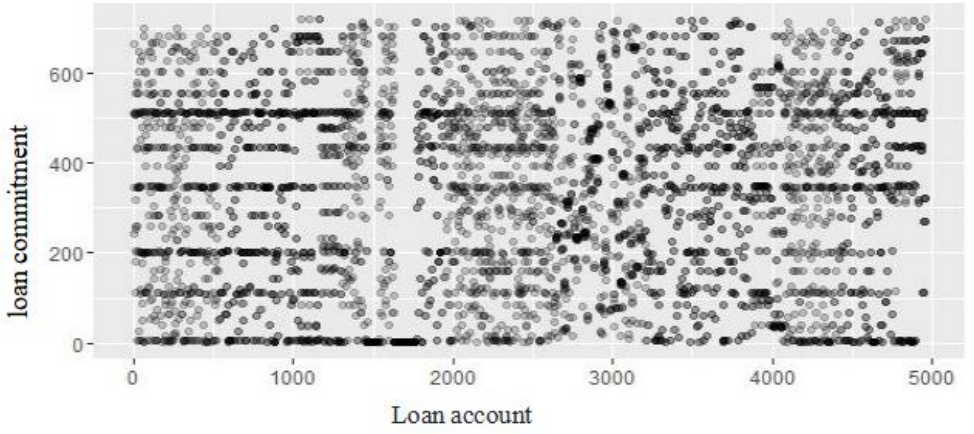

Figure 1. Scatter diagram of loan position for all loan accounts.

2.2. Face chart

The face chart is currently the most popular analysis tool in the R language for the display of financial data, and this piece of writing is a case study that mostly uses the face chart. The face chart presents the data in different dimensions by altering the form or size of a face, and each sample point can be used to design a face [6], such as figure 2. In addition, the face chart displays the data in several dimensions simultaneously. When compared to more traditional forms of financial reporting, the face chart visualization is both more cutting edge and more readily acceptable. Because of its unique format, it is capable of leaving a more profound impression on the human brain. This is particularly useful when dealing with a large amount of financial data, as the conventional method involves providing a large number of line graphs and bar charts to show the status of the financial operation of the corporation. This method is not only difficult to understand, but it is also easy to forget.

However, by utilizing the face chart format, not only is it impressive, but it also allows users to typically dismiss the worthless information and maintain the subjective information that they find valuable when confronted with complex financial data [7]. Therefore, by analyzing the efficacy of the use of face charts, we can see that face charts can not only show the changes in data more intuitively, but they can also deepen the impression of users. This helps improve the judgment of business operators and makes it easier for them to deal with potential crises in the future. The classic multi-chart analysis is more limited in scope than the face chart, which makes it easier and quicker for financial analysts to comprehend the current state of the company's finances.

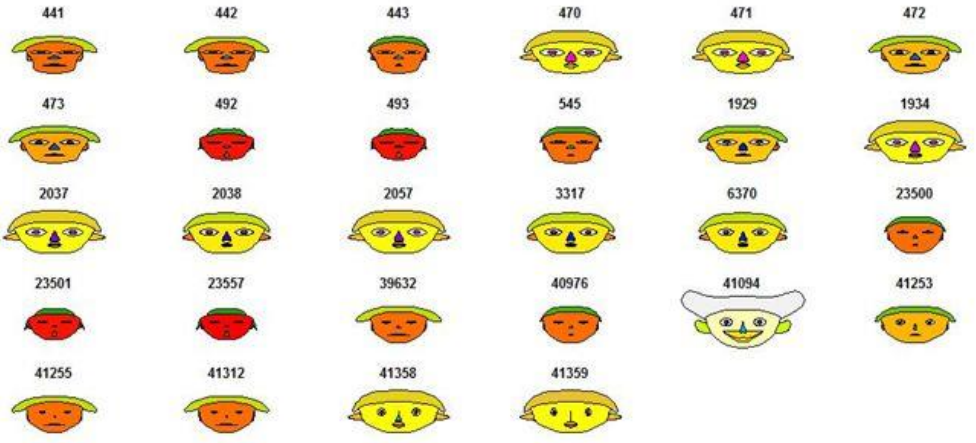

Figure 2. Financial Data Face Chart of a company Cluster diagram.

DOI: 10.54254/2755-2721/27/20230065

2.3. Cluster diagram

Cluster analysis is a method that begins with data from samples and then automatically groups all of the data together into reasonably similar clusters based on the degree of similarity between different persons or different samples, such as figure 3. Common approaches to cluster analysis include systematic clustering, decomposition, joining, dynamic clustering, ordered sample clustering, overlapping clustering, and fuzzy clustering [8]. Other approaches include fuzzy clustering, overlapping clustering, and overlapping ordered sample clustering. Auditors may find it difficult to detect potential problems by direct observation and inspection when confronted with large and complex financial data. As a result, auditors may wish to draw cluster diagrams, conduct exploratory studies of the data, and compare the rationality of homogeneous clusters with the actual situation [9]. The data can be further analyzed for anomalies for clusters that differ from the actual situation, which allows the audited unit to be ranked for problems in relevant business operations or internal control management [10]. This is completed so that the audited unit can be ranked for the severity of the problems. There are many functions in the R language that can be utilized for cluster analysis. Beginning with the data, clustering maps are drawn by various algorithms to explore similar results under different algorithms. This is helpful for auditors in that it allows them to lock the audit focus and summarize the issues.

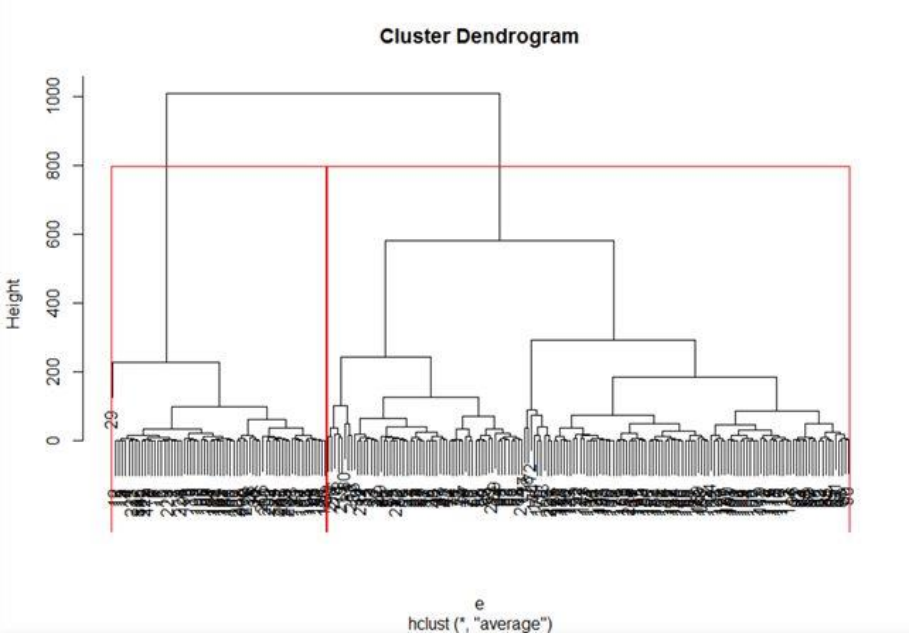

Figure 3. Systematic clustering diagram under the class averaging method.

3. Conclusion

The interpretation and application of business financial data has traditionally been the primary focus of theoretical study, as well as one of its most challenging aspects. Scholars from both inside and outside the country are interested in figuring out how to examine the financial risks that businesses face based on the information they have about their finances [11]. A new financial financial data data analysis model is reviewed in this paper by utilizing the one-of-a-kind benefits of the R language in the areas of data analysis and visualization. This model is derived through the statistics and analysis of enterprise financial data. This model makes it possible for more people to read and understand the financial data of enterprises with the help of the visualization ability of the R language. It also opens a new window for the research of related theories.

DOI: 10.54254/2755-2721/27/20230065

Second, the financial accounting personnel of enterprises are to some extent also considered to be a part of the enterprise management. Despite the fact that they will not participate in the decision-making work of enterprise development, they will still be responsible for the collection and collation of enterprise financial information. If visualization can be added to the process of collating and collecting financial data, it will make the enterprise financial personnel more familiar with how to judge the enterprise economic trends with the help of financial information [12]. This will allow the enterprise financial personnel to be more oriented toward meeting the needs of the enterprise decision makers when collating financial information. In the long term, corporate financial data can not only make the information that has been recorded more perfect after it has been arranged, but it can also allow for greater readability for workers who are not involved in finance.

Although computer-aided financial data analysis has received more attention as of late, and the implementation of data visualization software such as R language is being gradually promoted, the development and utilization of R language in the practice of financial data visualization is still manifestly insufficient. This paper merely provides a summary and discussion of the visualization of financial data for a few different private businesses. The impact of the study on the visualization of financial data in other businesses or administrative institutions has not yet been evaluated. On the one hand, because the R language has such a wide variety of extension packages, there are many function models that can be used in the application of financial data visualization. However, these models cannot be introduced and shown in this paper, and the visualization graphics that can be realized cannot be fully developed because of the limitations of the data. The author, on the other hand, holds the opinion that R-based financial data visualization offers a vast application field and a significant potential for further growth in the future.

References

[1] Zhang Yanfei. Analysis of financial informatization enhancement path based on visualization perspective [J]. Business Accounting,2016(4): 54.
[2] Chen T, Xia Cuijuan, Liu Wei, Zhang Lei. Research and implementation of visualization technology for correlated data[J]. Library Intelligence Work,2017(17): 33.
[3] Wu L.P. How financial analysis can effectively provide reference for business decision making [J]. China Business Theory, 2018(27).
[4] Zhang Jiping. Research on financial analysis and enterprise business decision-making[J]. Modern Business,2018(18):14.
[5] Zou Zongfeng, Xia Yun et al. The design of AHP analysis program based on R language[J]. Journal of Mathematical Medicine, 2015, 4: 468-470.
[6] Nir Kshetri. big data impact on privacy, security and consumer wel-fare[J].
[7] Jared P. Rand. R language: practical data analysis and visualization techniques [M]. Beijing: Machinery Industry Press,2015.4: 282-284.
[8] Ke F,Xin Jia Y. Research on financial risk prevention based on data mining technology[J]. Economic Forum,2015,7:100-102.
[9] Maria Brigida Ferraro, Paolo Giordani. A toolbox for fuzzy clustering using the Rprogramming language [J]. Fuzzy Sets and Systems,2015.05:1-16.
[10] Mc Kinsey Global Institute. Big data: the next frontier for innovation, competition, and productivity [EB/OL]. http://www.mckinsey.com/business-functions/business-technology/our-insights/big-data-the-next-frontier-for-innovation, 2011-5.
[11] Yu, X. S., Zhou, N. Research on high-dimensional data visualization methods[J]. Intelligence Science,2017(01):60.
[12] Peng Rongxuan. Analysis of trends and types of data visualization technology applications[J]. Western Resources,2017(06):30-44.