DATA VISUALIZATION USING THE RAPID MINER APPLICATION TO EVALUATE SALES PATTERNS

Naomi Sirait, Sahat Yeremia Situmeang, Widalius Zaluku, Wira Matias Panjaitan
STMIK Pelita Nusantara

Abstract. In today's digital era, data has become a valuable asset for many companies, including those in the retail and sales sectors. Using the right method to interpret and understand this data is the key to improving business strategy and making the right decisions. In this context, data visualization plays a crucial role as a tool for presenting complex information in an intuitive and easy-to-understand way. This study aims to utilize the RapidMiner application in visualizing data for evaluating sales patterns. Using a sales dataset from a retail company, we perform a series of analytical processes using RapidMiner. Various visualization techniques, such as bar graphs, line charts, heat maps and more, are applied to extract insights from sales data and identify certain patterns. The visualization results show several interesting findings, such as the season or period with the highest sales, the best-selling products, and the relationship between sales variables. Through this visualization, we are able to provide strategic recommendations to companies to increase revenue and optimize marketing strategies. The conclusion of this study is that RapidMiner, with its capabilities in data visualization, is an effective tool to help companies understand sales patterns and make data-driven decisions.

Keywords. Data Visualization, RapidMiner, Sales Patterns, Evaluation, Business Strategy.

INTRODUCTION

In the midst of the rapid development of information technology and digitalization, the business world, especially the sales sector, is facing different challenges and opportunities[1]. Data is a valuable source of information that can reveal important insights for business growth[2]. However, the large volume of data and the complexity of the information contained in it often makes it difficult for companies to interpret it. Therefore, data visualization is an important solution to facilitate understanding and interpretation of data.

RapidMiner, as one of the leading data analysis tools, offers visualization features that allow users to understand data more quickly and intuitively[3]. This application makes it easier for companies to identify trends, patterns and anomalies in sales data, which in turn supports more informed business decision making[4].

Sales pattern evaluation is the process of analyzing sales data to identify trends, patterns, and buyer behavior[5]. Using RapidMiner and data visualization can identify valuable information such as:

1. Sales trends over time.
2. Seasonal patterns, such as spikes in sales during certain holidays or seasons.
3. Buyer behavior based on geographic or demographic location.
4. Correlation between sales of certain products.
5. The influence of external factors, such as promotions, weather or industry trends.

Research by Zhou, W.[6] This study evaluates the effectiveness of RapidMiner in visualizing sales data. Researchers created interactive dashboards that allow users to see and understand sales trends and patterns better. This research shows that the resulting visualization can help in quick identification of problems and sales opportunities. This research aims to: present sales data with various visualization techniques using RapidMiner, identify relevant sales patterns to support business decisions and evaluate the effectiveness of RapidMiner as a visualization tool in the context of evaluating sales patterns.

Study by Hassan, MR,[7] In this study, RapidMiner is used for analysis and visualization of e-commerce transaction data. This study includes shopping basket analysis and identification of sales patterns. Researchers found that data visualization helps in exploring sales patterns and provides insights to improve sales strategies.

Research by Sharma, P.,[8] This study utilizes RapidMiner for the analysis and visualization of sales data in the retail sector. RapidMiner is used for data preprocessing and analysis, while data visualization helps in understanding sales patterns and figuring out factors influencing sales.

From this research, it is hoped that companies can understand their sales patterns more clearly and in depth, interested parties gain insight into the potential of RapidMiner as a visualization tool and provide strategic recommendations for companies to optimize revenue and marketing strategies.[9]

The results of these studies show that RapidMiner is an effective tool for sales data analysis and visualization. However, individual studies show that the effectiveness of RapidMiner is highly dependent on data quality, proper pre-processing, and selection of appropriate analytical techniques. Therefore, it is important to gain a good understanding of how to use RapidMiner effectively for sales data visualization.

**METHOD**

To conduct research on data visualization with the RapidMiner application to evaluate sales patterns, here are some research method steps that can be done[10][11]:

1. **Determination of Research Objectives:** Determine your research objectives. For example, you may want to evaluate the sales pattern of a particular product, evaluate the effect of a promotion on sales, or predict future sales based on past patterns.
2. **Data collection:** Gather relevant sales data. This data can come from a variety of sources, such as internal databases, sales management systems, or customer survey data. Make sure the data includes variables that are relevant for your research purposes[12].
3. **Preprocessing Data:** Import data into RapidMiner and do preprocessing. This may include steps such as removing duplicates, filling in missing values, or
performing data transformations. The goal is to make sure the data is ready for analysis[13].

4. **Data Exploratory Analysis:** Use the data visualization techniques in RapidMiner, such as histograms, scatter plots, and box plots, to understand the characteristics and structure of your data[14].

5. **Data Modeling:** Based on the exploratory analysis, choose the right model for your research purposes. RapidMiner provides a variety of machine learning algorithms, including regression, clustering, and classification.

6. **Result Visualization:** After modeling, visualize the results using various graphs and diagrams available in RapidMiner. This will help you interpret the results and find patterns or insights you might have missed[15].

7. **Evaluation and Interpretation of Results:** Evaluate your model or findings. Does the model have good performance? Are the results as expected? Interpret your findings in the context of research objectives and business practice.

8. **Application of Results:** Apply your findings in business decision making. For example, you might decide to improve your sales strategy based on identified sales patterns, or you might identify a product or service that needs to be improved.

9. **Documentation and Reporting:** Document your research process and findings. Create reports that include methods, results, and interpretations. This will help you and others understand your research and findings, and will be useful for future reference[16].

**RESULTS AND DISCUSSION**

Goods sales transaction data usually includes information about the date of the transaction, products sold, quantity of goods, price per unit, total price, and so on.

1. **Prepare initial data or raw data in the form of goods sales transaction data**

![Figure 1. Raw data in the form of goods sales transaction data](image-url)
2. Perform data cleaning process or data preprocessing

![Figure 2. Data Cleaning Process](image)

3. The data that has been cleaned is then converted into a tabular format and stored in a new excel file.

![Figure 3. Data in Tabular format](image)

4. Open the Rapid Miner application that has been downloaded.

![Figure 4. RapidMiner display](image)
5. The next step is to input data by clicking import data

![Figure 5. Import Data](image)

6. Click on the data that has been saved in the document then click next.

![Figure 6. Save Data](image)

7. Select the cell to be imported, namely the cell in the form of numeric data, while the item name is unnecessary, then click next.

![Figure 7. Cells in the form of numbers](image)
8. After obtaining the column format as shown below, click next.

![Column Format](image)

**Figure 8. Column Format**

9. Save the formatted data, then click finish.

![Save data](image)

**Figure 9. Save data that has been formatted**

10. Click Design to start the process

![Design for Process](image)

**Figure 10. Design for Process**
11. Drag the data format that was saved earlier to the process section.

![Figure 11. Drag formatted data to the processing section](image)

12. Next, look for the Numerical to Binomial operator, drag it to the process section and connect it to the data format section.

![Figure 12. Numerical to Binomial operator](image)

13. Look for the FP-growth operator then drag it to the process section and connect it to numerical to binomial.

![Figure 13. FP-growth operator](image)
14. Looked for the Create Association Rules operator and added to the process section.

The exa point of FP-growth is connected directly to the res point, while the fre point is connected to the ite point Create Association Rules[15]. Then the rul and ite points of the Create Association Rules operator are connected directly to the res point. Then set the support and confidence values and click the run button.

![Create Association Rules](image1.png)

**Figure 14. Create Association Rules**

15. After clicking the run button, the results are as shown below
Figure 15. Association Rules Results

16. Result of Association Rules from sales data.
Figure 16. Results of Association Rules from sales data

CONCLUSION

Data visualization, especially in a business context such as sales, is critical for interpreting and understanding complex information. This research emphasizes the advantages of RapidMiner as a powerful data visualization tool with an easy-to-use interface, which makes it possible to gain in-depth understanding of sales datasets. With proper visualization, previously hidden sales patterns such as seasonal trends, bestsellers, and price-volume relationships can be seen. Companies can optimize inventory management, pricing, and marketing strategies with powerful data, enhancing business insights. To stay competitive in the market, data-driven methods must be used as data analysis technology has been incorporated into daily business operations. Overall, results show that the use of RapidMiner exceeds aesthetic standards; it's about turning raw data into ideas for business progress.

REFERENCE


