International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Optimize Retail System Performance by Analyzing Big Data and Visualizing with Power BI

Huu Dang Quoc[1], Hue Man Thi*[2], Mai Nguyen Thi Ngoc[3], Huong Vu Dieu[4]

Submitted: 05/11/2023 **Revised**: 25/12/2023 **Accepted**: 03/01/2024

Abstract: Digital transformation significantly improves business performance and customer experience and adapts to the pace of change in the business environment. Applying digital transformation technologies to business causes the retail industry to gain data very quickly, creating a big data system requiring the ability to perform extensive data analysis and management. Businesses can improve their operational capacity and competitiveness by discovering the data source better. The article presents the method of designing and organizing data and the relationship of data objects in a data warehouse system. Data in the warehouse will be synthesized using computational algorithms to extract useful information and then visualized using data analysis tools, representing data with the output in the reporting charts. They help turn digital data into easy-to-understand visuals, helping retailers see patterns, trends and essential information. Businesses can achieve many meaningful benefits, such as improving operational performance, increasing revenue, predicting trends, etc... by using many charts. Data representation into understandable information is accomplished with the intelligent analysis tool Microsoft PowerBI.

Keywords: *Big data, Visualization reports, Power BI, Digital transformation*

1. Introduction

The rapid development and deployment of new technologies worldwide have im-pacted all sectors, including retail business. This situation requires businesses to change how they do business and interact with customers. Therefore, digital trans-formation in businesses is an inevitable and essential issue in technology. For the retail industry, data changes and arises continuously. The application of technologies helps business operators enhance management and control of business activities more easily through process automation and collection of data. Thanks to that, managers can capture information about revenue, profits, and business performance as quickly and accurately as possible. One tool that meets the above needs is Mi-crosoft Power BI, a data analysis and business reporting tool that helps executives better understand consumer trends and customer needs and make intelligent business decisions. In this article, we have built formulas for calculations such as costs, profits, revenue, etc... and synthesized visualized data by applying different methods based on the Microsoft Power BI tool. With data visualization, reports are easy to under-stand, impressive, use, and analyze data effectively. Data visualization helps identify trends, relationships, and patterns in data. It allows business executives to discover essential information and make the right decisions based on data analytics. The main contribution of this research is that we have proposed a database model to serve research on data analysis in management activities. This helps optimize business operations and comprehensively manage all data. In addition, we also provide some implications for managers in the current era of digital transformation.

The introduction of the research paper clearly states the context, reason for conduct-ing the research, and purpose of the research paper. Part 2 presents previous research and information related to the research content. Part 3 presents the structure of the data, the synthesis of retail data, and the method of synthesizing the data of the retail system. Part 4 presents the results of data analysis through Power BI visual reports and provides some recommendations. Part 5 provides conclusions and future research directions.

2. Literature Review

Organizing, effectively managing, and exploiting big data warehouses is increasingly essential for retail businesses. To solve this problem, businesses need to improve their digital transformation capabilities and apply intelligent analysis technologies to their businesses to increase competitiveness, improve productivity, and improve job per-formance. In the article [1], the authors pointed out the roles of digital transformation in enterprises in Vietnam, in which deploying information technology applications in enterprises significantly affects the effectiveness of corporate governance.

According to [2], the solution that meets the needs of modern organizations is busi-ness intelligence (BI). The

[1] *Thuongmai University, Ha Noi, Viet Nam, email: huudq@tmu.edu.vn*
[2] *Thuongmai University, Ha Noi, Viet Nam, email: hueman@tmu.edu.vn*
[3] *Thuongmai University, Ha Noi, Viet Nam, email: maingocnt@tmu.edu.vn*
[4] *Thuongmai University, Ha Noi, Viet Nam, email: huongvd@tmu.edu.vn*
* *Corresponding Author Email: hueman@tmu.edu.vn*

research findings affirm the pivotal role of Business Intelligence (BI) systems in fostering enterprise progress. These systems notably aid in mak-ing informed decisions, amplifying productivity, and enabling businesses to forecast economic trends effectively" [3-5]. The study conducted by Jasmin Praful Bharadiya [3] highlighted the significant role of both business intelligence and artificial intelli-gence (AI) technologies as crucial pillars of development for businesses. These tech-nologies play a pivotal role in supporting decision-making processes, facilitating accu-rate forecasting, and contributing to the overall corporate economy. The study of Carlo Caserio et al. [4] has combined ERP and BI systems to reduce the load and enhance the value of enterprise data, thereby helping to improve the quality of busi-ness information.

Azma et al. [5] demonstrated that managers have the capability to leverage Business Intelligence (BI) tools to discern customer behavior, conduct cost-benefit analyses, perform market assessments, and predict potential risks within their business opera-tions. Grabiska et al. [6] revealed that the integration of Business Intelligence (BI) in logistics yields several benefits, including cost-effectiveness, optimized routes, and efficient inventory management. Additionally, BI aids in offering effective solutions, which, when of high quality, significantly benefit businesses.

To illustrate the application of Business Intelligence (BI) in data analysis, Microsoft has introduced the Power BI suite - a comprehensive business analysis service equipped with data visualization and BI capabilities for users. Moyano et al. [7] em-ployed Power BI to forecast customer preferences, enabling businesses to promptly make informed decisions. This initiative resulted in the creation of a data mart, aiding in comprehending customer situations and the actual business landscape. Moreover, in [8], the utilization of Power BI as a business intelligence tool enabled the descrip-tive and predictive analysis of demand for GranSol products. Consequently, it facili-tated product distribution to customer branches, assisted in product planning, en-hanced market competitiveness, and provided estimations for future sales. Further-more, Edhya [9] utilized Kimball's ETL (Extract, Transform, Load) 4-step data ware-house modeling method to extract insights from data, designing the necessary struc-ture for the data warehouse.

2.1. Business Intelligence and Microsoft Power BI

Business Intelligence (BI) is a set of methods, processes, architectures, and technolo-gies that transform raw data into meaningful, practical information that can be used to inform strategies. The plan helps the entrepreneur run the business efficiently and make the best decisions. BI provides services to enterprise end users through collabo-rative data exploration between different mathematical models and methodological analysis. BI operations include queries, reporting, OLAP, statistical analysis, forecast-ing, and operational information systems.

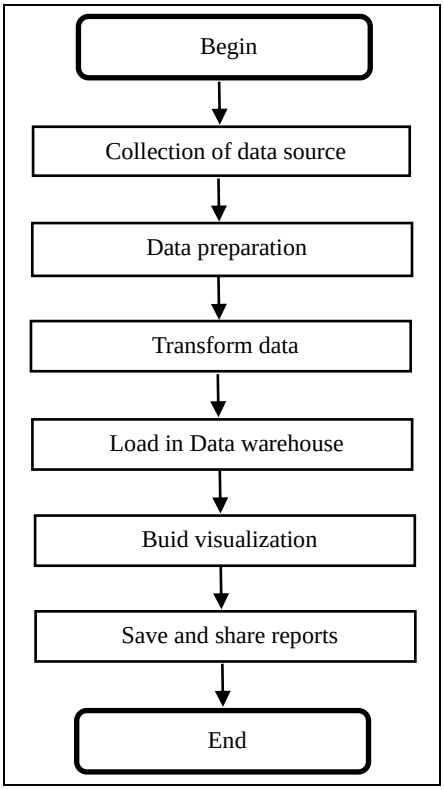

Fig 1. Operation diagram of Power BI

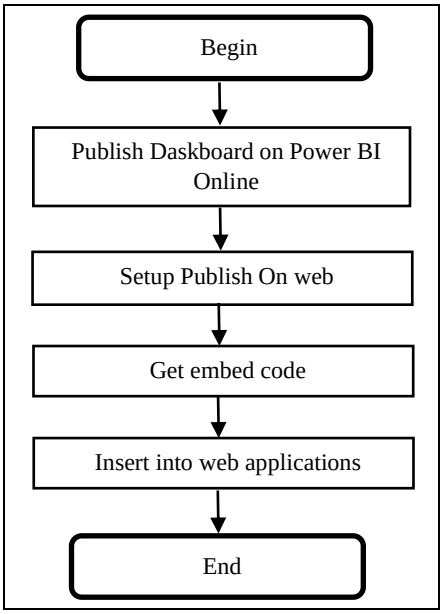

Fig 2. Power BI intergration diagram with other applications

Microsoft Power BI (MPBI) is a business analysis tool provided by Microsoft that is used to visualize reports on demand while allowing users to interact with other data sources. Each other and share the results with other users. MPBP's workflow is implemented as shown in Figure 1. In addition, MPBI can be integrated with web applications to enhance the ability to synthesize, analyze, and visualize

data; this integration process is shown in Figure 2.

2.2. Data warehouse

In the early 1990's, data warehouse was recommended as a general purpose solution to the issue of meeting organisational management information needs. The purpose of a data warehouse is to offer a specialized data source to assist applications used in decision-making [10]. According to Watson, data warehouse is a repository into which are placed all data relevant to the management of an organization and from which emerge the information and knowledge needed to effectively manage the organization [11]. A data warehouse is a location for electronically stored information within an organization; its purpose is to make reporting and analysis easier [12]. Data warehousing, in the context of enterprise data management, is the term to describe a group of decision-making tools designed to enable information specialists (administrators, analysts, or leaders) to make better decisions more quickly [13]. Therefore, data warehouses play a crucial role in facilitating business intelligence, enabling organizations to derive insights, spot trends, make predictions, and ultimately support data-driven decision-making processes across various departments and levels within an organization.

2.3. ETL process

ETL stands for Extract, Transform, Load. It is a process used to extract data from various sources, transform it to fit the structure and format of the target system, and load the transformed data into the target system. The ETL process is commonly used in building and maintaining data warehouses, where data is organized and stored for analysis and reporting purposes. ETL ensures the consistency and accuracy of data within the system. ETL process consists of three steps: extraction, transformation and loading.

Extraction: This is the first step in the process of ETL. Data is extracted from various sources such as databases, files, CRM systems, and other sources. This involves retrieving raw data from disparate sources, which might be structured, semi-structured, or unstructured and ensures that the necessary data is collected and prepared for the subsequent steps in the ETL process.

Transformation: Raw data often needs to be processed and transformed before it's ready for analysis. This phase involves cleaning, restructuring, filtering, aggregating, and converting data into a standardized format or schema suitable for the target system. It defines the grain of fact tables, the dimension tables, data warehouse schema (stare or snowflake), derived facts, slowly changing dimensions, factless fact tables [14].

Loading: Once the data has been extracted and transformed, it's loaded into a target system like a data warehouse, data mart, or database. Loading involves writing the transformed data into the destination in a way that facilitates efficient querying and analysis.

3. Research Methodology

This study used KimBall method to buid the dimensional modeling for designing data warehouse. The method was developed based on observations of practice, and in particular, of data vendors who are in the business of providing data in "user-friendly" form to their customers. In practice, dimensional modeling has become the most popular method for creating data marts and warehouses. It has made a significant contribution to the fields of database architecture and data modeling. This method consists of four steps [15]:

- Select the business process to model
- Declare the grain of the business process
- Choose the dimensions that apply to each fact table row
- Identify the numeric facts that will populate each fact table row

3.1. Select the business process to model

A process is a natural business activity performed in organization that typically is supported by a source data-collection system [15]. The business objective is to increase sales, assess the performance of stores, revenue by product, etc. The business process is a crucial step in designing the data warehouse, aiding in clearly defining the scope and grain of the system, determining data structures, dimensions, and fact tables accurately and in line with business requirements.

In this study, we will select the sales operation process as the primary business focus for investigation. This process significantly impacts the enterprise, providing the most essential and necessary information to help business managers gain the most accurate and objective understanding of the company's sales activities.

3.2. Determine the grain of the business process

In the Ralph Kimball methodology, the grain of a business process refers to specifying the level of detail in the model, which is manifested through two main types of tables: fact tables and dimension tables.

Fact Tables (events, numerical data): These store numerical data related to the chosen business activities. In this study, we utilize CHA store dataset, containing information on revenue, profit, quantity of items sold, etc. There will be a foreign key linking to the dimension tables to define the grain of the data.

Dimension Tables (size, time, location, product): These

tables hold detailed descriptions of various aspects within the Fact tables. Each dimension table focuses on a specific aspect. For instance, a time dimension table may contain details regarding day, month, year, and a detailed invoice table may contain information on total amount, discounts, etc. Dimension tables are linked to fact tables through foreign keys.

The grain of the business process in the Ralph Kimball method is evident through the clear separation between fact and dimension tables. Fact tables contain specific numerical data about business events, while dimension tables provide detailed descriptive information about aspects within the fact tables. This combination facilitates easy data retrieval, analysis, and reporting.

3.3. Choose the dimension that apply to each fact table row

When designing a fact table, each row in the table often relates to various aspects or different details. For example, in a fact table about orders, each row in the fact table will contain information about a specific order. This order may be associated with multiple aspects such as time, customer, product, delivery location, and various other facets.

This step in the Ralph Kimball methodology requires identifying which dimension tables will be linked to each row in the fact table to provide detailed information and descriptions about the event or numerical data recorded in the fact table.

For instance, if each row in the order fact table needs to include information about time, product, and customer, it requires selecting corresponding dimension tables for time, product, and customer to map with each row in the fact table. This helps establish a clear relationship between the rows in the fact table and detailed information from the dimension tables.

In summary, this step in the Ralph Kimball methodology involves deciding which dimension tables each row in the fact table should be linked to in order to supplement detailed information and descriptions for the event or numerical data recorded in the fact table.

3.4. Identify the numeric facts

During the construction of a data warehouse, the fact table contains information about business events or numerical data that an organization wants to track and analyze. This step involves identifying the outcomes of processes in step 1, often a performance metric within sales operations, i.e., valuable numeric information that can be measured, calculated, or statistically significant.

For instance, in a sales order fact table, numerical data might include:

- Quantity of products sold

- Sales revenue (monetary value)

- Applied discount values

- Total invoice value

This step requires the clear identification of these numerical data within the fact table, which is crucial for determining metrics and aiding in the construction of future data analysis and reports.

Once these numerical values are identified, analytical tools, reports, or dashboards can be developed based on this data to support business decision-making and gain a deeper understanding of organizational operations.

Revenue and profit: measure the two criteria CHA store are most interested in revenue and profit. Revenue and profit help enterprise owners know the business situation; based on that, they can adjust these two indicators for the best.

Quantity of available products: To know how large the enterprise's products are, CHA store owners can look at the number of available products to consider in inventory management.

Top-selling products/highest sales: The best-selling product has the most sales in the enterprise; this metric measures the product's popularity and the customer's preference. Based on this index, the enterprise owner can adjust the number of products imported to suit the consumption needs of customers. Unlike a best-selling product, the product with the highest sales usually has a higher selling price or pricing power, providing the most sales in terms of overall revenue and profit.

Top sales staff: This indicates employee performance in the enterprise. Sorting the sales index by top helps enterprise owners have the plan to reward the best employees, creating work motivation.

4. Data Analysis Model and Aggregation Method

The database must have a unified structure, complete and synchronous storage throughout the system to facilitate the organization, synthesis, and analysis of data in the retail system.

4.1. Data analysis model

With the retail system, we proposed 7 key entities to store in the database. These are important tables to help managers control information quickly and effectively.

Table 1. Table describing entities to store retail data

No.	Entity	Description
1	tblLocal	Information about the customers'address
2	tblCustomer	All detail information about customers
3	tblUser	Details of the store's staff
4	tblProduct	Store products information
5	tblBill	Include information about the store's import and export activities according to the invoice money
6	tblBillDetail	Detail information about bills and products
7	tblIncomeExpenses	Include store's revenue and expenses information

The data tables are related to each other to meet the constraints of the retail data system. Figure 3 shows the binding between data objects.

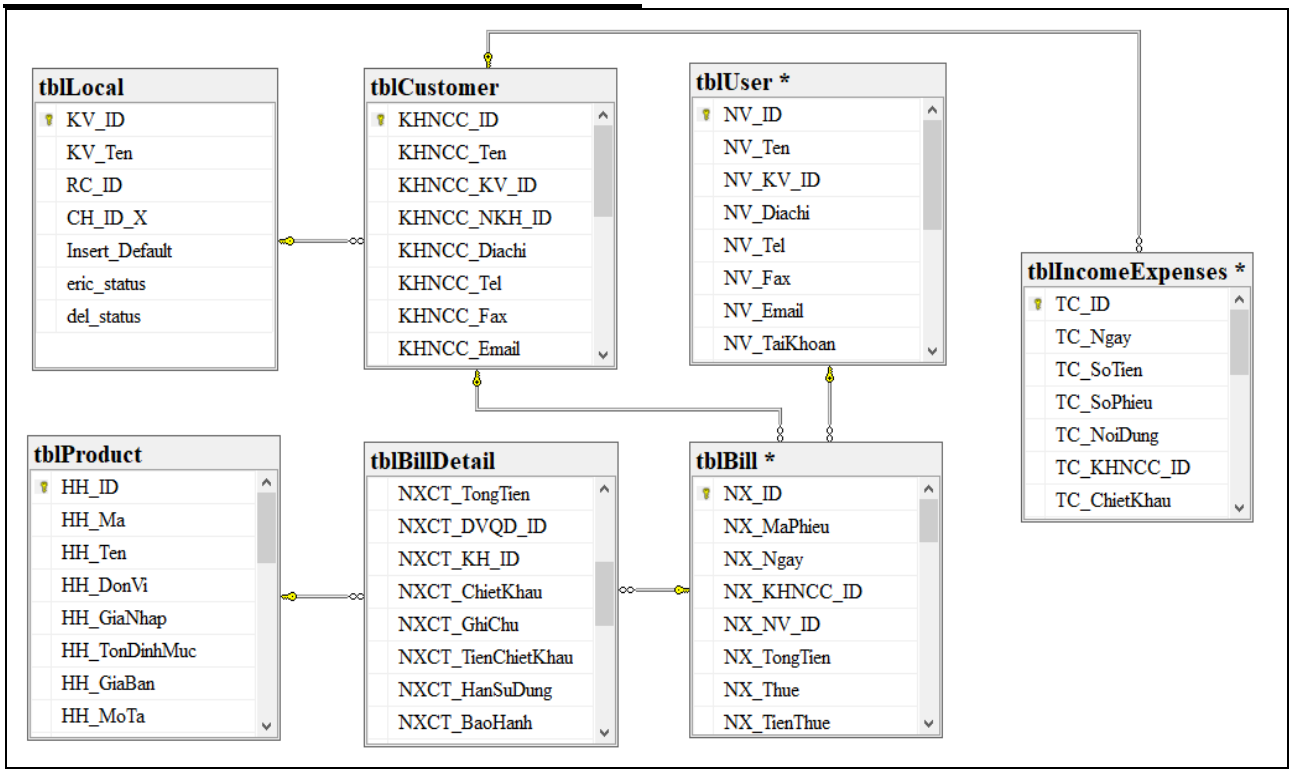

Fig 3. Relationship between data tables

4.2. Formulas for calculating and synthesizing data

In this study, we provided some calculation formulas to clarify the results calculated and used in Power BI. Table 2 presents the formulas for data aggregation in the retail system.

Table 2. The formulas

Formulas	Description
$S^k = \sum_{i=0}^{r} S_i$	Calculating the total product of a bill
$M^k = \sum_{i=0}^{r} M_i$	Calculating the total money of a bill
$T^k = \sum_{i=0}^{r} S_i . P_i . T_i \ \forall T_i: \in \{0.08;0.1\}$	Calculating the amount of tax on a bill, with a tax rate on each bill is 8% or 10%
$D^k = \sum_{i=0}^{r} S_i . P_i . D_i \ \forall D_i \in \mathbb{R}$	Discount money for a bill
$R^k = \sum_{i=n_{1,l}}^{n_2} M_l^i + T_l^i - D_l^i \ \forall l=20$	Caculating the revenue total between 2 dates
$A = \frac{\sum_{i=n_{1,l}}^{n_2} M^i}{\sum_{i=n_{1,l}}^{n_2} S^i} \ \forall l=10$	Calculating average product import price
$V = \sum_{k=n_{1,l}}^{n_2} \sum_{i=1}^{r} A_i . s_i^k \ \forall l=20$	Caculating total import money for the export products
N = R-V	Calculating profit

Where:

S^k: The total products of the k bill; S_i: The quantity of product in the i record in a bill; M^k: Total money of the k bill; M_i: The total money of a record i in a bill; P_i: The product price of the record i in a bill; T^k: Tax of the k bill; T_i: Tax at the record i in a bill; D: Discount; D_i: Discount at the record i in a bill; R: Revenue total; l: Import-export type, l = 10: import, l = 20: export; A: Average import price of product; r: Number of record in a bill; n1, n2: Caculating value from date 1 to date 2; N: Profit.

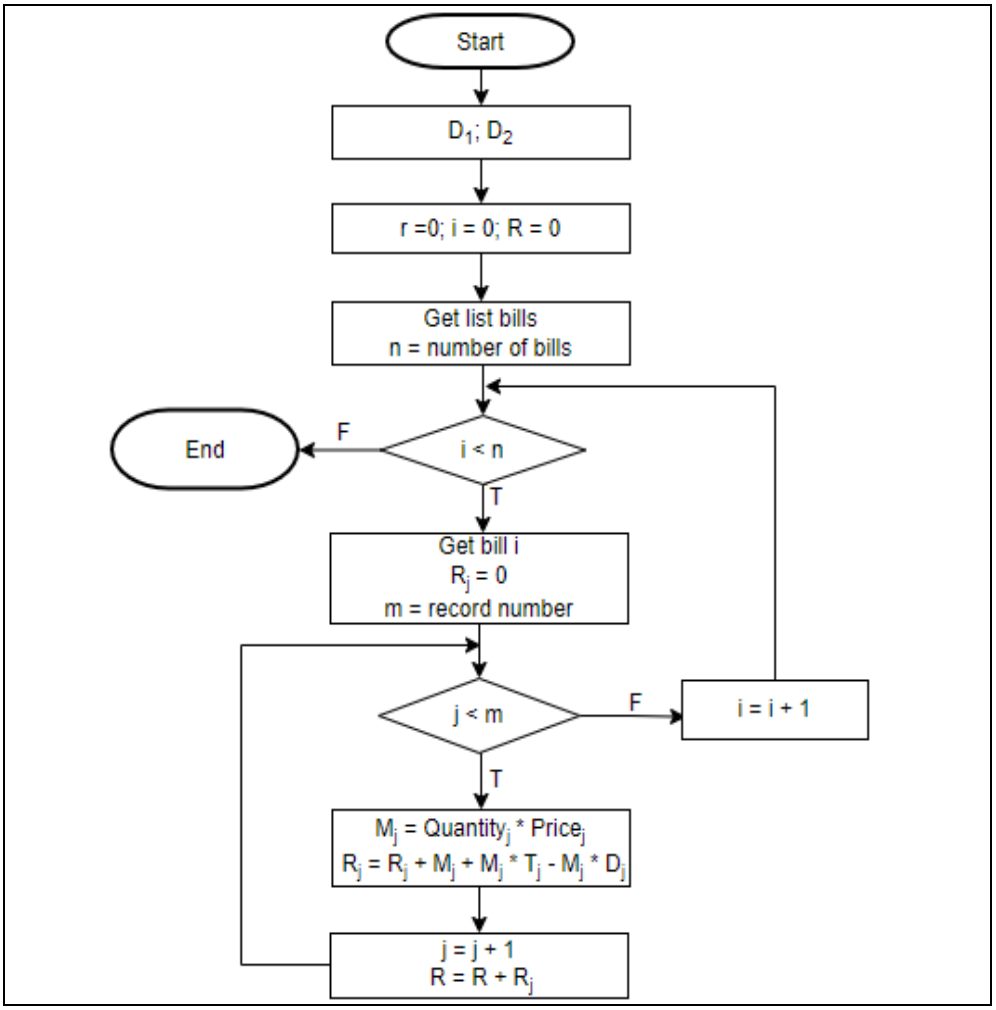

Fig 4. Revenue calculation algorithm

Figure 4 shows the revenue calculation algorithm diagram from day 1 (D1) to day 2 (D2). This revenue calculation algorithm is implemented through Power BI to automate and optimize the revenue calculation process.

4.3. The experimental dataset

This research article took data from CHA retail store. CHA store is one of the biggest retail stores in Hanoi - the capital of Vietnam. We collected store's sales data in 2022 on profits, revenue, regions, customers, etc. for this research.

Table 3. CHA's revenue dataset in 2022

Name	Month	Year	Revenue	%/year
RV.01.22	1	2022	11,946,385,276	9.15
RV.02.22	2	2022	7,008,843,219	5.37
RV.03.22	3	2022	8,399,197,167	6.44
RV.04.22	4	2022	7,709,974,898	5.91
RV.05.22	5	2022	7,958,309,811	6.10
RV.06.22	6	2022	7,024,551,307	5.38
RV.07.22	7	2022	9,048,200,252	6.93
RV.08.22	8	2022	11,129,417,672	8.53
RV.09.22	9	2022	11,698,859,021	8.97
RV.10.22	10	2022	12,804,057,253	9.81
RV.11.22	11	2022	15,804,231,316	12.11
RV.12.22	12	2022	19,961,895,533	15.30
	Total:		130,493,922,726	100.00

Table 4. CHA's profit dataset in 2022

Name	Month	Year	Profit	%/year
PR.01.22	1	2022	908,244,516	7.88
PR.02.22	2	2022	647,397,569	5.61
PR.03.22	3	2022	680,169,674	5.9
PR.04.22	4	2022	711,341,568	6.17
PR.05.22	5	2022	607,459,879	5.27
PR.06.22	6	2022	579,181,428	5.02
PR.07.22	7	2022	803,673,110	6.97
PR.08.22	8	2022	998,767,447	8.66
PR.09.22	9	2022	978,585,194	8.49
PR .10.22	10	2022	1,222,163,398	10.6
PR.11.22	11	2022	1,459,172,880	12.65
PR.12.22	12	2022	1,936,661,083	16.79
	Total:		11,532,817,746	100.00

5. The Result of Visualization and Implicants

After the data is aggregated and computed, it will be visualized using the Power BI data analysis tool. Viewing the real-time dashboard, managers can easily grasp the real-time business situation of the enterprise, enabling them to make precise and timely management decisions. Within the dashboard, all detailed charts are presented in the most visual report format. Users can access detailed figures and reports from this screen. Figure 4 illustrates an overall performance report on the retail situation of CHA store.

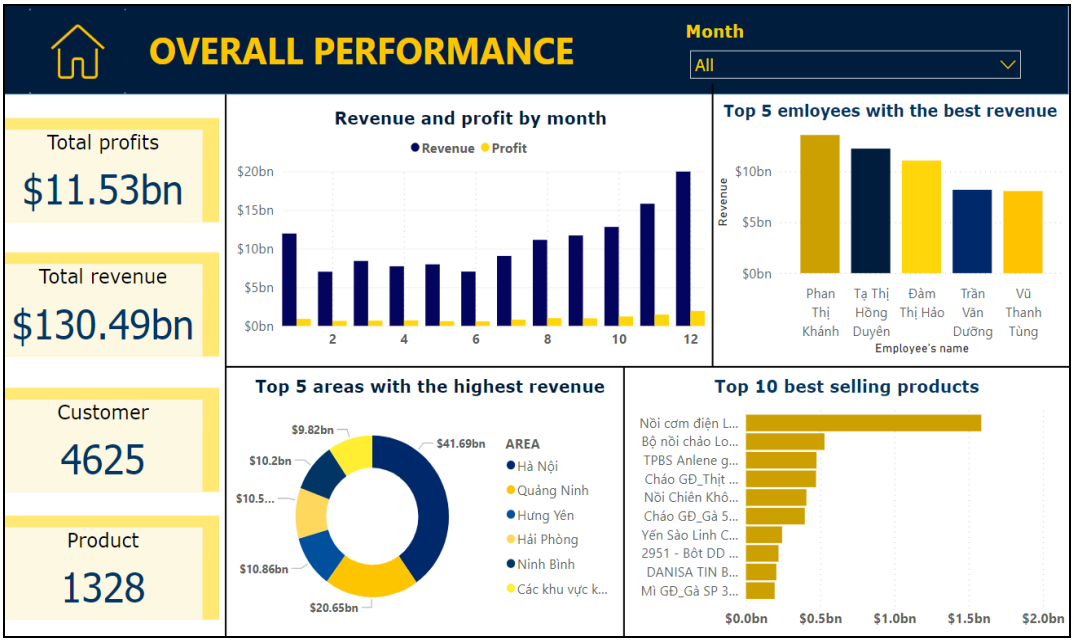

Fig 5. Power BI dashboard overview

The dashboard in 2022 provided a comprehensive and detailed view of the business performance of CHA store. By focusing on crucial metrics such as revenue, profit, customer count, and real-time updates, the dashboard offered an in-depth insight into the developmental trends of CHA store.

The reports on top-performing employees, regions, or products concentrated on the most significant factors influencing revenue. This allowed CHA store to identify decisive elements contributing to its business success. Applying a development strategy based on insights from this dashboard will serve as the foundation for optimizing business operations and shaping a sustainable long-term business strategy.

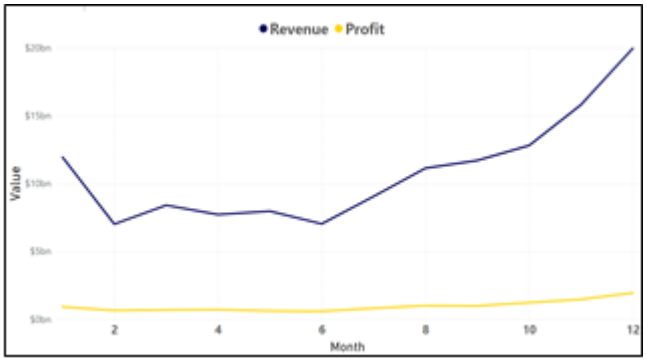

Fig 6. Revenue and profit by month

The report illustrates the revenue and profit for each month in 2022. Towards the latter part of the year, from October onwards, there was a significant surge in revenue for CHA store, peaking in December with nearly 20 billion in revenue and 1.9 billion in profit. There are notable variations in revenue among months, notably between February and June, where CHA store's revenue remained relatively low.

Market analysis indicates that the year-end months typically experience a peak period for all businesses, often marked by significant year-end shopping events and holidays. CHA store strategically amplifies its marketing efforts, advertising, and promotions in December to stimulate sales figures.

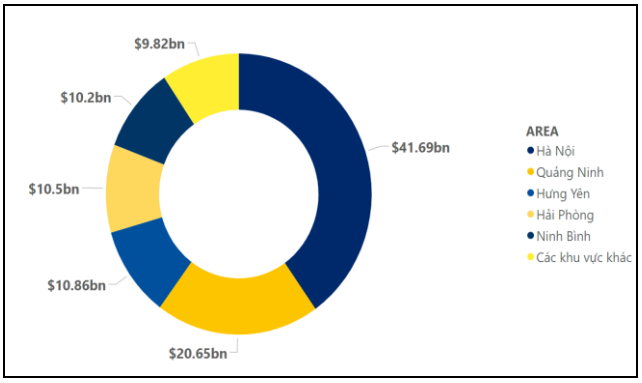

Fig 7. Top 5 areas with the highest revenue

The chart depicting the top 5 regions with the highest revenue illustrates crucial developmental trends. These areas represent targeted markets for CHA store, demonstrating substantial revenue-generating potential. Among them, Hanoi emerges as the primary market for CHA, accounting for approximately 32% of the total regional revenue. Multiple factors contribute to Hanoi's status as the most promising market for the store. The store's headquarters are situated in Hanoi, the capital city of Vietnam, which boasts a sizable population and a high standard of living. This chart assists CHA in focusing efforts on optimizing strategies to maximize the profitability potential of these regions.

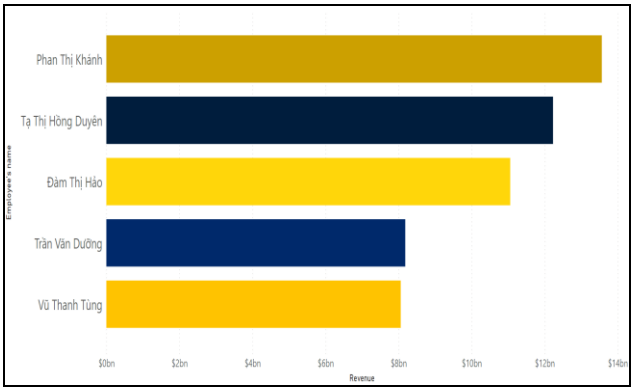

Fig 8. Top 5 employees with the highest revenue

The chart depicting the top 5 employees with the highest revenue provides a detailed insight into individual performance at CHA store. These top-performing employees significantly contribute to the store's revenue stream. According to the report, Phan Thi Khanh stands out as the employee bringing in the highest revenue for the store with over 13 billion VND, followed by Ta Thi Hong Duyen and others.

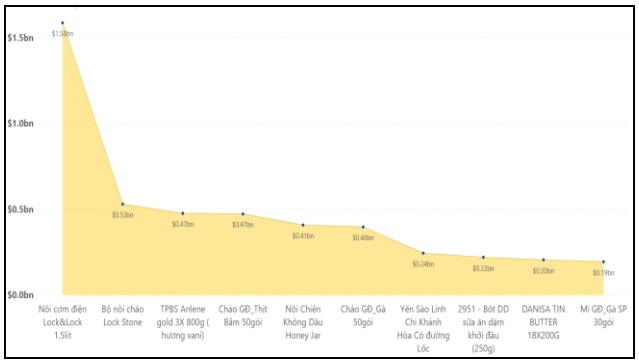

Fig 9. Top 10 best selling product

The top 10 best-selling products chart is a pivotal report that helps CHA store gain insights into the popularity and potential of each product within their business portfolio. This provides the basis for evaluating sales performance, shaping marketing strategies, and optimizing inventory management to capitalize on these high-performing products. Among these products, the "Noi com dien

Lock&Lock 1.5L" stands out as the highest-selling item, significantly surpassing other products with sales revenue exceeding 1.5 billion.

The application of MPBI brings a lot of efficiency to management, helping to capture situations and analyze data quickly and accurately. From this result, the article has the following implicants for managers to improve management efficiency:

The first, promoting digital transformation in enterprises, synchronizing data and data structures on retail activities need to be standardized according to clear structures, serving for exploitation and easy integration compatible with other applications.

The second, deploying the application of business intelligence management tools to visualize data, helping managers to quickly and accurately observe information about their retail business activities. From there, help managers make timely and accurate business management decisions.

The third, enterprises should learn and use a tool to help visualize data to facilitate analysis and observation. MPBI is a powerful tool that serves its purpose in this destination.

6. Conclusion and Future Works

The article presented how to apply data visualization and analysis tools in intelligent data analysis to enhance the management capabilities of enterprises, thereby permit-ting them to improve revenue, profit, competitiveness, and efficiency in operations. The article also proposes a data structure to store retail data in a standardized and unified manner and then proposes calculation formulas to aggregate the data. The Microsoft Power BI tool aggregates actual retail data to produce visual reports. From the visualization reports, the article also makes some recommendations on the man-agement.

In the coming time, the author will study some applied algorithms for analyzing and predicting business performance in retail systems to support businesses better.

Funding

This research is funded by Thuongmai University, Hanoi, Vietnam.

References

[1] Le Viet, H.; Dang Quoc, H., The Factors Affecting Digital Transformation in Vietnam Logistics Enterprises, *Electronics 2023*, 12, 1825. DOI: 10.3390/electronics 12081825.

[2] Olszak, CM (2022). Business intelligence systems for innovative development of organizations. *Procedia Computer Science*, 207, 1754-1762. DOI: 10.1016/j.procs.2022.09.233.

[3] Bharadiya, JP (2023). A Comparative Study of Business Intelligence and Artificial Intelligence with Big Data Analytics. American Journal of Artificial Intelligence, 7 (1), 24. DOI: 10.11648/j.ajai.20230701.14.

[4] Caserio, C., Trucco, S., Caserio, C., & Trucco, S. (2018). Business intelligence systems. Enterprise Resource Planning and Business Intelligence Systems for Information Quality: An Empirical Analysis in the Italian Setting, 43-73. DOI: 10.1007/978-3-319-77679-8_3.

[5] Azma, F., & Mostafapour, MA (2012). Business intelligence as a key strategy for development organisations. Procedia Technology, 1, 102-106. DOI: 10.1016/j.protcy. 2012.02.2020.

[6] Grabińska, A., & Ziora, L. (2019). The application of Business Intelligence systems in logistics. review of selected practical examples. System Safety: Human-Technical Facility-Environment, 1 (1). DOI: 10.2478/czoto-2019-0130.

[7] Moyano, Diego Marcelo Bermeo, and Milton Alfredo Campoverde Molina, Implementación de Data Mart, en Power BI, para el análisis de ventas a clientes, en los Econegocios “Gransol”, Polo del Conocimiento: Revista científico-profesional 5.1 (2020): 647: -673.

[8] Bermeo-Pérez, S. K., & Campoverde-Molina, M. A. (2020). Implementación de inteligencia de negocios, en el inventario de la Cooperativa GranSol, con la herramienta Power BI. Revista Científica FIPCAEC (Fomento de la investigación y publicación científico-técnica multidisciplinaria). ISSN: 2588-090X. Polo de Capacitación, Investigación y Publicación (POCAIP), 5(16), 240-266.

[9] Edhya, Billah Faktha Putra, Business intelligencce data marketing mengunakan metodekimball dan ETL dengan Power B, Kurawal-Jurnal Teknologi, Informasi dan Industri 5.2 (2022): 87-97, DOI: 10.33479/kurawal.v5i2.642.

[10] Gray, P. and H.J. Watson (1998) Decision Support in the Data Warehouse, Upper Saddle River, New Jersey, Prentice-Hall. DOI: 10.5555/572865.

[11] Watson, H. (2001). Recent Developments in Data Warehousing. Association for Information Systems “AMCIS 2001 Proceedings”, 2289-2292. DOI: 10.17705/1CAIS.00801.

[12] Nambiar, A., & Mundra, D. (2022). An Overview of Data Warehouse and Data Lake in Modern Enterprise Data Management. *Big Data and Cognitive Computing*, *6*(4), 132. DOI: 10.3390/bdcc6040132.

[13] *Inmon, W. H. (1995). What is a data warehouse. Prism Tech Topic, 1(1), 1-5*.

[14] El-Sappagh, S. H. A., Hendawi, A. M. A., & El Bastawissy, A. H. (2011). A pro-posed model for data warehouse ETL processes. Journal of King Saud University-Computer and Information Sciences, 23(2), 91-104. DOI: 10.1016/j.jksuci.2011.05.005.

[15] Kimball, R., & Ross, M. (2011). The data warehouse toolkit: the complete guide to dimensional modeling. John Wiley & Sons.