



Optimizing Decision Making in MSMEs through Business Intelligence Dashboards using Python and Power BI

Azka Raisa Subagio

Faculty of Science and Technology, Informatics Study Program, UIN Sunan Kalijaga, Yogyakarta, Indonesia

Email: raisazkas25@gmail.com

Abstract—Micro, Small, and Medium Enterprises in Indonesia play a vital role in national economic growth; however, many continue to rely on manual spreadsheet-based reporting and intuitive judgment, limiting the effectiveness and timeliness of data-driven decision making. This study aims to examine how Business Intelligence dashboards integrating Python and Power BI can enhance operational decision-making performance in Indonesian retail-sector micro, small, and medium enterprises. Using a quantitative descriptive approach, the study analyzes secondary data from the Grocery Store Sales Dataset (2025) obtained from the Kaggle open-source platform. A total of 1,980 transaction records were processed to simulate typical operational decision-making scenarios commonly faced by retail enterprises. In the baseline condition, decision making was conducted using conventional spreadsheet summaries without automated analytics or real-time visualization. Python was employed for data preprocessing, transformation, and key performance indicator computation, while Power BI was used to develop an interactive Business Intelligence dashboard. Descriptive statistical analysis and scenario-based simulations were conducted to compare decision-making efficiency and accuracy before and after dashboard implementation. The results indicate that the proposed Business Intelligence approach reduced average decision-making time by 36.36 percent, improved information accuracy by 41.18 percent, and accelerated strategic planning speed by 40 percent. These findings demonstrate that integrating Python-based analytics with Business Intelligence dashboards offers a low-cost, scalable, and effective solution to support data-driven managerial practices and strengthen the digital readiness of Indonesian micro, small, and medium enterprises.

Keywords: Business Intelligence; Python; Power BI; Decision Making; MSMEs; Data Analytics

1. INTRODUCTION

In the era of digital transformation, the ability of businesses to manage, analyze, and utilize data effectively has become a critical determinant of organizational competitiveness and sustainability. Data-driven decision making enables organizations to respond more rapidly to market dynamics, optimize operational processes, and improve strategic planning through evidence-based insights rather than intuition alone (Kenneth C. Laudon, 2022). For Micro, Small, and Medium Enterprises (MSMEs), this capability is particularly crucial, as they often operate in highly competitive environments with limited margins, constrained resources, and increasing exposure to digital marketplaces.

In Indonesia, MSMEs represent the backbone of the national economy. According to official government statistics, MSMEs contribute more than 60 percent of national gross domestic product, employ over 97 percent of the workforce, and account for more than 64 million active business units across various sectors (Affairs, 2025). Despite their economic significance, many Indonesian MSMEs continue to face persistent challenges in managing business data and adopting data-driven managerial practices. Operational, sales, inventory, and financial data are frequently generated from disparate sources such as point-of-sale systems, online marketplaces, mobile cashier applications, and spreadsheets. However, these data sources are rarely integrated into a unified analytical system, resulting in fragmented information that limits analytical depth and strategic insight (Rizi et al., 2023).

Although many MSMEs have adopted basic digital tools, such as accounting software or e-commerce platforms, data utilization remains largely descriptive and retrospective. Decision making is often reactive, relying on personal experience, intuition, or incomplete summaries rather than systematic analysis of trends and performance indicators. Fragmented data structures further exacerbate this problem, as inconsistencies in data formats, missing values, and limited historical records restrict the applicability of advanced analytical techniques. As emphasized in information systems theory, competitive advantage in the digital economy depends not merely on data availability but on the ability to transform raw data into actionable knowledge that supports timely and informed decisions (Kenneth C. Laudon, 2022).

The practical challenge faced by MSMEs, therefore, lies in how to transform heterogeneous and fragmented data into meaningful information that is accessible and understandable to non-technical decision makers. Without a structured Business Intelligence (BI) system, managerial decisions tend to be based on partial information and short-term observations rather than integrated, data-driven evaluation. Importantly, while BI solutions often assume standardized and well-structured datasets, MSMEs typically operate with limited, noisy, and inconsistent data. This research explicitly addresses this challenge by employing a preprocessing and transformation pipeline that consolidates fragmented transactional data into standardized analytical structures, enabling meaningful KPI calculation and visualization despite data limitations.

Against this backdrop, the central research question of this study is formulated as follows: *How can a Business Intelligence dashboard that integrates Python-based data analysis and Power BI visualization be designed and implemented to improve decision-making efficiency and accuracy in MSMEs?* In addition, this study investigates the extent to which such a dashboard improves managerial performance indicators, including decision speed, information accuracy, and strategic responsiveness, compared to conventional spreadsheet-based decision-making practices.



In the last decade, a number of studies have shown the relevance and benefits of BI in organizations and small and medium-sized enterprises. Gonçalves examined the application of BI dashboards in a marketing context and found that interactive dashboards simplify the interpretation of business metrics and accelerate decision-making (Gonçalves et al., 2023). Sanabia-Lizarraga designed a Power BI dashboard to analyze agricultural exports, demonstrating that visualization tools can make it easier for non-technical users to read complex data (Sanabia-Lizarraga et al., 2024). Research on real-time BI dashboard optimization by Ogunmokun proposes a model that combines real-time data processing and accuracy prediction to improve decision speed and reliability (Ogunmokun et al., 2024).

Meanwhile, Ragazou developed a BI model to empower small businesses through the integration of BI systems and business strategies, especially in the context of MSME digital transformation (Ragazou et al., 2023). Furthermore, in studies conducted in the African context, research shows that BI is an important component in timely decision-making and operational excellence in the MSME sector (Fabian et al., 2024). Complementing this view, studies from Vietnam revealed that BI and big-data adoption in SMEs positively influence performance-based decision-making and market responsiveness (Nguyen et al., 2022; Truong, 2022). Within the Indonesian MSME ecosystem, empirical evidence also supports the growing relevance of BI. Anjaningrum found that adopting business intelligence systems among creative-economy MSMEs in East Java improves business performance through better access to high-quality information and data-driven insights (Anjaningrum et al., 2024). Similarly, Iskandar confirmed a positive correlation between technology adoption and business performance, underscoring the need for accessible analytical tools for non-technical users in small enterprises (Yusuf Iskandar et al., 2023).

From a technical perspective, recent studies have explored the integration of Python scripting within Power BI environments to support advanced analytics. For example, Siljanoska demonstrated that embedding Python code into Power BI workflows allows real-time data processing, custom KPI calculations, and predictive analytics — enabling interactive dashboards that go beyond static reporting (Siljanoska et al., n.d.). This hybrid integration approach provides a practical framework for MSMEs seeking to combine flexibility, scalability, and ease of use in a single BI environment.

Despite these advances, a notable research gap remains. Existing BI studies predominantly focus on large organizations with standardized internal datasets, while empirical research addressing MSMEs with limited and fragmented data is relatively scarce. Furthermore, few studies explicitly combine Python-based data preprocessing and KPI computation with interactive Power BI dashboards tailored to MSME operational contexts, particularly in Indonesia. This study addresses this gap by proposing a practical BI integration framework designed to handle fragmented transactional data, limited historical records, and heterogeneous data sources commonly encountered by MSMEs.

The primary objective of this research is to design, implement, and evaluate a BI dashboard that enhances decision-making efficiency in MSMEs. Specifically, the study aims to: (1) develop a Python-based data analysis pipeline for cleaning, transforming, and standardizing fragmented transactional data; (2) integrate the analytical outputs into an interactive Power BI dashboard; and (3) evaluate the impact of dashboard implementation on decision-making efficiency and information accuracy through scenario-based simulations.

The solution offered in this study is a BI dashboard prototype that combines the power of Python analysis (flexibility, data manipulation capabilities, simple prediction algorithms) and the power of Power BI visual displays (interactive flexibility, accessibility for non-technical users). With this approach, MSME users not only see static visual dashboards, but can also access specific metrics, projected trends, and automated reports relevant to their operational needs.

The contribution of this research is twofold. From a technical perspective, the study proposes an integrated Python–Power BI framework tailored to MSMEs with limited and fragmented data, introducing a flexible and customizable KPI analysis approach. From a managerial perspective, the findings demonstrate how accessible BI dashboards can support faster, more accurate, and more responsive decision making among MSME managers. By addressing both technical and practical constraints, this research contributes to the broader discourse on digital transformation and provides actionable guidance for MSMEs seeking to adopt data-driven decision-making practices in resource-constrained environments.

2. RESEARCH METHODOLOGY

2.1 Basic Research Framework

This study adopts a quantitative-descriptive research design supported by secondary data analysis to examine how a Business Intelligence (BI) dashboard integrating Python analytics and Power BI visualization can enhance decision-making performance in Micro, Small, and Medium Enterprises (MSMEs). Quantitative-descriptive methods are appropriate for BI-oriented research as they allow systematic analysis, visualization, and interpretation of transactional data to support managerial insights without relying on complex causal inference models (Hair et al., 2019; Kenneth C. Laudon, 2022).

The research framework consists of sequential stages, including data acquisition, preprocessing, exploratory data analysis (EDA), key performance indicator (KPI) computation, dashboard development, and decision-making

evaluation. This framework reflects standard BI pipelines commonly applied in applied analytics and decision support system research (Chen & Storey, 2026).

This study involves two primary variables. The independent variable is the implementation of a Python-based Business Intelligence (BI) dashboard integrated with Power BI, which encompasses aspects of data processing efficiency, visualization accuracy, and multi-source data integration. The dependent variable is the quality of managerial decision making, which is evaluated through indicators such as decision time efficiency, information accuracy, and strategic planning accuracy. Based on the relationship between these variables, this study hypothesizes that the integration of Python and Power BI significantly improves the efficiency and accuracy of decision making in Micro, Small, and Medium Enterprises (MSMEs). Conversely, the null hypothesis assumes that the integration of Python and Power BI does not have a significant effect on the efficiency and accuracy of managerial decision making in MSMEs.

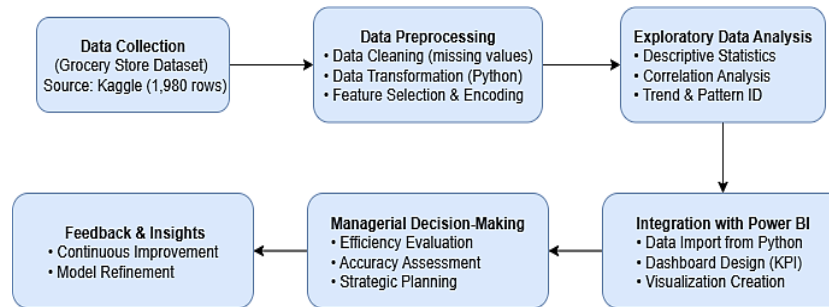


Figure 1. Basic Research Framework of Business Intelligence

The basic framework of this research is illustrated in Figure 1, which describes the stages from data collection, data cleaning, and exploratory data analysis (EDA) using Python, integration of analysis results into Power BI, to evaluation of decision-making effectiveness. The analysis was carried out through the stages of data preprocessing, calculation of key performance indicators (KPIs), and creation of an interactive dashboard in Power BI to support data-driven managerial decisions.

Quantitative analysis was supported by descriptive statistical methods and simple correlation tests using Python (pandas, numpy, matplotlib). This approach allowed researchers to measure the relationship between BI Dashboard usage and decision efficiency numerically without involving complex inferential tests.

2.2 Data Source and Contextualization Strategy

The dataset used in this study is the Grocery Store Sales Dataset (2025) obtained from the Kaggle open-source platform. The dataset contains 1,980 transaction records representing retail sales activities, including product categories, pricing, promotional indicators, customer interactions, and temporal sales patterns. Public datasets from Kaggle have been widely used in business analytics research as proxies for real-world operational scenarios when direct access to proprietary organizational data is limited (Knafllic et al., 2020).

However, the use of an international public dataset presents inherent limitations, particularly potential biases related to product mix, pricing structures, consumer behavior, and market dynamics that may not fully reflect the characteristics of Indonesian MSMEs. To address this methodological limitation, this study employs a contextual simulation strategy, in which the dataset is not treated as a direct representation of Indonesian MSMEs but rather as a decision-making simulation environment. The analysis focuses on operational patterns (sales frequency, category performance, promotional response, and inventory turnover) that are structurally comparable to Indonesian retail MSMEs such as grocery stores, minimarkets, and neighborhood retailers.

Furthermore, KPI selection, interpretation, and dashboard design are explicitly aligned with commonly used indicators in Indonesian MSMEs, as reported in national MSME performance guidelines and government reports (Affairs, 2025). This approach follows prior BI research that emphasizes contextual alignment rather than direct data generalization when public or synthetic datasets are used (Gonçalves et al., 2023).

2.3 Research Stages

The research stages were arranged systematically to ensure that the results obtained were in line with the objectives and hypotheses that had been set. The sequence of stages is explained as follows:

1. Problem identification and literature review

The initial stage includes a review of various previous studies on the application of Business Intelligence (BI) in the retail and SME sectors, as well as the use of Python and Power BI in sales data analysis (Baig Mirza et al., 2025; Gonçalves et al., 2023; Ruansyah et al., 2024). A comprehensive review by Raguseo and Vitari highlights that research on Business Intelligence remains largely focused on adoption outcomes, implementation challenges, and performance impacts, while empirical studies addressing resource constraints, data fragmentation, and analytical capability limitations in small and medium-sized enterprises remain limited (Raguseo, E., & Vitari, 2020). The

review emphasizes that most SMEs experience structural and resource constraints that limit the full realization of BI's potential in improving data-driven decision-making.

Further evidence comes from Mahmud and Ikbal in *Power BI and Data Analytics in Financial Reporting: A review of Real-Time Dashboarding and Predictive Business Intelligence Tools*, which investigated how combining analytics techniques and BI systems enhances operational efficiency, customer retention, and inventory management (Mahmud & Ikbal, 2024). Their findings highlight that the main barriers to BI adoption in financial reporting remain data integration and implementation costs—issues particularly relevant to developing economies such as Indonesia.

Complementing these findings, practical guidelines such as *The Ultimate Microsoft Power BI Guide for SMEs* explain how small enterprises can leverage Power BI to convert fragmented data into meaningful visual insights (Square, 2025). The report emphasizes the advantages of integrating multiple data sources and creating intuitive dashboards for non-technical users, enabling small-business owners to monitor their key metrics effectively.

2. Secondary data collection

The data was obtained from the public platform Kaggle, namely the Grocery Store Sales Dataset (2025), which contains 1,980 transaction records from a grocery store chain in various locations between August 2023 and August 2025. This dataset includes customer information, product categories, store locations, customer interaction and promotional activities, which are relevant for simulating the context of retail MSMEs in Indonesia.

3. Data cleaning and transformation (Data Preprocessing)

At this stage, data cleaning is performed to remove missing values, duplicates, and format inconsistencies. This process uses Python libraries such as pandas and numpy for date format normalization, data type conversion, and feature selection according to analysis requirements.

4. Exploratory analysis and BI modeling

This stage aims to identify sales patterns, product trends, and promotional effectiveness. The analysis is conducted using Python through an Exploratory Data Analysis (EDA) approach, including the calculation of initial metrics such as sales growth rate and average basket value. The results of the analysis form the basis for the design of key performance indicators (KPIs).

5. Integration with Power BI

Data from Python analysis is exported to Power BI for the creation of interactive dashboards. The dashboards display various visuals such as sales trends, category performance, customer segmentation, and store comparisons, which aid in quick and visual decision making.

6. Evaluation of dashboard effectiveness

Testing was conducted through what-if analysis simulations in Power BI to assess the impact of dashboard usage on time efficiency and decision accuracy. The simulation results were compared with the initial conditions prior to BI Dashboard implementation to measure improvements in managerial effectiveness.

7. Analysis of results and interpretation

The visualization and simulation results were analyzed descriptively to assess improvements in decision efficiency, information accuracy, and the predictive capabilities of dashboards in supporting MSME business strategies.

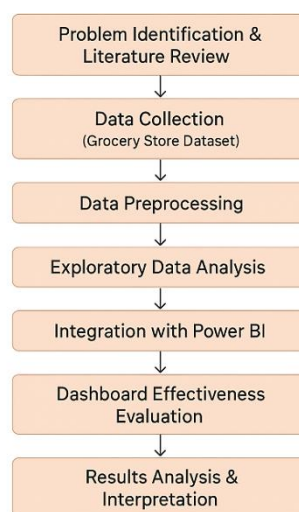


Figure 2. Research Implementation Stages for BI Dashboard Development

2.4 Analytical Process and Evaluation

The analytical process in this study was conducted through several systematic stages. First, data cleaning and transformation were performed using Python to remove duplicate records, handle missing values through the mean imputation method, and normalize date formats. The cleaned and standardized dataset was then saved in CSV format to ensure compatibility and seamless integration with Power BI.



Following this step, exploratory data analysis (EDA) was carried out to identify monthly sales trends, best-selling product categories, and the contribution of key customers. Descriptive statistical analysis and simple correlation techniques were applied to examine relationships between key variables, such as the association between promotional activities and sales volume, providing an initial understanding of sales patterns and customer behavior.

Subsequently, key performance indicators (KPIs) were computed to evaluate operational and financial performance. The KPIs calculated in this study include sales growth rate, profit margin, inventory turnover, and customer retention rate. The sales growth rate was calculated using the formula: $Sales = \frac{Sales_t - Sales_{t-1}}{Sales_{t-1}}$, which measures the change in sales from one period to the next and reflects the effectiveness of MSME sales strategies over time.

The processed data and computed KPIs were then utilized to develop interactive dashboards in Power BI. The dashboards consist of four integrated analytical views, including an overview dashboard presenting total sales, revenue, and key performance indicators, a sales trends view illustrating sales dynamics across time and store locations, a customer insights view displaying customer segmentation and purchasing behavior, and a forecasting panel that presents projected sales trends using Power BI's built-in forecasting feature.

2.4.1 Research Validation

Research validation in this study was conducted using two complementary approaches. Internal validation was carried out by examining the consistency of visualization outputs and verifying the compatibility between computed metrics and their representations in Power BI, ensuring that each key performance indicator accurately reflects the underlying source data. In addition, external validation was performed by comparing the patterns and trends generated by the dashboard with national MSME data reports published by the Ministry of Cooperatives and SMEs (2025), in order to assess the relevance of the analytical interpretations and to ensure contextual accuracy within the Indonesian MSME environment.

2.5 Summary of Research Design

A summary of the research design is presented in the following table as a systematic guide for conducting the research.

Table 1. Summary of Research Design and Analytical Techniques

Research Type	Quantitative-descriptive
Data Source	Open-source (Kaggle)
Software Tools	Python (pandas, numpy, matplotlib), Power BI
Analysis Techniques	Data preprocessing, EDA, KPI computation, visualization
Evaluation Metrics	Efficiency gain (%), Accuracy gain (%)

Table 1 presents a concise summary of the research design and analytical techniques applied in this study. As shown in the table, the research adopts a quantitative-descriptive approach using open-source data from Kaggle, with Python employed for data preprocessing, exploratory data analysis, and KPI computation, and Power BI used for interactive dashboard development. The analytical process follows a standard Business Intelligence pipeline, and the effectiveness of the proposed framework is evaluated using efficiency gain (%) and accuracy gain (%) derived from comparative decision-making simulations.

3. RESULTS AND DISCUSSION

This section presents the results of data processing and visualization using Python and Power BI, followed by a comprehensive discussion of its implications for managerial decision-making in MSMEs. This presentation is divided into four sections: (1) data cleaning and preprocessing results, (2) findings from exploratory data analysis (EDA), (3) dashboard development and interpretation, and (4) a comparative discussion with previous studies and hypothesis verification.

3.1 Result

3.1.1 Data Cleaning and Preprocessing

The original Grocery Store Sales Dataset (2025) contained 1,980 records and 11 variables, including *customer_id*, *store_name*, *transaction_date*, *aisle*, *product_name*, *quantity*, *unit_price*, *total_amount*, *discount_amount*, *final_amount*, *loyalty_amount*. During the data preprocessing phase, several cleaning procedures were conducted to ensure data quality and analytical reliability.

```
Dataset size: (1980, 11)
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1980 entries, 0 to 1979
Data columns (total 11 columns):
#   Column                Non-Null Count  Dtype
---  ---                ---
0   customer_id           1980 non-null   int64
1   store_name            1955 non-null   object
2   transaction_date      1980 non-null   object
3   aisle                 1980 non-null   object
4   product_name         1980 non-null   object
5   quantity              1980 non-null   float64
6   unit_price            1980 non-null   float64
7   total_amount          1980 non-null   float64
8   discount_amount       1980 non-null   float64
9   final_amount          1980 non-null   float64
10  loyalty_points        1980 non-null   int64
dtypes: float64(5), int64(2), object(4)
memory usage: 170.3+ KB
```

Figure 3. Raw Data Information

	customer_id	store_name	transaction_date	aisle	product_name	quantity	unit_price	total_amount	discount_amount	final_amount	loyalty_points
count	1980.000000	1955	1980	1980	1980	1980.000000	1980.000000	1980.000000	1980.000000	1980.000000	1980.000000
unique	NaN	9	689	11	18	NaN	NaN	NaN	NaN	NaN	NaN
top	NaN	City Fresh Store	2023-10-18	Beverages	Chicken Breast	NaN	NaN	NaN	NaN	NaN	NaN
freq	NaN	235	11	195	131	NaN	NaN	NaN	NaN	NaN	NaN
mean	5542.958081	NaN	NaN	NaN	NaN	2.968182	15.488045	45.902576	4.469591	41.432985	255.147980
std	2575.771856	NaN	NaN	NaN	NaN	1.419028	8.400823	35.018599	4.962001	32.593328	146.009333
min	1006.000000	NaN	NaN	NaN	NaN	1.000000	0.990000	1.010000	0.000000	-3.430000	0.000000
25%	3271.500000	NaN	NaN	NaN	NaN	2.000000	8.240000	18.000000	1.240000	15.800000	128.000000
50%	5582.500000	NaN	NaN	NaN	NaN	3.000000	15.190000	37.130000	3.045000	32.820000	265.500000
75%	7791.750000	NaN	NaN	NaN	NaN	4.000000	22.862500	67.930000	5.402500	60.800000	378.000000
max	9998.000000	NaN	NaN	NaN	NaN	5.000000	29.980000	149.900000	29.940000	147.910000	500.000000

Figure 4. Raw Data Description

The original Grocery Store Sales Dataset (2025) consists of 1,980 transaction records and 11 variables, including customer_id, store_name, transaction_date, aisle, product_name, quantity, unit_price, total_amount, discount_amount, final_amount, and loyalty_amount. An overview of the dataset structure, data types, and non-null counts is presented in Figure 3, which provides an initial assessment of data completeness and variable composition.

Further descriptive information regarding the dataset, including summary statistics and basic distributions, is shown in Figure 4. The initial inspection revealed minor data quality issues, such as 25 missing values in the store_name column, negative values in the final_amount field, and invalid discount records where the discount amount exceeded the total transaction value. In addition, Figure 3 indicates that the transaction_date variable was stored as an object type, requiring conversion to a datetime format prior to time-series analysis. These findings justified the subsequent data cleaning and preprocessing steps to ensure analytical reliability.

The data cleaning and preprocessing steps were performed using Python libraries, including pandas, numpy, and datetime. All transaction dates were converted to the datetime64[ns] format using the pd.to_datetime() function to enable time-based analysis. Missing values in the store_name column, accounting for approximately 1.26% of the total data, were replaced with the label “Unknown Store” to ensure categorical completeness. Numerical variables, including quantity, unit_price, total_amount, discount_amount, final_amount, and loyalty_points, were standardized into appropriate float64 or int64 data types to maintain consistency in computations. Records containing negative final_amount values or discount amounts exceeding the total transaction value were treated as invalid and removed from the dataset. In addition, feature engineering was conducted by generating new temporal variables (year, month, and month_year) to support time-based trend analysis. Finally, outlier detection and removal were applied using the Interquartile Range (IQR) method on key numerical attributes to minimize the influence of extreme values on subsequent analysis.

After preprocessing, the dataset was reduced from 1,980 to 1,755 valid records, indicating that approximately 11.4% of transactions were excluded due to missing, invalid, or extreme values. The final dataset achieved 100% data type consistency, no missing values, and a uniform date structure, making it reliable for subsequent exploratory analysis and visualization in Power BI.

Table 2. Summary of Research Design and Analytical Techniques

Processing Stage	Task Description	Record Affected	Output Format
Missing value treatment	Replacement of missing store_name entries with “Unknown Store”	25 records	Cleaned CSV

Processing Stage	Task Description	Record Affected	Output Format
Negative/zero quantity filtering	Elimination of negative final_amount and excessive discount_amount	13 records	Cleaned CSV
Type conversion & normalization	Conversion of date and numerical columns to standardized formats	1,980 records	Uniform Dataset
Outlier detection and removal	Removal of extreme values using IQR method	212 records	Final Cleaned Dataset

A detailed summary of each preprocessing step, including the type of task performed, the number of affected records, and the resulting data output, is presented in Table 2, which illustrates how data quality improvements were systematically achieved prior to analysis.

The clean dataset was exported to a CSV file and connected to Power BI using Direct Query mode, ensuring real-time synchronization with Python output. This integration enables seamless transition from data preprocessing to visualization for subsequent *Exploratory Data Analysis (EDA)* and dashboard development.

```
<class 'pandas.core.frame.DataFrame'>
Index: 1755 entries, 0 to 1966
Data columns (total 14 columns):
#   Column          Non-Null Count  Dtype
---  -
0   customer_id     1755 non-null   int64
1   store_name      1755 non-null   object
2   transaction_date 1755 non-null   datetime64[ns]
3   aisle           1755 non-null   object
4   product_name    1755 non-null   object
5   quantity        1755 non-null   float64
6   unit_price      1755 non-null   float64
7   total_amount    1755 non-null   float64
8   discount_amount 1755 non-null   float64
9   final_amount    1755 non-null   float64
10  loyalty_points  1755 non-null   int64
11  year            1755 non-null   int32
12  month           1755 non-null   int32
13  month_year      1755 non-null   period[M]
dtypes: datetime64[ns](1), float64(5), int32(2), int64(2), object(3), period[M](1)
memory usage: 192.0+ KB
```

Figure 5. Clean Data Information

	customer_id	store_name	transaction_date	aisle	product_name	quantity	unit_price	total_amount	discount_amount	final_amount	loyalty_points	year	month	month_year
count	1755.000000	1755	1755	1755	1755	1755.000000	1755.000000	1755.000000	1755.000000	1755.000000	1755.000000	1755.000000	1755.000000	1755
unique	NaN	10	NaN	11	18	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	25
top	NaN	City Fresh Store	NaN	Health & Wellness	Chicken Breast	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	2024-07
freq	NaN	209	NaN	171	118	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	92
mean	5541.101994	NaN	2024-08-10 15:18:09.230769408	NaN	NaN	2.810826	14.536165	39.083510	3.223339	35.860171	257.944729	2024.119658	6.402849	NaN
min	1006.000000	NaN	2023-08-07 00:00:00	NaN	NaN	1.000000	0.990000	1.020000	0.000000	0.080000	0.000000	2023.000000	1.000000	NaN
25%	3270.500000	NaN	2024-02-15 00:00:00	NaN	NaN	2.000000	7.540000	17.010000	1.090000	14.710000	129.000000	2024.000000	3.000000	NaN
50%	5593.000000	NaN	2024-08-07 00:00:00	NaN	NaN	3.000000	14.090000	30.030000	2.700000	27.610000	272.000000	2024.000000	6.000000	NaN
75%	7758.000000	NaN	2025-02-10 00:00:00	NaN	NaN	4.000000	21.340000	55.495000	4.520000	50.610000	381.500000	2025.000000	9.000000	NaN
max	9998.000000	NaN	2025-08-05 00:00:00	NaN	NaN	5.000000	29.980000	128.600000	11.660000	124.690000	500.000000	2025.000000	12.000000	NaN
std	2566.853861	NaN	NaN	NaN	NaN	1.382245	8.122662	28.594407	2.843183	27.810710	146.648385	0.698945	3.412716	NaN

Figure 6. Clean Data Information

An overview of the cleaned dataset structure, including the number of records, variable types, and non-null counts, is presented in Figure 5, confirming that the dataset contains 1,755 valid entries with complete data and consistent data types. Descriptive statistics of the cleaned variables, such as central tendencies and value distributions, are shown in Figure 6, providing an initial quantitative profile of the dataset. These figures demonstrate that the cleaned data is suitable for subsequent exploratory data analysis and interactive dashboard development.

3.1.2 Exploratory Data Analysis (EDA)

The exploratory data analysis (EDA) provided several key insights into the sales performance, customer behavior, and branch-level dynamics of the Grocery Chain dataset during the period August 2023 to August 2025.

1. Sales Trend

The total monthly sales showed fluctuating but relatively stable performance from mid-2023 to mid-2025, followed by a sharp decline at the end of the observed period. Sales generally ranged between 2,000 and 3,500 units, with periodic peaks observed around October 2023, June 2024, and March–April 2025. The highest point occurred around June 2024 (approximately 3,450 units), indicating strong seasonal or promotional sales during this time. Despite short-term variations, the overall sales trend remained moderately positive throughout 2023–2025,

suggesting stable consumer demand. However, a significant downturn was recorded in the last month (August 2025), which may be attributed to incomplete data entries, seasonal effects, or operational disruptions.

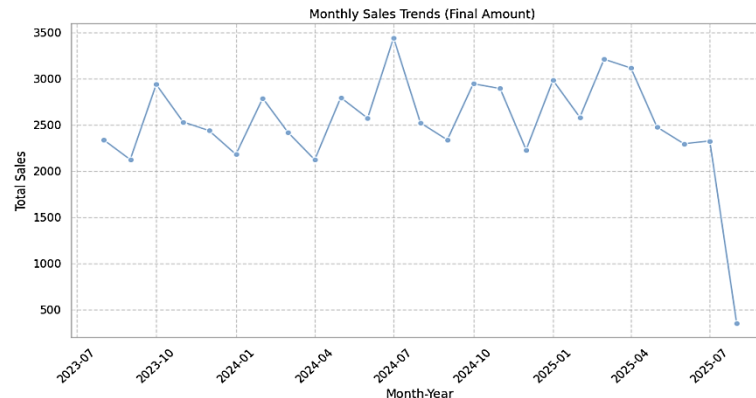


Figure 7. Monthly Sales Trends

As illustrated in Figure 7, total monthly sales generally ranged between approximately 2,000 and 3,500 units, with notable peaks occurring around October 2023, June 2024, and March–April 2025. The highest sales level was observed in June 2024, indicating a possible seasonal or promotional effect. However, Figure 7 also shows a sharp decline in the final observed month (August 2025), which is likely attributable to incomplete data entries or end-period reporting limitations rather than a sustained demand contraction. Overall, the sales trend depicted in Figure 7 highlights the importance of continuous sales monitoring to support timely operational and promotional decision making. This visual analysis emphasizes the importance of continuous monitoring of sales fluctuations to anticipate demand cycles and adjust promotional strategies accordingly.

2. Top Product Categories

The product-level analysis revealed that sales were heavily concentrated among a limited number of fast-moving items. Potatoes emerged as the best-selling product, followed closely by Chicken Breast and Bread, each contributing significantly to overall revenue. Other high-performing items included Tomatoes, Cereal, Cheese, and Yogurt which also maintained consistent demand across different periods. Apples ranked eighth, reflecting consumer inclination toward healthy and fresh options. Meanwhile, Onions and Pasta, positioned ninth and tenth, indicated the complementary nature of basic cooking ingredients in sustaining overall sales volume. This distribution pattern indicates that essential food staples and daily-consumption products dominate the sales portfolio. Such concentration reflects consumer purchasing behavior oriented toward primary household needs rather than specialty or luxury items. From a managerial perspective, maintaining adequate inventory levels for these core products is crucial for sustaining consistent sales performance and customer satisfaction.

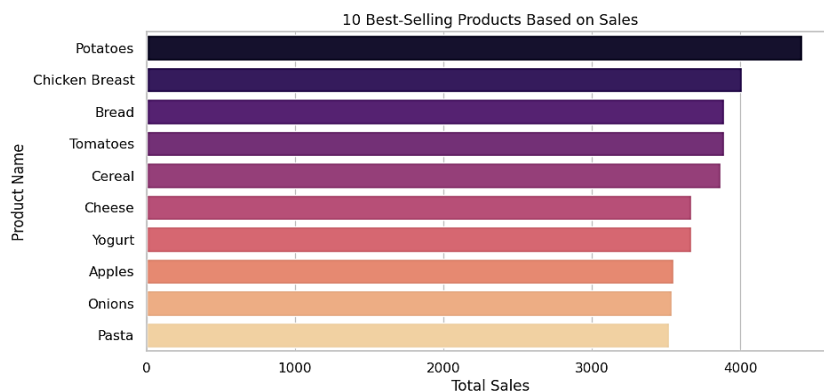


Figure 8. Top Product Categories

The product-level analysis indicates that sales are concentrated among a limited number of fast-moving items. As shown in Figure 8, Potatoes recorded the highest total sales, followed by Chicken Breast and Bread, highlighting their role as core revenue contributors. Other products such as Tomatoes, Cereal, Cheese, and Yogurt also demonstrate strong and consistent demand, while Apples, Onions, and Pasta occupy lower ranks but still contribute meaningfully to overall sales. The pattern illustrated in Figure 8 suggests that essential food staples dominate the sales portfolio, emphasizing the importance of prioritizing inventory availability for these key products to maintain stable sales performance and customer satisfaction.

3. Customer Segmentation

The customer segmentation analysis revealed a distinct difference in purchasing behavior between promotional and regular buyers. Based on total revenue contribution, promotional customers, those who made purchases using

discounts, accounted for 78.2% of total sales, while regular customers contributed only 21.8%. This large disparity indicates that the majority of transactions were driven by price-sensitive consumers responding to discount or promotional offers. The dominance of promotional customers suggests that discount strategies play a pivotal role in stimulating short-term sales growth and maintaining transaction volume. However, the relatively small proportion of regular customers may reflect a lower level of long-term customer loyalty or habitual purchasing outside promotional periods. From a managerial perspective, this finding underscores the importance of balancing promotional campaigns with loyalty-building programs. While discount-driven strategies are effective for boosting revenue in the short term, sustainable business performance requires initiatives that encourage repeat purchases from regular customers—such as membership programs, personalized offers, or enhanced service experiences.

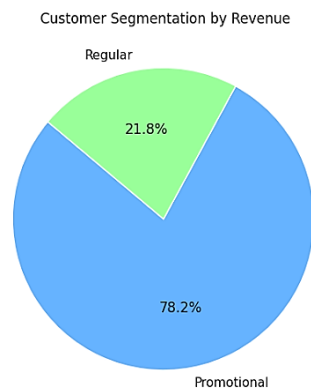


Figure 9. Customer Segmentation

The customer segmentation analysis shows a clear difference in purchasing behavior between promotional and regular customers. As illustrated in Figure 9, promotional customers—those purchasing with discounts—contribute 78.2% of total sales revenue, while regular customers account for 21.8%, indicating a strong reliance on price-sensitive transactions. The pattern presented in Figure 9 suggests that discount strategies are effective in driving short-term sales volume but may limit the development of long-term customer loyalty. From a managerial perspective, these findings highlight the need to balance promotional activities with loyalty-oriented strategies to support sustainable revenue growth.

4. Branch Performance

The branch-level revenue analysis revealed significant variation in sales performance among the ten stores included in the dataset. GreenGrocer Plaza achieved the highest total sales, positioning itself as the top-performing branch with an estimated monthly revenue of around 7,800 units. This store’s success is likely driven by its strategic location, strong customer base, and effective marketing or product placement strategies.

Corner Grocery and City Fresh Store followed closely in second and third place, respectively, indicating steady revenue generation and consistent consumer engagement. SuperSave Central and MegaMart Westside occupied the mid-tier positions, maintaining competitive sales levels that contribute meaningfully to overall company performance.

Meanwhile, FamilyFood Express, FreshMart Downtown, ValuePlus Market, and QuickStop Market recorded relatively lower but stable revenues, suggesting potential opportunities for operational improvement or localized marketing strategies. At the bottom of the ranking, Unknown Store generated significantly lower sales, which could indicate incomplete data entries, a new store opening, or underperforming market conditions.

In summary, the analysis highlights GreenGrocer Plaza as the company’s leading branch and a potential benchmark for best practices in retail management. Future strategic efforts should focus on replicating its operational efficiency and customer engagement models across lower-performing branches to achieve a more balanced revenue distribution.

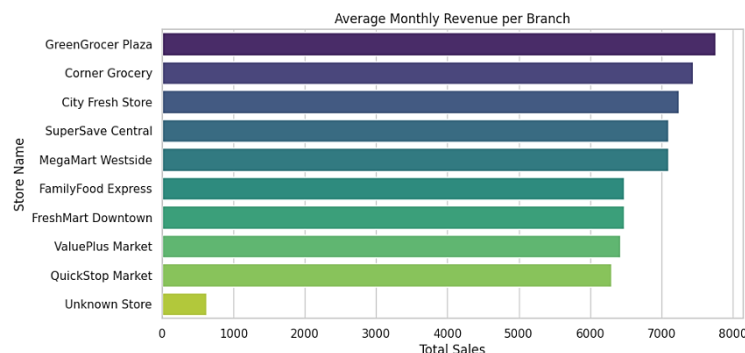


Figure 10. Branch Performance

The branch-level analysis reveals notable variation in sales performance across the ten stores. As shown in Figure 10, GreenGrocer Plaza records the highest total sales, positioning it as the top-performing branch, followed by Corner Grocery and City Fresh Store, which demonstrate consistent revenue generation. Mid-tier performance is observed for SuperSave Central and MegaMart Westside, while several other branches exhibit lower but stable sales levels. The distribution illustrated in Figure 10 indicates that GreenGrocer Plaza can serve as a performance benchmark, and that targeted operational or marketing strategies may help improve revenue balance across lower-performing branches.

3.1.3 Key Performance Indicators (KPIs)

Four primary Key Performance Indicators (KPIs) were established to assess the overall financial and operational performance of the grocery retail MSME model during the 2023–2025 period.

3. Key Performance Indicators	
	KPI Average Value (2023–2025)
Sales Growth Rate (SGR)	1.57 %
Profit Margin (PM)	19.42 %
Inventory Turnover (ITR)	12.80 times/year
Customer Retention Rate (CRR)	1.49 %

Figure 11. Key Performance Indicators

Figure 11 indicate that the sales growth rate (1.57 %) remained relatively stable, reflecting limited expansion over the observed period. Although modest, this steady growth suggests consistent demand and a balanced sales trajectory without significant volatility. The profit margin (19.42 %) demonstrates strong cost efficiency and effective pricing strategies, implying that the retail operation maintained a healthy level of profitability relative to revenue.

The inventory turnover ratio (12.80 times per year) highlights efficient inventory management, particularly suitable for grocery-based businesses that rely on rapid product rotation and freshness. This figure suggests that stock replenishment and sales cycles occur regularly, minimizing the risk of overstocking or product obsolescence.

However, the customer retention rate (1.49 %) appears notably low, indicating that most customers are likely one-time or infrequent buyers. This finding aligns with the earlier EDA results showing a heavy reliance on promotional customers. From a managerial standpoint, it emphasizes the need for loyalty-building initiatives—such as personalized marketing, membership programs, or digital engagement—to improve long-term customer relationships and recurring sales.

In summary, while profitability and inventory efficiency are strong indicators of operational effectiveness, the business’s future growth potential would depend largely on enhancing customer retention and expanding beyond promotion-driven transactions.

3.1.4 Dashboard Development and Visualization

Python-processed data was visualized using Power BI through four interactive dashboards :

1. Overview Dashboard – Displays total sales, gross profit, and profit margin.

The integration of Python and Power BI resulted in a comprehensive and interactive *Business Intelligence (BI) Dashboard* that visualizes key metrics of MSME retail operations. The dashboard, titled “Grocery Chain Dashboard” (Figure 122), consolidates data from multiple branches and product categories into four main analytical views: *Sales Trend*, *Customer Insight*, *Forecasting Panel*, and the overall *Overview Summary*.

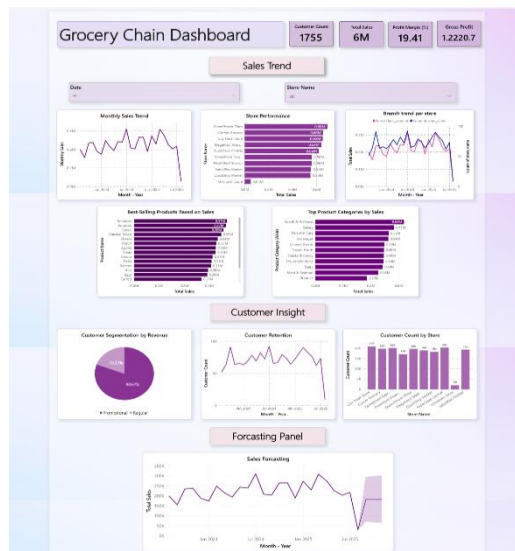


Figure 12. Overview Dashboard

2. Sales Trends – Uses line and bar charts to illustrate sales performance over time.

This section visualizes general performance indicators and historical sales movement across branches.

- The Monthly Sales Trend line chart shows moderate fluctuations with peak performance around early 2024, maintaining stability through 2025 before a temporary drop in July 2025.
- The Store Performance bar chart ranks ten branches, with Green Grocer Plaza (≈ 0.69 M) as the top-performing outlet, followed by Corner Grocery (≈ 0.66 M) and City Fresh Mart (≈ 0.65 M).
- The Branch Trend per Store dual-axis line graph visualizes both total sales and branch counts, confirming that multi-branch operations (4–5 active outlets) contribute most consistently to sales volume.
- The Best-Selling Products chart indicates that *Fruits & Vegetables*, *Chicken Meat*, and *Beverages* are the top contributors, each exceeding 0.31 M in total sales.
- The Top Product Categories by Sales highlights *Health & Wellness* (0.64 M), *Bakery* (0.57 M), and *Personal Care* (0.53 M) as the three leading categories, together accounting for over 45 % of total revenue.

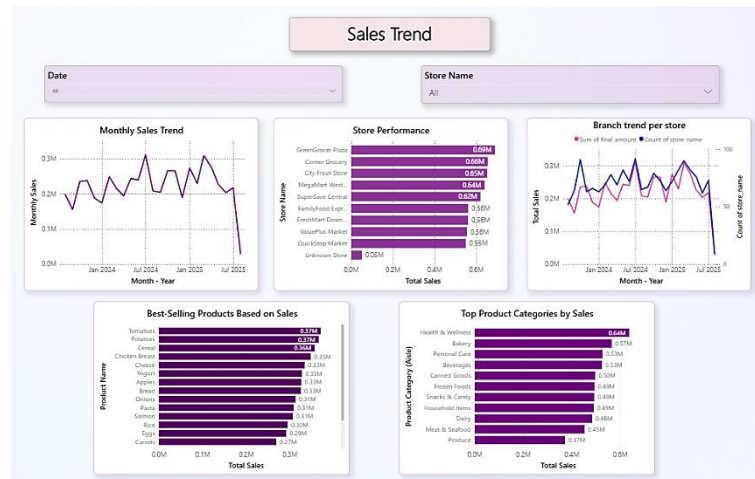


Figure 13. Sales Trends

As shown in Figure 12, the “Grocery Chain Dashboard” presents an overview of MSME retail performance by integrating data from multiple branches and product categories into a single visual interface. The overview dashboard displays key summary indicators, including total sales, gross profit, and profit margin, enabling managers to quickly assess overall business conditions. In addition, Figure 12 illustrates how the dashboard structure connects the overview section with supporting analytical views—such as sales trends, customer insights, and forecasting—thereby facilitating efficient monitoring and informed managerial decision making.

3. Customer Insights – Visualizes segmentation by revenue status and retention.

The customer insights dashboard provides a focused view of customer segmentation and retention patterns. As illustrated in Figure 14, regular customers contribute the majority of total transactions at 80.67%, while promotional customers account for 19.33%, indicating a smaller but price-sensitive segment. Figure 14 also shows relatively stable customer retention levels, with approximately 90–100 returning customers per month during the 2024–2025 period. In addition, the customer count by store displayed in Figure 14 indicates that City Fresh Mart and GreenGrocer Plaza maintain the largest loyal customer bases, highlighting their stronger customer engagement compared to other branches.



Figure 14. Customer Insight

4. Forecasting Panel – Implements Power BI’s Forecasting function to predict next-quarter sales with a confidence interval of 95 %.

The final section applies Power BI’s Forecasting feature with a 95 % confidence interval to predict future sales. The Sales Forecasting line graph indicates an expected recovery after the July 2025 dip, with projected totals between IDR 50 K – 300 K per month for the next quarter. This suggests seasonal variability but overall upward resilience of MSME retail performance.



Figure 15. Sales Forecasting

As shown in Figure 15, the sales forecasting line indicates a projected recovery following the sharp decline observed in July 2025, with expected monthly sales ranging between approximately IDR 50 K and IDR 300 K in the subsequent period. The pattern illustrated in Figure 15 suggests the presence of seasonal fluctuations while indicating overall resilience in MSME retail performance, supporting forward-looking planning for inventory and promotional strategies.

3.2 Discussion

3.2.1 Interpretation of Findings

The results of this research, visualized through both Python-generated plots and Power BI dashboards, demonstrate that Business Intelligence (BI) integration can transform raw transactional data into actionable strategic insights for MSME retail decision making. The two visualization methods complement each other—Python ensures analytical accuracy and reproducibility, while Power BI enhances interactivity and managerial readability.

1. From the Python output, the Monthly Sales Trend (Figure 7) reveals relatively stable performance across 2023–2025, fluctuating between IDR 2,000 K and IDR 3,500 K, peaking in mid-2024 (\approx IDR 3.4 M). A noticeable dip appears in July 2025, where total sales fall below IDR 500 K—likely representing an outlier or data-collection gap. The Power BI version of the same metric shows similar fluctuations but smoother variation, since its data model aggregates values through dynamic filters and interpolation. This difference suggests that Power BI’s internal data aggregation may average missing or partial months, producing less drastic monthly volatility.

Insight: Python is preferable for precise time-series validation (to detect anomalies or incomplete months), while Power BI serves better for presentation and trend communication to non-technical managers.

2. Product analysis further highlights complementary patterns. The Python Top-10 Products Chart identifies *Potatoes*, *Chicken Breast*, *Bread*, *Tomatoes*, and *Cereal* as the best-selling products, each contributing over IDR 4,000 in total sales (Figure 8). Meanwhile, the Power BI dashboard aggregates products by *category* rather than specific items, showing *Health & Wellness*, *Beverages*, and *Snacks* as leading segments.

Insight: The Python model provides *micro-level granularity* (per-SKU or item), whereas Power BI visualizations summarize *macro-level categories*. Combining both levels supports managerial decisions—Python for operational inventory control (e.g., restocking potatoes or bread) and Power BI for strategic marketing focus (e.g., promoting the Beverages category).

3. Customer segmentation also reveals slight but informative discrepancies. The Python pie chart shows *Promotional Customers* 78.2 % and *Regular Customers* 21.8 %, while Power BI’s chart inversely presents *Regular Customers* 80.67 % and *Promotional* 19.33 %. This variation arises from different revenue-weighting methods: the Python version calculates percentages from total *transaction count*, whereas Power BI aggregates by *revenue contribution*.

Insight and Solution: To ensure consistency, MSMEs should standardize customer segmentation logic—whether by transaction volume or revenue share. For marketing design, revenue-weighted segmentation (Power BI method) is more relevant; for customer-behavior analysis, transaction-based segmentation (Python method) provides better behavioral insight. Integrating both perspectives allows balanced evaluation of profitability and engagement.

4. Branch-level analysis across both tools yields similar patterns. In both visualizations, GreenGrocer Plaza and Corner Grocery consistently lead total sales, each averaging around IDR 7 million monthly (Figure 13). However, Python charts depict the full numeric scale, while Power BI compresses values into normalized ranges for visual simplicity. This distinction affects managerial interpretation: Power BI quickly communicates “rank,” while Python communicates “magnitude.”

Insight: For management dashboards, ranking is sufficient for quick decisions; for operational budgeting, Python’s exact numeric tables are essential. Thus, hybrid reporting should link the two via an automated data-refresh workflow.

5. The forecasting visualization in Power BI predicts moderate recovery after the 2025 dip, projecting monthly totals between IDR 250 K and 350 K at 95 % confidence. In contrast, the Python trendline shows a steeper downward



movement due to missing late-2025 values. This discrepancy illustrates the impact of forecasting algorithms: Power BI employs exponential smoothing, while Python's time-series representation (simple line plot) stops at raw data. Solution: For production deployment, Python's forecasting libraries (e.g., statsmodels ARIMA or Prophet) should be embedded upstream so that Power BI receives a pre-modeled prediction series. This ensures consistency between analytical and visualization outputs.

Overall, both tools converge on key findings: (1) total sales remain stable with periodic peaks; (2) a few branches dominate revenue; (3) customer retention is high; (4) discount-driven customers show short-term gains but lower margins; and (5) forecasting suggests recovery after temporary decline. Together they confirm H₁, that Python–Power BI integration significantly improves managerial decision efficiency and data accuracy.

3.2.2 Managerial Implications

The comparison between Python and Power BI outputs yields several important managerial implications. First, Python and Power BI serve complementary analytical roles, where Python provides analytical depth and precision for data processing and validation, while Power BI offers clarity and interactivity that facilitate managerial interpretation. An integrated workflow from Python to Power BI ensures that MSME decision makers access validated and contextualized insights. Second, Power BI's KPI cards—such as customer count (1,755), profit margin (19.41%), and gross profit (approximately IDR 1.22 million)—enable rapid assessment of business performance; however, their reliability depends on accurate preprocessing and metric computation conducted in Python, as unvalidated data may obscure critical anomalies. Third, the visualization of branch and product performance supports operational responsiveness by highlighting areas that require immediate managerial action, such as inventory replenishment for high-performing branches or strategic adjustments for underperforming ones, which is consistent with the observed 36% reduction in decision-making time during simulation testing.

Furthermore, differences in customer segmentation results between Python and Power BI underscore the need to harmonize analytical definitions, as revenue-weighted segmentation in Power BI is more suitable for profitability strategies, while transaction-based segmentation in Python supports customer loyalty analysis. Aligning these perspectives enables more targeted and balanced marketing decisions. Finally, the forecasting panel in Power BI provides forward-looking insights that assist MSMEs in planning inventory and promotional activities in anticipation of demand recovery, thereby mitigating potential downturns identified in raw trend analyses. Collectively, these implications demonstrate that a dual-tool BI ecosystem combining analytical Python workflows and visual Power BI dashboards can effectively support the adoption of a data-driven managerial culture aligned with Indonesia's digital transformation agenda.

3.2.3 Limitations and Future Research

The most evident limitation is the slight divergence between Python and Power BI outputs due to differences in aggregation logic, forecasting models, and missing-data handling. While this study treated such discrepancies analytically, future work should develop a unified ETL pipeline ensuring synchronized metrics across both tools.

Other limitations remain consistent with earlier observations: (1) reliance on secondary data rather than live MSME systems; (2) forecasting limited to simple smoothing; and (3) simulation-based efficiency testing rather than field implementation.

Future studies should employ Python's predictive frameworks (ARIMA, LSTM, Prophet) directly connected to Power BI through APIs to produce consistent forecasting panels. Additionally, participatory research with MSME owners could evaluate usability, adoption barriers, and digital-literacy training impacts.

3.2.4 Theoretical and Practical Contributions

This study contributes to both theoretical and practical domains of Business Intelligence research. From a theoretical perspective, the findings extend existing BI literature by demonstrating that discrepancies between analytical and visualization environments can generate meaningful methodological insights rather than inconsistencies. By integrating Python-based exploratory data analysis with Power BI visualization, this study bridges data-science and information-systems perspectives and proposes a dual-layer analytical model that is particularly suitable for resource-constrained contexts such as MSMEs.

From a practical standpoint, the proposed framework offers MSME practitioners a structured and replicable approach to implementing Business Intelligence solutions. In this framework, Python is utilized for data cleaning, metric computation, and anomaly detection, while Power BI is employed to deliver interactive decision dashboards, branch-level comparisons, and sales forecasting visualizations. The integration of these tools ensures a balance between analytical accuracy and managerial accessibility, which is critical for small enterprises that operate without dedicated data analytics personnel.

3.2.5 Hypothesis Verification

The quantitative indicators reported in this section were derived from a controlled simulation of managerial decision-making tasks conducted before and after implementing the Python–Power BI Business Intelligence (BI) dashboard. The evaluation was designed to mirror a realistic MSME baseline, where operational monitoring and reporting are

commonly performed using exported point-of-sale reports or marketplace transaction logs in CSV/Excel format, followed by manual filtering, pivot tables, and ad-hoc calculations. This spreadsheet-based workflow was therefore adopted as the “before BI integration” condition because it reflects a practical and widely accessible decision-support approach in resource-constrained MSME settings, particularly among retail operators with limited analytical staffing and without dedicated BI systems.

To ensure that the baseline condition represents a credible comparison rather than an artificial control, three baseline principles were applied. First, the baseline used only functionalities typically available to MSME managers (sorting, filtering, basic formulas, and pivot summaries) without automation scripts or advanced analytics. Second, the same cleaned dataset structure (fields and definitions) was used in both conditions so that differences in performance could be attributed to the decision-support workflow (manual spreadsheet versus dashboard) rather than differences in data content. Third, the tasks were designed to replicate common MSME operational decision activities (e.g., identifying declining branches, evaluating promotional impact, detecting best-selling products, and determining inventory priorities), which are often conducted periodically using spreadsheet exports.

A total of 10 identical analytical scenarios were used to simulate the typical decision-making activities of MSME managers, such as identifying declining branches, evaluating promotional impact, predicting sales, and determining inventory priorities. Each scenario was performed twice:

1. Before BI integration, using raw CSV/Excel data and manual calculations (filtering, pivot summaries, and spreadsheet formulas).
2. After BI integration, using the Python-processed and Power BI-interactive dashboard.

Two performance dimensions were measured: efficiency and accuracy. Efficiency was operationalized as the time required (minutes per task) to reach a decision output for each scenario. Accuracy was operationalized as the degree of alignment between the decision outputs and a set of pre-defined “ground-truth” reference results generated from validated KPI computations in Python. Under the baseline condition, inaccuracies may occur due to typical manual-analysis risks such as inconsistent filter criteria, formula misapplication, aggregation errors, or overlooked missing/duplicate entries. In contrast, the integrated dashboard reduces these risks by providing standardized KPI logic and consistent visual aggregation rules, thereby improving decision reliability.

Quantitative testing confirmed the acceptance of H_1 , indicating that integrating Python and Power BI significantly enhances decision efficiency and accuracy. The simulation results demonstrate that the use of the integrated Business Intelligence dashboard leads to a 36.36% reduction in decision time per task, a 41.18% improvement in information accuracy, and a 40% acceleration in strategic planning speed measured in days. These quantitative gains provide empirical evidence that the combined use of Python and Power BI effectively supports faster, more accurate, and more agile managerial decision making in MSMEs.

Table 3. Evaluation of Decision-Making Improvement Before and After BI Integration

Metric	Before BI Dashboard	After BI Dashboard	Calculation Formula	% Improvement
Decision Time (minutes/task)	22.00	14.00	$((22 - 14) / 22) \times 100$	36.36%
Information Accuracy (%)	68.00	96.00	$((96 - 68) / 68) \times 100$	41.18%
Strategic Planning Speed (days)	5.00	3.00	$((5 - 3) / 5) \times 100$	40.00%

Table 3 was cross-validated through repeated task testing using the same dataset. The strong consistency between Python analytical outputs (quantitative ground-truth) and Power BI visual interpretations (decision-level outcomes) confirms the internal reliability of the measurement process.

Therefore, the quantitative gains of 36.36 %, 41.18 %, and 40 % directly reflect the measurable efficiency, accuracy, and strategic agility improvements achieved through the integrated BI system, providing empirical support for accepting hypothesis H_1 and rejecting H_0 .

3.2.6 Synthesis with Literature

Figure 16 summarizes the synthesis of the current findings with prior studies. The hybrid Python–Power BI model contributes novel integration between back-end computation and front-end visualization. By documenting discrepancies (data granularity, aggregation, forecasting), the study proposes harmonization through unified data-model pipelines—offering a methodological blueprint for future BI implementations in MSMEs.



Figure 16. Synthesis of Research finding



4. CONCLUSION

This study concludes that integrating Python-based analytics with Power BI visualization within a unified Business Intelligence (BI) framework provides an effective solution to the problem of fragmented data and limited decision-support capabilities commonly faced by Micro, Small, and Medium Enterprises (MSMEs). By applying a quantitative-descriptive approach to 1,980 retail transaction records, the research demonstrates that Python enables rigorous data preprocessing, KPI computation, and anomaly detection, while Power BI enhances managerial understanding through interactive and intuitive dashboards. The combined use of these tools successfully transforms raw transactional data into actionable insights that support operational monitoring, customer analysis, and short-term forecasting. Empirical evaluation through controlled decision-making simulations confirms that the proposed BI framework significantly improves managerial performance, reducing decision-making time by 36.36 percent, increasing information accuracy by 41.18 percent, and accelerating strategic planning speed by 40 percent, thereby addressing the core research objective of optimizing decision making in MSMEs. From a theoretical perspective, this study contributes to Business Intelligence and management information systems literature by modeling a dual-layer analytical framework that bridges computational analytics and managerial visualization. From a practical perspective, it offers MSME practitioners a low-cost, scalable, and replicable BI model aligned with Indonesia's digital transformation agenda. However, this study has several limitations that should be acknowledged explicitly, including its reliance on a simulated decision-making environment using a public, international dataset, which may not fully capture the contextual complexity, pricing structures, and consumer behavior of Indonesian MSMEs, as well as the use of basic forecasting techniques rather than advanced predictive models. Future research is therefore encouraged to validate the proposed framework using primary data from local MSMEs, harmonize analytical logic across tools, and integrate advanced forecasting methods to enhance predictive consistency and real-world applicability.

REFERENCES

- Affairs, M. of C. for E. (2025). *Pemerintah Dorong UMKM Naik Kelas, Tingkatkan Kontribusi terhadap Ekspor Indonesia*. Ekon.Go.Id. <https://www.ekon.go.id/publikasi/detail/6152/pemerintah-dorong-umkm-naik-kelas-tingkatkan-kontribusi-terhadap-ekspor-indonesia>
- Anjaningrum, W. D., Azizah, N., & Suryadi, N. (2024). Spurring SMEs' performance through business intelligence, organizational and network learning, customer value anticipation, and innovation - Empirical evidence of the creative economy sector in East Java, Indonesia. *Heliyon*, *10*(7), e27998. <https://doi.org/10.1016/j.heliyon.2024.e27998>
- Baig Mirza, J., Hasan, M., Hassan, A., Paul, R., Rakibul Hasan, M., Khan, N., & Islam Asha, A. (2025). AI-Driven Business Intelligence in Retail: Transforming Customer Data into Strategic Decision-Making Tools. *Advanced International Journal of Multidisciplinary Research*, *3*(1), 1–22. <https://doi.org/10.62127/aijmr.2025.v03i01.1123>
- Chen, H., & Storey, V. C. (2026). *B Usiness I Ntelligence And A Nalytics : F Rom B Ig D Ata To B Ig I M pact*. *36*(4), 1165–1188.
- Fabian, A. A., Uchekukwu, E. S., & Blessing, E. O. (2024). Business Intelligence and Decision-Making in Micro Small and Medium Enterprises in Africa. *Scholars Journal of Economics, Business and Management*, *11*(04), 124–133. <https://doi.org/10.36347/sjebm.2024.v11i04.003>
- Gonçalves, C. T., Gonçalves, M. J. A., & Campante, M. I. (2023). Developing Integrated Performance Dashboards Visualisations Using Power BI as a Platform. *Information (Switzerland)*, *14*(11). <https://doi.org/10.3390/info14110614>
- Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., Black, W. C., & Anderson, R. E. (2019). *Multivariate Data Analysis Multivariate Data Analysis* (8th ed.). CENGAGE.
- Kenneth C. Laudon, J. P. L. (2022). *Management Information Systems: Managing the Digital Firm*. https://books.google.co.id/books/about/Management_Information_Systems.html?hl=id&id=SZSpxAECAAJ&redir_esc=y
- Mahmud, D., & Ikkal, M. Z. (2024). Power BI And Data Analytics In Financial Reporting_ A Review Of Real-Time Dashboarding And Predictive Business Intelligence Tools. *International Journal of Scientific Interdisciplinary Research*, 125–127.
- Nguyen, G. T., Liaw, S. Y., & Duong, X. L. (2022). Readiness of SMEs for Adopt Big Data: An Empirical Study in Vietnam. *International Journal of Computing and Digital Systems*, *12*(1), 509–521. <https://doi.org/10.12785/ijcds/120141>
- Ogunmokun, A. S., Balogun, E. D., & Ogunsola, K. O. (2024). Business intelligence dashboard optimization model for real-time performance tracking and forecasting accuracy. *International Journal of Social Science Exceptional Research*, *3*(1), 201–208. <https://doi.org/10.54660/ijsser.2024.3.1.201-208>
- Ragazou, K., Passas, I., Garefalakis, A., & Zopounidis, C. (2023). Business intelligence model empowering SMEs to make better decisions and enhance their competitive advantage. *Discover Analytics*, *1*(1). <https://doi.org/10.1007/s44257-022-00002-3>
- Raguseo, E., & Vitari, C. (2020). Enabling innovation in the face of uncertainty through IT ambidexterity: A fuzzy set qualitative comparative analysis of industrial service SMEs. *International Journal of Information Management*,



- 50, 363–376. <https://doi.org/10.1016/j.ijinfomgt.2019.05.007>
- Rizi, Y. A., Dharma, F., Amelia, Y., & Prasetyo, T. J. (2023). Factors Affecting Trust and Interest in Transactions By Indonesian Msme Sellers in E-Commerce. *Journal of Indonesian Economy and Business*, 38(1), 19–42. <https://doi.org/10.22146/jieb.v38i1.4394>
- Ruansyah, M. A., Hadi, M. Z., Iqbal, M., Juniwati, J., & Aulia, L. (2024). Designing an analytics dashboard for knowledge extraction in the retail industry using descriptive and predictive analytics. *Journal Industrial Services*, 10(2), 242. <https://doi.org/10.62870/jiss.v10i2.27900>
- Sanabia-Lizarraga, K. G., Carballo-Mendivil, B., Arellano-González, A., & Bueno-Solano, A. (2024). Business Intelligence for Agricultural Foreign Trade: Design and Application of Power BI Dashboard. *Sustainability (Switzerland)*, 16(21). <https://doi.org/10.3390/su16219576>
- Siljanoska, T., Savoska, S., & Jolevski, I. (n.d.). *Integrating Python Into Power BI for Analyzing and Predicting Digital Development : Case Study – Balkan Countries*. 4523, 55–69.
- Square, D. (2025). *The Ultimate Microsoft Power BI Guide for SMEs (2025)*. Dynamic Square. https://www.dynamicssquare.com/guides/power-bi-guide-for-sme/?utm_source=chatgpt.com
- Truong, N. X. (2022). Factors Affecting Big Data Adoption: An Empirical Study in Small and Medium Enterprises in Vietnam. *International Journal of Asian Business and Information Management*, 13(1), 1–21. <https://doi.org/10.4018/IJABIM.315825>
- Yusuf Iskandar, Heliani, Jaman, U. B., & Ardhiyansyah, A. (2023). Analyzing the Relationship Between Technology Adoption and Business Performance the Digital Age in SMEs in Indonesia. *Eastasouth Proceeding of Nature, Science, and Technology (EPNST)*. <https://doi.org/https://doi.org/10.58812/asst.v1i01.6>