



رواد مصر الرقمية

YAT402A – Group 3

Store Sales Analysis



Team Members

1	Arwa Saeed Sayed
2	Moustafa Yehia Abdallah
3	Ahmed Sayed Mahmoud
4	Mohamed Hammam Ibrahim
5	Mohamed Abdallah Abdo

Project Proposal: Store Sales Analysis

1) Overview

This project analyzes supermarket sales data using Excel and Power BI to uncover key business insights. It explores customer behavior, product performance, branch-wise sales, and payment trends. By identifying peak sales periods and revenue drivers, the analysis helps optimize pricing, stock management, and customer engagement strategies.

2) Objectives

- a. Analyze sales trends and customer behavior.
- b. Assess branch and product performance.
- c. Identify peak sales periods and revenue drivers.
- d. Provide data-driven insights for inventory and marketing.
- f. Visualize key metrics for better decision-making.

3) The Scope

The scope of this project involves examining supermarket sales data to extract meaningful insights that can support business decision-making. It includes analyzing customer demographics, sales distribution across branches, and product category performance. The project also explores revenue trends and transaction patterns using data analysis tools. However, it does not cover online sales, economic forecasting, or advanced machine learning models, keeping the focus on in-store sales insights.

- In-Scope:

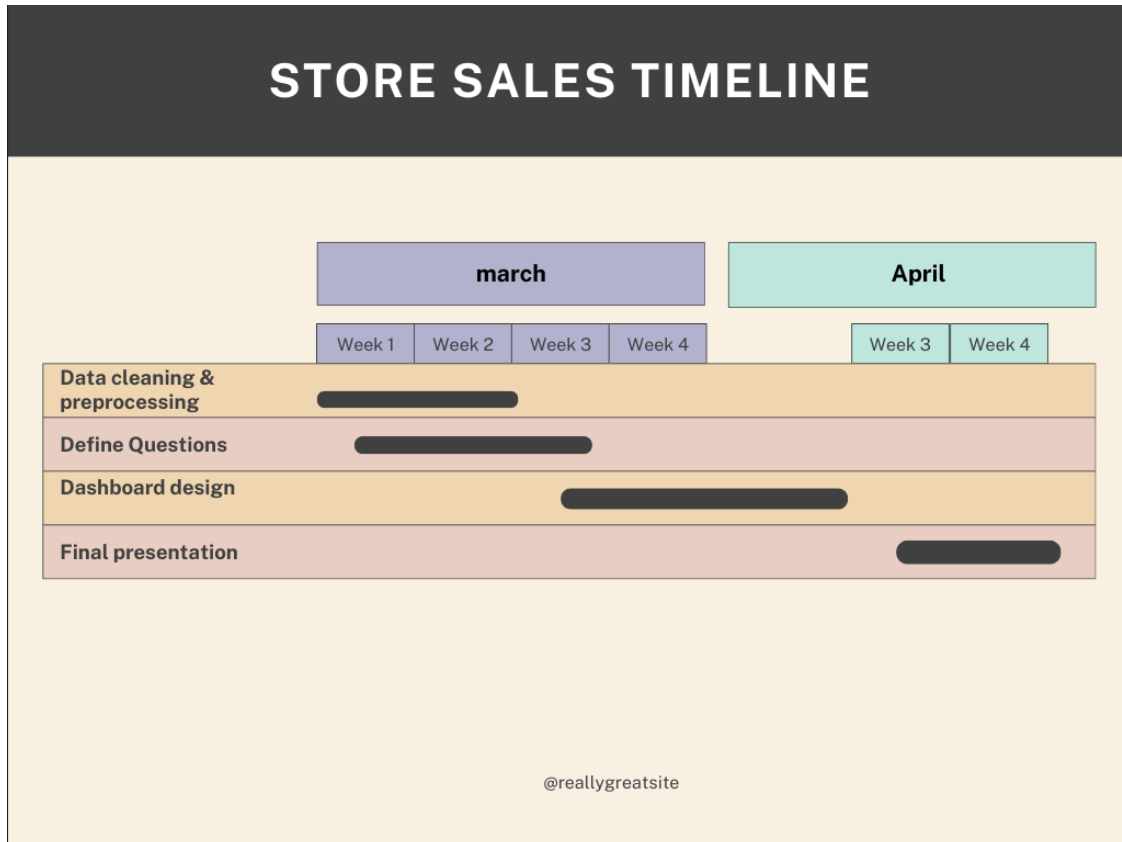
- a. Studying supermarket sales to identify key patterns and trends.
- b. Analyzing customer types, purchase behavior, and sales distribution.
- c. Evaluating the performance of different product categories and branches.
- d. Examining payment methods and their impact on sales.
- f. Using data visualization tools to present insights effectively.

- Out of Scope:

- a. Online sales and e-commerce transactions.
- b. External economic factors affecting sales.
- c. Advanced predictive modeling and machine learning.
- d. Detailed financial analysis beyond revenue and gross income.
- f. Supply chain and inventory management.

Project Plan

1)Timeline & Milestones



Week 1: Data Cleaning & Preprocessing

Tasks:

- Clean and preprocess the dataset to ensure accuracy and consistency.
- Handle missing values, duplicates, and incorrect data entries.
- Format the dataset for seamless analysis in Power BI.

Deliverables:

- **Cleaned and structured dataset ready for analysis.**
- Documentation of preprocessing steps and data handling techniques.
- Summary report detailing data quality issues and resolutions.
- Initial exploratory data analysis (EDA) report.

Week 2: Analysis Questions Development**Tasks:**

- Identify key business questions that can be answered using the dataset.
- Define relevant KPIs for sales performance evaluation.
- Categorize analysis focus areas (e.g., customer behavior, product performance, regional sales trends).

Deliverables:

- A well-defined list of analysis questions aligned with business needs.
- Justification for each question and how it contributes to decision-making
- A report outlining potential insights derived from the dataset.
- Visualization mockups for key metrics.

Week 3: Dashboard Development

Tasks:

- Create a Power BI dashboard to visualize key insights.
- Design interactive charts and graphs for easy data interpretation.
- Ensure the dashboard is user-friendly and provides clear, actionable insights.

Deliverables:

- Fully functional Power BI dashboard displaying sales insights.
- Documentation on dashboard structure and how to interpret visualizations.
- User guide on navigating and using the dashboard effectively.
- Performance evaluation report for dashboard accuracy.

Week 4: Final Report & Presentation

Tasks:

- Prepare a comprehensive report summarizing the project findings.
- Develop a presentation to communicate key insights and recommendations.
- Ensure the report and presentation are well-structured and professional.

Deliverables:

- Final project report including methodology, analysis, and insights.
- PowerPoint presentation summarizing key findings.
- Recommendations for business improvements based on data insights.

Roles	Responsibilities	Tools Used
Arwa Saeed Sayed	Design and structure the Power BI dashboard for visualizing insights	Power BI
Ahmed Sayed Mahmoud	Handle data cleaning, transformation data and providing insights	Excel, Power BI
Moustafa Yehia Abdallah	Analyze data, define key analytical questions, and generate meaningful insights	Excel, Power BI
Mohamed Hammam Ibrahim	Develop and optimize DAX formulas for calculated measures and KPIs.	Power BI
Mohamed Abdallah Abdo	Develop and optimize DAX formulas and presenting structured report	Power BI, Word, Powerpoint

Risk Assessment & Mitigation Plan

a. Data Quality Issues

- *Risk:* Missing values, duplicate records, or incorrect entries may impact analysis.
- *Mitigation:* Perform thorough data cleaning, apply validation checks, and document preprocessing steps.

b. Performance Limitations in Power BI

- *Risk:* Large datasets may slow down report generation and dashboard performance.
- *Mitigation:* Optimize data model, use **aggregations**, and apply **efficient DAX calculations**.

c. Unclear Business Questions

- *Risk:* Misalignment between analysis focus and business needs may result in irrelevant insights.
- *Mitigation:* Conduct stakeholder discussions early to define clear analysis objectives.

d. Dashboard Usability Challenges

- *Risk:* Users may find the dashboard difficult to navigate or interpret.
- *Mitigation:* Design an intuitive layout, provide a **user guide**, and gather feedback for improvements.

f. Time Constraints

- *Risk:* Project milestones may be delayed due to unforeseen issues.
- *Mitigation:* Establish a **realistic timeline**, prioritize critical tasks, and track progress weekly.

Key Performance Indicators

Week 1: Data Cleaning & Preprocessing

- **Data Accuracy Rate:** % of cleaned data without errors or inconsistencies.
- **Missing Data Handling Efficiency:** % of missing values resolved.
- **Data Processing Time:** Time taken to clean and preprocess the dataset.

Week 2: Analysis Questions Development

- **Number of Business Questions Identified:** Total relevant analysis questions formulated.
- **KPI Definition Accuracy:** % of KPIs aligned with business objectives.

- **Analysis Readiness Score:** Evaluation of dataset preparedness for insights extraction.

Week 3: Dashboard Development

- **Dashboard Functionality Score:** % of implemented features meeting requirements.
- **Visualization Clarity Rating:** User feedback on dashboard interpretability.
- **Performance Efficiency:** Dashboard loading time and responsiveness.

Week 4: Final Report & Presentation

- **Report Completeness:** % of project aspects covered in the report.
- **Stakeholder Satisfaction Score:** Feedback rating from decision-makers.
- **Presentation Effectiveness:** Engagement level and clarity of key insights.

Requirements Gathering

A. Stakeholder Analysis

1. Project Sponsor (The One Who Needs the Analysis)

- **What They Need:**

- Clean, structured sales data that is accurate and reliable.
- Insights that support decision-making rather than just raw numbers.
- A well-organized and interactive dashboard that simplifies data interpretation.
- A concise and actionable report highlighting key findings and recommendations.

- **How We Deliver:**

- Implement thorough data cleaning processes to eliminate errors and inconsistencies.
- Define key analytical questions to drive meaningful insights.
- Design a user-friendly and interactive dashboard with clear navigation.
- Provide a structured report focusing on essential takeaways without unnecessary details.

2. Data Analysts & Project Team (Us, the Ones Doing the Work)

- **What We Need:**

- A clearly defined project scope to understand the business problem.
- High-quality data that minimizes the need for extensive cleaning.
- The right analytical tools (e.g., Power BI) to enhance efficiency.
- A structured workflow to ensure an organized and systematic approach.

- **How We Make It Happen:**

- Establish clear analysis objectives to prevent wasted effort.
- Maintain detailed documentation of data cleaning and transformation steps for transparency.
- Utilize Power BI for automation and visualization instead of manual reporting.
- Distribute tasks efficiently to streamline workflow and improve productivity.

3. End Users (The Ones Who Use the Dashboard & Insights)

- **What They Need:**

- A dashboard that is intuitive and provides valuable insights at a glance.
- Quick access to key sales trends and performance indicators.

- Filters and drill-down capabilities to explore specific data segments.
- Actionable recommendations rather than overwhelming data.
- **How We Make It Work for Them:**
 - Ensure the dashboard has a clean and user-friendly design.
 - Highlight key performance metrics for quick insights.
 - Integrate filtering options to customize data views.
 - Summarize findings with clear recommendations for strategic action.

4. Business Decision Makers (Sales Managers, Executives, etc.)

- **What They Need:**
 - Insights into revenue trends, sales performance, and customer behavior.
 - The ability to track key performance indicators (KPIs) efficiently.
 - Data-driven recommendations to enhance sales strategy and operations.
 - A straightforward summary with actionable insights.
- **How We Help Them:**
 - Focus the analysis on real-world business decisions and operational needs.
 - Incorporate sales trends, customer segments, and performance metrics.

- Present insights in an accessible format to support informed decision-making.
- Provide scenario-based recommendations for strategic improvements.

B. User Stories & Use Cases

1. Project Sponsor (The One Who Needs the Analysis)

- **User Story:** As a project sponsor, I want a cleaned dataset and an interactive dashboard so I can quickly understand sales trends and make data-driven decisions.
- **Use Case:**
 - Opens the dashboard to review key sales performance indicators.
 - Uses insights to assess sales efficiency and market trends.
 - Prepares reports for stakeholders using summarized insights.

2. End Users (The People Using the Dashboard & Insights)

- **User Story:** As an end user, I want an easy-to-navigate sales dashboard so I can access relevant data without requiring technical expertise.
- **Use Case:**
 - Accesses the dashboard to check real-time or historical sales data.

- Uses filters to analyze sales trends by region, product, or time period.
- Exports relevant reports for internal presentations.

3. Business Decision Makers (Sales Managers, Executives, etc.)

- **User Story:** As a sales manager, I want to analyze revenue trends and customer behavior so I can optimize sales strategies and improve profitability.
- **Use Case:**
 - Reviews sales performance by product category and region.
 - Identifies top-performing sales channels and customer segments.
 - Uses predictive analytics to adjust future sales strategies.

C. Functional Requirements

1. Data Cleaning & Preprocessing

- Ensure all transaction IDs are unique and correctly formatted.
- Standardize date and time fields (e.g., Purchase Date, Transaction Time) to ensure consistency.
- Handle missing or incorrect values in fields such as Payment Method and Customer Category.
- Correctly classify and label product categories and sales channels.

- Remove duplicate entries and standardize customer demographic data.

2. Data Analysis & Metrics

- The system must provide key sales insights, including:
 - **Sales & Revenue Trends:**
 - Total revenue over time.
 - Average sales value by product category and region.
 - Most commonly used payment methods.
 - **Customer Behavior:**
 - Most frequent buyers and purchasing trends.
 - Peak sales periods and seasonal trends.
 - Preferred sales channels (e.g., online vs. in-store).
 - **Sales Performance:**
 - Conversion rates by marketing campaign.
 - Sales team performance based on targets and actual sales.
 - Customer retention and repeat purchase rates.

3. Interactive Dashboard

- The dashboard must enable users to:
 - View high-level sales performance metrics.
 - Filter data by date range, product category, and region.

- Compare current vs. previous sales periods.
- Drill down into specific sales transactions for deeper insights.
- Must include:
 - Clear visualizations such as trendlines, heatmaps, and bar charts.
 - Customer segmentation analysis for targeted marketing strategies.
 - Real-time sales tracking for immediate insights.

4. Reporting & Exporting

- Users must be able to:
 - Download reports summarizing key sales insights.
 - Export charts and tables for presentations and business reviews.
 - Generate reports on customer purchasing behavior.

5. User Experience

- The Dashboard should have:
 - A clean and intuitive interface for ease of use.
 - Fast data filtering and loading speeds.
 - Clear tooltips and labels to guide users.

D. Non-Functional Requirements

1. Performance (Speed & Efficiency)

- The system should load and process data efficiently without noticeable delays.
- It should handle large datasets without affecting performance.
- All interactions, such as filtering and navigation, should be smooth and responsive.

2. Security (Data Protection)

- Access should be restricted to authorized users.
- Data should be protected against accidental loss or unauthorized modifications.
- Customer and transaction data should be kept confidential.

3. Usability (User-Friendly Design)

- The system should be intuitive and require minimal training.
- Labels and instructions should be clear and helpful.
- The layout should be structured for quick and easy access to relevant information.

4. Reliability (System Stability)

- The system should operate consistently without frequent crashes or errors.
- It should handle various user inputs without breaking functionality.
- Mechanisms should be in place to recover from system failures without data loss.

Data Analysis & Design

- **Columns in the Dataset:**

- **Invoice ID:** Unique identifier for each transaction.
- **Branch:** The branch where the transaction occurred (A, B, C).
- **City:** The city corresponding to the branch (Yangon, Mandalay, Naypyitaw).
- **Customer type:** Type of customer (Member, Normal).
- **Gender:** Gender of the customer (Male, Female).
- **Product line:** Category of the product purchased (e.g., Health and beauty, Electronic accessories).
- **Unit price:** Price of one unit of the product.
- **Quantity:** Number of units purchased.
- **Tax 5%:** Tax applied to the transaction.
- **Total:** Total amount of the transaction (including tax).
- **Date:** Date of the transaction.
- **Time:** Time of the transaction.
- **Payment:** Payment method used (Cash, Credit card, Ewallet).
- **cogs:** Cost of goods sold.
- **gross margin percentage:** Gross margin percentage.
- **gross income:** Gross income from the transaction.
- **Rating:** Customer rating for the transaction (on a scale of 1-10).

- **Data Types:**

- **Numeric:** Unit price, Quantity, Tax 5%, Total, cogs, gross margin percentage, gross income, Rating.
- **Categorical:** Invoice ID, Branch, City, Customer type, Gender, Product line, Date, Time, Payment.

2. Key Analysis Questions

- **Sales Performance:**
 - What is the total sales revenue for each branch?
 - Which product line generates the highest revenue?
 - How does sales performance vary by city?
 - What is the trend in sales over time (daily, monthly)?
- **Customer Analysis:**
 - What is the distribution of customers by type (Member vs. Normal)?
 - How does gender influence purchasing behavior?
 - Which payment method is most popular among customers?
- **Product Analysis:**
 - Which product line has the highest and lowest unit prices?
 - What is the average quantity sold per product line?
 - How does the gross margin vary across product lines?
- **Geographical Analysis:**
 - Which city has the highest sales?
 - How does sales performance vary across branches?
- **Customer Satisfaction:**
 - What is the average customer rating for each branch?
 - How does customer rating vary by product line?
 - Are there any correlations between sales and customer ratings?

3. Insights and Visualizations

- **Sales Performance:**
 - **Total Sales by Branch:** We will use a bar chart to show total sales for each branch (A, B, C).
 - **Sales by Product Line:** We will use a pie chart or bar chart to show revenue distribution across product lines.
 - **Sales Trend Over Time:** We will use a line chart to visualize daily or monthly sales trends.
- **Customer Analysis:**
 - **Customer Type Distribution:** We will use a pie chart to show the proportion of Members vs. Normal customers.
 - **Gender Distribution:** We will use a bar chart to show the number of transactions by gender.
 - **Payment Method Popularity:** We will use a bar chart to show the most used payment methods.
- **Product Analysis:**
 - **Unit Price by Product Line:** We will use a bar chart to compare unit prices across product lines.
 - **Average Quantity Sold:** We will use a bar chart to show the average quantity sold per product line.
 - **Gross Margin by Product Line:** We will use a bar chart to compare gross margins across product lines.
- **Geographical Analysis:**
 - **Sales by City:** We will use a map visualization to show sales performance by city.
 - **Sales by Branch:** We will use a bar chart to compare sales across branches.

- **Customer Satisfaction:**
 - **Average Rating by Branch:** We will use a bar chart to show the average customer rating for each branch.
 - **Rating by Product Line:** We will use a bar chart to show average ratings for each product line.
 - **Correlation Between Sales and Ratings:** We will use a scatter plot to explore the relationship between sales and customer ratings.

4. Power BI Dashboard Layout

- **Page 1: Sales Overview**
 - Total Sales by Branch (Bar Chart)
 - Sales by Product Line (Pie Chart)
 - Sales Trend Over Time (Line Chart)
- **Page 2: Customer Insights**
 - Customer Type Distribution (Pie Chart)
 - Gender Distribution (Bar Chart)
 - Payment Method Popularity (Bar Chart)
- **Page 3: Product Analysis**
 - Unit Price by Product Line (Bar Chart)
 - Average Quantity Sold (Bar Chart)
 - Gross Margin by Product Line (Bar Chart)
- **Page 4: Geographical Analysis**
 - Sales by City (Map Visualization)
 - Sales by Branch (Bar Chart)
- **Page 5: Customer Satisfaction**

- Average Rating by Branch (Bar Chart)
- Rating by Product Line (Bar Chart)
- Correlation Between Sales and Ratings (Scatter Plot)

5. Additional

- **Filters:** Add filters for date, branch, city, and product line to allow users to drill down into specific data segments.
- **KPIs:** Include key performance indicators (KPIs) such as total sales, average rating, and gross margin percentage on the dashboard.

Entities and Attributes

1. Invoice

- **Attributes:** Invoice ID, Date, Time, Total, Tax 5%, cogs, gross margin percentage, gross income, Rating.
- **Primary Key:** Invoice ID.

2. Branch

- **Attributes:** Branch (A, B, C), City (Yangon, Mandalay, Naypyitaw).
- **Primary Key:** Branch.

3. Customer

- **Attributes:** Customer type (Member, Normal), Gender (Male, Female).
- **Primary Key:** Customer ID (can be derived or inferred from the dataset).

4. Product

- **Attributes:** Product line (e.g., Health and beauty, electronic accessories), Unit price.
- **Primary Key:** Product ID (can be derived or inferred from the dataset).

5. Payment

- **Attributes:** Payment method (Cash, Credit card, E-wallet).
- **Primary Key:** Payment ID (can be derived or inferred from the dataset).

Relationships

1. Invoice - Branch:

- Each invoice is associated with one branch.
- Relationship: **1-to-Many** (One branch can have many invoices).

2. Invoice - Customer:

- Each invoice is associated with one customer.
- Relationship: **1-to-Many** (One customer can have many invoices).

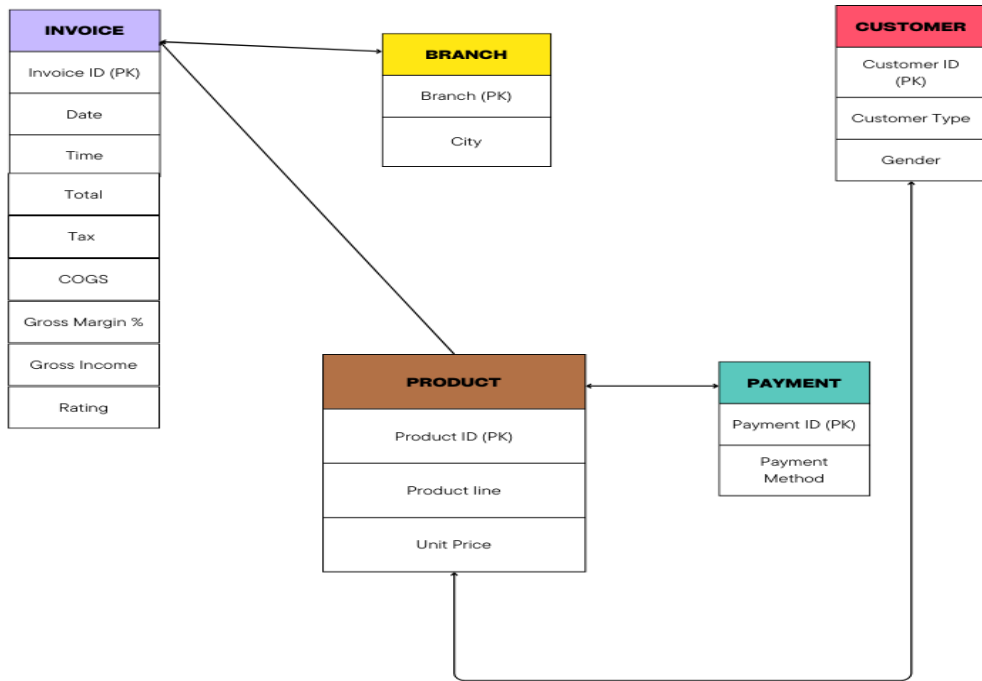
3. Invoice - Product:

- Each invoice can include multiple products.
- Relationship: **Many-to-Many** (An invoice can have multiple products, and a product can appear in multiple invoices).

4. Invoice - Payment:

- Each invoice is associated with one payment method.
- Relationship: **1-to-Many** (One payment method can be used in many invoices).

ER Diagram



Explanation of Relationships

1. Invoice - Branch:

- Each invoice is linked to a specific branch. For example, Invoice ID 750-67-8428 is associated with Branch A (Yangon).

2. Invoice - Customer:

- Each invoice is linked to a specific customer. For example, Invoice ID 750-67-8428 is associated with a Female Member customer.

3. Invoice - Product:

- Each invoice can include multiple products. For example, Invoice ID 750-67-8428 includes products from the "Health and beauty" product line.

4. Invoice - Payment:

- Each invoice is linked to a specific payment method. For example, Invoice ID 750-67-8428 uses the "E-wallet" payment method.