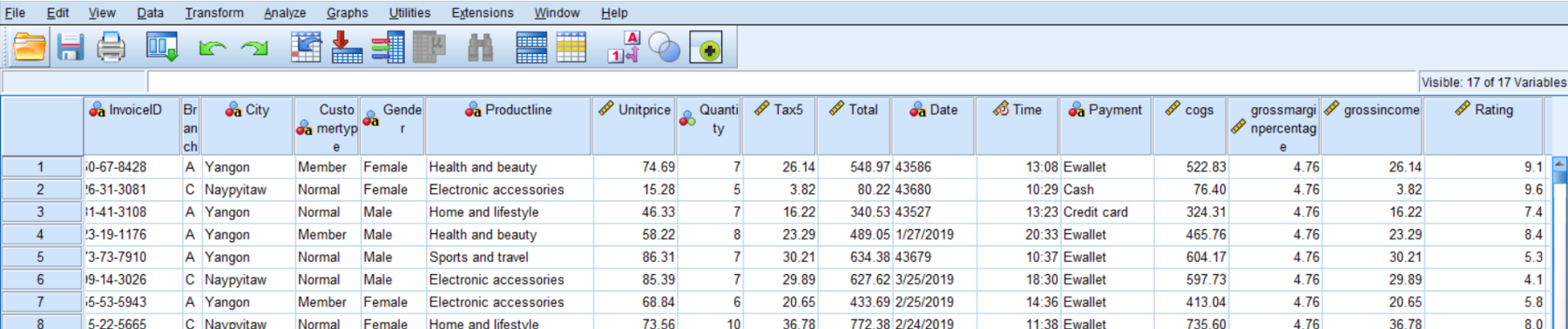
**Supermarket sales**

Historical record of sales data in 3 different supermarkets

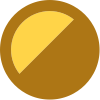


**Supermarket sales**

arrow\_drop\_up2450

**New Notebook**

file\_download**Download (37 kB)**arrow\_drop\_down



more\_vert

[Data Card](https://www.kaggle.com/datasets/aungpyaeap/supermarket-sales/data)[Code (392)](https://www.kaggle.com/datasets/aungpyaeap/supermarket-sales/code)[Discussion (26)](https://www.kaggle.com/datasets/aungpyaeap/supermarket-sales/discussion)[Suggestions (0)](https://www.kaggle.com/datasets/aungpyaeap/supermarket-sales/suggestions)

**About Dataset**

**Context**  
  
The growth of supermarkets in most populated cities are increasing and market competitions are also high. The dataset is one of the historical sales of supermarket company which has recorded in 3 different branches for 3 months data. Predictive data analytics methods are easy to apply with this dataset.

**Attribute information**  
  
Invoice id: Computer generated sales slip invoice identification number  
  
Branch: Branch of supercenter (3 branches are available identified by A, B and C).  
  
City: Location of supercenters  
  
Customer type: Type of customers, recorded by Members for customers using member card and Normal for without member card.  
  
Gender: Gender type of customer  
  
Product line: General item categorization groups - Electronic accessories, Fashion accessories, Food and beverages, Health and beauty, Home and lifestyle, Sports and travel  
  
Unit price: Price of each product in $  
  
Quantity: Number of products purchased by customer  
  
Tax: 5% tax fee for customer buying  
  
Total: Total price including tax  
  
Date: Date of purchase (Record available from January 2019 to March 2019)  
  
Time: Purchase time (10am to 9pm)  
  
Payment: Payment used by customer for purchase (3 methods are available – Cash, Credit card and Ewallet)  
  
COGS: Cost of goods sold  
  
Gross margin percentage: Gross margin percentage  
  
Gross income: Gross income  
  
Rating: Customer stratification rating on their overall shopping experience (On a scale of 1 to 10)

**Acknowledgements**  
  
Thanks to all who take time and energy to perform Kernels with this dataset and reviewers.

**Purpose**  
  
This dataset can be used for predictive data analytics purpose.

expand\_less**View less**

**Usability**

info

8.82

**License**

Other (specified in description)

**Expected update frequency**

Not specified

**Tags**

[Business](https://www.kaggle.com/datasets?tags=11102-Business)[Data Visualization](https://www.kaggle.com/datasets?tags=13208-Data+Visualization)[Economics](https://www.kaggle.com/datasets?tags=11205-Economics)

1. What is the relationship between the price of a product and the quantity sold? Do more expensive products tend to sell better or worse than cheaper ones?
2. Are there any differences in sales patterns between different cities or branches? Which city or branch is the most profitable?
3. How do sales vary by day of the week or time of day? Are there certain times that are busier than others?
4. Are there any differences in sales patterns between different customer types (e.g. members vs. non-members)?
5. Which product lines are the most popular, and which ones have the highest profit margins?
6. Is there a correlation between the rating of a product and its sales? Do customers tend to buy products with higher ratings more often?

**READ ME FOR PROJECT**

**# Supermarket Sales Analysis**

**## Project Overview**

**A detailed analysis of historical sales data from three different supermarket branches over a three-month period. The project involves data cleaning, exploratory data analysis (EDA), hypothesis testing, and visualization to uncover insights into sales performance, customer behavior, and pricing strategies.**

**## Dataset**

**- \*\*Source\*\*: [Kaggle](https://www.kaggle.com/datasets)**

**- \*\*Description\*\*: Historical sales data from three supermarket branches, including attributes such as invoice ID, branch, city, customer type, gender, product line, unit price, quantity, tax, total, date, time, payment, COGS, gross margin percentage, and gross income.**

**## Objectives**

**1. Data Analysis and Cleaning**

**2. Exploratory Data Analysis (EDA)**

**3. Research Questions**

**4. Hypothesis Testing**

**## Data Analysis and Cleaning**

**- Handled missing values and outliers.**

**- Preprocessed data for analysis.**

**## Exploratory Data Analysis (EDA)**

**- Created various visualizations to understand data patterns and trends.**

**## Research Questions**

**- What is the relationship between customer ratings and total sales?**

**- What is the relationship between product price and quantity sold?**

**## Hypothesis Testing**

**- Conducted hypothesis testing to assess differences in customer purchasing behavior.**

**## Usage**

**1. Clone this repository.**

**2. Install dependencies using `pip install -r requirements.txt`.**

**3. Run the analysis script or Jupyter Notebook.**

**## Installation**

**- Python 3.x**

**- Pandas**

**- NumPy**

**- Matplotlib**

**- Scipy**

**## Contributing**

**Contributions are welcome! Please submit a pull request or open an issue to discuss any improvements.**

**## License**

**This project is licensed under the MIT License - see the [LICENSE](LICENSE) file for details.**

**Option 1: Uploading a folder**

To upload a folder containing all your files, you can use the following commands:

1. Navigate to the parent directory of your project folder using **cd**.
2. Initialize a new Git repository by running **git init**.
3. Add the entire project folder to the Git repository by running **git add <project\_folder\_name>**.
4. Commit the changes with a meaningful commit message by running **git commit -m "Initial commit of Power BI project"**.
5. Create a new repository on GitHub and copy the repository URL.
6. Link the local repository to the GitHub repository by running **git remote add origin <your-github-repo-url>**.
7. Push the changes to the GitHub repository by running **git push -u origin master**.

This will upload the entire folder, including all files and subfolders, to your GitHub repository.

**Example:**

bash

VerifyOpen In EditorEditCopy code

1cd path/to/parent/directory

2git init

3git add ecommerce-sales-dashboard

4git commit -m "Initial commit of Power BI project"

5git remote add origin https://github.com/your-username/ecommerce-sales-dashboard.git

6git push -u origin master