

# Find a Dataset

Mukund Dhar  
UCF ID: 5499369

Dataset: Supermarket sales

Link: <https://www.kaggle.com/datasets/aungpyaeap/supermarket-sales>

This dataset includes information about supermarket sales from a retail chain in different stores and for different products. This dataset is interesting because it allows me to explore the factors that influence supermarket sales, which is a key business metric for retail chains. By analyzing this dataset, I can identify which factors are most strongly associated with sales amounts and make recommendations for marketing strategies that could improve sales. As someone who is interested in predictive and business analytics, this dataset aligns with my professional interests. By analyzing this dataset, I hope to gain insights into which factors have the greatest impact on supermarket sales and identify various patterns and trends to understand consumer behavior.

## List of features of the dataset and the response variable:

Response variable: Total (Ratio/Interval)

Features:

Invoice ID (Nominal)  
Branch (Nominal)  
City (Nominal)  
Product Line (Nominal)  
Gross Margin Percentage (Ratio/Interval)  
Gross Income (Ratio/Interval)  
Unit Price (Ratio/Interval)  
Quantity (Ratio/Interval)  
Tax (Ratio/Interval)  
Date (Ratio/Interval)  
Time (Ratio/Interval)  
Payment (Nominal)  
COGS (Ratio/Interval)  
Customer Type (Nominal)  
Rating (Ratio/Interval)  
Gender (Nominal)

## Descriptive 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

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)

Total: Total price including tax