# Supermart Grocery Sales

**AACS1573 Introduction to Data Science**

Presented by:
Yong Vin Sen
Chua Teng Hui
Wong Wai Kin
Siow Jun Jie

# Table of contents

1.0 Introduction

2.0 Case Study

3.0 Data Science Process

4.0 Result & Discussion

5.0 Conclusion

# 1.0 Introduction

# 1.1 Introduction to Data Science

- A branch of study that combines subject-matter expertise, programming prowess, and understanding of math and statistics.

- The insights produced by these technologies may then be transformed into real commercial value by analysts and business users.

# 1.2 Significance of Data Science

- To understand the pattern of datasets.

- Allows researchers to draw insights from the data that has been analysed.

- Make predictions from the data and know what is the trend in the future.

- Data science is useful in many industries.

# 2.0 Case Study

# 2.0 Case Study

**Case Study:** Supermart Grocery Sales

- Many aspects that will affect supermarket grocery sales.

**Purpose:** Find out the factors that affected the supermart grocery sales.

**Hypothesis factor:**
- Customers' region
- Supermarts' located city
- Profit margin of the supermarts' city
- Category and subcategory of goods
- Customers' order date
- Order's discount

# 3.0 Data Science Process

# 3.0 Data Science Process

## 3.1 Data Preparation

- Reading the data from various sources.
- Convert the data into the same format type.
- Cleansing the data Remove corrupt or problematic data.

## 3.2 Data Exploration

- Analyse data and get some useful data.
- Find out the similarities, differences, outliers.
- Hidden data can be figure out.

## 3.3 Data Representation

- Allows us to transform raw data into a dataset.
- Interpret the data more easily and accurately.

# 3.0 Data Science Process

## 3.4 Data Discovery

- Conducting initial data analysis.
- Collecting and evaluating data from different sources.
- To filter out less significant correlations.

## 3.5 Learning From Data

- Vital in data science process.
- Learn types of useful data from the datasets.
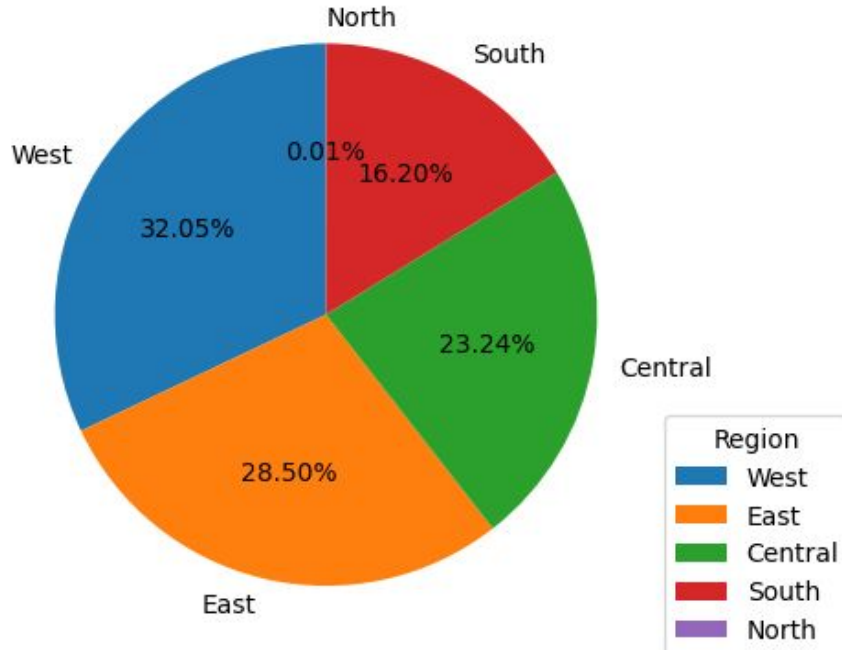- Can totally understand what the datasets are trying to interpret.
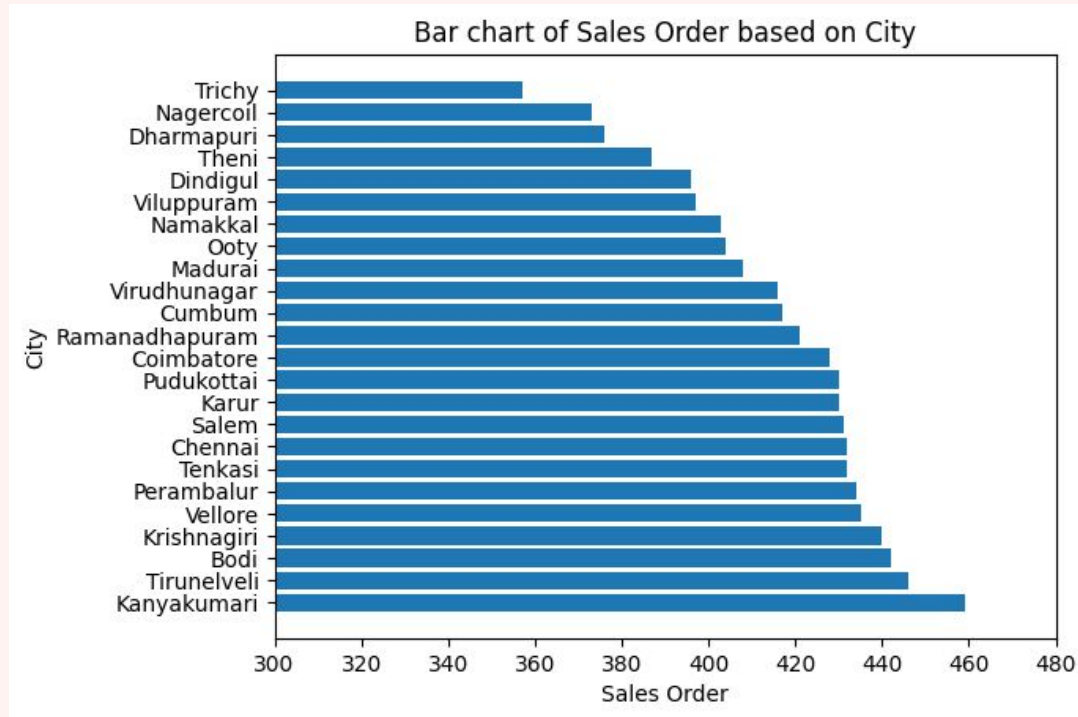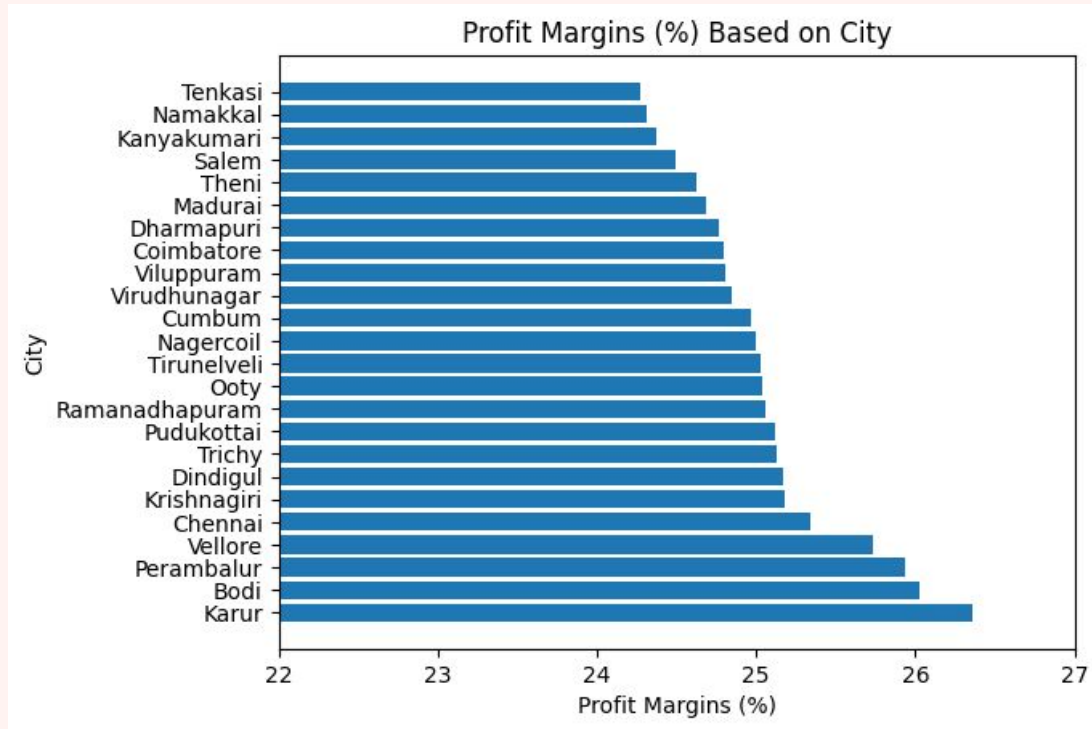
# 4.0 Result & Discussion

# 4.1 Customers' Region



Pie Chart of the Percentages of Sales Order by Customers' Region

# 4.2 Sales Order Based on City
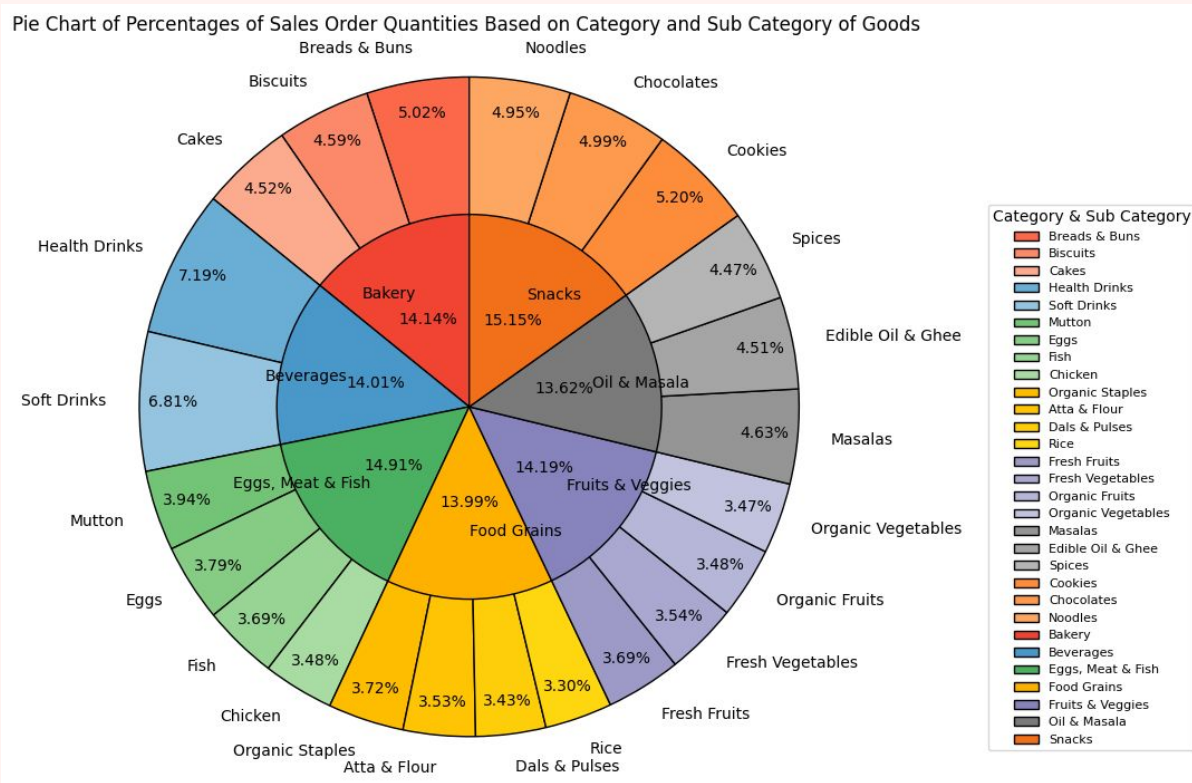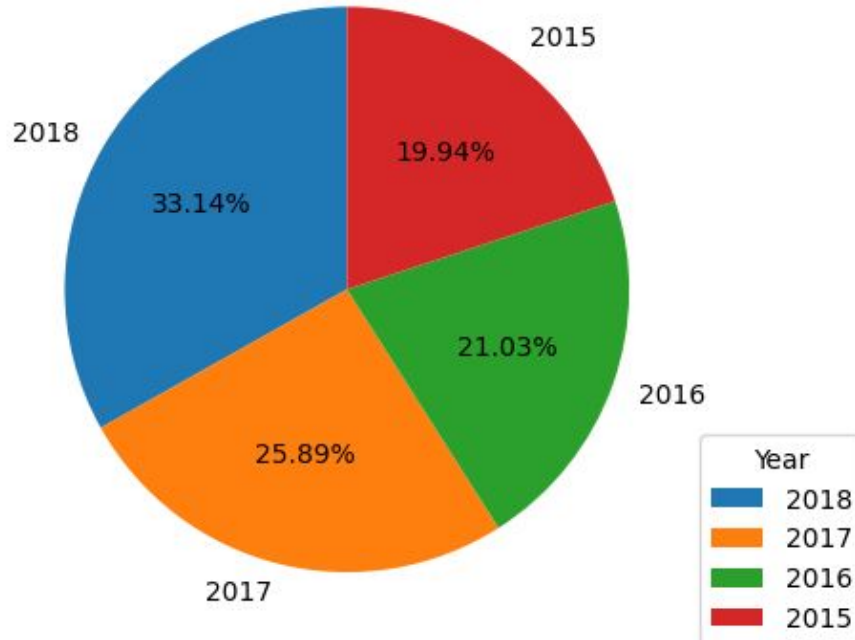


Bar chart of Sales Order based on City

# 4.3 Profit Margins Based on City



Profit Margins (%) Based on City

# 4.4 Category and Sub Category



Pie Chart of Percentages of Sales Order Quantities Based on Category and Sub Category of Goods

# 4.5 Order Date Based on Year



Pie Chart of Percentages of Sales Order by Different Years
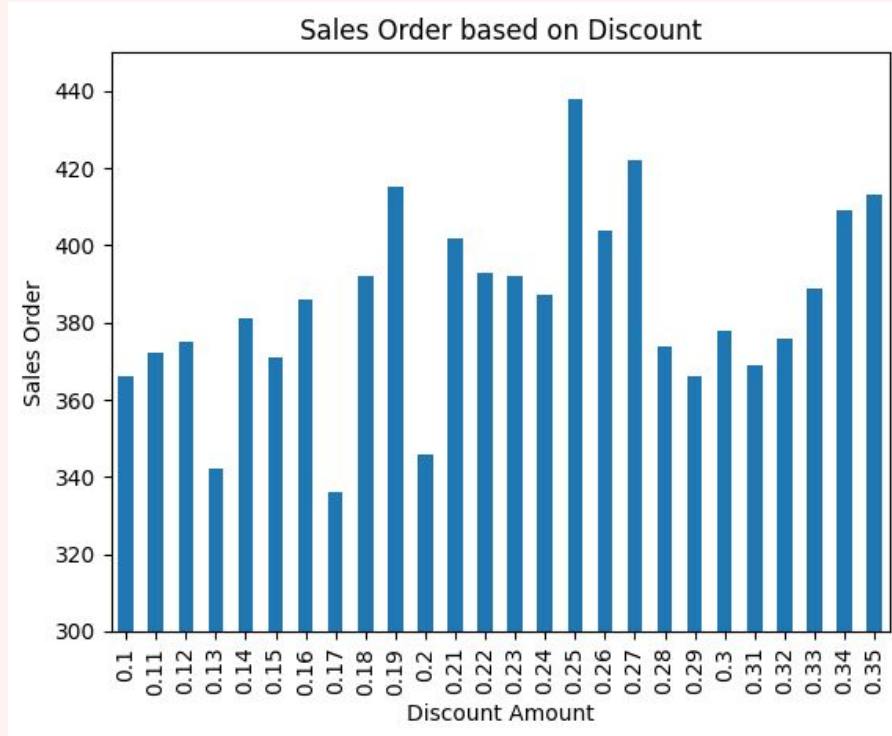
| Year |
| --- |
| 2018 |
| 2017 |
| 2016 |
| 2015 |

2018 — 33.14%
2015 — 19.94%
2016 — 21.03%
2017 — 25.89%

# 4.6 Discount Amount



Sales Order based on Discount

# 5.0 Conclusion

# 5.0 Conclusion

## Factors influence supermart sales:

- Region
- City
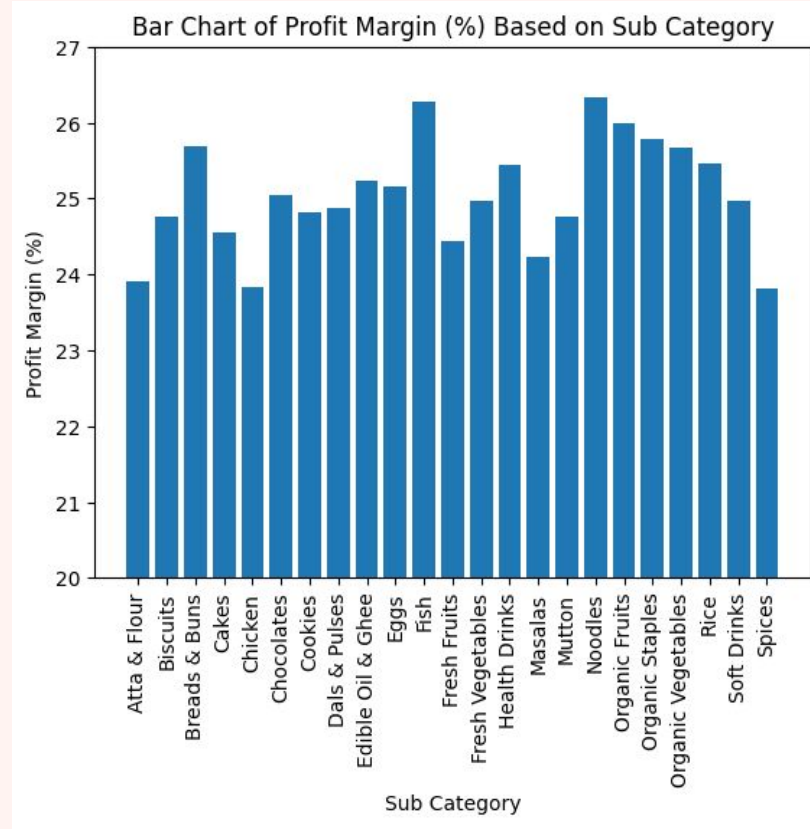- Category of goods
- Subcategory of goods

# 5.0 Conclusion

## Advantages:

- To prevent out of stock problems
- Increase the customer's satisfaction and retention towards the supermart.

## Disadvantages:

- The correlation between discount rate and sales orders is relatively weak.
- A large number of sales orders does not indicate that it will bring high profit to our supermart.

Bar Chart of Profit Margin (%) Based on Sub Category

# Thank You!