Supermarket Sales Data

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.

Format:

- Invoice id: Computer generated sales slip invoice identification number
- Branch: Branch of supercenter.
- 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
- Time: Purchase time
- Payment: Payment used by customer for purchase
- COGS: Cost of goods sold
- Gross margin percentage: Gross margin percentage
- Gross income: Gross income
- Rating: Customer stratification rating on their overall shopping experience

Task:

Tools Recommended- R, Python.

1. Calculate the customer stratification accurately.
2. Segment Gross income for Branch of Superstore.
3. Visualise correlation between City and Gross Income.
4. Can you predict the gender of customer by analysing data?
5. Predict the customer type by classification algorithm.