

Sales Performance Report

Contents

Introduction	2
Research Question	2
Descriptive Analysis	3
Correlation Matrix.....	6
Multiple Regression	6
➤ First Question :	6
➤ Second Question :	8
The decision.....	9
Conclusion	9
References & dataset.....	9

Introduction

This report should tell you the relationships between many variables in supermarket. As well as, Forecasting Quantity demanded based on unit price, taxes and other variables. Taking into consideration Changes in other variables can change the Quantity.

Giving an overview about customers gender, Products line which they interested in, preferences payment option, most active Branches. That can help in modifying targets of advertisements to increasing sales, Also achieve customer's satisfaction and Use the product performance report to determine which items are worth investing in and which ones shouldn't be re-ordered.

In this report, We tried to put hypothesis and determine if it true or not. We Analyze dataset Using R, PowerBI to summarize data and use Multiple linear Regression To determine the main factors which effect on our output.

Research Question

First, We will figure out if There is a relation between Quantity Demanded, Unit price and Taxes (Quantitative). Then, we will add Gender specifically "Males" and Payment Options "Credit Card, E-wallet" and then we will try to know the kind of this relationship.

$$Y = C + B_1 X_1 + B_2 X_2 + B_3 X_3 + B_4 X_4$$

$$\text{Quantity} = C + [B_1 \times \text{Unit price}] + [B_2 \times \text{Taxes}] + [B_3 \times \text{Gender (Male)}] + [B_4 \times \text{Payment Options "Credit Card"}] + [B_5 \times \text{Payment Options "E-wallet"}] + [B_6 \text{ Rating}]$$

For example :

- What are the potential determine affecting Quantity Demand?
- If we increase Unit Price or taxes is that will affect Quantity Demand or not?

Second, We will figure out if There is a relation between Gross income and Quantity, Unit price, Taxes, City, Gender, Payment option.

(City and Gross income)

Payment option & Gross income

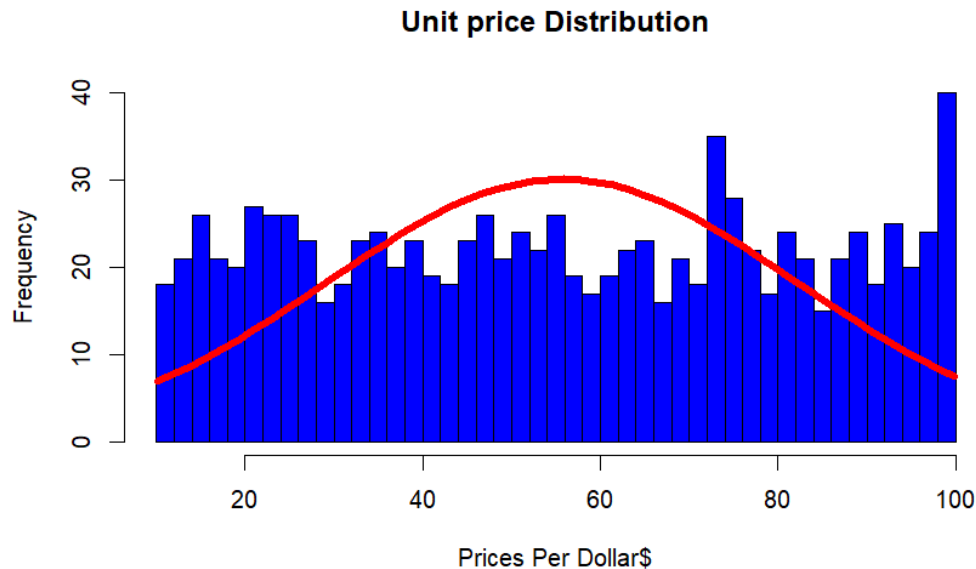
$$\text{Gross income} = C + [B_0 \text{ Quantity}] + [B_1 \times \text{Unit price}] + [B_2 \times \text{Taxes}] + [B_3 \times \text{Gender (Male)}] + [B_4 \times \text{Payment Options "Credit Card"}] + [B_5 \times \text{Payment Options "E-wallet"}]$$

Descriptive Analysis

1- Unit Price

Confidence Level(95.0%)	Count	Sum	Maximum	Minimum	Range	Skewness	Kurtosis	Sample Variance	Standard Deviation	Mode	Median	Standard Error	Mean
1.644115833	1000	55672	99.96	10.08	89.88	0.007077448	-1.218591428	701.9653313	26.49462835	83.77	55.23	0.837833713	55.67213

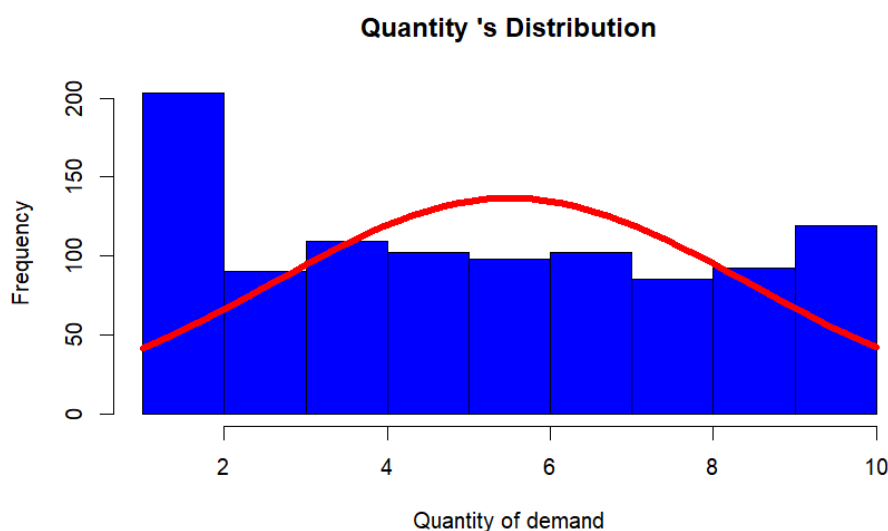
This data consists of a continuous variable and follow the normal distribution. We can see here that the mean is 55.67 , minimum is 10.08 and the maximum 99.96 . which means that the mean cluster between Min and Max “in the middle of the range “ for that no skewness 0.00707 .



2- Quantity

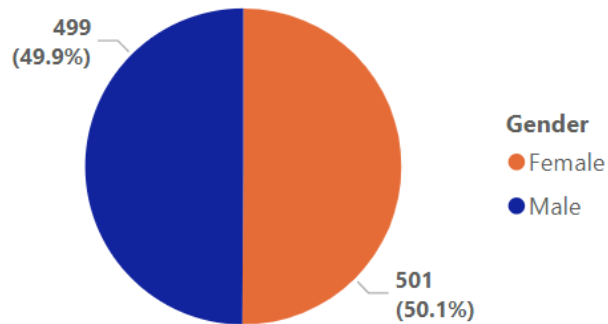
Confidence Level(95.0%)	Count	Sum	Maximum	Minimum	Range	Skewness	Kurtosis	Sample Variance	Standard Deviation	Mode	Median	Standard Error	Mean
0.181412567	1000	5510	10	1	9	0.012941048	-1.215547226	8.546446446	2.923430595	10	5	0.092446993	5.51

This data consists of a continuous variable and follow the normal distribution. We can see here that the mean is 5.51 , minimum is 1 and the maximum 10 . which means that the mean cluster between Min and Max “in the middle of the range “ .For that , there’s no skewness 0.0129410 .



3- Gender

Count of Gender by Gender

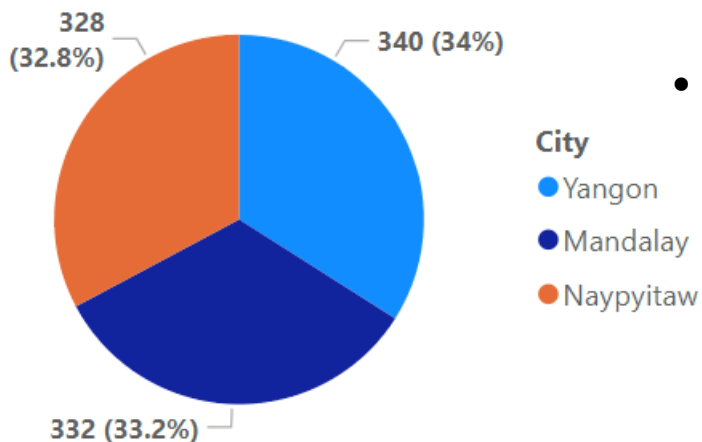


- We can notice that Count of Male and female are almost similar which female are greater than male by only 1 variable .

Total	Male	Female	Gender
1000	499	501	Count of Gender

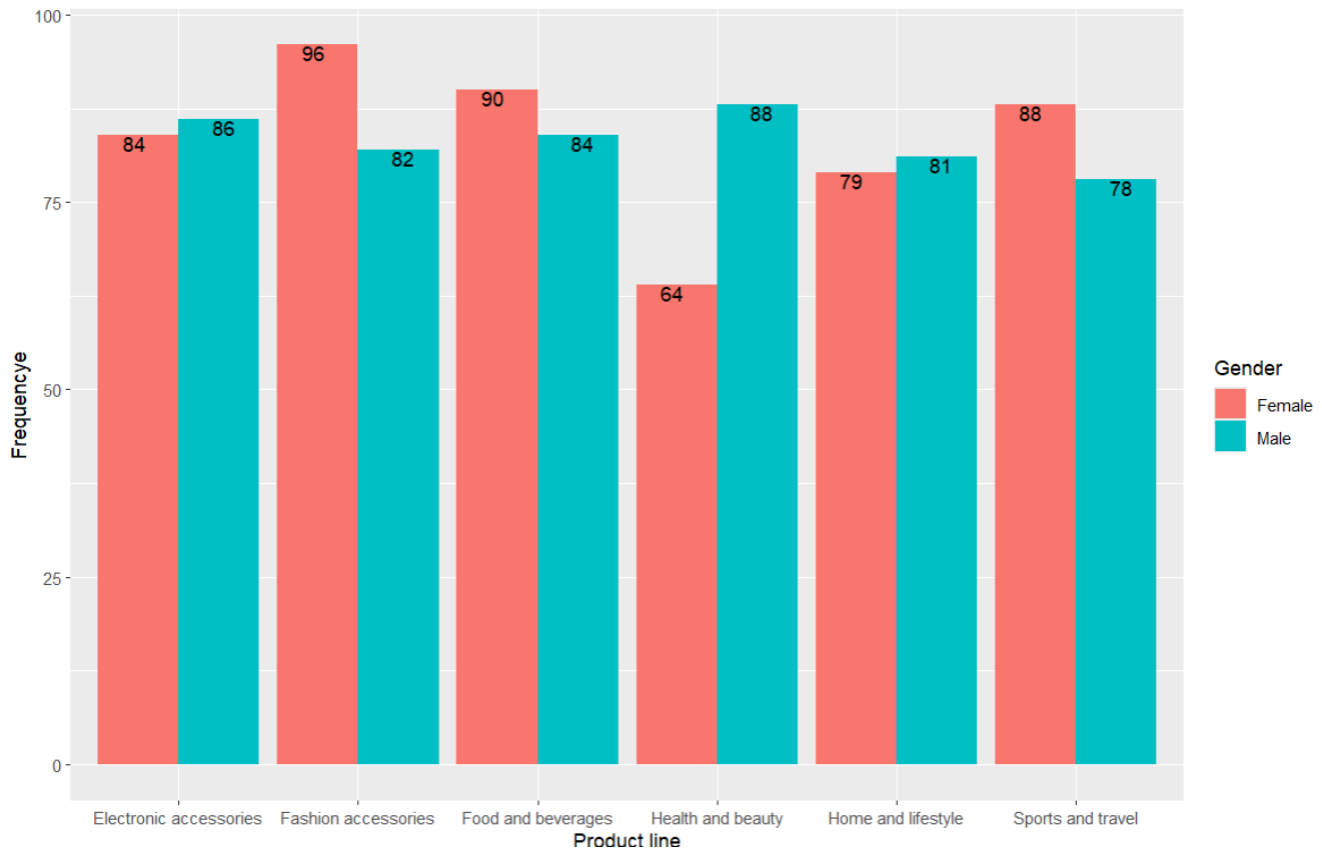
4- City

Count of City by City

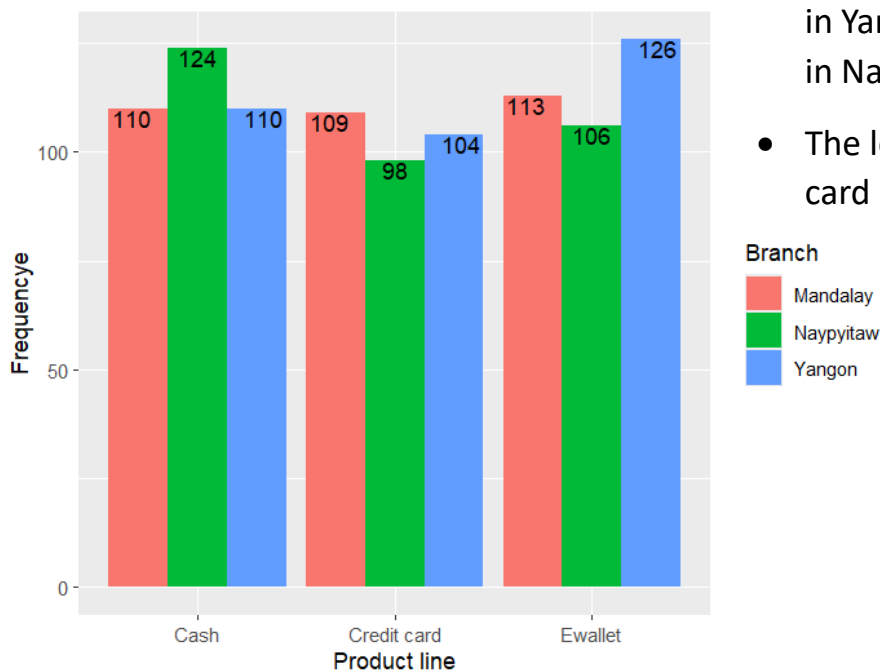


- We can see that Yangon city has the most percentage of variables (34%) .
- Mandalay has the second percentage (33.2%) .
- Naypyitaw has the lowest percentage (32.8 %)

Total	Yangon	Naypyitaw	Mandalay	City
1000	340	328	332	Count of City



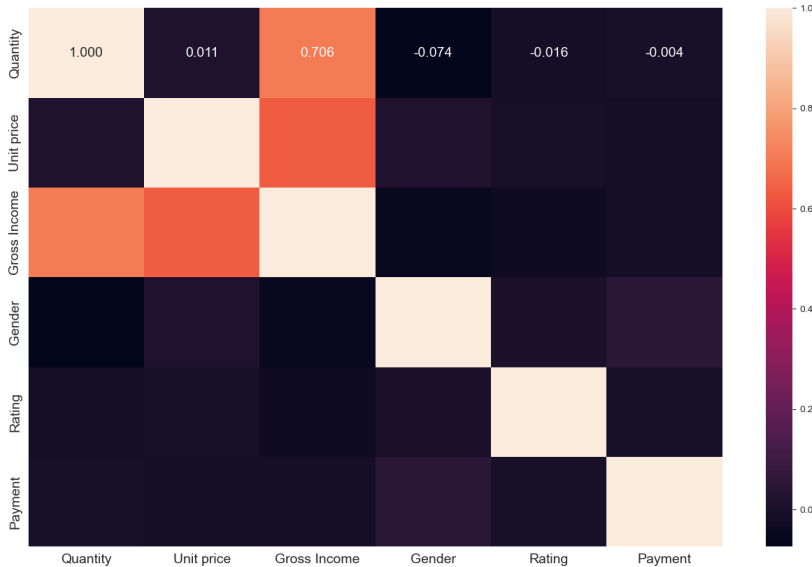
- Based on this figure we can see frequencies of product line based on Gender . As Electronic accessories the count of Male = 86 and female = 84 .
- The greatest Group in the product line is Females in Fashion accessories department which represent (96) .



- The highest Payment option is E-Wallet in Yangon Branch. The next one is Cash in Naypyitaw Branch.
- The lowest Payment option is Credit card in Naypyitaw Branch.

Correlation Matrix

	Quantity	Unit price	Gross Income	Gender	Rating	Payment
Quantity	1.000000	0.010778	0.705510	-0.074258	-0.015815	-0.003921
Unit price	0.010778	1.000000	0.633962	0.015445	-0.008778	-0.015941
Gross Income	0.705510	0.633962	1.000000	-0.049451	-0.036442	-0.012434
Gender	-0.074258	0.015445	-0.049451	1.000000	0.004800	0.044578
Rating	-0.015815	-0.008778	-0.036442	0.004800	1.000000	-0.005381
Payment	-0.003921	-0.015941	-0.012434	0.044578	-0.005381	1.000000



- Based on this correlation Matrix , We can see relations between different variables individually .
- For Example , There 's a positive relation between Quantity & Unit price (0.0107) . on the other side ,There's a negative relation between Quantity & Gender (-0.07425) .

Multiple Regression

➤ First Question :

Call:

```
lm(formula = Quantity ~ Unit.price + Tax.5. + Gender + Payment +
    Rating, data = df)
```

Residuals:

```
      Min       1Q   Median       3Q      Max
-3.8275 -0.6211 -0.0252  0.5809  3.7596
```

Coefficients:

```
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    5.326253   0.198197  26.873  <2e-16 ***
Unit.price     -0.080666   0.001942 -41.544  <2e-16 ***
Tax.5.         0.291996   0.004399  66.383  <2e-16 ***
GenderMale     -0.034269   0.079648  -0.430    0.667
PaymentCredit card -0.136580  0.098197  -1.391    0.165
PaymentEwallet -0.004777   0.095636  -0.050    0.960
Rating         0.035152   0.023102   1.522    0.128
---
```

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

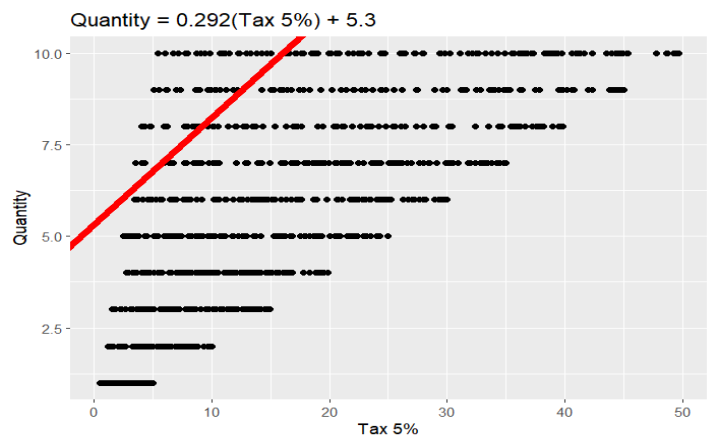
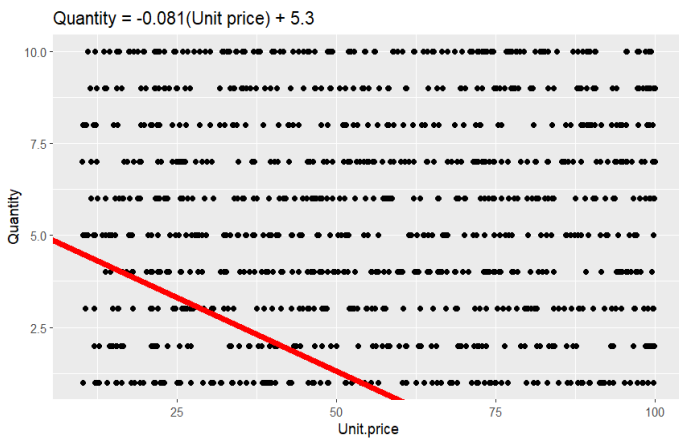
Residual standard error: 1.254 on 993 degrees of freedom

Multiple R-squared: 0.8172, Adjusted R-squared: 0.8161

F-statistic: 739.8 on 6 and 993 DF, p-value: < 2.2e-16

- We can see in this model that the function will be

$$\text{Quantity} = 5.3262 - 0.080666 (\text{Unit Price}) + 0.291996 (\text{Tax. 5\%}) - 0.034269 (\text{Gender "Male"}) - 0.136580 (\text{Payment "Credit card"}) - 0.004777 (\text{Payment "Ewallet"}) + 0.035152 (\text{Rating})$$
- From the output we will reject that there is relation between Quantity with Rating , Payment “wallet” , Payment “Credit Card “& Gender “Male” . So, P-value > 0.05 there ‘s no significant evidence that there is Relation .
- Also we can accept the relation between Quantity with Unit Price & Tax .5% .P-value < 0.05 so we accept the relation .
- We can see that R-Square (0.81) and it measures to a good model and it means that approximately (81.7 %) from the variance in dependent variable are explained by the independent .



- Based on the figure . If the increase Unit Price by 1 \$, Quantity will decrease By 0.08066 % for Unit . Also If the Taxes increases by 1 \$ The Quantity will increase by 0.291996 unit .
- For forecasting , We assume that the unit price = 25 \$ the Quantity demanded will be 3.275 ~ (3 units approximately)

➤ Second Question :

Call:

```
lm(formula = `Gross Income` ~ Quantity + `Unit price` + `Tax 5%` +  
Rating + Gender + Payment, data = df)
```

Residuals:

	Min	1Q	Median	3Q	Max
Residuals	-7.120e-15	-4.070e-16	-1.020e-16	1.670e-16	1.441e-13

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-2.461e-14	9.578e-16	-2.570e+01	<2e-16 ***
Quantity	4.267e-15	1.167e-16	3.656e+01	<2e-16 ***
`Unit price`	4.073e-16	1.181e-17	3.448e+01	<2e-16 ***
`Tax 5%`	1.000e+00	3.771e-17	2.652e+16	<2e-16 ***
Rating	-8.110e-18	8.504e-17	-9.500e-02	0.924
GenderMale	-2.555e-16	2.929e-16	-8.720e-01	0.383
PaymentCredit card	4.452e-16	3.614e-16	1.232e+00	0.218
PaymentEwallet	-1.299e-18	3.516e-16	-4.000e-03	0.997

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 4.61e-15 on 992 degrees of freedom

Multiple R-squared: 1, Adjusted R-squared: 1

F-statistic: 9.207e+32 on 7 and 992 DF, p-value: < 2.2e-16

- We can see in this model that the function will be

Gross income = -2.461e-14 + [4.267e-15 × Quantity] + [4.073e-16 × Unit price] + [1.00 × Taxes] + [2.555- e-16 × Gender (Male)] + [4.452e-16 × Payment Options "Credit Card"] + [-1.299e-18 × Payment Options" E-wallet"] + [-8.110e-18 × Rating]

- And this the Transformed Coefficient for the function

Gross income = -0.00000000000002461+ [0.00000000000004267 × Quantity]+[0.00000000000004073× Unit price] + [1.00 × Taxes] + [-0.000000000000002555 × Gender (Male)] + [0.000000000000004452 × Payment Options "Credit Card"] + [-0.0000000000000001299 × Payment Options" E-wallet"] + [-0.0000000000000008110× Rating]

- From the output we will reject that there is relation between Gross Income with Rating , Payment "wallet" , Payment "Credit Card "& Gender "Male " .
So, P-value > 0.05 There 's no significant evidence that there is a relation.
- Also we can accept the relation between Gross Income with Unit Price , Quantity and Tax .5% .P-value < 0.05 so we accept the hypothesis .
- We can see that R-Square (1) and it measures to a good model and it means that approximately (100 %) from the variance in dependent variable are explained by the independent .

The decision

- If we need to increase Gross Income, We can focus on Quantity, Unit price & Taxes. If we increasing any of them, the Gross Income will increase as well.
- On the other hand, If we need to increase Quantity we should focus on Unit price & Taxes and ignoring other variables.
- We can Also focus on Marketing for male and female by their desire in product line. As we see, Females are interested in fashion accessories & food more than males. So, we should keep our Marketing and advertises for women than men in those departments.
- On the other hand, Men interested in health and beauty & Electronic Accessories Than women. As well we should keep our advertises for men in those departments.
- If we look at payment options, we can see that we should modify and invest in E wallet payment in Yangon and Mandalay branches. We can buy more equipment, deal with more banks it may help to grabbing more customer and achieve satisfaction.

Conclusion

According to the result of out tests and summary, We should be more care about some variables which by that can help in improving sales. Also we have to be careful for our unit costs and before increasing any prices. As, It can effect on Quantities.

References & dataset

- 1- "Supermarket sales" , Historical record of sales data in 3 different supermarkets , Retrieved from the link <https://www.kaggle.com/datasets/aungpyaeap/supermarket-sales/code> .