

ICA

BIG DATA AND

BUSINESS INTELLIGENCE

SECTION 1: BUSINESS INTELLIGENCE DESIGN

1 . BI Data Source Description and BI Questions

1.1 Data Source Description

The dataset used in this assignment is titled "supermarket_sales - Sheet1.csv" which is provided on Kaggle for public and non-commercial purposes. The link to the dataset is as below:

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

Customers' purchase order information from a supermarket with three locations in Myanmar in 2019 is included in the dataset. The screenshot is attached below.

The screenshot shows a Power BI dataset view titled "Supermarket sales". The table has 17 columns: InvoiceID, Branch, City, CustomerType, Gender, Productline, UnitPrice, Quantity, Tax5, Total, Date, Time, Payment, COGS, GrossMarginPercentage, GrossIncome, and Rating. There are 17 rows of data. Row 1 shows an invoice for a female member from Yangon with a total of 549.97. Row 2 shows an invoice for a male member from Yangon with a total of 80.22. Row 3 shows an invoice for a male member from Yangon with a total of 340.53. Row 4 shows an invoice for a male member from Yangon with a total of 489.05. Row 5 shows an invoice for a male member from Yangon with a total of 634.38. Row 6 shows an invoice for a male member from Yangon with a total of 627.62. The bottom of the screen shows navigation links for Data, Code (98), Discussion (17), Activity, and Metadata, along with a Download (132 KB) button and a New Notebook button.

InvoiceID	Branch	City	CustomerType	Gender	Productline	UnitPrice	Quantity	Tax5	Total	Date	Time	Payment	COGS	GrossMarginPercentage	GrossIncome	Rating
1-0-0-8420	A	Yangon	Member	Female	Health and beauty	74.65	7	26.14	549.97	43586	13:00	Ewallet	522.83	4.76	26.14	9.1
1-1-1-108	A	Yangon	Normal	Male	Home and lifestyle	46.33	5	3.82	80.22	43680	10:29	Cash	76.40	4.76	3.82	9.6
3-1-1-108	A	Yangon	Normal	Male	Home and lifestyle	46.33	7	16.22	340.53	43527	13:23	Credit card	324.31	4.76	16.22	7.4
4-3-19-176	A	Yangon	Member	Male	Health and beauty	58.22	8	23.29	489.05	1/27/2019	20:33	Ewallet	465.76	4.76	23.29	8.4
5-3-73-7910	A	Yangon	Normal	Male	Sports and travel	86.31	7	30.21	634.38	43679	10:37	Ewallet	604.17	4.76	30.21	5.3
6-3-73-7910	A	Yangon	Normal	Male	Accessories	85.39	7	29.89	627.62	3/25/2019	18:30	Ewallet	597.73	4.76	29.89	4.1

Fig: Screenshot of dataset

1.2 Objectives and approach

Power BI and the report in this project are intended to display customer, sales, product, and financial information at the branch, product, customer levels.

The analysis is primarily concerned with the following issues:

- To compare the performance of 3 branches on sales, cost, gross income, and income margin.
- To compare the performance of 6 product categories on sales, cost, gross income, and income margin.
- To analyze the current customers' profiles and preferences.
- To investigate customers satisfaction based on the rating score.

1.3 Scope

This project scope comprises the above approach, but the focus is on data pre-processing, data cleaning, and data modeling in section 1. In section 2, the priority is to deliver an analytical report and interactive dashboard to support performance evaluation based on the above questions.

1.4 Rationale

Microsoft Power BI (business intelligence) is a powerful analytics tool that helps companies of all sizes analyze data and share insights. It can monitor the business closely and get instant answers with rich dashboards. So business data is exactly suitable for it.

And analyzing sales data is more appealing to me because it allows me to quickly see information about revenue, business performance, and financial health. This type of data helps managers in making strategic decisions and maximizing company profits.

1.5 Outcomes

In order to show the business performance, the following skills can be demonstrated:

- Data cleaning and data pre-processing: Importing data, adding/removing a column, removing NAs or errors, changing data types.
- Data modeling: creating and managing relationships, splitting big tables to normalize data.
- DAX and M language have both been used in the analysis.
- New calculated columns and new measures have been added to the model.
- Different visuals have been used in this project.
- A new type of chart (not covered in the lesson) has been included in the report.
- Designing and publishing the report.

1.6 Data description

"Supermarket_sales - Sheet1.csv" is a dataset containing details of products, customers, order dates, financial data. It has 17 columns and 1k records. In this project, 16 columns have been used to analyze the sales performance.

The details shows below:

Column Name	Description
Invoice id	Computer generated sales slip invoice identification number
Branch	3 branches are available identified by A, B and C
City	Location of branches
Customer type	Member or not member
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 without tax
Date	Date of purchase in 2019
Time	Purchase time (10am to 9pm)
Payment	3 methods are available – Cash, Credit card and E wallet
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)

CSV screenshot:

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
Invoice ID	Branch	City	Customer type	Gender	Product line	Unit price	Quantity	Tax %	Total	Date	Time	Payment	cogs	gross margin percentage	gross income	Rating
1 750-67-8428	A	Yangon	Member	Female	Health and beauty	74.69	7	26.1415	548.9715	01/05/2019	13:08	Ewallet	522.83	4.761904762	26.1415	9.1
3 226-31-3081	C	Naipyitaw	Normal	Female	Electronic accessories	15.28	5	3.82	80.22	03/08/2019	10:29	Cash	76.4	4.761904762	3.82	9.6
4 631-41-3108	A	Yangon	Normal	Male	Home and lifestyle	46.33	7	16.2155	340.5255	03/03/2019	13:23	Credit card	324.31	4.761904762	16.2155	7.4
5 123-19-1176	A	Yangon	Member	Male	Health and beauty	58.22	8	23.288	489.048	1/27/2019	20:33	Ewallet	465.76	4.761904762	23.288	8.4
6 373-73-7910	A	Yangon	Normal	Male	Sports and travel	86.31	7	30.2085	634.3785	02/08/2019	10:37	Ewallet	604.17	4.761904762	30.2085	5.3
7 699-14-3026	C	Naipyitaw	Normal	Male	Electronic accessories	85.39	7	29.8865	627.6165	3/25/2019	18:30	Ewallet	597.73	4.761904762	29.8865	4.1
8 385-53-5343	A	Yangon	Member	Female	Electronic accessories	68.84	6	20.652	433.692	2/25/2019	14:36	Ewallet	413.04	4.761904762	20.652	5.8
9 315-22-5626	C	Naipyitaw	Normal	Female	Home and lifestyle	72.56	10	36.78	727.38	2/24/2019	11:00	Ewallet	735.6	4.761904762	36.78	8
10 655-26-9167	A	Yangon	Member	Female	Health and beauty	36.26	2	3.626	76.146	01/10/2019	17:15	Credit card	72.52	4.761904762	3.626	7.2
11 693-92-5592	B	Mandalay	Member	Female	Food and beverages	54.04	3	2.046	172.746	2/20/2019	13:37	Credit card	164.52	4.761904762	2.046	5.9
12 351-62-0822	B	Mandalay	Member	Female	Fashion accessories	14.48	4	2.898	60.816	02/06/2019	18:07	Ewallet	57.92	4.761904762	2.898	4.5
13 529-56-3974	B	Mandalay	Member	Male	Electronic accessories	25.51	4	5.102	107.142	03/09/2019	17:03	Cash	102.04	4.761904762	5.102	6.8
14 365-64-0515	A	Yangon	Normal	Female	Electronic accessories	46.95	5	11.7375	246.4875	02/12/2019	10:25	Ewallet	234.75	4.761904762	11.7375	7.1
15 252-56-2699	A	Yangon	Normal	Male	Food and beverages	42.19	10	21.595	453.495	02/07/2019	16:48	Ewallet	431.9	4.761904762	21.595	4.5
16 829-34-3910	A	Yangon	Normal	Female	Health and beauty	71.38	10	35.69	749.49	3/29/2019	19:21	Cash	713.8	4.761904762	35.69	5.7
17 299-46-1805	B	Mandalay	Member	Female	Sports and travel	93.72	6	28.116	590.436	1/15/2019	16:19	Cash	562.32	4.761904762	28.116	4.5
18 656-95-9349	A	Yangon	Member	Female	Health and beauty	68.93	7	24.1256	506.6355	03/11/2019	11:03	Credit card	482.51	4.761904762	24.1255	4.6
19 765-26-6951	A	Yangon	Normal	Male	Sports and travel	72.61	6	21.783	457.443	01/01/2019	10:39	Credit card	435.66	4.761904762	21.783	6.9
20 329-62-1586	A	Yangon	Normal	Male	Food and beverages	54.67	3	8.2005	172.2105	1/21/2019	18:00	Credit card	164.01	4.761904762	8.2005	8.6
21 319-50-3348	B	Mandalay	Normal	Female	Home and lifestyle	40.3	2	4.03	84.63	03/11/2019	15:30	Ewallet	80.6	4.761904762	4.03	4.4

2 . BI Data Pre-Processing and Data Cleansing

2.1 Data loading

The very first step of data analysis and creating Business Intelligence is loading raw data.

First, launch power BI. I didn't notice my Power BI language was Chinese at the beginning, the language setting would be changed in few steps later.

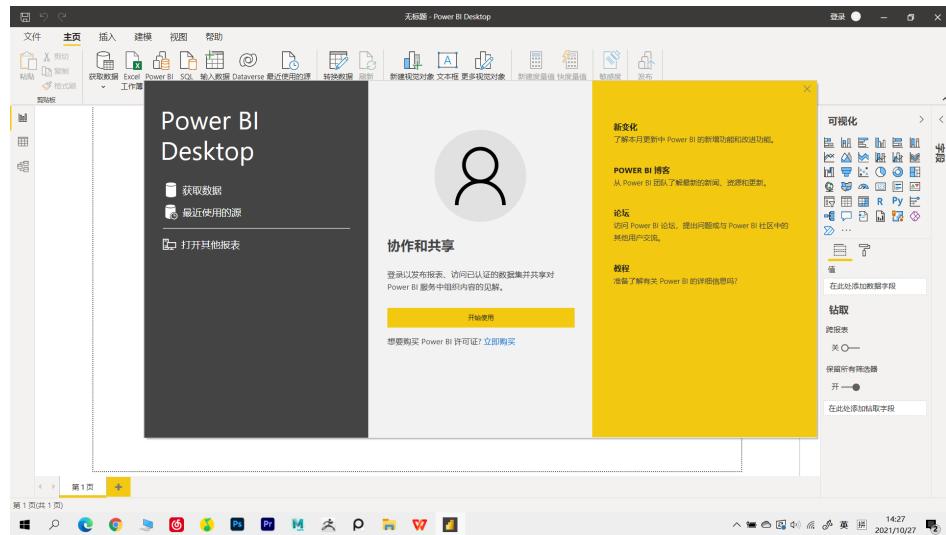


Fig: Power BI startup

Select the arrow next to “Get Data” in the Power BI Desktop ribbon's Home tab. In the Open dialog box, navigate to and select the "supermarket_sales - Sheet1.csv" file, and then select “Open”.

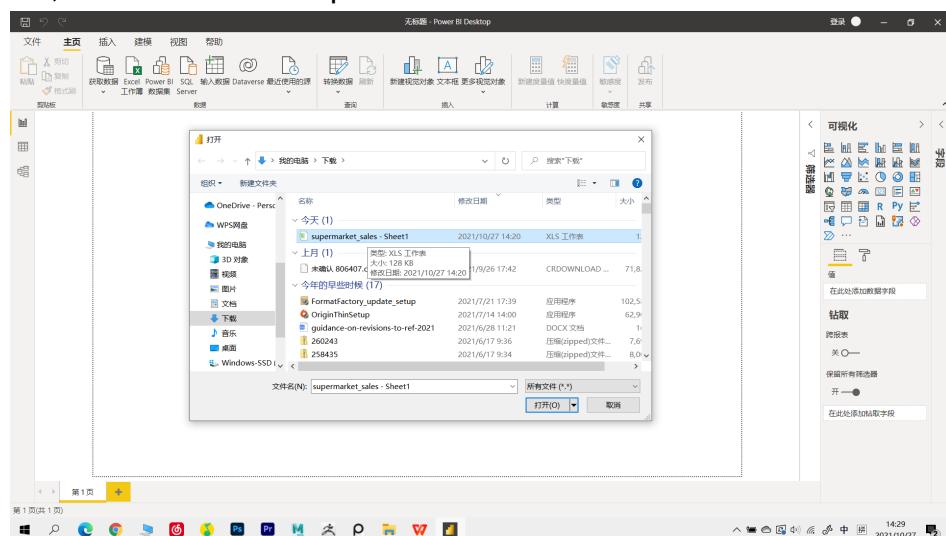


Fig: Choose file

In the Navigator, select “Load”.

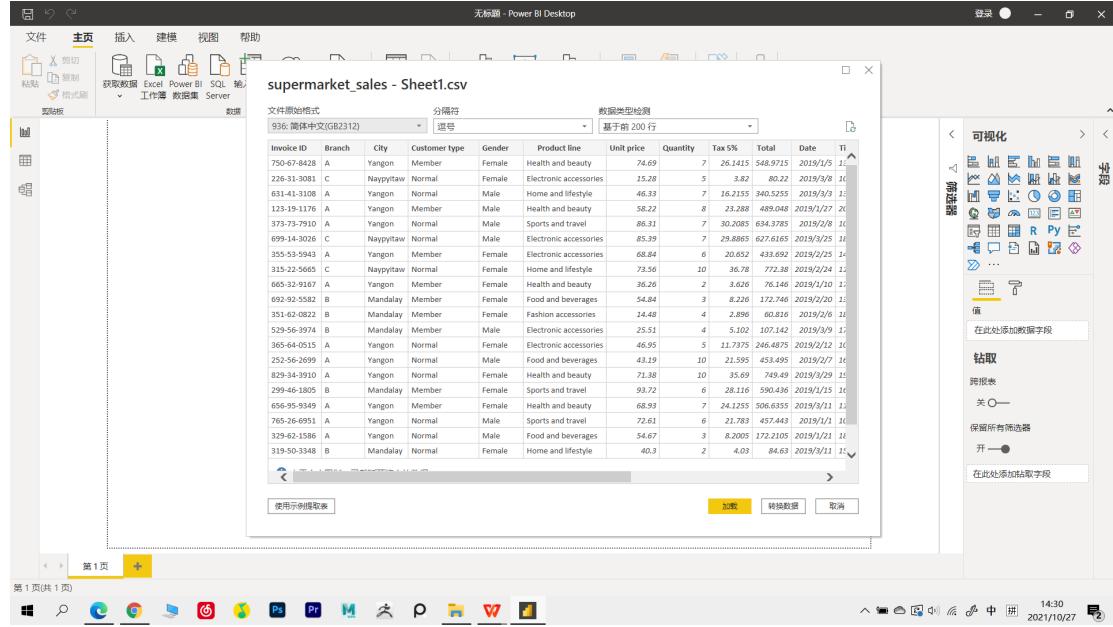


Fig: Load data

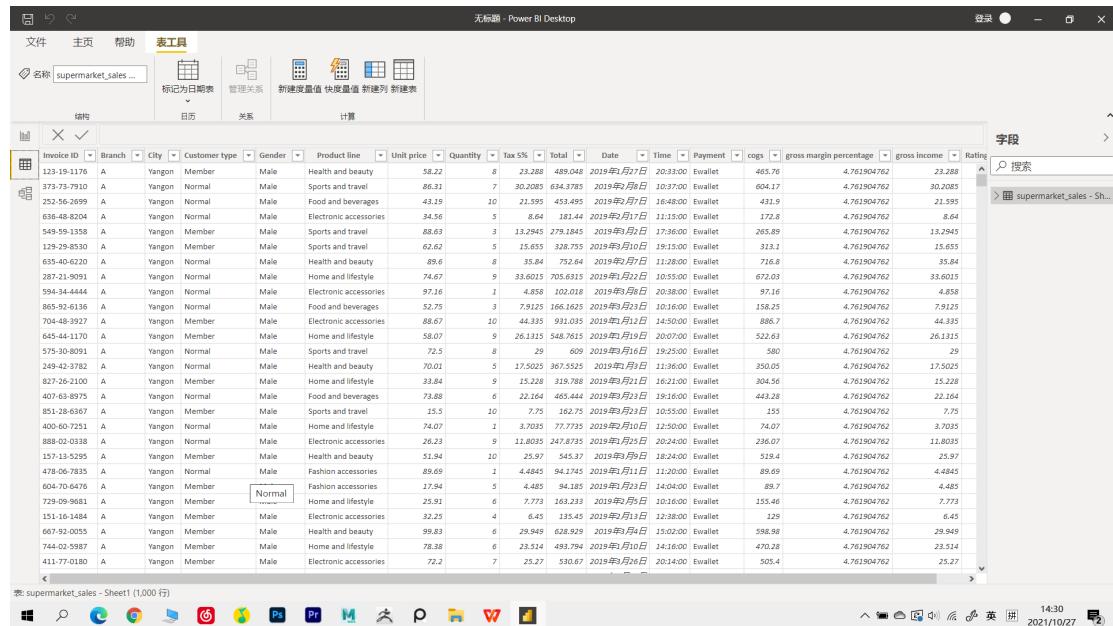


Fig: Screenshot of data loading in Power Query Editor

2.2 Quick Check

Before I start cleaning the data, I want to double-check that the dataset I chose is appropriate, so I ran a quick check on it.

I make two pie charts; one is total sales by different branches and another is gross income by different products. I discover that all of the items have a similar amount, which makes it difficult to get useful information. After checking with my tutor Annalisa, I decide to change the data by myself.

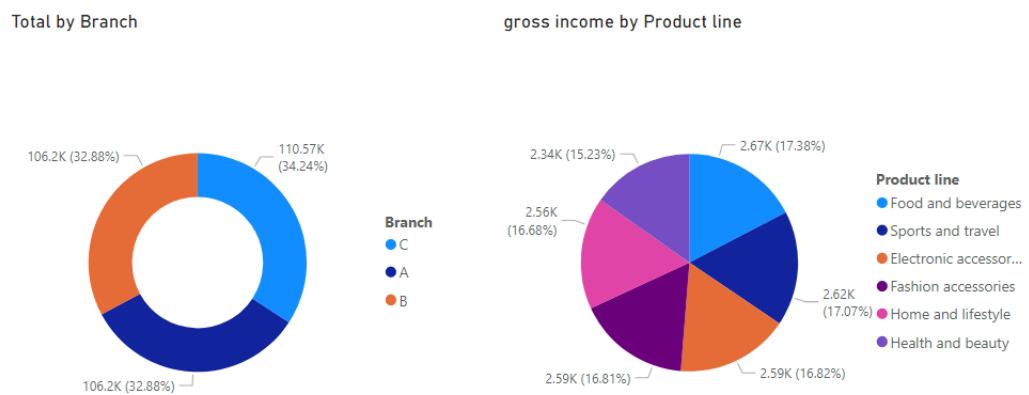


Fig: Screenshot of pie chart

2.3 Data pre-processing

I write VBA code to replace the current total cost and sell quantity with a random percentage and number. To maximize each item's difference, I have set different parameters for different products.

For the cost, if the product is food, the cost of food will be reduced by 1% to 50%, while the cost of other products will be reduced or increased by 20% only. For the sales quantity, If the product is food, the quantity will be added a random number between 5 and 20, while others will be given a number between 1 and 10.

```

supermarket_sales - macro.xlsm - Module1 (Code)
(General) Macro1

Dim LRandomNumber As Integer

For i = 2 To 1001
    Randomize
    If Cells(i, 6).Value = "Food and beverages" Then
        ' Int((upperbound - lowerbound + 1) * Rnd + lowerbound)
        LRandomNumber = -Int((50 - 1 + 1) * Rnd + 1) 'Generate random value between -50 and -1.
        LRandomNumber2 = Int((20 - 5 + 1) * Rnd + 5) 'Generate random value between 5 and 20.
        ' MsgBox (Cells(i, 6).Value)
    Else
        LRandomNumber = Int((20 + 20 + 1) * Rnd - 20) ' Generate random value between -20 and 20.
        LRandomNumber2 = Int((10 - 1 + 1) * Rnd + 1) ' Generate random value between 1 and 10.
    End If
    Cells(i, 18).Value = LRandomNumber * 0.01 + 1 'Generate percentage to change cost
    Cells(i, 20).Value = LRandomNumber2 'Generate number to change quantity
    Next i
End Sub

```

Fig: VBA code

After changing the data, I repeat the previous steps to reload the data again.

2.4 Data cleaning

2.4.1 Add calculated columns

After loading the data, the data cleaning is begun. As I have added two new columns called “Random %” and “Random Quantity”.

	% Random %	Random Quantity
1	82.00%	8
2	115.00%	6
3	97.00%	8
4	91.00%	20
5	85.00%	9

Fig: 2 new columns

As the cost and quantity changed, the columns “cogs”, “Total”, “gross margin percentage”, “gross income”, “gross margin percentage” should be changed accordingly.

To do so, I select the Transform Data button in the Power Query Editor window.

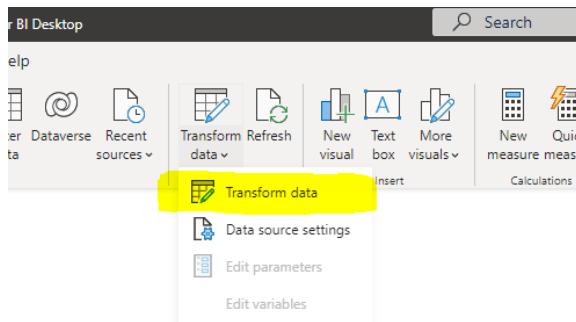


Fig: Click Transform data.

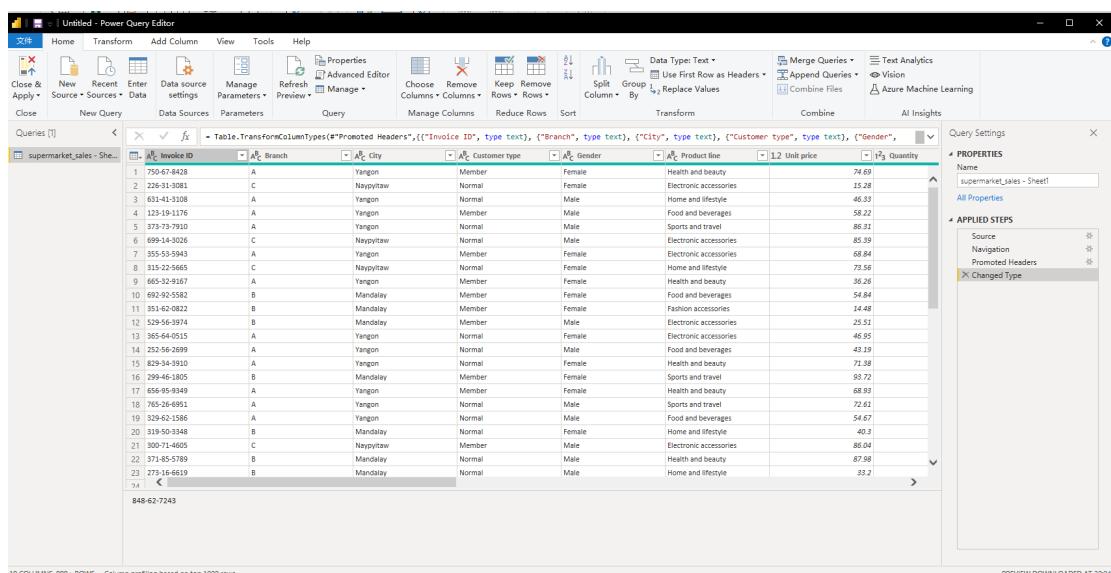


Fig: Open power query editor

The new columns should be calculated by following formulas:

New COGS = Random % * COGS

New Quantity = Random Quantity + Quantity

New Total = New Quantity * Unit Price

New Gross Income = New Total - New COGS

New Gross Margin = New Gross Income / New Total

In the Power Query Editor's Add Column ribbon tab, select "Custom Column".

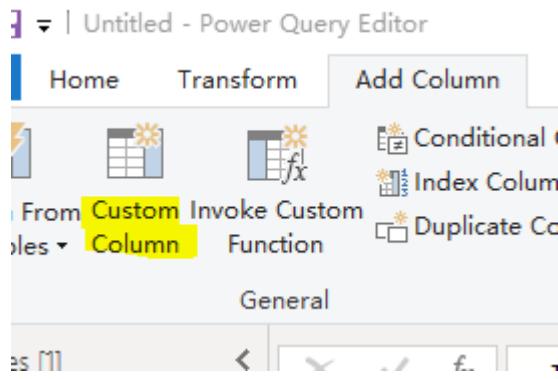


Fig: Select Custom Column

In the Custom column formula field, input the new column name, select the field names from the “Available columns” scroll box and select << Insert.

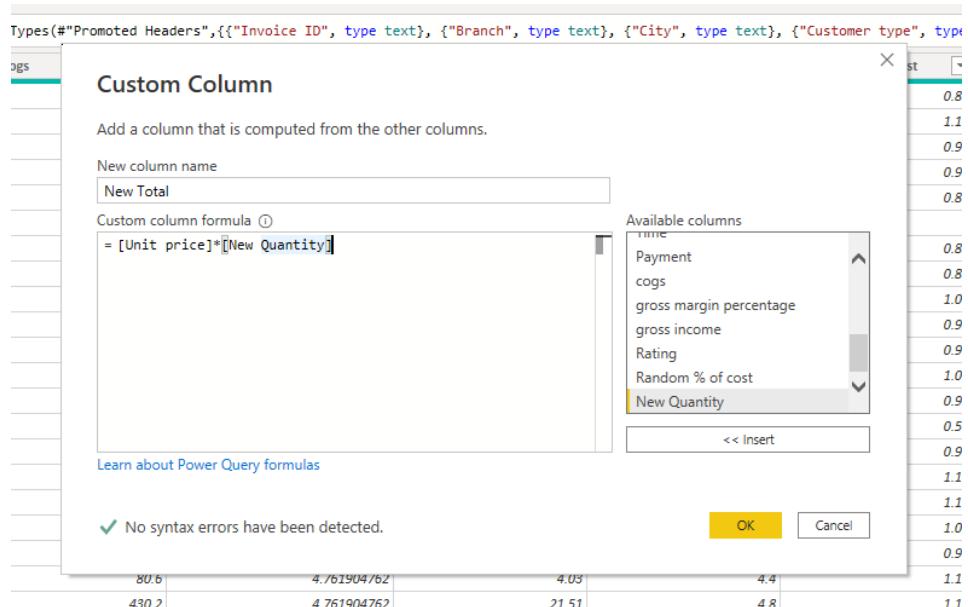


Fig: Add New Total

Instead of adding columns one by one, M language allows inserting bulk columns.
Go to view and click “Advanced Editor” .

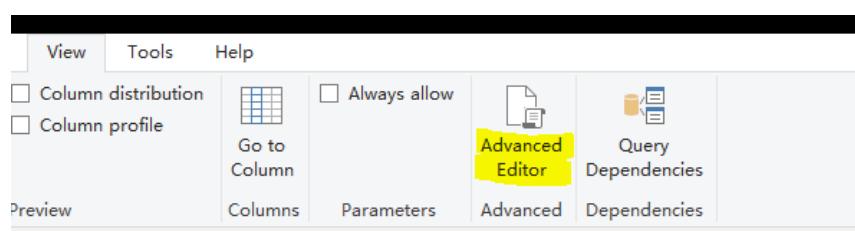


Fig: Advanced Editor

Copy the “Added Custom” in line 5, change the parameters.

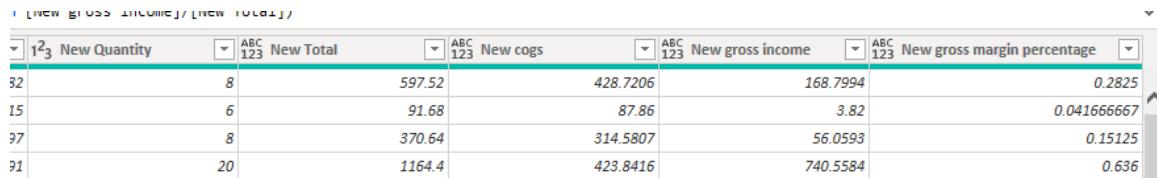


```

let
    Source = Excel.Workbook(File.Contents("C:\Users\liaosupermarket_sales.xlsx"), null, true),
    supermarket_sales = Source{[Item="supermarket_sales", Kind="Sheet"]}[Data],
    #Promoted Headers = Table.PromoteHeaders(#"supermarket_sales", [PromoteAllScalars=true]),
    #Changed Type = Table.TransformColumnTypes(#"Promoted Headers", {{"Invoice ID", type text}, {"Branch", type text}, {"City", type text}, {"Customer type", type text}, {"Gender", type text}, {"Product line", type text}, {"New Quantity", type number}, {"New Total", type number}, {"New cogs", type number}, {"New gross income", type number}, {"New gross margin percentage", type number}}),
    #Added Custom1 = Table.AddColumn(#"Changed Type", "New Total", each [Unit price]*[New Quantity]),
    #Added Custom2 = Table.AddColumn(#"Added Custom1", "New cogs", each #Random % of cost*[cogs]),
    #Added Custom3 = Table.AddColumn(#"Added Custom2", "New gross income", each [New Total]-[New cogs]),
    #Added Custom4 = Table.AddColumn(#"Added Custom3", "New gross margin percentage", each [New gross income]/[New Total])
in
    #Added Custom4

```

Fig: M language

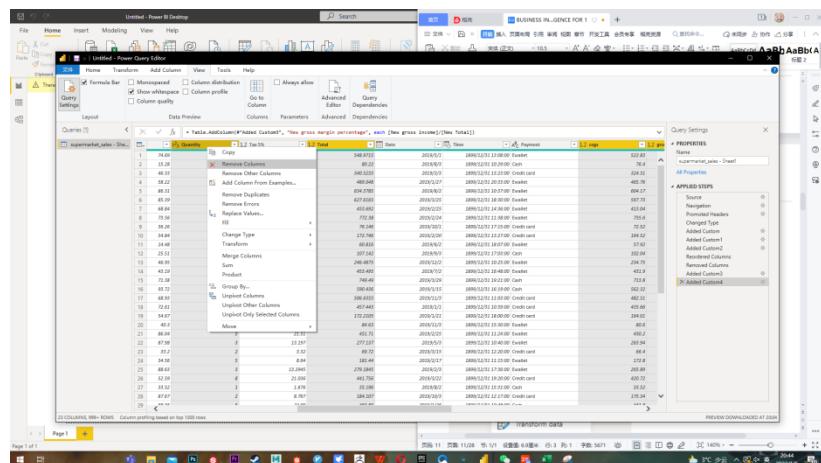


	New Quantity	ABC 123 New Total	ABC 123 New cogs	ABC 123 New gross income	ABC 123 New gross margin percentage
32	8	597.52	428.7206	168.7994	0.2825
15	6	91.68	87.86	3.82	0.041666667
97	8	370.64	314.5807	56.0593	0.15125
91	20	1164.4	423.8416	740.5584	0.636

Fig: New columns

2.4.2 Remove columns

As I have added new columns to replace the previous one, columns “Total”, “gross margin percentage”, “gross income”, “cogs”, “Quantity” should be removed. Select these columns by keeping pressing Ctrl. Then, right-click on one of the columns’ headers and select use “Remove Columns” to remove all the columns.



The screenshot shows the Power BI Data Editor interface. A context menu is open over the header of the 'New Total' column in the 'supermarket_sales' table. The menu path 'Remove Columns' is highlighted. Other options visible in the menu include 'Add Column from Examples...', 'Remove Duplicates', 'Remove Errors', 'Replace with Values...', 'Fill...', 'Change Type', 'Advanced', 'Transform', 'Delete Columns', 'Sum', 'Product', 'Group by', 'Unpivot All Columns', 'Unpivot Other Columns', 'Unpivot Only Selected Columns', and 'None'. The 'APPLIED STEPS' pane on the right shows the history of changes made to the query, including the addition of the 'New Total' column.

Fig: Remove columns

2.4.3 Change data type

To assign a data type, right-click the column “New gross income” header, select “Change Type” from the drop-down menu, and then select “Fixed decimal Number”.

The screenshot shows the Power Query Editor interface. In the center, there is a table with columns labeled 'Rating', 'New Total', 'Random % of cost', 'New Quantity', 'New Total', and 'New gross margin percentage'. The 'New gross margin percentage' column has its data type set to 'Fixed decimal number' with 12 decimal places. On the right side, the 'Applied Steps' pane is open, showing a list of steps. The first step is 'Changed Type1', which changes the type of the 'New gross margin percentage' column. The 'Properties' pane on the right shows the name of the query as 'supermarket_sales - Sheet1'.

Fig: Change type

Use M language to change other column types.

```
#"Changed Type1" = Table.TransformColumnTypes(#"Removed Columns1",{{"New Total", Currency.Type}, {"Random % of cost", Percentage.Type}}
#"Changed Type2" = Table.TransformColumnTypes(#"Renamed Columns",{{"gross margin percentage", Percentage.Type}})
```

Fig: M language

Change currency mark to U.S dollar.

The screenshot shows an Excel spreadsheet. At the top, the ribbon has tabs for Home, Insert, Page Layout, Formulas, Data, Sort & Filter, Review, and View. The 'Home' tab is selected. Below the ribbon, there is a 'Formatting' tab. A table is visible below the ribbon, containing columns for 'Type', 'Gender', 'Product line', and 'Unit price'. The 'Unit price' column contains values like '\$58.22', '\$86.31', and '\$43.19'. The 'Format' dropdown in the ribbon is set to 'Currency'.

Fig: Change Currency

2.4.4 Rename columns

As I have removed previous columns, change the name of new columns by right-clicking and selecting “Rename” from the drop-down menu.

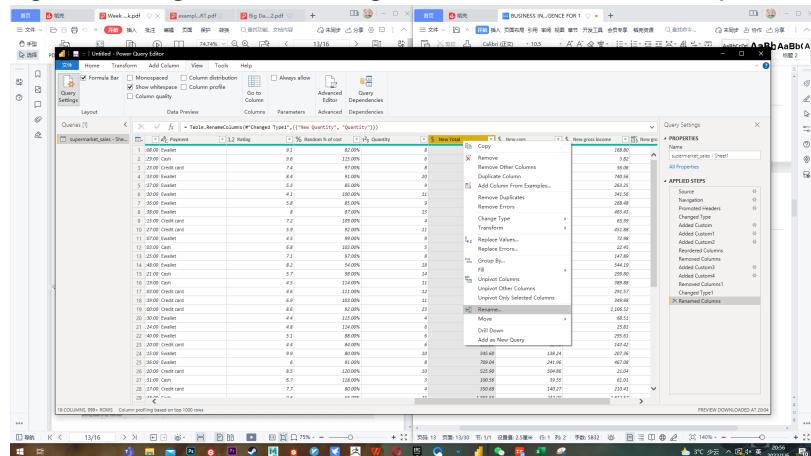


Fig: Change name

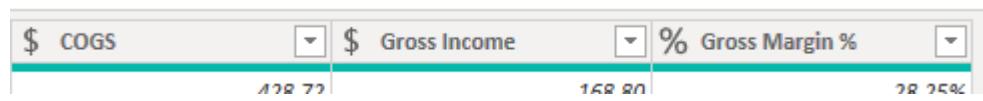


Fig: Change name back

2.4.5 Remove errors and NA

As I changed column name incorrectly, errors appears below. Change back column name solves this error.



Fig: Error

Remove NA column by click remove empty.

The screenshot shows the Power Query Editor interface. A table is being loaded from '1.2 cogs'. The table has 37 rows and 12 columns. A context menu is open over the cell at row 27, column 1, which contains the value 'null'. The menu path 'Column & Rows' -> 'Remove' is highlighted. Other options visible in the menu include 'Select All', 'Sort Ascending', 'Clear Sort', 'Remove', 'Remove Empty', and 'Number Filters'. The 'APPLIED STEPS' pane on the right shows the steps taken so far, including 'Source' and 'Promoted Headers'.

Fig: Remove empty value

	1.2 cogs
1	522.83
2	76.4
3	324.31
4	465.76
5	604.17
6	597.73
7	413.04
8	735.6
9	72.52
10	164.52
11	57.92
12	102.04
13	234.75
14	431.9
15	713.8
16	562.32
17	482.51
18	435.66
19	164.01
20	null
21	null
22	null
23	null
24	null
25	null
26	null
27	175.34
28	...

Fig: Null value

Select “Close & Apply” in the Home ribbon tab's Close group to save work.

2.4.6 Split the table

I split the main table into smaller tables by right-clicking on “supermarket_sales - Sheet1” “table on the left and selecting “Duplicate”.

The screenshot shows the Power BI Desktop interface with the 'supermarket_sales - Sheet1' table selected. A context menu is open, and the 'Duplicate' option is highlighted. The 'APPLIED STEPS' pane on the right shows the history of changes made to the table, including 'Changed Type1' and 'Changed Type2'. The table preview shows various columns like 'Invoice ID', 'Branch', 'Customer type', 'Gender', 'Product line', 'Unit price', and 'Tax%', with data rows corresponding to different customers and their purchases.

Fig: Duplicate table

Rename this table as “Customer Information” by right-clicking on it and selecting Rename

The screenshot shows the Power BI Desktop interface with the 'supermarket_sales - Sheet1' table selected. A context menu is open, and the 'Rename' option is highlighted. The 'APPLIED STEPS' pane on the right shows the history of changes made to the table, including 'Changed Type1' and 'Changed Type2'. The table preview shows the same data as before, but the table name has been changed to 'Customer Information'.

Fig: Rename table

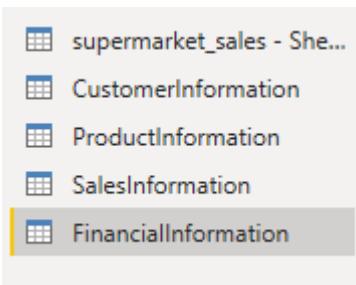


Fig: 4 new tables

Remove all the columns that are not needed. Click on the header of the Invoice ID column, then press Ctrl and click on the Gender and Customer type columns, right-click on any of the headers of the three columns and select “Remove Other Columns”.

Invoice ID	Customer type	Gender	Payment
750-67-8428	Member	Female	Ewallet
226-31-3081	Normal	Female	Cash
631-41-3108	Normal	Male	Credit card
123-19-1176	Member	Male	Ewallet
373-73-7910	Normal	Male	Ewallet
699-14-3026	Normal	Male	Ewallet

Fig: Remove Other Columns

Invoice ID	Customer type	Gender	Payment
750-67-8428	Member	Female	Ewallet
226-31-3081	Normal	Female	Cash
631-41-3108	Normal	Male	Credit card
123-19-1176	Member	Male	Ewallet
373-73-7910	Normal	Male	Ewallet
699-14-3026	Normal	Male	Ewallet

Fig: The result

Apply the same action to the other 3 tables.

The screenshot shows the Power BI Data Editor interface with the 'Product Information' table selected. The table has columns: Invoice ID, Product line, Unit price, and Rating. The data includes various product categories like Health and beauty, Electronic accessories, etc., with their respective unit prices and ratings.

	A ^b Invoice ID	A ^b Product line	1.2 Unit price	1.2 Rating
1	750-67-8428	Health and beauty	74.69	9.1
2	226-31-3081	Electronic accessories	15.28	9.6
3	631-41-3108	Home and lifestyle	46.33	7.4
4	123-19-1176	Food and beverages	58.22	8.4
5	373-73-7910	Sports and travel	86.31	5.3
6	699-14-3026	Electronic accessories	85.39	4.1
7	355-53-5943	Electronic accessories	68.84	5.8

Fig: Product Information

The screenshot shows the Power BI Data Editor interface with the Sales Information table selected. The table has columns: Invoice ID, Branch, City, Time, Date, Quantity, Total, and a few others removed. The data shows sales transactions across different branches and cities, with their respective quantities and total amounts.

	A ^b Invoice ID	A ^b Branch	A ^b City	Date	Time	Quantity	Total
1	750-67-8428	A	Yangon	2019/5/1	1899/12/31 13:08:00	8	597.52
2	226-31-3081	C	Naypyitaw	2019/8/5	1899/12/31 10:29:00	6	91.68
3	631-41-3108	A	Yangon	2019/3/9	1899/12/31 13:23:00	8	370.64
4	123-19-1176	A	Yangon	2019/1/27	1899/12/31 20:39:00	20	1,264.40
5	373-73-7910	A	Yangon	2019/8/2	1899/12/31 10:37:00	9	776.79
6	699-14-3026	C	Naypyitaw	2019/9/25	1899/12/31 18:30:00	11	839.29
7	355-53-5943	A	Yangon	2019/2/25	1899/12/31 14:36:00	9	619.56
8	315-22-5665	C	Naypyitaw	2019/9/24	1899/12/31 11:38:00	15	1,103.40
9	665-32-9167	A	Yangon	2019/10/1	1899/12/31 17:15:00	4	145.04

Fig: Sales Information

The screenshot shows the Power BI Data Editor interface with the Financial Information table selected. The table has columns: Invoice ID, cogs, gross income, gross margin percentage, and a few others removed. The data shows financial metrics for different invoices.

	A ^b Invoice ID	\$ cogs	\$ gross income	% gross margin percentage
1	750-67-8428	428.72	168.80	28.25%
2	226-31-3081	87.86	3.82	4.17%
3	631-41-3108	314.58	56.06	15.13%
4	123-19-1176	423.84	740.56	63.60%
5	373-73-7910	513.54	263.25	33.89%
6	699-14-3026	597.73	341.56	36.36%
7	355-53-5943	351.08	268.48	43.33%

Fig: Financial Information

As a backup, I will not delete the supermarket_sales table. Click on “Close & Apply” to apply the changes.

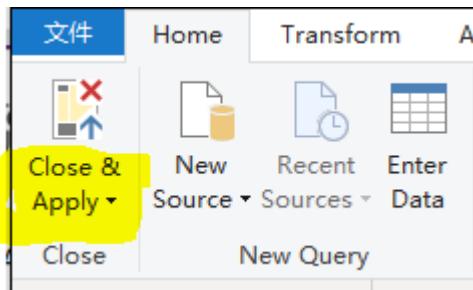


Fig: Close and apply

2.4.7 Replace value

This step happens when I use visual called “Map”, the city “Naypyitaw” can’t be found in map. It should be “Naypyidaw” after I google it.

Select column and click “Replace Values”.

The screenshot shows the Power Query Editor interface with a table of sales data. A 'Replace Values' dialog box is open, prompting the user to replace 'Naypyitaw' with 'Naypyidaw'. The 'OK' button is highlighted with a yellow box.

Invoice ID	Branch	City	Date	Time	Quantity	\$ Sales Total
750-07-8428	A	Yangoon	2019/5/1	12:08:00	8	597.52
226-31-2345	C	Naypyitaw	2019/5/7	10:29:00	6	91.69
312-03-2108	A	Yangoon	2019/5/9	12:31:00	8	370.04
132-19-1176	A	Yangoon	2019/5/17	20:33:00	20	1,164.40
5 373-73-7910	A	Yangoon	2019/8/2	10:27:00	9	776.79
6 699-14-3028	C	Naypyitaw	2019/8/25	12:30:00	11	839.29
7 355-53-5943	A	Yangoon	2019/2/25	14:36:00	9	619.56
8 315-22-5665	C	Naypyitaw				
9 665-32-1541	A	Yangoon				
10 355-02-5921	B	Mandalay				
11 351-02-0822	B	Mandalay				
12 528-56-3974	B	Mandalay				
13 365-64-0515	A	Yangoon				
14 252-56-2699	A	Yangoon				
15 829-34-3910	A	Yangoon				
16 209-46-1803	B	Mandalay				
17 656-95-3549	A	Yangoon				
18 319-02-5561	A	Yangoon				
19 319-02-1566	A	Yangoon				
20 319-50-3348	B	Mandalay				
21 309-71-4605	C	Naypyitaw				
22 371-85-5789	B	Mandalay	2019/5/3	10:40:00	6	527.88
23 271-14-6619	B	Mandalay	2019/5/15	12:20:00	6	199.20
24 630-85-3420	A	Yangoon	2019/5/17	12:30:00	10	345.60
25 308-99-1358	A	Yangoon	2019/5/20	12:38:00	4	300.04
26 327-03-5010	A	Yangoon	2019/5/22	19:29:00	10	525.99
27 649-29-6775	B	Mandalay	2019/8/2	12:31:00	3	200.55
28 189-17-4241	A	Yangoon	2019/8/3	12:17:00	4	350.68
29 145-94-9001	B	Mandalay	2019/3/25	19:48:00	21	2,855.56
30 848-62-7243	A	Yangoon	2019/3/15	12:36:00	12	298.68
31 871-79-8483	B	Mandalay	2019/3/25	19:39:00	8	733.04
32 149-71-6266	B	Mandalay	2019/3/28	12:43:00	10	780.70

Fig: Replace Values

3 . BI Data Modelling

3.1 Remove relationships

Select the Model tab. Power BI will automatically generate the proper relationship between tables. The picture below shows the entity-relationship diagram of raw unsupervised data and its interconnection.

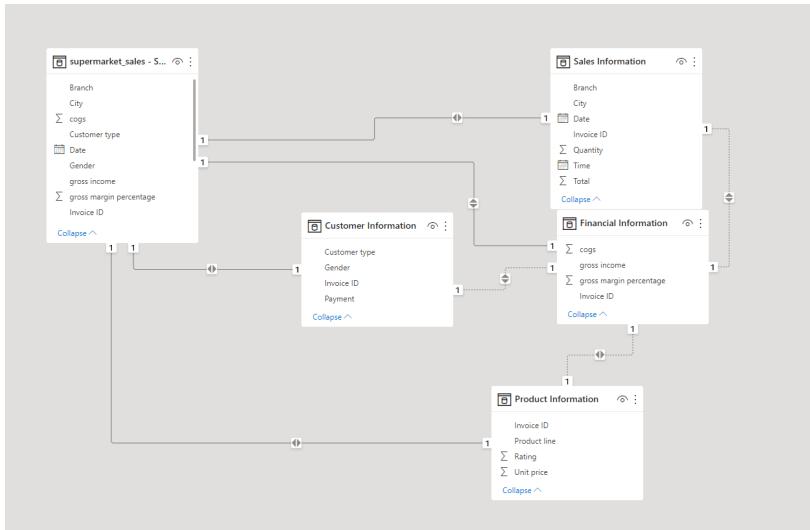


Fig: Original relationship

To remove a relationship, right-click on the link and select Delete (see below). In the pop-up window that asks if you want to delete the relationship, select Delete. Repeat this step to delete them all.

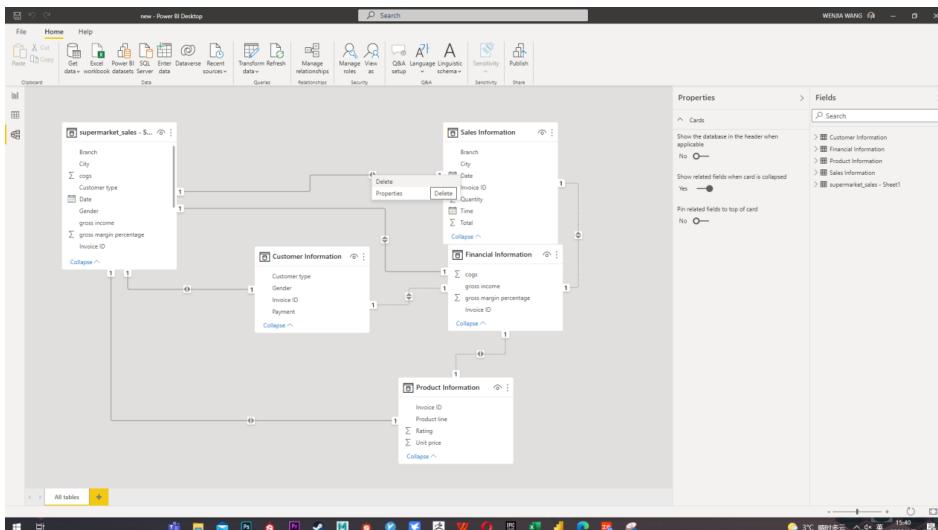


Fig: Delete relationship

3.2 Create new relationships

Relink these 4 tables via “Invoice ID”.

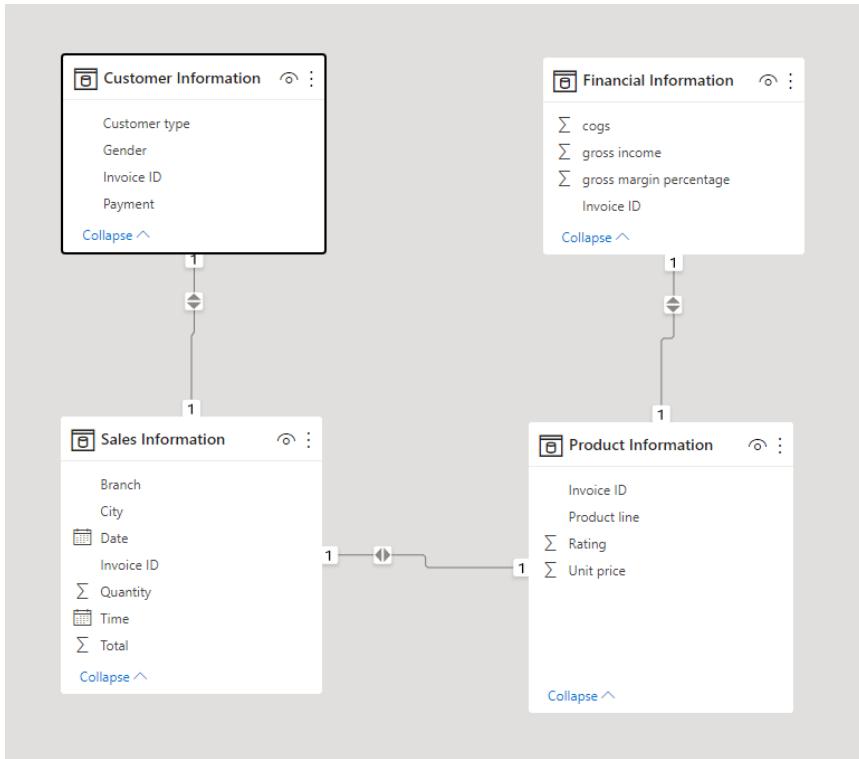


Fig: New relationship

Make sure all the relationships are one to one.

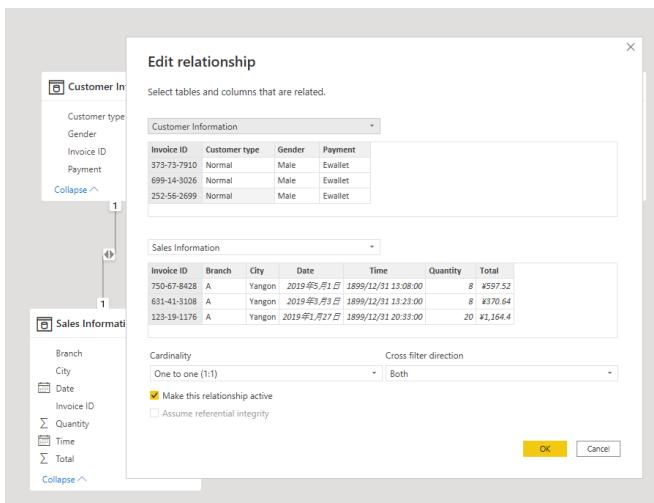


Fig: Edit relationship

3.3 Add new measures

To calculate the member percentage, use both COUNTROWS and FILTER formula.

The screenshot shows the Power BI ribbon with the 'Measures' tab selected. A new measure is being created with the name 'Member %'. The formula bar at the bottom contains the DAX formula: `1 Member % = COUNTROWS(FILTER(CustomerInformation,CustomerInformation[Customer Type] = "Member")) / COUNT([CustomerInformation[Customer Type]])`.

Fig: Member %

I also add below measures.

The screenshot shows the Power BI ribbon with the 'Measures' tab selected. A new measure is being created with the name 'Average COGS'. The formula bar at the bottom contains the DAX formula: `1 Average COGS = AVERAGE([FinancialInformation[COGS]])`.

Fig: Average COGS

The screenshot shows the Power BI ribbon with the 'Measures' tab selected. A new measure is being created with the name 'Average Gross Margin'. The formula bar at the bottom contains the DAX formula: `1 Average Gross Margin = AVERAGE([FinancialInformation[Gross Margin %]])`.

Fig: Average Gross Margin

The screenshot shows the Power BI ribbon with the 'Measures' tab selected. A new measure is being created with the name 'Average Rating'. The formula bar at the bottom contains the DAX formula: `1 Average Rating = AVERAGE([ProductInformation[Rating]])`.

Fig: Average Rating

The screenshot shows the Power BI ribbon with the 'Measures' tab selected. A new measure is being created with the name 'Income Total'. The formula bar at the bottom contains the DAX formula: `1 Income Total = SUM([FinancialInformation[Gross Income]])`.

Fig: Income Total

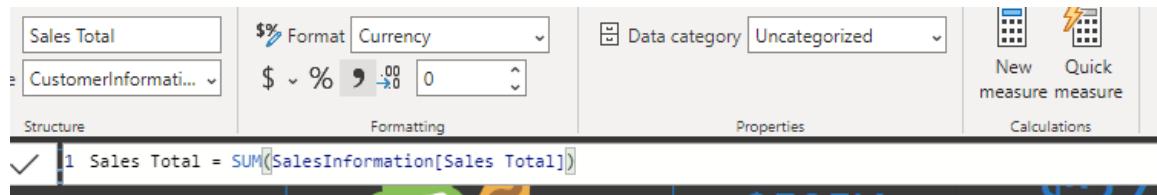


Fig: Sales Total

SECTION 2: BUSINESS REPORT

Executive Summary

1. Purpose of this report

Myanmar's supermarket sales data for 2019 has three branches. This report examines the differences in gross income, business performance, and financial health among these three branches. It helps executives in making strategic decisions and increasing profits.

The report addresses the following questions:

1. Which branch makes the most profit?
2. Which branch has the highest gross margin?
3. Which product makes the most profit?
4. Which product has the highest gross margin? the cost issue
5. Which payment method is most commonly for large deals?
6. The relationship between customer type and sales.
7. The relationship between rating and sales.
8. Which product attract female customers the most? What about male customers?

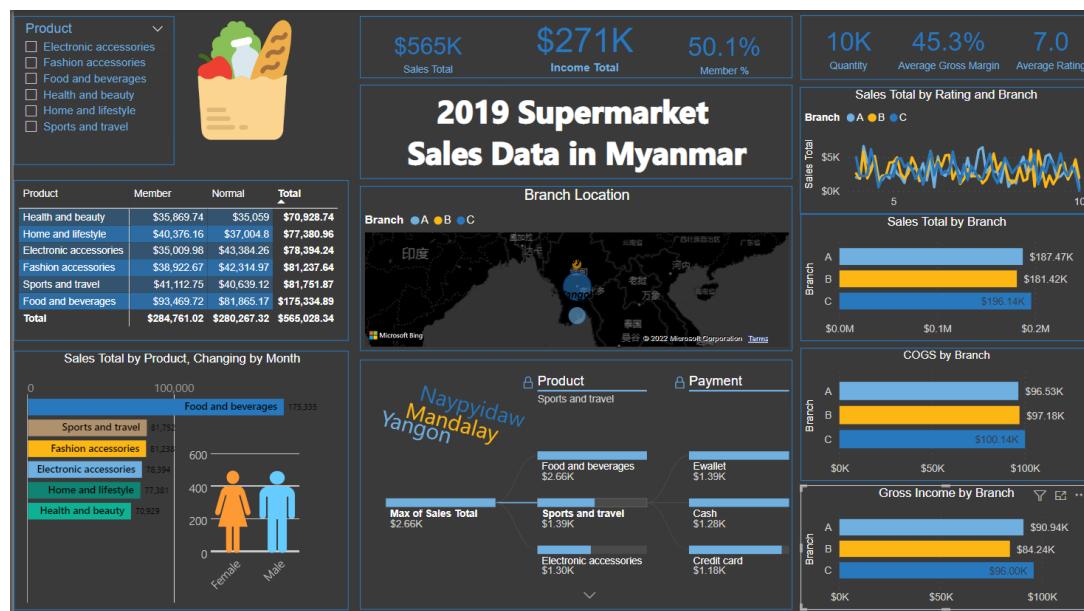


Fig: Screenshot of dashboard

2. Findings and Conclusions

Power BI Dashboard was set up to address the critical questions in the report

In 2019, Branch C located Naypyidaw posted the highest gross income in Myanmar, with profits of approximately 96K U.S. dollars. Followed by Branch A located in Yangon, with profits of 91K U.S. dollars. The next was Branch C located in Mandalay, with profits of 84K U.S. dollars.

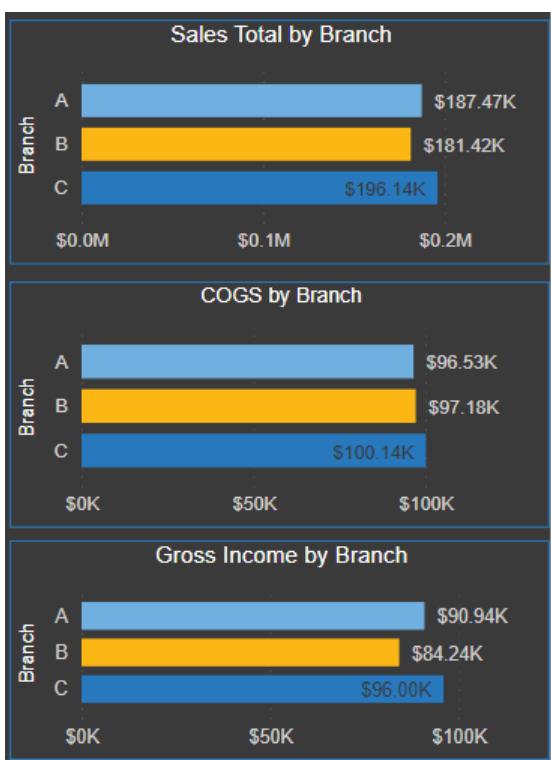


Fig: Sales Total, COGS, Gross Income of Branch A, B, C

In 2019, both Branch B and Branch C had the highest gross margin which was 76.7%, followed by Branch A with 76.6%.



Fig: Branch A



Fig: Branch B



Fig: Branch C

In 2019, “Food and beverages” was the most profitable product in all 3 branches, with a revenue of approximately 135K U.S. dollars.

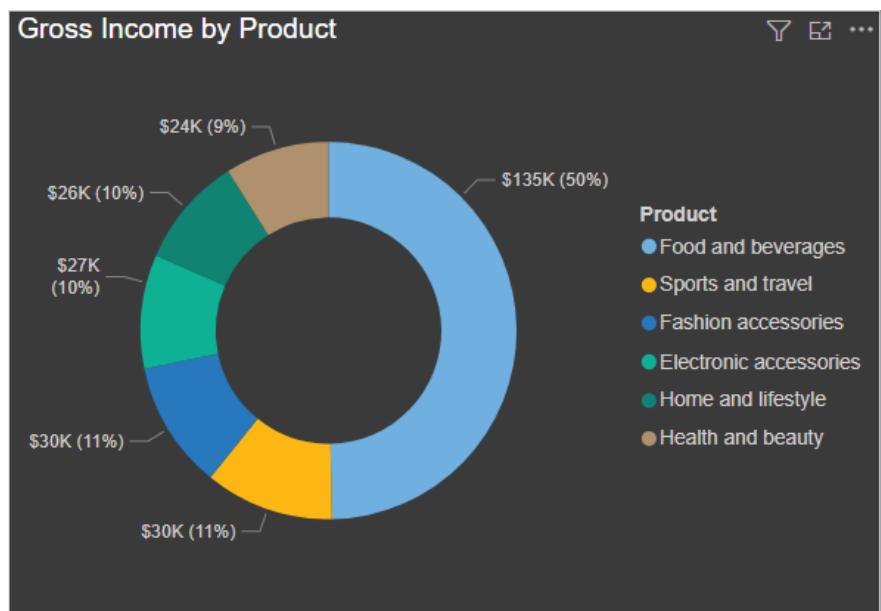


Fig: Pie chart

In 2019, “Food and beverages” had the highest gross margin which was 76.7%, followed by “Fashion accessories” with 40.9%

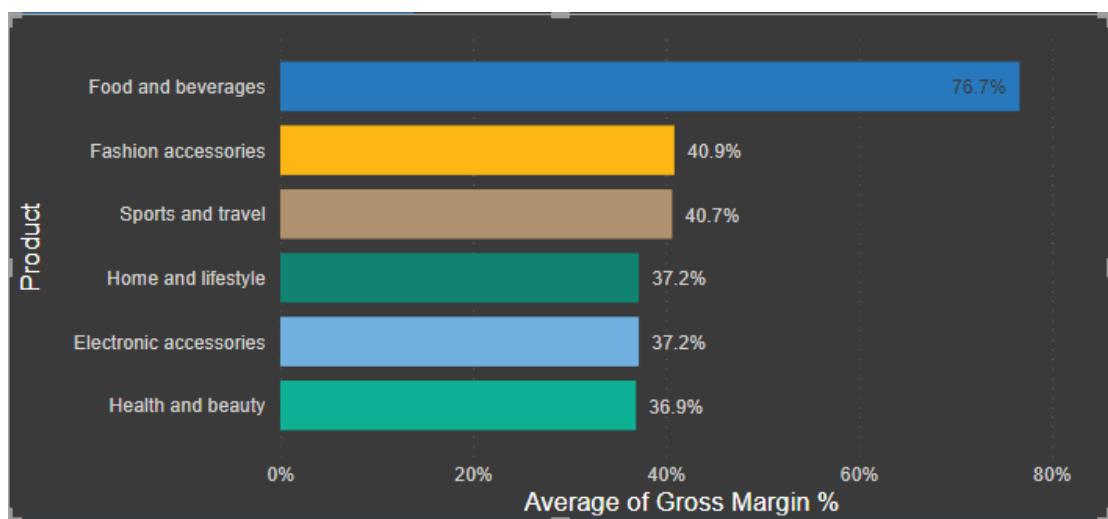


Fig: Animation bar

2/3 Max deal was paid by Ewallet and 1/3 max deal was paid by Credit card at the product level.

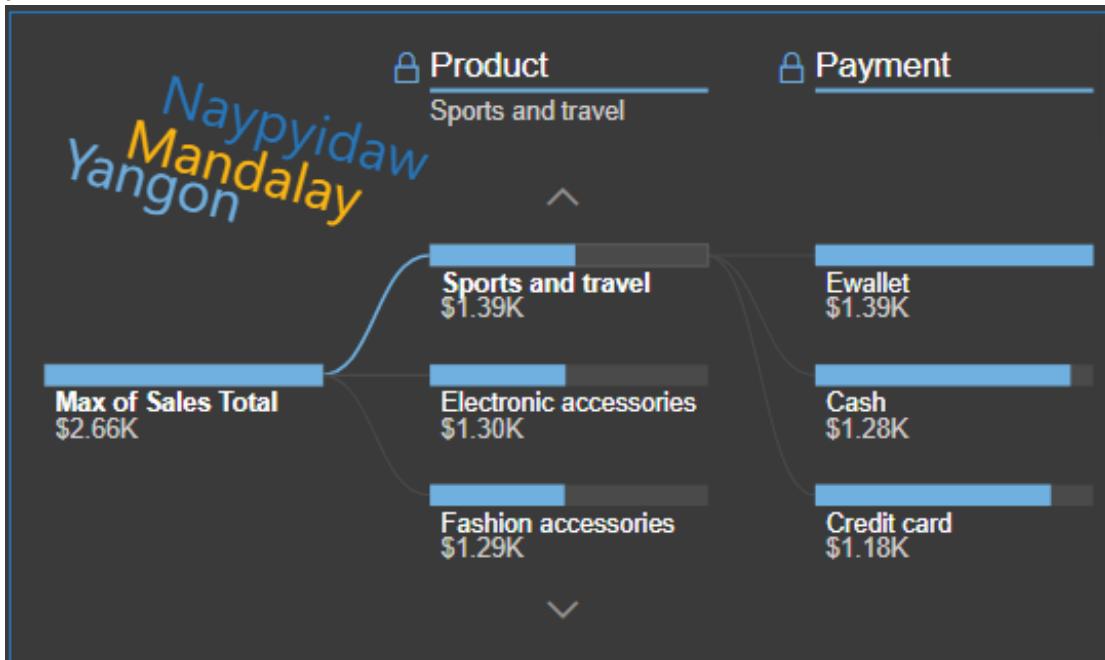


Fig: Decomposition tree

Total sales amount of Members is slightly higher than the Normal, converted to percentage was 8%.

Product	Member	Normal	Total
Food and beverages	\$93,469.72	\$81,865.17	\$175,334.89
Sports and travel	\$41,112.75	\$40,639.12	\$81,751.87
Fashion accessories	\$38,922.67	\$42,314.97	\$81,237.64
Electronic accessories	\$35,009.98	\$43,384.26	\$78,394.24
Home and lifestyle	\$40,376.16	\$37,004.8	\$77,380.96
Health and beauty	\$35,869.74	\$35,059	\$70,928.74
Total	\$284,761.02	\$280,267.32	\$565,028.34

Fig: Matix

No specific relationship was found between rating and sales according to the line chart shown below. The sales didn't increase while the rating is high.

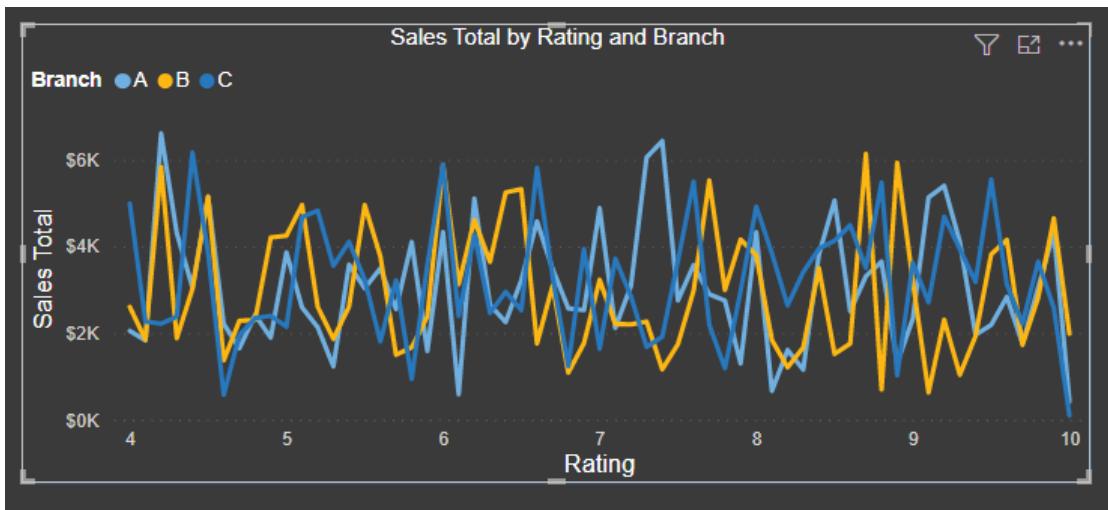


Fig: Line chart

In 2019, the product “Food and beverages” attracted female customers the most, with sales of 97K U.S. dollars, followed by “Fashion accessories”, with sales of 44K U.S. dollars. The product “Food and beverages” attracted male customers the most, with sales of 97K U.S. dollars, followed by “Health and beauty” with sales of 42K U.S. dollars.

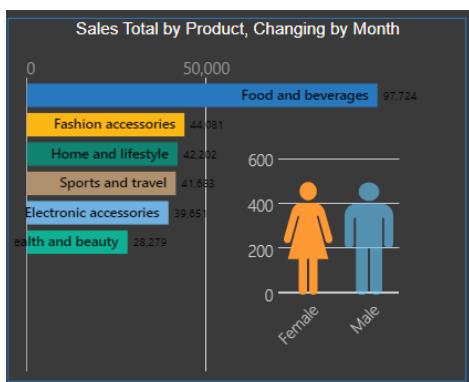


Fig: Choose Woman

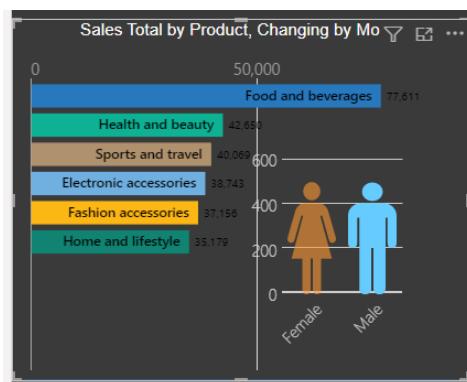


Fig: Choose Man

3. Recommendations for improving performance

At branch level:

- Branch C outperformed than other 3 branches in 2019. It had the highest sales amount and gross income compares to others. However, the gross margin was the same for Branch B and Branch C, implying that Branch C's cost is higher. It should focus on reducing the cost in the next fiscal year.
- Branch B needs to increase its sales and rating score.

At customer level:

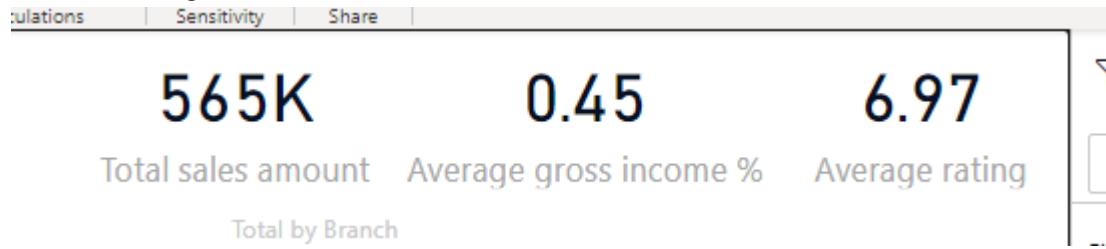
- Increase the number of customers. The purchasing power of members has not significantly increased according to the sales data. The policy should be made to keep customer engagement and loyalty to the company, like discounts, rewards, free delivery, etc.
- Identify customers' needs. Genders have different shopping hobbies. It should differentiate between male and female customers' needs and use bundled sales in different product categories.

At product level:

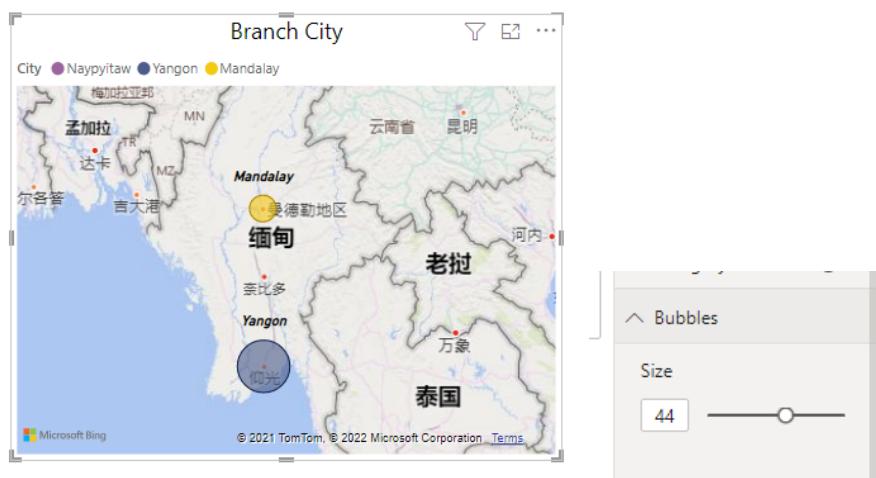
- It should add more categories in "Food and beverages". In 2019, "Food and beverages" provide approximately half profits for all the branches.
- "Electronic accessories" had the lowest gross margin in 2019. The sales were not satisfying compared to others. As this product isn't used very often in everyday life, so either reducing inventory or raising the unit price is a good way to maximize profits.

4. Steps of building the report

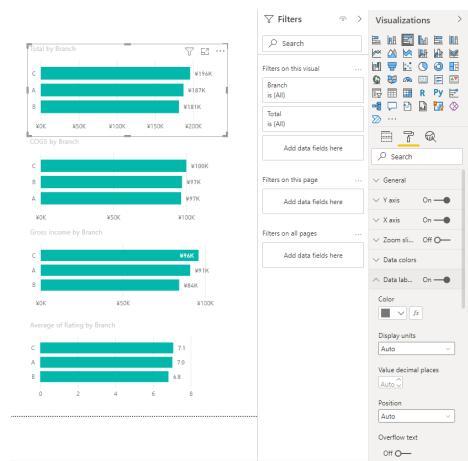
After creating measures, add measures into cards.



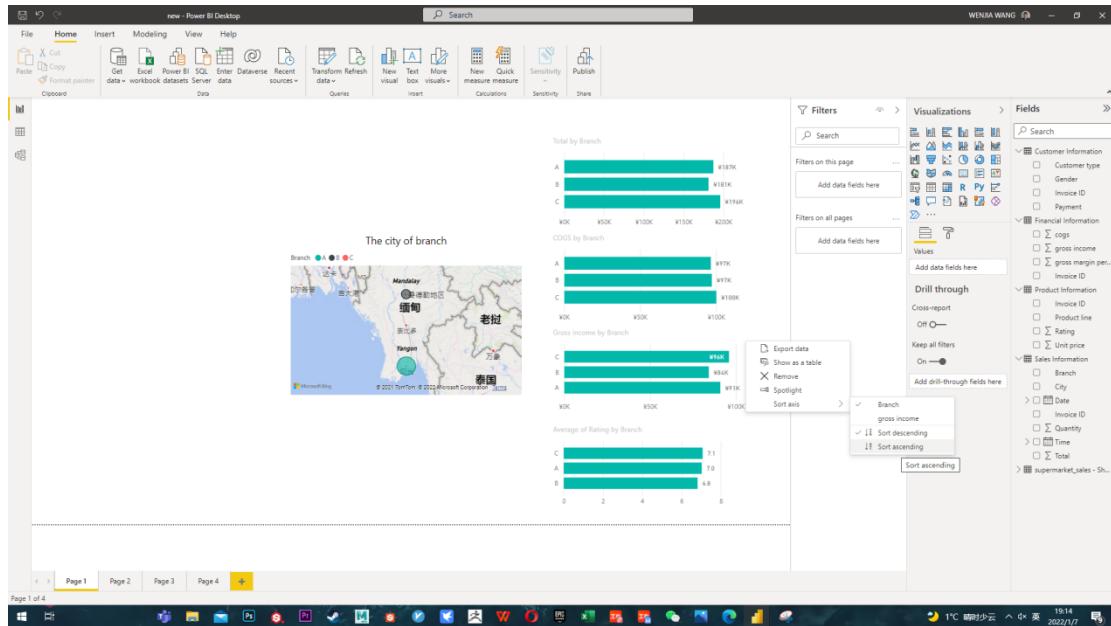
Create a map and adjust bubble size and color.



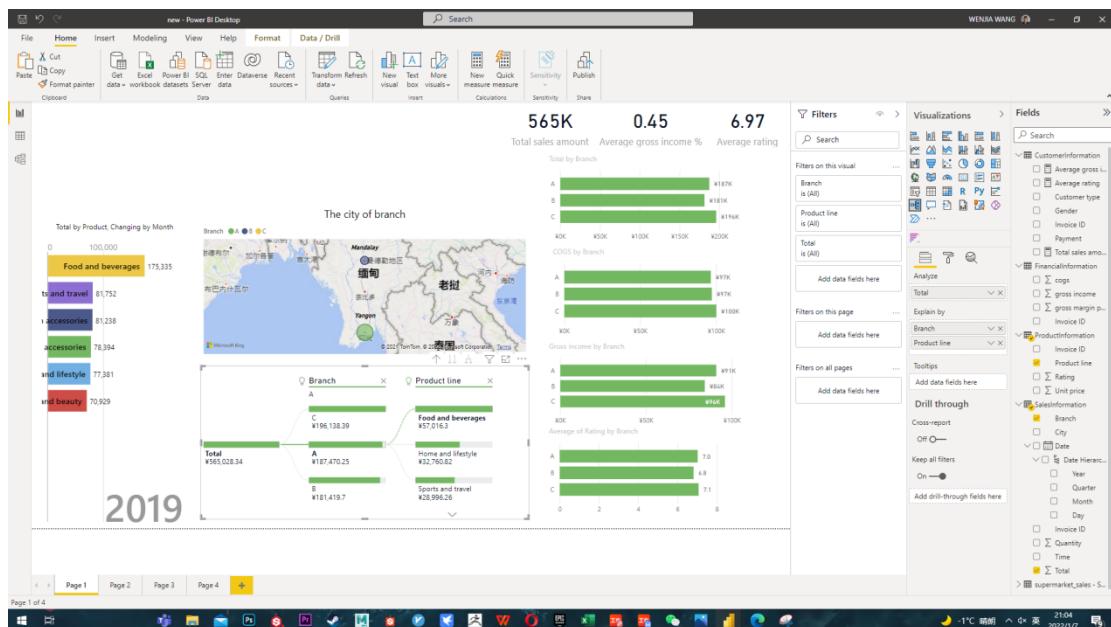
Add clustered bar chart.



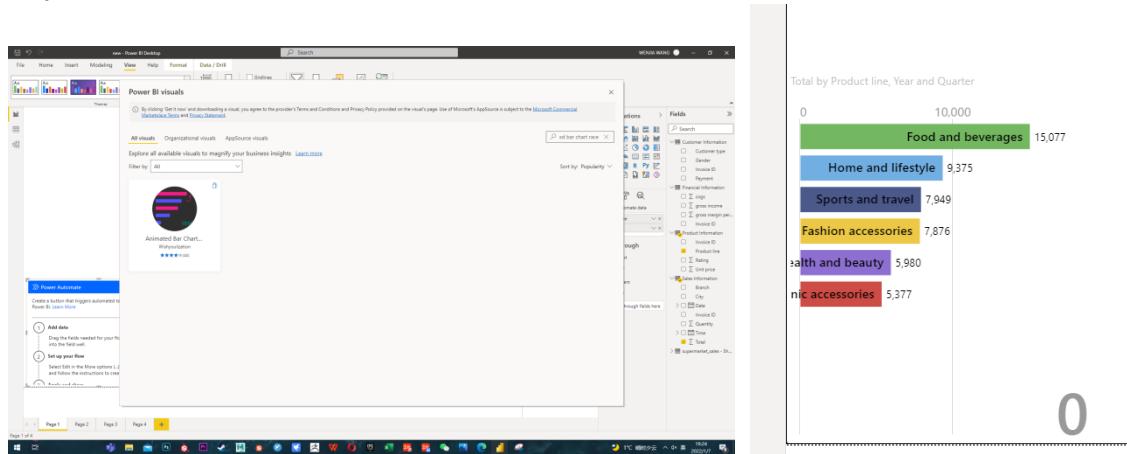
Uniform sorting



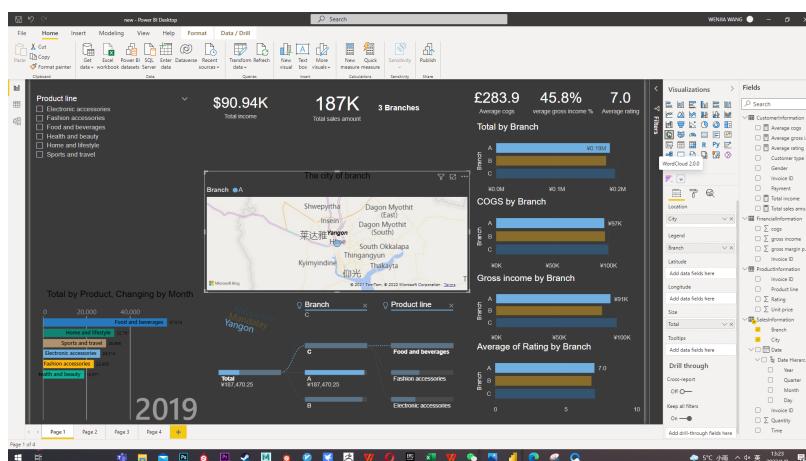
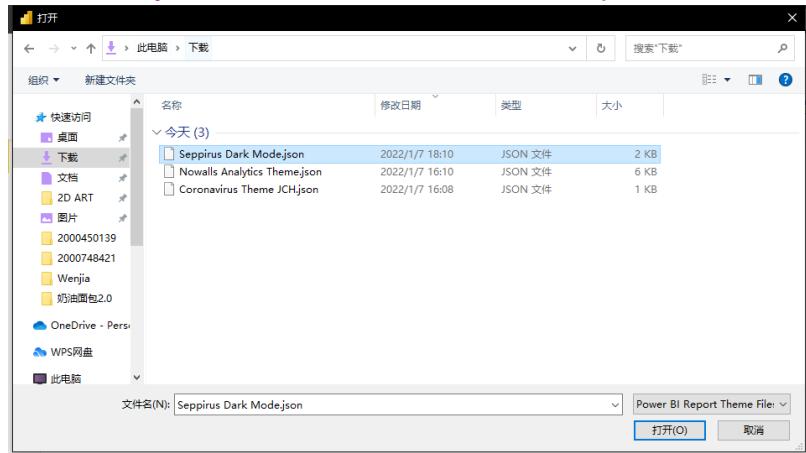
Add more visuals.



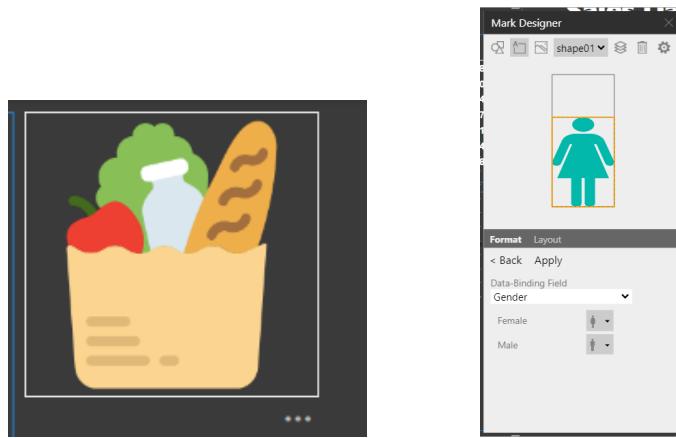
Import new visual called “Animated Bar Chart Race”.



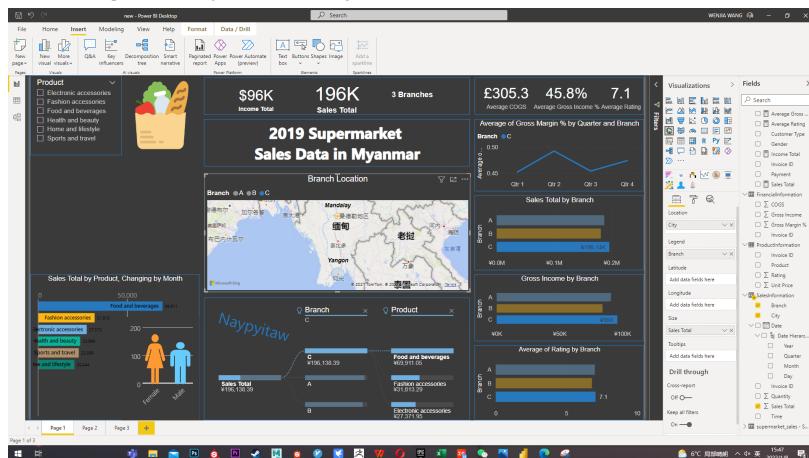
Import theme from website [Seppirus Dark Mode Theme - Microsoft Power BI Community](#). Choose the theme and click open



Import picture and use “Infographic and designer” to change the bar shape.



According to the questions adjust visuals.



Finish the final version of report.

