



UTM
UNIVERSITI TEKNOLOGI MALAYSIA

MCSD1123 BIG DATA MANAGEMENT

Assignment 1 :

Data Analysis Using Google Sheet

Case Study 1b :

Sales Performance

Lecturer's Name :

PM Dr. Mohd Shahizan Othman

Group Name :

F4

Group Members	Matrix Number
LYE KAH HOOI	MCS231010
LEE SEOW MING THERESA	MCS231013
THONG YEE MOON	MCS231001
SITI NORAFIZAH BINTI AB AZIZ	MCS231018

1.0 INTRODUCTION

A dataset named “dataset2.txt” was given with the purpose of sales performance analysis. The dataset contained 324 records. This dataset contains nine attributes such as *customer*, *product*, *salesperson*, etc as shown in Table 1.

Table 1: Attributes Information

Attributes	Sample Data
Customer	99 SpeedMart, Aeon Mall
Products	Flour, Sugar
Sales Person	Saharulnizam, Hadirah
Sales Region	North, West
Target	11681, 61690
Sales	15691.32,89032.32
Sales Year	2020, 2021
Sales Month	Apr, Jun
Sales QTR	Q2, Q3

In today's competitive business landscape, understanding and optimizing sales performance is paramount to the success and growth of any organization. Sales performance insight analysis is a strategic approach that empowers businesses to make data-driven decisions, enhance revenue generation, and improve customer relationships.

Data preprocessing and data visualization are essential steps in any data analysis or machine learning project for several reasons. Certainly, using Google Sheets for data preprocessing and data visualization is a practical choice, especially when dealing with smaller datasets.

Dashboards were created in Google Sheets that include a variety of graphs to represent different aspects of sales performance. There are three menus on the left side menu: sales year, sales region, and products. Customers and salespeople can also be found in the right side menu. To visualize information about sales performance, five charts of various shapes are used. Monthly sales, region, customer, salesperson, and sales trend are all represented in the chart.

2.0 METHODOLOGY - DATA PREPROCESSING

This dataset contained nine columns, *customer*, *product*, and *salesperson*. as shown in Table 1.

Table 2: Attributes Data Type

Attributes	Sample Data	Datatype
Customer	99 SpeedMart, Aeon Mall	Text
Products	Flour, Sugar	Text
Sales Person	Saharulnizam, Hadirah	Text
Sales Region	North, West	Text
Target	11681, 61690	Integer
Sales	15691.32, 89032.32	Integer
Sales Year	2020, 2021	Text
Sales Month	Apr, Jun	Text
Sales QTR	Q2, Q3	Text

2.1 Transform data type Text to Date.

- Attributes for 'Sales Month' and 'Sales Year' were needed to change their data type from Text to Date in order to create different charts.

A	B	C	D	E	F	G	H	I
CUSTOMER	PRODUCTS	SALES PERSON	Sales Region	Target	SALES	SALES YEAR	SALES MONTH	SALES QTR
99 SpeedMart	Flour	Saharulnizam	North	11681	15691.32	2020	Apr	Q2
99 SpeedMart	Flour	Saharulnizam	North	61690	89032.32	2020	Jun	Q2
99 SpeedMart	Flour	Saharulnizam	North	10000	13068.27	2020	Jul	Q3
99 SpeedMart	Flour	Saharulnizam	North	69399	90415.62	2020	Nov	Q4
99 SpeedMart	Flour	Saharulnizam	North	60190	74473.74	2020	Oct	Q4
99 SpeedMart	Flour	Saharulnizam	North	43948	48816.135	2020	Dec	Q4
99 SpeedMart	Flour	Saharulnizam	North	69301	65906.415	2020	May	Q2
99 SpeedMart	Flour	Saharulnizam	North	35258	24770.205	2020	Sept	Q3
Aeon Mall	Flour	Hadirah	West	64416	108643.86	2020	Nov	Q4
Aeon Mall	Flour	Hadirah	West	62753	96949.755	2020	Sept	Q3

Figure 1 : Dataset 1b

Dataset2 ☆ 📁 ☁

File Edit View Insert Format Data Tools Extensions Help

Q Menus 🔍 ↶ ↷ 🖨 📋 100% 🔍 \$ % .0 .00 123 Default... - 10 + B I 🔍 A 🗑 📊 📈 📉 📉 📉 📉 📉

J2 $\text{=DATE(G2,MONTH(DATEVALUE((Left(H2,3)\&"1"))),1)}$

	A	B	C	D	E	F	G	H	I	J
1	CUSTOMER	PRODUCTS	SALES PERSON	Sales Region	Target	SALES	SALES YE.	SALES MONTH	SALES QTR	SALES MONTH 2
2	99 SpeedMart	Flour	Saharulnizam	North	11681	15691.32	2020	Apr	Q2	Apr
3	99 SpeedMart	Flour	Saharulnizam	North	61690	89032.32	2020	Jun	Q2	Jun
4	99 SpeedMart	Flour	Saharulnizam	North	10000	13068.27	2020	Jul	Q3	Jul
5	99 SpeedMart	Flour	Saharulnizam	North	69399	90415.62	2020	Nov	Q4	Nov
6	99 SpeedMart	Flour	Saharulnizam	North	60190	74473.74	2020	Oct	Q4	Oct
7	99 SpeedMart	Flour	Saharulnizam	North	43948	48816.135	2020	Dec	Q4	Dec
8	99 SpeedMart	Flour	Saharulnizam	North	69301	65906.415	2020	May	Q2	May

Figure 2 : Formula

- b. One column named “Sales Month 2” was added. Key in this formula which aims to convert the type of data from text to month.

$\text{=DATE(G2,MONTH(DATEVALUE((Left(H2,3)\&"1"))),1)}$

- c. Understand the function that used in process (b)
- DATE function was used to take three separate values and combine them to form a date.
 - MONTH function returns the month of a date represented by a serial number. The month was given as an integer, ranging from 1 (January) to 12 (December).
 - DATEVALUE function assist on converting a date that is stored as text to a serial number that Excel recognizes as a date.
 - LEFT function returns the first character or characters in a text string, based on the number of characters have been specified.

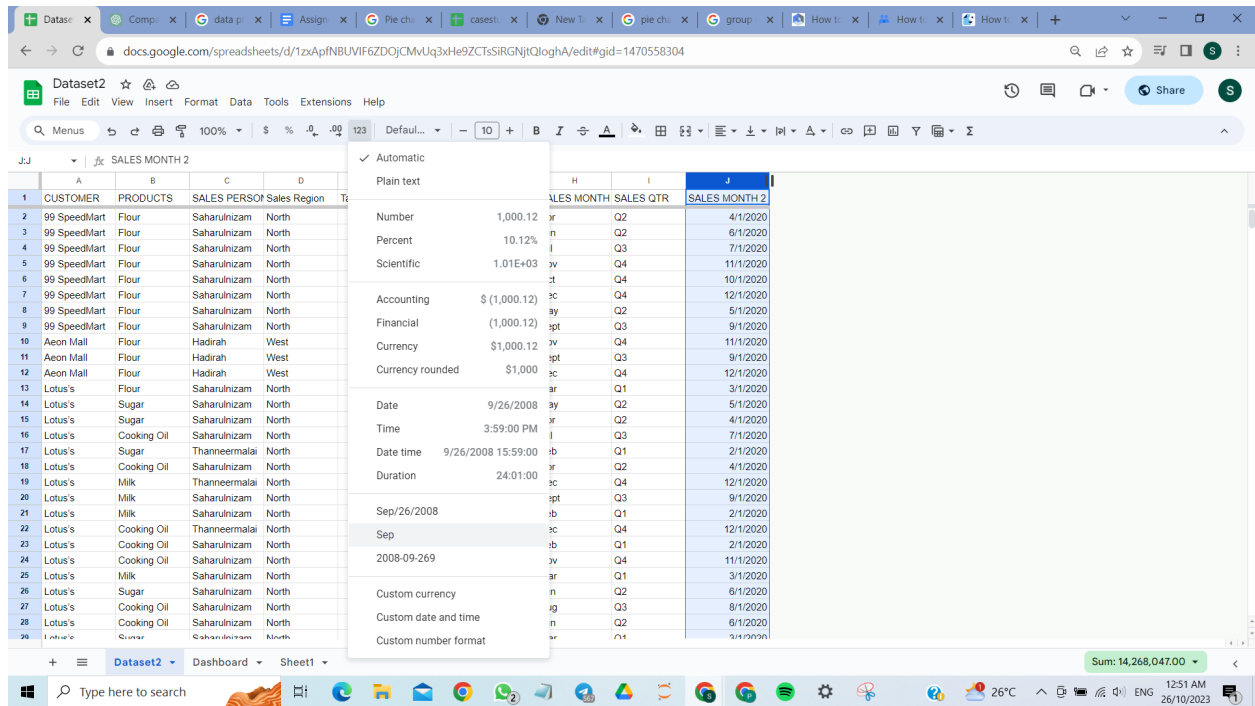


Figure 3 : Month value with date type

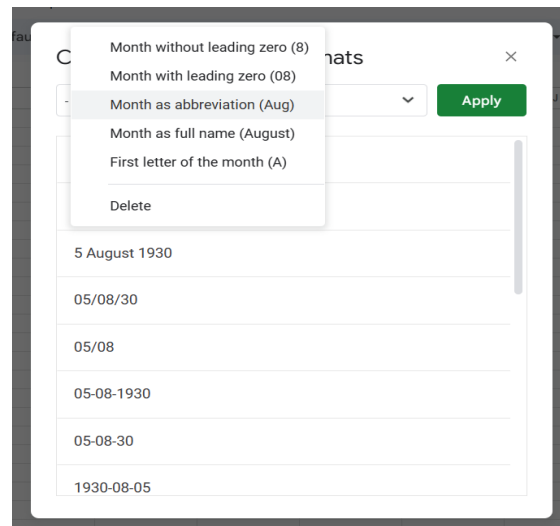
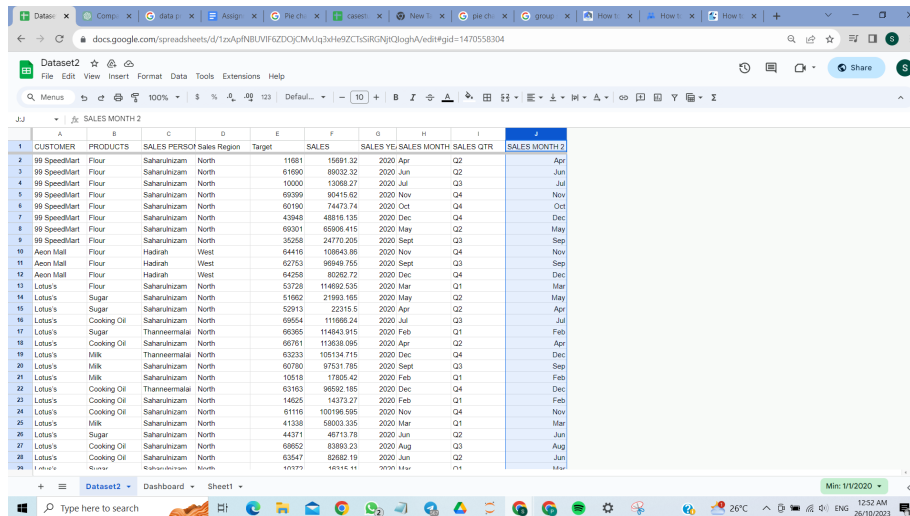


Figure 4 : Set Data Type Month as Month as abbreviation.



3.0 METHODOLOGY - DATA VISUALIZATION

Dashboards were created in Google Sheets that include a variety of graphs to represent different aspects of sales performance. There were three menus on the left side menu: sales year, sales region, and products. Customers and salespeople can also be found in the right side menu. To visualize information about sales performance, five charts of various shapes were used. Monthly sales, region, customer, salesperson, and sales trend were all represented in the chart.

3.1 Dashboard

[illegible]

Figure 6 : Dashboard Design Layout

3.2 Filter charts and tables with Slicers

A custom dashboard can be included a slicer to filter the tables, charts, or pivot tables.



Figure 7 : List of Menu Dashboard

Table 3 : Slicer Configuration

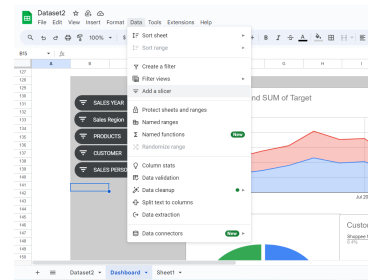
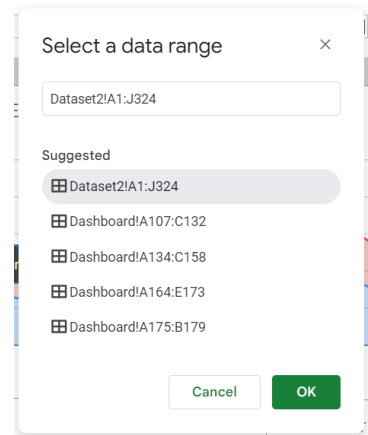
Figure	Configuration
	At the top, Data was clicked and then a Slicer can be added.
	A data range was selected from a chart or pivot table that needs to be filtered.
Sales Year	A column was chosen as an item that needed to filter.

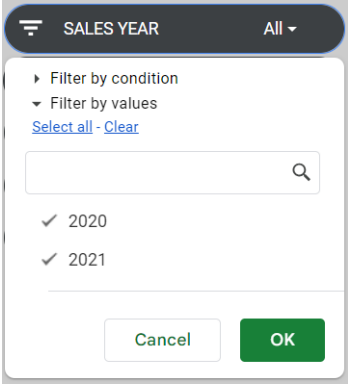
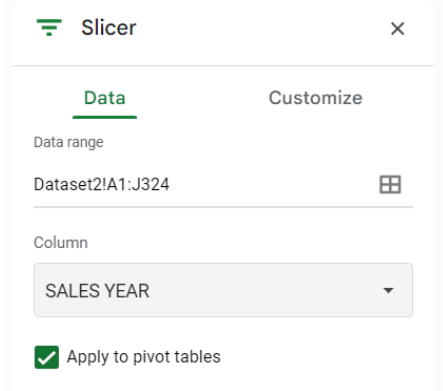
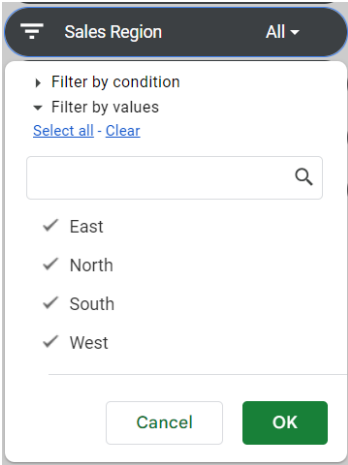
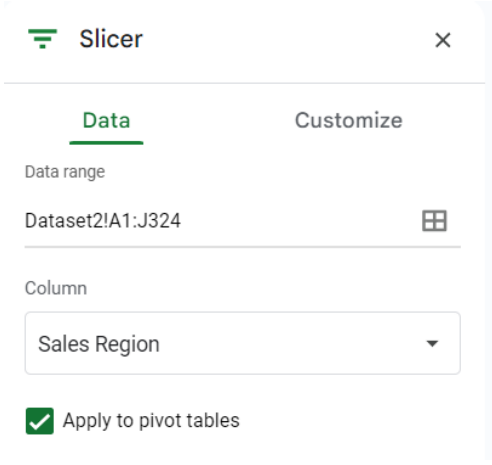
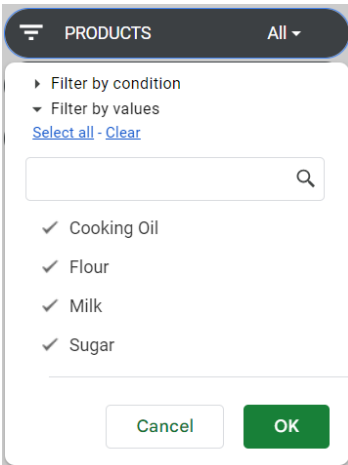
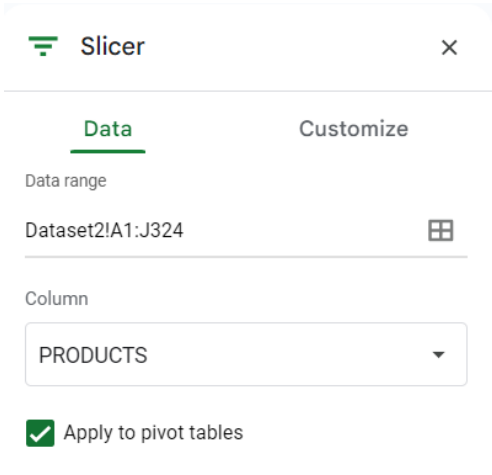
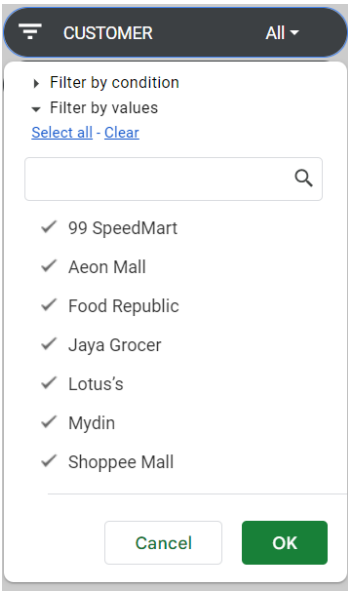
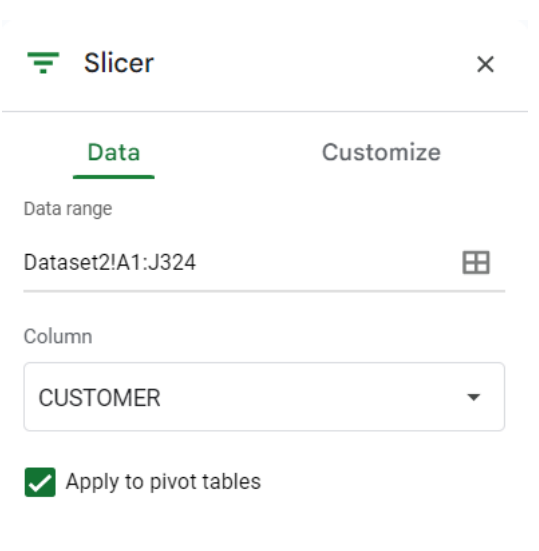
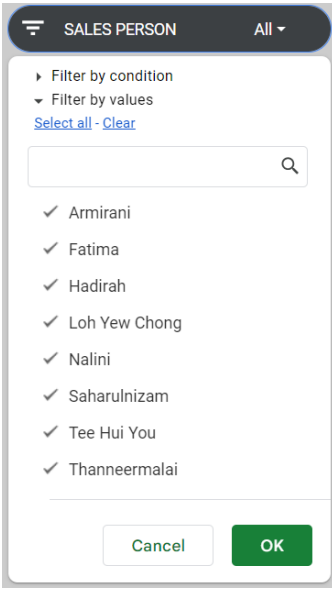
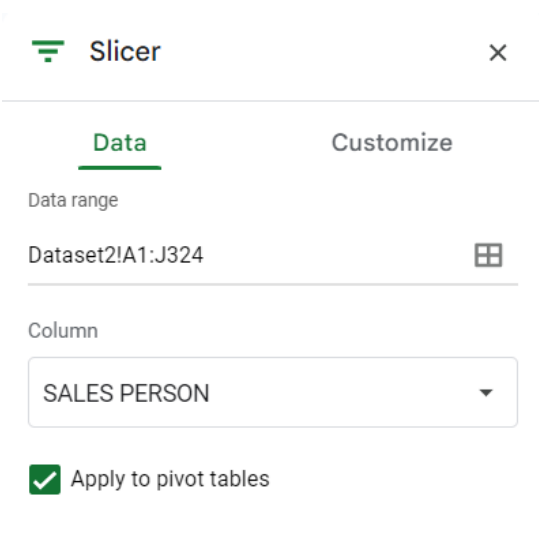
Figure	Configuration
	
<p>Sales Region</p> 	<p>A column was chosen as an item that needed to filter.</p> 
<p>Products</p> 	<p>A column was chosen as an item that needed to filter.</p> 

Figure	Configuration
<p>Customer</p> 	<p>A column was chosen as an item that needed to filter.</p> 
<p>Sales Person</p> 	<p>A column was chosen as an item that needed to filter.</p> 

3.3 Pie charts

Pie charts are used to compare parts of a single data series to the whole.

Table 4 : Pie Chart (Region) Configuration

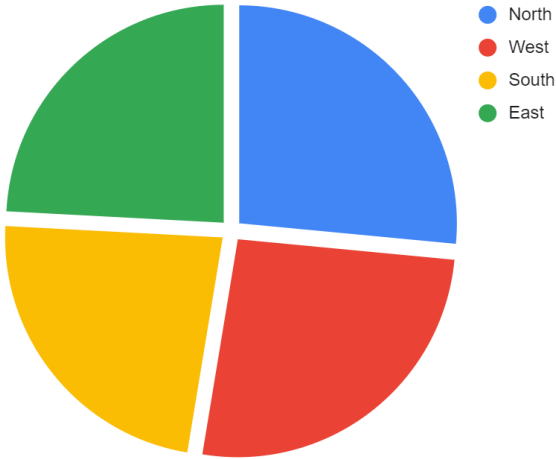
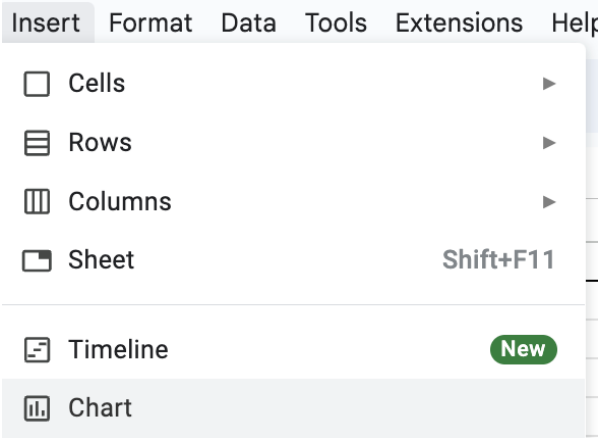
Figure	Configuration
<div>Regions</div> 	<p>At the top, “Insert” and “Chart” buttons were clicked to insert.</p> 

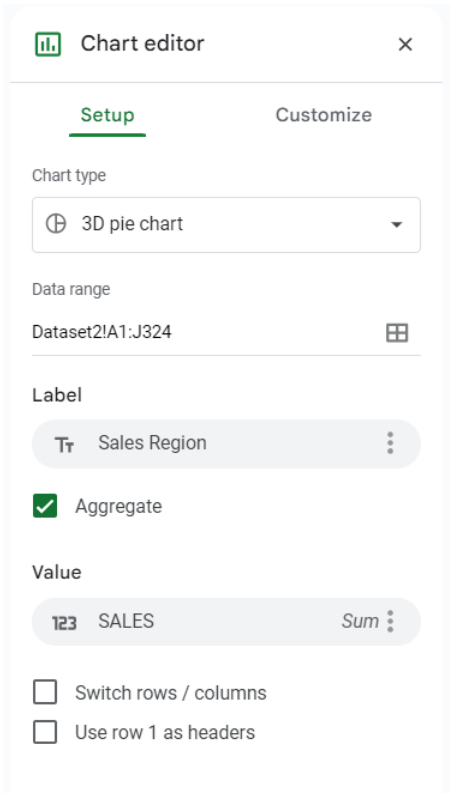
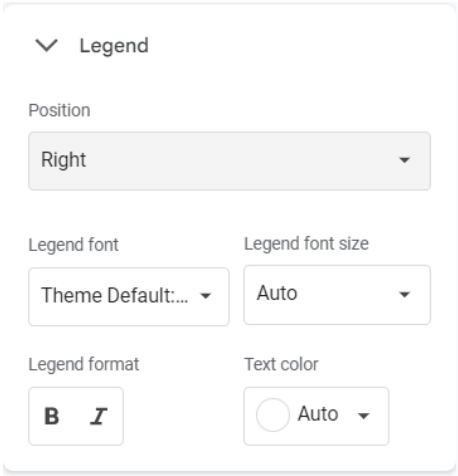
Figure	Configuration
 <p>Chart editor</p> <p>Setup Customize</p> <p>Chart type</p> <p>3D pie chart</p> <p>Data range</p> <p>Dataset2!A1:J324</p> <p>Label</p> <p>Sales Region</p> <p><input checked="" type="checkbox"/> Aggregate</p> <p>Value</p> <p>SALES Sum</p> <p><input type="checkbox"/> Switch rows / columns</p> <p><input type="checkbox"/> Use row 1 as headers</p>	<p>Next, with a default chart on a sheet, the chart editor can be used to set up.</p> <p>“Setup” on the chart editor was used , then “3D Pie Chart” was selected on the “Chart Type” Section</p>
 <p>Legend</p> <p>Position</p> <p>Right</p> <p>Legend font</p> <p>Theme Default...</p> <p>Legend font size</p> <p>Auto</p> <p>Legend format</p> <p>B <i>I</i></p> <p>Text color</p> <p>Auto</p>	<p>With the“Setup” on the chart editor, the legend can be set.</p>

Figure	Configuration
<div><div><div><div><div><div></div></div><div>Chart editor</div><div>×</div></div></div><div><div><div>Setup</div><div>Customize</div></div><div><div>Chart type</div><div><div>⊕ Pie chart</div><div>▼</div></div><div><div>Data range</div><div>Dataset2!A1:J324</div><div>⌵</div></div><div><div>Label</div><div><div>T CUSTOMER</div><div>⋮</div></div><div><div>✓ Aggregate</div></div><div><div>Value</div><div><div>123 SALES</div><div>Sum ⋮</div></div><div><div><input type="checkbox"/> Switch rows / columns</div><div><input type="checkbox"/> Use row 1 as headers</div></div></div></div></div></div></div></div>	<p>With a default chart on a sheet, and a chart editor showed up on the right.</p> <p>At the “Setup” on the chart editor, then “Pie Chart” was selected at the “Chart type” .</p>

Figure	Configuration
<div><div><div>▼ Legend</div><div>Position</div><div>Right ▼</div><div>Legend font</div><div>Theme Default:... ▼</div><div>Legend font size</div><div>Auto ▼</div><div>Legend format</div><div>B <i>I</i></div><div>Text color</div><div><input type="radio"/> Auto ▼</div></div></div>	At “Setup” on the chart editor, the legend can be set.
<div><div><div>▼ Chart & axis titles</div><div>Chart title</div><div>Title text</div><div>Customer </div><div>Title font</div><div>Theme Defaul... ▼</div><div>Title font size</div><div>Auto ▼</div><div>Title format</div><div>B <i>I</i> ≡ ▼</div><div>Title text color</div><div><input type="radio"/> Auto ▼</div></div></div>	At “Setup” on the chart & axis titles editor, the title text was added and edited.

3.4 A Stackable Chart was created.

SUM of SALES	PRODUCTS			
SALES PERSON	Cooking Oil	Flour	Milk	Sugar
Armirani	318982.455	420286.995	168425.91	634899.465
Fatima		313810.74		217972.845
Hadirah	1003417.11	1193477.31	712535.22	1225137.915
Loh Yew Chong	1236867.465	1176358.32	1306892.305	1020659.845
Nalini	1138519.845	474678.09	1005811.785	906653.97
Saharulnizam	1268227.7	1040760.99	839137.185	615445.195
Tee Hui You	300533.67	469151.415	514363.14	277962.39
Thanneermalai	351274.68	327641.13	496517.265	843601.59

Figure 8 : Pivot Table Analysis

Pivot table editor

Dataset2IA1:J324

Suggested

Rows

Add

SALES PERSON

OrderAscendingSort bySALES PERS...Show totals

Columns

Add

PRODUCTS

OrderAscendingSort byPRODUCTSShow totals

Values

Add

SALES

Summarize bySUMShow asDefault

Filters

Add

Search

CUSTOMER
PRODUCTS
SALES PERSON
Sales Region
Target
SALES
SALES MONTH
SALES YEAR
Sales Date
SALES QTR

Figure 9 : Pivot Table Configuration

Salesperson

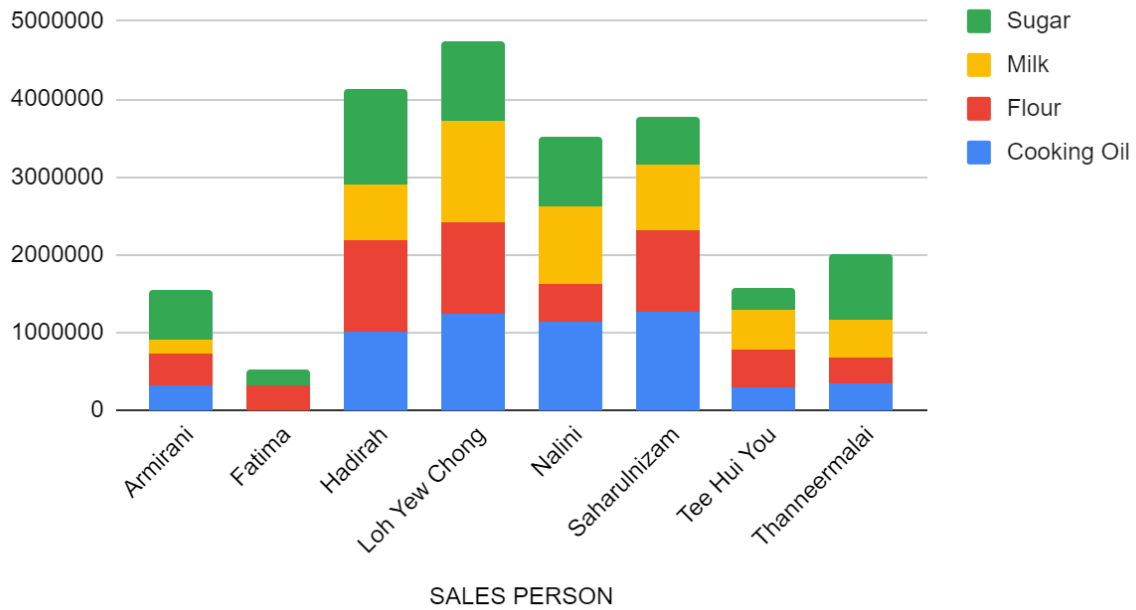


Figure 10 : Stackable Chart

Pivot table editor

Dataset2IA1:J324

Suggested

Rows

SALES PERSON

Order

Ascending

Sort by

SALES PERS...

☐ Show totals

Columns

PRODUCTS

Order

Ascending

Sort by

PRODUCTS

☐ Show totals

Values

SALES

Summarize by

SUM

Show as

Default

Q Search

CUSTOMER

PRODUCTS

SALES PERSON

Sales Region

Target

SALES

SALES YEAR

SALES MONTH

SALES QTR

SALES MONTH 2

Figure 11 : Stackable Chart Configuration

16

3.5 Stacked Area Chart was created

SALES YEAR	SALES MONTH	SUM of SALES	SUM of Target
2020	Jan	895121.455	732528
	Feb	952593.885	763810
	Mar	1097572.86	757446
	Apr	898809.615	660898
	May	987750.585	850575
	Jun	918419.85	731521
	Jul	996588.045	808361
	Aug	844524.225	676797
	Sep	1062939.465	869152
	Oct	1087126.325	820578
	Nov	1250867.85	975960
	Dec	1097634.195	865768
2021	Jan	905612.58	665570
	Feb	797595.12	663566
	Mar	886755.33	705943
	Apr	737216.685	602086
	May	846052.505	672309
	Jun	760373.91	698661
	Jul	822601.46	732865
	Aug	414286.605	397931
	Sep	954008.505	724404
	Oct	868699.35	618810
	Nov	682015.67	555424
	Dec	1054837.865	818925

Figure 12 : Pivot Table Analysis

Pivot table editor

×

Dataset2!A1:J324

Suggested

Rows

SALES YEAR

Order

Ascending

Sort by

SALES YEAR

☐ Show totals
☐ Repeat row labels

SALES MONTH 2

Order

Ascending

Sort by

SALES MONTH 2

☐ Show totals

Columns

Values as:

Columns

SALES

Summarize by

SUM

Show as

Default

Target

Summarize by

SUM

Show as

Default

Search

CUSTOMER

PRODUCTS

SALES PERSON

Sales Region

Target

SALES

SALES YEAR

SALES MONTH

SALES QTR

SALES MONTH 2

Figure 13: Pivot Table Configuration

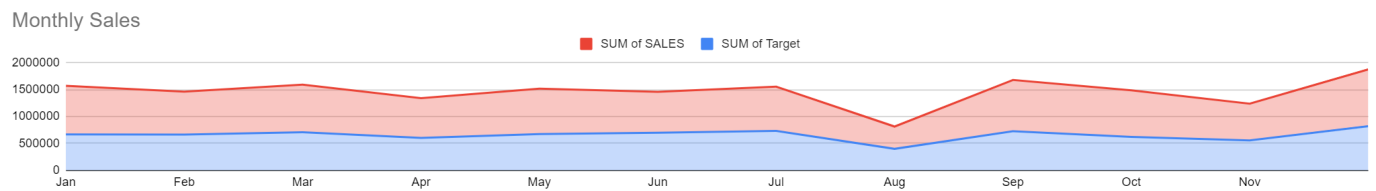




Figure 14 : Stacked Area Chart

 Chart editor ×

Setup

Customize

Chart type

 Stacked area chart


▼

Stacking


Standard

▼

Data range

A133:D157 

X-axis

 SALES MONTH 2

⋮

☐ Aggregate

Series

123

SUM of Target

⋮

123

SUM of SALES

⋮

Add Series

☐ Switch rows / columns

☒ Use row 133 as headers

☒ Use column B as labels

☒ Treat labels as text

Figure 15 : Stacked Area Chart Configuration

3.6 Smooth Line Chart was created

SALES YEAR	SALES MONTH	SUM of SALES
2020	Jan	895121.455
	Feb	952593.885
	Mar	1097572.86
	Apr	898809.615
	May	987750.585
	Jun	918419.85
	Jul	996588.045
	Aug	844524.225
	Sep	1062939.465
	Oct	1087126.325
	Nov	1250867.85
	Dec	1097634.195
2021	Jan	905612.58
	Feb	797595.12
	Mar	886755.33
	Apr	737216.685
	May	846052.505
	Jun	760373.91
	Jul	822601.46
	Aug	414286.605
	Sep	954008.505
	Oct	868699.35
	Nov	682015.67
	Dec	1054837.865

Figure 16 : Pivot Table Analysis

Pivot table editor

Dataset2!A1:J324

Suggested

Rows

Add

SALES YEAR

X

OrderAscending

Sort bySALES YEAR

Show totals

Repeat row labels

SALES MONTH 2

X

OrderAscending

Sort bySALES MONTH 2

Show totals

Columns

Add

Values

Add

SALES

X

Summarize bySUM

Show asDefault

Filters

Add

Search

CUSTOMER

PRODUCTS

SALES PERSON

Sales Region

Target

SALES

SALES YEAR

SALES MONTH

SALES QTR

SALES MONTH 2

Figure 17 : Pivot Table Configuration



Figure 18: Smooth Line Chart

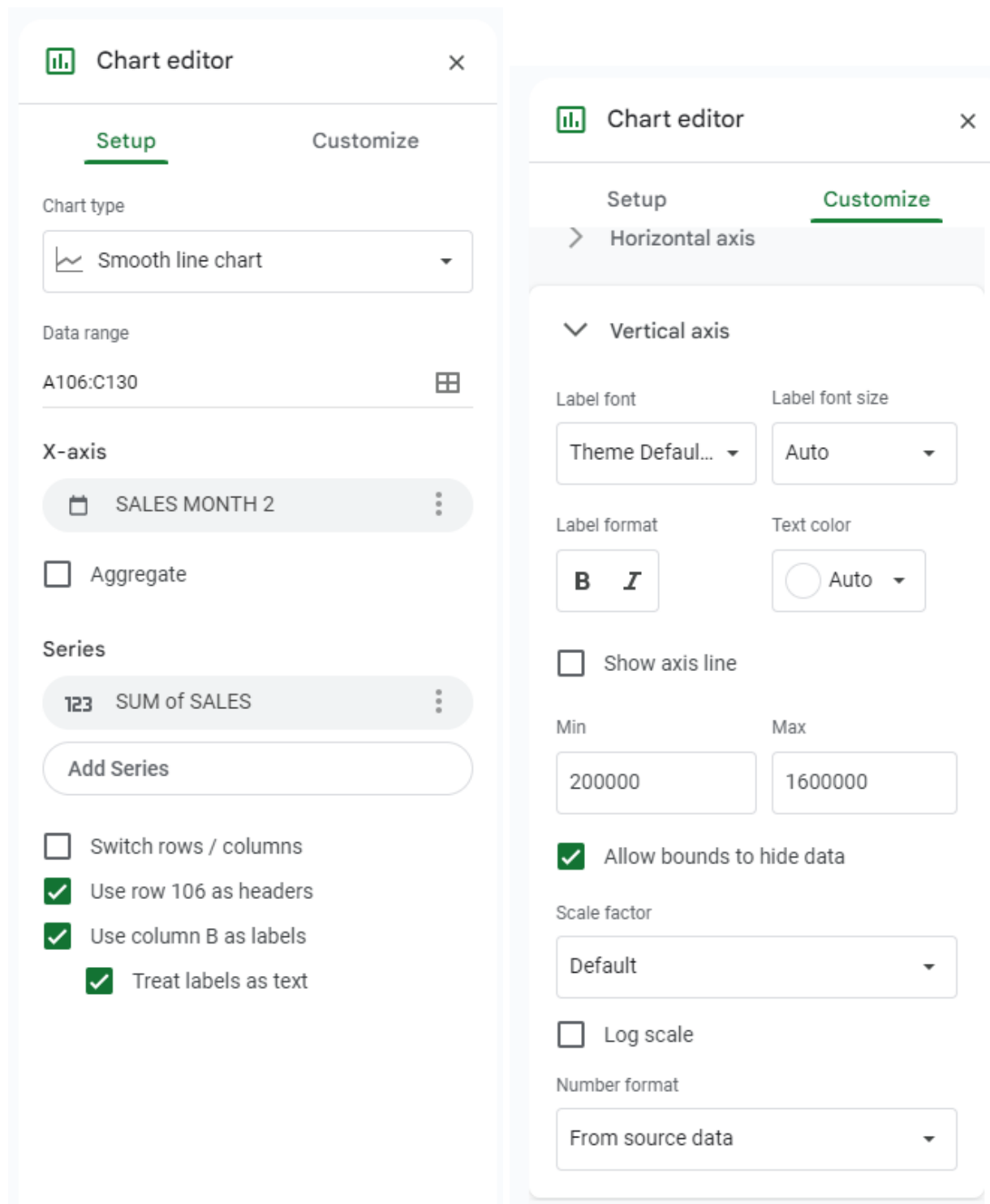
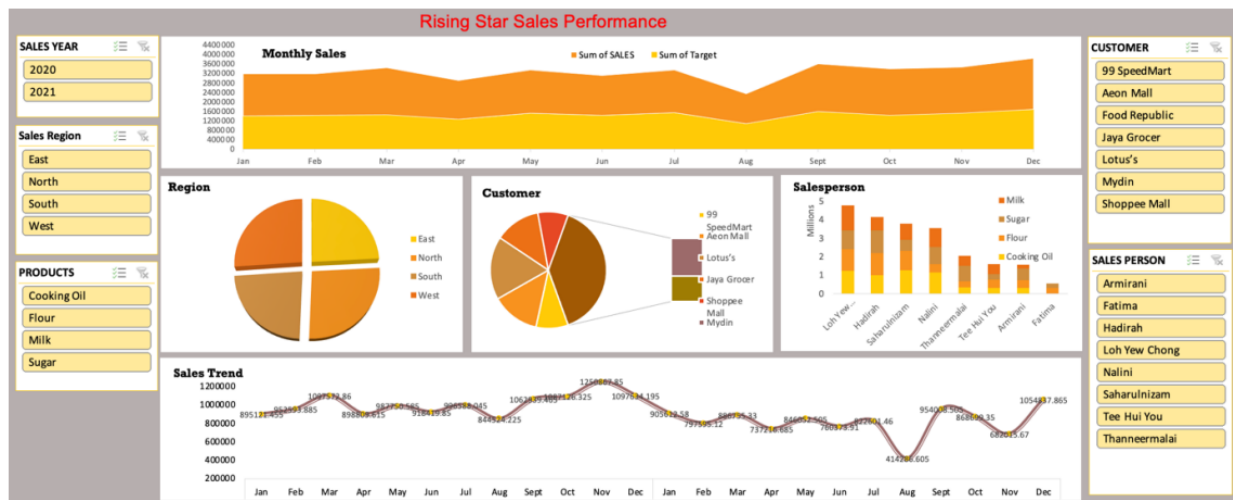


Figure 19 : Smooth Line Chart Configuration

4.0 OUTCOME

4.1 Question 1

3. You must create a dashboard similar to Figure 7. There are three menus on the left side menu: sales year, sales region, and products. Customers and salespeople can also be found in the right side menu. To visualise information about sales performance, five charts of various shapes are used. Monthly sales, region, customer, salesperson, and sales trend are all represented in the chart.



Answer for Question 1

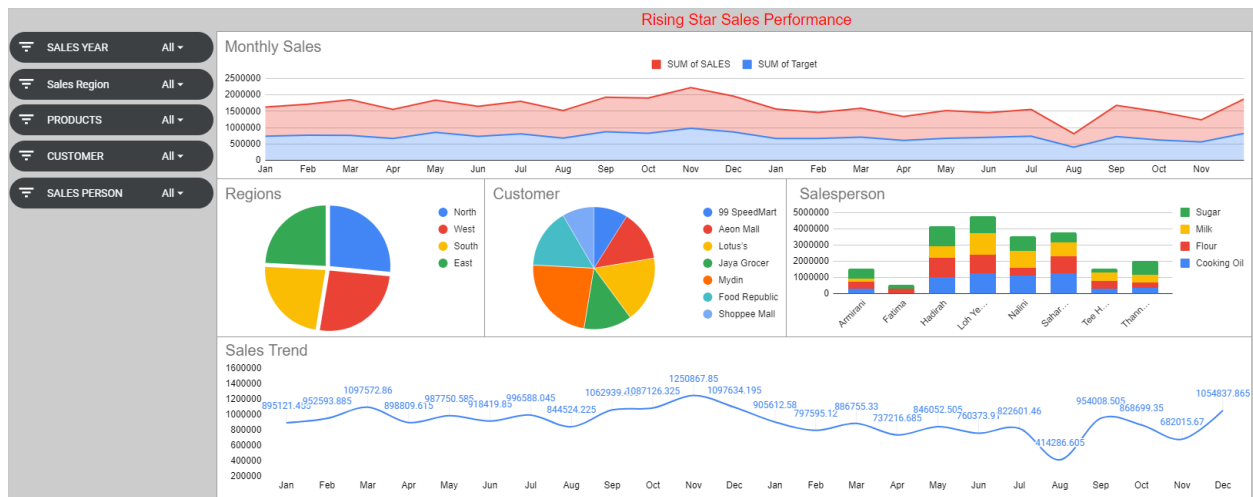
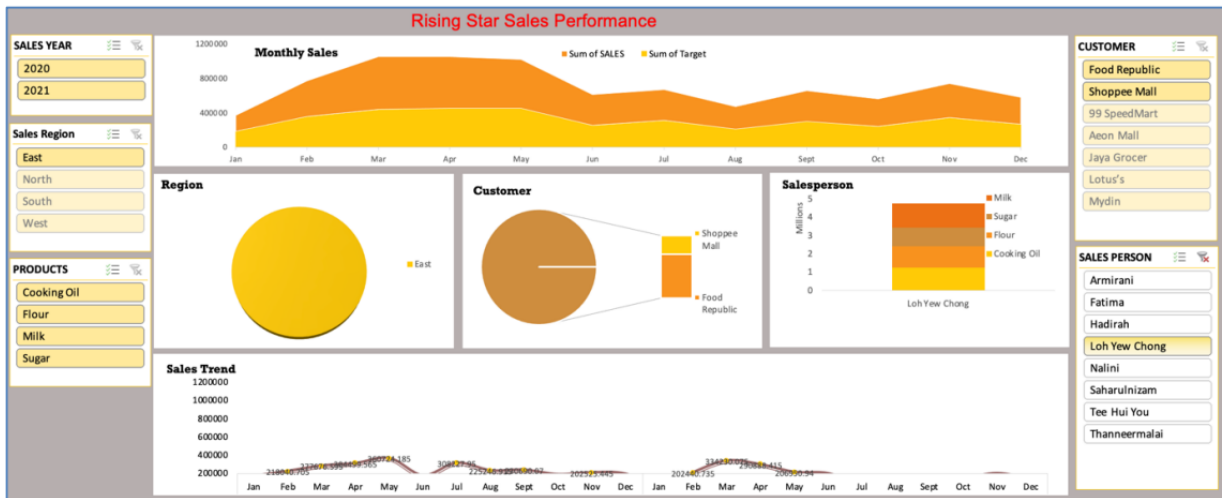


Figure 20 : Dashboard

4.2 Question 2

4. Each of these menus must be related to the others. If we select Loh Yew Chong as the salesperson, all charts will display information about Loh Yew Chong (refer Figure 8).



Answer for Question 2

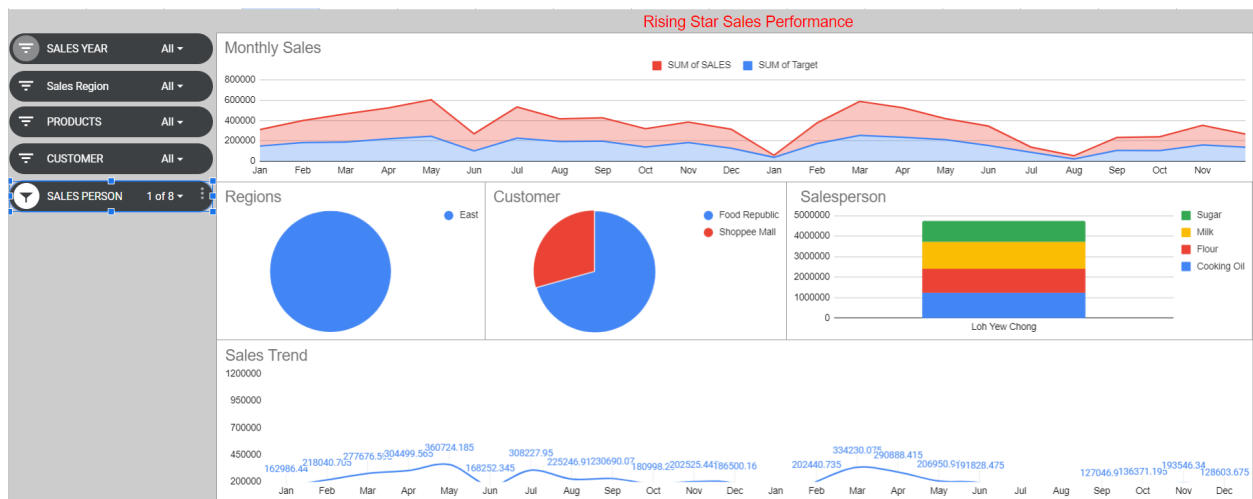
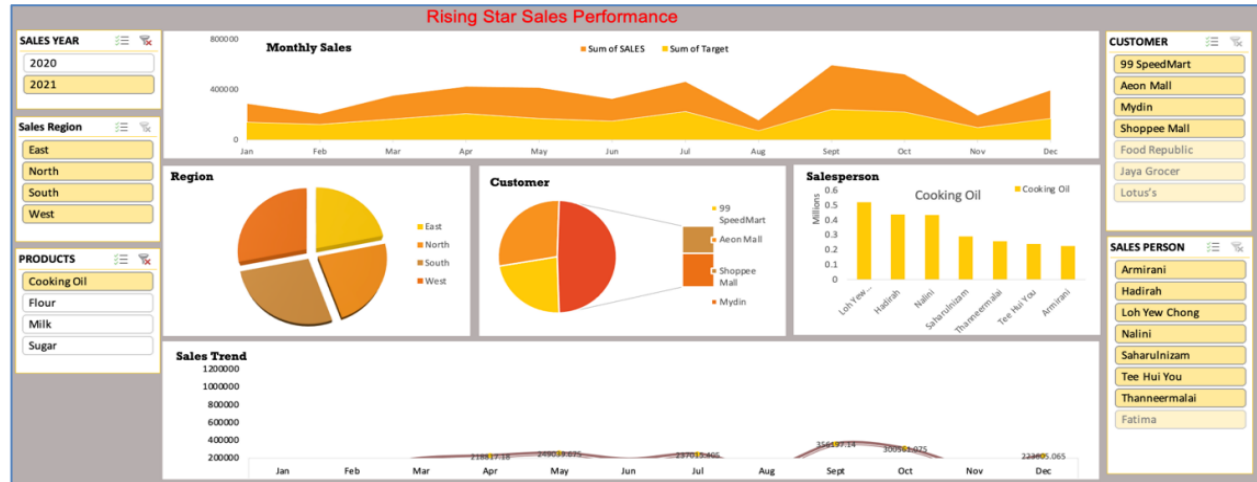


Figure 21 : Display after selecting Loh Yew Chong as the salesperson.

4.3 Question 3

5. If you select cooking oil as a product and the sales year is 2021, the display looks like Figure 9.



Answer for Question 3

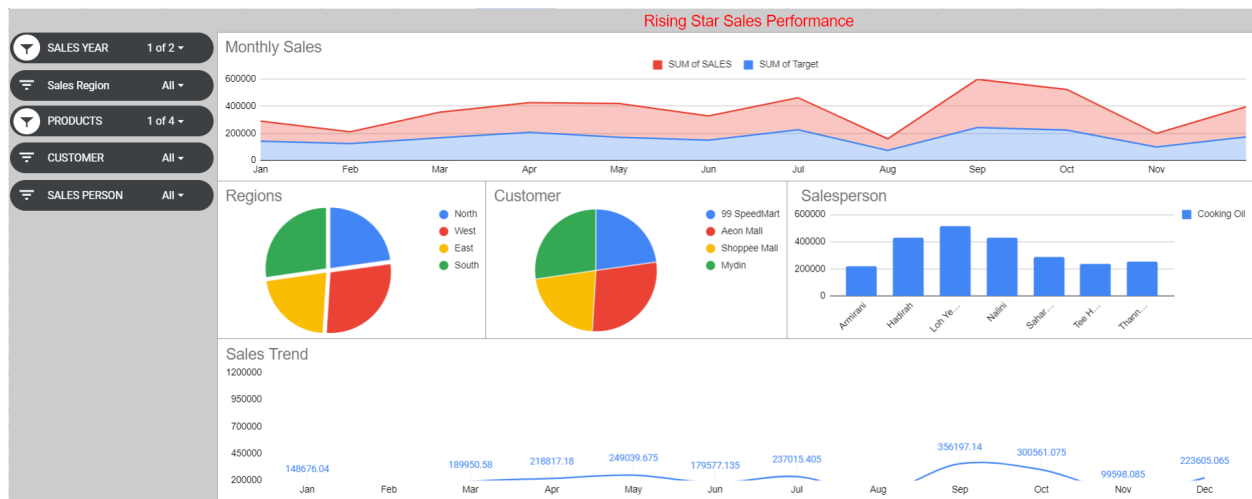


Figure 22 : Display if select cooking oil as a product and the sales year is 2021.

5.0 RECOMMENDATION

- a. Must explore and identify attribute data type
- b. Transform data follows a suitable data type.
- c. Pivot table
 - a. Transform data to pivot table in order to make a stackable chart/track multiple trend lines, especially when need to compare several data points.
 - b. It is often easier to see the comparison on a stacked bar chart versus a combined chart which are more clear when data are on top of each other.