

# **Apon Bazaar Data Analyst Assessment**

**Prepared By: Risad Raihan Malik**

**Tools Used: Excel, BigQuery, MySQL, Looker Studio**

Question:1:

You were given the 3 sales reports of 3 months of Apon Bazaar. You are asked to prepare a dynamic report to identify the angel customers for Apon for the 4th month who will get the special privilege offer from Apon.

Conditions to prepare the sheet are given below:

The sheet should be dynamic so that your manager can change the criteria of developing the angel list. Criteria are:

- a. Customers can purchase any of the 3 months or any 2 months with the 3 months.
- b. The purchase amount can be changed based on the or your manager ranging from 1,000 to 20,000+
- c. Your manager needs to see for each criterion set, the gender ratio of the angel customers. (You can find the gender data in other sheet named "Apon Main Database Gender Data")
- d. Divide the angel list shop wise considering the master database given separately named "Apon Main Database Gender Data".
- e. Your manager would be highly impressed if you can show any of your results with a graphical presentation.

Question:2:

You were given a previous 6 months sales report month by month of Apon Bazaar. You are asked to prepare

Big block wise dynamic dashboard on margin and sales contribution to see the month by month improvement / degrowth.

A dashboard on Big block wise shop wise margin and sales contribution.

Identify and suggest sub category wise 3 products with high margin to focus on sales.

Identify and suggest sub category wise 3 products with lower margin to remove from sales.

**\*\*Note:** Item code is the unique identifier of a product among all the sheets. All the additional information related to products is in the product info sheet.

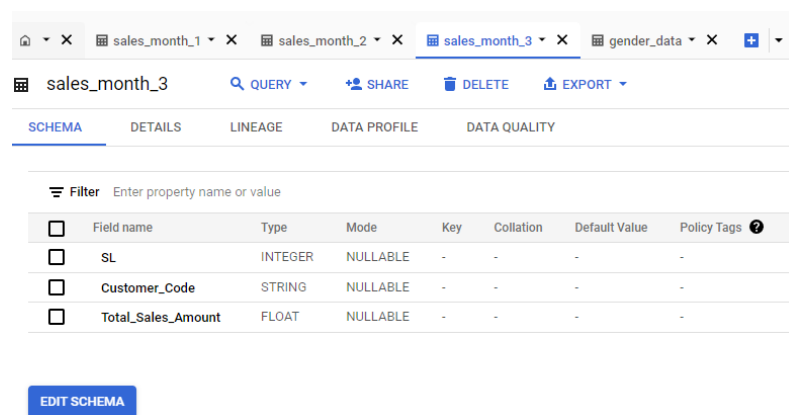
## Assessment 1 :

I was given three months of sale data of Apon Bazar and Apon Main DB gender data. My task was to make a dynamic dashboard/report to see Angel customers where the user can change the criteria of Angel customers.

To achieve this task, I had to do some data preprocessing and sorting and then finally visualization. I used my current company's GCP platform which is personally obtained by me. I will briefly describe here how I achieved my goal and source to the Dashboard.

### Step 1: Upload data BQ:

At first, I uploaded all four datasets that I have been provided. Here is the snippet.



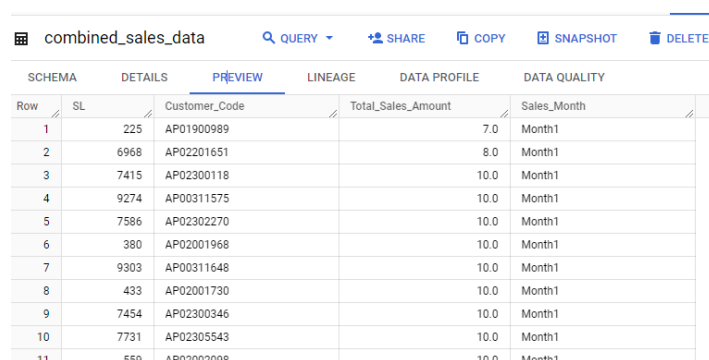
Field name	Type	Mode	Key	Collation	Default Value	Policy Tags
SL	INTEGER	NULLABLE	-	-	-	-
Customer_Code	STRING	NULLABLE	-	-	-	-
Total_Sales_Amount	FLOAT	NULLABLE	-	-	-	-

Figure 1 Successfully uploaded datasets.

### Step 2: Union three months data into one table:

In this step I unionid all three months' sale data into one table so that it gets easier to manipulate them. And assigned them as Month 1, Month 2, and Month 3

```
1 CREATE TABLE `tenms-userdb.temp_rrm.combined_sales_data` AS
2 SELECT
3   SL,
4   Customer_Code,
5   Total_Sales_Amount,
6   'Month1' AS Sales_Month
7 FROM `tenms-userdb.temp_rrm.sales_month_1`
8 UNION ALL
9 SELECT
10  SL,
11  Customer_Code,
12  Total_Sales_Amount,
13  'Month2'
14 FROM `tenms-userdb.temp_rrm.sales_month_2`
15 UNION ALL
16 SELECT
17  SL,
18  Customer_Code,
19  Total_Sales_Amount,
20  'Month3'
21 FROM `tenms-userdb.temp_rrm.sales_month_3`;
22
```



Row	SL	Customer_Code	Total_Sales_Amount	Sales_Month
1	225	AP01900989	7.0	Month1
2	6968	AP02201651	8.0	Month1
3	7415	AP02300118	10.0	Month1
4	9274	AP00311575	10.0	Month1
5	7586	AP02302270	10.0	Month1
6	380	AP02001968	10.0	Month1
7	9303	AP00311648	10.0	Month1
8	433	AP02001730	10.0	Month1
9	7454	AP02300346	10.0	Month1
10	7731	AP02305543	10.0	Month1
11	559	AP02002098	10.0	Month1

### Step 3: enriched Sale Data:

This is a crucial part, here I joined my combined 3months sale data with the gender data and joined them on customer code and apoc code.

Untitled

RUN

MORE

SAVE

DOWN

```
1 CREATE TABLE 'tenms-userdb.temp_rrm.enriched_sales_data' AS
2 SELECT
3   csd.Customer_Code,
4   csd.Total_Sales_Amount,
5   csd.Sales_Month,
6   agd.string_field_1 as Gender,
7   agd.string_field_2 as Shop_name
8 FROM 'tenms-userdb.temp_rrm.combined_sales_data' csd
9 JOIN 'tenms-userdb.temp_rrm.gender_data' agd
10 ON csd.Customer_Code = agd.string_field_0;
11
```

enriched\_sale...

QUERY

SHARE

COPY

SNAPSHOT

DELETE

EXPORT

SCHEMA

DETAILS

PREVIEW

LINEAGE

DATA PROFILE

DATA QUALITY

ow	Customer_Code	Total_Sales_Amount	Sales_Month	Gender	Shop_name
1	AP02104120	100.0	Month3	Male	Factory 1
2	AP02103724	100.0	Month3	Male	Factory 1
3	AP02103485	240.0	Month3	Male	Factory 1
4	AP02103964	272.0	Month3	Male	Factory 1
5	AP02103463	512.0	Month3	Male	Factory 1
6	AP02103679	960.0	Month3	Male	Factory 1
7	AP02103742	1728.0	Month3	Male	Factory 1
8	AP02103510	641.0	Month3	Male	Factory 1
9	AP02103681	897.0	Month3	Male	Factory 1
10	AP02103447	1410.0	Month3	Male	Factory 1
11	AP02103995	738.0	Month3	Male	Factory 1
12	AP02103788	145.0	Month3	Male	Factory 1
13	AP02103666	193.0	Month3	Male	Factory 1
14	AP02103735	225.0	Month3	Male	Factory 1
15	AP02103849	466.0	Month3	Male	Factory 1
16	AP02103661	933.0	Month3	Male	Factory 1
17	AP02104097	1357.0	Month3	Male	Factory 1
18	AP02103709	2330.0	Month3	Male	Factory 1
19	AP02103870	711.0	Month3	Male	Factory 1
20	AP02103574	1550.0	Month3	Male	Factory 1

Query results

JOB INFORMATION

RESULTS

EXECUTION DETAILS

EXECUTION

1

This statement created a new table named enriched\_sales\_data.

### Step 4: Making the Final report:

This is the final report that has been connected to the dashboard. Important thing to notice here is for **Months\_Purchased >= 2**

**AND Total\_Amount BETWEEN 1000 AND 20000**; which indicates the question where I was said that Angel customer has to buy more than 1 month and their total amount has to more than 1000 and less than 20000.

Untitled

RUN

SAVE

DOWN

SHARE

SCHEDULE

MORE

```
1 WITH customer_purchase_summary AS (
2   SELECT
3     Customer_Code,
4     COUNT(DISTINCT Sales_Month) AS Months_Purchased,
5     SUM(Total_Sales_Amount) AS Total_Amount,
6     Gender,
7     Shop_name
8 FROM 'tenms-userdb.temp_rrm.enriched_sales_data'
9 GROUP BY Customer_Code, Gender, Shop_name
10 )
11 SELECT
12   *
13 FROM customer_purchase_summary
14 WHERE
15   Months_Purchased >= 2
16   AND Total_Amount BETWEEN 1000 AND 20000; -- This range can be adjusted or made dynamic , i will do that later in looker.
17
```

Processing location: US

Query results

JOB INFORMATION

RESULTS

CHART

JSON

EXECUTION DETAILS

EXECUTION GRAPH

Row	Customer_Code	Months_Purchased	Total_Amount	Gender	Shop_name
1	AP02103463	2	1406.0	Male	Factory 1
2	AP02103906	2	1008.0	Male	Factory 1
3	AP02104810	2	3191.0	Male	Factory 1
4	AP02111238	2	3117.0	Male	Factory 1
5	AP02111239	2	8627.0	Male	Factory 1
6	AP02111048	2	1052.0	Female	Factory 1
7	AP02111234	2	8834.0	Female	Factory 1
8	AP02111245	2	1842.0	Female	Factory 1

## Step 5: Dashboard:

I then connected the final report with Looker Studio to make the interactive dynamic dashboard where viewer can see Angel customer and most importantly they can change the criteria as mentioned in the question.

Dashboard Link : <https://lookerstudio.google.com/reporting/78ff8ce2-cb4d-49e3-a189-68d338fdd369>



#### Insights:

1. Total\_amount slider can be changed to fix any amount of price from 1000 to 20000
2. Months purchased dropdown can be used to select last two months
3. Pie chart here shows that Men are most like to be the potential customers who buy the products.
4. The bar chart shows that Factory 5 had the most customers.
5. Total amount score card shows the total sale amount, customer\_code score card shows total customer and shop\_name score card shows total shops or factories.
6. All the data in the dashboard are interactive means if the user clicks any specific data, the dashboard will automatically sort to show that type of data, what I mean is, from the pie chart if the user clicks on 'Men' then the dashboard will show data that has been related to that value of the Gender

N.B: That was the assessment answer for Question 1. Question 2 answer starts from next page.

## Assessment 2 :

Second part of the assessment required more visualization and calculation knowledge.

### Step 1: Data cleaning, preprocessing and uploading on BQ

In this section I had clean all the data and checked all the data from 6 months sale, because the name was more than 600k. I uploaded them on BQ, unioned into one table and assigned there months.



```
1 CREATE TABLE `tenms-userdb.temp_rrm.combined_sales_data2` AS
2 SELECT 'August' AS Sales_Month, * FROM `tenms-userdb.temp_rrm.August_23`
3 UNION ALL
4 SELECT 'September' AS Sales_Month, Outlet, Challan, Customer_Code, Customer_Name, PARSE_DATE('%e-%b', Date) AS Date, Item_Code, Item_Name, Total_Quantity, MRP, Sales_Price, Total_Sales_Amount,
5 Total_Discount, NewCOGS, Total_COGS FROM `tenms-userdb.temp_rrm.september`
6 UNION ALL
7 SELECT 'October' AS Sales_Month, Outlet, Challan, Customer_Code, Customer_Name, PARSE_DATE('%e-%b', Date) AS Date, Item_Code, Item_Name, Total_Quantity, MRP, Sales_Price, Total_Sales_Amount,
8 Total_Discount, NewCOGS, Total_COGS FROM `tenms-userdb.temp_rrm.october`
9 UNION ALL
10 SELECT 'November' AS Sales_Month, Outlet, Challan, Customer_Code, Customer_Name, PARSE_DATE('%e-%b', Date) AS Date, Item_Code, Item_Name, Total_Quantity, MRP, Sales_Price, Total_Sales_Amount,
11 Total_Discount, NewCOGS, Total_COGS FROM `tenms-userdb.temp_rrm.November`
12 UNION ALL
13 SELECT 'December' AS Sales_Month, Outlet, Challan, Customer_Code, Customer_Name, PARSE_DATE('%e-%b', Date) AS Date, Item_Code, Item_Name, Total_Quantity, MRP, Sales_Price, Total_Sales_Amount,
14 Total_Discount, NewCOGS, Total_COGS FROM `tenms-userdb.temp_rrm.december`
15 UNION ALL
16 SELECT 'January' AS Sales_Month, Outlet, Challan, Customer_Code, Customer_Name, PARSE_DATE('%e-%b', Date) AS Date, Item_Code, Item_Name, Total_Quantity, MRP, Sales_Price, Total_Sales_Amount,
17 Total_Discount, NewCOGS, Total_COGS FROM `tenms-userdb.temp_rrm.january`;
```

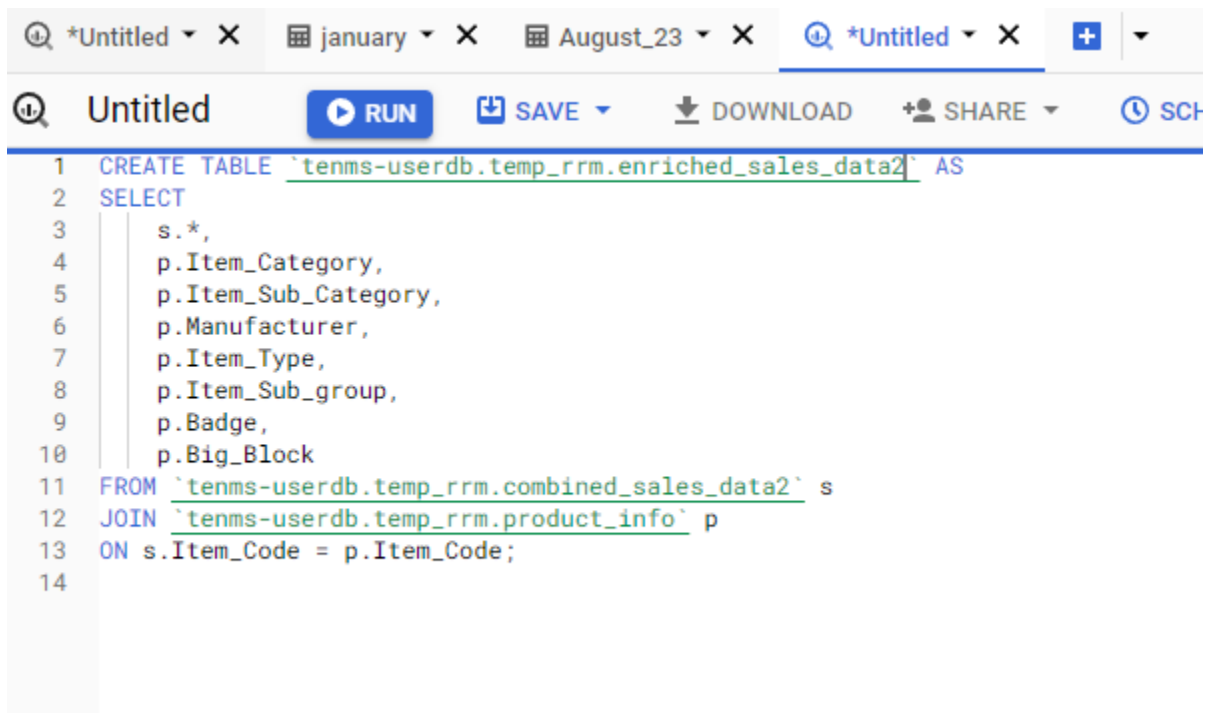
Query results

JOB INFORMATION RESULTS EXECUTION DETAILS EXECUTION GRAPH

This statement created a new table named combined\_sales\_data2.

### Step 2: Enriched Sale data with product info:

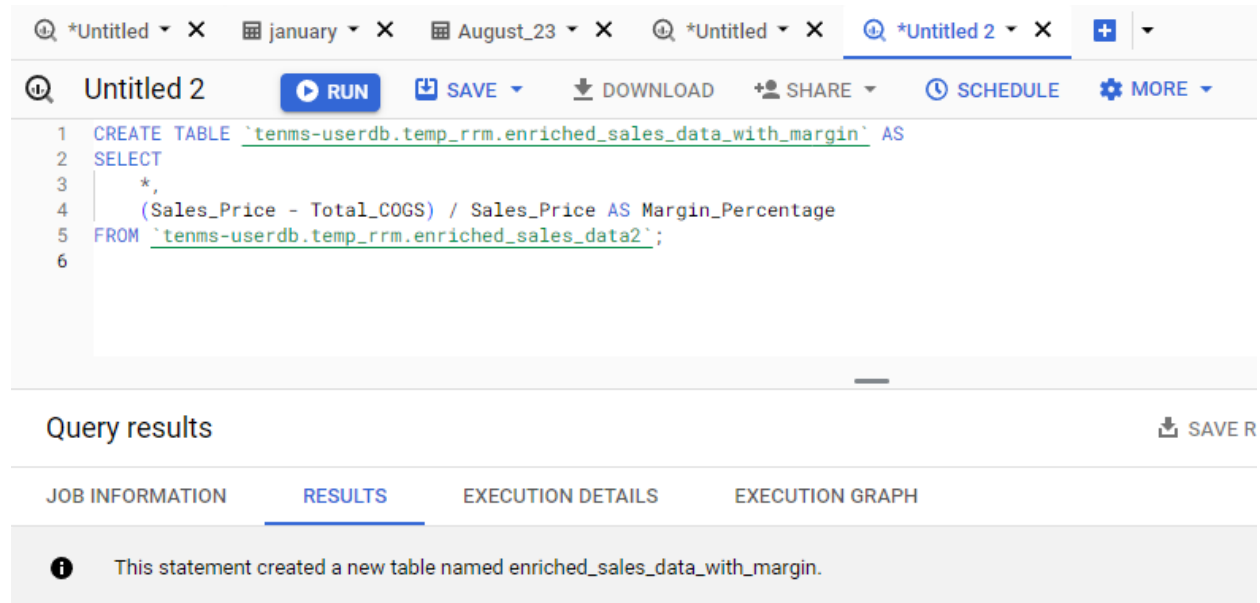
In this part I made a table called enriched sale data where I joined the combined table I created before with product info table on their Item code



```
1 CREATE TABLE `tenms-userdb.temp_rrm.enriched_sales_data2` AS
2 SELECT
3     s.*,
4     p.Item_Category,
5     p.Item_Sub_Category,
6     p.Manufacturer,
7     p.Item_Type,
8     p.Item_Sub_group,
9     p.Badge,
10    p.Big_Block
11 FROM `tenms-userdb.temp_rrm.combined_sales_data2` s
12 JOIN `tenms-userdb.temp_rrm.product_info` p
13 ON s.Item_Code = p.Item_Code;
```

### Step 3: Margin calculation:

For calculating the margin percentage, I used a calculation and then stored everything in a table for final table creation,



The screenshot shows a SQL query editor interface. The query is as follows:

```
1 CREATE TABLE `tenms-userdb.temp_rrm.enriched_sales_data_with_margin` AS
2 SELECT
3     *,
4     (Sales_Price - Total_COGS) / Sales_Price AS Margin_Percentage
5 FROM `tenms-userdb.temp_rrm.enriched_sales_data2`;
6
```

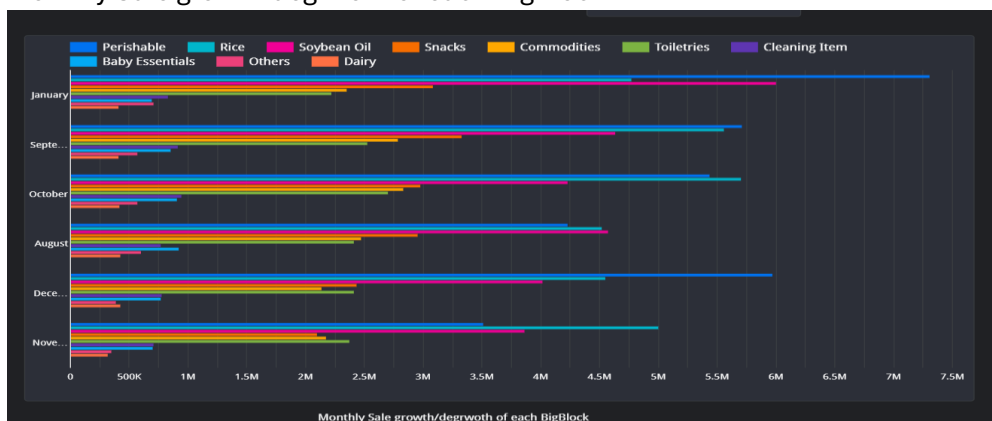
Below the query editor, there is a section titled "Query results" with a "SAVE R" button. The "RESULTS" tab is selected, showing a message: "This statement created a new table named enriched\_sales\_data\_with\_margin."

### Step 4: Dashboard:

Link: [https://lookerstudio.google.com/reporting/78ff8ce2-cb4d-49e3-a189-68d338fdd369/page/p\\_izu8no7ifd/edit](https://lookerstudio.google.com/reporting/78ff8ce2-cb4d-49e3-a189-68d338fdd369/page/p_izu8no7ifd/edit) , 2<sup>nd</sup> page

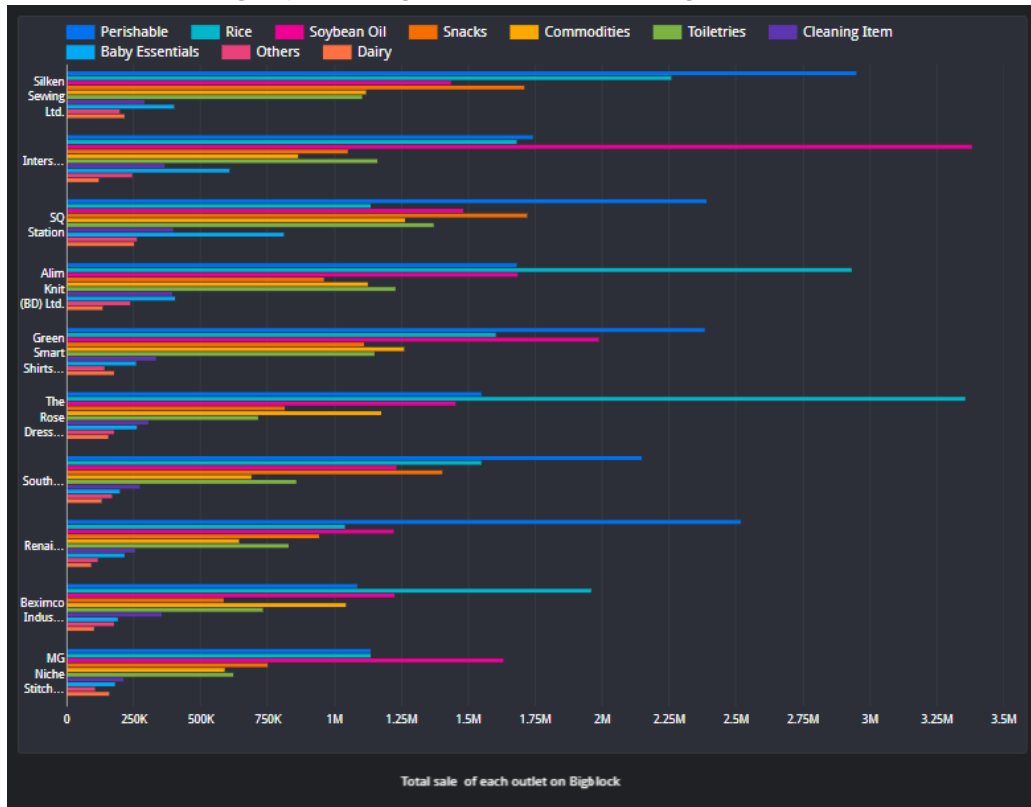
This is the main section of the report where I landed on some decisions:

1. Monthly Sale growth/degrwoth of each BigBlock:



According to the visualization we can see that in January Perishable Bigblock has the highest number of sales and in November Dairy BigBlock has the lowest amount of sale. This report is fully interactive.

2. Total sale and Margin\_percentage of each outlet on Bigblock:



	Big_Block	Outlet	Total_Sales_Amount ▾	Margin_Percentage
1.	Soybean Oil	Interstoff Apparels Ltd.	3,381,426	-710.79
2.	Rice	The Rose Dresses Limited	3,355,909.56	-58,804.84
3.	Perishable	Dekko Designs Limited	3,330,576.5	-2,024.46
4.	Perishable	Silken Sewing Ltd.	2,951,993.22	-2,670.71
5.	Rice	Alim Knit (BD) Ltd.	2,932,933.44	-46,614.31
6.	Perishable	Renaissance Apparels Ltd.	2,519,124.48	-4,543.06
7.	Perishable	SQ Station	2,391,653.22	-13,628.43
8.	Perishable	Green Smart Shirts Ltd.	2,383,017.28	-6,583.11
9.	Rice	Silken Sewing Ltd.	2,259,155.62	-35,432.94
10.	Perishable	Southern Garments Ltd.	2,150,922	-134.58
11.	Soybean Oil	Green Smart Shirts Ltd.	1,987,348.5	-147.6

1 - 100 / 244 < >

Total sale and Margin\_percentage of each outlet on Bigblock

From this report we can see Interstoff apparels has highest amount off sale from all the outlets . This table can be switched to ascending/descending order for both total sales amount and margin percentage



3. Top and Least 3 items by margin percentage:

	Item_Sub_Cat...	Item_Name	Margin_Percentage ▾
1.	Fruits	Apple Gala	1,887.95
2.	Fruits	Orange Komla	1,436.99
3.	Fruits	Dates Tunashia ...	1,230.69
1 - 100 / 1362 < >			
Top 3 item and item sub category by margin_percentage			
	Item_Sub_Cat...	Item_Name	Margin_Percentage ▴
1.	Rice	Half Bowel Rice ...	-195,665.07
2.	Rice	BR-28 Rice Pre...	-169,086.17
3.	Egg	Local Egg PC	-59,566.27
1 - 100 / 1362 < >			
Least 3 item and item sub category by margin_percentage			

I found this very important for business perspective as we can see fruits sub category and inside fruits sub category, Apple, Orange and Dates Tunisia has the highest margin percentage

And in Rice and Egg Half Bowel Rice, BR-28 Rice and Local Egg FC has the lowest Margin percentage

All the infographics in this dashboard are interactive, and since my report has few more columns, there are a lot other analytical scopes are there from the data that has been provided to me, but I only performed those analysis that has been asked in the assessment questions.