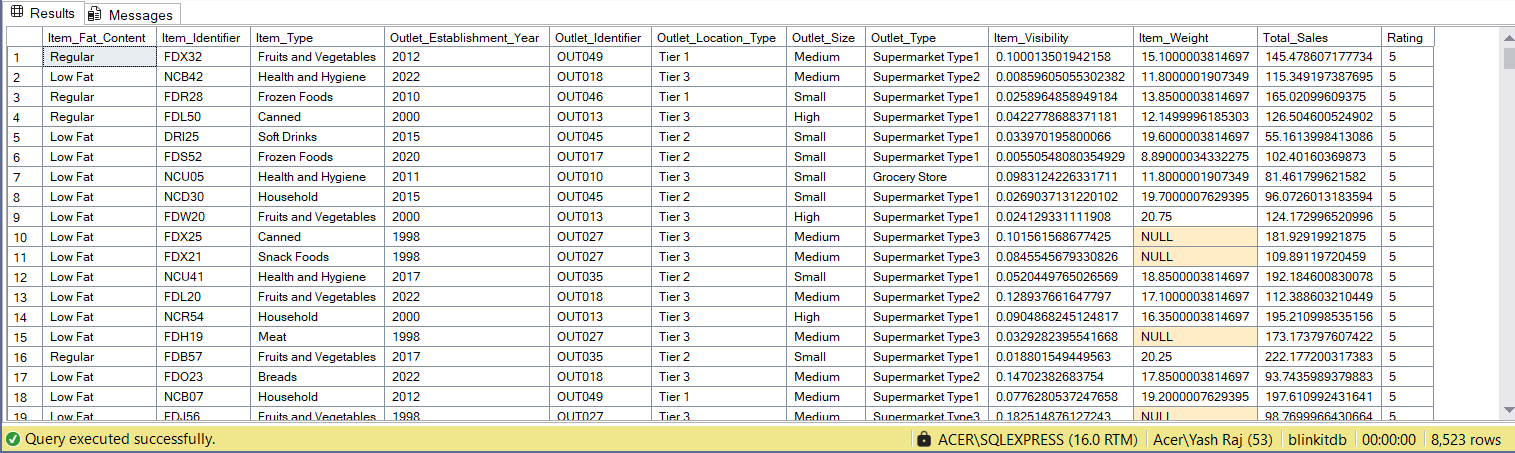
**Local Shop Store Analysis**

Before starting my project, Let’s look at the data by running simple DQL query:

select \* from localShop\_data



* **Data Cleaning**

Standardizing the Item\_Fat\_Content\*\* field helps maintain consistency and reliability in the data. Variations of the same category (such as LF, low fat, and Low Fat) can lead to inaccuracies in reporting, grouping, and filtering. By cleaning and unifying these values, we enhance data quality, simplify analysis, and ensure more accurate insights across the dataset.

update localShop\_data set Item\_Fat\_Content =

case

when Item\_Fat\_Content in ('low fat','LF') then 'Low Fat'

when Item\_Fat\_Content = 'reg' then 'Regular'

else Item\_Fat\_Content

end

After executing this query check the data has been cleaned or not using below query

select distinct (Item\_Fat\_Content) from localShop\_data

A screenshot of a computer

AI-generated content may be incorrect.

1. KPI’s

1.Total Sales

select CAST(round(sum(Total\_Sales/ 1000000),2) AS VARCHAR(10)) + ' M'

as "Total Sales"

from localShop\_data

A screenshot of a computer screen

AI-generated content may be incorrect.

2. Average Sales

select cast(AVG(total\_Sales) AS int) as "Average Sales"

from localShop\_data

A screenshot of a computer

AI-generated content may be incorrect.

3. Number of Items

select cast(round(COUNT(\*)/1000,2) AS varchar(10)) + ' K'

as "Number of Items Sold"

from localShop\_data

A screenshot of a computer

AI-generated content may be incorrect.

4. Average Rating

select round(AVG(rating),1) as "Average Rating"

from localShop\_data

A screenshot of a computer screen

AI-generated content may be incorrect.

Merging all KPI's in one go

select item\_fat\_content,

CAST(round(sum(Total\_Sales),2) AS VARCHAR(10)) as "Total Sales",

cast(AVG(total\_Sales) AS int) as "Average Sales" ,

cast(round(COUNT(\*),2) AS varchar(10)) as "Number of Items Sold",

round(AVG(rating),1) as "Average Rating"

from localShop\_data

group by item\_fat\_content

order by [Total Sales] desc

A screenshot of a graph

AI-generated content may be incorrect.

Check KPI's based on Item Type Top 5

select top 5 Item\_Type,

CAST(round(sum(Total\_Sales),2) AS VARCHAR(10)) as "Total Sales",

cast(AVG(total\_Sales) AS int) as "Average Sales" ,

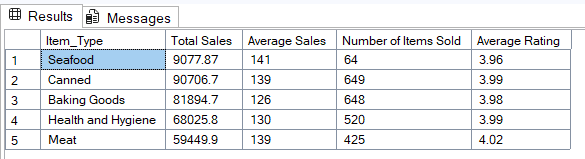
cast(round(COUNT(\*),2) AS varchar(10)) as "Number of Items Sold",

cast(AVG(rating) AS decimal(10,2)) as "Average Rating"

from localShop\_data

group by Item\_Type

order by [Total Sales] desc



1. Granular Requirements

Total Sales by Fat Content

select Item\_Fat\_Content,

cast(sum(total\_sales) as decimal(10,2)) as "Total Sales"

from localShop\_data

group by Item\_Fat\_Content

A screenshot of a computer

AI-generated content may be incorrect.

Total Sales by Item Type:

select Item\_Type,

cast(sum(total\_sales) as decimal(10,2)) as "Total Sales"

from localShop\_data

group by Item\_Type

order by [Total Sales] desc

A screenshot of a table

AI-generated content may be incorrect.

Fat Content by Outlet for Total Sales:

SELECT Outlet\_Location\_Type,

-- If there is no sales data for a particular category

-- (e.g., no "Low Fat" sales in an outlet), we get NULL

-- ISNULL(..., 0) replaces those NULLs with 0 for better readability.

ISNULL([Low Fat], 0) AS Low\_Fat,

ISNULL([Regular], 0) AS Regular

FROM

(

-- This part groups the data by Outlet\_Location\_Type and Item\_Fat\_Content and It calculates the total sales per combination

SELECT Outlet\_Location\_Type, Item\_Fat\_Content,

CAST(SUM(Total\_Sales) AS DECIMAL(10,2)) AS Total\_Sales

FROM localShop\_data

GROUP BY Outlet\_Location\_Type, Item\_Fat\_Content

) AS SourceTable

PIVOT

(

SUM(Total\_Sales) -- Fill total sales values in low fat and regular column

FOR Item\_Fat\_Content IN ([Low Fat], [Regular]) --Low Fat and Regular becomes column headers

) AS PivotTable

ORDER BY Outlet\_Location\_Type;

A screenshot of a computer

AI-generated content may be incorrect.

Total Sales by Outlet Establishment:

select Outlet\_Establishment\_Year,

CAST(SUM(total\_Sales) AS decimal(10,2)) as "Total Sales",

cast(AVG(total\_Sales) AS int) as "Average Sales" ,

cast(round(COUNT(\*),2) AS varchar(10)) as "Number of Items Sold",

cast(AVG(rating) AS decimal(10,2)) as "Average Rating"

from localShop\_data

group by Outlet\_Establishment\_Year

order by Outlet\_Establishment\_Year

A table with numbers and a number

AI-generated content may be incorrect.

D. Chart’s Requirements

Percentage of Sales by Outlet Size:

select Outlet\_Size,

CAST(SUM(total\_Sales) AS decimal(10,2)) as Total\_Sales,

cast((SUM(Total\_Sales)\*100/sum(sum(Total\_Sales)) over())

AS decimal(10,2)) as "Sales Percentage"

from localShop\_data

group by Outlet\_Size

order by Total\_Sales desc

A screenshot of a computer

AI-generated content may be incorrect.

Sales by Outlet Location:

select Outlet\_Location\_Type, CAST(SUM(total\_Sales) as decimal(10,2)) as "Total Sales"

from localShop\_data

group by Outlet\_Location\_Type

order by [Total Sales]

A screenshot of a computer

AI-generated content may be incorrect.

All Metrics by Outlet Type

select Outlet\_Type,

CAST(SUM(total\_Sales) AS decimal(10,2)) as "Total Sales",

cast(AVG(total\_Sales) AS int) as "Average Sales" ,

cast(round(COUNT(\*),2) AS varchar(10)) as "Number of Items Sold",

cast(AVG(rating) AS decimal(10,2)) as "Average Rating"

from localShop\_data

group by Outlet\_Type

order by Outlet\_Type

A screenshot of a data table

AI-generated content may be incorrect.

Top 10 Best-Selling Items

select top 10 Item\_Type, cast(sum(Total\_Sales) as decimal(10,2)) as Total\_Sales

from localShop\_data

group by Item\_Type

order by Total\_Sales desc -- based on total sales

A screenshot of a computer

AI-generated content may be incorrect.

Top 20 product based on customer rating

select top 20 Item\_Type, Rating

from localShop\_data

order by Rating desc

A table with text on it

AI-generated content may be incorrect.

Least-Selling Items

select top 5 Item\_Identifier, cast(sum(Total\_Sales) as decimal(10,2)) as Poor\_Sales

from localShop\_data

group by Item\_Identifier

order by Poor\_Sales

A screenshot of a computer screen

AI-generated content may be incorrect.

Monthly/Yearly Sales Trend (based on Outlet\_Establishment\_Year)

select

Outlet\_Establishment\_Year,

cast(sum(Total\_Sales) as decimal(10,2)) as Total\_Sales,

RANK() over(order by sum(Total\_Sales) desc) as Sales\_Rank

from localShop\_data

group by Outlet\_Establishment\_Year

order by Sales\_Rank;

A screenshot of a computer

AI-generated content may be incorrect.

Average Sales by Item Visibility Range

select

case

when Item\_Visibility < 0.05 then 'Low Visibility'

when Item\_Visibility between 0.05 and 0.15 then 'Medium Visibility'

else 'High Visibility'

end as Visibility\_Range,

cast(avg(Total\_Sales) as decimal(10,2)) as Average\_Sales

from localShop\_data

group by

case

when Item\_Visibility < 0.05 then 'Low Visibility'

when Item\_Visibility between 0.05 and 0.15 then 'Medium Visibility'

else 'High Visibility'

end

order by Average\_Sales desc;

A screenshot of a computer screen

AI-generated content may be incorrect.

Correlation between Item Weight & Sales

select

case

when Item\_Weight < 5 then 'Less than 5Kg'

when Item\_Weight between 5 and 10 then 'Weight Between 5Kg and 10Kg'

when Item\_Weight between 10 and 15 then 'Weight Between 10Kg and 15Kg'

else 'Above 15Kg'

end as Item\_weight,

CAST(AVG(total\_Sales) as decimal(10,2)) as Average\_Sales

from localShop\_data

group by case

when Item\_Weight < 5 then 'Less than 5Kg'

when Item\_Weight between 5 and 10 then 'Weight Between 5Kg and 10Kg'

when Item\_Weight between 10 and 15 then 'Weight Between 10Kg and 15Kg'

else 'Above 15Kg'

end

order by Item\_Weight

A screenshot of a message

AI-generated content may be incorrect.

Contribution of Each Outlet to Total Sales (in %)

select Outlet\_Type,

cast(round((SUM(Total\_Sales)\*100/sum(sum(Total\_Sales)) over()),2)

AS varchar(10))+'%' as "Sales Percentage",

RANK() over(order by sum(Total\_Sales) desc) as Sales\_Rank

from localShop\_data

group by Outlet\_Type

order by Sales\_Rank;

A screenshot of a data

AI-generated content may be incorrect.

Compare Sales by Outlet Type + Location Type

select

Outlet\_Type,

ISNULL([Tier 1],0) as 'Tier 1',

ISNULL([Tier 2],0) as 'Tier 2',

ISNULL([Tier 3],0) as 'Tier 3'

from

(

select Outlet\_Type, Outlet\_Location\_Type

, CAST(sum(total\_Sales) as decimal(10,2)) as Total\_Sales

from localShop\_data

group by Outlet\_Type, Outlet\_Location\_Type

) as SourceTable

pivot(

sum(total\_Sales)

for Outlet\_location\_type in ([Tier 1],[Tier 2],[Tier 3])

) as PivotTable

order by Outlet\_Type;

A screenshot of a computer

AI-generated content may be incorrect.

Sales vs. Ratings Analysis

SELECT

CASE

WHEN Rating BETWEEN 0 AND 2.5 THEN 'Low Rating'

WHEN Rating > 2.5 AND Rating <= 4 THEN 'Good Rating'

WHEN Rating > 4 AND Rating <= 5 THEN 'Excellent'

END AS Rating\_Category,

CAST(round(SUM(Total\_Sales)/100000,2) AS varchar(10))+ 'M' AS Total\_Sales

FROM localShop\_data

GROUP BY

CASE

WHEN Rating BETWEEN 0 AND 2.5 THEN 'Low Rating'

WHEN Rating > 2.5 AND Rating <= 4 THEN 'Good Rating'

WHEN Rating > 4 AND Rating <= 5 THEN 'Excellent'

END

ORDER BY Total\_Sales DESC;

A screenshot of a computer screen

AI-generated content may be incorrect.

Sales Distribution by Item Type (Pie Chart style data)

select Item\_Type,

CAST(round((SUM(Total\_Sales)\*100/sum(SUM(total\_sales)) over()),2) AS varchar(10))+ '%' AS "Sales Percentage"

from localShop\_data

group by Item\_Type

A table with numbers and percentages

AI-generated content may be incorrect.

Customer Basket Analysis (Simulated)

SELECT

a.Outlet\_Identifier,

a.Item\_Identifier AS Item1,

b.Item\_Identifier AS Item2,

COUNT(\*) AS Times\_Sold\_Together

FROM

localShop\_data a

INNER JOIN localShop\_data b

ON a.Outlet\_Identifier = b.Outlet\_Identifier

AND a.Item\_Identifier < b.Item\_Identifier -- Only unique unordered pairs

GROUP BY

a.Outlet\_Identifier,

a.Item\_Identifier,

b.Item\_Identifier

HAVING

COUNT(\*) > 1 -- Filter for pairs that appear together more than once

ORDER BY

a.Outlet\_Identifier,

Times\_Sold\_Together DESC;

A screenshot of a computer

AI-generated content may be incorrect.

No product found which sold together.

Fat Content Impact on Ratings

Do “Low Fat” vs “Regular” items have better customer satisfaction?

SELECT

Item\_Fat\_Content,

COUNT(\*) AS Num\_Items,

round(AVG(Rating \* 1.0),2) AS Avg\_Rating

FROM localShop\_data

WHERE Item\_Fat\_Content IN ('Low Fat', 'Regular')

GROUP BY Item\_Fat\_Content;

-- No both are of same rating

A screenshot of a computer

AI-generated content may be incorrect.

Pareto Analysis (80/20 Rule)

Find whether ~20% of items contribute to ~80% of sales.

WITH RankedSales AS (

SELECT

Item\_Identifier,

SUM(Total\_Sales) AS Total\_Sales

FROM localShop\_data

GROUP BY Item\_Identifier

),

SalesWithRank AS (

SELECT

Item\_Identifier,

Total\_Sales,

SUM(Total\_Sales) OVER (ORDER BY Total\_Sales DESC ROWS UNBOUNDED PRECEDING) AS Cumulative\_Sales,

SUM(Total\_Sales) OVER () AS Total\_Sales\_All

FROM RankedSales

),

SalesWithCumulativePercent AS (

SELECT

Item\_Identifier,

Total\_Sales,

Cumulative\_Sales,

Total\_Sales\_All,

100.0 \* Cumulative\_Sales / Total\_Sales\_All AS Cumulative\_Sales\_Percent

FROM SalesWithRank

)

SELECT

Item\_Identifier,

round(Total\_Sales,2),

ROUND(Cumulative\_Sales\_Percent, 2) AS Cumulative\_Sales\_Percent,

CASE

WHEN Cumulative\_Sales\_Percent <= 80 THEN 'Top 80% Sales'

ELSE '20% Sales'

END AS Pareto\_Category

FROM SalesWithCumulativePercent

ORDER BY Total\_Sales DESC;

A screenshot of a data sheet

AI-generated content may be incorrect.

Query Explanation:

1. RankedSales CTE → Calculates total sales per item.
2. SalesWithRank CTE →
   * Calculates cumulative sales as we go down from the highest-selling item.
   * Finds overall sales total.
3. SalesWithCumulativePercent CTE → Converts the cumulative sales into a percentage of total sales.
4. Final SELECT →
   * Shows each item’s sales, cumulative sales %,
   * Classifies them into “Top 80% Sales” or “20% Sales” (Pareto category).
   * Orders items by sales (highest first).