**BLINKIT Analysis**

Exploring the data using SQL

* See all data imported:

SELECT \* FROM blinkit\_data

* DATA CLEANING:

Cleaning the Item\_Fat\_Content field ensures data consistency and accuracy in analysis. The presence of multiple variations of the same category (e.g., LF, low fat vs. Low Fat) can cause issues in reporting, aggregations, and filtering. By standardizing these values, we improve data quality, making it easier to generate insights and maintain uniformity in our datasets.

UPDATE blinkit\_data

SET Item\_Fat\_Content =

CASE

WHEN Item\_Fat\_Content IN ('LF', 'low fat') THEN 'LowFat'

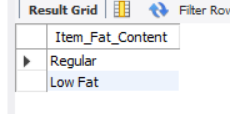
WHEN Item\_Fat\_Content = 'reg' THEN 'Regular'

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 blinkit\_data;



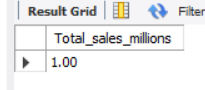
**A. KPI’s**

**1.Total Sales:**

SELECT CAST(SUM(Total\_Sales) / 1000000.0 AS DECIMAL(10,2)) AS Total\_Sales\_Million

FROM blinkit\_data;

This code displays the total sales in millions

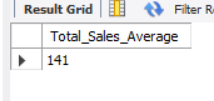
****

**2.Average Sales:**

SELECT CAST(AVG(Total\_Sales) AS INT) AS Avg\_Sales

FROM blinkit\_data;

This code displays the total average of sales

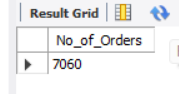


**3.Number of Sales**

SELECT COUNT(\*) AS No\_of\_Orders

FROM blinkit\_data;

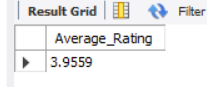
This code shows the total number of transactions in the dataset



**4.Average Rating**

SELECT CAST(AVG(Rating) AS DECIMAL(10,1)) AS Avg\_Rating

FROM blinkit\_data;

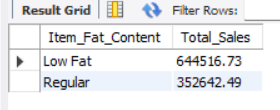


**B.Total Sales by Fat Content**

SELECT Item\_Fat\_Content, CAST(SUM(Total\_Sales) AS DECIMAL(10,2)) AS Total\_Sales

FROM blinkit\_data

GROUP BY Item\_Fat\_Content



**C. KPI’s by Item Type**

SELECT

Item\_Type,

COUNT(\*) AS total\_items,

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

CAST(AVG(Total\_Sales) as DECIMAL(10,0)) as Total\_Sales\_Average ,

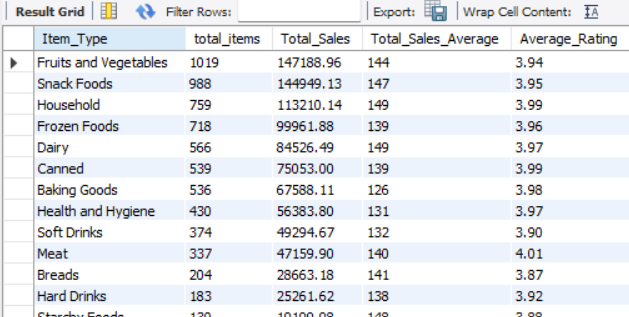
CAST(AVG(Rating) AS DECIMAL (10 , 2 )) AS Average\_Rating

FROM

blinkit\_data

GROUP BY Item\_Type

ORDER BY Total\_Sales DESC;

****

**D. Total sales of fat content on Outlet location type**

SELECT Outlet\_Location\_Type,

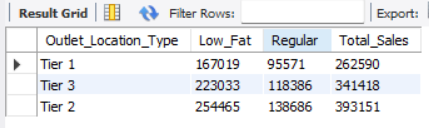
ROUND(SUM(CASE WHEN Item\_Fat\_Content='Low Fat' THEN Total\_Sales ELSE 0 END)) AS Low\_Fat,

ROUND(SUM(CASE WHEN Item\_Fat\_Content='Regular' THEN Total\_Sales ELSE 0 END)) AS Regular,

ROUND(SUM(CASE WHEN Item\_Fat\_Content='Low Fat' THEN Total\_Sales ELSE 0 END) + SUM(CASE WHEN Item\_Fat\_Content='Regular' THEN Total\_Sales ELSE 0 END)) AS Total\_Sales

FROM blinkit\_data

group by Outlet\_Location\_Type;



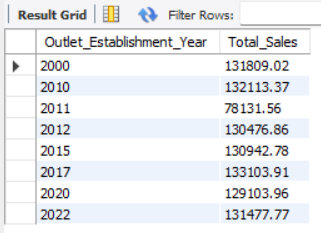
**E. Total sales by Outlet establishment**

SELECT Outlet\_Establishment\_Year, ROUND(SUM(Total\_Sales),2) AS Total\_Sales

FROM blinkit\_data

GROUP BY Outlet\_Establishment\_Year

ORDER BY Outlet\_Establishment\_Year;

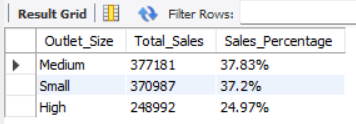


**F. Percentage of Sales by Outlet Size**

SELECT Outlet\_Size, round(sum(Total\_Sales)) as Total\_Sales,

concat(Round((sum(Total\_Sales)\*100 / sum(sum(Total\_Sales)) over()),2), '%') AS Sales\_Percentage from blinkit\_data

group by Outlet\_Size;



**E. All metrics on outlet type**

SELECT Outlet\_Type,

round(sum(Total\_Sales),2) AS Total\_Sales,

round(avg(Total\_sales),2) as Average\_Sales,

count(\*) AS No\_of\_Items,

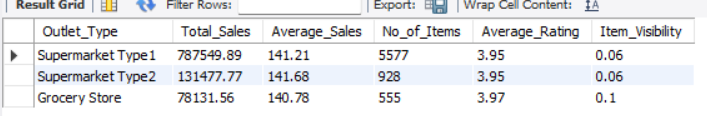
round(AVG(Rating),2) as Average\_Rating,

round(AVG(Item\_visibility),2) AS Item\_Visibility

FROM blinkit\_data

group by Outlet\_Type

order by Total\_Sales desc;



**Insights from The SQL Data Analysis**

**Customer Preference for Low-Fat Items**

Analyzing total sales by item fat content reveals a clear customer preference for "Low Fat" items. "Low Fat" products generated $644,516.73 in total sales, significantly outperforming "Regular" items which accounted for $352,642.49 in sales. This disparity strongly suggests that customers are more inclined to purchase products labeled as "Low Fat," indicating a potential trend in consumer behavior towards healthier options or a perception of such products being more desirable. This insight can inform inventory management, marketing strategies, and product development to align with customer preferences

**Dominance of Fruits, Vegetables, and Snack Foods Categories**

Fruits and Vegetables" and "Snack Foods" emerge as the top two performing categories in terms of total sales. "Fruits and Vegetables" generated $147,188.96 in sales, closely followed by "Snack Foods" with $144,949.13 in sales. This indicates a strong customer preference for ready-to-eat options and essential produce over meal-like ingredients or other item types. This trend suggests that consumers may be prioritizing convenience and healthier snacking or fresh produce choices in their purchases.

**Sales Performance Across Outlet Sizes**

When analyzing total sales by outlet size, the order of sales volume is "Medium" > "Small" > "High." "Medium" outlets generated the highest total sales at $377,181 (37.83%), closely followed by "Small" outlets with $370,987 (37.2%). "High" outlets recorded $248,992 (24.97%) in total sales. While "Medium" outlets lead, the sales difference between "Medium" and "High" outlets is substantial, at over $128,000. However, it's particularly noteworthy that "Medium" and "Small" outlets perform almost identically in terms of total sales and percentage contribution, indicating both sizes are highly effective sales channels.

**"Supermarket Type 1" as Key Sales Driver**

Among the outlet types, "Supermarket Type 1" significantly dominates in total sales and item volume. "Supermarket Type 1" generated $787,549.89 in total sales from 5,577 items, while "Supermarket Type 2" had $131,477.77 in total sales from 928 items, and "Grocery Store" recorded $78,131.56 in total sales from 555 items. Although the average sales per transaction and average ratings are remarkably similar across all three outlet types (around $141 and 3.95 respectively), the substantial difference in the number of items sold clearly highlights "Supermarket Type 1" as the primary driver of volume and overall revenue

**Quick overview of all the insights above**

* **Low-Fat Product Preference:** Customers show a clear preference for "Low Fat" items over "Regular" ones.
* **Top-Selling Categories:** "Fruits and Vegetables" and "Snack Foods" are the highest-performing categories, indicating a demand for convenience and fresh produce.
* **Tier 2 Outlets Lead Sales:** "Tier 2" outlets generate the highest total sales among all location types, with "Medium" and "Small" outlet sizes performing similarly well.
* **"Supermarket Type 1" Dominance:** "Supermarket Type 1" is the primary driver of sales volume and overall revenue, significantly outselling other outlet types