Blinkit Analysis

- See all the data imported:

```
SELECT * FROM BlinkIT_Grocery_Data;
```

- Data Cleaning – Standardizing Item_Fat_Content Field

  To ensure consistency and accuracy in our dataset, it's important to clean the Item_Fat_Content column. This field contains multiple variations of the same category, such as 'LF', 'low fat', and 'Low Fat', which all represent the same concept. Similarly, 'reg' and 'Regular' refer to the same category. Such inconsistencies can lead to incorrect reporting, faulty aggregations, and filtering issues.
  By standardizing these values, we enhance data quality, improve the reliability of insights, and maintain uniformity across the dataset.

```
UPDATE BlinkIT_Grocery_Data
SET Item_Fat_Content =
    CASE
        WHEN Item_Fat_Content IN ('LF', 'low fat') THEN 'Low Fat'
        WHEN Item_Fat_Content = 'reg' THEN 'Regular'
        ELSE Item_Fat_Content
           END;
```

- To confirm the field has been cleaned and standardized correctly, run the following query:

```
SELECT DISTINCT Item_Fat_Content FROM BlinkIT_Grocery_Data;
```

Results | Messages

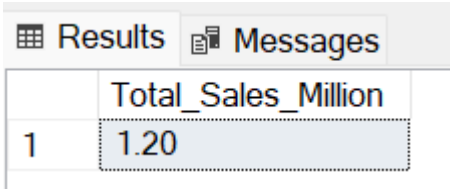
	Item_Fat_Content
1	Low Fat
2	Regular

akash-badhe

A. KPI's

1. TOTAL SALES:

```
SELECT CAST(SUM(Sales) / 1000000.0 AS DECIMAL(10,2))
AS Total_Sales_Million
FROM BlinkIT_Grocery_Data;
```

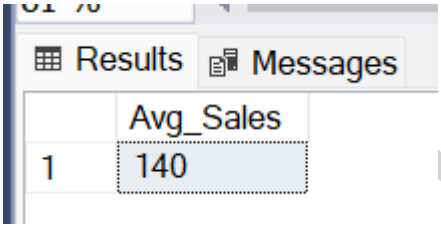


Results | Messages

	Total_Sales_Million
1	1.20

2. AVERAGE SALES

```
SELECT CAST(AVG(Sales) AS INT) AS Avg_Sales
FROM BlinkIT_Grocery_Data;
```

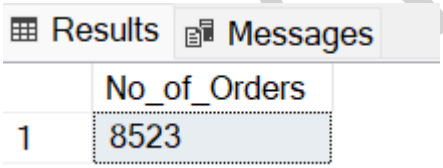


Results | Messages

	Avg_Sales
1	140

3. NO OF ITEMS

```
SELECT COUNT(*) AS No_of_Orders
FROM BlinkIT_Grocery_Data;
```



Results | Messages

	No_of_Orders
1	8523

4. AVG RATING

```
SELECT CAST(AVG(Rating) AS DECIMAL(10,1)) AS Avg_Rating
FROM BlinkIT_Grocery_Data;
```

Results | Messages

	Avg_Rating
1	4.0

ansh-badhe

B. Total Sales by Fat Content:

```
SELECT Item_Fat_Content, CAST(SUM(Sales) AS DECIMAL(10,2)) AS Total_Sales
FROM BlinkIT_Grocery_Data
GROUP BY Item_Fat_Content;
```

Results Messages

	Item_Fat_Content	Total_Sales
1	Low Fat	776319.69
2	Regular	425361.80

C. Total Sales by Item Type:

```
SELECT Item_Type, CAST(SUM(Sales) AS DECIMAL(10,2)) AS Total_Sales
FROM BlinkIT_Grocery_Data
GROUP BY Item_Type
ORDER BY Total_Sales DESC;
```

Results Messages

	Item_Type	Total_Sales
1	Fruits and Vegetables	178124.08
2	Snack Foods	175433.92
3	Household	135976.53
4	Frozen Foods	118558.88
5	Dairy	101276.46
6	Canned	90706.73
7	Baking Goods	81894.74
8	Health and Hygiene	68025.84
9	Meat	59449.86
10	Soft Drinks	58514.17
11	Breads	35379.12
12	Hard Drinks	29334.68
13	Others	22451.89
14	Starchy Foods	21880.03
15	Breakfast	15596.70
16	Seafood	9077.87

D. Fat Content by Outlet for Total Sales:

```
SELECT Outlet_Location_Type,
ISNULL([Low Fat], 0) AS Low_Fat,
ISNULL([Regular], 0) AS Regular
FROM
(
 SELECT Outlet_Location_Type, Item_Fat_Content,
 CAST(SUM(Sales) AS DECIMAL(10,2)) AS Total_Sales
 FROM BlinkIT_Grocery_Data
 GROUP BY Outlet_Location_Type, Item_Fat_Content
) AS SourceTable
PIVOT
(
 SUM(Total_Sales)
 FOR Item_Fat_Content IN ([Low Fat], [Regular])
) AS PivotTable
ORDER BY Outlet_Location_Type;
```

Results | Messages

	Outlet_Location_Type	Low_Fat	Regular
1	Tier 1	215047.91	121349.90
2	Tier 2	254464.78	138685.87
3	Tier 3	306807.00	165326.04

E. Total Sales by Outlet Establishment:

```
SELECT Outlet_Establishment_Year, CAST(SUM(Sales) AS DECIMAL(10,2)) AS
Total_Sales
FROM BlinkIT_Grocery_Data
GROUP BY Outlet_Establishment_Year
ORDER BY Outlet_Establishment_Year;
```

Results | Messages

	Outlet_Establishment_Year	Total_Sales
1	2011	78131.57
2	2012	130476.86
3	2014	131809.02
4	2015	130942.78
5	2016	132113.37
6	2017	133103.91
7	2018	204522.26
8	2020	129103.96
9	2022	131477.78

F. Percentage of Sales by Outlet Size:

```
SELECT Outlet_Size,
CAST(SUM(Sales) AS DECIMAL(10,2)) AS Total_Sales,
CAST((SUM(Sales) * 100.0 / SUM(SUM(Sales)) OVER()) AS DECIMAL(10,2)) AS
Sales_Percentage
FROM BlinkIT_Grocery_Data
GROUP BY Outlet_Size
ORDER BY Total_Sales DESC;
```

Results | Messages

	Outlet_Size	Total_Sales	Sales_Percentage
1	Medium	507895.74	42.27
2	Small	444794.17	37.01
3	High	248991.59	20.72

G. Sales by Outlet Location:

```
SELECT Outlet_Location_Type, CAST(SUM(Sales) AS DECIMAL(10,2)) AS Total_Sales
FROM BlinkIT_Grocery_Data
GROUP BY Outlet_Location_Type
ORDER BY Total_Sales DESC;
```

Results | Messages

	Outlet_Location_Type	Total_Sales
1	Tier 3	472133.03
2	Tier 2	393150.65
3	Tier 1	336397.81

H. All Metrics by Outlet Type:

```
SELECT Outlet_Type,
CAST(SUM(Sales) AS DECIMAL(10,2)) AS Total_Sales,
      CAST(AVG(Sales) AS DECIMAL(10,0)) AS Avg_Sales,
      COUNT(*) AS No_Of_Items,
      CAST(AVG(Rating) AS DECIMAL(10,2)) AS Avg_Rating,
      CAST(AVG(Item_Visibility) AS DECIMAL(10,2)) AS Item_Visibility
FROM BlinkIT_Grocery_Data
GROUP BY Outlet_Type
ORDER BY Total_Sales DESC;
```

Results | Messages

	Outlet_Type	Total_Sales	Avg_Sales	No_Of_Items	Avg_Rating	Item_Visibility
1	Supermarket Type1	787549.89	141	5577	3.96	0.06
2	Grocery Store	151939.15	140	1083	3.99	0.10
3	Supermarket Type2	131477.78	142	928	3.97	0.06
4	Supermarket Type3	130714.67	140	935	3.95	0.06