BlinkIT Data Analysis

- **See all the data imported:**

```
SELECT * FROM BlinkIT_data;
```

- **Check if all data is being imported or not:**

```
SELECT COUNT(*) FROM BlinkIT_data;
```

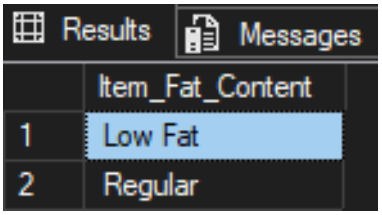
- **Data Cleaning:**

  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 '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

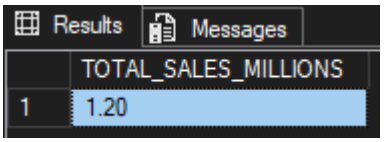
```
SELECT DISTINCT Item_Fat_Content FROM BlinkIT_data;
```



A. KPI's

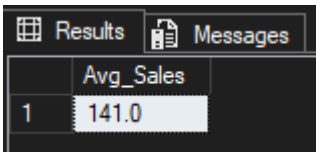
1. TOTAL SALES:

```
SELECT CAST(SUM(Total_Sales) / 1000000.0 AS DECIMAL(10,2)) AS
TOTAL_SALES_MILLIONS
FROM BlinkIT_data;
```



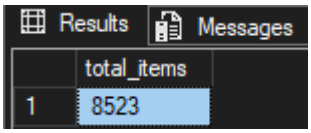
2. AVERAGE SALES

```
SELECT CAST(AVG(Total_Sales) AS INT) AS Avg_Sales
FROM BlinkIT_data;
```



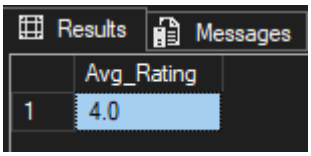
3. NUMBER OF ITEMS

```
SELECT COUNT(*) AS total_items
FROM BlinkIT_data;
```



4. AVG 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
Item_Fat_Content
FROM BlinkIT_data
GROUP BY Item_Fat_Content
```

Results | Messages

	Item_Fat_Content	Total_Sales
1	Low Fat	776319.68
2	Regular	425361.80

C. Total Sales by Item Type

```
SELECT Item_Type, CAST(SUM(Total_Sales) AS DECIMAL(10,2)) AS Item_Type
FROM BlinkIT_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	58514.16
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(Total_Sales) AS DECIMAL(10,2)) AS Total_Sales
    FROM BlinkIT_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	254464.77	138685.87
3	Tier 3	306806.99	165326.03

E. Total Sales, avg sales, total_items, avg rating by Outlet Establishment

```
SELECT Outlet_Establishment_Year,
       CAST(SUM(Total_Sales) AS DECIMAL(10,2)) AS Total_Sales,
       CAST(AVG(Total_Sales) AS DECIMAL(10,1)) AS Avg_Sales,
       COUNT(*) AS total_items,
                CAST(AVG(Rating) AS DECIMAL(10,2)) AS avg_rating
       FROM blinkit_data
       GROUP BY Outlet_Establishment_Year
```

ORDER BY Total_Sales DESC

Results | Messages

	Outlet_Establishment_Year	Total_Sales	Avg_Sales	total_items	avg_rating
1	1998	204522.26	139.8	1463	3.97
2	2017	133103.91	143.1	930	3.94
3	2010	132113.37	142.1	930	3.96
4	2000	131809.02	141.4	932	3.95
5	2022	131477.77	141.7	928	3.97
6	2015	130942.78	141.0	929	3.96
7	2012	130476.86	140.3	930	3.99
8	2020	129103.96	139.4	926	3.98
9	2011	78131.56	140.8	555	3.98

F. Percentage of Sales by Outlet Size

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

Results | Messages

	Outlet_Size	Total_Sales	Sales_Percentage
1	Medium	507895.73	42.27
2	Small	444794.17	37.01
3	High	248991.58	20.72

G. Sales by Outlet Location

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

Results | Messages

	Outlet_Location_Type	Total_Sales	Sales_Percentage	Avg_Sales	total_items	avg_rating
1	Tier 3	472133.03	39.29	140.9	3350	3.96
2	Tier 2	393150.64	32.72	141.2	2785	3.96
3	Tier 1	336397.81	27.99	140.9	2388	3.98

H. All Metrics by Outlet Type:

```
SELECT Outlet_Type,
            CAST(SUM(Total_Sales) AS DECIMAL(10,2)) AS Total_Sales,
            CAST(AVG(Total_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_data
GROUP BY Outlet_Type
ORDER BY Total_Sales DESC
```

Results | Messages

	Outlet_Type	Total_Sales	Sales_Percentage	Avg_Sales	total_items	avg_rating
1	Supermarket Type1	787549.89	65.54	141.2	5577	3.96
2	Grocery Store	151939.15	12.64	140.3	1083	3.99
3	Supermarket Type2	131477.77	10.94	141.7	928	3.97
4	Supermarket Type3	130714.67	10.88	139.8	935	3.95