

# Unlocking Business Insights at Blinkit



**A Data Analysis Project  
Using SQL**



LinkedIn  
[www.linkedin.com/in/zahraulayasifa/](https://www.linkedin.com/in/zahraulayasifa/)

Presented by  
**ZAHRA ULAYA SIFA**

# Introduction of **Company**



## About

Blinkit, formerly known as Grofers, is a leading Indian instant delivery service specializing in groceries and daily essentials. Founded in 2013, Blinkit has revolutionized the e-commerce grocery sector by adopting a quick commerce (q-commerce) model that delivers products to customers within 10–20 minutes. This ultra-fast delivery approach is supported by a network of micro-fulfillment centers, or dark stores, strategically located in urban areas.

### Vision

To become the most trusted and fastest delivery platform, providing essential services and products that enhance everyday life.

### Mission

Create a vibrant hyperlocal e-commerce economy that allows our partners – brands, merchants, delivery partners and customers – to benefit from the boom in digital transactions.

## BUSINESS

# REQUIREMENT

To conduct a comprehensive analysis of Blinkit's sales performance, customer satisfaction, and inventory distribution to identify key insights and opportunities for optimization using various KPIs and visualizations in Power BI.

**KPI's Requirements:**

- 1.Total Sales: The overall revenue generated from all items sold.
- 2.Average Sales: The average revenue per sale.
- 3.Number of Items: The total count of different items sold.
- 4.Average Rating: The average customer rating for items sold.

**Granular Requirements:**

- **1. Total Sales by Fat Content:**

Objective: Analyze the impact of fat content on total sales.

Additional KPI Metrics: Assess how other KPIs (Average Sales, Number of Items, Average Rating) vary with fat content.

- **2. Total Sales by Item Type:**

Objective: Identify the performance of different item types in terms of total sales.

Additional KPI Metrics: Assess how other KPIs (Average Sales, Number of Items, Average Rating) vary with item type.

- **3. Fat Content by Outlet for Total Sales:**

Objective: Compare total sales across different outlets segmented by fat content.

Additional KPI Metrics: Assess how other KPIs (Average Sales, Number of Items, Average Rating) vary with fat content.

- **4. Total Sales by Outlet Establishment:**

Objective: Evaluate how the age or type of outlet establishment influences total sales.

- **5. Percentage of Sales by Outlet Size:**

Objective: Analyze the correlation between outlet size and total sales.

- **6. Total Sales by Outlet Location:**

Objective: Assess the geographic distribution of sales across different locations.

- **7. All Metrics by Outlet Type:**

Objective: Provide a comprehensive view of all key metrics (Total Sales, Average Sales, Number of Items, Average Rating, Item Visibility) broken down by different outlet types.

# ABOUT DATASET

Link: [www.kaggle.com/datasets/arunkumaroraon/blinkit-grocery-dataset](https://www.kaggle.com/datasets/arunkumaroraon/blinkit-grocery-dataset)

Total Rows: 8523

No	Column	Description
1	Item Fat Content	Categorization of the fat content of the grocery items (e.g., Low Fat, Regular).
2	Item Identifier	Unique identifier for each grocery item in the dataset.
3	Item Type	Category or type of the grocery item (e.g., Dairy, Frozen Foods, Snacks).
4	Outlet Establishment Year	Year when the outlet (store) was established.
5	Outlet Identifier	Unique identifier for each outlet (store) in the dataset.
6	Outlet Location Type	Type of location where the outlet is situated (e.g., Urban, Rural).
7	Outlet Size	Size of the outlet (e.g., Small, Medium, High).
8	Outlet Type	Type of outlet (e.g., Grocery Store, Supermarket).
9	Item Visibility	Percentage of total display area of the item in the store.
10	Item Weight	Weight of the item.
11	Sales	Sales of the item in the given time period.
12	Rating	Customer rating or feedback score for the item or the outlet.

# Analysis TOOLS



Database Name: portfolio  
Schema Name: data  
Table Name: blinkit

## CLEANING DATA

```
SELECT * FROM data.blinkit;
```

	item_fat_content character varying (50)	item_identifier character varying (50)	item_type character varying (50)	outlet_establishment_year integer	outlet_identifier character varying (50)	outlet_location_type character varying (50)	o cl
1	Regular	FDX32	Fruits and Vegetables	2012	OUT049	Tier 1	N
2	Low Fat	NCB42	Health and Hygiene	2022	OUT018	Tier 3	N
3	Regular	FDR28	Frozen Foods	2016	OUT046	Tier 1	S
4	Regular	FDL50	Canned	2014	OUT013	Tier 3	F
5	Low Fat	DRI25	Soft Drinks	2015	OUT045	Tier 2	S
6	low fat	FDS52	Frozen Foods	2020	OUT017	Tier 2	S
7	Low Fat	NCU05	Health and Hygiene	2011	OUT010	Tier 3	S
8	Low Fat	NCD30	Household	2015	OUT045	Tier 2	S
9	Low Fat	FDW20	Fruits and Vegetables	2014	OUT013	Tier 3	F
10	Low Fat	FDX25	Canned	2018	OUT027	Tier 3	N
11	LF	FDX21	Snack Foods	2018	OUT027	Tier 3	N
12	Low Fat	NCU41	Health and Hygiene	2017	OUT035	Tier 2	S
13	Low Fat	FDL20	Fruits and Vegetables	2022	OUT018	Tier 3	N
14	Low Fat	NCR54	Household	2014	OUT013	Tier 3	F
15	Low Fat	FDH19	Meat	2018	OUT027	Tier 3	N
16	Regular	FDB57	Fruits and Vegetables	2017	OUT035	Tier 2	S
17	Low Fat	FDO23	Breads	2022	OUT018	Tier 3	N

Total rows: 8523    Query complete 00:00:01.227    CRLF    Ln 1, Col 29

```
SELECT DISTINCT(item_fat_content) FROM data.blinkit;
```

	item_fat_content character varying (50)
1	Low Fat
2	low fat
3	Regular
4	reg
5	LF

In the dataset, **inconsistent** labeling such as **low fat**, **LF**, and **reg** were standardized to **Low Fat** and **Regular** respectively to ensure data uniformity.

# CLEANING DATA

## 1. Standardizing Data

```
UPDATE data.blinkit
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;
```

	item_fat_content character varying (50)	item_identifier character varying (50)	item_type character varying (50)	outlet_establishment_year integer	outlet_identifier character varying (50)	outlet_location_type character varying (50)	cl
1	Regular	FDT27	Meat	2017	OUT035	Tier 2	S
2	Low Fat	FDP39	Meat	2020	OUT017	Tier 2	F
3	Low Fat	NCF06	Household	2011	OUT010	Tier 3	M
4	Low Fat	FDX15	Meat	2014	OUT013	Tier 3	F
5	Regular	FDQ27	Meat	2022	OUT018	Tier 3	M
6	Low Fat	NCC18	Household	2011	OUT010	Tier 3	F
7	Low Fat	DRC27	Dairy	2020	OUT017	Tier 2	S
8	Regular	FDE40	Dairy	2018	OUT027	Tier 3	M
9	Regular	FDH31	Meat	2018	OUT019	Tier 1	S
10	Low Fat	FDD51	Dairy	2018	OUT027	Tier 3	M
11	Regular	FDC16	Dairy	2015	OUT045	Tier 2	F
12	Low Fat	FDN13	Breakfast	2018	OUT019	Tier 1	S
13	Regular	FDA03	Dairy	2018	OUT027	Tier 3	M
14	Regular	FDT39	Meat	2020	OUT017	Tier 2	M
15	Regular	FDX32	Fruits and Vegetables	2012	OUT049	Tier 1	M
16	Low Fat	NCB42	Health and Hygiene	2022	OUT018	Tier 3	M
17	Regular	FDR28	Frozen Foods	2016	OUT046	Tier 1	S

### Cross Check:

```
SELECT DISTINCT(item_fat_content) FROM data.blinkit;
```

	item_fat_content character varying (50)
1	Regular
2	Low Fat

# RESULT (KPI)

## 1. Total Sales (milion)

```
SELECT CAST(SUM(sales)/1000000 AS DECIMAL(20,2))  
AS total_sales_milion FROM data.blinkit;
```

	total_sales_milion numeric (20,2) 🔒
1	1.20

## 2. Average Sales

```
SELECT ROUND(CAST(AVG(sales) AS NUMERIC), 0) AS avg_sales  
FROM data.blinkit;
```

	avg_sales numeric 🔒
1	141

## 3. Number of Items

```
SELECT COUNT(*) AS no_of_item FROM data.blinkit;
```

	no_of_item bigint 🔒
1	8523

## 4. Average Rating

```
SELECT ROUND(CAST(AVG(rating) AS NUMERIC), 2) AS avg_rating  
FROM data.blinkit;
```

	avg_rating numeric 🔒
1	3.97



# RESULT (Granular)

## 1. Total Sales (thousand) by Fat Content

```
SELECT item_fat_content,
       CAST(SUM(sales)/1000 AS DECIMAL(20,2)) AS Total_Sales,
       ROUND(CAST(AVG(sales) AS NUMERIC), 2) AS avg_sales,
       COUNT(*) AS no_of_item,
       ROUND(CAST(AVG(rating) AS NUMERIC), 2) AS avg_rating
FROM data.blinkit GROUP BY 1;
```

	item_fat_content character varying (50) 🔒	total_sales numeric (20,2) 🔒	avg_sales numeric 🔒	no_of_item bigint 🔒	avg_rating numeric 🔒
1	Regular	425.36	141.50	3006	3.97
2	Low Fat	776.32	140.71	5517	3.97

## 2. Total Sales (milion) by Item Type

```
SELECT item_type,
       CAST(SUM(sales)/1000000 AS DECIMAL(20,2)) AS Total_Sales,
       ROUND(CAST(AVG(sales) AS NUMERIC), 2) AS avg_sales,
       COUNT(*) AS no_of_item,
       ROUND(CAST(AVG(rating) AS NUMERIC), 2) AS avg_rating
FROM data.blinkit GROUP BY 1 ORDER BY 1;
```

	item_type character varying (50) 🔒	total_sales numeric (20,2) 🔒	avg_sales numeric 🔒	no_of_item bigint 🔒	avg_rating numeric 🔒
1	Baking Goods	0.08	126.38	648	3.98
2	Breads	0.04	140.95	251	3.88
3	Breakfast	0.02	141.79	110	3.93
4	Canned	0.09	139.76	649	3.99
5	Dairy	0.10	148.50	682	3.97
6	Frozen Foods	0.12	138.50	856	3.97
7	Fruits and Vegetables	0.18	144.58	1232	3.96
8	Hard Drinks	0.03	137.08	214	3.91
9	Health and Hygiene	0.07	130.82	520	3.99
10	Household	0.14	149.42	910	4.00
11	Meat	0.06	139.88	425	4.02
12	Others	0.02	132.85	169	3.95
13	Seafood	0.01	141.84	64	3.96
14	Snack Foods	0.18	146.19	1200	3.95
15	Soft Drinks	0.06	131.49	445	3.92
16	Starchy Foods	0.02	147.84	148	3.92



# RESULT (Granular)

## 3. Fat Content by Outlet for Total Sales (milion)

```
SELECT outlet_location_type,item_fat_content,
CAST(SUM(sales)/1000000 AS DECIMAL(20,2)) AS Total_Sales,
ROUND(CAST(AVG(sales) AS NUMERIC), 2) AS avg_sales,
COUNT(*) AS no_of_item,
ROUND(CAST(AVG(rating) AS NUMERIC), 2) AS avg_rating
FROM data.blinkit GROUP BY 1,2 ORDER BY 1;
```

	outlet_location_type character varying (50) 🔒	item_fat_content character varying (50) 🔒	total_sales numeric (20,2) 🔒	avg_sales numeric 🔒	no_of_item bigint 🔒	avg_rating numeric 🔒
1	Tier 1	Low Fat	0.22	139.64	1540	3.98
2	Tier 1	Regular	0.12	143.10	848	3.97
3	Tier 2	Low Fat	0.25	140.67	1809	3.97
4	Tier 2	Regular	0.14	142.10	976	3.95
5	Tier 3	Low Fat	0.31	141.52	2168	3.96
6	Tier 3	Regular	0.17	139.87	1182	3.97

## 4. Total Sales (thousand) by Outlet Establishment

```
SELECT outlet_establishment_year,
CAST(SUM(sales)/1000 AS DECIMAL(20,0)) AS Total_Sales
FROM data.blinkit GROUP BY 1 ORDER BY 1;
```

	outlet_establishment_year integer 🔒	total_sales numeric (20) 🔒
1	2011	78
2	2012	130
3	2014	132
4	2015	131
5	2016	132
6	2017	133
7	2018	205
8	2020	129
9	2022	131

# RESULT (Granular)

## 5. Percentage of Sales by Outlet Size

```
SELECT
  outlet_size,
  CAST(SUM(sales)/1000 AS DECIMAL(20,2)) AS total_sales,
  -- Persentase dari total_sales
  ROUND(
    (CAST(SUM(sales)/1000 AS NUMERIC) * 100.0) /
    SUM(CAST(SUM(sales)/1000 AS NUMERIC)) OVER (), 2
  ) AS pct_total_sales
FROM data.blinkit
GROUP BY 1
ORDER BY 3 DESC;
```

	outlet_size character varying (50) 🔒	total_sales numeric (20,2) 🔒	pct_total_sales numeric 🔒
1	Medium	507.90	42.27
2	Small	444.79	37.01
3	High	248.99	20.72

## 6. Total Sales by Outlet Location

```
SELECT outlet_location_type,
  CAST(SUM(sales)/1000 AS DECIMAL(20,2)) AS Total_Sales
FROM data.blinkit GROUP BY 1 ORDER BY 2 DESC;
```

	outlet_location_type character varying (50) 🔒	total_sales numeric (20,2) 🔒
1	Tier 3	472.13
2	Tier 2	393.15
3	Tier 1	336.40

# RESULT (Granular)

## 7. All Metrics by Outlet Type

```
SELECT outlet_type,  
       CAST(SUM(sales)/1000 AS DECIMAL(20,0)) AS Total_Sales,  
       COUNT(*) AS no_of_item,  
       ROUND(CAST(AVG(sales) AS NUMERIC), 0) AS avg_sales,  
       ROUND(CAST(AVG(rating) AS NUMERIC), 2) AS avg_rating,  
       ROUND(CAST(AVG(item_visibility) AS NUMERIC), 2) AS avg_item_visibility  
FROM data.blinkit GROUP BY 1 ORDER BY 1;
```

	outlet_type character varying (50) 🔒	total_sales numeric (20) 🔒	no_of_item bigint 🔒	avg_sales numeric 🔒	avg_rating numeric 🔒	avg_item_visibility numeric 🔒
1	Grocery Store	152	1083	140	3.99	0.10
2	Supermarket Type1	788	5577	141	3.96	0.06
3	Supermarket Type2	131	928	142	3.97	0.06
4	Supermarket Type3	131	935	140	3.95	0.06



# CONCLUSION

No	Description
01	The Low Fat category outperformed the Regular category in both total sales and number of items sold, contributing 64.73% of the total sales volume. Despite this, the Regular category had a slightly higher average sales per item (141.50 vs. 140.71). Both categories received the same average customer rating of 3.97, indicating similar levels of customer satisfaction.
02	The Fruits and Vegetables and Snack Foods categories recorded the highest total sales, driven by substantial item volumes, while Household and Dairy products exhibited the highest average sales per item, and Household and Meat received the highest customer satisfaction ratings at 4.00 and 4.02 respectively.
03	Tier 3 outlets showed the highest total sales and item volume for both Low Fat and Regular products, suggesting that consumer demand in smaller or semi-urban areas may be higher regardless of fat content.
04	Outlets established in 2018 achieved the highest total sales, indicating a possible correlation between newer infrastructure or strategy and improved sales performance.
05	Medium-sized outlets dominated total sales with a 42.27% contribution, outperforming both small and high-sized outlets, suggesting an optimal balance between capacity and efficiency.
06	outlets in Tier 3 areas (semi-urban or smaller towns) recorded the highest total sales (472.13) compared to Tier 2 and Tier 1, indicating that demand may be higher in less urbanized regions, possibly due to a larger consumer base or lower market competition in those areas.
07	Supermarket Type1 significantly outperformed other outlet types in total sales (788) and item volume (5,577), indicating its dominant role in product distribution, while Grocery Stores had the highest item visibility (0.10) and rating (3.99), suggesting better customer experience despite lower sales.



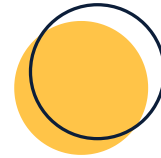
# RECOMMENDATION

No	Description
01	Focus on Low Fat products due to their dominant contribution to total sales and volume, as well as high customer satisfaction levels.
02	Strengthen promotional and distribution strategies for the Fruits, Vegetables, and Snacks categories with high sales volume, and leverage the high customer satisfaction in the Household and Meat categories.
03	Increase market penetration in Tier 3 areas, which show strong demand for both Low Fat and Regular products.
04	Analyze and replicate the strategies used by outlets opened after 2018, as they demonstrate better sales performance.
05	Prioritize medium-sized outlets as they offer an optimal balance between efficiency and sales performance.
06	Maximize the potential of Grocery Stores through improved conversion strategies, such as better product placement and targeted promotions.
07	Expand the Supermarket Type1 model to other regions due to its outstanding performance in total sales and product volume.

**For more detailed information, please refer to the Power BI dashboard available at [bit.ly/githubPBIblinkit](https://bit.ly/githubPBIblinkit) and [bit.ly/githubsqlblinkit](https://bit.ly/githubsqlblinkit) for detail query.**

# Thank You

Look forward to potential  
collaboration.



Linkedin

[www.linkedin.com/in/zahraulayasifa/](https://www.linkedin.com/in/zahraulayasifa/)

Gmail

[zahraulayasifa.work@gmail.com](mailto:zahraulayasifa.work@gmail.com)

Github

[github.com/zahraulayasifa-projects](https://github.com/zahraulayasifa-projects)