Blinkit Data Analysis

Date: [5/10/2024]

Prepared by: [Abhishek Munde]

---

1. Executive Summary

We analyzed the BlinkIT grocery dataset by examining sales trends across various item types, fat content, and outlet types. Key insights include high sales Item Types and slightly better performance for regular-fat items. We also found that Supermarket Type1 outlets tend to generate the highest total sales.

2. Objective

The purpose of this analysis is to identify key factors driving sales performance across different grocery items and outlets at BlinkIT. Specifically, we aim to understand which product categories, fat content types, and outlet types contribute the most to total sales. Additionally, we are analyzing item visibility and outlet size to uncover their potential impact on sales. Ultimately, the goal is to provide insights and recommendations to improve sales strategies, optimize product placement, and enhance customer satisfaction across outlets.

3. Data Overview

3.1 Data Source

The data comes from Kaggle website, capturing information on various grocery items sold across multiple outlet types. It appears to be a structured dataset provided by the company's database, containing detailed records of type of item which were sold over the years , outlet features, and corresponding sales performance and their rating, fat content.

Date Range:

The data spans outlets established from **2011 to 2022**, though the specific sales transactions in the dataset.

3.2 Data Description

The dataset contains information on various grocery items and their sales performance across BlinkIT outlets. The main variables include:

- **Item Information**:
    - *Item Identifier*, *Item Type*, *Item Fat Content*, *Item Weight*, and *Item Visibility*.
- **Outlet Information**:

- *Outlet Identifier, Outlet Type, Outlet Size, Outlet Location Type*, and *Outlet Establishment Year*.

- **Sales and Ratings**:

  - *Sales* (total revenue per item) and *Rating* (customer feedback score).

  Data was in Cleaned format we just done work on Item Fat Content column like sandardized inconsistent labels in *Item Fat Content* (e.g., "LF" and "Low Fat" were unified).

4. Methodology

Tools Used

The analysis was conducted using Python and pandas for data manipulation and aggregation. Microsoft Excel was used to initially view and organize the dataset.

Statistical Methods:

Descriptive statistics were used to summarize key metrics (e.g., total sales, item weights, visibility). We calculated overall and category-wise sales by grouping data by variables such as Item Type, Item Fat Content, Outlet Type and their ratings. Additionally, we explored relationships between variables like outlet typpe and Sales to identify patterns.

Approach:

- We started by standardizing categorical variables.

- Grouped and aggregated sales data to identify top-performing items and outlets.

- Explored sales trends by item category, outlet type, and fat content to uncover key drivers of performance.

- Done analysis and co-relation between columns to find out is there any relation between them.

- Visualizations and summary statistics were used to extract insights.

5. Analysis

1). Sales Distribution of Top 5 Item Types by Customer Rating

1. High Sales in Top-Rated Items

   - The chart shows that a significant portion of the total sales for top-performing item types comes from the "4 to 5" rating group. This reinforces the importance of customer satisfaction. High-rated items dominate sales, suggesting a strong correlation between customer approval and revenue generation.
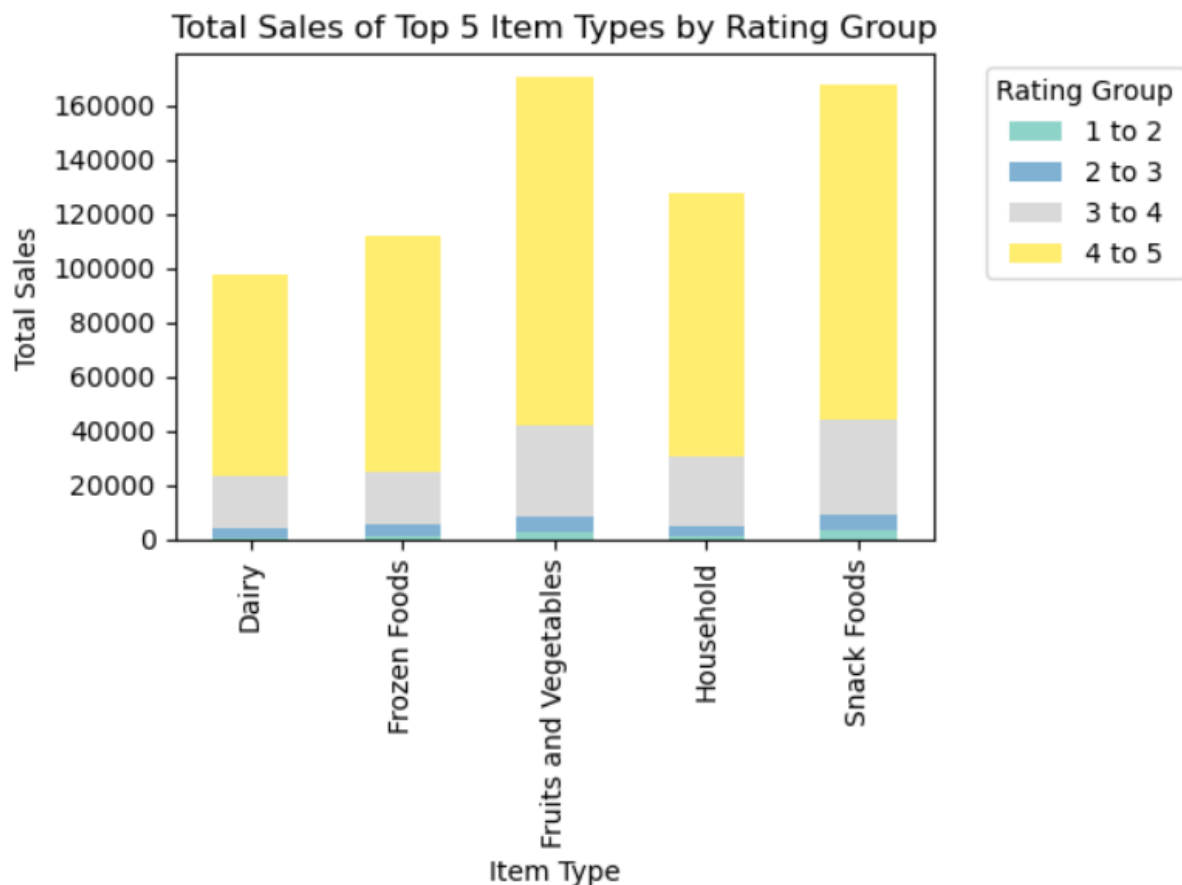
2. Low-Rated Items Have Minimal Impact:

- Sales contributions from items rated in the "1 to 2" and "2 to 3" groups are relatively small across the top item types. This implies that poor-rated products do not contribute much to overall sales, highlighting the importance of maintaining product quality.

Recommendation:

-By analysis of this chart we can say that top rated items have more sales than low rated items so we can increase sales of items which has lower ratings by encouraging customer feedback and quickly addressing issues for lower-rated items can help improve their ratings and, in turn, their sales potential.

-we can also increase sales by making marketing promotions by highlighting highly rated items in advertisements, loyalty programs, or discount offers can attract quality-conscious consumers.
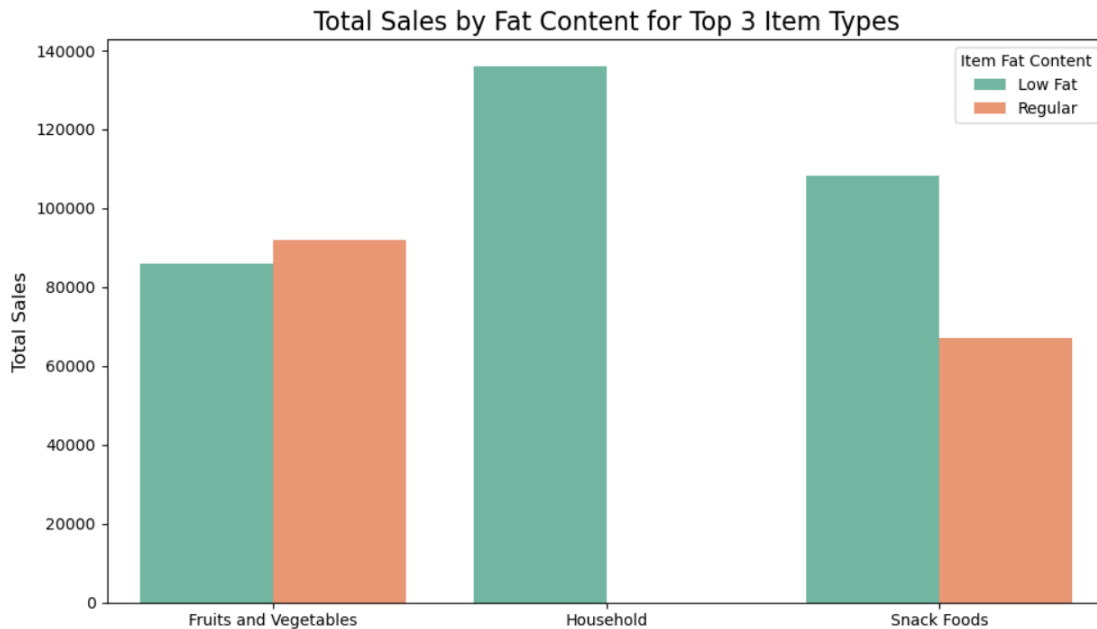
Total Sales of Top 5 Item Types by Rating Group

Rating Group
- 1 to 2
- 2 to 3
- 3 to 4
- 4 to 5

Item Type: Dairy, Frozen Foods, Fruits and Vegetables, Household, Snack Foods

Item Fat Content

Low Fat   5517

Regular   3006

2). Total sales by Fat Content top 3 Items:

Total Sales by Fat Content for Top 3 Item Types

1. High Sales in Low fat:
   - The chart shows that a significant portion of the total sales for top-performing item types comes from the low fat content in Items. This indicates that customer prefer low content items.
   - Household items have only low fat content that's why the household items most sale by fat content wise.

2. Customer preference:

   -In Snack foods more sales than fruits and vegetables by low fat content customer chosen low fat over regular in snacks.

Recommendation:

By analysis of this chart we can say top 3 items which has highest sales by content wise hat we can increase production of items which has low fat content.

if Low Fat products tend to receive higher ratings, BlinkIT can leverage this insight to focus on expanding its range of health-conscious offerings.

If Regular Fat products perform equally well, it suggests that consumers prioritize taste and indulgence, implying that promoting both health and flavor-oriented products can maximize sales.

0	Baking Goods	Low Fat	39653.0108
1	Baking Goods	Regular	42241.7256
2	Breads	Low Fat	18841.5058
3	Breads	Regular	16537.6140

4	Breakfast	Low Fat	6775.6466
5	Breakfast	Regular	8821.0500
6	Canned	Low Fat	47587.4086
7	Canned	Regular	43119.3204
8	Dairy	Low Fat	61494.4852
9	Dairy	Regular	39781.9764
10	Frozen Foods	Low Fat	62558.7706
11	Frozen Foods	Regular	56000.1108
12	Fruits and Vegetables	Low Fat	86031.5834
13	Fruits and Vegetables	Regular	92092.4976
14	Hard Drinks	Low Fat	29334.6806
15	Health and Hygiene	Low Fat	68025.8388
16	Household	Low Fat	135976.5254
17	Meat	Low Fat	22269.7940
18	Meat	Regular	37180.0698
19	Others	Low Fat	22451.8916
20	Seafood	Low Fat	4199.8166
21	Seafood	Regular	4878.0534
22	Snack Foods	Low Fat	108296.7374
23	Snack Foods	Regular	67137.1850
24	Soft Drinks	Low Fat	49868.4250
25	Soft Drinks	Regular	8645.7420
26	Starchy Foods	Low Fat	12953.5680
27	Starchy Foods	Regular	8926.4594

3). Outlet Establishment Year :

Number of Outlets Established by Year
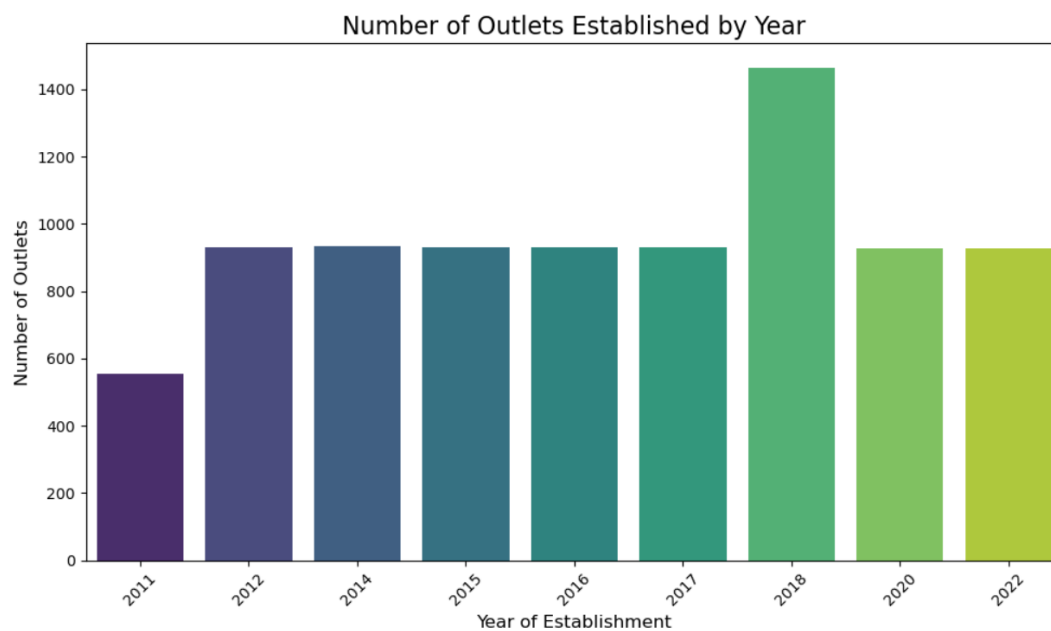
Significant Outlet Growth in 2018:

- The highest number of outlets was established in 2018, which indicates a period of rapid expansion. This suggests that the company experienced growth during this time, likely due to favorable market conditions or successful business strategies.

    From 2011 to 2017, the number of outlets opened each year remained relatively stable, indicating consistent but modest expansion. The business was likely building a strong foundation before the major growth spike in 2018.

    From 2018 to 2022 there is less outlet established in that 4 yrs.

Recommendation:

The sharp increase in 2018 followed by a slowdown suggests a potential revaluation of growth strategies. Understanding the factors behind the 2018 expansion could inform future decisions on whether to replicate that success or focus on optimizing the existing network of outlets.

Blinkit can observe the outlet establishment in 2018 and can work like that in future years to expand their outlets and grow sale more highly.

While growth slowed after 2018, the company could still identify underrepresented regions or explore new markets based on the success of earlier expansion years. A data-driven approach to identifying the most profitable areas for new outlets could help reignite growth.

4). Distribution of Sales by Outlet Types:

Dominance of Supermarket Type1:

- Supermarket Type1 accounts for a significant portion of sales, contributing 65.5%. This indicates that this format is the most successful and likely has the broadest customer appeal, product range, or operational efficiency.

    The other outlet types (Supermarket Type2, Supermarket Type3, and Grocery Stores) each contribute around 10-12% of total sales, indicating a more niche role in the overall business model. These outlets likely cater to specific customer segments or regions.

Recommendation:

Focus on Supermarket Type1:
- Given its overwhelming contribution to sales, the business should prioritize expanding Supermarket Type1 outlets. This could involve opening new locations, increasing the footprint in underrepresented areas, or further optimizing operations in existing locations.
- For the smaller outlet types (Supermarket Type2, Type3, and Grocery Stores), the business might adopt a targeted strategy. These outlet types could focus on specific products, regions, or customer demographics to improve their sales performance.
- Alternatively, if these outlet types are not delivering strong returns, the company could evaluate whether to downsize or refocus resources on better-performing formats.

By going throgh these insights, business decisions around resource allocation, outlet expansion, and market segmentation can be better informed, ultimately boosting profitability and customer satisfaction.

Sales Distribution by Outlet Type
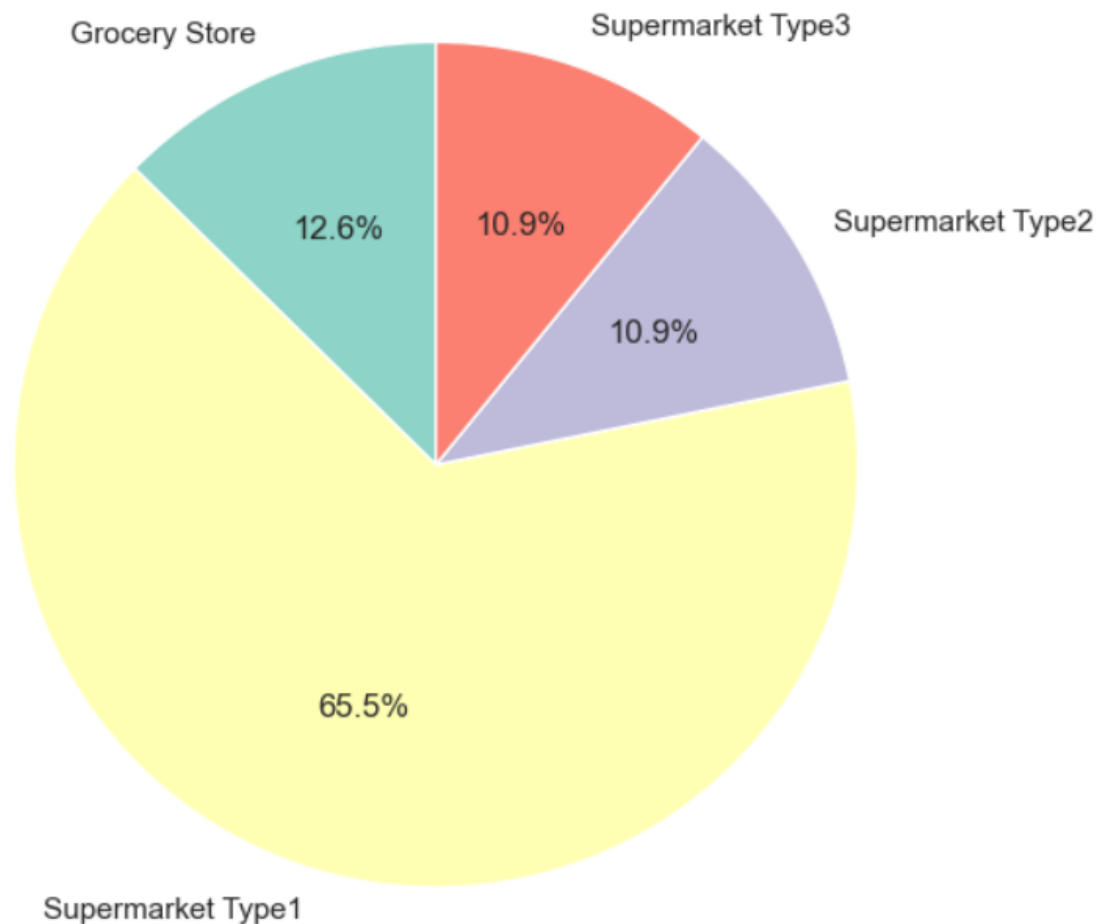
5). Most selling item type by Outlet Types:

As we see below bar chart we knew that there are most outlet types are Supermarket type 1 so the sales are also more in that sales **Fruits and vegetables** are most sold item type in Supermarket Type 1.

Also Fruits and vegetables dominate in sells in Grocery store and Supermarket Type 3 and in other hand Supermarket Type 2 has most sells of **Snack Foods**.

This indicates that each outlet type might serve different consumer needs or preferences.

Recommendation:

For **Supermarket Type1**, which clearly benefits from high sales in the Fruits and Vegetables category, the business should focus on optimizing inventory management and freshness in this

product category. Ensuring a reliable supply chain and offering promotions could further enhance sales.

Total Sales by Outlet Type with Most Sold Item Type
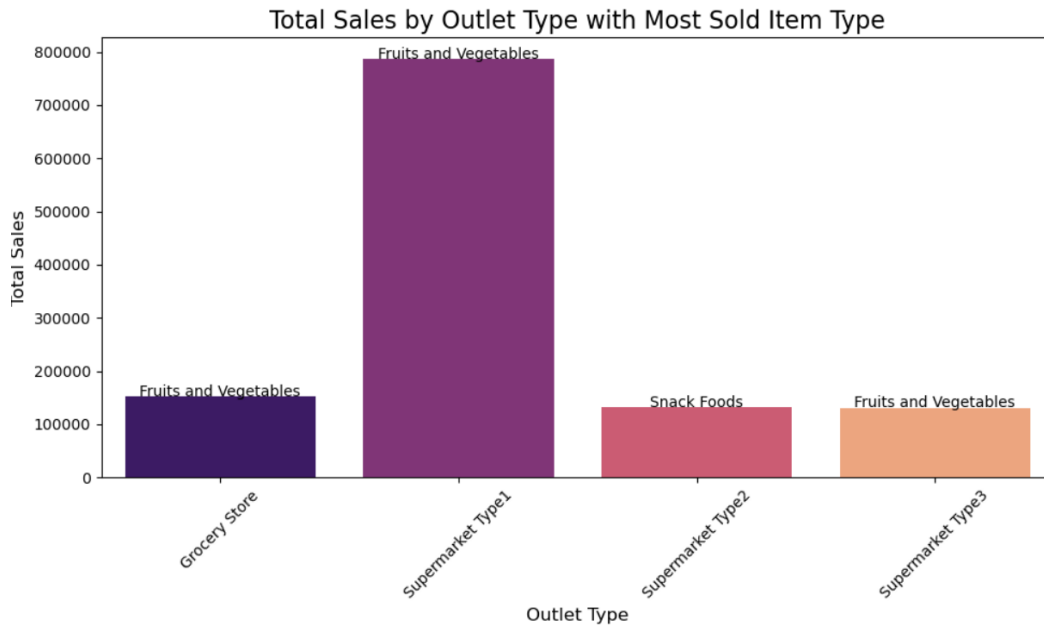
Since **Supermarket Type2** sees its highest sales in Snack Foods, marketing efforts could be tailored accordingly. Promotions, product placement, and partnerships with snack brands could improve performance further in these outlets.

In this we can focus on also other items to increase their sales we can promote marketing for other items increase their stocks find out the problems on their lower sales also give discount on the items.

These insights can guide **inventory planning, targeted marketing strategies, and supply chain decisions**, ensuring that each outlet type meets customer needs efficiently while maximizing sales.

## 6). Distribution of item type by Outlet Size:

As we can see on this Chart **Fruits and vegetables** have highest count of items in medium size outlets and also in small outlet sizes.

On other hand **snack foods items** have highest count in small outlet size.

Larger outlets (Medium or High size) may display a more diverse range of products, likely because they have the capacity to stock more item types. This can drive customer footfall as consumers may prefer these stores for the variety offered.

Smaller outlets may focus on essential or high-demand items, possibly limiting their product variety to optimize space and meet local demand efficiently.

Sea food , Starchy Food,Others and Breakfast Items has very low stock in all these high , medium and small outlet size.

Distribution of Item Type by Outlet Size
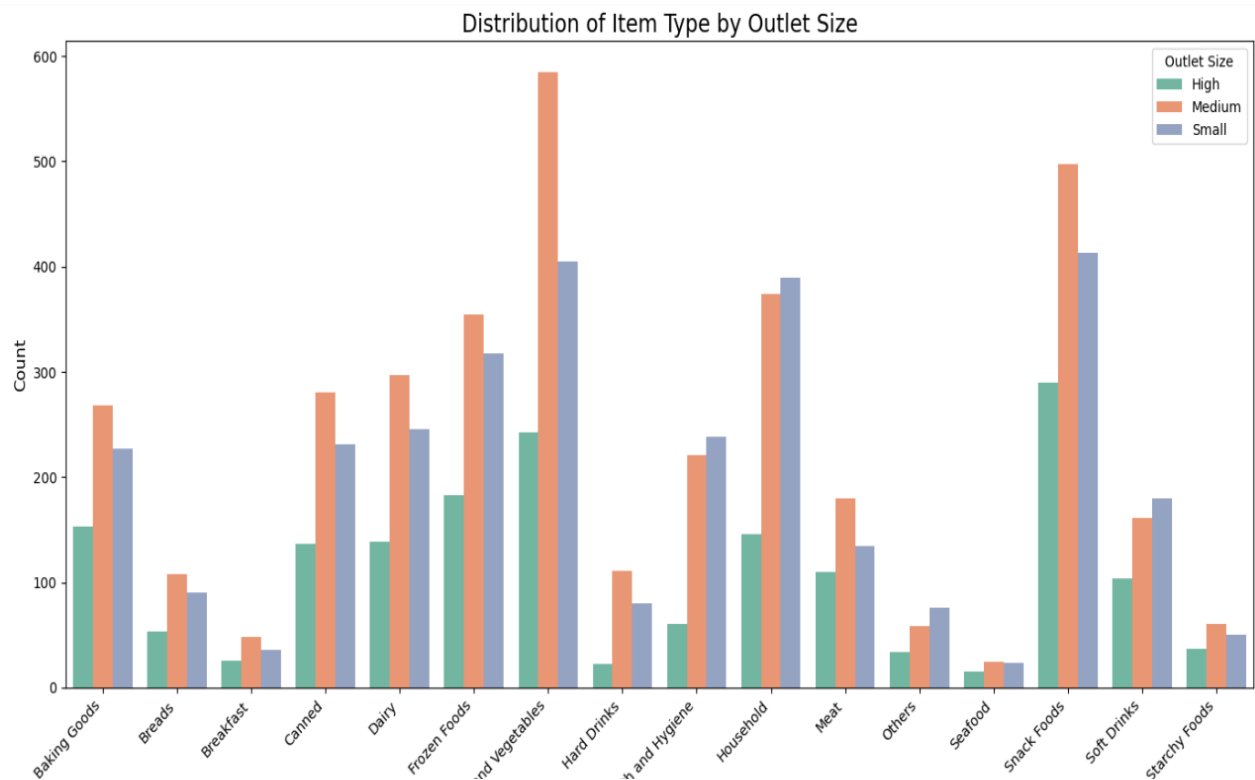
Recommendation:

Medium and large-sized outlets can use this data to adjust store layout and maximize the display of diverse items, ensuring popular categories are easy to find. Smaller outlets can optimize their layout to ensure efficiency for essential goods.

Product Mix Optimization: Retailers can adjust the product mix based on outlet size.Larger stores may prioritize offering a full range of items across all categories, while smaller stores focus on categories that have a high turnover, based on consumer preferences

Expansion Strategy: Businesses planning new outlet openings could use this data to determine the right product assortment based on store size, helping to cater effectively to customer needs and boost profitability.

These findings can help drive decisions on how to manage inventory, optimize product placement, and expand outlet sizes to better meet market demand.

6). Key Insights:

1) Sales by Outlet Size:

Larger outlets (Medium and High) show higher sales volume compared to smaller outlets. This suggests that larger outlets, with their ability to offer a wider variety of items, attract more customers and generate more revenue.

2) Item Type Distribution by Outlet Size:
   There is a greater diversity of item types in larger outlets. Categories like **Snacks** and **Fruits & Vegetables** dominate the offerings in these stores, which are in high demand and generate substantial sales. Smaller outlets, by contrast, focus more on essentials and less perishable goods.
   Fruits & Vegetables and Snacks has more count in this medium size outlets.
   Sea food , Starchy Food,Others and Breakfast Items has very low stock in all these high , medium and small outlet size.

3) Customer Ratings and Sales:

There is a direct positive relationship between customer ratings and sales. Higher-rated products tend to be associated with higher sales, indicating that customer satisfaction is critical for repeat business.

Rating for 4 to 5 has more sales and has **Fruits & Vegetables** has more sales in rating 4 to 5.

4) Total Sales by Fat Content for each Item Type:

	Item Type	Item Fat Content	Sales
0	Baking Goods	Low Fat	39653.0108
1	Baking Goods	Regular	42241.7256
2	Breads	Low Fat	18841.5058
3	Breads	Regular	16537.6140
4	Breakfast	Low Fat	6775.6466
5	Breakfast	Regular	8821.0500
6	Canned	Low Fat	47587.4086
7	Canned	Regular	43119.3204
8	Dairy	Low Fat	61494.4852
9	Dairy	Regular	39781.9764
10	Frozen Foods	Low Fat	62558.7706
11	Frozen Foods	Regular	56000.1108
12	Fruits and Vegetables	Low Fat	86031.5834
13	Fruits and Vegetables	Regular	92092.4976
14	Hard Drinks	Low Fat	29334.6806
15	Health and Hygiene	Low Fat	68025.8388
16	Household	Low Fat	135976.5254
17	Meat	Low Fat	22269.7940

18	Meat	Regular	37180.0698
19	Others	Low Fat	22451.8916
20	Seafood	Low Fat	4199.8166
21	Seafood	Regular	4878.0534
22	Snack Foods	Low Fat	108296.7374
23	Snack Foods	Regular	67137.1850
24	Soft Drinks	Low Fat	49868.4250
25	Soft Drinks	Regular	8645.7420
26	Starchy Foods	Low Fat	12953.5680
27	Starchy Foods	Regular	8926.4594

This shows that a significant portion of the total sales for top-performing item types comes from the low fat content in Items. This indicates that customer prefer low content items. Household items have only low fat content that's why the household items most sale by fat content wise.

Household items have most sells in low fat content by order of Snack food and Fruits and vegetables.

Conclusion:

Conclusion: BlinkIT Grocery Dataset Analysis

The BlinkIT grocery dataset analysis provides a comprehensive view of how product types, outlet characteristics, and customer preferences affect sales performance. Key insights reveal that larger outlets with higher product variety tend to outperform smaller outlets in both sales volume and customer engagement. Supermarket Type 1 outlets show stronger sales potential, making them ideal for business expansion.

Additionally, factors such as customer ratings play a significant role in driving sales. well-rated items consistently generate better sales, underscoring the importance of strategic product placement and customer satisfaction.

The analysis shows that a significant portion of the total sales for top-performing item types comes from the low fat content in Items. This indicates that customer prefer low content items.

To optimize business outcomes, retailers should focus on expanding medium and large outlets in urban locations, improving product assortment and visibility, and investing in customer satisfaction initiatives. By leveraging these insights, businesses can enhance their market position, increase sales, and ensure long-term profitability.

Actions To Be Taken:

Focus on Customer Satisfaction and Ratings:

Action: Enhance customer feedback mechanisms and improve product quality, focusing on top-rated items to encourage repeat purchases.

Reason: Higher customer ratings directly correlate with better sales, making customer satisfaction a key priority.

Risk/Consideration: Efforts to improve customer satisfaction could increase costs in the short term, but long-term loyalty and sales benefits justify the investment.

Promote Popular Categories in Marketing Campaigns:

Action: Focus marketing campaigns on high-demand categories like Frozen Foods and Fruits & Vegetables, and emphasize customer-rated top products in advertising.

Reason: These categories are strong performers, and promoting them could amplify sales growth, especially when tied to customer feedback.

Customer Feedback on Fat Content Preferences:

Action: Collect feedback from customers on their preferences for Low Fat vs. Regular products, and adjust inventory accordingly. Highlight top-rated products in each category.

Reason: Understanding which fat content options are preferred by the customer base will help optimize product stocking and improve sales. Top-rated products in each fat content category can be promoted to enhance customer satisfaction.