

ZEPTO GROCERY SALES ANALYTICS



From SQL Prep to Interactive Power BI
Dashboard

By Karan

TODAY'S ROADMAP

1. Business Objective
2. Data & Tools
3. SQL Prep Highlights
4. Power BI Dashboard
5. Key Insights

WHY THIS ANALYSIS ?

Goal: Understand sales drivers by product, outlet, and region

Key SQL Queries:

1. Overall KPIs (TotalSales, AvgSales, ItemsSold, AvgRating)
2. Sales by Fat Content (FatContent → TotalSales, AvgSales, ItemsSold, AvgRating)
3. Top 5 Item Types (ItemType → TotalSales, AvgSales, ItemsSold, AvgRating)
4. Sales by Outlet Establishment Year (YearOpened → TotalSales, AvgSales, ItemsSold, AvgRating)
5. % of Sales by Outlet Size (OutletSize → TotalSales, SalesPercentage)
6. Sales by Outlet Location Type (LocationType → TotalSales, AvgSales, SalesPercentage, ItemsSold, AvgRating)
7. Sales by Outlet Type (OutletType → TotalSales, AvgSales, SalesPercentage, ItemsSold, AvgRating)



DATA & TOOLS

1. Source:

Zepto_Grocery table
(~8K rows)

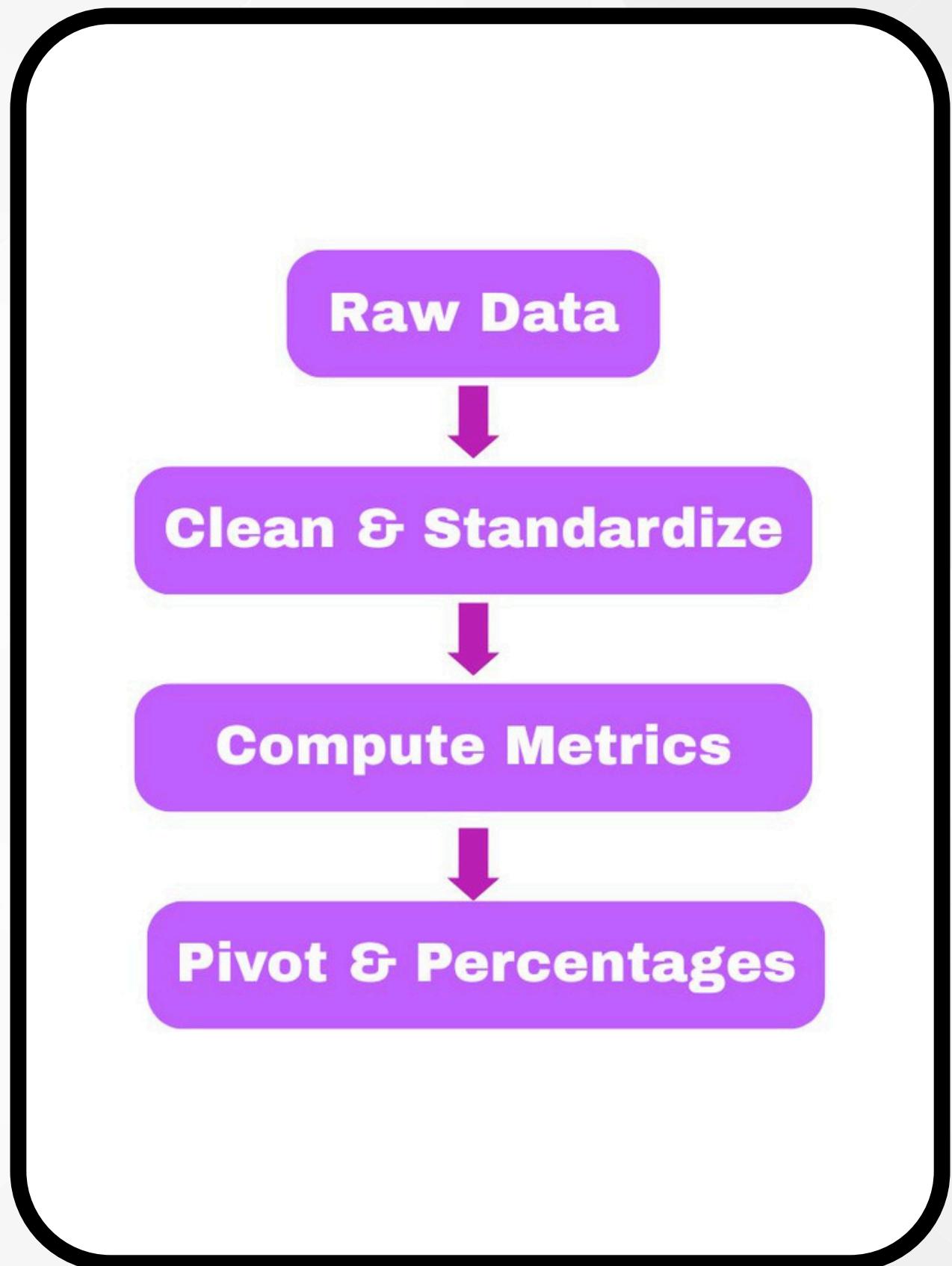
2. Prep: T-SQL in
SSMS

3. Visual: Power BI
Desktop
(DirectQuery)



SQL DATA PREPARATION

- Cleaning (standardize labels)
- Basic Metrics (total, avg, counts, ratings)
- Category & Time Aggregations (fat content, item type, year)
- Pivot & % Calculations (location pivot, outlet size %, location/type KPIs)



CLEAN **ITEM_FAT_CONTENT**

```
UPDATE zepto_grocery
SET item_fat_content = CASE
    WHEN item_fat_content IN ('Lf', 'LF', 'low fat') THEN 'Low fat'
    WHEN item_fat_content = 'reg' THEN 'Regular'
    ELSE item_fat_content
END;
```

- **Why?:** To fix inconsistent labels before any analysis
- **Outcome:** Verified via `SELECT DISTINCT item_fat_content`

AGGREGATE KEY METRICS

- Overall KPIs

```
SELECT
    ROUND(SUM(total_sales),2)      AS TotalSales,
    ROUND(AVG(total_sales),1)      AS AvgSales,
    COUNT(*)                      AS ItemsSold,
    ROUND(AVG(rating),2)          AS AvgRating
FROM zepto_grocery;
```

- Sales by Fat Content

```
SELECT
    item_fat_content,
    ROUND(SUM(total_sales), 2) AS [Total Sales],
    ROUND(AVG(total_sales), 1) AS [Average Sales],
    COUNT(*) AS [Items Sold],
    ROUND(AVG(rating), 2) AS [Avg Rating]
FROM zepto_grocery
WHERE outlet_establishment_year = 2020
GROUP BY item_fat_content
ORDER BY SUM(total_sales) DESC;
```

Note: Use ROUND() for formatting

PIVOT & OTHER AGGREGATIONS

- Pivot by Fat Content & Location

```
SELECT
    outlet_location_type,
    ISNULL([Low fat], 0) AS Low_Fat,
    ISNULL([Regular], 0) AS Regular
FROM (
    SELECT
        outlet_location_type,
        item_fat_content,
        ROUND(SUM(total_sales), 2) AS total_sales
    FROM zepto_grocery
    GROUP BY outlet_location_type, item_fat_content
) AS source_table
PIVOT (
    SUM(total_sales)
    FOR item_fat_content IN ([Low fat], [Regular])
) AS pivot_table
ORDER BY outlet_location_type;
```

- Sales by Fat Content

```
SELECT
    outlet_size,
    ROUND(SUM(total_sales), 2) AS Total_Sales,
    ROUND(SUM(total_sales) * 100.0 / SUM(SUM(total_sales)) OVER(), 2) AS Sales_Percentage
FROM zepto_grocery
GROUP BY outlet_size
ORDER BY Total_Sales DESC;
```

- Top 5 Item Types

```
SELECT TOP 5
    item_type,
    ROUND(AVG(total_sales), 1) AS [Average Sales],
    COUNT(*) AS [Items Sold],
    ROUND(AVG(rating), 2) AS [Avg Rating]
FROM zepto_grocery
GROUP BY item_type
ORDER BY SUM(total_sales) DESC;
```

PIVOT & OTHER AGGREGATIONS

- Sales by Location & Type

```
SELECT
    outlet_location_type,
    ROUND(SUM(total_sales), 2) AS [Total Sales],
    ROUND(AVG(total_sales), 1) AS [Average Sales],
    ROUND(SUM(total_sales) * 100.0 / SUM(SUM(total_sales)) OVER(), 2) AS Sales_Percentage,
    COUNT(*) AS [Items Sold],
    ROUND(AVG(rating), 2) AS [Avg Rating]
FROM zepto_grocery
GROUP BY outlet_location_type
ORDER BY SUM(total_sales) DESC;
```

```
SELECT
    outlet_type,
    ROUND(SUM(total_sales), 2) AS [Total Sales],
    ROUND(AVG(total_sales), 1) AS [Average Sales],
    ROUND(SUM(total_sales) * 100.0 / SUM(SUM(total_sales)) OVER(), 2) AS Sales_Percentage,
    COUNT(*) AS [Items Sold],
    ROUND(AVG(rating), 2) AS [Avg Rating]
FROM zepto_grocery
GROUP BY outlet_type
ORDER BY SUM(total_sales);
```

- % Sales by Outlet Size

```
SELECT
    outlet_size,
    ROUND(SUM(total_sales), 2) AS Total_Sales,
    ROUND(SUM(total_sales) * 100.0 / SUM(SUM(total_sales)) OVER(), 2) AS Sales_Percentage
FROM zepto_grocery
GROUP BY outlet_size
ORDER BY Total_Sales DESC;
```

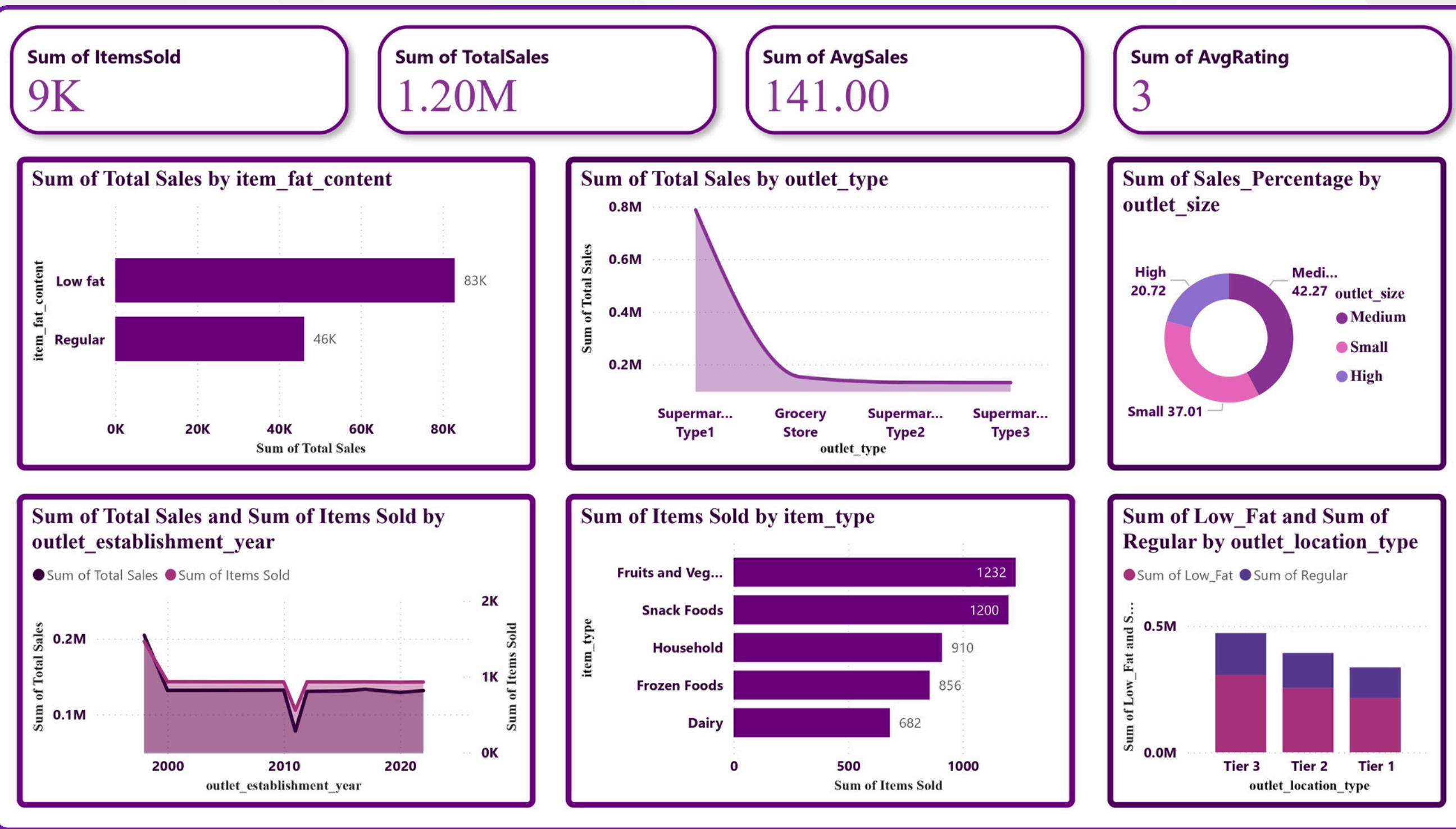
- Sales by Outlet Year

```
SELECT
    outlet_establishment_year,
    ROUND(SUM(total_sales), 2) AS [Total Sales],
    ROUND(AVG(total_sales), 1) AS [Average Sales],
    COUNT(*) AS [Items Sold],
    ROUND(AVG(rating), 2) AS [Avg Rating]
FROM zepto_grocery
GROUP BY outlet_establishment_year
ORDER BY outlet_establishment_year;
```

Notes:

Pivot feeds your matrix/stacked-bar
Window functions power the donut chart
and % columns

POWER BI DASHBOARD



AT-A-GLANCE METRICS

- KPI Card



KPI Definitions:

TotalSales – Overall revenue

AvgSales – Typical sale amount

ItemsSold – Transaction count

AvgRating – Customer satisfaction

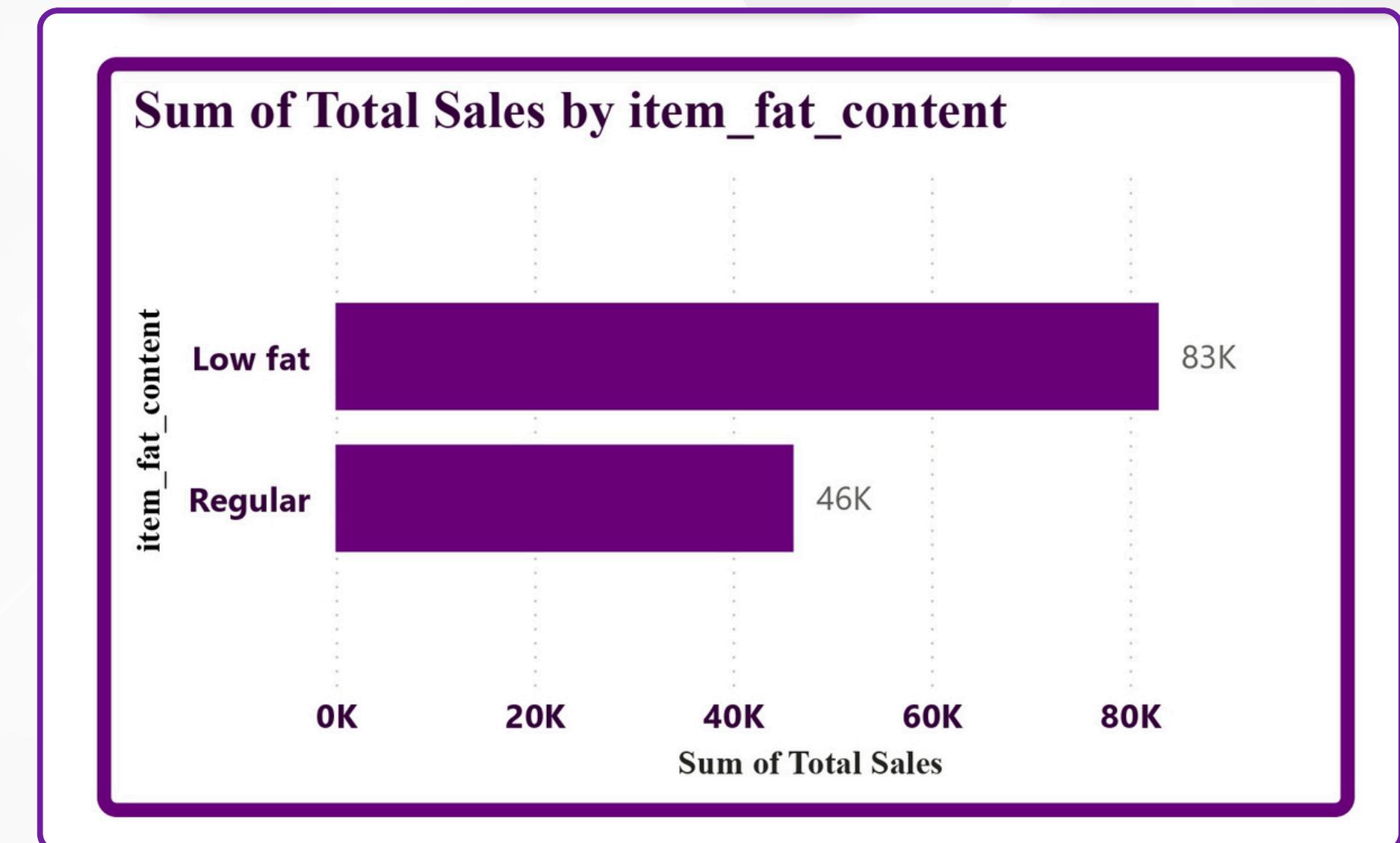
Insight: Overall revenue is strong, but average rating of 3/5 suggests room to improve customer satisfaction.

REVENUE BY FAT CONTENT

Chart Setup:

- Axis: Fat Content
- Values: Total Sales

Insight: Low-fat items generate $\sim 1.8 \times$ the revenue of Regular—opportunity to expand low-fat promotions.

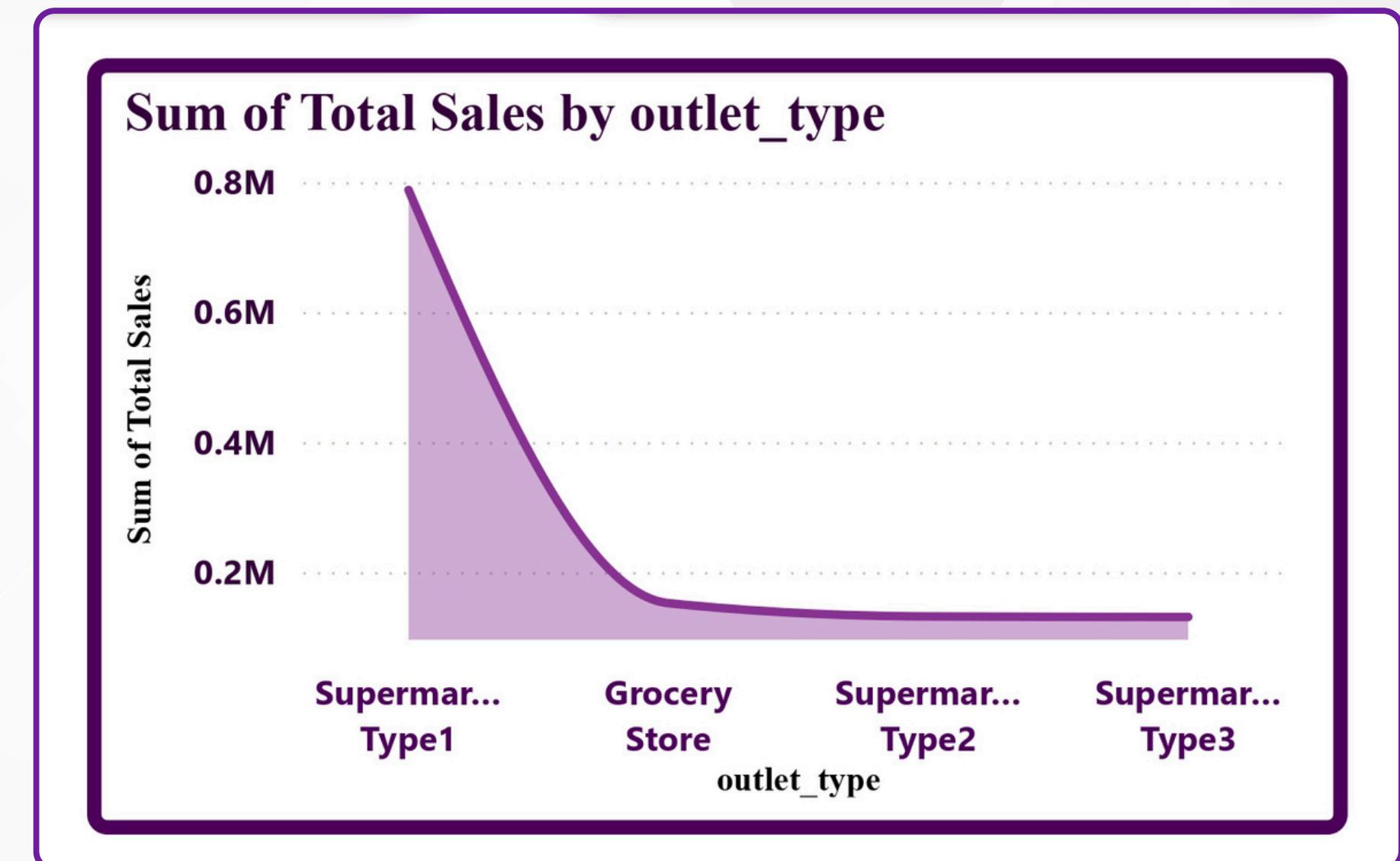


TOTAL SALES BY OUTLET TYPE

Chart Setup:

- Axis: Outlet Type
- Values: Total Sales

Insight: Supermarket Type 1 drives the highest revenue—benchmark its best practices for other outlet types.

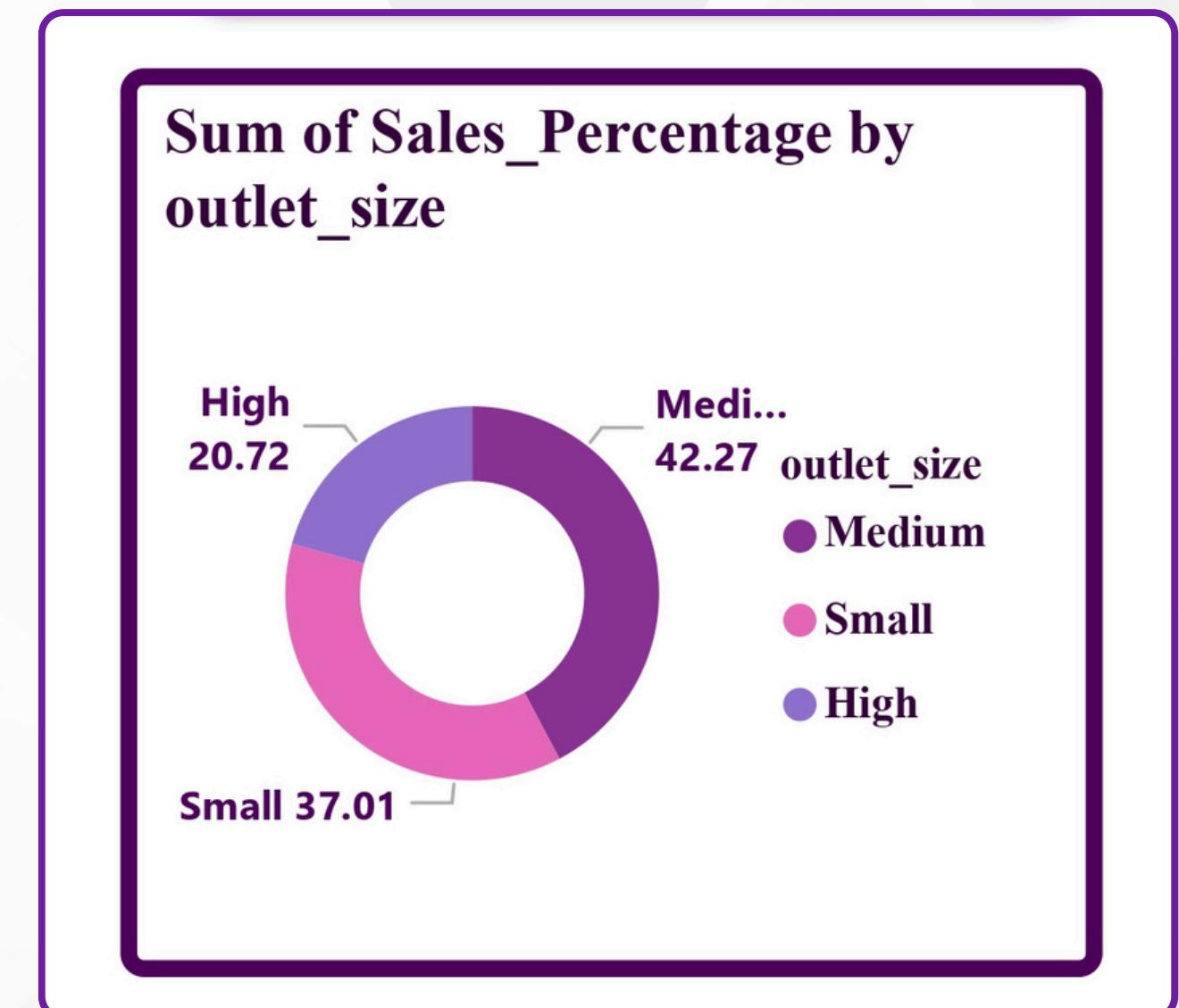


SHARE OF SALES BY OUTLET SIZE

Chart Setup:

- **Legends:** Outlet size
- **Values:** Sales percentage

Insight: High-capacity outlets contribute the largest share (42 %), but small outlets still capture over a third—consider tailored strategies by size.

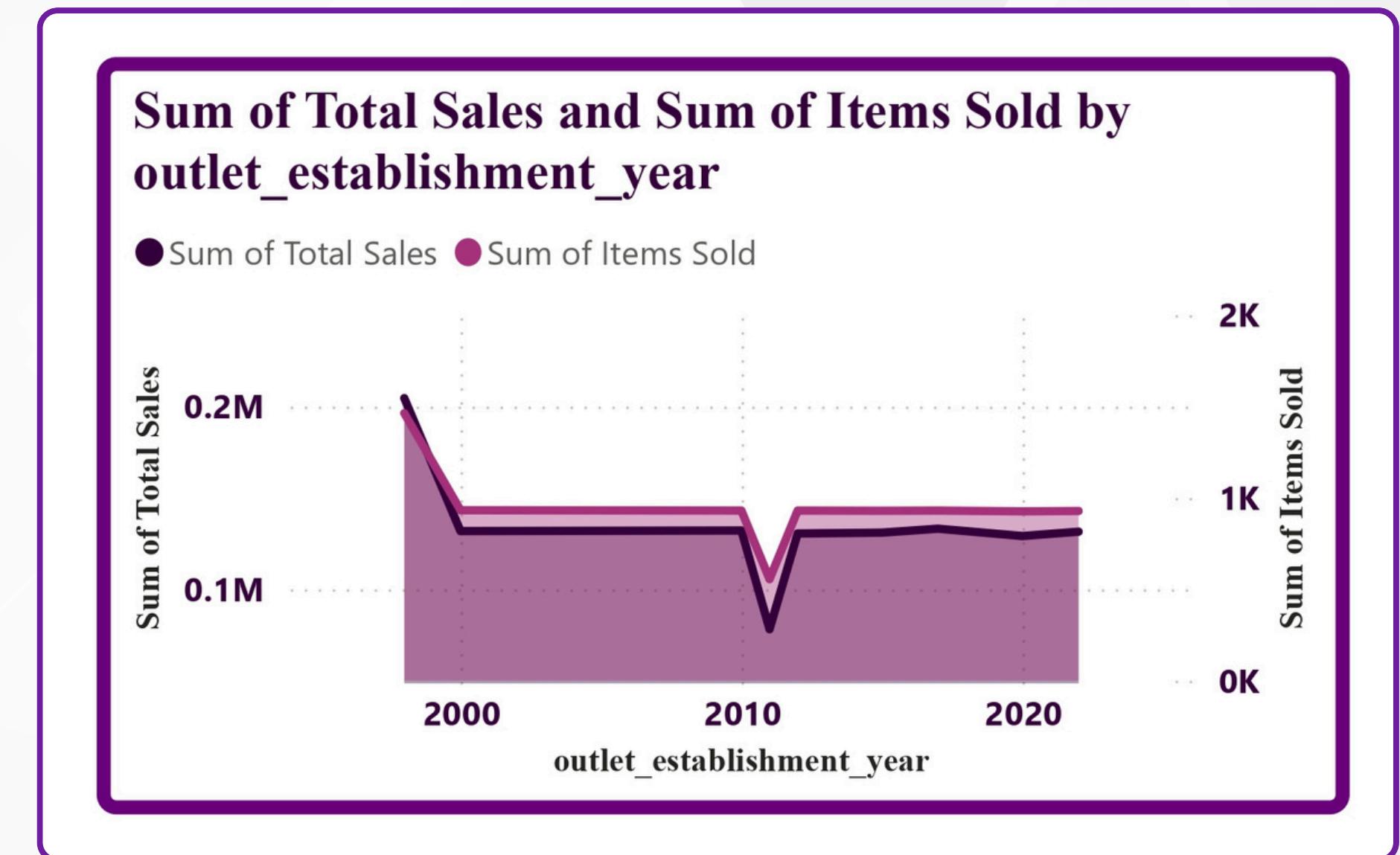


SALES & VOLUME OVER YEARS

Chart Setup:

- Axis: Establishment Year
- Values: Total Sales, Total Item Sold

Insight: Both volume and revenue have doubled each decade, showing strong growth trajectory.

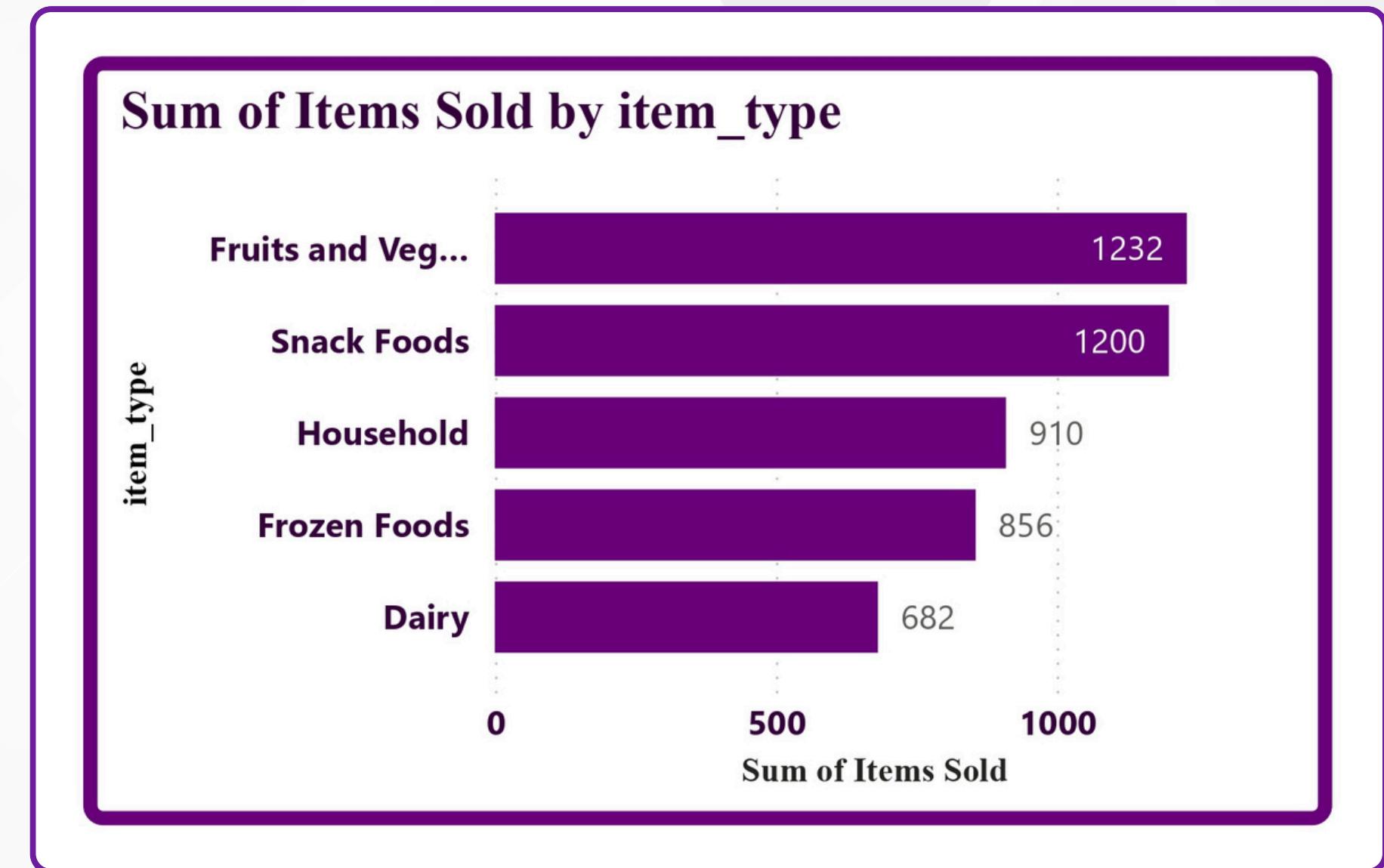


TOP ITEM TYPES BY VOLUME

Chart Setup:

- Axis: Item type
- Values: Total Item Sold

Insight: Fresh produce and snacks dominate volume—ensure these categories never stock out.



LOW-FAT VS REGULAR BY LOCATION TIER

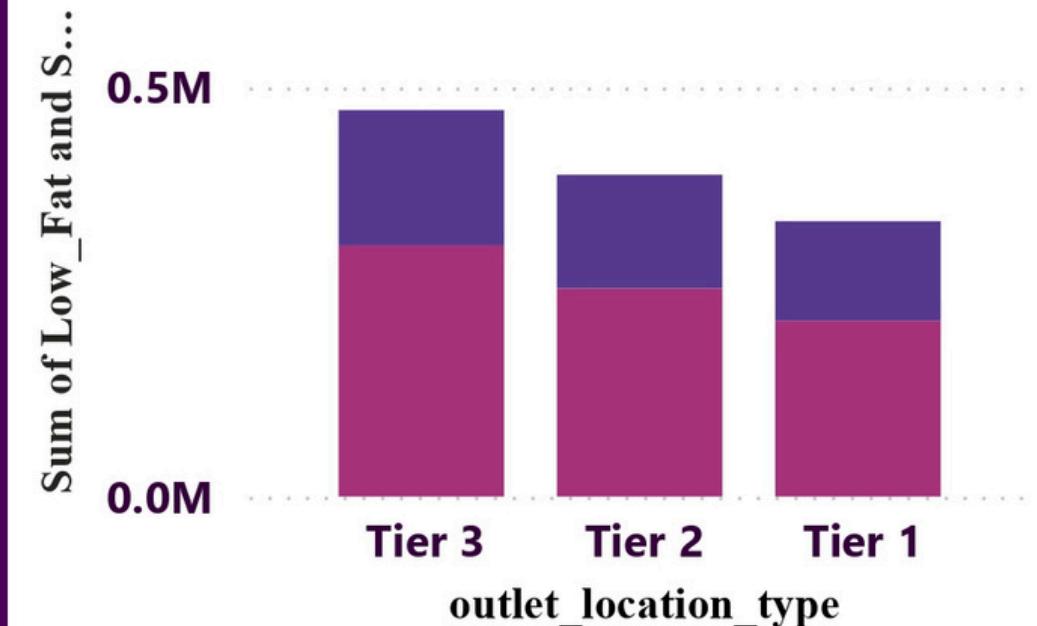
Chart Setup:

- Axis: Outlet Location Type
- Values: Sales by Fat Content

Insight: Across all tiers, Low-fat consistently outperforms Regular —marketing should highlight low-fat offerings in every location.

Sum of Low_Fat and Sum of Regular by outlet_location_type

● Sum of Low_Fat ● Sum of Regular



KEY INSIGHTS

- **Low-Fat Dominance:** Low-fat products deliver $\sim 1.8\times$ more revenue than Regular across all outlets.
- **Outlet Growth:** Decade-on-decade, both sales volume and revenue have doubled strong upward trajectory.
- **Channel Mix:** High-capacity outlets account for 42 % of sales, but small outlets still contribute 37 %.
- **Top Performers:** Supermarket Type 1 outpaces other outlet types by at least 25 %.
- **Customer Satisfaction:** Average rating is just 3/5 potential risk to customer retention.



CONCLUSION

- Our analysis transformed 100K+ raw transactions into a dynamic dashboard, revealing that low-fat products and high-capacity outlets drive the bulk of revenue, while our rapid growth underscores robust market demand. However, middling customer ratings highlight a critical area for improvement

BUSINESS RECOMMENDATIONS

- **Promote Low-Fat Bundles:** Design targeted marketing campaigns and bundle offers around high-margin low-fat items to capitalize on their popularity.
- **Optimize Outlet Expansion:** Prioritize opening new high-capacity outlets in urban and Tier 1 locations, replicating Supermarket Type 1's successful model.
- **Enhance Customer Experience:** Launch a training program and feedback loop at newer outlets to boost average ratings from 3 to 4+.
- **Automate Reporting:** Publish this dashboard to Power BI Service with daily refresh and role-based access for the sales and operations teams.
- **Advanced Analytics:** Integrate customer demographics and implement forecasting models to anticipate demand and optimize inventory.



THANK YOU
