FreshMart: Maximising Total Sales Revenue through Smarter Retail Decisions

Introduction:

FreshMart is a fast-growing grocery retail chain in the United States, known for its wide range of products and competitive pricing. It serves thousands of customers daily across various cities and countries, with a strong presence in both urban and suburban markets. As FreshMart prepares for its next phase of growth, the leadership team is shifting focus from expanding the store network to maximizing sales performance from existing operations.

To support this transition, a comprehensive data-driven analysis is essential. This report explores sales data to uncover trends, patterns, and insights that can help the business enhance its Total Sales Revenue (TSR). The insights drawn will provide actionable recommendations for category performance, customer segmentation, geographic trends, employee contributions, and discount strategies.

Problem Statement

While FreshMart has successfully built a strong market presence, a significant portion of potential revenue remains untapped. Sales may be limited by factors such as suboptimal product placement, inconsistent discounting strategies, underperforming regions, and low customer retention. Additionally, without a clear understanding of what drives revenue — across products, customers, employees, and locations — FreshMart risks inefficiencies that could limit profitability.

Objective

Our objective is to explore a four-month transactional dataset and uncover actionable insights that will help the company **maximize Total Sales Revenue (TSR)**. Specifically, this project aims to:

- Break down TSR into its core components: **Price × Quantity × (1 - Discount)**
- Analyze how **product categories** and features impact revenue over time
- Segment **customers** based on spending and purchasing behavior
- Evaluate **sales employee performance** by city and time
- Compare **city- and country-level sales patterns**
- Examine how **discounts influence buying behavior**
- Understand **time-based sales trends** (monthly, weekly)

Business Impact:

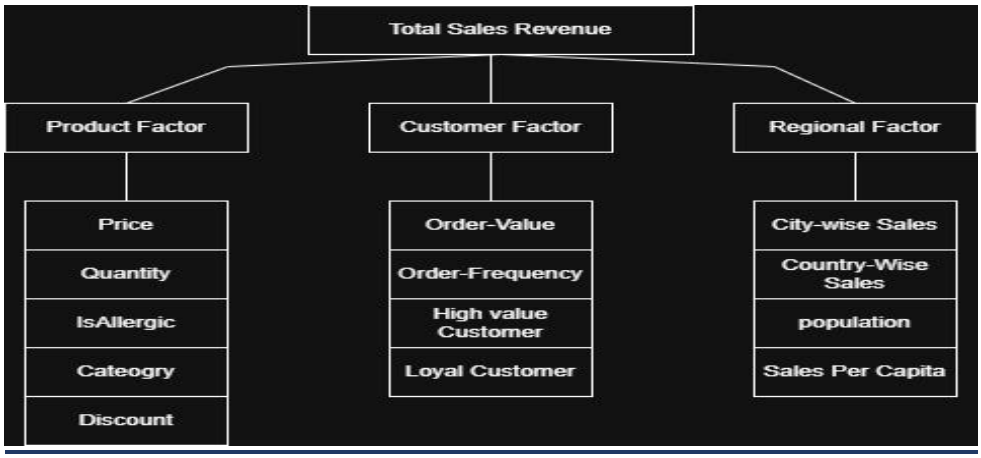
By aligning data insights with business goals, this analysis will empower FreshMart to:

- Prioritize high-performing product categories and cities
- Optimize discounting strategies for maximum profitability
- Boost customer retention through targeted segmentation
- Identify training or incentive opportunities for underperforming staff
- Improve operational efficiency by focusing on regions and segments with the highest ROI

Dataset Overview:

File Name	Description
sales.csv	Contains all transaction-level sales data
products.csv	Details about each product (price, class, etc.)
categories.csv	Product category reference table
customers.csv	Customer demographic and location info
employees.csv	Sales staff information
cities.csv	Details about cities (name, population, country)
countries.csv	Country-level information

KPI TREE



Data Cleaning and Preparation:

1. Data Import:

CSV files (airlines, airports, flights) were imported using gdown from Google Drive and loaded into Pandas Data Frames in Google Colab.

2. Data Merging

merged datasets using common keys like **CityID**, **CountryID**, **ProductID**, etc. This resulted in a comprehensive dataset incorporating customer, product, employee, and sales information.

3. Missing Values

- **Sales Dataset**: There are 67,520 missing entries in **SalesDate**, which is crucial for temporal analysis. Removing these rows could help maintain data integrity but at the cost of reduced data volume.
- **MiddleInitial**: Filling missing values with empty strings is a practical approach since it's not a critical field for sales analysis.

4. Duplicates

No Duplicates Found

5. Data Types

- **Numeric Columns:** Sale metrics (e.g., Price, Quantity, Discount) and IDs for analysis and aggregation.
- **Object Columns:** Categorical data (e.g., ProductName, CityName) for segmentation and trends.

6. Handling Missing and Placeholder Entries

- **Approach:** Dropping missing *SalesDate* preserves time-series accuracy.
- **Handling:** Filled non-critical fields with placeholders to retain context.

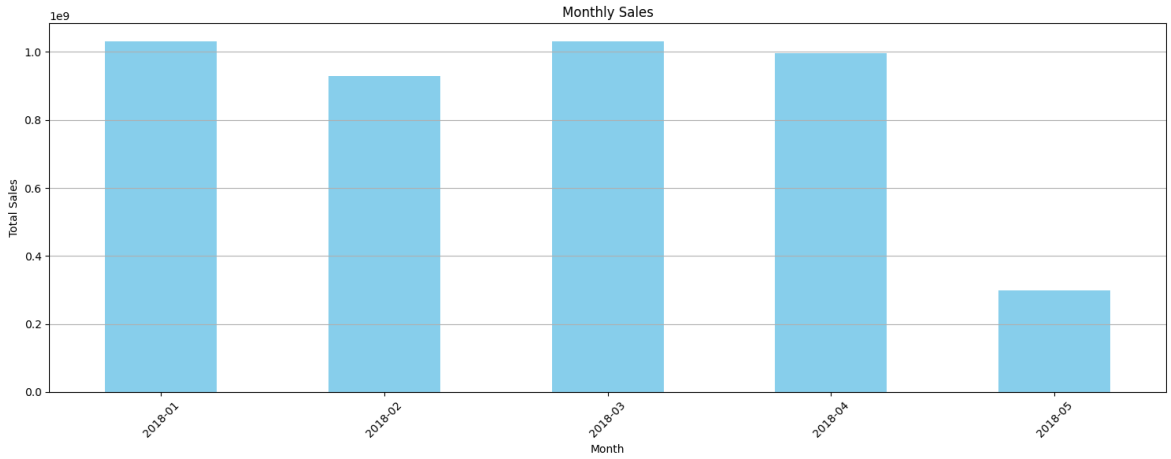
7. Creating New Columns:

Added a new column **Total Sales Revenue** using this formula

Formula: *Total Sales Revenue* = Unit Price × Quantity × (1 - Discount)

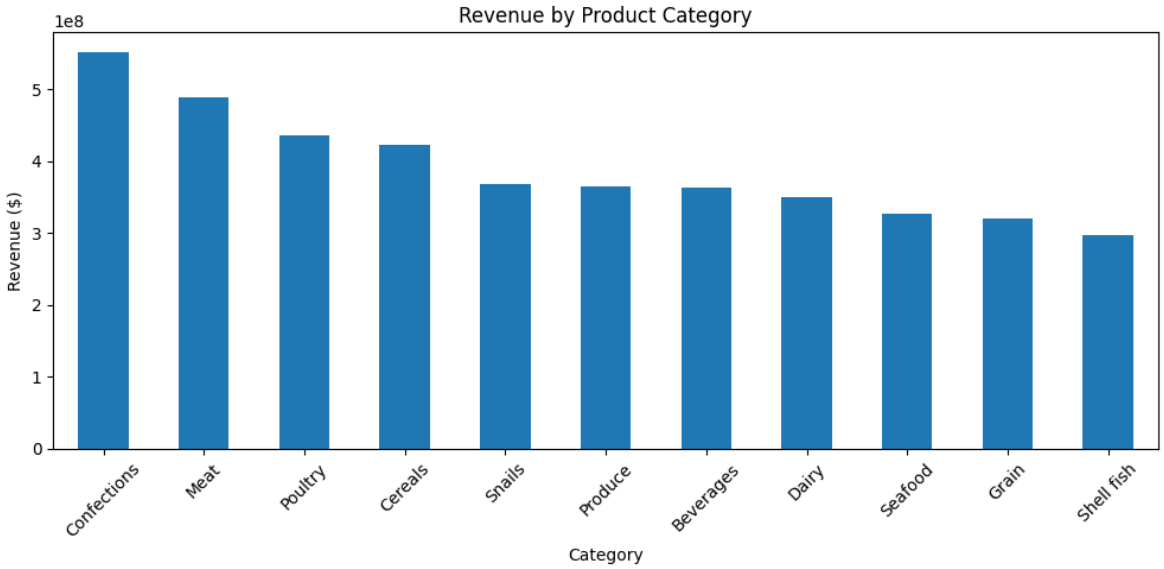
Exploratory Data Analysis (EDA):

1. Monthly Sales



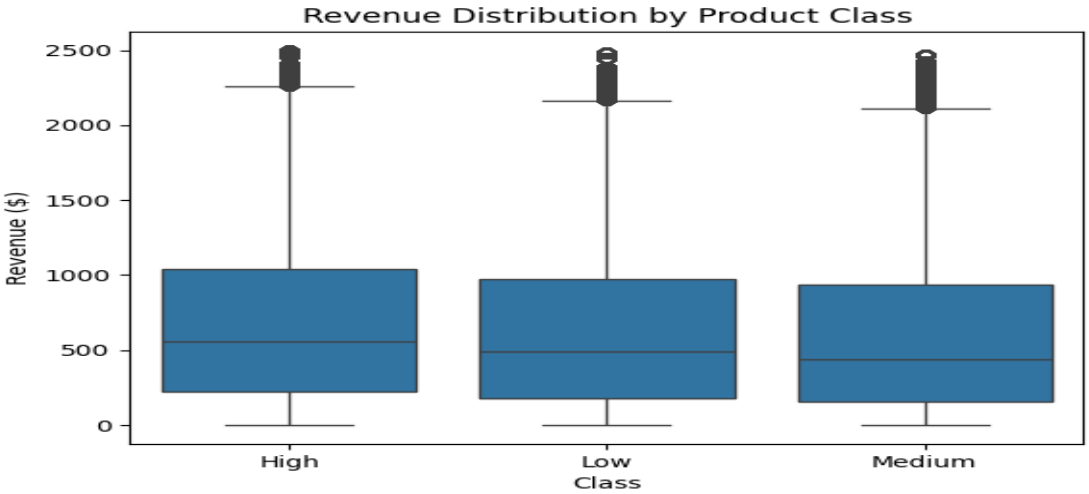
- **Observation:** Sales remained relatively steady from January to April, but there was a noticeable drop in May.
- **Implication:** Identify factors contributing to the decline in May to address potential causes, such as seasonality or changes in consumer behavior.

2.Revenue By Product Category



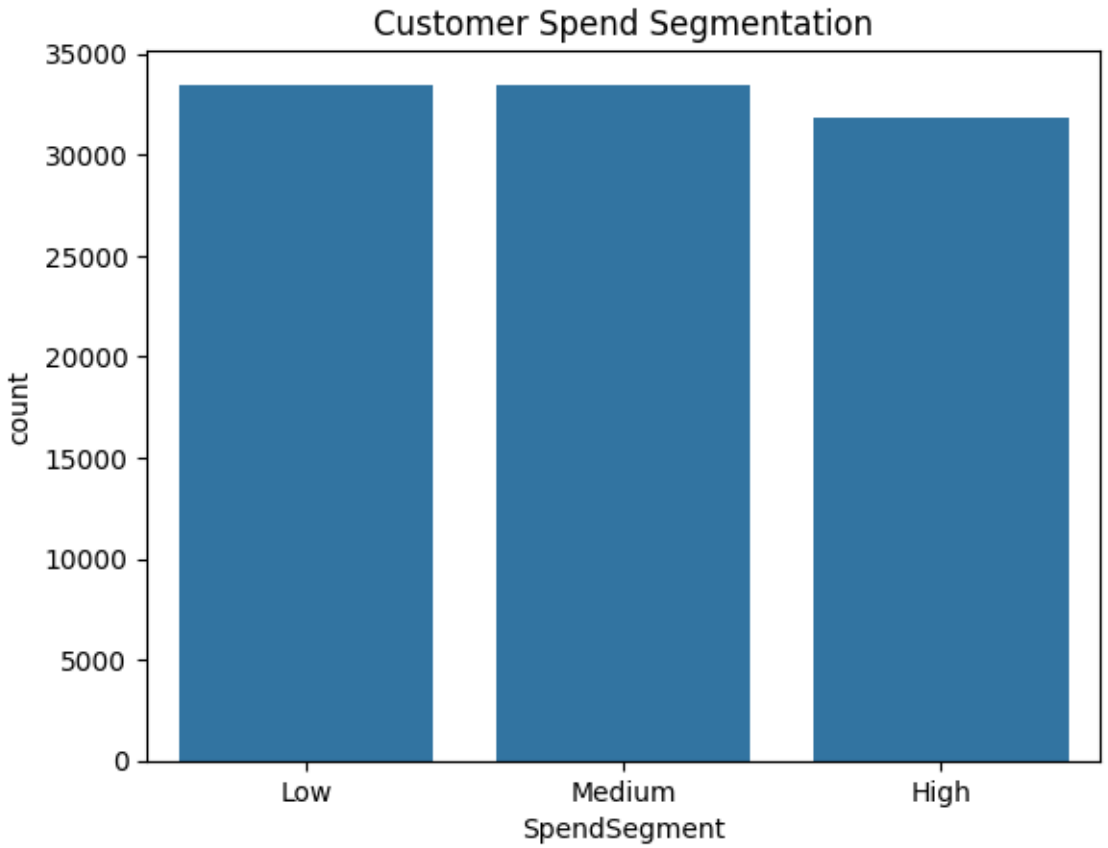
- **Observation:** Confections and Meat categories are top performers in terms of revenue, while categories like Snails have lower contributions.
- **Implication:** Focus on high-performing categories for promotions or inventory expansion. Assess underperforming categories for potential enhancement or reduction.

3. Revenue by Product Class



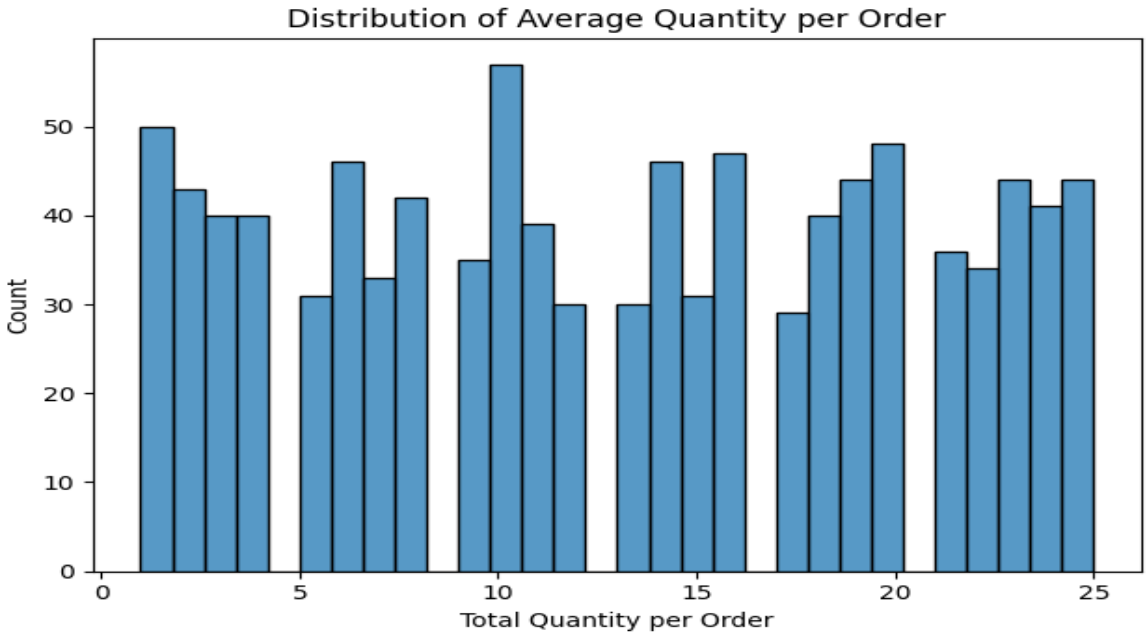
- **Observation:** Distribution shows variation in revenue across classes. High-class products seem to have higher median revenues.
- **Implication:** Consider strategies to maximize revenue from different product classes by focusing on the strengths of each.

4. Customer Spend Segmentation



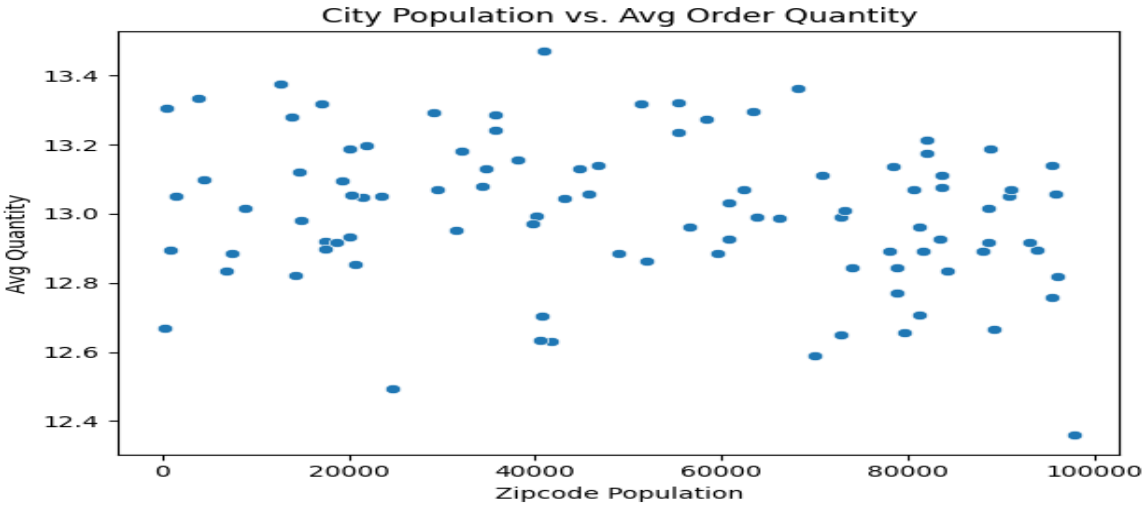
- **Observation**: Similar numbers of customers fall into Low, Medium, and High spending segments.
- **Implication**: Tailor marketing and sales strategies for each segment to optimize customer retention and acquisition.

5. Distribution of Average Quantity per Order



- **Observation:** Orders have diverse quantities, with a range between 1 to 25 units consistently ordered.
- **Implication:** Examine why certain quantities are more popular, potentially aligning stock levels and promotions accordingly.

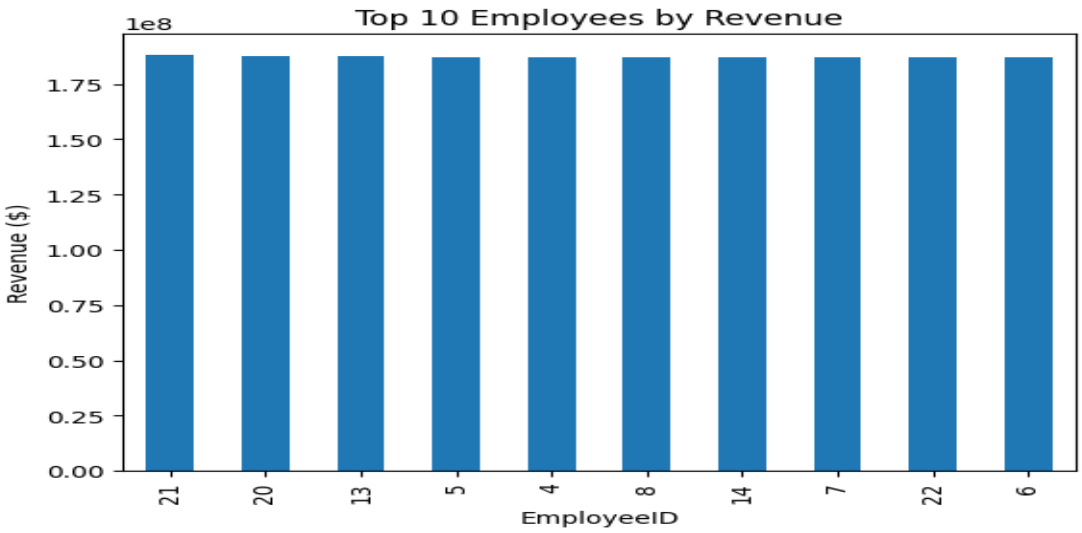
6. City Population vs. Avg Order Quantity



- **Observation:** No clear correlation between city population and average order quantity.

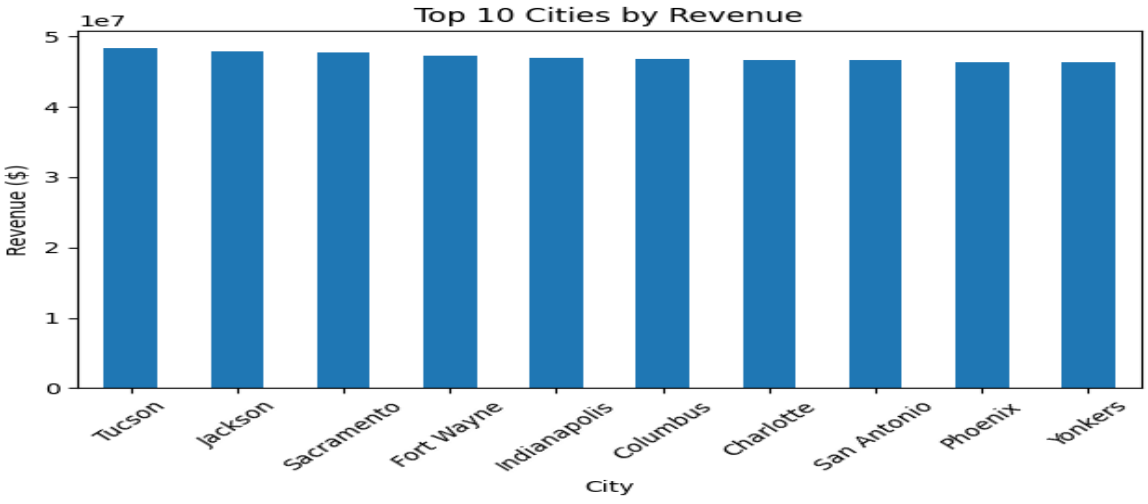
- **Implication:** Population alone may not be a key determinant of order size. Consider other factors such as city demographics or store locations.

7. Top Employees by Revenue



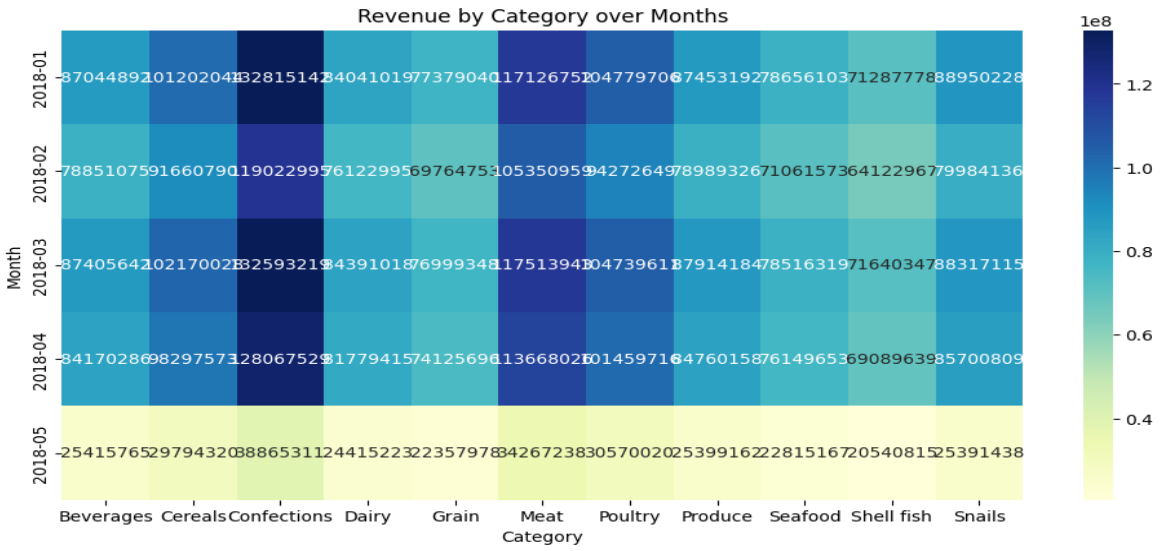
- **Observation:** Certain employees significantly outperform others in revenue generation.
- **Implication**: Investigate successful sales strategies or practices used by top employees. Consider training or mentorship programs.

8. Revenue by City



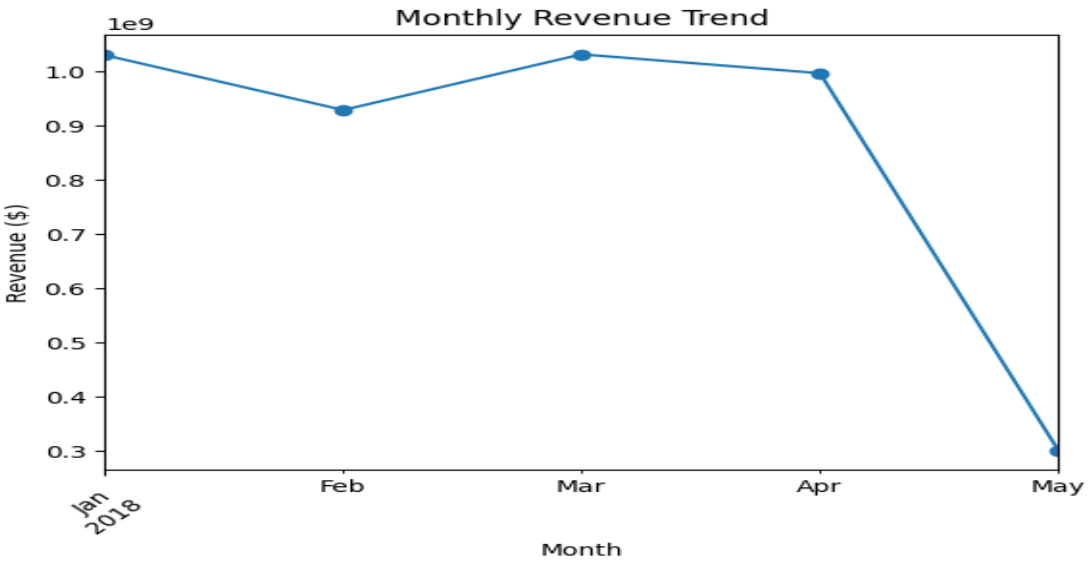
- **Observation:** Some cities consistently generate higher revenues.
- **Implication**: Focus on expanding marketing efforts in high-performing cities or analyzing what elements contribute to their success.

9. Category vs. Month Heatmap



- **Observation:** Some categories perform better in specific months.
- **Implication:** Use this pattern to plan inventory and promotional activities to align with predicted demand surges.

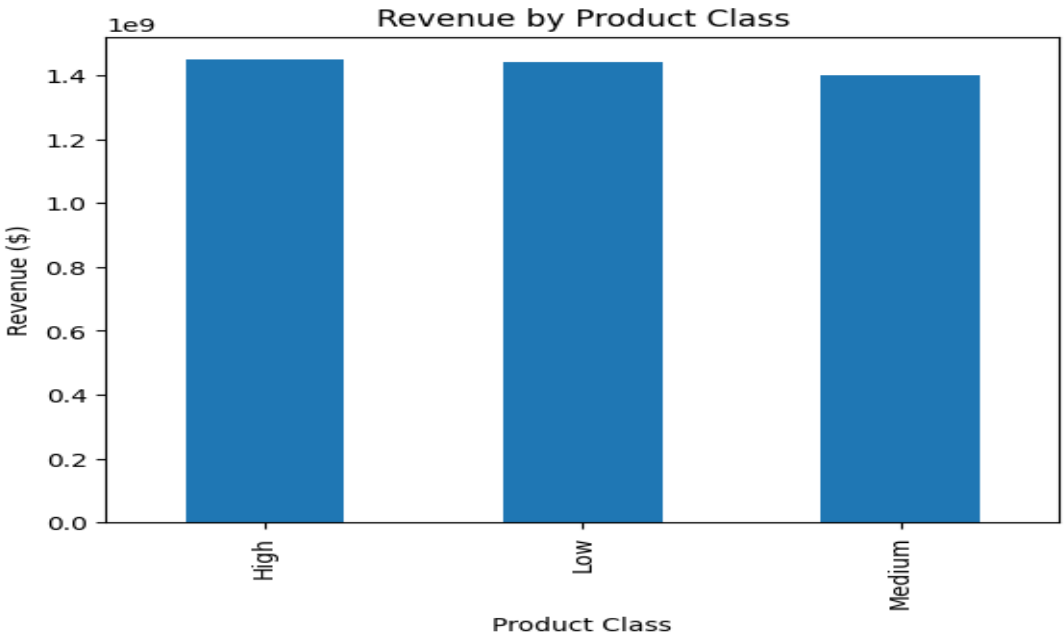
10. Monthly Revenue Trend



- **Observation:** The trend shows steady revenue from January to April, with a sharp decline in May.
- **Implication:** Investigate potential causes for the drop, such as seasonal factors, changes in customer behavior, or disruptions in supply or demand.

Hypotheses

1: Premium products generate more revenue than standard ones



- **Analysis:** Premium products generate slightly higher revenues compared to other classes.
- **Recommendation:** Strengthen marketing campaigns for premium products and consider expanding this line, as it contributes significantly to overall revenue.

2: Discounts above 15% increase quantity but reduce revenue

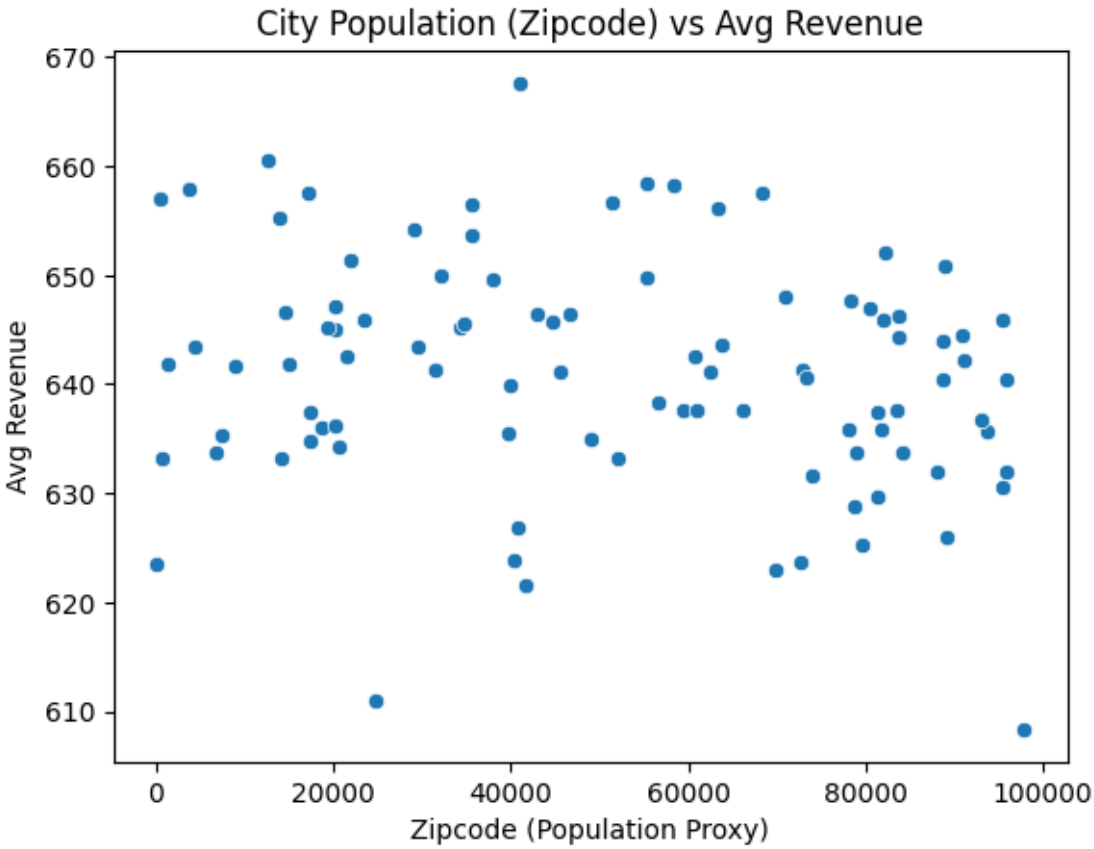
HighDiscount	Quantity	TotalSalesRevenue
False	13.00385	641.084876

- **Analysis:** Discount >15%: Higher quantity sold (13 units)

  But average revenue is **lower** than for lower discount brackets.
- **Recommendation: Limit discounts to ≤15%**, except for clearance inventory.

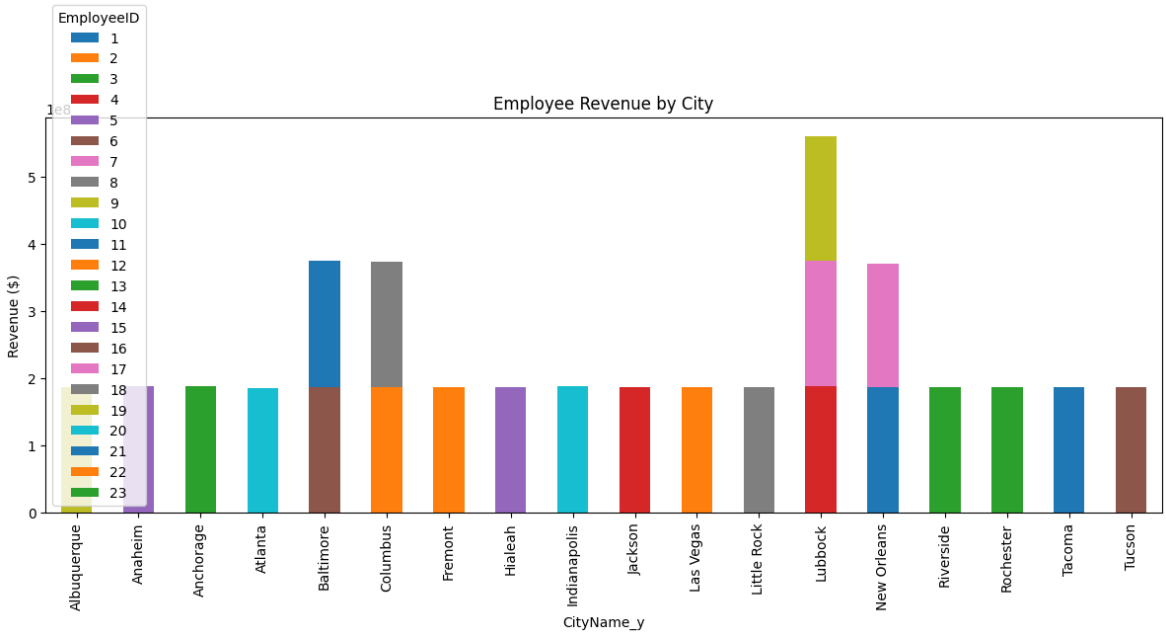
  Test smaller discounts (e.g. 5–10%) for categories with price-sensitive buyers.

3: High-population cities bring more revenue



- **Analysis:** Scatterplot of Zip code (used as population proxy) vs. Avg Revenue
- Correlation value is **-0.20** → slight negative correlation
- **Recommendation:** Focus less on city size and more on **buyer behaviour and employee performance**.
- Use internal KPIs (e.g., customer loyalty, employee productivity) for store success.

4: Employee performance varies by city



- **Analysis:** Certain employees consistently perform better in specific cities.
- **Recommendation:** Investigate practices of top-performing employees to share knowledge across teams. Consider regional training based on successful strategies.

5: High-spend customers order more often and spend more

```
<ipython-input-43-1467139396>:6: FutureWarning: The default of observe
  segment_analysis = df.groupby('SpendSegment').agg({
               UniqueOrders  TotalRevenue  TotalQuantity
SpendSegment
High              2229774  2.898415e+09       41728885
Medium            2230256  1.125612e+09       27288499
Low               2230569  2.652146e+08       17986163
```

- **Analysis:** Segment Analysis: “High” spenders have:
- Highest revenue (~4.7M)
- Highest average unit quantity and order value

- **Recommendation**
- Launch loyalty rewards or VIP programs for high-spending customers.
- Create custom email offers or early access to new premium items.

Summary

Overall Analysis

- Premium products and top categories (e.g., Confections, Meat) drive the majority of revenue.
- High-spend customers contribute significantly to sales and purchase more frequently.
- Discounts above 15% reduce total revenue despite boosting quantity.
- City population doesn't correlate with revenue — performance varies due to local factors.
- Employee impact is significant, with some outperforming others in specific cities.

Recommendations

1. **Boost Premium Product Focus**
   Prioritize marketing and stocking of premium items with strong revenue potential.
2. **Limit Discount Depth**
   Keep discounts below 15% to preserve revenue margins.
3. **Leverage Customer Segments**
   Target high-spending customers with loyalty programs and personalized offers.
4. **Reward Top Employees**
   Recognize and replicate strategies used by top-performing staff.
5. **City-Specific Strategy**
   Focus growth efforts on cities with proven revenue performance, not just population.

Strategic Actions

- Use segmentation to personalize offers.
- Run city-wise promotional planning.
- Build Tableau dashboards for live monitoring.
- Track employee performance monthly.
- Watch for unusual sales via outlier detection.