**Retail Pulse**

Parag Shrikant Tippannawar

Nuel Inc.

January 25, 2025

**Retail Pulse**

**Problem Definition:**

Retail businesses face challenges in understanding and addressing fluctuations in sales caused by a variety of factors, including customer demographics, weather, holidays, and product performance. Without actionable insights into these factors, it becomes difficult to optimize marketing strategies, manage inventory efficiently, and Improve overall profitability.

**Objectives:**

1. Analyze sales trends across categories, cities, and customer segments.
2. Assess the impact of holidays and weather conditions on sales performance.
3. Identify low-performing stores and products to inform decision-making.
4. Provide actionable recommendations to improve overall business efficiency and profitability.

**Assumptions:**

1. External factors such as holidays and weather influence customer purchasing behavior significantly.
2. The datasets provided (sales, holidays, and weather) are accurate, complete, and representative of real-world scenarios.
3. Weather data retrieved via the API aligns accurately with the cities and dates in the sales dataset.

**Retail Pulse**

To address these challenges, the problem is broken down into the following objectives:

1. Primary Factors Influencing Sales:
   * What categories, regions, and time periods contribute most significantly to sales performance?
   * Which products and stores are underperforming, and why?
2. Customer Demographics and Behavior:
   * How do customer demographics and spending behavior vary across different locations?
   * What are the key characteristics of high-value customers compared to low-value customers?
3. External Factors (Weather and Holidays):
   * How do weather conditions (e.g., precipitation, temperature) impact customer purchasing behavior?
   * What is the influence of public holidays on sales performance, and which holidays generate the highest sales?
4. Key Performance Indicators (KPIs):
   * Revenue growth across categories, time periods, and regions.
   * Sales distribution by category, region, and weather conditions.
   * Customer retention and segmentation by spending behavior.
   * Contribution of high-performing products to overall sales.

**Retail Pulse**

**Description of the Datasets Used and Their Relevance**

1. **Sales Dataset**:
   * Contains detailed transactional data, including order dates, customer information, sales figures, and product details.
   * Relevance: Forms the core dataset to analyze trends, segment customers, and assess product/store performance. Supplemented with **coordinates data** from the OpenCage API to enable geographic analysis and visualizations, such as mapping sales by city.
2. **Holiday Dataset**:
   * Contains information on public holidays across regions, including holiday names and dates.
   * Relevance: Used to correlate sales spikes with holiday periods and evaluate holiday-specific purchasing behavior.
3. **Weather Dataset**:
   * Retrieved via Visual Crossing API, providing weather attributes such as temperature and precipitation for cities.
   * Relevance: Helps determine the impact of weather conditions on sales patterns and customer activity.

**Assumptions:**

1. The holiday dataset includes all major holidays relevant to the analysis period.
2. Attributing holiday sales spikes entirely to holiday impacts.

**Retail Pulse**

**Challenges:**

1. **Data Integration**:
   * Standardizing city names and date formats to ensure proper merging of datasets.
   * Managing missing weather data for certain cities.
2. **Sparse Data**:
   * After executing a V-Lookup to check manually some cities lacked sufficient weather data, limiting insights for those locations, and were excluded for data accuracy.
3. **Opportunities for Further Analysis:**
   * Low-performing stores require further investigation to determine specific causes of underperformance (e.g., location, demographics, or operational inefficiencies).
   * Certain weather-related patterns may need more granular data (e.g., daily weather) for better insights.
4. **API Limitations**:
   * Daily API call limits necessitated a phased data collection approach, extending project timelines working with a trial API account further restricted access and limited the volume of data that could be retrieved, requiring careful planning and prioritization of requests

**Summary of Key Findings and Recommendations**

**Key Findings:**

1. **Sales Trends**:
   * Monthly sales exhibited consistent growth across the analysis period, with spikes observed during major holidays.
   * Technology and Furniture consistently outperformed Office Supplies in terms of revenue generation.
   * Products with low sales included niche items like "Eureka Disposable Bags".
   * Burlington and Arlington exhibited the highest sales, exceeding $20,000.
   * Cities like Chapel Hill and Antioch recorded sales below $20, indicating potential issues with market penetration or operational inefficiencies.
2. **Holiday Impact**:
   * Sales increased significantly during holidays such as Thanksgiving and Christmas.
   * On average, holiday sales were 20% higher compared to non-holiday periods.
3. **Weather Impact**:
   * Rainy weather had a slightly negative correlation with total sales, indicating decreased customer activity during adverse weather.
4. **Customer Segmentation**:
   * Medium spenders dominate the customer base, while high-value customers, though smaller in number, significantly impact total revenue.
   * Customer spending patterns vary across cities, with urban areas demonstrating higher average sales per customer compared to rural regions.
5. **KPIs for Stores and Products**:
   * In the context of revenue growth, monthly sales data displayed a consistent upward trend, with significant peaks during holiday months.
   * Observing customer retention, high-value customers exhibited strong repeat purchase behavior.
   * Product contribution towards sales, specific product categories (e.g., laptops and office chairs) accounted for the majority of sales within their respective categories.

**Business Recommendations:**

1. **Data Refinement**:
   * Standardize city and date formats in raw data to simplify future data integrations.
   * Address missing weather data by exploring alternative APIs or interpolating missing values.
2. **Scalability**:
   * Use higher API limits or automated batch processing for large-scale data retrieval to optimize project timelines.
3. **Data Augmentation**:

* Incorporate additional datasets, such as promotional campaigns and local events, to capture a more holistic view of sales drivers.

**Technical Recommendations:**

1. **Promotional Strategies**:
   * Target marketing campaigns and discounts during high-impact holidays to maximize sales.
   * Increase focus on high-value customer segments through loyalty programs or exclusive promotions.
2. **Inventory Optimization**:
   * Reduce stock for low-performing products and reallocate inventory toward high-demand items
3. **Data Augmentation**:
   * Incorporate additional datasets, such as promotional campaigns and local events, to capture a more holistic view of sales drivers.