**Project Report: Retail Sales & Inventory Intelligence System**

**1. Executive Summary**

This report outlines the design and implementation of a comprehensive data analysis solution for a retail company specializing in mobile phones and laptops. The primary goal was to analyze end-to-end operations—from sales and ordering to inventory and staff performance—to extract actionable insights.

By processing and analyzing 9 distinct datasets, we successfully built a system that identifies key performance indicators, uncovers operational inefficiencies, and provides a clear view of the company's health.

**Key Findings:**

* **Total Revenue:** The company generated **$7.69 million** in total sales.
* **Top Performers:** The **Baldwin Bikes (NY)** store is the clear leader, and **Mountain Bikes** is the most profitable category.
* **Critical Insight (Sales):** The business experienced strong, consistent sales in 2016-2017, followed by a **severe and immediate drop in sales and order volume starting in early 2018**.
* **Critical Insight (Operations):** The company faces a significant logistical challenge: **31.7% of all shipped orders were delayed**, arriving after the customer's required date.
* **Critical Insight (Inventory):** There are **25 product lines completely out of stock** and **141 product lines at critical low-stock levels** (1-4 units), representing significant lost revenue opportunities.

**2. Project Approach & Methodology**

The project was executed in three phases, following the plan outlined in the project brief.

**Phase 1: Data Preprocessing (Python & Pandas)**

The 9 raw CSV files (orders, customers, products, etc.) were loaded into a Python environment. The data was cleaned and transformed to ensure accuracy:

* **Date Formatting:** All date columns (order\_date, shipped\_date, required\_date) in the orders table were converted to a proper datetime format.
* **Null Value Handling:** Over 1,200 missing values in the customers[phone] column were filled with 'N/A' to maintain data integrity.
* **Feature Engineering:** A total\_price column was created in the order\_items table (quantity \* list\_price \* (1 - discount)) to serve as the primary metric for all sales analysis.
* The 9 cleaned datasets were saved as cleaned\_\*.csv files.

**Phase 2: Data Analysis & Modeling (Python & Power BI)**

This phase focused on answering the core business questions from the project brief.

1. **Data Loading:** All 9 cleaned CSVs were loaded into Power BI.
2. **Data Modeling:** A relational model was built in Power BI to connect all tables based on their primary and foreign keys (e.g., orders[store\_id] to stores[store\_id]), mirroring the provided database schema.
3. **DAX Measures:** Key performance indicators (KPIs) like Total Sales, Total Orders, and Total Customers were created using DAX formulas to ensure accurate and dynamic calculations.

**Phase 3: Visualization (Power BI)**

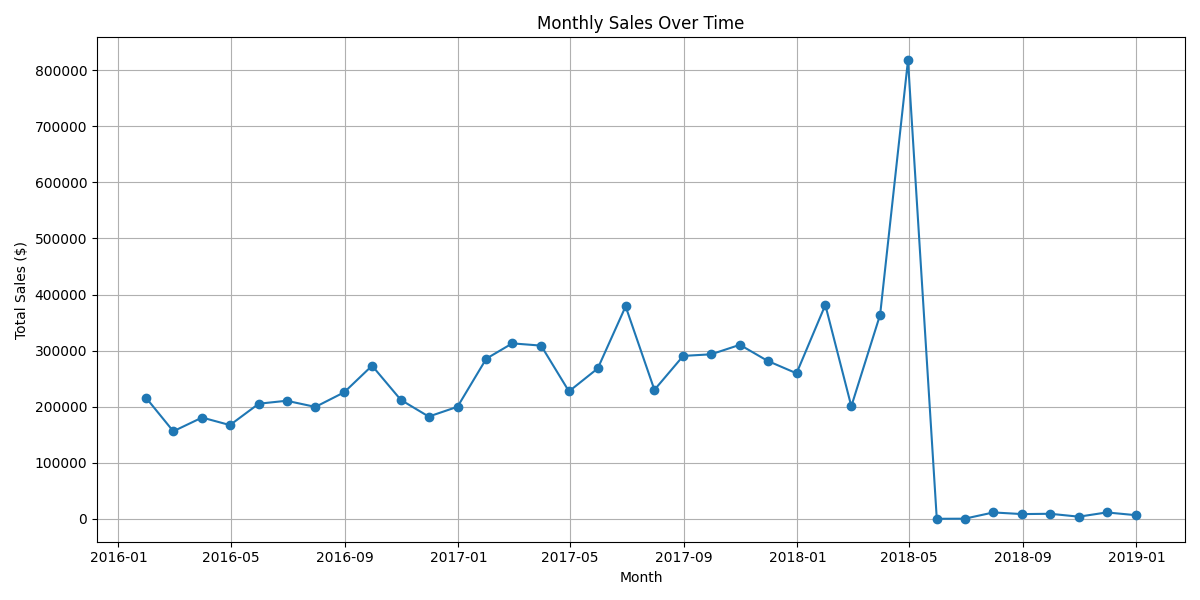
An interactive, multi-page dashboard was built in Power BI. This dashboard consolidates all findings into a single source of truth, allowing management to filter, drill down, and explore the data in real-time.

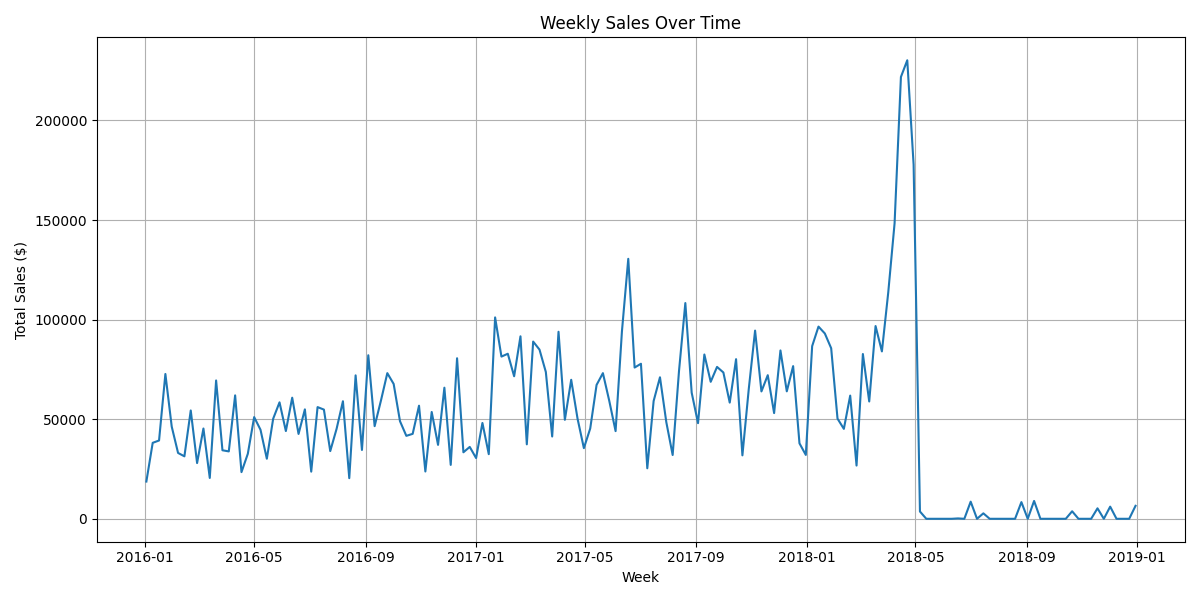
**3. Key Findings & Business Insights**

The analysis provided clear answers to all business use cases.

**Sales & Revenue Performance**

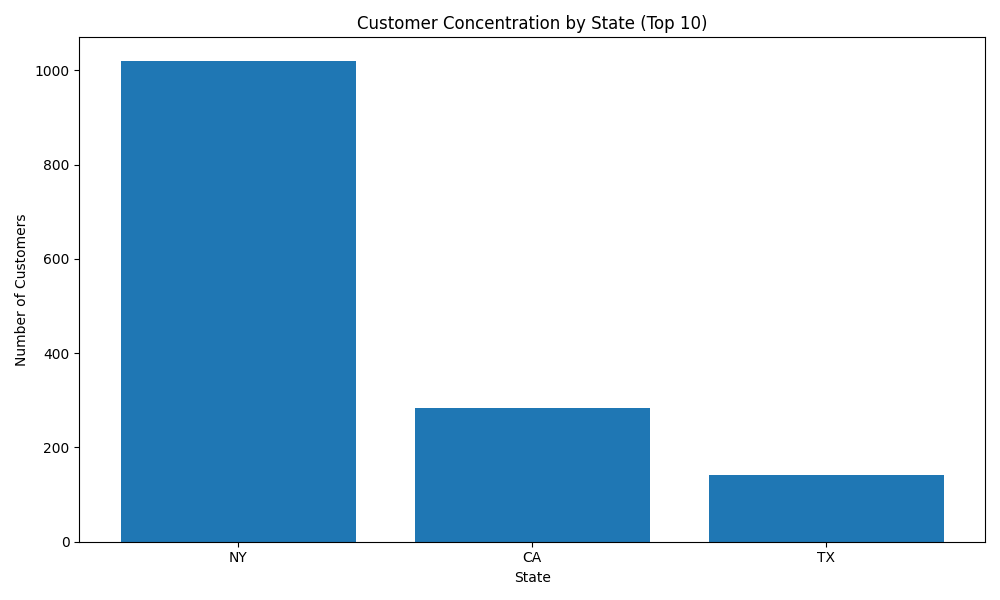
* **Overall Sales:** The business has generated **$7.69M** in total sales.
* **Sales Trend (Use Case 6):** Sales were strong and consistent through 2016 and 2017, but fell off a cliff in 2018. This is the single most important finding for the business to investigate.

****

****

* **Store Performance (Use Case 1):** The **Baldwin Bikes (NY)** store is the top performer, accounting for **$5.20M (67.6%)** of all sales.
  + Baldwin Bikes (NY): $5.20M
  + Santa Cruz Bikes (CA): $1.62M
  + Rowlett Bikes (TX): $0.86M
* **Staff Performance (Use Case 2):** Staff performance is directly correlated with store success. The top two staff, **Marcelene Boyer ($2.62M)** and **Venita Daniel ($2.59M)**, both work at the Baldwin Bikes store.

**Product & Customer Insights**

****

* **Profitable Categories (Use Case 4):** **Mountain Bikes** is the most profitable category, generating **$2.71M** in sales, followed by **Road Bikes ($1.66M)**. These two categories alone account for 57% of all revenue.
* **Brand Performance (Use Case 1):** **Trek** is the dominant brand across all stores, followed by **Electra** and **Surly**.
* **Customer Concentration (Use Case 7):** The customer base is heavily concentrated in **New York (1,019 customers)**, which explains the Baldwin store's success. This is followed by California (284) and Texas (142).
* **Top Customers:** The most valuable customer is **Sharyn Hopkins (NY)**, with $34,807 in lifetime sales. A full list of top customers has been generated.

**Operational & Inventory Analysis**

* **Order Fulfillment (Use Case 3):** There are **170 orders** in the system that are still unfulfilled (in "Pending," "Processing," or "Shipped" status).
* **Shipping Delays (Use Case 7):** A total of **458 orders (31.7% of all shipped orders)** were shipped *after* the required date. This indicates a major bottleneck in the fulfillment process that is likely damaging customer satisfaction.
* **Inventory Levels (Use Case 5):** The company has a critical inventory problem.
  + **Out-of-Stock:** **25** products are completely out of stock (quantity = 0).
  + **Low-Stock:** **141** products are at critically low stock levels (1-4 units). This combined 166-product deficit represents a significant and immediate loss of potential revenue.

**4. Recommendations & Next Steps**

Based on these findings, the following actions are recommended:

1. **Investigate the 2018 Sales Drop (Immediate Priority):** This is an existential threat. Management must immediately determine if this was a data-logging error (e.g., system change) or a real-world business collapse.
2. **Address Shipping Delays:** Conduct a root-cause analysis of the 31.7% delay rate. Is it tied to a specific store? A specific product category? A carrier? This is a correctable operational failure.
3. **Triage Inventory:** Immediately re-order the **25 out-of-stock items** and **141 low-stock items**, prioritizing those in the "Mountain Bike" and "Road Bike" categories to maximize revenue recapture.
4. **Develop Targeted Marketing:**
   * Launch a customer loyalty program for the high-value **New York** customer base.
   * Create targeted marketing campaigns for the **California** and **Texas** markets, focusing on the top-selling products (Trek Mountain Bikes) to grow market share.
5. **Adopt the Dashboard:** This Power BI dashboard should be adopted by management for weekly check-ins to monitor sales, track order fulfillment, and manage inventory levels proactively.

**5. Conclusion**

This project successfully transformed raw, siloed data into a unified intelligence system. The resulting Power BI dashboard provides a clear, 360-degree view of the business. By addressing the critical issues uncovered—the 2018 sales drop, the high shipping delay rate, and the widespread stock-outs—the company can correct its course and leverage its strengths to drive future growth.