Retail Sales Data Warehouse Analysis

End-to-end data analytics project leveraging SQL, Excel, and Power BI to analyze retail bike sales from a PostgreSQL data warehouse. The project covers data cleaning, customer segmentation, product performance analysis, and culminates in an interactive executive dashboard for dynamic insights.

Shivani Darshana
LinkedIn

Project Overview & Tools

This project involves time trend analysis , performance and category distribution analysis (part to whole), customer segmentation

Database

PostgreSQL

Tools Used

- Excel (Data Cleaning, EDA)
- PostgreSQL (SQL Queries)
- Power Bi

Data Cleaning & Preparation

Detailed validation and cleaning were performed to ensure data quality and consistency across the dimensional model, addressing missing values and maintaining referential integrity.

dim_customers

- 338 missing 'country' values replaced with 'Unknown'
- 15 missing 'gender' values replaced with 'Unknown'
- 17 blank 'birthdate' entries removed

dim_products

- 7 rows with missing category/sub_category/maintenance removed
- 17 rows missing 'product_line' dropped
- 2 products with 'cost = 0' set to NULL

fact_sales

- 19 missing 'order_date' entries removed
- Orphaned 'customer_key' values removed

Time-Based Sales Analysis

SQL aggregation was used to analyze year-wise sales and customer growth, revealing significant trends and seasonal patterns.

Yearly Sales Performance

- Sales increased exponentially from 2010 to 2013.
- Peak year: 2013 with 16.3M in sales
- 2014 shows a sudden drop, likely due to incomplete data or business downturn.

Customer Growth

- Strong increase in customer base indicates effective marketing and customer acquisition strategies.
- 17,427 customers in 2013 compared to 14 in 2010 is a major growth milestone.

Quantity Trends

- Strong growth in product movement (from 14 to 52,782 units/year).
- Demonstrates increasing market demand or better product availability.

Anomalies

- 2014 data is drastically low across all metrics. (due to partial availability of data

Monthly Seasonality Analysis

Seasonality Analysis

• Sales and customer numbers show strong seasonality.

• Peak months: November and December

• December had 3.21M in sales, 6,239 quantity entries, and 2,656 unique customers.

• Low activity in January–March, likely due to start of the year or due to post holiday season.

• Gradual ramp-up starting May, with consistent growth into Q4

Customer Behavior

• Customer counts align with sales trends, suggesting more customers—not just higher spending per customer.

• Sales and customer volume move in sync, showing that targeting marketing like promotional events – which includes the bike in a community race could have broader appeal to create brand awareness .

Category Contribution

Category Dependency Insight

Risks of Over-Reliance

01

Business

Revenue Vulnerability: A downturn in bike sales

02

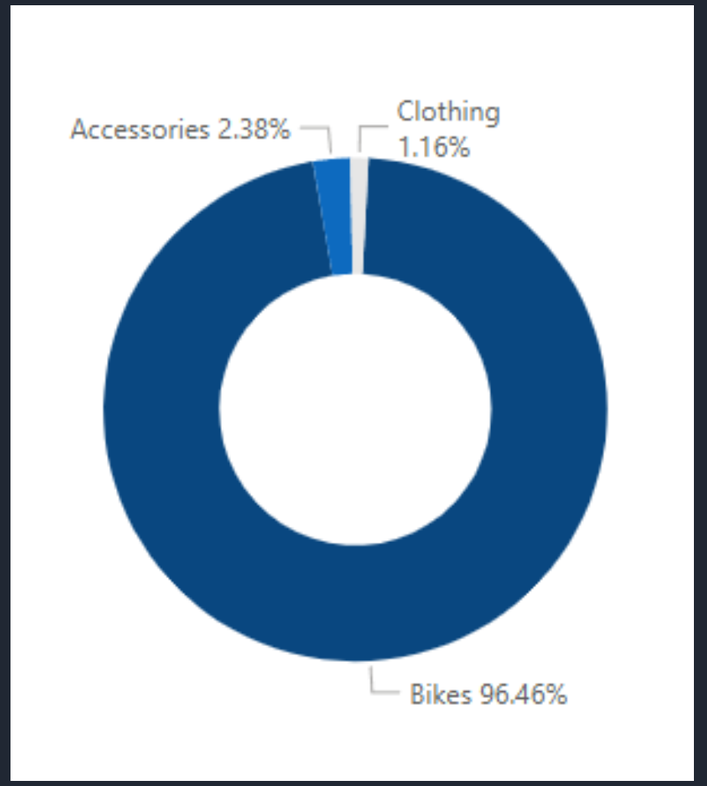
Brand Positioning Risk:

Risk in future if wanted to expand to other categories.

03

Growth

Growth Constraints: Heavy dependence on one



Category Risks & Recommendations

The heavy reliance on a single product category presents significant risks, necessitating strategic diversification and promotion efforts.

Over-reliance Risk

High dependency on 'Bikes' for 96.46% of revenue.

Diversify Offerings

Need to expand and promote other product lines.

Cross-sell & Bundle

Encourage sales of accessories and clothing through bundling.

Product Innovation

Invest in new product development to broaden appeal.

Customer Segmentation & Patterns

Customers were segmented based on spending and frequency, revealing a high percentage of new customers and potential churn risks.

High %

New Customers

Majority of customers are first-time buyers.

Churn

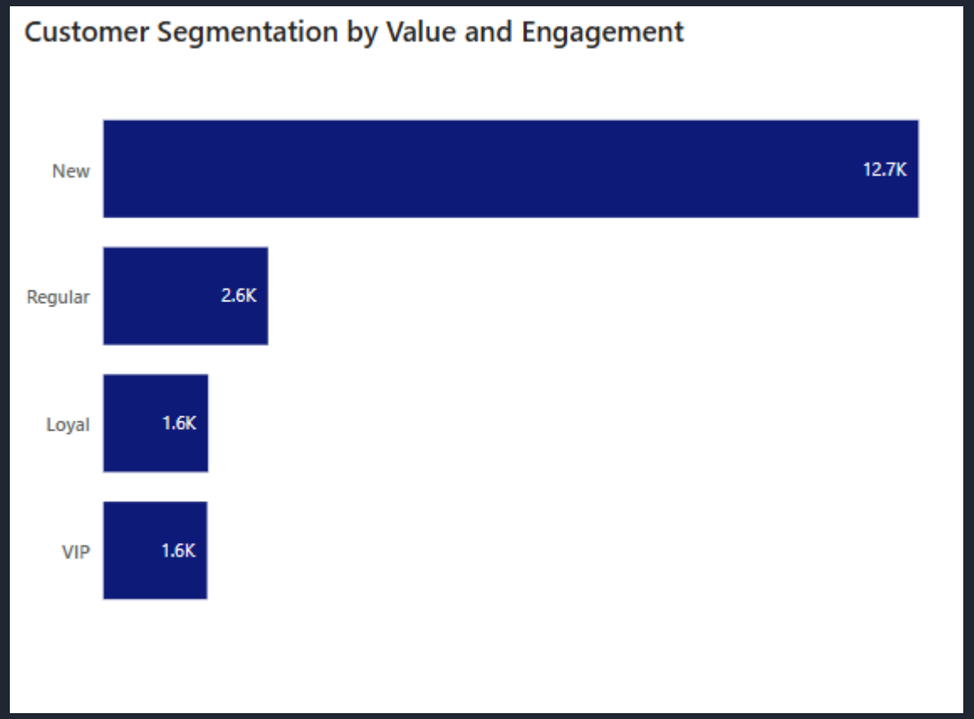
Risk Identified

High percentage of new customers suggests potential churn risk.

Retention

Strategy Needed

Focus on retention strategies for long-term growth.



Cumulative Analysis Insights

- Revenue is accumulating year after year, which indicates growth.
- Peak growth occurred between 2012–2013

Product Performance Evaluation

Each product’s yearly sales were evaluated against its historical average and the previous year’s value. This dual-metric approach highlights both relative performance and growth trends.

- Products like Classic Vest - S (2013) showed significant growth above average
- Others, like Fender Set - Mountain (2012), performed below expectations.

Key Takeaways & Next Steps

The analysis highlights significant sales growth, seasonal patterns, and critical areas for strategic focus to ensure sustained success.

- **Sales Growth:** Exponential growth from 2010–2013.
- **Seasonality:** November–December are peak sales months.
- **Top Product:** 'Classic Vest - S' performed exceptionally.
- **Category Risk:** 96% of sales from Bikes, requiring diversification.
- **Customer Pattern:** Majority are first-time buyers, indicating churn risk.