Globex Retail

Exploring how Globex Retail harnessed data engineering to transform customer data into actionable insights, driving business growth and innovation

10Alytics

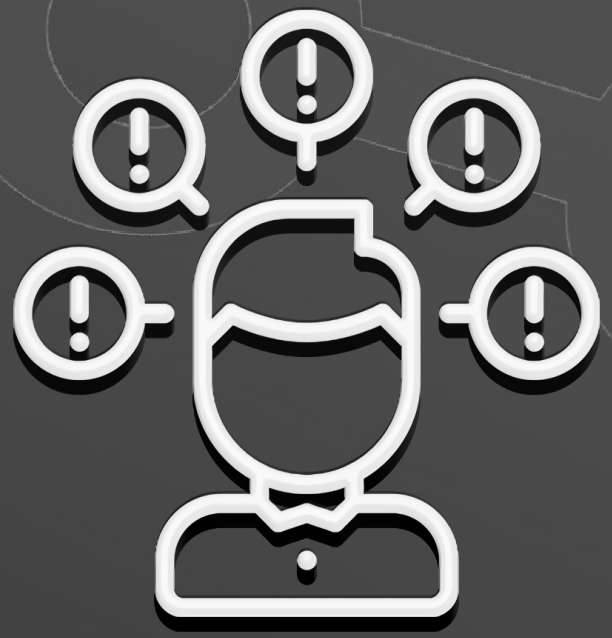
Business Introduction

Globex Retail, a dynamic leader in the retail industry, seamlessly integrates the convenience of e-commerce with the traditional appeal of brick-and-mortar stores. As the company continues to expand its reach, it recognizes the critical role that data engineering plays in driving success. By harnessing the power of advanced data engineering techniques, Globex Retail is committed to transforming raw data into actionable insights, optimizing inventory management, and ultimately boosting revenue. This data-driven approach positions Globex Retail at the forefront of innovation, ensuring that they not only meet but exceed the evolving demands of their diverse customer base.

10Alytics

PROBLEM STATEMENT

As a new Data Engineer at Globex Retail, a company specializing in e-commerce and brick-and-mortar sales, your first task involves analyzing a dataset to understand customer purchase behavior. The goal is to identify patterns in customer purchases that could lead to targeted marketing strategies and improved inventory management.



10Alytics

Rationale for the Project

Globex Retail seeks to

- **Enhance customer retention.**
- **Optimize inventory levels to meet customer demand.**
- **Increase revenue through targeted marketing.**
- **Make data-driven decisions that improve overall business performance.**

10Alytics

Data Description

The dataset includes information on customer purchases, such as:

1. Customer ID: Unique identifier for customers.
2. Order ID: Unique identifier for each order.
3. Order Date: Date the order was placed.
4. Product Category: Main category of the product.
5. Product Sub-Category: More specific category of the product.
6. Quantity: Number of units ordered.
7. Price: Price per unit of the product.
8. Discount: Discount applied to the product.
9. Customer Location: Location (state code) of the customer.

10Alytics

Tech Stack

Python

Pandas

Jupyter

10Alytics

Project Workflow

- Familiarize yourself with the Dataset
- Clean the data
- Engineer new features
- Perform Exploratory Data Analysis (EDA)
- Generating Insights with Pandas

10Alytics

Generating Insights with Pandas

- Which product category has generated the most revenue?
- Which product subcategories generate the most revenue?
- Which customer segment has the highest average order value?
- What is the average discount applied to high-value customers?
- Which geographic locations have the highest concentration of high-value customers?
- How does the discount rate impact the total sales across different product categories?
- Identify trends in purchase behavior over time (e.g., monthly).