

# Business Intelligence Project – Use Case Proposal

## Project Title

Global Retail Sales Intelligence: Customer Behavior & Performance Analysis

### STUDENTS:

lina chelaifa / chayma maagli /ghofrane mzoughi / Areej mighri

---

## 1. Introduction & Business Context

In today's highly competitive global retail environment, organizations must rely on data-driven insights to optimize operations, improve customer satisfaction, and increase profitability. Retail companies generate massive volumes of transactional and customer data on a daily basis; however, without proper Business Intelligence (BI) systems, this data remains underutilized.

This project aims to design and implement an end-to-end Business Intelligence solution that analyzes global retail sales transactions and customer demographics. By integrating structured and semi-structured data, the project provides meaningful insights into sales performance, customer behavior, product trends, and geographical patterns, enabling managers to make informed strategic decisions.

---

## 2. Business Problem

The retail company faces several challenges:

- Limited visibility into global and regional sales performance
- Difficulty identifying high-value customers and loyal segments
- Inefficient evaluation of product and category profitability
- Lack of understanding of seasonal and time-based sales trends
- Inability to measure the impact of promotions and discounts

As a result, decisions related to marketing, inventory, and expansion are often reactive rather than proactive.

---

## 3. Business Objectives

The main objectives of this BI project are to:

- Centralize sales and customer data into a data warehouse
- Track and monitor key performance indicators (KPIs)
- Analyze customer behavior and segmentation

- Evaluate product and regional performance
- Support strategic decisions through interactive dashboards and insights

## 4. Dataset Description

The project uses a large retail dataset (10,000+ records) collected from an open data source (e.g., Kaggle). The data represents transactional sales across multiple countries and includes both structured and semi-structured formats.

### Data Sources

- Transactional Sales Data (CSV):
  - Transaction ID
  - Date
  - Customer ID
  - Product Category (Beauty, Clothing, Electronics, etc.)
  - Quantity
  - Price per Unit
  - Total Amount
  - Country and City
- Customer Demographics Data (JSON):
  - Customer ID
  - Gender
  - Age
  - Country
  - City

## 5. Use Cases & Analytical Scope

The Business Intelligence analysis is organized around seven critical dimensions, using the same wording and structure as the reference example, to ensure full consistency with the defined use case tables.

### 1) Sales Performance

- Total Revenue: Monitor global earnings
- Growth Trends: Highlight Month-over-Month (MoM) and Year-over-Year (YoY) fluctuations

### 2) Customer Insights

- Customer Lifetime Value (CLV): Identify and prioritize high-value customers
- Demographics: Segment and target customers based on age and gender
- Repeat Rate: Measure customer loyalty and repeat purchasing behavior

### 3) Product Performance

- Top Categories: Analyze category dominance and contribution to revenue
- Regional Preferences: Identify product category performance across different regions

## 4) Geographical Insights

- Top Regions: Identify high-performing countries and cities
- Regional Growth: Detect growth opportunities and underperforming markets

## 5) Time-Based Analysis

- Seasonality: Analyze peak and low sales periods
- Daily Trends: Evaluate sales behavior during special events and promotions

## 6) Customer Segmentation

- Age Groups: Analyze purchasing behavior by age segments
- Gender Trends: Study gender-based purchasing patterns

## 7) Operational Efficiency

- Transaction Size: Analyze average number of units per order
- Revenue per Unit: Measure profitability per unit sold

---

## 6. Key Performance Indicators (KPIs)

The following KPIs are directly aligned with the seven use cases described above:

1. Total Revenue
2. Sales Growth Rate (MoM / YoY)
3. Customer Lifetime Value (CLV)
4. Customer Repeat Purchase Rate
5. Revenue by Product Category
6. Profit Margin (%)
7. Average Transaction Size
8. Revenue per Unit Sold

## 7. Data Preparation

To prepare the data for analysis, an ETL process was implemented using Python:

- Integration of CSV and JSON datasets
- Removal of duplicated records
- Handling missing or inconsistent values
- Standardization of data types (dates, numerical values)
- Creation of calculated fields such as total sales and profit

---

## 8. Data Storage & Modeling

### Database

- Relational database optimized for analytical queries

### Data Model

## Fact Table

### Sales\_Fact

Description: Central table storing transactional sales data. Fields:

- TransactionID (Primary Key)
- Date (Foreign Key to Date\_Dimension)
- CustomerID (Foreign Key to Customer\_Dimension)
- ProductCategory (Foreign Key to Product\_Dimension)
- Quantity
- PricePerUnit
- TotalAmount

## Dimension Tables

### Customer\_Dimension

- CustomerID (Primary Key)
- Gender
- Age
- GeographyID (Foreign Key to Geography\_Dimension)

### Product\_Dimension

- ProductCategory (Primary Key)

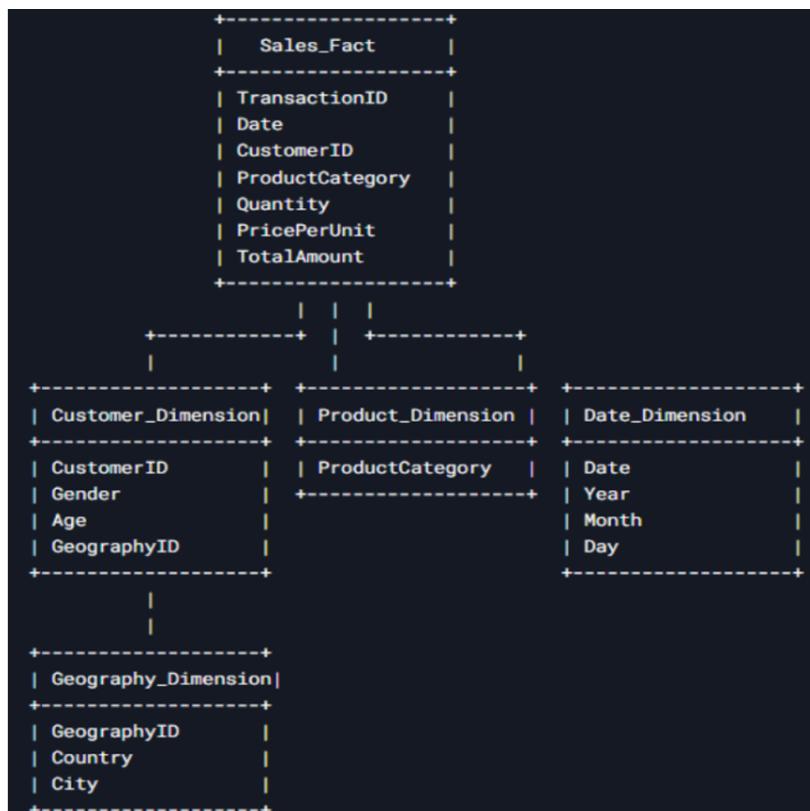
### Date\_Dimension

- Date (Primary Key)
- Year
- Month
- Day

### Geography\_Dimension

- GeographyID (Primary Key)
- Country
- City

## Data Model Diagram



## 9. Visualization & BI Tools

The cleaned and modeled data is visualized using a BI tool such as Power BI or

Tableau through interactive dashboards:

- Sales Overview Dashboard (revenue, growth, trends)
  - Customer Insights Dashboard (CLV, segmentation)
  - Product Performance Dashboard (categories, profitability)
  - Geographical Dashboard (sales by country and city)
- 

## 10. Insights & Expected Outcomes

The BI solution enables decision-makers to:

- Identify high-performing products and regions
- Target high-value customer segments
- Optimize marketing and promotional strategies
- Improve inventory and pricing decisions
- Detect seasonal opportunities and risks