Business Requirements Document (BRD)

Project Title: Retail Sales Data Analysis

Prepared By: Fusion Software Institute.
Date: 22/03/2025

---

1. Business Overview

This project aims to analyze a retail dataset to uncover insights into customer purchasing behavior, sales trends, and product preferences. The dataset comprises essential attributes such as transaction details, customer demographics, product categories, and sales figures.

---

2. Objectives and Goals

The key objectives of this analysis are:

- To understand how customer age and gender influence purchasing behavior.
- To identify patterns in sales across different time periods.
- To determine which product categories are most popular among customers.
- To analyze relationships between age, spending, and product preferences.
- To examine how customers adapt shopping habits based on seasonal trends.
- To identify distinct purchasing behaviors based on the number of items bought per transaction.
- To extract insights from the distribution of product prices within each category.

---

3. Data Description

The dataset includes the following key attributes:

Column Name	Description
Transaction ID	Unique identifier for each transaction
Date	Date of the transaction
Customer ID	Unique identifier for each customer
Gender	Gender of the customer
Age	Age of the customer
Product Category	Category of the purchased product
Quantity	Number of units bought in a transaction
Price per Unit	Cost of a single unit of the product
Total Amount	Total revenue from the transaction

4. Functional Requirements

4.1 Data Preprocessing

- Load the dataset and check for missing or inconsistent values.
- Handle missing values through imputation or removal as required. (if required)
- Convert data types appropriately (e.g., Date column to datetime format).
- Create new features if necessary (e.g., extracting year, month, or day from Date).

4.2 Exploratory Data Analysis (EDA)

- **Demographic Influence on Purchasing Behavior**
  - Analyze the relationship between age, gender, and spending behavior.
  - Compare average spending across different age groups and genders.
- **Sales Patterns Across Time Periods**
  - Identify trends in sales over different months, seasons, and years.
  - Detect peak and low sales periods.
- **Product Category Preferences**
  - Determine the most popular product categories by sales volume and revenue.
  - Analyze customer demographics for each product category.
- **Customer Spending and Product Preferences**
  - Study correlations between age, total spending, and product choices.
  - Identify high-value customer segments based on spending behavior.
- **Seasonal Shopping Trends**
  - Compare purchasing behavior during different seasons.
  - Identify changes in product preferences over time.
- **Transaction-Based Purchasing Behavior**
  - Analyze the impact of quantity bought per transaction on total spending.
  - Identify customer groups that buy in bulk vs. small quantities.
- **Product Price Distribution Insights**
  - Evaluate the price distribution within each product category.
  - Identify price-sensitive and premium product segments.

---

5. Tools and Technologies

The analysis will be performed using the following tools:

- **Programming Language:** Python
- **Libraries:** Pandas, NumPy, Matplotlib, Seaborn, Scipy, Plotly
- **Jupyter Notebook** for interactive analysis

6. Deliverables

The final deliverables for this project will include:

- A detailed EDA report with visualizations and insights.
- A cleaned and preprocessed dataset for further analysis.
- Data-driven recommendations based on findings.
- A presentation summarizing key insights for stakeholders.