**About Dataset**

**Dataset Description:**

* The dataset represents retail transactional data. It contains information about customers, their purchases, products, and transaction details. The data includes various attributes such as customer ID, name, email, phone, address, city, state, zipcode, country, age, gender, income, customer segment, last purchase date, total purchases, amount spent, product category, product brand, product type, feedback, shipping method, payment method, and order status.

**Key Points:**

**Customer Information:**

* Includes customer details like ID, name, email, phone, address, city, state, zipcode, country, age, and gender.  
  Customer segments are categorized into Premium, Regular, and New.

**Transaction Details:**

* Transaction-specific data such as transaction ID, last purchase date, total purchases, amount spent, total purchase amount, feedback, shipping method, payment method, and order status.

**Product Information:**

* Contains product-related details such as product category, brand, and type.  
  Products are categorized into electronics, clothing, grocery, books, and home decor.

**Geographic Information:**

* Contains location details including city, state, and country.  
  Available for various countries including USA, UK, Canada, Australia, and Germany.

**Temporal Information:**

* Last purchase date is provided along with separate columns for year, month, date, and time.  
  Allows analysis based on temporal patterns and trends.

**Data Quality:**

* Some rows contain null values, and others are duplicates, which may need to be handled during data preprocessing.  
  Null values are randomly distributed across rows.  
  Duplicate rows are available at different parts of the dataset.

**Potential Analysis:**

* Customer segmentation analysis based on demographics, purchase behavior, and feedback.  
  Sales trend analysis over time to identify peak seasons or trends.  
  Product performance analysis to determine popular categories, brands, or types.  
  Geographic analysis to understand regional preferences and trends.  
  Payment and shipping method analysis to optimize services.  
  Customer satisfaction analysis based on feedback and order status.

**Data Preprocessing:**

* Handling null values and duplicates.  
  Parsing and formatting temporal data.  
  Encoding categorical variables.  
  Scaling numerical variables if required.  
  Splitting data into training and testing sets for modeling.