**🛒 Online Retail RFM Analysis**

**📌 Project Overview**

This project applies **RFM (Recency, Frequency, Monetary)** analysis to the *Online Retail II* dataset. The goal is to identify customer segments based on their purchasing behavior and prepare the data for further use in **Power BI dashboards** and **predictive modeling**.

RFM is a widely used customer segmentation technique in marketing:

* **Recency** → How recently a customer made a purchase
* **Frequency** → How often a customer makes a purchase
* **Monetary** → How much money a customer spends

**⚙️ Workflow**

**1. Data Cleaning & Preprocessing**

* Loaded dataset from **CSV**.
* Added **region labels** by merging with an external Excel file.
* Filled missing numeric values using **Subregion-based median or mean**, depending on skewness.
* Removed negative values in *Quantity* and *Price*.
* Created **TotalPrice = Quantity × Price**.

**2. RFM Table Construction**

* Defined **snapshot\_date** (max date + 1).
* Calculated:
  + Recency = days since last purchase
  + Frequency = number of unique invoices
  + Monetary = total spend
* Applied **qcut** to assign quartile scores for Recency, Frequency, and Monetary.

**3. Logistic Regression (Feature Importance)**

* Built a dummy **classification target (y)** to simulate purchase prediction.
* Scaled RFM features using **StandardScaler**.
* Trained **Logistic Regression** model.
* Extracted feature coefficients → converted to **percentage importance**.

**4. Customer Segmentation**

* Created an **RFM Score** (multiplication of quartiles).
* Segmented customers into 4 groups:
  + **At Risk**
  + **Regular**
  + **Loyal**
  + **VIP**
* Exported cleaned dataset and RFM table to **Excel** for use in **Power BI**.

**📊 Outputs**

* **online\_retail\_cleaned.xlsx** → Cleaned transaction-level dataset
* **rfm\_table.xlsx** → Customer-level RFM table with segment labels

**🖼 Example Power BI Dashboard (to add screenshot)**

*(Add an image of your Power BI dashboard here if available)*

**📚 Technologies Used**

* **Python**: Pandas, NumPy, Scikit-learn, Scipy, Seaborn, Matplotlib
* **Excel**: Exporting RFM tables for BI integration
* **Power BI**: Dashboard creation and visualization