**Task Three Work Description**

**1. Introduction**

This project involves the application of **Recency-Frequency-Monetary (RFM) analysis** to the *Online Retail* dataset in order to segment customers based on purchasing behavior. The analysis aims to identify distinct customer groups and provide targeted marketing recommendations.

**2. Data Acquisition and Preparation**

* **Dataset Source**: *Online Retail.xlsx*
* **Tools Used**: Python, Pandas, NumPy, Matplotlib, Seaborn, Google Colab.
* **Initial Dataset**: 541,909 transactions with 8 columns:  
  InvoiceNo, StockCode, Description, Quantity, InvoiceDate, UnitPrice, CustomerID, Country.

**Data Cleaning Steps:**

1. **Missing Values Handling**
   * Removed rows with missing CustomerID or Description.
2. **Data Type Adjustments**
   * Converted CustomerID to string type for uniformity.
3. **Cancelled Orders Removal**
   * Filtered out invoices starting with 'C'.
4. **Total Price Calculation**
   * Created TotalPrice = Quantity × UnitPrice for monetary value computation.

**Final Clean Dataset Size**: Reduced after cleaning to only valid, complete transactions.

**3. RFM Metric Computation**

**Reference Date**: One day after the latest invoice date in the dataset.

**Metrics:**

* **Recency (R)**: Days since last purchase for each customer.
* **Frequency (F)**: Number of unique invoices per customer.
* **Monetary (M)**: Total purchase value per customer.

These metrics were computed via group aggregation, producing an **RFM table**.

**4. Scoring and Segmentation**

* **Quantile-Based Scoring**:
  + Recency scored inversely (low recency = high score).
  + Frequency & Monetary scored directly (higher value = higher score).
  + Scores range from **1 to 4**.
* **RFM Score**: Concatenation of R, F, and M scores (e.g., 444 = best in all three metrics).

**Segment Classification Rules:**

* 444 → **Best Customers**
* R = 4 → **Loyal Customers**
* F = 4 → **Frequent Buyers**
* M = 4 → **Big Spenders**
* All others → **Others**

**5. Customer Segment Distribution**

| **Segment** | **Count** | **% of Customers** |
| --- | --- | --- |
| Others | 2605 | Largest group |
| Loyal Customers | 664 |  |
| Best Customers | 462 |  |
| Frequent Buyers | 343 |  |
| Big Spenders | 265 | Smallest group |

**6. Visualization**

* **Pie Chart**: Customer segment distribution.
* **Count Plot**: Bar chart of segment counts.
* **Heatmap**: Average Recency, Frequency, and Monetary values per segment.

**7. Recommendations**

* **Best Customers**: Maintain strong relationships through VIP programs, early product access, and personalized offers.
* **Loyal Customers**: Enhance loyalty rewards, provide tailored recommendations.
* **Frequent Buyers**: Promote upselling via bundles or higher-value items.
* **Big Spenders**: Offer exclusive premium services and personalized engagement.
* **Others**: Implement re-engagement campaigns, time-sensitive offers, and feedback collection.

**8. Conclusion**

This RFM analysis successfully segmented the customer base into **five distinct behavioral groups**. These insights enable **data-driven marketing strategies** to improve customer retention, increase purchase value, and optimize targeting efforts.