In [11]:

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

In [12]:

```
df = pd.read_csv('Apriori Customer_Online_Transaction_data.csv')
df.head()
```

Out[12]:

	InvoiceNo	StockCode	Description	Quantity	InvoiceDate	UnitPrice	CustomerID	Country
0	536365	85123A	WHITE HANGING HEART T-LIGHT HOLDER	6	01/12/2010 08:26	2.55	17850.0	United Kingdom
1	536365	71053	WHITE METAL LANTERN	6	01/12/2010 08:26	3.39	17850.0	United Kingdom
2	536365	84406B	CREAM CUPID HEARTS COAT HANGER	8	01/12/2010 08:26	2.75	17850.0	United Kingdom
3	536365	84029G	KNITTED UNION FLAG HOT WATER BOTTLE	6	01/12/2010 08:26	3.39	17850.0	United Kingdom
4	536365	84029E	RED WOOLLY HOTTIE WHITE HEART.	6	01/12/2010 08:26	3.39	17850.0	United Kingdom

In [13]:

```
df['TotalPrice'] = df['Quantity'] * df['UnitPrice']
```

In [14]:

```
df['InvoiceDate'] = pd.to_datetime(df['InvoiceDate'], format='%d/%m/%Y %H:%M', errors='coerce')
```

In [15]:

```
df['InvoiceDate'].fillna(method='ffill', inplace=True)
```

```
C:\Users\User\AppData\Local\Temp\ipykernel_21892\3208964462.py:1: FutureWarning: A value is tr
ying to be set on a copy of a DataFrame or Series through chained assignment using an inplace
method.
The behavior will change in pandas 3.0. This inplace method will never work because the interm
ediate object on which we are setting values always behaves as a copy.

For example, when doing 'df[col].method(value, inplace=True)', try using 'df.method({col: valu
e}, inplace=True)' or df[col] = df[col].method(value) instead, to perform the operation inplac
e on the original object.


  df['InvoiceDate'].fillna(method='ffill', inplace=True)
C:\Users\User\AppData\Local\Temp\ipykernel_21892\3208964462.py:1: FutureWarning: Series.fillna
with 'method' is deprecated and will raise in a future version. Use obj.ffill() or obj.bfill()
instead.
  df['InvoiceDate'].fillna(method='ffill', inplace=True)
```

In [16]:

```
print(df.head())
```

```
  InvoiceNo StockCode                          Description  Quantity  \
0    536365    85123A   WHITE HANGING HEART T-LIGHT HOLDER         6   
1    536365     71053                  WHITE METAL LANTERN         6   
2    536365    84406B       CREAM CUPID HEARTS COAT HANGER         8   
3    536365    84029G  KNITTED UNION FLAG HOT WATER BOTTLE         6   
4    536365    84029E       RED WOOLLY HOTTIE WHITE HEART.         6   

          InvoiceDate  UnitPrice  CustomerID         Country  TotalPrice  
0 2010-12-01 08:26:00       2.55     17850.0  United Kingdom       15.30  
1 2010-12-01 08:26:00       3.39     17850.0  United Kingdom       20.34  
2 2010-12-01 08:26:00       2.75     17850.0  United Kingdom       22.00  
3 2010-12-01 08:26:00       3.39     17850.0  United Kingdom       20.34  
4 2010-12-01 08:26:00       3.39     17850.0  United Kingdom       20.34  
```

In [17]:

```
df['Recency'] = (df['InvoiceDate'].max() - df['InvoiceDate']).dt.days
```

In [18]:

```
frequency_df = df.groupby('CustomerID')['InvoiceNo'].count().reset_index()
frequency_df.rename(columns={'InvoiceNo': 'Frequency'}, inplace=True)
df = df.merge(frequency_df, on='CustomerID', how='left')
```

In [19]:

```
monetary_df = df.groupby('CustomerID')['TotalPrice'].sum().reset_index()
monetary_df.rename(columns={'TotalPrice': 'MonetaryValue'}, inplace=True)
df = df.merge(monetary_df, on='CustomerID', how='left')
```

In [21]:

```
rfm_df = df[['CustomerID', 'Recency', 'Frequency', 'MonetaryValue']]
print(rfm_df.head())
print(df.head())
```

```
   CustomerID  Recency  Frequency  MonetaryValue
0     17850.0      373      312.0        5288.63
1     17850.0      373      312.0        5288.63
2     17850.0      373      312.0        5288.63
3     17850.0      373      312.0        5288.63
4     17850.0      373      312.0        5288.63
  InvoiceNo StockCode                          Description  Quantity  \
0    536365    85123A   WHITE HANGING HEART T-LIGHT HOLDER         6
1    536365     71053                  WHITE METAL LANTERN         6
2    536365    84406B       CREAM CUPID HEARTS COAT HANGER         8
3    536365    84029G  KNITTED UNION FLAG HOT WATER BOTTLE         6
4    536365    84029E       RED WOOLLY HOTTIE WHITE HEART.         6

          InvoiceDate  UnitPrice  CustomerID         Country  TotalPrice  \
0 2010-12-01 08:26:00       2.55     17850.0  United Kingdom       15.30
1 2010-12-01 08:26:00       3.39     17850.0  United Kingdom       20.34
2 2010-12-01 08:26:00       2.75     17850.0  United Kingdom       22.00
3 2010-12-01 08:26:00       3.39     17850.0  United Kingdom       20.34
4 2010-12-01 08:26:00       3.39     17850.0  United Kingdom       20.34

   Recency  Frequency  MonetaryValue
0      373      312.0        5288.63
1      373      312.0        5288.63
2      373      312.0        5288.63
3      373      312.0        5288.63
4      373      312.0        5288.63
```

In [ ]: