In [1]:
```
import pandas as pd
import numpy as np
import sklearn as svm
```

In [3]:
```
cust_order= pd.read_csv('CLV_RFM_customer_orders.csv')
```

In [4]:
```
cust_order.head()
```

Out[4]:

	Order_id	Order_date	Customer	Amount
0	O-2018-100006	09-07-2018	Bharat	473
1	O-2018-100090	07-08-2018	Pearl	874
2	O-2018-100293	03-14-2018	Jahan	114
3	O-2018-100328	1-29-2018	Divsha	5
4	O-2018-100363	04-08-2018	Kasheen	26

In [5]:
```
cust_order.shape
```

Out[5]:
```
(2000, 4)
```

In [6]:
```
cust_order['Order_date'] = pd.to_datetime(cust_order['Order_date'])
```

In [7]:
```
from datetime import datetime
rfm_cust = cust_order.groupby('Customer').agg({'Order_date': lambda x:(datetime.t
                                               'Order_id': lambda x :len(x),
                                               'Amount': lambda x: x.sum()})
```

In [8]:
```
rfm_cust.head()
```

Out[8]:

Customer	Order_date	Order_id	Amount
Aakanksha	656	2	5391
Aarushi	90	17	9973
Aashna	456	4	995
Aastha	585	4	505
Aayush	475	4	1048

In [9]:
```
rfm_cust['Order_date'] = rfm_cust['Order_date'].astype(int)
rfm_cust.columns = ['recency','frequency','monetary']
rfm_cust.head()
```

Out[9]:

Customer	recency	frequency	monetary
Aakanksha	656	2	5391
Aarushi	90	17	9973
Aashna	456	4	995
Aastha	585	4	505
Aayush	475	4	1048

In [10]:
```
rfm_cust.head(4)
```

Out[10]:

Customer	recency	frequency	monetary
Aakanksha	656	2	5391
Aarushi	90	17	9973
Aashna	456	4	995
Aastha	585	4	505

In [11]:
```
cust_order[cust_order['Customer'] == 'Aakanksha']
```

Out[11]:

	Order_id	Order_date	Customer	Amount
108	O-2018-108147	2018-06-08	Aakanksha	2181
1066	O-11-20190677	2019-05-31	Aakanksha	3210

In [12]:
```
rfm_cust['R_quartile'] = pd.qcut(rfm_cust['recency'], q=4, labels = ['1','2','3',
rfm_cust['F_quartile'] = pd.qcut(rfm_cust['frequency'].rank(method = 'first'),q=4
rfm_cust['M_quartile'] = pd.qcut(rfm_cust['monetary'], q=4, labels = ['4','3','2'
```

In [13]:
```
rfmlist=[]
for i in range(0,len(rfm_cust)):
    rfmlist.append(str(rfm_cust.iloc[i]["R_quartile"])+str(rfm_cust.iloc[i]["F_qu
rfm_cust['RFM'] = rfmlist
```

In [14]:
```
rfm_cust.head()
```

Out[14]:

Customer	recency	frequency	monetary	R_quartile	F_quartile	M_quartile	RFM
Aakanksha	656	2	5391	4	4	1	441
Aarushi	90	17	9973	1	1	1	111
Aashna	456	4	995	2	4	3	243
Aastha	585	4	505	4	4	4	444
Aayush	475	4	1048	3	4	3	343

In [15]:
```
rfm_cust.tail()
```

Out[15]:

Customer	recency	frequency	monetary	R_quartile	F_quartile	M_quartile	RFM
Wale	726	2	129	4	4	4	444
Yaanvi	84	6	1121	1	2	3	123
Yash	540	4	8653	4	2	1	421
Yogesh	123	21	6922	1	1	1	111
Yohann	561	4	1139	4	2	3	423

In [16]:
```
rfm_cust[rfm_cust['RFM']=='114'].sort_values(by ='monetary',ascending = False).he
```

Out[16]:

Customer	recency	frequency	monetary	R_quartile	F_quartile	M_quartile	RFM

In [17]:
```
rfm_cust[rfm_cust['RFM']=='444'].sort_values(by ='monetary',ascending = False).he
```

Out[17]:

Customer	recency	frequency	monetary	R_quartile	F_quartile	M_quartile	RFM
Swetlana	583	2	836	4	4	4	444
Subhasmita	631	2	756	4	4	4	444
Anmol	564	4	733	4	4	4	444
Aastha	585	4	505	4	4	4	444
Inderpreet	718	2	500	4	4	4	444
Bathina	724	2	458	4	4	4	444
Parna	585	2	447	4	4	4	444
Bhawna	552	2	338	4	4	4	444
Samiksha	640	2	332	4	4	4	444
Manshul	804	2	318	4	4	4	444

In [18]:
```
rfm_cust[rfm_cust['RFM']=='111'].sort_values(by ='monetary',ascending = False).he
```

Out[18]:

Customer	recency	frequency	monetary	R_quartile	F_quartile	M_quartile	RFM
Mukesh	76	17	29905	1	1	1	111
Sanjay	133	17	22196	1	1	1	111
Anita	116	17	19425	1	1	1	111
Ramesh	121	21	17707	1	1	1	111
Ayush	89	10	16710	1	1	1	111
Sagar	88	21	14915	1	1	1	111
Deepak	81	20	14761	1	1	1	111
Bhishm	111	20	14543	1	1	1	111
Monisha	192	20	13715	1	1	1	111
Parth	146	12	13636	1	1	1	111

In [ ]:
```

```