Project 3 Code - Data Cleaning_Part1

February 12, 2021

```python
[1]: # Importing necessary modules
     import pandas as pd
     import numpy as np
     import matplotlib.pyplot as plt
     import datetime as dt
     import warnings
     warnings.filterwarnings("ignore")
```

```python
[2]: # Reading Card Transactions Raw Data
     data = pd.read_excel('card transactions.xlsx')
```

0.1 Exclusions

```python
[3]: #Retaining only P transaction type
     data = data[data['Transtype'] == 'P']
     data.head()
```

```
[3]:    Recnum      Cardnum       Date      Merchnum           Merch description  \
     0       1  5142190439  2010-01-01  5509006296254     FEDEX SHP 12/23/09 AB#
     1       2  5142183973  2010-01-01    61003026333   SERVICE MERCHANDISE #81
     2       3  5142131721  2010-01-01  4503082993600           OFFICE DEPOT #191
     3       4  5142148452  2010-01-01  5509006296254     FEDEX SHP 12/28/09 AB#
     4       5  5142190439  2010-01-01  5509006296254     FEDEX SHP 12/23/09 AB#

        Merch state  Merch zip Transtype  Amount  Fraud
     0          TN    38118.0         P    3.62      0
     1          MA     1803.0         P   31.42      0
     2          MD    20706.0         P  178.49      0
     3          TN    38118.0         P    3.62      0
     4          TN    38118.0         P    3.62      0
```

```python
[4]: #Removing the largest transaction (It is in Pesos)
     data = data[data['Amount'] < 3000000]
     data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 96397 entries, 0 to 96752
Data columns (total 10 columns):
```

```
     #   Column            Non-Null Count   Dtype
    ---  ------            --------------   -----
     0   Recnum            96397 non-null   int64
     1   Cardnum           96397 non-null   int64
     2   Date              96397 non-null   datetime64[ns]
     3   Merchnum          93199 non-null   object
     4   Merch description 96397 non-null   object
     5   Merch state       95377 non-null   object
     6   Merch zip         92097 non-null   float64
     7   Transtype         96397 non-null   object
     8   Amount            96397 non-null   float64
     9   Fraud             96397 non-null   int64
    dtypes: datetime64[ns](1), float64(2), int64(3), object(4)
    memory usage: 8.1+ MB
```

[5]: ```python
     #% of missing records
     data.count() * 100 /len(data)
     ```

[5]: ```
     Recnum              100.000000
     Cardnum             100.000000
     Date                100.000000
     Merchnum             96.682469
     Merch description   100.000000
     Merch state          98.941876
     Merch zip            95.539280
     Transtype           100.000000
     Amount              100.000000
     Fraud               100.000000
     dtype: float64
     ```

[6]: ```python
     #checking nulls
     data.isnull().sum()
     ```

[6]: ```
     Recnum                 0
     Cardnum                0
     Date                   0
     Merchnum            3198
     Merch description      0
     Merch state         1020
     Merch zip           4300
     Transtype              0
     Amount                 0
     Fraud                  0
     dtype: int64
     ```

0.2 Data Cleaning

```python
[7]: # Replacing 0s with nan
     data['Merchnum'] = data['Merchnum'].replace(np.nan,'0')
     data['Merch state'] = data['Merch state'].replace(np.nan,'0')
     data['Merch zip'] = data['Merch zip'].replace(np.nan,0)
     data.isnull().sum()
```

```
[7]: Recnum              0
     Cardnum             0
     Date                0
     Merchnum            0
     Merch description   0
     Merch state         0
     Merch zip           0
     Transtype           0
     Amount              0
     Fraud               0
     dtype: int64
```

0.2.1 Filling Merchnum

```python
[8]: data['Merchnum'].replace({'0':np.nan},inplace = True)
     data.isnull().sum()
```

```
[8]: Recnum                 0
     Cardnum                0
     Date                   0
     Merchnum            3251
     Merch description      0
     Merch state            0
     Merch zip              0
     Transtype              0
     Amount                 0
     Fraud                  0
     dtype: int64
```

```python
[9]: data['Merchnum']=data.groupby(['Merch description','Merch zip'])['Merchnum'].
     ↪transform(lambda x:x.fillna(x.mode()[0]) if not x.mode().empty else 'Empty')
     data['Merchnum'].replace({'Empty':np.nan},inplace = True)
     data['Merchnum']=data.groupby(['Merch description','Merch state'])['Merchnum'].
     ↪transform(lambda x:x.fillna(x.mode()[0]) if not x.mode().empty else 'Empty')
     data['Merchnum'].replace({'Empty':np.nan},inplace = True)
     data['Merchnum']=data.groupby('Merch description')['Merchnum'].transform(lambda␣
     ↪x:x.fillna(x.mode()[0]) if not x.mode().empty else 'Empty')
     data['Merchnum']=np.
     ↪where(data['Merchnum']=='Empty',data['Recnum']*(-1),data['Merchnum'])
```

```
data.isnull().sum()
```

[9]:
```
Recnum              0
Cardnum             0
Date                0
Merchnum            0
Merch description   0
Merch state         0
Merch zip           0
Transtype           0
Amount              0
Fraud               0
dtype: int64
```

0.2.2 Filling State & Zip

[10]:
```
data['Merch state'].replace({'0':np.nan},inplace = True)
data['Merch zip'].replace({0:np.nan},inplace = True)
data.isnull().sum()
```

[10]:
```
Recnum                 0
Cardnum                0
Date                   0
Merchnum               0
Merch description      0
Merch state         1020
Merch zip           4300
Transtype              0
Amount                 0
Fraud                  0
dtype: int64
```

[11]:
```
data['Merch state']=data.groupby('Merchnum')['Merch state'].transform(lambda x:
 →x.fillna(x.mode()[0]) if not x.mode().empty else 'Empty')
data['Merch zip']=data.groupby('Merchnum')['Merch zip'].transform(lambda x:x.
 →fillna(x.mode()[0]) if not x.mode().empty else 'Empty')
data['Merch state']=np.where(data['Merch␣
 →state']=='Empty',data['Recnum']*(-1),data['Merch state'])
data['Merch zip']=np.where(data['Merch␣
 →zip']=='Empty',data['Recnum']*(-1),data['Merch zip'])
data.isnull().sum()
```

[11]:
```
Recnum      0
Cardnum     0
Date        0
Merchnum    0
```

```
Merch description    0
Merch state          0
Merch zip            0
Transtype            0
Amount               0
Fraud                0
dtype: int64
```

[12]: `data.shape`

[12]: (96397, 10)

[13]: `#data.to_csv("Cleaned Data.csv")`

```
In [1]:  # Importing necessary modules
         import pandas as pd
         import numpy as np
         import matplotlib.pyplot as plt
         import datetime as dt
         import warnings
         warnings.filterwarnings("ignore")
```

```
In [2]:  # Reading Card Transactions Raw Data
         df1 = pd.read_csv('Cleaned Data.csv')
```

```
In [3]:  df1.shape
```

```
Out[3]:  (96397, 10)
```

```
In [4]:  # Converting to right format
         df1['Date'] = pd.to_datetime(df1['Date'])
         df1.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 96397 entries, 0 to 96396
Data columns (total 10 columns):
 #   Column            Non-Null Count  Dtype
---  ------            --------------  -----
 0   Recnum            96397 non-null  int64
 1   Cardnum           96397 non-null  int64
 2   Date              96397 non-null  datetime64[ns]
 3   Merchnum          96397 non-null  object
 4   Merch description 96397 non-null  object
 5   Merch state       96397 non-null  object
 6   Merch zip         96397 non-null  int64
 7   Transtype         96397 non-null  object
 8   Amount            96397 non-null  float64
 9   Fraud             96397 non-null  int64
dtypes: datetime64[ns](1), float64(1), int64(4), object(4)
memory usage: 7.4+ MB
```

```
In [5]:  df1['Cardnum'] = df1['Cardnum'].astype(str)
         df1['Merch zip'] = df1['Merch zip'].astype(str)
         df1.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 96397 entries, 0 to 96396
Data columns (total 10 columns):
 #   Column             Non-Null Count  Dtype
---  ------             --------------  -----
 0   Recnum             96397 non-null  int64
 1   Cardnum            96397 non-null  object
 2   Date               96397 non-null  datetime64[ns]
 3   Merchnum           96397 non-null  object
 4   Merch description  96397 non-null  object
 5   Merch state        96397 non-null  object
 6   Merch zip          96397 non-null  object
 7   Transtype          96397 non-null  object
 8   Amount             96397 non-null  float64
 9   Fraud              96397 non-null  int64
dtypes: datetime64[ns](1), float64(1), int64(2), object(6)
memory usage: 7.4+ MB
```

Variable Creation

```
In [6]:  df1 = df1.set_index('Date')
         df1.head()
```

Out[6]:

Date	Recnum	Cardnum	Merchnum	Merch description	Merch state	Merch zip	Transtype	Amount
2010-01-01	1	5142190439	5509006296254	FEDEX SHP 12/23/09 AB#	TN	38118	P	3.62
2010-01-01	2	5142183973	61003026333	SERVICE MERCHANDISE #81	MA	1803	P	31.42
2010-01-01	3	5142131721	4503082993600	OFFICE DEPOT #191	MD	20706	P	178.49
2010-01-01	4	5142148452	5509006296254	FEDEX SHP 12/28/09 AB#	TN	38118	P	3.62
2010-01-01	5	5142190439	5509006296254	FEDEX SHP 12/23/09 AB#	TN	38118	P	3.62

```
In [7]:  #Calculates all the amount variables (200)

         df1 = df1.reset_index()
         df1 = df1.sort_values(by = ['Cardnum', 'Date'])
         df1_index = df1.set_index('Date')
         for days in ['1d', '3d', '7d','14d','30d']:
             df1['mean_amt_{}_{}'.format('card',days)] = getattr(df1_index.groupby('C
         ardnum')['Amount'].rolling(days),'mean')().values
             df1['max_amt_{}_{}'.format('card',days)] = getattr(df1_index.groupby('Ca
         rdnum')['Amount'].rolling(days),'max')().values
             df1['median_amt_{}_{}'.format('card',days)] = getattr(df1_index.groupby(
         'Cardnum')['Amount'].rolling(days),'median')().values
             df1['sum_amt_{}_{}'.format('card',days)] = getattr(df1_index.groupby('Ca
         rdnum')['Amount'].rolling(days),'sum')().values
         df1.sort_values(by = ['Recnum'],inplace = True)


         df1 = df1.reset_index(drop = True)
         df1 = df1.sort_values(by = ['Merchnum', 'Date'])
         df1_index = df1.set_index('Date')
         for days in ['1d', '3d', '7d','14d','30d']:
             df1['mean_amt_{}_{}'.format('merch',days)] = getattr(df1_index.groupby(
         'Merchnum')['Amount'].rolling(days),'mean')().values
             df1['max_amt_{}_{}'.format('merch',days)] = getattr(df1_index.groupby('M
         erchnum')['Amount'].rolling(days),'max')().values
             df1['median_amt_{}_{}'.format('merch',days)] = getattr(df1_index.groupby
         ('Merchnum')['Amount'].rolling(days),'median')().values
             df1['sum_amt_{}_{}'.format('merch',days)] = getattr(df1_index.groupby('M
         erchnum')['Amount'].rolling(days),'sum')().values
         df1.sort_values(by = ['Recnum'],inplace = True)


         df1 = df1.reset_index(drop = True)
         df1 = df1.sort_values(by = ['Cardnum','Merchnum', 'Date'])
         df1_index = df1.set_index('Date')
         for days in ['1d', '3d', '7d','14d','30d']:
             df1['mean_amt_{}_{}'.format('cardmerch',days)] = getattr(df1_index.group
         by(['Cardnum','Merchnum'])['Amount'].rolling(days),'mean')().values
             df1['max_amt_{}_{}'.format('cardmerch',days)] = getattr(df1_index.groupb
         y(['Cardnum','Merchnum'])['Amount'].rolling(days),'max')().values
             df1['median_amt_{}_{}'.format('cardmerch',days)] = getattr(df1_index.gro
         upby(['Cardnum','Merchnum'])['Amount'].rolling(days),'median')().values
             df1['sum_amt_{}_{}'.format('cardmerch',days)] = getattr(df1_index.groupb
         y(['Cardnum','Merchnum'])['Amount'].rolling(days),'sum')().values
         df1.sort_values(by = ['Recnum'],inplace = True)


         df1 = df1.reset_index(drop = True)
         df1 = df1.sort_values(by = ['Cardnum','Merch zip', 'Date'])
         df1_index = df1.set_index('Date')
         for days in ['1d', '3d', '7d','14d','30d']:
             df1['mean_amt_{}_{}'.format('cardzip',days)] = getattr(df1_index.groupby
         (['Cardnum','Merch zip'])['Amount'].rolling(days),'mean')().values
             df1['max_amt_{}_{}'.format('cardzip',days)] = getattr(df1_index.groupby
         (['Cardnum','Merch zip'])['Amount'].rolling(days),'max')().values
             df1['median_amt_{}_{}'.format('cardzip',days)] = getattr(df1_index.group
```

```python
by(['Cardnum','Merch zip'])['Amount'].rolling(days),'median')().values
        df1['sum_amt_{}_{}'.format('cardzip',days)] = getattr(df1_index.groupby
(['Cardnum','Merch zip'])['Amount'].rolling(days),'sum')().values
df1.sort_values(by = ['Recnum'],inplace = True)


df1 = df1.reset_index(drop = True)
df1 = df1.sort_values(by = ['Cardnum','Merch state', 'Date'])
df1_index = df1.set_index('Date')
for days in ['1d', '3d', '7d','14d','30d']:
        df1['mean_amt_{}_{}'.format('cardstate',days)] = getattr(df1_index.group
by(['Cardnum','Merch state'])['Amount'].rolling(days),'mean')().values
        df1['max_amt_{}_{}'.format('cardstate',days)] = getattr(df1_index.groupb
y(['Cardnum','Merch state'])['Amount'].rolling(days),'max')().values
        df1['median_amt_{}_{}'.format('cardstate',days)] = getattr(df1_index.gro
upby(['Cardnum','Merch state'])['Amount'].rolling(days),'median')().values
        df1['sum_amt_{}_{}'.format('cardstate',days)] = getattr(df1_index.groupb
y(['Cardnum','Merch state'])['Amount'].rolling(days),'sum')().values
df1.sort_values(by = ['Recnum'],inplace = True)

for days in ['1d', '3d', '7d','14d','30d']:
        df1['ratio_actual_mean_amt_{}_{}'.format('card',days)] = df1['Amount']/d
f1['mean_amt_{}_{}'.format('card',days)]
        df1['ratio_actual_mean_amt_{}_{}'.format('merch',days)] = df1['Amount']/
df1['mean_amt_{}_{}'.format('merch',days)]
        df1['ratio_actual_mean_amt_{}_{}'.format('cardmerch',days)] = df1['Amoun
t']/df1['mean_amt_{}_{}'.format('cardmerch',days)]
        df1['ratio_actual_mean_amt_{}_{}'.format('cardzip',days)] = df1['Amount'
]/df1['mean_amt_{}_{}'.format('cardzip',days)]
        df1['ratio_actual_mean_amt_{}_{}'.format('cardstate',days)] = df1['Amoun
t']/df1['mean_amt_{}_{}'.format('cardstate',days)]

        df1['ratio_actual_max_amt_{}_{}'.format('card',days)] = df1['Amount']/df
1['max_amt_{}_{}'.format('card',days)]
        df1['ratio_actual_max_amt_{}_{}'.format('merch',days)] = df1['Amount']/d
f1['max_amt_{}_{}'.format('merch',days)]
        df1['ratio_actual_max_amt_{}_{}'.format('cardmerch',days)] = df1['Amoun
t']/df1['max_amt_{}_{}'.format('cardmerch',days)]
        df1['ratio_actual_max_amt_{}_{}'.format('cardzip',days)] = df1['Amount']
/df1['max_amt_{}_{}'.format('cardzip',days)]
        df1['ratio_actual_max_amt_{}_{}'.format('cardstate',days)] = df1['Amoun
t']/df1['max_amt_{}_{}'.format('cardstate',days)]

        df1['ratio_actual_median_amt_{}_{}'.format('card',days)] = df1['Amount']
/df1['median_amt_{}_{}'.format('card',days)]
        df1['ratio_actual_median_amt_{}_{}'.format('merch',days)] = df1['Amount'
]/df1['median_amt_{}_{}'.format('merch',days)]
        df1['ratio_actual_median_amt_{}_{}'.format('cardmerch',days)] = df1['Amo
unt']/df1['median_amt_{}_{}'.format('cardmerch',days)]
        df1['ratio_actual_median_amt_{}_{}'.format('cardzip',days)] = df1['Amoun
t']/df1['median_amt_{}_{}'.format('cardzip',days)]
        df1['ratio_actual_median_amt_{}_{}'.format('cardstate',days)] = df1['Amo
unt']/df1['median_amt_{}_{}'.format('cardstate',days)]

        df1['ratio_actual_sum_amt_{}_{}'.format('card',days)] = df1['Amount']/df
1['sum_amt_{}_{}'.format('card',days)]
        df1['ratio_actual_sum_amt_{}_{}'.format('merch',days)] = df1['Amount']/d
```

```
f1['sum_amt_{}_{}'.format('merch',days)]
        df1['ratio_actual_sum_amt_{}_{}'.format('cardmerch',days)] = df1['Amoun
t']/df1['sum_amt_{}_{}'.format('cardmerch',days)]
        df1['ratio_actual_sum_amt_{}_{}'.format('cardzip',days)] = df1['Amount']
/df1['sum_amt_{}_{}'.format('cardzip',days)]
        df1['ratio_actual_sum_amt_{}_{}'.format('cardstate',days)] = df1['Amoun
t']/df1['sum_amt_{}_{}'.format('cardstate',days)]
```

In [8]:
```
df1.to_csv("Amount_Variables_Final.csv")
```

In [14]:
```
pd.set_option('display.max_columns', 500)
df1.head()
```

Out[14]:

	Date	Recnum	Cardnum	Merchnum	Merch description	Merch state	Merch zip	Transtype	Amou
0	2010-01-01	1	5142190439	5509006296254	FEDEX SHP 12/23/09 AB#	TN	38118	P	3.
1	2010-01-01	2	5142183973	61003026333	SERVICE MERCHANDISE #81	MA	1803	P	31.
2	2010-01-01	3	5142131721	4503082993600	OFFICE DEPOT #191	MD	20706	P	178.
3	2010-01-01	4	5142148452	5509006296254	FEDEX SHP 12/28/09 AB#	TN	38118	P	3.
4	2010-01-01	5	5142190439	5509006296254	FEDEX SHP 12/23/09 AB#	TN	38118	P	3.

In [15]:
```python
df1 = pd.read_csv('Cleaned Data.csv')
df1['Date'] = pd.to_datetime(df1['Date'])
df1['Cardnum'] = df1['Cardnum'].astype(str)
df1['Merch zip'] = df1['Merch zip'].astype(str)
df1 = df1.set_index('Date')
df1.head()
```

Out[15]:

Date	Recnum	Cardnum	Merchnum	Merch description	Merch state	Merch zip	Transtype	Amount
2010-01-01	1	5142190439	5509006296254	FEDEX SHP 12/23/09 AB#	TN	38118	P	3.62
2010-01-01	2	5142183973	61003026333	SERVICE MERCHANDISE #81	MA	1803	P	31.42
2010-01-01	3	5142131721	4503082993600	OFFICE DEPOT #191	MD	20706	P	178.49
2010-01-01	4	5142148452	5509006296254	FEDEX SHP 12/28/09 AB#	TN	38118	P	3.62
2010-01-01	5	5142190439	5509006296254	FEDEX SHP 12/23/09 AB#	TN	38118	P	3.62

In [17]:
```python
#Calculates all the frequence variables - 25

df1 = df1.reset_index()
df1 = df1.sort_values(by = ['Cardnum', 'Date'])
df1_index = df1.set_index('Date')
for days in ['1d', '3d', '7d','14d','30d']:
    df1['vel_{}_{}'.format('card',days)] = getattr(df1_index.groupby('Cardnu
m')['Recnum'].rolling(days),'count')().values
df1.sort_values(by = ['Recnum'],inplace = True)

df1 = df1.reset_index(drop = True)
df1 = df1.sort_values(by = ['Merchnum', 'Date'])
df1_index = df1.set_index('Date')
for days in ['1d', '3d', '7d','14d','30d']:
    df1['vel_{}_{}'.format('merch',days)] = getattr(df1_index.groupby('Merchn
um')['Recnum'].rolling(days),'count')().values
df1.sort_values(by = ['Recnum'],inplace = True)

df1 = df1.reset_index(drop = True)
df1 = df1.sort_values(by = ['Cardnum','Merchnum', 'Date'])
df1_index = df1.set_index('Date')
for days in ['1d', '3d', '7d','14d','30d']:
    df1['vel_{}_{}'.format('cardmerch',days)] = getattr(df1_index.groupby(['C
ardnum','Merchnum'])['Recnum'].rolling(days),'count')().values
df1.sort_values(by = ['Recnum'],inplace = True)

df1 = df1.reset_index(drop = True)
df1 = df1.sort_values(by = ['Cardnum','Merch zip', 'Date'])
df1_index = df1.set_index('Date')
for days in ['1d', '3d', '7d','14d','30d']:
    df1['vel_{}_{}'.format('cardzip',days)] = getattr(df1_index.groupby(['Car
dnum','Merch zip'])['Recnum'].rolling(days),'count')().values
df1.sort_values(by = ['Recnum'],inplace = True)

df1 = df1.reset_index(drop = True)
df1 = df1.sort_values(by = ['Cardnum','Merch state', 'Date'])
df1_index = df1.set_index('Date')
for days in ['1d', '3d', '7d','14d','30d']:
    df1['vel_{}_{}'.format('cardstate',days)] = getattr(df1_index.groupby(['C
ardnum','Merch state'])['Recnum'].rolling(days),'count')().values
df1.sort_values(by = ['Recnum'],inplace = True)
```

In [18]:
```python
df1.to_csv("Frequency_Variables_Final.csv")
```

Project 3 Code - Variable Creation Part3

February 12, 2021

```python
[1]: # Importing necessary modules
     import pandas as pd
     import numpy as np
     import matplotlib.pyplot as plt
     import datetime as dt
     import warnings
     warnings.filterwarnings("ignore")
```

```python
[2]: amount = pd.read_csv("Amount_Variables_Final.csv")
     frequency= pd.read_csv("Frequency_Variables_Final.csv")
     frequency = frequency.iloc[:,10:36]
```

```python
[3]: df1 = pd.concat([amount, frequency], axis=1, join='inner')
     df1.head()
```

```
[3]:      Date  Recnum     Cardnum       Merchnum          Merch description  \
     0  1/1/2010       1  5142190439  5510000000000    FEDEX SHP 12/23/09 AB#
     1  1/1/2010       2  5142183973    61003026333  SERVICE MERCHANDISE #81
     2  1/1/2010       3  5142131721  4500000000000          OFFICE DEPOT #191
     3  1/1/2010       4  5142148452  5510000000000    FEDEX SHP 12/28/09 AB#
     4  1/1/2010       5  5142190439  5510000000000    FEDEX SHP 12/23/09 AB#

       Merch state  Merch zip Transtype  Amount  Fraud  …  vel_cardzip_1d  \
     0          TN      38118         P    3.62      0  …               1
     1          MA       1803         P   31.42      0  …               1
     2          MD      20706         P  178.49      0  …               1
     3          TN      38118         P    3.62      0  …               1
     4          TN      38118         P    3.62      0  …               2

        vel_cardzip_3d  vel_cardzip_7d  vel_cardzip_14d  vel_cardzip_30d  \
     0               1               1                1                1
     1               1               1                1                1
     2               1               1                1                1
     3               1               1                1                1
     4               2               2                2                2

        vel_cardstate_1d  vel_cardstate_3d  vel_cardstate_7d  vel_cardstate_14d  \
```

```
0                   1              1              1              1
1                   1              1              1              1
2                   1              1              1              1
3                   1              1              1              1
4                   2              2              2              2

   vel_cardstate_30
0                 1
1                 1
2                 1
3                 1
4                 2

[5 rows x 235 columns]
```

[4]: `df1.shape`

[4]: `(96397, 235)`

[6]:
```python
df1['Date'] = pd.to_datetime(df1['Date'])
df1['Cardnum'] = df1['Cardnum'].astype(str)
df1['Merch zip'] = df1['Merch zip'].astype(str)

import re
df1 = df1.rename(columns=lambda x: re.sub('1d','1',x))
df1 = df1.rename(columns=lambda x: re.sub('3d','3',x))
df1 = df1.rename(columns=lambda x: re.sub('7d','7',x))
df1 = df1.rename(columns=lambda x: re.sub('14d','14',x))
df1 = df1.rename(columns=lambda x: re.sub('30d','30',x))
df1.head()
```

[6]:
```
        Date  Recnum     Cardnum      Merchnum        Merch description  \
0 2010-01-01       1  5142190439  5510000000000   FEDEX SHP 12/23/09 AB#
1 2010-01-01       2  5142183973    61003026333  SERVICE MERCHANDISE #81
2 2010-01-01       3  5142131721  4500000000000          OFFICE DEPOT #191
3 2010-01-01       4  5142148452  5510000000000   FEDEX SHP 12/28/09 AB#
4 2010-01-01       5  5142190439  5510000000000   FEDEX SHP 12/23/09 AB#

  Merch state Merch zip Transtype  Amount  Fraud  …  vel_cardzip_1  \
0          TN     38118         P    3.62      0  …              1
1          MA      1803         P   31.42      0  …              1
2          MD     20706         P  178.49      0  …              1
3          TN     38118         P    3.62      0  …              1
4          TN     38118         P    3.62      0  …              2

   vel_cardzip_3  vel_cardzip_7  vel_cardzip_14  vel_cardzip_30  \
0              1              1               1               1
```

```
1               1               1               1               1
2               1               1               1               1
3               1               1               1               1
4               2               2               2               2

   vel_cardstate_1  vel_cardstate_3  vel_cardstate_7  vel_cardstate_14  \
0                1                1                1                 1
1                1                1                1                 1
2                1                1                1                 1
3                1                1                1                 1
4                2                2                2                 2

   vel_cardstate_30
0                 1
1                 1
2                 1
3                 1
4                 2

[5 rows x 235 columns]
```

[7]:
```python
# Calculates all the velocity change variables - 48

for day in [7,14,30]:
    df1['rel_vel_vel_card_card_{}'.format(day)] = df1['vel_card_1']/
 ↪df1['vel_card_{}'.format(day)]
    df1['rel_vel_vel_card_merch_{}'.format(day)] = df1['vel_card_1']/
 ↪df1['vel_merch_{}'.format(day)]
    df1['rel_vel_amt_card_card_{}'.format(day)] = df1['vel_card_1']/
 ↪df1['sum_amt_card_{}'.format(day)]
    df1['rel_vel_amt_card_merch_{}'.format(day)] = df1['vel_card_1']/
 ↪df1['sum_amt_merch_{}'.format(day)]

    df1['rel_amt_vel_merch_card_{}'.format(day)] = df1['sum_amt_merch_1']/
 ↪df1['vel_card_{}'.format(day)]
    df1['rel_amt_vel_merch_merch_{}'.format(day)] = df1['sum_amt_merch_1']/
 ↪df1['vel_merch_{}'.format(day)]
    df1['rel_amt_amt_merch_card_{}'.format(day)] = df1['sum_amt_merch_1']/
 ↪df1['sum_amt_card_{}'.format(day)]
    df1['rel_amt_amt_merch_merch_{}'.format(day)] = df1['sum_amt_merch_1']/
 ↪df1['sum_amt_merch_{}'.format(day)]

    df1['rel_vel_vel_merch_card_{}'.format(day)] = df1['vel_merch_1']/
 ↪df1['vel_card_{}'.format(day)]
    df1['rel_vel_vel_merch_merch_{}'.format(day)] = df1['vel_merch_1']/
 ↪df1['vel_merch_{}'.format(day)]
```

```
    df1['rel_vel_amt_merch_card_{}'.format(day)] = df1['vel_merch_1']/␣
→df1['sum_amt_card_{}'.format(day)]
    df1['rel_vel_amt_merch_merch_{}'.format(day)] = df1['vel_merch_1']/␣
→df1['sum_amt_merch_{}'.format(day)]

    df1['rel_amt_vel_card_card_{}'.format(day)] = df1['sum_amt_card_1']/␣
→df1['vel_card_{}'.format(day)]
    df1['rel_amt_vel_card_merch_{}'.format(day)] = df1['sum_amt_card_1']/␣
→df1['vel_merch_{}'.format(day)]
    df1['rel_amt_amt_card_card_{}'.format(day)] = df1['sum_amt_card_1']/␣
→df1['sum_amt_card_{}'.format(day)]
    df1['rel_amt_amt_card_merch_{}'.format(day)] = df1['sum_amt_card_1']/␣
→df1['sum_amt_merch_{}'.format(day)]
```

[8]: `df1.shape`

[8]: (96397, 283)

[9]: `df1.head()`

[9]:
```
        Date  Recnum    Cardnum     Merchnum      Merch description  \
0 2010-01-01       1 5142190439 5510000000000   FEDEX SHP 12/23/09 AB#
1 2010-01-01       2 5142183973   61003026333 SERVICE MERCHANDISE #81
2 2010-01-01       3 5142131721 4500000000000        OFFICE DEPOT #191
3 2010-01-01       4 5142148452 5510000000000   FEDEX SHP 12/28/09 AB#
4 2010-01-01       5 5142190439 5510000000000   FEDEX SHP 12/23/09 AB#

  Merch state Merch zip Transtype  Amount  Fraud  … \
0          TN     38118         P    3.62      0 …
1          MA      1803         P   31.42      0 …
2          MD     20706         P  178.49      0 …
3          TN     38118         P    3.62      0 …
4          TN     38118         P    3.62      0 …

   rel_amt_amt_merch_card_30  rel_amt_amt_merch_merch_30  \
0                        1.0                         1.0
1                        1.0                         1.0
2                        1.0                         1.0
3                        2.0                         1.0
4                        1.5                         1.0

   rel_vel_vel_merch_card_30  rel_vel_vel_merch_merch_30  \
0                        1.0                         1.0
1                        1.0                         1.0
2                        1.0                         1.0
3                        2.0                         1.0
4                        1.5                         1.0
```

```
     rel_vel_amt_merch_card_30  rel_vel_amt_merch_merch_30  \
0                     0.276243                    0.276243
1                     0.031827                    0.031827
2                     0.005603                    0.005603
3                     0.552486                    0.276243
4                     0.414365                    0.276243


   rel_amt_vel_card_card_30  rel_amt_vel_card_merch_30  \
0                      3.62                   3.620000
1                     31.42                  31.420000
2                    178.49                 178.490000
3                      3.62                   1.810000
4                      3.62                   2.413333


   rel_amt_amt_card_card_30  rel_amt_amt_card_merch_30
0                       1.0                   1.000000
1                       1.0                   1.000000
2                       1.0                   1.000000
3                       1.0                   0.500000
4                       1.0                   0.666667

[5 rows x 283 columns]
```

```
[19]:  # df1.to_csv("Velocity_Change_Variables_Final.csv")
```

```
[11]:  # df1 = pd.read_csv('Cleaned Data.csv')
       # df1['Date'] = pd.to_datetime(df1['Date'])
       # df1['Cardnum'] = df1['Cardnum'].astype(str)
       # df1['Merch zip'] = df1['Merch zip'].astype(str)
       # df1.head()
```

```
[12]:  # # Calculates all the days since variables - 5
       # base_date = pd.to_datetime('2010-01-01')

       # variable_name='dayssince_card'
       # df1[variable_name] = df1.groupby('Cardnum')['Date'].diff().dt.days
       # df1['temp'] = (df1.Date - base_date)
       # df1['temp'] = df1.apply(lambda x: x.temp.days, axis=1)
       # df1['temp'] = df1['temp'].astype(float)
       # df1['temp'][df1['temp'] <20] = 19
       # df1[variable_name] = df1[variable_name].replace(np.nan,df1['temp']+1)
       # df1 = df1.drop('temp',axis = 1)

       # variable_name='dayssince_merch'
       # df1[variable_name] = df1.groupby('Merchnum')['Date'].diff().dt.days
       # df1['temp'] = (df1.Date - base_date)
```

```python
# df1['temp'] = df1.apply(lambda x: x.temp.days, axis=1)
# df1['temp'] = df1['temp'].astype(float)
# df1['temp'][df1['temp'] <20] = 19
# df1[variable_name] = df1[variable_name].replace(np.nan,df1['temp']+1)
# df1 = df1.drop('temp',axis = 1)


# variable_name='dayssince_cardmerch'
# df1[variable_name] = df1.groupby(['Cardnum','Merchnum'])['Date'].diff().dt.
 ↪days
# df1['temp'] = (df1.Date - base_date)
# df1['temp'] = df1.apply(lambda x: x.temp.days, axis=1)
# df1['temp'] = df1['temp'].astype(float)
# df1['temp'][df1['temp'] <20] = 19
# df1[variable_name] = df1[variable_name].replace(np.nan,df1['temp']+1)
# df1 = df1.drop('temp',axis = 1)


# variable_name='dayssince_cardzip'
# df1[variable_name] = df1.groupby(['Cardnum','Merch zip'])['Date'].diff().dt.
 ↪days
# df1['temp'] = (df1.Date - base_date)
# df1['temp'] = df1.apply(lambda x: x.temp.days, axis=1)
# df1['temp'] = df1['temp'].astype(float)
# df1['temp'][df1['temp'] <20] = 19
# df1[variable_name] = df1[variable_name].replace(np.nan,df1['temp']+1)
# df1 = df1.drop('temp',axis = 1)


# variable_name='dayssince_cardstate'
# df1[variable_name] = df1.groupby(['Cardnum','Merch state'])['Date'].diff().dt.
 ↪days
# df1['temp'] = (df1.Date - base_date)
# df1['temp'] = df1.apply(lambda x: x.temp.days, axis=1)
# df1['temp'] = df1['temp'].astype(float)
# df1['temp'][df1['temp'] <20] = 19
# df1[variable_name] = df1[variable_name].replace(np.nan,df1['temp']+1)
# df1 = df1.drop('temp',axis = 1)
```

```python
[13]: # # Subtracts days since variable from max to keep it consistent (higher number␣
      ↪means bad and lower number means good now)
      # days_since_columns =␣
      ↪['dayssince_card','dayssince_merch','dayssince_cardmerch','dayssince_cardzip','dayssince_ca

      # for i in days_since_columns:
      #     maximum = df1[i].max()
      #     df1[i] = maximum - df1[i]
```

```python
[14]: # df1.shape
```

```
[15]:  # df1.head()
```

```
[16]:  # df1.to_csv("Dayssince_Variables_Final.csv")
```

```
In [2]:  # Importing necessary modules
         import pandas as pd
         import numpy as np
         import matplotlib.pyplot as plt
         import datetime as dt
         import warnings
         warnings.filterwarnings("ignore")
```

```
In [3]:  amount = pd.read_csv("Amount_Variables_Final.csv")
         frequency= pd.read_csv("Frequency_Variables_Final.csv")
         days = pd.read_csv("Dayssince_Variables_Final.csv")
         freq_change = pd.read_csv("Velocity_Change_Variables_Final.csv")
```

```
In [4]:  amount.shape
```
```
Out[4]:  (96397, 210)
```

```
In [5]:  frequency.shape
```
```
Out[5]:  (96397, 35)
```

```
In [6]:  days.shape
```
```
Out[6]:  (96397, 15)
```

```
In [7]:  freq_change.shape
```
```
Out[7]:  (96397, 58)
```

```
In [8]:  frequency = frequency.iloc[:,10:210]
         days = days.iloc[:,10:15]
         freq_change = freq_change.iloc[:,10:58]
```

In [9]:
```python
df1 = pd.concat([amount, frequency,freq_change,days], axis=1, join='inner')
df1.head()
```

Out[9]:

	Date	Recnum	Cardnum	Merchnum	Merch description	Merch state	Merch zip	Transtype	Am
0	1/1/2010	1	5142190439	5510000000000	FEDEX SHP 12/23/09 AB#	TN	38118	P	
1	1/1/2010	2	5142183973	61003026333	SERVICE MERCHANDISE #81	MA	1803	P	
2	1/1/2010	3	5142131721	4500000000000	OFFICE DEPOT #191	MD	20706	P	1
3	1/1/2010	4	5142148452	5510000000000	FEDEX SHP 12/28/09 AB#	TN	38118	P	
4	1/1/2010	5	5142190439	5510000000000	FEDEX SHP 12/23/09 AB#	TN	38118	P	

5 rows × 288 columns

In [10]:
```python
df1['Date'] = pd.to_datetime(df1['Date'])
df1['Cardnum'] = df1['Cardnum'].astype(str)
df1['Merch zip'] = df1['Merch zip'].astype(str)

import re
df1 = df1.rename(columns=lambda x: re.sub('1d','1',x))
df1 = df1.rename(columns=lambda x: re.sub('3d','3',x))
df1 = df1.rename(columns=lambda x: re.sub('7d','7',x))
df1 = df1.rename(columns=lambda x: re.sub('14d','14',x))
df1 = df1.rename(columns=lambda x: re.sub('30d','30',x))
df1.head()
```

Out[10]:

	Date	Recnum	Cardnum	Merchnum	Merch description	Merch state	Merch zip	Transtype	Amou
0	2010-01-01	1	5142190439	5510000000000	FEDEX SHP 12/23/09 AB#	TN	38118	P	3.
1	2010-01-01	2	5142183973	61003026333	SERVICE MERCHANDISE #81	MA	1803	P	31.
2	2010-01-01	3	5142131721	4500000000000	OFFICE DEPOT #191	MD	20706	P	178.
3	2010-01-01	4	5142148452	5510000000000	FEDEX SHP 12/28/09 AB#	TN	38118	P	3.
4	2010-01-01	5	5142190439	5510000000000	FEDEX SHP 12/23/09 AB#	TN	38118	P	3.

5 rows × 288 columns

In [11]:
```python
df1.shape
```

Out[11]: (96397, 288)

In [12]:
```python
# Creating day of week target encoding (excluding OOT)
df1['dayofweek'] = df1['Date'].dt.day_name()
risk = {'Monday': 0.0087,
        'Tuesday':0.0071,
        'Wednesday':0.0098,
        'Thursday':0.0186,
        'Friday':0.0260,
        'Saturday':0.0101,
        'Sunday':0.0096}
df1['risk_dayofweek'] = df1['dayofweek'].map(risk)
df1.head()
```

Out[12]:

	Date	Recnum	Cardnum	Merchnum	Merch description	Merch state	Merch zip	Transtype	Amou
0	2010-01-01	1	5142190439	5510000000000	FEDEX SHP 12/23/09 AB#	TN	38118	P	3.
1	2010-01-01	2	5142183973	61003026333	SERVICE MERCHANDISE #81	MA	1803	P	31.
2	2010-01-01	3	5142131721	4500000000000	OFFICE DEPOT #191	MD	20706	P	178.
3	2010-01-01	4	5142148452	5510000000000	FEDEX SHP 12/28/09 AB#	TN	38118	P	3.
4	2010-01-01	5	5142190439	5510000000000	FEDEX SHP 12/23/09 AB#	TN	38118	P	3.

5 rows × 290 columns

```
In [13]:   # Creating Benford Variables
           data = df1.copy()
           data2 = df1.copy()

           data.drop(data[data['Merch description'].str.contains('FEDEX',regex = False)].
           index,inplace = True)
           data.drop(data[data['Merch description'].str.contains('FED EX',regex = False)]
           .index,inplace = True)
           data.drop(data[data['Merch description'].str.contains('FEDERAL EXPRESS',regex
           = False)].index,inplace = True)
           data.drop(data[data['Merch description'].str.contains('FEDERAL EX',regex = Fal
           se)].index,inplace = True)

           data2.drop(data[data['Merch description'].str.contains('FEDEX',regex = False)]
           .index,inplace = True)
           data2.drop(data[data['Merch description'].str.contains('FED EX',regex = False
           )].index,inplace = True)
           data2.drop(data[data['Merch description'].str.contains('FEDERAL EXPRESS',regex
           = False)].index,inplace = True)
           data2.drop(data[data['Merch description'].str.contains('FEDERAL EX',regex = Fa
           lse)].index,inplace = True)


           # Cardnumber
           data['Amount']=data['Amount'].astype('str')
           grouped_df=data.groupby('Cardnum')['Amount']
           gb=grouped_df.groups

           c=pd.DataFrame()
           e=pd.DataFrame()
           for key,values in gb.items():
               d=pd.DataFrame(data.loc[values]['Amount'].transform(lambda x:x.startswith(
           '1')).cumsum())
               c=c.append(d)
               f=pd.DataFrame(data.loc[values]['Amount'].transform(lambda x:x.startswith(
           '2')).cumsum())
               e=e.append(f)

           both=c.merge(e,right_index=True,left_index=True)
           both.columns=['ones','twos']
           both['n_low']=both['ones']+both['twos']
           both['n']=data.groupby('Cardnum')['Amount'].cumcount()+1
           data=both.merge(data,right_index=True,left_index=True)
           data['n_high']=data['n']-data['n_low']
           data['n_low']=np.where(data['n_low']==0,1,data['n_low'])
           data['n_high']=np.where(data['n_high']==0,1,data['n_high'])
           data['R']=data['n_low']*1.096/data['n_high']
           data['1/R']=1/(data['R'])
           data['U']=data[['R','1/R']].max(axis=1)
           data['t']=(data['n']-15)/3
           data['U*']=(1+(data['U']-1)/(1+np.exp(-data['t'])))

           # Merchant Number
           data2['Amount']=data2['Amount'].astype('str')
           grouped_df=data2.groupby('Merchnum')['Amount']
           gb=grouped_df.groups
```

```python
c=pd.DataFrame()
e=pd.DataFrame()
for key,values in gb.items():
    d=pd.DataFrame(data2.loc[values]['Amount'].transform(lambda x:x.startswith
('1')).cumsum())
    c=c.append(d)
    f=pd.DataFrame(data2.loc[values]['Amount'].transform(lambda x:x.startswith
('2')).cumsum())
    e=e.append(f)

both=c.merge(e,right_index=True,left_index=True)
both.columns=['ones','twos']
both['n_low']=both['ones']+both['twos']
both['n']=data2.groupby('Merchnum')['Amount'].cumcount()+1
data2=both.merge(data2,right_index=True,left_index=True)
data2['n_high']=data2['n']-data2['n_low']
data2['n_low']=np.where(data2['n_low']==0,1,data2['n_low'])
data2['n_high']=np.where(data2['n_high']==0,1,data2['n_high'])
data2['R']=data2['n_low']*1.096/data2['n_high']
data2['1/R']=1/(data2['R'])
data2['U']=data2[['R','1/R']].max(axis=1)
data2['t']=(data2['n']-15)/3
data2['U*']=(1+(data2['U']-1)/(1+np.exp(-data2['t'])))

df1 = df1.merge(data[['Recnum','U*']],how='left')
df1 = df1.rename(columns={'U*': 'card_U*'})
df1 = df1.merge(data2[['Recnum','U*']],how='left')
df1 = df1.rename(columns={'U*': 'merch_U*'})
```

In [14]: `df1.head()`

Out[14]:

	Date	Recnum	Cardnum	Merchnum	Merch description	Merch state	Merch zip	Transtype	Amou
0	2010-01-01	1	5142190439	5510000000000	FEDEX SHP 12/23/09 AB#	TN	38118	P	3.
1	2010-01-01	2	5142183973	61003026333	SERVICE MERCHANDISE #81	MA	1803	P	31.
2	2010-01-01	3	5142131721	4500000000000	OFFICE DEPOT #191	MD	20706	P	178.
3	2010-01-01	4	5142148452	5510000000000	FEDEX SHP 12/28/09 AB#	TN	38118	P	3.
4	2010-01-01	5	5142190439	5510000000000	FEDEX SHP 12/23/09 AB#	TN	38118	P	3.

5 rows × 292 columns

In [15]:
```python
# Dropping original variables
original_variables = ['Cardnum','Merchnum','Merch description',
                      'Merch state','Merch zip','Transtype','dayofweek','Amoun
t']

df1.drop(original_variables,axis = 1,inplace = True)
df1.head()
```

Out[15]:

	Date	Recnum	Fraud	mean_amt_card_1	max_amt_card_1	median_amt_card_1	sum_amt_car
0	2010-01-01	1	0	3.62	3.62	3.62	
1	2010-01-01	2	0	31.42	31.42	31.42	3
2	2010-01-01	3	0	178.49	178.49	178.49	17
3	2010-01-01	4	0	3.62	3.62	3.62	
4	2010-01-01	5	0	3.62	3.62	3.62	

5 rows × 284 columns

In [16]:
```python
# moving fraud_label to last
df1['fraud_label1'] = df1['Fraud']*1
df1.drop('Fraud',axis = 1,inplace = True)
df1.rename(columns={'fraud_label1': 'Fraud'},inplace = True)
pd.set_option('display.max_columns', None)
df1.head()
```

Out[16]:

	Date	Recnum	mean_amt_card_1	max_amt_card_1	median_amt_card_1	sum_amt_card_1	m
0	2010-01-01	1	3.62	3.62	3.62	3.62	
1	2010-01-01	2	31.42	31.42	31.42	31.42	
2	2010-01-01	3	178.49	178.49	178.49	178.49	
3	2010-01-01	4	3.62	3.62	3.62	3.62	
4	2010-01-01	5	3.62	3.62	3.62	7.24	

In [17]:
```python
df1.fillna(1, inplace=True)
df1.isnull().sum()
```

Out[17]:
```
Date                  0
Recnum                0
mean_amt_card_1       0
max_amt_card_1        0
median_amt_card_1     0
                     ..
dayssince_cardstate   0
risk_dayofweek        0
card_U*               0
merch_U*              0
Fraud                 0
Length: 284, dtype: int64
```

In [18]:
```python
a = list(df1.columns)
a
```

```
Out[18]: ['Date',
          'Recnum',
          'mean_amt_card_1',
          'max_amt_card_1',
          'median_amt_card_1',
          'sum_amt_card_1',
          'mean_amt_card_3',
          'max_amt_card_3',
          'median_amt_card_3',
          'sum_amt_card_3',
          'mean_amt_card_7',
          'max_amt_card_7',
          'median_amt_card_7',
          'sum_amt_card_7',
          'mean_amt_card_14',
          'max_amt_card_14',
          'median_amt_card_14',
          'sum_amt_card_14',
          'mean_amt_card_30',
          'max_amt_card_30',
          'median_amt_card_30',
          'sum_amt_card_30',
          'mean_amt_merch_1',
          'max_amt_merch_1',
          'median_amt_merch_1',
          'sum_amt_merch_1',
          'mean_amt_merch_3',
          'max_amt_merch_3',
          'median_amt_merch_3',
          'sum_amt_merch_3',
          'mean_amt_merch_7',
          'max_amt_merch_7',
          'median_amt_merch_7',
          'sum_amt_merch_7',
          'mean_amt_merch_14',
          'max_amt_merch_14',
          'median_amt_merch_14',
          'sum_amt_merch_14',
          'mean_amt_merch_30',
          'max_amt_merch_30',
          'median_amt_merch_30',
          'sum_amt_merch_30',
          'mean_amt_cardmerch_1',
          'max_amt_cardmerch_1',
          'median_amt_cardmerch_1',
          'sum_amt_cardmerch_1',
          'mean_amt_cardmerch_3',
          'max_amt_cardmerch_3',
          'median_amt_cardmerch_3',
          'sum_amt_cardmerch_3',
          'mean_amt_cardmerch_7',
          'max_amt_cardmerch_7',
          'median_amt_cardmerch_7',
          'sum_amt_cardmerch_7',
          'mean_amt_cardmerch_14',
          'max_amt_cardmerch_14',
          'median_amt_cardmerch_14',
```

```
'sum_amt_cardmerch_14',
'mean_amt_cardmerch_30',
'max_amt_cardmerch_30',
'median_amt_cardmerch_30',
'sum_amt_cardmerch_30',
'mean_amt_cardzip_1',
'max_amt_cardzip_1',
'median_amt_cardzip_1',
'sum_amt_cardzip_1',
'mean_amt_cardzip_3',
'max_amt_cardzip_3',
'median_amt_cardzip_3',
'sum_amt_cardzip_3',
'mean_amt_cardzip_7',
'max_amt_cardzip_7',
'median_amt_cardzip_7',
'sum_amt_cardzip_7',
'mean_amt_cardzip_14',
'max_amt_cardzip_14',
'median_amt_cardzip_14',
'sum_amt_cardzip_14',
'mean_amt_cardzip_30',
'max_amt_cardzip_30',
'median_amt_cardzip_30',
'sum_amt_cardzip_30',
'mean_amt_cardstate_1',
'max_amt_cardstate_1',
'median_amt_cardstate_1',
'sum_amt_cardstate_1',
'mean_amt_cardstate_3',
'max_amt_cardstate_3',
'median_amt_cardstate_3',
'sum_amt_cardstate_3',
'mean_amt_cardstate_7',
'max_amt_cardstate_7',
'median_amt_cardstate_7',
'sum_amt_cardstate_7',
'mean_amt_cardstate_14',
'max_amt_cardstate_14',
'median_amt_cardstate_14',
'sum_amt_cardstate_14',
'mean_amt_cardstate_30',
'max_amt_cardstate_30',
'median_amt_cardstate_30',
'sum_amt_cardstate_30',
'ratio_actual_mean_amt_card_1',
'ratio_actual_mean_amt_merch_1',
'ratio_actual_mean_amt_cardmerch_1',
'ratio_actual_mean_amt_cardzip_1',
'ratio_actual_mean_amt_cardstate_1',
'ratio_actual_max_amt_card_1',
'ratio_actual_max_amt_merch_1',
'ratio_actual_max_amt_cardmerch_1',
'ratio_actual_max_amt_cardzip_1',
'ratio_actual_max_amt_cardstate_1',
'ratio_actual_median_amt_card_1',
'ratio_actual_median_amt_merch_1',
```

```
            'ratio_actual_median_amt_cardmerch_1',
            'ratio_actual_median_amt_cardzip_1',
            'ratio_actual_median_amt_cardstate_1',
            'ratio_actual_sum_amt_card_1',
            'ratio_actual_sum_amt_merch_1',
            'ratio_actual_sum_amt_cardmerch_1',
            'ratio_actual_sum_amt_cardzip_1',
            'ratio_actual_sum_amt_cardstate_1',
            'ratio_actual_mean_amt_card_3',
            'ratio_actual_mean_amt_merch_3',
            'ratio_actual_mean_amt_cardmerch_3',
            'ratio_actual_mean_amt_cardzip_3',
            'ratio_actual_mean_amt_cardstate_3',
            'ratio_actual_max_amt_card_3',
            'ratio_actual_max_amt_merch_3',
            'ratio_actual_max_amt_cardmerch_3',
            'ratio_actual_max_amt_cardzip_3',
            'ratio_actual_max_amt_cardstate_3',
            'ratio_actual_median_amt_card_3',
            'ratio_actual_median_amt_merch_3',
            'ratio_actual_median_amt_cardmerch_3',
            'ratio_actual_median_amt_cardzip_3',
            'ratio_actual_median_amt_cardstate_3',
            'ratio_actual_sum_amt_card_3',
            'ratio_actual_sum_amt_merch_3',
            'ratio_actual_sum_amt_cardmerch_3',
            'ratio_actual_sum_amt_cardzip_3',
            'ratio_actual_sum_amt_cardstate_3',
            'ratio_actual_mean_amt_card_7',
            'ratio_actual_mean_amt_merch_7',
            'ratio_actual_mean_amt_cardmerch_7',
            'ratio_actual_mean_amt_cardzip_7',
            'ratio_actual_mean_amt_cardstate_7',
            'ratio_actual_max_amt_card_7',
            'ratio_actual_max_amt_merch_7',
            'ratio_actual_max_amt_cardmerch_7',
            'ratio_actual_max_amt_cardzip_7',
            'ratio_actual_max_amt_cardstate_7',
            'ratio_actual_median_amt_card_7',
            'ratio_actual_median_amt_merch_7',
            'ratio_actual_median_amt_cardmerch_7',
            'ratio_actual_median_amt_cardzip_7',
            'ratio_actual_median_amt_cardstate_7',
            'ratio_actual_sum_amt_card_7',
            'ratio_actual_sum_amt_merch_7',
            'ratio_actual_sum_amt_cardmerch_7',
            'ratio_actual_sum_amt_cardzip_7',
            'ratio_actual_sum_amt_cardstate_7',
            'ratio_actual_mean_amt_card_14',
            'ratio_actual_mean_amt_merch_14',
            'ratio_actual_mean_amt_cardmerch_14',
            'ratio_actual_mean_amt_cardzip_14',
            'ratio_actual_mean_amt_cardstate_14',
            'ratio_actual_max_amt_card_14',
            'ratio_actual_max_amt_merch_14',
            'ratio_actual_max_amt_cardmerch_14',
            'ratio_actual_max_amt_cardzip_14',
```

```
                'ratio_actual_max_amt_cardstate_14',
                'ratio_actual_median_amt_card_14',
                'ratio_actual_median_amt_merch_14',
                'ratio_actual_median_amt_cardmerch_14',
                'ratio_actual_median_amt_cardzip_14',
                'ratio_actual_median_amt_cardstate_14',
                'ratio_actual_sum_amt_card_14',
                'ratio_actual_sum_amt_merch_14',
                'ratio_actual_sum_amt_cardmerch_14',
                'ratio_actual_sum_amt_cardzip_14',
                'ratio_actual_sum_amt_cardstate_14',
                'ratio_actual_mean_amt_card_30',
                'ratio_actual_mean_amt_merch_30',
                'ratio_actual_mean_amt_cardmerch_30',
                'ratio_actual_mean_amt_cardzip_30',
                'ratio_actual_mean_amt_cardstate_30',
                'ratio_actual_max_amt_card_30',
                'ratio_actual_max_amt_merch_30',
                'ratio_actual_max_amt_cardmerch_30',
                'ratio_actual_max_amt_cardzip_30',
                'ratio_actual_max_amt_cardstate_30',
                'ratio_actual_median_amt_card_30',
                'ratio_actual_median_amt_merch_30',
                'ratio_actual_median_amt_cardmerch_30',
                'ratio_actual_median_amt_cardzip_30',
                'ratio_actual_median_amt_cardstate_30',
                'ratio_actual_sum_amt_card_30',
                'ratio_actual_sum_amt_merch_30',
                'ratio_actual_sum_amt_cardmerch_30',
                'ratio_actual_sum_amt_cardzip_30',
                'ratio_actual_sum_amt_cardstate_30',
                'vel_card_1',
                'vel_card_3',
                'vel_card_7',
                'vel_card_14',
                'vel_card_30',
                'vel_merch_1',
                'vel_merch_3',
                'vel_merch_7',
                'vel_merch_14',
                'vel_merch_30',
                'vel_cardmerch_1',
                'vel_cardmerch_3',
                'vel_cardmerch_7',
                'vel_cardmerch_14',
                'vel_cardmerch_30',
                'vel_cardzip_1',
                'vel_cardzip_3',
                'vel_cardzip_7',
                'vel_cardzip_14',
                'vel_cardzip_30',
                'vel_cardstate_1',
                'vel_cardstate_3',
                'vel_cardstate_7',
                'vel_cardstate_14',
                'vel_cardstate_30',
                'rel_vel_vel_card_card_7',
```

```
                    'rel_vel_vel_card_merch_7',
                    'rel_vel_amt_card_card_7',
                    'rel_vel_amt_card_merch_7',
                    'rel_amt_vel_merch_card_7',
                    'rel_amt_vel_merch_merch_7',
                    'rel_amt_amt_merch_card_7',
                    'rel_amt_amt_merch_merch_7',
                    'rel_vel_vel_merch_card_7',
                    'rel_vel_vel_merch_merch_7',
                    'rel_vel_amt_merch_card_7',
                    'rel_vel_amt_merch_merch_7',
                    'rel_amt_vel_card_card_7',
                    'rel_amt_vel_card_merch_7',
                    'rel_amt_amt_card_card_7',
                    'rel_amt_amt_card_merch_7',
                    'rel_vel_vel_card_card_14',
                    'rel_vel_vel_card_merch_14',
                    'rel_vel_amt_card_card_14',
                    'rel_vel_amt_card_merch_14',
                    'rel_amt_vel_merch_card_14',
                    'rel_amt_vel_merch_merch_14',
                    'rel_amt_amt_merch_card_14',
                    'rel_amt_amt_merch_merch_14',
                    'rel_vel_vel_merch_card_14',
                    'rel_vel_vel_merch_merch_14',
                    'rel_vel_amt_merch_card_14',
                    'rel_vel_amt_merch_merch_14',
                    'rel_amt_vel_card_card_14',
                    'rel_amt_vel_card_merch_14',
                    'rel_amt_amt_card_card_14',
                    'rel_amt_amt_card_merch_14',
                    'rel_vel_vel_card_card_30',
                    'rel_vel_vel_card_merch_30',
                    'rel_vel_amt_card_card_30',
                    'rel_vel_amt_card_merch_30',
                    'rel_amt_vel_merch_card_30',
                    'rel_amt_vel_merch_merch_30',
                    'rel_amt_amt_merch_card_30',
                    'rel_amt_amt_merch_merch_30',
                    'rel_vel_vel_merch_card_30',
                    'rel_vel_vel_merch_merch_30',
                    'rel_vel_amt_merch_card_30',
                    'rel_vel_amt_merch_merch_30',
                    'rel_amt_vel_card_card_30',
                    'rel_amt_vel_card_merch_30',
                    'rel_amt_amt_card_card_30',
                    'rel_amt_amt_card_merch_30',
                    'dayssince_card',
                    'dayssince_merch',
                    'dayssince_cardmerch',
                    'dayssince_cardzip',
                    'dayssince_cardstate',
                    'risk_dayofweek',
                    'card_U*',
                    'merch_U*',
                    'Fraud']
```

In [17]:
```python
# df1.to_csv('Final Variable Data.csv')
```

In [90]:
```python
# Importing necessary modules
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import datetime as dt
```

In [91]:
```python
data=pd.read_csv('Final Variable Data.csv',index_col = 0)
```

In [92]:
```python
data.shape
```

Out[92]: (96397, 284)

In [60]:
```python
pd.set_option('display.max_columns',None)
data.head(100)
```

Out[60]:

	Date	Recnum	mean_amt_card_1	max_amt_card_1	median_amt_card_1	sum_amt_card_1	n
0	2010-01-01	1	3.620	3.620	3.620	3.620	
1	2010-01-01	2	31.420	31.420	31.420	31.420	
2	2010-01-01	3	178.490	178.490	178.490	178.490	
3	2010-01-01	4	3.620	3.620	3.620	3.620	
4	2010-01-01	5	3.620	3.620	3.620	7.240	
5	2010-01-01	6	3.670	3.670	3.670	3.670	
6	2010-01-01	7	3.620	3.620	3.620	3.620	
7	2010-01-01	8	230.320	230.320	230.320	230.320	
8	2010-01-01	9	62.110	62.110	62.110	62.110	
9	2010-01-01	10	3.620	3.620	3.620	10.860	
10	2010-01-01	11	3.620	3.620	3.620	3.620	
11	2010-01-01	12	3.620	3.620	3.620	14.480	
12	2010-01-01	13	3.850	3.850	3.850	3.850	
13	2010-01-01	14	106.890	106.890	106.890	106.890	
14	2010-01-01	15	137.450	137.450	137.450	137.450	
15	2010-01-01	16	3.680	3.740	3.680	7.360	
16	2010-01-01	17	3.720	3.800	3.740	11.160	
17	2010-01-01	18	3.670	3.670	3.670	3.670	
18	2010-01-01	19	3.670	3.670	3.670	3.670	
19	2010-01-01	20	218.570	218.570	218.570	218.570	
20	2010-01-01	21	3.620	3.620	3.620	18.100	
21	2010-01-01	22	3.620	3.620	3.620	21.720	
22	2010-01-01	23	3.627	3.670	3.620	25.390	

	Date	Recnum	mean_amt_card_1	max_amt_card_1	median_amt_card_1	sum_amt_card_1	r
23	2010-01-01	24	68.900	68.900	68.900	68.900	
24	2010-01-01	25	3.670	3.670	3.670	3.670	
25	2010-01-01	26	3.620	3.620	3.620	3.620	
26	2010-01-01	27	360.000	360.000	360.000	360.000	
27	2010-01-01	28	3.626	3.670	3.620	29.010	
28	2010-01-01	29	427.000	427.000	427.000	427.000	
29	2010-01-01	30	3.740	3.800	3.770	14.960	
30	2010-01-01	31	3.620	3.620	3.620	3.620	
31	2010-01-01	32	504.600	504.600	504.600	504.600	
32	2010-01-01	33	228.000	228.000	228.000	228.000	
33	2010-01-01	34	3.626	3.670	3.620	32.630	
34	2010-01-01	35	3.620	3.620	3.620	3.620	
35	2010-01-01	36	3.740	3.740	3.740	3.740	
36	2010-01-01	37	3.740	3.800	3.740	18.700	
37	2010-01-01	38	23.610	23.610	23.610	23.610	
38	2010-01-01	39	3.620	3.620	3.620	3.620	
39	2010-01-01	40	3.625	3.670	3.620	36.250	
40	2010-01-01	41	656.985	1083.650	656.985	1313.970	
41	2010-01-01	42	3.720	3.800	3.740	22.320	
42	2010-01-01	43	3.620	3.620	3.620	3.620	
43	2010-01-01	44	3.731	3.800	3.740	26.120	
44	2010-01-01	45	76.300	76.300	76.300	76.300	
45	2010-01-01	46	3.620	3.620	3.620	3.620	

	Date	Recnum	mean_amt_card_1	max_amt_card_1	median_amt_card_1	sum_amt_card_1	r
46	2010-01-01	47	3.620	3.620	3.620	3.620	
47	2010-01-01	48	3.625	3.670	3.620	39.870	
48	2010-01-01	49	32.800	32.800	32.800	32.800	
49	2010-01-01	50	259.800	259.800	259.800	259.800	
50	2010-01-01	51	11.290	11.290	11.290	11.290	
51	2010-01-02	52	20.150	20.150	20.150	20.150	
52	2010-01-02	53	530.980	530.980	530.980	530.980	
53	2010-01-02	54	36.700	36.700	36.700	36.700	
54	2010-01-02	55	23.900	23.900	23.900	23.900	
55	2010-01-02	56	19.950	19.950	19.950	19.950	
56	2010-01-02	57	70.150	70.150	70.150	70.150	
57	2010-01-02	58	30.000	30.000	30.000	30.000	
58	2010-01-02	59	20.150	20.150	20.150	40.300	
59	2010-01-02	60	20.150	20.150	20.150	60.450	
60	2010-01-02	61	20.150	20.150	20.150	80.600	
61	2010-01-02	62	20.150	20.150	20.150	100.750	
62	2010-01-02	63	27.410	27.410	27.410	27.410	
63	2010-01-02	64	2.090	2.090	2.090	2.090	
64	2010-01-02	65	20.150	20.150	20.150	120.900	
65	2010-01-02	66	20.150	20.150	20.150	141.050	
66	2010-01-02	67	19.950	19.950	19.950	39.900	
67	2010-01-02	68	31.200	31.200	31.200	31.200	
68	2010-01-02	69	25.803	37.510	19.950	77.410	

	Date	Recnum	mean_amt_card_1	max_amt_card_1	median_amt_card_1	sum_amt_card_1	r
69	2010-01-02	70	24.340	37.510	19.950	97.360	
70	2010-01-02	71	67.750	67.750	67.750	67.750	
71	2010-01-02	72	21.148	28.130	20.150	169.180	
72	2010-01-02	73	21.037	28.130	20.150	189.330	
73	2010-01-02	74	277.950	277.950	277.950	277.950	
74	2010-01-02	75	131.880	131.880	131.880	131.880	
75	2010-01-02	76	19.955	27.410	19.955	39.910	
76	2010-01-02	77	277.950	277.950	277.950	277.950	
77	2010-01-02	78	20.948	28.130	20.150	209.480	
78	2010-01-02	79	101.400	101.400	101.400	101.400	
79	2010-01-02	80	19.950	19.950	19.950	19.950	
80	2010-01-03	81	10.710	10.710	10.710	10.710	
81	2010-01-03	82	20.160	20.160	20.160	20.160	
82	2010-01-03	83	300.890	300.890	300.890	300.890	
83	2010-01-03	84	37.490	37.490	37.490	37.490	
84	2010-01-03	85	169.000	169.000	169.000	169.000	
85	2010-01-03	86	90.950	90.950	90.950	90.950	
86	2010-01-03	87	3.990	3.990	3.990	3.990	
87	2010-01-03	88	55.800	55.800	55.800	55.800	
88	2010-01-03	89	39.950	39.950	39.950	39.950	
89	2010-01-03	90	160.380	160.380	160.380	160.380	
90	2010-01-03	91	156.190	160.380	156.190	312.380	
91	2010-01-03	92	59.000	59.000	59.000	59.000	

	Date	Recnum	mean_amt_card_1	max_amt_card_1	median_amt_card_1	sum_amt_card_1	r
92	2010-01-03	93	549.000	549.000	549.000	549.000	
93	2010-01-03	94	2835.000	2835.000	2835.000	2835.000	
94	2010-01-03	95	1343.000	1343.000	1343.000	1343.000	
95	2010-01-03	96	557.200	557.200	557.200	557.200	
96	2010-01-03	97	201.750	201.750	201.750	201.750	
97	2010-01-03	99	319.000	319.000	319.000	319.000	
98	2010-01-03	100	549.000	549.000	549.000	1098.000	
99	2010-01-03	101	549.000	549.000	549.000	549.000	

In [62]:
```python
data['Date'] = pd.to_datetime(data['Date'])
data.head()
```

Out[62]:

	Date	Recnum	mean_amt_card_1	max_amt_card_1	median_amt_card_1	sum_amt_card_1	m
0	2010-01-01	1	3.620	3.620	3.620	3.620	
1	2010-01-01	2	31.420	31.420	31.420	31.420	
2	2010-01-01	3	178.490	178.490	178.490	178.490	
3	2010-01-01	4	3.620	3.620	3.620	3.620	
4	2010-01-01	5	3.620	3.620	3.620	7.240	

In [63]:
```python
#Random
from random import random
rand=[]
for i in range(0,len(data)):
    rand.append(random())

data['random']=rand
```

In [64]: `data.head()`

Out[64]:

	Date	Recnum	mean_amt_card_1	max_amt_card_1	median_amt_card_1	sum_amt_card_1	m
0	2010-01-01	1	3.620	3.620	3.620	3.620	
1	2010-01-01	2	31.420	31.420	31.420	31.420	
2	2010-01-01	3	178.490	178.490	178.490	178.490	
3	2010-01-01	4	3.620	3.620	3.620	3.620	
4	2010-01-01	5	3.620	3.620	3.620	7.240	

In [65]:
```python
# Removing OOT and Initial 15 days to calculate KS
data=data[(data['Date']>'2010-01-15') & (data['Date']<'2010-11-1')]
data.drop(columns='Date',inplace=True)
from scipy.stats import ks_2samp
KSFDR=pd.DataFrame()
for i in list(data.columns):
    KSFDR[i]=ks_2samp(data.loc[data['Fraud']==0,'{}'.format(i)],data.loc[data['Fraud']==1,'{}'.format(i)])
```

```
C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\pandas\cor
e\frame.py:3997: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/s
table/user_guide/indexing.html#returning-a-view-versus-a-copy
  errors=errors,
```

In [66]: `KSFDR=KSFDR.T`

In [67]:
```python
pd.set_option('display.float_format', lambda x: '%.15f' % x)
KSFDR.columns=['KSStatistic','PValue']
```

In [68]: `KSFDR.sort_values(by='KSStatistic',ascending=False,inplace=True)`

In [69]: KSFDR

Out[69]:

	KSStatistic	PValue
Fraud	1.000000000000000	0.000000000000000
sum_amt_cardzip_7	0.686477013602606	0.000000000000000
sum_amt_cardmerch_7	0.683539709039413	0.000000000000000
sum_amt_cardmerch_14	0.679401895904878	0.000000000000000
sum_amt_cardzip_14	0.676453162479227	0.000000000000000
sum_amt_cardzip_3	0.674958675753786	0.000000000000000
sum_amt_cardstate_3	0.674143793514959	0.000000000000000
sum_amt_cardstate_14	0.673269333441799	0.000000000000000
sum_amt_cardstate_7	0.671734215589368	0.000000000000000
sum_amt_cardmerch_3	0.670192362095565	0.000000000000000
sum_amt_cardmerch_30	0.659722133500064	0.000000000000000
max_amt_cardzip_14	0.657785223775823	0.000000000000000
sum_amt_cardzip_30	0.657413416372058	0.000000000000000
max_amt_cardmerch_14	0.656440876669363	0.000000000000000
max_amt_cardzip_7	0.655803815362772	0.000000000000000
max_amt_cardmerch_30	0.652617596259051	0.000000000000000
max_amt_cardmerch_7	0.651910049724677	0.000000000000000
max_amt_cardzip_30	0.651536112989124	0.000000000000000
max_amt_cardstate_7	0.650449501940603	0.000000000000000
max_amt_cardstate_3	0.642089136957491	0.000000000000000
max_amt_cardzip_3	0.638769943256350	0.000000000000000
sum_amt_cardstate_30	0.637847768768821	0.000000000000000
max_amt_cardmerch_3	0.636992603049012	0.000000000000000
max_amt_cardstate_14	0.635275709791024	0.000000000000000
sum_amt_merch_3	0.619470679405459	0.000000000000000
sum_amt_cardzip_1	0.614288711490926	0.000000000000000
sum_amt_cardmerch_1	0.613463834429418	0.000000000000000
sum_amt_cardstate_1	0.612531056738696	0.000000000000000
max_amt_merch_1	0.609878908812192	0.000000000000000
sum_amt_card_7	0.607419182950188	0.000000000000000
max_amt_cardzip_1	0.605173389749327	0.000000000000000
max_amt_cardstate_1	0.603546493351227	0.000000000000000
sum_amt_merch_7	0.603253340856393	0.000000000000000
max_amt_cardmerch_1	0.602031712790166	0.000000000000000
max_amt_cardstate_30	0.601027798035400	0.000000000000000

	KSStatistic	PValue
mean_amt_cardzip_30	0.596579450105715	0.000000000000000
sum_amt_card_3	0.595681089369620	0.000000000000000
mean_amt_cardzip_14	0.593795718165816	0.000000000000000
mean_amt_cardmerch_30	0.592032066519977	0.000000000000000
mean_amt_cardstate_7	0.591430725840536	0.000000000000000
mean_amt_cardmerch_14	0.591214620391854	0.000000000000000
max_amt_merch_3	0.591100462134527	0.000000000000000
mean_amt_cardmerch_7	0.590149172293969	0.000000000000000
mean_amt_cardstate_3	0.590093201286951	0.000000000000000
mean_amt_cardzip_7	0.587648815304976	0.000000000000000
max_amt_card_1	0.586042777667255	0.000000000000000
sum_amt_merch_1	0.584177265741563	0.000000000000000
mean_amt_cardmerch_3	0.583880636786666	0.000000000000000
mean_amt_cardzip_3	0.583390847019505	0.000000000000000
mean_amt_merch_1	0.582711242532890	0.000000000000000
mean_amt_cardstate_14	0.577239902773841	0.000000000000000
mean_amt_card_3	0.575215690446262	0.000000000000000
max_amt_card_3	0.573958950133396	0.000000000000000
mean_amt_cardmerch_1	0.573378728948839	0.000000000000000
mean_amt_cardzip_1	0.573007399558333	0.000000000000000
mean_amt_cardstate_1	0.572323797142645	0.000000000000000
median_amt_cardmerch_30	0.572312107545683	0.000000000000000
sum_amt_card_1	0.571683302831742	0.000000000000000
median_amt_cardmerch_3	0.569867460829203	0.000000000000000
median_amt_cardzip_3	0.569697896489618	0.000000000000000
mean_amt_card_1	0.569682208963583	0.000000000000000
median_amt_cardstate_3	0.569224663363508	0.000000000000000
mean_amt_merch_3	0.569026809330158	0.000000000000000
mean_amt_cardstate_30	0.564956048419093	0.000000000000000
median_amt_cardmerch_1	0.564368179022397	0.000000000000000
median_amt_cardzip_1	0.563290128756782	0.000000000000000
median_amt_cardstate_1	0.561196821785446	0.000000000000000
median_amt_cardzip_30	0.559637803270080	0.000000000000000
median_amt_card_1	0.558570052017403	0.000000000000000
sum_amt_card_14	0.558554972871878	0.000000000000000
median_amt_cardmerch_14	0.558523510908306	0.000000000000000

	KSStatistic	PValue
median_amt_cardmerch_7	0.556740477997615	0.000000000000000
max_amt_merch_7	0.556576128348124	0.000000000000000
median_amt_card_3	0.554438539966732	0.000000000000000
median_amt_cardzip_7	0.552987117890324	0.000000000000000
max_amt_card_7	0.552563033218358	0.000000000000000
median_amt_cardzip_14	0.551967559065274	0.000000000000000
median_amt_cardstate_7	0.551152416091943	0.000000000000000
median_amt_cardstate_30	0.549226370305399	0.000000000000000
mean_amt_card_7	0.547577659120402	0.000000000000000
median_amt_merch_1	0.540694442018165	0.000000000000000
median_amt_cardstate_14	0.536863730674467	0.000000000000000
mean_amt_merch_7	0.536651014774347	0.000000000000000
rel_vel_amt_merch_merch_7	0.535513690864654	0.000000000000000
max_amt_card_14	0.534951112714918	0.000000000000000
mean_amt_card_14	0.534622239592934	0.000000000000000
max_amt_merch_14	0.530915811696498	0.000000000000000
mean_amt_card_30	0.512778120064242	0.000000000000000
max_amt_card_30	0.505693485557090	0.000000000000000
sum_amt_merch_14	0.501714155545688	0.000000000000000
mean_amt_merch_14	0.500680603968883	0.000000000000000
median_amt_merch_3	0.499106462852243	0.000000000000000
rel_vel_amt_card_card_7	0.496204618181695	0.000000000000000
median_amt_card_7	0.492057766250995	0.000000000000000
sum_amt_card_30	0.486765029627696	0.000000000000000
max_amt_merch_30	0.484009022456802	0.000000000000000
median_amt_card_14	0.480440349290372	0.000000000000000
rel_amt_vel_merch_card_30	0.478321620704823	0.000000000000000
rel_vel_amt_merch_merch_14	0.471728253460360	0.000000000000000
rel_amt_vel_card_card_30	0.468445649501384	0.000000000000000
median_amt_card_30	0.459190487153741	0.000000000000000
mean_amt_merch_30	0.458363784950700	0.000000000000000
median_amt_merch_7	0.454719324954891	0.000000000000000
rel_amt_vel_card_card_14	0.454227188577187	0.000000000000000
rel_amt_vel_card_card_7	0.443613382181074	0.000000000000000
sum_amt_merch_30	0.443286986036885	0.000000000000000
rel_amt_vel_merch_card_14	0.442379716871878	0.000000000000000

	KSStatistic	PValue
median_amt_merch_14	0.427421552157721	0.000000000000000
rel_vel_amt_merch_card_7	0.419295674405875	0.000000000000000
median_amt_merch_30	0.417444242143439	0.000000000000000
rel_vel_amt_card_card_14	0.410858870757122	0.000000000000000
rel_amt_vel_merch_card_7	0.407188467678091	0.000000000000000
vel_card_7	0.405221747308807	0.000000000000000
rel_amt_vel_merch_merch_7	0.403935761275659	0.000000000000000
rel_vel_amt_merch_merch_30	0.400887948738206	0.000000000000000
rel_amt_vel_merch_merch_14	0.399744367200399	0.000000000000000
vel_card_3	0.395148226262274	0.000000000000000
rel_vel_amt_card_merch_7	0.395121935533046	0.000000000000000
rel_amt_vel_merch_merch_30	0.392910124729303	0.000000000000000
ratio_actual_mean_amt_merch_30	0.381505162934296	0.000000000000000
card_U*	0.369710881067955	0.000000000000000
vel_cardstate_7	0.363848178817286	0.000000000000000
dayssince_cardstate	0.361262213999235	0.000000000000000
vel_cardmerch_7	0.353513879810258	0.000000000000000
ratio_actual_mean_amt_card_30	0.353217989603125	0.000000000000000
ratio_actual_sum_amt_cardstate_3	0.351795813246999	0.000000000000000
vel_cardstate_3	0.351391892043400	0.000000000000000
ratio_actual_median_amt_merch_30	0.348921823712708	0.000000000000000
vel_cardzip_7	0.346270066887961	0.000000000000000
vel_card_14	0.345467091524851	0.000000000000000
rel_vel_amt_merch_card_14	0.342657546868983	0.000000000000000
vel_card_1	0.340226197612246	0.000000000000000
ratio_actual_mean_amt_merch_7	0.339527168405027	0.000000000000000
ratio_actual_mean_amt_merch_14	0.337910788298618	0.000000000000000
vel_cardstate_14	0.335026890853147	0.000000000000000
rel_amt_vel_card_merch_14	0.334848939553659	0.000000000000000
rel_vel_amt_card_merch_14	0.333660424769535	0.000000000000000
ratio_actual_max_amt_card_30	0.333220478748530	0.000000000000000
vel_cardmerch_14	0.332672458275309	0.000000000000000
rel_amt_vel_card_merch_7	0.328075491677919	0.000000000000000
vel_cardzip_14	0.327414920810151	0.000000000000000
dayssince_cardzip	0.326926782361519	0.000000000000000
rel_amt_vel_card_merch_30	0.326673522245911	0.000000000000000

	KSStatistic	PValue
ratio_actual_sum_amt_cardzip_3	0.325074654807071	0.000000000000000
ratio_actual_sum_amt_cardmerch_3	0.324476399485936	0.000000000000000
vel_cardmerch_3	0.324388488502086	0.000000000000000
vel_cardzip_3	0.322290314486661	0.000000000000000
rel_vel_amt_card_card_30	0.321765369017110	0.000000000000000
dayssince_cardmerch	0.321120094573620	0.000000000000000
rel_amt_amt_card_card_30	0.319490417007468	0.000000000000000
ratio_actual_sum_amt_card_1	0.317367038656585	0.000000000000000
ratio_actual_max_amt_cardstate_3	0.315393061176919	0.000000000000000
ratio_actual_sum_amt_cardstate_7	0.315236490106820	0.000000000000000
ratio_actual_median_amt_merch_14	0.314469626472618	0.000000000000000
ratio_actual_sum_amt_cardzip_7	0.312712319366394	0.000000000000000
ratio_actual_sum_amt_cardmerch_7	0.310640436080197	0.000000000000000
ratio_actual_median_amt_merch_7	0.302329958298992	0.000000000000000
ratio_actual_max_amt_cardzip_3	0.298858148001075	0.000000000000000
ratio_actual_max_amt_cardmerch_3	0.297875439652695	0.000000000000000
vel_cardmerch_30	0.296589844721304	0.000000000000000
ratio_actual_sum_amt_card_3	0.294785214302470	0.000000000000000
ratio_actual_sum_amt_cardstate_1	0.292579313480834	0.000000000000000
vel_cardstate_1	0.292403491513135	0.000000000000000
dayssince_card	0.292103951022935	0.000000000000000
ratio_actual_max_amt_cardmerch_7	0.290239134389256	0.000000000000000
ratio_actual_max_amt_cardzip_7	0.289606331746242	0.000000000000000
rel_vel_amt_merch_card_30	0.285333762328505	0.000000000000000
ratio_actual_median_amt_card_30	0.284988810578729	0.000000000000000
vel_cardzip_30	0.284190702259708	0.000000000000000
ratio_actual_max_amt_cardstate_7	0.283664279294630	0.000000000000000
vel_cardstate_30	0.280973064648164	0.000000000000000
ratio_actual_max_amt_card_1	0.280178172054702	0.000000000000000
ratio_actual_mean_amt_card_14	0.279957156106180	0.000000000000000
ratio_actual_sum_amt_cardmerch_14	0.276457316849003	0.000000000000000
vel_merch_7	0.274571380720439	0.000000000000000
rel_vel_amt_card_merch_30	0.273999068482526	0.000000000000000
vel_merch_3	0.273279919262691	0.000000000000000
ratio_actual_mean_amt_merch_3	0.272983203396293	0.000000000000000
ratio_actual_sum_amt_cardzip_1	0.272541649512509	0.000000000000000

	KSStatistic	PValue
vel_cardzip_1	0.272478855952616	0.000000000000000
ratio_actual_sum_amt_cardmerch_1	0.271590620406456	0.000000000000000
vel_cardmerch_1	0.271527826846563	0.000000000000000
vel_merch_14	0.269470848882887	0.000000000000000
vel_card_30	0.267859031630311	0.000000000000000
ratio_actual_sum_amt_cardzip_14	0.266845078597113	0.000000000000000
ratio_actual_max_amt_cardmerch_14	0.254716230905434	0.000000000000000
ratio_actual_mean_amt_card_7	0.253350416658084	0.000000000000000
ratio_actual_max_amt_card_3	0.246717895782159	0.000000000000000
ratio_actual_max_amt_cardzip_14	0.245983059036289	0.000000000000000
ratio_actual_median_amt_card_14	0.243722664703162	0.000000000000000
ratio_actual_max_amt_cardstate_1	0.243378973170159	0.000000000000000
ratio_actual_sum_amt_cardstate_14	0.242302139665566	0.000000000000000
ratio_actual_sum_amt_cardmerch_30	0.231950936371481	0.000000000000000
ratio_actual_max_amt_cardzip_1	0.231789020244035	0.000000000000000
ratio_actual_max_amt_cardmerch_1	0.230919905228222	0.000000000000000
ratio_actual_median_amt_merch_3	0.230796490895977	0.000000000000000
ratio_actual_max_amt_card_14	0.230373753352285	0.000000000000000
ratio_actual_sum_amt_merch_3	0.230150608071975	0.000000000000000
ratio_actual_median_amt_card_7	0.227564512886673	0.000000000000000
rel_amt_amt_card_merch_30	0.227446052510017	0.000000000000000
vel_merch_1	0.227228643388812	0.000000000000000
rel_amt_amt_merch_merch_30	0.226177709511690	0.000000000000000
vel_merch_30	0.225168275876574	0.000000000000000
ratio_actual_mean_amt_cardstate_30	0.218251467478976	0.000000000000000
ratio_actual_median_amt_cardstate_7	0.216855799130833	0.000000000000000
ratio_actual_max_amt_merch_30	0.216201703169932	0.000000000000000
ratio_actual_sum_amt_cardzip_30	0.215883302883889	0.000000000000000
ratio_actual_max_amt_cardstate_14	0.215520968833796	0.000000000000000
rel_amt_amt_card_card_14	0.212062934008879	0.000000000000000
ratio_actual_mean_amt_cardstate_14	0.211016258795335	0.000000000000000
ratio_actual_median_amt_cardmerch_7	0.209935731551926	0.000000000000000
ratio_actual_sum_amt_merch_1	0.209876979376857	0.000000000000000
ratio_actual_median_amt_cardstate_14	0.209677604392229	0.000000000000000
ratio_actual_sum_amt_merch_7	0.209326481925841	0.000000000000000
ratio_actual_sum_amt_card_7	0.209263210352689	0.000000000000000

	KSStatistic	PValue
Recnum	0.207678683311610	0.000000000000000
ratio_actual_max_amt_cardmerch_30	0.206356324815050	0.000000000000000
ratio_actual_mean_amt_cardstate_7	0.206343027355308	0.000000000000000
ratio_actual_median_amt_cardzip_7	0.204435580630103	0.000000000000000
ratio_actual_median_amt_card_3	0.203802951810093	0.000000000000000
ratio_actual_mean_amt_card_3	0.198677867473344	0.000000000000000
ratio_actual_median_amt_cardstate_3	0.197526290077392	0.000000000000000
ratio_actual_mean_amt_cardstate_3	0.196678816025473	0.000000000000000
ratio_actual_median_amt_cardstate_30	0.196360198460676	0.000000000000000
rel_amt_amt_merch_card_30	0.195019588548784	0.000000000000000
ratio_actual_sum_amt_card_30	0.194421811240908	0.000000000000000
ratio_actual_max_amt_cardzip_30	0.193581811578124	0.000000000000000
ratio_actual_median_amt_cardmerch_14	0.193302564923544	0.000000000000000
rel_vel_vel_merch_merch_7	0.192786440971403	0.000000000000000
rel_amt_amt_merch_merch_7	0.192296651204241	0.000000000000000
rel_vel_vel_card_card_30	0.190967426699057	0.000000000000000
rel_vel_vel_merch_card_7	0.190227679453348	0.000000000000000
ratio_actual_median_amt_cardmerch_3	0.190208602378750	0.000000000000000
rel_amt_amt_card_merch_14	0.189134550042209	0.000000000000000
ratio_actual_mean_amt_cardmerch_3	0.188895326134106	0.000000000000000
ratio_actual_median_amt_cardzip_14	0.188673527982071	0.000000000000000
ratio_actual_median_amt_cardzip_3	0.187887804552025	0.000000000000000
ratio_actual_mean_amt_cardmerch_7	0.187321793397976	0.000000000000000
ratio_actual_mean_amt_cardzip_3	0.185435683446410	0.000000000000000
ratio_actual_mean_amt_cardmerch_14	0.185326652967676	0.000000000000000
ratio_actual_mean_amt_cardzip_7	0.184932205117749	0.000000000000000
ratio_actual_mean_amt_merch_1	0.181833896997877	0.000000000000000
ratio_actual_mean_amt_cardzip_14	0.181415461573514	0.000000000000000
ratio_actual_max_amt_merch_14	0.180991637636052	0.000000000000000
dayssince_merch	0.171933547118245	0.000000000000000
rel_amt_amt_merch_merch_14	0.171597286518627	0.000000000000000
ratio_actual_mean_amt_card_1	0.171542314993877	0.000000000000000
ratio_actual_mean_amt_cardmerch_30	0.170610971342931	0.000000000000000
ratio_actual_mean_amt_cardstate_1	0.169277792364425	0.000000000000000
ratio_actual_mean_amt_cardzip_30	0.165502704381739	0.000000000000000
ratio_actual_sum_amt_merch_14	0.165088962178461	0.000000000000000

	KSStatistic	PValue
ratio_actual_median_amt_cardmerch_30	0.165075143249710	0.000000000000000
ratio_actual_median_amt_card_1	0.163526727937537	0.000000000000000
ratio_actual_sum_amt_merch_30	0.162393401956917	0.000000000000000
ratio_actual_mean_amt_cardzip_1	0.160725048772562	0.000000000000000
ratio_actual_median_amt_cardstate_1	0.160163339737842	0.000000000000000
ratio_actual_median_amt_cardzip_30	0.159937760935488	0.000000000000000
ratio_actual_mean_amt_cardmerch_1	0.159567083381243	0.000000000000000
ratio_actual_sum_amt_cardstate_30	0.156544040621741	0.000000000000000
ratio_actual_median_amt_cardzip_1	0.153611994204393	0.000000000000000
ratio_actual_median_amt_cardmerch_1	0.153578315997531	0.000000000000000
rel_vel_vel_card_merch_7	0.152643408975020	0.000000000000000
ratio_actual_median_amt_merch_1	0.152289244606077	0.000000000000000
ratio_actual_max_amt_merch_3	0.147746510785674	0.000000000000000
rel_vel_vel_merch_card_14	0.146333286314211	0.000000000000000
rel_amt_amt_card_merch_7	0.143989413483815	0.000000000000001
rel_vel_vel_merch_merch_14	0.143192043912558	0.000000000000001
ratio_actual_max_amt_merch_7	0.142769480191869	0.000000000000001
ratio_actual_max_amt_card_7	0.141413313121195	0.000000000000003
rel_vel_vel_card_card_14	0.139168823592857	0.000000000000007
risk_dayofweek	0.135249114871540	0.000000000000047
ratio_actual_max_amt_merch_1	0.130634809429585	0.000000000000387
ratio_actual_max_amt_cardstate_30	0.125031146909379	0.000000000004522
ratio_actual_sum_amt_card_14	0.122927758203859	0.000000000011062
rel_vel_vel_card_merch_30	0.119747970551254	0.000000000041561
rel_vel_vel_merch_card_30	0.115940290755476	0.000000000193744
rel_vel_vel_merch_merch_30	0.115448371656525	0.000000000235517
rel_vel_vel_card_merch_14	0.106461852760270	0.000000007204655
rel_amt_amt_card_card_7	0.104673344424979	0.000000013769589
rel_vel_vel_card_card_7	0.102711056542276	0.000000027674447
rel_amt_amt_merch_card_14	0.093182687889412	0.000000680101234
merch_U*	0.080267595298679	0.000031722510816
rel_amt_amt_merch_card_7	0.057916998994521	0.006341147813008
random	0.040835282106478	0.114477598228336

In [70]:
```python
#FDR Fraud Detection Rate
top=int(round(len(data)*0.03))
for i in list(data.columns):
    t1=data.sort_values(i,ascending=False).head(top)
    t2=data.sort_values(i,ascending=True).head(top)
    n1=t1.loc[:,'Fraud']
    n2=t2.loc[:,'Fraud']
    FDR1=sum(n1.astype('int'))/sum(data['Fraud'].astype('int'))
    FDR2=sum(n2.astype('int'))/sum(data['Fraud'].astype('int'))
    FDRate=max(FDR1,FDR2)
    KSFDR.loc['{}'.format(i),"FDR"]=FDRate
```

```
In [71]: pd.set_option('display.max_rows',None)
         KSFDR
```

Out[71]:

	KSStatistic	PValue	FDR
Fraud	1.000000000000000	0.000000000000000	1.000000000000000
sum_amt_cardzip_7	0.686477013602606	0.000000000000000	0.645905420991926
sum_amt_cardmerch_7	0.683539709039413	0.000000000000000	0.645905420991926
sum_amt_cardmerch_14	0.679401895904878	0.000000000000000	0.635524798154556
sum_amt_cardzip_14	0.676453162479227	0.000000000000000	0.635524798154556
sum_amt_cardzip_3	0.674958675753786	0.000000000000000	0.625144175317186
sum_amt_cardstate_3	0.674143793514959	0.000000000000000	0.618223760092272
sum_amt_cardstate_14	0.673269333441799	0.000000000000000	0.540945790080738
sum_amt_cardstate_7	0.671734215589368	0.000000000000000	0.615916955017301
sum_amt_cardmerch_3	0.670192362095565	0.000000000000000	0.617070357554787
sum_amt_cardmerch_30	0.659722133500064	0.000000000000000	0.566320645905421
max_amt_cardzip_14	0.657785223775823	0.000000000000000	0.476355247981546
sum_amt_cardzip_30	0.657413416372058	0.000000000000000	0.559400230680507
max_amt_cardmerch_14	0.656440876669363	0.000000000000000	0.475201845444060
max_amt_cardzip_7	0.655803815362772	0.000000000000000	0.461361014994233
max_amt_cardmerch_30	0.652617596259051	0.000000000000000	0.474048442906574
max_amt_cardmerch_7	0.651910049724677	0.000000000000000	0.461361014994233
max_amt_cardzip_30	0.651536112989124	0.000000000000000	0.483275663206459
max_amt_cardstate_7	0.650449501940603	0.000000000000000	0.489042675893887
max_amt_cardstate_3	0.642089136957491	0.000000000000000	0.460207612456747
max_amt_cardzip_3	0.638769943256350	0.000000000000000	0.468281430219146
sum_amt_cardstate_30	0.637847768768821	0.000000000000000	0.453287197231834
max_amt_cardmerch_3	0.636992603049012	0.000000000000000	0.468281430219146
max_amt_cardstate_14	0.635275709791024	0.000000000000000	0.487889273356401
sum_amt_merch_3	0.619470679405459	0.000000000000000	0.432525951557093
sum_amt_cardzip_1	0.614288711490926	0.000000000000000	0.558246828143022
sum_amt_cardmerch_1	0.613463834429418	0.000000000000000	0.561707035755479
sum_amt_cardstate_1	0.612531056738696	0.000000000000000	0.562860438292964
max_amt_merch_1	0.609878908812192	0.000000000000000	0.445213379469435
sum_amt_card_7	0.607419182950188	0.000000000000000	0.525951557093426
max_amt_cardzip_1	0.605173389749327	0.000000000000000	0.416378316032295
max_amt_cardstate_1	0.603546493351227	0.000000000000000	0.419838523644752
sum_amt_merch_7	0.603253340856393	0.000000000000000	0.379469434832757
max_amt_cardmerch_1	0.602031712790166	0.000000000000000	0.416378316032295
max_amt_cardstate_30	0.601027798035400	0.000000000000000	0.478662053056517

	KSStatistic	PValue	FDR
mean_amt_cardzip_30	0.596579450105715	0.000000000000000	0.296424452133795
sum_amt_card_3	0.595681089369620	0.000000000000000	0.557093425605536
mean_amt_cardzip_14	0.593795718165816	0.000000000000000	0.294117647058824
mean_amt_cardmerch_30	0.592032066519977	0.000000000000000	0.291810841983852
mean_amt_cardstate_7	0.591430725840536	0.000000000000000	0.310265282583622
mean_amt_cardmerch_14	0.591214620391854	0.000000000000000	0.292964244521338
max_amt_merch_3	0.591100462134527	0.000000000000000	0.441753171856978
mean_amt_cardmerch_7	0.590149172293969	0.000000000000000	0.294117647058824
mean_amt_cardstate_3	0.590093201286951	0.000000000000000	0.307958477508651
mean_amt_cardzip_7	0.587648815304976	0.000000000000000	0.296424452133795
max_amt_card_1	0.586042777667255	0.000000000000000	0.425605536332180
sum_amt_merch_1	0.584177265741563	0.000000000000000	0.567474048442907
mean_amt_cardmerch_3	0.583880636786666	0.000000000000000	0.301038062283737
mean_amt_cardzip_3	0.583390847019505	0.000000000000000	0.302191464821223
mean_amt_merch_1	0.582711242532890	0.000000000000000	0.309111880046136
mean_amt_cardstate_14	0.577239902773841	0.000000000000000	0.310265282583622
mean_amt_card_3	0.575215690446262	0.000000000000000	0.363321799307958
max_amt_card_3	0.573958950133396	0.000000000000000	0.440599769319493
mean_amt_cardmerch_1	0.573378728948839	0.000000000000000	0.319492502883506
mean_amt_cardzip_1	0.573007399558333	0.000000000000000	0.319492502883506
mean_amt_cardstate_1	0.572323797142645	0.000000000000000	0.324106113033449
median_amt_cardmerch_30	0.572312107545683	0.000000000000000	0.277970011534025
sum_amt_card_1	0.571683302831742	0.000000000000000	0.552479815455594
median_amt_cardmerch_3	0.569867460829203	0.000000000000000	0.291810841983852
median_amt_cardzip_3	0.569697896489618	0.000000000000000	0.294117647058824
mean_amt_card_1	0.569682208963583	0.000000000000000	0.328719723183391
median_amt_cardstate_3	0.569224663363508	0.000000000000000	0.297577854671280
mean_amt_merch_3	0.569026809330158	0.000000000000000	0.302191464821223
mean_amt_cardstate_30	0.564956048419093	0.000000000000000	0.324106113033449
median_amt_cardmerch_1	0.564368179022397	0.000000000000000	0.299884659746251
median_amt_cardzip_1	0.563290128756782	0.000000000000000	0.299884659746251
median_amt_cardstate_1	0.561196821785446	0.000000000000000	0.303344867358708
median_amt_cardzip_30	0.559637803270080	0.000000000000000	0.279123414071511
median_amt_card_1	0.558570052017403	0.000000000000000	0.309111880046136
sum_amt_card_14	0.558554972871878	0.000000000000000	0.475201845444060
median_amt_cardmerch_14	0.558523510908306	0.000000000000000	0.279123414071511

	KSStatistic	PValue	FDR
median_amt_cardmerch_7	0.556740477997615	0.000000000000000	0.280276816608997
max_amt_merch_7	0.556576128348124	0.000000000000000	0.457900807381776
median_amt_card_3	0.554438539966732	0.000000000000000	0.324106113033449
median_amt_cardzip_7	0.552987117890324	0.000000000000000	0.282583621683968
max_amt_card_7	0.552563033218358	0.000000000000000	0.480968858131488
median_amt_cardzip_14	0.551967559065274	0.000000000000000	0.280276816608997
median_amt_cardstate_7	0.551152416091943	0.000000000000000	0.282583621683968
median_amt_cardstate_30	0.549226370305399	0.000000000000000	0.297577854671280
mean_amt_card_7	0.547577659120402	0.000000000000000	0.388696655132641
median_amt_merch_1	0.540694442018165	0.000000000000000	0.288350634371396
median_amt_cardstate_14	0.536863730674467	0.000000000000000	0.294117647058824
mean_amt_merch_7	0.536651014774347	0.000000000000000	0.275663206459054
rel_vel_amt_merch_merch_7	0.535513690864654	0.000000000000000	0.227220299884660
max_amt_card_14	0.534951112714918	0.000000000000000	0.514417753171857
mean_amt_card_14	0.534622239592934	0.000000000000000	0.409457900807382
max_amt_merch_14	0.530915811696498	0.000000000000000	0.409457900807382
mean_amt_card_30	0.512778120064242	0.000000000000000	0.382929642445213
max_amt_card_30	0.505693485557090	0.000000000000000	0.380622837370242
sum_amt_merch_14	0.501714155545688	0.000000000000000	0.251441753171857
mean_amt_merch_14	0.500680603968883	0.000000000000000	0.282583621683968
median_amt_merch_3	0.499106462852243	0.000000000000000	0.277970011534025
rel_vel_amt_card_card_7	0.496204618181695	0.000000000000000	0.318339100346021
median_amt_card_7	0.492057766250995	0.000000000000000	0.317185697808535
sum_amt_card_30	0.486765029627696	0.000000000000000	0.410611303344867
max_amt_merch_30	0.484009022456802	0.000000000000000	0.251441753171857
median_amt_card_14	0.480440349290372	0.000000000000000	0.327566320645905
rel_amt_vel_merch_card_30	0.478321620704823	0.000000000000000	0.220299884659746
rel_vel_amt_merch_merch_14	0.471728253460360	0.000000000000000	0.139561707035755
rel_amt_vel_card_card_30	0.468445649501384	0.000000000000000	0.354094579008074
median_amt_card_30	0.459190487153741	0.000000000000000	0.302191464821223
mean_amt_merch_30	0.458363784950700	0.000000000000000	0.262975778546713
median_amt_merch_7	0.454719324954891	0.000000000000000	0.254901960784314
rel_amt_vel_card_card_14	0.454227188577187	0.000000000000000	0.325259515557093
rel_amt_vel_card_card_7	0.443613382181074	0.000000000000000	0.313725490196078
sum_amt_merch_30	0.443286986036885	0.000000000000000	0.134948096885813
rel_amt_vel_merch_card_14	0.442379716871878	0.000000000000000	0.201845444059977

	KSStatistic	PValue	FDR
median_amt_merch_14	0.427421552157721	0.000000000000000	0.237600922722030
rel_vel_amt_merch_card_7	0.419295674405875	0.000000000000000	0.339100346020761
median_amt_merch_30	0.417444242143439	0.000000000000000	0.201845444059977
rel_vel_amt_card_card_14	0.410858870757122	0.000000000000000	0.262975778546713
rel_amt_vel_merch_card_7	0.407188467678091	0.000000000000000	0.189158016147636
vel_card_7	0.405221747308807	0.000000000000000	0.122260668973472
rel_amt_vel_merch_merch_7	0.403935761275659	0.000000000000000	0.252595155709343
rel_vel_amt_merch_merch_30	0.400887948738206	0.000000000000000	0.066897347174164
rel_amt_vel_merch_merch_14	0.399744367200399	0.000000000000000	0.238754325259516
vel_card_3	0.395148226262274	0.000000000000000	0.227220299884660
rel_vel_amt_card_merch_7	0.395121935533046	0.000000000000000	0.167243367953409
rel_amt_vel_merch_merch_30	0.392910124729303	0.000000000000000	0.208765859284890
ratio_actual_mean_amt_merch_30	0.381505162934296	0.000000000000000	0.209919261822376
card_U*	0.369710881067955	0.000000000000000	0.205305651672434
vel_cardstate_7	0.363848178817286	0.000000000000000	0.077277970011534
dayssince_cardstate	0.361262213999235	0.000000000000000	0.074971164936563
vel_cardmerch_7	0.353513879810258	0.000000000000000	0.087658592848904
ratio_actual_mean_amt_card_30	0.353217989603125	0.000000000000000	0.078431372549020
ratio_actual_sum_amt_cardstate_3	0.351795813246999	0.000000000000000	0.096885813148789
vel_cardstate_3	0.351391892043400	0.000000000000000	0.163783160322953
ratio_actual_median_amt_merch_30	0.348921823712708	0.000000000000000	0.245674740484429
vel_cardzip_7	0.346270066887961	0.000000000000000	0.070357554786621
vel_card_14	0.345467091524851	0.000000000000000	0.027681660899654
rel_vel_amt_merch_card_14	0.342657546868983	0.000000000000000	0.307958477508651
vel_card_1	0.340226197612246	0.000000000000000	0.168396770472895
ratio_actual_mean_amt_merch_7	0.339527168405027	0.000000000000000	0.115340253748558
ratio_actual_mean_amt_merch_14	0.337910788298618	0.000000000000000	0.151095732410611
vel_cardstate_14	0.335026890853147	0.000000000000000	0.018454440599769
rel_amt_vel_card_merch_14	0.334848939553659	0.000000000000000	0.232987312572088
rel_vel_amt_card_merch_14	0.333660424769535	0.000000000000000	0.077277970011534
ratio_actual_max_amt_card_30	0.333220478748530	0.000000000000000	0.024221453287197
vel_cardmerch_14	0.332672458275309	0.000000000000000	0.009227220299885
rel_amt_vel_card_merch_7	0.328075491677919	0.000000000000000	0.243367935409458
vel_cardzip_14	0.327414920810151	0.000000000000000	0.032295271049596
dayssince_cardzip	0.326926782361519	0.000000000000000	0.066897347174164
rel_amt_vel_card_merch_30	0.326673522245911	0.000000000000000	0.230680507497116

	KSStatistic	PValue	FDR
ratio_actual_sum_amt_cardzip_3	0.325074654807071	0.000000000000000	0.122260668973472
ratio_actual_sum_amt_cardmerch_3	0.324476399485936	0.000000000000000	0.144175317185698
vel_cardmerch_3	0.324388488502086	0.000000000000000	0.169550173010381
vel_cardzip_3	0.322290314486661	0.000000000000000	0.164936562860438
rel_vel_amt_card_card_30	0.321765369017110	0.000000000000000	0.178777393310265
dayssince_cardmerch	0.321120094573620	0.000000000000000	0.047289504036909
rel_amt_amt_card_card_30	0.319490417007468	0.000000000000000	0.039215686274510
ratio_actual_sum_amt_card_1	0.317367038656585	0.000000000000000	0.046136101499423
ratio_actual_max_amt_cardstate_3	0.315393061176919	0.000000000000000	0.063437139561707
ratio_actual_sum_amt_cardstate_7	0.315236490106820	0.000000000000000	0.053056516724337
ratio_actual_median_amt_merch_14	0.314469626472618	0.000000000000000	0.200692041522491
ratio_actual_sum_amt_cardzip_7	0.312712319366394	0.000000000000000	0.064590542099193
ratio_actual_sum_amt_cardmerch_7	0.310640436080197	0.000000000000000	0.071510957324106
ratio_actual_median_amt_merch_7	0.302329958298992	0.000000000000000	0.143021914648212
ratio_actual_max_amt_cardzip_3	0.298858148001075	0.000000000000000	0.078431372549020
ratio_actual_max_amt_cardmerch_3	0.297875439652695	0.000000000000000	0.089965397923875
vel_cardmerch_30	0.296589844721304	0.000000000000000	0.012687427912341
ratio_actual_sum_amt_card_3	0.294785214302470	0.000000000000000	0.021914648212226
ratio_actual_sum_amt_cardstate_1	0.292579313480834	0.000000000000000	0.094579008073818
vel_cardstate_1	0.292403491513135	0.000000000000000	0.114186851211073
dayssince_card	0.292103951022935	0.000000000000000	0.040369088811995
ratio_actual_max_amt_cardmerch_7	0.290239134389256	0.000000000000000	0.072664359861592
ratio_actual_max_amt_cardzip_7	0.289606331746242	0.000000000000000	0.063437139561707
rel_vel_amt_merch_card_30	0.285333762328505	0.000000000000000	0.257208765859285
ratio_actual_median_amt_card_30	0.284988810578729	0.000000000000000	0.108419838523645
vel_cardzip_30	0.284190702259708	0.000000000000000	0.024221453287197
ratio_actual_max_amt_cardstate_7	0.283664279294630	0.000000000000000	0.038062283737024
vel_cardstate_30	0.280973064648164	0.000000000000000	0.001153402537486
ratio_actual_max_amt_card_1	0.280178172054702	0.000000000000000	0.026528258362168
ratio_actual_mean_amt_card_14	0.279957156106180	0.000000000000000	0.076124567474048
ratio_actual_sum_amt_cardmerch_14	0.276457316849003	0.000000000000000	0.051903114186851
vel_merch_7	0.274571380720439	0.000000000000000	0.021914648212226
rel_vel_amt_card_merch_30	0.273999068482526	0.000000000000000	0.041522491349481
vel_merch_3	0.273279919262691	0.000000000000000	0.010380622838370
ratio_actual_mean_amt_merch_3	0.272983203396293	0.000000000000000	0.093425605536332
ratio_actual_sum_amt_cardzip_1	0.272541649512509	0.000000000000000	0.104959630911188

	KSStatistic	PValue	FDR
vel_cardzip_1	0.272478855952616	0.000000000000000	0.113033448673587
ratio_actual_sum_amt_cardmerch_1	0.271590620406456	0.000000000000000	0.107266435986159
vel_cardmerch_1	0.271527826846563	0.000000000000000	0.111880046136102
vel_merch_14	0.269470848882887	0.000000000000000	0.024221453287197
vel_card_30	0.267859031630311	0.000000000000000	0.018454440599769
ratio_actual_sum_amt_cardzip_14	0.266845078597113	0.000000000000000	0.046136101499423
ratio_actual_max_amt_cardmerch_14	0.254716230905434	0.000000000000000	0.049596309111880
ratio_actual_mean_amt_card_7	0.253350416658084	0.000000000000000	0.065743944636678
ratio_actual_max_amt_card_3	0.246717895782159	0.000000000000000	0.012687427912341
ratio_actual_max_amt_cardzip_14	0.245983059036289	0.000000000000000	0.043829296424452
ratio_actual_median_amt_card_14	0.243722664703162	0.000000000000000	0.083044982698962
ratio_actual_max_amt_cardstate_1	0.243378973170159	0.000000000000000	0.083044982698962
ratio_actual_sum_amt_cardstate_14	0.242302139665566	0.000000000000000	0.036908881199539
ratio_actual_sum_amt_cardmerch_30	0.231950936371481	0.000000000000000	0.027681660899654
ratio_actual_max_amt_cardzip_1	0.231789020244035	0.000000000000000	0.096885813148789
ratio_actual_max_amt_cardmerch_1	0.230919905228222	0.000000000000000	0.110726643598616
ratio_actual_median_amt_merch_3	0.230796490895977	0.000000000000000	0.108419838523645
ratio_actual_max_amt_card_14	0.230373753352285	0.000000000000000	0.034602076124567
ratio_actual_sum_amt_merch_3	0.230150608071975	0.000000000000000	0.009227220299885
ratio_actual_median_amt_card_7	0.227564512886673	0.000000000000000	0.084198385236448
rel_amt_amt_card_merch_30	0.227446052510017	0.000000000000000	0.113033448673587
vel_merch_1	0.227228643388812	0.000000000000000	0.008073817762399
rel_amt_amt_merch_merch_30	0.226177709511690	0.000000000000000	0.008073817762399
vel_merch_30	0.225168275876574	0.000000000000000	0.012687427912341
ratio_actual_mean_amt_cardstate_30	0.218251467478976	0.000000000000000	0.069204152249135
ratio_actual_median_amt_cardstate_7	0.216855799130833	0.000000000000000	0.073817762399077
ratio_actual_max_amt_merch_30	0.216201703169932	0.000000000000000	0.011534025374856
ratio_actual_sum_amt_cardzip_30	0.215883302883889	0.000000000000000	0.043829296424452
ratio_actual_max_amt_cardstate_14	0.215520968833796	0.000000000000000	0.026528258362168
rel_amt_amt_card_card_14	0.212062934008879	0.000000000000000	0.027681660899654
ratio_actual_mean_amt_cardstate_14	0.211016258795335	0.000000000000000	0.068050749711649
ratio_actual_median_amt_cardmerch_7	0.209935731551926	0.000000000000000	0.094579008073818
ratio_actual_sum_amt_merch_1	0.209876979376857	0.000000000000000	0.009227220299885
ratio_actual_median_amt_cardstate_14	0.209677604392229	0.000000000000000	0.056516724336794
ratio_actual_sum_amt_merch_7	0.209326481925841	0.000000000000000	0.021914648212226
ratio_actual_sum_amt_card_7	0.209263210352689	0.000000000000000	0.013840830449827

	KSStatistic	PValue	FDR
Recnum	0.207678683311610	0.000000000000000	0.029988465974625
ratio_actual_max_amt_cardmerch_30	0.206356324815050	0.000000000000000	0.032295271049596
ratio_actual_mean_amt_cardstate_7	0.206343027355308	0.000000000000000	0.074971164936563
ratio_actual_median_amt_cardzip_7	0.204435580630103	0.000000000000000	0.094579008073818
ratio_actual_median_amt_card_3	0.203802951810093	0.000000000000000	0.040369088811995
ratio_actual_mean_amt_card_3	0.198677867473344	0.000000000000000	0.046136101499423
ratio_actual_median_amt_cardstate_3	0.197526290077392	0.000000000000000	0.064590542099193
ratio_actual_mean_amt_cardstate_3	0.196678816025473	0.000000000000000	0.059976931949250
ratio_actual_median_amt_cardstate_30	0.196360198460676	0.000000000000000	0.059976931949250
rel_amt_amt_merch_card_30	0.195019588548784	0.000000000000000	0.055363321799308
ratio_actual_sum_amt_card_30	0.194421811240908	0.000000000000000	0.024221453287197
ratio_actual_max_amt_cardzip_30	0.193581811578124	0.000000000000000	0.039215686274510
ratio_actual_median_amt_cardmerch_14	0.193302564923544	0.000000000000000	0.095732410611303
rel_vel_vel_merch_merch_7	0.192786440971403	0.000000000000000	0.026528258362168
rel_amt_amt_merch_merch_7	0.192296651204241	0.000000000000000	0.010380062283370
rel_vel_vel_card_card_30	0.190967426699057	0.000000000000000	0.044982698961938
rel_vel_vel_merch_card_7	0.190227679453348	0.000000000000000	0.151095732410611
ratio_actual_median_amt_cardmerch_3	0.190208602378750	0.000000000000000	0.083044982698962
rel_amt_amt_card_merch_14	0.189134550042209	0.000000000000000	0.110726643598616
ratio_actual_mean_amt_cardmerch_3	0.188895326134106	0.000000000000000	0.080738177623991
ratio_actual_median_amt_cardzip_14	0.188673527982071	0.000000000000000	0.083044982698962
ratio_actual_median_amt_cardzip_3	0.187887804552025	0.000000000000000	0.072664359861592
ratio_actual_mean_amt_cardmerch_7	0.187321793397976	0.000000000000000	0.081891580161476
ratio_actual_mean_amt_cardzip_3	0.185435683446410	0.000000000000000	0.070357554786621
ratio_actual_mean_amt_cardmerch_14	0.185326652967676	0.000000000000000	0.079584775086505
ratio_actual_mean_amt_cardzip_7	0.184932205117749	0.000000000000000	0.076124567474048
ratio_actual_mean_amt_merch_1	0.181833896997877	0.000000000000000	0.078431372549020
ratio_actual_mean_amt_cardzip_14	0.181415461573514	0.000000000000000	0.074971164936563
ratio_actual_max_amt_merch_14	0.180991637636052	0.000000000000000	0.019607843137255
dayssince_merch	0.171933547118245	0.000000000000000	0.066897347174164
rel_amt_amt_merch_merch_14	0.171597286518627	0.000000000000000	0.016147635524798
ratio_actual_mean_amt_card_1	0.171542314993877	0.000000000000000	0.051903114186851
ratio_actual_mean_amt_cardmerch_30	0.170610971342931	0.000000000000000	0.071510957324106
ratio_actual_mean_amt_cardstate_1	0.169277792364425	0.000000000000000	0.071510957324106
ratio_actual_mean_amt_cardzip_30	0.165502704381739	0.000000000000000	0.065743944636678
ratio_actual_sum_amt_merch_14	0.165088962178461	0.000000000000000	0.025374855824683

	KSStatistic	PValue	FDR
ratio_actual_median_amt_cardmerch_30	0.165075143249710	0.000000000000000	0.077277797001153
ratio_actual_median_amt_card_1	0.163526727937537	0.000000000000000	0.063437139561707
ratio_actual_sum_amt_merch_30	0.162393401956917	0.000000000000000	0.014994232987313
ratio_actual_mean_amt_cardzip_1	0.160725048772562	0.000000000000000	0.092272202998847
ratio_actual_median_amt_cardstate_1	0.160163339737842	0.000000000000000	0.072664359861592
ratio_actual_median_amt_cardzip_30	0.159937760935488	0.000000000000000	0.057670126874279
ratio_actual_mean_amt_cardmerch_1	0.159567083381243	0.000000000000000	0.106113033448674
ratio_actual_sum_amt_cardstate_30	0.156544040621741	0.000000000000000	0.024221453287197
ratio_actual_median_amt_cardzip_1	0.153611994204393	0.000000000000000	0.083044982698962
ratio_actual_median_amt_cardmerch_1	0.153578315997531	0.000000000000000	0.102652825836217
rel_vel_vel_card_merch_7	0.152643408975020	0.000000000000000	0.114186851211073
ratio_actual_median_amt_merch_1	0.152289244606077	0.000000000000000	0.071510957324106
ratio_actual_max_amt_merch_3	0.147746510785674	0.000000000000000	0.035755478662053
rel_vel_vel_merch_card_14	0.146333286314211	0.000000000000000	0.093425605536332
rel_amt_amt_card_merch_7	0.143989413483815	0.000000000000001	0.100346020761246
rel_vel_vel_merch_merch_14	0.143192043912558	0.000000000000001	0.020761245674740
ratio_actual_max_amt_merch_7	0.142769480191869	0.000000000000001	0.016147635524798
ratio_actual_max_amt_card_7	0.141413313121195	0.000000000000003	0.020761245674740
rel_vel_vel_card_card_14	0.139168823592857	0.000000000000007	0.050749711649366
risk_dayofweek	0.135249114871540	0.000000000000047	0.083044982698962
ratio_actual_max_amt_merch_1	0.130634809429585	0.000000000000387	0.023068050749712
ratio_actual_max_amt_cardstate_30	0.125031146909379	0.000000000004522	0.018454440599769
ratio_actual_sum_amt_card_14	0.122927758203859	0.000000000011062	0.011534025374856
rel_vel_vel_card_merch_30	0.119747970551254	0.000000000041561	0.099192618223760
rel_vel_vel_merch_card_30	0.115940290755476	0.000000000193744	0.046136101499423
rel_vel_vel_merch_merch_30	0.115448371656525	0.000000000235517	0.016147635524798
rel_vel_vel_card_merch_14	0.106461852760270	0.000000007204655	0.111880046136102
rel_amt_amt_card_card_7	0.104673344424979	0.000000013769589	0.048442906574394
rel_vel_vel_card_card_7	0.102711056542276	0.000000027674447	0.085351787773933
rel_amt_amt_merch_card_14	0.093182687889412	0.000000680101234	0.041522491349481
merch_U*	0.080267595298679	0.000031722510816	0.016147635524798
rel_amt_amt_merch_card_7	0.057916998994521	0.006341147813008	0.023068050749712
random	0.040835282106478	0.114477598228336	0.043829296424452

```
In [72]: KSFDR['rank_KS']=KSFDR['KSStatistic'].rank(ascending=True)
         KSFDR['rank_FDR']=KSFDR['FDR'].rank(ascending=True)
```

```
In [73]: KSFDR['average_rank']=(KSFDR['rank_KS']+KSFDR['rank_FDR'])/2
```

```
In [74]: pd.set_option('display.float_format', lambda x: '%.3f' % x)
```

```
In [75]: rank_file=KSFDR.sort_values(by='average_rank',ascending=False)
```

```
In [76]: rank_file[0:80].to_csv("P3_first_list_fdr.csv")
```

Wrapper for Feature Selection

```
In [77]: rank_file=rank_file[1:81]
```

```
In [78]: #Taking 50% of the columns based on top 50% columns of the sorted table
         top_80=list(rank_file.T.columns)
```

In [79]: top_80

```
Out[79]: ['sum_amt_cardzip_7',
          'sum_amt_cardmerch_7',
          'sum_amt_cardmerch_14',
          'sum_amt_cardzip_14',
          'sum_amt_cardzip_3',
          'sum_amt_cardstate_3',
          'sum_amt_cardstate_7',
          'sum_amt_cardmerch_3',
          'sum_amt_cardmerch_30',
          'sum_amt_cardstate_14',
          'sum_amt_cardzip_30',
          'max_amt_cardzip_14',
          'sum_amt_cardstate_1',
          'sum_amt_cardmerch_1',
          'max_amt_cardstate_7',
          'sum_amt_cardzip_1',
          'max_amt_cardzip_30',
          'max_amt_cardmerch_14',
          'max_amt_cardmerch_30',
          'max_amt_cardstate_14',
          'max_amt_cardzip_7',
          'sum_amt_card_7',
          'max_amt_cardmerch_7',
          'max_amt_cardzip_3',
          'sum_amt_card_3',
          'max_amt_cardmerch_3',
          'max_amt_cardstate_3',
          'sum_amt_merch_1',
          'sum_amt_cardstate_30',
          'max_amt_cardstate_30',
          'sum_amt_merch_3',
          'max_amt_merch_1',
          'max_amt_cardstate_1',
          'max_amt_cardzip_1',
          'sum_amt_card_1',
          'max_amt_cardmerch_1',
          'max_amt_merch_3',
          'sum_amt_merch_7',
          'max_amt_card_1',
          'max_amt_card_3',
          'sum_amt_card_14',
          'max_amt_card_7',
          'mean_amt_card_3',
          'max_amt_card_14',
          'mean_amt_cardstate_7',
          'max_amt_merch_7',
          'mean_amt_cardstate_3',
          'mean_amt_cardstate_1',
          'mean_amt_cardmerch_1',
          'mean_amt_card_1',
          'mean_amt_cardzip_1',
          'mean_amt_cardzip_30',
          'mean_amt_cardstate_14',
          'mean_amt_merch_1',
          'mean_amt_cardzip_14',
          'mean_amt_cardstate_30',
          'mean_amt_cardzip_3',
```

```
                    'mean_amt_cardmerch_3',
                    'mean_amt_cardzip_7',
                    'mean_amt_cardmerch_7',
                    'mean_amt_cardmerch_30',
                    'mean_amt_cardmerch_14',
                    'mean_amt_card_7',
                    'mean_amt_card_14',
                    'median_amt_card_3',
                    'max_amt_merch_14',
                    'mean_amt_card_30',
                    'mean_amt_merch_3',
                    'median_amt_card_1',
                    'max_amt_card_30',
                    'median_amt_cardstate_1',
                    'sum_amt_card_30',
                    'median_amt_cardstate_3',
                    'median_amt_cardmerch_1',
                    'median_amt_cardzip_1',
                    'median_amt_cardzip_3',
                    'median_amt_cardmerch_3',
                    'rel_amt_vel_card_card_30',
                    'median_amt_card_14',
                    'median_amt_cardmerch_30']
```

In [82]:
```python
from sklearn.linear_model import LogisticRegression
model=LogisticRegression()
#feature selection using RFECV ( Recursive feature elimination & Cross Validation)
from sklearn.feature_selection import RFECV
rfecv=RFECV(estimator=model,step=1,cv=2,verbose=3,n_jobs=-1,scoring='roc_auc')
rfecv.fit(data[top_80],data['Fraud'])
```

```
Fitting estimator with 80 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 79 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 78 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 77 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 76 features.
```

```
C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 75 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 74 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 73 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 72 features.
```

```
C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 71 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 70 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 69 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 68 features.
```

```
C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 67 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 66 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 65 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 64 features.
```

```
C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 63 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 62 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 61 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 60 features.
```

```
C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 59 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 58 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 57 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 56 features.
```

```
C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 55 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 54 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 53 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 52 features.
```

```
C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 51 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 50 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 49 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 48 features.
```

```
C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 47 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 46 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 45 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 44 features.
```

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 43 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 42 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 41 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 40 features.

```
C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 39 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 38 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 37 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 36 features.
```

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 35 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 34 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 33 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 32 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 31 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 30 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 29 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 28 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 27 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 26 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 25 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 24 features.

```
C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 23 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 22 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 21 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 20 features.
```

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 19 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 18 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 17 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 16 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 15 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 14 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 13 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 12 features.

```
C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 11 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 10 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 9 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 8 features.
```

```
C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 7 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 6 features.
Fitting estimator with 5 features.
Fitting estimator with 4 features.
Fitting estimator with 3 features.
Fitting estimator with 2 features.
```

Out[82]: RFECV(cv=2,
            estimator=LogisticRegression(C=1.0, class_weight=None, dual=False,
                                         fit_intercept=True, intercept_scaling=1,
                                         l1_ratio=None, max_iter=100,
                                         multi_class='auto', n_jobs=None,
                                         penalty='l2', random_state=None,
                                         solver='lbfgs', tol=0.0001, verbose=0,
                                         warm_start=False),
            min_features_to_select=1, n_jobs=-1, scoring='roc_auc', step=1,
            verbose=3)

In [83]:
```python
pd.set_option('display.max_rows',None)
selected=pd.DataFrame(sorted(zip(map(lambda x: round(x),rfecv.ranking_),data[t
op_80].columns)),columns=['ranking','variable'])
print(selected)
```

	ranking	variable
0	1	rel_amt_vel_card_card_30
1	2	max_amt_cardmerch_14
2	3	mean_amt_cardmerch_3
3	4	mean_amt_cardzip_3
4	5	max_amt_cardstate_14
5	6	max_amt_cardstate_30
6	7	max_amt_merch_14
7	8	mean_amt_cardzip_14
8	9	mean_amt_card_30
9	10	mean_amt_cardmerch_30
10	11	mean_amt_card_14
11	12	mean_amt_merch_3
12	13	median_amt_cardzip_1
13	14	mean_amt_cardstate_1
14	15	max_amt_cardstate_1
15	16	mean_amt_cardmerch_7
16	17	median_amt_cardmerch_3
17	18	sum_amt_cardmerch_7
18	19	sum_amt_merch_3
19	20	sum_amt_cardmerch_3
20	21	sum_amt_merch_7
21	22	median_amt_cardmerch_1
22	23	mean_amt_merch_1
23	24	mean_amt_card_1
24	25	sum_amt_cardzip_1
25	26	max_amt_card_1
26	27	max_amt_cardmerch_1
27	28	sum_amt_card_1
28	29	max_amt_cardzip_7
29	30	max_amt_cardzip_30
30	31	max_amt_merch_3
31	32	max_amt_merch_1
32	33	mean_amt_cardstate_3
33	34	median_amt_card_3
34	35	mean_amt_cardzip_7
35	36	median_amt_card_14
36	37	mean_amt_cardstate_30
37	38	max_amt_cardmerch_7
38	39	max_amt_cardmerch_3
39	40	sum_amt_card_14
40	41	max_amt_cardstate_3
41	42	median_amt_card_1
42	43	sum_amt_cardstate_14
43	44	sum_amt_cardstate_30
44	45	sum_amt_card_30
45	46	mean_amt_cardmerch_1
46	47	mean_amt_card_3
47	48	mean_amt_cardstate_7
48	49	max_amt_cardzip_1
49	50	max_amt_cardzip_3
50	51	max_amt_cardstate_7
51	52	sum_amt_cardzip_7
52	53	sum_amt_cardzip_3
53	54	max_amt_card_3
54	55	max_amt_card_7
55	56	mean_amt_cardmerch_14

```
56        57    median_amt_cardstate_1
57        58     median_amt_cardzip_3
58        59      mean_amt_cardzip_30
59        60     sum_amt_cardmerch_1
60        61          max_amt_card_14
61        62       sum_amt_cardzip_14
62        63          max_amt_merch_7
63        64    mean_amt_cardstate_14
64        65     sum_amt_cardmerch_14
65        66          max_amt_card_30
66        67     max_amt_cardmerch_30
67        68     sum_amt_cardmerch_30
68        69       mean_amt_cardzip_1
69        70       max_amt_cardzip_14
70        71           sum_amt_card_3
71        72       sum_amt_cardzip_30
72        73          sum_amt_merch_1
73        74          mean_amt_card_7
74        75    median_amt_cardstate_3
75        76      sum_amt_cardstate_3
76        77      sum_amt_cardstate_1
77        78      sum_amt_cardstate_7
78        79   median_amt_cardmerch_30
79        80           sum_amt_card_7
```

In [87]:
```python
for i in range(0,5):
    s2=list(selected[selected['ranking']<=30].variable)
    from sklearn.feature_selection import RFECV
    rfecv=RFECV(estimator=model,step=1,cv=2,verbose=3,n_jobs=-1,scoring='roc_a
uc')
    rfecv.fit(data[s2],data['Fraud'])
    pd.set_option('display.max_rows',None)
    selected=pd.DataFrame(sorted(zip(map(lambda x: round(x),rfecv.ranking_),da
ta[s2].columns)),columns=['ranking','variable'])
    print(selected)
```

Fitting estimator with 29 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 28 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 27 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 26 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 25 features.

```
C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 24 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 23 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 22 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 21 features.
```

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 20 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 19 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 18 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 17 features.

```
C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 16 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 15 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 14 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 13 features.
```

```
C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 12 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 11 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 10 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 9 features.
```

```
C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 8 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)
```

	ranking	variable
0	1	max_amt_cardmerch_14
1	1	max_amt_cardstate_14
2	1	mean_amt_card_30
3	1	mean_amt_cardmerch_3
4	1	mean_amt_cardzip_14
5	1	mean_amt_cardzip_3
6	1	rel_amt_vel_card_card_30
7	2	max_amt_cardstate_30
8	3	max_amt_merch_14
9	4	mean_amt_merch_3
10	5	mean_amt_card_14
11	6	mean_amt_cardstate_1
12	7	mean_amt_cardmerch_7
13	8	median_amt_cardmerch_3
14	9	mean_amt_cardmerch_30
15	10	median_amt_cardzip_1
16	11	max_amt_cardstate_1
17	12	sum_amt_cardmerch_7
18	13	sum_amt_merch_3
19	14	sum_amt_cardmerch_3
20	15	sum_amt_merch_7
21	16	median_amt_cardmerch_1
22	17	mean_amt_merch_1
23	18	mean_amt_card_1
24	19	sum_amt_cardzip_1
25	20	max_amt_card_1
26	21	max_amt_cardmerch_1
27	22	sum_amt_card_1
28	23	max_amt_cardzip_7

```
Fitting estimator with 29 features.
```

```
C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 28 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 27 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 26 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 25 features.
```

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 24 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 23 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 22 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 21 features.

```
C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 20 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 19 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 18 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 17 features.
```

```
C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 16 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 15 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 14 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 13 features.
```

```
C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 12 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 11 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 10 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 9 features.
```

```
C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 8 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)
```

	ranking	variable
0	1	max_amt_cardmerch_14
1	1	max_amt_cardstate_14
2	1	mean_amt_card_30
3	1	mean_amt_cardmerch_3
4	1	mean_amt_cardzip_14
5	1	mean_amt_cardzip_3
6	1	rel_amt_vel_card_card_30
7	2	max_amt_cardstate_30
8	3	max_amt_merch_14
9	4	mean_amt_merch_3
10	5	mean_amt_card_14
11	6	mean_amt_cardstate_1
12	7	mean_amt_cardmerch_7
13	8	median_amt_cardmerch_3
14	9	mean_amt_cardmerch_30
15	10	median_amt_cardzip_1
16	11	max_amt_cardstate_1
17	12	sum_amt_cardmerch_7
18	13	sum_amt_merch_3
19	14	sum_amt_cardmerch_3
20	15	sum_amt_merch_7
21	16	median_amt_cardmerch_1
22	17	mean_amt_merch_1
23	18	mean_amt_card_1
24	19	sum_amt_cardzip_1
25	20	max_amt_card_1
26	21	max_amt_cardmerch_1
27	22	sum_amt_card_1
28	23	max_amt_cardzip_7

```
Fitting estimator with 29 features.
```

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 28 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 27 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 26 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 25 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 24 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 23 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 22 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 21 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 20 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 19 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 18 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 17 features.

```
C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 16 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 15 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 14 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 13 features.
```

```
C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 12 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 11 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 10 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 9 features.
```

```
C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 8 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)
```

	ranking	variable
0	1	max_amt_cardmerch_14
1	1	max_amt_cardstate_14
2	1	mean_amt_card_30
3	1	mean_amt_cardmerch_3
4	1	mean_amt_cardzip_14
5	1	mean_amt_cardzip_3
6	1	rel_amt_vel_card_card_30
7	2	max_amt_cardstate_30
8	3	max_amt_merch_14
9	4	mean_amt_merch_3
10	5	mean_amt_card_14
11	6	mean_amt_cardstate_1
12	7	mean_amt_cardmerch_7
13	8	median_amt_cardmerch_3
14	9	mean_amt_cardmerch_30
15	10	median_amt_cardzip_1
16	11	max_amt_cardstate_1
17	12	sum_amt_cardmerch_7
18	13	sum_amt_merch_3
19	14	sum_amt_cardmerch_3
20	15	sum_amt_merch_7
21	16	median_amt_cardmerch_1
22	17	mean_amt_merch_1
23	18	mean_amt_card_1
24	19	sum_amt_cardzip_1
25	20	max_amt_card_1
26	21	max_amt_cardmerch_1
27	22	sum_amt_card_1
28	23	max_amt_cardzip_7

```
Fitting estimator with 29 features.
```

```
C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 28 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 27 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 26 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 25 features.
```

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 24 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 23 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 22 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 21 features.

```
C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 20 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 19 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 18 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 17 features.
```

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 16 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 15 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 14 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 13 features.

```
C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 12 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 11 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 10 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 9 features.
```

```
C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 8 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)
```

```
    ranking                   variable
0         1        max_amt_cardmerch_14
1         1        max_amt_cardstate_14
2         1             mean_amt_card_30
3         1        mean_amt_cardmerch_3
4         1          mean_amt_cardzip_14
5         1           mean_amt_cardzip_3
6         1   rel_amt_vel_card_card_30
7         2        max_amt_cardstate_30
8         3            max_amt_merch_14
9         4            mean_amt_merch_3
10        5            mean_amt_card_14
11        6        mean_amt_cardstate_1
12        7        mean_amt_cardmerch_7
13        8      median_amt_cardmerch_3
14        9        mean_amt_cardmerch_30
15       10        median_amt_cardzip_1
16       11         max_amt_cardstate_1
17       12         sum_amt_cardmerch_7
18       13             sum_amt_merch_3
19       14         sum_amt_cardmerch_3
20       15             sum_amt_merch_7
21       16      median_amt_cardmerch_1
22       17            mean_amt_merch_1
23       18             mean_amt_card_1
24       19           sum_amt_cardzip_1
25       20             max_amt_card_1
26       21        max_amt_cardmerch_1
27       22             sum_amt_card_1
28       23           max_amt_cardzip_7
Fitting estimator with 29 features.
```

```
C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 28 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 27 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 26 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 25 features.
```

```
C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 24 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 23 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 22 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 21 features.
```

```
C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 20 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 19 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 18 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 17 features.
```

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 16 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 15 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 14 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 13 features.

```
C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 12 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 11 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 10 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 9 features.
```

```
C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)

Fitting estimator with 8 features.

C:\Users\jayan\AppData\Local\Continuum\anaconda3\lib\site-packages\sklearn\li
near_model\_logistic.py:940: ConvergenceWarning: lbfgs failed to converge (st
atus=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regres
sion
  extra_warning_msg=_LOGISTIC_SOLVER_CONVERGENCE_MSG)
```

	ranking	variable
0	1	max_amt_cardmerch_14
1	1	max_amt_cardstate_14
2	1	mean_amt_card_30
3	1	mean_amt_cardmerch_3
4	1	mean_amt_cardzip_14
5	1	mean_amt_cardzip_3
6	1	rel_amt_vel_card_card_30
7	2	max_amt_cardstate_30
8	3	max_amt_merch_14
9	4	mean_amt_merch_3
10	5	mean_amt_card_14
11	6	mean_amt_cardstate_1
12	7	mean_amt_cardmerch_7
13	8	median_amt_cardmerch_3
14	9	mean_amt_cardmerch_30
15	10	median_amt_cardzip_1
16	11	max_amt_cardstate_1
17	12	sum_amt_cardmerch_7
18	13	sum_amt_merch_3
19	14	sum_amt_cardmerch_3
20	15	sum_amt_merch_7
21	16	median_amt_cardmerch_1
22	17	mean_amt_merch_1
23	18	mean_amt_card_1
24	19	sum_amt_cardzip_1
25	20	max_amt_card_1
26	21	max_amt_cardmerch_1
27	22	sum_amt_card_1
28	23	max_amt_cardzip_7

```
In [88]: selected['variable'].to_csv('P3_Top29Variables.csv')
```

```
In [89]: selected
```

Out[89]:

	ranking	variable
0	1	max_amt_cardmerch_14
1	1	max_amt_cardstate_14
2	1	mean_amt_card_30
3	1	mean_amt_cardmerch_3
4	1	mean_amt_cardzip_14
5	1	mean_amt_cardzip_3
6	1	rel_amt_vel_card_card_30
7	2	max_amt_cardstate_30
8	3	max_amt_merch_14
9	4	mean_amt_merch_3
10	5	mean_amt_card_14
11	6	mean_amt_cardstate_1
12	7	mean_amt_cardmerch_7
13	8	median_amt_cardmerch_3
14	9	mean_amt_cardmerch_30
15	10	median_amt_cardzip_1
16	11	max_amt_cardstate_1
17	12	sum_amt_cardmerch_7
18	13	sum_amt_merch_3
19	14	sum_amt_cardmerch_3
20	15	sum_amt_merch_7
21	16	median_amt_cardmerch_1
22	17	mean_amt_merch_1
23	18	mean_amt_card_1
24	19	sum_amt_cardzip_1
25	20	max_amt_card_1
26	21	max_amt_cardmerch_1
27	22	sum_amt_card_1
28	23	max_amt_cardzip_7

Models+FPR+Charts_Part6

February 12, 2021

```python
[226]:  # Import relevant libraries
        import pandas as pd
        import numpy as np
        import warnings
        warnings.filterwarnings("ignore")
```

```python
[227]:  # Load data
        data=pd.read_csv("Final Variable Data.csv")
        data.head()
```

```
[227]:    Unnamed: 0        Date  Recnum  mean_amt_card_1  max_amt_card_1  \
       0           0  2010-01-01       1             3.62            3.62
       1           1  2010-01-01       2            31.42           31.42
       2           2  2010-01-01       3           178.49          178.49
       3           3  2010-01-01       4             3.62            3.62
       4           4  2010-01-01       5             3.62            3.62

          median_amt_card_1  sum_amt_card_1  mean_amt_card_3  max_amt_card_3  \
       0               3.62            3.62             3.62            3.62
       1              31.42           31.42            31.42           31.42
       2             178.49          178.49           178.49          178.49
       3               3.62            3.62             3.62            3.62
       4               3.62            7.24             3.62            3.62

          median_amt_card_3  …  rel_amt_amt_card_merch_30  dayssince_card  \
       0               3.62  …                   1.000000             337
       1              31.42  …                   1.000000             337
       2             178.49  …                   1.000000             337
       3               3.62  …                   0.500000             337
       4               3.62  …                   0.666667             357

          dayssince_merch  dayssince_cardmerch  dayssince_cardzip  \
       0              345                  345                345
       1              345                  345                345
       2              345                  345                345
       3              365                  345                345
       4              365                  365                365
```

	dayssince_cardstate	risk_dayofweek	card_U*	merch_U*	Fraud
0	345	0.026	1.000000	1.000894	0
1	345	0.026	1.000894	1.000894	0
2	345	0.026	1.000894	1.000894	0
3	345	0.026	1.000000	1.010684	0
4	365	0.026	1.000000	1.031246	0

[5 rows x 285 columns]

1 Z Scaled Data

```
[228]: variables=list(["max_amt_cardmerch_14",
       "max_amt_cardstate_14",
       "mean_amt_card_30",
       "mean_amt_cardmerch_3",
       "mean_amt_cardzip_14",
       "mean_amt_cardzip_3",
       "max_amt_cardstate_30",
       "max_amt_merch_14",
       "mean_amt_merch_3",
       "mean_amt_card_14",
       "mean_amt_cardstate_1",
       "mean_amt_cardmerch_7",
       "median_amt_cardmerch_3",
       "mean_amt_cardmerch_30",
       "median_amt_cardzip_1",
       "max_amt_cardstate_1",
       "sum_amt_cardmerch_7",
       "sum_amt_merch_3",
       "sum_amt_cardmerch_3",
       "sum_amt_merch_7",
       "median_amt_cardmerch_1",
       "mean_amt_merch_1",
       "mean_amt_card_1",
       "sum_amt_cardzip_1",
       "max_amt_card_1",
       "max_amt_cardmerch_1",
       "sum_amt_card_1",
       "max_amt_cardzip_7"])

       len(variables)
```

[228]: 28

```
[229]:  #Z scaled data 28 variables
        df2=(data[variables]-np.mean(data[variables]))/np.std(data[variables])
```

```
[230]:  # Add fraud label & record number
        df2=df2.merge(data[['Fraud','Recnum']],right_index=True,left_index=True)
```

```
[231]:  cols=list(df2.columns)
```

```
[232]:  cols=cols[:-3:-1]+cols[:-2]
```

```
[233]:  df2=df2[cols]
```

```
[234]:  df2.head()
```

```
[234]:      Recnum  Fraud  max_amt_cardmerch_14  max_amt_cardstate_14  \
        0        1      0             -0.452618             -0.498582
        1        2      0             -0.425922             -0.475100
        2        3      0             -0.284692             -0.350873
        3        4      0             -0.452618             -0.498582
        4        5      0             -0.452618             -0.498582

            mean_amt_card_30  mean_amt_cardmerch_3  mean_amt_cardzip_14  \
        0          -0.816115             -0.487468             -0.499234
        1          -0.758403             -0.453101             -0.464396
        2          -0.453089             -0.271293             -0.280092
        3          -0.816115             -0.487468             -0.499234
        4          -0.816115             -0.487468             -0.499234

            mean_amt_cardzip_3  max_amt_cardstate_30  max_amt_merch_14  ...  \
        0            -0.489201             -0.540252         -0.623489  ...
        1            -0.454703             -0.518761         -0.604729  ...
        2            -0.272202             -0.405068         -0.505484  ...
        3            -0.489201             -0.540252         -0.623489  ...
        4            -0.489201             -0.540252         -0.623489  ...

            sum_amt_cardmerch_3  sum_amt_merch_7  median_amt_cardmerch_1  \
        0             -0.150503        -0.368934               -0.486834
        1             -0.143646        -0.364296               -0.452142
        2             -0.107372        -0.339763               -0.268611
        3             -0.150503        -0.368330               -0.486834
        4             -0.149610        -0.367726               -0.486834

            mean_amt_merch_1  mean_amt_card_1  sum_amt_cardzip_1  max_amt_card_1  \
        0          -0.505163        -0.536482          -0.200453       -0.479737
        1          -0.469345        -0.498234          -0.189846       -0.452772
        2          -0.279852        -0.295893          -0.133728       -0.310117
        3          -0.505163        -0.536482          -0.200453       -0.479737
```

```
4          -0.505163          -0.536482          -0.199072          -0.479737
```

```
   max_amt_cardmerch_1  sum_amt_card_1  max_amt_cardzip_7
0             -0.445986       -0.215078          -0.444981
1             -0.416271       -0.206977          -0.417856
2             -0.259068       -0.164117          -0.274359
3             -0.445986       -0.215078          -0.444981
4             -0.445986       -0.214023          -0.444981

[5 rows x 30 columns]
```

2 create Train Test OOT

```
[235]: df2=df2.merge(data['Date'],right_index=True,left_index=True)
```

```
[236]: df2['Date']=pd.to_datetime(df2['Date'])
```

```
[237]: train_test=df2[df2.Date<'2010-11-01']
       oot=df2[df2.Date>='2010-11-01']
```

```
[238]: train_test.drop(columns='Date',inplace=True)
       oot.drop(columns='Date',inplace=True)
```

```
[239]: oot
```

```
[239]:        Recnum  Fraud  max_amt_cardmerch_14  max_amt_cardstate_14  \
       83970   84300      0             -0.281322             -0.042827
       83971   84301      0             -0.240029             -0.311587
       83972   84302      0             -0.405199             -0.326876
       83973   84303      0             -0.432087             -0.052904
       83974   84304      0             -0.362783             -0.419562
       ...       ...    ...                   ...                   ...
       96392   96749      0             -0.374671             -0.430020
       96393   96750      0              0.119934              0.146652
       96394   96751      0             -0.106972             -0.194548
       96395   96752      0              1.658488              1.358370
       96396   96753      0              0.076520              2.152782

              mean_amt_card_30  mean_amt_cardmerch_3  mean_amt_cardzip_14  \
       83970         -0.255705             -0.266954            -0.275693
       83971         -0.408839             -0.276090            -0.253380
       83972         -0.354197             -0.426424            -0.437353
       83973         -0.623441             -0.461038            -0.472442
       83974         -0.137041             -0.371821            -0.382000
       ...                 ...                   ...                  ...
       96392         -0.647608             -0.387125            -0.397514
```

```
96393           0.480880              -0.345144              -0.053505
96394           0.094349              -0.042510              -0.048166
96395           0.973368               2.230207               2.255759
96396          -0.154800              -0.194513               0.706897


        mean_amt_cardzip_3  max_amt_cardstate_30  max_amt_merch_14  …  \
83970          -0.267846              -0.123143          -0.503115  …
83971          -0.277017              -0.369113          -0.048966  …
83972          -0.427924              -0.284541          -0.590166  …
83973          -0.462670              -0.132365          -0.609061  …
83974          -0.373113              -0.382255          -0.560360  …
…                    …                     …                 …   …
96392          -0.388476              -0.477503          -0.568714  …
96393          -0.346334               0.527630          -0.110691  …
96394          -0.042546              -0.261999           1.052991  …
96395           2.238841               1.159239           0.959847  …
96396           0.705131               1.886289          -0.251653  …


        sum_amt_cardmerch_3   sum_amt_merch_7   median_amt_cardmerch_1  \
83970          -0.106506           -0.339178               -0.264231
83971          -0.108329           -0.130892               -0.273453
83972          -0.138323           -0.360697               -0.425212
83973          -0.145229           -0.365367               -0.460154
83974          -0.127429           -0.353328               -0.370091
…                    …                  …                      …
96392          -0.130482           -0.355394               -0.385541
96393          -0.122106           -0.349729               -0.343161
96394          -0.061725            0.705117               -0.037659
96395           0.391727            0.584844                2.256599
96396           0.026634           -0.249131                0.200793


        mean_amt_merch_1  mean_amt_card_1  sum_amt_cardzip_1  max_amt_card_1  \
83970         -0.275330        -0.291064          -0.132389        -0.306712
83971         -0.284851        -0.301231          -0.135209        -0.313881
83972         -0.441540        -0.468544          -0.181612        -0.431840
83973         -0.477616        -0.507067          -0.192295        -0.458999
83974         -0.384629        -0.407774          -0.164758        -0.388996
…                   …                …                  …                …
96392         -0.400580        -0.424807          -0.169481        -0.401004
96393         -0.356824         0.011443          -0.156523         0.181186
96394          1.352268        -0.041271          -0.063111        -0.130603
96395          2.327376         2.488123           0.638393         1.652671
96396          0.204798         0.221620           0.009799         0.054740


        max_amt_cardmerch_1  sum_amt_card_1  max_amt_cardzip_7
83970          -0.255316         -0.163094          -0.270934
83971          -0.263215         -0.165248          -0.278145
```

```
83972          -0.393204          -0.200688          -0.396800
83973          -0.423133          -0.208848          -0.424120
83974          -0.345991          -0.187816          -0.353703
...                  ...                ...                ...
96392          -0.359224          -0.191423          -0.365783
96393          -0.322924           0.018099          -0.332648
96394          -0.061247          -0.110183          -0.093785
96395           1.903892           0.425591           1.700023
96396           0.142999          -0.054498           2.617665

[12427 rows x 30 columns]
```

3   Model starts :) Are you ready ? Funn !!!

```
[240]: train_test_backup=train_test.copy()
       oot_backup=oot.copy()
```

```
[241]: X,y=train_test.drop(axis=1,labels=['Fraud','Recnum']).
       ↪values,train_test['Fraud'].values.ravel()
```

```
[242]: from sklearn.model_selection import train_test_split
       X_train,X_test,y_train,y_test=train_test_split(X,y,test_size=0.3)
```

4   Add OoooOt

```
[243]: X_oot,y_oot=oot.drop(axis=1,labels=['Fraud','Recnum']).values,oot['Fraud'].
       ↪values.ravel()
```

5   Model result Evaluation :D

```
[244]: def FDR_top3perc_evaluation(model, X_test, y_test):
           '''calculate the percentage of fraud caught at top 3 percent in decreasing␣
       ↪order'''
           result = model.predict_proba(X_test)[:,1]
           test_pred =  pd.DataFrame(y_test, columns = ['Fraud'])
           test_pred['predicted']= result.tolist()
           test_pred = test_pred.sort_values(by='predicted', ascending = False)


           topRows= int(round(len(test_pred)*0.03)) # number of rows to slice for␣
       ↪evaluaiton
           top = test_pred.head(topRows)

           total_num_fraud= y_test.sum() # total number of fraud in test dataset
```

```
        FDR = top.loc[:,'Fraud'].sum() / total_num_fraud
        return FDR
```

6 Logistic Model

```
[245]:  # from sklearn import linear_model
        # c=[0.01,0.001,0.0001]
        # for i in c:
        #     for j in range(0,10):
        #         avg_value_train=[]
        #         avg_value_test=[]
        #         avg_value_oot=[]
        #         from sklearn.model_selection import train_test_split
        #         X_train,X_test,y_train,y_test=train_test_split(X,y,test_size=0.3)
        #         clf =linear_model.LogisticRegression(C=i,random_state=None).
        ↪fit(X_train, y_train)
        #         avg_value_train.append(FDR_top3perc_evaluation(clf, X_train, y_train))
        #         avg_value_test.append(FDR_top3perc_evaluation(clf, X_test, y_test))
        #         avg_value_oot.append(FDR_top3perc_evaluation(clf, X_oot, y_oot))
        #     print(f'Avg train at {i} is {np.mean(avg_value_train)}')
        #     print(f'Avg test at {i} is {np.mean(avg_value_test)}')
        #     print(f'Avg oot at {i} is {np.mean(avg_value_oot)}')
```

7 Boosted Tree

```
[246]:  # from sklearn.ensemble import GradientBoostingClassifier
        # n= list(np.arange(30,60,10))
        # learning=[0.01,0.1]
        # depth=[3,4,5]
        # for i in n :
        #     for j in learning:
        #         for k in depth:
        #             for t in range(0,10):
        #                 avg_value_train=[]
        #                 avg_value_test=[]
        #                 avg_value_oot=[]
        #                 from sklearn.model_selection import train_test_split
        #                    ␣
        ↪X_train,X_test,y_train,y_test=train_test_split(X,y,test_size=0.3)
        #                 bt=GradientBoostingClassifier(n_estimators=i,␣
        ↪learning_rate=j,max_depth=k).fit(X_train, y_train)
        #                 avg_value_train.append(FDR_top3perc_evaluation(bt, X_train,␣
        ↪y_train))
```

```
#                avg_value_test.append(FDR_top3perc_evaluation(bt, X_test,␣
 ↪y_test))
#                avg_value_oot.append(FDR_top3perc_evaluation(bt, X_oot,␣
 ↪y_oot))
#            print(f'Average train for N={i} LR={j} Depth={k} is {np.
 ↪mean(avg_value_train)}')
#            print(f'Average test for N={i} LR={j} Depth={k} is {np.
 ↪mean(avg_value_test)}')
#            print(f'Average oot for N={i} LR={j} Depth={k} is {np.
 ↪mean(avg_value_oot)}\n')
```

8   Random Forest

[247]:
```
# from sklearn.ensemble import RandomForestClassifier
# n=[100,200,300]
# max_depth=[2,3]
# for i in n:
#     for j in max_depth:
#         for p in range(0,10):
#             avg_value_train=[]
#             avg_value_test=[]
#             avg_value_oot=[]
#             from sklearn.model_selection import train_test_split
#             X_train,X_test,y_train,y_test=train_test_split(X,y,test_size=0.3)
#             rfc = RandomForestClassifier(max_depth=j, n_estimators = i,  ␣
 ↪n_jobs = -1,max_features = 5).fit(X_train, y_train)
#             avg_value_train.append(FDR_top3perc_evaluation(rfc, X_train,␣
 ↪y_train))
#             avg_value_test.append(FDR_top3perc_evaluation(rfc, X_test,␣
 ↪y_test))
#             avg_value_oot.append(FDR_top3perc_evaluation(rfc, X_oot, y_oot))
#         print(f'Average train for N={i} LR={j} Depth={k} is {np.
 ↪mean(avg_value_train)}')
#         print(f'Average test for N={i} LR={j} Depth={k} is {np.
 ↪mean(avg_value_test)}')
#         print(f'Average oot for N={i} LR={j} Depth={k} is {np.
 ↪mean(avg_value_oot)}')
```

9 Neural Network

```
[248]: # from sklearn.neural_network import MLPClassifier
       # size=[(4,2),(5,3),(6,4),(7,4)]
       # for i in size :
       #     for p in range(0,10):
       #         avg_value_train=[]
       #         avg_value_test=[]
       #         avg_value_oot=[]
       #         from sklearn.model_selection import train_test_split
       #         X_train,X_test,y_train,y_test=train_test_split(X,y,test_size=0.3)
       #         nn = MLPClassifier(solver='lbfgs', alpha=1e-5, hidden_layer_sizes=i).
       →fit(X_train, y_train)
       #         avg_value_train.append(FDR_top3perc_evaluation(nn, X_train, y_train))
       #         avg_value_test.append(FDR_top3perc_evaluation(nn, X_test, y_test))
       #         avg_value_oot.append(FDR_top3perc_evaluation(nn, X_oot, y_oot))
       #     print(f'Average train for N={i} LR={j} Depth={k} is {np.
       →mean(avg_value_train)}')
       #     print(f'Average test for N={i} LR={j} Depth={k} is {np.
       →mean(avg_value_test)}')
       #     print(f'Average oot for N={i} LR={j} Depth={k} is {np.
       →mean(avg_value_oot)}')
```

```
[249]: # X_train=train_test.iloc[:,2:]
       # y_train=train_test.iloc[:,1]
       # X_test=X_oot
       # y_test=y_oot
```

```
[250]: len(y_oot)
```

```
[250]: 12427
```

```
[251]: # def FDR_top3perc_evaluation(model, X_test, y_test):
       #     per=[]
       #     FDR_r=[]
       #     FPR_r=[]
       #     '''calculate the percentage of fraud caught at top 3 percent in␣
       →decreasing order'''
       #     for i in range(0,101,1):
       #         per.append(i)
       #         result = model.predict_proba(X_test)[:,1]
       #         result2=model.predict(X_test)
       #         test_pred =  pd.DataFrame(y_test, columns = ['Fraud'])
       #         test_pred['pred_val']=result2.tolist()
       #         test_pred['predicted']= result.tolist()
       # #         test_pred = test_pred.sort_values(by='predicted', ascending = False)
```

```
# #           test_pred['group'] =pd.qcut(test_pred['predicted'].
 →rank(method='first'), q=100)
#           #testing_tb = test_pred.groupby(['group']).sum().sort_index(␣
 →ascending = False).head(20)[['Fraud']].rename(columns={'Fraud_x':'# Bads'})
#           #testing_final = testing_tb.join(testing_tb.cumsum().rename(columns␣
 →={ '# Bads':'Cumulative Bads'}))
# #           topRows= int(round(len(test_pred)*i/100)) # number of rows to slice␣
 →for evaluaiton
# #           top = test_pred.head(topRows)
# #           total_num_fraud= y_test.sum() # total number of fraud in test␣
 →dataset
# #           FDR = (top.loc[:,'Fraud'].sum() / total_num_fraud)*100
# #           FDR_r.append(FDR)


#     return test_pred
```

```python
[252]: from sklearn.ensemble import GradientBoostingClassifier
       from sklearn.model_selection import train_test_split
       bt=GradientBoostingClassifier(n_estimators=40, learning_rate=0.1,max_depth=4).
        →fit(X_train, y_train)
```

```python
[253]: FDR_top3perc_evaluation(bt,X_train, y_train)
```

```
[253]: 0.8855799373040752
```

```python
[254]: FDR_top3perc_evaluation(bt,X_test, y_test)
```

```
[254]: 0.8181818181818182
```

```python
[255]: FDR_top3perc_evaluation(bt,X_oot, y_oot)
```

```
[255]: 0.5418994413407822
```

```python
[256]: result = bt.predict_proba(X_train)[:,1]
       result
```

```
[256]: array([0.00159051, 0.00159051, 0.00193992, …, 0.00193992, 0.00159051,
               0.00159051])
```

```python
[257]: train_pred =  pd.DataFrame(result, columns = ['fraud_label_predicted'])
       train_pred
```

```
[257]:      fraud_label_predicted
       0                  0.001591
       1                  0.001591
       2                  0.001940
       3                  0.001591
```

```
4                  0.001591
...                     ...
58774              0.001591
58775              0.001591
58776              0.001940
58777              0.001591
58778              0.001591

[58779 rows x 1 columns]
```

```
[278]: training = pd.DataFrame(y_train, columns =['fraud_label' ])\
       .join(train_pred).sort_values(by = 'fraud_label_predicted', ascending = False)
       training
```

```
[278]:        fraud_label  fraud_label_predicted
       2496             1               1.000000
       10835            1               1.000000
       36796            1               1.000000
       25025            1               1.000000
       20316            1               1.000000
       ...            ...                    ...
       9265             0               0.001437
       58437            0               0.001437
       34606            0               0.001339
       4048             0               0.001278
       28957            0               0.001278

[58779 rows x 2 columns]
```

```
[279]: training['group'] =pd.qcut(training['fraud_label_predicted'].
        →rank(method='first'), q=100)
       training
```

```
[279]:        fraud_label  fraud_label_predicted                group
       2496             1               1.000000  (58191.22, 58779.0]
       10835            1               1.000000  (58191.22, 58779.0]
       36796            1               1.000000  (58191.22, 58779.0]
       25025            1               1.000000  (58191.22, 58779.0]
       20316            1               1.000000  (58191.22, 58779.0]
       ...            ...                    ...                  ...
       9265             0               0.001437    (0.999, 588.78]
       58437            0               0.001437    (0.999, 588.78]
       34606            0               0.001339    (0.999, 588.78]
       4048             0               0.001278    (0.999, 588.78]
       28957            0               0.001278    (0.999, 588.78]

[58779 rows x 3 columns]
```

```
[280]: training_tb = training.groupby(['group']).sum().sort_index( ascending = False).
        →head(50)[['fraud_label']].rename(columns={'fraud_label':'# Bads'})
        training_tb
```

```
[280]:                      # Bads
        group
        (58191.22, 58779.0]     510
        (57603.44, 58191.22]     44
        (57015.66, 57603.44]     11
        (56427.88, 57015.66]      9
        (55840.1, 56427.88]       7
        (55252.32, 55840.1]       4
        (54664.54, 55252.32]      7
        (54076.76, 54664.54]      1
        (53488.98, 54076.76]      2
        (52901.2, 53488.98]       3
        (52313.42, 52901.2]       1
        (51725.64, 52313.42]      1
        (51137.86, 51725.64]      0
        (50550.08, 51137.86]      1
        (49962.3, 50550.08]       0
        (49374.52, 49962.3]       1
        (48786.74, 49374.52]      2
        (48198.96, 48786.74]      2
        (47611.18, 48198.96]      1
        (47023.4, 47611.18]       2
        (46435.62, 47023.4]       2
        (45847.84, 46435.62]      2
        (45260.06, 45847.84]      2
        (44672.28, 45260.06]      1
        (44084.5, 44672.28]       1
        (43496.72, 44084.5]       1
        (42908.94, 43496.72]      0
        (42321.16, 42908.94]      1
        (41733.38, 42321.16]      4
        (41145.6, 41733.38]       0
        (40557.82, 41145.6]       0
        (39970.04, 40557.82]      0
        (39382.26, 39970.04]      0
        (38794.48, 39382.26]      0
        (38206.7, 38794.48]       0
        (37618.92, 38206.7]       0
        (37031.14, 37618.92]      0
        (36443.36, 37031.14]      0
        (35855.58, 36443.36]      0
        (35267.8, 35855.58]       1
        (34680.02, 35267.8]       0
```

```
(34092.24, 34680.02]        2
(33504.46, 34092.24]        0
(32916.68, 33504.46]        0
(32328.9, 32916.68]         0
(31741.12, 32328.9]         0
(31153.34, 31741.12]        1
(30565.56, 31153.34]        0
(29977.78, 30565.56]        0
(29390.0, 29977.78]         0
```

[281]:
```python
training_final = training_tb.join(training_tb.cumsum().rename(columns ={ '#␣
↪Bads':'Cumulative Bads'}))
```

[282]:
```python
training_final.to_excel('trainingcumbad.xlsx')
training_final
```

[282]:
```
                        # Bads  Cumulative Bads
group
(58191.22, 58779.0]        510              510
(57603.44, 58191.22]        44              554
(57015.66, 57603.44]        11              565
(56427.88, 57015.66]         9              574
(55840.1, 56427.88]          7              581
(55252.32, 55840.1]          4              585
(54664.54, 55252.32]         7              592
(54076.76, 54664.54]         1              593
(53488.98, 54076.76]         2              595
(52901.2, 53488.98]          3              598
(52313.42, 52901.2]          1              599
(51725.64, 52313.42]         1              600
(51137.86, 51725.64]         0              600
(50550.08, 51137.86]         1              601
(49962.3, 50550.08]          0              601
(49374.52, 49962.3]          1              602
(48786.74, 49374.52]         2              604
(48198.96, 48786.74]         2              606
(47611.18, 48198.96]         1              607
(47023.4, 47611.18]          2              609
(46435.62, 47023.4]          2              611
(45847.84, 46435.62]         2              613
(45260.06, 45847.84]         2              615
(44672.28, 45260.06]         1              616
(44084.5, 44672.28]          1              617
(43496.72, 44084.5]          1              618
(42908.94, 43496.72]         0              618
(42321.16, 42908.94]         1              619
(41733.38, 42321.16]         4              623
```

```
(41145.6, 41733.38]          0              623
(40557.82, 41145.6]          0              623
(39970.04, 40557.82]         0              623
(39382.26, 39970.04]         0              623
(38794.48, 39382.26]         0              623
(38206.7, 38794.48]          0              623
(37618.92, 38206.7]          0              623
(37031.14, 37618.92]         0              623
(36443.36, 37031.14]         0              623
(35855.58, 36443.36]         0              623
(35267.8, 35855.58]          1              624
(34680.02, 35267.8]          0              624
(34092.24, 34680.02]         2              626
(33504.46, 34092.24]         0              626
(32916.68, 33504.46]         0              626
(32328.9, 32916.68]          0              626
(31741.12, 32328.9]          0              626
(31153.34, 31741.12]         1              627
(30565.56, 31153.34]         0              627
(29977.78, 30565.56]         0              627
(29390.0, 29977.78]          0              627
```

[283]: `training['fraud_label'].sum()`

[283]: 638

[284]:
```python
result = bt.predict_proba(X_test)[:,1]
test_pred =  pd.DataFrame(result, columns = ['fraud_label_predicted'])

testing = pd.DataFrame(y_test, columns =['fraud_label' ])\
.join(test_pred).sort_values(by = 'fraud_label_predicted', ascending = False)
testing['group'] =pd.qcut(testing['fraud_label_predicted'].
 →rank(method='first'), q=100)
testing
```

[284]:
```
        fraud_label  fraud_label_predicted              group
2652              0               1.000000  (24939.1, 25191.0]
22277             1               1.000000  (24939.1, 25191.0]
24803             0               1.000000  (24939.1, 25191.0]
17494             0               1.000000  (24939.1, 25191.0]
11894             0               1.000000  (24939.1, 25191.0]
...             ...                    ...                 ...
18491             0               0.001478     (0.999, 252.9]
24832             0               0.001478     (0.999, 252.9]
6859              0               0.001437     (0.999, 252.9]
11904             0               0.001278     (0.999, 252.9]
17072             0               0.001278     (0.999, 252.9]
```

14

```
[25191 rows x 3 columns]
```

```python
[285]: testing_tb = testing.groupby(['group']).sum().sort_index( ascending = False).
       →head(50)[['fraud_label']].rename(columns={'fraud_label':'# Bads'})
       testing_final = testing_tb.join(testing_tb.cumsum().rename(columns ={ '# Bads':
       →'Cumulative Bads'}))
       testing_final.to_excel('testingcumbad.xlsx')
       testing_final
```

```
[285]:                     # Bads  Cumulative Bads
       group
       (24939.1, 25191.0]     171              171
       (24687.2, 24939.1]      21              192
       (24435.3, 24687.2]       6              198
       (24183.4, 24435.3]       4              202
       (23931.5, 24183.4]       4              206
       (23679.6, 23931.5]       3              209
       (23427.7, 23679.6]       4              213
       (23175.8, 23427.7]       1              214
       (22923.9, 23175.8]       0              214
       (22672.0, 22923.9]       0              214
       (22420.1, 22672.0]       0              214
       (22168.2, 22420.1]       1              215
       (21916.3, 22168.2]       1              216
       (21664.4, 21916.3]       0              216
       (21412.5, 21664.4]       2              218
       (21160.6, 21412.5]       0              218
       (20908.7, 21160.6]       0              218
       (20656.8, 20908.7]       0              218
       (20404.9, 20656.8]       0              218
       (20153.0, 20404.9]       0              218
       (19901.1, 20153.0]       1              219
       (19649.2, 19901.1]       2              221
       (19397.3, 19649.2]       1              222
       (19145.4, 19397.3]       2              224
       (18893.5, 19145.4]       2              226
       (18641.6, 18893.5]       1              227
       (18389.7, 18641.6]       1              228
       (18137.8, 18389.7]       0              228
       (17885.9, 18137.8]       0              228
       (17634.0, 17885.9]       0              228
       (17382.1, 17634.0]       1              229
       (17130.2, 17382.1]       0              229
       (16878.3, 17130.2]       0              229
       (16626.4, 16878.3]       0              229
       (16374.5, 16626.4]       0              229
```

```
(16122.6, 16374.5]          0              229
(15870.7, 16122.6]          0              229
(15618.8, 15870.7]          0              229
(15366.9, 15618.8]          0              229
(15115.0, 15366.9]          0              229
(14863.1, 15115.0]          0              229
(14611.2, 14863.1]          0              229
(14359.3, 14611.2]          0              229
(14107.4, 14359.3]          0              229
(13855.5, 14107.4]          0              229
(13603.6, 13855.5]          0              229
(13351.7, 13603.6]          0              229
(13099.8, 13351.7]          0              229
(12847.9, 13099.8]          0              229
(12596.0, 12847.9]          0              229
```

[286]:
```python
result = bt.predict_proba(X_oot)[:,1]
oot_pred =  pd.DataFrame(result, columns = ['fraud_label_predicted'])

oot = pd.DataFrame(y_oot, columns =['fraud_label' ])\
.join(oot_pred).sort_values(by = 'fraud_label_predicted', ascending = False)
oot['group'] =pd.qcut(oot['fraud_label_predicted'].rank(method='first'), q=100)
oot
```

[286]:
```
        fraud_label  fraud_label_predicted                  group
12117             0               1.000000  (12302.74, 12427.0]
10261             1               1.000000  (12302.74, 12427.0]
5545              0               1.000000  (12302.74, 12427.0]
4068              0               0.996378  (12302.74, 12427.0]
5142              1               0.993669  (12302.74, 12427.0]
...             ...                    ...                  ...
8369              0               0.001571    (0.999, 125.26]
9925              0               0.001571    (0.999, 125.26]
11262             0               0.001571    (0.999, 125.26]
1062              0               0.001571    (0.999, 125.26]
1656              0               0.001278    (0.999, 125.26]

[12427 rows x 3 columns]
```

[293]:
```python
oot_tb = oot.groupby(['group']).sum().sort_index( ascending = False).
 ↪head(50)[['fraud_label']].rename(columns={'fraud_label':'# Bads'})
oot_final = oot_tb.join(oot_tb.cumsum().rename(columns ={ '# Bads':'Cumulative␣
 ↪Bads'}))
#oot_final.to_excel('ootcumbad.xlsx')

len(training),len(testing),len(oot)
```

```
[293]: (58779, 25191, 12427)
```

```
[296]: training['fraud_label'].sum(),testing['fraud_label'].sum(),oot['fraud_label'].
       ↪sum()
```

```
[296]: (638, 242, 179)
```

10   tIme charts

```
[22]: train_test_plot=p.merge(train_test,left_index=True,right_index=True)
```

```
[23]: train_test_plot.tail()
```

```
[23]:        Fraud_x  pred_val  predicted  Recnum  Fraud_y  max_amt_cardmerch_14  \
       83965        0         0   0.000299   84295        0             -0.168229
       83966        0         0   0.002399   84296        0              0.129048
       83967        0         0   0.000237   84297        0             -0.230426
       83968        0         0   0.001192   84298        0              0.120078
       83969        0         0   0.000163   84299        0             -0.427285

              max_amt_cardstate_14  mean_amt_card_30  mean_amt_cardmerch_3  \
       83965             -0.248430         -0.108704             -0.121367
       83966              0.013057         -0.004324             -0.004064
       83967             -0.303140         -0.363619             -0.201436
       83968              0.005168          0.421958              0.249778
       83969             -0.476300         -0.413292             -0.454857

              mean_amt_cardzip_14  …  sum_amt_cardmerch_3  sum_amt_merch_7  \
       83965            -0.128107  …            -0.077458        -0.031182
       83966            -0.009193  …             0.043287        -0.168508
       83967            -0.209275  …            -0.093434        -0.330336
       83968             0.248134  …            -0.003408        -0.269449
       83969            -0.466176  …            -0.143996        -0.026178

              median_amt_cardmerch_1  mean_amt_merch_1  mean_amt_card_1  \
       83965               -0.117263         -0.046701        -0.129034
       83966                0.001151         -0.159766         0.001517
       83967               -0.198091         -0.207042        -0.218146
       83968                0.257399          0.263242         0.284027
       83969               -0.453914         -0.188192        -0.500188

              sum_amt_cardzip_1  max_amt_card_1  max_amt_cardmerch_1  sum_amt_card_1  \
       83965          -0.087452       -0.192478            -0.129432       -0.128773
       83966           0.099346        0.107798             0.201467        0.013893
       83967          -0.112166       -0.255304            -0.198664       -0.147649
       83968           0.027107        0.098738             0.191484       -0.041279
```

```
83969           -0.190388        -0.454149           -0.417789        -0.207390

        max_amt_cardzip_7
83965           -0.156025
83966            0.146024
83967           -0.219222
83968            0.136911
83969           -0.419242

[5 rows x 33 columns]
```

[24]:
```python
d=c.merge(oot.reset_index(),left_index=True,right_index=True)
```

[25]:
```python
d.tail()
```

[25]:
```
       Fraud_x  pred_val  predicted  index  Recnum  Fraud_y  \
12422        0         0   0.000208  96392   96749        0
12423        0         0   0.000226  96393   96750        0
12424        0         0   0.001456  96394   96751        0
12425        0         0   0.002917  96395   96752        0
12426        0         0   0.003324  96396   96753        0

       max_amt_cardmerch_14  max_amt_cardstate_14  mean_amt_card_30  \
12422             -0.374671             -0.430020         -0.647608
12423              0.119934              0.146652          0.480880
12424             -0.106972             -0.194548          0.094349
12425              1.658488              1.358370          0.973368
12426              0.076520              2.152782         -0.154800

       mean_amt_cardmerch_3  …  sum_amt_cardmerch_3  sum_amt_merch_7  \
12422             -0.387125  …            -0.130482        -0.355394
12423             -0.345144  …            -0.122106        -0.349729
12424             -0.042510  …            -0.061725         0.705117
12425              2.230207  …             0.391727         0.584844
12426             -0.194513  …             0.026634        -0.249131

       median_amt_cardmerch_1  mean_amt_merch_1  mean_amt_card_1  \
12422               -0.385541         -0.400580        -0.424807
12423               -0.343161         -0.356824         0.011443
12424               -0.037659          1.352268        -0.041271
12425                2.256599          2.327376         2.488123
12426                0.200793          0.204798         0.221620

       sum_amt_cardzip_1  max_amt_card_1  max_amt_cardmerch_1  sum_amt_card_1  \
12422          -0.169481       -0.401004            -0.359224       -0.191423
12423          -0.156523        0.181186            -0.322924        0.018099
12424          -0.063111       -0.130603            -0.061247       -0.110183
```

```
12425           0.638393        1.652671           1.903892           0.425591
12426           0.009799        0.054740           0.142999          -0.054498


        max_amt_cardzip_7
12422          -0.365783
12423          -0.332648
12424          -0.093785
12425           1.700023
12426           2.617665

[5 rows x 34 columns]
```

[26]: ```python
d3=pd.read_excel('DQR3.xlsx')
```

[27]: ```python
d3_1=d3.merge(d,on='Recnum',how='inner')
```

[28]: ```python
d3_2=d3.merge(train_test_plot,on='Recnum',how='inner')
```

[29]: ```python
d3_1=d3_1[['Merchnum','Cardnum','predicted','Date','Recnum']]
```

[30]: ```python
d3_2=d3_2[['Merchnum','Cardnum','predicted','Date','Recnum']]
```

[31]: ```python
new=pd.concat([d3_2,d3_1])
```

[32]: ```python
d6=new.groupby(['Date','Merchnum'])['Merchnum'].size()
```

[33]: ```python
frauds_new=new[new['predicted']>0.5].sort_values(by='predicted')
```

[34]: ```python
frauds_new
```

[34]:
```
           Merchnum   Cardnum predicted   Date Recnum
6821   6899988049601 5142194136   0.500621 2010-12-06   91140
26244   6006333528866 5142189113   0.505546 2010-04-11   26309
53900   930090121224 5142212038   0.507688 2010-07-18   54089
53922   930090121224 5142212038   0.507688 2010-07-18   54111
35565   4620009957157 5142271065   0.508457 2010-05-14   35671
29411   6929 5142151962   0.514359 2010-04-23   29491
34243   9108234610000 5142189341   0.518019 2010-05-09   34347
6280   5000006000095 5142205500   0.522570 2010-12-02   90599
5087   4353000719908 5142235211   0.526487 2010-11-25   89401
3436   4503738417400 5142847398   0.529549 2010-01-15   3443
79536   8292309000040 5142241589   0.531455 2010-10-05   79861
1202   4503738417400 5142847398   0.554228 2010-01-07   1207
49735   410000971343 5142199728   0.557295 2010-06-30   49908
19562   1144000409955 5142126504   0.564936 2010-03-17   19595
1371   4503738417400 5142847398   0.566815 2010-01-07   1376
56974   6006333528866 5142189113   0.568334 2010-07-27   57170
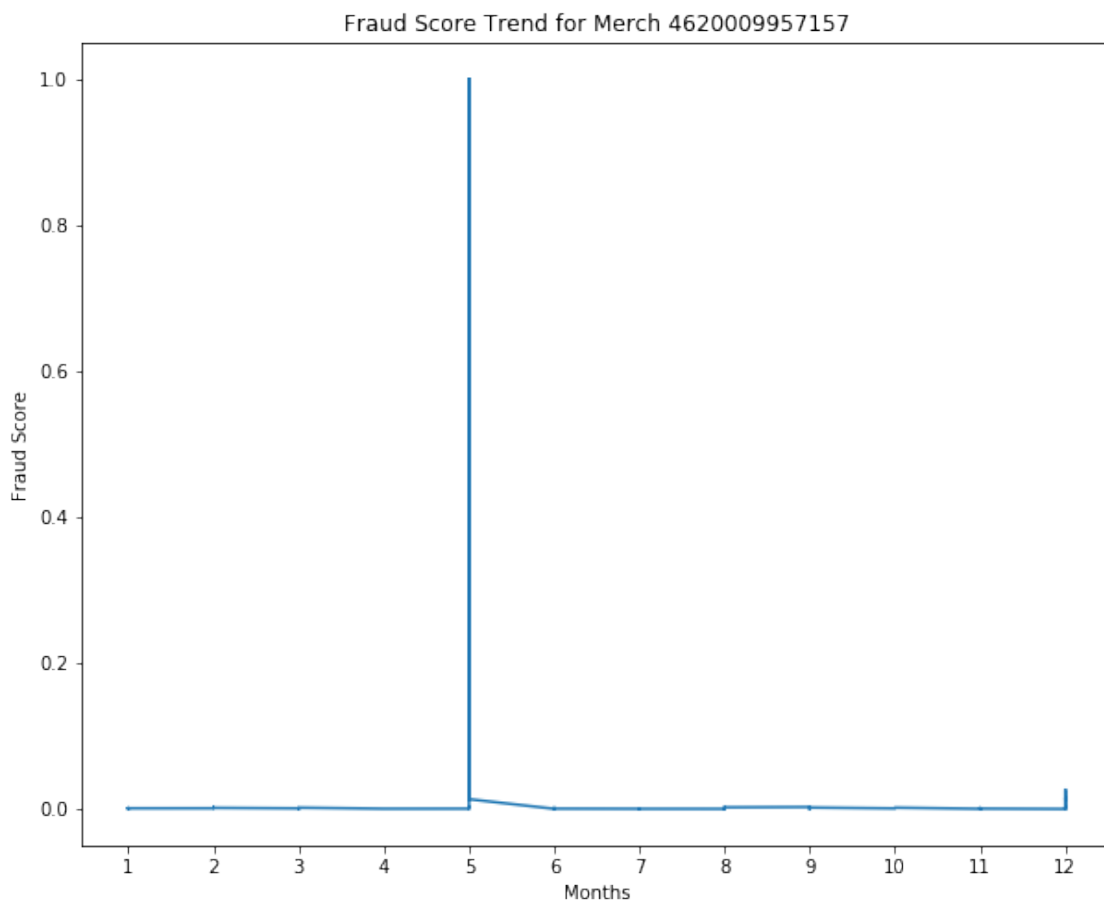```

```
69500  5000006000095  5142205500  0.576911  2010-09-02  69753
30189            6929  5142151962  0.577158  2010-04-26  30270
25273    997674930332  5142233702  0.596933  2010-04-06  25331
13657  4503738417400  5142847398  0.602361  2010-02-25  13677
59158    900009045549  5142138135  0.602470  2010-08-03  59360
36483  9900020006406  5142197711  0.608039  2010-05-17  36592
62498     92891948003  5142249750  0.616809  2010-08-12  62718
29460            6929  5142151962  0.625093  2010-04-24  29540
41907  4591200814444  5142111125  0.630242  2010-06-06  42042
35575  4620009957157  5142153880  0.634338  2010-05-14  35681
59857    602608969138  5142292976  0.647006  2010-08-05  60063
36487  9900020006406  5142197711  0.647122  2010-05-17  36596
35583  9108234610006  5142189341  0.647397  2010-05-14  35689
66879    938909877224  5142210575  0.649205  2010-08-25  67108
...              ...         ...       ...         ...     ...
54094    930090121224  5142212038  1.000000  2010-07-18  54283
9331   9900020006406  5142160007  1.000000  2010-12-15  93651
58242    938909877224  5142245297  1.000000  2010-08-01  58441
1042   4591200814444  5142152982  1.000000  2010-11-05  85349
43905  4503082476300  5142176939  1.000000  2010-06-12  44046
5909      680655463  5142184598  1.000000  2010-12-01  90225
53961    930090121224  5142212038  1.000000  2010-07-18  54150
53892    930090121224  5142212038  1.000000  2010-07-18  54081
53880    930090121224  5142212038  1.000000  2010-07-18  54069
53846    930090121224  5142212038  1.000000  2010-07-18  54035
53826    930090121224  5142212038  1.000000  2010-07-18  54015
53825    930090121224  5142212038  1.000000  2010-07-18  54014
6455   08-3500999175  5142228636  1.000000  2010-12-03  90774
59010  08-0963129334  5142197563  1.000000  2010-08-03  59211
60166  6000006800374  5142197563  1.000000  2010-08-07  60372
41050  6006333528866  5142189113  1.000000  2010-06-02  41181
63480    938909877224  5142245297  1.000000  2010-08-15  63703
36457    679613867334  5142197711  1.000000  2010-05-17  36566
65182    938909877224  5142210575  1.000000  2010-08-21  65407
65383    938909877224  5142245297  1.000000  2010-08-22  65608
66131    938909877224  5142243247  1.000000  2010-08-23  66358
33174  6006333528866  5142189113  1.000000  2010-05-05  33272
77695  6006333528866  5142189113  1.000000  2010-09-27  78010
79444  8292309000040  5142125194  1.000000  2010-10-05  79769
26238  6006333528866  5142189113  1.000000  2010-04-11  26303
25116    997674930332  5142234238  1.000000  2010-04-06  25174
3064   6899653660176  5142226204  1.000000  2010-11-16  87378
20943  6006333528866  5142189113  1.000000  2010-03-22  20979
59939  6887894784113  5142197563  1.000000  2010-08-05  60145
52895  6006333528866  5142189113  1.000000  2010-07-13  53079

[882 rows x 5 columns]
```

```
[35]:  # frauds_new[frauds_new['Merchnum']=='4620009957157']
```

```
[36]:  import matplotlib.pyplot as plt
       timeplot1=new[(new['Merchnum']=='4620009957157')][['Date','predicted']]
```

```
[37]:  plt.figure(figsize=(10,8))
       plt.plot(timeplot1['Date'].dt.month,timeplot1['predicted'])
       plt.xticks(np.arange(1,13,1))
       plt.ylabel('Fraud Score')
       plt.xlabel('Months')
       plt.title('Fraud Score Trend for Merch 4620009957157')
       plt.show()
```

Fraud Score Trend for Merch 4620009957157


```
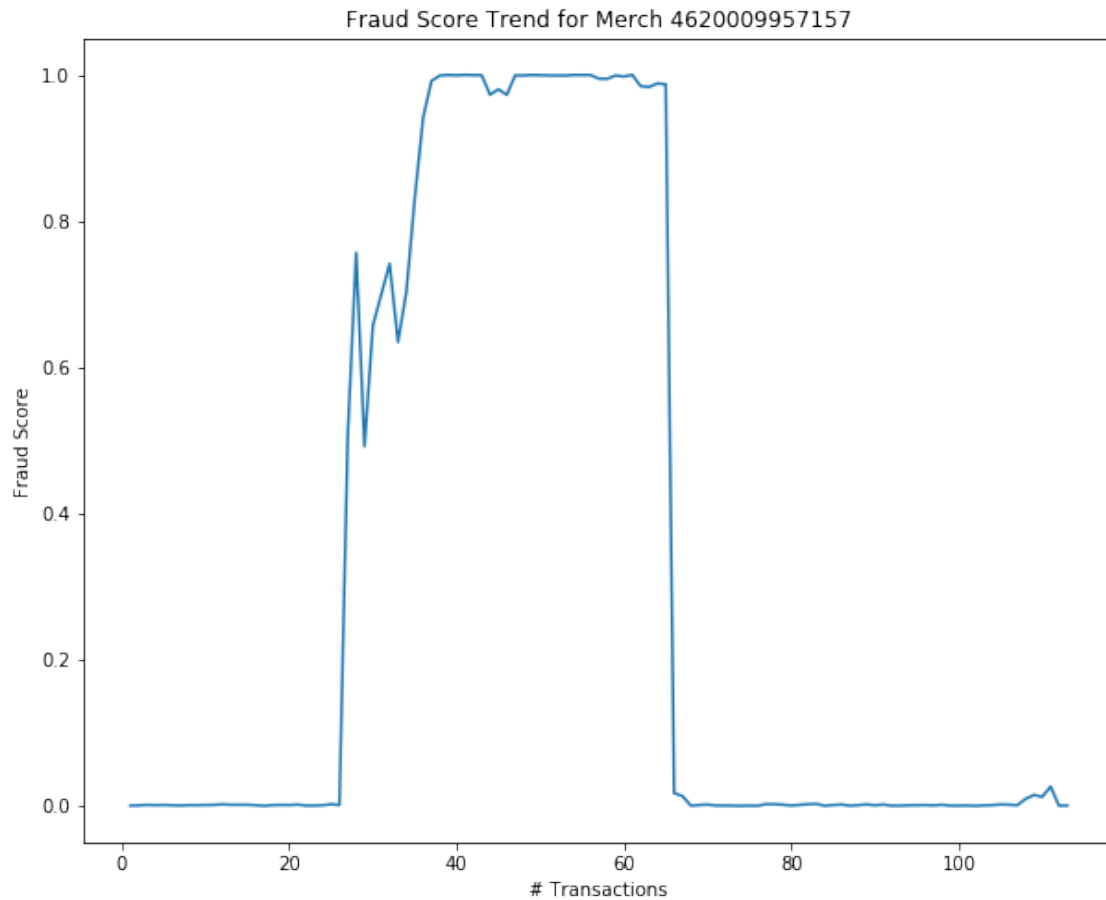[38]:  new['month']=new['Date'].dt.month
```

```
[39]:  month_card=new[new['Merchnum']=='4620009957157'].groupby('month').size()
```

```
[40]:  month_transact=new[new['Merchnum']=='4620009957157'].groupby('Merchnum').
       ↪cumcount()+1
```

```
[41]: plt.figure(figsize=(10,8))
      plt.plot(month_transact,timeplot1['predicted'])
      plt.ylabel('Fraud Score')
      plt.xlabel('# Transactions')
      plt.title('Fraud Score Trend for Merch 4620009957157')
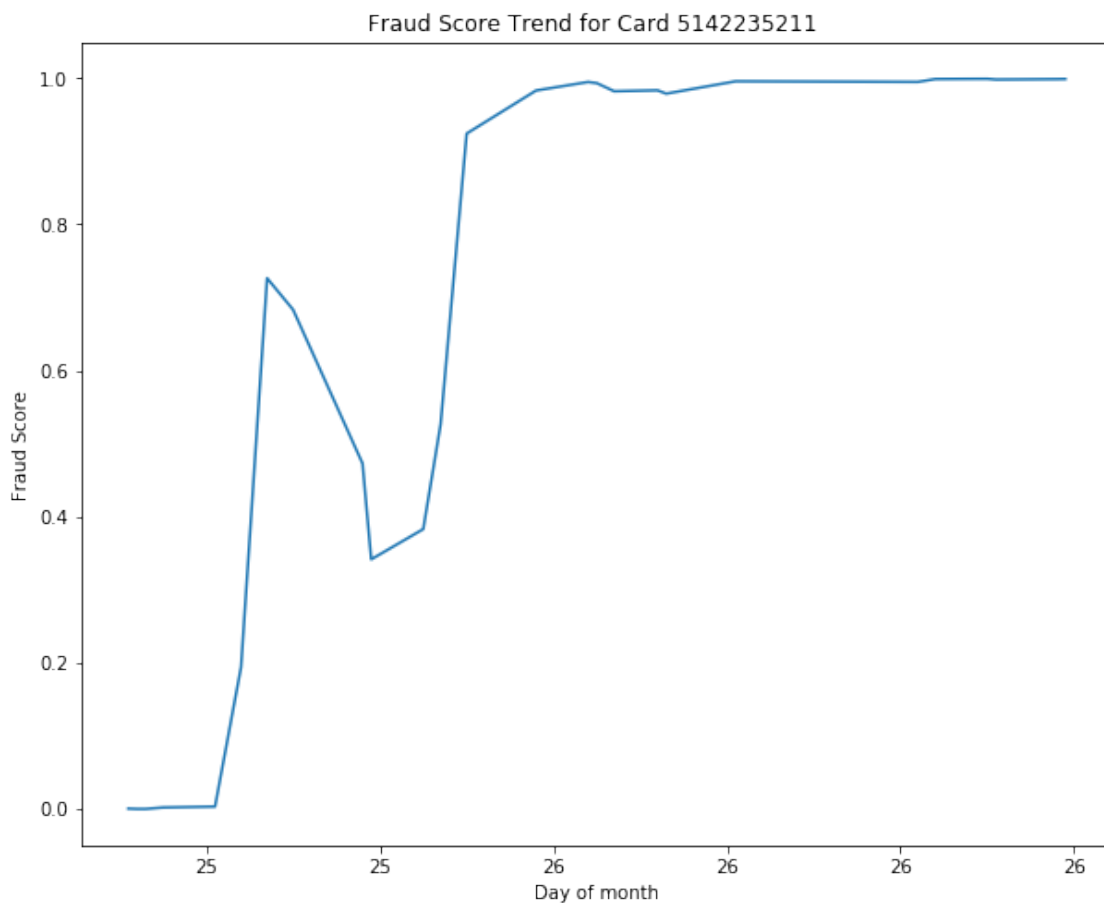      plt.show()
```


11 Card num time plots

```
[42]: import matplotlib.pyplot as plt
      timeplot2=new[(new['Cardnum']==5142235211)][['Date','predicted','Cardnum']]
```

```
[43]: timeplot2=timeplot2[timeplot2['Date'].dt.month==11][1:-5]
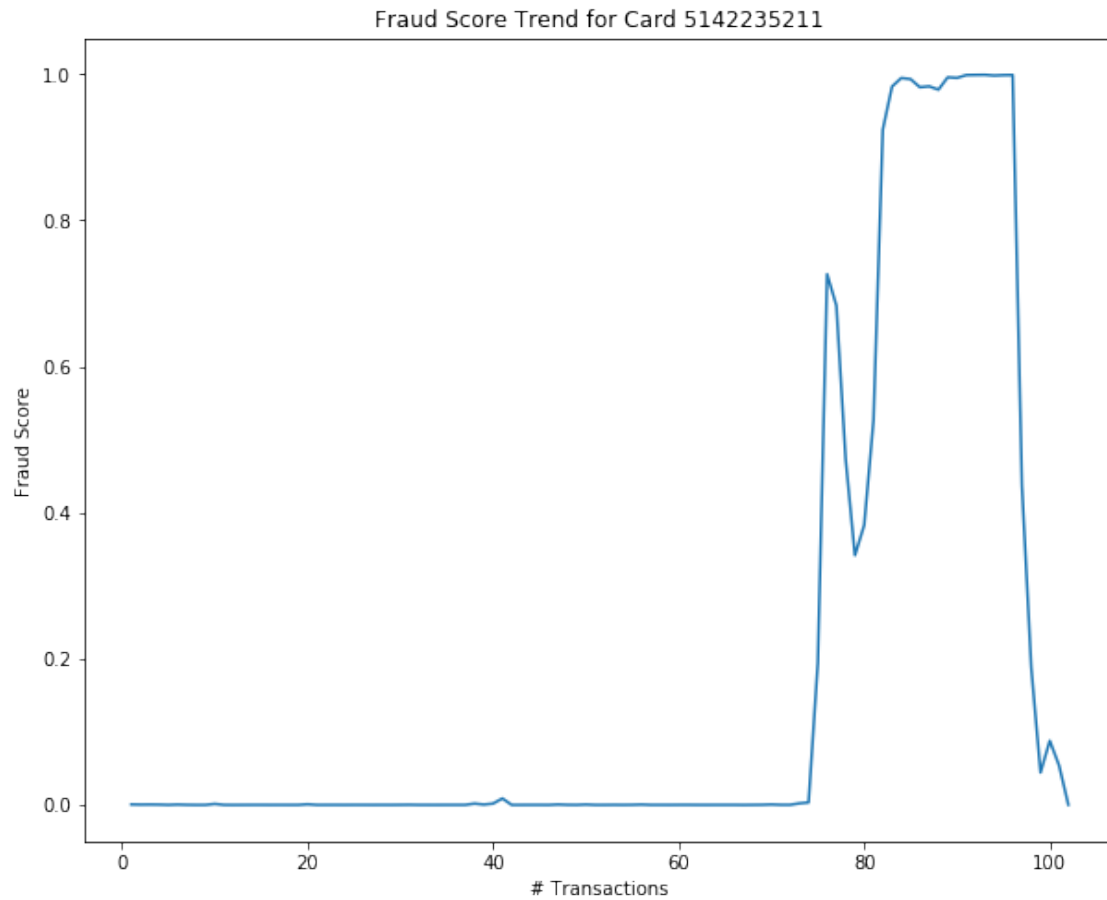```

```
[ ]:
```

```
[44]: plt.figure(figsize=(10,8))
      plt.plot(timeplot2['predicted'])
      plt.xticks([5060,5080,5100,5120,5140,5160],[25,25,26,26,26,26])
      plt.ylabel('Fraud Score')
      plt.xlabel('Day of month')
      plt.title('Fraud Score Trend for Card 5142235211 ')
      plt.show()
```



```
[45]: import matplotlib.pyplot as plt
      timeplot2=new[(new['Cardnum']==5142235211)][['Date','predicted','Cardnum']]
      month_transact2=timeplot2.groupby('Cardnum').cumcount()+1
```

```
[46]: plt.figure(figsize=(10,8))
      plt.plot(month_transact2,timeplot2['predicted'])
      plt.ylabel('Fraud Score')
      plt.xlabel('# Transactions')
      plt.title('Fraud Score Trend for Card 5142235211')
      plt.show()
```

Fraud Score Trend for Card 5142235211

[47]: `month_transact2`

[47]:
```
16216     1
16323     2
16568     3
16575     4
16928     5
17031     6
20623     7
22846     8
25337     9
25833    10
25842    11
26638    12
29356    13
34153    14
34854    15
35765    16
37144    17
```

```
37246     18
37546     19
37655     20
40783     21
43078     22
44313     23
45048     24
46175     25
46655     26
47837     27
48719     28
48947     29
49006     30
           ...
5055      73
5061      74
5064      75
5067      76
5070      77
5078      78
5079      79
5085      80
5087      81
5090      82
5098      83
5104      84
5105      85
5107      86
5112      87
5113      88
5121      89
5142      90
5144      91
5147      92
5150      93
5151      94
5158      95
5159      96
5160      97
5164      98
5204      99
5213     100
5216     101
5888     102
Length: 102, dtype: int64
```

[ ]: