

```
In [1]: import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sb
%matplotlib inline
```

```
In [2]: data = pd.read_csv("C:/Users/SHWETA/Pictures/interview documents/creditcard_csv.csv")
data.head(10)
```

Out[2]:

	Time	V1	V2	V3	V4	V5	V6	V7	V8	V9	...	V21	V22	V23	V24
0	0.0	-1.359807	-0.072781	2.536347	1.378155	-0.338321	0.462388	0.239599	0.098698	0.363787	...	-0.018307	0.277838	-0.110474	0.066928
1	0.0	1.191857	0.266151	0.166480	0.448154	0.060018	-0.082361	-0.078803	0.085102	-0.255425	...	-0.225775	-0.638672	0.101288	-0.339846
2	1.0	-1.358354	-1.340163	1.773209	0.379780	-0.503198	1.800499	0.791461	0.247676	-1.514654	...	0.247998	0.771679	0.909412	-0.689281
3	1.0	-0.966272	-0.185226	1.792993	-0.863291	-0.010309	1.247203	0.237609	0.377436	-1.387024	...	-0.108300	0.005274	-0.190321	-1.175575
4	2.0	-1.158233	0.877737	1.548718	0.403034	-0.407193	0.095921	0.592941	-0.270533	0.817739	...	-0.009431	0.798278	-0.137458	0.141267
5	2.0	-0.425966	0.960523	1.141109	-0.168252	0.420987	-0.029728	0.476201	0.260314	-0.568671	...	-0.208254	-0.559825	-0.026398	-0.371427
6	4.0	1.229658	0.141004	0.045371	1.202613	0.191881	0.272708	-0.005159	0.081213	0.464960	...	-0.167716	-0.270710	-0.154104	-0.780055
7	7.0	-0.644269	1.417964	1.074380	-0.492199	0.948934	0.428118	1.120631	-3.807864	0.615375	...	1.943465	-1.015455	0.057504	-0.649709
8	7.0	-0.894286	0.286157	-0.113192	-0.271526	2.669599	3.721818	0.370145	0.851084	-0.392048	...	-0.073425	-0.268092	-0.204233	1.011592
9	9.0	-0.338262	1.119593	1.044367	-0.222187	0.499361	-0.246761	0.651583	0.069539	-0.736727	...	-0.246914	-0.633753	-0.120794	0.385050

10 rows x 31 columns

In [3]: data.tail(10)

Out[3]:

	Time	V1	V2	V3	V4	V5	V6	V7	V8	V9	...	V21	V22	V23	
284797	172782.0	-0.241923	0.712247	0.399806	-0.463406	0.244531	-1.343668	0.929369	-0.206210	0.106234	...	-0.228876	-0.514376	0.279598	0.371
284798	172782.0	0.219529	0.881246	-0.635891	0.960928	-0.152971	-1.014307	0.427126	0.121340	-0.285670	...	0.099936	0.337120	0.251791	0.057
284799	172783.0	-1.775135	-0.004235	1.189786	0.331096	1.196063	5.519980	-1.518185	2.080825	1.159498	...	0.103302	0.654850	-0.348929	0.745
284800	172784.0	2.039560	-0.175233	-1.196825	0.234580	-0.008713	-0.726571	0.017050	-0.118228	0.435402	...	-0.268048	-0.717211	0.297930	-0.359
284801	172785.0	0.120316	0.931005	-0.546012	-0.745097	1.130314	-0.235973	0.812722	0.115093	-0.204064	...	-0.314205	-0.808520	0.050343	0.102
284802	172786.0	-11.881118	10.071785	-9.834783	-2.066656	-5.364473	-2.606837	-4.918215	7.305334	1.914428	...	0.213454	0.111864	1.014480	-0.509
284803	172787.0	-0.732789	-0.055080	2.035030	-0.738589	0.868229	1.058415	0.024330	0.294869	0.584800	...	0.214205	0.924384	0.012463	-1.016
284804	172788.0	1.919565	-0.301254	-3.249640	-0.557828	2.630515	3.031260	-0.296827	0.708417	0.432454	...	0.232045	0.578229	-0.037501	0.640
284805	172788.0	-0.240440	0.530483	0.702510	0.689799	-0.377961	0.623708	-0.686180	0.679145	0.392087	...	0.265245	0.800049	-0.163298	0.123
284806	172792.0	-0.533413	-0.189733	0.703337	-0.506271	-0.012546	-0.649617	1.577006	-0.414650	0.486180	...	0.261057	0.643078	0.376777	0.008

10 rows x 31 columns

In []:

Activate Windows
Go to Settings to activate Windows.



Type here to search



ENG
12:39 PM
13-01-2021



In [4]: `print(data.columns)`

```
Index(['Time', 'V1', 'V2', 'V3', 'V4', 'V5', 'V6', 'V7', 'V8', 'V9', 'V10',
      'V11', 'V12', 'V13', 'V14', 'V15', 'V16', 'V17', 'V18', 'V19', 'V20',
      'V21', 'V22', 'V23', 'V24', 'V25', 'V26', 'V27', 'V28', 'Amount',
      'Class'],
      dtype='object')
```

In [5]: `data.describe()`

Out[5]:

	Time	V1	V2	V3	V4	V5	V6	V7	V8	V9
count	284807.000000	2.848070e+05	2.848070e+05	2.848070e+05	2.848070e+05	2.848070e+05	2.848070e+05	2.848070e+05	2.848070e+05	2.848070e+05
mean	94813.859575	1.165980e-15	3.416908e-16	-1.373150e-15	2.086869e-15	9.604066e-16	1.490107e-15	-5.556467e-16	1.177556e-16	-2.406455e-15
std	47488.145955	1.958696e+00	1.651309e+00	1.516255e+00	1.415869e+00	1.380247e+00	1.332271e+00	1.237094e+00	1.194353e+00	1.098632e+00
min	0.000000	-5.640751e+01	-7.271573e+01	-4.832559e+01	-5.683171e+00	-1.137433e+02	-2.616051e+01	-4.355724e+01	-7.321672e+01	-1.343407e+01
25%	54201.500000	-9.203734e-01	-5.985499e-01	-8.903648e-01	-8.486401e-01	-6.915971e-01	-7.682956e-01	-5.540759e-01	-2.086297e-01	-6.430976e-01
50%	84692.000000	1.810880e-02	6.548556e-02	1.798463e-01	-1.984653e-02	-5.433583e-02	-2.741871e-01	4.010308e-02	2.235804e-02	-5.142873e-02
75%	139320.500000	1.315642e+00	8.037239e-01	1.027196e+00	7.433413e-01	6.119264e-01	3.985649e-01	5.704361e-01	3.273459e-01	5.971390e-01
max	172792.000000	2.454930e+00	2.205773e+01	9.382558e+00	1.687534e+01	3.480167e+01	7.330163e+01	1.205895e+02	2.000721e+01	1.559499e+01

8 rows x 30 columns

Activate Windows
Go to Settings to activate Windows.

Home

Untitled - Jupyter Notebook


+

localhost:8888/notebooks/Untitled.ipynb












jupyter


Untitled

Last Checkpoint: 3 minutes ago (unsaved changes)

 Logout

FileEditViewInsertCellKernelHelp

       Run    Code 




In [6]: data.info()

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 284807 entries, 0 to 284806
Data columns (total 31 columns):
#   Column  Non-Null Count  Dtype
---  -
0    Time    284807 non-null  float64
1    V1       284807 non-null  float64
2    V2       284807 non-null  float64
3    V3       284807 non-null  float64
4    V4       284807 non-null  float64
5    V5       284807 non-null  float64
6    V6       284807 non-null  float64
7    V7       284807 non-null  float64
8    V8       284807 non-null  float64
9    V9       284807 non-null  float64
10   V10      284807 non-null  float64
11   V11      284807 non-null  float64
12   V12      284807 non-null  float64
13   V13      284807 non-null  float64
14   V14      284807 non-null  float64
15   V15      284807 non-null  float64
16   V16      284807 non-null  float64
17   V17      284807 non-null  float64
18   V18      284807 non-null  float64
19   V19      284807 non-null  float64
20   V20      284807 non-null  float64
21   V21      284807 non-null  float64
22   V22      284807 non-null  float64
```

Activate Windows


Go to Settings to activate Windows.

Type here to search



12:41 PM

13-01-2021



Home

Untitled - Jupyter Notebook


+

localhost:8888/notebooks/Untitled.ipynb












jupyter

Untitled


Last Checkpoint: 5 minutes ago (unsaved changes)

 Logout

FileEditViewInsertCellKernelHelp

 Run

Code



In [8]:

```
print(data.dtypes)
```


Time	float64
V1	float64
V2	float64
V3	float64
V4	float64
V5	float64
V6	float64
V7	float64
V8	float64
V9	float64
V10	float64
V11	float64
V12	float64
V13	float64
V14	float64
V15	float64
V16	float64
V17	float64
V18	float64
V19	float64
V20	float64
V21	float64
V22	float64
V23	float64
V24	float64
V25	float64
V26	float64
V27	float64
V28	float64

Activate Windows

Go to Settings to activate Windows.


Windows logo


Type here to search



?

^






ENG

12:42 PM

13-01-2021

 14

```
In [9]: data.Class.unique()
```

```
Out[9]: array([''0'', ''1''], dtype=object)
```

```
In [10]: data.isnull().sum()
```

```
Out[10]: Time      0
         V1        0
         V2        0
         V3        0
         V4        0
         V5        0
         V6        0
         V7        0
         V8        0
         V9        0
         V10       0
         V11       0
         V12       0
         V13       0
         V14       0
         V15       0
         V16       0
         V17       0
         V18       0
         V19       0
         V20       0
         V21       0
         V22       0
         V23       0
         V24       0
```

Activate Windows
Go to Settings to activate Windows.

Home

Untitled - Jupyter Notebook


+

localhost:8888/notebooks/Untitled.ipynb











jupyter

Untitled


Last Checkpoint: 7 minutes ago (unsaved changes)

 Logout

FileEditViewInsertCellKernelHelp

 Run

Code



```
V28      0
Amount    0
Class     0
dtype: int64

In [11]: data.shape


Out[11]: (284807, 31)


In [12]: data['Amount'].value_counts()

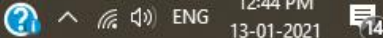
Out[12]: 1.00      13688
         1.98      6044
         0.89      4872
         9.99      4747
         15.00     3280
         ...
        192.63         1
        218.84         1
        195.52         1
        793.50         1
        1080.06         1
         Name: Amount, Length: 32767, dtype: int64

In [ ]:
```

Activate Windows
Go to Settings to activate Windows.

 Type here to search





12:44 PM
13-01-2021

Home

Untitled - Jupyter Notebook


+

localhost:8888/notebooks/Untitled.ipynb

jupyter












Untitled

Last Checkpoint: a few seconds ago (unsaved changes)

 Logout

FileEditViewInsertCellKernelHelp

TrustedPython 3

 Run

Code

In [13]:

data['Class'].value_counts()

Out[13]:

'0' 284315
'1' 492
Name: Class, dtype: int64

In [14]:


data['Time'].value_counts()


Out[14]:



163152.0 36
64947.0 26
68780.0 25
3767.0 21
3770.0 20
..
2088.0 1
64100.0 1
42068.0 1
119630.0 1
140344.0 1
Name: Time, Length: 124592, dtype: int64

In []:

Activate Windows
Go to Settings to activate Windows.

 Type here to search



 ENG 12:46 PM 13-01-2021 


```
In [15]: data.groupby('Class').agg(['min'])
```

Out[15]:

	Time	V1	V2	V3	V4	V5	V6	V7	V8	V9	...	V20	V21	V22	V23
	min	min	min	min	min	min	min	min	min	min	...	min	min	min	min
Class															
'0'	0.0	-56.40751	-72.715728	-48.325589	-5.683171	-113.743307	-26.160506	-31.764946	-73.216718	-6.290730	...	-54.497720	-34.830382	-10.933144	-8.887017
'1'	406.0	-30.55238	-8.402154	-31.103685	-1.313275	-22.105532	-6.406267	-43.557242	-41.044261	-13.434066	...	-4.128186	-22.797604	-8.887017	-8.887017

2 rows x 30 columns

```
In [16]: data.groupby('Class').agg(['max'])
```

Out[16]:

	Time	V1	V2	V3	V4	V5	V6	V7	V8	V9	...	V20	V21	V22	V23
	max	max	max	max	max	max	max	max	max	max	...	max	max	max	max
Class															
'0'	172792.0	2.454930	18.902453	9.382558	16.875344	34.801666	73.301626	120.589494	18.709255	15.594995	...	39.420904	22.614889	10.503090	22.52
'1'	170348.0	2.132386	22.057729	2.250210	12.114672	11.095089	6.474115	5.802537	20.007208	3.353525	...	11.059004	27.202839	8.361985	5.46

2 rows x 30 columns

```
In [ ]:
```

Activate Windows
Go to Settings to activate Windows.

```
In [15]: data.groupby('Class').agg(['min'])
```

Out[15]:

	Time	V1	V2	V3	V4	V5	V6	V7	V8	V9	...	V20	V21	V22	V23
	min	min	min	min	min	min	min	min	min	min	...	min	min	min	min
Class															
'0'	0.0	-56.40751	-72.715728	-48.325589	-5.683171	-113.743307	-26.160506	-31.764946	-73.216718	-6.290730	...	-54.497720	-34.830382	-10.933144	-10.933144
'1'	406.0	-30.55238	-8.402154	-31.103685	-1.313275	-22.105532	-6.406267	-43.557242	-41.044261	-13.434066	...	-4.128186	-22.797604	-8.887017	-8.887017

2 rows x 30 columns

```
In [16]: data.groupby('Class').agg(['max'])
```

Out[16]:

	Time	V1	V2	V3	V4	V5	V6	V7	V8	V9	...	V20	V21	V22	V23
	max	max	max	max	max	max	max	max	max	max	...	max	max	max	max
Class															
'0'	172792.0	2.454930	18.902453	9.382558	16.875344	34.801666	73.301626	120.589494	18.709255	15.594995	...	39.420904	22.614889	10.503090	22.52
'1'	170348.0	2.132386	22.057729	2.250210	12.114672	11.095089	6.474115	5.802537	20.007208	3.353525	...	11.059004	27.202839	8.361985	5.46

2 rows x 30 columns

```
In [17]: data.groupby('Amount').agg(['min'])
```

Activate Windows
Go to Settings to activate Windows.

```
In [17]: data.groupby('Amount').agg(['min'])
```

Out[17]:

	Time	V1	V2	V3	V4	V5	V6	V7	V8	V9	...	V20	V21	V22
	min	min	min	min	min	min	min	min	min	min	...	min	min	min
Amount														
0.00	282.0	-29.046615	-32.362448	-26.871339	-2.726250	-17.999630	-16.057255	-41.506796	-38.987263	-13.434066	...	-20.010641	-21.453736	-3.7
0.01	402.0	-28.709229	-11.380730	-27.855811	-4.507079	-18.983813	-7.858770	-43.557242	-41.044261	-13.320155	...	-2.339452	-22.797604	-3.6
0.02	1969.0	-4.514380	-2.108046	-1.824974	-3.416766	-1.986986	-1.698966	-4.460523	-7.526415	-1.689245	...	-0.874520	-3.356760	-1.7
0.03	47051.0	-2.627770	-1.470663	-1.047911	-3.140270	-0.112289	-2.412990	-1.846840	-2.017216	-1.191160	...	-0.194386	-1.222735	-0.6
0.04	25153.0	-7.608857	-10.770345	-1.134417	-2.865953	-1.115558	-6.837562	-7.096445	-0.133350	-2.567496	...	-0.253993	-0.262534	-0.7
...
11898.09	46253.0	-21.780665	-38.305310	-12.122469	9.752791	-12.880794	4.256017	14.785051	-2.818253	-0.667338	...	24.133894	7.437478	-5.6
12910.93	42951.0	-23.712839	-42.172688	-13.320825	9.925019	-13.945538	5.564891	15.710644	-2.844253	-1.580725	...	26.237391	7.921600	-6.3
18910.00	95286.0	-34.549296	-60.464618	-21.340854	16.875344	-19.229075	6.335259	24.422716	-4.964566	0.188912	...	38.117209	11.502580	-9.4
19656.53	48401.0	-36.802320	-63.344698	-20.645794	16.715537	-20.672064	7.694002	24.956587	-4.730111	-2.687312	...	39.420904	11.455313	-10.9
25691.16	166198.0	-35.548539	-31.850484	-48.325589	15.304184	-113.743307	73.301626	120.589494	-27.347360	-3.872425	...	-54.497720	-21.620120	5.7

32767 rows x 30 columns

In []:

Activate Windows
Go to Settings to activate Windows.



Type here to search



12:48 PM
13-01-2021




```
In [19]: data.groupby('Time').agg(['max'])
```

Out[19]:

	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	...	V21	V22	V23	V24
	max	max	max	max	max	max	max	max	max	max	...	max	max	max	max
Time															
0.0	1.191857	0.266151	2.536347	1.378155	0.060018	0.462388	0.239599	0.098698	0.363787	0.090794	...	-0.018307	0.277838	0.101288	0.0
1.0	-0.966272	-0.185226	1.792993	0.379780	-0.010309	1.800499	0.791461	0.377436	-1.387024	0.207643	...	0.247998	0.771679	0.909412	-0.6
2.0	-0.425966	0.960523	1.548718	0.403034	0.420987	0.095921	0.592941	0.260314	0.817739	0.753074	...	-0.009431	0.798278	-0.026398	0.1
4.0	1.229658	0.141004	0.045371	1.202613	0.191881	0.272708	-0.005159	0.081213	0.464960	-0.099254	...	-0.167716	-0.270710	-0.154104	-0.7
7.0	-0.644269	1.417964	1.074380	-0.271526	2.669599	3.721818	1.120631	0.851084	0.615375	1.249376	...	1.943465	-0.268092	0.057504	1.0
...
172785.0	0.120316	0.931005	-0.546012	-0.745097	1.130314	-0.235973	0.812722	0.115093	-0.204064	-0.657422	...	-0.314205	-0.808520	0.050343	0.1
172786.0	-11.881118	10.071785	-9.834783	-2.066656	-5.364473	-2.606837	-4.918215	7.305334	1.914428	4.356170	...	0.213454	0.111864	1.014480	-0.5
172787.0	-0.732789	-0.055080	2.035030	-0.738589	0.868229	1.058415	0.024330	0.294869	0.584800	-0.975926	...	0.214205	0.924384	0.012463	-1.0
172788.0	1.919565	0.530483	0.702510	0.689799	2.630515	3.031260	-0.296827	0.708417	0.432454	-0.399126	...	0.265245	0.800049	-0.037501	0.6
172792.0	-0.533413	-0.189733	0.703337	-0.506271	-0.012546	-0.649617	1.577006	-0.414650	0.486180	-0.915427	...	0.261057	0.643078	0.376777	0.0

124592 rows x 30 columns

Activate Windows
Go to Settings to activate Windows.

In []:



Type here to search



12:50 PM
13-01-2021




```
In [32]: filtereddata=data[['Time','Amount','Class']]
```

```
In [33]: filtereddata
```

Out[33]:

	Time	Amount	Class
0	0.0	149.62	'0'
1	0.0	2.69	'0'
2	1.0	378.66	'0'
3	1.0	123.50	'0'
4	2.0	69.99	'0'
...
284802	172786.0	0.77	'0'
284803	172787.0	24.79	'0'
284804	172788.0	67.88	'0'
284805	172788.0	10.00	'0'
284806	172792.0	217.00	'0'

284807 rows x 3 columns

```
In [24]:
```

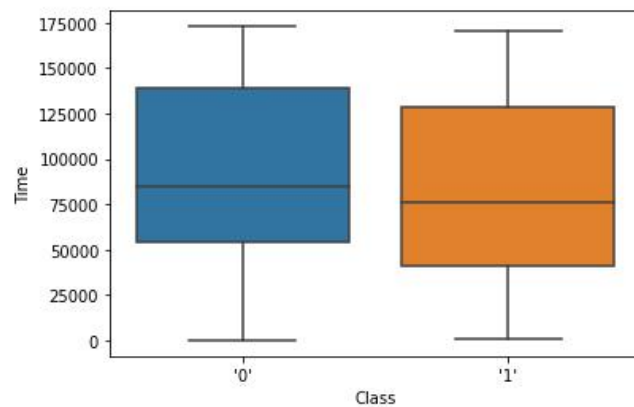
```
Out[24]: 0
```

Activate Windows
Go to Settings to activate Windows.

```
284804 172788.0 67.88 '0'
284805 172788.0 10.00 '0'
284806 172792.0 217.00 '0'
```

284807 rows x 3 columns

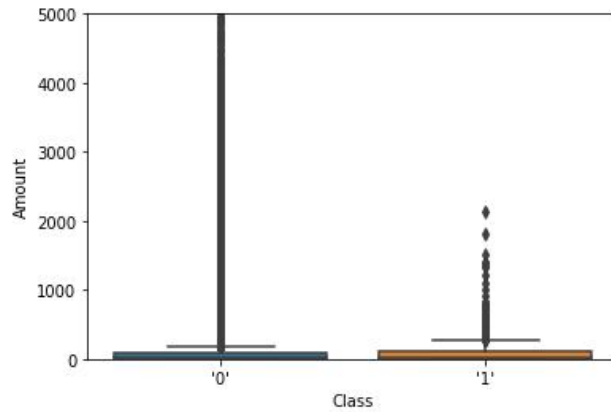
```
In [35]: sb.boxplot(x="Class", y="Time", data=data)
plt.show()
```



In []:

Activate Windows
Go to Settings to activate Windows.

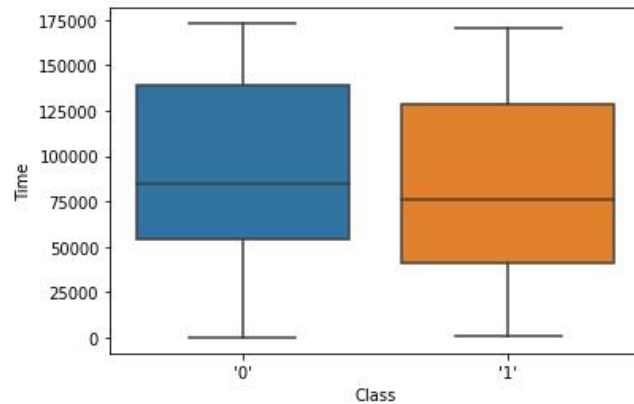
```
In [36]: sb.boxplot(x="Class", y="Amount", data=data)
plt.ylim(0, 5000)
plt.show()
```



In []: |

Activate Windows
Go to Settings to activate Windows.

```
In [35]: sb.boxplot(x="Class", y="Time", data=data)
plt.show()
```



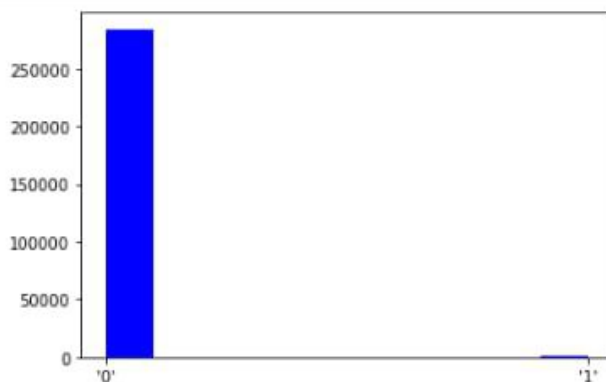
```
In [36]: sb.boxplot(x="Class", y="Amount", data=data)
plt.ylim(0, 5000)
plt.show()
```



Activate Windows
Go to Settings to activate Windows.

5 rows x 31 columns

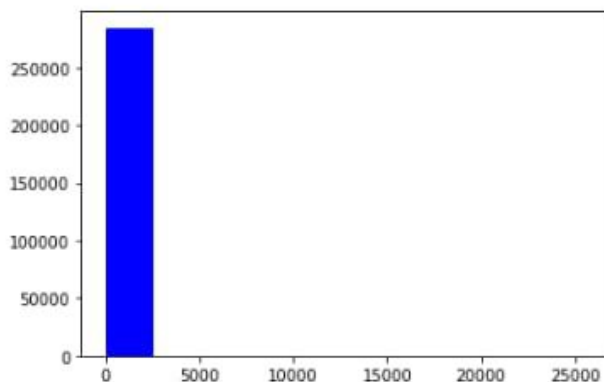
```
In [10]: #histogram
plt.hist(dd["Class"],color='blue')
plt.show()
```



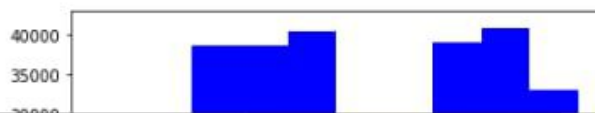
```
In [11]: plt.hist(dd["Amount"],color='green')
plt.show()
```

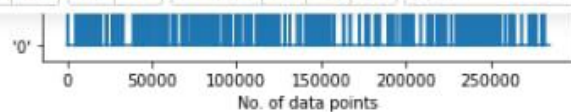


```
In [11]: plt.hist(dd["Amount"],color='green')
plt.show()
```

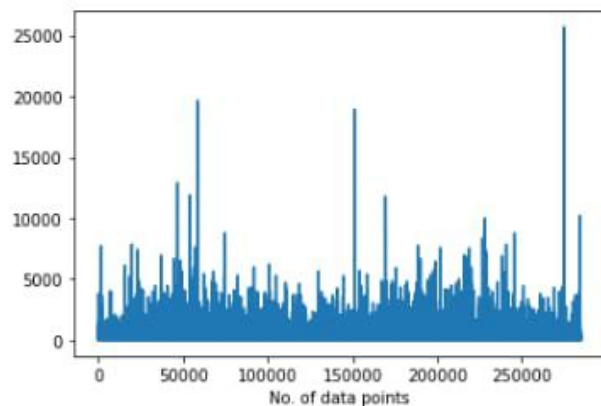


```
In [12]: plt.hist(dd["Time"],color='red')
plt.show()
```





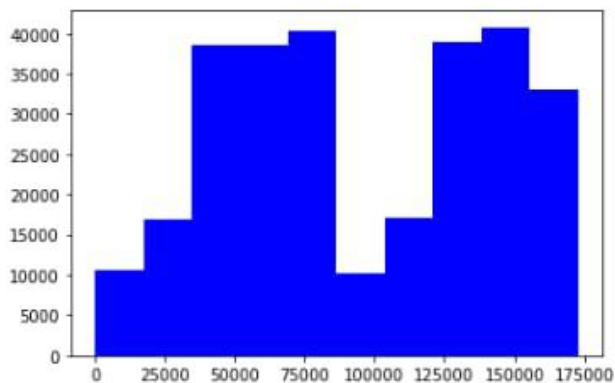
```
In [14]: plt.plot(dd["Amount"])
plt.xlabel("No. of data points")
plt.show()
```



```
In [15]: plt.plot(dd["Time"])
plt.xlabel("No. of data points")
plt.show()
```



```
In [12]: plt.hist(dd["Time"],color='red')
plt.show()
```

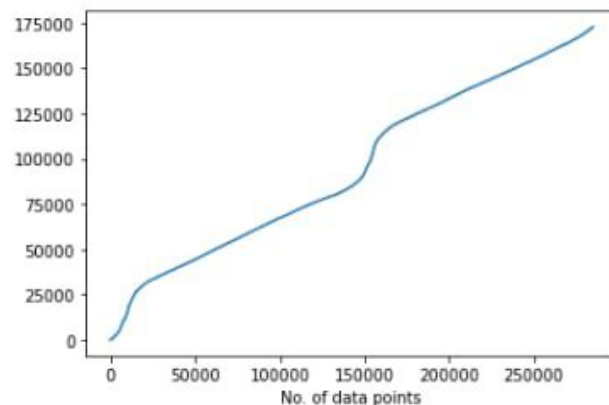


```
In [13]: plt.plot(dd["Class"])
plt.xlabel("No. of data points")
plt.show()
```

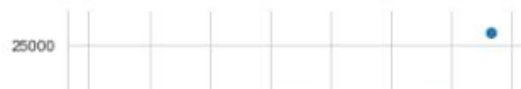


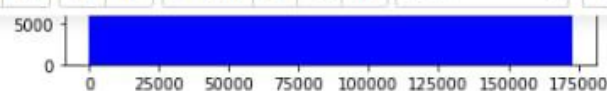
No. of data points

```
In [15]: plt.plot(dd["Time"])
plt.xlabel("No. of data points")
plt.show()
```

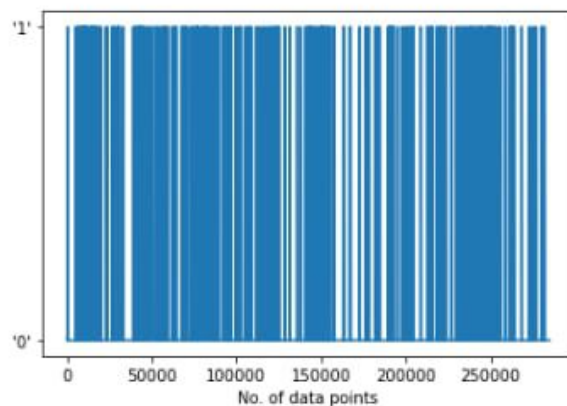


```
In [16]: sb.set_style('whitegrid');
sb.FacetGrid(dd,hue='Class',height=5).map(plt.scatter,'Time','Amount').add_legend();
plt.show();
```





```
In [13]: plt.plot(dd["Class"])
plt.xlabel("No. of data points")
plt.show()
```

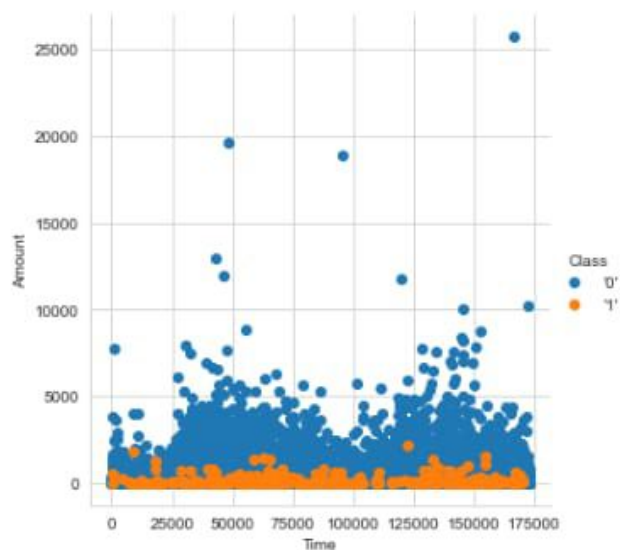


```
In [14]: plt.plot(dd["Amount"])
plt.xlabel("No. of data points")
plt.show()
```



0 50000 100000 150000 200000 250000
No. of data points

```
In [16]: sb.set_style('whitegrid');  
sb.FacetGrid(dd,hue='Class',height=5).map(plt.scatter,'Time','Amount').add_legend();  
plt.show();
```



```
In [*]: sb.boxplot(x='Time',y='Amount',data=dd)
```