ads-phase3

November 1, 2023

0.1 Date: 1/11/2023

0.2 Project Title: Credit Card Fraudlent Detection

TEAM MEMBERS:

VIGNESH S (2021506086)

SANJAY M (2021506123)

VIGNESH G (2021506121)

SISHAATH RA KRISHNA (20215060101)

SHREESH SHIVALINGAM (2021506097)

0.3 Importing required libraries

[5]:
```
import pandas as pd
import numpy as np
import seaborn as sns
from matplotlib import pyplot as plt
```

0.3.1 Set the jupyter notebook to show maximum number of columns

[6]:
```
pd.options.display.max_columns = None
```

[4]:
```
from google.colab import drive
drive.mount('/content/drive')
```

```
Mounted at /content/drive
```

0.3.2 Loading the datasets

[21]:
```
ccfd = pd.read_csv('drive/MyDrive/ColabNotebooks/creditcard.csv')
```

0.3.3 Displaying top 5 rows

[22]:
```
ccfd.head()
```

[22]:
```
   Time        V1        V2        V3        V4        V5        V6        V7  \
0   0.0 -1.359807 -0.072781  2.536347  1.378155 -0.338321  0.462388  0.239599
1   0.0  1.191857  0.266151  0.166480  0.448154  0.060018 -0.082361 -0.078803
2   1.0 -1.358354 -1.340163  1.773209  0.379780 -0.503198  1.800499  0.791461
3   1.0 -0.966272 -0.185226  1.792993 -0.863291 -0.010309  1.247203  0.237609
4   2.0 -1.158233  0.877737  1.548718  0.403034 -0.407193  0.095921  0.592941

         V8        V9       V10       V11       V12       V13       V14  \
0  0.098698  0.363787  0.090794 -0.551600 -0.617801 -0.991390 -0.311169
1  0.085102 -0.255425 -0.166974  1.612727  1.065235  0.489095 -0.143772
2  0.247676 -1.514654  0.207643  0.624501  0.066084  0.717293 -0.165946
3  0.377436 -1.387024 -0.054952 -0.226487  0.178228  0.507757 -0.287924
4 -0.270533  0.817739  0.753074 -0.822843  0.538196  1.345852 -1.119670

        V15       V16       V17       V18       V19       V20       V21  \
0  1.468177 -0.470401  0.207971  0.025791  0.403993  0.251412 -0.018307
1  0.635558  0.463917 -0.114805 -0.183361 -0.145783 -0.069083 -0.225775
2  2.345865 -2.890083  1.109969 -0.121359 -2.261857  0.524980  0.247998
3 -0.631418 -1.059647 -0.684093  1.965775 -1.232622 -0.208038 -0.108300
4  0.175121 -0.451449 -0.237033 -0.038195  0.803487  0.408542 -0.009431

        V22       V23       V24       V25       V26       V27       V28  \
0  0.277838 -0.110474  0.066928  0.128539 -0.189115  0.133558 -0.021053
1 -0.638672  0.101288 -0.339846  0.167170  0.125895 -0.008983  0.014724
2  0.771679  0.909412 -0.689281 -0.327642 -0.139097 -0.055353 -0.059752
3  0.005274 -0.190321 -1.175575  0.647376 -0.221929  0.062723  0.061458
4  0.798278 -0.137458  0.141267 -0.206010  0.502292  0.219422  0.215153

   Amount  Class
0  149.62      0
1    2.69      0
2  378.66      0
3  123.50      0
4   69.99      0
```

0.3.4 Displaying bottom 5 rows

[23]:
```python
ccfd.tail()
```

[23]:
```
            Time         V1         V2        V3        V4        V5  \
284802  172786.0 -11.881118  10.071785 -9.834783 -2.066656 -5.364473
284803  172787.0  -0.732789  -0.055080  2.035030 -0.738589  0.868229
284804  172788.0   1.919565  -0.301254 -3.249640 -0.557828  2.630515
284805  172788.0  -0.240440   0.530483  0.702510  0.689799 -0.377961
284806  172792.0  -0.533413  -0.189733  0.703337 -0.506271 -0.012546

              V6        V7        V8        V9       V10       V11       V12  \
```

```
284802 -2.606837 -4.918215  7.305334  1.914428  4.356170 -1.593105  2.711941
284803  1.058415  0.024330  0.294869  0.584800 -0.975926 -0.150189  0.915802
284804  3.031260 -0.296827  0.708417  0.432454 -0.484782  0.411614  0.063119
284805  0.623708 -0.686180  0.679145  0.392087 -0.399126 -1.933849 -0.962886
284806 -0.649617  1.577006 -0.414650  0.486180 -0.915427 -1.040458 -0.031513

             V13       V14       V15       V16       V17       V18       V19  \
284802 -0.689256  4.626942 -0.924459  1.107641  1.991691  0.510632 -0.682920
284803  1.214756 -0.675143  1.164931 -0.711757 -0.025693 -1.221179 -1.545556
284804 -0.183699 -0.510602  1.329284  0.140716  0.313502  0.395652 -0.577252
284805 -1.042082  0.449624  1.962563 -0.608577  0.509928  1.113981  2.897849
284806 -0.188093 -0.084316  0.041333 -0.302620 -0.660377  0.167430 -0.256117

             V20       V21       V22       V23       V24       V25       V26  \
284802  1.475829  0.213454  0.111864  1.014480 -0.509348  1.436807  0.250034
284803  0.059616  0.214205  0.924384  0.012463 -1.016226 -0.606624 -0.395255
284804  0.001396  0.232045  0.578229 -0.037501  0.640134  0.265745 -0.087371
284805  0.127434  0.265245  0.800049 -0.163298  0.123205 -0.569159  0.546668
284806  0.382948  0.261057  0.643078  0.376777  0.008797 -0.473649 -0.818267

             V27       V28  Amount  Class
284802  0.943651  0.823731    0.77      0
284803  0.068472 -0.053527   24.79      0
284804  0.004455 -0.026561   67.88      0
284805  0.108821  0.104533   10.00      0
284806 -0.002415  0.013649  217.00      0
```

0.3.5 Shows number of rows and columns

[24]:
```
print("Number of rows in given dataset ",ccfd.shape[0])
print("Number of columns in the given dataset ",ccfd.shape[1])
```

```
Number of rows in given dataset  284807
Number of columns in the given dataset  31
```

0.3.6 Getting basis information

[25]:
```
ccfd.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
 23  V23     284807 non-null  float64
 24  V24     284807 non-null  float64
 25  V25     284807 non-null  float64
 26  V26     284807 non-null  float64
 27  V27     284807 non-null  float64
 28  V28     284807 non-null  float64
 29  Amount  284807 non-null  float64
 30  Class   284807 non-null  int64
dtypes: float64(30), int64(1)
memory usage: 67.4 MB
```

0.3.7 Checking null values in the given data

[26]:
```
ccfd.isnull().sum()
```

[26]:
```
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
V25       0
V26       0
V27       0
V28       0
Amount    0
Class     0
dtype: int64
```

0.3.8 Scaling the Amount features, removing the independent columns

[51]:
```
#removing the column name Time, it is unnecessary to our training purposes
ccfd.head(20)
```

[51]:
```
          V1        V2        V3        V4        V5        V6        V7  \
0  -1.359807 -0.072781  2.536347  1.378155 -0.338321  0.462388  0.239599
1   1.191857  0.266151  0.166480  0.448154  0.060018 -0.082361 -0.078803
2  -1.358354 -1.340163  1.773209  0.379780 -0.503198  1.800499  0.791461
3  -0.966272 -0.185226  1.792993 -0.863291 -0.010309  1.247203  0.237609
4  -1.158233  0.877737  1.548718  0.403034 -0.407193  0.095921  0.592941
5  -0.425966  0.960523  1.141109 -0.168252  0.420987 -0.029728  0.476201
6   1.229658  0.141004  0.045371  1.202613  0.191881  0.272708 -0.005159
7  -0.644269  1.417964  1.074380 -0.492199  0.948934  0.428118  1.120631
8  -0.894286  0.286157 -0.113192 -0.271526  2.669599  3.721818  0.370145
9  -0.338262  1.119593  1.044367 -0.222187  0.499361 -0.246761  0.651583
10  1.449044 -1.176339  0.913860 -1.375667 -1.971383 -0.629152 -1.423236
11  0.384978  0.616109 -0.874300 -0.094019  2.924584  3.317027  0.470455
12  1.249999 -1.221637  0.383930 -1.234899 -1.485419 -0.753230 -0.689405
13  1.069374  0.287722  0.828613  2.712520 -0.178398  0.337544 -0.096717
14 -2.791855 -0.327771  1.641750  1.767473 -0.136588  0.807596 -0.422911
15 -0.752417  0.345485  2.057323 -1.468643 -1.158394 -0.077850 -0.608581
16  1.103215 -0.040296  1.267332  1.289091 -0.735997  0.288069 -0.586057
17 -0.436905  0.918966  0.924591 -0.727219  0.915679 -0.127867  0.707642
18 -5.401258 -5.450148  1.186305  1.736239  3.049106 -1.763406 -1.559738
19  1.492936 -1.029346  0.454795 -1.438026 -1.555434 -0.720961 -1.080664
```

	V8	V9	V10	V11	V12	V13	V14	\
0	0.098698	0.363787	0.090794	-0.551600	-0.617801	-0.991390	-0.311169	
1	0.085102	-0.255425	-0.166974	1.612727	1.065235	0.489095	-0.143772	
2	0.247676	-1.514654	0.207643	0.624501	0.066084	0.717293	-0.165946	
3	0.377436	-1.387024	-0.054952	-0.226487	0.178228	0.507757	-0.287924	
4	-0.270533	0.817739	0.753074	-0.822843	0.538196	1.345852	-1.119670	
5	0.260314	-0.568671	-0.371407	1.341262	0.359894	-0.358091	-0.137134	
6	0.081213	0.464960	-0.099254	-1.416907	-0.153826	-0.751063	0.167372	
7	-3.807864	0.615375	1.249376	-0.619468	0.291474	1.757964	-1.323865	
8	0.851084	-0.392048	-0.410430	-0.705117	-0.110452	-0.286254	0.074355	
9	0.069539	-0.736727	-0.366846	1.017614	0.836390	1.006844	-0.443523	
10	0.048456	-1.720408	1.626659	1.199644	-0.671440	-0.513947	-0.095045	
11	0.538247	-0.558895	0.309755	-0.259116	-0.326143	-0.090047	0.362832	
12	-0.227487	-2.094011	1.323729	0.227666	-0.242682	1.205417	-0.317631	
13	0.115982	-0.221083	0.460230	-0.773657	0.323387	-0.011076	-0.178485	
14	-1.907107	0.755713	1.151087	0.844555	0.792944	0.370448	-0.734975	
15	0.003603	-0.436167	0.747731	-0.793981	-0.770407	1.047627	-1.066604	
16	0.189380	0.782333	-0.267975	-0.450311	0.936708	0.708380	-0.468647	
17	0.087962	-0.665271	-0.737980	0.324098	0.277192	0.252624	-0.291896	
18	0.160842	1.233090	0.345173	0.917230	0.970117	-0.266568	-0.479130	
19	-0.053127	-1.978682	1.638076	1.077542	-0.632047	-0.416957	0.052011	

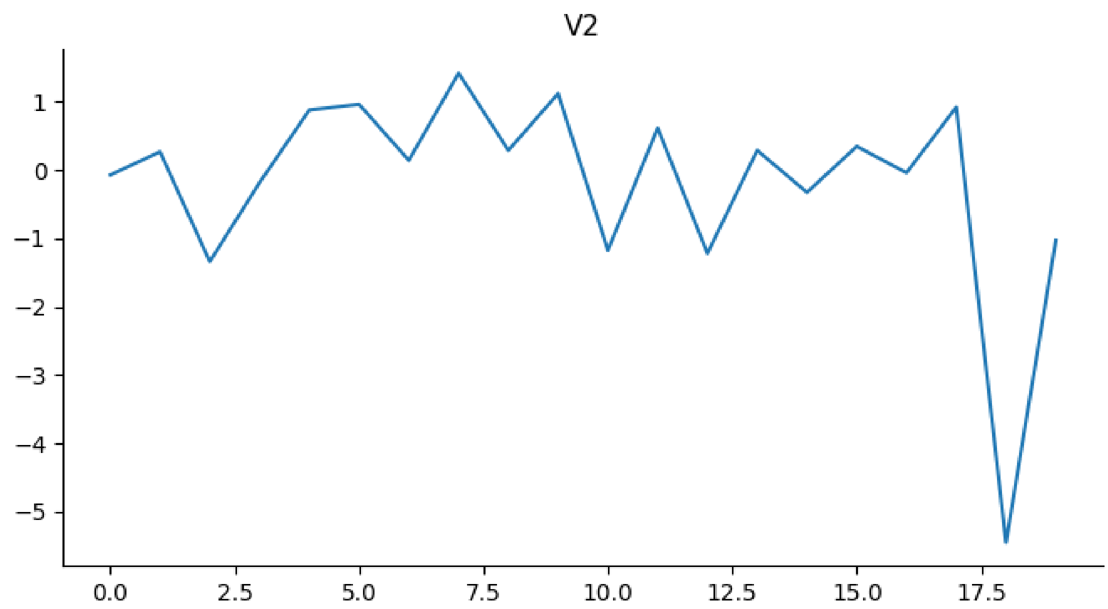
	V15	V16	V17	V18	V19	V20	V21	\
0	1.468177	-0.470401	0.207971	0.025791	0.403993	0.251412	-0.018307	
1	0.635558	0.463917	-0.114805	-0.183361	-0.145783	-0.069083	-0.225775	
2	2.345865	-2.890083	1.109969	-0.121359	-2.261857	0.524980	0.247998	
3	-0.631418	-1.059647	-0.684093	1.965775	-1.232622	-0.208038	-0.108300	
4	0.175121	-0.451449	-0.237033	-0.038195	0.803487	0.408542	-0.009431	
5	0.517617	0.401726	-0.058133	0.068653	-0.033194	0.084968	-0.208254	
6	0.050144	-0.443587	0.002821	-0.611987	-0.045575	-0.219633	-0.167716	
7	0.686133	-0.076127	-1.222127	-0.358222	0.324505	-0.156742	1.943465	
8	-0.328783	-0.210077	-0.499768	0.118765	0.570328	0.052736	-0.073425	
9	0.150219	0.739453	-0.540980	0.476677	0.451773	0.203711	-0.246914	
10	0.230930	0.031967	0.253415	0.854344	-0.221365	-0.387226	-0.009302	
11	0.928904	-0.129487	-0.809979	0.359985	0.707664	0.125992	0.049924	
12	0.725675	-0.815612	0.873936	-0.847789	-0.683193	-0.102756	-0.231809	
13	-0.655564	-0.199925	0.124005	-0.980496	-0.982916	-0.153197	-0.036876	
14	0.406796	-0.303058	-0.155869	0.778265	2.221868	-1.582122	1.151663	
15	1.106953	1.660114	-0.279265	-0.419994	0.432535	0.263451	0.499625	
16	0.354574	-0.246635	-0.009212	-0.595912	-0.575682	-0.113910	-0.024612	
17	-0.184520	1.143174	-0.928709	0.680470	0.025436	-0.047021	-0.194796	
18	-0.526609	0.472004	-0.725481	0.075081	-0.406867	-2.196848	-0.503600	
19	-0.042979	-0.166432	0.304241	0.554432	0.054230	-0.387910	-0.177650	

	V22	V23	V24	V25	V26	V27	V28	\
0	0.277838	-0.110474	0.066928	0.128539	-0.189115	0.133558	-0.021053	

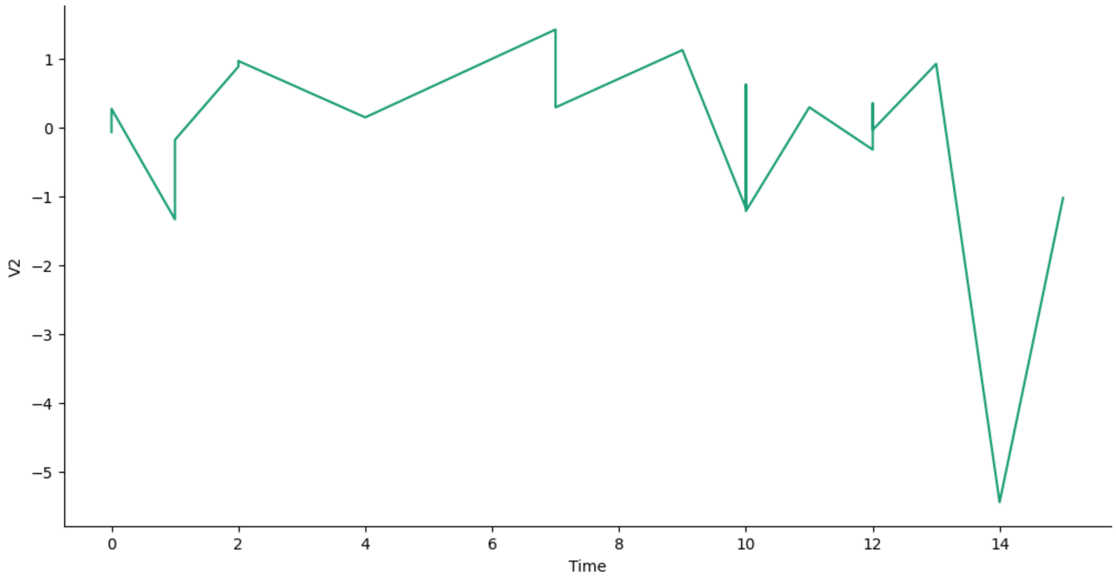
```
1  -0.638672  0.101288 -0.339846  0.167170  0.125895 -0.008983  0.014724
2   0.771679  0.909412 -0.689281 -0.327642 -0.139097 -0.055353 -0.059752
3   0.005274 -0.190321 -1.175575  0.647376 -0.221929  0.062723  0.061458
4   0.798278 -0.137458  0.141267 -0.206010  0.502292  0.219422  0.215153
5  -0.559825 -0.026398 -0.371427 -0.232794  0.105915  0.253844  0.081080
6  -0.270710 -0.154104 -0.780055  0.750137 -0.257237  0.034507  0.005168
7  -1.015455  0.057504 -0.649709 -0.415267 -0.051634 -1.206921 -1.085339
8  -0.268092 -0.204233  1.011592  0.373205 -0.384157  0.011747  0.142404
9  -0.633753 -0.120794 -0.385050 -0.069733  0.094199  0.246219  0.083076
10  0.313894  0.027740  0.500512  0.251367 -0.129478  0.042850  0.016253
11  0.238422  0.009130  0.996710 -0.767315 -0.492208  0.042472 -0.054337
12 -0.483285  0.084668  0.392831  0.161135 -0.354990  0.026416  0.042422
13  0.074412 -0.071407  0.104744  0.548265  0.104094  0.021491  0.021293
14  0.222182  1.020586  0.028317 -0.232746 -0.235557 -0.164778 -0.030154
15  1.353650 -0.256573 -0.065084 -0.039124 -0.087086 -0.180998  0.129394
16  0.196002  0.013802  0.103758  0.364298 -0.382261  0.092809  0.037051
17 -0.672638 -0.156858 -0.888386 -0.342413 -0.049027  0.079692  0.131024
18  0.984460  2.458589  0.042119 -0.481631 -0.621272  0.392053  0.949594
19 -0.175074  0.040002  0.295814  0.332931 -0.220385  0.022298  0.007602

    Class   Amounts
0       0  0.244964
1       0 -0.342475
2       0  1.160686
3       0  0.140534
4       0 -0.073403
5       0 -0.338556
6       0 -0.333279
7       0 -0.190107
8       0  0.019392
9       0 -0.338516
10      0 -0.322044
11      0 -0.313289
12      0  0.132538
13      0 -0.243282
14      0 -0.118142
15      0 -0.289300
16      0 -0.301294
17      0 -0.349671
18      0 -0.166119
19      0 -0.333239
```

[50]:

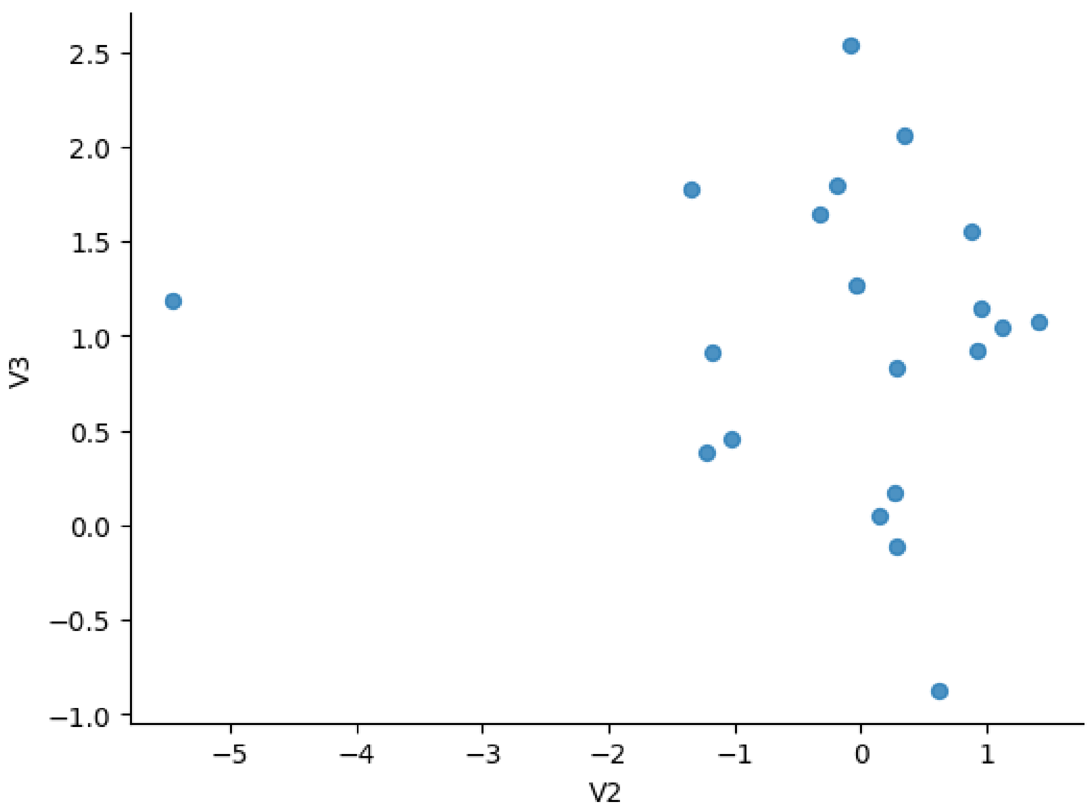
```
from matplotlib import pyplot as plt
_df_14['V2'].plot(kind='line', figsize=(8, 4), title='V2')
plt.gca().spines[['top', 'right']].set_visible(False)
```



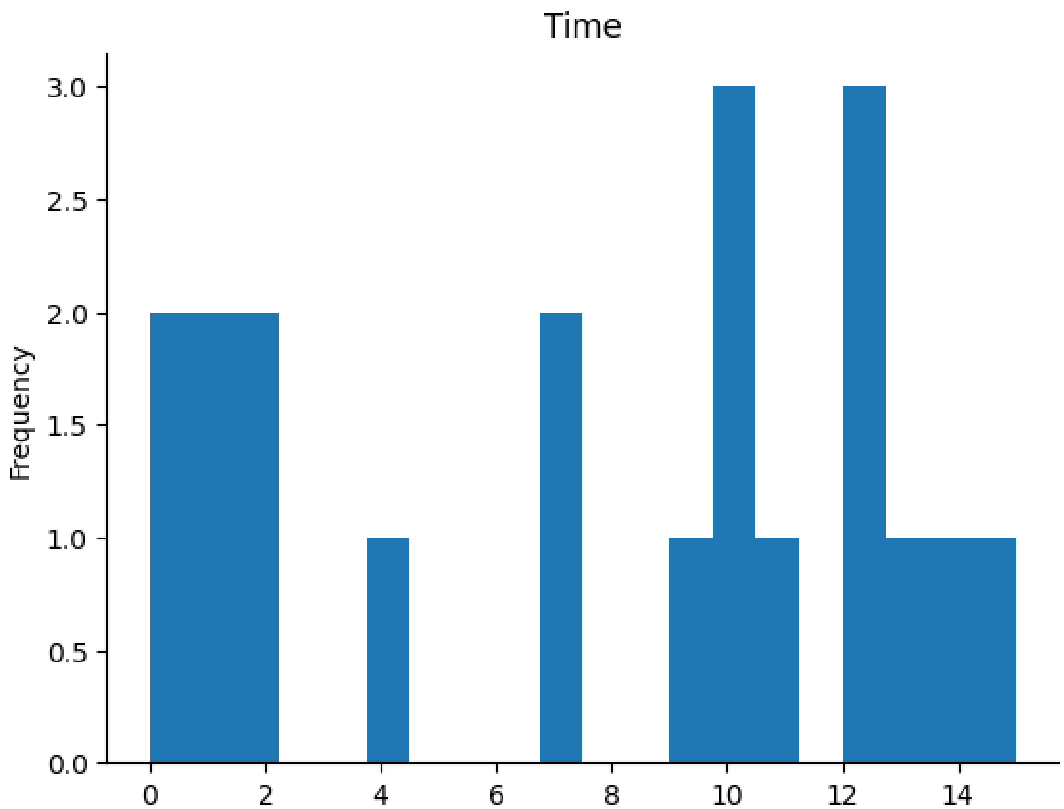
[49]:
```
from matplotlib import pyplot as plt
import seaborn as sns
def _plot_series(series, series_name, series_index=0):
  from matplotlib import pyplot as plt
  import seaborn as sns
  palette = list(sns.palettes.mpl_palette('Dark2'))
  xs = series['Time']
  ys = series['V2']

  plt.plot(xs, ys, label=series_name, color=palette[series_index %␣
↪len(palette)])

fig, ax = plt.subplots(figsize=(10, 5.2), layout='constrained')
df_sorted = _df_9.sort_values('Time', ascending=True)
_plot_series(df_sorted, '')
sns.despine(fig=fig, ax=ax)
plt.xlabel('Time')
_ = plt.ylabel('V2')
```



[48]:
```
from matplotlib import pyplot as plt
_df_6.plot(kind='scatter', x='V2', y='V3', s=32, alpha=.8)
plt.gca().spines[['top', 'right',]].set_visible(False)
```



[47]:
```
from matplotlib import pyplot as plt
_df_0['Time'].plot(kind='hist', bins=20, title='Time')
plt.gca().spines[['top', 'right',]].set_visible(False)
```



[31]:
```
#time features is unnecessary here
ccfd.drop('Time',axis=1,inplace=True)
```

[32]:
```
ccfd.head(20)
```

[32]:
```
         V1        V2        V3        V4        V5        V6        V7  \
0 -1.359807 -0.072781  2.536347  1.378155 -0.338321  0.462388  0.239599
1  1.191857  0.266151  0.166480  0.448154  0.060018 -0.082361 -0.078803
2 -1.358354 -1.340163  1.773209  0.379780 -0.503198  1.800499  0.791461
3 -0.966272 -0.185226  1.792993 -0.863291 -0.010309  1.247203  0.237609
4 -1.158233  0.877737  1.548718  0.403034 -0.407193  0.095921  0.592941
5 -0.425966  0.960523  1.141109 -0.168252  0.420987 -0.029728  0.476201
6  1.229658  0.141004  0.045371  1.202613  0.191881  0.272708 -0.005159
7 -0.644269  1.417964  1.074380 -0.492199  0.948934  0.428118  1.120631
```

8	-0.894286	0.286157	-0.113192	-0.271526	2.669599	3.721818	0.370145
9	-0.338262	1.119593	1.044367	-0.222187	0.499361	-0.246761	0.651583
10	1.449044	-1.176339	0.913860	-1.375667	-1.971383	-0.629152	-1.423236
11	0.384978	0.616109	-0.874300	-0.094019	2.924584	3.317027	0.470455
12	1.249999	-1.221637	0.383930	-1.234899	-1.485419	-0.753230	-0.689405
13	1.069374	0.287722	0.828613	2.712520	-0.178398	0.337544	-0.096717
14	-2.791855	-0.327771	1.641750	1.767473	-0.136588	0.807596	-0.422911
15	-0.752417	0.345485	2.057323	-1.468643	-1.158394	-0.077850	-0.608581
16	1.103215	-0.040296	1.267332	1.289091	-0.735997	0.288069	-0.586057
17	-0.436905	0.918966	0.924591	-0.727219	0.915679	-0.127867	0.707642
18	-5.401258	-5.450148	1.186305	1.736239	3.049106	-1.763406	-1.559738
19	1.492936	-1.029346	0.454795	-1.438026	-1.555434	-0.720961	-1.080664

	V8	V9	V10	V11	V12	V13	V14 \
0	0.098698	0.363787	0.090794	-0.551600	-0.617801	-0.991390	-0.311169
1	0.085102	-0.255425	-0.166974	1.612727	1.065235	0.489095	-0.143772
2	0.247676	-1.514654	0.207643	0.624501	0.066084	0.717293	-0.165946
3	0.377436	-1.387024	-0.054952	-0.226487	0.178228	0.507757	-0.287924
4	-0.270533	0.817739	0.753074	-0.822843	0.538196	1.345852	-1.119670
5	0.260314	-0.568671	-0.371407	1.341262	0.359894	-0.358091	-0.137134
6	0.081213	0.464960	-0.099254	-1.416907	-0.153826	-0.751063	0.167372
7	-3.807864	0.615375	1.249376	-0.619468	0.291474	1.757964	-1.323865
8	0.851084	-0.392048	-0.410430	-0.705117	-0.110452	-0.286254	0.074355
9	0.069539	-0.736727	-0.366846	1.017614	0.836390	1.006844	-0.443523
10	0.048456	-1.720408	1.626659	1.199644	-0.671440	-0.513947	-0.095045
11	0.538247	-0.558895	0.309755	-0.259116	-0.326143	-0.090047	0.362832
12	-0.227487	-2.094011	1.323729	0.227666	-0.242682	1.205417	-0.317631
13	0.115982	-0.221083	0.460230	-0.773657	0.323387	-0.011076	-0.178485
14	-1.907107	0.755713	1.151087	0.844555	0.792944	0.370448	-0.734975
15	0.003603	-0.436167	0.747731	-0.793981	-0.770407	1.047627	-1.066604
16	0.189380	0.782333	-0.267975	-0.450311	0.936708	0.708380	-0.468647
17	0.087962	-0.665271	-0.737980	0.324098	0.277192	0.252624	-0.291896
18	0.160842	1.233090	0.345173	0.917230	0.970117	-0.266568	-0.479130
19	-0.053127	-1.978682	1.638076	1.077542	-0.632047	-0.416957	0.052011

	V15	V16	V17	V18	V19	V20	V21 \
0	1.468177	-0.470401	0.207971	0.025791	0.403993	0.251412	-0.018307
1	0.635558	0.463917	-0.114805	-0.183361	-0.145783	-0.069083	-0.225775
2	2.345865	-2.890083	1.109969	-0.121359	-2.261857	0.524980	0.247998
3	-0.631418	-1.059647	-0.684093	1.965775	-1.232622	-0.208038	-0.108300
4	0.175121	-0.451449	-0.237033	-0.038195	0.803487	0.408542	-0.009431
5	0.517617	0.401726	-0.058133	0.068653	-0.033194	0.084968	-0.208254
6	0.050144	-0.443587	0.002821	-0.611987	-0.045575	-0.219633	-0.167716
7	0.686133	-0.076127	-1.222127	-0.358222	0.324505	-0.156742	1.943465
8	-0.328783	-0.210077	-0.499768	0.118765	0.570328	0.052736	-0.073425
9	0.150219	0.739453	-0.540980	0.476677	0.451773	0.203711	-0.246914
10	0.230930	0.031967	0.253415	0.854344	-0.221365	-0.387226	-0.009302

```
11  0.928904 -0.129487 -0.809979  0.359985  0.707664  0.125992  0.049924
12  0.725675 -0.815612  0.873936 -0.847789 -0.683193 -0.102756 -0.231809
13 -0.655564 -0.199925  0.124005 -0.980496 -0.982916 -0.153197 -0.036876
14  0.406796 -0.303058 -0.155869  0.778265  2.221868 -1.582122  1.151663
15  1.106953  1.660114 -0.279265 -0.419994  0.432535  0.263451  0.499625
16  0.354574 -0.246635 -0.009212 -0.595912 -0.575682 -0.113910 -0.024612
17 -0.184520  1.143174 -0.928709  0.680470  0.025436 -0.047021 -0.194796
18 -0.526609  0.472004 -0.725481  0.075081 -0.406867 -2.196848 -0.503600
19 -0.042979 -0.166432  0.304241  0.554432  0.054230 -0.387910 -0.177650

         V22       V23       V24       V25       V26       V27       V28  \
0   0.277838 -0.110474  0.066928  0.128539 -0.189115  0.133558 -0.021053
1  -0.638672  0.101288 -0.339846  0.167170  0.125895 -0.008983  0.014724
2   0.771679  0.909412 -0.689281 -0.327642 -0.139097 -0.055353 -0.059752
3   0.005274 -0.190321 -1.175575  0.647376 -0.221929  0.062723  0.061458
4   0.798278 -0.137458  0.141267 -0.206010  0.502292  0.219422  0.215153
5  -0.559825 -0.026398 -0.371427 -0.232794  0.105915  0.253844  0.081080
6  -0.270710 -0.154104 -0.780055  0.750137 -0.257237  0.034507  0.005168
7  -1.015455  0.057504 -0.649709 -0.415267 -0.051634 -1.206921 -1.085339
8  -0.268092 -0.204233  1.011592  0.373205 -0.384157  0.011747  0.142404
9  -0.633753 -0.120794 -0.385050 -0.069733  0.094199  0.246219  0.083076
10  0.313894  0.027740  0.500512  0.251367 -0.129478  0.042850  0.016253
11  0.238422  0.009130  0.996710 -0.767315 -0.492208  0.042472 -0.054337
12 -0.483285  0.084668  0.392831  0.161135 -0.354990  0.026416  0.042422
13  0.074412 -0.071407  0.104744  0.548265  0.104094  0.021491  0.021293
14  0.222182  1.020586  0.028317 -0.232746 -0.235557 -0.164778 -0.030154
15  1.353650 -0.256573 -0.065084 -0.039124 -0.087086 -0.180998  0.129394
16  0.196002  0.013802  0.103758  0.364298 -0.382261  0.092809  0.037051
17 -0.672638 -0.156858 -0.888386 -0.342413 -0.049027  0.079692  0.131024
18  0.984460  2.458589  0.042119 -0.481631 -0.621272  0.392053  0.949594
19 -0.175074  0.040002  0.295814  0.332931 -0.220385  0.022298  0.007602

    Amount  Class
0   149.62      0
1     2.69      0
2   378.66      0
3   123.50      0
4    69.99      0
5     3.67      0
6     4.99      0
7    40.80      0
8    93.20      0
9     3.68      0
10    7.80      0
11    9.99      0
12  121.50      0
13   27.50      0
```

```
14    58.80       0
15    15.99       0
16    12.99       0
17     0.89       0
18    46.80       0
19     5.00       0
```

0.3.9 Scaling the Amount column data

[33]:
```
from sklearn.preprocessing import StandardScaler
```

[34]:
```
ss = StandardScaler()
```

[35]:
```
ccfd['Amounts'] = ss.fit_transform(pd.DataFrame(ccfd['Amount']))
```

[36]:
```
ccfd.head()
```

[36]:
```
         V1        V2        V3        V4        V5        V6        V7  \
0 -1.359807 -0.072781  2.536347  1.378155 -0.338321  0.462388  0.239599
1  1.191857  0.266151  0.166480  0.448154  0.060018 -0.082361 -0.078803
2 -1.358354 -1.340163  1.773209  0.379780 -0.503198  1.800499  0.791461
3 -0.966272 -0.185226  1.792993 -0.863291 -0.010309  1.247203  0.237609
4 -1.158233  0.877737  1.548718  0.403034 -0.407193  0.095921  0.592941

         V8        V9       V10       V11       V12       V13       V14  \
0  0.098698  0.363787  0.090794 -0.551600 -0.617801 -0.991390 -0.311169
1  0.085102 -0.255425 -0.166974  1.612727  1.065235  0.489095 -0.143772
2  0.247676 -1.514654  0.207643  0.624501  0.066084  0.717293 -0.165946
3  0.377436 -1.387024 -0.054952 -0.226487  0.178228  0.507757 -0.287924
4 -0.270533  0.817739  0.753074 -0.822843  0.538196  1.345852 -1.119670

        V15       V16       V17       V18       V19       V20       V21  \
0  1.468177 -0.470401  0.207971  0.025791  0.403993  0.251412 -0.018307
1  0.635558  0.463917 -0.114805 -0.183361 -0.145783 -0.069083 -0.225775
2  2.345865 -2.890083  1.109969 -0.121359 -2.261857  0.524980  0.247998
3 -0.631418 -1.059647 -0.684093  1.965775 -1.232622 -0.208038 -0.108300
4  0.175121 -0.451449 -0.237033 -0.038195  0.803487  0.408542 -0.009431

        V22       V23       V24       V25       V26       V27       V28  \
0  0.277838 -0.110474  0.066928  0.128539 -0.189115  0.133558 -0.021053
1 -0.638672  0.101288 -0.339846  0.167170  0.125895 -0.008983  0.014724
2  0.771679  0.909412 -0.689281 -0.327642 -0.139097 -0.055353 -0.059752
3  0.005274 -0.190321 -1.175575  0.647376 -0.221929  0.062723  0.061458
4  0.798278 -0.137458  0.141267 -0.206010  0.502292  0.219422  0.215153

   Amount  Class   Amounts
0  149.62      0  0.244964
```

```
1     2.69      0 -0.342475
2   378.66      0  1.160686
3   123.50      0  0.140534
4    69.99      0 -0.073403
```

[37]:
```
ccfd.shape
```

[37]:
```
(284807, 31)
```

[38]:
```
ccfd.drop('Amount',axis=1,inplace=True)
```

[39]:
```
ccfd.shape
```

[39]:
```
(284807, 30)
```

0.3.10 Dropping the duplicate records

[40]:
```
ccfd.duplicated().any()
```

[40]:
```
True
```

[41]:
```
ccfd.drop_duplicates(inplace=True)
```

[42]:
```
ccfd.shape
```

[42]:
```
(275663, 30)
```

0.3.11 Exploring Class columns

[43]:
```
ccfd['Class'].unique()
```

[43]:
```
array([0, 1])
```

[44]:
```
ccfd['Class'].nunique()
```

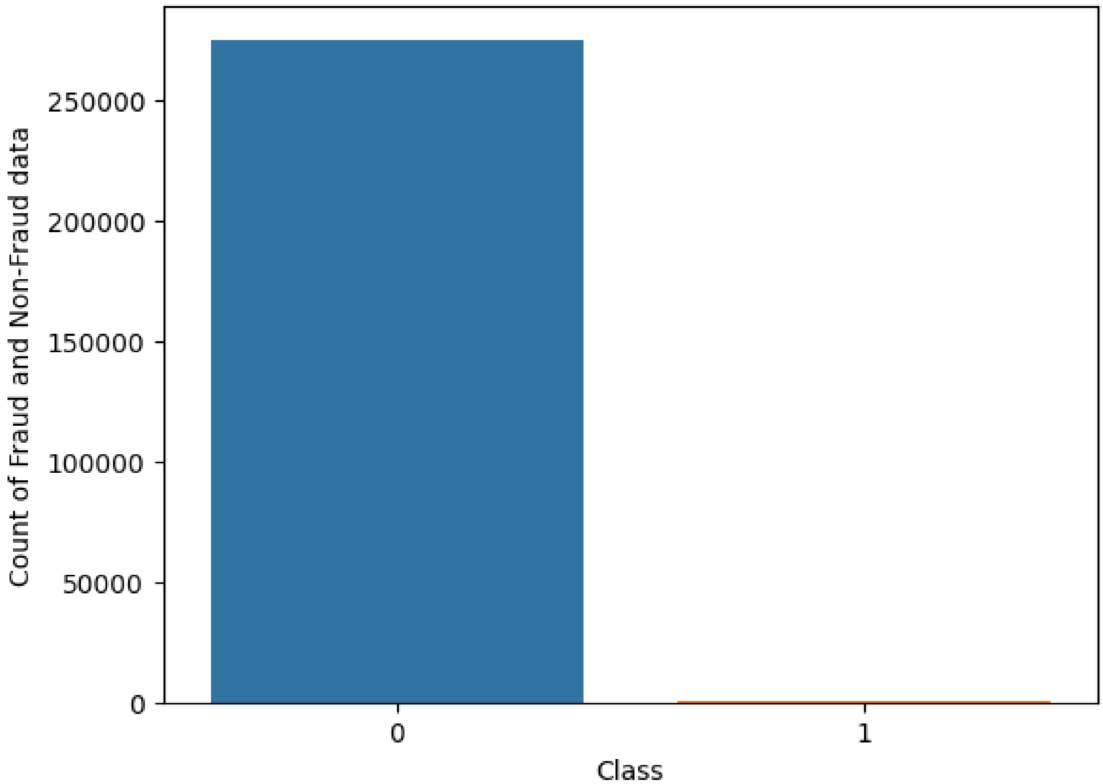
[44]:
```
2
```

[45]:
```
ccfd['Class'].value_counts()
```

[45]:
```
0    275190
1       473
Name: Class, dtype: int64
```

[46]:
```
#visualizing the distribution of 0 and 1 using seaborn countplot
sns.countplot(ccfd,x = ccfd['Class'])
plt.xlabel('Class')
```

```
plt.ylabel('Count of Fraud and Non-Fraud data')
plt.show()
```



From the above information, We can say that our data is high imbalanced, so need to apply oversampling and undersampling technique to train our model