TP3 Credit Card #Finance (partie 1)

Détection de fraude de carte bancaire

Analyse des données / Fouille des données

Mise en œuvre des algorithmes d'apprentissage automatique supervisé et non supervisé en respectant toutes les étapes de construction d'un modèle

In [1]:

```
# Import des librairies
import numpy as np # librairie de calcul numérique
import pandas as pd # librairie de statistiques
from pandas.plotting import scatter_matrix
import matplotlib.pyplot as plt # librairie de tracé de figures
import seaborn as sns
import matplotlib.patches as mpatches
```

In [2]:

```
df = pd.read_csv('creditcard.csv') # chargement de la base de données
df_init = df.copy() # sauvegarde du dataframe chargé
```

In [3]:

```
#Nombre de lignes, nombre de colonnes
df.shape
```

Out[3]:

```
(284807, 31)
```

Analyse univariable : Class

In [4]:

```
#Nombre de transactions bancaires par classe
df.groupby('Class').size()
```

Out[4]:

```
Class
0    284315
1       492
dtype: int64
```

In [5]:

```
#Répartition des transactions bancaires par classe
df['Class'].value_counts() / len(df) *100
```

Out[5]:

```
0    99.827251
1     0.172749
Name: Class, dtype: float64
```

Analyse univariable : Time

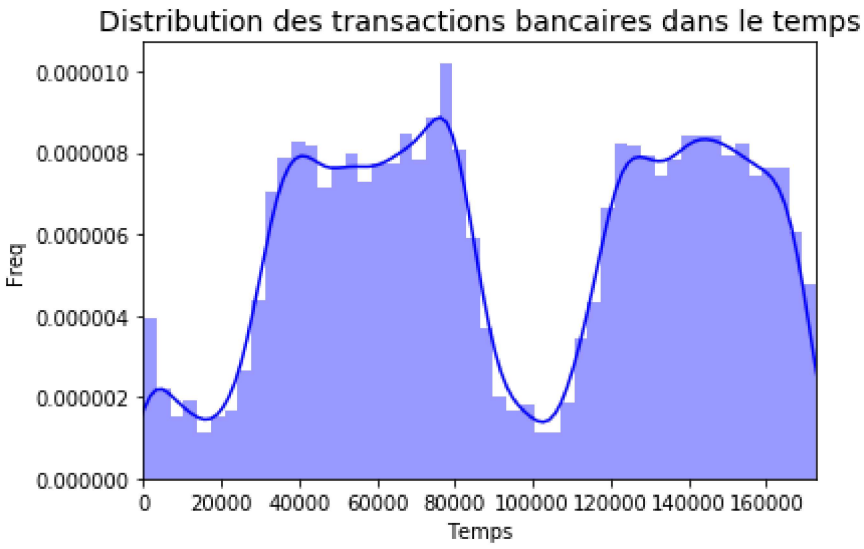
Distribution de toutes les transactions bancaires dans le temps

In [6]:

```
plt.figure()
time_val = df['Time'].values
sns.distplot(time_val, color='b')
plt.title('Distribution des transactions bancaires dans le temps', fontsize=14)
plt.xlim([min(time_val), max(time_val)])
plt.xlabel("Temps")
plt.ylabel("Freq")
```

Out[6]:

```
Text(0, 0.5, 'Freq')
```

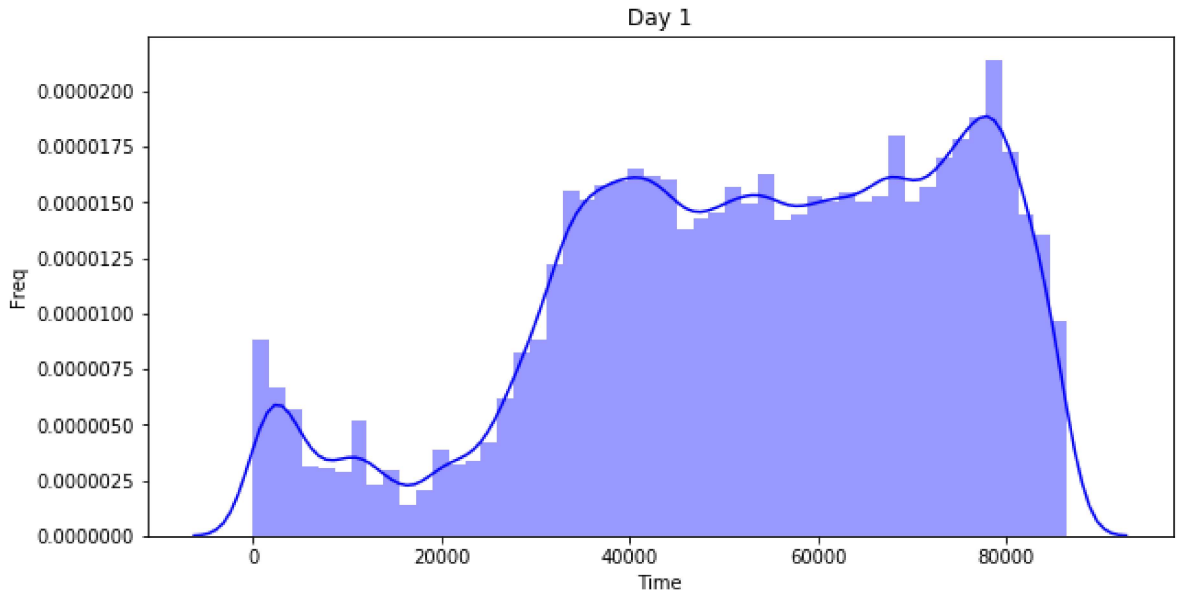


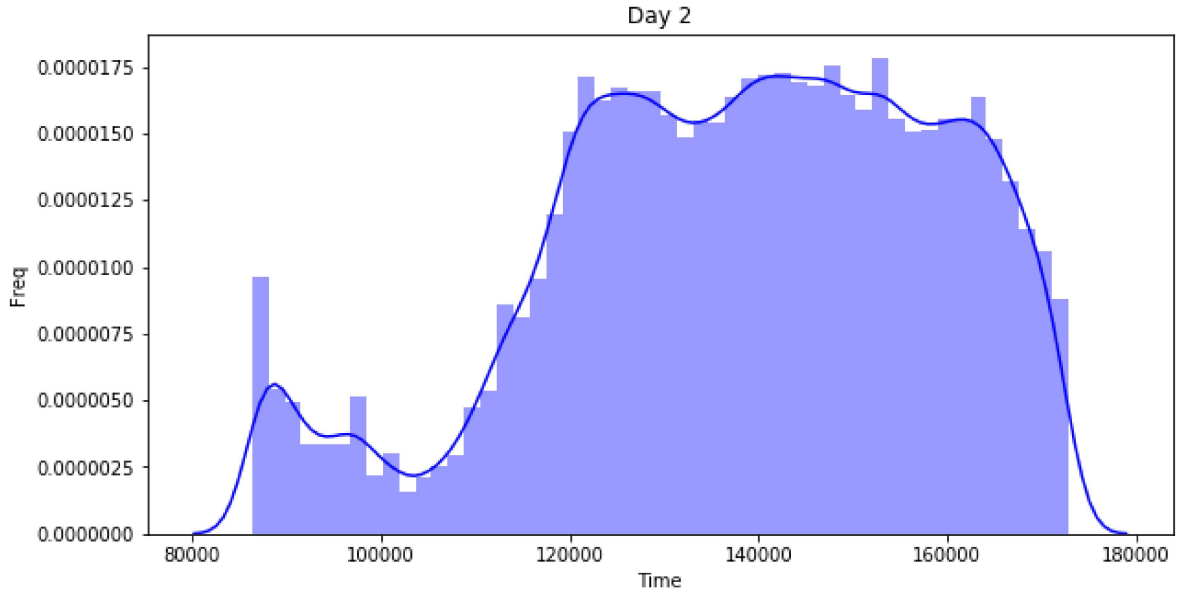
In [7]:

```
second_day = 60*60*24
index_day1 = np.where(df['Time'] < second_day)[0]
index_day2 = np.where(df['Time'] >= second_day)[0]

plt.figure(figsize=(10,5))
sns.distplot(df['Time'][index_day1], color='b')
plt.xlabel("Time")
plt.ylabel("Freq")
plt.title("Day 1")

plt.figure(figsize=(10,5))
sns.distplot(df['Time'][index_day2], color='b')
plt.xlabel("Time")
plt.ylabel("Freq")
plt.title("Day 2")
```

Out[7]:

```
Text(0.5, 1.0, 'Day 2')
```





Est-ce que les transactions bancaires frauduleuses sont toujours effectuées à la même heure ?

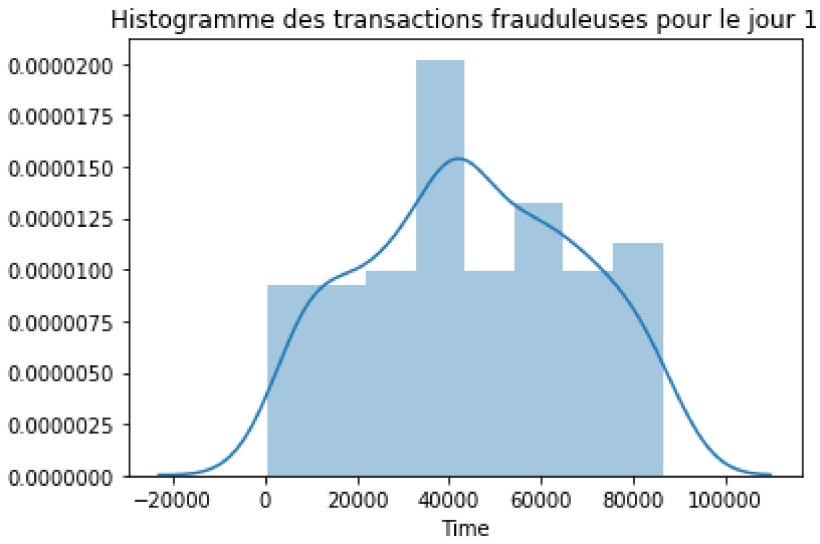
Distribution des transactions bancaires frauduleuses dans le temps

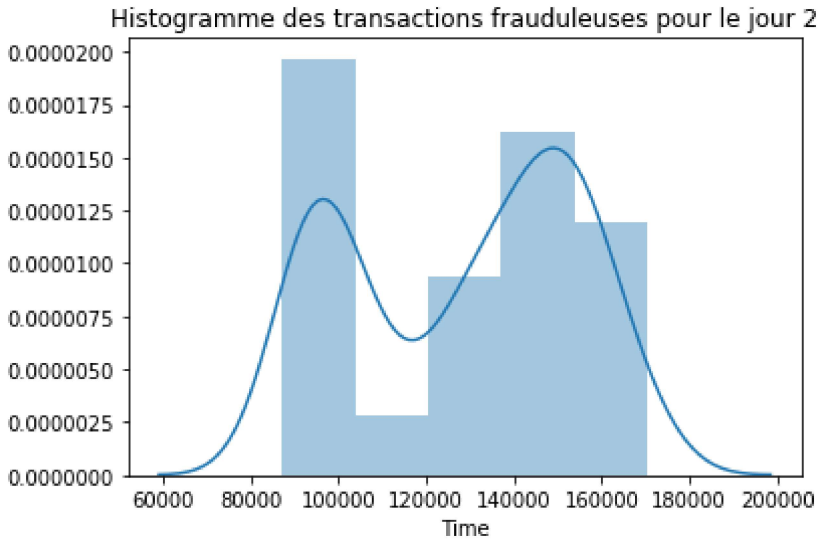
In [8]:

```
#Distribution des temps des transactions frauduleuses
x = df["Time"][df["Class"]==1]

plt.figure()
sns.distplot(x [x < second_day])
plt.title("Histogramme des transactions frauduleuses pour le jour 1")

plt.figure()
sns.distplot(x [x >= second_day])
plt.title("Histogramme des transactions frauduleuses pour le jour 2")
```

Out[8]:

```
Text(0.5, 1.0, 'Histogramme des transactions frauduleuses pour le jour 2')
```





Analyse uni-variable : Amount

In [9]:

```
df['Amount'].describe()
#np.median(df['Amount']) est le 50% du df.describe()
```

Out[9]:

```
count    284807.000000
mean         88.349619
std         250.120109
min           0.000000
25%           5.600000
50%          22.000000
75%          77.165000
max       25691.160000
Name: Amount, dtype: float64
```

Avez-vous remarqué que certaines transactions bancaires avaient un Amount à 0 ?

In [10]:

```
#Les transactions bancaires où Amount est à 0
zeroamount = df[ df['Amount']==0 ]
zeroamount.groupby("Class").size()
#Il y a 1825 transactions bancaires où Amount est à 0
```

Out[10]:

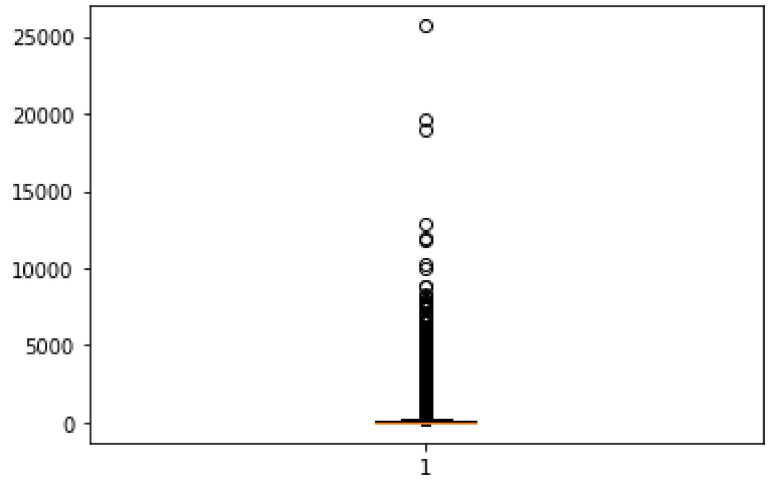
```
Class
0    1798
1      27
dtype: int64
```

In [11]:

```
plt.boxplot(df["Amount"])
```

Out[11]:

```
{'boxes': [<matplotlib.lines.Line2D at 0x171d1e398d0>],
 'caps': [<matplotlib.lines.Line2D at 0x171d1e48400>,
  <matplotlib.lines.Line2D at 0x171d1e48748>],
 'fliers': [<matplotlib.lines.Line2D at 0x171d1e48dd8>],
 'means': [],
 'medians': [<matplotlib.lines.Line2D at 0x171d1e48a90>],
 'whiskers': [<matplotlib.lines.Line2D at 0x171d1e39cf8>,
  <matplotlib.lines.Line2D at 0x171d1e39e10>]}
```



In [12]:

```
np.median(df['Amount'])
```

Out[12]:

```
22.0
```

In [13]:

```
np.median(df['Amount'][df['Class']==1])
```

Out[13]:

```
9.25
```

In [14]:

```
df.groupby('Class').describe()["Amount"]
```

Out[14]:

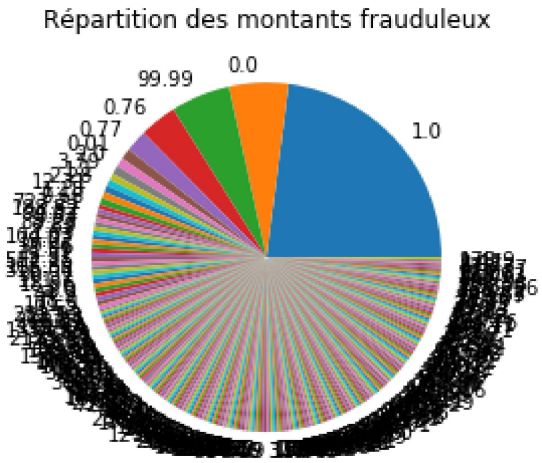
Class	count	mean	std	min	25%	50%	75%	max
0	284315.0	88.291022	250.105092	0.0	5.65	22.00	77.05	25691.16
1	492.0	122.211321	256.683288	0.0	1.00	9.25	105.89	2125.87

In [15]:

```
#fréquence des montants de transactions bancaires frauduleuses
amount_fraud = df['Amount'][df['Class']==1]
freq_amount_fraud = amount_fraud.value_counts()

#figure
plt.pie(freq_amount_fraud, labels=freq_amount_fraud.index)
plt.title('Répartition des montants frauduleux')
```

Out[15]:

```
Text(0.5, 1.0, 'Répartition des montants frauduleux')
```



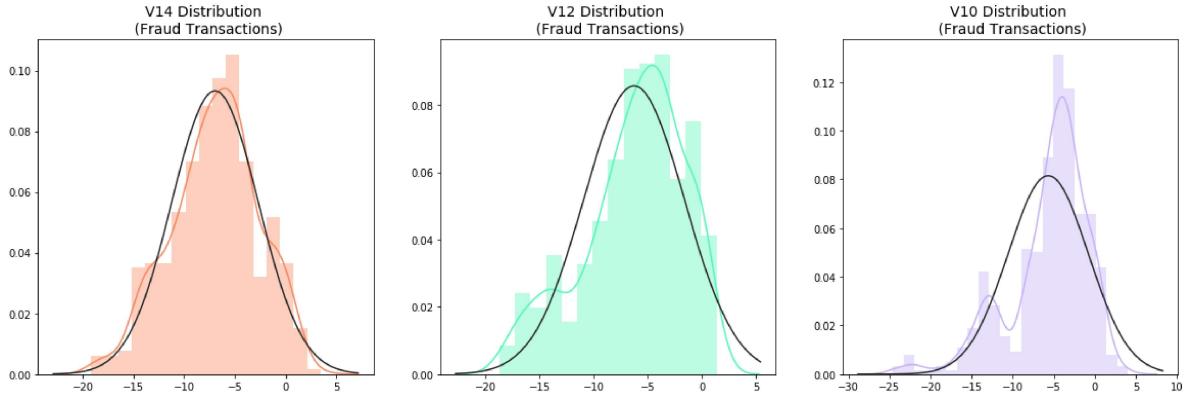
Variables quantitatives

In [16]:

```
from scipy.stats import norm

f, (ax1, ax2, ax3) = plt.subplots(1,3, figsize=(20, 6))

v10_fraud_dist = df['V10'].loc[df['Class'] == 1].values
sns.distplot(v10_fraud_dist,ax=ax3, fit=norm, color='#C5B3F9')
ax3.set_title('V10 Distribution \n (Fraud Transactions)', fontsize=14)

v12_fraud_dist = df['V12'].loc[df['Class'] == 1].values
sns.distplot(v12_fraud_dist,ax=ax2, fit=norm, color='#56F9BB')
ax2.set_title('V12 Distribution \n (Fraud Transactions)', fontsize=14)

v14_fraud_dist = df['V14'].loc[df['Class'] == 1].values
sns.distplot(v14_fraud_dist,ax=ax1, fit=norm, color='#FB8861')
ax1.set_title('V14 Distribution \n (Fraud Transactions)', fontsize=14)
```

Out[16]:

```
Text(0.5, 1.0, 'V14 Distribution \n (Fraud Transactions)')
```



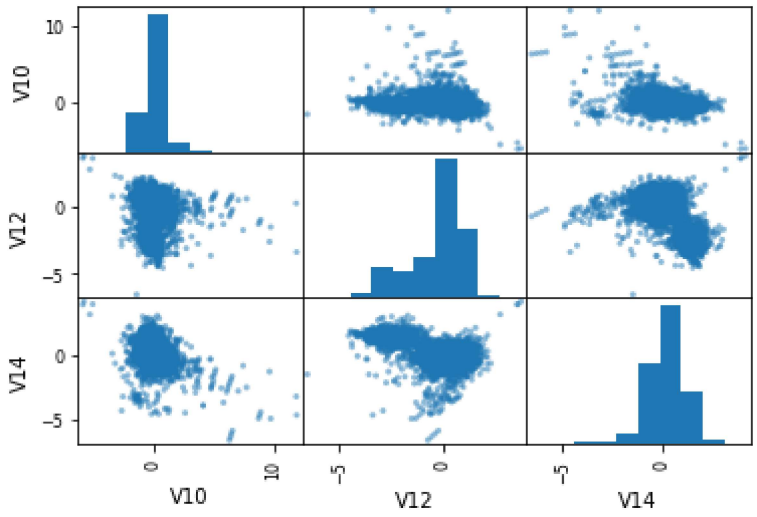
Analyse croisée des variables 2 à 2

In [17]:

```
#uniquement 3 variables et les 5000 premières lignes
sub_df = df[df.columns[[10,12,14]]][1:5000]

scatter_matrix(sub_df)
```

Out[17]:

```
array([[<matplotlib.axes._subplots.AxesSubplot object at 0x00000171D21CF470
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D22ED978
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D238DF60
>],
       [<matplotlib.axes._subplots.AxesSubplot object at 0x00000171D23C9518
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D23F9AC8
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D24370B8
>],
       [<matplotlib.axes._subplots.AxesSubplot object at 0x00000171D2466668
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D2497C50
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D2497C88
>]],
      dtype=object)
```

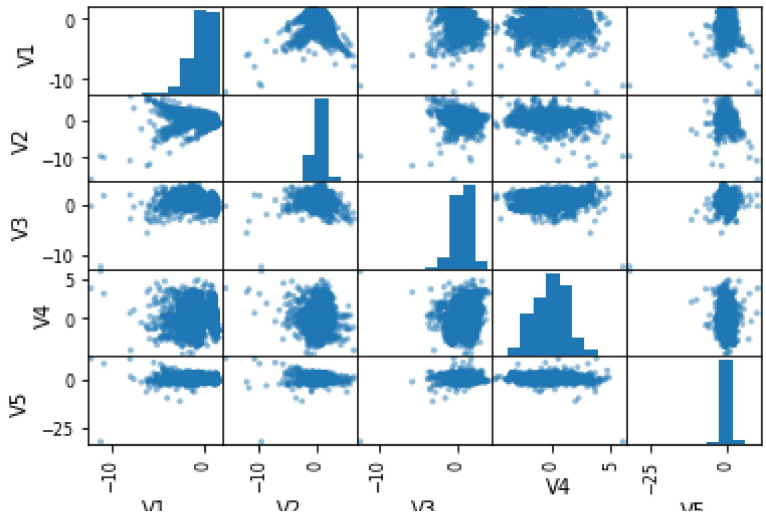


In [18]:

```
#uniquement V1 à V6 et les 1000 premières lignes
sub_df = df[df.columns[1:6]][1:5000]

scatter_matrix(sub_df)
```

Out[18]:

```
array([[<matplotlib.axes._subplots.AxesSubplot object at 0x00000171D35793C8
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D35E4240
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D360D780
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D363ECF8
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D367C2E8
>],
       [<matplotlib.axes._subplots.AxesSubplot object at 0x00000171D36AC898
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D36DDE48
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D391B470
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D391B4A8
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D3980F98
>],
       [<matplotlib.axes._subplots.AxesSubplot object at 0x00000171D39BB588
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D39EAB38
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D3A2D128
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D3A5C6D8
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D3A8FC88
>],
       [<matplotlib.axes._subplots.AxesSubplot object at 0x00000171D3ACB278
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D3AFD828
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D3B0CC50
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D3B4A240
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D3B7C7F0
>],
       [<matplotlib.axes._subplots.AxesSubplot object at 0x00000171D3BADDA0
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D3BEC390
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D3C1D940
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D3C4DEF0
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D3C8D4E0
>]],
      dtype=object)
```


V1
V2
V3
V4
V5
0
-10
0
-10
0
-10
5
0
0
-25
-10
0
-10
0
-10
0
0
5
-25
0
V4

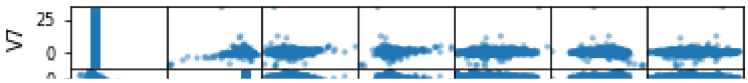
In [19]:

```
#uniquement V7 à V13 et les 1000 premières lignes
sub_df = df[df.columns[7:14]][1:5000]

scatter_matrix(sub_df)
```

Out[19]:

```
array([[<matplotlib.axes._subplots.AxesSubplot object at 0x00000171D3D1F668
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D3EB8278
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D3EDD630
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D3F0EBA8
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D3F4A198
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D3F7B748
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D3FACCF8
>],
       [<matplotlib.axes._subplots.AxesSubplot object at 0x00000171D3FE9320
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D3FE9358
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D404DE48
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D408A438
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D40BB9E8
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D40EFF98
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D4128588
>],
       [<matplotlib.axes._subplots.AxesSubplot object at 0x00000171D415BB38
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D419B128
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D41CA6D8
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D41FDC88
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D423A278
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D426D828
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D429EDD8
>],
       [<matplotlib.axes._subplots.AxesSubplot object at 0x00000171D42DB3C8
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D430D978
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D4341F28
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D437F518
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D43AFAC8
>,
```

```
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D43EE0B8
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D441D668
>],
       [<matplotlib.axes._subplots.AxesSubplot object at 0x00000171D444EC18
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D448F208
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D44BE7B8
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D44F1D68
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D4524EF0
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D45614E0
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D4590A90
>],
       [<matplotlib.axes._subplots.AxesSubplot object at 0x00000171D45D4080
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D4603630
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D4632BE0
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D46721D0
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D46A2780
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D46D4D30
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D4712320
>],
       [<matplotlib.axes._subplots.AxesSubplot object at 0x00000171D47448D0
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D4776E80
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D47B2470
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D47E5A20
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D4818FD0
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D48545C0
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D4886B70
>]],
      dtype=object)
```


V7
25
0

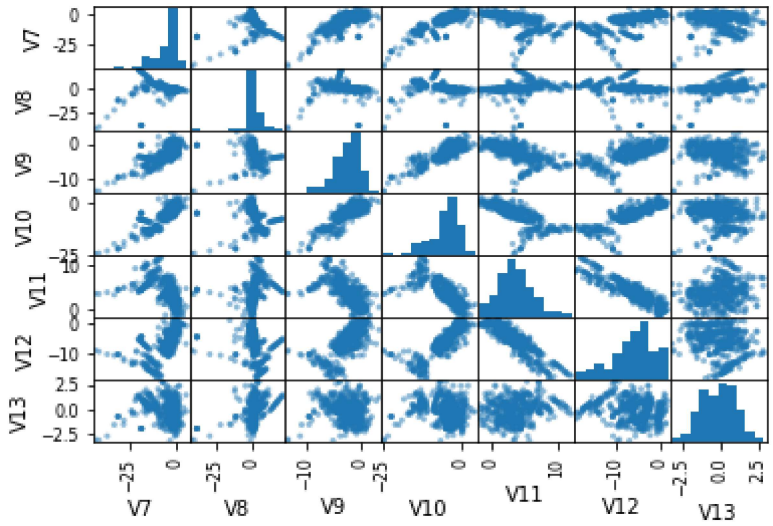
In [20]:

```
#Uniquement les transactions bancaires frauduleuses
scatter_matrix(df[df.columns[7:14]].loc[df["Class"]==1])
```

Out[20]:

```
array([[<matplotlib.axes._subplots.AxesSubplot object at 0x00000171D5B6E780
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D5B507B8
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D5BAB128
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D5BD53C8
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D5C04978
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D5C38F28
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D5C72518
>],
       [<matplotlib.axes._subplots.AxesSubplot object at 0x00000171D5CA3B00
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D5CA3B38
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D5D12668
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D5D45C18
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D5D83208
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D5DB37B8
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D5DE6D68
>],
       [<matplotlib.axes._subplots.AxesSubplot object at 0x00000171D5E23358
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D5E54908
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D5E87EB8
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D5EC6128
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D5EF46D8
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D5F28C88
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D5F68278
>],
       [<matplotlib.axes._subplots.AxesSubplot object at 0x00000171D5F99828
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D5FC9DD8
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D60073C8
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D6039978
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D606CF28
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D60A7518
>,
```

```
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D60DAAC8
>],
       [<matplotlib.axes._subplots.AxesSubplot object at 0x00000171D61180B8
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D6149668
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D617AC18
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D61B8208
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D61EA7B8
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D621BD68
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D625B358
>],
       [<matplotlib.axes._subplots.AxesSubplot object at 0x00000171D628B908
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D62BDEB8
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D62FD4A8
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D632DA58
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D636C048
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D639A5F8
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D63CDBA8
>],
       [<matplotlib.axes._subplots.AxesSubplot object at 0x00000171D640C198
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D643A748
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D646FCF8
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D64AB2E8
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D64DC898
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D650FE48
>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x00000171D654E438
>]],
      dtype=object)
```



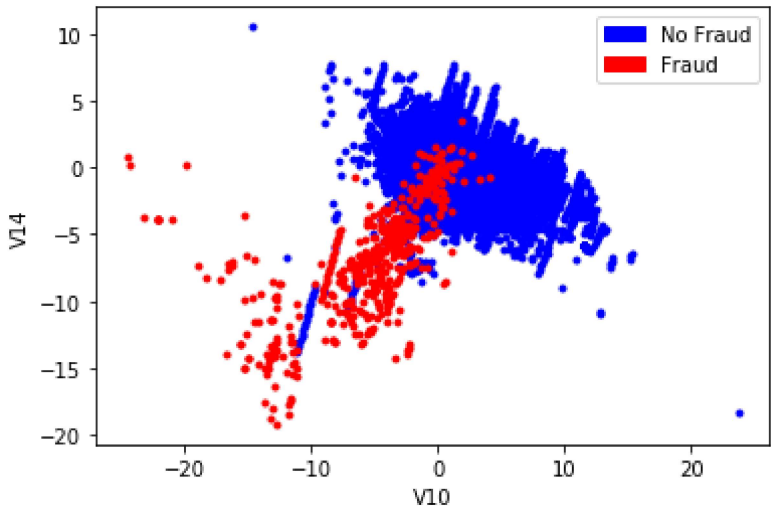
In [21]:

```
plt.plot(df[df.columns[10]].loc[df['Class'] == 0], df[df.columns[14]].loc[df['Class'] == 0]
plt.plot(df[df.columns[10]].loc[df['Class'] == 1], df[df.columns[14]].loc[df['Class'] == 1]
plt.xlabel("V10")
plt.ylabel("V14")
blue_patch = mpatches.Patch(color='blue', label='No Fraud')
red_patch = mpatches.Patch(color='red', label='Fraud')
plt.legend(handles=[blue_patch, red_patch])
```

Out[21]:

```
<matplotlib.legend.Legend at 0x171d9862b70>
```

```
d:\python\envs\myenvcnn\lib\site-packages\IPython\core\events.py:74: UserWar
ning: Creating legend with loc="best" can be slow with large amounts of dat
a.
  func(*args, **kwargs)
d:\python\envs\myenvcnn\lib\site-packages\IPython\core\pylabtools.py:125: Us
erWarning: Creating legend with loc="best" can be slow with large amounts of
data.
  fig.canvas.print_figure(bytes_io, **kw)
```



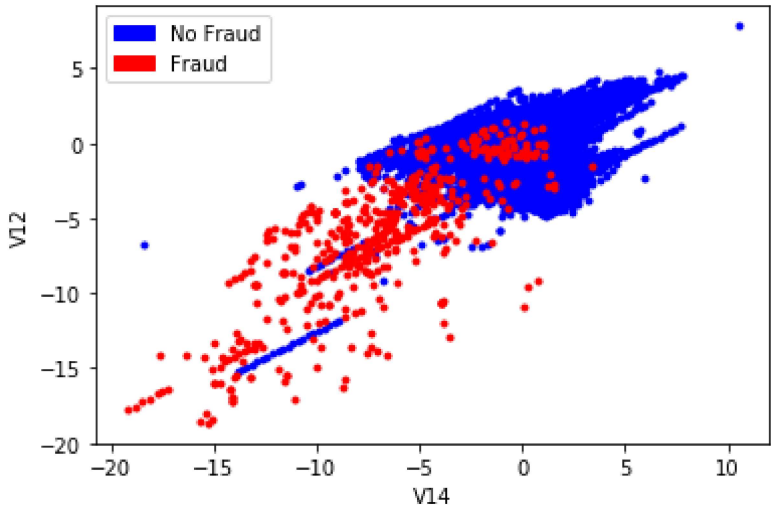
In [22]:

```
plt.plot(df[df.columns[14]].loc[df['Class'] == 0], df[df.columns[12]].loc[df['Class'] == 0]
plt.plot(df[df.columns[14]].loc[df['Class'] == 1], df[df.columns[12]].loc[df['Class'] == 1]
plt.xlabel("V14")
plt.ylabel("V12")
blue_patch = mpatches.Patch(color='blue', label='No Fraud')
red_patch = mpatches.Patch(color='red', label='Fraud')
plt.legend(handles=[blue_patch, red_patch])
```

Out[22]:

```
<matplotlib.legend.Legend at 0x171d223cf28>
```

```
d:\python\envs\myenvcnn\lib\site-packages\IPython\core\events.py:74: UserWar
ning: Creating legend with loc="best" can be slow with large amounts of dat
a.
  func(*args, **kwargs)
d:\python\envs\myenvcnn\lib\site-packages\IPython\core\pylabtools.py:125: Us
erWarning: Creating legend with loc="best" can be slow with large amounts of
data.
  fig.canvas.print_figure(bytes_io, **kw)
```



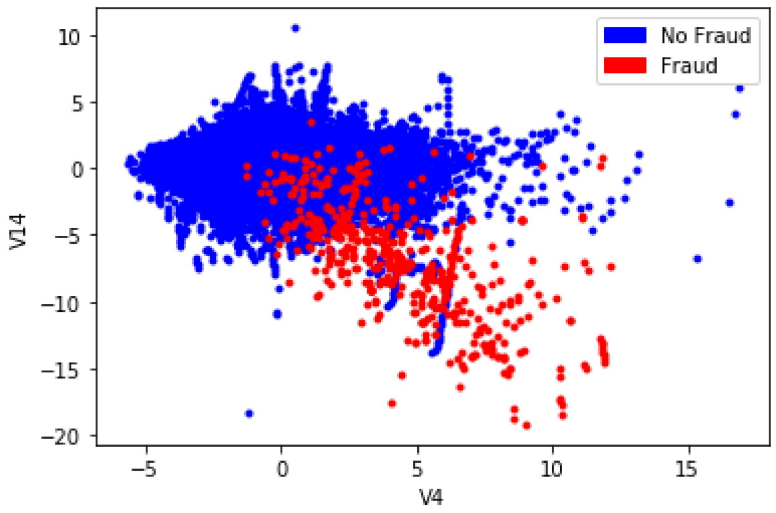
In [23]:

```
plt.plot(df[df.columns[4]].loc[df['Class'] == 0], df[df.columns[14]].loc[df['Class'] == 0],
plt.plot(df[df.columns[4]].loc[df['Class'] == 1], df[df.columns[14]].loc[df['Class'] == 1],
plt.xlabel("V4")
plt.ylabel("V14")
blue_patch = mpatches.Patch(color='blue', label='No Fraud')
red_patch = mpatches.Patch(color='red', label='Fraud')
plt.legend(handles=[blue_patch, red_patch])
```

Out[23]:

```
<matplotlib.legend.Legend at 0x171d6651710>
```

```
d:\python\envs\myenvcnn\lib\site-packages\IPython\core\events.py:74: UserWar
ning: Creating legend with loc="best" can be slow with large amounts of dat
a.
  func(*args, **kwargs)
d:\python\envs\myenvcnn\lib\site-packages\IPython\core\pylabtools.py:125: Us
erWarning: Creating legend with loc="best" can be slow with large amounts of
data.
  fig.canvas.print_figure(bytes_io, **kw)
```



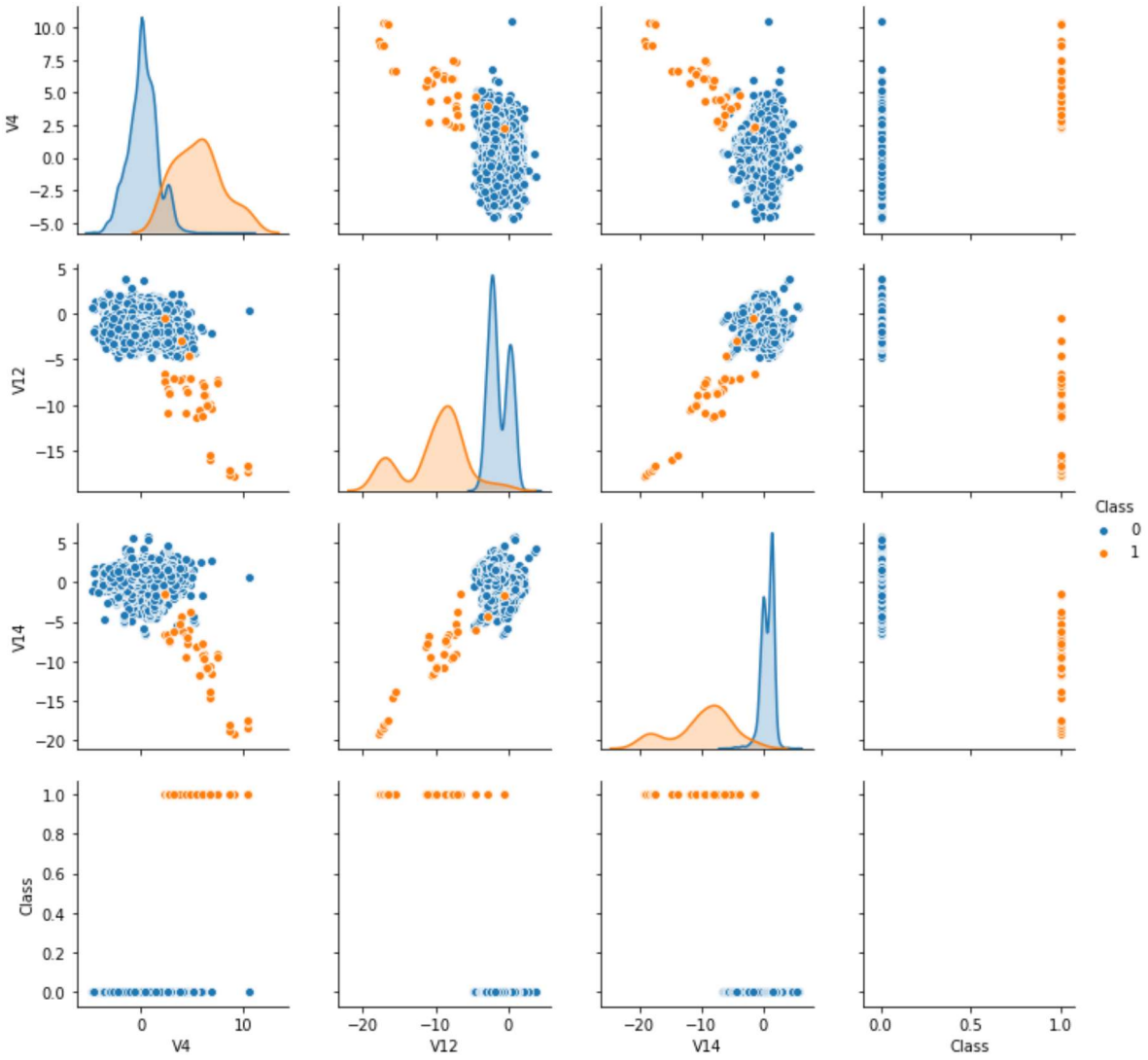
In [24]:

```
sns.pairplot(df.loc[0:10000,['V4', 'V12','V14','Class']], hue="Class")
```

```
d:\python\envs\myenvcnn\lib\site-packages\statsmodels\nonparametric\kde.py:4
87: RuntimeWarning: invalid value encountered in true_divide
  binned = fast_linbin(X, a, b, gridsize) / (delta * nobs)
d:\python\envs\myenvcnn\lib\site-packages\statsmodels\nonparametric\kdetool
s.py:34: RuntimeWarning: invalid value encountered in double_scalars
  FAC1 = 2*(np.pi*bw/RANGE)**2
```

Out[24]:

```
<seaborn.axisgrid.PairGrid at 0x171d98adba8>
```

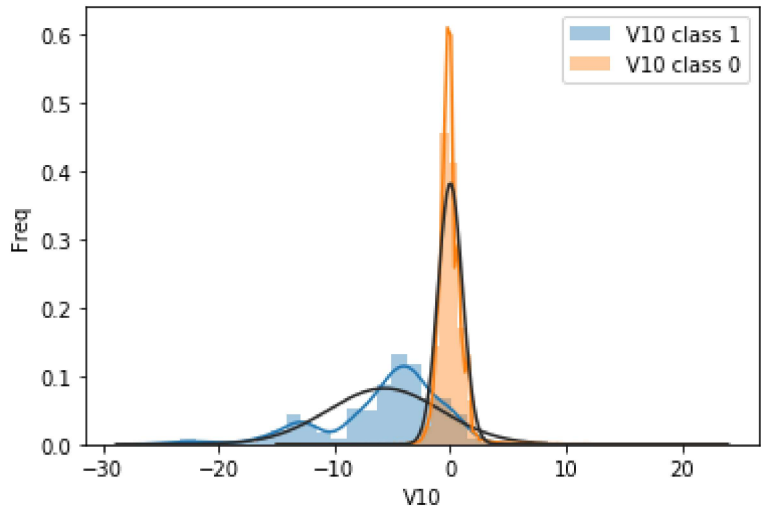


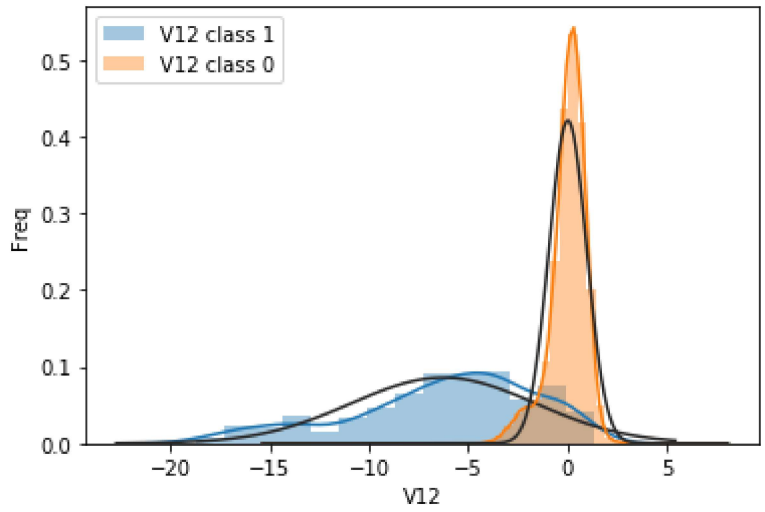
In [25]:

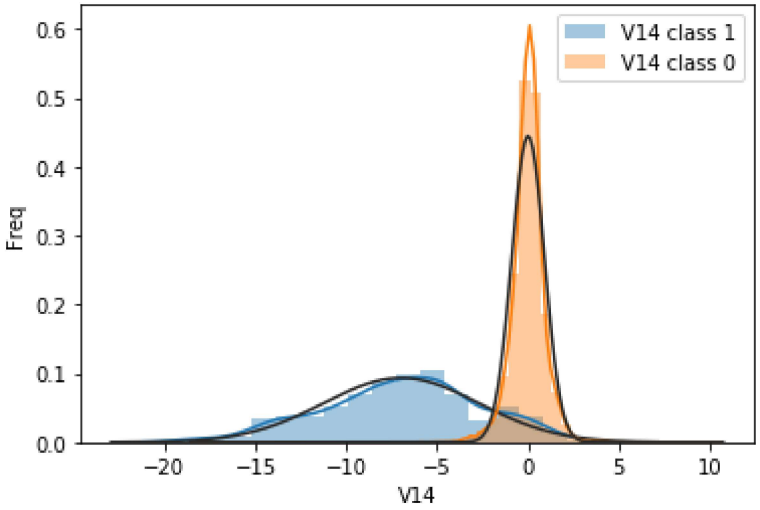
```
from scipy.stats import norm

plt.figure()
sns.distplot(df['V10'][df['Class']==1],fit=norm, label ='V10 class 1')
sns.distplot(df['V10'][df['Class']==0],fit=norm, label ='V10 class 0')
plt.ylabel("Freq")
plt.legend()

plt.figure()
sns.distplot(df['V12'][df['Class']==1],fit=norm, label ='V12 class 1')
sns.distplot(df['V12'][df['Class']==0],fit=norm, label ='V12 class 0')
plt.ylabel("Freq")
plt.legend()

plt.figure()
sns.distplot(df['V14'][df['Class']==1],fit=norm, label ='V14 class 1')
sns.distplot(df['V14'][df['Class']==0],fit=norm, label ='V14 class 0')
plt.ylabel("Freq")
plt.legend()
```

Out[25]:

<matplotlib.legend.Legend at 0x171dcc002b0>






V14 class 1
V14 class 0
Freq
V14
0.6
0.5
0.4
0.3
0.2
0.1
0.0
−20
−15
−10
−5
0
5
10

In [ ]: