In [32]:
```python
import pandas as pd
import numpy as np
import matplotlib as plot
import seaborn as sns
import matplotlib.pyplot as plt
```

In [34]:
```python
data = pd.read_csv('C:/Users/lione/OneDrive/Documents/creditcard.csv')
data.head
```

Out[34]:
```
<bound method NDFrame.head of             Time         V1         V2        V3        V4        V5  \
0            0.0  -1.359807  -0.072781  2.536347  1.378155 -0.338321
1            0.0   1.191857   0.266151  0.166480  0.448154  0.060018
2            1.0  -1.358354  -1.340163  1.773209  0.379780 -0.503198
3            1.0  -0.966272  -0.185226  1.792993 -0.863291 -0.010309
4            2.0  -1.158233   0.877737  1.548718  0.403034 -0.407193
...          ...        ...        ...       ...       ...       ...
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  ...       V21       V22  \
0       0.462388  0.239599  0.098698  0.363787  ... -0.018307  0.277838
1      -0.082361 -0.078803  0.085102 -0.255425  ... -0.225775 -0.638672
2       1.800499  0.791461  0.247676 -1.514654  ...  0.247998  0.771679
3       1.247203  0.237609  0.377436 -1.387024  ... -0.108300  0.005274
4       0.095921  0.592941 -0.270533  0.817739  ... -0.009431  0.798278
...          ...       ...       ...       ...  ...       ...       ...
284802 -2.606837 -4.918215  7.305334  1.914428  ...  0.213454  0.111864
284803  1.058415  0.024330  0.294869  0.584800  ...  0.214205  0.924384
284804  3.031260 -0.296827  0.708417  0.432454  ...  0.232045  0.578229
284805  0.623708 -0.686180  0.679145  0.392087  ...  0.265245  0.800049
284806 -0.649617  1.577006 -0.414650  0.486180  ...  0.261057  0.643078

             V23       V24       V25       V26       V27       V28  Amount  \
0      -0.110474  0.066928  0.128539 -0.189115  0.133558 -0.021053  149.62
1       0.101288 -0.339846  0.167170  0.125895 -0.008983  0.014724    2.69
2       0.909412 -0.689281 -0.327642 -0.139097 -0.055353 -0.059752  378.66
3      -0.190321 -1.175575  0.647376 -0.221929  0.062723  0.061458  123.50
4      -0.137458  0.141267 -0.206010  0.502292  0.219422  0.215153   69.99
...          ...       ...       ...       ...       ...       ...     ...
284802  1.014480 -0.509348  1.436807  0.250034  0.943651  0.823731    0.77
284803  0.012463 -1.016226 -0.606624 -0.395255  0.068472 -0.053527   24.79
284804 -0.037501  0.640134  0.265745 -0.087371  0.004455 -0.026561   67.88
284805 -0.163298  0.123205 -0.569159  0.546668  0.108821  0.104533   10.00
284806  0.376777  0.008797 -0.473649 -0.818267 -0.002415  0.013649  217.00

        Class
0           0
1           0
2           0
3           0
4           0
...       ...
284802      0
284803      0
284804      0
284805      0
284806      0

[284807 rows x 31 columns]>
```

In [35]:
```python
data.describe()
```

Out[35]:

	Time	V1	V2	V3	V4	V5	V6	V7	V8	V9	...	V21	V22	V23	V24	V25	V26	V27	
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	...	2.848070e+05	2.848070e+05	2.848070e+05	2.848070e+05	2.848070e+05	2.848070e+05	2.848070e+05	2.848
mean	94813.859575	1.168375e-15	3.416908e-16	-1.379537e-15	2.074095e-15	9.604066e-16	1.487313e-15	-5.556467e-16	1.213481e-16	-2.406331e-15	...	1.654067e-16	-3.568593e-16	2.578648e-16	4.473266e-15	5.340915e-16	1.683437e-15	-3.660091e-16	-1.227
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	...	7.345240e-01	7.257016e-01	6.244603e-01	6.056471e-01	5.212781e-01	4.822270e-01	4.036325e-01	3.300
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	...	-3.483038e+01	-1.093314e+01	-4.480774e+01	-2.836627e+00	-1.029540e+01	-2.604551e+00	-2.256568e+01	-1.543
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	...	-2.283949e-01	-5.423504e-01	-1.618463e-01	-3.545861e-01	-3.171451e-01	-3.269839e-01	-7.083953e-02	-5.295
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	...	-2.945017e-02	6.781943e-03	-1.119293e-02	4.097606e-02	1.659350e-02	-5.213911e-02	1.342146e-03	1.124
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	...	1.863772e-01	5.285536e-01	1.476421e-01	4.395266e-01	3.507156e-01	2.409522e-01	9.104512e-02	7.827
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	...	2.720284e+01	1.050309e+01	2.252841e+01	4.584549e+00	7.519589e+00	3.517346e+00	3.161220e+01	3.384

8 rows × 31 columns

In [36]:
```python
data.isnull().sum()
```

Out[36]:
```
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
```

In [37]:
```python
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
 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
```

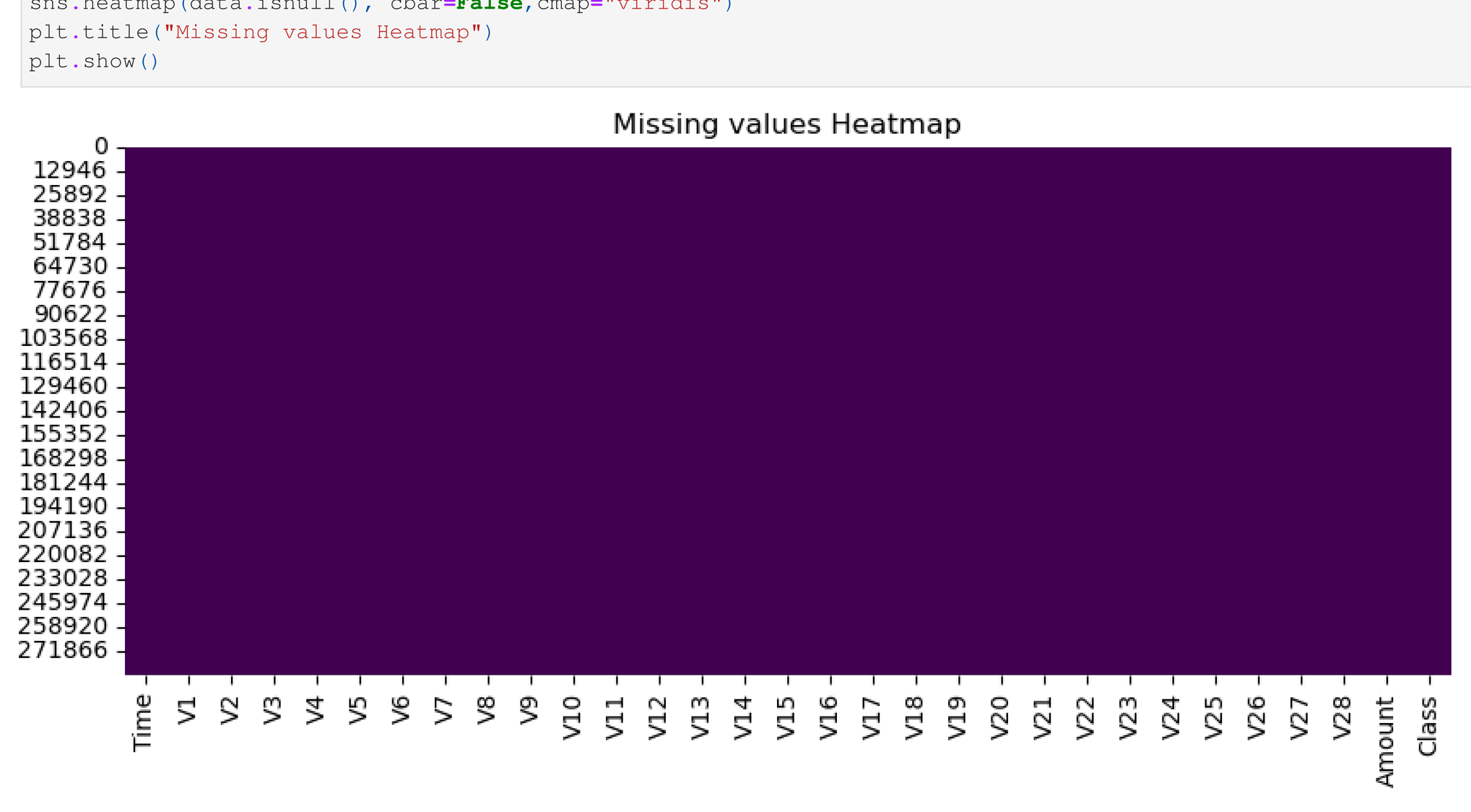
In [22]:
```python
correlation_matrix = data.corr(numeric_only=True) # Only for numeric features
correlation_matrix
```

Out[22]:

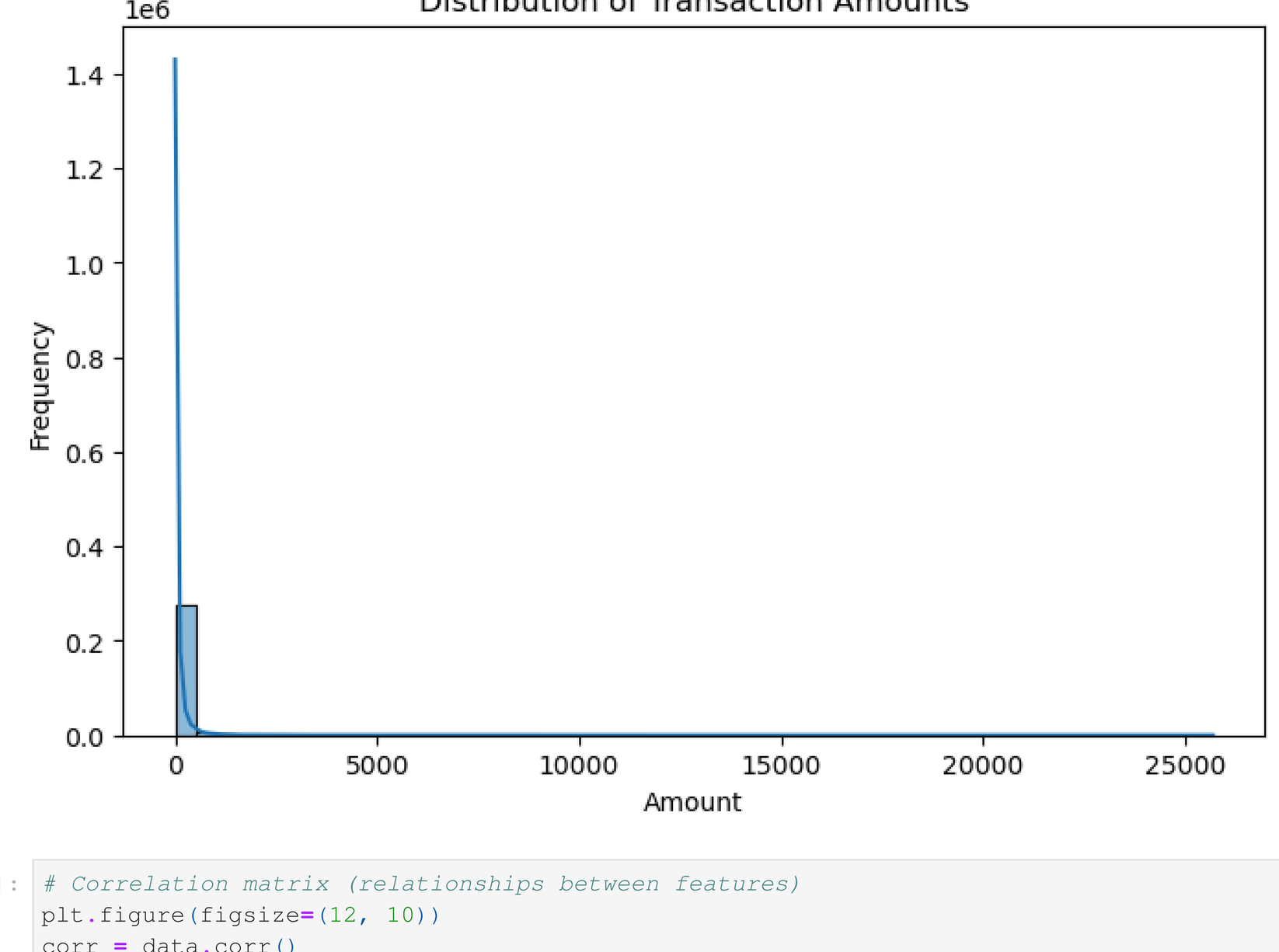
	Time	V1	V2	V3	V4	V5	V6	V7	V8	V9	...	V21	V22	V23	V24	V25	V26	V27	V28	Am
Time	1.000000	1.173963e-01	-1.059333e-02	-4.196182e-01	-1.052602e-01	1.730721e-01	-6.301647e-02	8.471437e-02	-3.694943e-02	-8.660434e-03	...	4.473573e-02	1.440591e-01	5.114236e-02	-1.618187e-02	-2.330828e-01	-4.140710e-02	-5.134591e-03	-9.412688e-03	-0.01
V1	0.117396	1.000000e+00	4.135835e-16	-1.227819e-15	-9.215150e-16	1.812612e-17	-6.506567e-16	-1.005191e-15	-2.433822e-16	-1.513678e-16	...	-2.457409e-16	-4.290944e-16	6.168652e-16	-4.425156e-17	-9.605737e-16	-1.581290e-17	1.198124e-16	2.083082e-15	-0.22
V2	-0.010593	4.135835e-16	1.000000e+00	3.243764e-16	-1.121065e-15	5.157519e-16	2.787346e-16	2.055934e-16	-5.377041e-17	1.978488e-17	...	-8.480447e-17	1.526333e-16	1.634231e-16	1.247925e-17	-4.478846e-16	2.057310e-16	-4.966953e-16	-5.093836e-16	-0.53
V3	-0.419618	-1.227819e-15	3.243764e-16	1.000000e+00	4.711293e-16	-6.539009e-17	1.627627e-15	4.895305e-16	-1.268779e-15	5.568367e-16	...	5.706192e-17	-1.133902e-15	-4.983035e-16	2.686834e-19	-1.104734e-15	-1.238062e-16	1.045747e-15	9.775546e-16	-0.21
V4	-0.105260	-9.215150e-16	-1.121065e-15	4.711293e-16	1.000000e+00	-1.719944e-15	-7.491959e-16	-4.104503e-16	5.697192e-16	6.923247e-16	...	-1.949553e-16	-6.276051e-17	9.164206e-17	1.584638e-16	6.070716e-16	-4.247268e-16	3.977061e-17	-2.761403e-18	0.09
V5	0.173072	1.812612e-17	5.157519e-16	-6.539009e-17	-1.719944e-15	1.000000e+00	2.408382e-16	2.715541e-16	7.437229e-16	7.391702e-16	...	-3.920976e-16	1.253751e-16	-8.428683e-18	-1.149255e-15	4.808532e-16	4.319541e-16	6.590482e-16	-5.613951e-18	-0.38
V6	-0.063016	-6.506567e-16	2.787346e-16	1.627627e-15	-7.491959e-16	2.408382e-16	1.000000e+00	1.191668e-16	-1.104219e-16	4.131207e-16	...	5.833316e-17	-4.705235e-19	1.046712e-16	-1.071589e-15	4.562861e-16	-1.357067e-16	-4.452461e-16	2.594754e-16	0.21
V7	0.084714	-1.005191e-15	2.055934e-16	4.895305e-16	-4.104503e-16	2.715541e-16	1.191668e-16	1.000000e+00	3.344412e-16	1.122501e-15	...	-2.027779e-16	-8.898922e-16	-4.387401e-16	7.434913e-18	-3.094082e-16	-9.657637e-16	-1.782106e-15	-2.776530e-16	0.39
V8	-0.036949	-2.433822e-16	-5.377041e-17	-1.268779e-15	5.697192e-16	7.437229e-16	-1.104219e-16	3.344412e-16	1.000000e+00	4.356078e-16	...	3.892798e-16	2.026927e-16	6.377260e-17	-1.047097e-16	-4.653279e-16	-1.727276e-16	1.299943e-16	-6.200930e-16	-0.10
V9	-0.008660	-1.513678e-16	1.978488e-17	5.568367e-16	6.923247e-16	7.391702e-16	4.131207e-16	1.122501e-15	4.356078e-16	1.000000e+00	...	1.936953e-16	-7.071869e-16	-5.214137e-16	-1.430343e-16	6.757763e-16	-7.888853e-16	-6.709655e-17	1.110541e-15	-0.04
V10	0.030617	7.388135e-17	-3.991394e-16	1.156587e-15	2.232685e-16	-5.202306e-16	5.932243e-17	-7.492834e-17	-2.801370e-16	-4.642274e-16	...	1.177547e-15	-6.418202e-16	3.214491e-16	-1.355885e-16	-2.846052e-16	-3.028119e-16	-2.197977e-16	4.864782e-17	-0.10
V11	-0.247689	2.125498e-16	1.975426e-16	1.576830e-15	3.459380e-16	7.203963e-16	1.980503e-15	1.425248e-16	2.487043e-16	1.354680e-16	...	-5.658364e-16	7.772895e-16	-4.505332e-16	1.933267e-15	-5.600475e-16	-1.003221e-16	-2.640281e-16	-3.792314e-16	0.00
V12	0.124348	2.053457e-16	-9.568710e-17	6.310231e-16	-5.625518e-16	7.412552e-16	2.375468e-16	-3.536655e-18	1.839891e-16	-1.079314e-15	...	7.300527e-16	1.644699e-16	1.800885e-16	4.436512e-16	-5.712973e-16	-2.359969e-16	-4.672391e-16	6.415167e-16	-0.00
V13	-0.065902	-2.425603e-17	6.295388e-16	2.807652e-16	1.303306e-16	5.886991e-16	-1.211182e-16	1.266462e-17	-2.921856e-16	2.251072e-15	...	1.008461e-16	6.747721e-17	-7.132064e-16	-1.397470e-16	-5.497612e-16	-1.769255e-16	-4.720898e-16	1.144372e-15	0.00
V14	-0.098757	-5.020280e-16	-1.730566e-16	4.739859e-16	2.282280e-16	6.565143e-16	2.621312e-16	2.607772e-16	-8.599156e-16	3.784757e-15	...	-3.356561e-16	3.740383e-16	3.883204e-16	2.003482e-16	-8.547932e-16	-1.660327e-16	1.044274e-16	2.289427e-15	0.03
V15	-0.183453	3.547782e-16	-4.995814e-17	9.068793e-16	1.377649e-16	-8.720275e-16	-1.531188e-15	-1.690540e-16	4.127777e-16	-1.051167e-15	...	6.605263e-17	-4.208921e-16	-3.912243e-16	-4.478263e-16	3.206423e-16	2.817791e-16	-1.143519e-15	-1.194130e-15	-0.00
V16	0.011903	7.212815e-17	1.177316e-17	8.299445e-16	-9.614528e-16	2.246261e-15	2.623672e-18	5.869302e-17	-5.254741e-16	-1.214086e-15	...	-4.715090e-16	-7.923387e-17	5.020770e-16	-3.005985e-16	-1.345418e-15	-7.290010e-16	6.789513e-16	7.588849e-16	-0.00
V17	-0.073297	-3.879840e-16	-2.685296e-16	7.614712e-16	-2.699612e-16	1.281914e-16	2.015618e-16	2.177192e-16	-2.269549e-16	1.113695e-15	...	-8.230527e-16	-8.743398e-16	3.706214e-16	-2.403828e-16	2.666806e-16	6.932833e-16	6.148525e-16	-5.534540e-17	0.00
V18	0.090438	3.230206e-17	3.284605e-16	1.509897e-16	-5.103644e-16	5.308590e-16	1.223814e-16	7.604126e-17	-3.667974e-16	4.993240e-16	...	-9.408680e-16	-4.819365e-16	-1.912006e-16	-8.986916e-17	-6.629212e-17	2.990167e-16	2.242791e-16	7.976796e-16	0.03
V19	0.028975	1.502024e-16	-7.118719e-18	3.463522e-16	-3.980557e-16	-1.450421e-16	-1.865597e-16	-1.881008e-16	-3.875186e-16	-1.376135e-16	...	5.115885e-16	-1.163768e-15	7.032035e-16	2.587708e-17	9.577163e-16	5.898033e-16	-2.959370e-16	-1.405379e-15	-0.05
V20	-0.050866	4.654551e-16	2.506675e-16	-9.316409e-16	-1.857247e-16	-3.554057e-16	-1.858755e-16	9.379684e-16	2.033737e-16	-2.343720e-16	...	-7.614597e-16	1.009285e-15	2.712885e-16	1.277215e-16	1.410054e-16	-2.803504e-16	-1.138829e-15	-2.436795e-16	0.33
V21	0.044736	-2.457409e-16	-8.480447e-17	5.706192e-17	-1.949553e-16	-3.920976e-16	5.833316e-17	-2.027779e-16	3.892798e-16	1.936953e-16	...	1.000000e+00	3.649908e-15	8.119580e-16	1.761054e-16	-1.686082e-16	-5.557329e-16	-1.211281e-15	5.278775e-16	0.10
V22	0.144059	-4.290944e-16	1.526333e-16	-1.133902e-15	-6.276051e-17	1.253751e-16	-4.705235e-19	-8.898922e-16	2.026927e-16	-7.071869e-16	...	3.649908e-15	1.000000e+00	-7.303916e-17	9.970809e-17	-5.018575e-16	-2.503187e-17	8.461337e-17	-6.627203e-16	-0.06
V23	0.051142	6.168652e-16	1.634231e-16	-4.983035e-16	9.164206e-17	-8.428683e-18	1.046712e-16	-4.387401e-16	6.377260e-17	-5.214137e-16	...	8.119580e-16	-7.303916e-17	1.000000e+00	2.130519e-17	-8.232727e-17	1.114524e-15	2.839721e-16	1.481903e-15	-0.11
V24	-0.016182	-4.425156e-17	1.247925e-17	2.686834e-19	1.584638e-16	-1.149255e-15	-1.071589e-15	7.434913e-18	-1.047097e-16	-1.430343e-16	...	1.761054e-16	9.970809e-17	2.130519e-17	1.000000e+00	1.015391e-15	1.343722e-16	-2.274142e-16	-2.819805e-16	0.00
V25	-0.233083	-9.605737e-16	-4.478846e-16	-1.104734e-15	6.070716e-16	4.808532e-16	4.562861e-16	-3.094082e-16	-4.653279e-16	6.757763e-16	...	-1.686082e-16	-5.018575e-16	-8.232727e-17	1.015391e-15	1.000000e+00	2.646517e-15	-6.406679e-16	-7.008939e-16	-0.04
V26	-0.041407	-1.581290e-17	2.057310e-16	-1.238062e-16	-4.247268e-16	4.319541e-16	-1.357067e-16	-9.657637e-16	-1.727276e-16	-7.888853e-16	...	-5.557329e-16	-2.503187e-17	1.114524e-15	1.343722e-16	2.646517e-15	1.000000e+00	-3.667715e-16	-2.782204e-16	-0.00
V27	-0.005135	1.198124e-16	-4.966953e-16	1.045747e-15	3.977061e-17	6.590482e-16	-4.452461e-16	-1.782106e-15	1.299943e-16	-6.709655e-17	...	-1.211281e-15	8.461337e-17	2.839721e-16	-2.274142e-16	-6.406679e-16	-3.667715e-16	1.000000e+00	-3.061287e-16	0.02
V28	-0.009413	2.083082e-15	-5.093836e-16	9.775546e-16	-2.761403e-18	-5.613951e-18	2.594754e-16	-2.776530e-16	-6.200930e-16	1.110541e-15	...	5.278775e-16	-6.627203e-16	1.481903e-15	-2.819805e-16	-7.008939e-16	-2.782204e-16	-3.061287e-16	1.000000e+00	0.01
Amount	-0.010596	-2.277087e-01	-5.314089e-01	-2.108805e-01	9.873167e-02	-3.863563e-01	2.159812e-01	3.973113e-01	-1.030791e-01	-4.424560e-02	...	1.059989e-01	-6.480065e-02	-1.126326e-01	5.146217e-03	-4.783686e-02	-3.208037e-03	2.882546e-02	1.025822e-02	1.00
Class	-0.012323	-1.013473e-01	9.128865e-02	-1.929608e-01	1.334475e-01	-9.497430e-02	-4.364316e-02	-1.872566e-01	1.987512e-02	-9.773269e-02	...	4.041338e-02	8.053175e-04	-2.685156e-03	-7.220907e-03	3.307706e-03	4.455398e-03	1.757973e-02	9.536041e-03	0.00

31 rows × 31 columns

In [43]:
```python
# Visualize missing value using heatmap
plt.figure(figsize=(10,4))
sns.heatmap(data.isnull(), cbar=False,cmap="viridis")
plt.title("Missing values Heatmap")
plt.show()
```

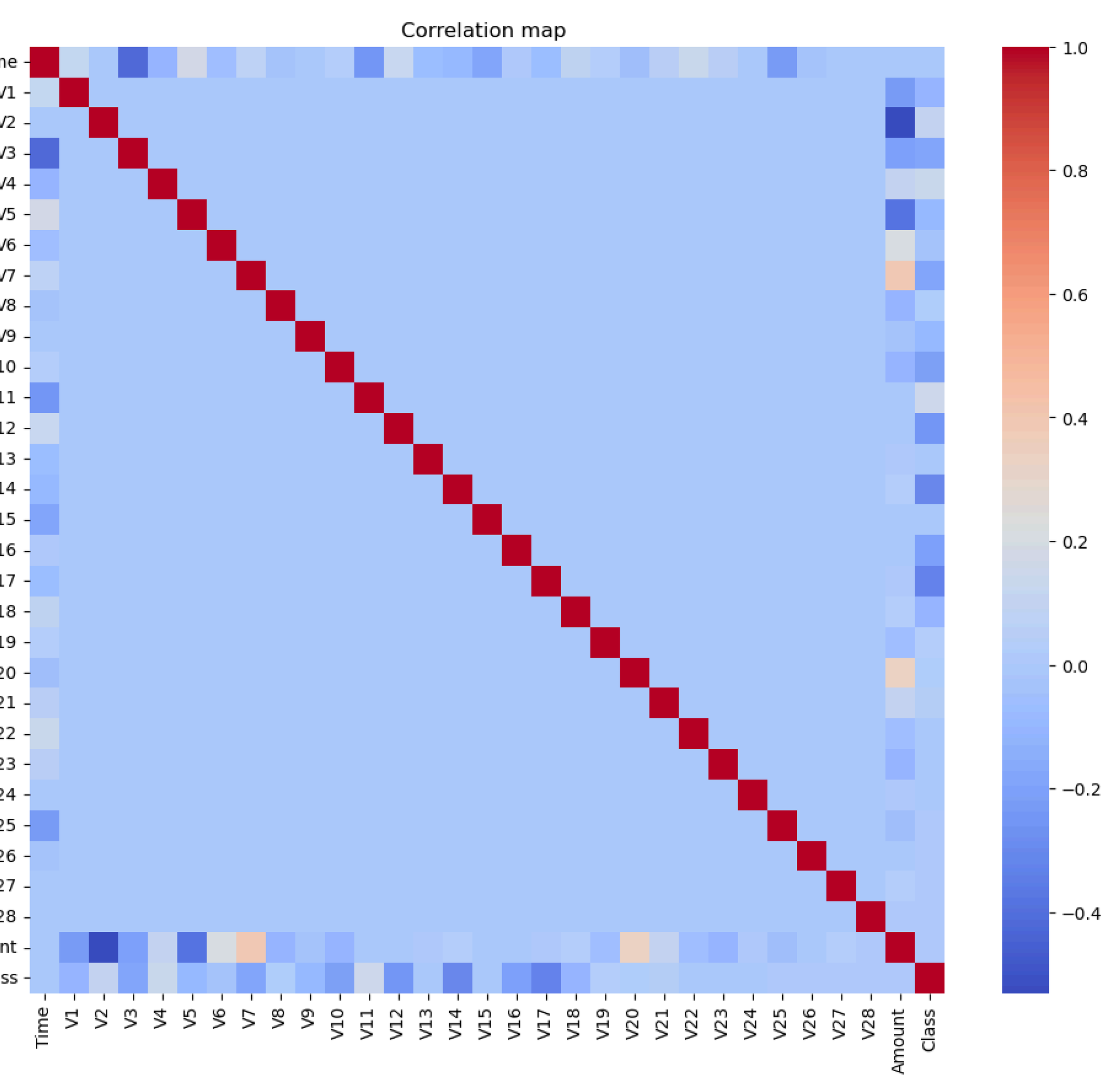


In [45]:
```python
plt.figure(figsize=(8, 5))
sns.histplot(data['Amount'], bins=50, kde=True)
plt.title("Distribution of Transaction Amounts")
plt.xlabel("Amount")
plt.ylabel("Frequency")
plt.show()
```



In [47]:
```python
# Correlation matrix (relationships between features)
plt.figure(figsize=(12, 10))
corr = data.corr()
sns.heatmap(corr, cmap="coolwarm", annot=False, fmt=".2f")
plt.title("Correlation map ")
```

Out[47]: `Text(0.5, 1.0, 'Correlation map ')`



In [49]:
```python
# Impact on model performance
# The biggest challenges in credit card fraud detection is the class imbalance. In this cases, a small amount of the fraction is around 0.2%.
```

```
# How to address it
# To address it, we can do an oversampling of the minority class, an undersampling of the majority class.
```

In [ ]: