In [1]:
```
pip install matplotlib
```

```
Requirement already satisfied: matplotlib in d:\virtualenv\myenv\lib\site-packages (3.9.0)
Requirement already satisfied: contourpy>=1.0.1 in d:\virtualenv\myenv\lib\site-packages (from matplotlib) (1.2.1)
Requirement already satisfied: cycler>=0.10 in d:\virtualenv\myenv\lib\site-packages (from matplotlib) (0.12.1)
Requirement already satisfied: fonttools>=4.22.0 in d:\virtualenv\myenv\lib\site-packages (from matplotlib) (4.53.
0)
Requirement already satisfied: kiwisolver>=1.3.1 in d:\virtualenv\myenv\lib\site-packages (from matplotlib) (1.4.5)
Requirement already satisfied: numpy>=1.23 in d:\virtualenv\myenv\lib\site-packages (from matplotlib) (1.26.3)
Requirement already satisfied: packaging>=20.0 in d:\virtualenv\myenv\lib\site-packages (from matplotlib) (24.1)
Requirement already satisfied: pillow>=8 in d:\virtualenv\myenv\lib\site-packages (from matplotlib) (10.2.0)
Requirement already satisfied: pyparsing>=2.3.1 in d:\virtualenv\myenv\lib\site-packages (from matplotlib) (3.1.2)
Requirement already satisfied: python-dateutil>=2.7 in d:\virtualenv\myenv\lib\site-packages (from matplotlib) (2.
9.0.post0)
Requirement already satisfied: six>=1.5 in d:\virtualenv\myenv\lib\site-packages (from python-dateutil>=2.7->matplo
tlib) (1.16.0)
Note: you may need to restart the kernel to use updated packages.
[notice] A new release of pip is available: 24.1 -> 24.2
[notice] To update, run: python.exe -m pip install --upgrade pip
```

In [2]:
```
pip install pandas
```

```
Requirement already satisfied: pandas in d:\virtualenv\myenv\lib\site-packages (2.2.2)
Requirement already satisfied: numpy>=1.23.2 in d:\virtualenv\myenv\lib\site-packages (from pandas) (1.26.3)
Requirement already satisfied: python-dateutil>=2.8.2 in d:\virtualenv\myenv\lib\site-packages (from pandas) (2.9.
0.post0)
Requirement already satisfied: pytz>=2020.1 in d:\virtualenv\myenv\lib\site-packages (from pandas) (2024.1)
Requirement already satisfied: tzdata>=2022.7 in d:\virtualenv\myenv\lib\site-packages (from pandas) (2024.1)
Requirement already satisfied: six>=1.5 in d:\virtualenv\myenv\lib\site-packages (from python-dateutil>=2.8.2->pand
as) (1.16.0)
Note: you may need to restart the kernel to use updated packages.
[notice] A new release of pip is available: 24.1 -> 24.2
[notice] To update, run: python.exe -m pip install --upgrade pip
```

In [3]:
```
!pip install seaborn
```

```
Requirement already satisfied: seaborn in d:\virtualenv\myenv\lib\site-packages (0.13.2)
Requirement already satisfied: numpy!=1.24.0,>=1.20 in d:\virtualenv\myenv\lib\site-packages (from seaborn) (1.26.
3)
Requirement already satisfied: pandas>=1.2 in d:\virtualenv\myenv\lib\site-packages (from seaborn) (2.2.2)
Requirement already satisfied: matplotlib!=3.6.1,>=3.4 in d:\virtualenv\myenv\lib\site-packages (from seaborn) (3.
9.0)
Requirement already satisfied: contourpy>=1.0.1 in d:\virtualenv\myenv\lib\site-packages (from matplotlib!=3.6.1,>=
3.4->seaborn) (1.2.1)
Requirement already satisfied: cycler>=0.10 in d:\virtualenv\myenv\lib\site-packages (from matplotlib!=3.6.1,>=3.4-
>seaborn) (0.12.1)
Requirement already satisfied: fonttools>=4.22.0 in d:\virtualenv\myenv\lib\site-packages (from matplotlib!=3.6.1,>
=3.4->seaborn) (4.53.0)
Requirement already satisfied: kiwisolver>=1.3.1 in d:\virtualenv\myenv\lib\site-packages (from matplotlib!=3.6.1,>
=3.4->seaborn) (1.4.5)
Requirement already satisfied: packaging>=20.0 in d:\virtualenv\myenv\lib\site-packages (from matplotlib!=3.6.1,>=
3.4->seaborn) (24.1)
Requirement already satisfied: pillow>=8 in d:\virtualenv\myenv\lib\site-packages (from matplotlib!=3.6.1,>=3.4->se
aborn) (10.2.0)
Requirement already satisfied: pyparsing>=2.3.1 in d:\virtualenv\myenv\lib\site-packages (from matplotlib!=3.6.1,>=
3.4->seaborn) (3.1.2)
Requirement already satisfied: python-dateutil>=2.7 in d:\virtualenv\myenv\lib\site-packages (from matplotlib!=3.6.
1,>=3.4->seaborn) (2.9.0.post0)
Requirement already satisfied: pytz>=2020.1 in d:\virtualenv\myenv\lib\site-packages (from pandas>=1.2->seaborn) (2
024.1)
Requirement already satisfied: tzdata>=2022.7 in d:\virtualenv\myenv\lib\site-packages (from pandas>=1.2->seaborn)
(2024.1)
Requirement already satisfied: six>=1.5 in d:\virtualenv\myenv\lib\site-packages (from python-dateutil>=2.7->matplo
tlib!=3.6.1,>=3.4->seaborn) (1.16.0)
[notice] A new release of pip is available: 24.1 -> 24.2
[notice] To update, run: python.exe -m pip install --upgrade pip
```

In [4]:
```
import os
import pandas as pd
import seaborn as sns
from matplotlib import pyplot as plt
import numpy as np
from config import Config
```

Definitions

Paths Declaration

Qui, di seguito, vengono riportati i path relativi al dataset di training, validation e testing. In particolare, il dataset di validation coincide con il primo dataset di testing contenuto nella cartella './data/test/'. Infatti, all'interno della directory './data/test/' sono presenti 5 possibili insiemi di dati così da poter essere sfruttati durante il testing.

In [5]:
```
config = Config(
    data_dir='./data',
    train_data_subdir='train/',
    val_data_subdir='test/',
    test_data_subdir='test/'
)
```

```
data_dir = ['test', 'train', 'val']
train_data_dir = ['TRAIN_LGHG2@n10degC_to_25degC_Norm_5Inputs.csv']
val_data_dir = ['01_TEST_LGHG2@n10degC_Norm_(05_Inputs).csv', '02_TEST_LGHG2@0degC_Norm_(05_Inputs).csv', '03_TEST_
LGHG2@10degC_Norm_(05_Inputs).csv', '04_TEST_LGHG2@25degC_Norm_(05_Inputs).csv']
test_data_dir = ['01_TEST_LGHG2@n10degC_Norm_(05_Inputs).csv', '02_TEST_LGHG2@0degC_Norm_(05_Inputs).csv', '03_TEST
_LGHG2@10degC_Norm_(05_Inputs).csv', '04_TEST_LGHG2@25degC_Norm_(05_Inputs).csv']
```

Dataset Analysis and Visualization

In [6]:
```
train_data_filename = os.listdir(config.get_train_data_dir())[0]
train_data_path = os.path.join(config.get_train_data_dir(), train_data_filename)

train_data = pd.read_csv(train_data_path)
train_data
```

Out[6]:

	V	I	Temp	V_avg	I_avg	SOC
0	0.385148	0.75102	0.303101	0.385148	0.75102	0.206417
1	0.385152	0.75102	0.304591	0.385150	0.75102	0.206417
2	0.385156	0.75102	0.306081	0.385152	0.75102	0.206417
3	0.385160	0.75102	0.307572	0.385154	0.75102	0.206417
4	0.385164	0.75102	0.309062	0.385156	0.75102	0.206417
...	...	...	...	...	...	...
669951	0.478843	0.75102	0.008477	0.459558	0.75102	0.283243
669952	0.478843	0.75102	0.008477	0.459699	0.75102	0.283243
669953	0.478843	0.75102	0.008477	0.459839	0.75102	0.283243
669954	0.478961	0.75102	0.008477	0.459979	0.75102	0.283243
669955	0.478961	0.75102	0.008477	0.460117	0.75102	0.283243

669956 rows × 6 columns

In [7]:
```
train_data.isna().sum()
```

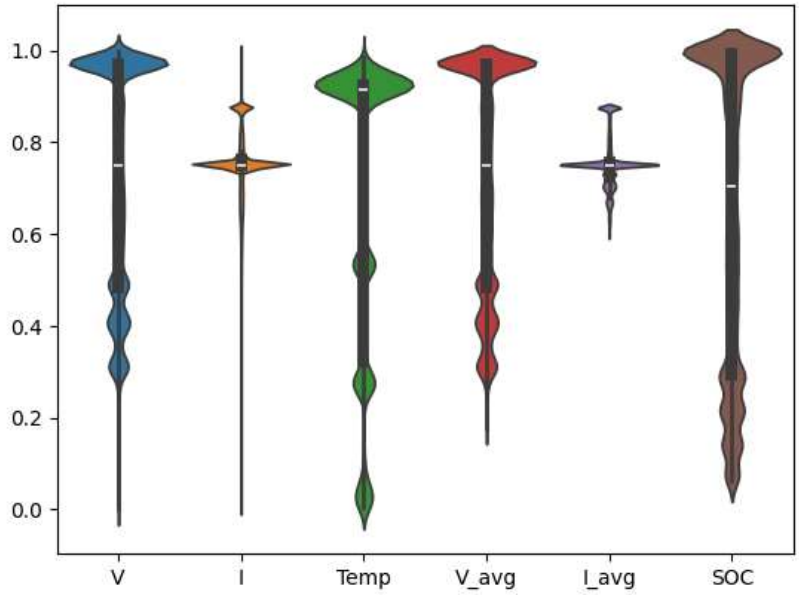
Out[7]:
```
V        0
I        0
Temp     0
V_avg    0
I_avg    0
SOC      0
dtype: int64
```

In [8]:
```
train_data.info()
```

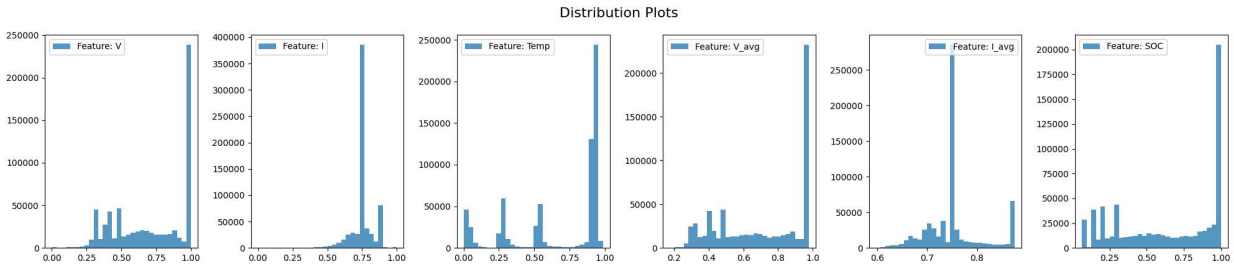
```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 669956 entries, 0 to 669955
Data columns (total 6 columns):
 #   Column  Non-Null Count   Dtype  
---  ------  --------------   -----  
 0   V       669956 non-null  float64
 1   I       669956 non-null  float64
 2   Temp    669956 non-null  float64
 3   V_avg   669956 non-null  float64
 4   I_avg   669956 non-null  float64
 5   SOC     669956 non-null  float64
dtypes: float64(6)
memory usage: 30.7 MB
```

In [9]:
```
sns.violinplot(data=train_data)
```

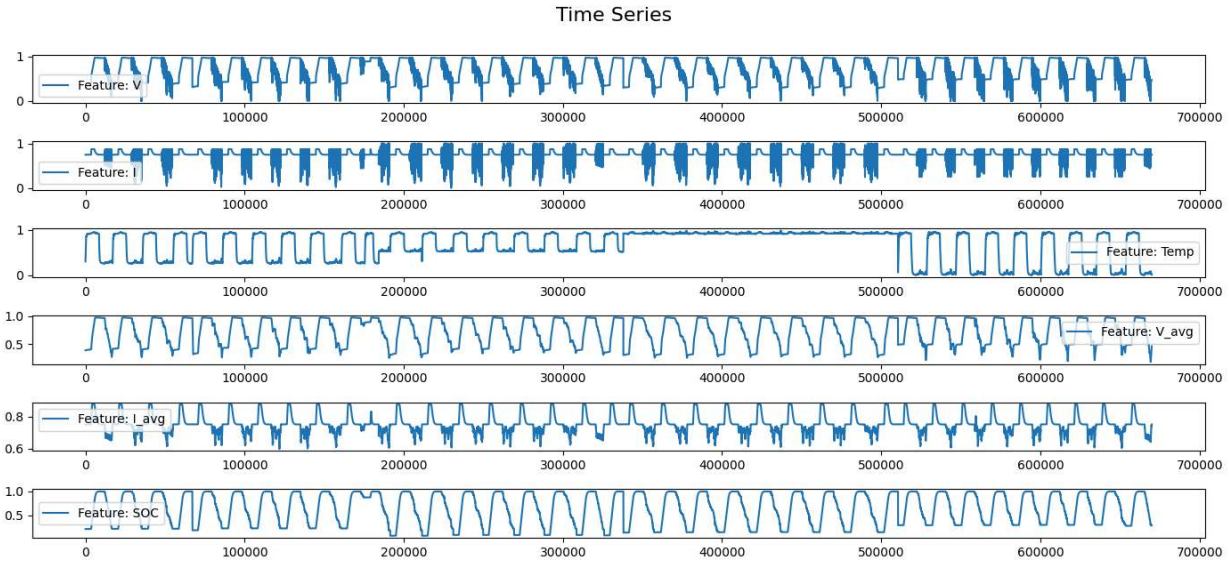
Out[9]: <Axes: >



In [10]:
```
fig, axs = plt.subplots(1, train_data.shape[1], figsize=(20, 4))
for i, key in enumerate(train_data.keys()):
    axs[i].hist(train_data[key], bins=30, alpha=0.75)
    axs[i].legend([f'Feature: {key}'])
plt.tight_layout()
plt.suptitle('Distribution Plots', fontsize=16, y=1.06)
plt.show()
```



In [11]:
```
fig, axs = plt.subplots(train_data.shape[1], 1, figsize=(14, 6))
for i, key in enumerate(train_data.keys()):
    plt.subplot(train_data.shape[1], 1, i + 1)
    plt.plot(train_data[key])
    plt.legend([f'Feature: {key}'])
plt.tight_layout()
plt.suptitle('Time Series', fontsize=16, y=1.06)
plt.show()
```



In [12]:
```
plt.figure(figsize=(16, 6))
sns.heatmap(train_data.corr(), annot=True)
```

```
plt.title('Correlation between the features')
plt.show()
```