•[11]:
```
import matplotlib as plt
import numpy as np
import pandas as pd
data=pd.read_csv('01.Data Cleaning and Preprocessing.csv')
data.info
```

[11]:
```
<bound method DataFrame.info of     Observation  Y-Kappa  ChipRate  BF-CMratio  BlowFlow  ChipLevel4   \
0      31-00:00    23.10    16.520     121.717  1177.607      169.805
1      31-01:00    27.60    16.810      79.022  1328.360      341.327
2      31-02:00    23.19    16.709      79.562  1329.407      239.161
3      31-03:00    23.60    16.478      81.011  1334.877      213.527
4      31-04:00    22.90    15.618      93.244  1334.168      243.131
..          ...      ...       ...         ...       ...          ...
319    10-16:00    23.75    12.667      93.450  1178.252      276.955
320     9-19:00    19.80    12.558      94.352  1184.119      297.071
321     9-20:00    23.01    12.550      90.842  1188.517      289.826
322     9-21:00    24.32    13.083      88.910  1192.879      318.006
323     9-22:00    25.75    13.417      85.451  1186.342      248.312

     T-upperExt-2   T-lowerExt-2    UCZAA  WhiteFlow-4   ...  SteamFlow-4   \
0         358.282        329.545  1.443       599.253  ...       67.122
1         351.050        329.067  1.549       537.201  ...       60.012
2         350.022        329.260  1.600       549.611  ...       61.304
3         350.938        331.142  1.604       623.362  ...       68.496
4         351.640        332.709    NaN       638.672  ...       70.022
..            ...            ...    ...           ...  ...          ...
```

Untitled.ipynb

Code

Notebook Python (Pyodide)

301 rows × 23 columns

[15]:
```
data=data.drop_duplicates()
data
```

[15]:

	Observation	Y-Kappa	ChipRate	BF-CMratio	BlowFlow	ChipLevel4	T-upperExt-2	T-lowerExt-2	UCZAA	WhiteFlow-4	...	SteamFlow-4	Lower-HeatT-3	Upper-HeatT-3	ChipMass-4	W
0	31-00:00	23.10	16.520	121.717	1177.607	169.805	358.282	329.545	1.443	599.253	...	67.122	329.432	303.099	175.964	
1	31-01:00	27.60	16.810	79.022	1328.360	341.327	351.050	329.067	1.549	537.201	...	60.012	330.823	304.879	163.202	
2	31-02:00	23.19	16.709	79.562	1329.407	239.161	350.022	329.260	1.600	549.611	...	61.304	329.140	303.383	164.013	
3	31-03:00	23.60	16.478	81.011	1334.877	213.527	350.938	331.142	1.604	623.362	...	68.496	328.875	302.254	181.487	
4	31-04:00	22.90	15.618	93.244	1334.168	243.131	351.640	332.709	NaN	638.672	...	70.022	328.352	300.954	183.929	
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	
298	12-09:00	20.90	15.167	84.640	1283.706	339.440	354.803	311.041	1.635	532.419	...	65.561	332.924	307.626	145.299	
299	12-10:00	24.98	NaN	85.034	1278.345	368.564	357.723	321.387	NaN	520.365	...	65.729	332.523	307.169	151.544	
300	12-11:00	21.00	NaN	88.013	1307.722	278.842	357.438	323.757	NaN	553.070	...	65.795	331.263	306.400	157.954	
301	12-12:00	21.40	NaN	85.490	1255.986	273.484	361.365	322.689	NaN	590.199	...	71.456	333.032	308.732	174.069	
307	31-05:00	20.89	14.308	94.172	1327.832	251.120	351.263	332.485	1.522	631.514	...	71.286	328.699	300.706	180.229	

Untitled.ipynb

Code

Notebook Python (Pyodide)

[4]:
```
data.isnull()
```

[4]:

	Observation	Y-Kappa	ChipRate	BF-CMratio	BlowFlow	ChipLevel4	T-upperExt-2	T-lowerExt-2	UCZAA	WhiteFlow-4	...	SteamFlow-4	Lower-HeatT-3	Upper-HeatT-3	ChipMass-4	V
0	False	False	False	False	False	False	False	False	False	False	...	False	False	False	False	
1	False	False	False	False	False	False	False	False	False	False	...	False	False	False	False	
2	False	False	False	False	False	False	False	False	False	False	...	False	False	False	False	
3	False	False	False	False	False	False	False	False	False	False	...	False	False	False	False	
4	False	False	False	False	False	False	False	False	True	False	...	False	False	False	False	
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	
298	False	False	False	False	False	False	False	False	False	False	...	False	False	False	False	
299	False	False	True	False	False	False	False	False	True	False	...	False	False	False	False	
300	False	False	True	False	False	False	False	False	True	False	...	False	False	False	False	
301	False	False	True	False	False	False	False	False	True	False	...	False	False	False	False	
307	False	False	False	False	False	False	False	False	False	False	...	False	False	False	False	

Collapse Output

Untitled.ipynb

Code

Notebook Python (Pyodide)

[16]:
```
data2=data.fillna(0)
data2
```

[16]:

	Observation	Y-Kappa	ChipRate	BF-CMratio	BlowFlow	ChipLevel4	T-upperExt-2	T-lowerExt-2	UCZAA	WhiteFlow-4	...	SteamFlow-4	Lower-HeatT-3	Upper-HeatT-3	ChipMass-4	V
0	31-00:00	23.10	16.520	121.717	1177.607	169.805	358.282	329.545	1.443	599.253	...	67.122	329.432	303.099	175.964	
1	31-01:00	27.60	16.810	79.022	1328.360	341.327	351.050	329.067	1.549	537.201	...	60.012	330.823	304.879	163.202	
2	31-02:00	23.19	16.709	79.562	1329.407	239.161	350.022	329.260	1.600	549.611	...	61.304	329.140	303.383	164.013	
3	31-03:00	23.60	16.478	81.011	1334.877	213.527	350.938	331.142	1.604	623.362	...	68.496	328.875	302.254	181.487	
4	31-04:00	22.90	15.618	93.244	1334.168	243.131	351.640	332.709	0.000	638.672	...	70.022	328.352	300.954	183.929	
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	
298	12-09:00	20.90	15.167	84.640	1283.706	339.440	354.803	311.041	1.635	532.419	...	65.561	332.924	307.626	145.299	
299	12-10:00	24.98	0.000	85.034	1278.345	368.564	357.723	321.387	0.000	520.365	...	65.729	332.523	307.169	151.544	
300	12-11:00	21.00	0.000	88.013	1307.722	278.842	357.438	323.757	0.000	553.070	...	65.795	331.263	306.400	157.954	
301	12-12:00	21.40	0.000	85.490	1255.986	273.484	361.365	322.689	0.000	590.199	...	71.456	333.032	308.732	174.069	
307	31-05:00	20.89	14.308	94.172	1327.832	251.120	351.263	332.485	1.522	631.514	...	71.286	328.699	300.706	180.229	

Untitled.ipynb

Code

Notebook Python (Pyodide)

[18]:
```
data3=data.fillna(method='pad')
data3
```

[18]:

	Observation	Y-Kappa	ChipRate	BF-CMratio	BlowFlow	ChipLevel4	T-upperExt-2	T-lowerExt-2	UCZAA	WhiteFlow-4	...	SteamFlow-4	Lower-HeatT-3	Upper-HeatT-3	ChipMass-4	W
0	31-00:00	23.10	16.520	121.717	1177.607	169.805	358.282	329.545	1.443	599.253	...	67.122	329.432	303.099	175.964	
1	31-01:00	27.60	16.810	79.022	1328.360	341.327	351.050	329.067	1.549	537.201	...	60.012	330.823	304.879	163.202	
2	31-02:00	23.19	16.709	79.562	1329.407	239.161	350.022	329.260	1.600	549.611	...	61.304	329.140	303.383	164.013	
3	31-03:00	23.60	16.478	81.011	1334.877	213.527	350.938	331.142	1.604	623.362	...	68.496	328.875	302.254	181.487	
4	31-04:00	22.90	15.618	93.244	1334.168	243.131	351.640	332.709	1.604	638.672	...	70.022	328.352	300.954	183.929	
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	
298	12-09:00	20.90	15.167	84.640	1283.706	339.440	354.803	311.041	1.635	532.419	...	65.561	332.924	307.626	145.299	
299	12-10:00	24.98	15.167	85.034	1278.345	368.564	357.723	321.387	1.635	520.365	...	65.729	332.523	307.169	151.544	
300	12-11:00	21.00	15.167	88.013	1307.722	278.842	357.438	323.757	1.635	553.070	...	65.795	331.263	306.400	157.954	
301	12-12:00	21.40	15.167	85.490	1255.986	273.484	361.365	322.689	1.635	590.199	...	71.456	333.032	308.732	174.069	
307	31-05:00	20.89	14.308	94.172	1327.832	251.120	351.263	332.485	1.522	631.514	...	71.286	328.699	300.706	180.229	

mple 0 3 Python (Pyodide) Idle Mode: Command Ln 1, Col 1 Untitled.ipynb 4

Untitled.ipynb

Code

Notebook Python (Pyodide)

[19]:
```
data4=data.fillna(method='bfill')
data4
```

[19]:

	Observation	Y-Kappa	ChipRate	BF-CMratio	BlowFlow	ChipLevel4	T-upperExt-2	T-lowerExt-2	UCZAA	WhiteFlow-4	...	SteamFlow-4	Lower-HeatT-3	Upper-HeatT-3	ChipMass-4	V
0	31-00:00	23.10	16.520	121.717	1177.607	169.805	358.282	329.545	1.443	599.253	...	67.122	329.432	303.099	175.964	
1	31-01:00	27.60	16.810	79.022	1328.360	341.327	351.050	329.067	1.549	537.201	...	60.012	330.823	304.879	163.202	
2	31-02:00	23.19	16.709	79.562	1329.407	239.161	350.022	329.260	1.600	549.611	...	61.304	329.140	303.383	164.013	
3	31-03:00	23.60	16.478	81.011	1334.877	213.527	350.938	331.142	1.604	623.362	...	68.496	328.875	302.254	181.487	
4	31-04:00	22.90	15.618	93.244	1334.168	243.131	351.640	332.709	1.436	638.672	...	70.022	328.352	300.954	183.929	
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	
298	12-09:00	20.90	15.167	84.640	1283.706	339.440	354.803	311.041	1.635	532.419	...	65.561	332.924	307.626	145.299	
299	12-10:00	24.98	14.308	85.034	1278.345	368.564	357.723	321.387	1.522	520.365	...	65.729	332.523	307.169	151.544	
300	12-11:00	21.00	14.308	88.013	1307.722	278.842	357.438	323.757	1.522	553.070	...	65.795	331.263	306.400	157.954	
301	12-12:00	21.40	14.308	85.490	1255.986	273.484	361.365	322.689	1.522	590.199	...	71.456	333.032	308.732	174.069	
307	31-05:00	20.89	14.308	94.172	1327.832	251.120	351.263	332.485	1.522	631.514	...	71.286	328.699	300.706	180.229	

mple 0 3 Python (Pyodide) | Idle Mode: Edit Ln 1, Col 1 Untitled.ipynb 4

[25]:
```
Q1=data2.quantile(0.25)
Q2=data2.quantile(0.75)
IQR=Q2-Q1
print(IQR)
```

```
Y-Kappa             4.550
ChipRate            2.233
BF-CMratio         10.912
BlowFlow           96.766
ChipLevel4        105.868
T-upperExt-2       11.994
T-lowerExt-2        7.609
UCZAA               0.152
WhiteFlow-4       100.098
AAWhiteSt-4         6.143
AA-Wood-4           1.486
ChipMoisture-4      2.186
SteamFlow-4         8.840
Lower-HeatT-3       8.585
Upper-HeatT-3       7.852
ChipMass-4         19.347
WeakLiquorF       180.613
BlackFlow-2       280.829
WeakWashF         267.219
SteamHeatF-3        6.903
T-Top-Chips-4       2.044
SulphidityL-4      30.420
dtype: float64
```

Untitled.ipynb

Code

Notebook Python (Pyodide)

[26]:
```
lower=Q1-1.5*IQR
upper=Q2+1.5*IQR
data2=data2[~((data2<lower) | (data2>upper)).any(axis=1)]
data2
```

[26]:

	Y-Kappa	ChipRate	BF-CMratio	BlowFlow	ChipLevel4	T-upperExt-2	T-lowerExt-2	UCZAA	WhiteFlow-4	AAWhiteSt-4	...	SteamFlow-4	Lower-HeatT-3	Upper-HeatT-3	ChipMass-4
1	27.60	16.810	79.022	1328.360	341.327	351.050	329.067	1.549	537.201	6.076	...	60.012	330.823	304.879	163.202
2	23.19	16.709	79.562	1329.407	239.161	350.022	329.260	1.600	549.611	0.000	...	61.304	329.140	303.383	164.013
3	23.60	16.478	81.011	1334.877	213.527	350.938	331.142	1.604	623.362	6.054	...	68.496	328.875	302.254	181.487
5	14.23	15.350	85.518	1171.604	198.538	344.014	325.195	1.436	628.245	6.020	...	65.225	322.103	298.517	165.814
6	13.49	13.700	98.186	1243.688	116.275	346.208	326.982	1.434	696.766	0.000	...	72.989	322.982	296.080	182.018
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
276	22.70	15.517	83.008	1288.010	306.886	350.155	322.485	1.590	568.752	6.170	...	67.678	331.854	309.346	160.061
296	20.50	13.358	97.662	1304.597	377.678	347.672	313.147	1.546	496.460	6.340	...	60.119	332.615	308.575	141.076
297	20.40	14.233	89.790	1278.006	379.458	354.290	315.558	1.515	491.374	0.000	...	60.424	331.980	308.078	140.301
298	20.90	15.167	84.640	1283.706	339.440	354.803	311.041	1.635	532.419	6.340	...	65.561	332.924	307.626	145.299
307	20.89	14.308	94.172	1327.832	251.120	351.263	332.485	1.522	631.514	0.000	...	71.286	328.699	300.706	180.229