In [2]:

```python
import pandas as pd
import numpy as np
df=pd.read_csv("Data Cleaning and Preprocessing.csv")
df
```

Out[2]:

	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	WeakLiquorF	BlackFlow-2	WeakWashF	SteamHeatF-3	T-Top-Chips-4	SulphidityL-4
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	1127.197	1319.039	257.325	54.612	252.077	NaN
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	665.975	1297.317	241.182	46.603	251.406	29.11
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	677.534	1327.072	237.272	51.795	251.335	NaN
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	767.853	1324.461	239.478	54.846	250.312	29.02
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	888.448	1343.424	215.372	54.186	249.916	29.01
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
319	10-16:00	23.75	12.667	93.450	1178.252	276.955	347.286	310.970	1.523	513.956	...	61.141	330.117	304.006	148.174	1027.201	1357.271	381.643	45.264	252.947	30.86
320	9-19:00	19.80	12.558	94.352	1184.119	297.071	399.135	319.576	1.451	570.058	...	67.667	330.848	304.616	165.178	906.962	1311.177	25.494	50.528	252.092	30.70
321	9-20:00	23.01	12.550	90.842	1188.517	289.826	373.633	314.591	1.457	549.306	...	66.446	330.226	304.686	160.841	887.125	1319.226	0.638	45.549	252.438	NaN
322	9-21:00	24.32	13.083	88.910	1192.879	318.006	364.081	308.559	1.523	504.852	...	61.054	327.346	304.363	147.589	804.423	1320.225	0.000	43.725	253.176	31.13
323	9-22:00	25.75	13.417	85.451	1186.342	248.312	356.289	310.482	1.474	497.375	...	58.247	328.092	304.093	144.218	828.328	1320.848	1.276	43.840	253.216	NaN

324 rows × 23 columns

In [5]:

```python
#checking the info of dataset
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 324 entries, 0 to 323
Data columns (total 23 columns):
 #   Column          Non-Null Count  Dtype  
---  ------          --------------  -----  
 0   Observation     324 non-null    object 
 1   Y-Kappa         324 non-null    float64
 2   ChipRate        319 non-null    float64
 3   BF-CMratio      307 non-null    float64
 4   BlowFlow        308 non-null    float64
 5   ChipLevel4      323 non-null    float64
 6   T-upperExt-2    322 non-null    float64
 7   T-lowerExt-2    322 non-null    float64
 8   UCZAA           299 non-null    float64
 9   WhiteFlow-4     323 non-null    float64
 10  AAWhiteSt-4     173 non-null    float64
 11  AA-Wood-4       323 non-null    float64
 12  ChipMoisture-4  323 non-null    float64
 13  SteamFlow-4     323 non-null    float64
 14  Lower-HeatT-3   322 non-null    float64
 15  Upper-HeatT-3   322 non-null    float64
 16  ChipMass-4      323 non-null    float64
 17  WeakLiquorF     323 non-null    float64
 18  BlackFlow-2     322 non-null    float64
 19  WeakWashF       323 non-null    float64
 20  SteamHeatF-3    322 non-null    float64
 21  T-Top-Chips-4   323 non-null    float64
 22  SulphidityL-4   173 non-null    float64
dtypes: float64(22), object(1)
memory usage: 58.3+ KB
```

In [6]:

```python
df.describe()
```

Out[6]:

	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	WeakLiquorF	BlackFlow-2	WeakWashF	SteamHeatF-3	T-Top-Chips-4	SulphidityL-4
count	324.000000	319.000000	307.000000	308.000000	323.000000	322.000000	322.000000	299.000000	323.000000	173.000000	...	323.000000	322.000000	322.000000	323.000000	323.000000	322.000000	323.000000	322.000000	323.000000	173.000000
mean	20.635370	14.347937	87.464456	1237.837614	258.164483	356.904295	324.020180	1.492010	591.732260	6.140410	...	66.668285	325.567820	300.525699	162.222322	873.828941	1175.917016	263.543068	49.696907	251.240087	30.411671
std	3.070036	1.499095	7.995012	100.593735	87.987452	9.209290	7.621402	0.105923	67.016351	0.081609	...	5.708587	4.609862	4.568484	14.160688	122.073521	149.334010	163.666942	4.551909	1.283432	0.701317
min	12.170000	9.983000	68.645000	0.000000	0.000000	339.168000	284.633000	1.182000	405.111000	5.890000	...	48.568000	318.051000	293.312000	113.922000	486.938000	838.948000	0.000000	35.510000	248.359000	29.010000
25%	18.382500	13.358000	81.823000	1193.215250	213.527000	350.241250	321.420000	1.431500	540.989500	6.089000	...	62.518000	321.385500	296.513250	153.032500	792.019500	1044.817500	134.649000	46.389750	250.312000	29.970000
50%	20.845000	14.308000	86.739000	1273.138500	271.792000	356.843000	325.669000	1.498000	592.895000	6.135000	...	67.429000	324.741000	299.126000	163.690000	865.254000	1150.221500	269.193000	50.277000	251.380000	30.370000
75%	23.032500	15.517000	92.372000	1289.196000	321.680000	362.242250	329.175000	1.560500	639.480500	6.199000	...	71.522000	329.845250	304.244750	172.555000	965.286500	1319.021250	405.563000	53.294250	252.323500	30.820000
max	27.600000	16.958000	121.717000	1351.240000	419.014000	399.135000	337.012000	1.747000	731.394000	6.340000	...	76.147000	333.854000	311.146000	189.268000	1226.277000	1395.767000	715.715000	63.332000	254.122000	32.840000

8 rows × 22 columns

In [7]:

```python
df=df.drop_duplicates()
df
```

Out[7]:

	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	WeakLiquorF	BlackFlow-2	WeakWashF	SteamHeatF-3	T-Top-Chips-4	SulphidityL-4
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	1127.197	1319.039	257.325	54.612	252.077	NaN
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	665.975	1297.317	241.182	46.603	251.406	29.11
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	677.534	1327.072	237.272	51.795	251.335	NaN
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	767.853	1324.461	239.478	54.846	250.312	29.02
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	888.448	1343.424	215.372	54.186	249.916	29.01
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
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	832.906	1344.708	388.911	49.524	251.833	30.29
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	905.639	1344.469	418.979	48.135	251.614	30.47
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	908.691	1344.588	462.712	54.373	251.197	NaN
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	986.206	1348.747	457.313	53.194	251.324	30.46
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	903.605	1323.082	232.729	54.503	250.084	NaN

301 rows × 23 columns

In [8]:

```python
df.isnull()
```

Out[8]:

	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	WeakLiquorF	BlackFlow-2	WeakWashF	SteamHeatF-3	T-Top-Chips-4	SulphidityL-4
0	False	False	False	False	False	False	False	False	False	False	...	False	False	False	False	False	False	False	False	False	True
1	False	False	False	False	False	False	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	False	False	False	False	False	True
3	False	False	False	False	False	False	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	False	False	False	False	False	False
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
298	False	False	False	False	False	False	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	False	False	False	False	False	False
300	False	False	True	False	False	False	False	False	True	False	...	False	False	False	False	False	False	False	False	False	True
301	False	False	True	False	False	False	False	False	True	False	...	False	False	False	False	False	False	False	False	False	False
307	False	False	False	False	False	False	False	False	False	False	...	False	False	False	False	False	False	False	False	False	True

301 rows × 23 columns

In [9]:

```python
# Check for missing values
df.isnull().sum()
```

Out[9]:

```
Observation         0
Y-Kappa             0
ChipRate            4
BF-CMratio         14
BlowFlow           13
ChipLevel4          1
T-upperExt-2        1
T-lowerExt-2        1
UCZAA              24
WhiteFlow-4         1
AAWhiteSt-4       141
AA-Wood-4           1
ChipMoisture-4      1
SteamFlow-4         1
Lower-HeatT-3       1
Upper-HeatT-3       1
ChipMass-4          1
WeakLiquorF         1
BlackFlow-2         1
WeakWashF           1
SteamHeatF-3        1
T-Top-Chips-4       1
SulphidityL-4     141
dtype: int64
```

In [10]:

```python
#filling null values using 0
df1=df.fillna(value=0)
df1
```

Out[10]:

	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	WeakLiquorF	BlackFlow-2	WeakWashF	SteamHeatF-3	T-Top-Chips-4	SulphidityL-4
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	1127.197	1319.039	257.325	54.612	252.077	0.00
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	665.975	1297.317	241.182	46.603	251.406	29.11
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	677.534	1327.072	237.272	51.795	251.335	0.00
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	767.853	1324.461	239.478	54.846	250.312	29.02
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	888.448	1343.424	215.372	54.186	249.916	29.01
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
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	832.906	1344.708	388.911	49.524	251.833	30.29
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	905.639	1344.469	418.979	48.135	251.614	30.47
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	908.691	1344.588	462.712	54.373	251.197	0.00
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	986.206	1348.747	457.313	53.194	251.324	30.46
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	903.605	1323.082	232.729	54.503	250.084	0.00

301 rows × 23 columns

In [12]:

```python
#checking if any null values are there or not
df2=df.fillna(method='pad')
df2
```

```
C:\Users\mvenu\AppData\Local\Temp\ipykernel_23208\3249974619.py:2: FutureWarning: DataFrame.fillna with 'method' is deprecated and will raise in a future version. Use obj.ffill() or obj.bfill() instead.
  df2=df.fillna(method='pad')
```

Out[12]:

	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	WeakLiquorF	BlackFlow-2	WeakWashF	SteamHeatF-3	T-Top-Chips-4	SulphidityL-4
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	1127.197	1319.039	257.325	54.612	252.077	NaN
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	665.975	1297.317	241.182	46.603	251.406	29.11
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	677.534	1327.072	237.272	51.795	251.335	29.11
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	767.853	1324.461	239.478	54.846	250.312	29.02
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	888.448	1343.424	215.372	54.186	249.916	29.01
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
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	832.906	1344.708	388.911	49.524	251.833	30.29
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	905.639	1344.469	418.979	48.135	251.614	30.47
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	908.691	1344.588	462.712	54.373	251.197	30.47
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	986.206	1348.747	457.313	53.194	251.324	30.46
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	903.605	1323.082	232.729	54.503	250.084	30.46

301 rows × 23 columns

In [13]:

```python
df3=df.fillna(method="bfill")
df3
```

```
C:\Users\mvenu\AppData\Local\Temp\ipykernel_23208\3092262244.py:1: FutureWarning: DataFrame.fillna with 'method' is deprecated and will raise in a future version. Use obj.ffill() or obj.bfill() instead.
  df3=df.fillna(method="bfill")
```

Out[13]:

	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	WeakLiquorF	BlackFlow-2	WeakWashF	SteamHeatF-3	T-Top-Chips-4	SulphidityL-4
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	1127.197	1319.039	257.325	54.612	252.077	29.11
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	665.975	1297.317	241.182	46.603	251.406	29.11
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	677.534	1327.072	237.272	51.795	251.335	29.02
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	767.853	1324.461	239.478	54.846	250.312	29.02
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	888.448	1343.424	215.372	54.186	249.916	29.01
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
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	832.906	1344.708	388.911	49.524	251.833	30.29
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	905.639	1344.469	418.979	48.135	251.614	30.47
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	908.691	1344.588	462.712	54.373	251.197	30.46
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	986.206	1348.747	457.313	53.194	251.324	30.46
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	903.605	1323.082	232.729	54.503	250.084	NaN

301 rows × 23 columns

In [14]:

```python
from scipy import stats
```

In [15]:

```python
df2.columns
```

Out[15]:

```
Index(['Observation', 'Y-Kappa', 'ChipRate', 'BF-CMratio', 'BlowFlow',
       'ChipLevel4 ', 'T-upperExt-2 ', 'T-lowerExt-2  ', 'UCZAA',
       'WhiteFlow-4 ', 'AAWhiteSt-4 ', 'AA-Wood-4  ', 'ChipMoisture-4 ',
       'SteamFlow-4 ', 'Lower-HeatT-3', 'Upper-HeatT-3 ', 'ChipMass-4 ',
       'WeakLiquorF ', 'BlackFlow-2 ', 'WeakWashF ', 'SteamHeatF-3 ',
       'T-Top-Chips-4 ', 'SulphidityL-4 '],
      dtype='object')
```

In [20]:

```python
df2.drop(["Observation"], axis=1, inplace=True)
df2.columns
```

Out[20]:

```
Index(['Y-Kappa', 'ChipRate', 'BF-CMratio', 'BlowFlow', 'ChipLevel4 ',
       'T-upperExt-2 ', 'T-lowerExt-2  ', 'UCZAA', 'WhiteFlow-4 ',
       'AAWhiteSt-4 ', 'AA-Wood-4  ', 'ChipMoisture-4 ', 'SteamFlow-4 ',
       'Lower-HeatT-3', 'Upper-HeatT-3 ', 'ChipMass-4 ', 'WeakLiquorF ',
       'BlackFlow-2 ', 'WeakWashF ', 'SteamHeatF-3 ', 'T-Top-Chips-4 ',
       'SulphidityL-4 '],
      dtype='object')
```

In [21]:

```python
Q1= df2.quantile(0.25)
Q3 = df2.quantile(0.75)
IQR=Q3-Q1
print(IQR)
```

```
Y-Kappa             4.55000
ChipRate            2.12500
BF-CMratio         10.13800
BlowFlow           96.76600
ChipLevel4        105.86800
T-upperExt-2       12.05500
T-lowerExt-2        7.58900
UCZAA               0.13900
WhiteFlow-4       100.09800
AAWhiteSt-4         0.10550
AA-Wood-4           1.47800
ChipMoisture-4      2.18600
SteamFlow-4         8.84000
Lower-HeatT-3       8.57600
Upper-HeatT-3       7.78500
ChipMass-4         19.34700
WeakLiquorF       180.61300
BlackFlow-2       280.07500
WeakWashF         263.33500
SteamHeatF-3        6.90100
T-Top-Chips-4       2.02900
SulphidityL-4       0.69525
dtype: float64
```

In [22]:

```python
df2=df2[~((df2 < (Q1-1.5*IQR)) | (df2 > (Q3+1.5*IQR))).any(axis=1)]
df2
```

Out[22]:

	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	WeakLiquorF	BlackFlow-2	WeakWashF	SteamHeatF-3	T-Top-Chips-4	SulphidityL-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	665.975	1297.317	241.182	46.603	251.406	29.11
2	23.19	16.709	79.562	1329.407	239.161	350.022	329.260	1.600	549.611	6.076	...	61.304	329.140	303.383	164.013	677.534	1327.072	237.272	51.795	251.335	29.11
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	826.243	907.641	595.875	52.807	249.580	30.34
6	13.49	13.700	98.186	1243.688	116.275	346.208	326.982	1.434	696.766	6.020	...	72.989	322.982	296.080	182.018	784.281	929.527	201.272	58.118	248.741	30.34
7	22.65	14.100	91.887	1307.852	288.989	352.321	331.162	1.468	625.549	6.143	...	71.298	329.662	301.539	179.886	837.178	1315.111	234.047	53.805	249.971	29.22
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
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	832.906	1344.708	388.911	49.524	251.833	30.29
299	24.98	15.167	85.034	1278.345	368.564	357.723	321.387	1.635	520.365	6.220	...	65.729	332.523	307.169	151.544	905.639	1344.469	418.979	48.135	251.614	30.47
300	21.00	15.167	88.013	1307.722	278.842	357.438	323.757	1.635	553.070	6.220	...	65.795	331.263	306.400	157.954	908.691	1344.588	462.712	54.373	251.197	30.47
301	21.40	15.167	85.490	1255.986	273.484	361.365	322.689	1.635	590.199	6.230	...	71.456	333.032	308.732	174.069	986.206	1348.747	457.313	53.194	251.324	30.46
307	20.89	14.308	94.172	1327.832	251.120	351.263	332.485	1.522	631.514	6.230	...	71.286	328.699	300.706	180.229	903.605	1323.082	232.729	54.503	250.084	30.46

235 rows × 22 columns

In [24]:

```python
df2.describe()
```

Out[24]:

	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	WeakLiquorF	BlackFlow-2	WeakWashF	SteamHeatF-3	T-Top-Chips-4	SulphidityL-4
count	235.000000	235.000000	235.000000	235.000000	235.000000	235.000000	235.000000	235.000000	235.000000	235.000000	...	235.000000	235.000000	235.000000	235.000000	235.000000	235.000000	235.000000	235.000000	235.000000	235.000000
mean	20.675489	14.602409	86.364787	1253.884400	265.499540	356.919094	324.992796	1.488749	601.162217	6.142443	...	67.402953	325.031523	299.996872	164.141877	879.282800	1159.761268	268.771315	49.889251	251.196153	30.488136
std	2.944999	1.296108	6.833130	47.605312	69.787154	7.548497	5.565984	0.103720	60.329201	0.067374	...	4.982040	4.680400	4.547736	11.473545	116.360066	151.553026	164.891435	4.065122	1.218812	0.620493
min	12.480000	10.833000	68.645000	1084.083000	61.783000	340.222000	310.421000	1.224000	468.841000	5.940000	...	52.962000	318.051000	293.312000	133.878000	596.446000	838.948000	0.000000	38.283000	248.359000	29.110000
25%	18.465000	13.687500	81.314500	1213.838000	224.714000	350.851500	321.829000	1.431000	549.484000	6.097000	...	63.819000	321.215000	296.345000	156.110000	795.064000	1017.102500	142.538000	46.778000	250.325000	30.120000
50%	20.800000	14.692000	85.534000	1278.006000	271.899000	357.438000	325.863000	1.488000	599.914000	6.140000	...	67.901000	322.412000	297.726000	164.250000	887.125000	1130.605000	284.877000	50.253000	251.253000	30.450000
75%	23.005000	15.604000	91.729500	1288.466000	315.738500	361.605000	329.161000	1.558500	652.770000	6.194000	...	71.707500	329.787500	303.981000	172.555000	970.082000	1318.663500	413.789000	53.051500	252.232000	30.910000
max	27.600000	16.958000	105.911000	1351.240000	394.234000	375.047000	337.012000	1.710000	731.394000	6.340000	...	75.974000	333.854000	311.146000	189.268000	1122.110000	1395.767000	715.715000	59.564000	254.122000	31.796000

8 rows × 22 columns

In [ ]: