Task 2

March 27, 2024

```python
[1]: import pandas as pd
```

```python
[2]: data = pd.read_csv('C:\\Users\\hp\\Downloads\\01.Data Cleaning and␣
      ↪Preprocessing.csv')
```

```python
[3]: type(data)
```

```
[3]: pandas.core.frame.DataFrame
```

```python
[4]: data.info
```

```
[4]: <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
     ..            ...           ...    ...          ...  …          ...
     319       347.286       310.970  1.523      513.956  …       61.141
     320       399.135       319.576  1.451      570.058  …       67.667
     321       373.633       314.591  1.457      549.306  …       66.446
     322       364.081       308.559  1.523      504.852  …       61.054
     323       356.289       310.482  1.474      497.375  …       58.247
```

```
        Lower-HeatT-3  Upper-HeatT-3  ChipMass-4  WeakLiquorF  BlackFlow-2  \
0             329.432        303.099     175.964     1127.197     1319.039
1             330.823        304.879     163.202      665.975     1297.317
2             329.140        303.383     164.013      677.534     1327.072
3             328.875        302.254     181.487      767.853     1324.461
4             328.352        300.954     183.929      888.448     1343.424
..                ...            ...         ...          ...          ...
319           330.117        304.006     148.174     1027.201     1357.271
320           330.848        304.616     165.178      906.962     1311.177
321           330.226        304.686     160.841      887.125     1319.226
322           327.346        304.363     147.589      804.423     1320.225
323           328.092        304.093     144.218      828.328     1320.848

     WeakWashF  SteamHeatF-3  T-Top-Chips-4  SulphidityL-4
0      257.325        54.612        252.077            NaN
1      241.182        46.603        251.406          29.11
2      237.272        51.795        251.335            NaN
3      239.478        54.846        250.312          29.02
4      215.372        54.186        249.916          29.01
..         ...           ...            ...            ...
319    381.643        45.264        252.947          30.86
320     25.494        50.528        252.092          30.70
321      0.638        45.549        252.438            NaN
322      0.000        43.725        253.176          31.13
323      1.276        43.840        253.216            NaN

[324 rows x 23 columns]>
```

[5]: `data.shape`

[5]: (324, 23)

[6]: `data.describe()`

[6]:
```
            Y-Kappa    ChipRate   BF-CMratio     BlowFlow   ChipLevel4  \
count   324.000000  319.000000   307.000000   308.000000   323.000000
mean     20.635370   14.347937    87.464456  1237.837614   258.164483
std       3.070036    1.499095     7.995012   100.593735    87.987452
min      12.170000    9.983000    68.645000     0.000000     0.000000
25%      18.382500   13.358000    81.823000  1193.215250   213.527000
50%      20.845000   14.308000    86.739000  1273.138500   271.792000
75%      23.032500   15.517000    92.372000  1289.196000   321.680000
max      27.600000   16.958000   121.717000  1351.240000   419.014000

        T-upperExt-2  T-lowerExt-2        UCZAA  WhiteFlow-4  AAWhiteSt-4  \
count     322.000000    322.000000   299.000000   323.000000   173.000000
mean      356.904295    324.020180     1.492010   591.732260     6.140410
```

std		9.209290	7.621402	0.105923	67.016351	0.081609
min		339.168000	284.633000	1.182000	405.111000	5.890000
25%		350.241250	321.420000	1.431500	540.989500	6.089000
50%		356.843000	325.669000	1.498000	592.895000	6.135000
75%		362.242250	329.175000	1.560500	639.480500	6.199000
max		399.135000	337.012000	1.747000	731.394000	6.340000

		SteamFlow-4	Lower-HeatT-3	Upper-HeatT-3	ChipMass-4	\
count	...	323.000000	322.000000	322.000000	323.000000
mean	...	66.668285	325.567820	300.525699	162.222322
std	...	5.708587	4.609862	4.568484	14.160688
min	...	48.568000	318.051000	293.312000	113.922000
25%	...	62.518000	321.385500	296.513250	153.032500
50%	...	67.429000	324.741000	299.126000	163.690000
75%	...	71.522000	329.845250	304.244750	172.555000
max	...	76.147000	333.854000	311.146000	189.268000

	WeakLiquorF	BlackFlow-2	WeakWashF	SteamHeatF-3	T-Top-Chips-4	\
count	323.000000	322.000000	323.000000	322.000000	323.000000
mean	873.828941	1175.917016	263.543068	49.696907	251.240087
std	122.073521	149.334010	163.666942	4.551909	1.283432
min	486.938000	838.948000	0.000000	35.510000	248.359000
25%	792.019500	1044.817500	134.649000	46.389750	250.312000
50%	865.254000	1150.221500	269.193000	50.277000	251.380000
75%	965.286500	1319.021250	405.563000	53.294250	252.323500
max	1226.277000	1395.767000	715.715000	63.332000	254.122000

	SulphidityL-4
count	173.000000
mean	30.411671
std	0.701317
min	29.010000
25%	29.970000
50%	30.370000
75%	30.820000
max	32.840000

[8 rows x 22 columns]

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

```
[9]: data
```

```
[9]:    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
..	…	…	…	…	…	…
298	12-09:00	20.90	15.167	84.640	1283.706	339.440
299	12-10:00	24.98	NaN	85.034	1278.345	368.564
300	12-11:00	21.00	NaN	88.013	1307.722	278.842
301	12-12:00	21.40	NaN	85.490	1255.986	273.484
307	31-05:00	20.89	14.308	94.172	1327.832	251.120

	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	
..	…	…	…	…	…	…	
298	354.803	311.041	1.635	532.419	…	65.561	
299	357.723	321.387	NaN	520.365	…	65.729	
300	357.438	323.757	NaN	553.070	…	65.795	
301	361.365	322.689	NaN	590.199	…	71.456	
307	351.263	332.485	1.522	631.514	…	71.286	

	Lower-HeatT-3	Upper-HeatT-3	ChipMass-4	WeakLiquorF	BlackFlow-2	\
0	329.432	303.099	175.964	1127.197	1319.039	
1	330.823	304.879	163.202	665.975	1297.317	
2	329.140	303.383	164.013	677.534	1327.072	
3	328.875	302.254	181.487	767.853	1324.461	
4	328.352	300.954	183.929	888.448	1343.424	
..	…	…	…	…	…	
298	332.924	307.626	145.299	832.906	1344.708	
299	332.523	307.169	151.544	905.639	1344.469	
300	331.263	306.400	157.954	908.691	1344.588	
301	333.032	308.732	174.069	986.206	1348.747	
307	328.699	300.706	180.229	903.605	1323.082	

	WeakWashF	SteamHeatF-3	T-Top-Chips-4	SulphidityL-4
0	257.325	54.612	252.077	NaN
1	241.182	46.603	251.406	29.11
2	237.272	51.795	251.335	NaN
3	239.478	54.846	250.312	29.02
4	215.372	54.186	249.916	29.01
..	…	…	…	…
298	388.911	49.524	251.833	30.29
299	418.979	48.135	251.614	30.47
300	462.712	54.373	251.197	NaN
301	457.313	53.194	251.324	30.46
307	232.729	54.503	250.084	NaN

```
[301 rows x 23 columns]
```

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

```
[10]:      Observation  Y-Kappa  ChipRate  BF-CMratio  BlowFlow  ChipLevel4  \
     0          False    False     False       False     False       False
     1          False    False     False       False     False       False
     2          False    False     False       False     False       False
     3          False    False     False       False     False       False
     4          False    False     False       False     False       False
     ..           ...      ...       ...         ...       ...         ...
     298        False    False     False       False     False       False
     299        False    False      True       False     False       False
     300        False    False      True       False     False       False
     301        False    False      True       False     False       False
     307        False    False     False       False     False       False

          T-upperExt-2  T-lowerExt-2   UCZAA  WhiteFlow-4  ...  SteamFlow-4  \
     0           False         False   False        False  ...        False
     1           False         False   False        False  ...        False
     2           False         False   False        False  ...        False
     3           False         False   False        False  ...        False
     4           False         False    True        False  ...        False
     ..            ...           ...     ...          ...  ...          ...
     298         False         False   False        False  ...        False
     299         False         False    True        False  ...        False
     300         False         False    True        False  ...        False
     301         False         False    True        False  ...        False
     307         False         False   False        False  ...        False

          Lower-HeatT-3  Upper-HeatT-3  ChipMass-4  WeakLiquorF  BlackFlow-2  \
     0            False          False       False        False        False
     1            False          False       False        False        False
     2            False          False       False        False        False
     3            False          False       False        False        False
     4            False          False       False        False        False
     ..             ...            ...         ...          ...          ...
     298          False          False       False        False        False
     299          False          False       False        False        False
     300          False          False       False        False        False
     301          False          False       False        False        False
     307          False          False       False        False        False

          WeakWashF  SteamHeatF-3  T-Top-Chips-4  SulphidityL-4
     0        False         False          False           True
     1        False         False          False          False
```

```
2        False           False           False              True
3        False           False           False             False
4        False           False           False             False
..         …               …               …                 …
298      False           False           False             False
299      False           False           False             False
300      False           False           False              True
301      False           False           False             False
307      False           False           False              True

[301 rows x 23 columns]
```

[11]: `data.isnull().sum()`

[11]:
```
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
```

[12]: `data.notnull()`

[12]:
```
     Observation  Y-Kappa  ChipRate  BF-CMratio  BlowFlow  ChipLevel4  \
0           True     True      True        True      True        True
1           True     True      True        True      True        True
2           True     True      True        True      True        True
3           True     True      True        True      True        True
4           True     True      True        True      True        True
```

```
..             …          …          …           …          …           …
298          True       True       True        True       True        True
299          True       True       False       True       True        True
300          True       True       False       True       True        True
301          True       True       False       True       True        True
307          True       True       True        True       True        True

        T-upperExt-2  T-lowerExt-2   UCZAA  WhiteFlow-4   …  SteamFlow-4   \
0             True          True     True         True    …         True
1             True          True     True         True    …         True
2             True          True     True         True    …         True
3             True          True     True         True    …         True
4             True          True     False        True    …         True
..             …            …          …           …      …  …          …
298           True          True     True         True    …         True
299           True          True     False        True    …         True
300           True          True     False        True    …         True
301           True          True     False        True    …         True
307           True          True     True         True    …         True

        Lower-HeatT-3  Upper-HeatT-3  ChipMass-4  WeakLiquorF  BlackFlow-2   \
0             True          True         True         True         True
1             True          True         True         True         True
2             True          True         True         True         True
3             True          True         True         True         True
4             True          True         True         True         True
..             …            …            …            …            …
298           True          True         True         True         True
299           True          True         True         True         True
300           True          True         True         True         True
301           True          True         True         True         True
307           True          True         True         True         True

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

[301 rows x 23 columns]
```

```
[13]: data.isnull().sum().sum()
```

```
[13]: 352
```

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

```
[14]:      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
      ..          ...      ...       ...         ...       ...         ...
      298    12-09:00    20.90    15.167      84.640  1283.706     339.440
      299    12-10:00    24.98     0.000      85.034  1278.345     368.564
      300    12-11:00    21.00     0.000      88.013  1307.722     278.842
      301    12-12:00    21.40     0.000      85.490  1255.986     273.484
      307    31-05:00    20.89    14.308      94.172  1327.832     251.120

           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  0.000      638.672  …       70.022
      ..            ...           ...    ...          ...  …          ...
      298       354.803       311.041  1.635      532.419  …       65.561
      299       357.723       321.387  0.000      520.365  …       65.729
      300       357.438       323.757  0.000      553.070  …       65.795
      301       361.365       322.689  0.000      590.199  …       71.456
      307       351.263       332.485  1.522      631.514  …       71.286

           Lower-HeatT-3  Upper-HeatT-3  ChipMass-4  WeakLiquorF  BlackFlow-2  \
      0          329.432        303.099     175.964     1127.197     1319.039
      1          330.823        304.879     163.202      665.975     1297.317
      2          329.140        303.383     164.013      677.534     1327.072
      3          328.875        302.254     181.487      767.853     1324.461
      4          328.352        300.954     183.929      888.448     1343.424
      ..             ...            ...         ...          ...          ...
      298        332.924        307.626     145.299      832.906     1344.708
      299        332.523        307.169     151.544      905.639     1344.469
      300        331.263        306.400     157.954      908.691     1344.588
      301        333.032        308.732     174.069      986.206     1348.747
      307        328.699        300.706     180.229      903.605     1323.082

           WeakWashF  SteamHeatF-3  T-Top-Chips-4  SulphidityL-4
```

```
0    257.325        54.612        252.077            0.00
1    241.182        46.603        251.406           29.11
2    237.272        51.795        251.335            0.00
3    239.478        54.846        250.312           29.02
4    215.372        54.186        249.916           29.01
..      …             …             …                …
298  388.911        49.524        251.833           30.29
299  418.979        48.135        251.614           30.47
300  462.712        54.373        251.197            0.00
301  457.313        53.194        251.324           30.46
307  232.729        54.503        250.084            0.00

[301 rows x 23 columns]
```

[15]: `data2.isnull().sum().sum()`

[15]: 0

[16]: `data`

[16]:
```
     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
..      …            …         …           …          …           …
298  12-09:00      20.90    15.167       84.640   1283.706     339.440
299  12-10:00      24.98       NaN       85.034   1278.345     368.564
300  12-11:00      21.00       NaN       88.013   1307.722     278.842
301  12-12:00      21.40       NaN       85.490   1255.986     273.484
307  31-05:00      20.89    14.308       94.172   1327.832     251.120

     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
..          …             …       …           …     …  …        …
298      354.803       311.041  1.635      532.419  …       65.561
299      357.723       321.387    NaN      520.365  …       65.729
300      357.438       323.757    NaN      553.070  …       65.795
301      361.365       322.689    NaN      590.199  …       71.456
307      351.263       332.485  1.522      631.514  …       71.286

     Lower-HeatT-3  Upper-HeatT-3  ChipMass-4  WeakLiquorF  BlackFlow-2   \
```

```
0          329.432        303.099         175.964       1127.197        1319.039
1          330.823        304.879         163.202        665.975        1297.317
2          329.140        303.383         164.013        677.534        1327.072
3          328.875        302.254         181.487        767.853        1324.461
4          328.352        300.954         183.929        888.448        1343.424
..             ...            ...             ...            ...             ...
298        332.924        307.626         145.299        832.906        1344.708
299        332.523        307.169         151.544        905.639        1344.469
300        331.263        306.400         157.954        908.691        1344.588
301        333.032        308.732         174.069        986.206        1348.747
307        328.699        300.706         180.229        903.605        1323.082

        WeakWashF   SteamHeatF-3   T-Top-Chips-4   SulphidityL-4
0        257.325         54.612         252.077             NaN
1        241.182         46.603         251.406           29.11
2        237.272         51.795         251.335             NaN
3        239.478         54.846         250.312           29.02
4        215.372         54.186         249.916           29.01
..           ...            ...             ...             ...
298      388.911         49.524         251.833           30.29
299      418.979         48.135         251.614           30.47
300      462.712         54.373         251.197             NaN
301      457.313         53.194         251.324           30.46
307      232.729         54.503         250.084             NaN

[301 rows x 23 columns]
```

[17]:
```python
data2 = data.fillna(value=0)
```

[18]:
```python
data2
```

[18]:
```
     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
..          ...      ...       ...         ...        ...         ...
298    12-09:00    20.90    15.167      84.640   1283.706     339.440
299    12-10:00    24.98     0.000      85.034   1278.345     368.564
300    12-11:00    21.00     0.000      88.013   1307.722     278.842
301    12-12:00    21.40     0.000      85.490   1255.986     273.484
307    31-05:00    20.89    14.308      94.172   1327.832     251.120

     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	0.000	638.672	...	70.022
..	...	...	...	...	...	...
298	354.803	311.041	1.635	532.419	...	65.561
299	357.723	321.387	0.000	520.365	...	65.729
300	357.438	323.757	0.000	553.070	...	65.795
301	361.365	322.689	0.000	590.199	...	71.456
307	351.263	332.485	1.522	631.514	...	71.286

	Lower-HeatT-3	Upper-HeatT-3	ChipMass-4	WeakLiquorF	BlackFlow-2 \
0	329.432	303.099	175.964	1127.197	1319.039
1	330.823	304.879	163.202	665.975	1297.317
2	329.140	303.383	164.013	677.534	1327.072
3	328.875	302.254	181.487	767.853	1324.461
4	328.352	300.954	183.929	888.448	1343.424
..	...	...	...	...	...
298	332.924	307.626	145.299	832.906	1344.708
299	332.523	307.169	151.544	905.639	1344.469
300	331.263	306.400	157.954	908.691	1344.588
301	333.032	308.732	174.069	986.206	1348.747
307	328.699	300.706	180.229	903.605	1323.082

	WeakWashF	SteamHeatF-3	T-Top-Chips-4	SulphidityL-4
0	257.325	54.612	252.077	0.00
1	241.182	46.603	251.406	29.11
2	237.272	51.795	251.335	0.00
3	239.478	54.846	250.312	29.02
4	215.372	54.186	249.916	29.01
..	...	...	...	...
298	388.911	49.524	251.833	30.29
299	418.979	48.135	251.614	30.47
300	462.712	54.373	251.197	0.00
301	457.313	53.194	251.324	30.46
307	232.729	54.503	250.084	0.00

[301 rows x 23 columns]

```
[20]: data4 = data.fillna(method='ffill')

      data4
```

```
C:\Users\hp\AppData\Local\Temp\ipykernel_10640\2159090985.py:1: FutureWarning:
DataFrame.fillna with 'method' is deprecated and will raise in a future version.
Use obj.ffill() or obj.bfill() instead.
  data4 = data.fillna(method='ffill')
```

```
[20]:      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
      ..         …         …          …          …         …           …
      298   12-09:00     20.90    15.167      84.640   1283.706     339.440
      299   12-10:00     24.98    15.167      85.034   1278.345     368.564
      300   12-11:00     21.00    15.167      88.013   1307.722     278.842
      301   12-12:00     21.40    15.167      85.490   1255.986     273.484
      307   31-05:00     20.89    14.308      94.172   1327.832     251.120


            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     1.604     638.672   …      70.022
      ..          …             …         …          …       …        …
      298      354.803       311.041     1.635     532.419   …      65.561
      299      357.723       321.387     1.635     520.365   …      65.729
      300      357.438       323.757     1.635     553.070   …      65.795
      301      361.365       322.689     1.635     590.199   …      71.456
      307      351.263       332.485     1.522     631.514   …      71.286


            Lower-HeatT-3  Upper-HeatT-3  ChipMass-4  WeakLiquorF  BlackFlow-2  \
      0        329.432        303.099      175.964    1127.197      1319.039
      1        330.823        304.879      163.202     665.975      1297.317
      2        329.140        303.383      164.013     677.534      1327.072
      3        328.875        302.254      181.487     767.853      1324.461
      4        328.352        300.954      183.929     888.448      1343.424
      ..          …             …            …           …            …
      298      332.924        307.626      145.299     832.906      1344.708
      299      332.523        307.169      151.544     905.639      1344.469
      300      331.263        306.400      157.954     908.691      1344.588
      301      333.032        308.732      174.069     986.206      1348.747
      307      328.699        300.706      180.229     903.605      1323.082


            WeakWashF   SteamHeatF-3   T-Top-Chips-4  SulphidityL-4
      0      257.325        54.612        252.077            NaN
      1      241.182        46.603        251.406          29.11
      2      237.272        51.795        251.335          29.11
      3      239.478        54.846        250.312          29.02
      4      215.372        54.186        249.916          29.01
      ..        …             …              …              …
      298    388.911        49.524        251.833          30.29
```

12

```
299    418.979         48.135          251.614         30.47
300    462.712         54.373          251.197         30.47
301    457.313         53.194          251.324         30.46
307    232.729         54.503          250.084         30.46
```

[301 rows x 23 columns]

[22]:
```python
import numpy as np
import matplotlib.pyplot as plt
from scipy import stats
```

Matplotlib is building the font cache; this may take a moment.

[23]:
```python
data2.columns
```

[23]:
```
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')
```

[25]:
```python
data2.drop(['Observation'], axis=1,inplace=True)
data2.columns
```

[25]:
```
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')
```

[27]:
```python
Q1=data2.quantile(0.25)
Q3=data2.quantile(0.75)
IQR=Q3-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
```

[28]:
```python
data2=data2[~((data2<(Q1-1.5*IQR))|(data2>(Q3+1.5*IQR))).any(axis=1)]
data2
```

[28]:

	Y-Kappa	ChipRate	BF-CMratio	BlowFlow	ChipLevel4	T-upperExt-2	\
1	27.60	16.810	79.022	1328.360	341.327	351.050	
2	23.19	16.709	79.562	1329.407	239.161	350.022	
3	23.60	16.478	81.011	1334.877	213.527	350.938	
5	14.23	15.350	85.518	1171.604	198.538	344.014	
6	13.49	13.700	98.186	1243.688	116.275	346.208	
..	…	…	…	…	…	…	
276	22.70	15.517	83.008	1288.010	306.886	350.155	
296	20.50	13.358	97.662	1304.597	377.678	347.672	
297	20.40	14.233	89.790	1278.006	379.458	354.290	
298	20.90	15.167	84.640	1283.706	339.440	354.803	
307	20.89	14.308	94.172	1327.832	251.120	351.263	

	T-lowerExt-2	UCZAA	WhiteFlow-4	AAWhiteSt-4	…	SteamFlow-4	\
1	329.067	1.549	537.201	6.076	…	60.012	
2	329.260	1.600	549.611	0.000	…	61.304	
3	331.142	1.604	623.362	6.054	…	68.496	
5	325.195	1.436	628.245	6.020	…	65.225	
6	326.982	1.434	696.766	0.000	…	72.989	
..	…	…	…	…	…	…	
276	322.485	1.590	568.752	6.170	…	67.678	
296	313.147	1.546	496.460	6.340	…	60.119	
297	315.558	1.515	491.374	0.000	…	60.424	
298	311.041	1.635	532.419	6.340	…	65.561	
307	332.485	1.522	631.514	0.000	…	71.286	

	Lower-HeatT-3	Upper-HeatT-3	ChipMass-4	WeakLiquorF	BlackFlow-2	\
1	330.823	304.879	163.202	665.975	1297.317	
2	329.140	303.383	164.013	677.534	1327.072	
3	328.875	302.254	181.487	767.853	1324.461	

```
5       322.103        298.517        165.814        826.243        907.641
6       322.982        296.080        182.018        784.281        929.527
..          ...            ...            ...            ...            ...
276     331.854        309.346        160.061        910.013       1381.389
296     332.615        308.575        141.076        997.904       1334.703
297     331.980        308.078        140.301        975.016       1344.835
298     332.924        307.626        145.299        832.906       1344.708
307     328.699        300.706        180.229        903.605       1323.082


        WeakWashF   SteamHeatF-3   T-Top-Chips-4   SulphidityL-4
1        241.182         46.603         251.406           29.11
2        237.272         51.795         251.335            0.00
3        239.478         54.846         250.312           29.02
5        595.875         52.807         249.580           30.34
6        201.272         58.118         248.741            0.00
..           ...            ...            ...             ...
276      441.934         51.466         252.216           29.59
296      389.497         46.206         252.423           30.43
297      388.676         47.803         252.311            0.00
298      388.911         49.524         251.833           30.29
307      232.729         54.503         250.084            0.00

[226 rows x 22 columns]
```

[29]:
```python
import scipy
import sklearn
from sklearn import preprocessing
from sklearn.preprocessing import scale
```

[30]:
```python
data2.describe()
```

[30]:
```
             Y-Kappa      ChipRate    BF-CMratio       BlowFlow    ChipLevel4   \
count     226.000000    226.000000    226.000000    226.000000    226.000000
mean       20.690487     14.673491     85.882181   1255.288916    264.664912
std         2.982916      1.297369      7.033155     47.896055     74.345135
min        12.480000     10.833000     68.645000   1084.083000     61.783000
25%        18.457500     13.850000     80.984000   1221.926000    220.356000
50%        20.775000     14.729000     84.967000   1280.291500    270.965000
75%        23.010000     15.708000     91.178750   1289.254000    322.492000
max        27.600000     16.958000    108.104000   1351.240000    419.014000


        T-upperExt-2   T-lowerExt-2          UCZAA    WhiteFlow-4    AAWhiteSt-4   \
count     226.000000     226.000000     226.000000     226.000000     226.000000
mean      356.861681     325.341124       1.487146     603.242482       3.098164
std         7.466897       5.557537       0.108054      61.052197       3.078138
min       340.222000     310.421000       1.182000     468.841000       0.000000
25%       350.704250     322.355500       1.429000     549.611000       0.000000
```

```
50%        357.560500       326.508500       1.492000       602.508000       5.904500
75%        361.555000       329.264500       1.556000       653.358500       6.140000
max        375.047000       337.012000       1.712000       731.394000       6.340000

           …   SteamFlow-4    Lower-HeatT-3   Upper-HeatT-3    ChipMass-4    \
count      …    226.000000     226.000000      226.000000     226.000000
mean       …     67.545478     324.752212      299.655420     164.220102
std        …      4.914301       4.526481        4.383788      11.423269
min        …     52.962000     318.051000      293.312000     133.878000
25%        …     63.954000     321.179500      296.338500     156.091000
50%        …     68.147000     322.380000      297.636500     164.333000
75%        …     71.760750     329.575000      303.777000     172.555000
max        …     75.974000     333.223000      309.854000     189.268000

           WeakLiquorF    BlackFlow-2    WeakWashF    SteamHeatF-3    T-Top-Chips-4    \
count       226.000000     226.000000   226.000000     226.000000      226.000000
mean        874.123035    1149.895257   273.739403      49.810239      251.177779
std         120.259977     150.321416   163.452307       4.143153        1.221296
min         596.446000     838.948000     0.000000      38.283000      248.359000
25%         784.366750    1014.977000   149.331750      46.639750      250.290750
50%         866.170000    1126.513500   283.079500      50.128500      251.233000
75%         968.683250    1302.847000   414.599750      52.889250      252.240000
max        1132.181000    1392.868000   715.715000      59.564000      254.122000

           SulphidityL-4
count        226.000000
mean          15.391987
std           15.297984
min            0.000000
25%            0.000000
50%           29.065000
75%           30.437500
max           32.840000

[8 rows x 22 columns]
```

```python
data2.matrix=data2.values.reshape(-1,1)
scaled=preprocessing.MinMaxScaler(feature_range=(0,10))
scaled_data=scaled.fit_transform(data2)

data2
```

```
[33]:    Y-Kappa   ChipRate   BF-CMratio   BlowFlow   ChipLevel4   T-upperExt-2    \
     1     27.60     16.810      79.022    1328.360      341.327       351.050
     2     23.19     16.709      79.562    1329.407      239.161       350.022
     3     23.60     16.478      81.011    1334.877      213.527       350.938
     5     14.23     15.350      85.518    1171.604      198.538       344.014
```

6	13.49	13.700	98.186	1243.688	116.275	346.208
..	…	…	…	…	…	…
276	22.70	15.517	83.008	1288.010	306.886	350.155
296	20.50	13.358	97.662	1304.597	377.678	347.672
297	20.40	14.233	89.790	1278.006	379.458	354.290
298	20.90	15.167	84.640	1283.706	339.440	354.803
307	20.89	14.308	94.172	1327.832	251.120	351.263

	T-lowerExt-2	UCZAA	WhiteFlow-4	AAWhiteSt-4	…	SteamFlow-4 \
1	329.067	1.549	537.201	6.076	…	60.012
2	329.260	1.600	549.611	0.000	…	61.304
3	331.142	1.604	623.362	6.054	…	68.496
5	325.195	1.436	628.245	6.020	…	65.225
6	326.982	1.434	696.766	0.000	…	72.989
..	…	…	…	…	…	…
276	322.485	1.590	568.752	6.170	…	67.678
296	313.147	1.546	496.460	6.340	…	60.119
297	315.558	1.515	491.374	0.000	…	60.424
298	311.041	1.635	532.419	6.340	…	65.561
307	332.485	1.522	631.514	0.000	…	71.286

	Lower-HeatT-3	Upper-HeatT-3	ChipMass-4	WeakLiquorF	BlackFlow-2 \
1	330.823	304.879	163.202	665.975	1297.317
2	329.140	303.383	164.013	677.534	1327.072
3	328.875	302.254	181.487	767.853	1324.461
5	322.103	298.517	165.814	826.243	907.641
6	322.982	296.080	182.018	784.281	929.527
..	…	…	…	…	…
276	331.854	309.346	160.061	910.013	1381.389
296	332.615	308.575	141.076	997.904	1334.703
297	331.980	308.078	140.301	975.016	1344.835
298	332.924	307.626	145.299	832.906	1344.708
307	328.699	300.706	180.229	903.605	1323.082

	WeakWashF	SteamHeatF-3	T-Top-Chips-4	SulphidityL-4
1	241.182	46.603	251.406	29.11
2	237.272	51.795	251.335	0.00
3	239.478	54.846	250.312	29.02
5	595.875	52.807	249.580	30.34
6	201.272	58.118	248.741	0.00
..	…	…	…	…
276	441.934	51.466	252.216	29.59
296	389.497	46.206	252.423	30.43
297	388.676	47.803	252.311	0.00
298	388.911	49.524	251.833	30.29
307	232.729	54.503	250.084	0.00

```
[226 rows x 22 columns]
```

[ ]: