```
In [3]:  ▶  import pandas as pd
             df = pd.read_csv('C:\\Users\\jagad\\Downloads\\01.Data Cleaning and Preprocessing.csv')
             print("First few rows of the DataFrame:")
             print(df.head())
```

```
First few rows of the DataFrame:
  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

   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

   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

   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

[5 rows x 23 columns]
```

```
In [24]:   filtered_df = df[df['BlowFlow'] > 10]
           print("\nFiltered DataFrame:")
           print(filtered_df.head())
```

```
Filtered DataFrame:
  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

   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

   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

   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

[5 rows x 23 columns]
```

```
In [21]:   df_filled = df.fillna(3)
```

```
In [7]:  ▶  print(df_filled)
```

```
     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       0.000  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        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
..      ...        ...           ...          ...
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        0.00
322    0.000    43.725        253.176        31.13
323    1.276    43.840        253.216        0.00

[324 rows x 23 columns]
```

```
In [8]:  df_dropped = df.dropna()
         print(df_dropped)
```

```
     Observation  Y-Kappa  ChipRate  BF-CMratio  BlowFlow  ChipLevel4  \
1      31-01:00    27.60    16.810      79.022  1328.360     341.327
3      31-03:00    23.60    16.478      81.011  1334.877     213.527
5       1-08:00    14.23    15.350      85.518  1171.604     198.538
7      31-06:00    22.65    14.100      91.887  1307.852     288.989
9      31-08:00    24.70    13.850      96.208  1334.892     362.511
..          ...      ...       ...         ...       ...         ...
312    31-10:00    24.40    14.117      85.998  1330.104     394.234
317     4-16:00    17.80    16.625      78.367  1276.082     202.744
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
322     9-21:00    24.32    13.083      88.910  1192.879     318.006

     T-upperExt-2  T-lowerExt-2  UCZAA  WhiteFlow-4  ...  SteamFlow-4  \
1         351.050       329.067  1.549      537.201  ...       60.012
3         350.938       331.142  1.604      623.362  ...       68.496
5         344.014       325.195  1.436      628.245  ...       65.225
7         352.321       331.162  1.468      625.549  ...       71.298
9         352.372       327.358  1.515      553.172  ...       64.249
..            ...           ...    ...          ...  ...          ...
312       348.089       319.027  1.429      540.558  ...       62.179
317       360.127       329.266  1.488      698.486  ...       75.296
319       347.286       310.970  1.523      513.956  ...       61.141
320       399.135       319.576  1.451      570.058  ...       67.667
322       364.081       308.559  1.523      504.852  ...       61.054

     Lower-HeatT-3  Upper-HeatT-3  ChipMass-4  WeakLiquorF  BlackFlow-2  \
1          330.823        304.879     163.202      665.975     1297.317
3          328.875        302.254     181.487      767.853     1324.461
5          322.103        298.517     165.814      826.243      907.641
7          329.662        301.539     179.886      837.178     1315.111
9          332.264        305.419     166.120      909.810     1318.725
..             ...            ...         ...          ...          ...
312        329.831        302.652     163.258      827.107     1312.372
317        321.658        297.088     180.438     1017.333     1052.785
319        330.117        304.006     148.174     1027.201     1357.271
320        330.848        304.616     165.178      906.962     1311.177
322        327.346        304.363     147.589      804.423     1320.225

     WeakWashF  SteamHeatF-3  T-Top-Chips-4  SulphidityL-4
1      241.182        46.603        251.406         29.110
3      239.478        54.846        250.312         29.020
5      595.875        52.807        249.580         30.340
7      234.047        53.805        249.971         29.220
9      180.375        48.842        251.121         29.210
..         ...           ...            ...            ...
312    132.163        49.446        251.406         29.292
317    433.089        54.675        251.747         30.320
319    381.643        45.264        252.947         30.860
320     25.494        50.528        252.092         30.700
322      0.000        43.725        253.176         31.130

[141 rows x 23 columns]
```

```
In [9]:  ▶ summary_stats = df.describe()
            print("\nSummary statistics:")
            print(summary_stats)
```

Summary statistics:
```
              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]
```

```
In [25]:  ▶| column_stats = df['BlowFlow'].describe()
          print("\nSummary statistics for a specific column:")
          print(column_stats)
```

```
Summary statistics for a specific column:
count      308.000000
mean      1237.837614
std        100.593735
min          0.000000
25%       1193.215250
50%       1273.138500
75%       1289.196000
max       1351.240000
Name: BlowFlow, dtype: float64
```