TASK 2:

DATA MANIPULATION WITH PANDAS

1. Load CSV file

```
In [5]: | import pandas as pd
import numpy as np

In [6]: | df = pd.read_csv('01.Data Cleaning and Preprocessing.csv')
```

2. Filtering data based on conditions.

```
In [7]: M filtered_df = df[df['ChipRate'] > 16.5]
            filtered_df.head()
   Out[7]:
                 Observation Y- ChipRate BF- CMratio BF- CMratio BlowFlow ChipLevel4 upperExt- lowerExt- lowerExt- 2 UCZAA WhiteFlow- 4 ... SteamFlow- Lower- Upper- ChipMass 4 HeatT-3 HeatT-3
                   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
                    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
                                                                                                       549.611 ...
                    31-02:00 23.19 16.709 79.562 1329.407
                                                                 239.161 350.022 329.260
                                                                                             1.600
                                                                                                                      61.304 329.140 303.383
                                                                                                                                               164.013
                                                                                                       537.811 ...
                    31-13:00 24.05 16.675 77.025 1284.386
                                                                 246.336
                                                                                             1.538
                                                                                                                      58.979 329.560 302.146
                                                                                                                                               160.587
                                                                           350.317
                                                                                    329.240
                                                                                                       622.751 ...
                     2-08:00 22.53 16.783 76.396 1281.536
                                                                 269.496 351.640 324.093 1.556
                                                                                                                      67.107 319.652 296.317
                                                                                                                                               171.703
             5 rows × 23 columns
```

3. Handling missing values.

```
In [12]:  df_filled_with_zero = df.fillna(0)
             df_filled_with_unknown = df.fillna('Unknown')
             df_filled_with_zero.head(), df_filled_with_unknown.head()
   Out[12]: ( Observation Y-Kappa ChipRate BF-CMratio BlowFlow ChipLevel4
                   31-00:00
                              23.10
                                       16.520
                                                  121.717 1177.607
                                                                          169.805
                   31-01:00
                                        16.810
                                                    79.022 1328.360
                                                                          341.327
                               27.60
              1
              2
                   31-02:00
                               23.19
                                        16.709
                                                    79.562 1329.407
                                                                          239,161
                   31-03:00
                               23.60
                                        16.478
                                                    81.011 1334.877
                                                                          213.527
              3
                                                    93.244 1334.168
              4
                   31-04:00
                               22.90
                                        15.618
                                                                          243.131
                               T-lowerExt-2
                                                UCZAA WhiteFlow-4
                                                                          SteamFlow-4
                 T-upperExt-2
              0
                       358.282
                                       329.545
                                               1.443
                                                            599.253 ...
                                                                                67.122
                                                            537.201 ...
              1
                       351.050
                                       329.067
                                                1.549
                                                                                60.012
              2
                       350.022
                                       329.260
                                                1.600
                                                            549.611 ...
                                                                                61.304
                       350.938
                                                1.604
                                                            623.362
                                                                                68.496
              3
                                       331,142
                                                                     . . .
              4
                                                            638.672 ...
                                                                                70.022
                       351.640
                                       332.709
                                                0.000
                 Lower-HeatT-3 Upper-HeatT-3
                                                ChipMass-4
                                                             WeakLiquorF
                                                                           BlackFlow-2
              0
                       329.432
                                       303.099
                                                    175.964
                                                                 1127.197
                                                                               1319.039
                       330.823
                                       304.879
                                                    163.202
                                                                  665.975
                                                                               1297.317
              1
                                       303.383
                                                                               1327.072
              2
                       329.140
                                                    164.013
                                                                  677.534
              3
                       328.875
                                       302.254
                                                    181.487
                                                                  767.853
                                                                               1324.461
              4
                                       300.954
                                                    183.929
                                                                  888.448
                                                                               1343.424
                       328.352
                 WeakWashF
                             SteamHeatF-3 T-Top-Chips-4 SulphidityL-4
              a
                    257.325
                                    54.612
                                                   252.077
                                                                      0.00
              1
                    241.182
                                    46.603
                                                   251.406
                                                                     29.11
                                    51.795
              2
                    237.272
                                                   251.335
                                                                      0.00
              3
                    239.478
                                    54.846
                                                                     29.02
                                                   250.312
              4
                    215.372
                                    54.186
                                                   249,916
                                                                     29.01
              [5 rows x 23 columns],
                Observation Y-Kappa ChipRate BF-CMratio BlowFlow ChipLevel4
              0
                                               121.717 1177.607
                   31-00:00
                               23.10
                                       16.52
                                                                       169.805
              1
                   31-01:00
                               27.60
                                        16.81
                                                  79.022
                                                          1328.36
                                                                       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. Calculating summary statistics.

```
summary_stats
    Out[10]:
                                                                                                   UCZAA WhiteFlow- AAWhiteSt-
                                                                                                                             St- SteamFlow-
                                                BF-
                                                      BlowFlow ChipLevel4 T- T- T- upperExt-2 lowerExt-2
                                                                                                                                                  Low
                      Y-Kappa ChipRate
                                            CMratio
                                                                                                                                                  Heat1
              count 324.000000 319.000000 307.000000
                                                     308.000000 323.000000 322.000000 322.000000 299.000000 323.000000
                                                                                                                      173.000000 ... 323.000000 322.0000
                     20.635370 14.347937 87.464456 1237.837614 258.164483 356.904295 324.020180
                                                                                                 1,492010 591,732260
                                                                                                                       6.140410 ... 66.668285 325.5678
              mean
                std
                      3.070036
                               1.499095
                                         7.995012 100.593735 87.987452 9.209290 7.621402
                                                                                                 0.105923 67.016351
                                                                                                                       0.081609 ... 5.708587 4.6098
                     12.170000
                                9.983000 68.645000
                                                       0.000000
                                                                0.000000 339.168000 284.633000
                                                                                                 1.182000 405.111000
                                                                                                                        5.890000 ... 48.568000 318.0510
                                                                                                                       6.089000 ... 62.518000 321.3855
               25%
                    18.382500 13.358000 81.823000 1193.215250 213.527000 350.241250 321.420000
                                                                                                 1.431500 540.989500
               50%
                     20.845000 \quad 14.308000 \quad 86.739000 \quad 1273.138500 \quad 271.792000 \quad 356.843000 \quad 325.669000
                                                                                                 1.498000 592.895000
                                                                                                                        6.135000 ... 67.429000 324.7410
                                                                                                                       6.199000 ... 71.522000 329.8452
                     23.032500 15.517000 92.372000 1289.196000 321.680000 362.242250 329.175000
                                                                                                  1.560500 639.480500
                                                                                                                       6.340000 ... 76.147000 333.8540
                    27.600000 16.958000 121.717000 1351.240000 419.014000 399.135000 337.012000
                                                                                                  1.747000 731.394000
             8 rows × 22 columns
```