MAINFLOW INTERNSHIP
DATA ANALYSIS WITH PYTHON
TASK-2

```
import pandas as pd
data=pd.read_csv('C:\\Users\\glady\\OneDrive\\Desktop\\DS with Python\\Task_2.csv')
data2=data.fillna(value=0)
```

```
data=data.drop_duplicates()
data
    Observation  Y-Kappa  ...  T-Top-Chips-4  SulphidityL-4
0      31-00:00    23.10  ...        252.077            NaN
1      31-01:00    27.60  ...        251.406          29.11
2      31-02:00    23.19  ...        251.335            NaN
3      31-03:00    23.60  ...        250.312          29.02
4      31-04:00    22.90  ...        249.916          29.01
..          ...      ...  ...            ...            ...
298    12-09:00    20.90  ...        251.833          30.29
299    12-10:00    24.98  ...        251.614          30.47
300    12-11:00    21.00  ...        251.197            NaN
301    12-12:00    21.40  ...        251.324          30.46
307    31-05:00    20.89  ...        250.084            NaN

[301 rows x 23 columns]
data.isnull()
     Observation  Y-Kappa  ...  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]
```

```
data.isnull().sum()
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
data.notnull()
     Observation  Y-Kappa  ...  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]
data.isnull().sum().sum()
np.int64(352)
```

```
data2
    Observation  Y-Kappa  ...  T-Top-Chips-4  SulphidityL-4
0      31-00:00    23.10  ...        252.077           0.00
1      31-01:00    27.60  ...        251.406          29.11
2      31-02:00    23.19  ...        251.335           0.00
3      31-03:00    23.60  ...        250.312          29.02
4      31-04:00    22.90  ...        249.916          29.01
..          ...      ...  ...            ...            ...
319    10-16:00    23.75  ...        252.947          30.86
320     9-19:00    19.80  ...        252.092          30.70
321     9-20:00    23.01  ...        252.438           0.00
322     9-21:00    24.32  ...        253.176          31.13
323     9-22:00    25.75  ...        253.216           0.00

[324 rows x 23 columns]
```

```
data2.columns
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')
```