This task involve using pandas libary to manipulate Data

Load Csv file

- Operation like filtering Data based on condition
- Handling missing Value
- Calculating summary statistics

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

	Observation	Y-Kappa	ChipRate	BF-CMratio	BlowFlow	ChipLevel4	T-upperExt-2	T-lowerExt-2	UCZAA	WhiteFlow-4	...	SteamFlow-4	Lower-HeatT-3	Uppe Heat
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	303.0
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	304.8
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	303.3
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	302.2
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	300.9
...	...	...	...	...	...	...	...	...	...	...	...	...	...	
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	304.0
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	304.6
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	304.6
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	304.3
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	304.0

324 rows × 23 columns

- Using loc and iloc Function to TO Filter Multiple Data

```
condition1=df["ChipRate"]>16.5
filter_df=df.loc[condition1]
filter_df
```

	Observation	Y-Kappa	ChipRate	BF-CMratio	BlowFlow	ChipLevel4	T-upperExt-2	T-lowerExt-2	UCZAA	WhiteFlow-4	...	SteamFlow-4	Lower-HeatT-3	Uppe Heat
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	303.0
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	304.8
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	303.3
14	31-13:00	24.05	16.675	77.025	1284.386	246.336	350.317	329.240	1.538	537.811	...	58.979	329.560	302.1
57	2-08:00	22.53	16.783	76.396	1281.536	269.496	351.640	324.093	1.556	622.751	...	67.107	319.652	296.3
64	2-15:00	25.40	16.958	74.405	1230.776	295.522	350.216	322.740	1.471	540.151	...	52.962	318.913	295.8
65	2-16:00	23.50	16.542	80.093	1246.777	198.217	350.521	323.513	1.474	596.506	...	65.556	319.151	296.0
69	2-20:00	23.30	16.742	73.689	1233.677	303.943	352.993	323.533	1.464	624.176	...	68.114	321.016	297.4
71	2-22:00	21.50	16.817	78.969	1237.844	285.892	353.644	324.500	1.437	621.887	...	68.750	321.235	296.6
72	2-23:00	18.13	16.825	84.836	1246.387	259.527	355.515	325.985	NaN	662.577	...	72.407	321.904	296.1
73	3-00:00	18.30	16.833	91.341	1236.911	287.397	356.787	327.856	1.540	661.532	...	71.891	323.285	297.7
81	3-08:00	23.20	16.700	73.381	1225.454	288.327	353.400	325.761	1.532	546.814	...	60.071	319.983	295.6
99	4-02:00	23.50	16.867	76.937	1284.849	334.170	364.132	329.250	1.551	620.208	...	72.448	322.259	297.5
100	4-03:00	24.20	16.700	78.434	1287.619	344.674	362.799	328.935	1.553	606.119	...	71.263	321.226	297.3
110	4-13:00	23.70	16.518	78.458	1277.550	342.110	363.287	325.232	1.433	571.686	...	66.242	320.164	296.9
113	4-16:00	17.80	16.625	78.367	1276.082	202.744	360.127	329.266	1.488	698.486	...	75.296	321.658	297.0
169	7-00:00	24.16	16.683	76.753	1280.487	267.235	355.892	327.226	1.678	621.684	...	64.686	320.424	295.8
170	7-01:00	20.40	16.683	76.838	1254.380	199.766	356.899	328.609	1.621	659.068	...	67.931	320.682	295.2
302	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	303.0
303	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	304.8
304	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	303.3
317	4-16:00	17.80	16.625	78.367	1276.082	202.744	360.127	329.266	1.488	698.486	...	75.296	321.658	297.0

22 rows × 23 columns

```
condition2=df["Y-Kappa"]>20
filter_df1=df.loc[condition2]
filter_df1
```

	Observation	Y-Kappa	ChipRate	BF-CMratio	BlowFlow	ChipLevel4	T-upperExt-2	T-lowerExt-2	UCZAA	WhiteFlow-4	...	SteamFlow-4	Lower-HeatT-3	Uppe Heat
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	303.0
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	304.8
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	303.3
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	302.2
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	300.9
...	...	...	...	...	...	...	...	...	...	...	...	...	...	
313	31-11:00	26.62	15.467	84.447	1334.255	386.971	349.392	321.021	1.428	531.250	...	59.407	330.284	303.2
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	304.0
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	304.6
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	304.3
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	304.0

195 rows × 23 columns

Using iloc function to filtering Data out of the given Csv file

1. Access the Data value from the csv file

```
df
```

	Observation	Y-Kappa	ChipRate	BF-CMratio	BlowFlow	ChipLevel4	T-upperExt-2	T-lowerExt-2	UCZAA	WhiteFlow-4	...	SteamFlow-4	Lower-HeatT-3	Uppe Heat
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	303.0
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	304.8
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	303.3
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	302.2
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	300.9
...	...	...	...	...	...	...	...	...	...	...	...	...	...	
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	304.0
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	304.6
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	304.6
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	304.3
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	304.0

324 rows × 23 columns

```
df_filter=df.iloc[0,3]
print("BF-CMratio   :",df_filter)
```

```
BF-CMratio      : 121.717
```

Handling Missing value checking Null data is there or not

```
df.isnull()
```

	Observation	Y-Kappa	ChipRate	BF-CMratio	BlowFlow	ChipLevel4	T-upperExt-2	T-lowerExt-2	UCZAA	WhiteFlow-4	...	SteamFlow-4	Lower-HeatT-3	Upper HeatT
0	False	False	False	False	False	False	False	False	False	False	...	False	False	Fals
1	False	False	False	False	False	False	False	False	False	False	...	False	False	Fals
2	False	False	False	False	False	False	False	False	False	False	...	False	False	Fals
3	False	False	False	False	False	False	False	False	False	False	...	False	False	Fals
4	False	False	False	False	False	False	False	False	True	False	...	False	False	Fals
...	...	...	...	...	...	...	...	...	...	...	...	...	...	.
319	False	False	False	False	False	False	False	False	False	False	...	False	False	Fals
320	False	False	False	False	False	False	False	False	False	False	...	False	False	Fals
321	False	False	False	False	False	False	False	False	False	False	...	False	False	Fals
322	False	False	False	False	False	False	False	False	False	False	...	False	False	Fals
323	False	False	False	False	False	False	False	False	False	False	...	False	False	Fals

324 rows × 23 columns

Checking the null value in particular columns

```
df.isnull().sum()
```

	0
Observation	0
Y-Kappa	0
ChipRate	5
BF-CMratio	17
BlowFlow	16
ChipLevel4	1
T-upperExt-2	2
T-lowerExt-2	2
UCZAA	25
WhiteFlow-4	1
AAWhiteSt-4	151
AA-Wood-4	1
ChipMoisture-4	1
SteamFlow-4	1
Lower-HeatT-3	2
Upper-HeatT-3	2
ChipMass-4	1
WeakLiquorF	1
BlackFlow-2	2
WeakWashF	1
SteamHeatF-3	2
T-Top-Chips-4	1
SulphidityL-4	151

dtype: int64

```
df.replace(to_replace=np.nan, value=0)
df
```

	Observation	Y-Kappa	ChipRate	BF-CMratio	BlowFlow	ChipLevel4	T-upperExt-2	T-lowerExt-2	UCZAA	WhiteFlow-4	...	SteamFlow-4	Lower-HeatT-3	Uppe Heat
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	303.0
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	304.8
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	303.3
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	302.2
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	300.9
...	...	...	...	...	...	...	...	...	...	...	...	...	...	
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	304.0
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	304.6
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	304.6
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	304.3
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	304.0

324 rows × 23 columns

```
df.isnull().sum()
```

	0
Observation	0
Y-Kappa	0
ChipRate	5
BF-CMratio	17
BlowFlow	16
ChipLevel4	1
T-upperExt-2	2
T-lowerExt-2	2
UCZAA	25
WhiteFlow-4	1
AAWhiteSt-4	151
AA-Wood-4	1
ChipMoisture-4	1
SteamFlow-4	1
Lower-HeatT-3	2
Upper-HeatT-3	2
ChipMass-4	1
WeakLiquorF	1
BlackFlow-2	2
WeakWashF	1
SteamHeatF-3	2
T-Top-Chips-4	1
SulphidityL-4	151

dtype: int64

Drop The Column which having NaN value

```
df.dropna(inplace=True)
df
```

	Observation	Y-Kappa	ChipRate	BF-CMratio	BlowFlow	ChipLevel4	T-upperExt-2	T-lowerExt-2	UCZAA	WhiteFlow-4	...	SteamFlow-4	Lower-HeatT-3	Uppe Heat
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	304.8
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	302.2
5	1-08:00	14.23	15.350	85.518	1171.604	198.538	344.014	325.195	1.436	628.245	...	65.225	322.103	298.5
7	31-06:00	22.65	14.100	91.887	1307.852	288.989	352.321	331.162	1.468	625.549	...	71.298	329.662	301.5
9	31-08:00	24.70	13.850	96.208	1334.892	362.511	352.372	327.358	1.515	553.172	...	64.249	332.264	305.4
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
312	31-10:00	24.40	14.117	85.998	1330.104	394.234	348.089	319.027	1.429	540.558	...	62.179	329.831	302.6
317	4-16:00	17.80	16.625	78.367	1276.082	202.744	360.127	329.266	1.488	698.486	...	75.296	321.658	297.0
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	304.0
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	304.6
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	304.3

141 rows × 23 columns

Summary statistics of Data Frame !!

```
df1=df.sum()
df1
```

	0
Observation	31-01:0031-03:001-08:0031-06:0031-08:0031-10:0...
Y-Kappa	2930.06
ChipRate	2031.426
BF-CMratio	12244.733
BlowFlow	175104.906
ChipLevel4	37673.834
T-upperExt-2	50314.694
T-lowerExt-2	45692.78
UCZAA	211.175
WhiteFlow-4	83750.881
AAWhiteSt-4	866.427
AA-Wood-4	2515.484
ChipMoisture-4	6590.415
SteamFlow-4	9457.234
Lower-HeatT-3	45893.343
Upper-HeatT-3	42359.938
ChipMass-4	22919.696
WeakLiquorF	121920.944
BlackFlow-2	164908.33