

Import labary

```
In [1]: import numpy as np  
import pandas as pd
```

Import dataset

```
In [2]: data=pd.read_csv(r"c:\Users\user\Downloads\bottle1.csv")  
data
```

Out[2]:

	Cst_Cnt	Btl_Cnt	Sta_ID	Depth_ID	Depthm	T_degC	Salnty	O2ml_L	STheta	O2Sat	...	R_NO2	R_NH4	R_CHLA	R_PHAEO
0	1	1	054.0 056.0	19-4903CR-HY-060-0930-05400560-0000A-3	0	10.50	33.440	NaN	25.649	NaN	...	NaN	NaN	NaN	NaN
1	1	2	054.0 056.0	19-4903CR-HY-060-0930-05400560-0008A-3	8	10.46	33.440	NaN	25.656	NaN	...	NaN	NaN	NaN	NaN
2	1	3	054.0 056.0	19-4903CR-HY-060-0930-05400560-0010A-7	10	10.46	33.437	NaN	25.654	NaN	...	NaN	NaN	NaN	NaN
3	1	4	054.0 056.0	19-4903CR-HY-060-0930-05400560-0019A-3	19	10.45	33.420	NaN	25.643	NaN	...	NaN	NaN	NaN	NaN
4	1	5	054.0 056.0	19-4903CR-HY-060-0930-05400560-0020A-7	20	10.45	33.421	NaN	25.643	NaN	...	NaN	NaN	NaN	NaN
...
5240	173	5241	069.0 144.0	19-4905CR-HY-125-1712-06901440-0279A-3	279	7.86	33.870	3.81	26.411	57.0	...	NaN	NaN	NaN	NaN

Cst_Cnt	Btl_Cnt	Sta_ID	Depth_ID	Depthm	T_degC	Salnty	O2ml_L	STheta	O2Sat	...	R_NO2	R_NH4	R_CHLA	R_PHAEO
5241	173	5242	069.0 144.0 19-4905CR-HY-125-1712-06901440-0300A-7	300	7.52	33.896	3.55	26.481	52.7	...	NaN	NaN	NaN	NaN
5242	173	5243	069.0 144.0 19-4905CR-HY-125-1712-06901440-0371A-3	371	6.54	33.930	2.67	26.642	38.8	...	NaN	NaN	NaN	NaN
5243	173	5244	069.0 144.0 19-4905CR-HY-125-1712-06901440-0400A-7	400	6.25	33.951	2.32	26.697	33.5	...	NaN	NaN	NaN	NaN
5244	173	5245	069.0 144.0 19-4905CR-HY-125-1712-06901440-0500A-7	500	5.44	34.027	1.30	26.858	18.4	...	NaN	NaN	NaN	NaN

5245 rows × 71 columns

Print head first 20 rows

In [3]: data.head(20)

			056.0	0930- 05400560- 0039A-3												
				19- 4903CR- HY-060- 0930- 05400560- 0050A-7												
7	1	8	054.0 056.0	50	10.24	33.424	NaN	25.682	NaN	...	NaN	NaN	NaN	NaN	NaN	NaN
				19- 4903CR- HY-060- 0930- 05400560- 0058A-3												
8	1	9	054.0 056.0	58	10.06	33.420	NaN	25.710	NaN	...	NaN	NaN	NaN	NaN	NaN	NaN
				19- 4903CR- HY-060- 0930- 05400560- 0075A-7												
9	1	10	054.0 056.0	75	9.86	33.494	NaN	25.801	NaN	...	NaN	NaN	NaN	NaN	NaN	NaN

Print tail last 7 rows

```
In [4]: data.tail(7)
```

Out[4]:

	Cst_Cnt	Btl_Cnt	Sta_ID	Depth_ID	Depthm	T_degC	Salnty	O2ml_L	STheta	O2Sat	...	R_NO2	R_NH4	R_CHLA	R_PHAEO
5238	173	5239	069.0 144.0	19- 4905CR- HY-125- 1712- 06901440- 0200A-7	200	9.44	33.602	4.60	25.957	71.2	...	NaN	NaN	NaN	NaN
5239	173	5240	069.0 144.0	19- 4905CR- HY-125- 1712- 06901440- 0250A-7	250	8.44	33.829	4.09	26.292	62.0	...	NaN	NaN	NaN	NaN
5240	173	5241	069.0 144.0	19- 4905CR- HY-125- 1712- 06901440- 0279A-3	279	7.86	33.870	3.81	26.411	57.0	...	NaN	NaN	NaN	NaN
5241	173	5242	069.0 144.0	19- 4905CR- HY-125- 1712- 06901440- 0300A-7	300	7.52	33.896	3.55	26.481	52.7	...	NaN	NaN	NaN	NaN
5242	173	5243	069.0 144.0	19- 4905CR- HY-125- 1712- 06901440- 0371A-3	371	6.54	33.930	2.67	26.642	38.8	...	NaN	NaN	NaN	NaN
5243	173	5244	069.0 144.0	19- 4905CR- HY-125- 1712- 06901440- 0400A-7	400	6.25	33.951	2.32	26.697	33.5	...	NaN	NaN	NaN	NaN
5244	173	5245	069.0 144.0	19- 4905CR- HY-125- 1712- 06901440- 0500A-7	500	5.44	34.027	1.30	26.858	18.4	...	NaN	NaN	NaN	NaN

7 rows × 71 columns

To print statistical data

```
In [5]: data.describe()
```

Out[5]:

	Cst_Cnt	Btl_Cnt	Depthm	T_degC	Salnty	O2ml_L	STheta	O2Sat	Oxy_μmol/Kg	BtlNum
count	5245.000000	5245.000000	5245.000000	5225.000000	5092.000000	3051.000000	5077.000000	2954.000000	2954.000000	0.0
mean	86.191992	2623.000000	345.400572	8.993144	33.827093	3.129839	26.119967	50.723798	137.281552	NaN
std	49.742962	1514.245412	356.298143	3.902631	0.492518	2.259226	0.923019	38.072596	98.237603	NaN
min	1.000000	1.000000	0.000000	2.700000	32.520000	0.210000	23.568000	3.100000	9.131386	NaN
25%	43.000000	1312.000000	54.000000	5.430000	33.490000	0.670000	25.117000	9.500000	29.570560	NaN
50%	86.000000	2623.000000	200.000000	8.620000	33.890000	2.990000	26.279000	46.150000	132.286550	NaN
75%	129.000000	3934.000000	600.000000	12.230000	34.255000	5.490000	26.998000	93.100000	240.036300	NaN
max	173.000000	5245.000000	1547.000000	19.760000	34.700000	6.630000	27.623000	112.900000	288.794800	NaN

8 rows × 69 columns



To print rows and coloum

```
In [6]: np.shape(data)
```

Out[6]: (5245, 71)

To print no. of elements

```
In [7]: np.size(data)
```

Out[7]: 372395

To print missing values


```
In [8]: data.isna()
```

Out[8]:

	Cst_Cnt	Btl_Cnt	Sta_ID	Depth_ID	Depthm	T_degC	Salnty	O2ml_L	STheta	O2Sat	...	R_NO2	R_NH4	R_CHLA	R_PHAEO	I
0	False	False	False	False	False	False	False	True	False	True	...	True	True	True	True	
1	False	False	False	False	False	False	False	True	False	True	...	True	True	True	True	
2	False	False	False	False	False	False	False	True	False	True	...	True	True	True	True	
3	False	False	False	False	False	False	False	True	False	True	...	True	True	True	True	
4	False	False	False	False	False	False	False	True	False	True	...	True	True	True	True	
...	
5240	False	False	False	False	False	False	False	False	False	False	...	True	True	True	True	
5241	False	False	False	False	False	False	False	False	False	False	...	True	True	True	True	
5242	False	False	False	False	False	False	False	False	False	False	...	True	True	True	True	
5243	False	False	False	False	False	False	False	False	False	False	...	True	True	True	True	
5244	False	False	False	False	False	False	False	False	False	False	...	True	True	True	True	

5245 rows × 71 columns



Filla a value 60 in missing place

```
In [9]: data.fillna(value=60)
```

Out[9]:

	Cst_Cnt	Btl_Cnt	Sta_ID	Depth_ID	Depthm	T_degC	Salnty	O2ml_L	STheta	O2Sat	...	R_NO2	R_NH4	R_CHLA	R_PHAEO
0	1	1	054.0 056.0	19- 4903CR- HY-060- 0930- 05400560- 0000A-3	0	10.50	33.440	60.00	25.649	60.0	...	60.0	60.0	60.0	60.0
1	1	2	054.0 056.0	19- 4903CR- HY-060- 0930- 05400560- 0008A-3	8	10.46	33.440	60.00	25.656	60.0	...	60.0	60.0	60.0	60.0
2	1	3	054.0 056.0	19- 4903CR- HY-060- 0930- 05400560- 0010A-7	10	10.46	33.437	60.00	25.654	60.0	...	60.0	60.0	60.0	60.0
3	1	4	054.0 056.0	19- 4903CR- HY-060- 0930- 05400560- 0019A-3	19	10.45	33.420	60.00	25.643	60.0	...	60.0	60.0	60.0	60.0
4	1	5	054.0 056.0	19- 4903CR- HY-060- 0930- 05400560- 0020A-7	20	10.45	33.421	60.00	25.643	60.0	...	60.0	60.0	60.0	60.0
...
5240	173	5241	069.0 144.0	19- 4905CR- HY-125- 1712- 06901440- 0279A-3	279	7.86	33.870	3.81	26.411	57.0	...	60.0	60.0	60.0	60.0

Cst_Cnt	Btl_Cnt	Sta_ID	Depth_ID	Depthm	T_degC	Salnty	O2ml_L	STheta	O2Sat	...	R_NO2	R_NH4	R_CHLA	R_PHAEO
5241	173	5242	069.0 144.0 19-4905CR-HY-125-1712-06901440-0300A-7	300	7.52	33.896	3.55	26.481	52.7	...	60.0	60.0	60.0	60.0
5242	173	5243	069.0 144.0 19-4905CR-HY-125-1712-06901440-0371A-3	371	6.54	33.930	2.67	26.642	38.8	...	60.0	60.0	60.0	60.0
5243	173	5244	069.0 144.0 19-4905CR-HY-125-1712-06901440-0400A-7	400	6.25	33.951	2.32	26.697	33.5	...	60.0	60.0	60.0	60.0
5244	173	5245	069.0 144.0 19-4905CR-HY-125-1712-06901440-0500A-7	500	5.44	34.027	1.30	26.858	18.4	...	60.0	60.0	60.0	60.0

5245 rows × 71 columns

```
In [10]: import matplotlib.pyplot as pp
```

```
In [11]: dd=data[['STheta', 'Salnty']]
dd
```

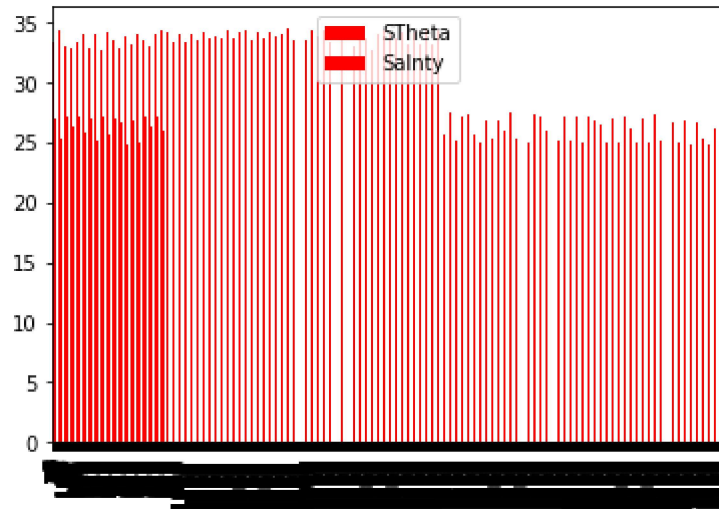
Out[11]:

	STheta	Salnty
0	25.649	33.440
1	25.656	33.440
2	25.654	33.437
3	25.643	33.420
4	25.643	33.421
...
5240	26.411	33.870
5241	26.481	33.896
5242	26.642	33.930
5243	26.697	33.951
5244	26.858	34.027

5245 rows × 2 columns

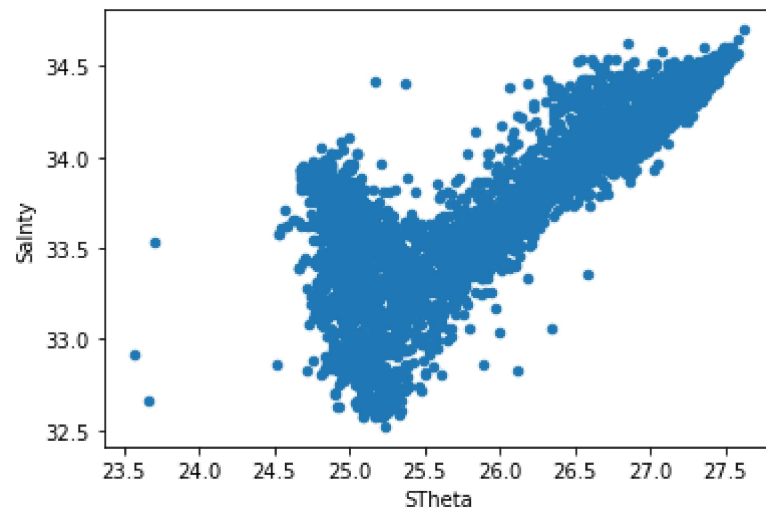
```
In [12]: dd.plot.bar(color='r')
```

```
Out[12]: <AxesSubplot:>
```



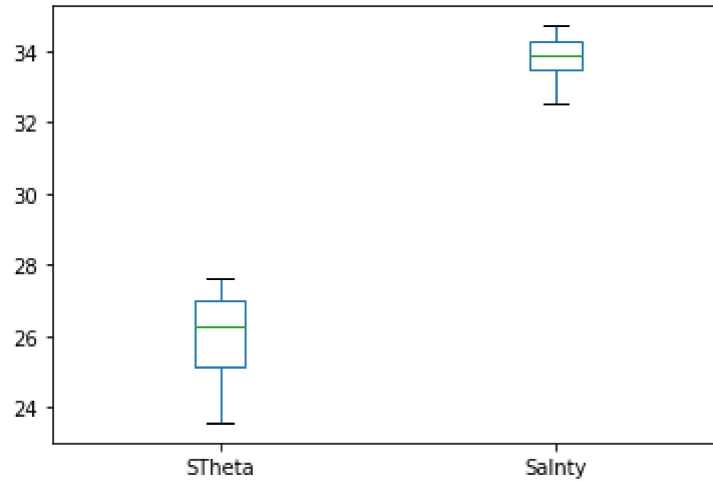
```
In [13]: dd.plot.scatter(x='STheta',y='Salnty')
```

```
Out[13]: <AxesSubplot:xlabel='STheta', ylabel='Salnty'>
```



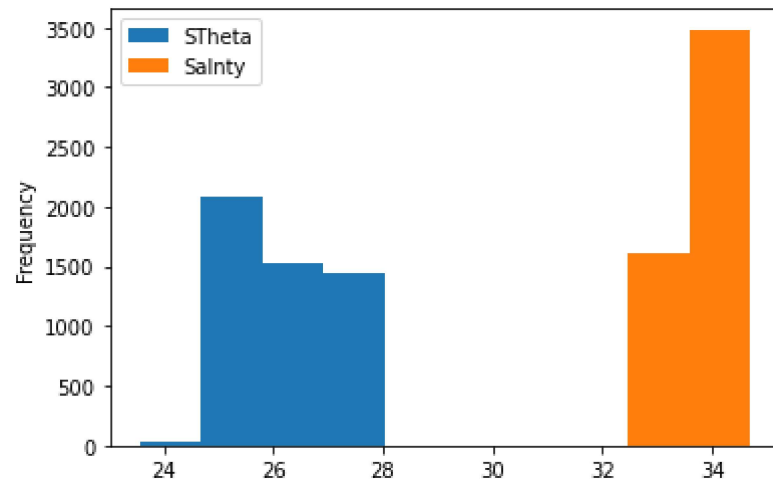
```
In [16]: dd.plot.box()
```

```
Out[16]: <AxesSubplot:>
```



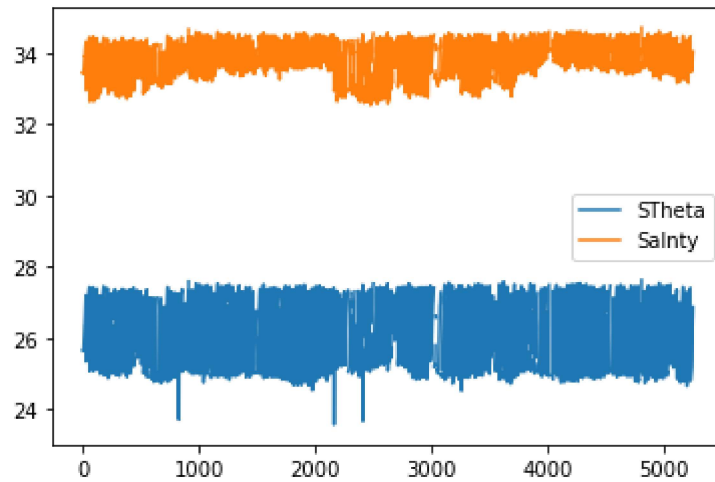
```
In [20]: dd.plot.hist()
```

```
Out[20]: <AxesSubplot:ylabel='Frequency'>
```



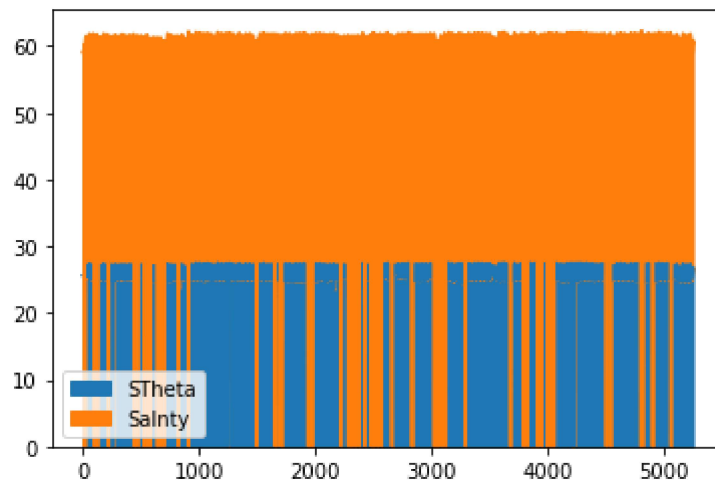
```
In [17]: dd.plot.line()
```

```
Out[17]: <AxesSubplot:>
```



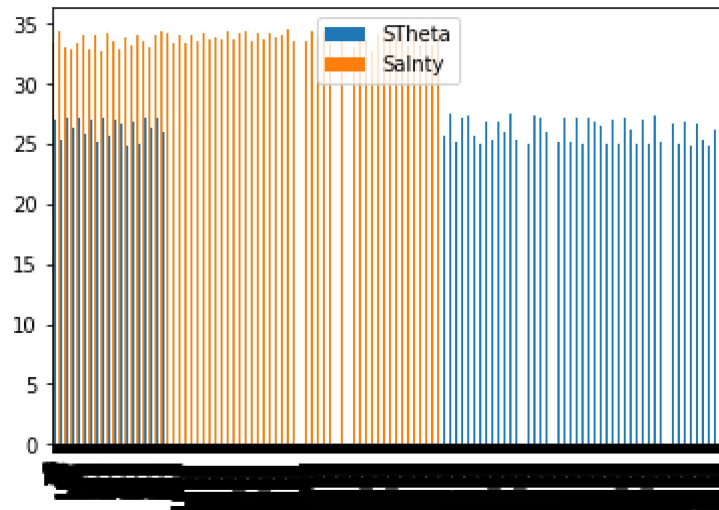
```
In [18]: dd.plot.area()
```

```
Out[18]: <AxesSubplot:>
```




```
In [19]: dd.plot.bar()
```

```
Out[19]: <AxesSubplot:>
```



```
In [ ]:
```