

# 1.IMPORTING LIBRARIES

In [1]:

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
```

# 2.Importing dataset

In [2]:

```
data=pd.read_csv(r"C:\Users\user\Downloads\bottle.csv")
data
```

C:\ProgramData\Anaconda3\lib\site-packages\IPython\core\interactiveshell.py:3165: DtypeWarning: Columns (47,73) have mixed types.Specify dtype option on import or set low\_memory=False.

```
has_raised = await self.run_ast_nodes(code_ast.body, cell_name,
```

Out[2]:

	Cst_Cnt	Btl_Cnt	Sta_ID	Depth_ID	Depthm	T_degC	Salnty	O2ml_L	STheta	O2Sat	...	R
0	1	1	054.0 056.0	19-4903CR-HY-060-0930-05400560-0000A-3	0	10.500	33.4400	NaN	25.64900	NaN	...	
1	1	2	054.0 056.0	19-4903CR-HY-060-0930-05400560-0008A-3	8	10.460	33.4400	NaN	25.65600	NaN	...	
2	1	3	054.0 056.0	19-4903CR-HY-060-0930-05400560-0010A-7	10	10.460	33.4370	NaN	25.65400	NaN	...	
3	1	4	054.0 056.0	19-4903CR-HY-060-0930-05400560-0019A-3	19	10.450	33.4200	NaN	25.64300	NaN	...	
4	1	5	054.0 056.0	19-4903CR-HY-060-0930-05400560-0020A-7	20	10.450	33.4210	NaN	25.64300	NaN	...	
...	...	...	...	...	...	...	...	...	...	...	...	

	Cst_Cnt	Btl_Cnt	Sta_ID	Depth_ID	Depthm	T_degC	Salnty	O2ml_L	STheta	O2Sat	...	R
864858	34404	864859	093.4 026.4	20- 1611SR- MX-310- 2239- 09340264- 0000A-7	0	18.744	33.4083	5.805	23.87055	108.74	...	
864859	34404	864860	093.4 026.4	20- 1611SR- MX-310- 2239- 09340264- 0002A-3	2	18.744	33.4083	5.805	23.87072	108.74	...	
864860	34404	864861	093.4 026.4	20- 1611SR- MX-310- 2239- 09340264- 0005A-3	5	18.692	33.4150	5.796	23.88911	108.46	...	
864861	34404	864862	093.4 026.4	20- 1611SR- MX-310- 2239- 09340264- 0010A-3	10	18.161	33.4062	5.816	24.01426	107.74	...	
864862	34404	864863	093.4 026.4	20- 1611SR- MX-310- 2239- 09340264- 0015A-3	15	17.533	33.3880	5.774	24.15297	105.66	...	

864863 rows × 74 columns

3.head

In [3]:

```
da=data.head(8)
da
```

Out[3]:

	Cst_Cnt	Btl_Cnt	Sta_ID	Depth_ID	Depthm	T_degC	Salnty	O2ml_L	STheta	O2Sat	...	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

Cst_Cnt	Btl_Cnt	Sta_ID	Depth_ID	Depthm	T_degC	Salnty	O2ml_L	STheta	O2Sat	...	R_PHAEO
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
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
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
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
5	1	6	054.0 056.0 19-4903CR-HY-060-0930-05400560-0030A-7	30	10.45	33.431	NaN	25.651	NaN	...	NaN
6	1	7	054.0 056.0 19-4903CR-HY-060-0930-05400560-0039A-3	39	10.45	33.440	NaN	25.658	NaN	...	NaN
7	1	8	054.0 056.0 19-4903CR-HY-060-0930-05400560-0050A-7	50	10.24	33.424	NaN	25.682	NaN	...	NaN

8 rows × 74 columns

4.tail

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

Out[4]:

	Cst_Cnt	Btl_Cnt	Sta_ID	Depth_ID	Depthm	T_degC	Salnty	O2ml_L	STheta	O2Sat	...	R
864856	34403	864857	093.3 120.0	20- 1611SR- MX-313- 2053- 09331200- 0500A-7	500	5.993	34.2160	0.456	26.94518	6.55	...	
864857	34403	864858	093.3 120.0	20- 1611SR- MX-313- 2053- 09331200- 0521A-3	521	5.818	34.2382	0.366	26.98477	5.23	...	
864858	34404	864859	093.4 026.4	20- 1611SR- MX-310- 2239- 09340264- 0000A-7	0	18.744	33.4083	5.805	23.87055	108.74	...	
864859	34404	864860	093.4 026.4	20- 1611SR- MX-310- 2239- 09340264- 0002A-3	2	18.744	33.4083	5.805	23.87072	108.74	...	
864860	34404	864861	093.4 026.4	20- 1611SR- MX-310- 2239- 09340264- 0005A-3	5	18.692	33.4150	5.796	23.88911	108.46	...	
864861	34404	864862	093.4 026.4	20- 1611SR- MX-310- 2239- 09340264- 0010A-3	10	18.161	33.4062	5.816	24.01426	107.74	...	
864862	34404	864863	093.4 026.4	20- 1611SR- MX-310- 2239- 09340264- 0015A-3	15	17.533	33.3880	5.774	24.15297	105.66	...	

7 rows × 74 columns



5.describe()

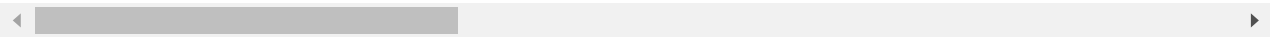
In [5]:

data.describe()

Out[5]:

	Cst_Cnt	Btl_Cnt	Depthm	T_degC	Salnty	O2ml_L	
count	864863.000000	864863.000000	864863.000000	853900.000000	817509.000000	696201.000000	81217
mean	17138.790958	432432.000000	226.831951	10.799677	33.840350	3.392468	2
std	10240.949817	249664.587267	316.050259	4.243825	0.461843	2.073256	
min	1.000000	1.000000	0.000000	1.440000	28.431000	-0.010000	2
25%	8269.000000	216216.500000	46.000000	7.680000	33.488000	1.360000	2
50%	16848.000000	432432.000000	125.000000	10.060000	33.863000	3.440000	2
75%	26557.000000	648647.500000	300.000000	13.880000	34.196900	5.500000	2
max	34404.000000	864863.000000	5351.000000	31.140000	37.034000	11.130000	25

8 rows × 70 columns



# 6.shape()

In [6]:

np.shape(data)

Out[6]: (864863, 74)

# 7.size()

In [7]:

np.size(data)

Out[7]: 63999862

# 8.isna()

In [8]:

data.isna()

Out[8]:

	Cst_Cnt	Btl_Cnt	Sta_ID	Depth_ID	Depthm	T_degC	Salnty	O2ml_L	STheta	O2Sat	...	R_PH
0	False	False	False	False	False	False	False	True	False	True	...	
1	False	False	False	False	False	False	False	True	False	True	...	
2	False	False	False	False	False	False	False	True	False	True	...	
3	False	False	False	False	False	False	False	True	False	True	...	

	Cst_Cnt	Btl_Cnt	Sta_ID	Depth_ID	Depthm	T_degC	Salnty	O2ml_L	STheta	O2Sat	...	R_PH
4	False	False	False	False	False	False	False	True	False	True	...	
...	...	...	...	...	...	...	...	...	...	...	...	
864858	False	False	False	False	False	False	False	False	False	False	...	
864859	False	False	False	False	False	False	False	False	False	False	...	
864860	False	False	False	False	False	False	False	False	False	False	...	
864861	False	False	False	False	False	False	False	False	False	False	...	
864862	False	False	False	False	False	False	False	False	False	False	...	

864863 rows × 74 columns

# 9.dropna

In [9]:

```
data.dropna()
```

Out[9]:

Cst_Cnt	Btl_Cnt	Sta_ID	Depth_ID	Depthm	T_degC	Salnty	O2ml_L	STheta	O2Sat	...	R_PHAEO	F
0 rows × 74 columns												

# 10.selecting specific column

In [10]:

```
d=data[["Sta_ID","Btl_Cnt"]]
d
```

Out[10]:

	Sta_ID	Btl_Cnt
0	054.0 056.0	1
1	054.0 056.0	2
2	054.0 056.0	3
3	054.0 056.0	4
4	054.0 056.0	5
...	...	...
864858	093.4 026.4	864859
864859	093.4 026.4	864860
864860	093.4 026.4	864861

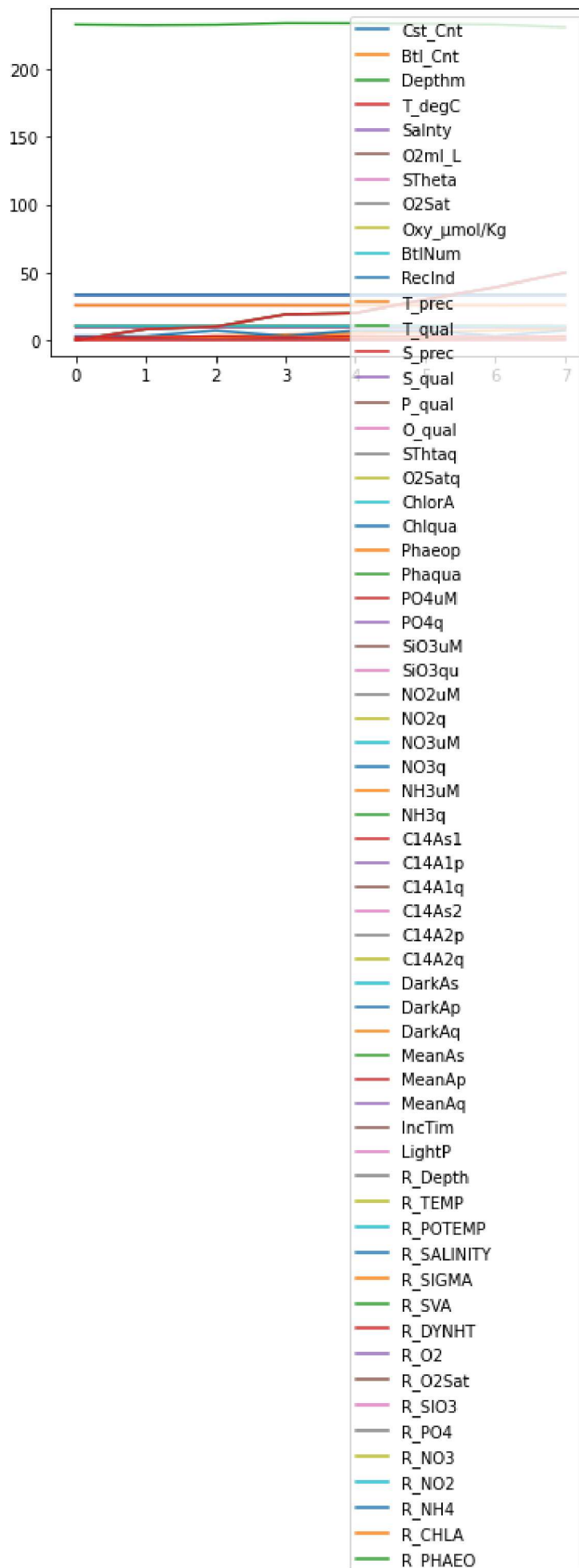
	Sta_ID	Btl_Cnt
<b>864861</b>	093.4 026.4	864862
<b>864862</b>	093.4 026.4	864863

864863 rows × 2 columns

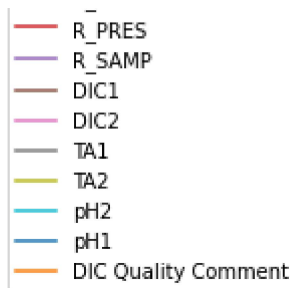
## 11.line plot

In [18]: `da.plot.line()`

Out[18]: <AxesSubplot:>





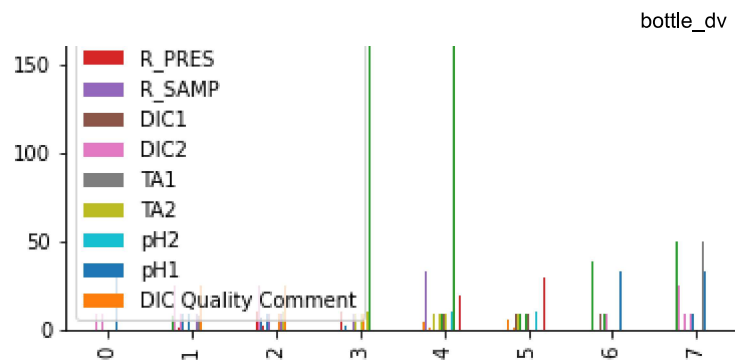


## 12.bar plot

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

Out[19]: <AxesSubplot:>

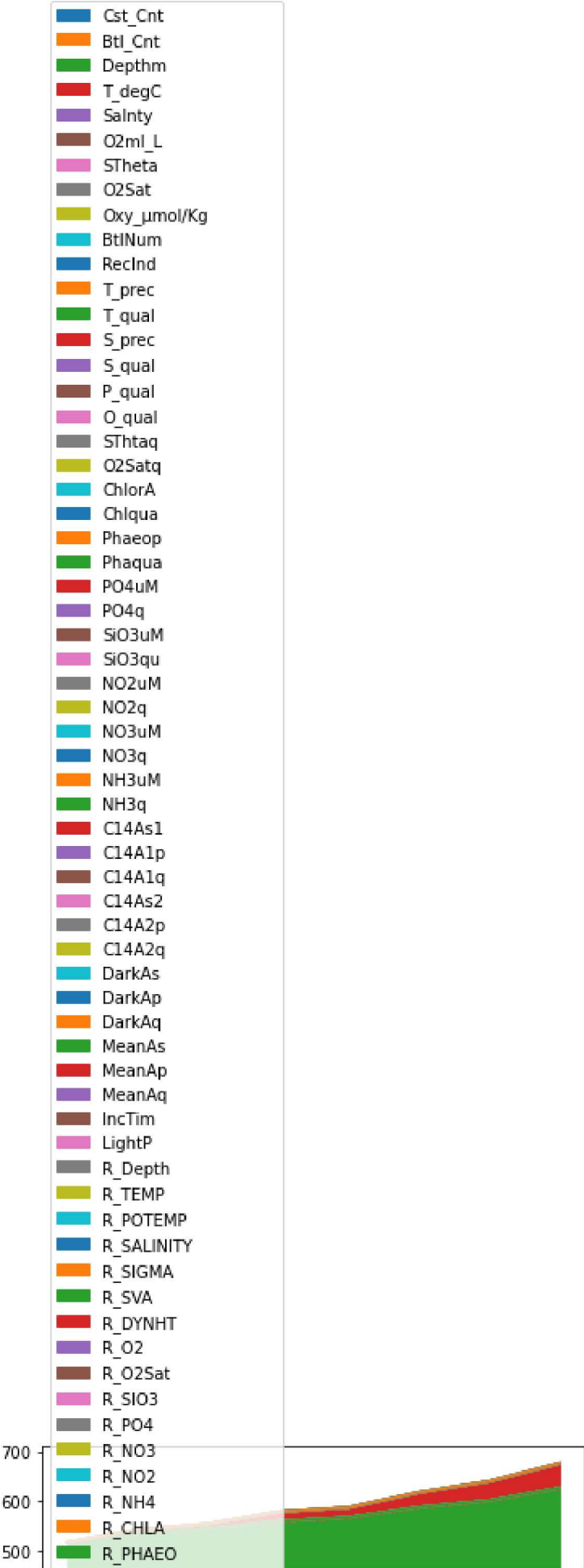


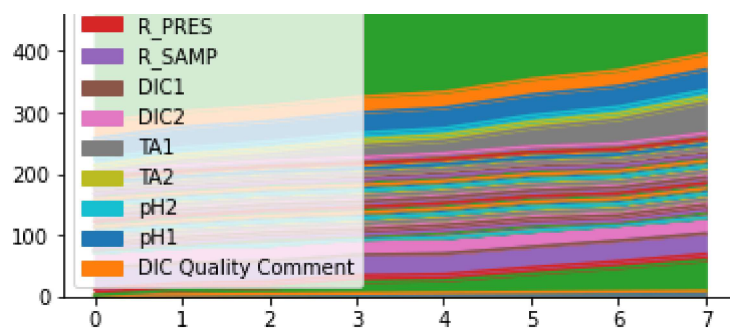


## 13.area plot

```
In [13]: da.plot.area()
```

```
Out[13]: <AxesSubplot:>
```

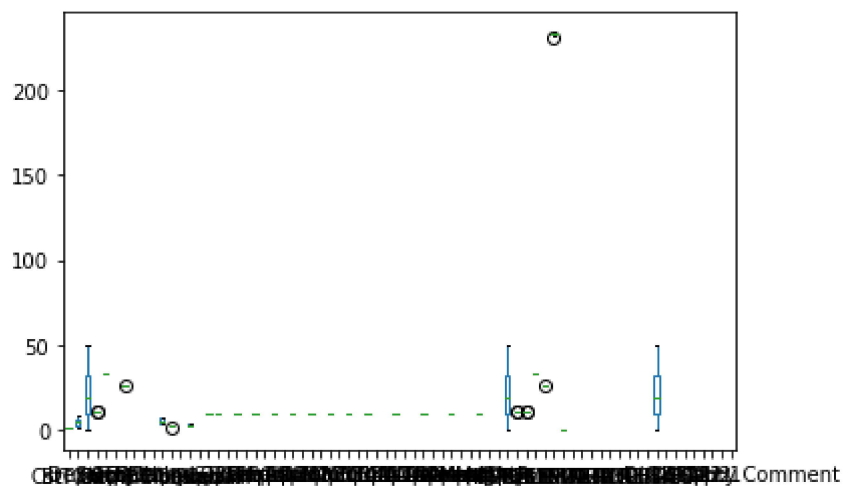




## 14.box plot

In [14]: `da.plot.box()`

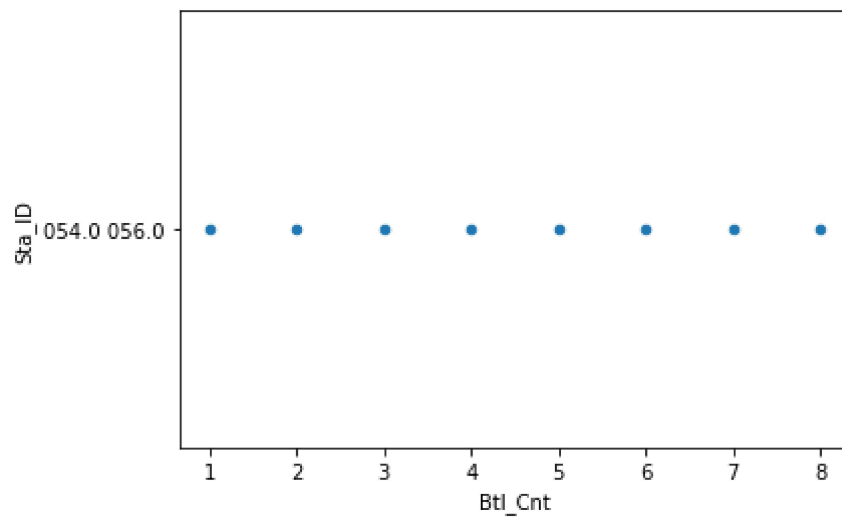
Out[14]: <AxesSubplot:>



## 15.scatter plot

In [15]: `da.plot.scatter("Btl_Cnt", "Sta_ID")`

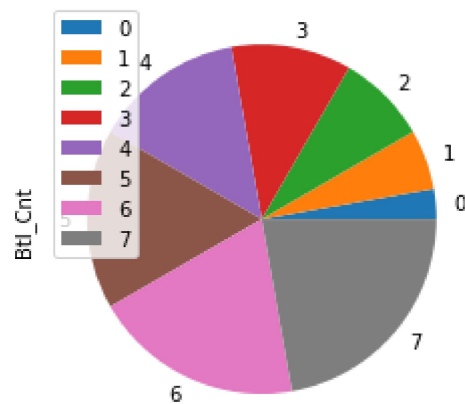
Out[15]: <AxesSubplot:xlabel='Btl\_Cnt', ylabel='Sta\_ID'>



## 16.pie plot

In [20]: `da.plot.pie(y="Btl_Cnt")`

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



## 17.histogram

In [17]: `da.plot.hist()`

Out[17]: `<AxesSubplot:ylabel='Frequency'>`

