27-07-2023

In [ ]:
```
# import libaries
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
```

In [6]:
```
x=pd.read_csv(r"C:\Users\user\Downloads\9_bottle.csv")
```

```
C:\ProgramData\Anaconda3\lib\site-packages\IPython\core\interactiveshell.py:3
165: DtypeWarning: Columns (47,73) have mixed types.Specify dtype option on i
mport or set low_memory=False.
  has_raised = await self.run_ast_nodes(code_ast.body, cell_name,
```

Out[6]:

	Cst_Cnt	Btl_Cnt	Sta_ID	Depth_ID	Depthm	T_degC	Salnty	O2n
0	1	1	054.0 056.0	19-4903CR-HY-060-0930-05400560-0000A-3	0	10.500	33.4400	
1	1	2	054.0 056.0	19-4903CR-HY-060-0930-05400560-0008A-3	8	10.460	33.4400	
2	1	3	054.0 056.0	19-4903CR-HY-060-0930-05400560-0010A-7	10	10.460	33.4370	
3	1	4	054.0 056.0	19-4903CR-HY-060-0930-05400560-0019A-3	19	10.450	33.4200	
4	1	5	054.0 056.0	19-4903CR-HY-060-0930-05400560-0020A-7	20	10.450	33.4210	
...	...	...	...	...	...	...	...	
864858	34404	864859	093.4 026.4	20-1611SR-MX-310-2239-09340264-0000A-7	0	18.744	33.4083	5
864859	34404	864860	093.4 026.4	20-1611SR-MX-310-2239-09340264-0002A-3	2	18.744	33.4083	5
864860	34404	864861	093.4 026.4	20-1611SR-MX-310-2239-09340264-0005A-3	5	18.692	33.4150	5
864861	34404	864862	093.4 026.4	20-1611SR-MX-310-2239-09340264-0010A-3	10	18.161	33.4062	5
864862	34404	864863	093.4 026.4	20-1611SR-MX-310-2239-09340264-0015A-3	15	17.533	33.3880	5

864863 rows × 74 columns

In [7]:
```
x=x.head(100)
```

Out[7]:

	Cst_Cnt	Btl_Cnt	Sta_ID	Depth_ID	Depthm	T_degC	Salnty	O2ml_L
0	1	1	054.0 056.0	19-4903CR-HY-060-0930-05400560-0000A-3	0	10.50	33.440	NaN
1	1	2	054.0 056.0	19-4903CR-HY-060-0930-05400560-0008A-3	8	10.46	33.440	NaN
2	1	3	054.0 056.0	19-4903CR-HY-060-0930-05400560-0010A-7	10	10.46	33.437	NaN
3	1	4	054.0 056.0	19-4903CR-HY-060-0930-05400560-0019A-3	19	10.45	33.420	NaN
4	1	5	054.0 056.0	19-4903CR-HY-060-0930-05400560-0020A-7	20	10.45	33.421	NaN
...	...	...	...	...	...	...	...	...
95	4	96	050.0 095.0	19-4903CR-HY-061-1042-05000950-0019A-3	19	10.14	32.660	NaN
96	4	97	050.0 095.0	19-4903CR-HY-061-1042-05000950-0020A-7	20	10.14	32.657	NaN
97	4	98	050.0 095.0	19-4903CR-HY-061-1042-05000950-0030A-7	30	10.07	32.641	NaN
98	4	99	050.0 095.0	19-4903CR-HY-061-1042-05000950-0040A-3	40	9.97	NaN	NaN
99	4	100	050.0 095.0	19-4903CR-HY-061-1042-05000950-0050A-7	50	9.72	32.699	NaN

100 rows × 74 columns

In [9]:

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 100 entries, 0 to 99
Data columns (total 74 columns):
 #   Column               Non-Null Count  Dtype  
---  ------               --------------  -----  
 0   Cst_Cnt              100 non-null    int64  
 1   Btl_Cnt              100 non-null    int64  
 2   Sta_ID               100 non-null    object 
 3   Depth_ID             100 non-null    object 
 4   Depthm               100 non-null    int64  
 5   T_degC               100 non-null    float64
 6   Salnty               98 non-null     float64
 7   O2ml_L               0 non-null      float64
 8   STheta               98 non-null     float64
 9   O2Sat                0 non-null      float64
 10  Oxy_µmol/Kg          0 non-null      float64
 11  BtlNum               0 non-null      float64
 12  RecInd               100 non-null    int64  
 13  T_prec               100 non-null    float64
 14  T_qual               0 non-null      float64
 15  S_prec               98 non-null     float64
 16  S_qual               2 non-null      float64
 17  P_qual               100 non-null    float64
 18  O_qual               100 non-null    float64
 19  SThtaq               2 non-null      float64
 20  O2Satq               100 non-null    float64
 21  ChlorA               0 non-null      float64
 22  Chlqua               100 non-null    float64
 23  Phaeop               0 non-null      float64
 24  Phaqua               100 non-null    float64
 25  PO4uM                0 non-null      float64
 26  PO4q                 100 non-null    float64
 27  SiO3uM               0 non-null      float64
 28  SiO3qu               100 non-null    float64
 29  NO2uM                0 non-null      float64
 30  NO2q                 100 non-null    float64
 31  NO3uM                0 non-null      float64
 32  NO3q                 100 non-null    float64
 33  NH3uM                0 non-null      float64
 34  NH3q                 100 non-null    float64
 35  C14As1               0 non-null      float64
 36  C14A1p               0 non-null      float64
 37  C14A1q               100 non-null    float64
 38  C14As2               0 non-null      float64
 39  C14A2p               0 non-null      float64
 40  C14A2q               100 non-null    float64
 41  DarkAs               0 non-null      float64
 42  DarkAp               0 non-null      float64
 43  DarkAq               100 non-null    float64
 44  MeanAs               0 non-null      float64
 45  MeanAp               0 non-null      float64
 46  MeanAq               100 non-null    float64
 47  IncTim               0 non-null      object 
 48  LightP               0 non-null      float64
```

```
 49  R_Depth              100 non-null    float64
 50  R_TEMP               100 non-null    float64
 51  R_POTEMP             100 non-null    float64
 52  R_SALINITY           98 non-null     float64
 53  R_SIGMA              98 non-null     float64
 54  R_SVA                98 non-null     float64
 55  R_DYNHT              100 non-null    float64
 56  R_O2                 0 non-null      float64
 57  R_O2Sat              0 non-null      float64
 58  R_SIO3               0 non-null      float64
 59  R_PO4                0 non-null      float64
 60  R_NO3                0 non-null      float64
 61  R_NO2                0 non-null      float64
 62  R_NH4                0 non-null      float64
 63  R_CHLA               0 non-null      float64
 64  R_PHAEO              0 non-null      float64
 65  R_PRES               100 non-null    int64
 66  R_SAMP               0 non-null      float64
 67  DIC1                 0 non-null      float64
 68  DIC2                 0 non-null      float64
 69  TA1                  0 non-null      float64
 70  TA2                  0 non-null      float64
 71  pH2                  0 non-null      float64
 72  pH1                  0 non-null      float64
 73  DIC Quality Comment  0 non-null      object
dtypes: float64(65), int64(5), object(4)
memory usage: 57.9+ KB
```

In [ ]:

In [10]:

Out[10]:
```
Index(['Cst_Cnt', 'Btl_Cnt', 'Sta_ID', 'Depth_ID', 'Depthm', 'T_degC',
       'Salnty', 'O2ml_L', 'STheta', 'O2Sat', 'Oxy_µmol/Kg', 'BtlNum',
       'RecInd', 'T_prec', 'T_qual', 'S_prec', 'S_qual', 'P_qual', 'O_qual',
       'SThtaq', 'O2Satq', 'ChlorA', 'Chlqua', 'Phaeop', 'Phaqua', 'PO4uM',
       'PO4q', 'SiO3uM', 'SiO3qu', 'NO2uM', 'NO2q', 'NO3uM', 'NO3q', 'NH3uM',
       'NH3q', 'C14As1', 'C14A1p', 'C14A1q', 'C14As2', 'C14A2p', 'C14A2q',
       'DarkAs', 'DarkAp', 'DarkAq', 'MeanAs', 'MeanAp', 'MeanAq', 'IncTim',
       'LightP', 'R_Depth', 'R_TEMP', 'R_POTEMP', 'R_SALINITY', 'R_SIGMA',
       'R_SVA', 'R_DYNHT', 'R_O2', 'R_O2Sat', 'R_SIO3', 'R_PO4', 'R_NO3',
       'R_NO2', 'R_NH4', 'R_CHLA', 'R_PHAEO', 'R_PRES', 'R_SAMP', 'DIC1',
       'DIC2', 'TA1', 'TA2', 'pH2', 'pH1', 'DIC Quality Comment'],
      dtype='object')
```

In [28]:
```
d=x[['R_PRES','RecInd','Depthm','Cst_Cnt']]
```

Out[28]:

	R_PRES	RecInd	Depthm	Cst_Cnt
0	0	3	0	1
1	8	3	8	1
2	10	7	10	1
3	19	3	19	1
4	20	7	20	1
...	...	...	...	...
95	19	3	19	4
96	20	7	20	4
97	30	7	30	4
98	40	3	40	4
99	50	7	50	4

100 rows × 4 columns

In [30]:

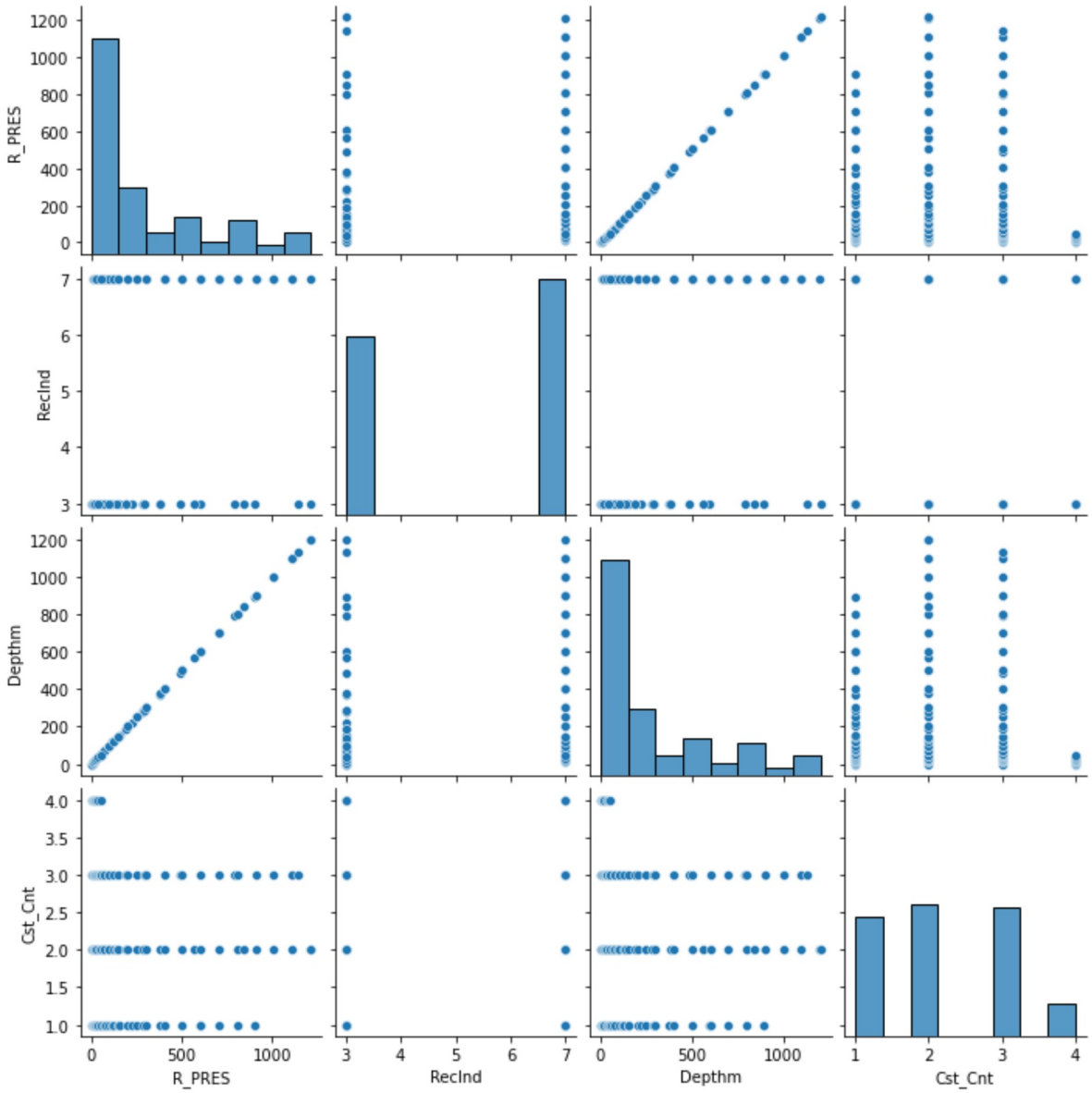
Out[30]:

	Cst_Cnt	Btl_Cnt	Depthm	T_degC	Salnty	O2ml_L	STheta	O2Sat
count	100.000000	100.000000	100.000000	100.000000	98.000000	0.0	98.000000	0.0
mean	2.180000	50.500000	302.130000	7.625300	33.621755	NaN	26.224337	NaN
std	0.946872	29.011492	338.487296	2.413229	0.580959	NaN	0.767899	NaN
min	1.000000	1.000000	0.000000	3.140000	32.630000	NaN	25.069000	NaN
25%	1.000000	25.750000	45.250000	5.297500	33.159000	NaN	25.628750	NaN
50%	2.000000	50.500000	150.000000	8.370000	33.810000	NaN	26.293000	NaN
75%	3.000000	75.250000	500.000000	9.862500	34.112250	NaN	26.970250	NaN
max	4.000000	100.000000	1203.000000	10.500000	34.450000	NaN	27.425000	NaN

8 rows × 69 columns

In [31]:

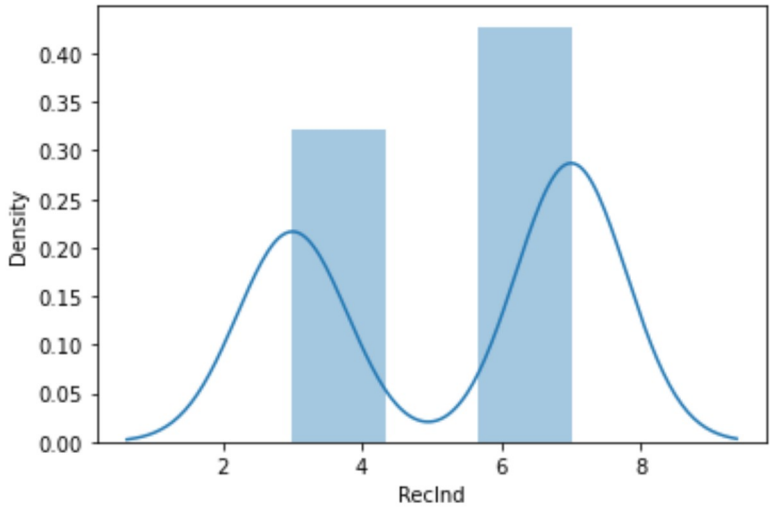
Out[31]: <seaborn.axisgrid.PairGrid at 0x26008350580>



In [61]:

```
C:\ProgramData\Anaconda3\lib\site-packages\seaborn\distributions.py:2557: Fut
ureWarning: `distplot` is a deprecated function and will be removed in a futu
re version. Please adapt your code to use either `displot` (a figure-level fu
nction with similar flexibility) or `histplot` (an axes-level function for hi
stograms).
  warnings.warn(msg, FutureWarning)
```

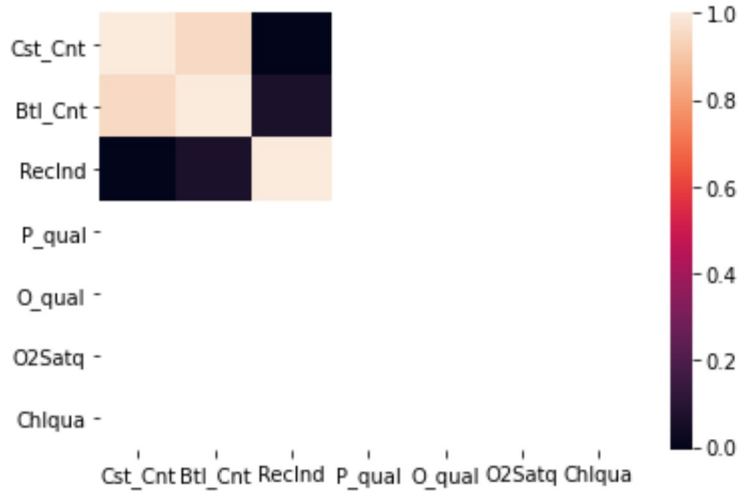
Out[61]: <AxesSubplot:xlabel='RecInd', ylabel='Density'>



In [63]:
```
x1=x[['Cst_Cnt', 'Btl_Cnt', 'Sta_ID', 'Depth_ID',  'RecInd',
```

In [64]:

Out[64]: <AxesSubplot:>



In [68]:
```
x=x1[['Cst_Cnt', 'Btl_Cnt']]
```

In [69]:
```
# to split my dataset into traning and test date

from sklearn.model_selection import train_test_split
```

In [70]:
```
from sklearn.linear_model import LinearRegression

lr=LinearRegression()
```

Out[70]:
```
LinearRegression()
```

In [71]:

```
1.4210854715202004e-14
```

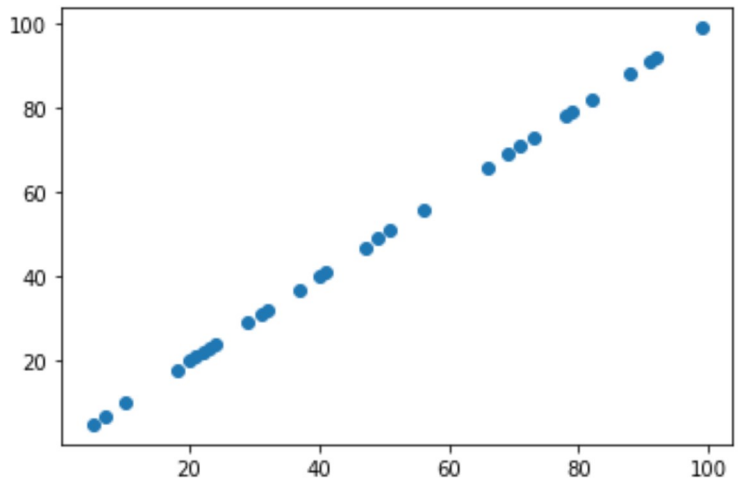
In [72]:
```
coeff=pd.DataFrame(lr.coef_,x.columns,columns=['Co-efficient'])
```

Out[72]:

	Co-efficient
Cst_Cnt	-6.248393e-15
Btl_Cnt	1.000000e+00

In [73]:
```
prediction=lr.predict(x_test)
```

Out[73]:
```
<matplotlib.collections.PathCollection at 0x2602b9abe80>
```



In [74]:

Out[74]:
```
1.0
```

In [ ]: