DATE:31-5-23_____**RELATIONSHIP B/W SALINITY & WATER TEMP.(ML)**

```
import numpy as np
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
from sklearn import preprocessing,svm
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression
```

->READ THE DATA SET

```
df=pd.read_csv("/content/bottle.csv")
df
```

```
df.isna().any()
```

```
Sal     True
Temp    True
dtype: bool
```

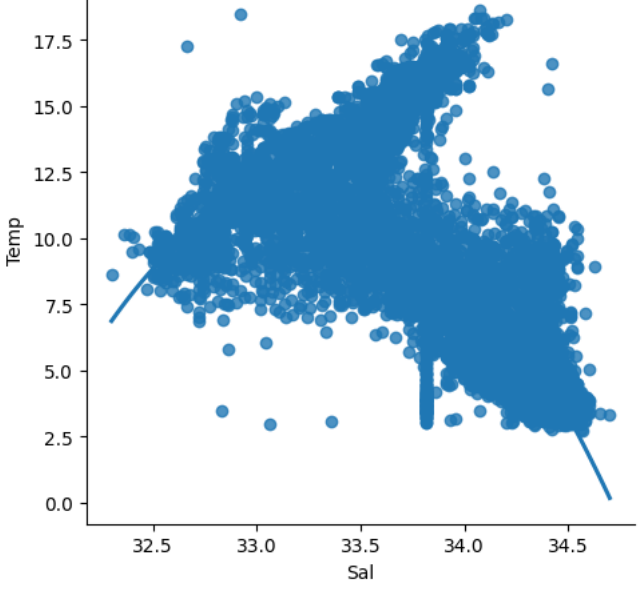
```
df
```

	Sal	Temp
0	33.440	10.50
1	33.440	10.46
2	33.437	10.46
3	33.420	10.45
4	33.421	10.45
...	...	...
7975	33.609	11.92
7976	33.600	11.04
7977	33.647	10.71
7978	33.930	9.42
7979	33.964	9.22

7980 rows × 2 columns

```
df1=df.dropna()
```

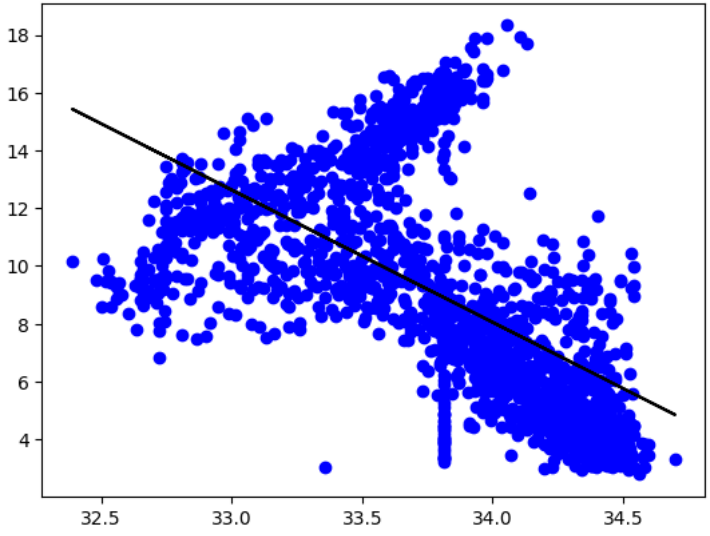
```
df1
```




```
df.describe()
```

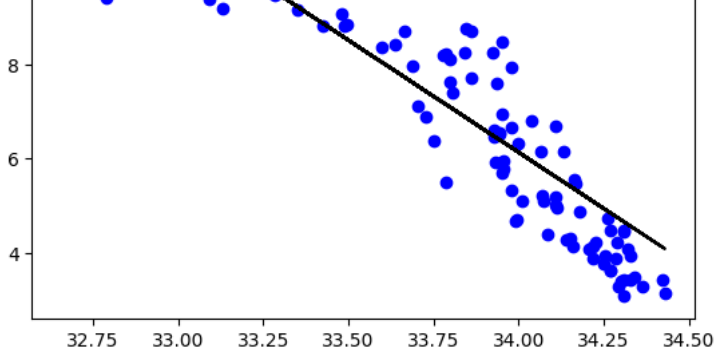
	Sal	Temp
count	7980.000000	7980.000000
mean	33.812703	8.869087
std	0.512919	3.915512
min	32.300000	2.700000
25%	33.492000	5.350000
50%	33.880000	8.420000
75%	34.251000	11.920000
max	34.700000	19.760000

```
df.fillna(method='ffill',inplace=True)
```




```
df500=df[:][:500]
```

```
sns.lmplot(x="Sal",y="Temp",data=df500,order=1,ci=None)
```




```
from sklearn.linear_model import LinearRegression
from sklearn.metrics import r2_score


model=LinearRegression()
model.fit(x_train,y_train)
```

▾ LinearRegression
LinearRegression()