water quality - ML project

1. dataset

- [ ] original data

1.1 original dataframe

	Date	Salinity (ppt)	DissolvedOxygen (mg/L)	pH	SecchiDepth (m)	WaterDepth (m)	W
0	1989-05-11	None	None	7.500000	0.300000	0.900000	
1	1989-05-18	None	12.000000	7.500000	0.200000	0.600000	
2	1989-05-25	None	None	8.000000	0.400000	0.800000	
3	1989-06-01	None	12.000000	8.000000	0.400000	0.900000	
4	1989-07-11	None	None	8.500000	0.300000	0.900000	

- [ ] null values

1.2 to check the null values:

```
Date                      5
Salinity (ppt)          130
DissolvedOxygen (mg/L)  851
pH                       95
SecchiDepth (m)          73
WaterDepth (m)           71
WaterTemp (C)           121
AirTemp (C)               0
dtype: int64
```

- [ ] null values after fill

1.3 null values after fill:

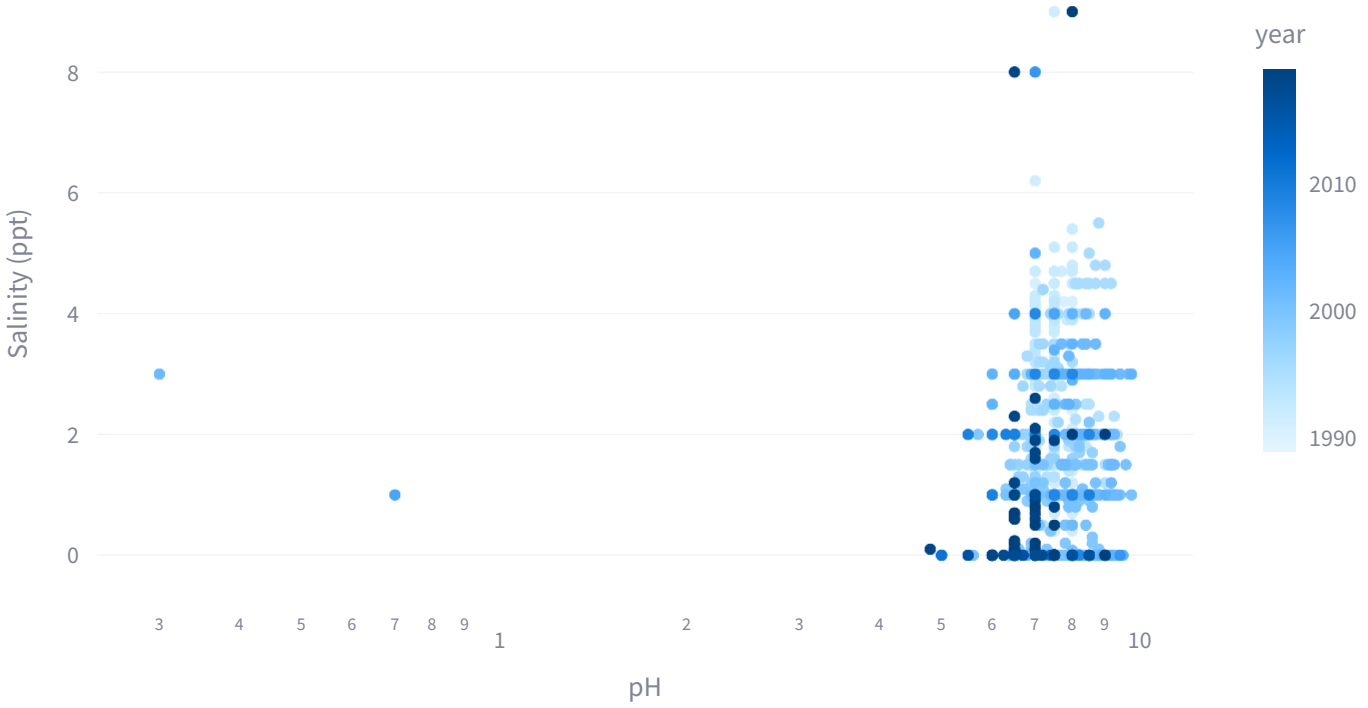
```
Date                    0
Salinity (ppt)          0
DissolvedOxygen (mg/L)  0
pH                      0
SecchiDepth (m)         0
```

```
WaterDepth (m)          0
WaterTemp (C)           0
AirTemp (C)             0
dtype: int64
```

☐ scatterplot

1.4 scatterplot

Streamlit theme (default) Plotly native theme



2. linear regression

ML models:

prediction

2.3 taking new inputs from user for the prediction

Enter dissolved oxygen:

6.00

pH:

3.99 − +

sacchi depth:

1.67 − +

water depth:

1.78 − +

water temp:

26.00 − +

air temp:

18.00 − +

year:

2017.02 − +

prediction:

[-0.77059419]