

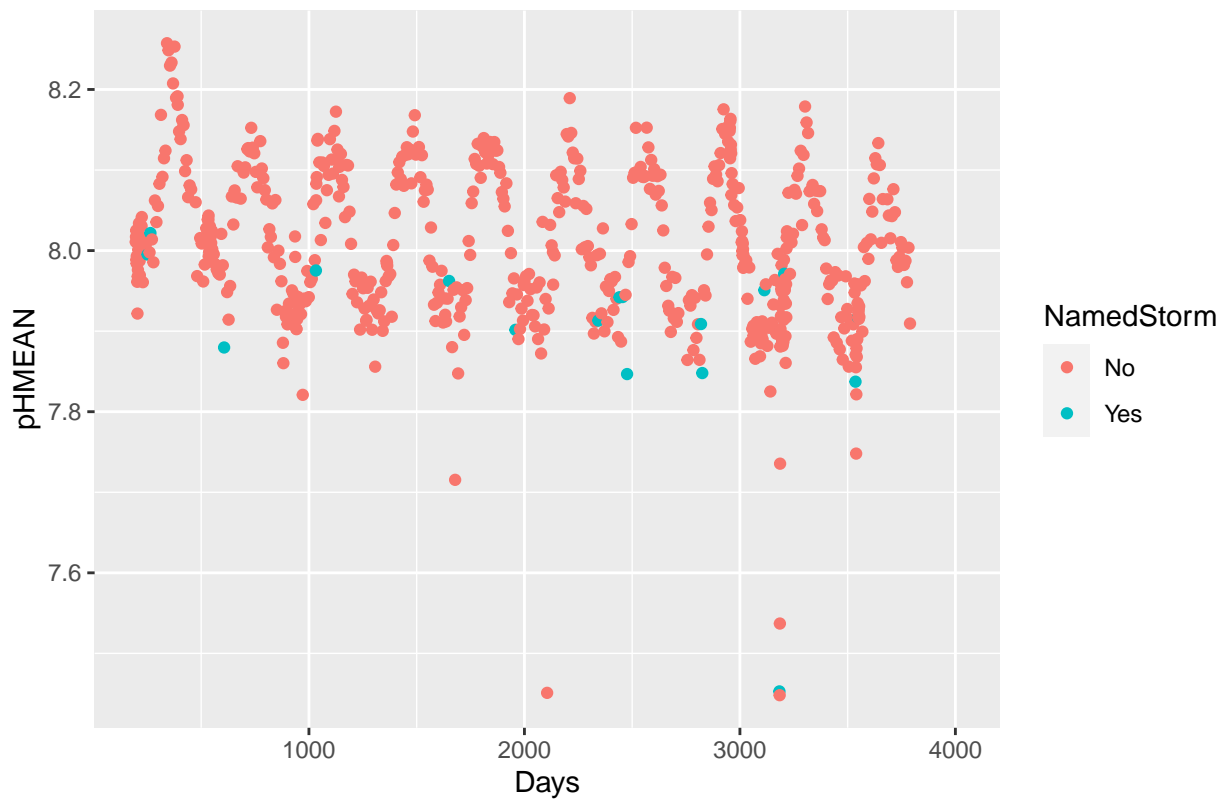
EDA

Alexandra Lawrence

Exploring pH

```
## Warning: Removed 48 rows containing missing values (geom_point).
```

pHMEAN Plotted against Days



Lower pH values observed when there is a named storm – is this because of the storm or just a coincidence because pH tends to lower in warmer seasons and storms happen more often in summer?

Table 1: Linear Model of pH and Days

term	estimate	std.error	statistic	p.value
(Intercept)	8.046527	0.0078787	1021.301971	0
Days	-0.000022	0.0000034	-6.440201	0

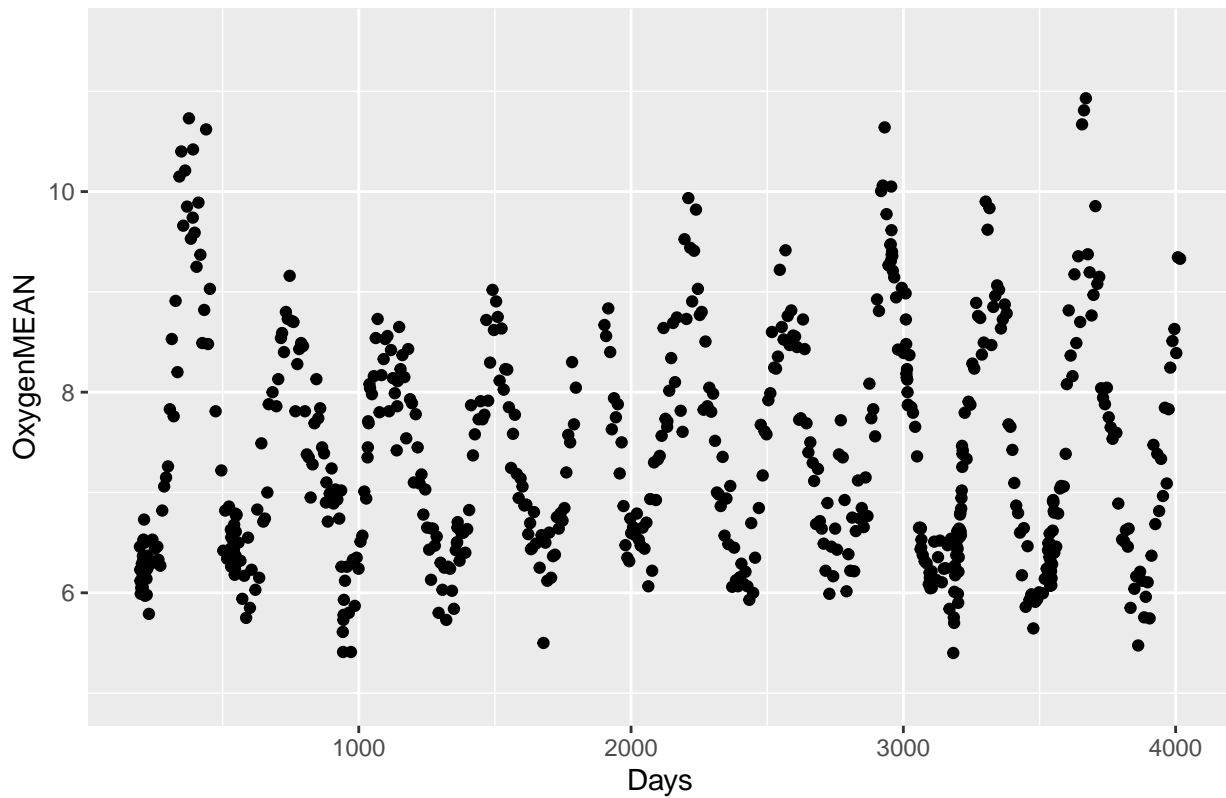
```
##      pHMEAN      Date
## 1 7.448091 9/20/2018
```

The smallest pH value was 9/20/2018 – About a week after Hurricane Florence hit NC

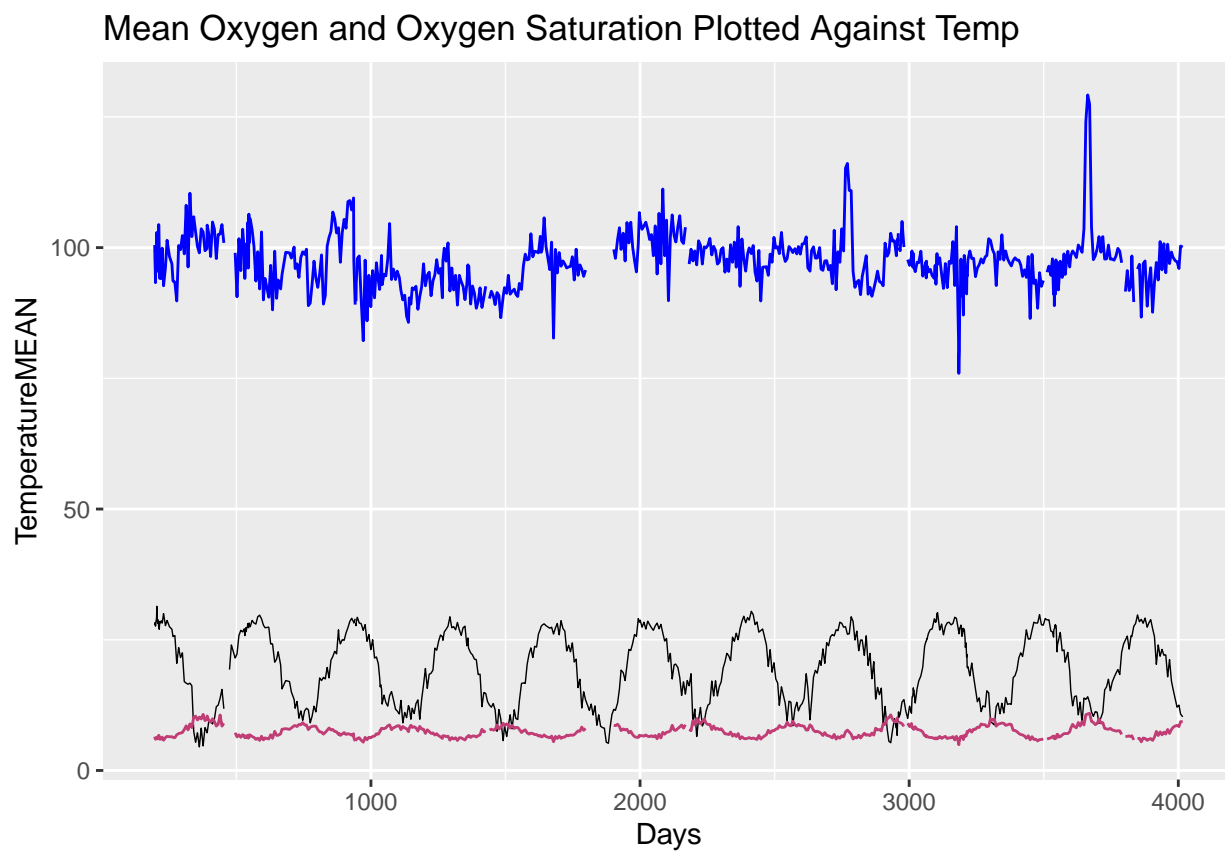
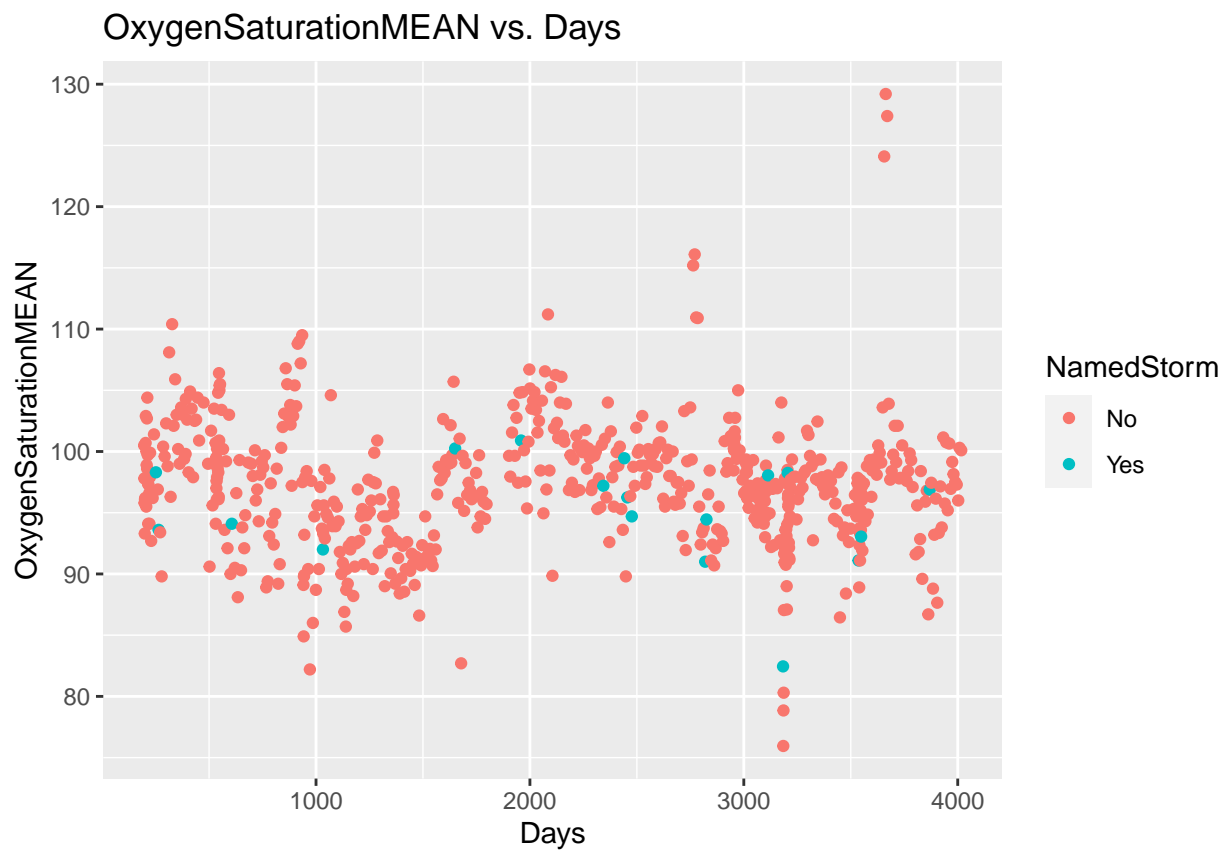
Exploring Oxygen

```
## Warning: Removed 31 rows containing missing values (geom_point).
```

OxygenMEAN vs. Days



```
## Warning: Removed 30 rows containing missing values (geom_point).
```



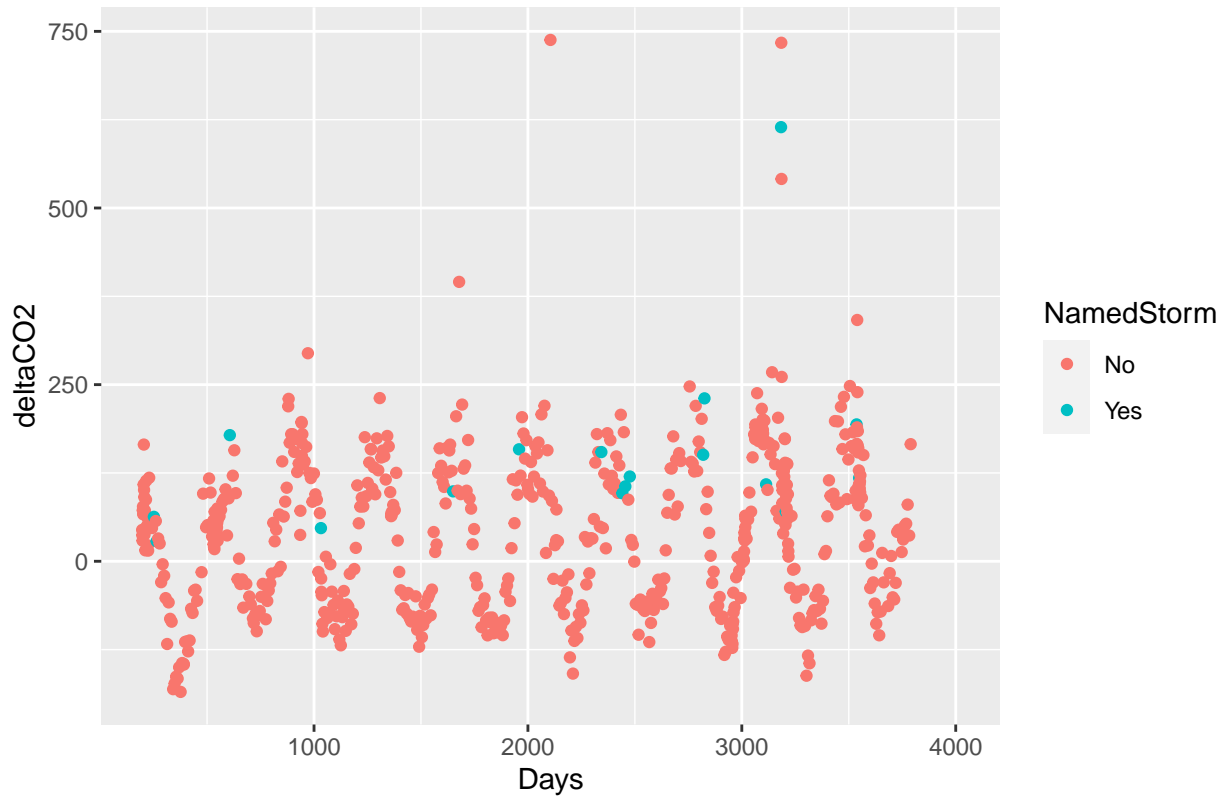
	OxygenSaturationMEAN	Date	OxygenMEAN	TemperatureMEAN
## 1	75.95	9/20/2018	4.93	27.15
## 2	129.20	1/12/2020	10.81	14.10

Max mean oxygen saturation: 129.20% on 1/12/2020 – cooler temp than min Min mean oxygen saturation: 75.95% on. 9/20/2018

Delta CO2

Warning: Removed 48 rows containing missing values (geom_point).

DeltaCO2 vs. Days



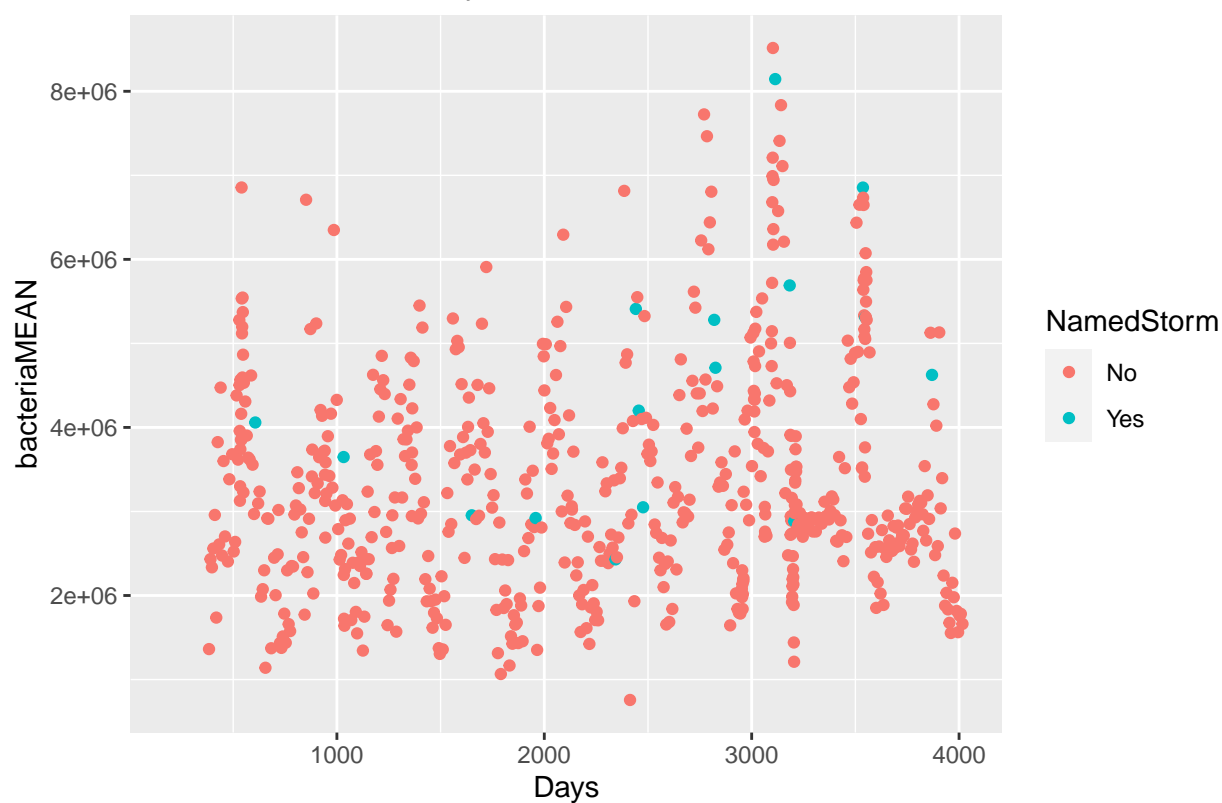
	deltaCO2	Date
## 1	737.8547	10/7/2015

Maximum value happened right around Hurricane Joaquin

Bacteria

Warning: Removed 54 rows containing missing values (geom_point).

BacteriaMEAN vs. Days



```
##   bacteriaMEAN      Date
## 1      8515000 6/30/2018
## Warning: Removed 55 rows containing missing values (geom_point).
## Warning: Removed 55 rows containing missing values (geom_point).
```

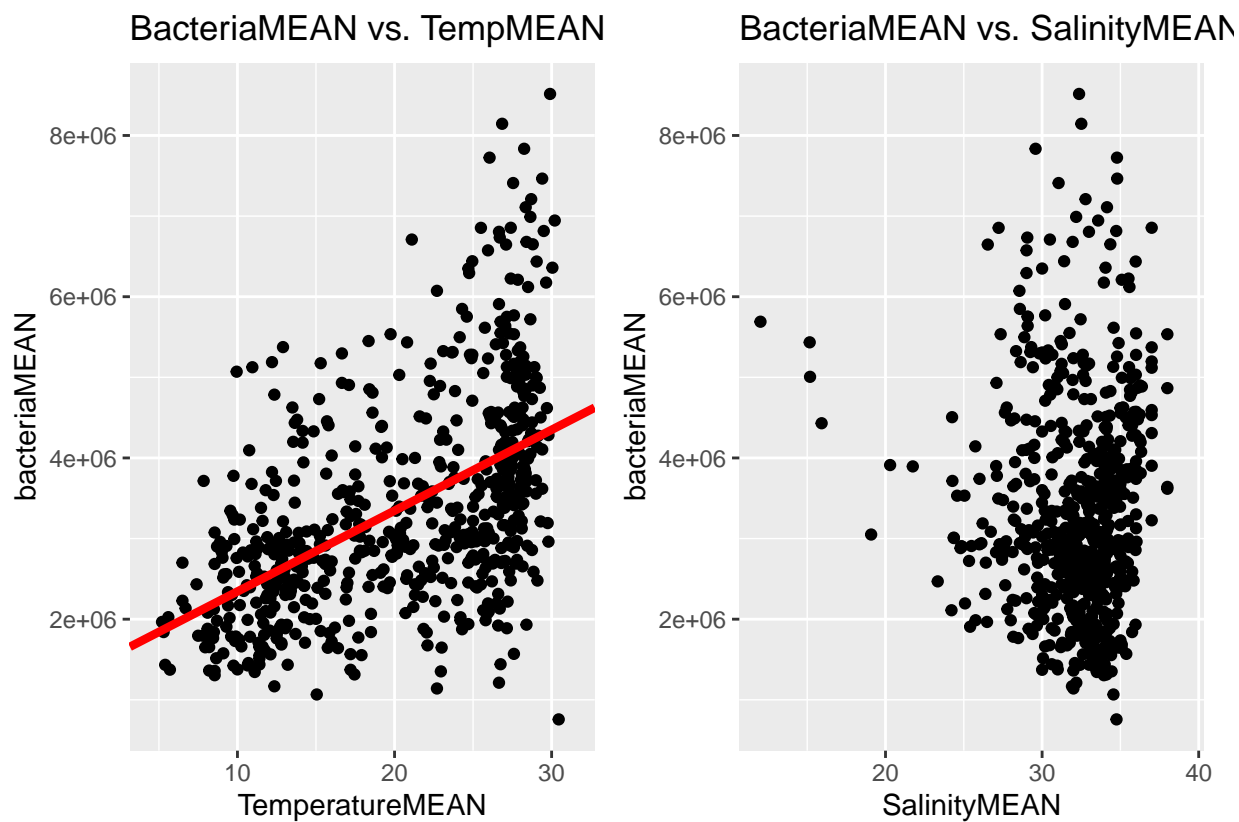
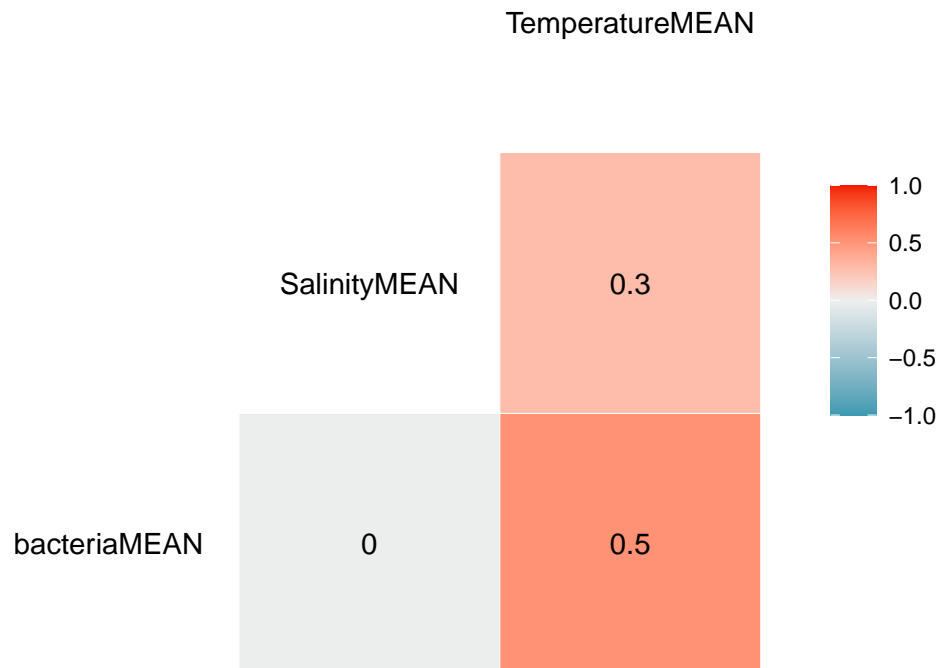


Table 2: Linear Model of Bacteria and Temperature

term	estimate	std.error	statistic	p.value
(Intercept)	1340430.7	140007.557	9.573988	0
TemperatureMEAN	100283.6	6565.585	15.274134	0

Table 3: Linear Model of Bacteria and Salinity

term	estimate	std.error	statistic	p.value
(Intercept)	3508647.531	548596.06	6.3956849	0.0000000
SalinityMEAN	-4518.938	16990.65	-0.2659661	0.7903534



The p-value for the linear regression shows that there is a relationship between mean bacteria and mean temperature, but there does not appear to be a strong correlation between the two

Table 4: Linear Model of Bacteria and pH

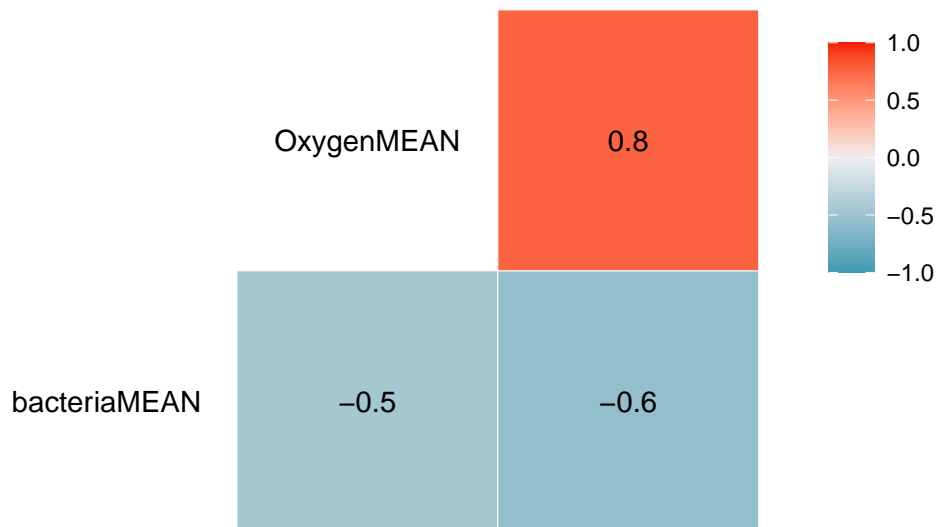
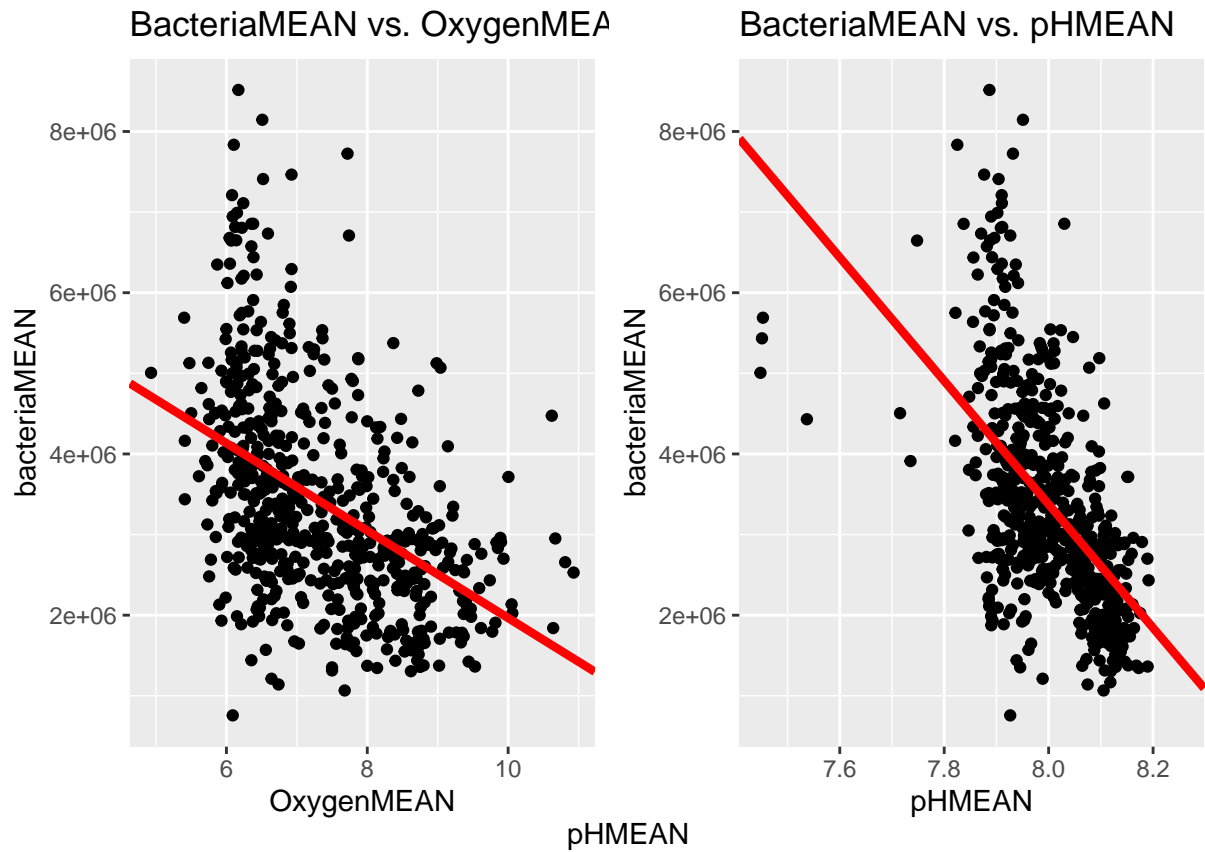
term	estimate	std.error	statistic	p.value
(Intercept)	64645987	3786766.3	17.07155	0
pHMEAN	-7658991	473416.1	-16.17814	0

Table 5: Linear Model of Bacteria and Oxygen

term	estimate	std.error	statistic	p.value
(Intercept)	7384236	319542.54	23.10877	0
OxygenMEAN	-541933	43042.07	-12.59077	0

Warning: Removed 77 rows containing missing values (geom_point).

Warning: Removed 95 rows containing missing values (geom_point).

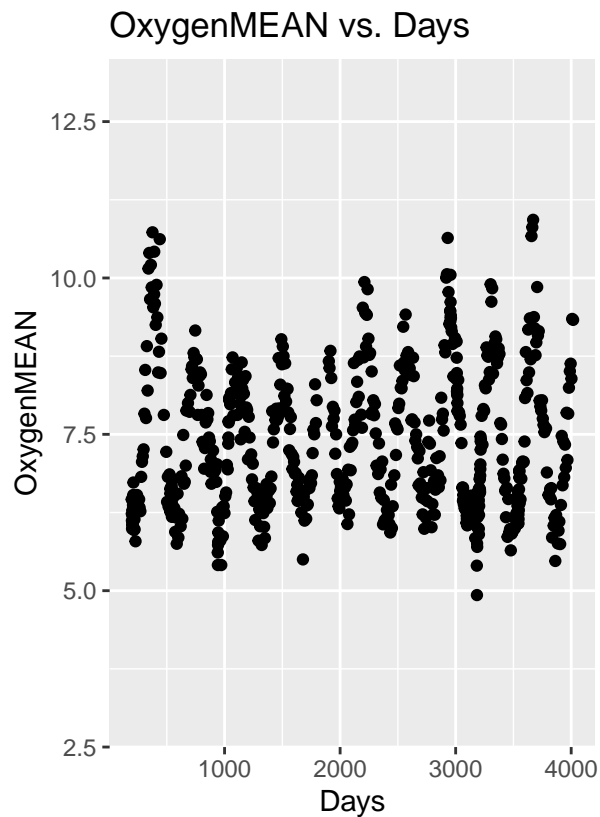
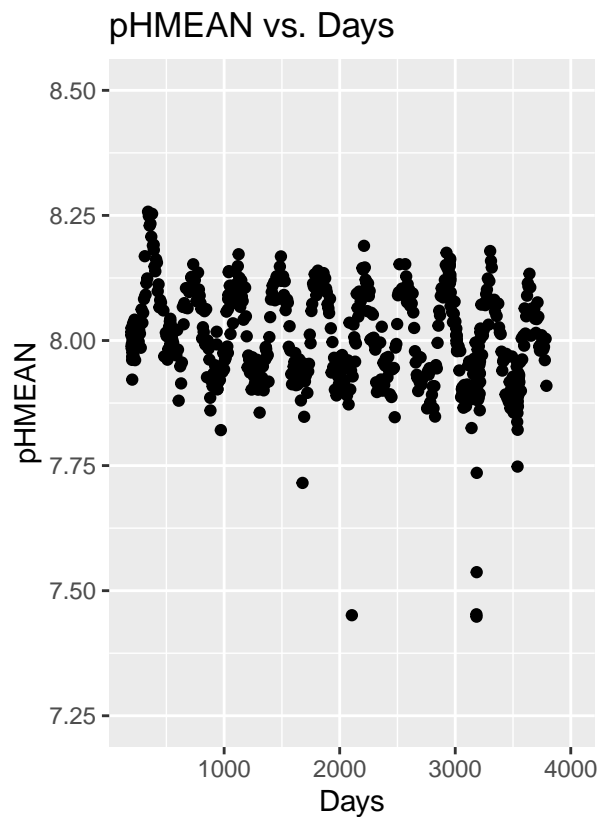


Again, there does not appear to be a strong correlation. between mean bacteria. and pH and oxygen, despite the p-values of the linear models suggesting otherwise

Looking at pH alongside Oxygen

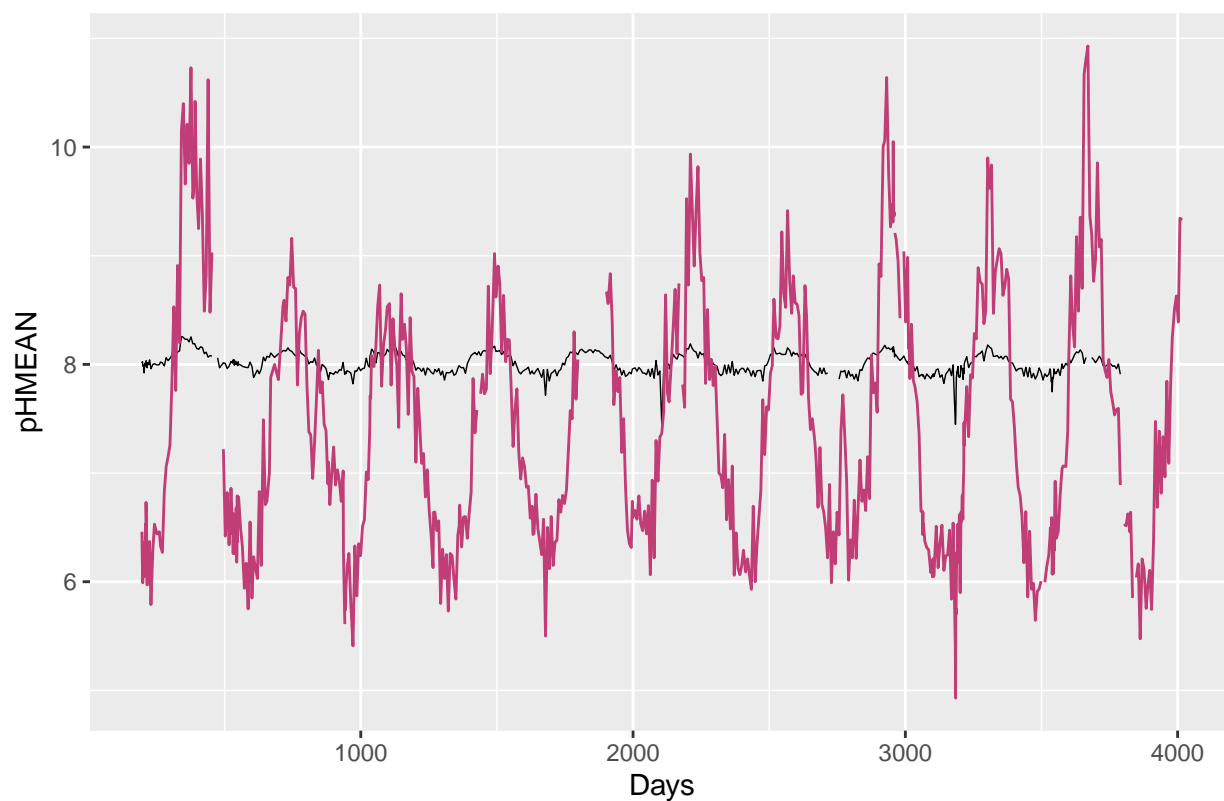
```
## Scale for 'y' is already present. Adding another scale for 'y', which will
## replace the existing scale.
## Scale for 'y' is already present. Adding another scale for 'y', which will
## replace the existing scale.

## Warning: Removed 48 rows containing missing values (geom_point).
## Warning: Removed 30 rows containing missing values (geom_point).
```



```
## Warning: Removed 33 row(s) containing missing values (geom_path).
```

Mean Oxygen Plotted with Mean pH



The mean oxygen measured by ProODO appears to follow a similar pattern of peaks as the mean pH calculated using CO2SYS.

```
## Warning: Removed 68 rows containing missing values (geom_point).
```

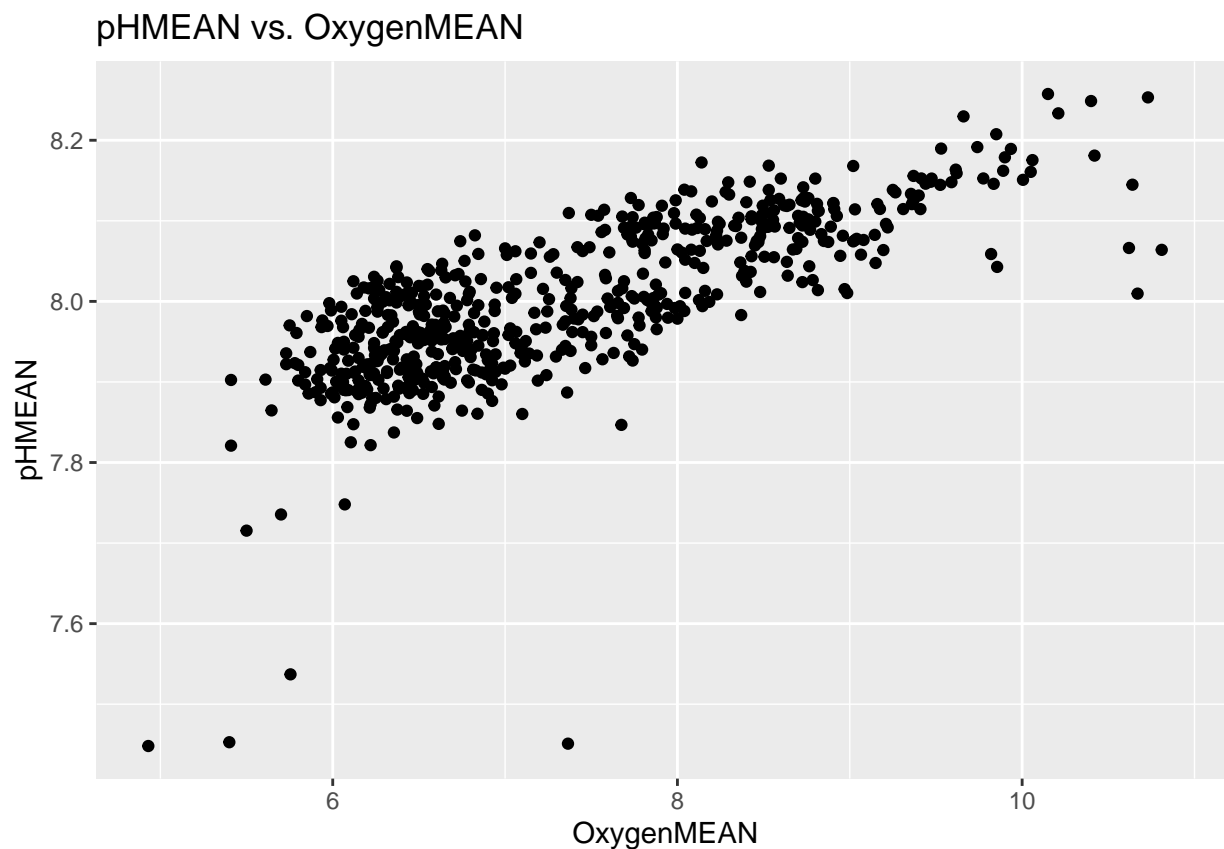
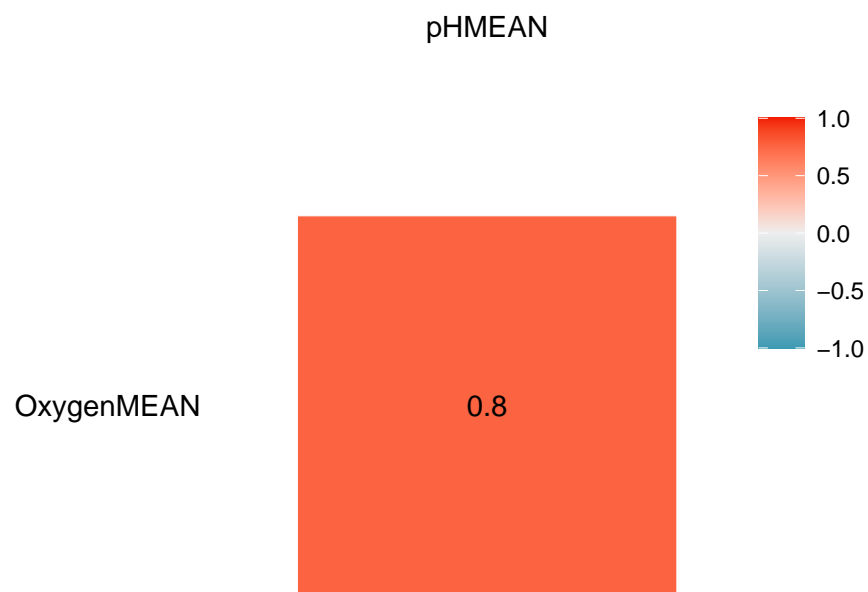


Table 6: Linear Model of pH and Oxygen

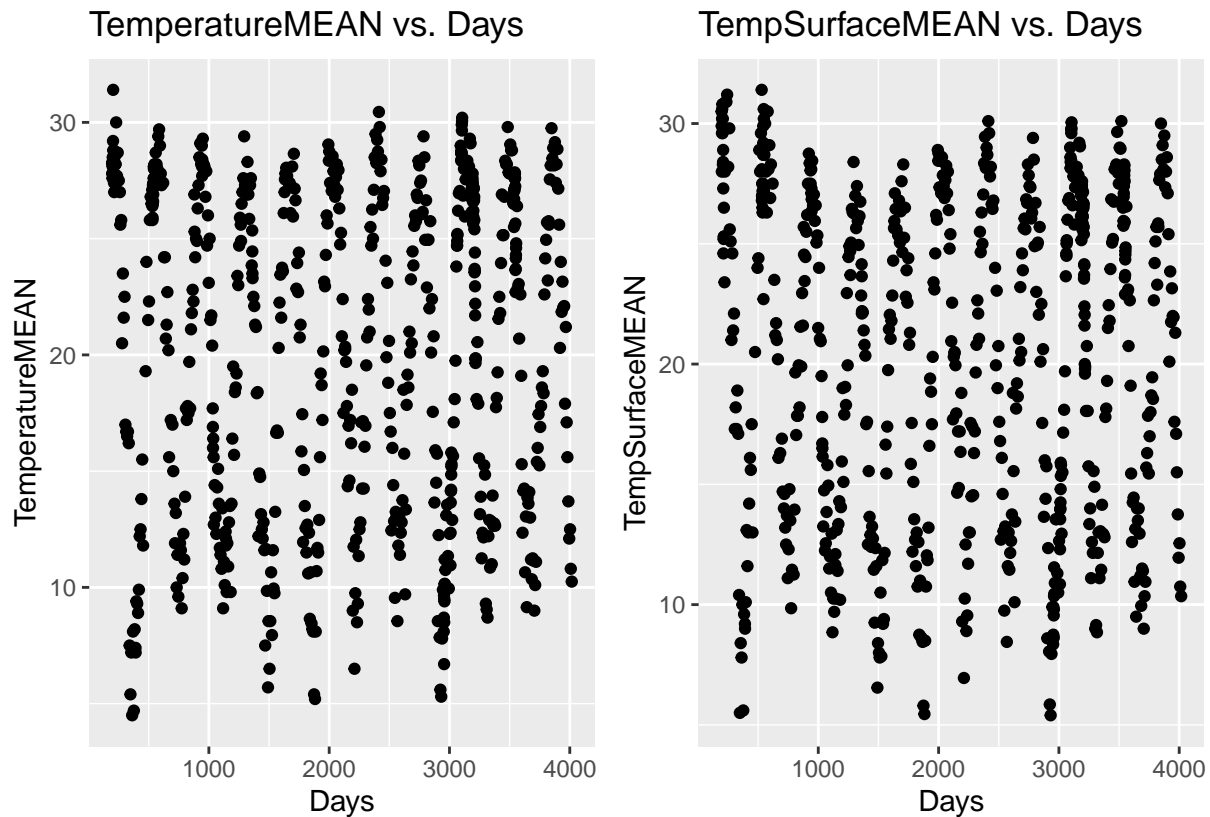
term	estimate	std.error	statistic	p.value
(Intercept)	7.5212403	0.0167333	449.47662	0
OxygenMEAN	0.0650988	0.0022531	28.89327	0



There appears to be a pretty strong positive correlation between the mean oxygen and pH levels.

```
## Warning: Removed 8 rows containing missing values (geom_point).
```

```
## Warning: Removed 13 rows containing missing values (geom_point).
```



```
## Warning: Removed 30 rows containing non-finite values (stat_density).
```

```
## Warning in ggally_statistic(data = data, mapping = mapping, na.rm = na.rm, :  
## Removed 68 rows containing missing values
```

```
## Warning in ggally_statistic(data = data, mapping = mapping, na.rm = na.rm, :  
## Removed 30 rows containing missing values
```

```
## Warning: Removed 68 rows containing missing values (geom_point).
```

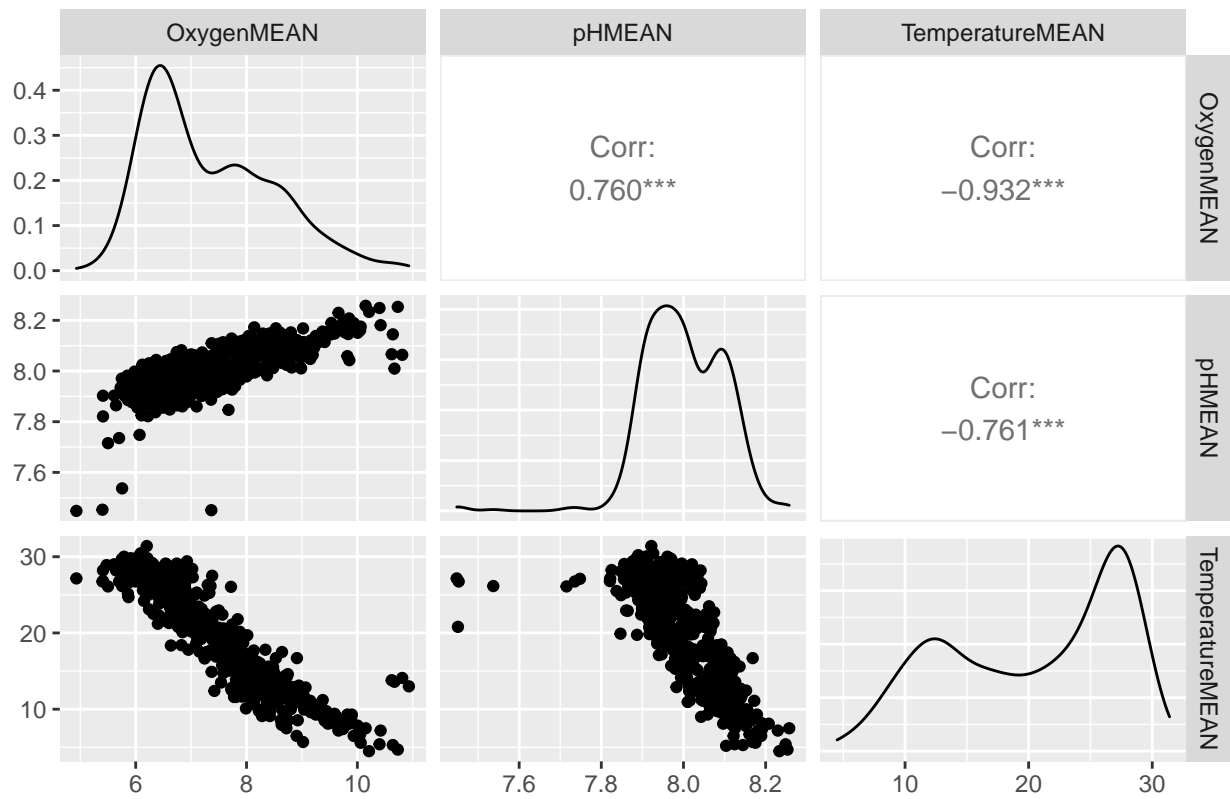
```
## Warning: Removed 48 rows containing non-finite values (stat_density).
```

```
## Warning in ggally_statistic(data = data, mapping = mapping, na.rm = na.rm, :  
## Removed 48 rows containing missing values
```

```
## Warning: Removed 30 rows containing missing values (geom_point).
```

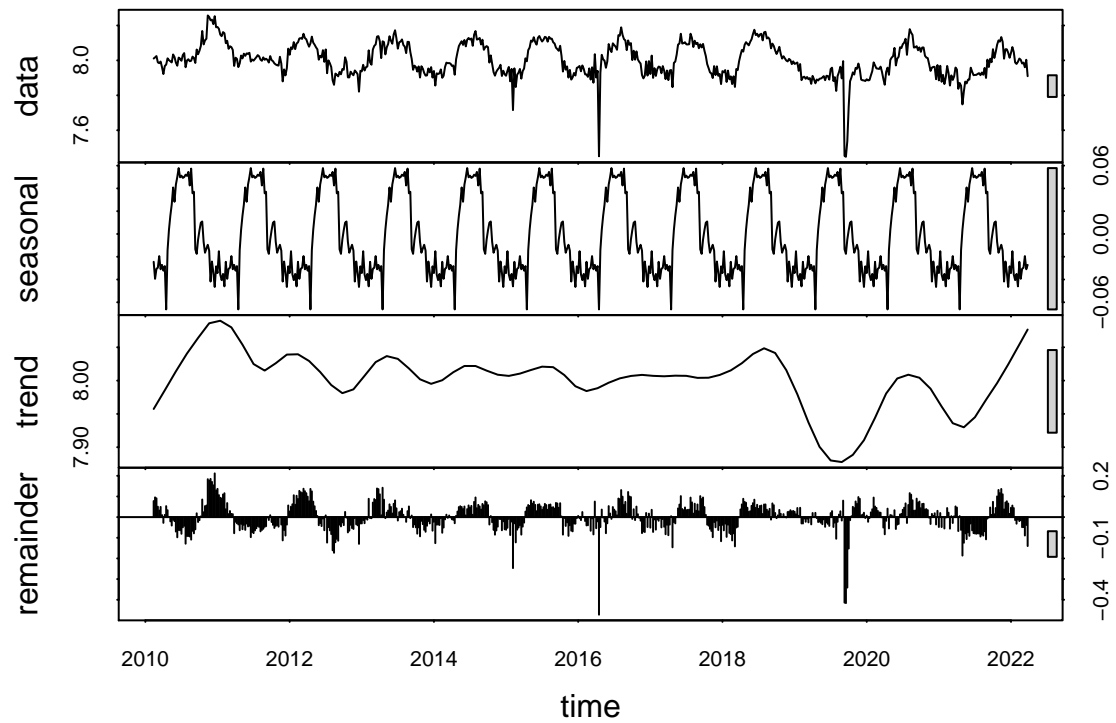
```
## Warning: Removed 48 rows containing missing values (geom_point).
```

```
## Warning: Removed 8 rows containing non-finite values (stat_density).
```

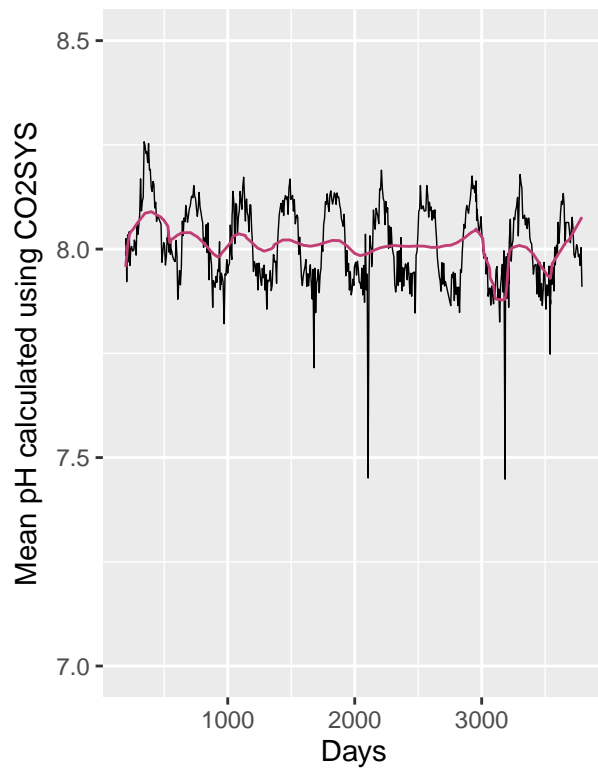


TIME SERIES

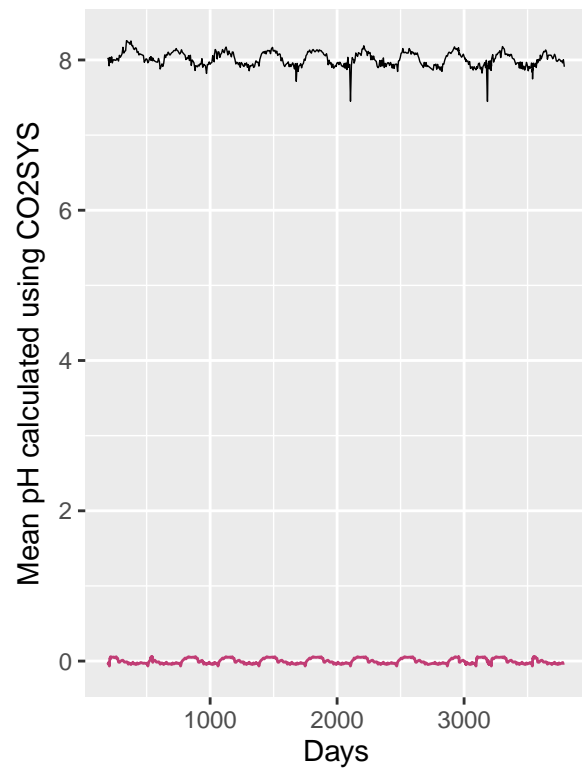
pH Mean



Trend Mapping onto Data

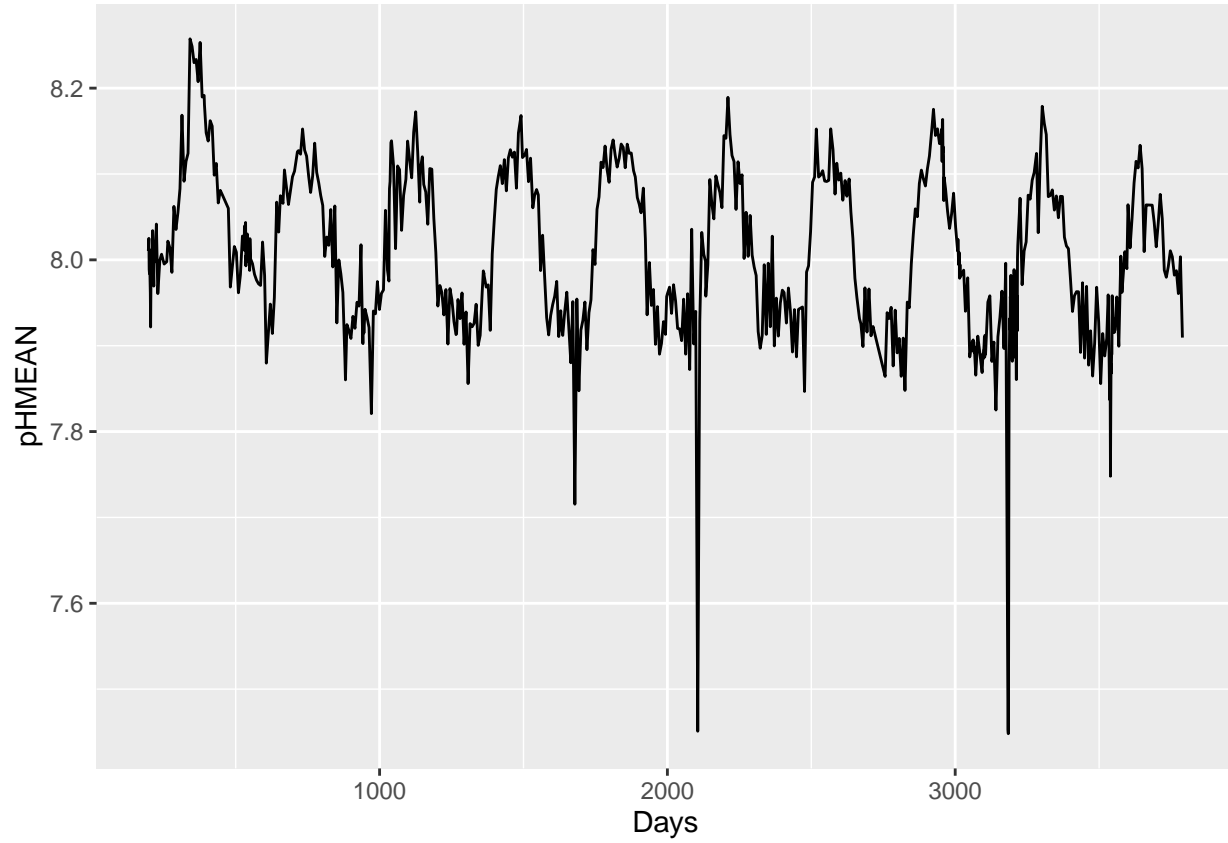


Seasonal Cycle Mapping onto Data

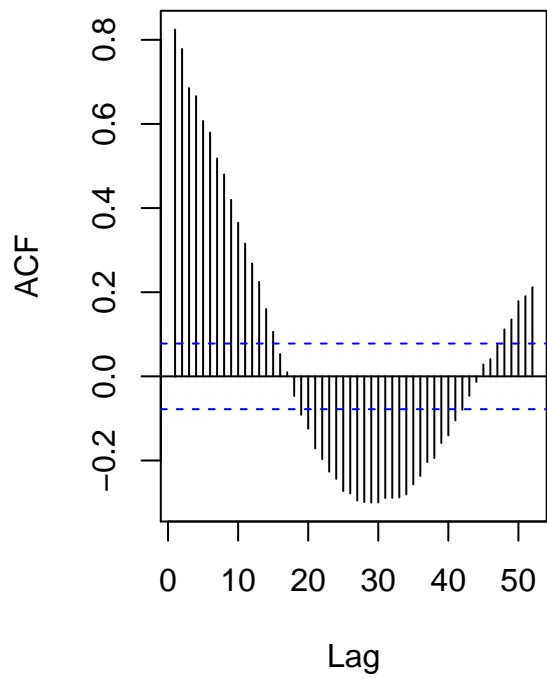


statistic	p.value	kendall_score	denominator	var_kendall_score
-0.0932878	0.0021753	-328	3516	11450.67

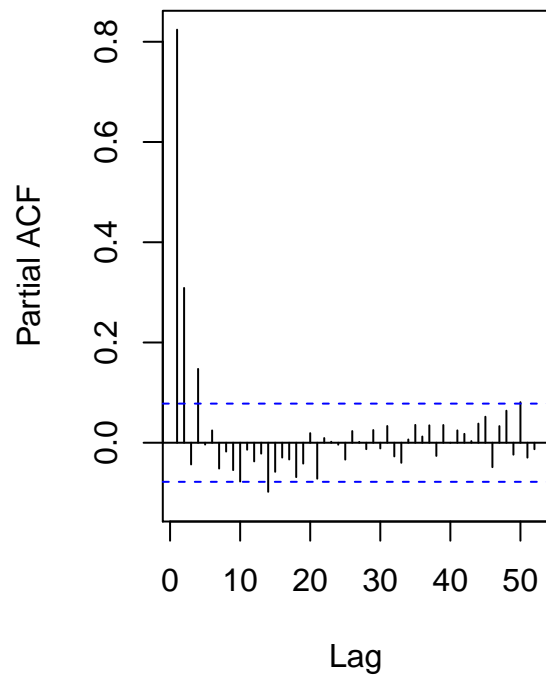
p-value is less than 0.05, so we can reject the null hypothesis meaning that there is a trend



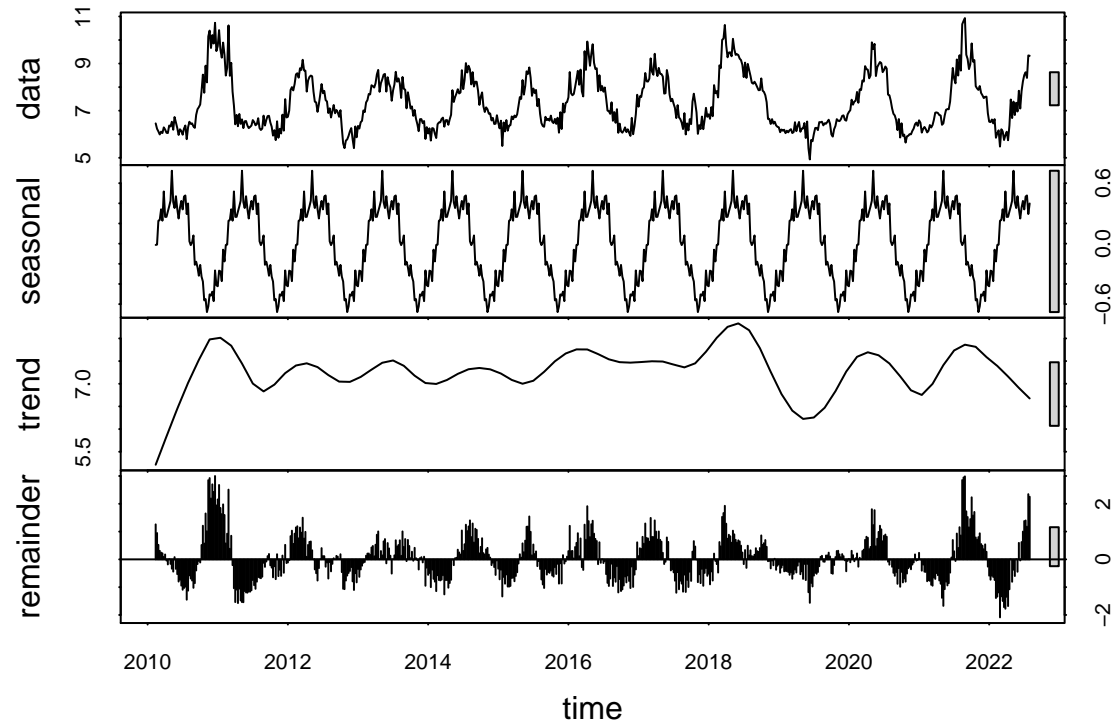
Series full_pHMEAN\$pHMEAN



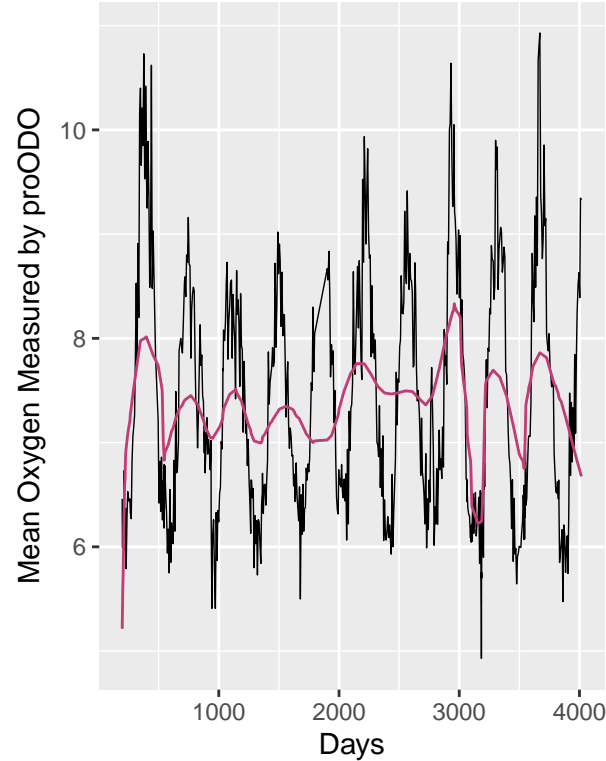
Series full_pHMEAN\$pHMEAN



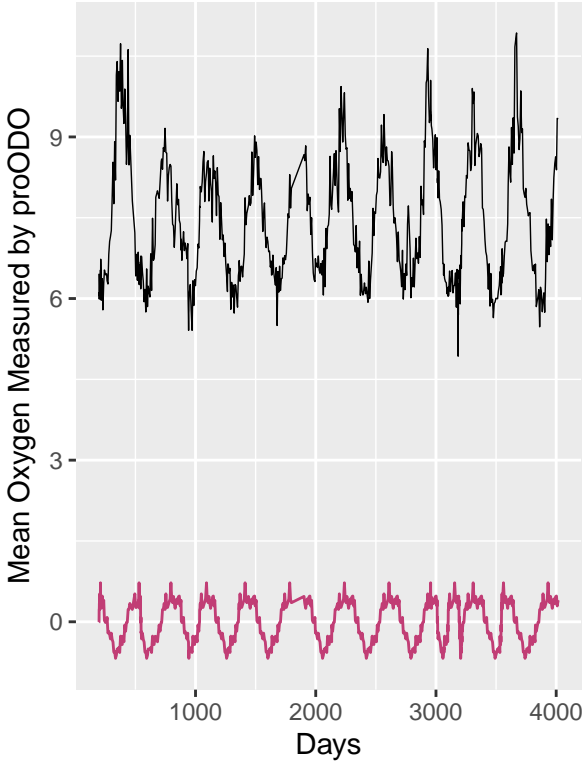
Oxygen



Trend Mapping onto Data



Seasonal Cycle Mapping onto Da



statistic	p.value	kendall_score	denominator	var_kendall_score
0.0724057	0.0155402	270	3728.989	12452.67