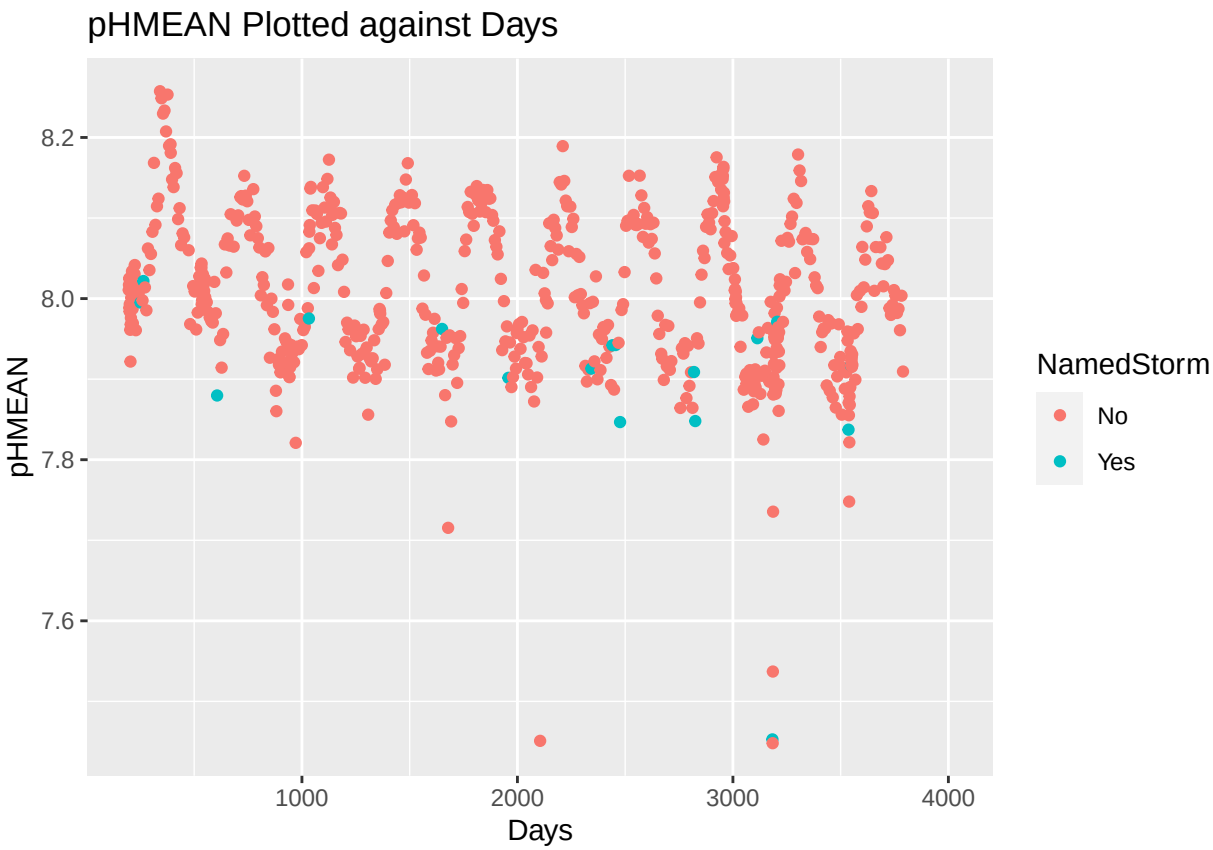
EDA

Alexandra Lawrence

Exploring pH

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



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?

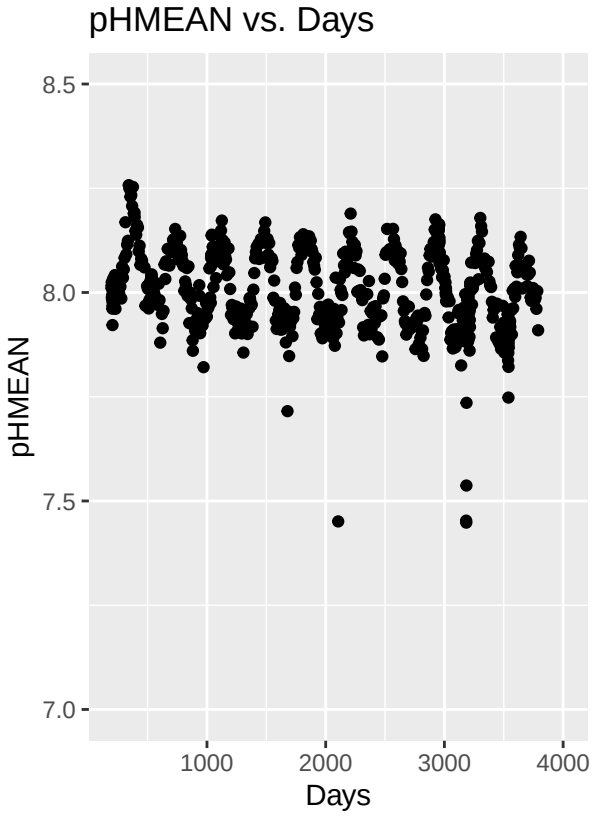
```
## # A tibble: 2 x 5
##   term          estimate  std.error statistic  p.value
##   <chr>            <dbl>      <dbl>     <dbl>    <dbl>
## 1 (Intercept)  8.05      0.00788      1021.   0.
## 2 Days        -0.0000220 0.00000342     -6.44 2.37e-10
```

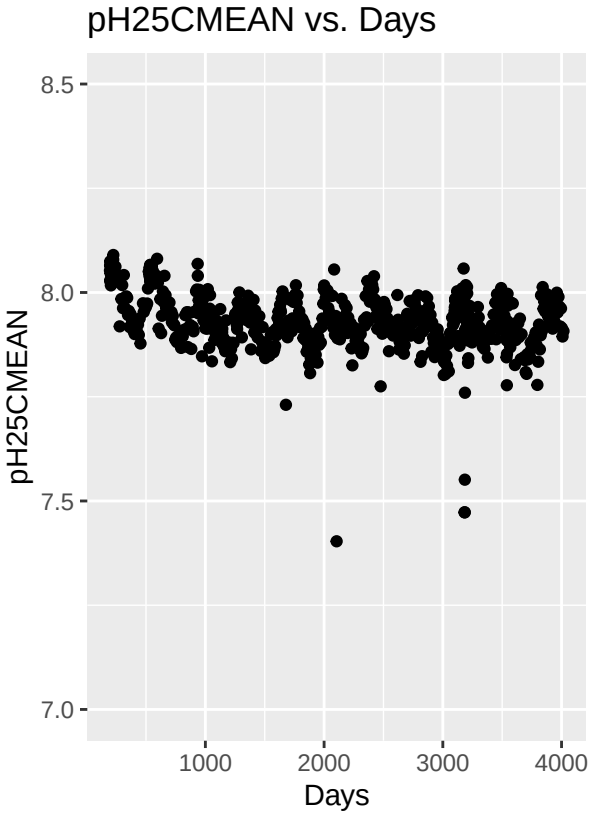
```
##     pHMEAN      Date
## 1 7.448091 9/20/2018
```

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

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

```
## Warning: Removed 10 rows containing missing values (geom_point).
```

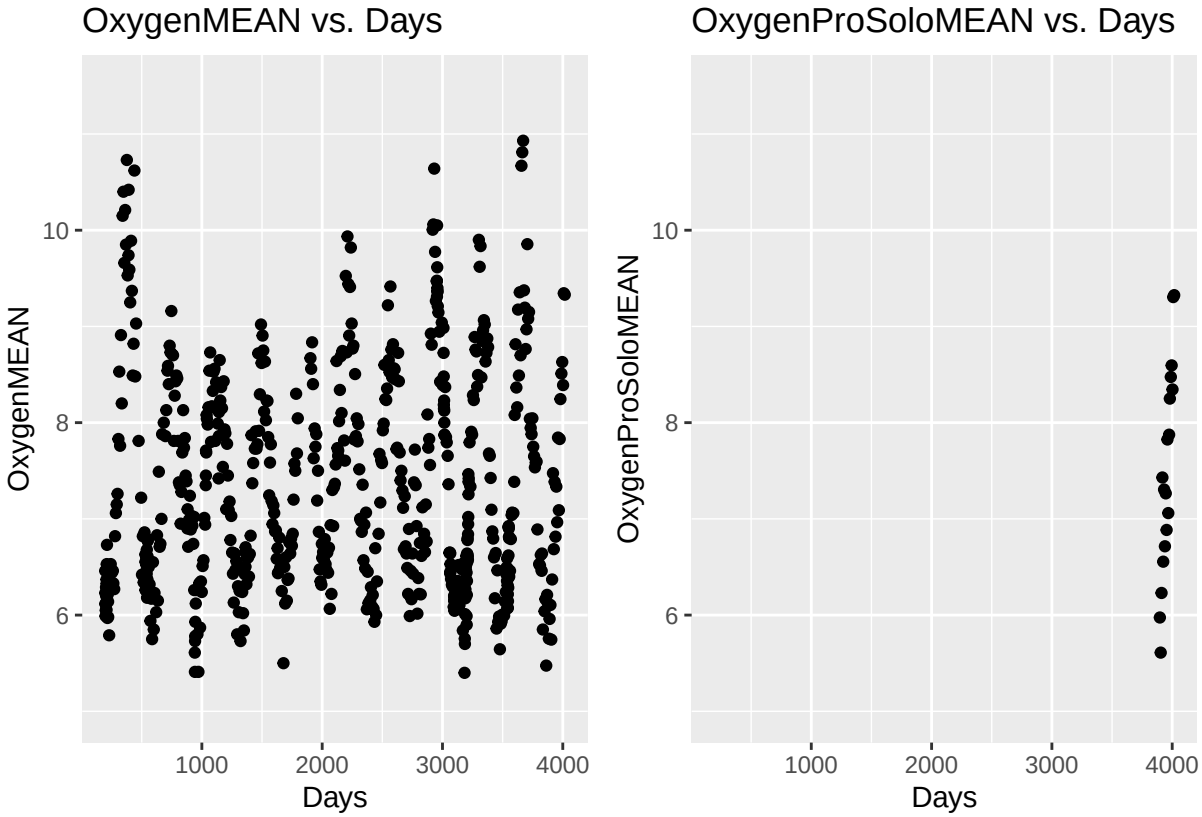




Exploring Oxygen

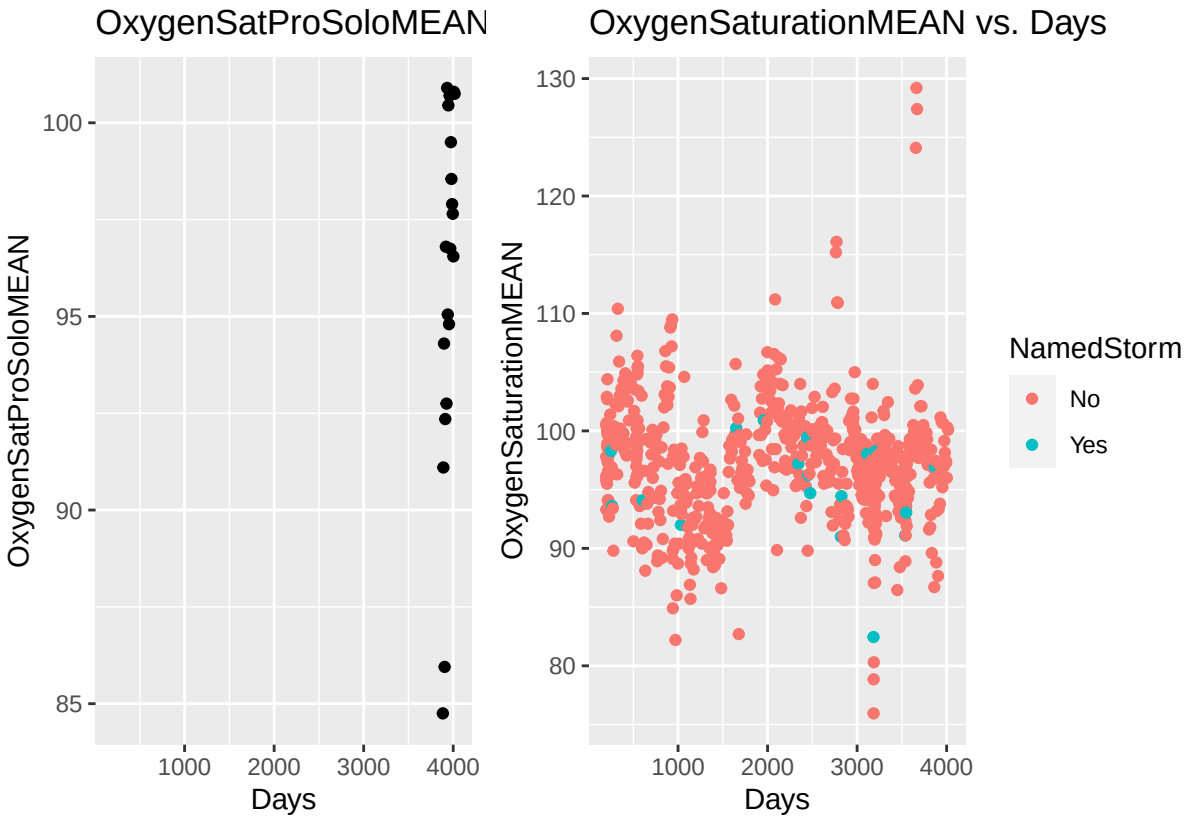
```
## Warning: Removed 31 rows containing missing values (geom_point).
```

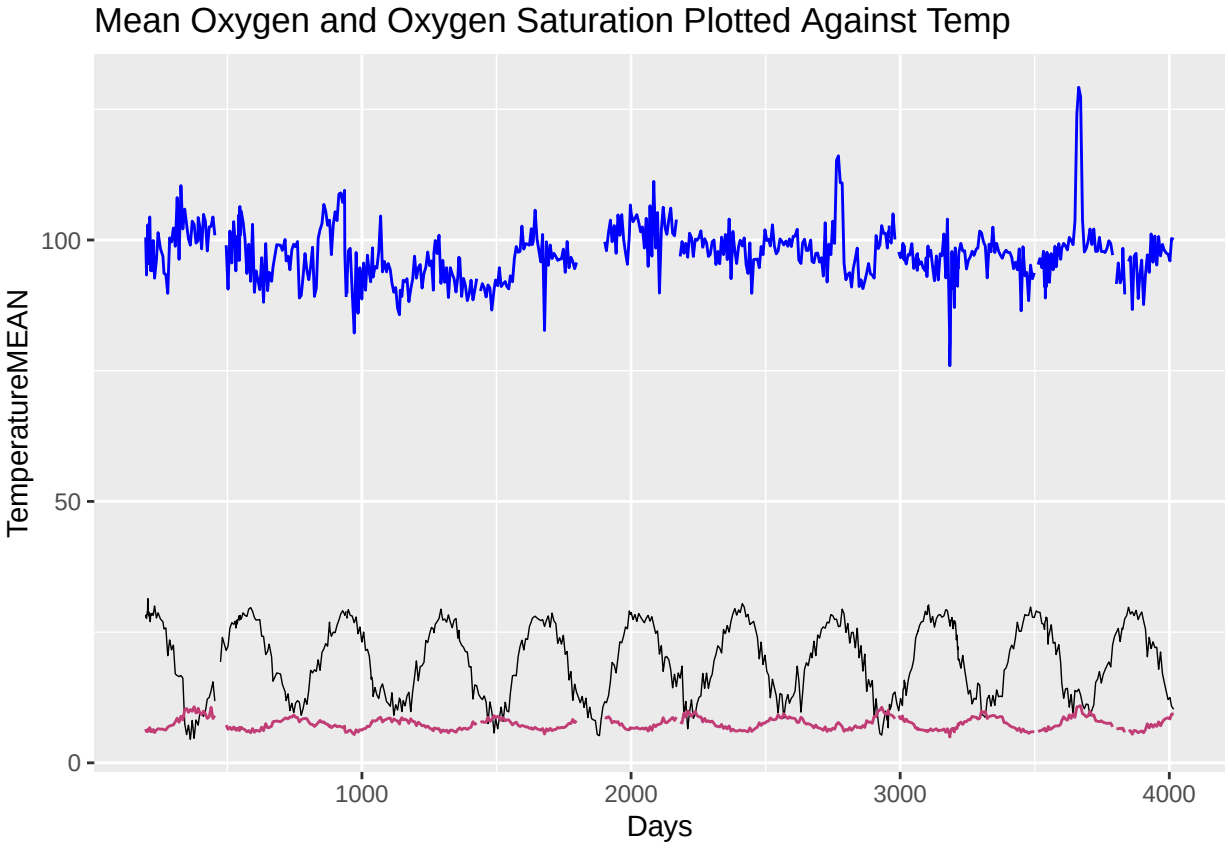
```
## Warning: Removed 661 rows containing missing values (geom_point).
```




```
## Warning: Removed 659 rows containing missing values (geom_point).
```

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





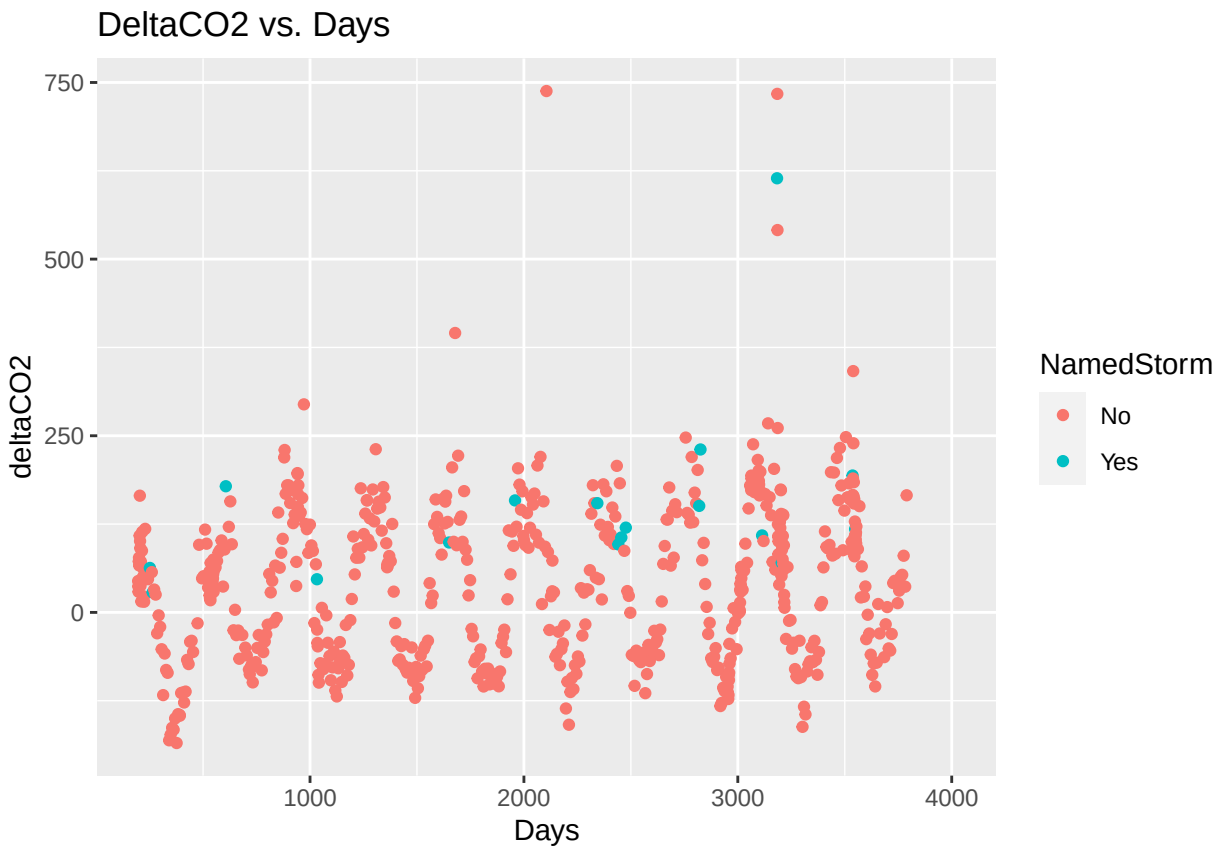

```
##   OxygenSaturationMEAN       Date OxygenMEAN TemperatureMEAN
```

```
## 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).
```

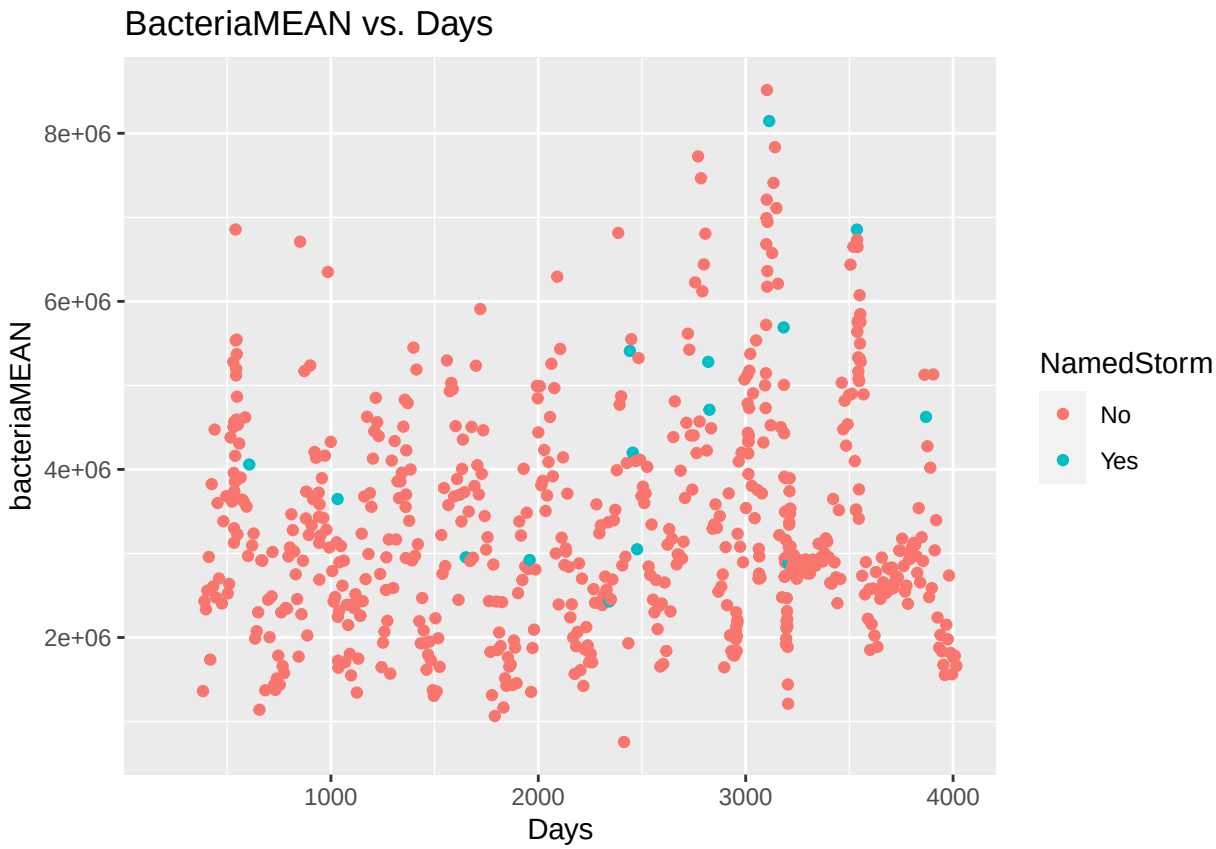



```
##   deltaCO2      Date
## 1 737.8547 10/7/2015
```

Maximum value happened right around Hurrricane Joaquin

Bacteria

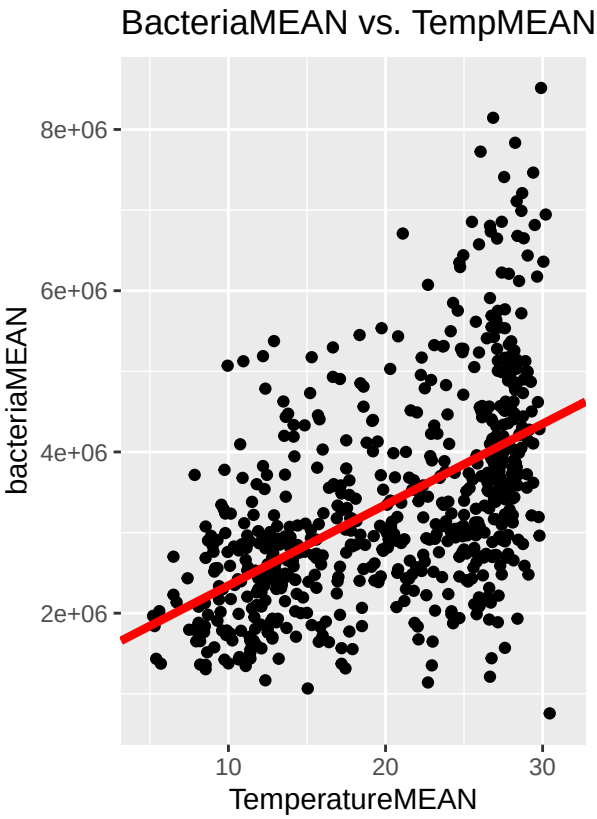
```
## Warning: Removed 54 rows containing missing values (geom_point).
```

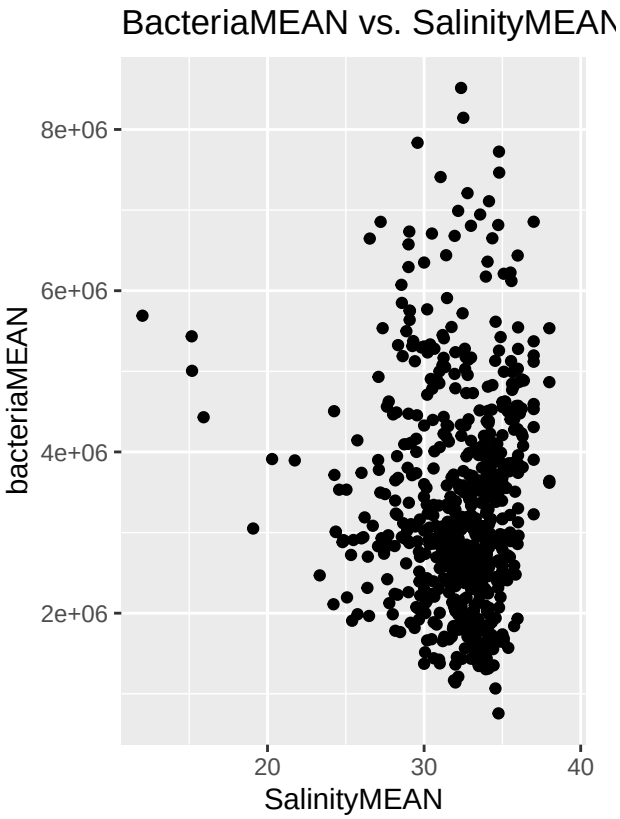



```
##   bacteriaMEAN      Date
## 1      8515000 6/30/2018
```

```
## Warning: Removed 55 rows containing missing values (geom_point).
```

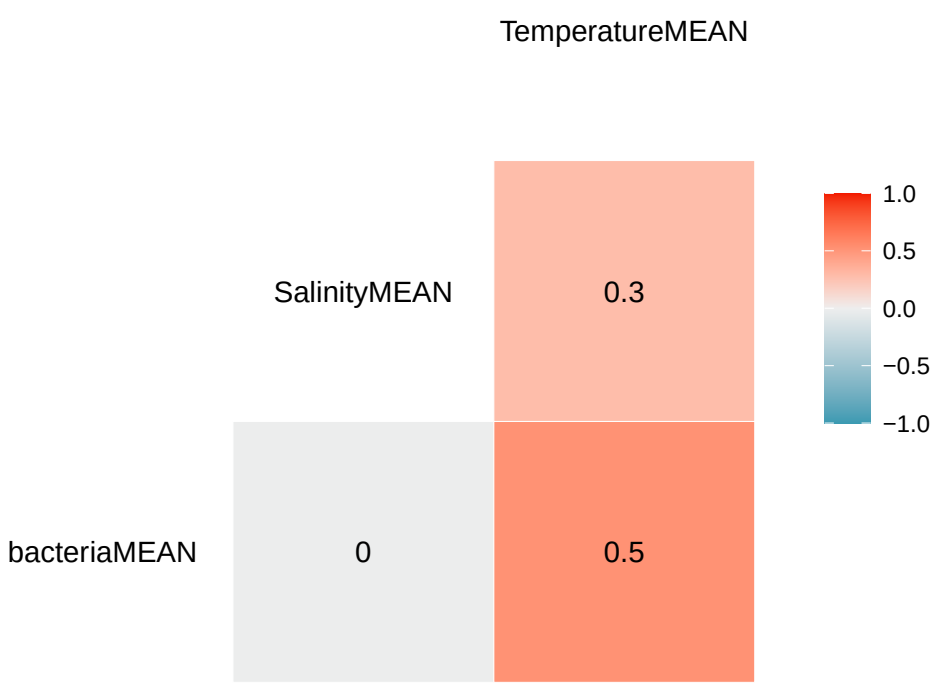
```
## Warning: Removed 55 rows containing missing values (geom_point).
```






```
## # A tibble: 2 x 5
##   term            estimate std.error statistic  p.value
##   <chr>              <dbl>     <dbl>     <dbl>    <dbl>
## 1 (Intercept)     1340431.   140008.      9.57 2.39e-20
## 2 TemperatureMEAN  100284.     6566.     15.3  5.82e-45
```

```
## # A tibble: 2 x 5
##   term          estimate std.error statistic  p.value
##   <chr>            <dbl>     <dbl>     <dbl>    <dbl>
## 1 (Intercept)  3508648.   548596.      6.40  3.14e-10
## 2 SalinityMEAN   -4519.    16991.    -0.266 7.90e- 1
```



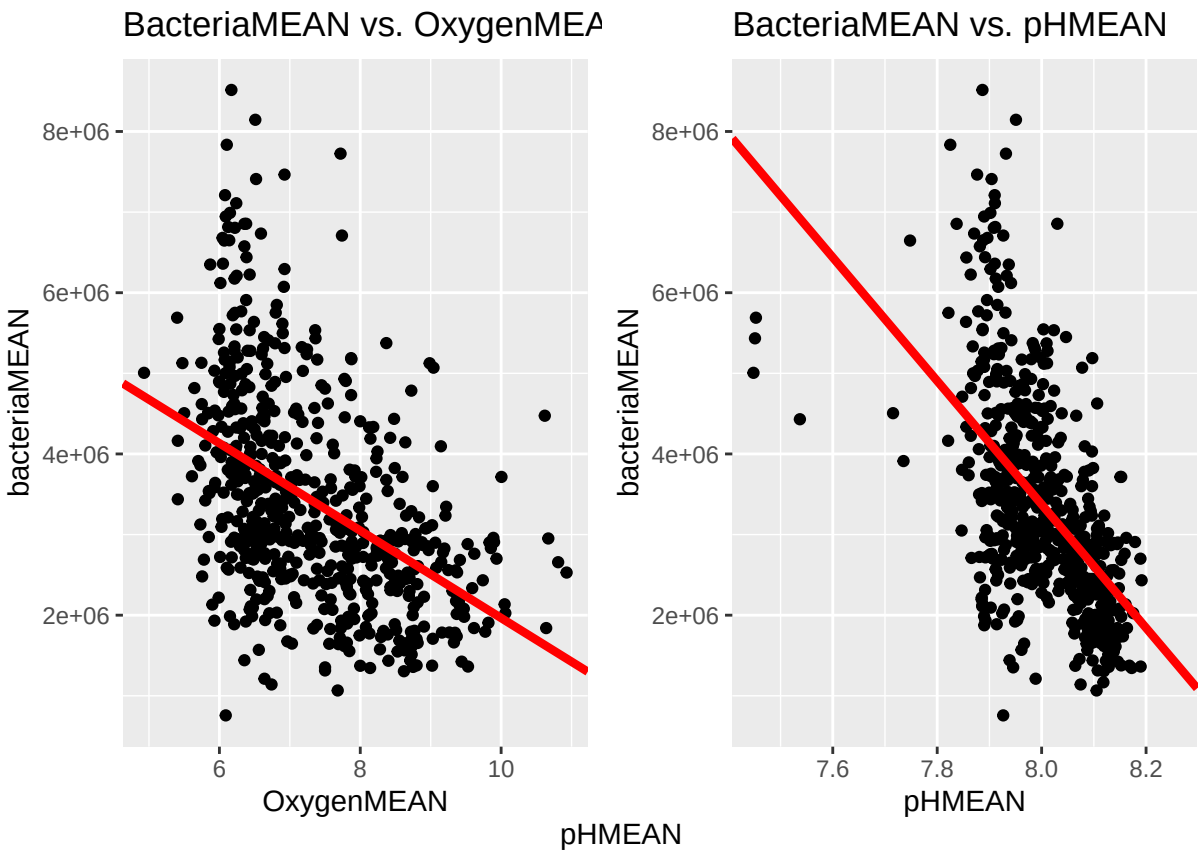
The p-value for the linear regression shows that there is a relationship between mean bacteria and mean temperature, but there does not appear to be a strong correlation between the two

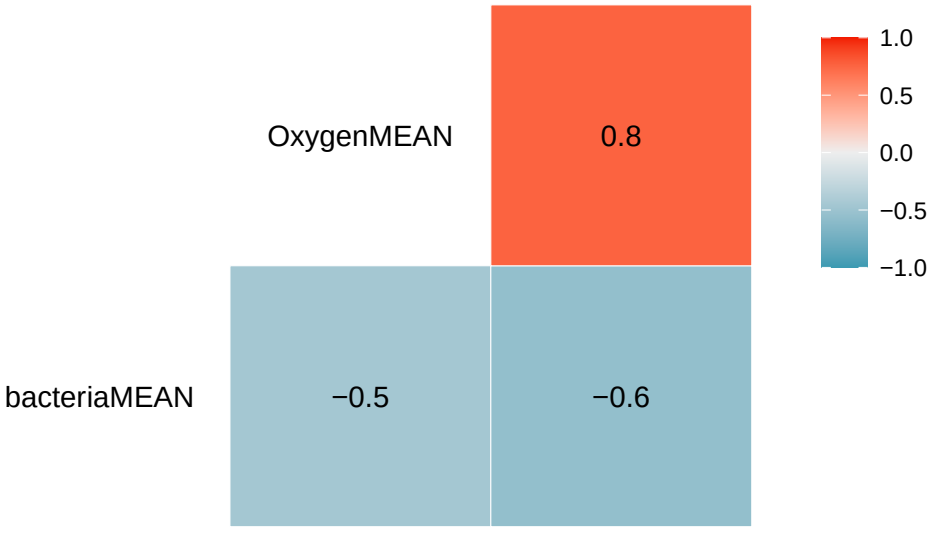
```
## # A tibble: 2 x 5
##   term         estimate std.error statistic  p.value
##   <chr>           <dbl>     <dbl>     <dbl>    <dbl>
## 1 (Intercept) 64645987.  3786766.      17.1 2.82e-53
## 2 pHMEAN      -7658991.   473416.     -16.2 6.90e-49
```

```
## # A tibble: 2 x 5
##   term        estimate std.error statistic  p.value
##   <chr>          <dbl>     <dbl>     <dbl>    <dbl>
## 1 (Intercept) 7384236.   319543.      23.1 5.43e-85
## 2 OxygenMEAN  -541933.    43042.     -12.6 2.01e-32
```

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

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





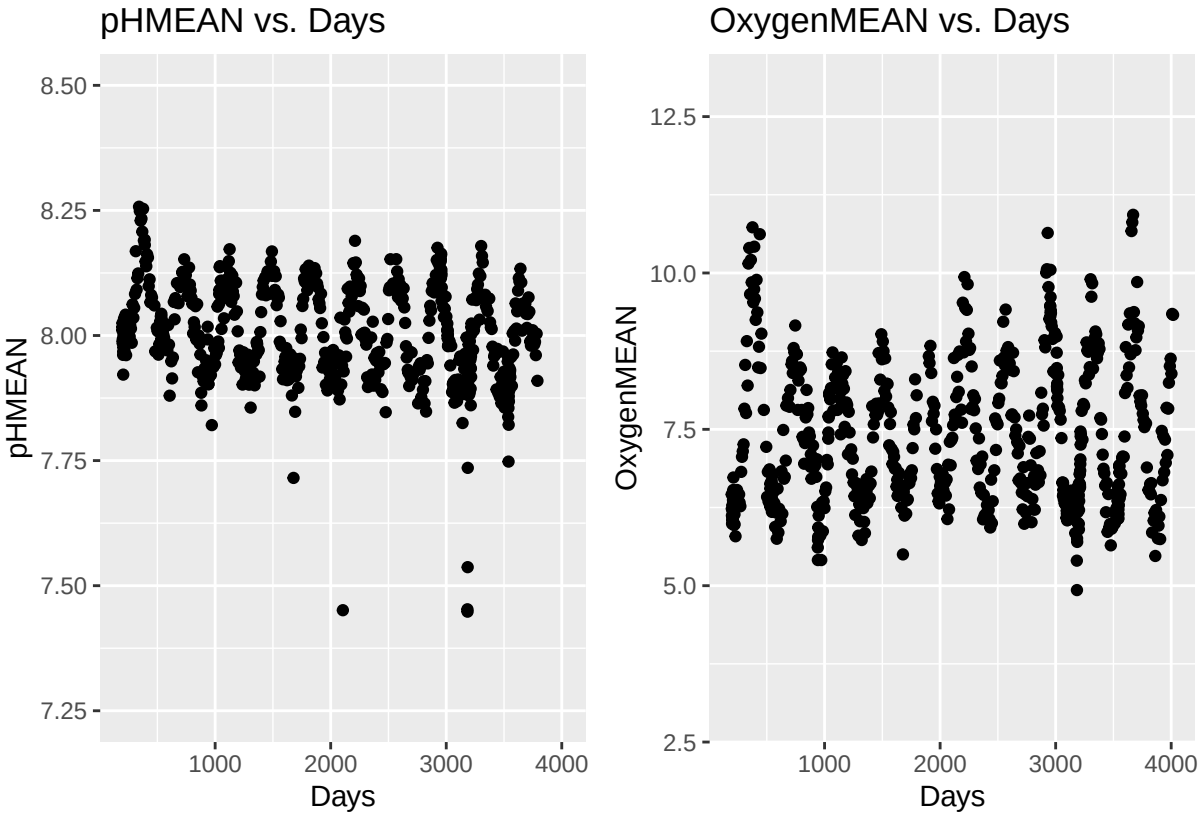
Again, there does not appear to be a strong correlation. between mean bacteria. and pH and oxygen, despite the p-values of the linear models suggesting otherwise

Looking at pH alongside Oxygen

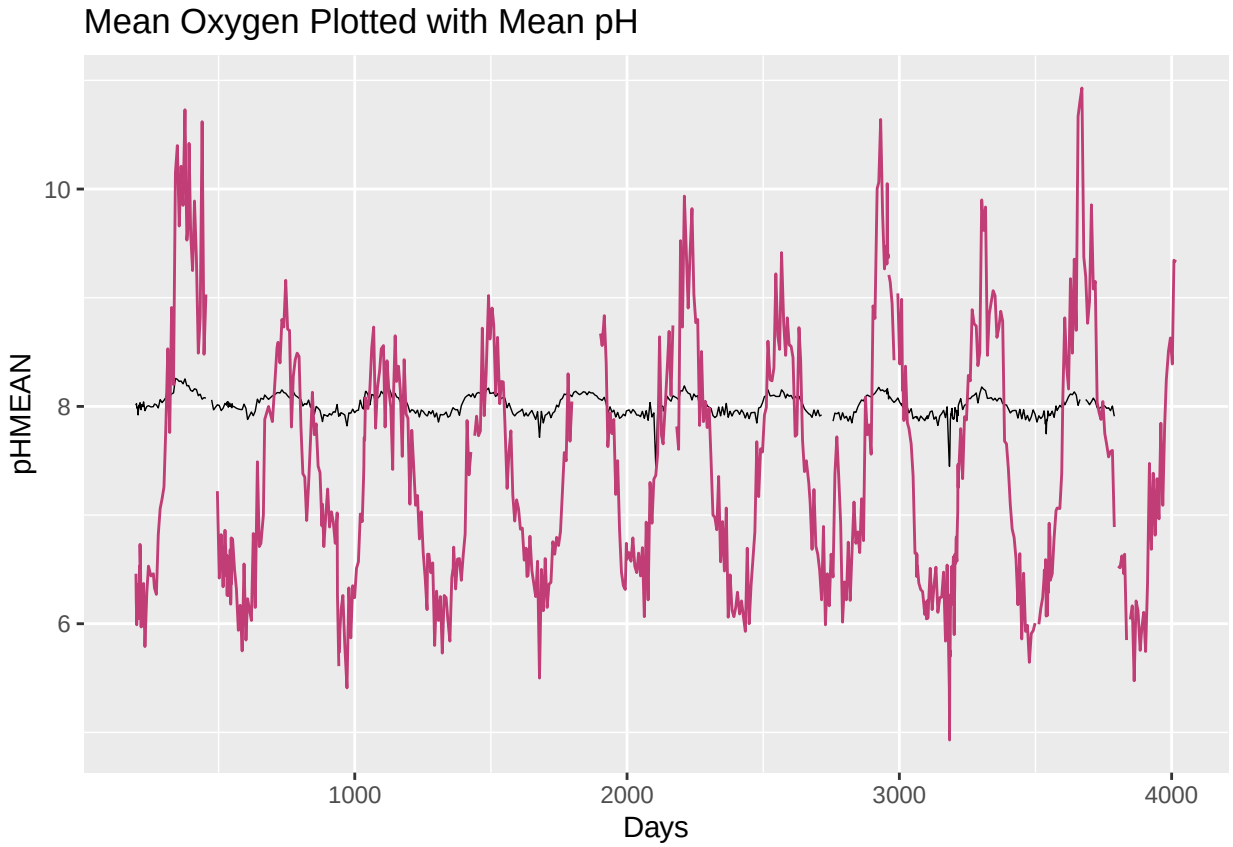
```
## Scale for 'y' is already present. Adding another scale for 'y', which will
## replace the existing scale.
## Scale for 'y' is already present. Adding another scale for 'y', which will
## replace the existing scale.
```

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

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

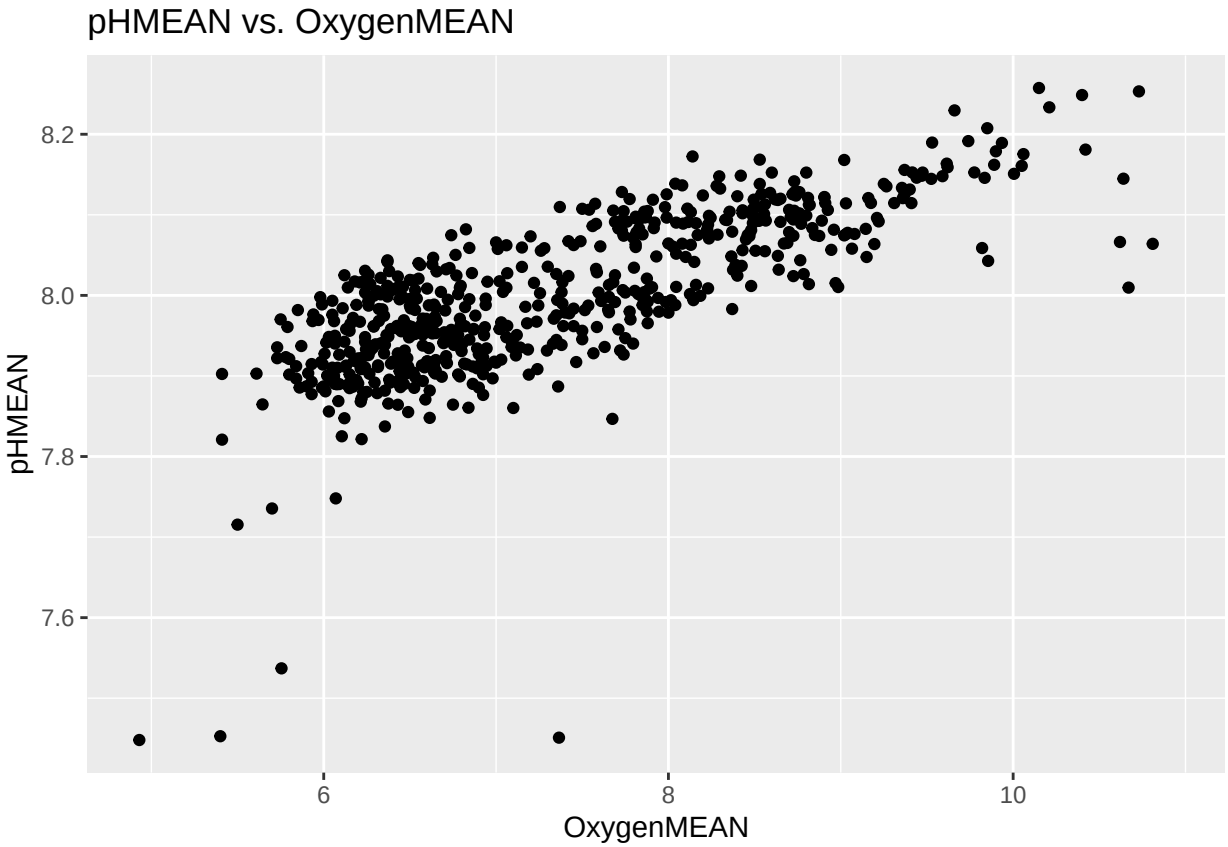



```
## Warning: Removed 33 row(s) containing missing values (geom_path).
```

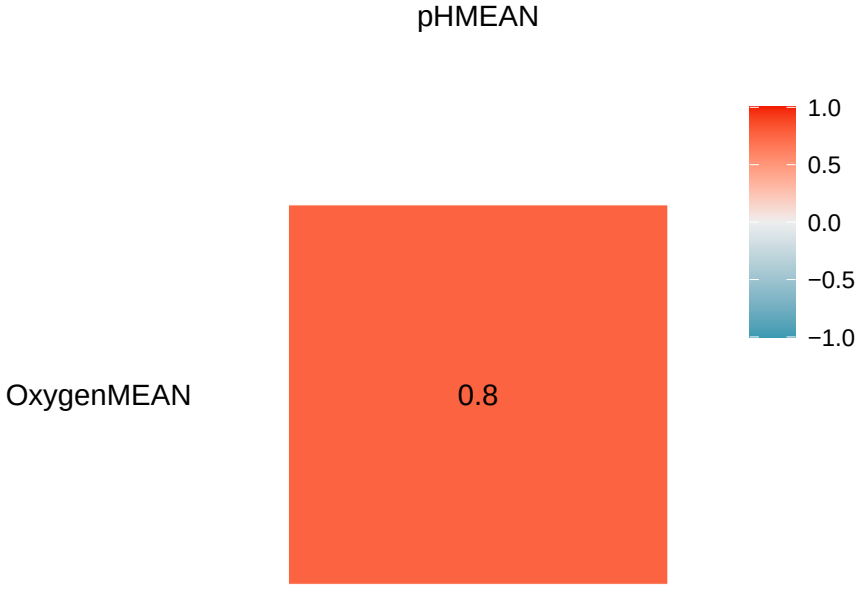


The mean oxygen measured by ProODO appears to follow a similar pattern of peaks as the mean pH calculated using CO2SYS.

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



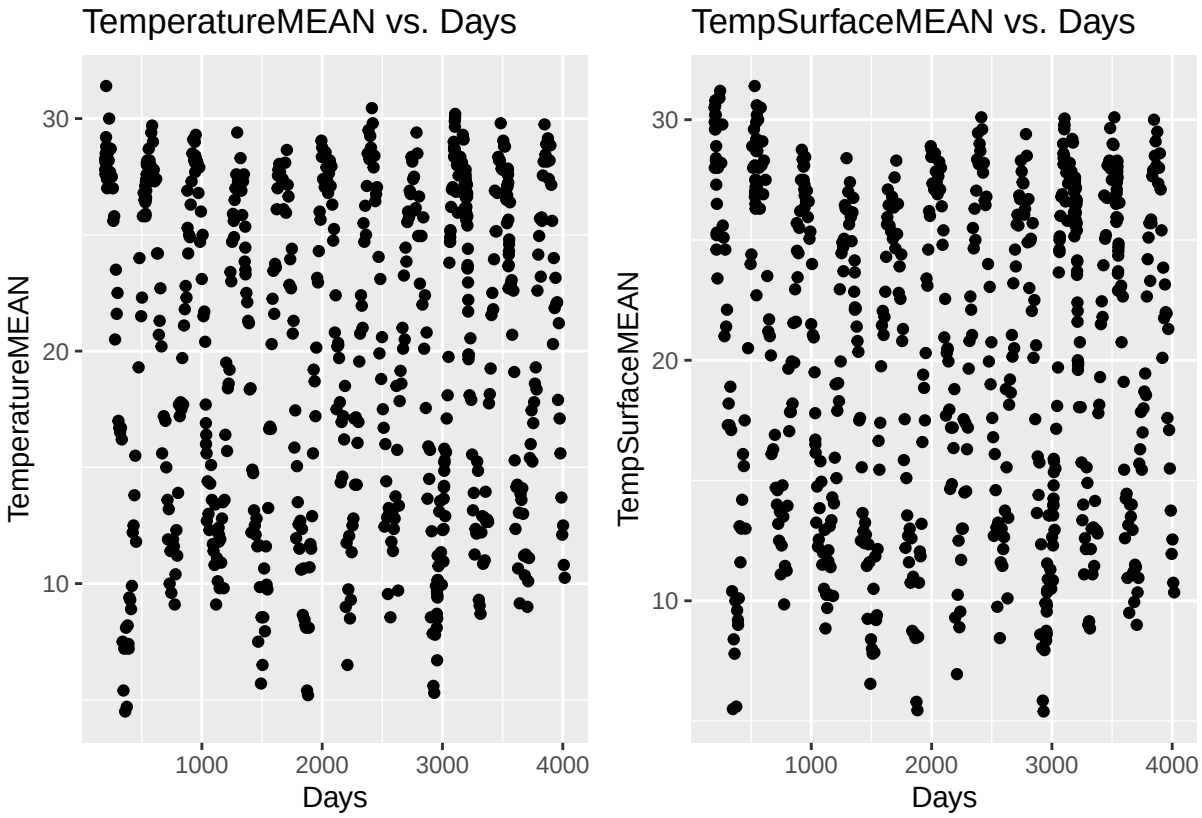

```
## # A tibble: 2 x 5
##   term        estimate std.error statistic   p.value
##   <chr>          <dbl>     <dbl>     <dbl>     <dbl>
## 1 (Intercept)   7.52     0.0167      449.  0.
## 2 OxygenMEAN    0.0651   0.00225      28.9 2.97e-116
```



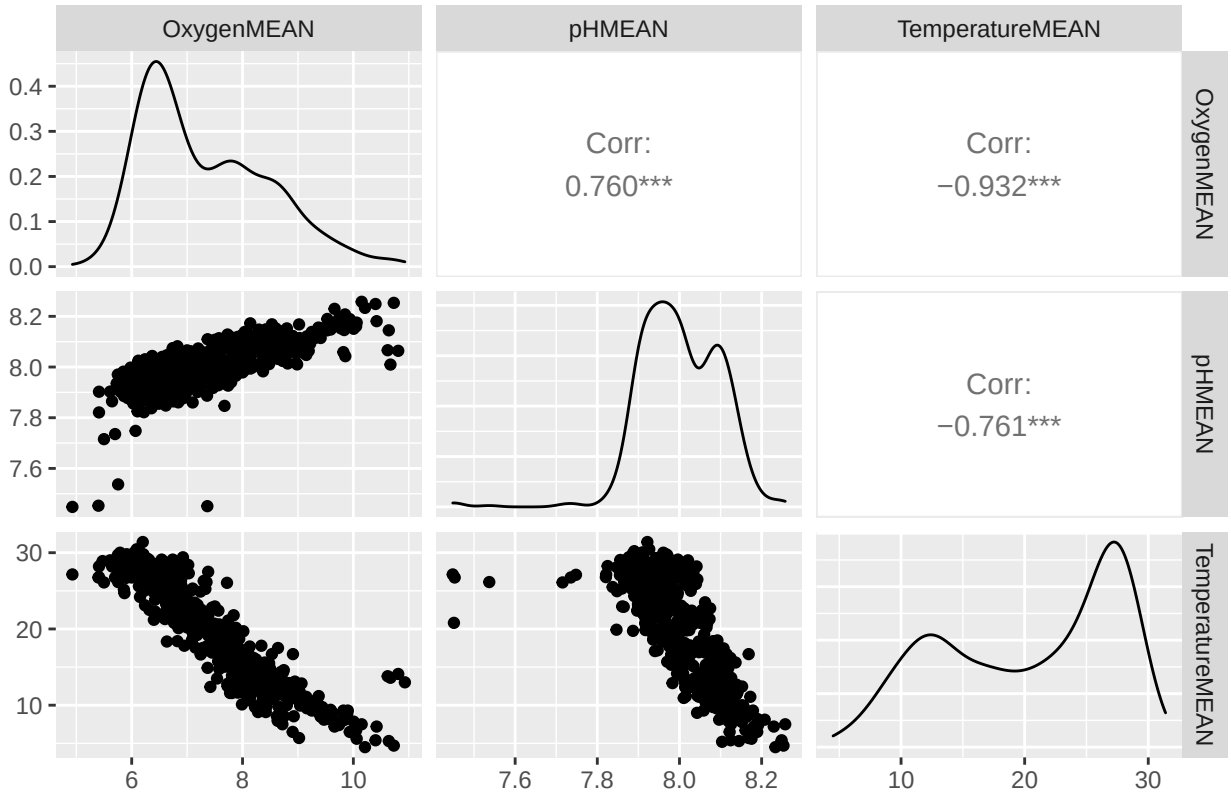
There appears to be a pretty strong positive correlation between the mean oxygen and pH levels.

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

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

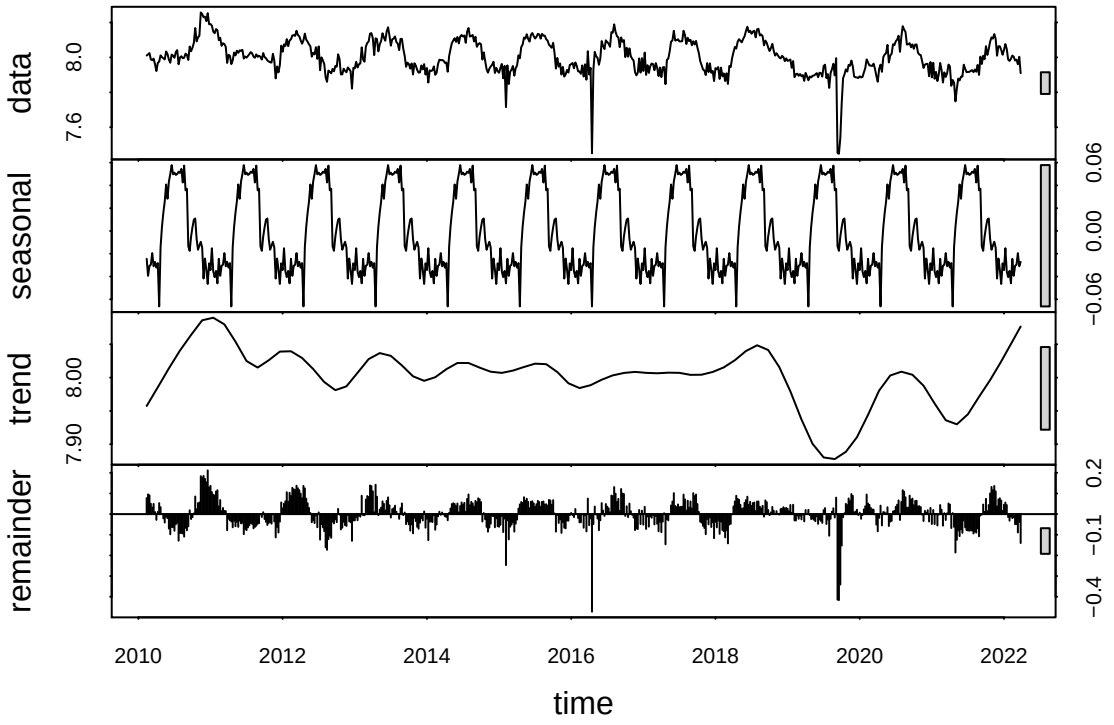


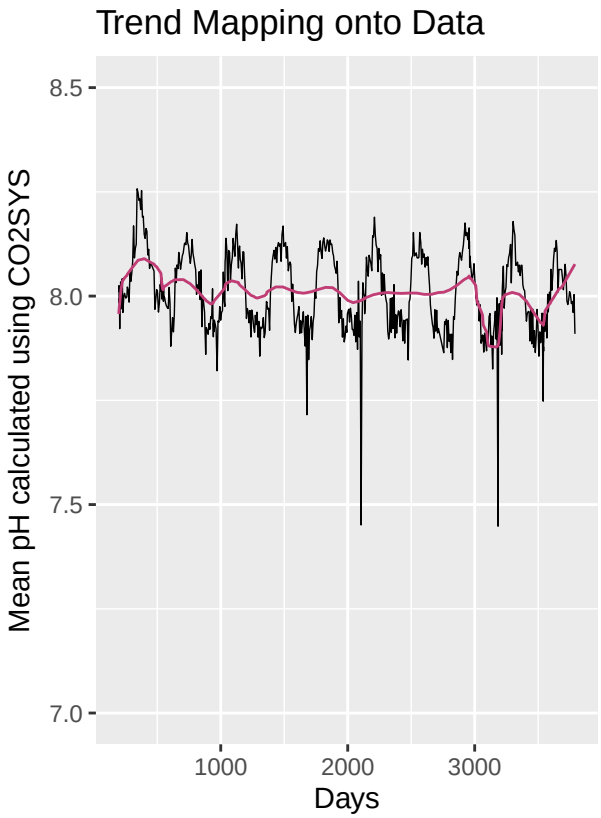

```
## Warning: Removed 30 rows containing non-finite values (stat_density).

## 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).
```

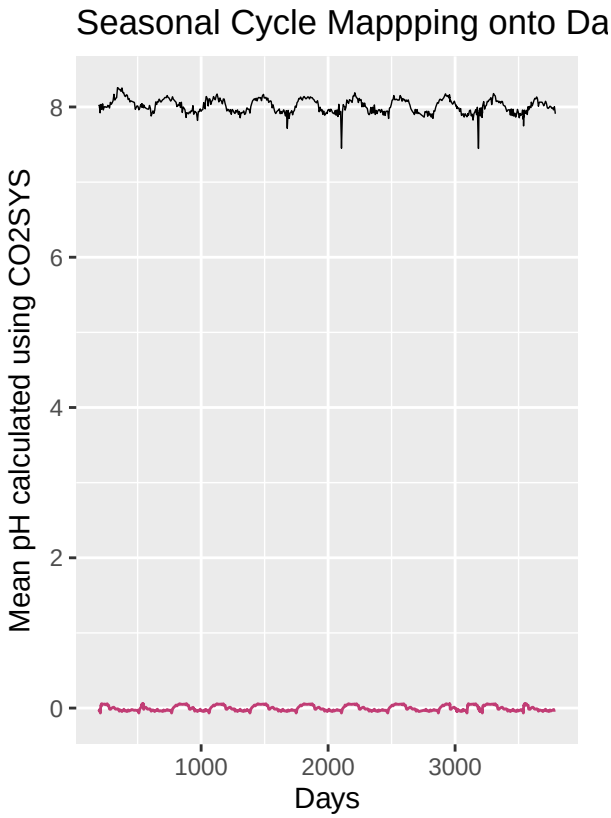

OxygenMEAN
pHMEAN
TemperatureMEAN
Corr:
0.760***
Corr:
−0.932***
Corr:
−0.761***
0.4
0.3
0.2
0.1
0.0
8.2
8.0
7.8
7.6
30
20
10
6
8
10
7.6
7.8
8.0
8.2
10
20
30

TIME SERIES

pH Mean



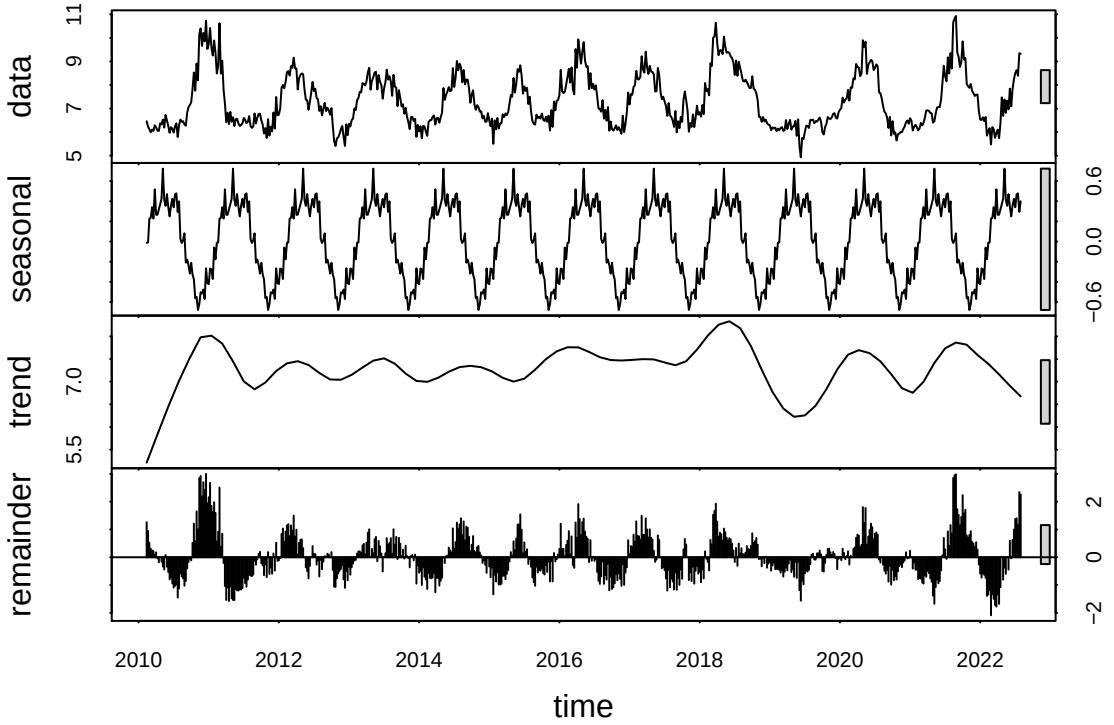


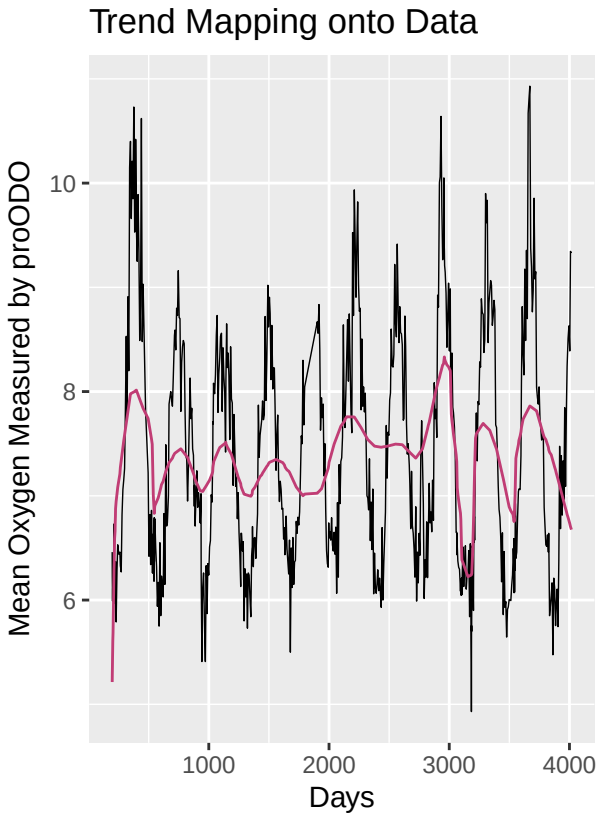


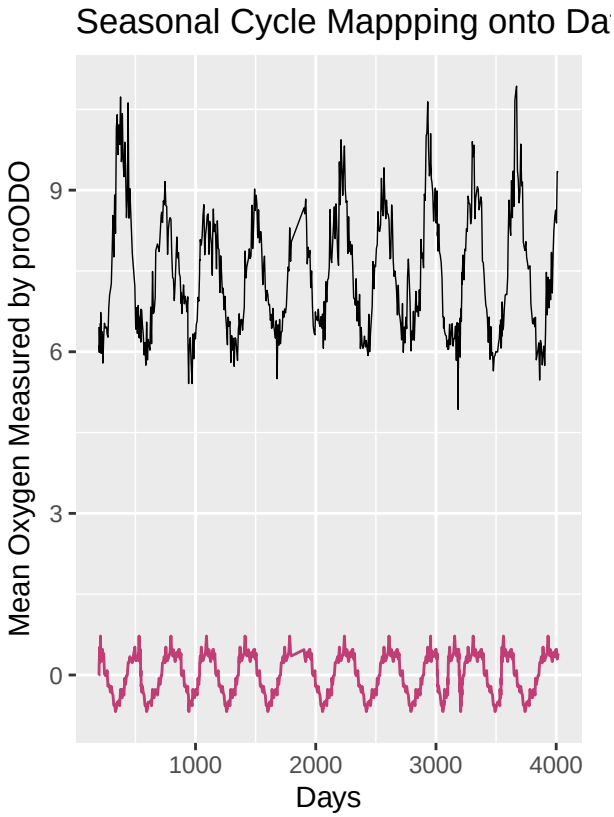

```
## tau = -0.0933, 2-sided pvalue =0.0021753
```

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

Oxygen








```
## tau = 0.0724, 2-sided pvalue =0.01554
```