Ammonia Model

Summer Heschong

2025-03-31

Setup

```
#load packages
library(here)
```

```
## here() starts at /Users/summerheschong/Stats_Group_Project
```

```
library(tidyverse)
```

```
## -- Attaching core tidyverse packages ------------------------ tidyverse 2.0.0 --
## v dplyr     1.1.4     v readr     2.1.5
## v forcats   1.0.0     v stringr   1.5.1
## v ggplot2   3.5.1     v tibble    3.2.1
## v lubridate 1.9.4     v tidyr     1.3.1
## v purrr     1.0.2
```

```
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

```
library(lme4)
```

```
## Loading required package: Matrix
##
## Attaching package: 'Matrix'
##
## The following objects are masked from 'package:tidyr':
##
##     expand, pack, unpack
```

```
wetland <- read.csv(here('Data/Processed/Combined_Data_NArm.csv'))
```

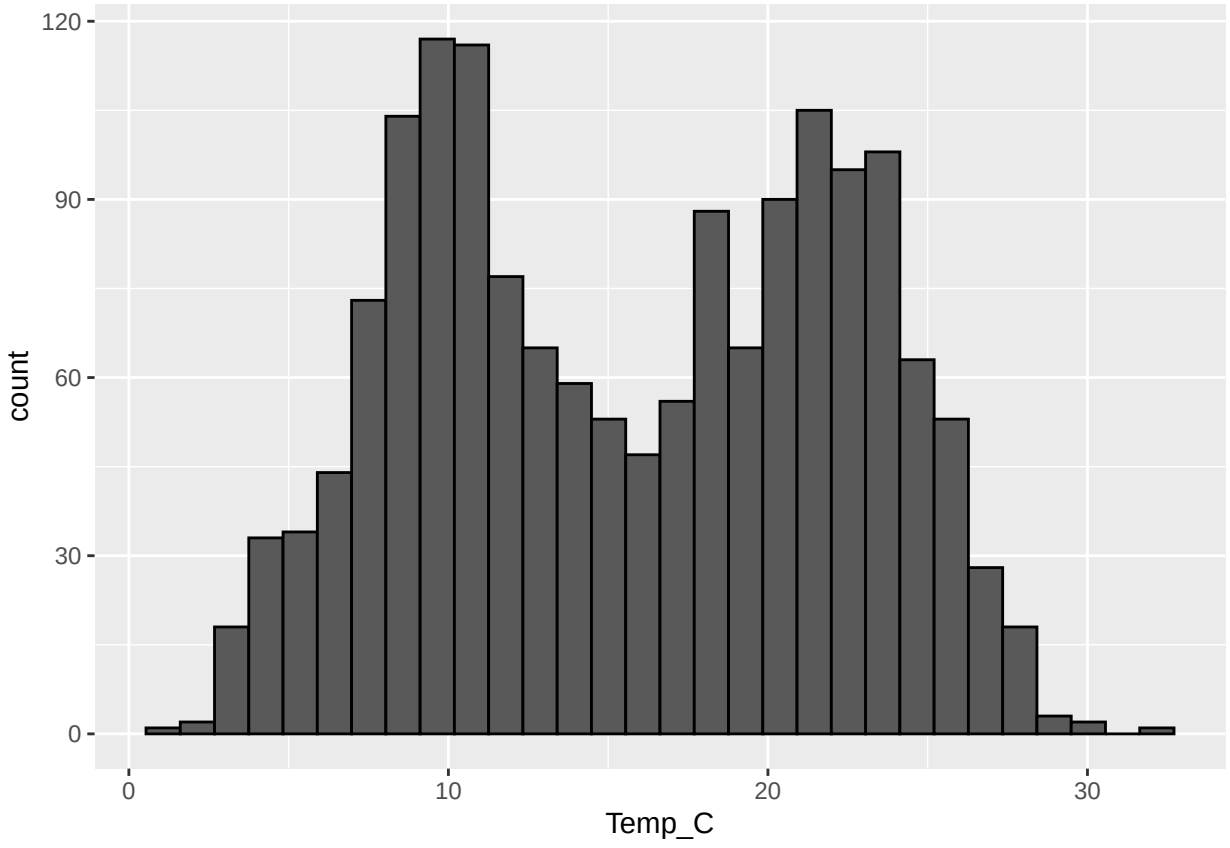
Research Question: What predicts ammonium levels in wetlands

Examine Data

display raw counts and distributions of data

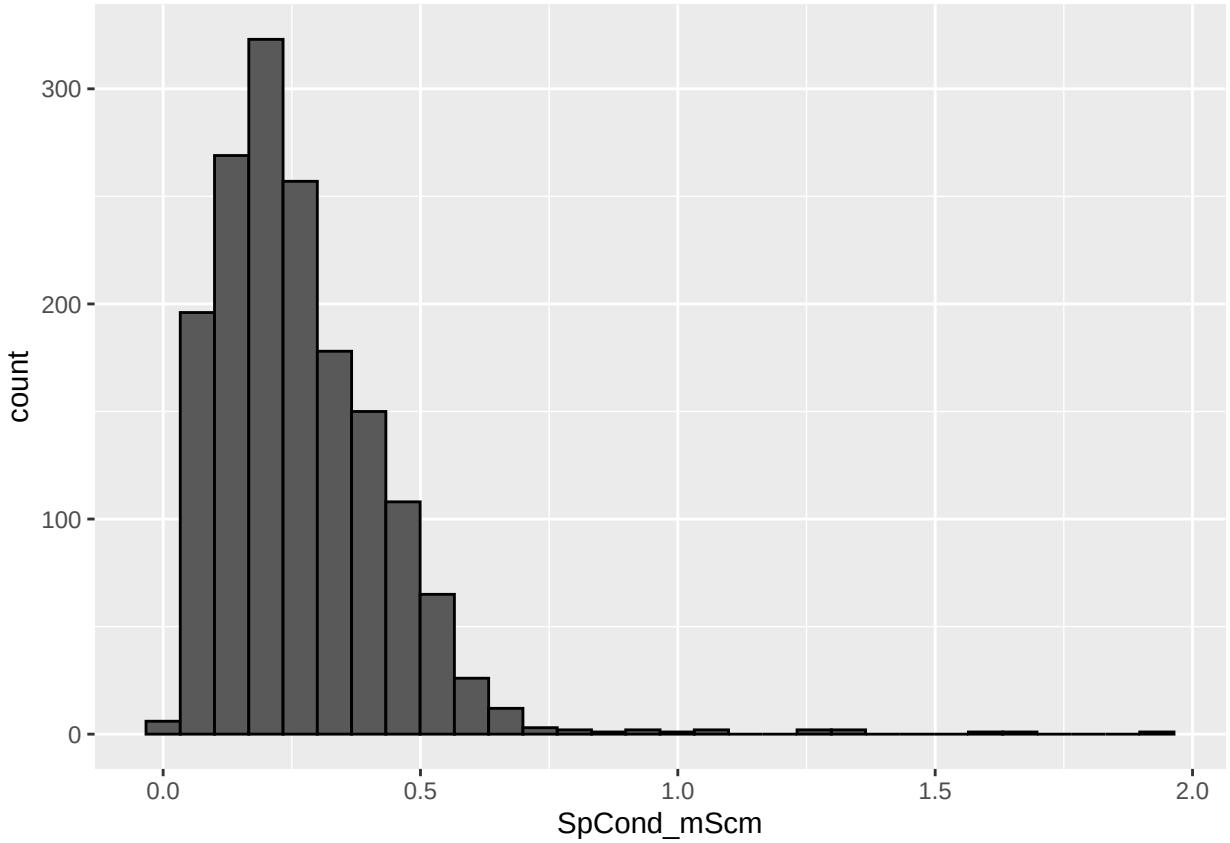
```
#remove outliers: TDS 2262, row 1039; DO 224, row 840; TSS 516, row 641; fecal_coliform rows 1545 and 1
wetland <- wetland [-c(147, 166, 641, 793, 840, 1035, 1039, 1545, 1549, 1614),]

#create histograms
ggplot(wetland, aes(x = Temp_C)) +
geom_histogram(color = 'black')
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



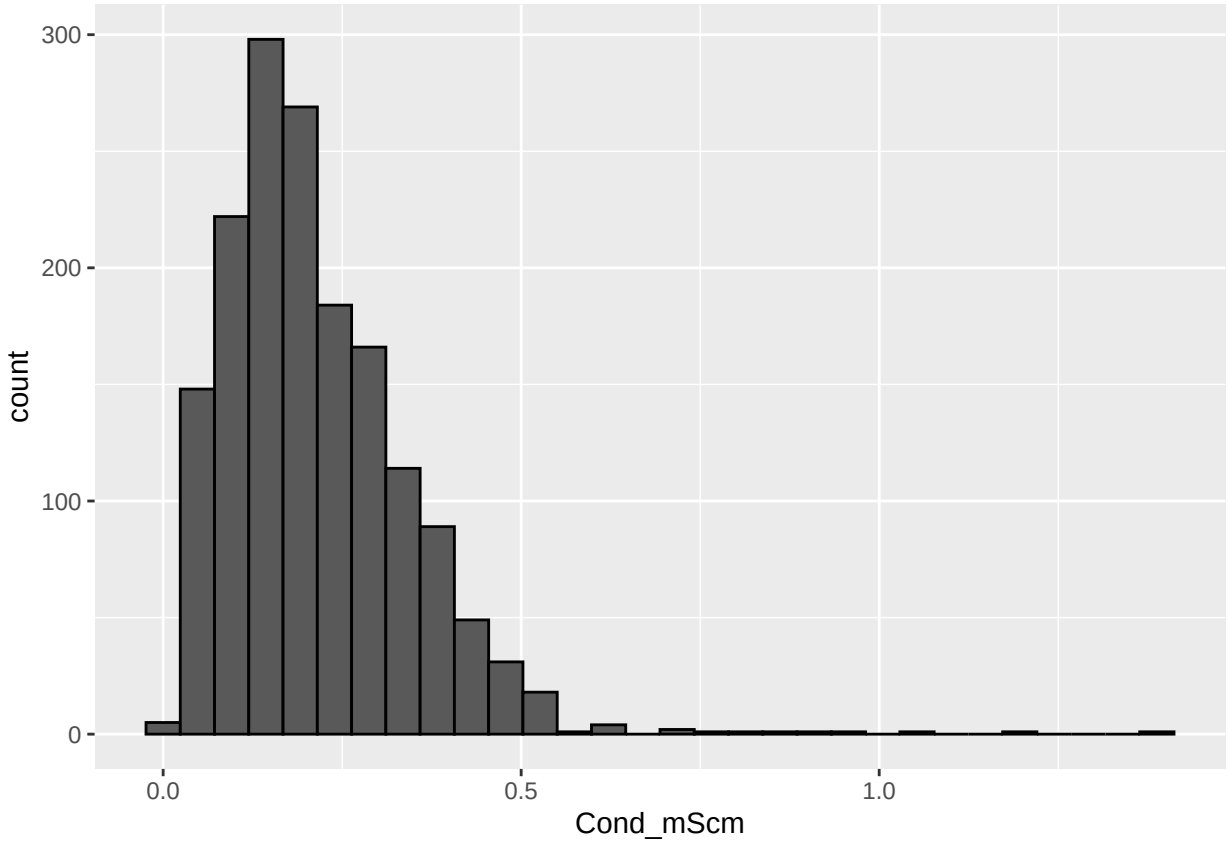

```
ggplot(wetland, aes(x = SpCond_mScm)) +
geom_histogram(color = 'black')
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



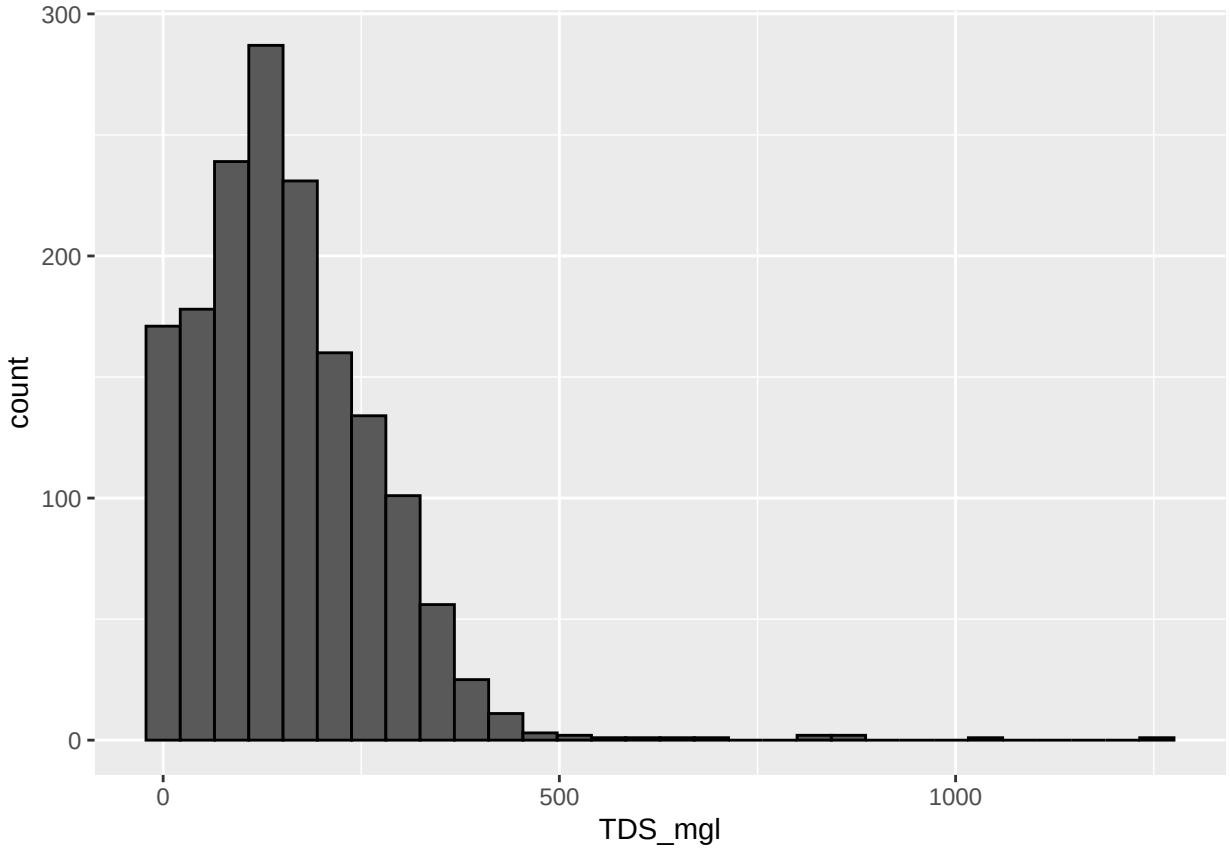

```
ggplot(wetland, aes(x = Cond_mScm)) +
geom_histogram(color = 'black')
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



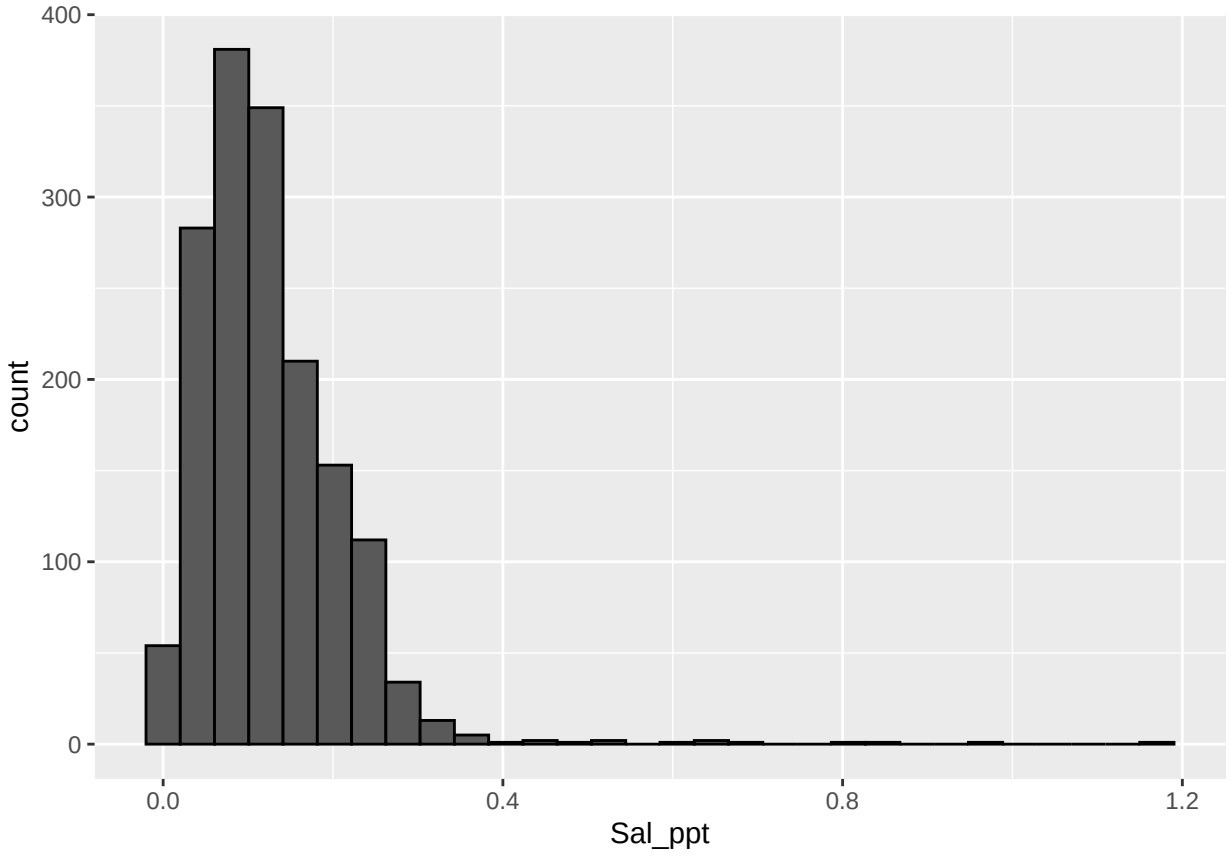

```
ggplot(wetland, aes(x = TDS_mgl)) +
geom_histogram(color = 'black')
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



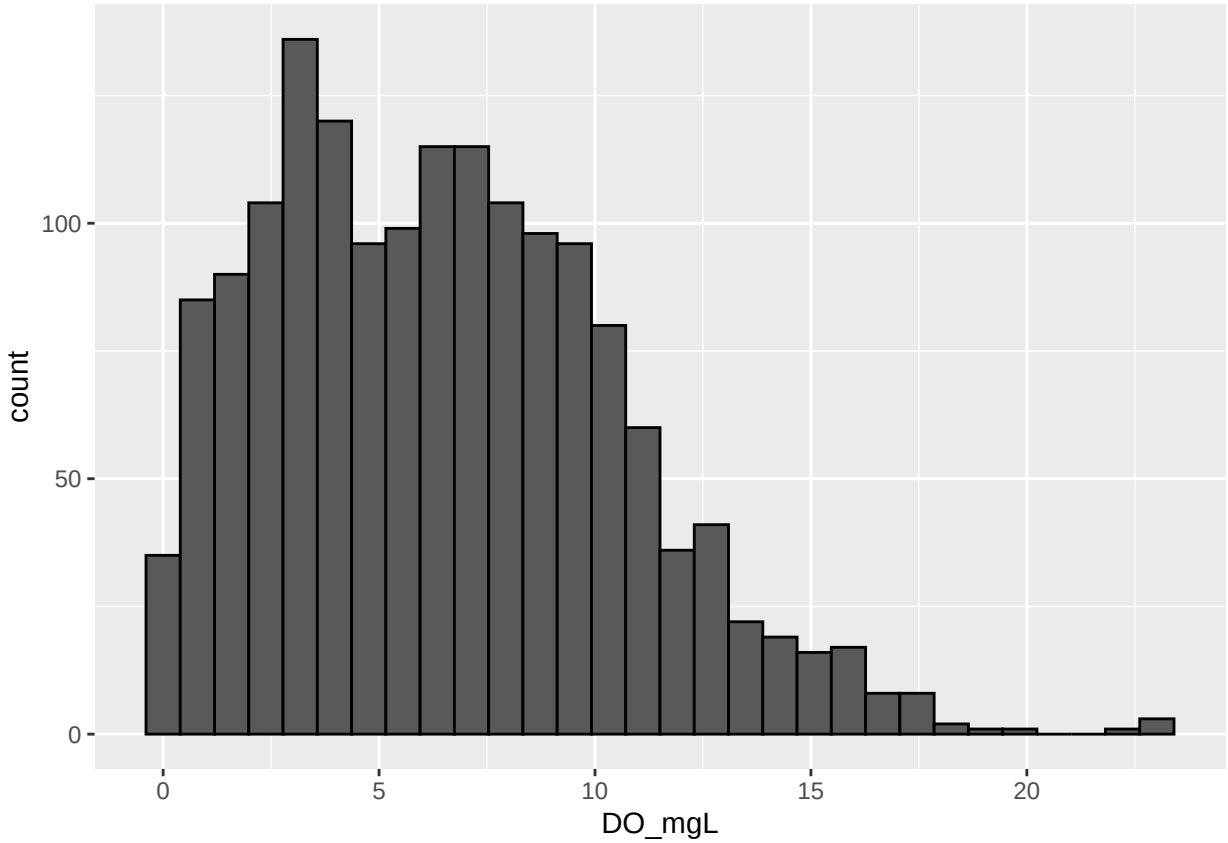

```
ggplot(wetland, aes(x = Sal_ppt)) +
geom_histogram(color = 'black')
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



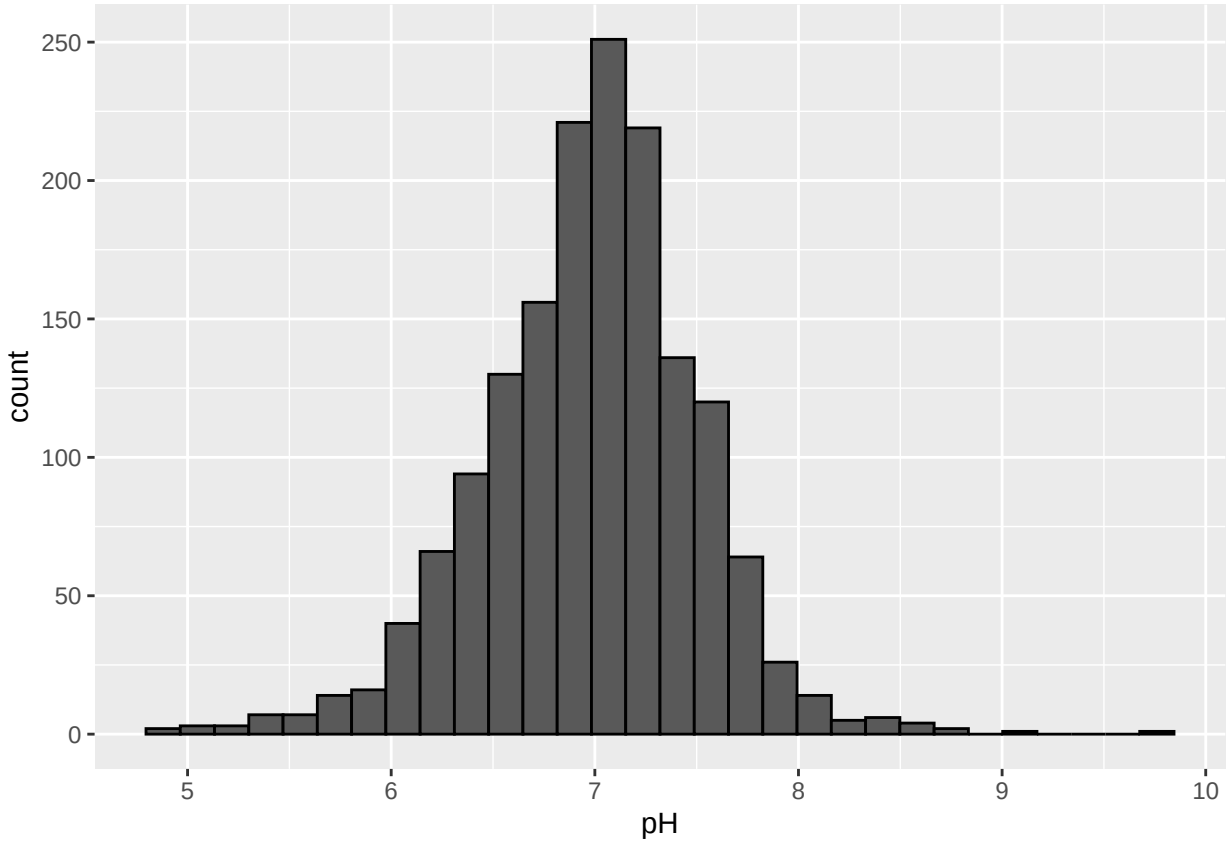

```
ggplot(wetland, aes(x = DO_mgL)) +
geom_histogram(color = 'black')
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



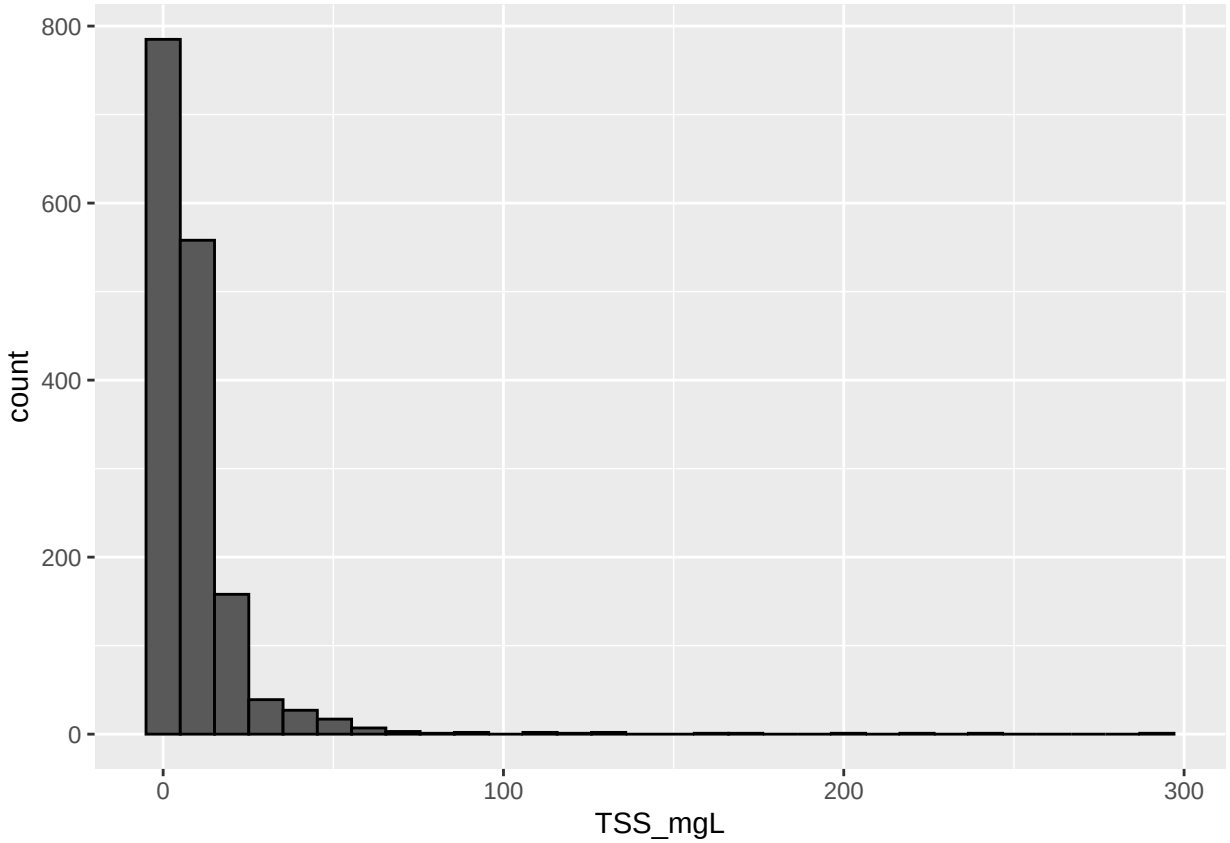

```
ggplot(wetland, aes(x = pH)) +
geom_histogram(color = 'black')
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



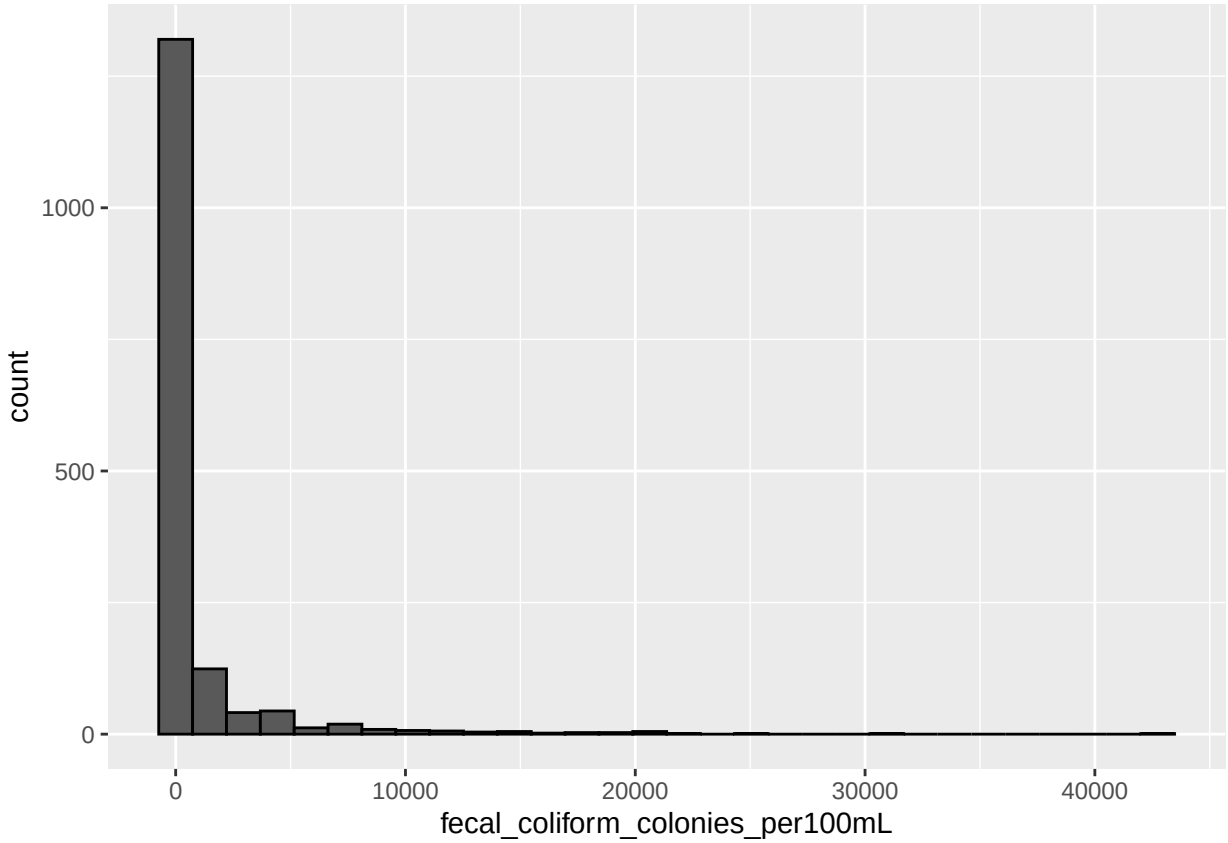

```
ggplot(wetland, aes(x = TSS_mgL)) +
geom_histogram(color = 'black')
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



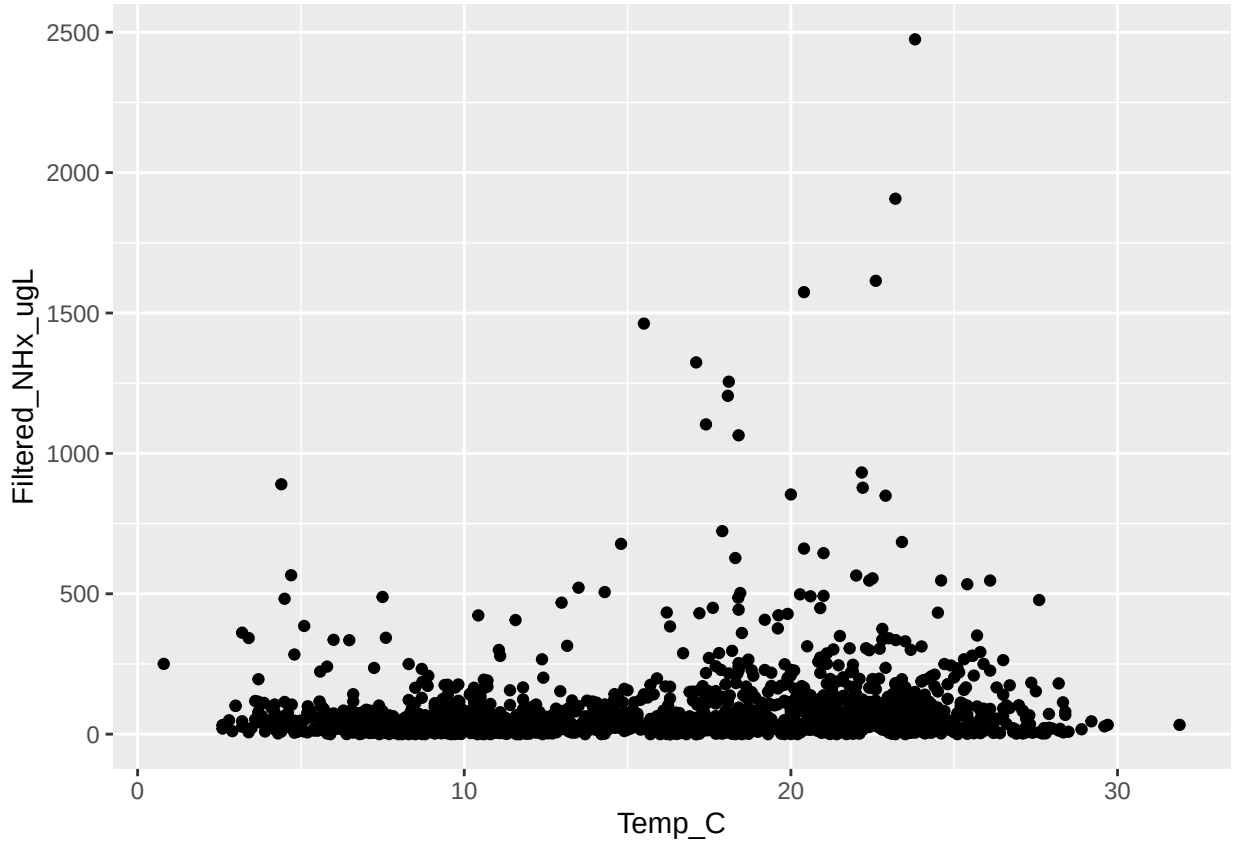

```
ggplot(wetland, aes(x = fecal_coliform_colonies_per100mL)) +
geom_histogram(color = 'black')
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

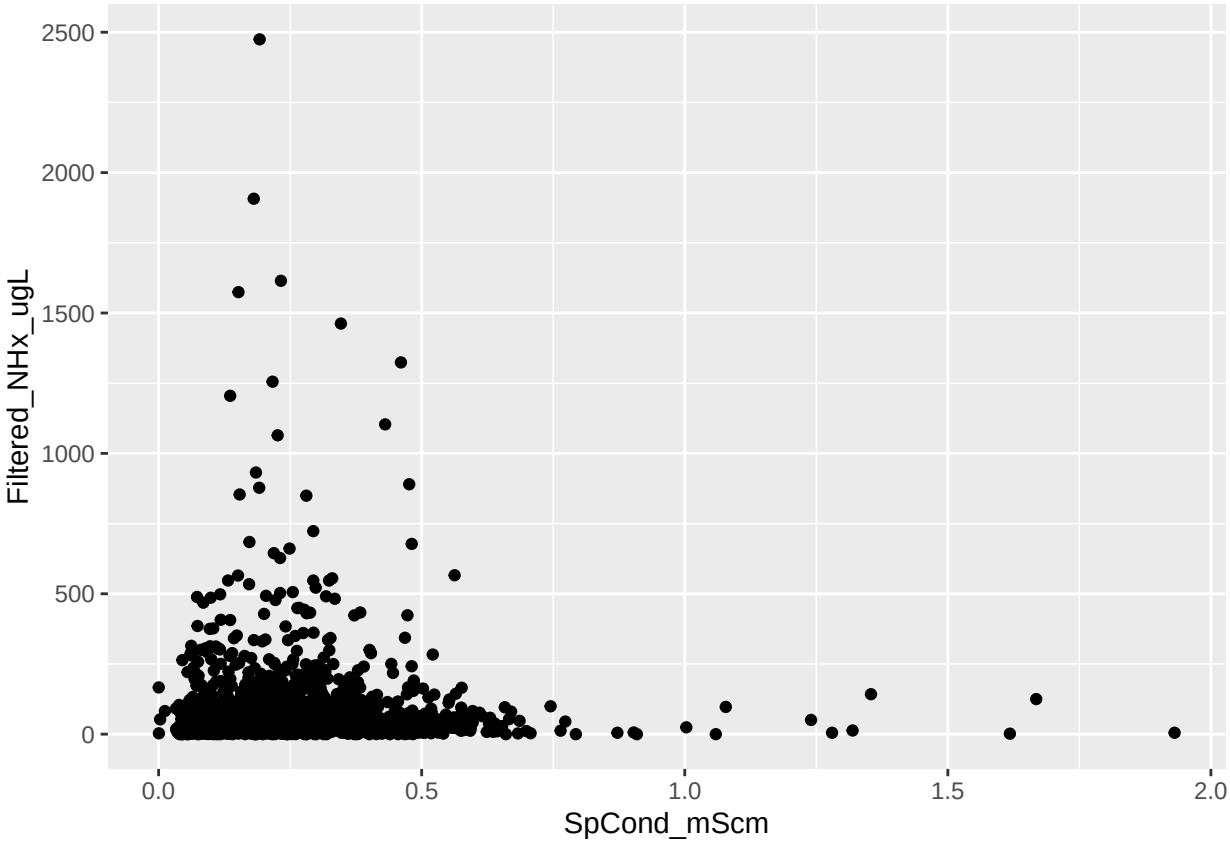


display relationships between predictor variables and outcome variable

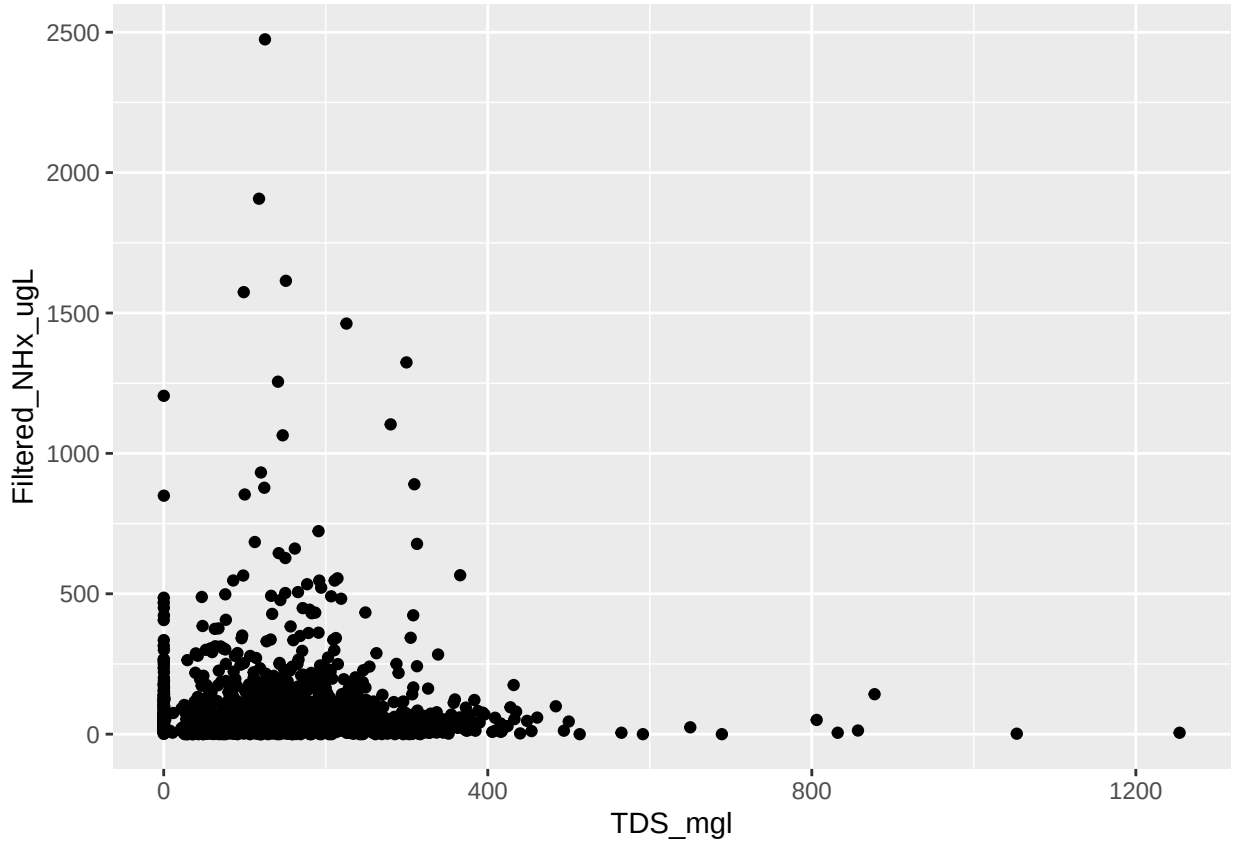
```
#create scatterplots
ggplot(wetland, aes(x = Temp_C, y = Filtered_NHx_ugL)) +
geom_point()
```



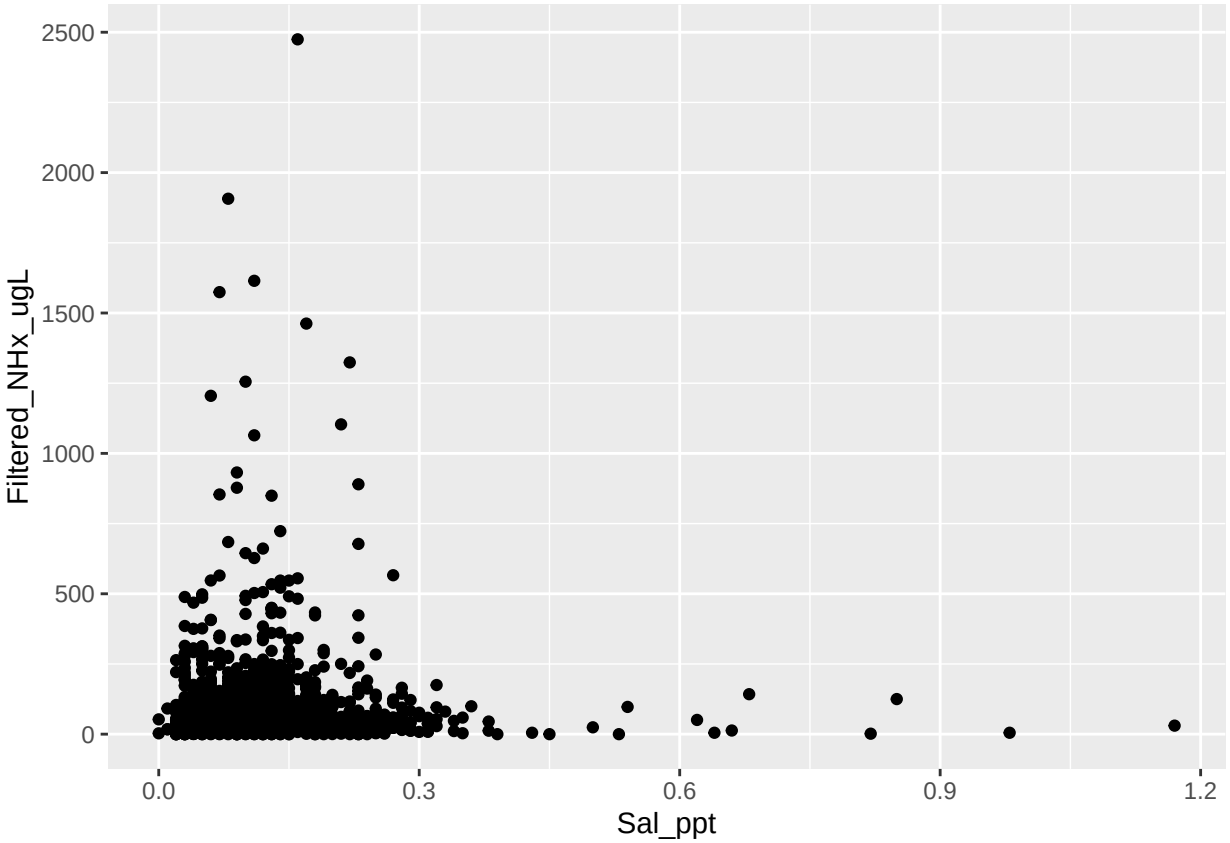

```
ggplot(wetland, aes(x = SpCond_mScm, y = Filtered_NHx_ugL)) +
geom_point()
```



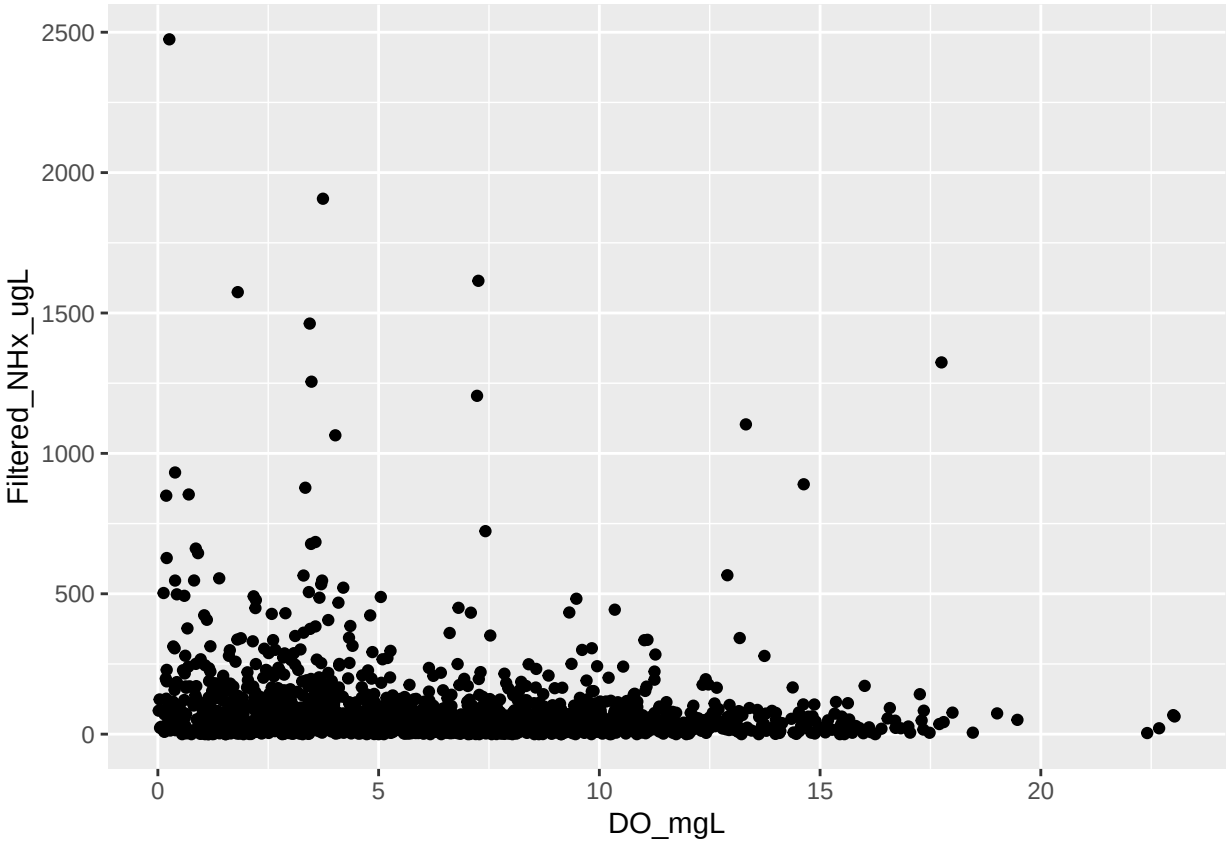

```
ggplot(wetland, aes(x = TDS_mgl, y = Filtered_NHx_ugL)) +
geom_point()
```



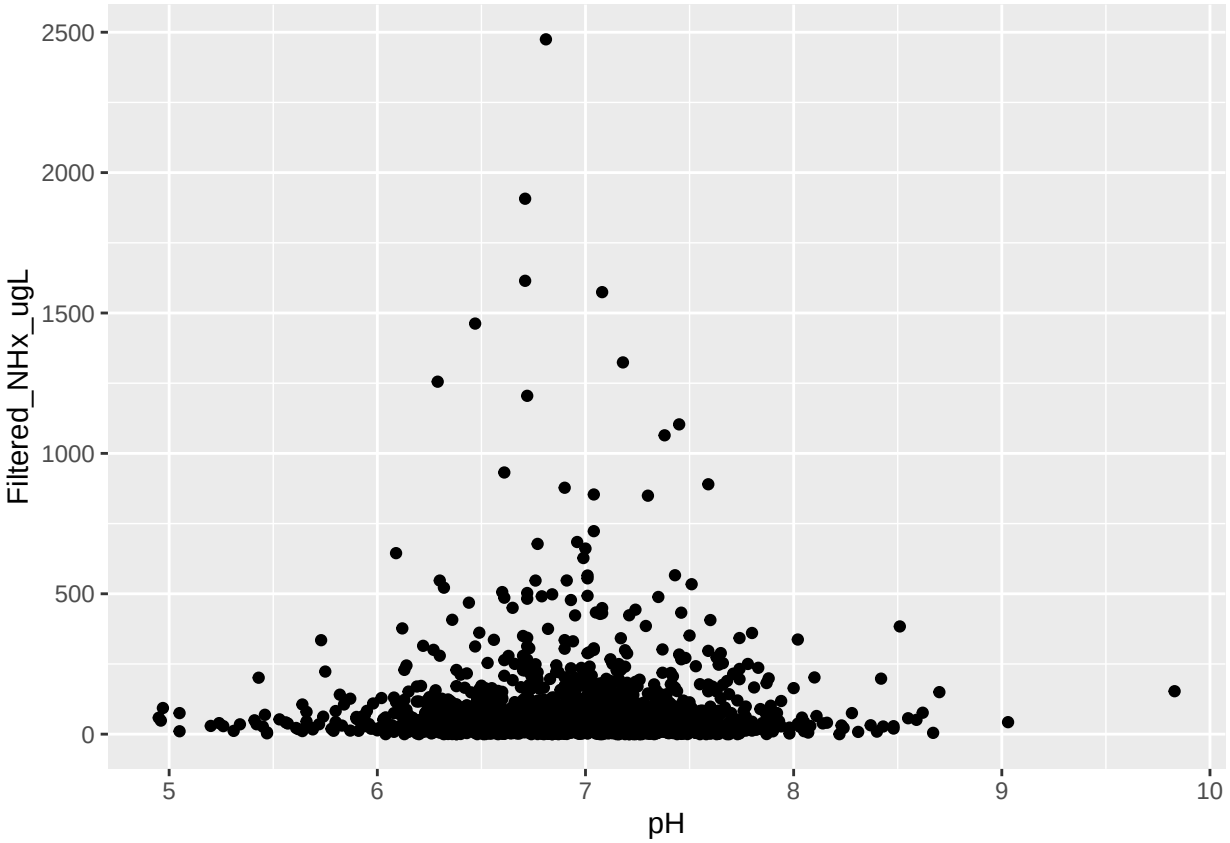

```
ggplot(wetland, aes(x = Sal_ppt, y = Filtered_NHx_ugL)) +
geom_point()
```



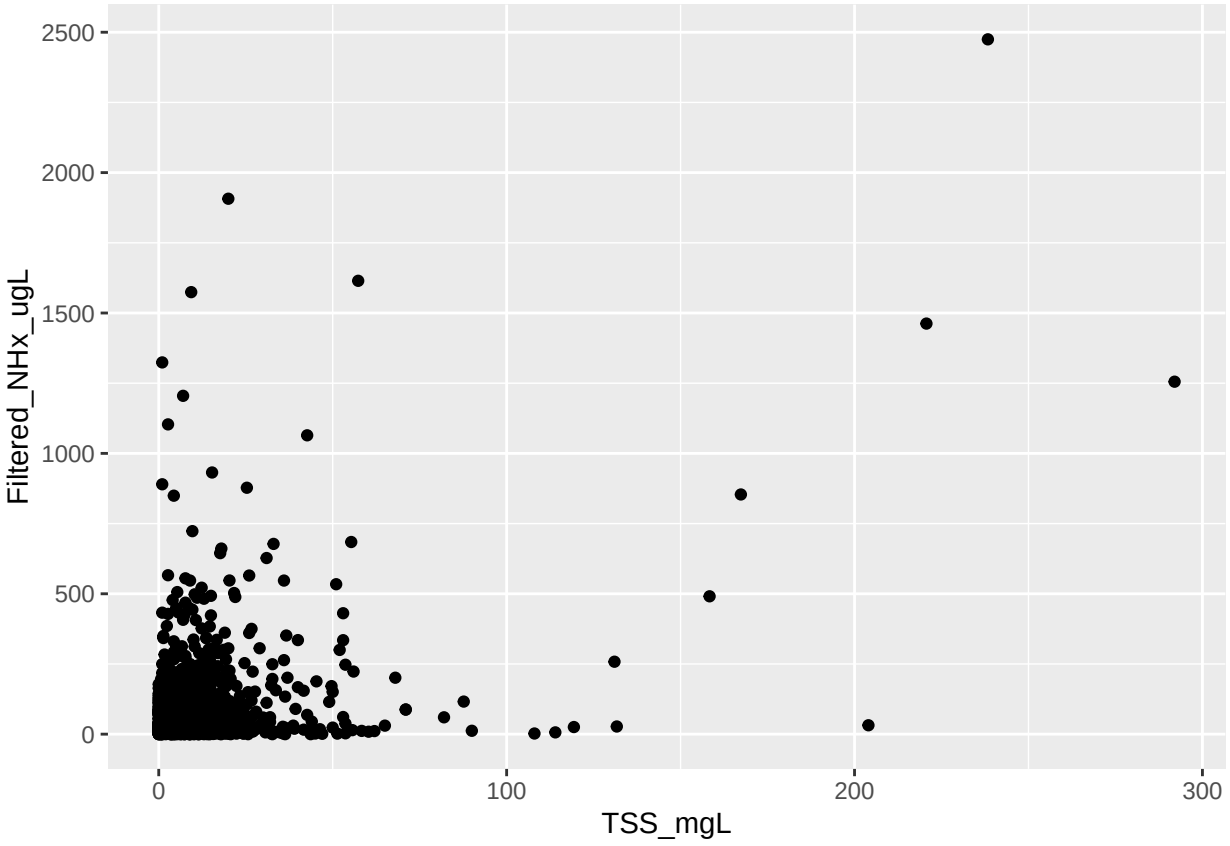

```
ggplot(wetland, aes(x = DO_mgL, y = Filtered_NHx_ugL)) +
geom_point()
```



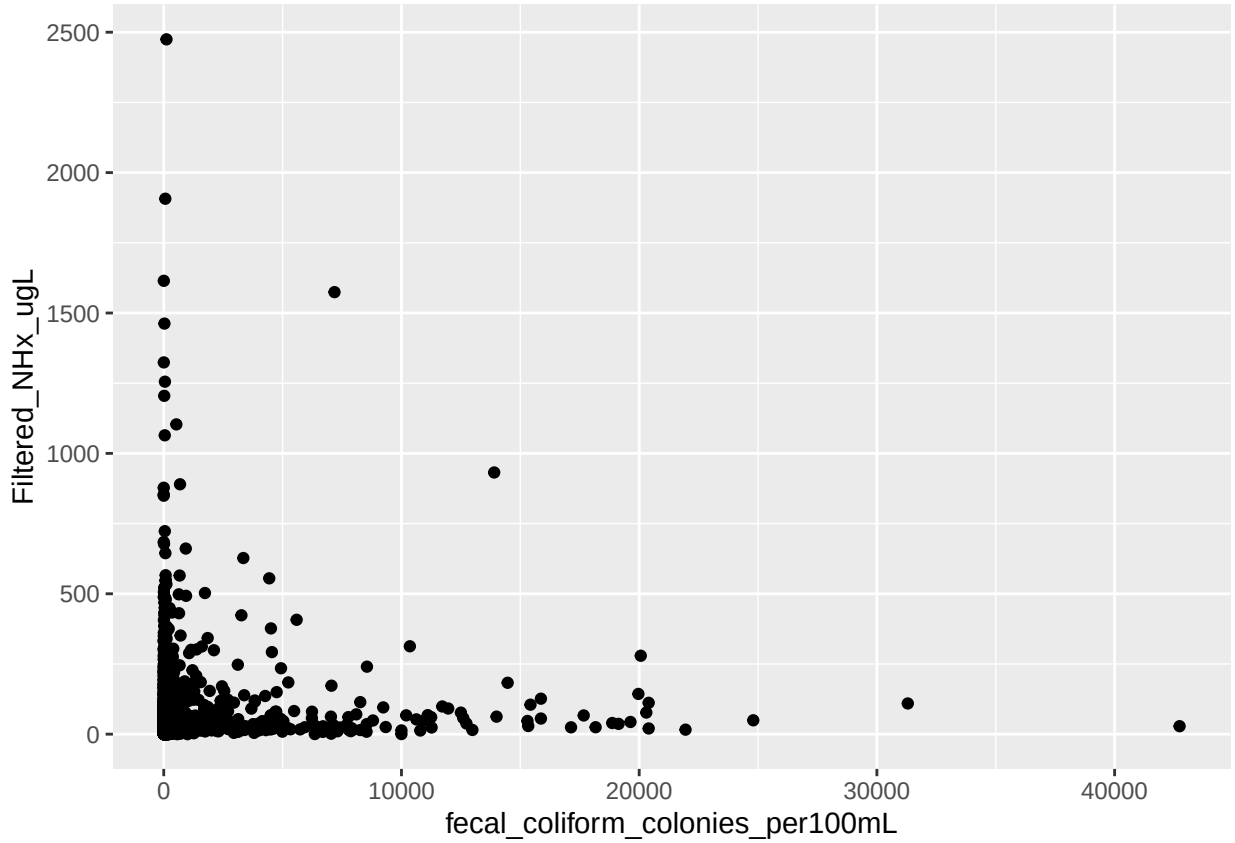

```
ggplot(wetland, aes(x = pH, y = Filtered_NHx_ugL)) +
geom_point()
```



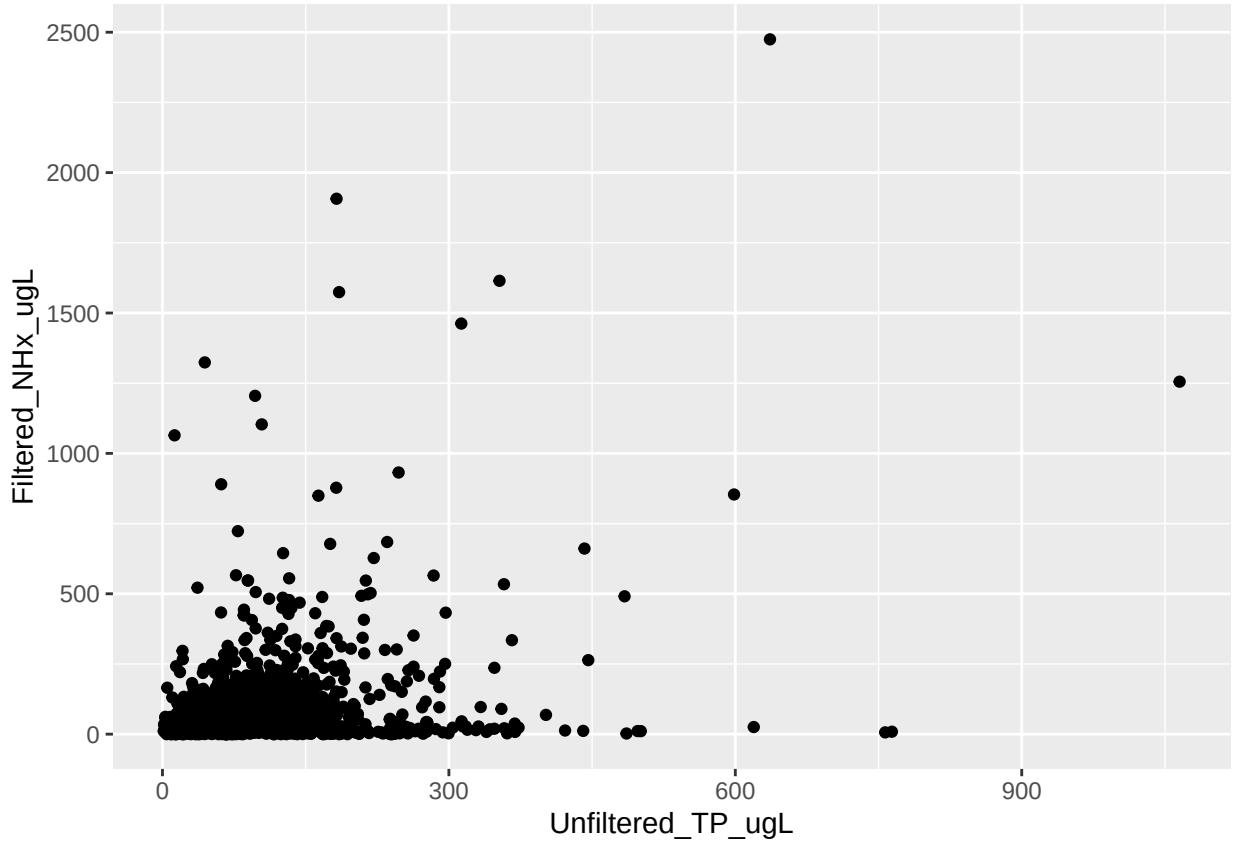

```
ggplot(wetland, aes(x = TSS_mgL, y = Filtered_NHx_ugL)) +
geom_point()
```




```
ggplot(wetland, aes(x = fecal_coliform_colonies_per100mL, y = Filtered_NHx_ugL)) +
geom_point()
```

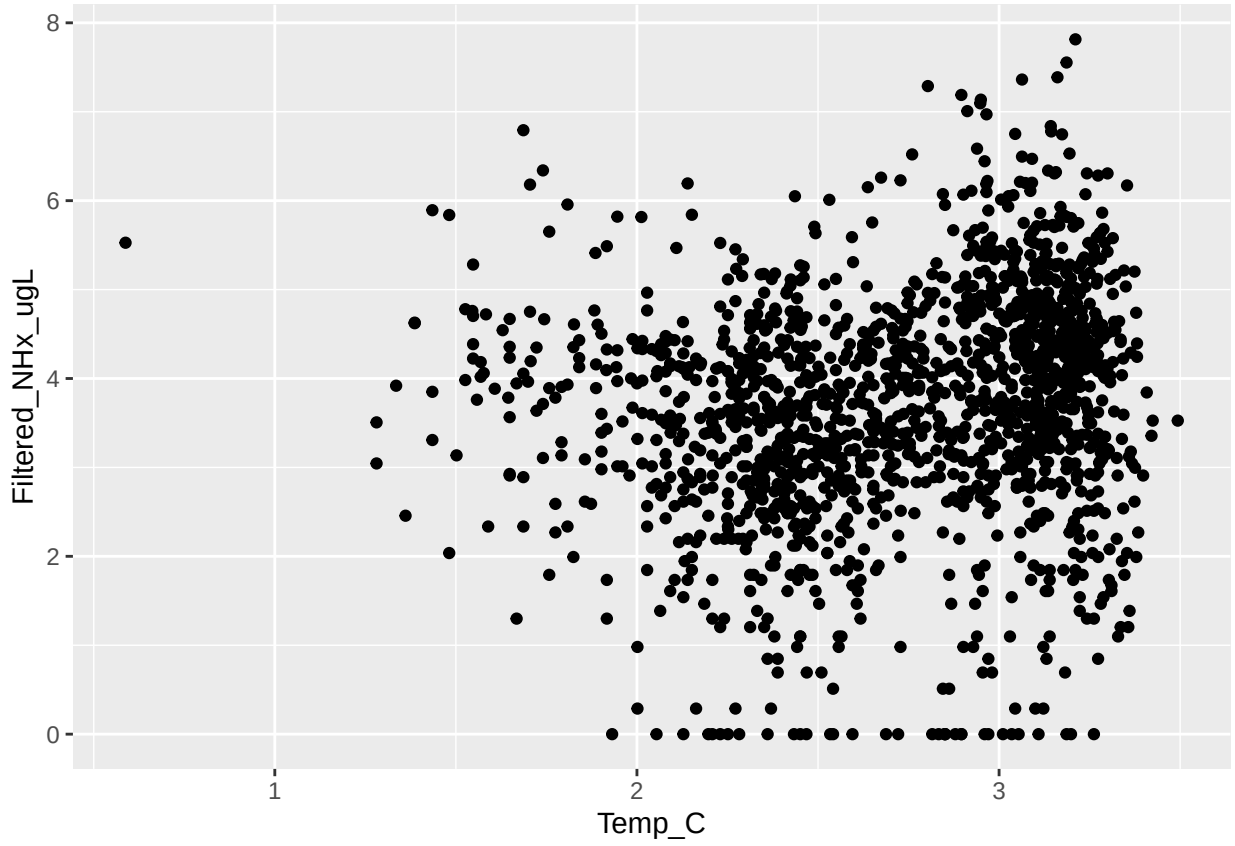



```
ggplot(wetland, aes(x = Unfiltered_TP_ugL, y = Filtered_NHx_ugL)) +
geom_point()
```

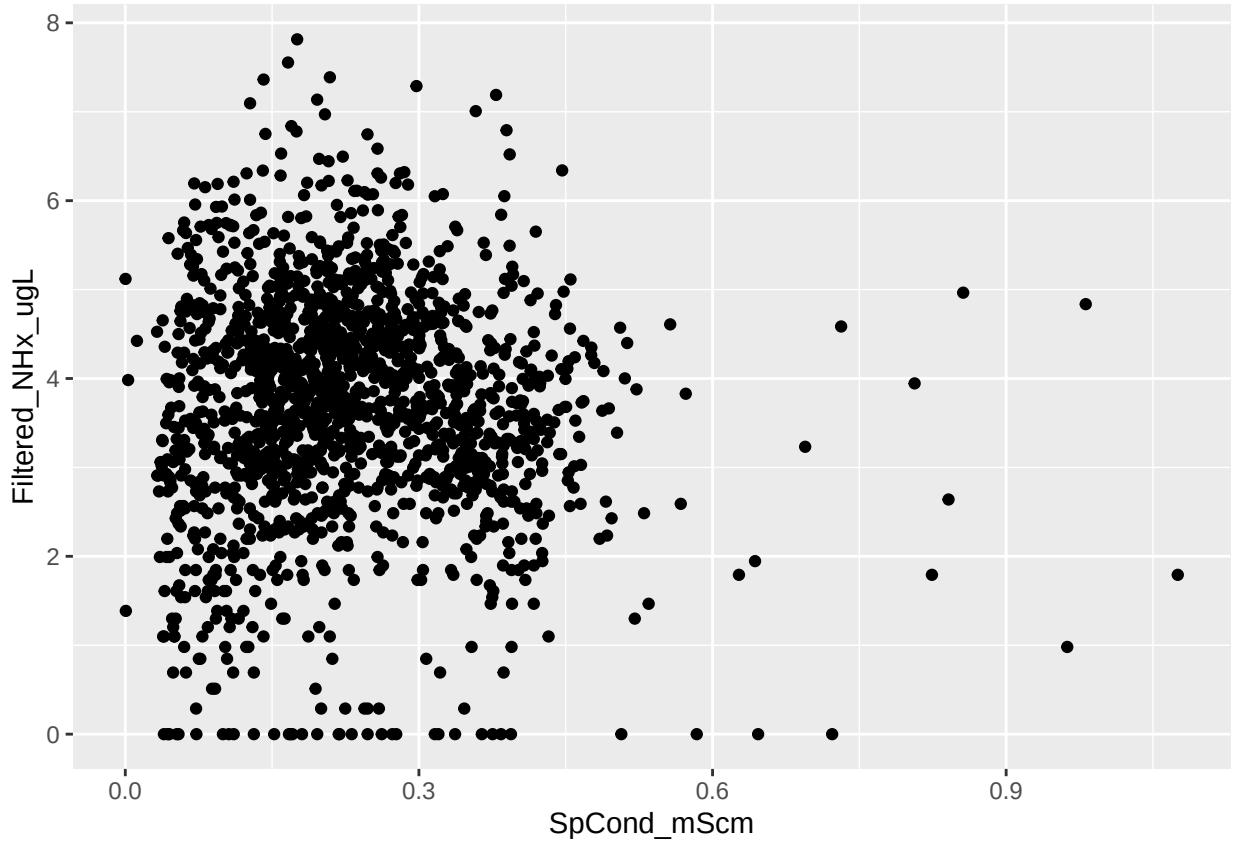


everything needs to be log transformed?

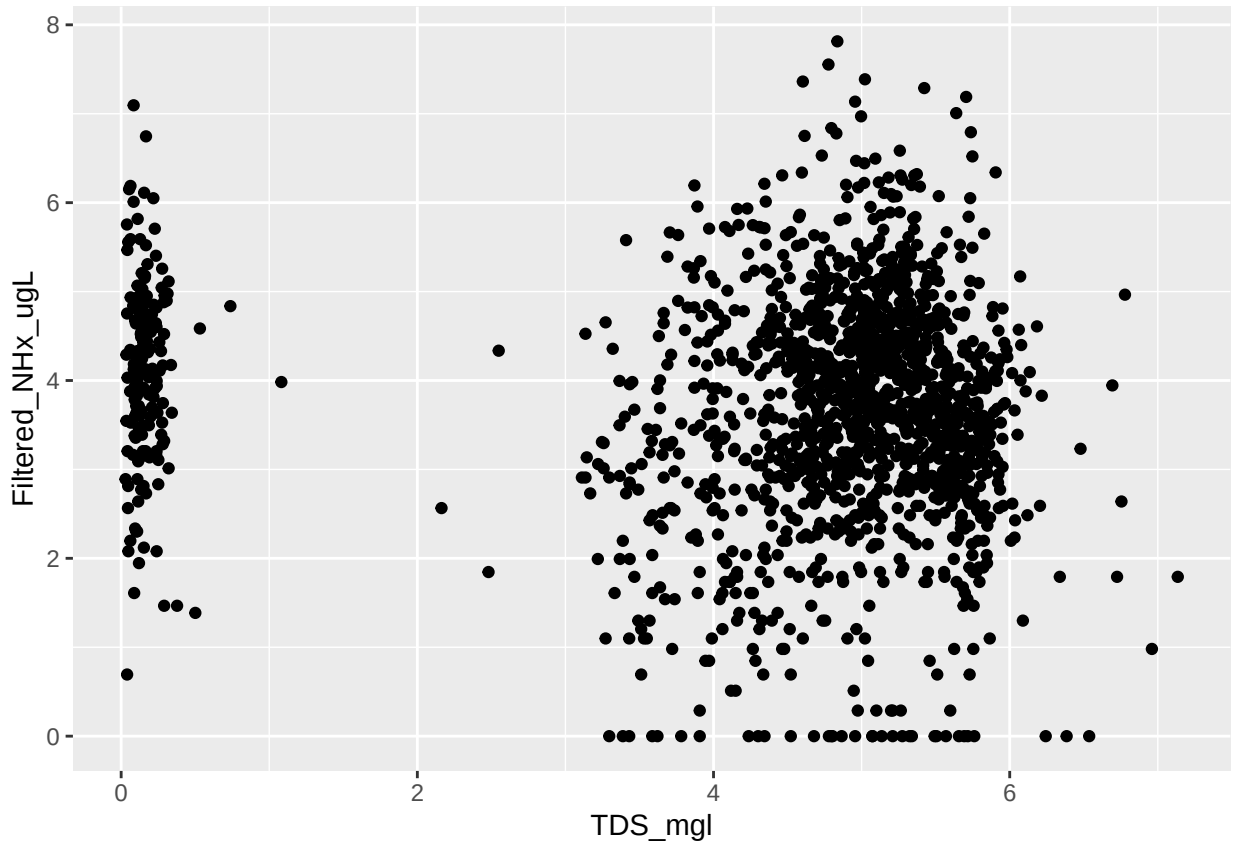
```
#log transform everything
log_wetland <- wetland %>%
  mutate(across(c(Temp_C, SpCond_mScm, Cond_mScm,
                  TDS_mgl, Sal_ppt, DO_mgL, pH,
                  Filtered_NHx_ugL, TSS_mgL,
                  fecal_coliform_colonies_per100mL, Unfiltered_TP_ugL),
                ~ log(.+1)))

#re-vistualize data

ggplot(log_wetland, aes(x = Temp_C, y = Filtered_NHx_ugL)) +
geom_point()
```



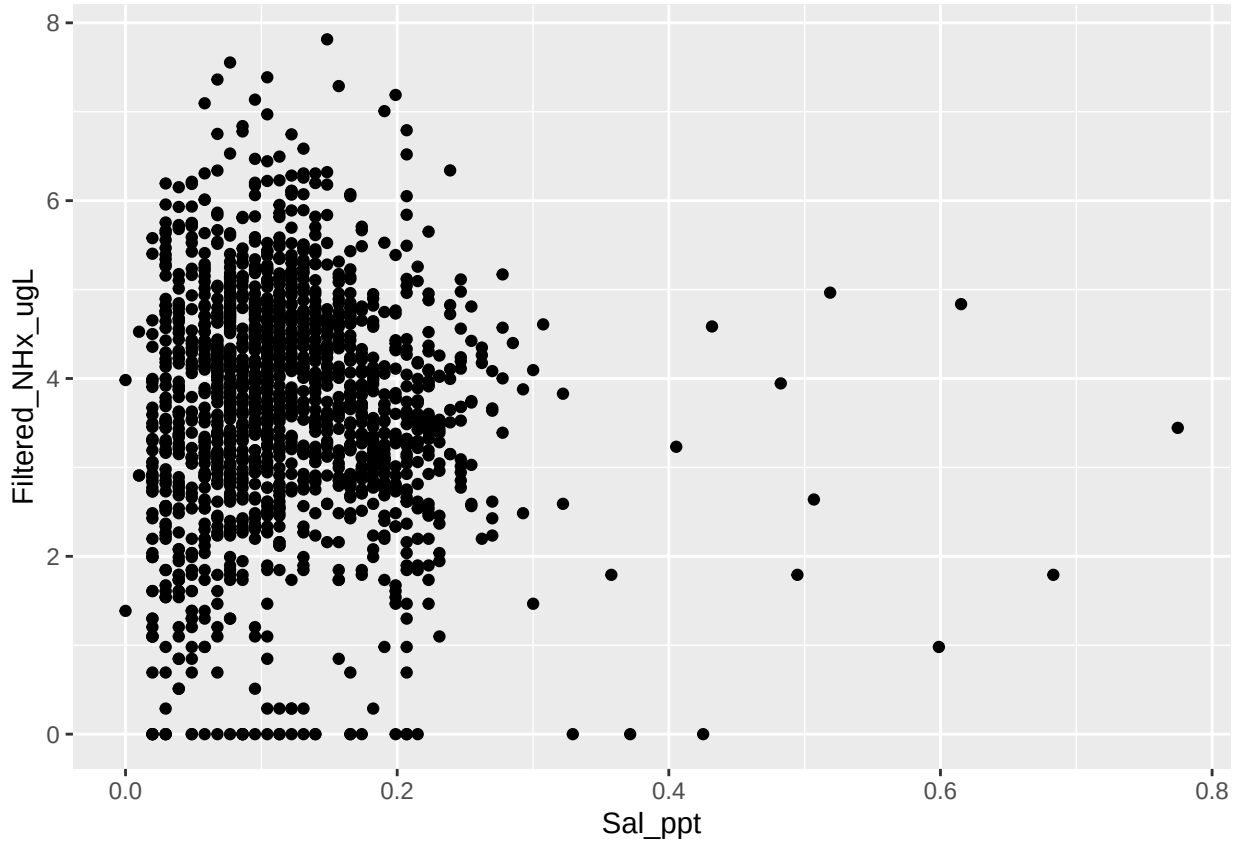

```
ggplot(log_wetland, aes(x = SpCond_mScm, y = Filtered_NHx_ugL)) +
geom_point()
```



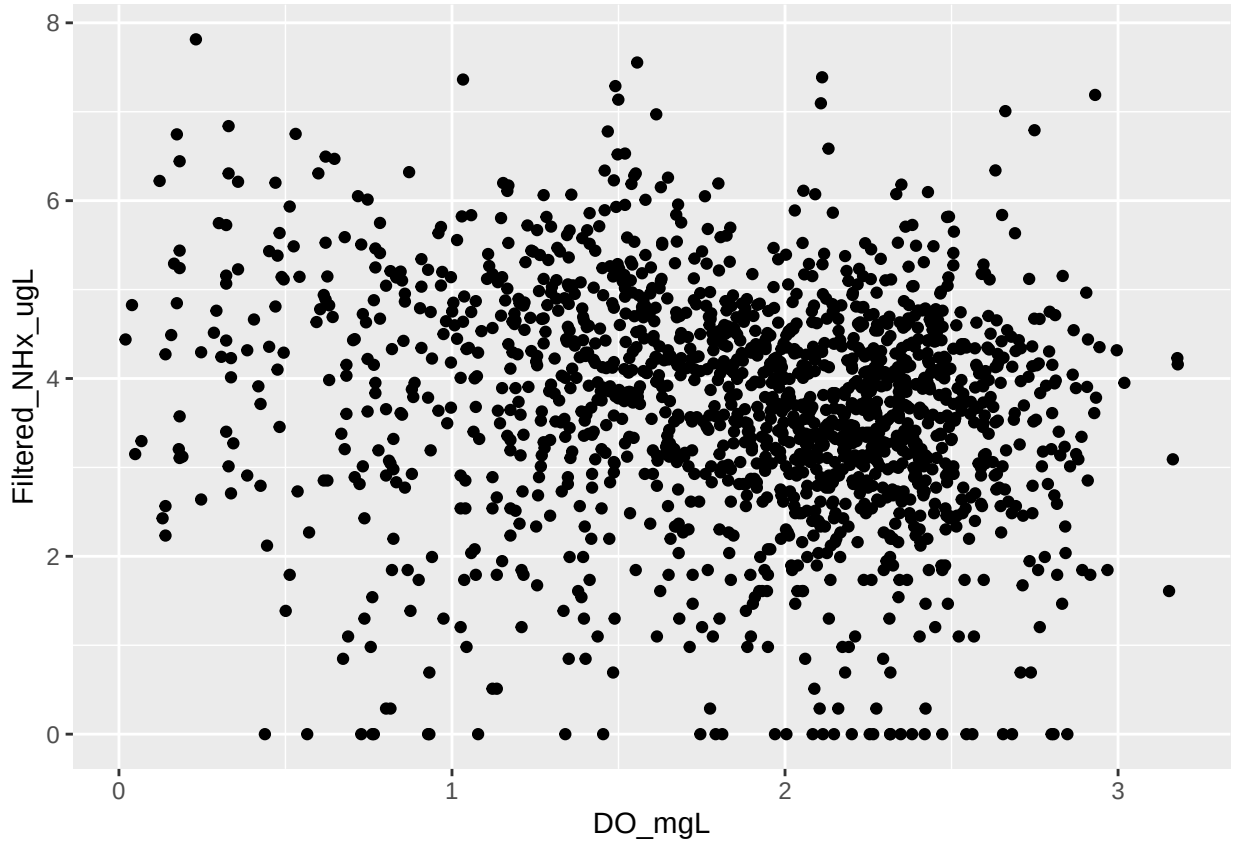

```
ggplot(log_wetland, aes(x = TDS_mgl, y = Filtered_NHx_ugL)) +
geom_point()
```



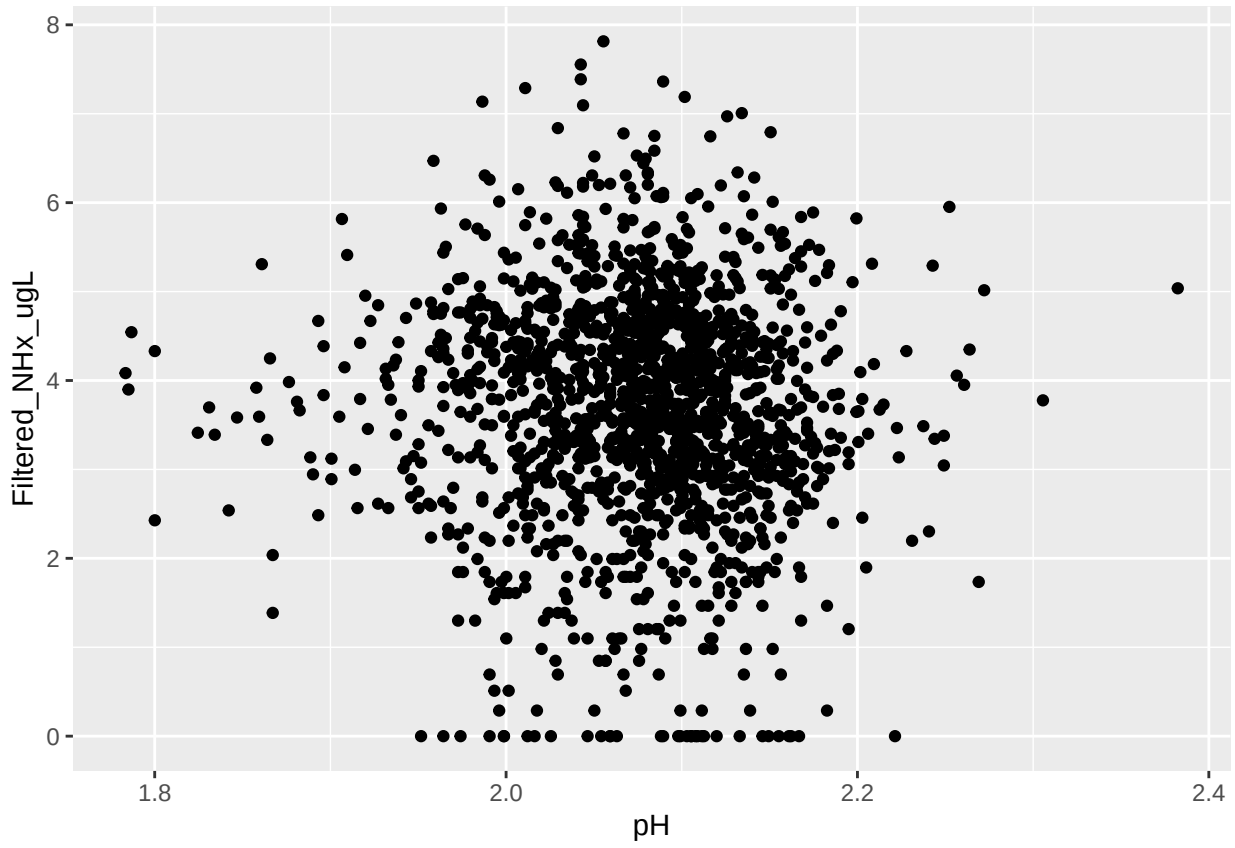

```
ggplot(log_wetland, aes(x = Sal_ppt, y = Filtered_NHx_ugL)) +
geom_point()
```



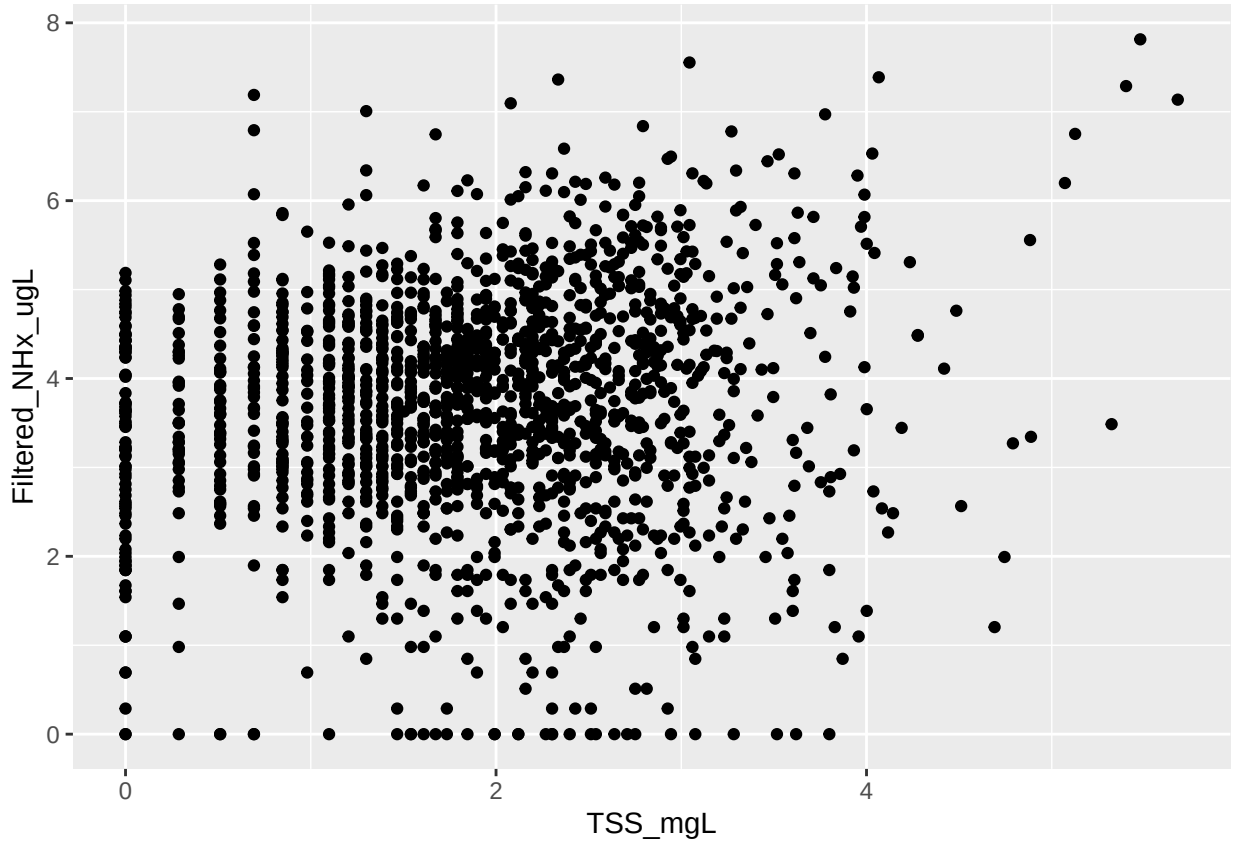

```
ggplot(log_wetland, aes(x = DO_mgL, y = Filtered_NHx_ugL)) +
geom_point()
```



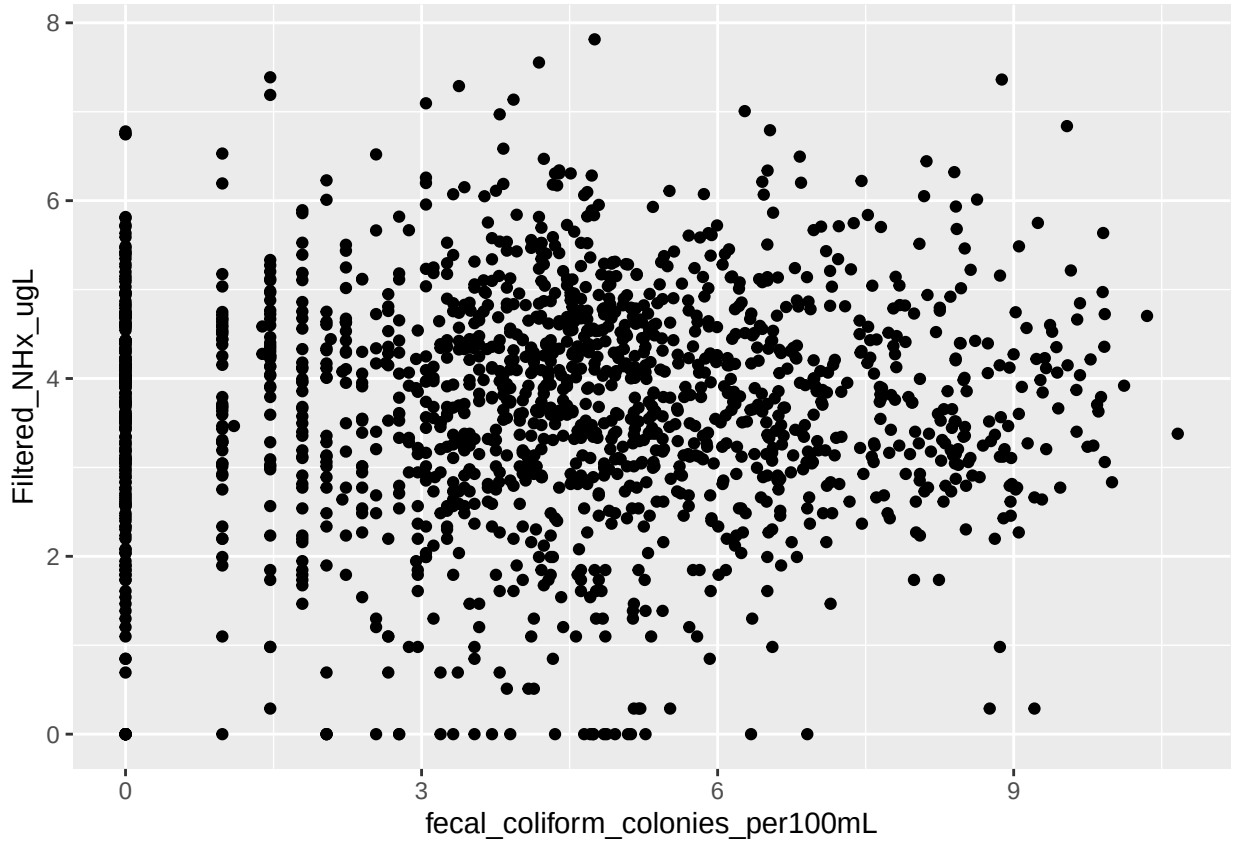

```
ggplot(log_wetland, aes(x = pH, y = Filtered_NHx_ugL)) +
geom_point()
```



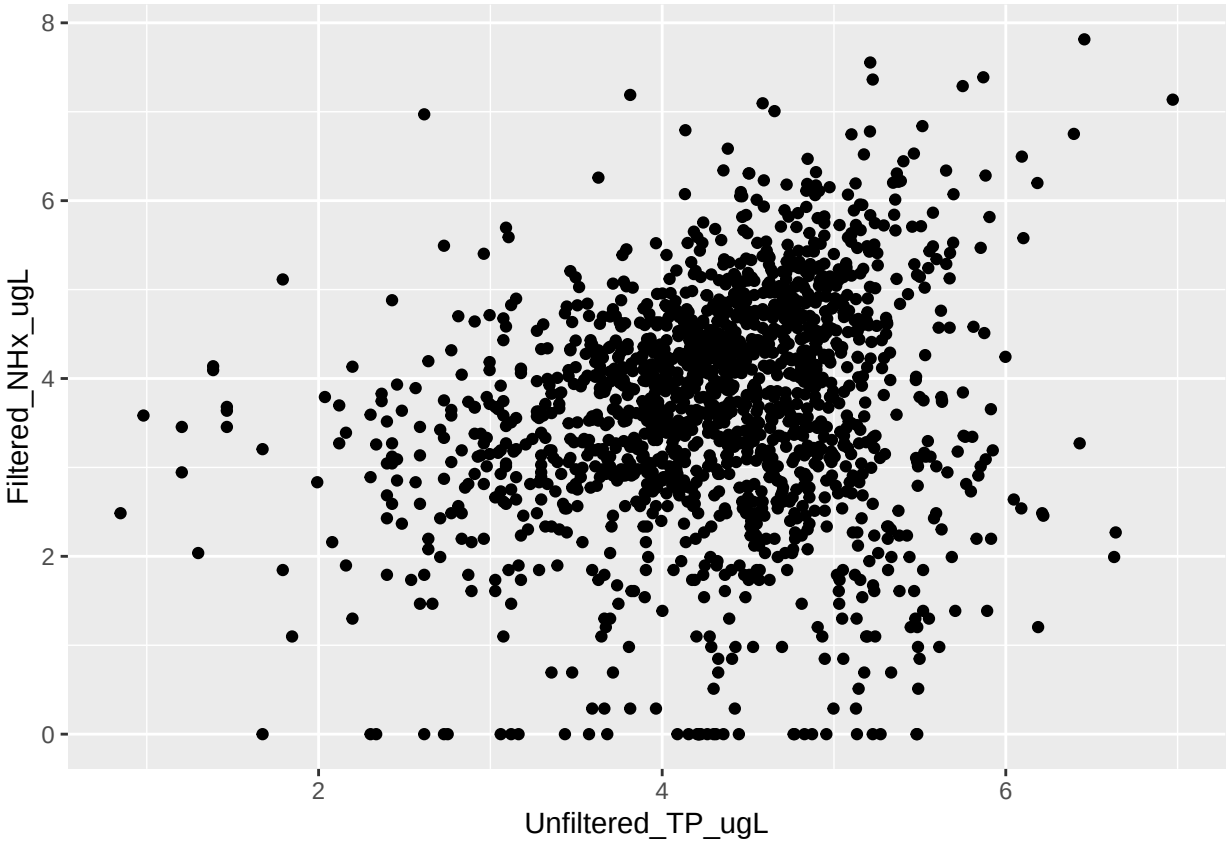

```
ggplot(log_wetland, aes(x = TSS_mgL, y = Filtered_NHx_ugL)) +
geom_point()
```




```
ggplot(log_wetland, aes(x = fecal_coliform_colonies_per100mL, y = Filtered_NHx_ugL)) +
geom_point()
```




```
ggplot(log_wetland, aes(x = Unfiltered_TP_ugL, y = Filtered_NHx_ugL)) +
geom_point()
```


Filtered_NHx_ugL
Unfiltered_TP_ugL
0
2
4
6
8
2
4
6