

Exploratory Plots 20190113

Whitney Beck

January 13, 2019

```
##Data Prep from 1) Data Clean Script
```

```
## 1) script to clean and wrangle data
library(tidyverse, verbose = F)
```

```
## -- Attaching packages ----- tidyverse 1.2.1 --
```

```
## v ggplot2 3.1.0      v purrr  0.2.5
## v tibble  1.4.2      v dplyr  0.7.8
## v tidyr   0.8.2      v stringr 1.3.1
## v readr   1.3.1      v forcats 0.3.0
```

```
## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()     masks stats::lag()
```

```
library(ggplot2)
theme_mod <- function(){theme_bw() %+replace% theme(panel.grid = element_blank())}
theme_set(theme_mod())
```

```
## Data import
```

```
Data_Beck=read.csv("raw-data/Beck Data/Beck Data.csv")
Data_Elser=read.csv("raw-data/Elser-et-al-2007_Global-N-P-limitation/data/doi_10.5063_AA_nceasadmin.910")
Data_Vanni=read.csv("raw-data/Vanni-et-al_2017_Animal-excretion/Aquatic_animal_excretion_data.csv")
Data_Vanni$Temperature..C.=as.numeric(Data_Vanni$Temperature..C.)
Data_Vanni$P.excretion.rate..ug.P.ind.h.=as.numeric(Data_Vanni$P.excretion.rate..ug.P.ind.h.)
Data_Vanni$N.excretion.rate..ug.N.ind.h.=as.numeric(Data_Vanni$N.excretion.rate..ug.N.ind.h.)
Data_Vanni$Excreted.N.P..molar.=as.numeric(Data_Vanni$Excreted.N.P..molar.)
```

```
## Elser: set variables of interest
```

```
var_names_Elser = c("id_study", "system", "strata", "habitat", "cat", "tax_resp_class",
                    "l.n.c", "l.p.c", "l.int.c", "temp", "n_avail", "n_total", "p_avail",
                    "p_total", "light", "location", "latitud", "longitud", "elev", "duration")
```

```
##Filter to variables of interest that have temperature values in aquatic systems
```

```
TempData_Elser <- Data_Elser %>%
  select(var_names_Elser) %>%
  filter(system %in% c("FRESHWATER", "MARINE")) %>%
  filter(!is.na(temp))
```

```
##Beck: set variables of interest
```

```
var_names_Beck=c("Ref", "Site.Name", "Latitude_GoogleEarth", "Longitude_GoogleEarth", "Days", "Nitrogen",
                 "Phosphorus.Molarity", "Elevation_m", "Depth_m", "Canopy_Percent", "Canopy_Qual", "Temp",
                 "NH4_ug_L", "NO3_ug_L", "DIN_ug_L", "TN_mg_L", "SRP_ug_L", "TP_ug_L", "NO3_SRP_Ratio",
                 "Control_mean_mg_m2", "Control.SE",
                 "N_mean_mg_m2", "N.SE", "N.LRR.Effect", "N.LRR.Variance",
                 "P_mean_mg_m2", "P.SE", "P.LRR.Effect", "P.LRR.Variance",
                 "NP_mean_mg_m2", "NP.SE", "NP.LRR", "NP.LRR.Variance")
```

```

#Filter to variables of interest that have temperature values
TempData_Beck <- Data_Beck %>%
  select(var_names_Beck) %>%
  filter(!is.na(Temp_C))

##Vanni: set variables of interest

#Many of these could be interesting to explore, and variable names are really strangely formatted
#so I'm going to leave as-is for now

TempData_Vanni <- Data_Vanni%>%
  filter(!is.na(Temperature..C.))

```

Elser Plots from 2) Exploratory script

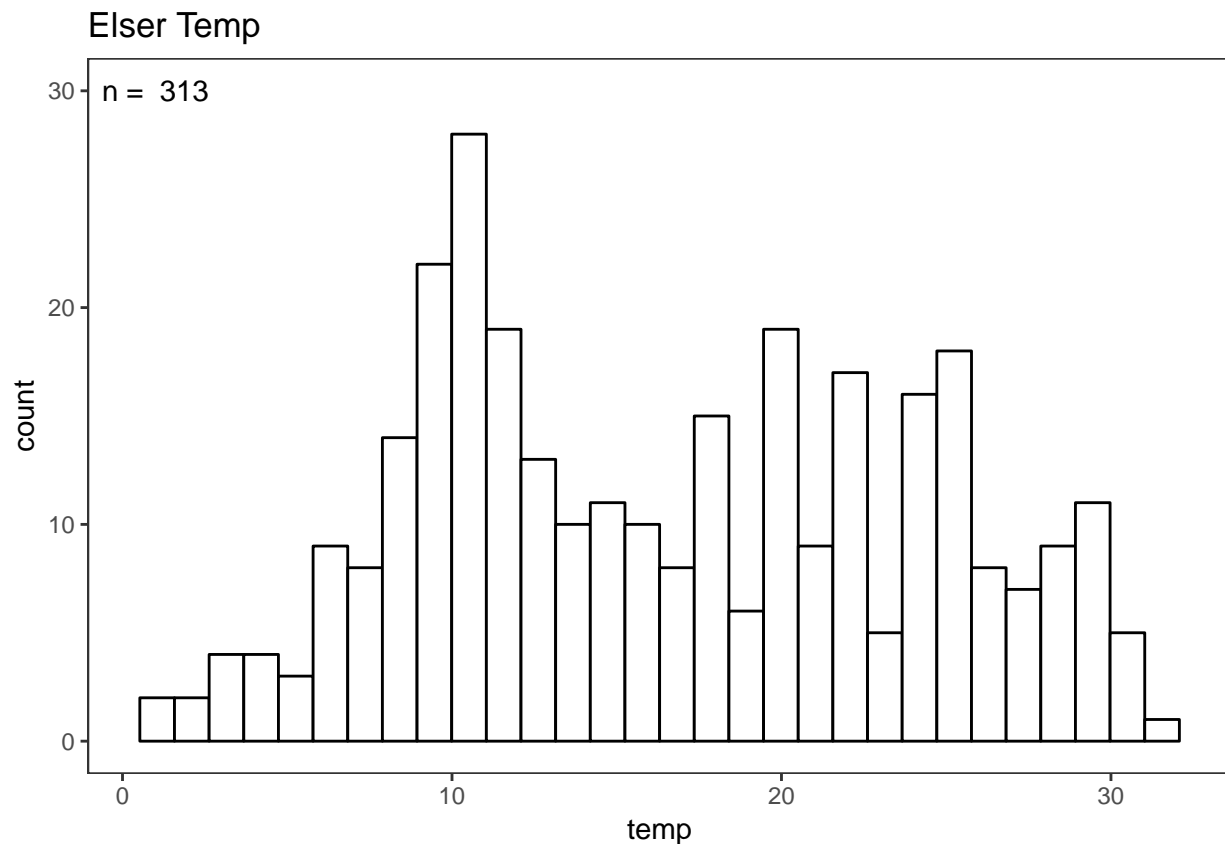
```

#Temperature Range

ggplot(TempData_Elser, aes(x = temp)) + geom_histogram(colour = 'black', fill = 'white')+
  annotate('text', x = 1, y = 30, label = paste("n = ", nrow(subset(TempData_Elser, temp!="NA"))))+
  ggtitle("Elser Temp")

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

```

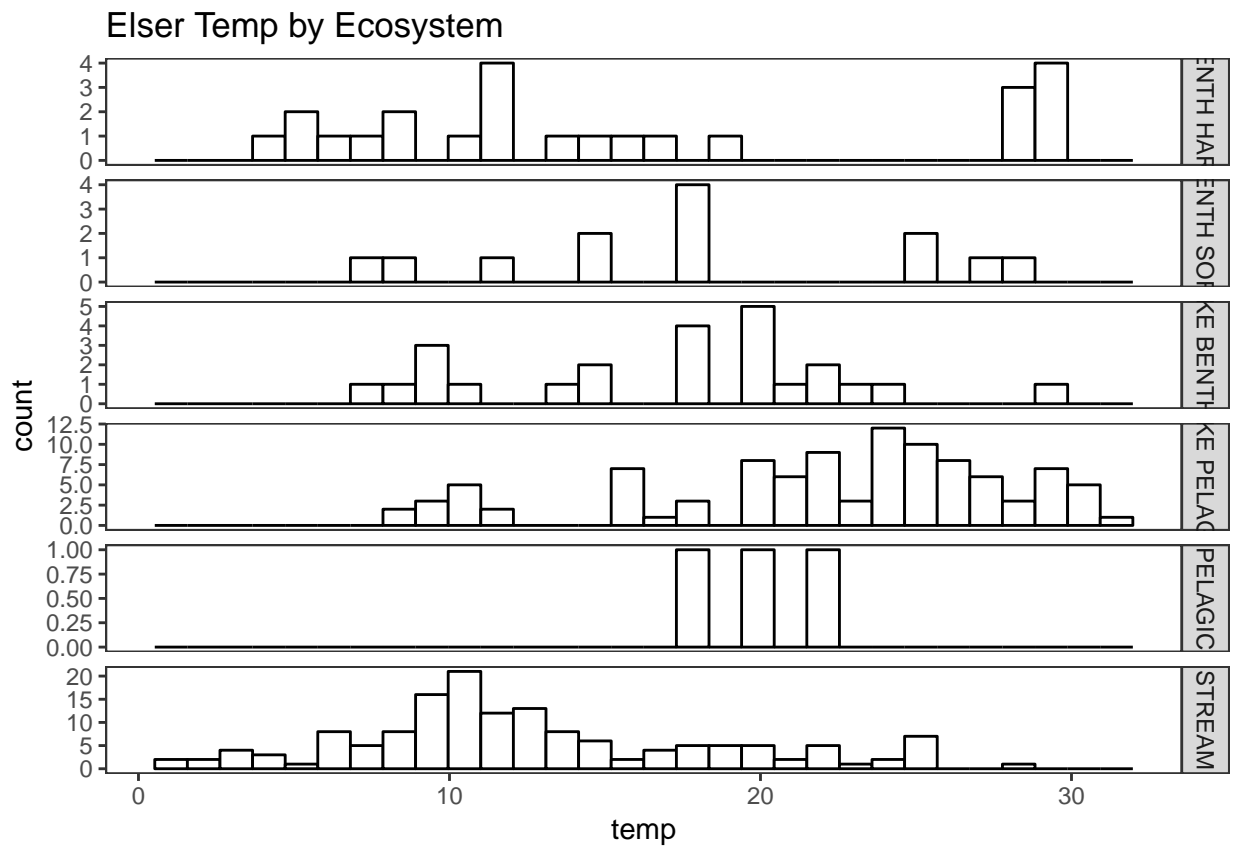


```

ggplot(TempData_Elser, aes(x = temp)) + geom_histogram(colour = 'black', fill = 'white') + facet_grid(ca
  ggtitle("Elser Temp by Ecosystem")

```

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

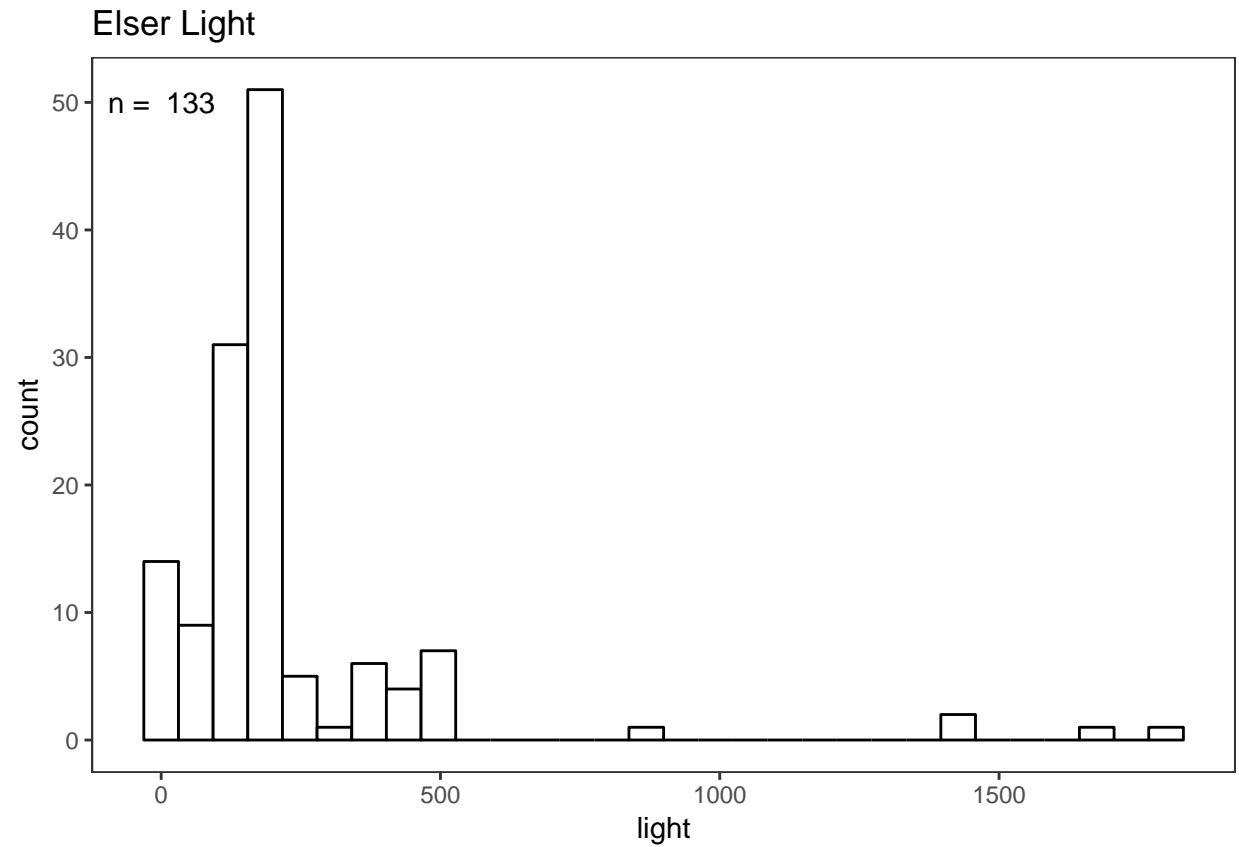


#Light Range

```
ggplot(TempData_Elser, aes(x = light)) + geom_histogram(colour = 'black', fill = 'white') +
  annotate('text', x = 1, y = 50, label = paste("n = ", nrow(subset(TempData_Elser, light != "NA")))) +
  ggtitle("Elser Light")
```

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

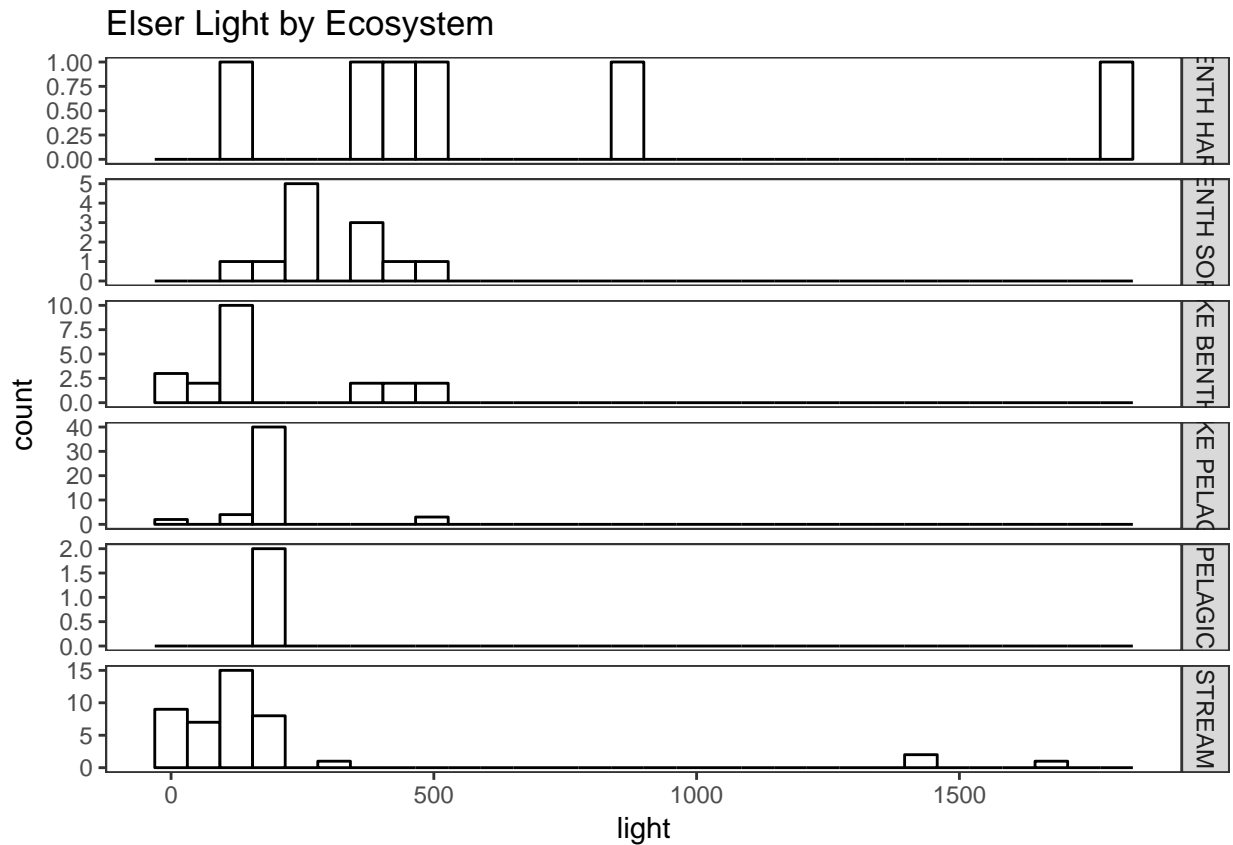
```
## Warning: Removed 180 rows containing non-finite values (stat_bin).
```



```
ggplot(TempData_Elser, aes(x = light)) + geom_histogram(colour = 'black',fill = 'white') + facet_grid(c~  
  ggtitle("Elser Light by Ecosystem"))
```

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

```
## Warning: Removed 180 rows containing non-finite values (stat_bin).
```



#Avail N Range

```
Elser_log_n_avail = ggplotGrob(ggplot(TempData_Elser, aes(x = log10(n_avail))) + geom_histogram(colour = 'black', fill = 'white'))
```

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

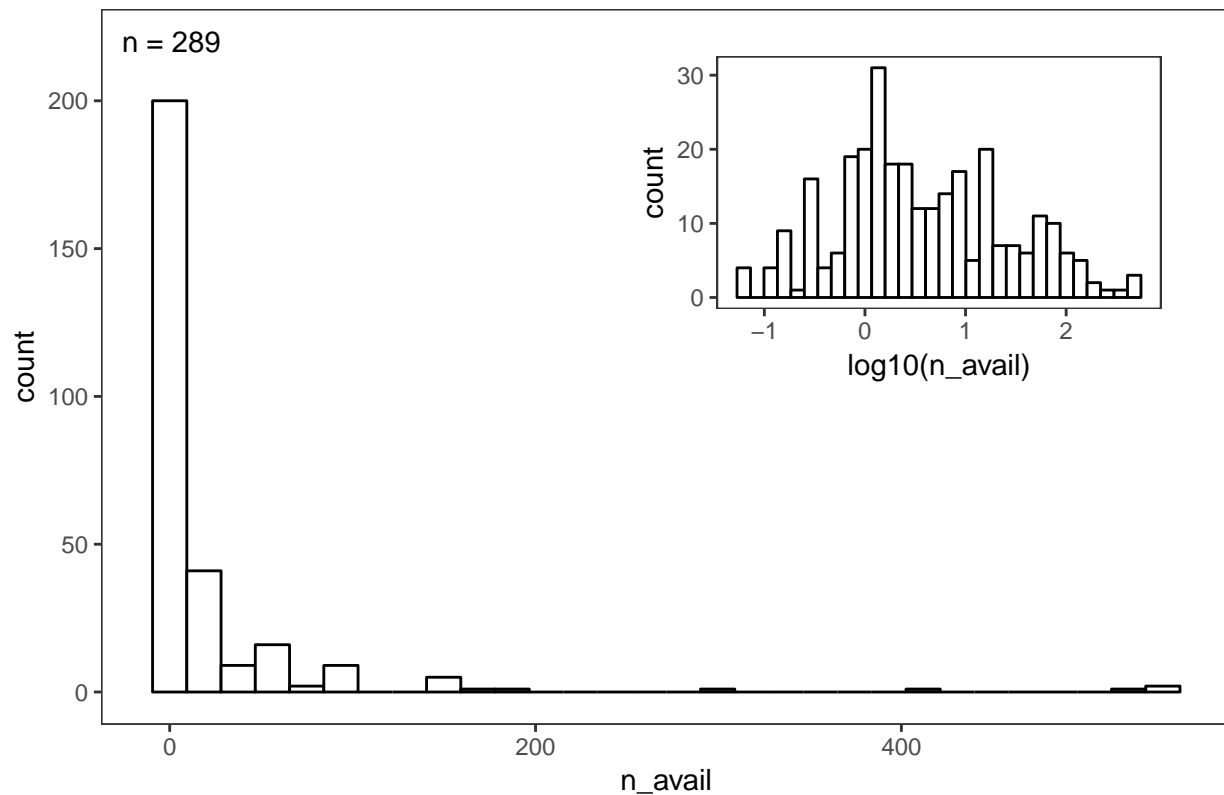
```
## Warning: Removed 24 rows containing non-finite values (stat_bin).
```

```
ggplot(TempData_Elser, aes(x = n_avail)) + geom_histogram(colour = 'black', fill = 'white') +
  annotate('text', x = 1, y = 220, label = paste("n =", nrow(subset(TempData_Elser, n_avail != "NA")))) +
  annotation_custom(grob = Elser_log_n_avail, xmin = 250, xmax = 550, ymin = 100, ymax = 220) +
  ggtitle("Elser Available N")
```

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

```
## Warning: Removed 24 rows containing non-finite values (stat_bin).
```

Elser Available N



#Total N Range

```
Elser_log_n_total = ggplotGrob(ggplot(TempData_Elser, aes(x = log10(n_total))) + geom_histogram(colour = 'black', fill = 'white'))
```

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

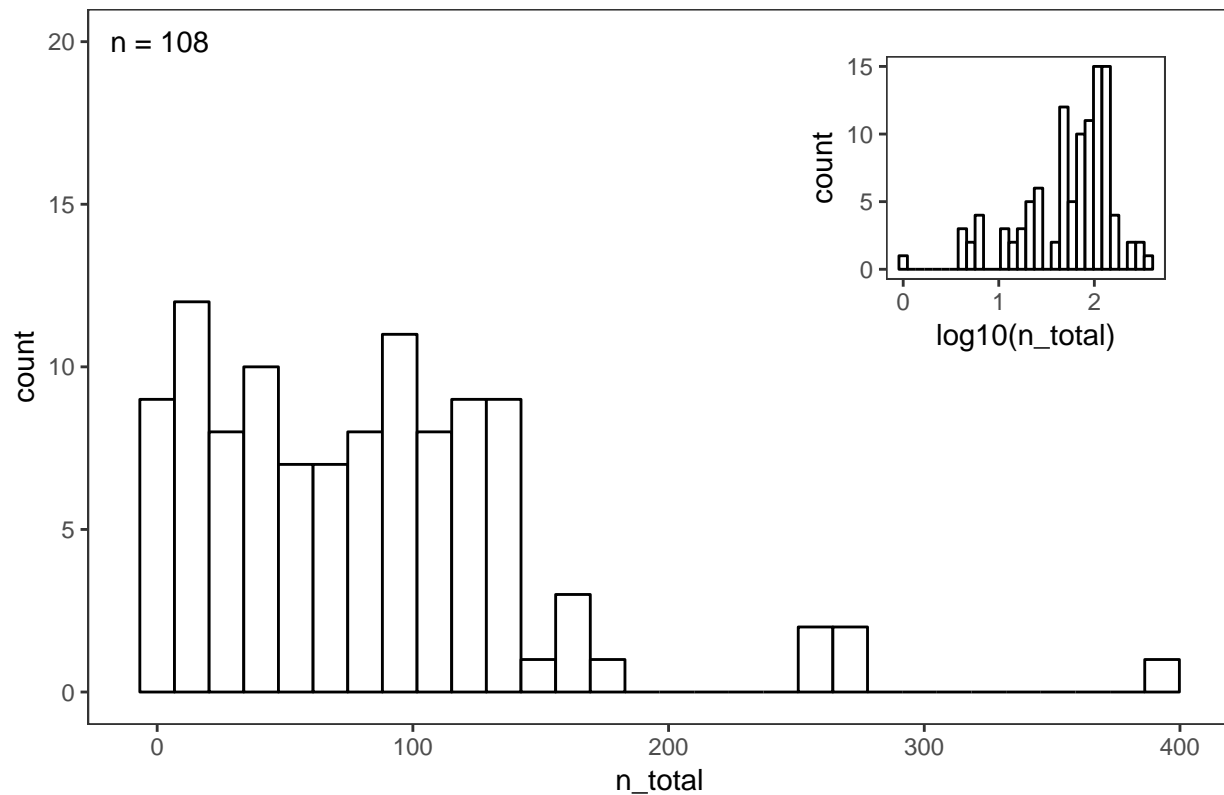
```
## Warning: Removed 205 rows containing non-finite values (stat_bin).
```

```
ggplot(TempData_Elser, aes(x = n_total)) + geom_histogram(colour = 'black', fill = 'white') +  
  annotate('text', x = 1, y = 20, label = paste("n =", nrow(subset(TempData_Elser, n_total!="NA")))) +  
  annotation_custom(grob = Elser_log_n_total, xmin = 250, xmax = 400, ymin = 10, ymax = 20)+  
  ggtitle("Elser Total N")
```

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

```
## Warning: Removed 205 rows containing non-finite values (stat_bin).
```

Elser Total N



#Avail P Range

```
Elser_log_p_avail = ggplotGrob(ggplot(TempData_Elser, aes(x = log10(p_avail))) + geom_histogram(colour = 'black', fill = 'white'))
```

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

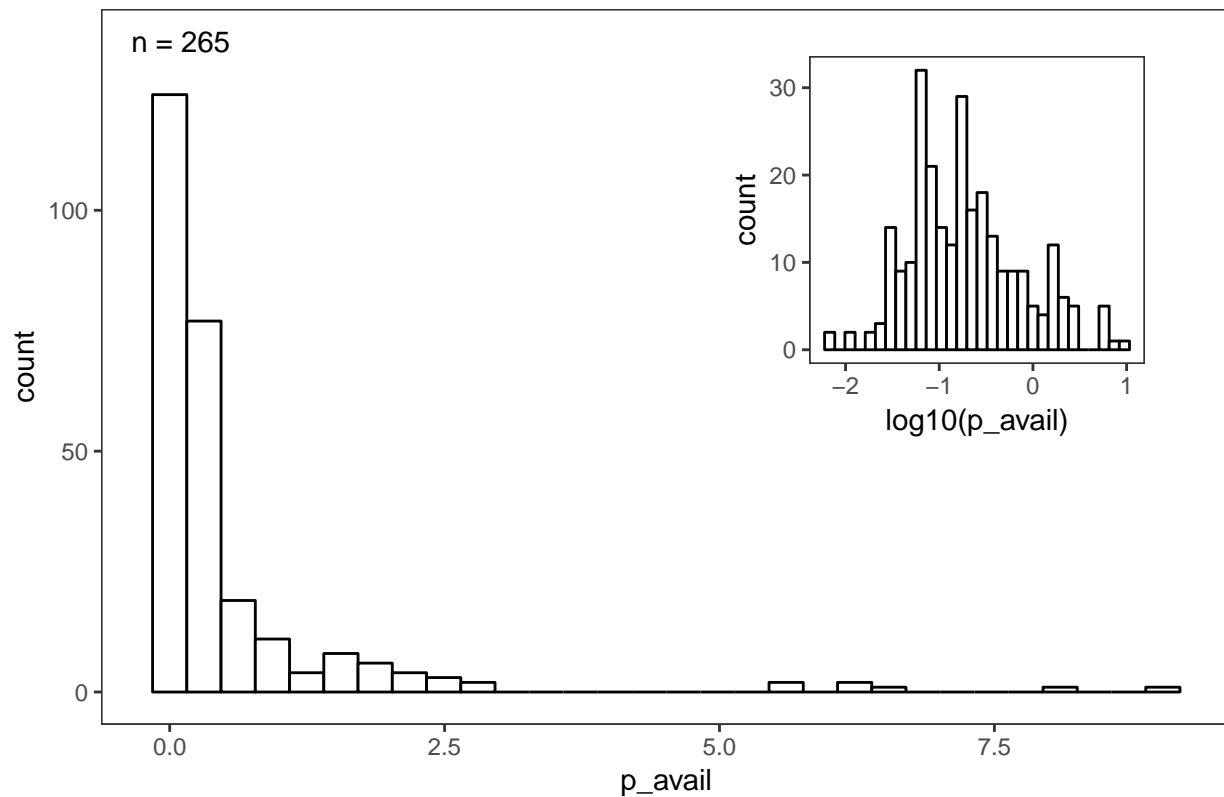
```
## Warning: Removed 50 rows containing non-finite values (stat_bin).
```

```
ggplot(TempData_Elser, aes(x = p_avail)) + geom_histogram(colour = 'black', fill = 'white') +
  annotate('text', x = .1, y = 135, label = paste("n =", nrow(subset(TempData_Elser, p_avail!="NA")))) +
  annotation_custom(grob = Elser_log_p_avail, xmin = 5, xmax = 9, ymin = 50, ymax = 135)+
  ggtitle("Elser Available P")
```

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

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

Elser Available P



#Total P Range

```
Elser_log_p_total = ggplotGrob(ggplot(TempData_Elser, aes(x = log10(p_total))) + geom_histogram(colour = 'black', fill = 'white'))
```

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

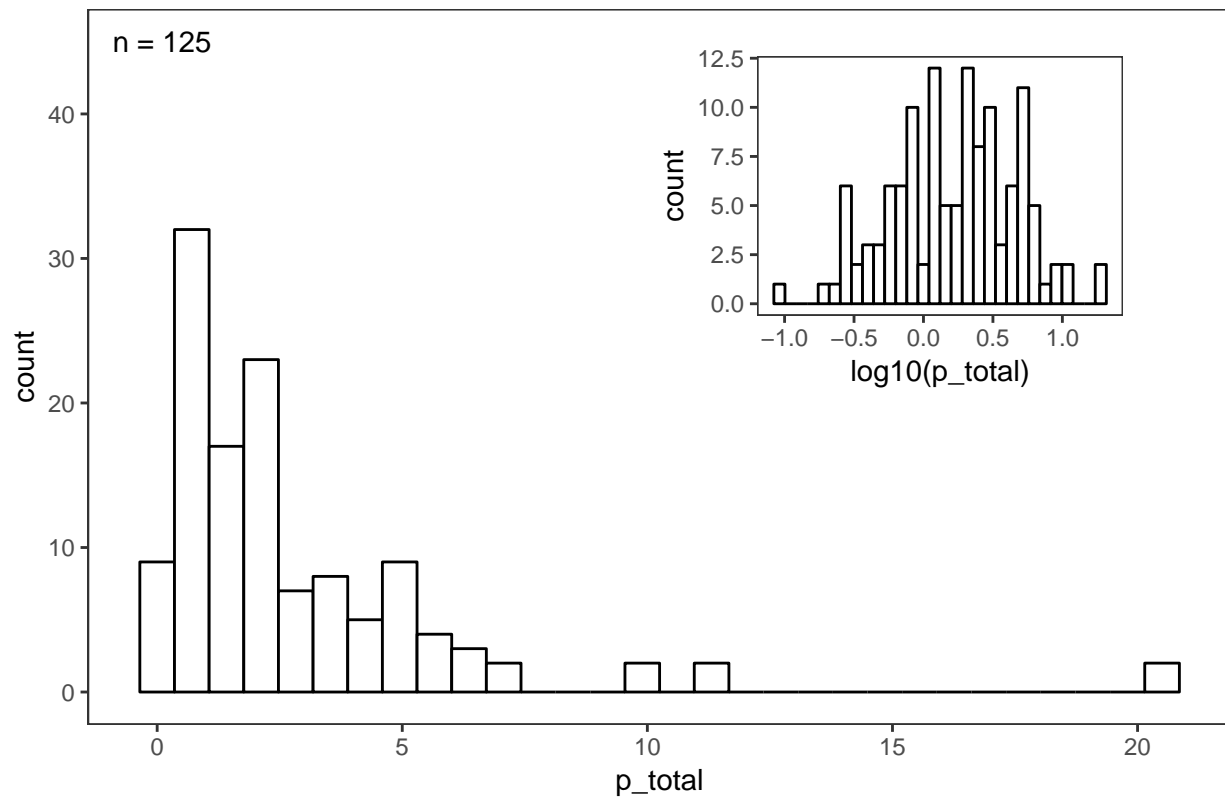
```
## Warning: Removed 188 rows containing non-finite values (stat_bin).
```

```
ggplot(TempData_Elser, aes(x = p_total)) + geom_histogram(colour = 'black', fill = 'white') +
  annotate('text', x = .1, y = 45, label = paste("n =", nrow(subset(TempData_Elser, p_total != "NA")))) +
  annotation_custom(grob = Elser_log_p_total, xmin = 10, xmax = 20, ymin = 20, ymax = 45) +
  ggtitle("Elser Total P")
```

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

```
## Warning: Removed 188 rows containing non-finite values (stat_bin).
```


Elser Total P



##Beck Plots from 2) Exploratory script

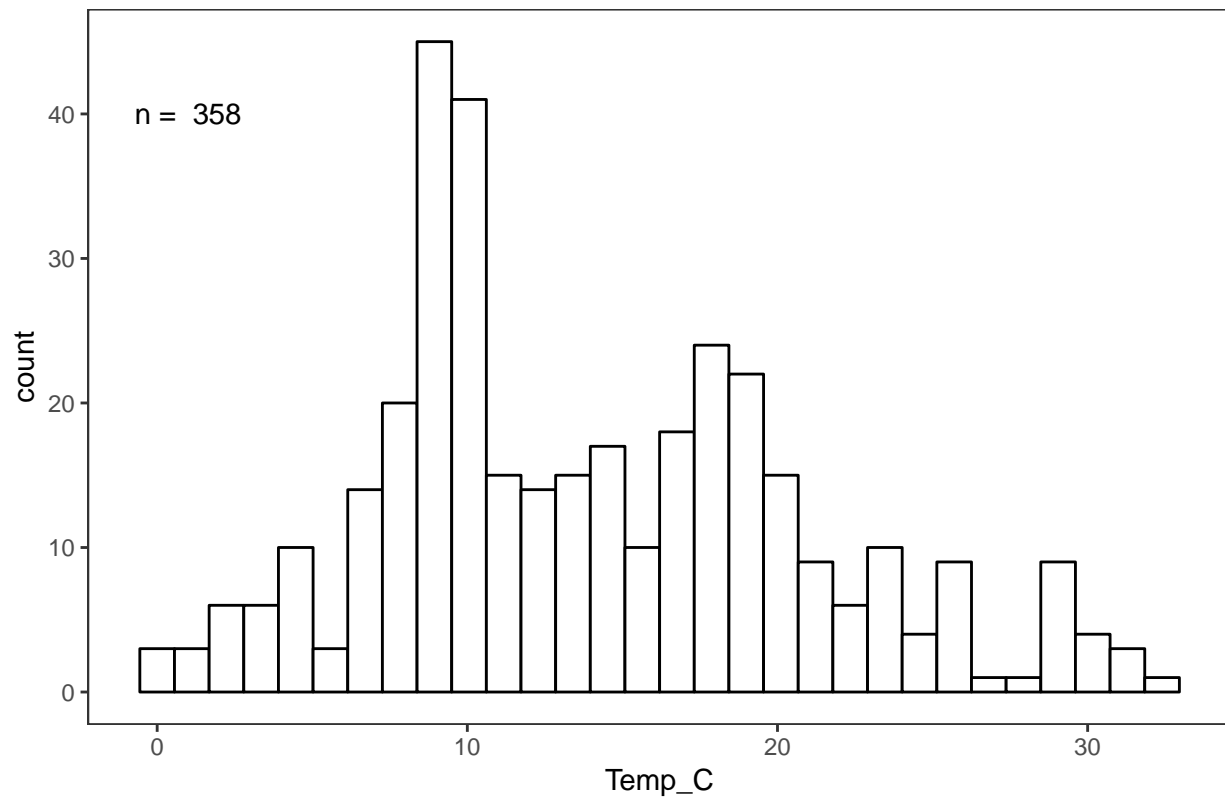
##Beck Plots

#Temp Range

```
ggplot(TempData_Beck, aes(x = Temp_C)) + geom_histogram(colour = 'black', fill = 'white') +
  annotate('text', x = 1, y = 40, label = paste("n = ", nrow(subset(TempData_Beck, Temp_C != "NA")))) +
  ggtitle("Beck Temp")
```

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

Beck Temp



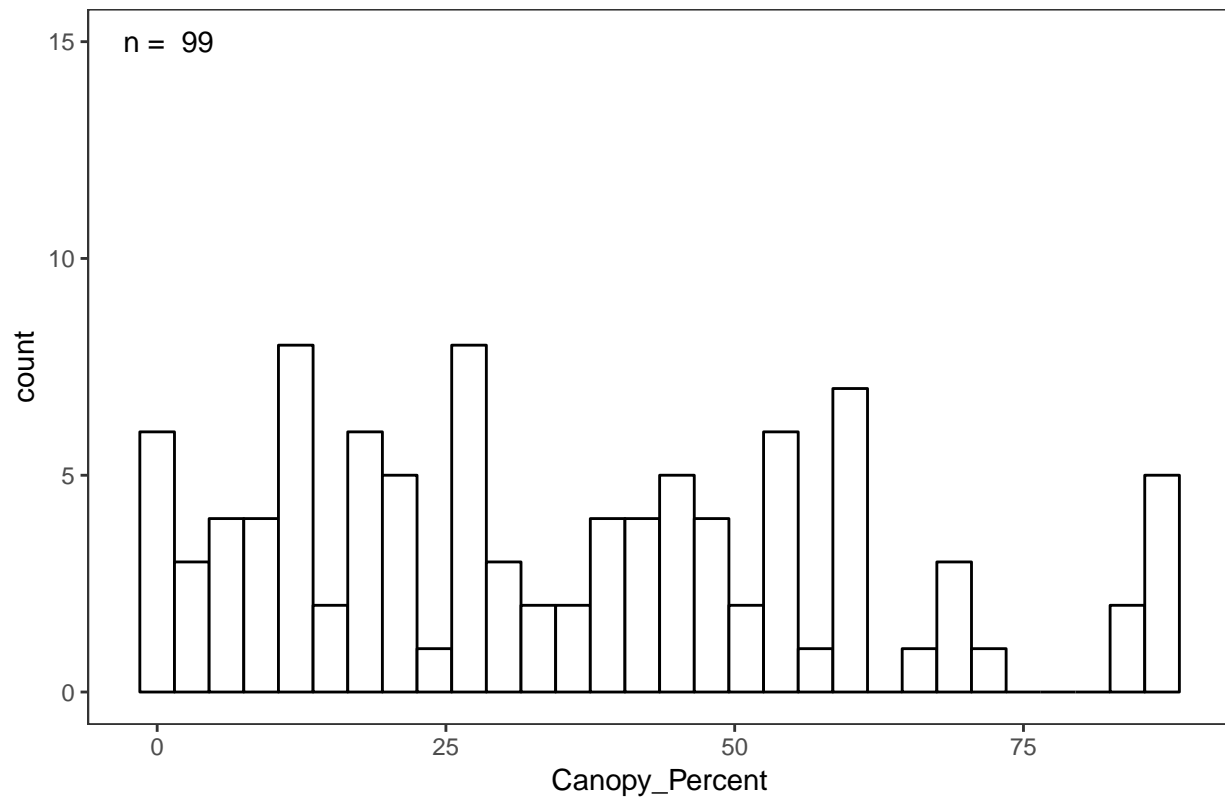
#Canopy % Range

```
ggplot(TempData_Beck, aes(x = Canopy_Percent)) + geom_histogram(colour = 'black', fill = 'white') +
  annotate('text', x = 1, y = 15, label = paste("n = ", nrow(subset(TempData_Beck, Canopy_Percent != "NA")))) +
  ggtitle("Beck Canopy Percent")
```

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

```
## Warning: Removed 259 rows containing non-finite values (stat_bin).
```

Beck Canopy Percent



#DIN Range

```
Beck_log_DIN = ggplotGrob(ggplot(TempData_Beck, aes(x = log10(DIN_ug_L))) + geom_histogram(colour = 'black', fill = 'white'))
```

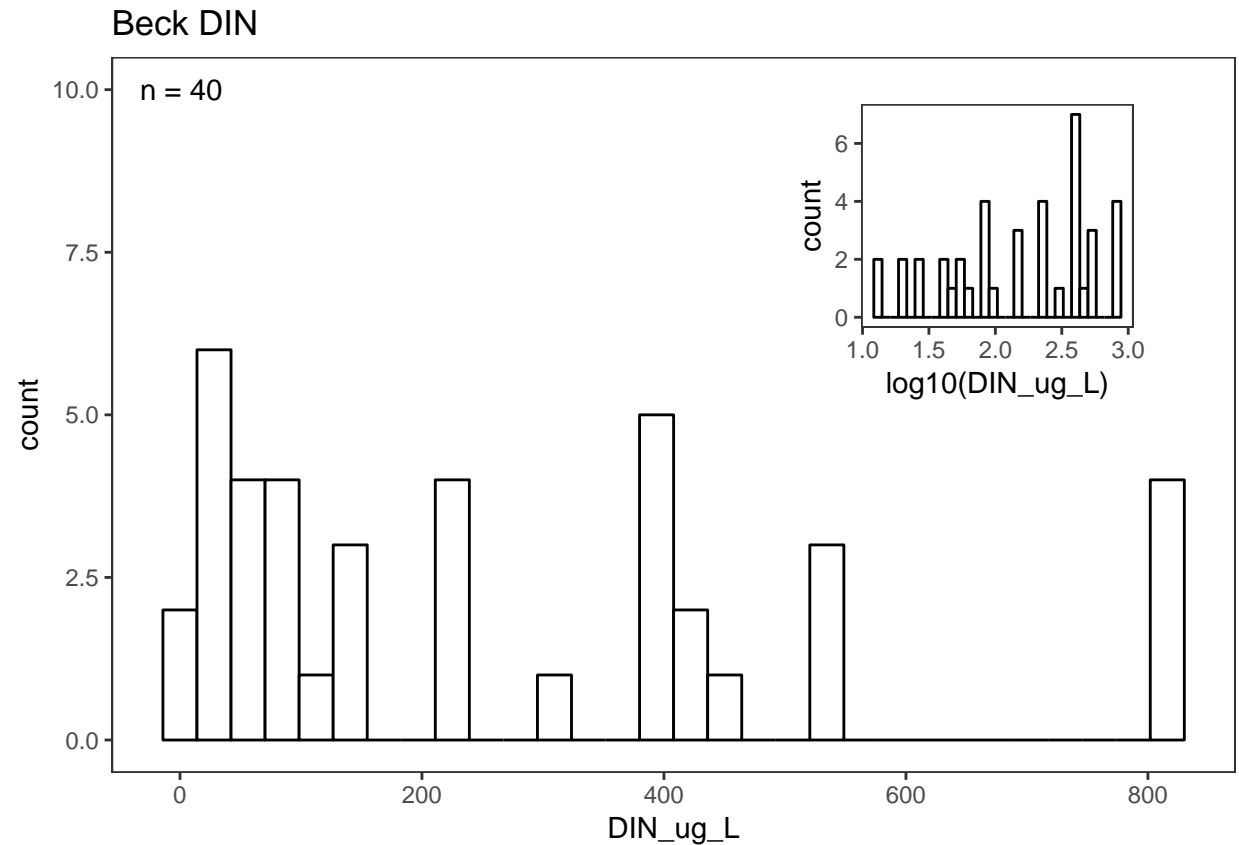
```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

```
## Warning: Removed 318 rows containing non-finite values (stat_bin).
```

```
ggplot(TempData_Beck, aes(x = DIN_ug_L)) + geom_histogram(colour = 'black', fill = 'white') +
  annotate('text', x = 1, y = 10, label = paste("n =", nrow(subset(TempData_Beck, DIN_ug_L!="NA")))) +
  annotation_custom(grob = Beck_log_DIN, xmin = 500, xmax = 800, ymin = 5, ymax = 10) +
  ggtitle("Beck DIN")
```

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

```
## Warning: Removed 318 rows containing non-finite values (stat_bin).
```



#TN Range

```
Beck_log_TN = ggplotGrob(ggplot(TempData_Beck, aes(x = log10(TN_mg_L))) + geom_histogram(colour = 'black', fill = 'white'))
```

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

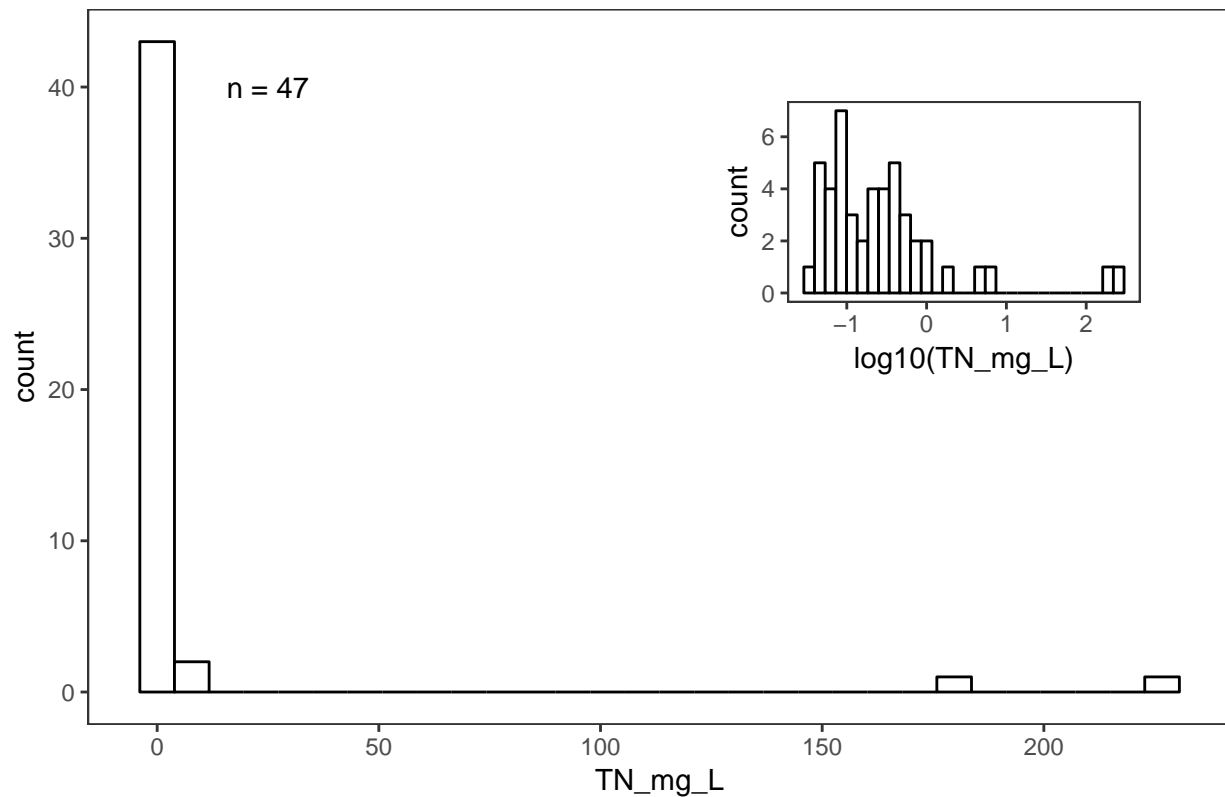
```
## Warning: Removed 311 rows containing non-finite values (stat_bin).
```

```
ggplot(TempData_Beck, aes(x = TN_mg_L)) + geom_histogram(colour = 'black', fill = 'white') +  
  annotate('text', x = 25, y = 40, label = paste("n =", nrow(subset(TempData_Beck, TN_mg_L!="NA")))) +  
  annotation_custom(grob = Beck_log_TN, xmin = 125, xmax = 225, ymin = 20, ymax = 40) +  
  ggtitle("Beck TN")
```

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

```
## Warning: Removed 311 rows containing non-finite values (stat_bin).
```

Beck TN



#SRP Range

```
Beck_log_SRP = ggplotGrob(ggplot(TempData_Beck, aes(x = log10(SRP_ug_L))) + geom_histogram(colour = 'black', fill = 'white'))
```

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

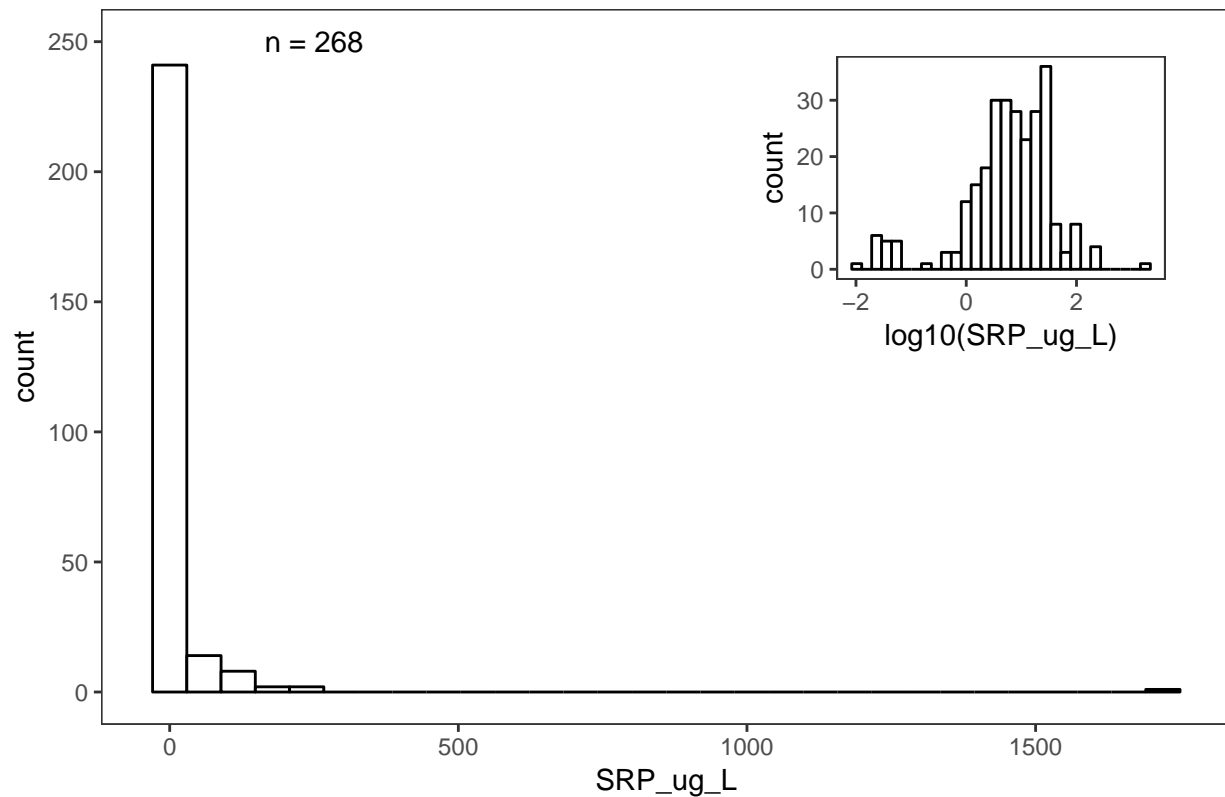
```
## Warning: Removed 90 rows containing non-finite values (stat_bin).
```

```
ggplot(TempData_Beck, aes(x = SRP_ug_L)) + geom_histogram(colour = 'black', fill = 'white') +  
  annotate('text', x = 250, y = 250, label = paste("n =", nrow(subset(TempData_Beck, SRP_ug_L != "NA")))) +  
  annotation_custom(grob = Beck_log_SRP, xmin = 1000, xmax = 1750, ymin = 125, ymax = 250) +  
  ggtitle("Beck SRP")
```

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

```
## Warning: Removed 90 rows containing non-finite values (stat_bin).
```

Beck SRP



#TP Range

```
Beck_log_TP = ggplotGrob(ggplot(TempData_Beck, aes(x = log10(TP_ug_L))) + geom_histogram(colour = 'black', fill = 'white'))
```

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

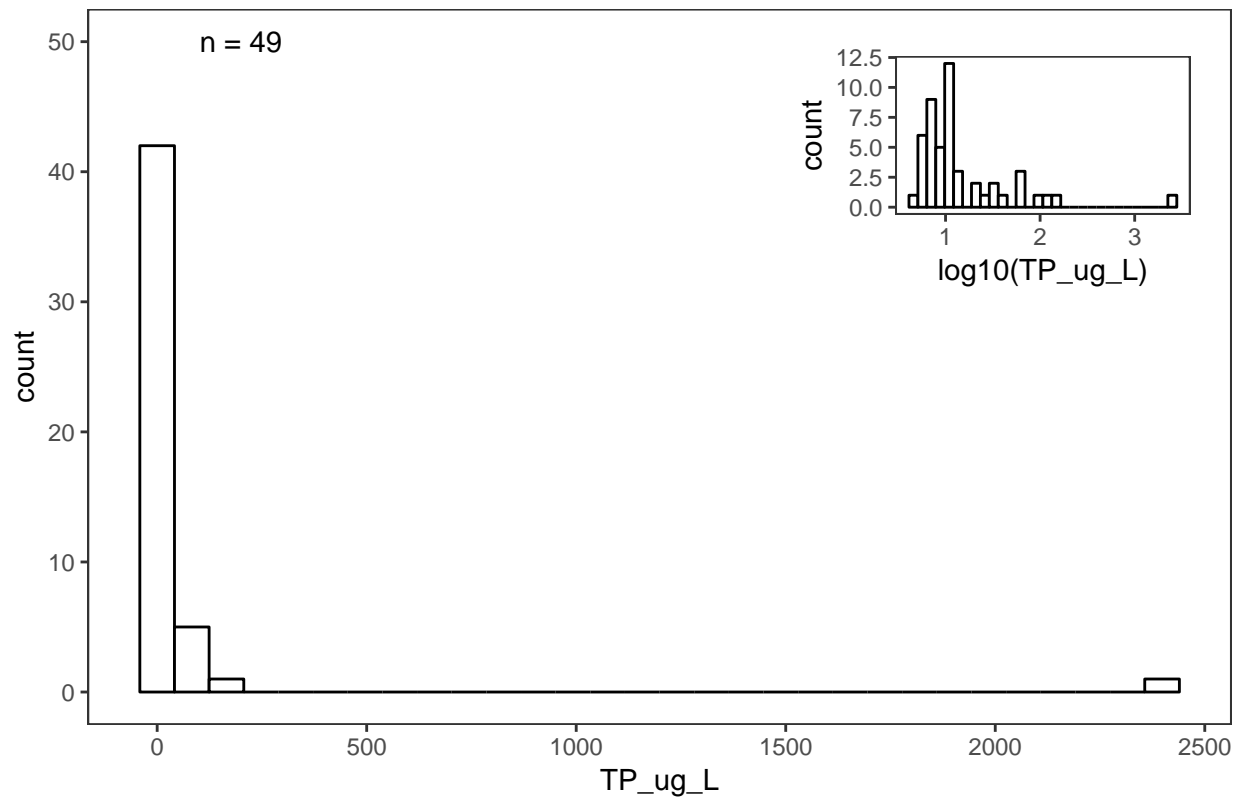
```
## Warning: Removed 309 rows containing non-finite values (stat_bin).
```

```
ggplot(TempData_Beck, aes(x = TP_ug_L)) + geom_histogram(colour = 'black', fill = 'white') +  
  annotate('text', x = 200, y = 50, label = paste("n =", nrow(subset(TempData_Beck, TP_ug_L!="NA")))) +  
  annotation_custom(grob = Beck_log_TP, xmin = 1500, xmax = 2500, ymin = 30, ymax = 50)+  
  ggtitle("Beck TP")
```

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

```
## Warning: Removed 309 rows containing non-finite values (stat_bin).
```

Beck TP

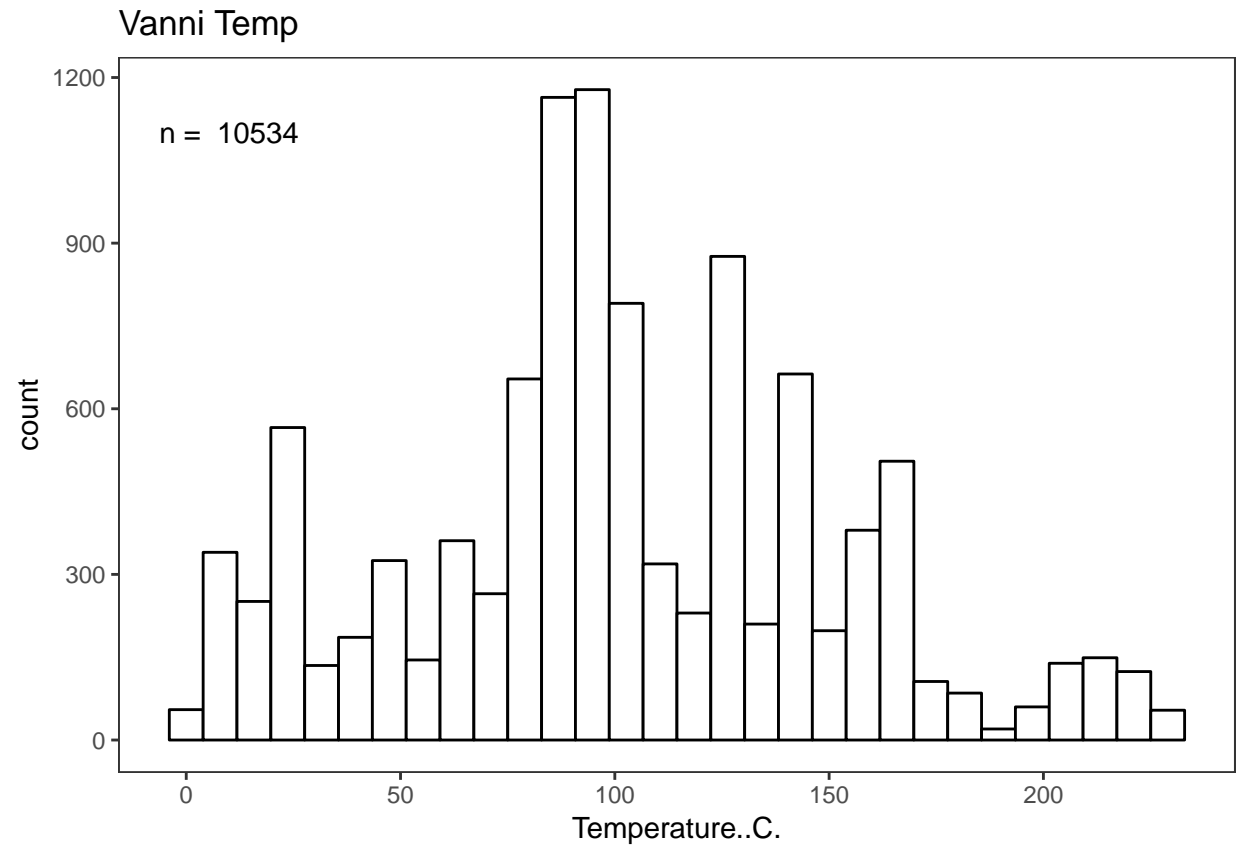


##Vanni Plots from 2) Exploratory script

#Temp Range

```
ggplot(TempData_Vanni, aes(x = Temperature..C.)) + geom_histogram(colour = 'black', fill = 'white') +
  annotate('text', x = 10, y = 1100, label = paste("n = ", nrow(subset(TempData_Vanni, Temperature..C. !=
  ggtitle("Vanni Temp")
```

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

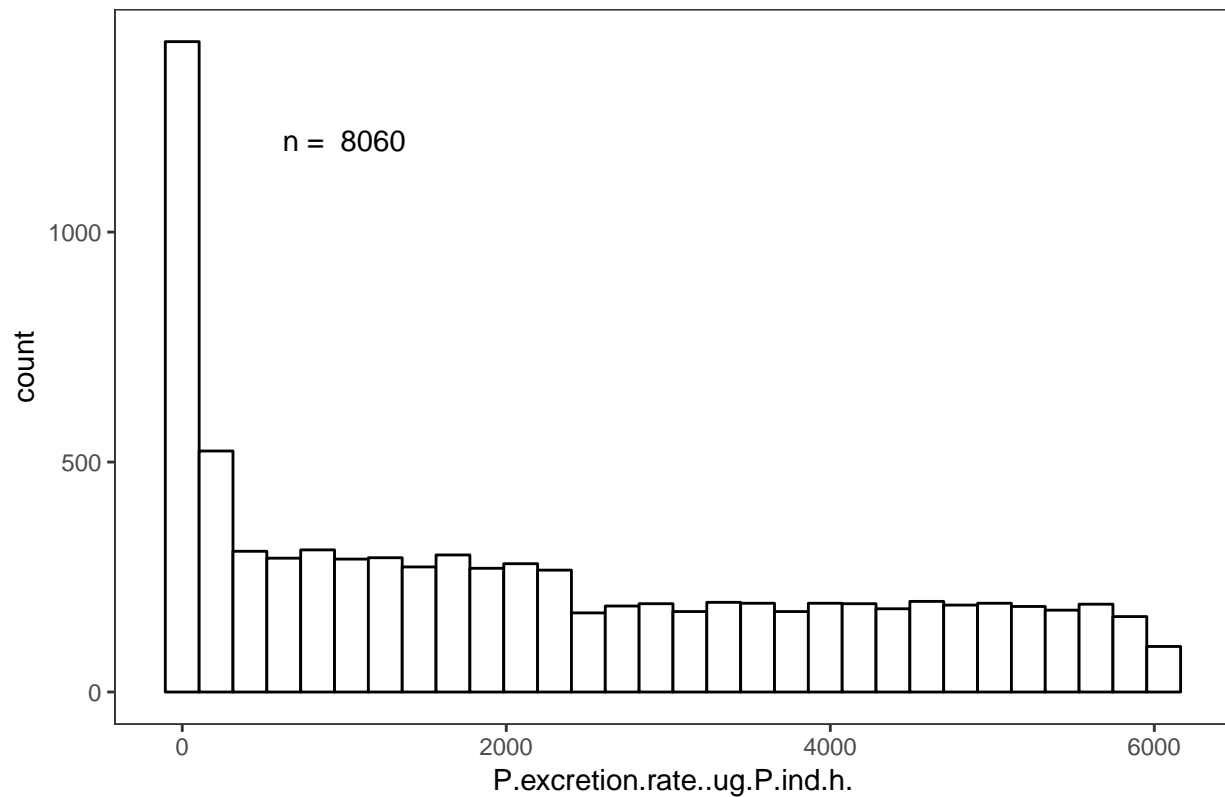


#Excretion SRP

```
ggplot(subset(TempData_Vanni, P.form=="SRP"), aes(x = P.excretion.rate..ug.P.ind.h.)) + geom_histogram(
  annotate('text', x = 1000, y = 1200, label = paste("n = ", nrow(subset(TempData_Vanni, P.form=="SRP")))) +
  ggtitle("Vanni SRP Excretion Rate")
```

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

Vanni SRP Excretion Rate

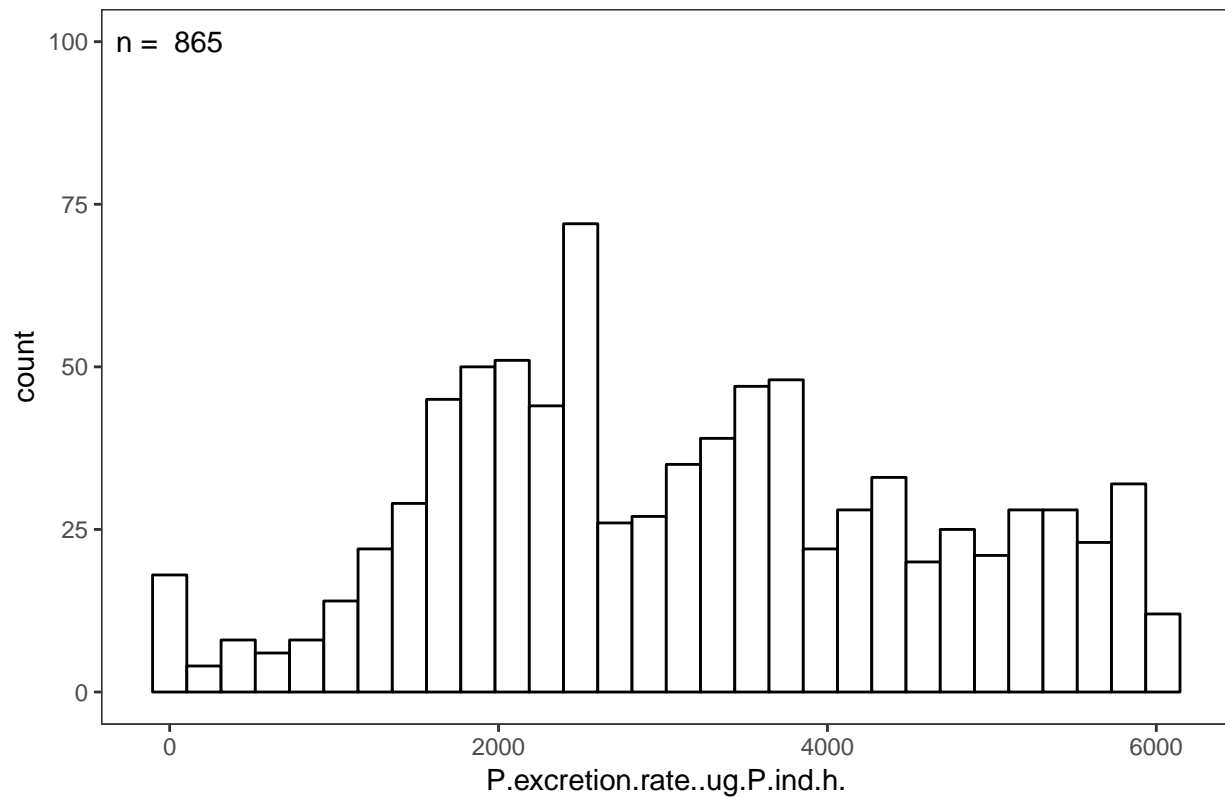


#Excretion TP

```
ggplot(subset(TempData_Vanni, P.form=="TP"), aes(x = P.excretion.rate..ug.P.ind.h.)) + geom_histogram(c
  annotate('text', x = 0, y = 100, label = paste("n = ", nrow(subset(TempData_Vanni, P.form=="TP"& P.exc
  ggtitle("Vanni TP Excretion Rate")
```

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

Vanni TP Excretion Rate

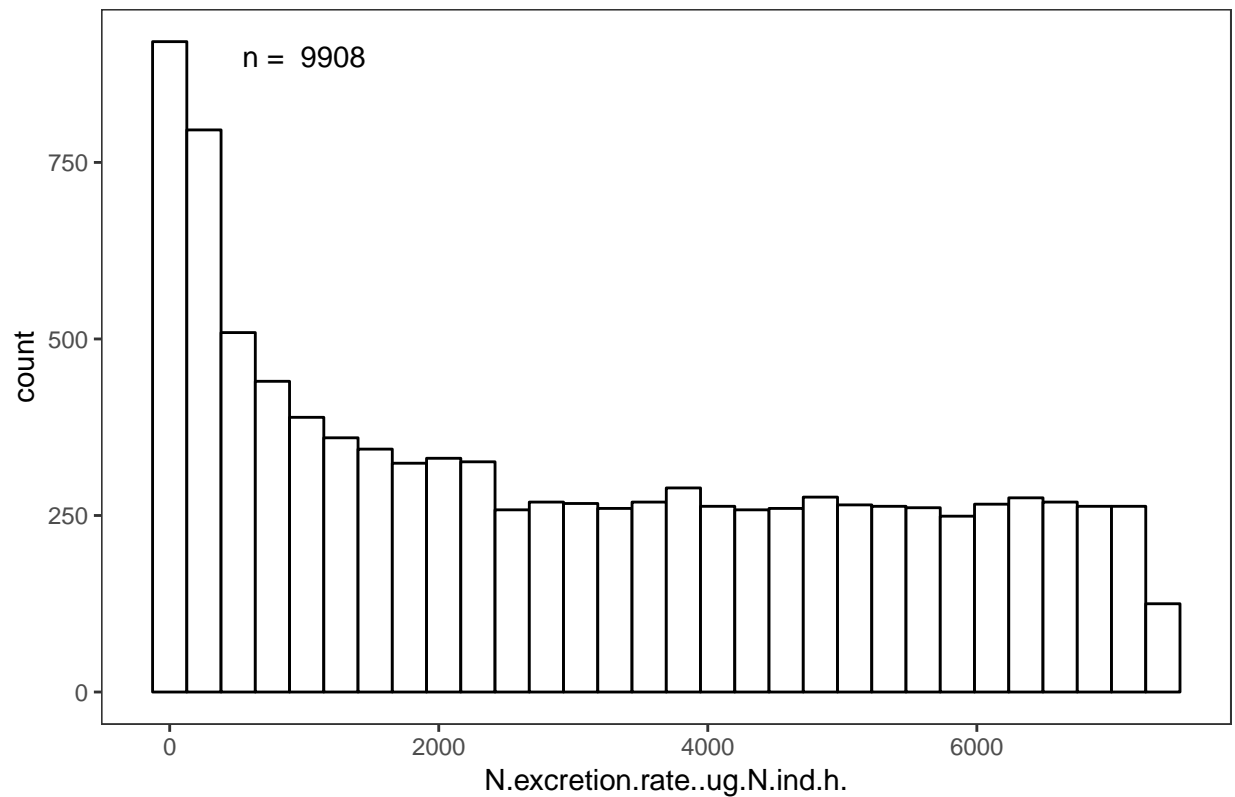


#Excretion Ammonium

```
ggplot(subset(TempData_Vanni, N.form=="NH4"), aes(x = N.excretion.rate..ug.N.ind.h.)) + geom_histogram(
  annotate('text', x = 1000, y = 900, label = paste("n = ", nrow(subset(TempData_Vanni, N.form=="NH4" & N
  ggtitle("Vanni NH4 Excretion Rate"))
```

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

Vanni NH4 Excretion Rate

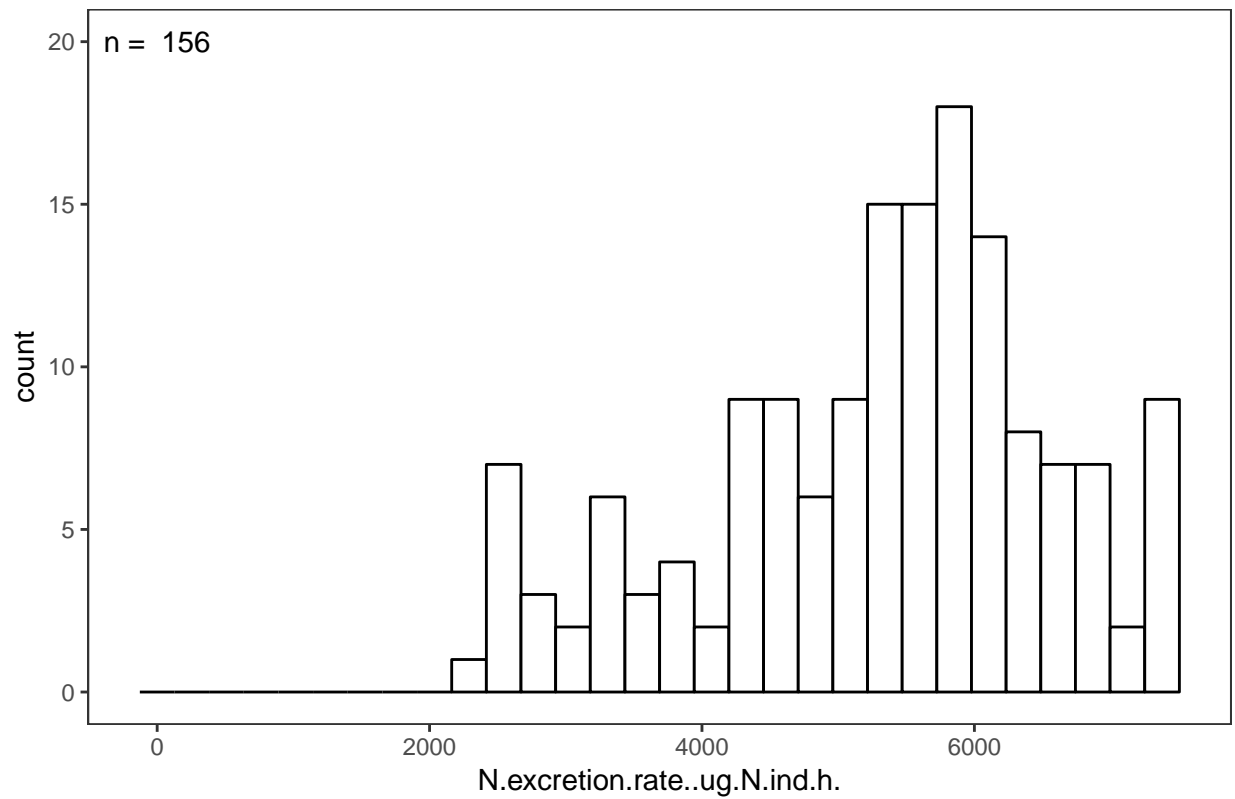


```
#Excretion TN
```

```
ggplot(subset(TempData_Vanni, N.form=="TN"), aes(x = N.excretion.rate..ug.N.ind.h.)) + geom_histogram(c
  annotate('text', x = 0, y = 20, label = paste("n = ", nrow(subset(TempData_Vanni, N.form=="TN" & N.excr
  ggtitle("Vanni TN Excretion Rate")
```

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

Vanni TN Excretion Rate



*#Excretion N:P (If N excretion or P excretion was NA, value was calculated as "1",
#so only included values >1)*

```
ggplot(subset(TempData_Vanni, Excreted.N.P..molar.>1), aes(x = Excreted.N.P..molar.)) + geom_histogram(  
  annotate('text', x = 500, y = 300, label = paste("n = ", nrow(subset(TempData_Vanni, Excreted.N.P..molar.>1))),  
  ggtitle("Vanni Excretion N:P Molar")
```

