

Project_DataAnalysis

Set Working Directory and Load Data

Format Dates

```
#Format Site Number to Character
```

```
#Ttest
```

```
#Is there a difference between upstream and downstream nutrient levels?  
### continuous response (N or P)  
### categorical explanatory
```

Nitrogen

```
summary(EC_Flow.Nutrients_Wide$Mixed_N)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.     NA's  
##      0.37   1.40    3.46   16.26   29.00   91.50   5490
```

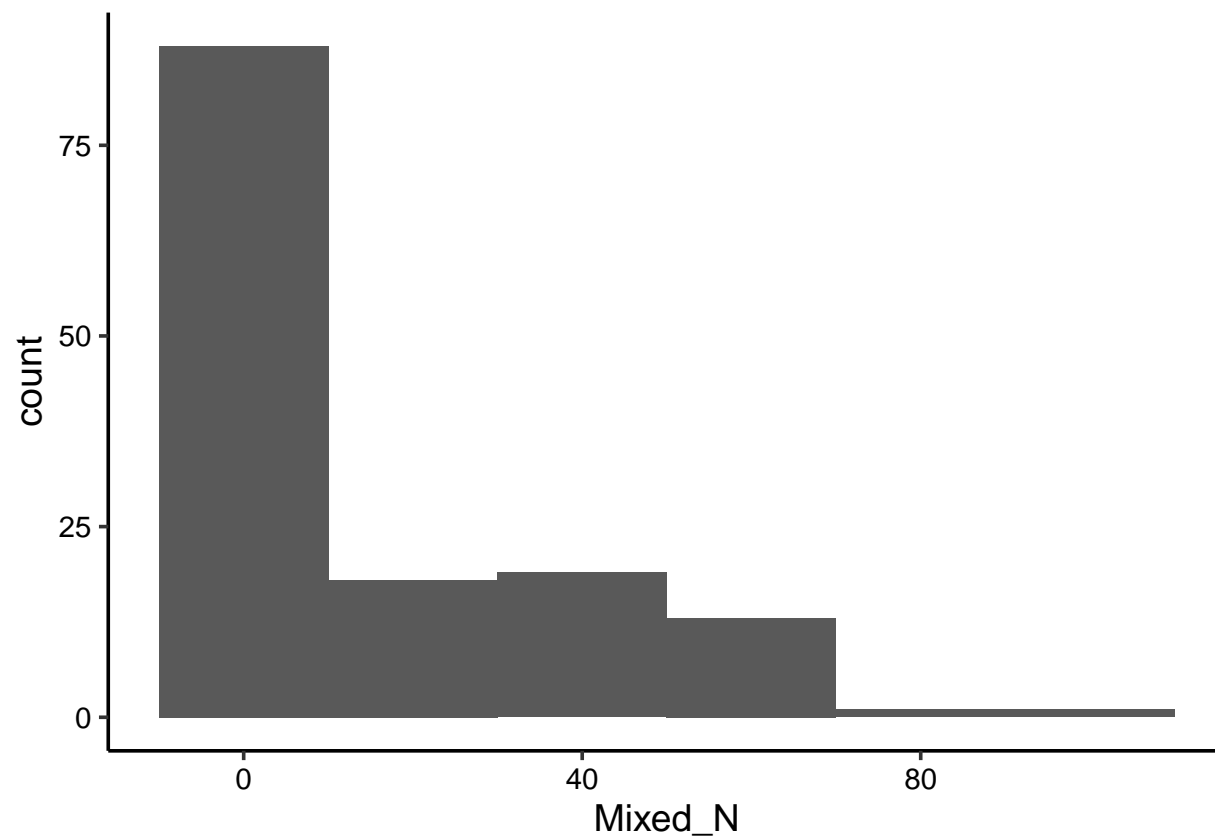
```
# Evaluate assumption of normal distribution  
shapiro.test((EC_Flow.Nutrients_Wide$Mixed_N))
```

```
##  
##  Shapiro-Wilk normality test  
##  
## data:  (EC_Flow.Nutrients_Wide$Mixed_N)  
## W = 0.75214, p-value = 4.311e-14
```

```
### shapiro wilke test the null hypothesis is that the data are a normal distribution  
### p value <0.05 shows we reject null and data is not normally distributed  
##### "not well approximated by a normal distribution"
```

```
ggplot(EC_Flow.Nutrients_Wide, aes(x = Mixed_N)) +  
  geom_histogram(binwidth = 20)
```

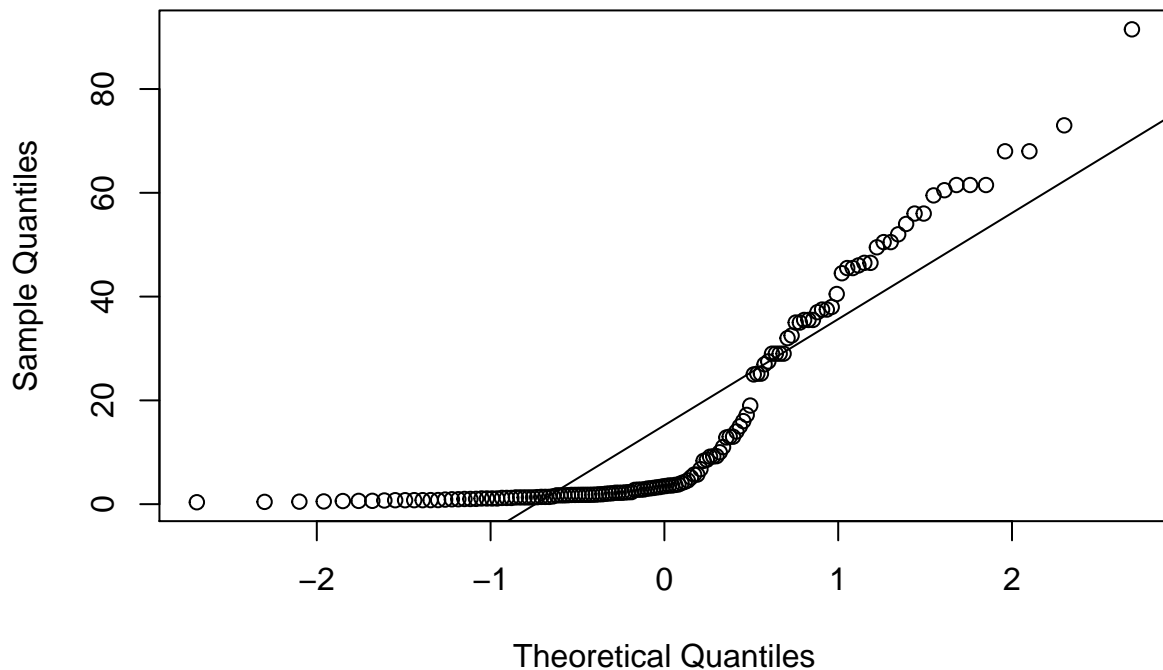
```
## Warning: Removed 5490 rows containing non-finite values (stat_bin).
```



histogram shows data is not very rightly skewed

```
qqnorm(EC_Flow.Nutrients_Wide$Mixed_N); qqline(EC_Flow.Nutrients_Wide$Mixed_N)
```

Normal Q-Q Plot



some samples are higher than normal and some are lower

```
N.onesample <- t.test(EC_Flow.Nutrients_Wide$Mixed_N, mu = 10, alternative = "greater")
N.onesample
```

```
##
## One Sample t-test
##
## data: EC_Flow.Nutrients_Wide$Mixed_N
## t = 3.5158, df = 139, p-value = 0.0002962
## alternative hypothesis: true mean is greater than 10
## 95 percent confidence interval:
## 13.31248 Inf
## sample estimates:
## mean of x
## 16.26161
```

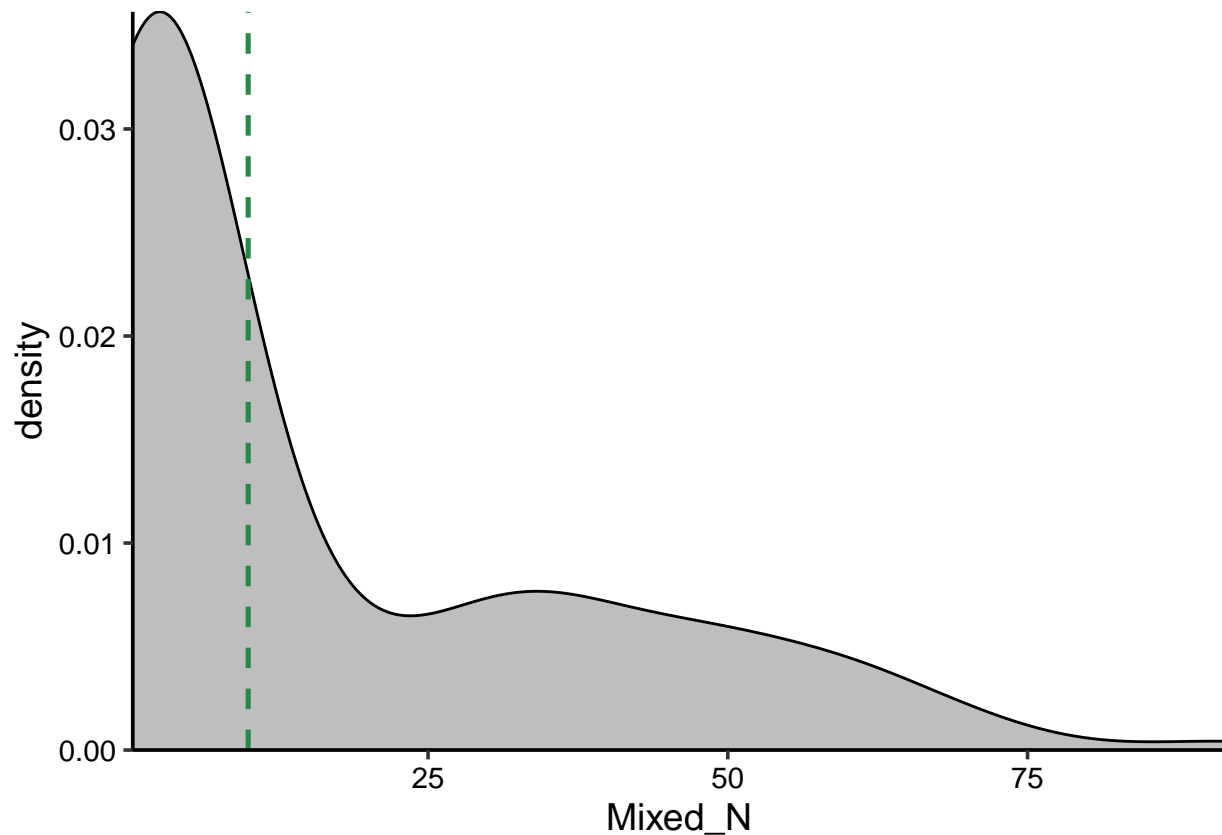
*# Null hypothesis is that mean = 10, alternative is mean is less than 10
p value >0.05 so we can reject the null the mean =10*

plot

```
N.plot <- ggplot(EC_Flow.Nutrients_Wide, aes(x = Mixed_N)) +
  #geom_density(stat = "count", fill = "gray") + #<-shows count so it's more jagged
  geom_density(fill = "gray") +
  geom_vline(xintercept = 10, color = "#238b45", lty = 2, size = 0.9) +
  scale_x_continuous(expand = c(0, 0)) + scale_y_continuous(expand = c(0, 0))
```

```
print(N.plot)
```

```
## Warning: Removed 5490 rows containing non-finite values (stat_density).
```



Nitrogen Results > Nitrogen measurements in Eller Creek were not significantly greater than 10 mg/L, the Maximum Contaminant Level for Nitrogen in drinking water (one sample t-test; $t = 3.515$, $df = 139$, $p < 0.001$)

Phosphorus

```
summary(EC_Flow.Nutrients_Wide$TP)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.     NA's
##      0.039  0.087   0.159   1.577  0.660   34.500    5387
```

```
# Evaluate assumption of normal distribution
```

```
shapiro.test((EC_Flow.Nutrients_Wide$TP))
```

```
##
```

```
## Shapiro-Wilk normality test
```

```
##
```

```
## data: (EC_Flow.Nutrients_Wide$TP)
```

```
## W = 0.47359, p-value < 2.2e-16
```

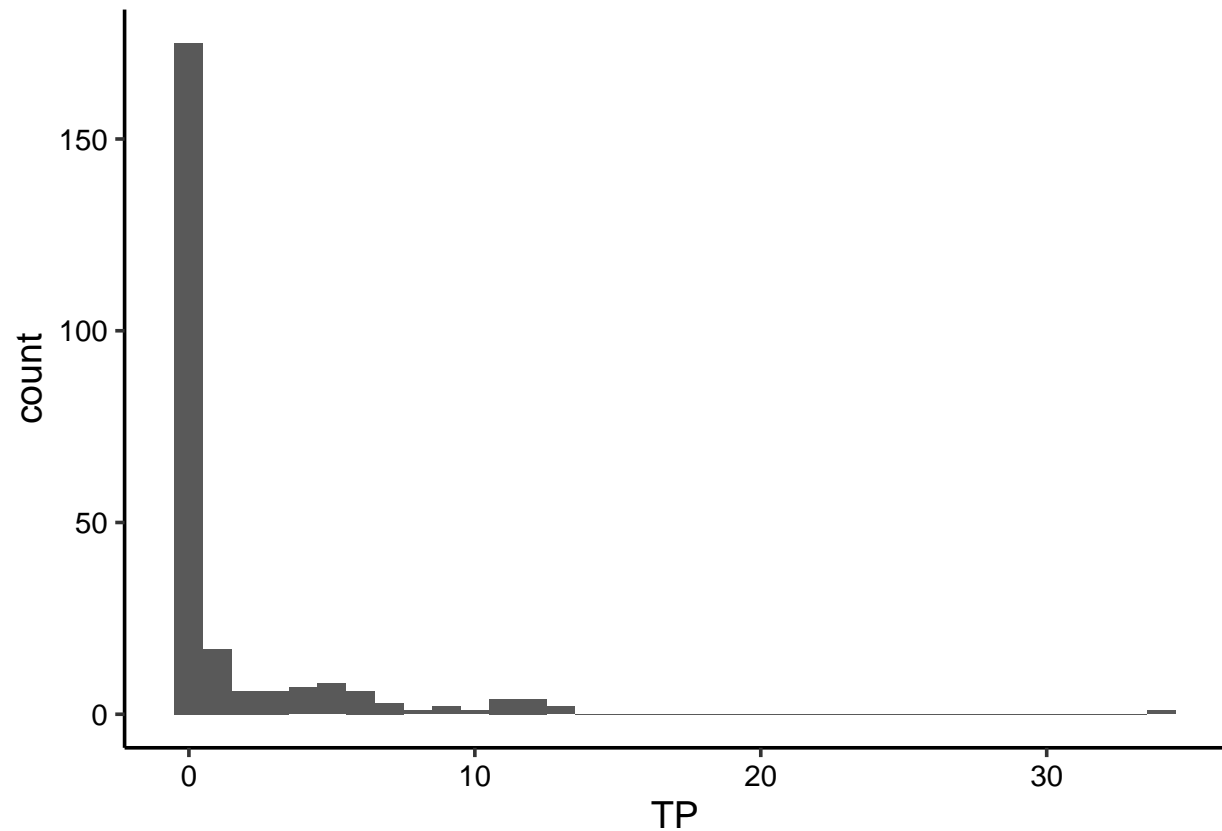
```
### shapiro wilke test the null hypothesis is that the data are a normal distribution
```

```
### p value < 0.05 shows we reject null and data is not normally distributed
```

```
##### "not well approximated by a normal distribution"
```

```
ggplot(EC_Flow.Nutrients_Wide, aes(x = TP)) +  
  geom_histogram(binwidth = 1)
```

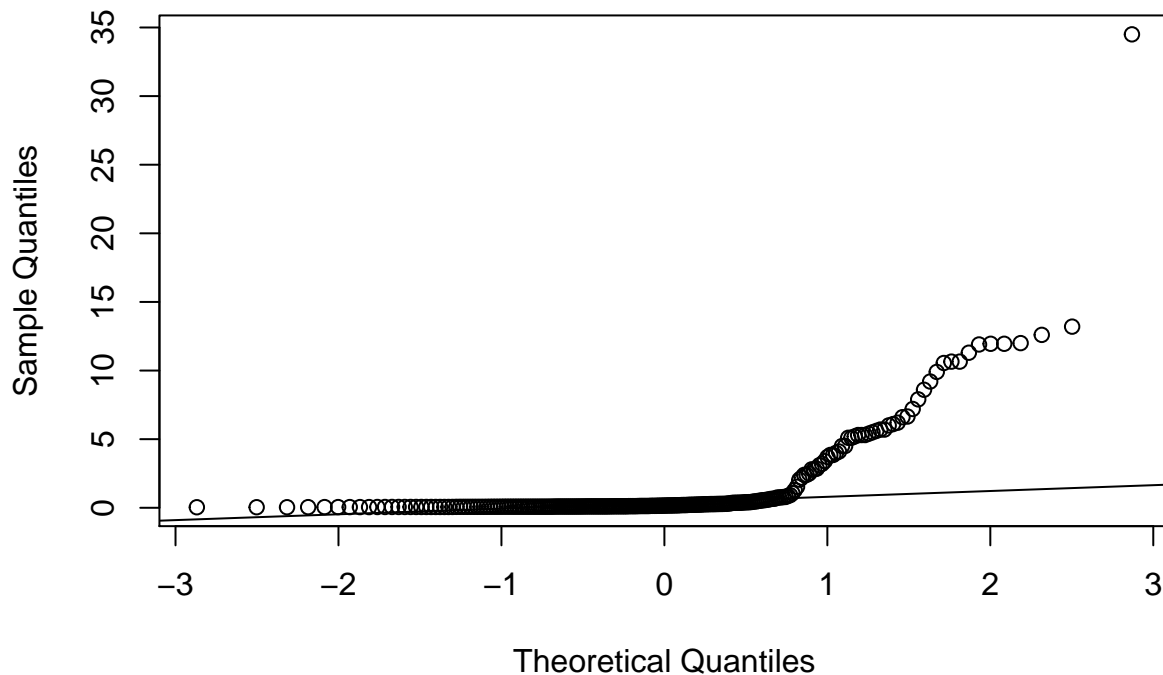
```
## Warning: Removed 5387 rows containing non-finite values (stat_bin).
```



```
### histogram shows data is not very rightly skewed
```

```
qqnorm(EC_Flow.Nutrients_Wide$TP); qqline(EC_Flow.Nutrients_Wide$TP)
```

Normal Q-Q Plot



samples are higher than normal distribution

```
P.onesample <- t.test(EC_Flow.Nutrients_Wide$TP, mu = 1, alternative = "greater")
P.onesample
```

```
##
## One Sample t-test
##
## data: EC_Flow.Nutrients_Wide$TP
## t = 2.5264, df = 242, p-value = 0.006081
## alternative hypothesis: true mean is greater than 1
## 95 percent confidence interval:
##  1.199743      Inf
## sample estimates:
## mean of x
##  1.576553
```

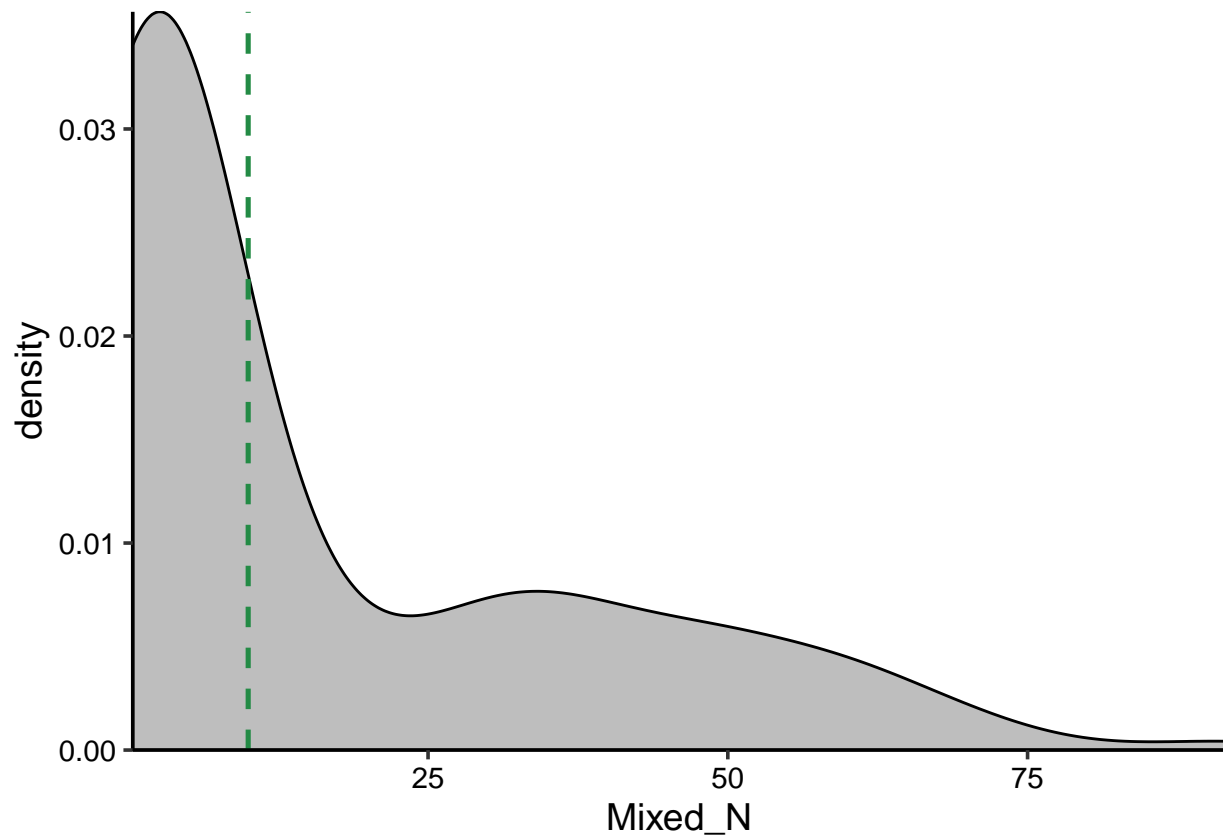
Null hypothesis is that mean = 1, alternative is mean is less than 1
p value < 0.05 so we can reject the null the mean = 1

plot

```
P.plot <- ggplot(EC_Flow.Nutrients_Wide, aes(x = TP)) +
  #geom_density(stat = "count", fill = "gray") + #<-shows count so it's more jagged
  geom_density(fill = "gray") +
  geom_vline(xintercept = 1, color = "#238b45", lty = 2, size = 0.9) +
  scale_x_continuous(expand = c(0, 0)) + scale_y_continuous(expand = c(0, 0))
```

```
print(N.plot)
```

```
## Warning: Removed 5490 rows containing non-finite values (stat_density).
```



```
##### Phosphorus Results > Phosphorus measurments in ellerbe creek were significantly greater than 1
mg/L, the Maximum Contaminant Level for phosphorus in water (one sample t-test; t= 2.5264, df = 242, p
= 0.006)
```

```
#Two Sample t-test (Upstream vs. Downstream) ### Nitrogen
```

```
shapiro.test(Club.Gorman_Flow.Nutrients_Wide$Mixed_N[Club.Gorman_Flow.Nutrients_Wide$Location == "Upstr
```

```
##
```

```
## Shapiro-Wilk normality test
```

```
##
```

```
## data: Club.Gorman_Flow.Nutrients_Wide$Mixed_N[Club.Gorman_Flow.Nutrients_Wide$Location == "Upst
```

```
## W = 0.93963, p-value = 0.06625
```

```
shapiro.test(Club.Gorman_Flow.Nutrients_Wide$Mixed_N[Club.Gorman_Flow.Nutrients_Wide$Location == "Downs
```

```
##
```

```
## Shapiro-Wilk normality test
```

```
##
```

```
## data: Club.Gorman_Flow.Nutrients_Wide$Mixed_N[Club.Gorman_Flow.Nutrients_Wide$Location == "Down
```

```
## W = 0.87157, p-value = 1.642e-07
```

```
var.test(Club.Gorman_Flow.Nutrients_Wide$Mixed_N ~ Club.Gorman_Flow.Nutrients_Wide$Location)
```

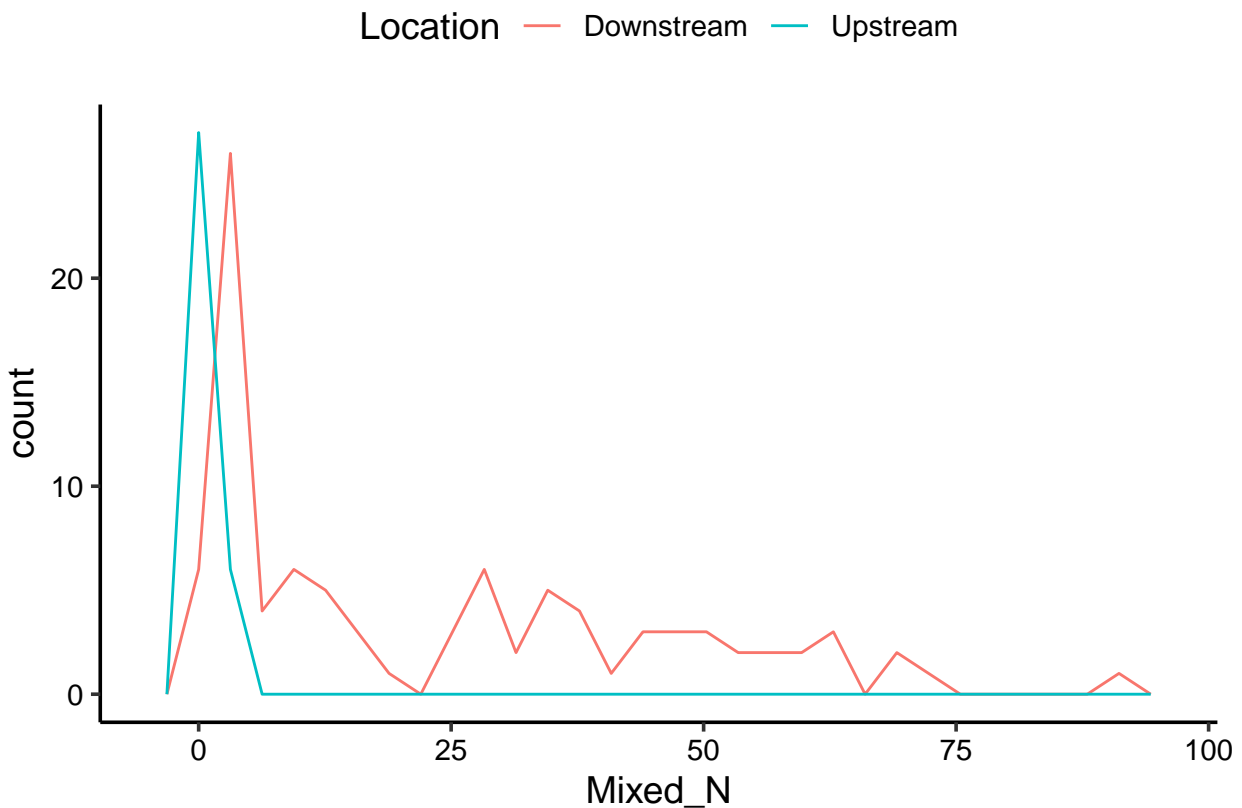
```
##
```

```
## F test to compare two variances
##
## data: Club.Gorman_Flow.Nutrients_Wide$Mixed_N by Club.Gorman_Flow.Nutrients_Wide$Location
## F = 1862.5, num df = 93, denom df = 32, p-value < 2.2e-16
## alternative hypothesis: true ratio of variances is not equal to 1
## 95 percent confidence interval:
## 1005.094 3181.440
## sample estimates:
## ratio of variances
## 1862.531
```

```
### "var.test" test variance
### asks are the variances equal - what is the diff between variances
### results: variances are significantly different
### results: violate assumption of normality and equal variance
```

```
ggplot(Club.Gorman_Flow.Nutrients_Wide, aes(x = Mixed_N, color = Location)) +
  geom_freqpoly()
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## Warning: Removed 9854 rows containing non-finite values (stat_bin).
```



```
N.twosample <- t.test(Club.Gorman_Flow.Nutrients_Wide$Mixed_N ~ Club.Gorman_Flow.Nutrients_Wide$Location)
N.twosample
```

```
##
## Welch Two Sample t-test
```



```
##
## data: Club.Gorman_Flow.Nutrients_Wide$Mixed_N by Club.Gorman_Flow.Nutrients_Wide$Location
## t = 9.7326, df = 93.284, p-value = 7.313e-16
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## 17.85949 27.01521
## sample estimates:
## mean in group Downstream mean in group Upstream
## 23.578564 1.141212
```

Phosphorus

```
shapiro.test(Club.Gorman_Flow.Nutrients_Wide$TP[Club.Gorman_Flow.Nutrients_Wide$Location == "Upstream"])
```

```
##
## Shapiro-Wilk normality test
##
## data: Club.Gorman_Flow.Nutrients_Wide$TP[Club.Gorman_Flow.Nutrients_Wide$Location == "Upstream"]
## W = 0.88509, p-value = 0.001368
```

```
shapiro.test(Club.Gorman_Flow.Nutrients_Wide$TP[Club.Gorman_Flow.Nutrients_Wide$Location == "Downstream"])
```

```
##
## Shapiro-Wilk normality test
##
## data: Club.Gorman_Flow.Nutrients_Wide$TP[Club.Gorman_Flow.Nutrients_Wide$Location == "Downstream"]
## W = 0.52423, p-value < 2.2e-16
```

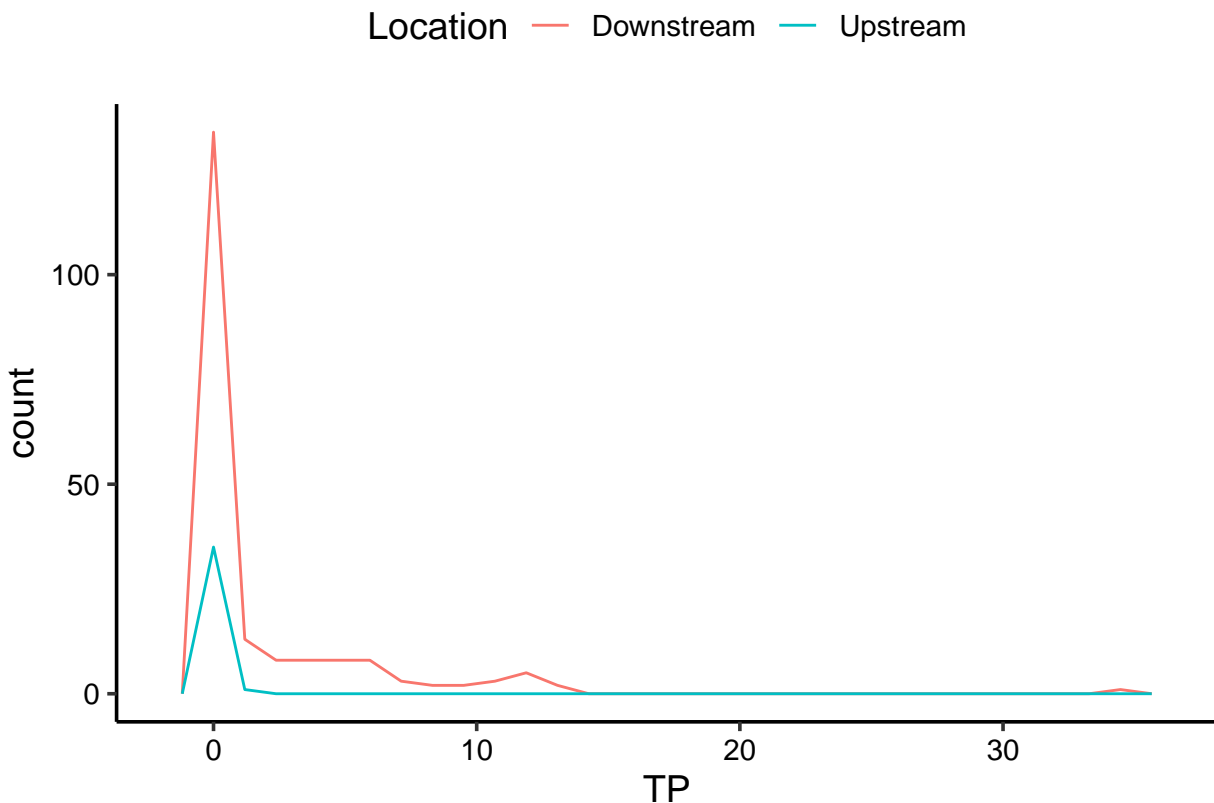
```
var.test(Club.Gorman_Flow.Nutrients_Wide$TP ~ Club.Gorman_Flow.Nutrients_Wide$Location)
```

```
##
## F test to compare two variances
##
## data: Club.Gorman_Flow.Nutrients_Wide$TP by Club.Gorman_Flow.Nutrients_Wide$Location
## F = 524.21, num df = 196, denom df = 35, p-value < 2.2e-16
## alternative hypothesis: true ratio of variances is not equal to 1
## 95 percent confidence interval:
## 298.9121 837.7713
## sample estimates:
## ratio of variances
## 524.2082
```

```
### "var.test" test variance
### asks are the variances equal - what is the diff between variances
### results: variances are significantly different
### results: violate assumption of normality and equal variance
```

```
ggplot(Club.Gorman_Flow.Nutrients_Wide, aes(x = TP, color = Location)) +
  geom_freqpoly()
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## Warning: Removed 9748 rows containing non-finite values (stat_bin).
```



```
P.twosample <- t.test(Club.Gorman_Flow.Nutrients_Wide$TP ~ Club.Gorman_Flow.Nutrients_Wide$Location)
P.twosample
```

```
##
## Welch Two Sample t-test
##
## data: Club.Gorman_Flow.Nutrients_Wide$TP by Club.Gorman_Flow.Nutrients_Wide$Location
## t = 6.0272, df = 199.99, p-value = 7.906e-09
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  1.127322 2.223637
## sample estimates:
## mean in group Downstream mean in group Upstream
##           1.899354           0.223875
```

Wilcoxon Test (Don't Assume Normality)

Nitrogen

```
N.onesample.wilcox <- wilcox.test(Club.Gorman_Flow.Nutrients_Wide$Mixed_N, mu = 10, alternative = "greater")
N.onesample.wilcox
```

```
##
## Wilcoxon signed rank test with continuity correction
##
## data: Club.Gorman_Flow.Nutrients_Wide$Mixed_N
```

```
## V = 4687, p-value = 0.04743
## alternative hypothesis: true location is greater than 10
N.twosample.wilcox <- wilcox.test(Club.Gorman_Flow.Nutrients_Wide$Mixed_N ~ Club.Gorman_Flow.Nutrients_Wide$Location, data = EC_Flow.Nutrients_Wide, na.rm = TRUE)
N.twosample.wilcox

##
## Wilcoxon rank sum test with continuity correction
##
## data: Club.Gorman_Flow.Nutrients_Wide$Mixed_N by Club.Gorman_Flow.Nutrients_Wide$Location
## W = 2988, p-value = 2.832e-15
## alternative hypothesis: true location shift is not equal to 0
```

Phosphorus

```
P.onesample.wilcox <- wilcox.test(Club.Gorman_Flow.Nutrients_Wide$TP, mu = 1, alternative = "greater")
P.onesample.wilcox

##
## Wilcoxon signed rank test with continuity correction
##
## data: Club.Gorman_Flow.Nutrients_Wide$TP
## V = 10454, p-value = 0.999
## alternative hypothesis: true location is greater than 1
P.twosample.wilcox <- wilcox.test(Club.Gorman_Flow.Nutrients_Wide$TP ~ Club.Gorman_Flow.Nutrients_Wide$Location, data = EC_Flow.Nutrients_Wide, na.rm = TRUE)
P.twosample.wilcox

##
## Wilcoxon rank sum test with continuity correction
##
## data: Club.Gorman_Flow.Nutrients_Wide$TP by Club.Gorman_Flow.Nutrients_Wide$Location
## W = 4136, p-value = 0.1128
## alternative hypothesis: true location shift is not equal to 0
```

Simple Linear Regression

Is flow a significant predictor of nutrient levels? continuous response (N or P) continuous predictor (discharge)

###Nitrogen

```
NitrogenClub.regression <- lm(data = EC_Flow.Nutrients_Wide, Mixed_N ~ Discharge.C)
summary(NitrogenClub.regression)
```

```
##
## Call:
## lm(formula = Mixed_N ~ Discharge.C, data = EC_Flow.Nutrients_Wide)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.38192 -0.75099 -0.09027  0.47413  2.81925
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  1.7519765  0.1382810  12.670   <2e-16 ***
```

```

## Discharge.C -0.0006604  0.0011685  -0.565    0.574
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.002 on 62 degrees of freedom
## (5566 observations deleted due to missingness)
## Multiple R-squared:  0.005126, Adjusted R-squared:  -0.01092
## F-statistic: 0.3195 on 1 and 62 DF, p-value: 0.574
## adj R squared says 1 % of variance is explained by depth

cor.test(EC_Flow.Nutrients_Wide$Mixed_N, EC_Flow.Nutrients_Wide$Discharge.C)

##
## Pearson's product-moment correlation
##
## data:  EC_Flow.Nutrients_Wide$Mixed_N and EC_Flow.Nutrients_Wide$Discharge.C
## t = -0.56521, df = 62, p-value = 0.574
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
## -0.3119172  0.1773328
## sample estimates:
## cor
## -0.07159742

NitrogenGorman.regression <-lm(data = EC_Flow.Nutrients_Wide, Mixed_N ~ Discharge.G)
summary(NitrogenGorman.regression)

##
## Call:
## lm(formula = Mixed_N ~ Discharge.G, data = EC_Flow.Nutrients_Wide)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.3766 -0.7452 -0.1024  0.4735  2.8263
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  1.7481842  0.1405379  12.439  <2e-16 ***
## Discharge.G -0.0001626  0.0003524  -0.461    0.646
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.003 on 62 degrees of freedom
## (5566 observations deleted due to missingness)
## Multiple R-squared:  0.003423, Adjusted R-squared:  -0.01265
## F-statistic: 0.2129 on 1 and 62 DF, p-value: 0.6461
## adj R squared says 1 % of variance is explained by depth

cor.test(EC_Flow.Nutrients_Wide$Mixed_N, EC_Flow.Nutrients_Wide$Discharge.G)

##
## Pearson's product-moment correlation
##
## data:  EC_Flow.Nutrients_Wide$Mixed_N and EC_Flow.Nutrients_Wide$Discharge.G

```

```

## t = -0.46144, df = 62, p-value = 0.6461
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
## -0.2999978 0.1900395
## sample estimates:
## cor
## -0.05850268

###Phosphorus
PhosphorusClub.regression <-lm(data = EC_Flow.Nutrients_Wide, TP ~ Discharge.C)
summary(PhosphorusClub.regression)

##
## Call:
## lm(formula = TP ~ Discharge.C, data = EC_Flow.Nutrients_Wide)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.30568 -0.07026 -0.04760  0.02954  0.59927
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.1450530  0.0106823  13.579  < 2e-16 ***
## Discharge.C 0.0005484  0.0001417   3.872 0.000158 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1286 on 157 degrees of freedom
## (5471 observations deleted due to missingness)
## Multiple R-squared:  0.08715,    Adjusted R-squared:  0.08134
## F-statistic: 14.99 on 1 and 157 DF,  p-value: 0.0001583

## adj R squared says 8 % of variance is explained by depth
## significant relationship p < 0.001

cor.test(EC_Flow.Nutrients_Wide$TP, EC_Flow.Nutrients_Wide$Discharge.C)

##
## Pearson's product-moment correlation
##
## data:  EC_Flow.Nutrients_Wide$TP and EC_Flow.Nutrients_Wide$Discharge.C
## t = 3.8716, df = 157, p-value = 0.0001583
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.1462907 0.4310563
## sample estimates:
## cor
## 0.2952159

PhosphorusGorman.regression <-lm(data = EC_Flow.Nutrients_Wide, TP ~ Discharge.G)
summary(PhosphorusGorman.regression)

##
## Call:
## lm(formula = TP ~ Discharge.G, data = EC_Flow.Nutrients_Wide)
##

```

```
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.24129 -0.07216 -0.04873  0.02568  0.63304
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.446e-01  1.117e-02  12.939  < 2e-16 ***
## Discharge.G 1.863e-04  4.457e-05   4.181 4.75e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1343 on 162 degrees of freedom
## (5466 observations deleted due to missingness)
## Multiple R-squared:  0.09738,    Adjusted R-squared:  0.0918
## F-statistic: 17.48 on 1 and 162 DF,  p-value: 4.749e-05

## adj R squared says 9 % of variance is explained by depth
## significant relationship p < 0.001

cor.test(EC_Flow.Nutrients_Wide$TP, EC_Flow.Nutrients_Wide$Discharge.G)

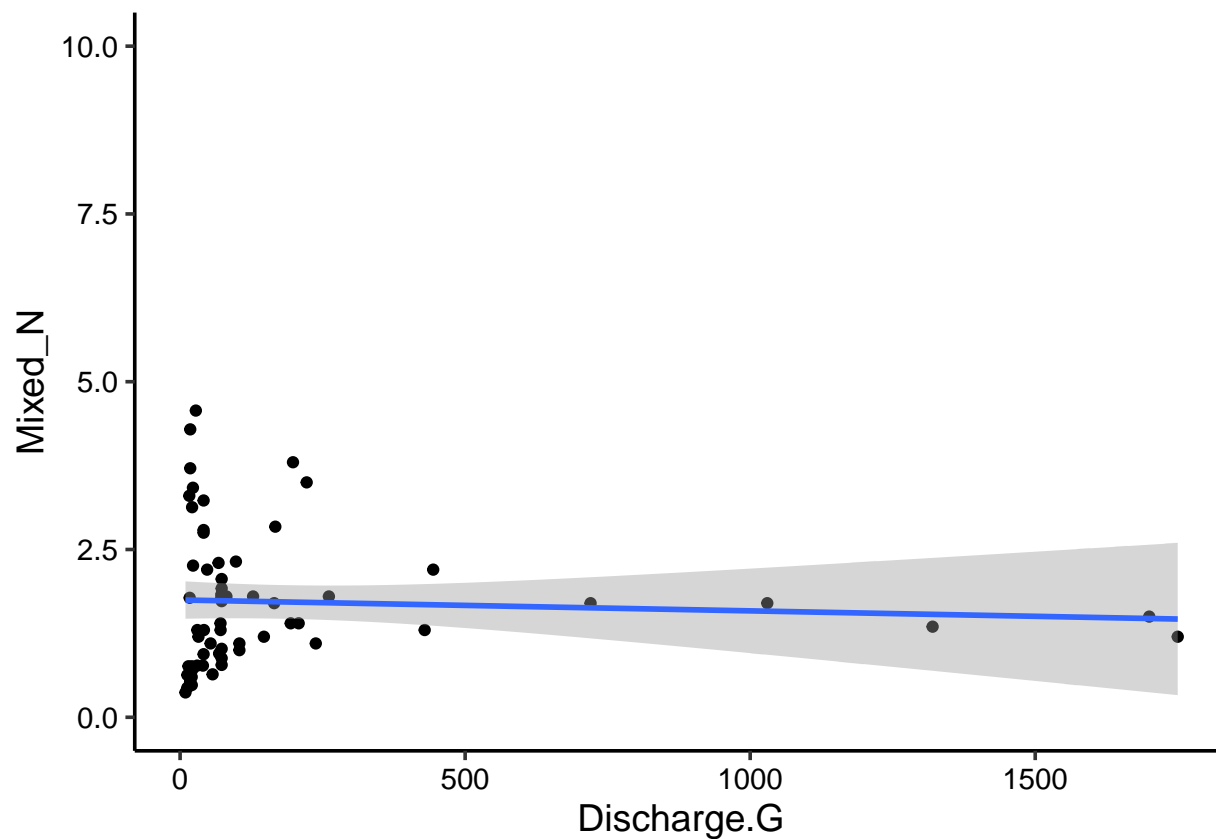
##
## Pearson's product-moment correlation
##
## data:  EC_Flow.Nutrients_Wide$TP and EC_Flow.Nutrients_Wide$Discharge.G
## t = 4.1805, df = 162, p-value = 4.749e-05
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.1667759 0.4440642
## sample estimates:
##          cor
## 0.3120502
```

Plot Regression

Nitrogen

```
Nitrogen_Discharge <-
  ggplot(EC_Flow.Nutrients_Wide, aes (x=Discharge.G, y = Mixed_N))+
  geom_point()+
  geom_smooth(method=lm)+
  ylim(0,10)
print(Nitrogen_Discharge)

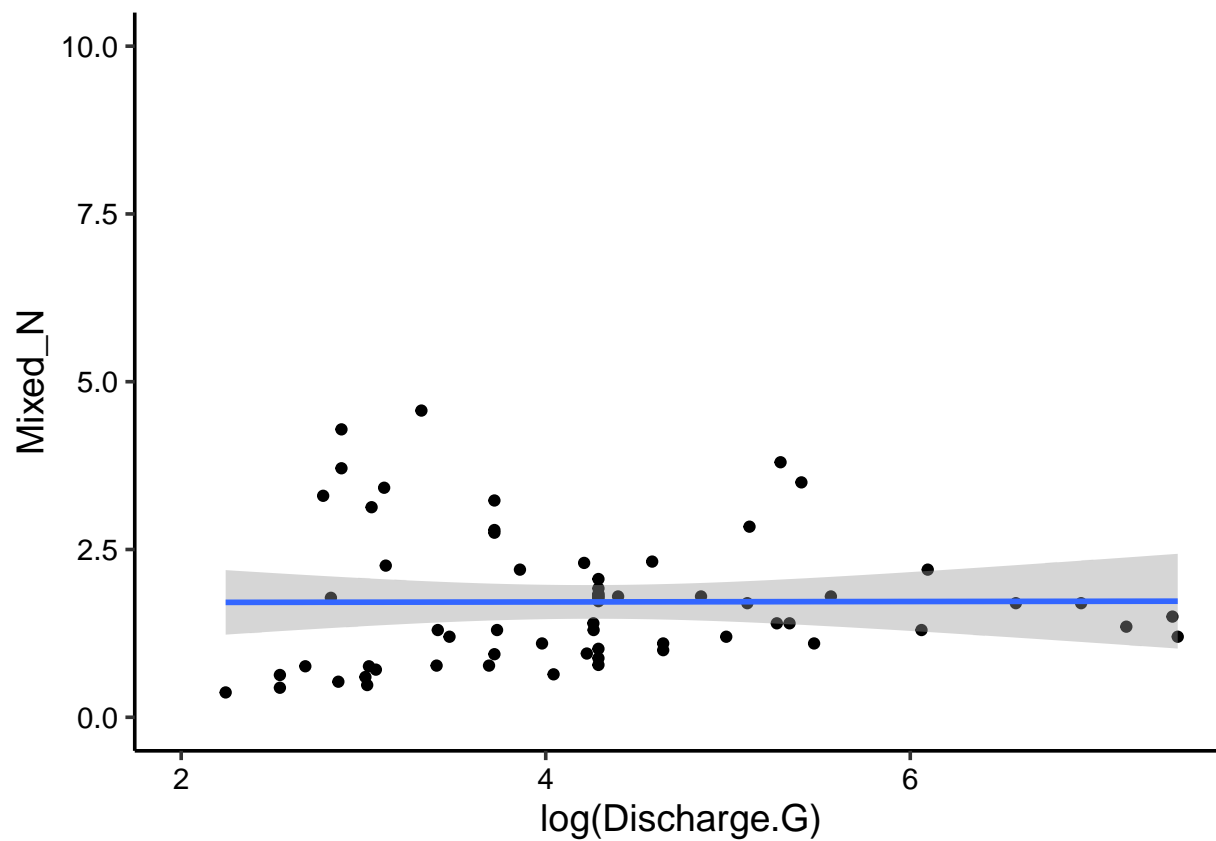
## Warning: Removed 5566 rows containing non-finite values (stat_smooth).
## Warning: Removed 5566 rows containing missing values (geom_point).
```



```
#log discharge
Nitrogen_logDischarge <-
  ggplot(EC_Flow.Nutrients_Wide, aes (x=log(Discharge.G), y = Mixed_N))+
  geom_point()+
  geom_smooth(method=lm)+
  ylim(0,10)
print(Nitrogen_logDischarge)
```

```
## Warning: Removed 5566 rows containing non-finite values (stat_smooth).
```

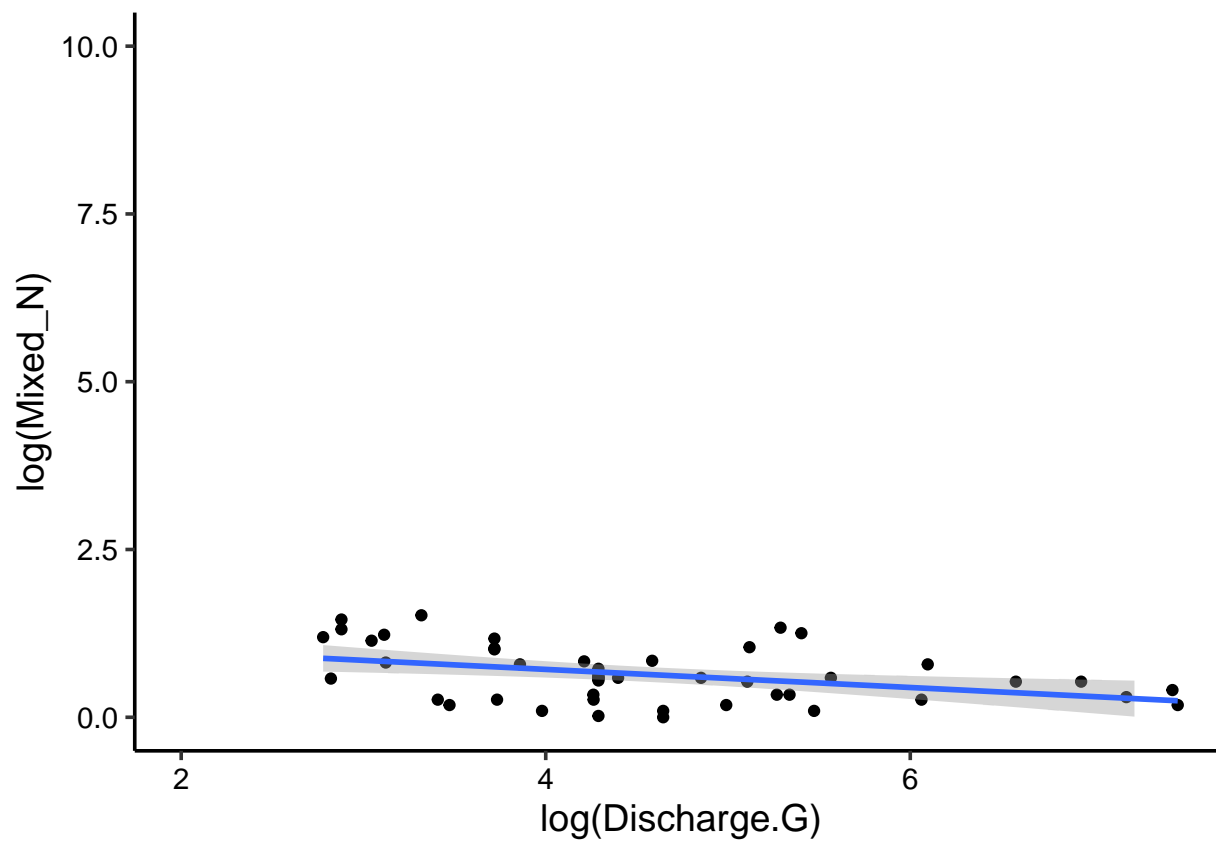
```
## Warning: Removed 5566 rows containing missing values (geom_point).
```



```
#log both
logNitrogen_logDischarge <-
  ggplot(EC_Flow.Nutrients_Wide, aes (x=log(Discharge.G), y = log(Mixed_N)))+
  geom_point()+
  geom_smooth(method=lm)+
  ylim(0,10)
print(logNitrogen_logDischarge)
```

```
## Warning: Removed 5582 rows containing non-finite values (stat_smooth).
```

```
## Warning: Removed 5582 rows containing missing values (geom_point).
```

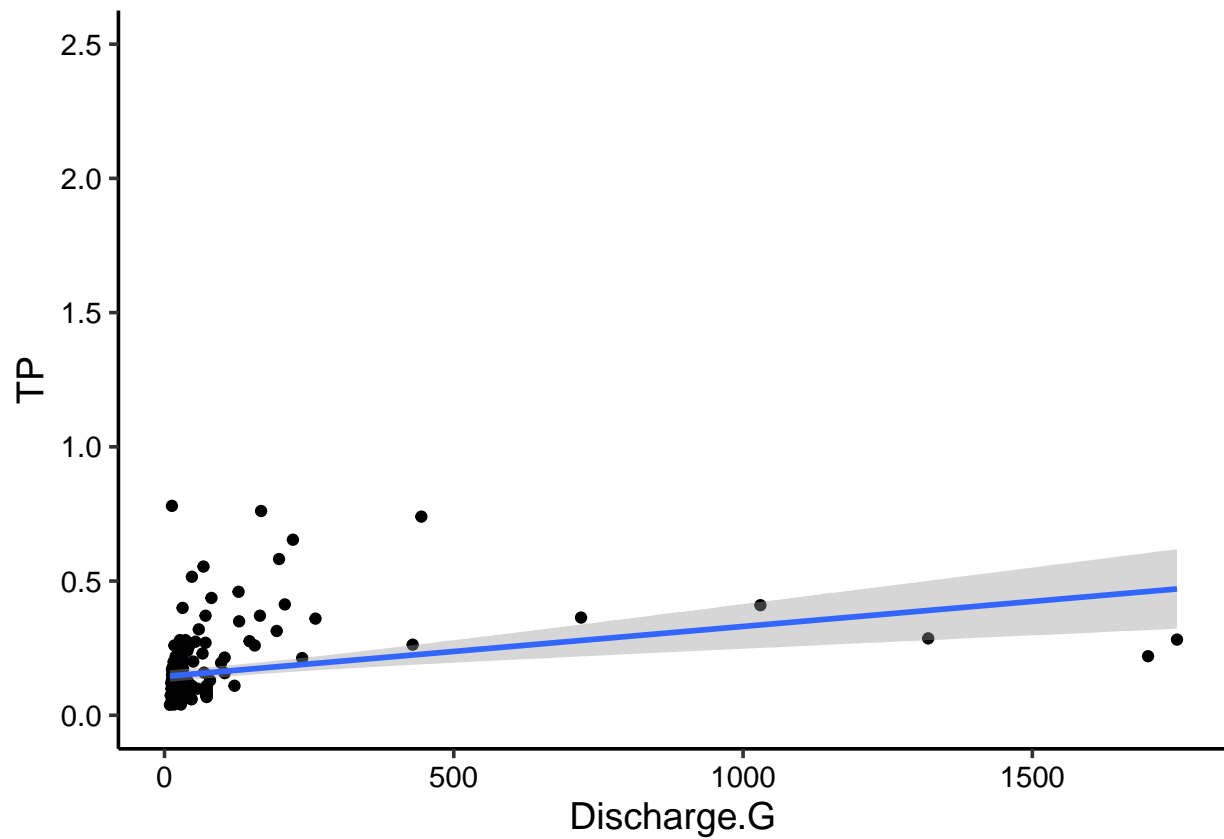



Phosphorus

```
Phosphorus_Discharge <-  
  ggplot(EC_Flow.Nutrients_Wide, aes (x=Discharge.G, y = TP))+  
  geom_point()+  
  geom_smooth(method=lm)+  
  ylim(0,2.5)  
print(Phosphorus_Discharge)
```

```
## Warning: Removed 5466 rows containing non-finite values (stat_smooth).
```

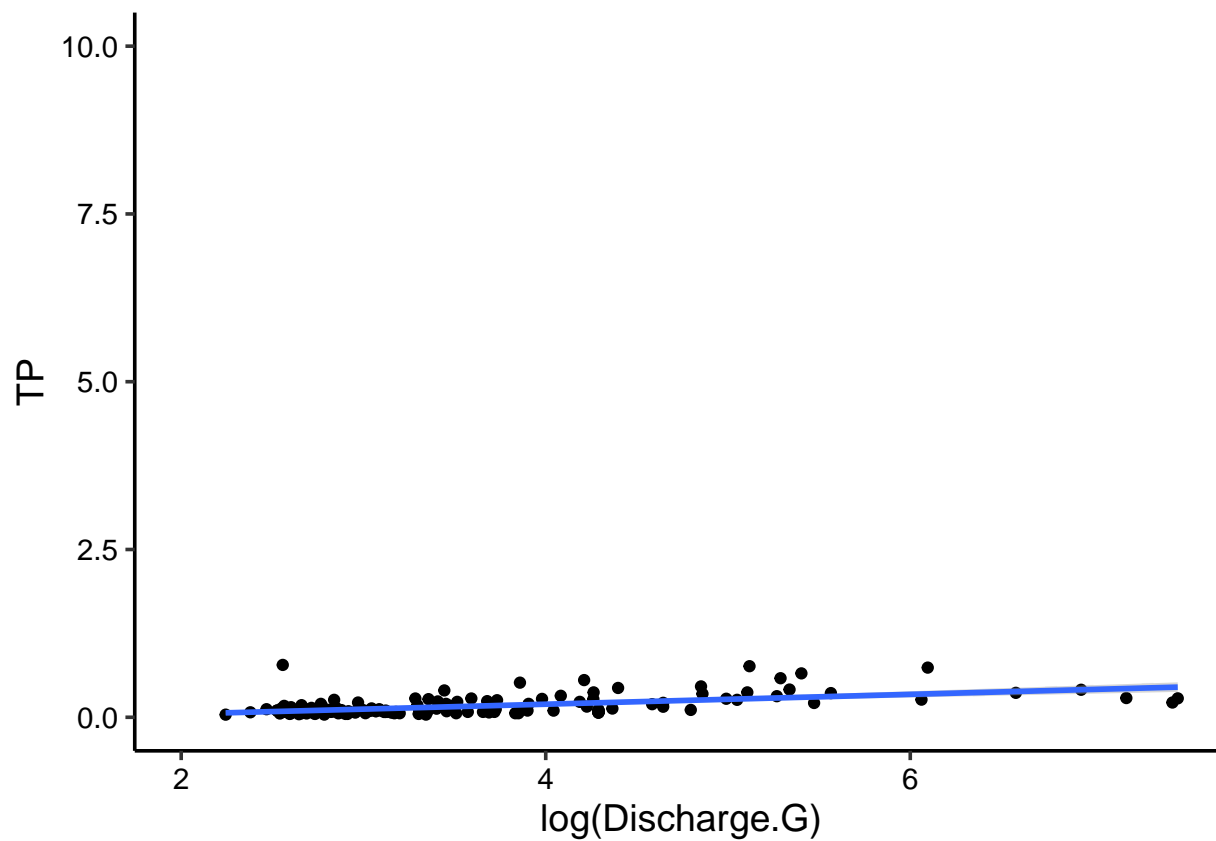
```
## Warning: Removed 5466 rows containing missing values (geom_point).
```



```
#log discharge
Phosphorus_logDischarge <-
  ggplot(EC_Flow.Nutrients_Wide, aes (x=log(Discharge.G), y = TP))+
  geom_point()+
  geom_smooth(method=lm)+
  ylim(0,10)
print(Phosphorus_logDischarge)
```

```
## Warning: Removed 5466 rows containing non-finite values (stat_smooth).
```

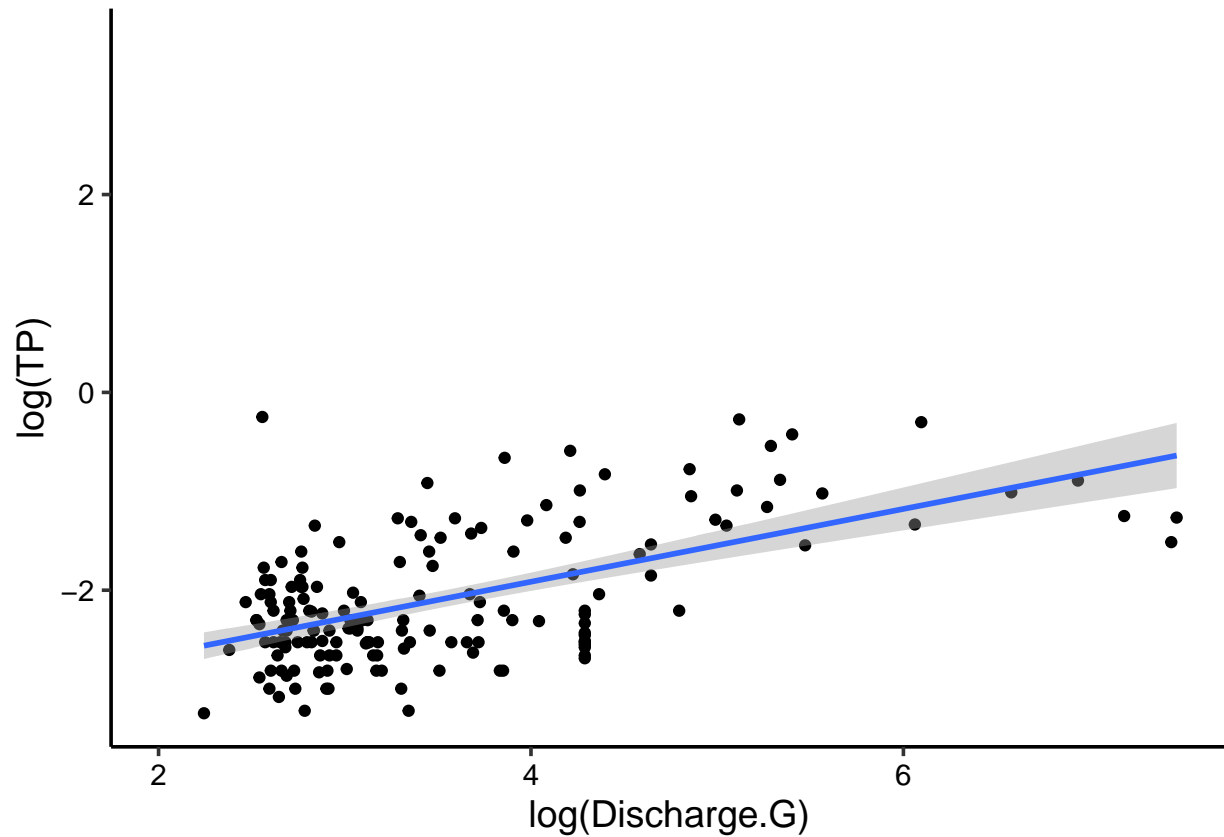
```
## Warning: Removed 5466 rows containing missing values (geom_point).
```



```
#log both
logPhosphorus_logDischarge <-
  ggplot(EC_Flow.Nutrients_Wide, aes (x=log(Discharge.G), y = log(TP)))+
  geom_point()+
  geom_smooth(method=lm)
print(logPhosphorus_logDischarge)
```

```
## Warning: Removed 5466 rows containing non-finite values (stat_smooth).
```

```
## Warning: Removed 5466 rows containing missing values (geom_point).
```

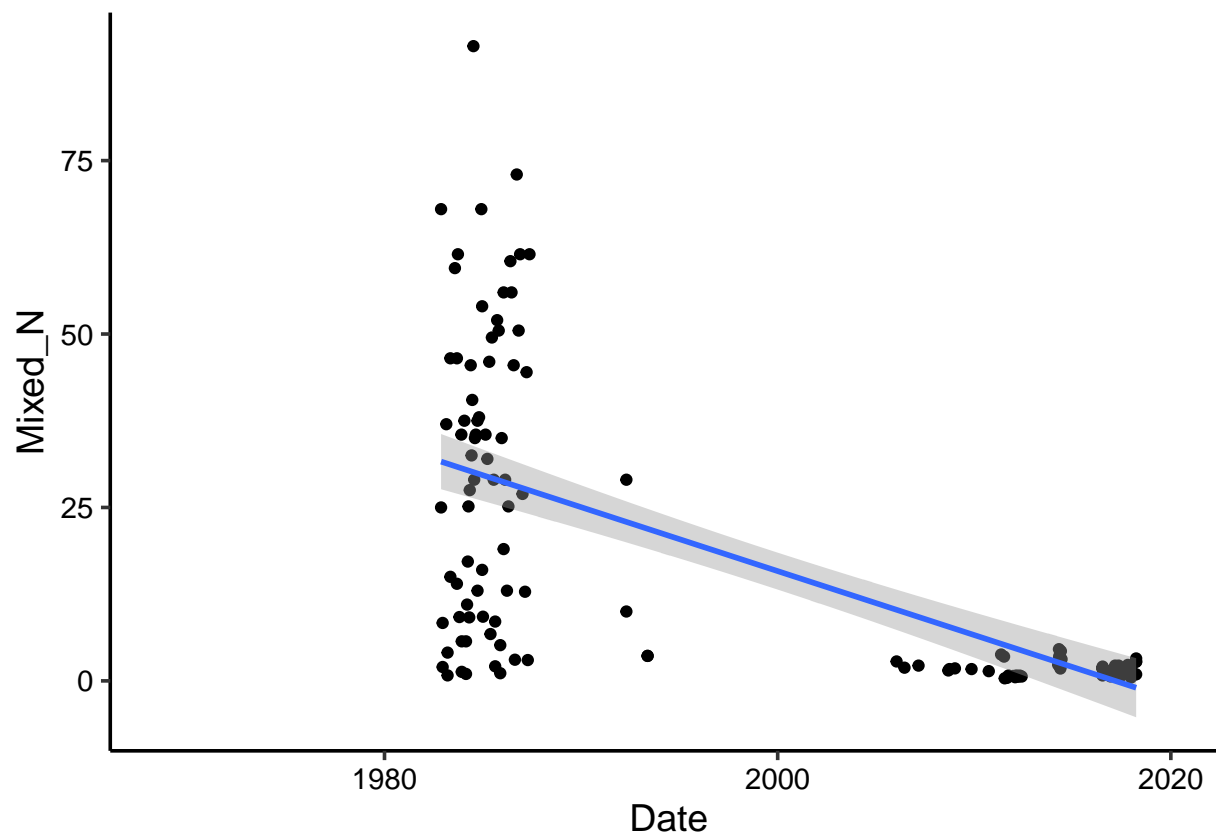


#Nutrients over Time

```
Nitrogen_Time <-  
  ggplot(EC_Flow.Nutrients_Wide, aes (x=Date, y = Mixed_N))+  
  geom_point()+  
  geom_smooth(method=lm)  
print(Nitrogen_Time)
```

Warning: Removed 5490 rows containing non-finite values (stat_smooth).

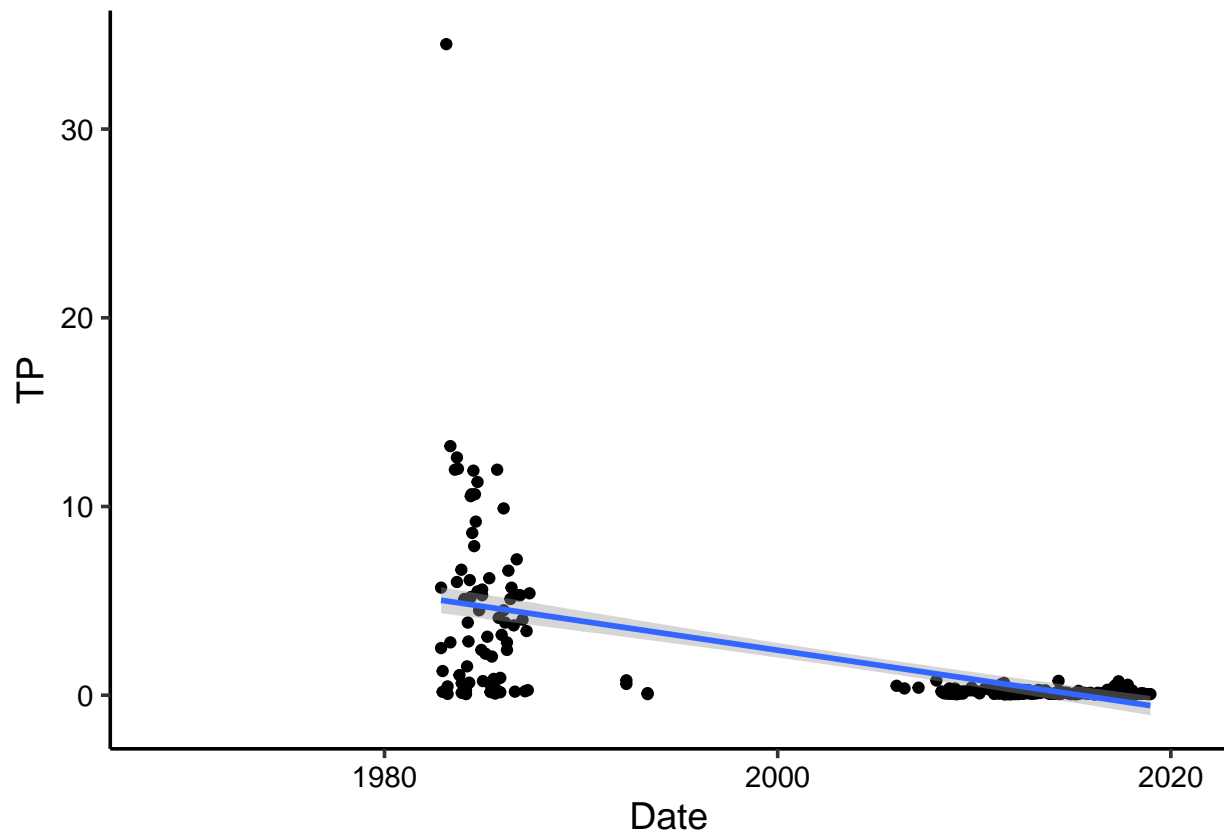
Warning: Removed 5490 rows containing missing values (geom_point).



```
Phosphorus_Time <-
  ggplot(EC_Flow.Nutrients_Wide, aes (x=Date, y = TP))+
  geom_point()+
  geom_smooth(method=lm)
print(Phosphorus_Time)
```

```
## Warning: Removed 5387 rows containing non-finite values (stat_smooth).
```

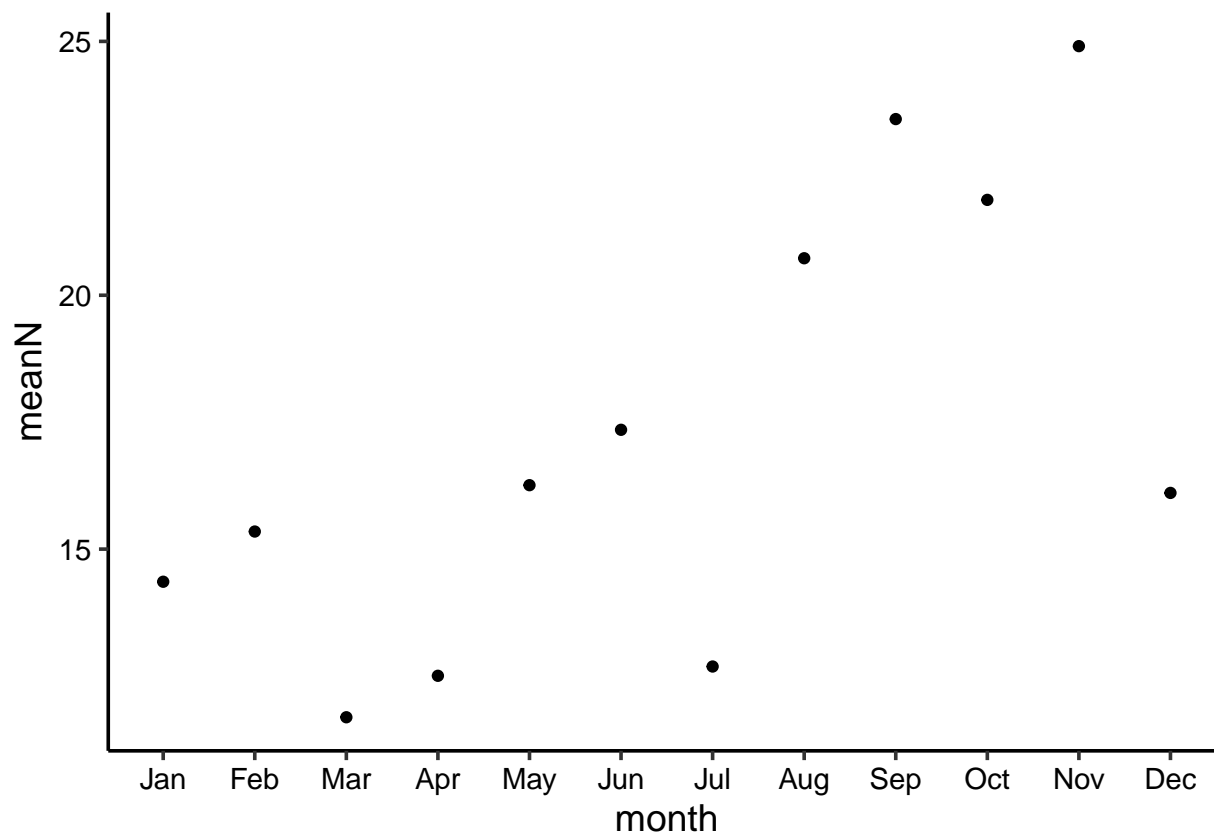
```
## Warning: Removed 5387 rows containing missing values (geom_point).
```



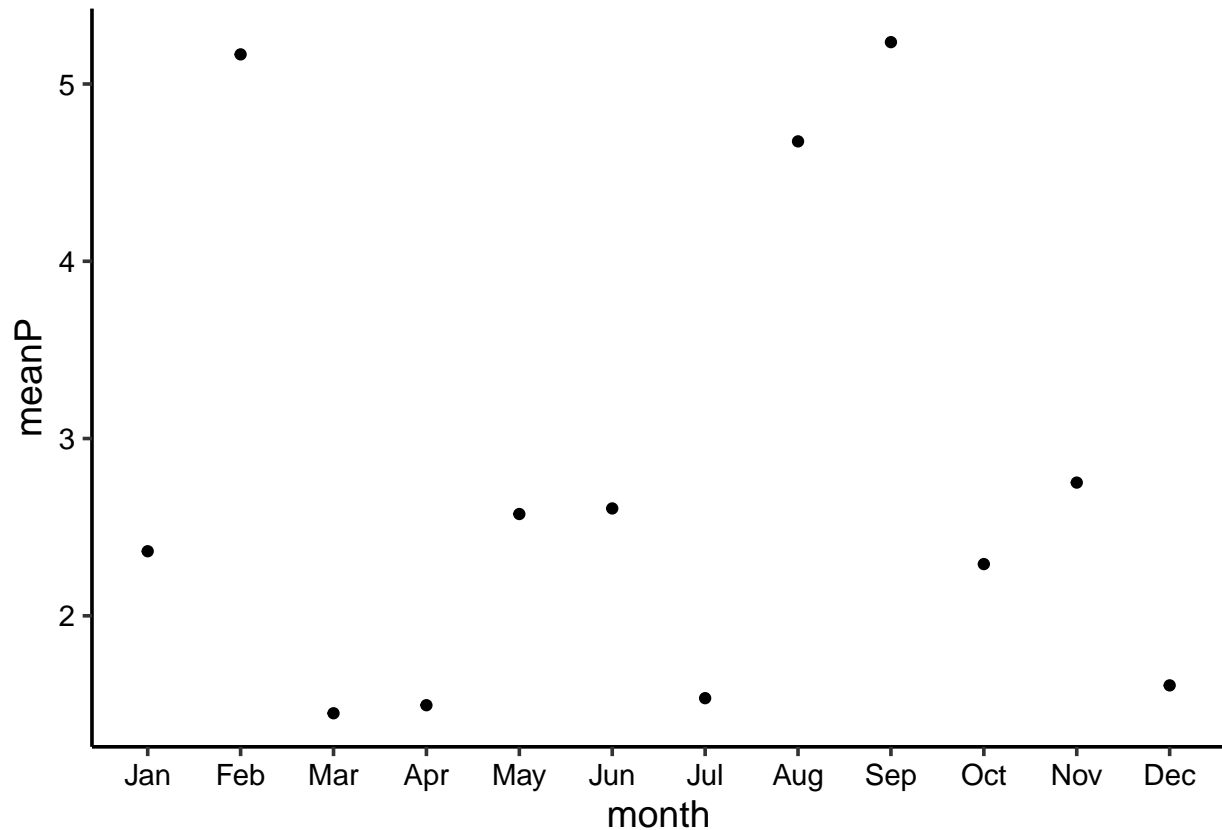
Nutrient Summaries

```
EC_Nutrient_MonthSummary <- EC_Flow.Nutrients_Wide %>%
  mutate(month = month(Date, label = TRUE)) %>%
  group_by(month) %>%
  filter(!is.na(Mixed_N), !is.na(TP)) %>%
  summarise(meanN = mean(Mixed_N),
            sdN = sd(Mixed_N),
            minN = min(Mixed_N),
            maxN = max(Mixed_N),
            meanP = mean(TP),
            sdP = sd(TP),
            minP = min(TP),
            maxP = max(TP))

ggplot(EC_Nutrient_MonthSummary)+
  geom_point(aes(y = meanN, x=month))
```

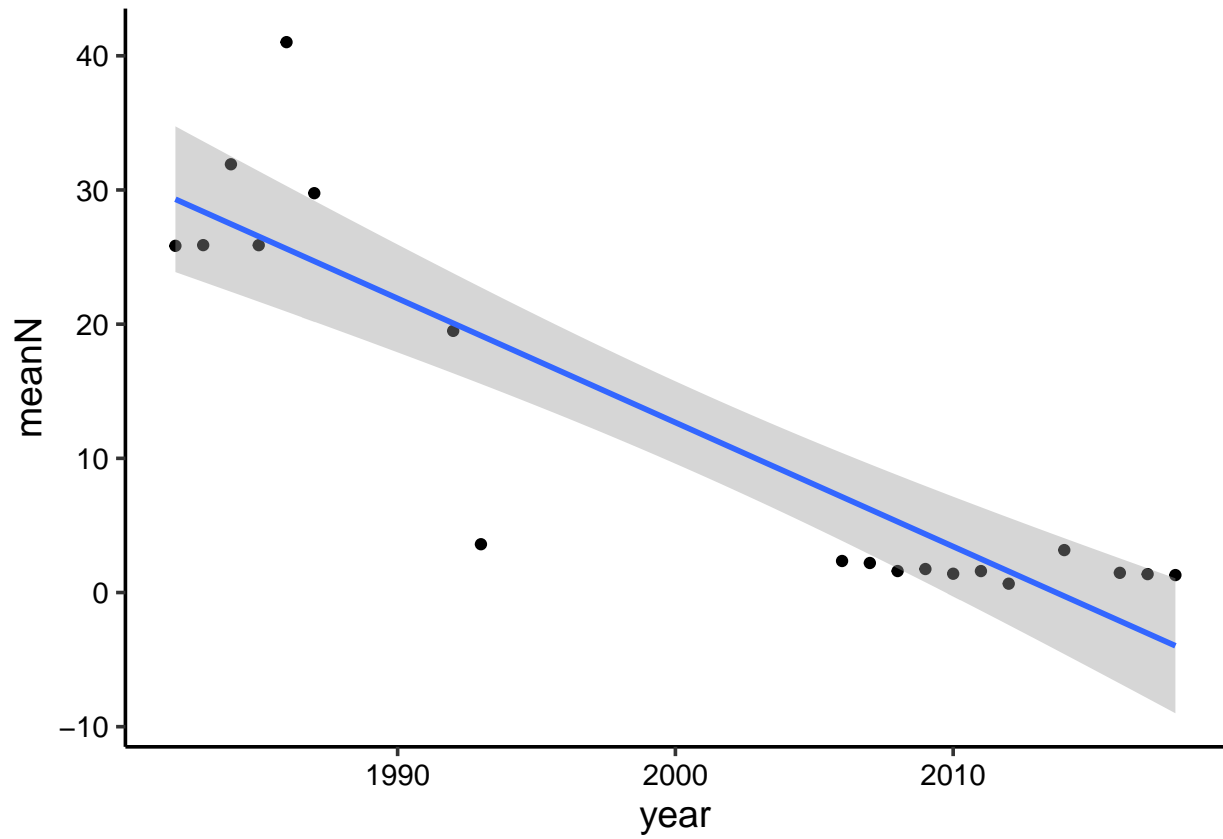


```
ggplot(EC_Nutrient_MonthSummary)+  
  geom_point(aes(y = meanP, x=month))
```

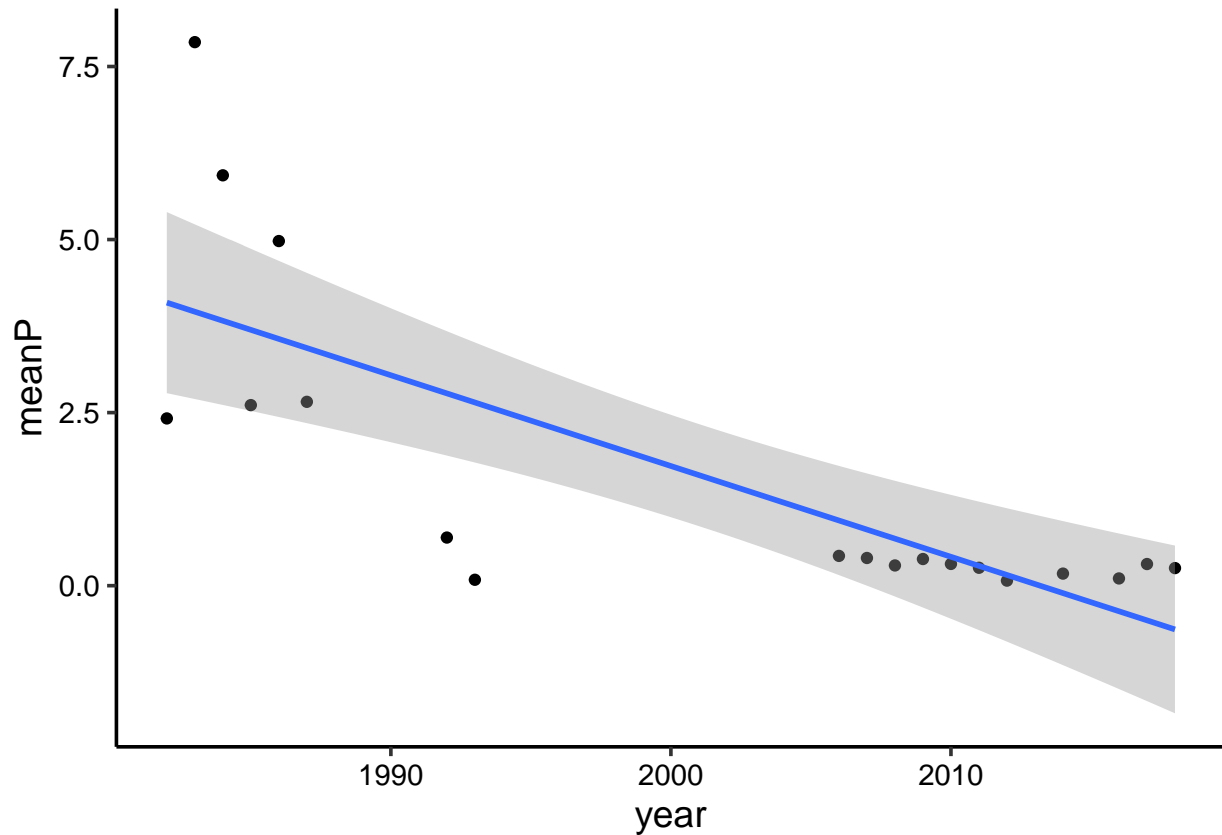


```
EC_Nutrient_YearSummary <- EC_Flow.Nutrients_Wide %>%
  mutate(year = year(Date)) %>%
  group_by(year) %>%
  filter(!is.na(Mixed_N), !is.na(TP)) %>%
  summarise(meanN = mean(Mixed_N),
            sdN = sd(Mixed_N),
            minN = min(Mixed_N),
            maxN = max(Mixed_N),
            meanP = mean(TP),
            sdP = sd(TP),
            minP = min(TP),
            maxP = max(TP))

ggplot(EC_Nutrient_YearSummary, aes(y = meanN, x=year))+
  geom_point()+
  geom_smooth(method=lm)
```

```
ggplot(EC_Nutrient_YearSummary, aes(y = meanP, x=year))+  
  geom_point()+  
  geom_smooth(method=lm)
```



```
#Test subset post 2000
```

```
Post2000Nutrients <- subset(EC_Flow.Nutrients_Wide, Date > "2000-01-01" & Date < "2020-12-31")
```

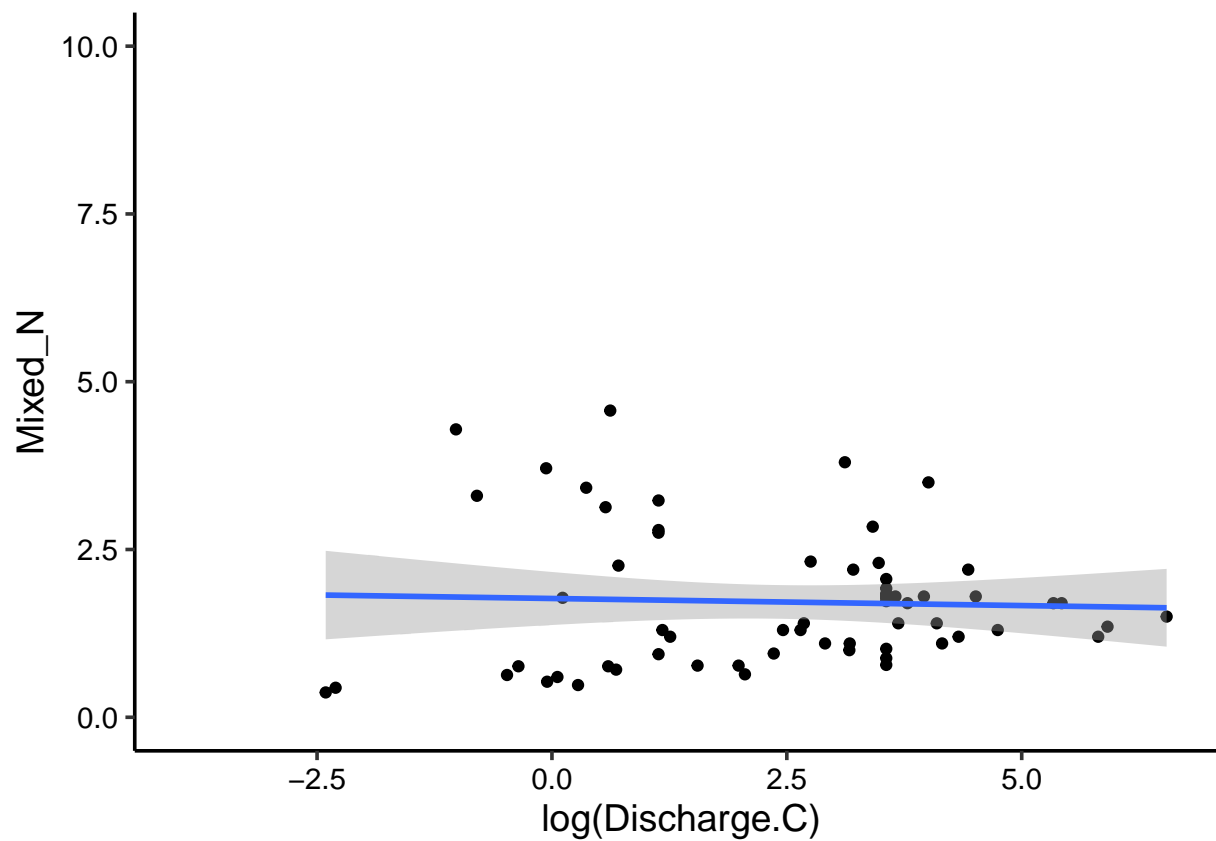
Plot

Nitrogen

```
Nitrogen_Discharge2000 <-  
  ggplot(Post2000Nutrients , aes (x=log(Discharge.C), y = Mixed_N))+  
  geom_point()+  
  geom_smooth(method=lm)+  
  ylim(0,10)  
print(Nitrogen_Discharge2000)
```

```
## Warning: Removed 4835 rows containing non-finite values (stat_smooth).
```

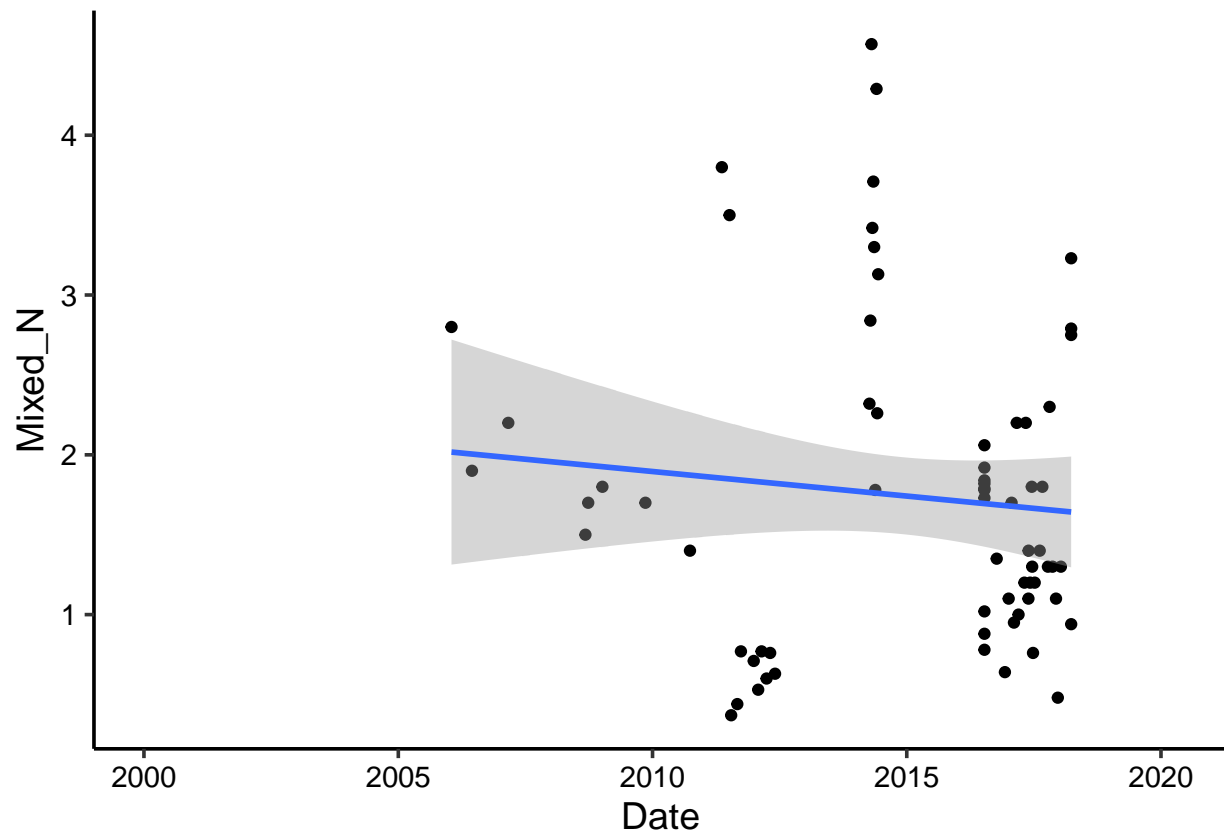
```
## Warning: Removed 4835 rows containing missing values (geom_point).
```



```
Nitrogen_Time2000 <-  
  ggplot(Post2000Nutrients, aes (x=Date, y = Mixed_N))+  
  geom_point()+  
  geom_smooth(method=lm)  
print(Nitrogen_Time2000)
```

```
## Warning: Removed 4832 rows containing non-finite values (stat_smooth).
```

```
## Warning: Removed 4832 rows containing missing values (geom_point).
```

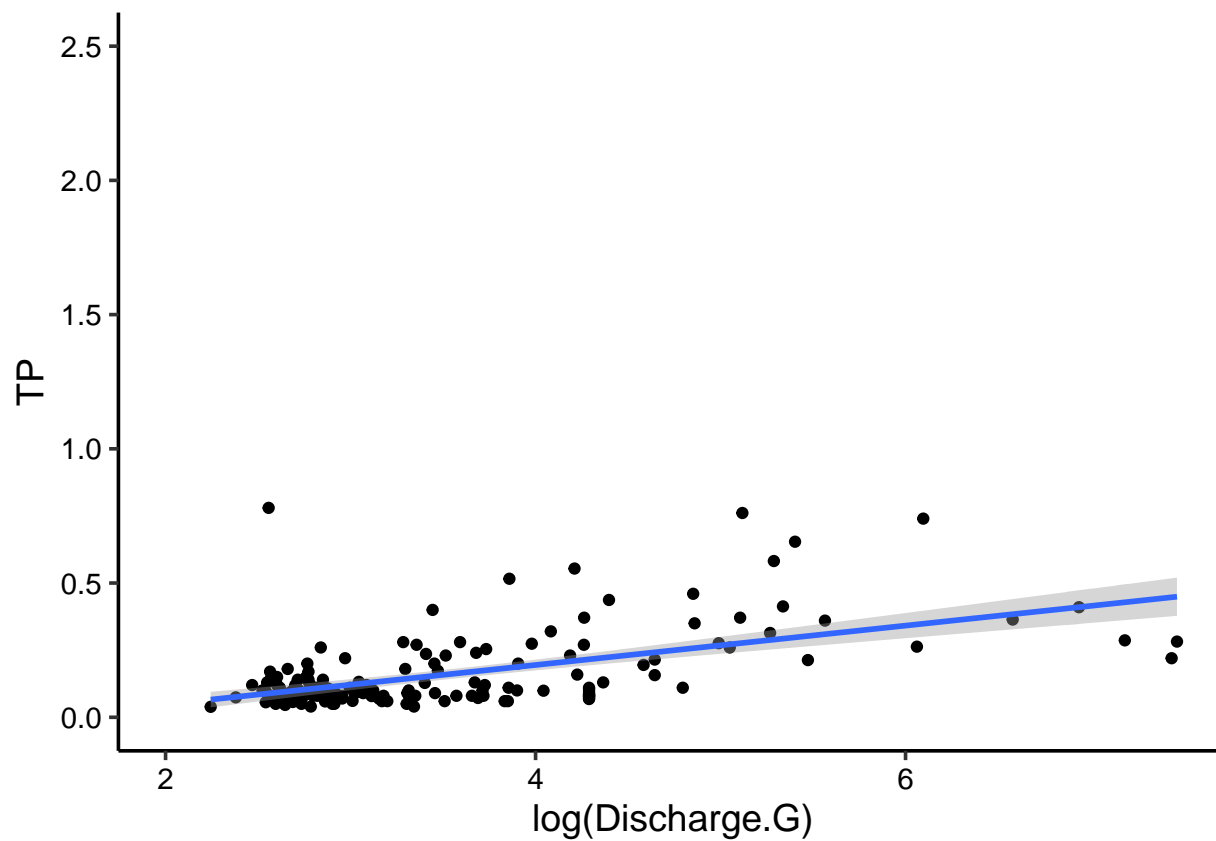


Phosphorus

```
Phosphorus_Discharge2000 <-
  ggplot(Post2000Nutrients, aes (x=log(Discharge.G), y = TP))+
  geom_point()+
  geom_smooth(method=lm)+
  ylim(0,2.5)
print(Phosphorus_Discharge2000)
```

Warning: Removed 4735 rows containing non-finite values (stat_smooth).

Warning: Removed 4735 rows containing missing values (geom_point).



```
Phosphorus_Time2000 <-  
  ggplot(Post2000Nutrients, aes (x=Date, y = TP))+  
  geom_point()+  
  geom_smooth(method=lm)  
print(Phosphorus_Time2000)
```

```
## Warning: Removed 4732 rows containing non-finite values (stat_smooth).
```

```
## Warning: Removed 4732 rows containing missing values (geom_point).
```

