Fish Abundance Analysis

Marc Los Huertos

2025-04-14

Read data into R

```
fish.csv = "/home/mwl04747/RTricks/13_coastal_fish/coastal_data.csv"
fish <- read.csv(fish.csv)

str(fish)
```

```
## 'data.frame':    200 obs. of  5 variables:
##  $ fish_abundance  : num  99.5 88.3 114.6 106 107.3 ...
##  $ water_temp      : num  20.9 21.5 25.1 22.1 22.3 ...
##  $ salinity        : num  38.3 37 34.6 35.8 34.4 ...
##  $ pH              : num  8.09 7.87 7.97 8.09 8.23 ...
##  $ dissolved_oxygen: num  7.07 5.97 5.97 4.48 6.79 ...
```

Multiple linear regression

```
# Fit a multiple linear regression model
fish.lm <- lm(fish_abundance ~ water_temp + salinity + pH + dissolved_oxygen, data = fish)

summary(fish.lm)
```

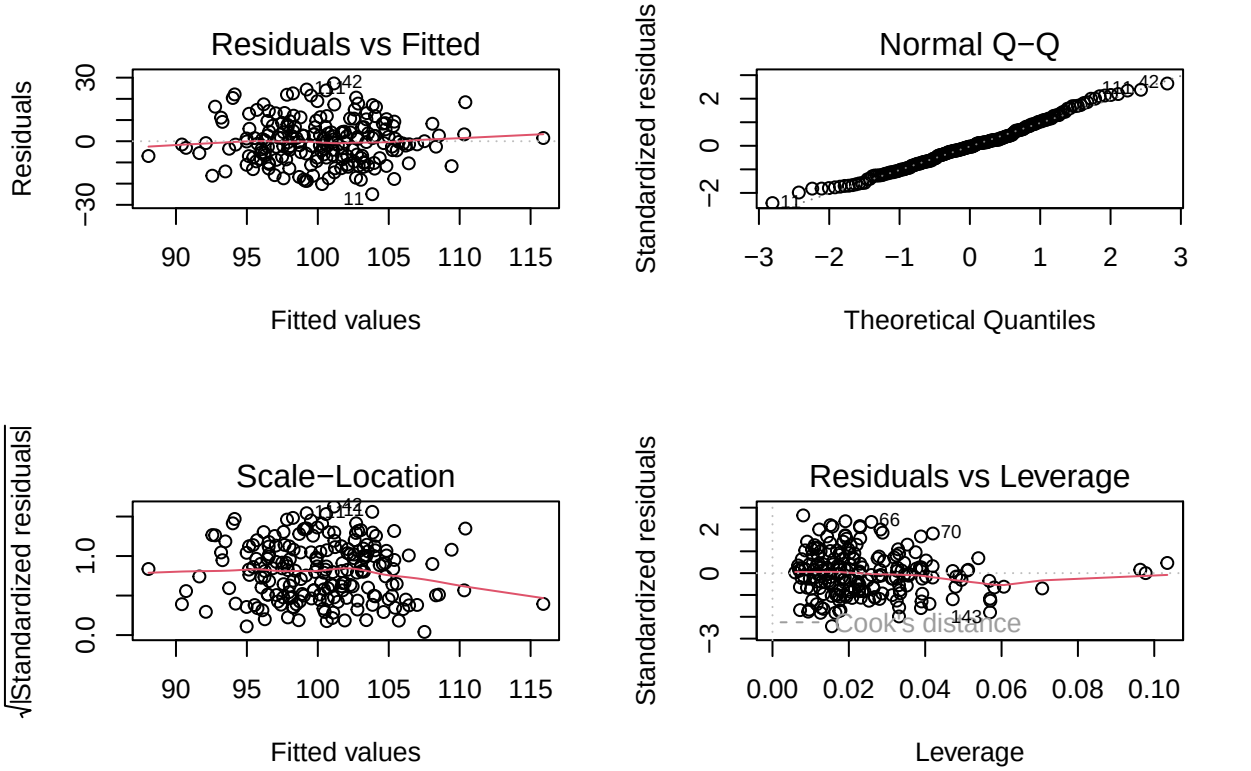
```
##
## Call:
## lm(formula = fish_abundance ~ water_temp + salinity + pH + dissolved_oxygen,
##     data = fish)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -24.9888  -7.1036  -0.3632   6.6081  27.2780
##
## Coefficients:
##                  Estimate Std. Error t value Pr(>|t|)
## (Intercept)       44.5298    37.5917   1.185  0.23763
## water_temp         1.7130     0.3900   4.393 1.84e-05 ***
## salinity          -0.7577     0.4940  -1.534  0.12671
## pH                 3.8156     3.8147   1.000  0.31843
## dissolved_oxygen   2.2847     0.7201   3.173  0.00175 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## Residual standard error: 10.36 on 195 degrees of freedom
## Multiple R-squared:  0.1383, Adjusted R-squared:  0.1206
## F-statistic: 7.823 on 4 and 195 DF,  p-value: 7.225e-06
```

```
par(mfrow=c(2,2))
# Diagnostic plots
```

Diagnostics – Checking Assumptions

```
# Diagnostic plots
par(mfrow = c(2, 2))
plot(fish.lm)
```




```
# Check multicollinearity (VIF)
library(car)
```

```
## Loading required package: carData
```

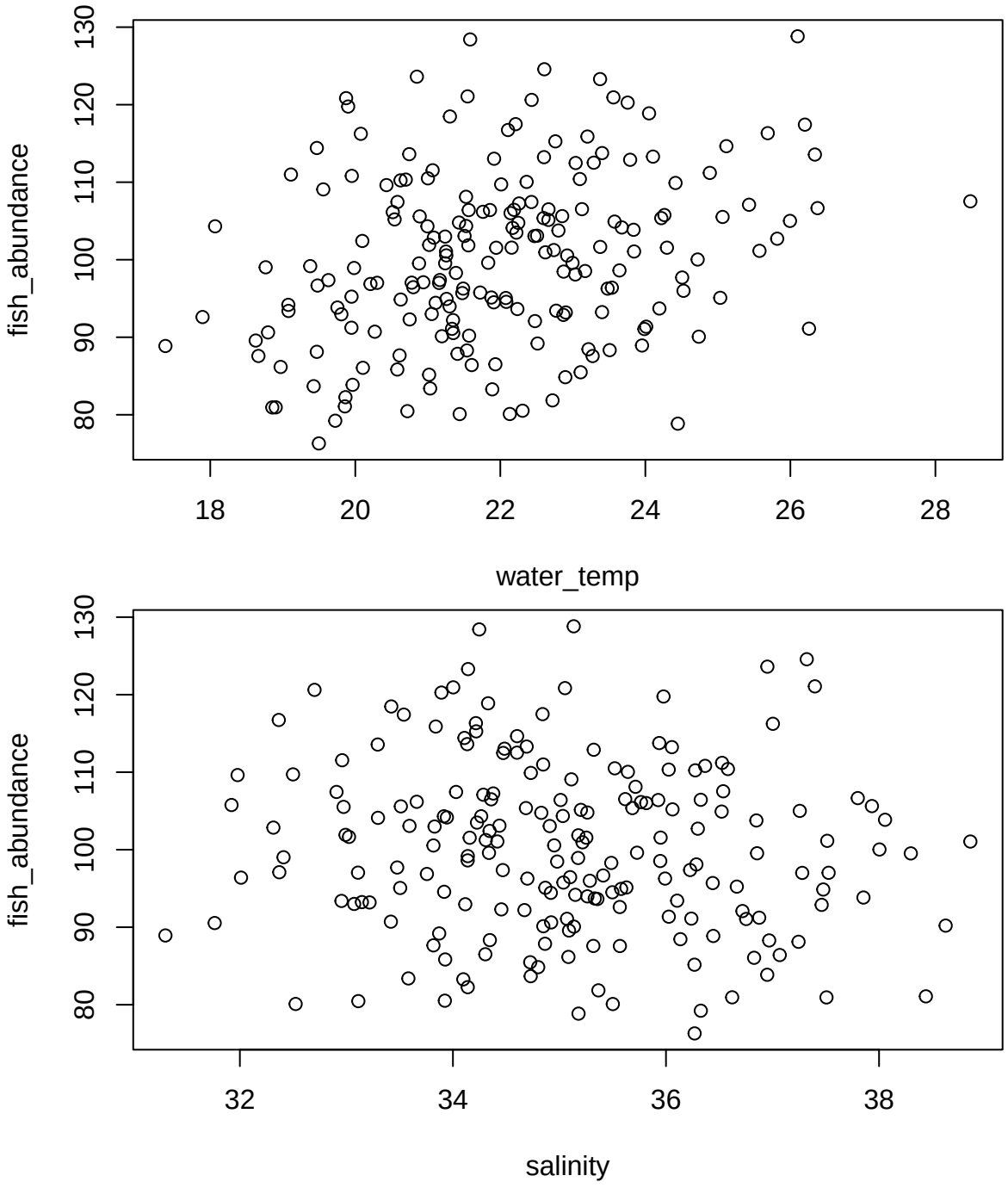
```
vif(fish.lm)
```

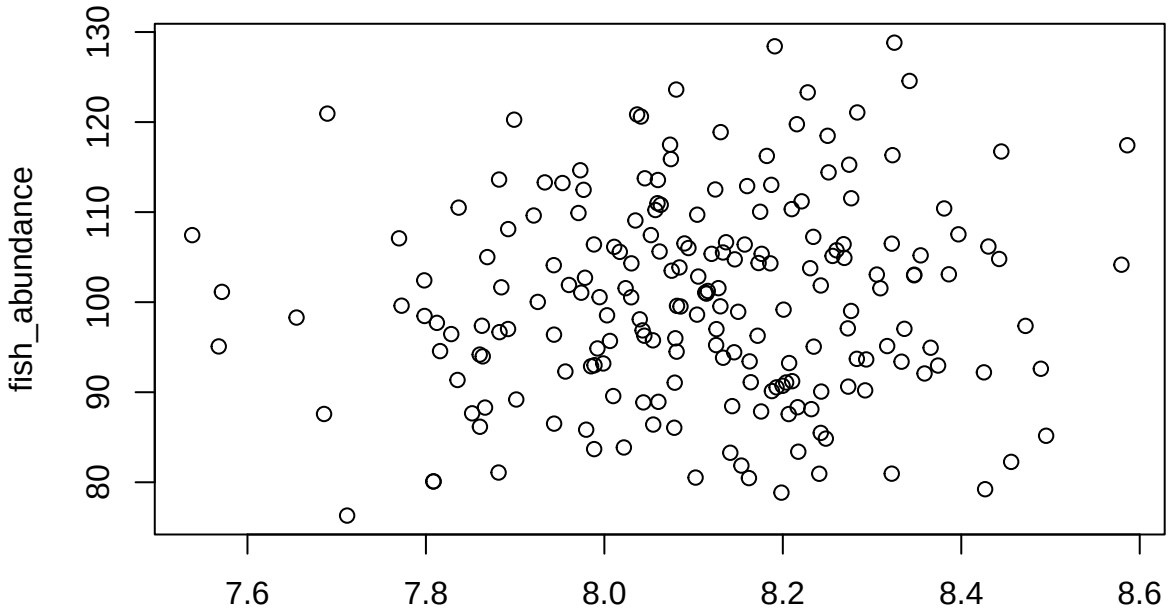
```
##       water_temp         salinity               pH dissolved_oxygen
##         1.003653         1.010096         1.004884         1.009680
```

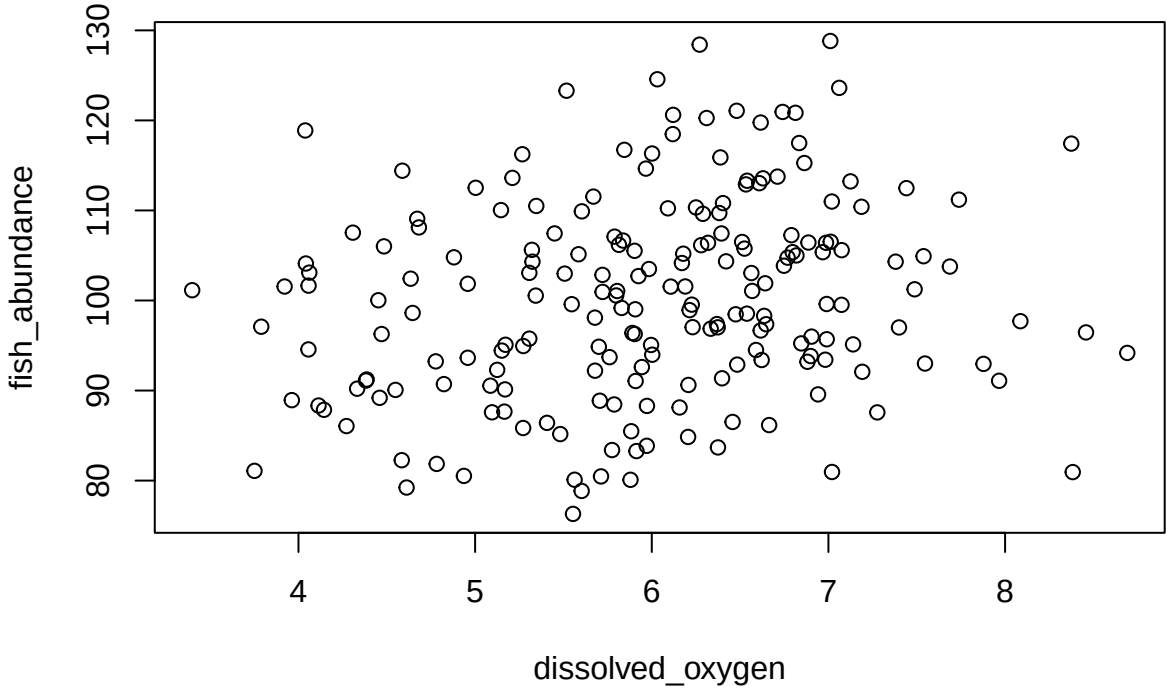
Plot Water temp and DO

```
par(mfrow= c(1,1))
```

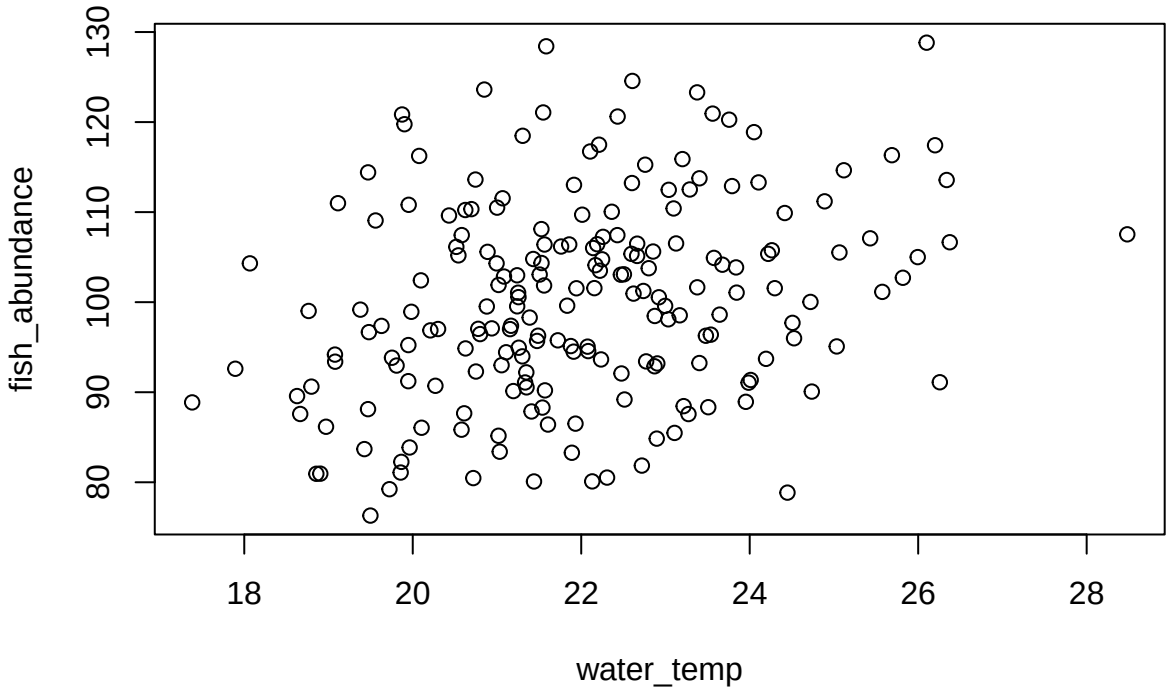
```r
plot(fish_abundance ~ water_temp + salinity + pH + dissolved_oxygen, data = fish)
```



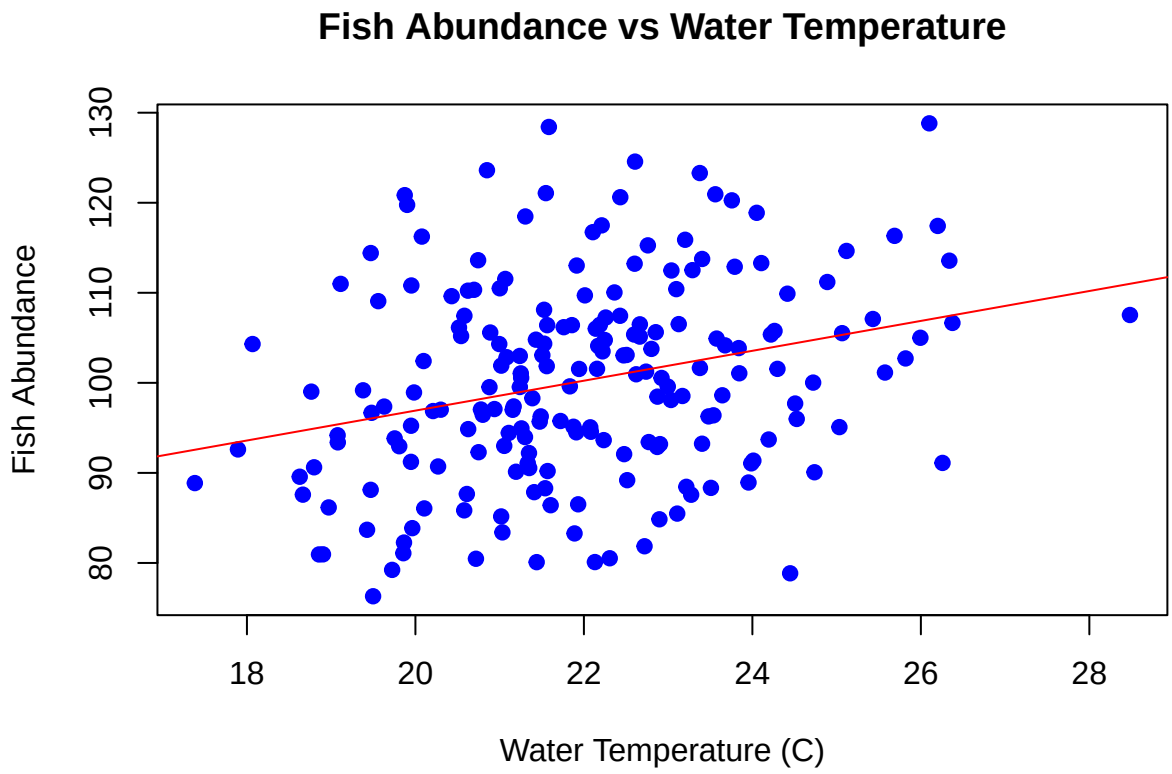





```
plot(fish_abundance ~ water_temp, data = fish)
```



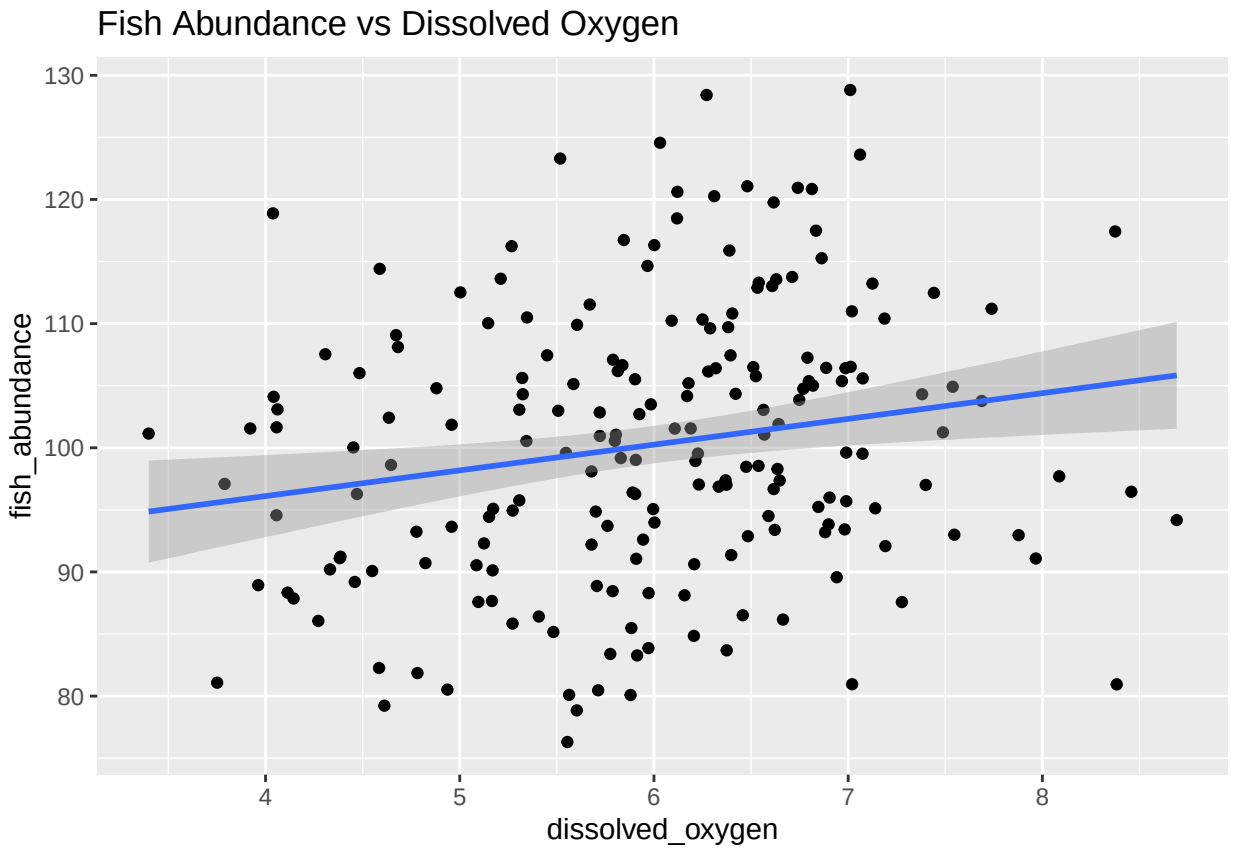

```
plot(fish_abundance ~ water_temp, data = fish,
     xlab = "Water Temperature (C)",
     ylab = "Fish Abundance",
     main = "Fish Abundance vs Water Temperature",
     pch = 19,
     col = "blue")
# Add regression line
temp.lm <- lm(fish_abundance ~ water_temp, data = fish)

abline(temp.lm, col = "red")
```



Fancy Plots!

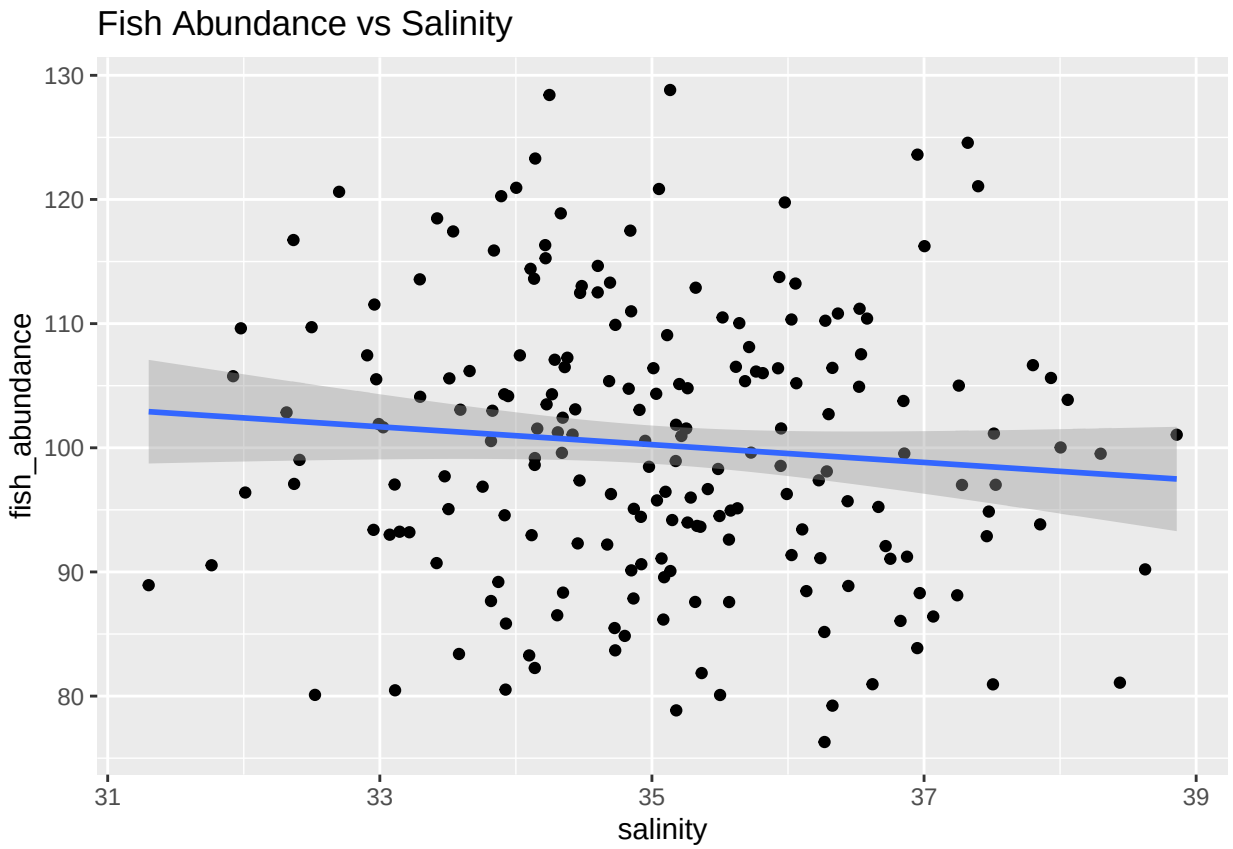
```
library(ggplot2)
# Visualizing relationships
ggplot(fish, aes(x = dissolved_oxygen, y = fish_abundance)) +
  geom_point() +
  geom_smooth(method = "lm") +
  labs(title = "Fish Abundance vs Dissolved Oxygen")
```

```
## `geom_smooth()` using formula = 'y ~ x'
```




```
ggplot(fish, aes(x = salinity, y = fish_abundance)) +
  geom_point() +
  geom_smooth(method = "lm") +
  labs(title = "Fish Abundance vs Salinity")
```

```
## `geom_smooth()` using formula = 'y ~ x'
```


Fish Abundance vs Salinity
fish_abundance
130
120
110
100
90
80
31
33
35
37
39
salinity