

Ly's Capstone Data Visualization

Directory and doc rules

```
knitr::opts_chunk$set(  
  echo = TRUE,      # Display code chunks  
  eval = TRUE,      # Evaluate code chunks  
  warning = FALSE,  # Hide warnings  
  message = FALSE,  # Hide messages  
  fig.width = 20,    # Set plot width in inches  
  fig.height = 9,    # Set plot height in inches  
  fig.align = "center" # Align plots to the center  
)
```

Load packages

```
library(tinytex)  
library(tidyr)  
library(tidyverse)  
library(vegan)
```

Load data

Weight = mg
Length, width, height = mm
p450, SOD = activity/ (mg/protein)
Condition factor= unitless

```
#this is important to know where you are working and where your outputs will land if you haven't set it  
getwd()
```

```
## [1] "/Users/cmantegna/Documents/WDFWmussels/code"
```

```
#data has all sites, coordinates, p450, sod, condition factor, economic factor data  
data<- read.csv("/Users/cmantegna/Documents/WDFWmussels/data/lycapstone.csv")
```

```
#review your data. Summary lets you see the values and data ranges. Str lets you see the data class of  
summary(data)
```

```
##      Region      latitude      longitude      site_name  
## Length:53      Min.   :47.26      Min.   : -123.0      Length:53  
## Class :character 1st Qu.:47.27      1st Qu.: -122.4      Class :character  
## Mode  :character Median :47.57      Median : -122.4      Mode  :character  
##                Mean   :47.49      Mean   : -122.5  
##                3rd Qu.:47.61      3rd Qu.: -122.4  
##                Max.   :48.21      Max.   : -122.3  
## site_number      sample      p450      SOD
```

```
## Min. : 3 Min. : 29.0 Min. : 0 Min. : -0.636
## 1st Qu.:17 1st Qu.: 70.0 1st Qu.: 1975307 1st Qu.: 1.531
## Median :29 Median :105.0 Median : 3125326 Median : 9.393
## Mean :32 Mean :100.8 Mean : 4054987 Mean : 16.042
## 3rd Qu.:46 3rd Qu.:122.0 3rd Qu.: 6222187 3rd Qu.: 20.437
## Max. :54 Max. :197.0 Max. :15260103 Max. :118.683
## weight_initial length width height
## Min. : 7.997 Min. :48.58 Min. :22.49 Min. :16.85
## 1st Qu.:12.463 1st Qu.:54.59 1st Qu.:24.60 1st Qu.:18.56
## Median :14.575 Median :57.47 Median :25.93 Median :19.39
## Mean :14.520 Mean :56.96 Mean :25.82 Mean :19.71
## 3rd Qu.:16.349 3rd Qu.:59.93 3rd Qu.:26.90 3rd Qu.:20.91
## Max. :20.214 Max. :61.58 Max. :30.94 Max. :23.97
## weight_final weight_change condition_factor avg_thickness
## Min. :3.316 Min. : 2.580 Min. :0.0430 Min. :0.5450
## 1st Qu.:4.191 1st Qu.: 8.410 1st Qu.:0.1541 1st Qu.:0.6950
## Median :4.552 Median :10.210 Median :0.1762 Median :0.8100
## Mean :4.594 Mean : 9.926 Mean :0.1735 Mean :0.7948
## 3rd Qu.:5.171 3rd Qu.:11.440 3rd Qu.:0.1944 3rd Qu.:0.8950
## Max. :6.059 Max. :14.270 Max. :0.2380 Max. :1.0400
```

```
str(data)
```

```
## 'data.frame': 53 obs. of 16 variables:
## $ Region : chr "Tacoma" "Tacoma" "Tacoma" "Tacoma" ...
## $ latitude : num 47.3 47.3 47.3 47.3 47.3 ...
## $ longitude : num -122 -122 -122 -122 -122 ...
## $ site_name : chr "Blair Waterway #2" "Blair Waterway #2" "Blair Waterway #2" "Blair Waterway #2" ...
## $ site_number : int 42 42 42 42 54 54 54 54 45 45 ...
## $ sample : int 141 142 143 144 194 195 196 197 37 38 ...
## $ p450 : int 3885225 2469960 3506452 4998033 2165448 9793068 7655768 9992165 6882929 6222929 ...
## $ SOD : num 26.424 12.854 2.637 1.269 0.0025 ...
## $ weight_initial : num 9.83 12.9 13.32 11.97 16.11 ...
## $ length : num 50.1 51.5 55.2 55.2 59.2 ...
## $ width : num 22.5 24.3 25.5 22.8 25.9 ...
## $ height : num 17.8 20.2 18.8 20 19.6 ...
## $ weight_final : num 3.35 4.96 4.46 4.19 5.49 ...
## $ weight_change : num 6.48 7.94 8.86 7.78 10.62 ...
## $ condition_factor: num 0.129 0.154 0.16 0.141 0.179 ...
## $ avg_thickness : num 0.745 0.97 0.9 0.815 0.84 0.8 0.82 0.935 0.695 0.8 ...
```

```
#remove any p450 values that are 0 - those are true 0's not non-detectable. I am replacing with na so I
data$p450[data$p450 <= 0] <- NA
1
```

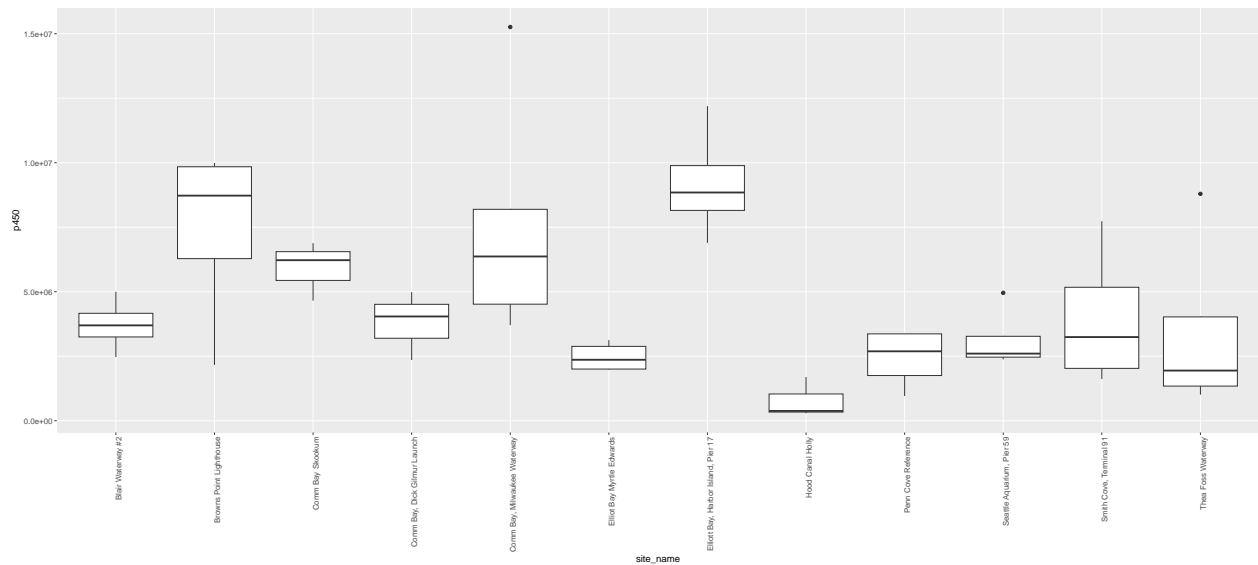
```
## [1] 1
```

Boxplots

p450 - these are non ranked from lowest to highest activity, all sites listed alphabetically.

```
#plot with ordered site names
p450ly<- ggplot(data, aes(x = site_name, y = p450)) +
  geom_boxplot() +
```

```
theme(axis.text.x = element_text(angle = 90, hjust = 1)) # Rotate x labels if needed
print(p450ly)
```



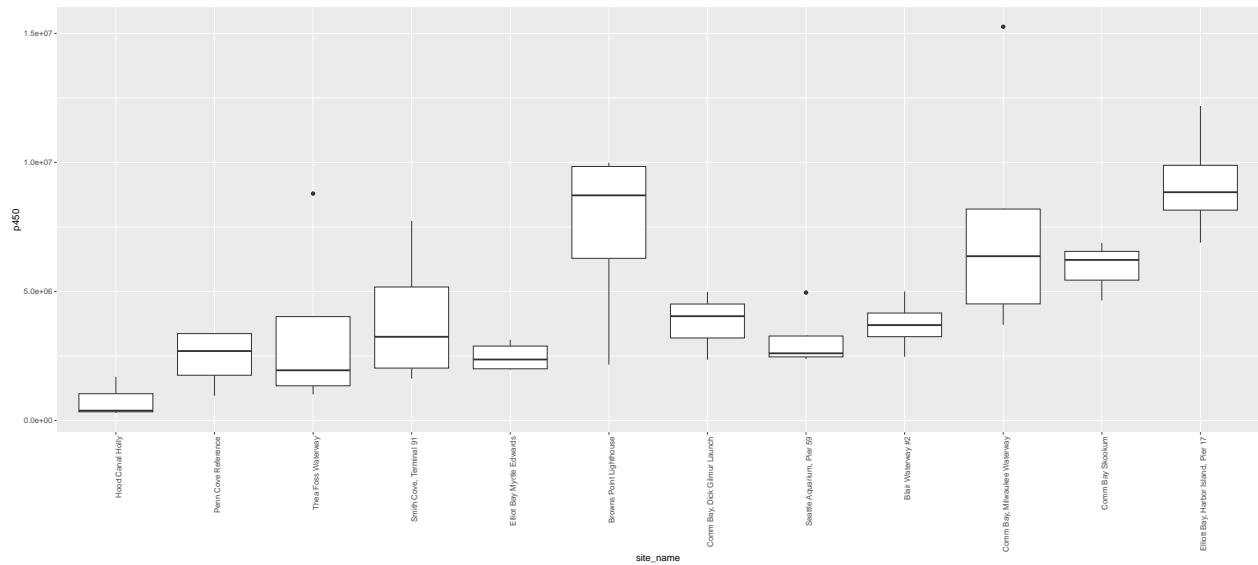
```
#ggsave(plot=p450ly, filename="/Users/cmantegna/Documents/WDFWmussels/output/lycapstone/p450boxplot.png")
```

p450 - these are ranked from lowest to highest activity.

```
#order the sites by value
data_ordered <- data[order(data$p450),]

#create a factor with the ordered site names
data_ordered$site_name <- factor(data_ordered$site_name, levels = unique(data_ordered$site_name))

#plot with ordered site names
lyrankp<- ggplot(data_ordered, aes(x = site_name, y = p450)) +
  geom_boxplot() +
  theme(axis.text.x = element_text(angle = 90, hjust = 1)) # Rotate x labels if needed
print(lyrankp)
```

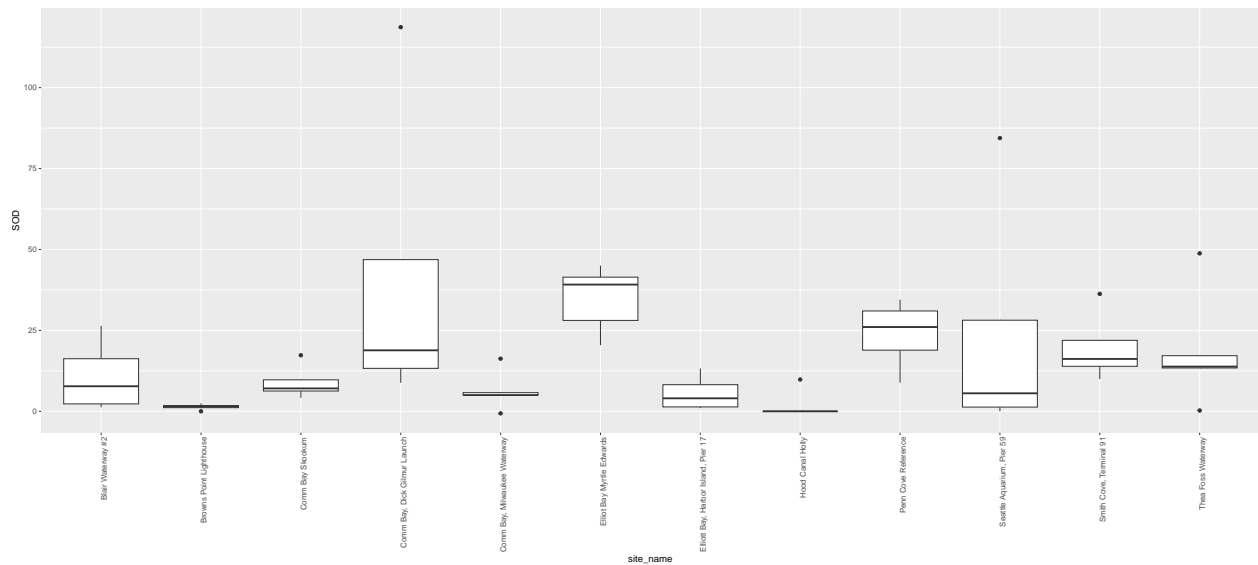


```
#ggsave(plot=lyrankp, filename="/Users/cmantegna/Documents/WDFWmussels/output/lycapstone/p450boxplotran
```

SOD - these are non ranked from lowest to highest activity, all sites listed alphabetically.

```
#plot with ordered site names
sodly<- ggplot(data, aes(x = site_name, y = SOD)) +
  geom_boxplot() +
  theme(axis.text.x = element_text(angle = 90, hjust = 1)) # Rotate x labels if needed

print(sodly)
```



```
#ggsave(plot=sodly, filename="/Users/cmantegna/Documents/WDFWmussels/output/lycapstone/sodboxplot.png",
```

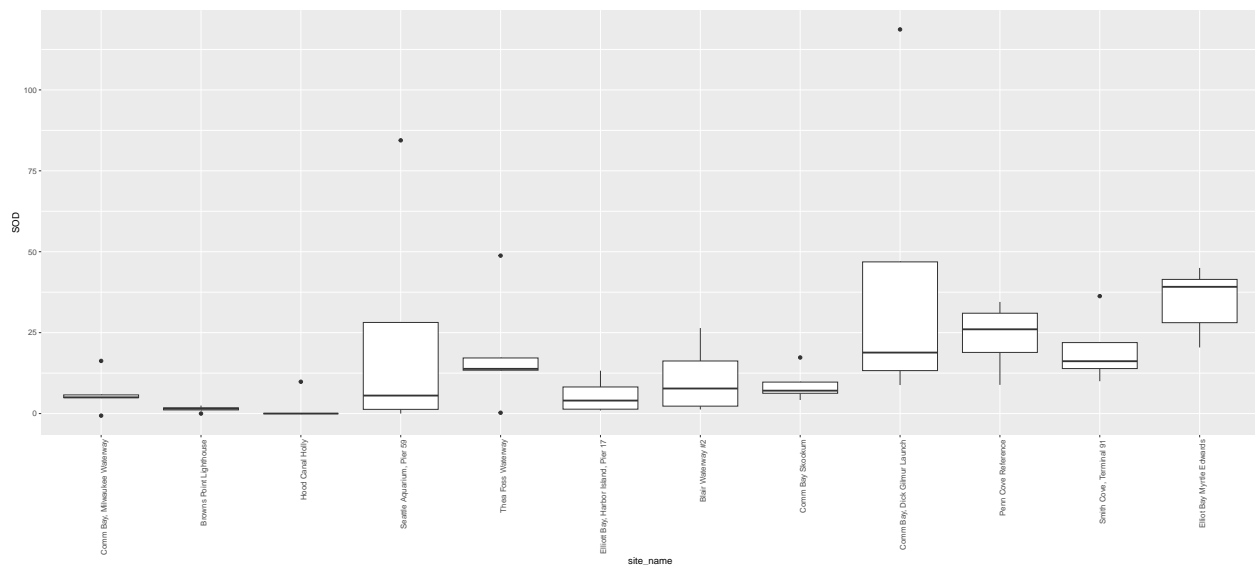
SOD - these are ranked from lowest to highest activity.

```
#order the sites by value
data_ordered <- data[order(data$SOD),]

#create a factor with the ordered site names
data_ordered$site_name <- factor(data_ordered$site_name, levels = unique(data_ordered$site_name))

#plot with ordered site names
lyranks<- ggplot(data_ordered, aes(x = site_name, y = SOD)) +
  geom_boxplot() +
  theme(axis.text.x = element_text(angle = 90, hjust = 1)) # Rotate x labels if needed

print(lyranks)
```



```
#ggsave(plot=lyranks, filename="/Users/cmantegna/Documents/WDFWmussels/output/lycapstone/sodboxplotrank
```

Maps

If you want to keep the full view of Washington state instead of the Puget Sound view we can do that easily.

```
#base map of washington state
library(sf)
library(viridis)
library(rnaturalearth)
library(rnaturalearthdata)

world <- ne_states(country = "united states of america", returnclass = "sf")
washington_map <- world[world$name == "Washington", ]
```

p450 - all sites (puget sound)

```

#zoom into puget sound region & note the legend, lighter colors are higher values
xlim <- c(-124, -122) # longitude bounds
ylim <- c(47, 49) # latitude bounds

lyp450pugetsound<- ggplot() +
  geom_sf(data = washington_map, fill = "lightgrey", color = "white") +
  geom_point(data = data, aes(x = longitude, y = latitude, color = p450), size = 3) +
  scale_color_viridis(option = "C", name = "p450") +
  coord_sf(xlim = xlim, ylim = ylim, expand = FALSE)+
  theme_minimal() +
  labs(title = "Puget Sound - P450 Values")

print(lyp450pugetsound)

```

```
#ggsave(plot=lyp450pugetsound, filename="/Users/cmantegna/Documents/WDFWmussels/output/lycapstone/lyp450pugetsound.png")
```

SOD - all sites (puget sound)

```
#zoom into puget sound region & note the legend, lighter colors are higher values
xlim <- c(-124, -122) # longitude bounds
ylim <- c(47, 49) # latitude bounds

lySODpugetsound<- ggplot() +
  geom_sf(data = washington_map, fill = "lightgrey", color = "white") +
  geom_point(data = data, aes(x = longitude, y = latitude, color = SOD), size = 3) +
  scale_color_viridis(option = "C", name = "SOD") +
  coord_sf(xlim = xlim, ylim = ylim, expand = FALSE)+
  theme_minimal() +
  labs(title = "Puget Sound - SOD Values")

print(lySODpugetsound)
```


Puget Sound – SOD Values

