06- Explore(01) through Analytes(05) with one average value per site

Setup

```
knitr::opts_chunk$set(
  echo = TRUE,          # Display code chunks
  eval = TRUE,          # Evaluate code chunks
  warning = FALSE,      # Hide warnings
  message = FALSE,      # Hide messages
  fig.width = 8,        # Set plot width in inches
  fig.height = 5,       # Set plot height in inches
  fig.align = "center" # Align plots to the center
)
```

Load packages

```
#install.packages("/Users/cmantegna/Downloads/spdep_1.2-8.tar.gz", repos = NULL, type = "source")
library(tidyr)
library(tidyverse)
library(ggplot2)
library(vegan)
library(tinytex)
```

Load data

Note:

For *data* the units are listed below. Weight = g
Length, width, height = mm
p450, SOD = activity/ (mg/protein)
Condition factor, economic factor = unitless
For *pah* and *indv* the units are ng/g

```
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/biomarkerfull.csv")

#pah has complete site values and different summed pah analyte whole tissue values
pah<- read.csv("/Users/cmantegna/Documents/Biomarker Data Analysis/sum_analytes.csv")

#indv has complete site values and individual named pah analyte whole tissue values
indv<- read.csv("/Users/cmantegna/Documents/Biomarker Data Analysis/individual_analytes.csv")
```

```r
# Review data frame structure

str(data)
```

```
## 'data.frame':    312 obs. of  16 variables:
##  $ latitude        : num  48.7 48.7 48.7 48.7 47.5 ...
##  $ longitude       : num  -123 -123 -123 -123 -122 ...
##  $ site_name       : chr  "Aiston Preserve" "Aiston Preserve" "Aiston Preserve" "Aiston Preserve" ..
##  $ site_number     : int  77 77 77 77 13 13 13 13 13 13 ...
##  $ sample          : int  239 240 241 242 281 282 283 284 285 286 ...
##  $ p450            : int  5965780 1508156 4674882 2861653 3448794 6485447 3563340 1813227 1987132 958
##  $ SOD             : num  0 4.88 8.87 0.01 7.08 ...
##  $ weight_initial  : chr  "11.6884" "10.833" "14.7041" "14.6121" ...
##  $ length          : chr  "53.9" "53.49" "55.99" "58.55" ...
##  $ width           : chr  "22.73" "23.92" "27.79" "28.38" ...
##  $ height          : chr  "18.59" "18.36" "19.57" "19.55" ...
##  $ weight_final    : num  3.28 3.48 4.73 4.45 4.62 ...
##  $ weight_change   : chr  "8.41" "7.35" "9.98" "10.17" ...
##  $ condition_factor: chr  "0.1560" "0.1374" "0.1782" "0.1737" ...
##  $ avg_thickness   : num  0.7 0.79 0.825 0.93 0.92 0.965 0.86 0.955 0.875 0.645 ...
##  $ economic_index  : chr  "0.0018" "0.002" "0.002" "0.0021" ...
```

```r
str(pah)
```

```
## 'data.frame':    390 obs. of  10 variables:
##  $ latitude   : num  48.2 48.7 48.2 48.2 47.6 ...
##  $ longitude  : num  -123 -123 -123 -123 -123 ...
##  $ LabSampleID: chr  "119-5763" "119-5753" "119-5764" "119-5765" ...
##  $ SampleID   : chr  "22WB_PCB1-MTW01" "22SAM1115-MTW01" "22WB_PCB2-MTW01" "22WB_PCB3-MTW01" ...
##  $ SiteName   : chr  "Penn Cove Baseline 1" "Aiston Preserve" "Penn Cove Baseline 2" "Penn Cove Basel
##  $ Analyte    : chr  "SumPAHs" "SumPAHs" "SumPAHs" "SumPAHs" ...
##  $ Units      : chr  "ng/g" "ng/g" "ng/g" "ng/g" ...
##  $ PctSolids  : num  20 16.4 17.9 18.8 14.6 ...
##  $ DryValue   : num  89.9 97.4 100.7 101.3 116.1 ...
##  $ Comments   : chr  NA NA NA NA ...
```

```r
str(indv)
```

```
## 'data.frame':    2886 obs. of  12 variables:
##  $ Latitude       : num  47.5 47.5 47.5 47.5 47.5 ...
##  $ Longitude      : num  -122 -122 -122 -122 -122 ...
##  $ LabSampleID    : chr  "119-5708" "119-5708" "119-5708" "119-5708" ...
##  $ SampleID       : chr  "22SAM002-MTW01" "22SAM002-MTW01" "22SAM002-MTW01" "22SAM002-MTW01" ...
##  $ SiteName       : chr  "Arroyo Beach" "Arroyo Beach" "Arroyo Beach" "Arroyo Beach" ...
##  $ Analyte        : chr  "acenaphthene" "acenaphthylene" "anthracene" "benz[a]anthracene" ...
##  $ Units          : chr  "ng/g" "ng/g" "ng/g" "ng/g" ...
##  $ PctSolids      : num  16.7 16.7 16.7 16.7 16.7 ...
##  $ DryValue       : num  -3.9 -3.9 -3.9 13.2 -3 ...
##  $ Qualifier      : chr  "U" "U" "U" "D" ...
##  $ LabMethodConcat: chr  "GC/MS_ASE_MeCl2" "GC/MS_ASE_MeCl2" "GC/MS_ASE_MeCl2" "GC/MS_ASE_MeCl2" ...
##  $ Comments       : chr  "dNPH Surrogate recovery below minimum QA value (60%)" "dNPH Surrogate recov
```

```r
# Review basic data types and stats

summary(data)
```

```
##     latitude       longitude       site_name          site_number
##  Min.   :47.05   Min.   :-123.5   Length:312         Min.   : 1.00
##  1st Qu.:47.33   1st Qu.:-122.7   Class :character   1st Qu.:21.00
##  Median :47.61   Median :-122.6   Mode  :character   Median :40.00
##  Mean   :47.71   Mean   :-122.6                      Mean   :39.66
##  3rd Qu.:48.02   3rd Qu.:-122.4                      3rd Qu.:59.00
##  Max.   :48.82   Max.   :-122.2                      Max.   :77.00
##      sample            p450               SOD          weight_initial
##  Min.   :  1.00   Min.   :       0   Min.   : -0.636   Length:312
##  1st Qu.: 78.75   1st Qu.: 2309270   1st Qu.:  1.201   Class :character
##  Median :156.50   Median : 3746310   Median :  5.730   Mode  :character
##  Mean   :156.50   Mean   : 4430731   Mean   : 10.345
##  3rd Qu.:234.25   3rd Qu.: 5729620   3rd Qu.: 13.752
##  Max.   :312.00   Max.   :52717198   Max.   :133.268
##     length             width              height           weight_final
##  Length:312         Length:312         Length:312         Min.   : 2.484
##  Class :character   Class :character   Class :character   1st Qu.: 3.801
##  Mode  :character   Mode  :character   Mode  :character   Median : 4.417
##                                                           Mean   : 4.656
##                                                           3rd Qu.: 5.003
##                                                           Max.   :20.625
##  weight_change      condition_factor   avg_thickness    economic_index
##  Length:312         Length:312         Min.   :0.3550   Length:312
##  Class :character   Class :character   1st Qu.:0.6800   Class :character
##  Mode  :character   Mode  :character   Median :0.7975   Mode  :character
##                                        Mean   :0.7835
##                                        3rd Qu.:0.8762
##                                        Max.   :1.2600
```

```
summary(pah)
```

```
##     latitude       longitude      LabSampleID          SampleID
##  Min.   :47.05   Min.   :-123.5   Length:390         Length:390
##  1st Qu.:47.34   1st Qu.:-122.7   Class :character   Class :character
##  Median :47.62   Median :-122.6   Mode  :character   Mode  :character
##  Mean   :47.74   Mean   :-122.6
##  3rd Qu.:48.05   3rd Qu.:-122.4
##  Max.   :48.99   Max.   :-122.2
##    SiteName           Analyte             Units             PctSolids
##  Length:390         Length:390         Length:390         Min.   : 7.726
##  Class :character   Class :character   Class :character   1st Qu.:15.149
##  Mode  :character   Mode  :character   Mode  :character   Median :15.855
##                                                           Mean   :16.037
##                                                           3rd Qu.:16.953
##                                                           Max.   :21.245
##     DryValue          Comments
##  Min.   :   11.61   Length:390
##  1st Qu.:  119.14   Class :character
##  Median :  201.58   Mode  :character
##  Mean   :  659.37
##  3rd Qu.:  409.00
##  Max.   :15313.22
```

```r
summary(indv)
```

```
##     Latitude       Longitude      LabSampleID          SampleID
##  Min.   :47.05   Min.   :-123.5   Length:2886        Length:2886
##  1st Qu.:47.34   1st Qu.:-122.7   Class :character   Class :character
##  Median :47.62   Median :-122.6   Mode  :character   Mode  :character
##  Mean   :47.74   Mean   :-122.6
##  3rd Qu.:48.05   3rd Qu.:-122.4
##  Max.   :48.99   Max.   :-122.2
##    SiteName           Analyte             Units             PctSolids
##  Length:2886        Length:2886        Length:2886        Min.   : 7.726
##  Class :character   Class :character   Class :character   1st Qu.:15.149
##  Mode  :character   Mode  :character   Mode  :character   Median :15.855
##                                                           Mean   :16.037
##                                                           3rd Qu.:16.953
##                                                           Max.   :21.245
##     DryValue          Qualifier         LabMethodConcat      Comments
##  Min.   : -12.502   Length:2886        Length:2886        Length:2886
##  1st Qu.:  -3.491   Class :character   Class :character   Class :character
##  Median :   4.481   Mode  :character   Mode  :character   Mode  :character
##  Mean   :  16.329
##  3rd Qu.:  13.843
##  Max.   :1776.333
```

```r
# Data contains 0's and must be adjusted in this order to preserve all usable data.

#sod
#replace any SOD values at or below 0 with half of the lower detection limit of .005 (.005*.5). Lower d
data$SOD[data$SOD <= 0] <- 0.0025

#p450
#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
```

Explore

These are basic statistical tests to explore the current data post cleaning.

```r
#Mean, median, sd, and variance

#stats p450
mean_p450 <- mean(data$p450, na.rm = TRUE)
median_p450 <- median(data$p450, na.rm = TRUE)
sd_p450 <- sd(data$p450, na.rm = TRUE)
var_p450 <- var(data$p450, na.rm = TRUE) #just to see, using sd

#make it a list
stats_p450 <- list(mean = mean_p450, median = median_p450, sd = sd_p450, variance = var_p450)

# Print the results
print(stats_p450)
```

```
## $mean
## [1] 4547329
```

```
##
## $median
## [1] 3873418
##
## $sd
## [1] 4027626
##
## $variance
## [1] 1.622177e+13
```

```
#Mean, median, sd, and variance

#stats sod
mean_SOD <- mean(data$SOD, na.rm = TRUE)
median_SOD <- median(data$SOD, na.rm = TRUE)
sd_SOD <- sd(data$SOD, na.rm = TRUE)
var_SOD <- var(data$SOD, na.rm = TRUE) #just to see, using sd

stats_sod <- list(mean = mean_SOD, median = median_SOD, sd = sd_SOD, variance = var_SOD)

print(stats_sod)
```

```
## $mean
## [1] 10.34759
##
## $median
## [1] 5.73
##
## $sd
## [1] 15.66503
##
## $variance
## [1] 245.3932
```

```
#Quartiles

#p450
quartiles_p450 <- quantile(data$p450, probs = c(0.25, 0.5, 0.75), na.rm = TRUE)
print(quartiles_p450)
```

```
##     25%     50%     75%
## 2413951 3873418 5787570
```

```
#Quartiles

#SOD
quartiles_SOD <- quantile(data$SOD, probs = c(0.25, 0.5, 0.75), na.rm = TRUE)
print(quartiles_SOD)
```

```
##     25%     50%     75%
##  1.2010  5.7300 13.7525
```

```
#Reduction of dataset to the average of biomarker values by site. Individual value analyses were perfor

library(dplyr)

#simplifying the dataframe for joining with next steps
```

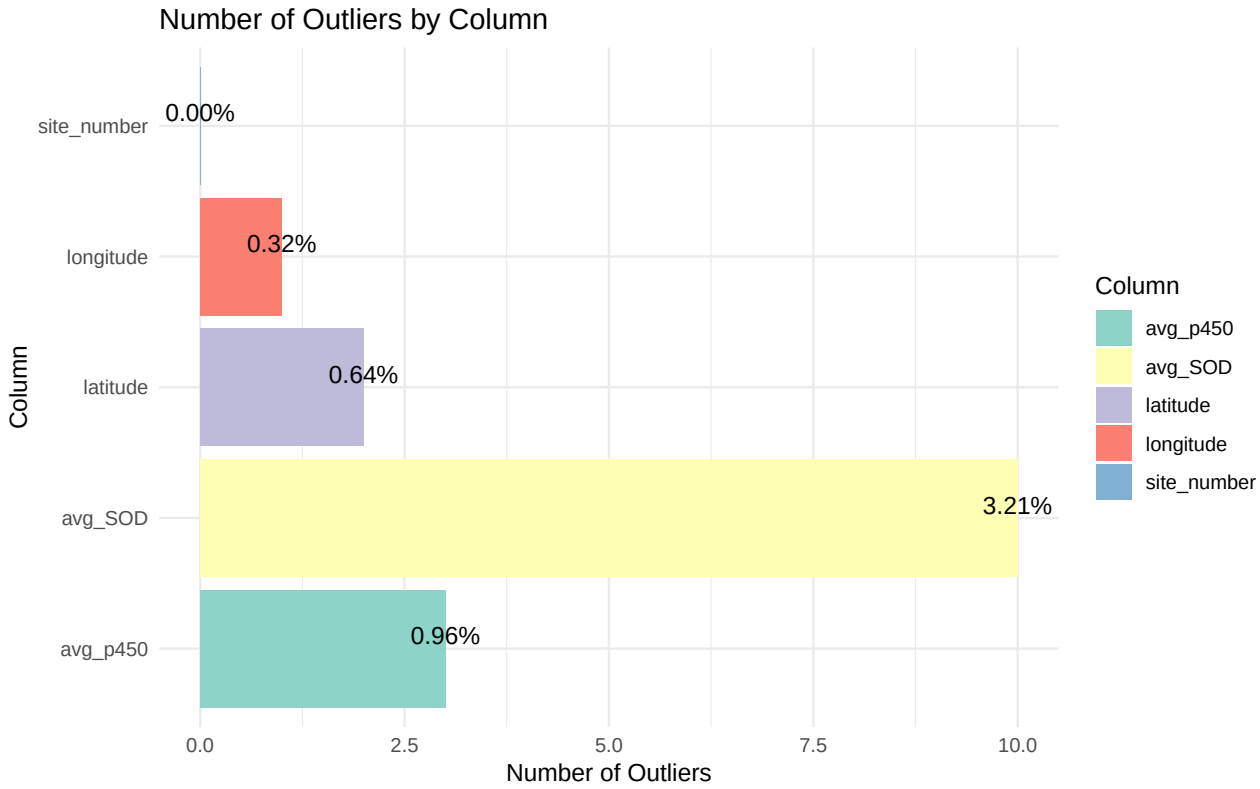
```r
averaged_data <- data %>%
  group_by(site_number, latitude, longitude, site_name) %>%
  summarise(
    avg_p450 = mean(p450, na.rm = TRUE),
    avg_SOD = mean(SOD, na.rm = TRUE)
  ) %>%
  ungroup()  # Remove grouping for the new dataframe

# View the new dataframe with averaged values
averaged_data
```

```
## # A tibble: 74 x 6
##    site_number latitude longitude site_name                avg_p450 avg_SOD
##          <int>    <dbl>     <dbl> <chr>                       <dbl>   <dbl>
##  1           1     48.1     -123. Port Angeles Yacht Club  5751355    7.39
##  2           2     48.0     -123. Jamestown                3263515   24.5
##  3           3     48.2     -123. Penn Cove Reference      2427656.  23.9
##  4           7     48.3     -123. North Camano            12290521    0.752
##  5           8     48.0     -123. Chimacum Creek delta     2641574.   2.19
##  6           9     48.0     -123. S of Skunk Island        3556923.  11.3
##  7          10     48.0     -123. Oak Bay County Park      2335145   19.8
##  8          11     48.0     -123. Maristone Island         4772561.   5.68
##  9          12     48.1     -123. Discovery Bay            4029898.   8.74
## 10          13     47.5     -122. Arroyo Beach             4480860.   8.83
## # i 64 more rows
```

Outliers

```r
# Outliers

# Detect outliers and plot them. Averaging the values reduced the SOD outliers from approximately 9% to
library(ggplot2)
detect_outliers_mad <- function(averaged_data, accuracy = 0.99) {
  # Calculate z-score equivalent for the given accuracy
  z_threshold <- qnorm(accuracy + (1 - accuracy) / 2)

  # Initialize a list to store outlier indices for each numeric column
  outliers_list <- list()

  # Initialize a vector to keep track of rows with outliers
  rows_with_outliers <- rep(FALSE, nrow(averaged_data))

  # Loop through each column in the dataframe
  for (col_name in names(averaged_data)) {
    # Check if the column is numeric
    if (is.numeric(averaged_data[[col_name]])) {
      # Calculate MAD and median for the column
      mad_value <- median(abs(averaged_data[[col_name]] - median(averaged_data[[col_name]])))
      median_value <- median(averaged_data[[col_name]])

      # Calculate the deviation scores (using a modified z-score formula)
      deviation_scores <- 0.6745 * (averaged_data[[col_name]] - median_value) / mad_value
```

```
      # Identify indices of outliers
      outlier_indices <- which(abs(deviation_scores) > z_threshold)

      # Store the indices in the list
      outliers_list[[col_name]] <- outlier_indices

      # Update rows with outliers
      rows_with_outliers[outlier_indices] <- TRUE
    }
  }

  # Return the list of outliers and rows with outliers
  list(outliers_list = outliers_list, rows_with_outliers = rows_with_outliers)
}

outliers_info <- detect_outliers_mad(averaged_data)

# Convert the list of outliers to a named vector of counts
num_outliers_each_col <- sapply(outliers_info$outliers_list, length)
num_rows_with_outliers <- sum(outliers_info$rows_with_outliers)

# Check if there are any outliers
if (all(num_outliers_each_col == 0)) {
  print("There are no outliers in any columns.")
} else {
  # Create a data frame for plotting
  outliers_data_df <- data.frame(
    Column = names(num_outliers_each_col),
    Outliers = as.integer(num_outliers_each_col),
    OutlierPercentage = (as.integer(num_outliers_each_col) / nrow(data)) * 100
  )

  # Plot the number of outliers for all columns
  outlier_plot <- ggplot(outliers_data_df, aes(x = Column, y = Outliers, fill = Column)) +
    geom_bar(stat = "identity") +
    geom_text(aes(label = sprintf("%.2f%%", OutlierPercentage)), position = position_dodge(width = 0.9)
    coord_flip() +
    labs(title = "Number of Outliers by Column", x = "Column", y = "Number of Outliers") +
    scale_fill_brewer(palette = "Set3") +
    theme_minimal()

  print(outlier_plot)
}
```

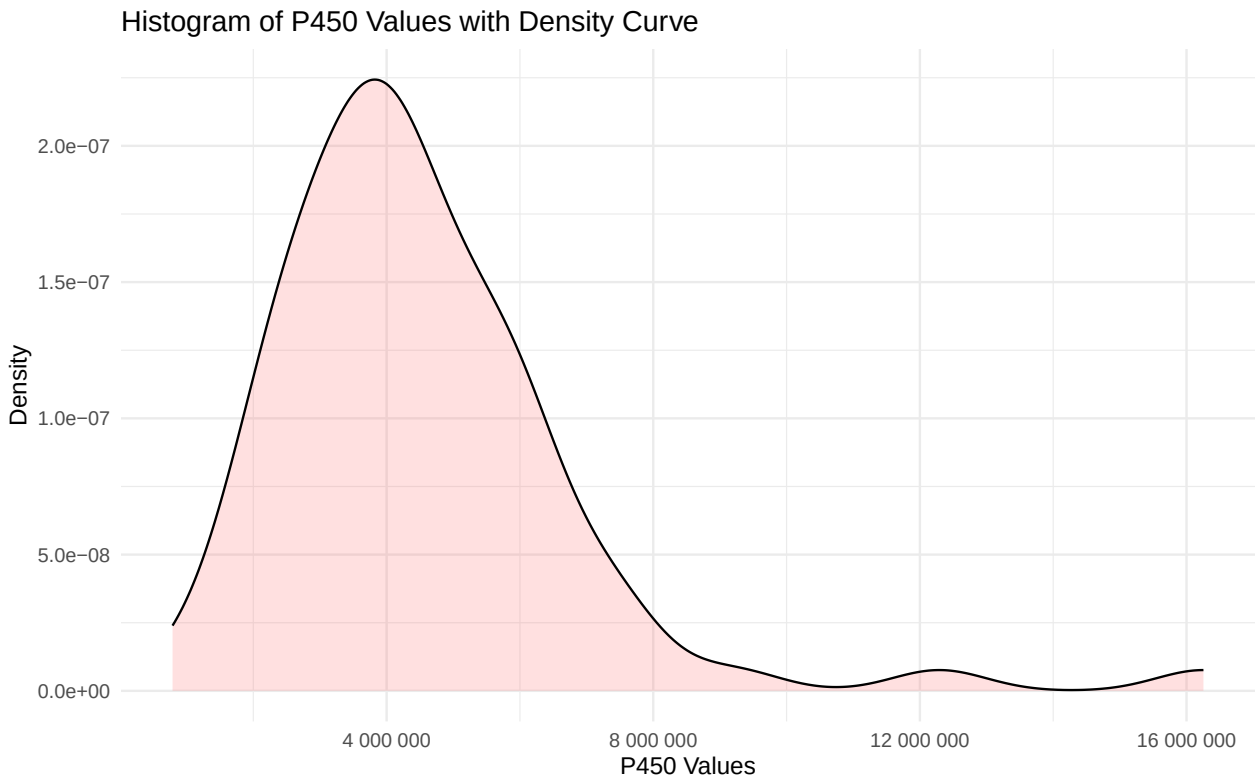



```
#ggsave(plot= outlier_plot, filename="/Users/cmantegna/Documents/WDFWmussels/output/outliersavg.png", w
```

Histograms

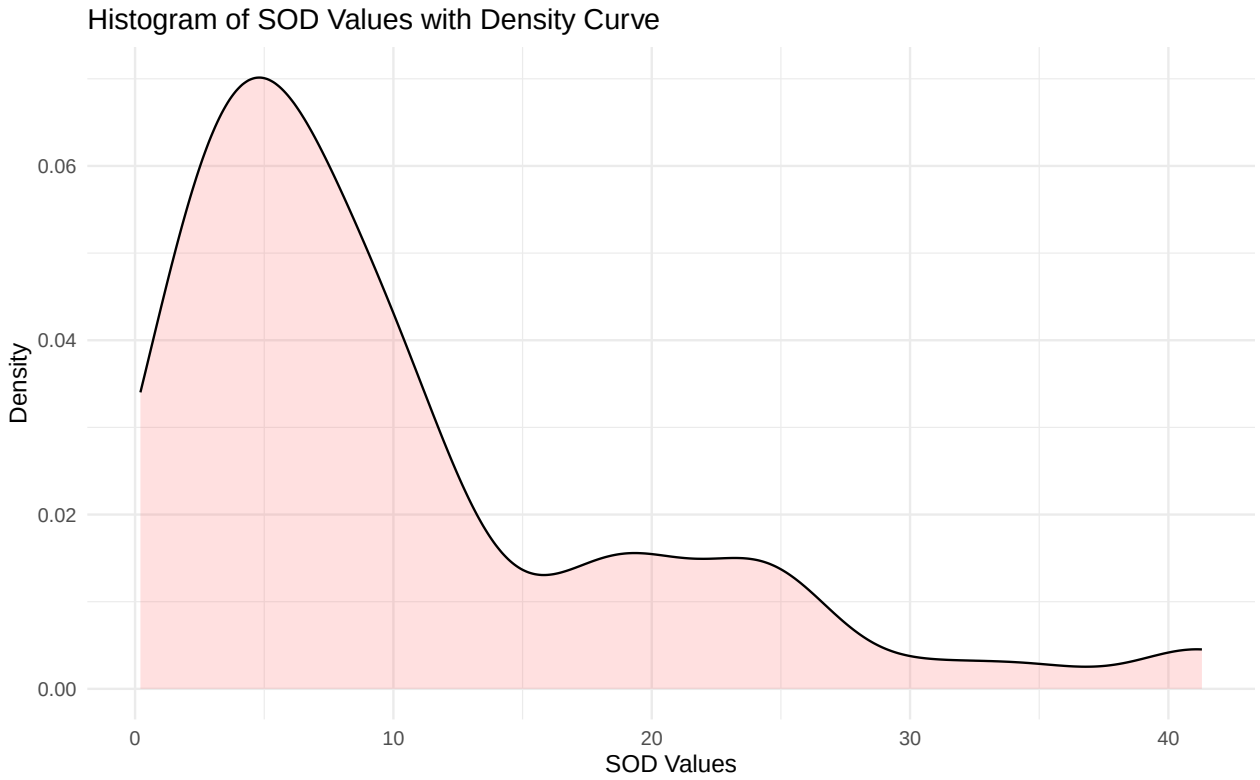
```
#p450 Histogram

#basic histogram + basic density plot
#hist(averaged_data$p450)
#plot(density(averaged_data$p450), main="Density Plot", xlab="p450 Value")

#ggplot histogram with density curve
library(scales) # Make sure this package is loaded for label_number()

phist<- ggplot(averaged_data, aes(x = avg_p450)) +
  geom_histogram(aes(y = after_stat(density)), binwidth = diff(range(data$avg_p450))/30, colour = "blac
  geom_density(alpha = .2, fill = "#FF6666") +
  labs(x = "P450 Values", y = "Density", title = "Histogram of P450 Values with Density Curve") +
  theme_minimal() +
  scale_x_continuous(labels = label_number()) # This line adjusts the x-axis labels

print(phist)
```

Interpretation - data appears not to be normally distributed, but could be transformed via log transformation to get closer if necessary. As a result, any tests for normality will be run in pairs (normal distribution and Poisson distribution).




```
#ggsave(plot=phist, filename="/Users/cmantegna/Documents/WDFWmussels/output/p450avghistogram.png", widt
```

```
# SOD histogram

#basic histogram + basic density plot
#hist(averaged_data$SOD)
#plot(density(averaged_data$SOD), main="Density Plot", xlab="SOD Value")

#ggplot histogram with density curve
library(scales) # Make sure this package is loaded for label_number()

shist <- ggplot(averaged_data, aes(x = avg_SOD)) +
  geom_histogram(aes(y = after_stat(density)), binwidth = diff(range(data$avg_SOD))/30, colour = "black
  geom_density(alpha = .2, fill = "#FF6666") +
  labs(x = "SOD Values", y = "Density", title = "Histogram of SOD Values with Density Curve") +
  theme_minimal() +
  scale_x_continuous(labels = label_number()) # This line adjusts the x-axis labels

print(shist)
```

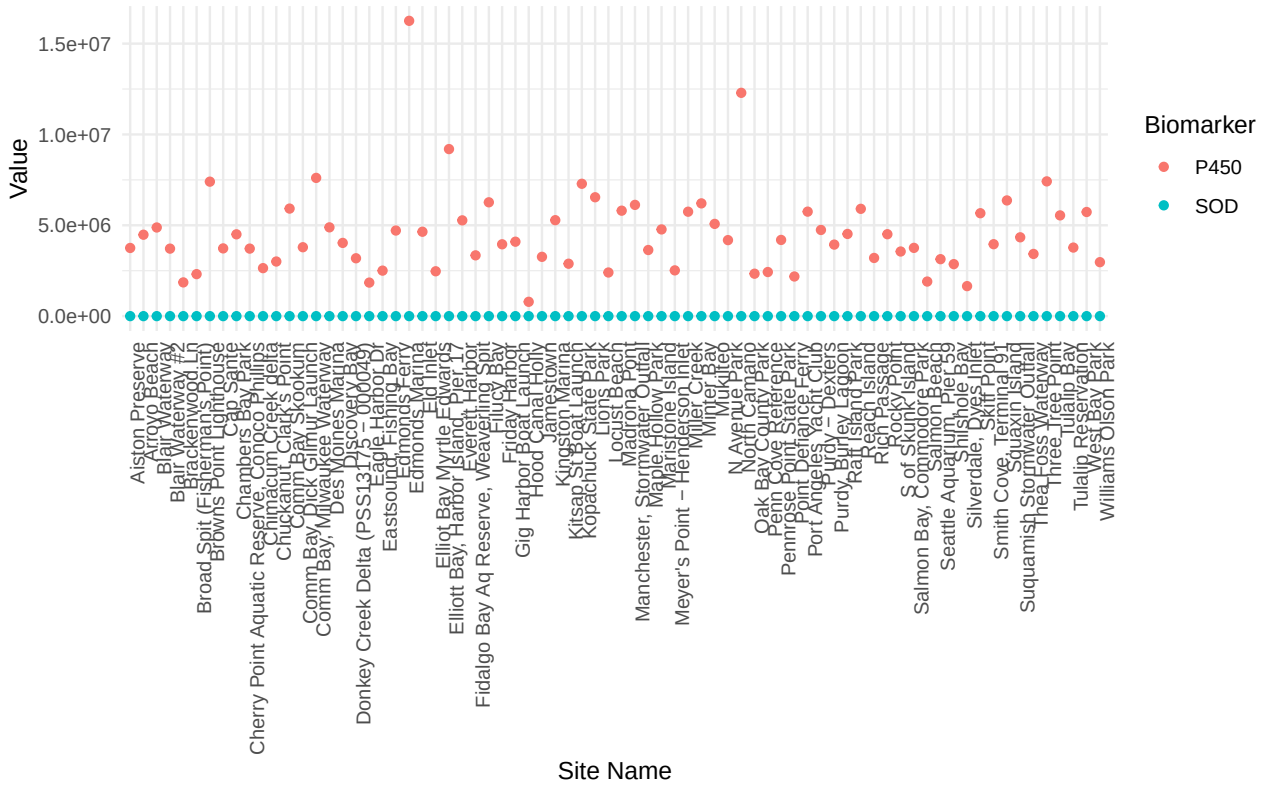



```
#ggsave(plot=shist, filename="/Users/cmantegna/Documents/WDFWmussels/output/SODavghistogram.png", width
```

Boxplots

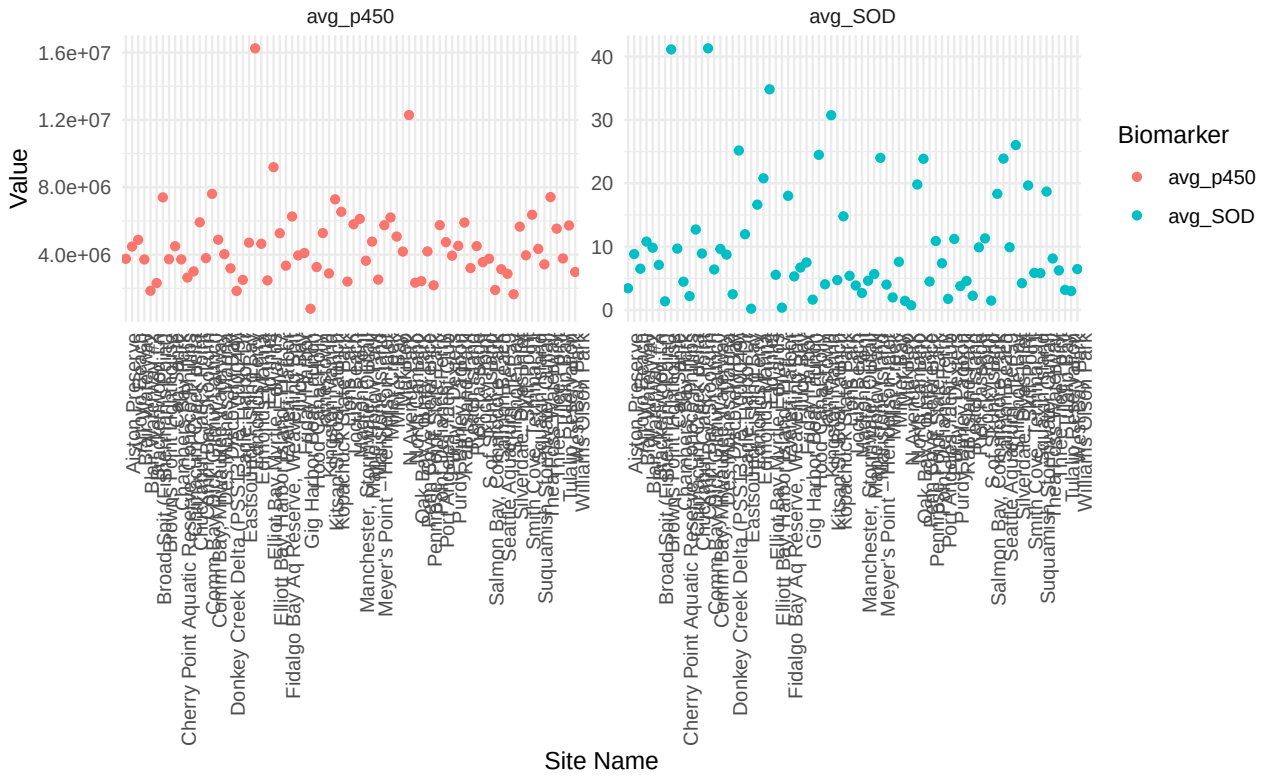
```
# Plotting both biomarkers in a box plot
all<-plot<- ggplot(averaged_data) +
  geom_point(aes(x = site_name, y = avg_SOD, color = "SOD")) +
  geom_point(aes(x = site_name, y = avg_p450, color = "P450")) +
  labs(x = "Site Name", y = "Value", color = "Biomarker") +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 90, hjust = 1))


print(all)
```

Interpretation - two clear outliers exist (~2.7% of the 3.21% outliers) with a third outlier less obvious.



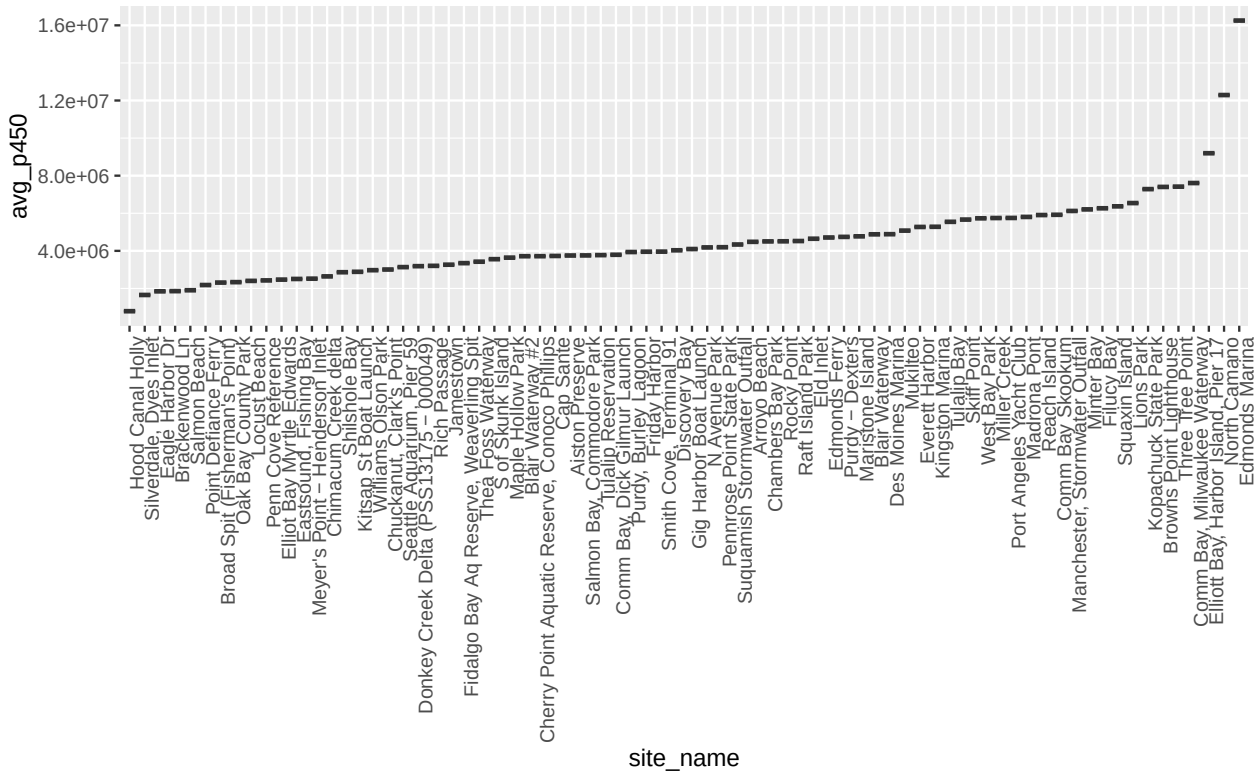

```
#ggsave(plot=all, filename="/Users/cmantegna/Documents/WDFWmussels/output/allavgBoxPlot.png", width=15,
```

```
# Individual boxplots for comparison since the scale for each biomarker is very different.

library(tidyr)

# Reshape
long_data <- pivot_longer(averaged_data, cols = c(avg_SOD, avg_p450), names_to = "Biomarker", values_to

# Faceted plot with separate scales.
plotf<- ggplot(long_data, aes(x = site_name, y = Value)) +
  geom_point(aes(color = Biomarker)) +
  facet_wrap(~ Biomarker, scales = "free_y") +
  labs(x = "Site Name", y = "Value", color = "Biomarker") +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 90, hjust = 1))

print(plotf)
```



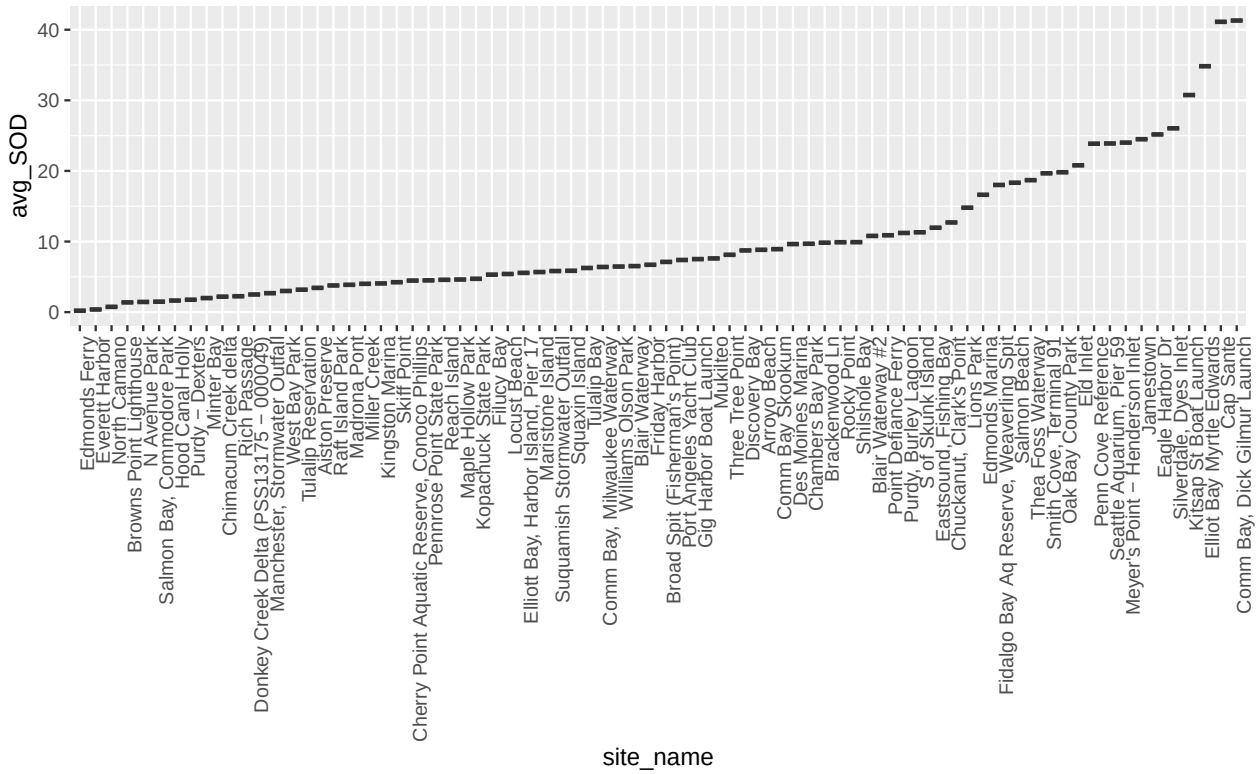

```
#ggsave(plot=plotf, filename="/Users/cmantegna/Documents/WDFWmussels/output/allavgBoxPlotFacetPanels.pn
```

```
# Plotting p450 values ranked from smallest to largest

#order the sites by value
data_ordered <- averaged_data[order(averaged_data$avg_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
rankp<- ggplot(data_ordered, aes(x = site_name, y = avg_p450)) +
  geom_boxplot() +
  theme(axis.text.x = element_text(angle = 90, hjust = 1)) # Rotate x labels if needed

print(rankp)
```




```
#ggsave(plot=rankp, filename="/Users/cmantegna/Documents/WDFWmussels/output/avgp450ranked.png", width=1
```

```
# Plotting SOD values ranked from smallest largest

#order the sites by value
data_ordered <- averaged_data[order(averaged_data$avg_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
ranks<- ggplot(data_ordered, aes(x = site_name, y = avg_SOD)) +
  geom_boxplot() +
  theme(axis.text.x = element_text(angle = 90, hjust = 1)) # Rotate x labels if needed

print(ranks)
```




```
#ggsave(plot=ranks, filename="/Users/cmantegna/Documents/WDFWmussels/output/avgSODranked.png", width=15
```

Statistics

These are more descriptive statistics to test for normality and correlation.

```
library(tidyr)
library(tidyverse)
library(ggplot2)
library(vegan)
```

Shapiro- Wilkes

```
#Test for data normality

shapiro.test(averaged_data$avg_p450)
```

Interpretation - both biomarkers are close to a normal distribution (W=1) but are not normally distributed based on both W and the p-value.

```
##
##  Shapiro-Wilk normality test
##
## data:  averaged_data$avg_p450
## W = 0.82368, p-value = 5.781e-08
```

```
#Test for data normality

shapiro.test(averaged_data$avg_SOD)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  averaged_data$avg_SOD
## W = 0.82966, p-value = 8.711e-08
```

Correlation, Pearson Correlation

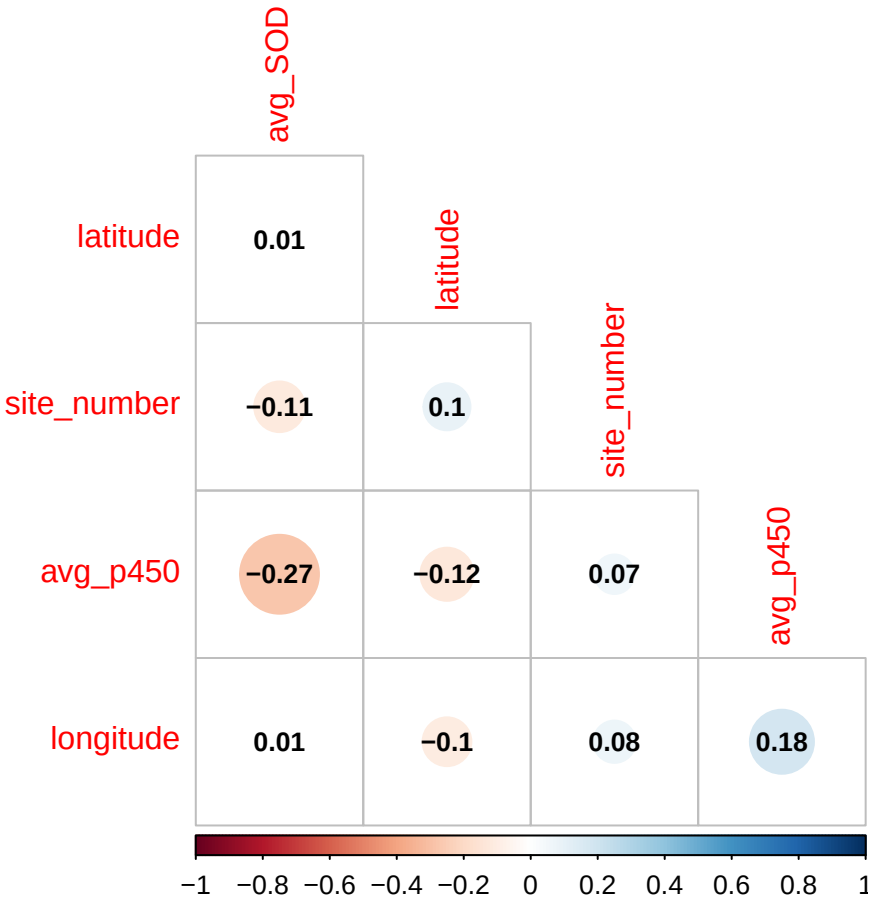
```
# Correlation, Pearson Correlation

#add condition_factor back to the dataframe


#individual correlation test between the biomarkers

cor.test(averaged_data$avg_p450, averaged_data$avg_SOD)
```

Interpretation - the biomarkers show a statistically significant *weak* negative correlation. Where one biomarker increases, the other decreases.

```
##
##  Pearson's product-moment correlation
##
## data:  averaged_data$avg_p450 and averaged_data$avg_SOD
## t = -2.4138, df = 72, p-value = 0.01833
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  -0.47256554 -0.04812135
## sample estimates:
##        cor
## -0.2736115
```

```
# Pearson Correlation Plot

library(corrplot)

cor.mtest <- function(mat, conf.level = 0.95){
  mat <- as.matrix(mat)
  n <- ncol(mat)
  p.mat <- matrix(NA, n, n)
  diag(p.mat) <- 0

  for(i in 1:(n-1)){
    for(j in (i+1):n){
      tmp <- cor.test(mat[,i], mat[,j], conf.level = conf.level)
      p.mat[i,j] <- tmp$p.value
      p.mat[j,i] <- tmp$p.value
    }
  }
  list(p = p.mat, conf.level = conf.level)
}

# Assuming 'mdata' and 'data' are defined and 'data' is used to subset 'mdata'
```

```r
df2 <- averaged_data[sapply(averaged_data, is.numeric)] # Ensure 'mdata' is used here
df2 <- na.omit(df2)
M <- cor(df2)
testRes <- cor.mtest(df2, conf.level = 0.95)

# Visualization with corrplot
corrplot::corrplot(M,
  method = "circle", type = "lower", insig = "blank",
  addCoef.col = "black", number.cex = 0.8, order = "AOE", diag = FALSE
)
```




```r
print(cor)
```

```
## function (x, y = NULL, use = "everything", method = c("pearson",
##     "kendall", "spearman"))
## {
##     na.method <- pmatch(use, c("all.obs", "complete.obs", "pairwise.complete.obs",
##         "everything", "na.or.complete"))
##     if (is.na(na.method))
##         stop("invalid 'use' argument")
##     method <- match.arg(method)
##     if (is.data.frame(y))
##         y <- as.matrix(y)
##     if (is.data.frame(x))
##         x <- as.matrix(x)
```

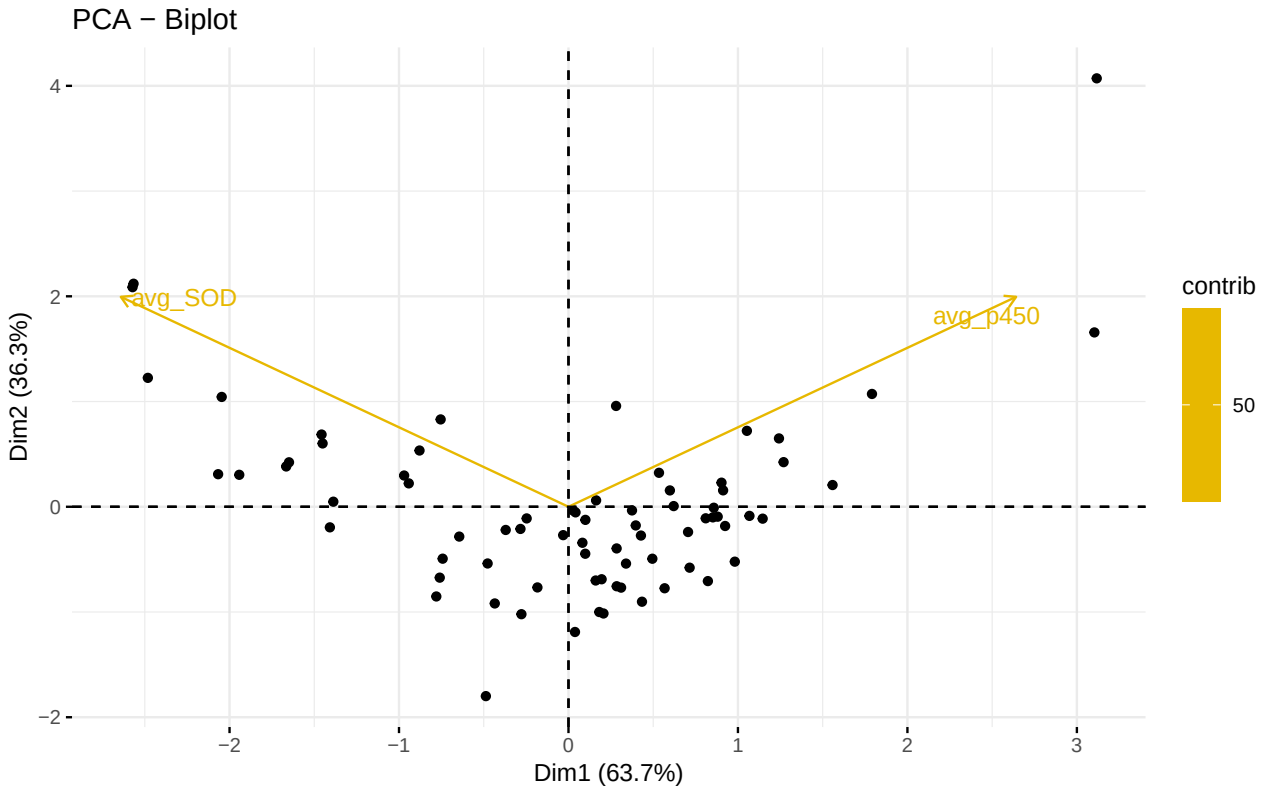
```
##     if (!is.matrix(x) && is.null(y))
##         stop("supply both 'x' and 'y' or a matrix-like 'x'")
##     if (!(is.numeric(x) || is.logical(x)))
##         stop("'x' must be numeric")
##     stopifnot(is.atomic(x))
##     if (!is.null(y)) {
##         if (!(is.numeric(y) || is.logical(y)))
##             stop("'y' must be numeric")
##         stopifnot(is.atomic(y))
##     }
##     Rank <- function(u) {
##         if (length(u) == 0L)
##             u
##         else if (is.matrix(u)) {
##             if (nrow(u) > 1L)
##                 apply(u, 2L, rank, na.last = "keep")
##             else row(u)
##         }
##         else rank(u, na.last = "keep")
##     }
##     if (method == "pearson")
##         .Call(C_cor, x, y, na.method, FALSE)
##     else if (na.method %in% c(2L, 5L)) {
##         if (is.null(y)) {
##             .Call(C_cor, Rank(na.omit(x)), NULL, na.method, method ==
##                 "kendall")
##         }
##         else {
##             nas <- attr(na.omit(cbind(x, y)), "na.action")
##             dropNA <- function(x, nas) {
##                 if (length(nas)) {
##                   if (is.matrix(x))
##                     x[-nas, , drop = FALSE]
##                   else x[-nas]
##                 }
##                 else x
##             }
##             .Call(C_cor, Rank(dropNA(x, nas)), Rank(dropNA(y,
##                 nas)), na.method, method == "kendall")
##         }
##     }
##     else if (na.method != 3L) {
##         x <- Rank(x)
##         if (!is.null(y))
##             y <- Rank(y)
##         .Call(C_cor, x, y, na.method, method == "kendall")
##     }
##     else {
##         if (is.null(y)) {
##             ncy <- ncx <- ncol(x)
##             if (ncx == 0)
##                 stop("'x' is empty")
##             r <- matrix(0, nrow = ncx, ncol = ncy)
##             for (i in seq_len(ncx)) {
```

```
##                 for (j in seq_len(i)) {
##                   x2 <- x[, i]
##                   y2 <- x[, j]
##                   ok <- complete.cases(x2, y2)
##                   x2 <- rank(x2[ok])
##                   y2 <- rank(y2[ok])
##                   r[i, j] <- if (any(ok))
##                     .Call(C_cor, x2, y2, 1L, method == "kendall")
##                   else NA
##                 }
##             }
##             r <- r + t(r) - diag(diag(r))
##             rownames(r) <- colnames(x)
##             colnames(r) <- colnames(x)
##             r
##         }
##         else {
##             if (length(x) == 0L || length(y) == 0L)
##                 stop("both 'x' and 'y' must be non-empty")
##             matrix_result <- is.matrix(x) || is.matrix(y)
##             if (!is.matrix(x))
##                 x <- matrix(x, ncol = 1L)
##             if (!is.matrix(y))
##                 y <- matrix(y, ncol = 1L)
##             ncx <- ncol(x)
##             ncy <- ncol(y)
##             r <- matrix(0, nrow = ncx, ncol = ncy)
##             for (i in seq_len(ncx)) {
##                 for (j in seq_len(ncy)) {
##                   x2 <- x[, i]
##                   y2 <- y[, j]
##                   ok <- complete.cases(x2, y2)
##                   x2 <- rank(x2[ok])
##                   y2 <- rank(y2[ok])
##                   r[i, j] <- if (any(ok))
##                     .Call(C_cor, x2, y2, 1L, method == "kendall")
##                   else NA
##                 }
##             }
##             rownames(r) <- colnames(x)
##             colnames(r) <- colnames(y)
##             if (matrix_result)
##                 r
##             else drop(r)
##         }
##     }
## }
## <bytecode: 0x136d61eb0>
## <environment: namespace:stats>
#ggsave(plot=cor, filename="/Users/cmantegna/Documents/WDFWmussels/output/avgpearson.png", width=15, he
```

PCA Plot

```r
# PCA Plot with biomarkers
#install.packages("FactoMineR")
#install.packages("factoextra")
library('FactoMineR')
library("factoextra")


# Remove NAs from the dataset
df_clean <- na.omit(averaged_data)

# Selecting the relevant variables for PCA
pca_data <- df_clean[, c("avg_SOD", "avg_p450")]

# Performing PCA
pca_res <- PCA(pca_data, scale.unit = TRUE, graph = FALSE)

# Plotting the PCA
pcaplot<- fviz_pca_biplot(pca_res, label = "var", col.var = "contrib",
                gradient.cols = c("#00AFBB", "#E7B800", "#FC4E07"),
                repel = TRUE)  # Avoid text overlapping (slow if many points)



print(pcaplot)
```

Interpretation - This plot has to be interpreted in space, each axis represents a different variance and the attahced percentage to each axis is the total variance captured by that axis; the x-axis is the max variation in the data and should have the higher percentage. The y-axis is the second most variation. The points are observations, more distance mean more difference. The plot is not particularly helpful.




```
#ggsave(plot=pcaplot, filename="/Users/cmantegna/Documents/WDFWmussels/output/pca.png", width=15, heigh
```

K-means cluster/ relationship test

```
#fixing data for kmeans

sum(is.na(averaged_data$avg_p450)) # Checks for NA values in p450
```

Interpretation- there are three clusters that vary in size and that is an indicator of a significant number of outliers, see above for the code for the plot of the outliers. There are significant outliers for SOD, this may be rendering this test unhelpful. "The means of each cluster give you an idea of the centroid values around which the data points in each cluster are aggregated."

```
## [1] 0
```

```
sum(is.nan(averaged_data$avg_p450)) # Checks for NaN values in p450
```

```
## [1] 0
```

```
sum(is.infinite(averaged_data$avg_p450)) # Checks for Inf values in p450
```

```
## [1] 0
```

```
sum(is.na(averaged_data$avg_SOD))
```

```
## [1] 0
```

```
sum(is.nan(averaged_data$avg_SOD))
```

```
## [1] 0
```

```r
sum(is.infinite(averaged_data$avg_SOD))
```

```
## [1] 0
```

```r
#sum(is.na(mdata$condition_factor))
#sum(is.nan(mdata$condition_factor))
#sum(is.infinite(mdata$condition_factor))

averaged_data$avg_p450 <- as.numeric(averaged_data$avg_p450)
averaged_data$avg_SOD <- as.numeric(averaged_data$avg_SOD)
#mdata$condition_factor <- as.numeric(mdata$condition_factor)

# Remove rows with NA or NaN values in specified columns
clean_data <- averaged_data[complete.cases(averaged_data[, c("avg_p450", "avg_SOD")]), ]

# Ensure all data is numeric for the Inf check to make sense
clean_data <- transform(clean_data,
                        avg_p450 = as.numeric(avg_p450),
                        avg_SOD = as.numeric(avg_SOD))

# Now check for and handle Inf values
clean_data <- clean_data[!is.infinite(clean_data$avg_p450) & !is.infinite(clean_data$avg_SOD), ]

kmeans_result <- kmeans(clean_data[, c("avg_p450", "avg_SOD")], centers = 3)

print(kmeans_result)
```

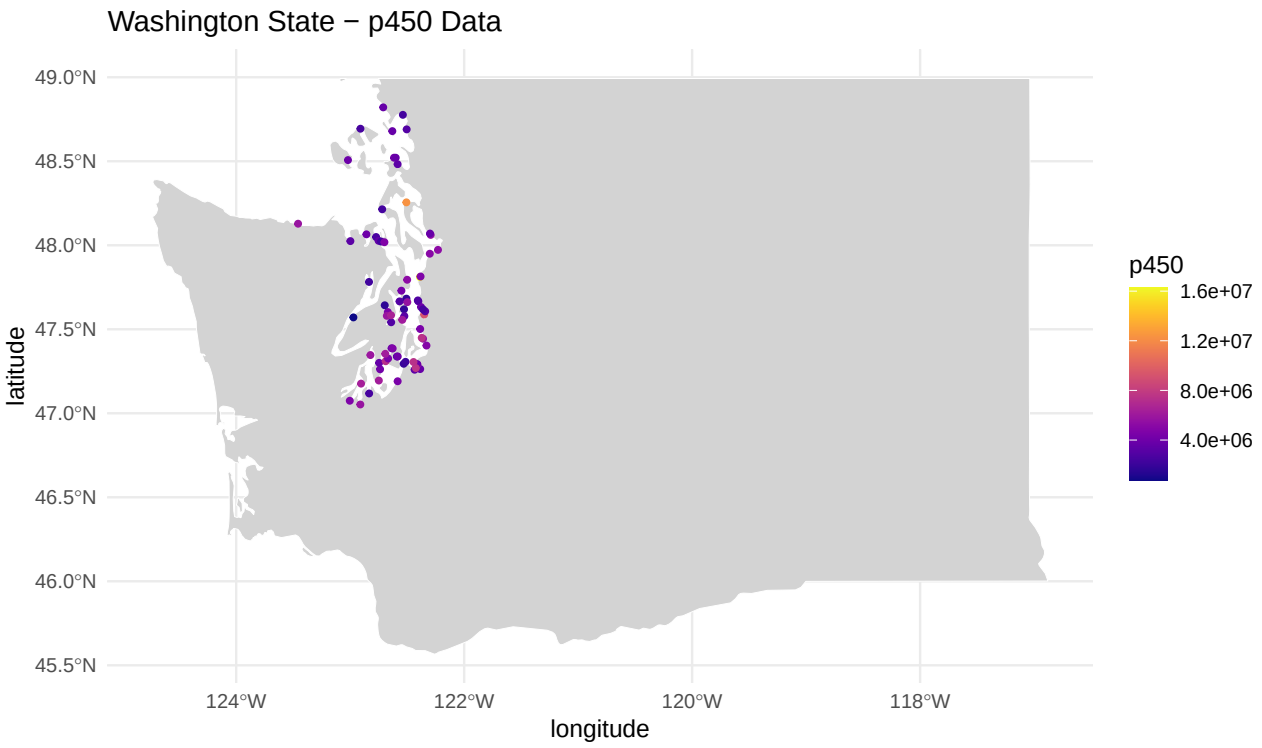
```
## K-means clustering with 3 clusters of sizes 2, 41, 31
##
## Cluster means:
##   avg_p450   avg_SOD
## 1 14272130  8.688250
## 2  3103566 13.610124
## 3  5766556  6.062117
##
## Clustering vector:
##  1  2  3  4  5  6  7  8  9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26
##  3  2  2  1  2  2  2  3  2  3  3  2  2  2  2  2  3  3  3  2  3  2  2  3  3  2
## 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52
##  2  2  2  3  2  2  3  3  3  3  2  3  2  2  2  3  2  2  2  3  2  2  3  3  3  3
## 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74
##  3  3  2  3  2  2  2  2  2  3  3  3  1  3  2  2  3  3  2  2  2  2
##
## Within cluster sum of squares by cluster:
## [1] 7.853548e+12 2.798817e+13 3.779290e+13
##  (between_SS / total_SS =  81.3 %)
##
## Available components:
##
## [1] "cluster"      "centers"      "totss"        "withinss"     "tot.withinss"
## [6] "betweenss"    "size"         "iter"         "ifault"
```

Mapping

```
library(tidyr)
#library(tidyverse)
library(ggplot2)
library(vegan)
library(sf)
library(viridis)
library(rnaturalearth)
library(rnaturalearthdata)
```

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

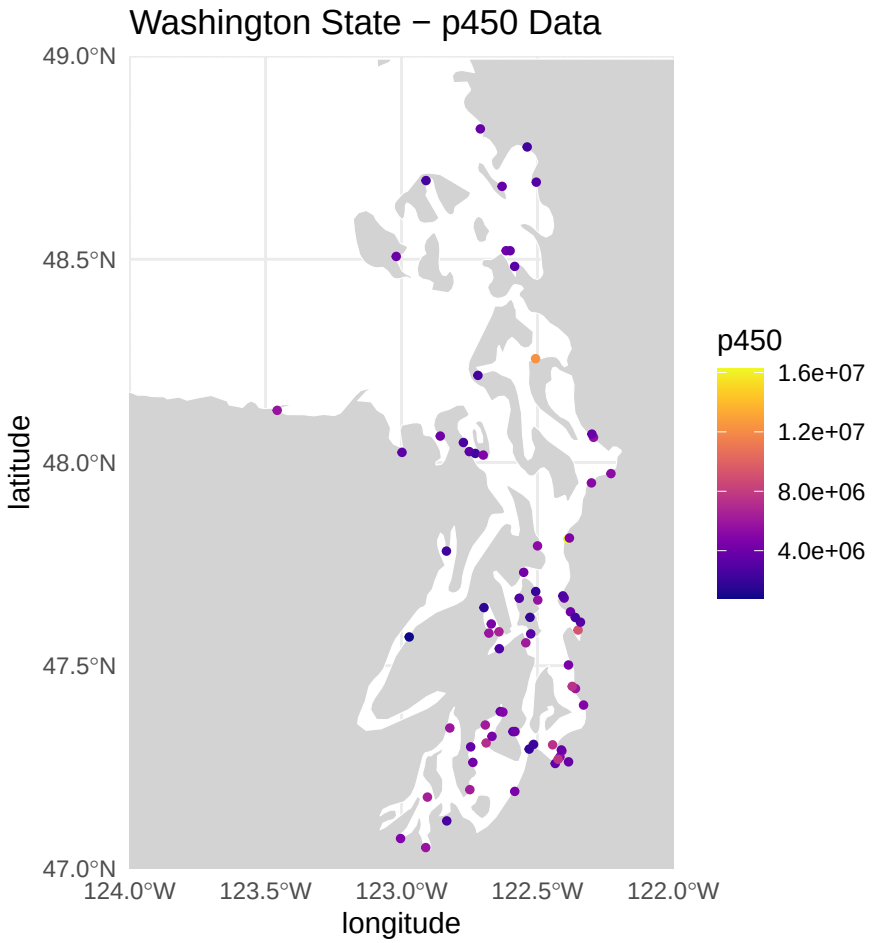
```
# p450 map
pmap<- ggplot() +
  geom_sf(data = washington_map, fill = "lightgrey", color = "white") +
  geom_point(data = averaged_data, aes(x = longitude, y = latitude, color = avg_p450), size = 1) +
  scale_color_viridis(option = "C", name = "p450") +
  theme_minimal() +
  labs(title = "Washington State - p450 Data")

print(pmap)
```




```
#ggsave(plot=pmap, filename="/Users/cmantegna/Documents/WDFWmussels/output/avgp450map.png", width=15, h
```

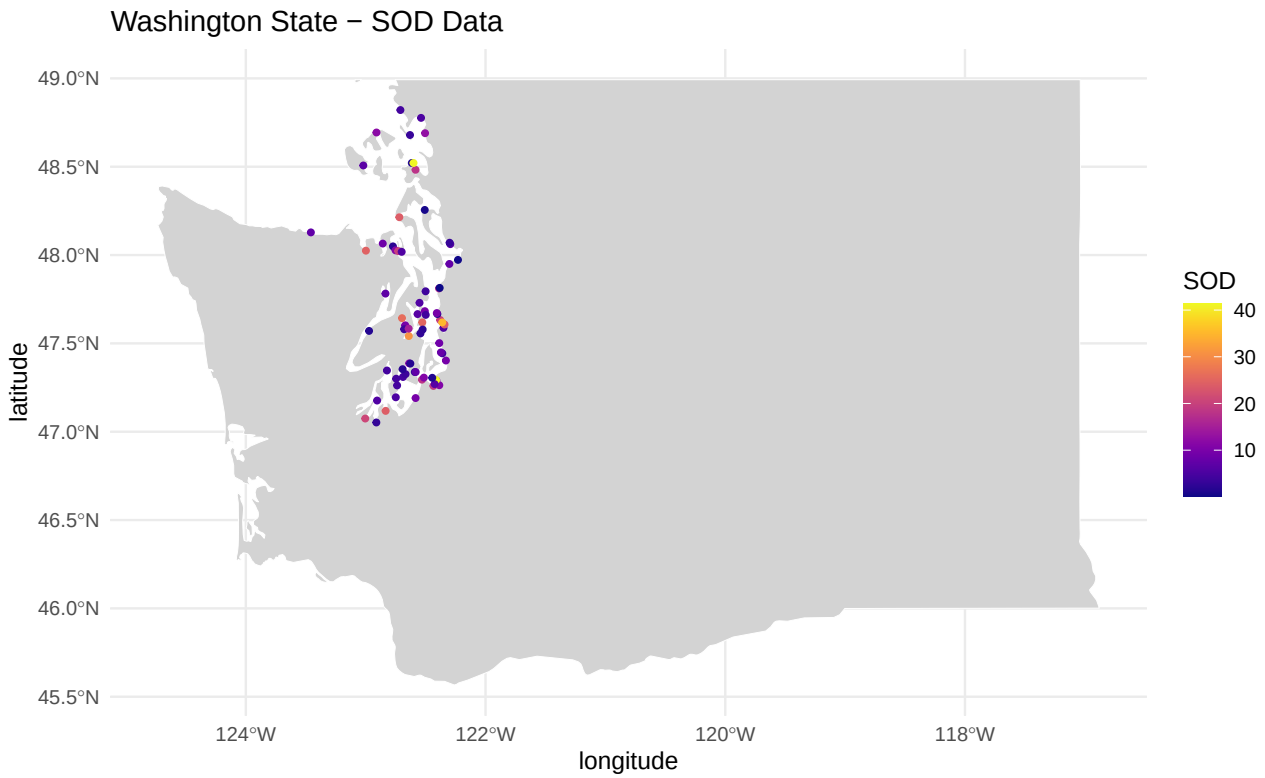
```
#zoom into puget sound region
xlim <- c(-124, -122)  # longitude bounds
ylim <- c(47, 49)  # latitude bounds

pmap<- ggplot() +
```

```
  geom_sf(data = washington_map, fill = "lightgrey", color = "white") +
  geom_point(data = averaged_data, aes(x = longitude, y = latitude, color = avg_p450), size = 1) +
  scale_color_viridis(option = "C", name = "p450") +
  coord_sf(xlim = xlim, ylim = ylim, expand = FALSE)+
  theme_minimal() +
  labs(title = "Washington State - p450 Data")

print(pmap)
```



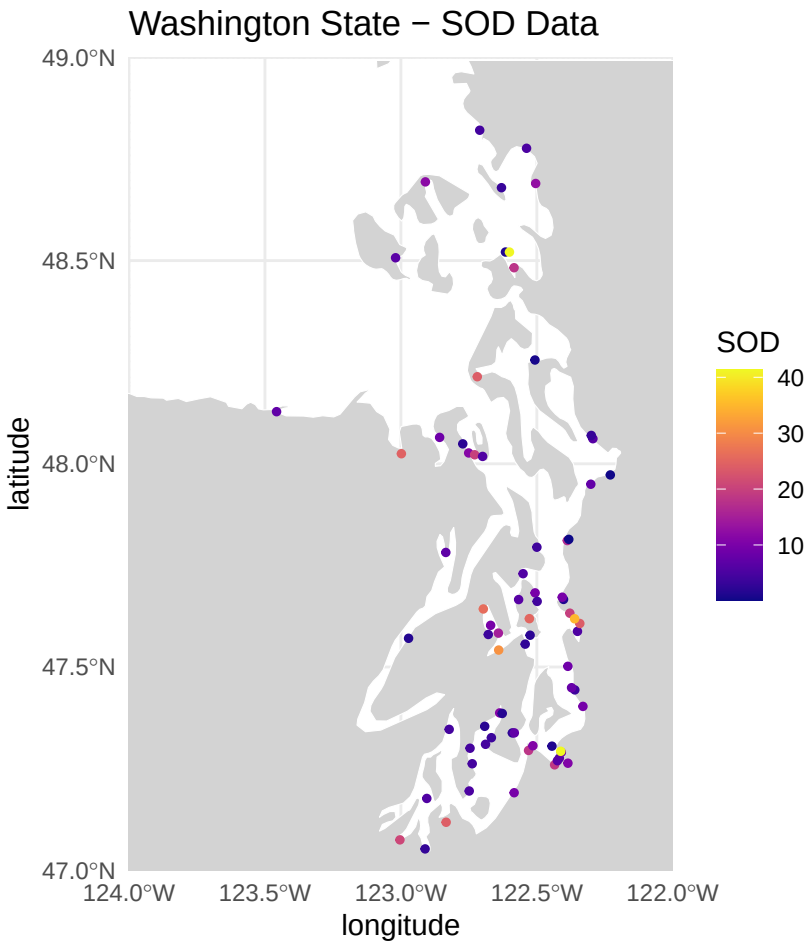

```
#ggsave(plot=pmap, filename="/Users/cmantegna/Documents/WDFWmussels/output/avgp450psmap.png", width=15,
```

```
#SOD map
smap<- ggplot() +
  geom_sf(data = washington_map, fill = "lightgrey", color = "white") +
  geom_point(data = averaged_data, aes(x = longitude, y = latitude, color = avg_SOD), size = 1) +
  scale_color_viridis(option = "C", name = "SOD") +
  theme_minimal() +
  labs(title = "Washington State - SOD Data")

print(smap)
```



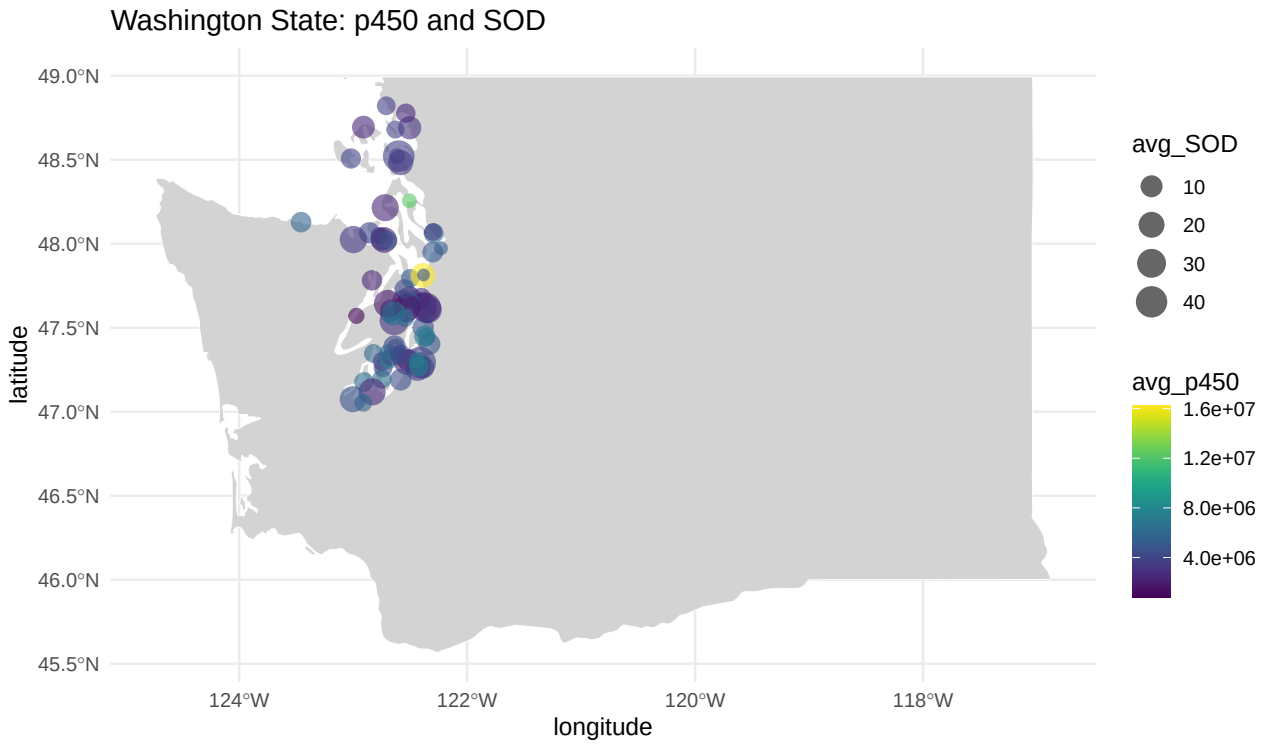

```
#ggsave(plot=smap, filename="/Users/cmantegna/Documents/WDFWmussels/output/avgsodmap.png", width=15, he
```

```
#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

pmap<- ggplot() +
  geom_sf(data = washington_map, fill = "lightgrey", color = "white") +
  geom_point(data = averaged_data, aes(x = longitude, y = latitude, color = avg_SOD), size = 1) +
  scale_color_viridis(option = "C", name = "SOD") +
  coord_sf(xlim = xlim, ylim = ylim, expand = FALSE)+
  theme_minimal() +
  labs(title = "Washington State - SOD Data")

print(pmap)
```

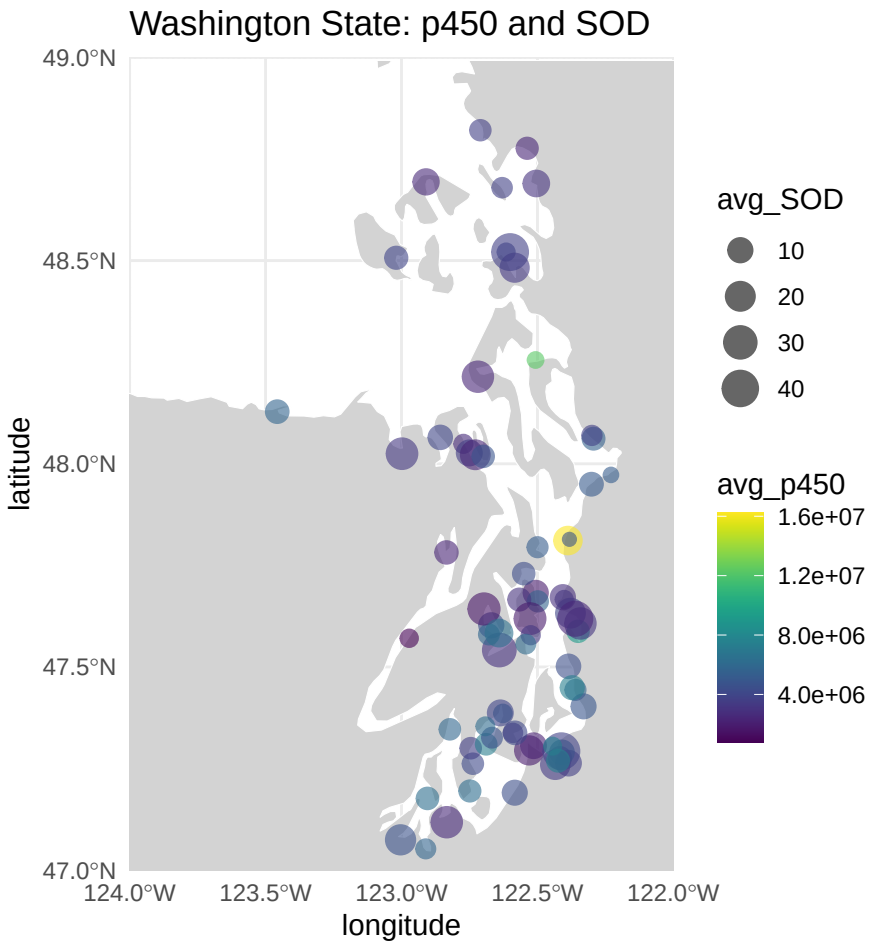



```
#ggsave(plot=pmap, filename="/Users/cmantegna/Documents/WDFWmussels/output/avgSODpsmap.png", width=15,
```

```
#mapping both
# Assuming washington_map is your sf object for Washington State
ggplot(data = washington_map) +
  geom_sf(fill = "lightgrey", color = "white") +
  geom_point(data = averaged_data, aes(x = longitude, y = latitude, color = avg_p450, size = avg_SOD), a
  scale_color_viridis(name = "avg_p450", option = "D") +
  scale_size_continuous(name = "avg_SOD", range = c(2, 6)) +
  theme_minimal() +
  labs(title = "Washington State: p450 and SOD")
```




```
#still not a helpful visualization
xlim <- c(-124, -122)  # longitude bounds
ylim <- c(47, 49)  # latitude bounds

# Assuming washington_map is your sf object for Washington State
ggplot(data = washington_map) +
  geom_sf(fill = "lightgrey", color = "white") +
  geom_point(data = averaged_data, aes(x = longitude, y = latitude, color = avg_p450, size = avg_SOD), a
  scale_color_viridis(name = "avg_p450", option = "D") +
  scale_size_continuous(name = "avg_SOD", range = c(2, 6)) +
  coord_sf(xlim = xlim, ylim = ylim, expand = FALSE)+
  theme_minimal() +
  labs(title = "Washington State: p450 and SOD")
```



Spatial Analysis

Note this workflow is used to determine if there are spatial data patterns that wouldn't stand out in the exploratory or relational statistics.

```
# Data prep for geospatial analysis

#install.packages("spatstat")
library(spatstat)

# Create the kind of data frame that works in spatstat package
sites_pp <- ppp(x = averaged_data$longitude, y = averaged_data$latitude,
                window = owin(xrange = range(averaged_data$longitude),
                              yrange = range(averaged_data$latitude)))
```

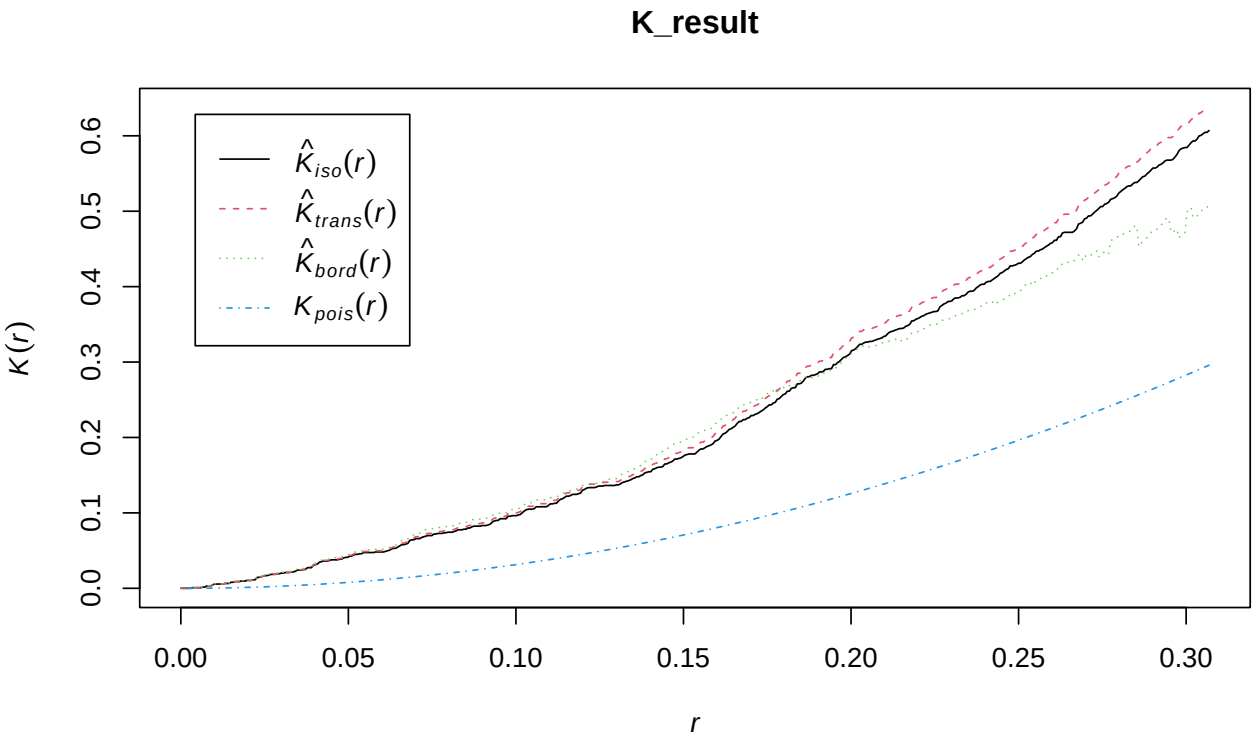
```
#Add biomarker layer to the geographical data
sites_pp$marks <- data.frame(p450 = averaged_data$avg_p450,
                             SOD = averaged_data$avg_SOD)
```

K test, 1 of 2.

```
# average values result's mirror the results returned in the full biomarker run. see 04-spatial for int
```

```
K_result <- Kest(sites_pp)
print(plot(K_result))
```

Interpretation- the clustering of the black, red, and green lines (isotropic, and two transformations) lie significanlty above the blue line, the Poisson distribution line. This means that the biomarker data is more clustered than random. The proximity of the black, red and green lines show agreement across the isotropic and transformation analyses. This result is in agreement with the Shapiro- Wilkes tests for normality.




```
##        lty col    key                    label
## iso      1   1    iso   italic(hat(K)[iso](r))
## trans    2   2  trans italic(hat(K)[trans](r))
## border   3   3 border  italic(hat(K)[bord](r))
## theo     4   4   theo       italic(K[pois](r))
##                                      meaning
## iso      isotropic-corrected estimate of K(r)
## trans  translation-corrected estimate of K(r)
## border      border-corrected estimate of K(r)
## theo                theoretical Poisson K(r)
```

Point Process Model (PPM) model, 2 of 2

```
# Fit a point process model
#ppm_model <- ppm(sites_pp ~ p450)
#summary(ppm_model)


#option 2
sites_pp$marks <- as.factor(cut(sites_pp$marks$p450, breaks = quantile(sites_pp$marks$p450, probs = 0:3,
```

```r
# Now run the ppm model
ppm_model <- ppm(sites_pp ~ marks, covariates = NULL)
print(summary(ppm_model))
```

Interpretation - this model measures the independence and intensity of the occurence of something in space or time based on whatever marks you determine. In this case the occurences are the data point locations, the intensity is the value of the marks and the marks are the biomarkers. This model is confusing to me and will be removed as I do not understand it well enough to interpret the results.

```
## Point process model
## Fitted to data: sites_pp
## Fitting method: maximum likelihood (Berman-Turner approximation)
## Model was fitted using glm()
## Algorithm converged
## Call:
## ppm.formula(Q = sites_pp ~ marks, covariates = NULL)
## Edge correction: "border"
##  [border correction distance r = 0 ]
## --------------------------------------------------------------------------------
## Quadrature scheme (Berman-Turner) = data + dummy + weights
##
## Data pattern:
## Marked planar point pattern:  74 points
## Average intensity 34.1 points per square unit
## Multitype:
##                     frequency proportion intensity
## [7.88e+05,3.58e+06]        25      0.338      11.5
## (3.58e+06,4.84e+06]        24      0.324      11.1
## (4.84e+06,1.63e+07]        25      0.338      11.5
##
## Window: rectangle = [-123.45715, -122.22977] x [47.05236, 48.8208] units
##                     (1.227 x 1.768 units)
## Window area = 2.17055 square units
##
## Dummy quadrature points:
##      32 x 32 grid of dummy points, plus 4 corner points
##      dummy spacing: 0.03835562 x 0.05526369 units
##
## Original dummy parameters: =
## Marked planar point pattern:  3232 points
## Average intensity 1490 points per square unit
## Multitype:
##                     frequency proportion intensity
## [7.88e+05,3.58e+06]      1080      0.333       496
## (3.58e+06,4.84e+06]      1080      0.334       497
## (4.84e+06,1.63e+07]      1080      0.333       496
##
## Window: rectangle = [-123.45715, -122.22977] x [47.05236, 48.8208] units
##                     (1.227 x 1.768 units)
## Window area = 2.17055 square units
## Quadrature weights:
##      (counting weights based on 32 x 32 array of rectangular tiles)
## All weights:
```

```
##  range: [0.00053, 0.00212]   total: 6.51
## Weights on data points:
##  range: [0.00053, 0.00106]   total: 0.0684
## Weights on dummy points:
##  range: [0.00053, 0.00212]   total: 6.44
## ------------------------------------------------------------------------------
## FITTED :
##
## Stationary multitype Poisson process
## Possible marks:
## [7.88e+05,3.58e+06] (3.58e+06,4.84e+06] (4.84e+06,1.63e+07]
## ---- Intensity: ----
##
## Log intensity: ~marks
##
## Intensities:
## beta_[7.88e+05,3.58e+06] beta_(3.58e+06,4.84e+06] beta_(4.84e+06,1.63e+07]
##                 11.51784                 11.05713                 11.51784
##
##                               Estimate      S.E.    CI95.lo   CI95.hi Ztest
## (Intercept)                2.443897e+00 0.2000000  2.0519045 2.8358901   ***
## marks(3.58e+06,4.84e+06] -4.082199e-02 0.2857738 -0.6009284 0.5192844
## marks(4.84e+06,1.63e+07] -9.960261e-14 0.2828427 -0.5543615 0.5543615
##                                  Zval
## (Intercept)                1.221949e+01
## marks(3.58e+06,4.84e+06] -1.428472e-01
## marks(4.84e+06,1.63e+07] -3.521484e-13
##
## ----------- gory details -----
##
## Fitted regular parameters (theta):
##              (Intercept) marks(3.58e+06,4.84e+06] marks(4.84e+06,1.63e+07]
##             2.443897e+00           -4.082199e-02           -9.960261e-14
##
## Fitted exp(theta):
##              (Intercept) marks(3.58e+06,4.84e+06] marks(4.84e+06,1.63e+07]
##                 11.51784                 0.96000                  1.00000
```

Spatial Autocorrelation Analysis, Global

Note: there are global and local tests that assess the likelihood of the biomarker values to be randomly 'assigned'. My hypothesis is that the biomarkers values are not random.

```
#load packages & prep dataframe for "sf" package

#install.packages(spdep)
#install.packages("/Users/cmantegna/Downloads/spdep_1.2-8.tar.gz", repos = NULL, type = "source")
library(sf)
library(sp)
library(spdep)

sf_data <- st_as_sf(averaged_data, coords = c("longitude", "latitude"), crs = 4326)
```

```r
# Take a look at your sf object
print(sf_data)
```

```
## Simple feature collection with 74 features and 4 fields
## Geometry type: POINT
## Dimension:     XY
## Bounding box:  xmin: -123.4572 ymin: 47.05236 xmax: -122.2298 ymax: 48.8208
## Geodetic CRS:  WGS 84
## # A tibble: 74 x 5
##    site_number site_name              avg_p450 avg_SOD             geometry
##  *       <int> <chr>                     <dbl>   <dbl>          <POINT [°]>
##  1           1 Port Angeles Yacht Cl~   5.75e6   7.39  (-123.4571 48.12823)
##  2           2 Jamestown                3.26e6  24.5   (-122.9981 48.02479)
##  3           3 Penn Cove Reference      2.43e6  23.9    (-122.719 48.21423)
##  4           7 North Camano             1.23e7   0.752  (-122.507 48.25536)
##  5           8 Chimacum Creek delta     2.64e6   2.19  (-122.7724 48.04906)
##  6           9 S of Skunk Island        3.56e6  11.3   (-122.7508 48.02667)
##  7          10 Oak Bay County Park      2.34e6  19.8   (-122.7287 48.02218)
##  8          11 Maristone Island         4.77e6   5.68  (-122.6995 48.01813)
##  9          12 Discovery Bay            4.03e6   8.74  (-122.8575 48.06496)
## 10          13 Arroyo Beach             4.48e6   8.83  (-122.3859 47.50161)
## # i 64 more rows
```

```r
# Create a spatial weights matrix to assess nearest neighbors and distance-based neighbors to define sp
# You can extract the matrix of coordinates using st_coordinates
coords <- st_coordinates(sf_data)

# Now use the knearneigh function from the spdep package directly on the coordinates
neighbors <- knn2nb(knearneigh(coords, k = 4))

# Then convert the neighbors into spatial weights with nb2listw
weights <- nb2listw(neighbors, style = "W")
```

Moran's I and Geary's C determine if the distribution of the biomarker values across the entire Puget Sound region are random or not- random.

Interpretation - Moran's I statistic's, used for normally distributed data, ranges from 0 - 1. A value close to 0 mean randomness and a value close to 1 means less randomness. A statistic of .04 means points that are close to each other are more likely to have similar *p450* values than you would expect by chance but this is not statistically significant. The *SOD* values are random and not statistically significant.
Geary's C statistic, used for non- normal data, ranges from 0-2. A value close to 0 is a positive spatial correlation, a value close to 1 is an indication of randomness, and a value close to 2 is an indication of dispersion (negative correlation). A statistic of .8 with a p-value below .05 shows a statistically significant weak indication of negative autocorrelation of *p450* values. The *SOD* values are just above 1 and are not statistically significant. Since Geary's C is influenced by "local" and "global' neighbors, this should not be used.

```r
## Run Moran's I and Geary's C for p450

moran_result <- moran.test(sf_data$avg_p450, weights)
print(moran_result)
```

```
##
##  Moran I test under randomisation
##
```

```
## data:  sf_data$avg_p450
## weights: weights
##
## Moran I statistic standard deviate = 0.75836, p-value = 0.2241
## alternative hypothesis: greater
## sample estimates:
## Moran I statistic       Expectation          Variance
##       0.038455120      -0.013698630       0.004729608
```

```
geary_result <- geary.test(sf_data$avg_p450, weights)
print(geary_result)
```

```
##
##  Geary C test under randomisation
##
## data:  sf_data$avg_p450
## weights: weights
##
## Geary C statistic standard deviate = 1.9348, p-value = 0.02651
## alternative hypothesis: Expectation greater than statistic
## sample estimates:
## Geary C statistic       Expectation          Variance
##         0.7858923         1.0000000         0.0122455
```

```
## Run Moran's I and Geary's C for SOD
moran_result <- moran.test(sf_data$avg_SOD, weights)
print(moran_result)
```

```
##
##  Moran I test under randomisation
##
## data:  sf_data$avg_SOD
## weights: weights
##
## Moran I statistic standard deviate = -0.66613, p-value = 0.7473
## alternative hypothesis: greater
## sample estimates:
## Moran I statistic       Expectation          Variance
##      -0.061916602      -0.013698630       0.005239556
```

```
geary_result <- geary.test(sf_data$avg_SOD, weights)
print(geary_result)
```

```
##
##  Geary C test under randomisation
##
## data:  sf_data$avg_SOD
## weights: weights
##
## Geary C statistic standard deviate = -1.9389, p-value = 0.9737
## alternative hypothesis: Expectation greater than statistic
## sample estimates:
## Geary C statistic       Expectation          Variance
##       1.172658406       1.000000000       0.007929774
```

Spatial Regression Analysis

Note: Each of the models analyzes if the observations (biomarker values) are not independent of each other but rather influenced by their location and the spatial arrangement.

```
#library(spdep)
#install.packages("sphet")
library(sphet)
```

```
# Spatial Lag Model. Is the value of one point influencing the value of a neighboring point?
#Interpretation - No. No statistically significant result.

library(spatialreg)
#accounts for spatial dependence in the dependent variable
slm_model <- lagsarlm(avg_p450 ~ avg_SOD, data = sf_data, listw = weights)
print(summary(slm_model))
```

```
##
## Call:lagsarlm(formula = avg_p450 ~ avg_SOD, data = sf_data, listw = weights)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -4271767 -1218409  -442173   835360 12194381
##
## Type: lag
## Coefficients: (numerical Hessian approximate standard errors)
##             Estimate Std. Error z value  Pr(>|z|)
## (Intercept)  4675566     834605  5.6021 2.117e-08
## avg_SOD       -68781      27521 -2.4992   0.01245
##
## Rho: 0.12696, LR test value: 0.50819, p-value: 0.47592
## Approximate (numerical Hessian) standard error: 0.17385
##     z-value: 0.73029, p-value: 0.46521
## Wald statistic: 0.53333, p-value: 0.46521
##
## Log likelihood: -1186.116 for lag model
## ML residual variance (sigma squared): 4.8805e+12, (sigma: 2209200)
## Number of observations: 74
## Number of parameters estimated: 4
## AIC: 2380.2, (AIC for lm: 2378.7)
```

```
# Spatial Error Model. This accounts for the spatial autocorrelation in the error term. Could statistic
# No. There is a trend of a mild negative correlation in the data set but it is not statistically signi
sem_model <- errorsarlm(avg_p450 ~ avg_SOD, data = sf_data, listw = weights)
print(summary(sem_model))
```

```
##
## Call:errorsarlm(formula = avg_p450 ~ avg_SOD, data = sf_data, listw = weights)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -4259627 -1212073  -416963   792277 12298927
##
## Type: error
## Coefficients: (asymptotic standard errors)
```

```
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  5269555     413117 12.7556  < 2e-16
## avg_SOD       -70421      26944 -2.6136  0.00896
##
## Lambda: 0.17124, LR test value: 0.87442, p-value: 0.34974
## Approximate (numerical Hessian) standard error: 0.17817
##     z-value: 0.96115, p-value: 0.33648
## Wald statistic: 0.9238, p-value: 0.33648
##
## Log likelihood: -1185.933 for error model
## ML residual variance (sigma squared): 4.8437e+12, (sigma: 2200800)
## Number of observations: 74
## Number of parameters estimated: 4
## AIC: 2379.9, (AIC for lm: 2378.7)
```

```r
# Spatial Durbin Model. An all-in-one model that combines both SLM and SEM. No expected result change.

#install.packages("spatialreg")
library(spatialreg)

#combines SLM and SEM
sdm_model <- spatialreg::lagsarlm(avg_p450 ~ avg_SOD, data = sf_data, listw = weights, type="mixed")
print(summary(sdm_model))
```

```
##
## Call:spatialreg::lagsarlm(formula = avg_p450 ~ avg_SOD, data = sf_data,
##     listw = weights, type = "mixed")
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -4207088 -1182408  -366986   848985 12511496
##
## Type: mixed
## Coefficients: (numerical Hessian approximate standard errors)
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  3960266    1112014  3.5613 0.000369
## avg_SOD       -66343      27412 -2.4202 0.015512
## lag.avg_SOD    46095      48212  0.9561 0.339020
##
## Rho: 0.16963, LR test value: 0.86994, p-value: 0.35097
## Approximate (numerical Hessian) standard error: 0.17709
##     z-value: 0.95791, p-value: 0.33811
## Wald statistic: 0.91759, p-value: 0.33811
##
## Log likelihood: -1185.666 for mixed model
## ML residual variance (sigma squared): 4.8095e+12, (sigma: 2193100)
## Number of observations: 74
## Number of parameters estimated: 5
## AIC: 2381.3, (AIC for lm: 2380.2)
```

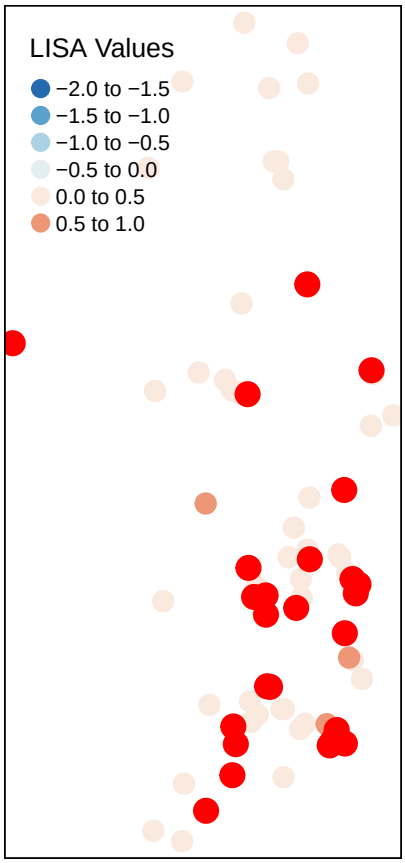
Spatial Autocorrelation Analysis, Local

Local Indicator of Spatial Association (LISA)

```r
# Prep data

#install.packages("spdep")
#install.packages("sf")  # for spatial data handling
#install.packages("tmap")  # for visualization
library(spdep)
library(sf)
library(tmap)
```

```r
lisa_values <- localmoran(sf_data$avg_p450, weights)
```

```r
# Add LISA values and p-values to your spatial data

sf_data$lisa <- lisa_values[,1]  # Local Moran's I values
sf_data$p.value <- lisa_values[,4]  # p-values for significance

# Use tmap for plotting
library(tmap)

# Define breaks for significance levels, e.g., 0.05 for 95% confidence
sig_breaks <- c(0, 0.05, 1)  # Change according to your significance level

# Create a map
tm_shape(sf_data) +
  tm_dots(col = "lisa", size = 0.5, palette = "-RdBu", title = "LISA Values") +
  tm_layout(legend.position = c("left", "top")) +
  tm_shape(sf_data[sf_data$p.value <= 0.05, ]) +  # Add a layer for significant points only
  tm_dots(col = "red", size = 0.7, title = "Significant Clusters") +
  tm_layout(main.title = "LISA Cluster Map", main.title.position = "center")
```

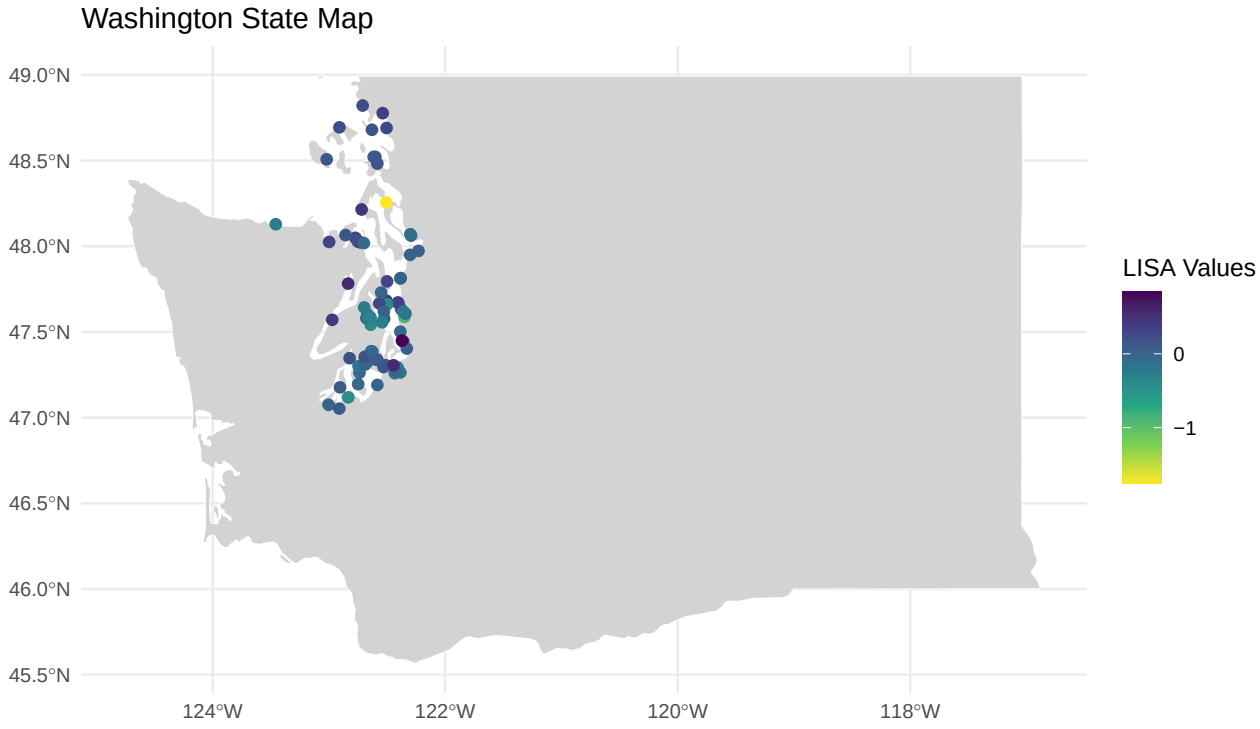



```
# Plot the LISA data over the map from the base Washington State map found in file: 03-map.rmd

library(sf)
library(viridis)
library(rnaturalearth)
library(rnaturalearthdata)
library(ggplot2)
```

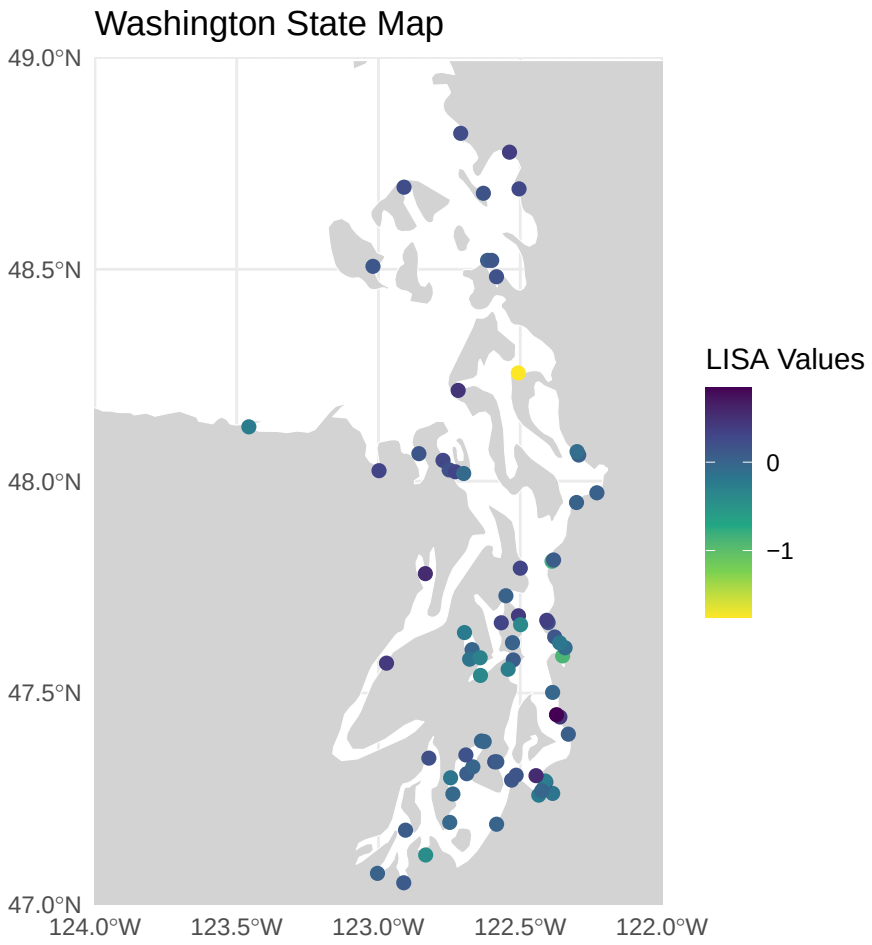
```
world <- ne_states(country = "united states of america", returnclass = "sf")
washington_map <- world[world$name == "Washington", ]

pmap <- ggplot() +
  geom_sf(data = washington_map, fill = "lightgrey", color = "white") +
  theme_minimal() +
  labs(title = "Washington State Map")

# Ensure CRS compatibility
sf_data <- st_transform(sf_data, st_crs(washington_map))

# Prepare the base map
pmap <- ggplot() +
  geom_sf(data = washington_map, fill = "lightgrey", color = "white") +
  theme_minimal() +
```

```
  labs(title = "Washington State Map")

# Add the LISA points layer to the map
complete_map <- pmap +
  geom_sf(data = sf_data, aes(color = lisa), size = 2) +
  scale_color_viridis_c(option = "D", direction = -1, name = "LISA Values")

# Display the combined map
print(complete_map)
```

Interpretation - LISA uses the global application of the Moran's I statistic. This is used to determine any point clustering or significant outliers. There are no sites that are clustered significantly.




```
#zoom in on the puget sound region
# Assuming you know the bounding box coordinates you want to zoom in on
# For example: xmin, xmax, ymin, ymax
xlim <- c(-124, -122)  # longitude bounds
ylim <- c(47, 49)  # latitude bounds

complete_map <- pmap +
  geom_sf(data = sf_data, aes(color = lisa), size = 2) +
  scale_color_viridis_c(option = "D", direction = -1, name = "LISA Values") +
  coord_sf(xlim = xlim, ylim = ylim, expand = FALSE)

print(complete_map)
```



Analytes

```
# Have to merge and reshape for a date frame that lets me compare biomarker and analytes in the models

library(reshape2)
#merge data frames and reshape for input.
colnames(pah)[colnames(pah) == "SiteName"] <- "site_name"
merged_df <- merge(averaged_data, pah, by = c("site_name"), all.x = TRUE)


#reshape to get the analytes into their own columns with the DryValue as their values
reshaped_df <- dcast(merged_df, site_name + site_number +latitude.x + longitude.x + avg_p450 + avg_SOD

print(reshaped_df)
```

```
##                                 site_name site_number latitude.x
## 1                          Aiston Preserve          77   48.67938
## 2                             Arroyo Beach          13   47.50161
## 3                           Blair Waterway          41   47.27568
## 4                         Blair Waterway #2          42   47.26324
## 5                            Brackenwood Ln          23   47.68234
## 6          Broad Spit (Fisherman's Point)          30   47.78184
## 7                  Browns Point Lighthouse          54   47.30515
```

```
## 8                                    Cap Sante         64   48.52097
## 9                                Chambers Bay Park         56   47.19057
## 10 Cherry Point Aquatic Reserve, Conoco Phillips         75   48.82080
## 11                            Chimacum Creek delta          8   48.04906
## 12                       Chuckanut, Clark's Point         74   48.68975
## 13                               Comm Bay Skookum         45   47.29000
## 14                   Comm Bay, Dick Gilmur Launch         48   47.29255
## 15                   Comm Bay, Milwaukee Waterway         49   47.26940
## 16                               Des Moines Marina         21   47.40301
## 17                                   Discovery Bay         12   48.06496
## 18          Donkey Creek Delta (PSS13175 - 000049)         44   47.33775
## 19                                Eagle Harbor Dr         31   47.61889
## 20                          Eastsound, Fishing Bay         60   48.69368
## 21                                   Edmonds Ferry         69   47.81407
## 22                                   Edmonds Marina         68   47.81110
## 23                                       Eld Inlet         72   47.07460
## 24                       Elliot Bay Myrtle Edwards         19   47.61862
## 25             Elliott Bay, Harbor Island, Pier 17         14   47.58766
## 26                                   Everett Harbor         66   47.97260
## 27         Fidalgo Bay Aq Reserve, Weaverling Spit         63   48.48245
## 28                                     Filucy Bay         59   47.19483
## 29                                   Friday Harbor         61   48.50694
## 30                          Gig Harbor Boat Launch         58   47.33785
## 31                                Hood Canal Holly         29   47.57060
## 32                                       Jamestown          2   48.02479
## 33                                  Kingston Marina         36   47.79469
## 34                             Kitsap St Boat Launch         25   47.54167
## 35                              Kopachuck State Park         52   47.30973
## 36                                     Lions Park         37   47.58335
## 37                                   Locust Beach         76   48.77637
## 38                                   Madrona Pont         28   47.57996
## 39                  Manchester, Stormwater Outfall         33   47.55622
## 40                               Maple Hollow Park         51   47.30008
## 41                                Maristone Island         11   48.01813
## 42                  Meyer's Point - Henderson Inlet         71   47.11795
## 43                                   Miller Creek         20   47.44360
## 44                                     Minter Bay         55   47.35397
## 45                                      Mukilteo         65   47.94968
## 46                                   N Avenue Park         62   48.52108
## 47                                   North Camano          7   48.25536
## 48                               Oak Bay County Park         10   48.02218
## 49                                Penn Cove Reference          3   48.21423
## 50                         Pennrose Point State Park         50   47.26190
## 51                              Point Defiance Ferry         43   47.30620
## 52                            Port Angeles Yacht Club          1   48.12823
## 53                                    Purdy - Dexters         57   47.38566
## 54                               Purdy, Burley Lagoon         47   47.38698
## 55                                   Raft Island Park         53   47.32598
## 56                                       Reach Island         39   47.34654
## 57                                       Rich Passage         32   47.57812
## 58                                        Rocky Point         27   47.60255
## 59                                S of Skunk Island          9   48.02667
## 60                         Salmon Bay, Commodore Park         16   47.66630
## 61                                       Salmon Beach         40   47.29464
```

```
## 62                     Seattle Aquarium, Pier 59          17   47.60700
## 63                                 Shilshole Bay          18   47.67168
## 64                    Silverdale, Dyes Inlet          35   47.64279
## 65                                 Skiff Point          24   47.66142
## 66                   Smith Cove, Terminal 91          15   47.63237
## 67                             Squaxin Island          38   47.17650
## 68                  Suquamish Stormwater Outfall          34   47.72961
## 69                           Thea Foss Waterway          46   47.25919
## 70                             Three Tree Point          22   47.44896
## 71                                  Tulalip Bay          67   48.06170
## 72                           Tulalip Reservation          70   48.06979
## 73                                 West Bay Park          73   47.05236
## 74                           Williams Olson Park          26   47.66586
##    longitude.x avg_p450   avg_SOD     SumPAHs  SumPAHs16 SumPAHs42_DMNcorrected
## 1    -122.6301  3752618  3.440125    97.40305   31.65599               97.40305
## 2    -122.3859  4480860  8.832583   408.22636  168.09321              408.22636
## 3    -122.4173  4879642  6.517750   715.09344  247.53234              715.09344
## 4    -122.3857  3714918 10.796000  1038.61553  299.36565             1038.61553
## 5    -122.5064  1857012  9.835125   389.36704  176.98502              389.36704
## 6    -122.8347  2311731  7.116250          NA         NA                     NA
## 7    -122.4444  7401612  1.378875   438.80107  175.52043              438.80107
## 8    -122.6007  3728576 41.120250   375.61898   63.66423              375.61898
## 9    -122.5835  4502338  9.687875   182.85775   70.78365              182.85775
## 10   -122.7101  3716209  4.467750   325.98083  126.77032              325.98083
## 11   -122.7724  2641574  2.193375   168.45933   68.63158              168.45933
## 12   -122.5043  3006724 12.702500   206.57357   80.33416              206.57357
## 13   -122.4100  5919817  8.918500   857.31981  342.92792              857.31981
## 14   -122.4122  3792068 41.297750   781.27358  332.04127              781.27358
## 15   -122.4243  7608742  6.397700   599.01078  195.18329              599.01078
## 16   -122.3306  4886368  9.639750 14700.69093 9800.46062            15313.21971
## 17   -122.8575  4029898  8.739125   235.35381  117.67691              235.35381
## 18   -122.5901  3185175  2.500625          NA         NA                     NA
## 19   -122.5275  1846201 25.167500  1031.19294  577.46805             1031.19294
## 20   -122.9099  2504169 11.961125   512.49588  292.06755              518.00659
## 21   -122.3825  4710937  0.206375   532.00981  268.92804              526.16355
## 22   -122.3880 16253739 16.624750   702.84765  374.85208              702.84765
## 23   -123.0035  4644881 20.791000   171.59812   64.34930              171.59812
## 24   -122.3611  2469548 34.825200 12078.42170 8052.28113            12078.42170
## 25   -122.3507  9194072  5.563250  3195.16749 1882.86656             3195.16749
## 26   -122.2298  5271456  0.375875   372.81540  173.09287              372.81540
## 27   -122.5839  3345616 18.015000   699.56200  165.35102              699.56200
## 28   -122.7487  6265040  5.313125   148.27311   65.24017              148.27311
## 29   -123.0194  3954243  6.709600   532.69368  338.41716              532.69368
## 30   -122.5828  4097010  7.509375   418.64682  233.94969              418.64682
## 31   -122.9717   788472  1.640417   116.10138   24.58618               88.78341
## 32   -122.9981  3263515 24.487333   119.31910   31.15554              119.31910
## 33   -122.4999  5281103  4.064000   273.01097  123.50496              273.01097
## 34   -122.6403  2885042 30.745000   758.67818  303.47127              758.67818
## 35   -122.6887  7285132  4.726000   141.69495   59.37693              141.69495
## 36   -122.6415  6543254 14.793250   289.59279  118.46978              289.59279
## 37   -122.5379  2401828  5.401500   147.20503   46.08157              147.20503
## 38   -122.6786  5804701  3.871500   259.13965  108.47706              259.13965
## 39   -122.5428  6122830  2.682375   369.42136  171.02841              369.42136
## 40   -122.7457  3639570  4.621500   176.66667   71.97531              176.66667
```

```
## 41   -122.6995  4772561  5.678750   134.05496   41.62759              134.05496
## 42   -122.8336  2519276 24.012500   200.73694   66.16884              200.73694
## 43   -122.3603  5747266  4.015375   292.31683  114.38485              285.96212
## 44   -122.6917  6206749  1.991375   164.47704   78.94898              164.47704
## 45   -122.3016  5076323  7.606500   440.19763  207.15183              440.19763
## 46   -122.6153  4183302  1.443625  2333.52357  976.82382             2333.52357
## 47   -122.5070 12290521  0.751750   165.04623   63.47932              165.04623
## 48   -122.7287  2335145 19.807750   141.06998   47.45081              141.06998
## 49   -122.7190  2427656 23.860000   160.93092   64.37237              160.93092
## 50   -122.7376  4193870  4.499500          NA         NA                     NA
## 51   -122.5145  2181857 10.884250   694.82398  335.83159              694.82398
## 52   -123.4571  5751355  7.385500  1983.40206 1081.85567             1983.40206
## 53   -122.6273  4743078  1.760875   149.15677   68.84159              149.15677
## 54   -122.6367  3937679 11.214000   215.42133   86.16853              215.42133
## 55   -122.6672  4523052  3.778125   147.74866   70.35651              147.74866
## 56   -122.8224  5905498  4.583750   192.65779   81.73361              192.65779
## 57   -122.5248  3203053  2.245000   338.95875  142.71947              338.95875
## 58   -122.6699  4508401  9.895500   310.52017  133.94988              310.52017
## 59   -122.7508  3556923 11.307000   294.28413  144.13917              294.28413
## 60   -122.4018  3756618  1.481750  1041.76471  409.26471             1041.76471
## 61   -122.5305  1904019 18.330875   428.65764  214.32882              428.65764
## 62   -122.3420  3135234 23.887625  3304.96063 1879.29134             3304.96063
## 63   -122.4067  2862824  9.906750  2029.62742 1052.39940             1954.45604
## 64   -122.6967  1652091 26.031000   363.26584  137.08145              370.11991
## 65   -122.4988  5664099  4.236250   449.87987  213.10099              449.87987
## 66   -122.3787  3958924 19.661250  7153.27706 4161.90665             7153.27706
## 67   -122.9046  6368736  5.858750   174.78708   71.20955              174.78708
## 68   -122.5506  4338556  5.819250          NA         NA                     NA
## 69   -122.4347  3423038 18.690800  1103.39463  454.33897             1103.39463
## 70   -122.3723  7417010  8.125125   312.23111  129.60537              312.23111
## 71   -122.2939  5544824  6.253833   356.89176  181.69035              356.89176
## 72   -122.2998  3773556  3.182125   253.49047  123.79767              253.49047
## 73   -122.9109  5731487  2.996125   391.41916  161.96655              384.67055
## 74   -122.5670  2970859  6.457625   301.25306  150.62653              301.25306
##    SumPAHsHMW SumPAHsLMW NA
## 1    19.48061   79.13998 NA
## 2   186.10320  216.11984 NA
## 3   280.53666  456.55966 NA
## 4   403.22721  610.95031 NA
## 5   176.98502  212.38202 NA
## 6          NA         NA NA
## 7   181.37111  257.42996 NA
## 8   120.96204  254.65693 NA
## 9    76.68228  106.17547 NA
## 10  150.91705  169.02709 NA
## 11   56.77703  112.30622 NA
## 12   74.59601  131.97756 NA
## 13  369.30699  461.63374 NA
## 14  338.55189  442.72170 NA
## 15  215.37466  390.36658 NA
## 16 8575.40304 6125.28789 NA
## 17   98.84860  136.50521 NA
## 18         NA         NA NA
## 19  653.08886  364.35484 NA
```

```
## 20  270.02471  242.47117 NA
## 21  268.92804  257.23551 NA
## 22  374.85208  304.56731 NA
## 23   75.07418   96.52394 NA
## 24 9394.32799 3220.91245 NA
## 25 2054.03624 1141.13125 NA
## 26  166.43545  206.37995 NA
## 27  254.38618  438.81616 NA
## 28   59.30925   88.96387 NA
## 29  319.61621  206.81049 NA
## 30  246.26283  172.38398 NA
## 31   11.61014  109.27189 NA
## 32   15.24633  106.06143 NA
## 33  123.50496  149.50601 NA
## 34  354.04982  419.80193 NA
## 35   53.97903   87.71592 NA
## 36  131.63309  157.95970 NA
## 37   27.52094  121.60416 NA
## 38  120.53007  138.60958 NA
## 39  171.02841  198.39295 NA
## 40   64.77778  111.23457 NA
## 41   23.98878  112.88839 NA
## 42   52.78638  148.69403 NA
## 43  120.73956  171.57727 NA
## 44   65.79082   98.68622 NA
## 45  194.20484  245.99279 NA
## 46 1085.35980 1248.16377 NA
## 47   54.59221  107.91484 NA
## 48   33.98504  102.59635 NA
## 49   46.81627  109.72563 NA
## 50         NA         NA NA
## 51  335.83159  370.57279 NA
## 52 1202.06186  781.34021 NA
## 53   57.36799   91.78878 NA
## 54   86.16853  129.25280 NA
## 55   61.91373   91.46346 NA
## 56   81.73361  110.92418 NA
## 57  142.71947  190.29263 NA
## 58  133.94988  176.57029 NA
## 59  132.12757  162.15656 NA
## 60  535.76471  476.23529 NA
## 61  226.23598  202.42166 NA
## 62 2203.30709 1101.65354 NA
## 63 1277.91356  714.12817 NA
## 64  150.78959  212.47624 NA
## 65  213.10099  236.77888 NA
## 66 4682.14498 2145.98312 NA
## 67   77.68315   97.10393 NA
## 68         NA         NA NA
## 69  545.20676  584.15010 NA
## 70  141.38767  176.73459 NA
## 71  168.71247  188.17929 NA
## 72  117.90254  135.58792 NA
## 73  202.45819  182.21237 NA
```

```
## 74  156.65159  144.60147 NA
```

Linear Regression - only here to show that the non- normal data passed through this model is unsuccessful.

```
# Linear Regression- p450 as response
lm_model <- lm(avg_p450 ~ avg_SOD + SumPAHs + SumPAHs16 + SumPAHs42_DMNcorrected + SumPAHsHMW + SumPAHsL
print(summary(lm_model))
```

```
##
## Call:
## lm(formula = avg_p450 ~ avg_SOD + SumPAHs + SumPAHs16 + SumPAHs42_DMNcorrected +
##     SumPAHsHMW + SumPAHsLMW, data = reshaped_df)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -4234121 -1358333  -398182   761952 12118565
##
## Coefficients:
##                           Estimate Std. Error t value Pr(>|t|)
## (Intercept)             5420352.5   579886.3   9.347 1.63e-13 ***
## avg_SOD                  -73508.2    32849.9  -2.238   0.0288 *
## SumPAHs                   -5306.1    28666.3  -0.185   0.8537
## SumPAHs16                 -1482.1     8273.8  -0.179   0.8584
## SumPAHs42_DMNcorrected     7481.0    30613.3   0.244   0.8077
## SumPAHsHMW                 -674.5     6797.7  -0.099   0.9213
## SumPAHsLMW                -2572.0     6519.6  -0.395   0.6945
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2366000 on 63 degrees of freedom
##   (4 observations deleted due to missingness)
## Multiple R-squared:  0.08903,    Adjusted R-squared:  0.002275
## F-statistic: 1.026 on 6 and 63 DF,  p-value: 0.4167
```

Generalized Linear Model - used for non- normal data.

```
# GLM - p450 as response
glm_model <- glm(avg_p450 ~ avg_SOD + SumPAHs + SumPAHs16 + SumPAHs42_DMNcorrected + SumPAHsHMW + SumPAH
print(summary(glm_model))
```

Interpretation - All categories of PAH's show a statistically significant relationship to the p450 biomarker result.

```
##
## Call:
## glm(formula = avg_p450 ~ avg_SOD + SumPAHs + SumPAHs16 + SumPAHs42_DMNcorrected +
##     SumPAHsHMW + SumPAHsLMW, family = poisson(), data = reshaped_df)
##
## Coefficients:
##                            Estimate Std. Error  z value Pr(>|z|)
## (Intercept)              1.553e+01  1.140e-04 136245.3   <2e-16 ***
## avg_SOD                 -1.811e-02  7.311e-06  -2476.4   <2e-16 ***
## SumPAHs                 -1.048e-03  5.764e-06   -181.7   <2e-16 ***
```

```
## SumPAHs16              -3.171e-04  1.656e-06   -191.5   <2e-16 ***
## SumPAHs42_DMNcorrected  1.646e-03  6.132e-06    268.5   <2e-16 ***
## SumPAHsHMW             -2.747e-04  1.411e-06   -194.7   <2e-16 ***
## SumPAHsLMW             -6.914e-04  1.302e-06   -531.0   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for poisson family taken to be 1)
##
##     Null deviance: 71801382  on 69  degrees of freedom
## Residual deviance: 63582448  on 63  degrees of freedom
##   (4 observations deleted due to missingness)
## AIC: Inf
##
## Number of Fisher Scoring iterations: 5
```

```r
# LR- p450 as response
#### Replicating with individual analytes to see if anything really sticks out.

#merge and reshape just like above
colnames(indv)[colnames(indv) == "SiteName"] <- "site_name"
merged_df2 <- merge(averaged_data, indv, by = c("site_name"), all.x = TRUE)


#reshape to get the analytes into their own columns with the DryValue as their values
reshaped2_df <- dcast(merged_df2, site_name + site_number +latitude + longitude + avg_p450 + avg_SOD ~ A

print(reshaped2_df)
```

```
##                                       site_name site_number latitude longitude
## 1                                Aiston Preserve          77 48.67938 -122.6301
## 2                                   Arroyo Beach          13 47.50161 -122.3859
## 3                                  Blair Waterway          41 47.27568 -122.4173
## 4                                Blair Waterway #2          42 47.26324 -122.3857
## 5                                  Brackenwood Ln          23 47.68234 -122.5064
## 6                  Broad Spit (Fisherman's Point)          30 47.78184 -122.8347
## 7                         Browns Point Lighthouse          54 47.30515 -122.4444
## 8                                       Cap Sante          64 48.52097 -122.6007
## 9                                Chambers Bay Park          56 47.19057 -122.5835
## 10 Cherry Point Aquatic Reserve, Conoco Phillips          75 48.82080 -122.7101
## 11                           Chimacum Creek delta           8 48.04906 -122.7724
## 12                       Chuckanut, Clark's Point          74 48.68975 -122.5043
## 13                              Comm Bay Skookum          45 47.29000 -122.4100
## 14                  Comm Bay, Dick Gilmur Launch          48 47.29255 -122.4122
## 15                  Comm Bay, Milwaukee Waterway          49 47.26940 -122.4243
## 16                               Des Moines Marina          21 47.40301 -122.3306
## 17                                  Discovery Bay          12 48.06496 -122.8575
## 18        Donkey Creek Delta (PSS13175 - 000049)          44 47.33775 -122.5901
## 19                                Eagle Harbor Dr          31 47.61889 -122.5275
## 20                         Eastsound, Fishing Bay          60 48.69368 -122.9099
## 21                                  Edmonds Ferry          69 47.81407 -122.3825
## 22                                  Edmonds Marina          68 47.81110 -122.3880
## 23                                       Eld Inlet          72 47.07460 -123.0035
## 24                     Elliot Bay Myrtle Edwards          19 47.61862 -122.3611
## 25           Elliott Bay, Harbor Island, Pier 17          14 47.58766 -122.3507
```

```
## 26                                Everett Harbor          66 47.97260 -122.2298
## 27      Fidalgo Bay Aq Reserve, Weaverling Spit          63 48.48245 -122.5839
## 28                                  Filucy Bay          59 47.19483 -122.7487
## 29                                Friday Harbor          61 48.50694 -123.0194
## 30                       Gig Harbor Boat Launch          58 47.33785 -122.5828
## 31                            Hood Canal Holly          29 47.57060 -122.9717
## 32                                  Jamestown           2 48.02479 -122.9981
## 33                               Kingston Marina          36 47.79469 -122.4999
## 34                         Kitsap St Boat Launch          25 47.54167 -122.6403
## 35                          Kopachuck State Park          52 47.30973 -122.6887
## 36                                  Lions Park          37 47.58335 -122.6415
## 37                               Locust Beach          76 48.77637 -122.5379
## 38                               Madrona Pont          28 47.57996 -122.6786
## 39               Manchester, Stormwater Outfall          33 47.55622 -122.5428
## 40                           Maple Hollow Park          51 47.30008 -122.7457
## 41                            Maristone Island          11 48.01813 -122.6995
## 42              Meyer's Point - Henderson Inlet          71 47.11795 -122.8336
## 43                               Miller Creek          20 47.44360 -122.3603
## 44                                 Minter Bay          55 47.35397 -122.6917
## 45                                   Mukilteo          65 47.94968 -122.3016
## 46                               N Avenue Park          62 48.52108 -122.6153
## 47                               North Camano           7 48.25536 -122.5070
## 48                          Oak Bay County Park          10 48.02218 -122.7287
## 49                          Penn Cove Reference           3 48.21423 -122.7190
## 50                    Pennrose Point State Park          50 47.26190 -122.7376
## 51                         Point Defiance Ferry          43 47.30620 -122.5145
## 52                    Port Angeles Yacht Club           1 48.12823 -123.4571
## 53                            Purdy - Dexters          57 47.38566 -122.6273
## 54                       Purdy, Burley Lagoon          47 47.38698 -122.6367
## 55                           Raft Island Park          53 47.32598 -122.6672
## 56                               Reach Island          39 47.34654 -122.8224
## 57                               Rich Passage          32 47.57812 -122.5248
## 58                                Rocky Point          27 47.60255 -122.6699
## 59                           S of Skunk Island           9 48.02667 -122.7508
## 60                  Salmon Bay, Commodore Park          16 47.66630 -122.4018
## 61                               Salmon Beach          40 47.29464 -122.5305
## 62                  Seattle Aquarium, Pier 59          17 47.60700 -122.3420
## 63                               Shilshole Bay          18 47.67168 -122.4067
## 64                  Silverdale, Dyes Inlet          35 47.64279 -122.6967
## 65                               Skiff Point          24 47.66142 -122.4988
## 66                  Smith Cove, Terminal 91          15 47.63237 -122.3787
## 67                           Squaxin Island          38 47.17650 -122.9046
## 68                Suquamish Stormwater Outfall          34 47.72961 -122.5506
## 69                           Thea Foss Waterway          46 47.25919 -122.4347
## 70                            Three Tree Point          22 47.44896 -122.3723
## 71                                 Tulalip Bay          67 48.06170 -122.2939
## 72                          Tulalip Reservation          70 48.06979 -122.2998
## 73                               West Bay Park          73 47.05236 -122.9109
## 74                          Williams Olson Park          26 47.66586 -122.5670
##    avg_p450   avg_SOD acenaphthene acenaphthylene anthracene benz[a]anthracene
## 1   3752618  3.440125    -3.165599      -3.165599  -3.165599         -3.104722
## 2   4480860  8.832583    -3.902164      -3.902164  -3.902164         13.207324
## 3   4879642  6.517750     5.500719      -2.915381   6.050791         17.052228
## 4   3714918 10.796000     7.942354      -4.948698   5.315268         17.717559
```

```
## 5   1857012  9.835125    -3.598695      -3.598695  -3.598695         12.978901
## 6   2311731  7.116250           NA             NA         NA                NA
## 7   7401612  1.378875     4.212490      -3.334888   5.031586         12.286430
## 8   3728576 41.120250    -4.774818      -4.711153  -4.774818          4.902146
## 9   4502338  9.687875    -2.713373      -2.713373  -2.713373          5.426746
## 10  3716209  4.467750    -3.682376      -3.622009  -3.622009          9.055023
## 11  2641574  2.193375    -3.868325      -3.868325  -3.868325          3.493971
## 12  3006724 12.702500    -3.155985      -3.098603  -3.155985          6.311970
## 13  5919817  8.918500     7.913721      -5.671500   7.913721         31.654885
## 14  3792068 41.297750     8.463797      -4.752748   9.765920         20.182901
## 15  7608742  6.397700     6.124717      -4.307493   4.913235         12.787871
## 16  4886368  9.639750   257.262091       5.083989 287.888531        673.781667
## 17  4029898  8.739125    -1.976972      -1.929901   2.541821          9.414152
## 18  3185175  2.500625           NA             NA         NA                NA
## 19  1846201 25.167500     6.393396      -4.331010   6.874620         33.685636
## 20  2504169 11.961125    -3.747282      -3.747282   4.573888         20.389621
## 21  4710937  0.206375    -4.267771      -4.267771   6.430888         23.385047
## 22 16253739 16.624750     4.275657      -3.397097   9.957008         33.385263
## 23  4644881 20.791000    -2.252225      -2.252225  -2.252225          2.466723
## 24  2469548 34.825200    54.352898      -5.099778  80.522811        597.210851
## 25  9194072  5.563250    24.534322      -3.594563  51.350906        165.464031
## 26  5271456  0.375875    -5.126212      -5.126212  -5.126212         11.983352
## 27  3345616 18.015000    -6.168865      -6.105268  -6.168865          7.631585
## 28  6265040  5.313125    -3.439936      -3.380627  -3.439936          3.677173
## 29  3954243  6.709600     3.509511      -2.882813   6.204315         19.427652
## 30  4097010  7.509375    -3.878640      -3.878640  -3.878640         12.928799
## 31   788472  1.640417    -3.824516      -3.824516  -3.824516         -3.961106
## 32  3263515 24.487333    -4.640187      -4.573899  -4.640187         -4.905341
## 33  5281103  4.064000    -3.445138      -3.445138  -3.445138          8.450339
## 34  2885042 30.745000     4.046284      -2.680663   5.563640         18.714062
## 35  7285132  4.726000    -3.238742      -3.171268  -3.171268          3.171268
## 36  6543254 14.793250    -4.541341      -4.541341  -4.541341          7.897985
## 37  2401828  5.401500    -4.352149      -4.352149  -4.352149         -4.416151
## 38  5804701  3.871500    -3.615902      -3.555637  -3.555637          4.760938
## 39  6122830  2.682375    -3.967859      -3.967859   4.788795         12.998159
## 40  3639570  4.621500    -3.337037      -3.271605  -3.271605          4.318519
## 41  4772561  5.678750    -4.797757      -4.797757  -4.797757         -5.150533
## 42  2519276 24.012500    -5.724720      -5.724720  -5.724720         -5.724720
## 43  5747266  4.015375    -3.622187      -3.622187  -3.622187          9.532071
## 44  6206749  1.991375    -3.750077      -3.684286  -3.684286          4.079031
## 45  5076323  7.606500    -4.401976      -4.337241   7.768194         14.889038
## 46  4183302  1.443625     8.682878      -3.310347  17.908437         70.548387
## 47 12290521  0.751750    -3.554842      -3.491363  -3.491363         -3.047007
## 48  2335145 19.807750    -4.232099      -4.167977  -4.232099         -4.488590
## 49  2427656 23.860000    -3.291769      -3.291769  -3.291769         -3.291769
## 50  4193870  4.499500           NA             NA         NA                NA
## 51  2181857 10.884250     5.326984      -3.358316  11.580400         21.423739
## 52  5751355  7.385500    10.818557      -3.005155  32.455670         78.134021
## 53  4743078  1.760875    -3.384711      -3.327343  -3.327343          3.384711
## 54  3937679 11.214000    -4.123780      -4.062231  -4.062231          4.554623
## 55  4523052  3.778125    -3.728895      -3.728895  -3.728895          3.799251
## 56  5905498  4.583750    -3.035820      -2.977439  -3.035820          4.261824
## 57  3203053  2.245000    -3.627453      -3.627453  -3.627453         10.703960
## 58  4508401  9.895500    -2.983429      -2.922543  -2.922543          6.088631
```

```
## 59  3556923 11.307000    -3.543421      -3.543421  -3.543421          7.206958
## 60  3756618  1.481750    -5.655294      -5.655294   5.729706         29.764706
## 61  1904019 18.330875    -5.179613      -5.120077  -5.120077         17.860735
## 62  3135234 23.887625    14.904724      -4.341811  57.674803        162.007874
## 63  2862824  9.906750    12.027422      -5.261997  11.275708        105.239940
## 64  1652091 26.031000    -3.564118      -3.564118  -3.564118          6.031584
## 65  5664099  4.236250    -5.682693      -5.682693  -5.682693         17.166469
## 66  3958924 19.661250    71.532771       3.511609  84.538729        273.125124
## 67  6368736  5.858750    -2.589438      -2.589438  -2.589438          3.689949
## 68  4338556  5.819250           NA             NA         NA                NA
## 69  3423038 18.690800    12.981113      -5.062634  11.683002         29.856561
## 70  7417010  8.125125    -4.123807      -4.123807  -4.123807         10.604076
## 71  5544824  6.253833    -4.477369      -4.412480  -4.477369         11.031200
## 72  3773556  3.182125    -3.713930      -3.654979  -3.654979          6.484640
## 73  5731487  2.996125     6.073746      -4.656538  -4.656538          9.448049
## 74  2970859  6.457625    -3.494535      -3.494535  -3.494535          7.832579
##    benzo[a]pyrene benzo[b]fluoranthene benzo[e]pyrene benzo[ghi]perylene
## 1       -3.104722            -3.104722      -3.104722          -3.104722
## 2       -3.001664            11.406325       8.404660          -2.941631
## 3        2.915381             9.901294      12.101581           2.915381
## 4       -5.009793            17.106609      26.270863           6.720453
## 5       -2.713770             9.439201       6.489451          -2.713770
## 6              NA                   NA             NA                 NA
## 7       -3.217875             7.605885       7.605885          -3.159368
## 8       -4.838482            -4.838482      -4.838482          -4.774818
## 9       -2.536414             5.308773       2.890332          -2.536414
## 10      -3.501276             5.312280       4.768979          -3.440909
## 11      -3.119617             3.182010      -3.057225          -3.057225
## 12      -2.467406             4.533142      -2.467406          -2.410025
## 13      -5.671500            17.146396      17.805873          -5.605553
## 14      -4.622535            11.068042      18.229717          -4.557429
## 15      -4.509407             7.403504       8.749596          -4.442102
## 16     165.382773           483.897743     349.141410          34.914141
## 17      -1.553335             7.060614       2.588892          -1.506264
## 18             NA                   NA             NA                 NA
## 19      15.124163            61.871576      43.310103          11.686853
## 20      -2.755354            16.532125      11.021417          -2.700247
## 21       4.326234            16.369533      11.107897          -4.150846
## 22       8.199889            22.256842      14.642659           3.338526
## 23      -1.769606             4.289953       2.895718          -1.769606
## 24     174.466091           442.875462     456.295931          48.313687
## 25      46.215815           148.347062     108.407468          18.828666
## 26      -5.192786             6.524270       5.592231          -5.126212
## 27      -6.359655             7.631585       8.903516          -6.359655
## 28      -3.380627             3.855101      -3.380627          -3.380627
## 29       3.196162            15.040763       9.400477          -2.318784
## 30       4.494297            16.007084      12.313142           5.233085
## 31      -4.029401            -3.961106      -3.961106          -3.961106
## 32      -4.971629            -4.971629      -4.905341          -4.905341
## 33      -3.185128             7.800313       5.070204          -3.185128
## 34       6.069425            25.795058      18.714062           6.069425
## 35      -3.036320             3.373689      -3.036320          -3.036320
## 36      -4.541341             9.214316       6.581654          -4.475525
## 37      -4.480153            -4.480153      -4.416151          -4.416151
```

```
## 38      -3.676167            9.642406       7.834455          -3.615902
## 39      -3.625802           10.261704       7.525250          -3.557391
## 40      -3.206173            4.907407      -3.140741          -3.140741
## 41      -5.221088           -5.221088      -5.150533          -5.150533
## 42      -5.724720           -5.724720      -5.724720          -5.650373
## 43      -3.431545            7.625656       4.194111          -3.367998
## 44      -3.618495            4.079031      -3.618495          -3.552704
## 45      -4.401976            9.062892       6.473495          -4.337241
## 46      17.365757           81.401985      59.694789          12.481638
## 47      -3.047007           -3.047007      -3.047007          -3.047007
## 48      -4.488590           -4.488590      -4.488590          -4.424468
## 49      -3.291769           -3.291769      -3.291769          -3.218618
## 50             NA                  NA             NA                 NA
## 51       5.790200           17.370599      13.317460           3.705728
## 52      15.626804           66.113402      52.289691           7.212371
## 53      -3.269975            3.958391      -3.269975          -3.212607
## 54      -4.369976            5.724053      -4.308427          -4.308427
## 55      -3.588182            4.221390      -3.588182          -3.588182
## 56      -2.335246            6.421926       3.386107          -2.276865
## 57      -3.686920            9.514631       5.233047          -3.627453
## 58      -2.191907           10.350672       7.915220           3.348747
## 59      -2.702609            7.807538       4.984813          -2.642551
## 60       9.673529           29.764706      38.694118          11.905882
## 61       4.107969           17.860735       8.930368          -3.750754
## 62      62.859055          142.566929     116.645669          23.329134
## 63      26.309985           75.171386      75.171386          15.034277
## 64      -3.289955            8.224887       8.910294           3.495577
## 65      -4.853967           14.206733       8.287261          -4.794772
## 66     104.047666          240.610228     273.125124          36.416683
## 67      -2.006815            4.207837       3.042590          -1.942079
## 68             NA                  NA             NA                 NA
## 69       8.437724           25.313171      27.260338           7.788668
## 70      -4.241630            8.247614       5.478772          -4.182719
## 71      -4.477369            6.488941       5.191153          -4.412480
## 72      -3.654979            5.541419       4.244492          -3.596028
## 73      -4.791510           10.122909       7.423467          -4.724024
## 74      -2.711278           10.242604       7.832579          -2.651027
##    benzo[k]fluoranthene C1-benzanthracenes/chrysenes C1-dibenzothiophenes
## 1             -3.104722                   -2.252446           -3.104722
## 2             13.207324                    9.605326           -3.842130
## 3             10.451366                   12.651653            3.905510
## 4             16.495658                   28.103714           -4.887603
## 5             11.799001                    7.079401           -3.539700
## 6                    NA                          NA                  NA
## 7              9.946158                    9.946158           -3.276381
## 8             -4.902146                    7.639708           -4.711153
## 9              5.131814                    3.893101           -2.713373
## 10             5.795215                    5.855581           -3.622009
## 11             3.993110                   -2.495694           -3.868325
## 12             5.393865                    3.557656           -3.098603
## 13            19.124826                   21.762734           -5.605553
## 14            13.021226                   16.276533           -4.752748
## 15             7.403504                   14.807008           -4.307493
## 16           343.016122                  202.134500           85.754030
```

```
## 17          7.060614          3.200812         -1.929901
## 18                NA                NA                NA
## 19         75.620816         28.185940         -4.331010
## 20         21.491763          9.919275         -3.692175
## 21         14.615654          9.354019         -4.209308
## 22         23.428255         14.056953         -3.397097
## 23          5.147944          2.949343         -2.252225
## 24        543.528977        328.801480         43.616523
## 25        148.347062         85.584843         10.270181
## 26          9.986127          6.657418         -5.126212
## 27          6.995620         13.991240         -6.105268
## 28          5.041286          3.439936         -3.380627
## 29         17.547557          7.520381         -2.882813
## 30         23.394969          9.850513         -3.878640
## 31         -4.029401         -2.868387         -3.824516
## 32         -4.971629         -3.446996         -4.573899
## 33          9.100366          5.590225         -3.380136
## 34         27.818200         15.173564          2.781820
## 35          4.858112          3.103794         -3.171268
## 36          9.872481          7.239820         -4.475525
## 37         -4.480153         -3.200109         -4.352149
## 38         10.245056          6.629154         -3.555637
## 39         11.629932          8.209363         -3.899448
## 40          6.085185          3.795062         -3.271605
## 41         -5.221088         -3.598317         -4.727201
## 42         -5.799067         -3.940392         -5.650373
## 43          8.261128          6.990185         -3.558640
## 44          6.579082          3.618495         -3.684286
## 45         12.299640          8.415543         -4.337241
## 46         75.975186         40.700993          7.054839
## 47          3.935718         -2.475693         -3.491363
## 48         -4.488590         -3.077890         -4.167977
## 49          4.023273         -2.340813         -3.218618
## 50                NA                NA                NA
## 51         19.107659         14.475500          3.474120
## 52         72.123711         45.678351          5.349175
## 53          6.310479          3.614183         -3.327343
## 54          6.154895          4.800818         -4.062231
## 55          5.909947         -2.673547         -3.728895
## 56          7.589549          3.561250         -2.977439
## 57         10.109296          7.730638         -3.567987
## 58         11.568399          6.027744         -2.922543
## 59          9.008698          4.564407         -3.543421
## 60         39.438235         29.020588         -5.655294
## 61         22.028240         10.121083         -5.120077
## 62        155.527559        103.685039         10.368504
## 63         97.722802         60.888823          5.412340
## 64          9.595701          8.224887         -3.564118
## 65         15.390627          9.471155         -5.623498
## 66        292.634062        214.598312         22.110129
## 67          5.308348          3.172062         -2.589438
## 68                NA                NA                NA
## 69         27.260338         29.856561         -5.062634
## 70          8.836730          7.658499         -4.064896
```

```
## 71             8.435624                   5.840047           -4.412480
## 72             7.663665                   4.185540           -3.654979
## 73            12.147491                  11.472631           -4.656538
## 74            11.447616                   5.543056           -3.494535
##    C1-fluoranthenes/pyrenes C1-fluorenes C1-naphthalenes
## 1                 -3.104722   -3.104722       -3.104722
## 2                 15.008322    3.902164       -6.003329
## 3                 22.002875    5.115668       -3.630474
## 4                 28.714665    6.109503       -4.887603
## 5                 15.338702   -3.539700       -5.309551
## 6                        NA          NA              NA
## 7                 15.796839   -3.276381        3.568915
## 8                 10.186277   -4.711153       -4.838482
## 9                  6.488501   -2.713373        2.949319
## 10                10.262359   -3.622009       -3.501276
## 11                 5.864880   -3.868325       -4.554641
## 12                 6.885786   -3.098603       -3.844564
## 13                32.314362    6.001239       -5.539605
## 14                29.297759    5.403809        6.510613
## 15                19.518329    4.644016        5.451671
## 16               673.781667  104.129894       61.252879
## 17                 8.472737   -1.929901       -2.118184
## 18                       NA          NA              NA
## 19                39.185332   -4.331010        4.468503
## 20                19.287479   -3.692175       -7.714992
## 21                22.800421   -4.209308       -4.033921
## 22                31.042438    4.158515       -3.279956
## 23                 5.308817   -2.252225       -2.573972
## 24               738.125771   35.564242       20.801726
## 25               159.758375   13.123009        6.276222
## 26                13.980578   -5.126212       -4.926489
## 27                18.442998   -6.105268       -6.168865
## 28                 5.634378   -3.380627       -3.321318
## 29                23.814541   -2.882813       -2.945483
## 30                16.622741   -3.878640       -3.755508
## 31                -3.824516   -3.824516       -3.756221
## 32                -4.573899   -4.573899       -4.573899
## 33                11.050444   -3.380136        3.380136
## 34                27.312415    5.057855       -3.388763
## 35                 5.262955   -3.171268       -3.103794
## 36                10.530647   -4.475525       -4.475525
## 37                -4.352149   -4.352149       -4.416151
## 38                 8.437105   -3.555637       -3.555637
## 39                14.366386   -3.899448        4.104682
## 40                 5.954321   -3.271605        3.598765
## 41                -4.727201   -4.727201       -4.727201
## 42                 6.616884   -5.650373       -6.170802
## 43                 9.532071   -3.558640        3.685734
## 44                 5.921173   -3.684286       -3.618495
## 45                18.125785    4.401976       -4.143037
## 46                86.828784   14.652357       -5.318263
## 47                 4.760949   -3.491363       -3.618321
## 48                 4.937449   -4.167977       -4.167977
## 49                 5.925184   -3.218618       -3.291769
```

```
## 50                      NA          NA             NA
## 51               25.476879    5.095376       4.226846
## 52               90.154639    6.611340      -4.928454
## 53                5.105751   -3.327343      -3.269975
## 54                6.770385   -4.062231      -4.185329
## 55                5.276738   -3.728895      -3.588182
## 56                6.421926   -2.977439      -3.327725
## 57               12.487954   -3.567987       4.400517
## 58                3.957610   -2.922543      -4.140269
## 59               11.411017   -3.543421      -5.525335
## 60               32.741176   -5.655294      -5.432059
## 61               16.670019   -5.120077     -12.502515
## 62              174.968504   13.608661       4.536220
## 63               90.205663    7.517139       6.464739
## 64               10.281109   -3.564118       4.043903
## 65               18.350363   -5.623498      -8.879208
## 66              390.178749   29.913704      16.907746
## 67                5.437820   -2.589438      -3.431006
## 68                      NA          NA             NA
## 69               41.539563    7.139612      -4.932823
## 70               11.782306   -4.064896      -4.241630
## 71               14.275671   -4.412480      -4.282701
## 72                9.432203   -3.654979      -4.185540
## 73               14.172073   -4.656538      -4.589052
## 74               11.447616   -3.494535      -4.639297
##    C1-phenanthrenes/anthracenes C2-benzanthracenes/chrysenes
## 1                      8.522767                    -2.252446
## 2                     25.814314                     6.603662
## 3                     45.105894                    11.001437
## 4                     45.210323                    29.325615
## 5                     27.137703                     3.775680
## 6                            NA                           NA
## 7                     27.498200                     7.605885
## 8                     18.462628                    12.096204
## 9                     12.387138                     2.890332
## 10                    19.921050                     7.847687
## 11                    13.102392                    -2.495694
## 12                    14.345387                    -2.008354
## 13                    47.482328                    14.508489
## 14                    42.970047                    12.370165
## 15                    34.325337                    13.460916
## 16                   918.793183                    73.503455
## 17                    17.416182                    -1.223840
## 18                           NA                           NA
## 19                    40.560256                    17.186549
## 20                    29.757825                     5.180066
## 21                    30.400561                     5.203173
## 22                    38.070914                     7.028476
## 23                    12.333615                     2.359474
## 24                   624.051788                   114.073983
## 25                   154.052718                    39.369028
## 26                    25.298188                     5.192786
## 27                    33.706169                    19.078964
## 28                    11.861849                    -2.490988
```

```
## 29                   32.588319                    3.070822
## 30                   22.779312                    6.772228
## 31                    6.488018                   -2.868387
## 32                    9.280375                   -3.446996
## 33                   16.900679                   -2.145086
## 34                   39.957051                   11.127280
## 35                   11.470543                   -2.226635
## 36                   19.086797                    5.594406
## 37                   12.800437                   -3.200109
## 38                   16.874210                    5.905973
## 39                   23.259863                    4.788795
## 40                   13.086420                   -2.355556
## 41                   10.583287                   -3.598317
## 42                   15.612873                   -3.940392
## 43                   19.064141                    4.575394
## 44                   12.500255                   -2.697423
## 45                   30.425425                    4.855121
## 46                  130.243176                   32.018114
## 47                   14.600243                   -2.475693
## 48                   10.900862                   -3.077890
## 49                   13.898580                   -2.340813
## 50                          NA                          NA
## 51                   44.005519                    9.843340
## 52                  102.175258                   27.647423
## 53                   11.473598                   -2.409456
## 54                   14.771749                   -3.015899
## 55                   11.257041                   -2.673547
## 56                   15.179098                   -1.868197
## 57                   22.597250                    4.341051
## 58                   20.701345                    5.236222
## 59                   20.419715                   -2.162088
## 60                   37.205882                   26.044118
## 61                   26.195745                    5.001006
## 62                  162.007874                   60.266929
## 63                   75.171386                   38.337407
## 64                   20.562217                    6.511369
## 65                   31.373201                    4.913162
## 66                  325.148957                   97.544687
## 67                   12.947191                    2.395230
## 68                          NA                          NA
## 69                   53.222565                   27.260338
## 70                   21.208151                    5.537684
## 71                   26.604659                   -2.855134
## 72                   18.864407                   -2.593856
## 73                   20.245819                    7.423467
## 74                   19.280196                    3.253533
##    C2-dibenzothiophenes C2-fluoranthenes/pyrenes C2-fluorenes C2-naphthalenes
## 1             -3.104722               -3.104722   -3.104722       7.305229
## 2             -3.842130                9.004993    4.742630      10.205659
## 3              9.351222               14.851941    9.901294      17.437278
## 4             13.440907               23.827062   17.717559      18.328509
## 5             -3.539700                8.259301    4.719600      10.029151
## 6                    NA                      NA          NA             NA
## 7              4.622038                8.776021    5.850681      12.286430
```

```
## 8            5.793445          10.822920    -4.711153         8.912993
## 9           -2.713373           3.657155    -2.713373         5.839651
## 10          -3.622009           8.451355    -3.622009         9.055023
## 11          -3.868325          -3.868325    -3.868325         8.111005
## 12          -3.098603           4.188853    -3.098603         8.607232
## 13           7.913721          17.146396     9.232675        18.465350
## 14           7.812736          16.276533     9.765920        20.833962
## 15           6.730458          11.441779     8.076550        20.864420
## 16          91.879318         202.134500    98.004606       179.470935
## 17          -1.929901           3.906873     2.024043         5.648491
## 18                 NA                 NA           NA               NA
## 19           8.249544          21.998783     8.249544        11.686853
## 20          -3.692175           9.368204    -3.692175         9.368204
## 21          -4.209308           8.769393     4.793935        11.107897
## 22           3.689950          12.299834     5.857064        11.128421
## 23          -2.252225           3.700085     2.466723         6.971174
## 24          59.721085         275.119605    53.681874        39.590382
## 25          16.546403          79.879187    18.828666        18.600439
## 26          -5.126212           5.325934    -5.126212         8.654643
## 27          13.355275          21.622825     8.267551        13.355275
## 28          -3.380627          -3.380627    -3.380627         6.524017
## 29          -2.882813           8.147080     3.196162         9.400477
## 30          -3.878640           7.387885     4.063337        -3.878640
## 31          -3.824516          -3.824516    -3.824516        14.341935
## 32          -4.573899          -4.573899    -4.573899         8.617491
## 33          -3.380136           5.460219     3.510141         9.100366
## 34           5.563640          16.690920    12.138851        10.621495
## 35          -3.171268          -3.171268    -3.171268         6.072641
## 36          -4.475525           5.857672    -4.475525         9.214316
## 37          -4.352149          -4.352149    -4.352149        10.240350
## 38          -3.555637           5.544383    -3.555637         7.834455
## 39          -3.899448           6.841136     4.241504        10.261704
## 40          -3.271605          -3.271605    -3.271605         7.197531
## 41          -4.727201          -4.727201    -4.727201         9.877734
## 42          -5.650373          -5.650373    -5.650373        11.895522
## 43          -3.558640           5.465054     3.812828         8.896599
## 44          -3.684286          -3.684286    -3.684286         6.447500
## 45          -4.337241           7.120844     4.725651        11.004941
## 46          14.109677          46.670471    46.127792        16.280397
## 47          -3.491363          -3.491363    -3.491363         7.617518
## 48          -4.167977          -4.167977    -4.167977         8.335953
## 49          -3.218618          -3.218618    -3.218618         8.778050
## 50                 NA                 NA           NA               NA
## 51           4.747964          13.896480     7.527260        14.475500
## 52          10.818557          46.880412     9.015464        11.960515
## 53          -3.327343          -3.327343    -3.327343         6.310479
## 54          -4.062231          -4.062231    -4.062231         7.385874
## 55          -3.728895          -3.728895    -3.728895         6.683868
## 56          -2.977439           3.794775    -2.977439         8.173361
## 57          -3.567987           6.541309     3.686920        11.298625
## 58          -2.922543          -2.922543     4.505587         7.306357
## 59          -3.543421           5.225045     3.663537         9.609278
## 60           8.185294          26.044118     9.673529        14.138235
## 61          -5.120077           9.525725    -5.120077        11.311799
```

```
## 62            17.496850              97.204724    22.033071       15.811969
## 63            11.275708              46.606259    13.530849       19.544560
## 64            -3.564118               6.854072     4.866391        9.595701
## 65            -5.623498               8.287261     5.919472       11.246997
## 66            33.815492             188.586395    42.269364       35.116087
## 67            -2.589438               3.366270    -2.589438        7.120955
## 68                   NA                     NA          NA              NA
## 69             9.086779              29.856561    12.332058       20.769781
## 70            -4.064896               5.537684    -4.064896        8.836730
## 71            -4.412480               5.061374    -4.412480       10.382306
## 72            -3.654979               4.303443    -3.654979        8.253178
## 73            -4.656538               7.423467    -4.656538        8.773188
## 74            -3.494535               5.723808     3.615037        7.230073
##    C2-phenanthrenes/anthracenes C3-benzanthracenes/chrysenes
## 1                      9.740305                    -2.252446
## 2                     40.222304                     3.181764
## 3                     66.008625                     5.500719
## 4                    103.861553                    15.884708
## 5                     38.346754                    -2.182815
## 6                            NA                           NA
## 7                     35.104086                     4.212490
## 8                     37.561898                     9.549635
## 9                     17.695911                    -1.887564
## 10                    25.354064                     6.640350
## 11                    22.461244                    -2.495694
## 12                    20.657357                    -2.008354
## 13                    65.947678                     7.254245
## 14                    53.387028                     6.445507
## 15                    46.440162                     6.730458
## 16                   673.781667                    28.788853
## 17                    24.006089                    -1.223840
## 18                           NA                           NA
## 19                    48.122337                     8.249544
## 20                    38.023888                    -2.204283
## 21                    39.169953                    -2.689280
## 22                    42.170859                    -2.049972
## 23                    18.232300                     1.823230
## 24                   570.369914                    32.880148
## 25                   176.875343                    15.975837
## 26                    31.955606                    -3.328709
## 27                    63.596545                    13.991240
## 28                    15.420404                    -2.490988
## 29                    35.095113                    -1.942765
## 30                    27.704569                    -2.770457
## 31                     4.575760                    -2.868387
## 32                    14.583446                    -3.446996
## 33                    21.450862                    -2.145086
## 34                    70.809964                     6.575211
## 35                    16.193708                    -2.226635
## 36                    27.642948                    -3.290827
## 37                    17.280590                    -3.200109
## 38                    24.106014                    -2.591397
## 39                    29.416886                    -2.462809
## 40                    17.666667                    -2.355556
```

```
## 41                   13.405496                   -3.598317
## 42                   19.330224                   -3.940392
## 43                   27.325269                   -2.541885
## 44                   15.789796                   -2.697423
## 45                   33.662172                   -2.783603
## 46                  227.925558                   20.079156
## 47                   20.948175                   -2.475693
## 48                   14.106998                   -3.077890
## 49                   10.972563                   -2.340813
## 50                          NA                          NA
## 51                   50.374738                    5.095376
## 52                  114.195876                   14.424742
## 53                   16.063037                   -2.409456
## 54                   19.080175                   -3.015899
## 55                   15.478431                   -2.673547
## 56                   23.352459                   -1.868197
## 57                   25.570572                   -2.557057
## 58                   34.096333                    3.409633
## 59                   27.026094                   -2.162088
## 60                   69.947059                   14.138235
## 61                   37.507544                   -3.036325
## 62                  181.448819                   25.273228
## 63                   97.722802                   17.289419
## 64                   30.843326                    3.975362
## 65                   43.804092                   -3.906851
## 66                  325.148957                   35.116087
## 67                   21.362865                   -1.618399
## 68                          NA                          NA
## 69                   84.377237                   15.577336
## 70                   30.044881                   -3.004488
## 71                   31.795812                   -2.855134
## 72                   21.222458                   -2.593856
## 73                   26.994425                   -3.374303
## 74                   28.920293                   -2.169022
##    C3-dibenzothiophenes C3-fluoranthenes/pyrenes C3-fluorenes C3-naphthalenes
## 1             -3.104722                -3.104722   -3.104722        9.131536
## 2             -3.842130                 6.003329    5.883262       13.807656
## 3             10.451366                13.751797   17.602300       34.104456
## 4             18.939460                26.881814   40.933671       34.824168
## 5             -3.539700                 4.424625   10.619101       14.158801
## 6                    NA                       NA          NA              NA
## 7              5.558147                 7.020817   10.531226       16.381907
## 8             14.006131                15.279416    8.276350       12.732847
## 9             -2.713373                 2.831346    3.893101        7.668228
## 10            -3.622009                14.488037   -3.622009       14.488037
## 11            -3.868325                -3.868325   -3.868325       11.230622
## 12            -3.098603                -3.098603    3.155985       10.902494
## 13             9.892152                13.849012   12.530059       32.314362
## 14            11.068042                12.370165   16.276533       29.297759
## 15             8.749596                10.768733   16.153100       28.940970
## 16            67.378167               104.129894  110.255182      336.890834
## 17            -1.929901                 2.071114    2.400609        7.531322
## 18                   NA                       NA          NA              NA
## 19            10.999391                18.561473   14.436701       17.186549
```

```
## 20    4.022817   6.612850   5.124959  12.123558
## 21   -4.209308   6.430888   5.437023  19.877290
## 22    3.865662   8.785596   5.857064  17.571191
## 23   -2.252225   2.788469   2.788469   9.652394
## 24   51.668804 161.045623  80.522811  80.522811
## 25   19.399231  57.056562  31.381109  37.086766
## 26   -5.126212  -5.126212  -5.126212  16.643545
## 27   29.254411  25.438618  17.171067  21.622825
## 28   -3.380627  -3.380627  -3.380627   7.710202
## 29   -2.882813   4.637569   3.760191  11.907271
## 30   -3.878640   6.095005   7.387885  13.544456
## 31   -3.824516  -3.824516  -3.824516   8.878341
## 32   -4.573899  -4.573899  -4.573899   9.280375
## 33   -3.380136  -3.380136   4.095164  11.050444
## 34    6.575211  16.185135  23.771916  22.760346
## 35   -3.171268  -3.171268  -3.171268   7.422116
## 36   -4.475525   5.265323  -4.475525  13.163309
## 37   -4.352149  -4.352149  -4.352149  12.160416
## 38   -3.555637   5.122528  -3.555637  11.450357
## 39   -3.899448   4.720384   5.130852  13.682272
## 40   -3.271605  -3.271605  -3.271605   8.506173
## 41   -4.727201  -4.727201  -4.727201  10.583287
## 42   -5.650373  -5.650373  -5.650373  12.638993
## 43   -3.558640   4.130564   4.829582  11.438485
## 44   -3.684286  -3.684286  -3.684286   7.894898
## 45   -4.337241  -4.337241   4.596181  18.125785
## 46   22.792556  36.902233  81.401985  86.828784
## 47   -3.491363  -3.491363  -3.491363   9.521898
## 48   -4.167977  -4.167977  -4.167977   8.977181
## 49   -3.218618  -3.218618  -3.218618  10.241059
## 50          NA         NA         NA         NA
## 51    7.527260  10.422360  14.475500  20.265699
## 52   16.828866  40.269072  19.834021  21.036082
## 53   -3.327343  -3.327343  -3.327343   6.884159
## 54   -4.062231  -4.062231   4.123780   9.847832
## 55   -3.728895  -3.728895  -3.728895   7.739216
## 56   -2.977439  -2.977439  -2.977439   9.924795
## 57   -3.567987   4.519450   5.411447  13.677283
## 58    3.287861  -2.922543   6.088631  12.177262
## 59   -3.543421  -3.543421   3.603479  12.612177
## 60   12.650000  25.300000  15.626471  25.300000
## 61   -5.120077   5.953578  -5.120077  16.074662
## 62   21.385039  71.283465  24.625197  39.529921
## 63   13.530849  36.082265  28.565127  26.309985
## 64    5.414717   6.648450   7.539480  13.708145
## 65   -5.623498  -5.623498   7.103366  15.982574
## 66   38.367577 123.556604  65.029791  78.035750
## 67   -2.589438  -2.589438  -2.589438   9.710393
## 68          NA         NA         NA         NA
## 69   14.928280  27.909394  16.875447  31.154672
## 70   -4.064896   4.712922   5.773330  11.782306
## 71   -4.412480  -4.412480  -4.412480  17.520141
## 72   -3.654979  -3.654979  -3.654979  11.200742
## 73   -4.656538   6.748606  -4.656538  12.147491
```

```
## 74            -3.494535                4.097042     4.277793       12.050122
##    C3-phenanthrenes/anthracenes C4-benzanthracenes/chrysenes
## 1                      5.600675                    -2.252446
## 2                     34.819308                    -2.401332
## 3                     71.509344                    -1.815237
## 4                    128.299566                    -3.543512
## 5                     30.087453                    -2.182815
## 6                            NA                           NA
## 7                     29.253405                    -2.340272
## 8                     50.931387                    -3.501533
## 9                     11.797274                    -1.887564
## 10                    19.921050                    -2.233572
## 11                    12.478469                    -2.495694
## 12                    13.197756                    -2.008354
## 13                    65.947678                    -3.627122
## 14                    52.735967                    -3.190200
## 15                    49.805391                    -3.096011
## 16                   361.391985                     7.350345
## 17                    16.474767                    -1.223840
## 18                           NA                           NA
## 19                    48.122337                    -2.681102
## 20                    30.308896                    -2.204283
## 21                    26.308178                    -2.689280
## 22                    35.142382                    -2.049972
## 23                    10.724883                    -1.447859
## 24                   342.221948                     8.723305
## 25                   148.347062                     4.336299
## 26                    25.963930                    -3.328709
## 27                    76.315854                     6.995620
## 28                     9.489479                    -2.490988
## 29                    22.561144                    -1.942765
## 30                    22.779312                    -2.770457
## 31                    -3.824516                    -2.868387
## 32                     6.363686                    -3.446996
## 33                    16.250653                    -2.145086
## 34                    75.867818                    -1.618513
## 35                     8.771592                    -2.226635
## 36                    25.668452                    -3.290827
## 37                     8.960306                    -3.200109
## 38                    21.695413                    -2.591397
## 39                    22.575750                    -2.462809
## 40                     9.814815                    -2.355556
## 41                     7.761077                    -3.598317
## 42                    10.408582                    -3.940392
## 43                    20.335084                    -2.541885
## 44                     8.552806                    -2.697423
## 45                    26.541328                    -2.783603
## 46                   195.364764                     7.597519
## 47                    11.426277                    -2.475693
## 48                     9.618408                    -3.077890
## 49                     8.778050                    -2.340813
## 50                           NA                           NA
## 51                    38.215319                    -2.084472
## 52                   120.206186                     5.409278
```

```
## 53                    9.178878                    -2.409456
## 54                   11.694301                    -3.015899
## 55                    7.739216                    -2.673547
## 56                   12.260041                    -1.868197
## 57                   19.623927                    -2.557057
## 58                   29.834291                    -1.765703
## 59                   18.617976                    -2.162088
## 60                  104.176471                    -3.497353
## 61                   29.172534                    -3.036325
## 62                  155.527559                     5.637874
## 63                  105.239940                    11.275708
## 64                   31.528733                    -2.193303
## 65                   32.557096                    -3.906851
## 66                  260.119166                    14.306554
## 67                   12.299831                    -1.618399
## 68                          NA                           NA
## 69                  103.848907                    -3.115467
## 70                   25.921074                    -3.004488
## 71                   21.413506                    -2.855134
## 72                   14.737818                    -2.593856
## 73                   23.620122                    -3.374303
## 74                   20.485208                    -2.169022
##    C4-dibenzothiophenes C4-fluoranthenes/pyrenes C4-naphthalenes
## 1             -3.104722               -3.104722        3.530861
## 2             -3.842130               -3.842130        9.004993
## 3             11.551509                5.005654       26.403450
## 4             15.884708                8.553304       32.991317
## 5             -3.539700               -3.539700        9.439201
## 6                    NA                      NA              NA
## 7              4.270997               -3.276381       11.116294
## 8             14.006131                5.602453       10.822920
## 9             -2.713373               -2.713373       -2.654387
## 10            -3.622009                4.286044        7.244018
## 11            -3.868325               -3.868325        6.863158
## 12            -3.098603               -3.098603        6.885786
## 13             6.001239               -5.605553       23.741164
## 14             7.812736                5.338703       19.531840
## 15             8.749596               -4.307493       21.537466
## 16            34.301612               50.227361      196.009212
## 17            -1.929901               -1.929901        4.377581
## 18                   NA                      NA              NA
## 19             6.874620                8.937005        8.937005
## 20             4.684102               -3.692175        7.714992
## 21            -4.209308               -4.209308       12.277150
## 22            -3.397097                3.807091        9.957008
## 23            -2.252225               -2.252225        5.898685
## 24            30.867078               73.812577       66.431319
## 25            16.546403               26.816584       26.816584
## 26            -5.126212               -5.126212        7.988901
## 27            21.622825                8.267551       18.442998
## 28            -3.380627               -3.380627       -3.380627
## 29            -2.882813               -2.882813        6.893683
## 30            -3.878640               -3.878640        4.432731
## 31            -3.824516               -3.824516        4.985530
```

```
## 32            -4.573899              -4.573899        7.291723
## 33            -3.380136              -3.380136        5.070204
## 34             5.563640               7.080996       21.242989
## 35            -3.171268              -3.171268       -3.171268
## 36            -4.475525              -4.475525        5.660223
## 37            -4.352149              -4.352149        5.440186
## 38            -3.555637              -3.555637        4.881468
## 39            -3.899448              -3.899448        6.635902
## 40            -3.271605              -3.271605        3.402469
## 41            -4.727201              -4.727201        7.055524
## 42            -5.650373              -5.650373        7.434701
## 43            -3.558640              -3.558640        3.939923
## 44            -3.684286              -3.684286       -3.684286
## 45            -4.337241              -4.337241       10.357591
## 46            23.335236              16.280397       97.682382
## 47            -3.491363              -3.491363        5.840097
## 48            -4.167977              -4.167977        5.322186
## 49            -3.218618              -3.218618       -3.218618
## 50                   NA                     NA              NA
## 51             4.747964               4.111042       13.317460
## 52            16.227835              21.036082       13.823711
## 53            -3.327343              -3.327343       -3.327343
## 54            -4.062231              -4.062231        6.770385
## 55            -3.728895              -3.728895       -3.658538
## 56            -2.977439              -2.977439        6.421926
## 57            -3.567987              -3.567987        7.730638
## 58             3.044315              -2.922543        7.915220
## 59            -3.543421              -3.543421        8.408118
## 60             6.771471               9.673529       19.347059
## 61            -5.120077              -5.120077       10.121083
## 62            13.608661              31.753543       27.865354
## 63            15.034277              16.537705       21.799702
## 64             5.414717              -3.564118        6.648450
## 65            -5.623498              -5.623498        9.471155
## 66            35.766385              53.974727       59.827408
## 67            -2.589438              -2.589438        4.466781
## 68                   NA                     NA              NA
## 69             9.735835              11.033946       22.716948
## 70            -4.064896              -4.064896        6.480268
## 71            -4.412480              -4.412480        9.084518
## 72            -3.654979              -3.654979        5.777225
## 73            -4.656538              -4.656538        5.398885
## 74            -3.494535              -3.494535        7.832579
##    C4-phenanthrenes/anthracenes   chrysene dibenz[a,h]anthracene
## 1                      6.026814   4.139630             -3.104722
## 2                     30.616977  23.412983             -3.001664
## 3                     55.007187  24.753234             -2.200287
## 4                     85.533044  39.711770             -4.948698
## 5                     23.008052  21.828152             -2.713770
## 6                            NA         NA                    NA
## 7                     28.083268  21.647519             -3.159368
## 8                     36.925255  10.822920             -4.838482
## 9                     14.156729  10.617547             -2.536414
## 10                    19.317382  13.884368             -3.501276
```

```
## 11                  12.478469   8.111005           -3.119617
## 12                  24.100249  12.050125           -2.467406
## 13                  46.822851  44.844421           -5.671500
## 14                  39.063679  33.855189           -4.622535
## 15                  41.728841  23.556604           -4.442102
## 16                 140.881621 735.034546           18.375864
## 17                  13.179813  11.296983           -1.506264
## 18                         NA         NA                  NA
## 19                  27.498478  59.121729           -3.987279
## 20                  21.491763  35.268534           -2.755354
## 21                  32.154439  28.646682           -4.150846
## 22                  24.599668  42.170859           -3.221385
## 23                  10.724883   7.507418           -1.769606
## 24                 154.335388 671.023428           16.775586
## 25                  91.290500 171.169687            7.987919
## 26                  21.303737  19.306512           -5.126212
## 27                  54.693029  21.622825           -6.359655
## 28                   9.489479  11.268757           -3.380627
## 29                  13.787366  23.814541           -2.381454
## 30                  16.007084  31.398511           -3.755508
## 31                   4.644055   5.873364           -3.961106
## 32                   5.435648   6.098532           -4.905341
## 33                  11.700470  16.250653           -3.185128
## 34                  50.072760  30.852913           -1.972563
## 35                   8.771592  10.121068           -3.036320
## 36                  18.428632  18.428632           -4.541341
## 37                  17.920612   5.568190           -4.416151
## 38                  15.066259  15.066259           -3.615902
## 39                  15.050500  22.575750           -3.625802
## 40                  12.432099  12.432099           -3.140741
## 41                   5.926640   6.843859           -5.150533
## 42                   8.921642  10.408582           -5.724720
## 43                  19.064141  18.428670           -3.367998
## 44                   9.868622  13.816071           -3.618495
## 45                  21.362532  22.657231           -4.337241
## 46                  97.682382  70.548387            4.992655
## 47                  16.504623   8.887105           -3.047007
## 48                   6.412272   8.977181           -4.488590
## 49                  13.167076   9.509555           -3.291769
## 50                         NA         NA                  NA
## 51                  26.055899  30.109039           -3.010904
## 52                  96.164948 102.175258            2.884948
## 53                   9.752558  12.047278           -3.269975
## 54                  12.309791  15.387238           -4.308427
## 55                   6.543155  11.257041           -3.588182
## 56                  16.346721  12.260041           -2.276865
## 57                  13.677283  19.623927           -3.686920
## 58                  20.092482  15.221577           -2.191907
## 59                  14.413917  15.014497           -2.642551
## 60                  74.411765  45.391176           -5.506471
## 61                  19.051451  26.791103           -3.750754
## 62                  90.724409 162.007874            9.072441
## 63                  75.171386 135.308495           -5.412340
## 64                  21.933032  17.135181           -3.289955
```

```
## 65                   24.269835  28.413465             -4.853967
## 66                  136.562562 383.675770             14.956852
## 67                   12.299831  10.357753             -1.942079
## 68                          NA         NA                    NA
## 69                   71.396123  48.030119             -4.867917
## 70                   20.029920  20.619036             -4.182719
## 71                   -4.412480  18.817929             -4.477369
## 72                   11.790254  12.379767             -3.596028
## 73                   31.718449  16.871515             -4.724024
## 74                   12.652628  17.472677             -2.711278
##    dibenzothiophene indeno[1,2,3-cd]pyrene phenanthrene NA
## 1         -3.104722              -3.104722     12.17538 NA
## 2         -3.842130              -3.001664     30.61698 NA
## 3         -2.915381              -2.200287     37.95496 NA
## 4         -4.887603              -5.009793     39.10082 NA
## 5         -3.539700              -2.713770     37.16685 NA
## 6                NA                     NA           NA NA
## 7         -3.276381              -3.217875     32.17875 NA
## 8         -4.647489              -4.902146     19.09927 NA
## 9         -2.654387              -2.536414     15.33646 NA
## 10        -3.561642              -3.501276     33.20175 NA
## 11        -3.805933              -3.119617     18.71770 NA
## 12        -3.098603              -2.467406     22.37880 NA
## 13        -5.539605              -5.671500     57.37448 NA
## 14        -4.687642              -4.622535     59.89764 NA
## 15        -4.240189              -4.509407     39.03666 NA
## 16       140.881621              39.201842   1776.33349 NA
## 17        -1.929901              -1.553335     33.89095 NA
## 18               NA                     NA           NA NA
## 19        -4.262264              13.061777     57.74680 NA
## 20        -3.692175              -2.755354     60.61779 NA
## 21        -4.209308              -4.209308     43.26234 NA
## 22        -3.338526              -3.221385     64.42770 NA
## 23        -2.198601              -1.769606     12.86986 NA
## 24        36.906289              42.274476    657.60296 NA
## 25        11.411312              18.828666    199.69797 NA
## 26        -5.059638              -5.192786     41.94173 NA
## 27        -6.041672              -6.359655     36.88600 NA
## 28        -3.380627              -3.380627     17.19968 NA
## 29        -2.820143              -2.381454     52.64267 NA
## 30        -3.817074               4.124902     32.62983 NA
## 31        -3.756221              -4.029401     12.97604 NA
## 32        -4.507611              -4.971629     15.90921 NA
## 33        -3.380136              -3.185128     25.35102 NA
## 34        -2.630084               4.754383     37.93391 NA
## 35        -3.171268              -3.036320     14.16949 NA
## 36        -4.475525              -4.541341     22.37762 NA
## 37        -4.288147              -4.480153     18.56063 NA
## 38        -3.495372              -3.676167     19.28481 NA
## 39        -3.899448              -3.625802     32.15334 NA
## 40        -3.271605              -3.206173     17.01235 NA
## 41        -4.727201              -5.221088     17.63881 NA
## 42        -5.576026              -5.724720     19.33022 NA
## 43        -3.558640              -3.431545     27.96074 NA
```

```
## 44       -3.618495             -3.618495     18.42143 NA
## 45       -4.272506             -4.401976     42.72506 NA
## 46        5.969479             14.652357    113.96278 NA
## 47       -3.427883             -3.047007     19.04380 NA
## 48       -4.103854             -4.488590     17.95436 NA
## 49       -3.218618             -3.291769     24.13964 NA
## 50              NA                    NA           NA NA
## 51       -3.300414             -3.068806     63.69220 NA
## 52        5.830000              7.813402    144.24742 NA
## 53       -3.269975             -3.269975     16.06304 NA
## 54       -4.000682             -4.369976     18.46469 NA
## 55       -3.658538             -3.588182     15.47843 NA
## 56       -2.977439             -2.335246     15.17910 NA
## 57       -3.567987             -3.686920     33.30121 NA
## 58       -2.861656              2.618111     21.91907 NA
## 59       -3.483363             -2.702609     30.62957 NA
## 60       -5.580882              8.185294     38.69412 NA
## 61       -5.060542             -3.750754     35.72147 NA
## 62        9.720472             22.681102    187.92913 NA
## 63        6.013711             11.275708    105.23994 NA
## 64       -3.495577             -3.289955     26.04548 NA
## 65       -5.564304             -4.853967     44.39604 NA
## 66       22.110129             31.864598    390.17875 NA
## 67       -2.524702             -2.006815     13.59455 NA
## 68              NA                    NA           NA NA
## 69       -4.997729             -4.932823     61.01123 NA
## 70       -4.064896             -4.241630     28.86665 NA
## 71       -4.347591             -4.477369     40.88033 NA
## 72       -3.596028             -3.654979     27.70710 NA
## 73       -4.589052             -4.791510     27.66929 NA
## 74       -3.434285             -2.711278     25.90776 NA
```

```r
#get the column names so I don't have to individually type each one

all_columns <- names(reshaped2_df)

# Remove the columns you don't want to include in the model
excluded_columns <- c('avg_p450', 'latitude', 'longitude', 'avg_SOD', 'site_name', 'site_number', 'NA')
independent_columns <- all_columns[!all_columns %in% excluded_columns]

# Enclose each column name in backticks to handle special characters
independent_columns <- sapply(independent_columns, function(x) paste0("`", x, "`"))

# Create a string representing the formula
formula_str <- paste("avg_p450 ~", paste(independent_columns, collapse = " + "))

# Convert the string to a formula object
formula <- as.formula(formula_str)

# Now you can use this formula in your model
#lm_model <- lm(formula, data = reshaped_df)
#summary(lm_model)
```

```r
indvlm_model <- lm(formula, data = reshaped2_df)
print(summary(indvlm_model))
```

```
##
## Call:
## lm(formula = formula, data = reshaped2_df)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -3998260  -925684   -46998   893887  6773539
##
## Coefficients:
##                                Estimate Std. Error t value Pr(>|t|)
## (Intercept)                     8253120    2893118   2.853  0.00754 **
## acenaphthene                     103342     182869   0.565  0.57593
## acenaphthylene                  1822594    1027471   1.774  0.08560 .
## anthracene                        44812     148264   0.302  0.76442
## `benz[a]anthracene`              152252     108074   1.409  0.16854
## `benzo[a]pyrene`                -121907     266145  -0.458  0.65001
## `benzo[b]fluoranthene`           -40744     206181  -0.198  0.84460
## `benzo[e]pyrene`                -212536     209781  -1.013  0.31860
## `benzo[ghi]perylene`            -161066     412598  -0.390  0.69885
## `benzo[k]fluoranthene`            38340     212082   0.181  0.85768
## `C1-benzanthracenes/chrysenes`  -139353     230364  -0.605  0.54949
## `C1-dibenzothiophenes`          -358848     479115  -0.749  0.45934
## `C1-fluoranthenes/pyrenes`      -130668     185601  -0.704  0.48651
## `C1-fluorenes`                   -81602     226397  -0.360  0.72089
## `C1-naphthalenes`                201789     149565   1.349  0.18675
## `C1-phenanthrenes/anthracenes`    78012     233322   0.334  0.74030
## `C2-benzanthracenes/chrysenes`   300661     226475   1.328  0.19371
## `C2-dibenzothiophenes`           195132     455744   0.428  0.67140
## `C2-fluoranthenes/pyrenes`       -23132     238330  -0.097  0.92328
## `C2-fluorenes`                    -9409     216261  -0.044  0.96557
## `C2-naphthalenes`               -217272     226659  -0.959  0.34495
## `C2-phenanthrenes/anthracenes`     7161     217795   0.033  0.97398
## `C3-benzanthracenes/chrysenes`  -346058     300052  -1.153  0.25732
## `C3-dibenzothiophenes`           562430     601522   0.935  0.35679
## `C3-fluoranthenes/pyrenes`      -159452     168664  -0.945  0.35155
## `C3-fluorenes`                  -209370     174778  -1.198  0.23975
## `C3-naphthalenes`                 76862     177662   0.433  0.66819
## `C3-phenanthrenes/anthracenes`    83530     151012   0.553  0.58402
## `C4-benzanthracenes/chrysenes`    69312     424837   0.163  0.87143
## `C4-dibenzothiophenes`          -694340     512308  -1.355  0.18481
## `C4-fluoranthenes/pyrenes`       148775     216692   0.687  0.49730
## `C4-naphthalenes`                109153     167164   0.653  0.51844
## `C4-phenanthrenes/anthracenes`     2694      70794   0.038  0.96988
## chrysene                         241653     155819   1.551  0.13077
## `dibenz[a,h]anthracene`          414524     876871   0.473  0.63961
## dibenzothiophene                -195285     630474  -0.310  0.75876
## `indeno[1,2,3-cd]pyrene`         -35786     415577  -0.086  0.93191
## phenanthrene                    -103317      98558  -1.048  0.30236
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2383000 on 32 degrees of freedom
```

```
##   (4 observations deleted due to missingness)
## Multiple R-squared:  0.5308, Adjusted R-squared:  -0.01167
## F-statistic: 0.9785 on 37 and 32 DF,  p-value: 0.5285
```

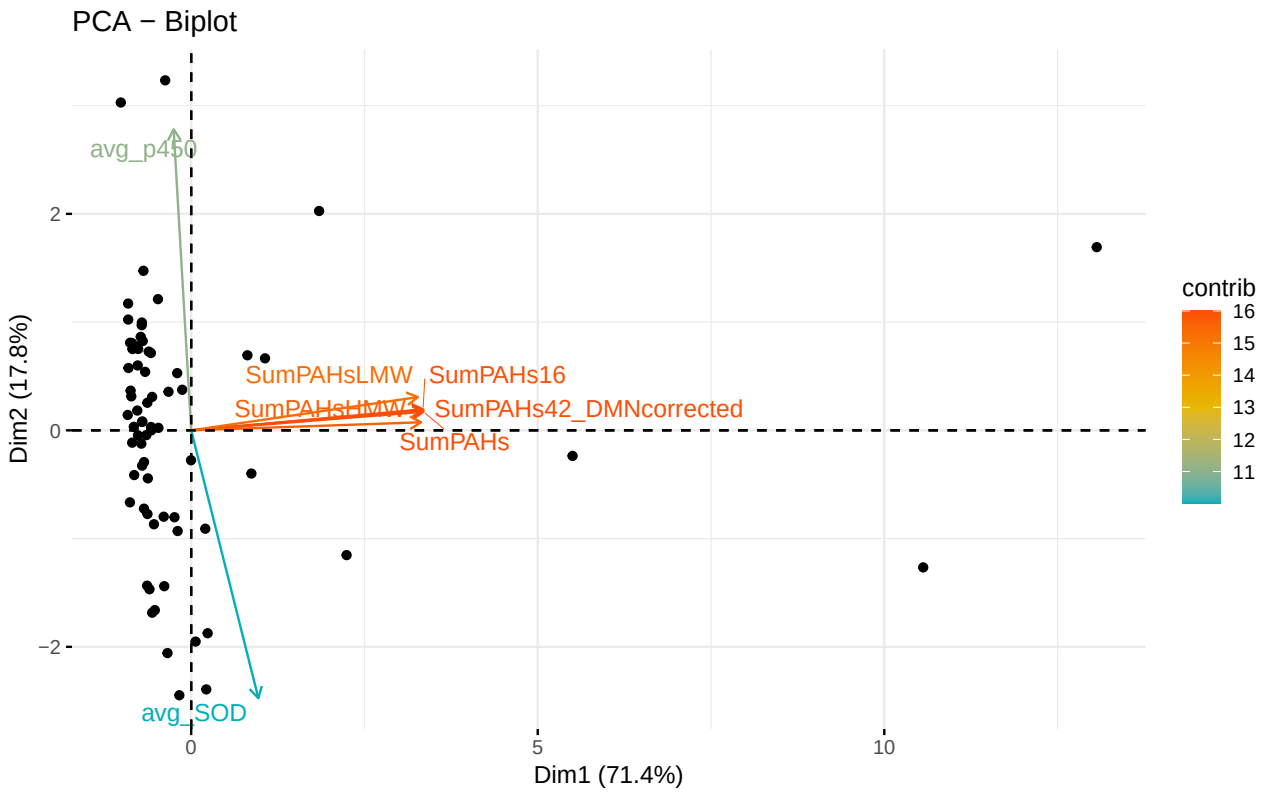
```
# GLM- p450 as response
indvglm_model<- glm(formula, data = reshaped2_df, family = poisson())
print(summary(indvglm_model))
```

```
##
## Call:
## glm(formula = formula, family = poisson(), data = reshaped2_df)
##
## Coefficients:
##                                  Estimate Std. Error   z value Pr(>|z|)
## (Intercept)                     1.600e+01  5.888e-04 27175.889  < 2e-16 ***
## acenaphthene                    2.481e-02  3.538e-05   701.326  < 2e-16 ***
## acenaphthylene                  3.656e-01  2.102e-04  1739.497  < 2e-16 ***
## anthracene                      1.233e-02  2.874e-05   428.867  < 2e-16 ***
## `benz[a]anthracene`             3.089e-02  2.274e-05  1358.295  < 2e-16 ***
## `benzo[a]pyrene`               -4.739e-03  5.589e-05   -84.792  < 2e-16 ***
## `benzo[b]fluoranthene`         -1.615e-02  4.240e-05  -380.945  < 2e-16 ***
## `benzo[e]pyrene`               -3.192e-02  4.153e-05  -768.562  < 2e-16 ***
## `benzo[ghi]perylene`           -6.793e-02  8.911e-05  -762.254  < 2e-16 ***
## `benzo[k]fluoranthene`          1.211e-02  4.504e-05   269.006  < 2e-16 ***
## `C1-benzanthracenes/chrysenes` -1.913e-02  4.366e-05  -438.086  < 2e-16 ***
## `C1-dibenzothiophenes`         -9.704e-02  1.057e-04  -917.955  < 2e-16 ***
## `C1-fluoranthenes/pyrenes`     -3.781e-02  4.000e-05  -945.323  < 2e-16 ***
## `C1-fluorenes`                 -2.006e-02  4.436e-05  -452.288  < 2e-16 ***
## `C1-naphthalenes`               3.969e-02  2.883e-05  1376.795  < 2e-16 ***
## `C1-phenanthrenes/anthracenes`  2.623e-02  4.712e-05   556.590  < 2e-16 ***
## `C2-benzanthracenes/chrysenes`  5.556e-02  4.394e-05  1264.525  < 2e-16 ***
## `C2-dibenzothiophenes`          5.796e-02  9.751e-05   594.422  < 2e-16 ***
## `C2-fluoranthenes/pyrenes`     -3.606e-03  4.630e-05   -77.876  < 2e-16 ***
## `C2-fluorenes`                 -3.688e-04  4.281e-05    -8.616  < 2e-16 ***
## `C2-naphthalenes`              -5.017e-02  4.767e-05 -1052.447  < 2e-16 ***
## `C2-phenanthrenes/anthracenes`  6.099e-03  4.318e-05   141.254  < 2e-16 ***
## `C3-benzanthracenes/chrysenes` -6.968e-02  6.057e-05 -1150.397  < 2e-16 ***
## `C3-dibenzothiophenes`          6.209e-02  1.233e-04   503.586  < 2e-16 ***
## `C3-fluoranthenes/pyrenes`     -3.720e-02  3.300e-05 -1127.241  < 2e-16 ***
## `C3-fluorenes`                 -3.880e-02  3.323e-05 -1167.788  < 2e-16 ***
## `C3-naphthalenes`               1.292e-02  3.620e-05   356.898  < 2e-16 ***
## `C3-phenanthrenes/anthracenes`  2.101e-02  3.114e-05   674.450  < 2e-16 ***
## `C4-benzanthracenes/chrysenes`  7.781e-03  9.168e-05    84.869  < 2e-16 ***
## `C4-dibenzothiophenes`         -9.891e-02  1.006e-04  -983.087  < 2e-16 ***
## `C4-fluoranthenes/pyrenes`      2.872e-02  4.420e-05   649.876  < 2e-16 ***
## `C4-naphthalenes`               6.489e-03  3.153e-05   205.792  < 2e-16 ***
## `C4-phenanthrenes/anthracenes`  9.562e-04  1.409e-05    67.867  < 2e-16 ***
## chrysene                        3.736e-02  2.932e-05  1274.052  < 2e-16 ***
## `dibenz[a,h]anthracene`         3.968e-02  1.731e-04   229.260  < 2e-16 ***
## dibenzothiophene                5.975e-03  1.316e-04    45.395  < 2e-16 ***
## `indeno[1,2,3-cd]pyrene`       -4.502e-04  8.860e-05    -5.081 3.76e-07 ***
## phenanthrene                   -2.266e-02  1.931e-05 -1173.937  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## (Dispersion parameter for poisson family taken to be 1)
##
##     Null deviance: 71801382  on 69  degrees of freedom
## Residual deviance: 30482761  on 32  degrees of freedom
##   (4 observations deleted due to missingness)
## AIC: Inf
##
## Number of Fisher Scoring iterations: 5
```

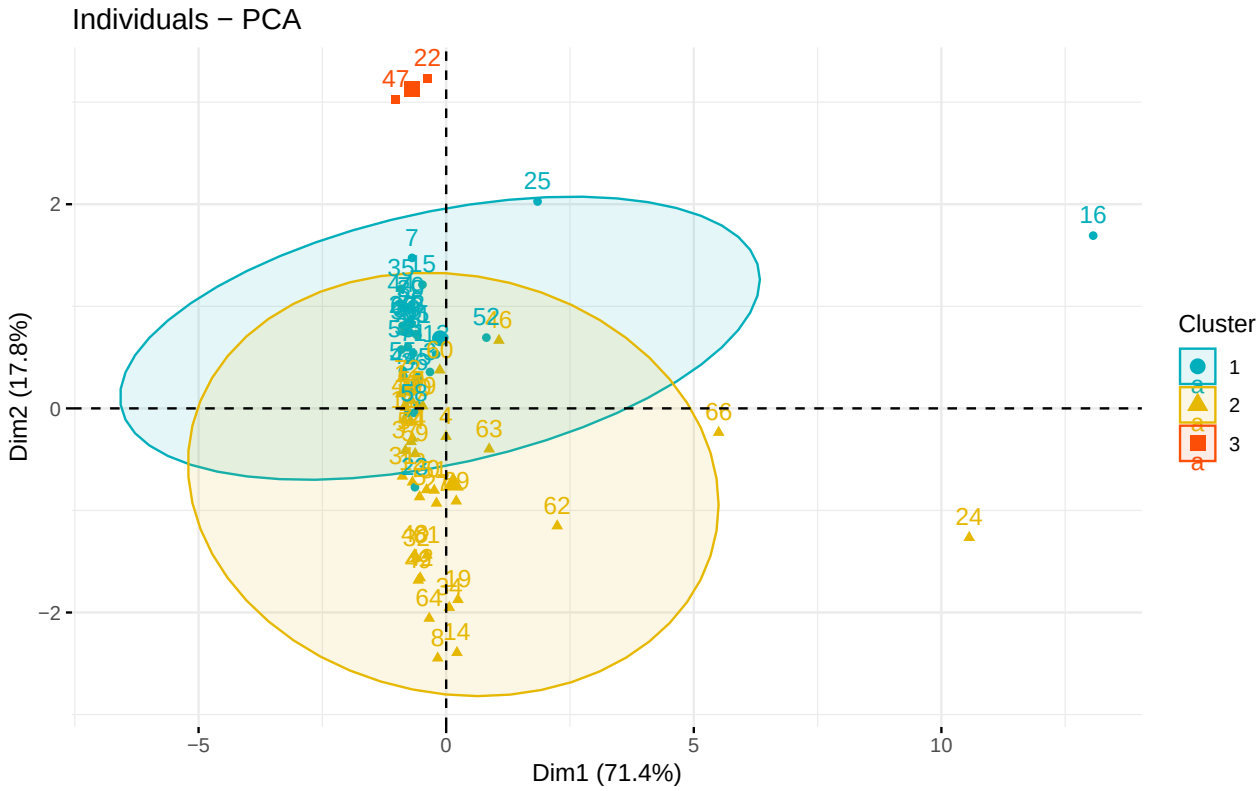
PCA of biomarkers and analyte data

```
#pca for pah and biomarkers, using reshaped_df
#install.packages("FactoMineR")
#install.packages("factoextra")
library('FactoMineR')
library('factoextra')

# Specify the columns you want to use for PCA
pca_columns <- c("avg_SOD", "avg_p450", "SumPAHs", "SumPAHs16", "SumPAHs42_DMNcorrected", "SumPAHsHMW",

# Remove rows with NAs in the specified columns only
df_clean <- reshaped_df[complete.cases(reshaped_df[, pca_columns]), ]

# Selecting the relevant variables for PCA
pca_data <- df_clean[, pca_columns]

# Performing PCA
pca_res <- PCA(pca_data, scale.unit = TRUE, graph = FALSE)

# Plotting the PCA
analytepcaplot<- fviz_pca_biplot(pca_res, label = "var", col.var = "contrib",
                gradient.cols = c("#00AFBB", "#E7B800", "#FC4E07"),
                repel = TRUE)  # Avoid text overlapping (slow if many points)



print(analytepcaplot)
```




```
ggsave(plot=analytepcaplot, filename="/Users/cmantegna/Documents/WDFWmussels/output/analytepca.png", wi
```

```
# Clustering based on k-means results. Needs to be revisited to verify it is the best option for this p

#different pca option
library(FactoMineR)
library(factoextra)

# Assuming pca_data is prepared for PCA
pca_result <- PCA(pca_data, graph = FALSE)

# If you need to perform clustering (e.g., k-means)
set.seed(123)  # for reproducibility
clusters <- kmeans(pca_data, centers = 3)  # change centers according to your data

# Add the cluster assignments to your PCA data
pca_data$cluster <- as.factor(clusters$cluster)

# Now plot the PCA with fviz_pca_ind
fviz_pca_ind(pca_result,
             col.ind = pca_data$cluster, # color by cluster
             palette = c("#00AFBB", "#E7B800", "#FC4E07"), # customize colors
             addEllipses = TRUE, # add confidence ellipses around clusters
             ellipse.level = 0.95, # 95% confidence ellipse
             legend.title = "Cluster")
```


Individuals − PCA
Dim2 (17.8%)
Dim1 (71.4%)
Cluster
1
2
3
22
47
25
16
7
52
26
66
63
62
24
64
14
2
0
−2
−5
0
5
10