02_EDA_Aqua_Aerobic_Primary_Filtration_Lab

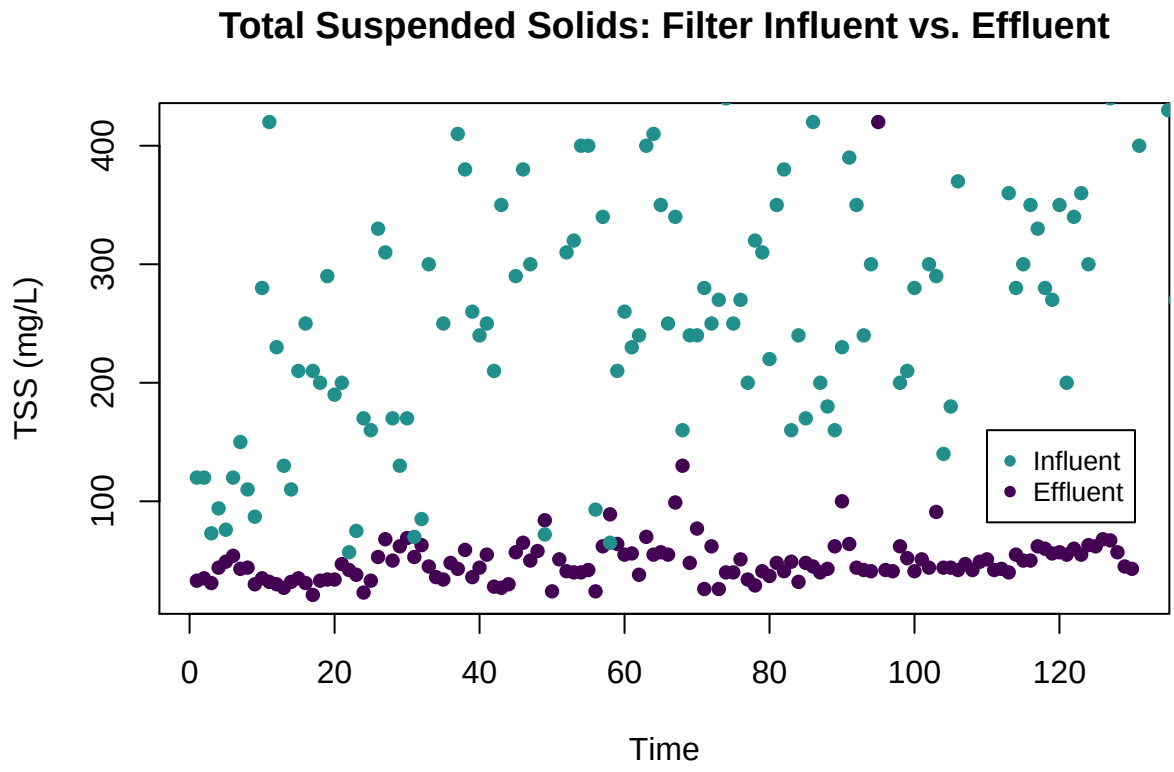
Maggie Bailey

5/30/2020

```
##-------------------------------------
## Clear working memory
##-------------------------------------
rm(list=ls())

##-------------------------------------
## Install and load any needed libraries
##-------------------------------------

library(lubridate)
library(xts)
library(glmnet)
library(viridis)

##-------------------------------------
## Load the data
##-------------------------------------

setwd("/Users/maggiebailey/Documents/Mines/MOWATER /Aqua Aerobic/data/clean")
setwd("~/Documents/Mines/MOWATER /Aqua Aerobic/data/clean")


clar_eff <- read.csv(file = "clarifier_effluent.csv")
clar_inf <- read.csv(file = "clarifier_influent.csv")
filter_eff <- read.csv(file = "filter_effluent.csv")
filter_inf <- read.csv(file = "filter_influent.csv")

# remove first column
clar_eff <- clar_eff[,-1]
clar_inf <- clar_inf[,-1]
filter_eff <- filter_eff[,-1]
filter_inf <- filter_inf[,-1]

str(filter_eff) # as an example
```

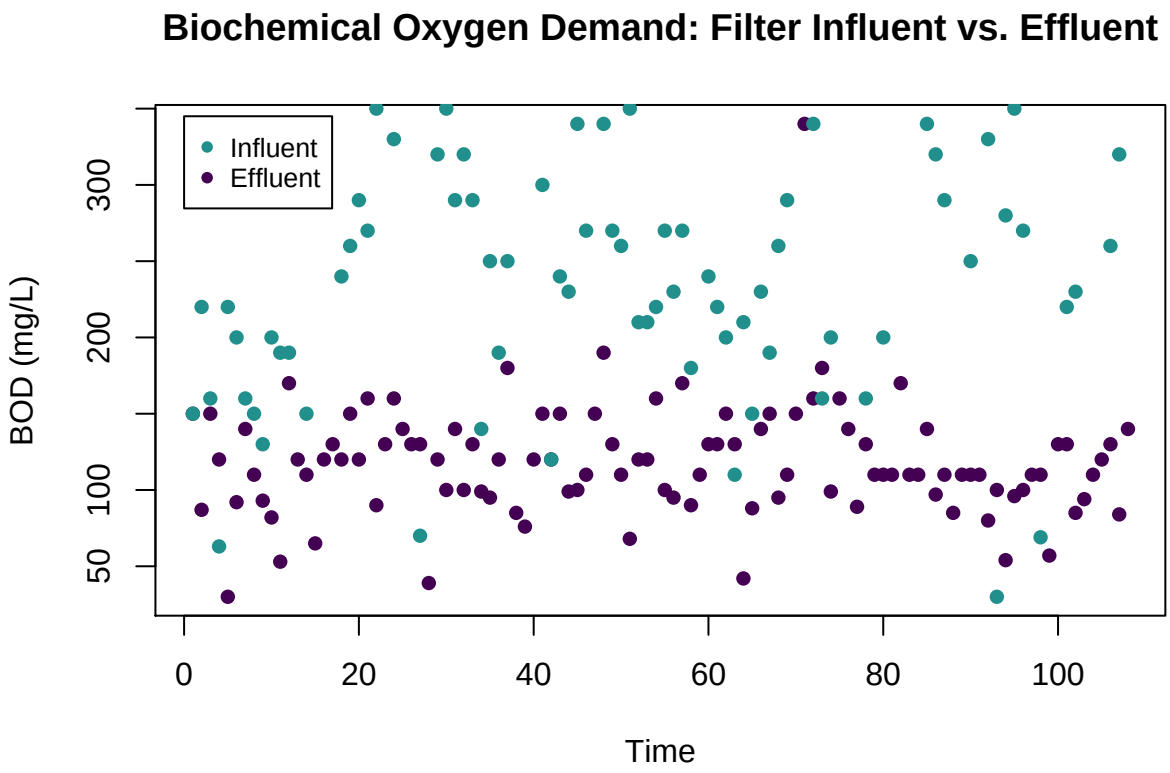
```
## 'data.frame':    625 obs. of  5 variables:
##  $ date       : Factor w/ 115 levels "2017-07-05","2017-07-10",..: 1 1 1 1 2 2 2 2 3 3 ...
##  $ time       : Factor w/ 17 levels "03:00:00","07:30:00",..: 6 6 6 6 6 6 6 6 6 6 ...
##  $ sample_type: Factor w/ 5 levels "1-hr Composite",..: 5 5 5 5 5 5 5 5 5 5 ...
##  $ parameter  : Factor w/ 11 levels "Alkalinity","Ammonia",..: 10 11 3 6 10 11 3 6 10 11 ...
##  $ value      : num  33 33 150 230 35 31 87 230 31 31 ...
```

```r
##-----------------------------------------
## Look at TSS: filter influent vs. effluent
##-----------------------------------------
cols <- viridis(3)

filt_eff_tss <- subset(filter_eff, filter_eff$parameter == "TSS (mg/L)")
filt_inf_tss <- subset(filter_inf, filter_inf$parameter == "TSS (mg/L)")

plot(filt_eff_tss$value, col = cols[1], pch = 16,
     ylab = "TSS (mg/L)",
     xlab = "Time",
     main = "Total Suspended Solids: Filter Influent vs. Effluent")
points(filt_inf_tss$value, col = cols[2], pch = 16)
legend(110, 160, c("Influent", "Effluent"), col = c(cols[2], cols[1]), cex = 0.8, pch = 16)
```

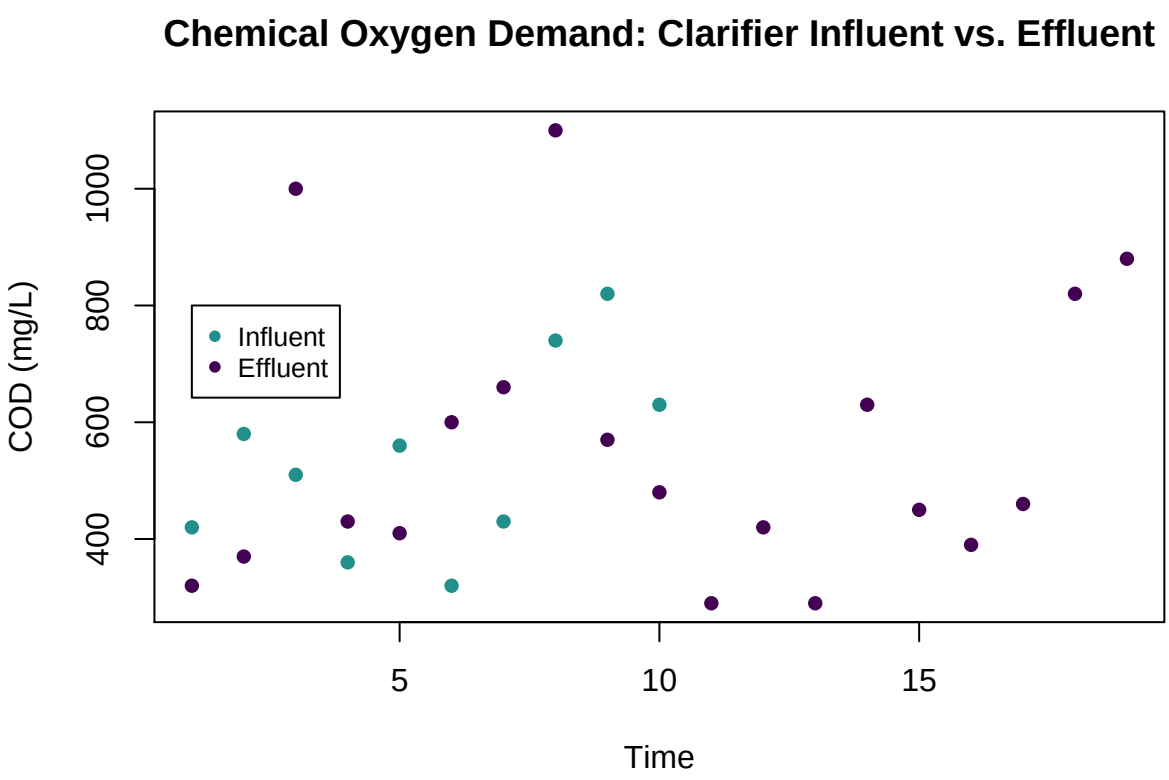



```r
##-----------------------------------------
## Look at BOD: filter influent vs. effluent
##-----------------------------------------

filt_eff_bod <- subset(filter_eff, filter_eff$parameter == "BOD (mg/L)")
filt_inf_bod <- subset(filter_inf, filter_inf$parameter == "BOD (mg/L)")

plot(filt_eff_bod$value, col = cols[1], pch = 16,
     ylab = "BOD (mg/L)",
     xlab = "Time",
```

```
     main = "Biochemical Oxygen Demand: Filter Influent vs. Effluent")
points(filt_inf_bod$value, col = cols[2], pch = 16)
legend(0, 345, c("Influent", "Effluent"), col = c(cols[2], cols[1]), cex = 0.8, pch = 16)
```




```
##-------------------------------------------
## Look at COD: clarifier influent vs. effluent
##-------------------------------------------

clar_eff_cod <- subset(clar_eff, clar_eff$parameter == "COD (mg/L)")
clar_inf_cod <- subset(clar_inf, clar_inf$parameter == "COD (mg/L)")

plot(clar_eff_cod$value, col = cols[1], pch = 16,
     ylab = "COD (mg/L)",
     xlab = "Time",
     main = "Chemical Oxygen Demand: Clarifier Influent vs. Effluent")
points(clar_inf_cod$value, col = cols[2], pch = 16)
legend(1, 800, c("Influent", "Effluent"), col = c(cols[2], cols[1]), cex = 0.8, pch = 16)
```




```
##-------------------------------------------
## Quick glimpse at compiled data
##-------------------------------------------

load("~/Documents/Mines/MOWATER /Aqua Aerobic/data/raw/compiledDBF.RData")
dim(rawData)
```

```
## [1] 17181    25
```

```
colnames(rawData)
```

```
##  [1] "ADFS\\BASIN_LEVEL\\PROCESS_VALUE"
##  [2] "ADFS\\EFFLUENT_TSS\\PROCESS_VALUE"
##  [3] "ADFS\\INFLUENT_FLOW\\PROCESS_VALUE"
##  [4] "ADFS\\WASTE_FLOW\\PROCESS_VALUE"
##  [5] "ADFS\\WASTE_VACUUM\\PROCESS_VALUE"
##  [6] "ADFS\\CURRENT_MODE"
##  [7] "SEPARATOR\\WASTE_FLOW\\PROCESS_VALUE"
##  [8] "SEPARATOR\\TANK_LEVEL\\PROCESS_VALUE"
##  [9] "SEPARATOR\\CURRENT_MODE"
## [10] "ADFS\\INFLUENT_VALVE\\POSITION_FB\\PROCESS_VALUE"
## [11] "SEPARATOR\\SR_CURRENT_MODE"
## [12] "SEPARATOR\\INFLUENT_FLOW\\PROCESS_VALUE"
## [13] "SEPARATOR\\INFLUENT_TANK_LEVEL\\PROCESS_VALUE"
## [14] "ADFS\\WASTE_PUMP_1\\SPEED_OUT"
```

```
## [15] "ADFS\\SCUM_REMOVAL_MODE"
## [16] "ADFS\\BASIN_PH\\PROCESS_VALUE"
## [17] "ADFS\\INFLUENT_FLOW\\TOTAL_FOREVER"
## [18] "ADFS\\WASTE_TSS\\PROCESS_VALUE"
## [19] "SEPARATOR\\INFLUENT_FLOW\\TOTAL_FOREVER"
## [20] "SEPARATOR\\WASTE_FLOW\\TOTAL_FOREVER"
## [21] "ADFS\\INFLUENT_FLOW\\PID\\OUTPUT"
## [22] "ADFS\\INFLUENT_TSS\\PROCESS_VALUE"
## [23] "ADFS\\INFLUENT_TURBIDITY\\PROCESS_VALUE"
## [24] "ADFS\\EFFLUENT_TURBIDITY\\PROCESS_VALUE"
## [25] "ADFS\\INHIBIT_WASTE_FLOW"
```