Aqua Aerobics EDA

Nate Rowan

6/2/2020

```r
library(tidyverse)
library(lubridate)
library(here) # eliminates working directory/filepath stuff

# Read in the data
raw_clar_eff <- read_csv(here("datasets", "clarifier_effluent.csv"))
raw_clar_inf <- read_csv(here("datasets", "clarifier_influent.csv"))
raw_filter_eff <- read_csv(here("datasets", "filter_effluent.csv"))
raw_filter_inf <- read_csv(here("datasets", "filter_influent.csv"))
```

```r
# Shows what is measured in clarifier data
unique(raw_clar_eff$parameter)
```

```
## [1] "TSS (mg/L)"       "BOD (mg/L)"       "COD (mg/L)"       "VSS (mg/L)"
## [5] "TKN"              "Ammonia"          "Phosphorus"       "cBOD"
## [9] "ortho-Phosphate"
```

Here are the 9 parameters measured from the clarifier data

```r
# Shows what water quality things are measured in filter data
unique(raw_filter_eff$parameter)
```

```
## [1] "TSS (mg/L)"       "VSS (mg/L)"       "BOD (mg/L)"       "COD (mg/L)"
## [5] "Ammonia"          "TKN"              "cBOD"             "ortho-Phosphate"
## [9] "Phosphorus"       "Alkalinity"       "COD (mg/kg)"
```

It appears there are 11 different water quality values measured. We should define these and have knowledge of what they mean to us. Two of them, cBOD and COD (mg/kg) aren't listed in the presentation, what are these variables? Are they just in different units from other stuff? ALSO, included in this are Alkalinity and the COD (mg/kg), which are not in the clarifier data.