Raspberry Pi Analysis

Marc Los Huertos

2025-04-02

Read Data (csv, fix dates, and subset to pm25)

Find the path to the data and read the csv

```
filepath.csv <- "/home/mwl04747/beginnersluck/EA30SP25/RPi_corrected.csv"

rpidata = read.csv(filepath.csv)

str(rpidata)
```

```
## 'data.frame':    597706 obs. of  6 variables:
##  $ Datetime          : chr  "2025-03-05 15:35:00" "2025-03-05 15:44:00" "2025-03-05 15:44:00" "2025-0
##  $ Datetime_corrected: chr  "2025-03-05 12:28:00" "2025-03-05 12:37:00" "2025-03-05 12:37:00" "2025-0
##  $ Param             : chr  "START" "temp:" "pres:" "humi:" ...
##  $ Value             : num  0 20.1 969.7 46.8 1 ...
##  $ Units             : chr  "" "Â°C" "hPa" "%" ...
##  $ ID                : chr  "PiZ1" "PiZ1" "PiZ1" "PiZ1" ...
```

```
head(rpidata)
```

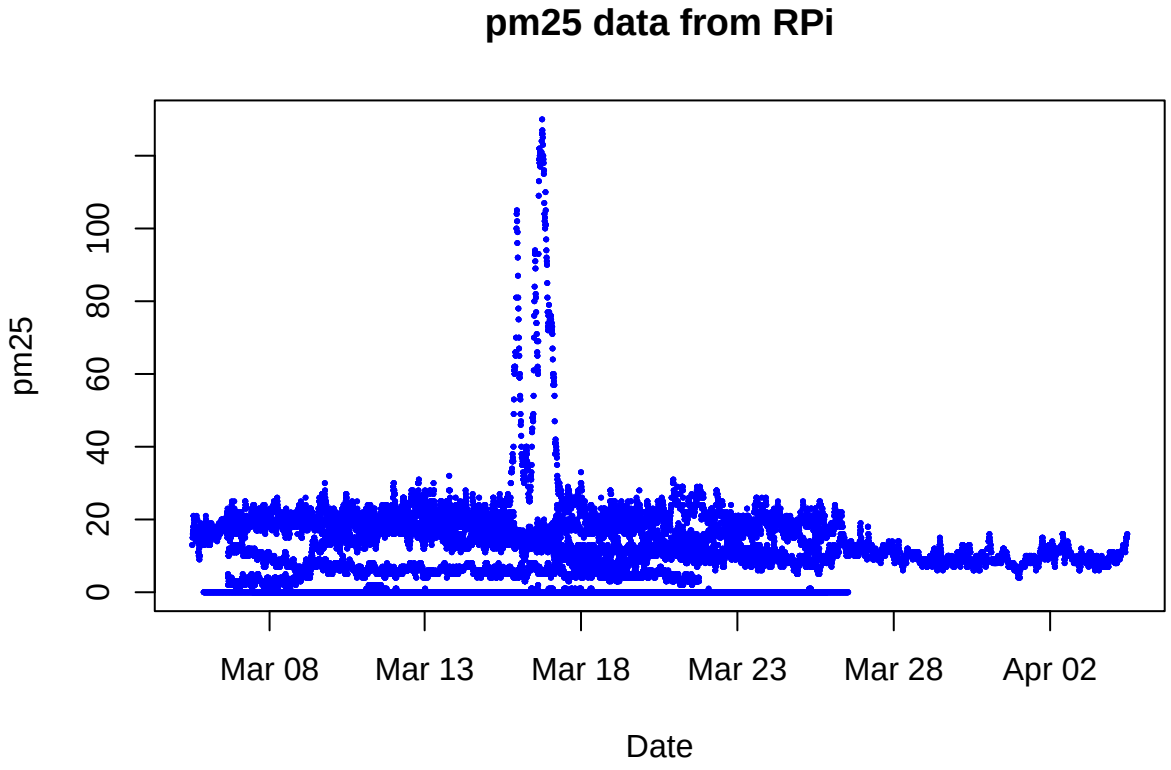
```
##              Datetime  Datetime_corrected Param Value Units   ID
## 1 2025-03-05 15:35:00 2025-03-05 12:28:00 START   0.0       PiZ1
## 2 2025-03-05 15:44:00 2025-03-05 12:37:00 temp:  20.1   Â°C PiZ1
## 3 2025-03-05 15:44:00 2025-03-05 12:37:00 pres: 969.7   hPa PiZ1
## 4 2025-03-05 15:44:00 2025-03-05 12:37:00 humi:  46.8     % PiZ1
## 5 2025-03-05 15:44:00 2025-03-05 12:37:00 ligh:   1.0   Lux PiZ1
## 6 2025-03-05 15:44:00 2025-03-05 12:37:00 oxid:  12.0    kO PiZ1
```

```
rpidata$Date <- as.POSIXct(rpidata$Datetime_corrected, format = "%Y-%m-%d %H:%M:%S")

rpidata_pm25 <- subset(rpidata, subset = Param == "pm25:")
```

Plot pm25 data

```
plot(Value ~ Date, rpidata_pm25, type = "p", col = "blue", lwd = 2, xlab = "Date", ylab = "pm25", main =
```



Create a table of location for each pi

```
IDs = unique(rpidata_pm25$ID)

rpi_mean <- aggregate(Value ~ ID, rpidata_pm25, FUN = mean)

rpi_mean$Location <- NA

rpi_mean$Location[rpi_mean$ID == "PiZ1"] <- "outside"
rpi_mean$Location[rpi_mean$ID == "PiZ2"] <- "inside"
rpi_mean$Location[rpi_mean$ID == "PiZ4"] <- "outside"
rpi_mean$Location[rpi_mean$ID == "PiZ2W2"] <- "inside"
rpi_mean$Location[rpi_mean$ID == "PiZ2W1"] <- "inside"
rpi_mean$Location[rpi_mean$ID == "PiZ2W3"] <- "outside"
rpi_mean$Location[rpi_mean$ID == "PiZ8"] <- "outside"
rpi_mean$Location[rpi_mean$ID == "PiZ9"] <- "inside"
rpi_mean$Location[rpi_mean$ID == "PiZ6"] <- "inside"
rpi_mean$Location[rpi_mean$ID == "PiZ15"] <- "outside"
rpi_mean$Location[rpi_mean$ID == "PiZ5"] <- NA


rpi_mean
```

```
##        ID        Value Location
## 1    PiZ1 17.911803279  outside
## 2   PiZ15 24.351176270  outside
## 3    PiZ2  0.000000000   inside
```

```
## 4  PiZ2W2  0.016481548   inside
## 5  PiZ2W3  6.345735027  outside
## 6    PiZ4  0.055897436  outside
## 7    PiZ5  2.000000000     <NA>
## 8    PiZ6 10.363061564   inside
## 9    PiZ8 11.648389571  outside
## 10   PiZ9  0.007385524   inside
```

```
rpi_aov <- aov(Value ~ Location, data = rpi_mean)

summary(rpi_aov)
```

```
##             Df Sum Sq Mean Sq F value Pr(>F)
## Location     1  199.1  199.12   3.149  0.119
## Residuals    7  442.7   63.24
## 1 observation deleted due to missingness
```

stuff that hasn't been finised.

f1a = data.frame(ID = IDs[1:4], factor1 = "indoor") f1a = data.frame(ID = IDs[1:4], factor1 = "indoor")

data.frame(Location = Location[ID="PiZ1"] = "outside"

IDs[1] == "outside" IDs[2] == "inside" IDs[3] == "outside" IDs[4] == "inside" IDs[5] == "outside" "'