

# Introdução à Linguagem de Programação em R para tratamento de dados de poluição do ar

openair e R do dia dia

---

Mario Gavidia-Calderón, Rafaela Squizzato, Thiago Nogueira

06/02/2024

Universidade de São Paulo

Média móvel

Juntar dados

Combinar data frames

Transformar tipos de objetos

Mapas de estações

Boas praticas do R

Introdução RMarkdown

- Nesta aula vamos falar sobre situações que acontecem quando trabalhamos com dados de qualidade do ar.
- Também falaremos sobre boas praticas de R.
- E resolver as suas dúvidas.

# Média móvel

---

# Média móvel

- O padrão qualidade do ar em São Paulo de  $O_3$  é a **média móvel de 8 horas**.
- Outro padrão da **WHO** é a **MDA8** (*average of daily maximum 8-hour*) e também **Peak season**.
- openair conta com a função `rollingMean()` para fazer esses padrões.

# Média móvel

- Vamos calcular o padrão de qualidade do ar de O<sub>3</sub>% para CETESB no 2021.

```
library(openair)
# Lendo o arquivo
pin <- readRDS('../../data/pin_openair_ex.rds')
pin_2021 <- selectByDate(pin, year = 2021)
# Média móvel
pin_2021 <- rollingMean(
  pin_2021,
  pollutant = 'o3',
  width = 8,
  new.name = "o3_8h",
  data.thresh = 0.75)
```

# Média móvel

- Foi criada a coluna o3\_8h.

##		date	aqi	wd	pm25	no	no2	o3	ws	o3_8h
## 1	2021-01-01	00:00:00	Pinheiros	346	NA	0	1	27	1.7	NA
## 2	2021-01-01	01:00:00	Pinheiros	331	NA	0	0	28	1.2	NA
## 3	2021-01-01	02:00:00	Pinheiros	338	NA	2	10	24	1.1	NA
## 4	2021-01-01	03:00:00	Pinheiros	NA	NA	0	11	18	0.0	24.25
## 5	2021-01-01	04:00:00	Pinheiros	272	NA	1	17	13	0.8	22.00
## 6	2021-01-01	05:00:00	Pinheiros	359	NA	1	12	20	1.2	23.25

# Exercício 1

A MDA8 é a média móvel **máxima diária**. Como seria calculada?



## Script do exercício 1: MDA8

```
library(openair)
pin_mda8 <- timeAverage(pin_2021[c("date", "o3_8h")],
                        avg.time = "day", # Diária
                        statistic = "max") # Máxima
head(pin_mda8)
```

```
## # A tibble: 6 x 2
##   date          o3_8h
##   <dtm>         <dbl>
## 1 2021-01-01 00:00:00 52.4
## 2 2021-01-02 00:00:00 49.6
## 3 2021-01-03 00:00:00 62.4
## 4 2021-01-04 00:00:00 67.9
```

## Exercício 2

Quantos dias foi superado o padrão de  $O_3$  na estação Pinheiros?

## Script exercício 2

Podemos usar subset. O padrão é  $130 \mu\text{gm}^{-3}$ .

```
dias_o3 <- subset(pin_mda8,  
                  subset = o3_8h >= 130)  
print(  
  paste("0 padrão foi superado ", nrow(dias_o3), "dias")  
)
```

```
## [1] "0 padrão foi superado 3 dias"
```

## Juntar dados

---

## Juntar data.frames multiples archivos

# Combinar data frames

---

## Combinar data frames

- Muitas vezes precisamos combinar duas tabelas que tem uma coluna comun. Por exemplo se temos uma tabela com a média anual de  $O_3$  das estações e outra tabela com a média anual de  $PM_{2.5}$ , para isso usamos `merge()`.

## Combinar data frames

```
o3 <- data.frame(aqs = c("pin", "ibu", "usp", "fsp"),  
                 o3 = runif(4, 0, 160))  
pm25 <- data.frame(aqs = c("pin", "ibu", "usp", "fsp"),  
                   pm25 = runif(4, 10, 60))  
dados <- merge(o3, pm25)  
dados
```

```
##   aqs      o3    pm25  
## 1 fsp  12.48010 59.15224  
## 2 ibu 120.90312 52.23340  
## 3 pin  98.50181 54.43695  
## 4 usp 125.92254 33.06977
```



## Completar dados faltantes

- Podemos usar o `merge` para completar com NA se um data frame não tem uma linha de dados.
- Precisamos adicionar o argumento `all = TRUE`.
- Serve para detectar quantos dados faltantes existem na nossa base de dados.

## Completar dados faltantes

```
o3 <- data.frame(aqs = c("pin", "ibu", "usp", "fsp"),
                 o3 = runif(4, 0, 160))
pm25 <- data.frame(aqs = c("pin", "ibu", "usp"),
                  pm25 = runif(3, 10, 60))
dados <- merge(o3, pm25, all = TRUE)
dados
```

```
##   aqs      o3    pm25
## 1 fsp  16.29005      NA
## 2 ibu 136.59050 41.10601
## 3 pin  27.31374 52.52071
## 4 usp  87.83188 12.80002
```

# Transformar tipos de objetos

---

## character para numeric

- Às vezes quando usamos `read.table` ou `read_excel`, se uma coluna tem o dado faltante como um character especial (e.g “M”). Para poder operar precisamos forçar a transformação para `numeric`.

## character para numeric

```
o3 <- c(90, 89, 76, 83, "M")  
class(o3)
```

```
## [1] "character"
```

```
o3 <- as.numeric(o3) # Atualizamos o valor do vetor o3
```

```
## Warning: NAs introduced by coercion
```

```
class(o3)
```

```
## [1] "numeric"
```

```
o3
```

# Mapas de estações

---

# Boas praticas do R

---

# Boas praticas do R

- Para cada trabalho usar RStudio Projects:
  - Mais fácil de compartilhar. Não precisa definir o caminho dos inputs.
- É bom olhar guias de estilo do R.



# Introdução RMarkdown

---

# Pratica

---

**Mais recursos.**

---