

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

Mario Gavidia-Calderón, Rafaela Squizzato, Thiago Nogueira

06/02/2024

Universidade de São Paulo

openair

Introdução

Instalar e desinstalar pacotes

Importando os dados

4. Organizando os dados

7. Funções do pacote openair

openair

Manual do openair

Introdução

- Software gratuito, muito utilizado entre a comunidade científica;
- Muito material disponível na internet;
- Ótimo para armazenar e manipular dados; realizar cálculos e testes estatísticos; produção de gráficos etc;
- Ótima opção para quem trabalha com banco de dados grandes; otimização do tempo.

R ou Rstudio?

- **Rstudio** - interface mais amigável e funcional;
- Para usar o Rstudio é necessário ter instalado no computador o software R.
- <https://posit.co/download/rstudio-desktop/>

Relembrando:

- Menu de ajuda: `?sum`, `help(sum)`
- R é sensível a maiúscula e minúsculas, ou seja, `"a" ≠ "A"`.
- Nunca deixe espaços entre nomes de objetos no R (ex: `o3 noturno`), coloque símbolos no lugar (ex: `o3_noturno`).
- Não é possível usar um número isolado para nomear objetos.
- Evite usar acentos, cedilhas, apóstrofes, aspas, etc. em nomes de objetos no R.

Instalar e desinstalar pacotes

Instalar e desinstalar pacotes

```
install.packages("openair")  
remove.packages("openair")  
rm(a) #remove o data.frame "a"  
rm(list=ls()) #remove todos os data.frames no ambiente
```

Para que o pacote funcione ele precisa ser carregado

```
library(openair)  
library(readxl) #para ler arquivos em excel (.xls)
```

Importando os dados

```
file <- 'G:/Meu Drive/QUALAMET_2023/QUALAMET_FMUSP_gases.xlsx'  
dados <- read_excel(file)
```

OBS:

Inverter a barra \ para /.

Em .csv

```
# Definir diretório trabalho
setwd("G:/Meu Drive/QUALAMET_2023/NH3")
dado <- read.csv("nh3_all_hour.csv", header = TRUE)
#ou
dado <-
  read.csv("G:/Meu Drive/QUALAMET_2023/NH3/nh3_all_hour.csv",
           header = TRUE)
getwd() # mostra o diretório de trabalho atual
```

OBS:

O argumento `header = FALSE` indica que os dados não contêm cabeçalho. Caso contrário, use `header = TRUE`.

3.3 Usando o qualR

```
install.packages('qualR',  
  repos = c('https://ropensci.r-universe.dev',  
            'https://cloud.r-project.org'))
```

```
library(qualR)
```

```
cetesb_aqs #nome,código, lat lon
```

```
cetesb_param #parâmetro, unidade e código
```

```
my_user_name <- "e-mail"
```

```
my_password <- "senha"
```

```
start_date <- "01/01/2020"
```

```
end_date <- "31/12/2021"
```

1º Exercício

- 1) Baixar a partir do pacote qualR as variáveis O3, PM2.5, VV e DV para a estação Pinheiros, para os anos de 2020 e 2021.


```
dado$wd<- replace(dado$wd,  dado$wd > 360, NA)
```

4. Organizando os dados

Organizando os dados

```
dado$date <- as.POSIXct(strptime(dado$date, format = "%Y-%m-%d %H:%M:%S",  
                                tz = "UTC"))
```

```
dado$hour <- as.numeric((dado$hour)) #transformando a coluna em numérico
```

```
names(dado) <- c("date", "estacao", "wd", "pm25", "o3", "ws") #trocar o nome
```

```
dado["hour"] <- format(dado$date, "%H") #criando coluna só com hora
```

```
novo <- dado[,c(1,5)] #novo dataframe com as variáveis selecionadas
```

```
dado$aqs <- "Pinheiros" # criou nova coluna
```

`cbind` # faz a junção dos data.frames pela coluna

`rbind` # faz a junção dos data.frames pela linha

```
df4 <- cbind(df1,df2,df3)
```

7. Funções do pacote openair

7. Funções do pacote openair

As funções presentes no pacote Openair são dedicadas à análise de dados de poluição atmosférica.

As funções usam como padrão uma forma de análise mais simples e rápida, porém análises mais detalhadas também são possíveis.

Entre as principais funções, temos: `summaryPlot`, `timePlot`, `calendarPlot`, `timeVariation`, `windRose`, `percentilRose`, `polarPlot`, `polarAnnulus`, `scatterPlot`, `corPlot`.

7.1 `summaryPlot`

Interessante para monitoramento, pois resume rapidamente aspectos importantes dos dados, apresentando resumos estatísticos (mín, máx, média, mediana etc).

Comando simples: `summaryPlot(dado)`

```
dado$wd <- replace(dado$wd, dado$wd > 360, NA)
```