

Lista de Exercícios 1

Thaís Paiva

19/03/2018

Comandos Básicos no R

Exercício 1

O objeto `data.table` herda as funcionalidade de um objeto `data.frame`, mas oferece acesso mais rápido e mais eficiente. A diferença principal é a maneira de acessar os objetos. Com um objeto `data.table`, é possível realizar uma operação do tipo:

```
DT[i, j, by]
```

que significa pegar a tabela `DT`, selecionar as linhas `i`, e computar `j` agrupado por `by`. Esse mesmo comando exigiria uma sintaxe mais complexa ao trabalhar com objetos do tipo `data.frame`.

Exercício 2

É possível imprimir o valor de $\log(4)$ com o comando abaixo:

```
print( log(4), digits = 15)
```

```
## [1] 1.38629436111989
```

Exercício 3

A função `intersect` encontra a interseção entre dois vetores, ou seja, os elementos que estão em ambos.

```
seq(4,28,by=7)
```

```
## [1] 4 11 18 25
```

```
seq(3,31,by=2)
```

```
## [1] 3 5 7 9 11 13 15 17 19 21 23 25 27 29 31
```

```
intersect( seq(4,28,by=7), seq(3,31,by=2) )
```

```
## [1] 11 25
```

Exercício 4

O comando abaixo irá fazer a operação **E** elemento a elemento dos dois vetores de tamanho quatro. Como os vetores são da classe `logical`, o resultado da operação será `TRUE` se ambos os elementos forem `TRUE`, e `FALSE` caso contrário.

```
c(TRUE,TRUE,FALSE,FALSE) & c(TRUE,FALSE,FALSE,TRUE)
```

```
## [1] TRUE FALSE FALSE FALSE
```

Importando Dados

Exercício 5

Primeiro, vamos carregar a base de dados completa.

```
file = "../datasets/extremedatasince1899.csv"
StormMax = read.table(file, header=T, sep=",")
head(StormMax, 3)
```

```
##      Yr Region      Wmax      nao      soi      sst      sstmda sun split
## 1 1899   Basin 96.64138 -0.64 -0.21 0.05193367 -0.03133333 8.4      0
## 2 1899   East 90.19791 -0.64 -0.21 0.05193367 -0.03133333 8.4      0
## 3 1899   Basin 90.35300 -0.64 -0.21 0.05193367 -0.03133333 8.4      0
```

Agora podemos selecionar os eventos que ocorreram apenas no século XX. São 1782 dados de ciclones que ocorreram entre os anos 1900 e 2000.

```
sec20 = subset(StormMax, subset=StormMax$Yr>1900 & StormMax$Yr<=2000)
dim(sec20)
```

```
## [1] 1782      9
```

Exercício 6

Para ordenar os dados em ordem crescente da velocidade do vento (variável Wmax), podemos executar o seguinte comando:

```
sec20 = sec20[ order(sec20$Wmax),]
head(sec20)
```

```
##      Yr Region      Wmax      nao      soi      sst      sstmda sun
## 1553 1987   East 10.00000 -0.695 -1.1633333 0.30893367 0.29500000 33.9
## 1559 1987    US 10.00000 -0.695 -1.1633333 0.30893367 0.29500000 33.9
## 1541 1986 Florida 10.01247  0.435 -0.2800000 -0.13673300 0.06766667  3.8
## 1618 1990   East 15.00000 -0.385 -0.5533333 0.24493367 0.28233333 125.2
## 1078 1960   Gulf 15.00052 -0.170  0.3266667 0.33193367 0.05566667 127.2
## 1678 1994   East 15.00240  0.780 -1.7700000 0.01293367 0.12233333  25.7
##      split
## 1553      1
## 1559      1
## 1541      1
## 1618      1
## 1078      1
## 1678      1
```

```
tail(sec20)
```

```
##      Yr Region      Wmax      nao      soi      sst      sstmda sun
##  877 1950   Basin 160.8074 -0.235  1.1500000 0.004600333 -0.09100000  51.3
## 1585 1988   Basin 160.8934 -1.995  1.5433333 0.047933667 0.09000000 120.1
## 1438 1980   Basin 165.0000 -1.230 -0.3133333 0.093267000 0.01533333 155.0
## 1226 1969   Basin 171.3285 -0.700 -1.0533333 -0.055399667 0.04400000  91.3
## 1236 1969   Gulf 171.3285 -0.700 -1.0533333 -0.055399667 0.04400000  91.3
## 1239 1969    US 171.3285 -0.700 -1.0533333 -0.055399667 0.04400000  91.3
##      split
##  877      1
```

```
## 1585    1
## 1438    1
## 1226    1
## 1236    1
## 1239    1
```

Exercício 7

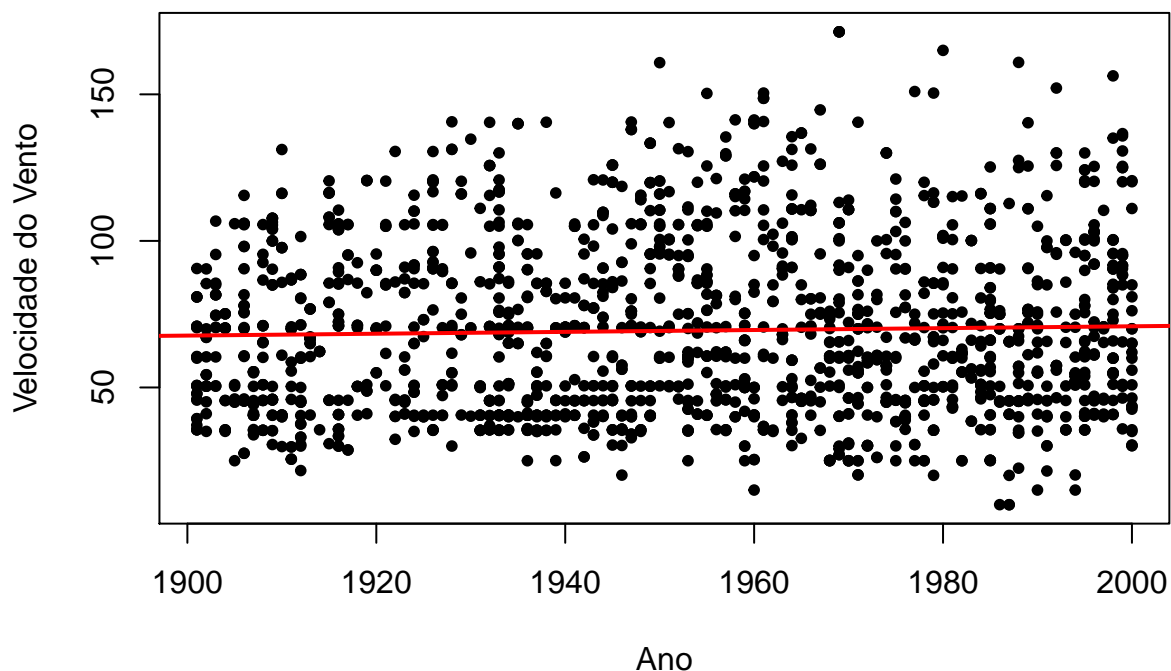
```
media = mean(sec20$Wmax[sec20$Region=="Basin"])
```

A média da velocidade do vento Wmax na região Basin é 74.9437991.

Exercício 8

O gráfico de dispersão entre o ano e a velocidade do vento pode ser gerado com os comandos abaixo. Adicionamos também uma linha com a regressão ajustada entre as duas variáveis. Notamos que não há nenhuma tendência de crescimento ou decrescimento da velocidade ao longo do século XX.

```
plot(sec20$Yr, sec20$Wmax, pch=20, xlab="Ano", ylab="Velocidade do Vento")
abline(lm(sec20$Wmax ~ sec20$Yr), col=2, lwd=2)
```



O gráfico também poderia ser gerado usando o pacote `ggplot2`, que facilita a inclusão de uma linha de tendência suavizada.

```
library(ggplot2)
ggplot(sec20, aes(Yr, Wmax)) + geom_point() + geom_smooth(method="lm")
```

