

Iniciando Nossos Trabalhos

James R. Hunter

12 de fevereiro de 2017

MAD-CB

Figure 1:

- Dados médicos sobre tratamentos para artrite e inflamação
- Com isso, podemos aprender
 - ▶ Como inserir dados nos objetos R
 - ▶ Como manipular esses dados
 - ▶ Gráficos básicos
 - ▶ Resumos descritivos de dados
 - ▶ Funções em R
 - ▶ Loops
- Muito deste lição e os dados vêm do site “Software Carpentry”
 - ▶ Agradeço eles

- Queremos estar trabalhando na pasta para matéria
- Pode usar a função `getwd()` para ver o que é o working directory

```
## [1] "/Users/James/Documents/UNIFESP/MAD Course/MAD-CBt1"
```

Dados para Working Directory

- Coloque os dados no working directory
- Salvé-los da Github para o working directory
- Nome de dados: “r-novice-inflammation-data.zip”
- No OS, expande o arquivo ‘zip’ com um dupla-clique
- Vai criar uma nova pasta chamada “data”
- Mude o nome desta pasta para “artrite_data”

- Precisa abrir alguns pacotes para ler e manipular os dados
- Uso da função `library()`
- Dados são carregados no disco no formato “.csv”
- “.csv” (“Comma Separated Values”) – Formato de Excel
- A função `read_csv()` faz parte do pacote `readr`
- `readr` faz parte do “tidyverse”
 - ▶ Podemos chamar isso e outros carregando o pacote `tidyverse`

Pacotes Carregados com o Pacote tidyverse

```
sessionInfo() ## Comando para mostrar o estado do sistema R neste momento
```

```
## R version 3.3.2 (2016-10-31)
## Platform: x86_64-apple-darwin13.4.0 (64-bit)
## Running under: OS X El Capitan 10.11.6
##
## locale:
## [1] en_US.UTF-8/en_US.UTF-8/en_US.UTF-8/C/en_US.UTF-8/en_US.UTF-8
##
## attached base packages:
## [1] stats      graphics  grDevices  utils      datasets  methods   base
##
## loaded via a namespace (and not attached):
## [1] backports_1.0.5 magrittr_1.5      rprojroot_1.2    tools_3.3.2
## [5] htmltools_0.3.5 yaml_2.1.14      Rcpp_0.12.9      stringi_1.1.2
## [9] rmarkdown_1.3   knitr_1.15.1     stringr_1.1.0    digest_0.6.12
## [13] evaluate_0.10
```

```
library(tidyverse)
```

```
## Loading tidyverse: ggplot2
## Loading tidyverse: tibble
## Loading tidyverse: tidyr
## Loading tidyverse: readr
## Loading tidyverse: purrr
## Loading tidyverse: dplyr
```

```
## Conflicts with tidy packages -----
```

```
## filter(): dplyr, stats
## lag():    dplyr, stats
```



```
## R version 3.3.2 (2016-10-31)
## Platform: x86_64-apple-darwin13.4.0 (64-bit)
## Running under: OS X El Capitan 10.11.6
##
## locale:
## [1] en_US.UTF-8/en_US.UTF-8/en_US.UTF-8/C/en_US.UTF-8/en_US.UTF-8
##
## attached base packages:
## [1] stats      graphics  grDevices  utils      datasets  methods   base
##
## other attached packages:
## [1] dplyr_0.5.0      purrr_0.2.2      readr_1.0.0      tidyr_0.6.1
## [5] tibble_1.2       ggplot2_2.2.1    tidyverse_1.1.1
##
## loaded via a namespace (and not attached):
## [1] Rcpp_0.12.9      plyr_1.8.4       forcats_0.2.0    tools_3.3.2
## [5] digest_0.6.12    jsonlite_1.2     lubridate_1.6.0  evaluate_0.10
## [9] nlme_3.1-130     gtable_0.2.0     lattice_0.20-34  psych_1.6.12
## [13] DBI_0.5-1        yaml_2.1.14      parallel_3.3.2   haven_1.0.0
## [17] xml2_1.1.1       stringr_1.1.0    httr_1.2.1       knitr_1.15.1
## [21] hms_0.3          rprojroot_1.2    grid_3.3.2       R6_2.2.0
## [25] readxl_0.1.1     foreign_0.8-67   rmarkdown_1.3    modelr_0.1.0
## [29] reshape2_1.4.2   magrittr_1.5     backports_1.0.5  scales_0.4.1
## [33] htmltools_0.3.5  rvest_0.3.2      assertthat_0.1   mnormt_1.5-5
## [37] colorspace_1.3-2 stringi_1.1.2     lazyeval_0.2.0   munsell_0.4.3
```

Carregar os Dados em Memória

```
dados <- read_csv(file = "artrite_data/inflammation-01.csv",  
                  col_names = FALSE)
```

```
## Parsed with column specification:
```

```
## cols(  
##   .default = col_integer()  
## )
```

```
## See spec(...) for full column specifications.
```

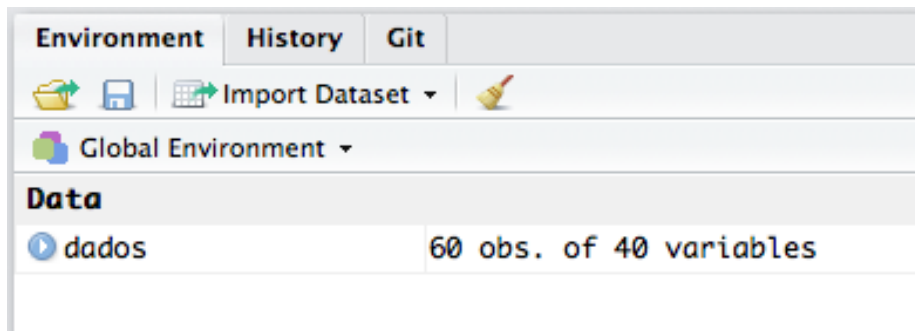


Figure 2:

O Que Significa Dados

- Dando um nome para os dados na planilha
- `<-` = Assignment
- `<-` \neq = (igual)
- `<-` quer dizer que o valor a direita está sendo associado ao nome a esquerda

Exemplo Mais Simples

```
## 1ª Versão  
peso <- 55 ## Pessoa pesa 55 kg.
```

```
## 2ª Versão  
peso_kg <- 55 ## Mais claro
```

```
## Pode Converter à Libra  
(peso_lb <- peso_kg * 2.2)
```

```
## [1] 121
```

```
peso_lb
```

```
## [1] 121
```

Qual Tipo de Dados Temos?

- Pacientes são linhas
- Dias são colunas

```
class(dados)
```

```
## [1] "tbl_df"      "tbl"          "data.frame"
```

```
dim(dados)
```

```
## [1] 60 40
```

Pegar Dados Específicos do Conjunto

```
dados[1, 1]
```

```
## # A tibble: 1 × 1
##       X1
##   <int>
## 1     0
```

```
dados[20, 20] # valor de linha 20, coluna 20
```

```
## # A tibble: 1 × 1
##       X20
##   <int>
## 1    16
```

```
dados[1:3, 1:5] # valores dos primeiros 3 pacientes para primeiro 5 dias
```

```
## # A tibble: 3 × 5
##       X1     X2     X3     X4     X5
##   <int> <int> <int> <int> <int>
## 1     0     0     1     3     1
## 2     0     1     2     1     2
## 3     0     1     1     3     3
```