

Pesquisa
repro-
duzível

Frederico
Bertholini

Nunca
esqueça

Recapitulando

Pesquisa
repro-
duzível

Versionando
projetos

Gerenciamento
de arquivos

ggplot

Pesquisa reproduzível

Métodos Quantitativos Aplicados à Ciência Política

Frederico Bertholini

05.out.2020

Pesquisa
repro-
duzível

Frederico
Bertholini

Nunca
esqueça

Recapitulando

Pesquisa
repro-
duzível

Versionando
projetos

Gerenciamento
de arquivos

ggplot

1 Nunca esqueça

2 Recapitulando

3 Pesquisa reproduzível

4 Versionando projetos

5 Gerenciamento de arquivos

6 ggplot

Pesquisa
repro-
duzível

Frederico
Bertholini

Nunca
esqueça

Recapitulando

Pesquisa
repro-
duzível

Versionando
projetos

Gerenciamento
de arquivos

ggplot

Nunca esqueça

Pacotes e diretório de trabalho

Pesquisa
repro-
duzível

Frederico
Bertholini

Nunca
esqueça

Recapitulando

Pesquisa
repro-
duzível

Versionando
projetos

Gerenciamento
de arquivos

ggplot

```
# rotina para carregar pacotes
```

```
lista.de.pacotes = c("tidyverse", "haven", "lubridate",  
                     "janitor", "readxl",  
                     "stringr", "magrittr",  
                     "survey", "srvyr")
```

```
novos.pacotes <-
```

```
  lista.de.pacotes[!(lista.de.pacotes %in%  
                     installed.packages()[, "Package"])]
```

```
if(length(novos.pacotes) > 0) {install.packages(novos.pacotes)}
```

```
lapply(lista.de.pacotes, require, character.only=T)
```

```
rm(list = ls())
```

```
gc()
```

```
# Definindo o diretorio de trabalho como do arquivo local
```

```
setwd(dirname(rstudioapi::getActiveDocumentContext()$path))
```

Pesquisa
repro-
duzível

Frederico
Bertholini

Nunca
esqueça

Recapitulando

Pesquisa
repro-
duzível

Versionando
projetos

Gerenciamento
de arquivos

ggplot

Recapitulando

Live coding de manipulação de dados

Pesquisa
repro-
duzível

Frederico
Bertholini

Nunca
esqueça

Recapitulando

Pesquisa
repro-
duzível

Versionando
projetos

Gerenciamento
de arquivos

ggplot

Reproduzindo trabalho sobre COVID e estados

**Pesquisa
repro-
duzível**

**Frederico
Bertholini**

Nunca
esqueça

Recapitulando

**Pesquisa
repro-
duzível**

Versionando
projetos

Gerenciamento
de arquivos

ggplot

Pesquisa reproduzível

Por quê?

Pesquisa
repro-
duzível

Frederico
Bertholini

Nunca
esqueça

Recapitulando

Pesquisa
repro-
duzível

Versionando
projetos

Gerenciamento
de arquivos

ggplot

■ Pra ciência

■ Pra você

Ferramentas

Pesquisa
repro-
duzível

Frederico
Bertholini

Nunca
esqueça

Recapitulando

Pesquisa
repro-
duzível

Versionando
projetos

Gerenciamento
de arquivos

ggplot

- R e RStudio (ok)
- Github
- knitr e rmarkdown
- LaTeX

Fluxo de trabalho

Pesquisa
repro-
duzível

Frederico
Bertholini

Nunca
esqueça

Recapitulando

Pesquisa
repro-
duzível

Versionando
projetos

Gerenciamento
de arquivos

ggplot

1 Coleta

2 Análise

3 Comunicação

Pesquisa
repro-
duzível

Frederico
Bertholini

Nunca
esqueça

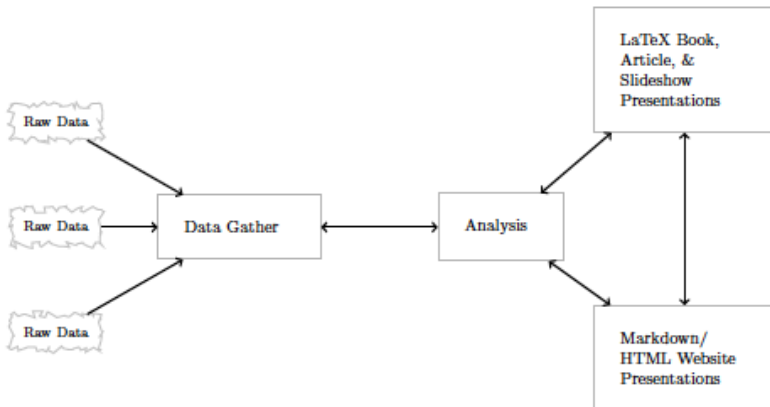
Recapitulando

Pesquisa
repro-
duzível

Versionando
projetos

Gerenciamento
de arquivos

ggplot



Dicas

Pesquisa
repro-
duzível

Frederico
Bertholini

Nunca
esqueça

Recapitulando

Pesquisa
repro-
duzível

Versionando
projetos

Gerenciamento
de arquivos

ggplot

- 1 Documento tudo!
- 2 Tudo é um arquivo (de texto).
- 3 Todos os arquivos devem ser legíveis (por humanos).
- 4 Relacione explicitamente seus arquivos.
- 5 Tenha um plano para organizar, armazenar e disponibilizar seus arquivos.

Pesquisa
repro-
duzível

Frederico
Bertholini

Nunca
esqueça

Recapitulando

Pesquisa
repro-
duzível

**Versionando
projetos**

Gerenciamento
de arquivos

ggplot

Versionando projetos

Pesquisa
repro-
duzível

Frederico
Bertholini

Nunca
esqueça

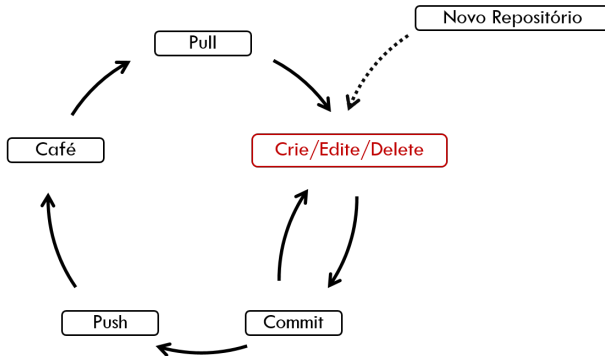
Recapitulando

Pesquisa
repro-
duzível

Versionando
projetos

Gerenciamento
de arquivos

ggplot



Passo-a-passo

Pesquisa
repro-
duzível

Frederico
Bertholini

Nunca
esqueça

Recapitulando

Pesquisa
repro-
duzível

Versionando
projetos

Gerenciamento
de arquivos

ggplot

- 1 Repositório: Criação de repositório do projeto no Github
- 2 .Rproj: Criação do Projeto no RStudio
- 3 Commit: Editando e “Commitando” as mudanças no código
- 4 Push: Subindo os commits para o Github
- 5 Pull: Baixando o estado atual do projeto

Pesquisa
repro-
duzível

Frederico
Bertholini

Nunca
esqueça

Recapitulando

Pesquisa
repro-
duzível

Versionando
projetos

Gerenciamento
de arquivos

ggplot

Gerenciamento de arquivos

Pesquisa
repro-
duzível

Frederico
Bertholini

Nunca
esqueça

Recapitulando

Pesquisa
repro-
duzível

Versionando
projetos

Gerenciamento
de arquivos

ggplot

ggplot

O que você quer mostrar?

Pesquisa
repro-
duzível

Frederico
Bertholini

Nunca
esqueça

Recapitulando

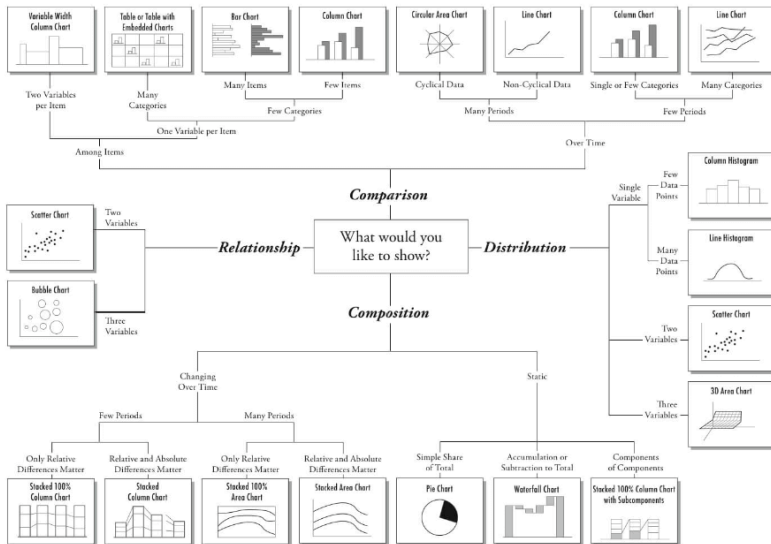
Pesquisa
repro-
duzível

Versionando
projetos

Gerenciamento
de arquivos

ggplot

Chart Suggestions—A Thought-Starter



Princípios

Pesquisa
repro-
duzível

Frederico
Bertholini

Nunca
esqueça

Recapitulando

Pesquisa
repro-
duzível

Versionando
projetos

Gerenciamento
de arquivos

ggplot

- O que você quer mostrar?
 - Elementos que podem **destacar** ou **confundir** o que você quer mostrar.
- Exemplo Codeplan
- vamos tentar alternar “teoria” com live code

Recursos

Pesquisa
repro-
duzível

Frederico
Bertholini

Nunca
esqueça

Recapitulando

Pesquisa
repro-
duzível

Versionando
projetos

Gerenciamento
de arquivos

ggplot

- R Cookbook <http://www.cookbook-r.com/Graphs/>
- STHDA <http://www.sthda.com/english/wiki/be-awesome-in-ggplot2-a-practical-guide-to-be-highly-effective-r-software-and-data-visualization>
- R Graph Gallery <https://www.r-graph-gallery.com/>
- Extensões <http://www.ggplot2-exts.org/>

Elementos do ggplot

Pesquisa
repro-
duzível

Frederico
Bertholini

Nunca
esqueça

Recapitulando

Pesquisa
repro-
duzível

Versionando
projetos

Gerenciamento
de arquivos

ggplot

- Dados
- Geometrias
- Estéticas
- Escalas (estética)
- Escalas (eixos)
- Tema
- Facet

Dados data =

Pesquisa
repro-
duzível

Frederico
Bertholini

Nunca
esqueça

Recapitulando

Pesquisa
repro-
duzível

Versionando
projetos

Gerenciamento
de arquivos

ggplot

- Dado empilhado?
- Cada coluna será uma entrada!

Geometrias geom_

Pesquisa
repro-
duzível

Frederico
Bertholini

Nunca
esqueça

Recapitulando

Pesquisa
repro-
duzível

Versionando
projetos

Gerenciamento
de arquivos

ggplot

- `geom_tipo_de_geometria`
- Recursos +
- cheat sheet <https://www.rstudio.com/wp-content/uploads/2016/03/ggplot2-cheatsheet-portuguese.pdf>
- manual ggplot <https://ggplot2.tidyverse.org/reference/>

Estéticas `aes()`

Pesquisa
repro-
duzível

Frederico
Bertholini

Nunca
esqueça

Resumindo

Pesquisa
repro-
duzível

Versionando
projetos

Gerenciamento
de arquivos

ggplot

■ `x` (`xmax` e `xmin`)

■ `y`

■ `color`

■ `fill`

■ `shape`

■ `group`

■ `size`

Escalas (estética) scale_

Pesquisa
repro-
duzível

Frederico
Bertholini

Nunca
esqueça

Recapitulando

Pesquisa
repro-
duzível

Versionando
projetos

Gerenciamento
de arquivos

ggplot

■ `scale_color_xx`

■ `scale_fill_xx`

■ `scale_shape_xx`

Escalas (eixos) `scale_x`

Pesquisa
repro-
duzível

Frederico
Bertholini

Nunca
esqueça

Recapitulando

Pesquisa
repro-
duzível

Versionando
projetos

Gerenciamento
de arquivos

ggplot

- Contínua `scale_x_continuous`
- Discreta `scale_x_discrete`
- Série de tempo zoo → `scale_yearmon`

Tema

Pesquisa
repro-
duzível

Frederico
Bertholini

Nunca
esqueça

Recapitulando

Pesquisa
repro-
duzível

Versionando
projetos

Gerenciamento
de arquivos

ggplot

- Customização total da visualização
- Eixos
- Texto `element_text`
- linhas de grade

Facet

Pesquisa
repro-
duzível

Frederico
Bertholini

Nunca
esqueça

Recapitulando

Pesquisa
repro-
duzível

Versionando
projetos

Gerenciamento
de arquivos

ggplot

■ facet_grid

■ facet_wrap

Gráficos com interatividade

Pesquisa
repro-
duzível

Frederico
Bertholini

Nunca
esqueça

Recapitulando

Pesquisa
repro-
duzível

Versionando
projetos

Gerenciamento
de arquivos

ggplot

■ ggiraph

■ plotly (ggplotly)

Exercício

Pesquisa
repro-
duzível

Frederico
Bertholini

Nunca
esqueça

Recapitulando

Pesquisa
repro-
duzível

Versionando
projetos

Gerenciamen-
to de arquivos

ggplot

- Carregue os dados de exemplo do pacote `survey data(api)`, use o `data.frame` `apisrs`
- Crie o objeto `tbl_svy amostra_expandida` expandindo a amostra aleatória simples usando apenas a variável (coluna) “`pw`”, contendo o peso amostral. Dica: execute `as_survey(weight=pw)`.
- Usando a variável `stype` crie uma nova variável indicando se a escola é de nível fundamental (categorias **E** e **M** de `stype`) ou de nível médio (categoria *H* de `stype`). Dica: use `mutate` e `case_when`.
- Faça um gráfico de barras comparando a variação média das notas de 1999 (`api99`) e 2000 (`api00`) por nível e utilize as estimativas intervalares. Dica: olhe o código da aula 07, utilize `geom_errorbar` para a estimativa intervalar.

Resolução

Pesquisa
repro-
duzível

Frederico
Bertholini

Nunca
esqueça

Recapitulando

Pesquisa
repro-
duzível

Versionando
projetos

Gerenciamento
de arquivos

ggplot

```
data(api)

amostra_expandida <- apisrs %>%
  as_survey(weight = pw) %>%
  mutate(nivel=case_when(
    stype=="E"~"Fundamental",
    stype=="M"~"Fundamental",
    stype=="H"~"Médio"
  ))
```

Pesquisa
repro-
duzível

Frederico
Bertholini

Nunca
esqueça

Recapitulando

Pesquisa
repro-
duzível

Versionando
projetos

Gerenciamento
de arquivos

ggplot

```
out <- amostra_expandida %>%  
  group_by(nivel) %>%  
  summarise(api_diff =  
    survey_mean(api00 - api99, vartype = "ci"))
```


Pesquisa
repro-
duzível

Frederico
Bertholini

Nunca
esqueça

Recapitulando

Pesquisa
repro-
duzível

Versionando
projetos

Gerenciamento
de arquivos

ggplot

```
grafico <- out %>%  
  ggplot(aes(x = nivel, y = api_diff,  
             fill = nivel,color=nivel,  
             ymax = api_diff_upp,  
             ymin = api_diff_low)) +  
  geom_bar(stat = "identity",alpha=0.6) +  
  geom_errorbar(width = 0,size=3)
```

Pesquisa
repro-
duzível

Frederico
Bertholini

Nunca
esqueça

Resumindo

Pesquisa
repro-
duzível

Versionando
projetos

Gerenciamento
de arquivos

ggplot

grafico

