

**Pesquisa
repro-
duzível**

**Frederico
Bertholini**

Nunca
esqueça

Recapitulando
manipu-
lação

Pesquisa
repro-
duzível

Versionando
projetos

ggplot

Pesquisa reproduzível

Métodos Quantitativos Aplicados à Ciência Política

Frederico Bertholini

05.out.2020

Pesquisa
repro-
duzível

Frederico
Bertholini

Nunca
esqueça

Recapitulando
manipu-
lação

Pesquisa
repro-
duzível

Versionando
projetos

ggplot

1 Nunca esqueça

2 Recapitulando manipulação

3 Pesquisa reproduzível

4 Versionando projetos

5 ggplot

Pesquisa
repro-
duzível

Frederico
Bertholini

Nunca
esqueça

Recapitulando
manipu-
lação

Pesquisa
repro-
duzível

Versionando
projetos

ggplot

Nunca esqueça

Pacotes e diretório de trabalho

Pesquisa
repro-
duzível

Frederico
Bertholini

Nunca
esqueça

Recapitulando
manipu-
lação

Pesquisa
repro-
duzível

Versionando
projetos

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
manipu-
lação

Pesquisa
repro-
duzível

Versionando
projetos

ggplot

Recapitulando manipulação

Live coding de manipulação de dados

Pesquisa
repro-
duzível

Frederico
Bertholini

Nunca
esqueça

Recapitulando
manipu-
lação

Pesquisa
repro-
duzível

Versionando
projetos

ggplot

Reproduzindo trabalho sobre COVID e estados

**Pesquisa
repro-
duzível**

**Frederico
Bertholini**

Nunca
esqueça

Recapitulando
manipu-
lação

**Pesquisa
repro-
duzível**

Versionando
projetos

ggplot

Pesquisa reproduzível

Por quê?

Pesquisa
repro-
duzível

Frederico
Bertholini

Nunca
esqueça

Recapitulando
manipu-
lação

Pesquisa
repro-
duzível

Versionando
projetos

ggplot

■ Pra ciência

■ Pra você

Ferramentas

Pesquisa
repro-
duzível

Frederico
Bertholini

Nunca
esqueça

Recapitulando
manipu-
lação

Pesquisa
repro-
duzível

Versionando
projetos

ggplot

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

Fluxo de trabalho

Pesquisa
repro-
duzível

Frederico
Bertholini

Nunca
esqueça

Recapitulando
manipu-
lação

Pesquisa
repro-
duzível

Versionando
projetos

ggplot

- 1 Coleta
- 2 Análise
- 3 Comunicação

Pesquisa
repro-
duzível

Frederico
Bertholini

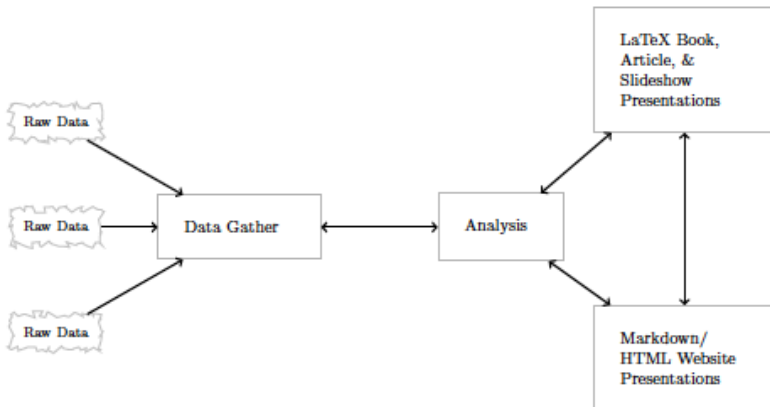
Nunca
esqueça

Recapitulando
manipu-
lação

Pesquisa
repro-
duzível

Versionando
projetos

ggplot



Dicas

Pesquisa
repro-
duzível

Frederico
Bertholini

Nunca
esqueça

Recapitulando
manipu-
lação

Pesquisa
repro-
duzível

Versionando
projetos

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.

Trabalhando com projetos

Pesquisa
repro-
duzível

Frederico
Bertholini

Nunca
esqueça

Recapitulando
manipu-
lação

Pesquisa
repro-
duzível

Versionando
projetos

ggplot

Gerenciamento de arquivos: caminhos relativos e não absolutos

- Onde guardar seus projetos?

Pesquisa
repro-
duzível

Frederico
Bertholini

Nunca
esqueça

Recapitulando
manipu-
lação

Pesquisa
repro-
duzível

Versionando
projetos

ggplot

Versionando projetos

Pesquisa
repro-
duzível

Frederico
Bertholini

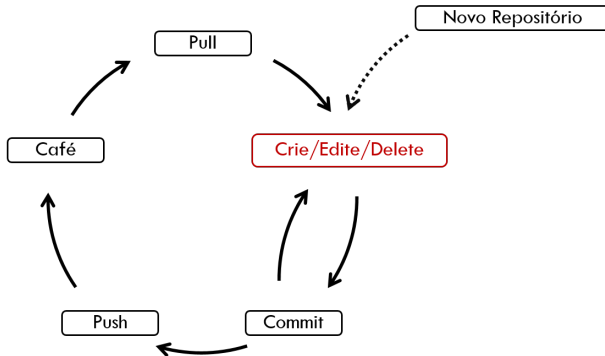
Nunca
esqueça

Recapitulando
manipu-
lação

Pesquisa
repro-
duzível

Versionando
projetos

ggplot



Passo-a-passo

Pesquisa
repro-
duzível

Frederico
Bertholini

Nunca
esqueça

Recapitulando
manipu-
lação

Pesquisa
repro-
duzível

Versionando
projetos

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

Caso tenha dificuldade

Pesquisa
repro-
duzível

Frederico
Bertholini

Nunca
esqueça

Recapitulando
manipu-
lação

Pesquisa
repro-
duzível

Versionando
projetos

ggplot

Beatriz Milz - RLadies

Curso R

Pesquisa
repro-
duzível

Frederico
Bertholini

Nunca
esqueça

Recapitulando
manipu-
lação

Pesquisa
repro-
duzível

Versionando
projetos

ggplot

ggplot

Uma exibição gráfica deve (1/2)

Pesquisa
repro-
duzível

Frederico
Bertholini

Nunca
esqueça

Recapitulando
manipu-
lação

Pesquisa
repro-
duzível

Versionando
projetos

ggplot

Mostrar os dados

Induzir o observador a pensar em sua substância, não em metodologia ou tecnologia de produção

Evitar distorcer o que os dados dizem

Apresentar muitos números em pequenos espaços

Tornar grandes conjuntos de dados coerentes

Uma exibição gráfica deve (2/2)

Pesquisa
repro-
duzível

Frederico
Bertholini

Nunca
esqueça

Recapitulando
manipu-
lação

Pesquisa
repro-
duzível

Versionando
projetos

ggplot

Encorajar o observador a comparar diferentes partes dos dados

Revelar diferentes níveis de detalhamento dos dados

Servir a um propósito claro e razoável: descrição, exploração, tabulação ou decoração

Estar integrada com as descrições estatísticas e verbais do conjunto de dados

4 princípios (Edward Tufte)

Pesquisa
repro-
duzível

Frederico
Bertholini

Nunca
esqueça

Recapitulando
manipu-
lação

Pesquisa
repro-
duzível

Versionando
projetos

ggplot

Miniaturas Múltiplas

Menor diferença efetiva

Causalidade (Respondendo a pergunta: “Comparado com o quê?”)

Contexto

Pesquisa
repro-
duzível

Frederico
Bertholini

Nunca
esqueça

Recapitulando
manipu-
lação

Pesquisa
repro-
duzível

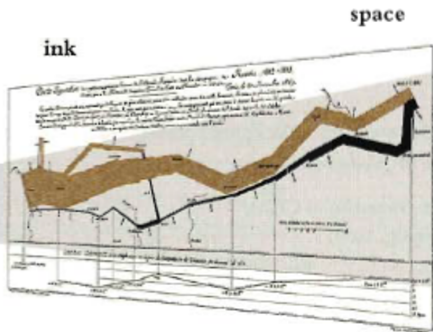
Versionando
projetos

ggplot



ideas

time



O que você quer mostrar?

Pesquisa
repro-
duzível

Frederico
Bertholini

Nunca
esqueça

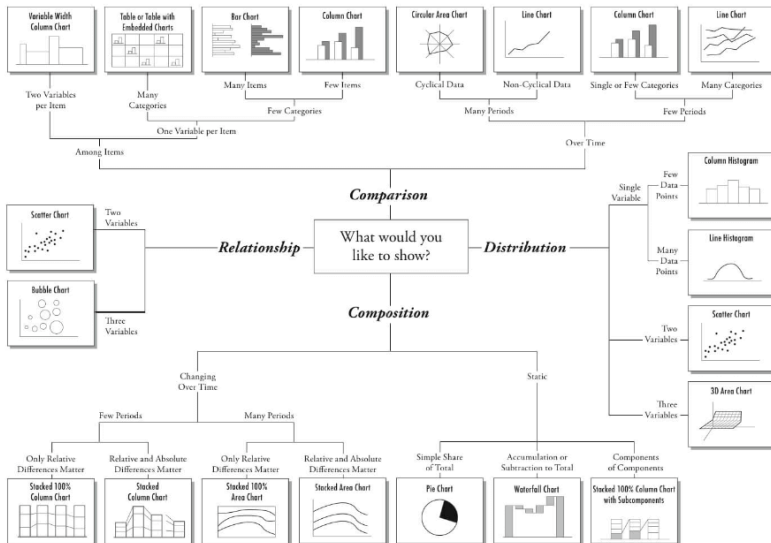
Recapitulando
manipu-
lação

Pesquisa
repro-
duzível

Versionando
projetos

ggplot

Chart Suggestions—A Thought-Starter



Pesquisa
repro-
duzível

Frederico
Bertholini

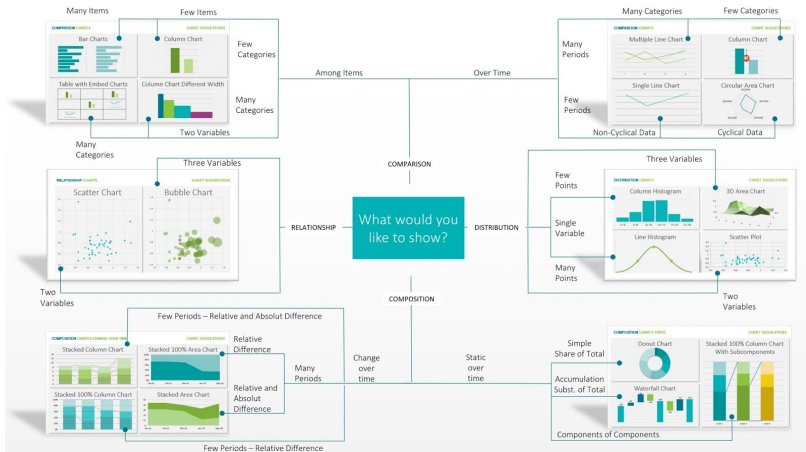
Nunca
esqueça

Recapitulando
manipulação

Pesquisa
repro-
duzível

Versionando
projetos

ggplot



Princípios

Pesquisa
repro-
duzível

Frederico
Bertholini

Nunca
esqueça

Recapitulando
manipu-
lação

Pesquisa
repro-
duzível

Versionando
projetos

ggplot

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

Recursos

Pesquisa
repro-
duzível

Frederico
Bertholini

Nunca
esqueça

Recapitulando
manipu-
lação

Pesquisa
repro-
duzível

Versionando
projetos

ggplot

- R Cookbook
- STHDA
- R Graph Gallery
- Extensões

Elementos do ggplot

Pesquisa
repro-
duzível

Frederico
Bertholini

Nunca
esqueça

Recapitulando
manipu-
lação

Pesquisa
repro-
duzível

Versionando
projetos

ggplot

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

Dados data =

Pesquisa
repro-
duzível

Frederico
Bertholini

Nunca
esqueça

Recapitulando
manipu-
lação

Pesquisa
repro-
duzível

Versionando
projetos

ggplot

- Dado empilhado?
- Cada coluna será uma entrada!

Geometrias geom_

Pesquisa
repro-
duzível

Frederico
Bertholini

Nunca
esqueça

Recapitulando
manipu-
lação

Pesquisa
repro-
duzível

Versionando
projetos

ggplot

- geom_tipo_de_geometria
- Recursos +
- cheat sheet
- manual ggplot

Estéticas aes()

Pesquisa
repro-
duzível

Frederico
Bertholini

Nunca
esqueça

Recapitulando
manipu-
lação

Pesquisa
repro-
duzível

Versionando
projetos

ggplot

■ x (xmax e xmin)

■ y

■ color

■ fill

■ shape

■ group

■ size

Escalas (estética) scale_

Pesquisa
repro-
duzível

Frederico
Bertholini

Nunca
esqueça

Recapitulando
manipu-
lação

Pesquisa
repro-
duzível

Versionando
projetos

ggplot

■ `scale_color_xx`

■ `scale_fill_xx`

■ `scale_shape_xx`

Escalas (eixos) `scale_x`

Pesquisa
repro-
duzível

Frederico
Bertholini

Nunca
esqueça

Recapitulando
manipu-
lação

Pesquisa
repro-
duzível

Versionando
projetos

ggplot

- Contínua `scale_x_continuous`
- Discreta `scale_x_discrete`
- Série de tempo `zoo` e `lubridate` -> `scale_yearmon`

Tema

Pesquisa
repro-
duzível

Frederico
Bertholini

Nunca
esqueça

Recapitulando
manipu-
lação

Pesquisa
repro-
duzível

Versionando
projetos

ggplot

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

Facet

Pesquisa
repro-
duzível

Frederico
Bertholini

Nunca
esqueça

Recapitulando
manipu-
lação

Pesquisa
repro-
duzível

Versionando
projetos

ggplot

■ `facet_grid`

■ `facet_wrap`

Adicionais

Pesquisa
repro-
duzível

Frederico
Bertholini

Nunca
esqueça

Recapitulando
manipu-
lação

Pesquisa
repro-
duzível

Versionando
projetos

ggplot

Gráficos com interatividade:

- ggiraph
- plotly (ggplotly)

Combinação de gráficos

- patchwork
- cowplot

Exercício

Pesquisa
repro-
duzível

Frederico
Bertholini

Nunca
esqueça

Recapitulando
manipu-
lação

Pesquisa
repro-
duzível

Versionando
projetos

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 `mutatee 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
manipu-
lação

Pesquisa
repro-
duzível

Versionando
projetos

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
manipu-
lação

Pesquisa
repro-
duzível

Versionando
projetos

ggplot

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

Pesquisa
repro-
duzível

Frederico
Bertholini

Nunca
esqueça

Resumindo
manipu-
lação

Pesquisa
repro-
duzível

Versionando
projetos

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
manipu-
lação

Pesquisa
repro-
duzível

Versionando
projetos

ggplot

grafico

