

# Oficina de Introdução ao Software R

R e Desafios no Projeto Temático

Beatriz Milz

17/10/2018

# Sobre mim



- Beatriz Milz, 25 anos, Diadema.
- Bacharel em Gestão Ambiental (EACH/USP);
- Mestre em Ciências no PPG-Análise Ambiental Integrada (UNIFESP/Diadema)
- Atualmente candidata ao Doutorado em Ciência Ambiental no PROCAM/IEE/USP.
- Aprendendo R desde agosto de 2018.
- Participa da comunidade R-Ladies São Paulo desde agosto de 2018.
- Email: [beatriz.milz@hotmail.com](mailto:beatriz.milz@hotmail.com)
- Github: [@beatrizmilz](https://github.com/beatrizmilz) (repositório de código)

# Macrometrópole Paulista

Projeto Temático Fapesp: Governança ambiental da macrometrópole paulista face à variabilidade climática

# Delimitação territorial da Macrometrópole Paulista (MMP)

- Diferentes delimitações:
- Empresa Paulista de Planejamento Metropolitano (EMPLASA): 174 municípios
- Departamento de Águas e Energia Elétrica (DAEE): 180 municípios



# Como o R pode ajudar?

R is a free software environment for statistical computing and graphics.  
(<https://www.r-project.org/>)

R é um ambiente de software livre para computação estatística e gráficos.

- O R é open-source



# O R é uma linguagem de programação.

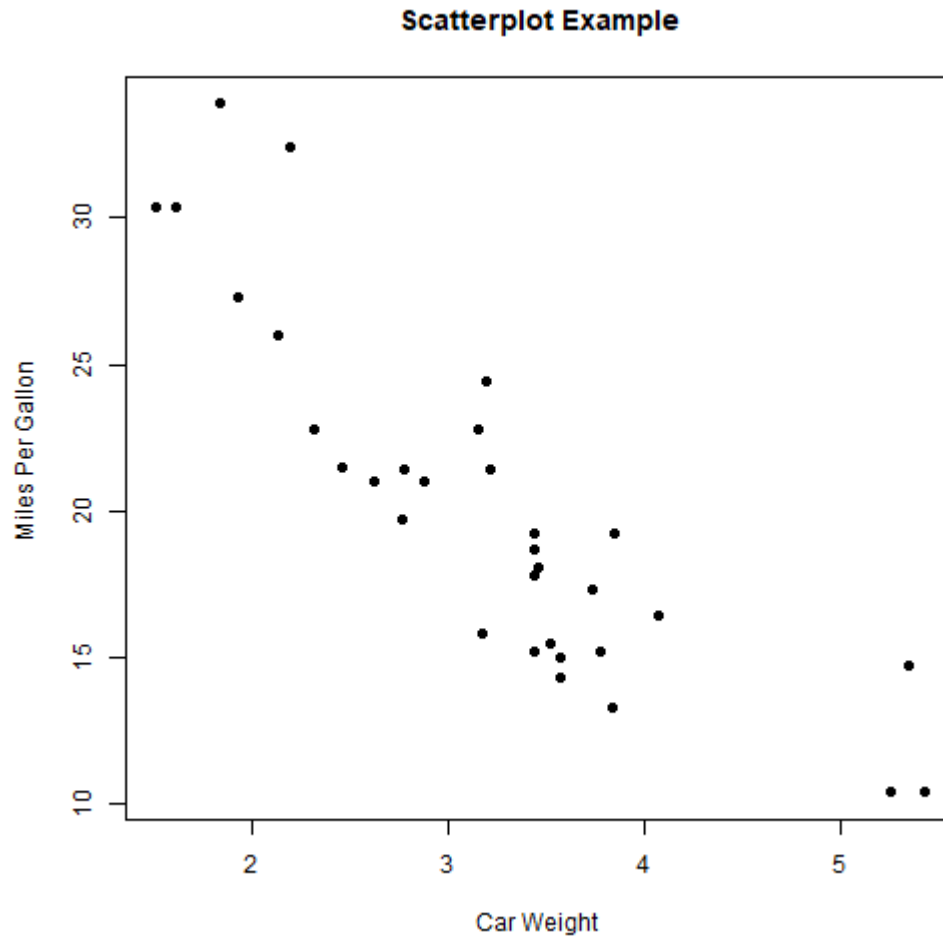
Qual é a vantagem?

É um texto

É reprodutível

Dá para compartilhar!

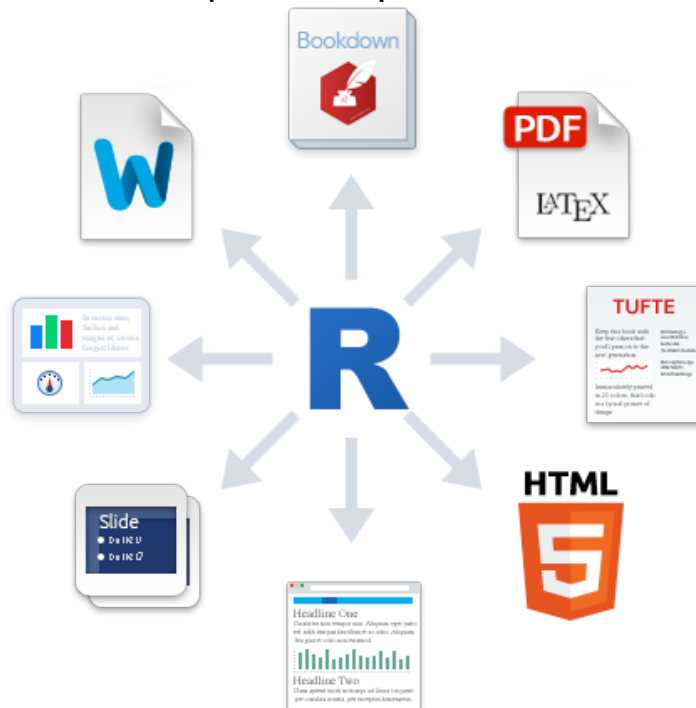
```
attach(mtcars)  
plot(wt, mpg, main="Scatterplot Example",  
xlab="Car Weight ", ylab="Miles Per Gallon ", pch=19)
```





# Rmarkdown

- É um tipo de arquivo que suporta códigos em R, texto, markdown e outros formatos.
- O markdown é uma linguagem de marcação simples.
- [Rmarkdown Cheatsheet](#)
- Possibilita exportar diferentes tipos de arquivos.



# Rmarkdown

É possível elaborar relatórios no Rmarkdown. Porque é interessante?

## Exemplo:

Pergunta: Qual é a população da MMP?

- Depende da delimitação territorial!
- Neste exemplo usaremos a delimitação da EMPLASA.

# Tidyverse

Primeiro precisamos carregar o pacote tidyverse.

```
#install.packages("tidyverse") # Instalar o pacote Tidyverse  
library(tidyverse) #Carregar o pacote tidyverse
```

Fonte: <http://r4ds.had.co.nz/>

# Quais são as divisões administrativas?

## Importar a tabela para o R:

obs: Agradeço à Nana Laura e Ana Lia pela elaboração e disponibilização da tabela que contém a divisão administrativa, e os municípios pertencentes à MMP.

```
#Carregar a tabela de divisão administrativa
divisao_adm <- readxl::read_xlsx("dados/divisao_adm.xlsx")
#Renomear as colunas
names(divisao_adm) <- c("codigo_ibge", "municipio", "regiao_adm")
#Transformar a coluna do código IBGE em número
divisao_adm$codigo_ibge <- as.integer(divisao_adm$codigo_ibge)
```

# Delimitação Emplasa

- Importar as tabelas excel contendo os municípios que fazem parte da MMP:

```
# Importar a tabela para o R
mmp_emplasa <- readxl::read_xlsx("dados/delimitacaoterritorial")

# Renomear as colunas para padronizar os nomes
names(mmp_emplasa) <- c("municipio", "codigo_ibge")

# Padronizar e deixar o código do IBGE configurado como número
mmp_emplasa$codigo_ibge <- as.integer(mmp_emplasa$codigo_ibge)

# Juntar os dados de divisão administrativa baseados nos municípios
mmp_emplasa <- left_join(mmp_emplasa, divisao_adm, by="codigo_ibge")

# Deletar a coluna que fica repetida
mmp_emplasa$municipio.y <- NULL
```

**E agora?**

**Podemos cruzar com outros bancos de dados!**

---

# Cruzar com outras informações

Para esse exemplo, vamos cruzar com dados de população por município:

- Informações dos Municípios Paulistas - <http://www.imp.seade.gov.br/frontend/>

```
# Importar a tabela para o R  
imp_2018 <- read_csv2("dados/imp_2018-10-02_01-05.csv")  
#Renomear as colunas  
names(imp_2018) <- c("municipio", "ano", "populacao", "popula
```

# Left Join - Emplasa

```
# Cria a tabela e faz o left join baseado no código IBGE
imp_2018_mmp_emplasa <- left_join(mmp_emplasa, imp_2018, by="municipio")
# A coluna de município está repetida, então apago uma delas.
imp_2018_mmp_emplasa$municipio.y <- NULL

head(imp_2018_mmp_emplasa)
```

```
## # A tibble: 6 x 16
##   municipio.x codigo_ibge regioao_administ~ regioao_de_gover~
##   <chr>          <int> <chr>          <chr>
## 1 ÁGUAS DE S~      3500600 Região Administ~ Região de Gover~
## 2 ÁGUAS DE S~      3500600 Região Administ~ Região de Gover~
## 3 ÁGUAS DE S~      3500600 Região Administ~ Região de Gover~
## 4 ÁGUAS DE S~      3500600 Região Administ~ Região de Gover~
## 5 ÁGUAS DE S~      3500600 Região Administ~ Região de Gover~
## 6 ALAMBARI        3500758 Região Administ~ Região de Gover~
## # ... with 12 more variables: regioao_metropolitana <chr>,
## #   aglomeracoes_urbanas <chr>, rm_au <chr>, municipio <chr>, ano
## #   populacao <int>, populacao_masculina <int>, populacao_feminina
## #   razao_de_sexos <dbl>, populacao_urbana <int>, populacao_rural
## #   grau_urbanizacao_porcentagem <dbl>
```



Vamos aos dados!

# População na MMP entre 2014 e 2018 - Delimitação EMPLASA

Fonte: - Informações dos Municípios Paulistas -

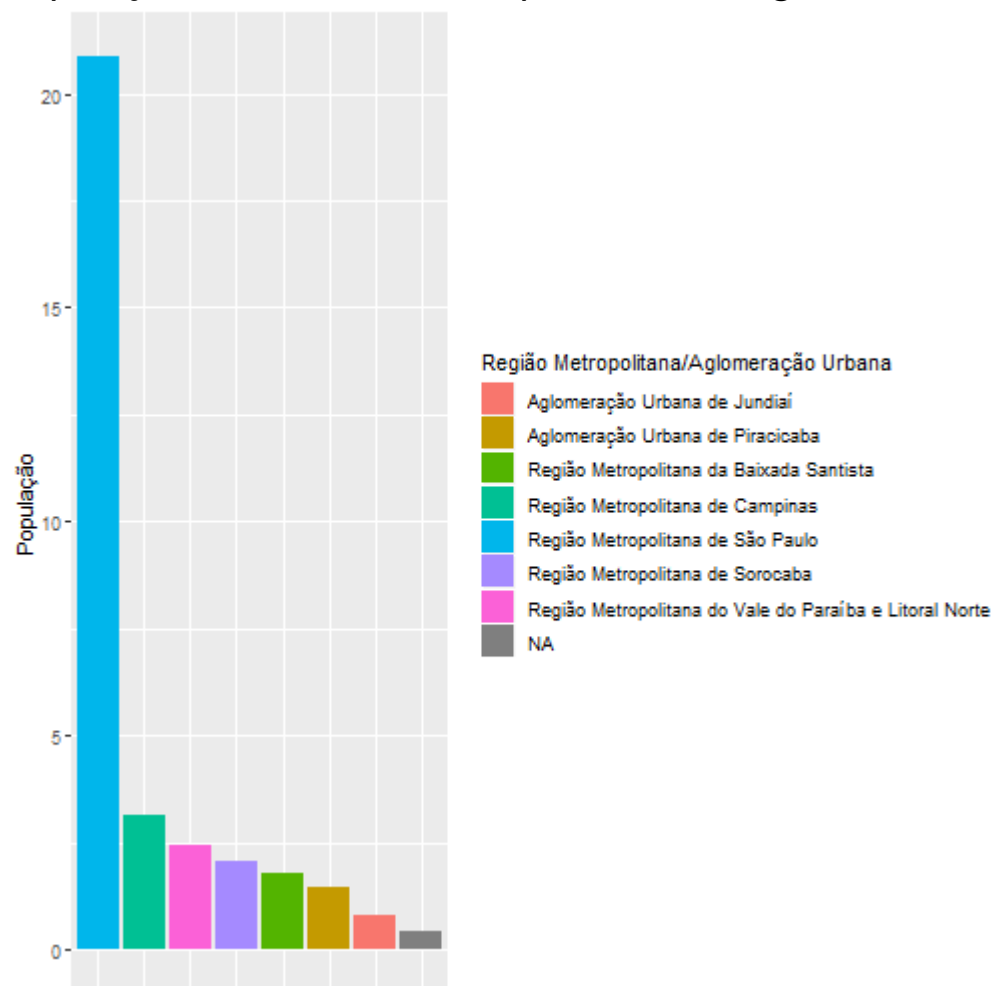
<http://www.imp.seade.gov.br/frontend/>

```
imp_2018_mmp_emplasa %>%  
group_by(ano) %>%  
  summarise(PopulacaoTotal = sum(populacao))
```

```
## # A tibble: 5 x 2  
##   ano PopulacaoTotal  
##   <int>         <int>  
## 1  2014         31828473  
## 2  2015         32126172  
## 3  2016         32376430  
## 4  2017         32629239  
## 5  2018         32884608
```

Segundo os dados da SEADE, em 2018 a MMP tem 32,88 milhões de habitantes.

População na MMP em 2018, por RM/AU, segundo os dados da SEADE:



# Vantagens de trabalhar com o R e o Rmarkdown

- É possível escrever os códigos que geram as tabelas e gráficos.
- Quando o banco de dados for atualizado (por exemplo, houver alguma modificação na lista de municípios pertencentes à MMP) não será necessário refazer os gráficos. Apenas necessário compilar novamente.
- Facilita o compartilhamento dos dados e análises com outros pesquisadores.
- É possível exportar em diversos formatos, inclusive transformar o relatório em apresentações.

# Materiais de estudo

- [R for Data Science](#) - Livro em inglês, disponibilizado gratuitamente online. Conteúdo muito bom sobre R.
- [Curso-R](#) - Cursos presenciais sobre R.
- [Material Curso-R](#) - Material disponibilizado pela Curso-R para estudo.
- [Blog Curso-R](#) - Blog do Curso-R com diversos conteúdos interessantes.

# Links interessantes

- [Dados de execução orçamentária da PMSP](#)
- [Rpollution](#) - Blog onde é publicado análise de dados sobre poluição do ar;

# Obrigada!

Apresentação elaborada com **Xaringan**