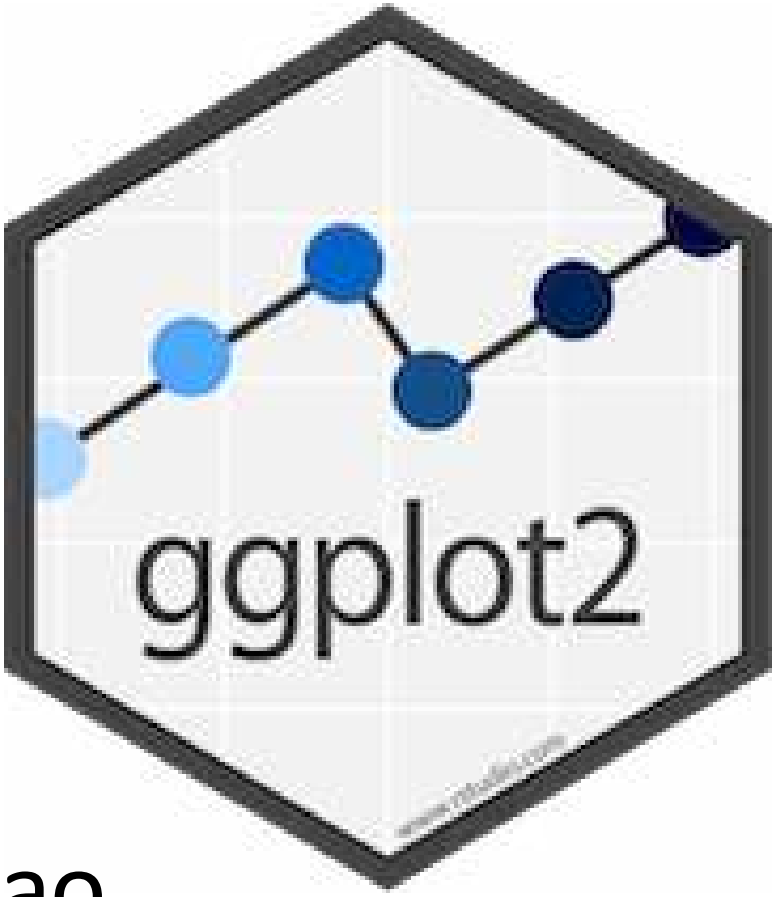
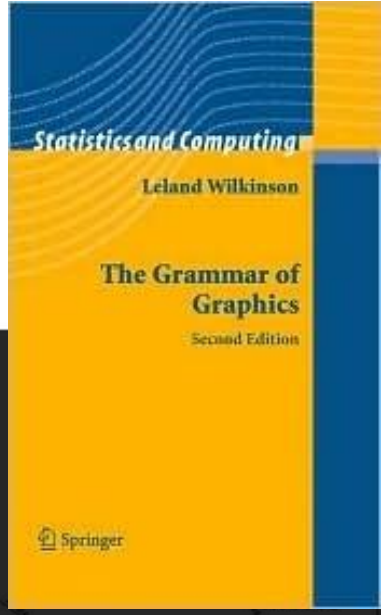


Introdução ao
** Tidyverse **
gráficos com ggplo2

Grammar of Graphics by L. Wilkinson



Describes non-data ink. Design elements!

The plotting space you are using

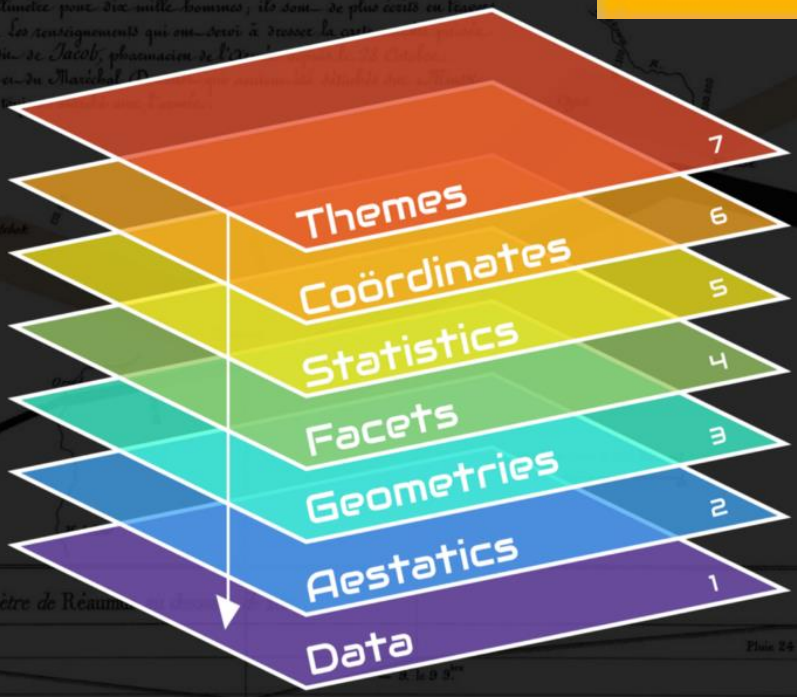
Statistical models & summaries

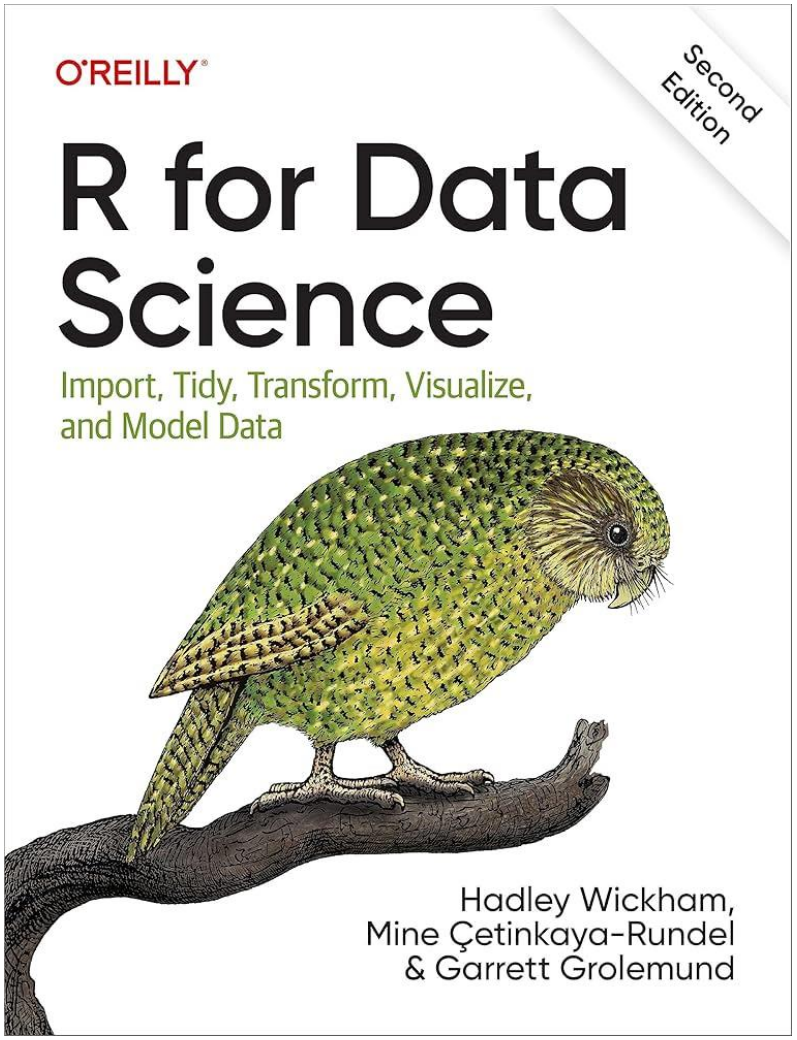
Rows and columns of sub-plots

Shapes used to represent your data

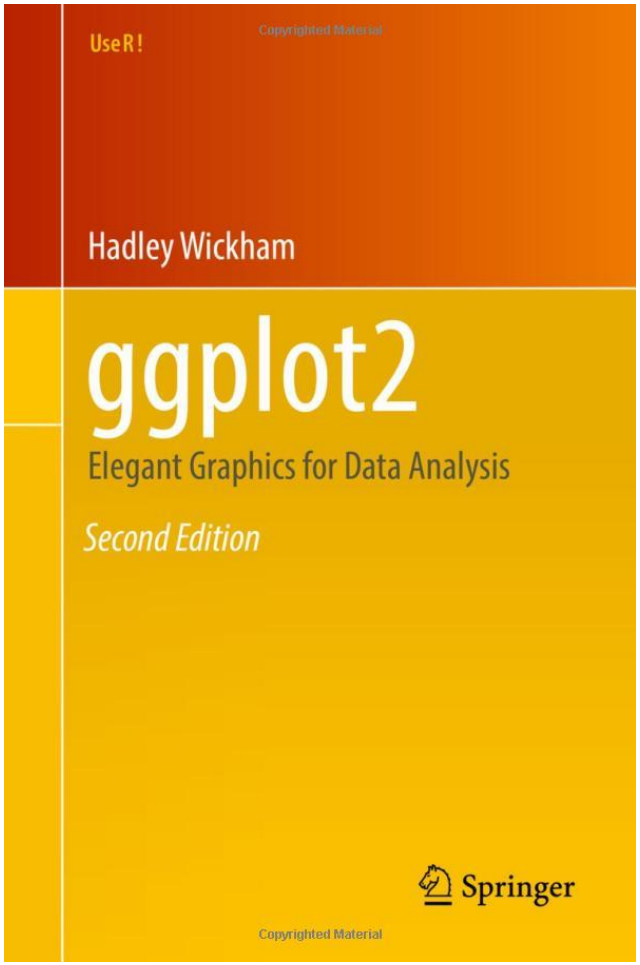
The scales on which the data is mapped

The actual variables to be plotted



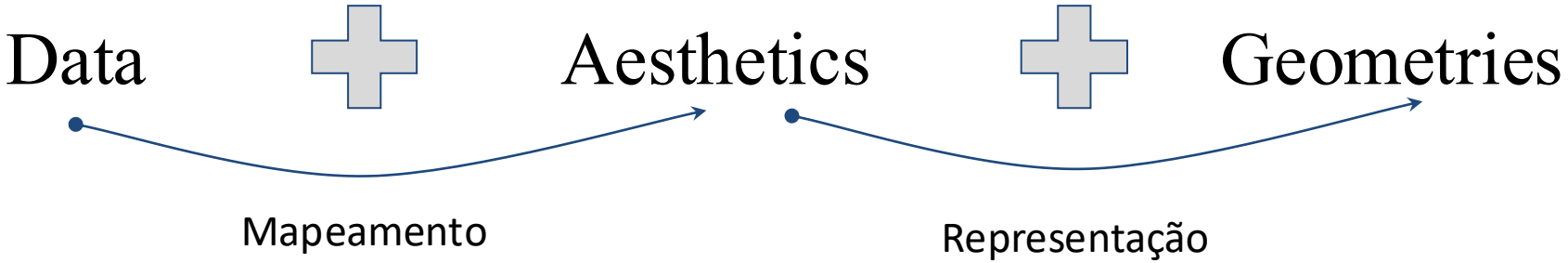


https://r4ds.hadley.nz



https://ggplot2-book.org

O que é um gráfico



https://paternogbc.github.io/disciplina-graficos-r/

A lógica do ggplot2

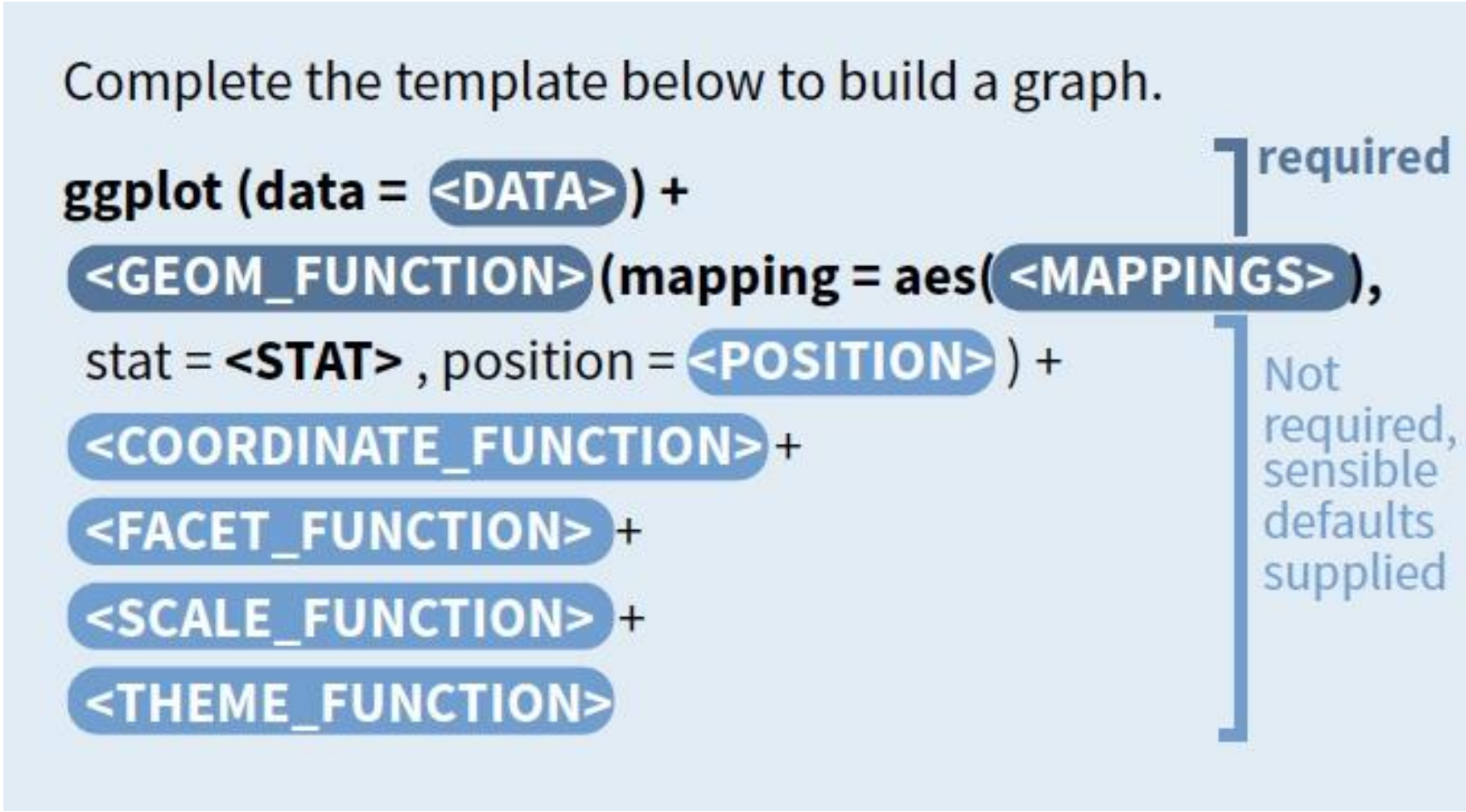
```
ggplot(data = <DATASET>) +

    <GEOM_FUNCTION>(mapping = aes(<MAPPING>), stat = <STAT>,
    position = <POSITION>) +

        <COORDINATE_FUNCTION> +

            <FACET_FUNCTION>
```

A lógica do ggplot2



esquisse

> The purpose of this add-in is to let you explore your data quickly to extract the information they hold. You can create visualization with {ggplot2}, filter data with {dplyr} and retrieve generated code.

This addin allows you to interactively explore your data by visualizing it with the ggplot2 package. It allows you to draw bar plots, curves, scatter plots, histograms, boxplot and sf objects, then export the graph or retrieve the code to reproduce the graph.

See online documentation : https://dreamrs.github.io/esquisse/index.html

If you find bugs, please open an issue

Installation

Install from CRAN with :

```
install.packages("esquisse")
```

Or install development version from GitHub :

```
remotes::install_github("dreamRs/esquisse")
```

Then launch the addin via the RStudio menu or with `esquisse::esquisser()`.

Links

View on CRAN

Browse source code

Report a bug

License

GPL-3 | file LICENSE

Citation

Citing esquisse

Developers

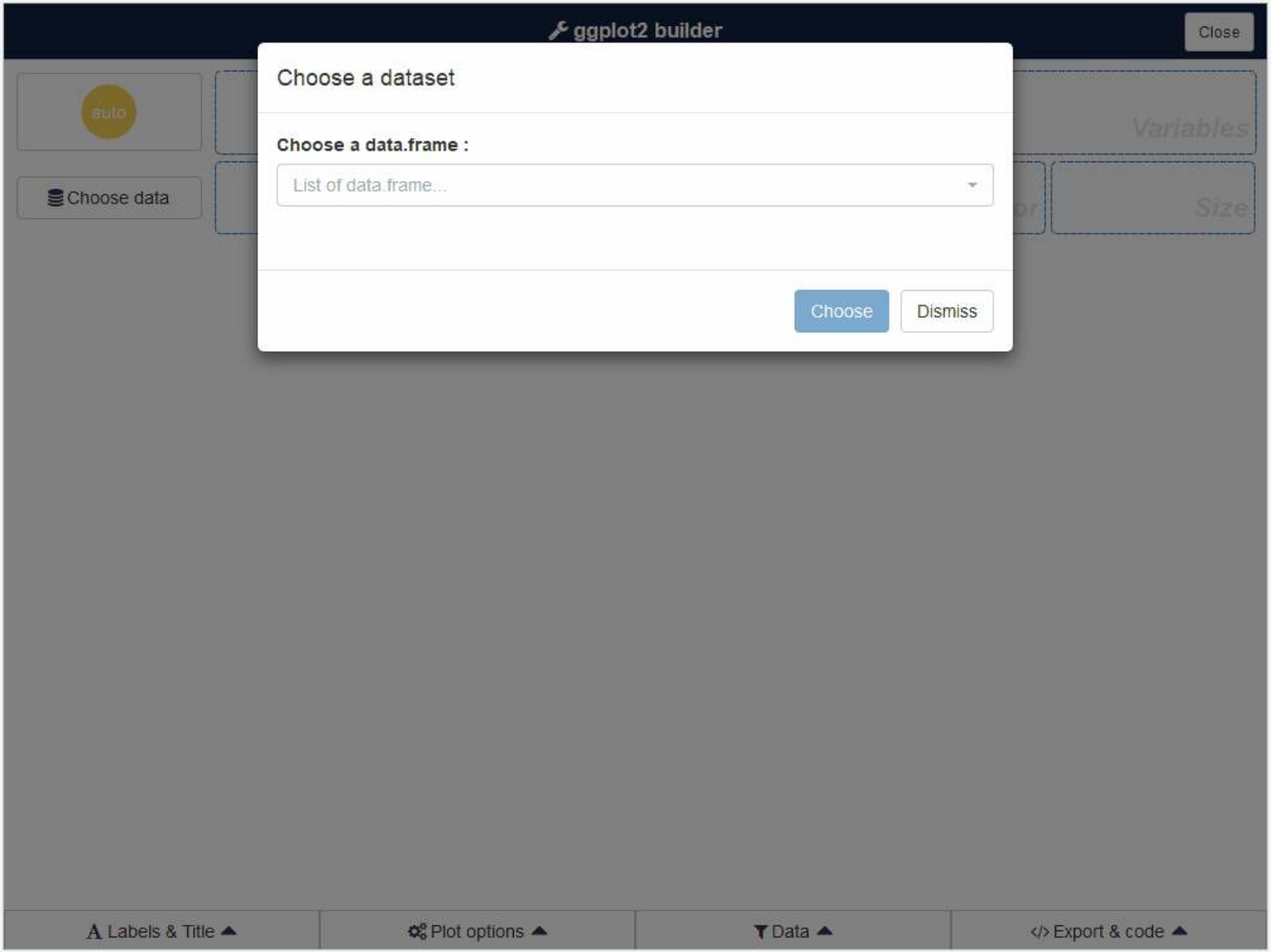

Fanny Meyer
Author


Victor Perrier
Author, maintainer

More about authors...

Dev status

Explore your data with ggplot2



Funcionamento do ggplot2 - Duas variáveis contínuas

Dados

TOTAL	HIGH	MEDIUM	LOW
81	2	71	8
44	3	30	11
61	1	30	30
34	1	30	3
34	0	34	0
34	0	34	0
28	5	5	18
32	0	32	0
28	0	28	0
26	1	20	5
22	0	22	0
18	1	14	3
18	0	18	0
9	3	0	6
3	3	0	0

https://paternogbc.github.io/disciplina-graficos-r/

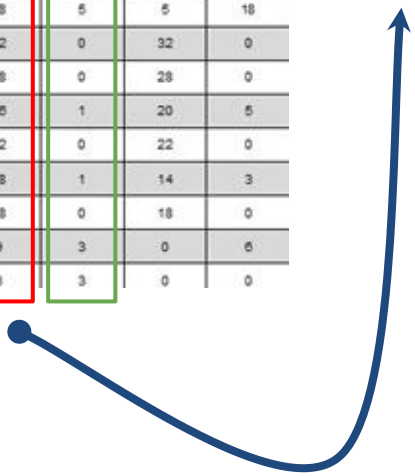
Funcionamento do ggplot2

Dados + Mapeamento

TOTAL	HIGH	MEDIUM	LOW
81	2	71	8
44	3	30	11
61	1	30	30
34	1	30	3
34	0	34	0
34	0	34	0
28	5	5	18
32	0	32	0
28	0	28	0
26	1	20	5
22	0	22	0
18	1	14	3
18	0	18	0
9	3	0	6
3	3	0	0

Y = TOTAL

X = HIGH



https://paternogbc.github.io/disciplina-graficos-r/

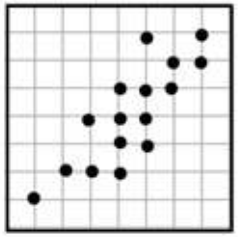
Funcionamento do ggplot2

Dados + Mapeamento + Geometria

TOTAL	HIGH	MEDIUM	LOW
81	2	71	8
44	3	30	11
61	1	30	30
34	1	30	3
34	0	34	0
34	0	34	0
28	5	5	18
32	0	32	0
28	0	28	0
26	1	20	5
22	0	22	0
18	1	14	3
18	0	18	0
9	3	0	6
3	3	0	0

Y = TOTAL

X = HIGH



point

https://paternogbc.github.io/disciplina-graficos-r/

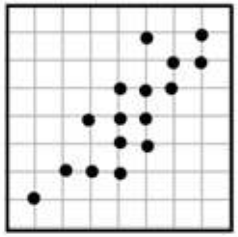
Funcionamento do ggplot2

Dados + Mapeamento + Geometria

TOTAL	HIGH	MEDIUM	LOW
81	2	71	8
44	3	30	11
61	1	30	30
34	1	30	3
34	0	34	0
34	0	34	0
28	5	5	18
32	0	32	0
28	0	28	0
26	1	20	5
22	0	22	0
18	1	14	3
18	0	18	0
9	3	0	6
3	3	0	0

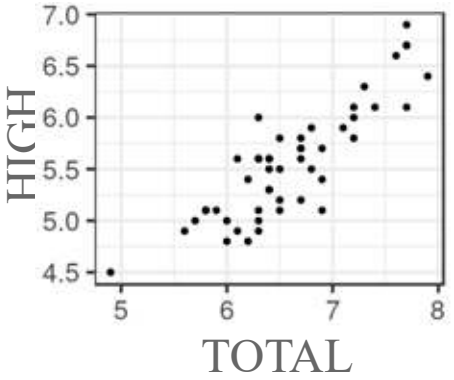
Y = TOTAL

X = HIGH



point

Gráfico



https://paternogbc.github.io/disciplina-graficos-r/

Funcionamento do ggplot2 - Uma variável contínua

Dados

TOTAL	HIGH	MEDIUM	LOW
81	2	71	8
44	3	30	11
61	1	30	30
34	1	30	3
34	0	34	0
34	0	34	0
28	5	5	18
32	0	32	0
28	0	28	0
26	1	20	5
22	0	22	0
18	1	14	3
18	0	18	0
9	3	0	6
3	3	0	0

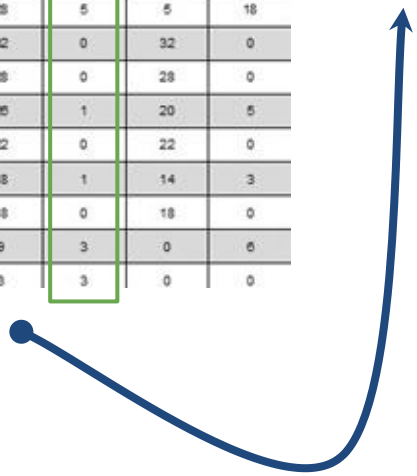
https://paternogbc.github.io/disciplina-graficos-r/

Funcionamento do ggplot2

Dados + Mapeamento

TOTAL	HIGH	MEDIUM	LOW
81	2	71	8
44	3	30	11
61	1	30	30
34	1	30	3
34	0	34	0
34	0	34	0
28	5	5	18
32	0	32	0
28	0	28	0
26	1	20	5
22	0	22	0
18	1	14	3
18	0	18	0
9	3	0	6
3	3	0	0

X = HIGH



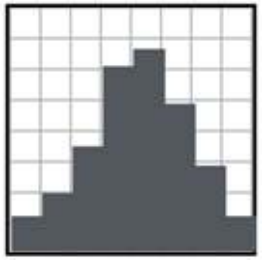
https://paternogbc.github.io/disciplina-graficos-r/

Funcionamento do ggplot2

Dados + Mapeamento + Geometria

TOTAL	HIGH	MEDIUM	LOW
81	2	71	8
44	3	30	11
61	1	30	30
34	1	30	3
34	0	34	0
34	0	34	0
28	5	5	18
32	0	32	0
28	0	28	0
26	1	20	5
22	0	22	0
18	1	14	3
18	0	18	0
9	3	0	6
3	3	0	0

X = HIGH



histogram

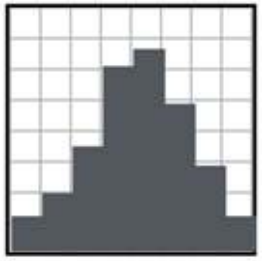
https://paternogbc.github.io/disciplina-graficos-r/

Funcionamento do ggplot2

Dados + **Mapeamento** + **Geometria**

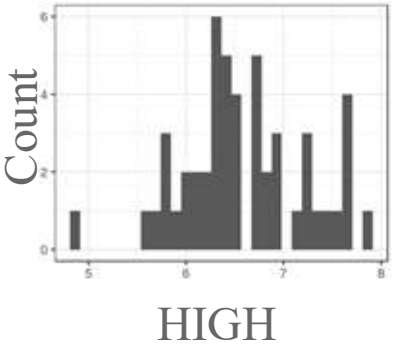
TOTAL	HIGH	MEDIUM	LOW
81	2	71	8
44	3	30	11
61	1	30	30
34	1	30	3
34	0	34	0
34	0	34	0
28	5	5	18
32	0	32	0
28	0	28	0
26	1	20	5
22	0	22	0
18	1	14	3
18	0	18	0
9	3	0	6
3	3	0	0

X = HIGH



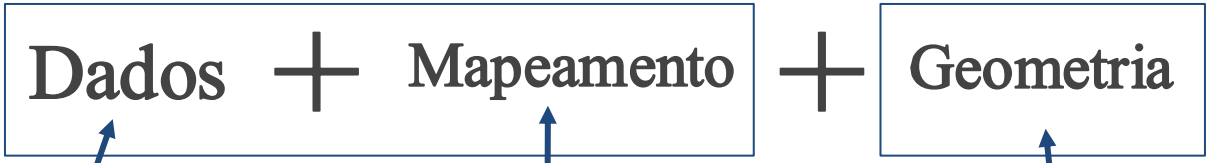
histogram

Gráfico



https://paternogbc.github.io/disciplina-graficos-r/

ggplot2 no R



Objeto
Classe
“data.frame”

colunas
do data.frame Y e X
Grupos, Fill e Cor

função
família geom_

https://paternogbc.github.io/disciplina-graficos-r/

ggplot2 no R


```
ggplot(data, aes(x, y)) + geom_point()
```

`data` = Define o objeto com a tabela de dados

`aes` = Define qual coluna é X e qual é Y

`geom_` = Define qual geometria plotar

https://paternogbc.github.io/disciplina-graficos-r/

```
ggplot(data, aes(x, y)) +
 geom_point() +
 theme() +
 etc
```


O sinal de mais é utilizado para adicionar mais camadas. O recomendado é colocar apenas uma camada por linha para facilitar a leitura e o pensamento.

Mapeamento das variáveis no ggplot

Tabela de dados

Petal.Length <dbl>	Petal.Width <dbl>	Species <fctr>
1.4	0.2	setosa
1.4	0.2	setosa
1.3	0.2	setosa
1.5	0.2	setosa
1.4	0.2	setosa
1.7	0.4	setosa

https://paternogbc.github.io/disciplina-graficos-r/

Mapeamento das variáveis no ggplot

Tabela de dados

https://paternogbc.github.io/disciplina-graficos-r/

Mapeamento das variáveis no ggplot

Tabela de dados

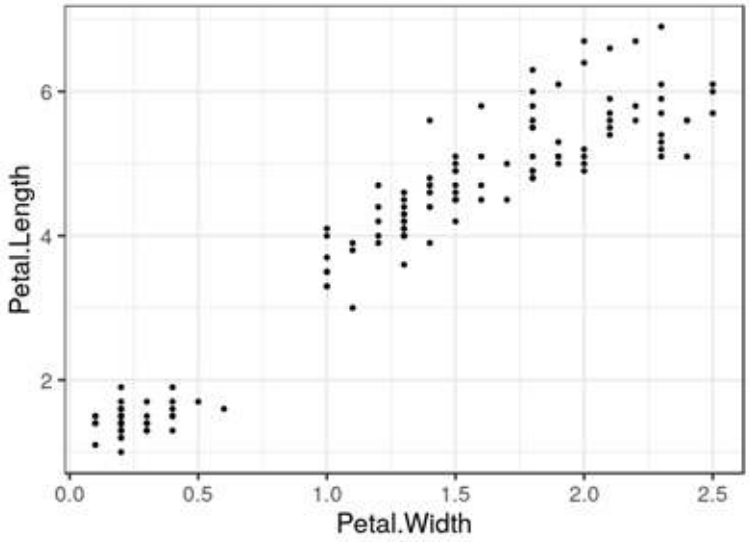
Petal.Length <dbl>	Petal.Width <dbl>	Species <fctr>
1.4	0.2	setosa
1.4	0.2	setosa
1.3	0.2	setosa
1.5	0.2	setosa
1.4	0.2	setosa
1.7	0.4	setosa

Mapeamento

Y

X

Gráfico



Geometria = ●

https://paternogbc.github.io/disciplina-graficos-r/

Mapeamento das variáveis no ggplot

Tabela de dados

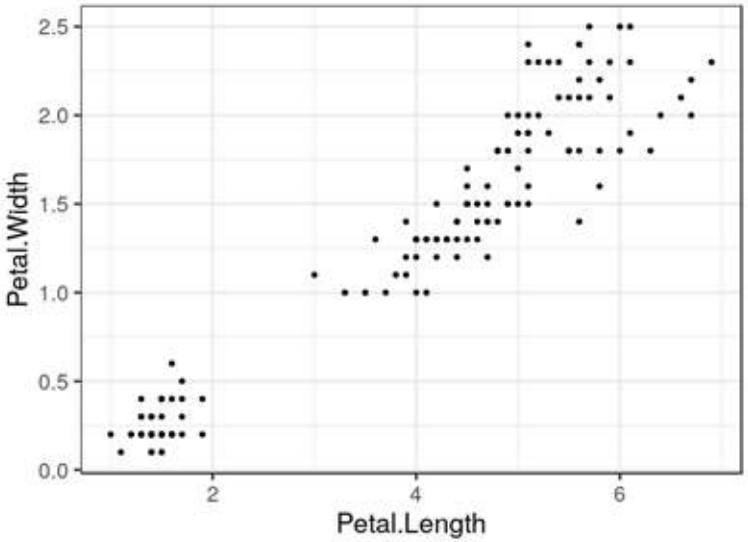
Petal.Length <dbl>	Petal.Width <dbl>	Species <fctr>
1.4	0.2	setosa
1.4	0.2	setosa
1.3	0.2	setosa
1.5	0.2	setosa
1.4	0.2	setosa
1.7	0.4	setosa

Mapeamento

X

Y

Gráfico



Geometria = ●

https://paternogbc.github.io/disciplina-graficos-r/

Mapeamento das variáveis no ggplot

Tabela de dados

Petal.Length <dbl>	Petal.Width <dbl>	Species <fctr>
1.4	0.2	setosa
1.4	0.2	setosa
1.3	0.2	setosa
1.5	0.2	setosa
1.4	0.2	setosa
1.7	0.4	setosa

Mapeamento

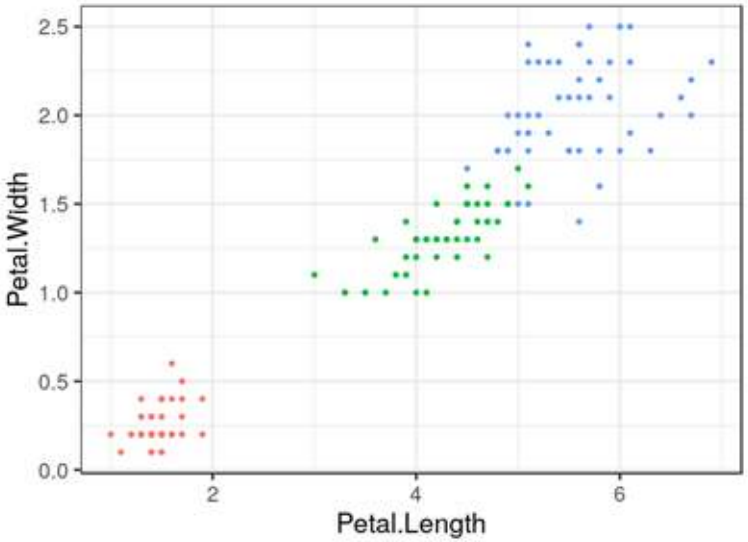
X

Y

color

Geometria = ●

Gráfico



https://paternogbc.github.io/disciplina-graficos-r/

Mapeamento das variáveis no ggplot

Tabela de dados

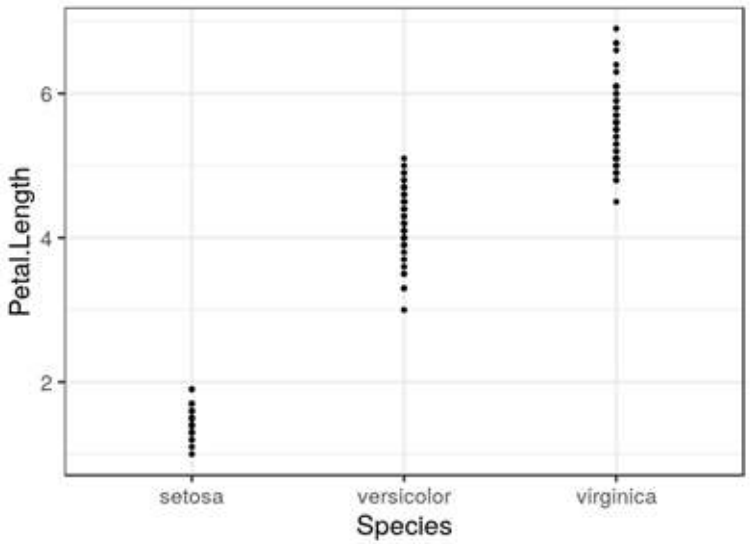
Petal.Length <dbl>	Petal.Width <dbl>	Species <fctr>
1.4	0.2	setosa
1.4	0.2	setosa
1.3	0.2	setosa
1.5	0.2	setosa
1.4	0.2	setosa
1.7	0.4	setosa

Mapeamento

Y

X

Gráfico



Geometria = ●

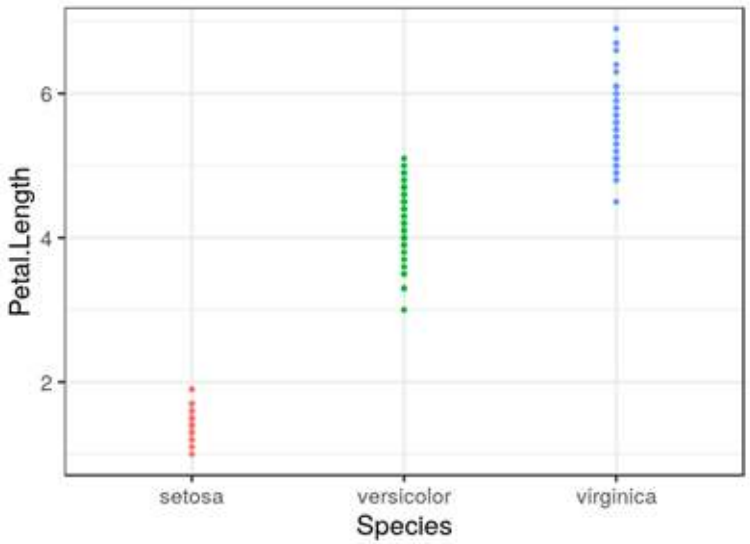
https://paternogbc.github.io/disciplina-graficos-r/

Mapeamento das variáveis no ggplot

Tabela de dados

Petal.Length <dbl>	Petal.Width <dbl>	Species <fctr>
1.4	0.2	setosa
1.4	0.2	setosa
1.3	0.2	setosa
1.5	0.2	setosa
1.4	0.2	setosa
1.7	0.4	setosa

Gráfico



Mapeamento

Y

X

color

Geometria = ●

https://paternogbc.github.io/disciplina-graficos-r/

Variáveis Discretas vs Variáveis Contínuas

O ggplot interpreta qualquer variável **categórica como discreta**

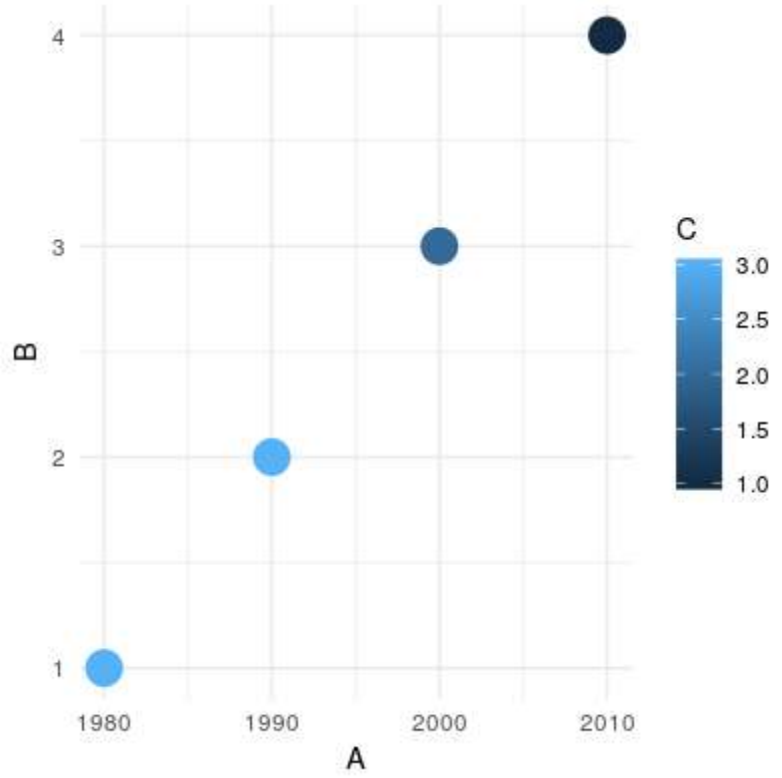
variáveis de texto geralmente são convertidas em fatores pelo R

Fatores são considerados variáveis discretas pelo ggplot

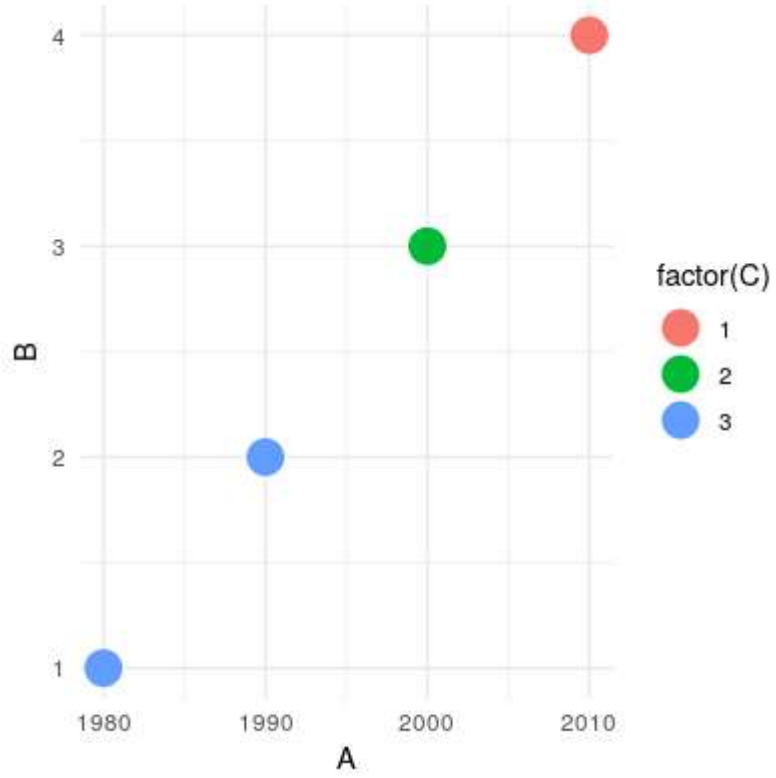
Variáveis discretas são mapeadas por níveis e não de forma contínua

Variáveis Discretas vs Variáveis Contínuas

C é uma variável contínua
color = C



C é uma variável discreta
color = factor(C)



https://paternogbc.github.io/disciplina-graficos-r/

Os cinco Gráficos Básicos

https://paternogbc.github.io/disciplina-graficos-r/

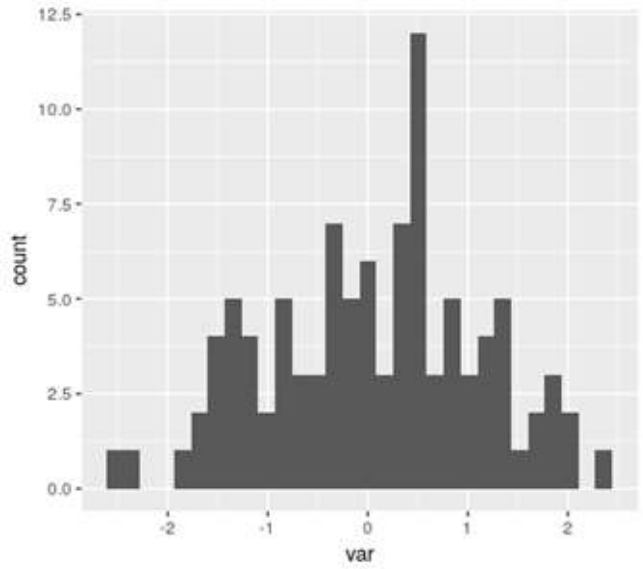
Uma Variável - Numérica

Histograma é uma representação da distribuição de uma variável numérica

Histograma → **geom_histogram()**

```
ggplot(dados, aes(x = var)) +
  geom_histogram()
```

Variável numérica



https://paternogbc.github.io/disciplina-graficos-r/

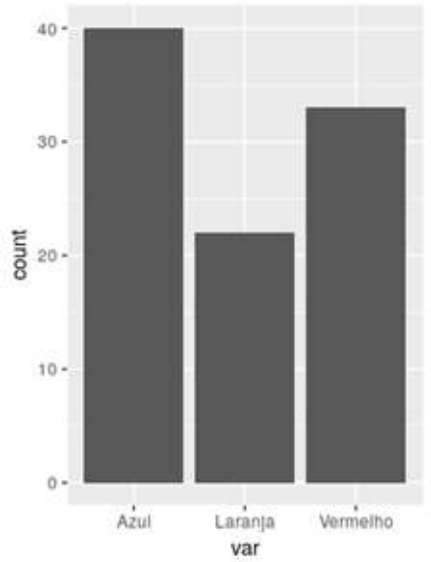
Uma Variável - Categórica

Gráfico de Barras é um gráfico que representa variáveis categóricas com barras retangulares e comprimento proporcional aos valores que ele representa

Barras → **geom_bar()**

```
ggplot(dados, aes(x = var)) +
  geom_bar()
```

Variável categórica



https://paternogbc.github.io/disciplina-graficos-r/

Duas variáveis - [Y = numérica | X = categórica]

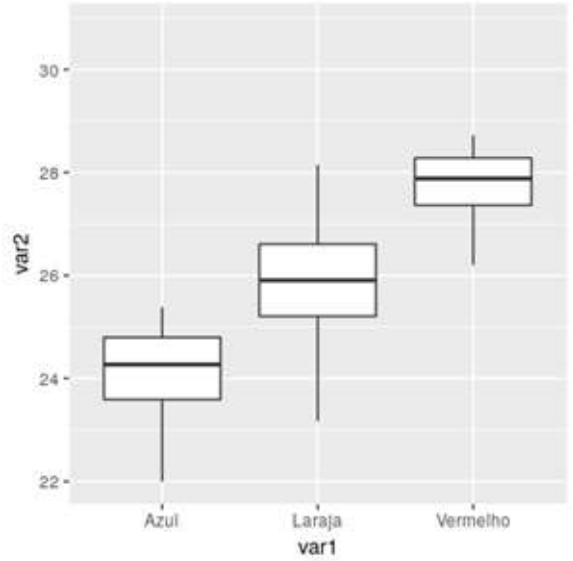
Boxplot é um gráfico que representa quartis de uma variável numérica em função de uma variável categórica (grupos)

Boxplot → **geom_boxplot()**

```
ggplot(dados, aes(x = var1, y = var2)) +
  geom_boxplot()
```

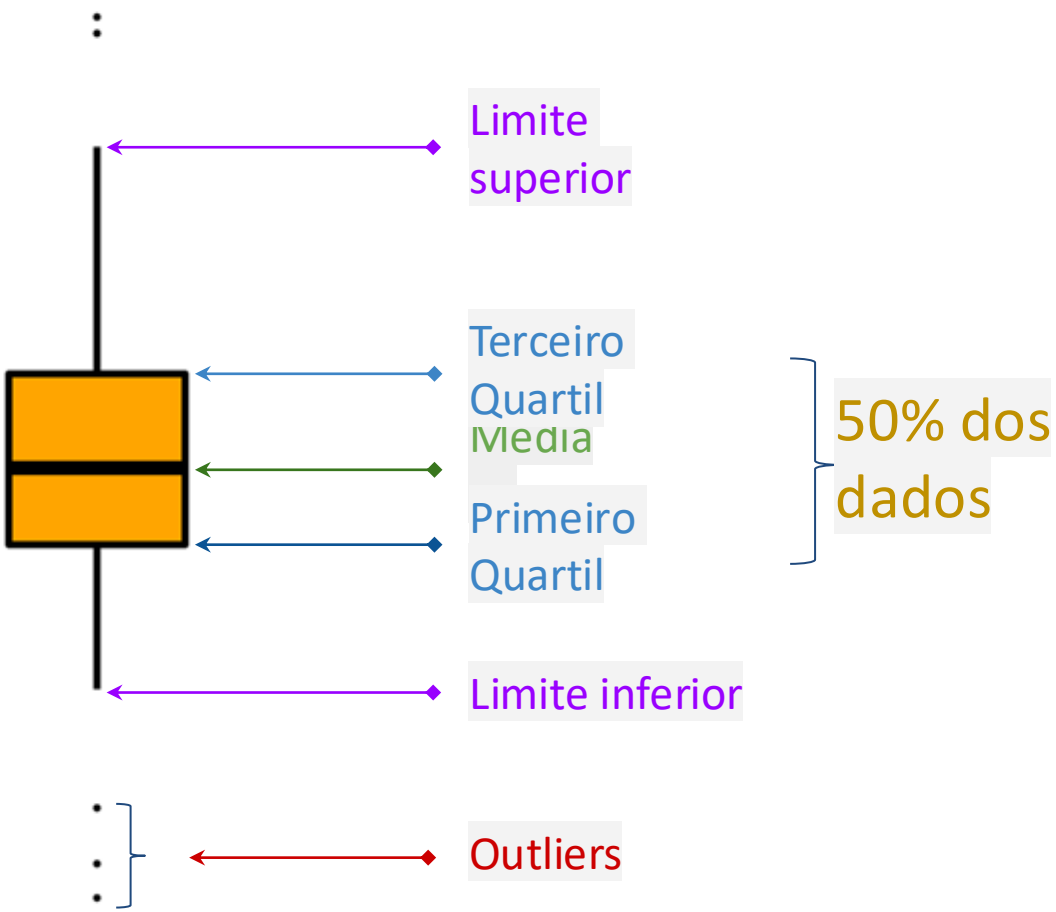
Variável categórica

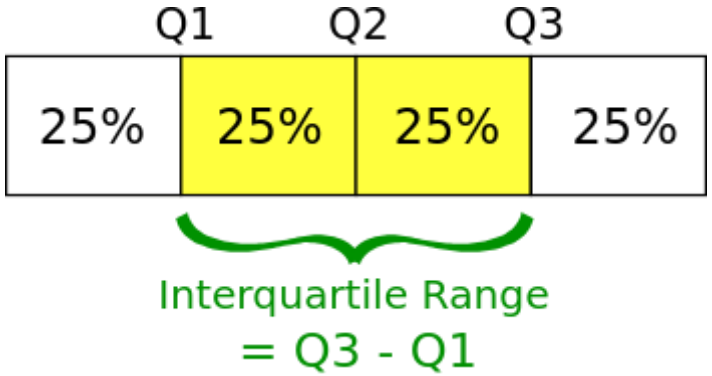
Variável numérica



https://paternogbc.github.io/disciplina-graficos-r/

Anatomia do boxplot





https://paternogbc.github.io/disciplina-graficos-r/

Duas variáveis - [Y = numérica | X = numérica]

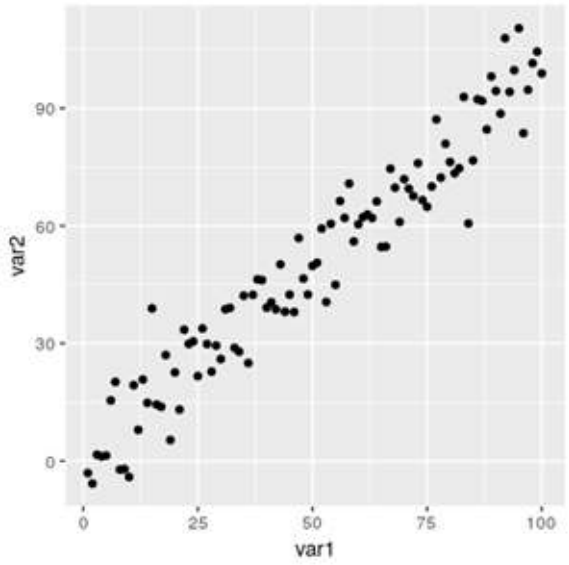
Gráfico de dispersão é um gráfico que representa a relação entre duas variáveis numéricas

Dispersão → **geom_point()**

```
gplot(dados, aes(x = var1, y = var2)) +
  geom_point()
```

Variável numérica

Variável numérica



https://paternogbc.github.io/disciplina-graficos-r/

Duas variáveis - [Y = numérica | X = numérica] | Um Y por X

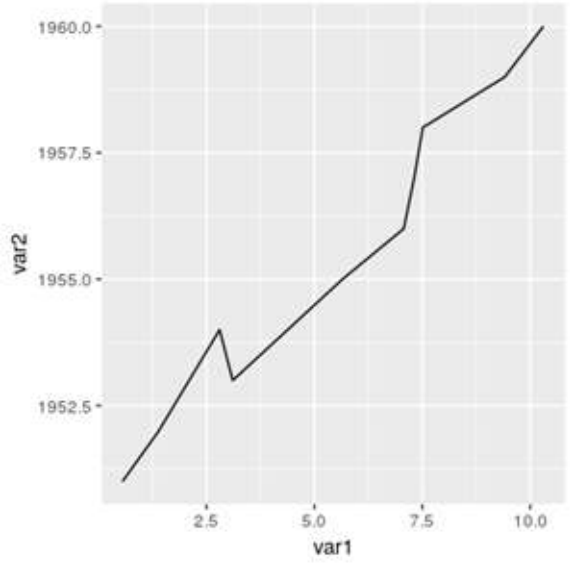
Gráfico de Linha representa informação através de dados em série, conectados por segmentos de linha reta

Linha → **geom_line()**

```
ggplot(dados, aes(x = var1, y = var2)) +
  geom_line()
```

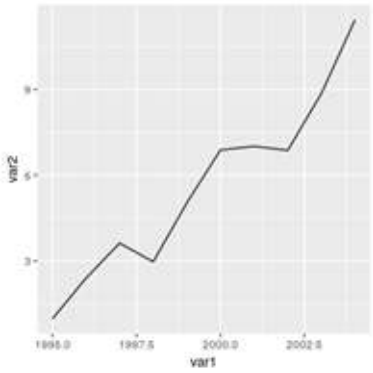
Variável numérica

Variável numérica



https://paternogbc.github.io/disciplina-graficos-r/

É possível combinar os dois?



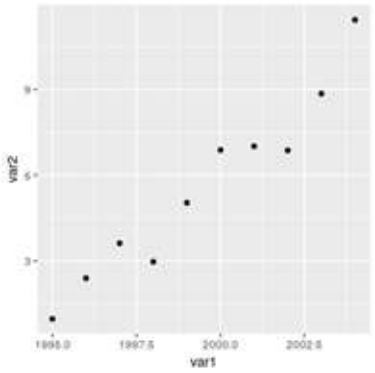
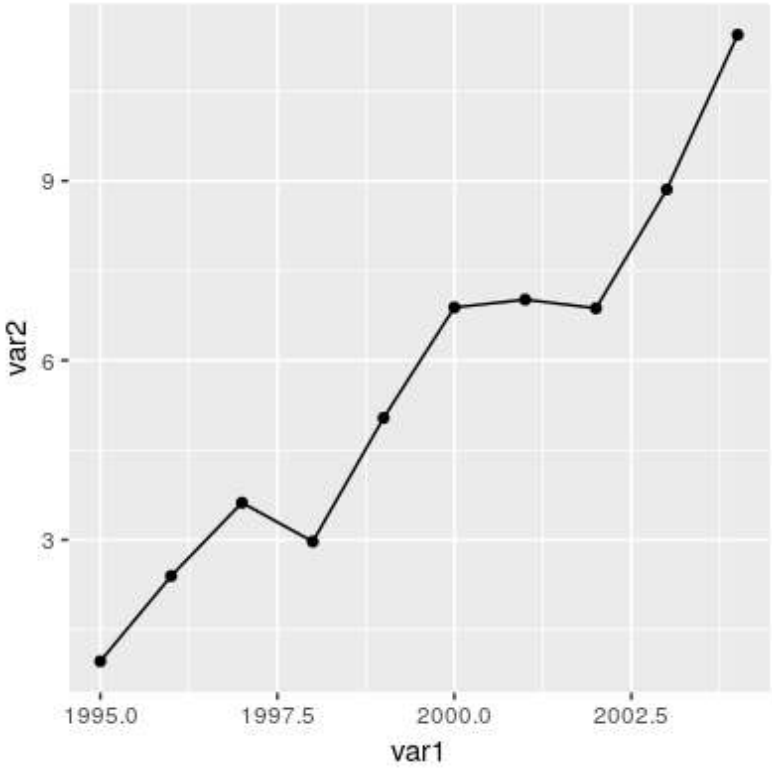
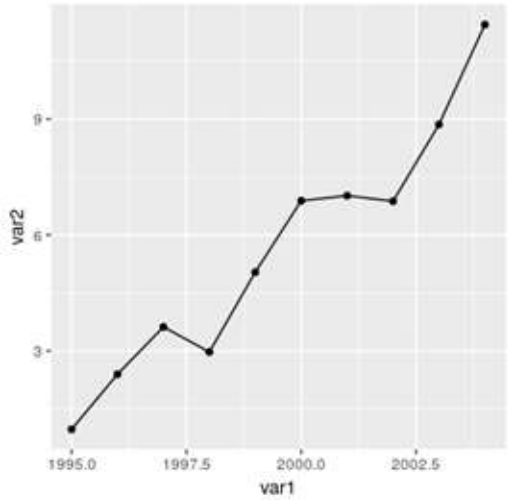


Gráfico de Dispersão com Linha



https://paternogbc.github.io/disciplina-graficos-r/

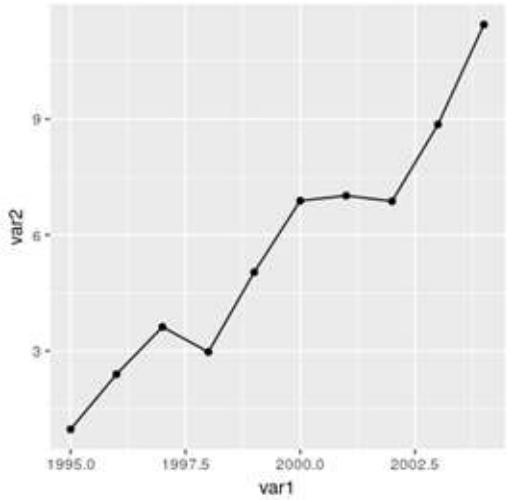
Desafio: Como seria o comando para juntar os dois?



https://paternogbc.github.io/disciplina-graficos-r/

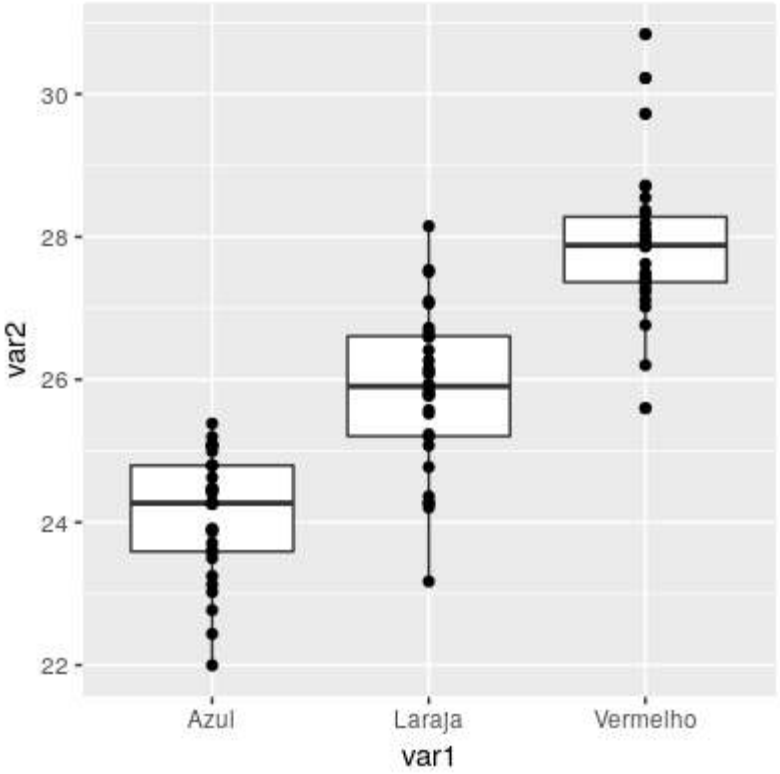
Desafio: Como seria o comando para juntar os dois?

```
ggplot(dados, aes(x = var1, y = var2)) +
  geom_line() +
  geom_point()
```



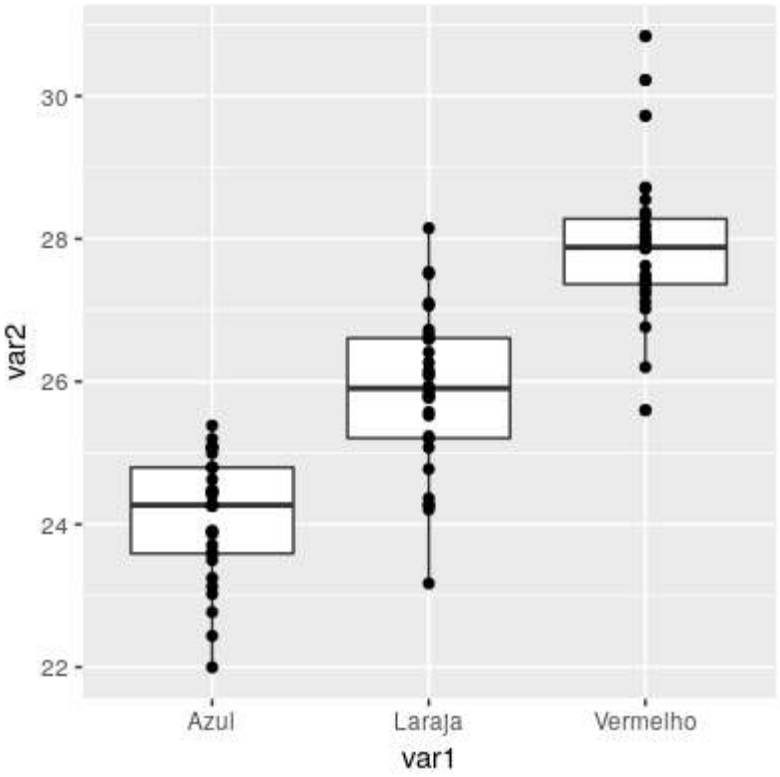
https://paternogbc.github.io/disciplina-graficos-r/

Desafio 2: Como este gráfico foi feito?



https://paternogbc.github.io/disciplina-graficos-r/

Desafio 2: Como este gráfico foi feito?

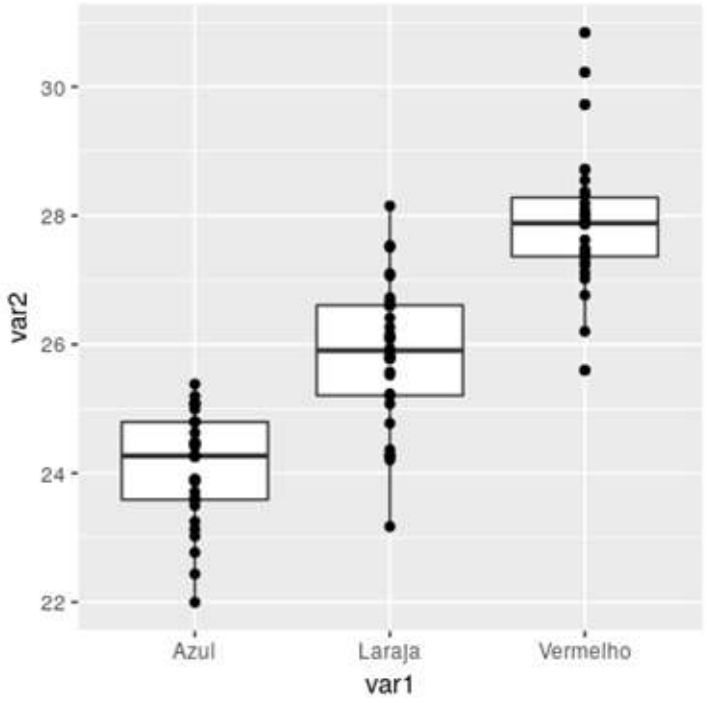



```
ggplot(dados, aes(x = var1, y = var2)) +
  geom_boxplot() +
  geom_point()
```

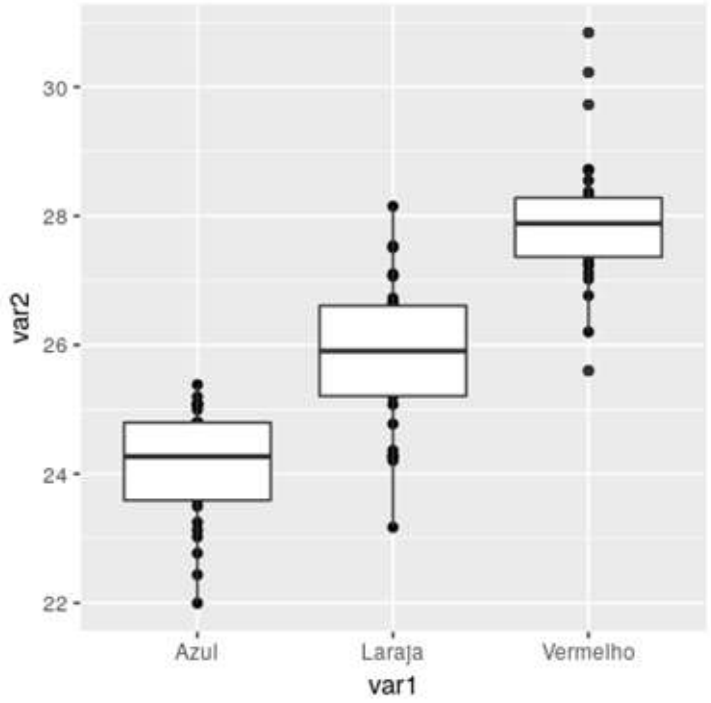
https://paternogbc.github.io/disciplina-graficos-r/

A ordem das camadas importa!

```
ggplot(dados, aes(x = var1, y = var2)) +
  geom_boxplot() +
  geom_point()
```

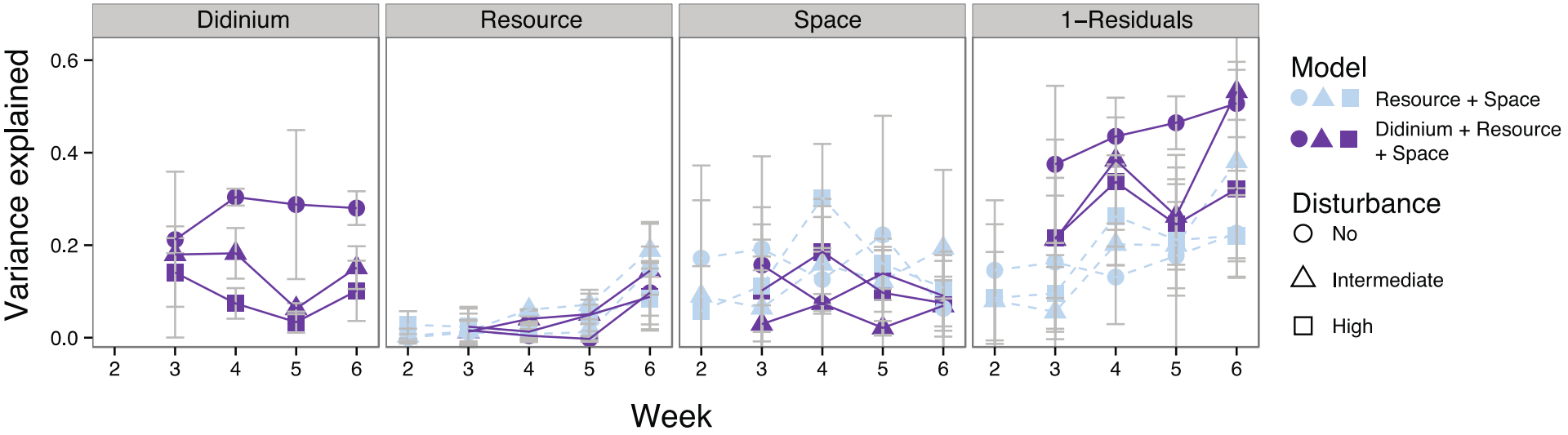


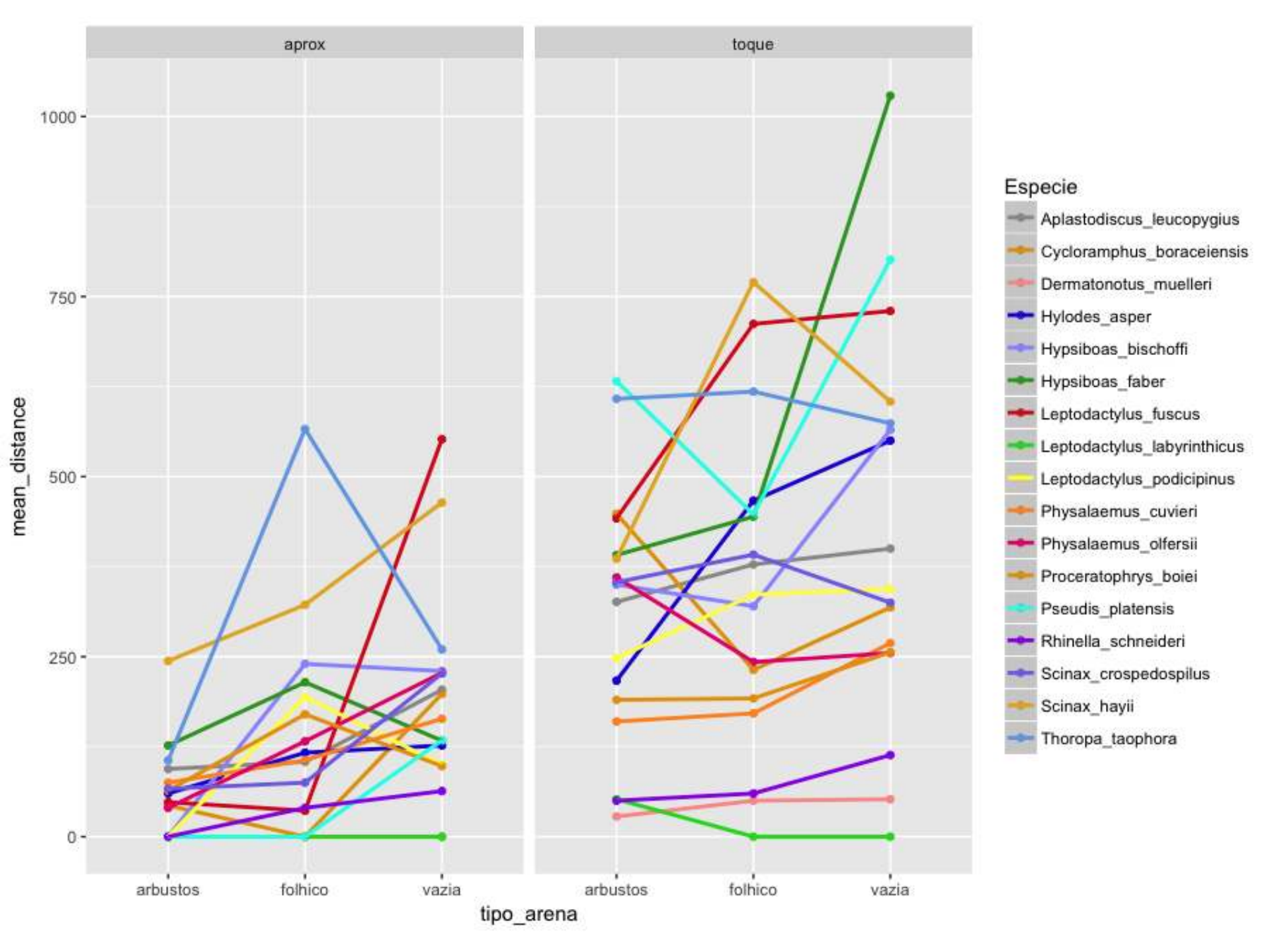

```
ggplot(dados, aes(x = var1, y = var2)) +
  geom_point() +
  geom_boxplot()
```

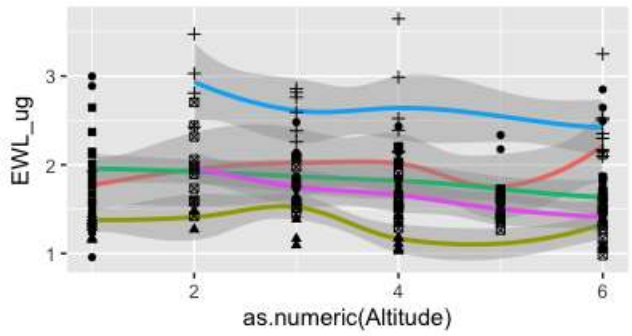


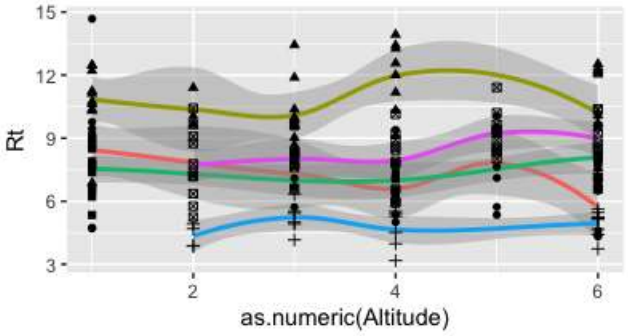
https://paternogbc.github.io/disciplina-graficos-r/

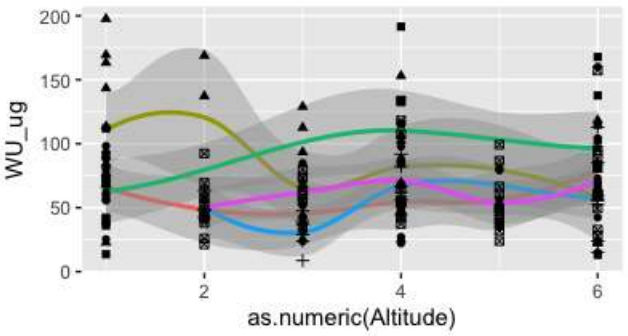
Exemplos do que se pode fazer com o ggplot2

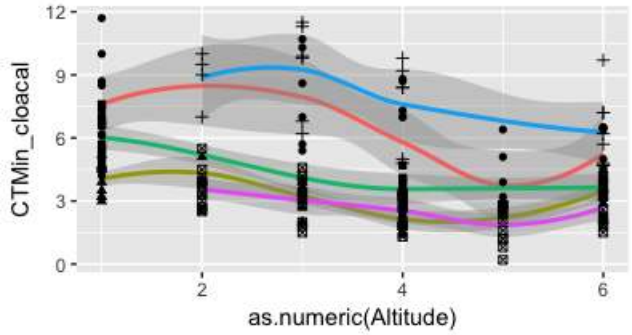

Didinium
Resource
Space
1−Residuals
Variance explained
0.0
0.2
0.4
0.6
2
3
4
5
6
Week
Model
Resource + Space
Didinium + Resource + Space
Disturbance
No
Intermediate
High

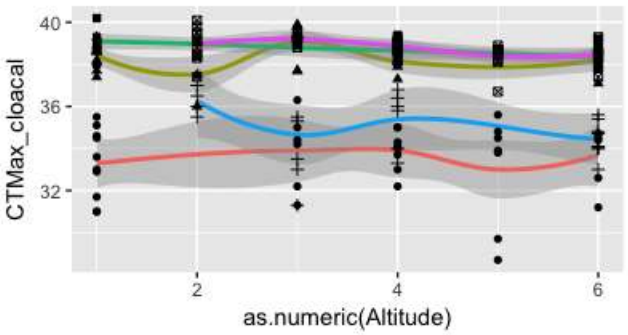

aprox
toque
1000
750
500
250
0
mean_distance
arbustos
folhico
vazia
tipo_arena
Especie
Aplastodiscus_leucopygius
Cycloramphus_boraceiensis
Dermatonotus_muelleri
Hylodes_asper
Hypsiboas_bischoffi
Hypsiboas_faber
Leptodactylus_fuscus
Leptodactylus_labyrinthicus
Leptodactylus_podicipinus
Physalaemus_cuvieri
Physalaemus_olfersii
Proceratophrys_boiei
Pseudis_platensis
Rhinella_schneideri
Scinax_crospedospilus
Scinax_hayii
Thoropa_taophora

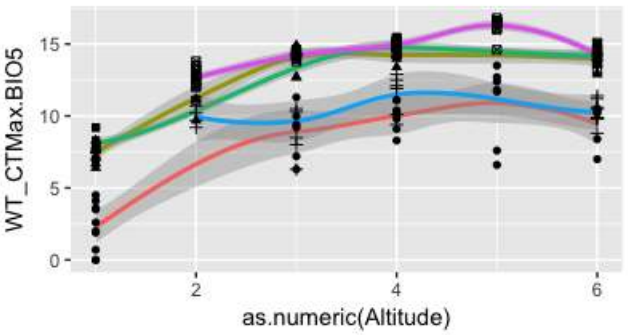

EWL_ug
as.numeric(Altitude)

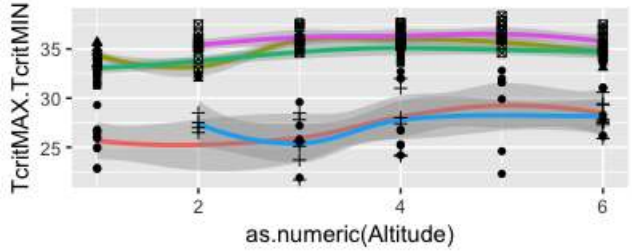

Rt
as.numeric(Altitude)


WU_ug
as.numeric(Altitude)

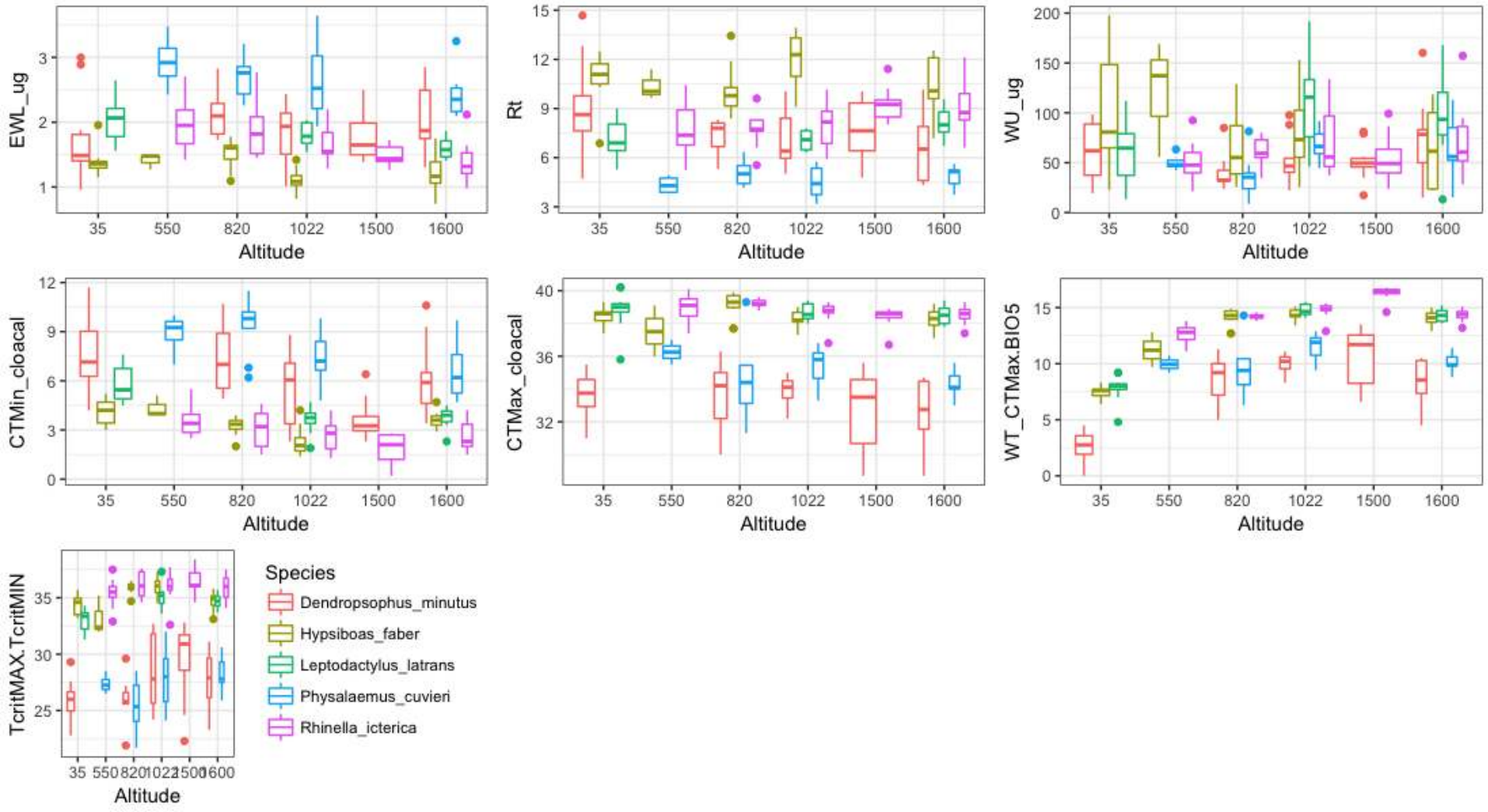

CTMin_cloacal
as.numeric(Altitude)

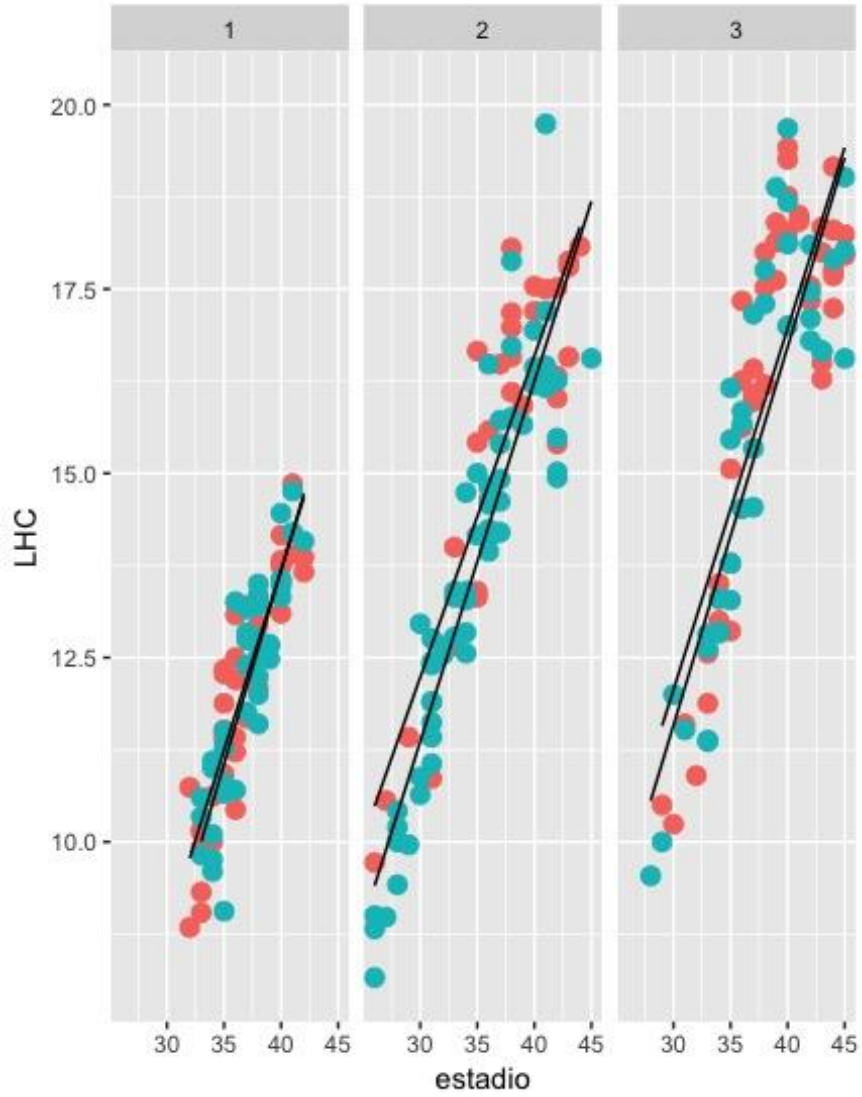

CTMax_cloacal
as.numeric(Altitude)

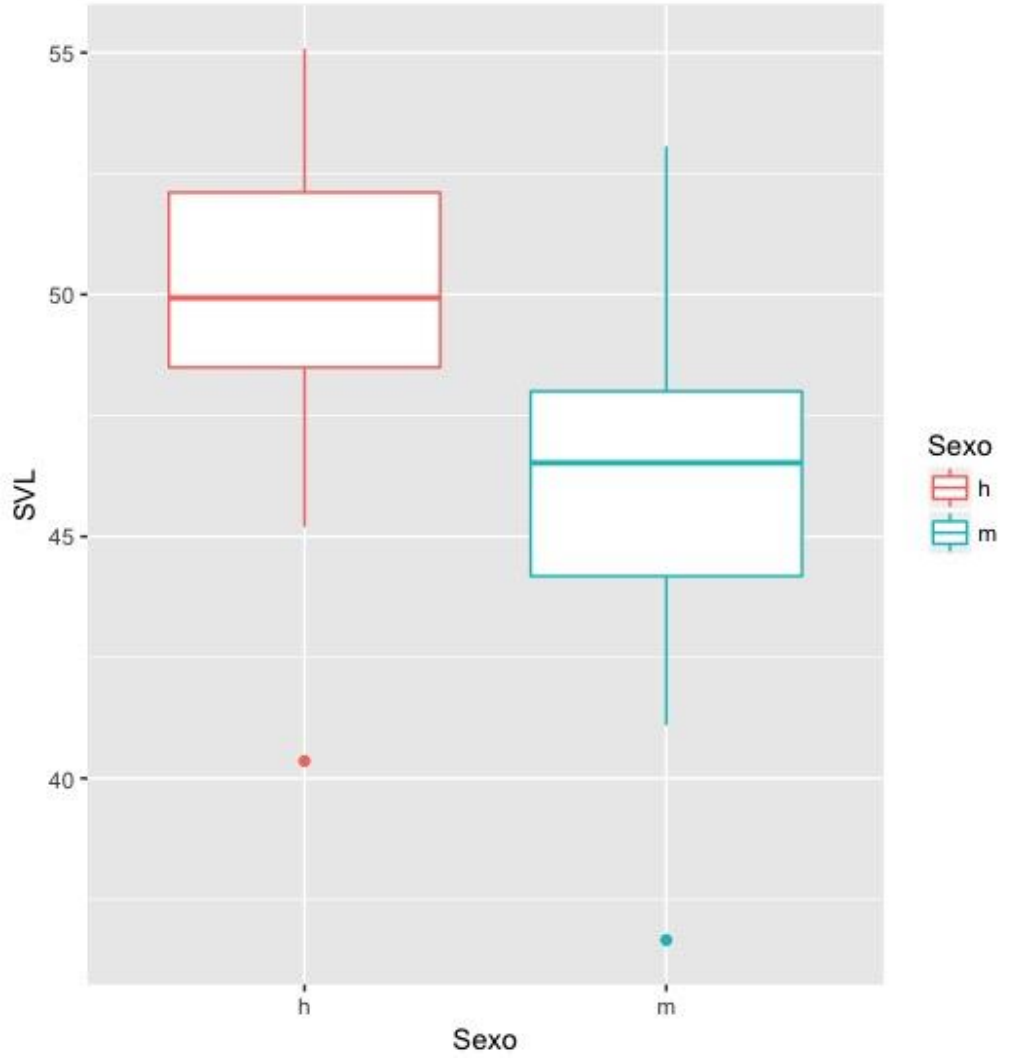

WT_CTMax.BIO5
as.numeric(Altitude)

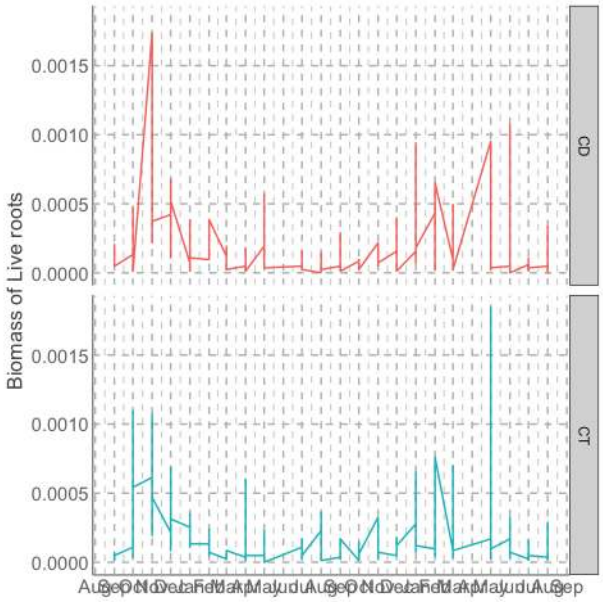

TcritMAX.TcritMIN
as.numeric(Altitude)

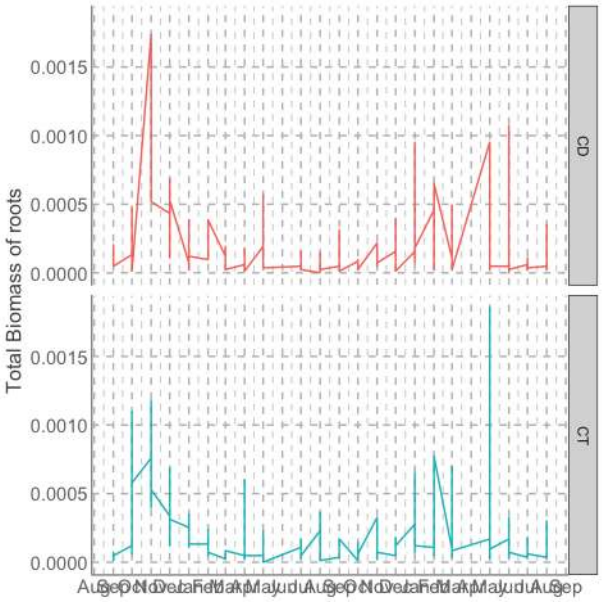
hus_minutus
Hypsiboas_faber
Leptodactylus_latrans
Physalaemus_cuvieri
Rhinella_icterica

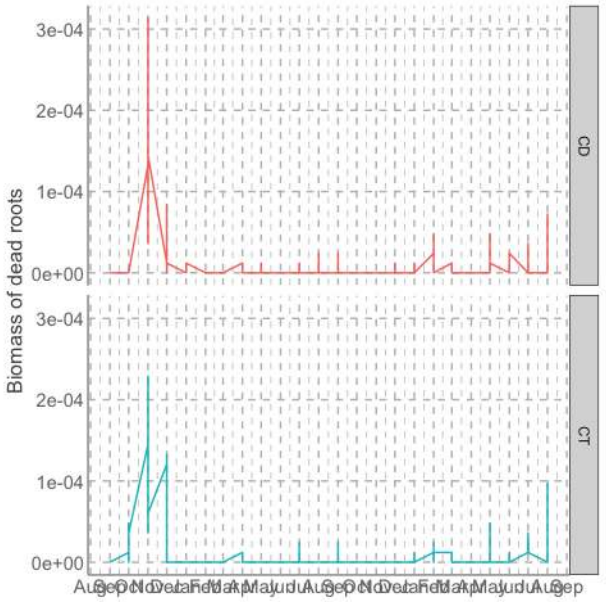

EWL_ug
Rt
WU_ug
CTMin_cloacal
CTMax_cloacal
WT_CTMax.BIO5
TcritMAX.TcritMIN
Altitude
35
550
820
1022
1500
1600
Species
Dendropsophus_minutus
Hypsiboas_faber
Leptodactylus_latrans
Physalaemus_cuvieri
Rhinella_icterica


1
2
3
LHC
20.0
17.5
15.0
12.5
10.0
30
35
40
45
estadio


SVL
55
50
45
40
h
m
Sexo
Sexo
h
m


Biomass of Live roots
0.0015
0.0010
0.0005
0.0000
CD
CT


Total Biomass of roots
0.0015
0.0010
0.0005
0.0000
CD
CT


Biomass of dead roots
3e-04
2e-04
1e-04
0e+00
CD
CT


YES!
AGORA BORA BOTAR A MÃO NA MASSA!

https://analises-ecologicas.com/cap6.html#grámatica-dos-gráficos

https://analises-ecologicas.com/cap6.html#exerc%C3%ADcios-2