

Introduction to ggplot2

08/12/2021

MAGDALENA DROŻDŻ

Try right now:
Open Rstudio
Run “library(tidyverse)“
Tell me about any errors
Go to: github.com/magdadrozdz/ggplot2-workshop

Workshop plan

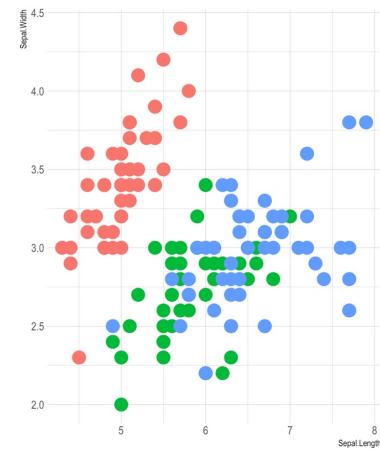
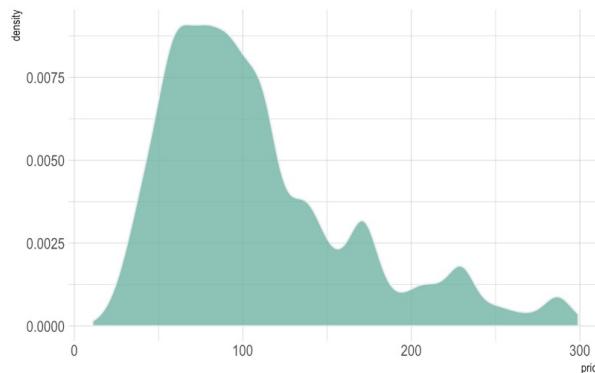
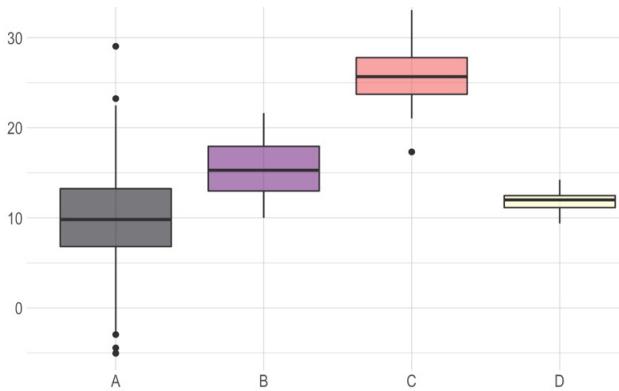
1. Introduction & what is ggplot2?
2. Plotting the first graphs
3. Troubleshooting and tips!



Artwork by @allison_horst

What is ggplot2?

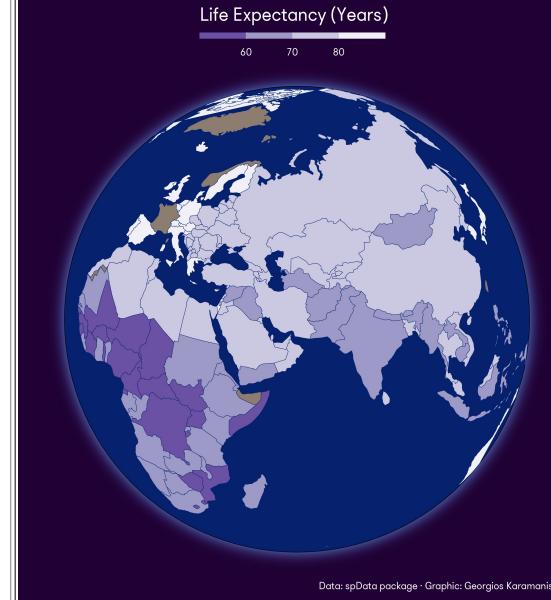
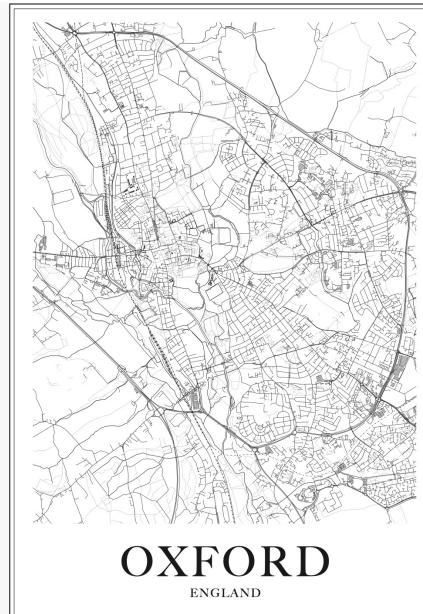
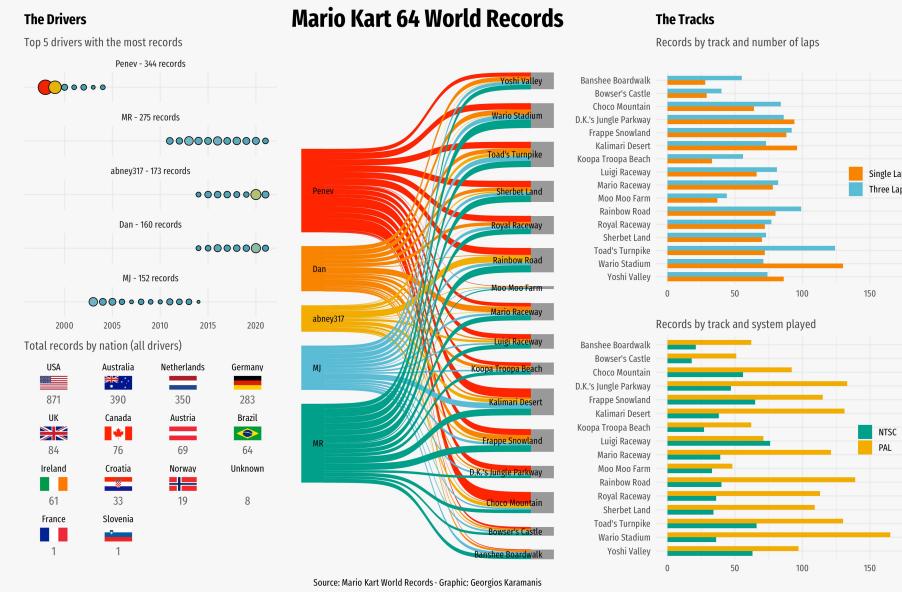
An R package used to create plots based on the grammar of graphics.



source: <https://www.r-graph-gallery.com>

What is ggplot2?

An R package used to create plots based on the grammar of graphics.



1st and 3rd graph by Georgios Karamanis

The Grammar of Graphics

A framework consisting of layered approach to visualize the data. In case of ggplot2 the components usually are:

1. Data – our datasets
2. Aesthetics – what do we want on the axes? Do we want to encode by colour/shape?
3. Scale – do the values need to be scaled?
4. Geometric objects (geoms) – points, bars, lines, or maybe a map?
5. Statistics
6. Facets – do we need subplots?
7. Coordinate system



Why not base R?

- much easier to make the plots pretty
- even defaults are more aesthetic
- more understandable syntax (...once you get used to it)
- easy faceting

Sometimes it's still a controversial topic 😊



David Robinson @drob · Feb 11, 2016

Short version of why @jtleek uses base plotting instead of ggplot2:

simplystatistics.org/2016/02/11/why... #rstats



Download Github code

<https://github.com/magdadrozdz/ggplot2-workshop>

The screenshot shows a GitHub repository page for 'magdadrozdz/ggplot2-workshop'. At the top right, there are three buttons: 'Go to file', 'Add file ▾', and a green 'Code ▾' button. Below these are two main cloning options: 'Clone' (with links for HTTPS, SSH, and GitHub CLI) and 'Open with GitHub Desktop'. At the bottom, a large button labeled 'Download ZIP' is highlighted with a green border.

Go to file Add file ▾ Code ▾

Clone

HTTPS SSH GitHub CLI

<https://github.com/magdadrozdz/ggplot2-workshop>

Use Git or checkout with SVN using the web URL.

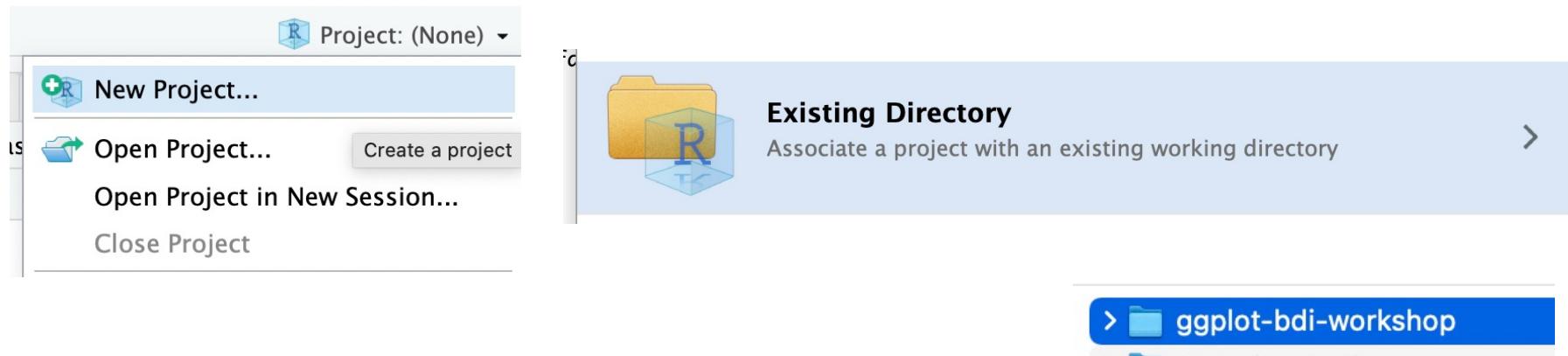
Open with GitHub Desktop

Download ZIP

- Intro to ggplot2.pdf
- Intro to ggplot2.pptx
- cars-worksheet-final.Rmd
- cars-worksheet.Rmd
- first-plot-final.Rmd
- first-plot.Rmd
- gene-expression.Rmd
- ggplot-bdi-workshop.Rproj
- tooth-growth-final.Rmd
- tooth-growth.Rmd

Rstudio & Rmarkdown

Right up corner



Rstudio & Rmarkdown

Plots inline

Easier to incorporate explanatory text

Can run one chunk at a time

Can create beautiful reports (but
that's another workshop material)

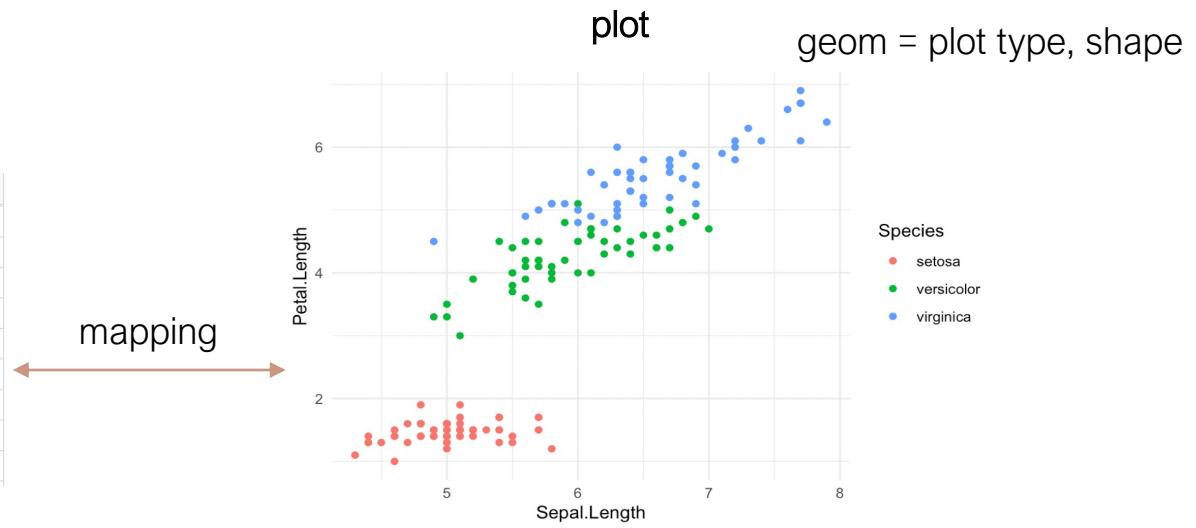


Artwork by @allison_horst

The basic elements of ggplot2 visualisations

data

Sepal.Length	Sepal.Width	Petal.Length	Petal.Width	Species
4.8	3.0	1.4	0.3	setosa
5.1	3.8	1.6	0.2	setosa
4.6	3.2	1.4	0.2	setosa
5.3	3.7	1.5	0.2	setosa
5.0	3.3	1.4	0.2	setosa
7.0	3.2	4.7	1.4	versicolor
6.4	3.2	4.5	1.5	versicolor
6.9	3.1	4.9	1.5	versicolor
5.5	2.3	4.0	1.3	versicolor



practical part

Long vs. wide data

genes <chr>	EV_01 <dbl>	EV_02 <dbl>	EV_03 <dbl>	KD_01 <dbl>	KD_02 <dbl>	KD_03 <dbl>
ENS000100	392.94135	395.02210	376.74416	464.18756	465.4274	475.21717
ENS000101	90.92910	71.32363	84.31325	99.95752	110.3702	81.14716
ENS000102	155.58151	171.62968	157.43093	378.40128	387.5866	374.06325
ENS000103	33.40566	41.43457	71.87467	250.78897	212.0394	228.09377

4 rows



genes <chr>	sample <chr>	replicate <chr>	expression <dbl>
ENS000100	EV	01	392.94135
ENS000100	EV	02	395.02210
ENS000100	EV	03	376.74416
ENS000100	KD	01	464.18756
ENS000100	KD	02	465.42735
ENS000100	KD	03	475.21717
ENS000101	EV	01	90.92910
ENS000101	EV	02	71.32363
ENS000101	EV	03	84.31325
ENS000101	KD	01	99.95752
ENS000101	KD	02	110.37022
ENS000101	KD	03	81.14716
ENS000102	EV	01	155.58151
ENS000102	EV	02	171.62968



Artwork by @allison_horst

Troubleshooting and resources



Artwork by @allison_horst

Error messages

```
> mean$x
```

```
Error in mean$x : object of type 'closure' is not subsettable
```

`Roses are red,
violets are blue,
object of type closure
is not subsettable`

Some are famously unclear...



Error messages

```
> ggplot(data = iris) +  
+     geom_point()  
Error: geom_point requires the following missing aesthetics: x and y  
Run `rlang::last_error()` to see where the error occurred.
```

```
> ggplot(data = iris,  
+         mapping = aes(x = Sepal.Length,  
+                           y = Sepal.Width)) %>%  
+     geom_point()  
Error: `mapping` must be created by `aes()`  
Did you use %>% instead of +?
```

```
> ggplot(data = iris,  
+         mapping = aes(x = Sepal.Length,  
+                           y = Sepal.With)) +  
+     geom_point()  
Error in FUN(X[[i]], ...) : object 'Sepal.With' not found
```

...but the ones in ggplot2 tend to be great!

Looking for help



“Rotating x axis labels ggplot2”

“How to add horizontal line ggplot2”

“How to centre a title in ggplot2”

or paste a relevant part of an error message

There are way too many options for us to remember. No shame in googling things!

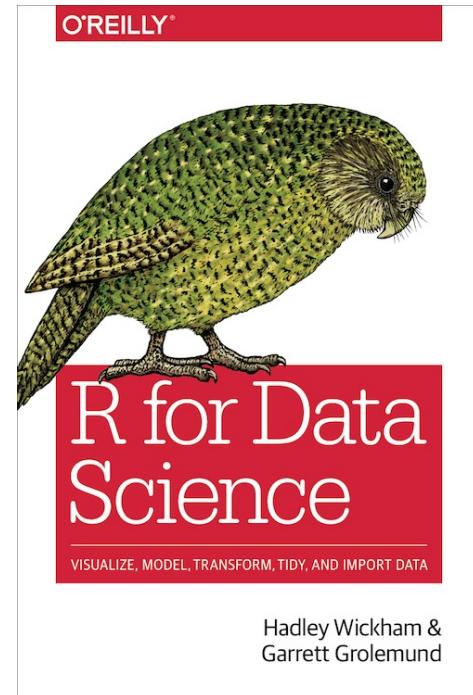
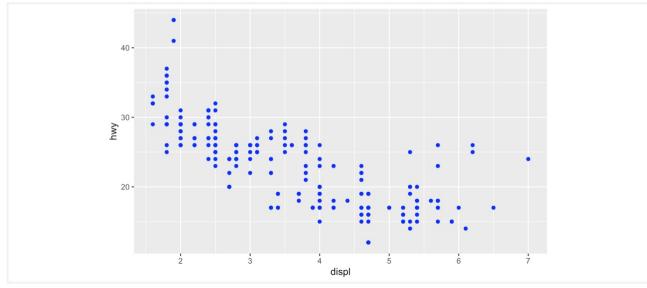
General R & ggplot2 resource

<https://r4ds.had.co.nz>

Amazing introduction to R through the Tidyverse.
Includes a whole chapter about ggplot2.

You can also set the aesthetic properties of your geom manually. For example, we can make all of the points in our plot blue:

```
ggplot(data = mpg) +  
  geom_point(mapping = aes(x = displ, y = hwy), color = "blue")
```



Cheatsheets – useful as a reference

Data visualization with ggplot2 :: CHEAT SHEET

Basics

ggplot2 is based on the **grammar of graphics**, which means that you can build every graph from the same components: a **data set**, a **coordinate system**, and **geoms** – **visual marks** that represent data points.



To display values, map variables in the data to visual properties of the geom (**aesthetics**) like size, color, and x and y positions.



Complete the template below to build a graph.

```
ggplot(data = DATA) +  
  GEOM_FUNCTIONS mapping = aes(MAPPINGS)  
  stat = STAT position = POSITION) +  
  COORD_FUNCTIONS +  
  SCALE_FUNCTIONS +  
  THEME_FUNCTIONS
```

`ggplot(mtcars, aes(x = cyl, y = hwy))` begins a plot that you finish by adding layers to. Add one geom function per layer.

`last_plot()` Returns the last plot.

`ggsave("plot.png", width = 5, height = 3)` saves last plot as 5" x 5" file named "plot.png" in working directory. Matches file type to file extension.

Aes

Common aesthetic values.

color and fill – string ("red", "#FFCCBC")

linetype – integer or string ("solid", "dotted", "dashed", "longdash", "longdash-dash", "twodash")

linejoin – string ("round", "butt", "square")

size – integer (line width in mm)

shape – integer/shape name or a single character ("a")

`geom_bar()`

Geoms

Use a geom function to represent data points, use the geom's aesthetic properties to represent variables. Each function returns a layer.

GRAPHICAL PRIMITIVES

```
b <- ggplot(economics, aes(date, unemploy))  
b + geom_point()
```

Ensure limits include values across all plots.

```
b + geom_curve(aes(yend = lat + 1,  
                    xend = long + 1), curvature = 1) -> yend,  
                    y, alpha, angle, curvature, linetype, size
```

```
b + geom_rect(aes(min = lat + 1,  
                    max = long + 1, ymin = lat,  
                    ymax = long + 1)) -> min, max,  
                    ymin, ymax, alpha, color, group, linetype, size
```

```
b + geom_ribbons(aes(unemploy = 400,  
                    ymax = unemploy + 400)) -> ymax, ymin,  
                    alpha, color, fill, group, linetype, size
```

Complete the template below to build a graph.

```
ggplot(data = DATA) +  
  GEOM_FUNCTIONS mapping = aes(MAPPINGS)  
  stat = STAT position = POSITION) +  
  COORD_FUNCTIONS +  
  SCALE_FUNCTIONS +  
  THEME_FUNCTIONS
```

`ggplot(mtcars, aes(x = cyl, y = hwy))` begins a plot that you finish by adding layers to. Add one geom function per layer.

`last_plot()` Returns the last plot.

`ggsave("plot.png", width = 5, height = 3)` saves last plot as 5" x 5" file named "plot.png" in working directory. Matches file type to file extension.

ONE VARIABLE

continuous

```
c <- ggplot(mtcars, aes(hwy)); c +  
  geom_area(stat = "bin")
```

```
c + geom_bar()
```

```
c + geom_density(kernel = "gaussian")
```

```
c + geom_dotplot()
```

```
c + geom_freqpoly()
```

```
c + geom_histogram(binwidth = 5)
```

```
c + geom_jitter(height = 2, width = 2)
```

```
c + geom_parallel()
```

```
c + geom_violin(scale = "area")
```

discrete

```
d <- ggplot(mpg, aes(rat))
```

```
d + geom_bar()
```

Three Variables

```
seals <- with(seals, sqrt(delta_long^2 + delta_lat^2)); l <- ggplot(seals, aes(long, lat))
```

```
l + geom_contour(aes(z = 2))
```

```
x, y, alpha, color, group, linetype, size, weight
```

```
l + geom_contour_filled(aes(z = 2))
```

```
x, y, alpha, color, fill, group, linetype, size, subgroup
```

discrete

```
d <- ggplot(mpg, aes(rat))
```

```
d + geom_bar()
```

geom that the mapping can

RStudio® is a trademark of RStudio, PBC • CC BY SA RStudio • info@rstudio.com • 844-448-1212 • rstudio.com • Learn more at ggplot2.tidyverse.org • ggplot2 3.3.5 • Updated: 2021-08-06



Asking the right questions

99% of the time, someone has already asked the same question. The trick is to find it. If we really can't find the answer, we can ask on Stack Overflow.

Some tips for getting a quick answer:

1. Include a reproducible chunk of data (or use the inbuild datasets),
2. Describe what happens is happening and what code you are currently using,
3. Show what you would like to achieve (if it's not obvious)



Rotating and spacing axis labels in ggplot2

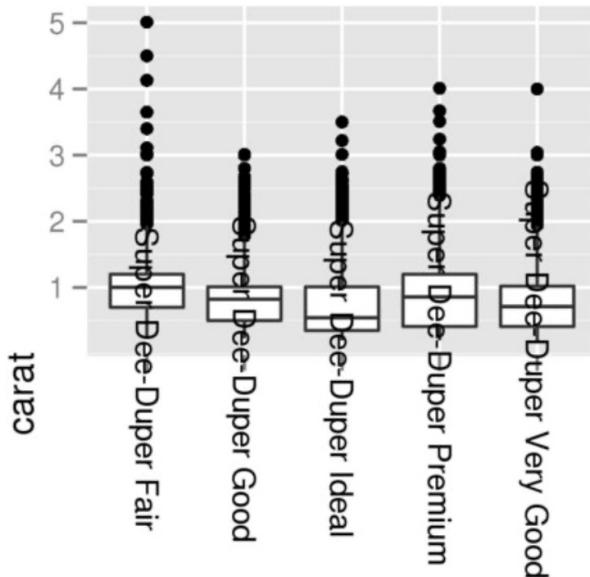
Asked 12 years, 3 months ago Active 11 months ago Viewed 1.0m times

clear title

(look how many views!)

I have a plot where the x-axis is a factor whose labels are long. While probably not an ideal visualization, for now I'd like to simply rotate these labels to be vertical. I've figured this part out with the code below, but as you can see, the labels aren't totally visible.
780

```
data(diamonds)
diamonds$cut <- paste("Super Dee-Duper",as.character(diamonds$cut))
q <- qplot(cut,carat,data=diamonds,geom="boxplot")
q + opts(axis.text.x=theme_text(angle=-90))
```



short description

showing the current code

using inbuilt dataset – easy to reproduce

obvious what needs to be corrected

Got an answer quickly!

Change the last line to

```
q + theme(axis.text.x = element_text(angle = 90, vjust = 0.5, hjust=1))
```

Chart ideas

The R Graph Gallery

<https://www.r-graph-gallery.comx>

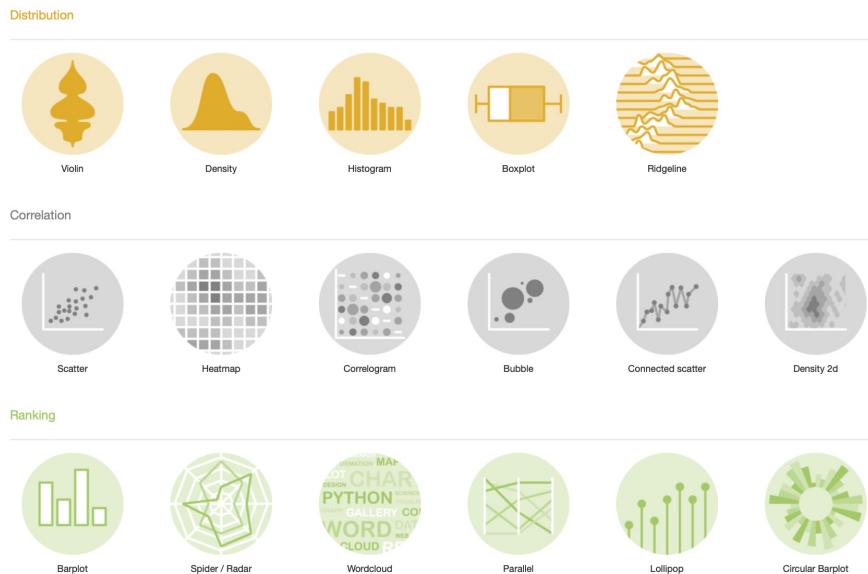
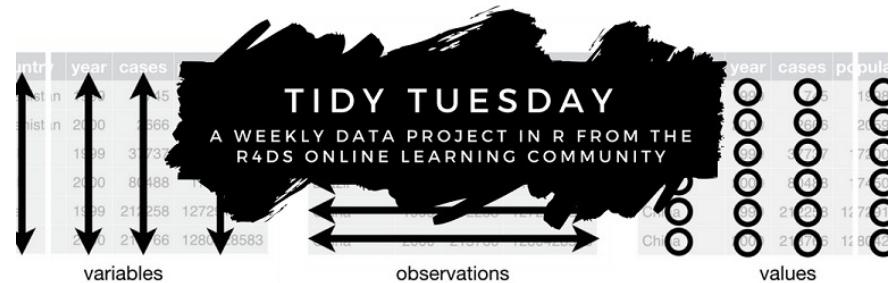


Chart ideas

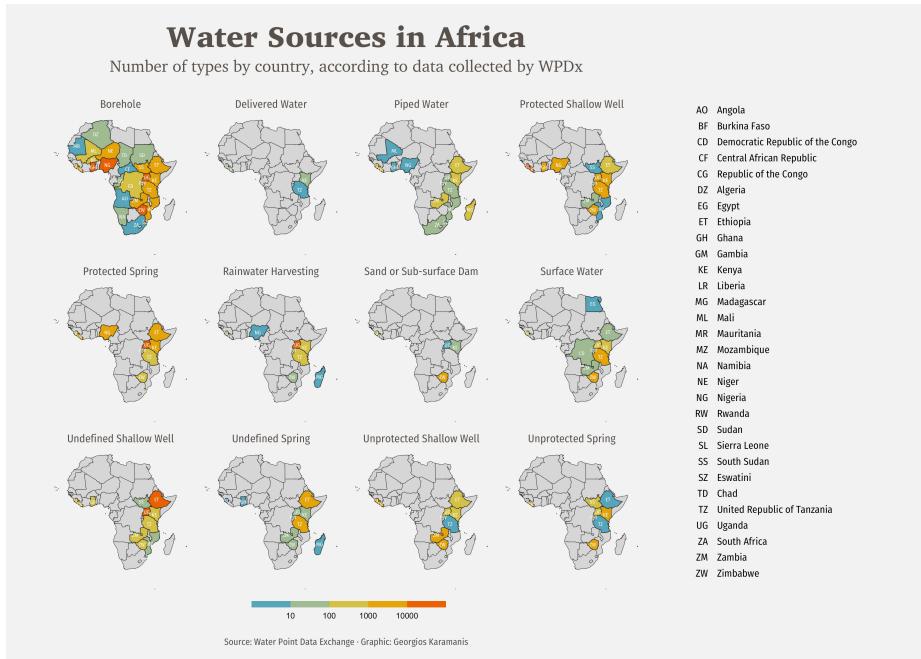
- Code attached to some publications (still quite rare)
 - TidyTuesday projects

<https://github.com/rfordatascience/tidyTuesday>

<https://github.com/qkaramanis/tidytuesday>



TidyTuesday examples



↔

```

39  cl <- ww %>
40  distinct(country_name, iso2) %>%
41  arrange(iso2) %>%
42  mutate(n = row_number())
43
44  pal <- wesanderson::wes_palette("Zissou1", 6, type = "continuous")
45
46  f1 <- "Fira Sans Condensed"
47  f2 <- "Charter Roman"
48  f2b <- "Charter Bold"
49  f2bb <- "Charter Black"
50
51  p <- ggplot(ww, aes(geom = "point", fill = "grey95", size = 0.2, color = "grey20") +
52    geom_sf(data = africa, fill = "#58514B", size = 1))
53  geom_sf(aes(geometry = geometry, fill = n, color = "grey20", size = 0.3)) +
54  geom_sf_text(aes(geometry = geometry, label = iso2), size = 2, family = f1, color = "white") +
55  # scale_fill_stepsn(colors = pal, n.breaks = 7, labels = unit_format(unit = "K", scale = 1e-3)) +
56  # scale_fill_stepsn(colors = pal, n.breaks = 7, labels = unit_format(unit = "K", scale = 1e-3)) +
57  # facet_wrap(vars(water_source)) +
58  theme_void()
59
60  plot.background = element_rect(fill = "grey95", color = NA)
61  strip.text = element_text(family = f1, size = 12, margin = margin(0, 0, 5, 0), color = "#58514B"),
62  plot.title = element_text(hjust = 0.5, family = f2bb, size = 32, color = "#58514B"),
63  plot.subtitle = element_text(hjust = 0.5, family = f2, size = 18, margin = margin(15, 0, 30, 0), color = "grey30"),
64  plot.caption = element_text(hjust = 0.5, family = f1, color = "grey30", size = 10, margin = margin(20, 310, 10, 45))
65
66  c <- ggplot(cl, aes(0, -n)) +
67  geom_text(aes(label = country_name), hjust = 0, family = f1) +
68  geom_text(aes(label = iso2), hjust = 1, nudge_x = -0.02, family = f1) +
69  xlim(-1, 0.5) +
70  ylim(-30, 0) +
71  theme_void()
72
73  ggdraw(p) +
74  draw_plot(c, scale = 0.8, x = 0.15) +
75  theme(
76    plot.background = element_rect(fill = "grey95", color = NA)
77  )

```

source: Georgios Karamanis

Thank you!



Artwork by @allison_horst