

Introduction to \mathcal{R}

Session 4: ggplot2 Graphics

Dag Tanneberg¹

Potsdam Center for Quantitative Research
University of Potsdam, Germany
October 11/12, 2018

¹Chair of Comparative Politics, UP, dag.tanneberg@uni-potsdam.de

Before we start

- 1 Please, run the following code.

```
install.packages(c("tidyverse", "gapminder"))  
for(p in c("tidyverse", "gapminder")) {  
  library(p, character.only = TRUE)  
}
```

- 2 I could not replicate the `data$[abbreviated var name]` issue.
- 3 Mac users with a locale warning on startup: Go to [Stack Overflow](#) or [rstudio.com](#)

Introduction

A Taste of ggplot2

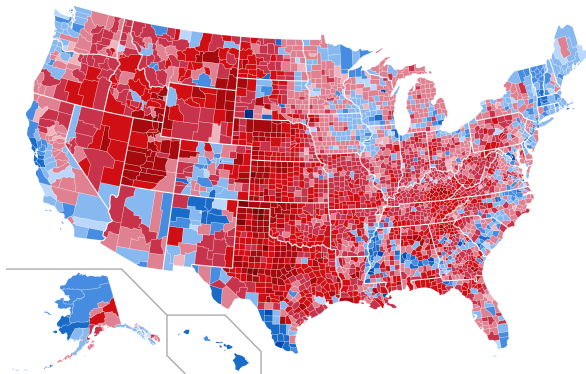


Figure 1: Vote Share in the 2012 Presidential Elections by County²

²Healan, K. forthc. Data Visualization. A Practical Introduction. <https://socviz.co/maps.html> (last access: 10/03/2018).

More Tasting: London Bicycle Routes³



³Cheshire, J. 2012. Great maps with ggplot2.

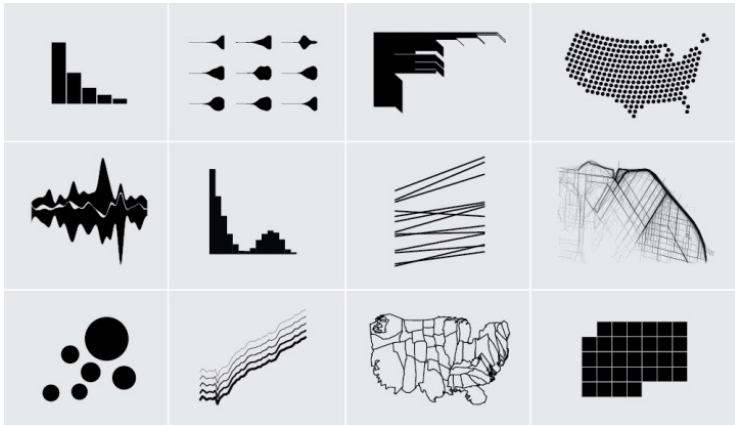
<http://spatial.ly/2012/02/great-maps-ggplot2/> (last access: 10/06/2018).

Even More Tasting: Hierarchical Edge Bundling⁴



⁴Holtz, Y. 2017. The R Graph Gallery. <https://bit.ly/2QKdhLU> (last access: 10/06/2018).

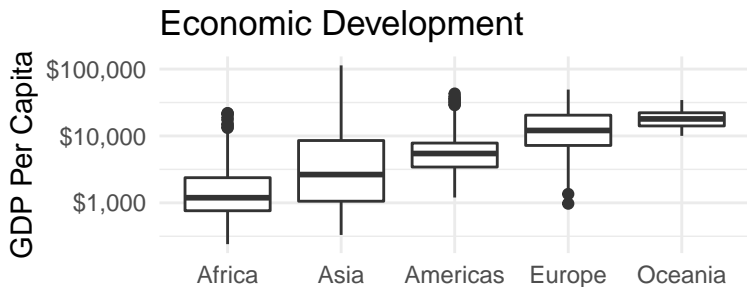
An Embarrassing Riches of Options⁵



⁵Yau, N. 2016. 5 Tips for Learning to Code for Visualization.

<https://bit.ly/2yu0fe0> (last access: 10/11/2018).

We will create this modest graph together.



Source: Gapminder.

Outline

- 1 Introduction
- 2 What is ggplot2?
- 3 How do you talk to ggplot2?
- 4 Summary

What is ggplot2?

Graphics Engines in \mathcal{R}

- Three major graphics systems in \mathcal{R}
 - base graphics: ships with \mathcal{R} ⁶
 - lattice developed by Deepayan Sarkar
 - ggplot2 developed by Hadley Wickham
- All build on Paul Murrell's grid Graphics
- All differ remarkably on usability & quality of output

⁶Check out Nathan Yau's tutorials at [Flowing Data](#).

What's the buzz about ggplot2?

- Implements the Grammar of Graphics⁷

“In brief, the grammar tells us that a statistical graphic is a mapping from data to aesthetic attributes (colour, shape, size) of geometric objects (points, lines, bars).”⁸

- Very generic, schematic approach to data viz
- Flexible themes & fast, high quality results
- Many users, active community, numerous extension
 - Create animations, networks, maps, dags, etc.

⁷Wilkinson, L. 1999/2005. The Grammar of Graphics.

New York: Springer Science + Business.

⁸Wickham, H. 2009. ggplot2. New York: Springer Science + Business.

How do you talk to ggplot2?

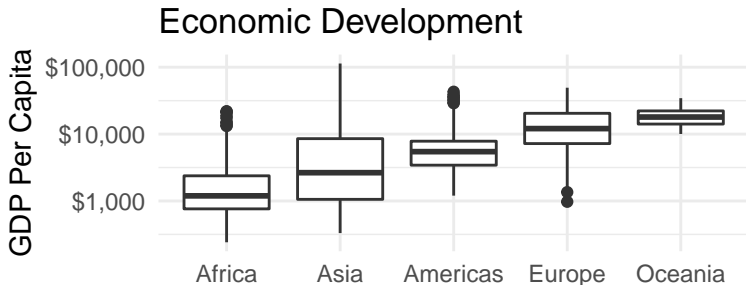
Intuition

- Idea: logical connection b/w data (e.g., gender) and plot elements (e.g., shapes)
 - a.k.a. aesthetic mapping
- Answer the following questions for ggplot2:
 - 1 What is your data?
 - 2 What relationships you want to see?
 - 3 How do you want to see it?
 - 4 What additional information do you want to see?
 - 5 What scales, axes, labels should be shown?
- Complete reference:
 - Hadley's book: **ggplot2**
 - ggplot2 Online Reference: **tidyverse**
 - Kieran Healy's book: **Data Visualization**

A schematic for making a plot

```
p <- ggplot(  
  data = <data>, # 1. What is your ("tidy") data?  
  mapping = aes( # 2. What var's map unto the plot?  
    <aesthetic> = <variable>, # x  
    <aesthetic> = <variable>, # y  
    <...> = <...> # colour, fill, shape, size,  
    # alpha, ...  
  )  
  ) + # Add layers to your plot  
  geom_<type>(<...>) + # Define your plot type  
  scale_<mapping>_<type>(<...>) + # Adjust scales  
  coord_<type>(<...>) + # Adjust co-ordinates  
  labs(<...>) # Label plot elements  
ggsave(file = <...>, plot = <...>, ...) # Save ur work
```

Our target graph



Source: Gapminder.

- Install and load packages `ggplot2` and `gapminder`.
- Answer questions 1-5 on paper.
- Use the generic schematic and **tidyverse** to generate the plot.

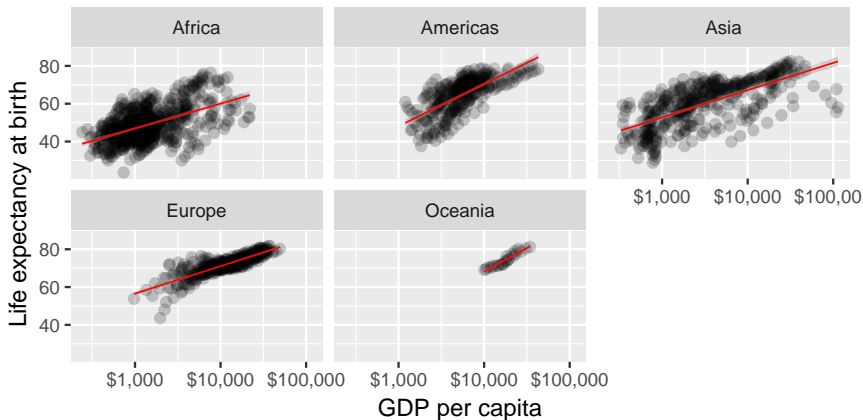
Summary

Summary

- There are three major graphics engines: base, ggplot, lattice.
- ggplot2's attraction
 - very generic, schematic approach to viz
 - fast, visually pleasing results
- Principle strategie:
 - a. Connect (map) data to plot elements
 - b. Layer plot elements
- Read Healan, K. forthc. Data Visualization. A Practical Introduction. <https://socviz.co/maps.html> (last access: 10/03/2018).

Self-practice: Replicate these figures.

Life Expectancy vs. Economic Development



Self-practice: Replicate this figure.

Life Expectancy vs. Economic Development

Graph shows 1997, 2003, and 2007.

