Data visualization and Communication

Ewha GSIS Computational Social Science Workshop II

Iegor Vyshnevskyi
Woosong University | KDI School
Saturday, November 19, 2022

Agenda

1. Introduction
2. The Roots
3. Some practical tips

1. Introduction

About me

Hello! My name is *Iegor*.

- Assistant Professor, Woosong University
- Ph.D candidate, KDIS
- MSc, KDIS (2014)
- MSc, KNEU (2009)
- Background: international finance & central banking
- Research interests: banking and central banking, computational data science

Some Info:

- Google Scholar
- Linkedin
- ResearchGate
- GitHub
- email: ievysh@kdis.ac.kr

Data visualization

> "*Data doesn't speak for itself*, so you need to analyze it and make the findings accessible by presenting them with effective visualization."

- Dr. Jae Yeon Kim

> "Data visualization is a great way to simplify data and show it in a form that is understandable, insightful, and actionable.
>
> **Data visualization is being increasingly seen as the vital final step of any successful data-driven analytics plan.**"

- Caroline Lee

Jesse Maegan
@kierisi

Following

My #rstats learning path:

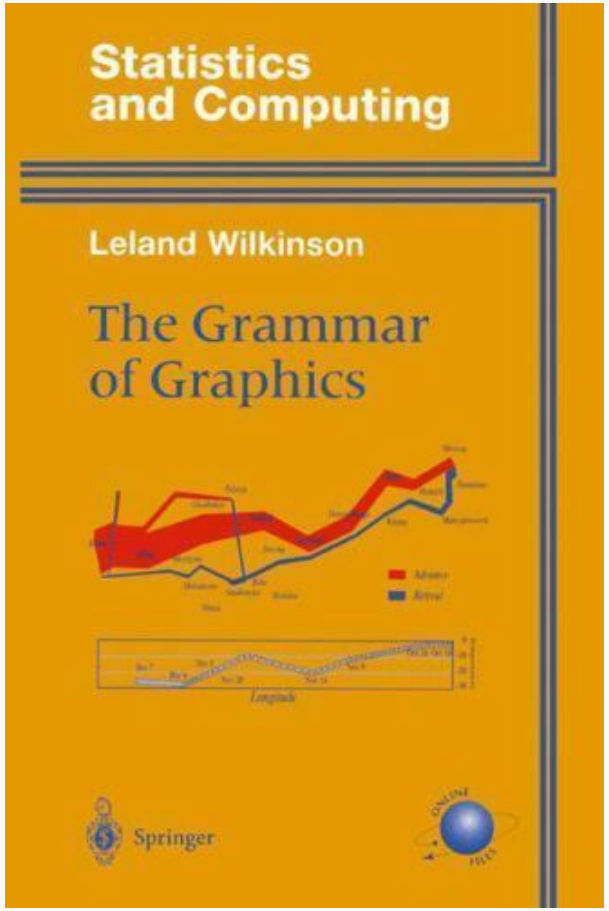
1. Install R
2. Install RStudio
3. Google "How do I [THING I WANT TO DO] in R?"

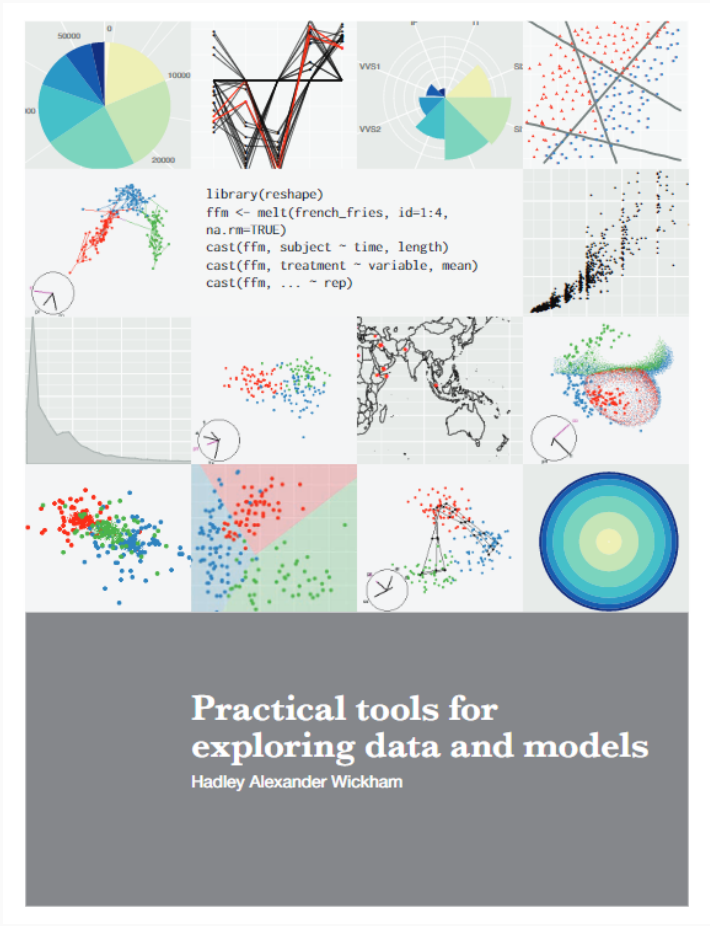
Repeat step 3 ad infinitum.

7:19 AM - 18 Aug 2017

2. The Roots

The beginnings of ggplot and reshape





3. Some practical tips

To remember

> "In good information visualization, there are no rules, no guidelines, no templates, no standard technologies, no stylebooks...*You must simply do whatever it takes.*"

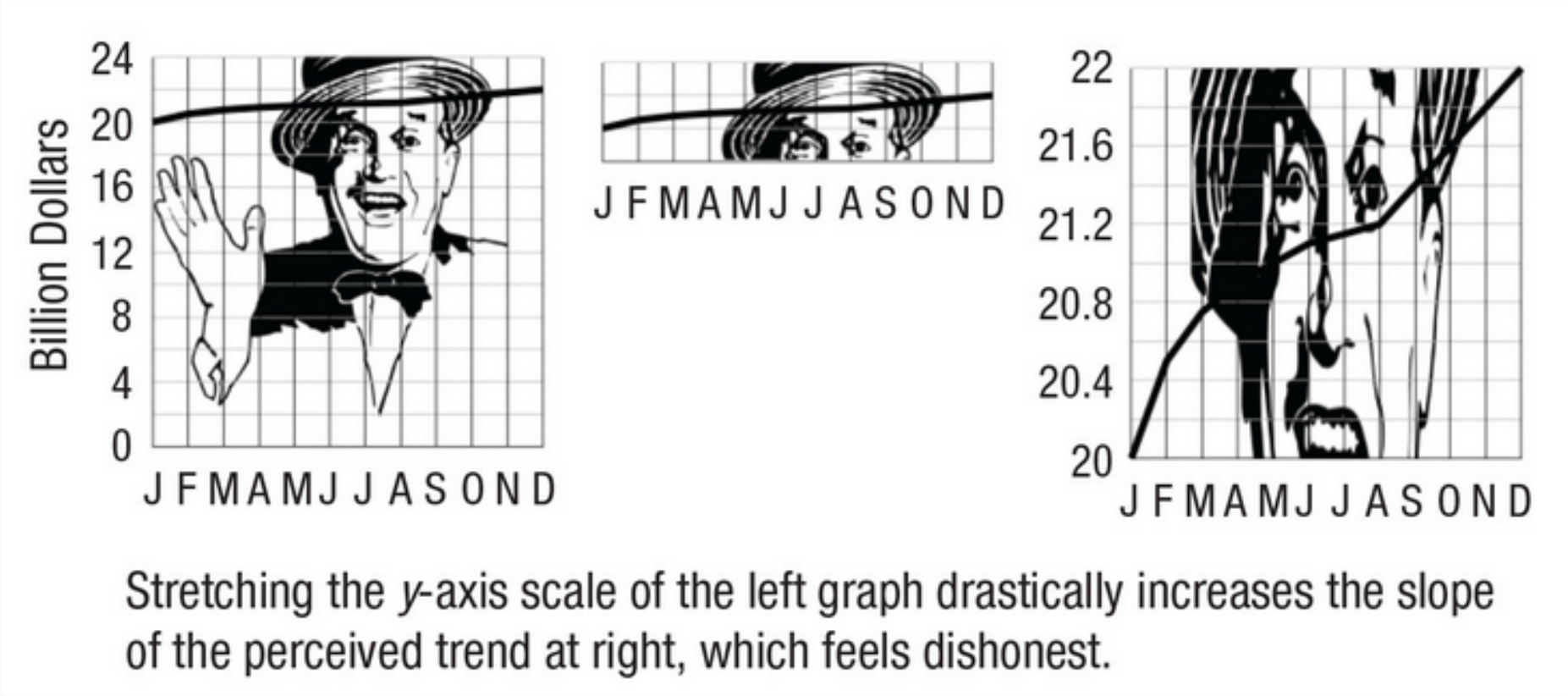
- Edward Tufte

> "Whenever possible, visualize your data with *solid, colored shapes* rather than with lines that outline those shapes. Solid shapes are more easily perceived, are less likely to create visual artifacts or optical illusions, and do more immediately convey amounts than do outlines."

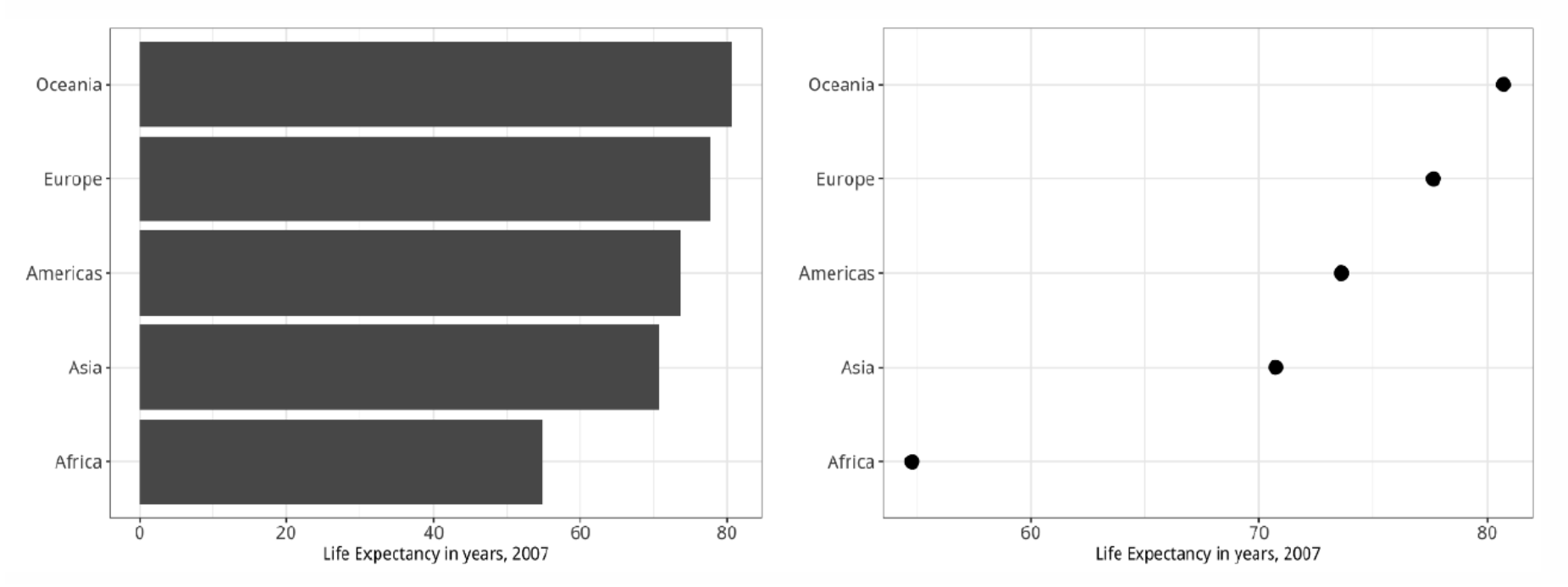
- Claus Wilke

Don't play with y-axis scale

Franconeri et al (2021). The Science of Visual Data Communication: What Works.



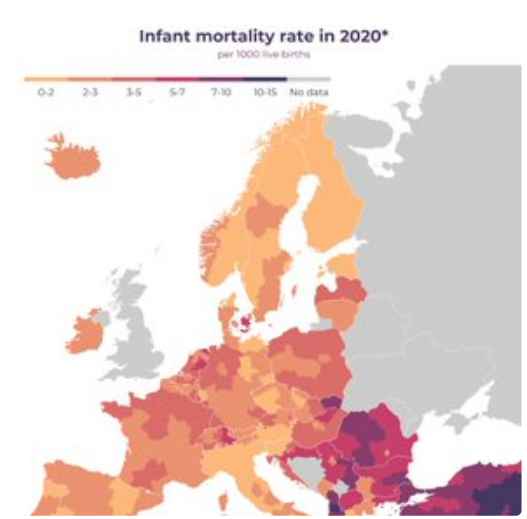
Don't play with x-axis scale



Use color to your advantage

Milos Popovic
@milos_agathon

My new map shows infant mortality rate according to the most recent Eurostat data#infant #mortality #Europe #RStats #DataScience #dataviz #maps

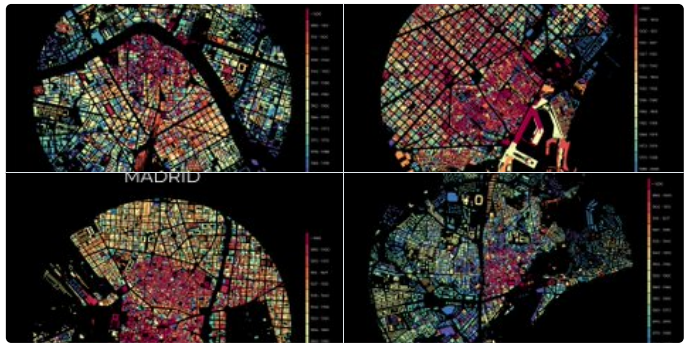


♡ 46 5:14 PM - Nov 15, 2022

18 people are talking about this

Dr. Dominic Royé (mstdn.social/@dr_xeo)
@dr_xeo

El éxito inesperado del mapa 🗺️ de Valencia con muchas peticiones de todos vosotros me ha llevado a ofreceros un formato pdf de Valencia, Barcelona, Madrid y Málaga. ¡Gracias a todos! 🤩



♡ 5,487 11:14 PM - May 4, 2021

1,395 people are talking about this

There are many more dos and don'ts

1. Drop legends, use annotations & bigger labels
2. Never use dual axes (some exceptions: Fahrenheit vs Celsius)
3. Sort bar charts (i.e., ascending/descending order)
4. Use shapes & colors (double-encoding data)
5. Use 3D only when really necessary
6. Use pie charts with few categories (otherwise, bar charts)

Some readings

1. Franconeri et al (2021). The Science of Visual Data Communication: What Works
2. Kieran Healy (2019). Data Visualization
3. Claus Wilke (2019). Fundamentals of Data Visualization
4. Hadley Wickham et al (2009). ggplot2: Elegant Graphics for Data Analysis
5. Garrick Aden-Buie. A Gentle Guide to the Grammar of Graphics with ggplot2

Let's have some fun!