

How to make graphs in R

From default to publication-quality graphics

@anchu

2th Hanoi UseRs Meetup

Thank you for coming!



About me

Download this PDF

<https://speakerdeck.com/chuvanan>

A word cloud centered around the word "visualization". The words are colored in shades of blue, green, and red. The size of each word indicates its frequency or importance in the context of visualization.

The most prominent words are:

- visualization
- information
- processing
- scientific
- perception
- understandable
- application
- computer
- network
- presentation
- interaction
- big
- data
- image
- users
- statistical
- experience
- graphs
- diagrams
- journalism
- plot
- design
- display
- communication
- analysis
- inforgraphics
- graphics
- objects

This talk is . . .

This talk is . . .

- ▶ not really for beginners

This talk is . . .

- ▶ not really for beginners
- ▶ not really for experts either

This talk is . . .

- ▶ not really for beginners
- ▶ not really for experts either
- ▶ not a comprehensive treatment of data visualization

This talk is . . .

- ▶ not really for beginners
- ▶ not really for experts either
- ▶ not a comprehensive treatment of data visualization
- ▶ not about design

This talk is . . .

- ▶ not really for beginners
- ▶ not really for experts either
- ▶ not a comprehensive treatment of data visualization
- ▶ not about design
- ▶ not about R's ggplot2 (sorry ggplot2's folks)

This talk is . . .

This talk is . . .

- ▶ a gentle introduction to graphing data with R

This talk is . . .

- ▶ a gentle introduction to graphing data with R
- ▶ (mostly) about statistical graphics

This talk is . . .

- ▶ a gentle introduction to graphing data with R
- ▶ (mostly) about statistical graphics
- ▶ (entirely) about R's base graphics

This talk is . . .

- ▶ a gentle introduction to graphing data with R
- ▶ (mostly) about statistical graphics
- ▶ (entirely) about R's base graphics
- ▶ (hopefully) helping you to improve your visualization skills

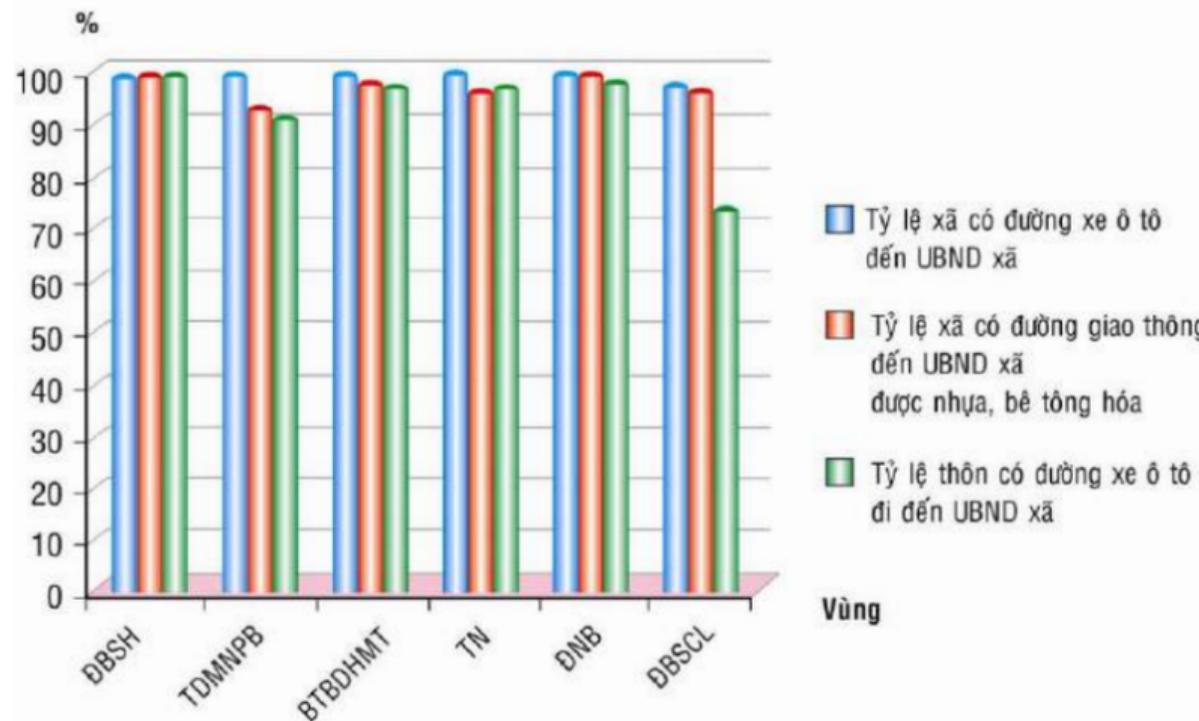
This talk is . . .

- ▶ a gentle introduction to graphing data with R
- ▶ (mostly) about statistical graphics
- ▶ (entirely) about R's base graphics
- ▶ (hopefully) helping you to improve your visualization skills
- ▶ (hopefully) fun and entertaining

Statistical Graphics

Some (frequent) encounters from the web/publications

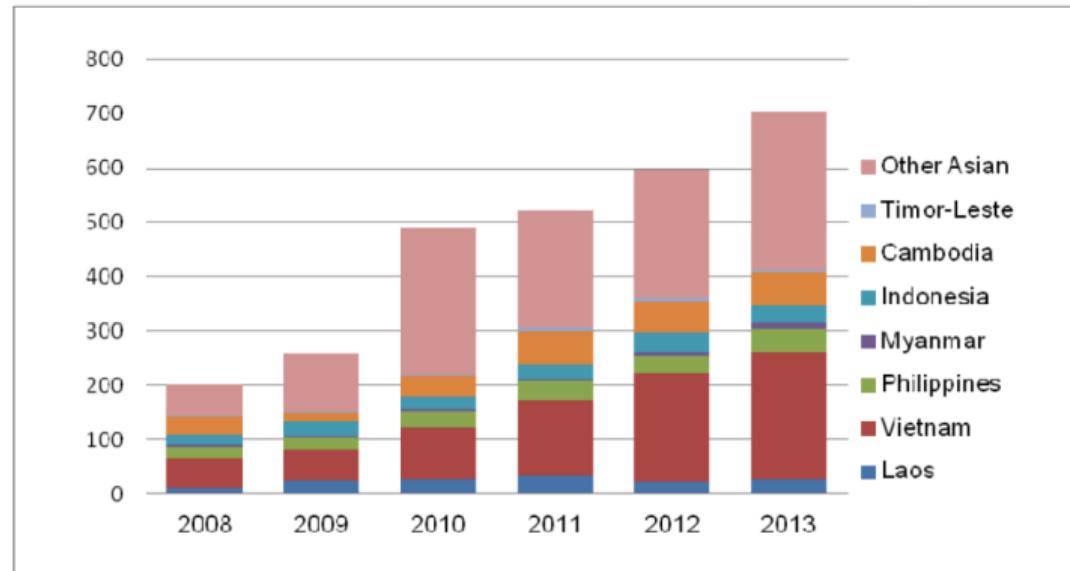
Hình 1. Tỷ lệ hệ thống đường giao thông trên địa bàn xã



Some (frequent) encounters from the web/publications

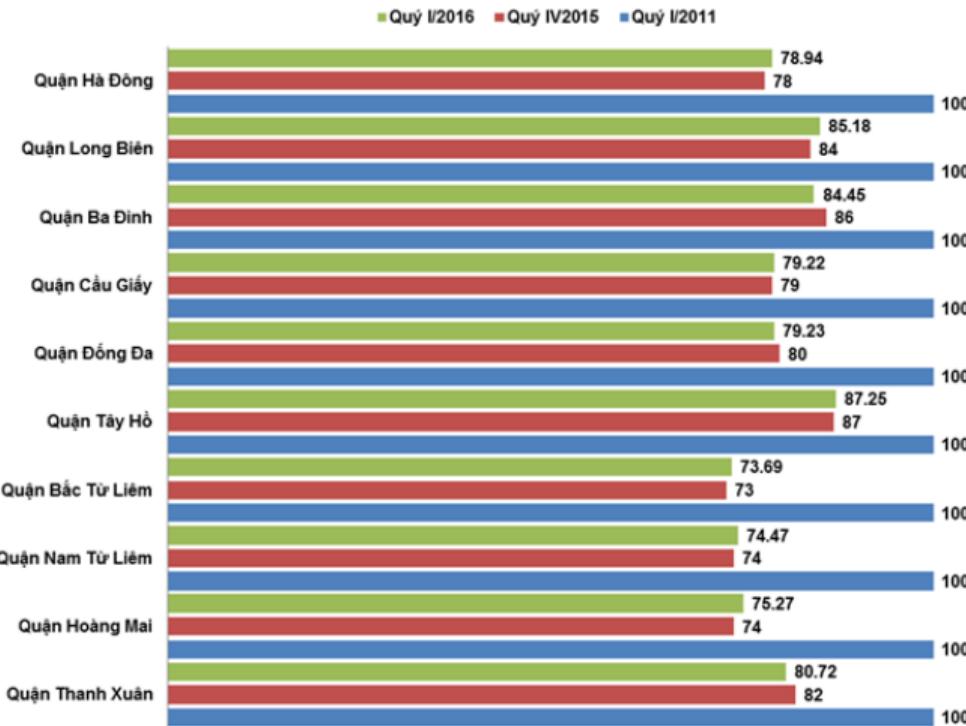
Figure 11. Korea's ODA for Asian countries

(Unit: US\$ mill.)

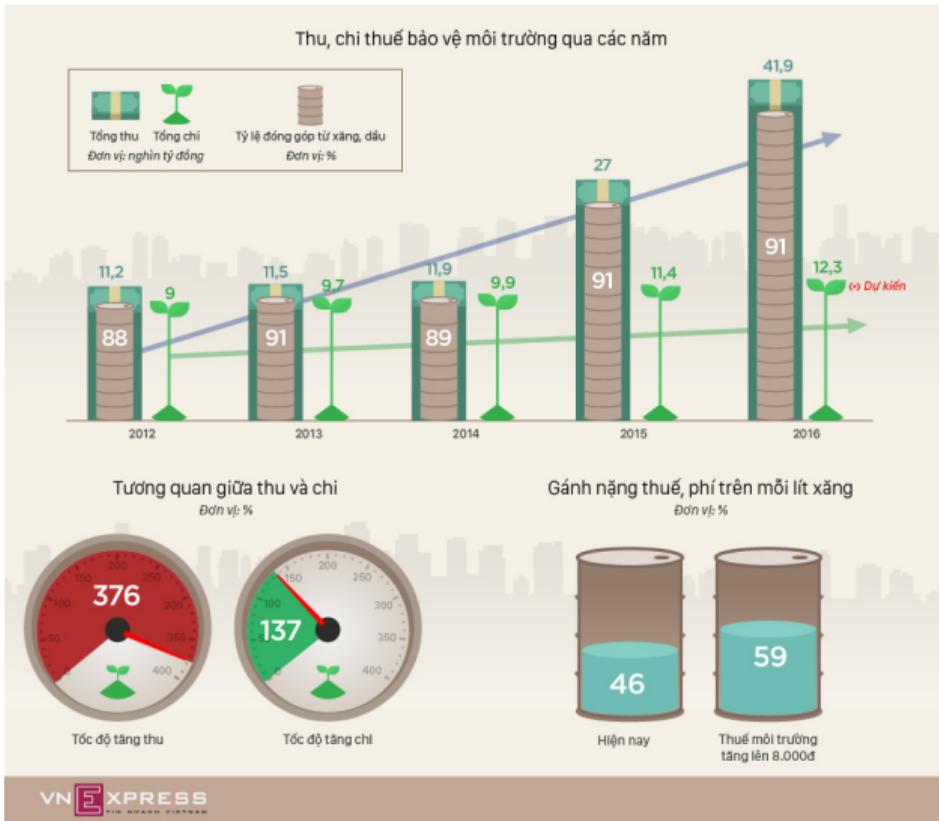


Source: http://www.odakorea.go.kr/eng/result.RegionCountry_Asia.do

Some (frequent) encounters from the web/publications



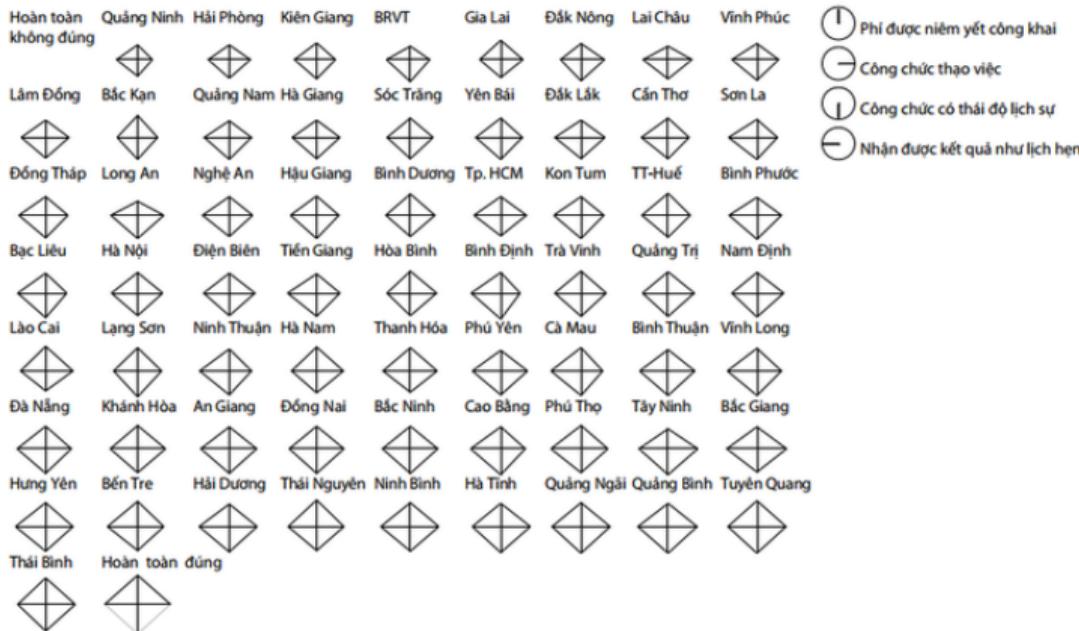
Some (frequent) encounters from the web/publications



Some (frequent) encounters from the web/publications

Biểu đồ 3.5c: Tổng chất lượng dịch vụ hành chính được cấp ở UBND xã/phường, 2016

(Các cạnh từ tâm hình thoi thể hiện tỉ lệ % người trả lời đồng ý với các nhận định trong chú giải. 'Hoàn toàn đúng' = 100% đồng ý; 'Hoàn toàn không đúng' = 0% đồng ý.)



We can do better

We can do better with R

Basic principles of visualization

Visualization components

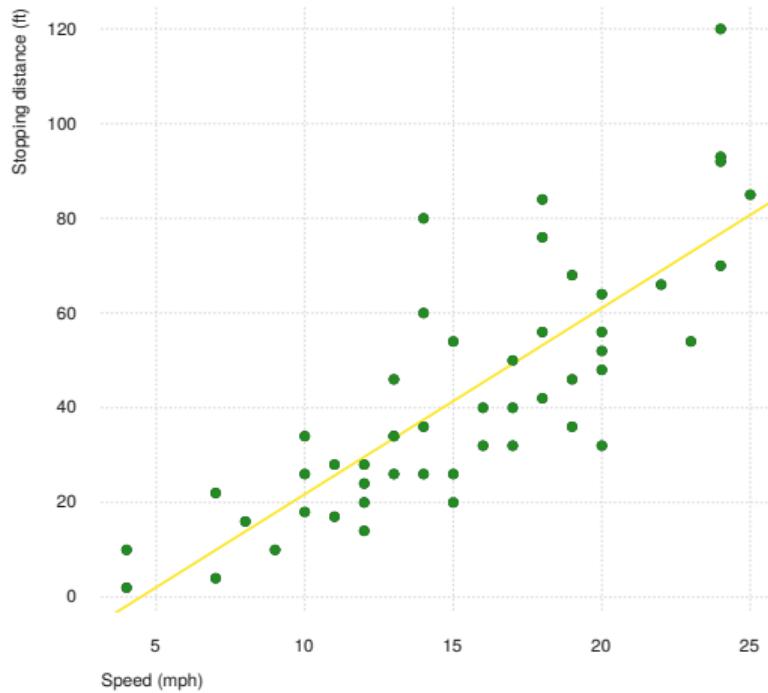
- ▶ Visual encodings
- ▶ Coordinate system
- ▶ Scale
- ▶ Context

(Data Points, Nathan Yau)

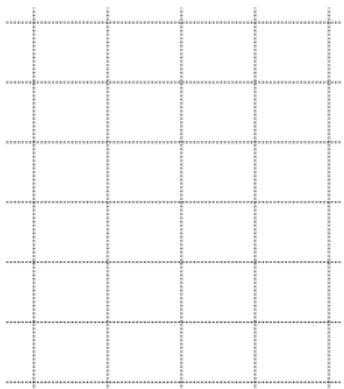
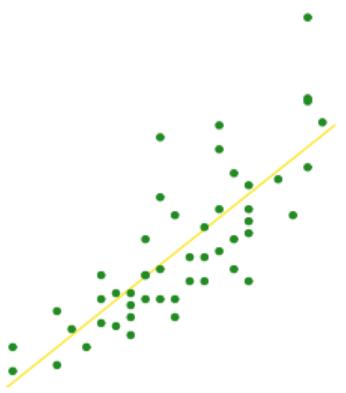
Decomposing visualization

Drivers: Keep your speed down!

Higher speed, longer distance taken to stop



Source: Ezekiel, M. (1930). Wiley.



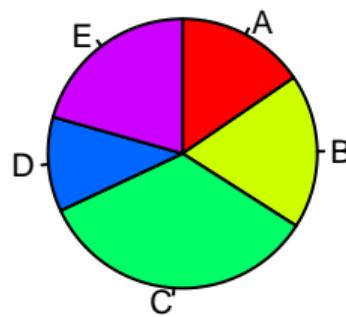
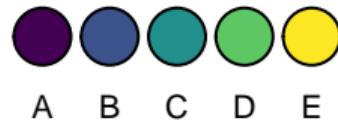
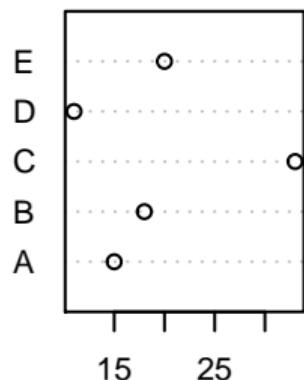
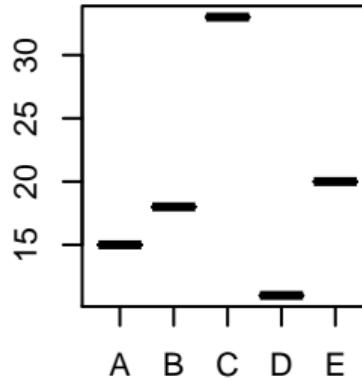
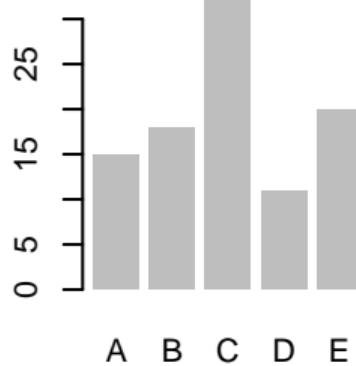
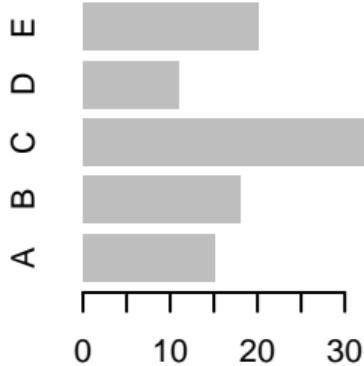
Drivers: Keep your speed down!
Higher speed, longer distance taken to stop

Source: Ezekiel, M. (1930). Wiley.

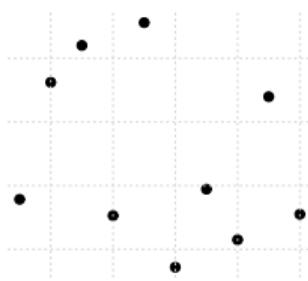
Visual encodings

- ▶ Mapping data into visual properties

Different methods of encoding the same data set



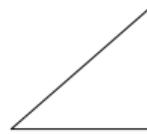
Position



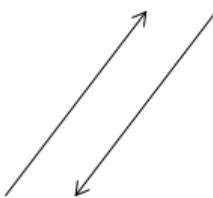
Length



Angle



Direction



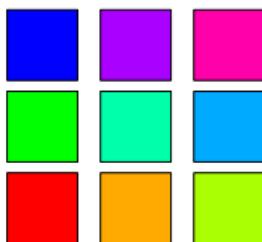
Shapes



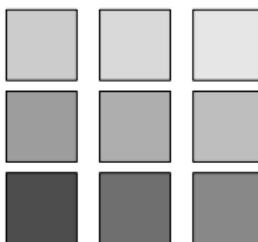
Area



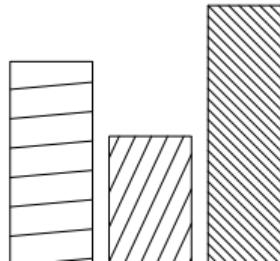
Color hue



Color saturation



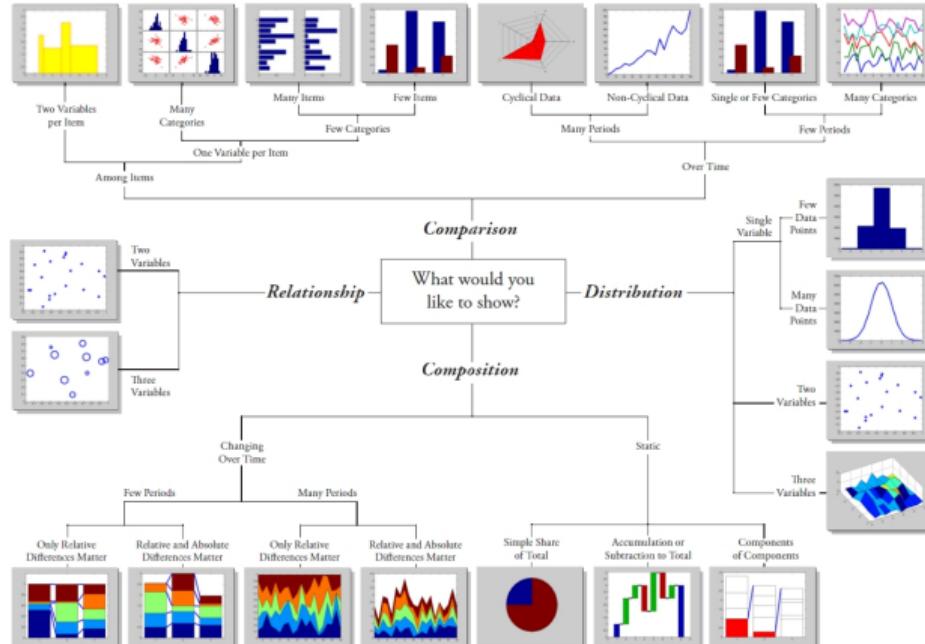
Shading



How to Choose Appropriate Visual Encodings

What is your question?

Chart Suggestions—A Thought-Starter



Graphical perception

Graphical Perception and Graphical Methods for Analyzing Scientific Data

William S. Cleveland; Robert McGill

Science, New Series, Vol. 229, No. 4716 (Aug. 30, 1985), 828-833.

Stable URL:

<http://links.jstor.org/sici?&sici=0036-8075%2819850830%293%3A229%3A4716%3C828%3AGPAGMF%3E2.0.CO%3B2-D>

Science is currently published by American Association for the Advancement of Science.



Graphical perception

Crowdsourcing Graphical Perception: Using Mechanical Turk to Assess Visualization Design

Jeffrey Heer and Michael Bostock

Computer Science Department
Stanford University
`{jheer, mbostock}@cs.stanford.edu`

ABSTRACT

Understanding perception is critical to effective visualization design. With its low cost and scalability, crowdsourcing presents an attractive option for evaluating the large design space of visualizations; however, it first requires validation. In this paper, we assess the viability of Amazon's Mechanical Turk as a platform for graphical perception experiments. We replicate previous studies of spatial encoding and luminance contrast and compare our results. We also conduct new experiments to validate the use of Mechanical Turk for graphical perception. Our results indicate that Mechanical Turk is a promising platform for graphical perception experiments.

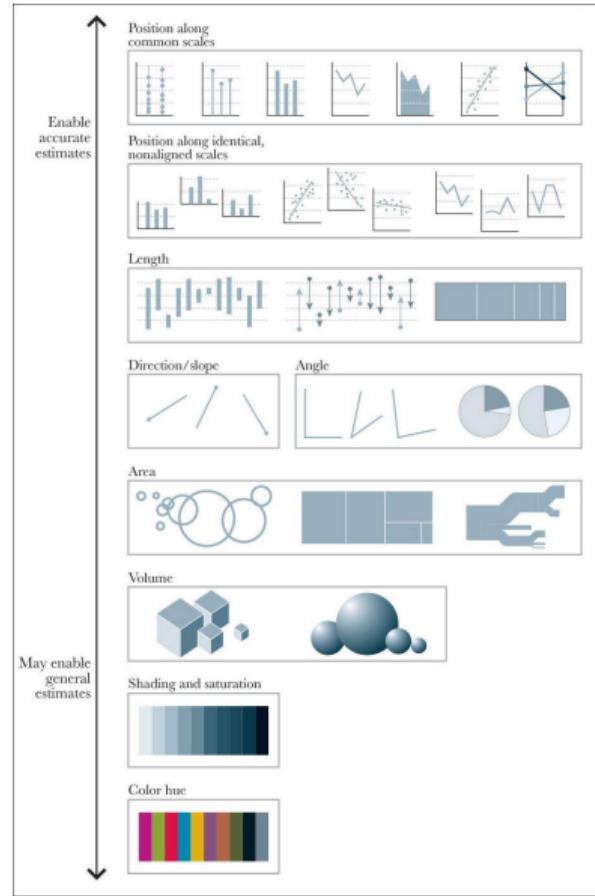
for ecological validity. Crowdsourced experiments may also substantially reduce both the cost and time to result.

Unfortunately, crowdsourcing introduces new concerns to be addressed before it is credible. Some concerns, such as ecological validity, subject motivation and expertise, apply to any study and have been previously investigated [13, 14, 23]; others, such as display configuration and viewing environment, are specific to visual perception. Crowdsourced perception experiments lack control over many experimental

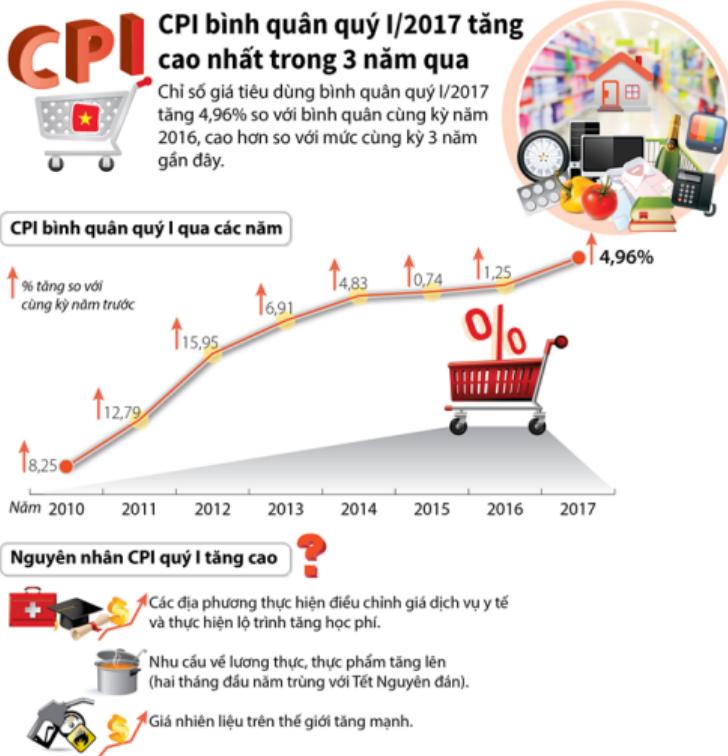
Graphical perception

1. Position along a common scale
2. Position on identical but nonaligned scales
3. Length
4. Angle. Slope
5. Direction
6. Area
7. Volume. Density. Color saturation
8. Color hue

(W. Cleveland and R. McGrill, 1985)



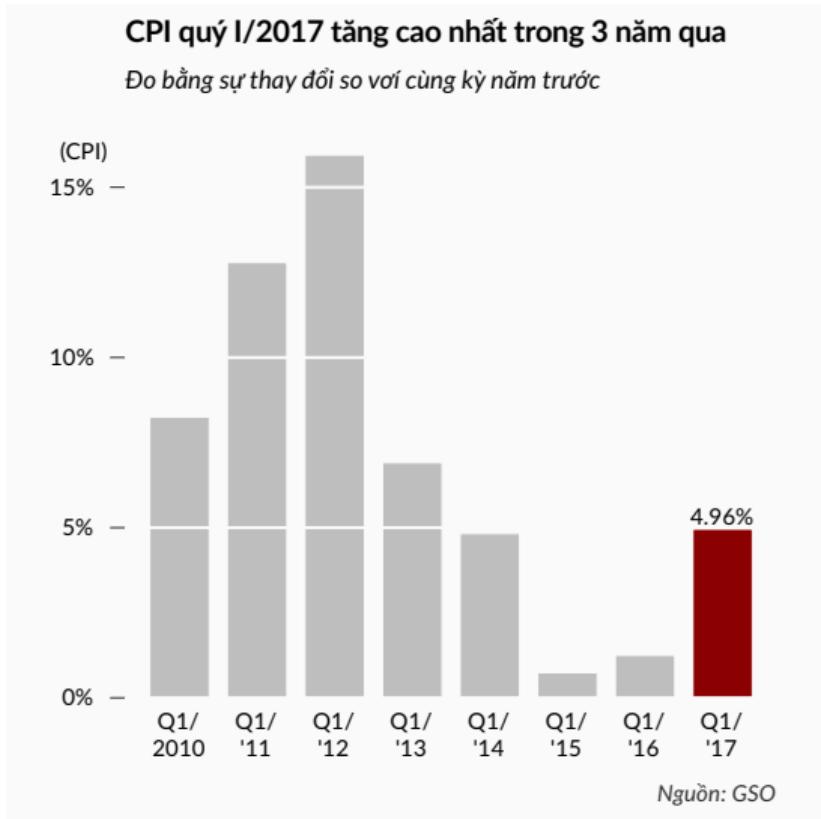
Example



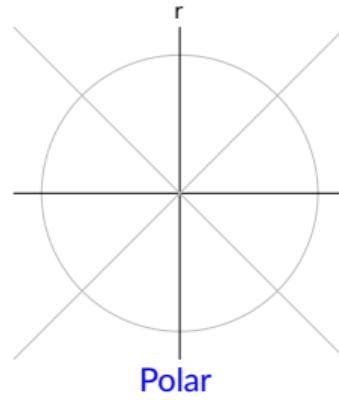
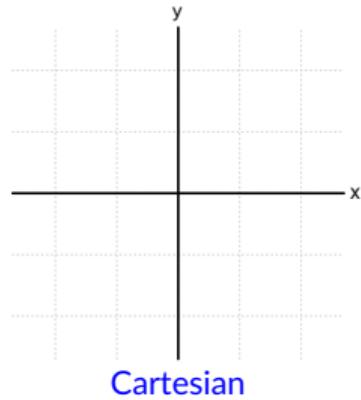
Nguồn: Tổng Cục Thống kê.

<http://infographics.vn>

Redesign



Coordinate system

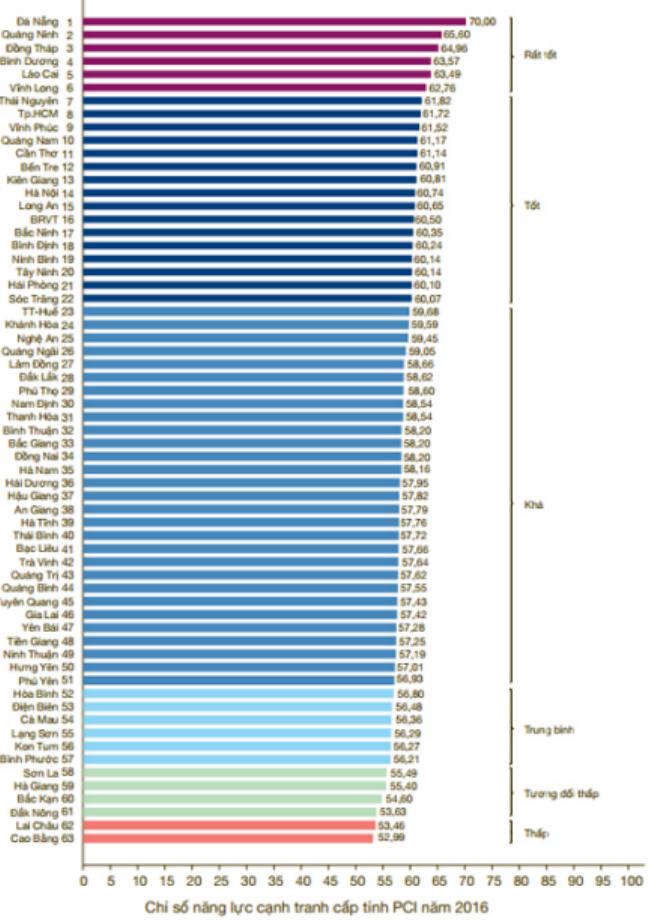


Cartesian

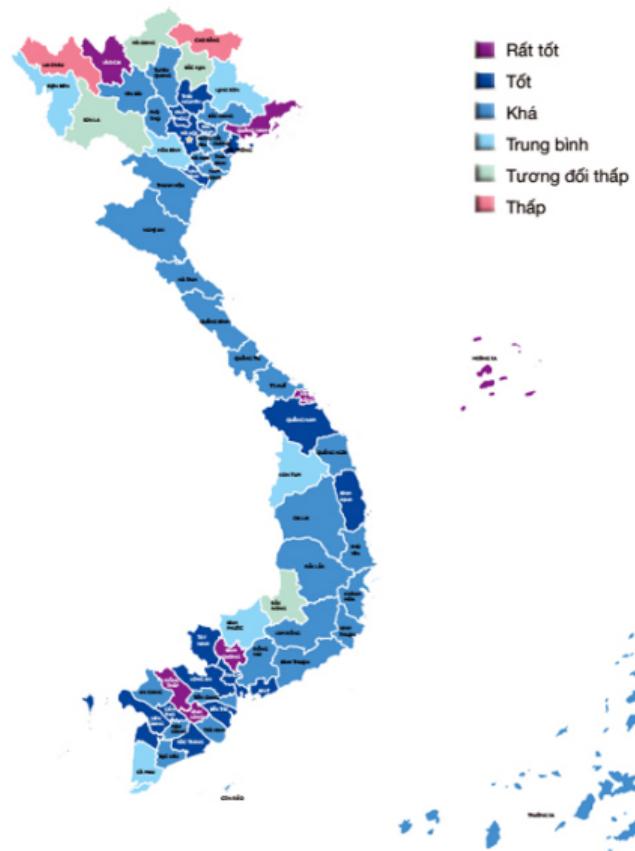
Polar



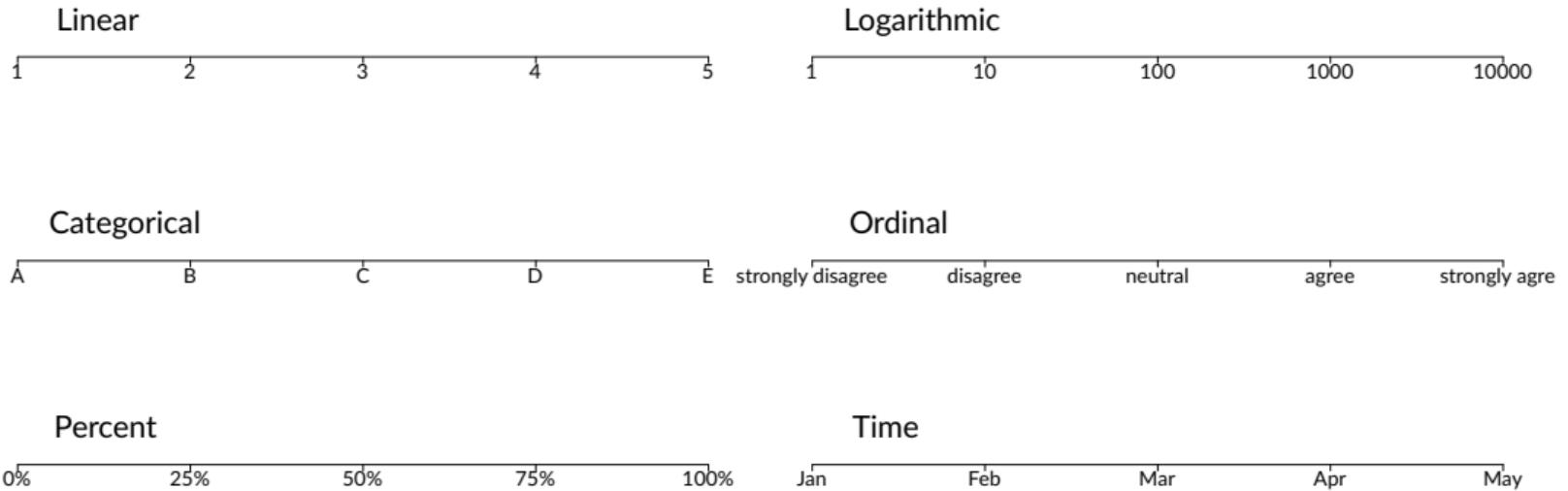
Geographic



Hình 1.4: Bản đồ PCI năm 2016



Scales

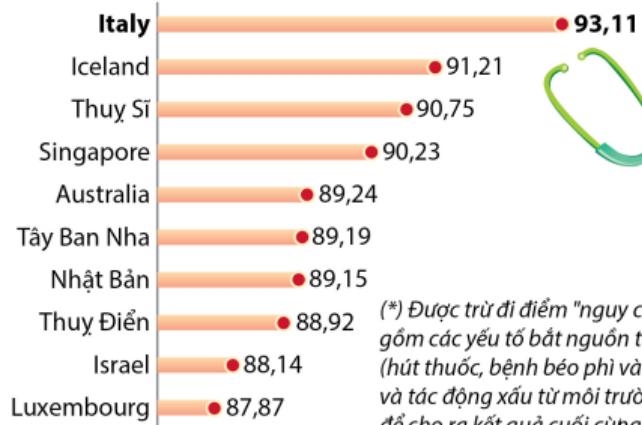


Example

Người dân Italy khỏe mạnh nhất thế giới

Theo Chỉ số Sức khỏe toàn cầu mới nhất do Bloomberg công bố, Italy là quốc gia có người dân khỏe mạnh nhất thế giới. "Điểm số sức khỏe" (*) dựa vào việc khảo sát số liệu ở 163 quốc gia, tính dựa trên tỷ lệ tử vong do bệnh tật hoặc thương tật, tuổi thọ đối với các độ tuổi khác nhau và xác suất sống sót vào những thời điểm then chốt.

Điểm số sức khỏe của 10 nước đứng đầu

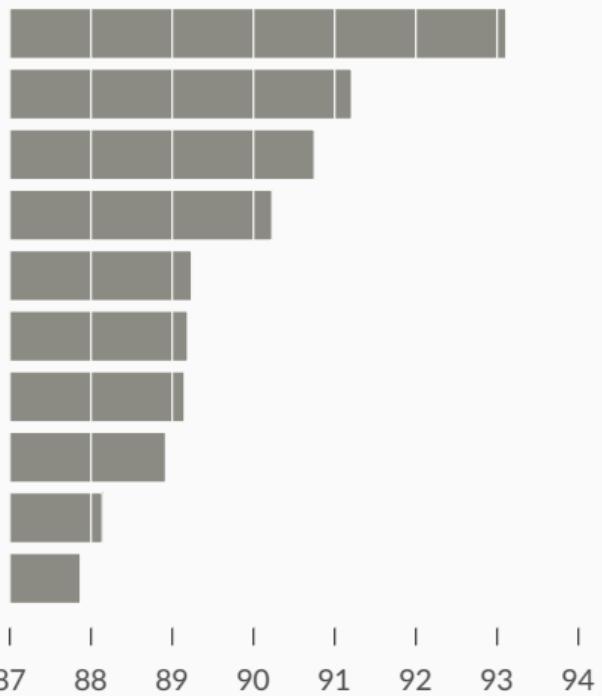


(*) Được trừ đi điểm "nguy cơ sức khỏe",
gồm các yếu tố bắt nguồn từ những hành vi gây hại
(hút thuốc, bệnh béo phì và suy dinh dưỡng ở trẻ em)
và tác động xấu từ môi trường (ô nhiễm và thiếu nước sạch)
để cho ra kết quả cuối cùng.

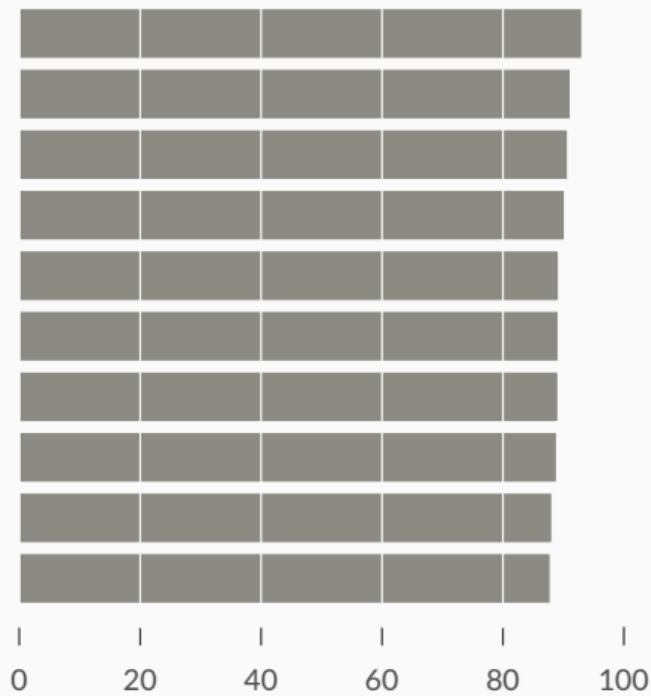
Nguồn: TTXVN, Bloomberg.

<http://infographics.vn>

Truncated axis



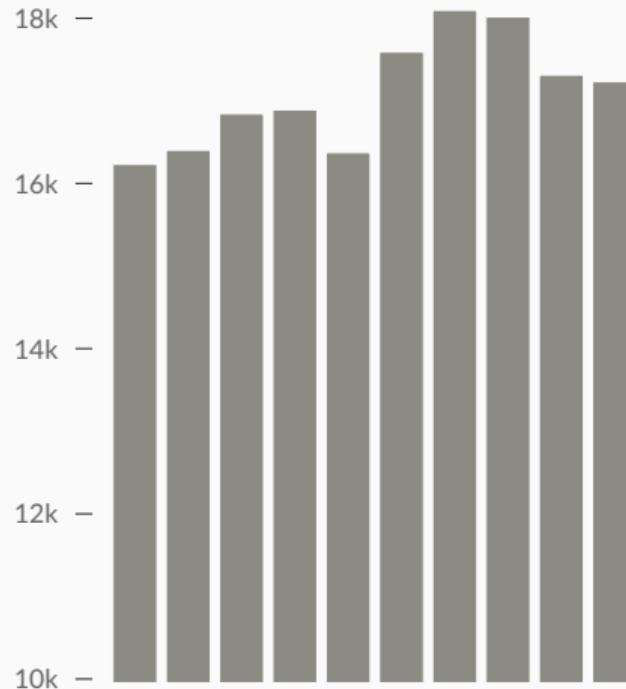
Full-scale axis



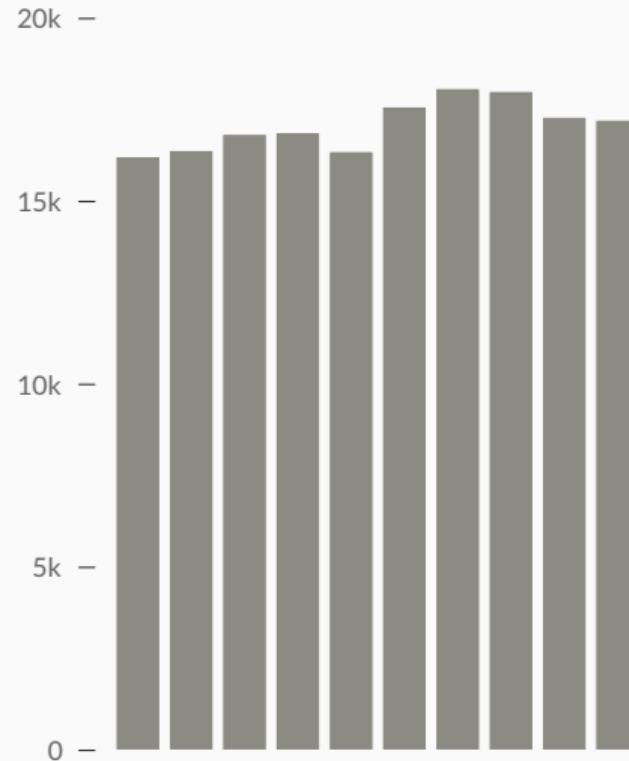
Example



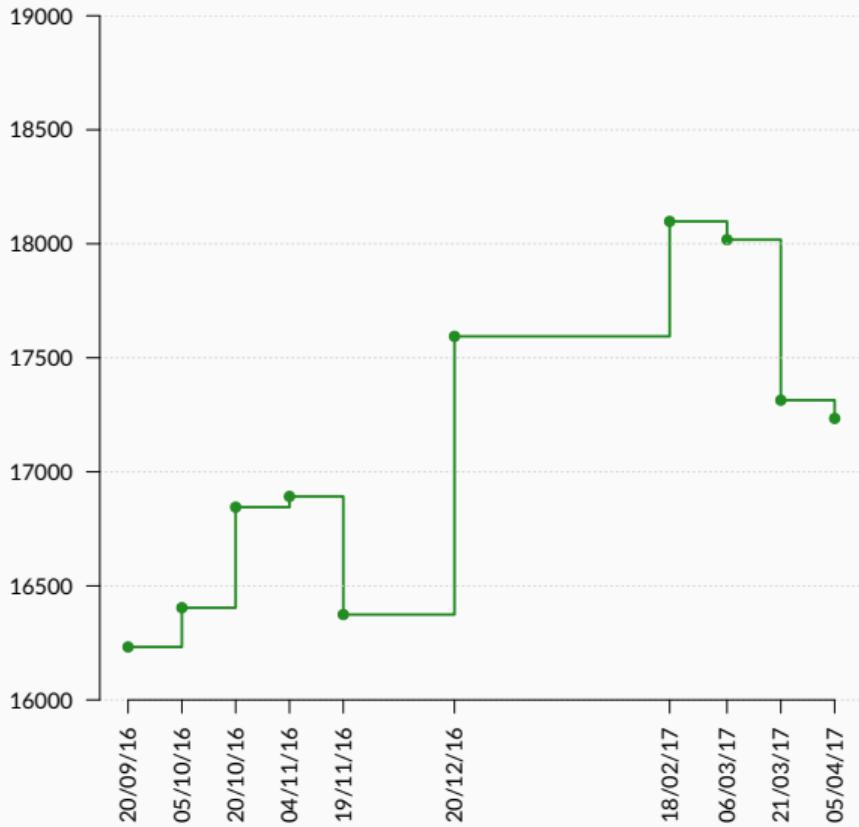
None-zero baseline



Zero baseline



Biểu đồ giá xăng (VND)



Annotation

“The annotation layer is the most important thing we do... otherwise it's a case of here it is, you go figure it out.”

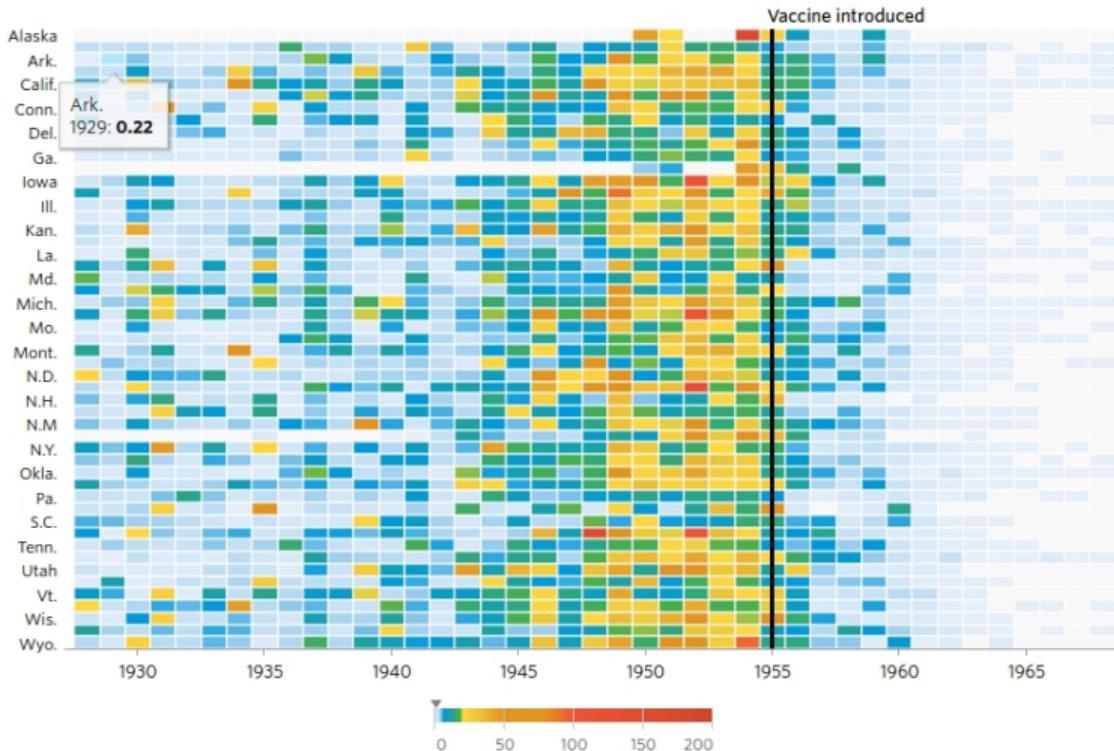
(Amanda Cox, Graphics Editor, NYTimes)

Annotation

- ▶ Title
- ▶ Subtitle
- ▶ Axix labels
- ▶ Source
- ▶ User guide

Example

Polio



Example

The New York Times

LOG IN

MAINE
VT.
N.H.
MASS.
R.I.

NOAA tide gauges

Boston

A Sharp Increase In 'Sunny Day' Flooding

By JONATHAN CORUM SEPT. 3, 2016

Global warming and rising seas are increasing the amount of tidal flooding on the Atlantic and Gulf Coasts. Flood levels are different from city to city, but the trends are similar. [Related Article](#)

Boston

The city has not been hit by tidal flooding as hard as cities farther south, but it is working on a plan to combat flooding and sea-level rise.

Mean sea level rise, in inches

Days of nuisance flooding

10" rise ...
5" ...
10 days

1950 1975 2000 2015

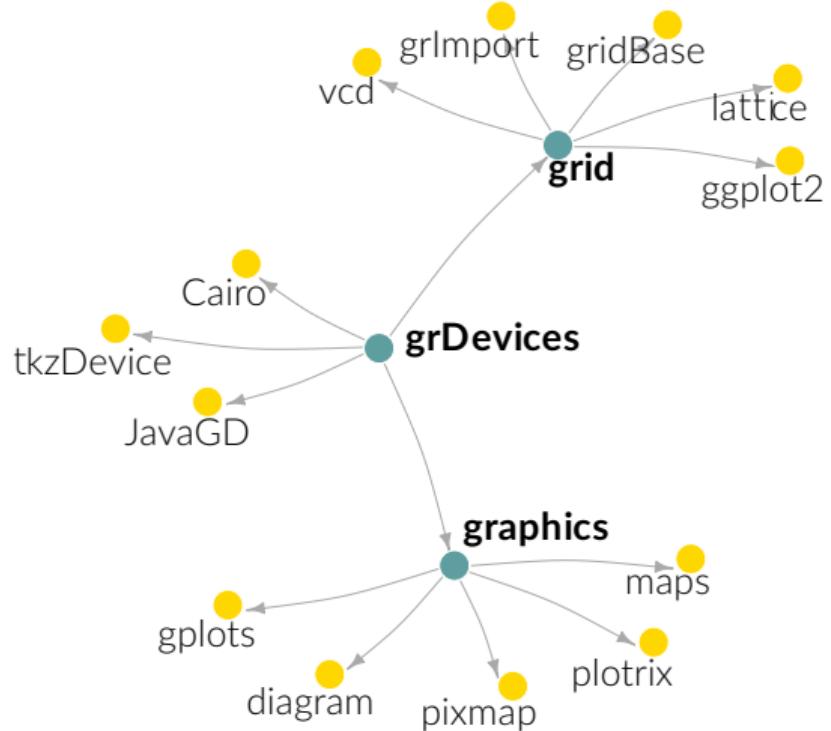
The Battery, New York City

In 2012, Hurricane Sandy laid bare the city's vulnerability to storm surges and tidal flooding. The city is spending some \$20 billion on a resilience plan.

10" ...
5" ...

The R Plotting System

R graphics system



R graphics system

- ▶ **grDevices**: graphics engine which provides facilities such as selecting colors, fonts and output formats.

Two (largely incompatible) packages built on top of the graphics engine:

- ▶ **graphics**: (aka base graphics) S's legacy. `graphics` provides both high-level and low-level functions for creating plots.
- ▶ **grid**: unique to R. `grid` offers low-level tools used for building `ggplot2` and `lattice`.

Base graphics vs ggplot2

A screenshot of a Google search results page titled "base vs ggplot2 - Google Search - Google Chrome". The search bar shows the query "base vs ggplot2". Below the search bar, there are filters for "All", "Images", "Videos", "News", and "More". There are also "Settings" and "Tools" buttons. The search results section indicates "About 143,000 results (0.33 seconds)". The first result is a link to "Comparing ggplot2 and R Base Graphics | FlowingData" from flowingdata.com. The second result is "Why I don't use ggplot2 · Simply Statistics" from simplystatistics.org. The third result is "Base versus ggplot2 | DIRKMJK" from dirkmjk.nl. The fourth result is "Why I use ggplot2 – Variance Explained" from varianceexplained.org. At the bottom left, there is a "People also ask" section with the question "What is a Ggplot?".

base vs ggplot2 - Google Search - Google Chrome

Secure | https://www.google.com.vn/search?q=base+vs+ggplot2&oq=bas&aqs=chrome.0.69|59|2j69|57j69|65l2j69|59.1796j0j1&sourceid=chrome&ie=UTF-8

base vs ggplot2

All Images Videos News More Settings Tools

About 143,000 results (0.33 seconds)

Comparing ggplot2 and R Base Graphics | FlowingData
flowingdata.com/2016/03/22/comparing-ggplot2-and-r-base-graphics/ ▾
Mar 22, 2016 - These days, people tend to either go by way of base graphics or with ggplot2. It's one or the other. Rarely both. I use base graphics. I don't use ...

Why I don't use ggplot2 · Simply Statistics
simplystatistics.org/2016/02/11/why-i-dont-use-ggplot2/ ▾
Feb 11, 2016 - When making graphs that are production ready or fit for publication, you can do ... You can do it with ggplot2, with lattice, with base R graphics.

Base versus ggplot2 | DIRKMJK
dirkmjk.nl/en/2016/02/base-versus-ggplot2 ▾
Feb 12, 2016 - Base versus ggplot2. Yesterday, stats guru Jeff Leek confessed the ultimate unpopular opinion in data science: «I don't use ggplot2 and I get ...

Why I use ggplot2 – Variance Explained
varianceexplained.org/r/why-i-use-ggplot2/ ▾
Feb 12, 2016 - In ggplot2, you add a group = stock or group = gene aesthetic. In base plotting, you write a loop, subset the data each time, and call lines - but ...

People also ask

What is a Ggplot?

Base graphics vs ggplot2

base graphics:

- ▶ Accept many types of input (vectors, data frames, matrix,...)
- ▶ Quicker to get going (lots of single-call functions)
- ▶ Better performance (speedier)
- ▶ Easier to customize (mimics 'painters model': output occurs in steps)
- ▶ Awkward workflow (sometimes)

ggplot2:

- ▶ Dataframe-centered
- ▶ Steeper learning curve (to master conceptual framework)
- ▶ Better default (generally)
- ▶ Highly extensible, more efficient in the long run (thanks to paradigm-based design)
- ▶ Seamlessly integration with tidyverse

Why I choose base graphics

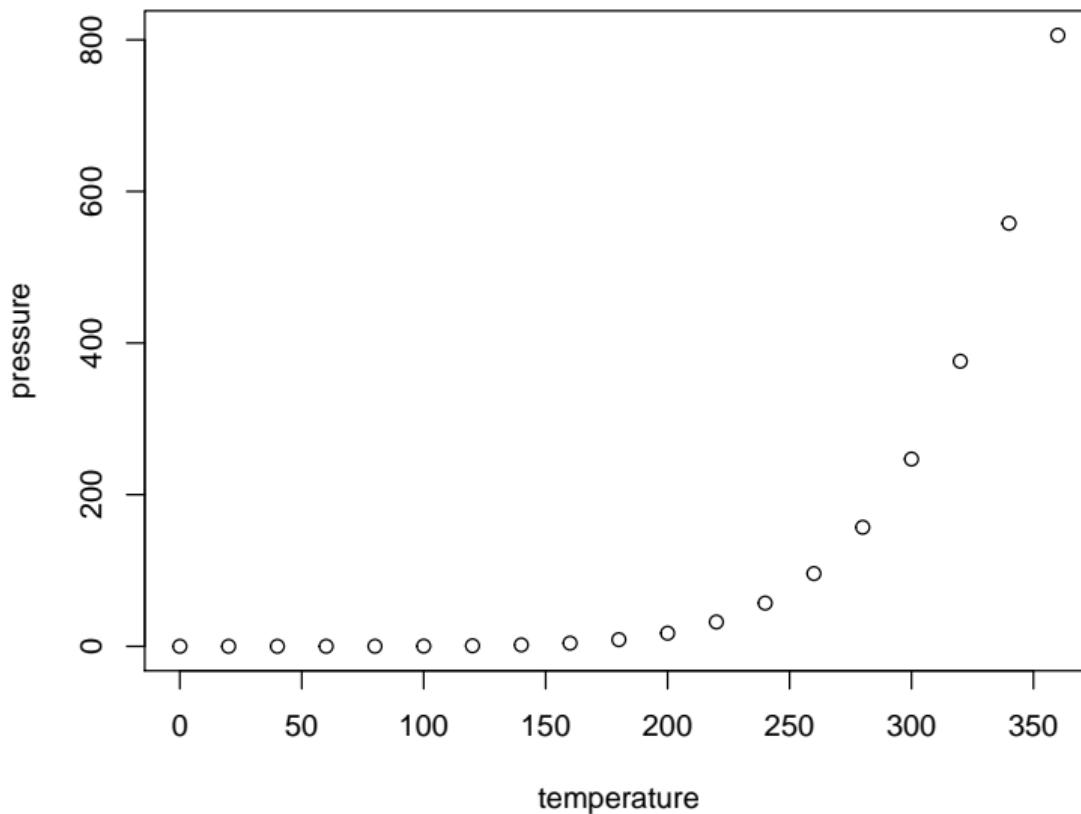
- ▶ Both packages are super!
- ▶ But, base graphics might be better for beginners to make graphs quickly with as little as mental efforts.

Graphical Functions

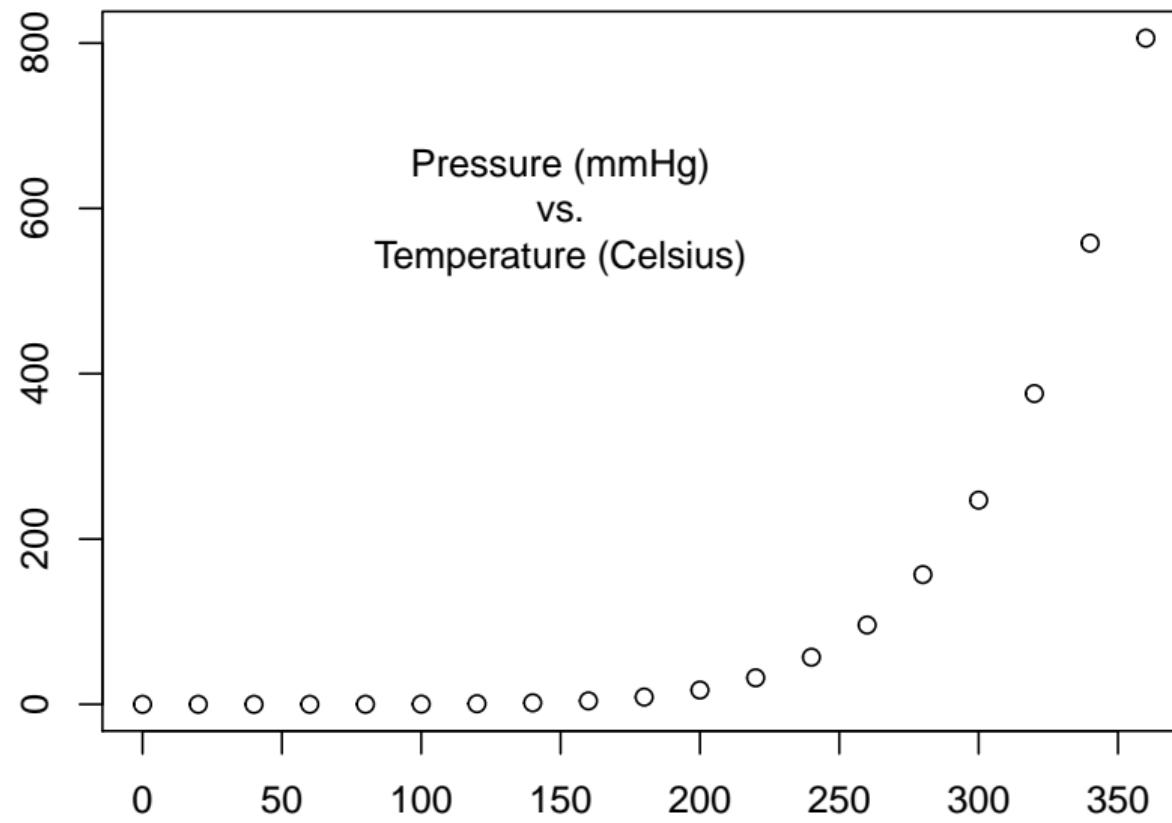
The base graphics model

- ▶ First, call a high-level functions to make a complete plot.
- ▶ Then, call low-level functions to add more output (if necessary).

```
plot(x = pressure$temperature, y = pressure$pressure,  
      xlab = "temperature", ylab = "pressure")
```



```
plot(x = pressure$temperature, y = pressure$pressure, ann = F)
text(150, 600, "Pressure (mmHg)\nvs.\nTemperature (Celsius)")
```



High-level plotting functions

- ▶ `plot()` : scatterplot (lines plot)
- ▶ `barplot()` : bar chart
- ▶ `pie()` : pie chart
- ▶ `boxplot` : boxplot
- ▶ `hist()` : histogram
- ▶ `stripchart()` : 1-d scatterplot

Visualizing distributions

Demo `visualize_distributions.R`

Revealing changes

Demo reveal_changes.R:

Showing relationships

Demo show_relationships.R

Making comparisions

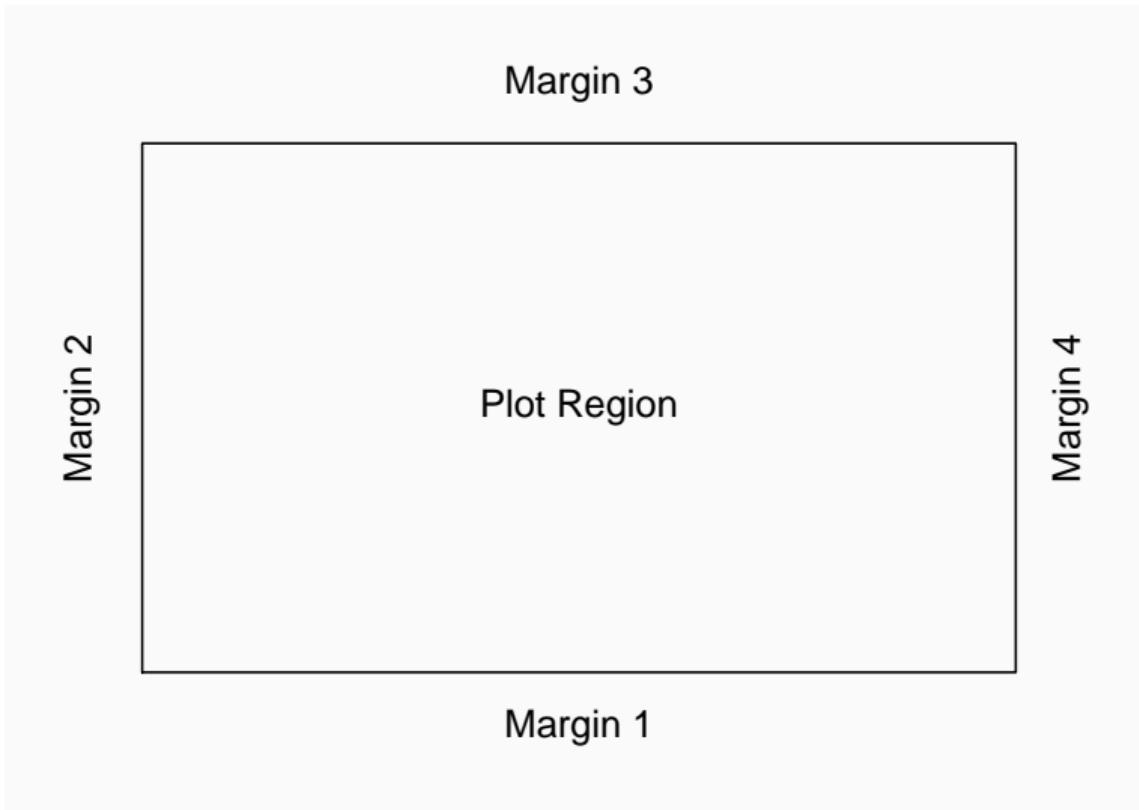
Demo make_comparisions.R

Some arguments are accepted by many high-level functions:

- ▶ `xlab`: x-axis label
- ▶ `ylab`: y-axis label
- ▶ `main`: plot title
- ▶ `sub`: plot subtitle
- ▶ `col`: color
- ▶ `lty`: line type
- ▶ `lwd`: line width
- ▶ `xlim`: x-axis scale limits
- ▶ `ylim`: y-axis scale limits

Going beyond the default

The layout of graphics



Control the margins

`par()` is used to control low-level graphics by setting *graphical parameters*

- ▶ set the margin sizes in inches

```
par(mai = c(2, 2, 1, 1))
```

- ▶ set the margin sizes in lines of text

```
par(mar = c(4, 4, 2, 2))
```

line 2
line 1

line 4
line 3
line 2
line 1

line 1
line 2

line 1
line 2
line 3
line 4

Multifigure layouts

par can be used to set up multiple figures on the page.

- ▶ row-by-row:

```
par(mfrow = c(2, 2))
```

- ▶ column-by-column:

```
par(mfcol = c(2, 2))
```

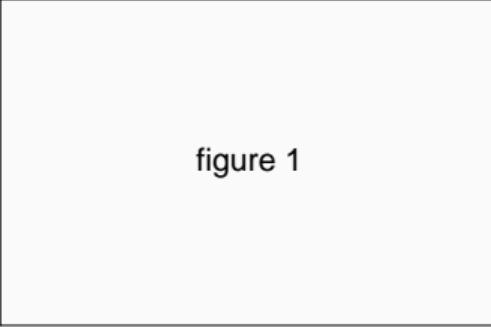


figure 1

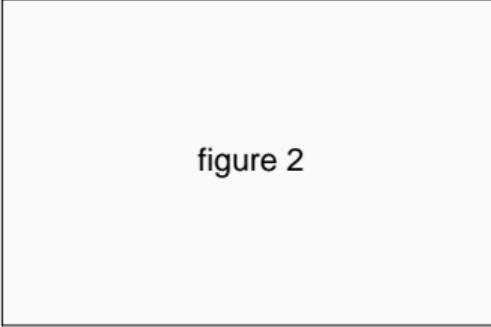


figure 2

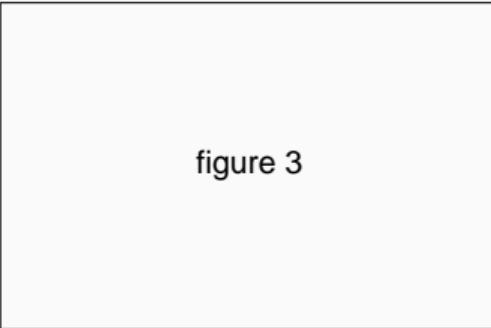


figure 3

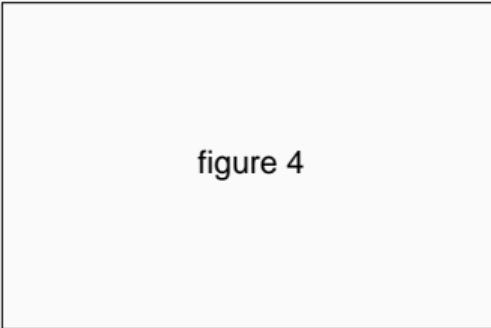


figure 4

Draw axes

`axis()` can be used to draw axes at any of the four side of a plot.

This is a default call:

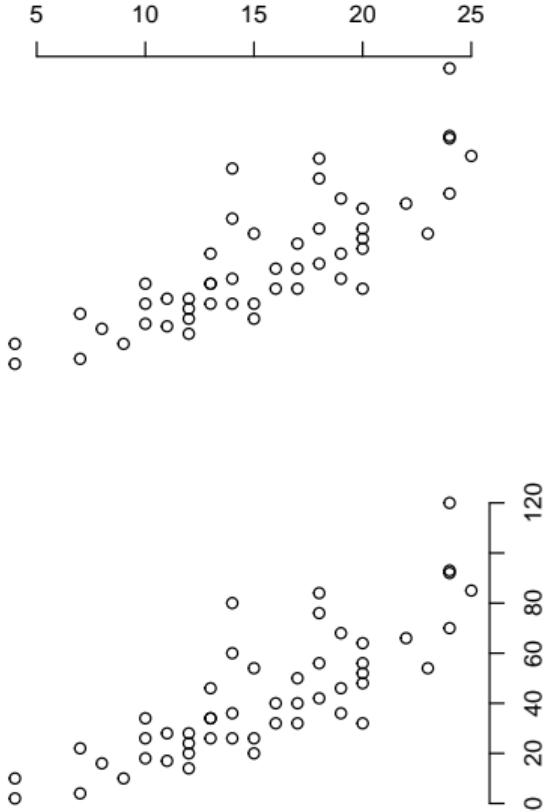
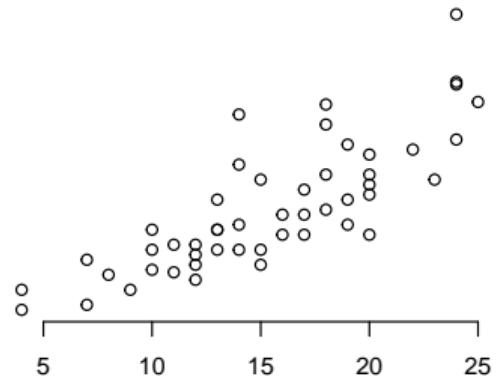
```
axis(side)
```

Possible values of `side` include:

- ▶ `side = 1`: bottom
- ▶ `side = 2`: left
- ▶ `side = 3`: top
- ▶ `side = 4`: right

axis() example

```
par(mfrow = c(2, 2))
plot(cars, ann = F, axes = F)
axis(1)
plot(cars, ann = F, axes = F)
axis(2)
plot(cars, ann = F, axes = F)
axis(3)
plot(cars, ann = F, axes = F)
axis(4)
```



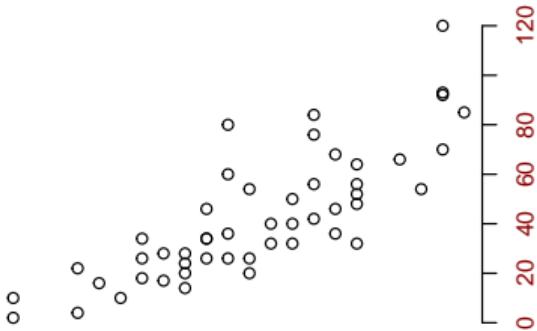
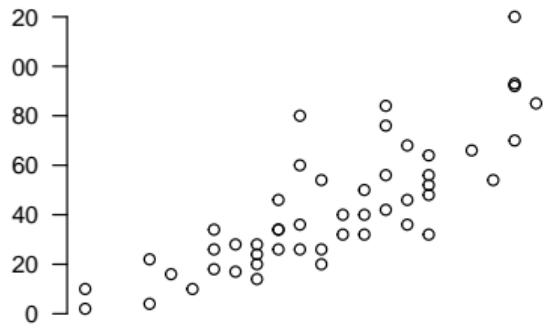
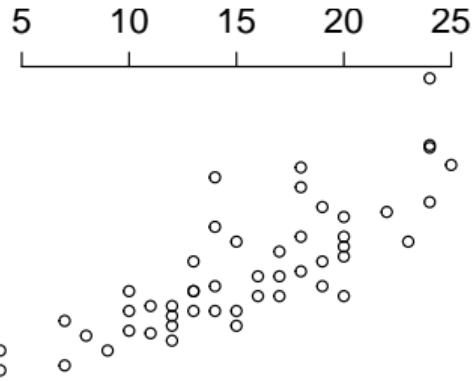
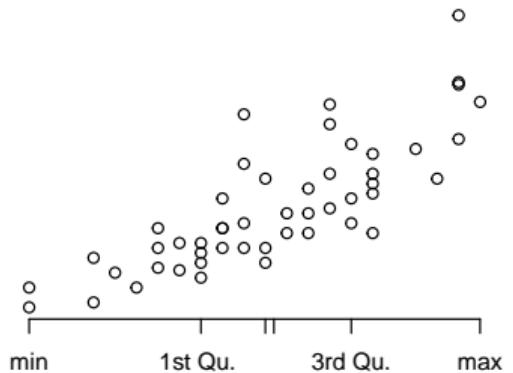
Axis customization

```
axis(at = , labels = , las = , line = , cex.axis = , col.axis = )
```

- ▶ at: tick marks position
- ▶ labels: texts are to be made the tick marks
- ▶ las: labels rotation (0, 1, 2, 3)
- ▶ line: the number of lines into the margin at which the axis line will be drawn
- ▶ cex.axis and col.axis: size and color of axis labels

axis() example (cont)

```
par(mfrow = c(2, 2), mar = c(2, 2, 2, 2))
plot(cars, ann = F, axes = F)
axis(1, at = c(4, 12, 15, 15.4, 19, 25),
     labels = c("min", "1st Qu.", "median", "mean", "3rd Qu.", "ma
plot(cars, ann = F, axes = F)
axis(3, cex.axis = 2)
plot(cars, ann = F, axes = F)
axis(2, las = 1)
plot(cars, ann = F, axes = F)
axis(4, col.axis = "gray40")
```



Plot annotation

`mtext()` function can be used to place labels in the margins of a plot.

```
mtext(text = , side = )
```

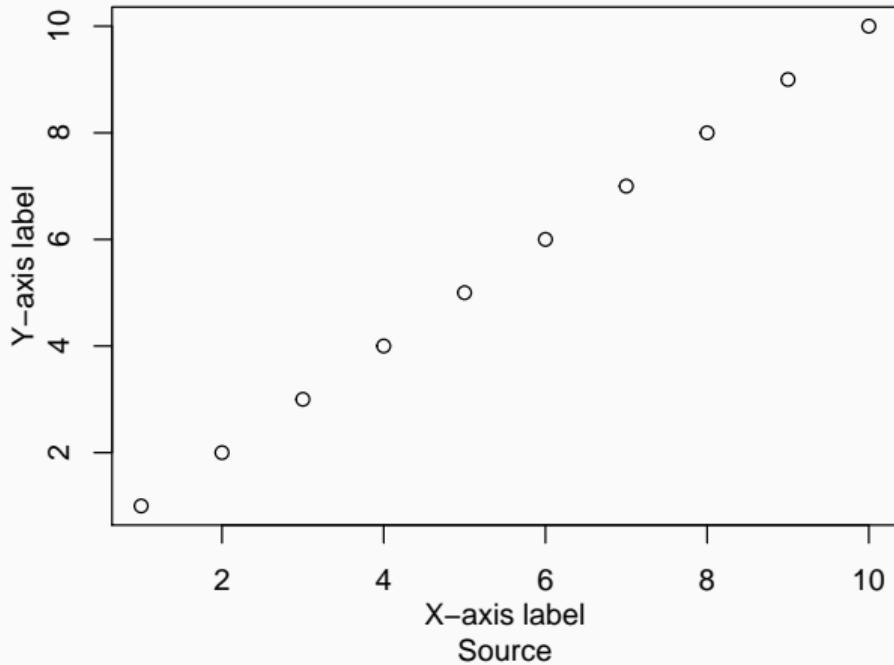
- ▶ `side`: on which side of the plot (1=b, 2=l, 3=t, 4=r)

mtext() example

```
par(mar = c(4, 4, 4, 4))
plot(1:10, 1:10, ann = F)
mtext("X-axis label", side = 1, line = 2)
mtext("Y-axis label", side = 2, line = 2)
mtext("Title", side = 3, line = 3)
mtext("Subtitle", side = 3, line = 1.5)
mtext("Source", side = 1, line = 3)
```

Title

Subtitle



Customizing plot annotation

- ▶ `line`: the number of lines into the margin at which the axis line will be drawn
- ▶ `adj`: adjustment direction [0, 1]
- ▶ `outer`: use outer margins if available (TRUE or FALSE)
- ▶ `cex`: size (expansion factor)
- ▶ `col`: color
- ▶ `font`: 1=regular, 2=bold, 3=italic, 4=bold-italic

Recap

- ▶ Setting margins by `par(mar)`
- ▶ Draw axes by `axis()`
- ▶ Add annotation by `mtext()`

Special topic: Color

Choosing color is hard

Choosing color is hard

R has:

- ▶ 657 named colors (run: `colors()`)

Choosing color is hard

R has:

- ▶ 657 named colors (run: `colors()`)
- ▶ 7 default color sets (`rainbow()`, `heat.colors()`, `terrain.colors()`,
`topo.colors()`, `cm.colors()`, `gray.colors()`)

Choosing color is hard

R has:

- ▶ 657 named colors (run: `colors()`)
- ▶ 7 default color sets (`rainbow()`, `heat.colors()`, `terrain.colors()`,
`topo.colors()`, `cm.colors()`, `gray.colors()`)
- ▶ and a bunch of color packages (`viridis`, `RColorBrewer`,
`colorspace`,...)

When choosing color:

Above all, do no harm.

(Edward R. Tufte)

Think carefully what you are using color for

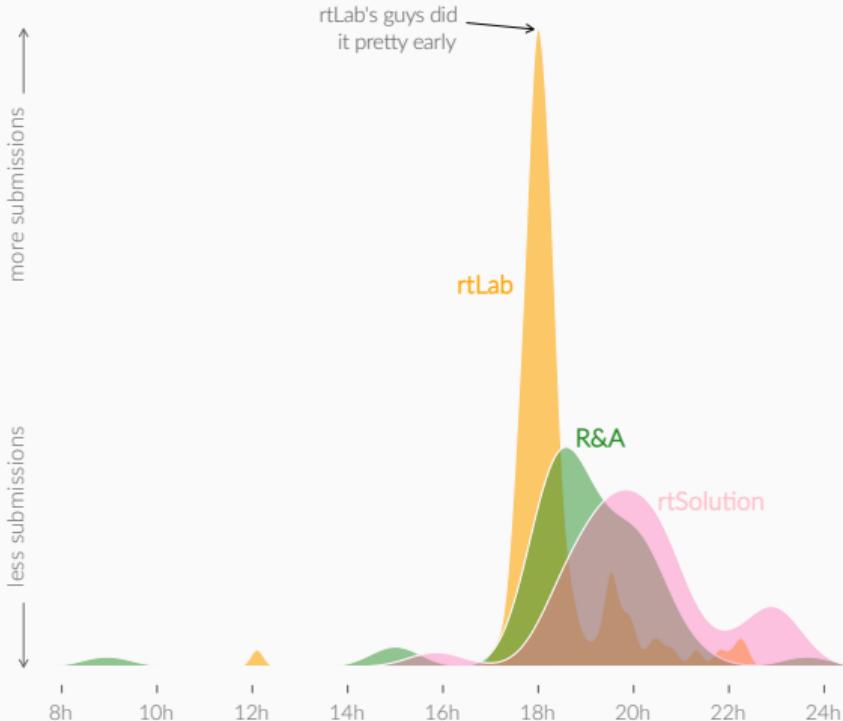
Fundamental use of color in visualization:

- ▶ to label (color as noun)
- ▶ to measure (color as quantity)
- ▶ to represent and imitate reality (color as representation)
- ▶ to decorate (color as beauty)

(Edward R. Tufte)

When do you submit your daily report?

A survey of RTA's employees



Source: Form RTA_Daily_Report

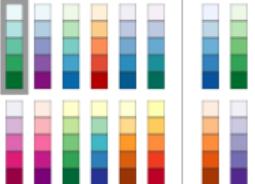
Consult experts:

Number of data classes: 3 [i](#)

Nature of your data: [i](#)

sequential diverging qualitative

Pick a color scheme:

Multi-hue: 

Single hue: 

Only show:

colorblind safe [i](#)
 print friendly [i](#)
 photocopy safe [i](#)

Context:

roads
 cities
 borders [i](#)

Background:

solid color [i](#)
 terrain [i](#)

color transparency

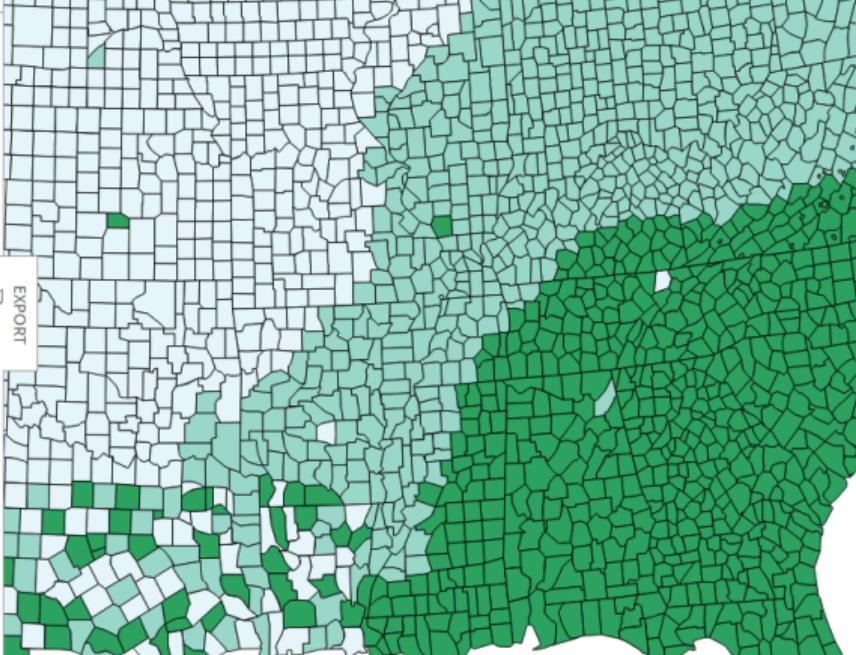
3-class BuGn [i](#)

HEX [i](#)

#e5f5f9
#99d8c9
#2ca25f

EXPORT



© Cynthia Brewer, Mark Harrower and The Pennsylvania State University
[Source code and feedback](#)
[Back to Flash version](#)
[Back to ColorBrewer 1.0](#)



Consult experts:

I want hue Tutorials Examples Theory Experiment Old version GitHub Issues + Médialab Tools

i want hue

Colors for data scientists. Generate and refine palettes of optimally distinct colors.

Color space

Default preset

H 0 C 30 L 35

360 80 80

Improve for the colorblind (slow)
 Dark background



Palette

5 colors soft (k-Means)

Reroll palette

Sort by diff hue chroma lightness random

Colors

	#7aa444 122,164,68
HEX	RGB

	#946fc7 148,111,199
HEX	RGB

	#c57b3d 197,123,61
HEX	RGB

JSON

HEX json

```
[#"7aa444",  
 "#946fc7",  
 "#c57b3d",  
 "#4cab98",  
 "#ca5477"]
```

RGB json

```
[{[122,164,68],  
 [148,111,199],  
 [197,123,61],  
 [76,171,152],  
 [202,84,119]}]
```

HCL json

```
[{[114,53.689,62.596],  
 [309,32.084,53.802],  
 [63,80.637,38.392],  
 [114,53.689,62.596],  
 [309,32.084,53.802]}]
```

CSS

HEX list for CSS

```
#7aa444  
#946fc7  
#c57b3d  
#4cab98  
#ca5477
```

RGB list for CSS

```
rgb(122,164,68)  
rgb(148,111,199)  
rgb(197,123,61)  
rgb(76,171,152)  
rgb(202,84,119)
```

Going beyond the default

Demo `customize_traditional_graphics.R`

Adding details with graphical primitives

Some functions for drawing basic graphical primitives

- ▶ `points()`: draw data symbols at (x, y)
- ▶ `lines()`: draw lines between locations (x, y)
- ▶ `abline()`: draw straight lines
- ▶ `segments()`: draw line segments between (x0, y0) and (x1, y1)
- ▶ `arrows()`: draw line segments with arrowheads
- ▶ `rect()`: draw rectangles
- ▶ `polygon()`: draw one or more polygons
- ▶ `text()`: draw text at locations (x, y)
- ▶ `legends()`: draw legends

Drawing points

Basic call:

```
points(x, y, pch = , col = )
```

- ▶ pch: plotting symbols
- ▶ col: color of point

Plotting symbols

O
pch = 1

△
pch = 2

+

pch = 3

X
pch = 4

◇
pch = 5

▽
pch = 6

⊗
pch = 7

*
pch = 8

◇
pch = 9

⊕
pch = 10

⊗
pch = 11

田
pch = 12

⊗
pch = 13

□
pch = 14

■
pch = 15

●
pch = 16

▲
pch = 17

◆
pch = 18

●
pch = 19

●
pch = 20

O
pch = 21

Drawing connected line

Basic call:

```
lines(x, y, lty = , lwd = , col =, type = )
```

- ▶ lty: line texture (“blank”, “solid”, “dashed”, “dotted”, “dotdash”, “longdash”, “twodash”)
- ▶ lwd: line width
- ▶ col: color of lines
- ▶ type: type of lines

Line texture

`lty = 6`



`lty = 5`



`lty = 4`



`lty = 3`



`lty = 2`



`lty = 1`

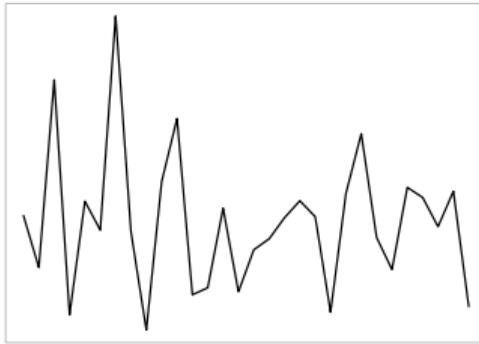


Line graph variations

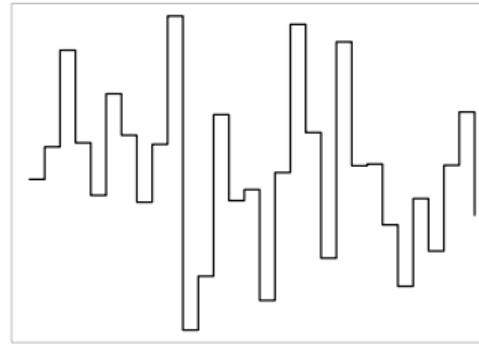
Other forms can be made by `lines()`:

- ▶ `type="l"`: line graph (default)
- ▶ `type="s"`: step - horizontal first
- ▶ `type="S"`: step - vertical first
- ▶ `type="h"`: high density plot
- ▶ `type="b"`: both points and lines
- ▶ `type="o"`: over-plotting of points and lines

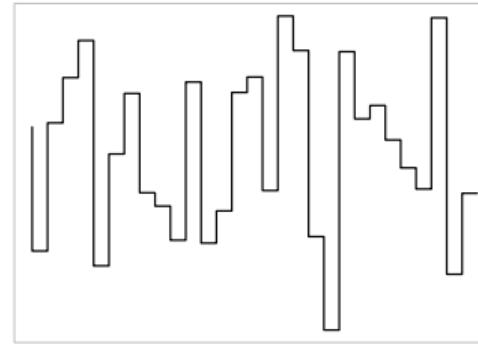
type = l



type = s



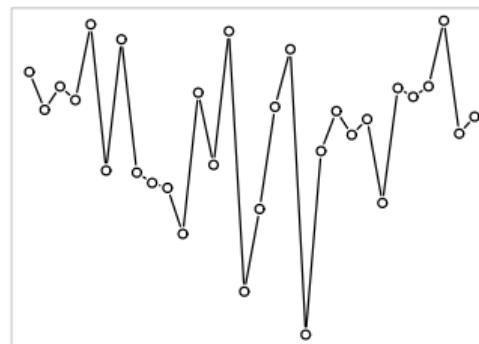
type = S



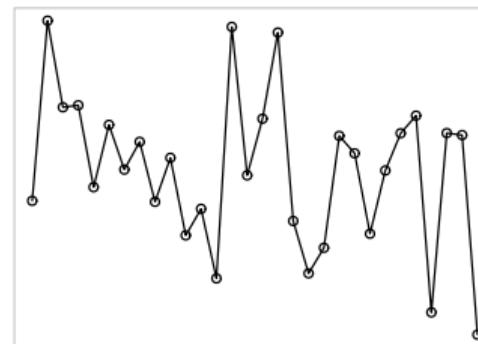
type = h



type = b



type = o



Drawing straight lines

Basic calls:

```
abline(a = , b = )  
abline(h = )  
abline(v = )
```

- ▶ a and b: specifies a line intercept and slope
- ▶ h: horizontal lines
- ▶ v: vertical lines
- ▶ Other arguments: lty, col, lwd

Drawing line segments

Basic call:

```
segments(x0, y0, x1, y1)
```

- ▶ (x_0, y_0, x_1, y_1) gives the locations of the star and end point of the segments.
- ▶ Other arguments: lty, col, lwd

Drawing arrows

Basic calls:

```
arrows(x0, y0, x1, y1, code = , length = , angle = )
```

- ▶ (x_0, y_0, x_1, y_1) gives the locations of the start and end point of the arrows.
- ▶ head=1: head at the start; head=2: head at the end and head=3: head at both ends.
- ▶ length: length of arrow head.
- ▶ angle: angle to the shaft.

Drawing rectangles

Basic call:

```
rect(x0, y0, x1, y1, col = , border = )
```

- ▶ (x_0, y_0, x_1, y_1) gives the locations of opposite corners of the rectangles.
- ▶ `col` and `border` give color of the interior and border.
- ▶ Others: `lty` and `lwd`

Drawing polygon

Basic call:

```
polygon(x, y, col = , border = )
```

- ▶ x, y gives the coordinates of the polygon vertexes.
- ▶ col and border give color of the interior and border.
- ▶ Others: lty and lwd.

Drawing text

```
text(x, y, labels)
```

- ▶ x, y: locations of the text
- ▶ labels: actual strings
- ▶ Others: font, col, adj

Drawing legend

Basic form:

```
legend(x, y, legend = ,...)
```

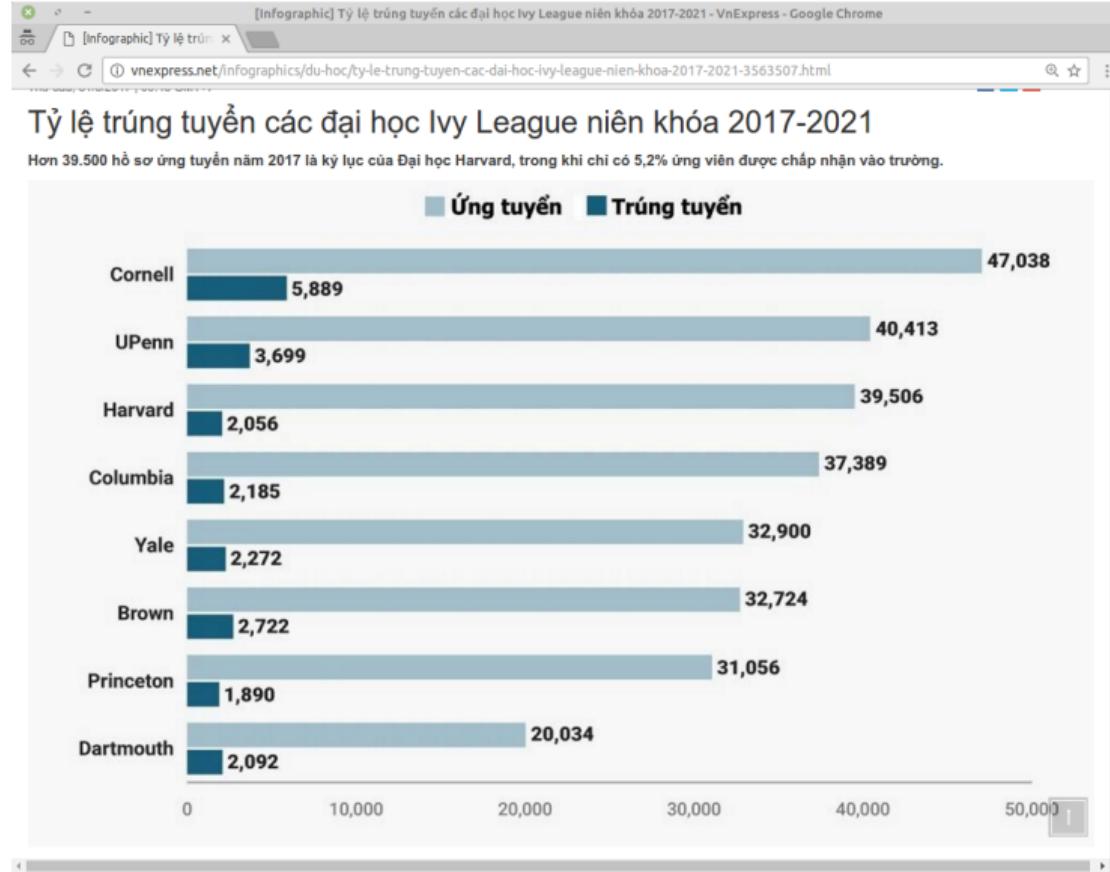
Adding details

Demo add_details.R

Practice

“Critique by redesign”

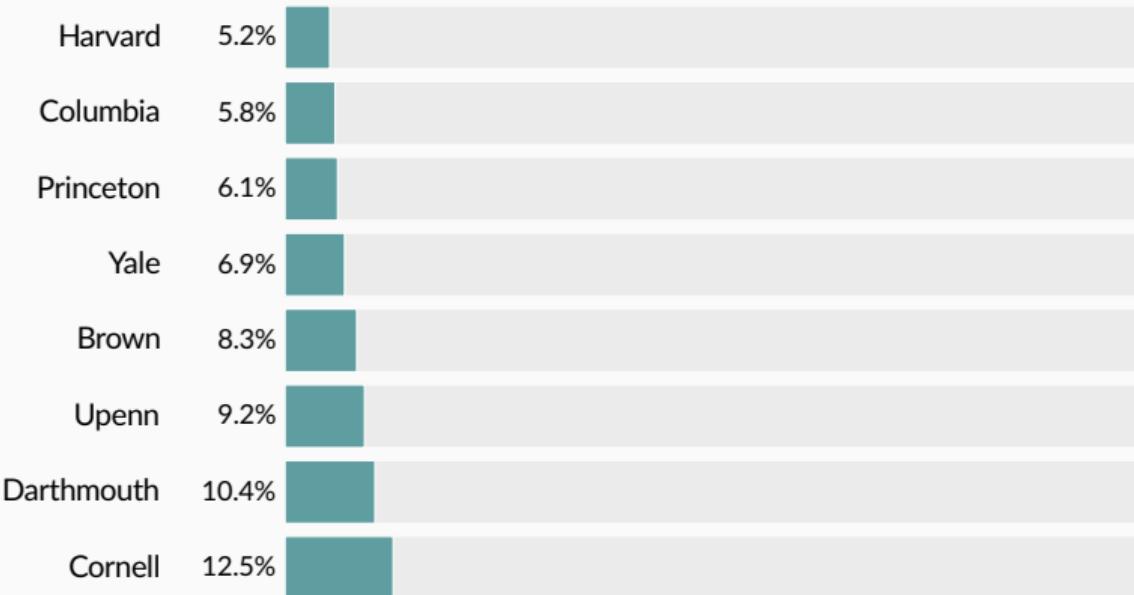
Case 1



Case 1 - Redesign

Các trường đại học cạnh tranh nhất nước Mỹ

So sánh dựa trên tỷ lệ trúng tuyển niên khóa 2017-2021



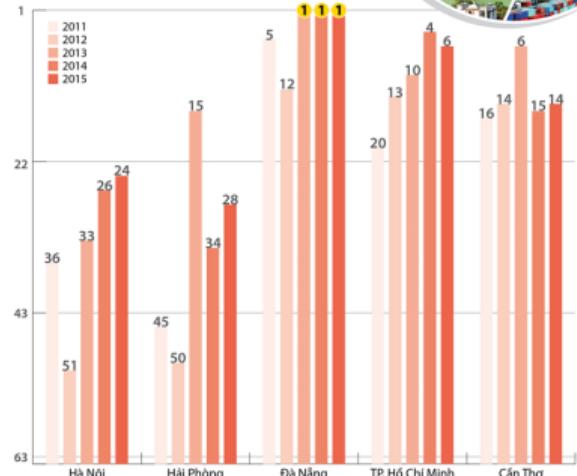
Nguồn: Business Insider

Case 2

Nhìn lại xếp hạng PCI của 5 thành phố trực thuộc Trung ương

Chi số năng lực cạnh tranh cấp tỉnh (PCI) nhằm đánh giá, xếp hạng môi trường kinh doanh, chất lượng điều hành kinh tế và nỗ lực cải cách hành chính của chính quyền 63 tỉnh/thành phố.

Thứ hạng PCI của 5 thành phố trực thuộc Trung ương
(so với 63 tỉnh/thành)



10 chỉ tiêu xếp hạng

- Chỉ phí giao nhập thị trường
- Tiếp cận đất đai
- Tinh minh bạch
- Chi phí thời gian
- Chi phí không chính thức
- Cảnh tranh bình đẳng
- Dịch vụ hỗ trợ doanh nghiệp
- Đào tạo lao động
- Thiết chế pháp lý
- Tính năng động

Nguồn: Phòng Thương mại và Công nghiệp Việt Nam (VCCI).

<http://infographics.vn>

Case 2 - Redesign

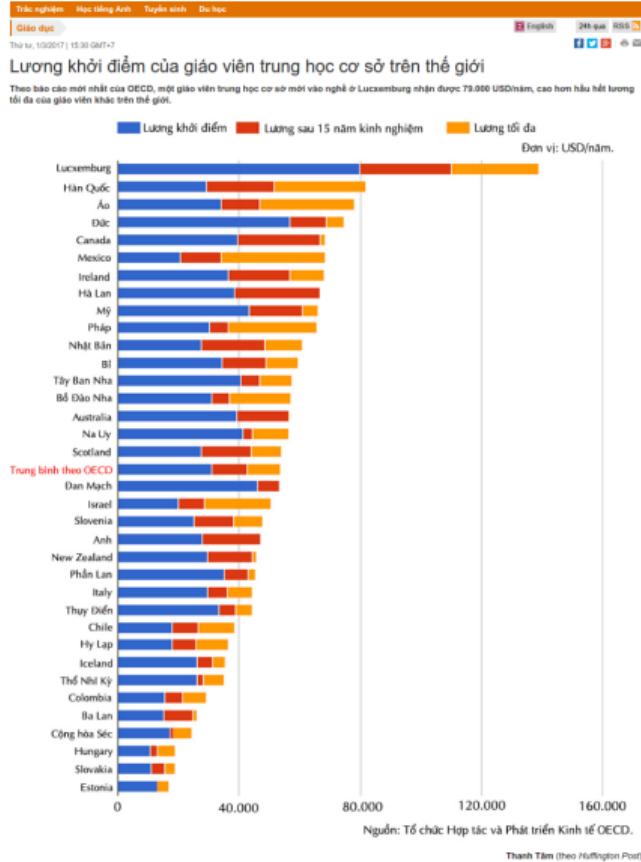
Thứ hạng PCI của 5 thành phố trực thuộc TW

Chỉ số PCI do lường chất lượng môi trường kinh doanh, điều hành kinh tế
và cải cách hành chính của chính quyền 63 tỉnh/thành phố



Nguồn: VCCI

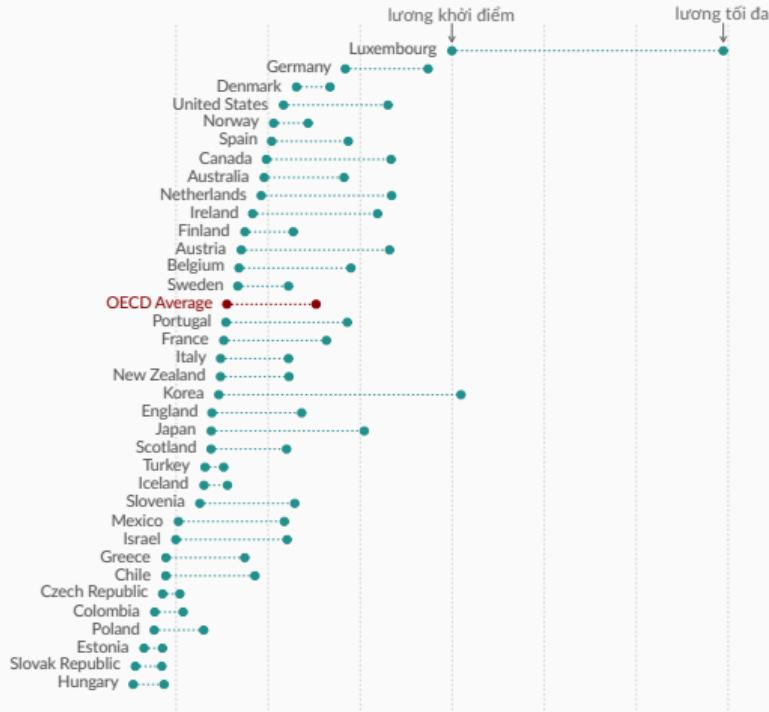
Case 3



Case 3 - Redesign

Cách biệt về lương giáo viên các quốc gia OECD

So sánh lương khởi điểm và lương tối đa của giáo viên trung học cơ sở, 2013



Nguồn: OECD

Where to go
from here

Blogs to follow

- ▶ <http://flowingdata.com/>, [Nathan Yau](#) (lots of R tutorials)
- ▶ <https://eagereyes.org/>, [Robert Kosara](#)
- ▶ <http://junkcharts.typepad.com/>, [Kaiser Fung](#)
- ▶ <http://www.perceptualedge.com/library.php>, [Stephen Few](#)
- ▶ <http://www.thefunctionalart.com/>, [Alberto Cairo](#)
- ▶ <http://www.visualisingdata.com/>, [Andy Kirk](#)
- ▶ <http://www.randalolson.com/blog/>, [Randal S. Olson](#)

Books to read. Classic

- ▶ The Visual Display Of Quantitative Information, [Edward R.Tufte](#)
- ▶ Visual Explanations, [Edward R.Tufte](#)
- ▶ Envisioning Information, [Edward R.Tufte](#)
- ▶ Beautiful Evidence, [Edward R.Tufte](#)
- ▶ The Elements of Graphing Data, [William Cleveland](#)
- ▶ Visualizing Data, [William Cleveland](#)
- ▶ Semiology of Graphics, [Jacques Bertin](#)
- ▶ Exploratory Data Analysis, [John W. Tukey](#)

Books to read. Accessible and Recent

- ▶ The Truthful Art, [Albert Cairo](#)
- ▶ The Functional Art, [Albert Cairo](#)
- ▶ Visualize This, [Nathan Yau](#)
- ▶ Data Points, [Nathan Yau](#)
- ▶ Information Dashboard Design, [Stephen Few](#)
- ▶ Show Me The Numbers, [Stephen Few](#)
- ▶ Now you see it, [Stephen Few](#)
- ▶ Signal, [Stephen Few](#)
- ▶ Storytelling With Data, [Cole Nussbaumer Knaflic](#)
- ▶ Creating More Effective Graphs, [Naomi B. Robbins](#)
- ▶ The Wall Street Journal Guide to Information Graphics, [Dona M. Wong](#)

Books to read. R-centric

- ▶ R Graphics, [Paul Murrell](#)
- ▶ ggplot2 - Elegant Graphics for Data Analysis, [Hadley Wickham](#)
- ▶ lattice - Multivariate Data Visualization with R, [Deepayan Sarkar](#)
- ▶ R Graphics Cookbook: Practical Recipes for Visualizing Data, [Winston Chang](#)
- ▶ Data Visualisation with R - 100 Examples, [Thomas Rahlf](#)
- ▶ Graphing Data With R, [John Jay Hilfiger](#)
- ▶ Graphics for Statistics and Data Analysis with R, [Kevin J. Keen](#)
- ▶ Graphical Data Analysis with R, [Antony Unwin](#)

Books to read. Design

- ▶ Data Visualisation: A Handbook for Data Driven Design, [Andy Kirk](#)
- ▶ Data Visualization: A Successful Design Process, [Andy Kirk](#)
- ▶ Information Visualization: Perception for Design, [Colin Ware](#)
- ▶ Visual Thinking for Design, [Colin Ware](#)
- ▶ Designing Data Visualizations: Representing Informational Relationships, [Noah Iliinsky](#)
- ▶ Visualization Analysis and Design, [Tamara Munzner](#)
- ▶ Design for Information, [Isabel Meirelles](#)
- ▶ The Non-designer's Design Book, [Robin Williams](#)