02 Legible Graphics

36-721 Statistical Graphics and Visualization

Jerzy Wieczorek

9/3/15

Last time

- Stat graphs + vis in larger context
- Syllabus
- (Historical examples: review on your own)

Today

- HW 1: assignment, submissions, grading
- Intro to base R graphs and documentation
- Legible and reproducible: image formats, resolution, saving images from R
- Comprehensible: titles, labels, legends, axes
- The core charts: scatter, line, bar
- Statistical summaries: hist, density, box, regression line, loess
- Other tips: jittering, log scales

Today

Follow along:

- ▶ Editable code in `02_Legible_code.R`
- ▶ Code with output examples in `02_Legible_code.html`

Base R graphs

- Read in data from a CSV file
- Scatterplots with `plot(x, y)`
- Documentation:
  - `?` and `example()`
  - Quick-R
  - Chang, *R graphics cookbook* (CMU, Amazon)

Legible graphics

Image formats, resolution, saving images from R

Vector vs bitmap

Vector vs bitmap explained

Bitmap: common formats

- **jpg/jpeg** is lossy, designed for photos but not text/charts
- **png** is lossless, good for text/charts, common on web

Vector: common formats

- **svg** can display in browser, used in D3.js
- **pdf** is for standalone doc (or to put inside another pdf)

Recommended formats and resolutions

Software	Recommended graphics device
Illustrator	svg
latex	ps
pdflatex	pdf, png (600 ppi)
Office	png (600 ppi)
web	png (72 ppi)

(ppi = Pixels Per Inch)

Table 8.3 in `ggplot2` book

Save images at intended final size

Better quality than changing size after saving

- LaTeX article default textwidth is around 5.4 inches
- Blogs may have a default width, e.g. 500 pixels

Saving images from R

- `png`, `pdf`
- `dev.off`
- `width`, `height`, `units`, `pointsize`, `res`

Saving images from R

Why bother saving files via code?
Why not just click "Export" in RStudio?

Export is handy, esp. for previewing image dimensions. But when
something changes in your huge analysis (revised data, new model)
and you need to remake 30 plots at once, then
automated code beats clicking the Export menu 30 times.

Comprehensible

- `plot(main, xlab, ylab, las, col, pch, xaxt)`
- `legend`
- `axis`

par cheat sheet
Also useful:

- `text`
- `par(mfrow, cex, lty, lwd)`

The core charts in base R

- Plotting data directly: scatter, line, bar
- Plotting summaries: histogram, box, trend line
- Other tips: jittering, log scales

Plotting data directly

- `plot(x, y, type, xlim, ylim)`
- `points, lines`
- `barplot, barplot(table(...))`

Also useful:

- `matplot`
- `curve`

Plotting summaries

- `hist`
- `plot(density(...))`
- `boxplot, boxplot(y ~ x)`
- `abline(lm(...))`
- `scatter.smooth`

Other tricks

- jitter
- plot(log)

For next time

- HW 1 due Saturday at 5pm, through Blackboard
- Read Cairo Ch 5-7
- We'll cover human visual perception and cognition, and how to apply findings to your graphics