Ggplot2 Chapter 3 – Mastering the Grammar

Vincenzo

September 29, 2014

Why learn the Grammar?

- More customizability
- Enable you to discover new graphic types
  - One that effectively conveys your story.

  "the grammar effectively defines the parameter space of statistical graphics" – Hadley

- Ch3 = an overview. Best to return to relate subsequent chapters back to the grammar.

Six Components

- Seemingly not explicitly listed in ggplot2 book.
- From Wilkinson's Grammar of Graphics (2e), pg. 19:

Aesthetics

Geometry

Scales

Statistics

Coordinates

~~Algebra~~   **Layers?**
(hence **Layered** gg)

1. Aesthetics

- These are ways to perceive the data.
- We can't "see" data; we can see aesthetic attributes.
- Examples:
  - Horizontal or Vertical position
  - Radial position
  - Colour
  - Shape
  - Size

1. Aesthetics

- Mapping between **Variables** and **Aesthetics**.
  - Not between the data and aesthetics!
    - (Table 3.2 – only the headings have been mapped)

- Example:
  - What are the mappings in this Pie graph?

2. Geometric Objects

- The "type" of object that's plotted.
  - Points
  - Bars
  - Boxplots
  - Lines

- For simple plots, geom's determine **plot name.**
  - Table 3.3 for example.

3. Scales

- Maps *data values* to "physical units"
  - i.e. units understood by the computer.

- As opposed to aesthetics (maps *variable*)

- Example: "Position" aesthetic
  - Data value → [0,1]

3. Scales

- Another example: Colour
  - 3 types of cones in the eye ==> 3d colour space

  - Continuous variable → 3d colour space

  - Discrete data values → evenly spaced hues
    - Figure 3.4

4. Statistics

- A statistical transformation of the data.

- Examples:
  - Mean Regression
    - (x,y) data → (x, E(Y)) data
  - Box Plots
    - *x* data → some quantiles
  - Includes Wilkinson's "Algebra" component?
    - (x,y) data → x/y, as an example.

5. Coordinates

- Mapping the position of objects on the plane.
  - Cartesian (some sort of identity mapping)
  - Polar
  - Map projections (sphere --> plane)

- Determines the axes/grid lines.
  - Example: Rose diagram
    - Binned angle and quantity.

6. Layers

- Basically a "single plot" with one set of:
  - aesthetics
  - stat
  - Geom
  - position adjustment (described next)

- Form a 3d array
  - Faceting (2d) and stacking (1d).

4. Scales (Again?)

- There's more to scaling when there are layers!
- Coordinates need to match up.
- 3 Steps:
  1. Transforming
     - Sounds just like "Coordinate" component to me.
  2. Training
     - Matching up scales across layers (so they match up)
  3. Mapping
     - Confusing – "map the data values into aesthetic values?"

Data Structures
(Not part of the Grammar)

- Section 3.6 discusses what you can do with ggplot2 objects (i.e. the plots)
  - print() to view it (only in a loop or function)
  - ggsave() to save the image to disk
  - save() to save the object to disk
    - Load with load()
  - As usual, describe the object with summary().