R Lesson 6: Data Visualization

vanderbi.lt/r

Steve Baskauf

DISC DIGITAL SCHOLARSHIP AND COMMUNICATIONS

Recommended for this lesson:

- chapter 3 of "R For Data Science" https://r4ds.had.co.nz/

- Data Visualization Cheatsheet https://rstudio.com/resources/cheatsheets/

Grammar of Graphics ("gg")
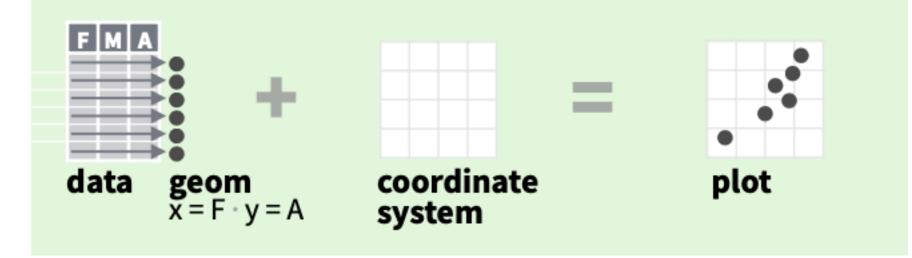
G of G: Aspects of a graphic

- data – the underlying dataset to be visualized
- geometry – the type and shape of the graph
- aesthetic – the colors, markers, size (visual aspects)

- plus optional:
  - stat
  - scale
  - facet
  - coord
  - theme
  - …

(data -> geom) + coord = plot

ggplot2 is based on the **grammar of graphics**, the idea that you can build every graph from the same components: a **data** set, a **coordinate system**, and geoms—visual marks that represent data points.

data   geom
x = F · y = A

coordinate
system

plot

(figures from data viz cheat sheet)

aesthetic controls visual properties

To display values, map variables in the data to visual properties of the geom (**aesthetics**) like **size, color**, and **x** and **y** locations.

data

geom
x = F · y = A
color = F
size = A

coordinate system

=

plot

Multiple layers can be built on the same plot

format of ggplot reflects G of G

Complete the template below to build a graph.

ggplot (data = <DATA> **) +**   ⌐required

<GEOM_FUNCTION> **(mapping = aes(** <MAPPINGS> **),**

stat = <STAT> , position = <POSITION> **) +**   Not required, sensible defaults supplied

<COORDINATE_FUNCTION> +

<FACET_FUNCTION> +

<SCALE_FUNCTION> +

<THEME_FUNCTION>

ggplot(data = mpg, **aes(**x = cty, y = hwy**))** Begins a plot that you finish by adding layers to. Add one geom function per layer.

Play with the schools data

DiSC DIGITAL SCHOLARSHIP AND COMMUNICATIONS