

NIH National Institutes of Health
Office of Management

Introduction to Data Visualization with R's ggplot2

Lisa Federer, MLIS, MA, AHIP
Research Data Informationist

NIH Library | <http://nihlibrary.nih.gov>

Office of Research Services
National Institutes of Health
U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES

Office of Research Services
Serving the NIH Community

Class outline

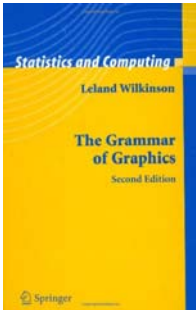
- The Grammar of Graphics: an intro
 - Components of visualizations
 - Practical considerations for aesthetic choices
- Using RStudio and ggplot2
- Your questions!

NIH National Institutes of Health
Office of Management

NIH Library | <http://nihlibrary.nih.gov>

Office of Research Services
Serving the NIH Community

The Grammar of Graphics



"A language consisting of words and no grammar expresses only as many ideas as there are words. By specifying how words are combined in statements, a grammar expands a language's scope...The grammar of graphics takes us beyond a limited set of charts (words) to an almost unlimited world of graphical forms (statements)."

NIH National Institutes of Health
Office of Management

NIH Library | <http://nihlibrary.nih.gov>

Office of Research Services
Serving the NIH Community

GG components

- **Data:** what is being visualized.
- **Mappings:** mappings between variables in the data and components of the chart.
- **Geometric Objects:** geometric objects that are used to display the data, such as points, lines, or shapes.
- **Aesthetic Properties:** properties determining how the plot looks, such as, typeface sizes, label locations, and tick marks.
- **Scales:** control how variables are mapped to aesthetics.
- **Coordinates:** describe how data is mapped to the plot
- **Statistical Transformations:** applied to the data to summarize it.
- **Facets:** describe how the data is partitioned into subsets and how these different subsets are plotted.

NIH National Institutes of Health
Office of Management

NIH Library | <http://nihlibrary.nih.gov>

Office of Research Services
Serving the NIH Community

Data and Mapping

Observation	Subject	Time	Concentration
2	1	0.50	0.94
3	1	0.75	0.78
4	1	1.00	0.48
5	1	1.25	0.37
6	1	2.00	0.19
7	1	3.00	0.12
8	1	4.00	0.11

NIH National Institutes of Health
Office of Management

NIH Library | <http://nihlibrary.nih.gov>

Office of Research Services
Serving the NIH Community

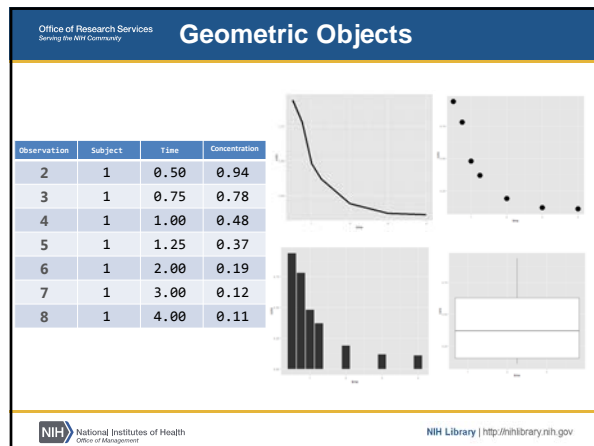
Statistical Transformations

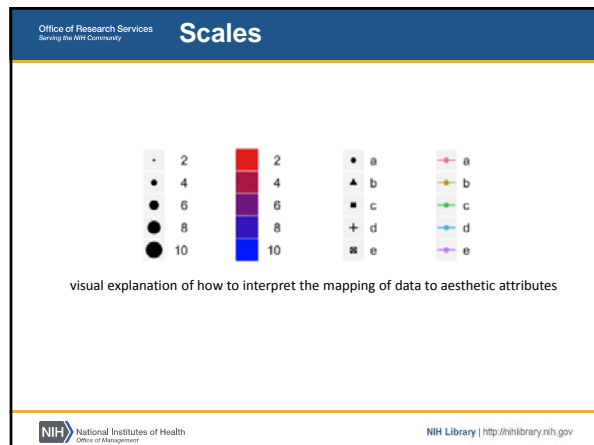
Name	Description
Bin	divide continuous range into bins, and count number of points in each
Boxplot	compute statistics necessary for boxplot
Contour	Calculate contour lines
Density	Compute 1d density estimate
Identity	Identity transformation, $f(x) = x$
Jitter	Jitter values by adding small random value
Qq	Calculate values for quantile-quantile plot
Quantile	Quantile regression
Smooth	Smooother conditional mean of y given x
Summary	Aggregate values of y for given x
Unique	Remove duplicated observations

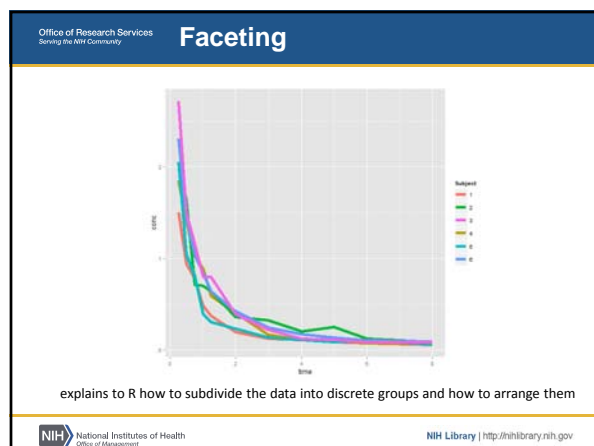
For more information, see <http://vita.had.co.nz/papers/layered-grammar.pdf>

NIH National Institutes of Health
Office of Management

NIH Library | <http://nihlibrary.nih.gov>







Office of Research Services
Serving the NIH Community

Suggested Reading

- Wickham, Hadley. *ggplot2: Elegant Graphics for Data Analysis*.
- Wickham, Hadley. "A Layered Grammar of Graphics." <http://vita.had.co.nz/papers/layered-grammar.pdf>
- Wilkinson, Leland. *The Grammar of Graphics*. Ebook available through NIH Library: <http://search.ebscohost.com/login.aspx?direct=true&db=nlebk&AN=155836&site=ehost-live>
- Winston, Chang. *R Graphics Cookbook*.

NIH National Institutes of Health
Office of Management

NIH Library | <http://nihlibrary.nih.gov>

Office of Research Services
Serving the NIH Community

Want to make a heatmap?

Check out materials from the basic heatmaps session:
<http://nihlibrary.campusguides.com/data-services/heatmaps>

Register for the upcoming advanced heatmaps webinar:
<https://attendee.gotowebinar.com/register/8281982639931948033>

NIH National Institutes of Health
Office of Management

NIH Library | <http://nihlibrary.nih.gov>
