

MassMutual DSDP/DEDP 2020:

CRASH COURSE IN ggplot2

July 6, 2020

R. Jordan Crouser

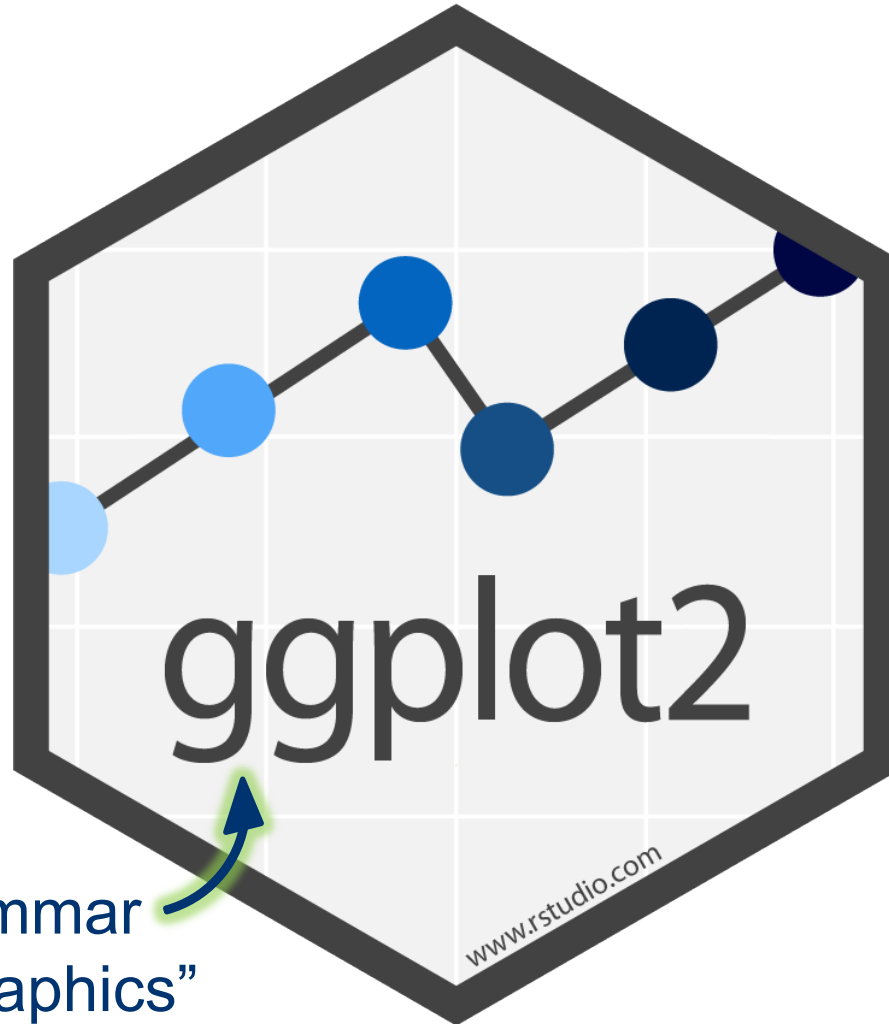
Assistant Professor of Computer Science

Smith College

Outline

- ✓ Introductions
- ✓ Visualization overview
 - ✓ Flashback to early experiences in data wrangling
 - ✓ Visualization (def.)
 - ✓ Data (def.)
- ✓ Graphical primitives
- ✓ Visual dimensions
- ✓ Pre-lunch activity: mapping visual to data dimensions
- After lunch: ggplot2 crash course

10-minute crash course



“(g)rammar
of (g)raphics”

What is the “Grammar of Graphics”?

- **Big idea:** independently specify plot *building blocks* and combine them to create graphical displays
- Building blocks include:
 - data (obvi.)
 - geometric objects (the literal stuff we draw)
 - aesthetic mappings (how we draw that stuff)
 - statistical transformations (underlying model)
 - scales (range of values, colors, etc.)
 - faceting (small multiples)

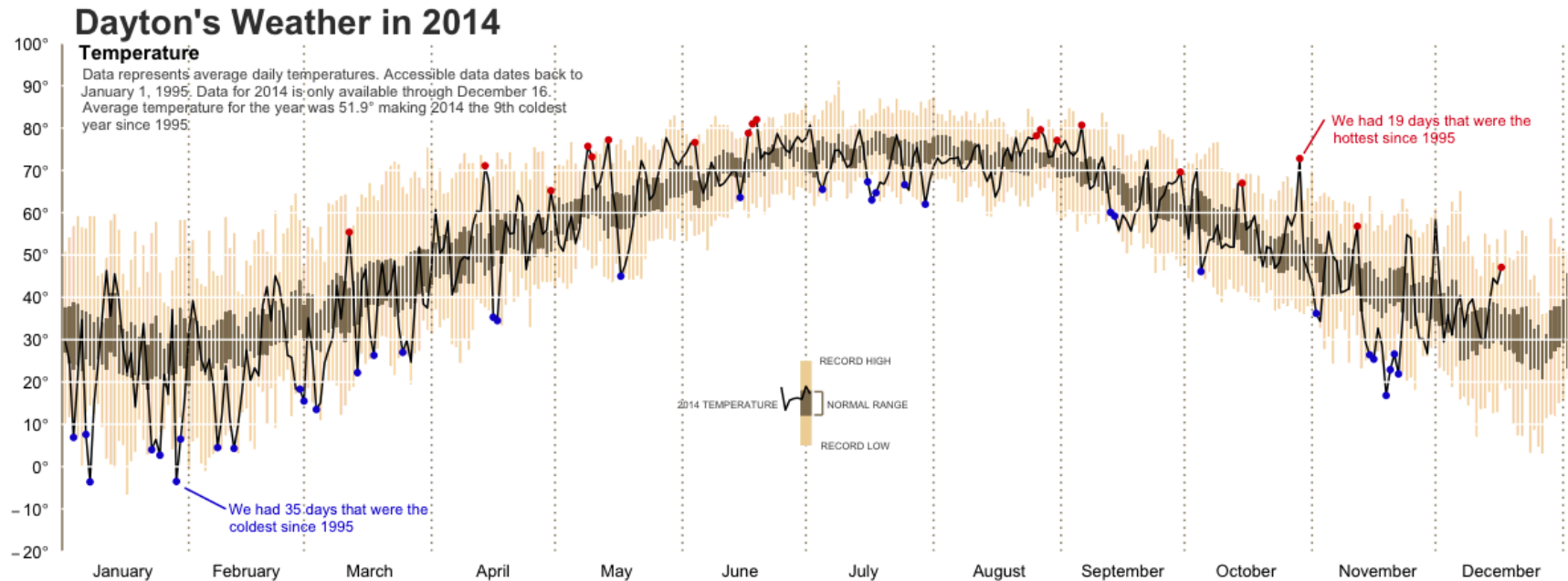
Geometric objects (geom)

- In `ggplot2`, the actual marks we put on a plot are called *geometric objects* or *geoms*
- Examples:
 - points (`geom_point`, for scatter plots, dot plots, etc.)
 - lines (`geom_line`, for time series, trend lines, etc.)
 - boxplot (`geom_boxplot`, for, well, boxplots!)
 - ... and many more!

Aesthetic mapping (aes)

- In `ggplot2`, an *aesthetic* means “something you can see”
- Aesthetic mappings are set with the `aes ()` function
- Each type of geom accepts only a subset of all aesthetics.
- Examples include:
 - position (i.e., on the x and y axes)
 - color (“outside” color)
 - fill (“inside” color)
 - shape (of points)
 - line type
 - size
 - ... and many more!

All you really need to know...



`geoms + aes()`

Lab: ggplot2 crash course

Crash Course in ggplot2

Introduction to Data Visualization - Summer 2020 Home Labs/Activities Datasets Resources

Introduction

- Geometric Objects and Aesthetics
- Scales: Controlling Aesthetics
- Faceting
- The #1 FAQ
- Your turn!

Crash Course in ggplot2 Introduction

Goal: by the end of this lab, you will be able to use `ggplot2` to build several different data-driven graphics.

Setting up

Remember: before we can use a library like `ggplot2`, we have to load it:

```
library(ggplot2)
```

Why ggplot2?

Advantages of `ggplot2`

- consistent underlying **grammar of graphics** (Wilkinson, 2005)
- plot specification at a high level of abstraction
- very flexible
- `theme` system for polishing plot appearance (more on this later)
- mature and complete graphics system
- many users, active mailing list

What Is The Grammar Of Graphics?

The big idea: independently specify plot building blocks and combine them to create just about any kind of graphical display you want. Building blocks of a graph include: