

Introduction to ggplot2

Part 2 (of 2)

Intro to Programming in R
Week 4

Myfanwy Johnston, PhD

The grammar that guides ggplot2

“Grammar of Graphics”: layerable graphics

Objective: understand the additional components of a ggplot, and how to use them to build more complex plots.

Source materials:

ggplot2: Elegant Graphics for Data Analysis (Hadley Wickham, PhD)
“Myfanwy Johnston on graphing with ggplot.” Davis R Users Group, 2015.

Key components of a `ggplot`:

1. Data
2. Aesthetic mappings
3. At least one layer*

*(usually created with a `geom` function)

Key components of a `ggplot`:

1. { Data
 2. { Aesthetic mappings
 3. At least one layer*
- Controlled by scales
- Statistical transformation
 - Geometry (“geom”)
 - Scale
 - Position adjustment

Layers

Points, lines, polygons

How you build your plot up

Typically done with a `geom_*()` call

```
ggplot(data = cod, mapping = aes(x = Depth, y = Intensity)) +  
  geom_point(aes(color = factor(Year))) + ...
```

statistical transformation

= "stat".

Transforms the data, usually by summarizing it in some way.

```
stat_identity()  
stat_bin()  
stat_summary()
```

geometry

geoms control the type of plot that you create.

They often have an accompanying underlying **statistical transformation**.

```
geom_point(), geom_bar(), geom_line(),  
  geom_boxplot(stat = "boxplot"),  
  geom_area(), geom_smooth()
```


scales

These control the mapping from data to
aesthetics.

When you want to change the appearance of
stuff on the **axes and guides** of your plots,
you will probably be doing so via the **scale**.

scales

Position scales are used for numeric and date/time data.

Color scales are used to map data to colours.

Manual scales are used to map *discrete* variables to custom aesthetics

position adjustment

These allow you to control the position of elements within a layer.

```
position_dodge(),  
position_stack(),  
position_fill(),  
position_jitterdodge()
```

facets

“Facetting” lets you create multiple small plots
by subgroup.

It’s like subsetting your data down to separate
groups, plotting each one, and lining up the plots
next to each other.

```
facet_wrap(~variable); facet_grid(~variable)
```

All together:

```
ggplot(cod,
       aes(x = Depth, y = Intensity)) +
  geom_point(aes(color = factor(Prevalence))) +
  scale_color_manual(values = c("blue",
                                "yellow")) +
  geom_smooth(lty = "dashed", size = 0.5) +
  facet_wrap(~Year, scales = "free_y")
```

data
mapping to aes

point layer: stat = "identity",
color scale = discrete

color scale adjustment

layer; fixed aesthetics (no
mapping)

facetting specification

Course GitHub repository:

<https://github.com/fishsciences/2020-R-Course>