Introduction
○○

Basics
○○○○○○○○○

Exploring Diamonds
○○

"New" Plots
○

Where to Learn More?
○

Graphics with R using ggplot2

Drew Schmidt

June 28, 2012

Introduction
oo

Basics
oooooooooo

Exploring Diamonds
oo

"New" Plots
o

Where to Learn More?
o

Contents

Introduction   Basics   Exploring Diamonds   "New" Plots   Where to Learn More?

○●   ○○○○○○○○○   ○○   ○   ○

Plotting Data in R

Plotting Data in R

1. Base Graphics: Easy for simple things, hard for complex things

Plotting Data in R

1. Base Graphics: Easy for simple things, hard for complex things
2. Grid: "Assembly language for graphics"

Plotting Data in R

1. Base Graphics: Easy for simple things, hard for complex things
2. Grid: "Assembly language for graphics"
3. ggplot2: Powerful and flexible, takes some time to learn

Plotting Data in R

1. Base Graphics: Easy for simple things, hard for complex things
2. Grid: "Assembly language for graphics"
3. ggplot2: Powerful and flexible, takes some time to learn
4. Lattice: "Like" ggplot2 in scope, but very different

Plotting Data in R

1. Base Graphics: Easy for simple things, hard for complex things
2. Grid: "Assembly language for graphics"
3. ggplot2: Powerful and flexible, takes some time to learn
4. Lattice: "Like" ggplot2 in scope, but very different
   . . .

Plotting Data in R

1. Base Graphics: Easy for simple things, hard for complex things
2. Grid: "Assembly language for graphics"
3. ggplot2: Powerful and flexible, takes some time to learn
4. Lattice: "Like" ggplot2 in scope, but very different
   . . .
5. And about 20 other packages:
   http://cran.r-project.org/web/views/Graphics.html

Ways of Plotting in ggplot2

Ways of Plotting in ggplot2

1. qplot syntax — Shorthand, good to learn eventually

Ways of Plotting in ggplot2

1. qplot syntax — Shorthand, good to learn eventually
2. layer syntax — Extremely rare 'in the wild'

Ways of Plotting in ggplot2

1. qplot syntax — Shorthand, good to learn eventually
2. layer syntax — Extremely rare 'in the wild'
3. geom/stat syntax — Probably the most popular way

Ways of Plotting in ggplot2

1. qplot syntax — Shorthand, good to learn eventually

2. layer syntax — Extremely rare 'in the wild'

3. geom/stat syntax — Probably the most popular way

4. autoplot syntax — Methods for plotting objects; mimics base graphics functionality

Ways of Plotting in ggplot2

1. qplot syntax — Shorthand, good to learn eventually
2. layer syntax — Extremely rare 'in the wild'
3. geom/stat syntax — Probably the most popular way
4. autoplot syntax — Methods for plotting objects; mimics base graphics functionality

See also: plot builder with Deducer (info on handout)

Introduction
○○

Basics
●○○○○○○○○○

Exploring Diamonds
○○

"New" Plots
○

Where to Learn More?
○

General Process to Produce a Plot with ggplot2

Introduction
○○

Basics
●○○○○○○○○○

Exploring Diamonds
○○

"New" Plots
○

Where to Learn More?
○

General Process to Produce a Plot with ggplot2

1 Declare which dataframe ggplot should use and set your aesthetics

General Process to Produce a Plot with ggplot2

1. Declare which dataframe ggplot should use and set your aesthetics

2. Add layers via geom and/or stat functions

General Process to Produce a Plot with ggplot2

1. Declare which dataframe ggplot should use and set your aesthetics
2. Add layers via geom and/or stat functions
3. Options and extras

General Process to Produce a Plot with ggplot2

1. Declare which dataframe ggplot should use and set your aesthetics
2. Add layers via geom and/or stat functions
3. Options and extras

General Process to Produce a Plot with ggplot2

1. Declare which dataframe ggplot should use and set your aesthetics
2. Add layers via geom and/or stat functions
3. Options and extras

ggplot2 requires that your data be stored in a dataframe

?????

Geoms? Aesthetics? Facets?

The Four Components to a Graphic in the Grammar of Graphics Schema

Introduction
○○

Basics
○○●○○○○○○○

Exploring Diamonds
○○

"New" Plots
○

Where to Learn More?
○

The Four Components to a Graphic in the Grammar of Graphics Schema

1. Geoms: "Physical components", e.g. point, path, line, polygon, . . .

The Four Components to a Graphic in the Grammar of Graphics Schema

1. Geoms: "Physical components", e.g. point, path, line, polygon, . . .

2. Aesthetics: "Visual cues", e.g. size, rotation, thickness, gradient, shape, color, . . .

The Four Components to a Graphic in the Grammar of Graphics Schema

1. Geoms: "Physical components", e.g. point, path, line, polygon, . . .

2. Aesthetics: "Visual cues", e.g. size, rotation, thickness, gradient, shape, color, . . .

3. Coordinates: Just what it sounds like: rectangular, polar, . . .

## The Four Components to a Graphic in the Grammar of Graphics Schema

1. Geoms: "Physical components", e.g. point, path, line, polygon, . . .

2. Aesthetics: "Visual cues", e.g. size, rotation, thickness, gradient, shape, color, . . .

3. Coordinates: Just what it sounds like: rectangular, polar, . . .

4. Faceting: Coplotting — more on this later

Customization

Everything in a ggplot2 plot can be customized:

Customization

Everything in a ggplot2 plot can be customized:
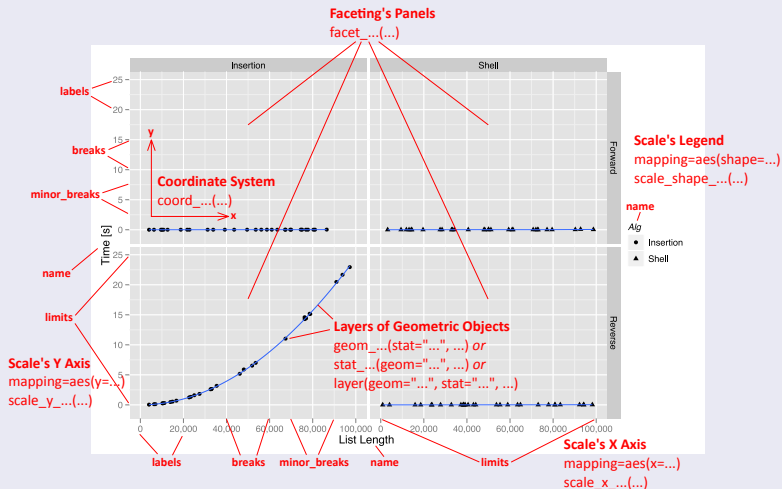
Geom and Stat Functions

There are many "layering" functions:

Geom and Stat Functions

There are many "layering" functions:

Geom Functions				
geom_abline	geom_bar	geom_blank	geom_contour	geom_density
geom_errorbar	geom_freqpoly	geom_histogram	geom_jitter	geom_linerange
geom_point	geom_polygon	geom_rect	geom_rug	geom_smooth
geom_text	geom_vline	geom_area	geom_bin2d	geom_boxplot
geom_crossbar	geom_density2d	geom_errorbarh	geom_hex	geom_hline
geom_line	geom_path	geom_pointrange	geom_quantile	geom_ribbon
geom_segment	geom_step	geom_tile		

Stat Functions				
stat_abline	stat_bin2d	stat_boxplot	stat_density	stat_function
stat_identity	stat_quantile	stat_spoke	stat_summary	stat_vline
stat_bin	stat_binhex	stat_contour	stat_density2d	stat_hline
stat_qq	stat_smooth	stat_sum	stat_unique	

Geom and Stat Functions

There are many "layering" functions:

Geom Functions				
geom_abline	geom_bar	geom_blank	geom_contour	geom_density
geom_errorbar	geom_freqpoly	geom_histogram	geom_jitter	geom_linerange
geom_point	geom_polygon	geom_rect	geom_rug	geom_smooth
geom_text	geom_vline	geom_area	geom_bin2d	geom_boxplot
geom_crossbar	geom_density2d	geom_errorbarh	geom_hex	geom_hline
geom_line	geom_path	geom_pointrange	geom_quantile	geom_ribbon
geom_segment	geom_step	geom_tile		

Stat Functions				
stat_abline	stat_bin2d	stat_boxplot	stat_density	stat_function
stat_identity	stat_quantile	stat_spoke	stat_summary	stat_vline
stat_bin	stat_binhex	stat_contour	stat_density2d	stat_hline
stat_qq	stat_smooth	stat_sum	stat_unique	

For explanations and examples other than those provided here, see
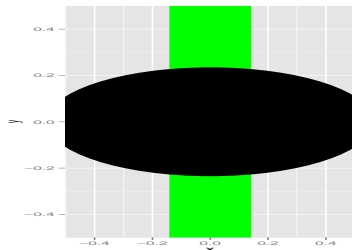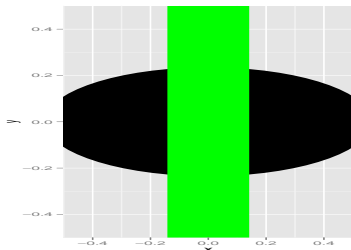the ggplot2 reference manual http://had.co.nz/ggplot2/

Adding Layers is Not Necessarily (though usually) Commutative

Introduction
oo

Basics
oooooo●ooo

Exploring Diamonds
oo

"New" Plots
o

Where to Learn More?
o

Adding Layers is Not Necessarily (though usually) Commutative

```
1  # Plot points layer then lines layer on top
2  g + geom_point() + geom_line()
3
4  # Plot lines layer then points layer on top
5  g + geom_line() + geom_point()
```

Adding Layers is Not Necessarily (though usually) Commutative

```
1  # Plot points layer then lines layer on top
2  g + geom_point() + geom_line()
3
4  # Plot lines layer then points layer on top
5  g + geom_line() + geom_point()
```

Some Simple Plots

Refer to Section 2, lines 5–23 of the file Rcode_ggplot2.R

Introduction
○○

Basics
○○○○○○○●○

Exploring Diamonds
○○

"New" Plots
○

Where to Learn More?
○

Saving Your Plots

Introduction
oo

Basics
ooooooo●o

Exploring Diamonds
oo

"New" Plots
o

Where to Learn More?
o

Saving Your Plots

```
1  # Option 1:  ggplot2 independent
2  pdf ("location/filename.pdf") # see help ("device") for
        other filetypes
3  g  # or last_plot() to save the last plot created by
        ggplot2
4  dev.off ()
5
6  # Option 2:  only for ggplot2 plots
7  ggsave ("location/filename.pdf", g)
```

Saving Your Plots

```
1  # Option 1: ggplot2 independent
2  pdf("location/filename.pdf") # see help("device") for
       other filetypes
3  g # or last_plot() to save the last plot created by
       ggplot2
4  dev.off()
5
6  # Option 2: only for ggplot2 plots
7  ggsave("location/filename.pdf", g)
```

See also

```
1  ?pdf
2  ?ggsave
```

Exercises

1. Create a histogram of the "carat" variable.

2. Save the plot you just made as both a pdf and a png.

3. Begin with:

```
1 g <- ggplot (data=diamonds , aes(x=clarity) )
```

Produce a barplot and a histogram with g (remember, "clarity" is categorical). Is there a difference?

Introduction
oo

Basics
oooooooooo

Exploring Diamonds
●o

"New" Plots
o

Where to Learn More?
o

Graphically Exploring the Diamonds Dataset

Refer to Section 3, lines 29–59 of the file Rcode_ggplot2.R

Exercises

1. Create scatterplots of price by carat faceted by color. How would you describe the relationship between price and carat across groups?

2. Every plot should tell a story. What story do our scatterplots tell about a diamond's carat and its price? (Just a short, one sentence explanation)

3. Refer to the subset plot above where we restricted the data only to those diamonds with color "J". Produce a scatterplot with a LOESS fit in the same plot. Do you notice anything striking in this plot (you may have noticed it in another plot above)?

Graphically Exploring the Diamonds Dataset

Refer to Section 4, lines 70–193 of the file Rcode_ggplot2.R

Where to Learn More?

Reference Manual:	http://had.co.nz/ggplot2/
CRAN page:	http://cran.r-project.org/web/packages/ggplot2
Wiki:	https://github.com/hadley/ggplot2/wiki/
Google Group:	https://groups.google.com/group/ggplot2
Tag on stackoverflow:	http://stackoverflow.com/questions/tagged/ggplot2
Blog:	http://blog.ggplot2.org/
Official Book:	http://tinyurl.com/ggplot2-book

Thanks for coming!

Questions?