

# 04 Grammar of Graphics

36-721 Statistical Graphics and Visualization

Jerzy Wiecek

9/10/15

## Last time

- ▶ Visual perception: quantitative comparisons, grouping and search, cognition, consistency
- ▶ Plotting symbols/colors/line types and layouts in base R
- ▶ HW 2 is coming up
- ▶ HW 1 is graded

## HW 2

To clarify:

- ▶ It's OK to just compare STEM vs non-STEM as the “fields”; larger dataset is **not** required, just there in case you're curious
- ▶ It's OK to use RMarkdown, knitr, Sweave, etc. instead of saving plots with `pdf()` and `png()` like in HW 1

# HW 1

- ▶ A score of 1 means **Competent**: no need to revise/resubmit.
- ▶ Any lower score means **Not yet competent**: please revise following the TA's comments and resubmit.

# Today

- ▶ Grammar of Graphics concept
- ▶ Tableau
- ▶ ggplot2

# Today

Follow along:

- ▶ Editable code in `04_GoG_code.R`
- ▶ Code with output examples in `04_GoG_code.html`

# Grammar of Graphics

- ▶ Examples
- ▶ Why bother
- ▶ History
- ▶ Components

# Grammar of Graphics: examples

Let's demonstrate on a small subset of diamonds dataset that comes with `ggplot2`  
(with black-and-white theme, and larger font)



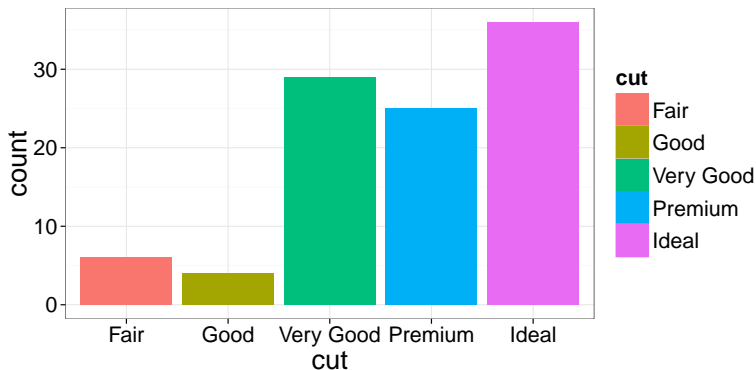
## Grammar of Graphics: examples

```
library(ggplot2)
theme_set(theme_bw() +
           theme(text = element_text(size = 24)))
dsmall = diamonds[sample(nrow(diamonds), 100),]
```

## Grammar of Graphics: examples

“Bar chart”: map discrete variable to x-axis and to color; compute counts-by-category, and map them to bar heights

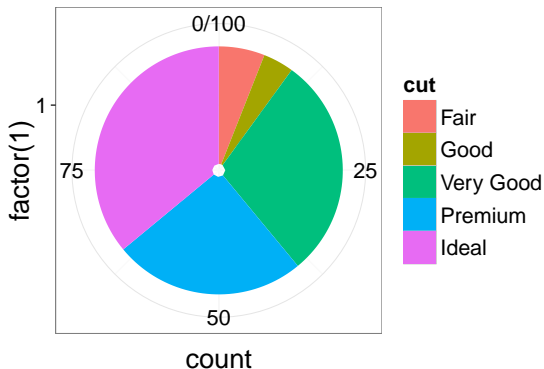
```
ggplot(data = dsmall, aes(x = cut, fill = cut)) +  
  geom_bar(stat = "bin") + coord_cartesian()
```



## Grammar of Graphics: examples

“Pie chart”: set a constant radius of 1 on polar coordinates; map discrete variable to color; compute counts-by-category, and map them to angles

```
ggplot(data = dsmall, aes(x = factor(1), fill = cut)) +  
  geom_bar(stat = "bin") + coord_polar(theta = "y")
```

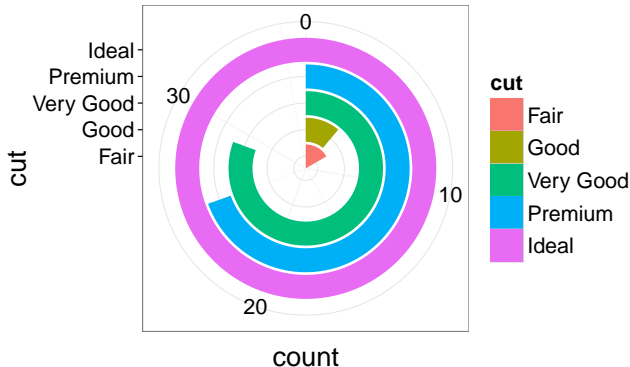


# Grammar of Graphics: examples

So... what if we map discrete variable to color **and radius** instead?

# Grammar of Graphics: examples

```
ggplot(data = dsmall, aes(x = cut, fill = cut)) +  
  geom_bar(stat = "bin") + coord_polar(theta = "y")
```



## Grammar of Graphics: examples

“Race track plot”: terrible idea, considering last lecture :) but nifty example of GoG’s flexibility.

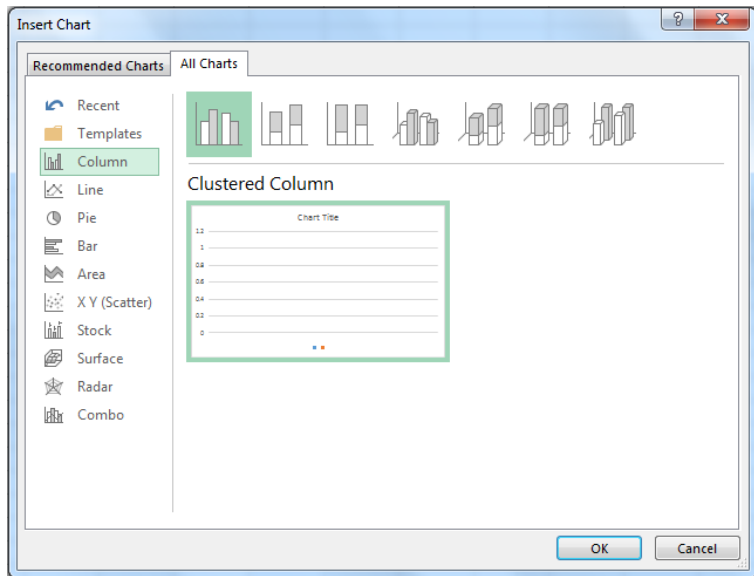
“This system is capable of producing some hideous graphics . . . This system cannot produce a meaningless graphic, however.”

–Leland Wilkinson, *The Grammar of Graphics*

We’ll see nicer examples in R demo soon!

# Grammar of Graphics: why bother

Expressing a graph from the ground up is more flexible than “chart zoo” approach (like Excel’s chart wizard)



## Grammar of Graphics: why bother

“The grammar is useful for you both as a user and as a potential developer of statistical graphics. As a user, it makes it easier for you to iteratively update a plot, changing a single feature at a time. The grammar is also useful because it suggests the high-level aspects of a plot that *can* be changed, giving you a framework to think about graphics, and hopefully shortening the distance from mind to paper. It also encourages the use of graphics customised to a particular problem, rather than relying on generic named graphics.”

–Hadley Wickham, `ggplot2`



# Grammar of Graphics: history and influence

- ▶ Leland Wilkinson, *The Grammar of Graphics*
- ▶ Hadley Wickham, ggplot2, popular R implementation
- ▶ [yeroon.net/ggplot2](http://yeroon.net/ggplot2), web GUI for ggplot2
- ▶ Tableau (Wilkinson now works there)
- ▶ SPSS Graphics Production Language (GPL) and Visualization Designer
- ▶ IBM VizJSON
- ▶ ...

# Grammar of Graphics: components

Wilkinson's grammar:

- ▶ data
- ▶ trans: variable transformation (identity, bin, smooth, quantile...)
- ▶ scale: scale transformation (axis limits, log scale, color mapping...)
- ▶ coord: Cartesian, polar, map projection...
- ▶ element: graphic element (point, line, bar...) with attributes (color, symbol, length...)
- ▶ guide: axes, legends, titles...

# Grammar of Graphics: components

ggplot2 specifications:

- ▶ data
- ▶ aes: aesthetic attributes (position, length, color, symbol...)
- ▶ stat: statistical variable transformation (identity, bin, smooth, quantile...)
- ▶ geom: geometric element (point, line, bar...)
- ▶ scale: scale transformation (axis limits, log scale, color mapping...)
- ▶ coord: Cartesian, polar, map projection...
- ▶ facet: divide into subplots / small multiples using a discrete variable

# Grammar of Graphics: components

More on ggplot2 specifications:

- ▶ Each layer has its own data, aes, stat, and geom ... then the scale and coord are coordinated across facets
- ▶ Finer control over stat summaries with group: see [documentation](#), [Oxboys example](#)
- ▶ Of course control over guides (axes, legends, titles...) is also possible

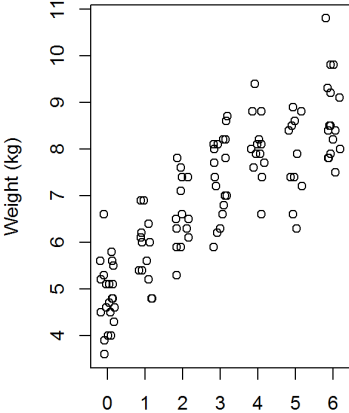
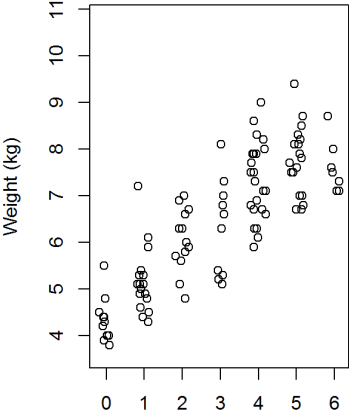
# Grammar of Graphics: practice

Example plot from last lecture:

What data map to which aes here? What stat, geom, scale, coord are used? Any facet?

## Weight vs Age, by Gender

Female Male



# Grammar of Graphics: practice

WHO Child Growth Standards, charts of **Length-for-age, percentiles, by gender**

What data map to which aes here? What stat, geom, scale, coord are used?

Any facet, if we consider Boy and Girl plots side-by-side?

## Grammar of Graphics: more resources

- ▶ ggplot2 [official documentation](#)
- ▶ ggplot2 [cheat sheet](#)
- ▶ [StackOverflow help for ggplot2](#)
- ▶ A nice ggplot2 [tutorial](#)
- ▶ Wickham's book ggplot2, especially Ch 3-4; free PDF on [Springer Link](#) through CMU
- ▶ Wilkinson's book *The Grammar of Graphics*, esp. last chapter "Coda"; free PDF on [Springer Link](#) through CMU

# Tableau

Polished implementation of Wilkinson's "graphboard" idea

Student license (1 year free)



# ggplot2

Follow along in R code.

We won't cover `qplot()`, a `ggplot2` wrapper function that acts more like base R, because I find it doesn't help explain the GoG concept.

## For next time

- ▶ We'll learn basic concepts of (Graphic) Design, how to apply them to your visualizations, and how to implement them in Inkscape/Illustrator
- ▶ **Install Inkscape** if you want to follow along
- ▶ Readings: Cairo Ch 8
- ▶ I strongly recommend *The Non-Designer's Design Book* (newest edition is great, but old editions OK)
- ▶ HW 2 due Saturday at 5pm, through Blackboard