Conceptual introduction to `ggplot`

Author: Nicholas G Reich

Choices for R graphics

You have three central choices for making graphics in R:

- "Base R graphics"
- `ggplot2`
- lattice

I use `ggplot` because:

1. it is integrated with the `tidyverse`
2. it is actively developed/maintained
3. there are a ton of extensions (see more later)

Understanding the “grammar” of ggplot2

The grammar ...

- layers (a ‘geom’, a ‘stat’, an ‘annotation’)
- aesthetics (‘aes‘)
- scales
- facets
- data
- ... and more here: http://ggplot2.tidyverse.org/reference/

What is a layer?

Layers define the basic structure of the elements on the plot

- Geoms: point, line, tile, boxplot, ribbon, ...
- Stats: histogram, smooth, density, ...
- Annotation: hline, vline, text, ...

For more info check out the documentation:
http://ggplot2.tidyverse.org/reference/.

What are "aesthetics"?

Aesthetics define a mapping between **tidy data** and the information required to create a specific graphic[1]

length	width	depth	trt
2	3	4	a
1	2	1	a
4	5	15	b
9	10	80	b

x	y	colour
2	3	a
1	2	a
4	5	b
9	10	b

[1] Figure credits: Hadley Wickham

geom_point

Each geom has a different set of aesthetics.
What information do we need to draw a scatterplot?
Or, asked another way, what aesthetics do we need for `geom_point`?

geom_point

Each geom has a different set of aesthetics.
What information do we need to draw a scatterplot?
Or, asked another way, what aesthetics do we need for `geom_point`?

- x (required)
- y (required)
- alpha
- color
- fill
- shape
- size

geom_line

What information do we need to draw a line of connected points?
Or, asked another way, what aesthetics do we need for `geom_line`?

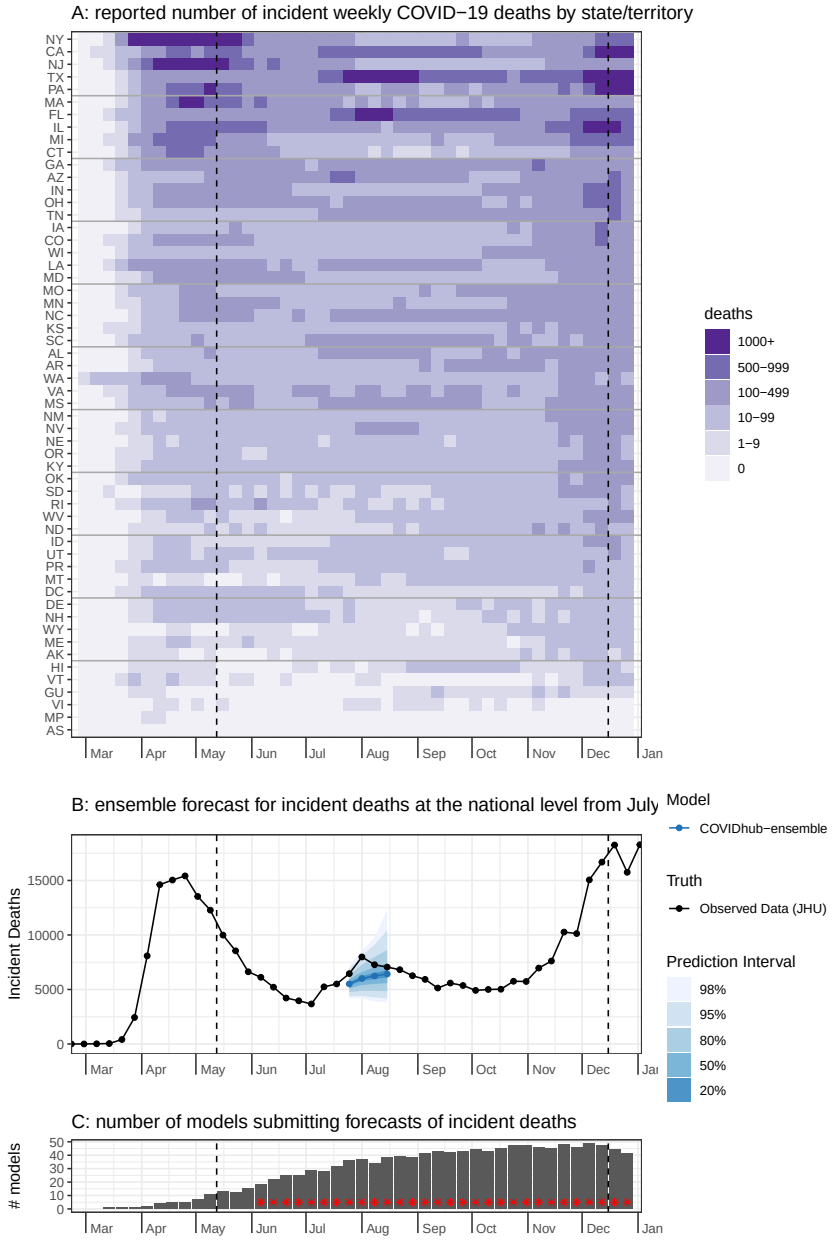
geom_line

What information do we need to draw a line of connected points?
Or, asked another way, what aesthetics do we need for geom_line?

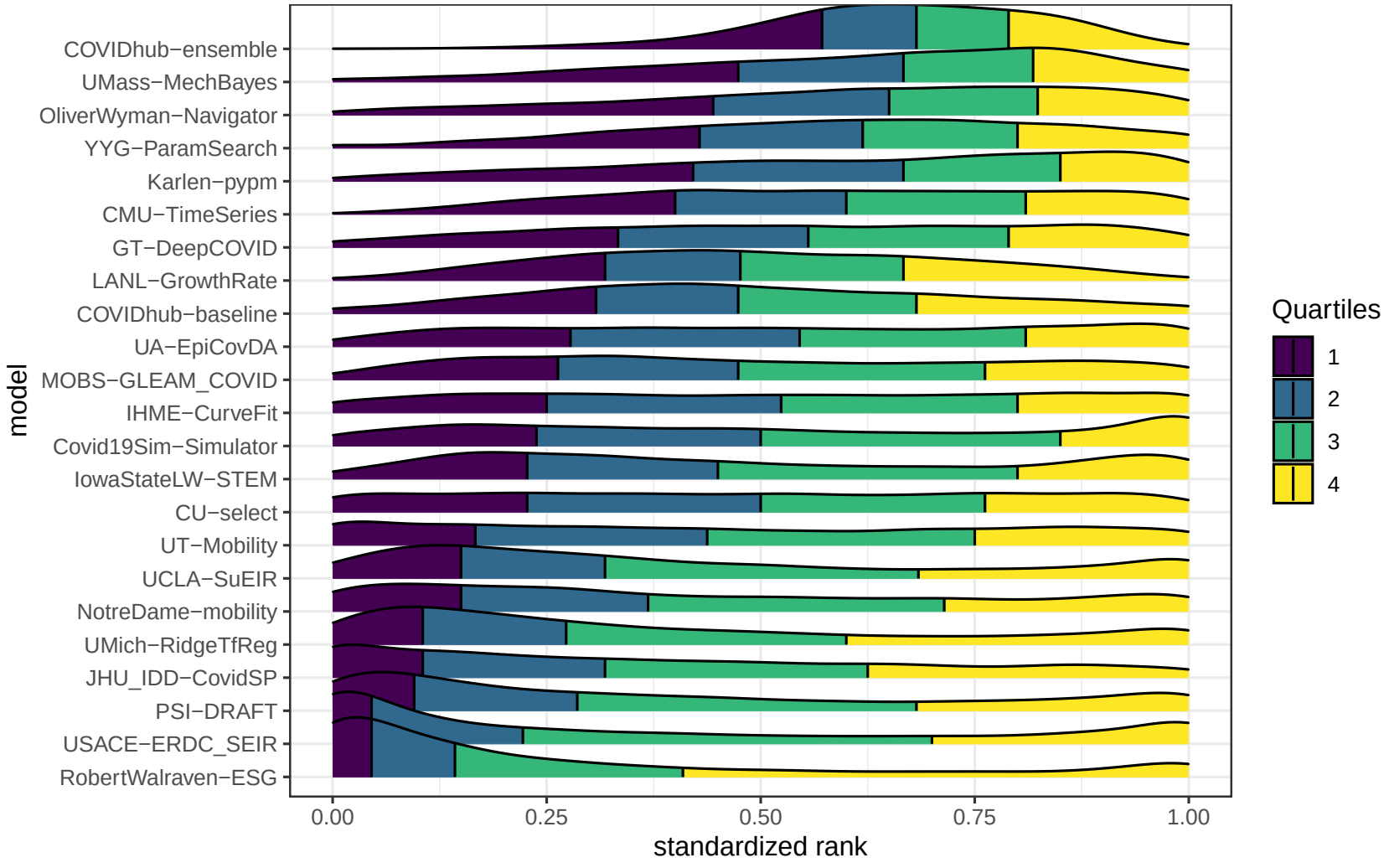
- x (required)
- y (required)
- alpha
- color
- linetype
- size

`ggplot` extensions that I used in a recent paper

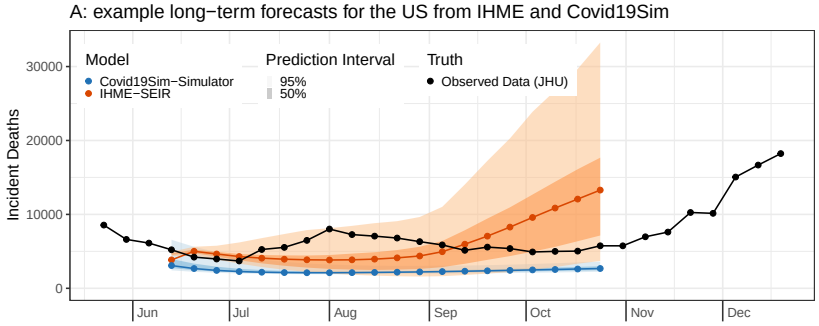
gridExtra or cowplot for multi-plot alignment

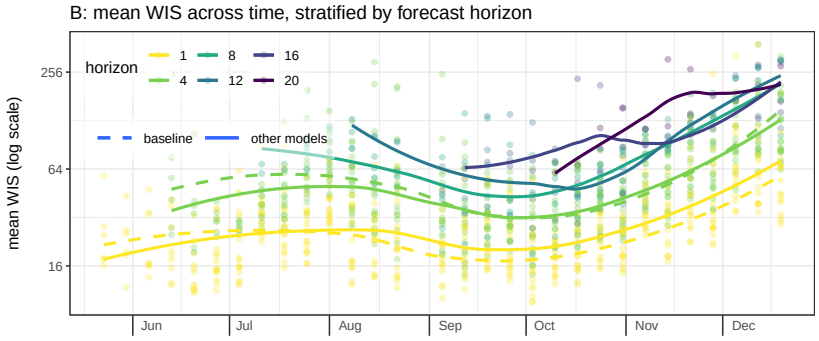


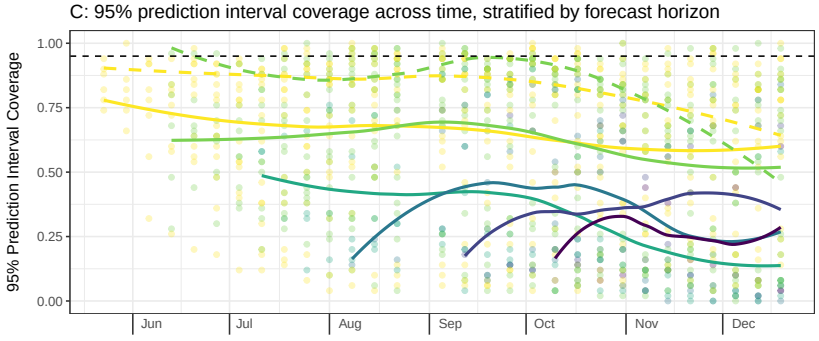
ggrides for ridgeplots

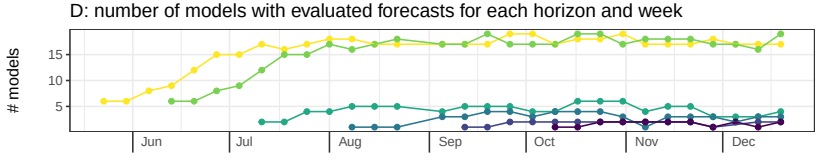


RColorBrewer and viridis for colors









Note-catcher

A figure from "Cities, traffic, and CO2: A multidecadal assessment of trends, drivers, and scaling relationships", Gately et al, PNAS, 2015. Original paper on Moodle.

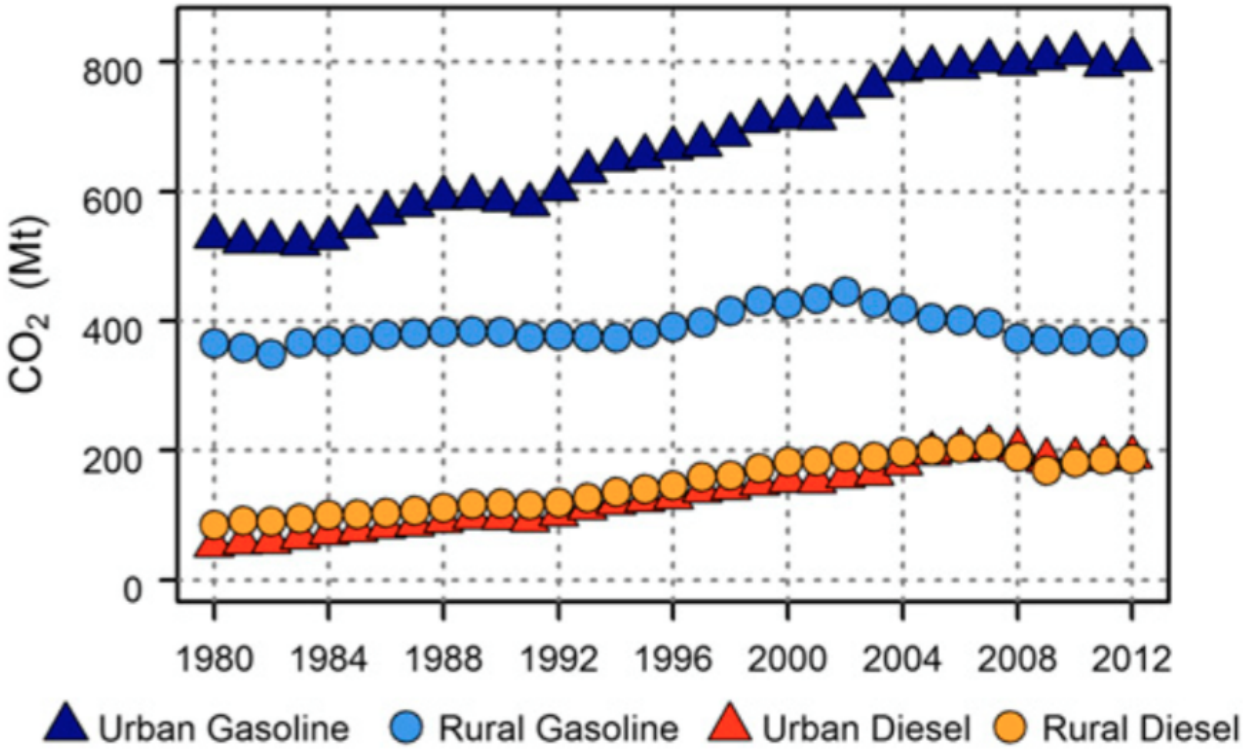


Fig. 2. Time series of US on-road CO_2 emissions. Urban roads accounted for 80% of total emissions growth since 1980. Rural road emissions have been declining since 2002.

Note-catcher

We have made the data from the CO2 emissions figure available on Moodle. As a group, you will be asked to complete the following tasks:

1. Recreate the figure as close as possible to the original.
2. Improve the figure. Make some changes that you think make the figure more clear.
3. Post your final figures on the Note-catcher document.

At the end of class, the class will vote on which figure is (1) closest to the original and (2) the best improvement.

NOTE

Everyone should be doing some coding here, and having a version of the graphic working on their laptop! Make sure it's not just 1-2 people typing the code and having it work for them.