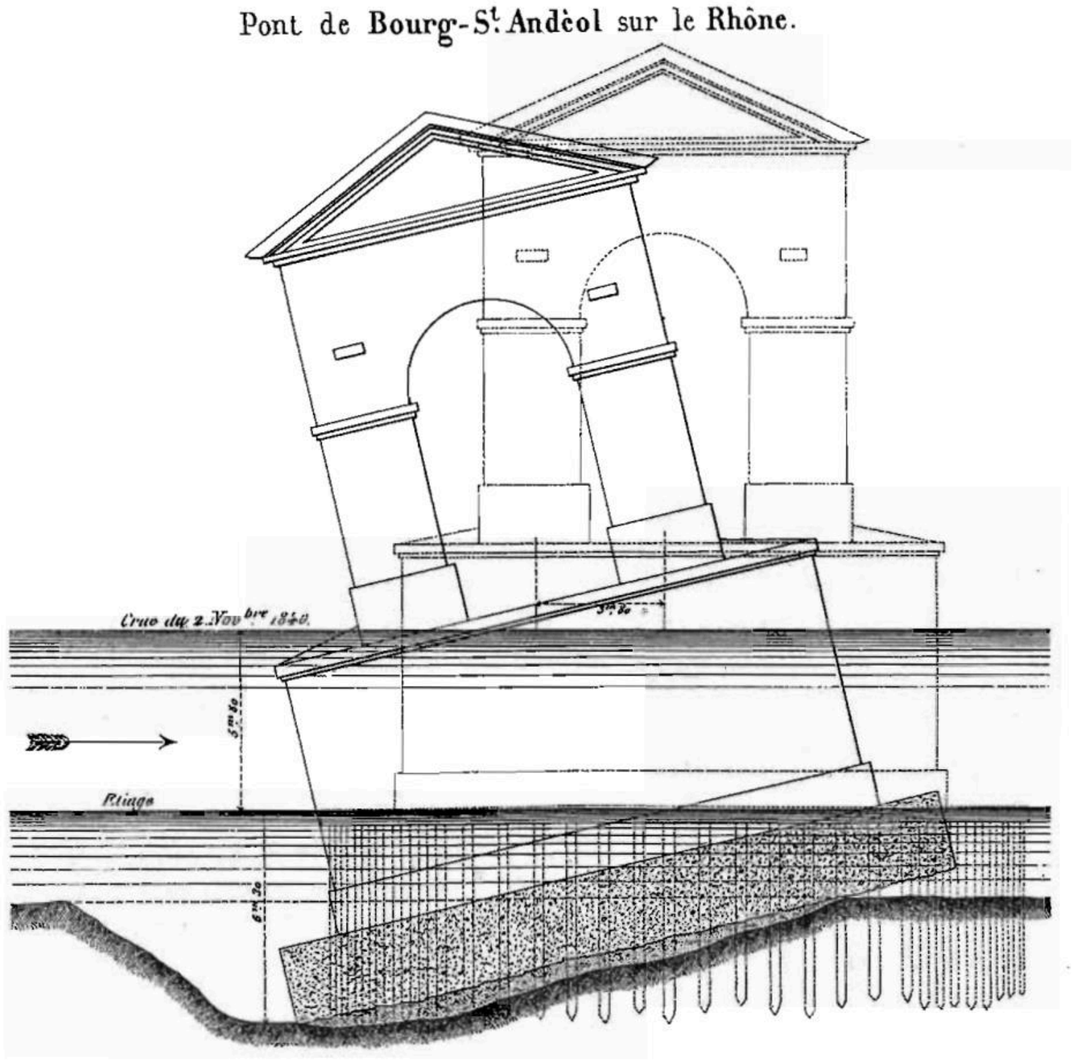
Data Visualization with ggplot2

Coffee, Cookie and Coding (C3) Workshop supported by the Public Health Data Science and Data Equity team

Shelby Golden, M.S.

November 24th, 2025



The Visual Display of Quantitative Information 2nd Edition by Edward R. Tufte. Charles Minard's before/after bridge collapse on the Rhône 1840 (pg. 39). Accessed March 14th, 2025.

Yale SCHOOL OF PUBLIC HEALTH
Data Science and Data Equity

Shelby Golden, M.S.

- Worked 7 years as a Molecular Biologist and Biochemist.
- Received a Masters in Applied Computational Mathematics from Johns Hopkins University in 2024.

Today's Learning Objectives

01 Classify the Grammar of Graphics layers used in `ggplot` syntax (~ 15 minutes)

02 Applications of different geometries, effective use of layering, and polishing the result (~ 45 minutes)

03 Interactive plots, map projections, and leverage AI assisted coding (~ 10 minutes)

Our Choice Resources

- *R for Data Science (2e)* by Hadley Wickham, Mine Çetinkaya-Rundel, and Garrett Grolemund
- *ggplot2: Elegant Graphics for Data Analysis (3e)* by Hadley Wickham, Danielle Navarro, and Thomas Lin Pedersen
- ggplot2 package documentation and cheat sheets by tidyverse. Specifically, the function references page
- "ggplot2 workshop" part 1 and part 2 by Thomas Lin Pedersen

Accessing the Codespaces

In this workshop, you'll need to access the R code prepared for the workshop discussions and challenge questions. If you haven't already, please download the latest version of R to your device. We also recommend using the latest version of RStudio as your Integrated Development Environment (IDE) [1].

> **Attribution and Ownership**
>
> Please note that all materials provided in this workshop, including any code added to your personal repository, belongs to DSDE. When using or referencing this material, please ensure to cite it correctly to give proper credit to the original authors.

> **Settings Used in Development**

Initializing the Environment

1. Download the prepared codespace, which contains a comprehensive code environment configured with the RStudio IDE for in-class discussions and challenge questions with solutions.

   Codespace

2. Unzipped the downloaded directory and move it to the file location you wish to house the project.

Command-Line Application

```
cd "file_path/Downloads/"                    # Open the directory the file was downloaded
unzip Data-Visualization-with-ggplot2.zip    # Unzip the file.
mv Data-Visualization-with-ggplot2 /new_path/    # Move the unzipped directory to the new loc
```

Worked-Through Example

Environment Set-Up and Data Description

First, we will load the necessary libraries and any special functions used in the script.

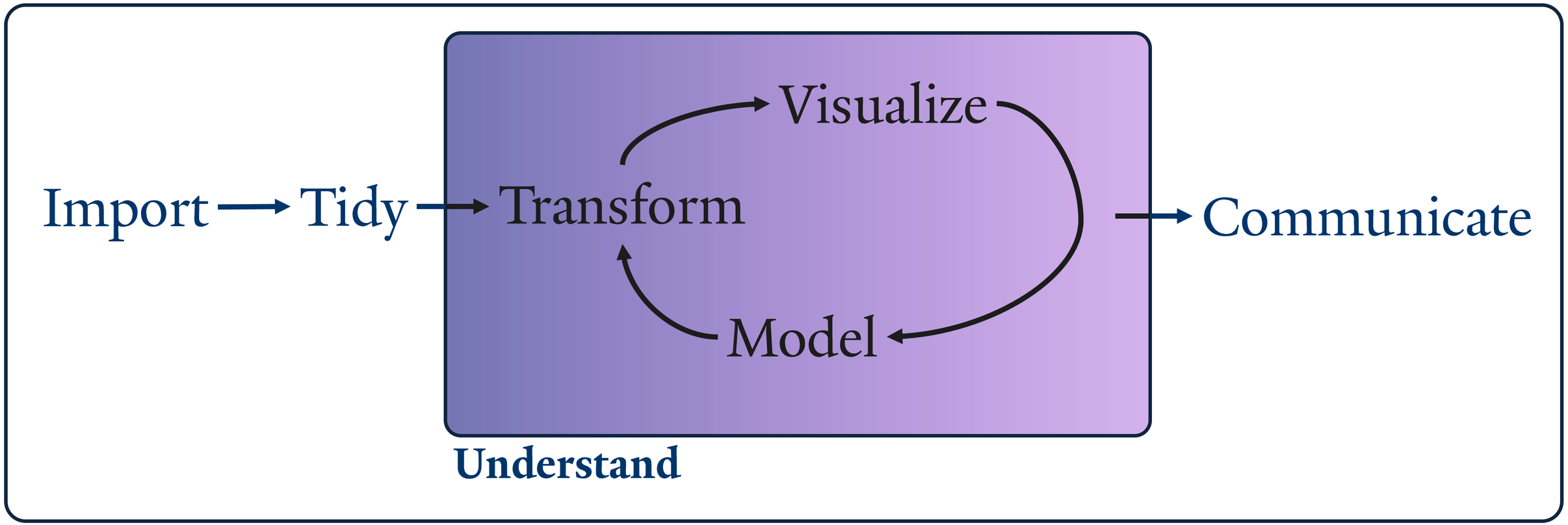
```
# NOTE: renv initialization might need to be run twice after the repo is
#       first copied.
#renv::init()
renv::restore()

suppressPackageStartupMessages({
  library("arrow")          # For reading in the data
  library("dplyr")          # For data manipulation
  library("ggplot2")        # For creating static visualizations
  library("plotly")         # For interactive plots
  library("cowplot")        # ggplot add on for composing figures
  library("tigris")         # Imports TIGER/Line shapefiles from the Census Bureau
  library("sf")             # Handles "Special Features": spatial vector data
  library("RColorBrewer")   # Load Color Brewer color palettes
  library("viridis")        # Load the Viridis color pallet
})

# Function to select "Not In"
'%!in%' <- function(x,y)!('%in%'(x,y))
```

Now we will import our cleaned and tidy data, which is ready for plotting. Students who would like to find out more about how to get their data into the plottable, tabular form you will see here can explore our A Journey into the World of tidyverse workshop.


Welcome to the tidyverse
(the abridged version)



R for Data Science (2e) - Introduction Figure 1 by Hadley Wickham, Mine Çetinkaya-Rundel, and Garrett Grolemund. Accessed November 15th, 2024.

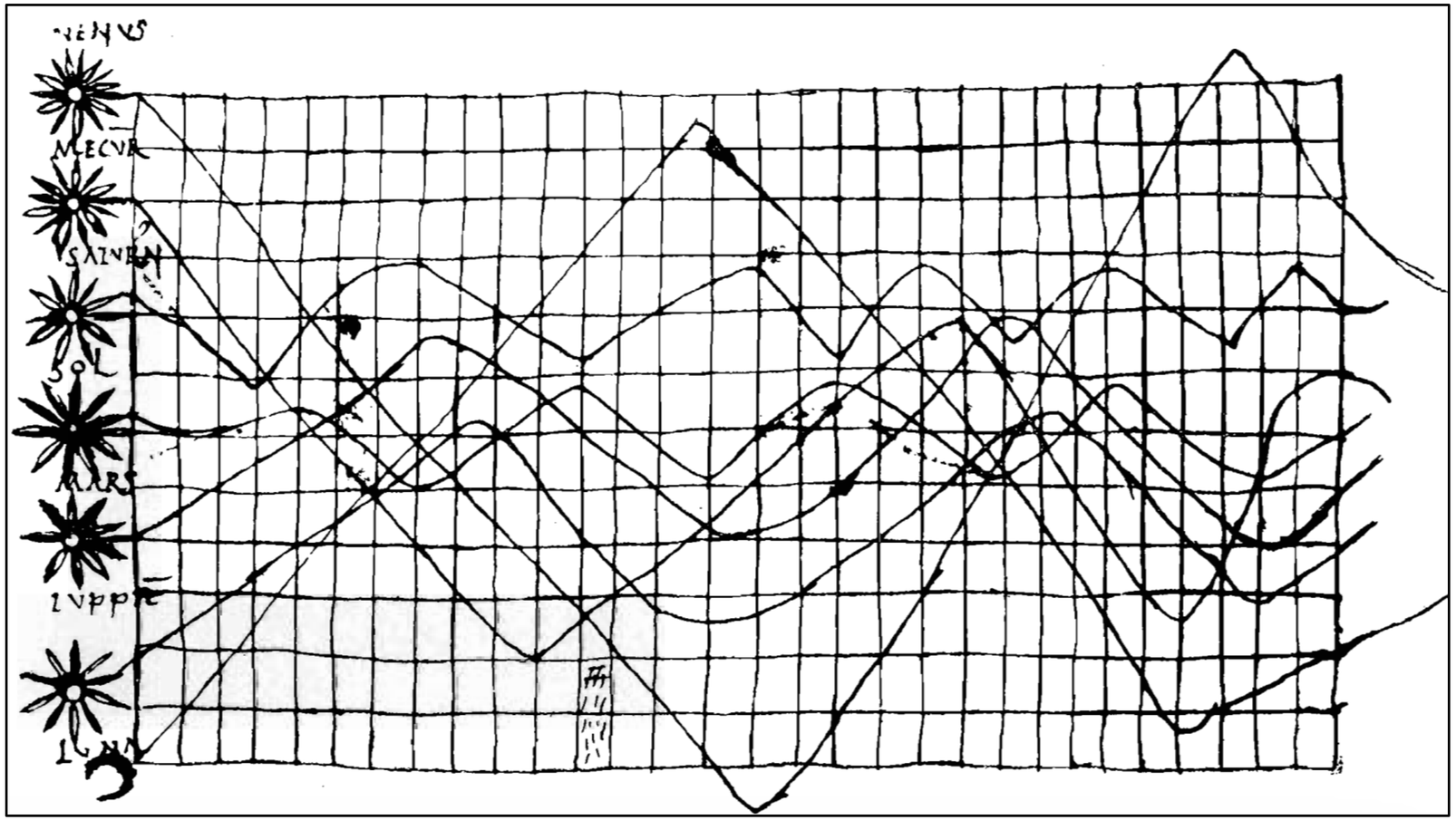
 # Part of the tidyverse
 # Core Packages

```
ggplot(data) +
  geom_*(aes(x, y),
         stat,
         position) +
  coordinates() +
  scale_*()+
  facet_*() +
  theme()
```


Tidyverse Package Graphic. Accessed November 15th, 2024.

The Layered Grammar of Graphics

Yale SCHOOL OF PUBLIC HEALTH
Data Science and Data Equity

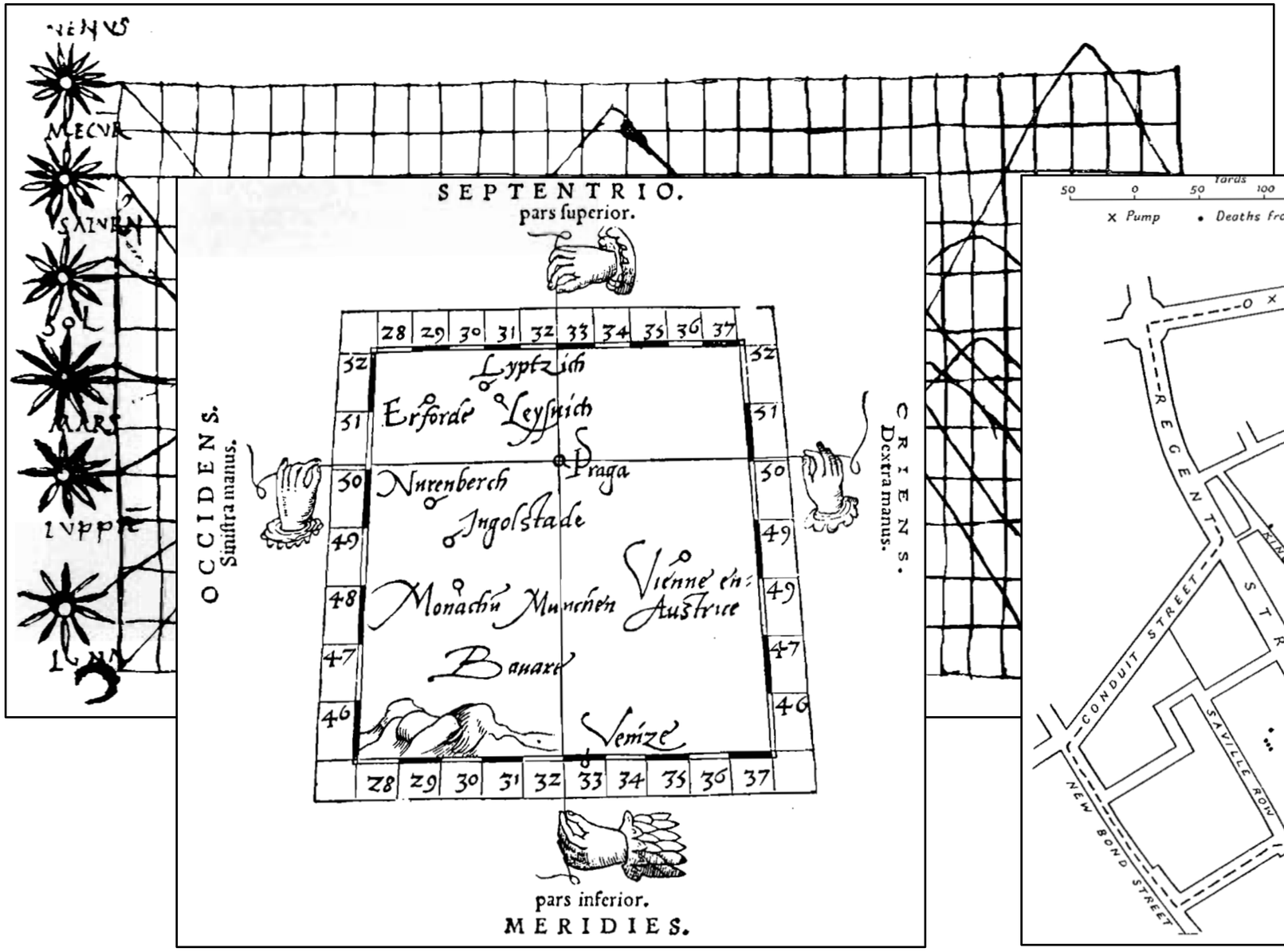


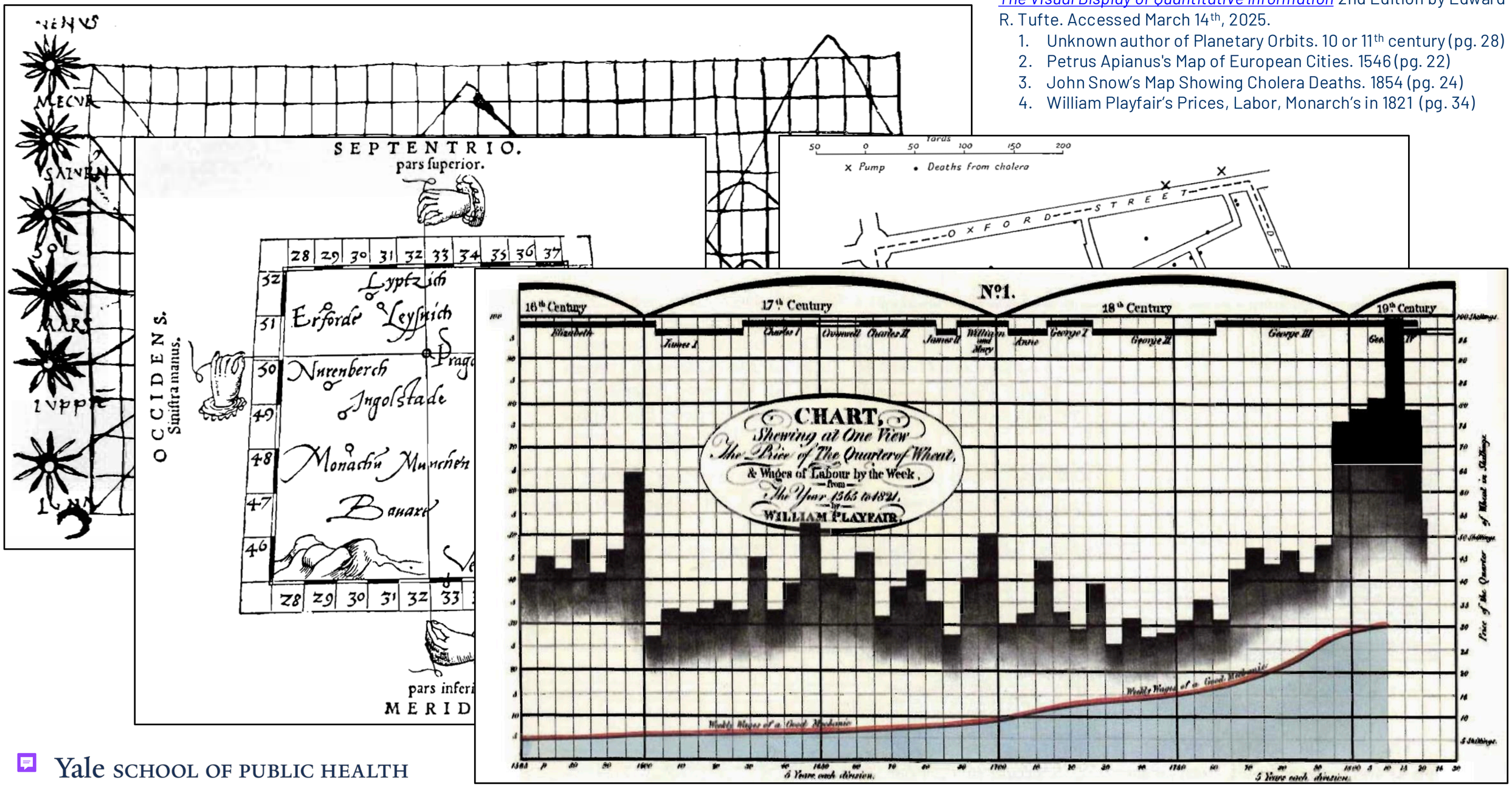
The Visual Display of Quantitative Information 2nd Edition by Edward R. Tufte. Accessed March 14th, 2025.

1. Unknown author of Planetary Orbits. 10 or 11th century (pg. 28)
2. Petrus Apianus's Map of European Cities. 1546 (pg. 22)
3. John Snow's Map Showing Cholera Deaths. 1854 (pg. 24)
4. William Playfair's Prices, Labor, Monarch's in 1821 (pg. 34)

The Visual Display of Quantitative Information 2nd Edition by Edward R. Tufte. Accessed March 14th, 2025.

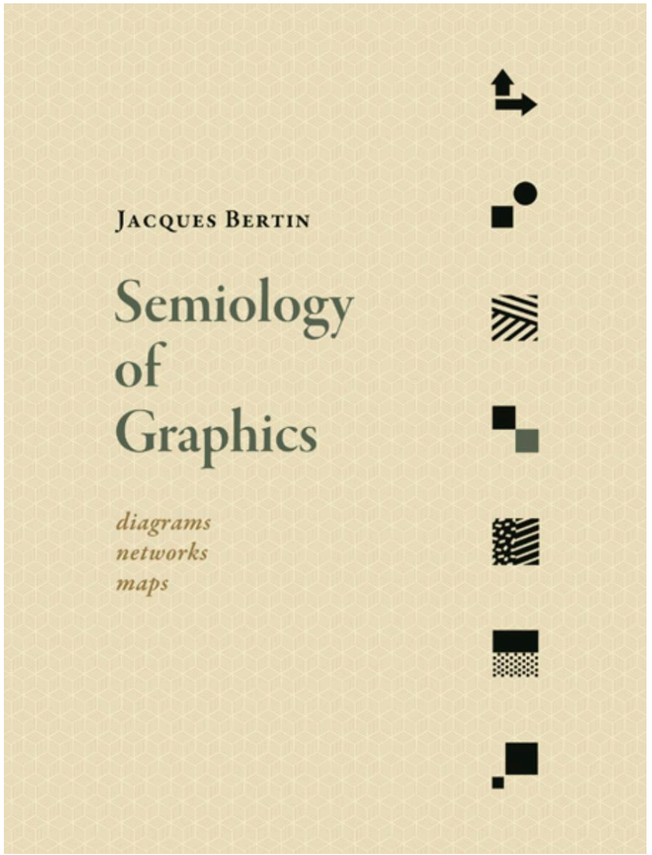
1. Unknown author of Planetary Orbits. 10 or 11th century (pg. 28)
2. Petrus Apianus's Map of European Cities. 1546 (pg. 22)
3. John Snow's Map Showing Cholera Deaths. 1854 (pg. 24)
4. William Playfair's Prices, Labor, Monarch's in 1821 (pg. 34)

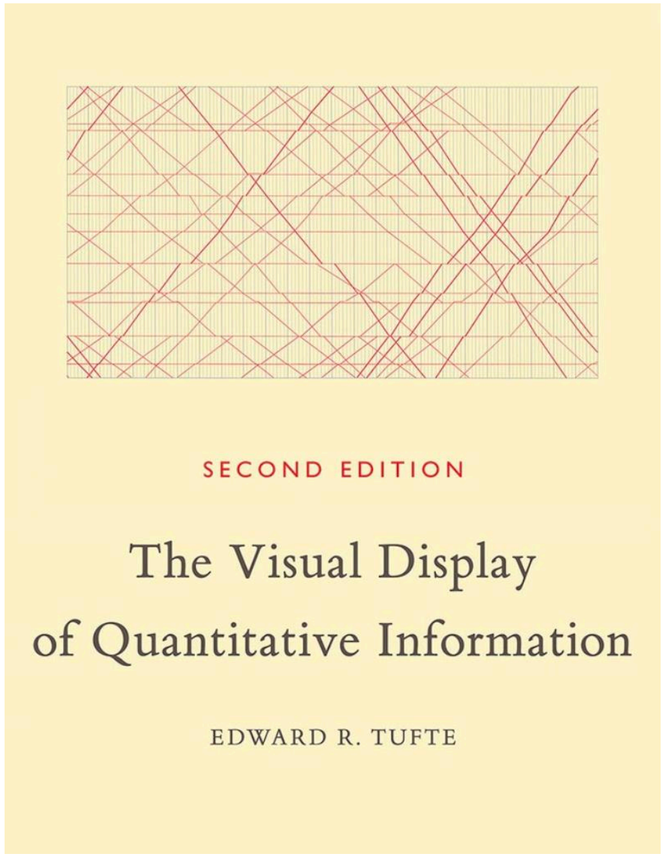


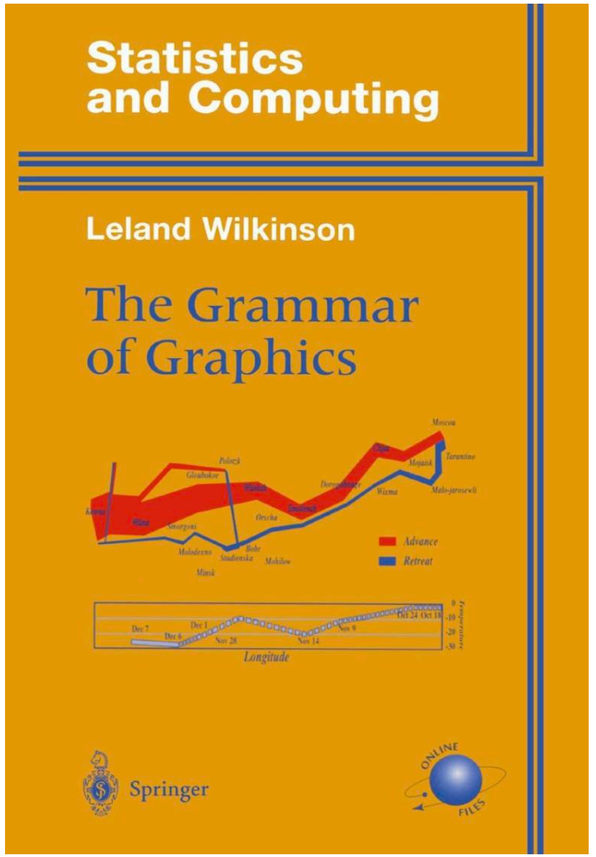


The Visual Display of Quantitative Information 2nd Edition by Edward R. Tufte. Accessed March 14th, 2025.

1. Unknown author of Planetary Orbits. 10 or 11th century (pg. 28)
2. Petrus Apianus's Map of European Cities. 1546 (pg. 22)
3. John Snow's Map Showing Cholera Deaths. 1854 (pg. 24)
4. William Playfair's Prices, Labor, Monarch's in 1821 (pg. 34)







1967
Principles

1983
Clarity

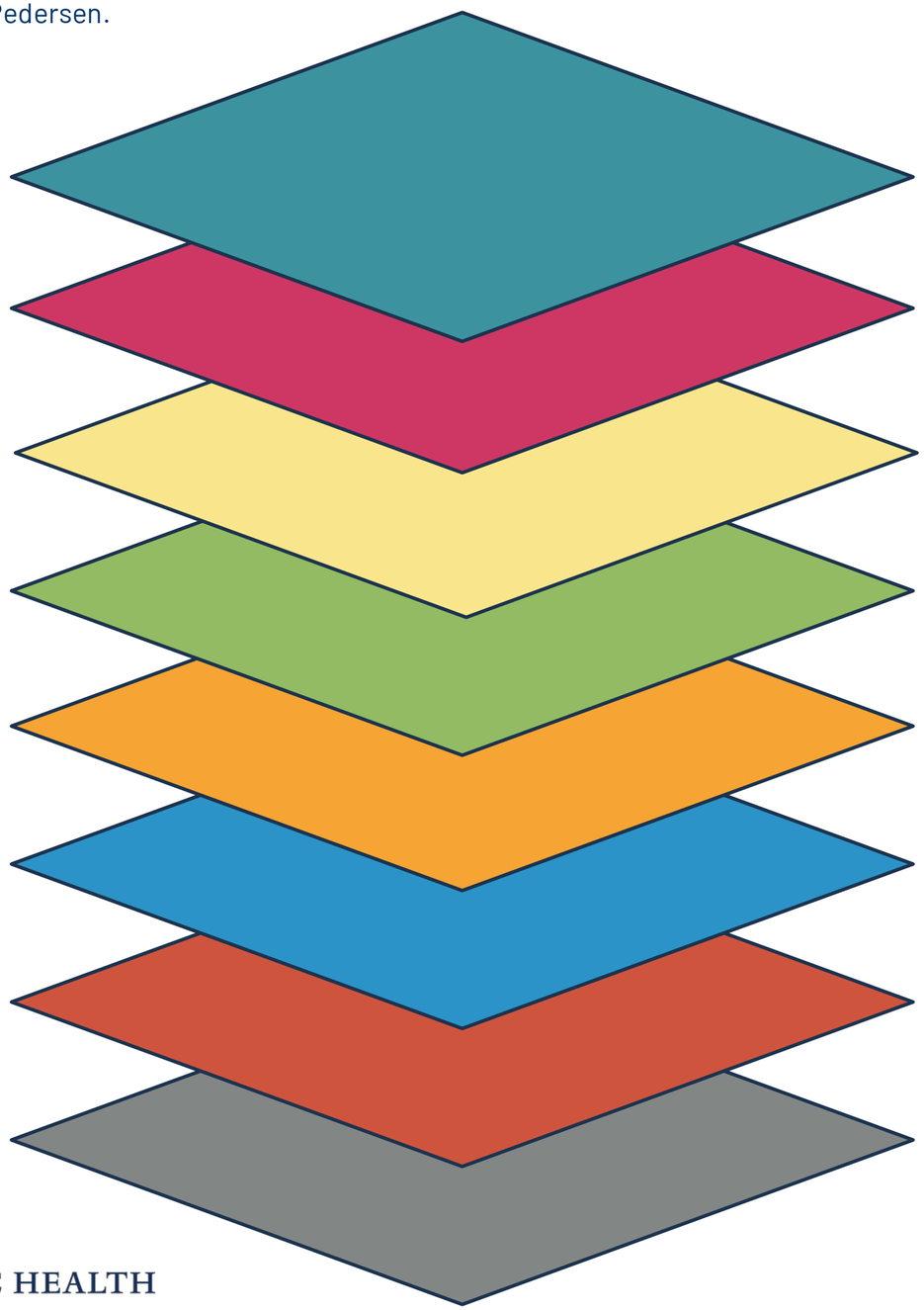
1999
Formalize

Semiology of Graphics: Diagrams, Networks, Maps by Jacques Bertin. Accessed March 2nd, 2025.

The Visual Display of Quantitative Information 2nd Edition by Edward R. Tufte. Accessed March 2nd, 2025.

The Grammar of Graphics 1st Edition by Leland Wilkinson. Accessed March 13th, 2025.

ggplot2 workshop part 1 by Thomas Lin Pedersen. Accessed March 15th, 2025.



Theme

Coordinates

Facets

Scales

Statistics

Geometries

Mapping

Data

“… good grammar is just the first step to creating a good sentence.” - Hadley Wickham

- *The Visual Display of Quantitative Information* by Professor Edward Tufte
- *Ten guidelines for effective data visualization in scientific publications* by Christa Kelleher and Thorsten Wagener
- Recommendations from Professor Wilkinson’s *The Grammar of Graphics*: William Cleveland, 1985, 1995 and Edward Tufte, 1990, and 1997

Crafting a Graphic: A Layer-Wise Walkthrough

Yale SCHOOL OF PUBLIC HEALTH
Data Science and Data Equity

Variables

Country	Year	Cases	Pop
AFG	1999	745	20 M
AFG	2000	2666	20.5 M
Brazil	1999	37737	172 M
Brazil	2000	80488	174.5 M
China	1999	212258	1,272 M
China	2000	213766	1,280 M

Observations

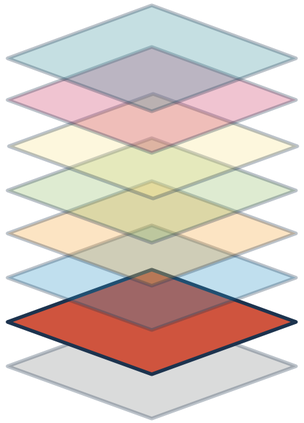
Country	Year	Cases	Pop
AFG	1999	745	20 M
AFG	2000	2666	20.5 M
Brazil	1999	37737	172 M
Brazil	2000	80488	174.5 M
China	1999	212258	1,272 M
China	2000	213766	1,280 M

Values

Country	Year	Cases	Pop
AFG	1999	745	20 M
AFG	2000	2666	20.5 M
Brazil	1999	37737	172 M
Brazil	2000	80488	174.5 M
China	1999	212258	1,272 M
China	2000	213766	1,280 M

Region	Season	Week Observed	MMWRweek	Level	Positives Detected	Spline	Kernel	Crude Rate
Georgia	2018-19	2018-11-10	19	18-49 Years	3	8.68	12.51	0.1
Georgia	2019-20	2020-04-04	40	Male	3	0.95	0.81	0.1
Michigan	2024-25	2024-07-27	4	Female	0	0.00	0.00	0.0
Oregon	2021-22	2021-08-07	5	American Indian or Alaska Native	0	2.49	0.00	0.0
Oregon	2019-20	2019-12-28	26	White	25	46.73	49.54	7.8
Utah	2021-22	2022-02-05	31	Hispanic or Latino	6	8.42	7.90	2.6

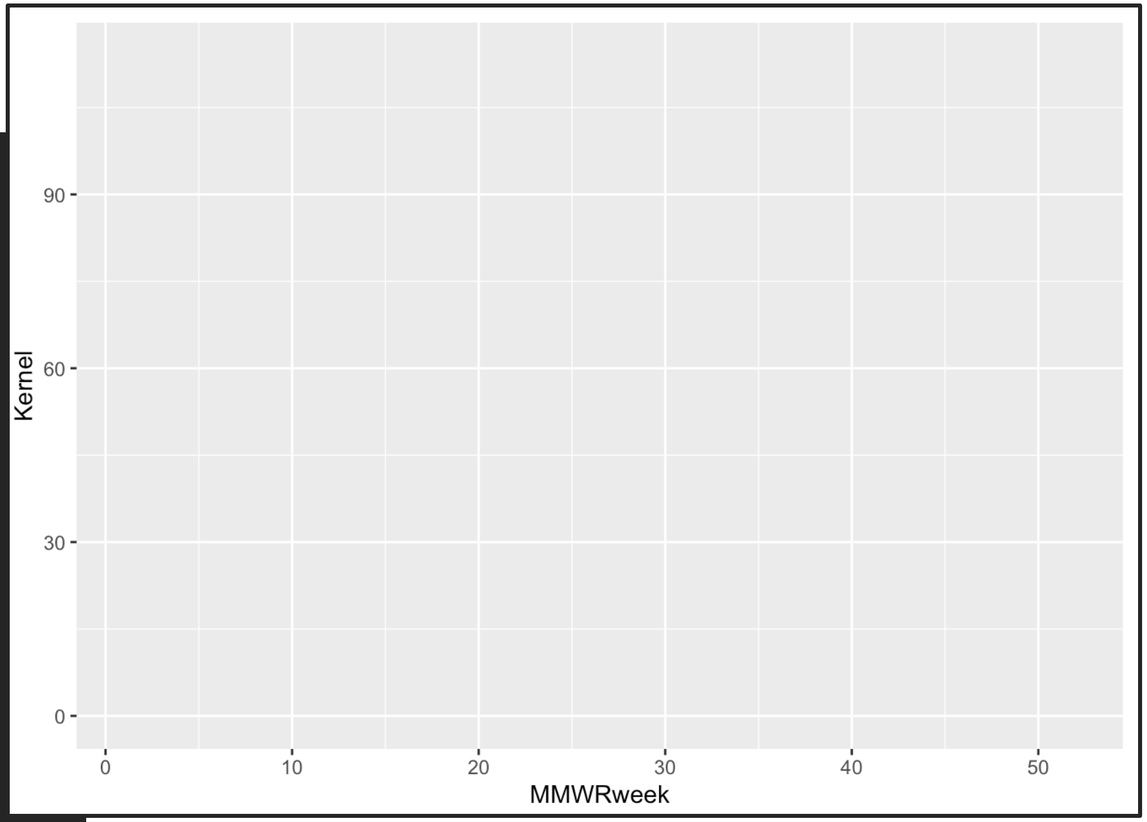
R for Data Science (2e) - Introduction Figure 5.1 by Hadley Wickham, Mine Çetinkaya-Rundel, and Garrett Grolemund. Accessed March 16th, 2025.

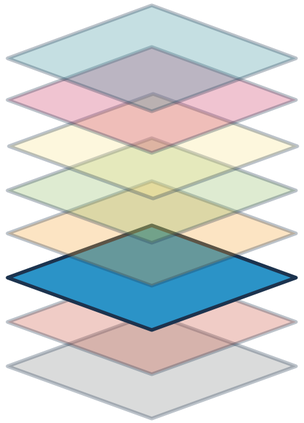


Mapping tells the function which variables get used for which aesthetic feature.

```
ggolot(data, aes(x, y, group)) +
  [geom_*|stat_*](aes(color, fill, [size|linewidth],
                      [shape|linetype]))
```

```
ggplot(data = df,
       aes(x = MMWRweek, y = Kernel))
```

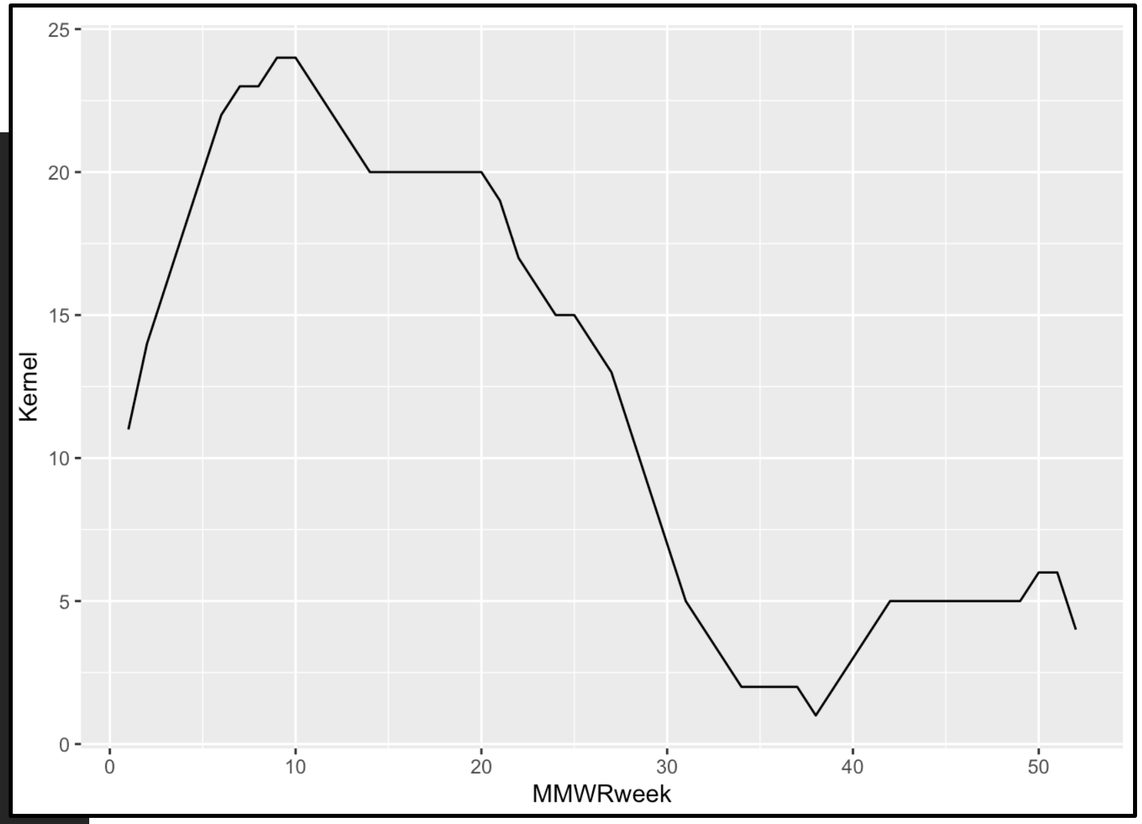




Geometry functions are the engines that transform data mappings into the desired plot type.

```
ggolot() +
    geom_*(data, aes(x, y, group), position, statistics)
```

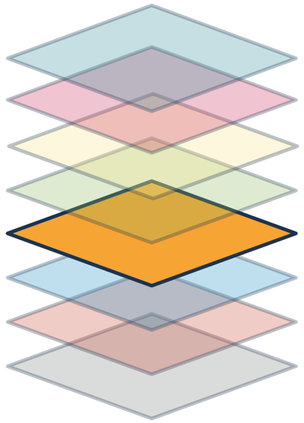
```
df |>
  ggplot() +
    geom_line(aes(x = MMWRweek,
                  y = Kernel))
```



Discussion:

We can choose to place the aesthetics and data in either `ggplot()` or the `geom_*()` layer.

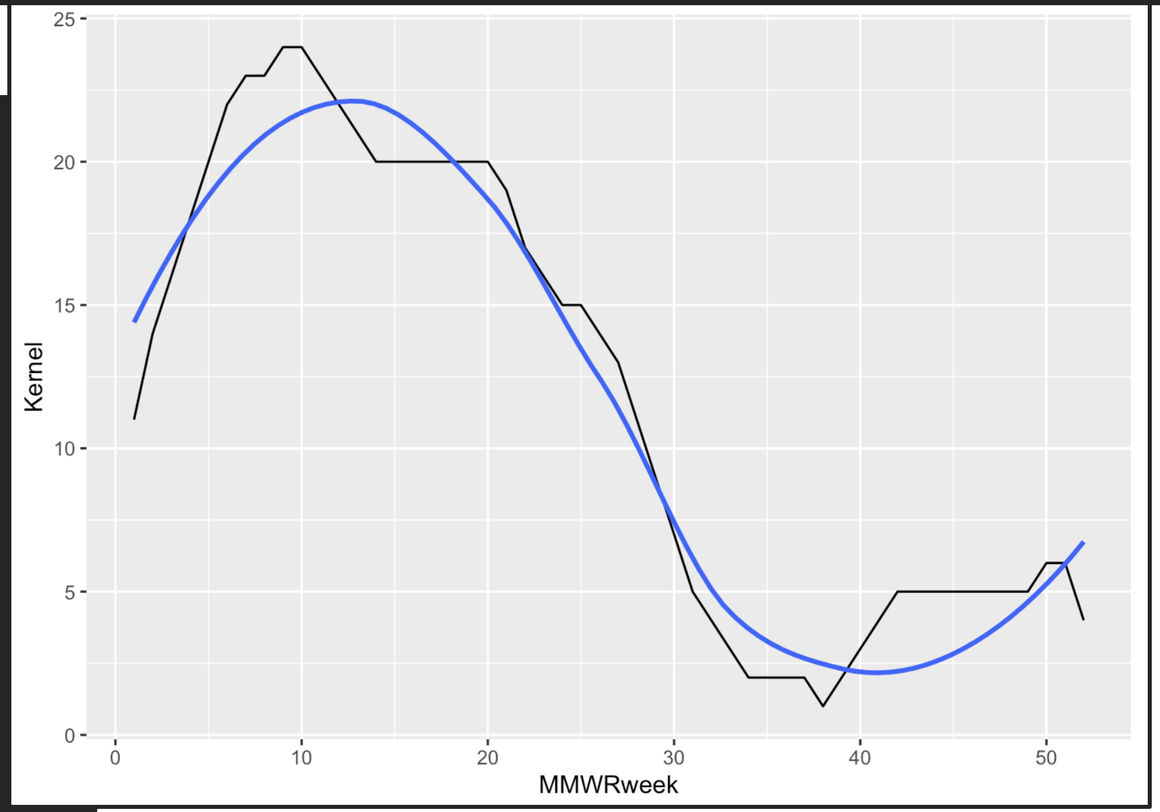
Can you think of reasons why you might choose one over the other?

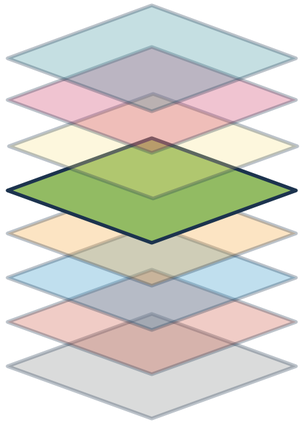


Statistics layers do the statistical calculations used in geometric engines. Many are interchangeable with a comparable geometry.

```
ggolot() +
    stat_*(data, aes(x, y, group), geometry, parameters)
```

```
df |>
  ggplot() +
    geom_line(aes(x = MMWRweek,
                  y = Kernel)) +
    stat_smooth(aes(x = MMWRweek,
            y = `Scaled Positives`),
                geom = "smooth",
                method = "loess",
                se = FALSE)
```

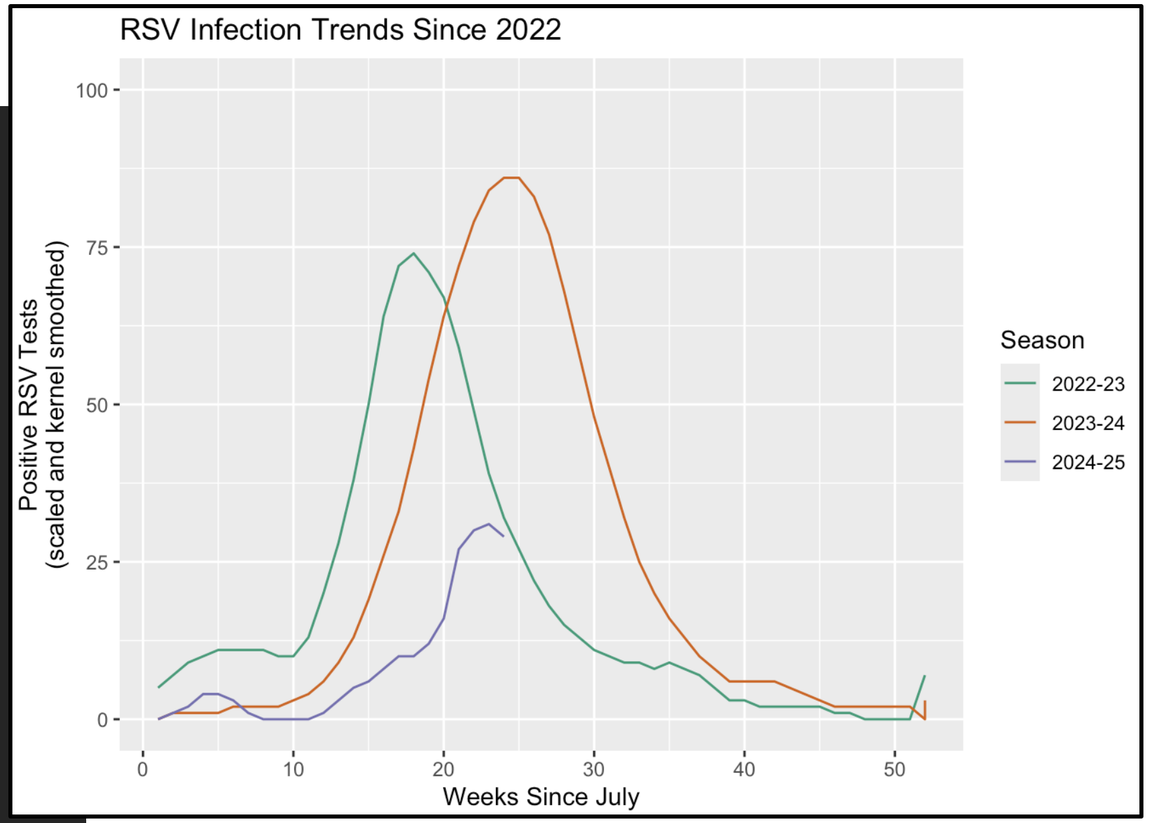




Scales interpret aesthetic mapping into plottable values and offer the most customization options.

```
ggolot(data, aes(x, y, group)) + [geom_*|stat_*]() +
   scale_*(parameters)
```

```
df |>
  ggplot(aes(x = MMWRweek, y = Kernel,
             color = Season)) +
    geom_line() +
    labs(title, x, y) +
    ylim(0, 100) +
    scale_color_brewer(type = "qual",
                       palette = "Dark2")
```



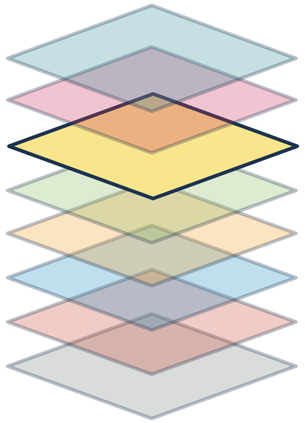
Discussion:

We see that associating a variable with `aes(color)` also groups by that variable.

a) What happens if we only group the outcome?

b) Does adding a `scale_color_*()` help?

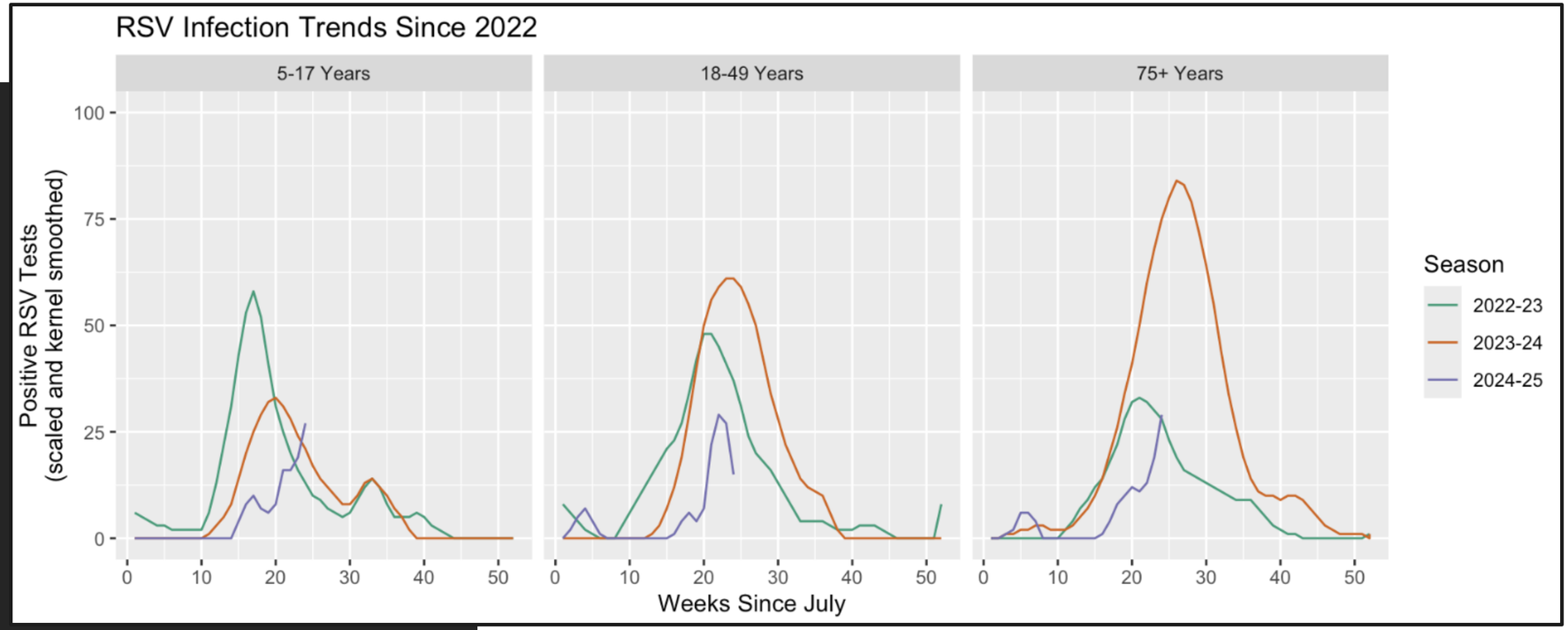
c) Why do we get the result we do?

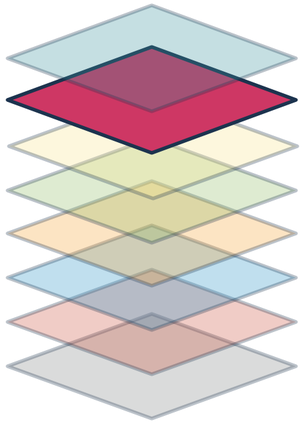


Facets divide the same plot across subgroups of a discrete variable.

```
ggolot(data, aes(x, y, group)) + [geom_*|stat_*]() +
   facet_*(discrete_variable)
```

```
df |>
  ggplot(aes(…)) +
    geom_line() +
    …
    facet_grid(~Level)
```

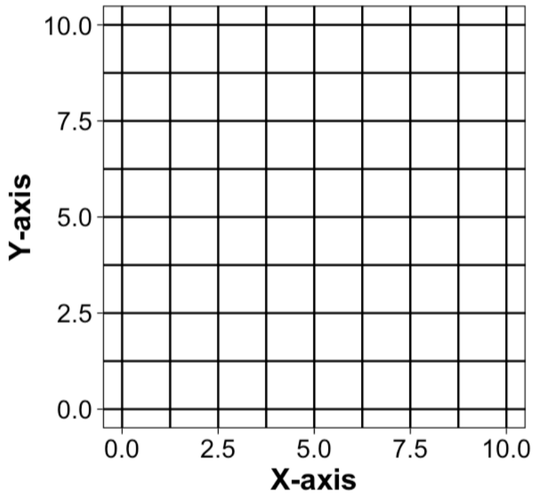




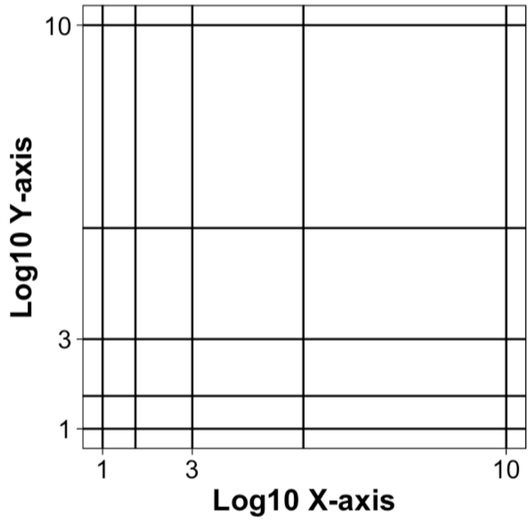
Coordinates plot the values prepared by the preceding layers in a specified coordinate system.

```
ggolot(data, aes(x, y, group)) + [geom_*|stat_*]() +
    coord_*()
```

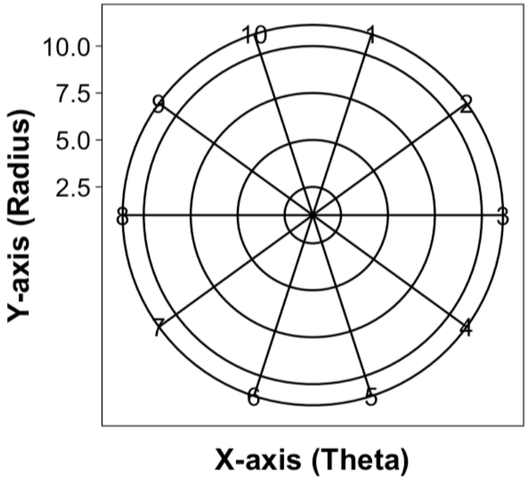
Cartesian



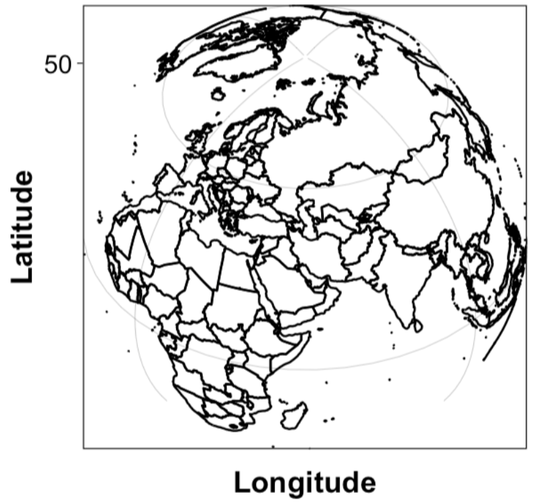
Transformed



Polar



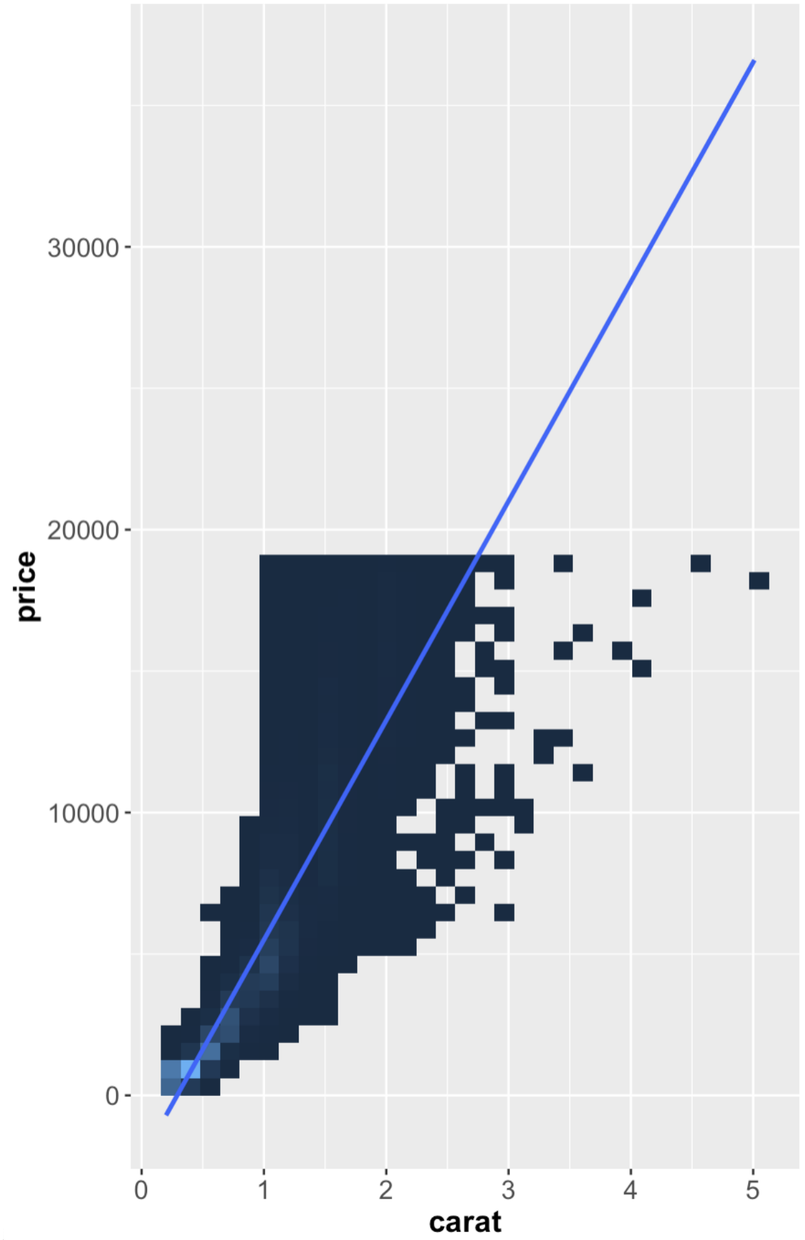
Map

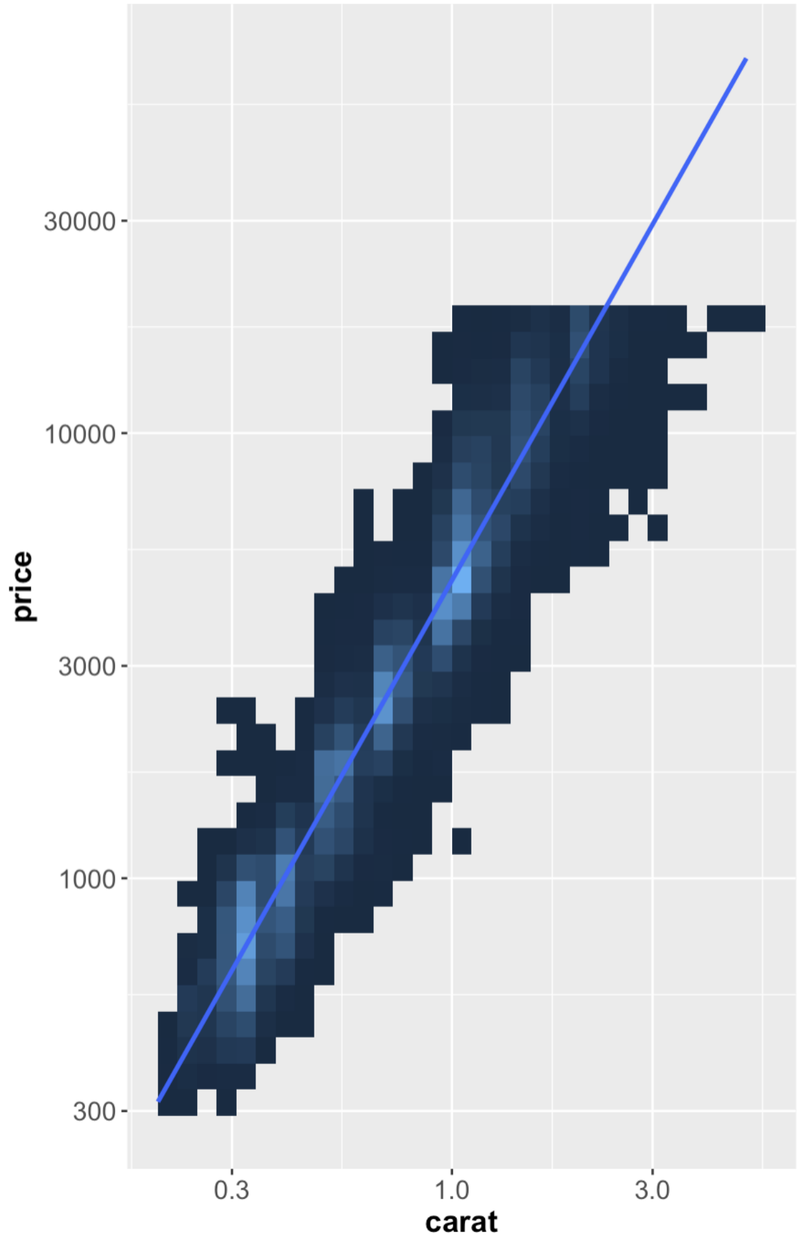


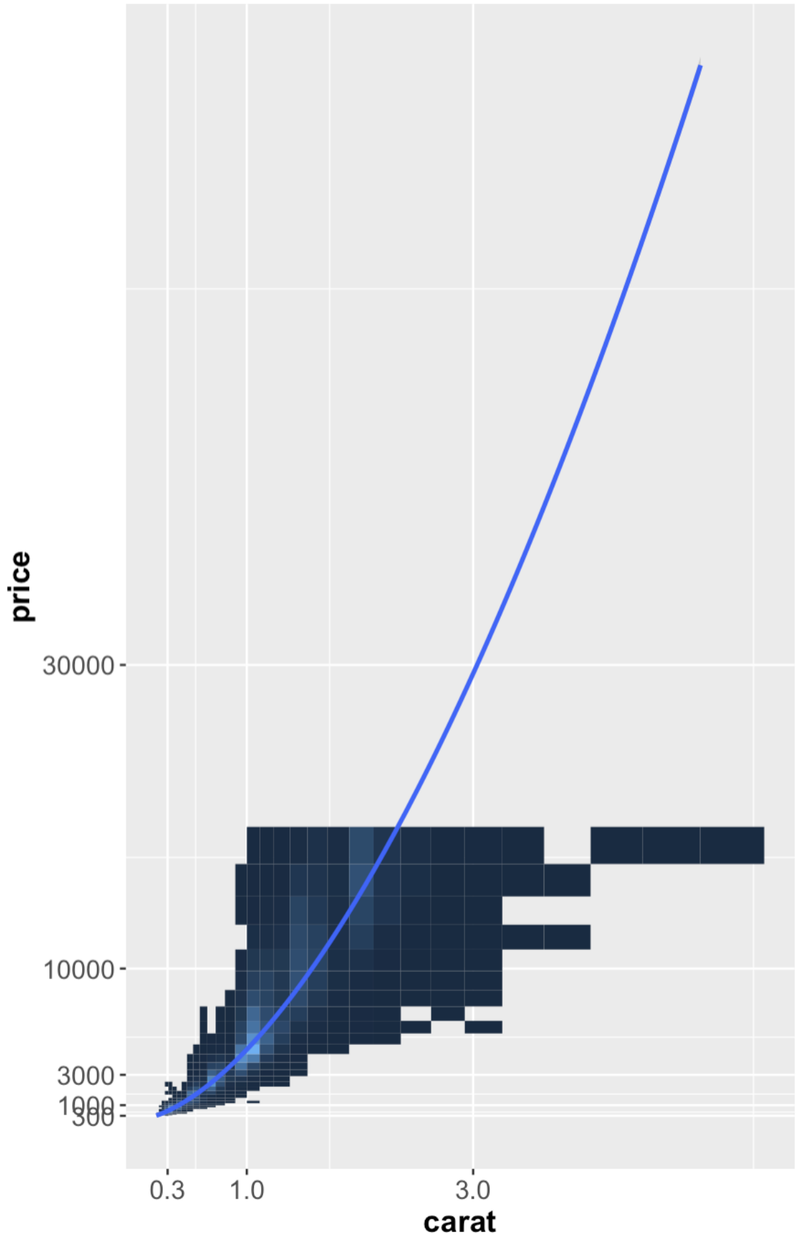
Based on *Data Visualisation: From Theory to Practice Figure 5.5* by James Baglin. Accessed March 22nd, 2025.

Discussion:

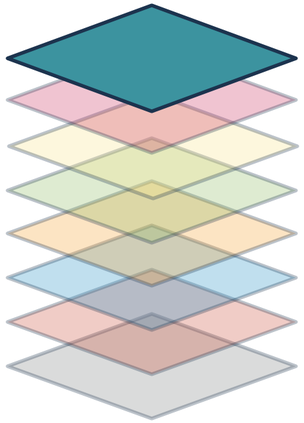
If we apply each option individually, do they all plot the same way? If not, what do you think is happening and how would you fix the code?







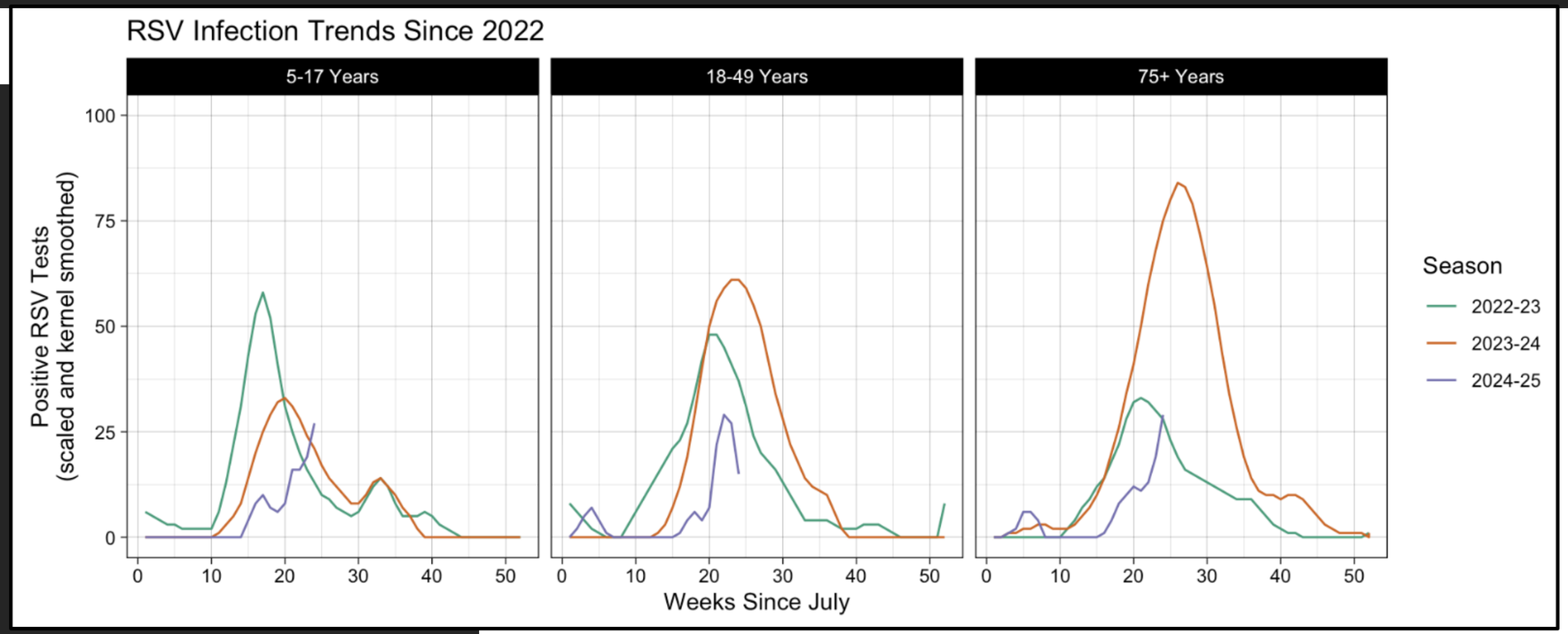
15.2.1 Transformations with coord_trans() from *ggplot2: Elegant Graphics for Data* by Hadley Wickham, Danielle Navarro, and Thomas Lin Pedersen. Accessed March 24th, 2025.



The Theme layer allows users to change non-data aspects of their plot to improve its visual appeal and styling.

```
ggolot(data, aes(x, y, group)) + [geom_*|stat_*]() +
    theme_*()
```

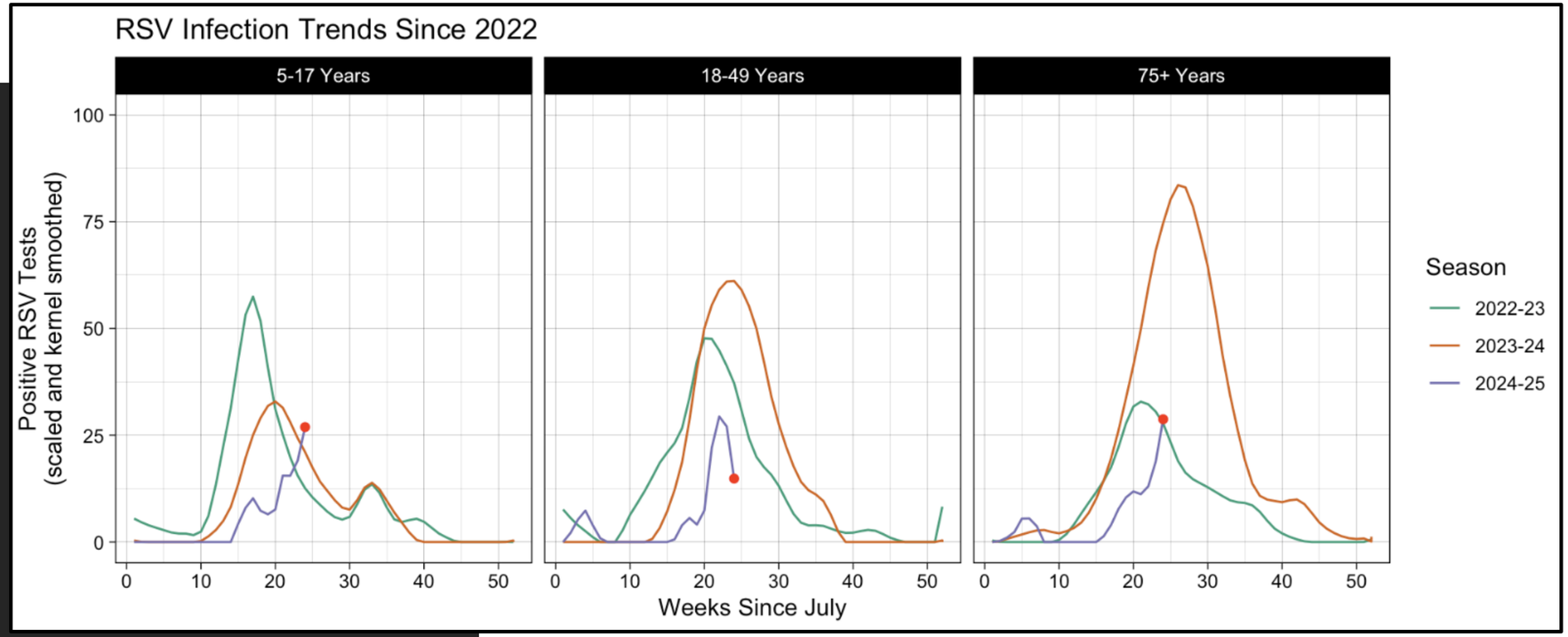
```
df |>
  ggplot(aes(…)) +
    geom_line() +
    …
    theme_linedraw()
```



Overlay layers onto an existing plot to achieve more complex visualizations.

- Each layer is an independent object, enabling adaptive and functional programming.
- Data and aesthetic mappings can be inherited from `ggplot()`.

```
df |>
  ggplot(aes(…)) +
    geom_line() +
    geom_point(df2)
    …
```



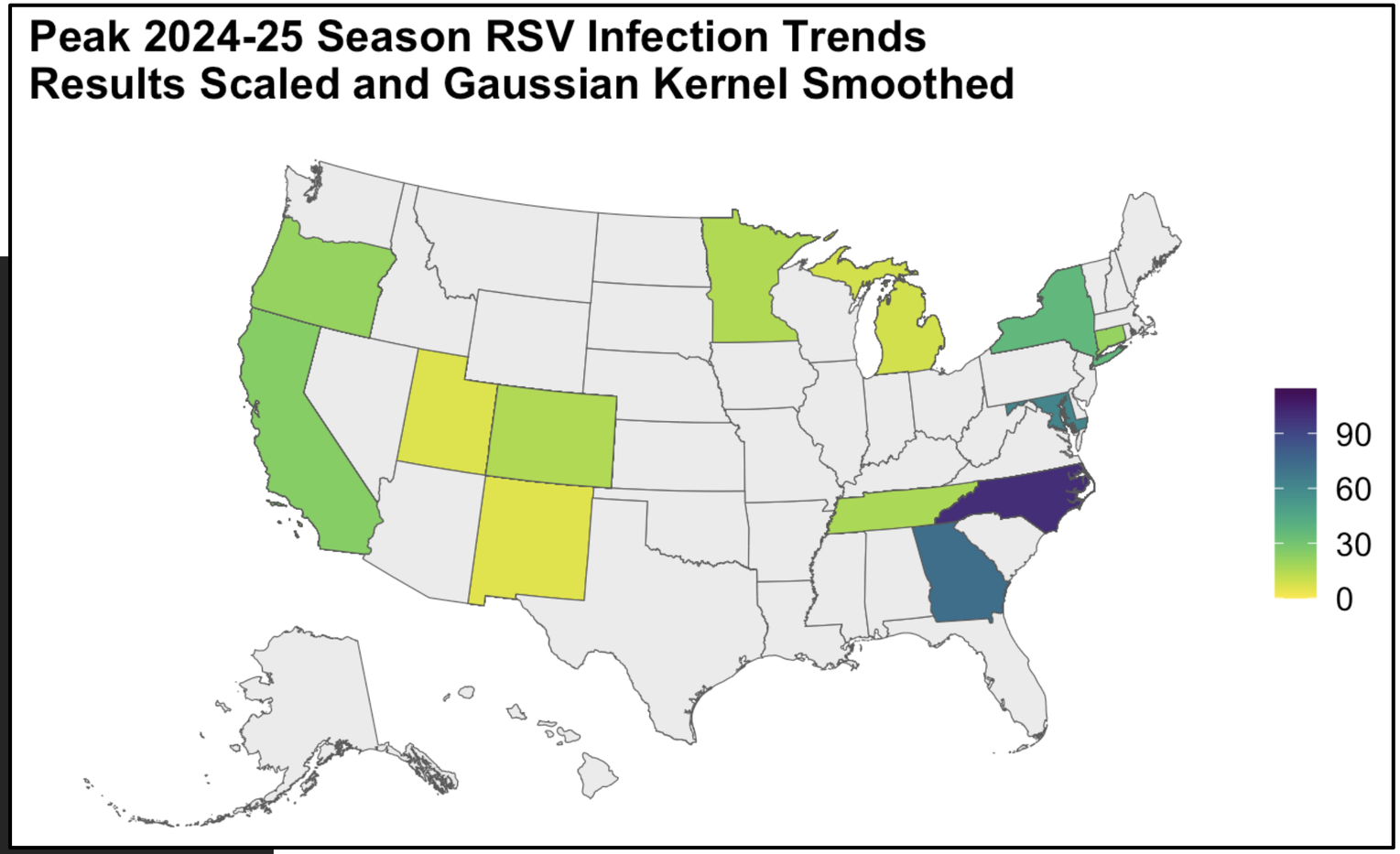
Advanced use of ggplot2

Going beyond the basics

- Community managed "ggplot2 extensions gallery"
- "R Graph Gallery" – fucuses on applications with ggplot2
- Additional topics from *ggplot2: Elegant Graphics for Data Analysis (3e)*: "Chapter 5 Statistical summaries", "Chapter 7 Networks", "Chapter 8 Annotations", "Chapter 18 Programming with ggplo2", and "Chapter 20 Extending ggplot2"
- "Extending ggplot2" vignette
- "Using ggplot2 in packages" vignette

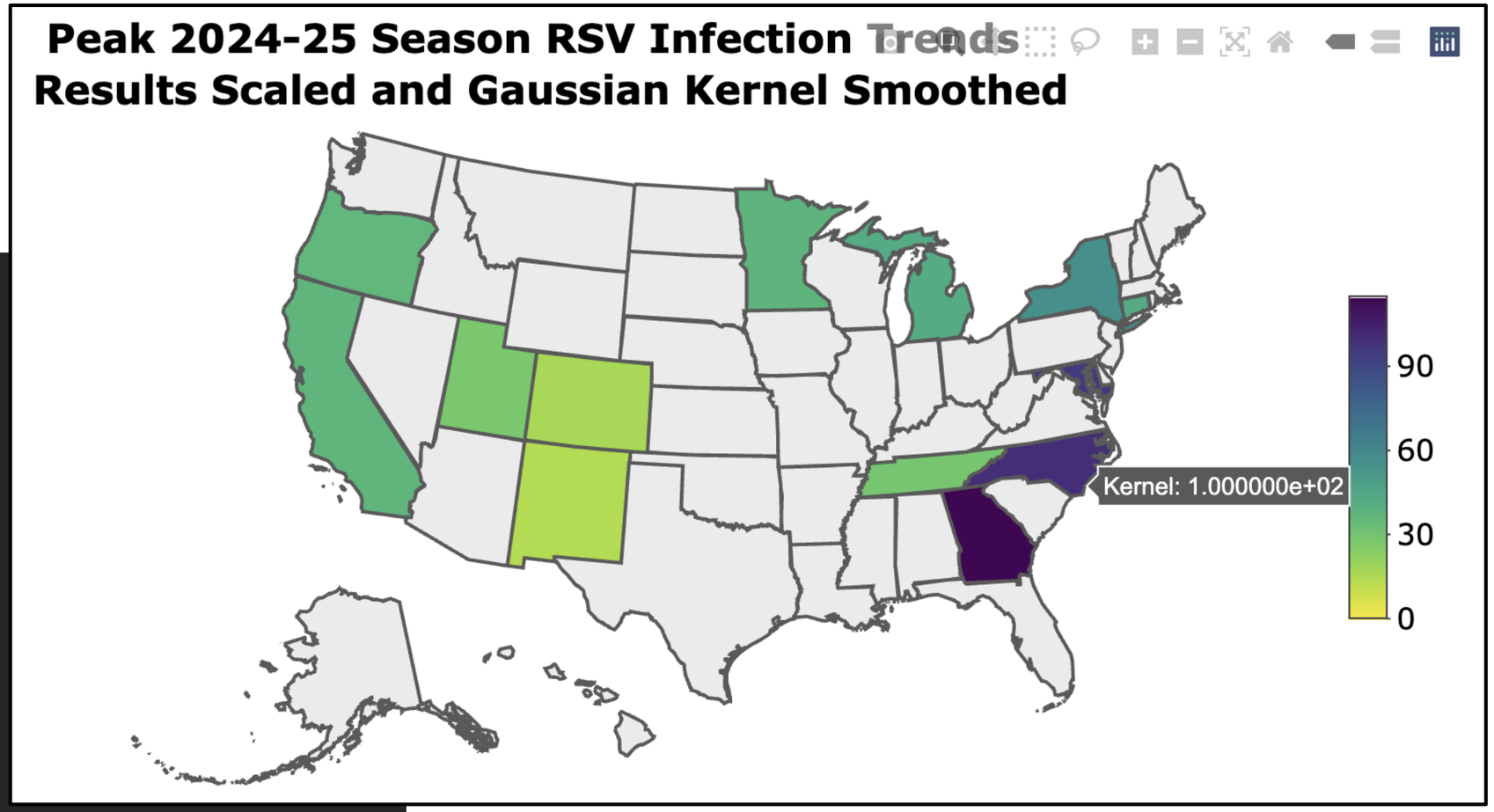
Use vector data (sf) to tell `geom_sf()` where on a map the data should be filled.

```
df |>
  ggplot(aes(fill = Kernel)) +
    geom_sf() +
    …
    coord_sf() +
    theme_map()
```



`plotly` allows you to quickly render an HTML-based, interactive `ggplot`.

```
ggplotly(plot_name)
```



Yale's AI Clarity

1. Take a screenshot of one of the plots from the workshop or one of your own.
2. Then, navigate to https://ai-chat.yale.edu/signin-oidc
3. Upload the screenshot into the chat and ask, "Suggest R code that could make this plot in ggplot2."

Discussion:

How did the AI chatbot do?

ysph.yale.edu
sph.yale.edu/dsde

@YaleSPH

Public Health Data Science and Data Equity
Yale School of Public Health
60 College Street, New Haven, CT 06510

Yale SCHOOL OF PUBLIC HEALTH

Appendix

Glossary

Grammer of Graphics	Definition for the distinct elements that make up all graphical representations of relational data in tabular form. First created by Professor Leland Wilkinson in 1999.
Layered Grammer of Graphics	The modified version of Grammar of Graphics that stores each element as an independent object. These objects get added together to generate a comprehensive plot.
Layer: Data	A "tidy" data frame with the necessary columns of information to generate the plot you intend.
Layer: Mapping	Assigns variables in the data frame to aesthetic features on the plot (i.e. shape, color, etc.).

Glossary

Layer: Geometry Engines (composite of operations) that process the data into a defined plot type (i.e. boxplot, histogram, line, etc.).

Layer: Statistics Statistical transformations that generate a geometry. Sometimes interchangeable with geom_*() objects.

Layer: Scales Interpret aesthetic Mappings into plottable values (i.e. axis scaling, color scaling, etc.).

Layer: Facets Spreads out the same plot into new subplots, each showing distinct instantiations of a variable.

Glossary

Layer: Coordinates	Defines the coordinate plane of the plot: i.e. Cartesian, polar, transformed, or a map projection.
Layer: Theme	Controls for the non-data elements of the plot.
Map Projection	Interpreting the curved surface of the earth into a flat plane for 2D plotting.
Simple Feature (SF)	Standard vector data produced by the Open Geospatial Consortium (OGC) that translates projection data into plottable polygons.

References

Slide 1

1. E. Tufte, The Visual Display of Quantitative Information, Second Edition. Cheshire, CT: Graphics Press, LLC, 2001. Accessed: Mar. 13, 2025. [Online]. Available: https://www.edwardtufte.com/book/the-visual-display-of-quantitative-information

Slide 4

1. Hadley Wickham et al., "Function references," ggplot2 Documentation. Accessed: Mar. 09, 2025. [Online]. Available: https://ggplot2.tidyverse.org/reference/index.html
2. T. L. Pedersen, "ggplot2 workshop part 1," YouTube. Accessed: Mar. 12, 2025. [Online]. Available: https://www.youtube.com/watch?v=h29g21z0a68
3. T. L. Pedersen, "ggplot2 workshop part 2," YouTube. Accessed: Mar. 12, 2025. [Online]. Available: https://www.youtube.com/watch?v=0m4yywqNPVY

References

Slide 4 continued

4. Posit Contributors, Data visualization with ggplot2: Cheat Sheet. Springer-Verlag, 2024. Accessed: Mar. 20, 2025. [Online]. Available: https://rstudio.github.io/cheatsheets/html/data-visualization.html

5. H. Wickham et al., ggplot2: Elegant Graphics for Data Analysis Documentation. New York: Springer-Verlag, 2016. Accessed: Mar. 12, 2025. [Online]. Available: https://ggplot2.tidyverse.org/

6. H. Wickham, M. Çetinkaya-Rundel, and G. Grolemund, R for Data Science (2e), Second. O'Reilly Media, 2023. Accessed: Mar. 12, 2025. [Online]. Available: https://r4ds.hadley.nz/

7. H. Wickham, D. Navarro, and T. L. Pedersen, ggplot2: Elegant Graphics for Data Analysis (3e), Third. Springer, 2010. Accessed: Mar. 20, 2025. [Online]. Available: https://ggplot2-book.org/

Slide 7

1. "ggplot2," Wikipedia. Accessed: Mar. 13, 2025. [Online]. Available: https://en.wikipedia.org/wiki/Ggplot2

References

Slide 7 continued

2. H. Wickham, M. Çetinkaya-Rundel, and G. Grolemund, R for Data Science (2e), Second. O'Reilly Media, 2023. Accessed: Mar. 12, 2025. [Online]. Available: https://r4ds.hadley.nz/

Slide 8

1. H. Wickham et al., ggplot2: Elegant Graphics for Data Analysis Documentation. New York: Springer-Verlag, 2016. Accessed: Mar. 12, 2025. [Online]. Available: https://ggplot2.tidyverse.org/

2. H. Wickham, "Tidyverse." Accessed: Nov. 14, 2024. [Online]. Available: https://www.tidyverse.org/

Slide 10

1. E. Tufte, The Visual Display of Quantitative Information, Second Edition. Cheshire, CT: Graphics Press, LLC, 2001. Accessed: Mar. 13, 2025. [Online]. Available: https://www.edwardtufte.com/book/the-visual-display-of-quantitative-information/

References

Slide 11

1. J. Bertin, Semiology of Graphics: Diagrams, Networks, Maps. Esri Press, 1967. Accessed: Mar. 13, 2025. [Online]. Available: https://www.esri.com/en-us/esri-press/browse/semiology-of-graphics-diagrams-networks-maps
2. E. Meeks, "3rd Wave Data Visualization," Medium. Accessed: Mar. 12, 2025. [Online]. Available: https://medium.com/nightingale/3rd-wave-data-visualization-824c5dc84967
3. J. M. Norman, "Chomsky's Hierarchy of Syntactic Forms : History of Information." Accessed: Mar. 14, 2025. [Online]. Available: https://www.historyofinformation.com/detail.php?id=745
4. E. Tufte, The Visual Display of Quantitative Information, Second Edition. Cheshire, CT: Graphics Press, LLC, 2001. Accessed: Mar. 13, 2025. [Online]. Available: https://www.edwardtufte.com/book/the-visual-display-of-quantitative-information/
5. L. Wilkinson, The Grammar of Graphics. Springer-Verlag, 1999. doi: 10.1007/0-387-28695-0.

References

Slide 11 continued

6. "Edward Tufte," Wikipedia. Accessed: Mar. 13, 2025. [Online]. Available: https://en.wikipedia.org/wiki/Edward_Tufte

7. "Leland Wilkinson," Wikipedia. Accessed: Mar. 14, 2025. [Online]. Available: https://en.wikipedia.org/wiki/Leland_Wilkinson

Slide 12

1. T. L. Pedersen, "ggplot2 workshop part 1," YouTube. Accessed: Mar. 12, 2025. [Online]. Available: https://www.youtube.com/watch?v=h29g21z0a68

2. H. Wickham, "A Layered grammar of graphics," Journal of Computational and Graphical Statistics, vol. 19, no. 1, pp. 3–28, Mar. 2010, doi: 10.1198/JCGS.2009.07098.

3. "ggplot2," Wikipedia. Accessed: Mar. 13, 2025. [Online]. Available: https://en.wikipedia.org/wiki/Ggplot2

References

Slide 13

1. W. S. Cleveland, Visualizing Data, First Edition. Summit, NJ: Hobart Press, 1995. Accessed: Mar. 14, 2025. [Online]. Available: https://www.science.org
2. W. Cleveland, The Elements of Graphing Data, First Edition., no. 814. Belmont, CA: Wadsworth Advanced Books and Software, 1985. Accessed: Mar. 14, 2025. [Online]. Available: https://www.amazon.com/exec/obidos/ASIN/0534037305/acmorg-20
3. C. Kelleher and T. Wagener, "Ten guidelines for effective data visualization in scientific publications," Environmental Modelling & Software, vol. 26, no. 6, pp. 822–827, Jun. 2011, doi: 10.1016/J.ENVSOFT.2010.12.006.
4. E. Tufte, Envisioning Information. Cheshire, CT: Graphics Press, LLC, 1990. Accessed: Mar. 14, 2025. [Online]. Available: https://www.edwardtufte.com/book/envisioning-information/

References

Slide 13 continued

5. E. Tufte, Visual Explanations: Images and Quantities, Evidence and Narrative. Cheshire, CT: Graphics Press, LLC, 1997. Accessed: Mar. 14, 2025. [Online]. Available: https://www.edwardtufte.com/book/visual-explanations-images-and-quantities-evidence-and-narrative/

6. E. Tufte, The Visual Display of Quantitative Information, First Edition. Cheshire, CT: Graphics Press, LLC, 1983. Accessed: Nov. 20, 2025. [Online]. Available: https://www.edwardtufte.com/book/the-visual-display-of-quantitative-information/

7. L. Wilkinson, The Grammar of Graphics. Springer-Verlag, 1999. doi: 10.1007/0-387-28695-0.

Slides 15-27

1. M. Freeman and J. Ross, "Chapter 13 The gglot2 Library | Technical Foundations of Informatics." Accessed: Mar. 20, 2025. [Online]. Available: https://info201.github.io/ggplot2.html

References

Slides 15-27 continued

2. Hadley Wickham et al., “Function references,” ggplot2 Documentation. Accessed: Mar. 09, 2025. [Online]. Available: https://ggplot2.tidyverse.org/reference/index.html

3. T. L. Pedersen, “ggplot2 workshop part 1,” YouTube. Accessed: Mar. 12, 2025. [Online]. Available: https://www.youtube.com/watch?v=h29g21z0a68

4. Posit Contributors, Data visualization with ggplot2: Cheat Sheet. Springer-Verlag, 2024. Accessed: Mar. 20, 2025. [Online]. Available: https://rstudio.github.io/cheatsheets/html/data-visualization.html

5. H. Wickham et al., ggplot2: Elegant Graphics for Data Analysis Documentation. New York: Springer-Verlag, 2016. Accessed: Mar. 12, 2025. [Online]. Available: https://ggplot2.tidyverse.org/

6. H. Wickham, M. Çetinkaya-Rundel, and G. Grolemund, R for Data Science (2e), Second. O’Reilly Media, 2023. Accessed: Mar. 12, 2025. [Online]. Available: https://r4ds.hadley.nz/

References

Slides 15-27 continued

7. H. Wickham, D. Navarro, and T. L. Pedersen, ggplot2: Elegant Graphics for Data Analysis (3e), Third. Springer, 2010. Accessed: Mar. 20, 2025. [Online]. Available: https://ggplot2-book.org/

Slide 16

1. H. Wickham et al., "Aesthetic specifications vignette," ggplot2 Documentation. Accessed: Mar. 22, 2025. [Online]. Available: https://ggplot2.tidyverse.org/articles/ggplot2-specs.html

Slide 23

1. J. Baglin, "Data Visualisation: From Theory to Practice." Accessed: Mar. 21, 2025. [Online]. Available: https://data-visualisation.stem.melbourne/grammar-and-vocabulary#coord

Slides 24-25

1. H. Wickham, D. Navarro, and T. L. Pedersen, 15 Coordinate systems – ggplot2: Elegant Graphics for Data Analysis (3e), Springer. Springer, 2010. Accessed: Nov. 22, 2025. [Online]. Available: https://ggplot2-book.org/coord.html#transformations-with-coord_trans

References

Slide 26

1. M. Pickens, “Learning to create custom themes in ggplot2,” RPubs. Accessed: Mar. 22, 2025. [Online]. Available: https://rpubs.com/mclaire19/ggplot2-custom-themes

Slide 29

1. D. Emaasit and Various, “ggplot2 extensions gallery.” Accessed: Mar. 23, 2025. [Online]. Available: https://exts.ggplot2.tidyverse.org/
2. H. Wickham, D. Navarro, and T. L. Pedersen, ggplot2: Elegant Graphics for Data Analysis (3e), Third. Springer, 2010. Accessed: Mar. 20, 2025. [Online]. Available: https://ggplot2-book.org/
3. H. Wickham et al., “Extending ggplot2 vignette,” ggplot2 Documentation. Accessed: Mar. 22, 2025. [Online]. Available: https://ggplot2.tidyverse.org/articles/extending-ggplot2.html
4. H. Wickham et al., “Using ggplot2 in packages vignette,” ggplot2 Documentation. Accessed: Mar. 23, 2025. [Online]. Available: https://ggplot2.tidyverse.org/articles/ggplot2-in-packages.html

References

Slide 29 continued

5. "The R Graph Gallery." Accessed: Mar. 23, 2025. [Online]. Available: https://r-graph-gallery.com/index.html

Slide 30

1. "The R Graph Gallery - Interactive charts." Accessed: Mar. 24, 2025. [Online]. Available: https://r-graph-gallery.com/interactive-charts.html
2. "Plotly r graphing library in R," plotly Documentation. Accessed: Mar. 24, 2025. [Online]. Available: https://plotly.com/r/
3. "Plotly graphing library for ggplot2 in ggplot2," plotly Documentation. Accessed: Mar. 24, 2025. [Online]. Available: https://plotly.com/ggplot2/
4. "Plotly Open Source Graphing Libraries," plotly Documentation. Accessed: Mar. 24, 2025. [Online]. Available: https://plotly.com/graphing-libraries/

References

Slide 31

1. U. C. Bureau, “Understanding Geographic Identifiers (GEOIDs),” Aug. 2024, Accessed: Mar. 23, 2025. [Online]. Available: https://www.census.gov/programs-surveys/geography/guidance/geo-identifiers.html
2. U. C. Bureau, “Cartographic Boundary Files,” Apr. 2024, Accessed: Mar. 23, 2025. [Online]. Available: https://www.census.gov/geographies/mapping-files/time-series/geo/cartographic-boundary.html
3. U. C. Bureau, “TIGER/Line Shapefiles,” Oct. 2024, Accessed: Mar. 23, 2025. [Online]. Available: https://www.census.gov/geographies/mapping-files/time-series/geo/tiger-line-file.html
4. J. Cheng, B. Schloerke, B. Karambelkar, and Y. Xie, “Create Interactive Web Maps with the JavaScript ‘Leaflet’ Library,” CRAN: Contributed Packages, vol. version 2.2.2, Mar. 2024, doi: 10.32614/CRAN.PACKAGE.LEAFLET.
5. Office of Intergovernmental and External Affairs (IEA), “HHS Regional Offices,” HHS.gov. Accessed: Mar. 23, 2025. [Online]. Available: https://www.hhs.gov/about/agencies/iea/regional-offices/index.html

References

Slide 31 continued

6. US Census Bureau, "Legal/Statistical Area Description (LSAD) Codes." Accessed: Mar. 23, 2025. [Online]. Available: https://www2.census.gov/geo/pdfs/reference/LSADCodes.pdf

7. Various, "Topologically Integrated Geographic Encoding and Referencing," Wikipedia. Accessed: Mar. 23, 2025. [Online]. Available: https://en.wikipedia.org/wiki/Topologically_Integrated_Geographic_Encoding_and_Referencing

About the Data

1. Centers for Disease Control and Prevention (CDC), "Respiratory Syncytial Virus Hospitalization Surveillance Network (RSV-NET)," cdc.gov. Accessed: Mar. 24, 2025. [Online]. Available: https://www.cdc.gov/rsv/php/surveillance/rsv-net.html

2. Centers for Disease Control and Prevention (CDC), "About RSV." Accessed: Jul. 15, 2025. [Online]. Available: https://www.cdc.gov/rsv/about/index.html

References

About the Data continued

3. Centers for Disease Control and Prevention (CDC), "Respiratory Virus Hospitalization Surveillance Network (RESP-NET)." Accessed: Jul. 15, 2025. [Online]. Available: https://www.cdc.gov/resp-net/dashboard/

4. Yale School of Public Health, "PopHIVE." Accessed: Jul. 15, 2025. [Online]. Available: https://www.pophive.org/

5. "Weekly Rates of Laboratory-Confirmed RSV Hospitalizations from the RSV-NET Surveillance System," CDC's data.gov. Accessed: Jan. 28, 2025. [Online]. Available: https://data.cdc.gov/Public-Health-Surveillance/Weekly-Rates-of-Laboratory-Confirmed-RSV-Hospitali/29hc-w46k/about_data

Data Processing

1. Answered by Hong Ooi, "lm() within mutate() in group_by()," Stack Overflow. Accessed: Mar. 17, 2025. [Online]. Available: https://stackoverflow.com/questions/40060828/lm-within-mutate-in-group-by

References

Data Processing continued

2. Centers for Disease Control and Prevention (CDC), “MMWR Weeks Definition”, Accessed: Mar. 14, 2025. [Online]. Available: https://ndc.services.cdc.gov/wp-content/uploads/MMWR_week_overview.pdf

3. R. Zhu, “Statistical Learning and Machine Learning with R.” Accessed: Mar. 17, 2025. [Online]. Available: https://teazrq.github.io/SMLR/kernel-smoothing.html

Glossary

1. Posit Contributors, Data visualization with ggplot2: Cheat Sheet. Springer-Verlag, 2024. Accessed: Mar. 20, 2025. [Online]. Available: https://rstudio.github.io/cheatsheets/html/data-visualization.html

2. H. Wickham et al., ggplot2: Elegant Graphics for Data Analysis Documentation. New York: Springer-Verlag, 2016. Accessed: Mar. 12, 2025. [Online]. Available: https://ggplot2.tidyverse.org/

References

Glossary continued

3. H. Wickham, M. Çetinkaya-Rundel, and G. Grolemund, *R for Data Science (2e)*, Second. O'Reilly Media, 2023. Accessed: Mar. 12, 2025. [Online]. Available: https://r4ds.hadley.nz/

Glossary

1. Centers for Disease Control and Prevention (CDC), "MMWR Weeks Definition", Accessed: Mar. 14, 2025. [Online]. Available: https://ndc.services.cdc.gov/wp-content/uploads/MMWR_week_overview.pdf

2. C. Wilke, "Streamlined Plot Theme and Plot Annotations for ggplot2 • cowplot." Accessed: Mar. 20, 2025. [Online]. Available: https://wilkelab.org/cowplot/