

# BASIC CHART TYPES

---

Hans-Jörg Schulz



1. INTRODUCTION

2. VISUAL ENCODING

3. BASIC CHART TYPES

4. INTERACTION

5. VISUALIZATION DESIGN

6. DATA PREPROCESSING

7. RECAP 1<sup>st</sup> Half

8. MULTIVARIATE DATAVIS

9. TEMPORAL DATAVIS

10. GEOSPATIAL DATAVIS

11. GRAPH DATAVIS

12. 3D DATAVIS

13. VISUAL ANALYTICS

14. RECAP 2<sup>nd</sup> Half

Basics

Visualization  
Building Blocks  
& Processes

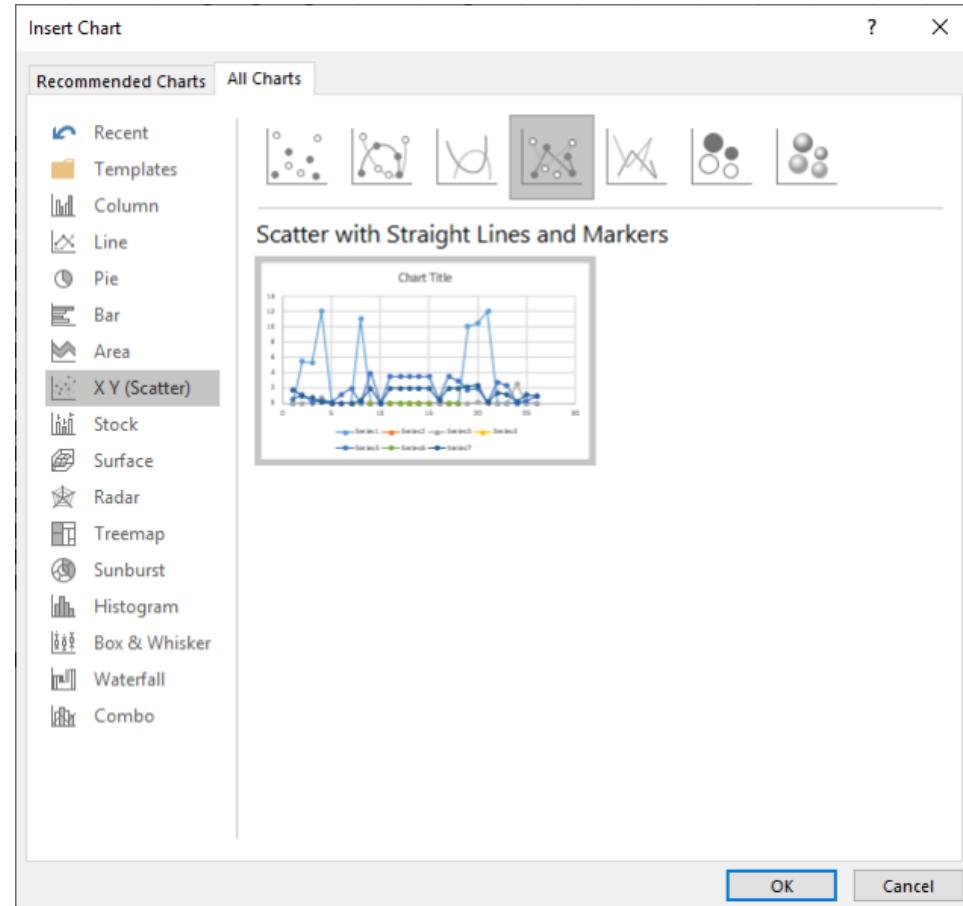
Visualization  
Techniques

Visualization  
Applications

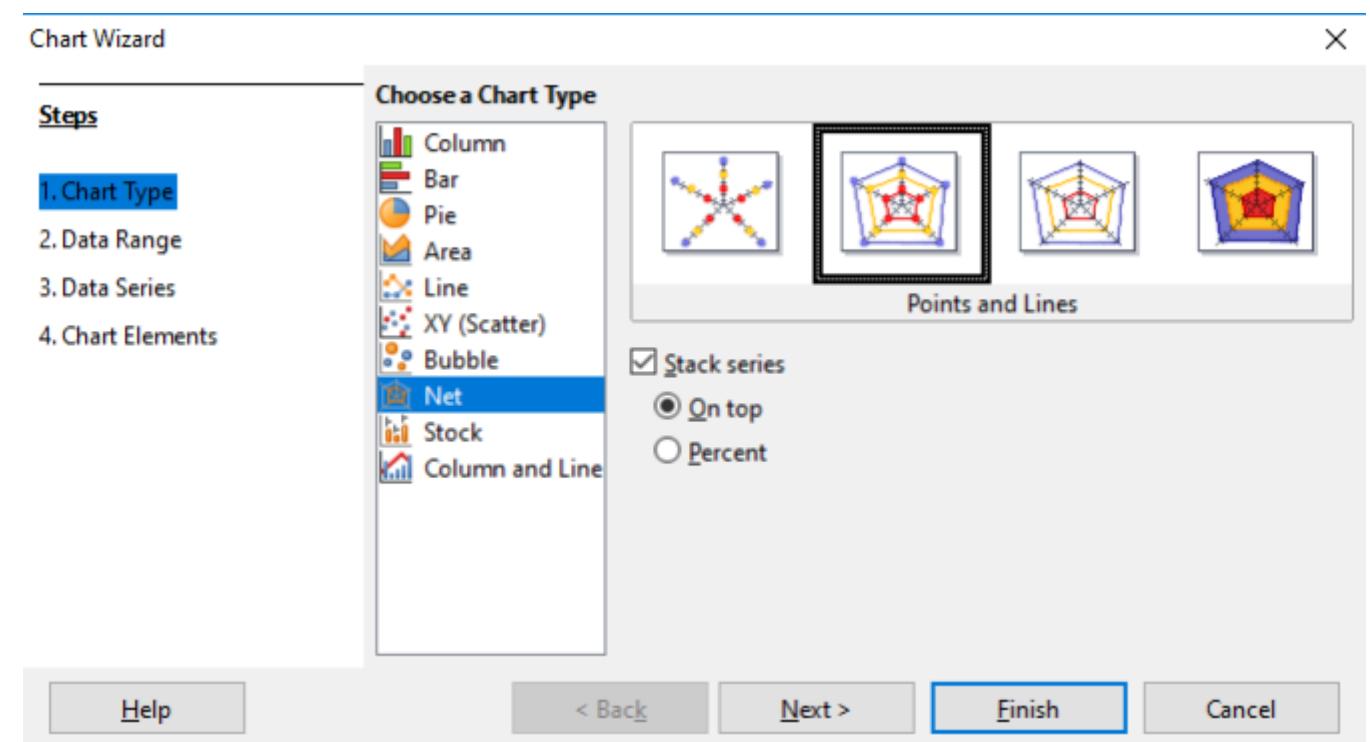


# WHY TALK ABOUT BASIC CHARTS?

---



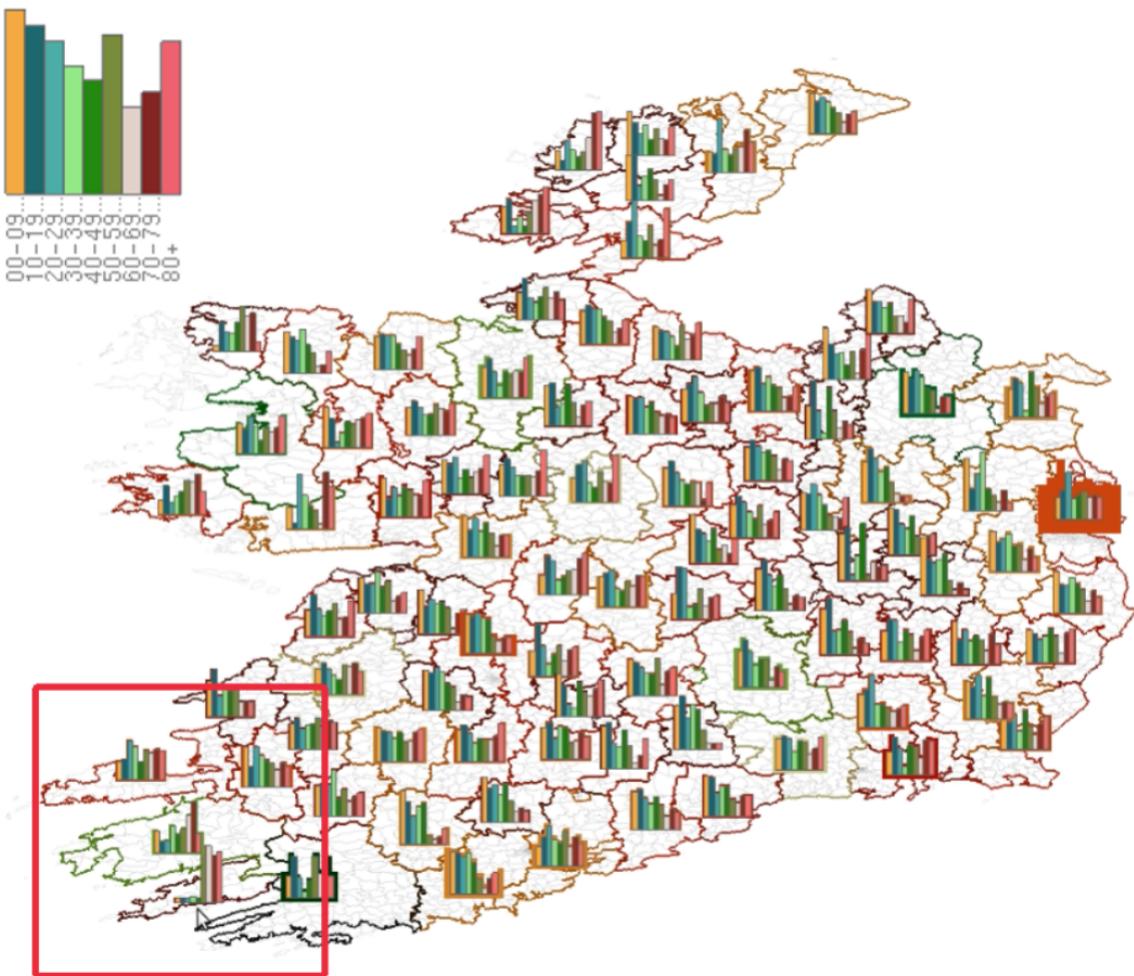
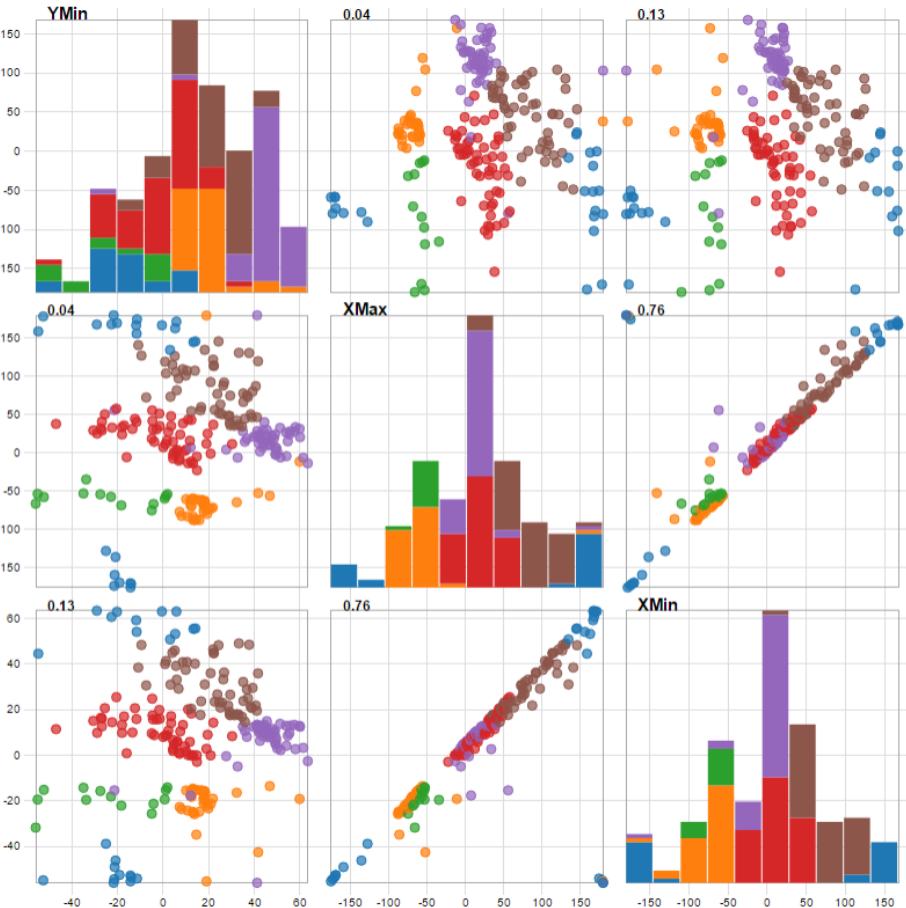
Charting Wizard MS Excel



Charting Wizard Libre Calc

# BASIC CHARTS AS BUILDING BLOCKS

---



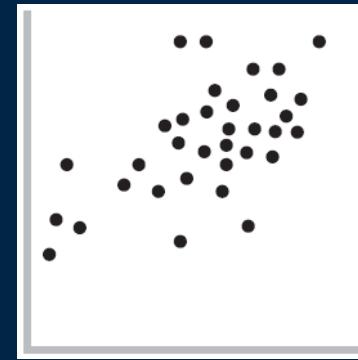
# OVERVIEW

---

- Scatterplots incl. Binned Scatterplots, Splatterplots, Connected Scatterplots
- Line Charts, Stepped Line Charts
- Area Charts and Stacked Area Charts
- Bar / Column Charts incl. Stacked Bar Charts, Grouped Bar Charts, Histograms
- Radial Variants incl. Radar Chart, Coxcomb

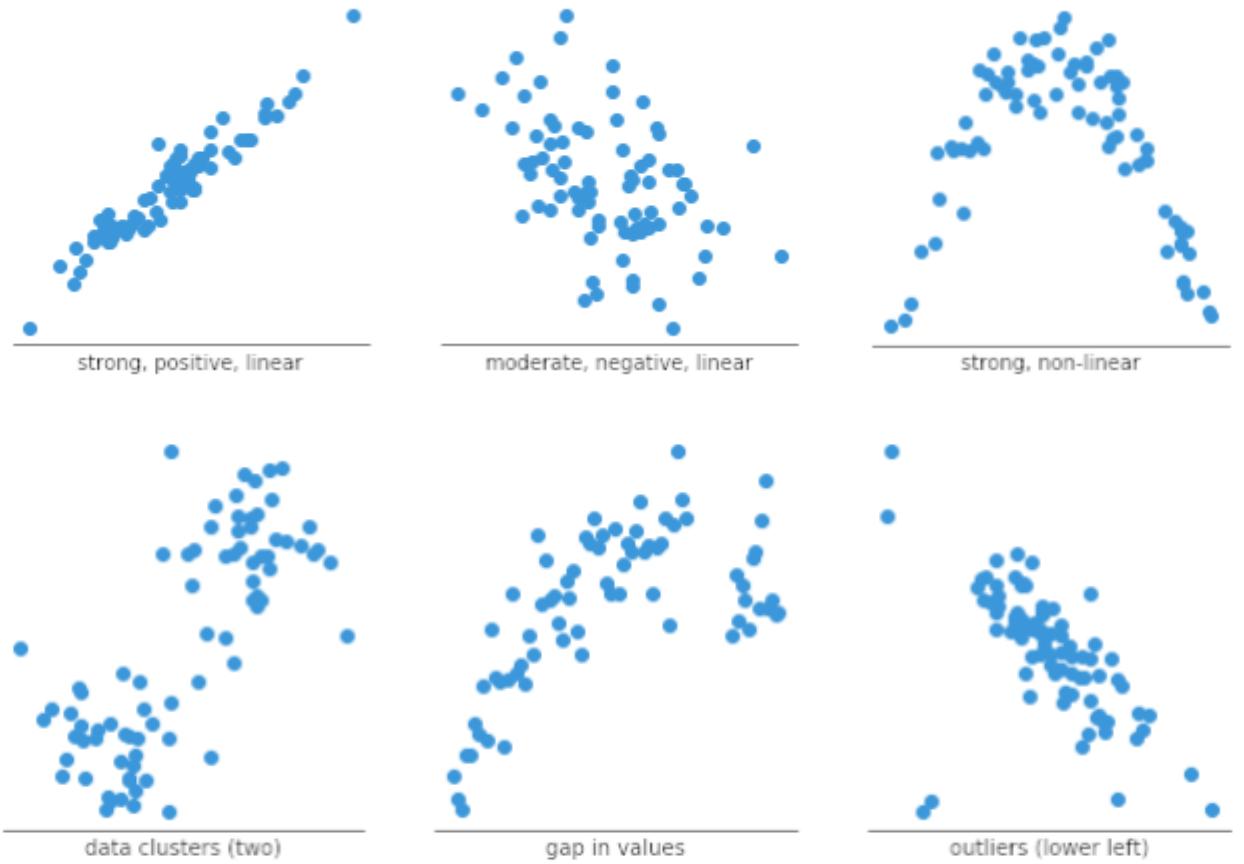
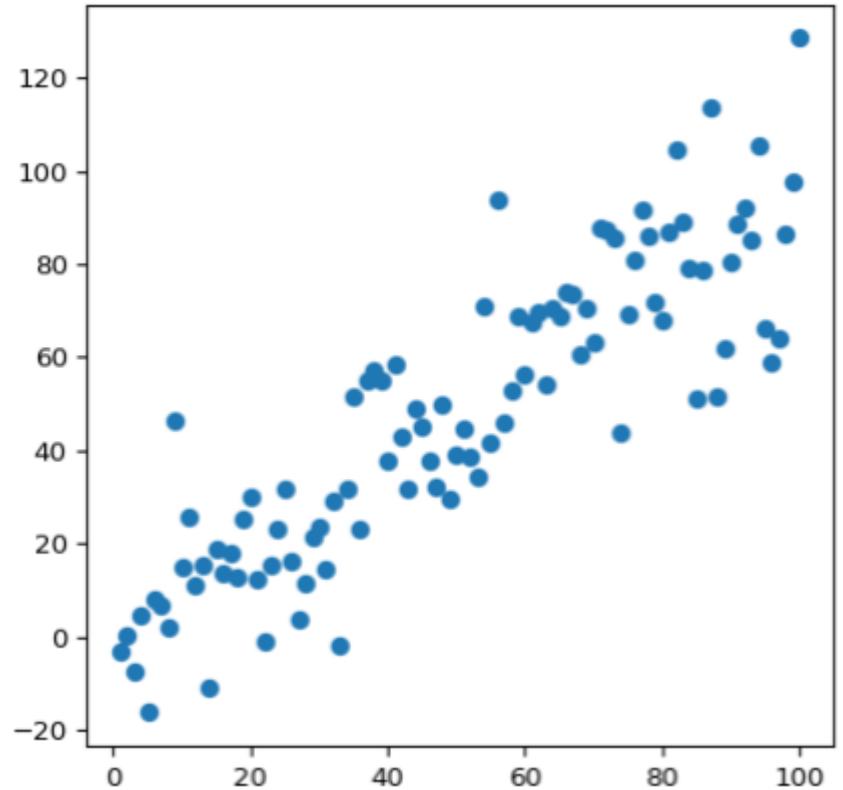


# THE SCATTERPLOT



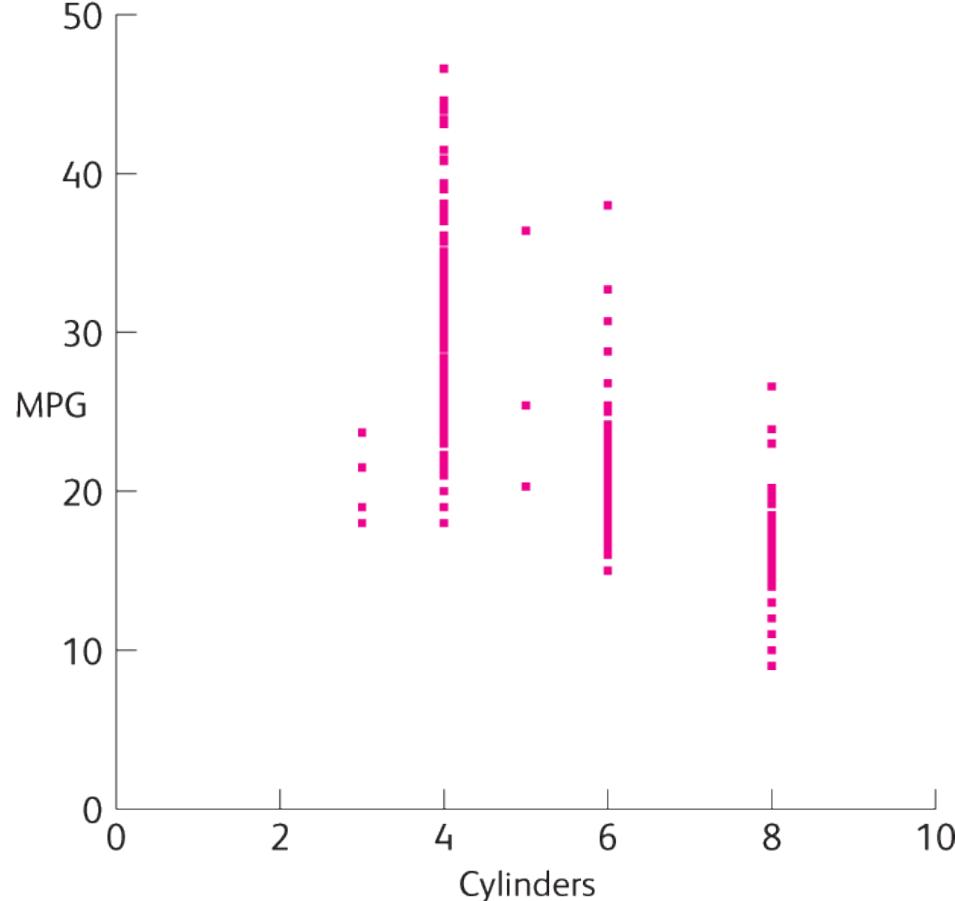
# THE SCATTERPLOT

---



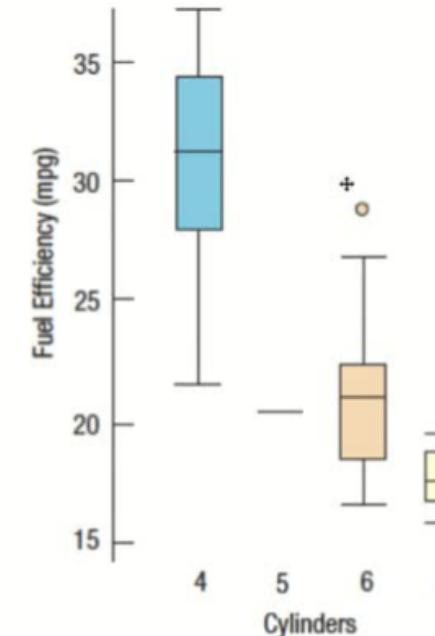
# THE SCATTERPLOT

---



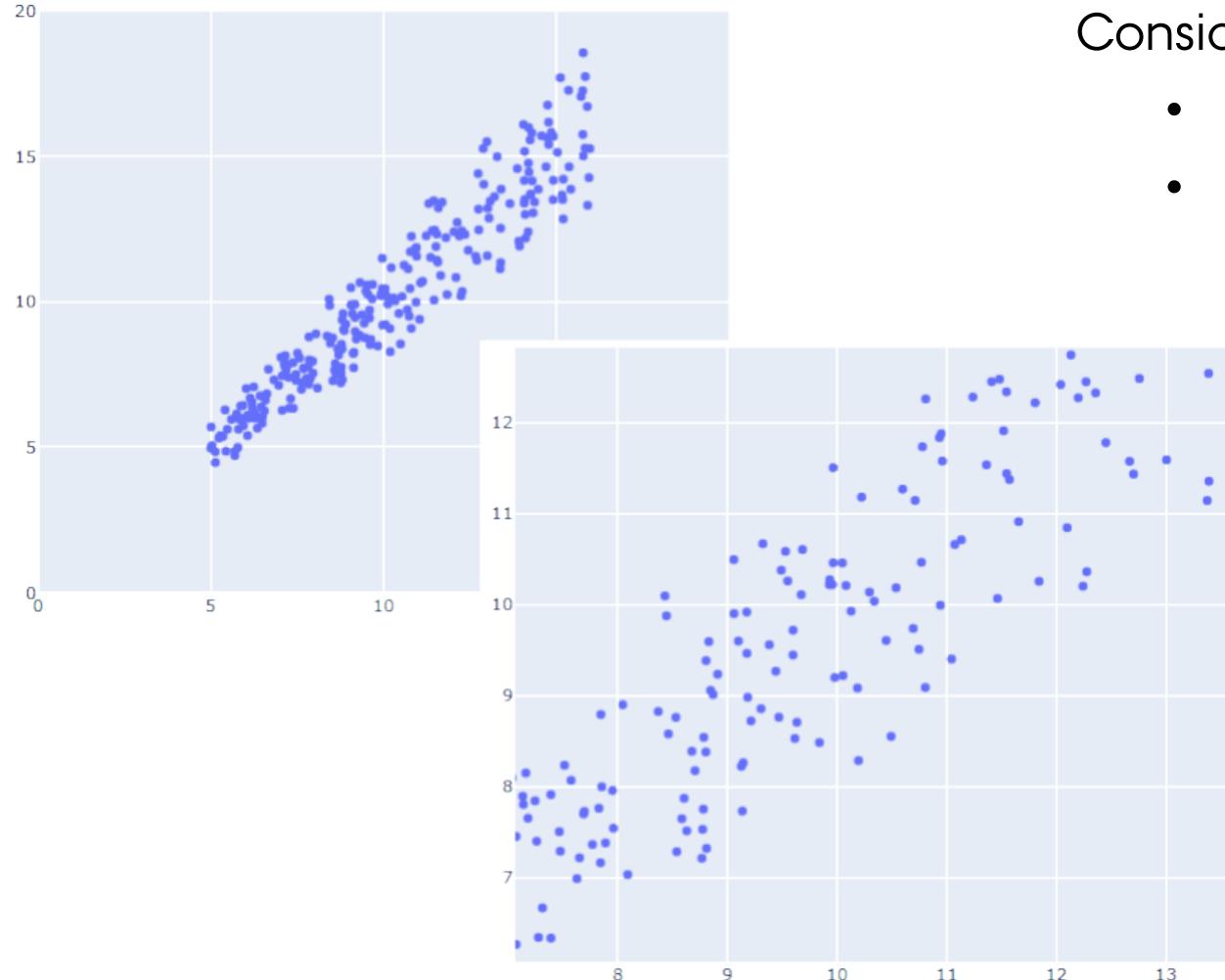
Considerations:

- Does it make sense?



# THE SCATTERPLOT

---



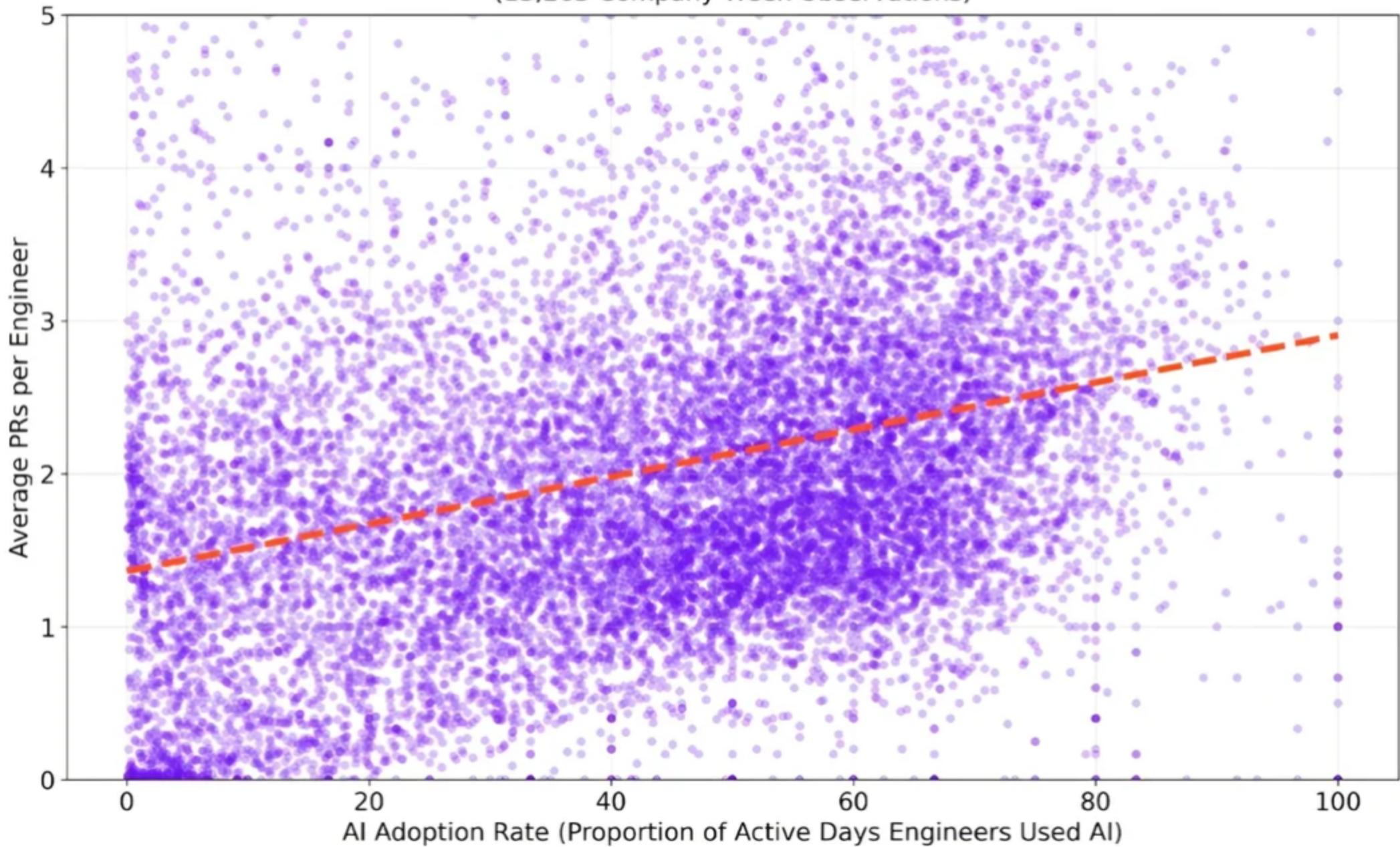
Considerations:

- Does it make sense?
- How to scale it?



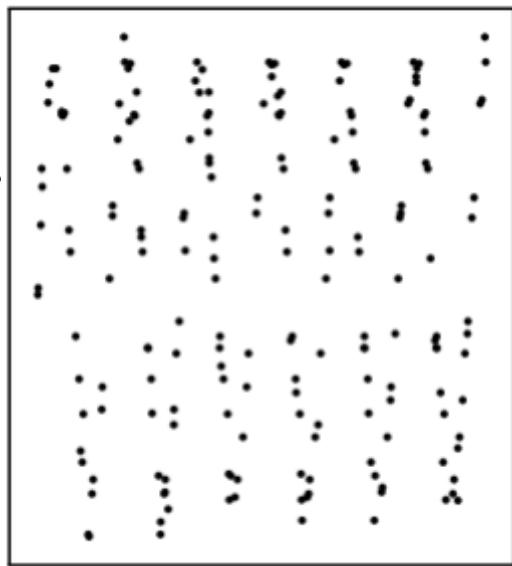
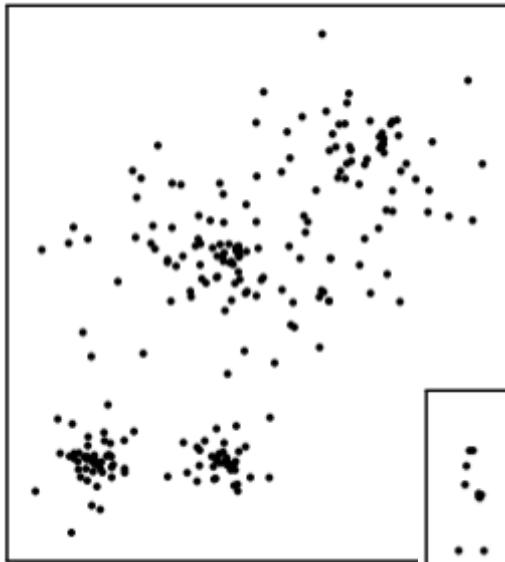
# Companies with Higher AI Use Merge More PRs per Engineer

(13,263 Company-Week Observations)



# THE SCATTERPLOT

---



Considerations:

- Does it make sense?
- How to scale it?
- Which aspect ratio to use?

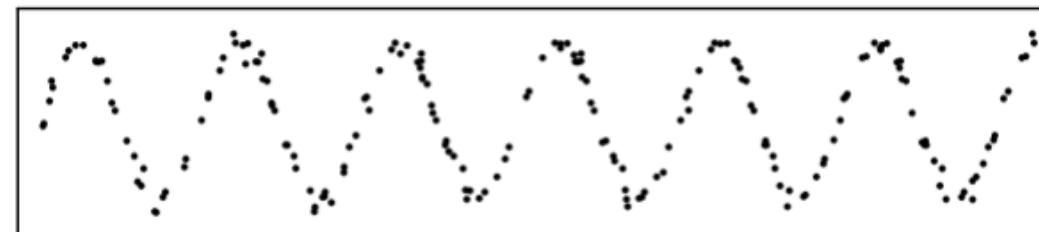
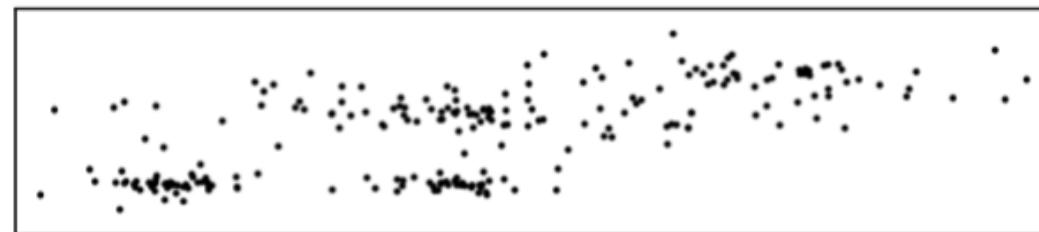
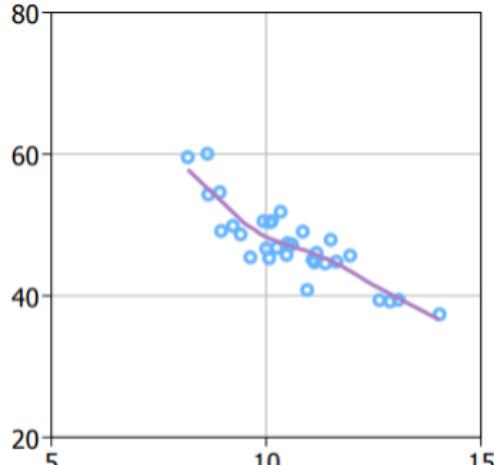


Image source:  
[doi:10.1109/TVCG.2013.187](https://doi.org/10.1109/TVCG.2013.187)

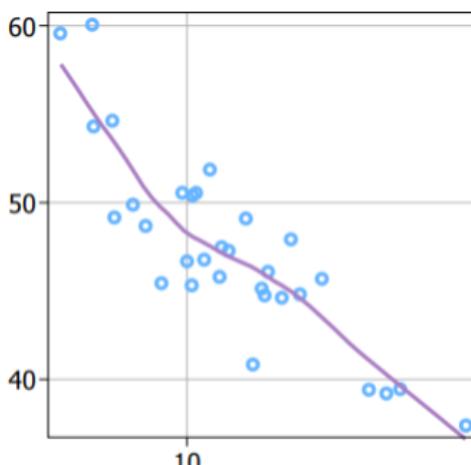


# THE SCATTERPLOT

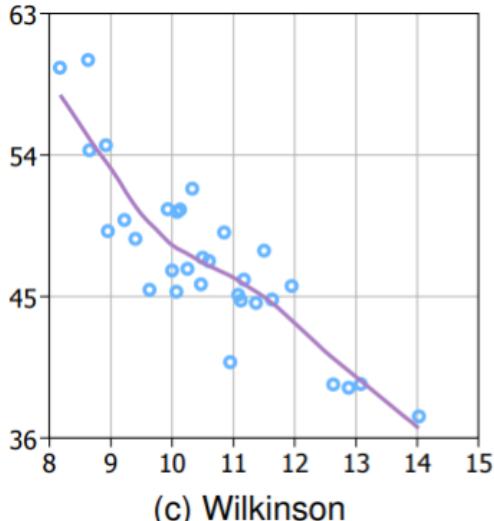
---



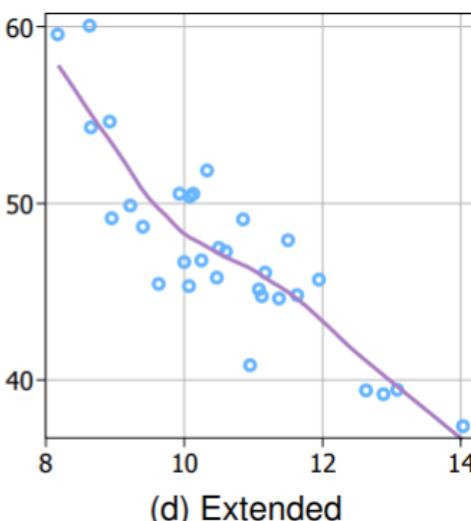
(a) Heckbert



(b) R's pretty



(c) Wilkinson



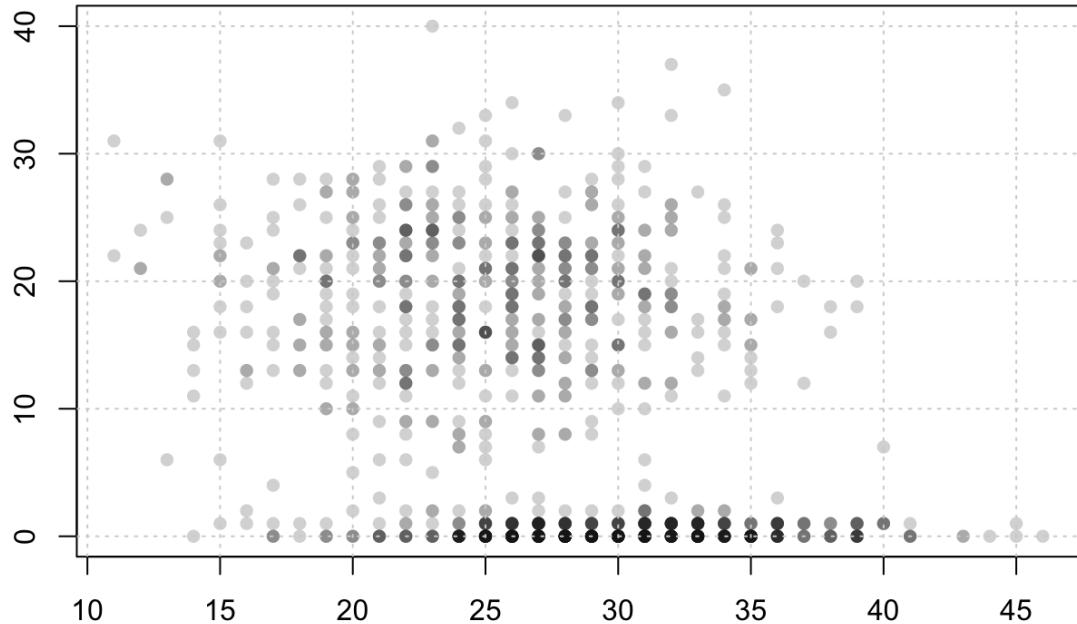
(d) Extended

Considerations:

- Does it make sense?
- How to scale it?
- Which aspect ratio to use?
- How to place the tick marks on the axes?

# THE SCATTERPLOT

---



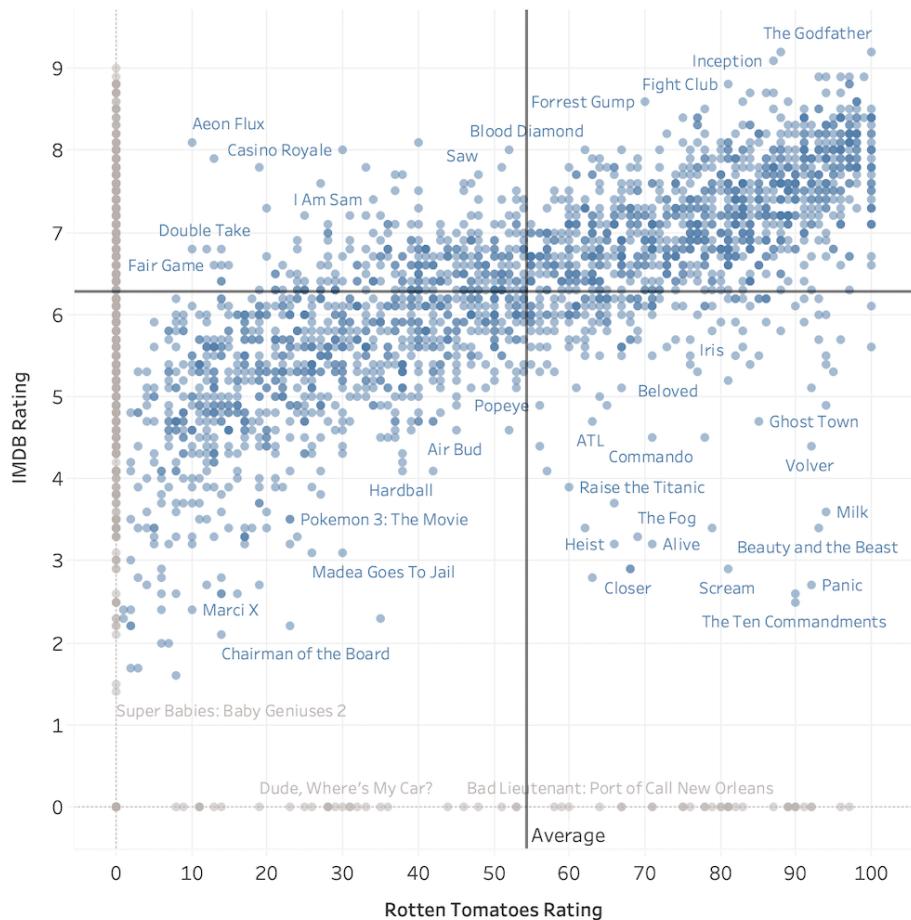
Considerations:

- Does it make sense?
- How to scale it?
- Which aspect ratio to use?
- How to place the tick marks on the axes?
- Should I add grid lines in the background?

# THE SCATTERPLOT

---

Movie ratings

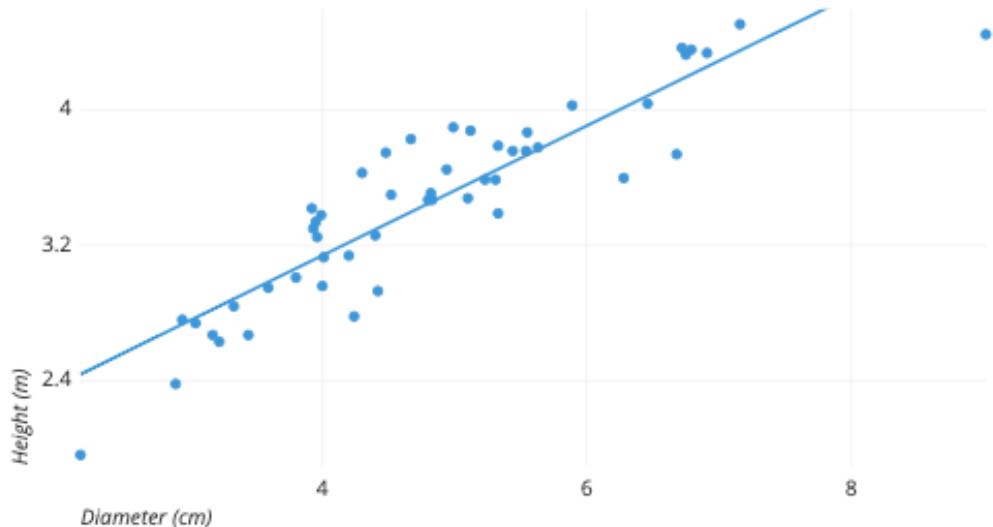


## Considerations:

- Does it make sense?
- How to scale it?
- Which aspect ratio to use?
- How to place the tick marks on the axes?
- Should I add grid lines in the background?
- Should I show quadrants / average lines?

# THE SCATTERPLOT

---

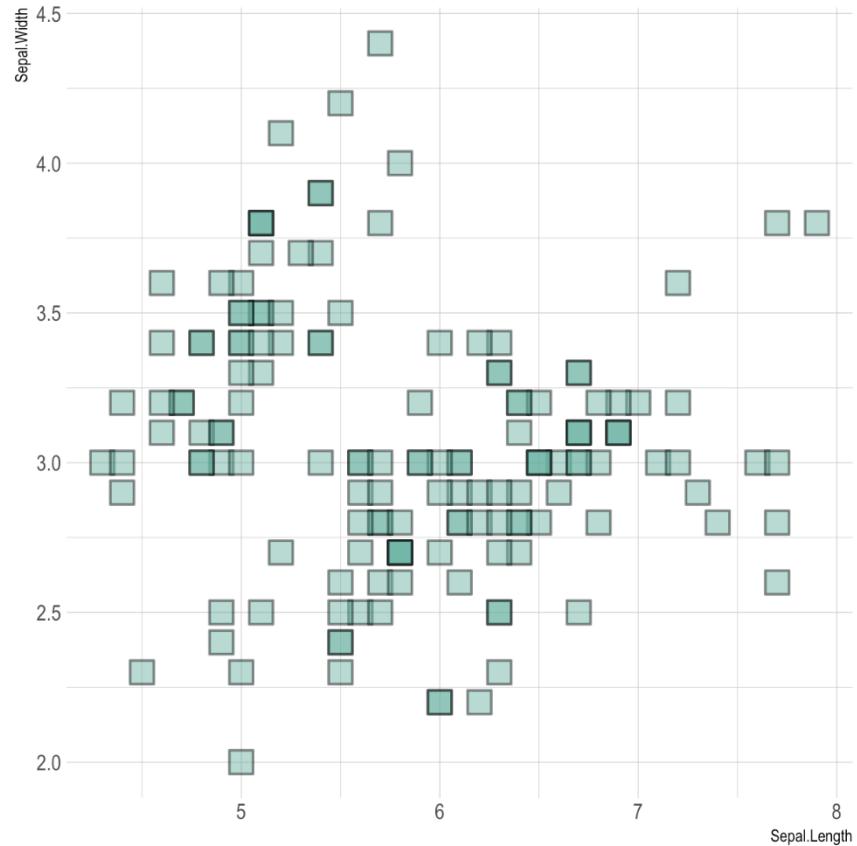


Considerations:

- Does it make sense?
- How to scale it?
- Which aspect ratio to use?
- How to place the tick marks on the axes?
- Should I add grid lines in the background?
- Should I show quadrants / average lines?
- Should I add a trend line?

# THE SCATTERPLOT

---



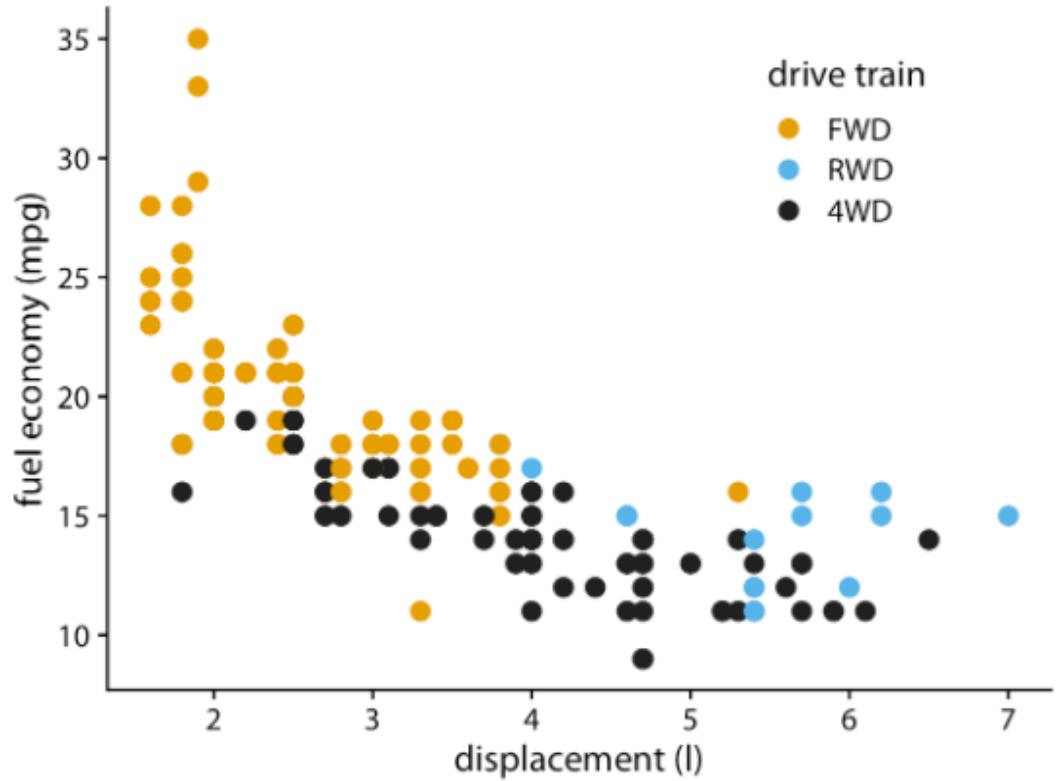
## Considerations:

- Does it make sense?
- How to scale it?
- Which aspect ratio to use?
- How to place the tick marks on the axes?
- Should I add grid lines in the background?
- Should I show quadrants / average lines?
- Should I add a trend line?
- What shape and size should the dots have?



# THE SCATTERPLOT

---

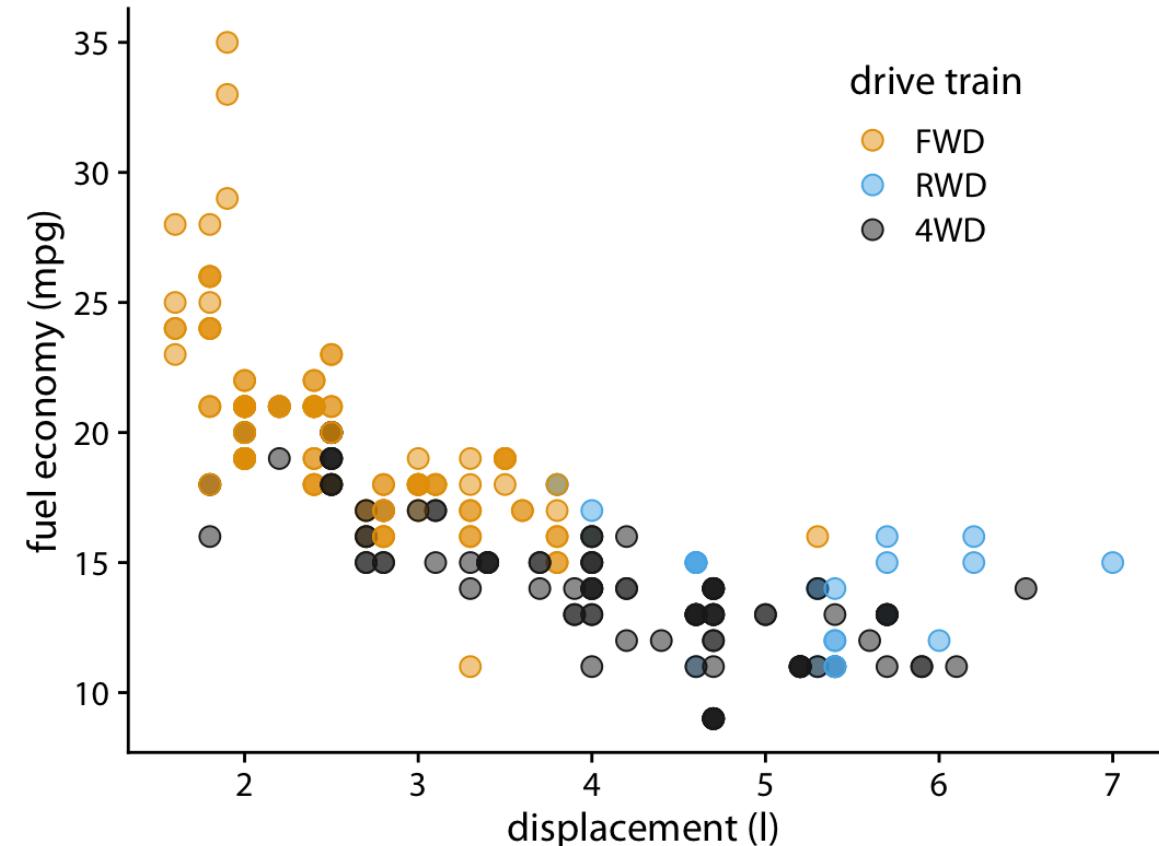
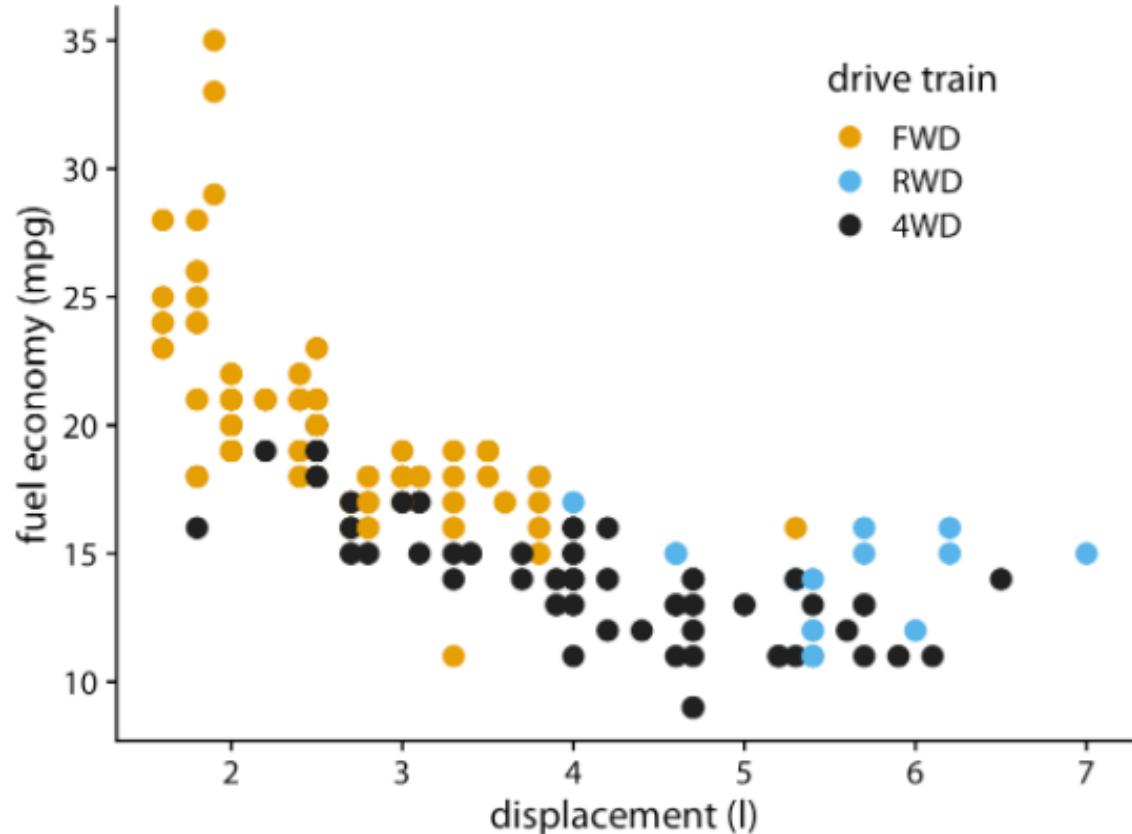


Considerations:

- Does it make sense?
- How to scale it?
- Which aspect ratio to use?
- How to place the tick marks on the axes?
- Should I add grid lines in the background?
- Should I show quadrants / average lines?
- Should I add a trend line?
- What shape and size should the dots have?
- What if the plot is heavily overplotted?  
-> Transparency and Jittering

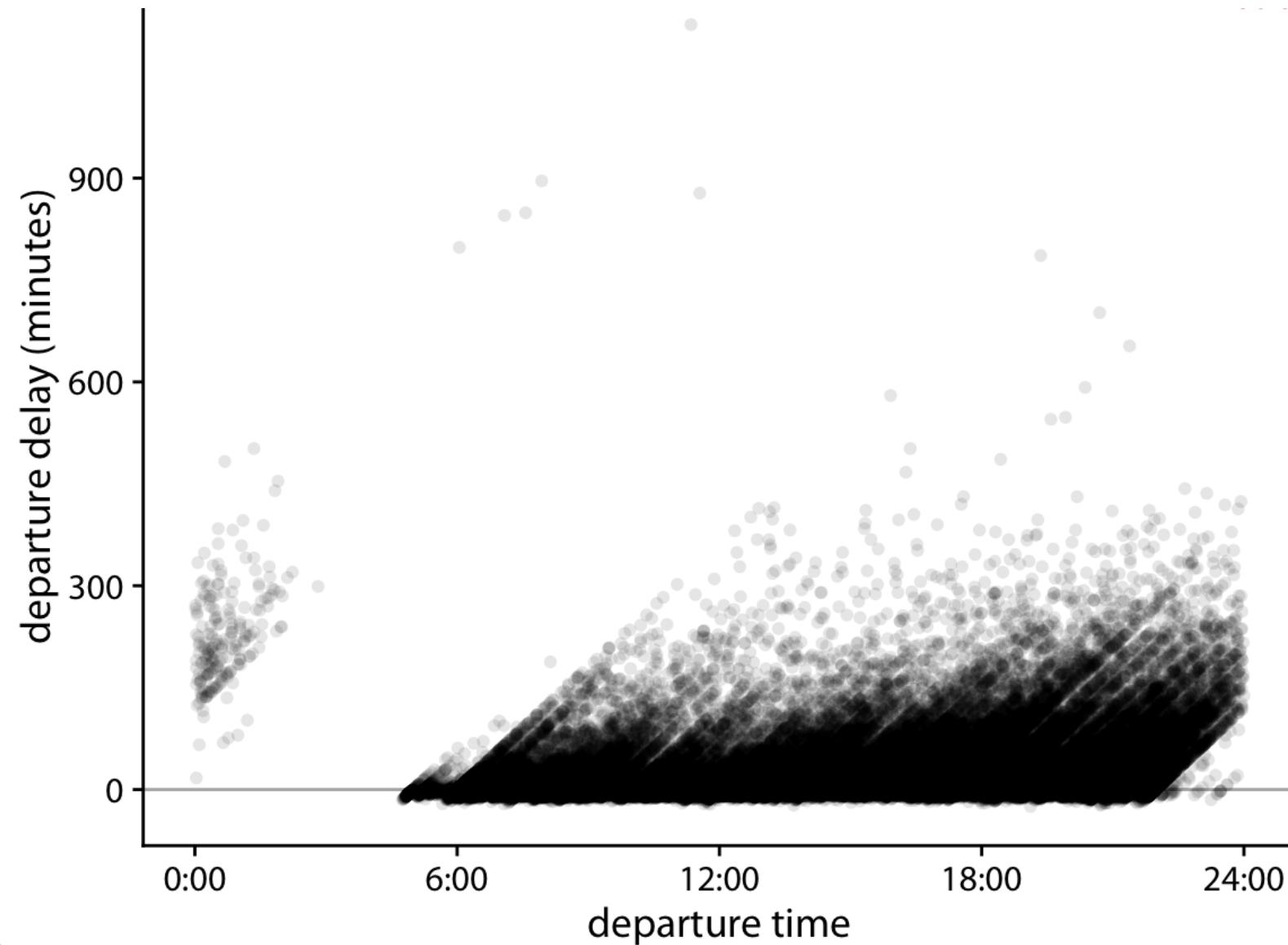
# ADD TRANSPARENCY

---



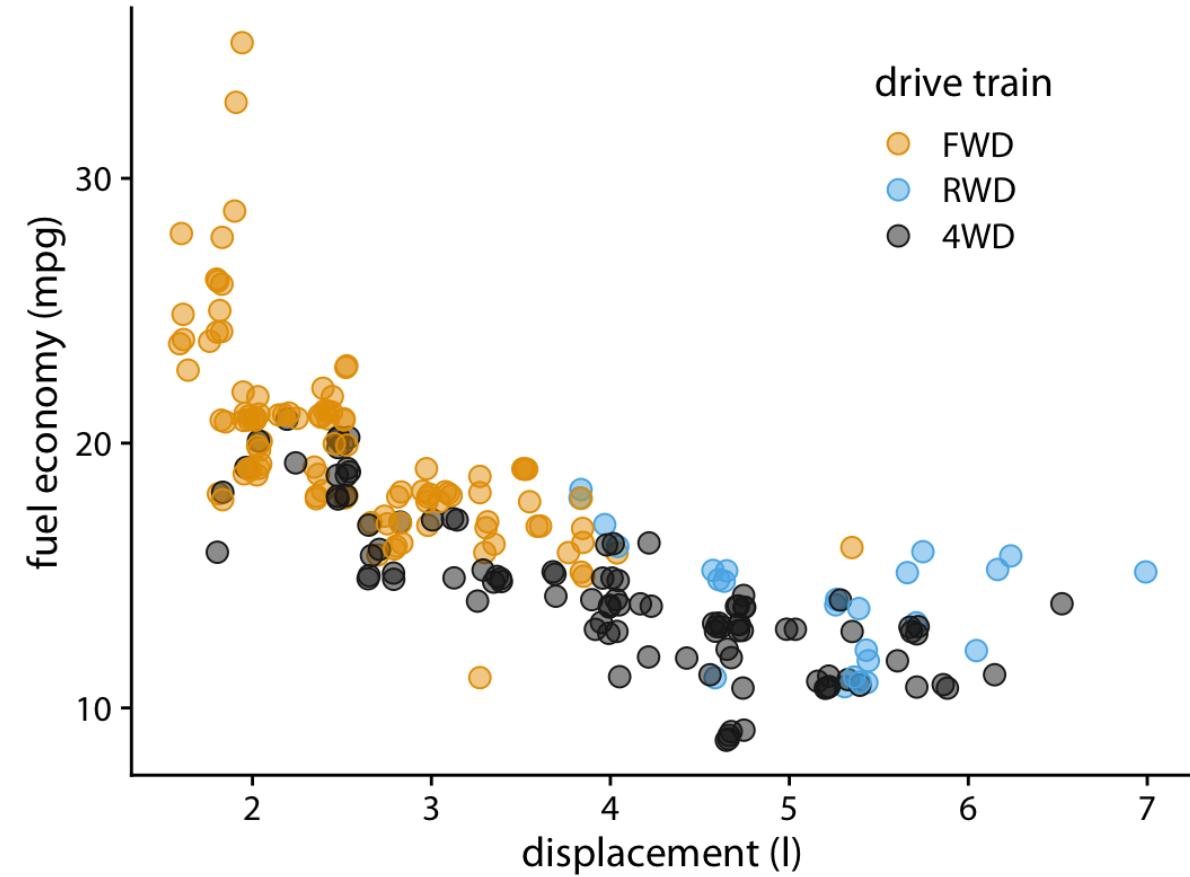
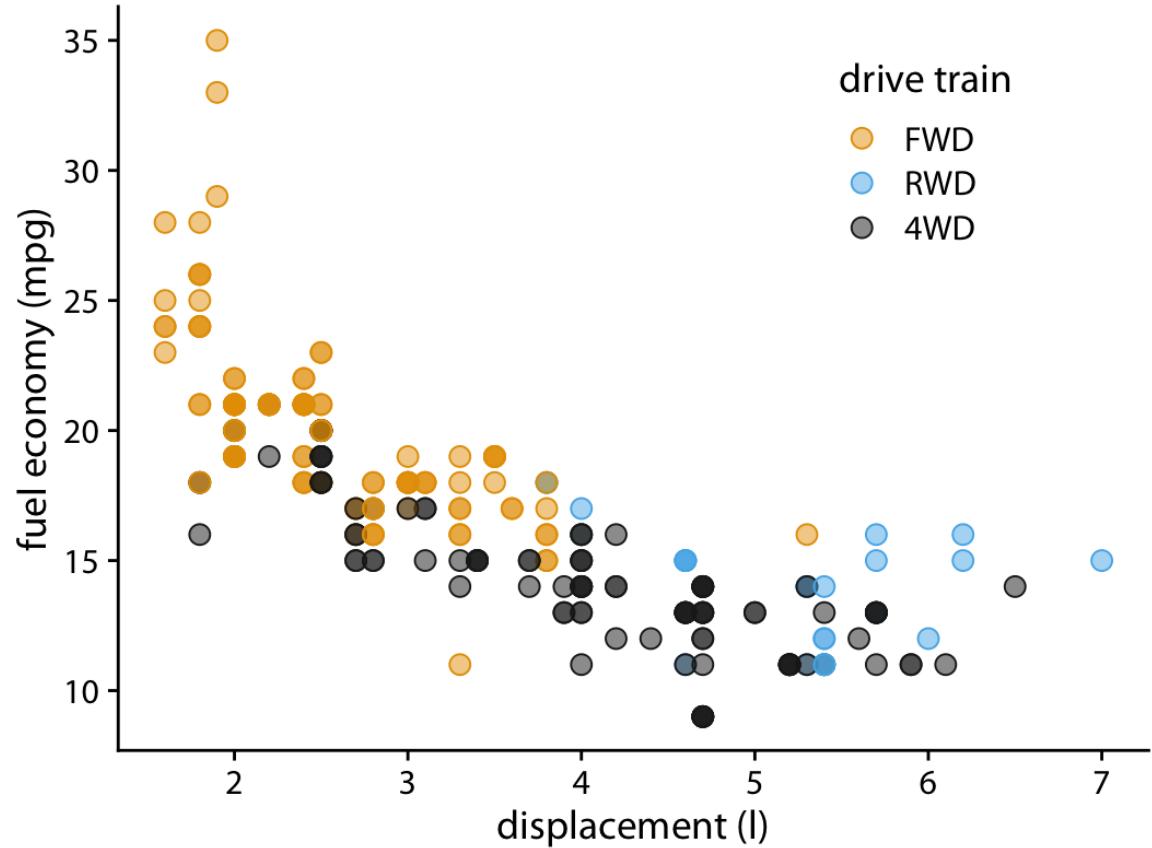
# THE PROBLEM WITH TRANSPARENCY

---



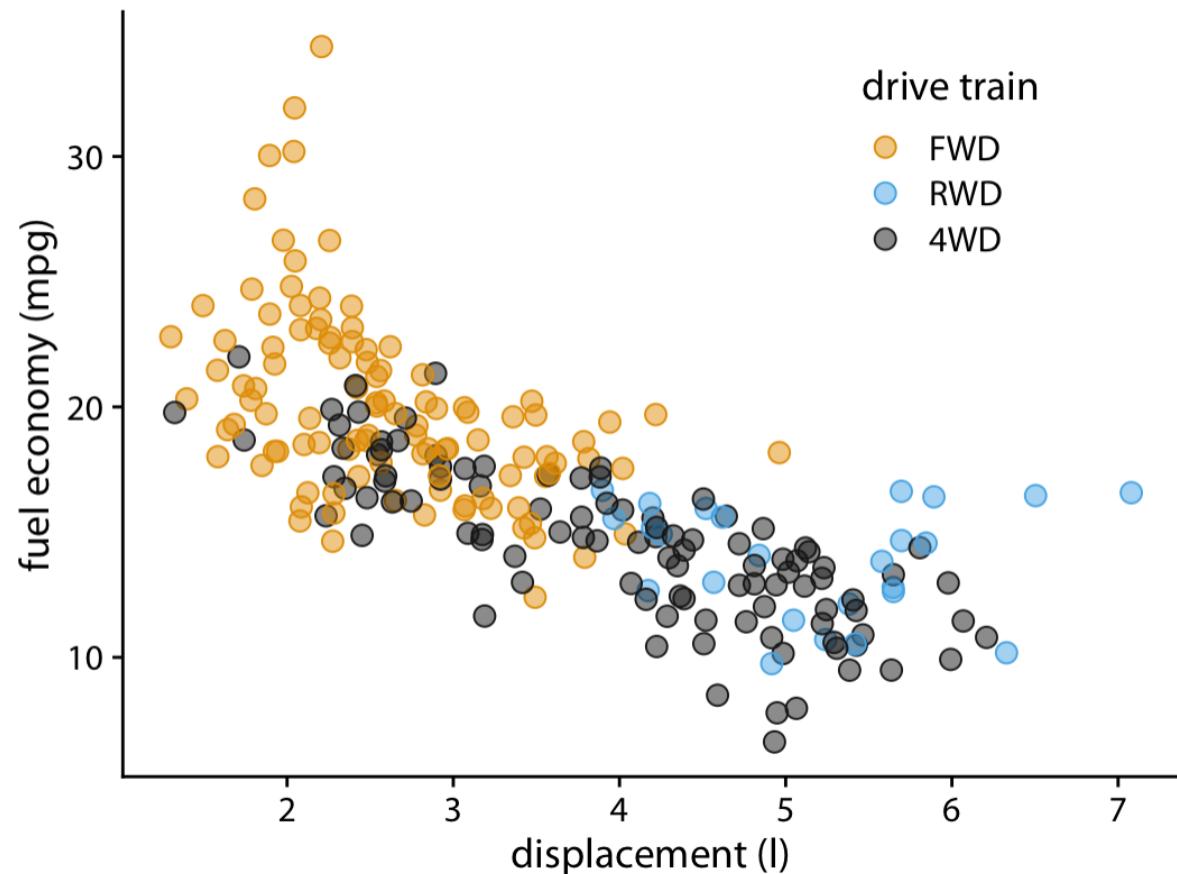
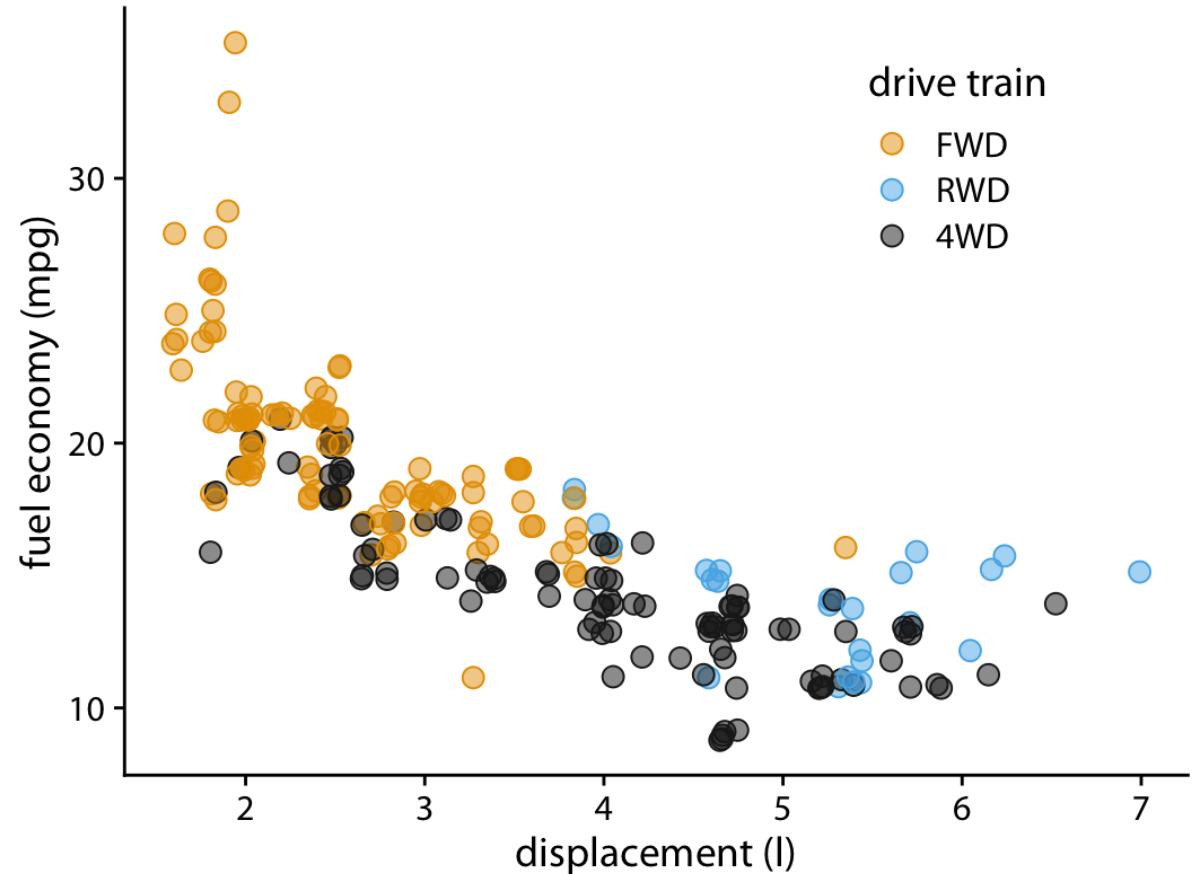
# ADD JITTERING

---



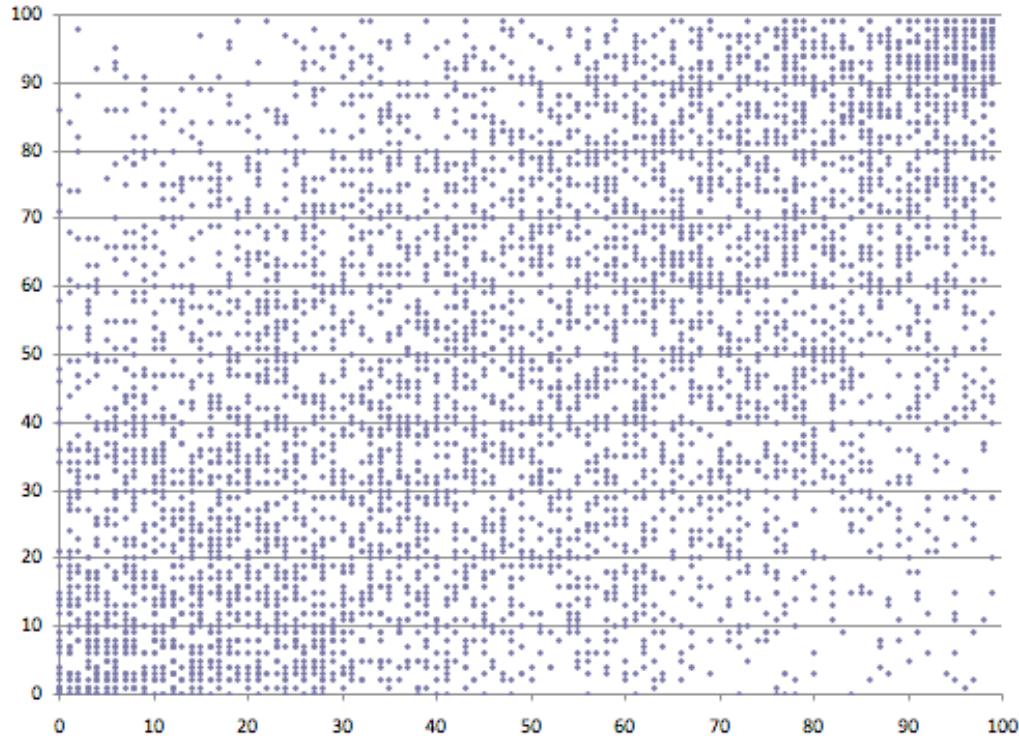
# ...BUT NOT TOO MUCH!

---



# THE SCATTERPLOT

---



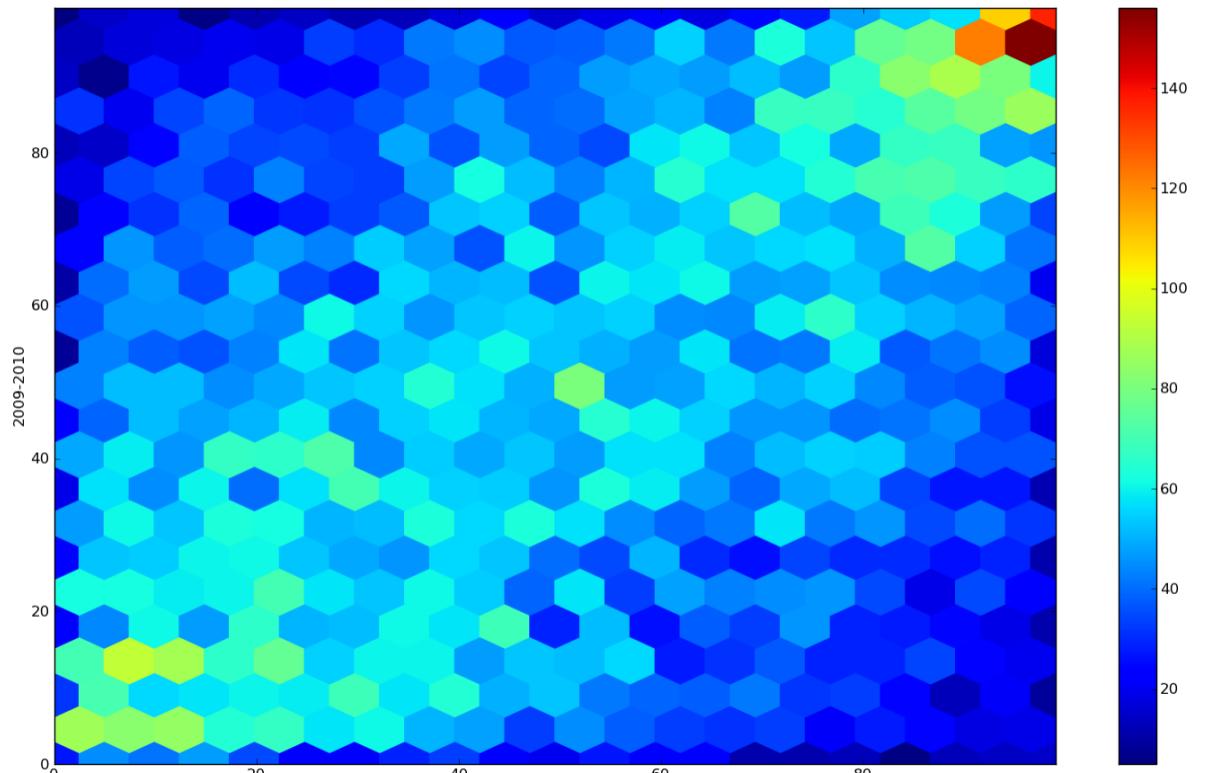
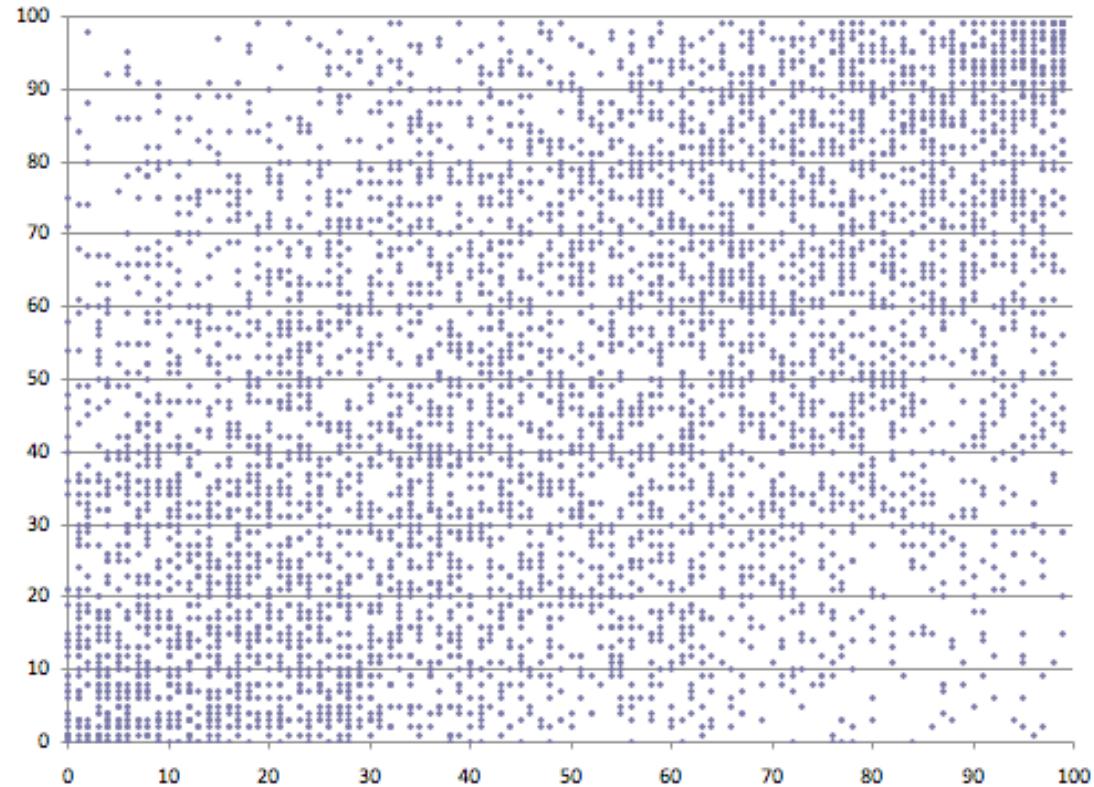
Considerations:

- Does it make sense?
- How to scale it?
- Which aspect ratio to use?
- How to place the tick marks on the axes?
- Should I add grid lines in the background?
- Should I show quadrants / average lines?
- Should I add a trend line?
- What shape and size should the dots have?
- What if the plot is heavily overplotted?  
-> Transparency and Jittering
- What if the plot is heavily overcrowded?  
-> Binning, Splatting, Heatmap



# ADD BINNING

---

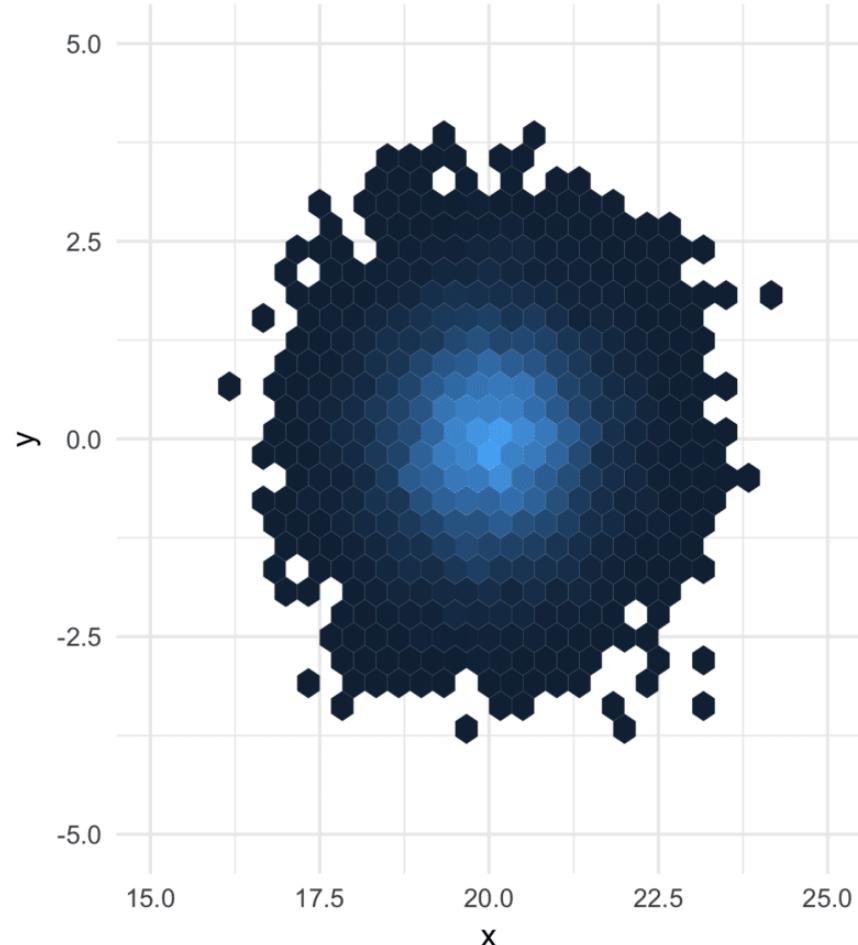


# THE BINNED SCATTERPLOT

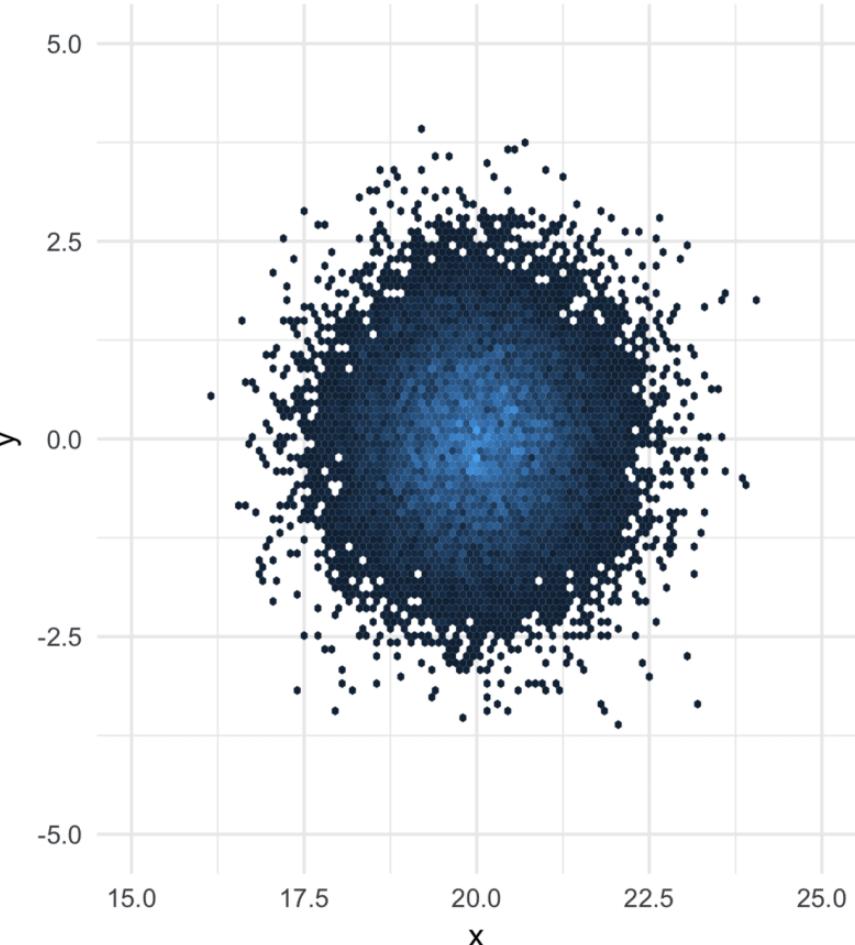
---

Counteracting overplotting with `geom_hex`

With default bins



With  $0.1 \times 0.1$  bins

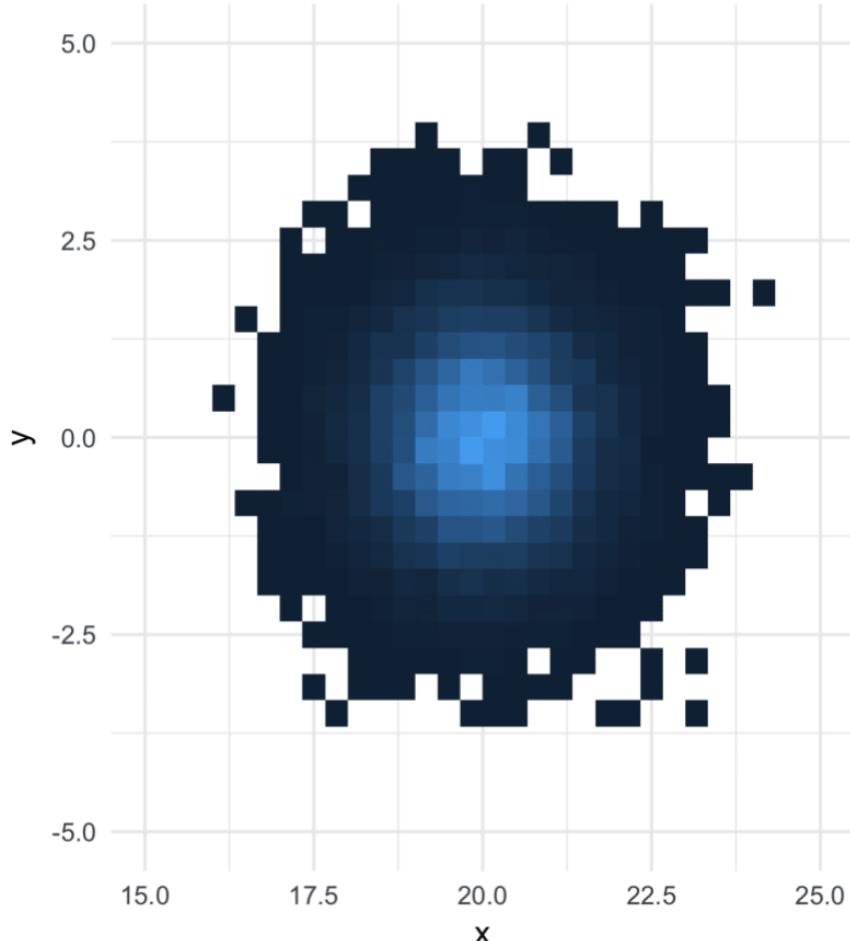


# THE BINNED SCATTERPLOT

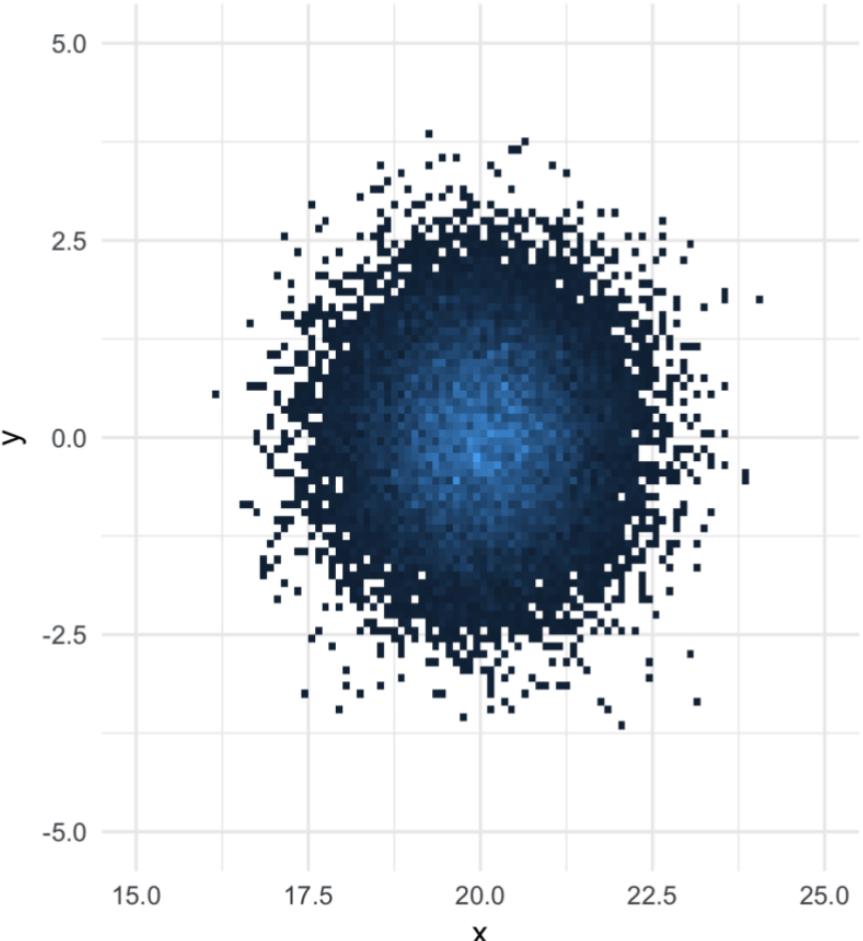
---

Counteracting overplotting with `geom_bin2d`

With default number of bins

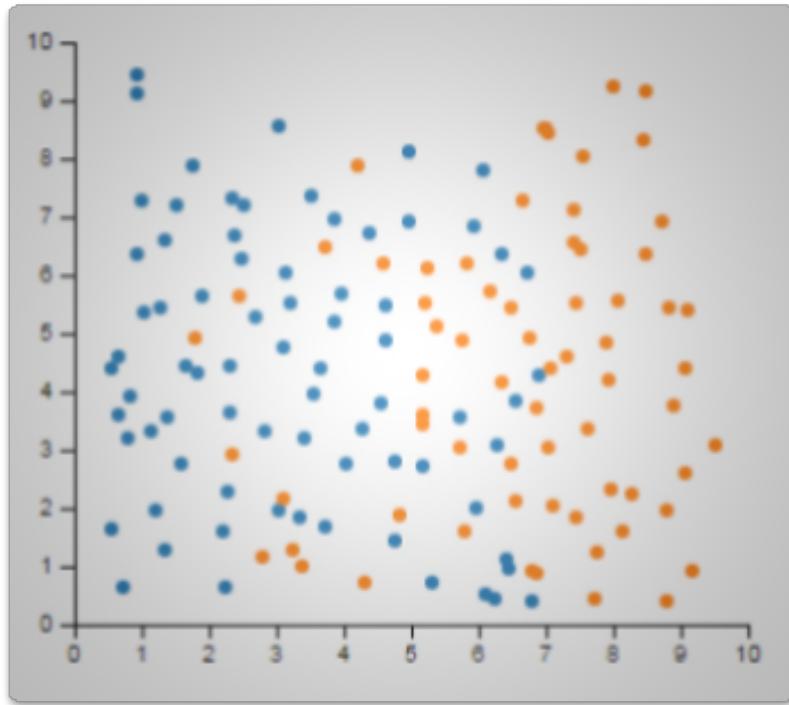


With 100 bins

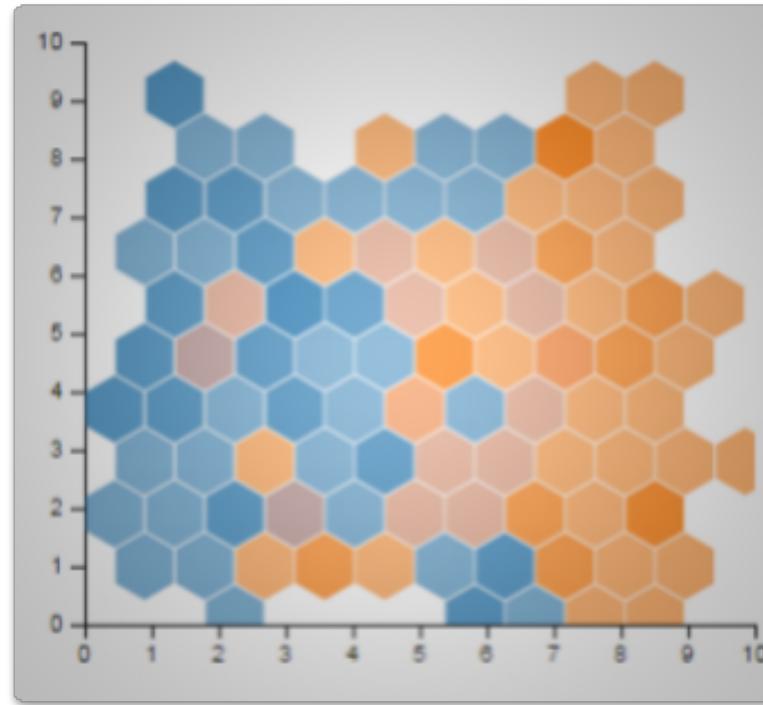


# THE SPLATTERPLOT

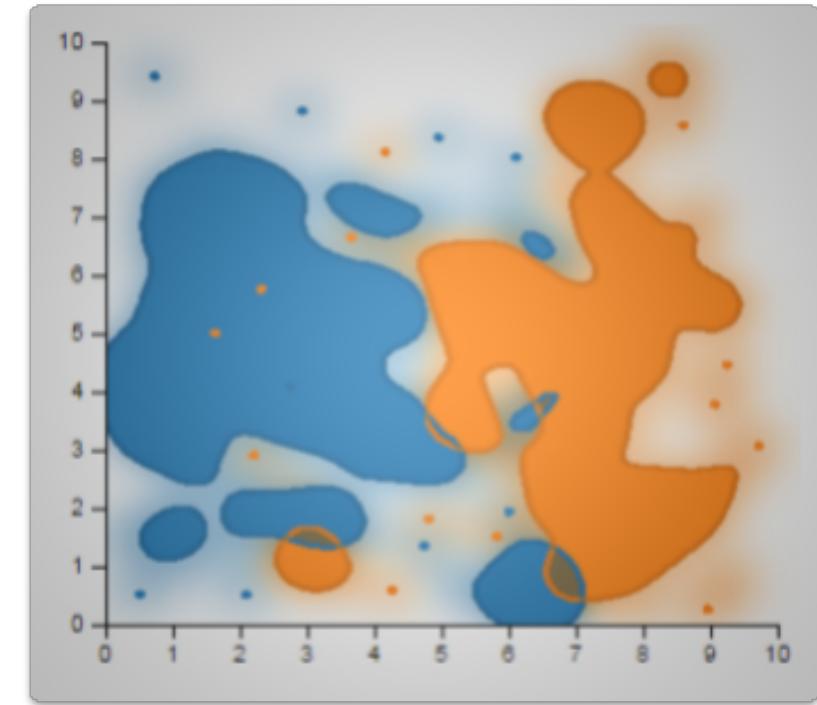
---



Scatterplot



Binned Scatterplot

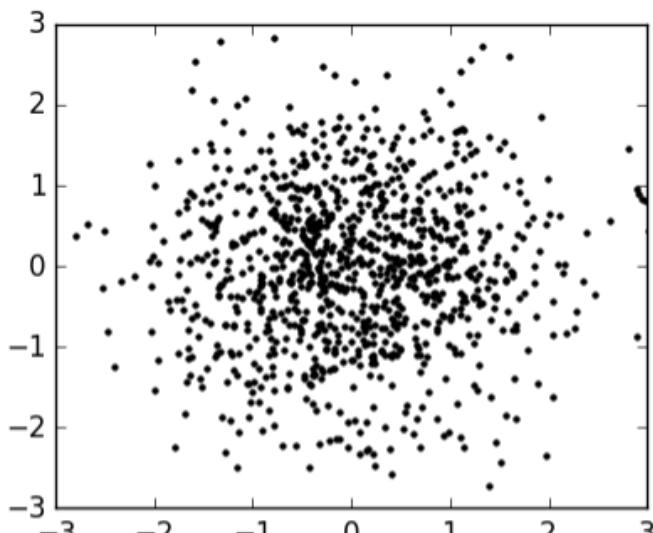


Splatterplot

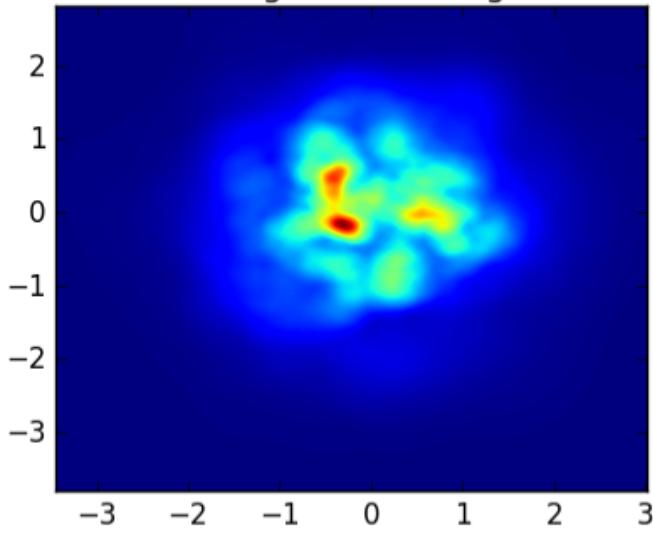


# THE HEATMAP

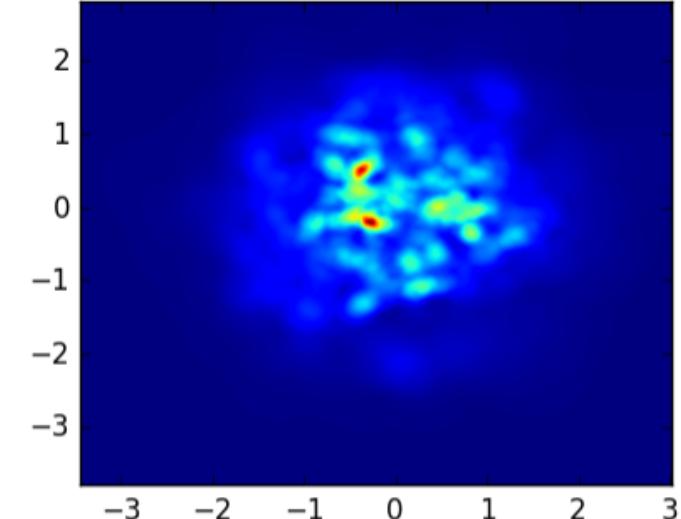
---



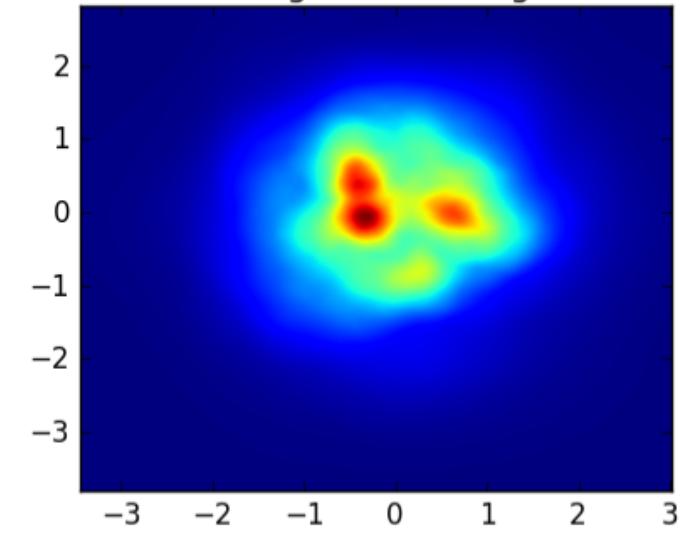
Smoothing over 32 neighbors



Smoothing over 16 neighbors

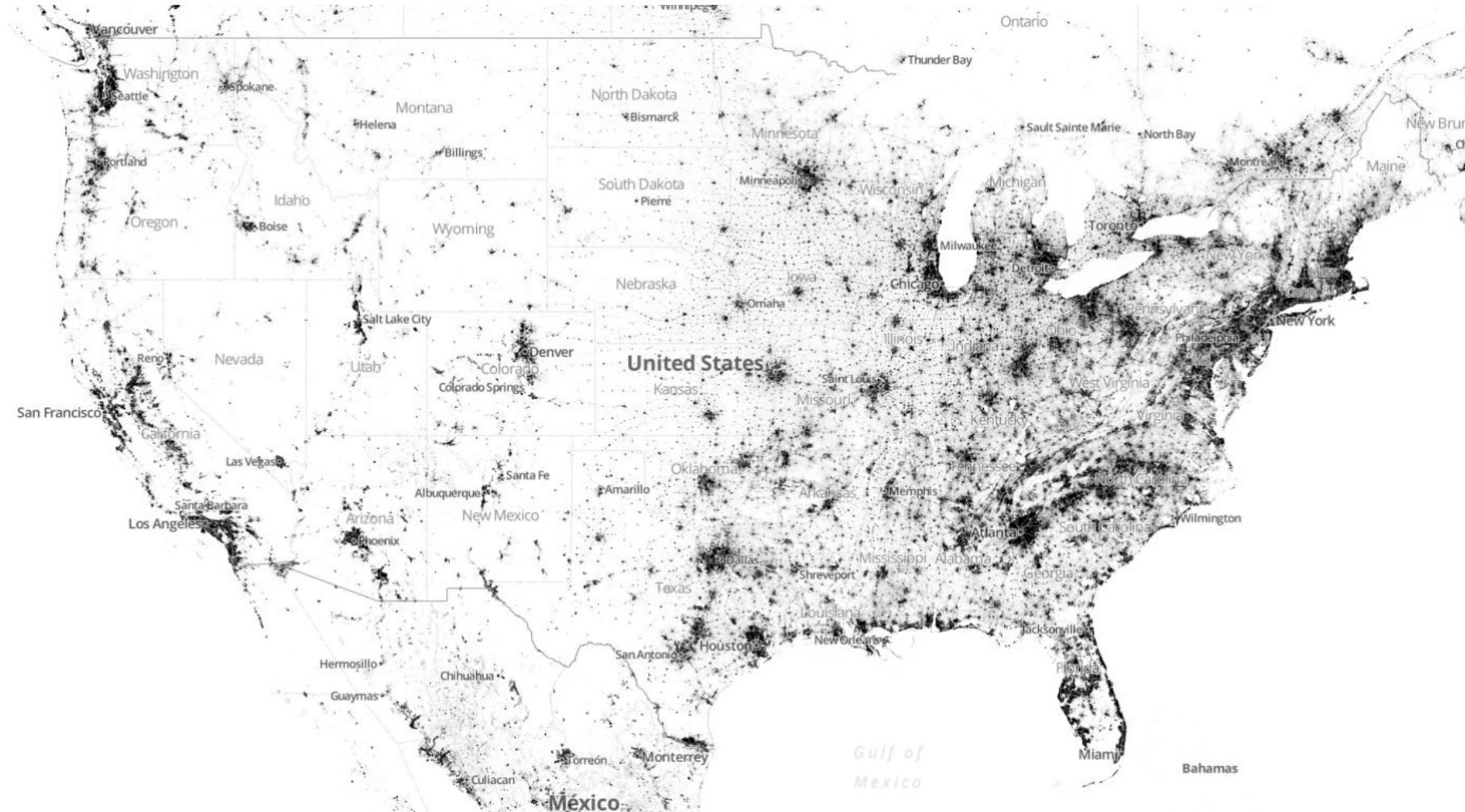


Smoothing over 64 neighbors



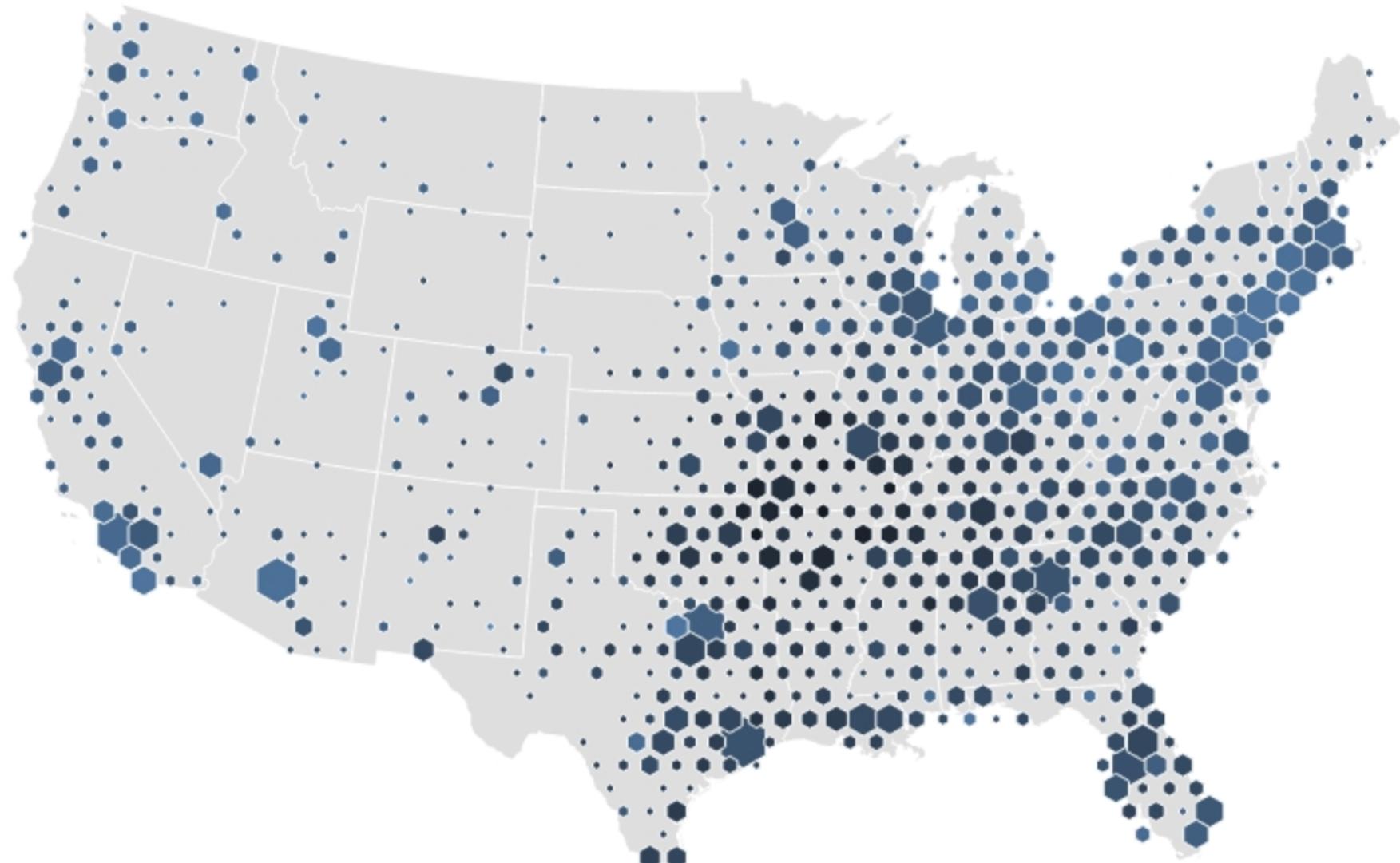
# ALSO WORKS FOR DOTMAPS...

---



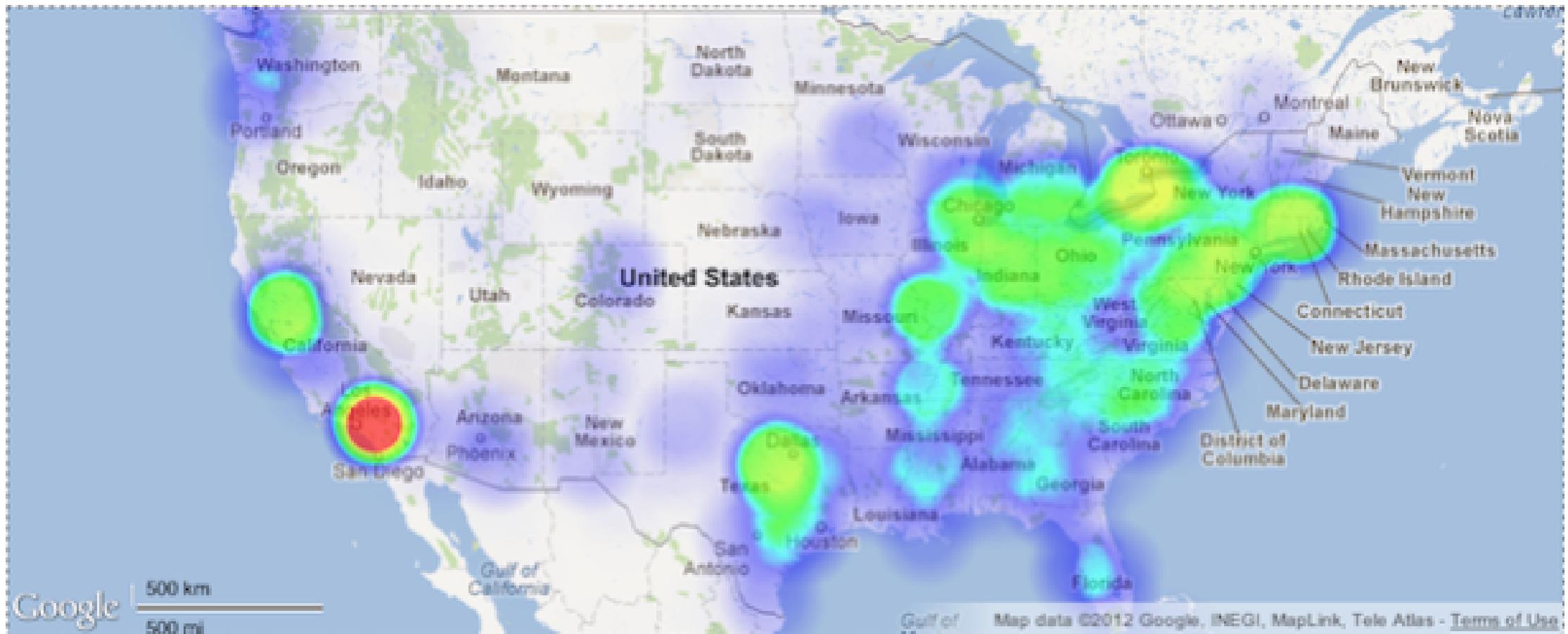
# ...BINNED DOTMAP

---



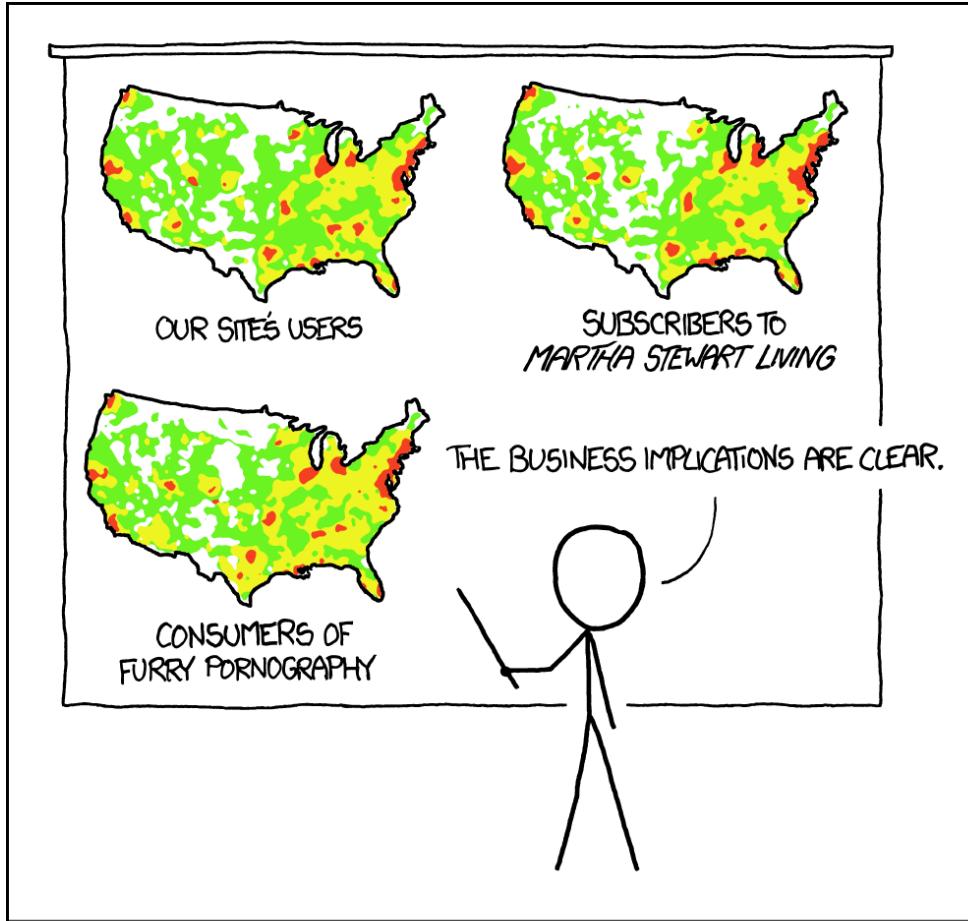
# ...HEATMAP

---



# NORMALIZE BY POPULATION DENSITY!

---



# Why Shouldnt All Charts Be Scatter Plots? Beyond Precision-Driven Visualizations

Enrico Bertini  
New York University

Michael Correll  
Tableau Research

Steven Franconeri  
Northwestern University

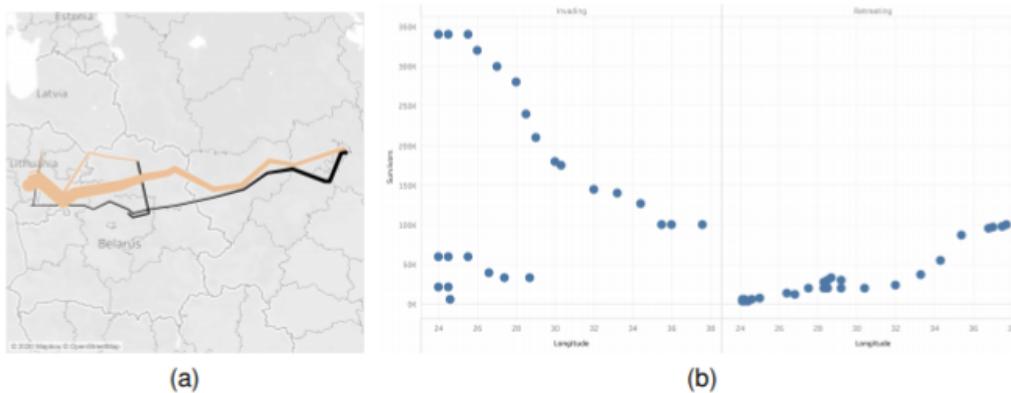


Figure 1: Two re-creations of Charles Minard’s map of the invasion of Russia in Tableau. In (1a) we hew as closely as possible to Minard’s design. In (1b) we attempt to encode all of the same data about the three groups of the invasion force using the most efficient channel: position on a common axis. In which circumstances might we prefer (1a) over (1b)?

## ABSTRACT

A central tenet of information visualization research and practice is the notion of visual variable *effectiveness*, or the perceptual precision at which values are decoded given visual channels of encoding. Formative work from Cleveland & McGill has shown that position along a common axis is the most effective visual variable for comparing individual values. One natural conclusion is that *any* chart that is not a dot plot or scatterplot is deficient and should be avoided. In this paper we refute a caricature of this “scatterplots only” argument as a way to call for new perspectives on how information visualization is researched, taught, and evaluated.

**Index Terms:** Human-centered computing—Visualization—Visualization theory, concepts and paradigms

## 1 INTRODUCTION

This example highlights an apparent contradiction at the heart of information visualization. On one hand, our exemplars of good visualizations can be diverse, complex, and reward contemplation [19]. On the other hand, our foundational empirical results and rules of thumb are often simple and minimalist. These rules are typically evaluated in terms of how quickly and accurately people extract specific information from charts, including formative psychophysical studies showing that viewers extract data values most precisely when they are encoded via position on shared axes [7].

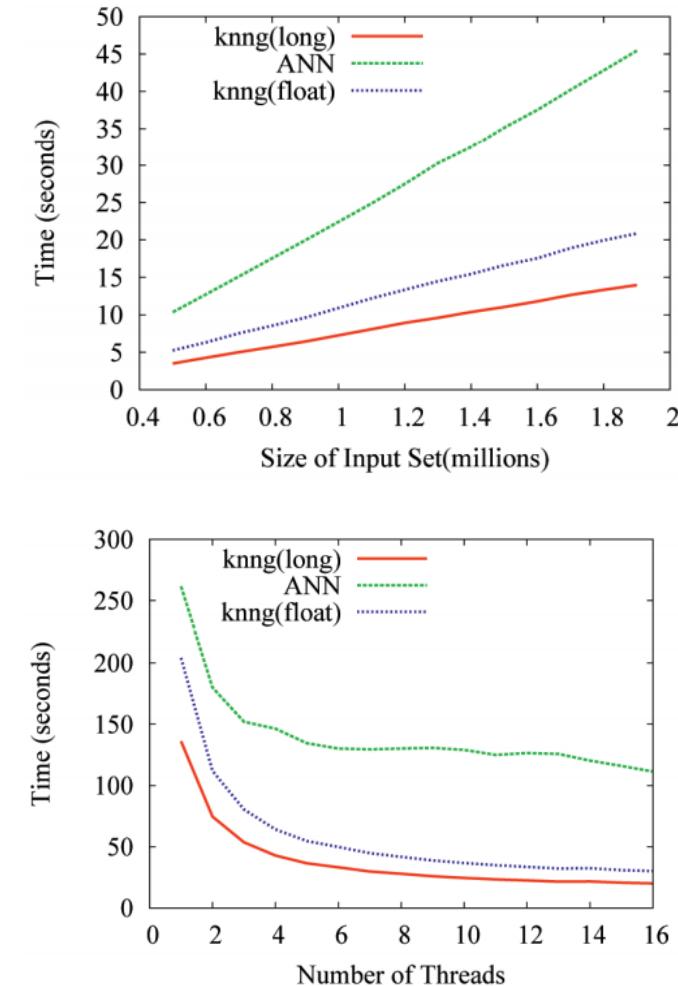
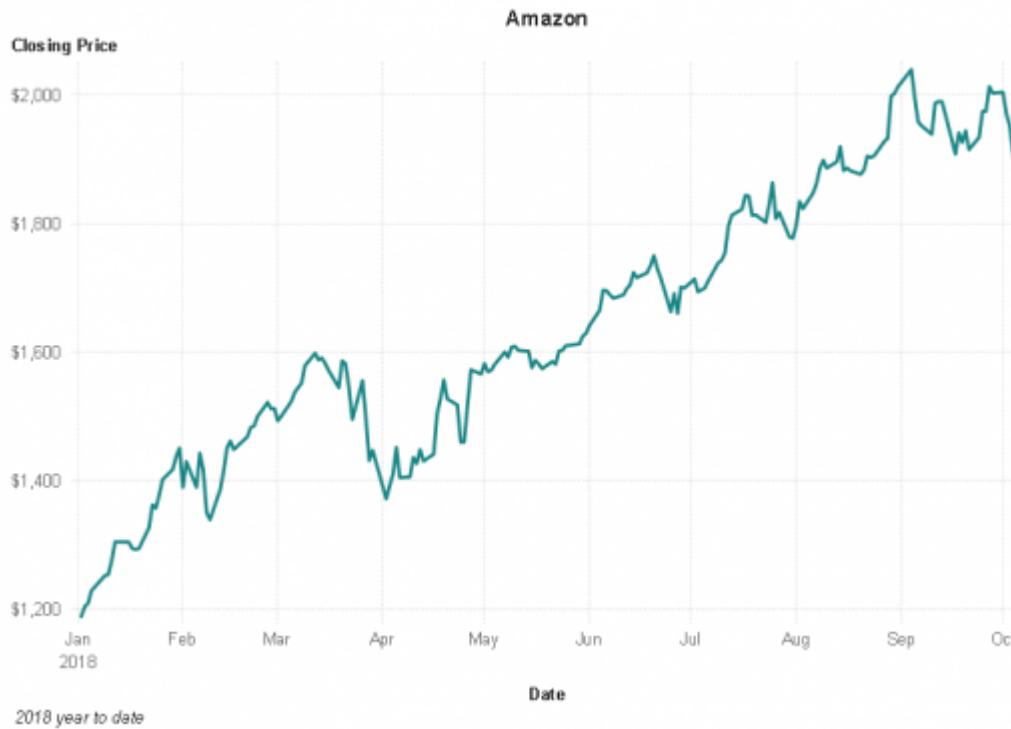
Given these constraints, a natural conclusion is that *quantitative data should almost always be depicted in a dot plot or scatter plot*, perhaps breaking data into SPLOMs, small multiples, or employing brushing and linking when there are too many variables for one view. While this argument is a strawman, its premises lie at the heart of foundational visualizations books by authors like Bertin and Tufte,

# THE LINE CHART



# THE LINE CHART

---

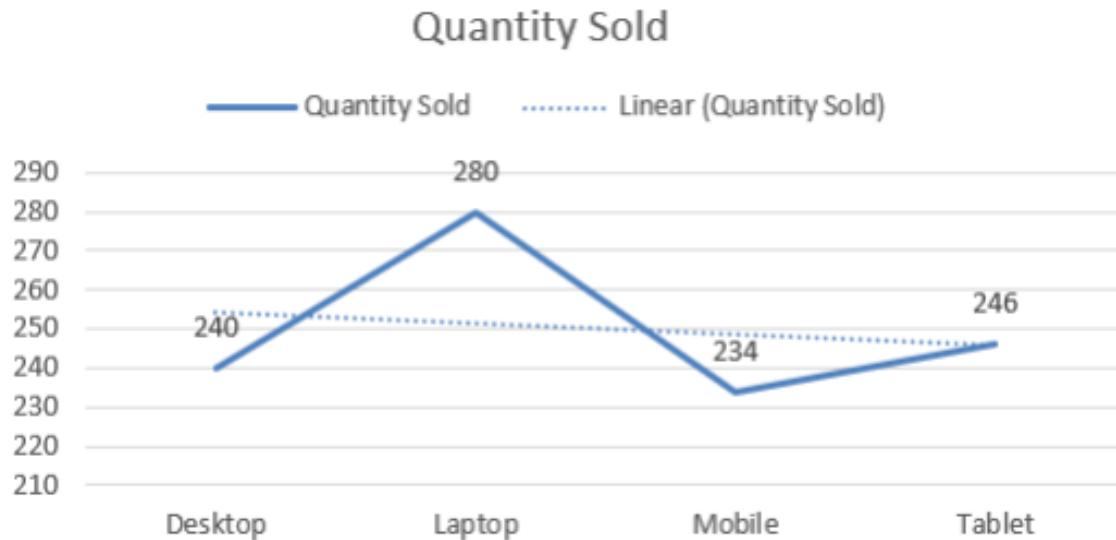


# THE LINE CHART

---

Considerations:

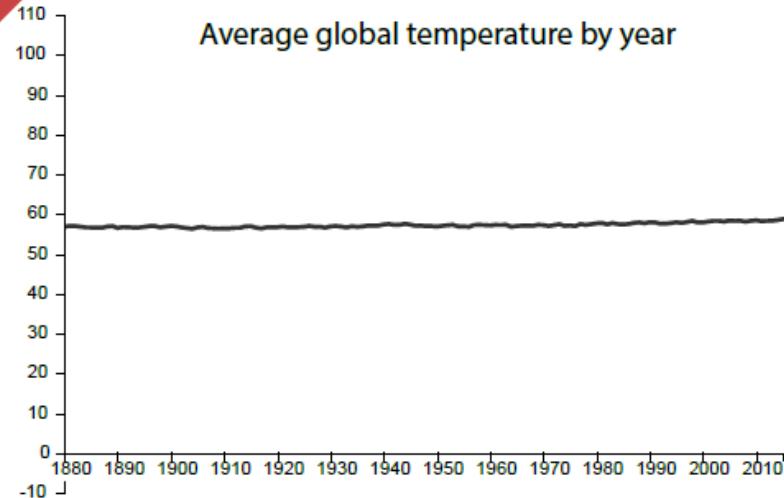
- Does it make sense?



# THE LINE CHART

---

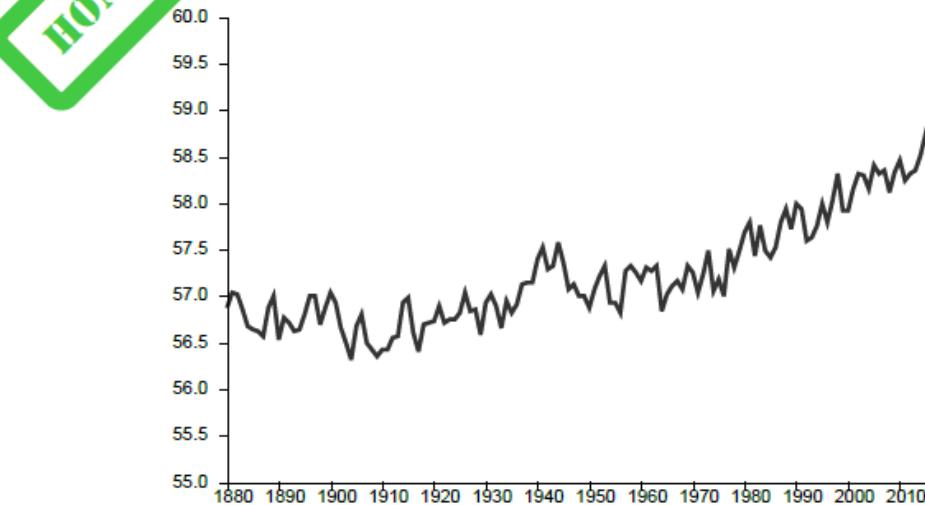
DECEPTIVE



Considerations:

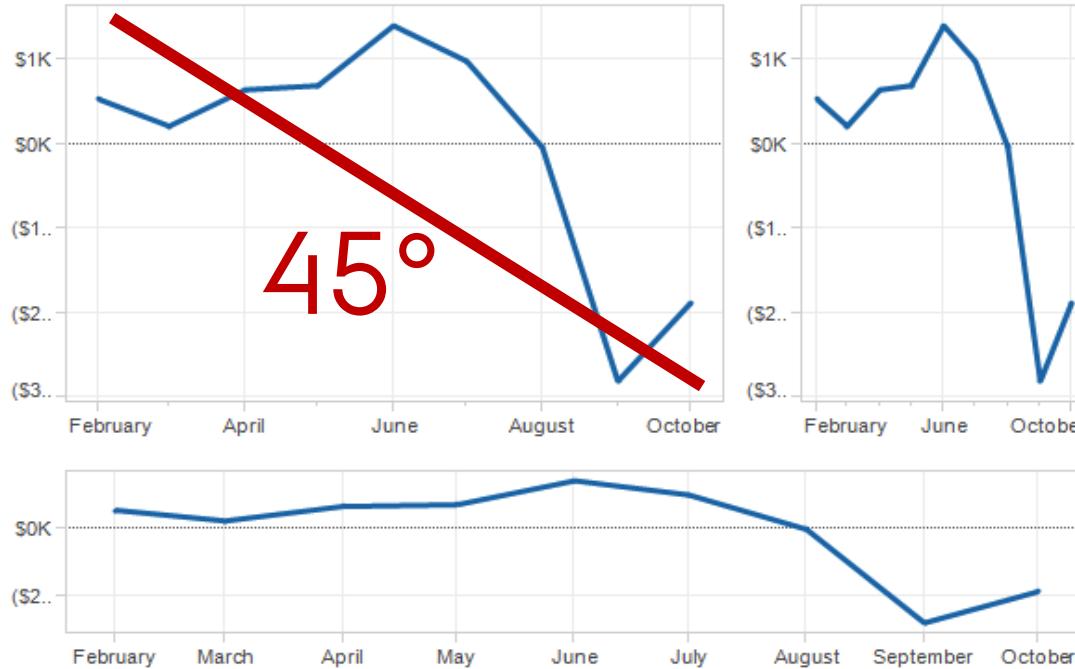
- Does it make sense?
- How to scale the Y-axis?

HONEST



# THE LINE CHART

---



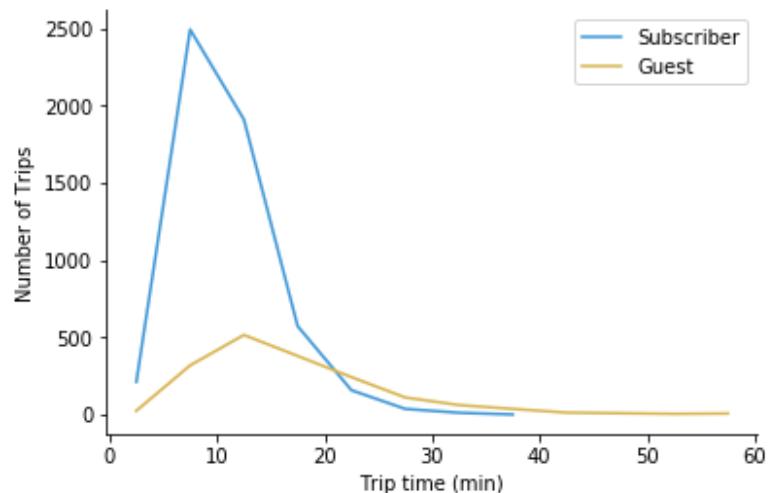
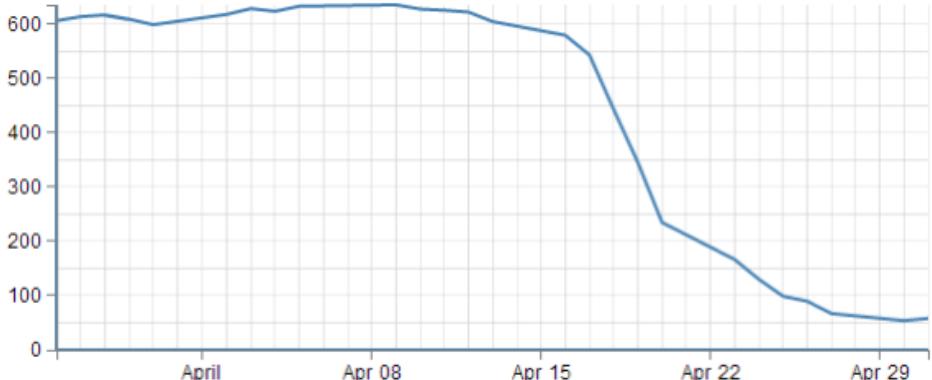
Considerations:

- Does it make sense?
- How to scale the Y-axis?
- What's the right aspect ratio?

Banking to  $45^\circ$   
[Cleveland et al. 1988]

# THE LINE CHART

---



## Considerations:

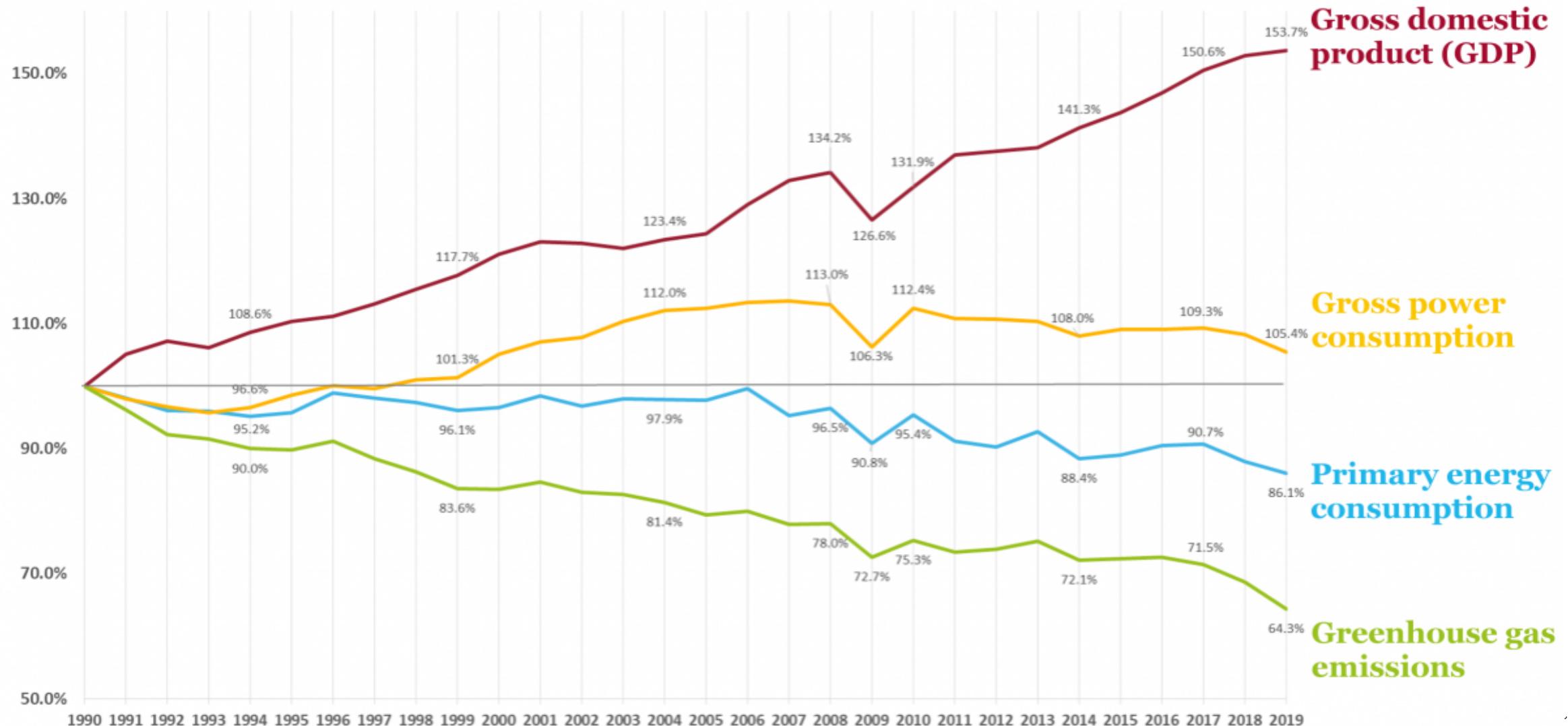
- Does it make sense?
- How to scale the Y-axis?
- What's the right aspect ratio?
- Should I use grid lines?



# Economic growth, power & energy consumption, GHG emissions 1990 - 2019.

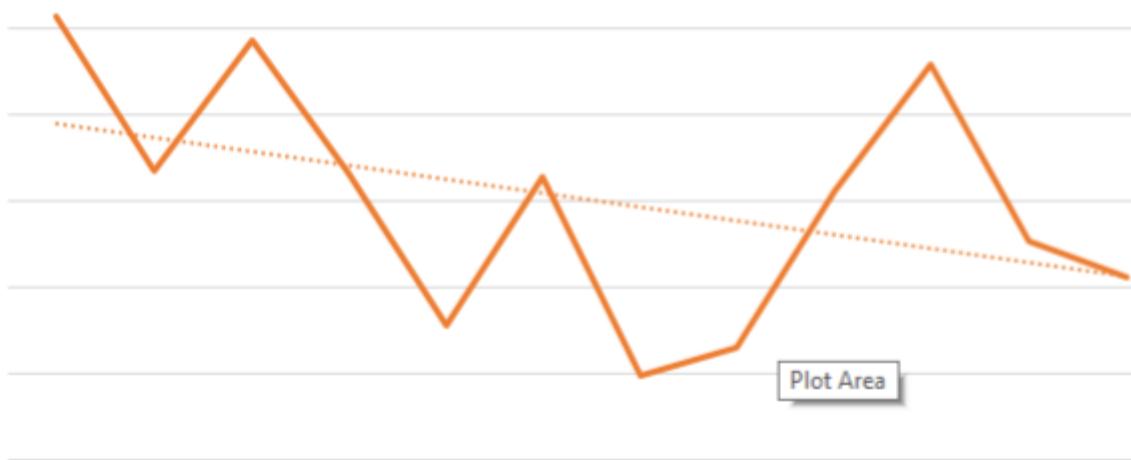
Data: BMWi 2020, UBA 2020.

CLEAN  
ENERGY  
WIRE



# THE LINE CHART

---



Considerations:

- Does it make sense?
- How to scale the Y-axis?
- What's the right aspect ratio?
- Should I use grid lines?
- Should I add a trend line?



# THE LINE CHART

---



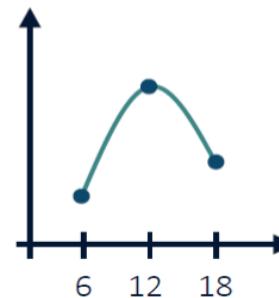
## Considerations:

- Does it make sense?
- How to scale the Y-axis?
- What's the right aspect ratio?
- Should I use grid lines?
- Should I add a trend line?
- Which line style to use?

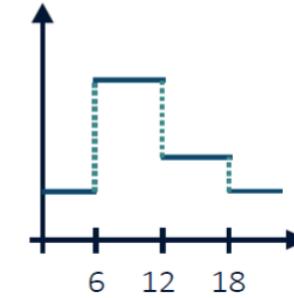
Straight Line



Curved Line

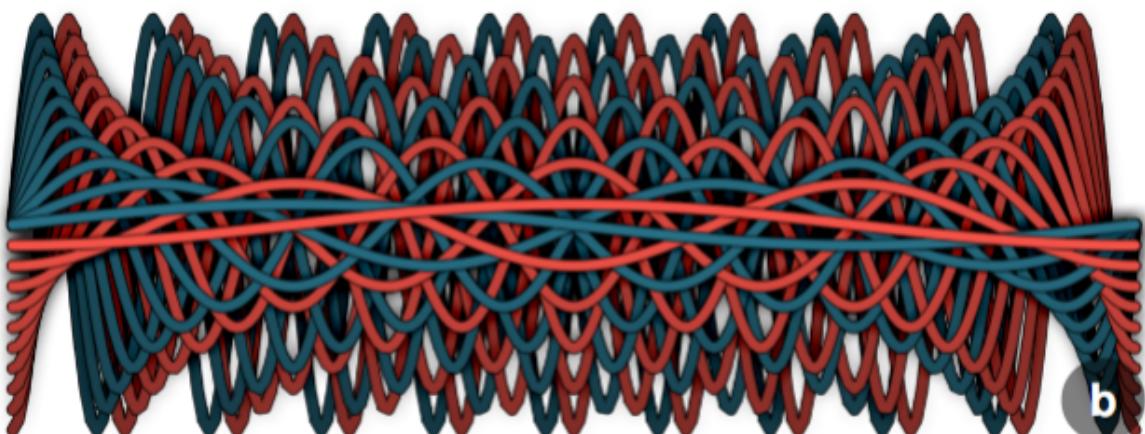
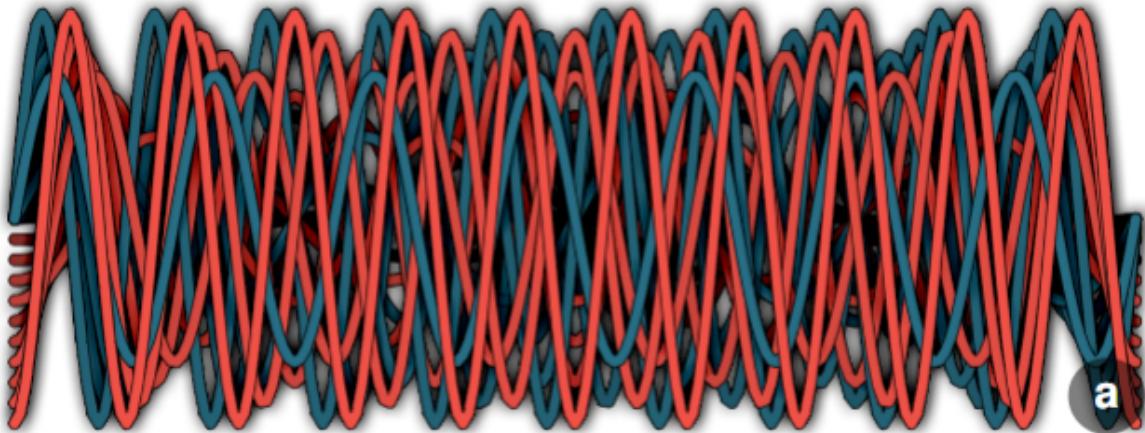


Flat Line



# THE LINE CHART

---



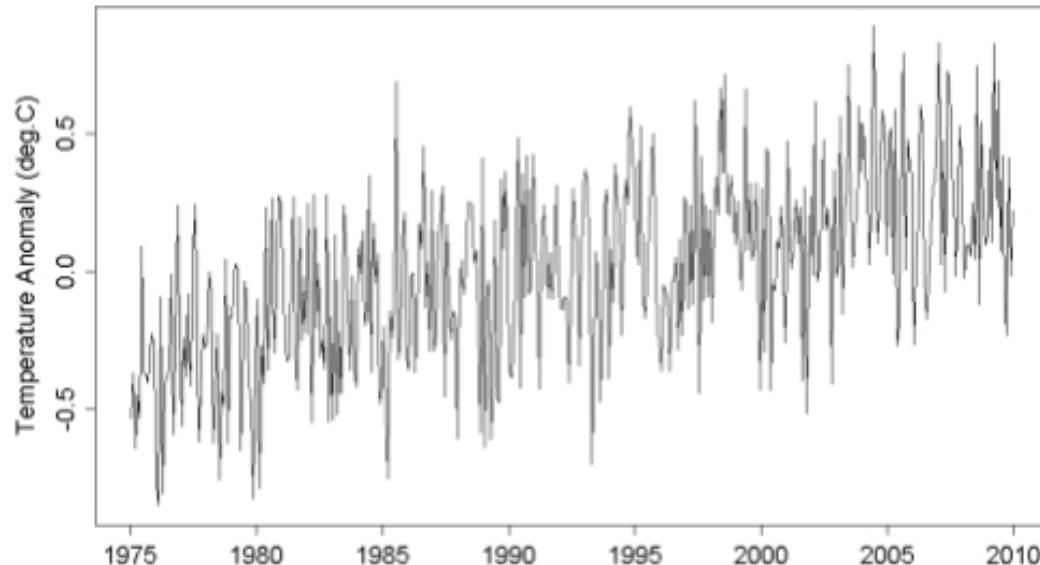
Considerations:

- Does it make sense?
- How to scale the Y-axis?
- What's the right aspect ratio?
- Should I use grid lines?
- Should I add a trend line?
- Which line style to use?
- In which order to plot multiple lines?



# THE LINE CHART

---



Considerations:

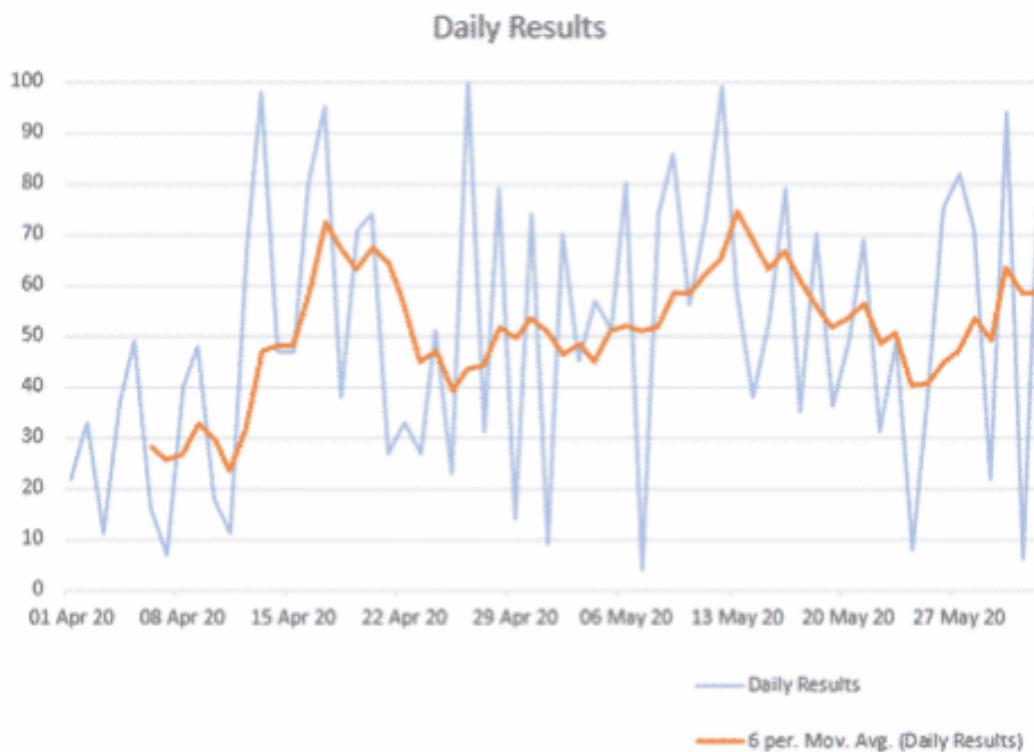
- Does it make sense?
- How to scale the Y-axis?
- What's the right aspect ratio?
- Should I use grid lines?
- Should I add a trend line?
- Which line style to use?
- In which order to plot multiple lines?
- How to deal with noisy data?

1. Don't deal with it at all.



# THE LINE CHART

---

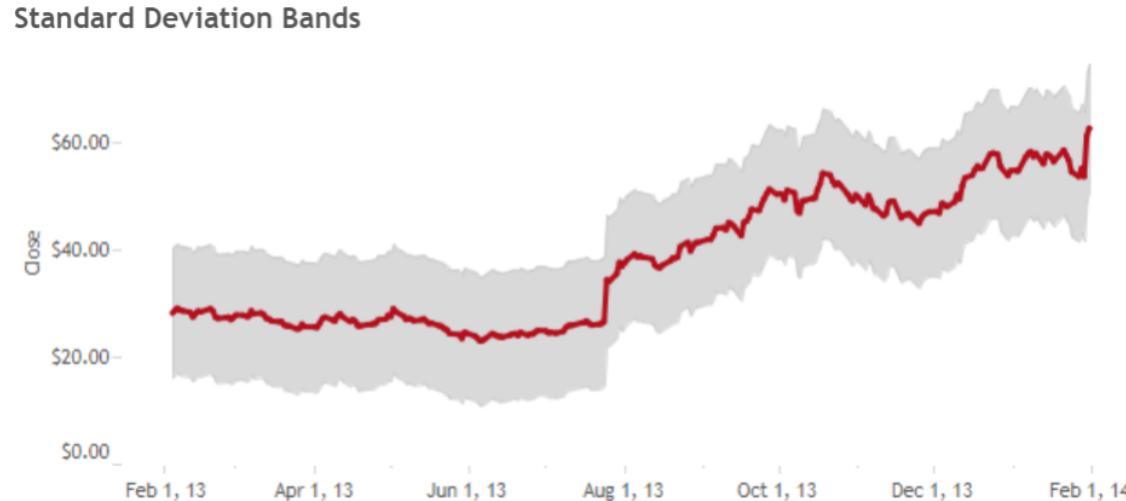


## Considerations:

- Does it make sense?
- How to scale the Y-axis?
- What's the right aspect ratio?
- Should I use grid lines?
- Should I add a trend line?
- Which line style to use?
- In which order to plot multiple lines?
- How to deal with noisy data?
  1. Don't deal with it at all.
  2. Smooth the data (Aggregation)

# THE LINE CHART

---



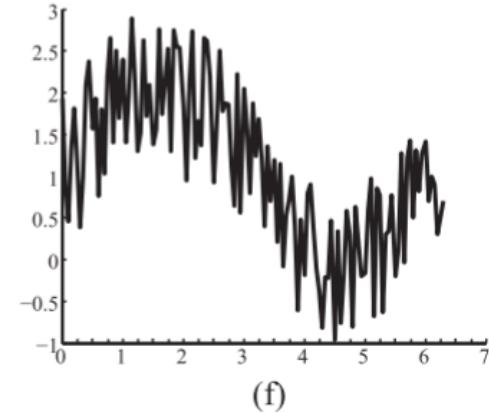
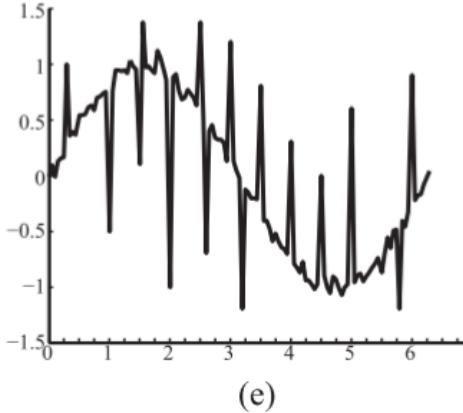
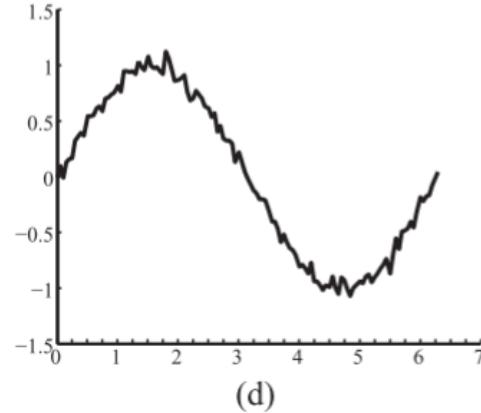
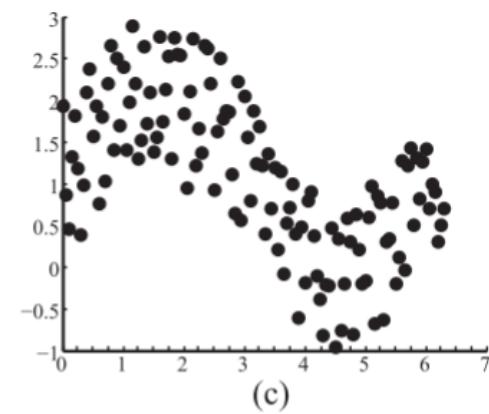
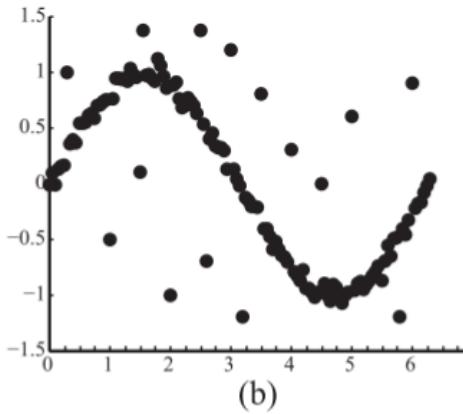
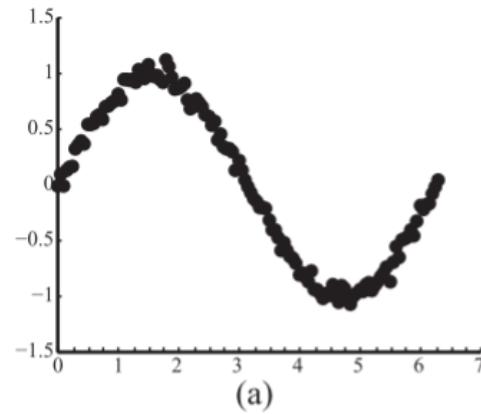
## Considerations:

- Does it make sense?
- How to scale the Y-axis?
- What's the right aspect ratio?
- Should I use grid lines?
- Should I add a trend line?
- Which line style to use?
- In which order to plot multiple lines?
- How to deal with noisy data?
  1. Don't deal with it at all.
  2. Smooth the data (Aggregation)

# SCATTERPLOTS VS LINE CHARTS

---

Image: 10.1109/TVCG.2017.2653106

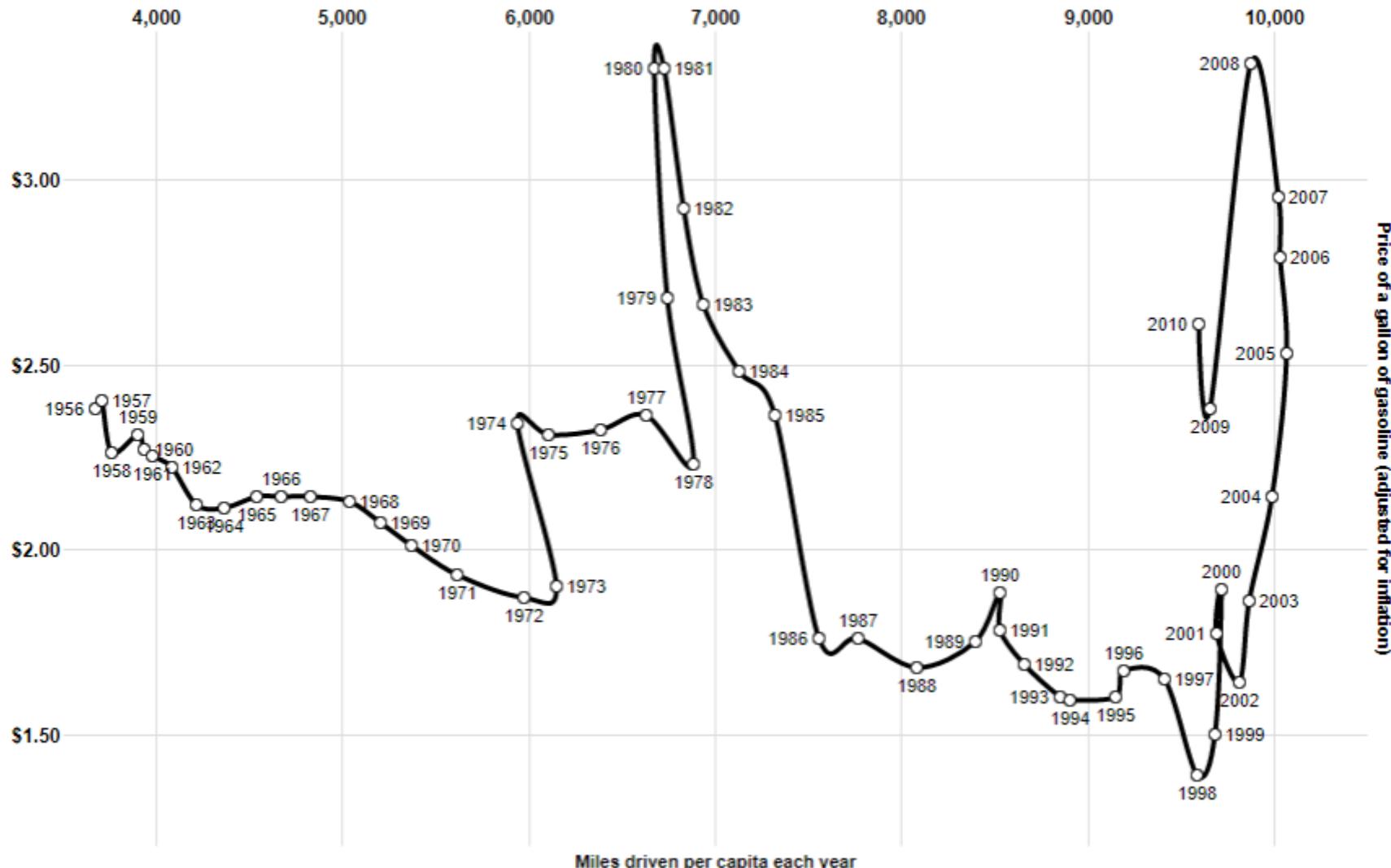


Less noise, clear trend  
=> line charts recommended

More noise, unclear trend  
=> scatterplots recommended



# THE CONNECTED SCATTERPLOT



Price of a gallon of gasoline (adjusted for inflation)



# THE AREA CHART

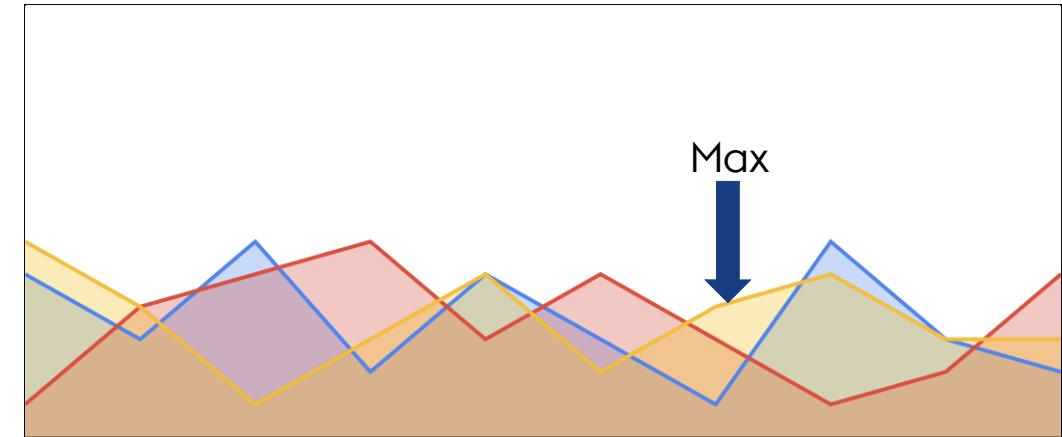


# THE AREA CHART

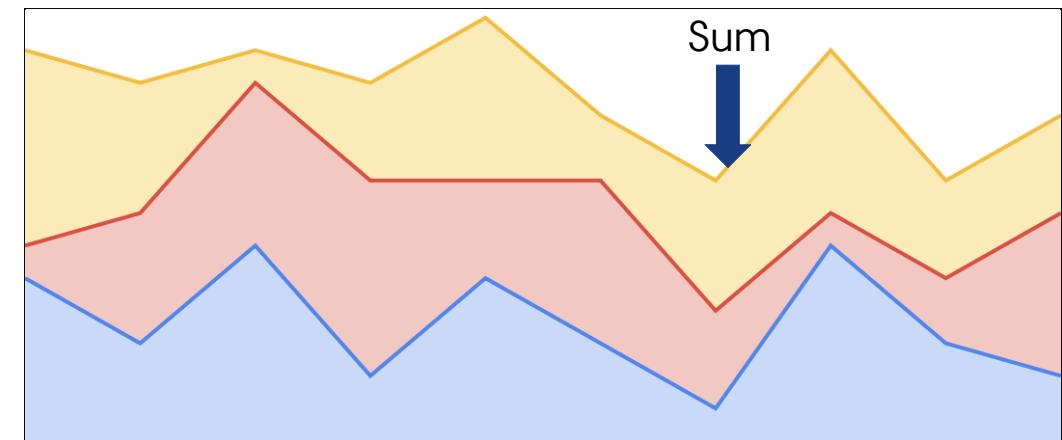
---

Fundamental Analysis of Stocks

Price Movements



"Ordinary" Area Chart

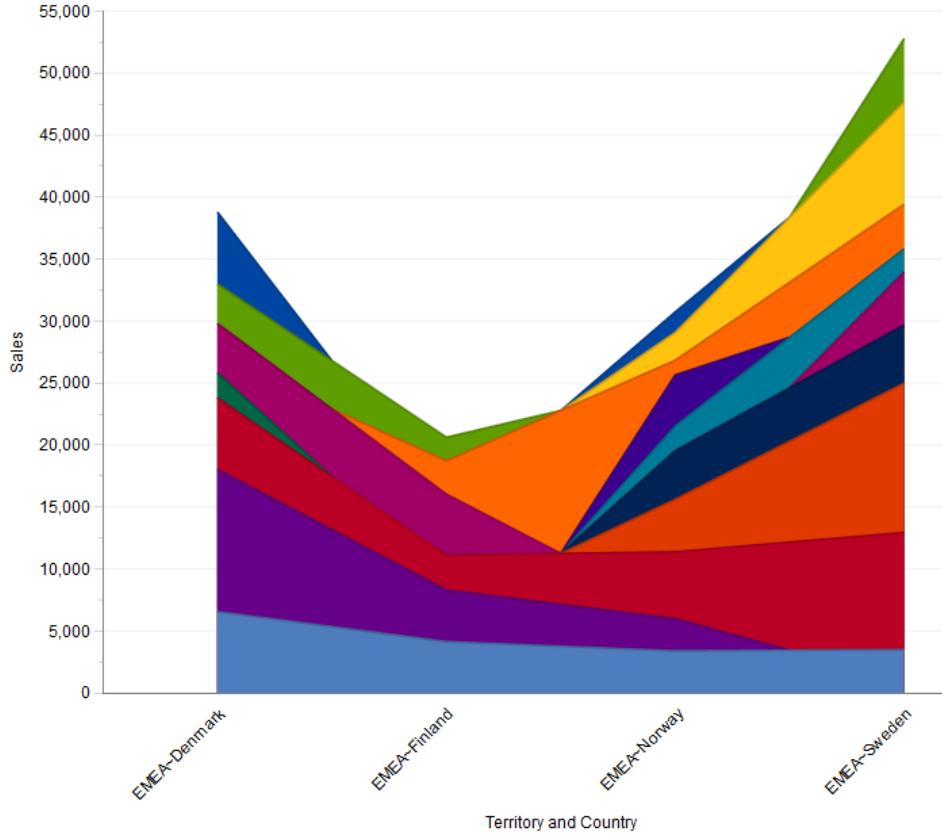


Stacked Area Chart



# THE AREA CHART

---



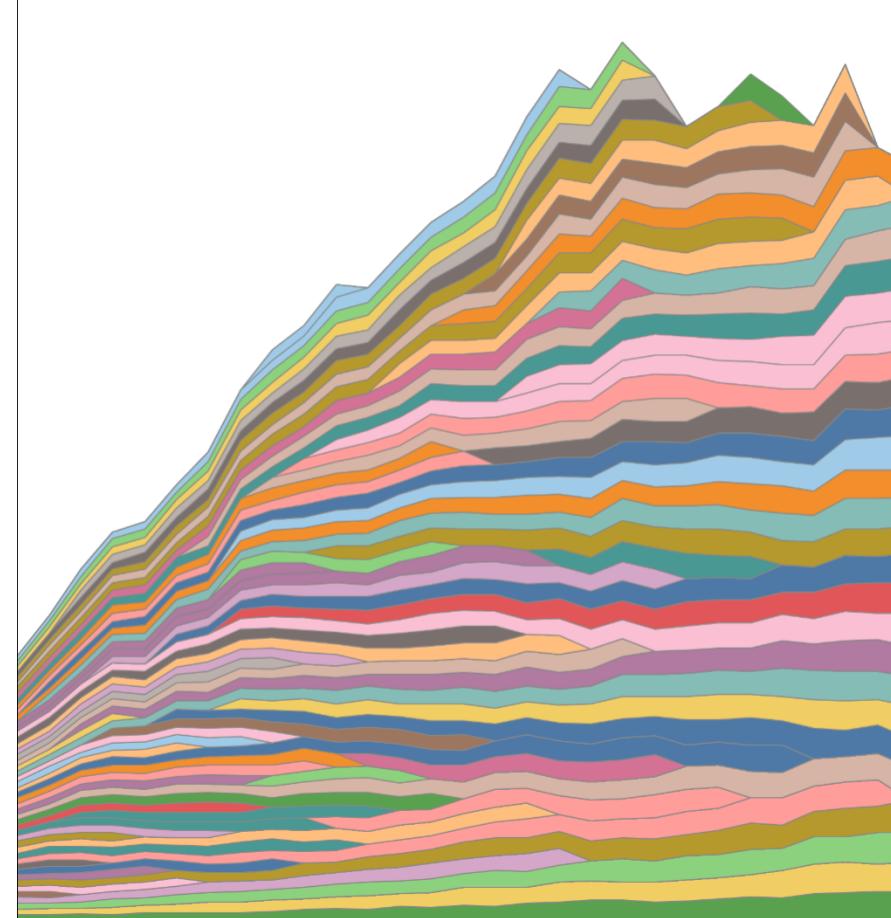
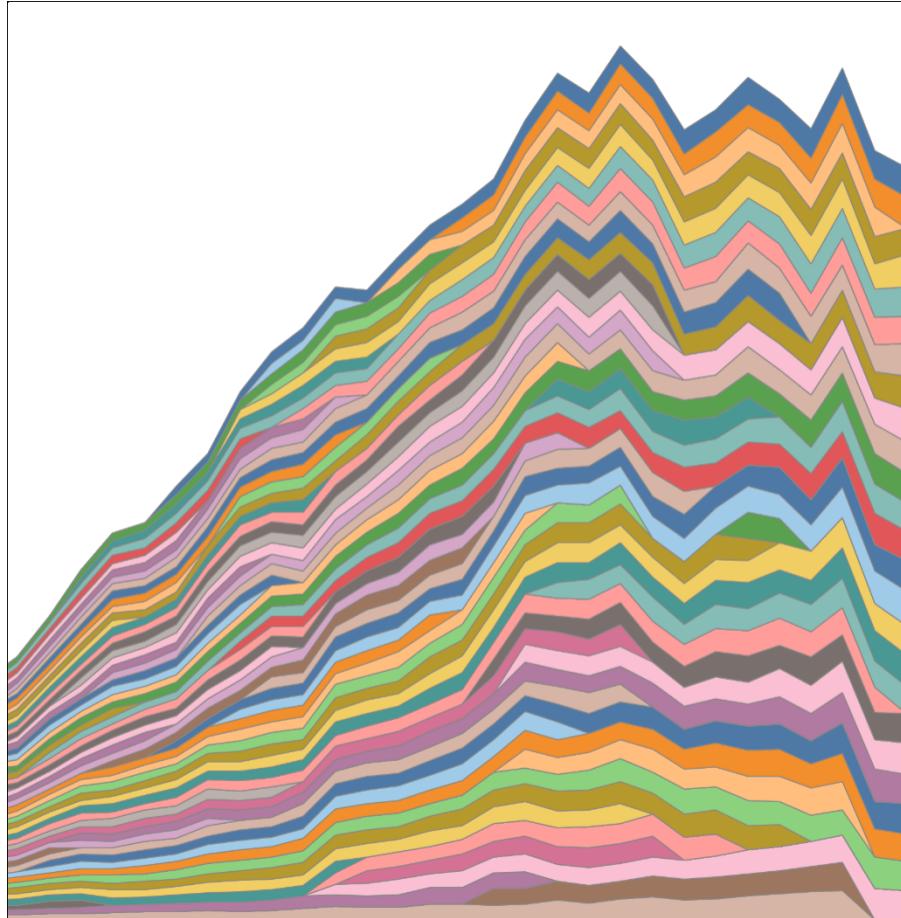
## Considerations:

- Most of those for line charts
- For stacked area charts: Order



# THE EFFECT OF DIFFERENT ORDERINGS

---

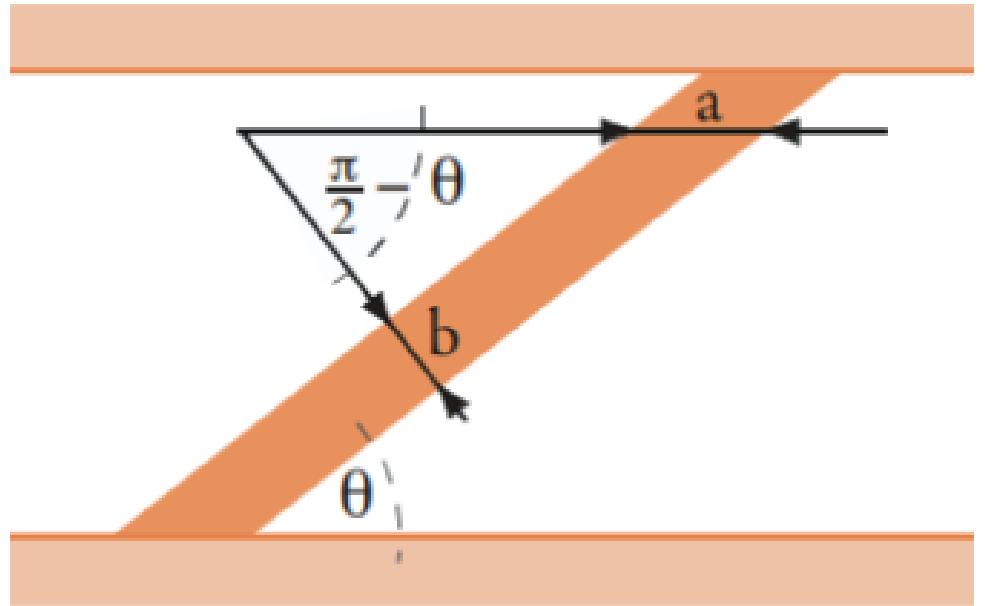


Images Courtesy of Steffen Strunge Mathiesen

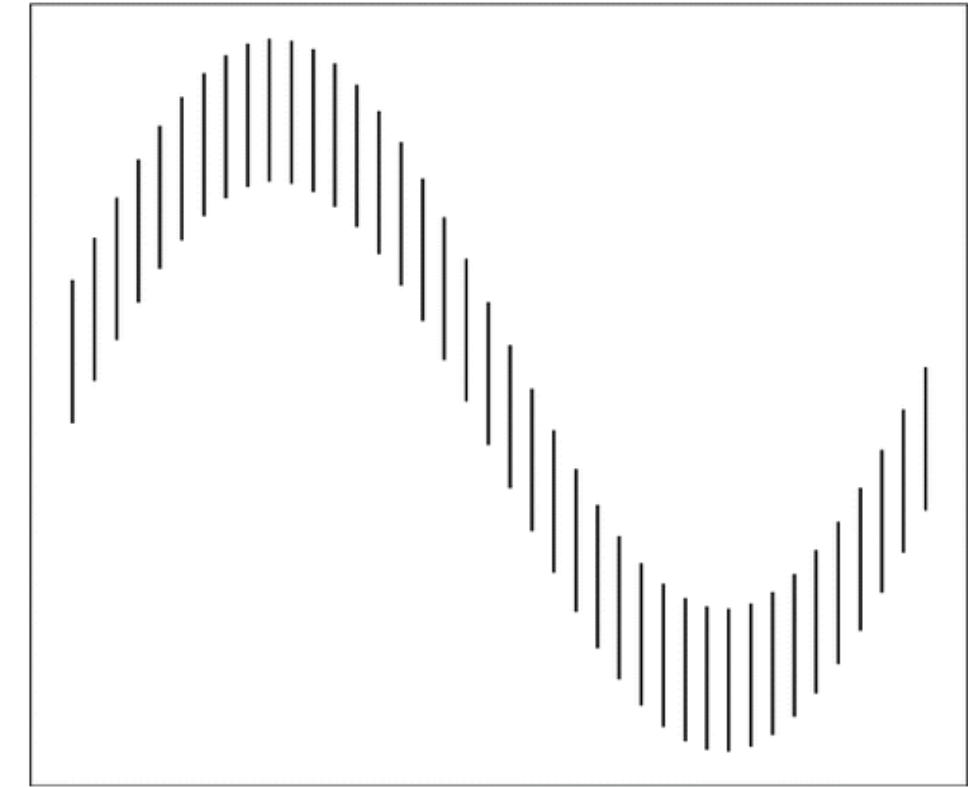


# COGNITIVE DISTORTION

---



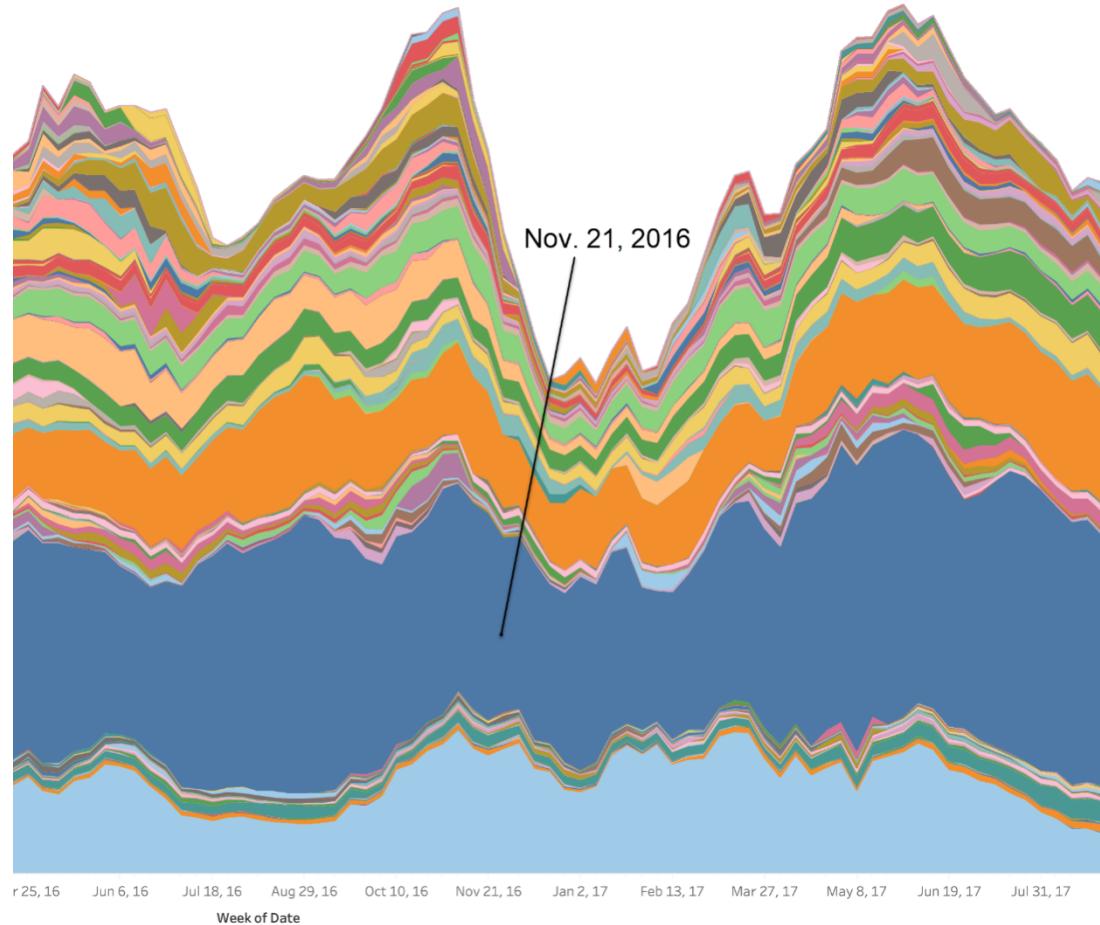
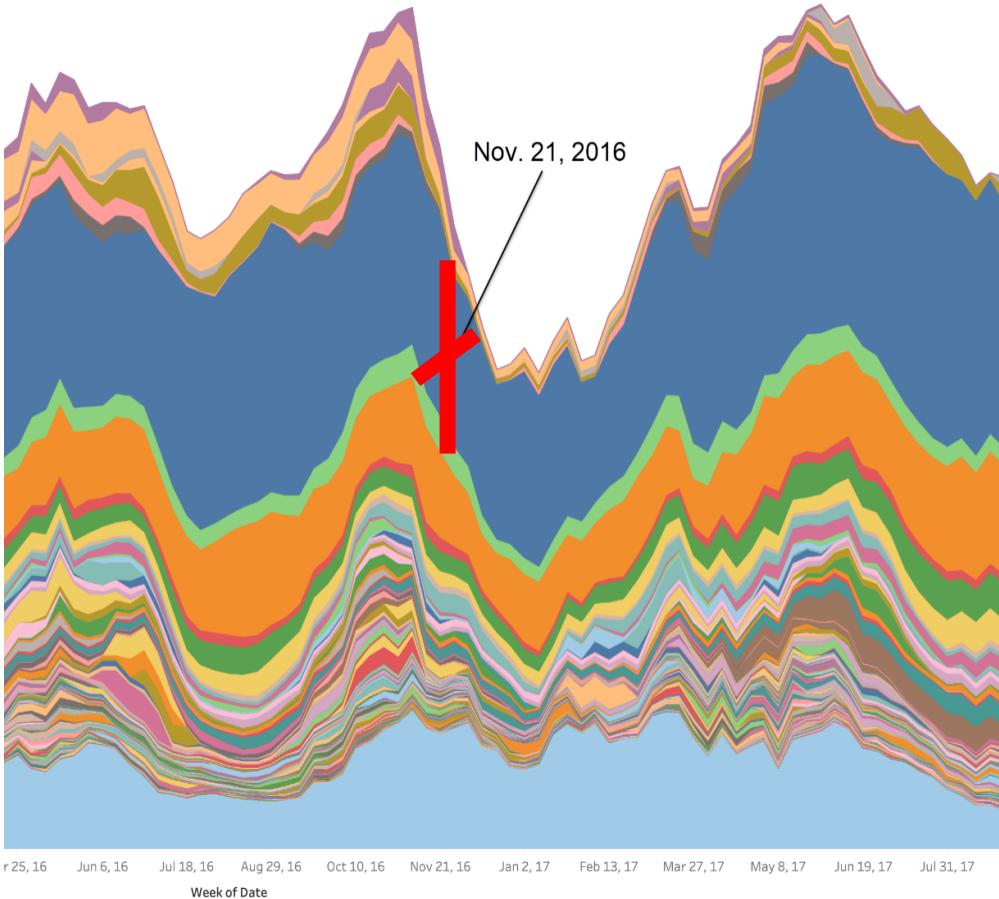
Line Width Illusion



Sine Illusion

# REASON: LINE WIDTH ILLUSION (ALSO: LABELS)

---



## Algorithm 1 UpwardsOpt

```
1: procedure UPWARDSOPT(init, minImpr)
2:   order  $\leftarrow$  init
3:   repeat
4:     oldCost  $\leftarrow$  costchart(order)    wiggle, bump, ...
5:     i  $\leftarrow$  0
6:     done  $\leftarrow$  []
7:     while i  $<$  length(order) do
8:       if i  $\in$  done then
9:         i ++
10:        continue
11:       end if
12:
13:       Remove current
14:       layer from stack,
15:       find best position,
16:       and reinsert it
17:
18:       Lots of
19:       bookkeeping to
20:       make sure we only
21:       optimize each
22:       layer once
23:
24:       fi  $\leftarrow$  pop(order, i)
25:       newIdx  $\leftarrow$  FindBestPosition(order, fi)
26:       insert(order, newIdx, fi)
27:
28:       for d = 0 to length(done) - 1 do
29:         if (newIdx > i) && (done[d]  $\in$  [i ... newIdx]) then
30:           done[d] --           $\triangleright$  layer i moved up
31:         end if
32:         if (newIdx < i) && (done[d]  $\in$  [newIdx ... i]) then
33:           done[d] ++           $\triangleright$  layer i moved down
34:         end if
35:       end for
36:       done.add(newIndex)
37:
38:     end while
39:     until costchart(order)  $\geq$  oldCost  $\cdot$  (100% - minImpr)
40:   end procedure
```

Repeat this until improvement  
is below parameter minImpr

## Algorithm 2 FindBestPosition

```
1: procedure FINDBESTPOSITION(order, fi)
2:   /* Preprocessing Stage */
3:   gBelow  $\leftarrow$  0
4:   gAbove  $\leftarrow$  fi
5:   costBelow, costAbove, costLayer  $\leftarrow$  []
6:   for pos = 0 to length(order) - 1 do
7:     costBelow.add(costlayer(order[pos], gBelow))
8:     costAbove.add(costlayer(order[pos], gAbove))
9:     costLayer.add(costlayer(fi, gBelow))
10:    gBelow  $\leftarrow$  gBelow + order[pos]
11:    gAbove  $\leftarrow$  gAbove + order[pos]
12:   end for
13:   costLayer.add(costlayer(fi, gBelow))
14:
15:   /* Testing Stage */
16:   currentCost  $\leftarrow$  costLayer[0] + suml=0j-2 costAbove[l]
17:   bestIndex  $\leftarrow$  0, bestCost  $\leftarrow$  currentCost
18:   for pos = 1 to length(order) - 1 do
19:     currentCost += costBelow[pos - 1]
20:     currentCost -= costAbove[pos - 1]
21:     currentCost += costLayer[pos]
22:     currentCost -= costLayer[pos - 1]
23:     if currentCost < bestCost then
24:       bestIndex  $\leftarrow$  pos, bestCost  $\leftarrow$  currentCost
25:     end if
26:   end for
27:   return bestIndex
28: end procedure
```

Precompute  
three types of  
**layer** costs:  
above and  
below the  
layer to be  
positioned,  
and the layer  
placed at this  
position itself

Test all  
possible  
positions  
using these  
precomputed  
costs

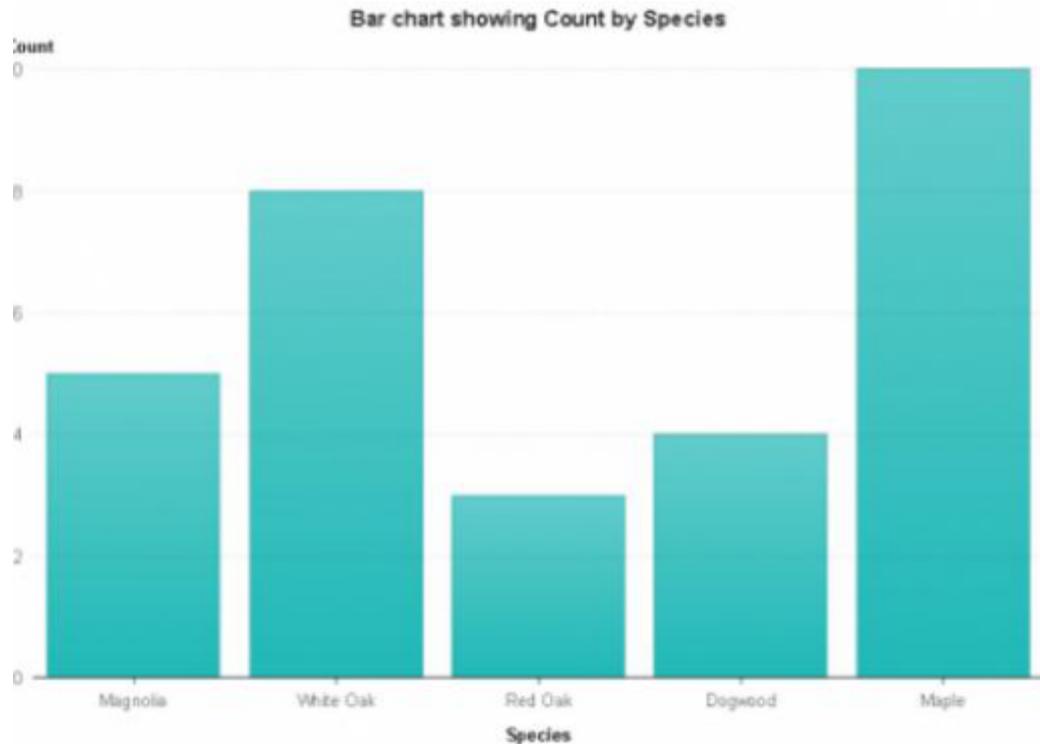


# THE BAR CHART

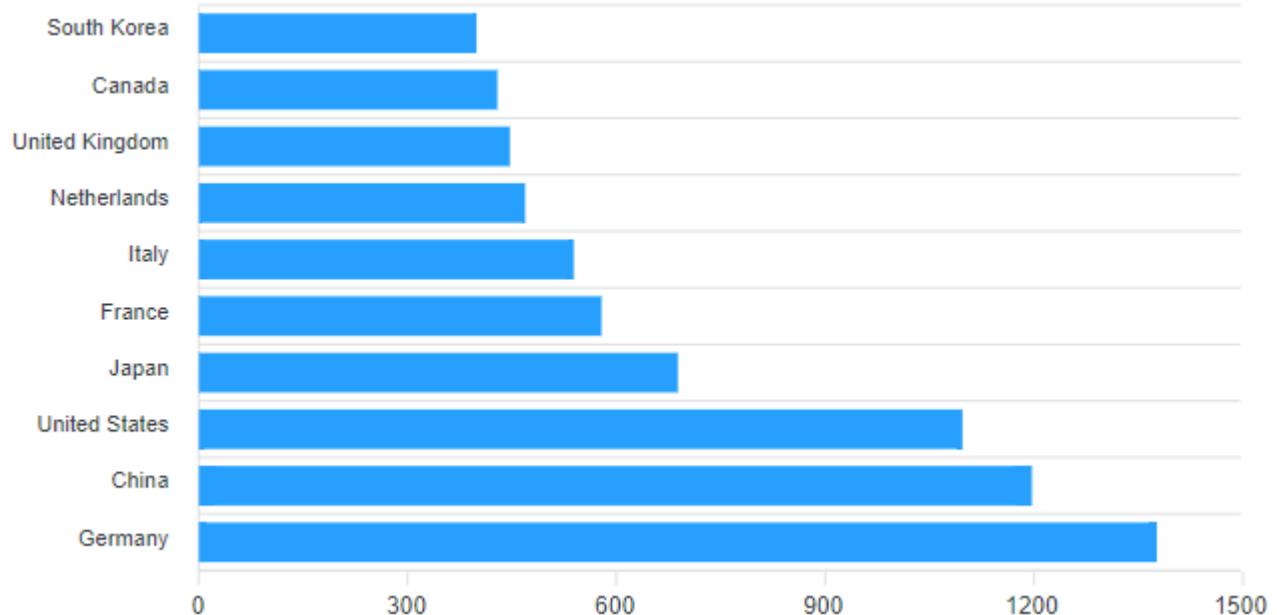


# THE BAR / COLUMN CHART

---



Column Chart



Bar Chart



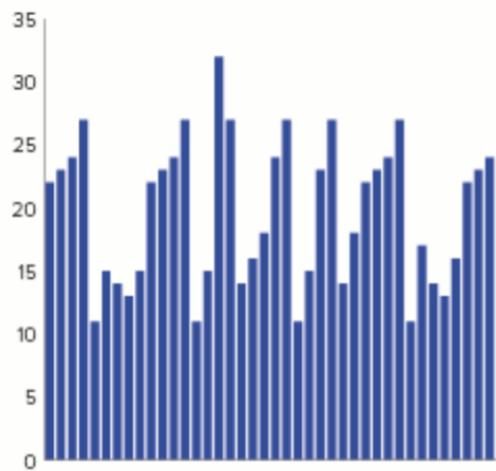
# THE BAR / COLUMN CHART

---

Considerations:

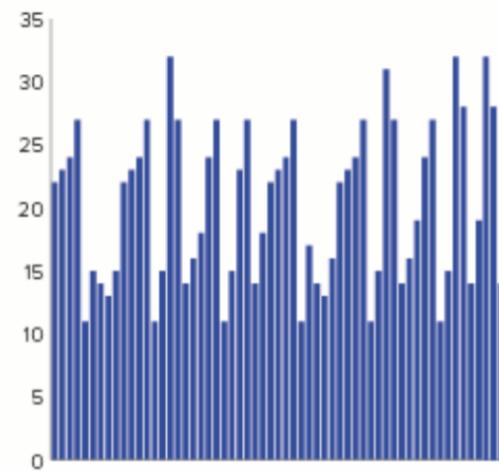
- Does it make sense?

40 bars



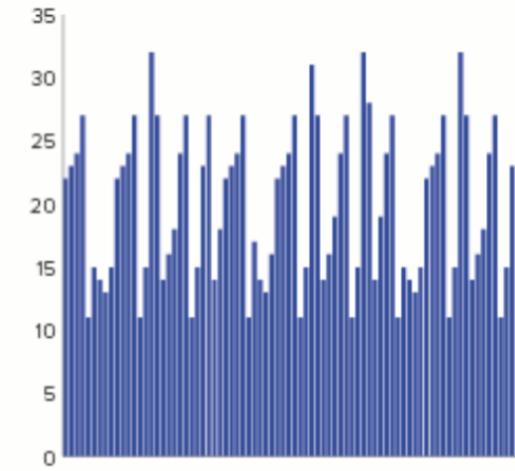
Pushing it

60 bars



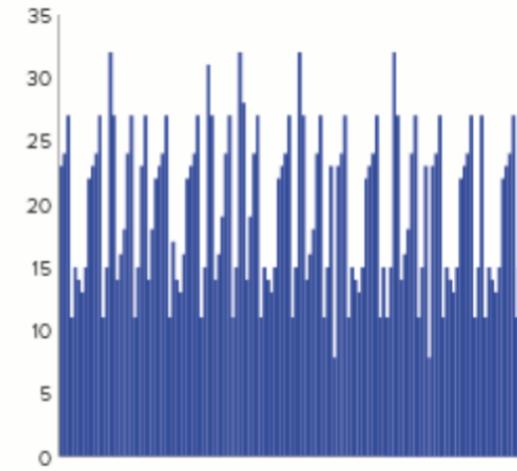
Starting to shimmer

80 bars



Becoming a blur

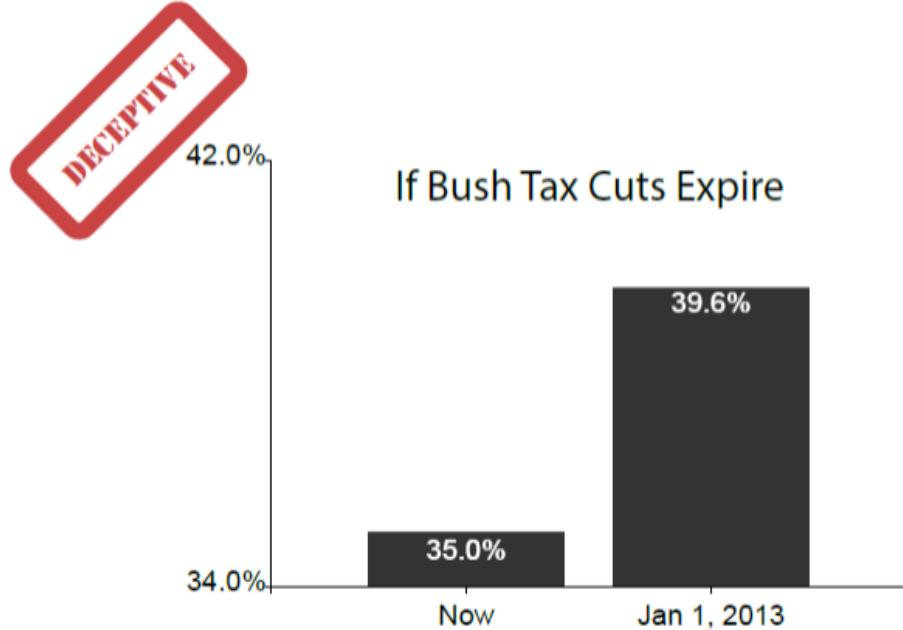
100+ bars



Not a bar chart anymore

<https://www.addtwodigital.com/add-two-blog/2021/6/16/rule-17-not-too-many-bars>

# THE BAR / COLUMN CHART



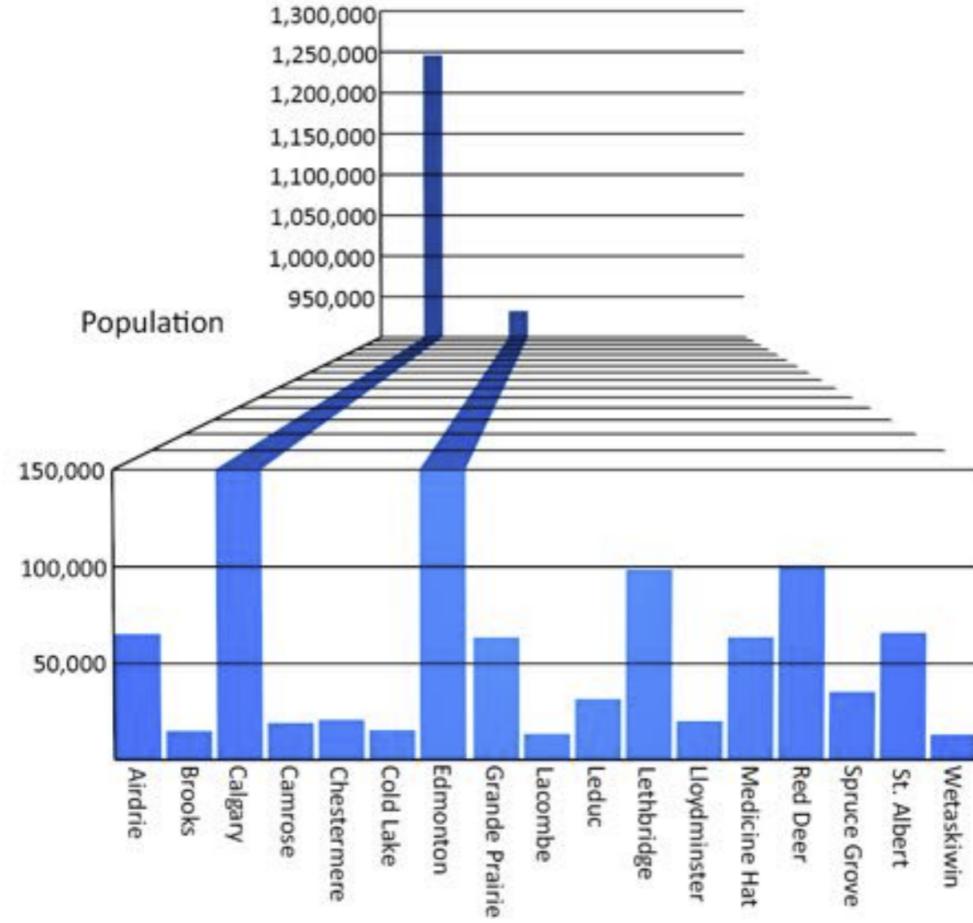
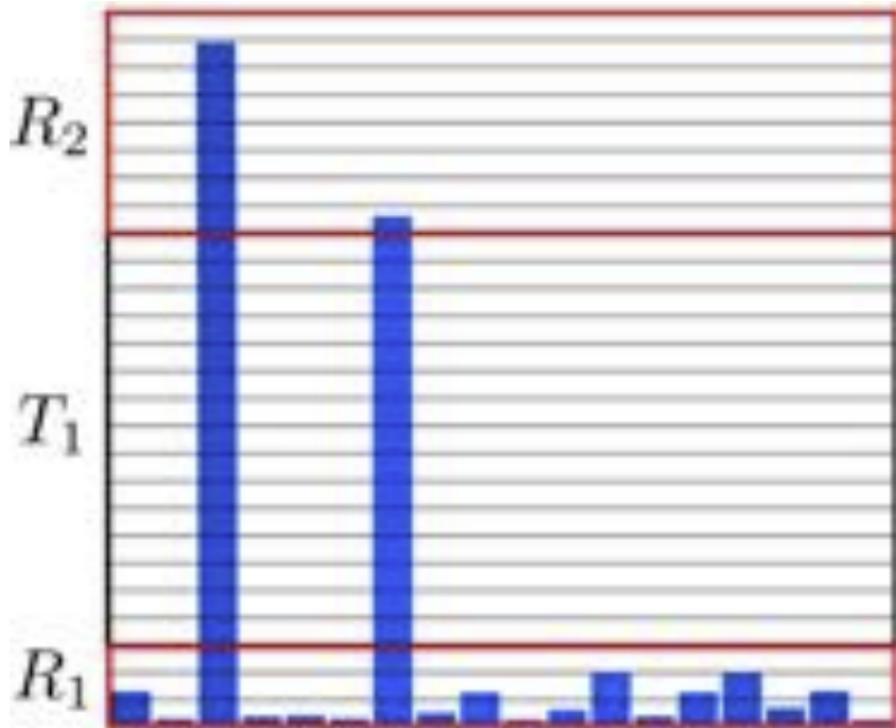
Considerations:

- Does it make sense?
- Always show full Y axis!



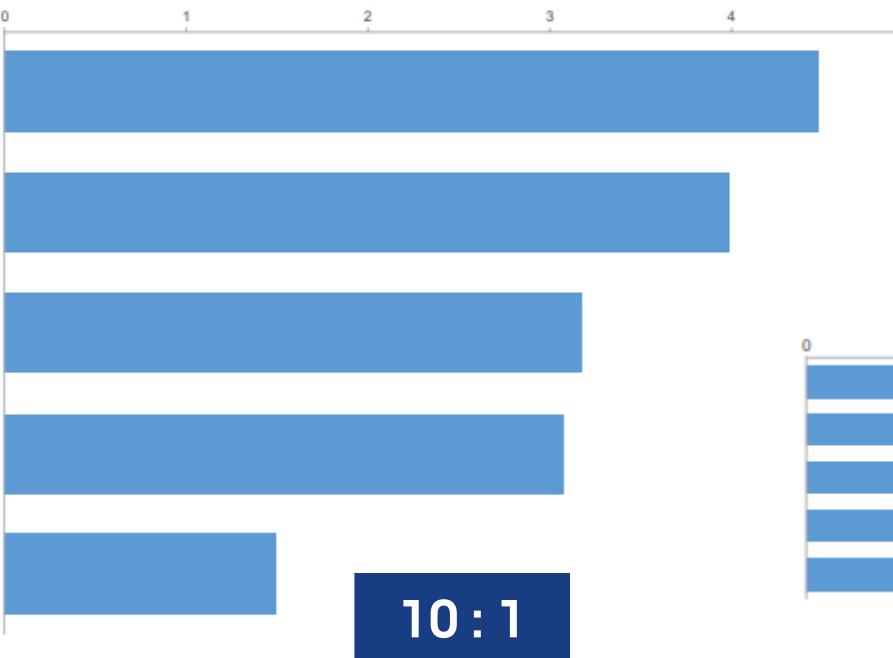
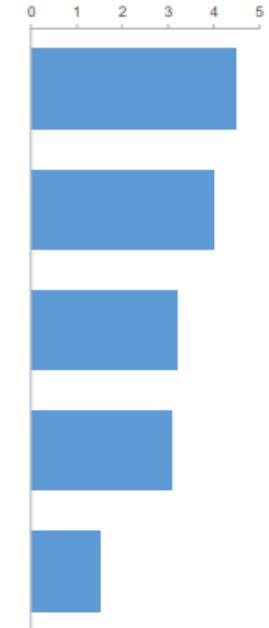
# THE BAR / COLUMN CHART

[MacTavish et al. 2021]



# THE BAR / COLUMN CHART

---

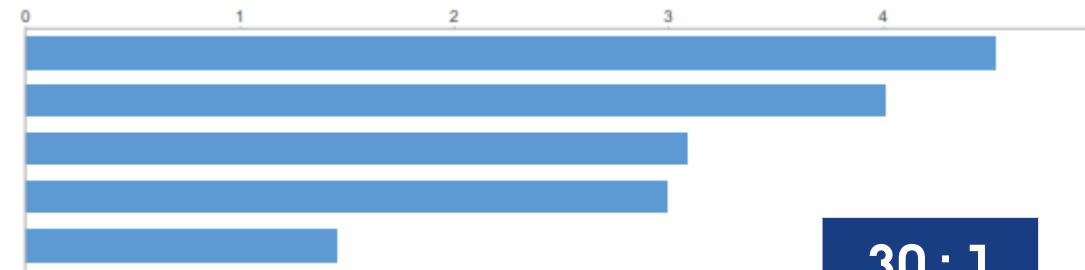


Considerations:

- Does it make sense?
- Always show full Y axis!
- Which aspect ratio to choose?  $\geq 10 : 1$

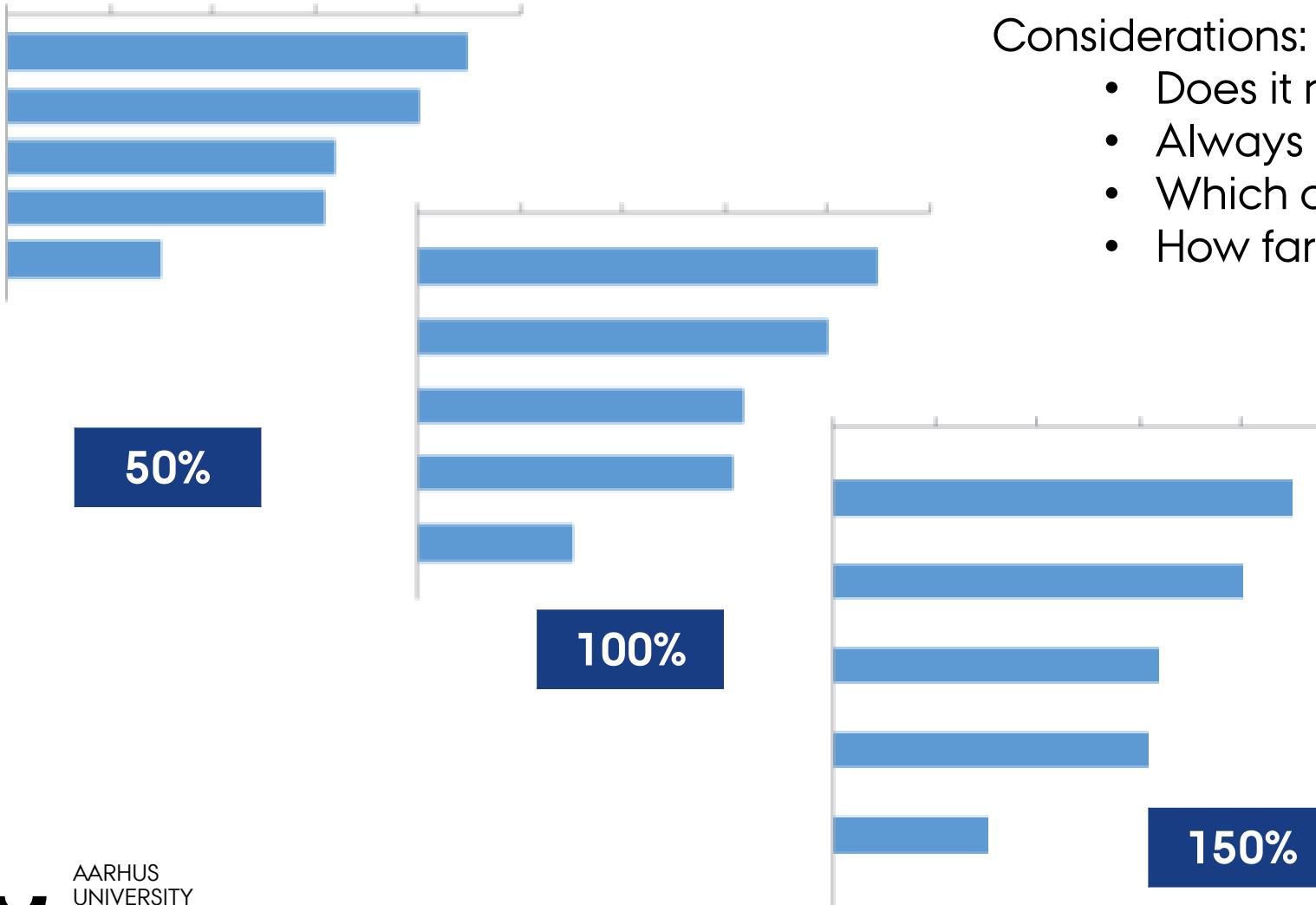
Source:

[https://www.perceptualedge.com/articles/visual\\_business\\_intelligence/bar\\_widths.pdf](https://www.perceptualedge.com/articles/visual_business_intelligence/bar_widths.pdf)



# THE BAR / COLUMN CHART

---



## Considerations:

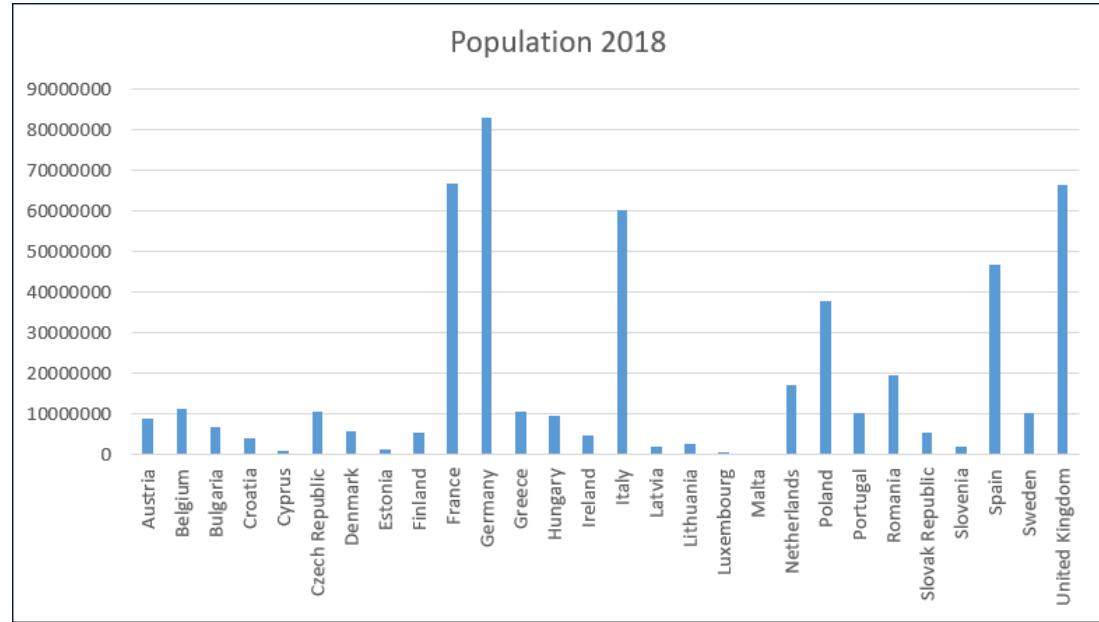
- Does it make sense?
- Always show full Y axis!
- Which aspect ratio to choose?
- How far to space the bars? 50%-75%

Source:

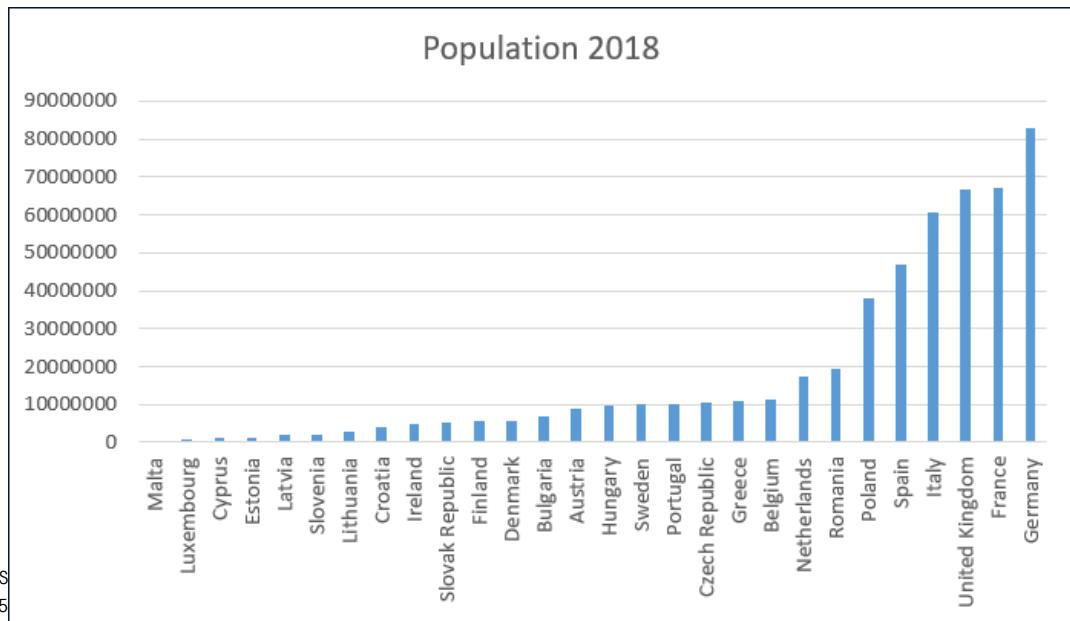
[https://www.perceptualedge.com/articles/visual\\_business\\_intelligence/bar\\_widths.pdf](https://www.perceptualedge.com/articles/visual_business_intelligence/bar_widths.pdf)

# THE BAR / COLUMN CHART

---

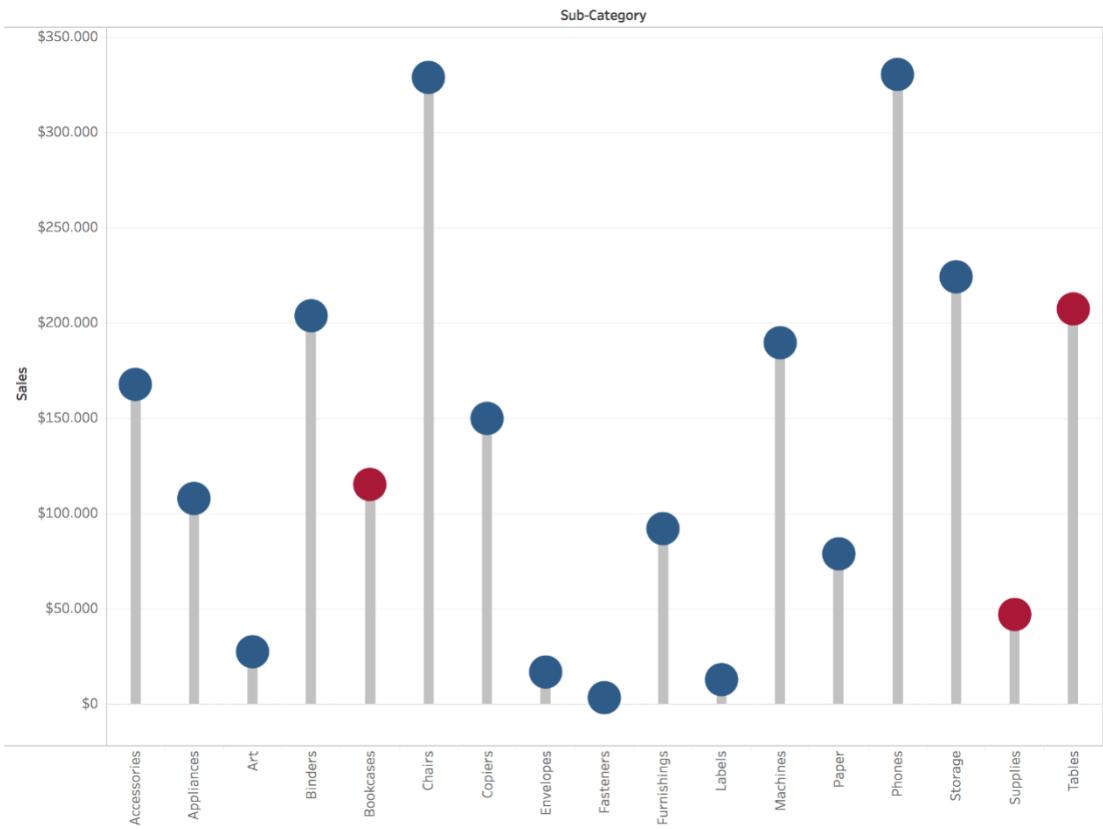


BASIC CHARTS  
DATAVIS FALL 2025



# THE BAR / COLUMN CHART

---



## Considerations:

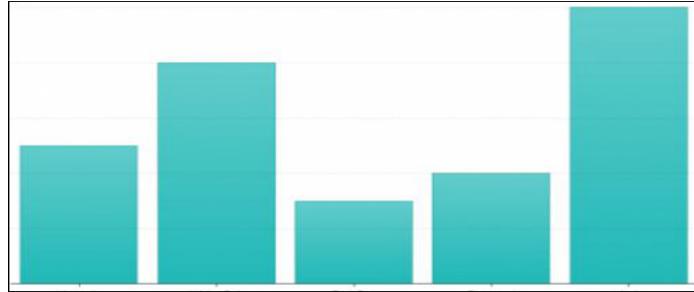
- Does it make sense?
- Always show full Y axis!
- Which aspect ratio to choose?
- How far to space the bars?
- Order of the categories
- Shape of the bars

Lollipop Chart

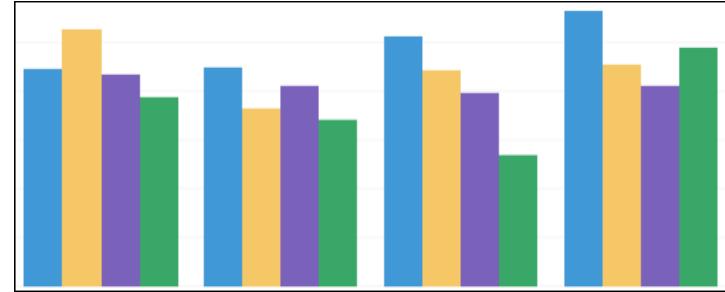


# GROUPED AND STACKED BAR CHARTS

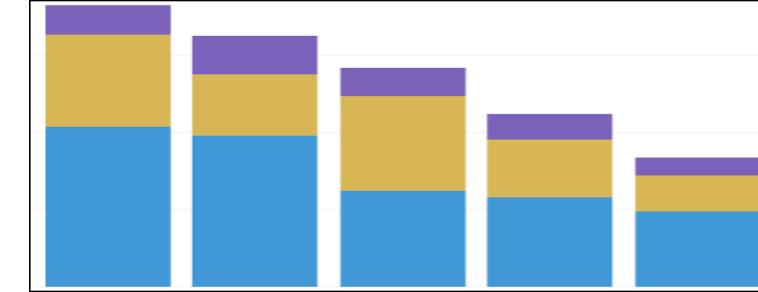
---



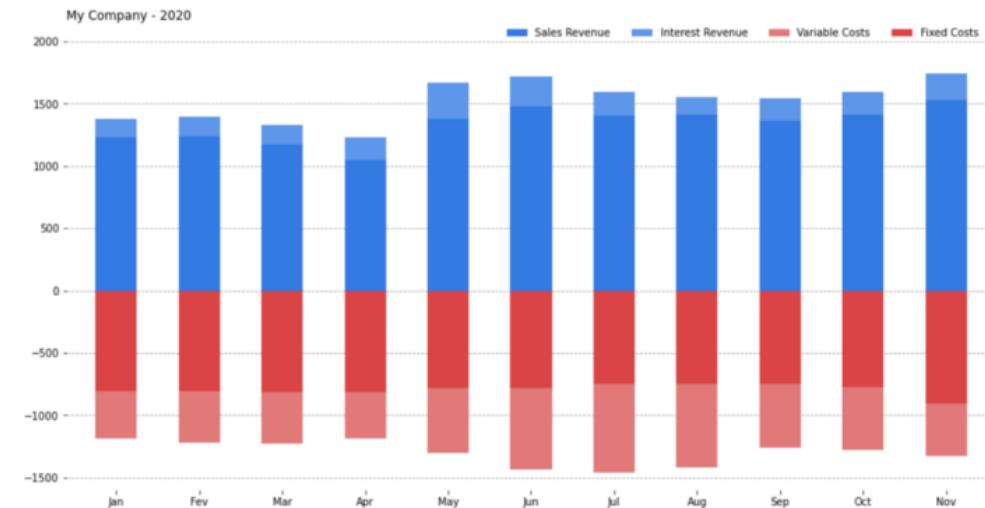
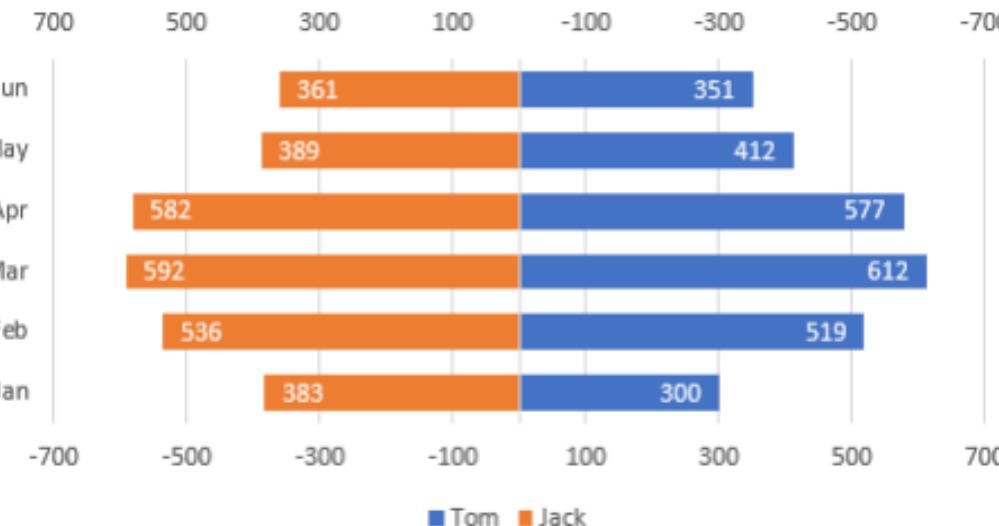
Column Chart



Grouped Column Chart

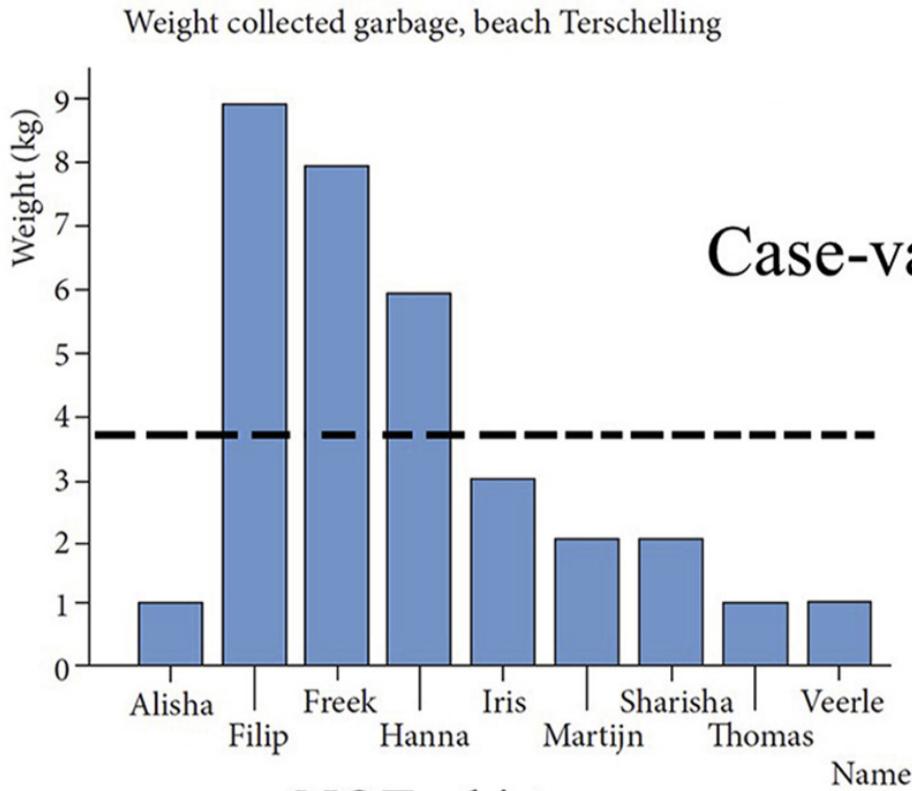


Stacked Column Chart



# THE HISTOGRAM

---



Thick black dotted  
lines: estimations  
of arithmetic mean

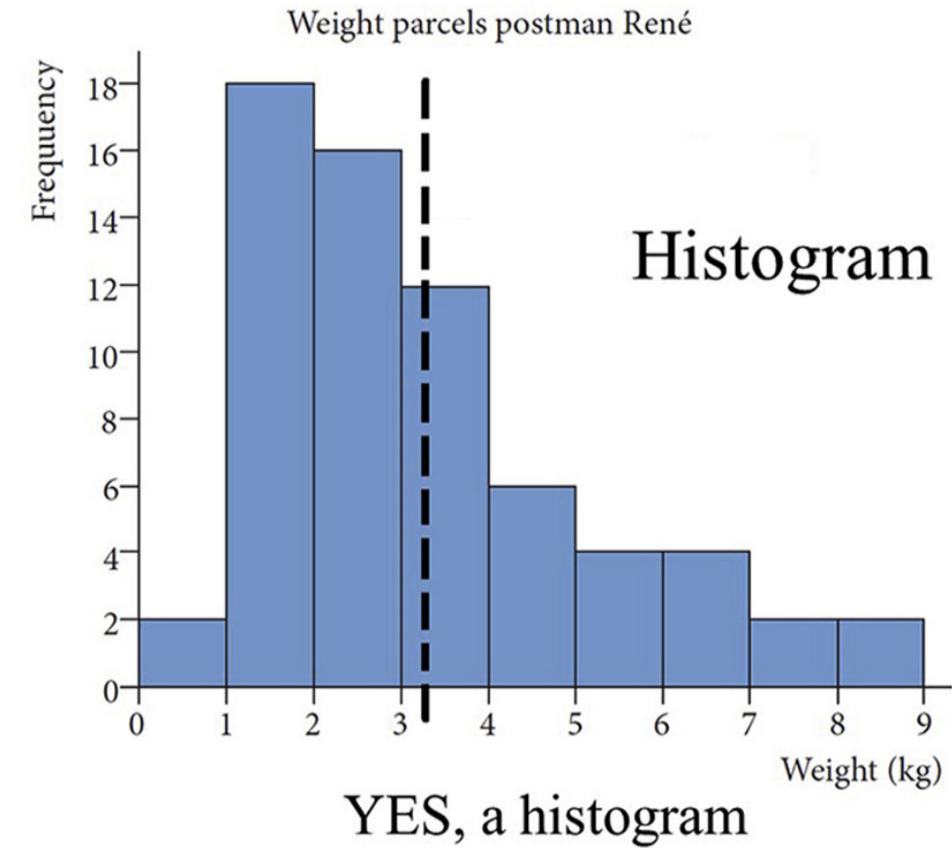
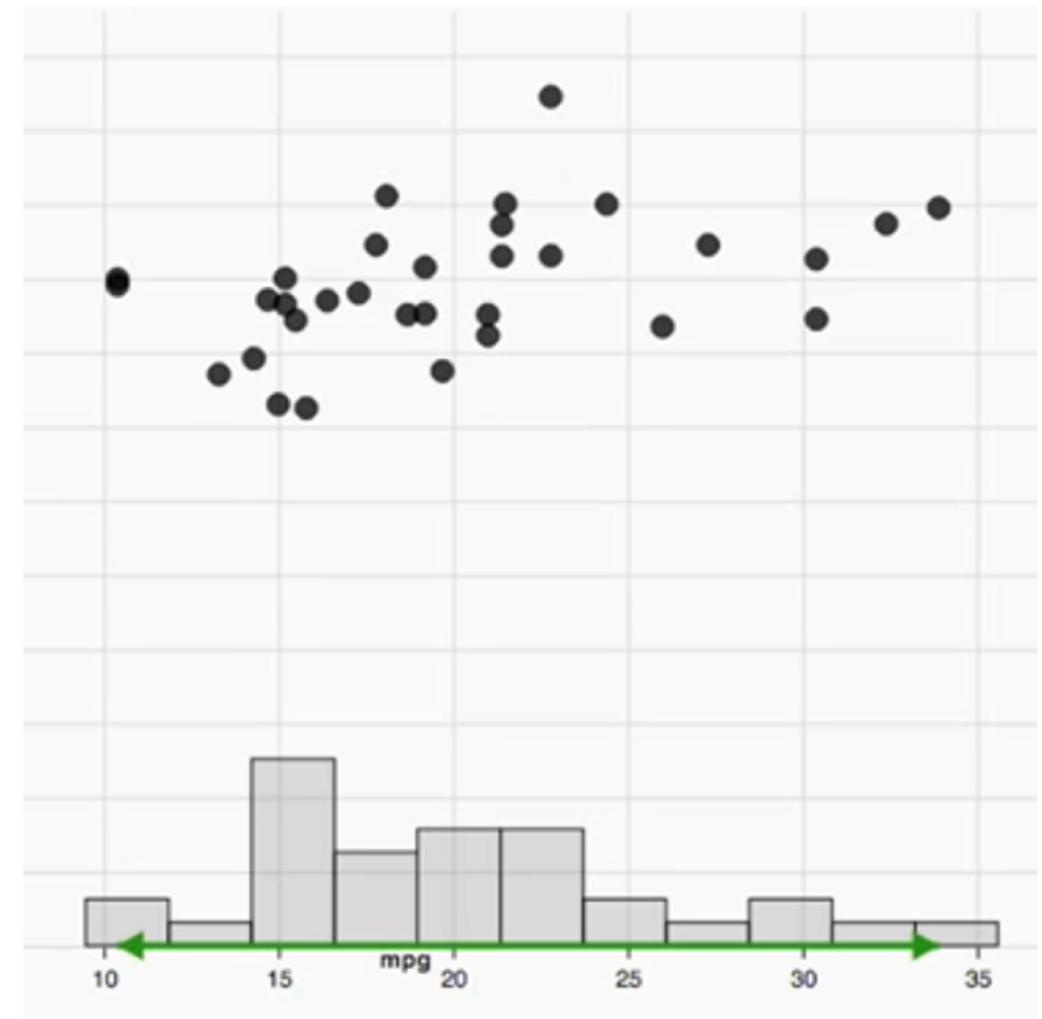
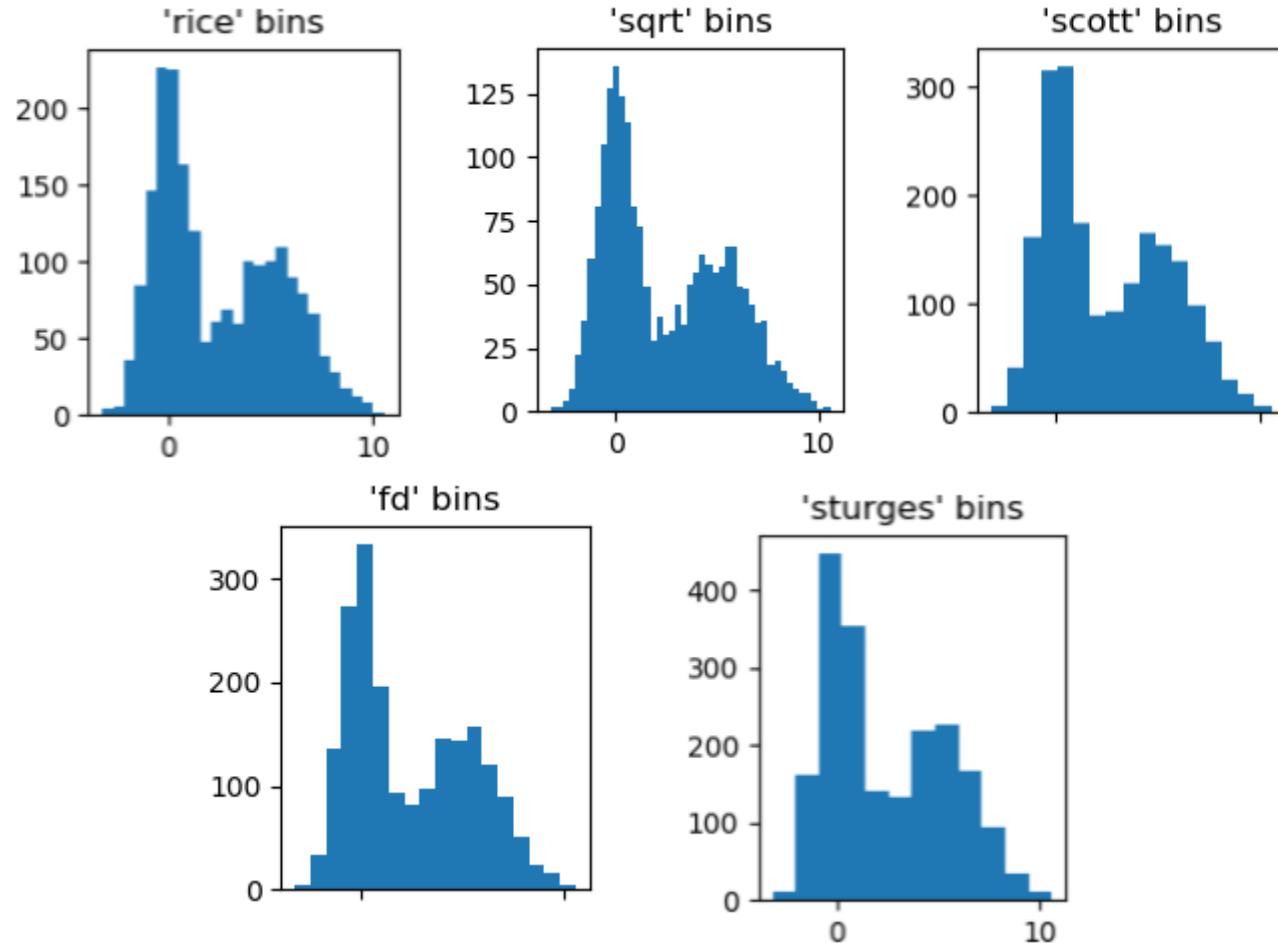


Image: 10.1016/j.edurev.2019.100291

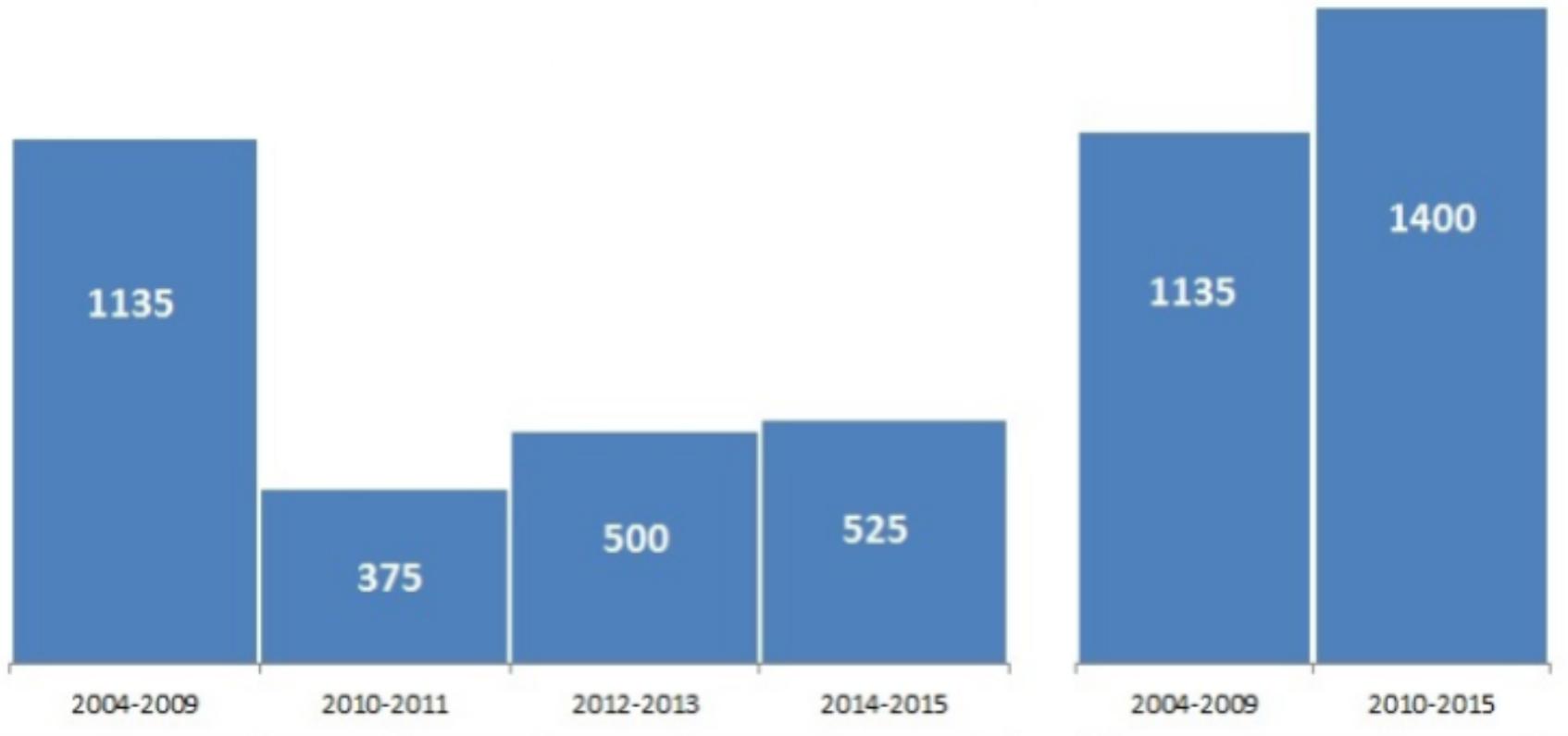
# THE HISTOGRAM - BIN SIZES & OFFSET

---



# THE HISTOGRAM - BIN SIZES

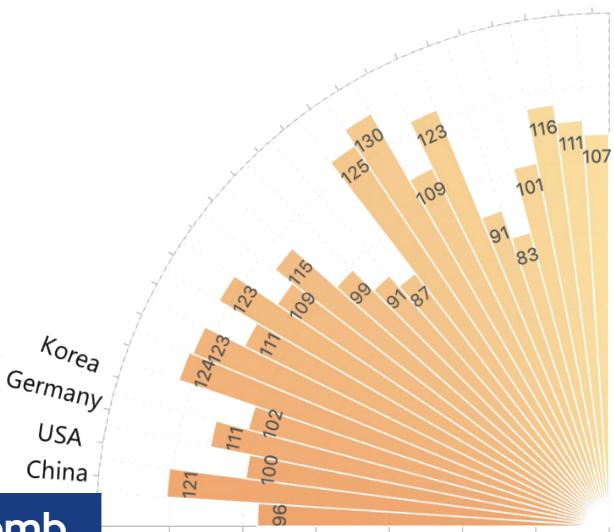
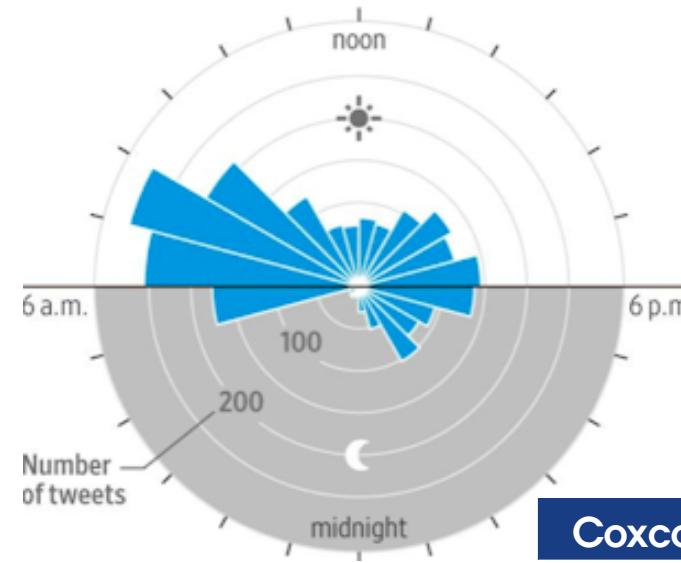
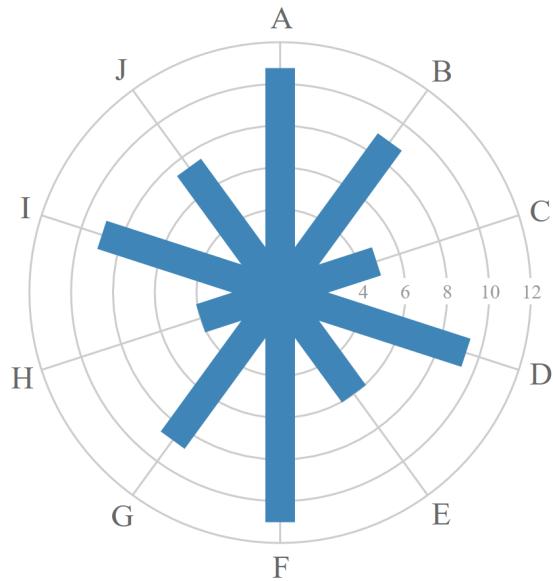
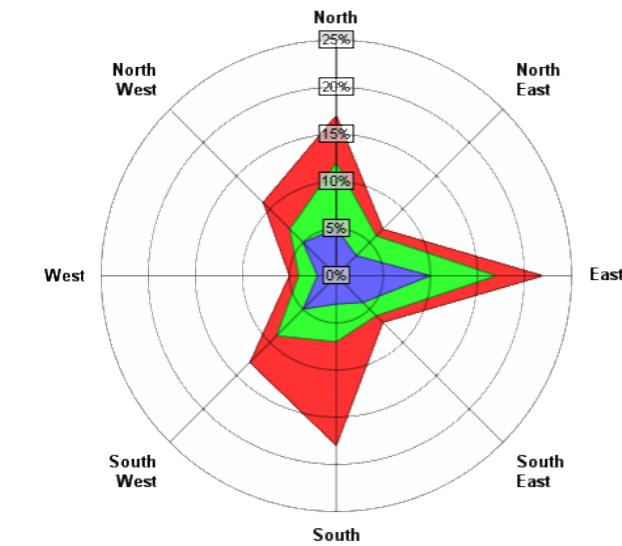
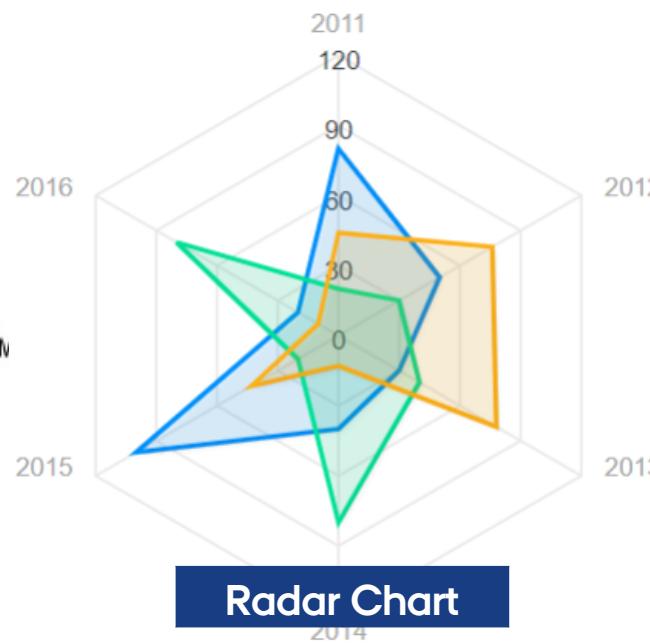
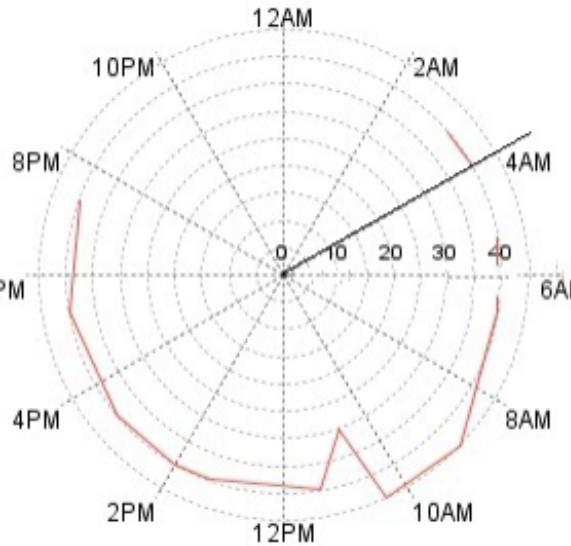
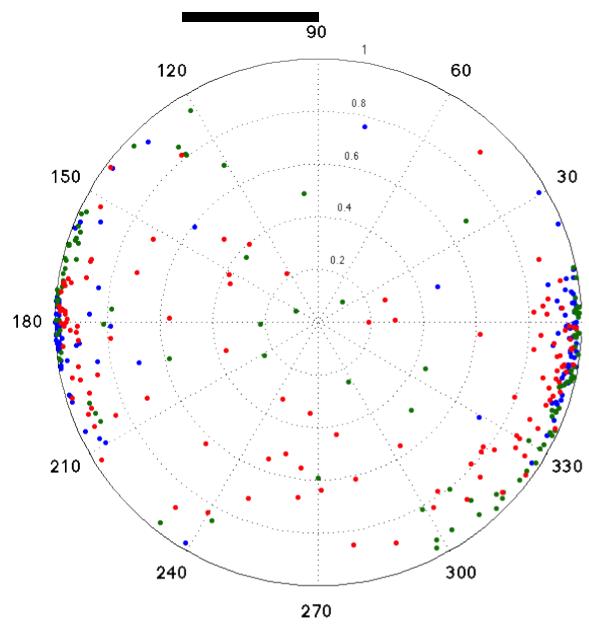
---



# THE RADIAL VARIANTS



# POLAR / RADIAL / CIRCULAR VARIANTS



Coxcomb

# LIST OF LITERATURE SOURCES

---

- Scatterplots – Scales: <https://doi.org/10.1126/science.216.4550.1138>
- Scatterplots – Aspect Ratio: <https://doi.org/10.1109/TVCG.2013.187>
- Scatterplots – Tick Marks: <https://doi.org/10.1109/TVCG.2010.130>
- Binned Scatterplots: <https://doi.org/10.2307/2289444>
- Splatterplots: <https://doi.org/10.1109/TVCG.2013.65>
- All charts scatterplots: <https://doi.org/10.1109/VIS47514.2020.00048>
- Line charts – Y-axis: <https://doi.org/10.1145/3313831.3376222>
- Line charts – Aspect Ratio: <https://doi.org/10.2307/2288843> & <https://doi.org/10.1109/TVCG.2012.196>
- Line charts – Rendering order: <https://doi.org/10.1111/cgf.14316>
- Line charts vs. Scatterplots: <https://doi.org/10.1109/TVCG.2017.2653106>
- Connected Scatterplot: <https://doi.org/10.1109/TVCG.2015.2502587>
- Stacked Area Charts: [https://doi.org/10.1007/978-3-030-86062-2\\_1](https://doi.org/10.1007/978-3-030-86062-2_1)
- Line Width Illusion: <https://doi.org/10.1109/TVCG.2013.140>
- Sine Illusion: <https://doi.org/10.1080/10618600.2014.951547>
- Stepped Perspective Bar Charts: <https://doi.org/10.20380/GI2021.28>
- Bar charts vs. Histograms: <https://doi.org/10.1016/j.edurev.2019.100291>

