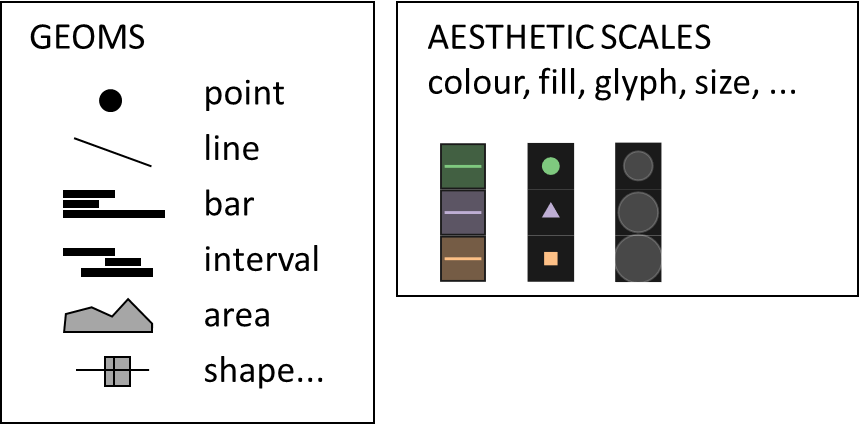
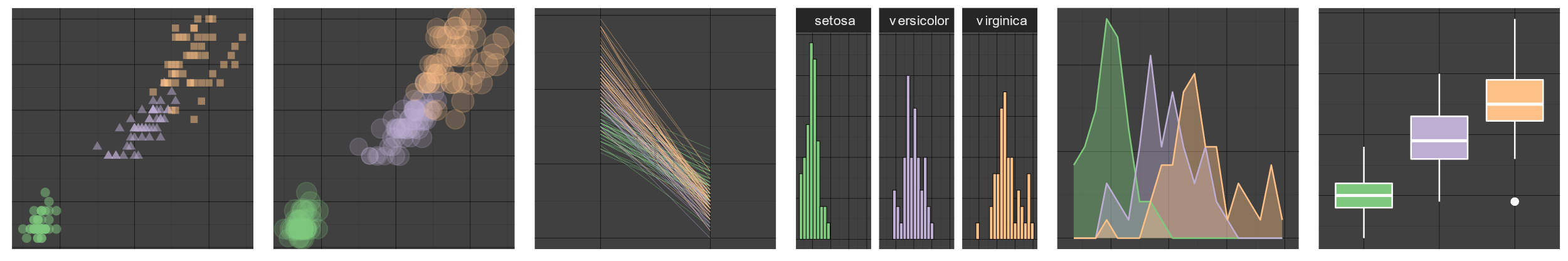
Visualization

# 1. Anatomy of a plot

* A plot consists of geoms.
* Often, there is one geom per row of the dataset, though they can also come from groups of rows.
* Data columns are mapped to aesthetic attributes of the geom.





# 2. Scale theory

There are four basic types of data scale:

* Nominal: no comparison is meaningful
* Ordinal: we can ask which is greater, but not measure how much
* Interval: we can subtract one value from another
* Ratio: we can divide one value by another

The four data scales work naturally with certain aesthetic scales.

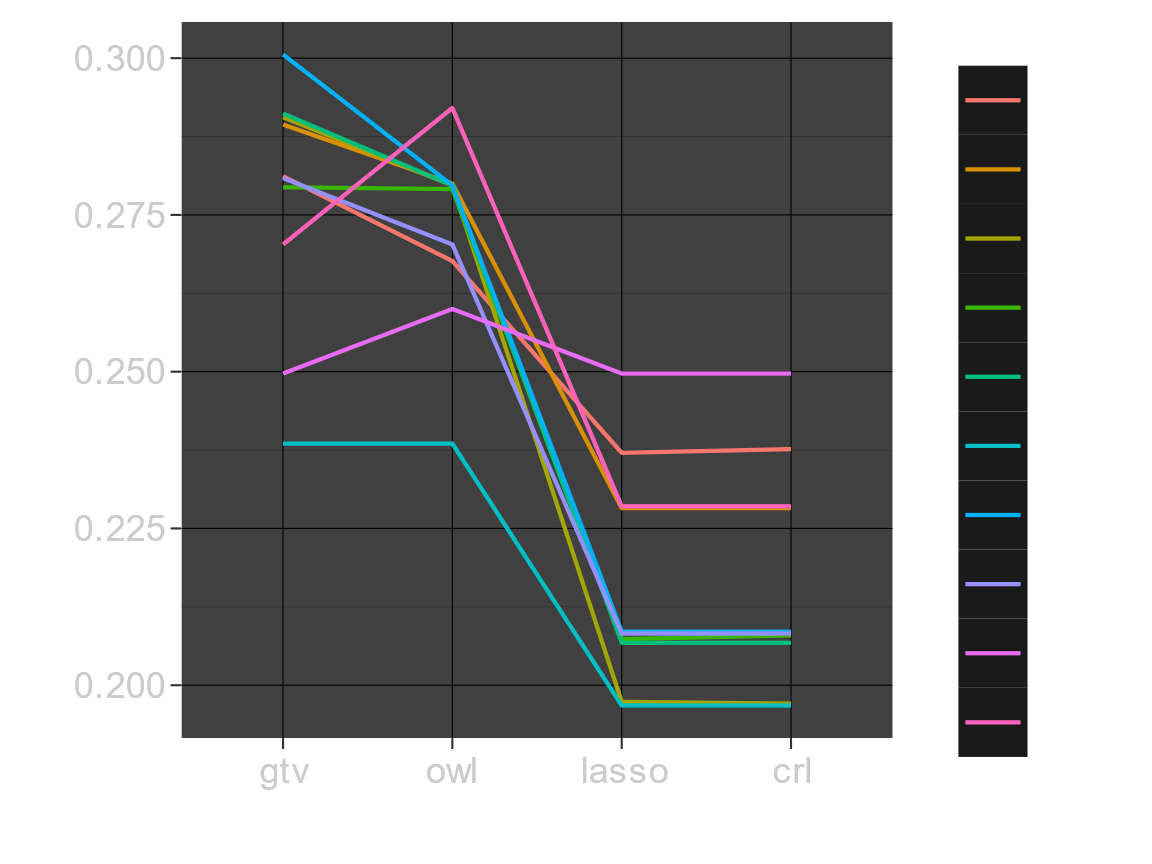
|  |  |
| --- | --- |
| nominal | shape, colour choice |
| ordinal | position index, colour sequence |
| interval | location, extent, colour gradient |
| ratio | size, area, divergent colour scale |

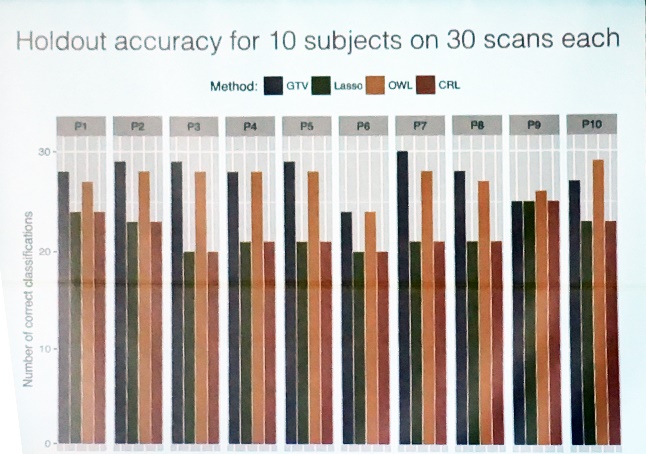
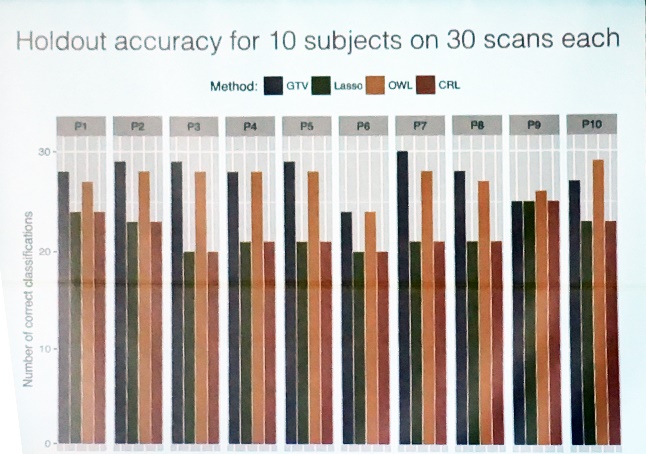
# 3. Scale perception

* We can see colour differences more easily than glyph.
* Less is more. Fechner/Weber: we notice %difference in a sensation, not absolute difference.
* Differences in or coordinate are the easiest to perceive
* Note that the subplot (in a multipanel plot) is also a type of location scale
* Area is dangerous. Stevens exponent: perceived area = (drawn area)0.8.
* Memory is also an aesthetic scale, used in user-hostile slideshows

# 4. Making comparisons

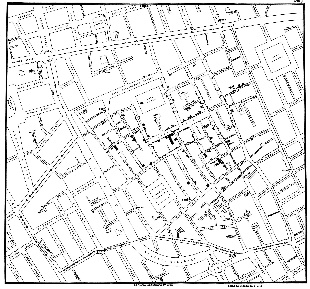
* Plots encourage the viewer to make comparisons   
  (how does feature depend on , , or ?)
* Have you put your primary comparators  
  on the best-perceived scales?

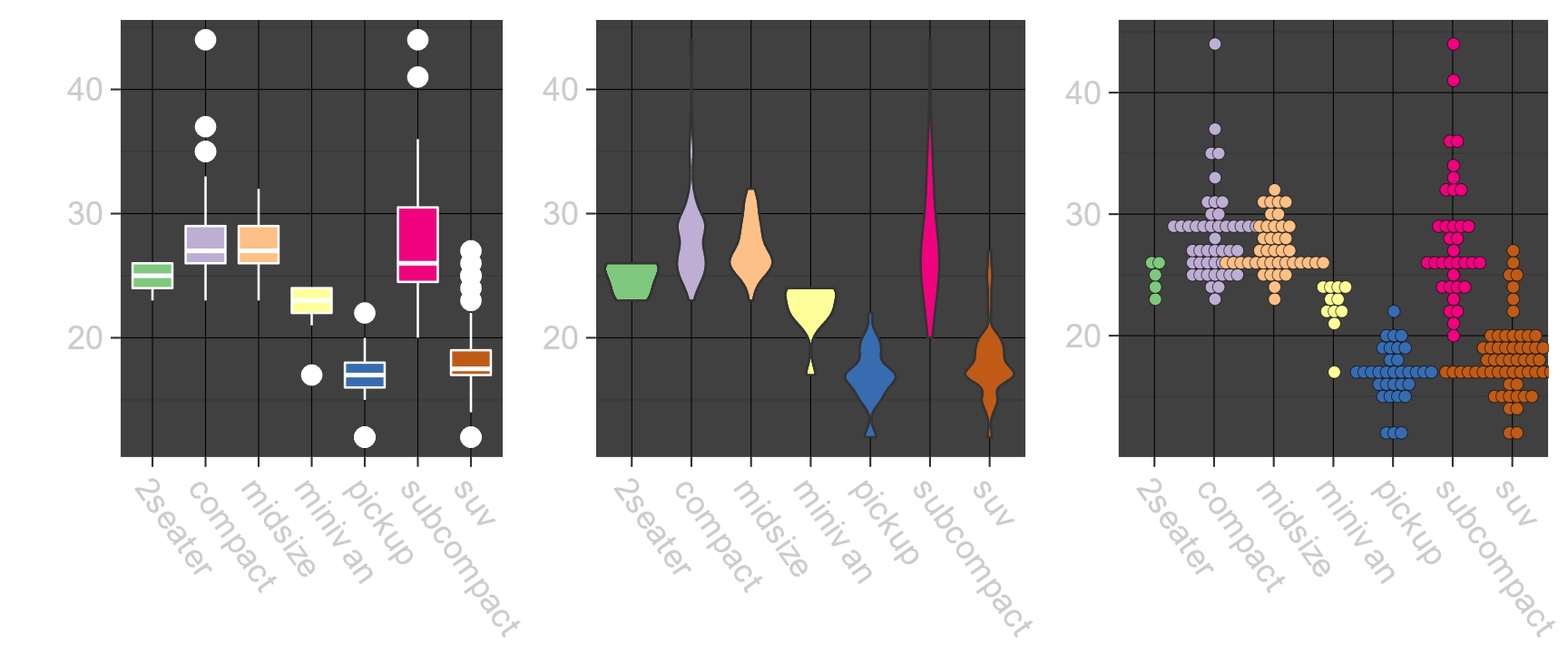




# 5. Atomic plots

In the best plots, every dot of ink is a datapoint. (This is why histograms are easy to read.)

John Snow, 1854 https://www.theguardian.com/news/datablog/  
2013/mar/15/john-snow-cholera-map.



# Next steps

* Style… Tufte, *The visual display of quantitative information.*
* How to tell a story… Berinato, *Good Charts*
* Software libraries… ggplot2 is unrivalled, and it’s not too hard to use it from Python.