Making effective graphs

Outline

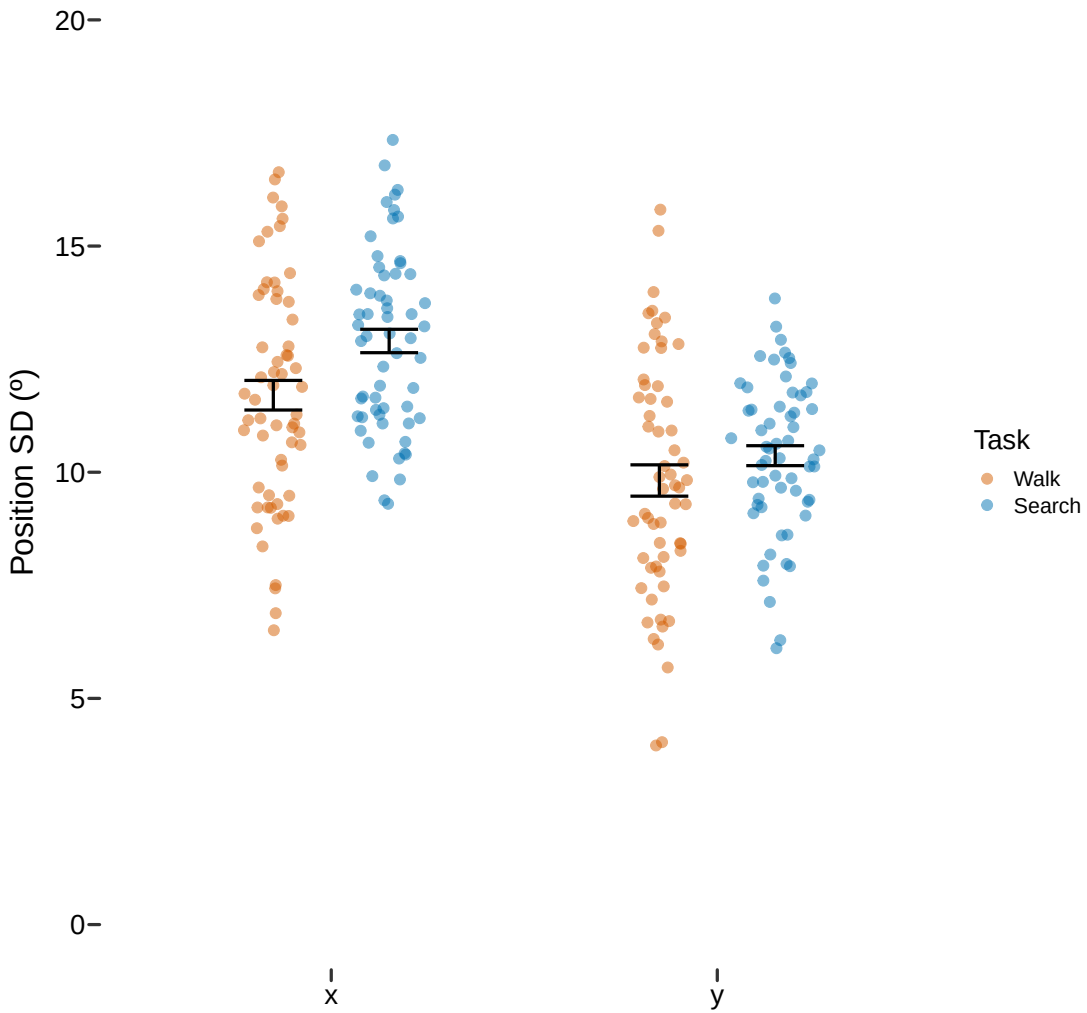
- What makes a good graph?
- What makes a good series of graphs?
- What graphical elements are appropriate for different types of data?

Making effective visualizations

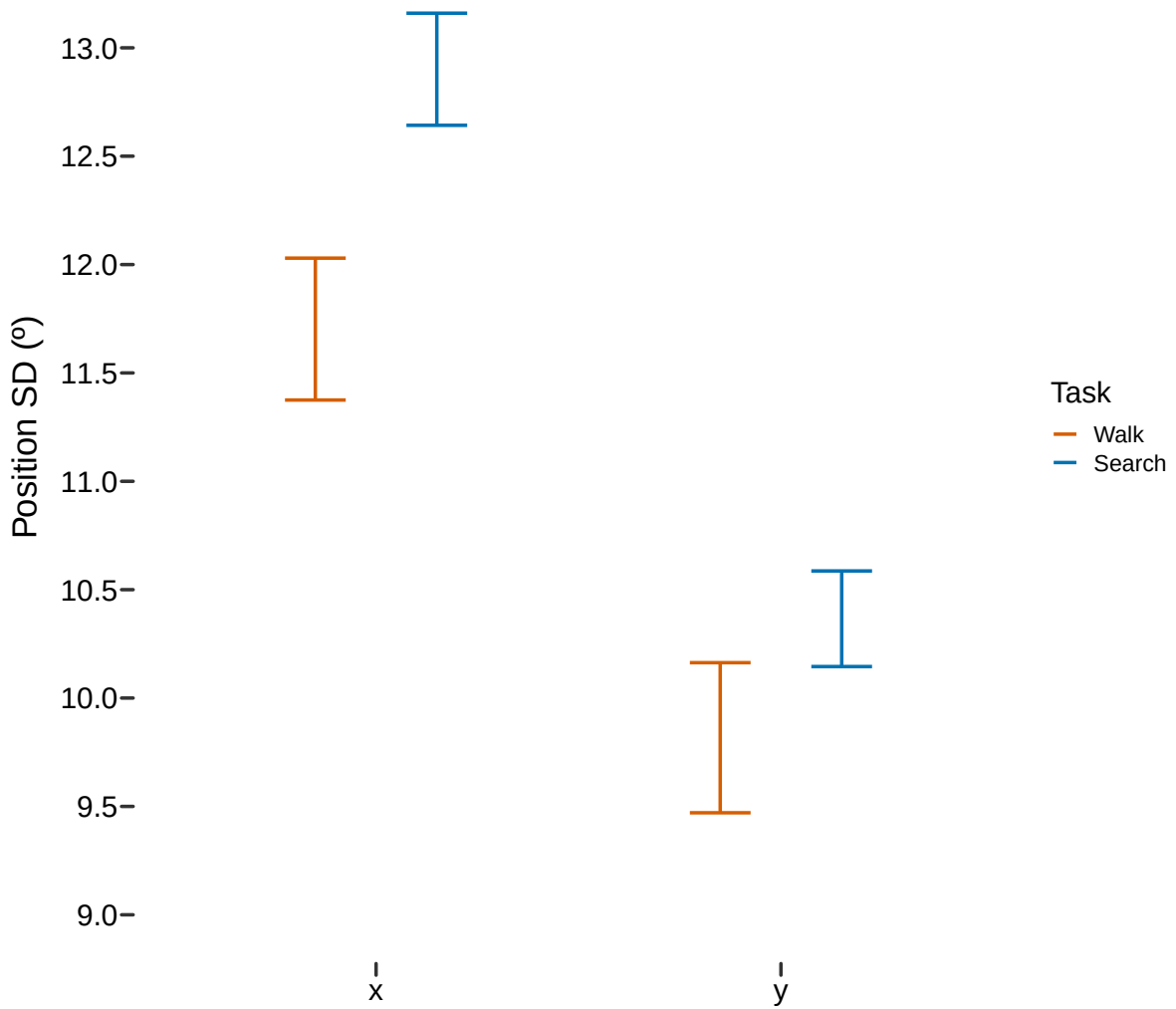
- What makes a good graph?

Graph from last time

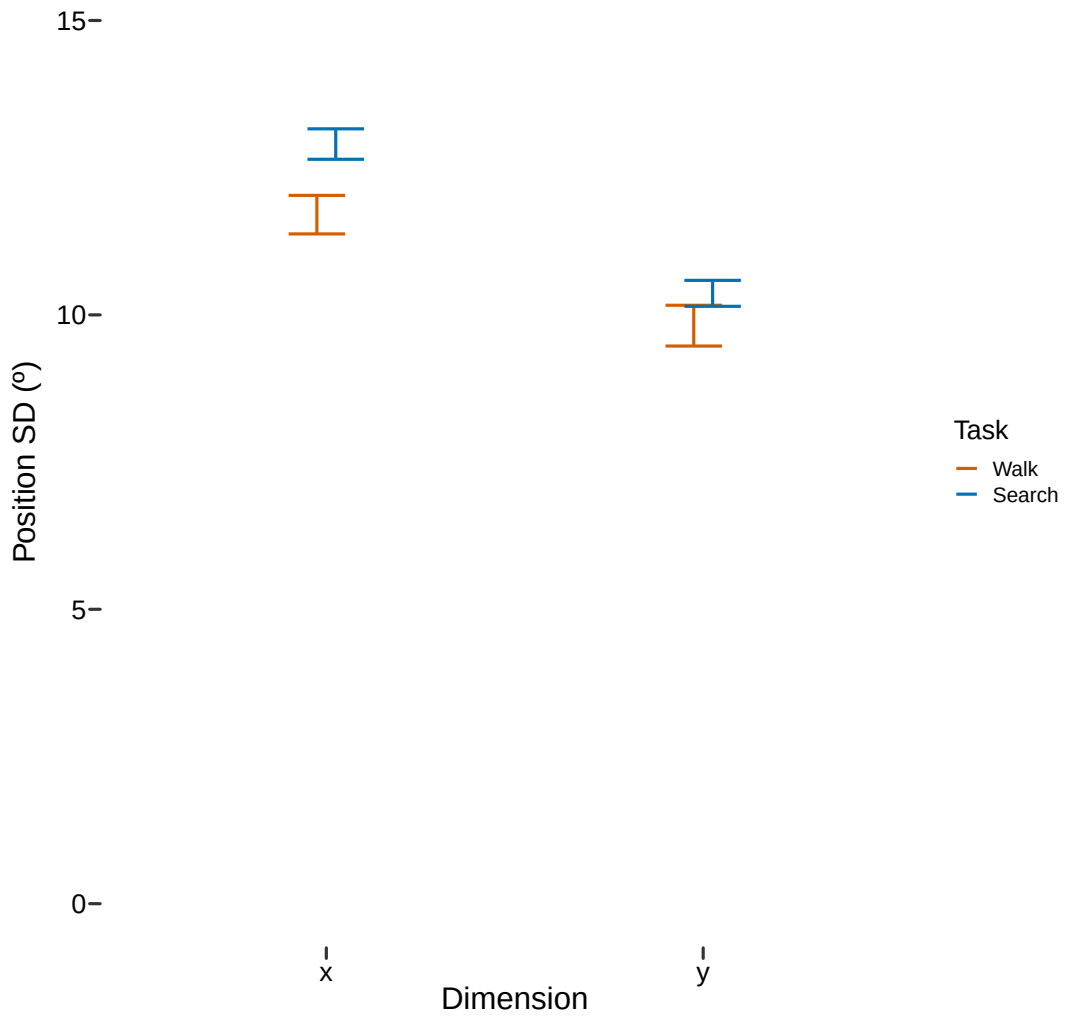
- What makes this a good graph?
- What makes this a bad graph?



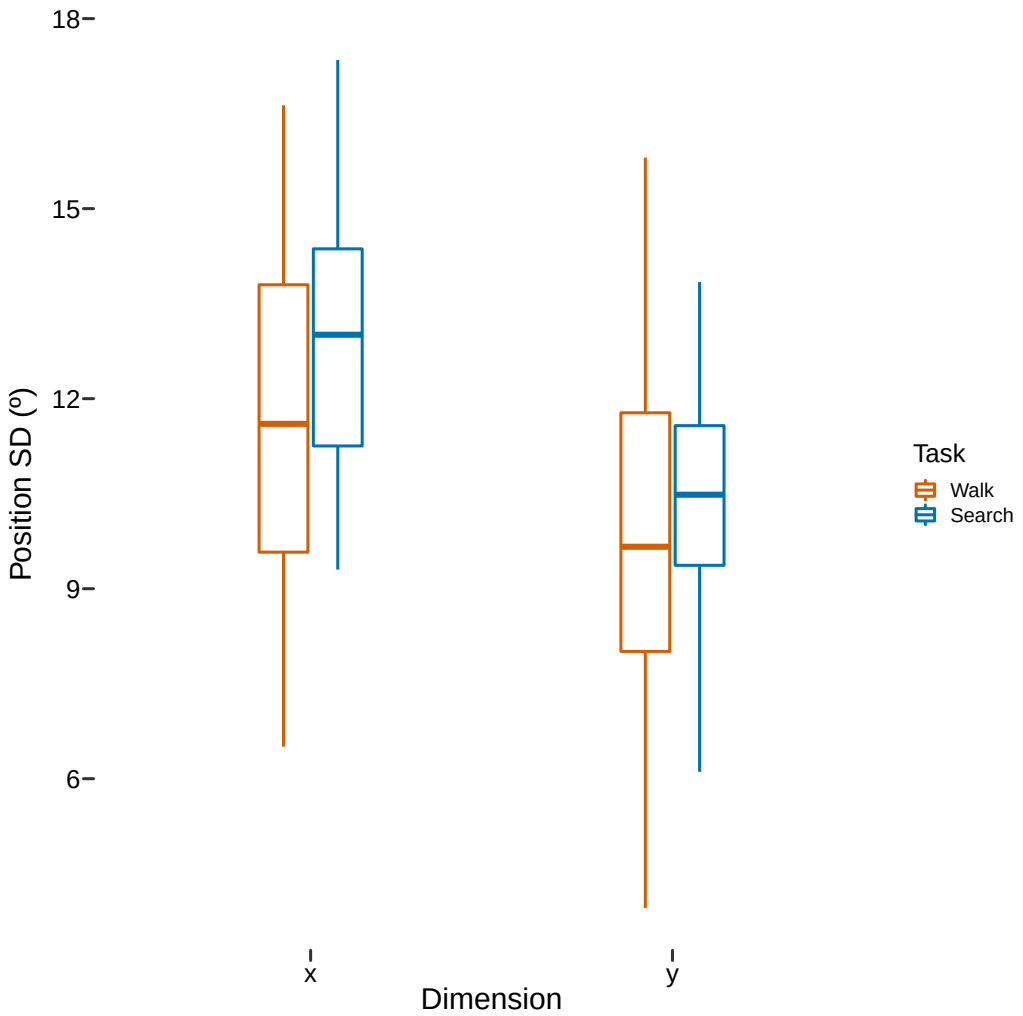
How about this one?



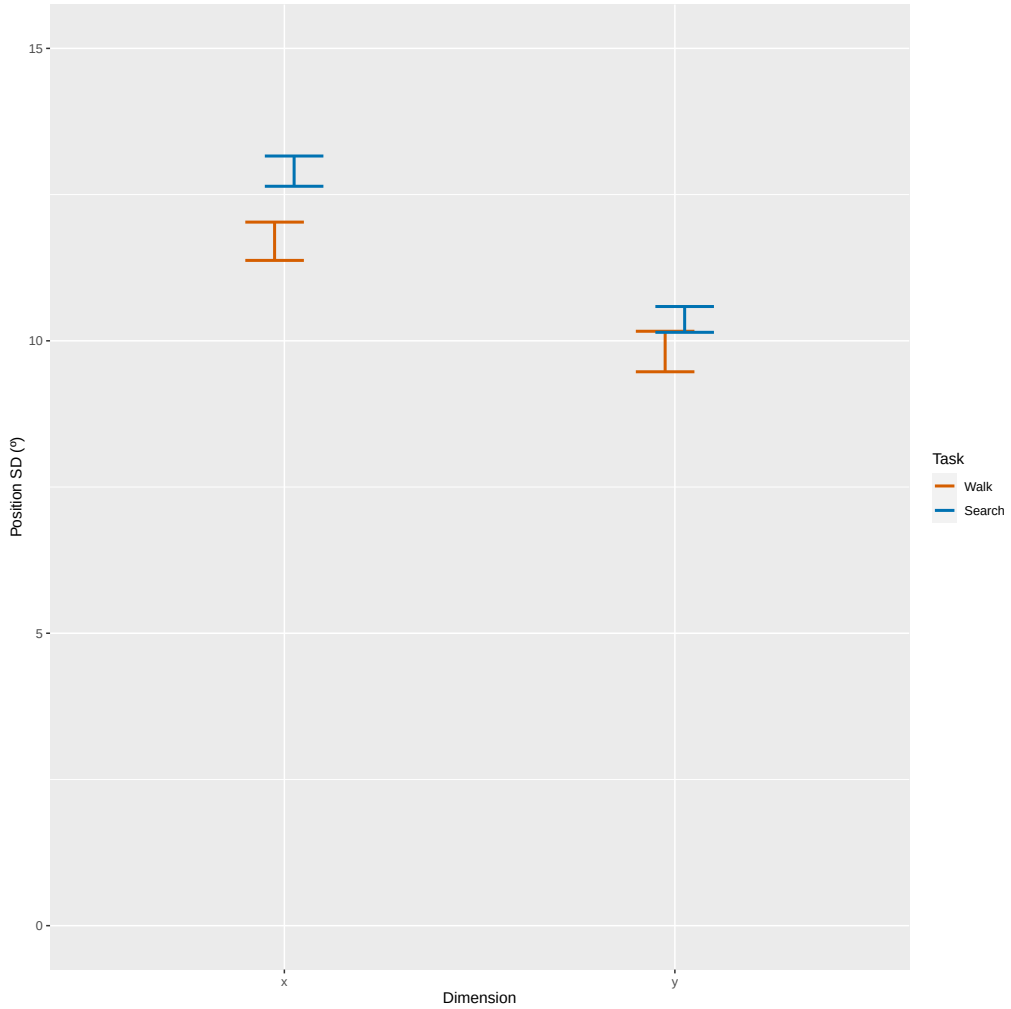
How about this one?



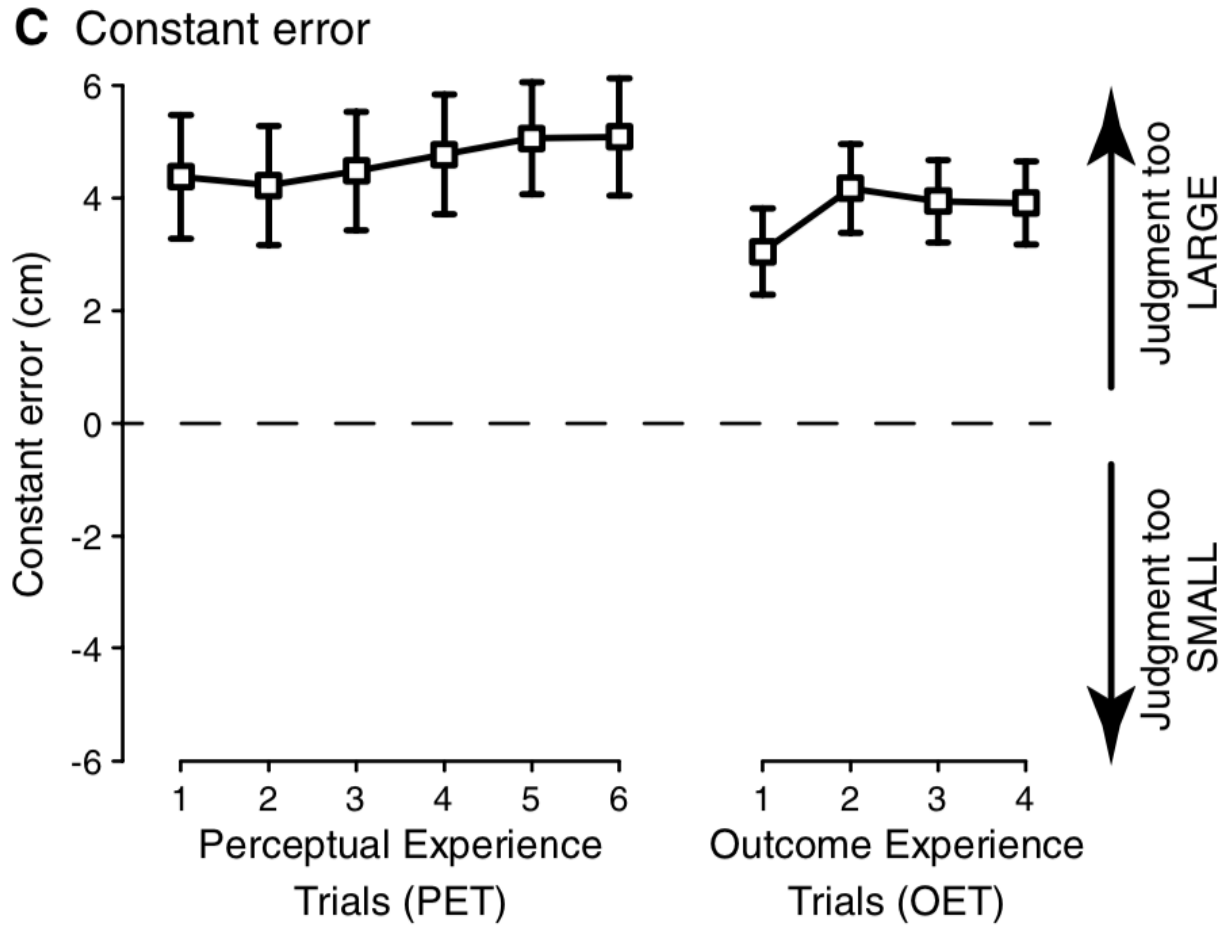
How about this one?



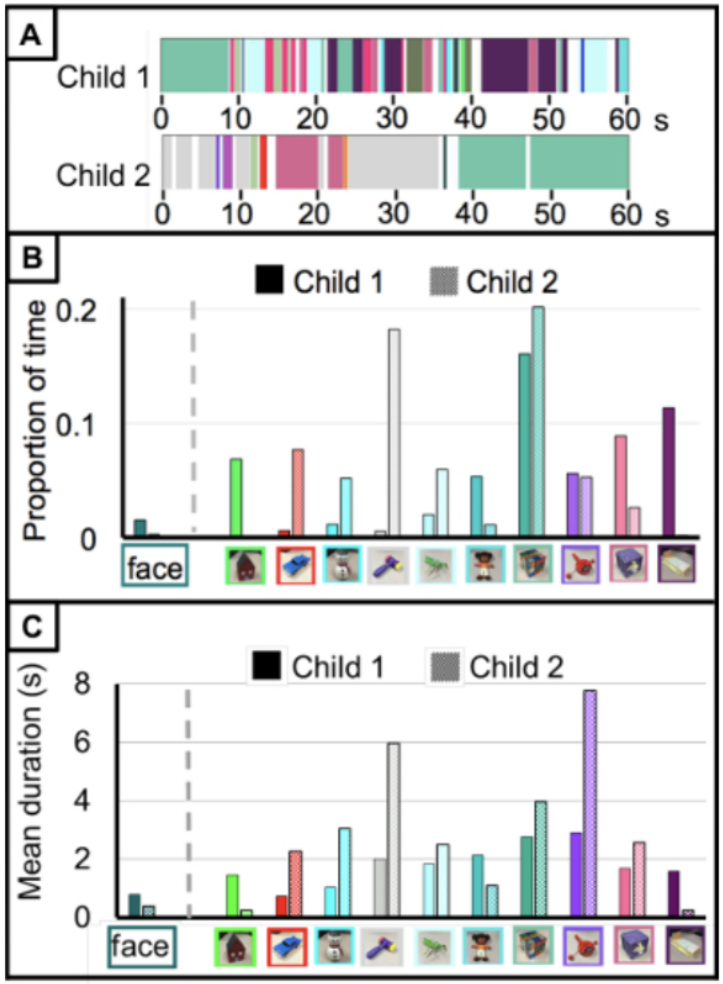
How about this one?



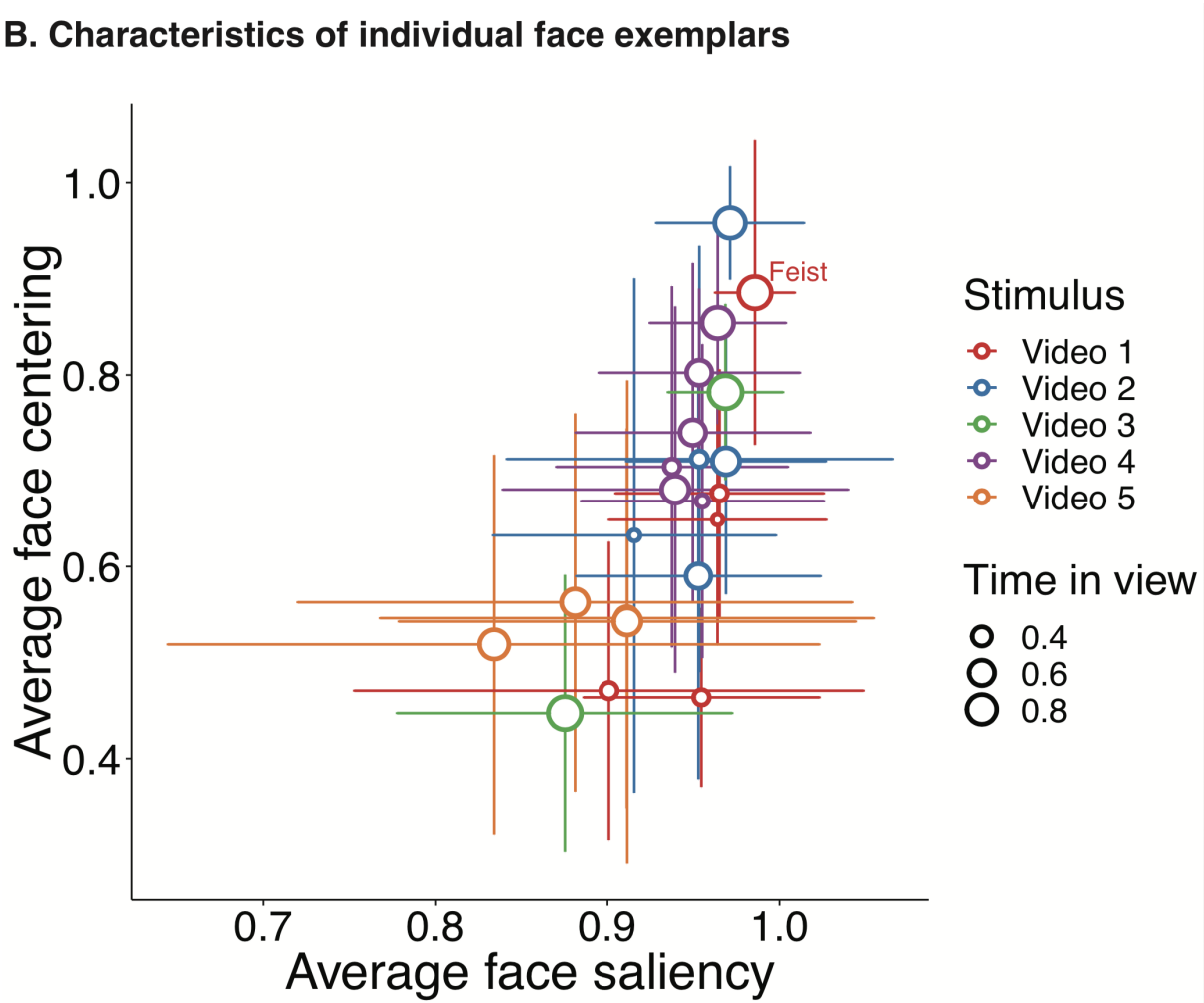
How about this one?



How about this one?



How about this one?



Making effective visualizations

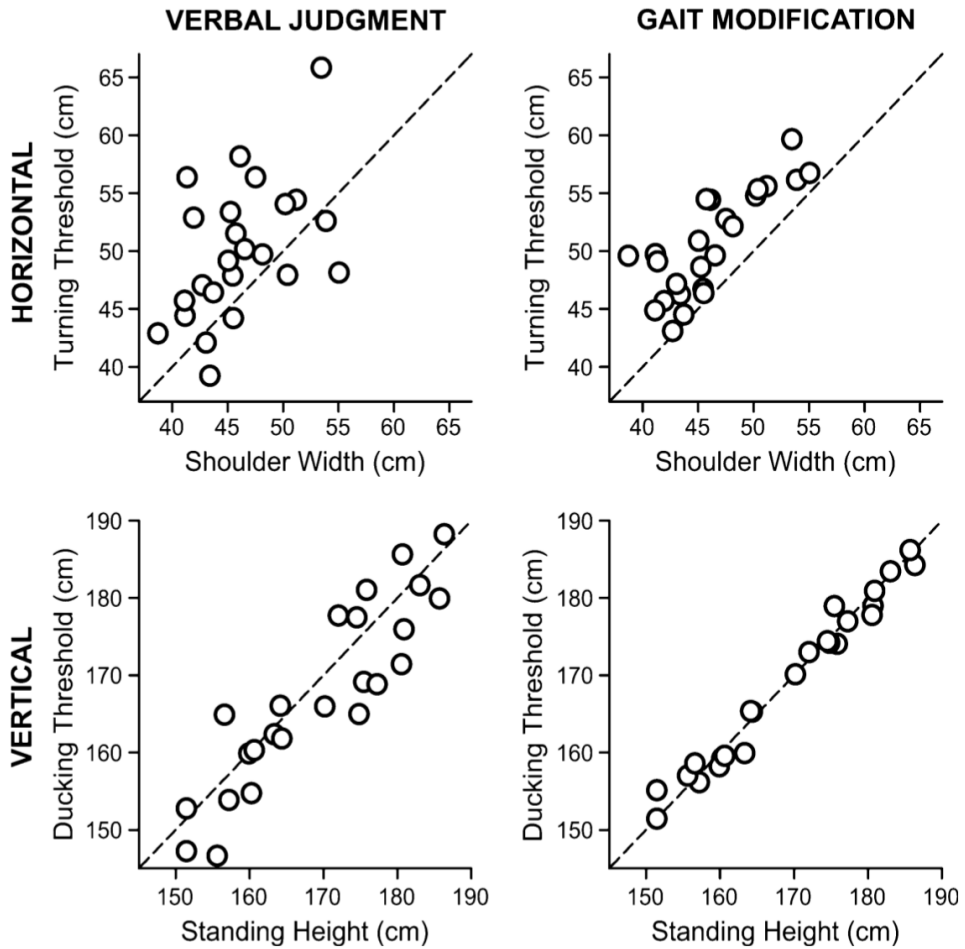
What makes a good graph?

- Highlights the most important information
- Reduces visual clutter (remove unnecessary backgrounds, lines, etc)
- Clear, legible elements (text is big enough, colors/shapes are distinguishable)
- Graphical elements map onto conceptual elements (graph mean +/- SE if showing inferential statistics, don't show min/max if it's not relevant)
- Graph does not mislead the audience (sensible axes, doesn't hide individual data/variability)
- Scaffolds for interpreting figure (annotations, reference lines, scales)

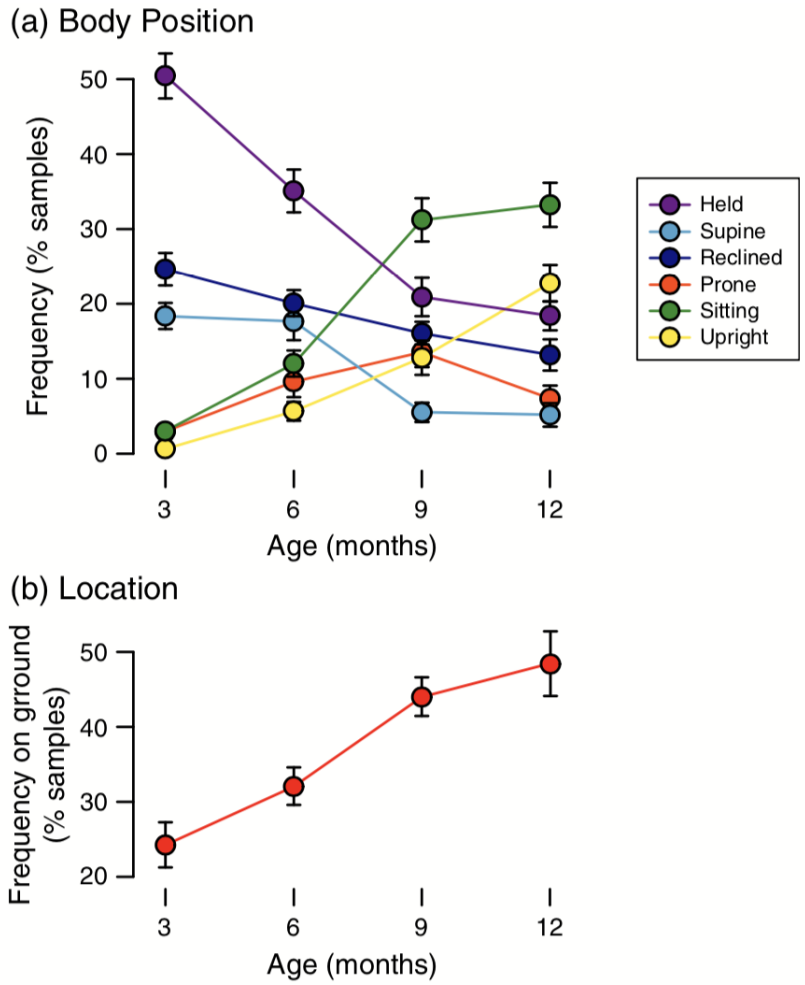
Making effective sets of visualizations

- What makes a good series of graphs in a paper?

How about this one?



How about this one?



How about this one?

Body mass index, a measure of body proportion that takes both height and weight into account, correlated with threshold opening width in the entrapment condition, $r(13) = .66$, $p = .008$, but not in the falling condition, $r(12) = -.04$, $p = .99$. Presumably, thresholds in the falling condition depended more on walking skill than on body dimensions. Indeed, three infants managed to navigate openings less than 10 cm in width by holding onto the moving wall with their hands to keep balance. Infants with earlier walking onsets tended to navigate smaller openings; however, the correla-

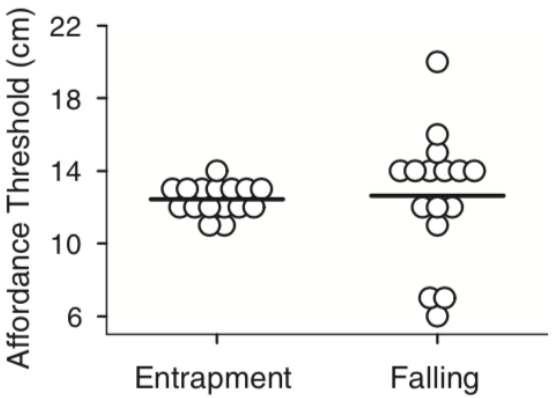


Figure 3. Infants' success thresholds in the entrapment and falling conditions. Each circle shows data for one infant. Horizontal bars indicate means.

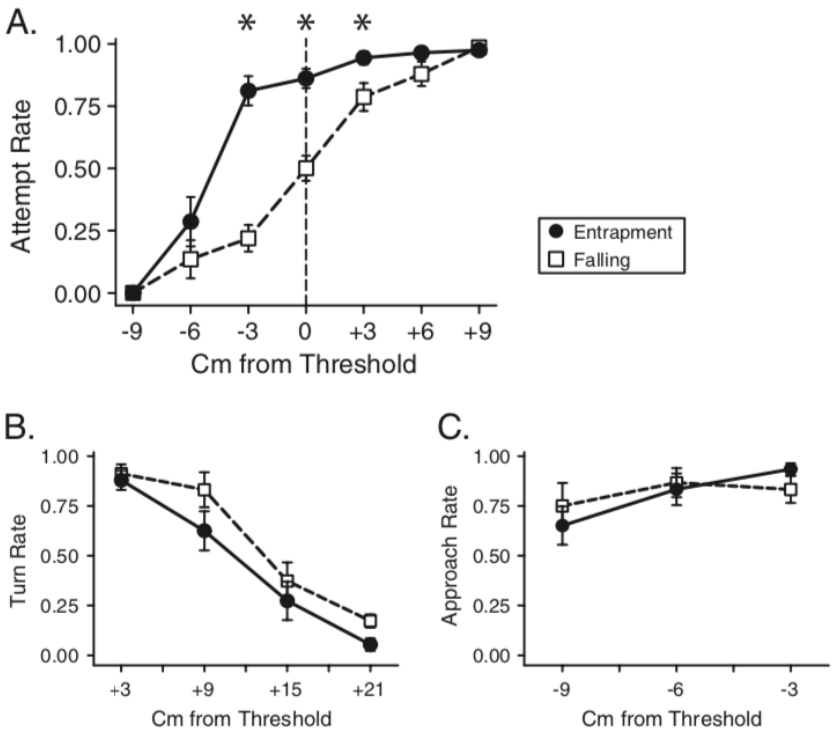


Figure 4. Infants' rates of (A) attempts, (B) prospective turning, and (C) approach behaviors in entrapment (filled circles) and falling (open squares) conditions. Vertical dashed line in (A) represents each infant's success threshold. Negative numbers on the *x*-axis denote impossible openings; positive numbers indicate possible openings.

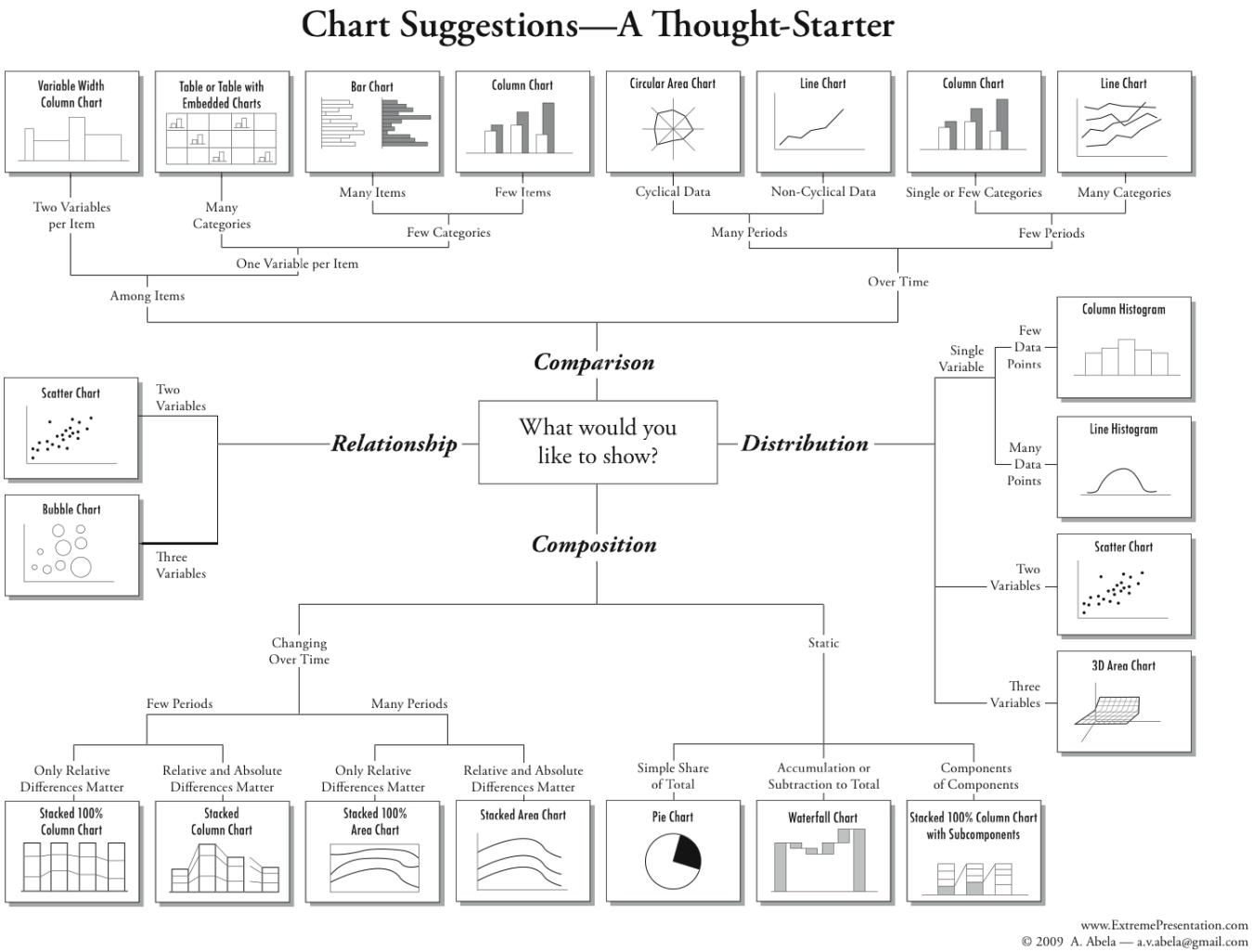
Making effective sets of visualizations

What makes a good series of graphs in a paper?

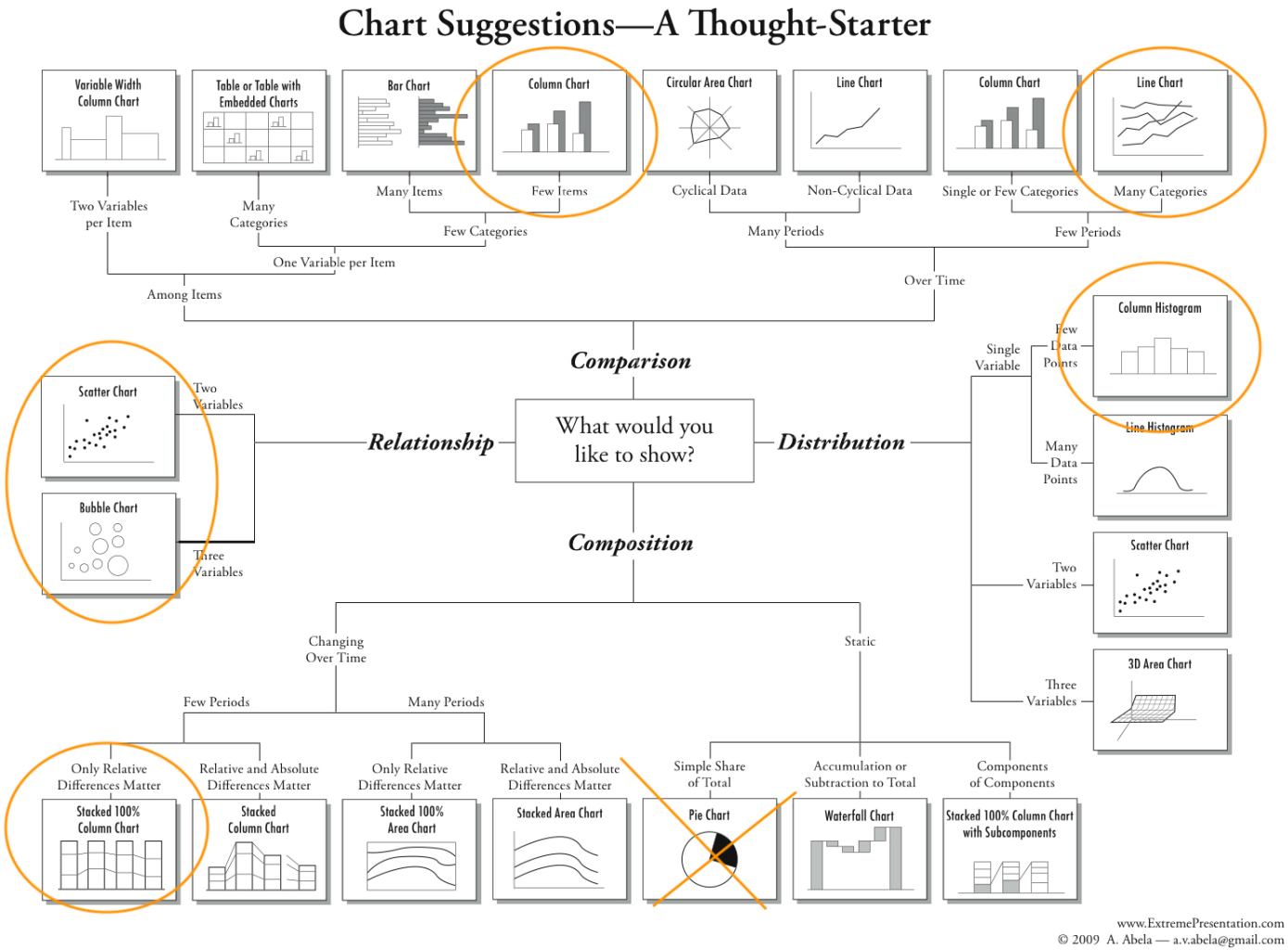
- Consistency scaling across graphs
- Same mappings for shape/color/line styles across graphs
- Consistent graphical styles (font sizes, line weights, capitalization, etc)

Following examples from: * http://r-statistics.co/Top50-Ggplot2-Visualizations-MasterList-R-Code.html

What graphical elements are appropriate for different types of data?

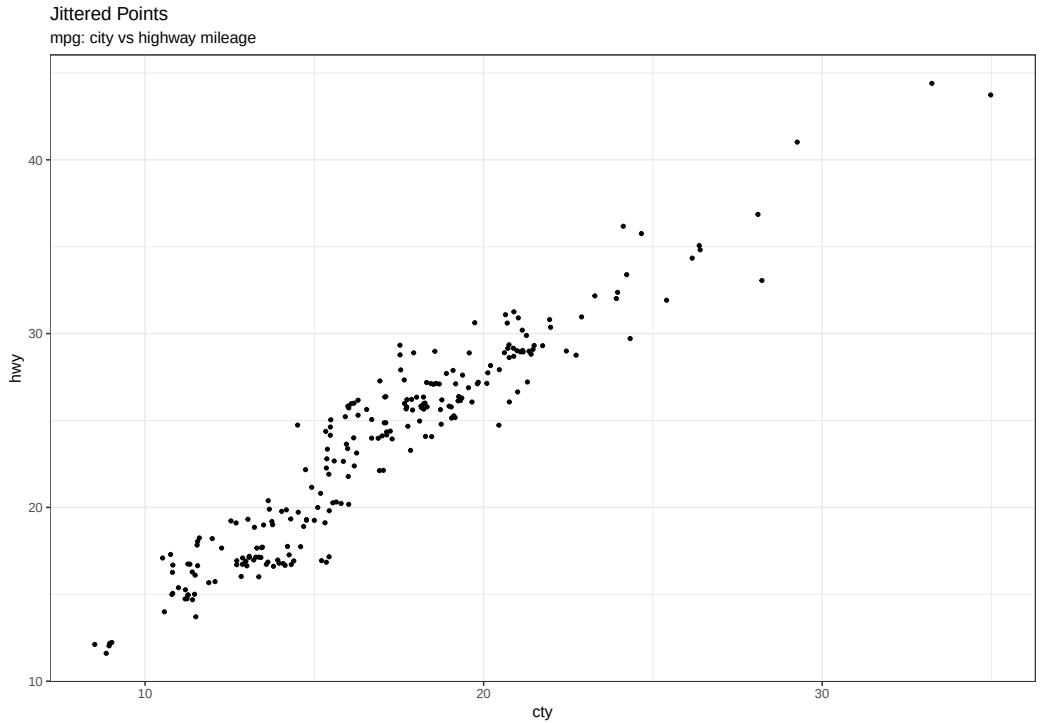


What graphical elements are appropriate for different types of data?



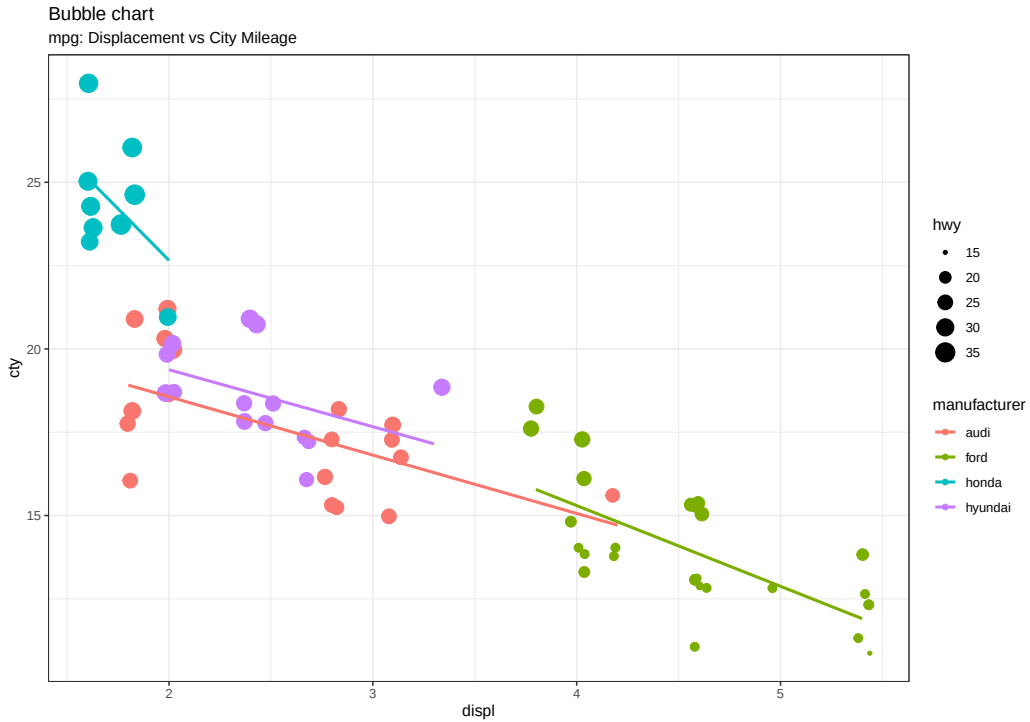
What graphical elements are appropriate for different types of data?

- Relationship
  - Scatterplot = continous variables, at least 4-5 levels each
  - May need to jitter if variables are discrete
  - Be careful of misleading axes scales (like this example)



What graphical elements are appropriate for different types of data?

- Relationship
- Use other aesthetics – size, color, lines – to represent subgroups



What graphical elements are appropriate for different types of data?

- Comparison
  - Bar/column chart, line and scatter, point range, boxplot
  - When to use a line vs bar?

What graphical elements are appropriate for different types of data?

- Comparison
  - Bar/column chart, line and scatter, point range, boxplot
  - Line typically means w/in subject data (data over time/treatments)
  - Point range is a better stand in for a bar chart: