



Affaires mondiales
Canada

Global Affairs
Canada

Canada

CANADIAN
FOREIGN
SERVICE
INSTITUTE

L'INSTITUT
CANADIEN
DU SERVICE
EXTÉRIEUR



Introduction to Data Analysis

DATA VISUALIZATION BASICS

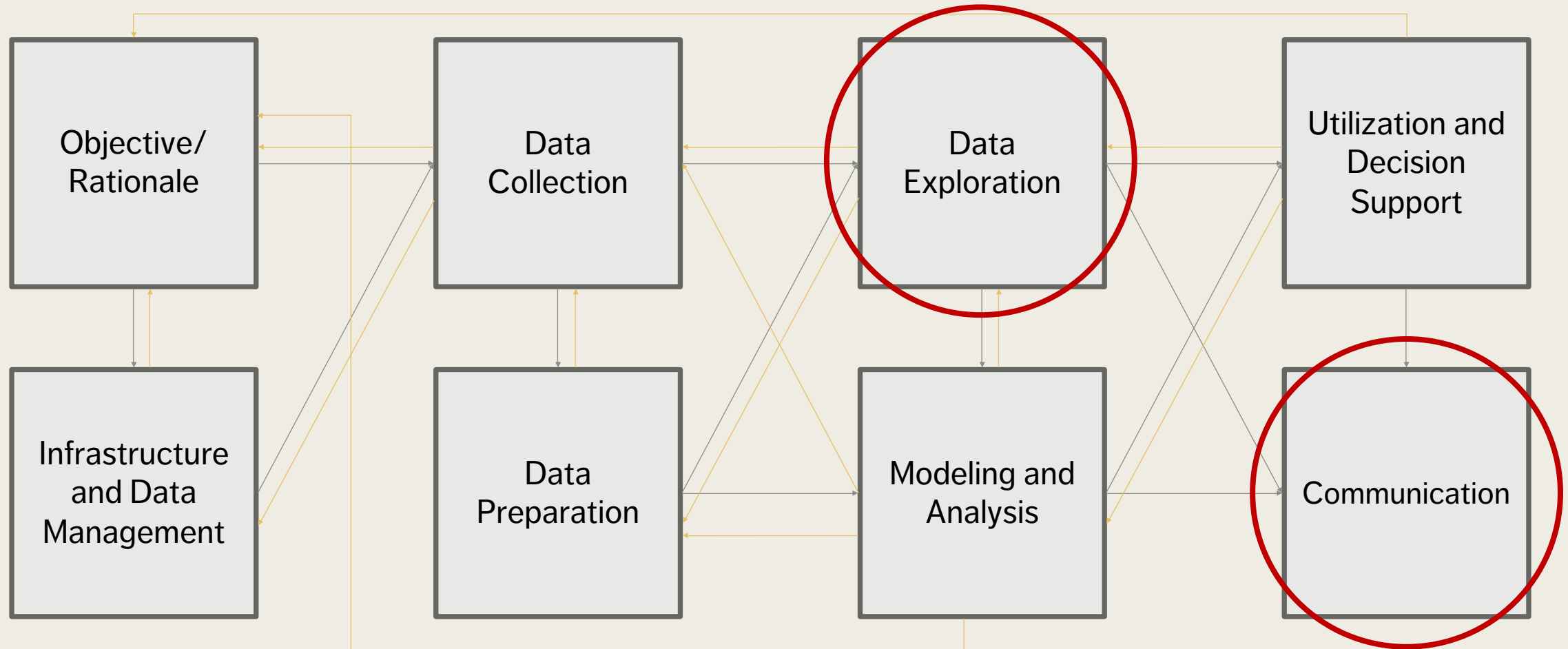
Patrick Boily

Data Action Lab | uOttawa | Idlewyld Analytics

pboily@uottawa.ca

[with files from Stephen Davies | DAVHILL]

THE (MESSY) ANALYSIS PROCESS



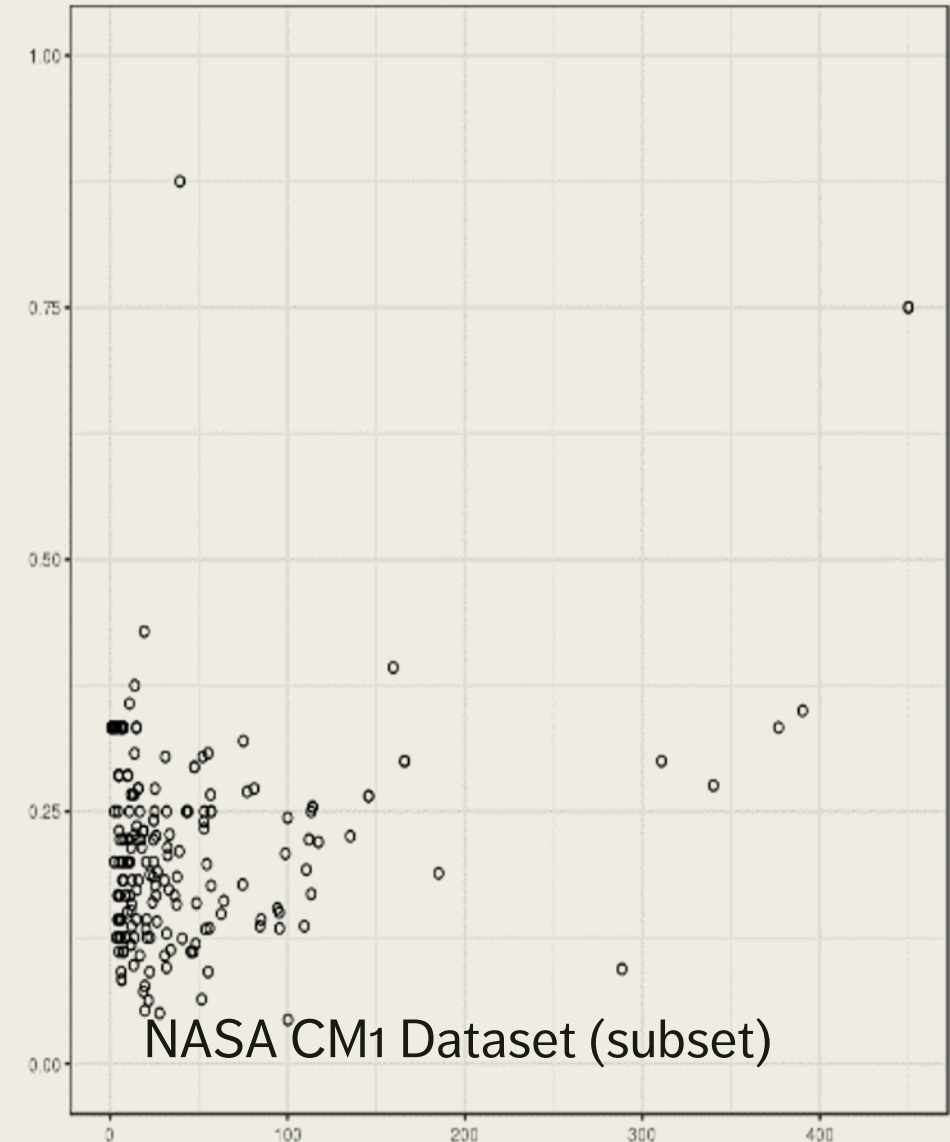


REPRESENTING OBSERVATIONS

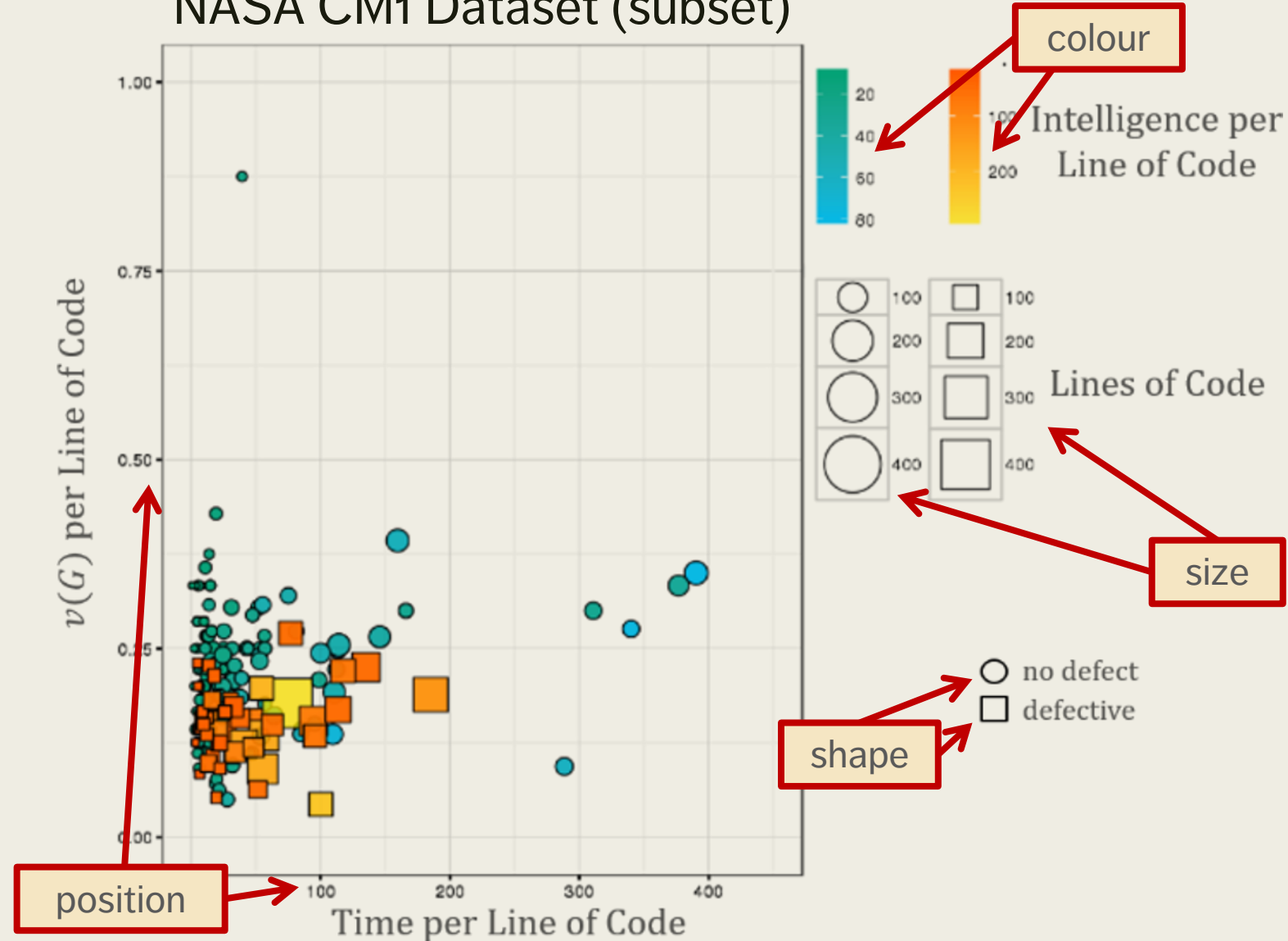
2 variables can be represented by position in the plane.

Additional factors can be depicted through:

- size
- color
- value
- texture
- line orientation
- shape
- (motion?)



NASA CM1 Dataset (subset)



WORKHORSE VISUALIZATIONS

Line Chart/Rug Chart/Number Line

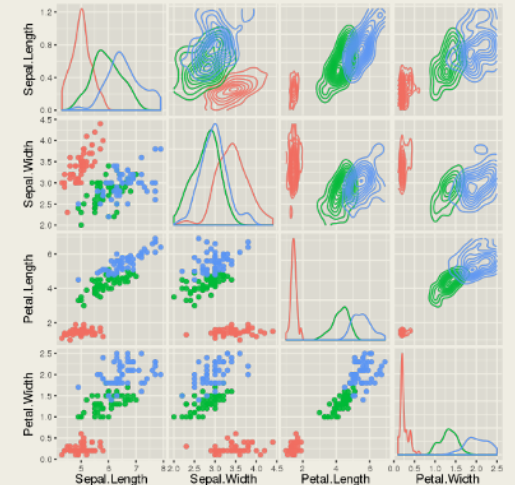
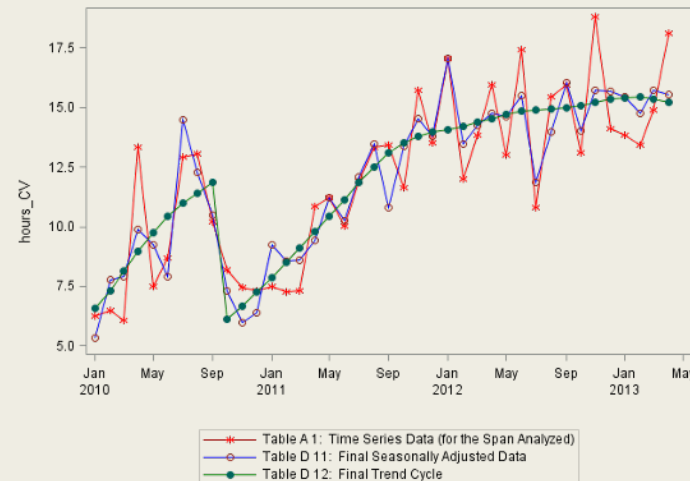
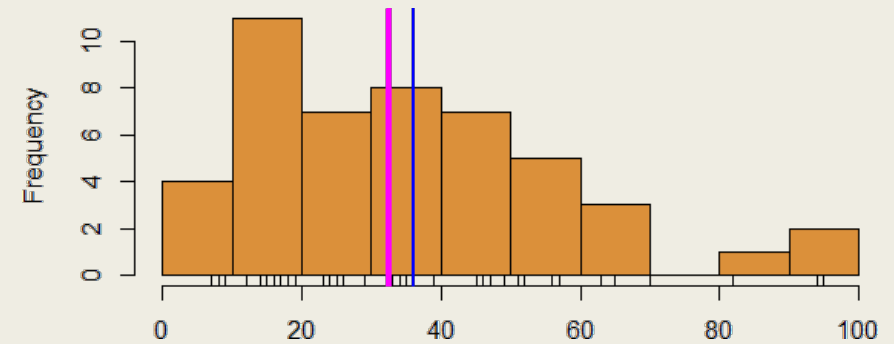
Histogram

Line Graph

Boxplot

Bar Chart

Scatterplot





FUNDAMENTAL PRINCIPLES OF ANALYTICAL DESIGN

Reasoning and communicating our thoughts are intertwined with our lives in a causal and dynamic multivariate Universe.

Symmetry to visual displays of evidence: consumers should be seeking exactly what producers should be providing, namely

- meaningful comparisons
- causal networks and underlying structure
- multivariate links
- integrated and relevant data
- honest documentation
- primary focus on content

Is the point getting across?
Is the message being conveyed?

A WORD ABOUT ACCESSIBILITY

Charts cannot usually be translated to Braille. Describing the features and emerging structures in a visualization is a possible solution... **if they can be spotted.**

Analysts must produce clear and meaningful visualizations, but they must also describe them and their features in a fashion that allows all to "see" the insights. This requires analysts to have "seen" all the insights, which is not always possible.

Conditions: colourblindness, low vision, motor impairment, cognitive disability, ADHD, etc.

Best Practices: high contrast text/elements, zoom/magnifications, keyboard navigation, assistive design, short summaries, undo/redo functionality, etc. [F. Elavsky]

A CLASSIFICATION OF CHART TYPES

Data comparison charts

Data reduction charts

Comparison

Bars



Dot plot



Bullet



ID Scatterplot



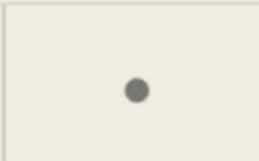
Heat map



Slope

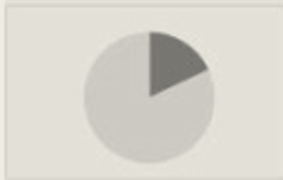


Alert



Composition

Pie



Pareto



Multidimensional Pie



Distribution

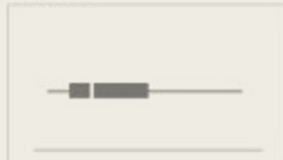
Histogram



ID Scatterplot



Boxplot



Evolution

Line



Horizon



Step



Connected Scatterplot



Relationship

Scatterplot



Connected Scatterplot



Bubble



Connected Scatterplot



Profiling

Grouped bars



Cycle plot



Scatterplot matrix



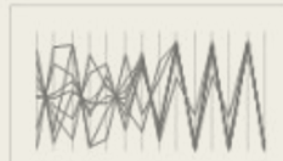
Reorderable matrix



Horizon



Parallel Plot



Trellis



CHART TYPES

Simple Text and Tables

Scatterplot

Line Chart

Bar Charts

Stacked Bar Charts

100% Bar Charts

Area Charts

Treemaps

Gauge Charts

Heatmaps and Choropleth Maps

Geographical Maps

Parallel Coordinates

Chernoff Faces

Word Clouds

Network Diagrams

Dendrograms and Trees

Sparklines

Interactive Charts

Small Multiples

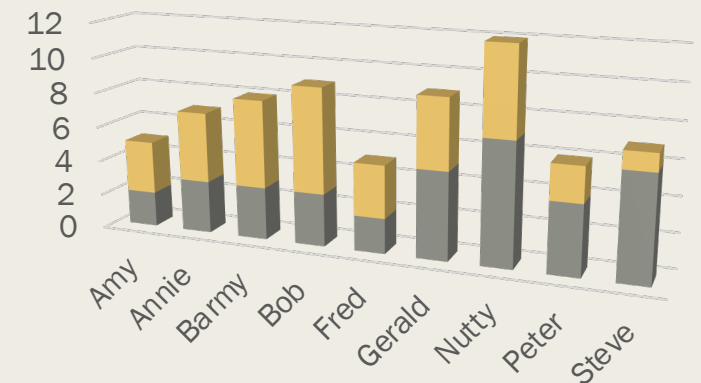
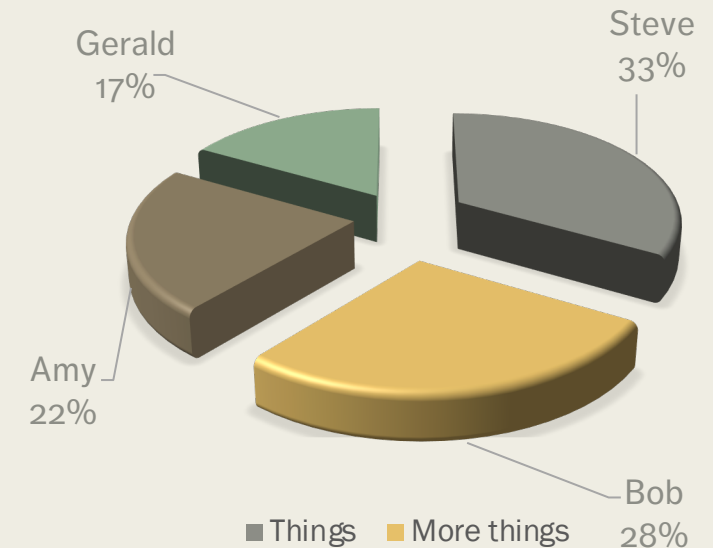
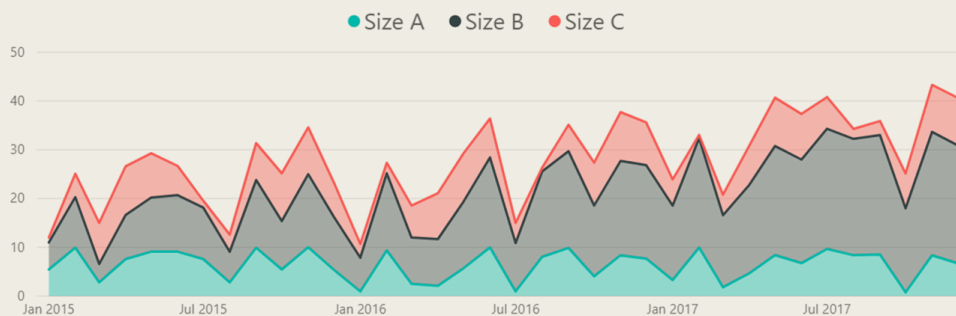
etc.

CHARTS TO AVOID

AVOID (?) anything with an arc (except gauge charts): pie, donut, etc: human brains have a hard time **comparing arcs** -- without labels, how different are Steve & Bob?

AVOID 3D charts: it is difficult to compare them visually (and they add **too much** clutter).

AVOID stacked area charts: way too confusing.



TAKE-AWAYS

Effective data visualizations **provide insights** and **facilitate understanding**.

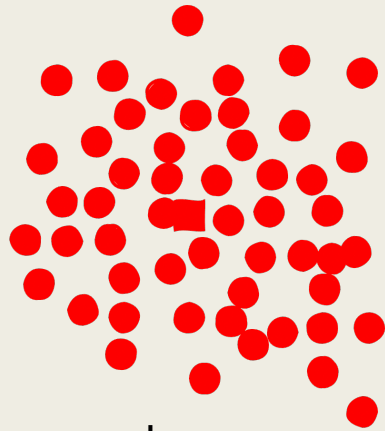
The basic principles can guide your visualization design and consumption.

Be **creative** but keep your data and your representations **honest**.

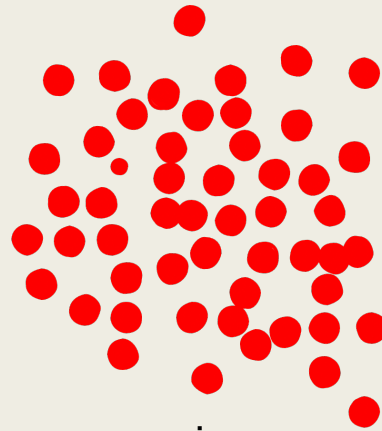
Be mindful of attempts to distort trends and conclusions with flashy visuals.

Data and code should be made available along with the displays.

PRE-ATTENTIVE FEATURES



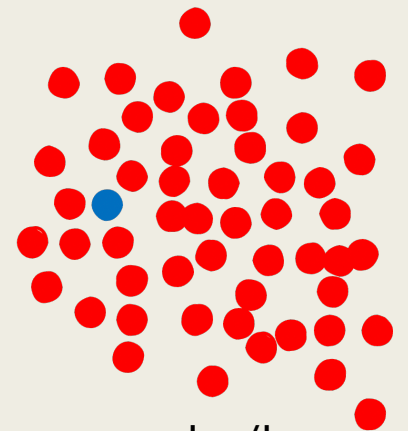
shape



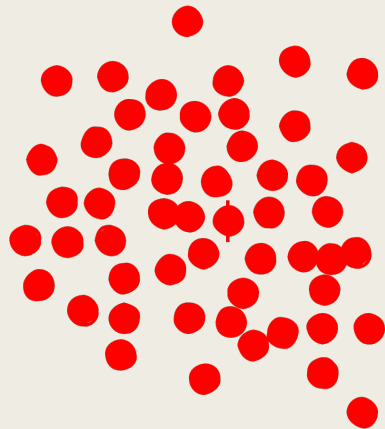
size



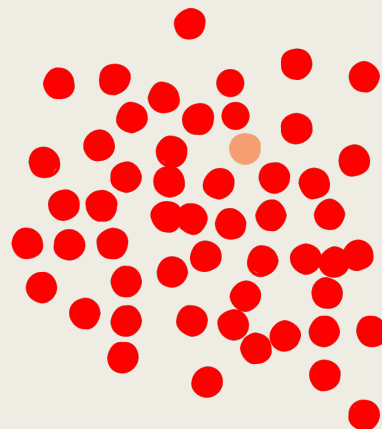
sharpness



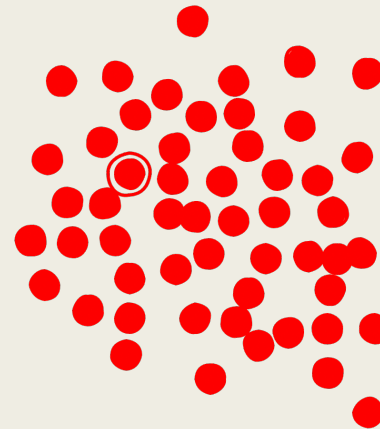
color/hue



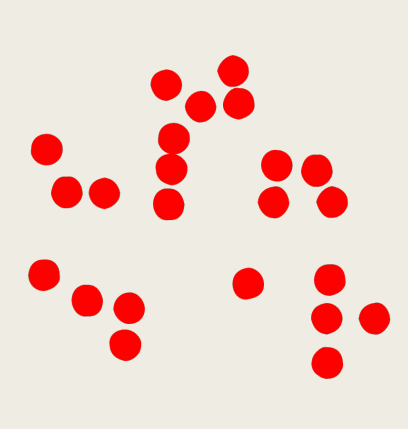
markings



intensity/value



enclosure



numerosity

DECLUTTERING

CLUTTER IS THE ENEMY!

- every element on a page adds **cognitive load**
- identify anything that isn't adding value and **remove**
- think of cognitive load as mental effort required to process information (lower is better)
- Tufte refers to the **data to ink ratio** – “the larger the share of a graphic's ink devoted to data, the better”
- in Resonate, Duarte refers to this as “**maximizing the signal-to-noise ratio**” where the signal is the information or the story we want to communicate.

DECLUTTERING

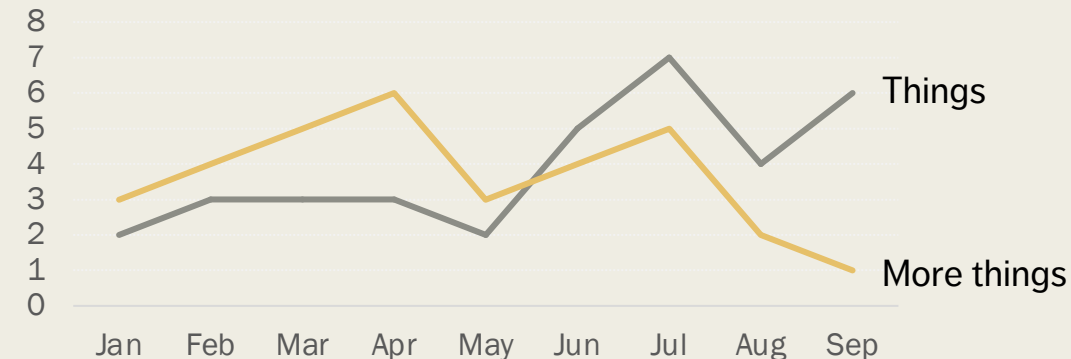
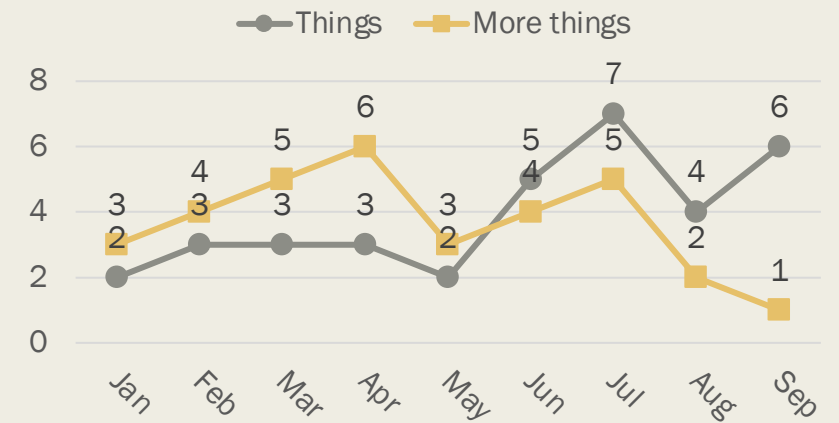
Use **Gestalt Principles** to organize/highlight data in a chart.

Align all the elements (graphs, text, lines, titles, etc.)

- DON'T rely on eye, use position boxes and values

Charts:

- remove border, gridlines, data markers
- clean up axis labels
- label data directly



DECLUTTERING

Use **consistent** font, font size, colour and alignment.

Don't rotate text to anything other than 0 or 90 degrees.

Use **white space**:

- margins should remain free of text and visuals
- don't stretch visuals to edge of page or too close to other visuals
- think of white space as a border

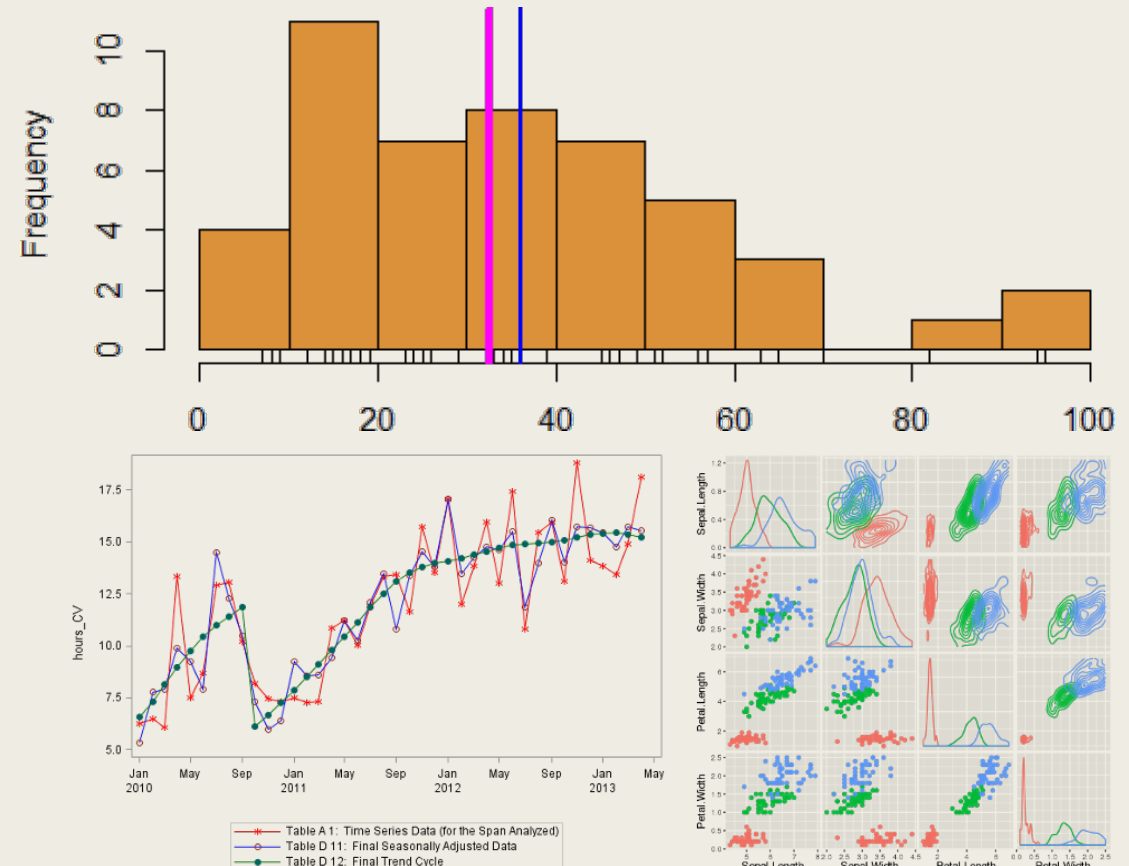
CHART SIZES

Assuming that the chart has been decluttered:

- things of equal importance size **similarly**;
- other things scale to **importance**.

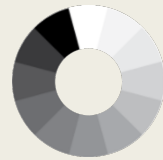
As one rarely puts more than 3-4 charts on a page, there are limited size options.

Perennial exception: **geographical maps** may require more space.

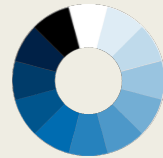


COLOUR SCHEMES

Achromatic



Monochromatic



Complementary



Split complementary



Split-Left/Right Complementary



Analogous



Colour Diad



Colour Triad



Colour Tetrad



COLOUR SCHEMES

When it comes to colour, **less is more**: use it sparingly (graphic designers are taught to “get it right, in black and white”).

Based on the Gestalt Principles, **monochrome** schemes can be particularly effective.

When appropriate, pick scheme based on corporate identity (this maximizes buy in).

Create a template (and stick to it).

Upload images to see what charts look like in various flavours of colour-blindness:

- <https://www.color-blindness.com/coblis-color-blindness-simulator> (there are other tools)