Data Visualization 101

WWCode

Hannah Yan Han

Why Visualize

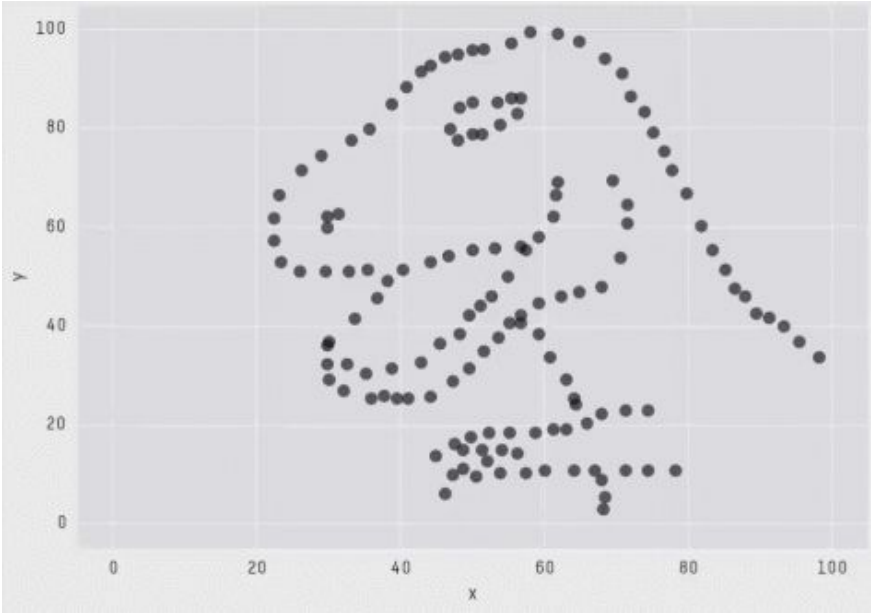
- A picture is worth a thousand words
- Our brains are made to process visual information in parallel and pre-attentively
- Intuitive communication tool

Usage of Datavis

- Explore data
- Communicate findings
- Persuasion
- Just for fun

Why not raw data

Same stats, different data



X Mean: 54.2659224
Y Mean: 47.8313999
X SD : 16.7649829
Y SD : 26.9342120
Corr. : -0.0642526

Data types

Quantitative -- numbers

- -1.5, 3.14

Categorical

- Singapore, France, Argentina

Ordinal -- categories with intrinsic ordering

- Awesome, average, lackluster

Networks - nodes and edges

- Social connections

Spatial

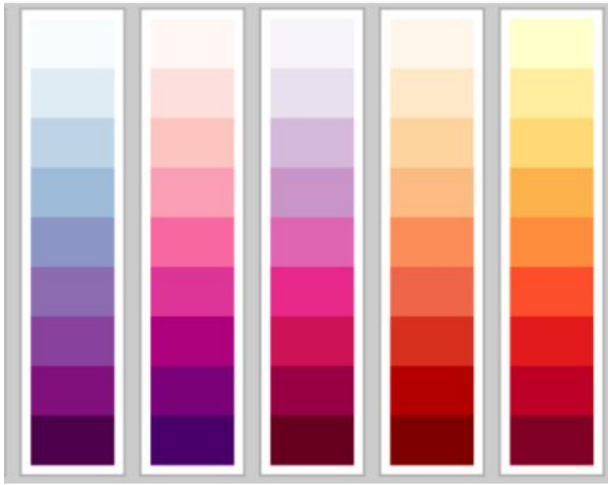
- Maps, physical topologies, etc

Time series

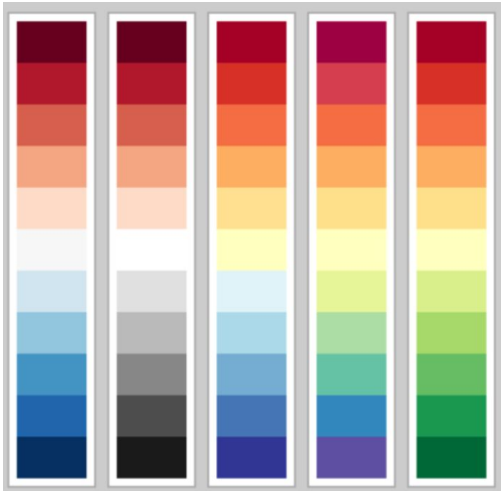
- 2016, 2026, 2036, 2046

Color palette

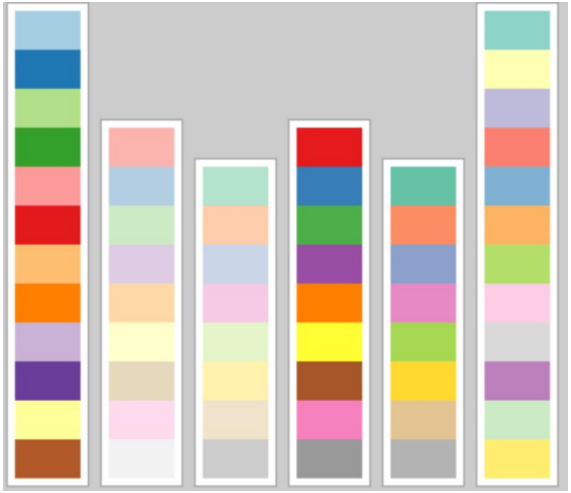
sequential



diverging



qualitative



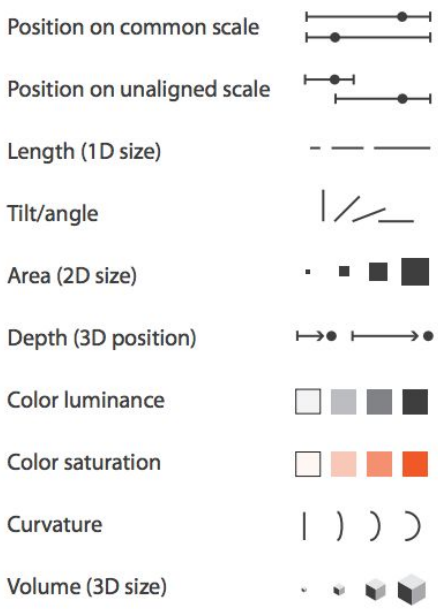
Geometric primitives

- Points

- Lines

- Areas

Visual encodings

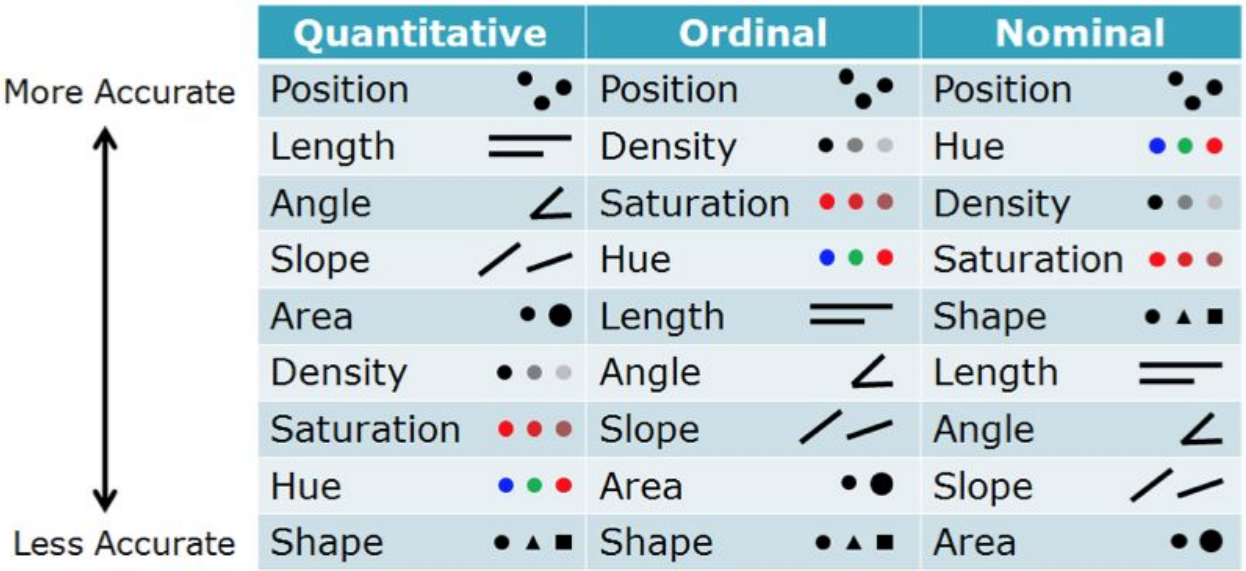


Magnitude channels

Category channels

Adapted from *Tamara Munzner. Visualization Analysis and Design*

Effectiveness



Finding the right chart

The data u have + The question u ask

Independent quantities

Comparing the values of independent variables

Continuous quantities

For data that is continuous, for example when visualizing data over a period of time.

Correlations

These are used when each piece of data has two quantifiable variables which can be plotted on a grid.

Hierarchies

Used when the data has a strict hierarchy that needs to be communicated.

Networks

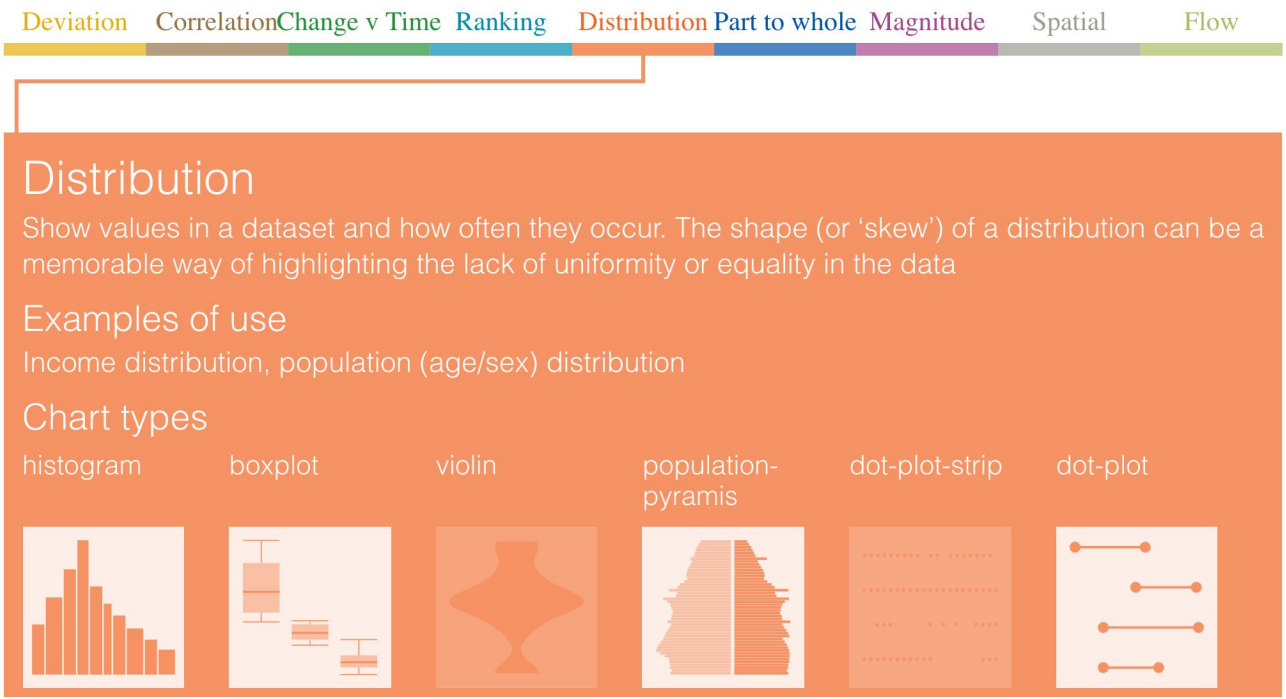
Network visualizations are used when the most important feature of the visualization is to show which data is connected to each other, as opposed to how.

Cartographic

For data that is relevant to specific locations or regions which can be plotted on a map.

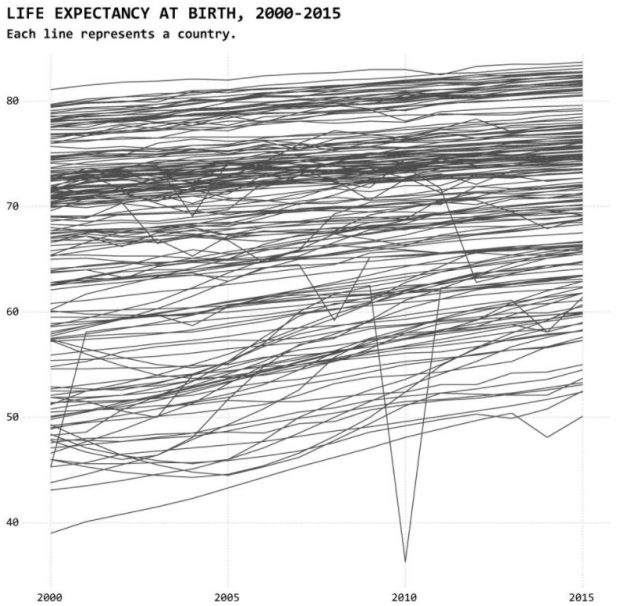
From Joel Laumans An Introduction to Visualizing Data

Designing with Data



From FT Visual Vocabulary

Exploring + Iterating



Example from Nathan Yau [One Dataset, Visualized 25 Ways]

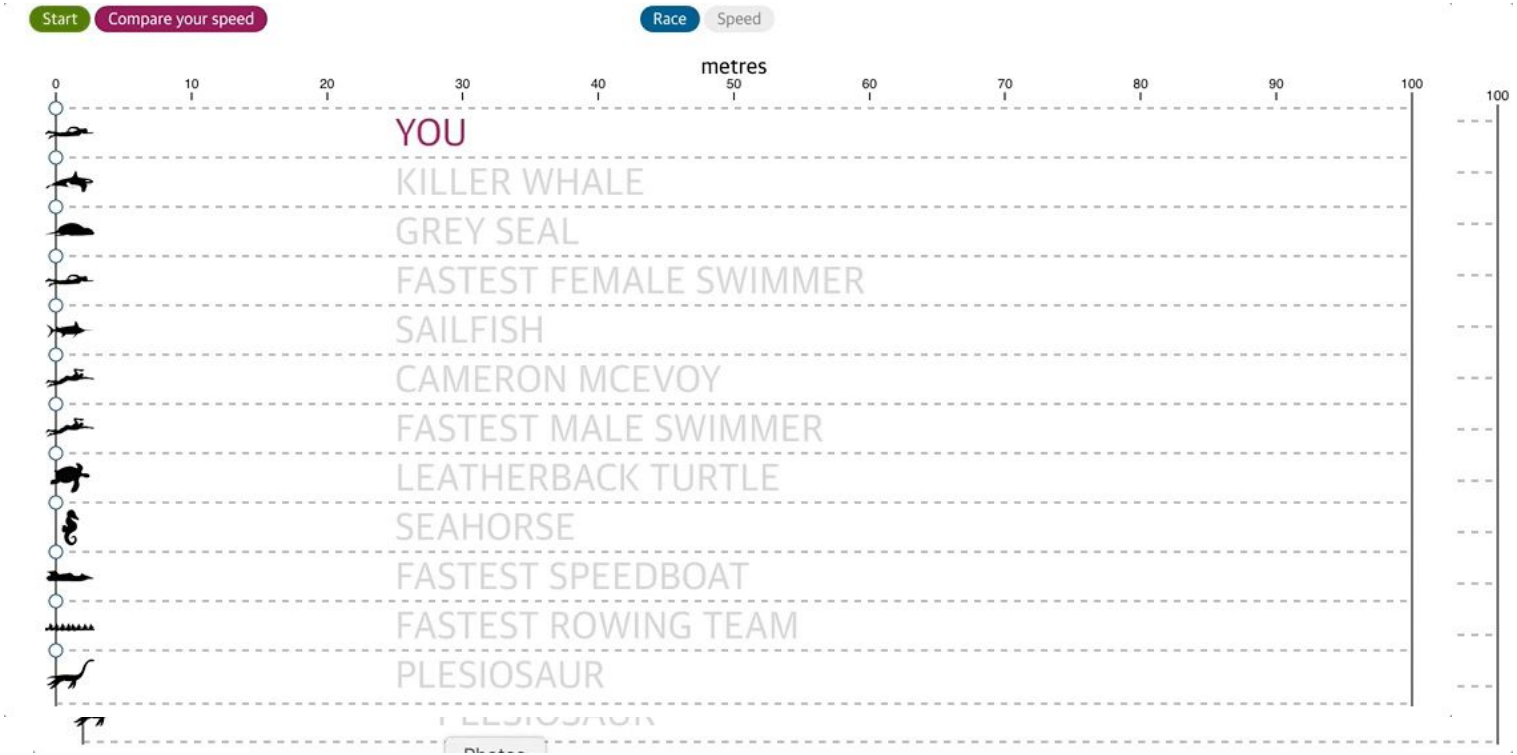
Let the data speak

“ Numbers have an important story to tell. They rely on you to give them a clear and convincing voice ”

Quote from data visualization expert Stephen Few

Storytelling with data

Storytelling with personalized data



Thank you


Hannahyanhan

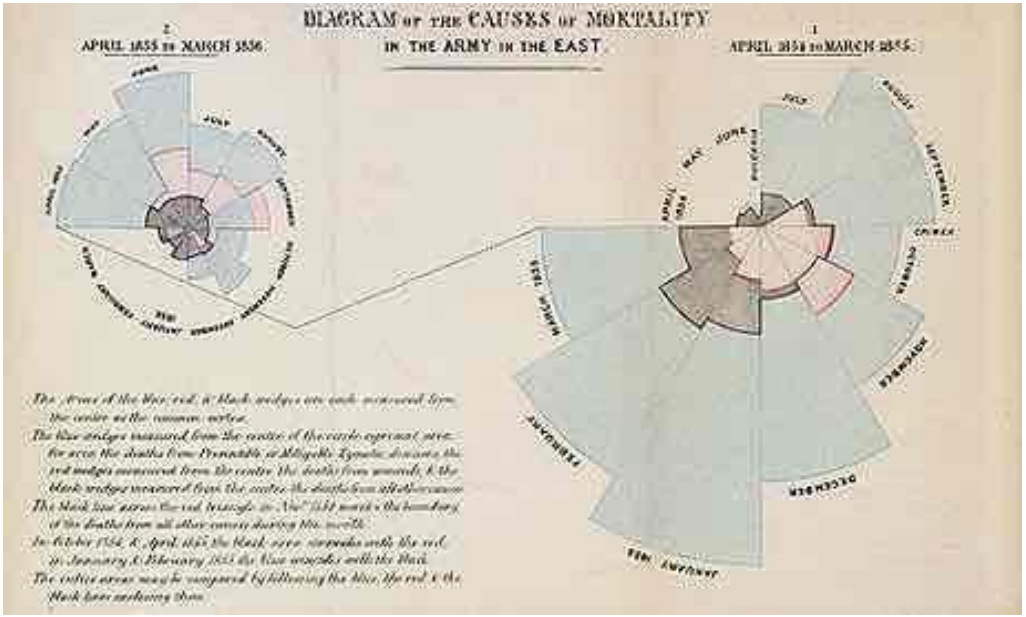

bit.ly/100daysofviz


@hannahyan

Feedback form http://bit.ly/wwcodeworkshop

Appendix

The invention of visual form



[Read more](#)about Nightingale's coxcomb chart

Chart types

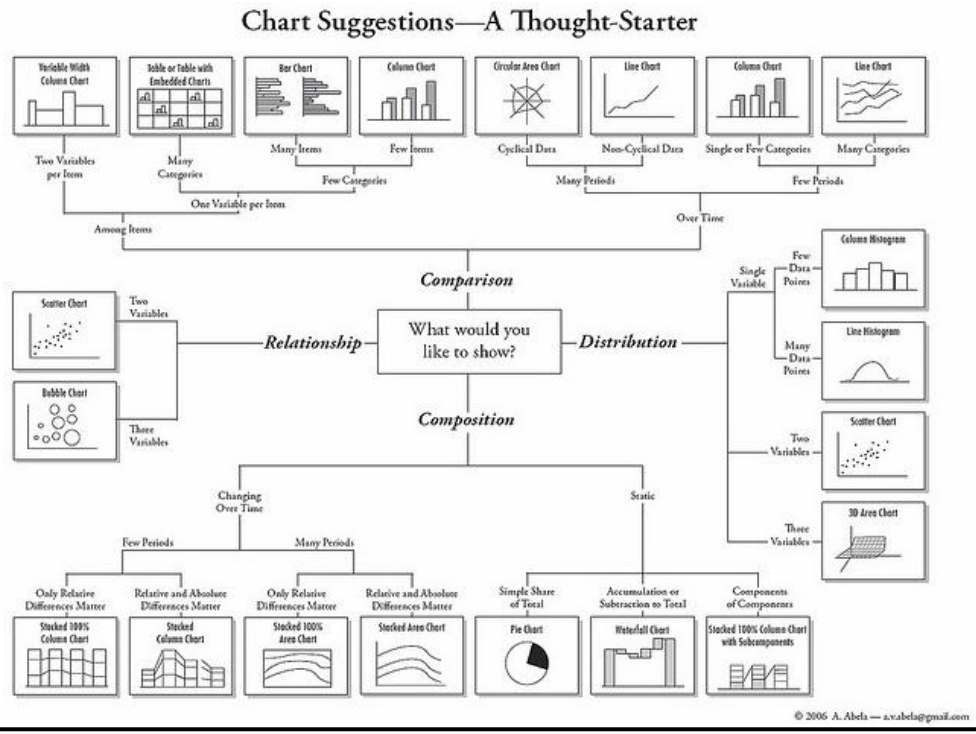
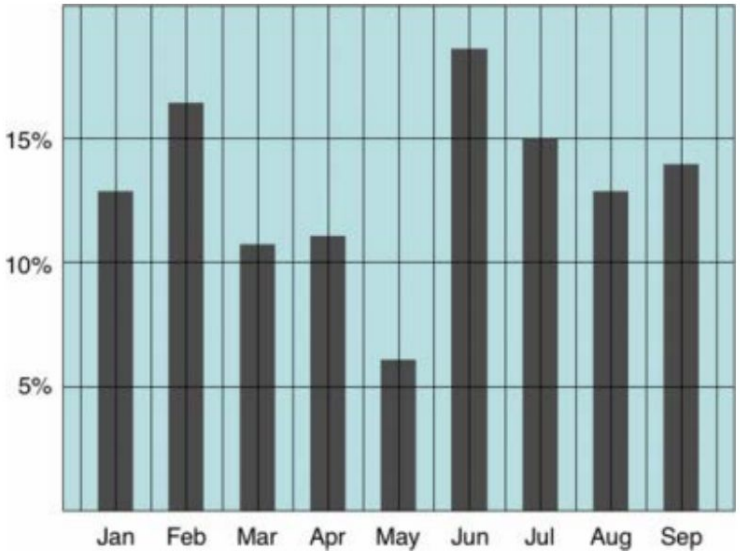
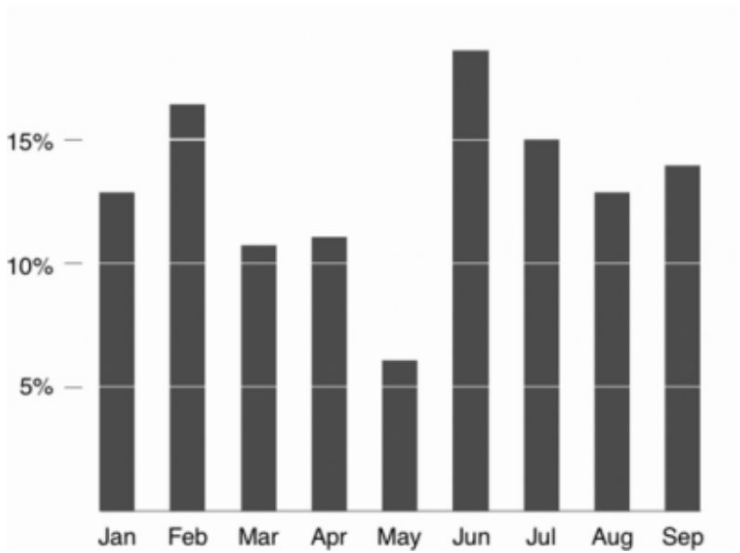


Chart Suggestions—A Thought-Starter
Variable Width Column Chart
Table or Table with Embedded Charts
Bar Chart
Column Chart
Circular Area Chart
Line Chart
Column Chart
Line Chart
Two Variables per Item
Many Categories
Many Items
Few Items
Few Categories
Cyclical Data
Non-Cyclical Data
Single or Few Categories
Many Categories
Many Periods
Few Periods
One Variable per Item
Among Items
Over Time
Comparison
What would you like to show?
Relationship
Distribution
Composition
Scatter Chart
Bubble Chart
Two Variables
Three Variables
Single Variable
Few Data Points
Many Data Points
Column Histogram
Line Histogram
Scatter Chart
3D Area Chart
Two Variables
Three Variables
Changing Over Time
Static
Few Periods
Many Periods
Only Relative Differences Matter
Relative and Absolute Differences Matter
Only Relative Differences Matter
Relative and Absolute Differences Matter
Simple Share of Total
Accumulation or Subtraction to Total
Components of Components
Stacked 100% Column Chart
Stacked Column Chart
Stacked 100% Area Chart
Stacked Area Chart
Pie Chart
Waterfall Chart
Stacked 100% Column Chart with Subcomponents
© 2006 A. Abela — a.v.abela@gmail.com

Chart junks



Low Data Ink (BAD)



High Data Ink (GOOD)

Some principles of graphics

- Principle 1: Show comparisons
- Principle 2: Show causality, mechanism, explanation
- Principle 3: Show multivariate data
- Principle 4: Integrate multiple modes of evidence
- Principle 5: Describe and document the evidence
- Principle 6: Content is king

From Edward Tufte *Beautiful Evidence*