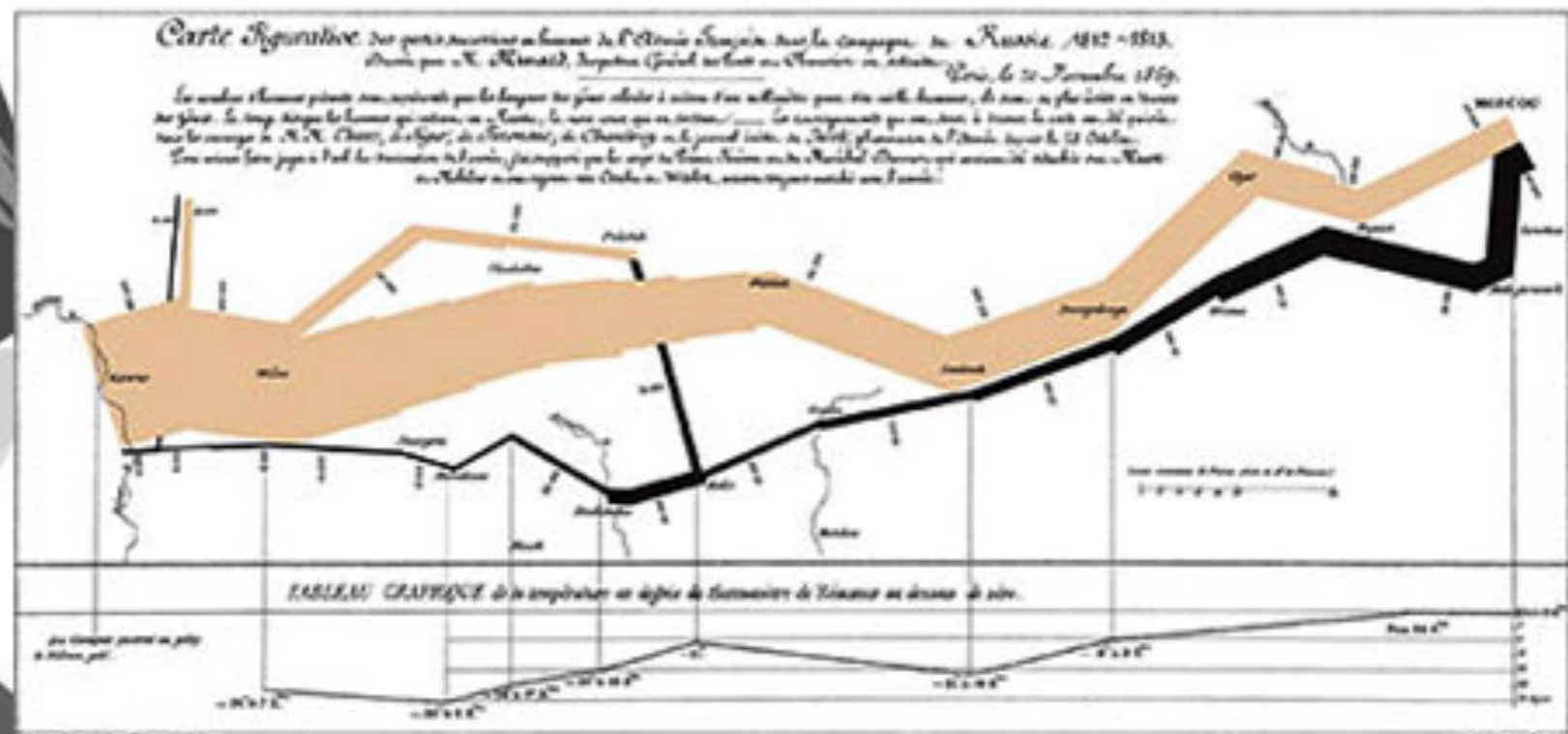# LESSONS FROM
EDWARD TUFTE

WHO IS EDWARD TUFTE?

an analytical design theorist, educator, and landscape sculptor best known for his self-published books on analytical design

Illustrated by Merchant for the Brunswick Review
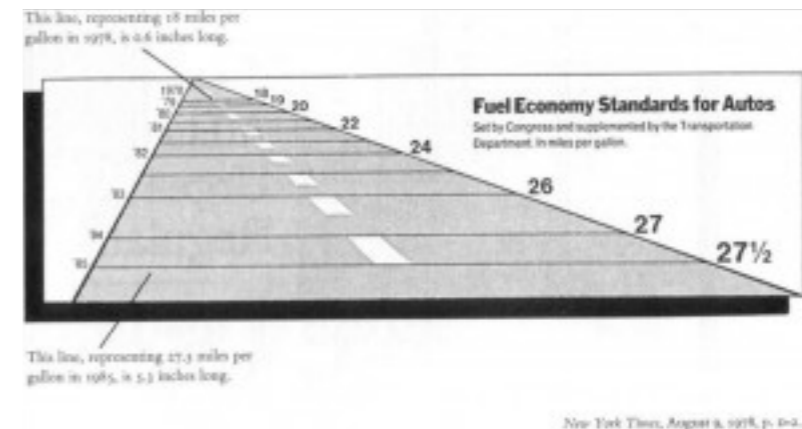
1

STRIVE FOR
GRAPHICAL INTEGRITY

Visual representations of data

must tell the truth.

THE LIE FACTOR

Calculated by dividing the size of the

effect shown in the graphic by the size

of the effect in the data.

If the Lie Factor is greater than 1
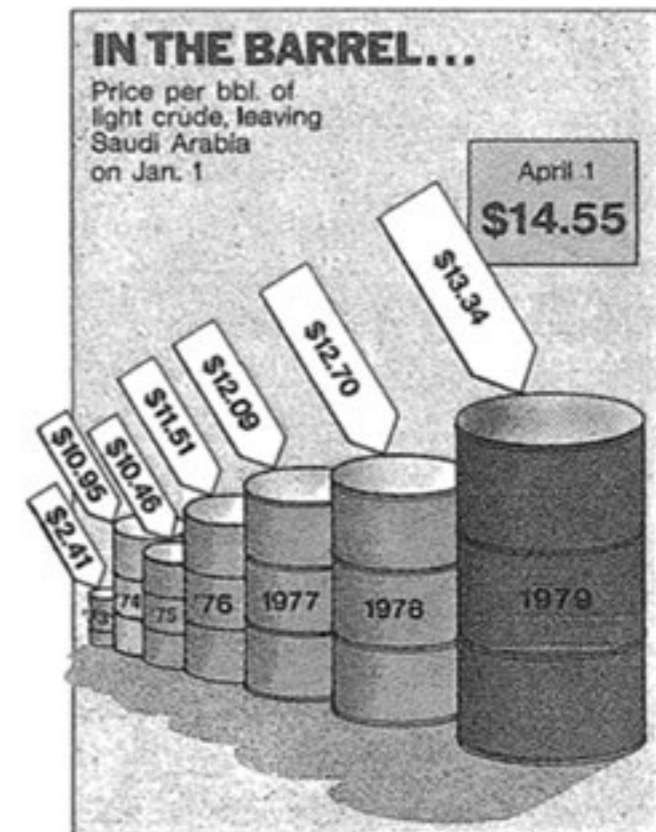
the graph overstates the effect.

According to Tufte the Lie Factor of this graph is 14.8. A numerical change of 53% is represented by a graphical change (size of horizontal lines) of 783%.
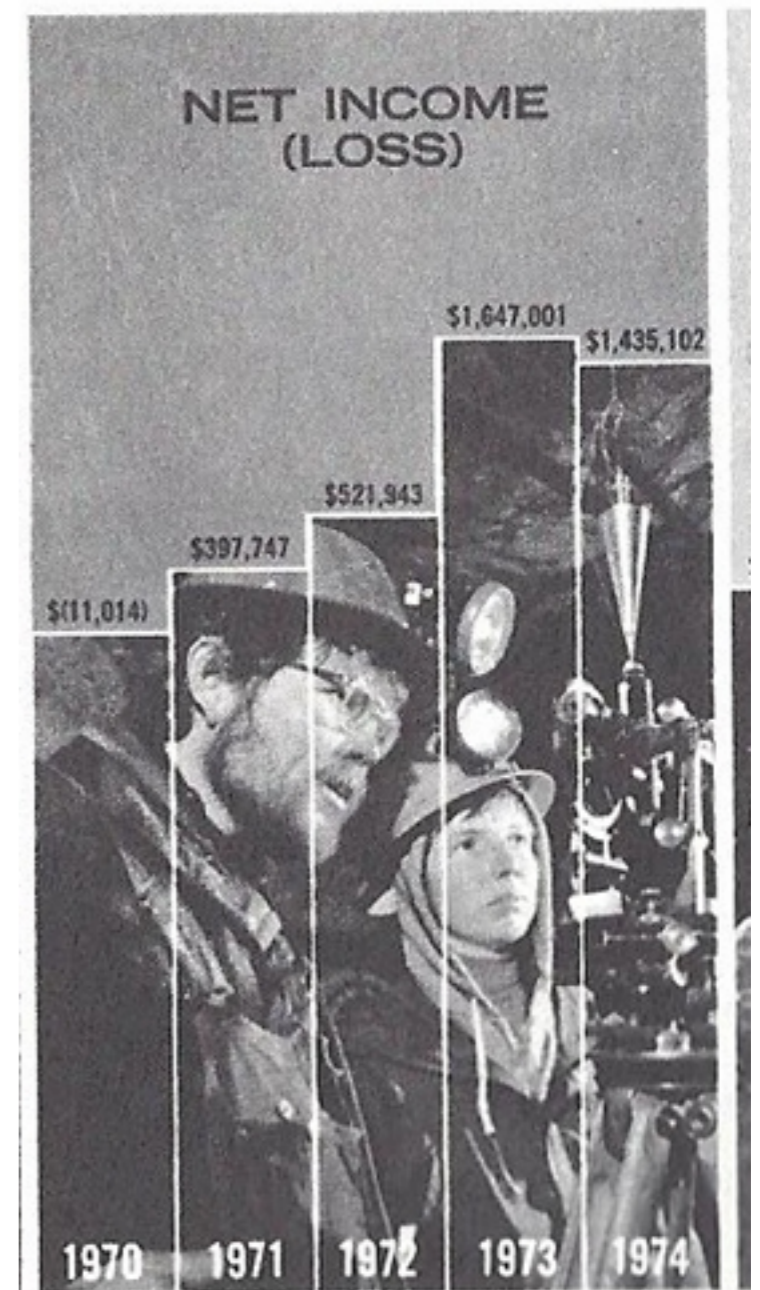
1ST PRINCIPLE

The representation of numbers, as physically measured on the surface of the graph itself, should be directly proportional to the numerical quantities represented

IN THE BARREL...
Price per bbl. of
light crude, leaving
Saudi Arabia
on Jan. 1

April 1
$14.55

$13.34

$12.70

$12.09

$11.51

$10.46

$10.95
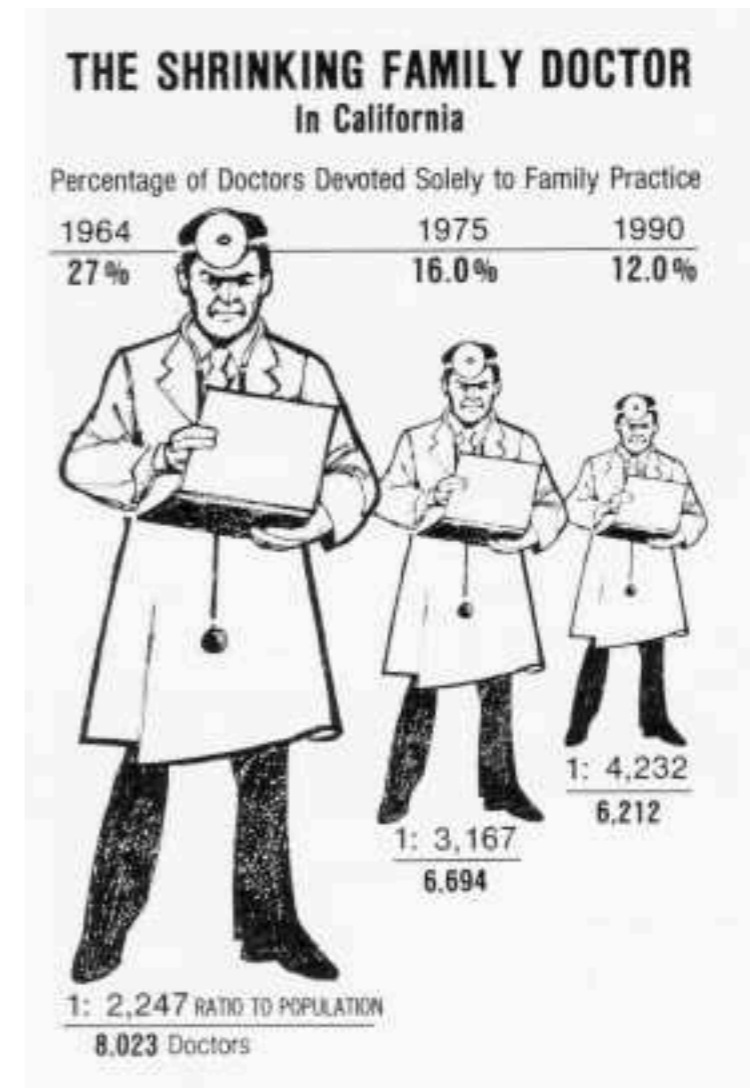
$2.41

73 74 75 76 1977 1978 1979

2ND PRINCIPLE

Clear, detailed and thorough labeling should be used to defeat graphical distortion and ambiguity. Write out explanations of the data on the graph itself. Label important events in the data.
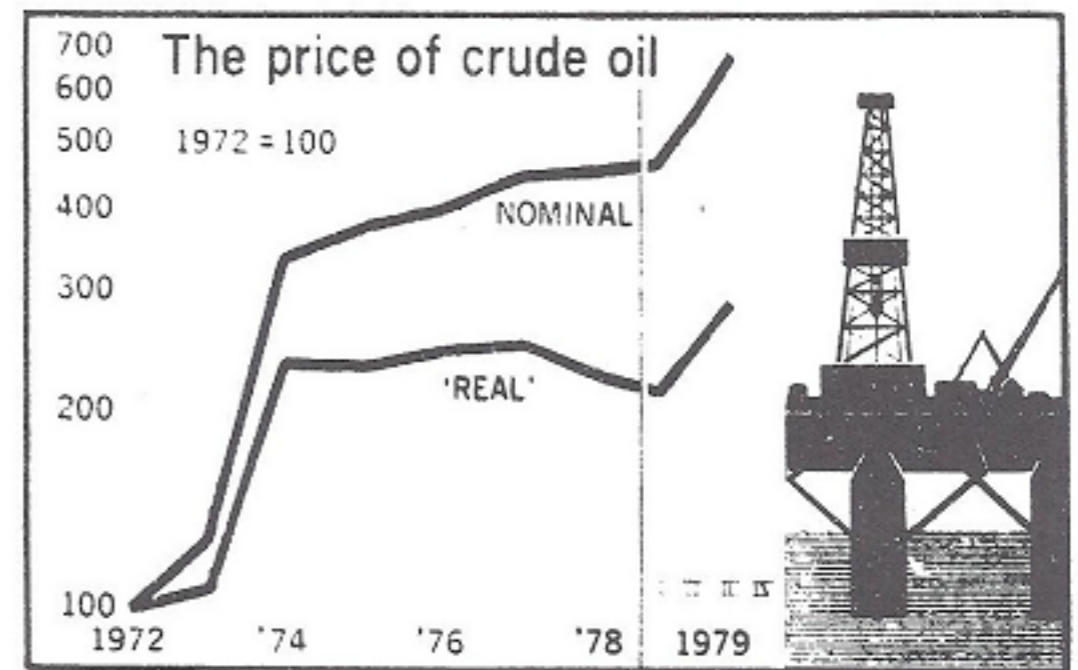
3RD PRINCIPLE

Show data variation, not design variation.

THE SHRINKING FAMILY DOCTOR
In California

Percentage of Doctors Devoted Solely to Family Practice

1964	1975	1990
27%	16.0%	12.0%

1: 2,247 RATIO TO POPULATION
8,023 Doctors

1: 3,167
6,694

1: 4,232
6,212

4TH PRINCIPLE

In time-series displays of money, deflated and standardized units of monetary measurement are nearly always better than nominal units.
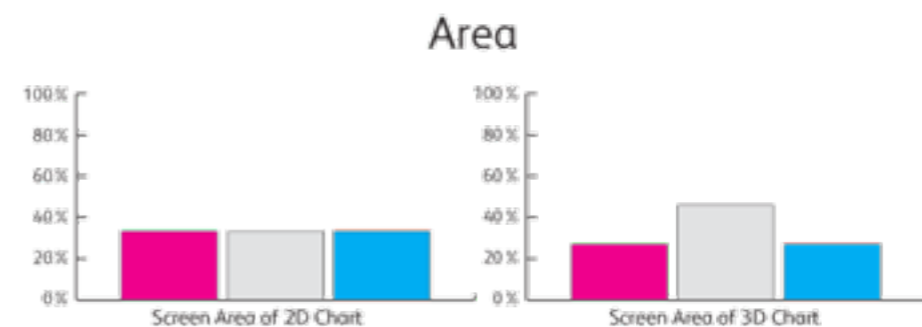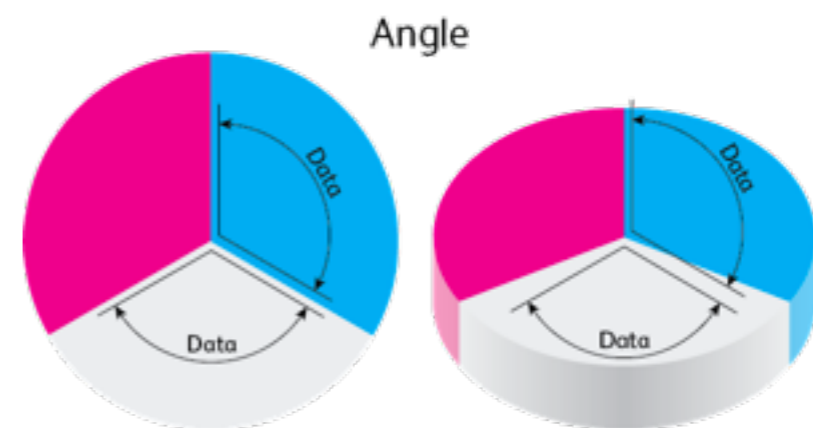

The price of crude oil
1972 = 100
NOMINAL
'REAL'

5TH PRINCIPLE

The number of information carrying (variable) dimensions depicted should not exceed the number of dimensions in the data.
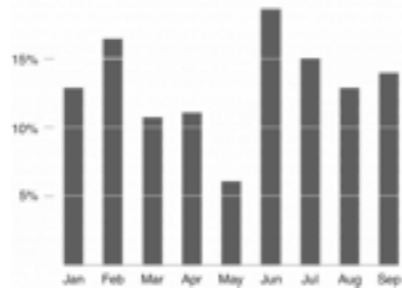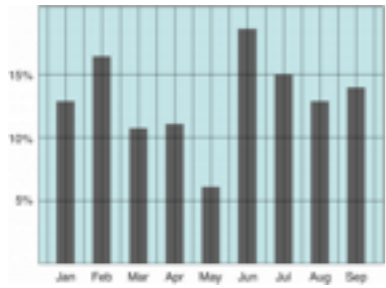
Graphics must not quote data out of context.

Angle

Area

2

MAXIMIZE
DATA INK

The ink on a graph that represents data. Good graphical representations

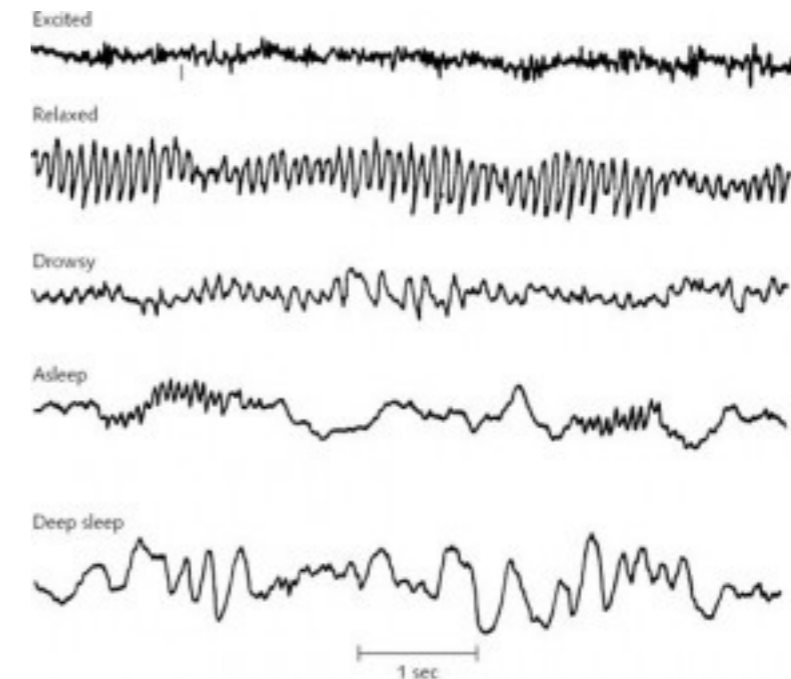maximize data-ink and erase as much non-data-ink as possible.

The data-ink ratio is calculated by 1 minus the proportion of the graph

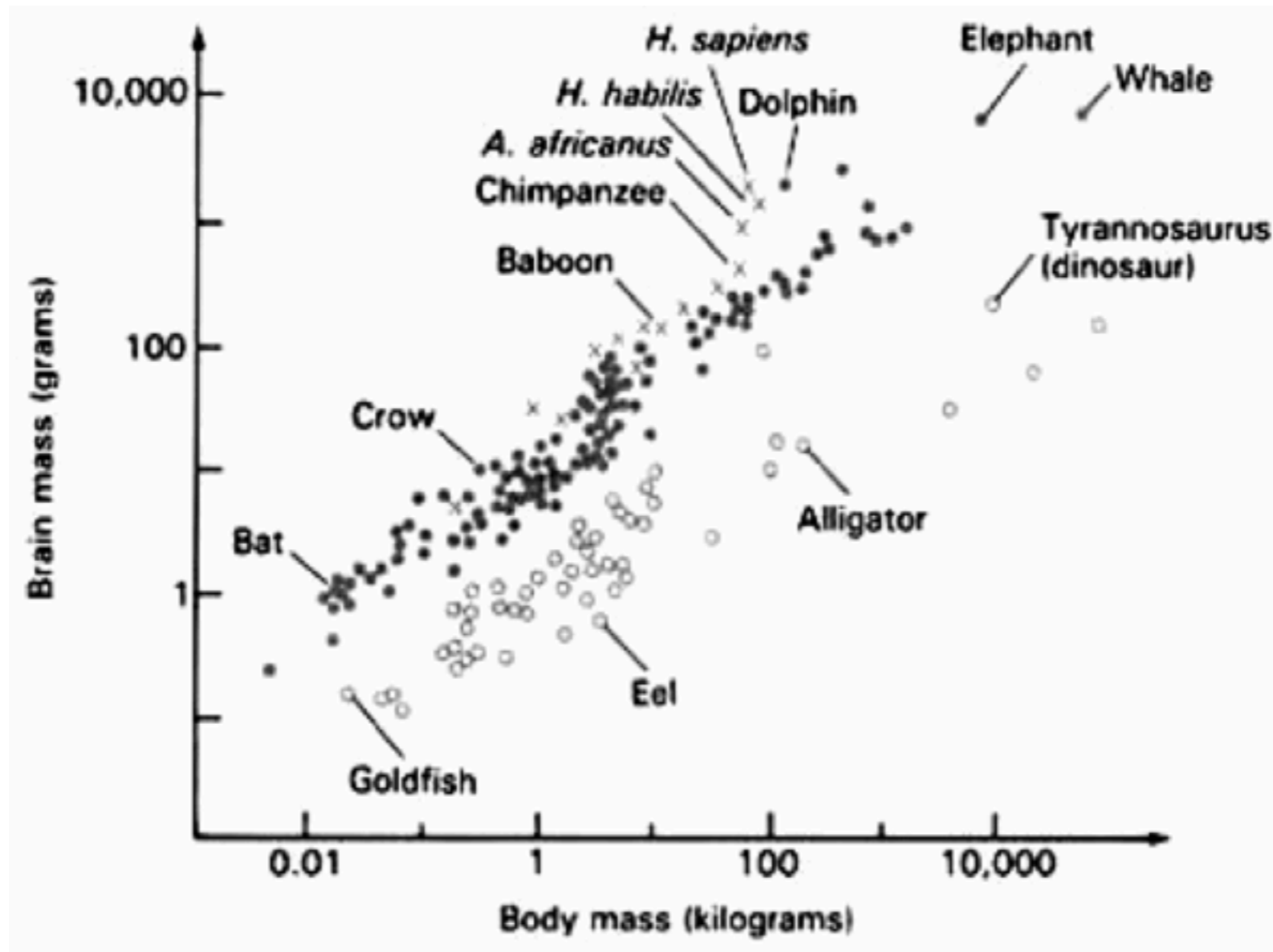that can be erased without loss of data-information.

5 PRINCIPLES

1. Above all else show data.

2. Maximize the data-ink ratio.

3. Erase non-data-ink.

4. Erase redundant data-ink.

5. Revise and edit

Excited

Relaxed
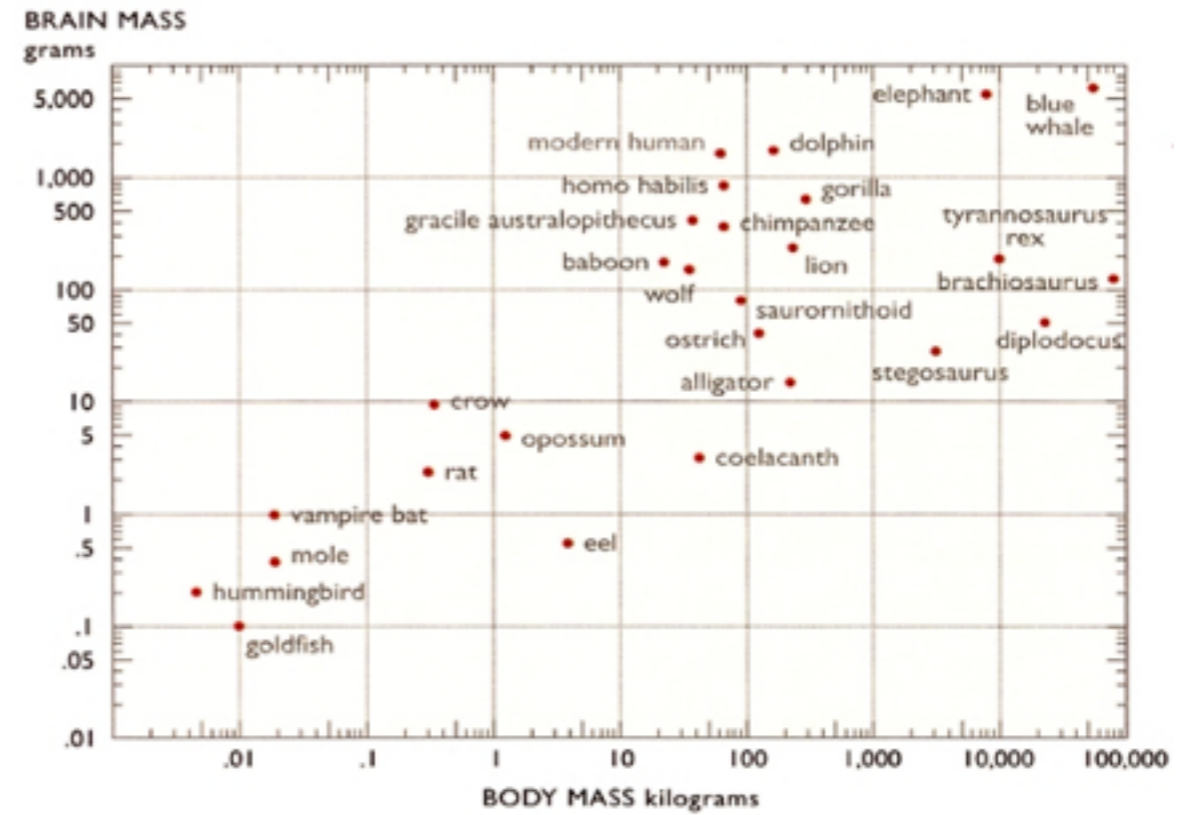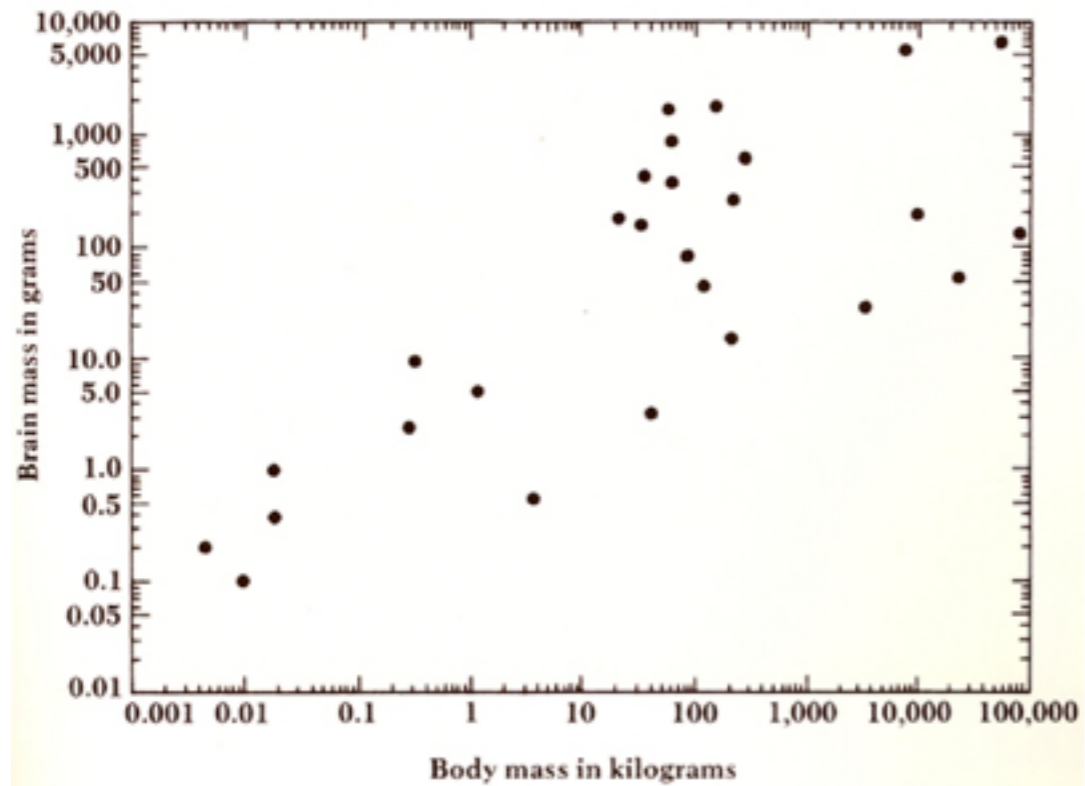
Drowsy

Asleep

Deep sleep

1 sec

It's an electroencephalogram – a graph that records the electrical activity from the brain. This graph would have a very high data-ink ratio of 1.
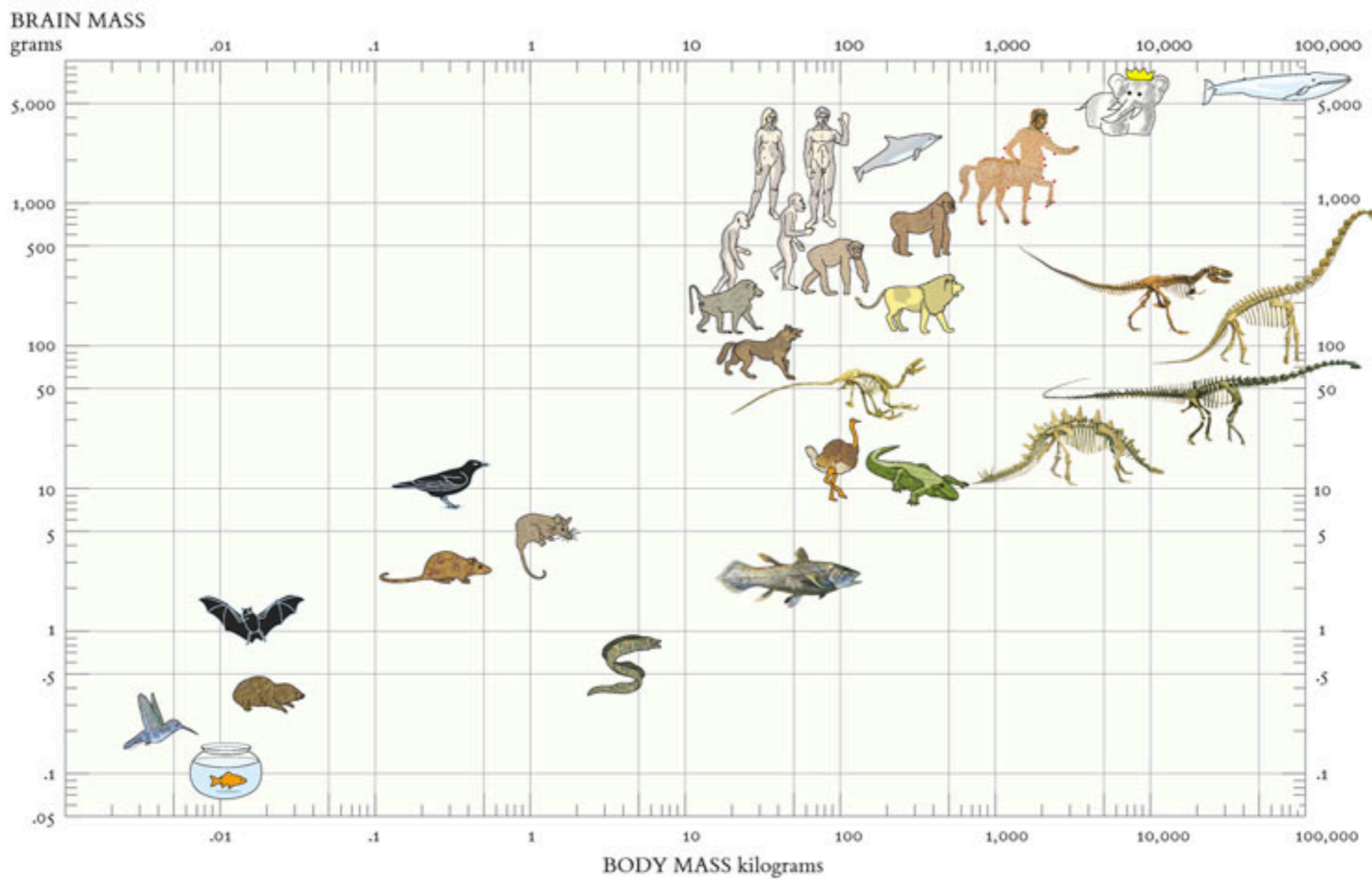
2 DATA INK

Graph from Carl Sagan's The Dragons of Eden

2 DATA INK

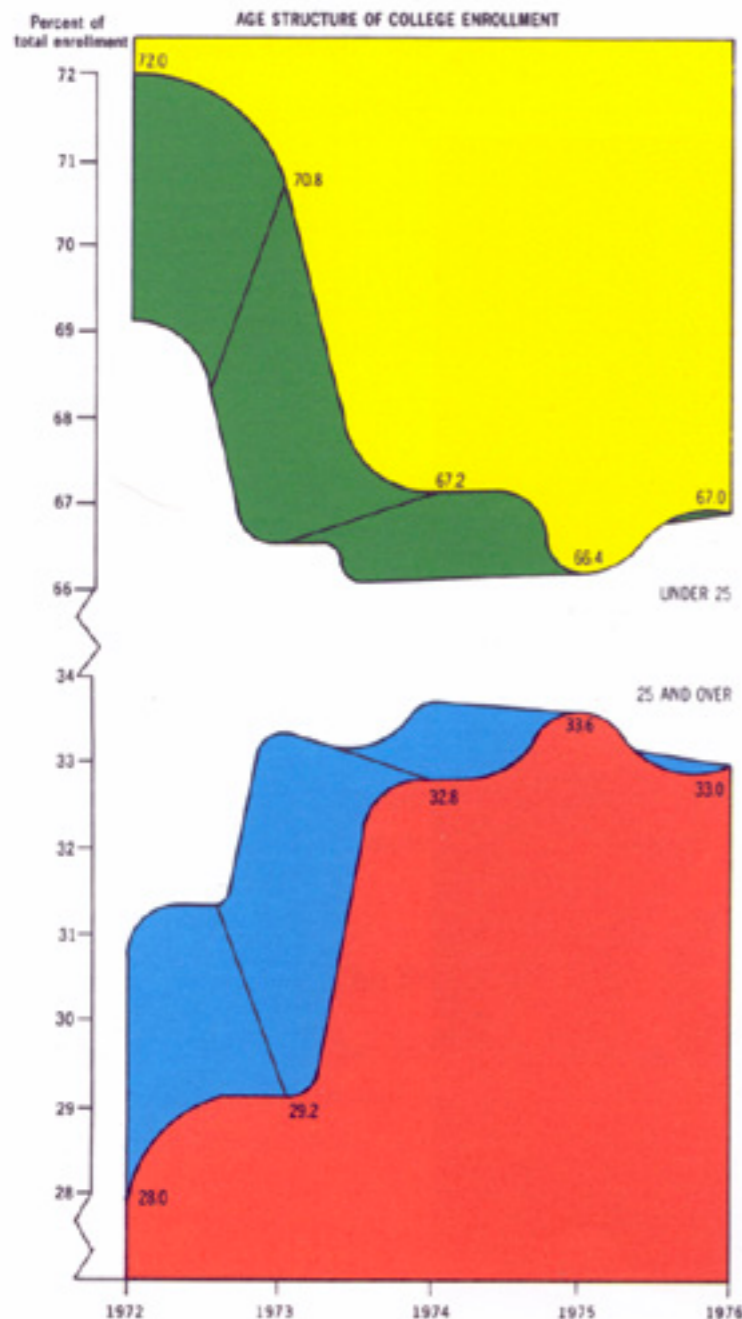Tufte's redesign

2 DATA INK

Tufte's redesign from Beautiful Evidence

3

AVOID
CHART JUNK

The excessive and unnecessary use of graphical effects in graphs used to demonstrate the graphic ability of the designer rather than display the data.

3 CHART JUNK

This is according to Tufte possibly the worst graph ever :

"*A series of weird three-dimensional displays appearing in the magazine of American Education in the 1970's delighted the connoisseurs of the graphically preposterous. Here five colors report, almost by happenstance, only 5 pieces of data (since the division within each adds to 100%). This may well be the worst graphic ever to find its way into print.*"

4

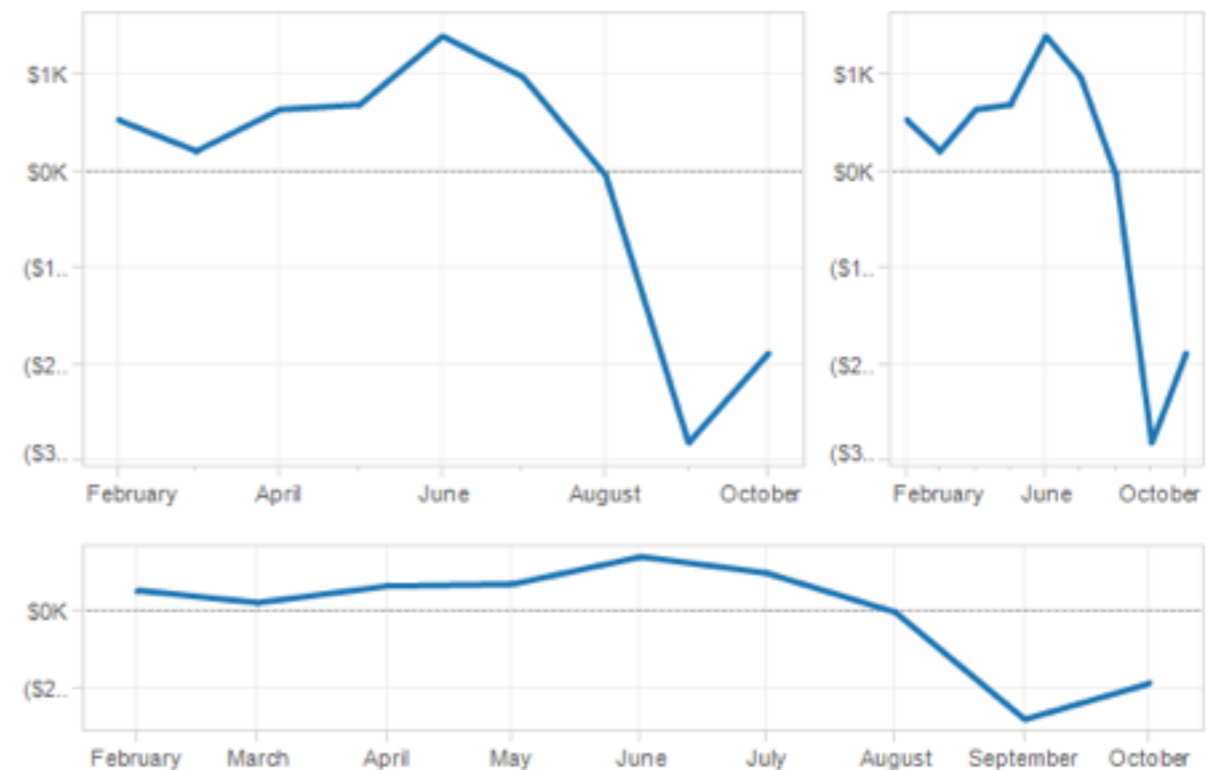## AIM FOR HIGH
DATA DENSITY

The proportion of the total size of the graph that

is dedicated displaying data.

4 DATA DENSITY

SHRINK PRINCIPLE

Maximize data density and the size of the data matrix within reason. One way of achieving this is through the Shrink Principle. Most graphs can be shrunk way down without losing legibility or information.

5

USE
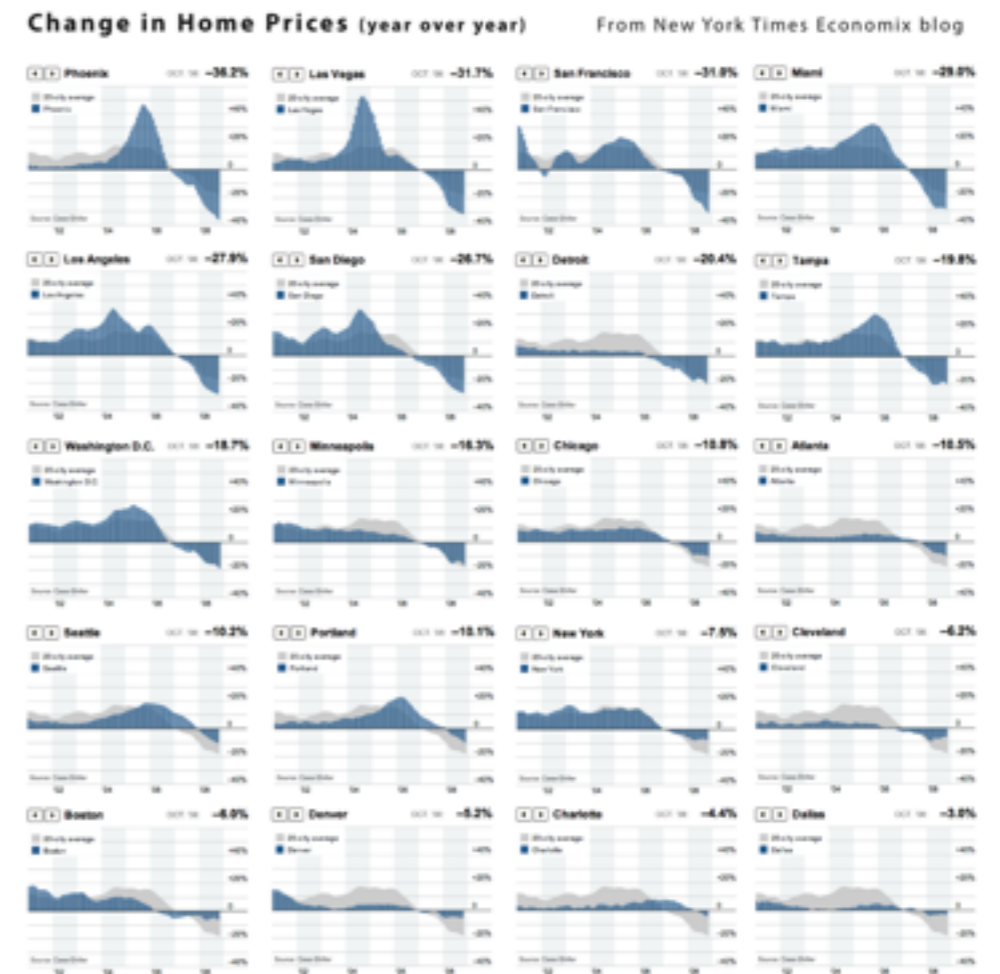# CLASSIC DESIGN SOLUTIONS

5 DESIGN SOLUTIONS

SMALL MULTIPLES

Series of the same small

graph repeated in one visual.

Change in Home Prices (year over year)   From New York Times Economix blog
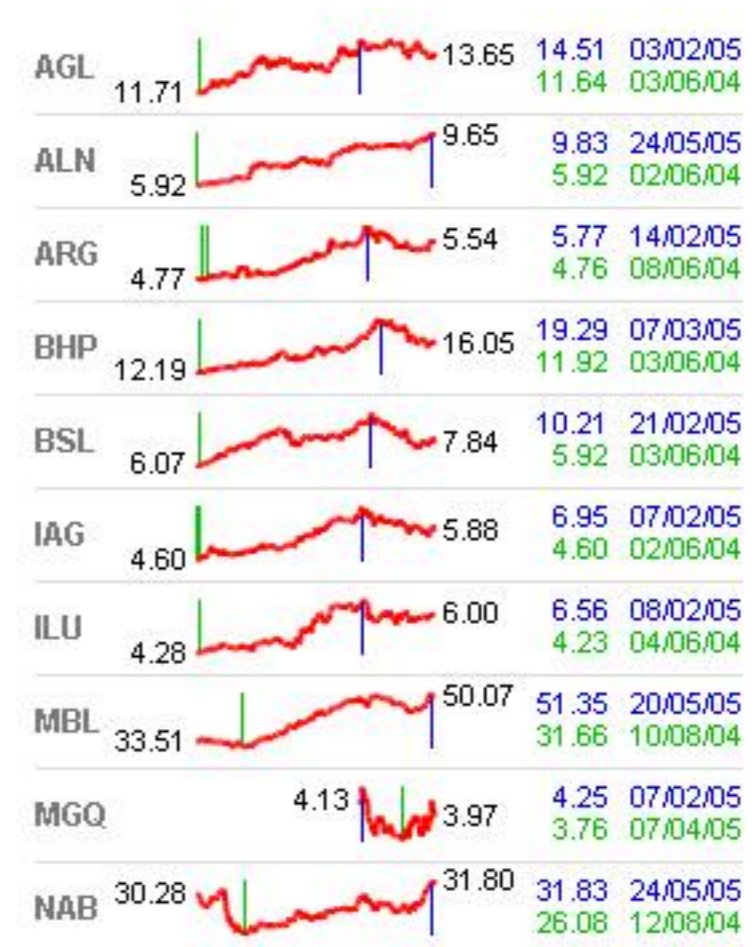
SPARKLINES

Small multiples are a great tool to visualize large quantities of data and with a high number of dimensions.

Sparklines are data-intense, design-simple, word-sized graphics.

5 DESIGN SOLUTIONS

TIME SERIES

The most frequently used form of graphic design.

One dimension, usually the horizontal, is time, and the graphics march along showing variation as time proceeds. Most visualization time-series works are videos, which show time by changing the picture, requiring the user to remember what came before.

MICRO/MACRO COMPOSITION

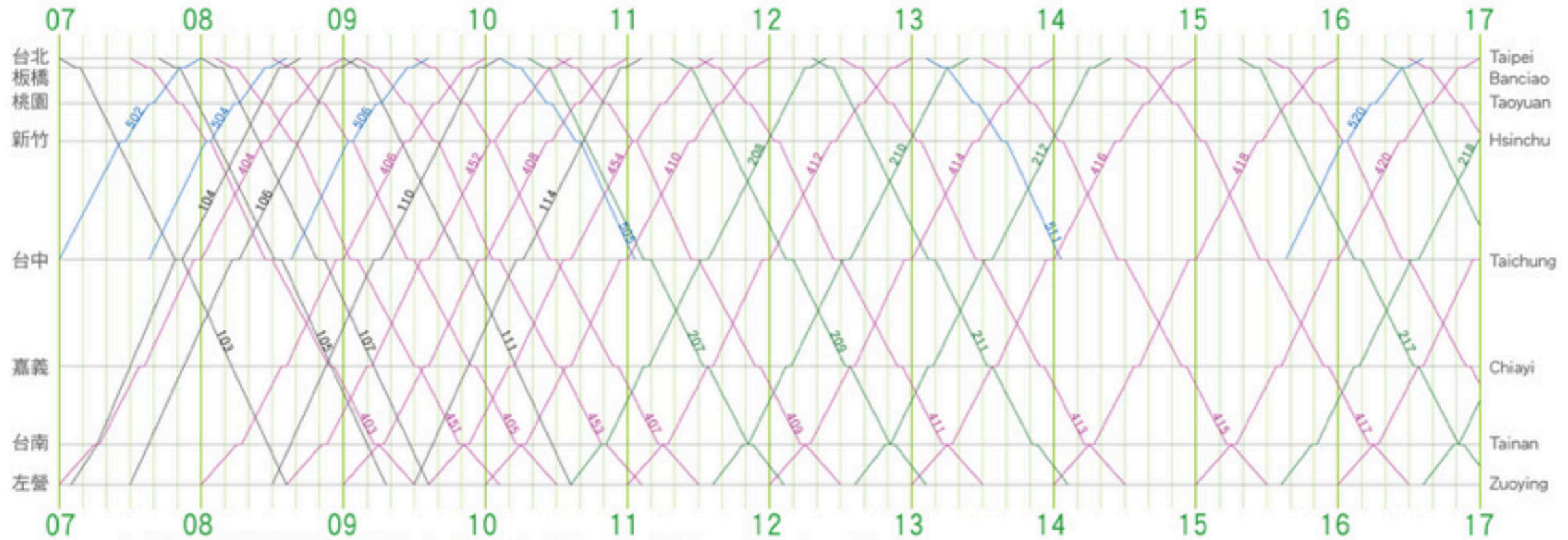Refers to an approach where a visualization contains enormous detail, but an overall pattern emerges. Panorama, vista, and prospect deliver to viewers the freedom of choice that derives from an overview, a capacity to compare and sift through detail. And that micro-information, like smaller texture in landscape perception, provides a credible refuge where the pace of visualization is condensed, slowed, and personalized.

台灣高鐵列車運行圖 | Train Operation Diagram of Taiwan High Speed Rail
2007-July-27

6

APPLY
AESTHETICS & TECHNIQUE

6 AESTHETICS & TECHNIQUE

- have a properly chosen format and design

- use words, numbers, and drawing together

- reflect a balance, a proportion, a sense of relevant scale

- display an accessible complexity of detail

- have a narrative quality, a story to tell about the data

- draw elements in a professional manner, with the

  technical details of production done with care

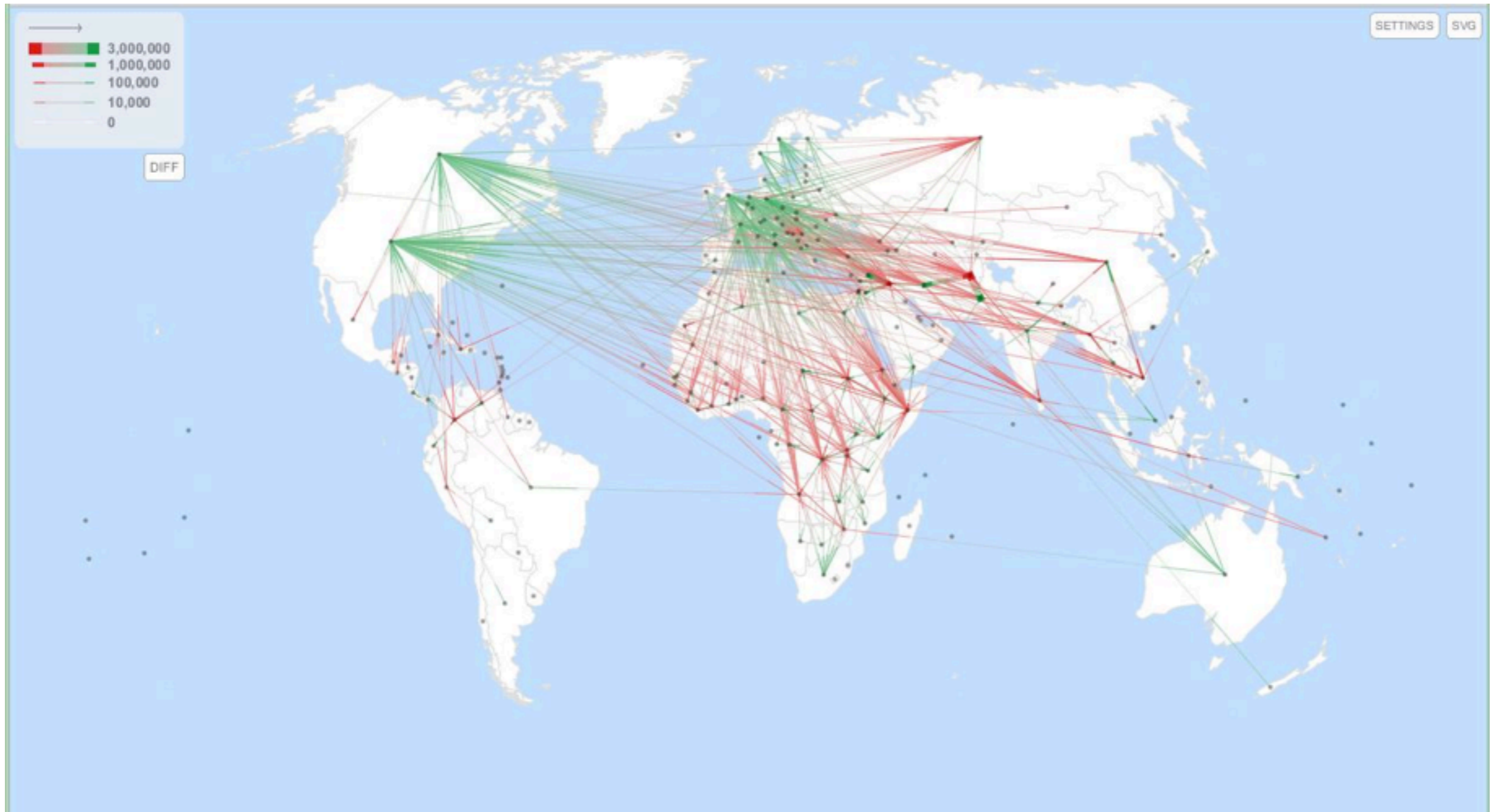- avoid content-free decoration, including chartjunk

Juan Velasco
Cornell ornithologist Edwing Scholes and biologist/photographer Tim Laman
Senior Graphics Editor Fernando Baptista, Graphics Specialist Maggie Smith and freelance researcher
Fanna Gebreyesus
National Geographic

CLUTTER

- If a visualization is too cluttered, don't remove data, change the design. Credibility comes from detail and in many cases one can clarify a design by adding detail.

- High-density designs also allow viewers to select, to narrate, to recast and personalize data for their own uses.

- Data-thin, forgetful displays move viewers toward ignorance and passivity, and at the same time diminish the credibility of the source.

JFlowMap is a research prototype developed at the University of Fribourg that experiments various visualization techniques for spatial interactions, i.e. interactions between pairs of geographic locations. These can be migrations, movement of goods and people, network traffic, or any kind of entities "flowing" between locations.

Hyperakt and Ekene Ijeoma visualized migrations over time and space in The Refugee Project
http://www.therefugeeproject.org
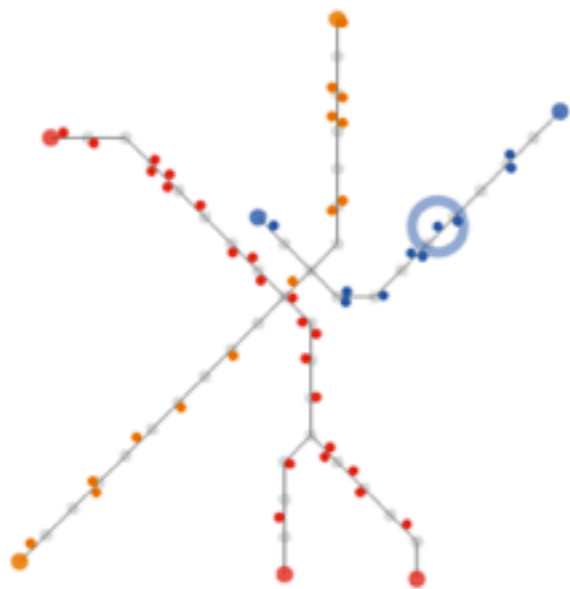
CLUTTER

- Empty space may reduce clutter, but it is not how much empty space there is, but rather how it is used. It is not how much information there is, but rather how effectively it is arranged.

- Low density computer displays lead to spreading information out over many screens or dialog boxes. Place information adjacent in space, not stacked in time, to avoid the `Where am I?' problem.

For a graduate project, Michael Barry and Brian Card explored the Boston subway system through a set of annotated interactives that show train routes, usage, and scheduling.
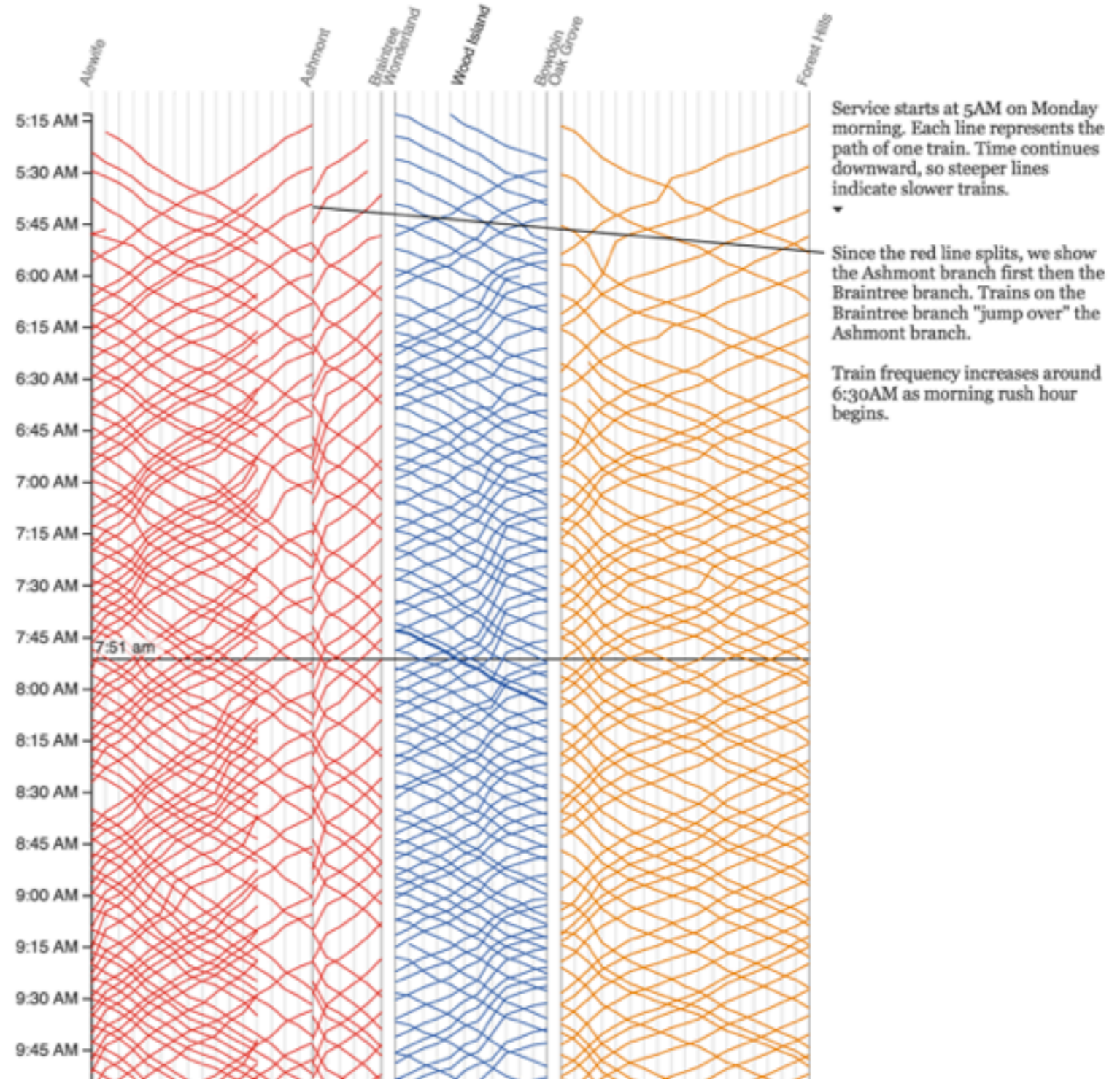http://mbtaviz.github.io/

Subway Trips on Monday February 3, 2014

Alewife

Ashmont

Braintree
Wonderland

Wood Island

Bowdoin
Oak Grove

Forest Hills

5:15 AM
5:30 AM
5:45 AM
6:00 AM
6:15 AM
6:30 AM
6:45 AM
7:00 AM
7:15 AM
7:30 AM
7:45 AM 7:51 am
8:00 AM
8:15 AM
8:30 AM
8:45 AM
9:00 AM
9:15 AM
9:30 AM
9:45 AM

Service starts at 5AM on Monday morning. Each line represents the path of one train. Time continues downward, so steeper lines indicate slower trains.

Since the red line splits, we show the Ashmont branch first then the Braintree branch. Trains on the Braintree branch "jump over" the Ashmont branch.

Train frequency increases around 6:30AM as morning rush hour begins.

Locations of each train on the red, blue, and orange lines at 7:51 am. Hover over the diagram to the right to display trains at a different time.

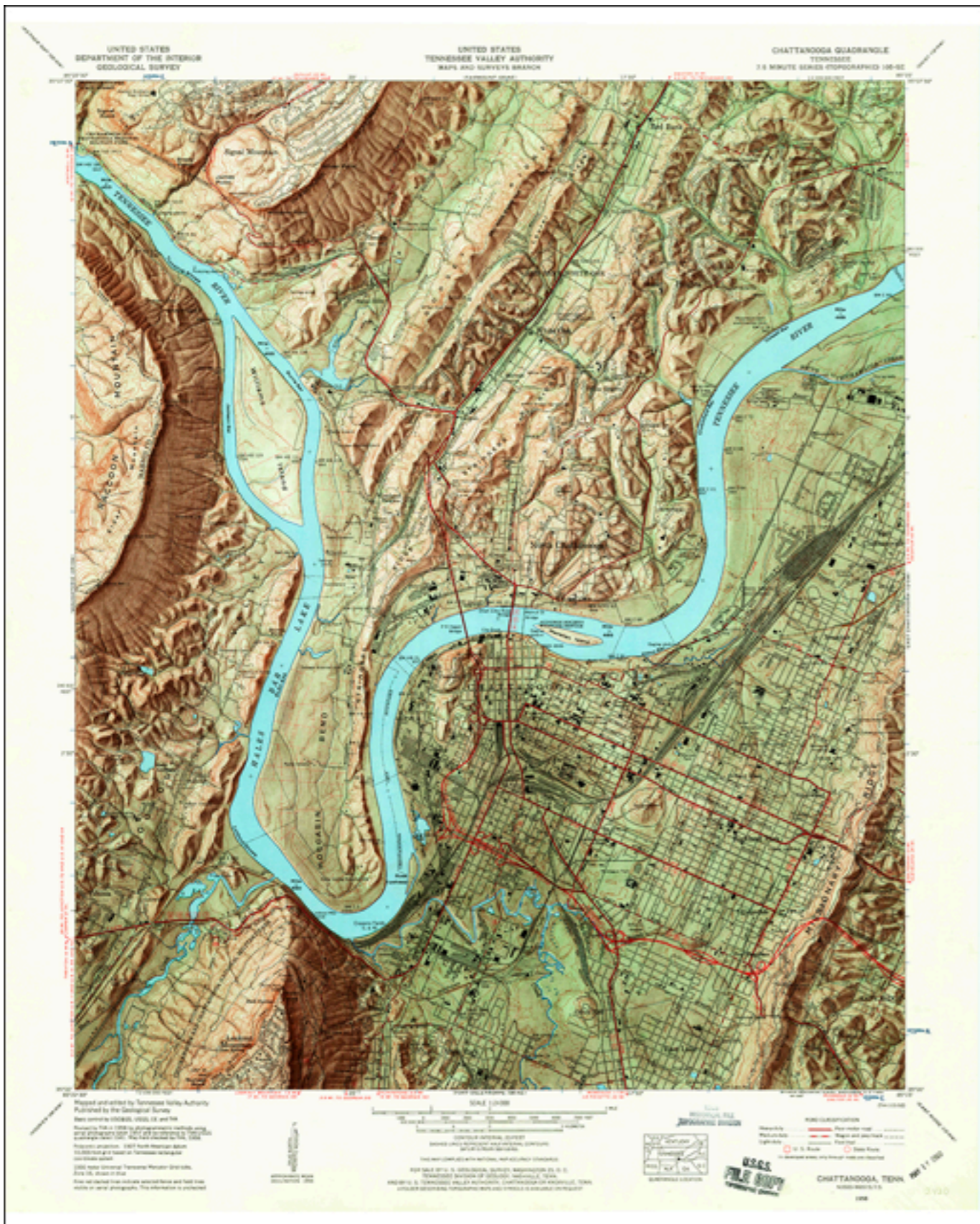Trains are on the right side of the track relative to the direction they are moving.

See the morning rush-hour, midday lull, afternoon rush-hour, and the evening lull.

6 AESTHETICS & TECHNIQUE

LAYERING AND SEPARATION

Implies using color or other differentiation to separate important classes of information. Consider a colormapped surface that requires annotation. If the colormap uses all possible colors, positioning annotation will be difficult because of color clashes. A better approach might be to use intensity of a single hue for the colormap, leaving visual space for addition information; i.e., the annotation.
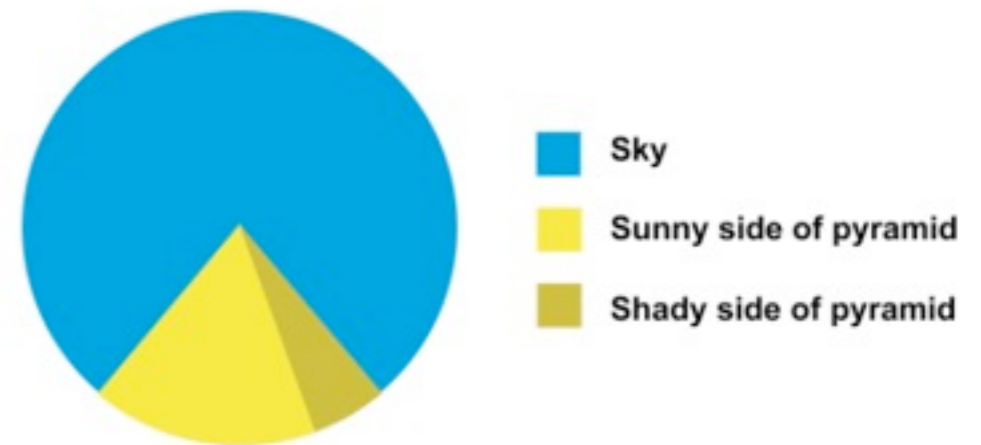
Muted colors, subtle shading and thin contour lines allow multiple types of data to be layered together in this 1958 topographic map of Chattanooga, Tennessee.

1 + 1 = 3 OR MORE

Effective layering of information is often

difficult because an omnipresent, yet

subtle, design issue is involved:

the various elements collected together

interact, creating non-information
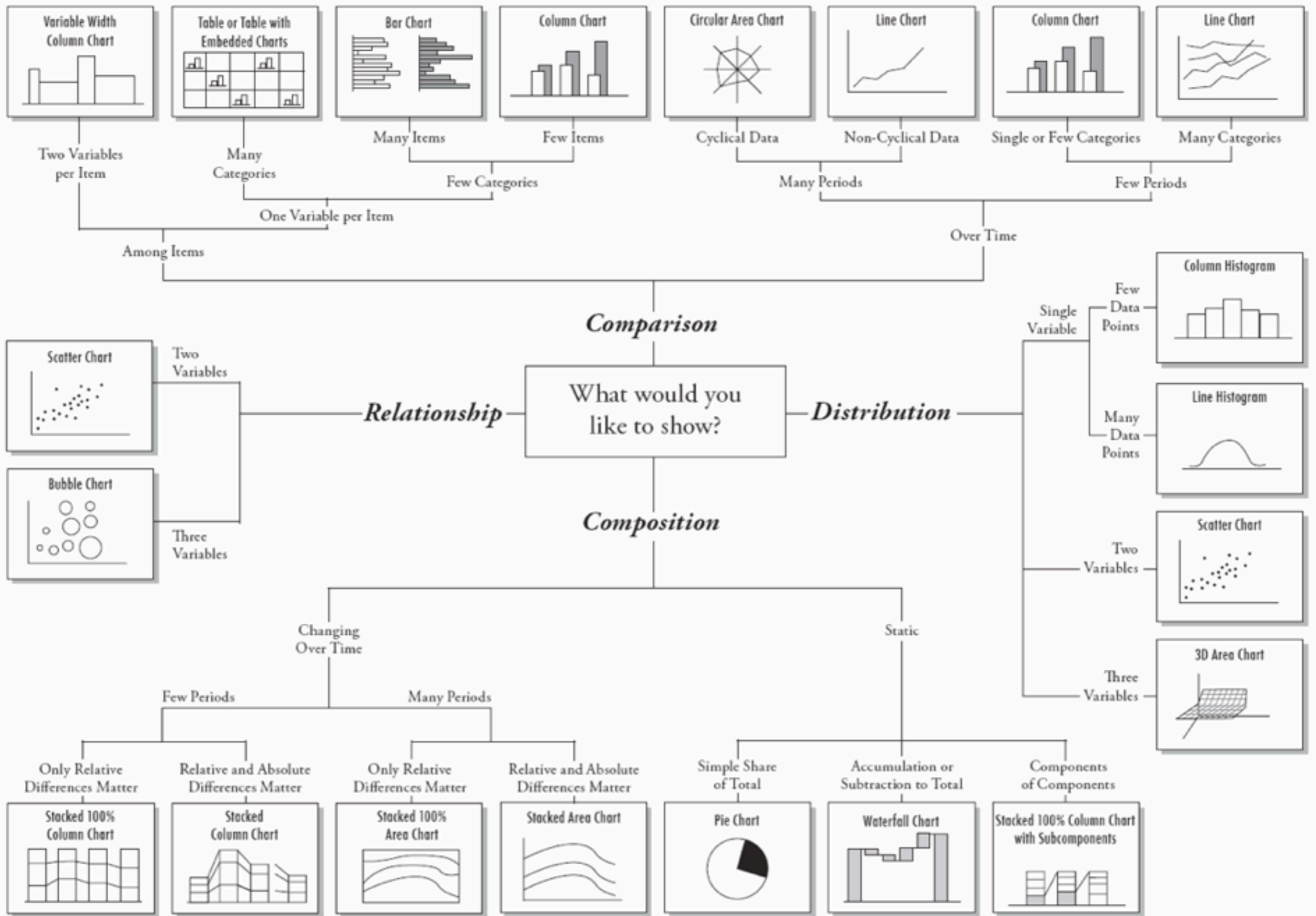
patterns and texture.

Chart by Andrew Abela

FOR MORE TIPS

HTTP://WWW.SLIDESHARE.NET/MCGARRAHJESSEE/MCJEDWARD-TUFTENOTES