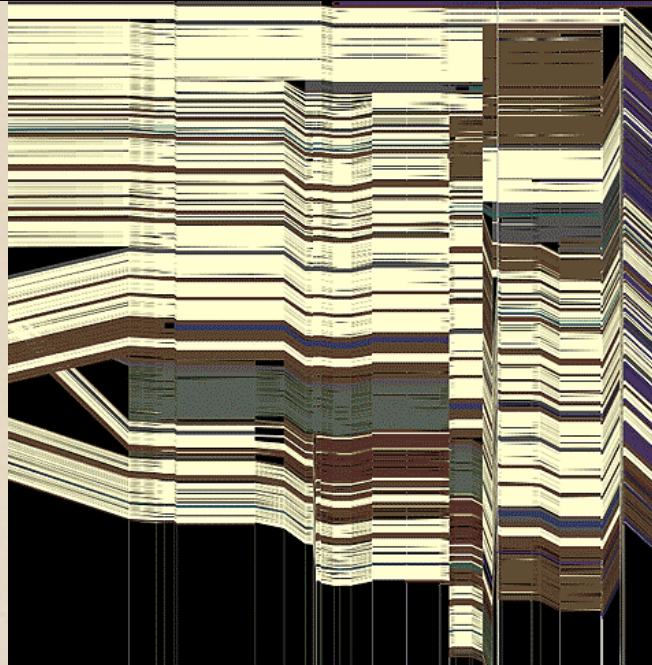
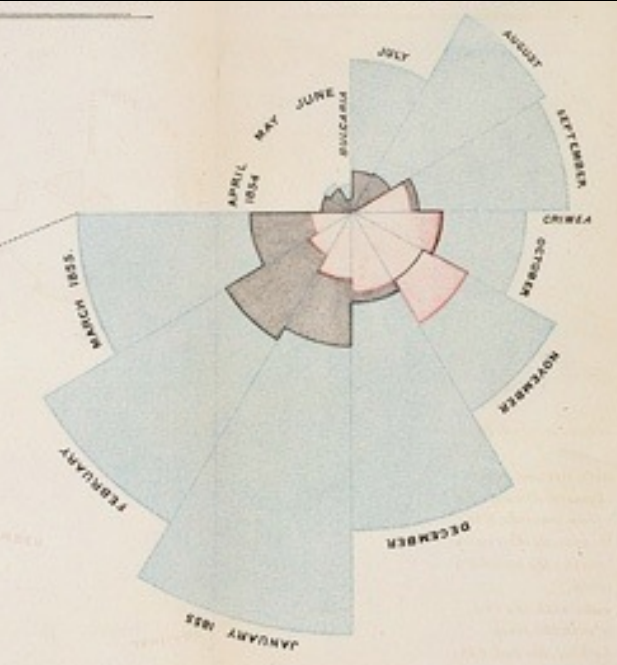


# CSE 412 - Intro to Data Visualization

# Image Models



Jane Hoffswell University of Washington

# Zoom Poll: Tools & Expertise

# Image Models



# Visual Language is a Sign System



Jacques Bertin

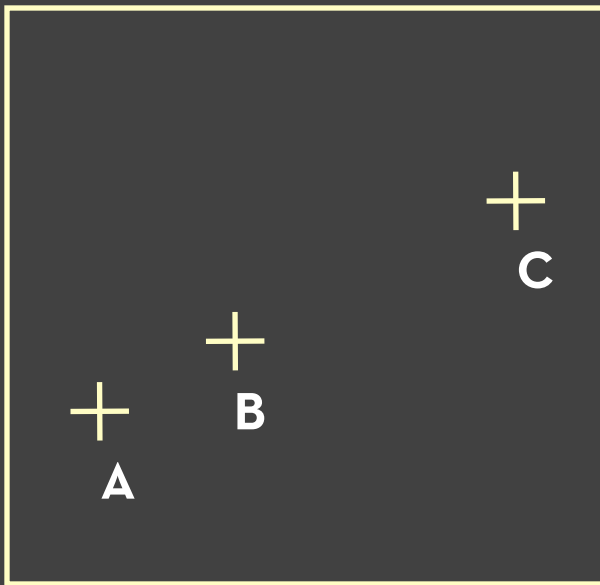
Images perceived as a set of signs

Sender encodes information in signs

Receiver decodes information from signs

Sémiologie Graphique, 1967

# Bertin's Semiology of Graphics



1. A, B, C are distinguishable
2. B is between A and C.
3. BC is twice as long as AB.

∴ Encode quantitative variables

*"Resemblance, order and proportional are the three signfields in graphics."* - Bertin



# LES VARIABLES DE L'IMAGE

	POINTS			LIGNES			ZONES	
XY 2 DIMENSIONS DU PLAN								
Z TAILLE								
VALEUR								

# LES VARIABLES DE SÉPARATION DES IMAGES

GRAIN								
COULEUR								
ORIENTATION								
FORME								

# Visual Encoding Variables

Position (x 2)

Size

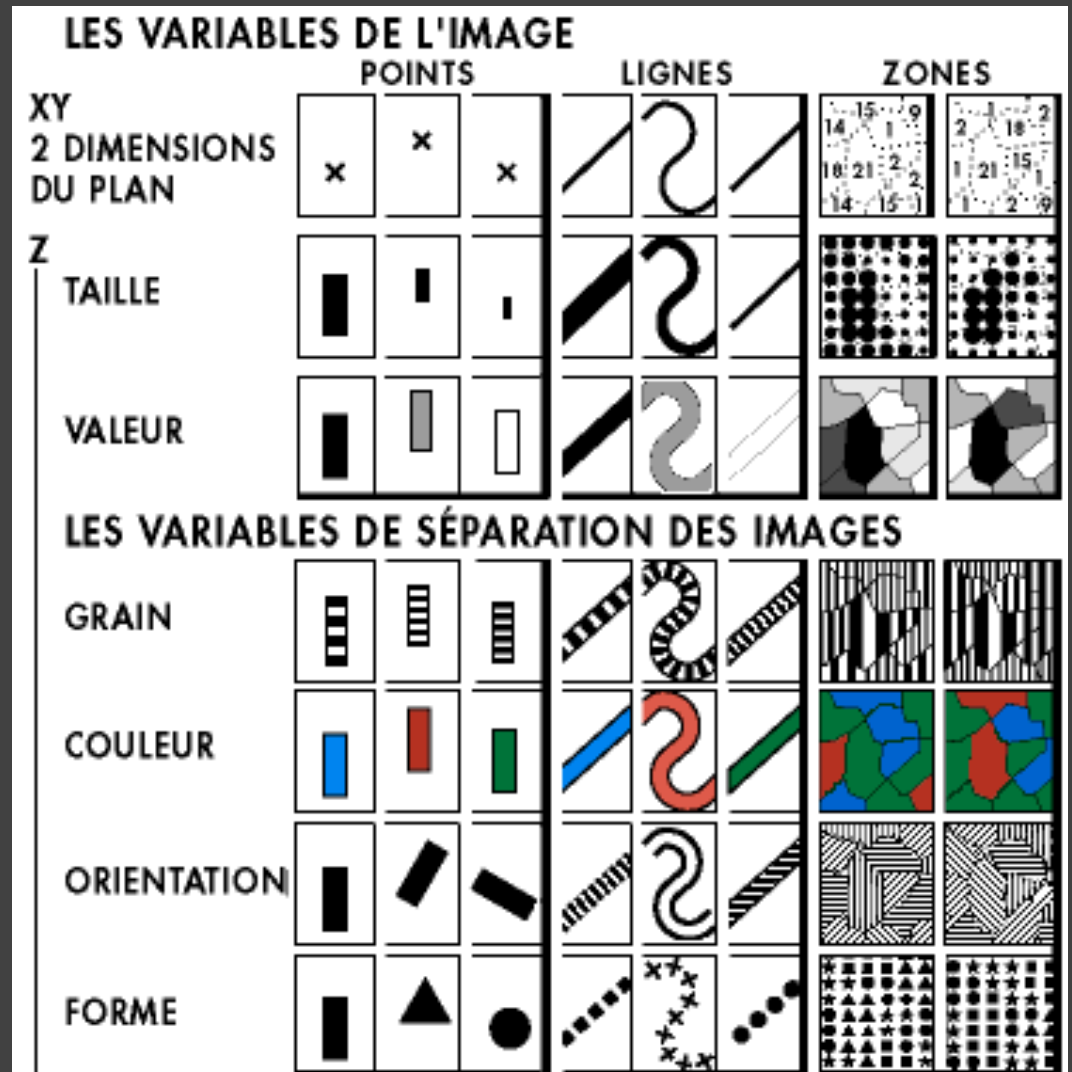
Value

Texture

Color

Orientation

Shape





# Visual Encoding Variables

Position

**Length**

**Area**

**Volume**

Value

Texture

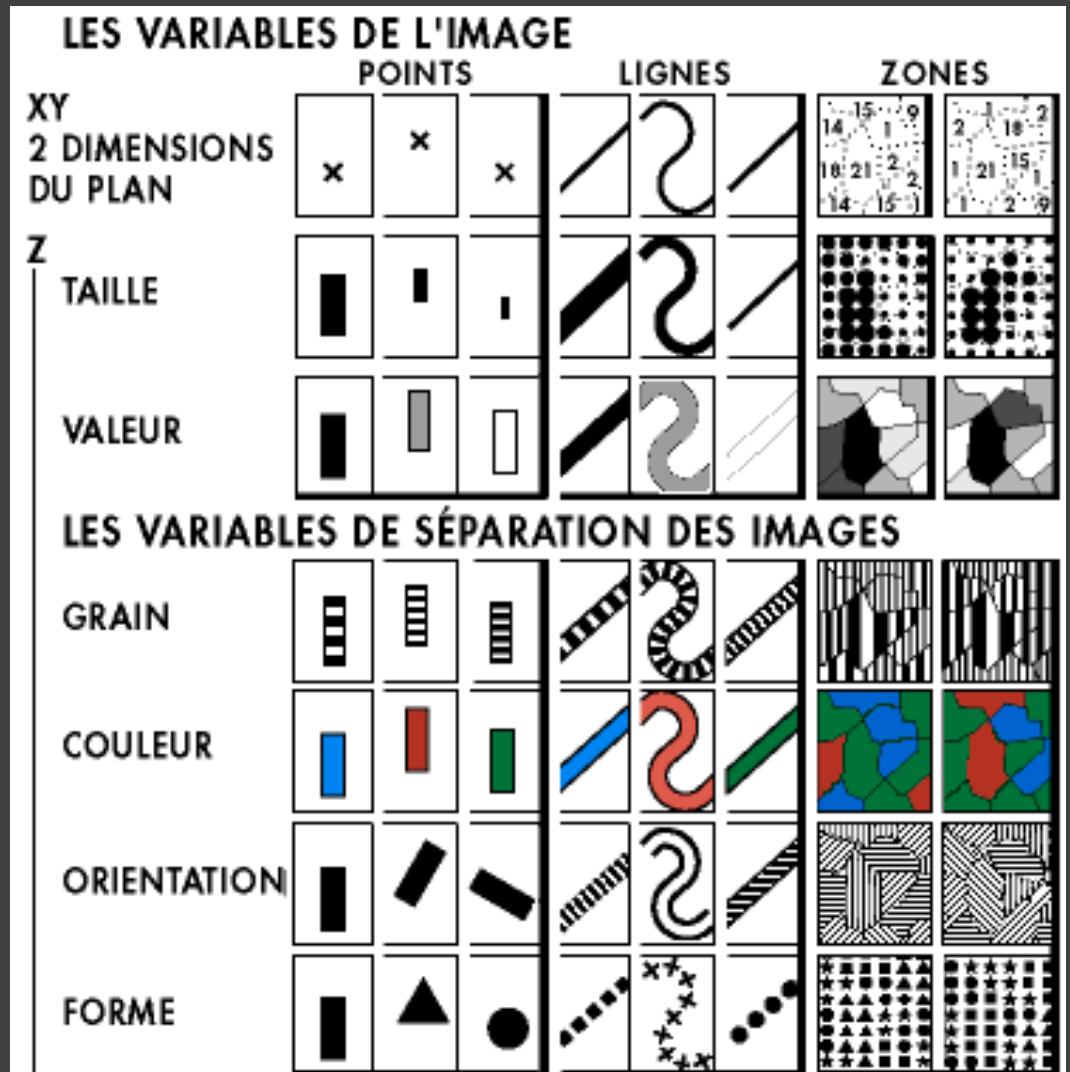
Color

Orientation

Shape

**Transparency**

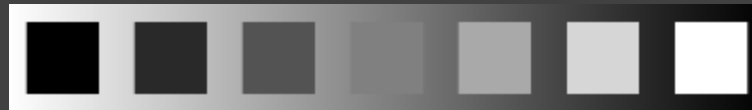
**Blur / Focus ...**



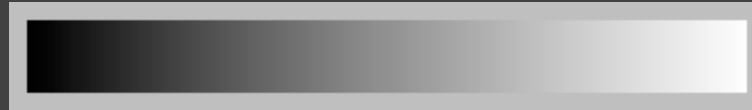
# Information in Hue and Value

Value is perceived as ordered

∴ Encode ordinal variables (O)



∴ Encode continuous variables (Q) [not as well]



Hue is normally perceived as unordered

∴ Encode nominal variables (N) using color



# Bertin's Levels of Organization

Position

N	O	Q
---	---	---

Nominal

Size

N	O	Q
---	---	---

Ordinal

Value

N	O	Q
---	---	---

Quantitative

Note: **Q**  $\subset$  **O**  $\subset$  **N**

Texture

N	o	
---	---	--

Color

N		
---	--	--

Orientation

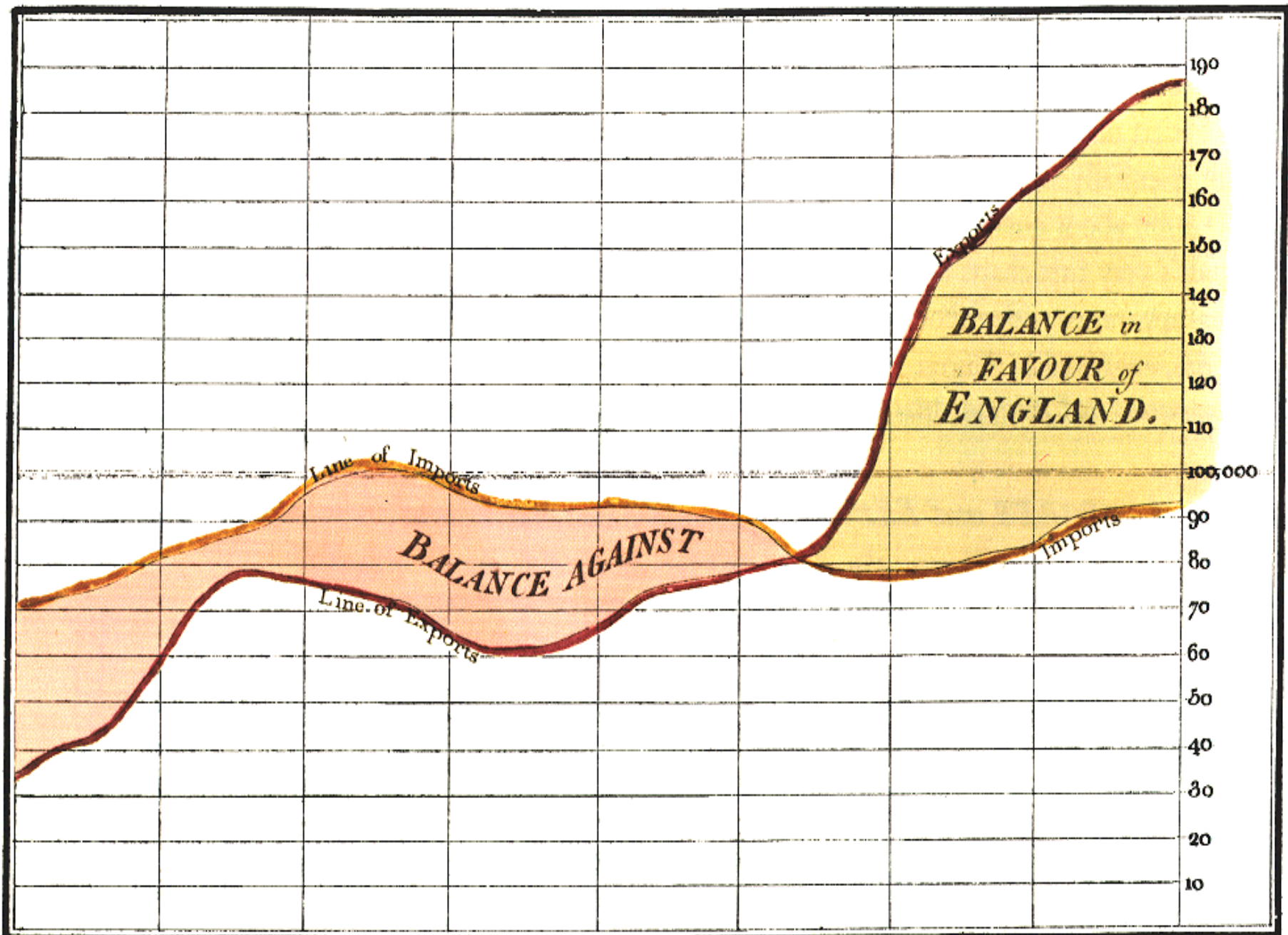
N		
---	--	--

Shape

N		
---	--	--

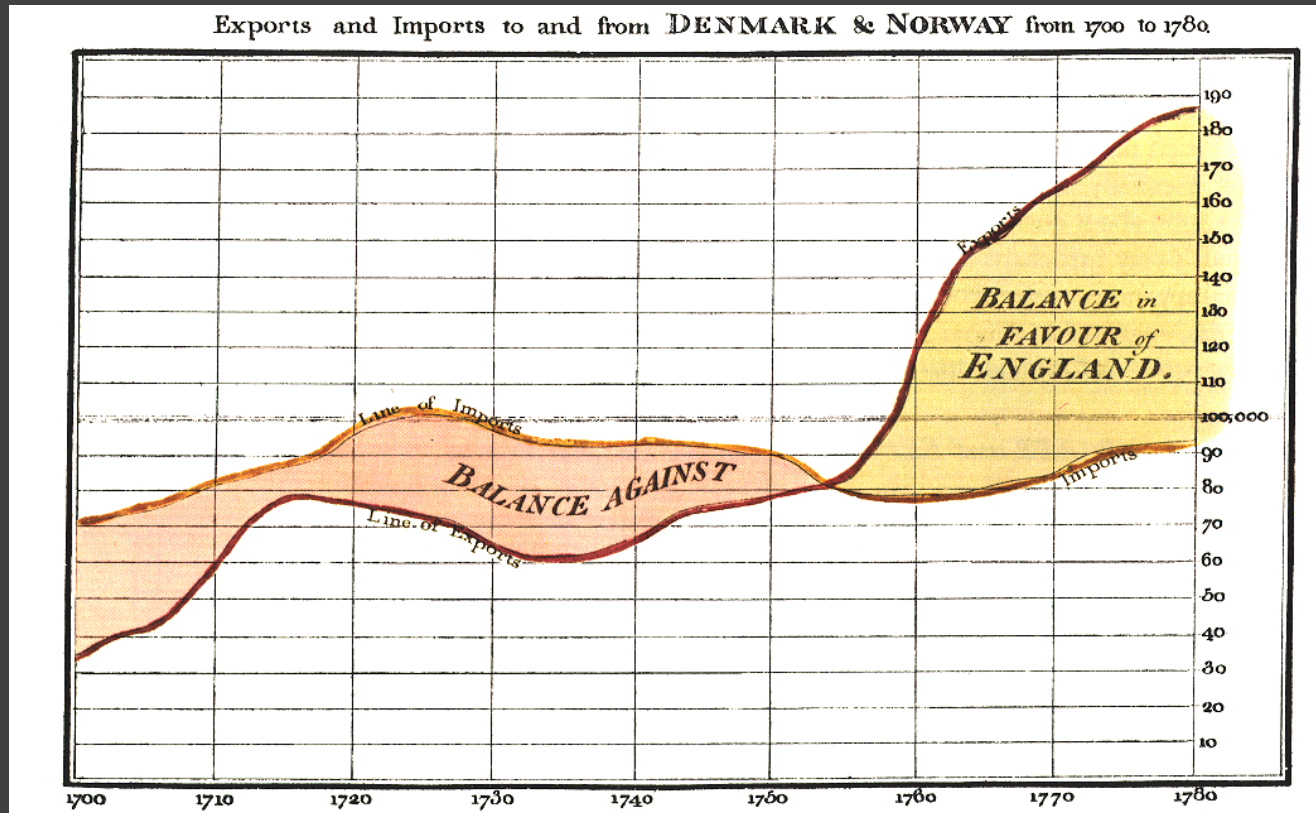
# Deconstructions

Exports and Imports to and from DENMARK & NORWAY from 1700 to 1780.



1700 1710 1720 1730 1740 1750 1760 1770 1780

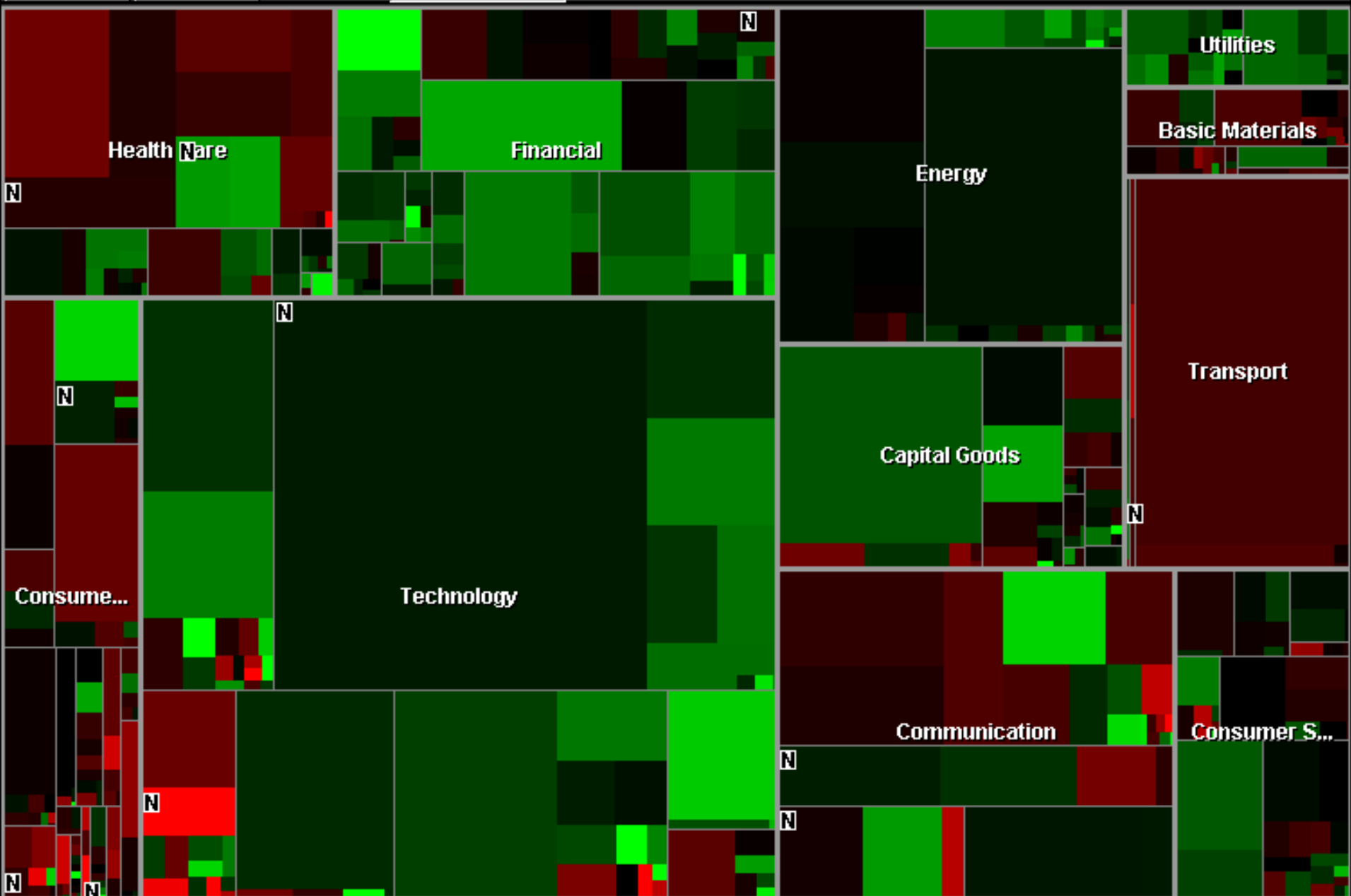
# William Playfair, 1786



X-axis: year (Q)

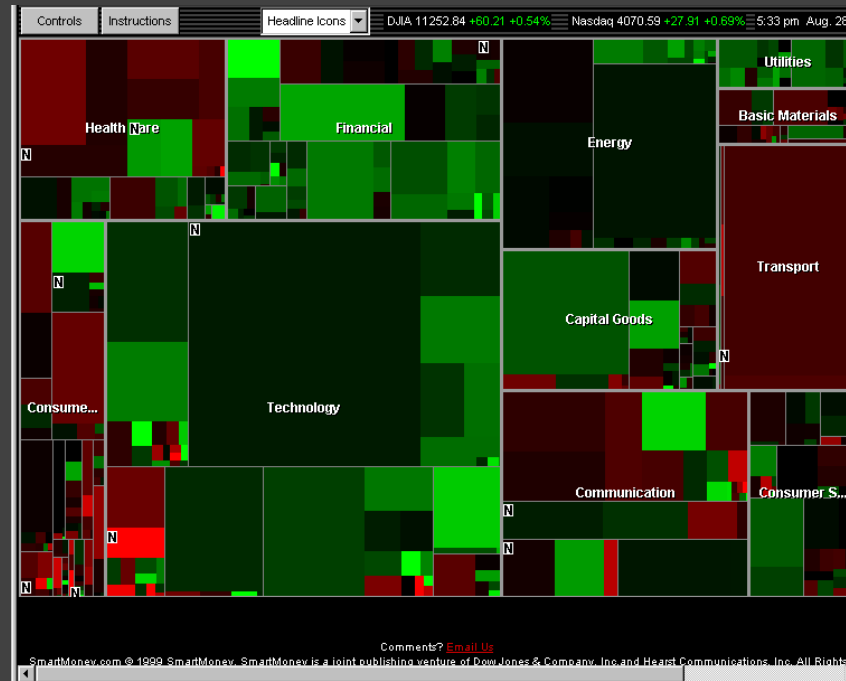
Y-axis: currency (Q)

Color: imports/exports (N, O)





# Wattenberg's Map of the Market



Rectangle Area: market cap (Q)

Rectangle Position: market sector (N), market cap (Q)

Color Hue: loss vs. gain (N, O)

Color Value: magnitude of loss or gain (Q)

# Minard 1869: Napoleon's March

## Carte Figurative des pertes successives en hommes de l'Armée Française dans la campagne de Russie 1812-1813.

Dressée par M. Minard, Inspecteur Général des Ponts et Chaussées en retraite Paris, le 20 Novembre 1869.

Les nombres d'hommes présents sont représentés par les largeurs des zones colorées à raison d'un millimètre pour dix mille hommes; ils sont de plus écrits en travers des zones. Le rouge désigne les hommes qui entrent en Russie, le noir ceux qui en sortent. Les renseignements qui ont servi à dresser la carte ont été puisés dans les ouvrages de M. M. Chiers, de Légar, de Fezensac, de Chambray et le journal inédit de Jacob, pharmacien de l'Armée depuis le 28 Octobre.

Pour mieux faire juger à l'œil la diminution de l'armée, j'ai supposé que les corps du Prince Jérôme et du Maréchal Davout, qui avaient été détachés sur Minsk et Mohilow et qui s'étaient joints vers Orscha et Witebsk, avaient toujours marché avec l'armée.

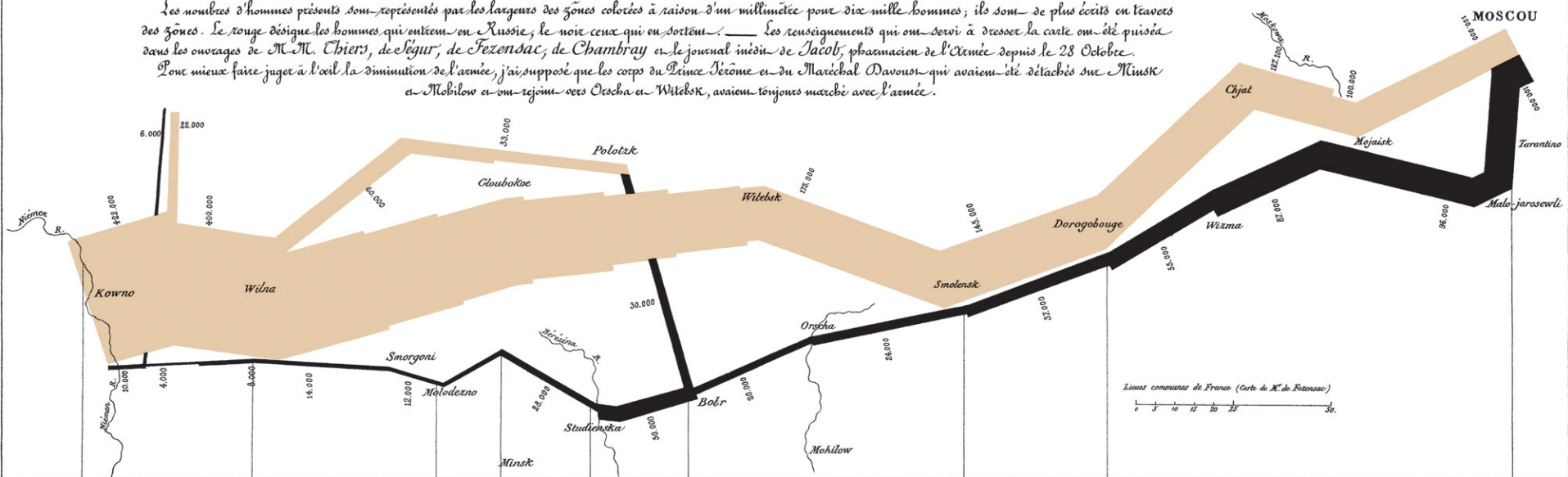
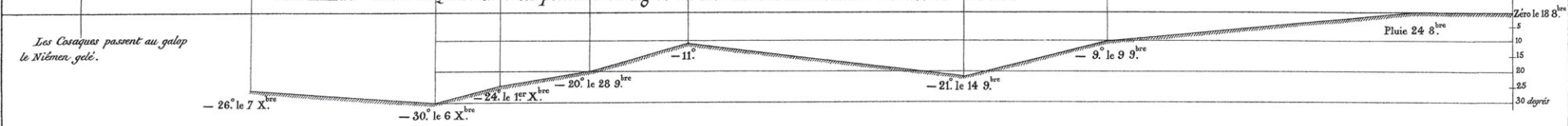
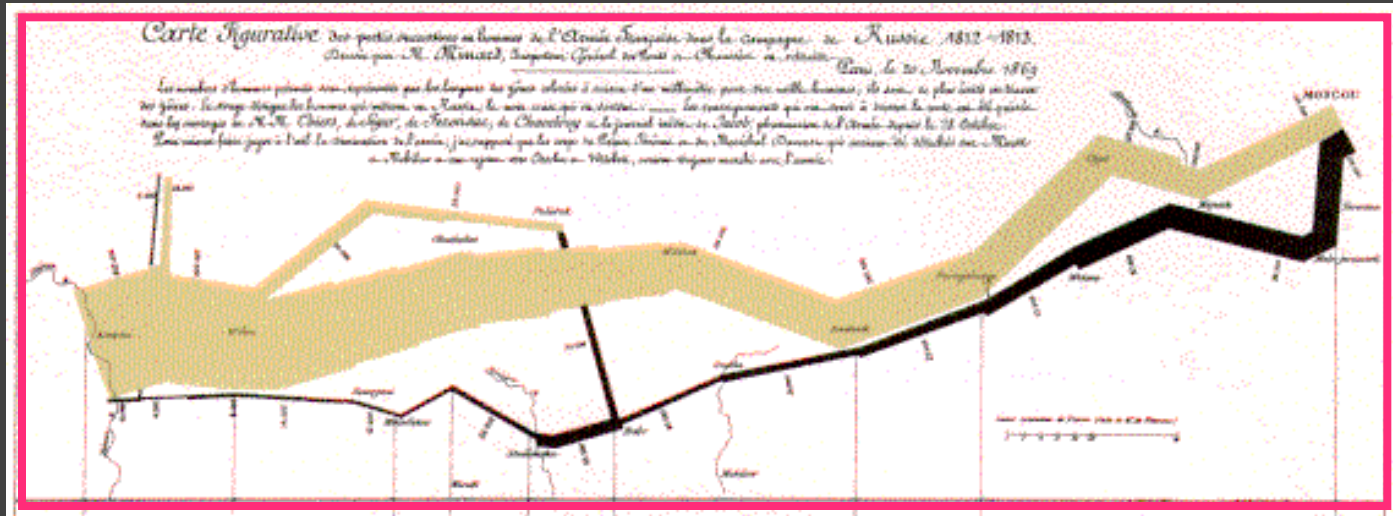


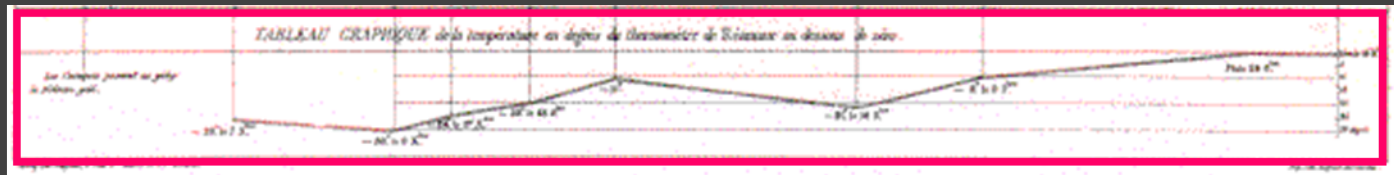
TABLEAU GRAPHIQUE de la température en degrés du thermomètre de Réaumur au dessous de zéro.



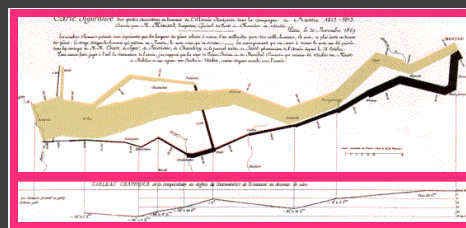
# Single-Axis Composition



+



=



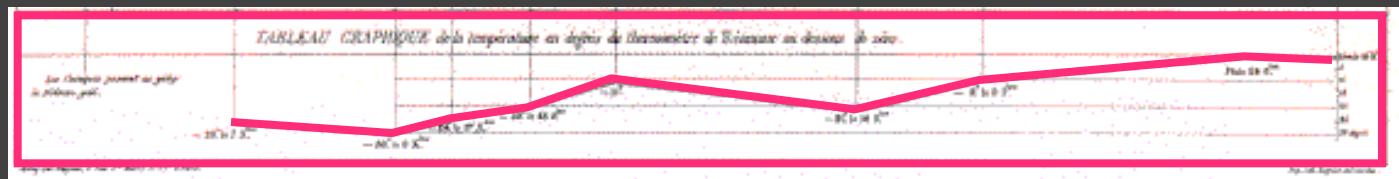
# Mark Composition

Y-axis: temperature (Q)

+

X-axis: longitude (Q) / time (O)

=



Temp over space/time (Q x Q)

# Mark Composition

Y-axis: latitude (Q)

+

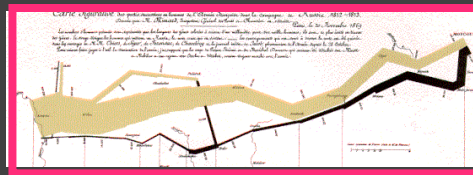
X-axis: longitude (Q)

+

Width: army size (Q)



=

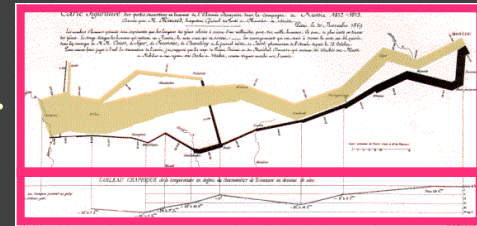
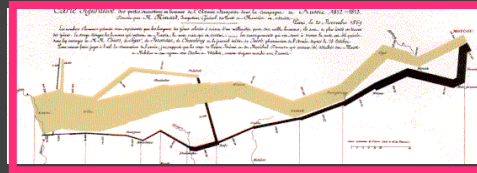


Army position (Q x Q) and army size (Q)

latitude (Q)

longitude (Q)

army size (Q)



temperature (Q)

longitude (Q) / time (O)



# Minard 1869: Napoleon's March

## Carte Figurative des pertes successives en hommes de l'Armée Française dans la campagne de Russie 1812-1813.

Dressée par M. Minard, Inspecteur Général des Ponts et Chaussées en retraite Paris, le 20 Novembre 1869.

Les nombres d'hommes présents sont représentés par les largeurs des zones colorées à raison d'un millimètre pour dix mille hommes; ils sont de plus écrits en travers des zones. Le rouge désigne les hommes qui entrent en Russie, le noir ceux qui en sortent. Les renseignements qui ont servi à dresser la carte ont été puisés dans les ouvrages de M. M. Chiers, de Légar, de Fezensac, de Chambrey et le journal inédit de Jacob, pharmacien de l'Armée depuis le 28 Octobre.

Pour mieux faire juger à l'œil la diminution de l'armée, j'ai supposé que les corps du Prince Jérôme et du Maréchal Davout qui avaient été détachés sur Minsk et Mohilow et qui rejoignent vers Orscha et Witebsk, avaient toujours marché avec l'armée.

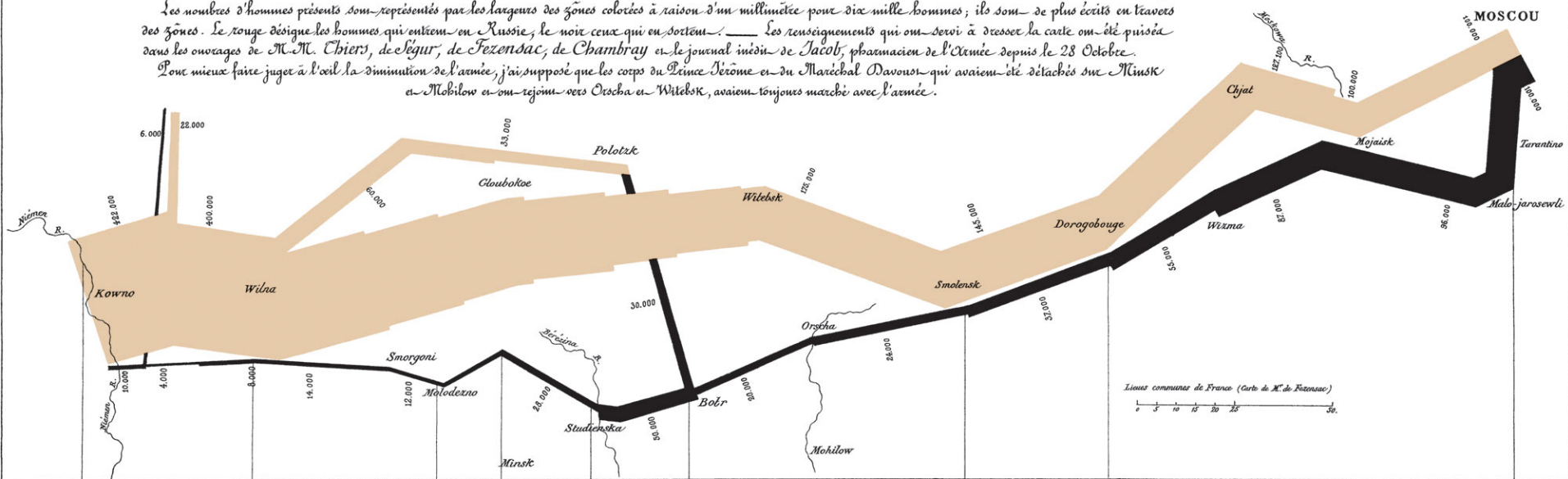
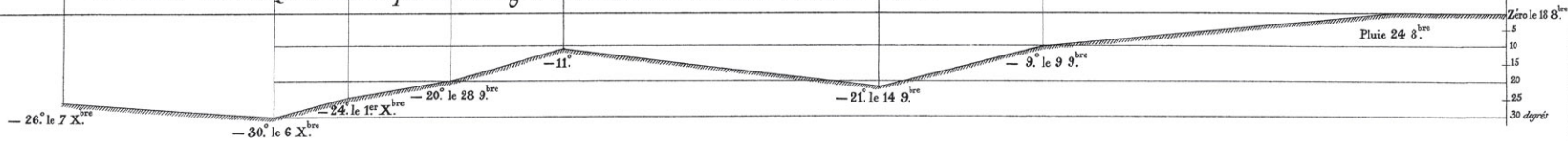


TABLEAU GRAPHIQUE de la température en degrés du thermomètre de Réaumur au dessous de zéro.



Les Cosaques passent au galop le Niémen gelé.

Depicts at least 5 quantitative variables. Any others?



# Formalizing Design

# Choosing Visual Encodings

Assume  $k$  visual encodings and  $n$  data attributes. We would like to pick the “best” encoding among a combinatorial set of possibilities of size  $(n+1)^k$

## Principle of Consistency

The properties of the image (visual variables) should match the properties of the data.

## Principle of Importance Ordering

Encode the most important information in the most effective way.

# Design Criteria [Mackinlay 86]

## Expressiveness

*A set of facts is expressible in a visual language if the sentences (i.e. the visualizations) in the language express all the facts in the set of data, and only the facts in the data.*

## Effectiveness

*A visualization is more effective than another visualization if the information conveyed by one visualization is more readily perceived than the information in the other visualization.*

# Design Criteria [Mackinlay 86]

## Expressiveness

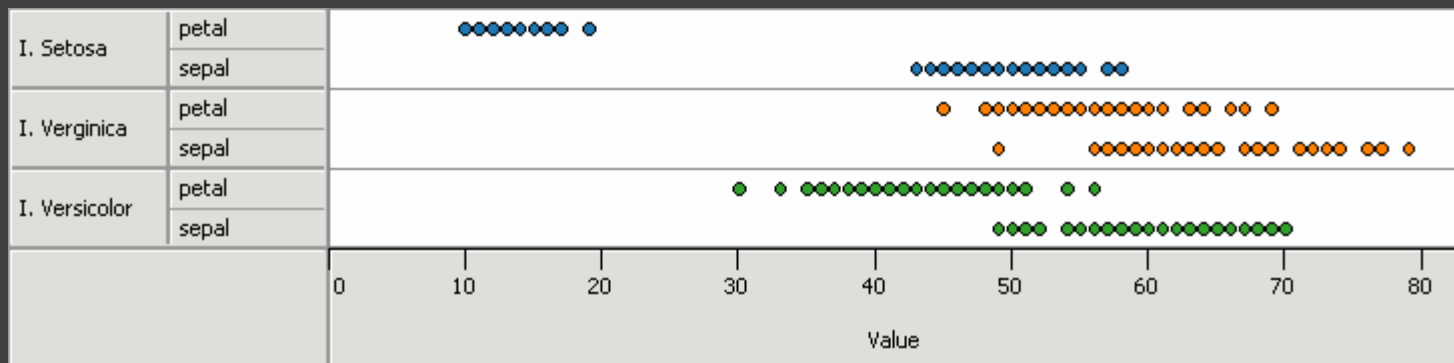
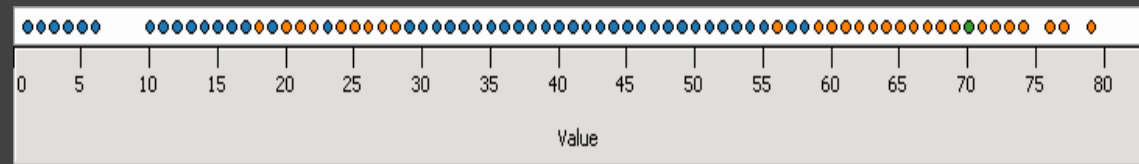
A set of facts is *expressible* in a visual language if the sentences (i.e. the visualizations) in the language express all the facts in the set of data, and only the facts in the data.

## Effectiveness

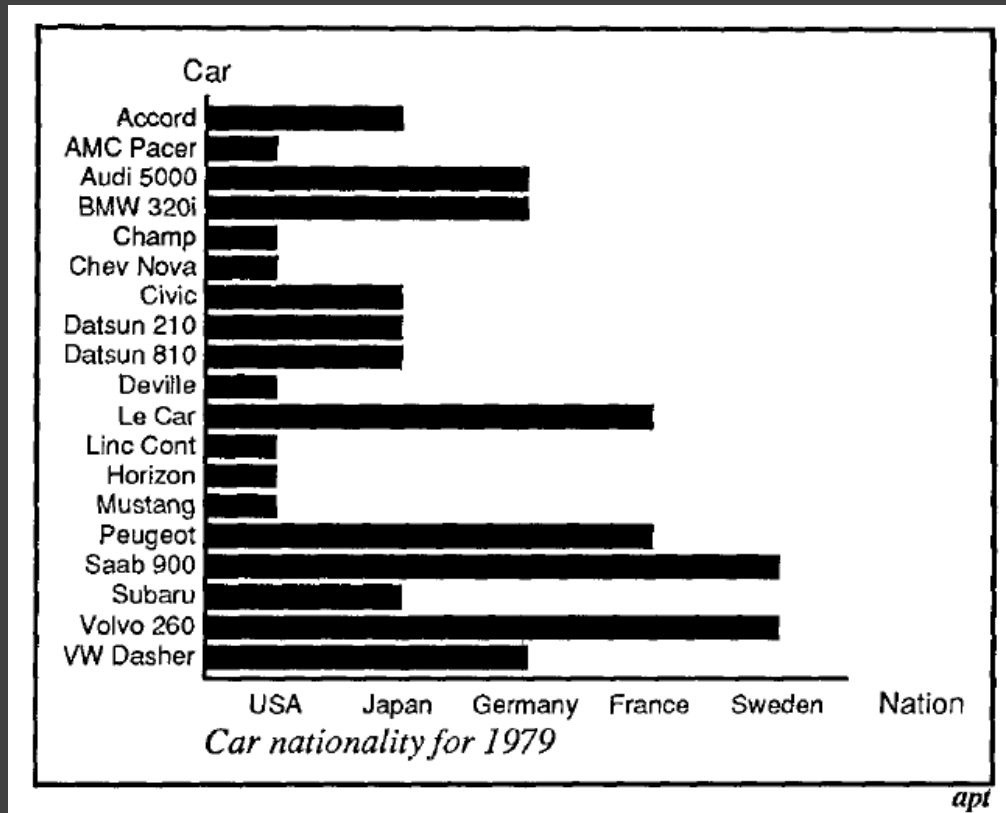
Visualization A is more effective than another visualization if the information conveyed by one is perceived more effectively than the other in the other visualization.

# Can not express the facts

A multivariate relation may be *inexpressive* in a single horizontal dot plot because multiple records are mapped to the same position.



# Expresses facts not in the data



A length is interpreted as a quantitative value.

# Design Criteria [Mackinlay 86]

## Expressiveness

A set of facts is *expressible* in a visual language if the sentences (i.e. the visualizations) in the language express all the facts in the set of data, and only the facts in the data.

## Effectiveness

Visualization A is more effective than another visualization if the information conveyed by one is perceived more effectively than the other in the other visualization.



# Design Criteria [Mackinlay 86]

## Expressiveness

A set of facts is *expressible* in a visual language if the sentences (i.e. the visualizations) in the language express all the facts in the set of data, and only the facts in the data.

## Effectiveness

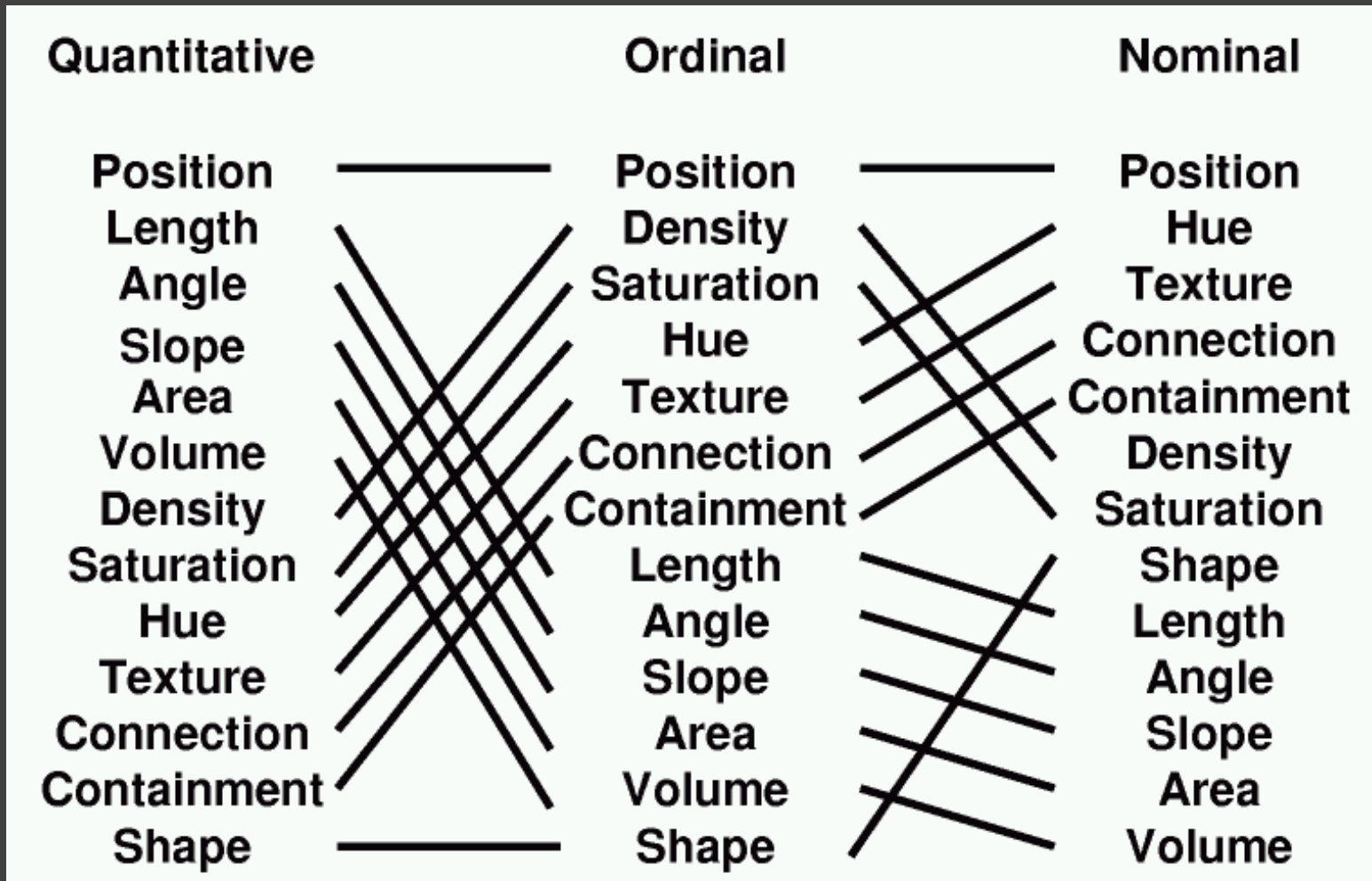
A visualization is more *effective* than another visualization if the information conveyed by one visualization is more readily perceived than the information in the other visualization.

# Design Criteria *Translated*

**Tell the truth and nothing but the truth**  
(don't lie, and don't lie by omission)

**Use encodings that people decode better**  
(where better = faster and/or more accurate)

# Mackinlay's Ranking



Conjectured *effectiveness* of encodings by data type

# Mackinlay's Design Algorithm

**APT** - "A Presentation Tool", 1986

**User formally specifies data model and type**

Input: ordered list of data variables to show

**APT searches over design space**

Test expressiveness of each visual encoding

Generate encodings that pass test

Rank by perceptual effectiveness criteria

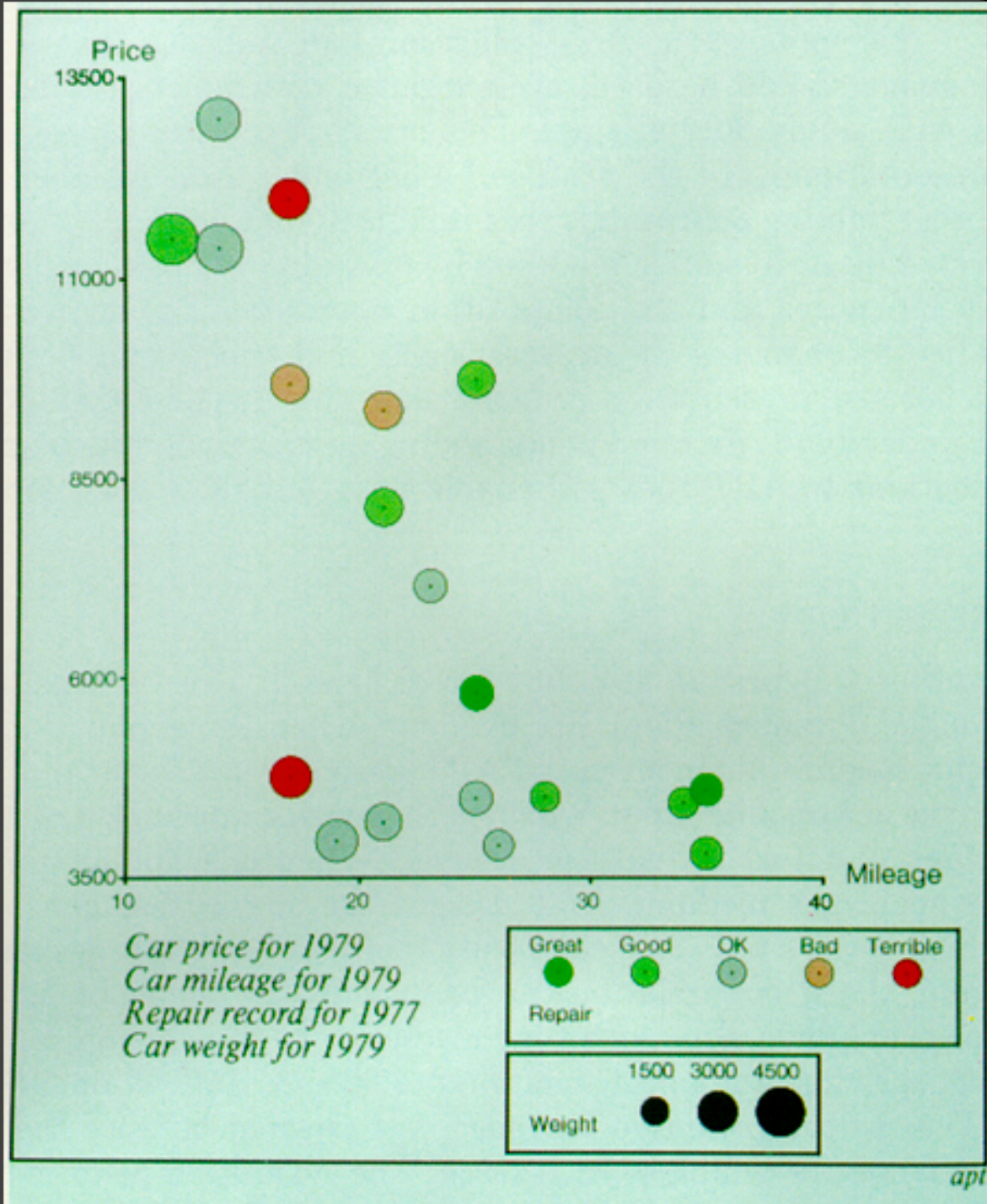
**Output the "most effective" visualization**

# APT

Automatically  
generate chart  
for car data

Input variables:

1. Price
2. Mileage
3. Repair
4. Weight



# Limitations of APT?

# Limitations of APT

**Does not cover many visualization techniques**

Networks, hierarchies, maps, diagrams

Also: 3D structure, animation, illustration, ...

**Does not consider interaction**

**Does not consider semantics / conventions**

**Assumes single visualization as output**

Recent related work:

**Draco visualization design knowledge base**

# Administrivia



# A1: Visualization Design

Pick a **guiding question**, use it to title your vis.  
Design a **static visualization** for that question.  
You are free to **use any tools** (inc. pen & paper).

**Deliverables** (upload via Canvas; see A1 page)

Image of your visualization (PNG or JPG format)

Short description + design rationale ( $\leq 4$  paragraphs)

Due by **11:59 pm PT, Monday April 5th.**

# Course Participation

Quiz & discussion comments on class forum (Ed).

Both are due each Monday, by 11:59pm PT up through week 8 of the quarter.

You have 1 “pass” (quiz + comment) for the quarter.

First discussion and quiz are now posted on Ed,  
Due by **11:59 pm PT, Monday April 5th.**

# Required Readings for Mon 4/5

**Database Schema:**  
The user drags fields from the database schema to shelves to define the visual specification.

**Layer Tabs:**  
Each layer has its own tab; different transformations and mappings can be specified for each layer.

**Axis Shelves:**  
The fields placed here determine the structure of the table and the types of graphs in each table pane.

**Context Menu:**  
The context menu provides access to the data transformation and interaction capabilities of Polaris such as sorting, filtering, and aggregation.

**Layer Shelf:**  
The fields placed here determine how records are partitioned into layers.

**Grouping and Sorting Shelves:**  
The fields placed here determine how records are grouped and sorted within the table panes.

**Mark Pulldown:**  
Relations in each pane are mapped to marks of the selected type.

**Retinal Property Shelves:**  
The fields placed here determine how data is encoded in the retinal properties of the marks.

**Legends:**  
Legends enable the user to see and modify the mappings from data to retinal properties.

# Optional Readings for Week 2

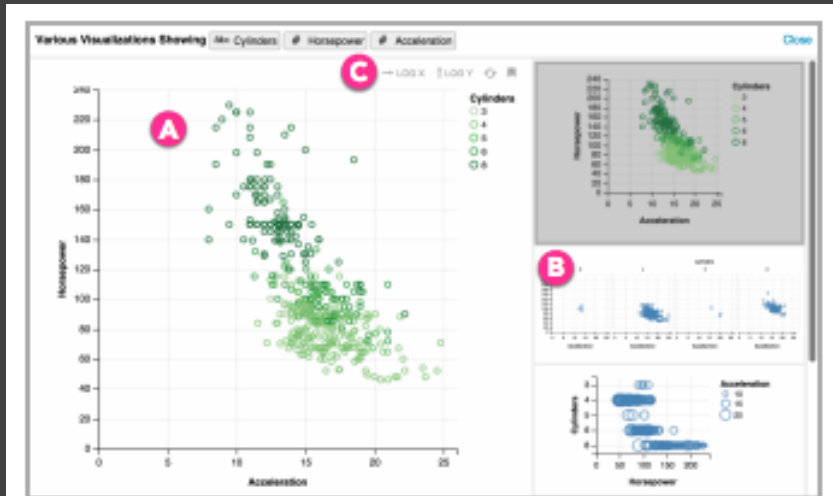


Fig. 4. The expanded gallery for cylinder, horsepower, and acceleration. (a) The main panel presents the selected chart in an enlarged view. (b) The sidebar shows alternative encodings for the expanded data.

Voyager: Exploratory Analysis via Faceted Browsing of Visualization Recommendations. Kanit Wongsuphasawat et al. IEEE TVCG. 2016.

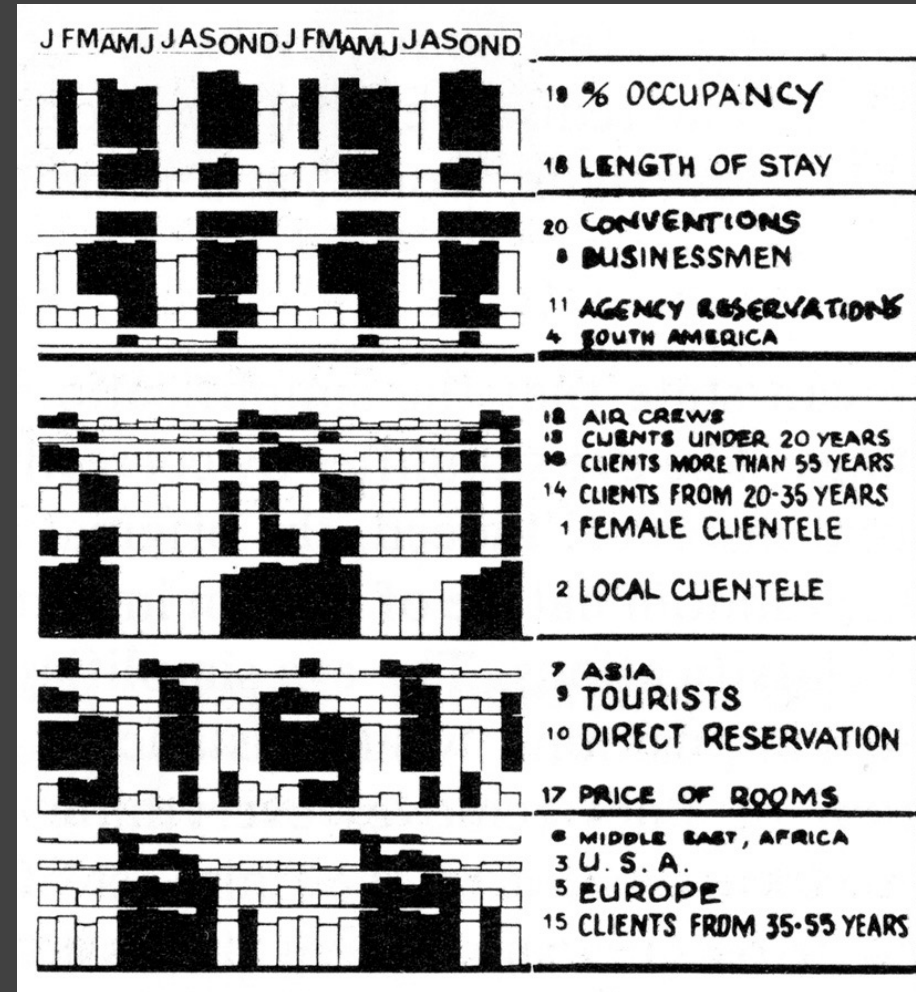
ENGINEERING STATISTICS HANDBOOK

HOME
TOOLS & AIDS
SEARCH
BACK
NEXT

## 1. Exploratory Data Analysis

This chapter presents the assumptions, principles, and techniques necessary to gain insight into data via EDA--exploratory data analysis.

Exploratory Data Analysis, NIST Engineering Statistics Handbook.



Postmortem of an Example. Jacques Bertin. Graphics and Graphic Information-Processing. 1981.

# Design Exercise

# Visual Encoding Exercise

5 17

How many visualizations can you think of for conveying these two numbers? Feel free to invent tasks or contexts. **Sketch as many as you can!**

*Don't stress over quality, go for quantity.*

Time: ~5 minutes

# Visual Encoding Exercise

5 17

We will assign you to breakout rooms. Introduce yourselves! Then compare your designs. (*You can hold drawings up to the camera to share.*) How many ideas are the same? How many are different?

Capture your favorite images and post them on the Ed thread "In-Class Design Activity".

# Visual Encoding Exercise

5 17

How many visualizations can you think of for conveying these two numbers? Feel free to invent tasks or contexts. **Sketch as many as you can!**

Time permitting, let's share back with the class.

What were the most common designs?

The most surprising / creative / innovative?



# Summary: Data & Image Models

## Formal specification

Data model: relational data; N,O,Q types

Image model: visual encoding channels

Encodings map data to visual variables

## Choose expressive and effective encodings

Rule-based tests of expressiveness

Perceptual effectiveness rankings

**Question:** how do we establish effectiveness criteria? *Subject of perception lectures...*