

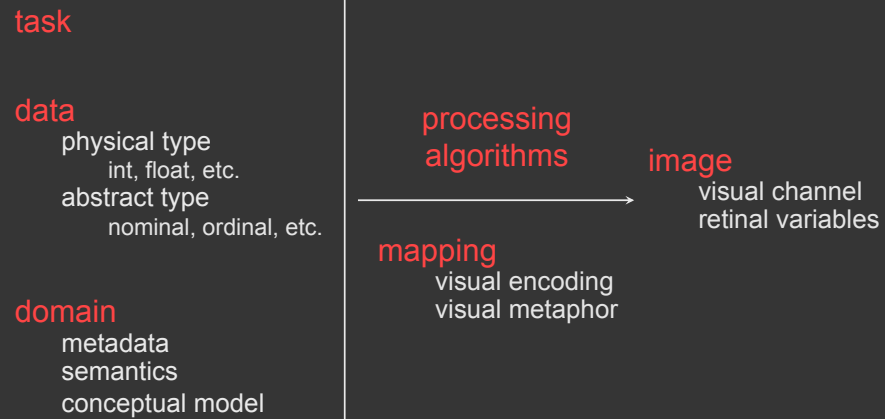
Visualization Designs

Maneesh Agrawala

CS 294-10: Visualization
Fall 2013

**Last Time: Data and Image
Models**

The big picture



[based on slide from Munzner]

Nominal, ordinal and quantitative

N - Nominal (labels)

- Operations: =, \neq

O - Ordered

- Operations: =, \neq , $<$, $>$, \leq , \geq

Q - Interval (Location of zero arbitrary)

- Operations: =, \neq , $<$, $>$, \leq , \geq , -
- Can measure distances or spans

Q - Ratio (zero fixed)

- Operations: =, \neq , $<$, $>$, \leq , \geq , -, \div
- Can measure ratios or proportions

S. S. Stevens, On the theory of scales of measurements, 1946

Visual variables

- Position (x 2)
- Size
- Value
- Texture
- Color
- Orientation
- Shape

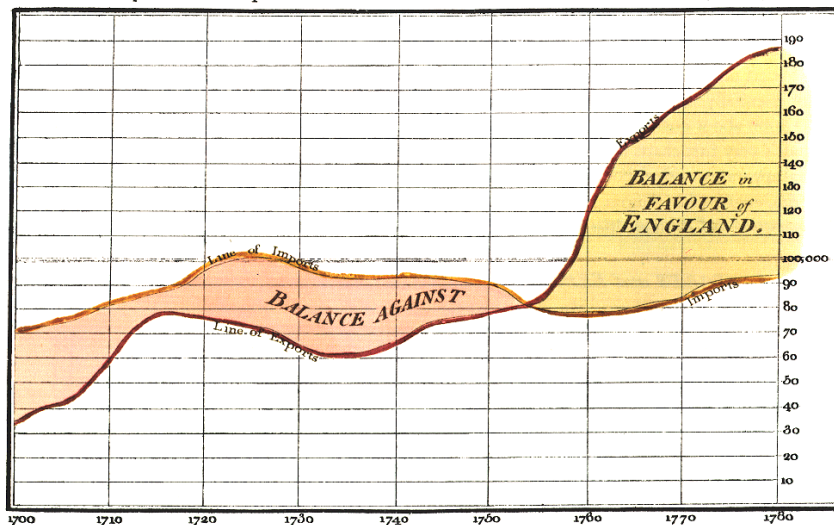
LES VARIABLES DE L'IMAGE					
		POINTS	LIGNES	ZONES	
XY 2 DIMENSIONS DU PLAN		x x x	/ \	16 15 14 13 12 11 10 9 8 7 6 5 4 3 2 1	2 1 18 17 21 20 19 22 23 24 25 26
Z TAILLE		■ ■ ■	/ \	■ ■ ■	■ ■ ■
VALEUR		■ ■ ■	/ \	■ ■ ■	■ ■ ■
LES VARIABLES DE SÉPARATION DES IMAGES					
GRAIN		■ ■ ■	/ \	■ ■ ■	■ ■ ■
COULEUR		■ ■ ■	/ \	■ ■ ■	■ ■ ■
ORIENTATION		■ ■ ■	/ \	■ ■ ■	■ ■ ■
FORME		■ ■ ■	/ \	■ ■ ■	■ ■ ■

Note: Bertin does not consider 3D or time

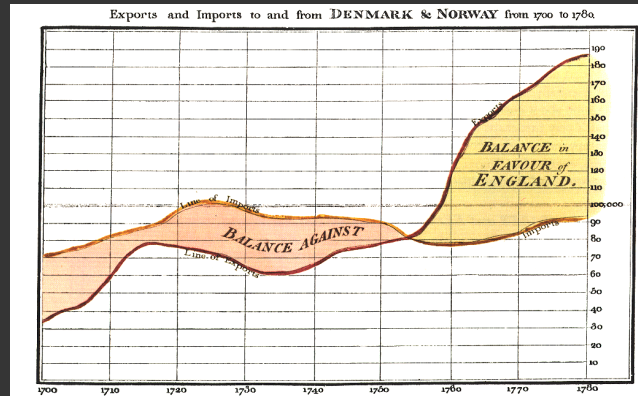
Note: Card and Mackinlay extend the number of vars.

Playfair 1786

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

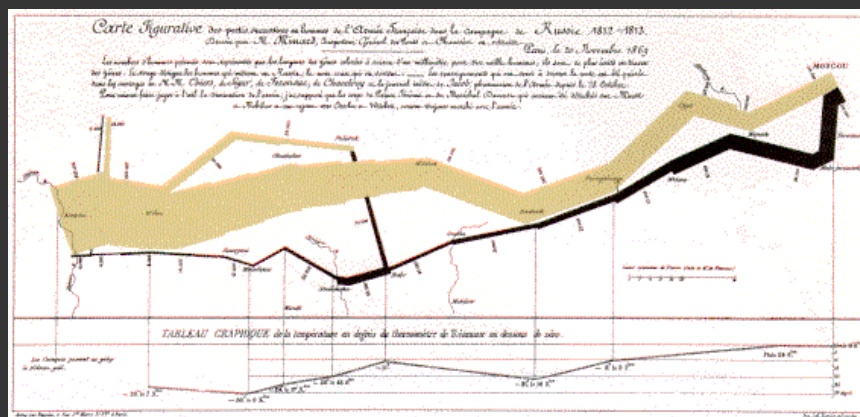


Playfair 1786

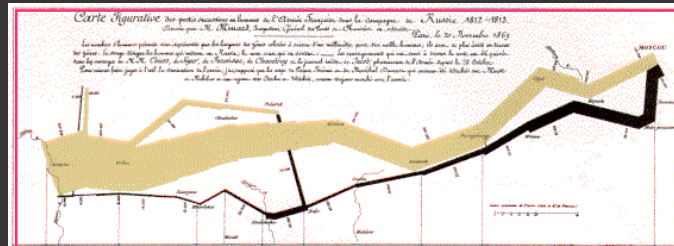


- x-axis: year (Q)
- y-axis: currency (Q)
- color: imports/exports (N, O)

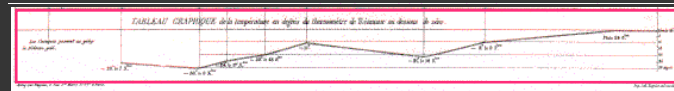
Minard 1869: Napoleon's march



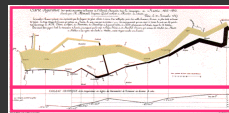
Single axis composition



+



=



[based on slide from Mackinlay]

Mark composition

y-axis: temperature (Q)

+ x-axis: time (Q)

=



temp over time (Q x Q)

[based on slide from Mackinlay]

Mark composition

y-axis: longitude (Q)

+ x-axis: latitude (Q)

+ width: army size (Q)

=



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

[based on slide from Mackinlay]

longitude (Q)

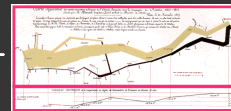
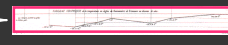
latitude (Q)

army size (Q)



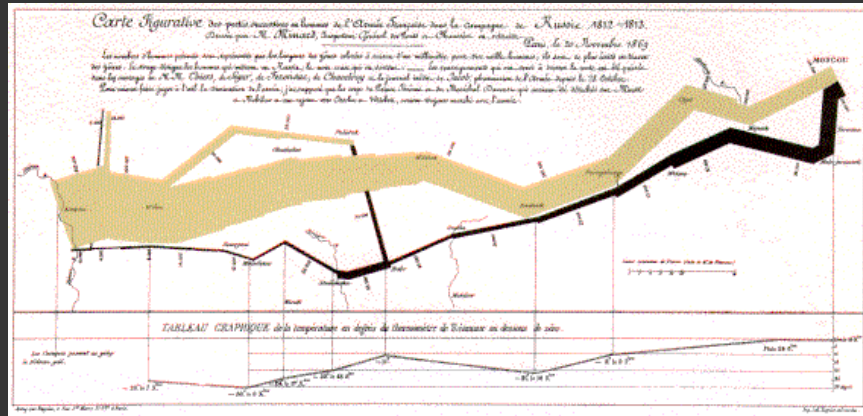
temperature (Q)

time (Q)



[based on slide from Mackinlay]

Minard 1869: Napoleon's march



Depicts at least 5 quantitative variables
Any others?

Automated design

Jock Mackinlay's APT 86



Combinatorics of encodings

Challenge:

Assume 8 visual encodings and n data attributes

Pick the best encoding from the exponential number of possibilities $(n+1)^8$

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

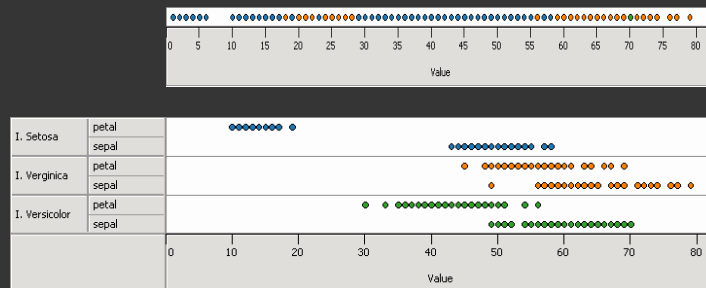
Mackinlay's expressiveness criteria

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.

Cannot express the facts

A one-to-many ($1 \rightarrow N$) relation cannot be expressed in a single horizontal dot plot because multiple tuples are mapped to the same position



Expresses facts not in the data

A length is interpreted as a quantitative value;
 \therefore Length of bar says something untrue about N data

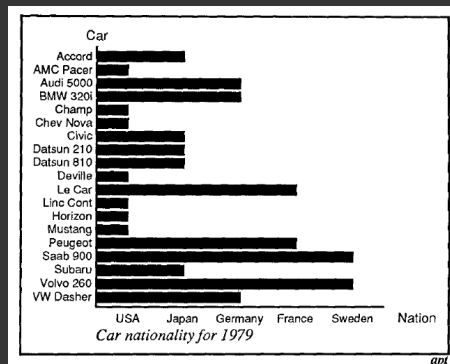


Fig. 11. Incorrect use of a bar chart for the *Nation* relation. The lengths of the bars suggest an ordering on the vertical axis, as if the USA cars were longer or better than the other cars, which is not true for the *Nation* relation.

[Mackinlay, APT, 1986]

Mackinlay's effectiveness criteria

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.

Subject of perception lecture

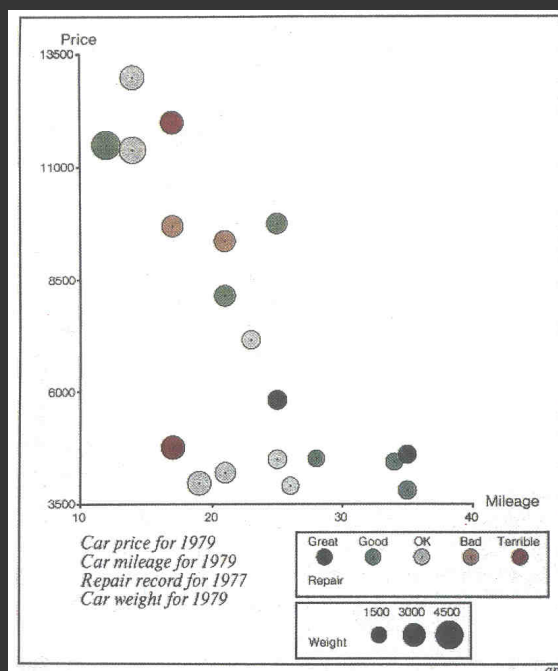
Mackinlay's ranking

Quantitative		Ordinal		Nominal
Position	————	Position	————	Position
Length		Density		Hue
Angle		Saturation		Texture
Slope		Hue		Connection
Area		Texture		Containment
Volume		Connection		Density
Density		Containment		Saturation
Saturation		Length		Shape
Hue		Angle		Length
Texture		Slope		Angle
Connection		Area		Slope
Containment		Volume		Area
Shape	————	Shape		Volume

Conjectured *effectiveness* of the encoding

Mackinlay's design algorithm

- User formally specifies data model and type
- APT searches over design space
 - Tests expressiveness of each visual encoding
 - Generates image for encodings that pass test
 - Tests perceptual effectiveness of resulting image
- Outputs most effective visualization



Limitations

Does not cover many visualization techniques

- Bertin and others discuss networks, maps, diagrams
- They do not consider 3D, animation, illustration, photography, ...

Does not model interaction

Summary

Formal specification

- Data model
- Image model
- Encodings mapping data to image

Choose expressive and effective encodings

- Formal test of expressiveness
- Experimental tests of perceptual effectiveness

Announcements

Announcements

Auditors, *please* enroll in the class (1 unit, P/NP)

- Requirements: Come to class and participate (online as well)
- Requirements: Assignment 1

Class participation requirements

- Complete readings before class
- In-class discussion
- Post at least 1 discussion substantive comment/question by 3pm on day of lecture

All, add yourself to participants page on the wiki

Class wiki

<http://vis.berkeley.edu/courses/cs294-10-fa13/wiki/>

Assignment 2: Exploratory Data Analysis

Use existing software to formulate & answer questions

First steps

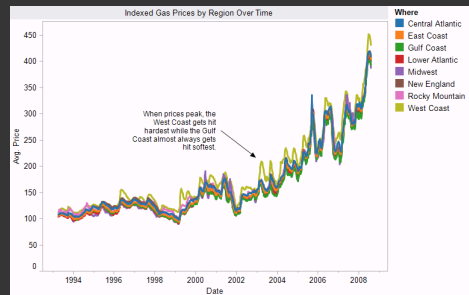
- Step 1: Pick a domain
- Step 2: Pose questions
- Step 3: Find data
- Iterate

Create visualizations

- Interact with data
- Question will evolve
- Tableau

Make wiki notebook

- Keep record of all steps you took to answer the questions



Due before class on Sep 30, 2013

Announcements

9/18: Investigative journalist Nate Halverson will describe a data set on real-estate buyers.

Hopes some of you will analyze it for Assignment 2

Assignment 1: Visualization Design



TEXAS DEPARTMENT OF CRIMINAL JUSTICE

Home Public Resources Employment About TDCJ Online Services Search

[HOME](#) | [DEATH ROW](#) | Death Row Information

Executed Offenders

Execution	Link	Link	Last Name	First Name	TDCJ Number	Age	Date	Race	County
503	Offender Information	Last Statement	Feldman	Douglas	999326	55	07/31/2013	White	Dallas
502	Offender Information	Last Statement	Ross	Vaughn	999429	41	07/18/2013	Black	Lubbock
501	Offender Information	Last Statement	Quintanilla	John	999491	36	07/16/2013	Hispanic	Victoria
500	Offender Information	Last Statement	McCarthy	Kimberly	999287	52	06/26/2013	Black	Dallas
499	Offender Information	Last Statement	Chester	Elroy	999280	43	06/12/2013	Black	Jefferson
498	Offender Information	Last Statement	Williams	Jeffrey	999350	37	05/15/2013	Black	Harris
497	Offender Information	Last Statement	Parr	Carroll	999479	35	05/07/2013	Black	McLennan
496	Offender Information	Last Statement	Cobb	Richard	999467	29	04/26/2013	White	Cherokee

Design Considerations

Title, labels, legend, captions, source!

Expressiveness and Effectiveness

Avoid unexpressive marks (lines? bars? gradients?)

Use perceptually effective encodings

Don't distract: faint gridlines, pastel highlights/fills

The "elimination diet" approach – start minimal

Support comparison and pattern perception

Between elements, to a reference line, or to counts

Design Considerations

Group / sort data by meaningful dimensions

Transform data (e.g., invert, log, normalize)

Are model choices (regression lines) appropriate?

Reduce cognitive overhead

Minimize visual search, minimize ambiguity

Avoid legend lookups if direct labeling works

Avoid color mappings with indiscernible colors

Be consistent! Visual inferences should consistently support data inferences

Design Space of A1 Submissions

Spatial Encoding

Bar charts, Line charts, Area charts
Scatterplots, Maps

Color Encoding

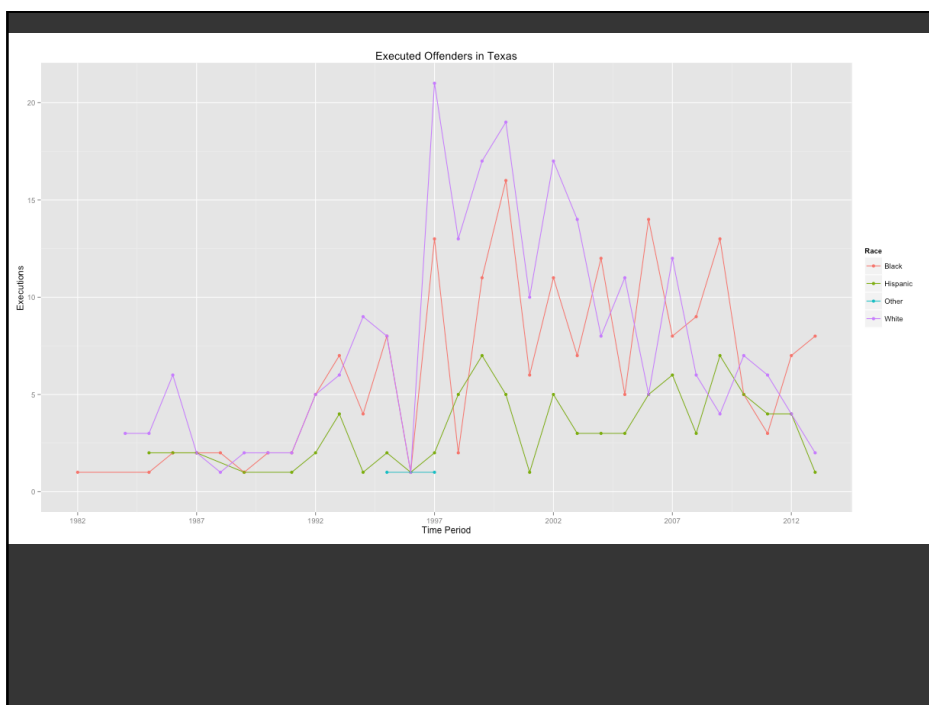
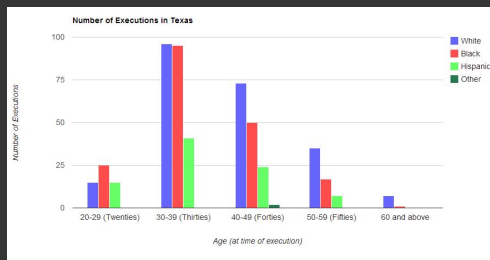
Nominal, Highlights, Never quantitative

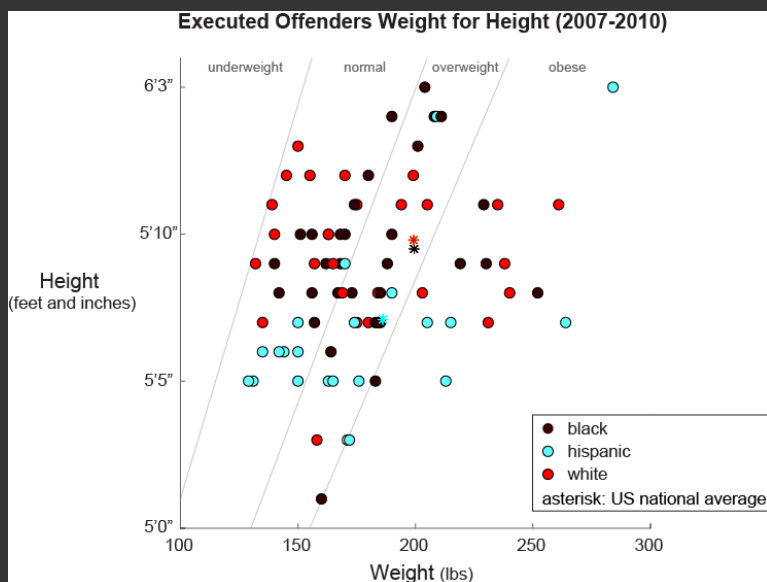
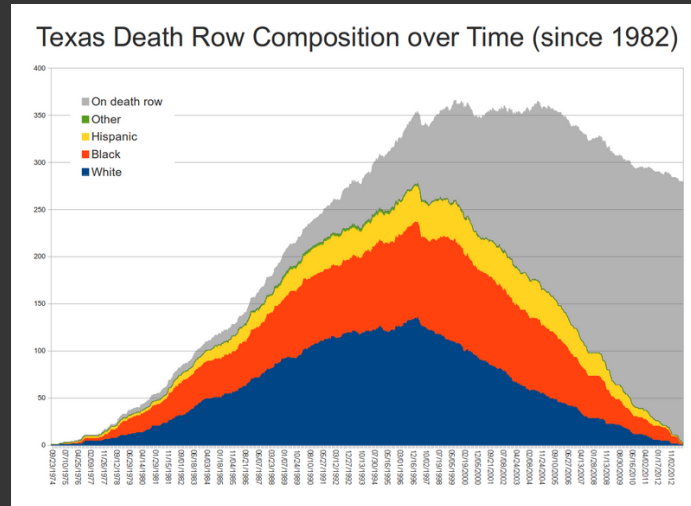
Data Transformation

Often raw counts grouped by
(county, age, weight, etc.)
Text of last statement in some cases

Labeling

Title, Caption, Axis labels
Annotations, photographs of faces



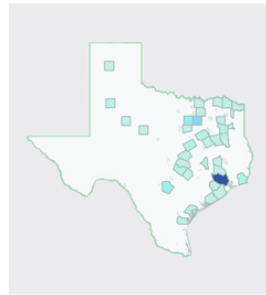


Executed Texas Offenders since 1982

Broken down by Race / Details shown for County

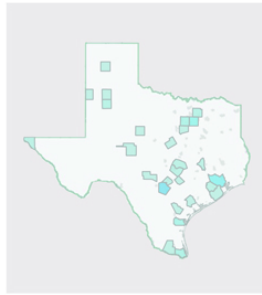


BLACK



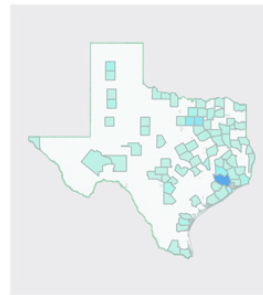
SUBTOTAL: 188

HISPANIC



SUBTOTAL: 87

WHITE



SUBTOTAL: 226

