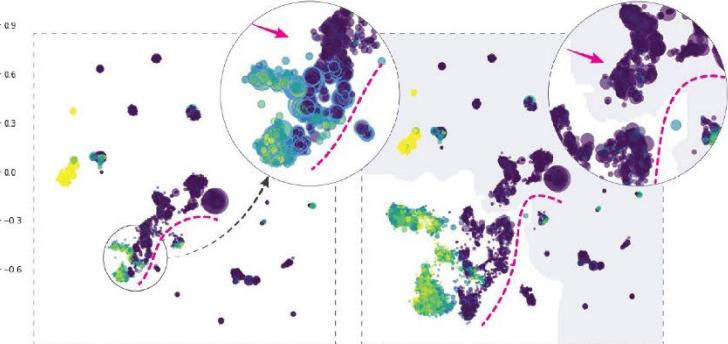
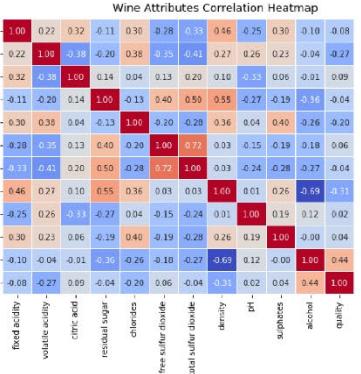
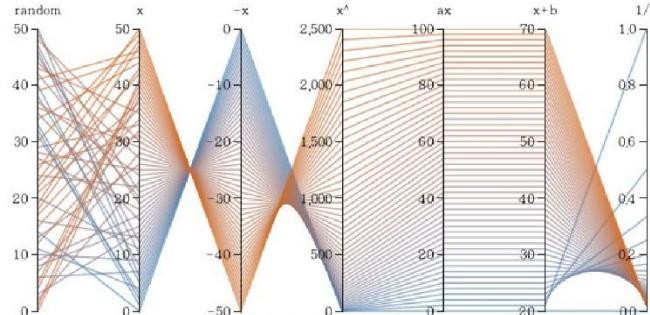


JBI100 Visualization

VISUALIZATION DESIGN AND ENCODINGS

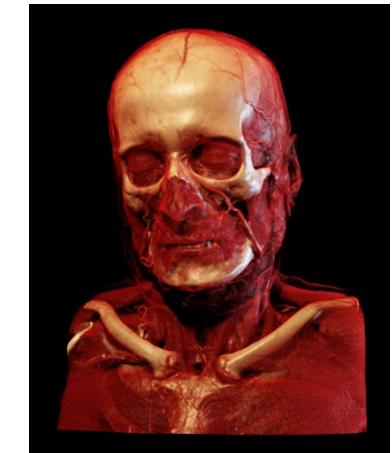
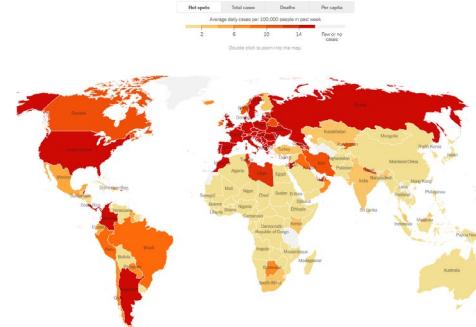
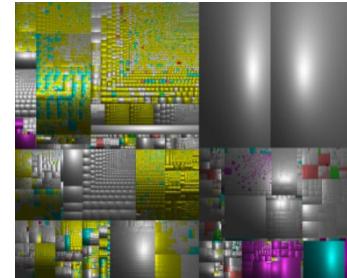
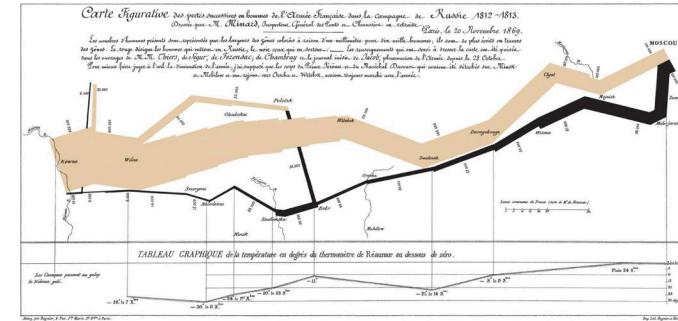
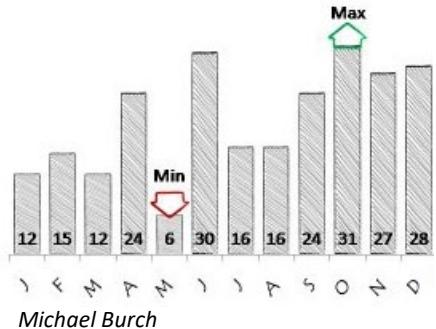


Anna Vilanova, Prof. dr.

Mathematics & Computer Science, Visualization Cluster

What did we see last time ?

- Introduction to data visualization
- What is data visualization?
- When is visualization useful?
- Main types of goals in visualization



Visualizations – Three Types of Goals

Visualization, ...

- ... to **explore**

Nothing is known,

Vis. used for data **Exploration**



- ... to **analyze**

There are hypotheses,

Vis used **Verification or Falsification**



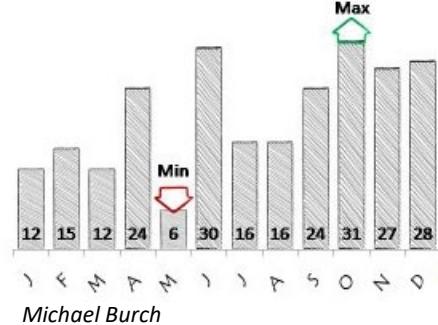
- ... to **present**

“everything” known about the data,

Vis. used for **Communication** of results

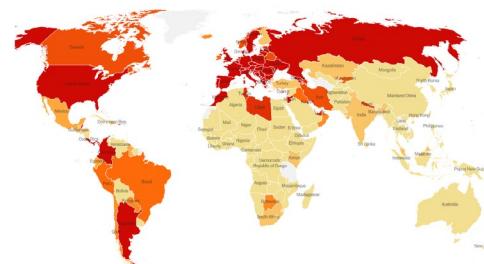
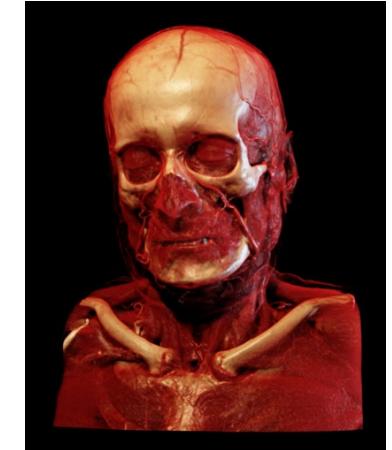
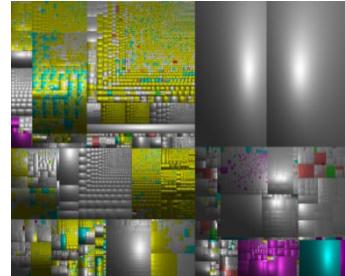
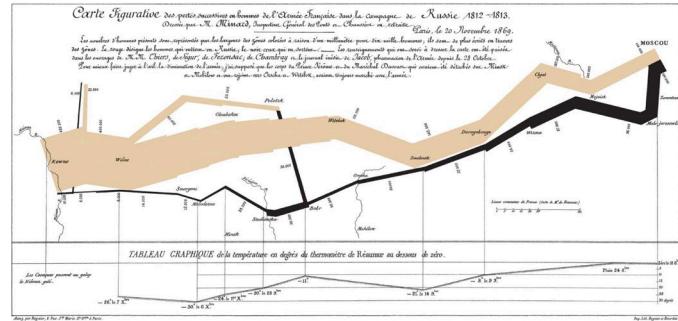
What we saw?

- Introduction to data visualization
- What is data visualization?
- When is visualization useful?
- Main types of goals in visualization
- Visualization design



4

Michael Burch



Visualization Design

- Huge space of design alternatives
 - Tradeoffs abound
- Many possibilities now known to be ineffective
 - avoid random walk through the possibilities
 - avoid some of known mistakes
 - extensive experimentation has already been done
- Guidelines continue to evolve
 - we reflect on lessons learned in design studies
 - iterative refinement usually wise

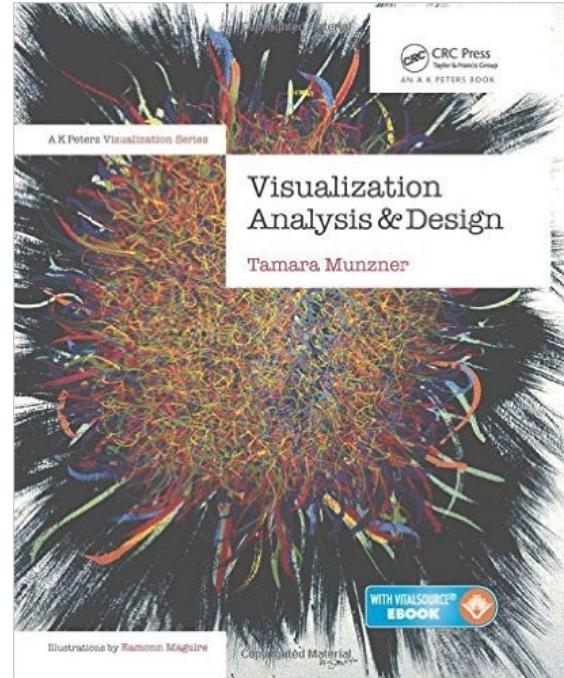
Lectures based on Tamara Munzner's book

Visualization analysis & design

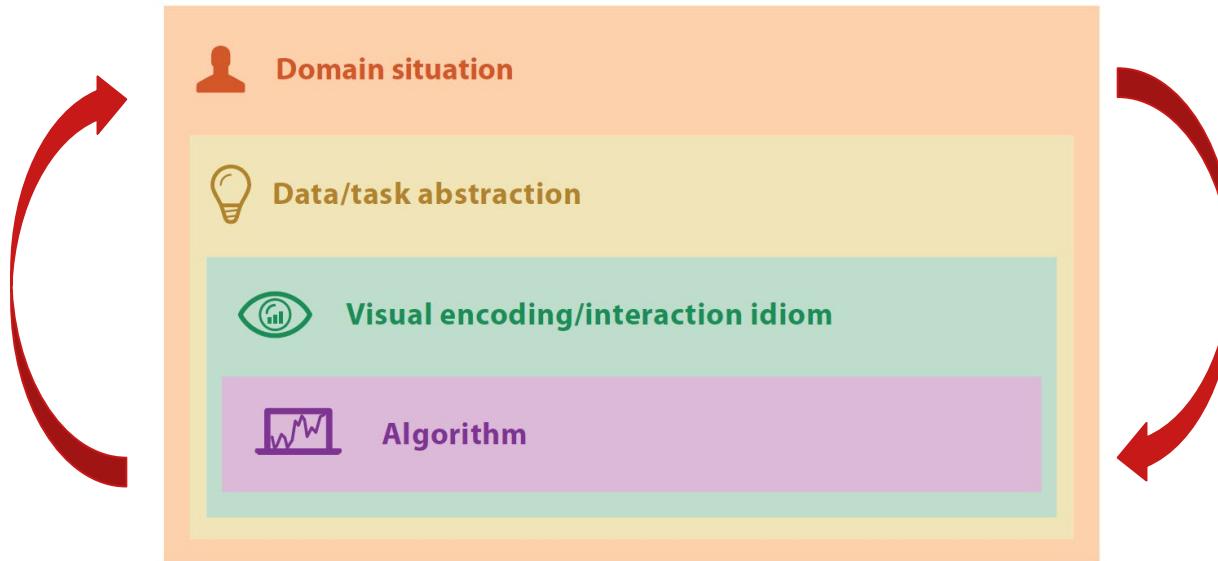
by Tamara Munzner,
eBook 2015
ISBN: 9781466508934



You can find the book at the library:
<https://tue.on.worldcat.org/oclc/897069361>



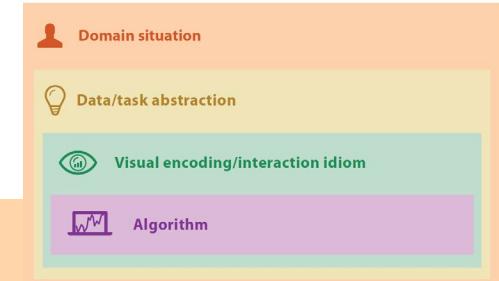
Visualization Design (Nested model)



In practice it will be an iterative/refinement process

Nested model (domain)

domain situation - problem characterization



- Understand the user, the data and tasks

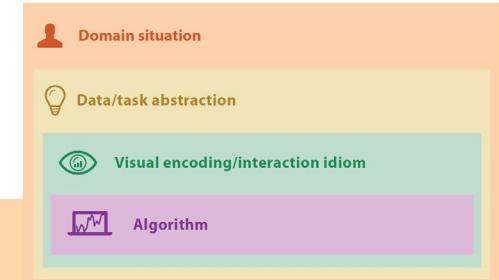
Know your users

- What are their needs/wants/limitations/skills?
- What is their workflow?
- How to provide **actionable knowledge**?
 - What decisions must be made?
 - What information is relevant for that?
- How to make them happy/enthusiastic/satisfied?

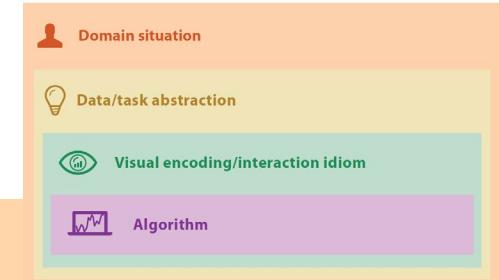
Nested model (domain)

domain situation - problem characterization

- Understand the user, the data and tasks
- Use domain specific vocabulary
- Produce a set of tasks/questions of target users on the data, on different levels
- Information obtained through interviews, observation, reading, ...



Nested model (data/task abstraction)



domain situation - problem characterization

data/task abstraction design

- Data described in generic (Visualization) terms:
table, hierarchy, sets, ...
- Tasks described in generic (Visualization) terms:
search, compare, see trend, ...

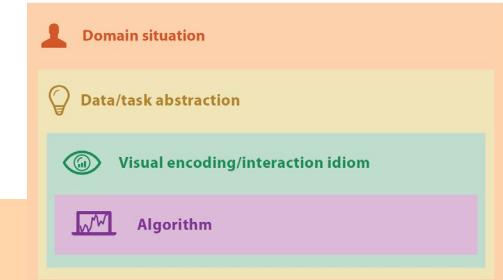
Nested model (Visual encoding)

domain situation - problem characterization

data/task abstraction design

visual encoding/interaction idiom

Design space, select visual encodings, define interactions, etc.



Visualization/interaction design

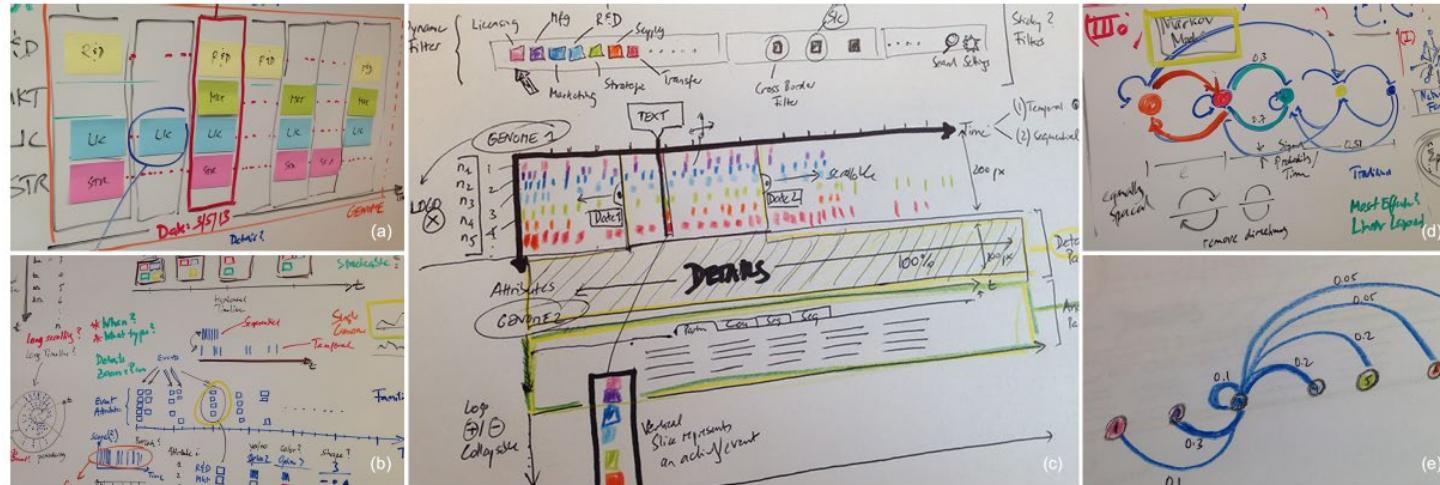


Fig. 1. Following guidelines by [McKenna et al. 2014], the development of our visualization involved an iterative process of design ideation, sketching, prototyping, and refinement between members of the research team and three prototypical corporate users. As corporate analysts are often not familiar or clear with potential visualization alternatives, we used dry-erase boards (a,b, and d), color paper sketches (c,e) and post-it notes (a) to brainstorm potential visualization ideas.

Generating solutions

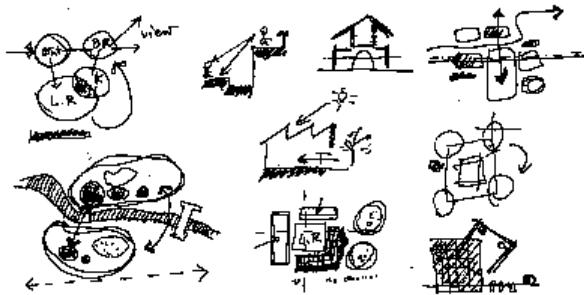
- Engineering, systematic approach
- Creative approach
- Exploring design space!

Approach:

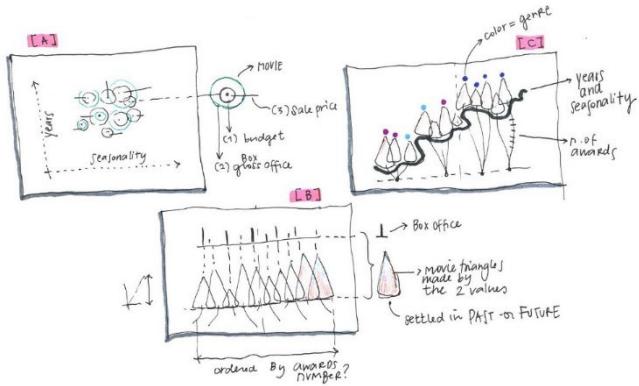
- Generate as many as you can, as fast as you can
- Don't be critical
- Look back
- Use sketching

Sketching!

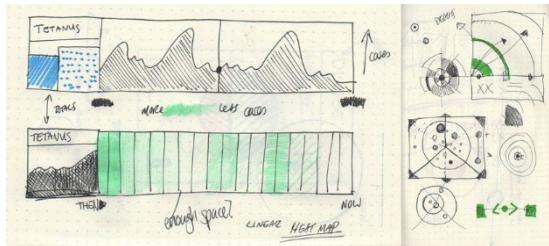
- Explore design space
- Find solutions
- Find structure
- Find and break constraints
- Inspire yourself



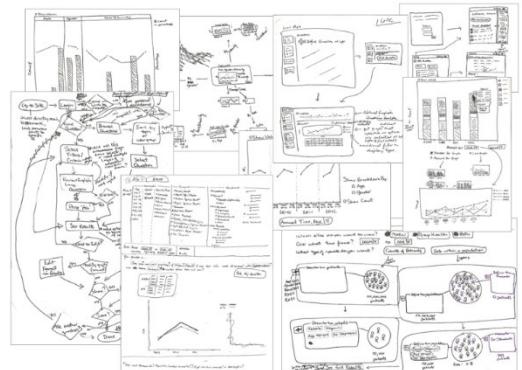
<http://depts.washington.edu/napkin/drawing.html>



<https://medium.com/accurat-studio/sketching-with-data-opens-the-mind-s-eye-92d78554565>



<http://beyondwordsstudio.com/our-work/global-health-check/>



<http://karminmauritz.com/data-viz-for-group-health-research-institute/>

Generating solutions

- Engineering, systematic approach
- Creative approach
- Exploring design space!

Approach:

- Generate as many as you can, as fast as you can
- Don't be critical
- Look back
- Use sketching

Nested model (algorithm)

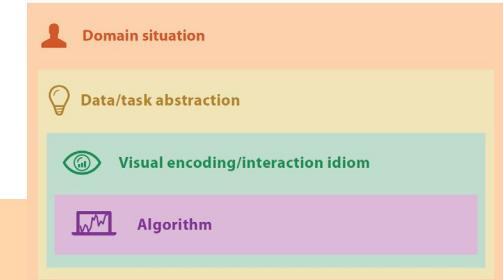
domain situation - problem characterization

data/task abstraction design

visual encoding/interaction idiom

algorithm

layout algorithm, ordering, rendering ...



Nested model - Dangers at each level

A mistake at the higher level cannot be corrected on the lower

👤 Domain situation

You misunderstood their needs

💡 Data/task abstraction

You're showing them the wrong thing

👁️ Visual encoding/interaction idiom

The way you show it doesn't work

💻 Algorithm

Your code is too slow

Nested model - Validation

👤 Domain situation

Observe target users using existing tools

💡 Data/task abstraction

👁️ Visual encoding/interaction idiom

Justify design with respect to alternatives

💻 Algorithm

Measure system time/memory

Analyze computational complexity

Analyze results qualitatively

Measure human time with lab experiment (*lab study*)

Observe target users after deployment (*field study*)

Measure adoption



Analysis framework

- **What** is shown?
data abstraction
- **Why** is the user looking at it?
task abstraction
- **How** is it shown?
visual encoding and interaction

domain situation - problem characterization

data/task abstraction design

domain situation - problem characterization

data/task abstraction design

visual encoding/interaction idiom

Analysis framework

- **What** is shown?
data abstraction
- **Why** is the user looking at it?
task abstraction
- **How** is it shown?
visual encoding and interaction

domain situation - problem characterization

data/task abstraction design

Munzner's Categorization (Data Abstraction)

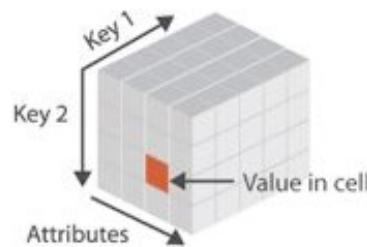
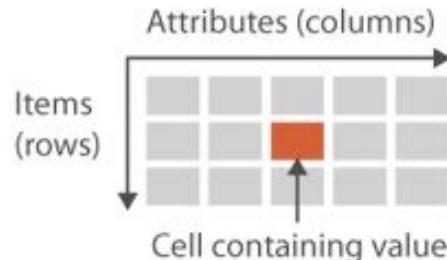
Datasets

④ Data Types

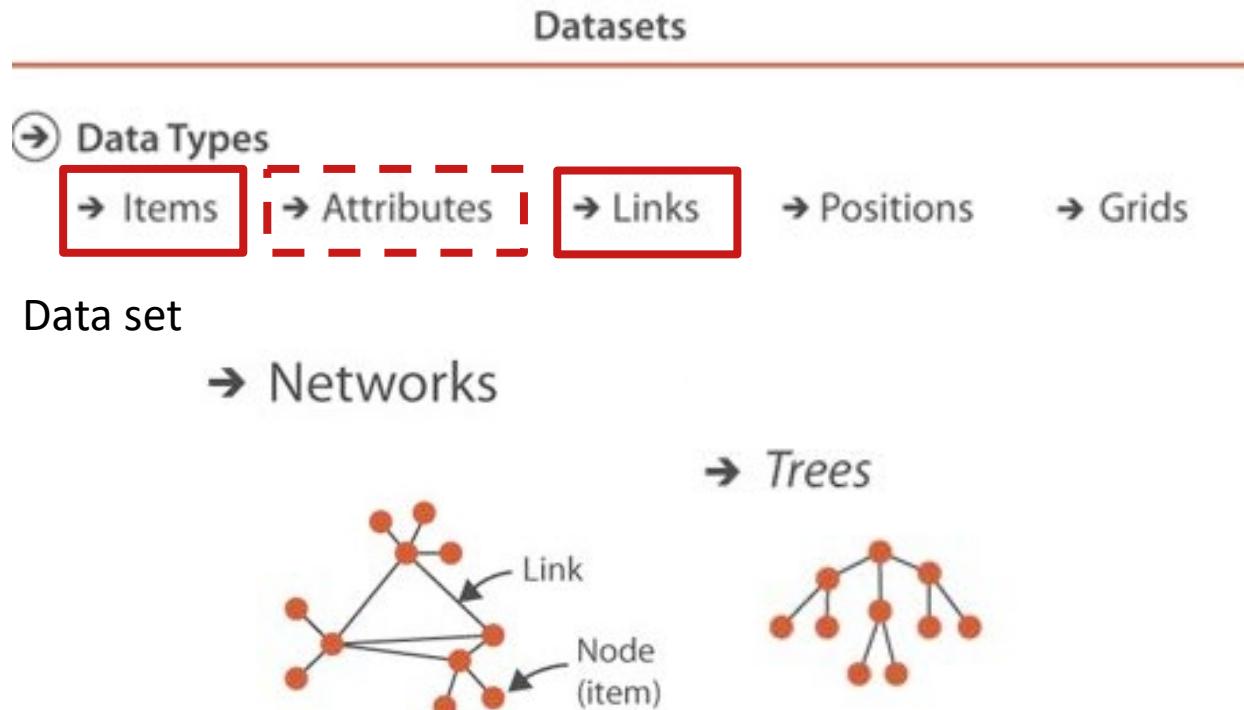
- Items → Attributes
- Links → Positions → Grids

→ Tables

→ *Multidimensional Table*



Munzner's Categorization (Data Abstraction)



Munzner's Categorization (Data Abstraction)

Datasets

④ Data Types

→ Items

→ Attributes

→ Links

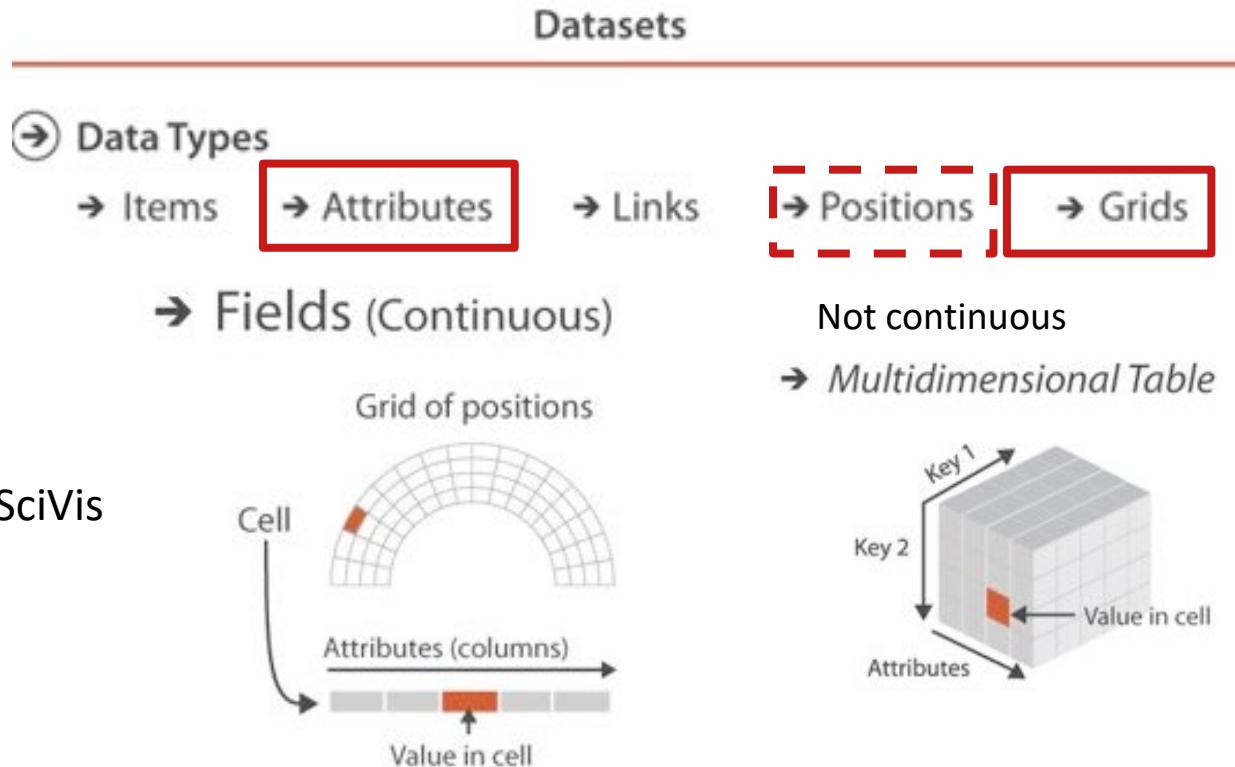
→ Positions

→ Grids

→ Geometry (Spatial)



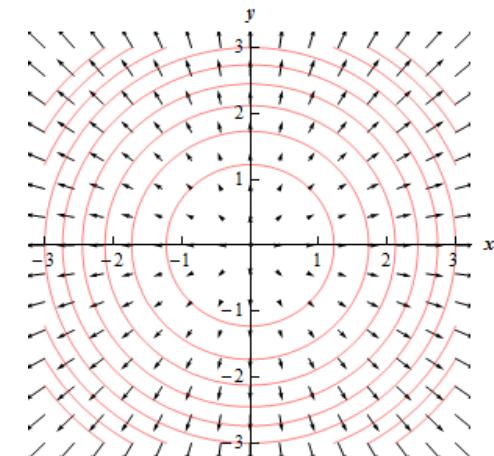
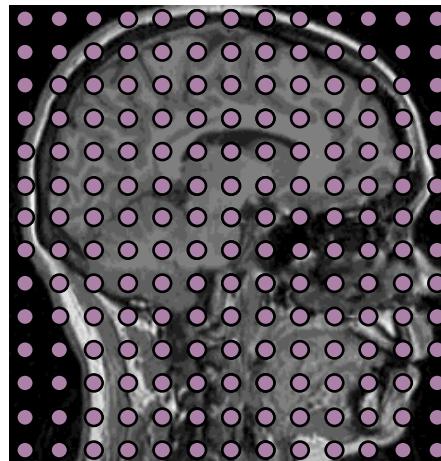
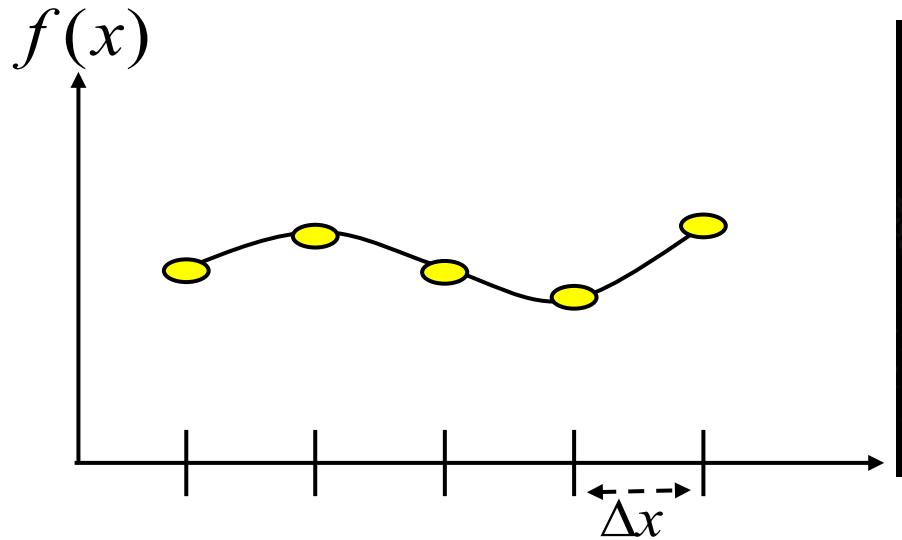
Munzner's Categorization (Data Abstraction)

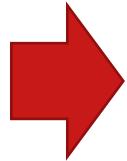
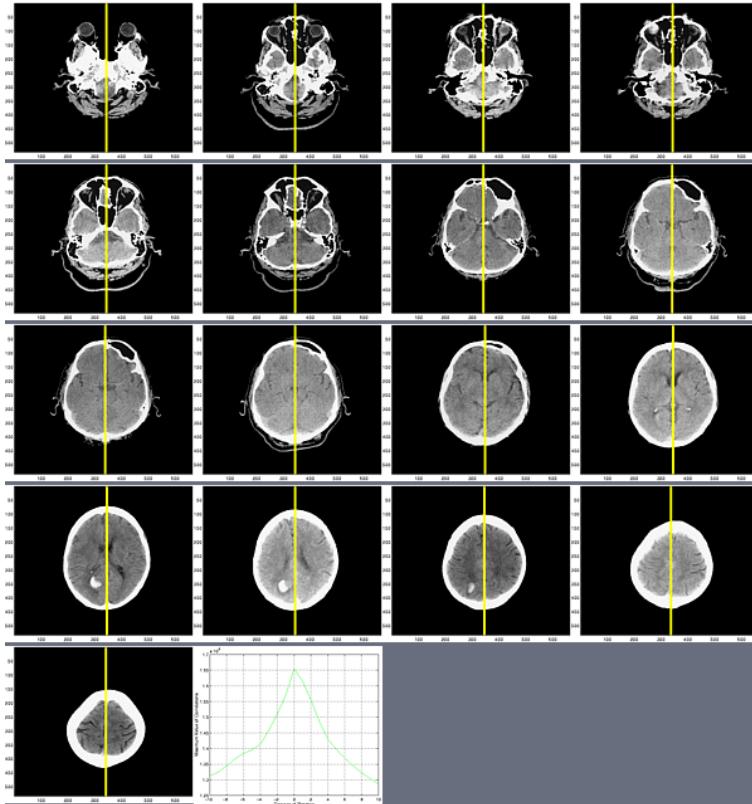


Table

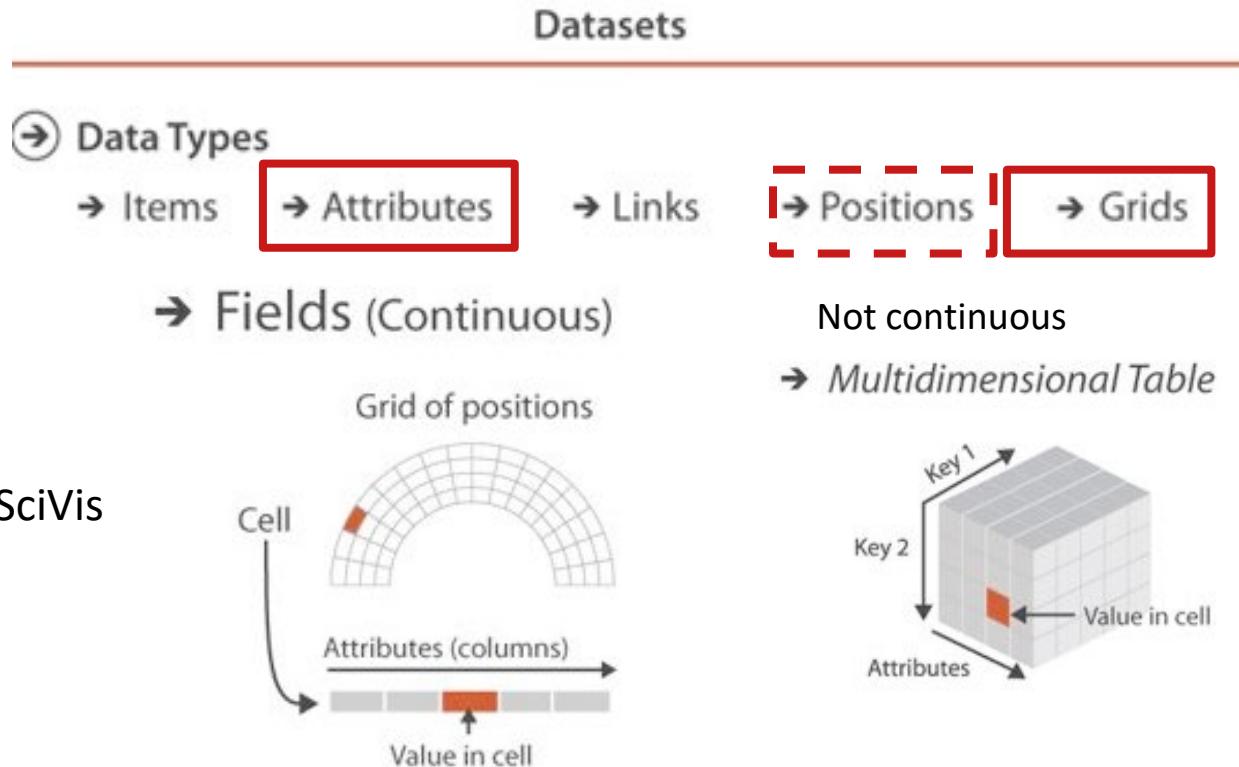
ID	Name	Age	Shirt Size	Favorite Fruit
1	Amy	8	S	Apple
2	Basil	7	S	Pear
3	Clara	9	M	Durian
4	Desmond	13	L	Elderberry
5	Ernest	12	L	Peach
6	Fanny	10	S	Lychee
7	George	9	M	Orange
8	Hector	8	L	Loquat
9	Ida	10	M	Pear
10	Amy	12	M	Orange

Field





Munzner's Categorization (Data Abstraction)



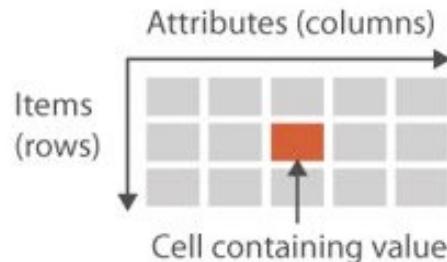
We will focus mainly on tabular data

Datasets

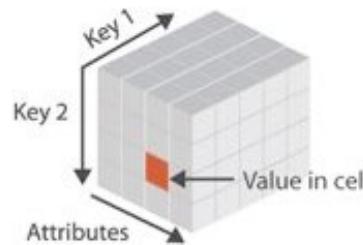
④ Data Types

- Items → Attributes
- Links → Positions → Grids

→ Tables



→ *Multidimensional Table*



Example – tabular data

Items and attributes (key)

Key	Attribute				
	ID	Name	Age	Shirt Size	Favorite Fruit
1	Amy		8	S	Apple
2	Basil		7	S	Pear
3	Clara		9	M	Durian
4	Desmond		13	L	Elderberry
5	Ernest		12	L	Peach
6	Fanny		10	S	Lychee
7	George		9	M	Orange
8	Hector		8	L	Loquat
9	Ida		10	M	Pear
10	Amy		12	M	Orange

item

Munzner's Categorization

Attributes

➔ Attribute Types

➔ Categorical

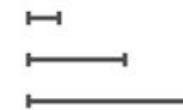


➔ Ordered

➔ Ordinal



➔ Quantitative



➔ Ordering Direction

➔ Sequential



➔ Diverging



➔ Cyclic



Attributes

Attribute Types

→ Categorical

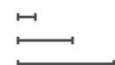


→ Ordered

→ Ordinal



→ Quantitative



Ordering Direction

→ Sequential



→ Diverging



→ Cyclic



ID	Name	Age	Shirt Size	Favorite Fruit
1	Amy	8	S	Apple
2	Basil	7	S	Pear
3	Clara	9	M	Durian
4	Desmond	13	L	Elderberry
5	Ernest	12	L	Peach
6	Fanny	10	S	Lychee
7	George	9	M	Orange
8	Hector	8	L	Loquat
9	Ida	10	M	Pear
10	Amy	12	M	Orange

A	B	C	S	T	U
Order ID	Order Date	Order Priority	Product Container	Product Base Margin	Ship Date
3	10/14/06	5-Low	Large Box	0.8	10/21/06
6	2/21/08	4-Not Specified	Small Pack	0.55	2/22/08
32	7/16/07	2-High	Small Pack	0.79	7/17/07
32	7/16/07	2-High	Jumbo Box	0.72	7/17/07
32	7/16/07	2-High	Medium Box	0.6	7/18/07
32	7/16/07	2-High	Medium Box	0.65	7/18/07
35	10/23/07	4-Not Specified	Wrap Bag	0.52	10/24/07
35	10/23/07	4-Not Specified	Small Box	0.58	10/25/07
36	11/3/07	1-Urgent	Small Box	0.55	11/3/07
65	3/18/07	1-Urgent	Small Pack	0.49	3/19/07
66	1/20/05	5-Low	Wrap Bag	0.56	1/20/05
69	6/4/05	4-Not Specified	Small Pack	0.44	6/6/05
69	6/4/05	4-Not Specified		0.6	6/6/05
70	12/18/06	5-Low		0.59	12/23/06
70	12/18/06	5-Low		0.82	12/23/06
96	4/17/05	2-High		0.55	4/19/05
97	1/29/06	3-Medium		0.38	1/30/06
129	11/19/08	5-Low		0.37	11/28/08
130	5/8/08	2-High	Small Box	0.37	5/9/08
130	5/8/08	2-High	Medium Box	0.38	5/10/08
130	5/8/08	2-High	Small Box	0.6	5/11/08
132	6/11/06	3-Medium	Medium Box	0.6	6/12/06
132	6/11/06	3-Medium	Jumbo Box	0.69	6/14/06
134	5/1/08	4-Not Specified	Large Box	0.82	5/3/08
135	10/21/07	4-Not Specified	Small Pack	0.64	10/23/07
166	9/12/07	2-High	Small Box	0.55	9/14/07
193	8/8/06	1-Urgent	Medium Box	0.57	8/10/06
194	4/5/08	3-Medium	Wrap Bag	0.42	4/7/08

quantitative
ordinal
categorical

Munzner's Categorization (Data Abstraction)

→ Dataset Availability

→ Static



→ Dynamic



Offline (most cases)

vs Online (streaming – time varying)

New Add/Remove/Change attribute and items.

For Example: Financial data sensor data, etc.

Munzner's Categorization (Data Abstraction)

Starting point!

Real data sets might be more complex than the basic types presented:

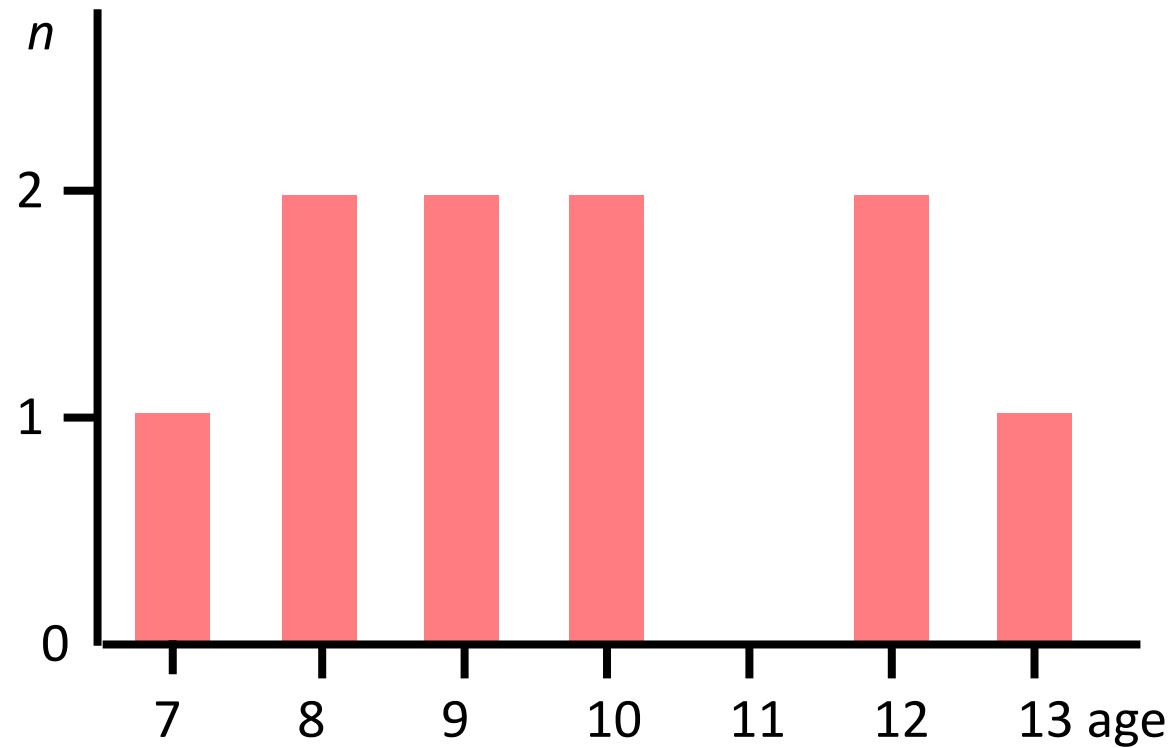
- Sets/ Groups
- Combinations of previous defined data sets
- New types
- ...

One data set can be looked at in different way

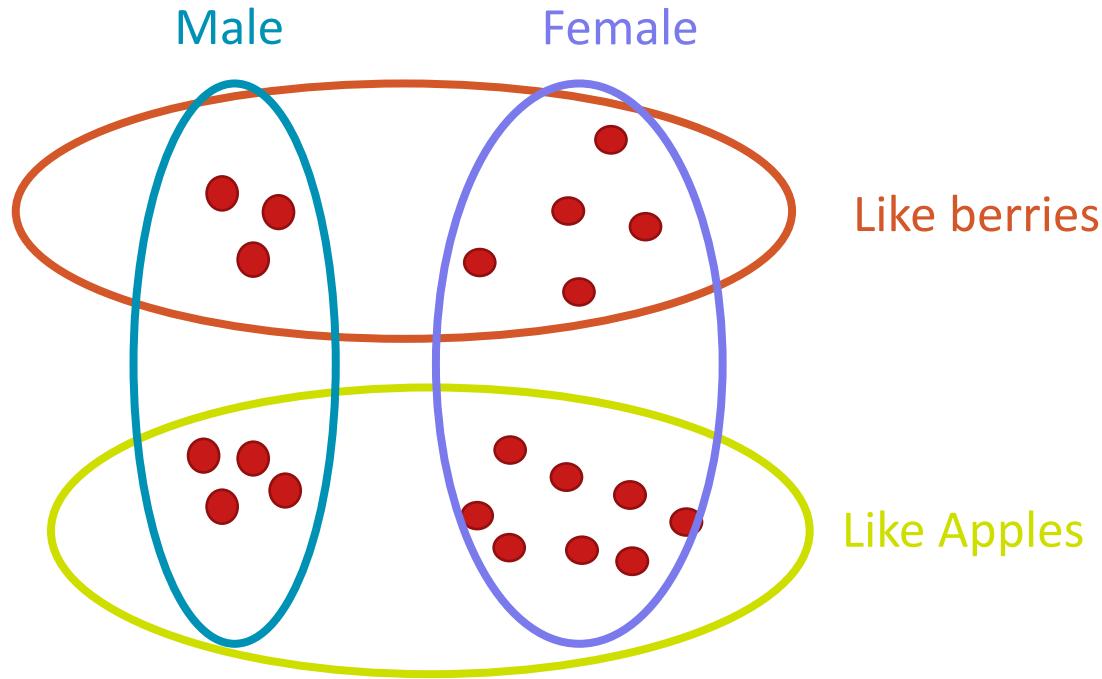
Multivariate data: tables

ID	Name	Age	Shirt Size	Favorite Fruit
1	Amy	8	S	Apple
2	Basil	7	S	Pear
3	Clara	9	M	Durian
4	Desmond	13	L	Elderberry
5	Ernest	12	L	Peach
6	Fanny	10	S	Lychee
7	George	9	M	Orange
8	Hector	8	L	Loquat
9	Ida	10	M	Pear
10	Amy	12	M	Orange

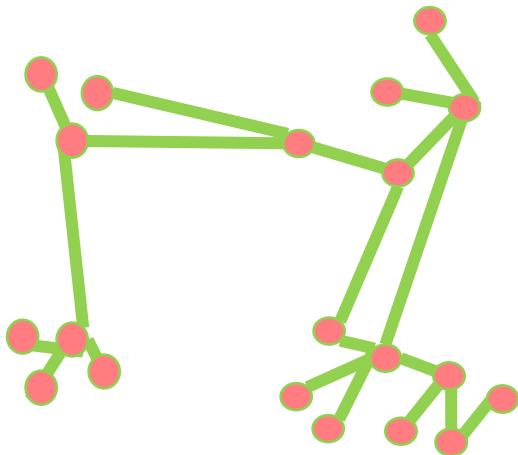
Distribution per attribute



Sets



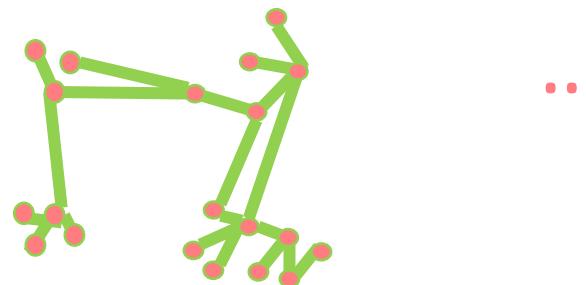
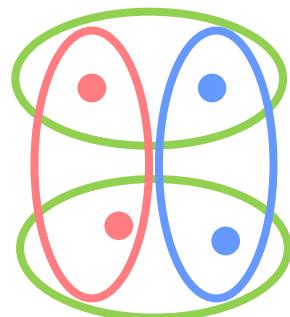
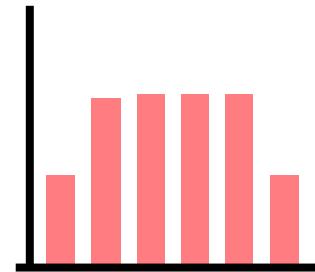
Network



Measure a distance
(e.g. similar ages)

One data set, multiple views

ID	Name	Age	Shirt Size	Favorite Fruit
1	Amy	8	S	Apple
2	Basil	7	S	Pear
3	Clara	9	M	Durian
4	Desmond	13	L	Elderberry
5	Ernest	12	L	Peach
6	Fanny	10	S	Lychee
7	George	9	M	Orange
8	Hector	8	L	Loquat
9	Ida	10	M	Pear
10	Amy	12	M	Orange





Analysis framework

- What is shown?
data abstraction
- Why is the user looking at it?
task abstraction
- How is it shown?
visual encoding and interaction

domain situation - problem characterization

data/task abstraction design

visual encoding/interaction idiom

