Data Visualization

DATA ABSTRACTION

Outline

Motivation

Dataset types

Attribute types

Attribute semantics

Implications for design

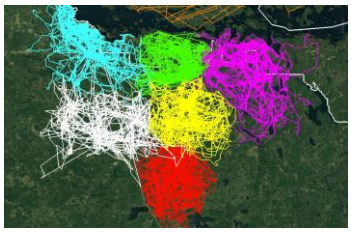
Motivation

Data is typically described with the domain language

In order to visualize it, it needs to be translated to more abstract structures that we know how to represent

Examples

- Retweets with the hashtag #GiletsJaunes: network data
- Movement of wolves: spatial data

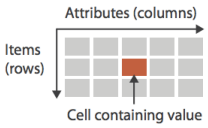


https://earthlymission.com/gps-tracking-shows-how-much-wolf-packs-avoid-each-others-range/
https://twitter.com/fs0c131y/status/1070978229224267776

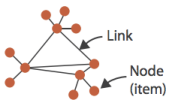
3

Dataset types

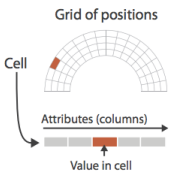
→ Tables



→ Networks

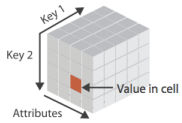


→ Fields (Continuous)



→ Geometry (Spatial)

→ *Multidimensional Table*



→ *Trees*

A dataset is any collection of information that is the target of analysis

T. Munzner. *Visualization Analysis & Design*. CRC Press, Boca Raton, 2014

4

Continuous fields

Each cell in a field contains measurements or computations from a continuous domain

Scientific visualization

Multivariate (#attributes)
- Scalar field
- Vector field
- Tensor field

Multidimensional (#keys)
- 2-D
- 3-D

Dataset types

Data and Dataset Types

Tables	Networks & Trees	Fields	Geometry	Clusters, Sets, Lists
Items	Items (nodes)	Grids	Items	Items
Attributes	Links	Positions	Positions	
	Attributes	Attributes		

Data Types

→ Items → Attributes → Links → Positions → Grids

Dataset types

- **Dataset Availability**
  - Static
  - Dynamic

Static = offline (entire dataset available all at once)

Dynamic = online (dataset information trickles in over time)

Attribute types

- **Attribute Types**
  - Categorical
  - Ordered
    - *Ordinal*
    - *Quantitative*

- **Ordering Direction**
  - Sequential
  - Diverging
  - Cyclic

Attribute types

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 Spec		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

Attribute semantics (meaning)

Attribute type does not tell us about its semantics

Key vs. value

- Keys are unique attributes that act as an index to look up values

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 Spec		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

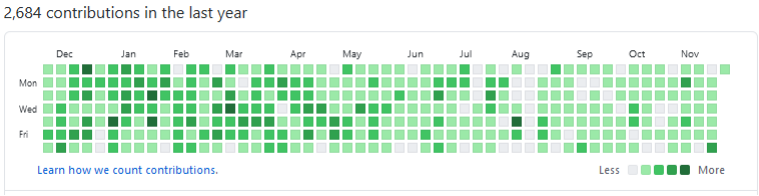
quantitative
ordinal
categorical

Temporal semantics

Temporal = relates to time

Complicated to handle
- Hierarchical structure
- Cyclic
- Transformations and aggregations can be challenging (weeks do not fit neatly into months)

Can be values or keys

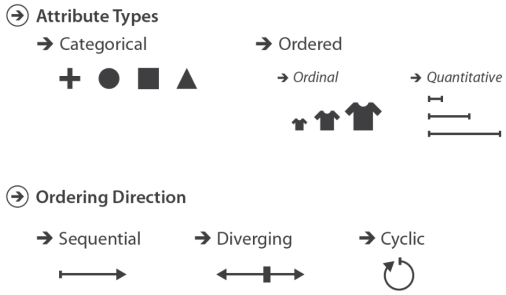


Hierarchical attributes

Some attributes may have an internal hierarchical structure
- Dates (individual days, weeks, months, …, centuries)
- Spatial regions
- Taxonomies

Implications for design

Design choices highly depend on the type and values of the data (color, chart type, …)

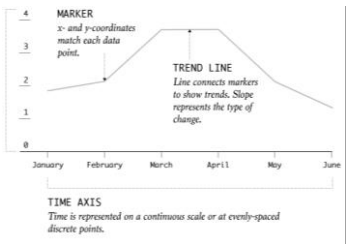


T. Munzner. *Visualization Analysis & Design*. CRC Press, Boca Raton, 2014 13

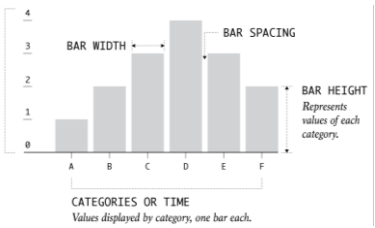
Implications for design

Line chart: Time or other continuous value

What can be placed on the x axis?



Bar chart: Category or discrete time



https://flowingdata.com/charttype/line-chart-type/
https://flowingdata.com/charttype/bar-chart-type/

14