paul.rosen@utah.edu @paulrosenphd https://cspaul.com

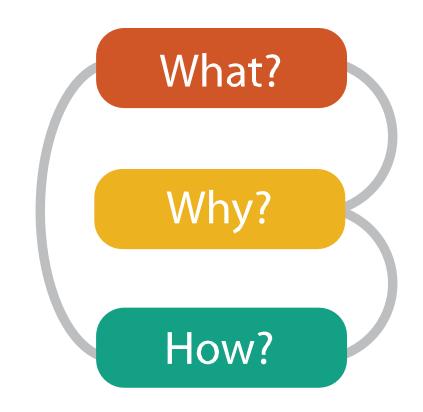


# Visualization for Data Science DS-4630 / CS-5630 / CS-6630

DATA ABSTRACTION

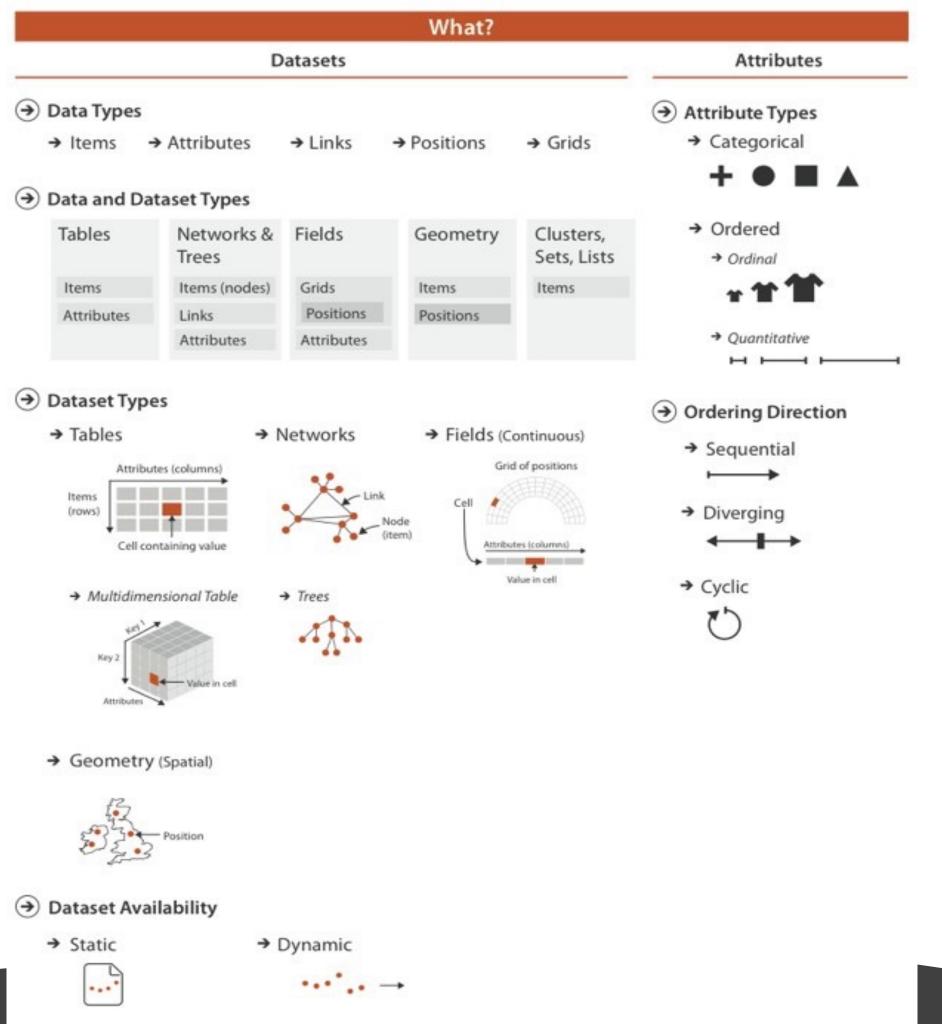
#### analysis: what, why, and how

- •what is shown?
- •why is the user looking at it?
- •how is it shown?



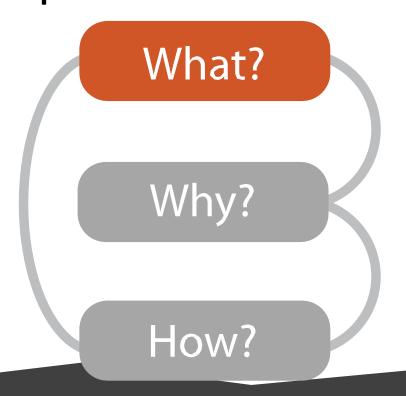
- abstract vocabulary avoids domain-specific terms
- what-why-how analysis framework as scaffold to think systematically about design space





#### data abstraction

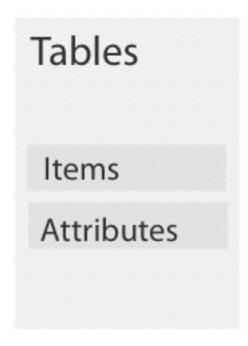
- the *what* part of an analysis that pertains to the data
- translation of domain-specific terms into words that are as generic as possible



## data types

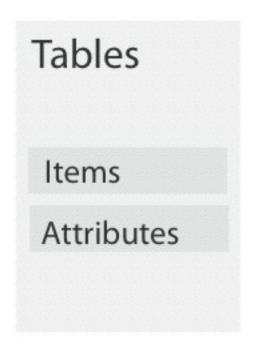
- Items
- Attributes
- Links
- Positions
- Grids

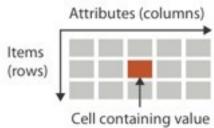




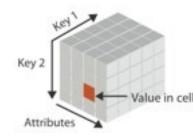


Α	В	С	S	Т	U
Order ID	Order Date	Order Priority	Product Container	Product Base Margin	Ship Date
3	10/14/06	5-Low	Large Box	0.8	10/21/06
6	2/21/08	4-Not Specified	Small Pack	0.55	2/22/08
32	7/16/07	2-High	Small Pack	0.79	7/17/07
32	7/16/07	2-High	Jumbo Box	4.4 • 1 4	7/17/07
32	7/16/07	2-High	Medium Box	attribute	7/18/07
32	7/16/07	2-High	Medium Box	0.03	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	itom 5	4-Not Specified	Small Pack	0.44	6/6/05
69	item 5	4-Not Specified	Wrap Bag	0.6	6/6/05
70	12/18/06	5-Low	Small Box	0.59	12/23/06
70	12/18/06	5-Low	Wrap Bag	0.82	12/23/06
96	4/17/05	2-High	Small Box	0.55	4/19/05
97	1/29/06	3-Medium	Small Box	0.38	1/30/06
129	11/19/08	5-Low	Small Box	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

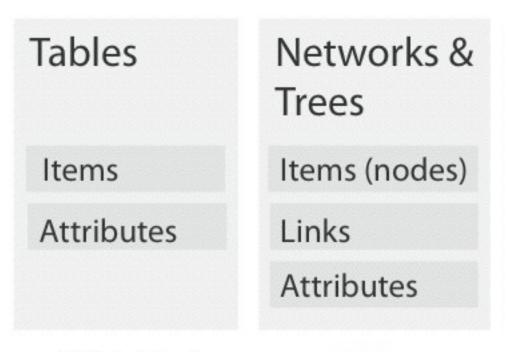


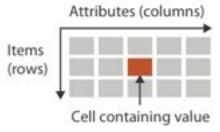


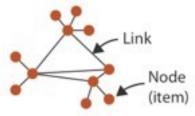
#### → Multidimensional Table



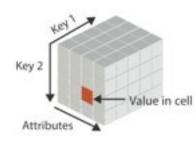








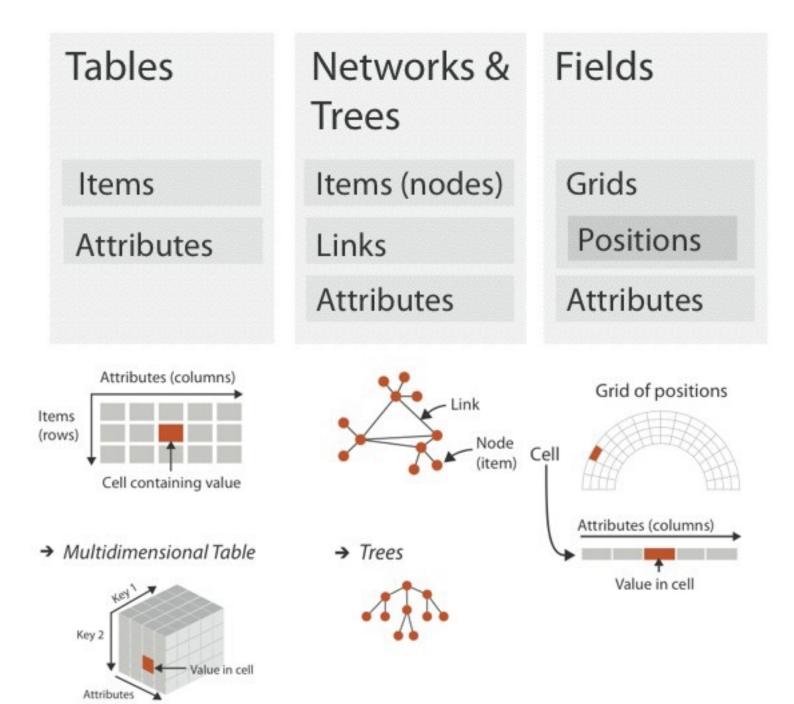
→ Multidimensional Table



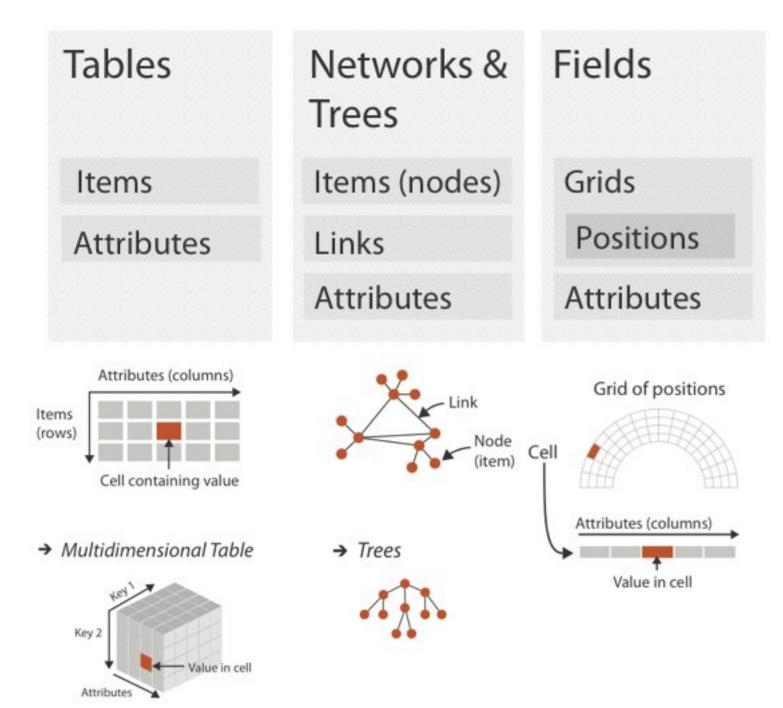






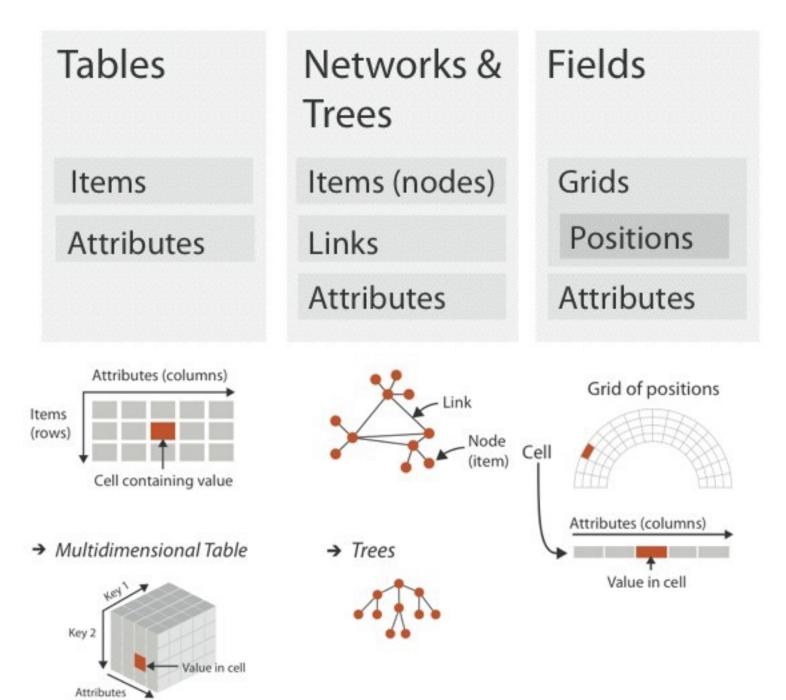


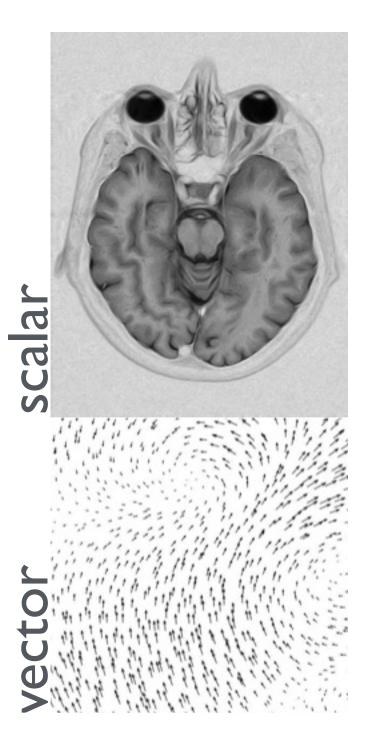






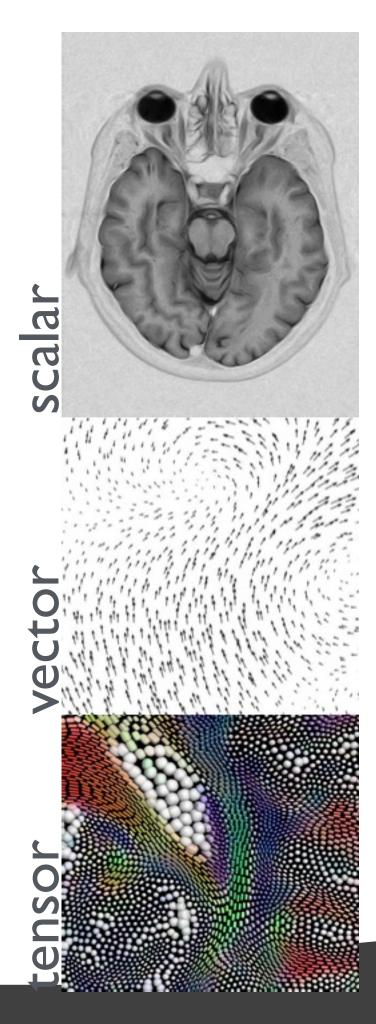








**Tables** Networks & **Fields** Trees Items (nodes) Items Grids **Positions** Attributes Links Attributes Attributes Attributes (columns) Grid of positions Items (rows) Node (item) Cell containing value Attributes (columns) → Multidimensional Table → Trees Value in cell Key 2

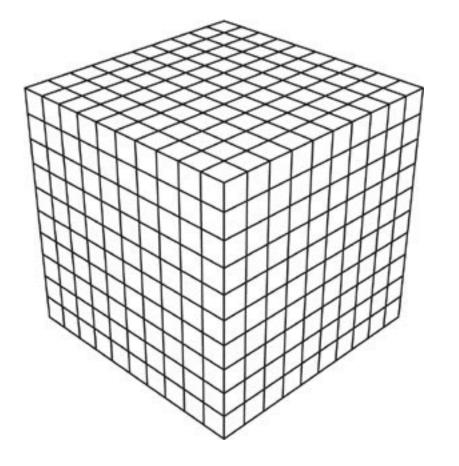




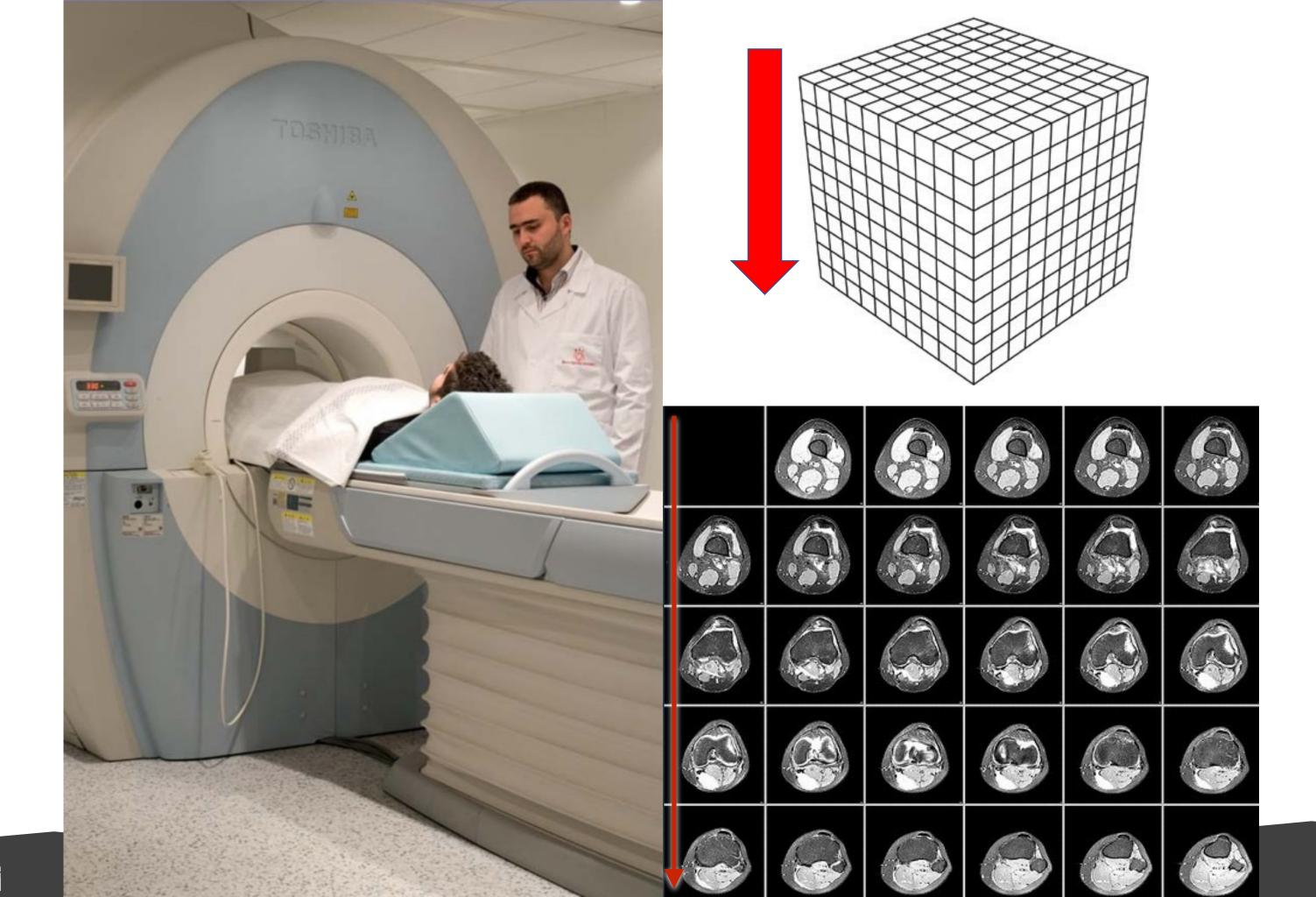






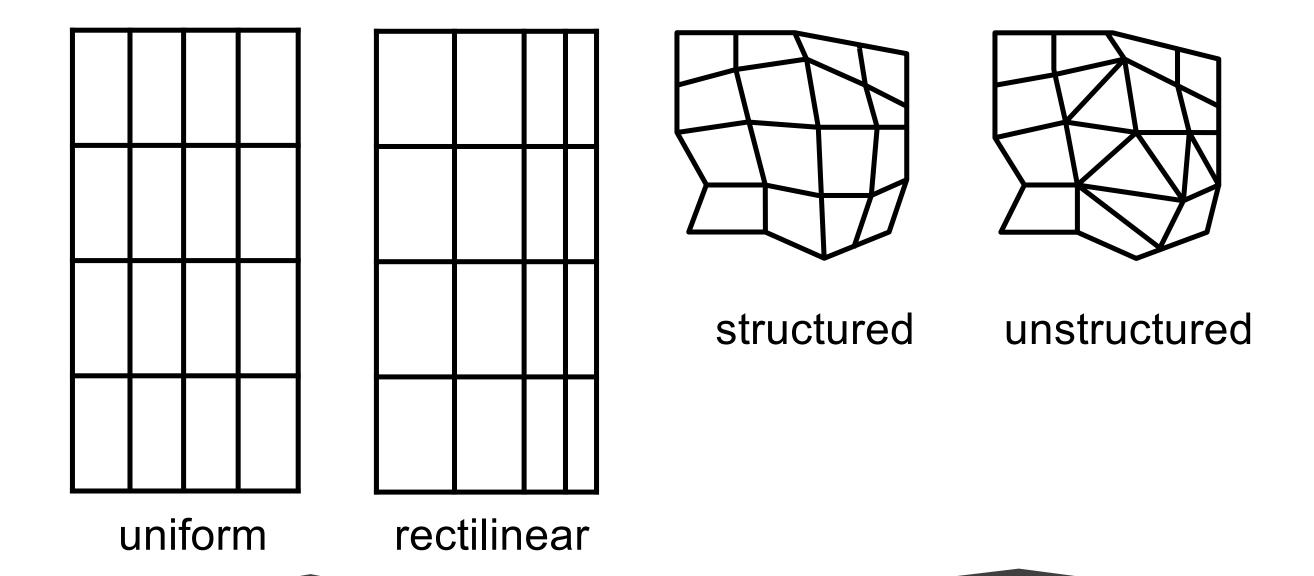








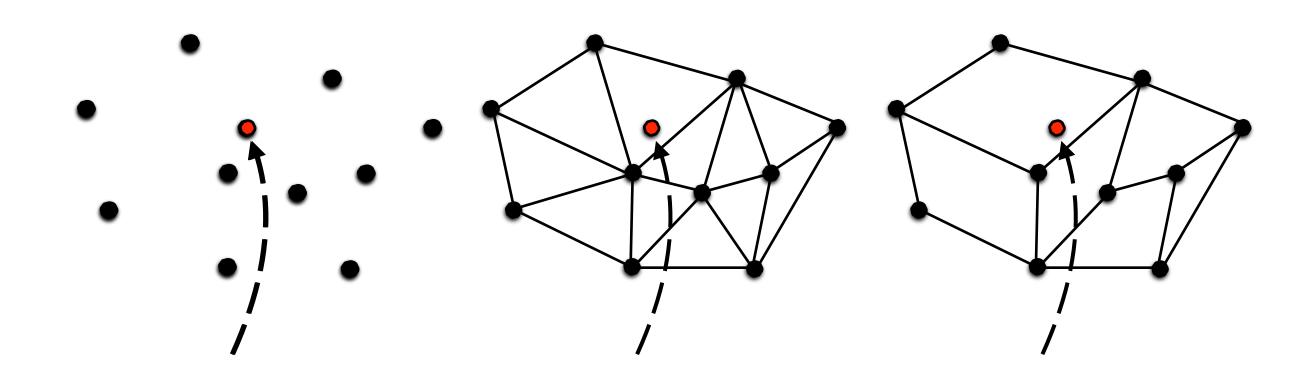
## grid types



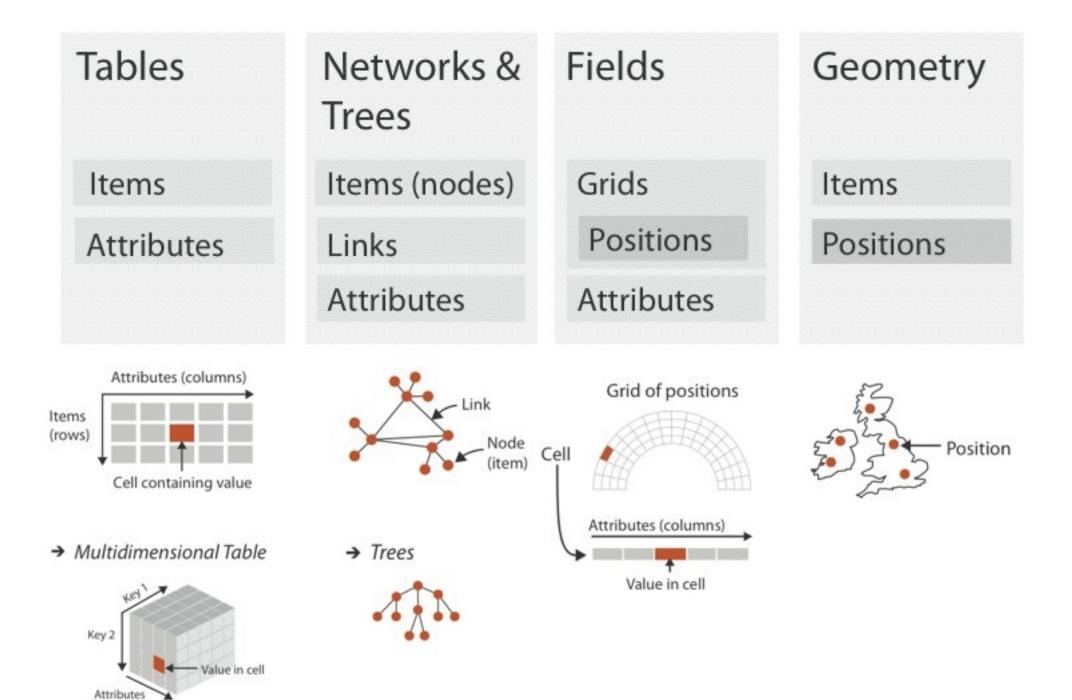


## grid choices impact how continuous data is interpreted

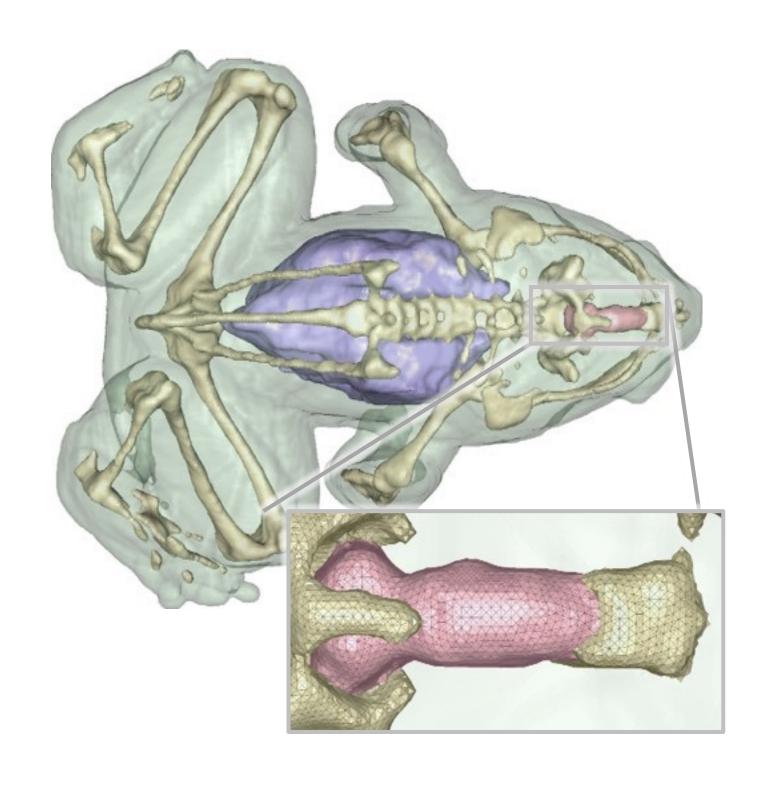
- two key considerations:
  - Sampling the choice of where attributes are measured
  - Interpolation how to model the attributes in the rest of space

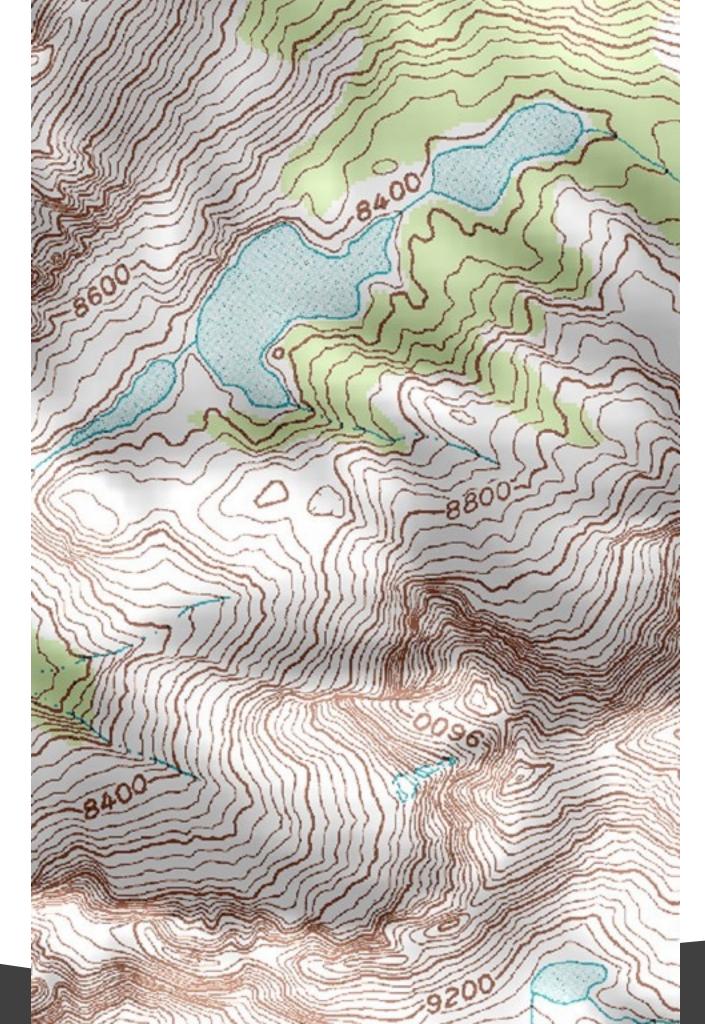






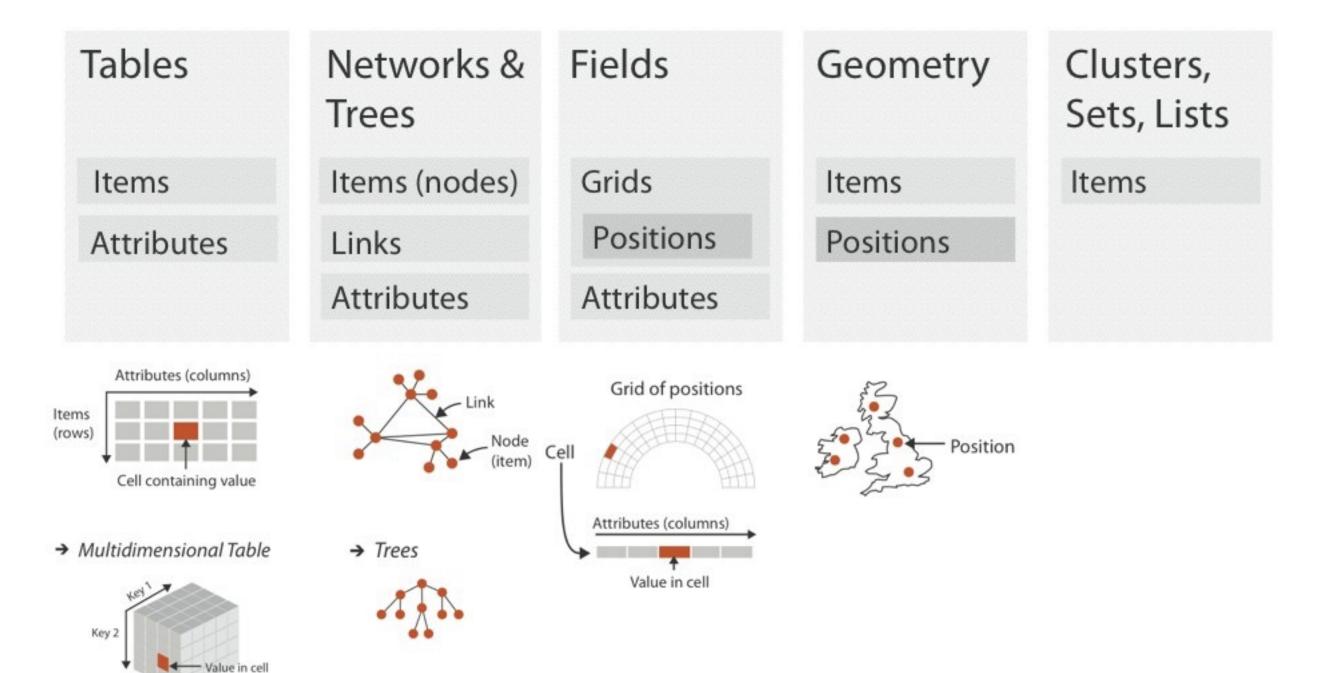








Attributes





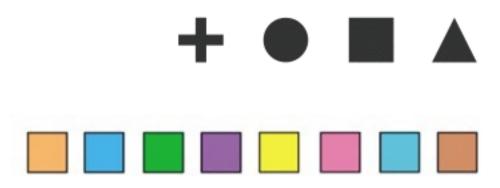
#### **Categorical**

no implicit ordering



#### **Categorical**

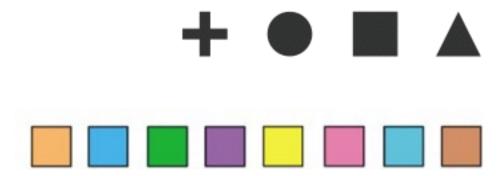
no implicit ordering





#### **Categorical**

no implicit ordering



#### **Ordered**

<u>Ordinal</u>

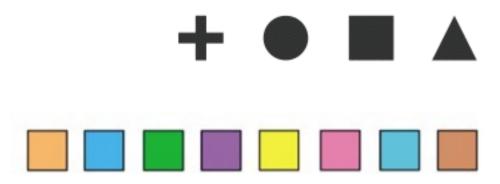
**Quantitative** 





#### **Categorical**

no implicit ordering



#### **Ordered**

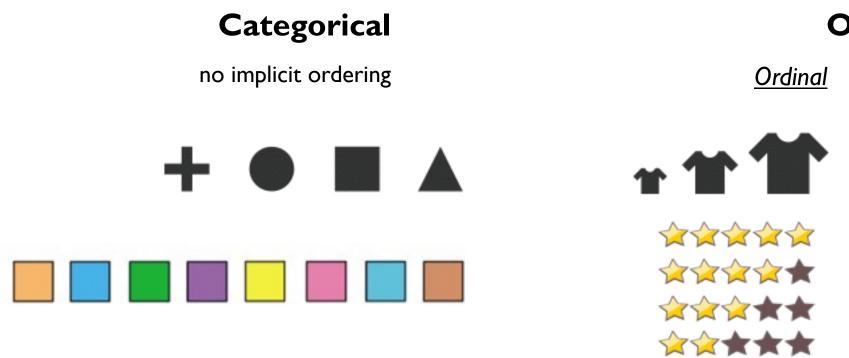
<u>Ordinal</u>

**Quantitative** 

meaningful magnitude (can do arithmetic)





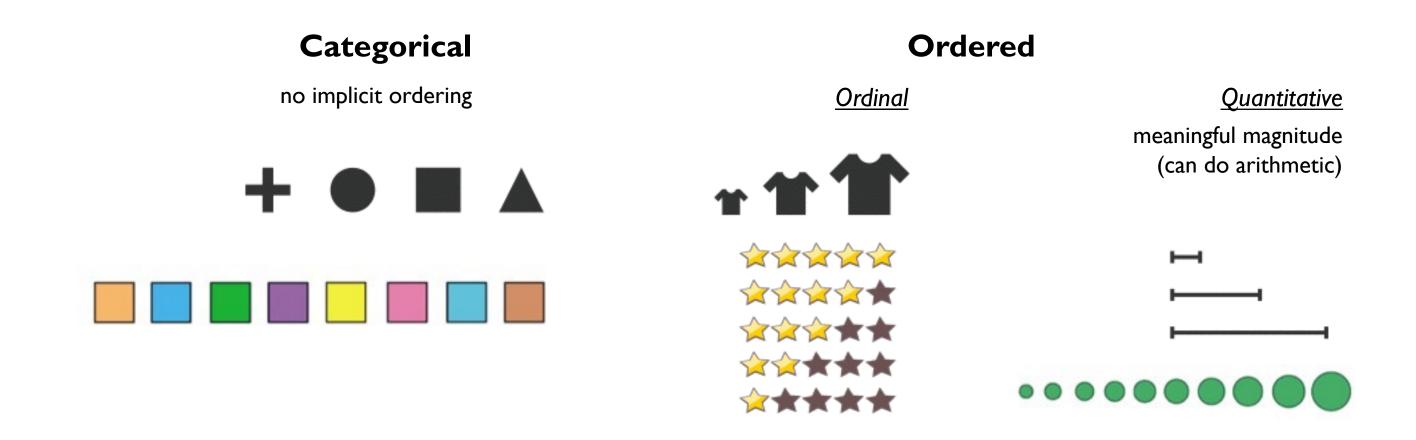


#### **Ordered**

<u>Quantitative</u>

meaningful magnitude (can do arithmetic)





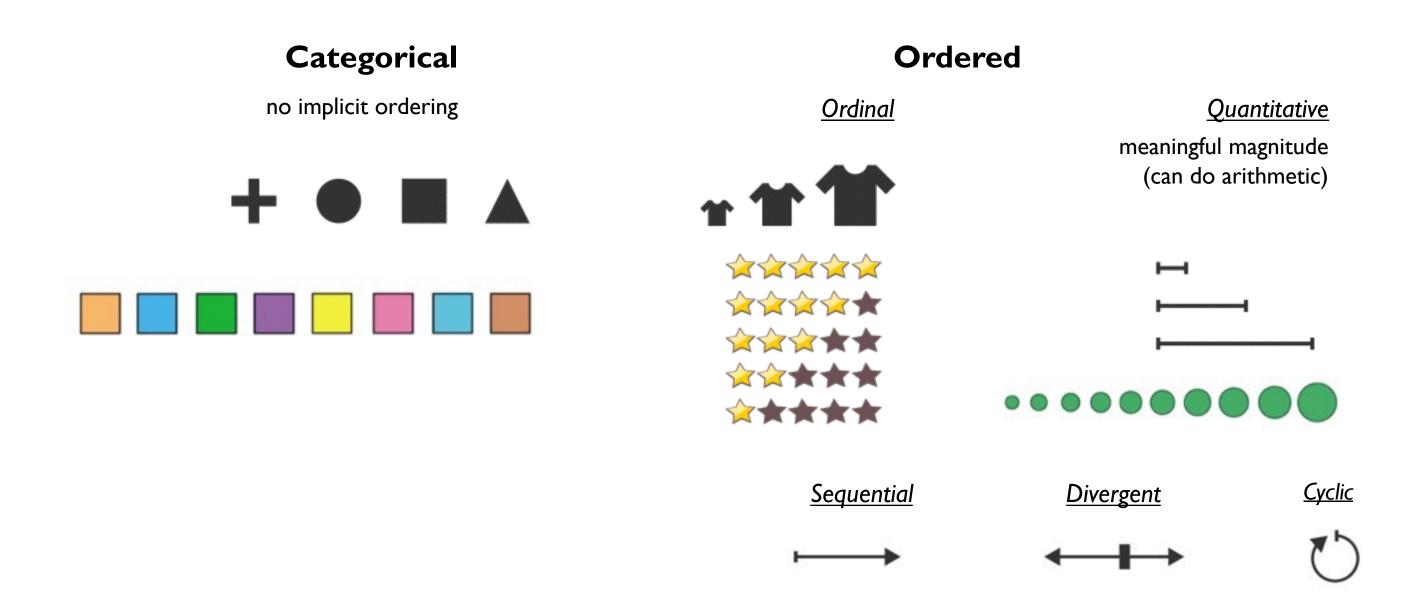


6         2/21/08         4-Not Specified         Small Pack         0.55         2/22/0           32         7/16/07         2-High         Small Pack         0.79         7/17/0           32         7/16/07         2-High         Jumbo Box         0.6         7/18/0           32         7/16/07         2-High         Medium Box         0.65         7/18/0           32         7/16/07         2-High         Medium Box         0.65         7/18/0           35         10/23/07         4-Not Specified         Wrap Bag         0.52         10/24/0           35         10/23/07         4-Not Specified         Small Box         0.58         10/25/0           36         11/3/07         1-Urgent         Small Box         0.55         11/3/0           65         3/18/07         1-Urgent         Small Box         0.55         11/20/0           66         1/20/05         5-Low         Wrap Bag         0.56         1/20/0           69         6/4/05         4-Not Specified         Small Box         0.44         6/6/0           70         12/18/06         5-Low         Ordinal         0.82         12/23/0           96         4/17/05         2-Hig	Α	В	С		S	T	U
6         2/21/08         4-Not Specified         Small Pack         0.55         2/22/0           32         7/16/07         2-High         Small Pack         0.79         7/17/0           32         7/16/07         2-High         Jumbo Box         0.6         6/18/0           32         7/16/07         2-High         Medium Box         0.65         7/18/0           32         7/16/07         2-High         Medium Box         0.65         7/18/0           35         10/23/07         4-Not Specified         Wrap Bag         0.52         10/24/0           35         10/23/07         4-Not Specified         Small Box         0.58         10/25/0           36         11/3/07         1-Urgent         Small Box         0.55         11/3/0           65         3/18/07         1-Urgent         Small Box         0.55         11/20/0           66         1/20/05         5-Low         Wrap Bag         0.56         1/20/0           69         6/4/05         4-Not Specified         Small Box         0.44         6/6/0           70         12/18/06         5-Low         Ordinal         0.82         12/23/0           96         4/17/05         2-Hig	Order ID	Order Date	Order Priority		<b>Product Container</b>	Product Base Margin	Ship Date
32	3	10/14/06	5-Low		Large Box	0.8	10/21/06
32	6	2/21/08	4-Not Specified		Small Pack	0.55	2/22/08
32	32	7/16/07			Small Pack	0.79	7/17/07
32         7/16/07         2-High         Medium Box         0.65         7/18/0           35         10/23/07         4-Not Specified         Wrap Bag         0.52         10/24/0           35         10/23/07         4-Not Specified         Small Box         0.58         10/25/0           36         11/3/07         1-Urgent         Small Box         0.55         11/3/0           65         3/18/07         1-Urgent         Small Box         0.49         3/19/0           66         1/20/05         5-Low         Wrap Bag         0.56         1/20/0           69         6/4/05         4-Not Specified         Small Bok         0.56         1/20/0           69         6/4/05         4-Not Specified         Small Bok         0.44         6/6/0           70         12/18/06         5-Low         Ordinal         0.82         12/23/0           96         4/17/05         2-High         O.82         12/23/0           97         1/29/06         3-Medium         O.38         1/30/0           130         5/8/08         2-High         Small Box         0.37         5/9/0           130         5/8/08         2-High         Medium Box         0.6	32	7/16/07	2-High		Jumbo Box	0.72	7/17/07
35	32	7/16/07	2-High		Medium Box	0.6	7/18/07
35	32	7/16/07	2-High		Medium Box	0.65	7/18/07
36	35	10/23/07	4-Not Specified		Wrap Bag	0.52	10/24/07
Small Pack   O.49   3/19/0	35	10/23/07			Small Box	0.58	10/25/07
1/20/05   5-Low   Wrap Bag   0.56   1/20/06   6/4/05   4-Not Specified   Carpell Pack   0.44   6/6/07   6/6/06   6/4/05   4-Not Specified   Carpell Pack   0.66   6/6/07   12/18/06   5-Low   0.82   12/23/07   12/18/06   5-Low   0.82   12/23/07   12/29/06   3-Medium   0.38   1/30/07   1/29/06   3-Medium   0.38   1/30/07   1/29   11/19/08   5-Low   0.37   11/28/07   130   5/8/08   2-High   Small Box   0.37   5/9/07   130   5/8/08   2-High   Medium Box   0.38   5/10/07   130   5/8/08   2-High   Medium Box   0.38   5/10/07   132   6/11/06   3-Medium   Medium Box   0.66   6/12/07   132   6/11/06   3-Medium   Medium Box   0.69   6/14/07   134   5/1/08   4-Not Specified   Large Box   0.82   5/3/07   135   10/21/07   4-Not Specified   Small Box   0.55   9/14/07   193   8/8/06   1-Urgent   Medium Box   0.57   8/10/07   193	36	11/3/07	1-Urgent		Small Box	0.55	11/3/07
69 6/4/05 4-Not Specified quantitative 0.6 6/6/0 70 12/18/06 5-Low 70 12/18/06 5-Low 96 4/17/05 2-High 97 1/29/06 3-Medium 129 11/19/08 5-Low 11/19/08 5-Low 129 11/19/08 5-Low 130 5/8/08 2-High 97 130 5/8/08 2-High 98 180x 10.37 5/9/08 130 5/8/08 2-High 1	65	3/18/07	1-Urgent		Small Pack	0.49	3/19/07
69 6/4/05 4-Not Spec 70 12/18/06 5-Low 70 12/18/06 5-Low 96 4/17/05 2-High 97 1/29/06 3-Medium 129 11/19/08 5-Low 130 5/8/08 2-High Small Box 130 5/8/08 2-High Medium Box 130 5/8/08 2-High Medium Box 130 5/8/08 2-High Small Box 130 5/8/08 2-High Medium Box 130 5/8/08 2-High Small Box 130 5/8/08 2-High Small Box 130 5/8/08 2-High Medium Box 131 5/8/08 2-High Small Box 132 6/11/06 3-Medium Medium Box 133 6/11/06 3-Medium Jumbo Box 134 5/1/08 4-Not Specified Large Box 135 10/21/07 4-Not Specified Small Pack 166 9/12/07 2-High Small Box 193 8/8/06 1-Urgent Medium Box 194 0.66 6/6/0 0.59 12/23/0 0.38 1/30/0 0.37 5/9/0 0.37 5/9/0 0.38 5/10/0 0.6 5/11/0 0.6 6/12/0 0.6 6/12/0	66	1/20/05	5-Low		Wrap Bag	0.56	1/20/05
70 12/18/06 5-Low 96 4/17/05 2-High 97 1/29/06 3-Medium 129 11/19/08 5-Low 130 5/8/08 2-High 130 5/8/0	69	6/4/05	4-Not Specified			0.44	6/6/05
70 12/18/06 5-Low 96 4/17/05 2-High 97 1/29/06 3-Medium 129 11/19/08 5-Low 130 5/8/08 2-High 130 5/8/0	69	6/4/05	4-Not Spec	ดเเลเ	ntitative	0.6	6/6/05
96 4/17/05 2-High 97 1/29/06 3-Medium 129 11/19/08 5-Low 130 5/8/08 2-High Small Box 130 5/8/08 2-High Medium Box 130 5/8/08 2-High Small Box 130 5/8/08 2-High Medium Box 130 5/8/08 2-High Medium Box 130 5/8/08 2-High Small Box 130 5/8/08 2-High Small Box 131 5/8/08 2-High Small Box 132 6/11/06 3-Medium Medium Box 133 6/11/06 3-Medium Jumbo Box 134 5/1/08 4-Not Specified Large Box 135 10/21/07 4-Not Specified Small Pack 136 9/12/07 2-High Small Box 193 8/8/06 1-Urgent Medium Box 10.55 4/19/0 1.36 0.38 1/30/0 1.37 0.37 0.37 0.37 0.38 0.37 0.38 0.37 0.38 0.39 0.39 0.39 0.39 0.39 0.39 0.39 0.39	70	12/18/06	5-Low			0.59	12/23/06
96 4/17/05 2-High 97 1/29/06 3-Medium 129 11/19/08 5-Low 130 5/8/08 2-High Small Box 130 5/8/08 2-High Medium Box 130 5/8/08 2-High Small Box 130 5/8/08 2-High Medium Box 130 5/8/08 2-High Medium Box 130 5/8/08 2-High Medium Box 130 5/8/08 2-High Small Box 131 5/8/08 2-High Small Box 132 6/11/06 3-Medium Medium Box 133 6/11/06 3-Medium Jumbo Box 134 5/1/08 4-Not Specified Large Box 135 10/21/07 4-Not Specified Small Pack 136 9/12/07 2-High Small Box 193 8/8/06 1-Urgent Medium Box 10.55 4/19/0 1.36 0.38 1/30/0 1.37 11/28/0 0.38 1/30/0	70	12/18/06	5-Low			0.82	12/23/06
129       11/19/08       5-Low       0.37       11/28/0         130       5/8/08       2-High       Small Box       0.37       5/9/0         130       5/8/08       2-High       Medium Box       0.38       5/10/0         130       5/8/08       2-High       Small Box       0.6       5/11/0         132       6/11/06       3-Medium       Medium Box       0.6       6/12/0         132       6/11/06       3-Medium       Jumbo Box       0.69       6/14/0         134       5/1/08       4-Not Specified       Large Box       0.82       5/3/0         135       10/21/07       4-Not Specified       Small Pack       0.64       10/23/0         166       9/12/07       2-High       Small Box       0.55       9/14/0         193       8/8/06       1-Urgent       Medium Box       0.57       8/10/0	96	4/17/05	2-High			0.55	4/19/05
129       11/19/08       5-Low       0.37       11/28/0         130       5/8/08       2-High       Small Box       0.37       5/9/0         130       5/8/08       2-High       Medium Box       0.38       5/10/0         130       5/8/08       2-High       Small Box       0.6       5/11/0         132       6/11/06       3-Medium       Medium Box       0.6       6/12/0         132       6/11/06       3-Medium       Jumbo Box       0.69       6/14/0         134       5/1/08       4-Not Specified       Large Box       0.82       5/3/0         135       10/21/07       4-Not Specified       Small Pack       0.64       10/23/0         166       9/12/07       2-High       Small Box       0.55       9/14/0         193       8/8/06       1-Urgent       Medium Box       0.57       8/10/0	97	1/29/06	3-Medium			0.38	1/30/06
130       5/8/08       2-High       Medium Box       0.38       5/10/0         130       5/8/08       2-High       Small Box       0.6       5/11/0         132       6/11/06       3-Medium       Medium Box       0.6       6/12/0         132       6/11/06       3-Medium       Jumbo Box       0.69       6/14/0         134       5/1/08       4-Not Specified       Large Box       0.82       5/3/0         135       10/21/07       4-Not Specified       Small Pack       0.64       10/23/0         166       9/12/07       2-High       Small Box       0.55       9/14/0         193       8/8/06       1-Urgent       Medium Box       0.57       8/10/0	129	11/19/08	5-Low		0	0.37	11/28/08
130       5/8/08       2-High       Small Box       0.6       5/11/0         132       6/11/06       3-Medium       Medium Box       0.6       6/12/0         132       6/11/06       3-Medium       Jumbo Box       0.69       6/14/0         134       5/1/08       4-Not Specified       Large Box       0.82       5/3/0         135       10/21/07       4-Not Specified       Small Pack       0.64       10/23/0         166       9/12/07       2-High       Small Box       0.55       9/14/0         193       8/8/06       1-Urgent       Medium Box       0.57       8/10/0	130	5/8/08	2-High		Small Box	0.37	5/9/08
132       6/11/06       3-Medium       Medium Box       0.6       6/12/0         132       6/11/06       3-Medium       Jumbo Box       0.69       6/14/0         134       5/1/08       4-Not Specified       Large Box       0.82       5/3/0         135       10/21/07       4-Not Specified       Small Pack       0.64       10/23/0         166       9/12/07       2-High       Small Box       0.55       9/14/0         193       8/8/06       1-Urgent       Medium Box       0.57       8/10/0	130	5/8/08			Medium Box	0.38	5/10/08
132       6/11/06       3-Medium       Jumbo Box       0.69       6/14/0         134       5/1/08       4-Not Specified       Large Box       0.82       5/3/0         135       10/21/07       4-Not Specified       Small Pack       0.64       10/23/0         166       9/12/07       2-High       Small Box       0.55       9/14/0         193       8/8/06       1-Urgent       Medium Box       0.57       8/10/0	130	5/8/08			Small Box	0.6	5/11/08
134       5/1/08       4-Not Specified       Large Box       0.82       5/3/0         135       10/21/07       4-Not Specified       Small Pack       0.64       10/23/0         166       9/12/07       2-High       Small Box       0.55       9/14/0         193       8/8/06       1-Urgent       Medium Box       0.57       8/10/0	132	6/11/06			Medium Box	0.6	6/12/06
135       10/21/07 4-Not Specified       Small Pack       0.64       10/23/0         166       9/12/07 2-High       Small Box       0.55       9/14/0         193       8/8/06 1-Urgent       Medium Box       0.57       8/10/0	132	6/11/06	3-Medium		Jumbo Box	0.69	6/14/06
166       9/12/07       2-High       Small Box       0.55       9/14/0         193       8/8/06       1-Urgent       Medium Box       0.57       8/10/0	134	5/1/08	4-Not Specified		Large Box	0.82	5/3/08
193 8/8/06 1-Urgent Medium Box 0.57 8/10/0	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
194 4/5/08 3-Medium Wrap Bag 0.42 4/7/0	193	8/8/06	1-Urgent		Medium Box	0.57	8/10/06
To the state of th	194	4/5/08	3-Medium		Wrap Bag	0.42	4/7/08



Α	В	С		S	Т	U
Order ID	Order Date	Order Priority		<b>Product Container</b>	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		4-Not Speci		Small Pack	0.44	6/6/05
69	6/4/05	4-Not Spec	anar	ntitative	0.6	6/6/05
70	12/18/06	5-Low	_		0.59	12/23/06
70	12/18/06	5-Low	ordi	nal	0.82	12/23/06
96	4/17/05	2-High			0.55	4/19/05
97	1/29/06	3-Medium	cates	gorical	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

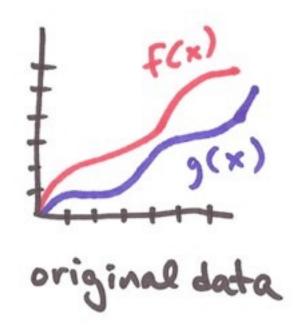


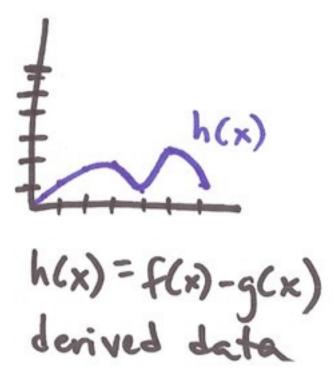




#### DERIVED ATTRIBUTES

- derived attribute: computed from originals
  - simple change of type
  - acquire additional data
  - complex transformation
  - transformation is abstraction choice







#### TYPE vs SEMANTIC

- data model: mathematical abstraction (data abstraction)
  - set with operations, e.g., floats with \* / +
- conceptual model: mental construction (semantics)
  - includes semantics, supports reasoning
- conceptual model motivates data abstraction choices



#### EXAMPLE

- from data model . . .
  - - 32.52, 54.06, -17.35, . . . (floats)
- using conceptual model . . .
  - temperature
- to new data abstraction...
  - continuous to 2 significant figures (quantitative)
  - hot, warm, cold (ordinal)
  - above freezing, below freezing (categorical)



