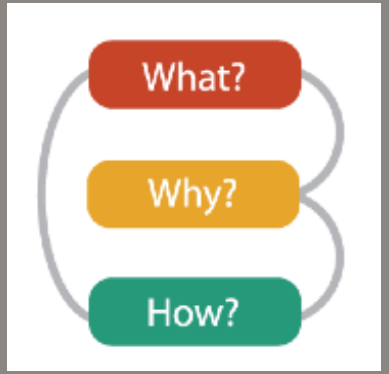
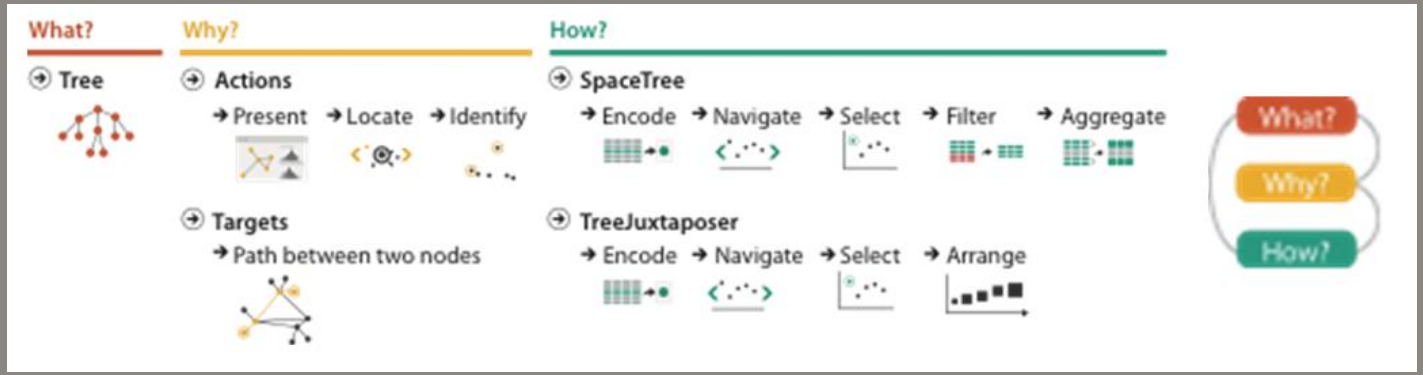
Foundation:
What, Why and How

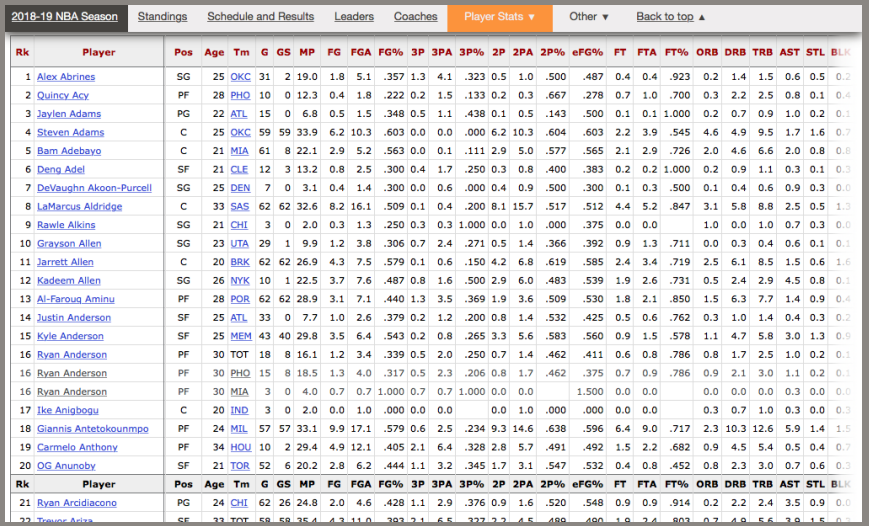
What, Why and How

- What is shown?
  - Data abstraction
- Why is the user looking at it?
  - Task abstraction
- How is it shown?
  - Idiom: visual encoding and interaction



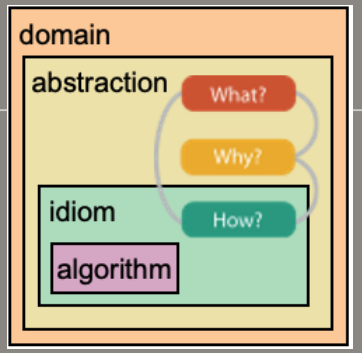


Nested Model: Four Levels of Vis Design

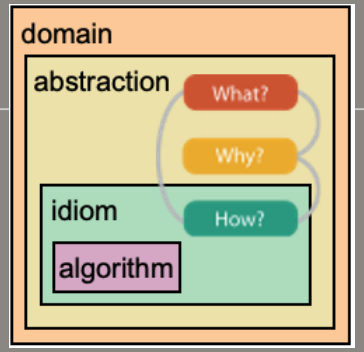


2018-19 NBA Season | Standings | Schedule and Results | Leaders | Coaches | Player Stats ▾ | Other ▾ | Back to top ▲

Rk	Player	Pos	Age	Tm	G	GS	MP	FG	FGA	FG%	3P	3PA	3P%	2P	2PA	2P%	eFG%	FT	FTA	FT%	ORB	DRB	TRB	AST	STL	BLK
1	Alex Abrines	SG	25	OKC	31	2	19.0	1.8	5.1	.357	1.3	4.1	.323	0.5	1.0	.500	.487	0.4	0.4	.923	0.2	1.4	1.5	0.6	0.5	0.2
2	Quincy Acy	PF	28	PHO	10	0	12.3	0.4	1.8	.222	0.2	1.5	.133	0.2	0.3	.667	.278	0.7	1.0	.700	0.3	2.2	2.5	0.8	0.1	0.4
3	Jaylen Adams	PG	22	ATL	15	0	6.8	0.5	1.5	.348	0.5	1.1	.438	0.1	0.5	.143	.500	0.1	0.1	1.000	0.2	0.7	0.9	1.0	0.2	0.1
4	Steven Adams	C	25	OKC	59	59	33.9	6.2	10.3	.603	0.0	0.0	.000	6.2	10.3	.604	.603	2.2	3.9	.545	4.6	4.9	9.5	1.7	1.6	0.7
5	Bam Adebayo	C	21	MIA	61	8	22.1	2.9	5.2	.563	0.0	0.1	.111	2.9	5.0	.577	.565	2.1	2.9	.726	2.0	4.6	6.6	2.0	0.8	0.8
6	Deng Adel	SF	21	CLE	12	3	13.2	0.8	2.5	.300	0.4	1.7	.250	0.3	0.8	.400	.383	0.2	0.2	1.000	0.2	0.9	1.1	0.3	0.1	0.3
7	DeVaughn Akoon-Purcell	SG	25	DEN	7	0	3.1	0.4	1.4	.300	0.0	0.6	.000	0.4	0.9	.500	.300	0.1	0.3	.500	0.1	0.4	0.6	0.9	0.3	0.0
8	LaMarcus Aldridge	C	33	SAS	62	62	32.6	8.2	16.1	.509	0.1	0.4	.200	8.1	15.7	.517	.512	4.4	5.2	.847	3.1	5.8	8.8	2.5	0.5	1.3
9	Rawle Alkins	SG	21	CHI	3	0	2.0	0.3	1.3	.250	0.3	0.3	1.000	0.0	1.0	.000	.375	0.0	0.0		1.0	0.0	1.0	0.7	0.3	0.0
10	Grayson Allen	SG	23	UTA	29	1	9.9	1.2	3.8	.306	0.7	2.4	.271	0.5	1.4	.366	.392	0.9	1.3	.711	0.0	0.3	0.4	0.6	0.1	0.1
11	Jarrett Allen	C	20	BRK	62	62	26.9	4.3	7.5	.579	0.1	0.6	.150	4.2	6.8	.619	.585	2.4	3.4	.719	2.5	6.1	8.5	1.5	0.6	1.6
12	Kadeem Allen	SG	26	NYK	10	1	22.5	3.7	7.6	.487	0.8	1.6	.500	2.9	6.0	.483	.539	1.9	2.6	.731	0.5	2.4	2.9	4.5	0.8	0.1
13	Al-Farouq Aminu	PF	28	POR	62	62	28.9	3.1	7.1	.440	1.3	3.5	.369	1.9	3.6	.509	.530	1.8	2.1	.850	1.5	6.3	7.7	1.4	0.9	0.4
14	Justin Anderson	SF	25	ATL	33	0	7.7	1.0	2.6	.379	0.2	1.2	.200	0.8	1.4	.532	.425	0.5	0.6	.762	0.3	1.0	1.4	0.4	0.3	0.2
15	Kyle Anderson	SF	25	MEM	43	40	29.8	3.5	6.4	.543	0.2	0.8	.265	3.3	5.6	.583	.560	0.9	1.5	.578	1.1	4.7	5.8	3.0	1.3	0.9
16	Ryan Anderson	PF	30	TOT	18	8	16.1	1.2	3.4	.339	0.5	2.0	.250	0.7	1.4	.462	.411	0.6	0.8	.786	0.8	1.7	2.5	1.0	0.2	0.1
16	Ryan Anderson	PF	30	PHO	15	8	18.5	1.3	4.0	.317	0.5	2.3	.206	0.8	1.7	.462	.375	0.7	0.9	.786	0.9	2.1	3.0	1.1	0.2	0.1
16	Ryan Anderson	PF	30	MIA	3	0	4.0	0.7	0.7	1.000	0.7	0.7	1.000	0.0	0.0		1.500	0.0	0.0		0.0	0.0	0.0	0.3	0.0	0.0
17	Ike Anigbogu	C	20	IND	3	0	2.0	0.0	1.0	.000	0.0	0.0		0.0	1.0	.000	.000	0.0	0.0		0.3	0.7	1.0	0.3	0.0	0.3
18	Giannis Antetokounmpo	PF	24	MIL	57	57	33.1	9.9	17.1	.579	0.6	2.5	.234	9.3	14.6	.638	.596	6.4	9.0	.717	2.3	10.3	12.6	5.9	1.4	1.5
19	Carmelo Anthony	PF	34	HOU	10	2	29.4	4.9	12.1	.405	2.1	6.4	.328	2.8	5.7	.491	.492	1.5	2.2	.682	0.9	4.5	5.4	0.5	0.4	0.7
20	OG Anunoby	SF	21	TOR	52	6	20.2	2.8	6.2	.444	1.1	3.2	.345	1.7	3.1	.547	.532	0.4	0.8	.452	0.8	2.3	3.0	0.7	0.6	0.3
Rk	Player	Pos	Age	Tm	G	GS	MP	FG	FGA	FG%	3P	3PA	3P%	2P	2PA	2P%	eFG%	FT	FTA	FT%	ORB	DRB	TRB	AST	STL	BLK
21	Ryan Arcidiacono	PG	24	CHI	62	26	24.8	2.0	4.6	.428	1.1	2.9	.376	0.9	1.6	.520	.548	0.9	0.9	.914	0.2	2.2	2.4	3.5	0.9	0.0
22	Trevor Ariza	SF	33	TOT	58	58	35.4	4.3	11.0	.393	2.1	6.5	.327	2.2	4.5	.489	.490	1.9	2.4	.803	0.7	4.9	5.6	3.9	1.5	0.9



Nested Model: Four Levels of Vis Design



- Domain situation
  - Who are the target user?

The weakness of players?

Can starters' skills match well?

coach

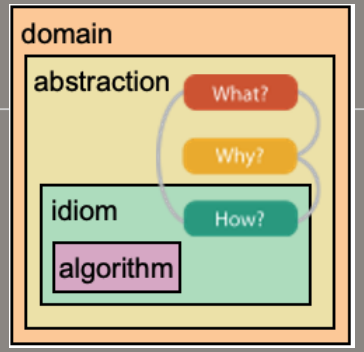
basketball data

commentator

Observe whether a team's win-loss ratio change a lot after all-star week

The best FG% of a players by shot position

Nested Model: Four Levels of Vis Design



- Domain situation
  - Who are the target user?
- Abstraction
  - Translate from specifics of domain to vocabulary of vis
    - What is shown? **data** abstraction
    - Why is the user looking at it? **task** abstraction

Rank	AreaName	St1	Population 7/1/2010	Population 7/1/2015	Change 2010-15	%Change 2010-15
1	New York city	NY	8,192,426	8,550,405	357,979	4.4
2	Los Angeles city	CA	3,796,575	3,971,883	175,308	4.6
3	Chicago city	IL	2,697,650	2,720,546	22,896	0.8
4	Houston city	TX	2,114,761	2,296,224	181,463	8.6
5	Philadelphia city	PA	1,528,338	1,567,442	39,104	2.6
6	Phoenix city	AZ	1,450,267	1,563,025	112,758	7.8
7	San Antonio city	TX	1,333,953	1,469,845	135,892	10.2
8	San Diego city	CA	1,306,080	1,394,928	88,848	6.8
9	Dallas city	TX	1,200,699	1,300,092	99,393	8.3
10	San Jose city	CA	955,399	1,026,908	71,509	7.5

Non-order order

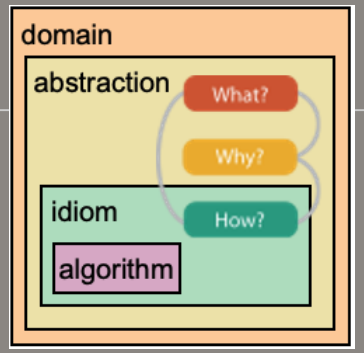
Rk	Player	Pos	Age	Tm	G	GS	MP	FG	FGA	FG%	3P
1	Alex Abrines	SG	25	OKC	31	2	19.0	1.8	5.1	.357	1.3
2	Quincy Acy	PF	28	PHO	10	0	12.3	0.4	1.8	.222	0.2
3	Jaylen Adams	PG	22	ATL	15	0	6.8	0.5	1.5	.348	0.5
4	Steven Adams	C	25	OKC	59	59	33.9	6.2	10.3	.603	0.0
5	Bam Adebayo	C	21	MIA	61	8	22.1	2.9	5.2	.563	0.0
6	Deng Adel	SF	21	CLE	12	3	13.2	0.8	2.5	.300	0.4
7	DeVaughn Akoon-Purcell	SG	25	DEN	7	0	3.1	0.4	1.4	.300	0.0
8	LaMarcus Aldridge	C	33	SAS	62	62	32.6	8.2	16.1	.509	0.1
9	Rawle Alkins	SG	21	CHI	3	0	2.0	0.3	1.3	.250	0.3
10	Grayson Allen	SG	23	UTA	29	1	9.9	1.2	3.8	.306	0.7
11	Jarrett Allen	C	20	BRK	62	62	26.9	4.3	7.5	.579	0.1
12	Kadeem Allen	SG	26	NYK	10	1	22.5	3.7	7.6	.487	0.8
13	Al-Farouq Aminu	PF	28	POR	62	62	28.9	3.1	7.1	.440	1.3
14	Justin Anderson	SF	25	ATL	33	0	7.7	1.0	2.6	.379	0.2
15	Kyle Anderson	SF	25	MEM	43	40	29.8	3.5	6.4	.540	0.2

Non-order order

"State name" and "position" are both non-order data. We might be able to use similar way to visualize them

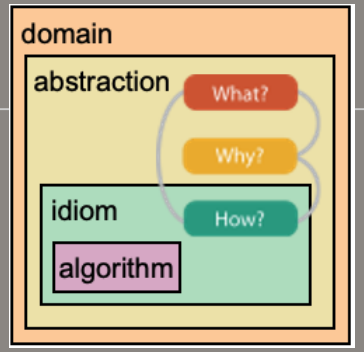
Understand the abstraction of data could help us to pick up better way to visualize the data

Nested Model: Four Levels of Vis Design



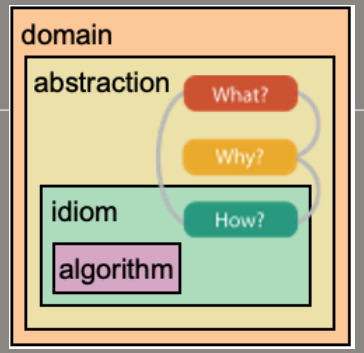
- Domain situation
  - Who are the target user?
- Abstraction
  - Translate from specifics of domain to vocabulary of vis
    - What is shown? **data** abstraction
    - Why is the user looking at it? **task** abstraction
- Idiom (approach)
  - How is it shown?
    - **Visual encoding** idiom: how to draw the picture
    - **Interaction** idiom: how to manipulate the picture

Nested Model: Four Levels of Vis Design



- Domain situation
  - Who are the target user?
- Abstraction
  - Translate from specifics of domain to vocabulary of vis
    - What is shown? **data** abstraction
    - Why is the user looking at it? **task** abstraction
- Idiom (approach)
  - How is it shown?
    - **Visual encoding** idiom: how to draw the picture
    - **Interaction** idiom: how to manipulate the picture
- Algorithm
  - Efficient computation

Nested Model: Four Levels of Vis Design



- Domain situation
  - Who are the target user?
- Abstraction
  - Translate from specifics of domain to vocabulary of vis
    - What is shown? **data** abstraction
    - Why is the user looking at it? **task** abstraction
- Idiom (approach)
  - How is it shown?
    - **Visual encoding** idiom: how to draw the picture
    - **Interaction** idiom: how to manipulate the picture
- Algorithm
  - Efficient computation

Value to think about these four level differently: help us to debug when the visual design fails

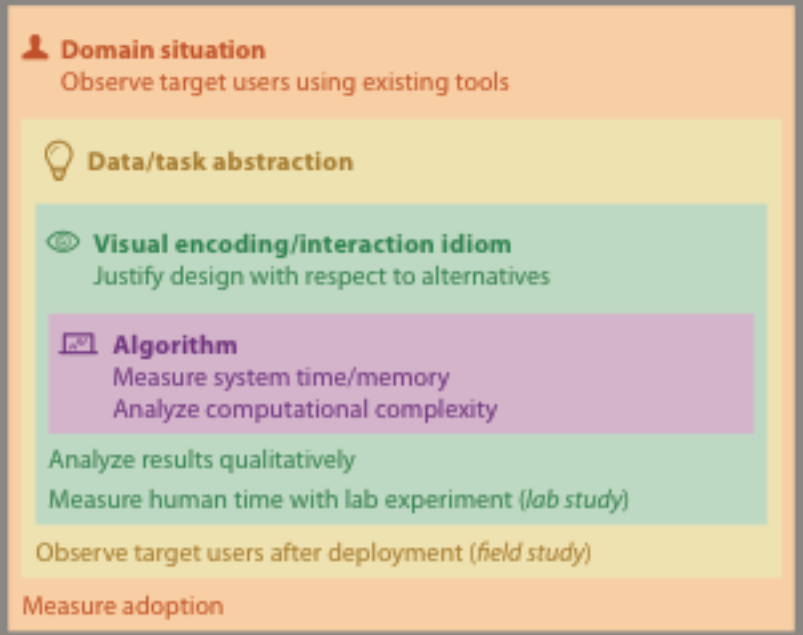
Why is Validation Difficult?

- After the visualization is designed and implemented, we would like to evaluate it. However, it is not trivial.
  - Different ways to get the vis wrong/ineffective at each level

Why is Validation Difficult?

- Solution: use methods from different fields at each level

Computer science



Technique-driven work. e.g. scalability

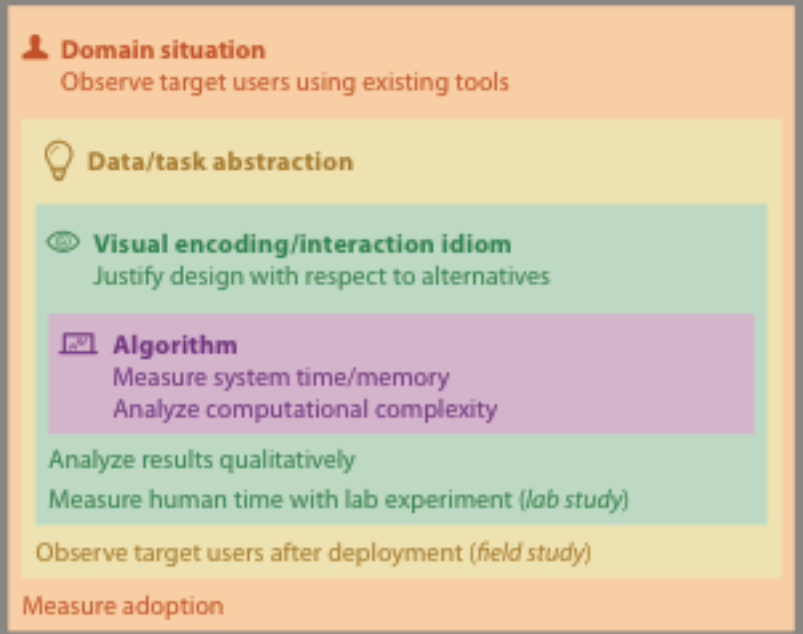
Why is Validation Difficult?

- Solution: use methods from different fields at each level

Design

Computer science

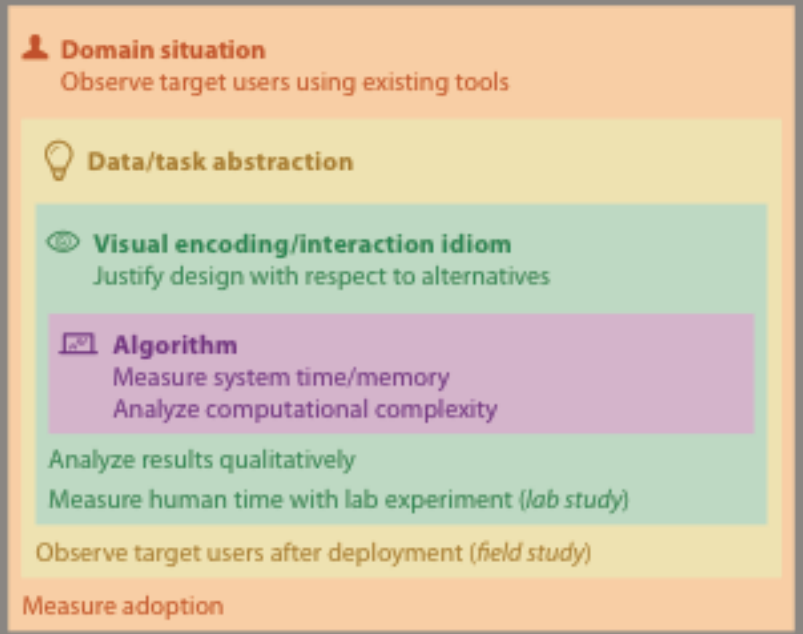
Cognitive psychology



Technique-driven work. e.g. scalability

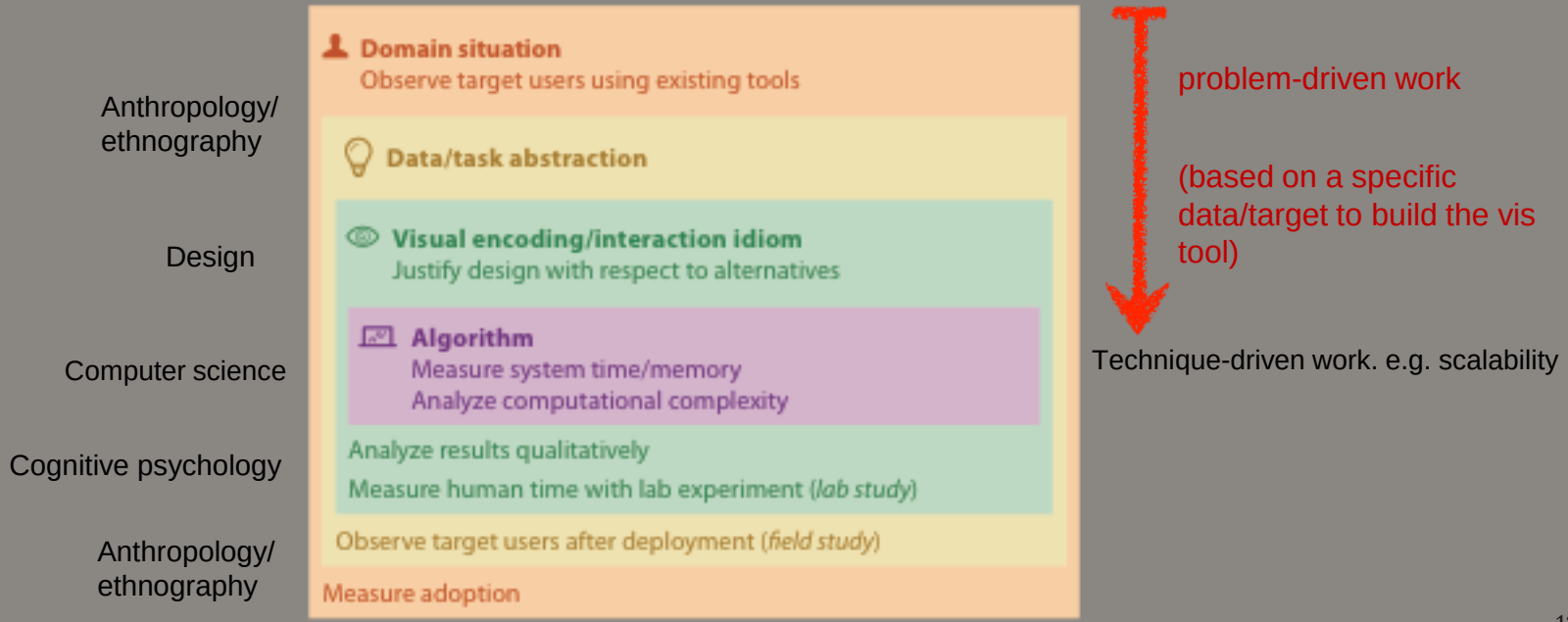
Why is Validation Difficult?

- Solution: use methods from different fields at each level



Why is Validation Difficult?

- Solution: use methods from different fields at each level



Abstraction for Data, Action, Target

- Easy for vis designer to choose effective design
  - The space of the design choice is too huge
  - Abstraction helps vis designer to ignore most of the improper vis design choices (improper color encoding, style of plots...)
- Different domain data/questions could be abstracted to similar vis design language
  - When does my salary increase most?
  - Which value of the parameter could have the most impact of classification accuracy?
  - Abstraction: need to observe the **pattern** of **one attribute** according to change of **another attribute**
  - Pie chart(x), heap map(x), line chart(o), bar chart(?)

What?

Datasets

Data Types

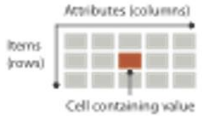
→ Items → Attributes → Links → Positions → Grids

Data and Dataset Types

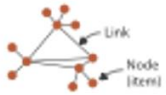
Tables	Networks & Trees	Fields	Geometry	Clusters, Sets, Lists
Items	Items (nodes)	Grids	Items	Items
Attributes	Links	Positions	Positions	
	Attributes	Attributes		

Dataset Types

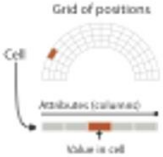
→ Tables



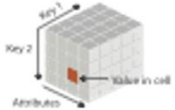
→ Networks



→ Fields (Continuous)



→ *Multidimensional Table*



→ *Trees*

→ Geometry (Spatial)

Dataset Availability

→ Static

→ Dynamic

Attributes

Attribute Types

→ Categorical

→ Ordered

→ *Ordinal*

→ *Quantitative*

Ordering Direction

→ Sequential

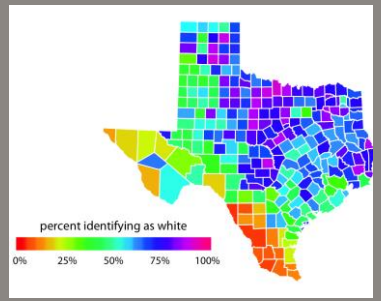
→ Diverging

→ Cyclic

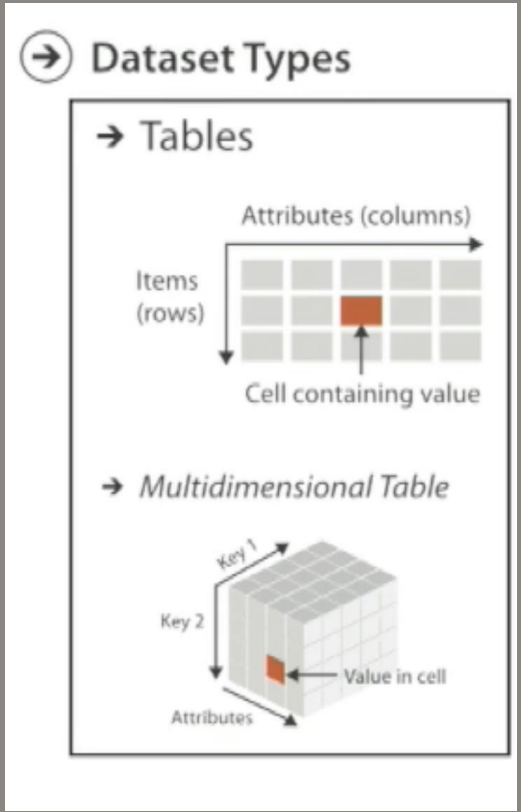
What – Data Abstraction

- Only a few visual design fit a specific type or attribute well
  - Reduce the design search space
  - Easy to know something wrong in the visual design

- Example:
  - Student score is “ordered” attribute:
    - what visual designs are the candidates? Then make the design decision
  - Department name is “non-ordered” attribute:
    - what visual designs are the candidates? Then make the design decision



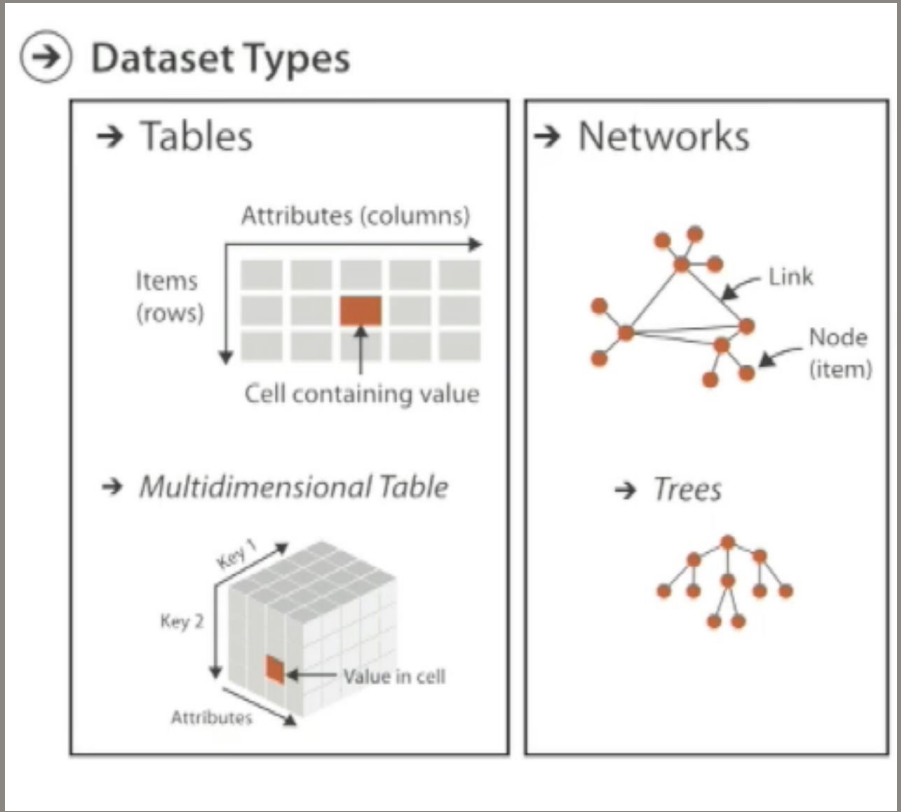
Three Major Dataset Types



Code	Quiz 12	Quiz 13	Quiz 14	Quiz 15	Quiz 16	Quiz 17	Quiz 18	Quiz 19	Quiz 20	Quiz 21	Quiz 22	Quiz 23	Quiz 24	Quiz Total	Drop 5	Quiz %	Quiz	Exam I	Exam II	Final	Term Paper	Total Grade	
	10	10	10	10	10	10	10	10	10	10	10	10	10	240	190	100.0%	100.0	100%	100%	100%	100.0%	100.0%	Grade
030767	10	4	0	10	0	0	8	1	10	10	9	10	5	166	163	85.8%	85.8	88%	82%	87%	96.5%	87.7%	B+
11291129	1	10	0	6	9	10	6	2	10	8	9	9	10	155	152	80.0%	80.0	74%	87%	71%	95.0%	82.4%	B
12345	10	10	10	10	10	10	4	10	10	10	5	9	10	211	187	98.4%	98.4	94%	98%	89%	100.0%	96.0%	A
2182	0	10	9	10	3	10	8	3	10	6	0	10	10	164	158	83.2%	83.2	88%	94%	99%	98.0%	93.4%	A
2663	10	10	10	0	6	10	10	8	10	9	3	10	10	200	185	97.4%	97.4	97%	71%	86%	98.0%	91.2%	A-
3196	10	5	10	10	7	10	8	8	10	10	10	10	10	203	181	95.3%	95.3	88%	90%	93%	100.0%	92.5%	A
3885	10	10	10	10	8	8	10	10	10	10	10	10	10	230	190	100.0%	100.0	95%	101%	110%	95.0%	100.5%	A
4598379	10	10	10	10	9	8	0	10	10	10	4	5	10	193	179	94.2%	94.2	87%	102%	107%	98.0%	98.5%	A
623562	10	10	10	10	10	10	10	10	10	10	7	10	10	208	190	100.0%	100.0	81%	98%	100%	100.0%	96.3%	A
8765309	10	5	0	10	7	10	4	10	10	8	4	5	5	164	154	81.1%	81.1	80%	88%	91%	96.5%	86.7%	B
9648	10	3	10	6	5	10	6	1	10	5	4	2	5	148	139	73.2%	73.2	93%	85%	79%	96.5%	85.5%	B
9792	10	1	0	2	4	0	0	4	0	5	2	5	5	73	73	38.4%	38.4	62%	52%	68%	96.5%	66.8%	C
Cayuga	7	10	10	10	10	10	8	10	10	10	8	10	5	215	187	98.4%	98.4	98%	103%	92%	96.5%	96.8%	A
emm9899	10	10	10	10	10	10	6	8	8	10	10	10	10	217	187	98.4%	98.4	96%	98%	100%	95.0%	97.2%	A
Junior	10	10	10	10	10	10	8	10	10	10	10	10	10	236	190	100.0%	100.0	96%	103%	103%	98.0%	100.2%	A
Mako	0	2	11	10	10	8	10	10	10	10	8	5	10	177	173	91.1%	91.1	94%	93%	98%	98.0%	94.7%	A
nonays	10	5	6	1	0	0	8	3	10	6	9	9	10	157	153	80.5%	80.5	81%	68%	68%	99.0%	75.4%	B-
ride	10	3	10	10	8	9	10	0	0	10	3	9	10	172	169	88.9%	88.9	82%	88%	84%	95.0%	87.8%	B+
SW363	4	10	10	10	10	10	10	1	10	10	5	8	10	195	178	93.7%	93.7	90%	101%	97%	100.0%	96.4%	A
Tiger	10	3	0	10	6	8	2	5	10	6	0	0	10	124	124	65.3%	65.3	74%	95%	73%	95.0%	81.9%	B
TURKEYBOY	10	0	9	8	10	10	10	1	10	10	7	10	10	188	179	94.2%	94.2	77%	95%	93%	100.0%	92.3%	A
							5.1																

Usually, it is infovis data

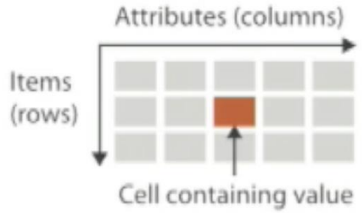
Three Major Datatypes



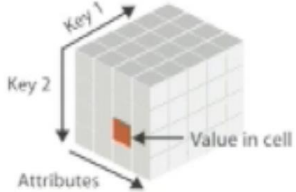
Three Major Datatypes

Dataset Types

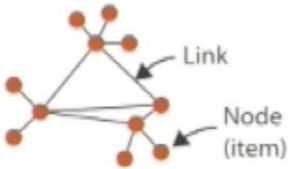
Tables



Multidimensional Table



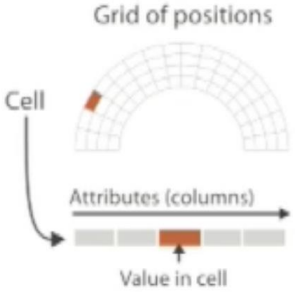
Networks



Trees

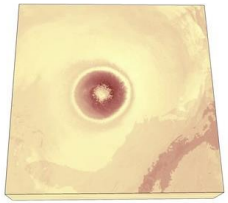
Spatial

Fields (Continuous)



Geometry (Spatial)

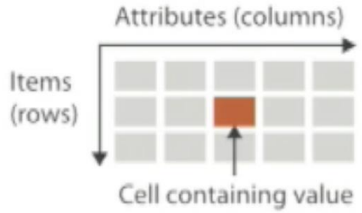
Usually, it is scivis data



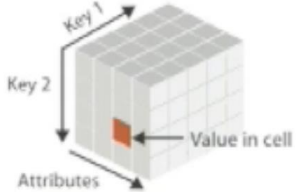
Three Major Datatypes

Dataset Types

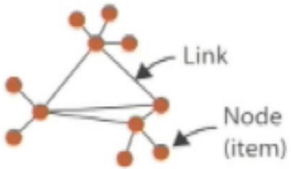
Tables



Multidimensional Table



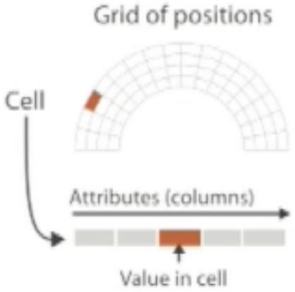
Networks



Trees

Spatial

Fields (Continuous)



Geometry (Spatial)

- Visualization vs computer graphics?
- Usually, people consider visualization is a sub-domain of computer graphics
- Visualization usually requires design decision

Dataset and Data 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

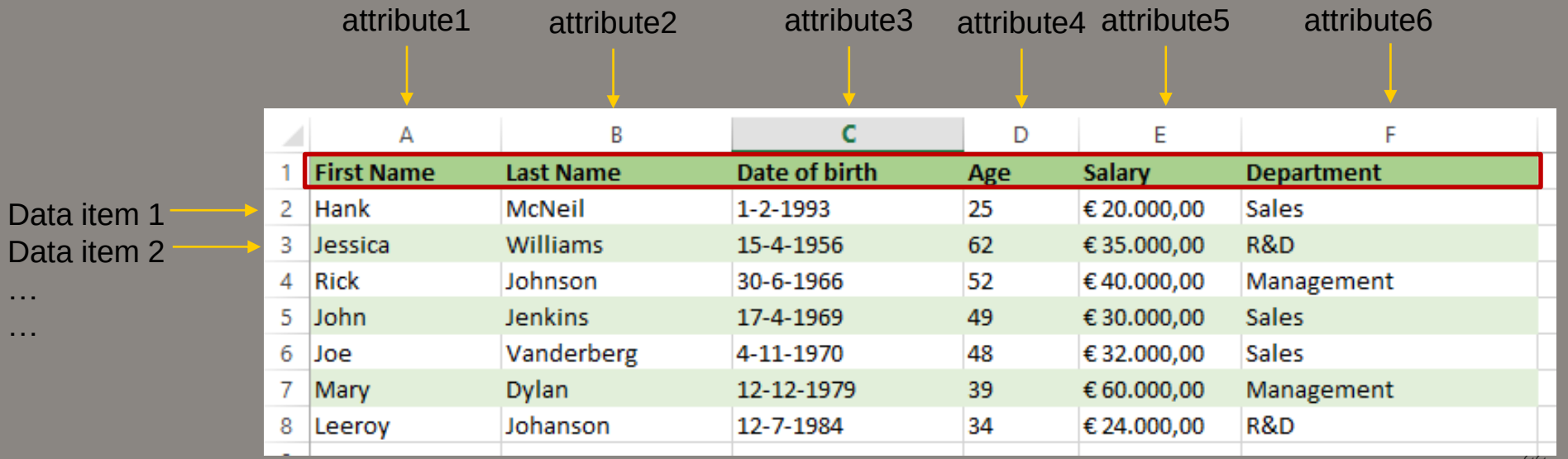
→ Items → Attributes → Links → Positions → Grids

Dataset Availability

→ Static → Dynamic

Attribute Types

- Attribute? (variable)
  - Example 1: spreadsheet

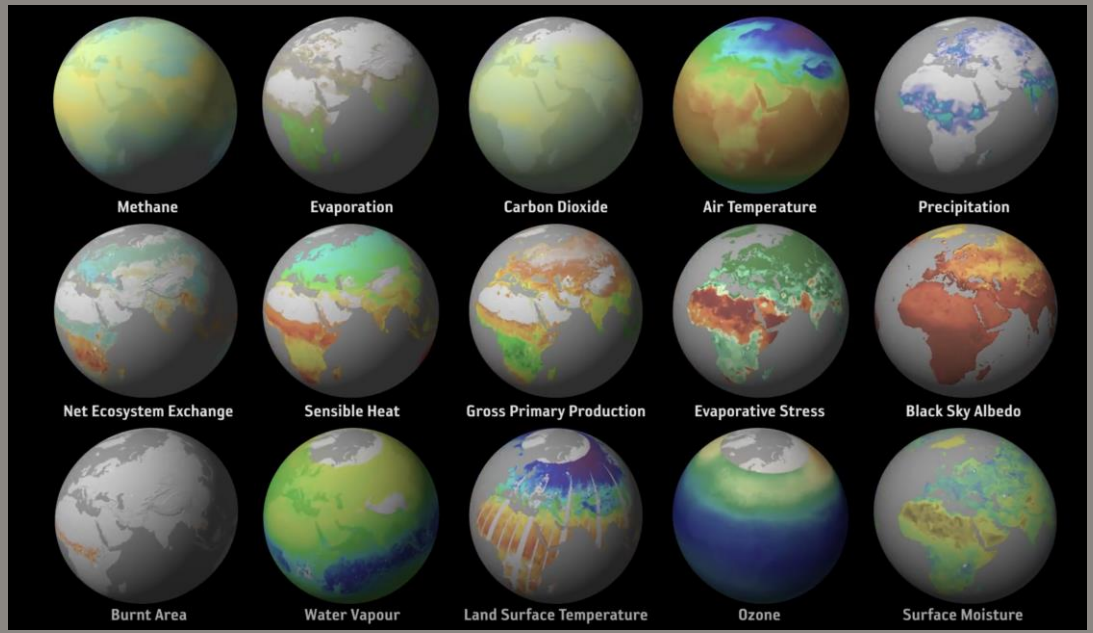


	A	B	C	D	E	F
1	First Name	Last Name	Date of birth	Age	Salary	Department
2	Hank	McNeil	1-2-1993	25	€ 20.000,00	Sales
3	Jessica	Williams	15-4-1956	62	€ 35.000,00	R&D
4	Rick	Johnson	30-6-1966	52	€ 40.000,00	Management
5	John	Jenkins	17-4-1969	49	€ 30.000,00	Sales
6	Joe	Vanderberg	4-11-1970	48	€ 32.000,00	Sales
7	Mary	Dylan	12-12-1979	39	€ 60.000,00	Management
8	Leeroy	Johanson	12-7-1984	34	€ 24.000,00	R&D

S02-01

Attribute Types

- Attribute? (variable)
  - Example 2: spatial dataset



Attribute Types

- Categorical v.s. ordered
  - Can you sort the data item by the attribute?
    - Yes: ordered attribute
    - No: categorical attribute

Attribute Types

Categorical attribute

Ordered attribute

	A	B	C	D	E	F
1	First Name	Last Name	Date of birth	Age	Salary	Department
2	Hank	McNeil	1-2-1993	25	€ 20.000,00	Sales
3	Jessica	Williams	15-4-1956	62	€ 35.000,00	R&D
4	Rick	Johnson	30-6-1966	52	€ 40.000,00	Management
5	John	Jenkins	17-4-1969	49	€ 30.000,00	Sales
6	Joe	Vanderberg	4-11-1970	48	€ 32.000,00	Sales
7	Mary	Dylan	12-12-1979	39	€ 60.000,00	Management
8	Leeroy	Johanson	12-7-1984	34	€ 24.000,00	R&D

Attribute Types

- Ordinal v.s. Quantitative
  - Is the result of subtraction between two values of the attribute meaningful?
    - Yes: quantitative attribute
    - No: ordinal attribute

Attribute Types

Ordered attribute (all quantitative attribute)

	A	B	C	D	E	F
1	First Name	Last Name	Date of birth	Age	Salary	Department
2	Hank	McNeil	1-2-1993	25	€ 20.000,00	Sales
3	Jessica	Williams	15-4-1956	62	€ 35.000,00	R&D
4	Rick	Johnson	30-6-1966	52	€ 40.000,00	Management
5	John	Jenkins	17-4-1969	49	€ 30.000,00	Sales
6	Joe	Vanderberg	4-11-1970	48	€ 32.000,00	Sales
7	Mary	Dylan	12-12-1979	39	€ 60.000,00	Management
8	Leeroy	Johanson	12-7-1984	34	€ 24.000,00	R&D

- Example of **ordinal** attribute
  - Cloth size (S, M, L, XL)
    - They have order
    - S->M->L->XL
  - However, can you calculate XL – M? or is the result meaningful?

Attribute Types

- Sequential attribute: from min to max

	A	B	C	D	E	F
1	First Name	Last Name	Date of birth	Age	Salary	Department
2	Hank	McNeil	1-2-1993	25	€ 20.000,00	Sales
3	Jessica	Williams	15-4-1956	62	€ 35.000,00	R&D
4	Rick	Johnson	30-6-1966	52	€ 40.000,00	Management
5	John	Jenkins	17-4-1969	49	€ 30.000,00	Sales
6	Joe	Vanderberg	4-11-1970	48	€ 32.000,00	Sales
7	Mary	Dylan	12-12-1979	39	€ 60.000,00	Management
8	Leeroy	Johanson	12-7-1984	34	€ 24.000,00	R&D

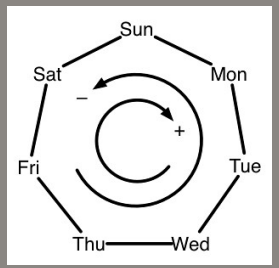
Attribute Types

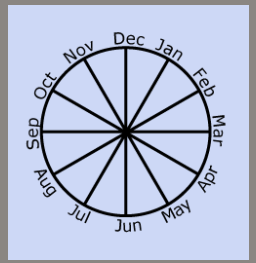
- Sequential attribute: from min to max
- Diverging attribute: can be divided into two sequences

Survey data

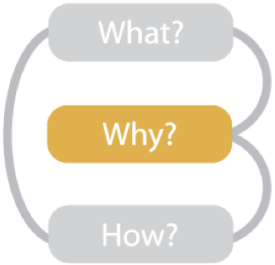
Attribute Types

- Sequential attribute: from min to max
- Diverging attribute: can be divided into two sequences
- Cyclic attribute: values wrap around back to a starting point





S02-02



Why?

Actions

Analyze

→ Consume

→ Discover → Present → Enjoy

→ Produce

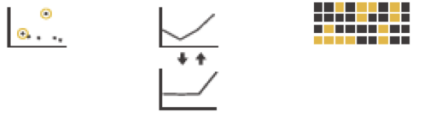
→ Annotate → Record → Derive

Search

	Target known	Target unknown
Location known	Lookup	Browse
Location unknown	Locate	Explore

Query

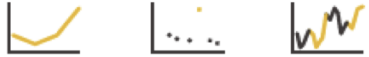
→ Identify → Compare → Summarize



Targets

All Data

→ Trends → Outliers → Features



Attributes

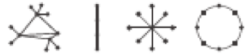
→ One → Many

→ Distribution → Dependency → Correlation → Similarity

→ Extremes

Network Data

→ Topology



→ Paths

Spatial Data

→ Shape

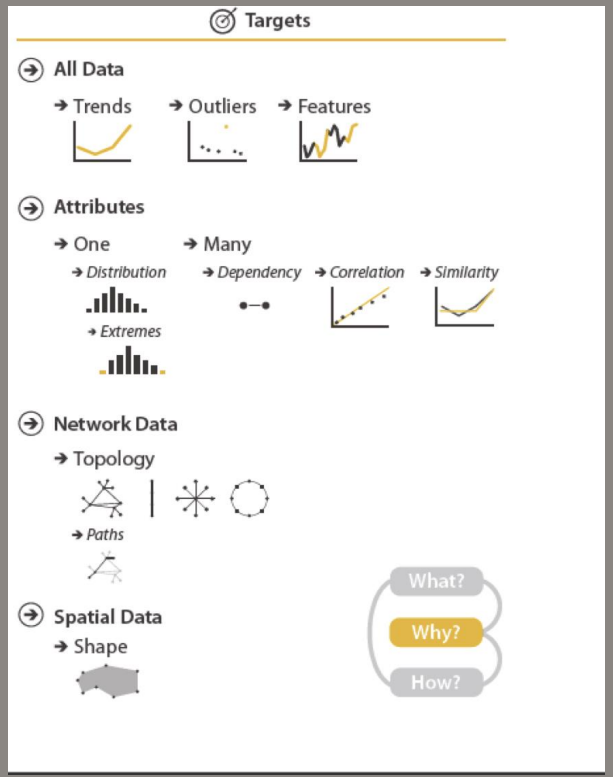
“Why?” - Task Abstract

- Motivation
  - When a domain expert describe his/her “task”, it probably sounds very complicated and hard to understand. So, this may increase the difficulty of visual design.
    - Actually, the task usually can be abstracted, and the action of the task will be just one of a few common actions.
    - Reduce our search in the design space

S02-03

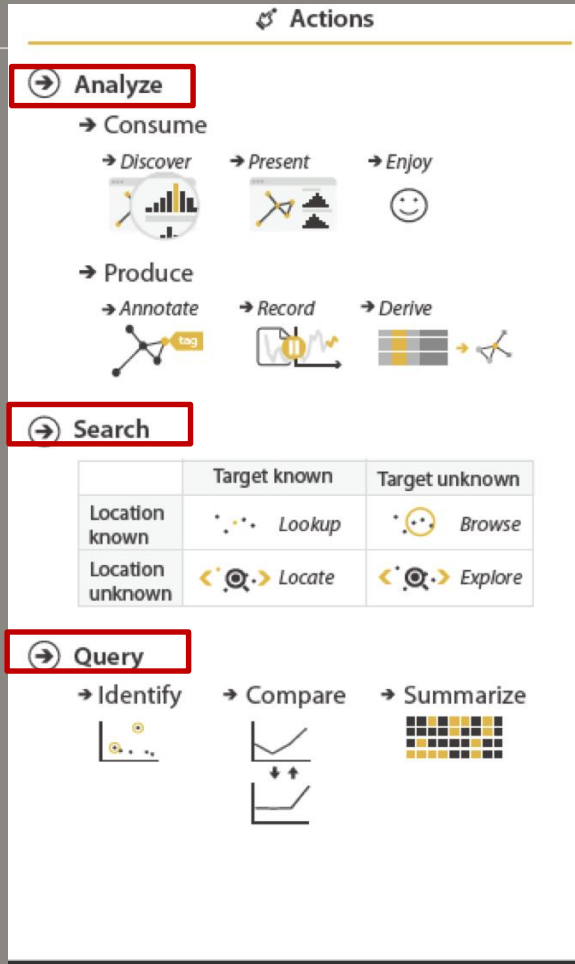
Task Abstraction: Action and Target Pattern

- A task can be described by one or multiple {action, target} pairs
  - Compare attributes
  - Discover distribution
  - Identify outlier
  - ……



Three Categories of Actions

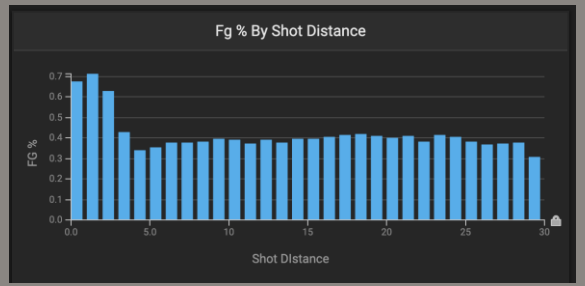
- Analyze
- Search
- Query



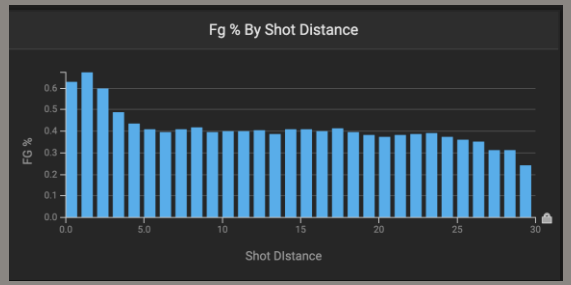
Action: Analyze

- Consume
  - Discover
    - Discover new knowledge, that was not previously known
      - generate hypothesis, verify hypothesis

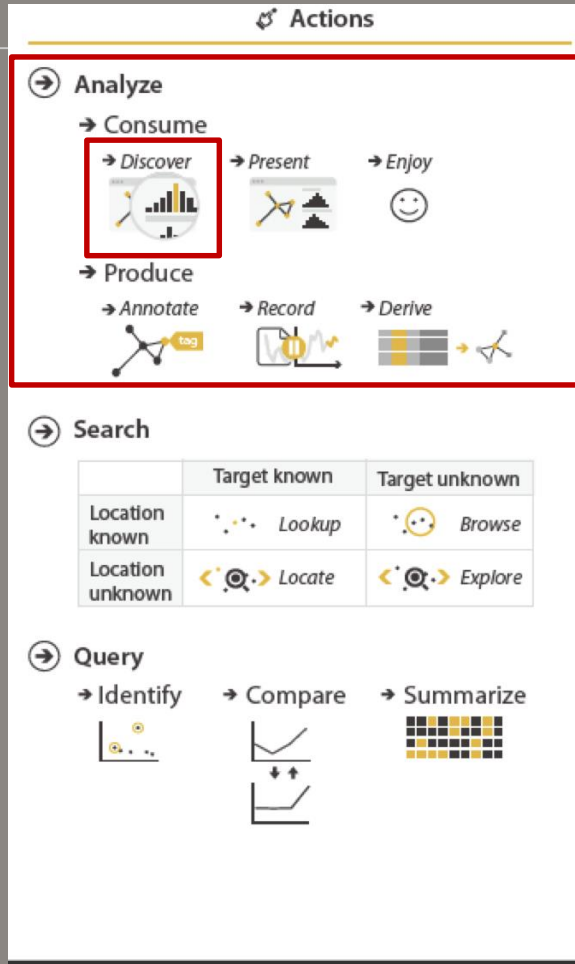
Does GSW have a better FG% around 3pts line? (we want to verify hypothesis)



GSW

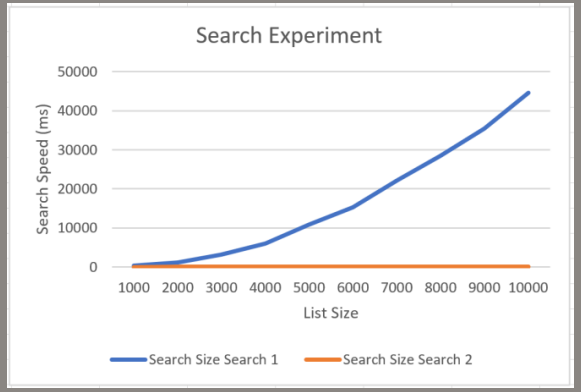


LAL

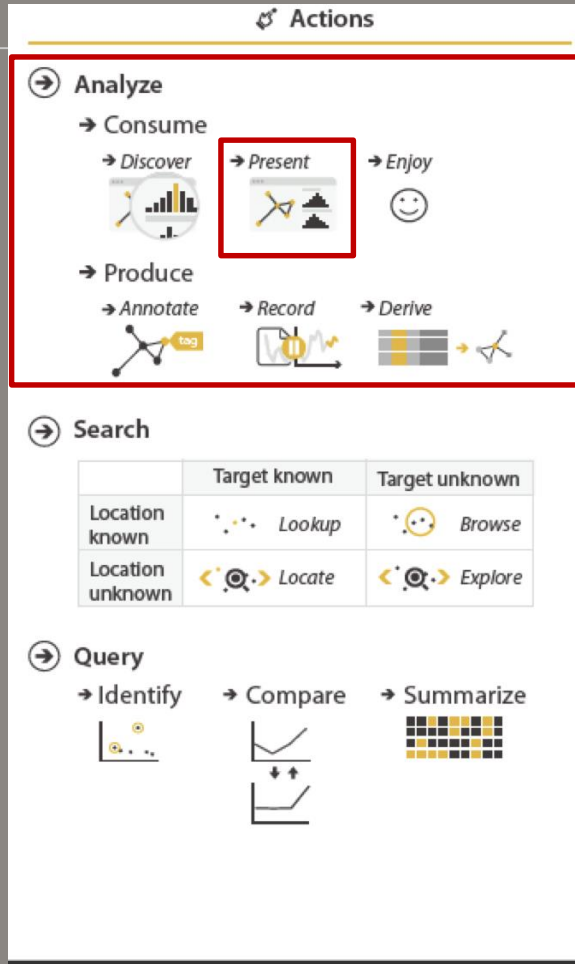


Action: Analyze

- Consume
  - Discover
  - Present (convey information, storytelling)
    - The information that the visualization creator already knows



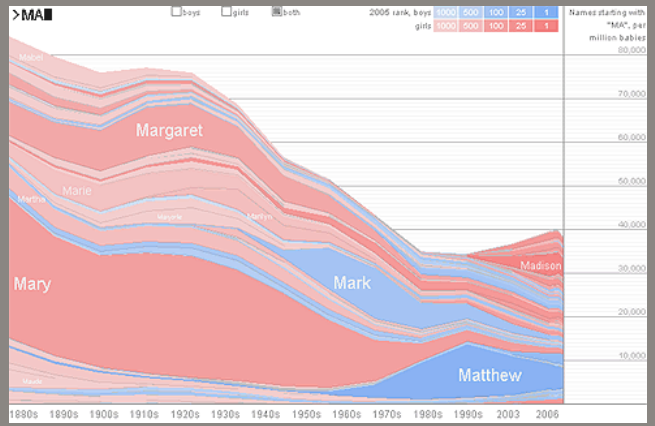
Experiment result

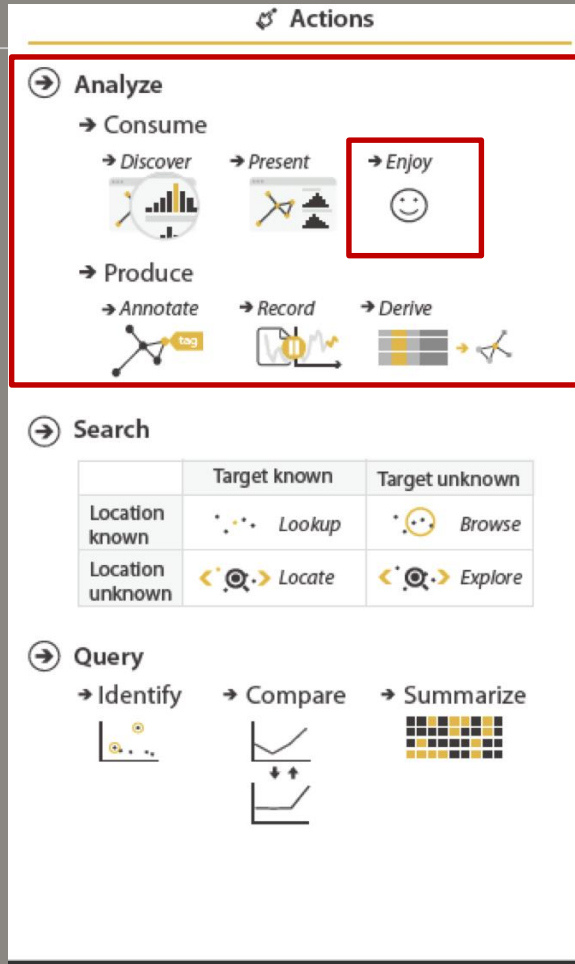


Action: Analyze

- Consume
  - Discover
  - Present
  - Enjoy
    - For fun, get simple information

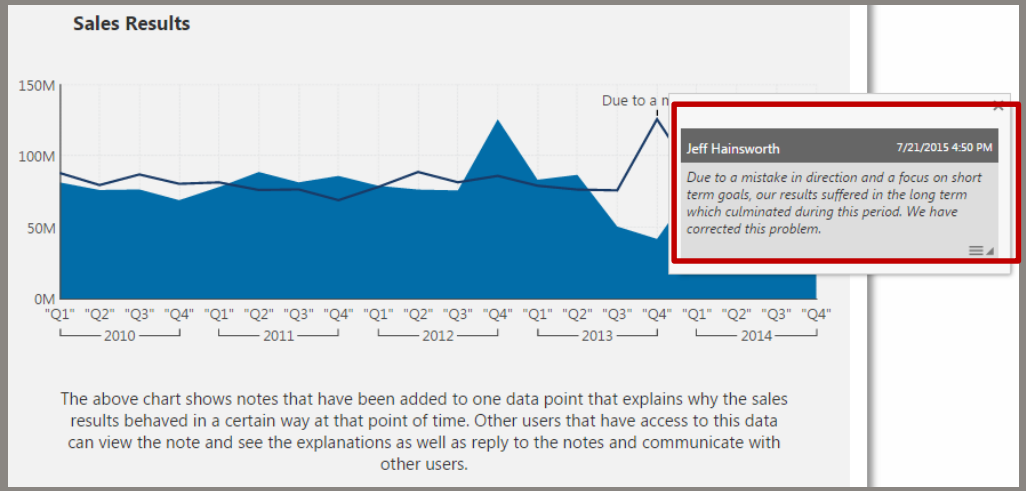
Baby name vis

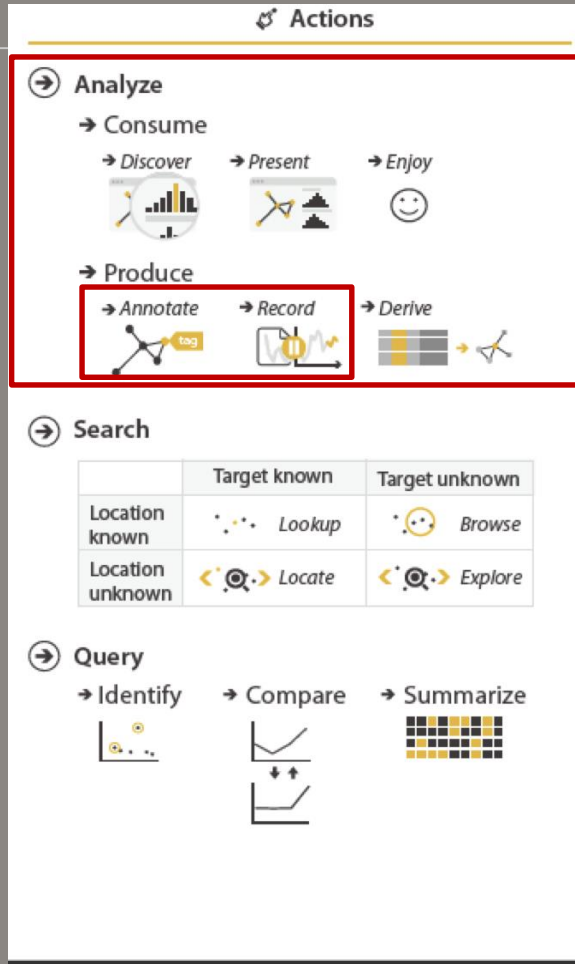




Action: Analyze

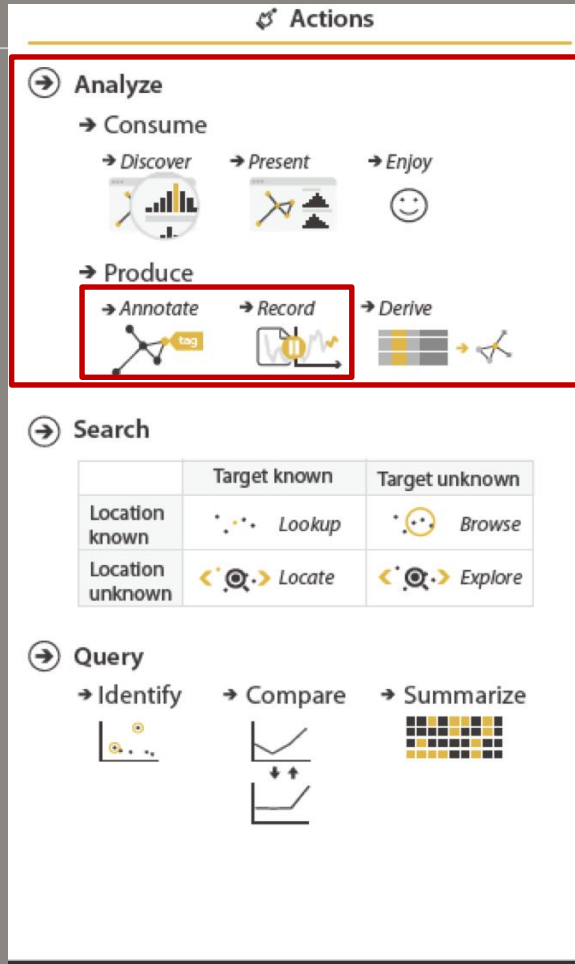
- Produce (the vis tool take input)
  - Annotate: add comment/note for the data





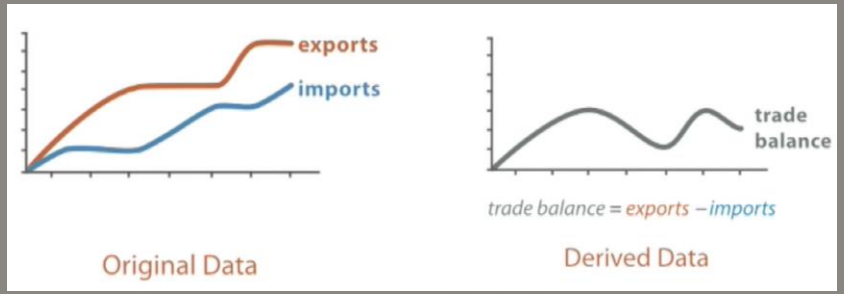
Action: Analyze

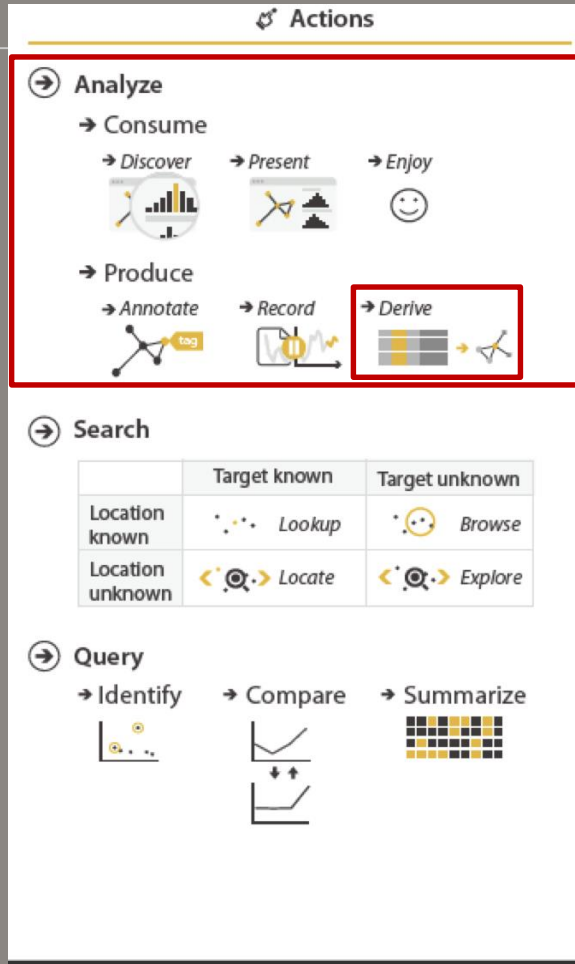
- Produce (the vis tool take input)
  - Annotate
  - Record: save each steps of interaction on the vis tool
    - As the reference of other users



Action: Analyze

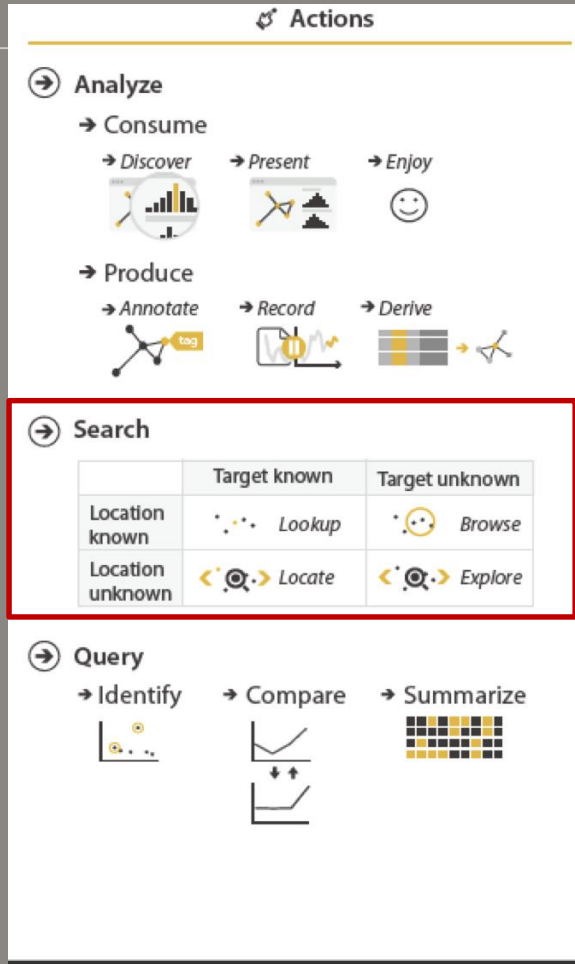
- Produce (the vis tool take input)
  - Annotate
  - Record
  - Derive: derive new data from the tool or derive new data to show
    - Don't always just draw what you're given!
    - Transform city name to latitude and longitude
    - Transform temperature (floating) to categories (hot or cold)
    - Transform attributes to relation among attributes





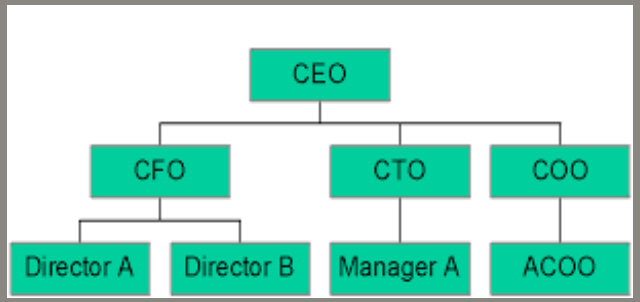
Actions: Search

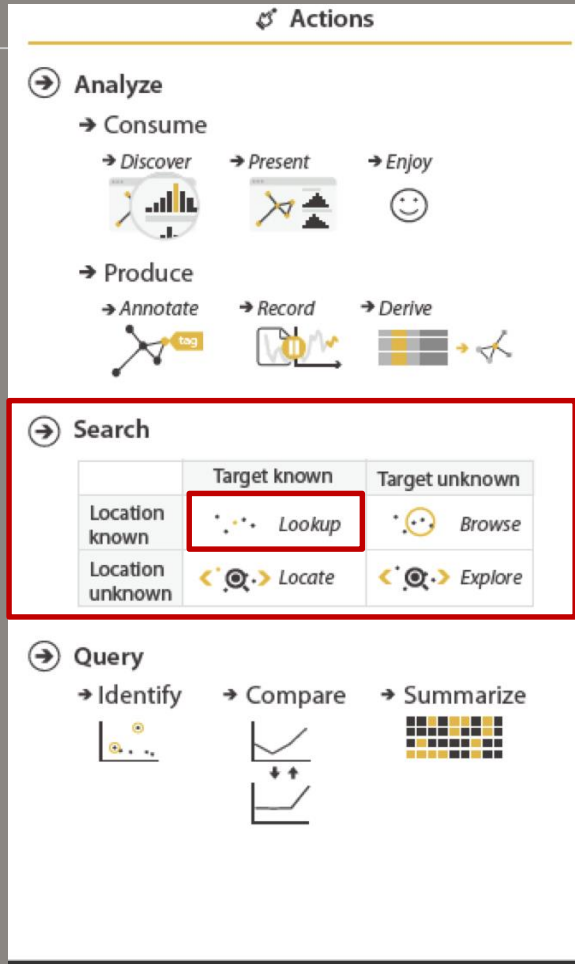
- These two question determine what users will do in search action
  - Do users know the target?
  - Do users know the location?



Actions: Search

- These two question determine what users will do in search action
  - Do users know the target?
  - Do users know the location?

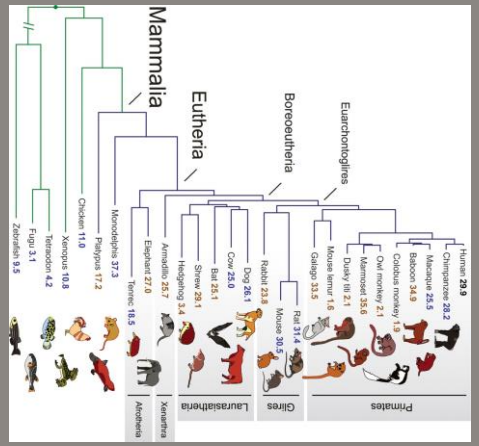


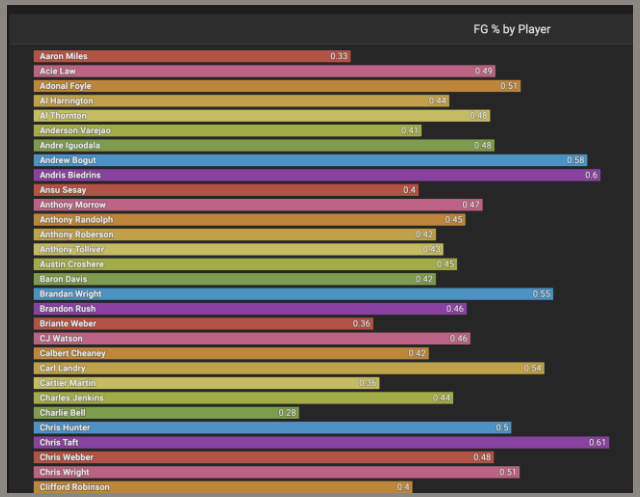


Actions: Search

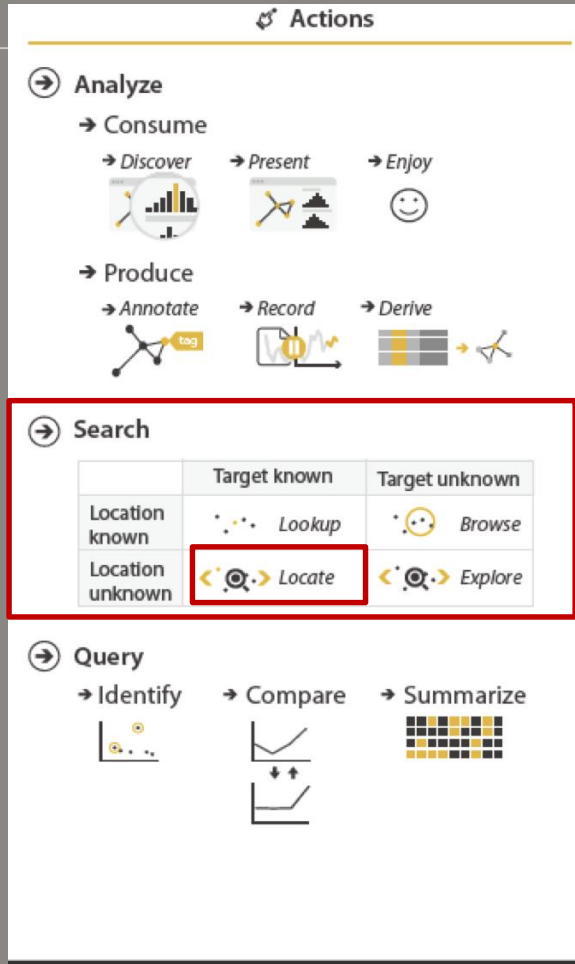
- These two question determine what users will do in search action
  - Do users know the target?
  - Do users know the location?

human?



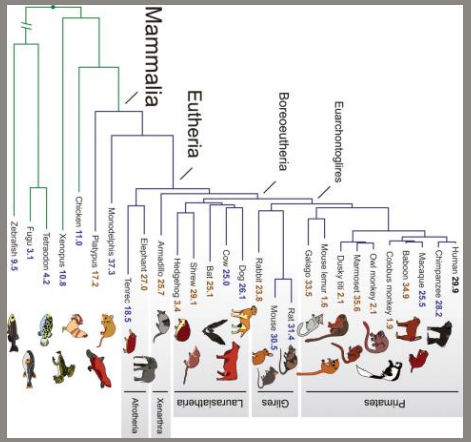


Alphabetical order (first name)

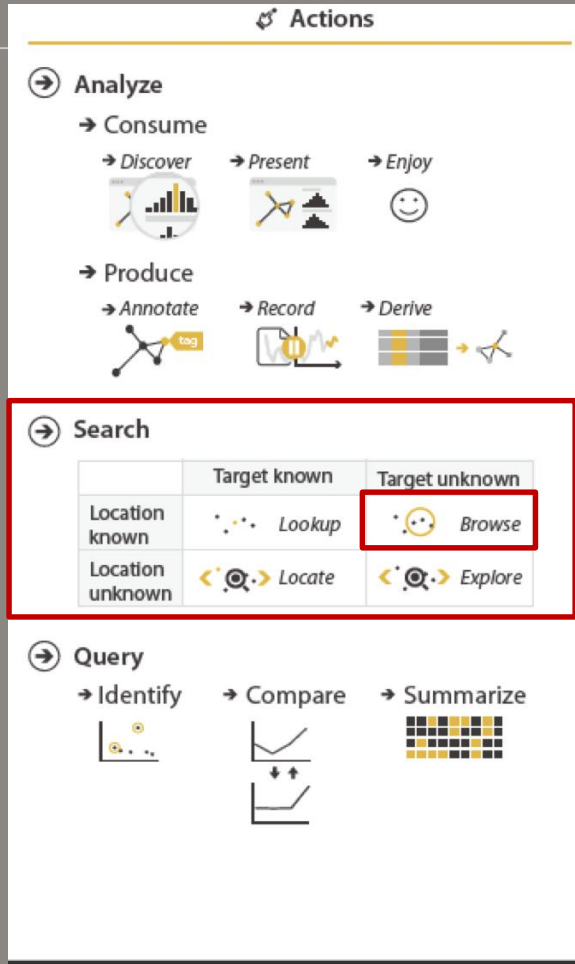


Actions: Search

- These two question determine what users will do in search action
  - Do users know the target?
  - Do users know the location?

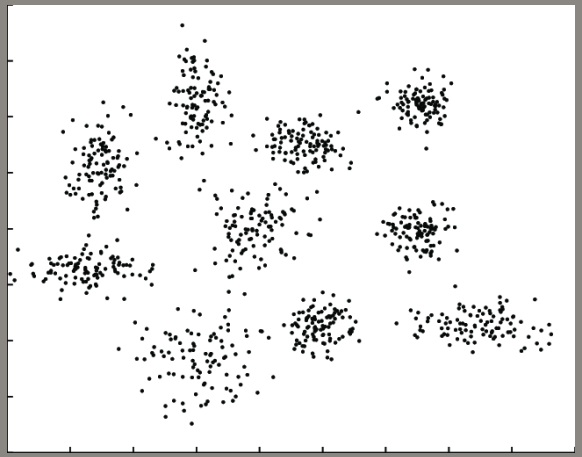


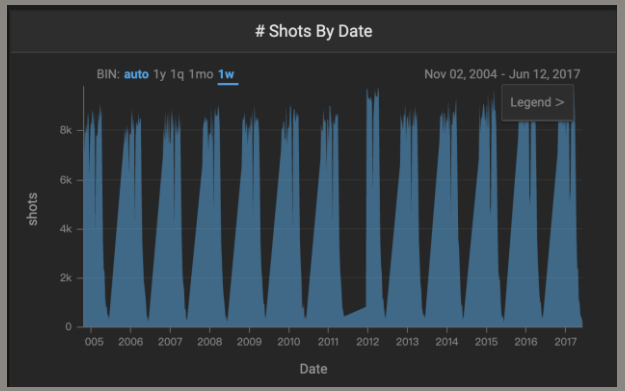
Closest to human?



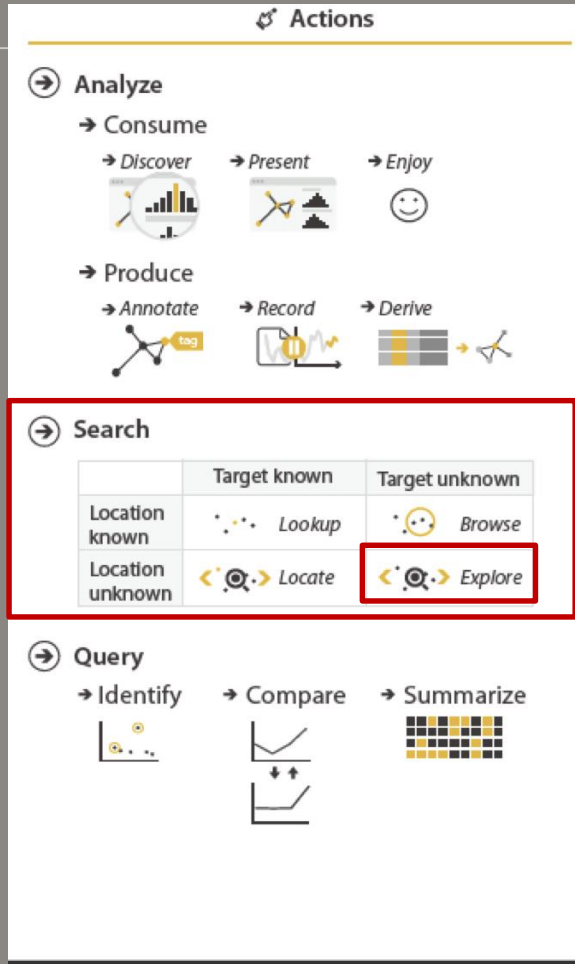
Actions: Search

- These two question determine what users will do in search action
  - Do users know the target?
  - Do users know the location?



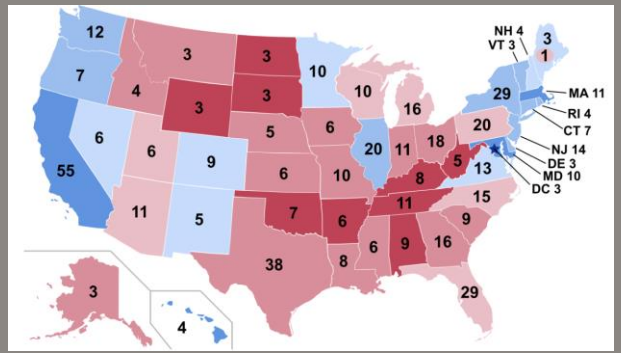


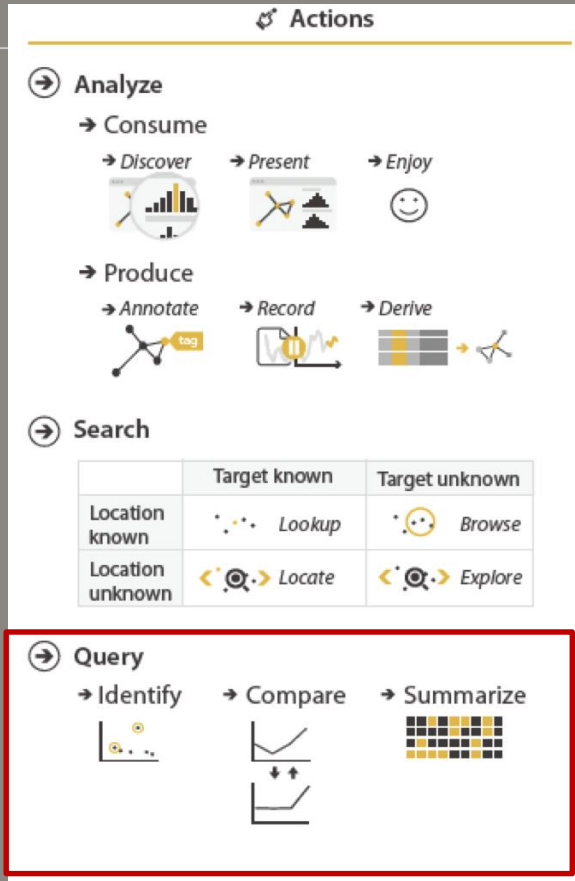
No specific goal, search special pattern for future exploration



Actions: Query

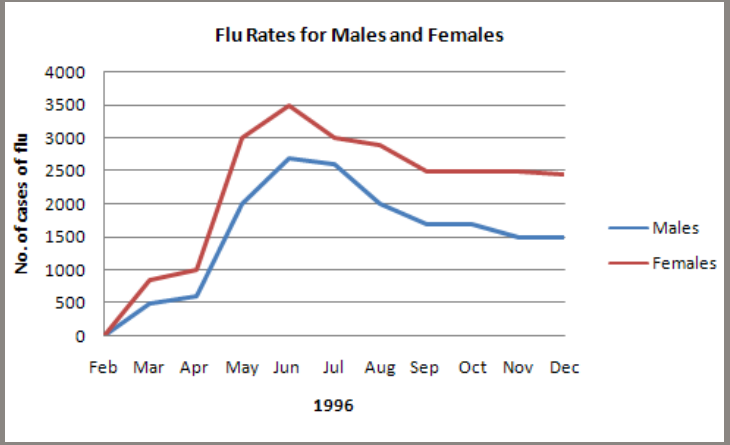
- Identify: after search
  - We are interested in California, We search and locate it. Then, we can identity its winning party and margin

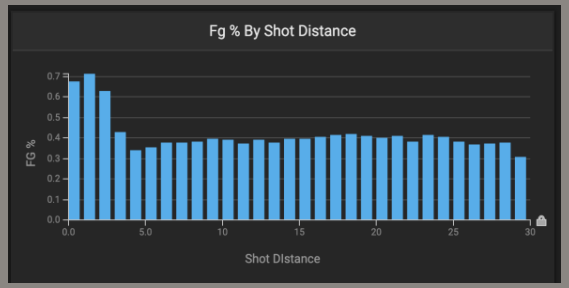


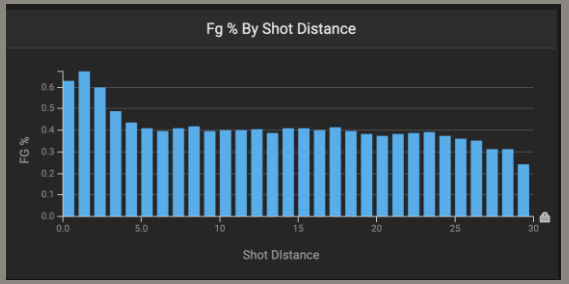


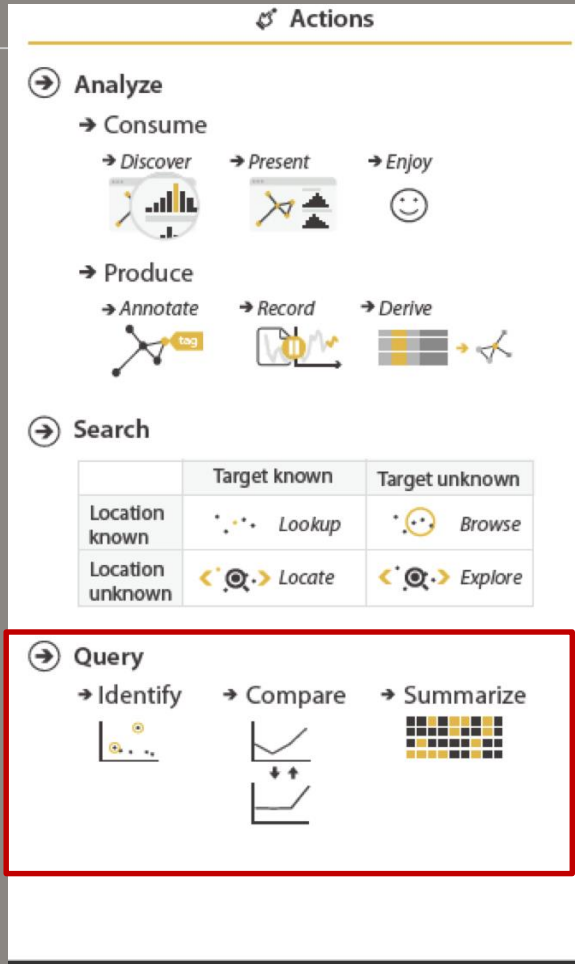
Actions: Query

- Identify
- Compare: among items, among groups





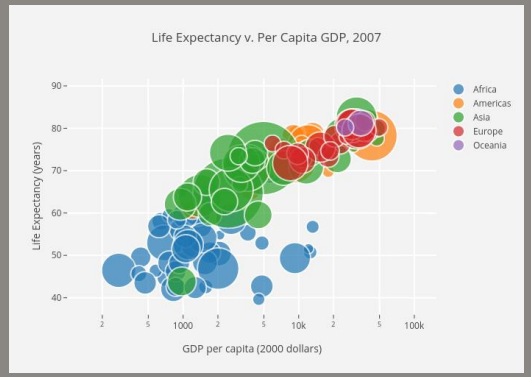


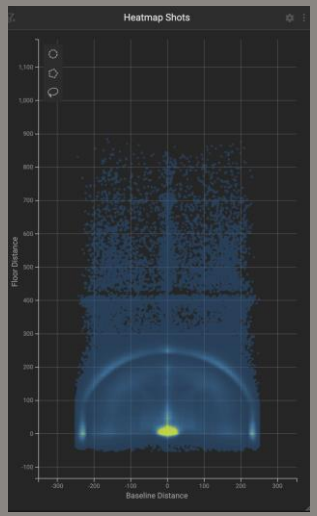


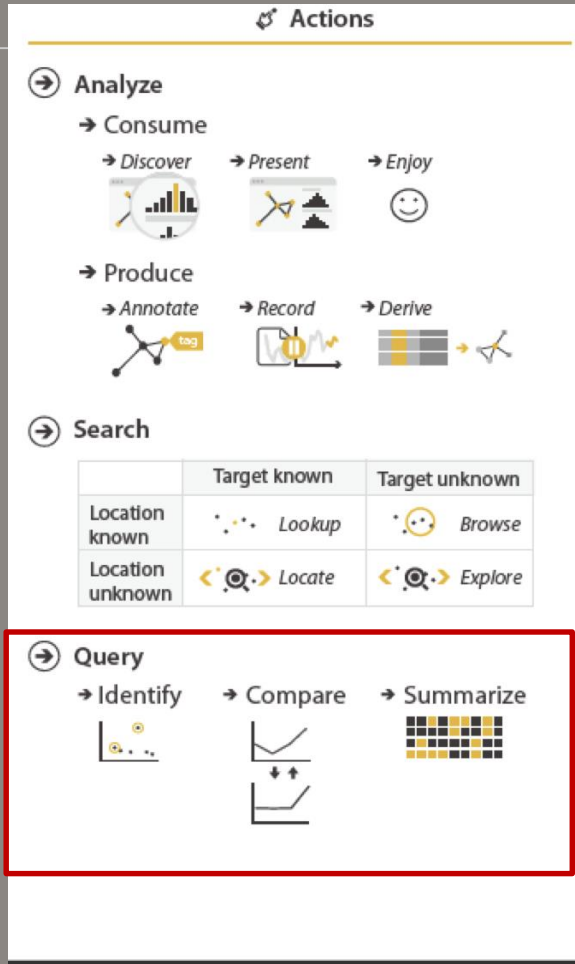
Actions: Query

- Identify
- Compare
- Summarize: provide overview. Lead users to find out more information by more interaction

Find/explore the pattern in the dataset



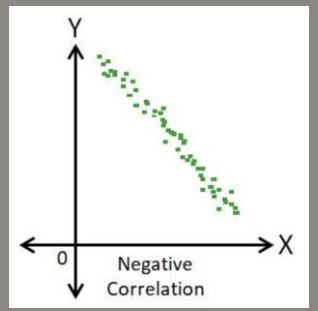


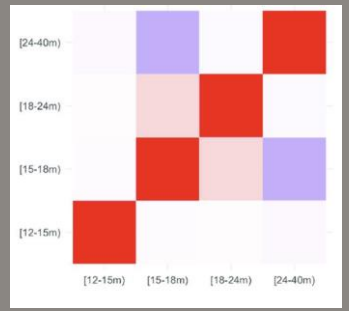


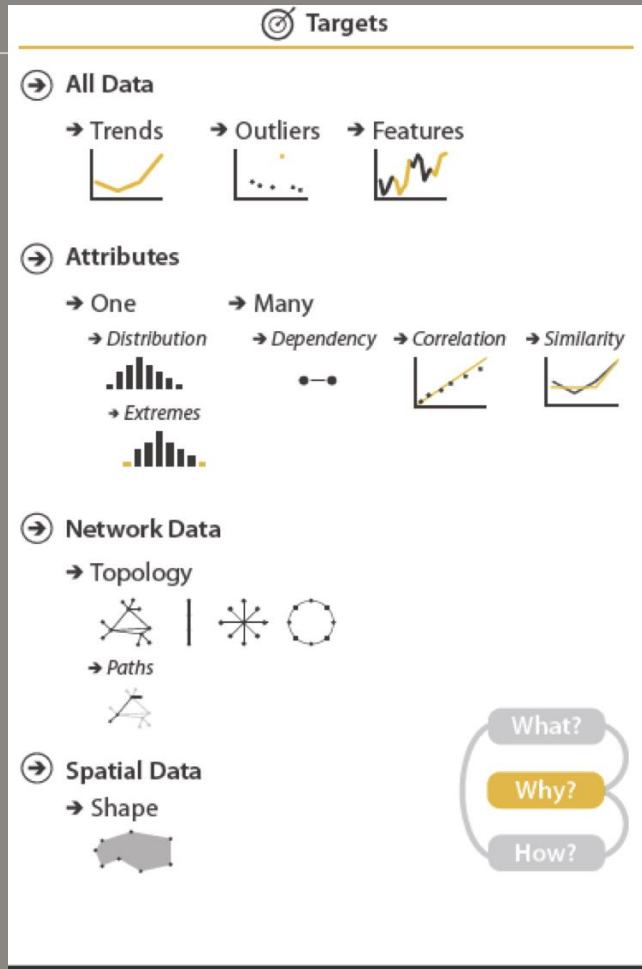
S02-04

Target

- To know what portions of the dataset you need to use to consider the visual design
  - e.g. A attribute is related to B or not?
  - You may need all data items of attributes A and B

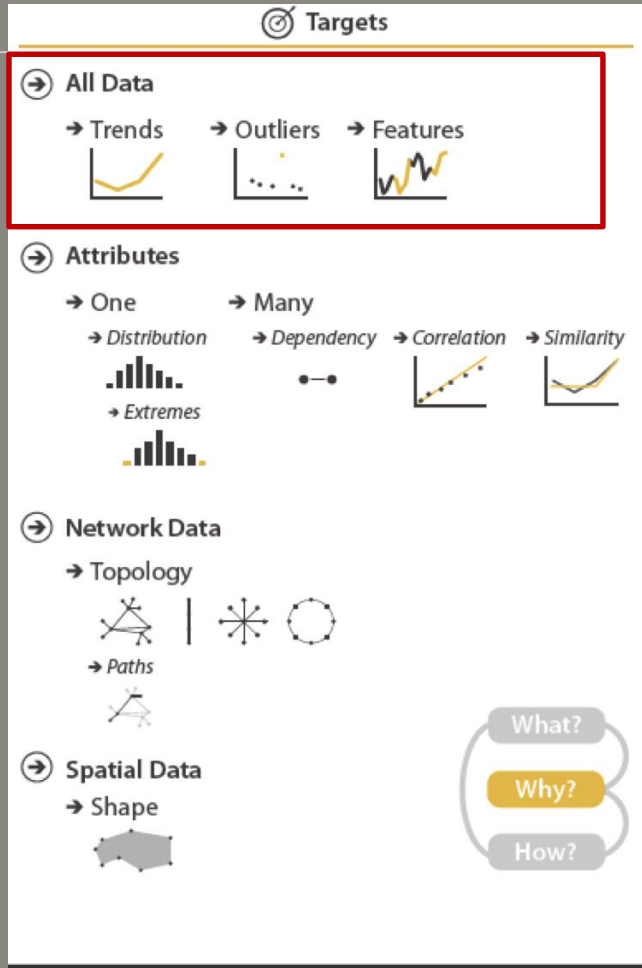






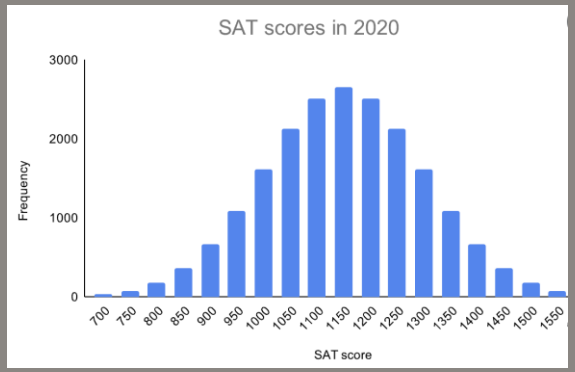
Target

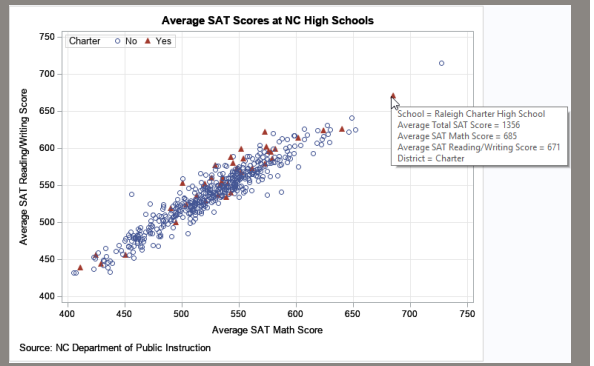
- All data
  - If your target is trends, outliers, features, you need all data to achieve the goal.
  - At least, it is impossible to visualize only one data item to see the trend

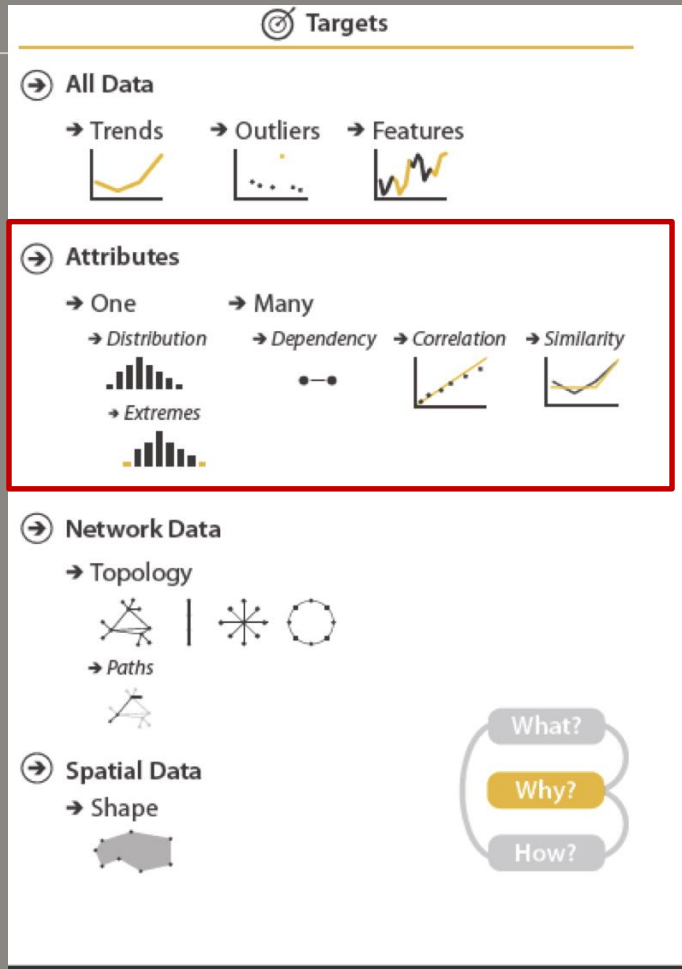


Target

- All data
- Attribute
  - Example
    - single
      - observe distribution of salary
      - Find min, max salary
    - Multiple
      - Correlation between salary and age







Target

- All data
- Attribute
- Network data

Targets

All Data
- Trends
- Outliers
- Features

Attributes
- One
- Many
  - Distribution
  - Dependency
  - Correlation
  - Similarity
  - Extremes

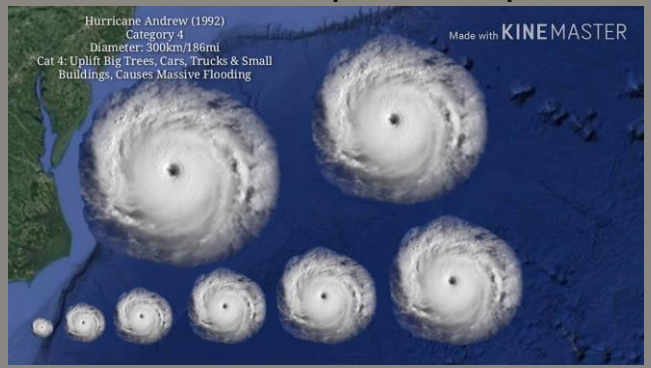
Network Data
- Topology
- Paths

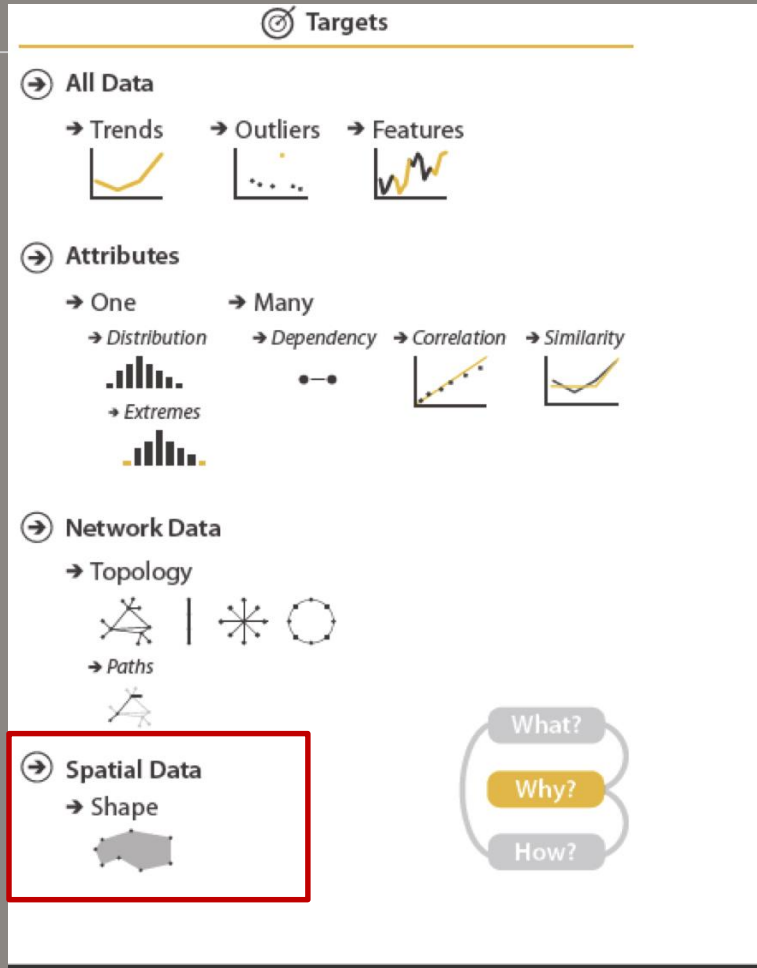
Spatial Data
- Shape

What?
Why?
How?

Target

- All data
- Attribute
- Network data
- Spatial data
  - Shape comparison





How?

Encode

Arrange

- Express
- Separate
- Order
- Align
- Use

Map

from categorical and ordered attributes

- Color
  - *Hue*
  - *Saturation*
  - *Luminance*
- Size, Angle, Curvature, ...
- Shape
- Motion

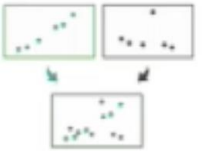
Direction, Rate, Frequency, ...

Manipulate

- Change
- Select
- Navigate

Facet

- Juxtapose
- Partition
- Superimpose



Reduce

- Filter
- Aggregate
- Embed