

CS49000-VIZ - Fall 2020

Introduction to Data Visualization

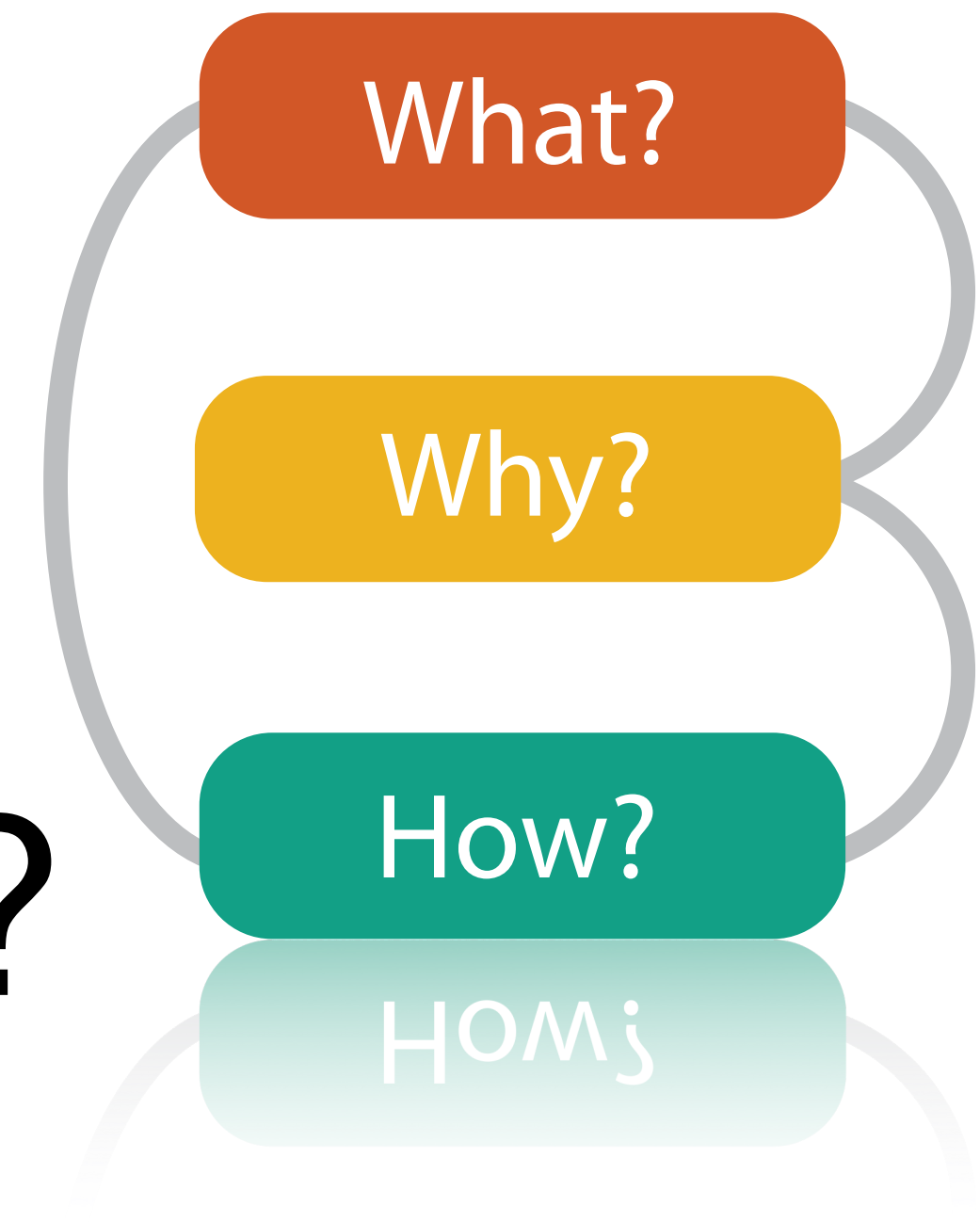
Methodology

Lecture 4

September 7, 2020

Analysis: 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



Analysis: What, why, and how

- **Abstract** vocabulary avoids domain-specific terms
- Translation process iterative, tricky
- What-why-how analysis framework as scaffold to **think systematically about design space**



How: Idiom design choices

How?

Encode

➔ Arrange

➔ Express



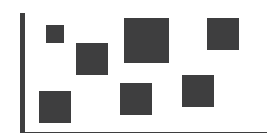
➔ Order



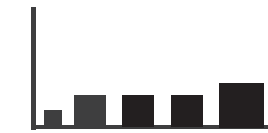
➔ Use



➔ Separate



➔ Align



➔ Map

from **categorical** and **ordered** attributes

➔ Color

➔ Hue



➔ Saturation



➔ Luminance



➔ Size, Angle, Curvature, ...

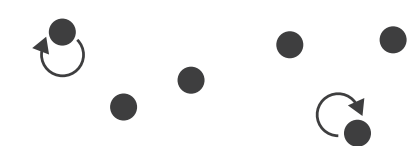


➔ Shape



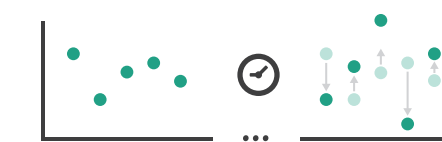
➔ Motion

Direction, Rate, Frequency, ...

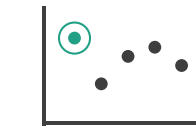


Manipulate

➔ Change



➔ Select



➔ Navigate

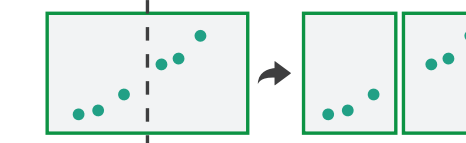


Facet

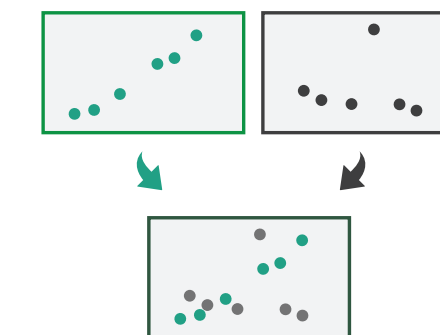
➔ Juxtapose



➔ Partition

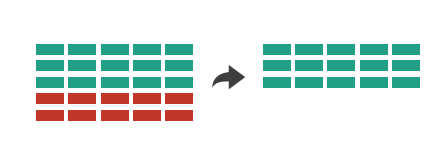


➔ Superimpose

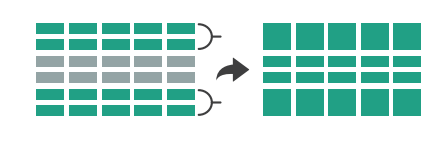


Reduce

➔ Filter



➔ Aggregate



➔ Embed



What?

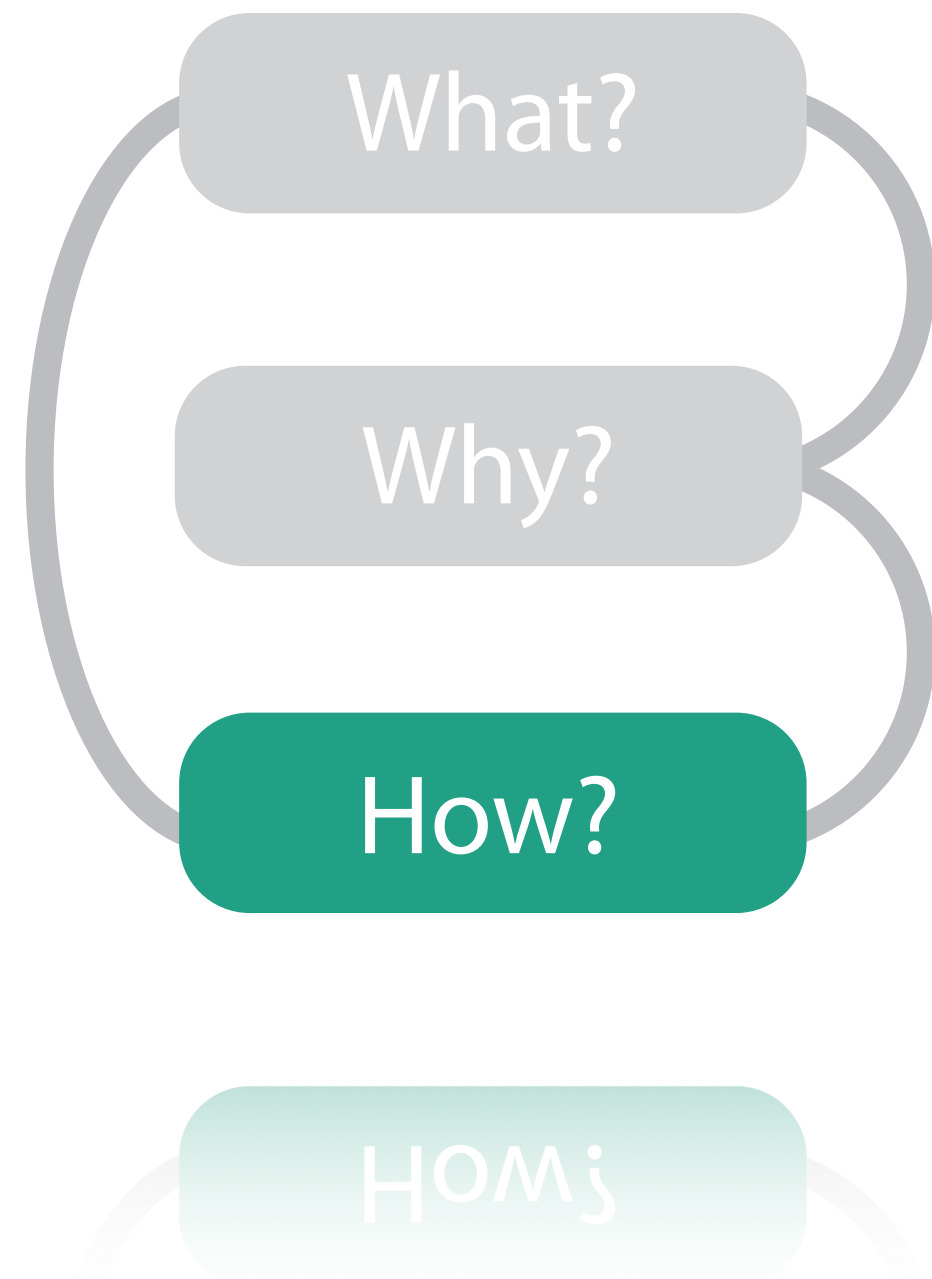
Why?

How?

How?

Encode

Encode

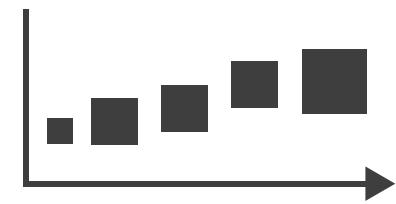


→ Arrange

→ Express



→ Order



→ Use



→ Separate



→ Align



→ Map

from **categorical** and **ordered** attributes

→ Color

→ Hue



→ Saturation



→ Luminance



→ Size, Angle, Curvature, ...

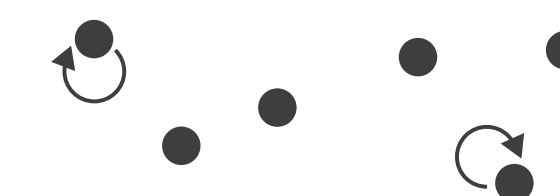


→ Shape



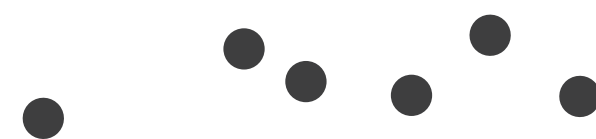
→ Motion

Direction, Rate, Frequency, ...



Marks and channels

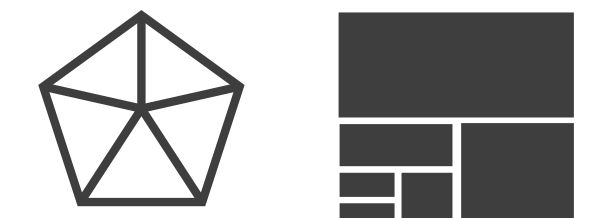
→ Points



→ Lines



→ Areas



→ Position

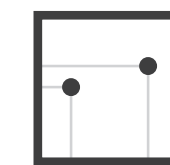
→ Horizontal



→ Vertical



→ Both



→ Color



→ Shape



→ Tilt



→ Size

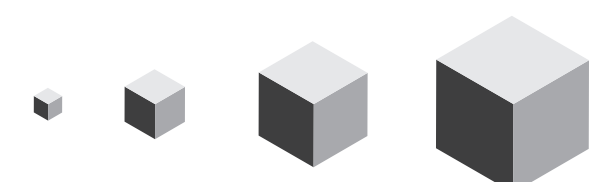
→ Length



→ Area



→ Volume

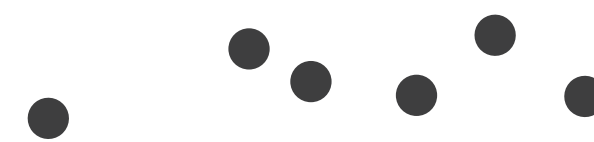


Marks and channels

- Marks

- geometric primitives

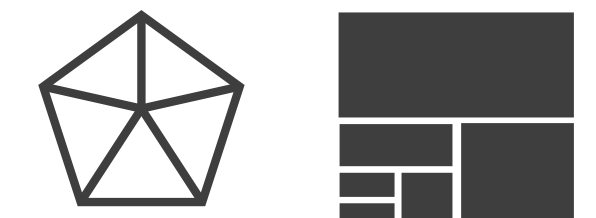
→ Points



→ Lines



→ Areas



→ Position

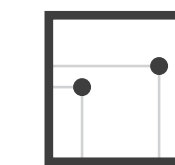
→ Horizontal



→ Vertical



→ Both



→ Color



→ Shape



→ Tilt



→ Size

→ Length



→ Area



→ Volume

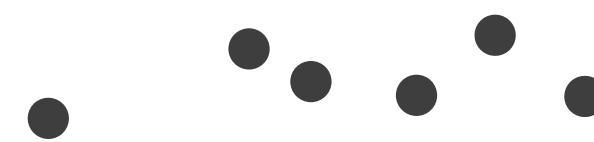


Marks and channels

- Marks

- geometric primitives

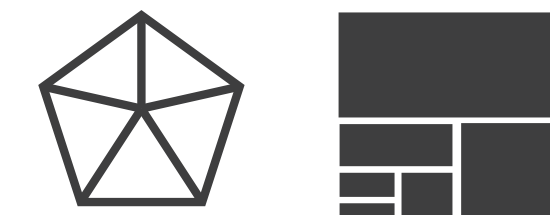
→ Points



→ Lines



→ Areas



→ Position

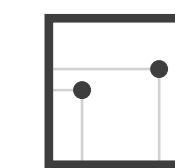
→ Horizontal



→ Vertical



→ Both



→ Color



→ Shape



→ Tilt



→ Size

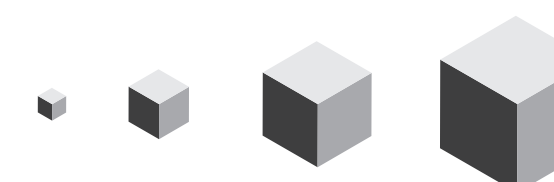
→ Length



→ Area



→ Volume



Marks and channels

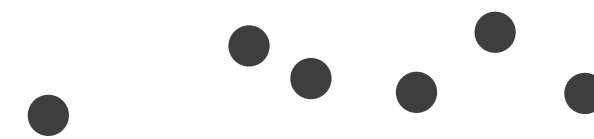
- Marks

- geometric primitives

- Channels

- control appearance of marks

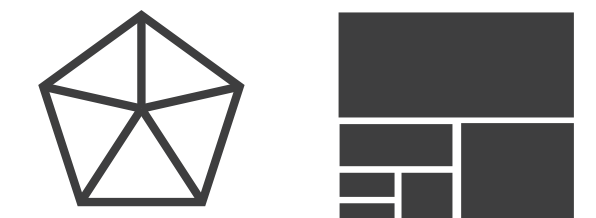
→ Points



→ Lines



→ Areas



→ Position

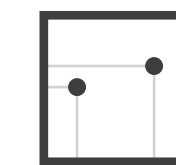
→ Horizontal



→ Vertical



→ Both



→ Color



→ Shape



→ Tilt



→ Size

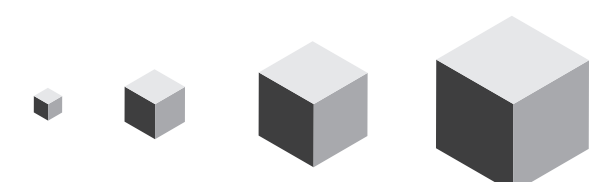
→ Length



→ Area

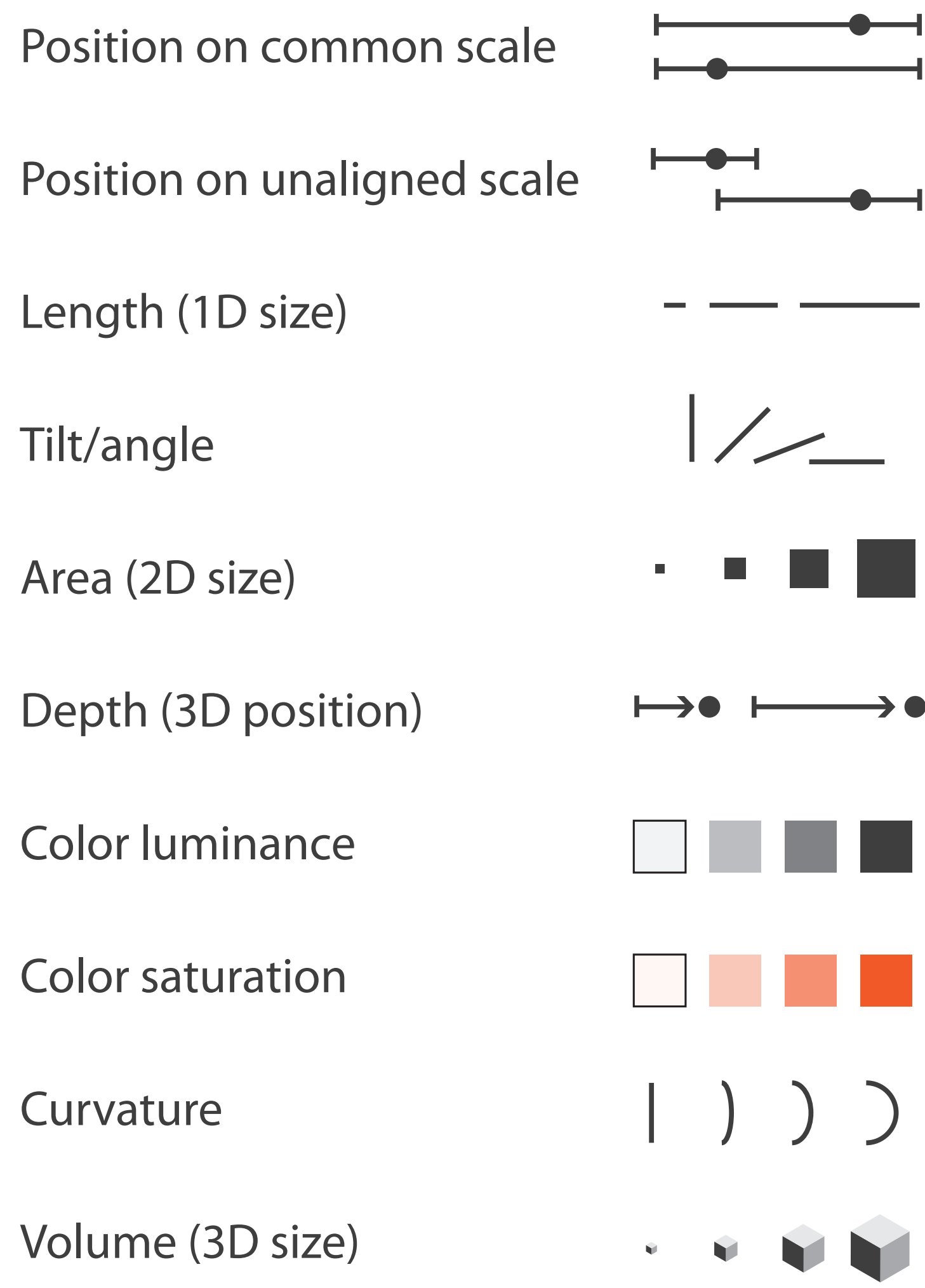


→ Volume



Channels: Types and rankings

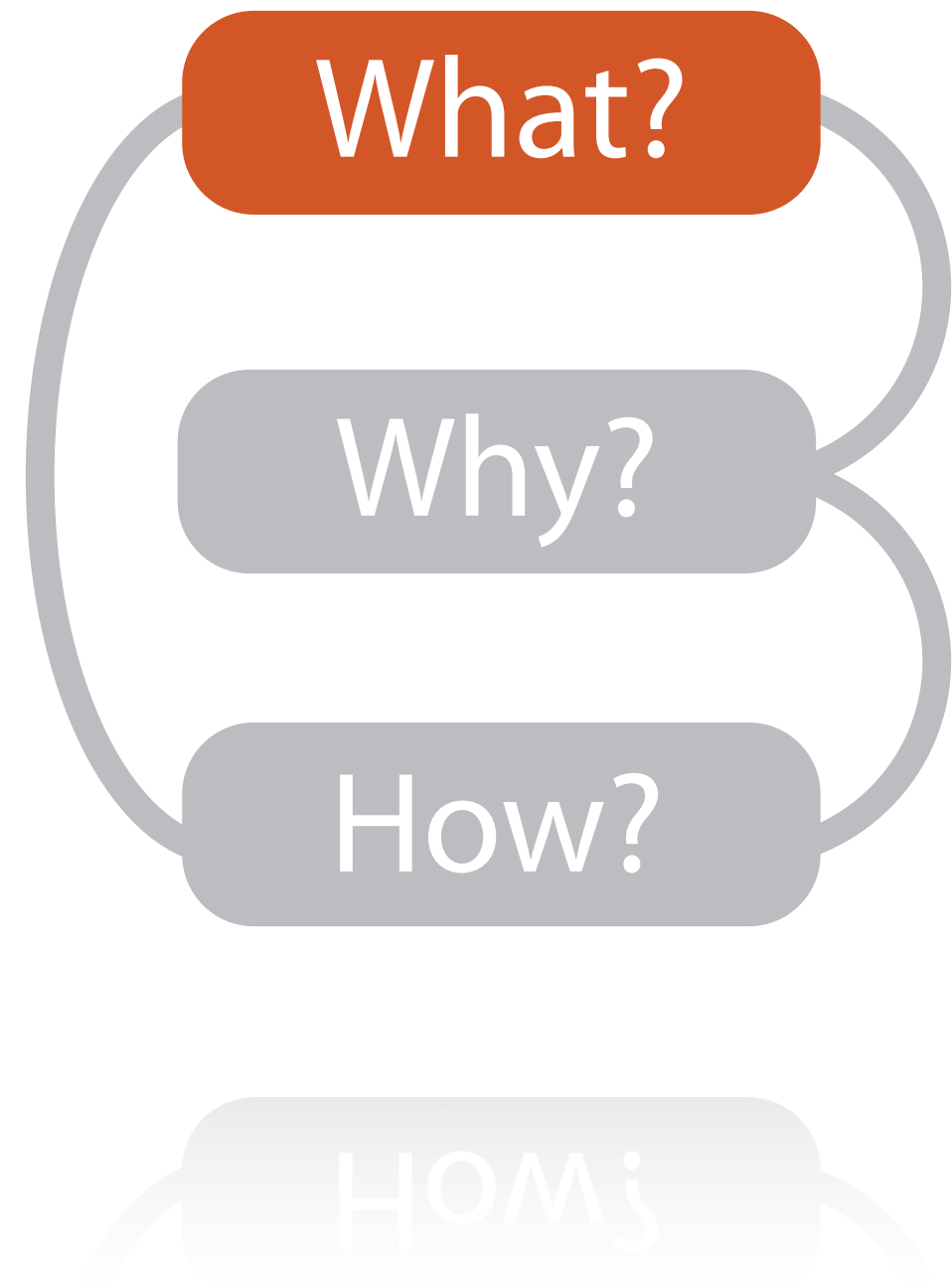
➔ Magnitude Channels: Ordered Attributes



➔ Identity Channels: Categorical Attributes



What: Input Data

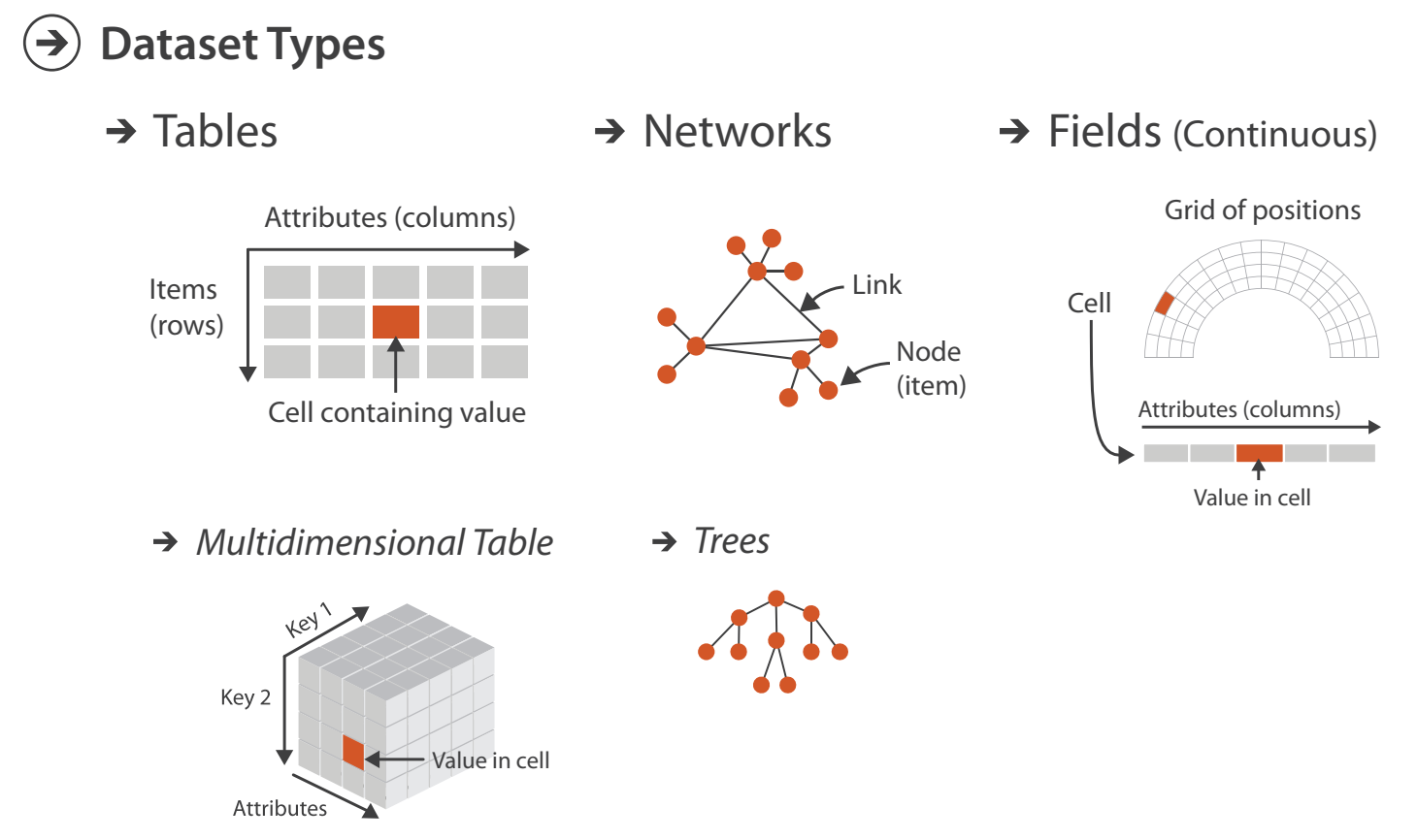
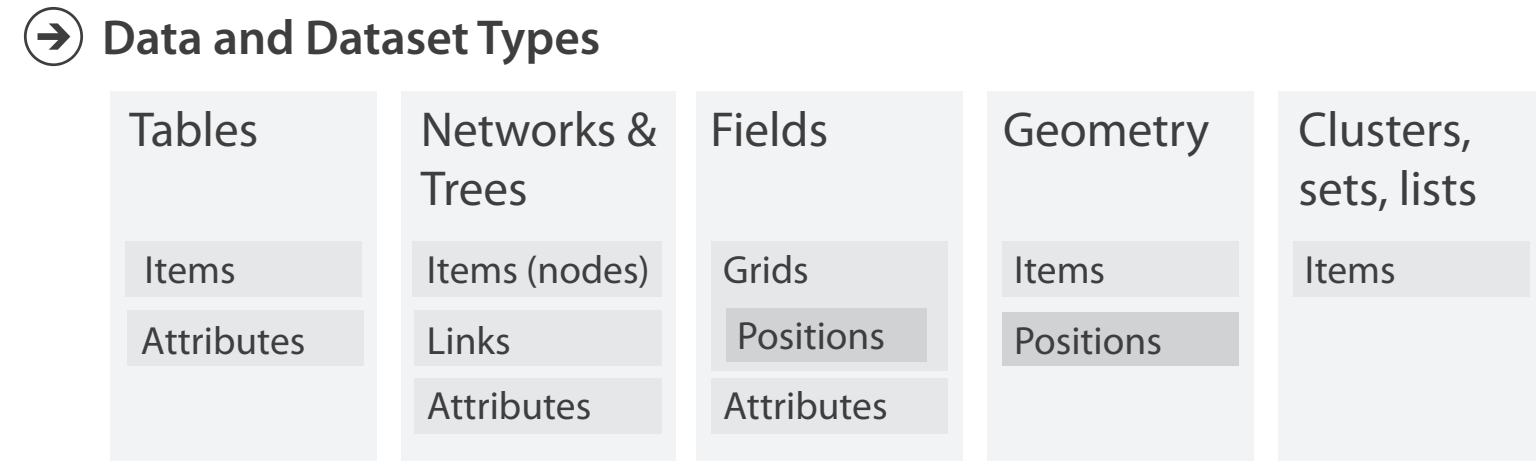


What?

Datasets Attributes

- ➔ **Data Types**
- ➔ Items
 - ➔ Attributes
 - ➔ Links
 - ➔ Positions
 - ➔ Grids

- ➔ **Attribute Types**
- ➔ Categorical
 - + ● ■ ▲
 - ➔ Ordered
 - ➔ Ordinal
 - ↑ ↑↑ ↑↑↑
 - ➔ Quantitative
 - ——— —————



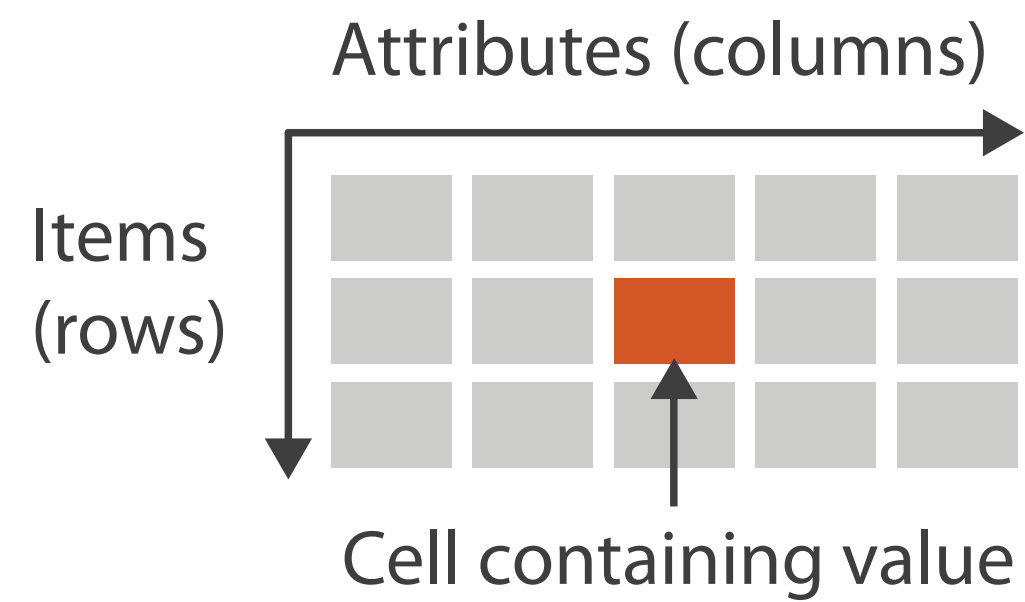
- ➔ **Ordering Direction**
- ➔ Sequential
 -
 - ➔ Diverging
 - ← →
 - ➔ Cyclic
 - ↻

- ➔ **Geometry (Spatial)**
-
- ➔ **Dataset Availability**
- ➔ Static
 - 📄
 - ➔ Dynamic
 - ⋯ →

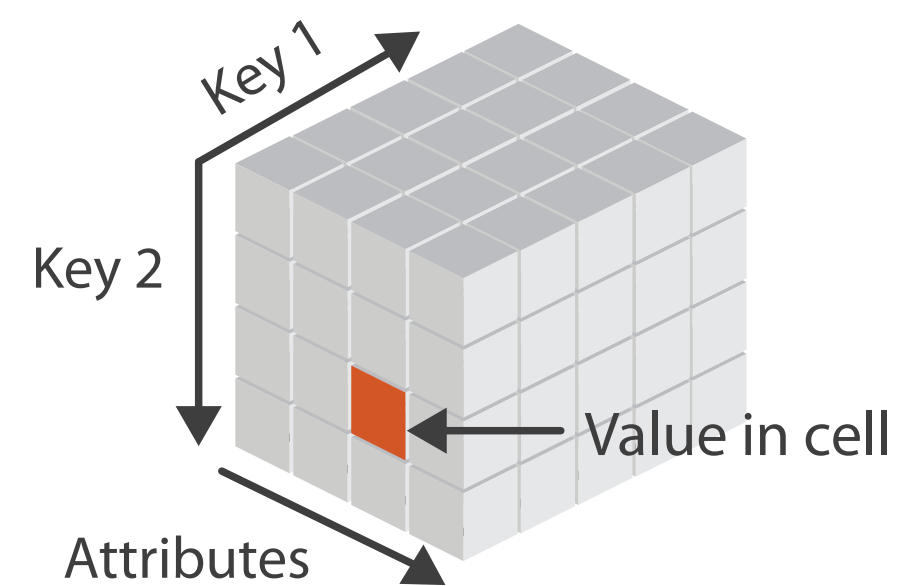
Dataset types

→ Dataset Types

→ Tables



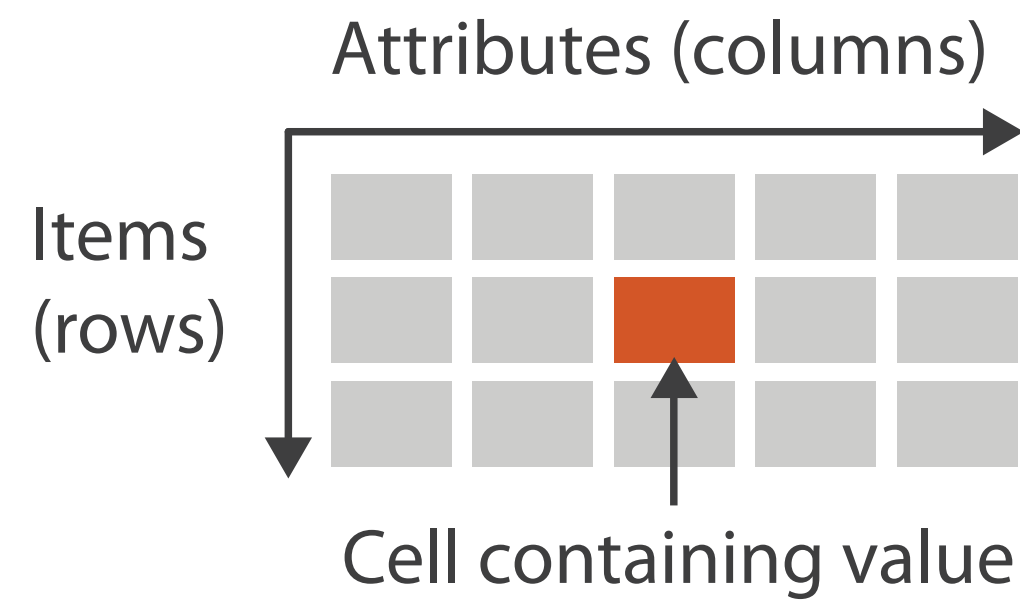
→ *Multidimensional Table*



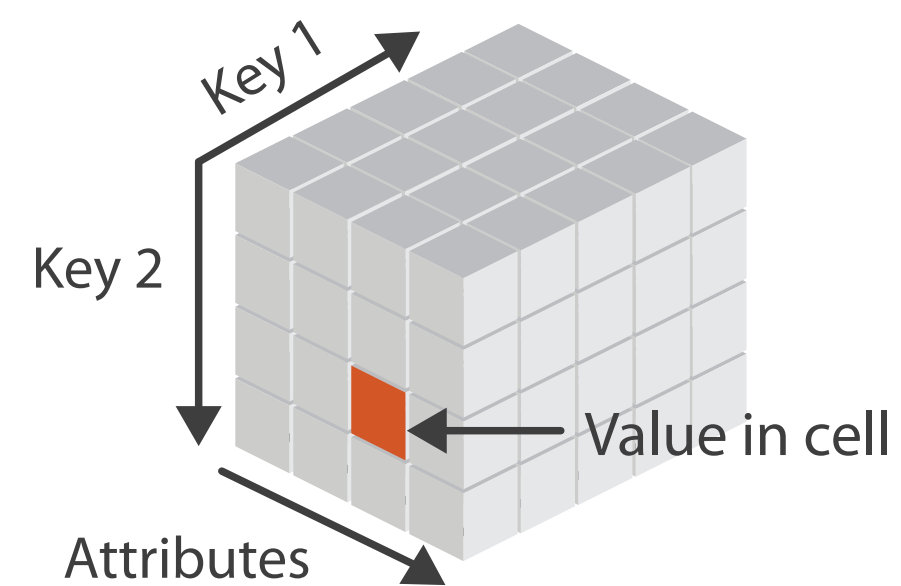
Dataset types

→ Dataset Types

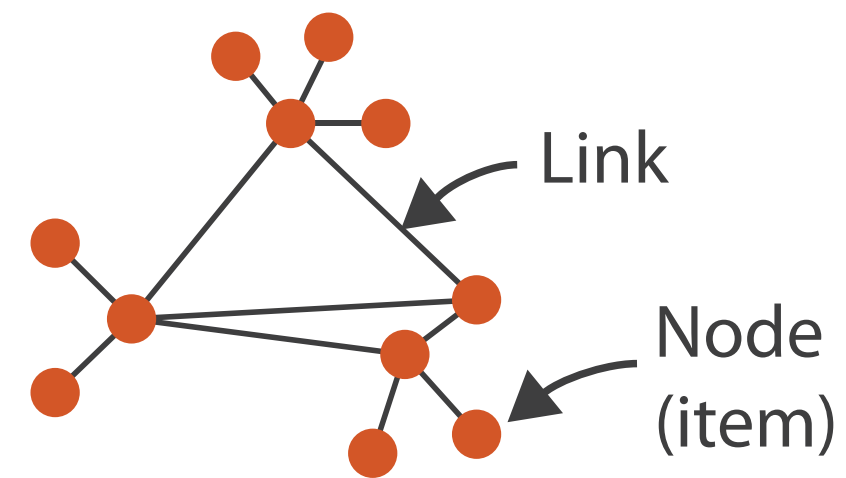
→ Tables



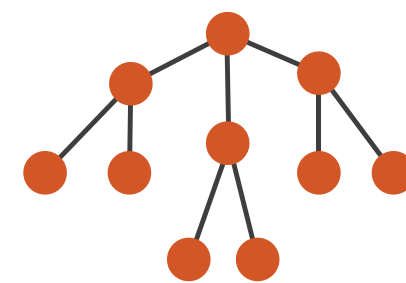
→ *Multidimensional Table*



→ Networks



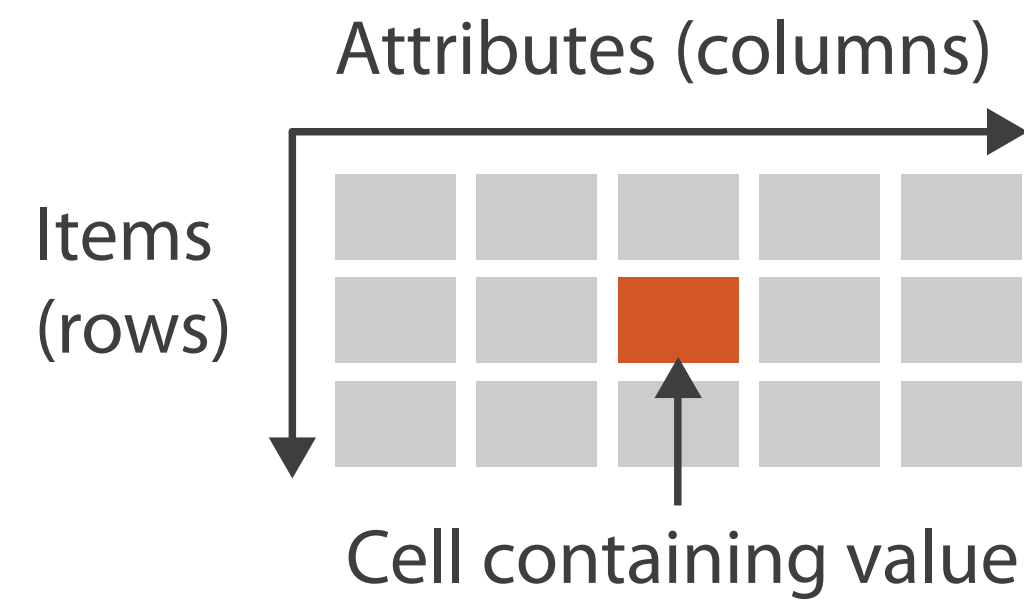
→ *Trees*



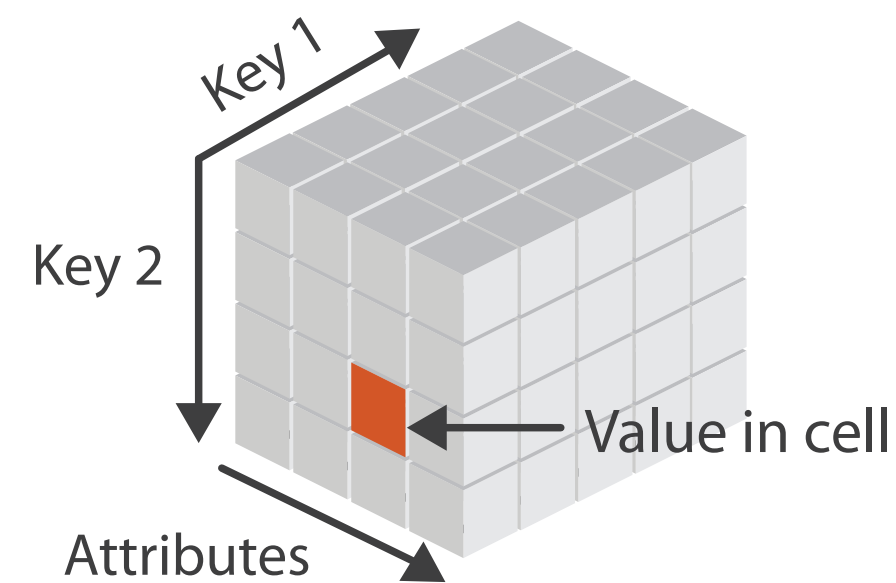
Dataset types

→ Dataset Types

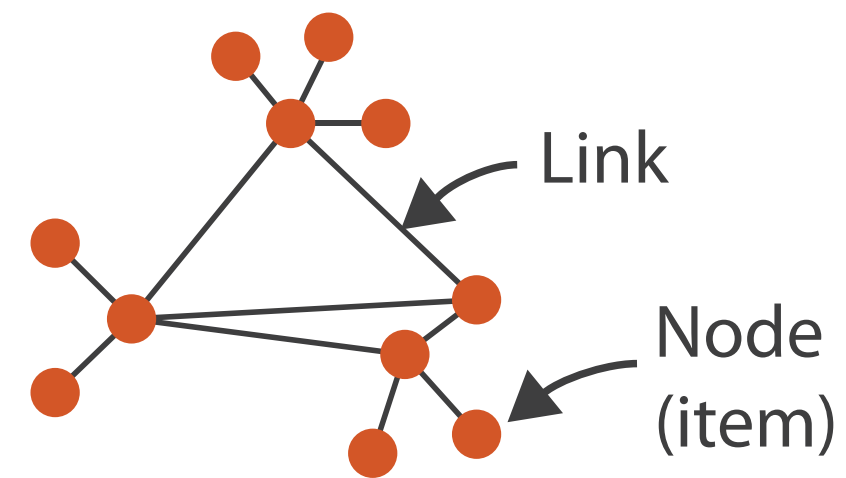
→ Tables



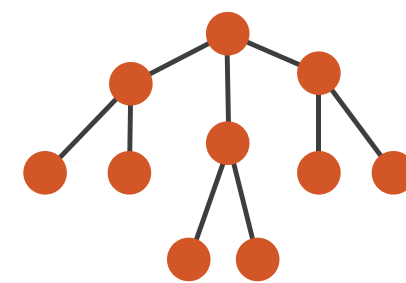
→ *Multidimensional Table*



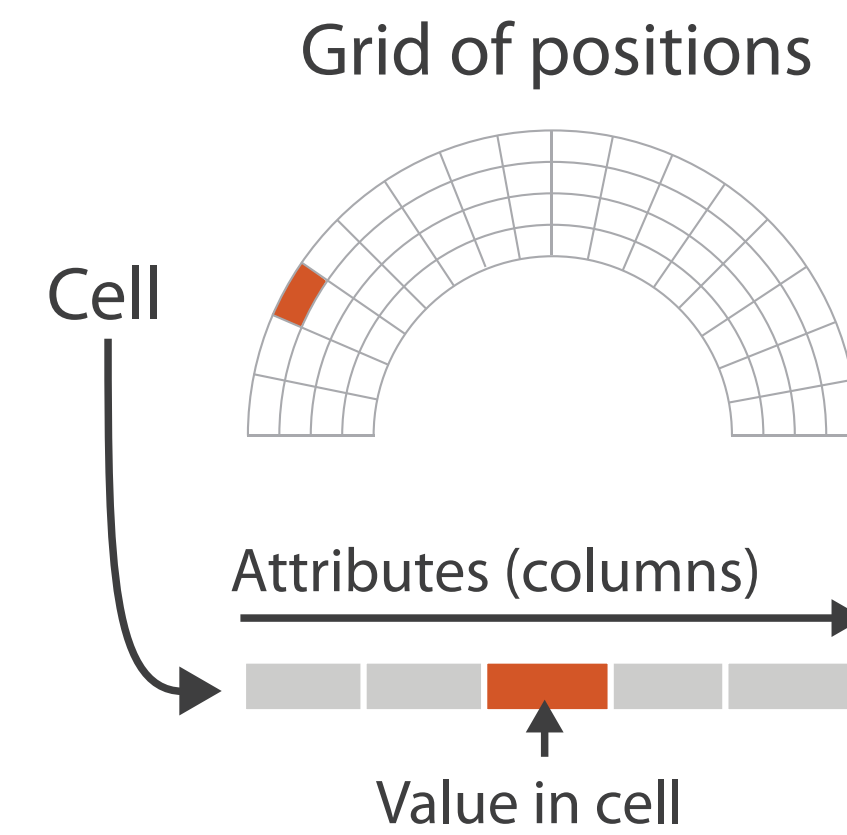
→ Networks



→ Trees



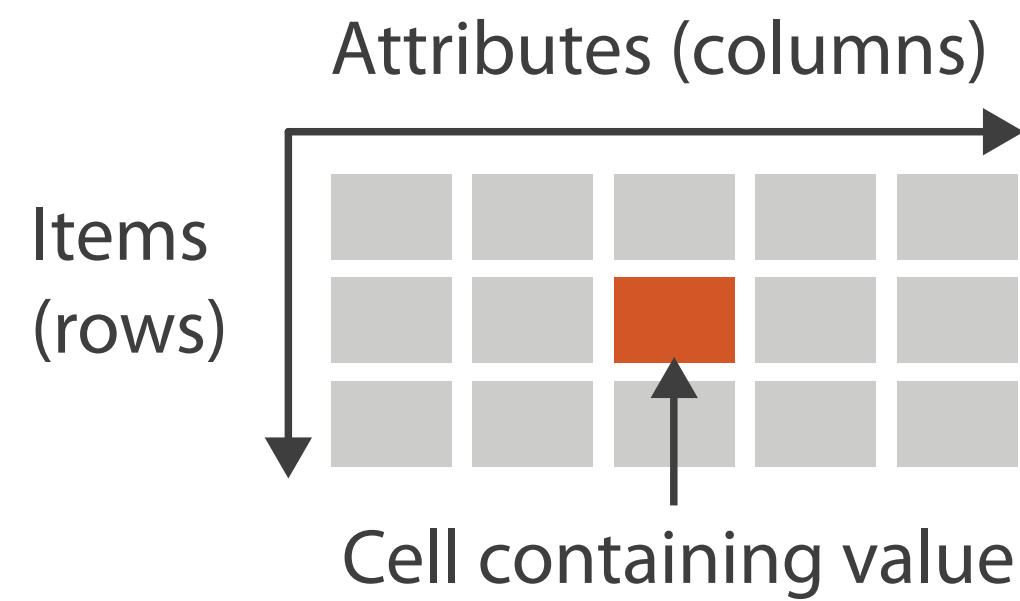
→ Fields (Continuous)



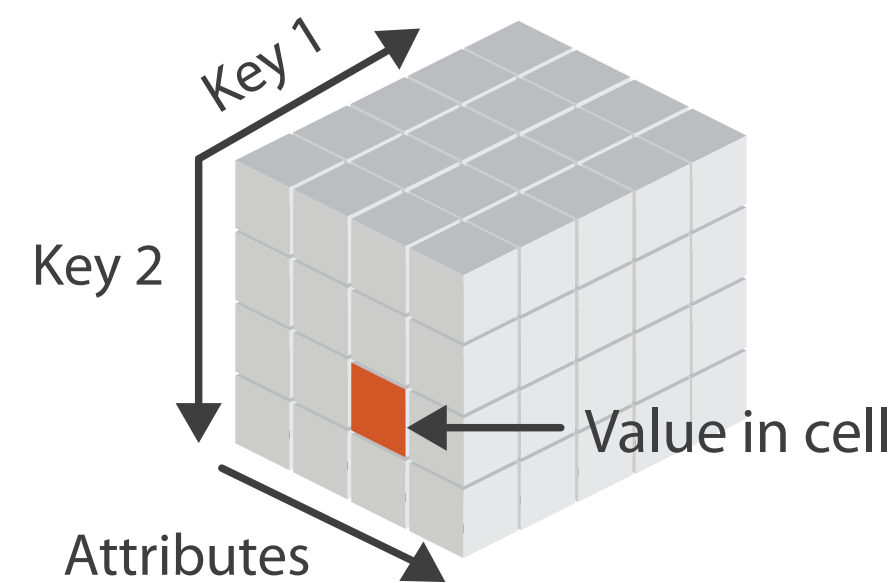
Dataset types

→ Dataset Types

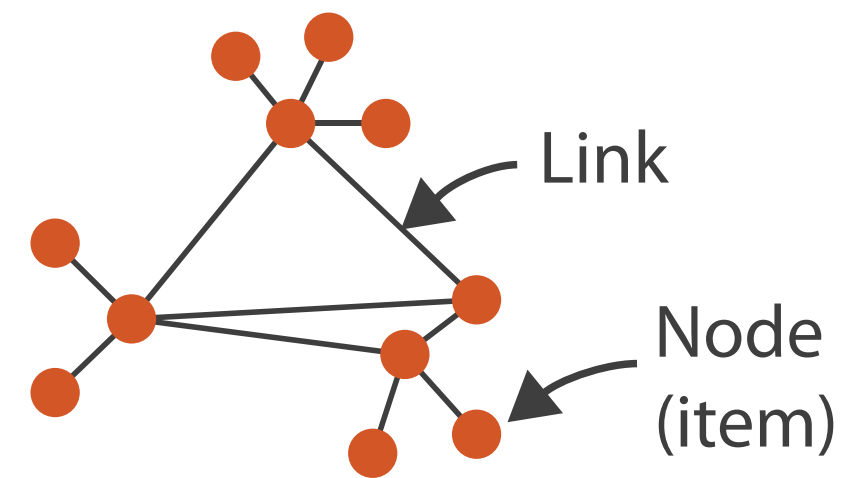
→ Tables



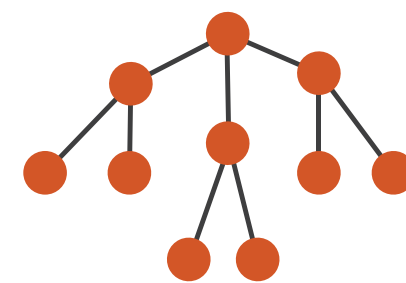
→ Multidimensional Table



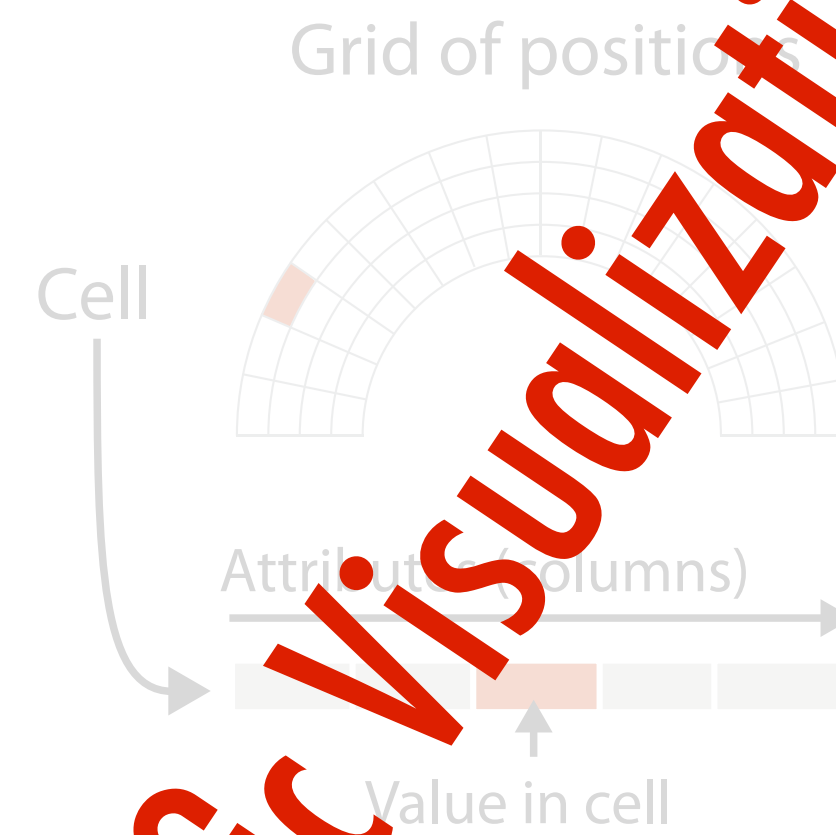
→ Networks



→ Trees



→ Fields (Continuous)

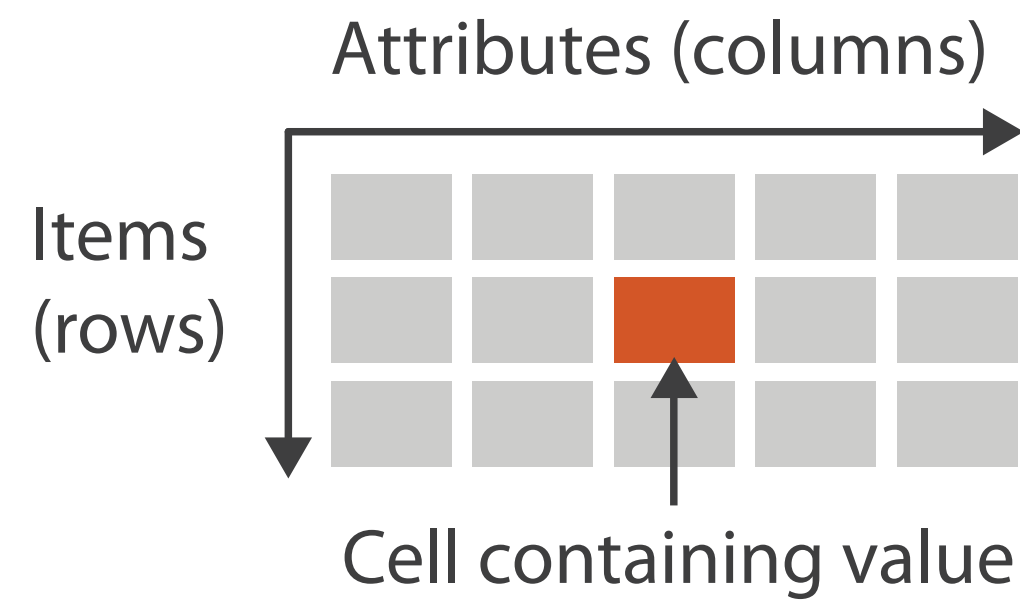


Scientific Visualization

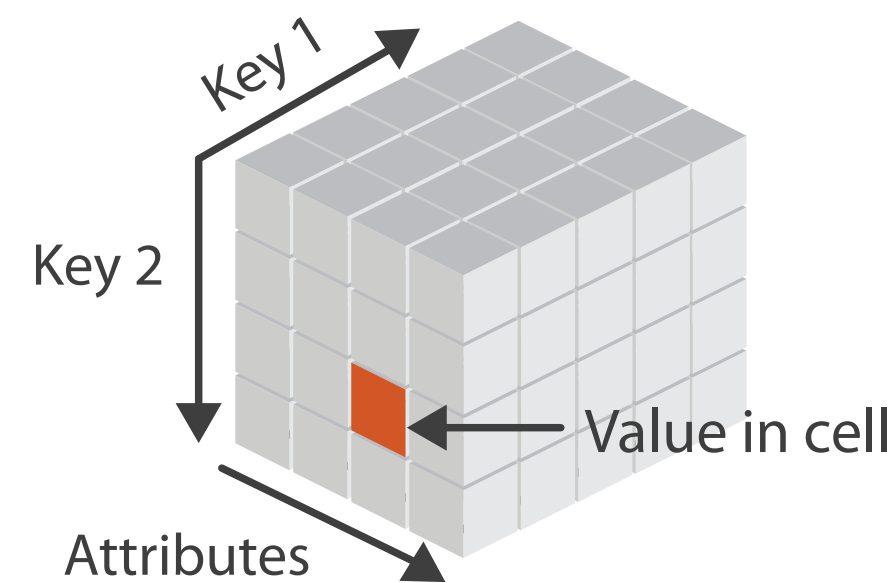
Dataset types

→ Dataset Types

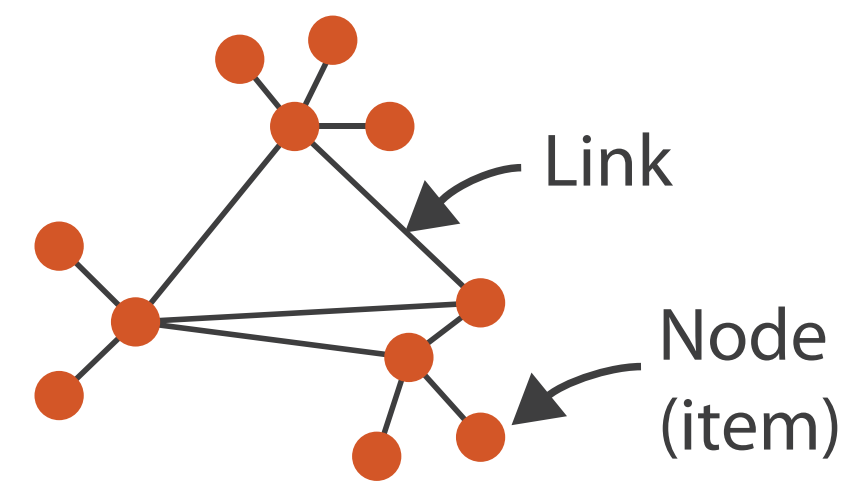
→ Tables



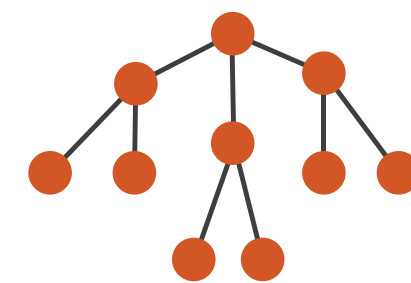
→ *Multidimensional Table*



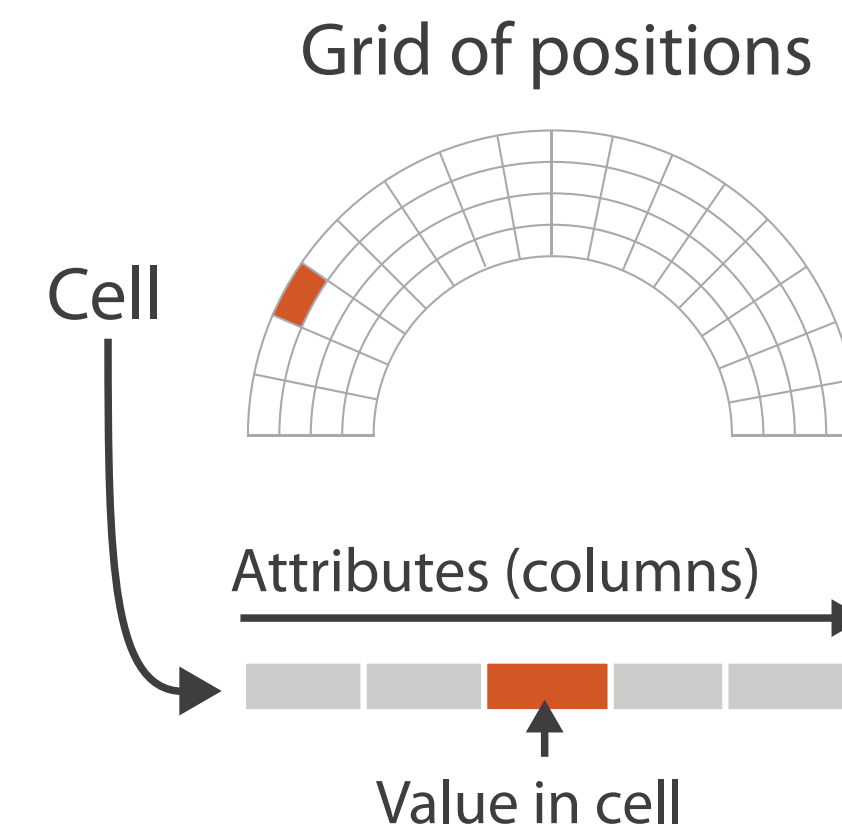
→ Networks



→ Trees



→ Fields (Continuous)



→ Geometry (Spatial)



Attribute types

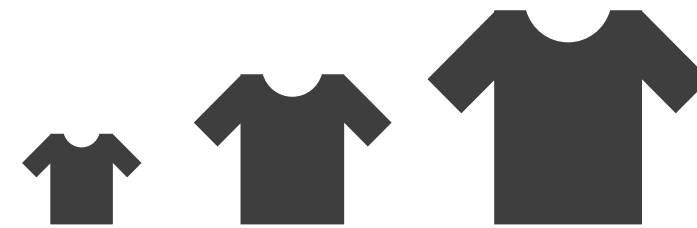
→ Attribute Types

→ Categorical



→ Ordered

→ *Ordinal*



→ *Quantitative*



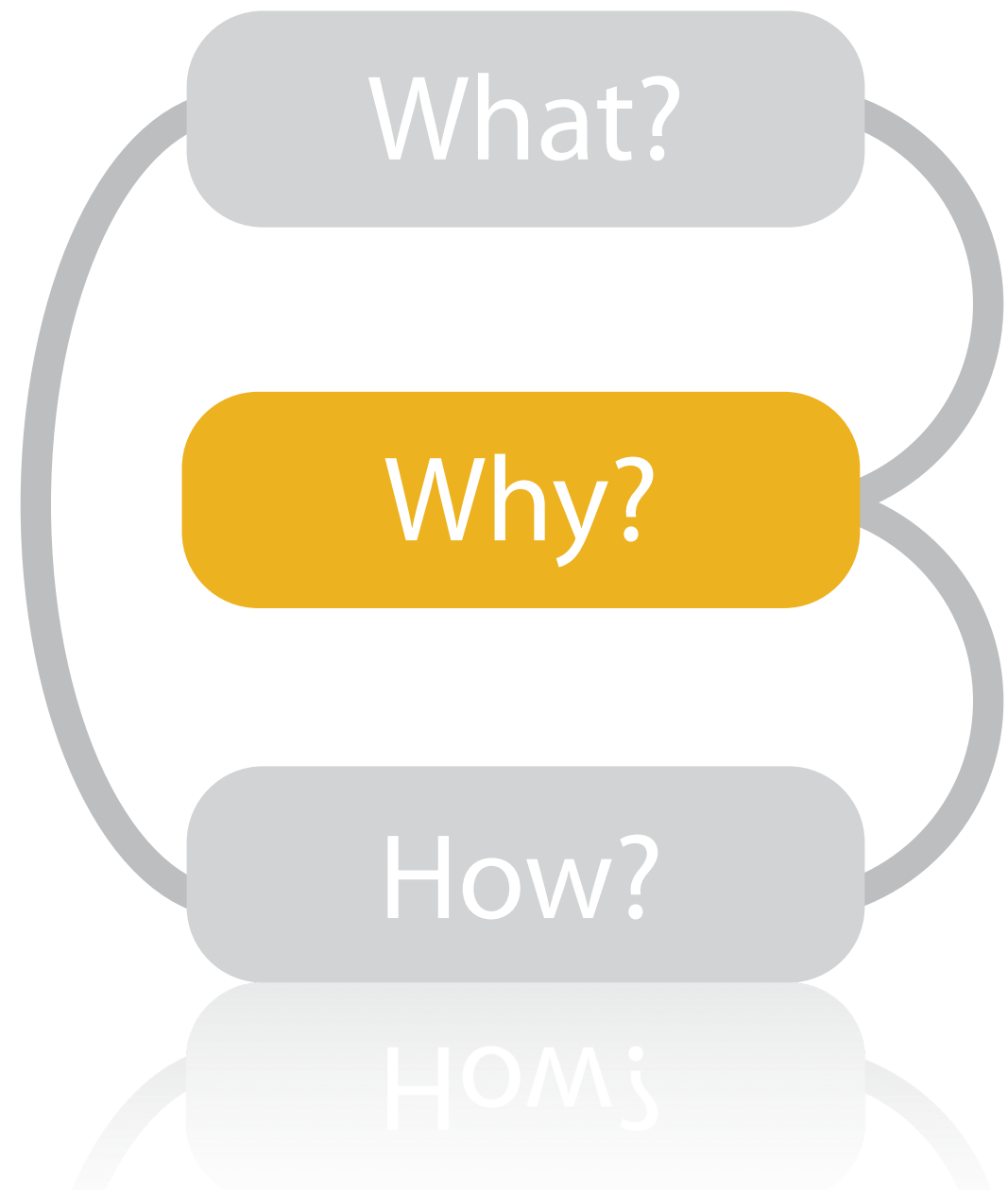
Why: Targets



Why?

👉 Actions	🎯 Targets									
<p>➔ Analyze</p> <ul style="list-style-type: none"> ➔ Consume <ul style="list-style-type: none"> ➔ Discover ➔ Present ➔ Enjoy ➔ Produce <ul style="list-style-type: none"> ➔ Annotate ➔ Record ➔ Derive 	<p>➔ All Data</p> <ul style="list-style-type: none"> ➔ Trends ➔ Outliers ➔ Features <p>➔ Attributes</p> <ul style="list-style-type: none"> ➔ One <ul style="list-style-type: none"> ➔ Distribution ➔ Extremes ➔ Many <ul style="list-style-type: none"> ➔ Dependency ➔ Correlation ➔ Similarity 									
<p>➔ Search</p> <table border="1"> <thead> <tr> <th></th> <th>Target known</th> <th>Target unknown</th> </tr> </thead> <tbody> <tr> <td>Location known</td> <td> <i>Lookup</i></td> <td> <i>Browse</i></td> </tr> <tr> <td>Location unknown</td> <td> <i>Locate</i></td> <td> <i>Explore</i></td> </tr> </tbody> </table>		Target known	Target unknown	Location known	<i>Lookup</i>	<i>Browse</i>	Location unknown	<i>Locate</i>	<i>Explore</i>	<p>➔ Network Data</p> <ul style="list-style-type: none"> ➔ Topology <ul style="list-style-type: none"> ➔ Paths <ul style="list-style-type: none">
	Target known	Target unknown								
Location known	<i>Lookup</i>	<i>Browse</i>								
Location unknown	<i>Locate</i>	<i>Explore</i>								
<p>➔ Query</p> <ul style="list-style-type: none"> ➔ Identify ➔ Compare ➔ Summarise 	<p>➔ Spatial Data</p> <ul style="list-style-type: none"> ➔ Shape 									

Why: Targets



{action, target} pairs

- *discover distribution*
- *compare trends*
- *locate outliers*
- *browse topology*

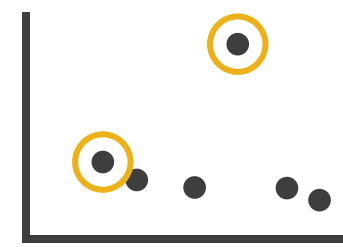
Why?

👉 Actions	🎯 Targets										
<p>👉 Analyze</p> <p>→ Consume</p> <p>→ Discover → Present → Enjoy </p> <p>→ Produce</p> <p>→ Annotate → Record → Derive </p>	<p>👉 All Data</p> <p>→ Trends → Outliers → Features </p>										
<p>👉 Search</p> <table border="1"> <thead> <tr> <th></th> <th>Target known</th> <th>Target unknown</th> </tr> </thead> <tbody> <tr> <td>Location known</td> <td> Lookup</td> <td> Browse</td> </tr> <tr> <td>Location unknown</td> <td> Locate</td> <td> Explore</td> </tr> </tbody> </table>		Target known	Target unknown	Location known	Lookup	Browse	Location unknown	Locate	Explore	<p>👉 Attributes</p> <p>→ One</p> <p>→ Distribution → Extremes </p> <p>→ Many</p> <p>→ Dependency → Correlation → Similarity </p>	
	Target known	Target unknown									
Location known	Lookup	Browse									
Location unknown	Locate	Explore									
<p>👉 Query</p> <p>→ Identify → Compare → Summarise </p>	<p>👉 Network Data</p> <p>→ Topology </p> <p>→ Paths </p>										
	<p>👉 Spatial Data</p> <p>→ Shape </p>										

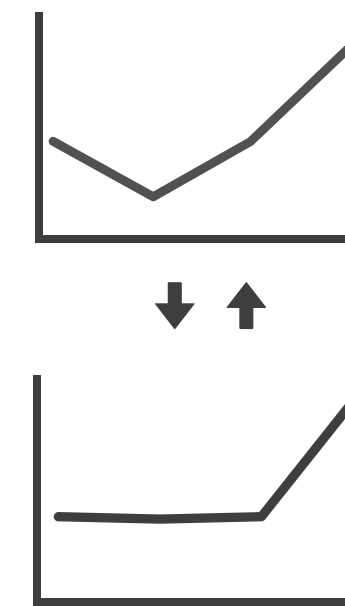
Actions: low-level query

→ Query

→ Identify



→ Compare



→ Summarise

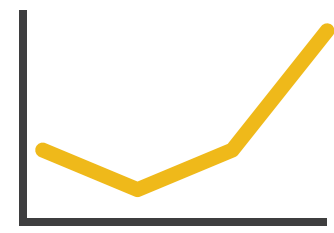


- How much of the data matters?
 - one, some, all

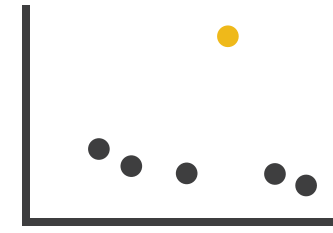
Why: Targets

→ ALL DATA

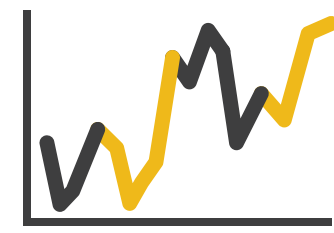
→ Trends



→ Outliers

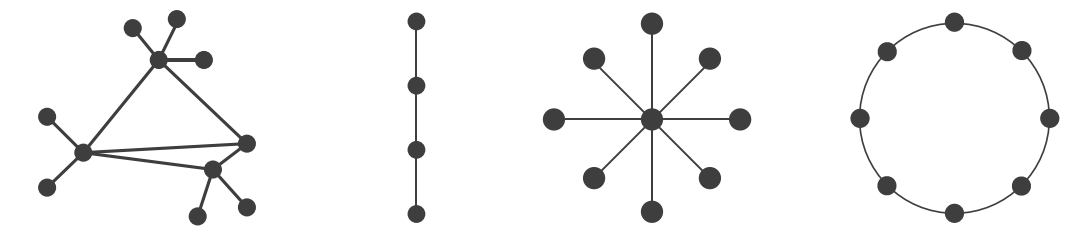


→ Features

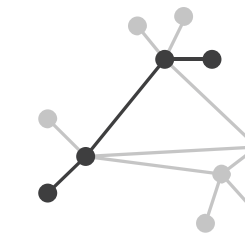


→ NETWORK DATA

→ Topology



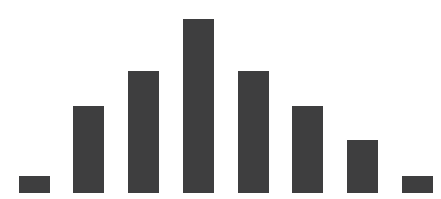
→ Paths



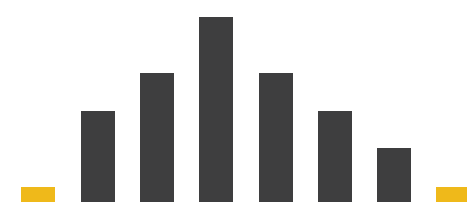
→ ATTRIBUTES

→ One

→ Distribution



↓ Extremes

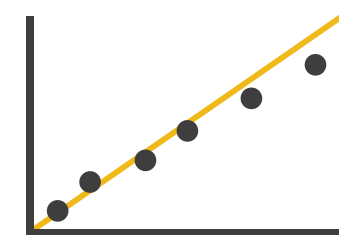


→ Many

→ Dependency



→ Correlation

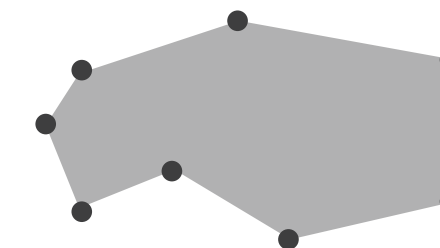


→ Similarity



→ SPATIAL DATA

→ Shape

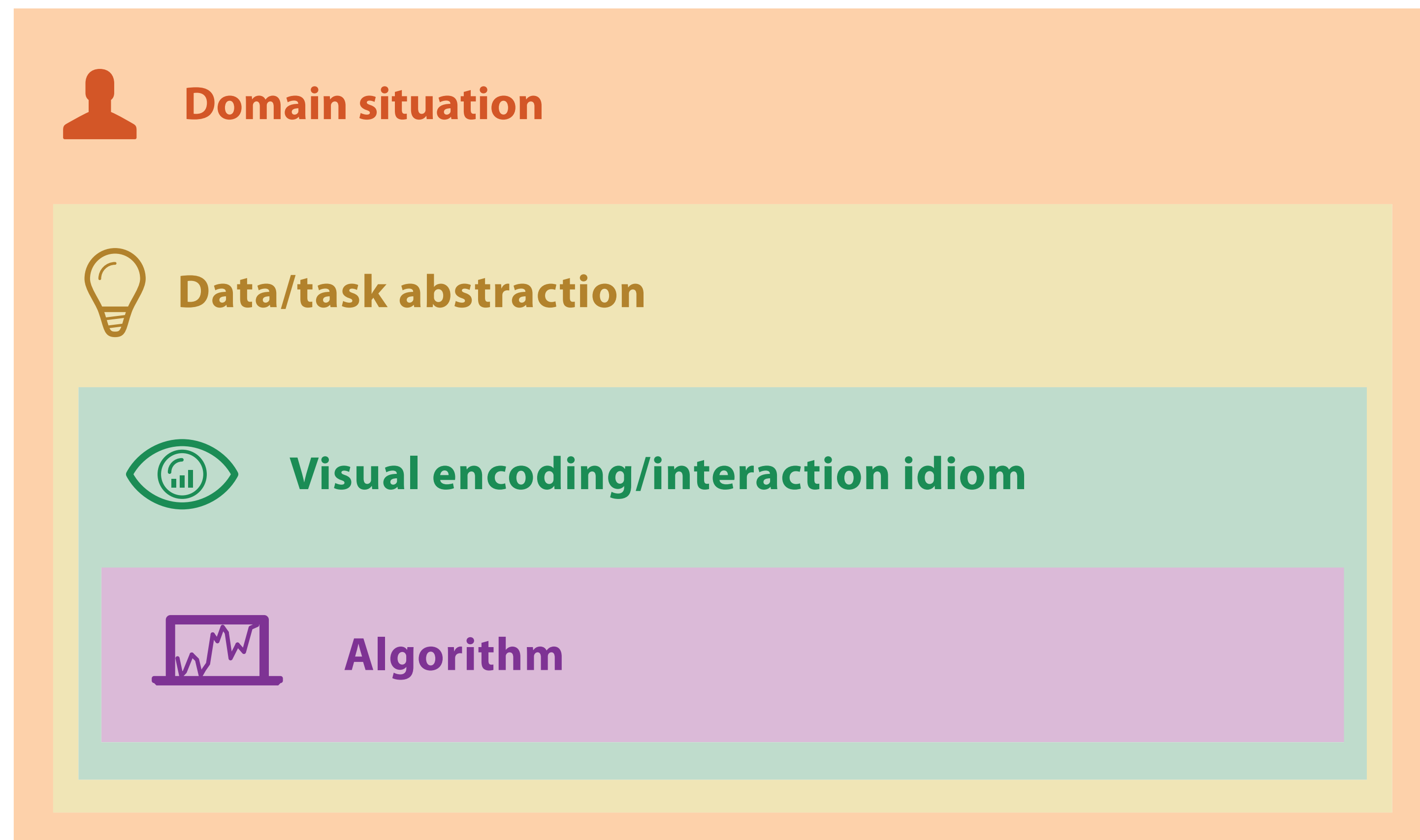


Rules of Thumb

- No unjustified 3D
- Eyes beat memory
- Resolution over immersion
- Overview first, zoom and filter, details on demand
- Function first, form next
- ...

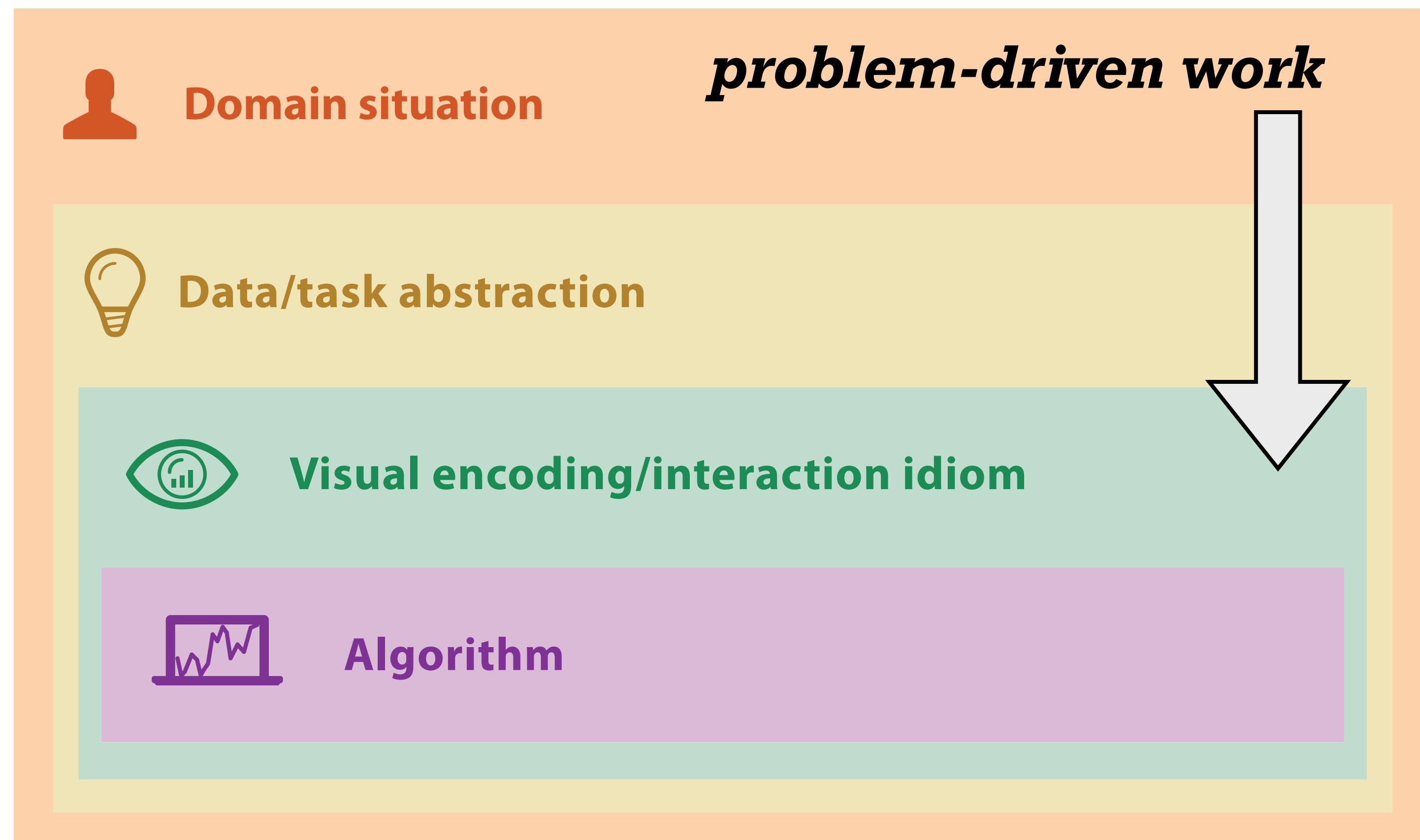
Four Levels of Design

- Domain situation: all aspects of user context
- Data/task abstraction: why/what
- Encoding/interaction idioms: how
- Algorithm: efficient implementation of idioms



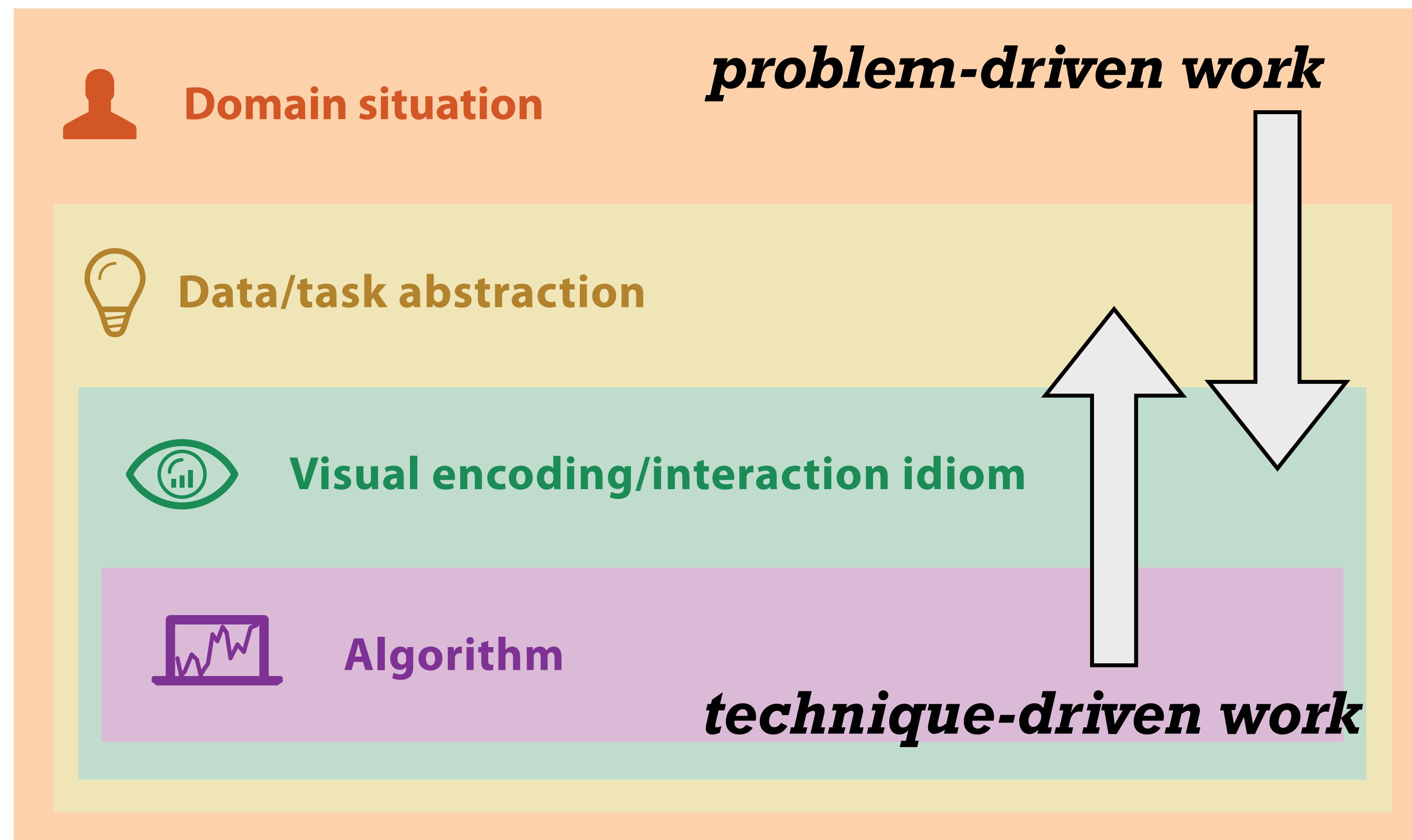
Four Levels of Design

- Domain situation: all aspects of user context
- Data/task abstraction: why/what
- Encoding/interaction idioms: how
- Algorithm: efficient implementation of idioms



Four Levels of Design

- Domain situation: all aspects of user context
- Data/task abstraction: why/what
- Encoding/interaction idioms: how
- Algorithm: efficient implementation of idioms




Nested Levels of Design and 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

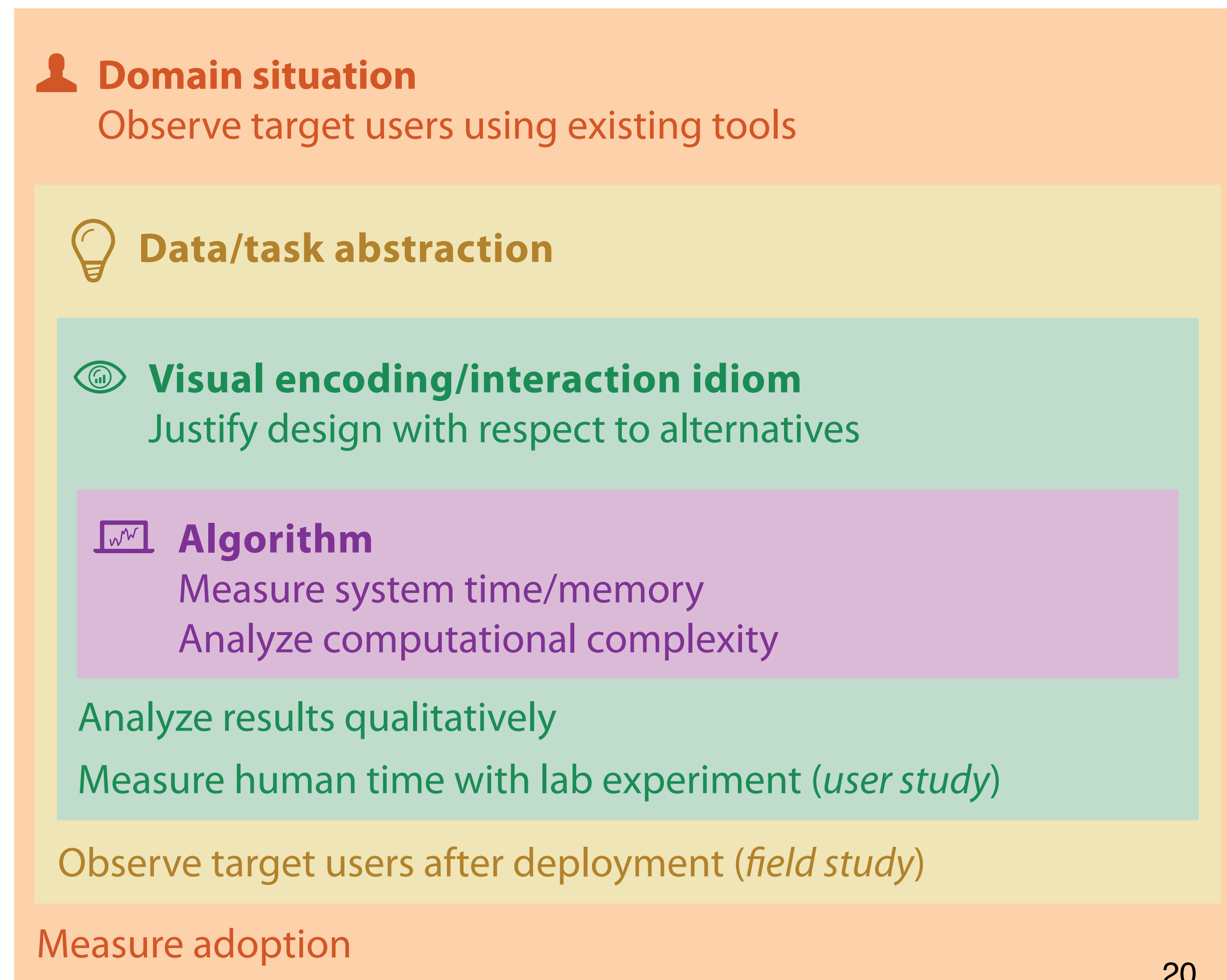
Measure human time with lab experiment (*user study*)

Observe target users after deployment (*field study*)

Measure adoption

Nested Levels of Design and Validation

- Mismatch: cannot show idiom good with system timings
- Mismatch: cannot show abstraction good with lab study



How?

Encode

→ Arrange

→ Express



→ Order



→ Use



→ Separate



→ Align



→ Map

from **categorical** and **ordered** attributes

→ Color

→ Hue



→ Saturation



→ Luminance



→ Size, Angle, Curvature, ...

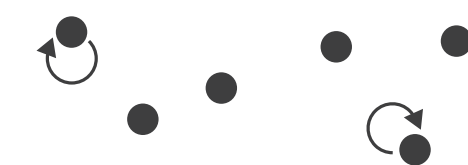


→ Shape



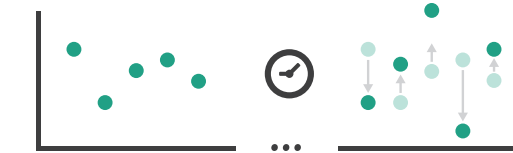
→ Motion

Direction, Rate, Frequency, ...

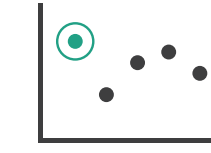


Manipulate

→ Change



→ Select

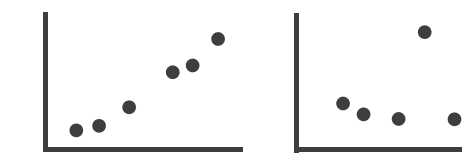


→ Navigate

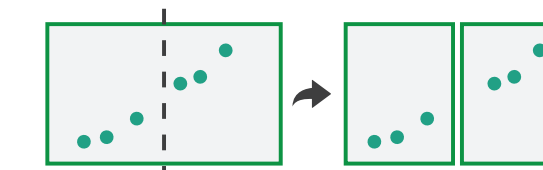


Facet

→ Juxtapose



→ Partition

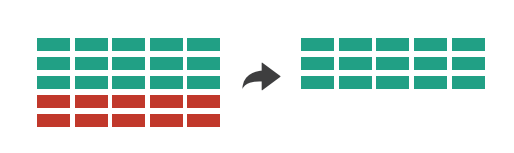


→ Superimpose

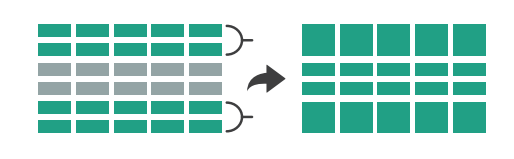


Reduce

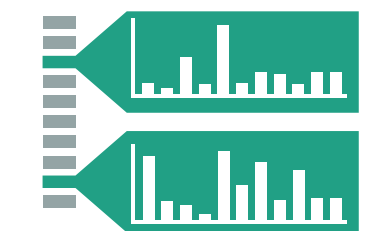
→ Filter



→ Aggregate



→ Embed



What?

Why?

How?

Arrange space

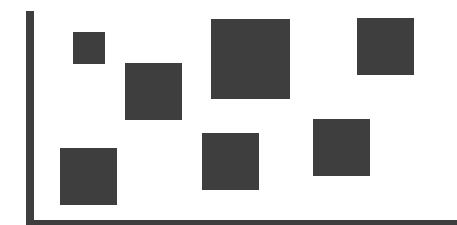
Encode

→ Arrange

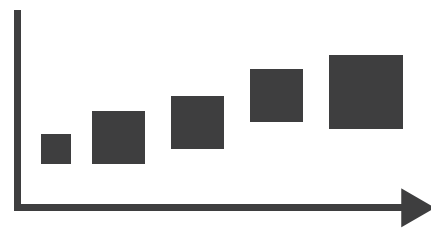
→ Express



→ Separate



→ Order



→ Align



→ Use



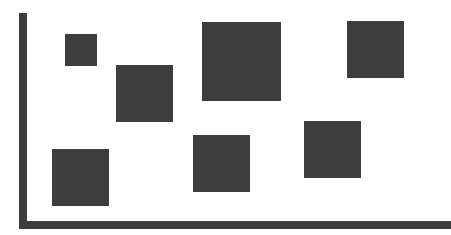
Arrange tables

→ Express Values



→ Separate, Order, Align Regions

→ Separate



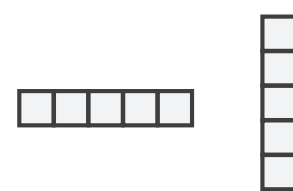
→ Order



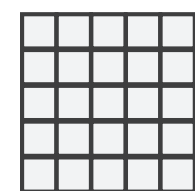
→ Align



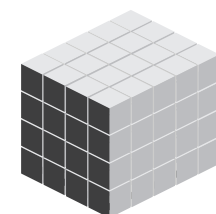
→ 1 Key
List



→ 2 Keys
Matrix



→ 3 Keys
Volume

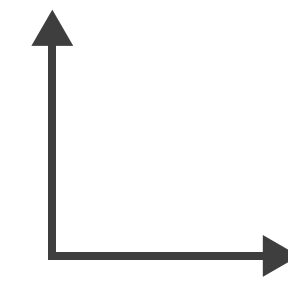


→ Many Keys
Recursive Subdivision

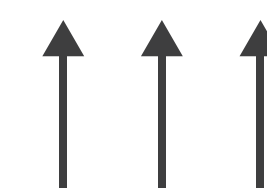


→ Axis Orientation

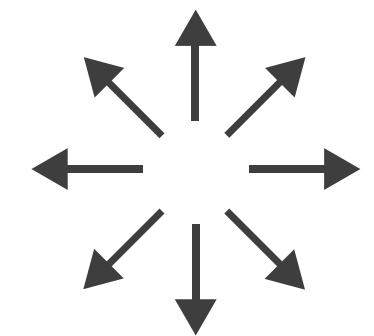
→ Rectilinear



→ Parallel

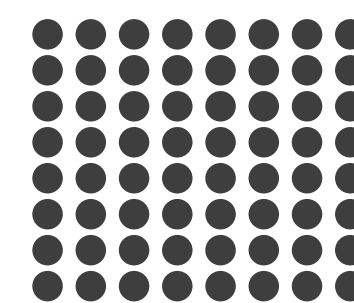


→ Radial



→ Layout Density

→ Dense



→ Space-Filling



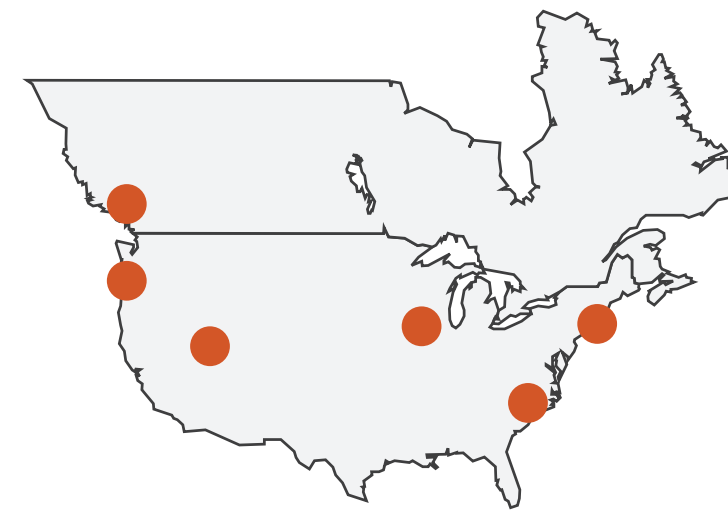
Arrange spatial data

→ Use Given

→ Geometry

→ *Geographic*

→ *Other Derived*

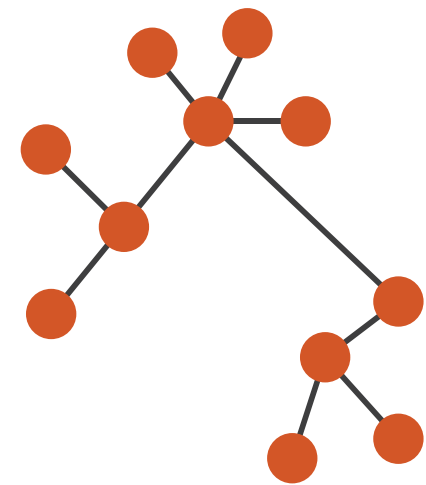


Arrange networks and trees

→ Node-link Diagrams Connections and Marks

✓ NETWORKS

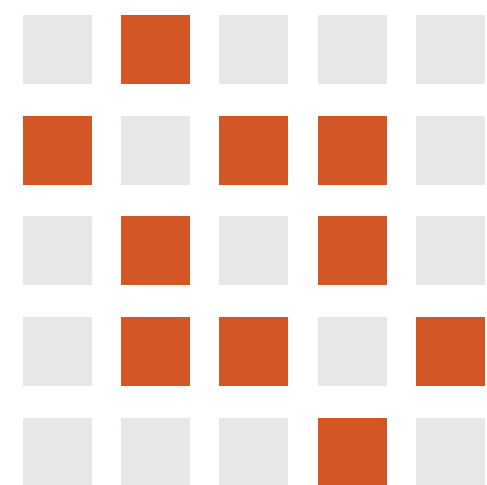
✓ TREES



→ Adjacency Matrix Derived Table

✓ NETWORKS

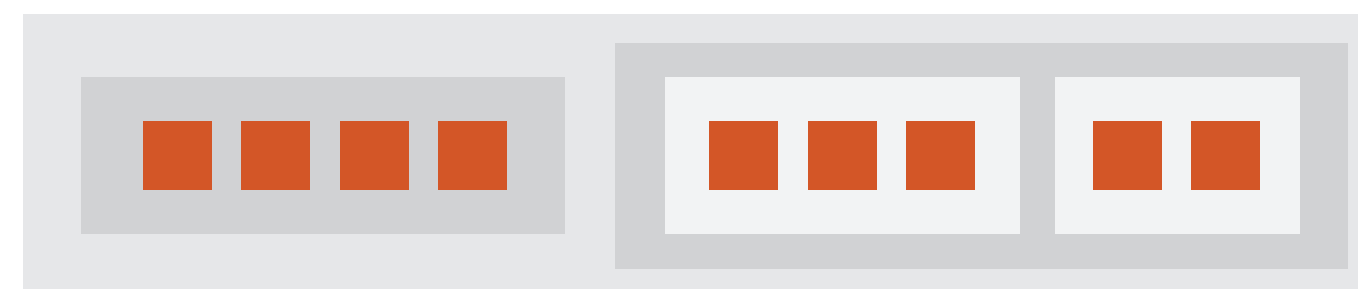
✓ TREES



→ Enclosure Containment Marks

✗ NETWORKS

✓ TREES



Color: Luminance, saturation, hue

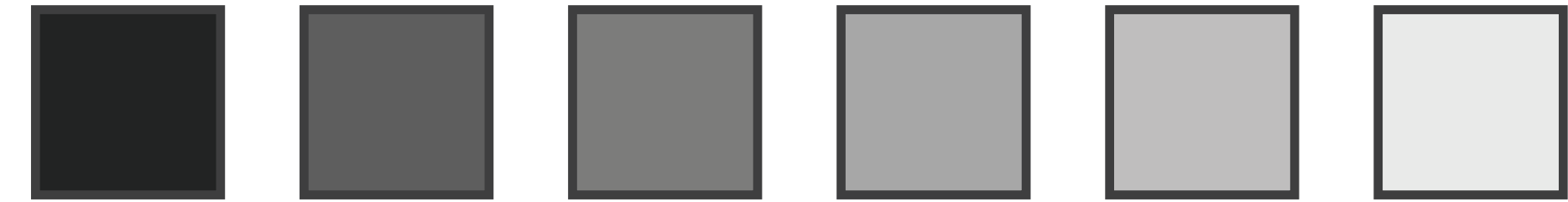
- **3 channels**

- identity for categorical
 - hue
- magnitude for ordered
 - luminance
 - saturation

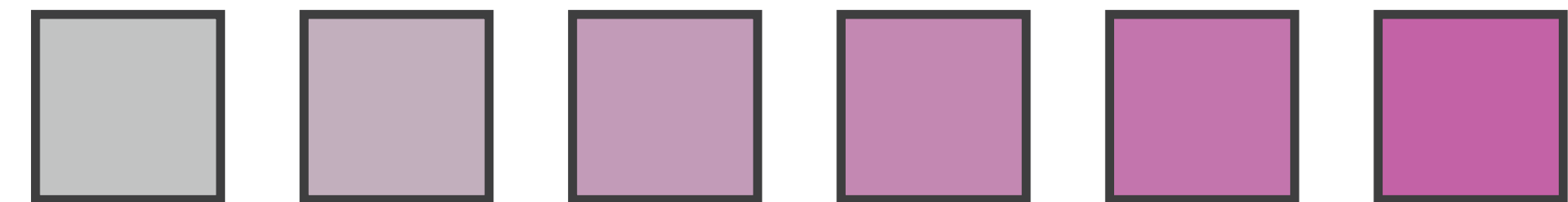
- **Other common color spaces**

- RGB: poor choice for visual encoding
- HSL: better, but beware
 - lightness \neq luminance

Luminance



Saturation



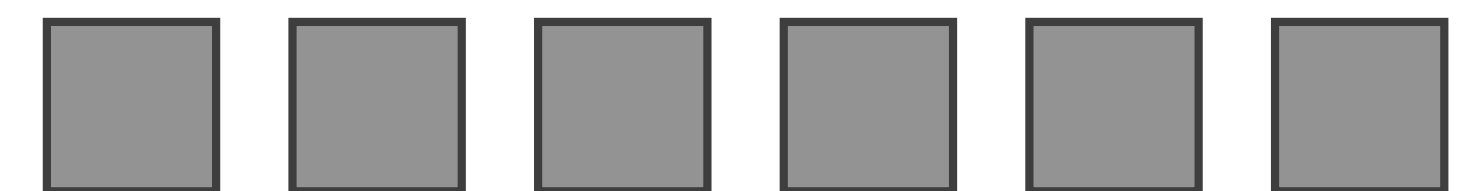
Hue



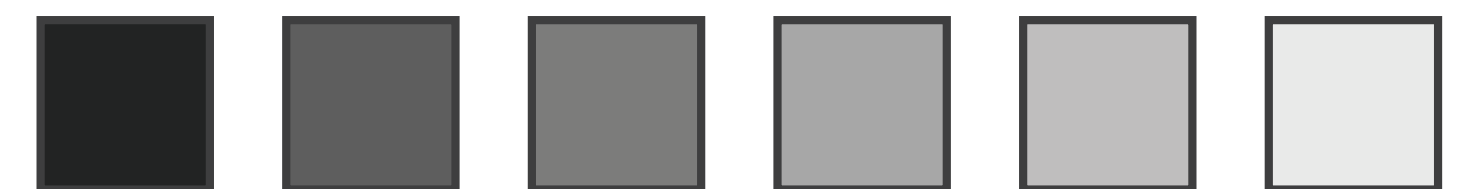
Corners of the RGB
color cube



L from HLS
All the same

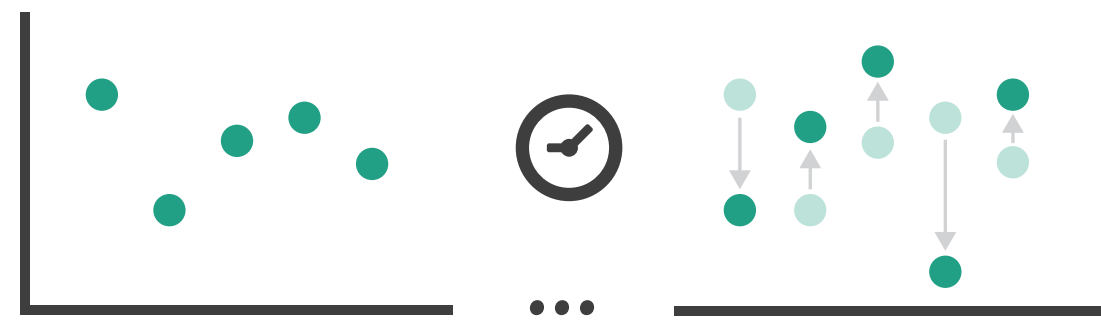


Luminance values

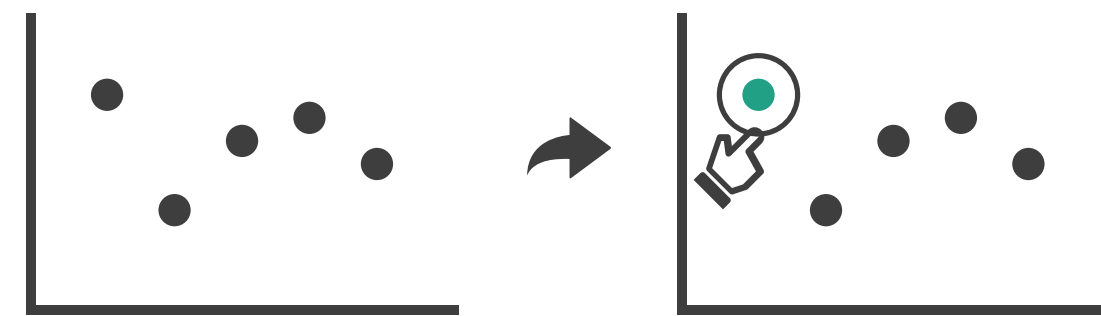


Manipulate

→ Change View Over Time



→ Select

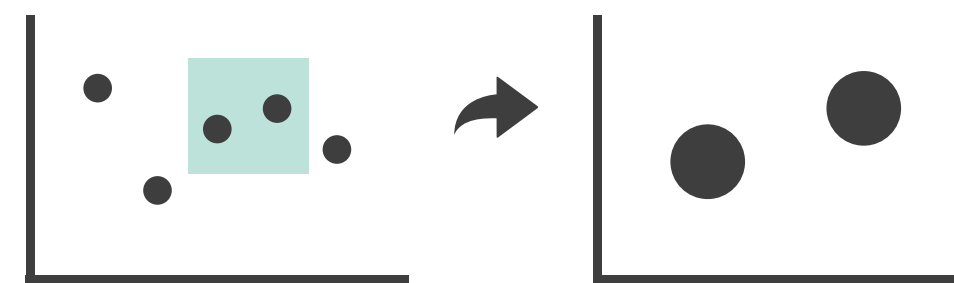


→ Navigate

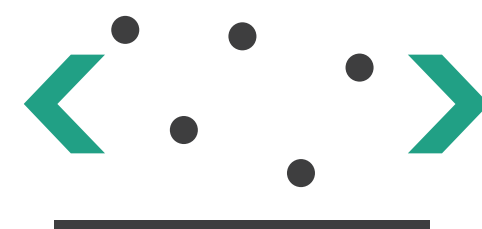
→ Item Reduction

→ Zoom

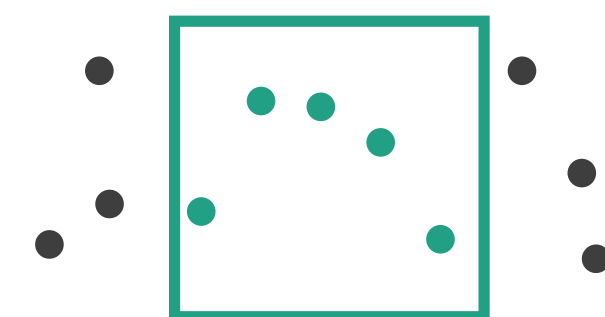
Geometric or Semantic



→ Pan/Translate

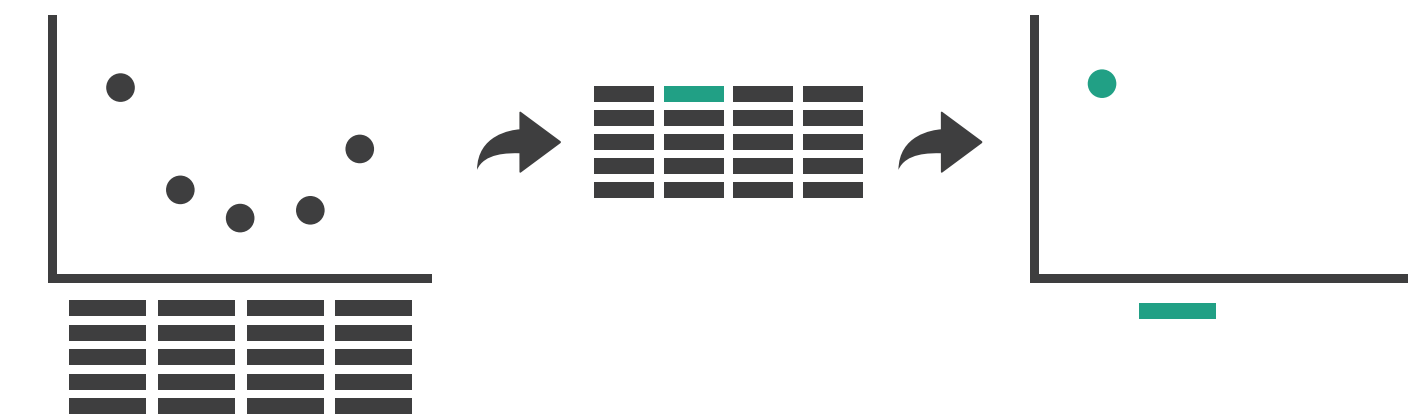


→ Constrained

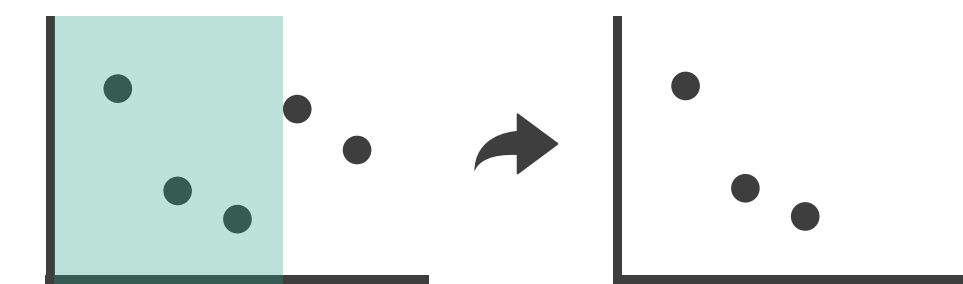


→ Attribute Reduction

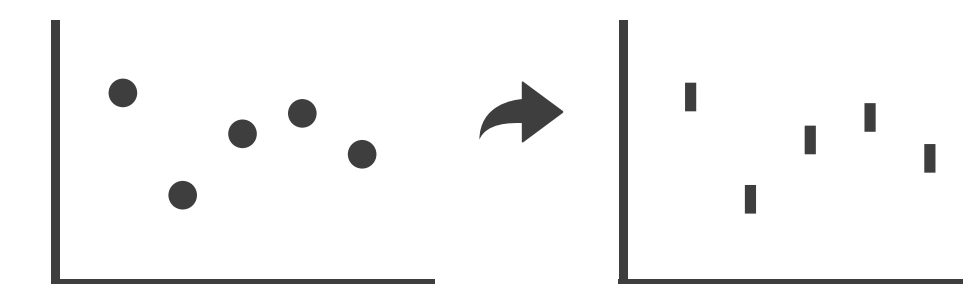
→ Slice



→ Cut

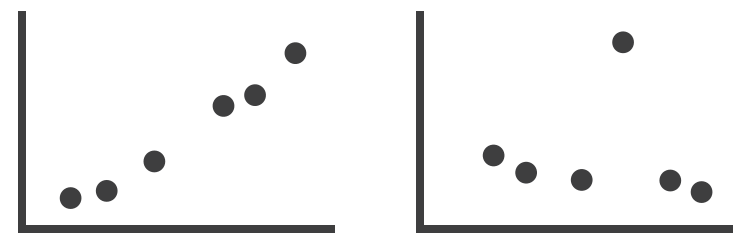


→ Project

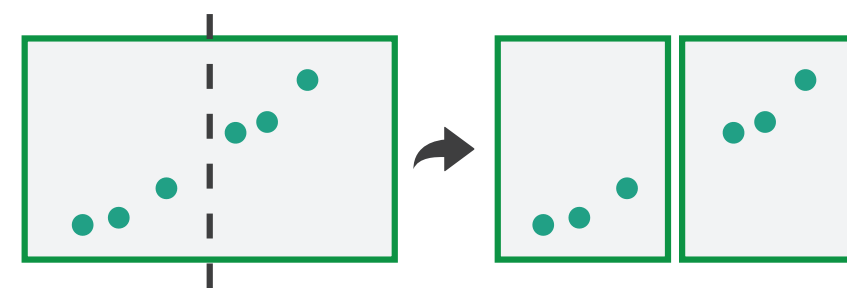


Facet

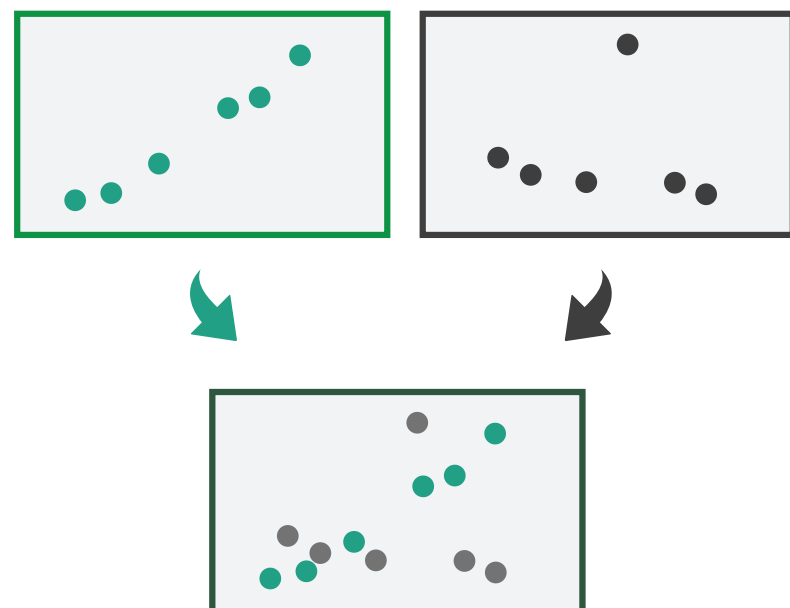
→ Juxtapose



→ Partition



→ Superimpose



→ Share Encoding: Same/Different

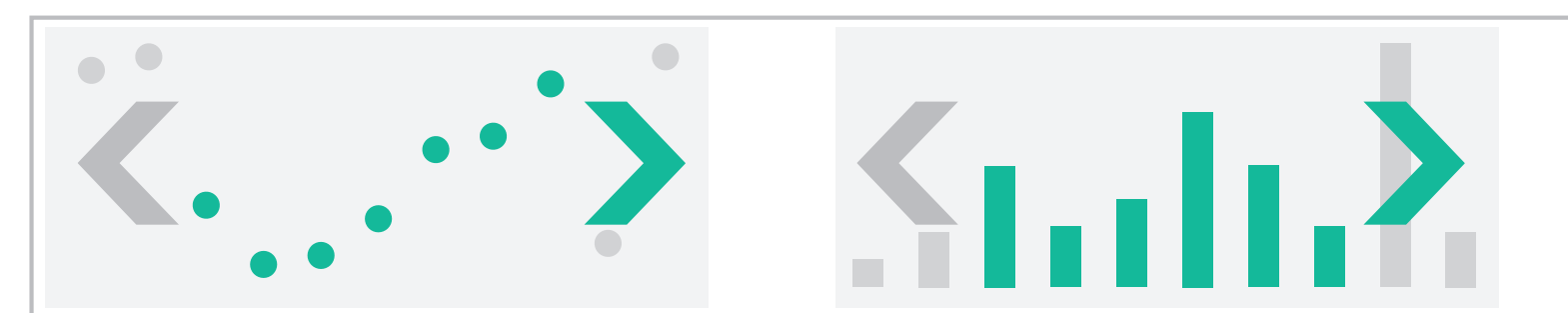
→ *Linked Highlighting*



→ Share Data: All/Subset/None



→ Share Navigation



Juxtapose and coordinate views

→ Share Encoding: Same/Different

→ *Linked Highlighting*



→ Share Data: All/Subset/None



→ Share Navigation

