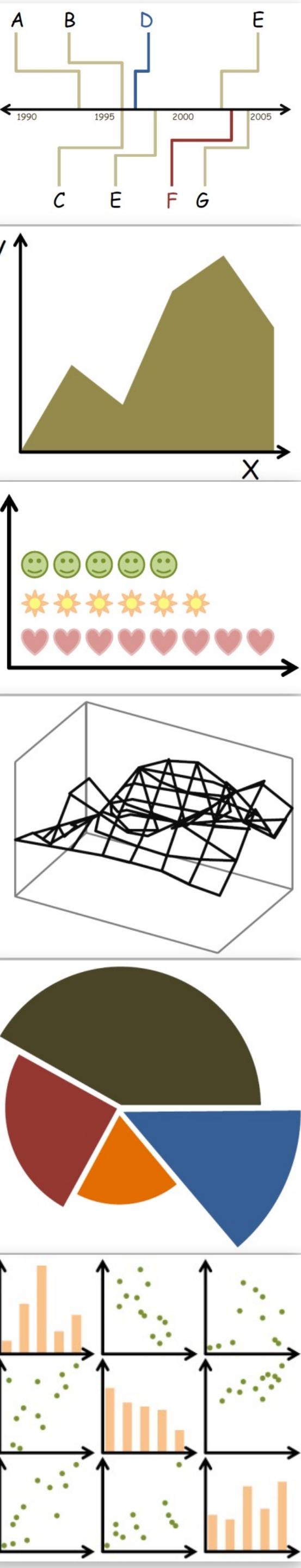


# Lecture 3: Data Abstraction

DS 4200  
FALL 2022

Prof. Ab Mosca (*they/them*)  
NORTHEASTERN UNIVERSITY



Slides and inspiration from Cody Dunne, Michelle Borkin, Dylan Cashman, Krzysztof Gajos, Hanspeter Pfister, Miriah Meyer, Jonathan Schwabish, and David Sprague, Ashley Suh

# Last Class

**We:**

- Reviewed Munzner's Nested Model of Visualization Design
- Finished ic-01, setting up Git

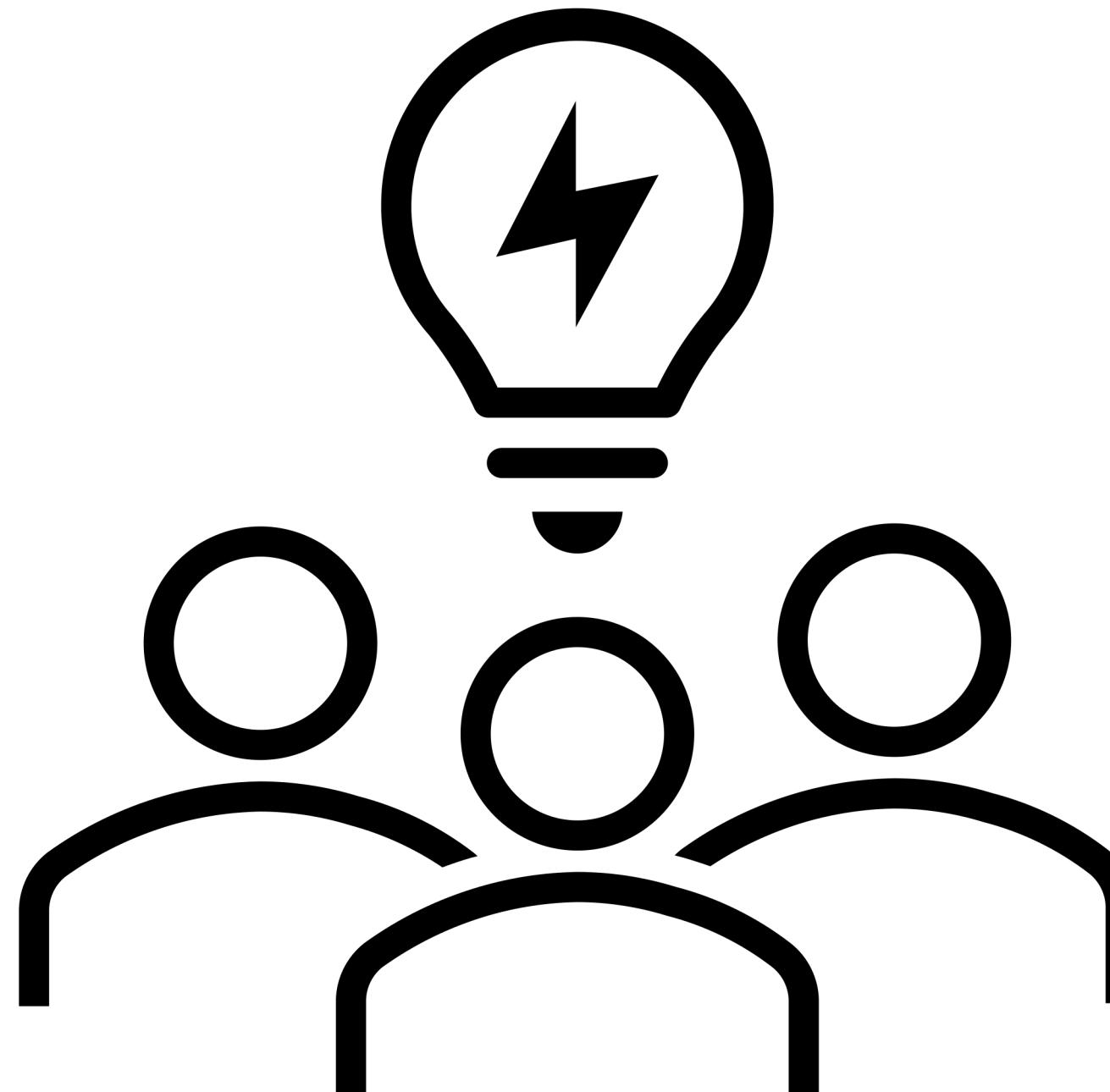
**Any Questions?**

# Today

- Visualization building blocks
- Data Abstraction
- HTML & CSS

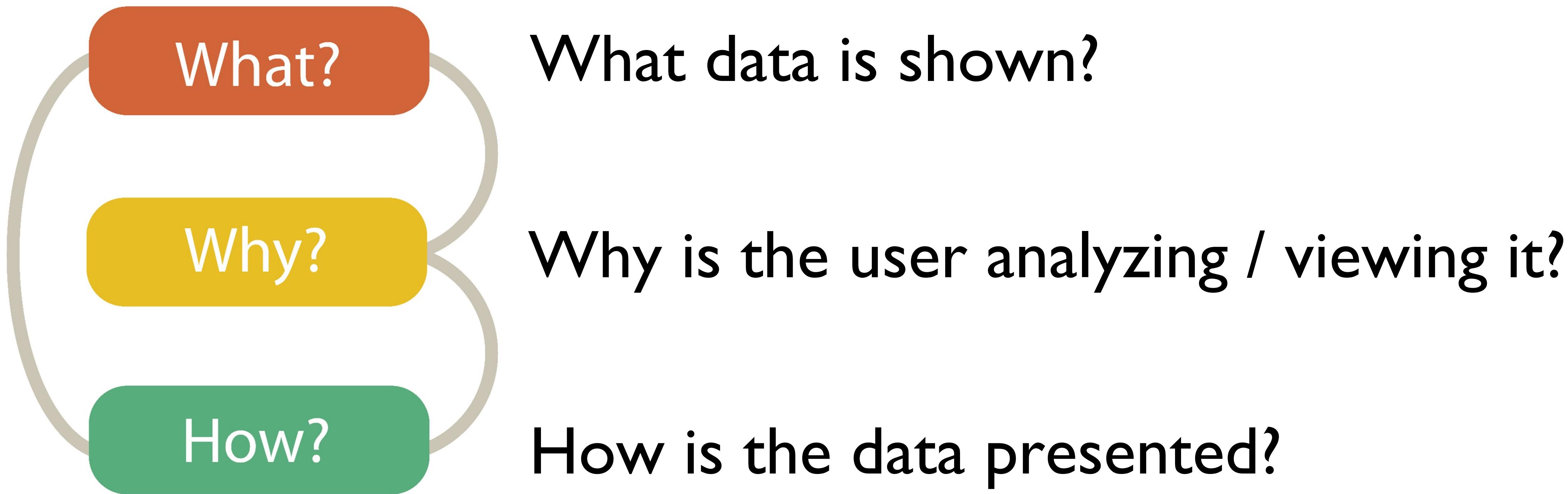
# VISUALIZATION BUILDING BLOCKS

# Visualization Building Blocks

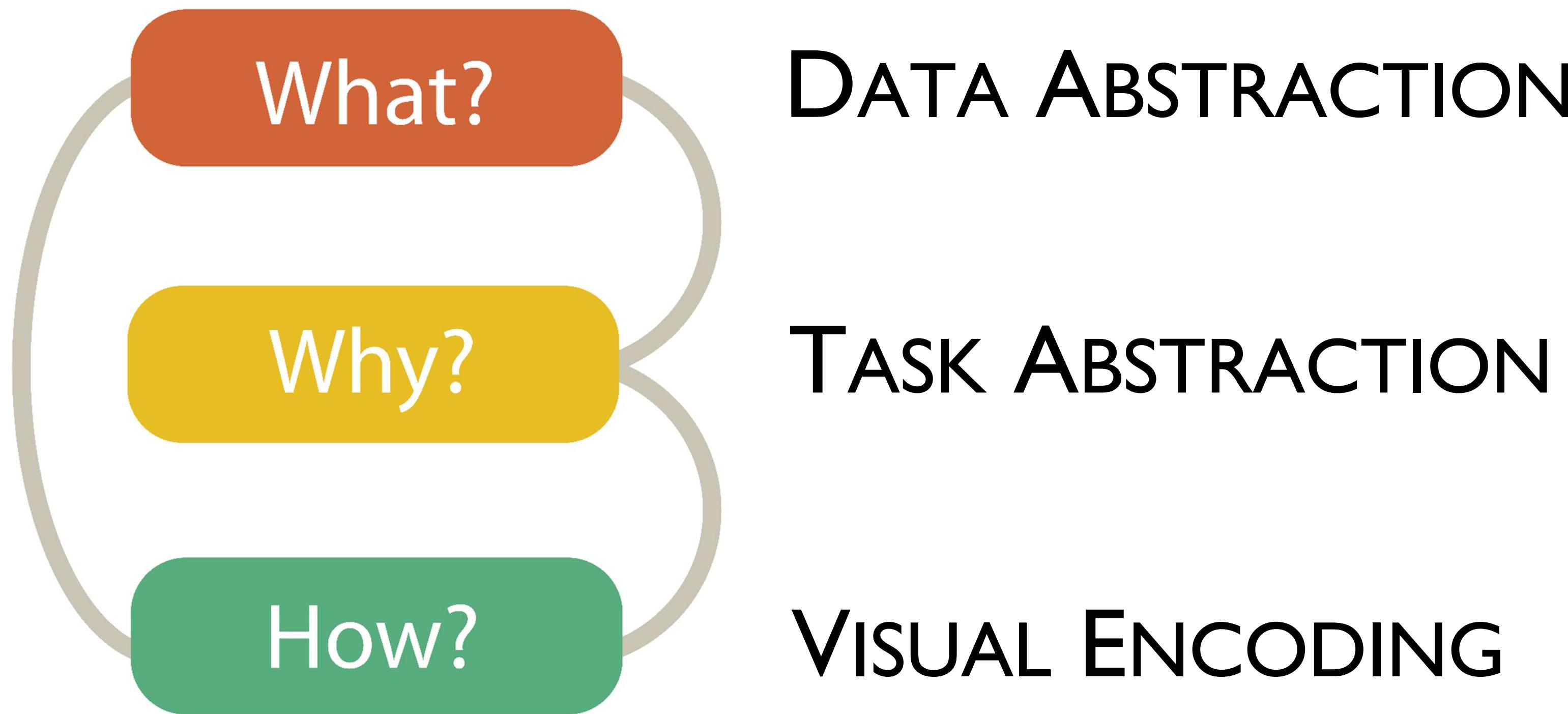


- Talk to your neighbors and **brainstorm what you think the essential pieces for building a visualization are**
- Add your thoughts to the [Vis-Building-Blocks](#) Jamboard

# Visualization Building Blocks

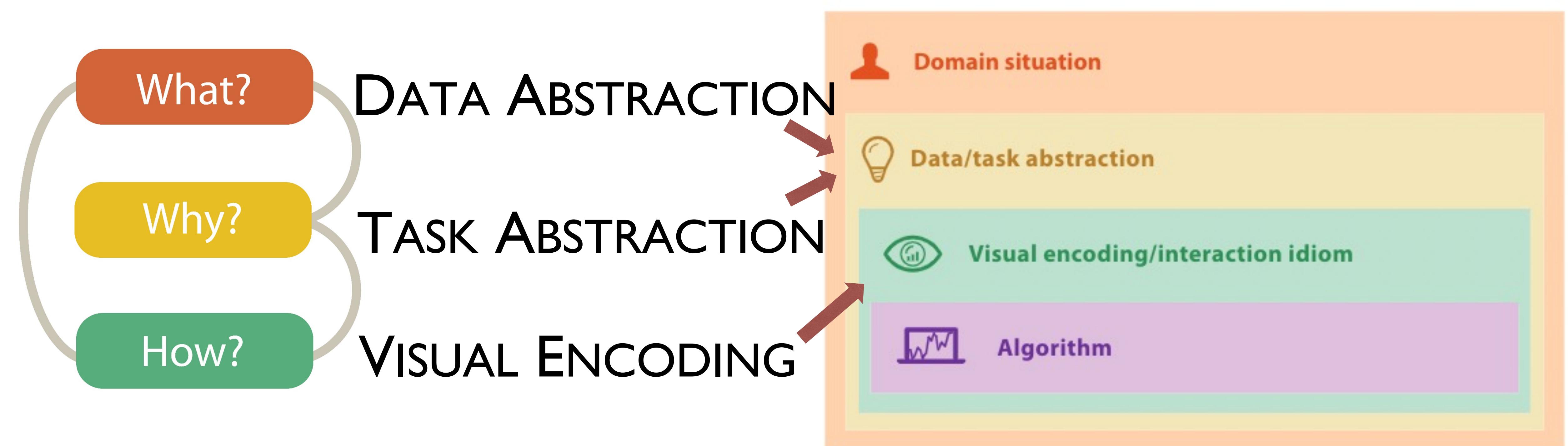


# Visualization Building Blocks

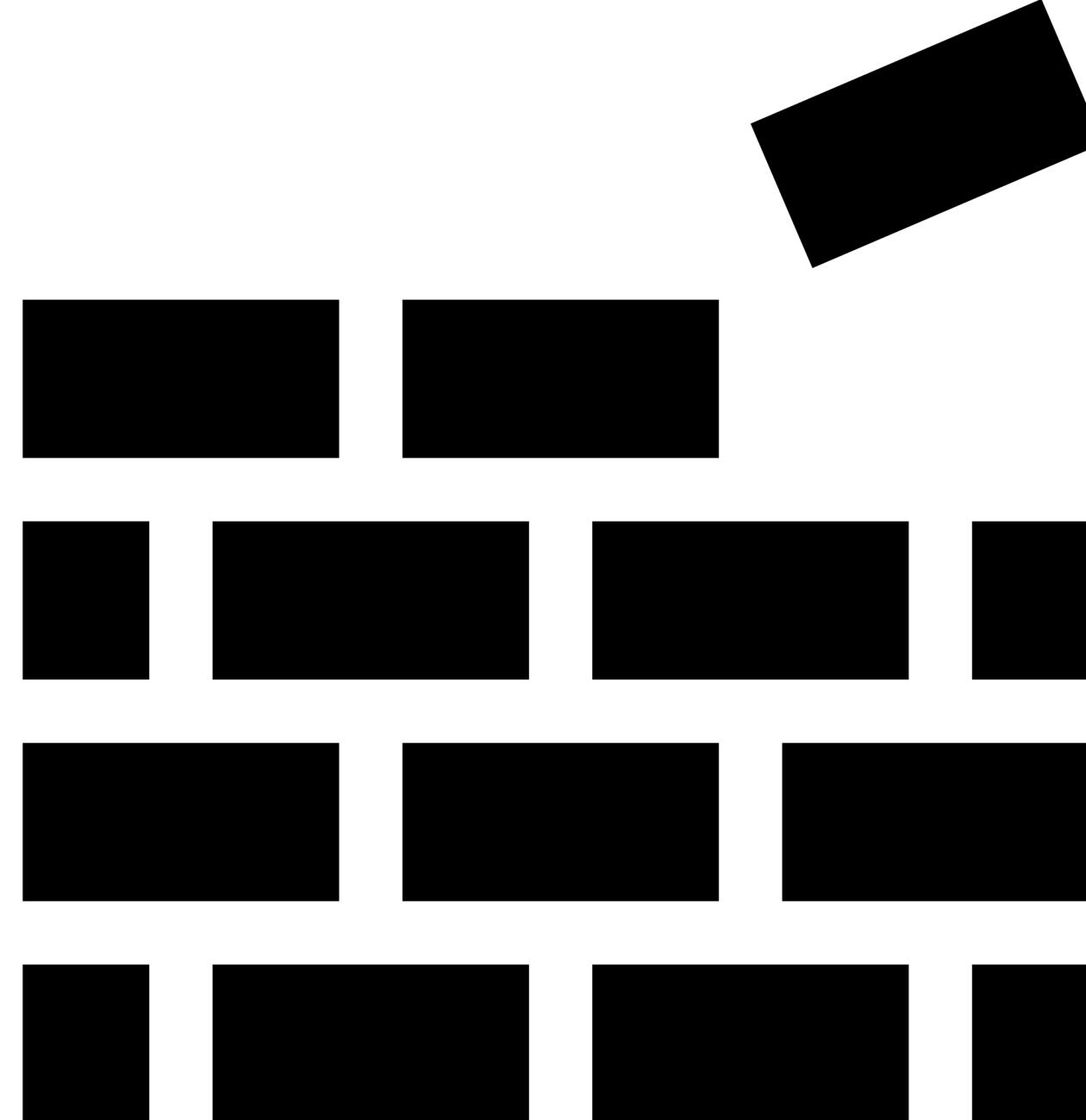


# Visualization Building Blocks

## Munzner's Nested Model

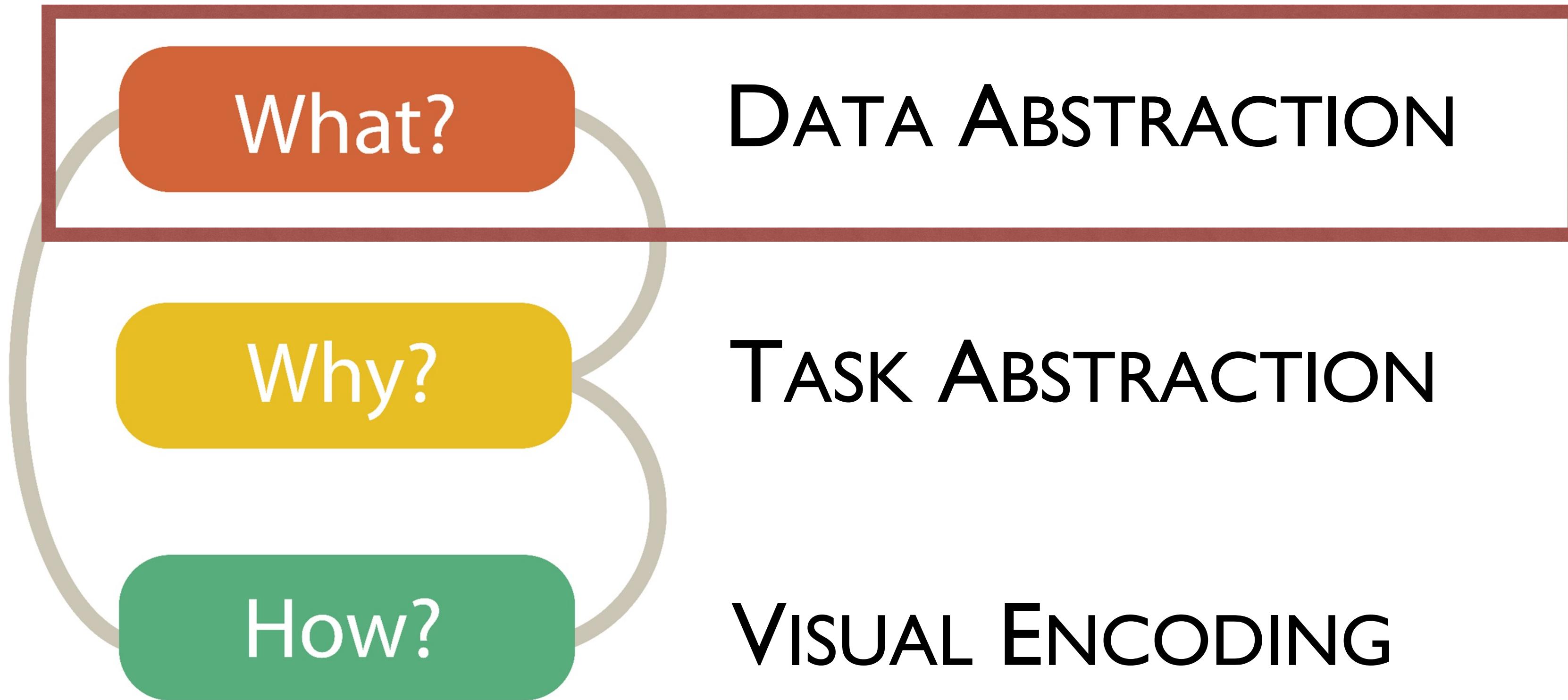


# Visualization Building Blocks



- **Data Abstraction (what)**
  - Source, biases, ethics, quality
  - Types, sets
- **Task Abstraction (why)**
  - What does the user need to do?
  - What does the visualization need to do?
- **Visual Encodings (how)**
  - How is data represented on the screen?

# Visualization Building Blocks



# DATA ABSTRACTION

From Munzner's book

# Data Types

**DATA TYPE** = structural or mathematical interpretation of the data

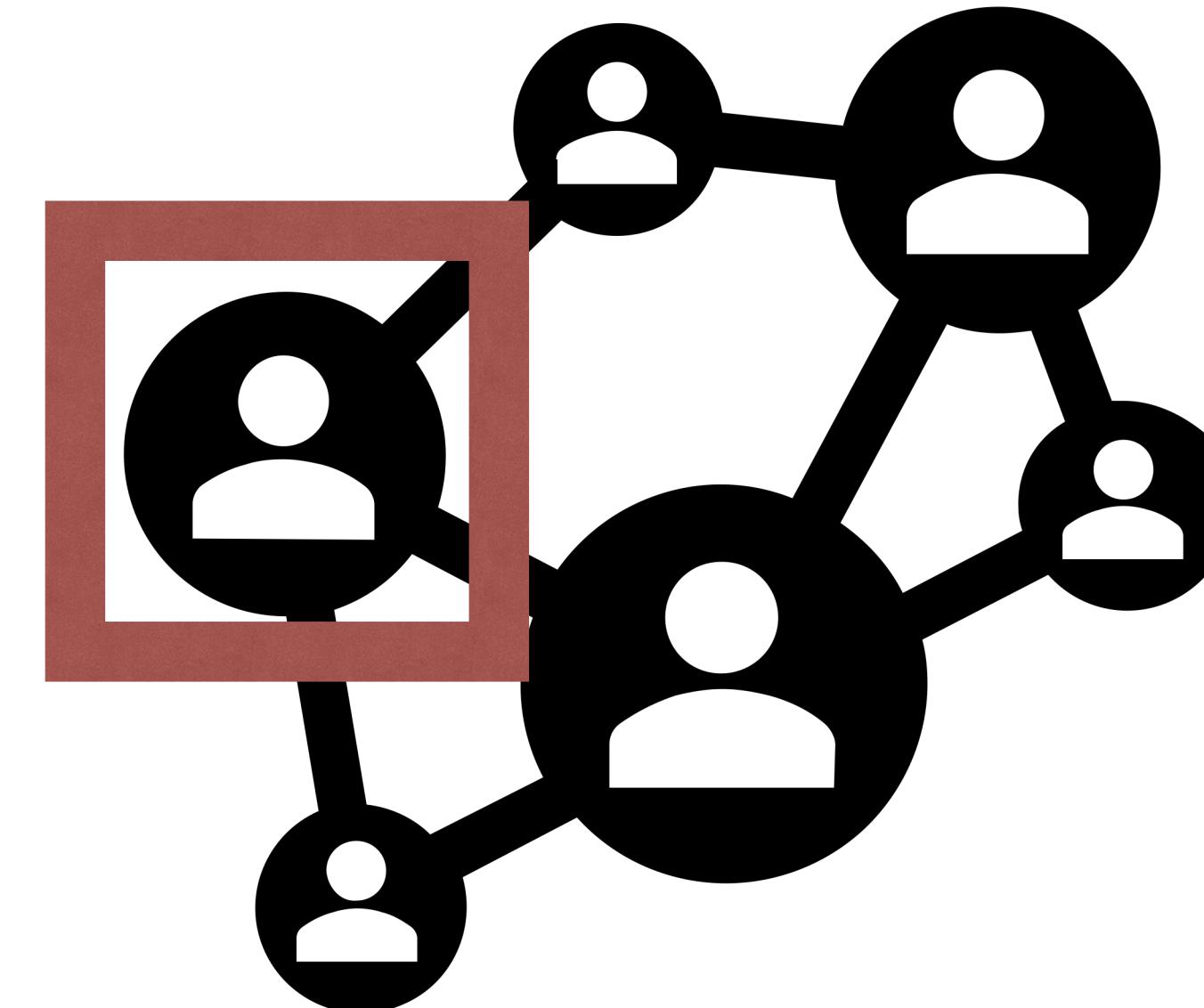
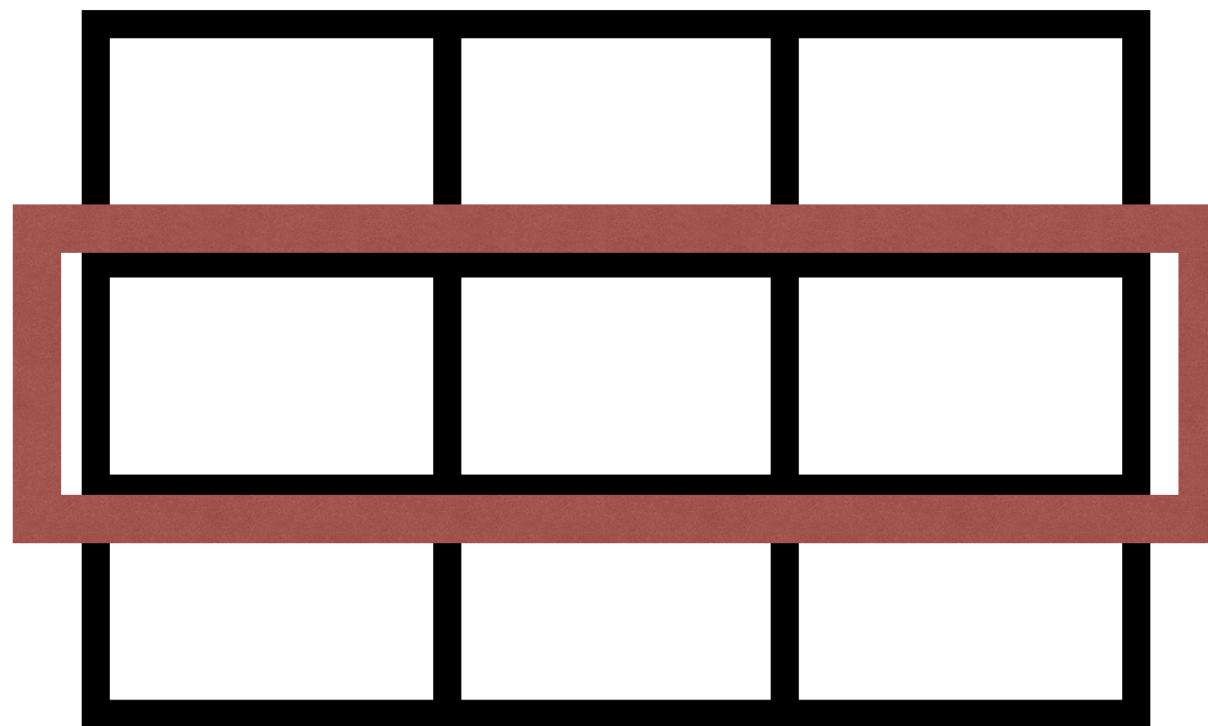
→ Data Types

→ Items      → Attributes      → Links      → Positions      → Grids

# Data Types

## DATA TYPE: Items

- An individual entity that is discrete
- Ex. Row in table, node in network



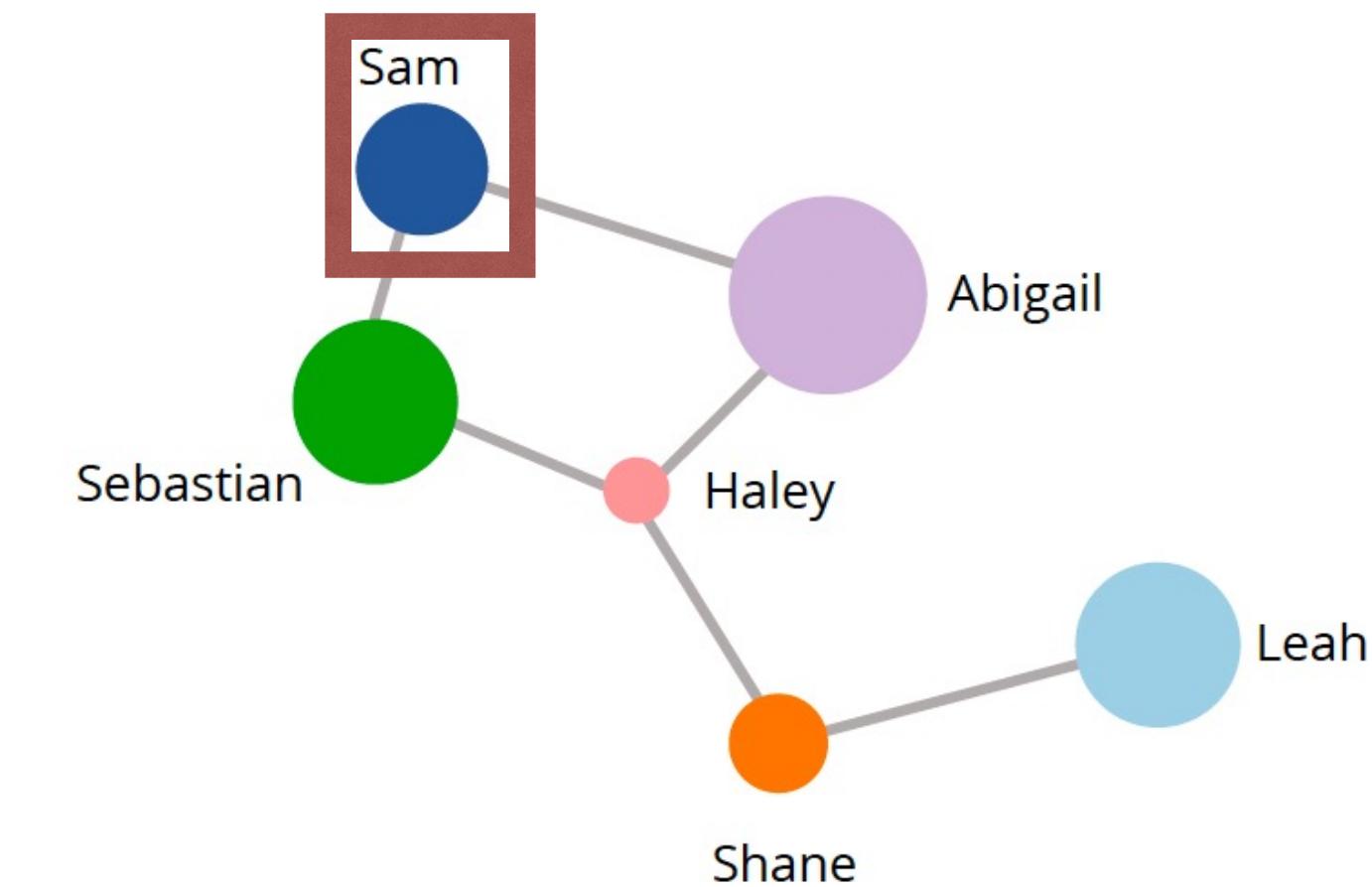
# Data Types

## DATA TYPE: Items

- An individual entity that is discrete
- Ex. Row in table, node in network

Museum ID	Museum Name	Museum Type
8400200098	ALASKA AVIATION HERITAGE MUSEUM	HISTORY MUSEUM
8400200117	ALASKA BOTANICAL GARDEN	ARBORETUM, BOTANICAL GARDEN, OR NATUR
8400200153	ALASKA CHALLENGER CENTER FOR SPACE SCIENCE TECHNOLOGY	SCIENCE & TECHNOLOGY MUSEUM OR PLANET
8400200143	ALASKA EDUCATORS HISTORICAL SOCIETY	HISTORIC PRESERVATION
8400200027	ALASKA HERITAGE MUSEUM	HISTORY MUSEUM
8400200096	ALASKA HISTORICAL MUSEUM	HISTORIC PRESERVATION
8400200078	ALASKA JEWISH MUSEUM	GENERAL MUSEUM
8400200084	ALASKA LIGHTHOUSE ASSOCIATION	HISTORIC PRESERVATION
8400200107	ALASKA MASONIC LIBRARY AND MUSEUM FOUNDATION	GENERAL MUSEUM
8400200073	ALASKA MINING HALL OF FAME FOUNDATION	HISTORY MUSEUM
8400200131	ALASKA NATIONAL GUARD HISTORICAL HOLDING-MUSEUM	HISTORIC PRESERVATION
8400200123	ALASKA NATIVE HERITAGE CENTER	GENERAL MUSEUM

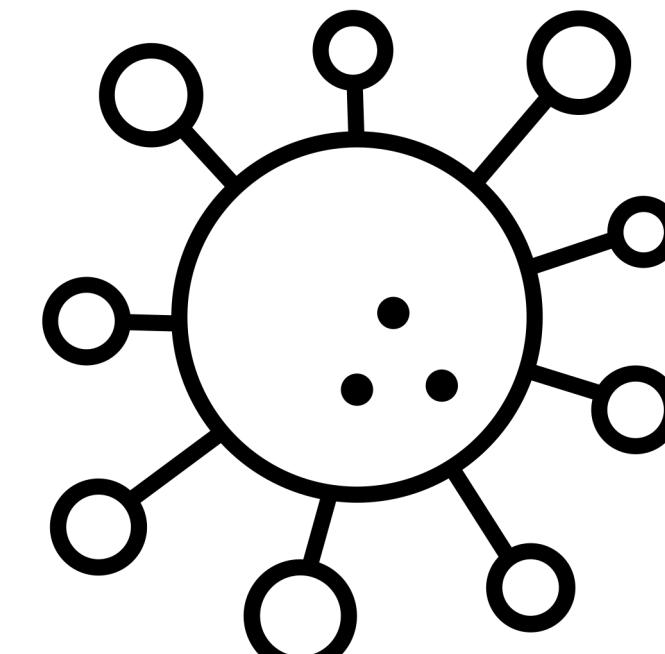
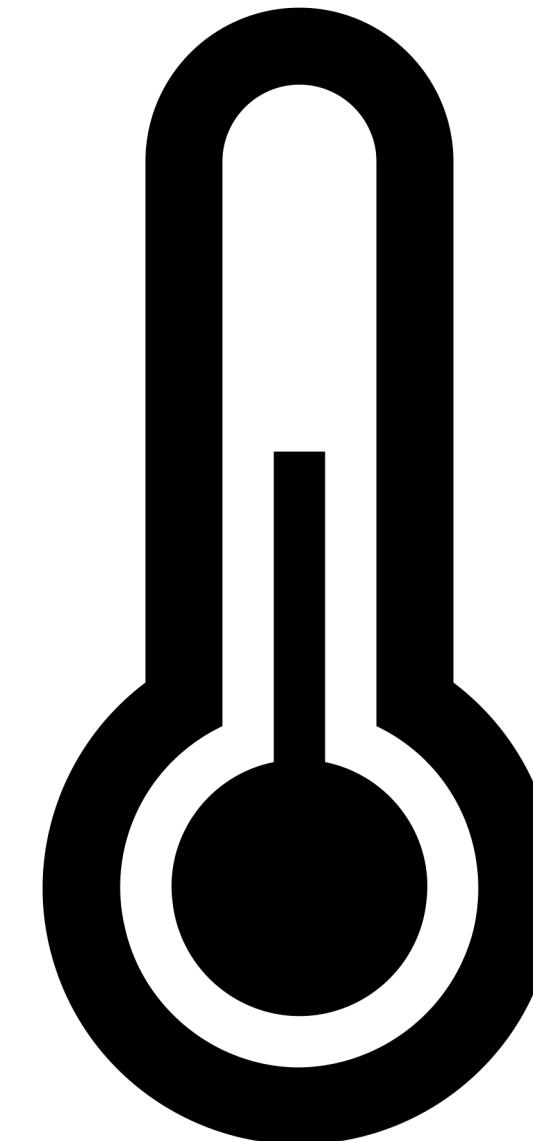
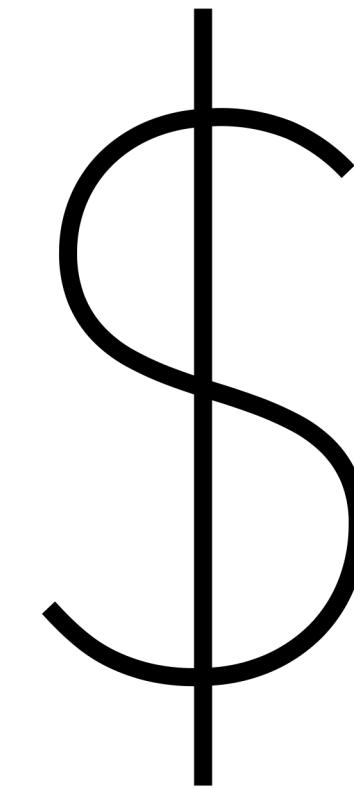
Vertex id	Name	Favorite color	Popularity
1	Sam	Blue	6
2	Sebastian	Green	7
3	Abigail	Purple	9
4	Haley	Pink	2
5	Shane	Orange	4
6	Leah	Purple	7



# Data Types

## DATA TYPE: Attributes

- Specific property that can be measured, observed, or logged
- Ex. Salary, temperature, protein expression levels



# Data Types

## DATA TYPE: Attributes

- Specific property that can be measured, observed, or logged
- Ex. Salary, temperature, protein expression levels

Museum ID	Museum Name	Museum Type	Employer ID Number	Tax Period	Income	Revenue
8400200098	ALASKA AVIATION HERITAGE MUSEUM	HISTORY MUSEUM	920071852	201312	602912	550236
8400200117	ALASKA BOTANICAL GARDEN	ARBORETUM, BOTANICAL GARDEN, OR NATUR	920115504	201312	1379576	1323742
8400200153	ALASKA CHALLENGER CENTER FOR SPACE SCIENCE TECHNOLOGY	SCIENCE & TECHNOLOGY MUSEUM OR PLANET	921761906	201312	740030	729080
8400200143	ALASKA EDUCATORS HISTORICAL SOCIETY	HISTORIC PRESERVATION	920165178	201412	0	0

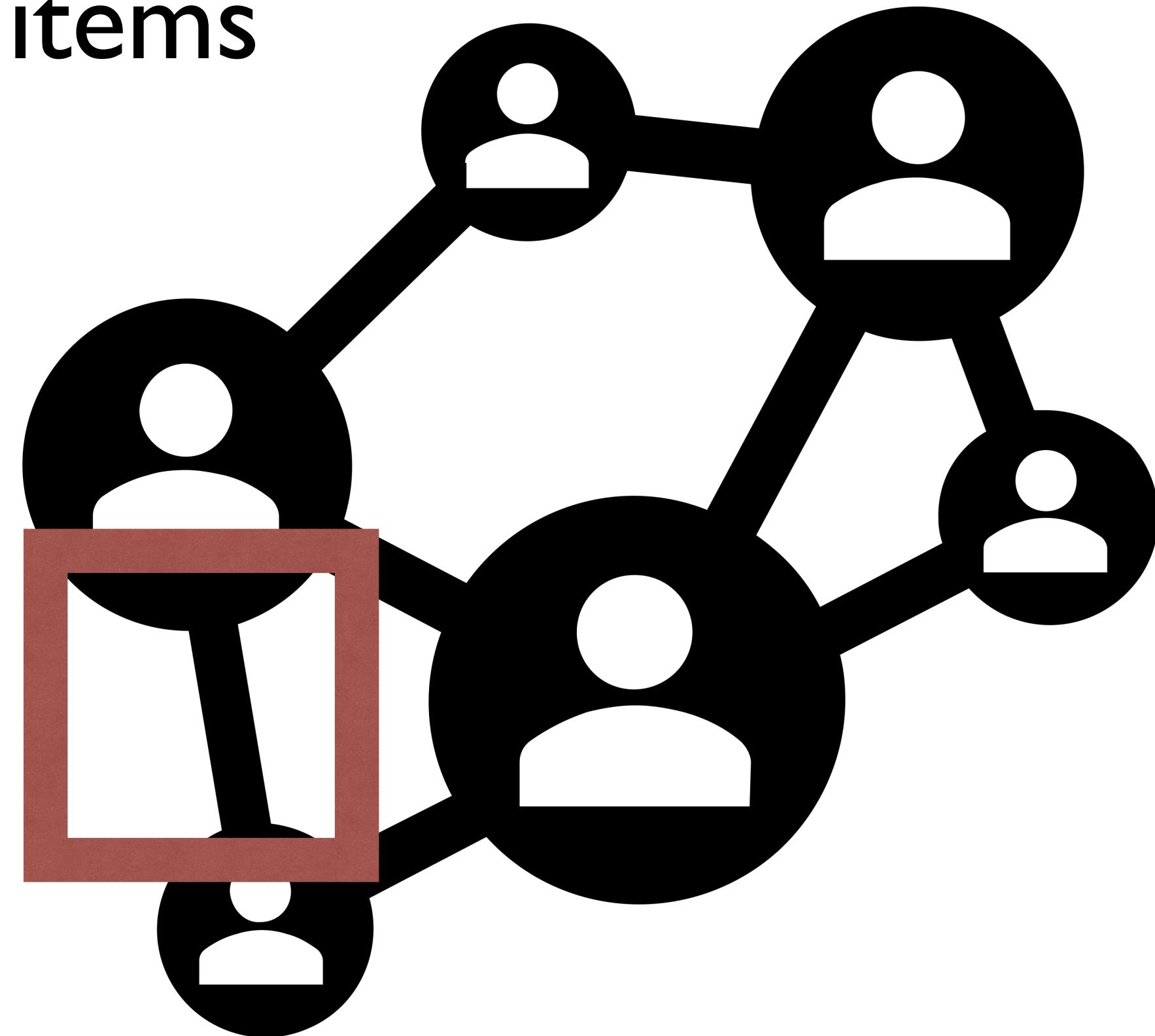
<https://www.kaggle.com/imls/museum-directory>

Vertex id	Name	Favorite color	Popularity
1	Sam	Blue	6
2	Sebastian	Green	7
3	Abigail	Purple	9
4	...	...	2

# Data Types

## DATA TYPE: Link

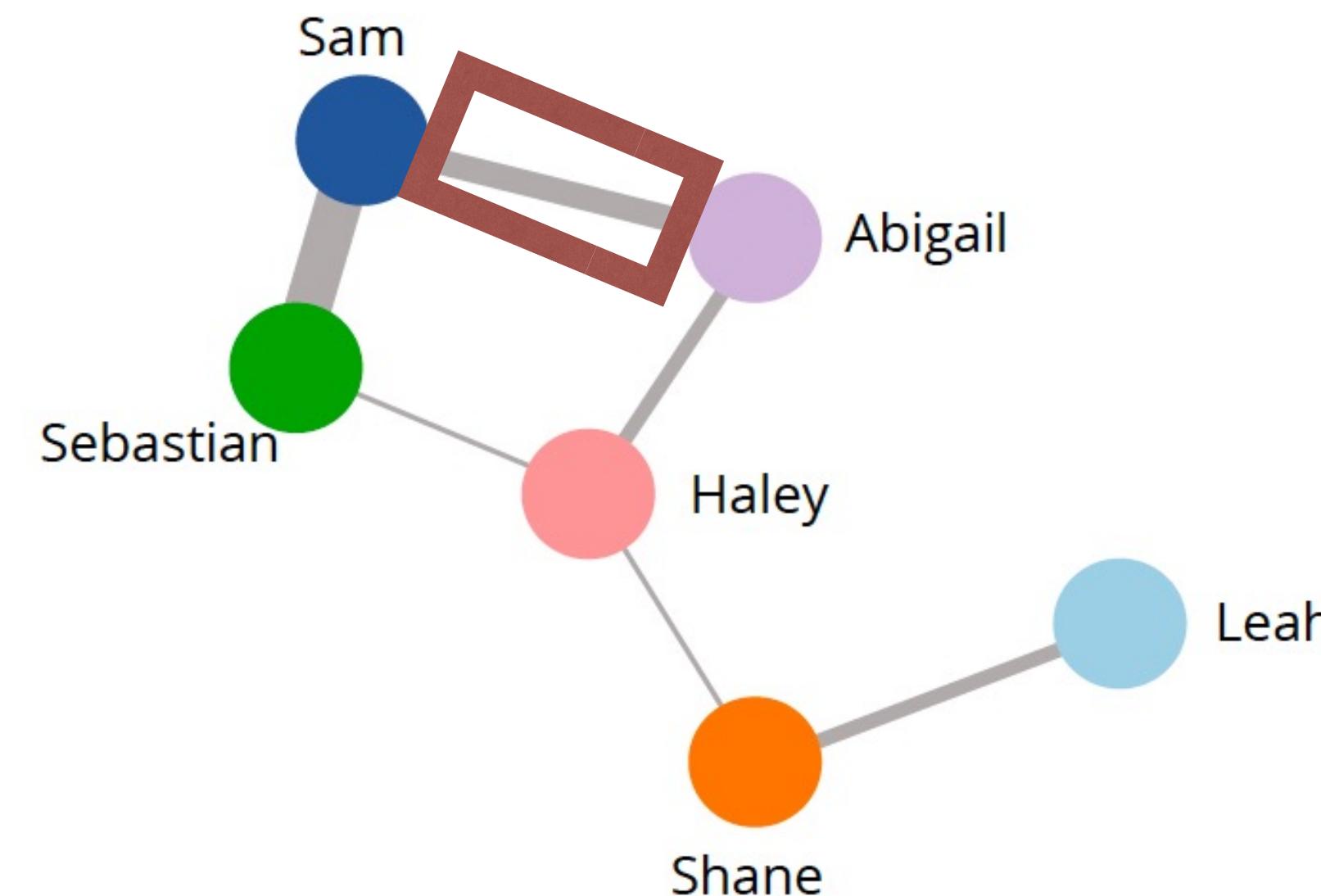
- Relationship between two items
- Ex. Edge in a network



# Data Types

## DATA TYPE: Link

- Relationship between two items
- Ex. Edge in a network

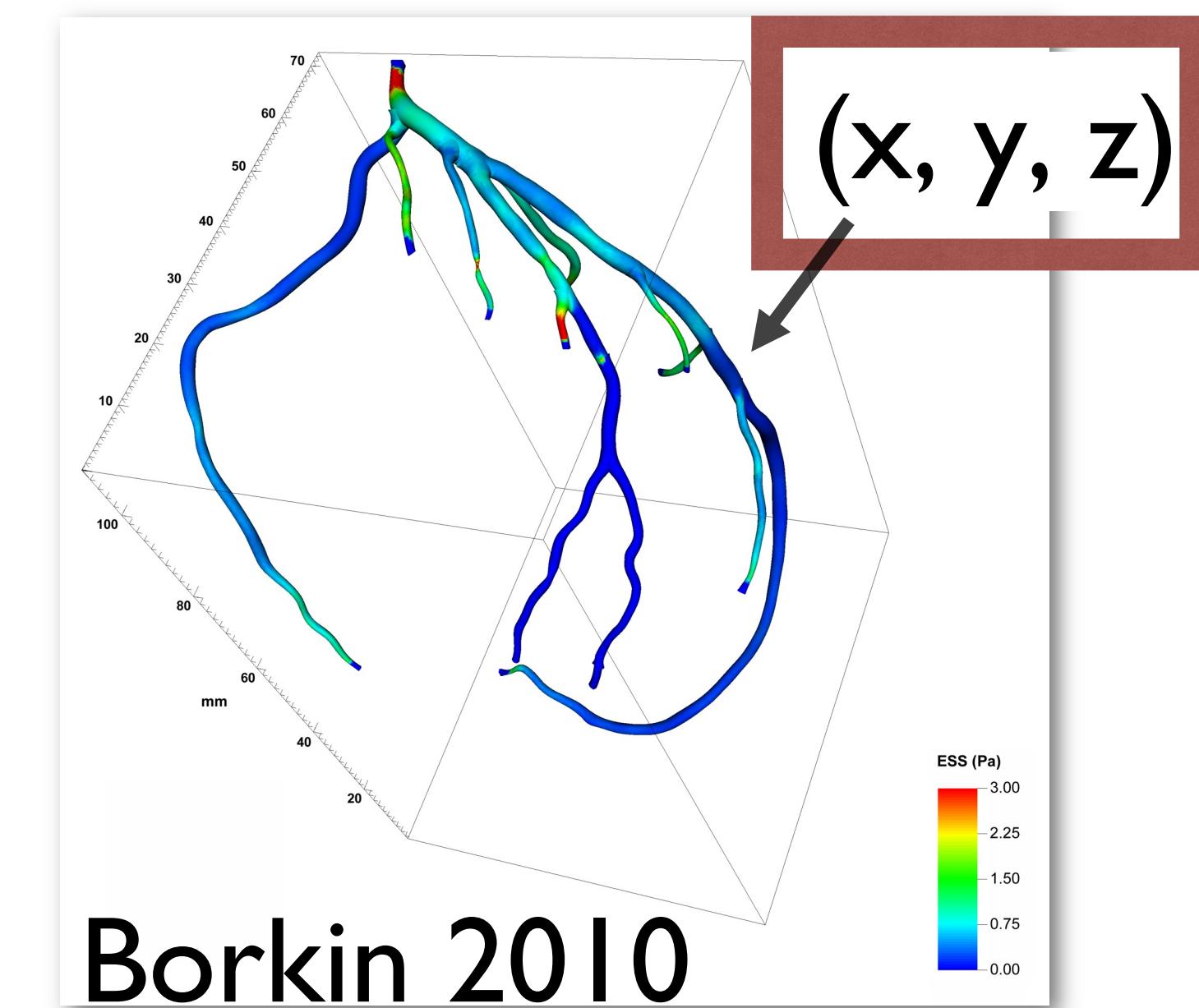
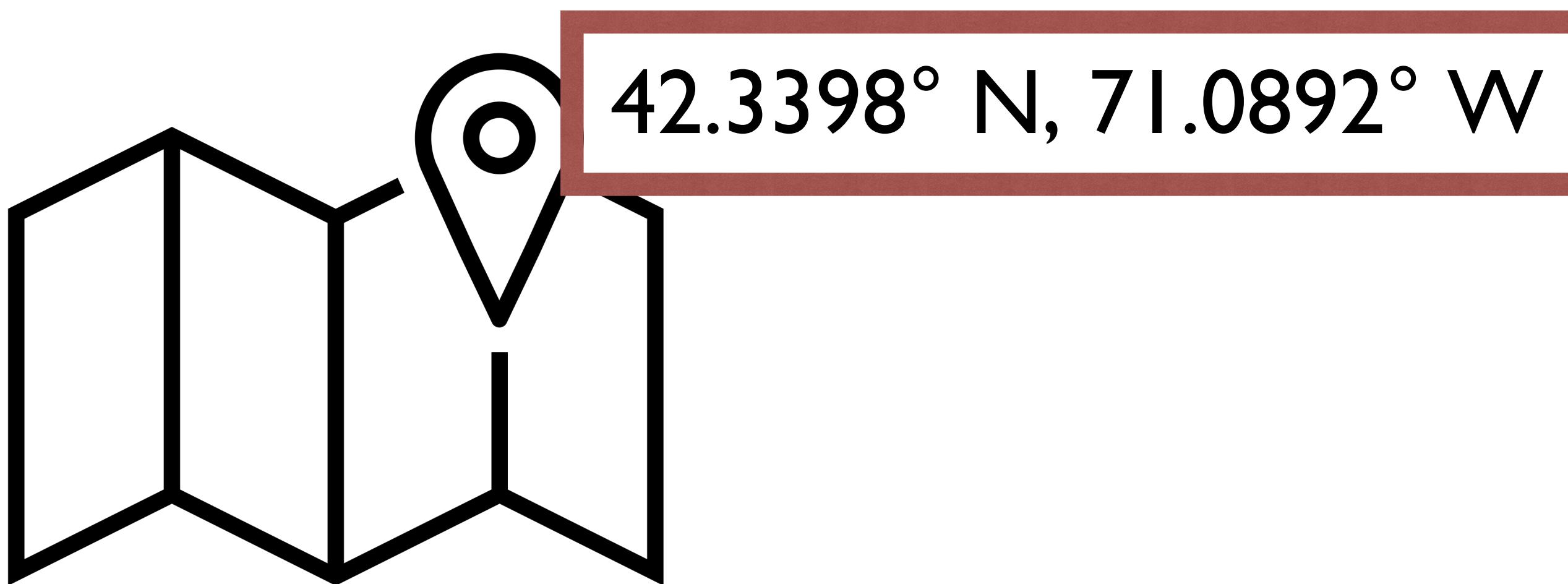


Edge id	Source	Target	Friend value
1	Sam	Sebastian	10
2	Sam	Abigail	6
3	Sebastian	Haley	1
4	Abigail	Haley	2
5	Haley	Shane	1
6	Shane	Leah	2

# Data Types

## DATA TYPE: Position

- Spatial data providing a location in 2D or 3D
- Ex. Latitude & longitude, three numbers describing a region in space measured by a medical scanner



# Data Types

## DATA TYPE: Position

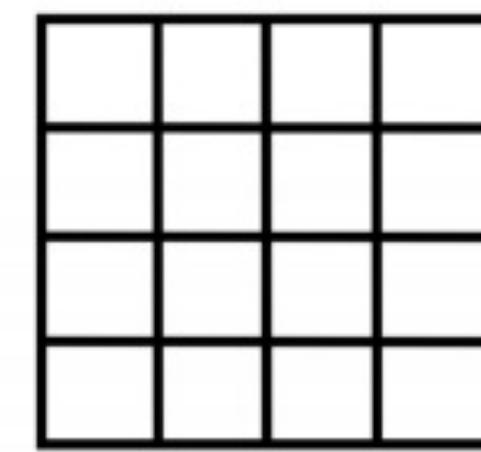
- Spatial data providing a location in 2D or 3D
- Ex. Latitude & longitude, three numbers describing a region in space measured by a medical scanner

Museum ID	Museum Name	Museum Type	Latitude	Longitude
8400200098	ALASKA AVIATION HERITAGE MUSEUM	HISTORY MUSEUM	61.17925	-149.97254
8400200117	ALASKA BOTANICAL GARDEN	ARBORETUM, BOTANICAL GARDEN, OR NATURE CENTER	61.1689	-149.76708
8400200153	ALASKA CHALLENGER CENTER FOR SPACE SCIENCE TECHNOLOGY	SCIENCE & TECHNOLOGY MUSEUM OR PLANETARIUM	60.56149	-151.21598
8400200143	ALASKA EDUCATORS HISTORICAL SOCIETY	HISTORIC PRESERVATION	60.5628	-151.26597
8400200027	ALASKA HERITAGE MUSEUM	HISTORY MUSEUM	61.17925	-149.97254
8400200096	ALASKA HISTORICAL MUSEUM	HISTORIC PRESERVATION	61.21785	-149.85049
8400200078	ALASKA JEWISH MUSEUM	GENERAL MUSEUM	61.18946	-149.86071
8400200084	ALASKA LIGHTHOUSE ASSOCIATION	HISTORIC PRESERVATION	58.28299	-134.40583
8400200107	ALASKA MASONIC LIBRARY AND MUSEUM FOUNDATION	GENERAL MUSEUM	61.21833	-149.89456
8400200073	ALASKA MINING HALL OF FAME FOUNDATION	HISTORY MUSEUM	64.85079	-147.82945
8400200131	ALASKA NATIONAL GUARD HISTORICAL HOLDING-MUSEUM	HISTORIC PRESERVATION	61.25315	-149.68711
8400200123	ALASKA NATIVE HERITAGE CENTER	GENERAL MUSEUM		

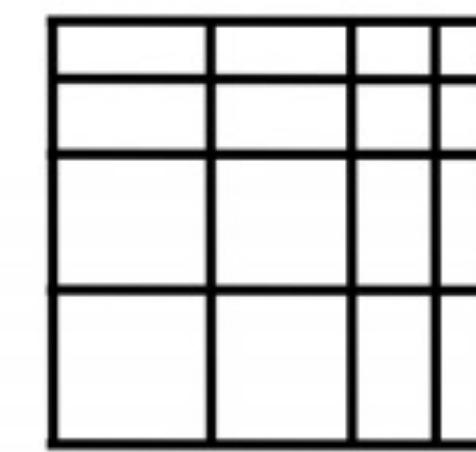
# Data Types

## DATA TYPE: Grid

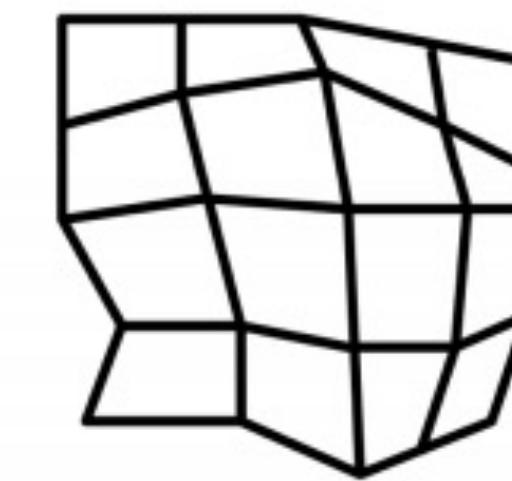
- Specifies the strategy for sampling continuous data in terms of geometric and topological relationships between cells



uniform



rectilinear



structured

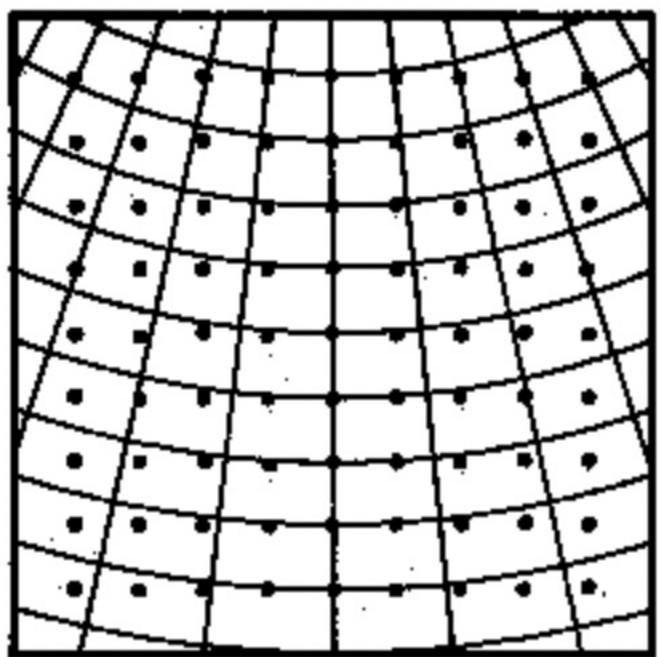
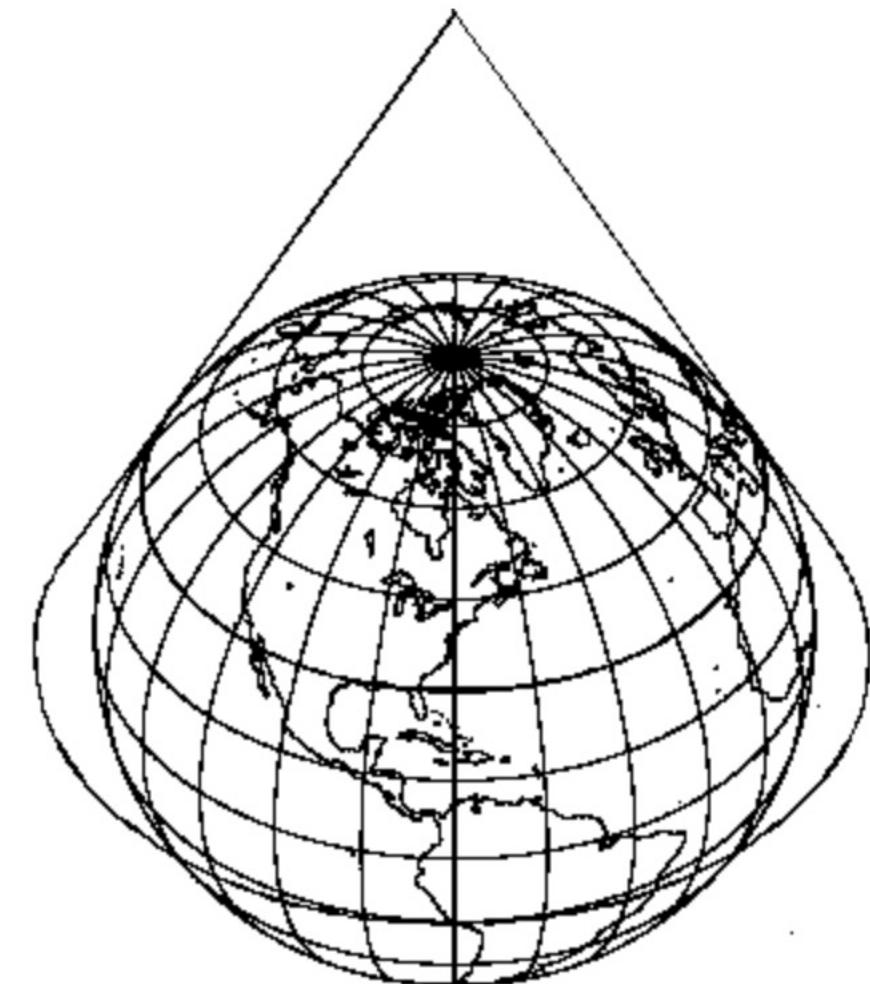
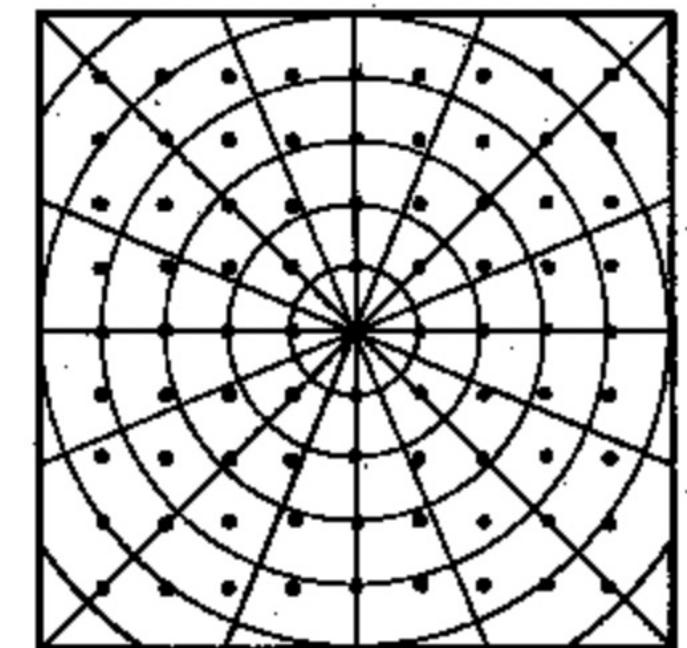
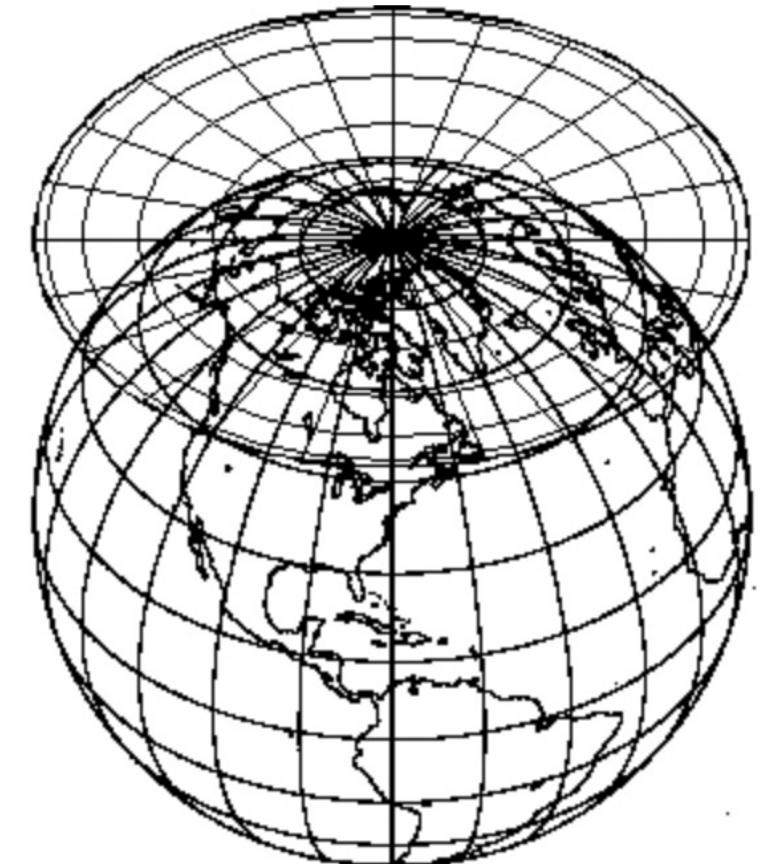
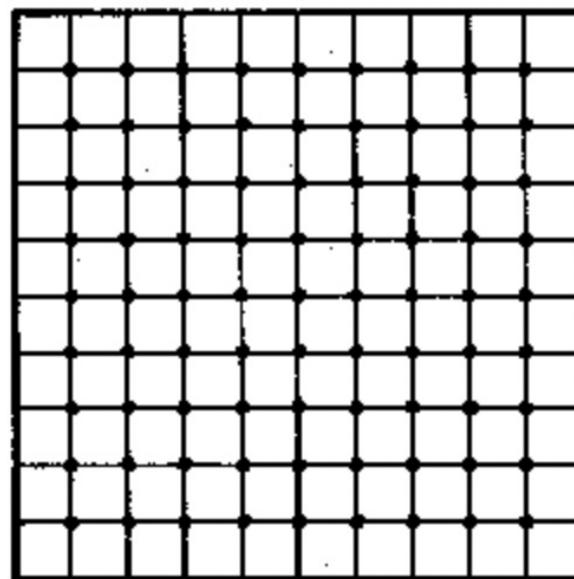
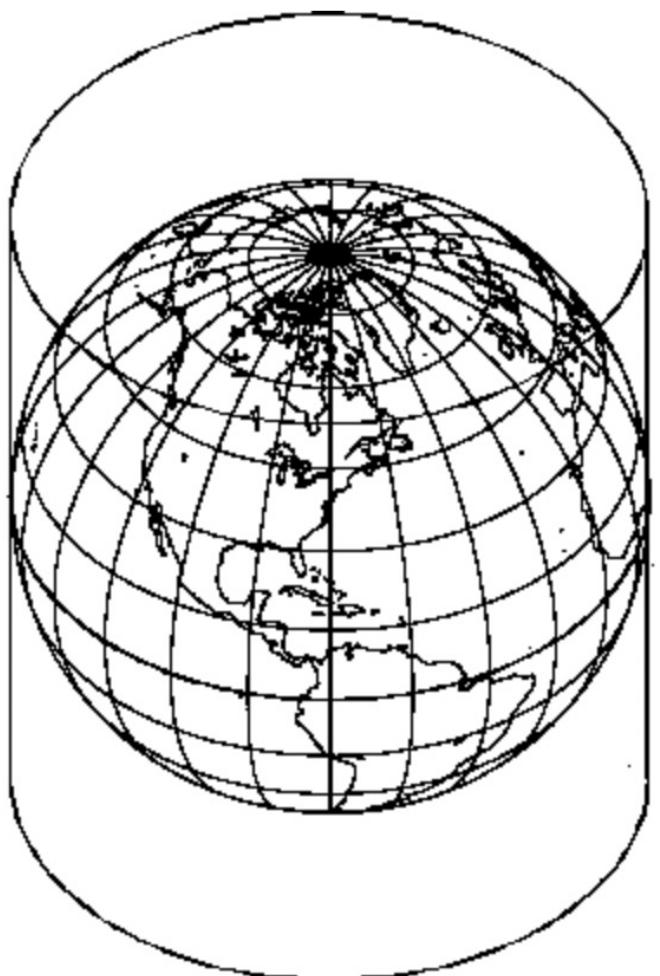


unstructured

# Data Types

## DATA TYPE: Grid

- Specifies the strategy for sampling continuous data in terms of geometric and topological relationships between cells



# Datasets

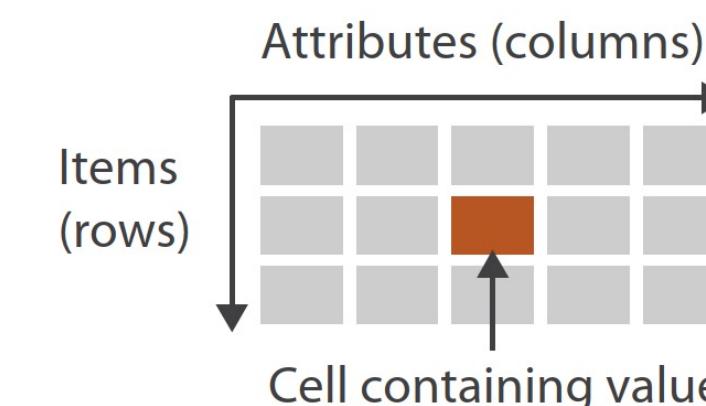
**DATASET** = collection of information that is the target of analysis

# Dataset Types

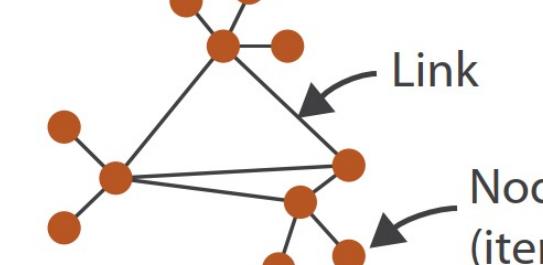
**DATASET** = collection of information that is the target of analysis  
**DATASET TYPE** is determined by the types of data making up the set

## → Dataset Types

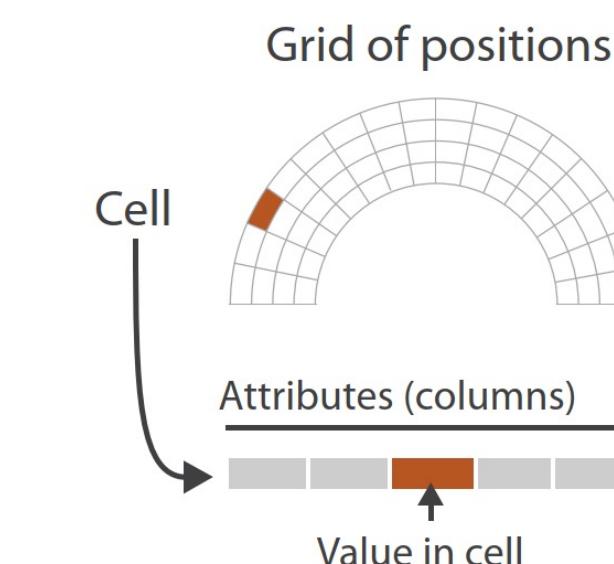
### → Tables



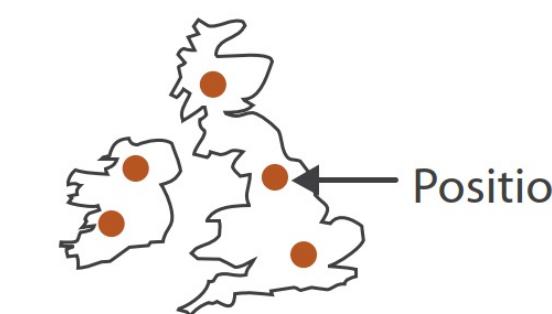
### → Networks



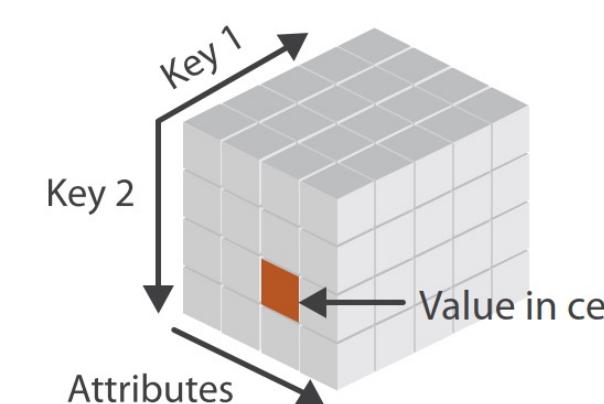
### → Fields (Continuous)



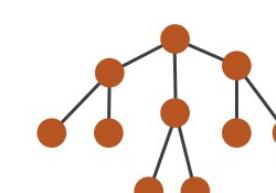
### → Geometry (Spatial)



### → Multidimensional Table



### → Trees



# Data & Dataset Types

**DATASET** = collection of information that is the target of analysis

**DATASET TYPE** is determined by the types of data making up the set

## → Data and Dataset Types

Tables

Items

Attributes

Networks &  
Trees

Items (nodes)

Links

Attributes

Fields

Grids

Positions

Attributes

Geometry

Items

Positions

Clusters,  
Sets, Lists

Items

# Data & Dataset Types

**DATASET** = collection of information that is the target of analysis

**DATASET TYPE** is determined by the types of data making up the set

## → Data and Dataset Types

Tables

Items

Attributes

Networks &  
Trees

Items (nodes)

Links

Attributes

Fields

Grids

Positions

Attributes

Geometry

Items

Positions

Clusters,  
Sets, Lists

Items

# Attributes

**ATTRIBUTE** = Specific property that can be measured, observed, or logged

- There are different types of attributes

## ➔ Attribute Types

→ Categorical

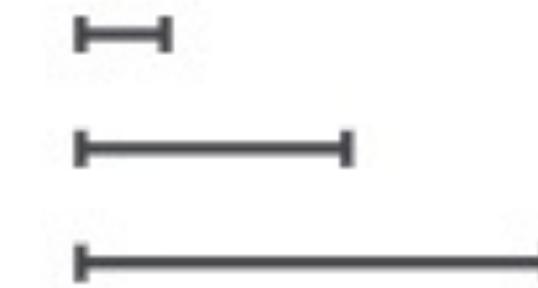


→ Ordered

→ *Ordinal*



→ *Quantitative* (i.e. continuous)



Ex.

**Fruit:** apple, pear, grape

**Colleges:** CAMD, Khoury, COE

Ex.

**Sizes:** s, m, l

**Months:** Jan., Feb.

Ex.

**Lengths:** 1', 2.5', 5'

**Population:** 1m., 11.75 m., 2.2 m.

# Attributes

**ATTRIBUTE** = Specific property that can be measured, observed, or logged

- There are different types of attributes

## → Attribute Types

→ Categorical



Ex.

**Fruit:** apple, pear, grape

**Colleges:** CAMD, Khoury, COE

→ Ordered

→ *Ordinal*

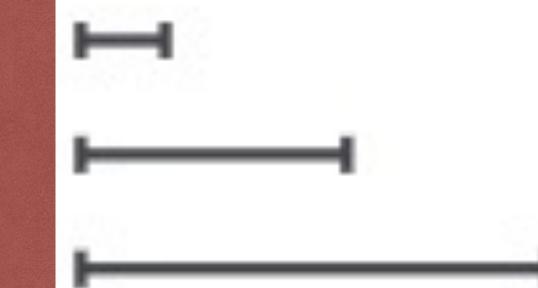


Ex.

**Sizes:** s, m, l

**Months:** Jan., Feb.

→ *Quantitative* (i.e. continuous)



Ex.

**Lengths:** 1', 2.5', 5'

**Population:** 1m., 11.75 m., 2.2 m.

# Attributes

**ATTRIBUTE** = Specific property that can be measured, observed, or logged

- **LEVELS** are the different groups in **CATEGORICAL** and **ORDINAL** attributes

Unique Squirrel ID	Hectare	Shift	Date	Hectare Squirrel Number	Age	Primary Fur Color	Highlight Fur Color	Location	Specific Location	Running	Chasing	Climbing	Eating	Foraging
37F-PM-1014-03	37F	PM	10142018	3						FALSE	FALSE	FALSE	FALSE	FALSE
21B-AM-1019-04	21B	AM	10192018	4						FALSE	FALSE	FALSE	FALSE	FALSE
11B-PM-1014-08	11B	PM	10142018	8		Gray		Above Ground		FALSE	TRUE	FALSE	FALSE	FALSE
32E-PM-1017-14	32E	PM	10172018	14	Adult	Gray				FALSE	FALSE	FALSE	TRUE	TRUE
13E-AM-1017-05	13E	AM	10172018	5	Adult	Gray	Cinnamon	Above Ground	on tree stump	FALSE	FALSE	FALSE	FALSE	TRUE
11H-AM-1010-03	11H	AM	10102018	3	Adult	Cinnamon	White			FALSE	FALSE	FALSE	FALSE	TRUE
36H-AM-1010-02	36H	AM	10102018	2	Adult	Gray		Ground Plane		FALSE	FALSE	FALSE	FALSE	TRUE
33F-AM-1008-02	33F	AM	10082018	2	Adult	Gray		Ground Plane		FALSE	FALSE	FALSE	FALSE	TRUE
12I-AM-1013-01	12I	AM	10132018	1	Adult	Cinnamon	White	Ground Plane	on tree roots	FALSE	FALSE	FALSE	FALSE	FALSE
39G-PM-1013-06	39G	PM	10132018	6	Adult	Gray	Cinnamon	Ground Plane		TRUE	FALSE	FALSE	FALSE	FALSE
29I-PM-1007-01	29I	PM	10072018	1		Gray	Cinnamon	Ground Plane		FALSE	FALSE	FALSE	TRUE	TRUE
07E-AM-1006-02	07E	AM	10062018	2	Adult	Gray		Ground Plane	under a tree	FALSE	FALSE	FALSE	TRUE	FALSE
17C-PM-1013-05	17C	PM	10132018	5	Adult	Gray		Ground Plane		FALSE	FALSE	FALSE	TRUE	FALSE

AM,  
PM

Adult,  
Juvenile

Gray,  
Cinnamon

White

Cinnamon,

Above Ground,

Ground Plane

# Attributes

**ATTRIBUTE** = Specific property that can be measured, observed, or logged

- There are different types of attributes

## ➔ Attribute Types

→ Categorical



Ex.

**Fruit:** apple, pear, grape

**Colleges:** CAMD, Khoury, COE

→ Ordered

→ *Ordinal*

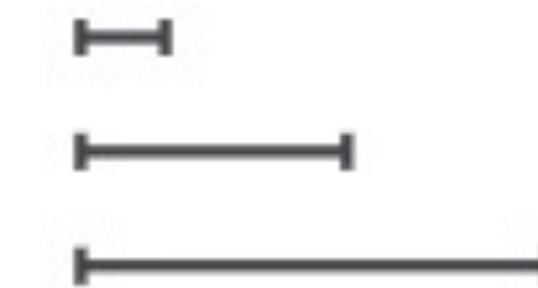


Ex.

**Sizes:** s, m, l

**Months:** Jan., Feb.

→ *Quantitative* (i.e. continuous)



Ex.

**Lengths:** 1', 2.5', 5'

**Population:** 1m., 11.75 m., 2.2 m.

# Attributes

**ATTRIBUTE** = Specific property that can be measured, observed, or logged

- There are **TYPES OF ORDERINGS** for ordered attributes



## Ordering Direction

→ Sequential



Ex.

**Height:** > 0  
**Counts**

→ Diverging



Ex.

**Elevation:** below  
sea level, at sea level,  
above sea level  
**Temperature**

→ Cyclic



Ex.

**Time:** hour of day  
**Moon Phases**

# Data & Dataset Types

**DATASET** = collection of information that is the target of analysis

**DATASET TYPE** is determined by the types of data making up the set

## → Data and Dataset Types

Tables

Items

Attributes

Networks &  
Trees

Items (nodes)

Links

Attributes

Fields

Grids

Positions

Attributes

Geometry

Items

Positions

Clusters,  
Sets, Lists

Items

Let's take a break! Stretch, go  
for a walk, be social ☺  
Be back here in 10 mins.

# INTRO TO HTML + CSS

# HTML

## HTML

- Hypertext Markup Language
- A markup language that **defines the structure of webpages**

```
<!DOCTYPE html>
<html>
<head>
<title>Page Title</title>
</head>
<body>

<h1>This is a Heading</h1>
<p>This is a paragraph.</p>

</body>
</html>
```

**This is a Heading**

This is a paragraph.

# HTML

## HTML

→ Defines webpage structure as a series of **elements** identified by **tags** and filled with **content**

```
<!DOCTYPE html>
<html>
<head>
<title>Page Title</title>
</head>
<body>

<h1>This is a Heading</h1>
<p>This is a paragraph.</p>

</body>
</html>
```

## This is a Heading

This is a paragraph.

```
<html>

<head>
    <title>Page title</title>
</head>

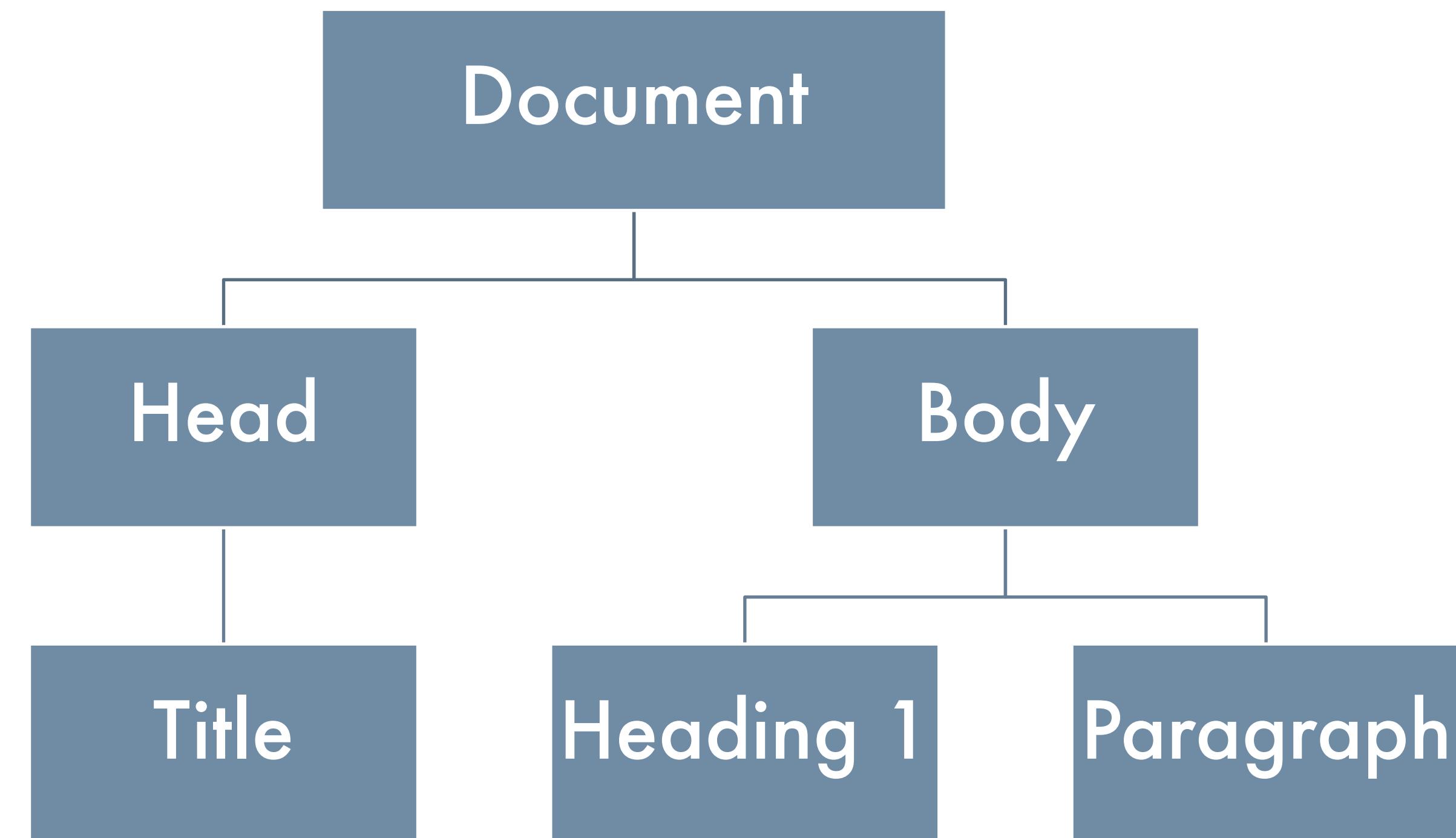
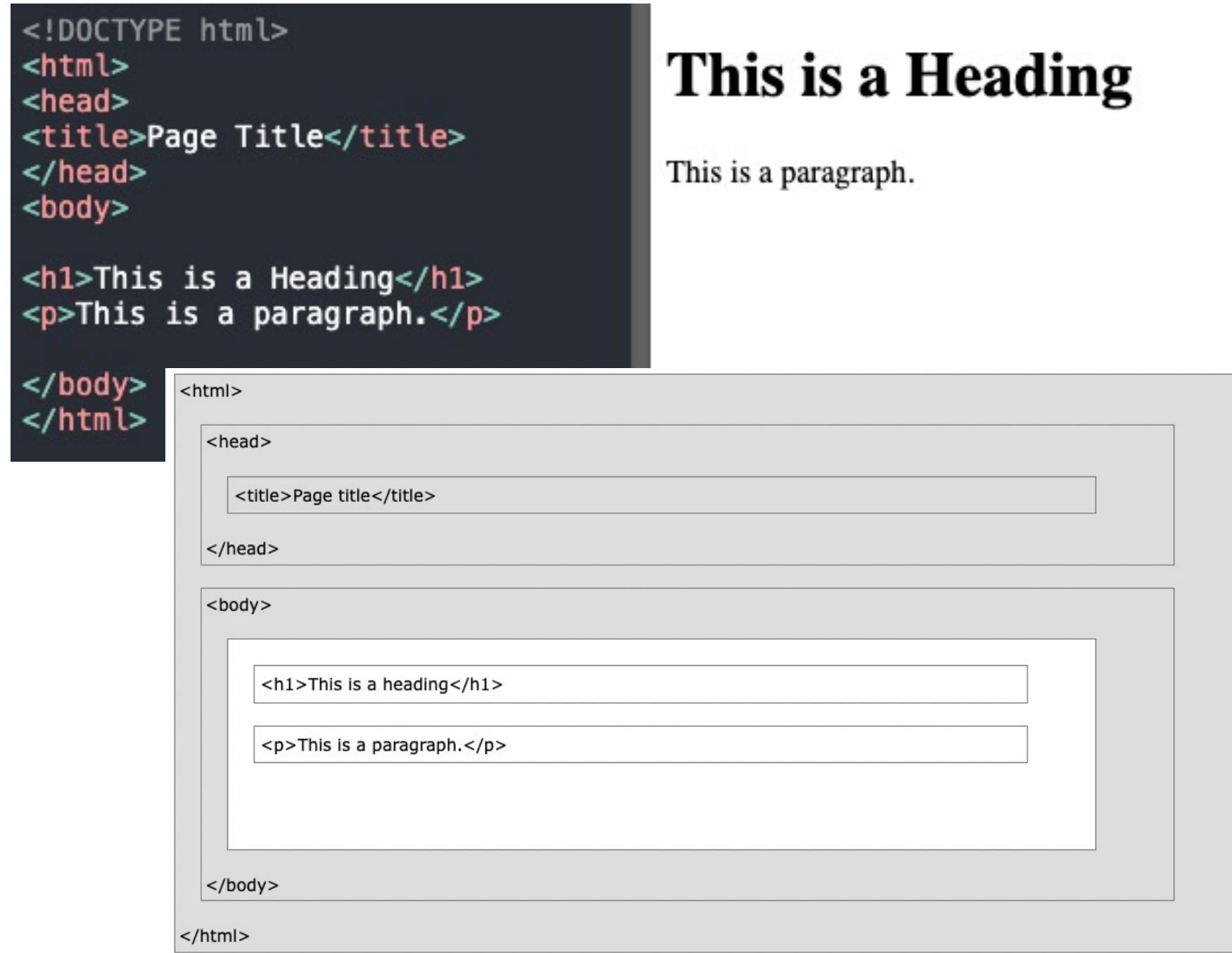
<body>
    <h1>This is a heading</h1>
    <p>This is a paragraph.</p>
</body>

</html>
```

# HTML

## HTML

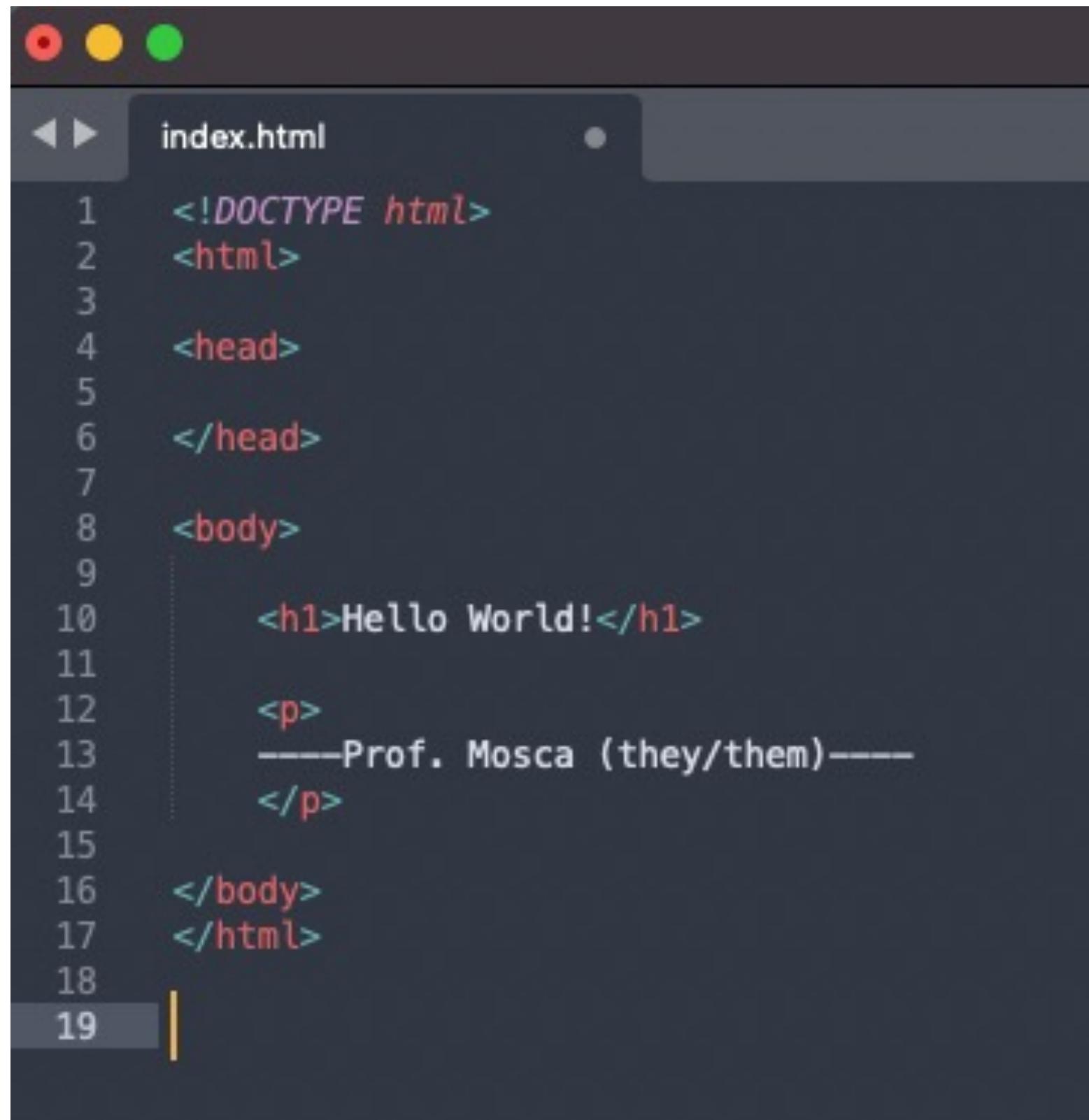
→ We conceptualize the structure of webpages using the **Document Object Model (DOM)**



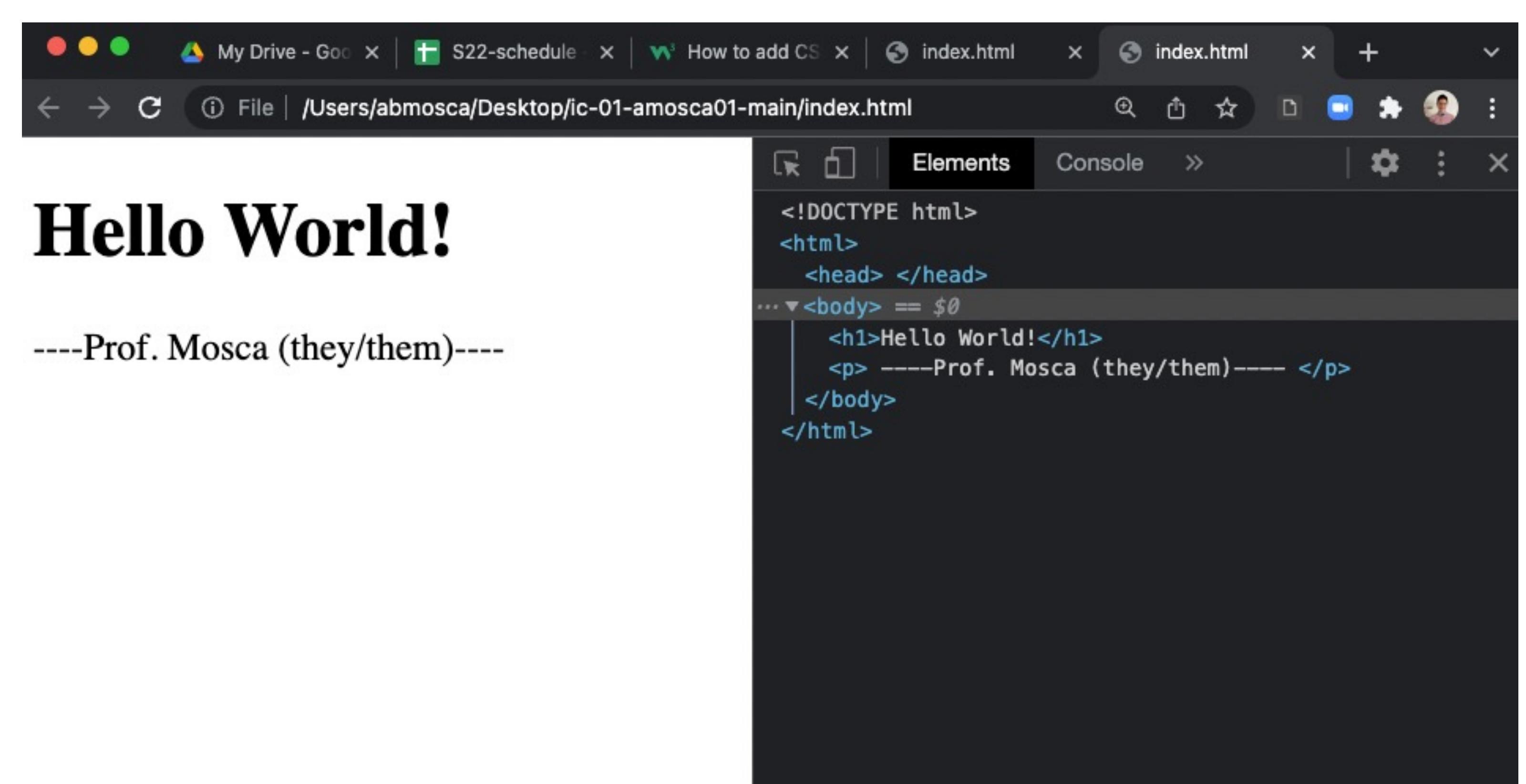
# HTML

## HTML

→ You can view the HTML for any webpage using the developer console



```
index.html
1 <!DOCTYPE html>
2 <html>
3 <head>
4 </head>
5 <body>
6   <h1>Hello World!</h1>
7
8   <p>
9     ----Prof. Mosca (they/them)-----
10  </p>
11
12 </body>
13 </html>
```



The browser window shows the file path: /Users/abmosca/Desktop/ic-01-amosca01-main/index.html. The developer tools are open, displaying the page content "Hello World!" and "----Prof. Mosca (they/them)----". The developer tools interface includes tabs for Elements, Console, and a gear icon.

```
<!DOCTYPE html>
<html>
  <head> </head>
... <body> == $0
  |   <h1>Hello World!</h1>
  |   <p> ----Prof. Mosca (they/them)---- </p>
  |   </body>
</html>
```

# HTML

## HTML

→ Distinguish between elements with **id's** and group them with **classes**

The image shows a code editor on the left and a browser window on the right. The code editor displays the following HTML code:

```
index.html
1 <!DOCTYPE html>
2 <html>
3 <head>
4 </head>
5 <body>
6   <h1>Hello World!</h1>
7   <p id="name" class="important">
8     ---Prof. Mosca (they/them)---
9   </p>
10  <div id="containter1">
11    <p class="important">This is my paragraph inside of a div</p>
12  </div>
13 </body>
14 </html>
```

The browser window shows the rendered HTML with the following content:

Hello World!

----Prof. Mosca (they/them)---

This is my paragraph inside of a div

The browser's developer tools Elements tab shows the DOM structure:

```
<!DOCTYPE html>
<html>
  <head> </head>
  ... <body> == $0
    <h1>Hello World!</h1>
    <p id="name" class="important"> ----Prof. Mosca
      (they/them)--- </p>
    ><div id="containter1">...</div>
  </body>
</html>
```

# HTML + CSS

## CSS

- Cascading Style Sheet
- Used to **style** HTML elements

The image shows a file structure and a browser output. On the left, two code editors are open: one for `index.html` and one for `style.css`. The `index.html` file contains basic HTML structure and a paragraph with a class. The `style.css` file contains a single rule for the `body` element. To the right, a browser window displays the rendered content. The page title is "Hello World!". Below it is a paragraph with the text "----Prof. Mosca (they/them)----". Inside a `div` with id="containter1", there is another paragraph with the text "This is my paragraph inside of a div". The browser's developer tools are open, showing the DOM tree and the styles tab where the rule from `style.css` is listed under the `body` element.

```
index.html
1 <!DOCTYPE html>
2 <html>
3 <head>
4   <link rel="stylesheet" href="style.css">
5 </head>
6 <body>
7   <h1>Hello World!</h1>
8   <p id="name" class="important">
9     Prof. Mosca (they/them)
10    </p>
11   <div id="containter1">
12     <p class="important">This is my paragraph inside of a div</p>
13   </div>
14 </body>
15 </html>
```

```
style.css
1 body {
2   font-family: Arial;
3 }
```

Hello World!

----Prof. Mosca (they/them)----

This is my paragraph inside of a div

```
<!DOCTYPE html>
<html>
  <head>...</head>
... <body> == $0
  <h1>Hello World!</h1>
  <p id="name" class="important"> ----Prof. Mosca
  (they/them)---- </p>
  <div id="containter1">...</div>
</body>
</html>
```

html body

Styles Computed Layout Event Listeners »

Filter :hov .cls + □

element.style {  
}  
body {  
 font-family: Arial;  
}

# HTML + CSS

CSS

→ Apply style to elements, id's, or classes

The image shows a development environment with two code editors and a browser window.

**Code Editors:**

- index.html:** Contains the following HTML code:

```
<!DOCTYPE html>
<html>
  <head>
    <link rel="stylesheet" href="style.css">
  </head>
  <body>
    <h1>Hello World!</h1>
    <p id="name" class="important">----Prof. Mosca (they/them)----</p>
    <div id="containter1">
      <p class="important">This is my paragraph inside of a div</p>
    </div>
  </body>
</html>
```
- style.css:** Contains the following CSS code:

```
body {
  font-family: Arial;
}

#name {
  font-size: 40px;
  color: green;
}

.important {
  font-style: italic;
}
```

**Browser Preview:**

The browser window displays the rendered HTML with the following visual output:

**Hello World!**

**----Prof. Mosca  
(they/them)----**

*This is my paragraph inside of a div*

**Elements Panel:**

- Shows the DOM tree:

```
<!DOCTYPE html>
<html>
  <head>...</head>
  <body>
    <h1>Hello World!</h1>
    <p id="name" class="important">----Prof. Mosca (they/them)----</p>
    <div id="containter1">...</div>
  </body>
</html>
```
- Shows the selected element: **p#name.important**
- Shows the CSS styles applied to the selected element:

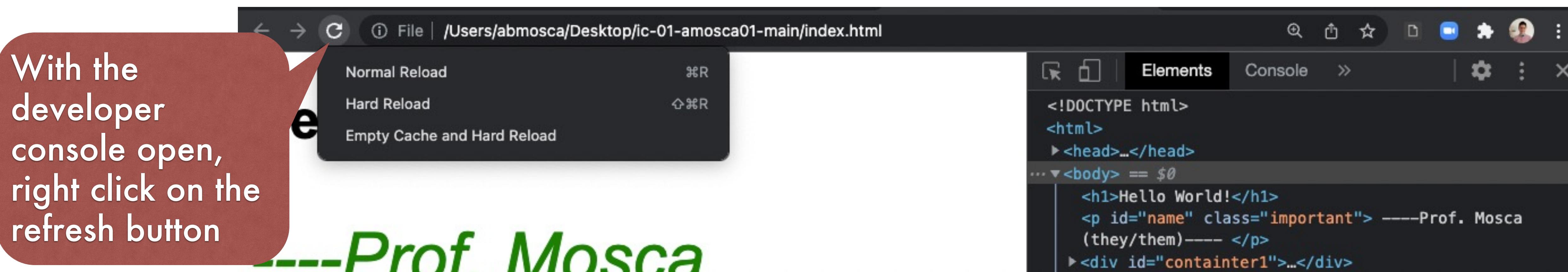
```
#name {
  font-size: 40px;
  color: green;
}

.important {
  font-style: italic;
}
```

# HTML + CSS

## General Tips

- W3Schools is a great resource. Note all of the topics linked on the left-hand side of the page.
- If you update your HTML/CSS and see no changes in the browser, try clearing your cache.



# IC-02: HTML, CSS

# Summary

## **Today we:**

- Reviewed building blocks of visualization
- Reviewed data abstraction
- Intro'd HTML and CSS

**ic-02 is DUE today.**