DIGITAL INFORMATION

In a computer, information is represented as bits (0's and 1's). We typically quantify data as bytes (8 bits): There are numerous ways to translate human readable data to binary, such as ASCII:

- Kilobyte (kB) = 1,000 bytes
- Megabyte (MB) = 1,000,000 bytes
- Gigabyte (GB) = 1,000,000,000 bytes

DIGITAL INFORMATION

There are numerous ways to translate human readable data to binary, such as ASCII, where each character is represented as one byte. There are
$2^8 = 256$ unique combinations of 0's and 1's in a byte.

- "A" : 01000001
- "CAT": 01000011 01000001 01010100
- "31": 00110011 00110001

DIGITAL INFORMATION

Modern computers use 64-bit "architecture". That is, the central processing unit (CPU) can handle 64 bits (8 bytes) of information at a time.

- "Word" Length is 64 bits
- 2^{64} (>18 quintillion) possible unique values
- CPUs can be stacked in parallel to handle more information at one time

REPRESENTING SPATIAL PHENOMENA

Within the context of a GIS, every piece of information describing a phenomenon is referred to as an **Attribute**. Broadly speaking each attribute can address one of three questions:

- **Where?**
- **What?**
- **When?**

TYPES OF ATTRIBUTES

- **Spatial Data**: Attributes that describe where.
- **Non-Spatial Data**: Attributes that describe what or when.

ATTRIBUTES

All data, spatial and non-spatial, can be either **qualitative** or **quantitative**. The types of analysis we can do with qualitative data are more limited, but that does not make quantitative data “better”.

QUALITATIVE DATA

Categorical: strictly descriptive and lack any meaningful numeric value.

- Textual or coded numerals
- Measured on either a **Nominal** or **Ordinal** scale.

NOMINAL SCALE

- Names or categories with no ranking or direction
- Categories are not more/less, better/worse, just different.

NOMINAL SCALE

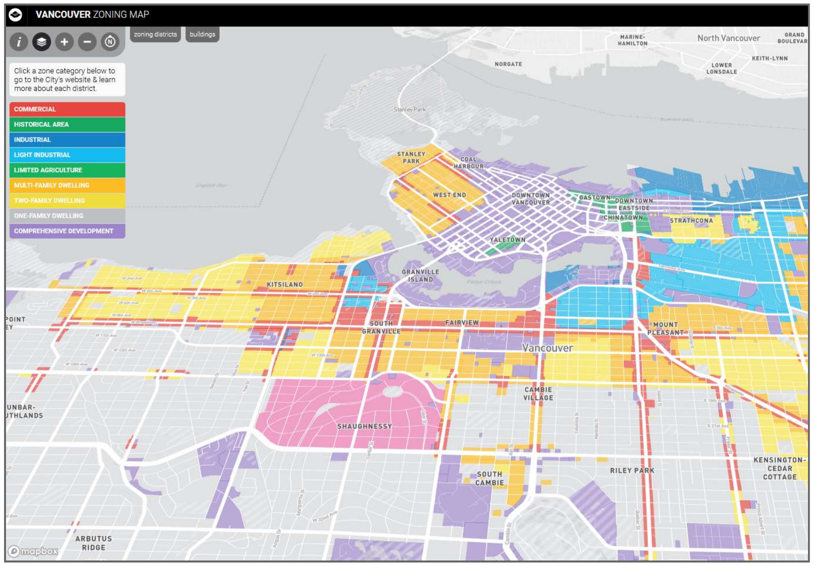
- Flower Species

Examples:

NOMINAL SCALE

Examples:

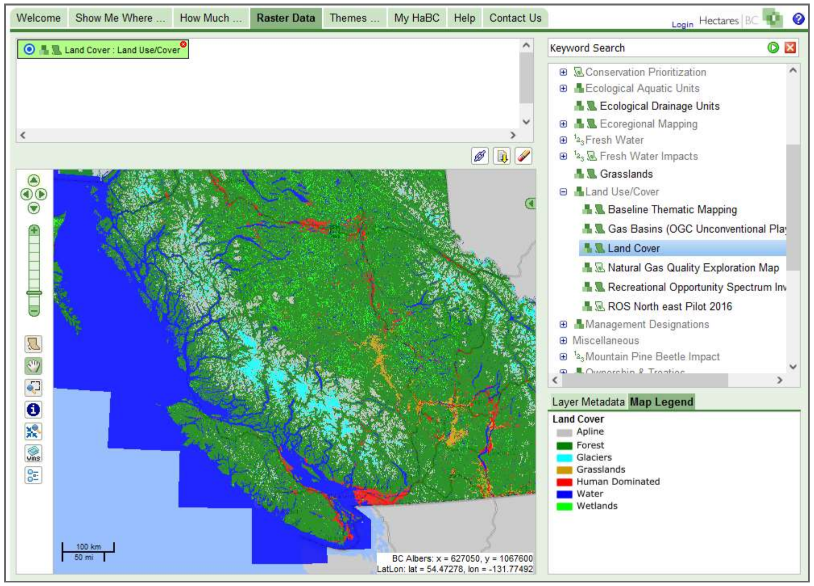
- Flower Species
- Zoning Categories



NOMINAL SCALE

Examples:

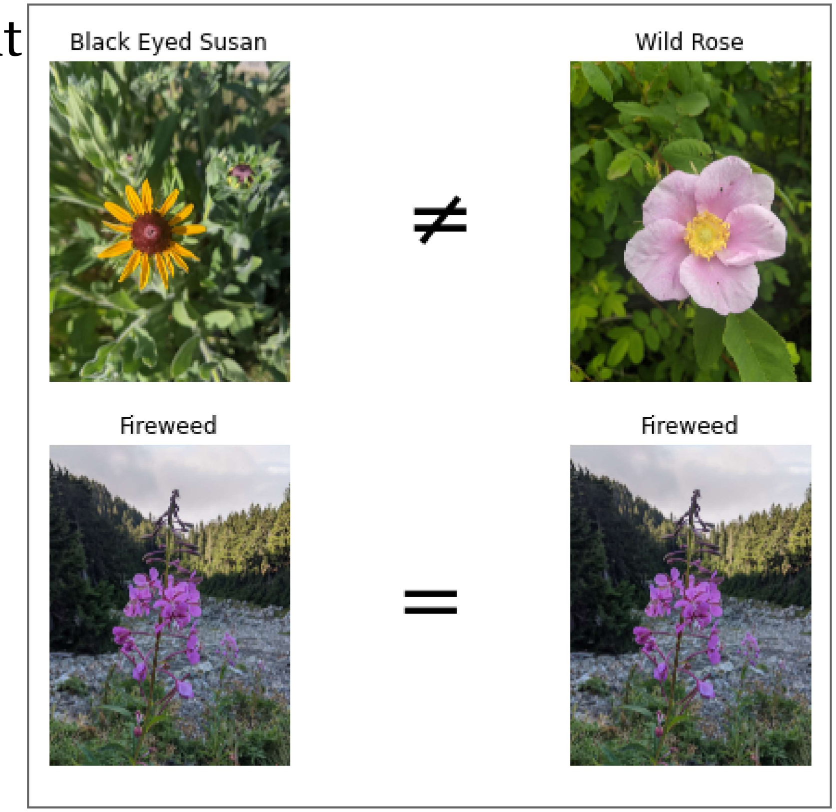
- Flower Species
- Zoning Categories
- Landcover Classification



NOMINAL OPERATIONS

Se can check equivalency and count frequencies.

Nothing else



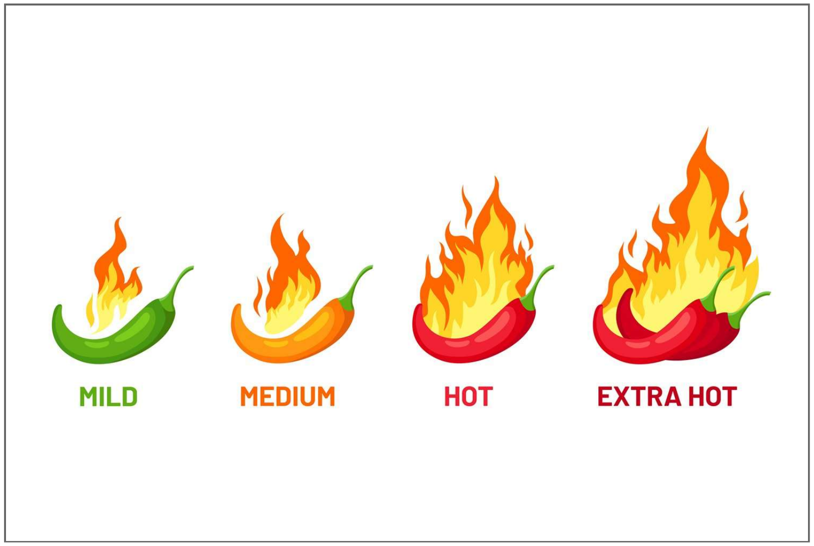
NOMINAL SCALE

- Names or categories
- Some ranking or directionality

ORDINAL SCALE

- Spice levels

Examples:



ORDINAL SCALE

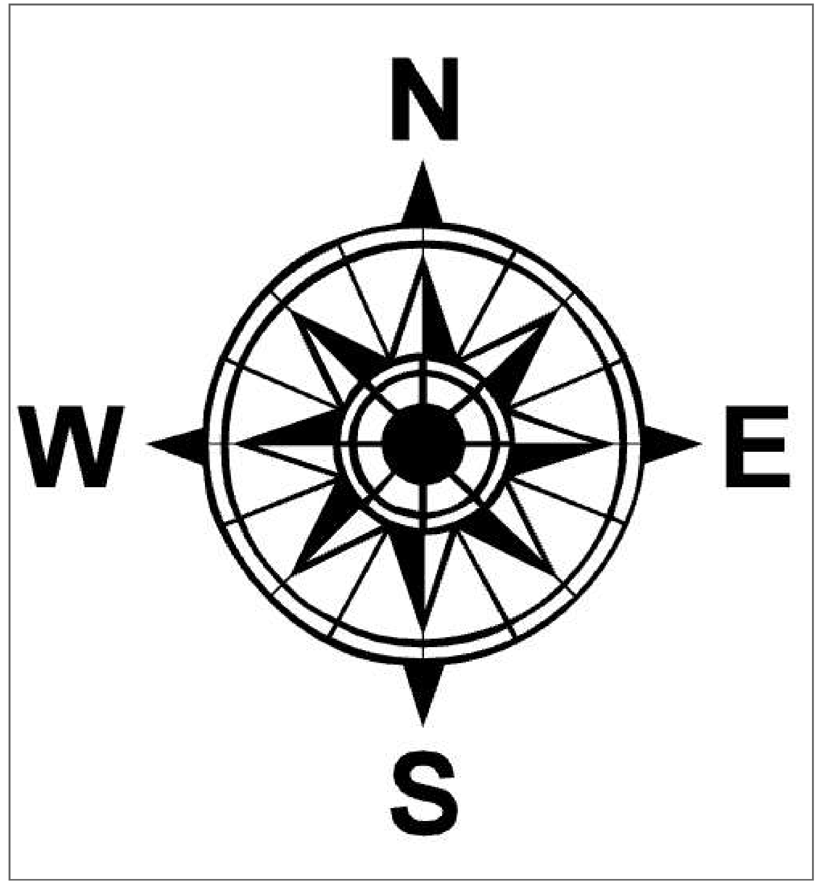
Examples:

- Spice levels
- Relative heights

ORDINAL SCALE

Examples:

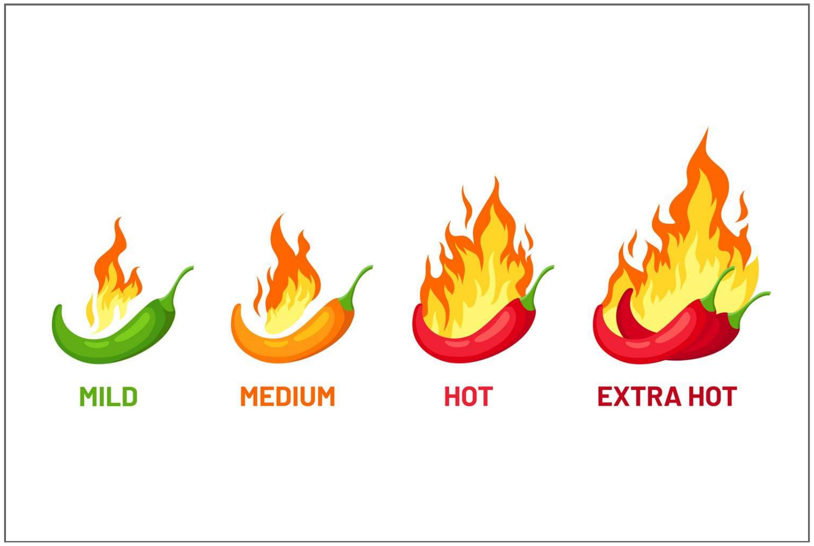
- Spice levels
- Relative heights
- Compass Direction



ORDINAL OPERATIONS

Same operations as nominal data + more.

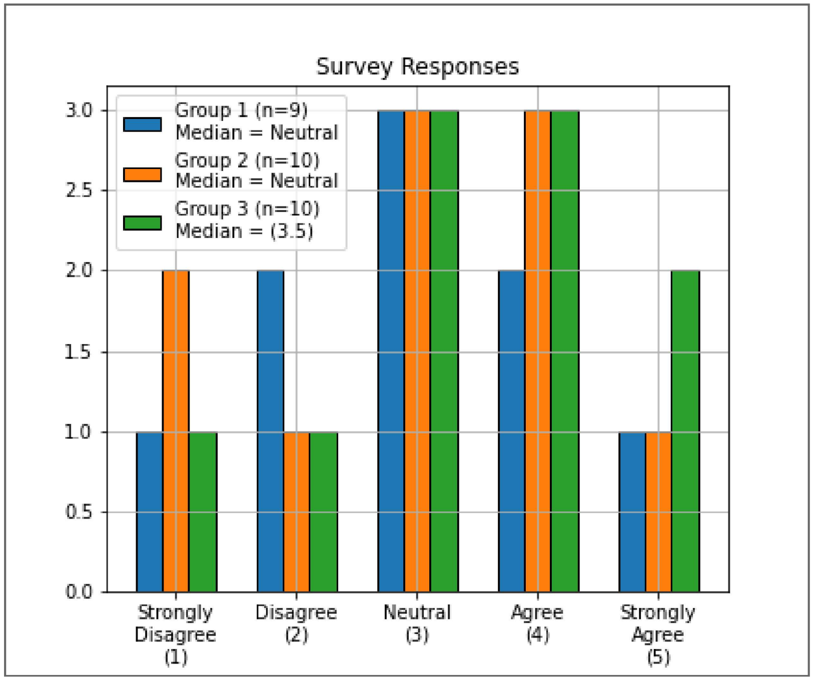
- Check order/rank: greater/less than



ORDINAL OPERATIONS

Sometimes we can calculate the median.

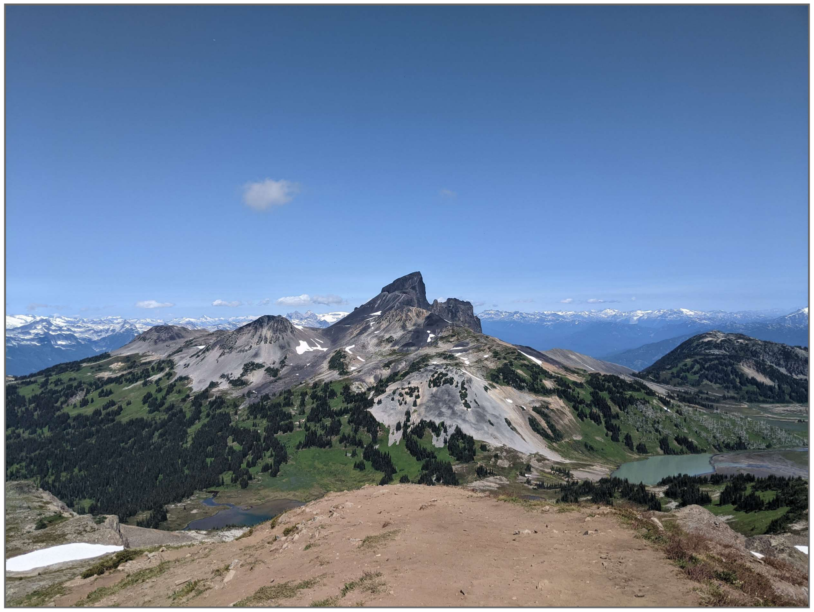
- Odd sets the median is the middle.
- Even sets, average of the middle two.
- One solution, *arbitrarily* assign a numeric score.



GRADED MEMBERSHIP

Exceptions that blur the lines.

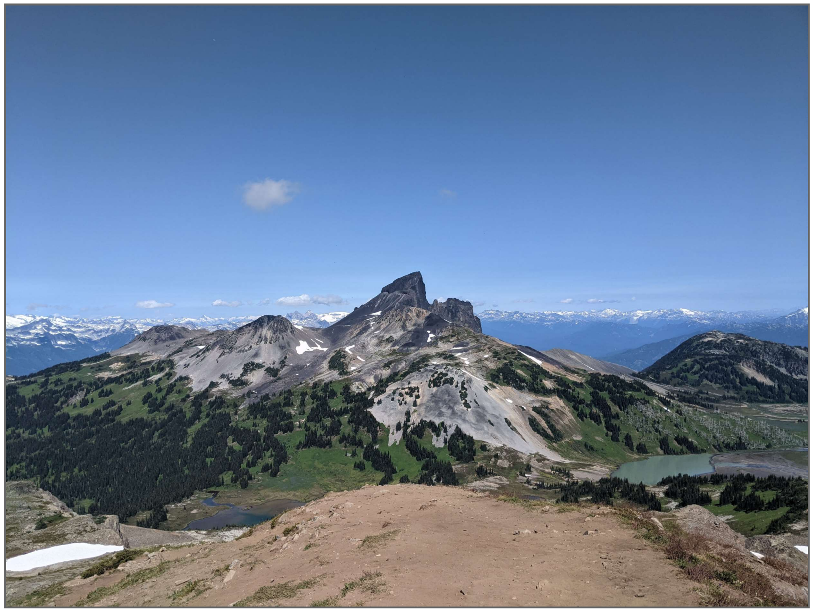
- Grade membership to assign categories
- Where to draw the line between forest/alpine?



GRADED MEMBERSHIP

Winner take all: alpine meadow

- 45% alpine meadow
- 40% forest
- 5% bare rock



GRADED MEMBERSHIP

The Downside: variability within the area is lost.

- In practice, lots of qualitative data we work with, especially for *natural phenomena*, are actually graded membership.

QUANTITATIVE DATA

Numeric; describe the quantities associated with a phenomenon.

- Values separated by a meaningful unit.
- More arithmetic operations possible.
- Can be **Discrete** or **Continuous** numbers.
- Measured on either a **Ratio** or **Interval** scale.

KINDS OF NUMBERS

Discrete:

- Whole numbers
- Counts
- Not infinitely divisible
- Integer, Long

Continuous:

- Decimals
- Measurements
- Infinitely divisible
- Float, Double

KINDS OF NUMBERS

Discrete:

- Countable
- Ex. **Population**

Continuous:

- Non-countable
- Ex. **Temperature**

QUANTITATIVE DATA

Discrete *and* continuous data can be measured on an **Interval** or **Ratio** scale.

- These types of quantitative data are closely related, but have **one important distinction**.

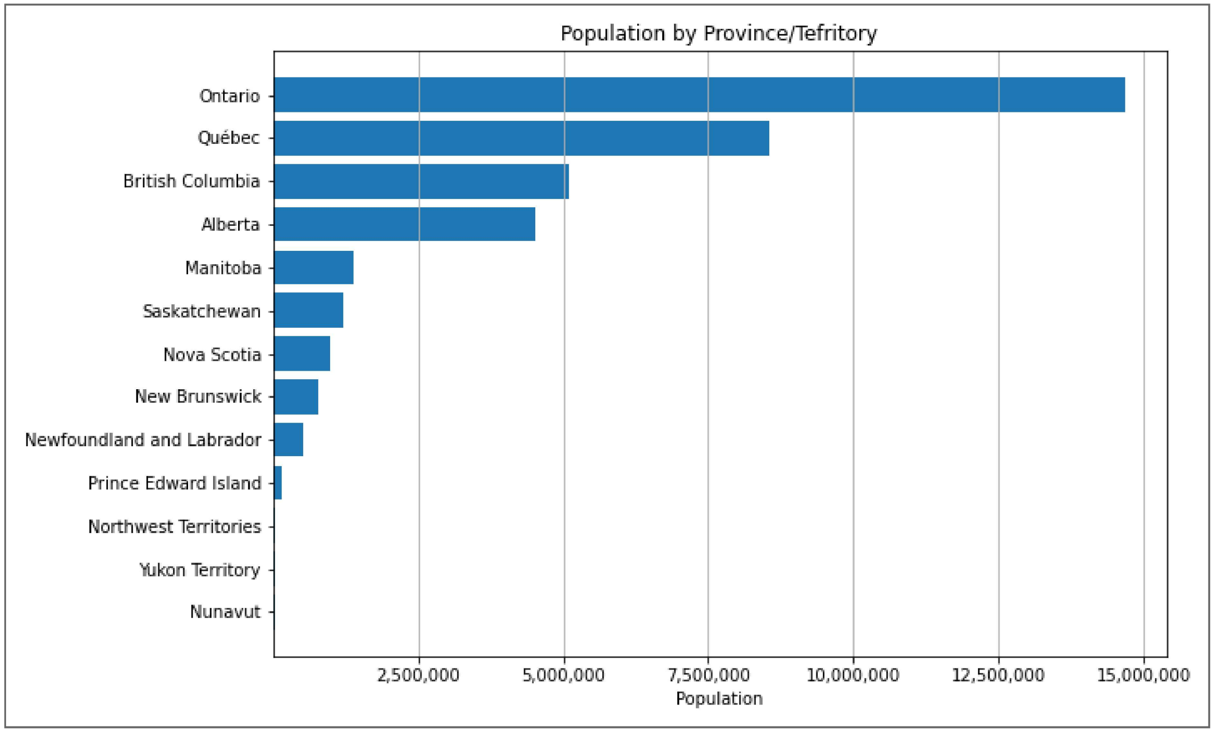
RATIO DATA

Fixed, **absolute zero point**.

- **Cannot** take negative values
- **Can** multiply/divide

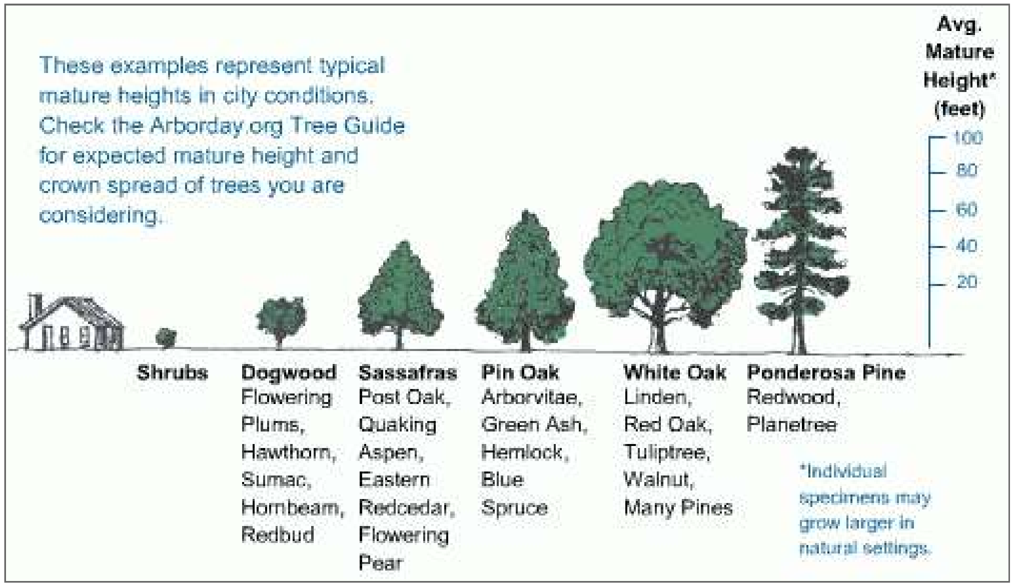
RATIO DATA

Population is a good example of *discrete* ratio data.



RATIO DATA

Tree height is a good example of *continuous* ratio data.



RATIO DATA

Other examples of ratio data include:

- Temperature in **Kelvin** (Continuous)
- Precipitation (Continuous)
- Units of time (Continuous)
- Rental cost (Discrete-ish)
- Popular Vote Totals (Discrete)

INTERVAL DATA

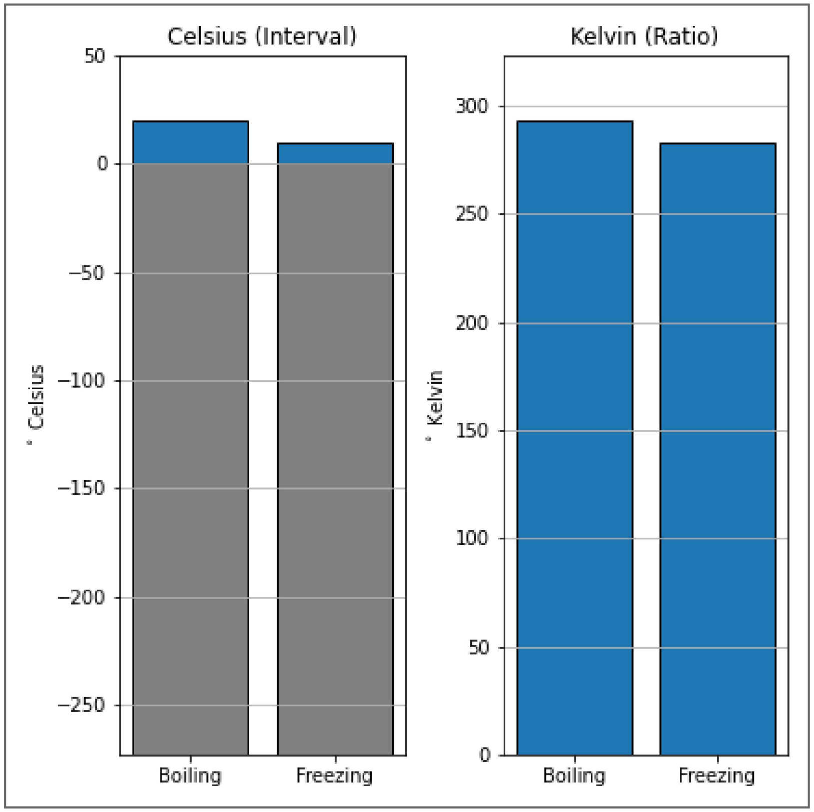
Arbitrary zero point

- **Can** take negative values
- **Cannot** multiply/divide

THE DIFFERENCE

Celsius (interval) vs. **Kelvin** (ratio).

- °C = °K-273.15.
- 0 °K: "Absolute Zero"
- 0 °C: Freezing point of water



INTERVAL DATA

Other examples include:

- ph scale (continuous)
- Dates (discrete)
- Times (discrete-ish)

DERIVED RATIO

If we want to account for the influence of one variable when analyzing another. Referred to as **Normalizing** or **Standardizing**.

- Formula is C=A/B
- A is the variable of interest
- B is the "confounding" variable
- C is the **Derived Ratio**

DERIVED RATIO

There are many circumstances where we might need to do this. ie. Housing affordability.

- **A**: My rent $1,018/mo
- **B**: I make ~2300/mo
- **C**: 44% my income goes to rent :(
- Income and rent (in $) are both discrete, Housing affordability is continuous.

DERIVED RATIO

Another would be population proportions