**Session (19-11-24):**

<https://native-land.ca/>

<https://www.whose.land/en/>

Is the visualization pleasing to look at? → **Aesthetic**

Does the visualization accurately and honestly present data? → **Substantive**

Can we understand what message the maker of the visualization is attempting to

convey? → **Perceptual**

We need to consider all of these

**Context** → where and how will our visualization be used? (eg. academic journal,

poster, infographic)

**Audience** → who is intended to use our visualization? (eg. subject experts,

general public)

**Data structure** → what information do our data capture? (eg. quantities,

relationships)

39:00

**Assignment 1:** Class code (Participation)

**Assignment 2:** (Look at some visualizations and evaluate if they are good or bad+how to make them good)

**Assignment 3:** (Look at some vis and determine if they are Ethical, Accessible, reproduceable … etc

**Assignment 4:** Own vis (Python + Another tool)

Coding: 43:10 (Video)

Libraries:

* Numpy
* Matplotlib
* Pandas
* Scipy
* seaborn

Coding Scope:

* Learn about matplotlib
* Produce our first data viz in Python
* Begin to modify elements of our data viz, including
  + Colour
  + Line type
  + Marker size
* Apply our visualization evaluation skills to matplotlib images

***Python Graph Gallery:*** <https://python-graph-gallery.com/>

Plot is called axes in MPL *(Not the same thing as the x and y axes on our graph)*

**Session (20-11-24): Reproducibility**

**Reproducible work** is “capable of being checked because the data, code, and

methods of analysis are available to other researchers”

**That is:** someone could repeat the steps we took to generate a particular result or

image from our data

**Reproducibility is ethical:** Good statistical practice is fundamentally based on transparent assumptions,

reproducible results, and valid interpretations

**Reproducibility is practical:** Reproducibility is practical, helps us to draw on previous work to make new graphics more easily, and useful for version control.

***Notes:***

* The ability to reproduce a result **does not** necessarily **indicate correctness**, nor does the inability to do so mean a result is incorrect.
* But “science is incremental: it is only through transparency and by enabling **reproducibility** that scientific **knowledge evolves**.”
* Customizing our plots
  + Adding legends
  + Annotating our plots
  + Working with colour
  + Pre-made styles
* Choosing the right graph for a given situation (Professional skills)

**Data visualizations as rhetorical objects**

* **Rhetoric** is the act of communicating effectively and persuasively
* From D’Ignazio and Klein (2020),

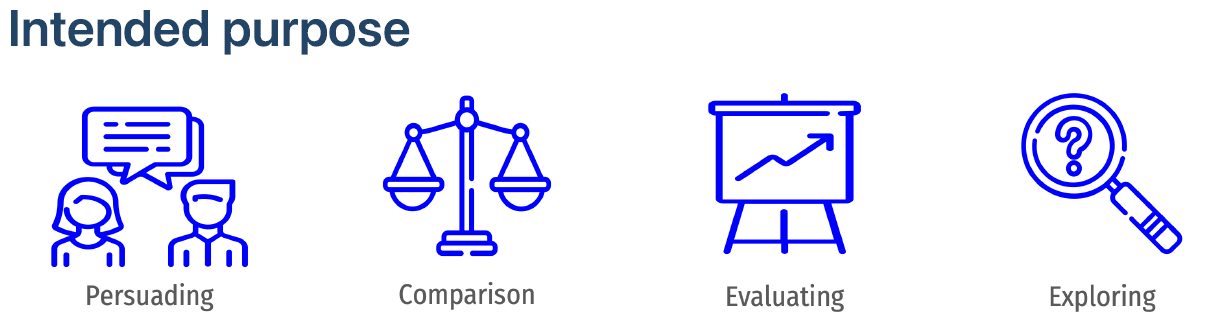
“Any communicating object that reflects choices about the selection and

representation of reality ***is a rhetorical object***. Whether or not it is rhetorical (it

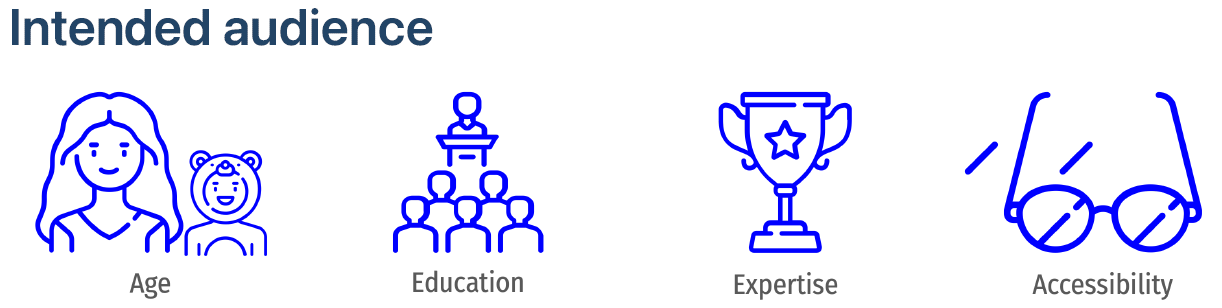
always is) has nothing to do with whether or not it is true (it may or may not be).”

* That is, we make choices about how to visualize our data, so these visualizations are not neutral…
* …BUT data visualizations can be factual without being neutral

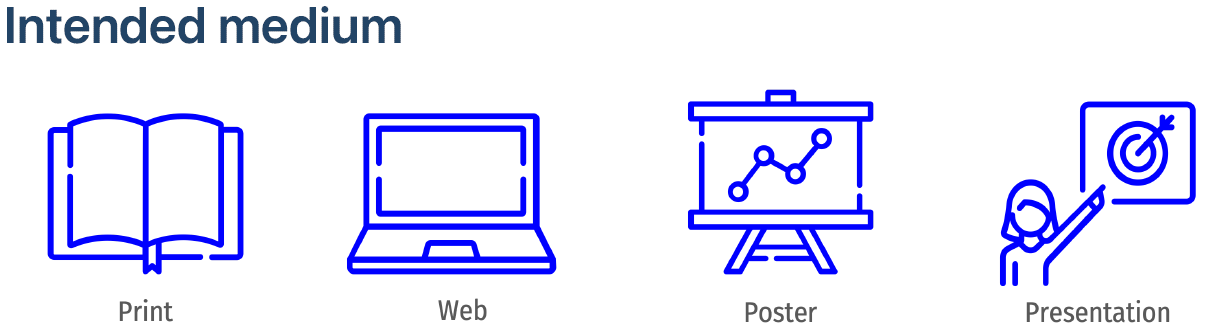
**Purpose** of a visualization (When designing one):



**Audience:**



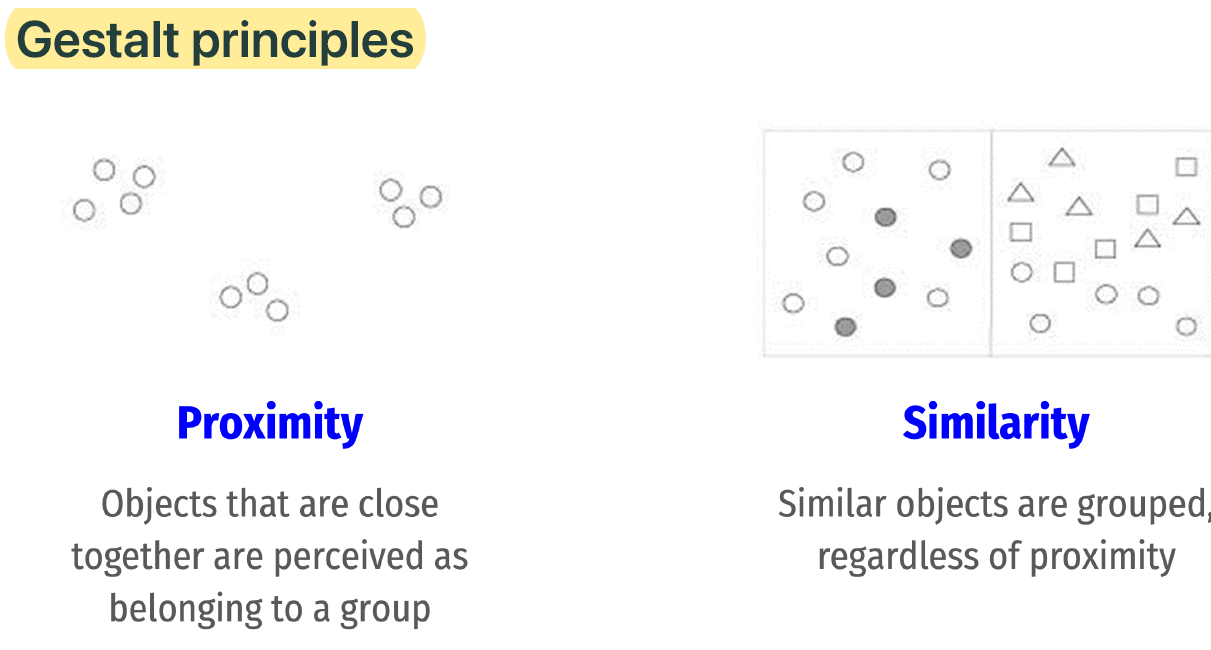
**Medium:**

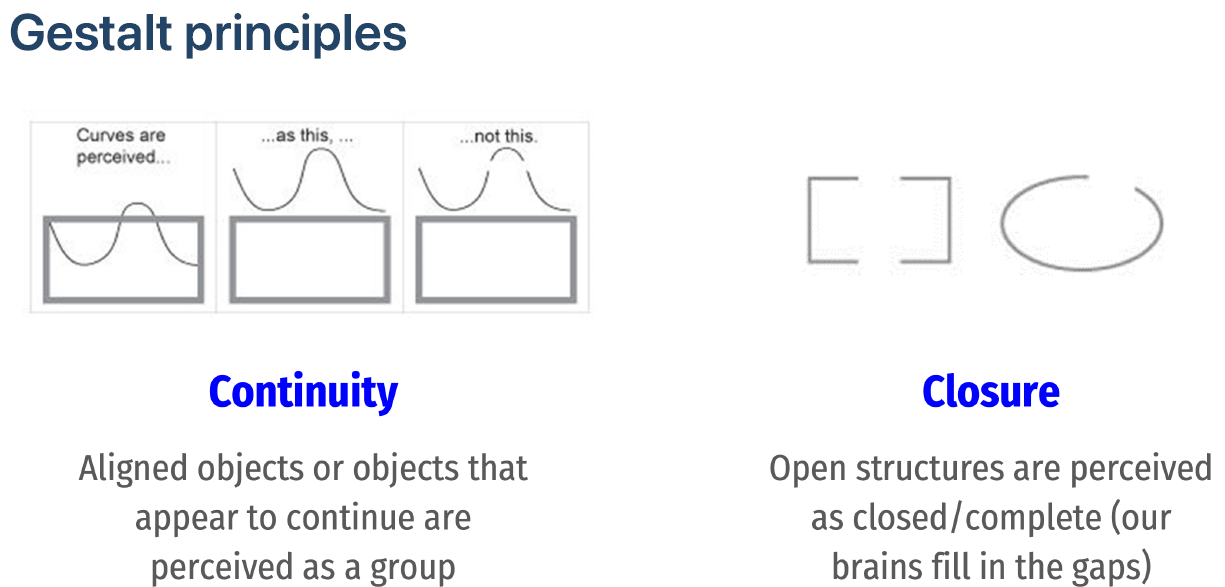


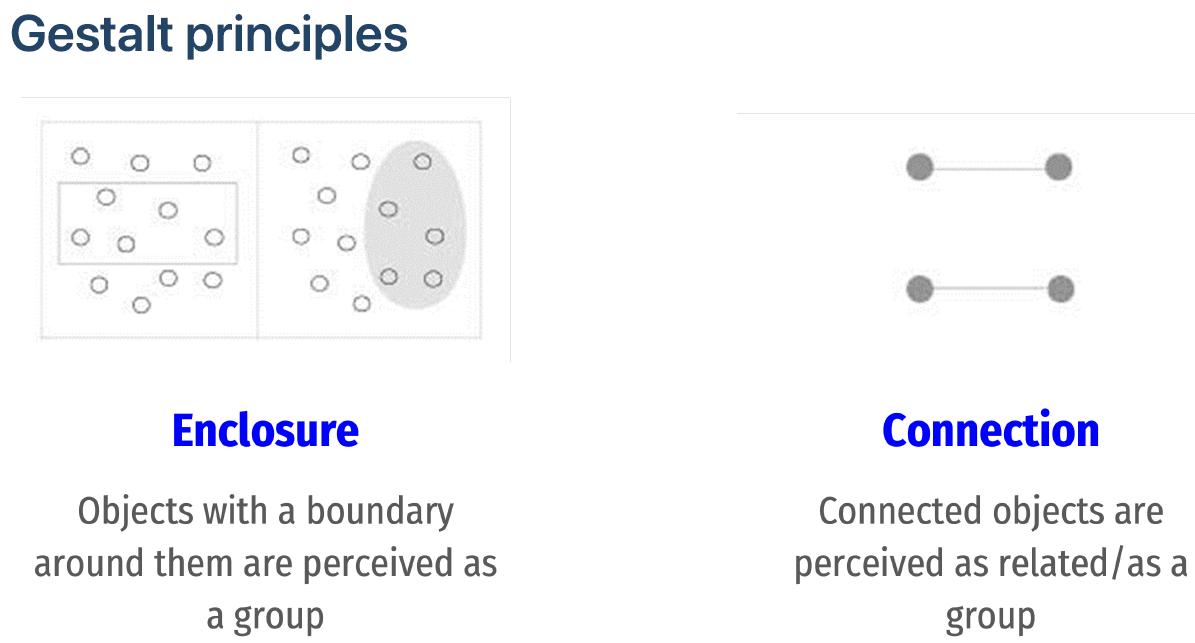
**Gestalt principles** (Gestalt is German for shape) are a set of cognitive theories for

how people tend to organize visual information; and are commonly used in UX design

and data visualization:







**Cognitive Load**

**Cognitive load** the amount of work required to take in new information:

* **Intrinsic** (the intrinsic complexity of the new information)
* **Germane** (the audience’s familiarity with the information)
* **Extraneous** (complexity from how the information is presented)

*In a data visualization context, extraneous cognitive load is most* ***within our control***

**Elements** of a visualization that can affect cognitive load include:

* **Familiar vs. Rare chart types** → rare types increase cognitive load
* **Accurate vs. Approximate interpretation** → relational values or areas (approximate) increase cognitive load compared to absolute values or position (accurate)
* **Concise vs. Detailed composition** → more visual elements increase cognitive load
* **Explanatory vs. Exploratory composition** → a chart that the audience navigates alone increases cognitive load compared to a chart that they are guided through step-by-step

**Perceived factual basis:** Conventions that help to perceive visualizations more objective and factual

* Two-dimensional image
* Clean layouts
* Geometric shapes and lines
* Inclusion of data sources at the bottom of the image (“transparency and trustworthiness”: makes the visual more persuasive - helps people to trust)

**Decision making tools (To choose the right visualization): include vis purpose and cognitive principles**

* The data visualization catalogue (Severino Ribecca): <https://datavizcatalogue.com/>
* Financial Times Visual Vocabulary sheet: <https://community.fabric.microsoft.com/t5/Data-Stories-Gallery/FT-Visual-Vocabulary-Power-BI-Edition/td-p/584460>

**Assignment 2:** Find 2 visualization (1 good and 1 bad: 3 reasons to classify them like this + 2 suggestions for improvement)

Ex. It’s good because it takes the advantage of Animation to convey emotion or because it takes advantage of providence rhetoric to convey that it’s very trustworthy.

How Vis is now and how can be better (Ex. Reduce cognitive load,

***No Coding in this session***

**Session (21-11-24): Customizing Plots**

***Starting with coding session***

Modifications to plots using Matplotlib

* Legends
* Annotations (text, shapes, and labels)
* Axis Labels
* Styles
* Learn about subplot notation in matplotlib
* Put multiple visualizations on the same axes objects
* Show errors
* Adjust figure layout
* Add images to plots

**How does matplotlib work?**

* A figure is like a container that holds a set of axes
* The axes is our actual plot or graph
* A figure can hold multiple axes (like subplots)
* Every visual element of our plots – colour, legends, axis titles and scales, text – is called an artist and belongs to an axes (not to a figure)

**Layouts**

* We can use layouts to make sure that our subplots fit neatly in our figure area
* There are two main kinds of layout we consider:
  + **Tight layout** adjusts subplots so tick labels, axis labels, and titles don't overlap or leave the figure area
  + **Constrained layout** works similarly except it also fits things like legends or colorbars

***Session ended in 06\_subplots\_and\_combining\_viz (21+26-11-24) – Slide 17***

**Session (26-11-24): (On Slides)**

**Coding at the beginning**



<https://python-graph-gallery.com/web-scatterplot-with-images-in-circles/>

<https://python-graph-gallery.com/web-scatterplot-with-images-in-circles/>



Two people discussing an experiment in a lab

**Session (27-11-24):**

Packages for class:

plotly  
wordcloud  
plottable  
seaborn  
matplotlib-venn