

Brief Intro to Geospatial Data

Modified from “[Overview of Geospatial Data](#)”

by Patty Frontiera, Drew Hart, and Hikari Murayama of UC Berkeley’s D-Lab,
used under [CC BY-NC-SA 4.0](#).

Geographic data



A single unit of geographic data includes:

Location (where) : Anatone

Attributes (what) : *data that describe the location*

Also great to have metadata:

When : 2003

Who : Anatone 4-H

How : local census?

Geospatial data

Encodes location geometrically with coordinates:

Anatone, WA: 46.130479, -117.134167



BUT, how do we know that those coordinates reference that specific location?

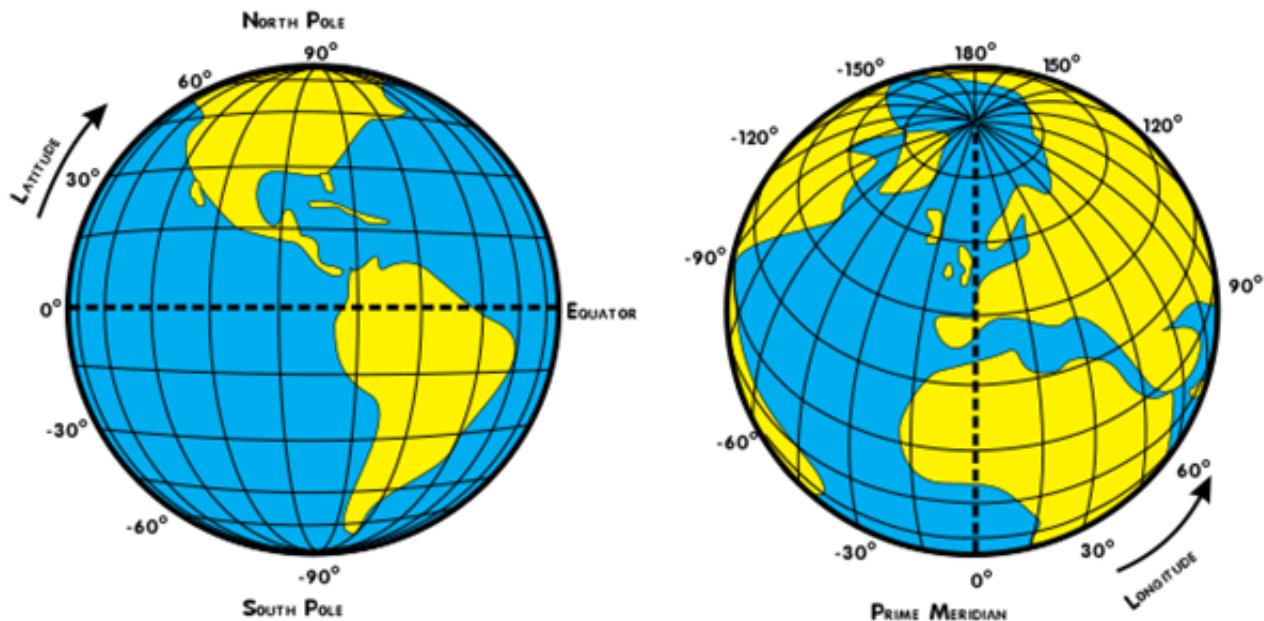
Coordinate Reference Systems (CRS)

A Coordinate Reference System, or CRS, is a system for associating coordinates with a **specific, unambiguous** location on the surface of the Earth.



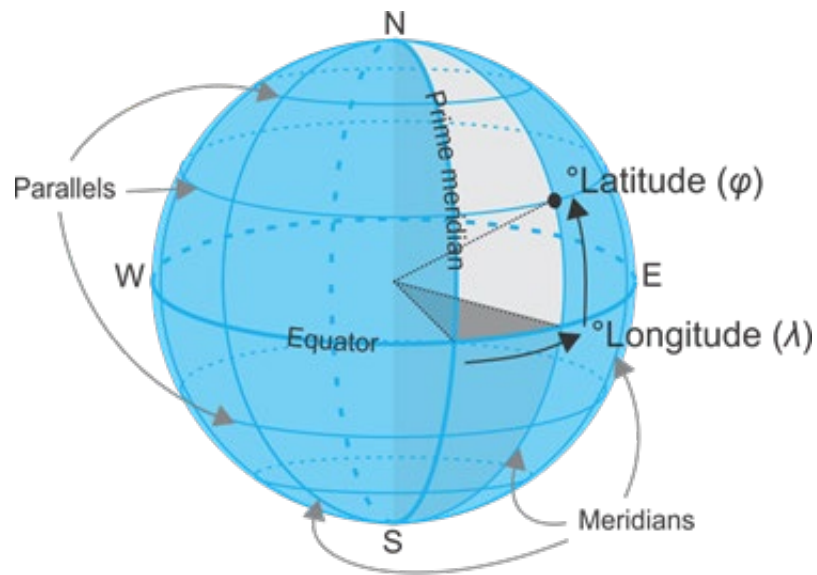
New Orleans = 30N, 90W

Coordinate Reference Systems (CRS)



Geographic Coordinates: *Latitude* and *Longitude*

Coordinate Reference Systems (CRS)

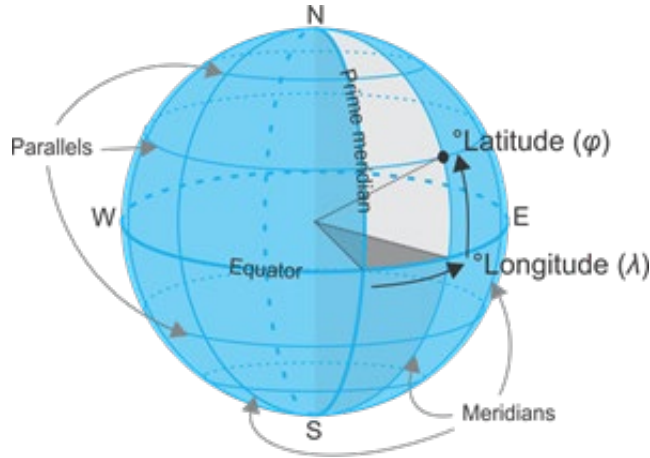


There are **many** CRSs, not just one!

Why? Because our understanding of and ability to measure the shape of the earth has changed over time.

Two Types of Coordinate Reference Systems

Geographic CRS



Angular Units = Degrees (DMS or DD)

Projected CRS



Cartesian Units = Feet or Meters
Good for local & regional mapping & analysis

Geographic Coordinate Systems (GCS)

Widely used! Expressed as latitude & longitude

WGS84 (EPSG: 4326)

Based on satellites, used by cell phones, GPS

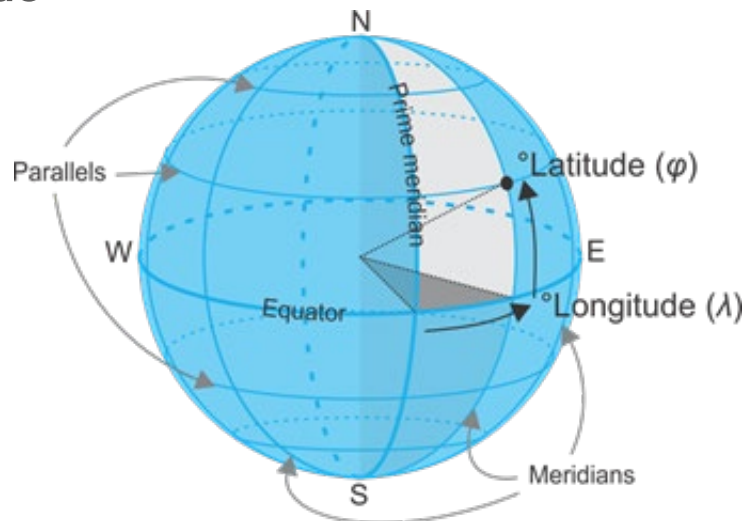
Best overall fit for most places on earth

NAD83 (EPSG: 4269)

Based on satellites and survey data

Best fit for USA

Used by many federal data products, like Census data



CRSs are referenced in software by numeric codes, often called **EPSG codes**

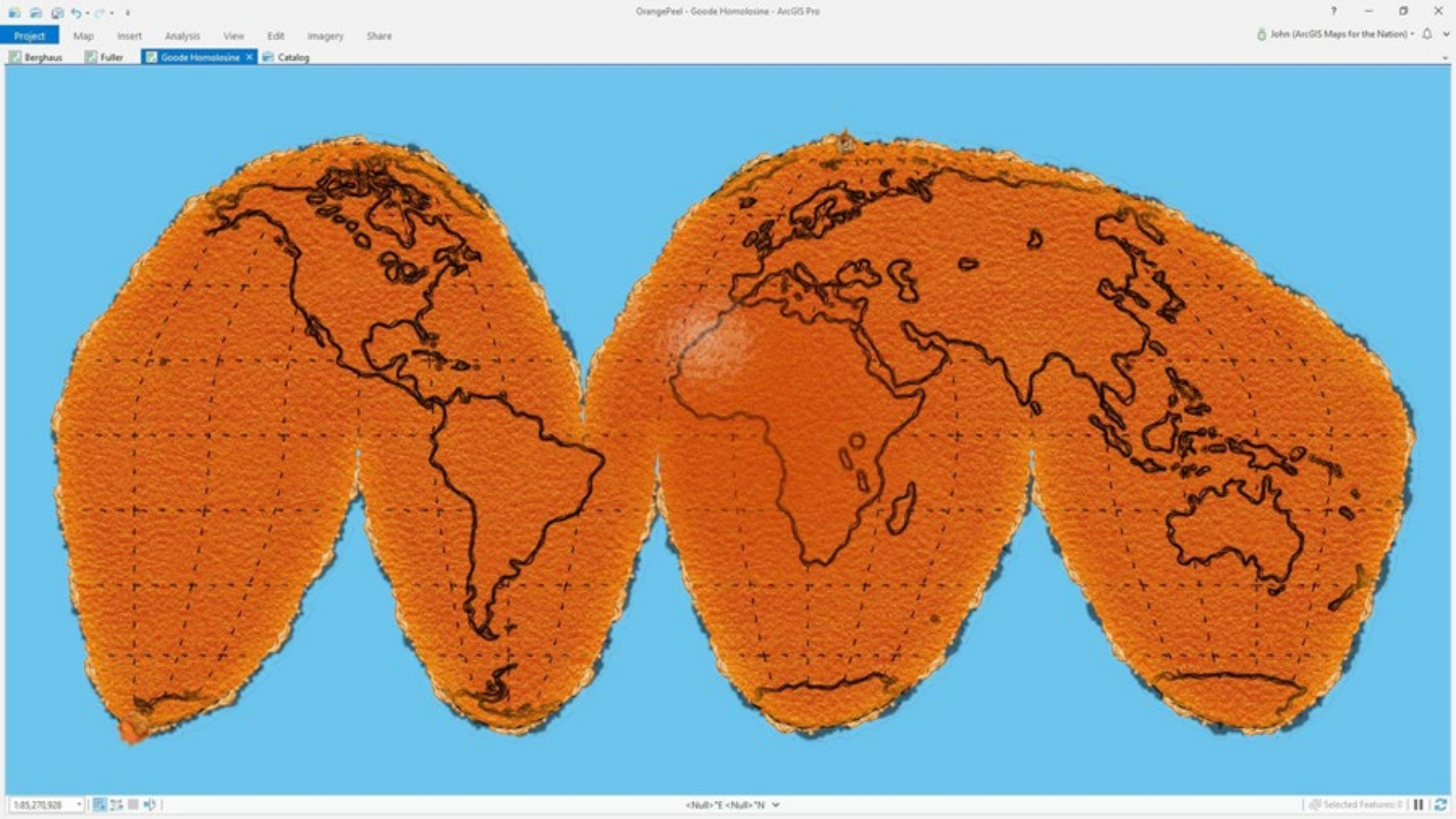
HI THERE, I'M AN ORANGE!

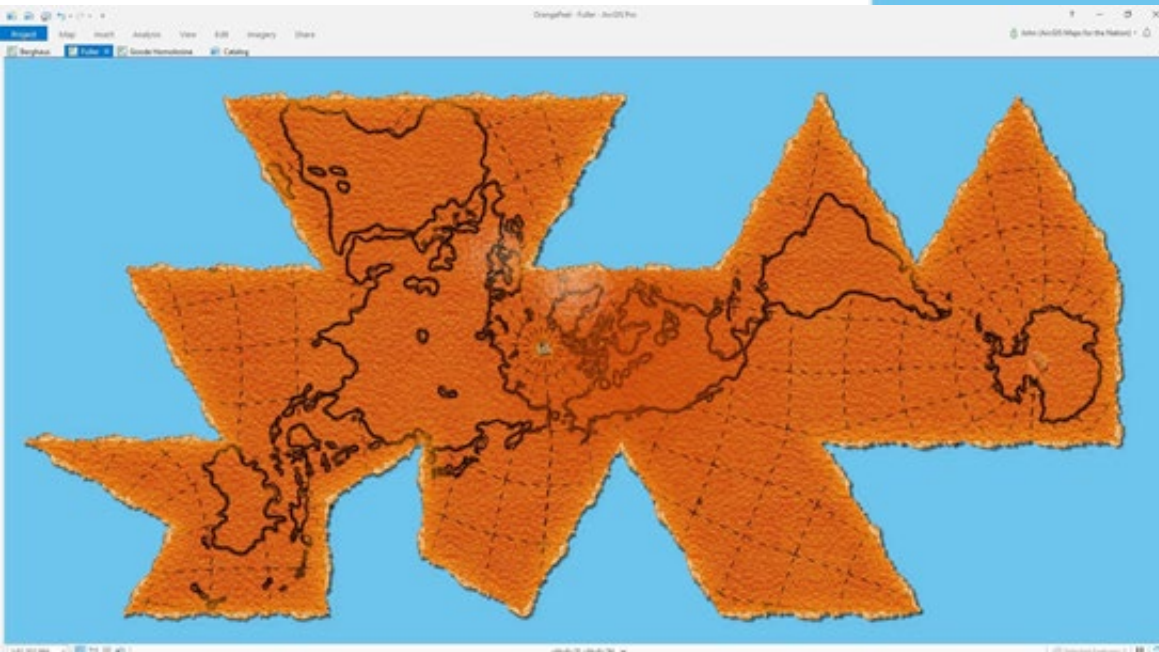
YOU MIGHT RECOGNIZE ME FROM EVERY
PROJECTION METAPHOR EVER.

IF ONLY THERE WAS A WAY I COULD LIVE
IN YOUR COMPUTER, AND WARP MYSELF
INTO ALL YOUR VIRTUAL PROJECTIONS!
AND MAKE THEM SMELL AWESOME.









Projected Coordinate Systems (PCS)

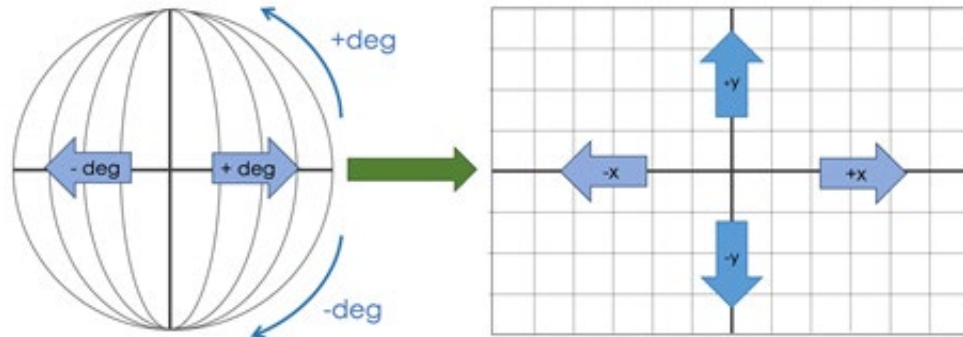
Map Projections transform geographic coordinates (lat/lon) to 2D coordinates (X/Y)

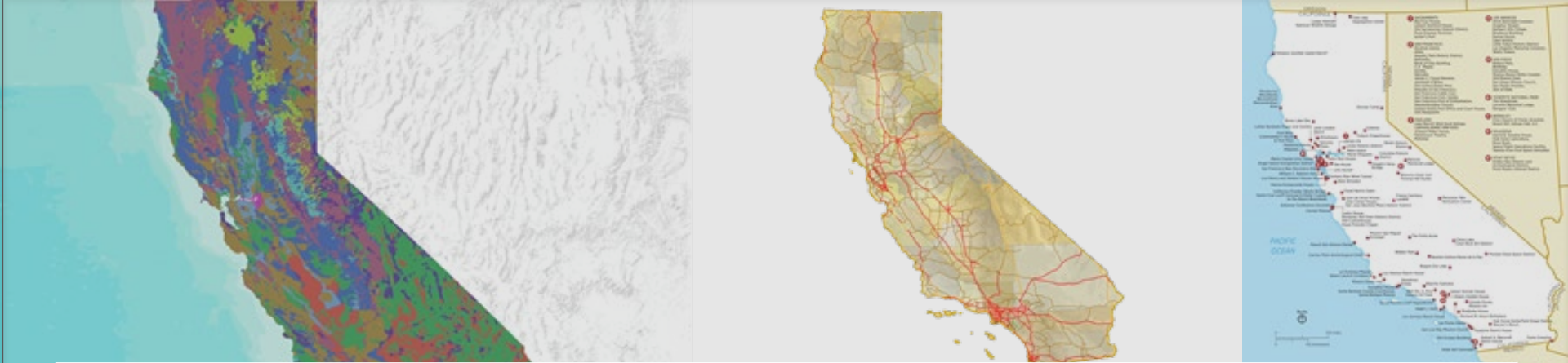
All map projections introduce **distortion** in area, shape, distance or direction.

Specific map projections minimize distortion in one or more properties

You need to know the coordinate reference system of your input data

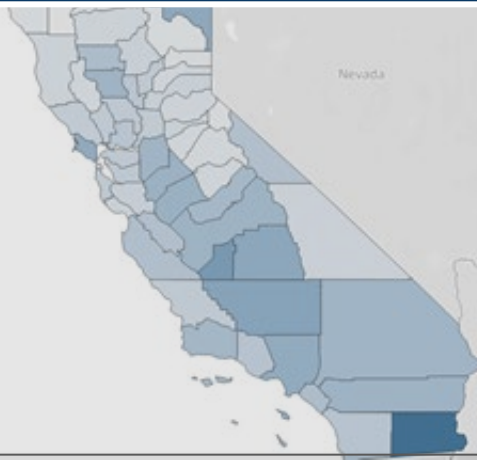
You need to select the CRS that is most suitable for your data and application.





Types of Spatial Data

**California
coronavirus map**



Types of Spatial Data

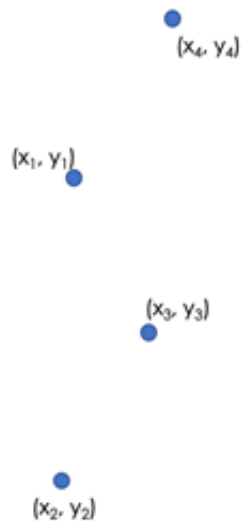
There are two fundamental spatial data models:

- Vector
- Raster

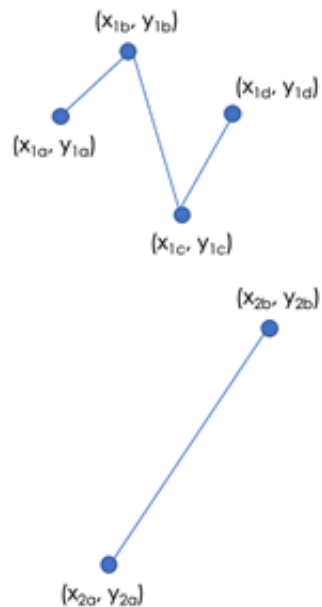
Vector Data

“Connect the dots”

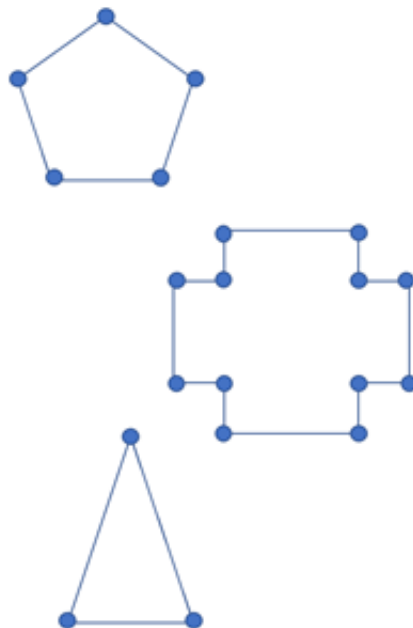
Points



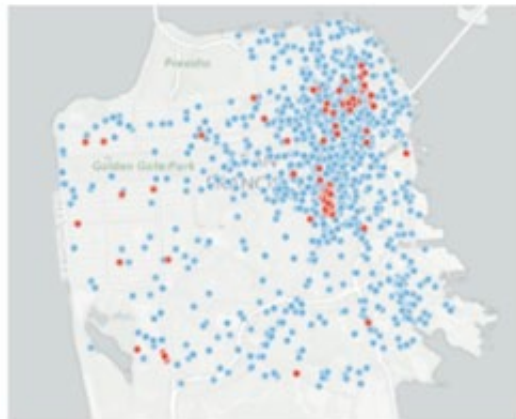
Lines



Polygons



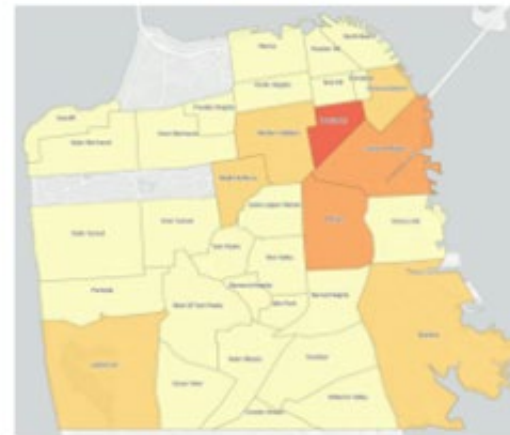
Points, Lines, Polygons



Crime locations



City freeways



Neighborhoods

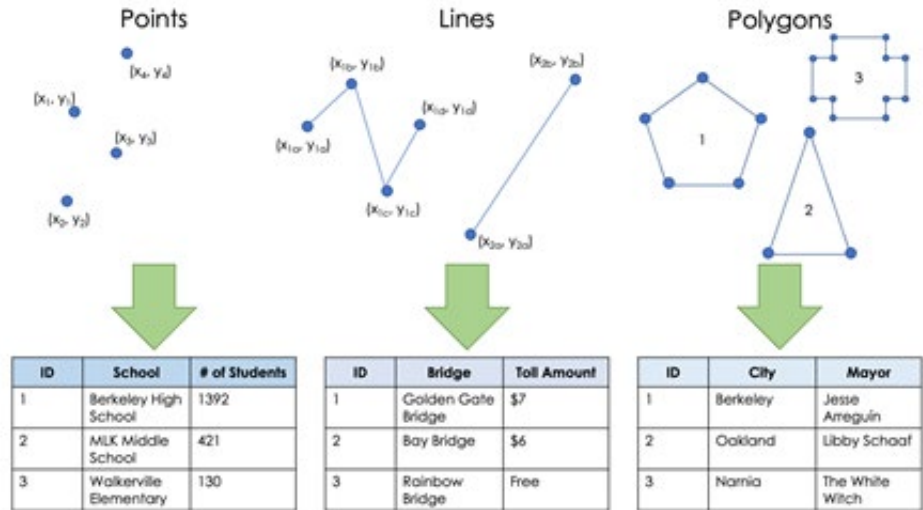
Vector Data with Attributes

Each row represents one geospatial
feature

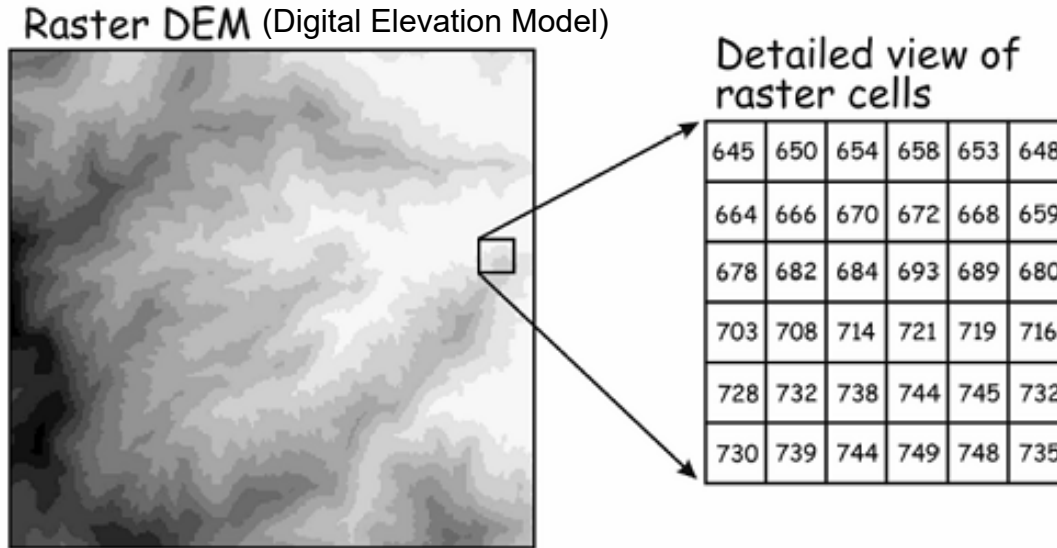
Attributes describe the features
(*fields or columns*)

Each feature has an associated
geometry or geometry collection

A group of features is called a **layer**



Raster Data - regular grids



A location is represented by a grid cell

Cells have regular size, eg 30x30m

Grid has dimension - fixed number of rows and columns

Each cell has a value that represents the attribute of interest, e.g. elevation

Imagery Data are Raster Data



Note:

Aerial imagery, satellite data and other remotely sensed geographic data are commonly used as sources of vector data

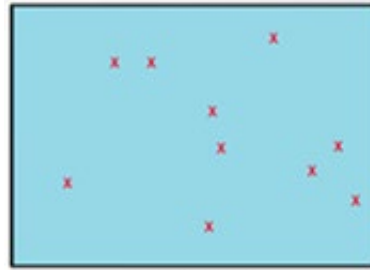
In other words, the building footprints or streets can be digitized off of the imagery and saved as vector data.



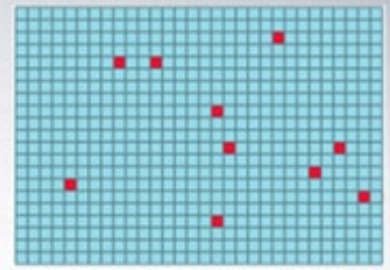
Vector vs Raster

Vector data are better for
discretely bounded data
e.g. political boundaries, fire
hydrants, rivers, roads, etc.

Raster data are better for
continuous data
e.g. temperature, elevation,
rainfall, etc.



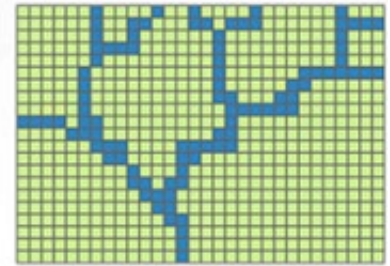
Point features



Raster point features



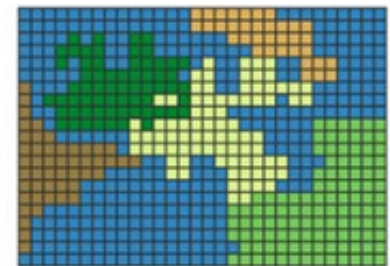
Line features



Raster line features



Polygon features



Raster polygon features

Some Common File Formats

Vector Data

- Shapefile (.shp...)
- GeoJSON, JSON
- KML
- GeoPackage

Raster Data

- GeoTIFF
- netCDF
- DEM

Georeferencing

Data layers in the same coordinate reference system can be linked dynamically to explore associations and build complex models of the real world

