# Brief Intro to Geospatial Data

# Geographic data


A single unit of geographic data includes:

**Location (where)**: Anatone
**Attributes (what)**: *data that describe the location*

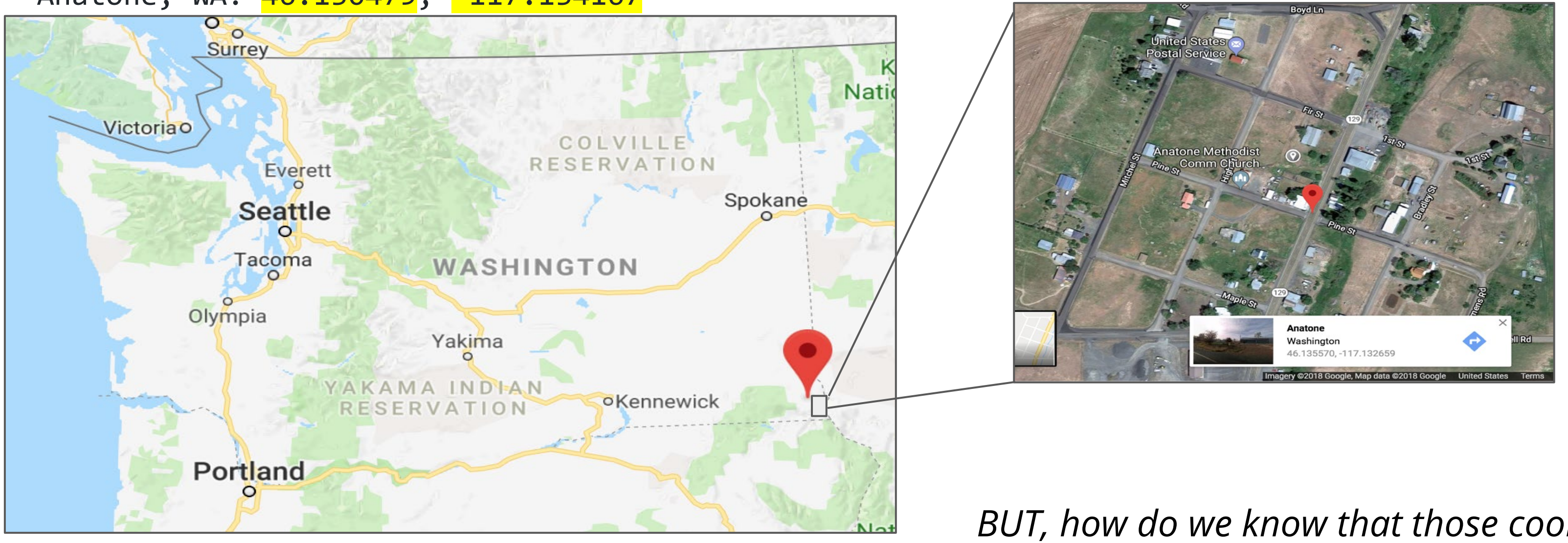
Also great to have metadata:

**When**: 2003
**Who**: Anatone 4-H
**How**: local census?

# Geospatial data

Encodes location geometrically with coordinates:

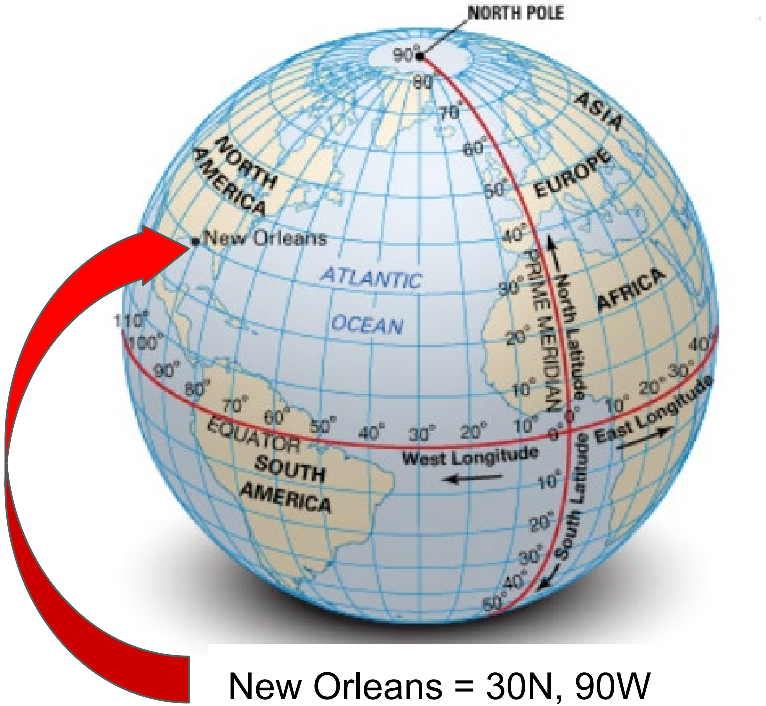
`Anatone, WA: 46.130479, -117.134167`




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

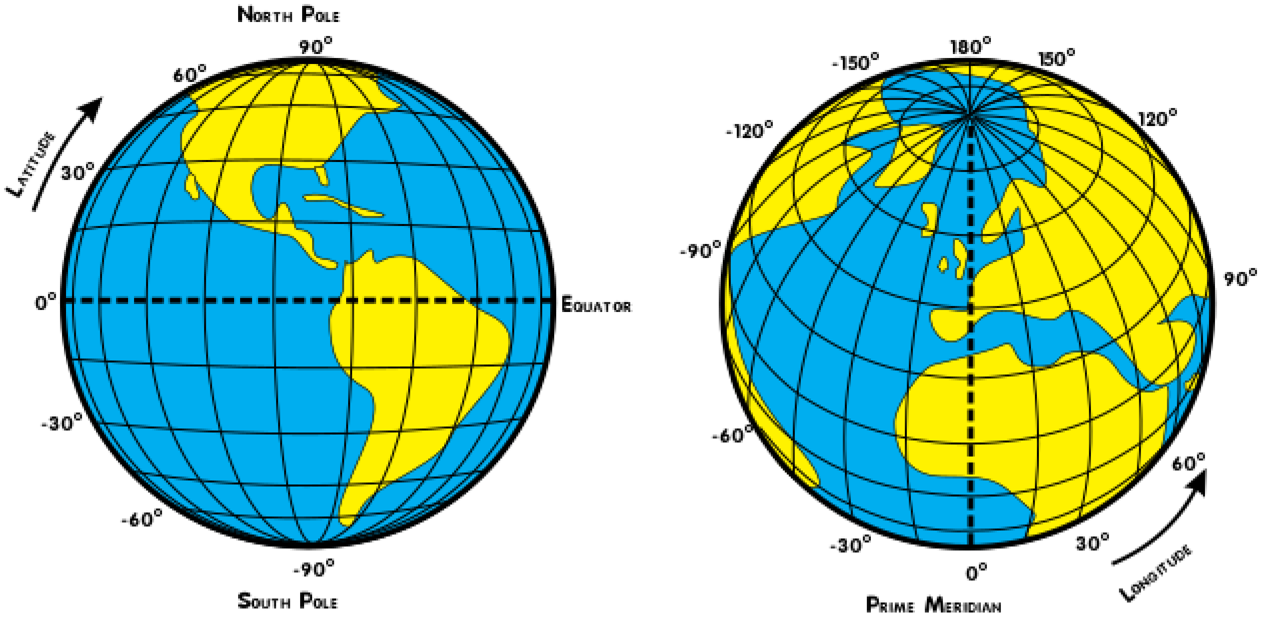
# Coordinate Reference Systems (CRS)

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

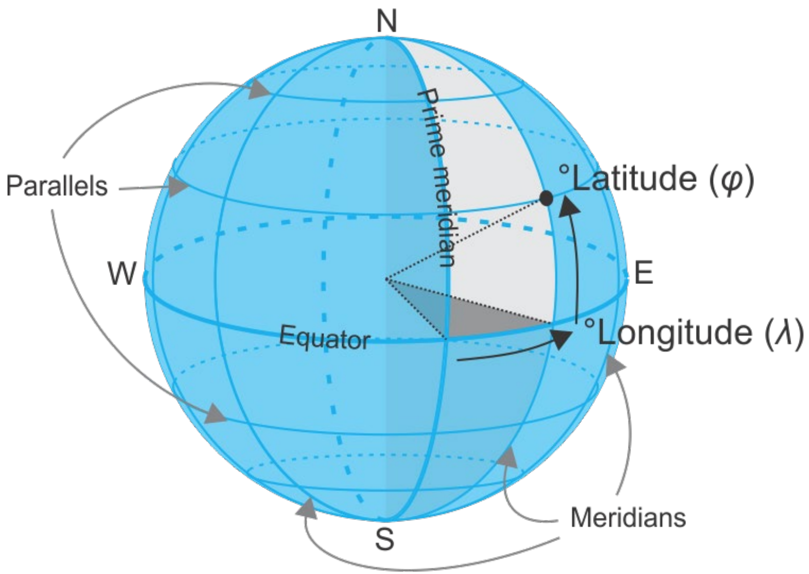
# Coordinate Reference Systems (CRS)




Geographic Coordinates: *Latitude* and *Longitude*

http://latitude-longitude.net

# Coordinate Reference Systems (CRS)

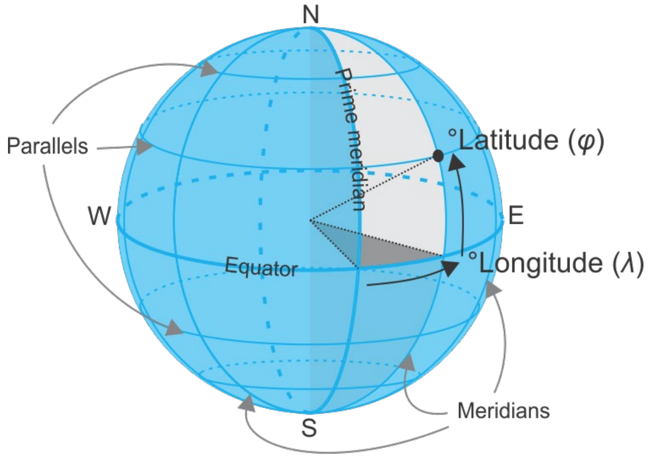



There are **many** CRSs, not just one!

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

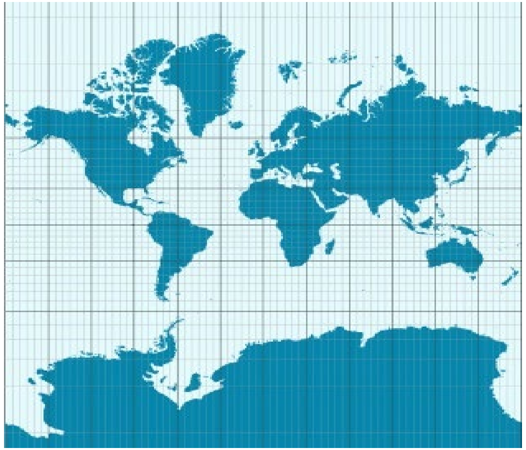
# Two Types of Coordinate Reference Systems

## 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

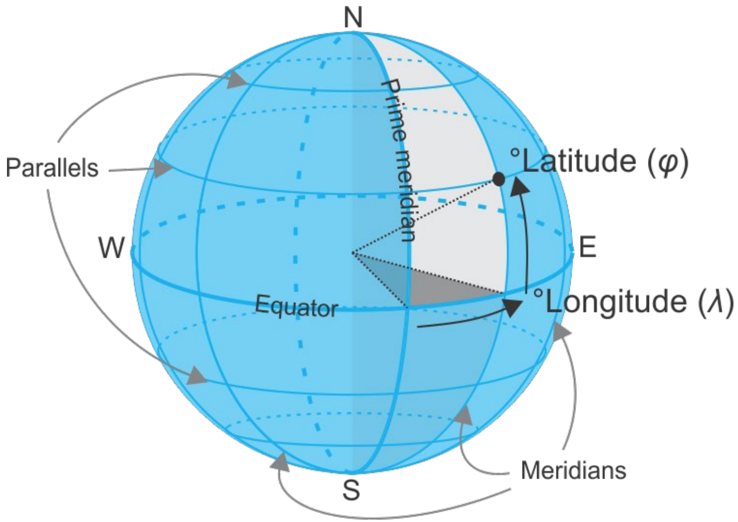
Best overall fit for most places on earth

**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**

# Projected Coordinate Systems (PCS)

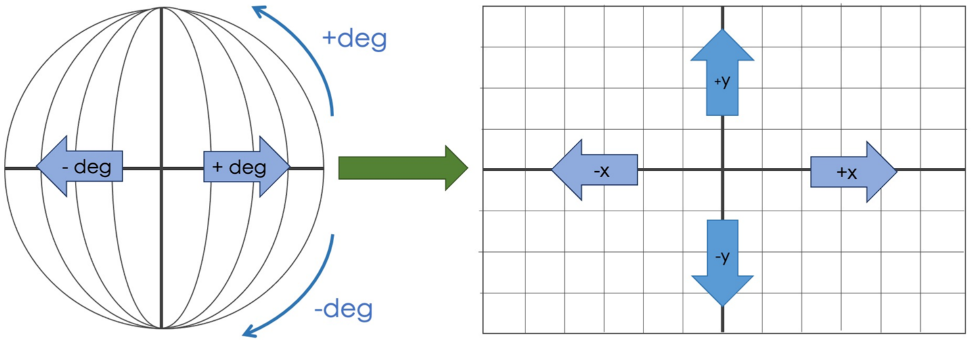
**Map Projections** transform geographic coordinates (lat/lon) to 2D coordinates (X/Y)

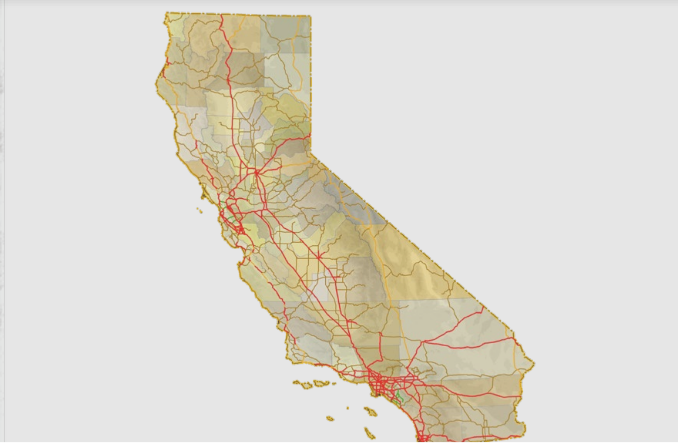
All map projections introduce **distortion** in area, shape, distance or direction.

Specific map projections minimize distortion in one or more properties

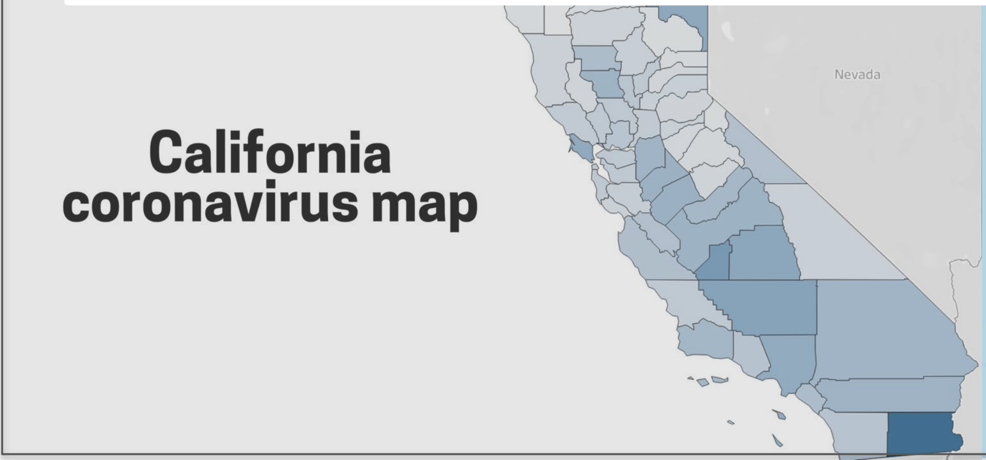
You need to know the coordinate reference system of your input data

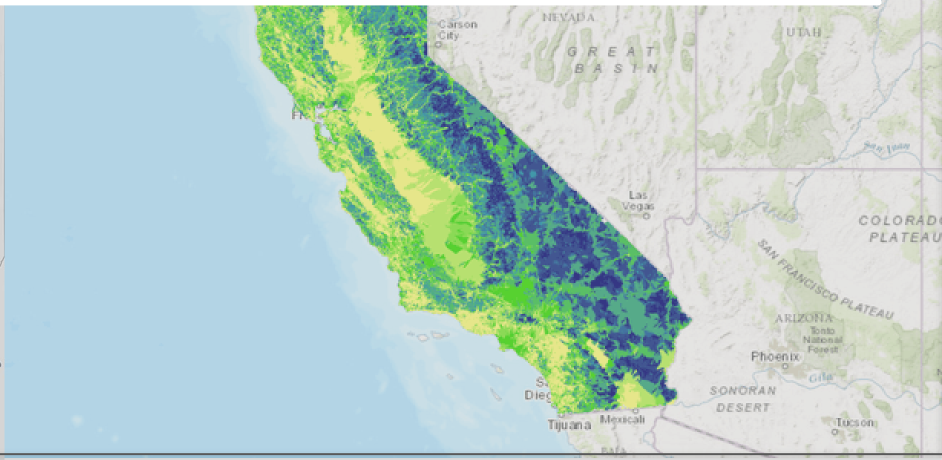
You need to select the CRS that is most suitable for your data and application.






# Types of Spatial Data



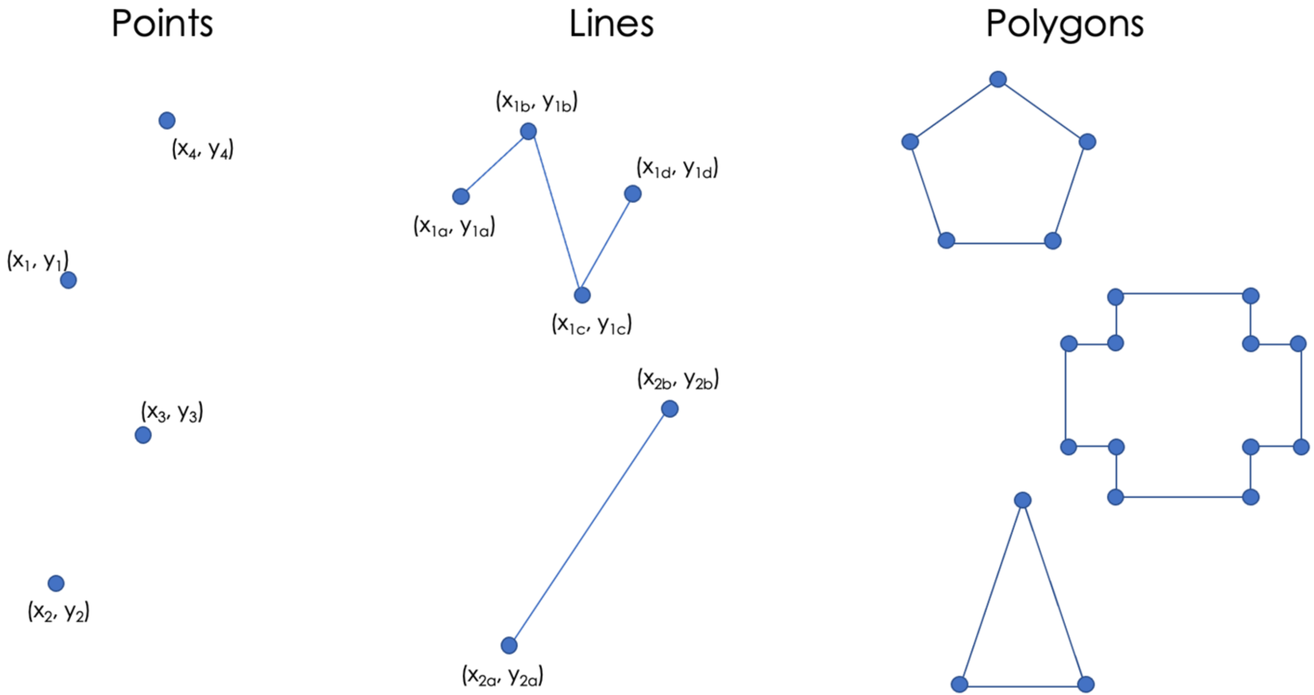
# Types of Spatial Data

There are two fundamental spatial data models:

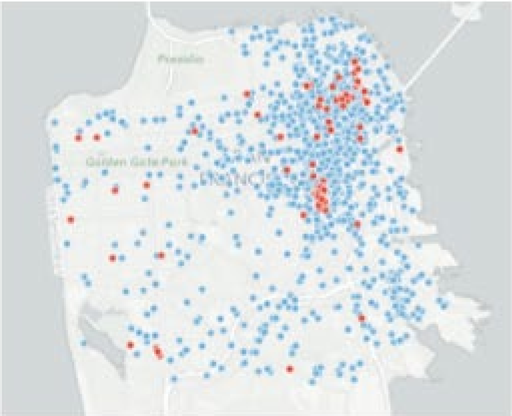
- Vector
- Raster

# Vector Data

"Connect the dots"

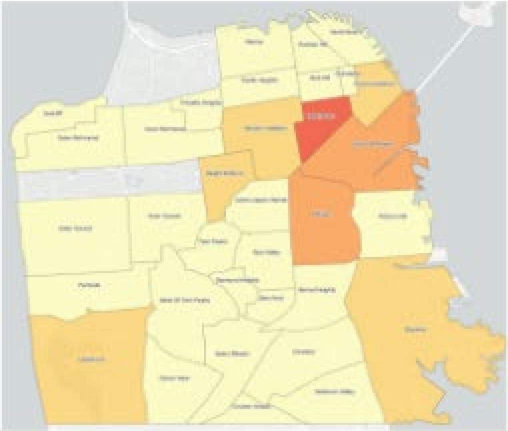


# Points, Lines, Polygons




Crime locations


City freeways



Neighborhoods

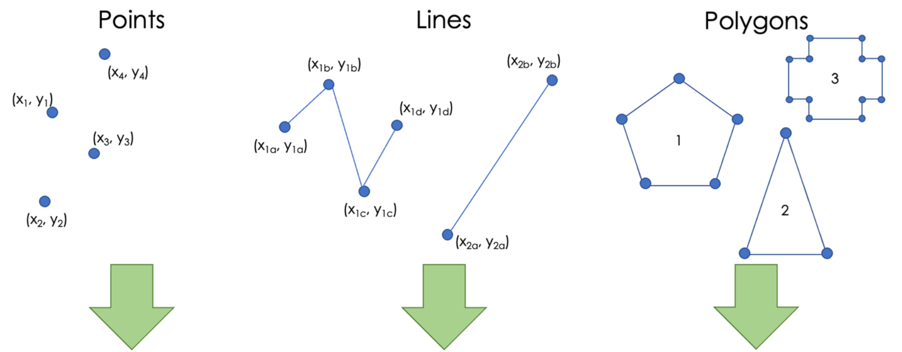
# Vector Data with Attributes

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**

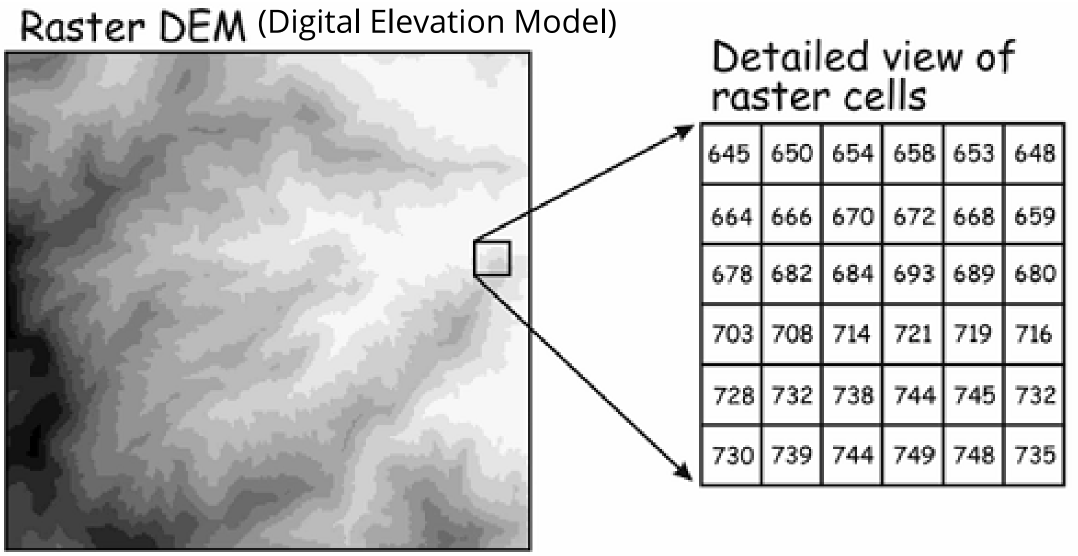



| ID | School | # of Students |
|---|---|---|
| 1 | Berkeley High School | 1392 |
| 2 | MLK Middle School | 421 |
| 3 | Walkerville Elementary | 130 |

| ID | Bridge | Toll Amount |
|---|---|---|
| 1 | Golden Gate Bridge | $7 |
| 2 | Bay Bridge | $6 |
| 3 | Rainbow Bridge | Free |

| ID | City | Mayor |
|---|---|---|
| 1 | Berkeley | Jesse Arreguín |
| 2 | Oakland | Libby Schaaf |
| 3 | Narnia | The White Witch |

# **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

Source: BING Aerial imagery

**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

<u>Vector data are better for **discretely bounded** data</u>
e.g. political boundaries, fire hydrants, rivers, roads, etc.

<u>Raster data are better for **continuous** data</u>
e.g. temperature, elevation, rainfall, etc.

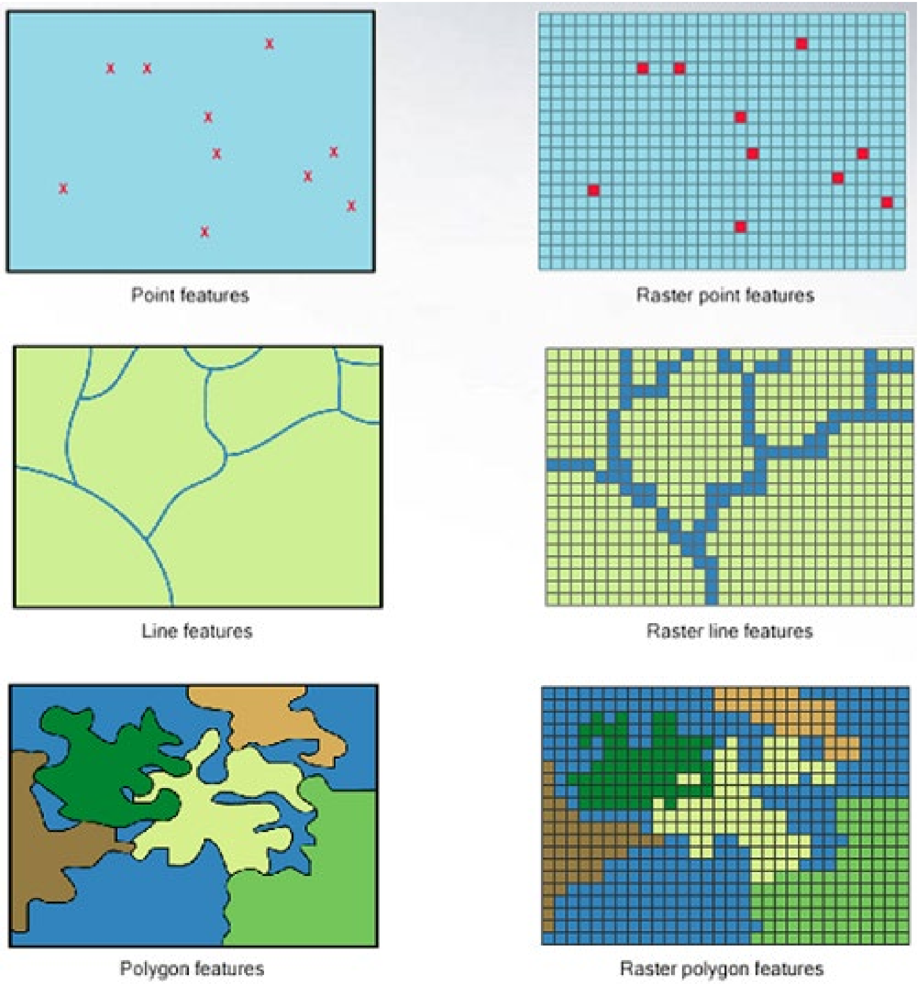



Image: http://gsp.humboldt.edu/

# 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