DATA 301
Introduction to Data Analytics
Geographic Information Systems

Dr. Ramon Lawrence
University of British Columbia Okanagan
ramon.lawrence@ubc.ca

Why learn Geographic Information Systems?

Geographic Information Systems (GIS) are used in a wide variety of areas for the analysis and display of spatial and geographical data:

- City and infrastructure planning
- Business development, planning, forecasting
  - Store placement, sales trends
- Population forecasting and analysis
- Environmental and water
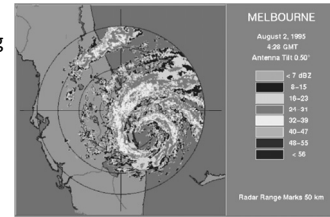
What is a Geographic Information System?

Geographic Information Systems are systems designed for storing, manipulating, analyzing, and displaying spatial and geographical data.

A GIS will contain components:

- for importing data from various sources in different formats
- organizing the data into layers or groups
- composing and integrating data to produce new information
- displaying the data visually as maps or 3D visualizations to help interpret the data

---

GIS History

The technique of *overlaying* information on maps dates back long before computers.

In the 1960s geographic mapping software developed the basic GIS concepts. First GIS developed by Dr. Roger Tomlinson for Canadian Department of Forestry and Rural Development.

In 1969, Environmental Systems Research Institute (ESRI) founded by Laura and Jack Dangermond and developed suite of products.

- Software linked spatial representation of features with table attributes
- ESRI now the de facto standard for commercial GIS products (e.g. ArcGIS, ArcView).

GIS Features

A GIS allows a user to add (or overlay) data on a map including:

- Annotation (text) - name or description of item/feature (e.g. city name)
- Point - a single (x,y) co-ordinate on the map
- Line - a connected pair of two (x,y) points
- Polygon - three or more (x,y) points connected to form a closed shape

Each feature can have one or more additional attributes (data items) describing it.

For example, a city could be represented as a point on the map (with coordinates) and additional attributes include its name and population.

GIS Data Types

The GIS feature consists of coordinates placing it on the map.

Each feature can have attributes providing additional information.

These attributes may have data types such as:

- Text - for names and labels
- Categories - for grouping similar features/classes (road, land use, etc.)
- Numbers - for measurements (population, rainfall, etc.)

Data is often placed in categories for display. Each item in the category has the same symbol.

- Measurement data is also grouped into categories to ease understanding even if the data is continuous. For example, ordinal data has categories ranked according to a scale: low, medium, high

GIS Interval Data

Interval data has values along a regular numeric scale.
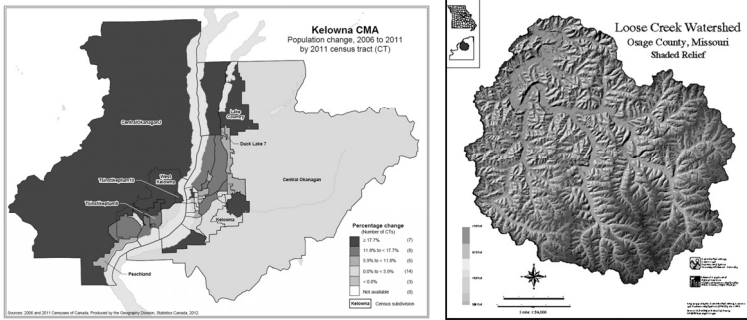
GIS Terminology: Scale and Precision

Scale is the ratio of size on the ground to size on the map.

Precision is a measure of how accurate the map representation is compared to the real-world.

- If map symbol is 1 point (72 points=1 inch) and scale is 1:100000, then symbol S = 100,000 points = 1388 inches = 115 feet is uncertainty in placement based on representation.

Resolution is the sampling distance of the stored x-y values.
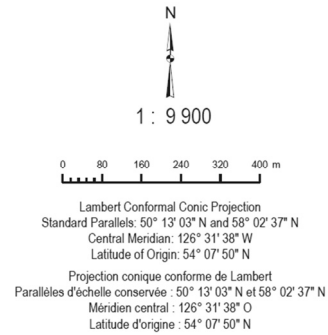
Scale Question

Question: If the scale of the map is 1:100000, and the feature on the map is 3 cm long, how long is the feature in the real-world?

A) 1 cm   **B)** 3 cm   **C)** 300000 m   **D)** 300 m   **E)** 3 km

Precision Question

Question: If the scale of the map is 1:100000, and a road is represented on the map by a line that is 2 mm think, what is the error in representation (i.e. how far off can the road be in real-life)?

A) 2 km   **B)** 200 m   **C)** 20 m   **D)** 2 m   **E)** No error

Feature Classes and Layers

A ***feature class*** is a collection of objects with the same attributes.
- May be stored as individual rows of a single table
- Have the same geometry (e.g. all points or all polygons).
- Example classes: states, cities, rivers

A ***layer*** is a grouping of features that can be added or removed from the map (and its display visualization).
- A layer will often reference or use a feature class.

Feature Class and Layer Screenshot
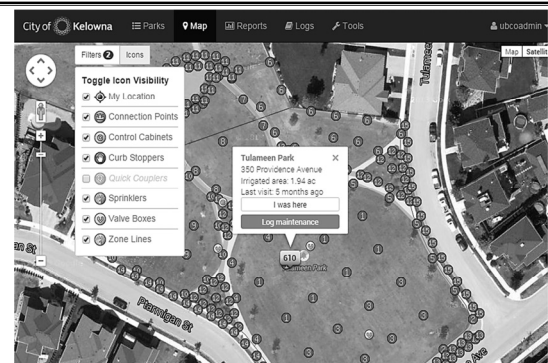
Representing GIS Data: Raster and Vector

There are two common methods for representing GIS data.

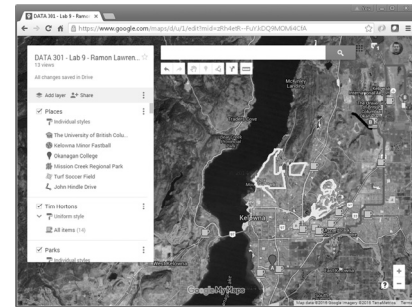*Raster representation* uses a matrix of data values.

Vector representation adds features (points, polygons) onto a map each with its own coordinates and attributes.

Vector Representation

Vector representation adds features onto a map with their own coordinates and attributes.

- Features stored as series of x-y coordinates and may be points, lines, polygons.
- Features are linked to a row in a data table which may have multiple attributes describing it.

Allows for very precise specification of features by coordinates which may have multiple attributes.

Raster Representation

A *raster* stores data as a matrix of data points that is georeferenced to earth's surface.

- The value at each data point may be discrete or continuous.
- Scanned images are continuous.
- Often used to store continuously changing values such as elevation.

Resolution measured by cell size.

- Since space is N^2, smaller cell sizes result in much large raster files
- Raster formats: GRID, geoTIFF (both georeferenced), TIFF

urban area — raster data

river — vector data
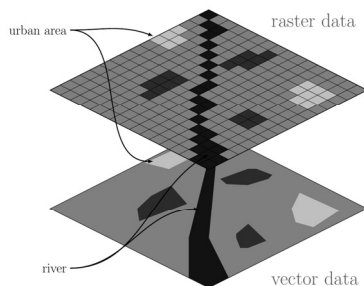
Raster vs. Vector

Vector - Advantages
- precision of coordinates
- may have multiple attributes per feature
- less storage space required
- flexible cartography

Disadvantages:
- hard to perform surface analysis
- not easy for continuous data storage

Raster - Advantages:
- simple, robust format
- implicit georeferencing
- stores continuous data
- surface analysis, faster analysis

Disadvantages:
- storage space
- lower precision

Raster versus Vector Question

Question: How many of the following statements are **TRUE**?

1) In a raster if the cell size decreases by half, the space increases by 4 times.
2) Rasters are often useful for scanning or remote sensing applications.
3) A raster typically stores only one numeric data value per cell.
4) A vector format may allow multiple attributes to describe each feature.
5) Vector representation is better suited for continuous data than rasters.

A) 1   **B)** 2   **C)** 3   **D)** 4   **E)** 5

Representing Geographical Data

A geographic data set requires a description of its coordinate system for display and analysis, often called the spatial reference.
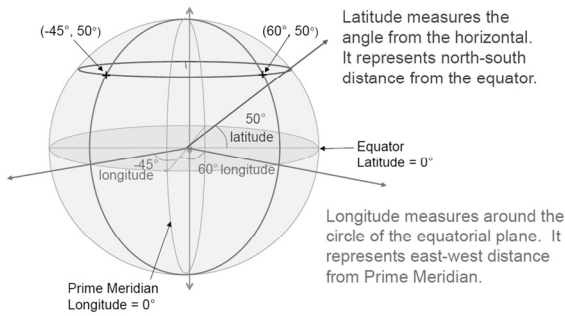
Components:
- Geographical coordinate system (GCS) / datum - for assigning coordinates to points on the earth's surface
- Projection - for mapping 3D spherical view to 2D plane
- Storage units (degrees, meters, etc.)
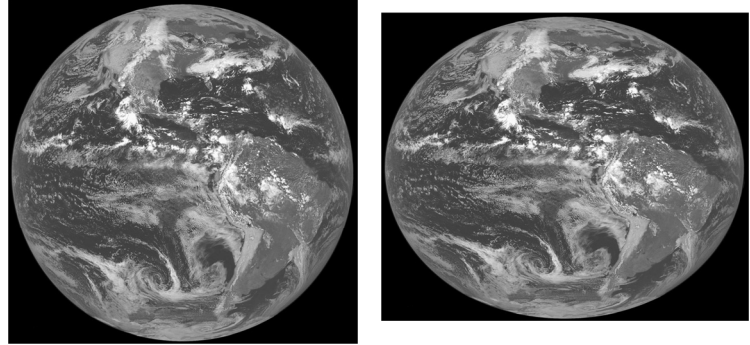- Resolution - accuracy of the measurements

Geographic Coordinate System

Latitude and longitude:

Latitude measures the angle from the horizontal. It represents north-south distance from the equator.

Longitude measures around the circle of the equatorial plane. It represents east-west distance from Prime Meridian.

## GCS - Earth is not a Perfect Sphere

Earth is Not Even a Perfect Ellipsoid

The earth has been approximated by various ellipsoids over time. Current standard is: (Maling, 1989)

Still not perfect as affected by topography which is the height of surface features.

A *geoid* is an earth model that takes into account surface height from the centre of earth. (defined by gravity measurements - see https://en.wikipedia.org/wiki/Geoid)
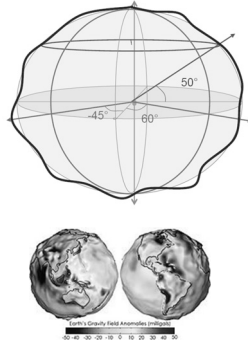
Datum

A *datum* is a mapping to minimize the difference between geoid and ellipsoid. Shifts ellipsoid relative to geoid for a particular location.

Datum components:
- ellipsoid used
- adjustment or fit (translation of center)

Note this means different datums are incompatible. Make sure you know your datum.
- Example: WGS 84 - reference coordinate system for GPS

Projections

A *projection* transforms a spherical coordinate system to a planar coordinate system. Each projection has different benefits and distortions.

Cylindrical   Conic   Azimuthal
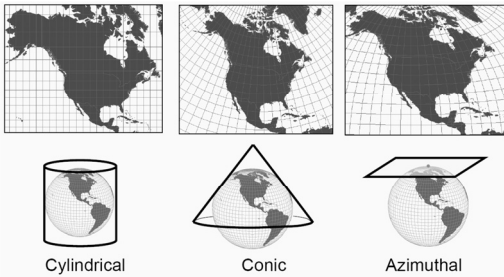
GPS Data

GPS units collect points with a datum and projection. (eg. Lat-Lon NAD 1983).

Important to record datum when perform analysis later.

Map Design Process

Determine objectives
- Know audience and purpose, use case

Decide on data and layers required
- What types of data: points, line, area, volume, temporal?

Plan map layout

Choose colors and symbols
- Use colors consistent with understanding (red/green) and real-world
- Use bold colors sparingly. Color/size use strategically for emphasis.

Create!

Coordinate Systems Question

Question: How many of the following statements are **TRUE**?

1) The earth can be modeled as a perfect spheroid.
2) Latitude measures the angle from the horizontal.
3) The zero degree for longitude is the equator.
4) A projection will cause a distortion when representing 3D as 2D.
5) A datum consists of a mapping between a geoid and an ellipsoid.

A) 0   **B)** 1   **C)** 2   **D)** 3   **E)** 4

Google Maps

There are a variety of GIS tools and software to use. We will use Google Maps, specifically Google My Maps, as it is easy to use and handles many of the details for us.

Google My Maps Link: https://www.google.com/maps/d/u/1

Representation is using vector format and layers consisting of single or groups of objects (features classes) can be easily added.

Supports importing data from files, including KML, CSV, and others as well as entering data via searching or map exploring.

KML

KML or Keyhole Markup Language uses XML to represent geographic information for visualization.
- Supported by Google and international standard in 2008.
- KML file represents features for display on maps with latitude/longitude coordinates.
- Data file is often in zipped form (KMZ files).

Example:
```
<?xml version="1.0" encoding="UTF-8"?>
<kml xmlns="http://www.opengis.net/kml/2.2">
<Document>
<Placemark>
  <name>New York City</name>
  <description>New York City</description>
  <Point>
    <coordinates>-74.006393,40.714172,0</coordinates>
  </Point>
</Placemark>
</Document>
</kml>
```
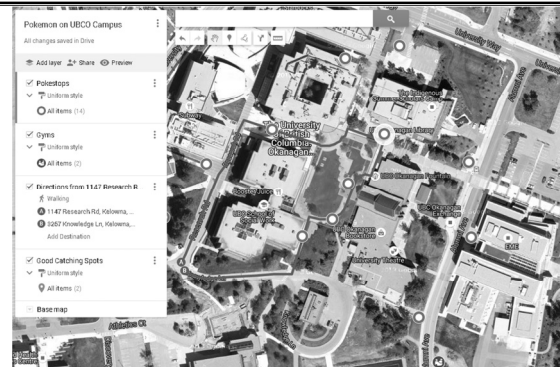
Google My Maps
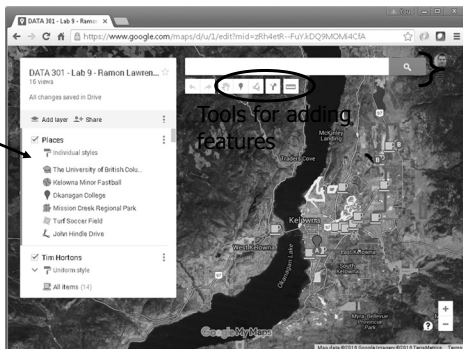
Define layers with features

Search to find features

Try it: Google My Maps - Pokémon!

Build a map that looks like this.

Note the library has a picture of the library.

Use walking path and custom icons.

Try it: Google My Maps (Create Your Own)

Build a map that has the following:
- Two layers
- Three features per layer
- At least one marker with an image
- An area
- A driving route
- Import an open data set. City of Kelowna data:
  http://www.kelowna.ca/CM/Page3936.aspx

Suggestion: Add a map with places you have been or would like to visit.

Google Maps API with Python

The Google Maps API can be used with Python to access and manipulate geographical data using a Python program.
- https://developers.google.com/maps/web-services/client-library

Services and features:
- Geocoding and reverse geocoding
- Directions (walking, driving, transit)
- Distance calculations and routes
- Elevations
- Geolocation (based on WIFI and cell towers)
- Road information and speed limits
- Times zones and places (points of interest)

Google Maps API - Getting an API Key

The first step is to get an API key that allows access to the Google services. This API key should be kept private and not shared!
- To get a key you will need a Google account.
- Securing API keys: https://support.google.com/cloud/answer/6310037

Get an API key using Google Developer Console.
- https://developers.google.com/maps/documentation/directions/get-api-key

With directions API, test with:
- https://maps.googleapis.com/maps/api/directions/json?origin=Toronto&destination=Montreal&key=yourkey

Installing Google Maps API for Python

Command:
```
pip install -U googlemaps
```

Python Google Maps API Example

```
import googlemaps
from datetime import datetime

# TODO: Replace the API key below with a valid API key.
gmaps = googlemaps.Client(key='yourkey')

# Use Geocoding API to look up latitude, longitude
address = '3333 University Way, Kelowna, BC, Canada'
geocode_result = gmaps.geocode(address)

print("Geocoding address...")
print("Address:",address,
"Coordinates:",geocode_result[0]["geometry"]["location"])
```

Python Google Maps API Example (2)

```
# Look up an address with reverse geocoding (UBC Van)
lat = 49.2683043
lon = -123.2489377
reverse_geocode_result=gmaps.reverse_geocode((lat, lon))

print("\nReverse geocoding...")
print("Coordinates: ",lat,lon,"Address:",
reverse_geocode_result[0]["formatted_address"])
```

Python Google Maps API Example (3)

```
# Request driving directions between UBCO and UBCV
directions_result = gmaps.directions(address,
        reverse_geocode_result[0]["formatted_address"],
        mode="driving", departure_time=datetime.now())
print("\nDriving directions...")
leg = directions_result[0]['legs'][0]
print("Start address:",leg['start_address'],
      "\nDestination address:",leg['end_address'])
print("Distance:",leg['distance']['text'],
      "Time:",leg['duration']['text'])

for step in leg['steps']:
    print("Step:",step['duration']['text'],
          step['html_instructions'])
```

Conclusion

Geographic Information Systems are systems designed for storing, manipulating, analyzing, and displaying spatial and geographical data.

A GIS supports:

- importing data from various sources in different formats
- organizing the data into layers or groups and integrating data from sources
- displaying the data visually as maps or 3D visualizations to help interpret the data

Understanding how GIS data is encoded using a geographical coordinate system and datums is important when interpreting and combining data from sources.

Google My Maps is an easy-to-use tool for displaying geographical information. The Google Maps API can be used with Python.

Objectives

- Provide examples where a GIS is used
- Define GIS and list some of its features/components
- Appreciate history of GIS including Canadian connection
- List and use GIS features: text, point, line, polygon
- Explain the relationship between features, coordinates, and attributes
- Provide an example on how interval and categorical data is displayed
- Define: scale, precision, resolution and perform simple calculations
- Define: feature class, layer
- Compare and contrast raster versus vector representations
- Define and use latitude and longitude
- Explain the challenge in modeling a point on the earth's surface given that it is not a perfect sphere and has topography
- Explain role and connection between a geoid, spheroid, datum

Objectives (2)

- Explain the purpose of a projection and understand different projections have different benefits and distortions
- Apply a map design process to produce visually appealing maps
- Define and use KML
- Create a map with Google My Maps with various features
- Write a program to access the Google Maps API using Python