Chapter 1.
GIS Data Concepts

Gervais W Tabopda, PhD

School of City and Regional Planning

Georgia Tech

Course Objectives

- Understanding how **GIS data represent real-world features**
- Knowing fundamental differences between the **raster** and **vector** data models
- Using and discussing **map scales** and **source scales**
- Recognizing local and Internet **sources** of GIS data
- Cultivating **awareness** of the types of shortcomings associated with GIS data
- Learning to correctly **cite** data sets on a map or in a report
- Learning the basics of the **ArcGIS Pro interface** and viewing **map data**

GIS as a Database

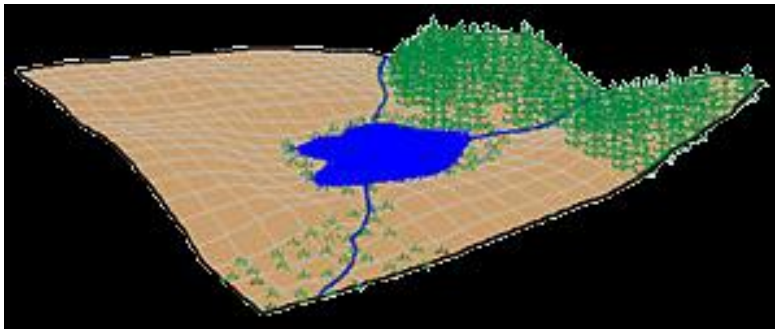
- **GIS** does not save information as maps or pictures
- **GIS** holds information in a database, but the information has a **spatial context**
- To go beyond making pretty pictures, you need to know:
  - 1) What it is
  - 2) **Where it is**
  - 3) How it relates to other features

Organization of GIS Data

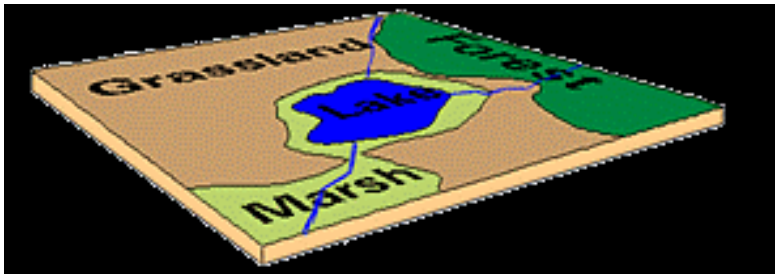
- **Objects and Fields**
  - Objects are spatial entities like **lines**, **points**, and **polygons**
  - Fields are **attributes** describing these entities

- **Vector vs. raster data**
  - Vector Data uses **geometry** to specify the size and shape of the object
  - Raster Data uses **"pixels"** or **grids** arranged in a two-dimensional array that specifies the location of each pixel

Vector and Raster Representation

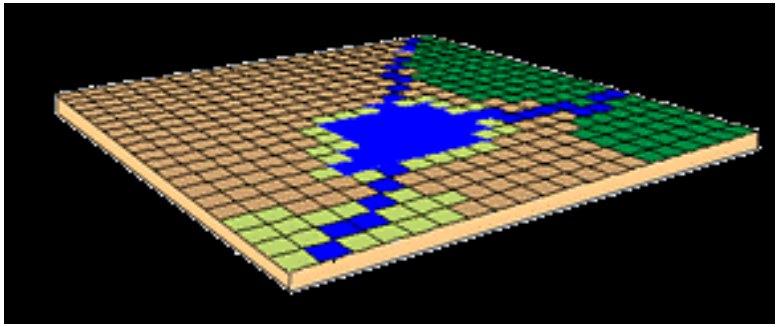


In this **example**, our landscape consists of rolling hills, grassland, lake, rivers, forest and marshland.



VECTOR DATA

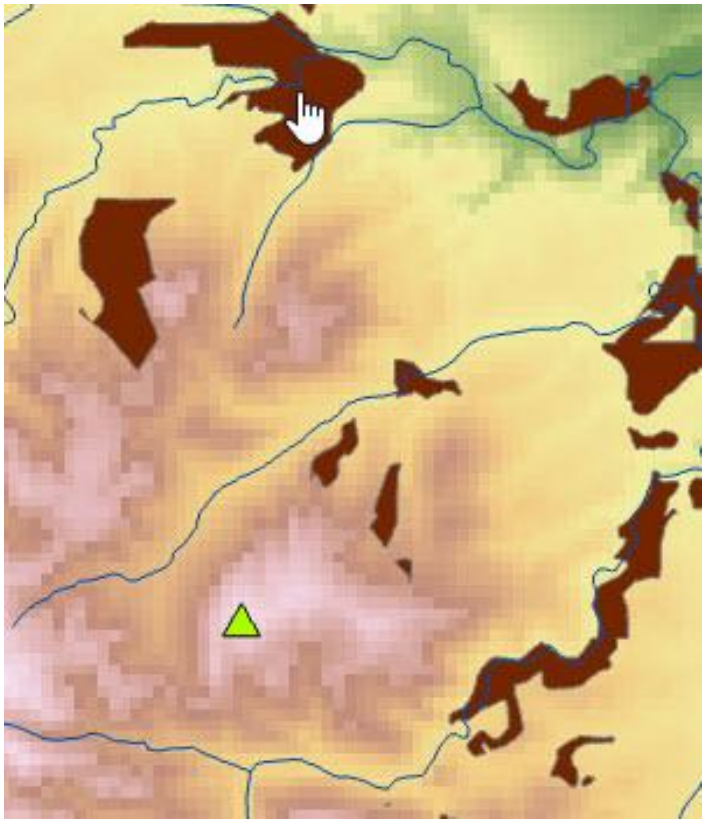
In the **Vector Data Model**, features on the earth are represented as **points**, **lines/routes,** and **polygons/regions**



RASTER DATA

In the **Raster Data Model**, land cover is represented as **single square cells** containing a value corresponding to its land cover type.

Types of map data



Source: USGS

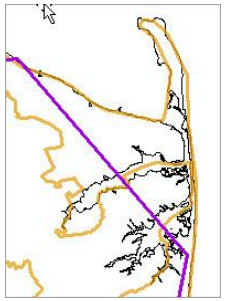
Discrete objects exist in a defined location/space

- Lookout tower (point).
- Stream (line).
- Snail habitat (polygon).

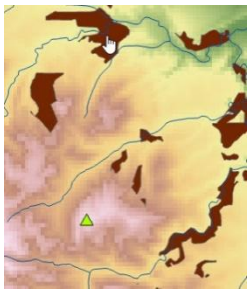
Continuous data exist everywhere

- Elevation (raster).

GIS data models



Source: Esri



Source: USGS

Source: Google Earth and Tele Atlas

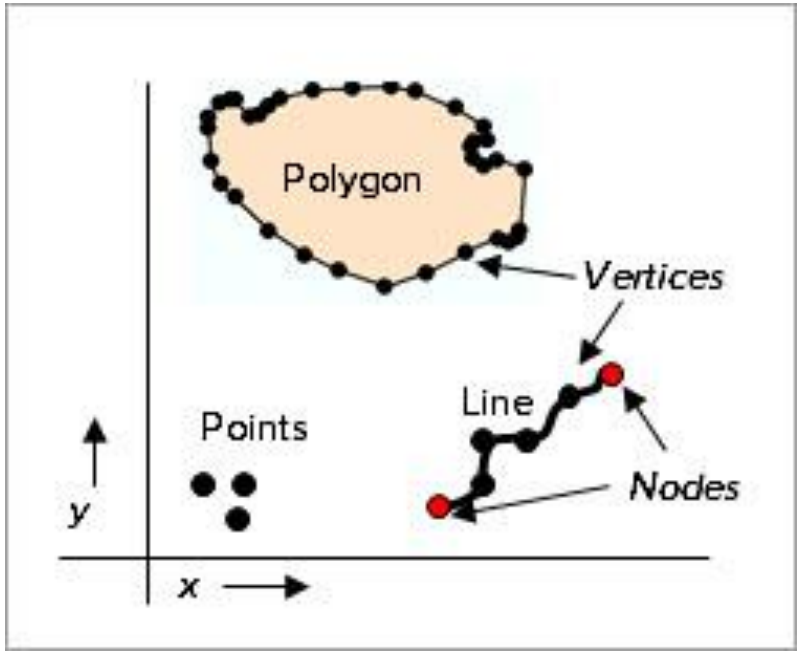
The vector model is best for <u>discrete data</u>

- state boundaries.
- streams.
- snail habitat.

The raster model is used for imagery and is best for <u>continuous data</u>

- elevation.
- aerial photography.

Vector data model



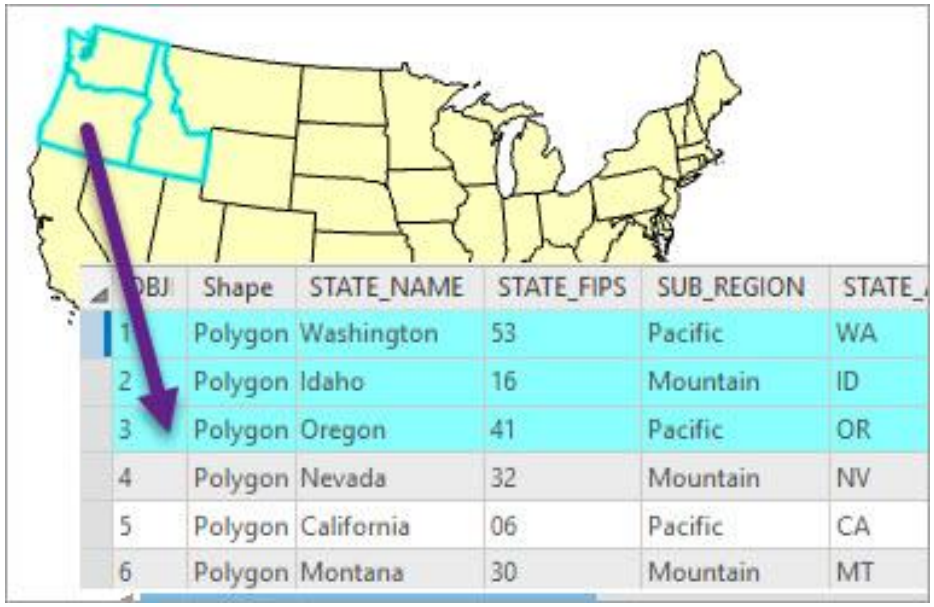
The vector data model is best for discrete data

Features are stored map objects

- Points, lines, polygons.

A **feature class** is a collection of similar features stored together, like states or rivers

Attribute tables



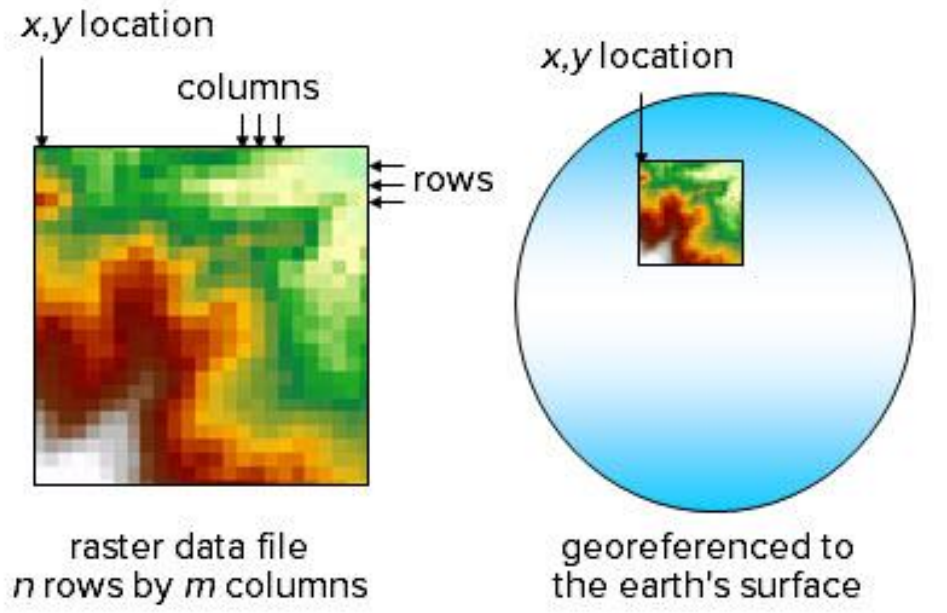
OBJ	Shape	STATE_NAME	STATE_FIPS	SUB_REGION	STATE_
1	Polygon	Washington	53	Pacific	WA
2	Polygon	Idaho	16	Mountain	ID
3	Polygon	Oregon	41	Pacific	OR
4	Polygon	Nevada	32	Mountain	NV
5	Polygon	California	06	Pacific	CA
6	Polygon	Montana	30	Mountain	MT

Source: Esri

Features are linked to tables containing information about the **spatial objects**

The map object and the table data are connected by a **unique integer**

Raster data model

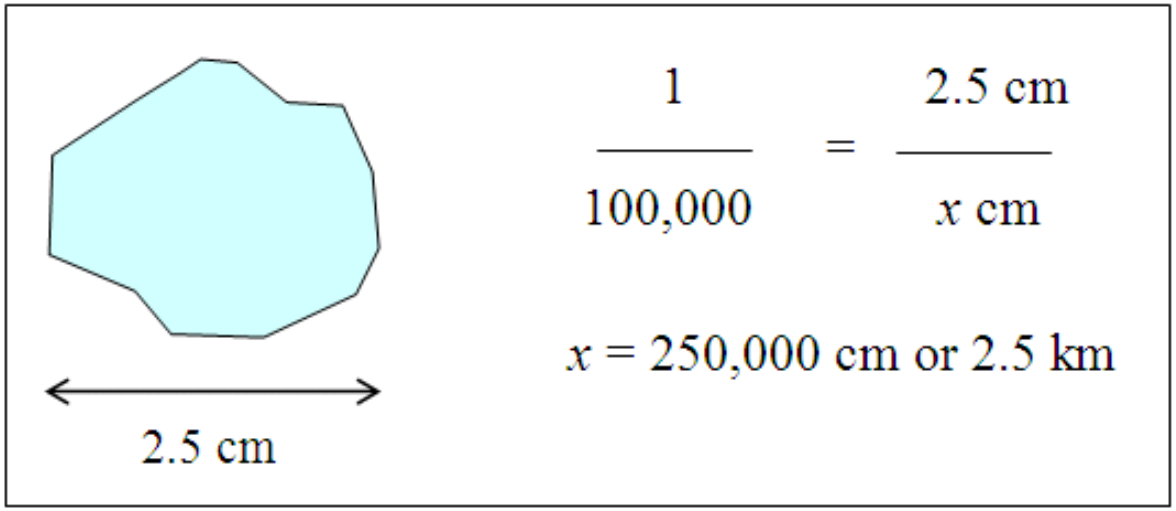


Source: USGS

The raster model breaks map areas into small squares known as **cells** or **pixels**

A single numeric value is stored in each cell, such as elevation

Map scale

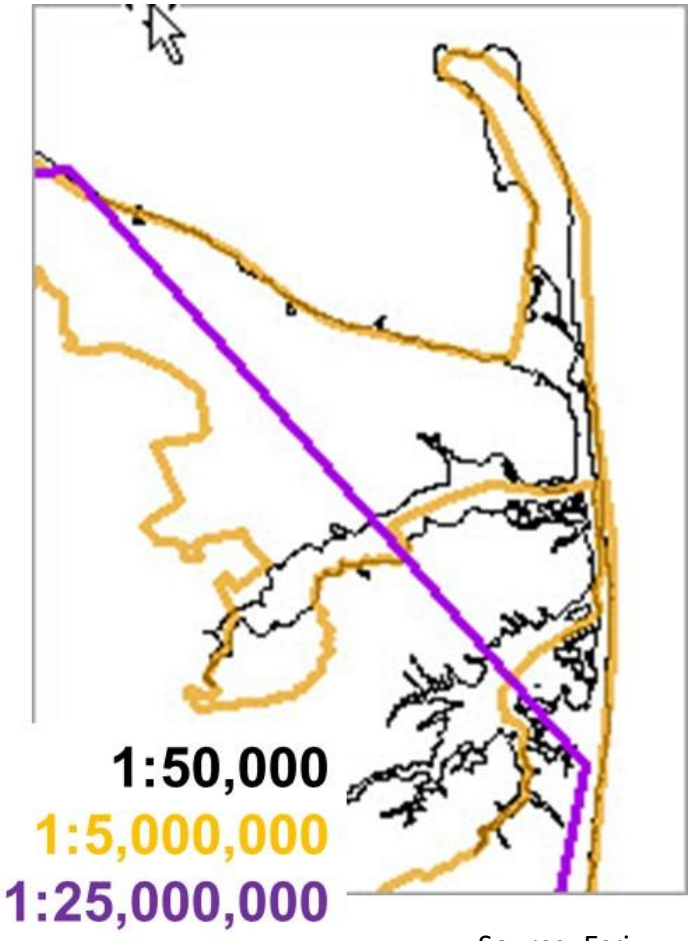


1 *cm* on map = 100,000 *cm* on ground

Map scale is the ratio of distance on the map to distance on the ground

It is dimensionless and can be expressed in any unit: **cm** or **inches** or **mm**…

Generalization



Source: Esri

Generalization is used to simplify map features for clear display

- Small-scale maps present less detailed versions of objects.
- Large-scale maps present more detailed versions.

Source scale is the scale at which data are captured

- **It impacts the detail and accuracy of the data set.**

Data Quality

Data sets are rarely perfect

No absolute quality standard exists

Quality is defined as the fitness of a data set for a particular purpose

The same data set may be unsuitable for one use but adequate for another

A user has **ethical** and **legal** responsibility to determine whether a data set is sufficient for its intended use

Geometric accuracy

Is it where it says it is?

- Depends on the level of error in the original source.
- Additional errors may be incurred or propagated during processing.
- Assessed by comparing the data to another data set with known **high accuracy**.

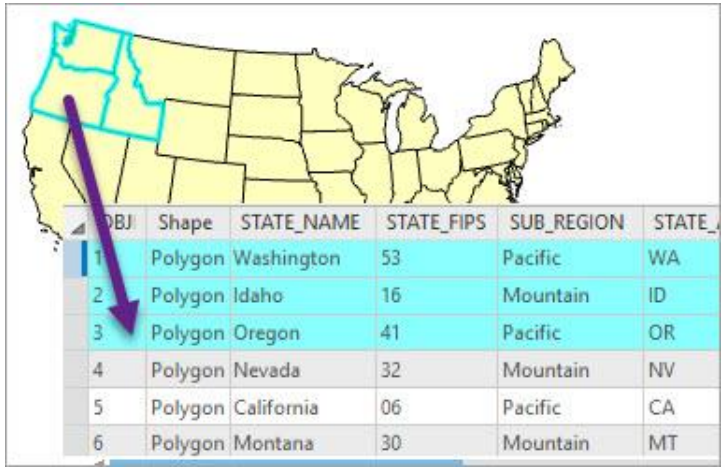
Which road is “correct”? What errors might occur in the locations of both?

Source: Google Earth and Tele Atlas

Thematic accuracy

How accurate are the attributes?

- One might ask questions about, for example, tree crown density.
  - How is it measured?
  - How accurate are the measurements?
  - What are the possible sources of error?



OBJ	Shape	STATE_NAME	STATE_FIPS	SUB_REGION	STATE_
1	Polygon	Washington	53	Pacific	WA
2	Polygon	Idaho	16	Mountain	ID
3	Polygon	Oregon	41	Pacific	OR
4	Polygon	Nevada	32	Mountain	NV
5	Polygon	California	06	Pacific	CA
6	Polygon	Montana	30	Mountain	MT

Source: Esri

Source: GoogleEarth and TeleAtlas

Resolution

Resolution is the sampling interval of measurements during data collection

- **Spatial sampling resolution**.
  - Distance between GPS points along a road.
  - Size of pixel for elevation or satellite image raster.
- **Thematic data resolution**.
  - How fine was the measuring scale?
  - Were the data classified after measurement?
- **Temporal data resolution**.
  - How frequently were data sampled?
  - Daily, monthly, every decade?

Source: Google Earth and Tele Atlas

Precision

Precision has two meanings in science

It is NOT the same as accuracy!

Meaning One: the number of significant digits in a measurement

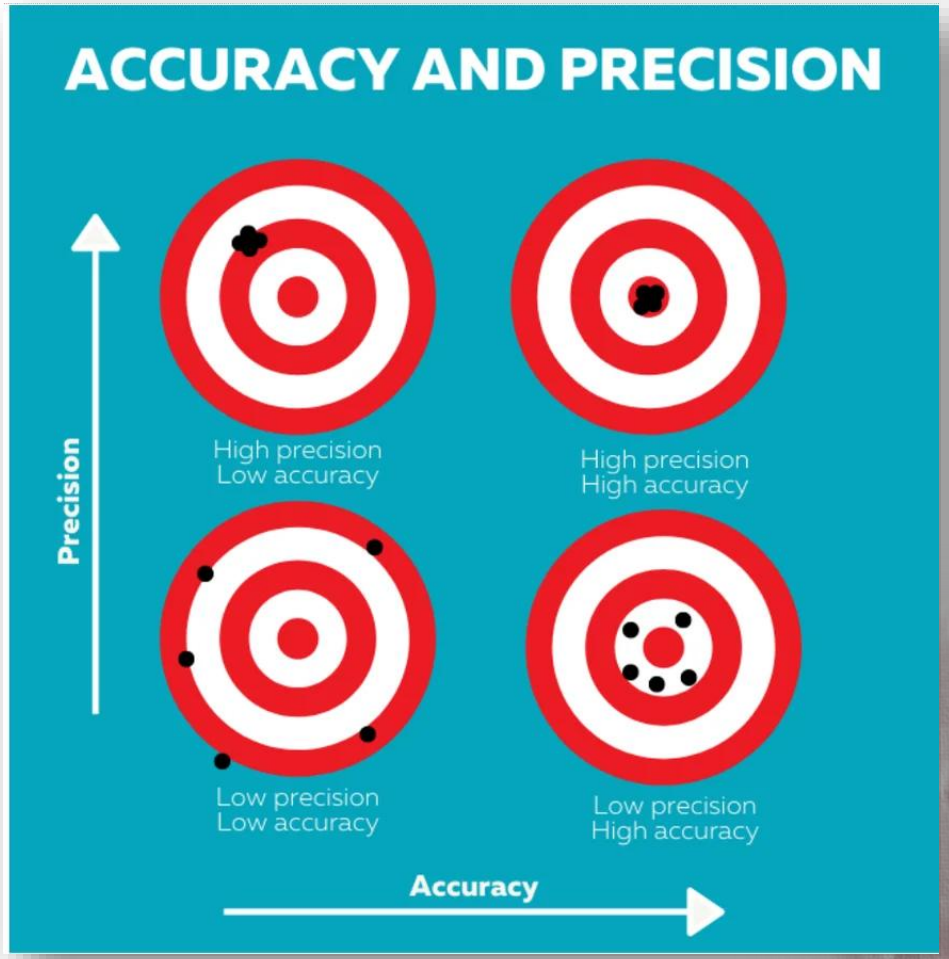
- Example: a GPS unit reports locations to the nearest meter. The precision is about 1 meter.

Meaning Two: The statistical variability of a repeated measurement

- Example: 20 GPS measurements at the same spot have a standard deviation of 5 meters. The precision is about ±10 meters.

Source: Google Earth and TeleAtlas

Accuracy vs Precision

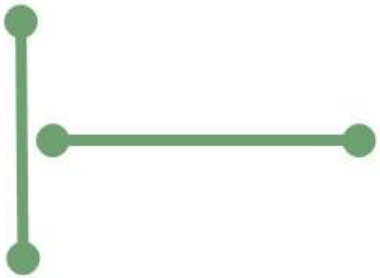


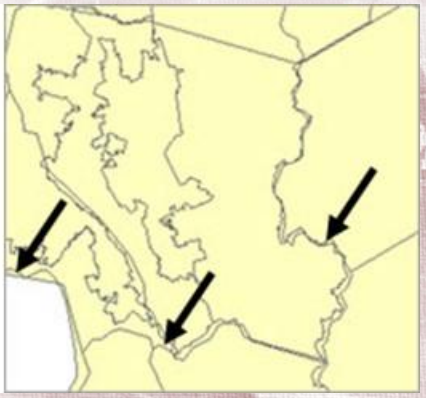
Source: [Accuracy, Precision, & Resolution - What Do They Mean for IAQ Sensors?]

Logical consistency

Logical consistency assesses how well a data set represents real-world relationships.

- Are common boundaries identical?
- Do roads actually connect?
- Do voting district and county boundaries align?





Source: Esri

Citing GIS data

Formats may vary for different agencies or companies

A good citation enables someone to find the source should they wish to obtain a copy

Citing data is a professional and ethical responsibility

Data set name (Year published) [source type]. Producer name, producer contact information. Resource *URL: [Date accessed].*

Citation examples

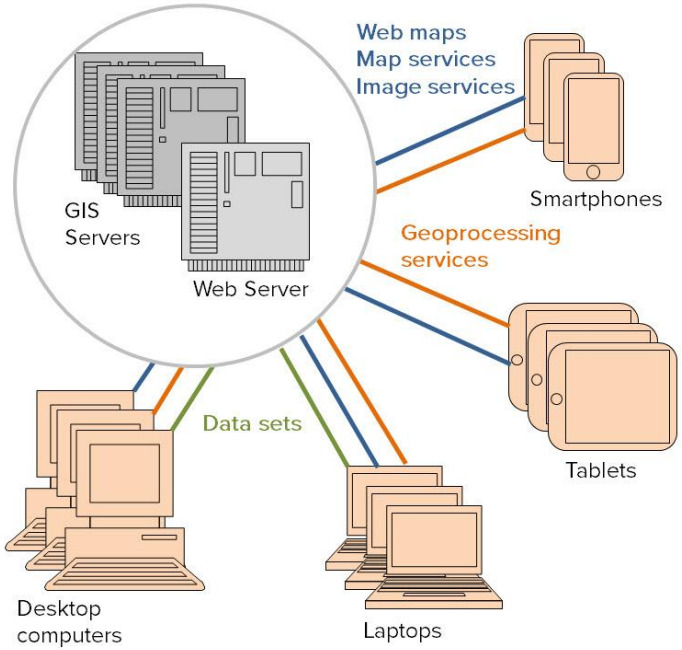
- Black Hills RIS Vegetation Database (2008) [downloaded file]. Black Hills National Forest, Custer, SD. URL: http://www.fs.usda.gov/main/blackhills/landmanagement/gis [August, 2010].

- Esri™ Data and Maps (2012) [DVD]. Esri™, Inc., Redlands, CA.

- National Hydrography Dataset (2015) [downloaded file]. United States Geological Survey on the National Map Viewer. URL: http://viewer.nationalmap.gov/viewer/ [July 23, 2015].

- USA Topo Maps (2009) [map service]. Esri™ on ArcGIS Online. URL: http://server.arcgisonline.com/arcgis/services/USA_Topo_Maps/MapServer [January 1, 2012].

GIS Services



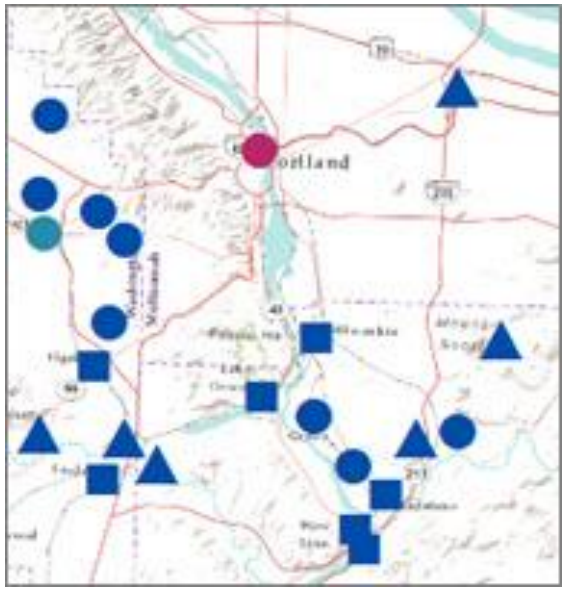
GIS servers provide data over the Internet

Accessible to many types of **platforms**

Clients only consume what is needed

Possible to do **real-time updates**

Web maps

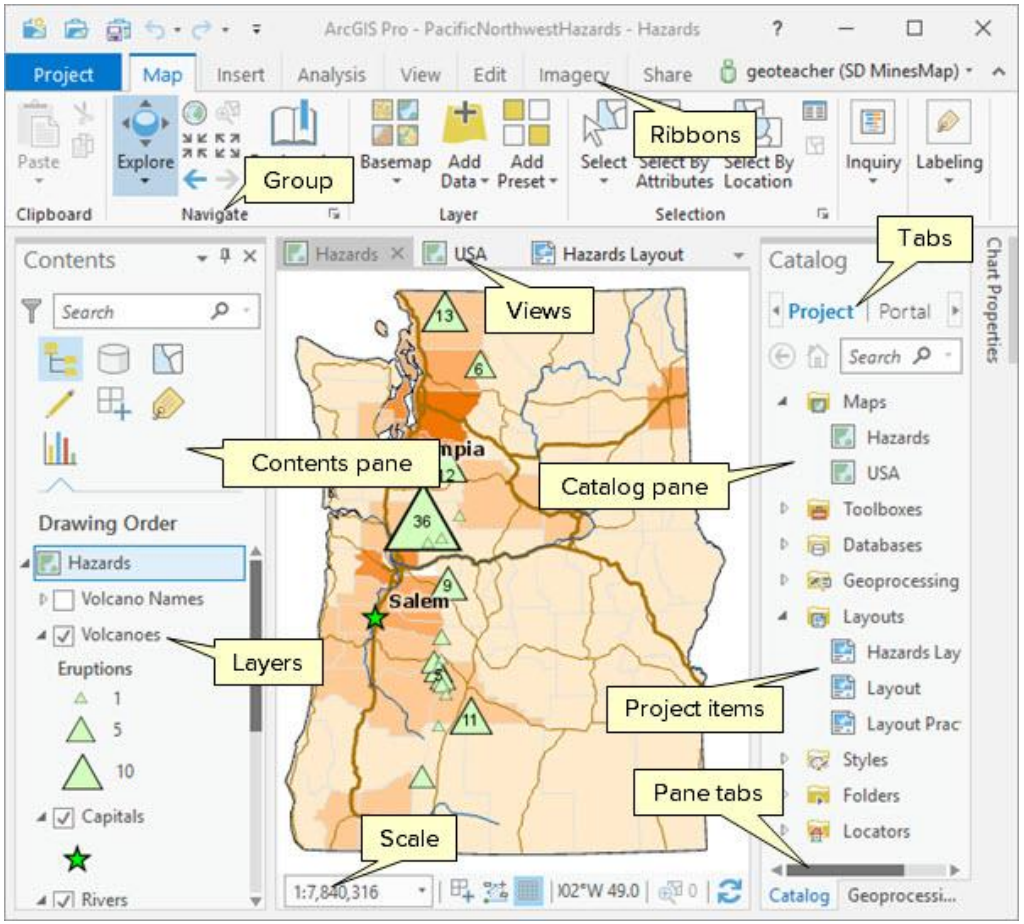


Source: Esri

Web maps are maps that only use GIS services

They can be shared from platform to platform

Pro GUI

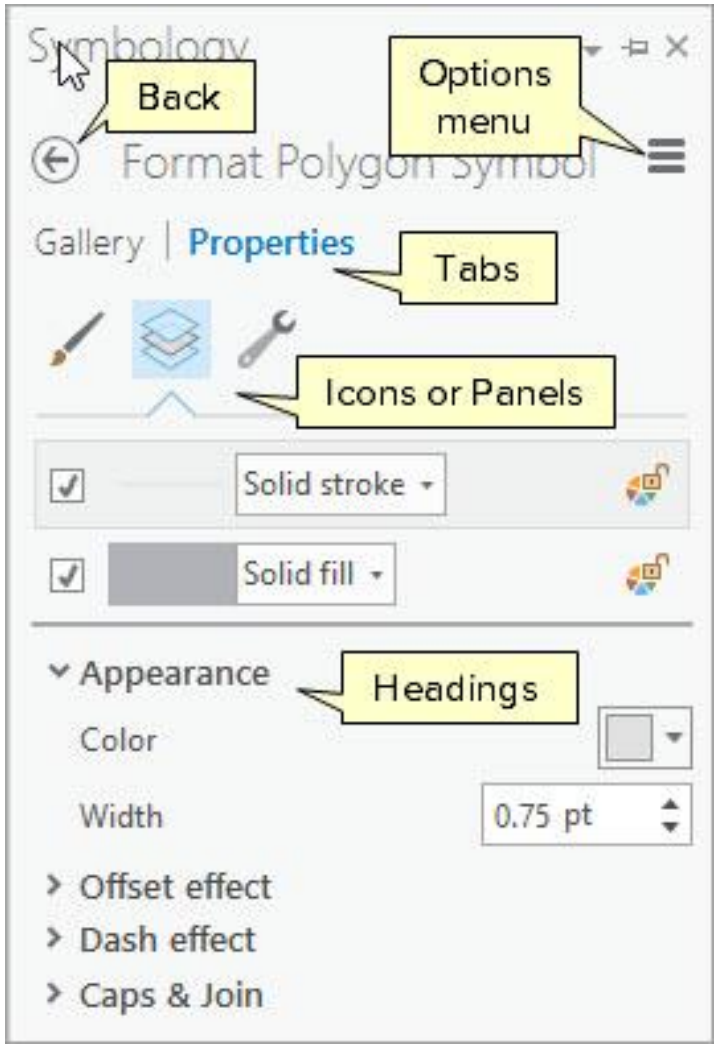


Source: Esri

Projects contain

- Maps.
- Tools.
- Databases.
- Layouts.
- Styles.
- Folders.
- Models.
- Tasks.
- Locators.

Panes



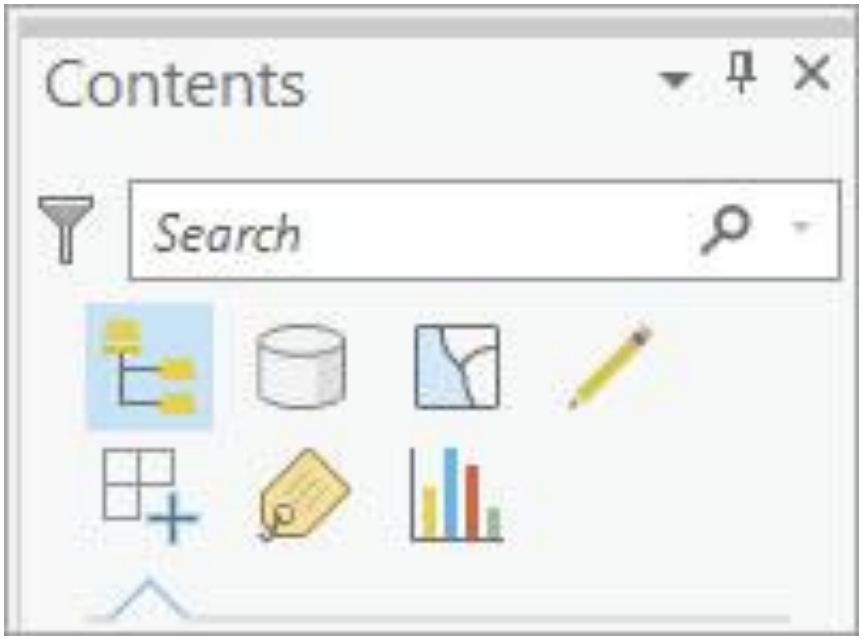
Panes contain **commands** and **settings**

Movable and stackable

It may have tabs and panels with more functions

Source: Esri

Contents pane panels



Source: Esri

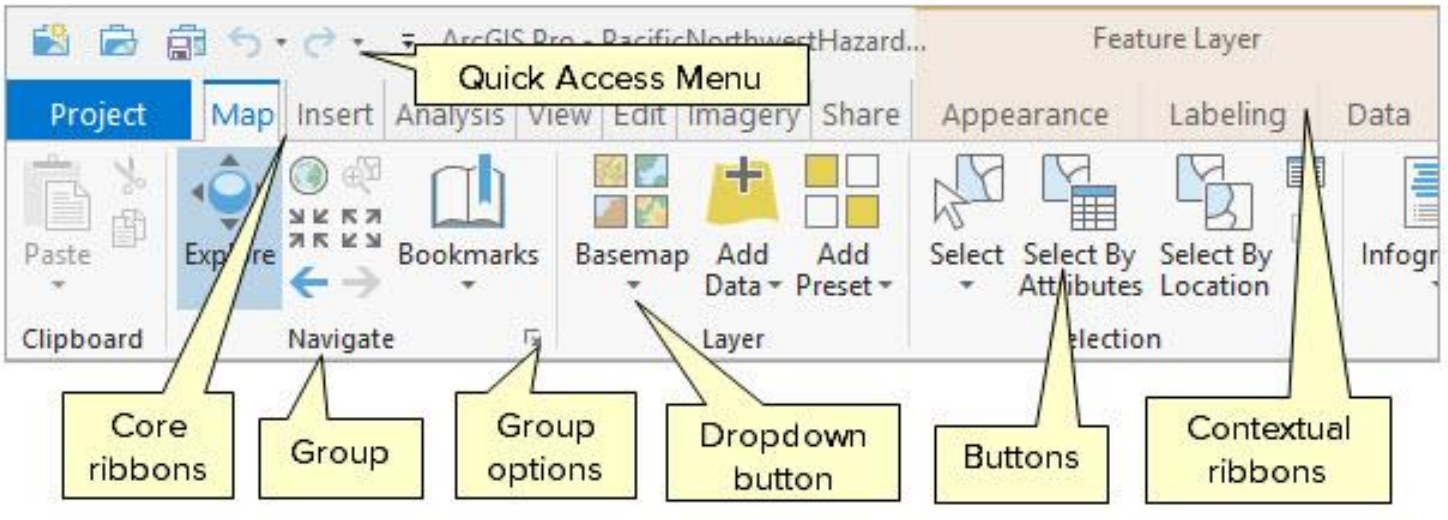
Panels (or modes) of the Contents pane

- List by Drawing Order.
- List by Source.
- List by Selection.
- List by Editing.
- List by Snapping.
- List by Labeling.
- List by Charts.

Referring to ribbon items

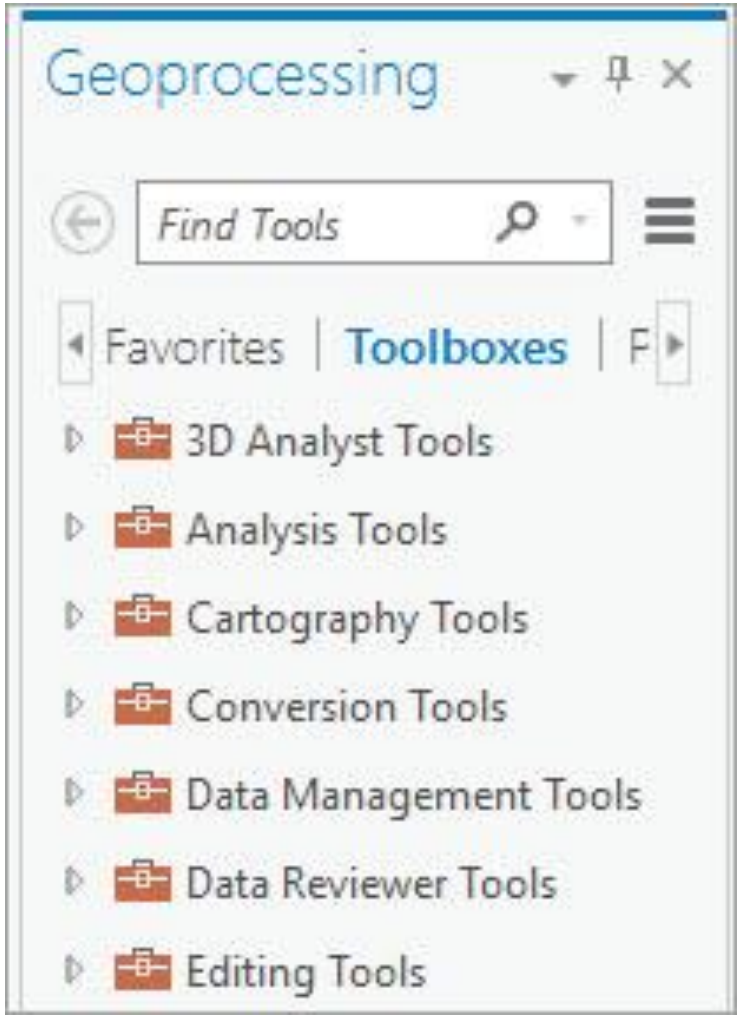
Each **ribbon** contains **buttons** and **functions** organized into logical **groups**

Some ribbons are always present; others appear only when appropriate



Source: Esri

Geoprocessing pane

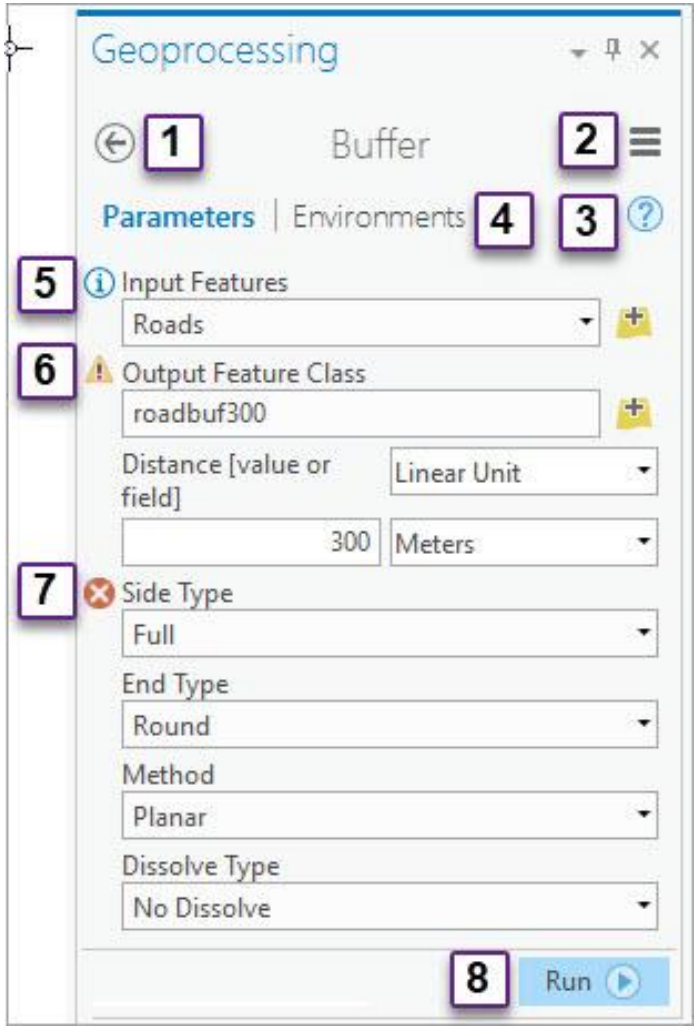


The **geoprocessing pane** contains **toolboxes** and **tools** for added functions

Extends capabilities far beyond the GUI

Source: Esri

Running a tool

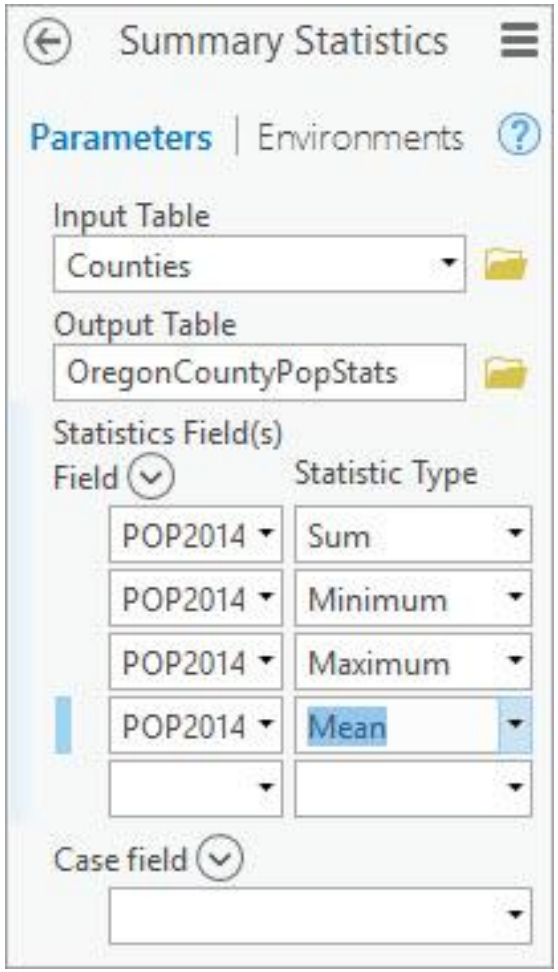


Icons, tabs, and menus available on tools include:

1. Go back to the main pane
2. Options menu
3. View Help
4. Environments tab
5. Parameter Help
6. Warning
7. Error
8. Run

Source: Esri

Summary statistics tool



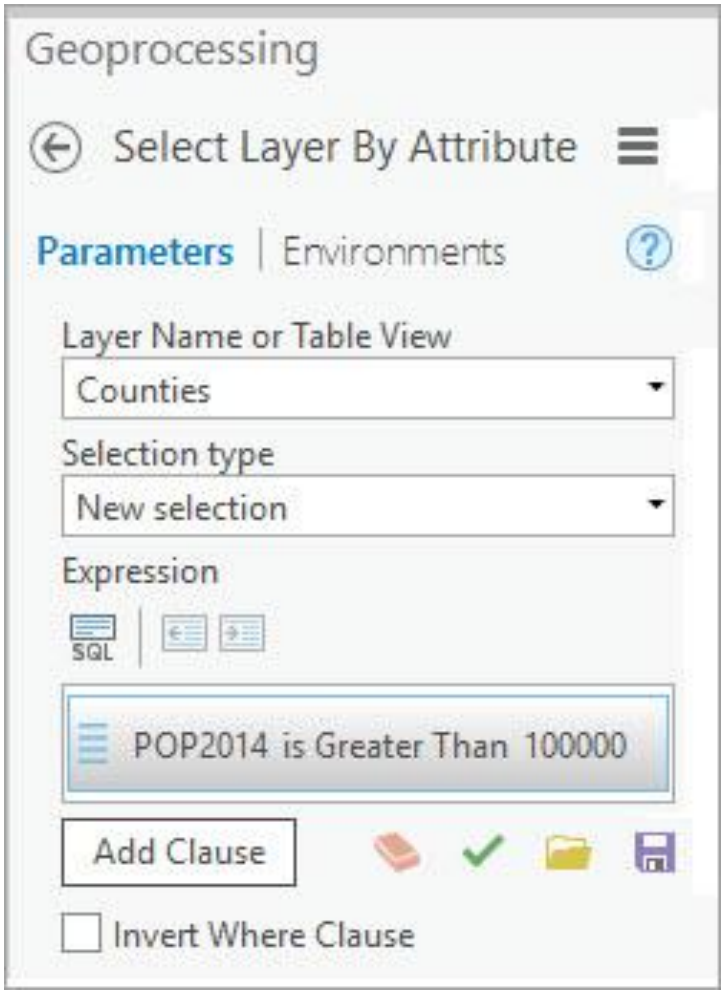
Source: Esri

Calculates statistics from table attributes

The user chooses **fields** and **statistics types**

Creates a new table as its output

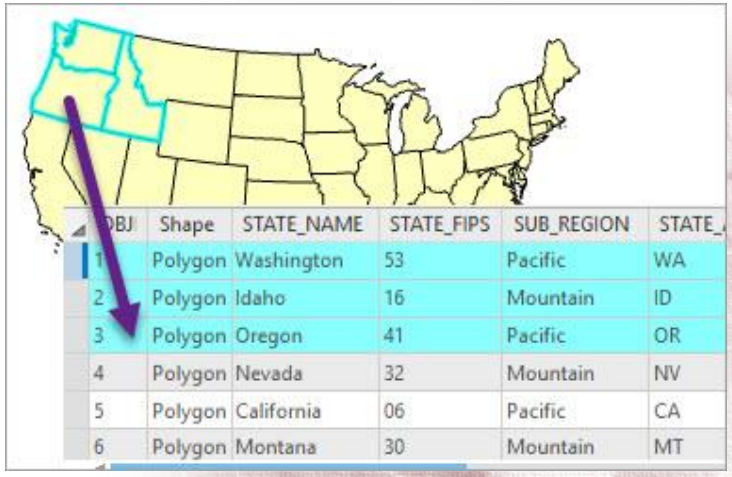
Select Layer By Attribute tool



Source: Esri

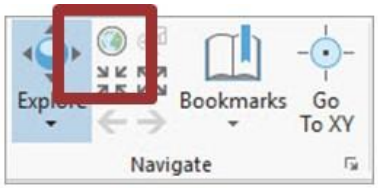
Selects features based on values in the table

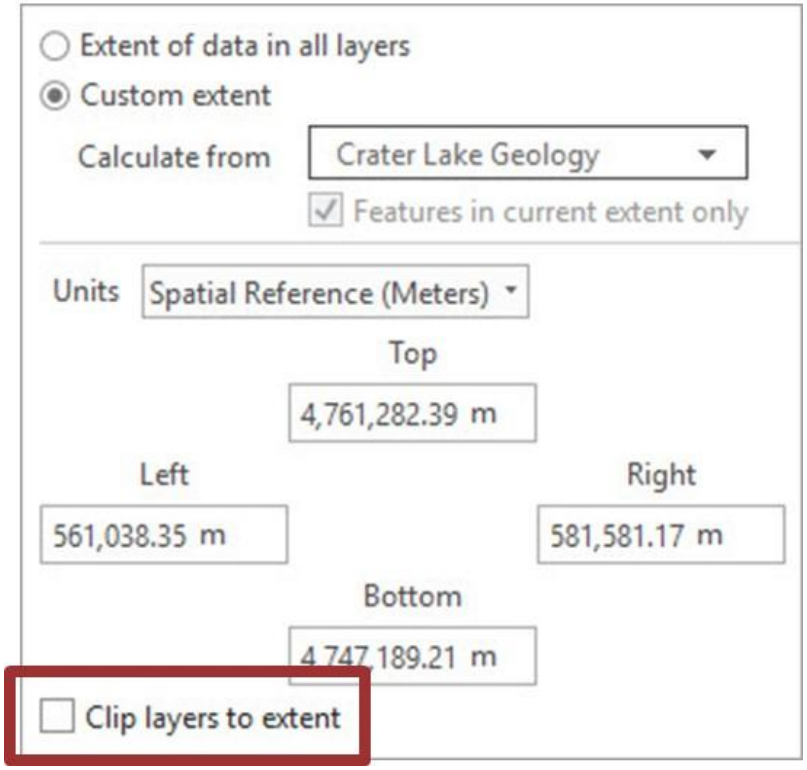
Highlights feature in the table and on the map



Source: Esri

Setting a custom full map extent





Source: Esri

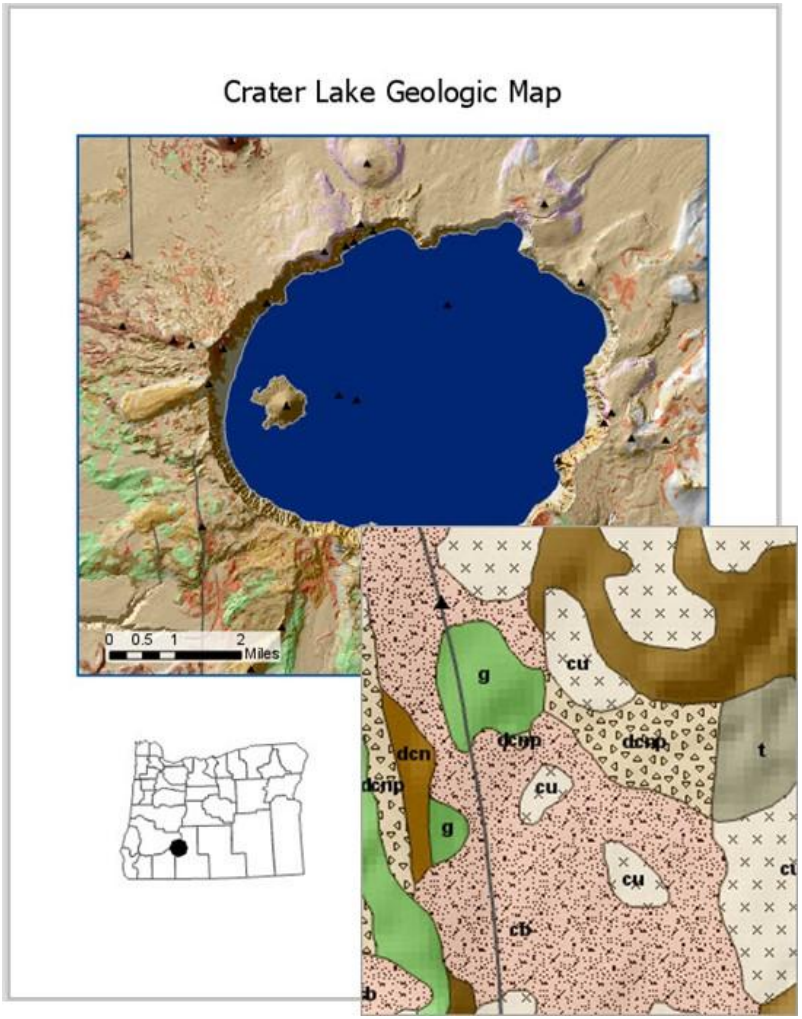
The **Full Extent button** goes to the extent set in the map properties

Defaults to the world

You may set it to a layer or the current display extent

You can also **turn off** the display outside the extent (**clip**)

Scale ranges



Source: Esri, USGS

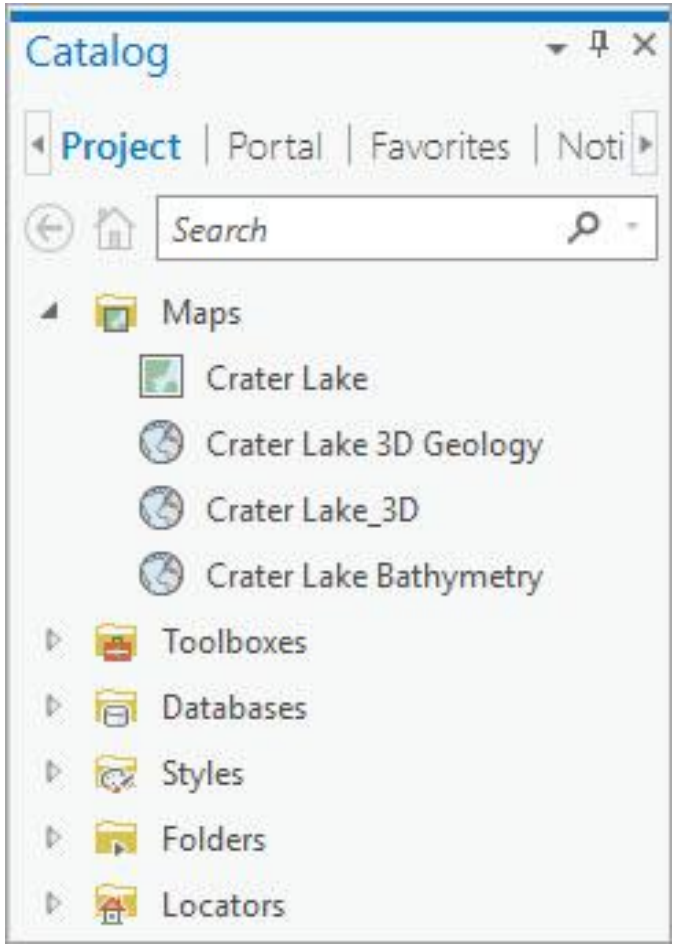
Scale ranges are properties of a layer

They are set to show a layer only at the appropriate scales

Detailed symbols and labels for large-scale

Generalized symbols and labels for small scales

Catalog pane



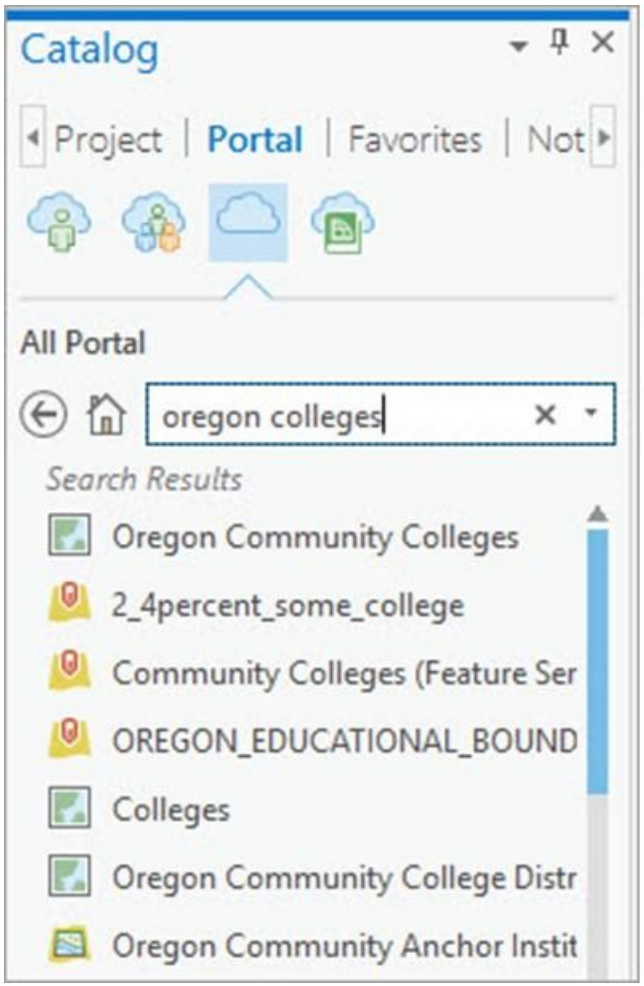
Source: Esri

The **Catalog pane** accesses many types of items in the project

It has four tabs

- **Project**, for local resources.
- **Portal**, for online resources.
- **Favorites**, to store locations.
- **Notifications**, to receive info about topics such as software updates.

Portal resources



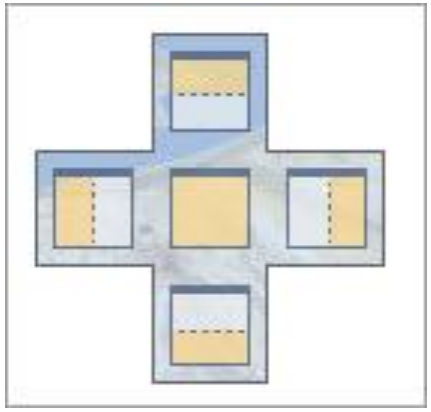
The **Portal tab** accesses ArcGIS Online content or data from other servers

There are four main types of services

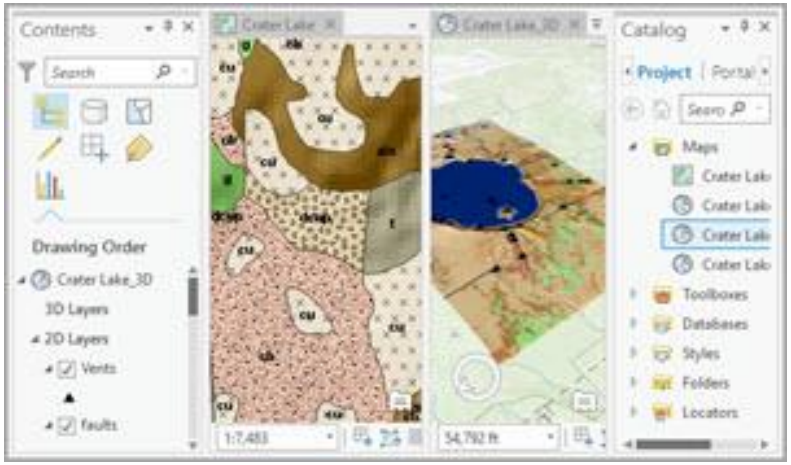
- web maps.
- feature layers.
- imagery layer.
- map image layer.

Source: Esri

Managing views and panes



Source: Esri



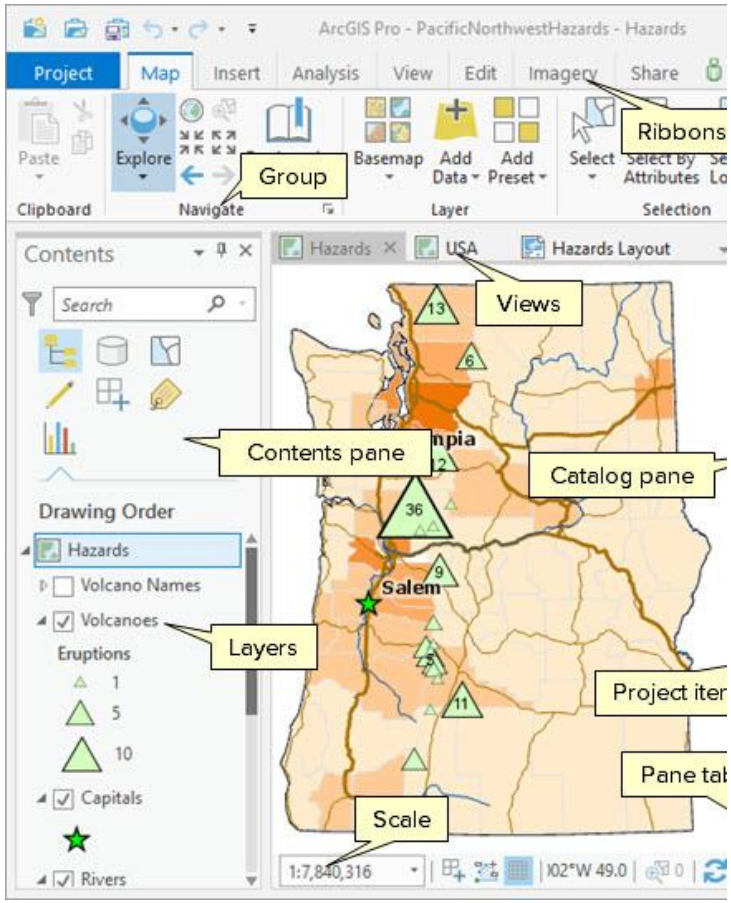
Source: Esri

A **docking icon** is used to manage the placement of panes and views in the GUI

They can be positioned as

- Adjacent.
- Stacked.
- Floating.

Layers



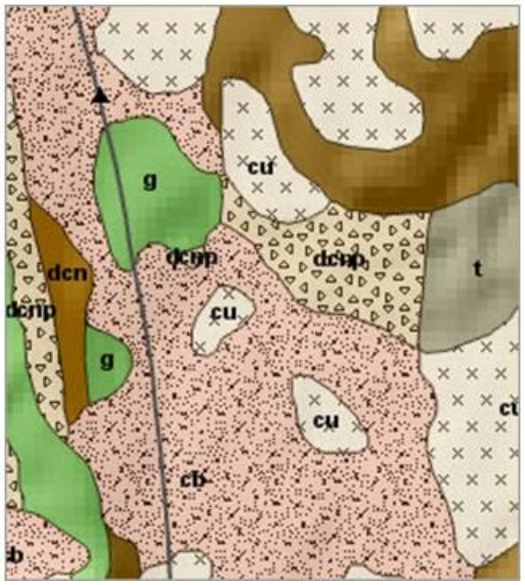
Source: Esri

Source data become layers when added to a map

Layers store properties and settings for display

Each layer's symbols and settings can be customized to appear and behave in the desired fashion.

Group layers



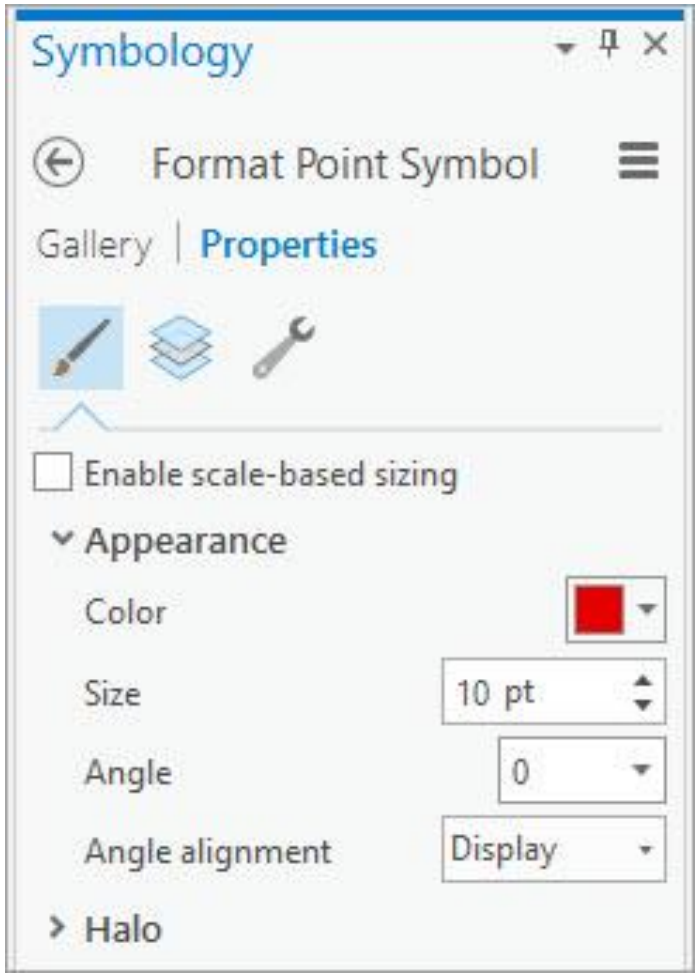
Source: USGS

A **group layer** is composed of multiple other layers

The layers may be turned on/off together or separately

It can be saved as layer files to quickly transfer display settings to new layers

Setting symbols

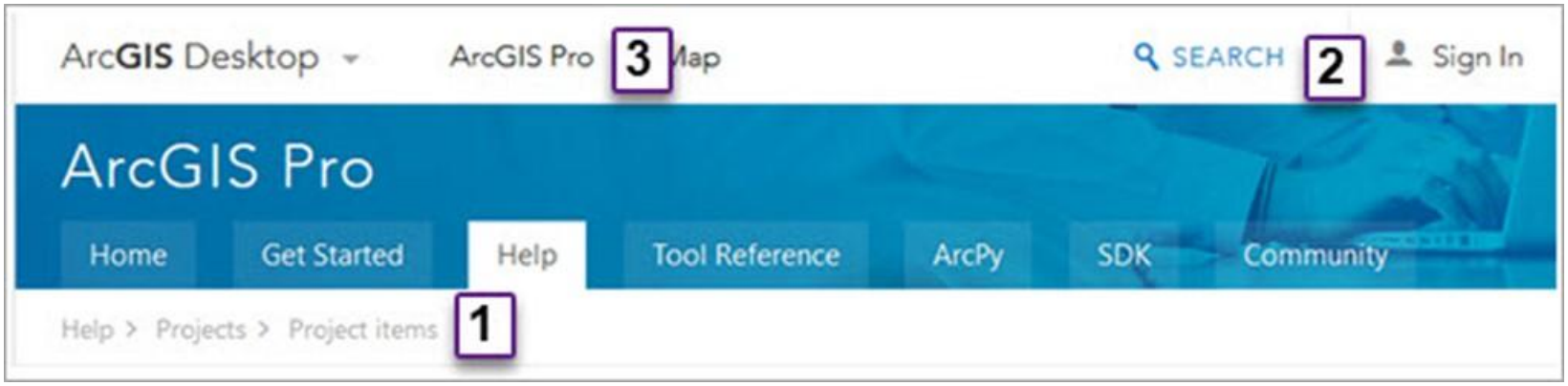


The **Symbology pane** is used to pick and modify symbols for layers

- The **Gallery tab** is used to choose a symbol from the available styles.
- The **Properties tab** is used to manipulate the settings of the chosen symbol.

Source: Esri

Using the Help



Source: Esri

Help is accessed through a web browser

Tabs and links on the Help bar include:

1. Breadcrumb trail to show the outline of topics
2. Search box to find a topic
3. Tabs to different sections of the Help

Course Objectives

- Understanding how **GIS data represent real-world features**
- Knowing fundamental differences between the **raster** and **vector** data models
- Using and discussing **map scales** and **source scales**
- Recognizing local and Internet **sources** of GIS data
- Cultivating **awareness** of the types of shortcomings associated with GIS data
- Learning to correctly **cite** data sets on a map or in a report
- Learning the basics of the **ArcGIS Pro interface** and viewing map data

Chapter 1 Lab exercise

- Exploring an ArcGIS Pro **Project**
- Navigating **2D** and **3D** Map
- Using the **Help**
- Managing **Windows**
- Exploring Project **Resources**
- Working with **local** GIS Data
- Using **ArcGIS Online** Content
- Using **Geoprocessing**