Geographic Information Systems

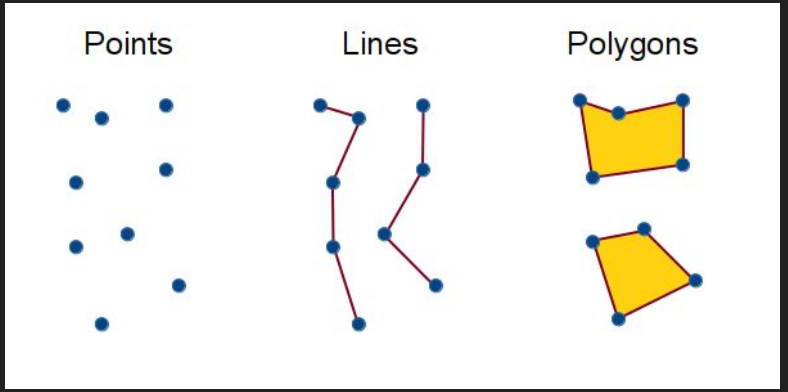
ESM 263 - Winter 2023

Vector Data

Why Vector Data?

- Recall: features vs fields
  - features: discrete entities with specific locations
  - fields: continuous functions of (x, y)
- “Vector” is GIS-speak for feature representations
  - dimensionality: point, line, area
  - topology: preserve/ignore connectivity
  - simple vs. composite

Simple Features



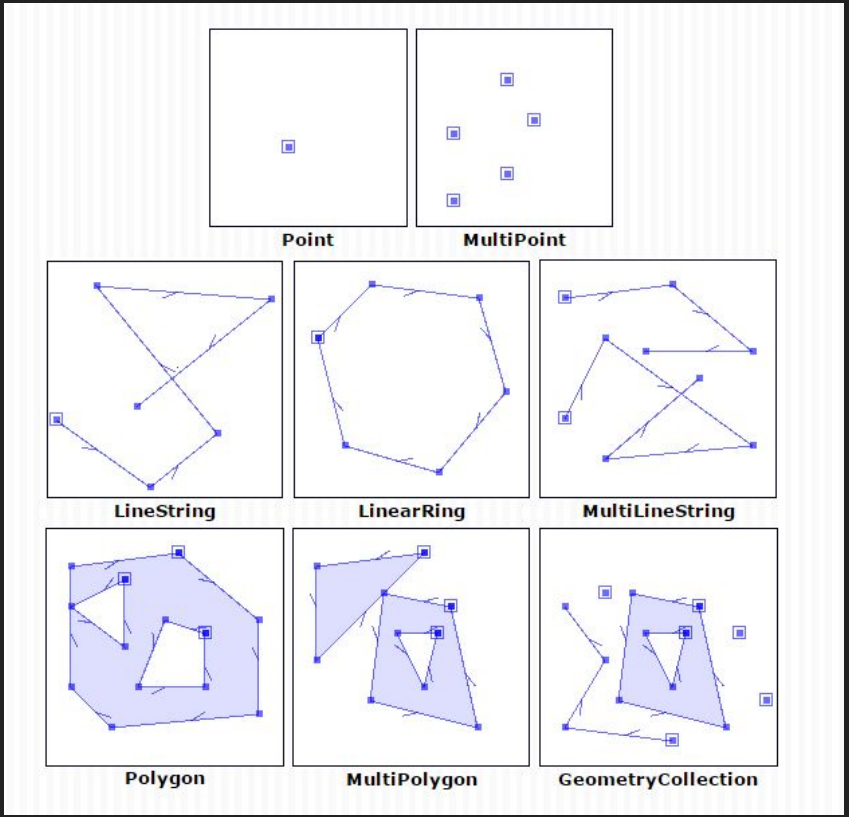
Dimensionality

- 0: point
- 1: line
- 2: area

Composition

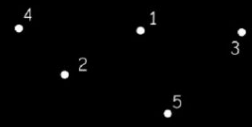
- line: sequence of points:
  - implicitly connected
- area: sequence of lines
  - boundary
  - implicit or explicit closure

Simple Features



Georelational Data Model

Points

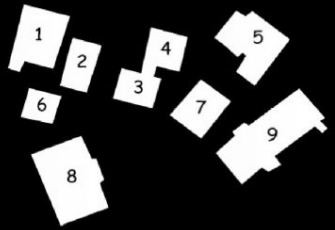


ID	Tower Name	Height	Format
1	WKRP	101.0	Pop
2	WYOU	55.5	Oldies
3	TPT	486.0	Public TV
4	WQXR	99.5	Classical
5	BBC	212.1	News

Lines

ID	Name
1	Tuckaseegee River
2	Pigeon Branch
3	Poplar Run
4	Shope Fork
5	Mel's Brook
6	Merdesansrame Creek
7	Longue Arm
8	Arroyo Grande

Polygons



ID	Building Name	Floors	Roof Type
1	Hodson Hall	6.0	flat, sealed tar
2	Borlaug Hall	5.5	pitched 9/12, tile
3	Guilford Technology Bldg.	4.0	flat, gasket
4	Shop Annex	2.5	flat, sealed tar
5	Animal Sciences Bldg.	1.0	pitched 12/12, tile
6	Administration Bldg.	14.0	pitched 6/12, metal
7	Climate Sciences Center	6.0	flat, sealed tar
8	Grantham Tower	1.0	pitched 9/12, tile
9	Biological Sciences Bldg.	9.0	pitched 12/12, tile

- Separation of geometry and attributes
  - related by feature ID
- May or may not represent topology (connectivity)
  - explicit topology: coverage (obsolete, but still used)
  - no topology: shapefile

Non-Topological Vector Data

- Lists of simple features
  - no explicit connectivity: features that share geometry, duplicate the geometry
- Advantages
  - easier to draw/display: don’t have to look up arcs
  - simpler file formats
  - easier to extract subsets of features
- Disadvantages
  - can’t tell if duplicate geometry is shared geometry
  - editing features with shared geometry can introduce inconsistencies:
    e.g. boundary between counties

Shapefile

- foo.shp
  - Geometry:
  - feature ID: coordinate list
- foo.shx
  - geometry index
  - feature ID: offset in bytes: from beginning of foo.shp
- foo.dbf
  - Attributes:
  - feature ID: attributes…
- foo.prj
  - coordinate system: geographic/projected
- foo.xml
  - Metadata
- foo.{anything else}:
  - (probably ESRI-specific)

Databases

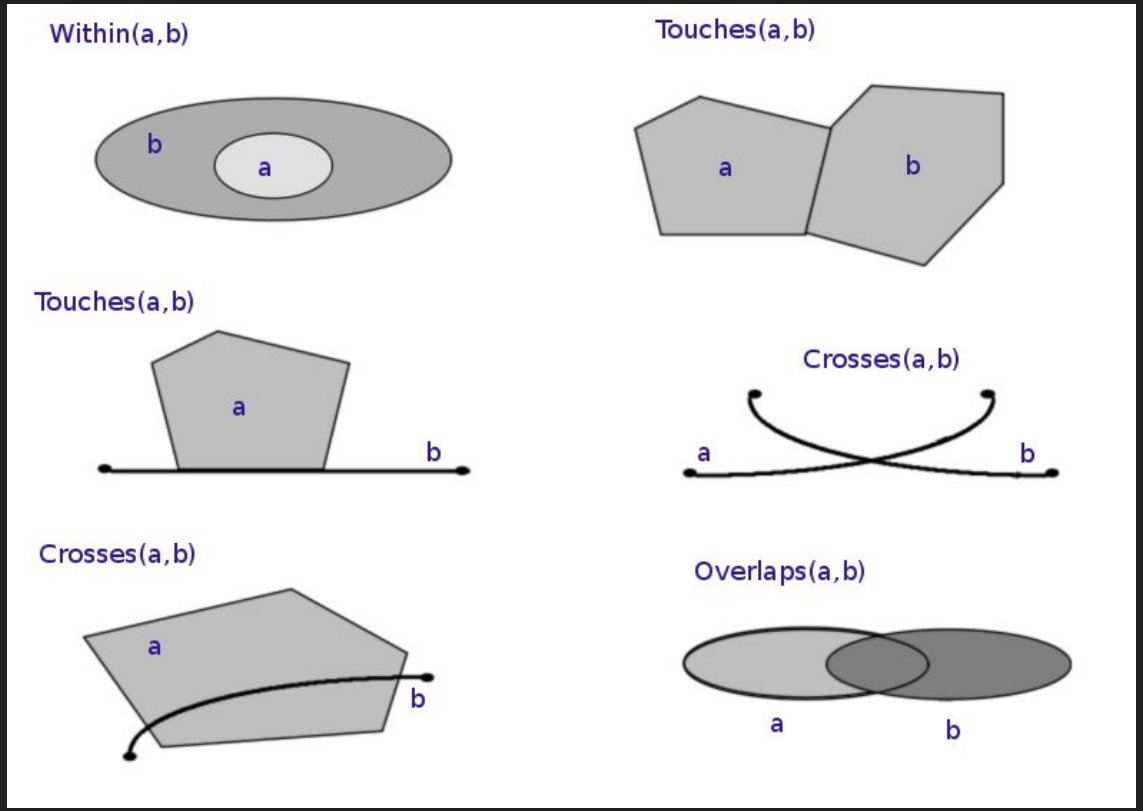
- Layers = database tables
  - geometry stored directly in the database
- GeoPackage
  - file containing SQLite database
  - ".gpkg" filename extension
  - open standard: anyone can implement it
- Geodatabase
  - folder containing database tables as files
  - ".gdb" folder name extension
  - ESRI proprietary
    - file format / database schema not documented
    - QGIS can read (sometimes) but not write

WKT/WKB

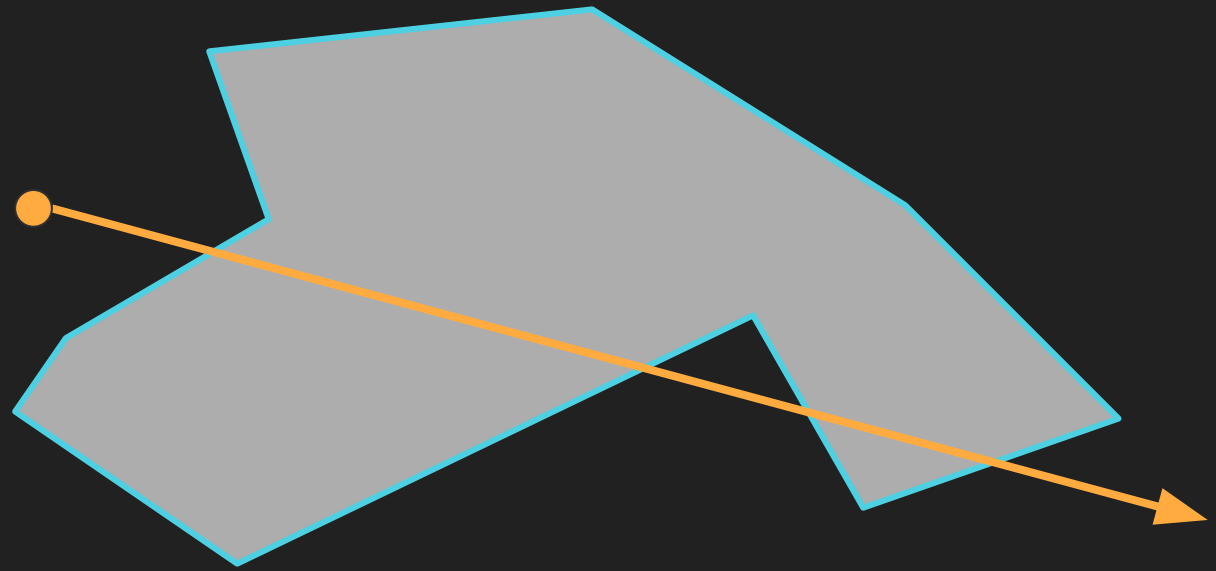
- **POINT:** `POINT(123.45 543.21)`
- **LINESTRING:** `LINESTRING(100.0 200.0, 201.5 102.5, 1234.56 123.89)`
- **POLYGON:**
  - `POLYGON((101.23 171.82, 201.32 101.5, 215.7 201.953, 101.23 171.82))`
    exterior ring, no interior rings
  - `POLYGON((10 10, 20 10, 20 20, 10 20, 10 10),`
    `(13 13, 17 13, 17 17, 13 17, 13 13))`
    exterior ring, one interior ring
- **MULTIPOINT:** `MULTIPOINT(1234.56 6543.21, 1 2, 3 4, 65.21 124.78)`
  three points
- **MULTILINESTRING:** `MULTILINESTRING((1 2, 3 4), (5 6, 7 8, 9 10), (11 12, 13 14))`
  first and last linestrings have 2 vertices each one; the second linestring has 3 vertices
- **MULTIPOLYGON:** `MULTIPOLYGON(((0 0,10 20,30 40,0 0),(1 1,2 2,3 3,1 1)),`
  `((100 100,110 110,120 120,100 100)))`
  two polygons: the first one has an interior ring

Spatial Relations

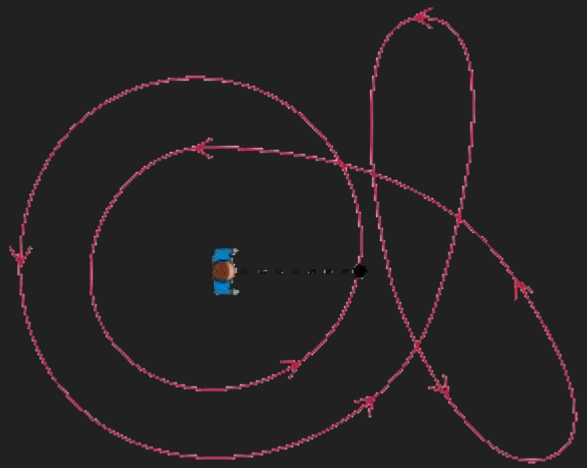
1. Equals
2. Disjoint
3. Touches
4. Contains
5. Covers
6. Intersects
7. Within
8. Covered by
9. Crosses
10. Overlaps



Point-in-Polygon test (ray-casting)



Point-in-Polygon test (winding number)

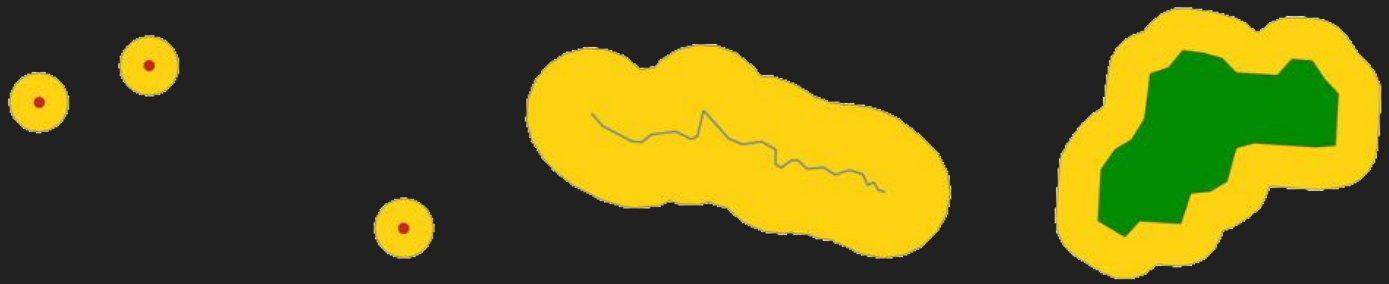


Vector Data Operations

- Buffering
- Overlay
- Editing

Proximity

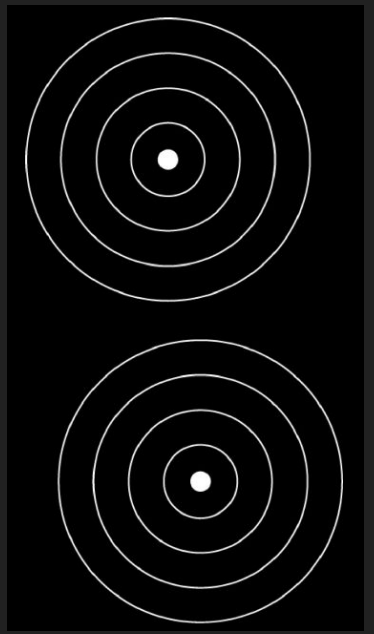
- Buffering
  - feature of interest + distance D → Buffer →region w/in D vs region beyond D
  - w/in distance D? = w/in buffer?
- Ubiquitous in GIS
  - exclusion zone
  - impact area
  - uncertainty
  - etc...



Buffering

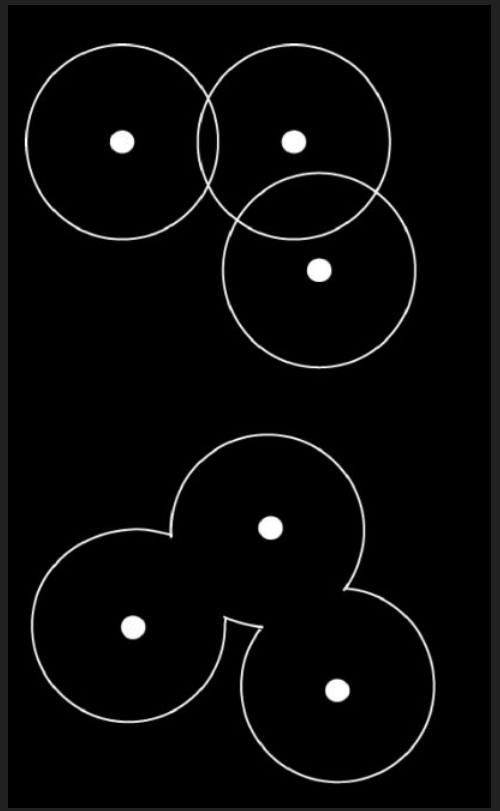
Variable Distance Buffers
(e.g. stream gradients)

Concentric Buffers
E.g. distance from well



Buffering with Dissolve

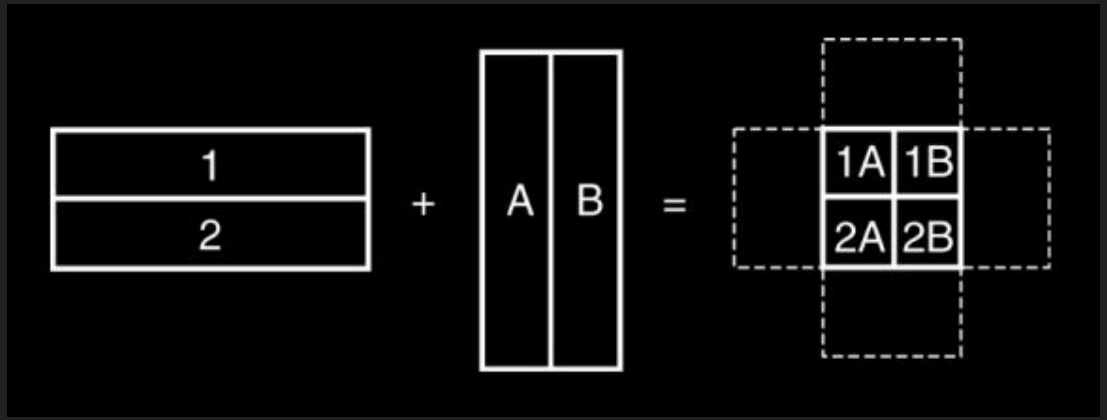
- Dissolve = remove overlap
  - Useful if same process governs creation of all the buffers
  - turns “each area” into “all areas”
- e.g. “drug-free school zone”
  - “No drugs here” more important than “no drugs near school X”



Overlay

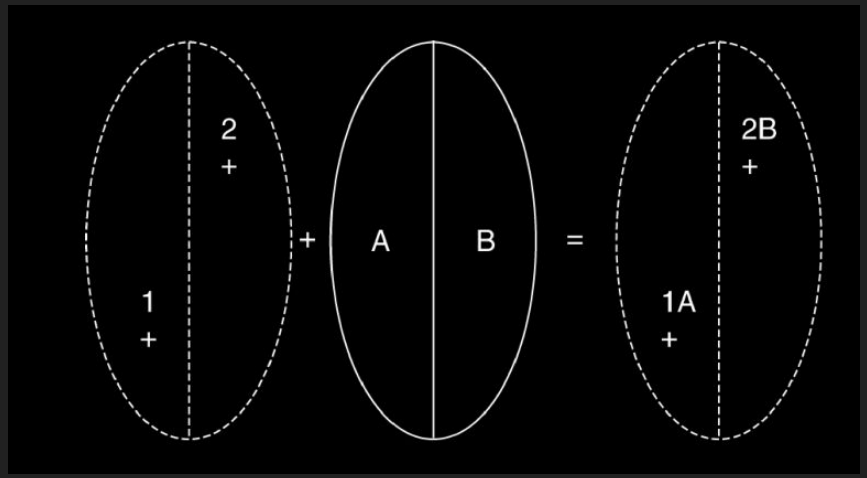
Combine feature layers

- New_Layer = Layer_1 {op} Layer_2
  - new geometry based on intersection of old geometries
  - old attributes distributed over new geometry
- Example: polygon AND polygon



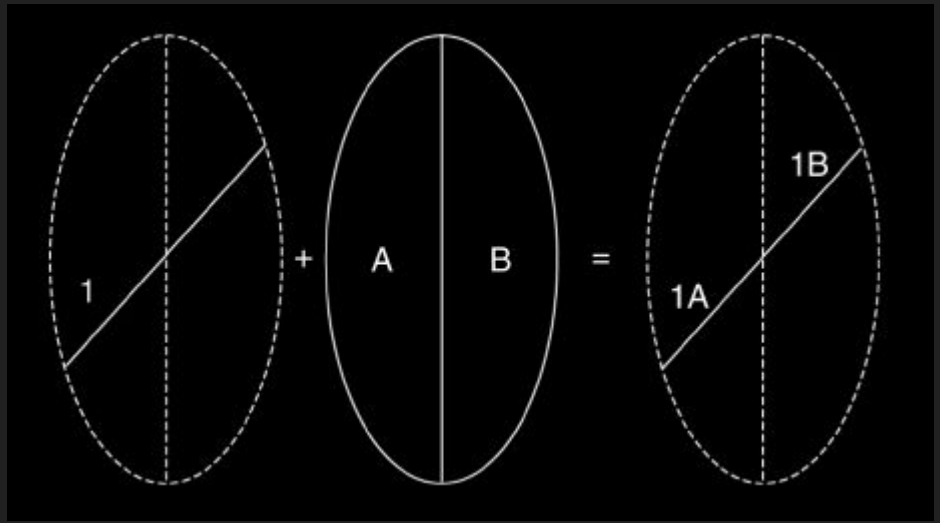
Overlay: Point-in-Polygon

- Points receive attributes of containing polygon
- ... not the reverse: why?
  - hint: what if >1 point in a polygon...



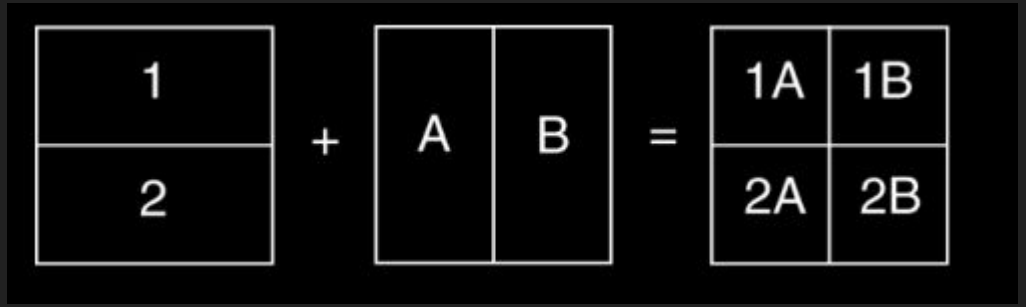
Overlay: Line-in-Polygon

- Old line → 2 new lines
  - 1 per polygon
- New lines receive attributes of containing polygon



Overlay: Polygon-on-Polygon

Polygons broken up as needed so attributes distribute correctly



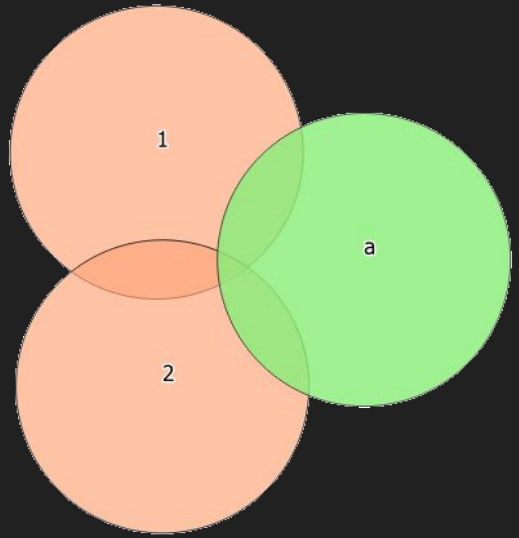
Overlay Operations

- Intersection
- Union
- Symmetrical Difference

- Difference
- Identity
- Update

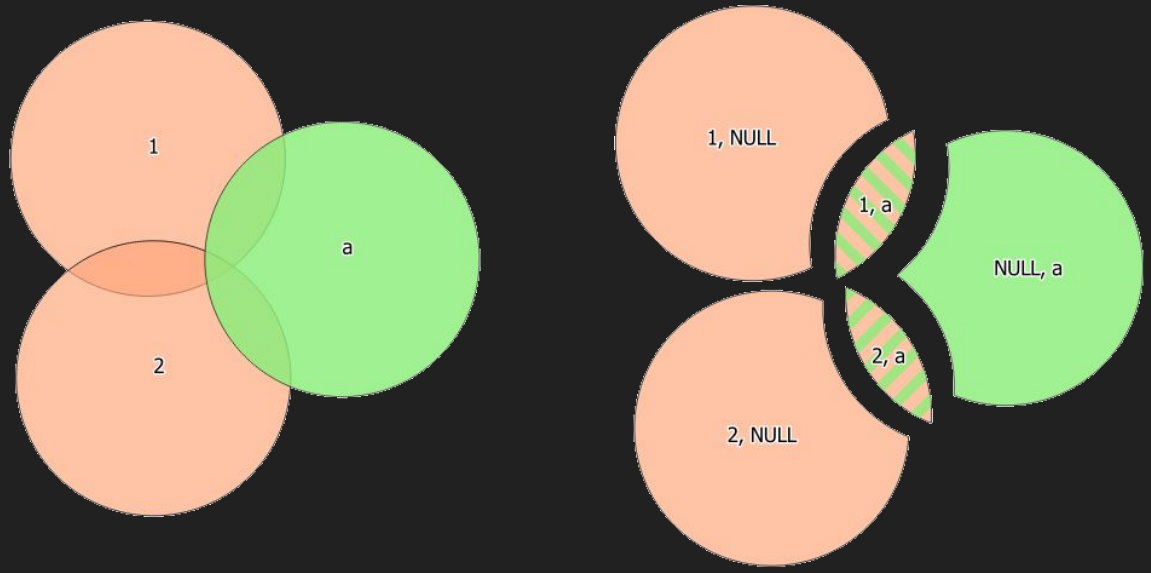
Intersection (AND)

- only areas common to both input layers
- output = input AND intersect



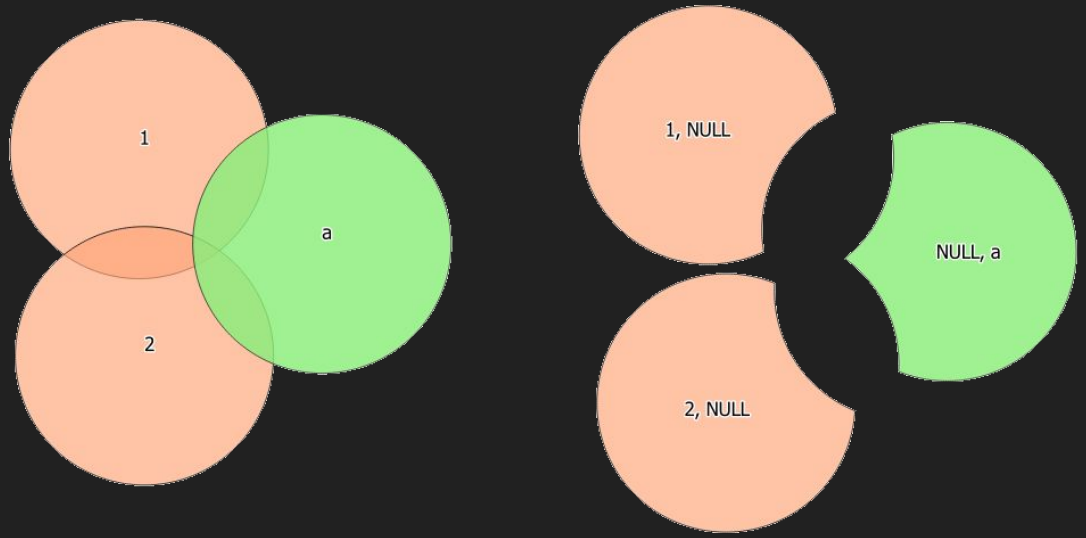
Union (OR)

- All areas from both input layers
- output = input 1 OR input2 (OR: inclusive “or” 1 or 2 or both)



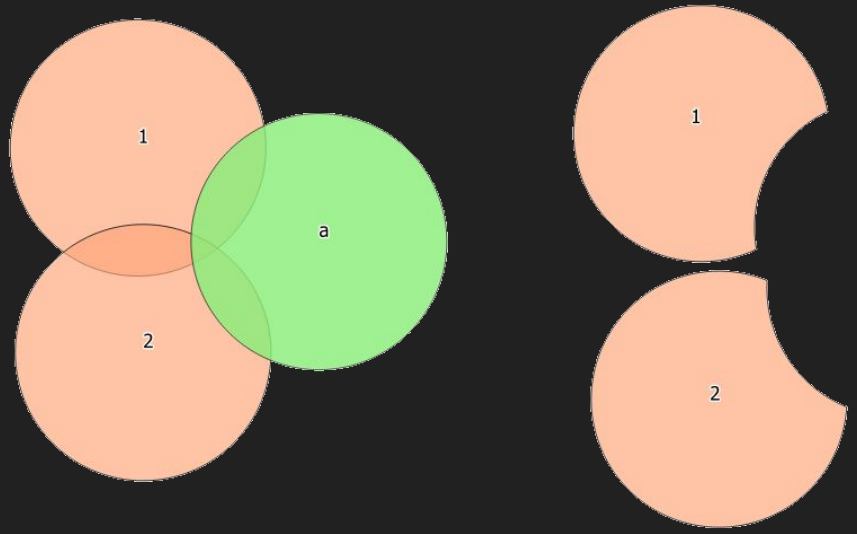
Symmetrical Difference (XOR)

- Only areas unique to 1 input layer
- output = input_1 XOR input_2 (XOR: exclusive “or” 1 or 2 but not both)



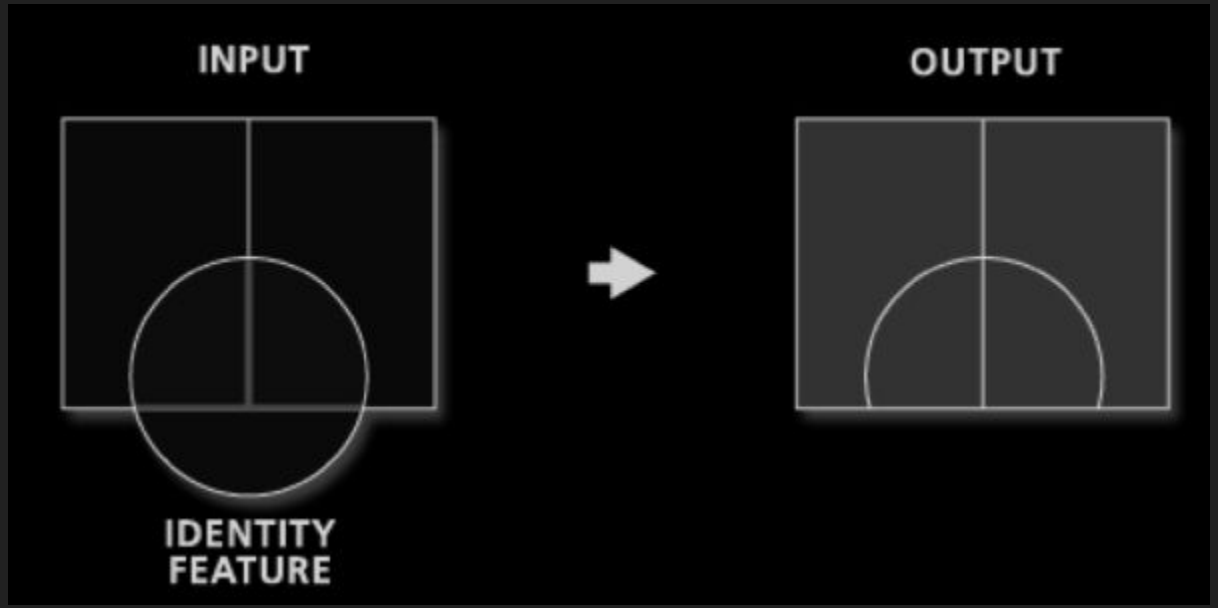
Difference

- Like symmetrical difference, but only keeps input layer
- output = (input XOR erase) AND input



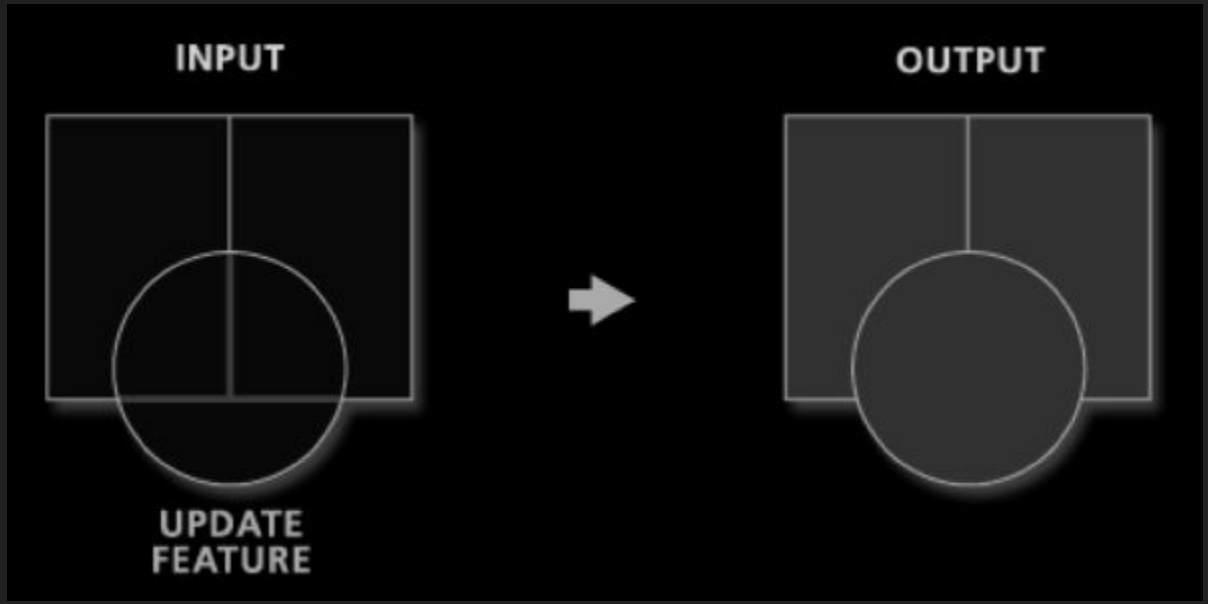
Identity

- Like intersect, but keeps all of 1st input layer
- output = (input AND identity) OR input



Update

- Like difference, but keeps update layer
- output = ((input XOR update) AND input) OR update



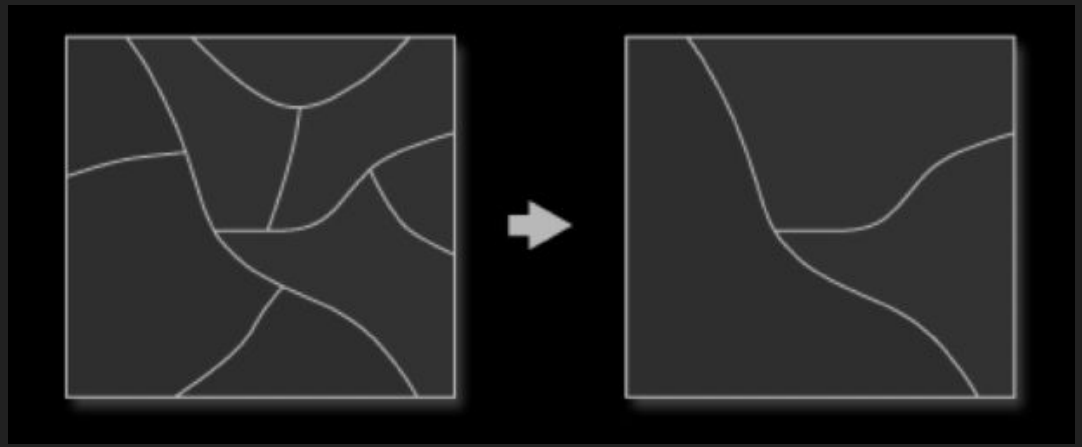
Vector Editing Operations

- Apply combinations of …
  - overlay operators
  - attribute queries
- ... to create new feature layers

- Dissolve
- Clip
- Merge
- Select
- Eliminate

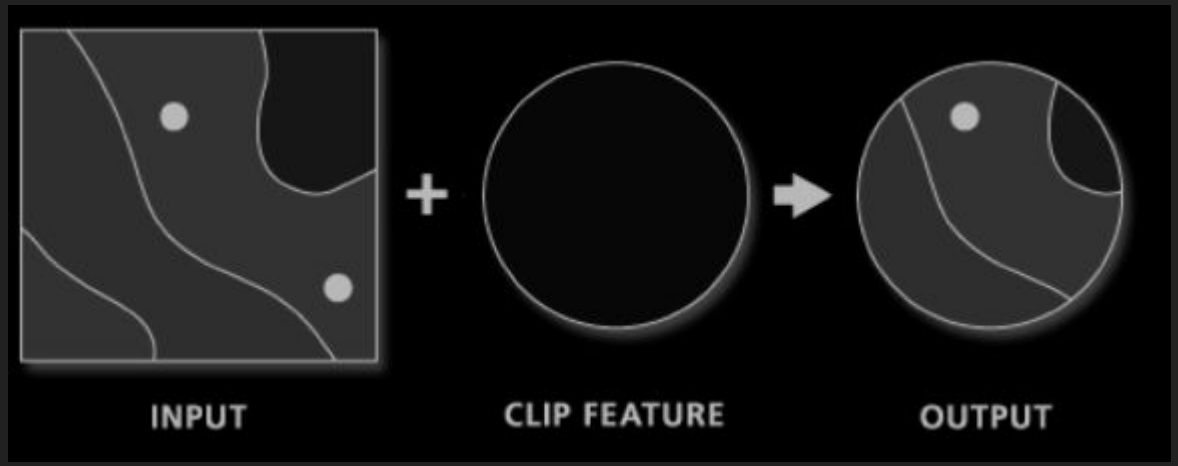
Dissolve

- Combine adjacent polygons based on shared attribute value
- i.e. remove unnecessary boundaries
  - simplifies analysis
  - smaller dataset



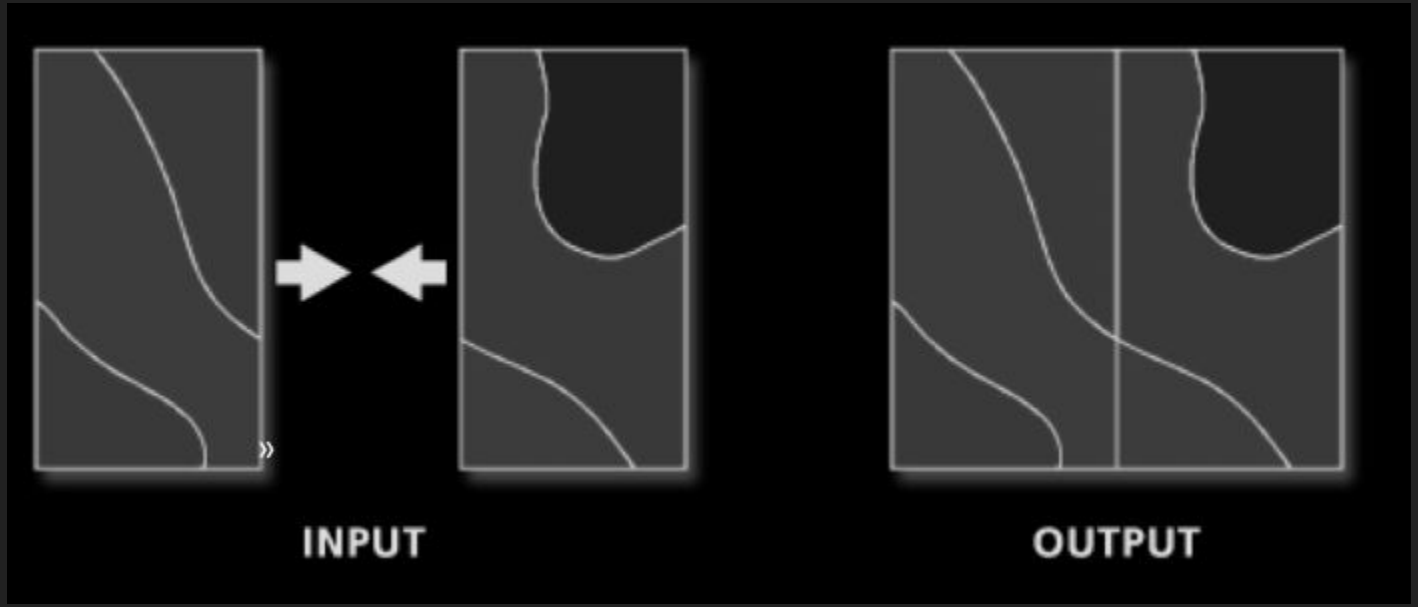
Clip

- “Cut out” input layer using feature(s) from clip layer
- Think “stencil” ...



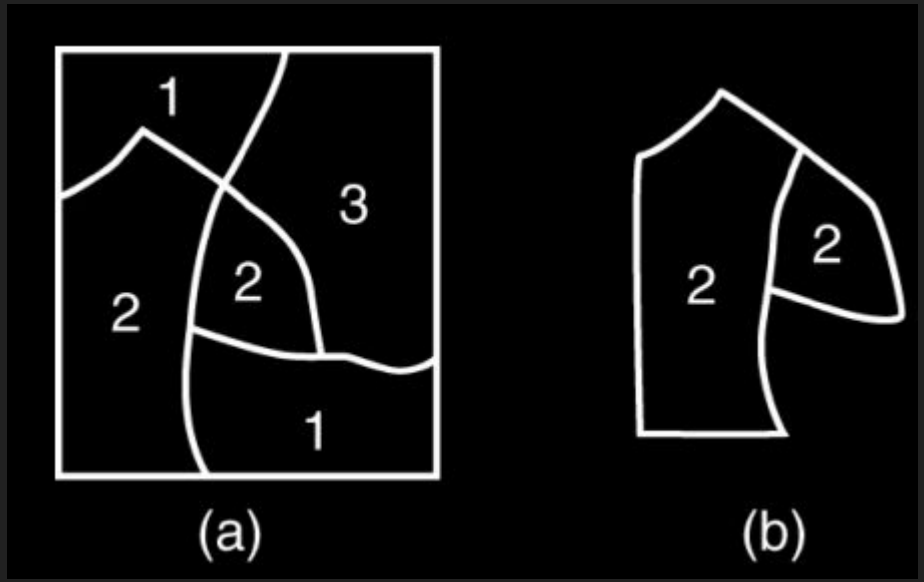
Merge

Combine Vector Datasets



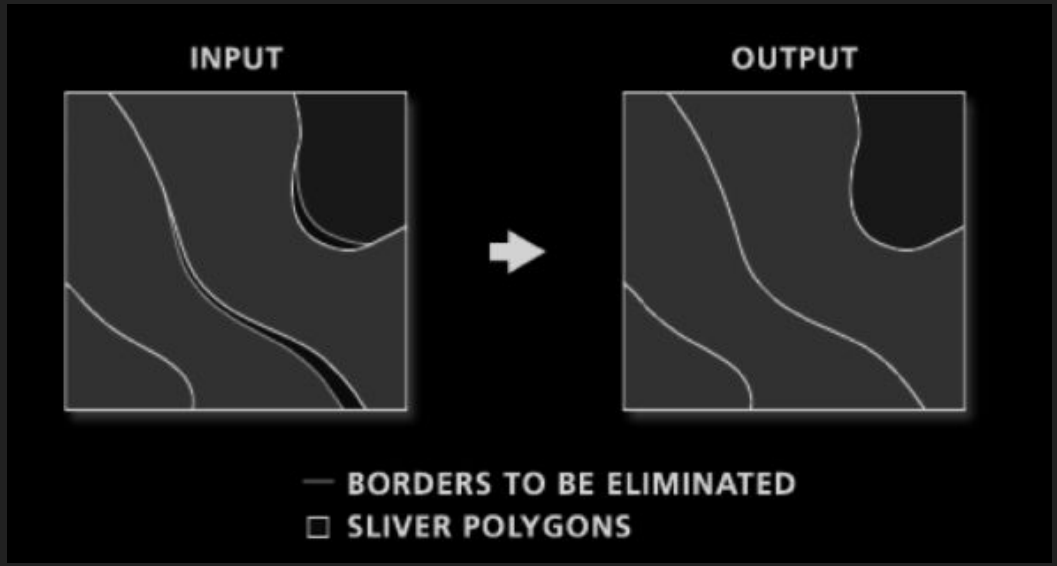
Select

Extract selected features into new layer



Eliminate

- Get rid of “spurious” polygons…
  - e.g. digitization errors
- … by merging into neighbors
  - Largest
  - Longest common boundary



Graphics Credits

- Introduction to Geographic Information Systems, 5/e
- ArcMap Help
- GIS Fundamentals, 6/e