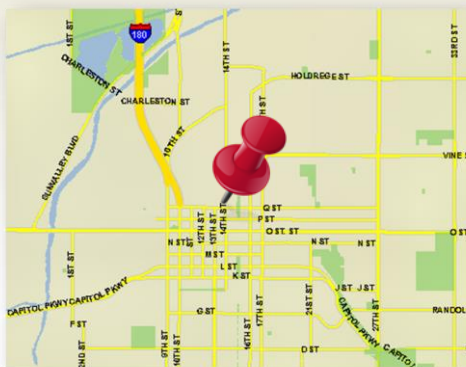




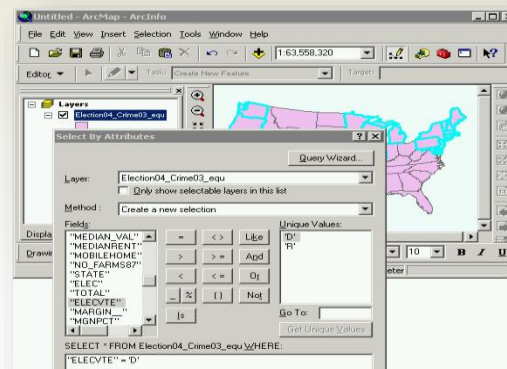
# Working with tabular data in GIS

*Patricia Carbajales-Dale*  
[clemsongis.org](http://clemsongis.org)

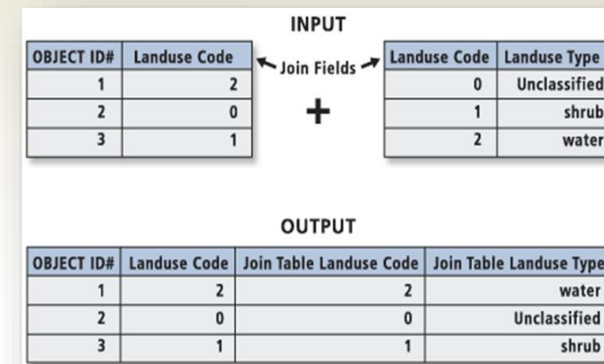
# Overview



Geocode



Query

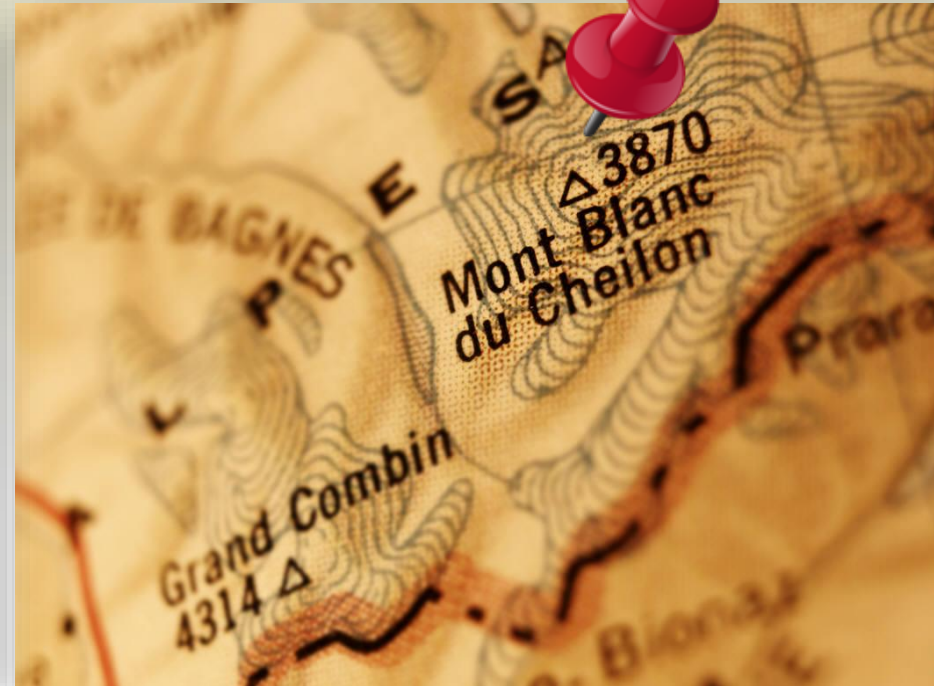
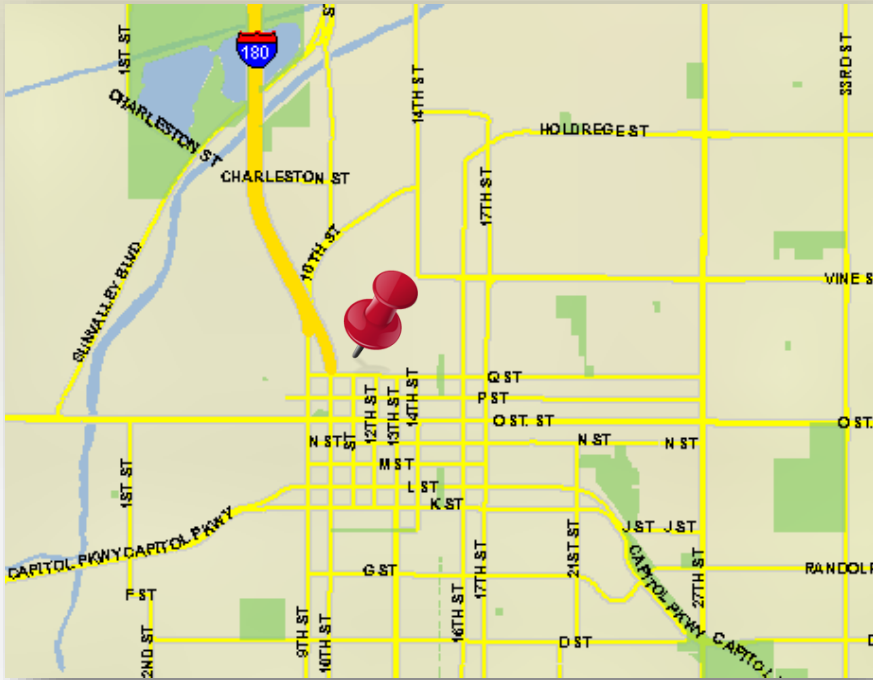


Join

# What is Geocoding?

## □ Description of a Location → Location on a Map

- Street Address, Intersection
- Place name (park, monument)
- Buildings, zip codes, ip addresses



# Geocoding

Geocoding is the process of **assigning a location**, usually in the form of coordinate values (points), to an **address** by comparing the descriptive location elements in the address to those present in the reference materia

Original address: 127 West Birmingham Drive, 92373

Address parsed: 127 | West | Birmingham | Drive | 92373

Abbreviations standardized: 127 | W | Birmingham | Dr | 92373

Elements assigned to match keys: [HN]: 127 [ST]: Dr [SD]: W [ZP]: 92373 [SN]: Birmingham

Index values calculated: [HN]: 127 [ST]: Dr [SD]: W [ZP]: 92373 (index #92373) [SN]: Birmingham (Soundex index #B655)

Search address locator and identify candidates.

Score of each potential match established:

Street	Number	Direction	MatchScore
Birmingham	129	W	90
Birmingham	125	W	85
Burnington	1100	W	60
Brunton	129	N	70
Broomstick	145	S	30

List of candidates filtered:

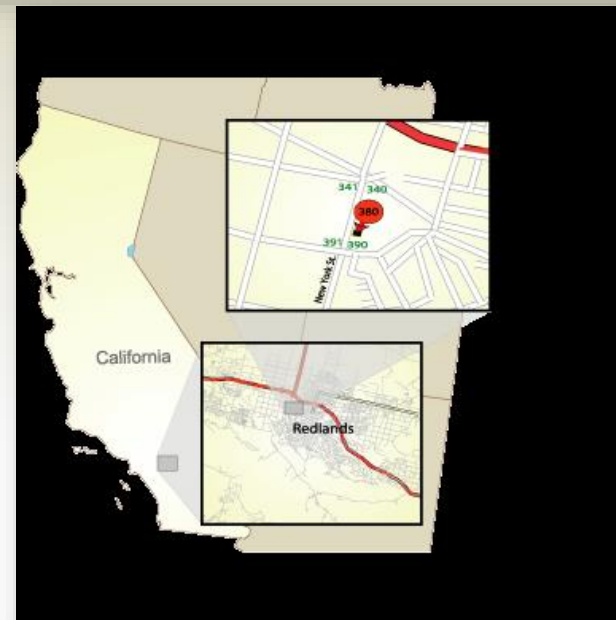
Street	Number	Direction	MatchScore
Birmingham	129	W	90
Birmingham	125	W	85

Best candidate matched: 129 W Birmingham Dr. , 92373

Matched feature indicated:

**26376 Alpine Lane, Twin Peaks, CA 92391**

House number | Street name | Street type | City | State | Postal code



# Geocoding Options

## ❑ Google

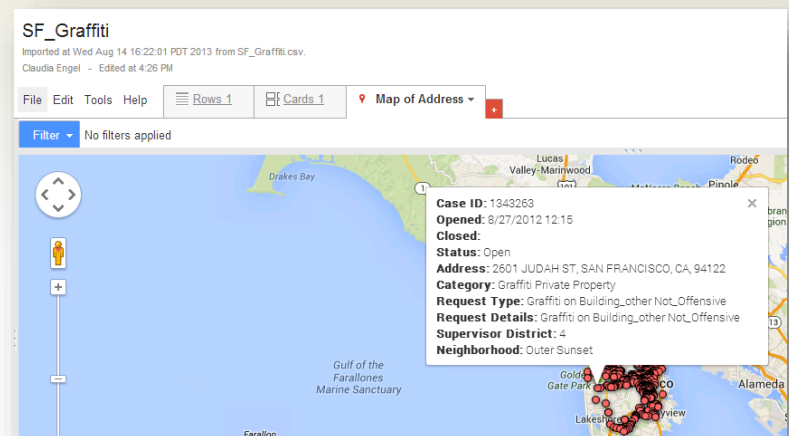
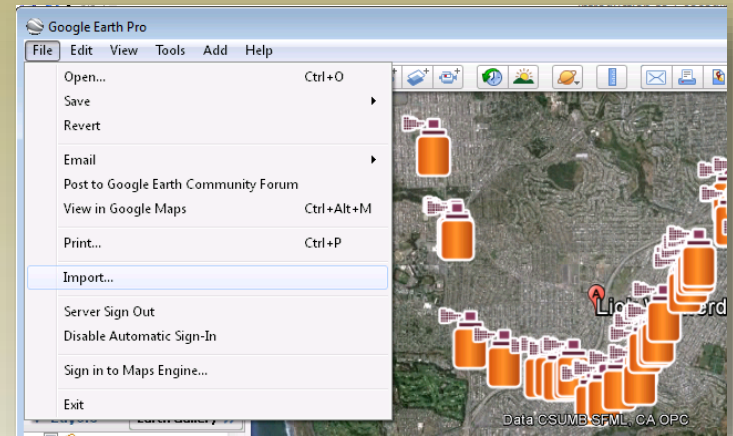
- Google Earth Pro
- Google Fusion Tables
- Google Maps API

## ❑ ArcGIS

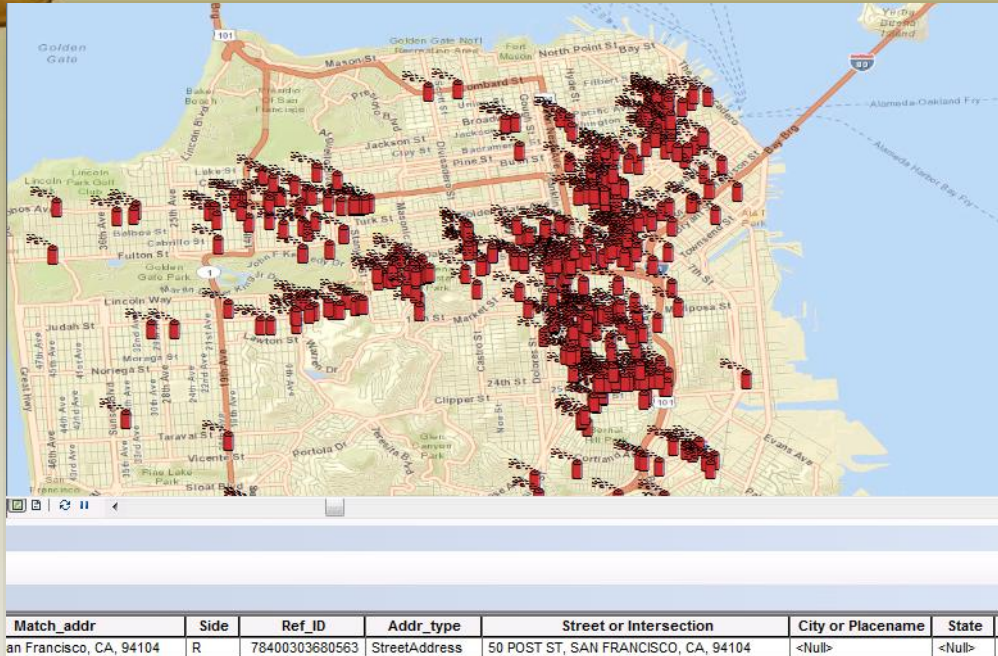
- ArcMap
- ArcGIS Online
- Building your own address locator

## ❑ Automated Processes

## ❑ Web Services



# Geocoding in ArcGIS



## Specifications

- Import .csv or Excel files
- No limit on number available
- Max. control on the engine
- Flexible entry format
- No spaces or special characters on field names

## Results (out of 500 records):

- 496 records geocoded in place
- 4 records rematched manually

# How does it work?

- ❑ Standardize your Address Data
- ❑ Compare to Reference Data (geographic data)
- ❑ Return best match of results

Original address: 127 West Birmingham Drive, 92373

Address parsed: 127 | West | Birmingham | Drive | 92373

Abbreviations standardized: 127 | W | Birmingham | Dr | 92373

Elements assigned to match keys: [HN]: 127 [ST]: Dr [SD]: W [ZP]: 92373 [SN]: Birmingham

Index values calculated: [HN]: 127 [ST]: Dr [SD]: W [ZP]: 92373 (index #92373) [SN]: Birmingham (Soundex index #B655)

Search address locator and identify candidates.

Score of each potential match established

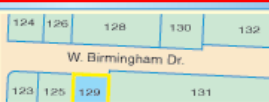
Street	Number	Direction	MatchScore
Birmingham	129	W	90
Birmingham	125	W	85
Burnington	1100	W	60
Brunton	129	N	70
Broomstick	145	S	30

List of candidates filtered

Street	Number	Direction	MatchScore
Birmingham	129	W	90
Birmingham	125	W	85

Best candidate matched: 129 W Birmingham Dr. , 92373

Matched feature indicated



**26376 Alpine Lane, Twin Peaks, CA 92391**

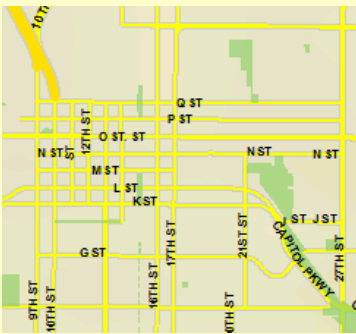
House number    Street name    Street type    City    State    Postal code



# Address Locator

## Address Locator

### Reference Data



### Rules

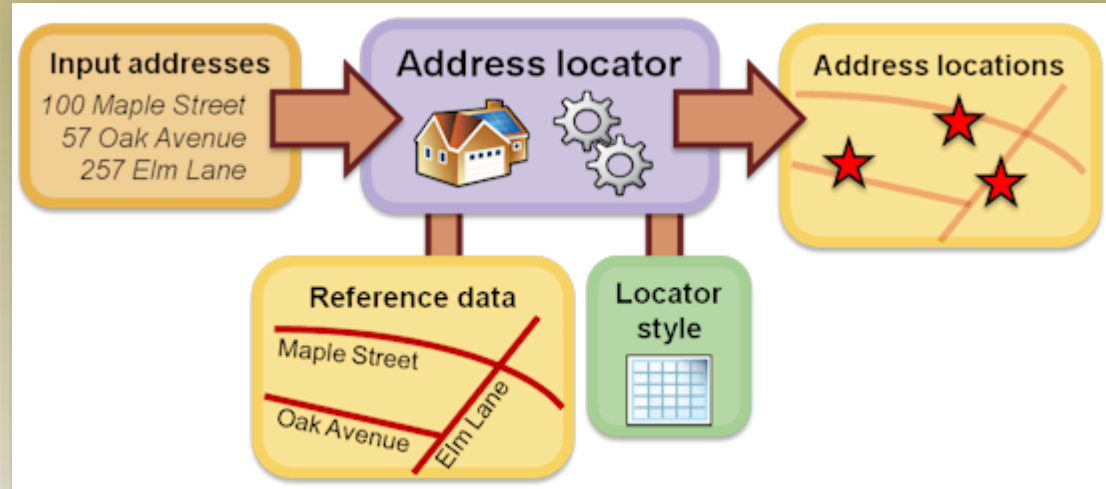
Even house numbers  
are on the right side  
of the street

### Street Components

Avenue = Ave  
Street = St  
Highway = Hwy



# Geocoding Workflow in ArcGIS

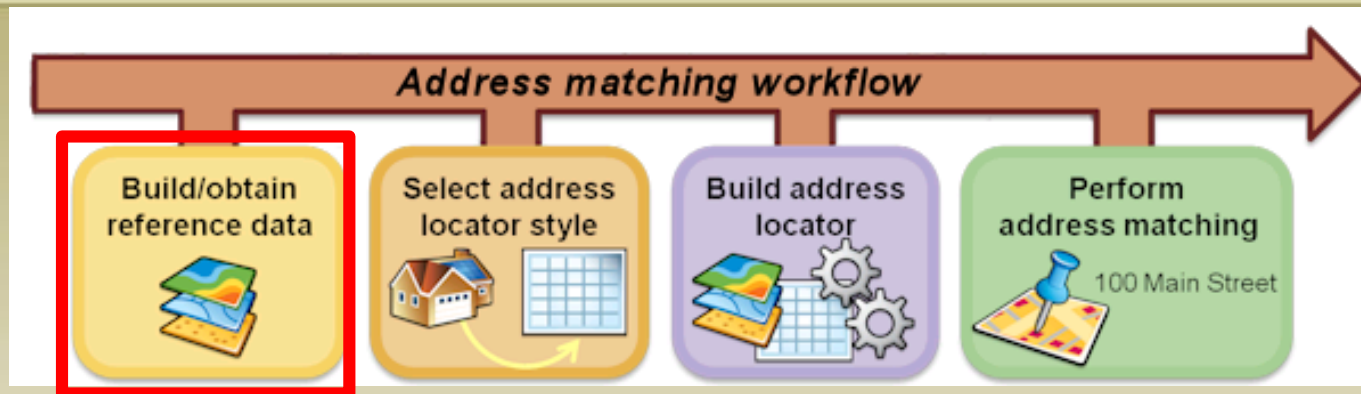


Build or obtain  
reference data

Determine  
address  
locator style

Locate addresses  
& rematch remaining  
addresses

# Reference Data



## Snapshot of Geographic Information

- ESRI Street Data
- US Census TIGER roads
- Parcels
- Gazetteers
- General Place names
- Zip Codes/ZCTAs
- NavTec, TeleAtlas

House	PreDir	PreType	Name	SufType	SufDir	City
10	N	Ave	Bluewater	Ave	N	Bel Air
23	E		Heathercliff	Ct	NE	Beverly Hills
100	S		Mapleton	Ln	E	Brentwood
594	W		Mulholland	Pkwy	SE	Malibu
1085			Napoli	Rd	S	Pacific Palisades
			Rexford	St	W	Santa Monica
				Way		Sherman Oaks



# Querying Tables

- Queries are the most common operations in a spatial database
  - used to **find features** that meet certain criteria
  - used to **isolate features** for future analysis (subset or filter data)
- Structured Query Language (**SQL**)
  - standard language for retrieving and updating information in a database
- Most common operation is the **SELECT**
  - retrieves data from a table or multiple related tables
  - `SELECT * FROM "Table" WHERE "Field = Value"`
  - returns a subset of records, e.g. restricts records based on some condition
- Can be simple or complex
  - simple: `"area" > 20`
  - complex: `("area" > 20) AND ("area" < 50)`

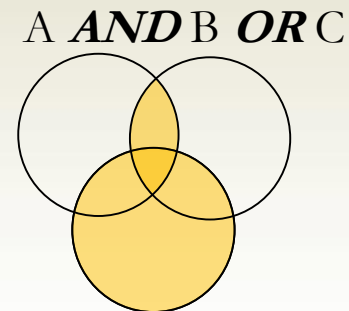
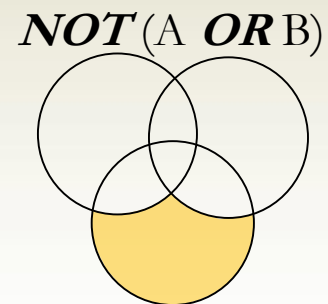
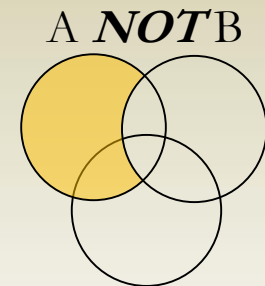
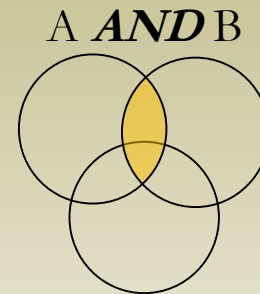
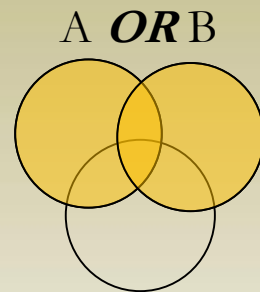
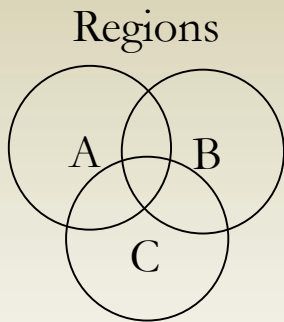
- **Set Algebra**

- basis for simple query expressions
  - = (equal)
  - > (greater than)
  - < (less than)
  - <> (not equal to)
- can be used on both **strings** and **numbers**

- **Boolean Algebra**

- basis for complex query expressions
  - **NOT**: accepts value of one input and outputs opposite value
  - **AND**: accepts two values as input and outputs the intersection of both
  - **OR**: accepts two values as input and outputs the sum of both
- **parentheses** may be required and the **order** of precedence is important

# Boolean Algebra



# Simple Queries

ID	Area	Landuse	Municipality
1	10.5	Urban	City
2	330.3	Farm	County
3	2.4	Suburban	Township
4	96.0	Suburban	County
5	22.1	Urban	City
6	30.2	Farm	Township
7	4.4	Urban	County

- Find all records with **Area** greater than 20.0      ←..... Area > 20.0
- Find all **Urban Landuse**      ←..... Landuse = Urban
- Find all records with **Area** less than or equal to 55.0      ←..... Area <= 55.0
- Find all **Non-Urban Landuse**      ←..... Landuse <> Urban



# Compound Queries

Find all the municipalities that are not urban cities

**NOT [(Landuse = Urban) AND (Municipality = County)]**

- Landuse = Urban ←..... Set1
- Municipality = County ←..... Set2
- Set1 **AND** Set2 ←..... Set3
- **NOT** (Set3)

ID	Area	Landuse	Municipality
1	10.5	Urban	City
2	330.3	Farm	County
3	2.4	Suburban	Township
4	96.0	Suburban	County
5	22.1	Urban	City
6	30.2	Farm	Township
7	4.4	Urban	County



# Compound Queries

Excluding counties, which municipalities are not urban?

**[NOT (Landuse = Urban)] AND [NOT (Municipality = County)]**

- Landuse = Urban ←..... Set1
- **NOT** (Set1) ←..... Set2
- Municipality = County ←..... Set3
- **NOT** (Set3) ←..... Set4
- Set2 **AND** Set4

ID	Area	Landuse	Municipality
1	10.5	Urban	City
2	330.3	Farm	County
3	2.4	Suburban	Township
4	96.0	Suburban	County
5	22.1	Urban	City
6	30.2	Farm	Township
7	4.4	Urban	County





# Querying Tables

ID	Area	Landuse	Municipality
1	10.5	Urban	City
2	330.3	Farm	County
3	2.4	Suburban	Township
4	96.0	Suburban	County
5	22.1	Urban	City
6	30.2	Farm	Township
7	4.4	Urban	County

Find all **Suburban** and **Urban** Landuse **greater than 20.0** and **less than 5.0**

Find all **County Non-Urban** Landuse **less than 100.0**

Find all **Urban Cities** **greater than 20.0**



# Other SQL Statements

- IN

- search for **several strings or values** in a field; alternative to multiple ORs

```
state IN ('Alabama', 'Alaska', 'California', 'Florida')
```

- LIKE

- used with **wildcards** to build a **partial string search**

```
state LIKE 'Mi%'
```

- % (shapefile); \* (geodatabase)

- **wildcard** representing any **multiple** characters; used with LIKE

```
place_name LIKE '%ort%'
```

- \_ (shapefile); ? (geodatabase)

- **wildcard** representing any **single** character; used with LIKE

```
first_name LIKE '_arry'
```

# Querying Tables

**Select By Attributes**

Layer: Landuse

Only show selectable layers in this layer

Method: Create a new selection

Fields:

- "FID"
- "AREA"
- "PERIMETER"
- "LU\_CODE"

Unique Values:

- 11
- 12
- 13
- 14
- 16
- 17
- 21
- 22

SELECT \* FROM landuse\_query WHERE:  
"LU\_CODE" = 22

SELECT \* FROM landuse\_query WHERE:  
"LU\_CODE" = [ ]

OR NOT

Clear Verify Help Load.. Save... Apply Close

# Joins

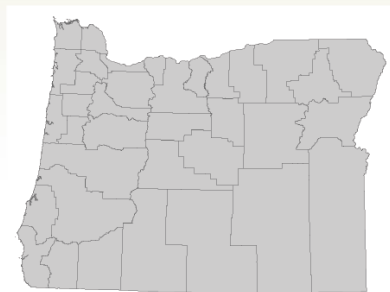
Joins allow us to **append** new data to existing geometry or tables, extending their informational range.

Geometry and attribute data are not always stored together.

Any data that shares a common attribute with a spatial dataset can be mapped.

- Map new attributes to existing features
- Extend analytical range of spatial data
- Based on common attributes between tables
- Everything is spatial!

**Boring Geometry**

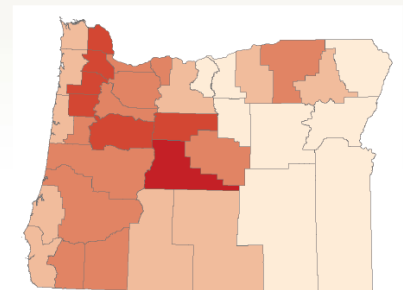


**Boring Statistics**

OBJECTID	Unit	F1	F2	Number	Percent
1	OREGON	342139	280107	496675	0.119739
2	Baker	18741	18134	467	-0.036258
3	Benton	78153	88579	7426	0.095019
4	Clatsop	336361	376662	37661	0.111117
5	Clatsop	33630	37039	1409	0.039445
6	Columbia	43660	48351	5791	0.132943
7	Cook	62779	53043	264	-0.004205
8	Crook	19152	20978	1786	0.093629
9	Curry	21137	22364	1227	0.05805
10	Deschutes	115367	157733	42366	0.367226
11	Douglas	100288	107027	7266	0.072391
12	Gilliam	1016	1071	44	-0.022977
13	Grant	7936	7445	490	-0.061752
14	Harney	7039	7422	483	-0.024570
15	Hood River	20411	22346	1935	0.094802
16	Jackson	181289	203206	21931	0.121819
17	Jefferson	19009	21720	2711	0.142617
18	Josephine	75725	82713	6587	0.092267
19	Klamath	63775	66380	2605	0.040847
20	Lake	7422	7895	473	0.063729
21	Lane	320569	351715	28758	0.089939
22	Linn	44479	48034	1555	0.03496
23	Linn	103009	110072	15063	0.13190



**Awesomeness!**



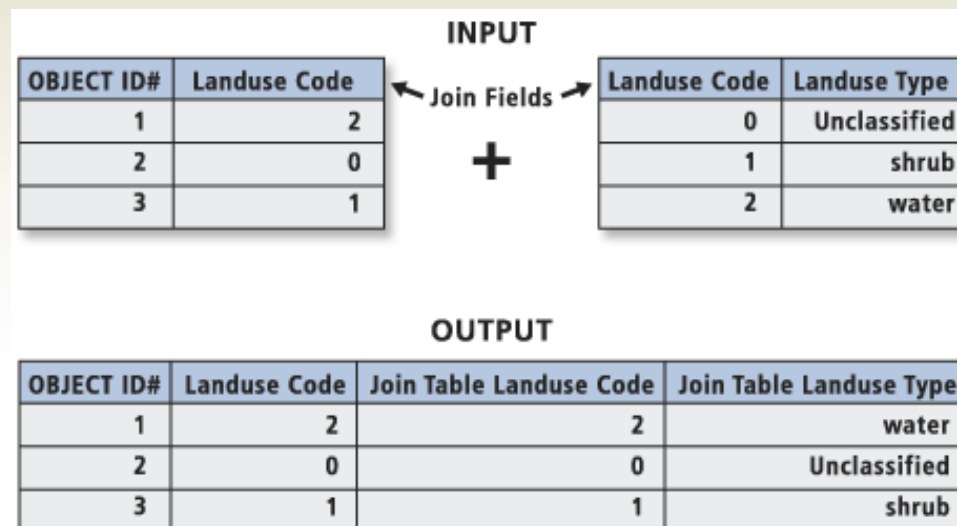
# Joins

Join one table to another table or layer based a common field.

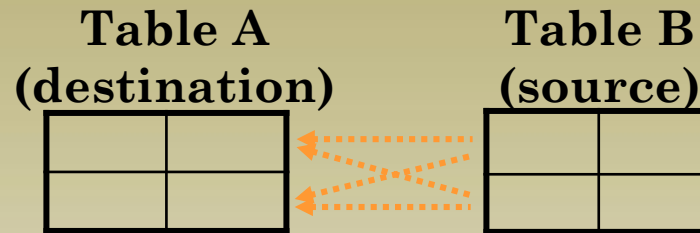
Attributes of the join table are appended to the input table.

A match is made when the input field and output field values are equal.

- Common identifier or “key”
- Field names do not need to match
- Field type must be the same (numbers, text, etc.)
- Field data should be the same
- The join is dynamic



# Cardinality of Relationships



- ***One-to-one (join)***
  - every record in A matches exactly one record in B; every record in B matches exactly one record in A.
- ***Many-to-one (join)***
  - every record in table B matches one or more records in table A.
- ***One-to-many (relate)***
  - one or more records in table B matches exactly one record in table A.
- ***Many-to-many (relate)***
  - every record in A matches one or more records in B; every record in B matches one or more records in A.



# Joining and Relating Tables

one-to-one (join)

<i>Country</i>	<i>Capital</i>
France	Paris
Sweden	Stockholm
United Kingdom	London

many-to-one (join)

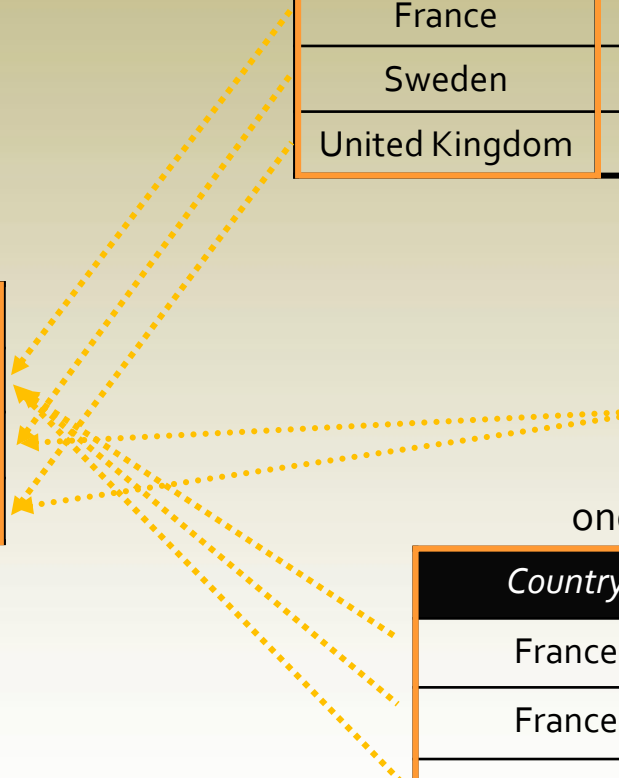
<i>Gov-ID</i>	<i>Government</i>
1	Republic
2	Const. Monarchy

destination table

<i>Country</i>	<i>Gov-ID</i>
France	1
Sweden	2
United Kingdom	2

one-to-many (relate)

<i>Country</i>	<i>Major City</i>
France	Bordeaux
France	Marseille
France	Lyon
United Kingdom	London
United Kingdom	Manchester



# Joining Tables

Join Data

Join lets you append additional data to this layer's attribute table so you can, for example, symbolize the layer's features using this data.

What do you want to join to this layer?

Join attributes from a table

1. Choose the field in this layer that the join will be based on:

from disk:

Joining table. Fields being

Keep only matching records  
If a record in the target table doesn't have a match in the join table, that record is removed from the resulting target table.

About Joining Data

OK Cancel

Attributes of census

OBJECTID	Shape *	STFID	STFID *	POP2000	PDENSITY	HOUSEHOLDS	HSE_UNITS
1	Polygon	060816134003	060816134003	313	709	123	126
2	Polygon	060816138001	060816138001	218	15	83	102
3	Polygon	060816134001	060816134001	618	638	214	225
4	Polygon	060816133002	060816133002	403	678	143	146
5	Polygon	060816133001	060816133001	525	732	196	201
6	Polygon	060816111003	060816111003	1090	3278	455	468
7	Polygon	060816111001	060816111001	1631	4072	613	620
8	Polygon	060816134002	060816134002	1923	142	711	758
9	Polygon	060816097002	060816097002	2323	4037	847	862
10	Polygon	060816098004	060816098004	981	5382	367	372
11	Polygon	060816097001	060816097001	1831	1200	691	705

Record: 1 Show: All Selected Records (0 out of 233 Selected) Options

Attributes of census

060816111003	1090	3278
060816111001	1631	4072
060816134002	1923	142
060816097002	2323	4037
060816098004	981	5382

Record: 1 Show



# Join Validation

Check for field names that **start** with an invalid character:

~@#\$%^&\*()-+=|\ \, <> ? / { } . ! [ ] ; \_ 0123456789

Check for field names that **contain** an invalid character:

~@#\$%^&\*()-+=|\ \, <> ? / { } . ! [ ] ;

Check for field names that match:

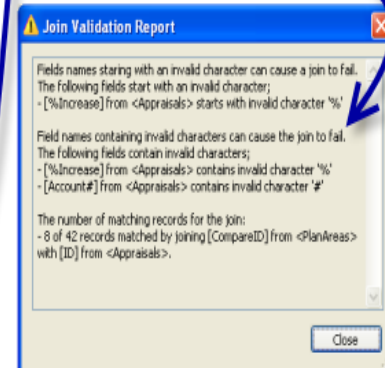
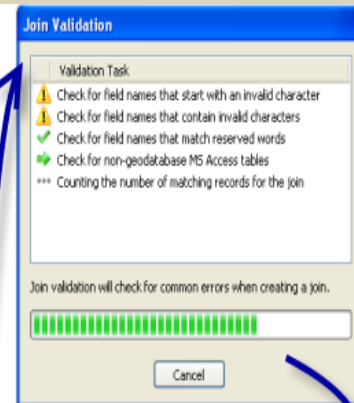
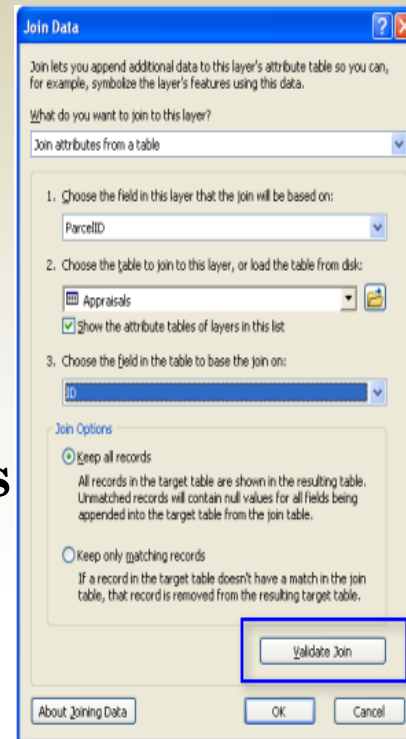
“California” <> “ California”

“California” <> “california”

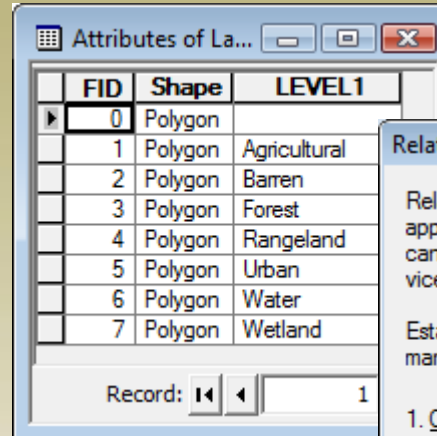
1 <> “1”

Check for field names in the table that  
Are Microsoft **Access reserved words**  
date, day, month, table, text, user, when,  
where, year, and zone

Use **Validate Join** tool!

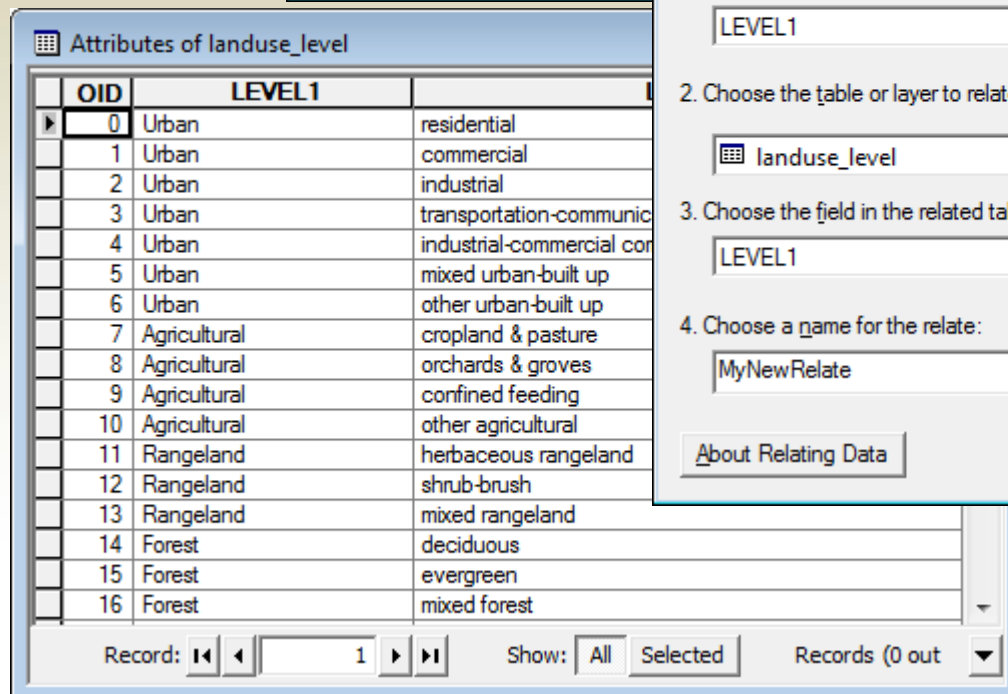


# Relating Tables



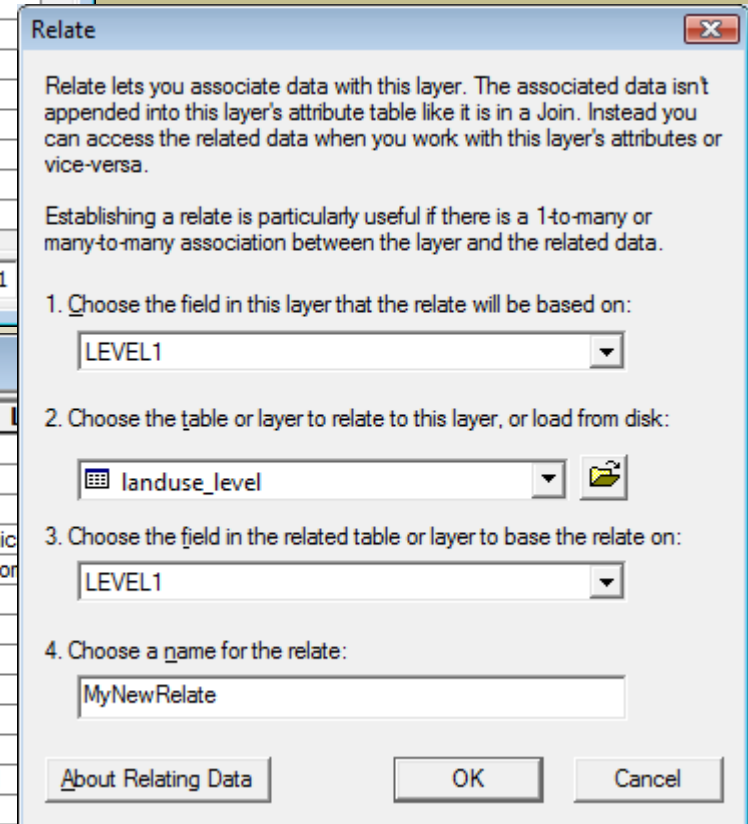
Attributes of Layer dialog box showing a table with columns FID, Shape, and LEVEL1. The table contains 8 rows of data. An orange box highlights the LEVEL1 column header, and a dashed orange arrow points from it to the 'Attributes of landuse\_level' dialog box.

FID	Shape	LEVEL1
0	Polygon	
1	Polygon	Agricultural
2	Polygon	Barren
3	Polygon	Forest
4	Polygon	Rangeland
5	Polygon	Urban
6	Polygon	Water
7	Polygon	Wetland



Attributes of landuse\_level dialog box showing a table with columns OID, LEVEL1, and another column. The table contains 17 rows of data. An orange box highlights the LEVEL1 column header, and a dashed orange arrow points from it to the 'Attributes of Layer' dialog box.

OID	LEVEL1	
0	Urban	residential
1	Urban	commercial
2	Urban	industrial
3	Urban	transportation-communic
4	Urban	industrial-commercial cor
5	Urban	mixed urban-built up
6	Urban	other urban-built up
7	Agricultural	cropland & pasture
8	Agricultural	orchards & groves
9	Agricultural	confined feeding
10	Agricultural	other agricultural
11	Rangeland	herbaceous rangeland
12	Rangeland	shrub-brush
13	Rangeland	mixed rangeland
14	Forest	deciduous
15	Forest	evergreen
16	Forest	mixed forest



Relate dialog box with instructions and configuration options.

Relate lets you associate data with this layer. The associated data isn't appended into this layer's attribute table like it is in a Join. Instead you can access the related data when you work with this layer's attributes or vice-versa.

Establishing a relate is particularly useful if there is a 1-to-many or many-to-many association between the layer and the related data.

1. Choose the field in this layer that the relate will be based on:  
LEVEL1
2. Choose the table or layer to relate to this layer, or load from disk:  
landuse\_level
3. Choose the field in the related table or layer to base the relate on:  
LEVEL1
4. Choose a name for the relate:  
MyNewRelate

Buttons: About Relating Data, OK, Cancel

# Next workshop: Field Data Collection



# THANK YOU!!