GEOCODING

Sandeep Talasila, GISP

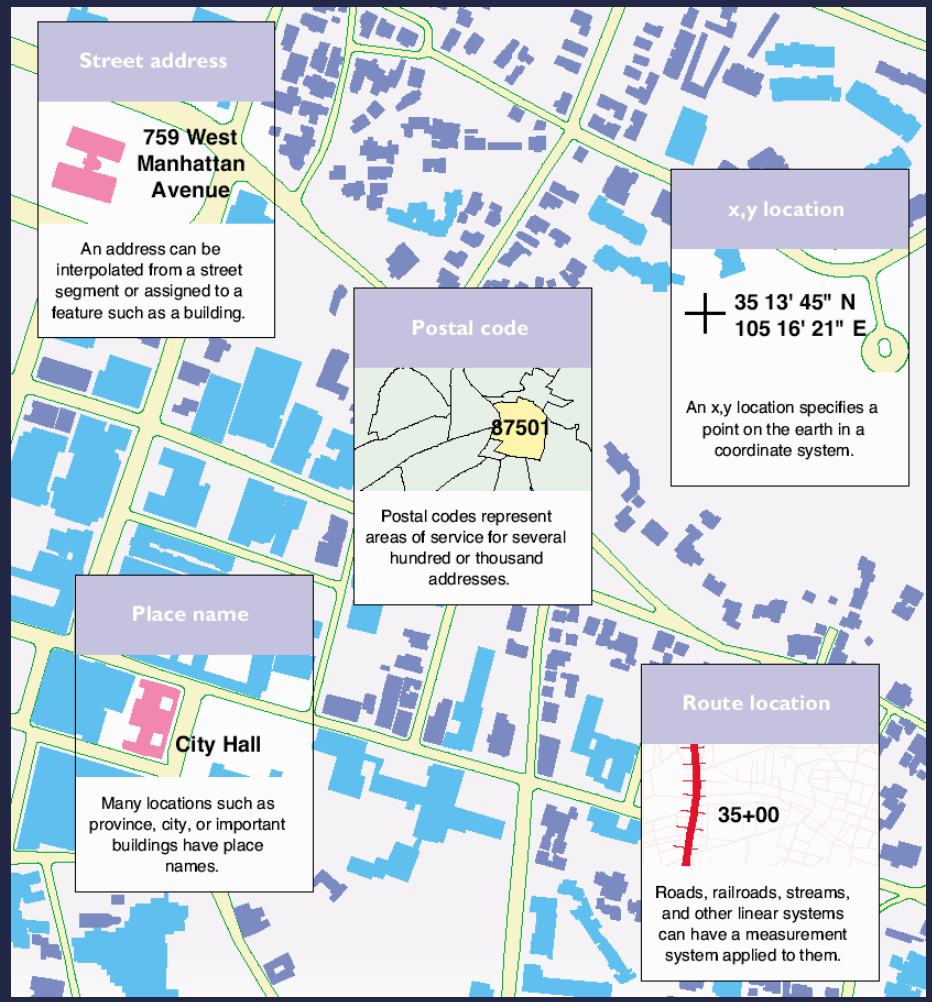
GEOCODING

- *Geocoding is the process of transforming a description of a location – such as a pair of coordinates, an address, or a name of a place – to a location on the earth's surface.* – ArcGIS

medium.com

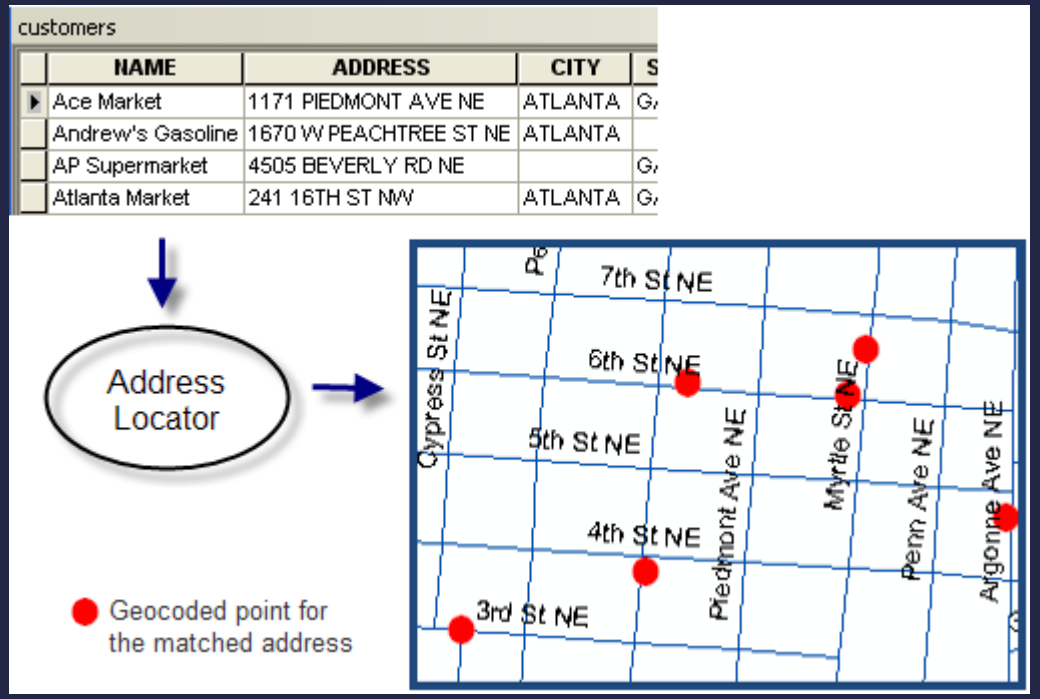
TYPES OF LOCATIONS

- Coordinates (x, y)
- Street addresses
- Postal codes
- Place names
- Route location



COMPONENTS OF GEOCODING

- Location Data
- Reference Data
- Address Locator
- Results



ESRI

CONVERTING (X,Y) LOCATIONS

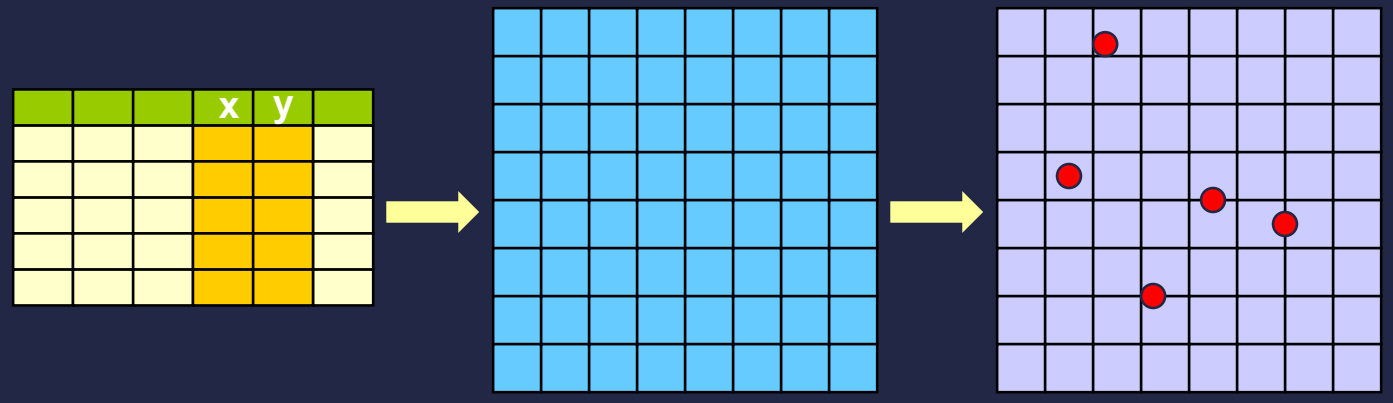


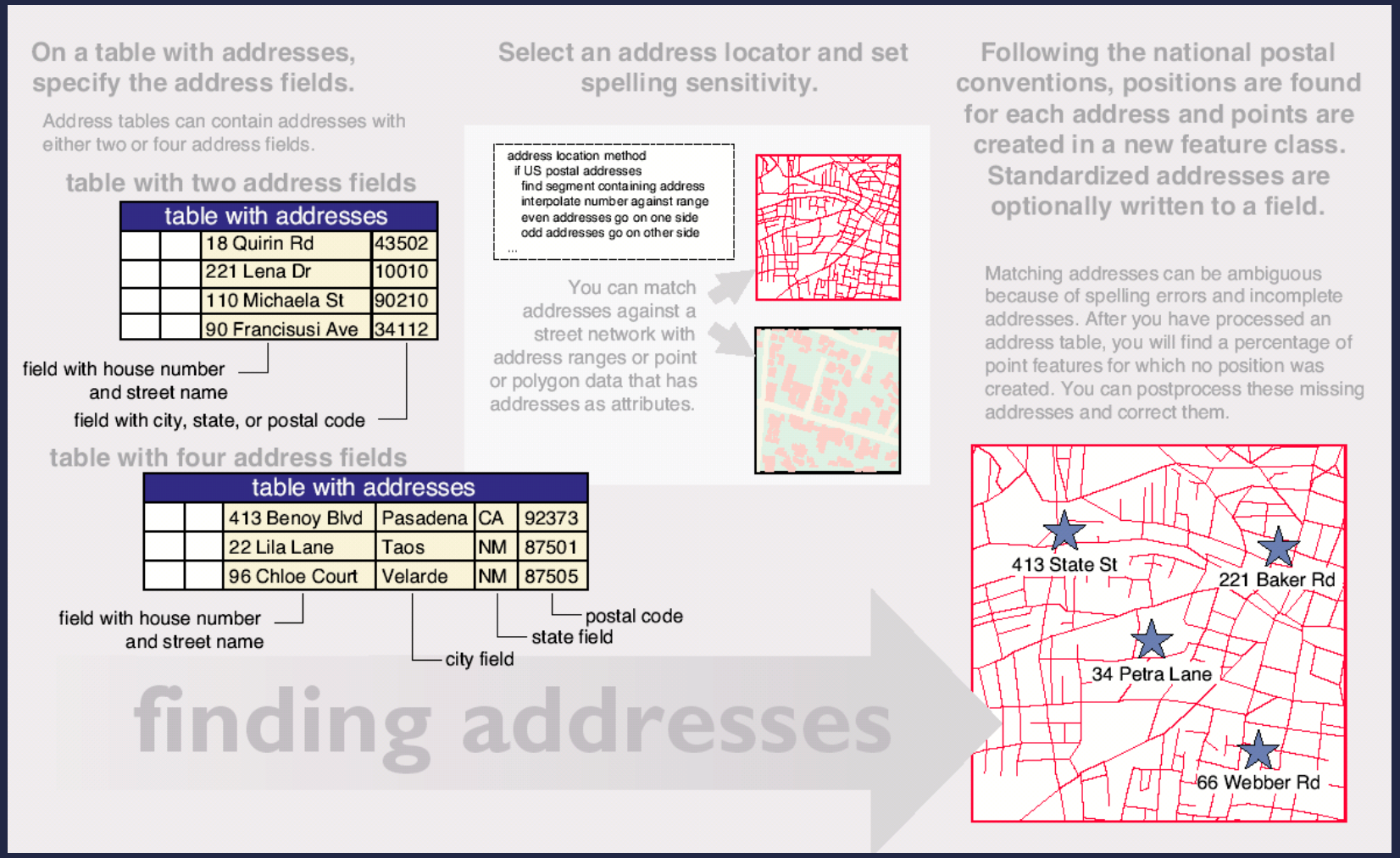
Table with fields for x and y coordinates

Spatial reference (coordinate system)

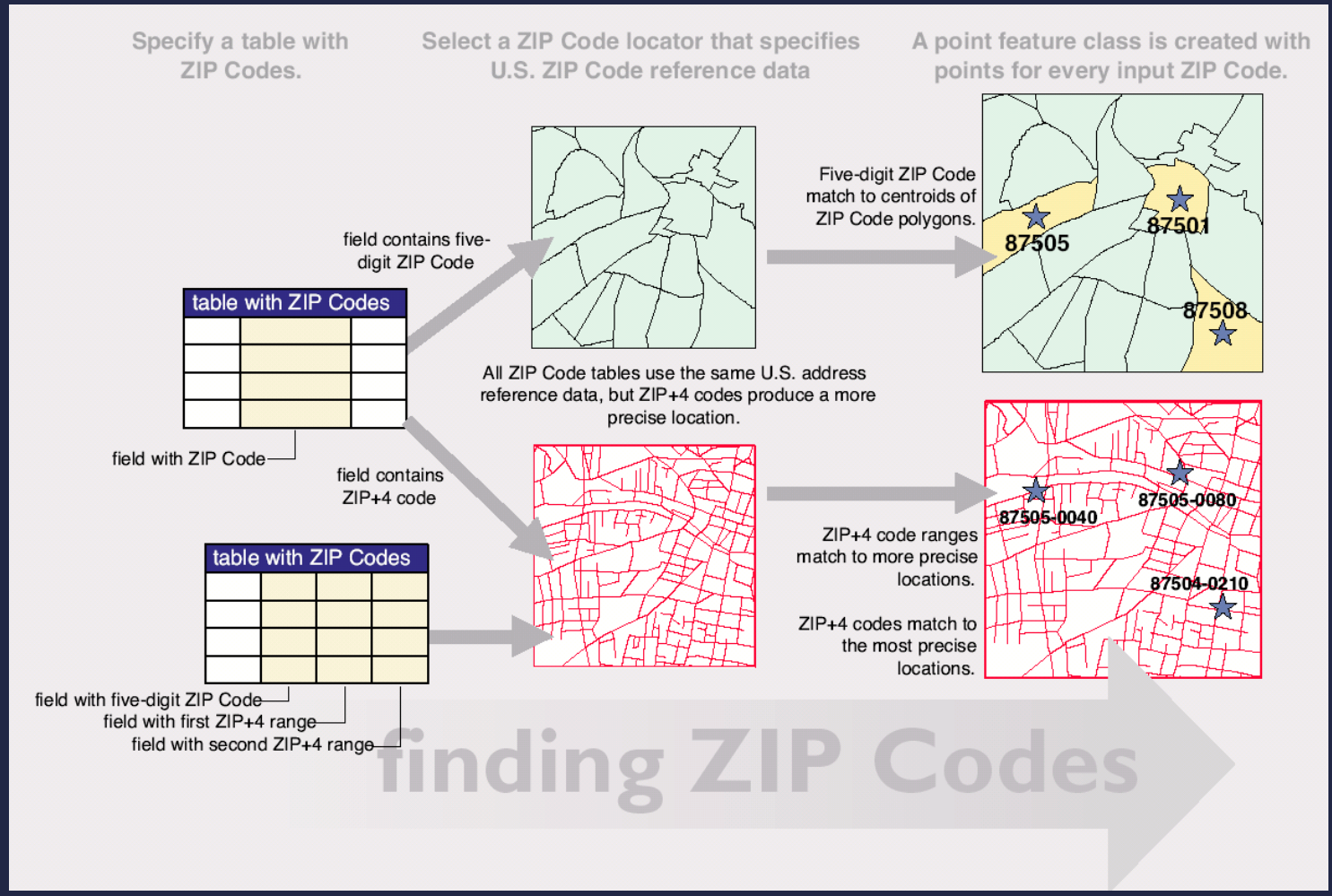
New data layer with a point for each input (x,y) value

This is how we map GPS data!

CONVERTING ADDRESSES

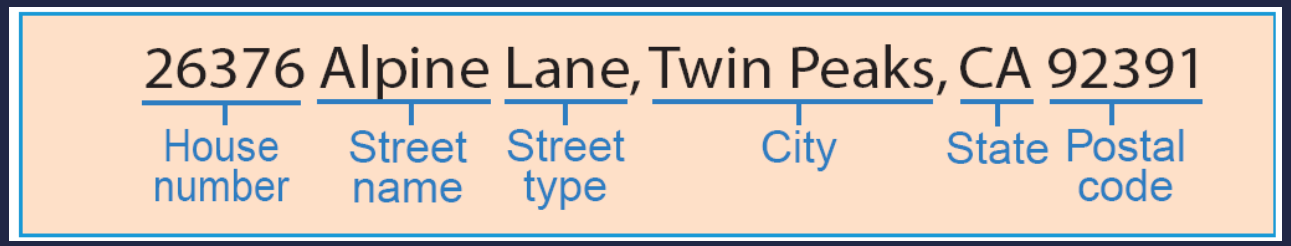


CONVERTING ZIP CODES



ADDRESS ELEMENTS

- Address (house or building) number
- Prefix direction
- Prefix type
- Street name
- Street type
- Suffix direction
- Zone
- Postal Code
- …



ADDRESS FORMATS

76 - 20 34th Ave, Jackson Heights, Queens, NY 11372

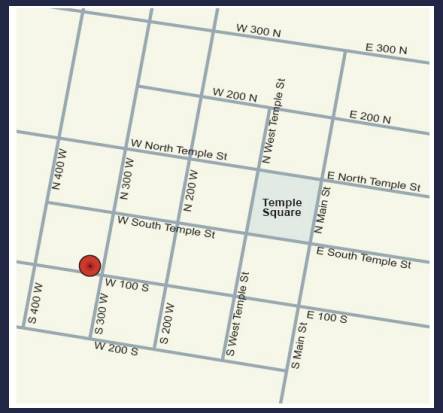
Cross street name | House number | Street name | Street type | City | Borough | State | Postal code

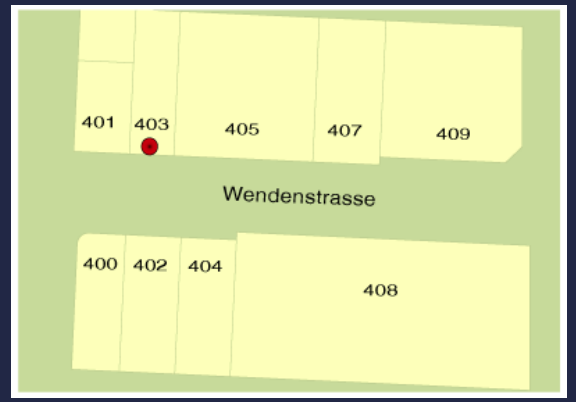
Wendenstrasse 403, 20537 Hamburg

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

305 W 100 S, Salt Lake City, UT 84119

House number | Prefix direction | Street name | Suffix direction | City | State | Postal code





GECODING WORKFLOW

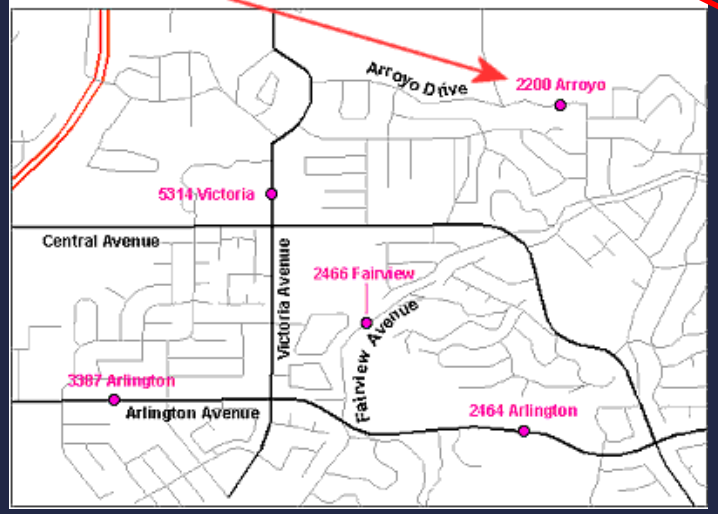
- Build or obtain reference data
- Determine address locator style
- Build an address locator
- Locate addresses
- Publish or maintain the address locator

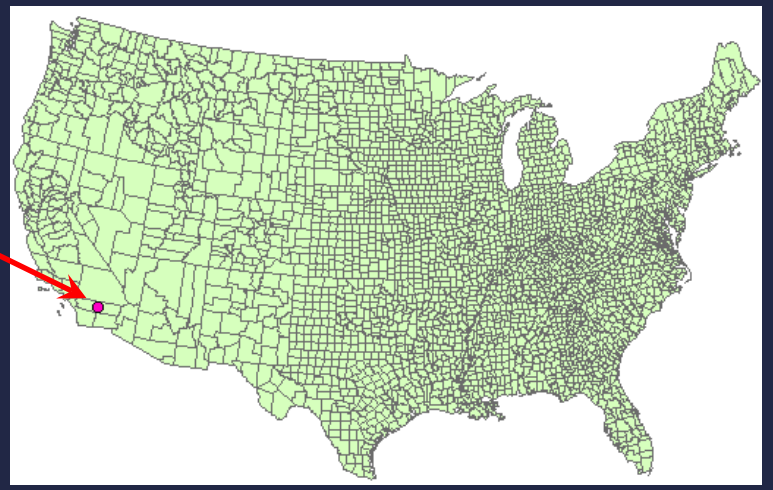
REFERENCE DATA

- Topologically Integrated Geographic Encoding and Referencing system (TIGER)
- Other Street data – Local entities or Commercial
- Parcels
- Zip code
- …
- Note: address locator needs to rebuild once reference data are updated.

USING STREET REFERENCE FILE

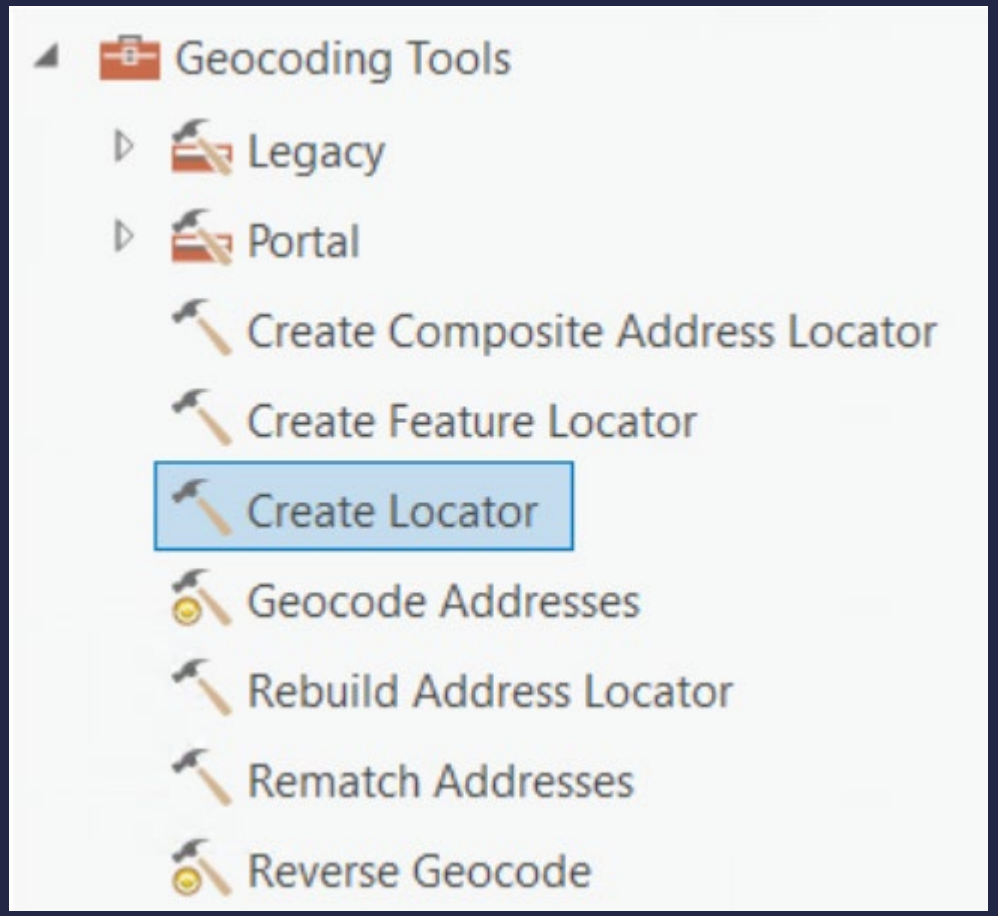
Address	City	State	Zip
3387 Arlington Avenue	Riverside	California	92506
5314 Victoria Avenue	Riverside	California	92506
2466 Fairview Avenue	Riverside	California	92506
2464 Arlington Avenue	Riverside	California	92506
2200 Arroyo Drive	Riverside	California	92506



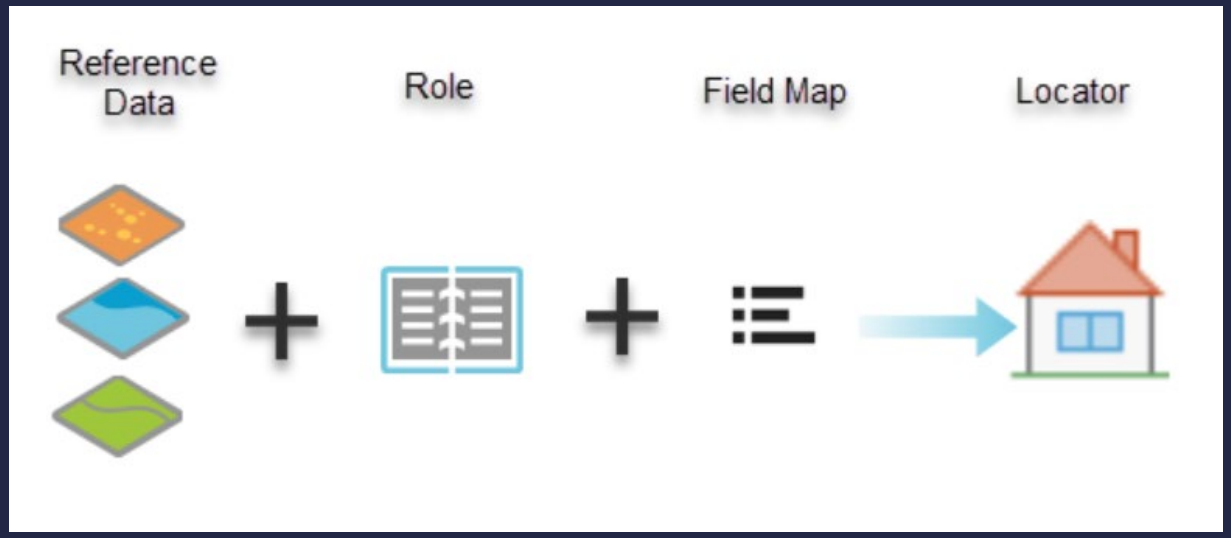


ADDRESS LOCATOR

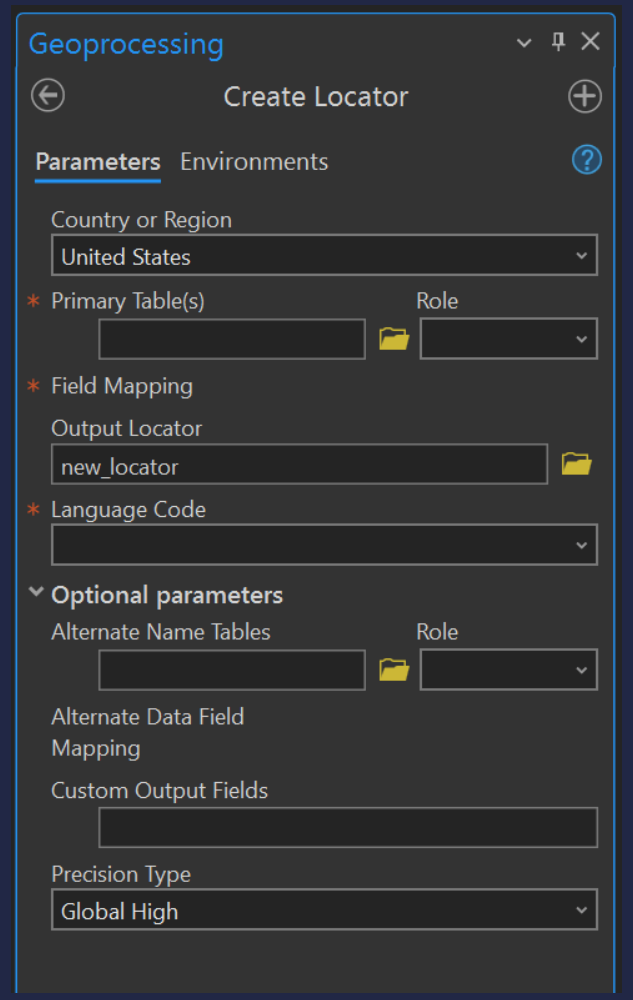
- It's an essential geocoding tool.
- It contains a snapshot of the reference data that is used for geocoding
- Address locator file is created with following extensions
  - *.loc, *.loz – stores reference information
- Can be shared as a Locator Package or as a Service



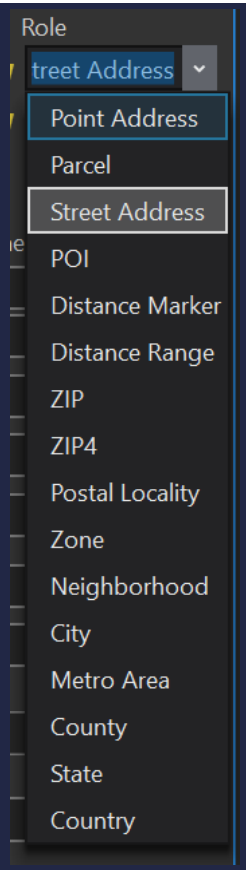
CREATE ADDRESS LOCATOR



pro.arcgis.com



ADDRESS LOCATOR ROLES



https://pro.arcgis.com/en/pro-app/latest/help/data/geocoding/introduction-to-locator-roles.htm

Role	Typical reference dataset geometry	Typical reference dataset representation	Address search parameters	Examples	Applications
Point Address	Points or polygons	Each feature represents an address	Address elements in a single field	71 Cherry Ln	Finding parcels, buildings, or address points
Parcel	Points or polygons	Each feature represents a parcel	All address elements in a single field	320 Madison St. 1760820300	Finding parcels or address points
Street Address	Lines	Each feature has the address range for both sides of the street segment	All address elements in a single field	2 Summit Rd. N5200 County Rd PP 115-19 Post St.	Finding a house on a specific side of the street or street intersections
POI	Points or polygons	Each feature represents a particular geographic place name or landmark	All place name elements in a single field	Leeds Castle, England Sapporo, Japan	Finding geographic place names or landmarks in an area or the world
Distance Marker	Points or polygons	Each feature represents sequentially nymbered markers placed along roads at regular intervals	Distance marker in a single field	Mile 25 I-25N, Raton, NM	Finding a distance marker sign on highway
Distance Range	Lines	Each feature represents the distance marker range for each line segment.	Distance marker range in a single field	Carr 682 KM 4.4, Barceloneta, 00617	Finding an approximate distance along a highway
Postal	Points or polygons	Each feature represents a single postal code region or centroid	Postal code in a single field	96822	Finding a specific postal code location
Postal Extension	Points	Each feature represents a single postal extension centroid	Five-digit ZIP Codes and four-digit extension in separate field	63703-0078	Finding a specific postal extension location
Postal Locality	Points or polygons	Each feature represents the union of postal code and city in a postal code boundary or centroid.	Postal code and city in a single field	7132 Frauenkirchen	Finding a specific locality
Administrative Areas	Points or polygons	Each feature represents a particular administrative area such as city, neighborhood, metro area, territory, region, and so on.	Administrative area name in a single field	British Columbia North Park, San Diego	Finding a specific administrative zone

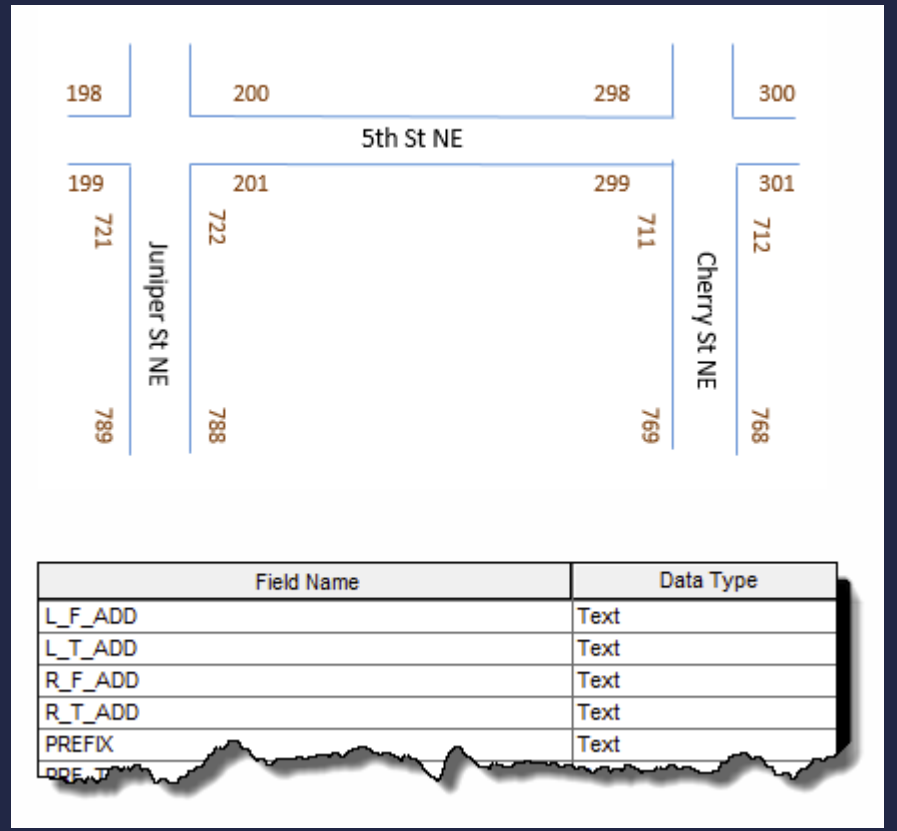
POINT ADDRESS

- Reference data can be polygon or point geometry
- Each feature in the reference data corresponds to a single address such as a parcel or a building
- Data must have individual fields that contain a street number and street name information.

ESRI

STREET ADDRESS

- Reference data must contain dual ranges for roads
- Can provide a range of house number values
- Can determine side of the road segment the address is located
- Supports normal block ranges, alphanumeric addresses with grid zone, or hyphenated addresses containing cross-street information in the house number

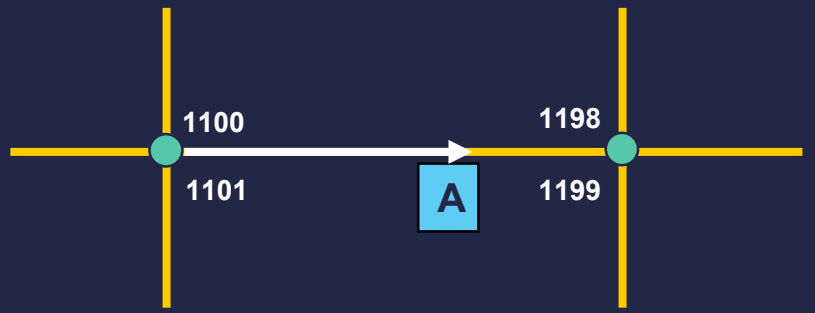


Field Name	Data Type
L_F_ADD	Text
L_T_ADD	Text
R_F_ADD	Text
R_T_ADD	Text
PREFIX	Text

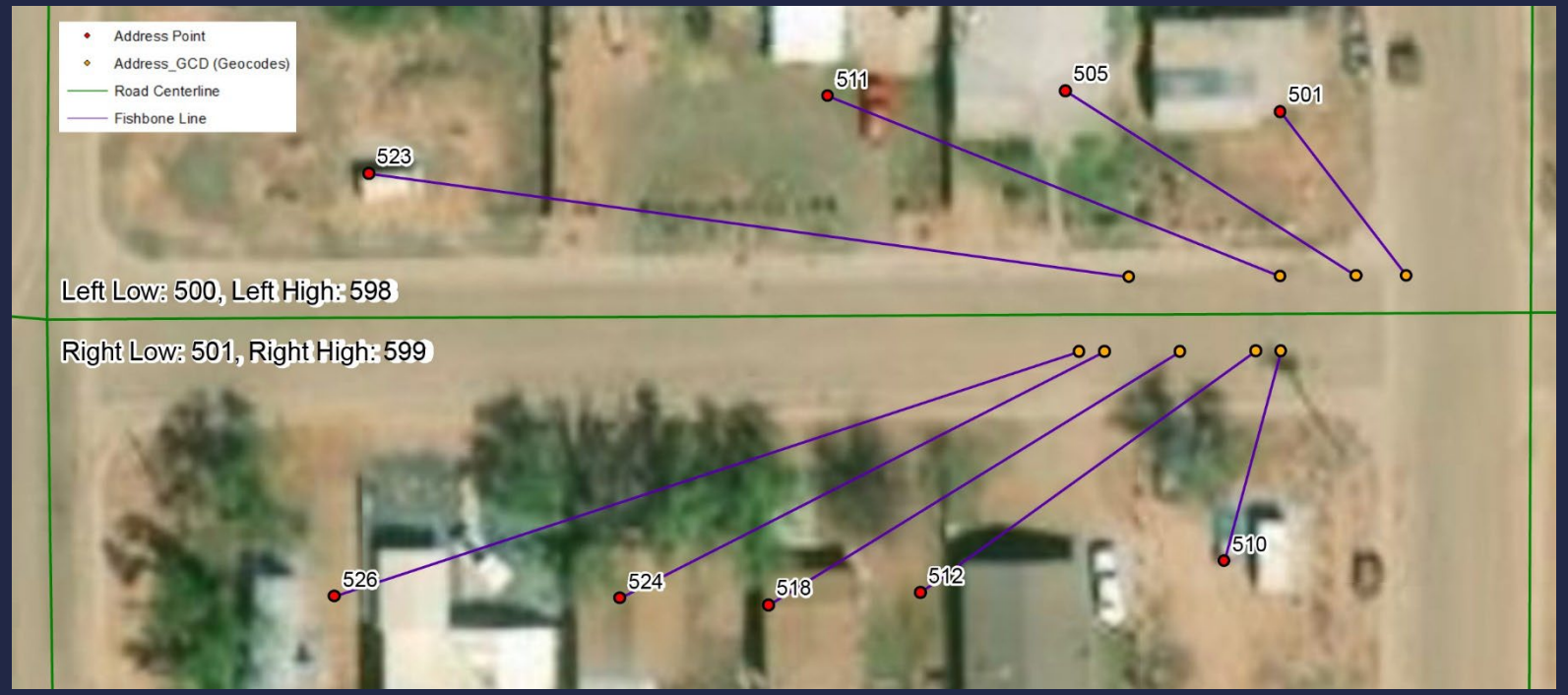
ESRI

REFERENCE DATA – ROADS

- A road network is represented by line segments
- Each intersection is represented by a node
- The coordinates of an address are interpolated linearly between the two nodes of a segment
- Example: 1167 Main St.

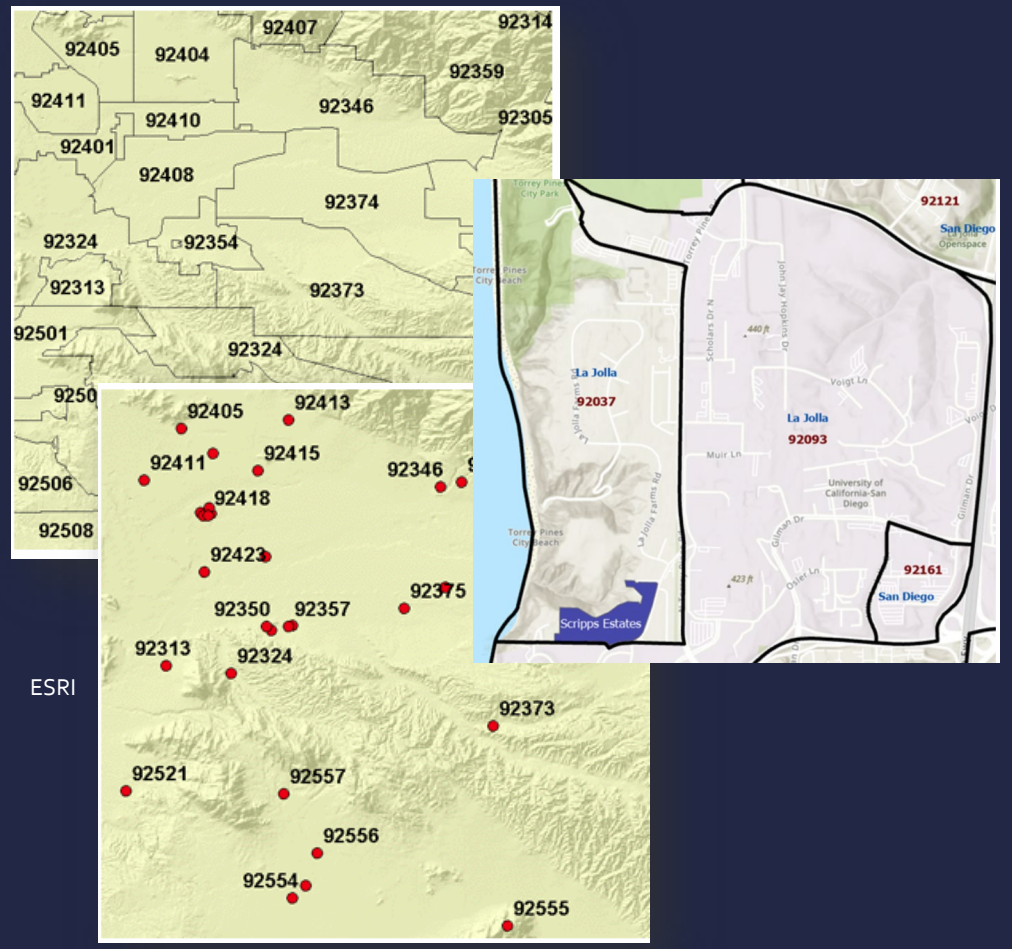


ADDRESS INTERPOLATION



POSTAL

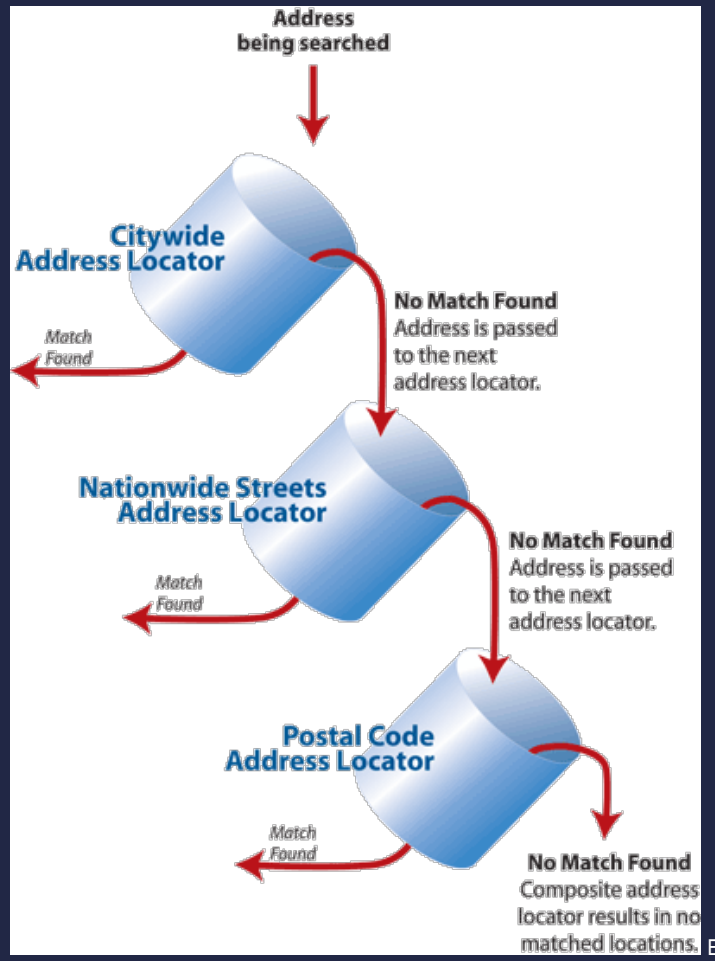
- Reference data can be point or polygon geometry
- Each feature must contain a field that specifies the 5-digit postal code
- For polygon data the geocoded points will be at the centroid of the zip code polygon
- Postal Extension enables finding a specific postal extension location
- Postal locality enables finding a specific locality



ESRI

COMPOSITE ADDRESS LOCATORS

- Consists of two or more individual address locators
- Data are automatically matched against each of the individual locators
- Selection criteria can be set for the order of the locators



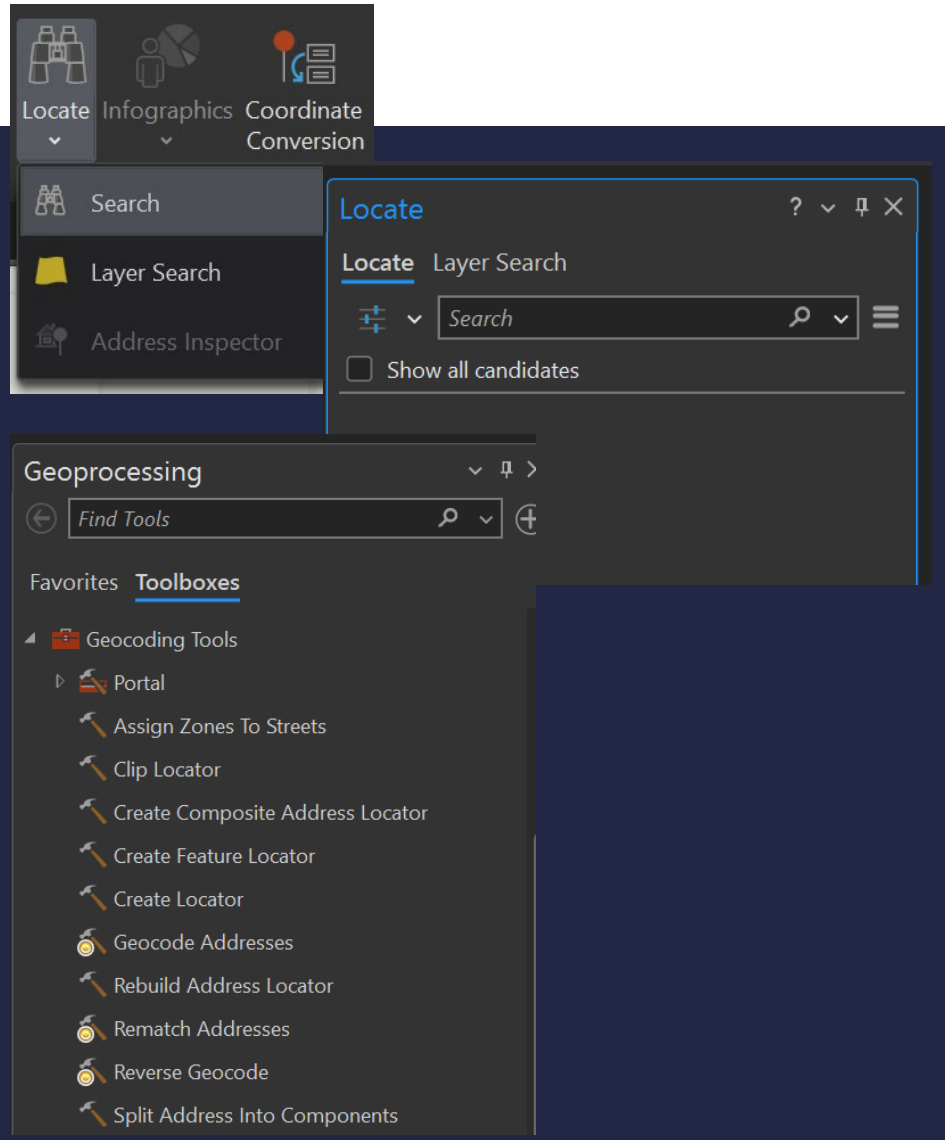
ESRI

MULTIROLE LOCATORS

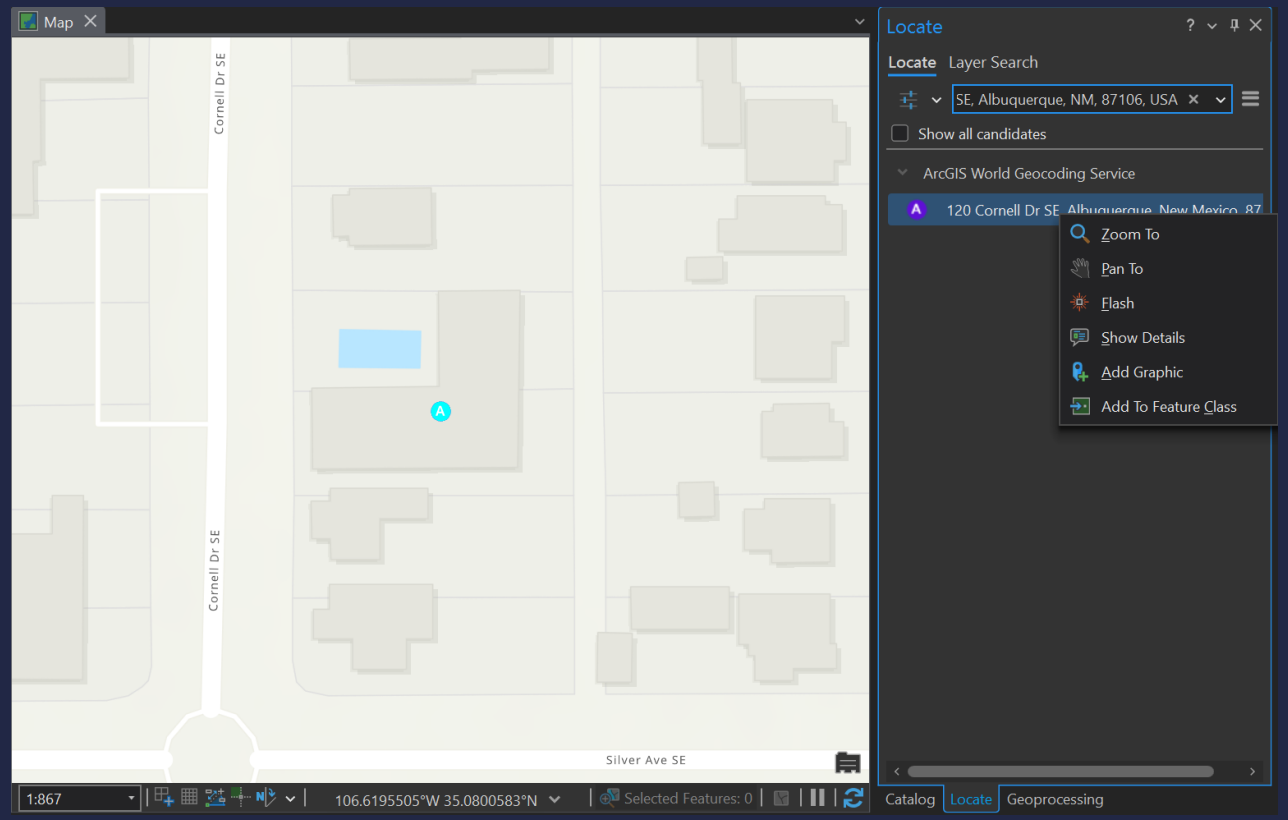
- Allows to combine multiple reference data layers and roles into a single locator to search for multiple types of locations at once
- Ex: rooftop locations, interpolated street locations, points of interest, etc.
- Reduces redundant information and candidates, beneficial for performance and reducing data size
- Minimize duplicate geocode results with multirole locators

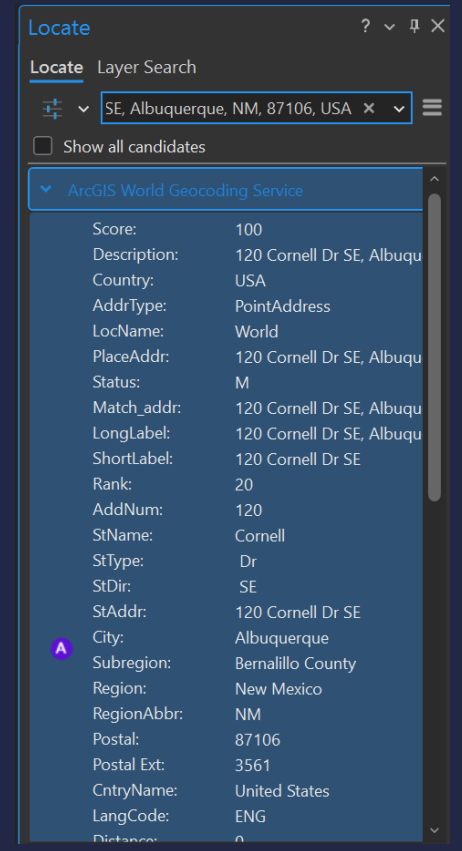
GEOCODING IN ARCGIS

- Create an Address Locator in ArcCatalog
- Locate Tool
- Geocoding Process
  - Add Address Locator to ArcMap
  - Select the Address Table to be geocoded
  - Save results as a feature class
  - Review results, including unmatched addresses
  - Decide on rematching strategy

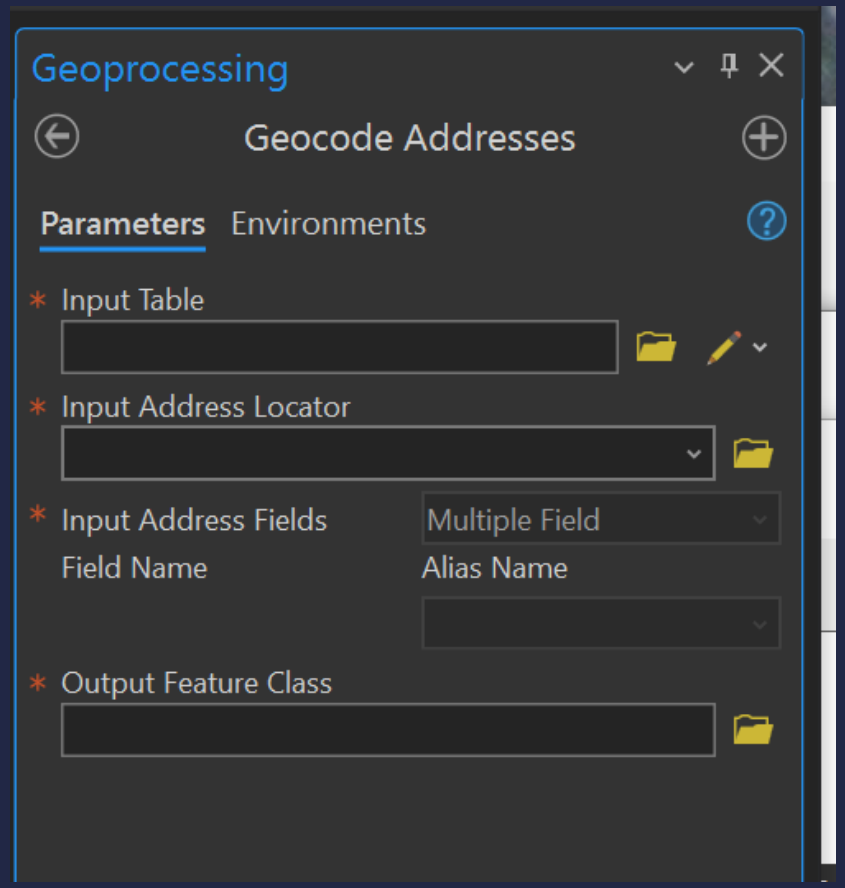


GEOCODING USING THE LOCATE TOOL





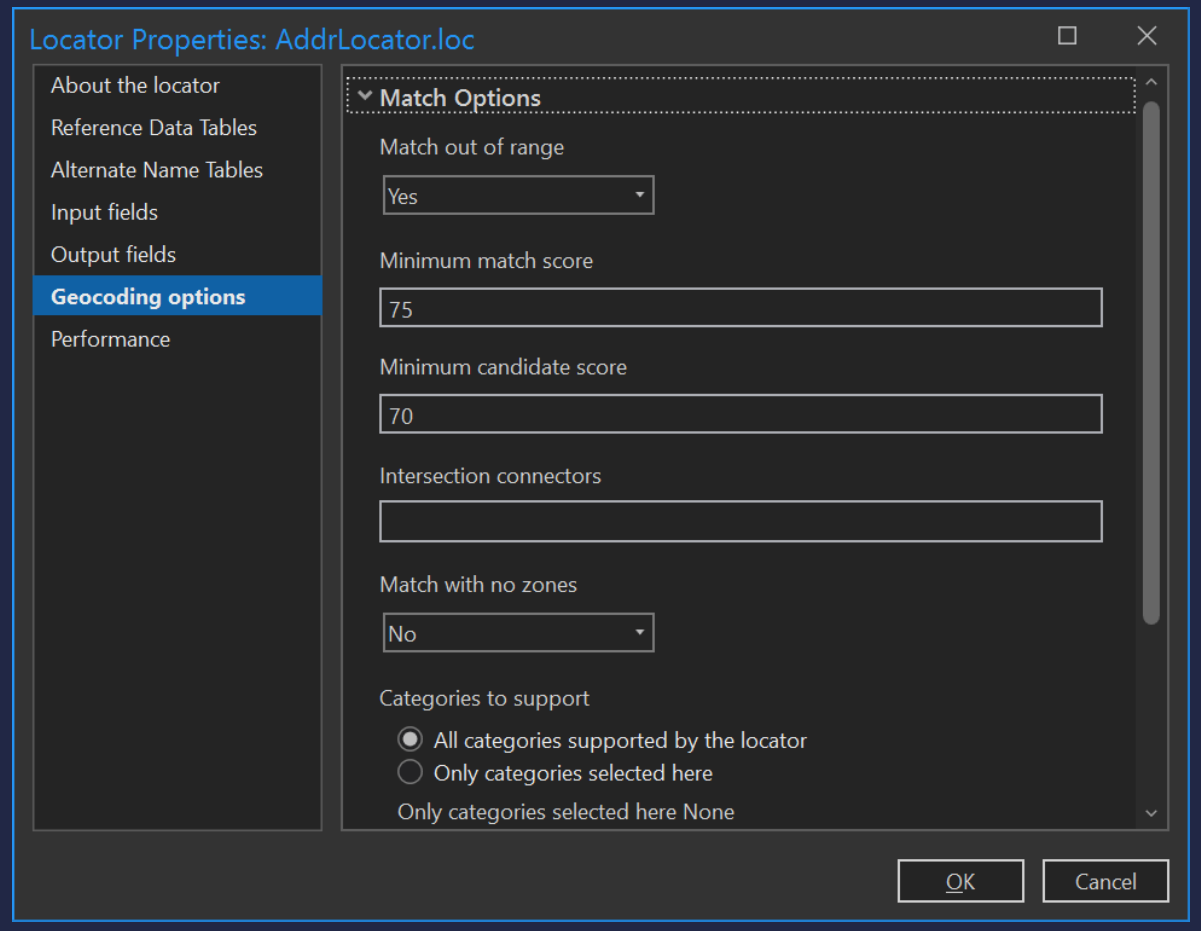
GEOCODE ADDRESS TOOL



OUTPUT DATA

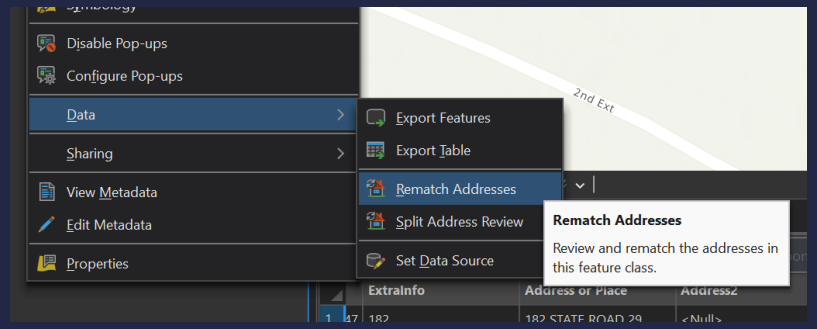
- In addition to the user provided data, following are a few columns that are created by the software based on the locator style…
- **Status**—A code indicating whether the address was matched. Possible values include the following:
  - **M**—Matched. **U**—Unmatched. **T**—Tied (more than one match but different locations).
- **Score**—The match score of the candidate to which the address was matched. The score can be in a range of 0 to 100, where 100 indicates that the candidate is a perfect match.
- **Match_type**—A code showing how an address was matched. You can group the results based on this attribute to show how the addresses were matched or use the grouping to select records for rematching. The codes are as follows:
  - **A**—Automatically matched or rematched.
  - **M**—Manually matched or unmatched.
  - **PP**—Pick by Point. The address was matched to the click point using the **Pick Address from Map** tool in the **Interactive Rematch** dialog box in ArcMap.
- **Match_addr**—The address where the matched location actually resides based on the information of the matched candidate. For example, an input address of 123 Main St N is matched to a candidate with the suffix direction NW and all other components matched correctly. The **Match_addr** field will contain 123 Main St NW as the actual address that was matched.
- **Addr_type**—The type of address that was geocoded. This attribute indicates to what kind of feature the address was matched. You can study the accuracy of the matched addresses and pattern of the matches based on the values.
  - PointAddress, StreetAddress, Parcel,…

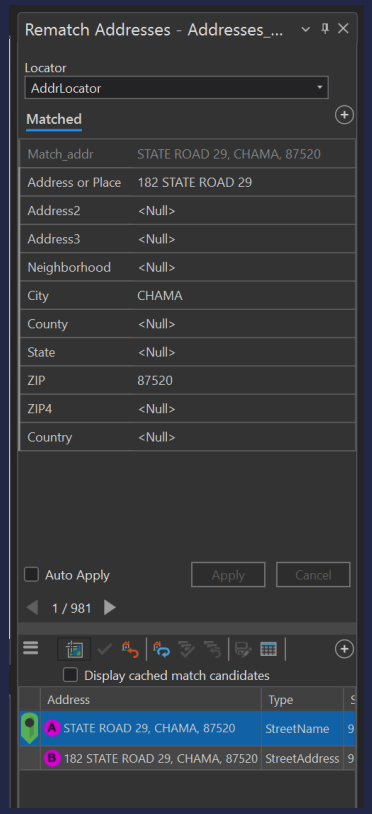
ADDRESS LOCATOR PROPERTIES

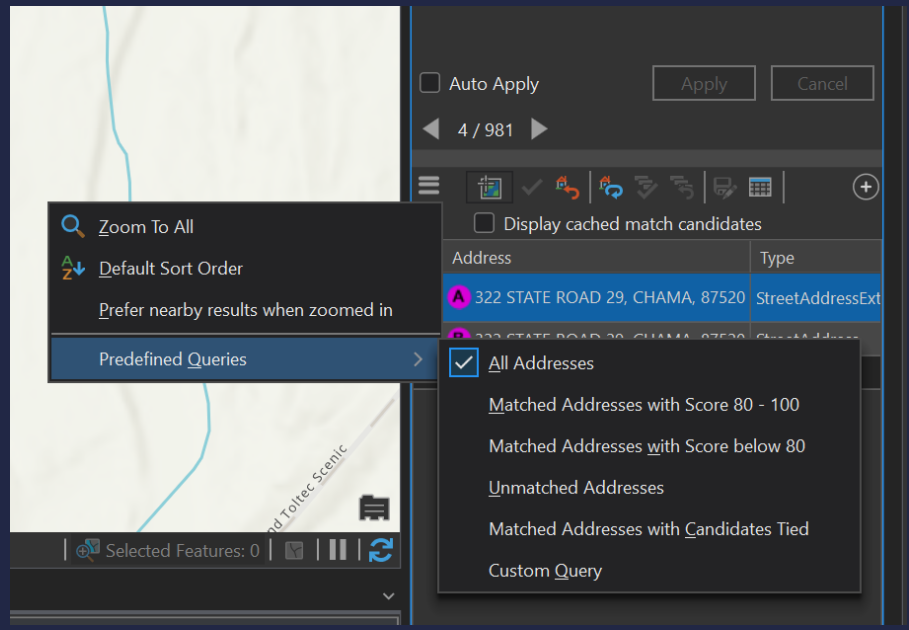


Modify parameters

REMATCH PROCESS

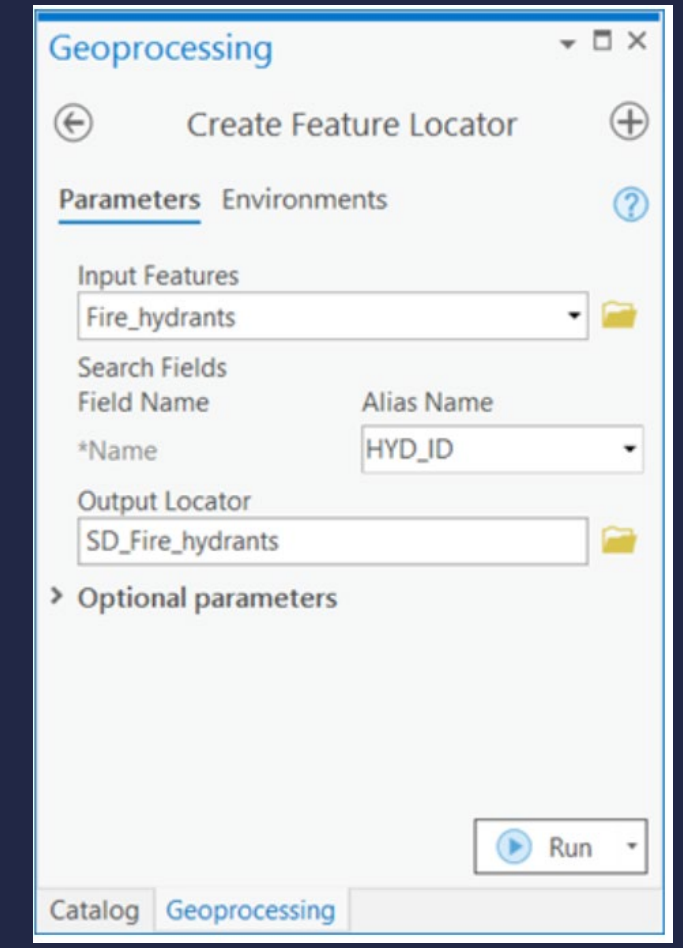






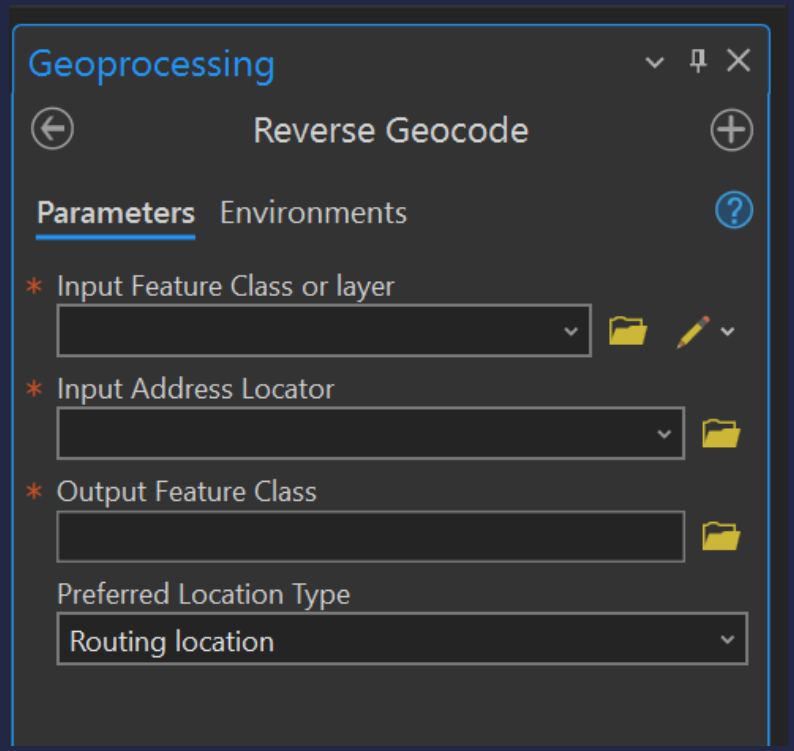
CUSTOM DATA SEARCH – FEATURE LOCATOR

- Feature Locator – Search by features
  - Water meters, cell towers, oil wells, etc.

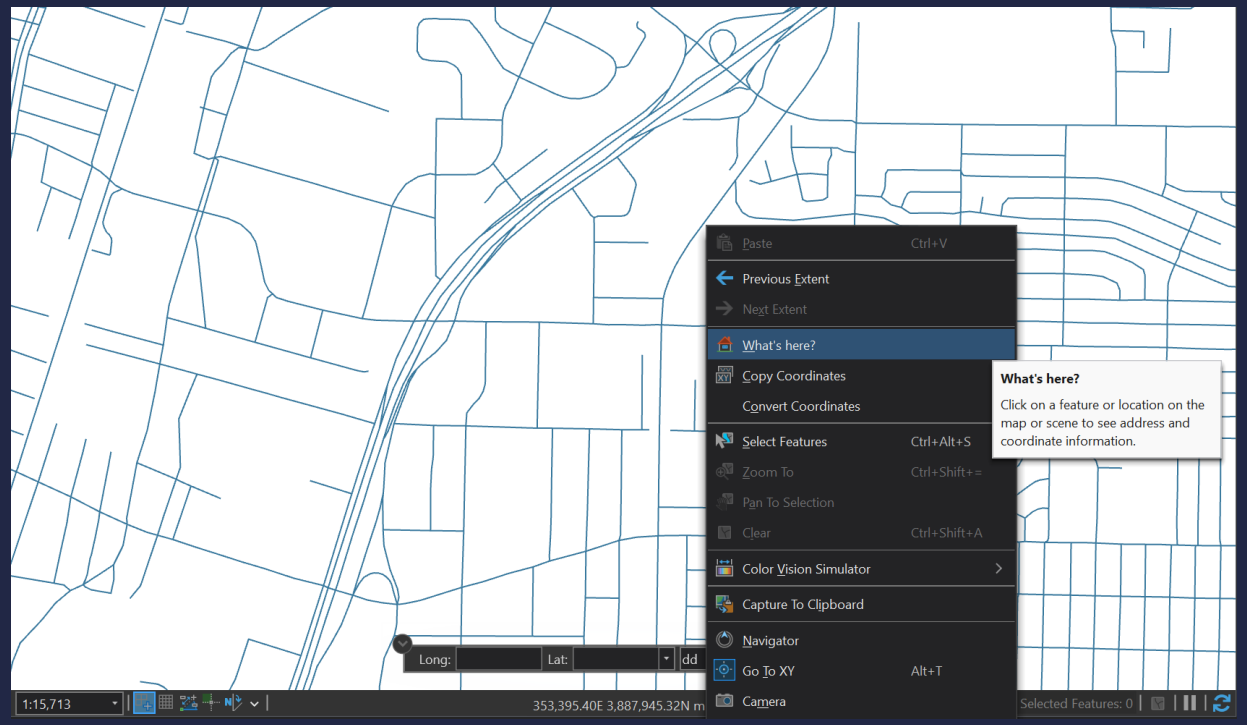


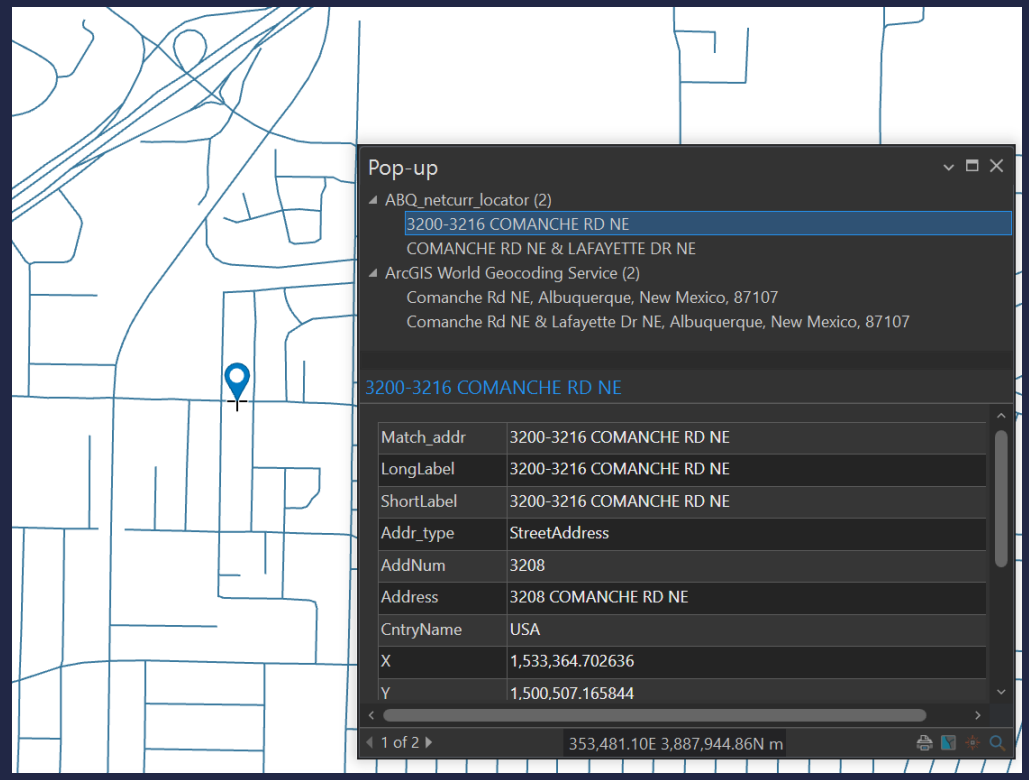
REVERSE GEOCODING

- Searches for the nearest address, place, or intersection for the point location
- Input feature class must contain point shapes with valid coordinates



REVERSE GEOCODING

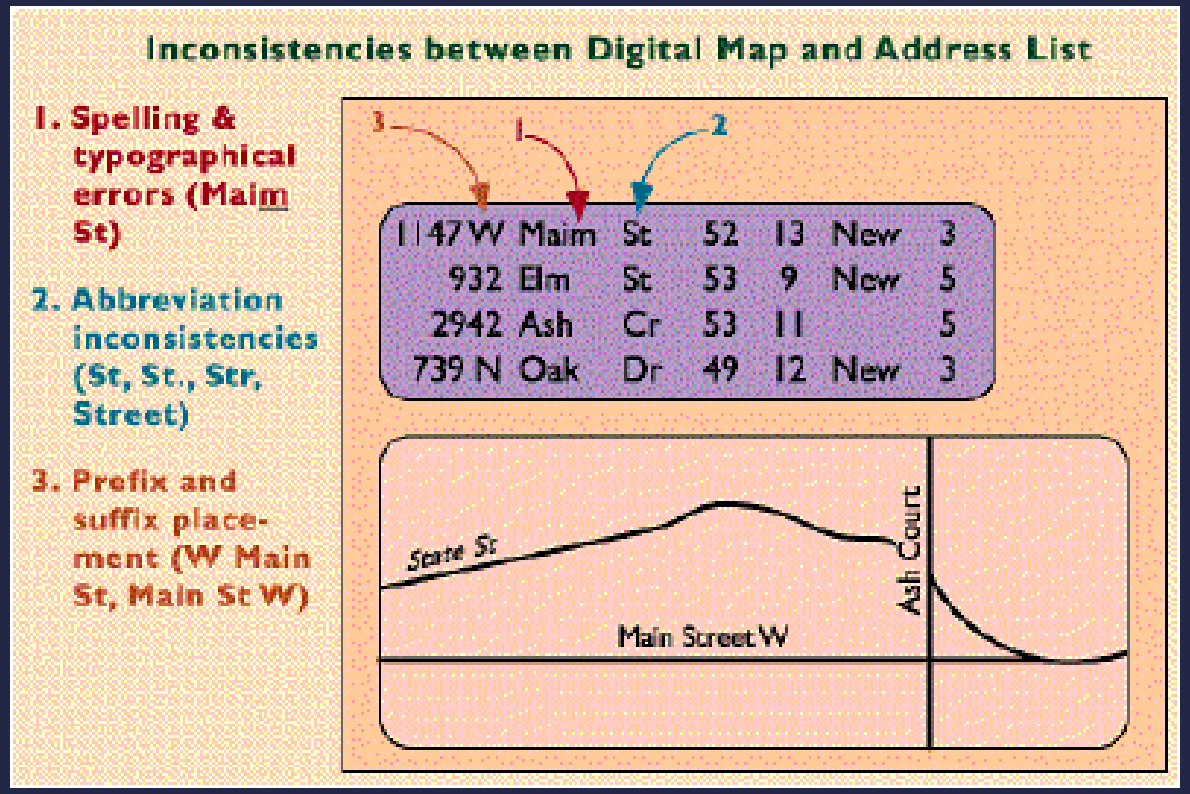


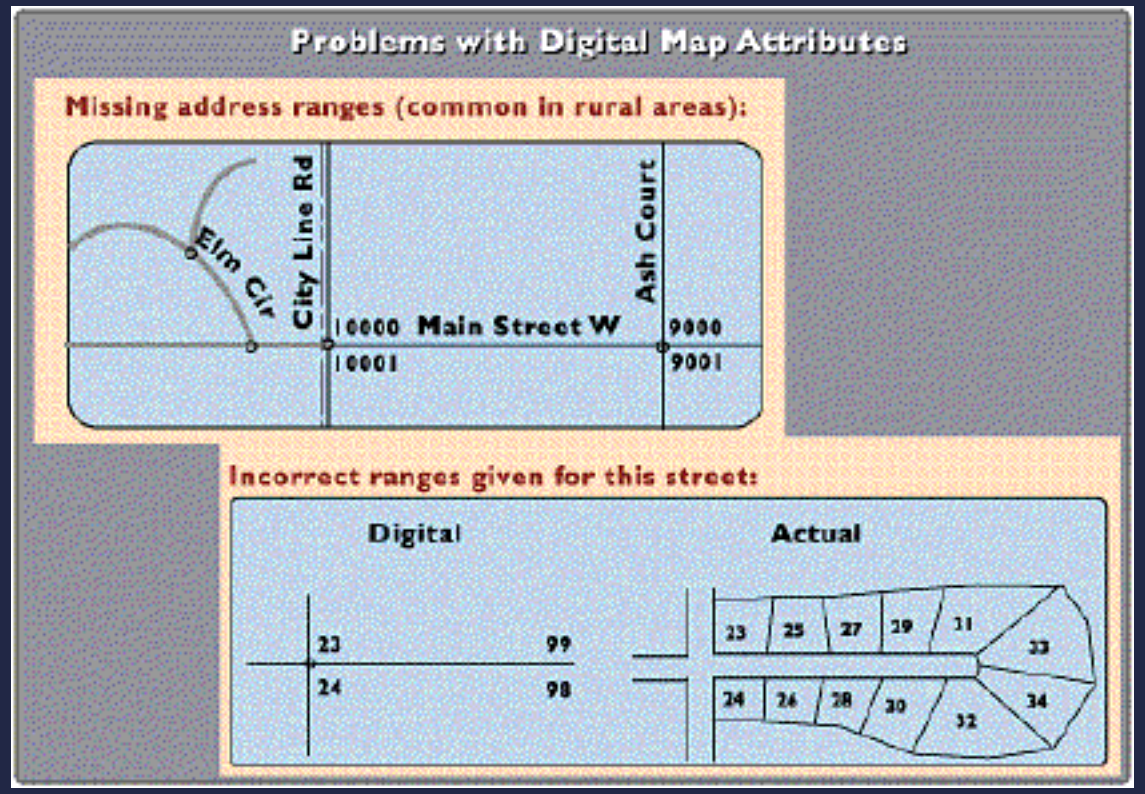


GEOCODING PROBLEMS

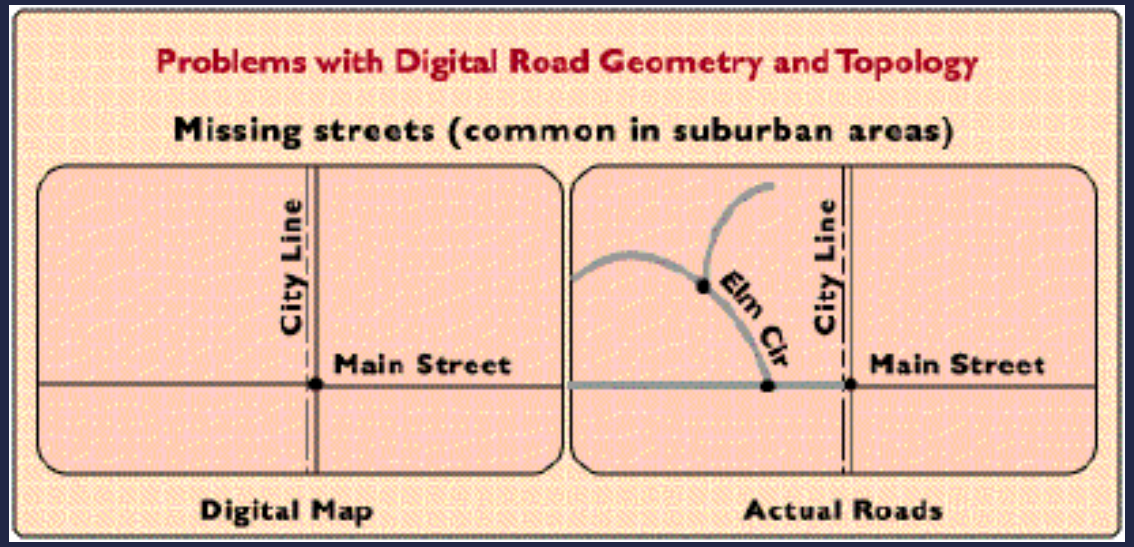
- Address-matching is rarely a fully automated process.
- An address may not address-match because of inaccuracies or inconsistencies in the digital street map or in the file of addresses to be matched.
- Address-matching rates (the proportion of addresses that are correctly matched) will increase if efforts are made to increase the quality of the digital street map, the address file, or both.

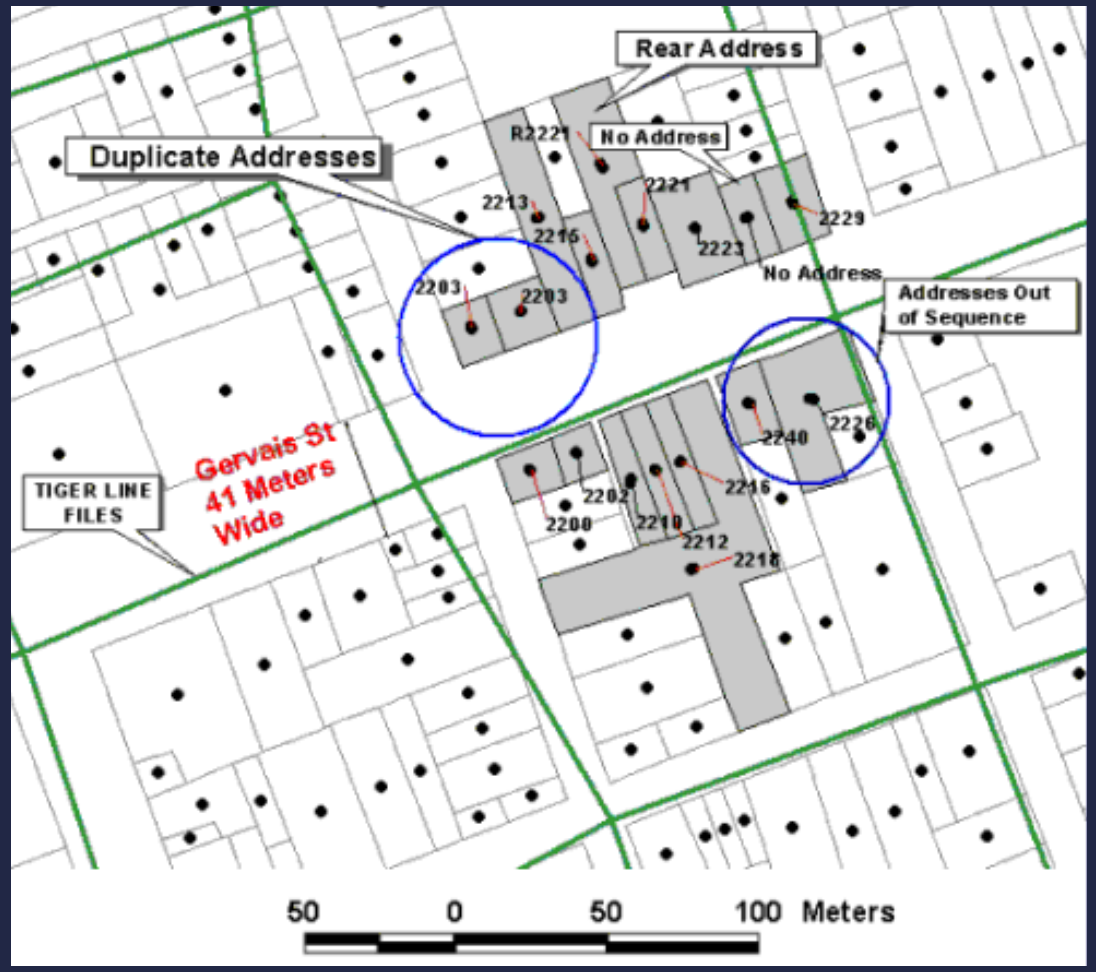
PROBLEMS – CONTINUED





PROBLEMS – CONTINUED





PRACTICE

- Geocode a table of addresses
  - https://pro.arcgis.com/en/pro-app/latest/help/data/geocoding/tutorial-geocode-a-table-of-addresses.htm
- Rematch addresses from a geocoded feature class
  - https://pro.arcgis.com/en/pro-app/latest/help/data/geocoding/tutorial-rematch-addresses-from-a-geocoded-feature-class.htm
- Create a Locator
  - https://pro.arcgis.com/en/pro-app/latest/help/data/geocoding/tutorial-create-a-locator.htm