

GIS Institute Geocoding Lab Exercise
By Jeff Blossom: jblossom@cga.harvard.edu
January 9, 2019

Introduction

The ability to assign specific geographic locations to textual information (the process known as geocoding) is available to anyone with a computer and internet access. The relative ease of geocoding and resulting accuracy can vary widely depending on a number of factors. What is the nature of the data? How 'clean' is it and what format is it in? What geocoding technique will be used? Determining a geocoding strategy that best suits a particular need is not always clear.

The process of geocoding begins with comparing data in text or tabular form to a reference data table in geographic format. The reference table is a dataset that has already been mapped, with established map coordinates. When matches between the input data and the reference data are found, the corresponding map coordinates are assigned from the reference data to the input features, thus geocoding them. A geocoding service is a program that allows for a user to input a batch of data contained in a table, search for matches as compared to a reference table, and output the result in a map or GIS layer format. The key to confidently geocoding data lies in understanding the reference table which the data is being matched to, how a match is found, and the resulting spatial accuracy.

Objective: The goal of this exercise is to give everyone practice performing geocoding using different methods, and to gain insight regarding how to successfully geocode.

Geocoding lab

In this exercise you will geocode single addresses and cities using Bing Maps and Google Maps. You will geocode lists of multiple addresses and cities using ArcGIS Online and ArcGIS Desktop (ArcMap). You will create GIS datasets, online maps, and desktop maps of the results. You will learn how to evaluate the accuracy of the data using base maps.

Lab Setup: From the GIS Institute Winter 2019 page, download and unzip the **(lab data)** file under the **Day 2 Geocoding** topic. This ZIP file contains a folder named *Geocoding_lab*. Unzip/Extract this folder onto your Desktop.

1) Geocode one address with Bing Maps

In an internet browser, go to <http://www.bing.com/maps/>

Type an address, such as the CGA's address "1737 Cambridge St. Cambridge, MA " into the

window at the top, and click the Search button: 

The map will zoom to this location, and the latitude, longitude values will be displayed at the bottom of the information window (see below). You can click Copy, and paste into a word or other document.



United States · MA · Middlesex Co.
Square
42.375583, -71.113257

To find the latitude/longitude of any point on the map, just right click anywhere on the map. The two numbers listed at the bottom of the pop-up window are the latitude, longitude of that location. *This could be helpful if you want to find the location of specific points, (for example a*

park bench) instead of just the location the geocoder returns. Right click somewhere on the Bing Map and copy the coordinates. Then paste these coordinates into the search window of Google Maps at <http://maps.google.com>. The Google map will zoom to this same latitude/longitude coordinate. To view the latitude/longitude location of an address on Google Maps, you must right click on the location, and choose “What’s here?”. Try searching for other places, like city names, water bodies, or famous locations (such as the Eiffel Tower). Google Maps, Bing Maps and most online geocoders will be able to locate “non-address” data as well.

2) Geocode multiple cities with the World Geocoder in ArcMap

In this exercise you will geocode 50 worldwide cities. In Windows File Explorer, browse to the Geocoding_lab folder and double click the World_Livable_Cities.csv file. This will open it in Excel. Notice that this contains a list of city, country locations, and a “livability” ranking and index. This list was generated by Business Week magazine, and downloaded in 2012.

2.1 Open a new blank map in ArcMap. Click **File > Add Data > Geocoding > Geocode Addresses**. Choose the ArcGIS World Geocoding Service and click OK. You will be prompted to enter your ArcGIS Online account information.

2.2 If you have a Harvard Key, click on the ENTERPRISE LOGIN button on the lower right. If not, use the Username **GIS_Institute** with password **ARCGIS2019** and click SIGN IN. Then skip to step 2.5.

2.3 For the ENTERPRISE LOGIN option, enter “harvard-cga” under where it says “Enter your ArcGIS organization's URL below.” And click Continue.

2.4 Click the HARVARD UNIVERSITY button. Enter your Harvard Key information and you’ll be logged into the Harvard account.

2.5 In the Geocode Addresses dialog box, click the browse button to specify your address table as the **World_Livable_Cities.csv**. Choose **Single Field**, and specify the Single Line Input as **City_Country**. Specify the Output shapefile or feature class to be **World_Livable_Cities.shp**, saving it in your Geocoding lab folder. Your screen should look like the one on the right →

2.6 Click **OK**, and the list of cities will be geocoded.

2.7 Inspect the geocoded results.

In ArcMap, click **File > Add Data > Add Basemap**, and choose the **Streets** basemap, and choose Add. Review the results by highlighting individual addressees in the attribute table, and seeing how they compare to street names listed on the map,

Geocode Addresses: World

Address table:
World_Livable_Cities.csv

Address Input Fields
☒ Single Field ☐ Multiple Fields
Single Line Input: City_Country

Output
☒ Create static snapshot of table inside new feature class
☐ Create dynamic feature class related to table
Output shapefile or feature class:
C:\Geocoding_lab\World_Livable_Cities.shp


Config Keyword:

Advanced Geometry Options...

Geocoding Options...

About geocoding a table of addresses

OK Cancel

or by selecting a point using the Identify tool:  Do the address locations match the correct cities as shown on the Streets basemap?

2.8 Refine the geocoded results.

In the table of contents, right click **Geocoding Result: ArcGIS_Geocoding** and choose **Data > Review/Rematch Addresses**. Note the Status (M=matched, T=Tied, U=Unmatched) and Score (a value between 0 and 100 that represents how closely the text of the input address matched an address in the reference dataset), X (longitude) and Y (latitude) fields.

Input locations can be changed by selecting the City_Country from the list, entering different information in the Single Line Input field, clicking Search, highlighting one of the search results, and clicking Match.

3) Geocode multiple addresses with the World Geocoder in ArcMap

Now you'll geocode a list of addresses in Atlanta, GA. These are locations of service stations and restaurants pulled from the internet, with fictional sales volume information added.ns.xls spreadsheet, within your Geocoding_lab folder.

3.1 Open the Customer_Locations.xls spreadsheet in Excel to inspect the data. Note that the address information is formatted two different ways: with a Full_Address field, and also with individual fields for Street, City, and State. These are the two standard formats for address geocoding - some geocoders require one format or the other.

3.2 Click **File > Add Data > Geocoding > Geocode Addresses**. Specify your Customer_Locations.xls file as the Address Table, Single Field, and choose Full_Address as the single line input. Click OK to run the geocoder. Inspect and refine the results as you did with the World Livable Cities.

4) Geocode to an existing dataset.

So far you've geocoded addresses and cities into point locations on the map. Often you may want to geocode data to non-point features, such as administrative areas. In this step, you'll geocode a list of US ski resort counts per state to an existing States shapefile. This list of ski resorts is found in the Ski_Areas_for_mapping.xls spreadsheet. Open it in Excel to examine the contents.

4.1 In ArcMap, **Add** the States shapefile using **File > Add Data > Add Data** and browsing to the Geocoding_Lab location on disk. Double click the **States** file to add it to the map.

4.2 **Right click** the **States** layer and choose **Open Attribute Table**. Note that the STATE_NAME field contains the name of the states. This field is what you'll use to geocode the Ski Areas to. Close the States attribute table.

Right click the **States** Layer and choose **Joins and Relates > Join**. In the Join Data dialog, specify to "Join attributes from a table" on the top menu choice. Then specify:

1. STATE_NAME
2. Browse to the Ski_Areas_for_mapping.xls file and choose the Ski_Areas workbook.
3. State

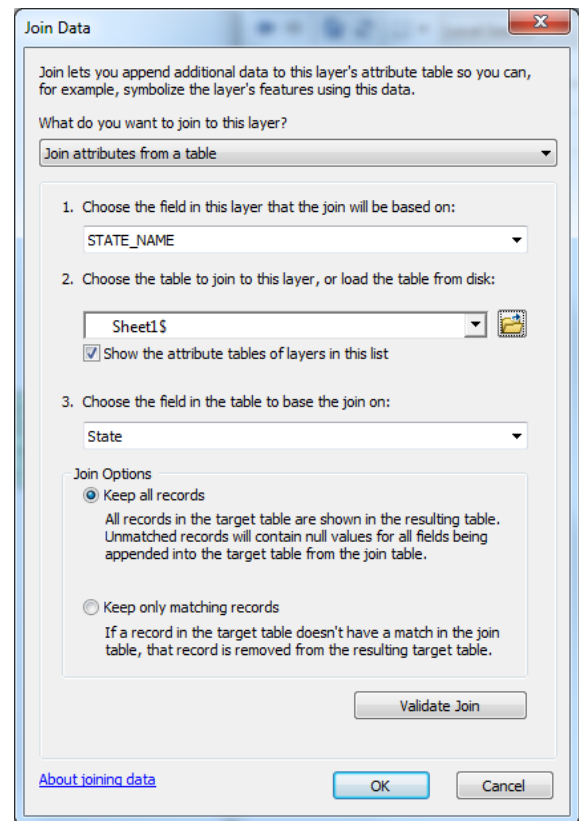
Your Join Data window should look like this →
Click **OK** to run the Join.

4.3 Right click the **States** layer and choose **Open Attribute Table**.

Scroll over to the far right of the table, and you should see the **Ski_areas** field at the very end.

Now you can symbolize your map by this statistic by **right clicking the States layer**, choosing **Properties**, **Symbology** tab, **Quantities** choice on the left, specifying the Value Field as **Ski_areas** and clicking **OK**. To save this geocoded data as a permanent shapefile, right click States and choose **Data > Export Data**. Specify a name and location for the new shapefile.

For tips on adding additional items to your map layout (like a legend, scalebar, title, etc.), and exporting it into an image or PDF format, refer to the **ArcMap Map Creation Layout Tips.doc** contained in the **Geocoding_lab** folder.

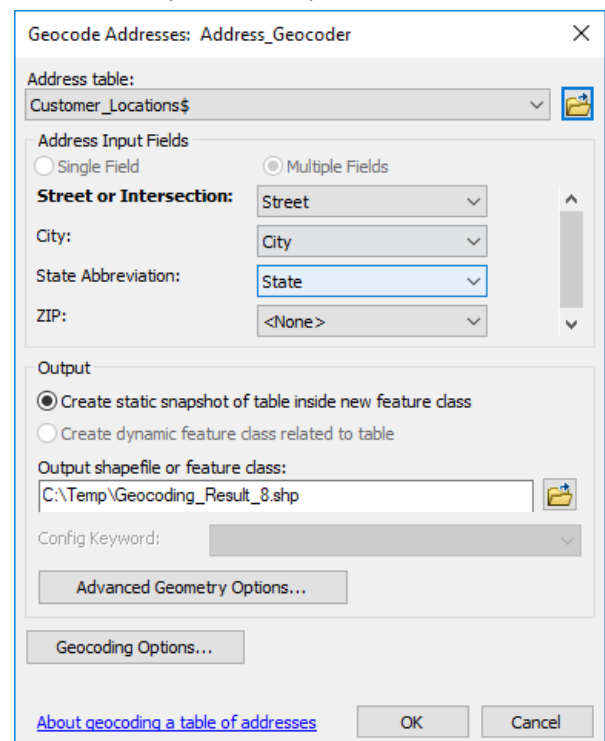


5) Geocode addresses using a local geocoder (optional)

The World geocoding service in ArcMap actually passes each address to Esri's server in California when searching for an address match, and if you have protected health information such as addresses, you may be restricted from sending this information online. Esri also makes local geocoding services that install on one's desktop. This local service has been clipped out for the Atlanta area, and is contained in your **Geocoding_lab** folder. The steps below explain how to use this local geocoder.

5.1) Click **File > Add Data > Geocoding > Geocode Addresses.** In the Choose an Address Locator to use window, click the **Add** button.

5.2) Browse to the **Geocoding_lab/Address_Locator folder.** Choose the **Address_Geocoder** and click **Add**. Now, highlight this geocoder, and click **OK**. This is a multiple field geocoder, so you'll have to specify the proper fields, as shown on the screenshot on the right. Click **OK** to geocode your addresses. Inspect the results. To see the streets that this local geocoder uses, add the **Atlanta_Streets.lyr** file from the **Geocoding_lab/GIS_Data** folder to ArcMap.



6) Geocode multiple cities with ArcGIS Online (optional)

You may want to geocode your information directly to an interactive webmap to share. If so, the directions below outline how to do this.

6.1 Go to <http://www.arcgis.com> and click Sign In in the upper right.

6.2 If you have a Harvard Key, click on the ENTERPRISE LOGIN button on the lower right. If not, use the Username **GIS_Institute** with password **ARCGIS2019** and click SIGN IN. Then skip to step 2.5.

6.3 For the ENTERPRISE LOGIN option, enter “harvard-cga” under where it says “Enter your ArcGIS organization's URL below.” And click Continue.

6.4 Click the HARVARD UNIVERSITY button. Enter your Harvard Key information and you’ll be logged into the Harvard account.

6.5 Click **Map** at the top – this will create a new blank map for you.

6.6 Click **Add > Add Layer** from file, and click the **Choose File** button.

Browse to your Geocoding_lab folder, and choose the World_Livable_Cities.csv file and choose IMPORT LAYER. This will open the Add CSV Layer dialog box.

6.7 At the top of the Add CSV Layer box, specify **Addresses or Places** for the Locate features by choice. Next to the City_County field name, click under Location Fields where it says “Not used”, and change this to “Address or Place”. Make sure the other fields are specified as “Not used”. Click ADD LAYER, and the geocoder will run.

6.9 The map will now prompt you to choose an attribute to show. Click the drop down arrow and choose “Show Location only”. Your map will now display a black dot at each of the geocoded locations. Inspect some of these to make sure they are in the correct location. You can save this map using the Save button at the top, and then you’ll be able to share the link using the Share button.

Conclusion - General tips to keep in mind when geocoding:

- Clearly determine your geocoding purpose, and the accuracy level required to meet this need. If you just need to geocode your data to the city or country level, don’t complicate things with address information. However, if address level accuracy is necessary, take the proper steps to ensure confidence in your result.
- Have an ‘iterative’ mindset. Geocoding may reveal errors or typos in your data, or expose the shortcomings of the geocoding method used. Be prepared to re-geocode, and refine your data and geocoding process accordingly – several geocoding iterations may be necessary to achieve the desired result.
- Don’t treat one basemap as the gospel. Compare your results to more than one basemap. For example, if geocoded in ArcMap, import the results to Google Earth to see if they match GE’s basemap.

Additional notes:

If you have over 5,000 US addresses, using the local geocoder is recommended as it will be run much faster.

Batch geocoding is possible using Google Maps, however it requires obtaining an API key, and using a Python script. For more information on this go here:

<https://gis.harvard.edu/Google-Maps-API-Premium>

Batchgeo.com is a real quick, simple way to geocode up to 250 addresses

For GIS and geocoding questions and help contact the CGA at: <http://gis.harvard.edu/contactus>