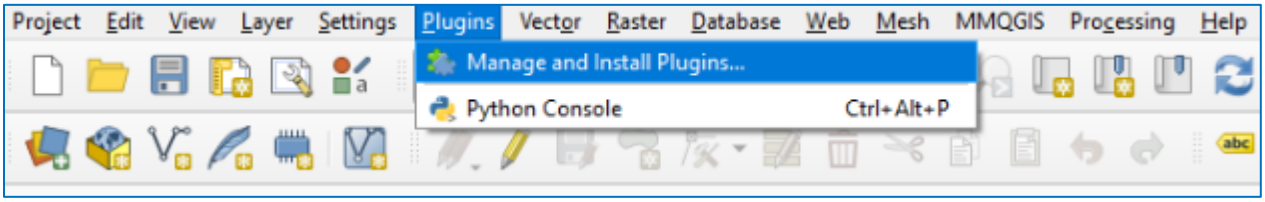
Geocoding in QGIS with the MMQGIS Plugin

Background

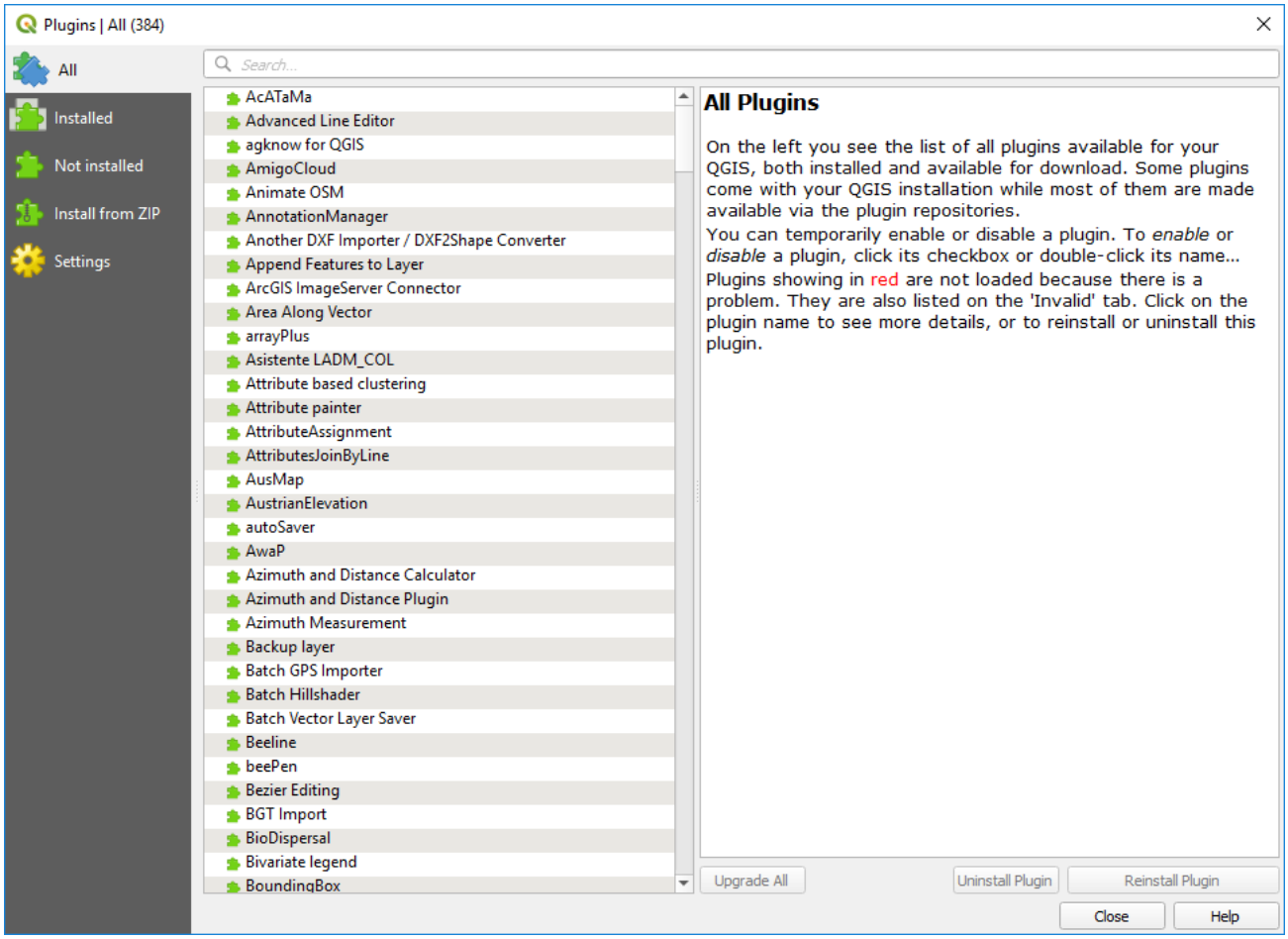
The primary use of Geospatial Data Service web services has been through geocoding operations, however prior to October of 2019 geocoding in QGIS with the Geospatial Data Services geocoding service was not an option. In early 2019 Geospatial Data Services reached out to Michael Minn, a professor at Farmingdale University (now at the University of Illinois), and the developer of the MMQGIS plugin. The MMQGIS plugin allowed for multiple geocoding options such as BING, OpenStreetMap and Google, however at the time it did not allow for the addition of a custom ESRI geocoding option. Dr. Minn incorporated ESRI options for geocoding in QGIS, and the successful integration of this work was presented at the QGIS North America 2020 Virtual Conference. Recording available here - https://www.youtube.com/watch?v=fT1OPJasv3U&t

Step1: Installing the MMQGIS Plugin

To install a QGIS plugin, click **Plugins** and then **Manage and Install Plugins…** from the top QGIS toolbar



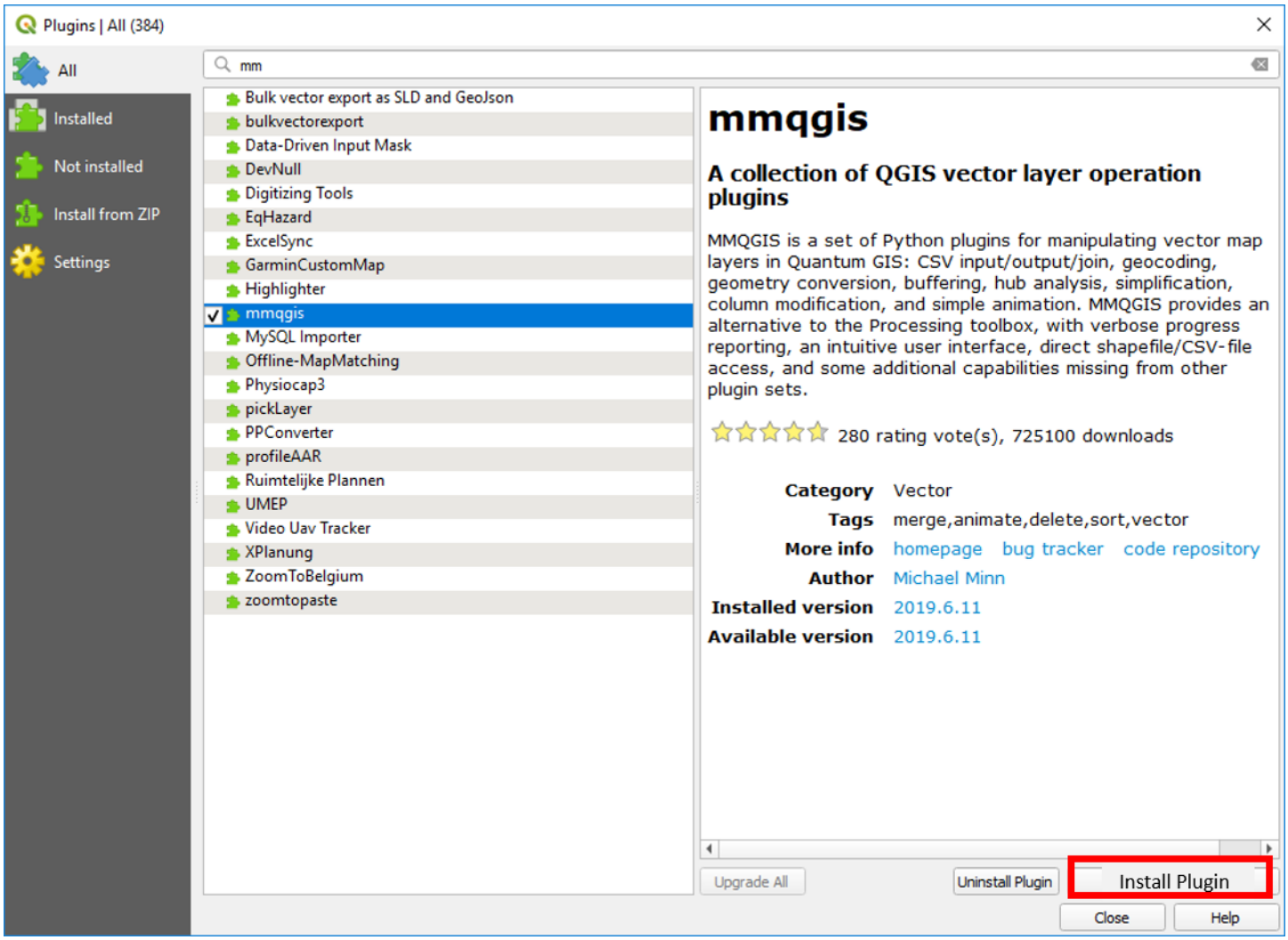
This will open a plugins window where you can add various plugins provided through the QGIS repository. You may be prompted that QGIS is “fetching repository information,” however just wait for the plugin window to open. It should look like this…



Notice that on the left-hand side of this window there are options to view plugins that are “Installed”, “Not Installed,” etc. For the purpose of this exercise, make sure that the option for “**All**” is selected. You should see an alphabetized list of plugins in the middle column.

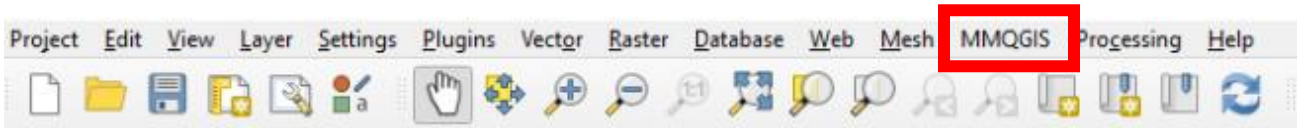
Either in the search bar at the top of the window, or by scrolling down, locate the MMQGIS plugin.

Highlighting the MMQGIS plugin should change the dialogue on the right-hand side of the window to include a description of the MMQGIS plugin. In the lower right-hand corner, you should also see a button labeled **Install Plugin**. Click the button and wait for the plugin to install (please note that if you have previously installed the MMQGIS plugin the button will say “Reinstall Plugin,” which you should do in order to use the updated geocoding features). You are now ready to start geocoding.



Step 2: Geocoding with the MMQGIS plugin

After installing the plugin, you should now see a MMQGIS option in the top QGIS toolbar.



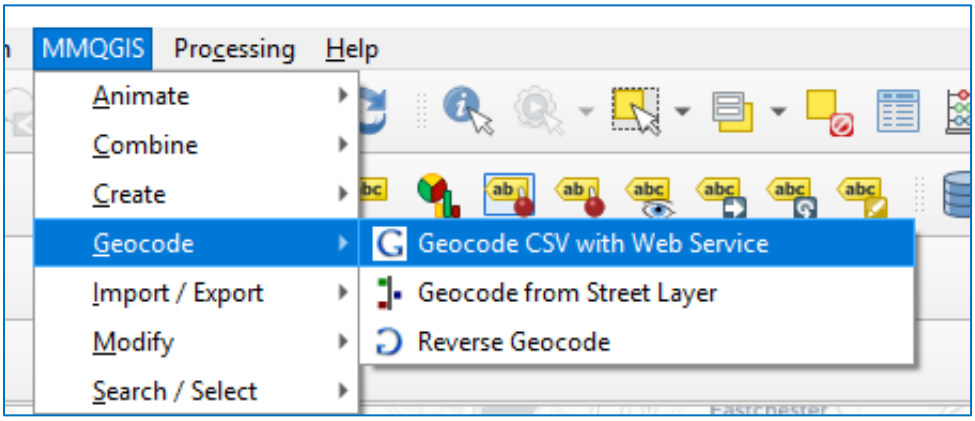
Clicking MMQGIS will provide a dropdown list of supported activities; select Geocode to see the geocoding options available. The MMQGIS plugin provides three options for geocoding:

1) **Geocode CSV with Web Service** – Tool that imports addresses from a CSV file and uses a web geocoding service to geocode addresses to a point output file. The web service tools use the Python urllib module to make https requests to the respective geocoding APIs.

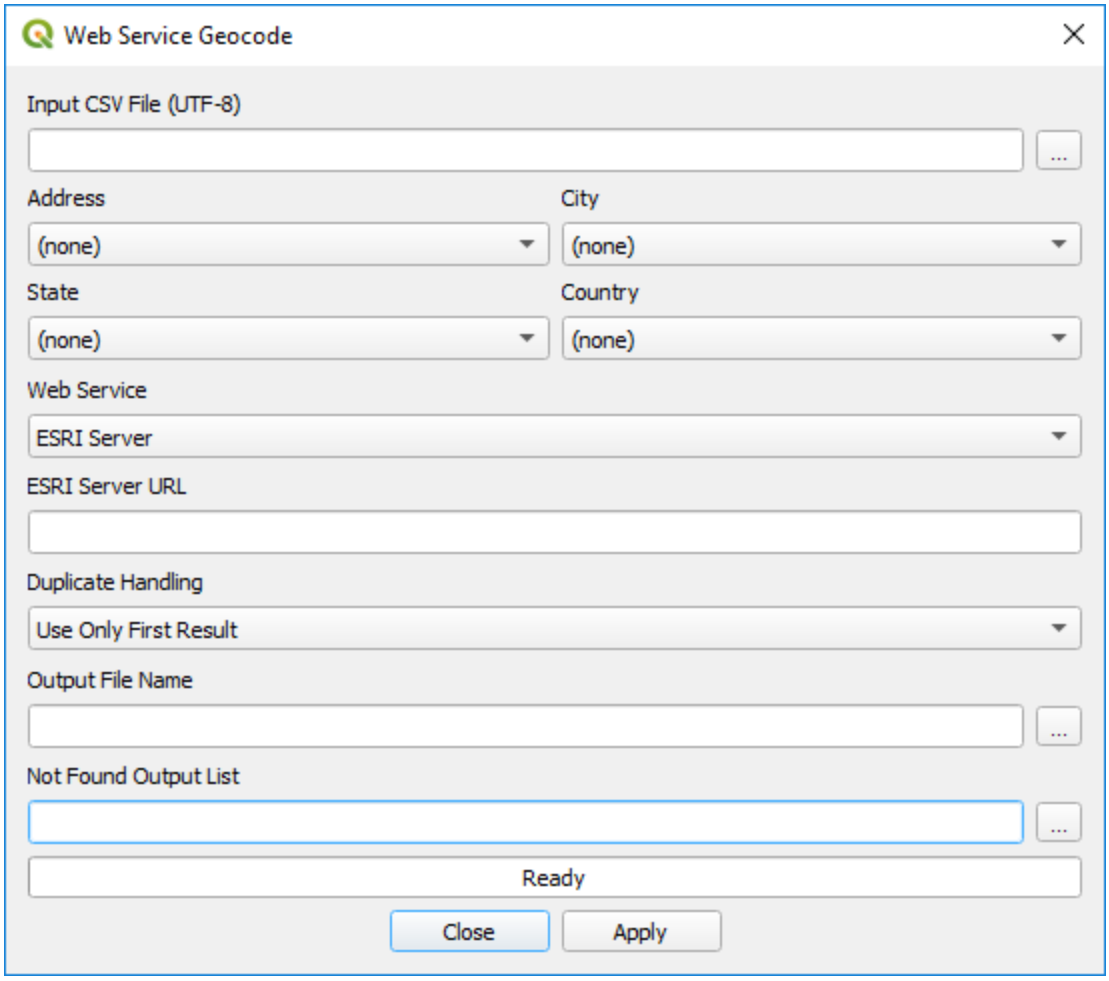
2) **Geocode Street Layer** – Tool that geocodes addresses from a CSV file using an address locator layer with street centerline features and attributes indicating the range of addresses associated with each feature.

3) **Reverse Geocode** – Tool that uses Google or Nominatim (OpenStreetMap) API to find addresses associated with point feature locations. If features are lines or polygons, the centroid of each feature will be used for querying.

Select the **Geocode CSV with Web Service** option.



This will open the **Web Service Geocode** window.



The inputs for this window are as follows:

- **Input CSV File:** CSV table of addresses to geocode. This file should be encoded in the UTF-8 character set. Although other 8-bit encodings will work if only ASCII characters are present, non-ASCII characters may cause unpredictable behavior.
- **Address, City, State, and Country:** Selected columns from the input CSV file are added as attributes in the output shapefile. Addresses may be spread across as many as four different columns. However, these fields are concatenated into a single address to query the API, so only one meaningful column is absolutely required (such as for a city/state combination).

  PLEASE NOTE: This means that even though the Address, City, State and Country dropdowns appear to require a like named corresponding column from your CSV file, ***any CSV field may be entered into any of the dropdowns***. For example, a single address field may be entered into the State dropdown and all the other dropdown can be set to "none." The plugin concatenates everything entered into all of the dropdowns and uses it to geocode.

- **Web Service:** The type of geocoder you wish to use. Choose from one of the four options provided: Google; OpenStreetMap/Nominatim; US Census Bureau; or ESRI Server. You will select the ESRI Server option.

- **API Key:** This field appears when either the Google, OpenStreetMap or US Census Bureau Geocoders are selected. To use these geocoders, you would need to get an API key and include it in the API Key dialog box.
- **ESRI Server URL: When using an ESRI Server for geocoding, a server URL must be provided. Publicly** accessible servers include:
    - ArcGIS World Geocoding Service: https://geocode.arcgis.com/arcgis/rest/services/World/GeocodeServer/findAddressCandidates
    - **NY State ITS Geospatial Data Services' Geocoding Service:**

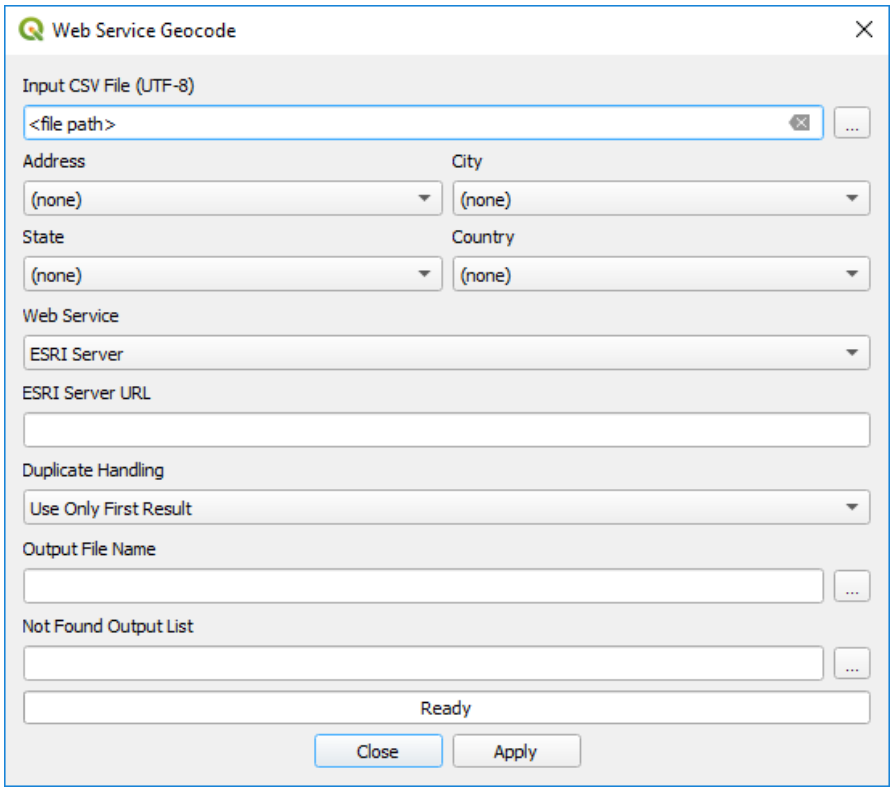
      https://nysgeohub.ny.gov/arcgis/rest/services/Geocoder/NYS_Geocoder/GeocodeServer/findAddressCandidates

      PLEASE NOTE: For the geocoding function to work, the url must be entered all the way down to the ESRI supported operation level. In the case of geocoding the url must therefore end at the **"findAddressCandidates"** operation.
- **Duplicate Handling:** This parameter indicates how to handle conditions where multiple results are returned for the same address.
    - **Use Only First Result:** As it suggests, this option retains only the first geocoded result. The result returned represents the best possible match determined by the geocoder after cascading through a series of locators.

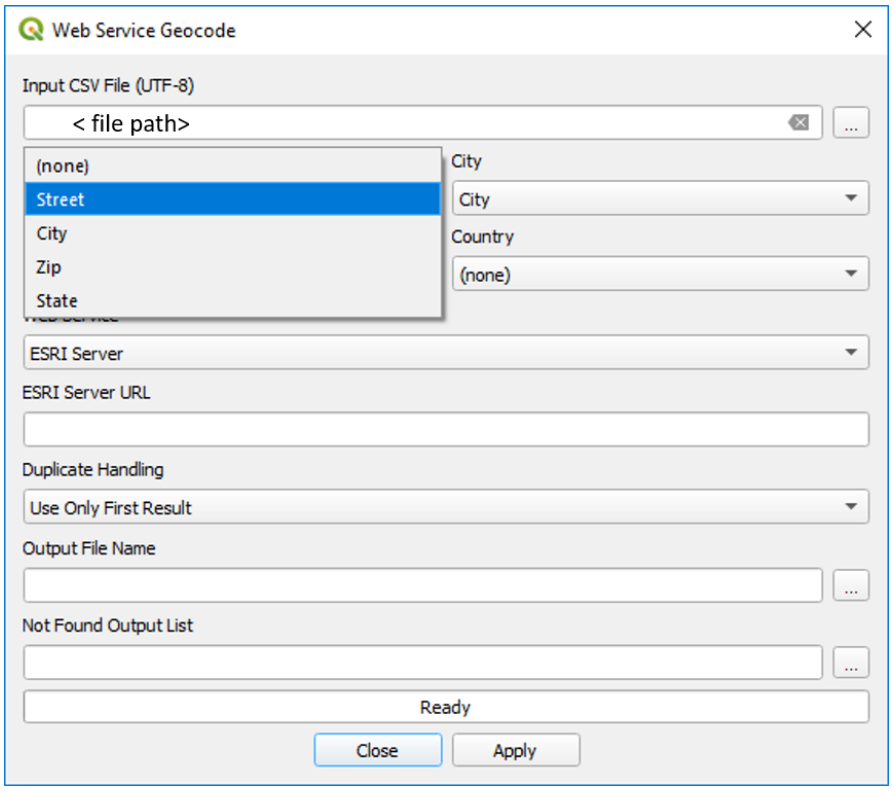
      PLEASE NOTE: The score for this match may not always be 100% even though it is the best match returned.
    - **Multiple Features For Multiple Results**: This option retains all of the potential matches determined for a single address based on matches to different locators and the information stored within. Geocoders often return multiple locations representing different accuracy levels and spatial scales, for example: rooftop accuracy, street segment, parcel centroid, primary entrance, driveway, etc.). It is possible that since all these options exist inside the GPO's locators you may see a match return for each. In the case of an apartment complex or other addressing anomaly, a large number of matches could be returned. Therefore, use of this option will likely require extensive editing of the results to cull unneeded points.
- **Output File Name:** The name of the output spatial / feature file that will be added to the QGIS map.
- **Not Found Output List:** The name of the output CSV file that will contain input addresses that could not be geocoded.

To begin geocoding:

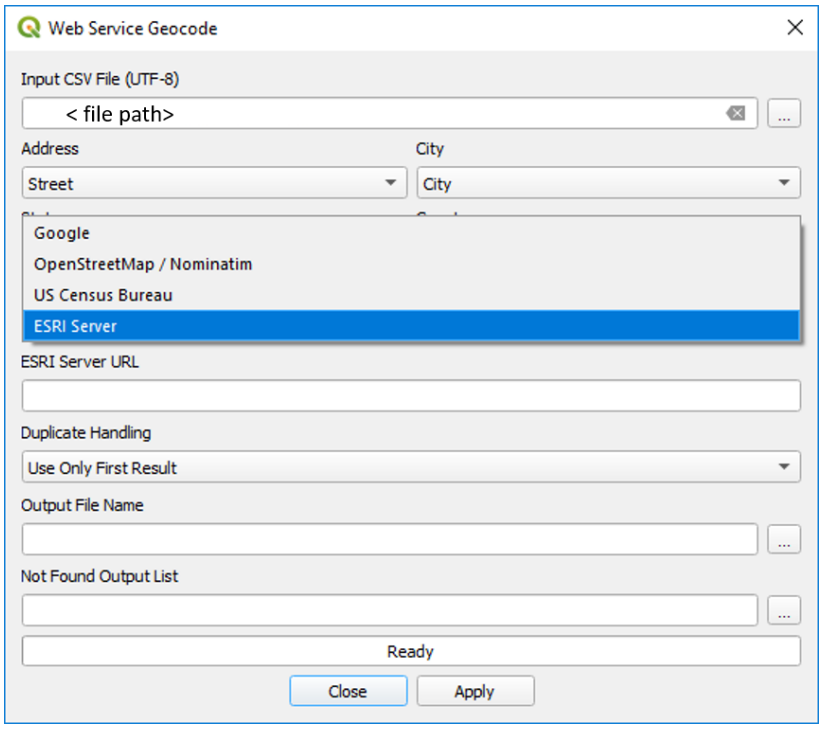
1) Browse for and select an input CSV file to be geocoded.



2) Determine which field in your input file will be used to geocode an address, for example Address, Street, SingleLine, or some combination of fields. Enter these fields into the MMQGIS window using the dropdown options, remembering that the dropdown name and your CSV field do not need to be the same.

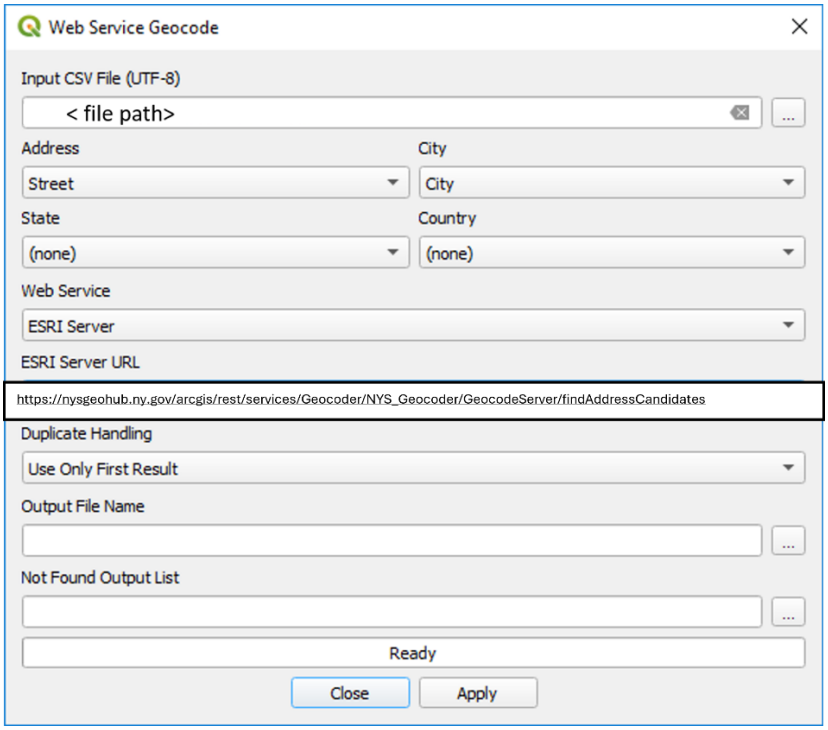


3) Select the Web Service you want to use to geocode; in this case choose “ESRI Server”

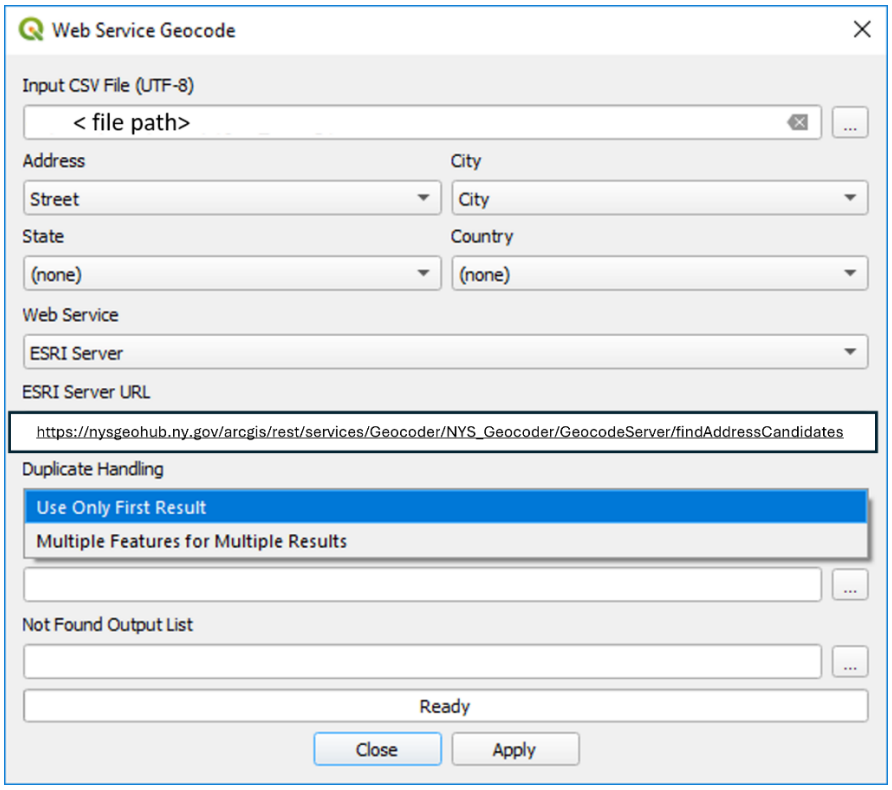


4) Enter the NYS Geospatial Data Services url for the Streets and Addresses composite locator into the ESRI Server URL box

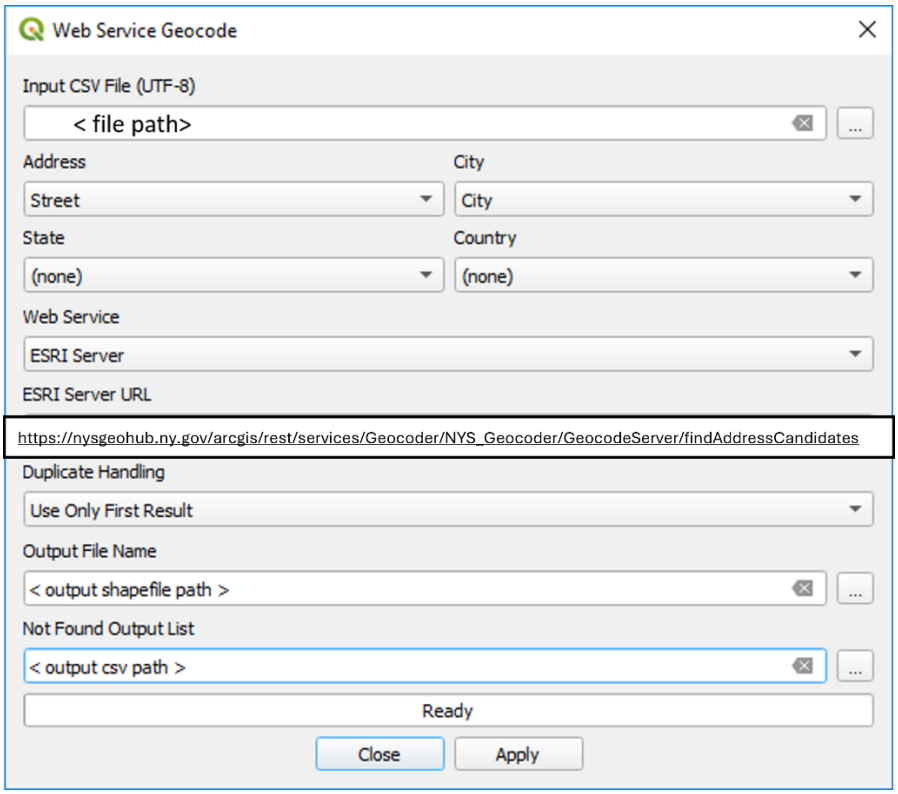
https://nysgeohub.ny.gov/arcgis/rest/services/Geocoder/NYS_Geocoder/GeocodeServer/findAddressCandidates



5) Determine if you want only the first result returned or multiple results per address.



6) Give a file name and output location for your spatial output file and not found list.



The output spatial file may take any of the following forms:

- ESRI Shapefile (*.shp)
- GEOJSON (*.geojson)
- KML (*.kml)
- Spatialite (*.sqlite)
- GPKG (*.gpkg)

The not found address file can take the form of either a .csv or .txt. file.

7) When you are ready to geocode click "**Apply**"
This will initiate the geocoding operation. The "Ready" bar visible on the dialogue box will begin to will with the total number of successful geocodes completed until it run through 100% of the input addresses. You may see a spinning icon for particularly large datasets, just wait for the operation to conclude.