

Chapter 6: Data Acquisition Methods, Procedures, and Issues

In this Exercise:

- Data Acquisition
- Downloading Geographic Data
- Accessing Data Via Web Map Service
- Using Data from a Text File or Spreadsheet w/Coordinates
- Joining Tabular Data
- Downloading Data from a GPS
- Creating and Editing Data
- Online Data Resources

This exercise requires an active Internet connection!

Data Acquisition

The data used in previous exercises has been provided on a CD or loaded directly on the computer.

In many situations users will want to build projects that include data acquired from other sources. In addition, users will frequently add data that they themselves have created to overlay with available geographic data.

The process of **Data Acquisition** involves a number of specific tasks that can be challenging. Procedures typically include:

- Locating and reviewing the properties of available geographic data
- Downloading files
- Extracting contents from compressed files (used to reduce file size)
- Transferring all relevant files to directories for use in the GIS application

There are a number of issues to consider when beginning the Data Acquisition process that can be identified by viewing available **Metadata**.

Data Acquisition and Metadata

Metadata should always accompany geographic data. As mentioned in previous chapters, metadata is the information that will tell potential users about the specific properties of geographic data layers. All VCGI metadata follows a federal format, however any metadata is better than none! Even a brief paragraph about who created it, why, and when as well as definitions of codes can be invaluable. Ask for information about the data when you get it from a source other than VCGI, and create simple metadata if you create new GIS data.

1. GIS File Formats

Geographic data may be available in a variety of file formats. Thus far we have used ESRI **Shapefiles** and **Imagery** in JPEG2000 or TIFF format.

QGIS can use data and data services of the following types:

- PostGIS and SpatiaLite
- Most vector formats supported by the OGR library, including ESRI shapefiles, MapInfo, SDTS and GML
- Raster formats supported by the GDAL library, such as digital elevation models, aerial photography or landsat imagery
- GRASS locations and mapsets
- Online spatial data served as OGC-compliant WMS (web map service), WMS-C (tile cache), WFS (web feature service) and WFS-T (tile cache).

A growing collection of open source software and utilities for GIS that may be useful for file conversions can be found on the Open Source Geospatial Foundation Website (<http://www.osgeo.org/>).

Metadata, as well as any online description, should include **Distribution Information** describing the format of the available digital data.

Metadata:

- [Identification Information](#)
- [Data Quality Information](#)
- [Spatial Data Organization Information](#)
- [Spatial Reference Information](#)
- [Entity and Attribute Information](#)
- [Distribution Information](#)
- [Metadata Reference Information](#)

Standard_Order_Process:

Digital_Form:

Digital_Transfer_Information:

Format_Name: SHP

Format_Specification: ESRI Shapefile format (compressed to ZIP format)

Format_Information_Content: geospatial data

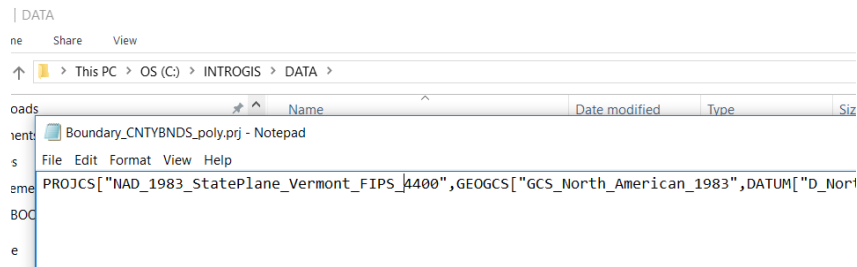
File-Decompression_Technique: ZIP decompression software

2. Coordinate Reference System (CRS)

In Chapter 5, we set the Coordinate Reference System to VT State Plane Meters, NAD 83. We also turned on the "Enable on the fly CRS transformation" option because this allows the software to transform data from whatever coordinate system and project it is natively in to the coordinate system and projection that we have set for our project. By doing so, we ensure that the data lines up and scales properly.

The software is able to perform these functions because of a projection file associated with each shapefile with the following extension: .prj.

- a) Open your file browser (outside of QGIS) and navigate to the C:\INTROGIS\DATA FOLDER
- b) Right click on a file with a .prj extension and choose Open With and then Notepad. Note that this file is simply a text file defining the projection, coordinate system, units, and datum for this shapefile.



- c) Projection files give GIS software the information they need to transform the data into whatever projection, coordinate system, and datum the project is set to. This allows all data layers to line up and scale properly.
- d) **TIP:** if you have a shapefile that does not have a .prj file, but you actually do know that it is the same projection and coordinate system as other data you have (e.g. VT State Plane Meters NAD 83...) you can simply make a copy of an existing .prj file and "save as" with a name that exactly matches the new data's name. This makes the new .prj file part of the new shapefile and indicates to GIS software what its coordinate system and reference system are.

The **Spatial Reference Information** section of metadata will list information describing the coordinate system.

Metadata:

- ♦ [Identification Information](#)
- ♦ [Data Quality Information](#)
- ♦ [Spatial Data Organization Information](#)
- ♦ [Spatial Reference Information](#)
- ♦ [Entity and Attribute Information](#)
- ♦ [Distribution Information](#)
- ♦ [Metadata Reference Information](#)

For example, the Vermont State Plane Coordinate System is a Coordinate System that has been projected. This means that the geographic data has been translated from a spheroid surface (the earth) onto a flat map. Metadata will include the projection (Transverse Mercator), the datum (NAD83), map units (Meters), as well as other specific values.

Understanding the specific values defined by a Coordinate System and Projection is critical when adding and assembling data themes in a GIS project and application (see below).

Downloading Geographic Data

Geographic data is available online from a wide variety of sources and websites.

1. Vermont

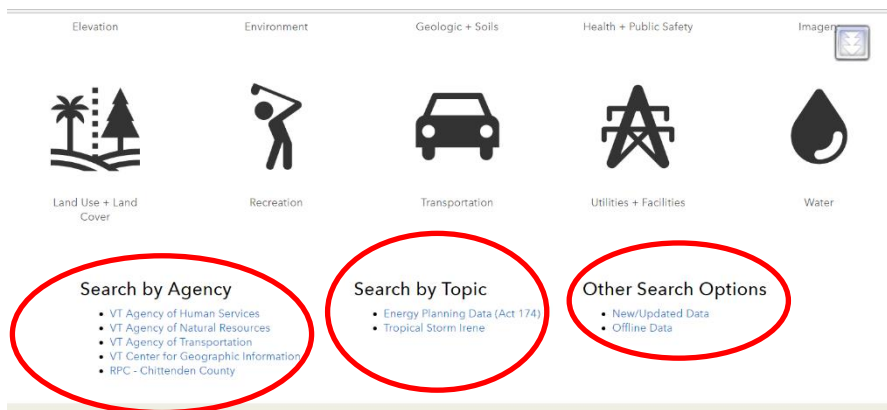
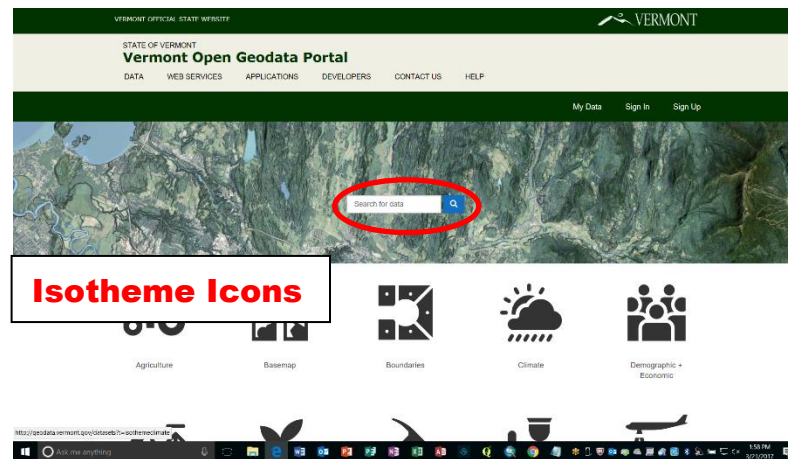
The main source of public GIS data in Vermont is the **Vermont Open Geodata Portal** (<http://geodata.vermont.gov/>). This online resource is available to users who wish to search for and download free Vermont GIS data from VCGI. Many of the data sets created by regional, state, and federal organizations are found here. Available data varies in content, scale, age, and accuracy.

Here is an overview of what we will encounter at the Portal:

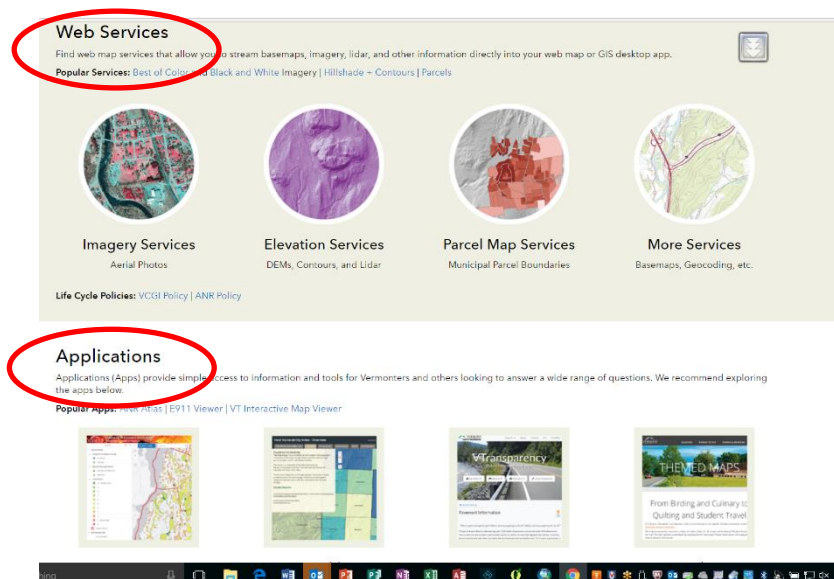
Link to How-to videos at bottom of page!

Different ways to search

- Keyword
- Isothemes – click on icon
- Agency – click on name of agency
- Topic – click on name of topic
- New/Updated – click on
- Offline



- Web Services Categories
- Mapping Application Categories



All methods of search result in a Search Results Table. The results listed may be quite different, however and lead to different options

- Data that can be explored and downloaded from within the portal – click on the title and you will go to the information page for that item. Note that the thumbnail to the right of the item title confirms that it is a DATA item (not a service or application)

The image shows two screenshots of the Vermont Open Geodata Portal. The top screenshot is a search results page with a map of Vermont and surrounding areas. A red circle highlights the search result for "VT Green Mountain National Forest Mgmt Areas 2003 - polygons". The bottom screenshot is the detailed view of this dataset. It includes a map, a description, and a table of attributes. A red circle highlights the "(Link to Metadata)" text in the description, and another red circle highlights the "View Metadata" link in the "About" section.

Search Results Page:

Filter By: 1 - 10 of 71 results. Relevance

Tags: isothemeboundary (58), isothemecadastral (12), isothemecother (56), nodevcgi (52), vcgi open data (52)

VT Green Mountain National Forest Mgmt Areas 2003 - polygons
 Shared by Opendata_VCGI
 (Link to Metadata) The EnvironMangareas_MAREA2003 data layer includes management area boundaries administered by the Green Mountain National Forest and partner organizations. EnvironMangareas_MAREA2003 is an updated version of EnvironMangareas_MAREA2001 including all lands
 Custom License 12/16/2016 Spatial Dataset 725 Rows

Item Details Page:

Overview Data API Explorer

VT Green Mountain National Forest Mgmt Areas 2003 - polygons

(Link to Metadata) The EnvironMangareas_MAREA2003 data layer includes management area boundaries administered by the Green Mountain National Forest and part VT House Districts 2012 - polygons. EnvironMangareas_MAREA2003 is an updated version of EnvironMangareas_MAREA2001.

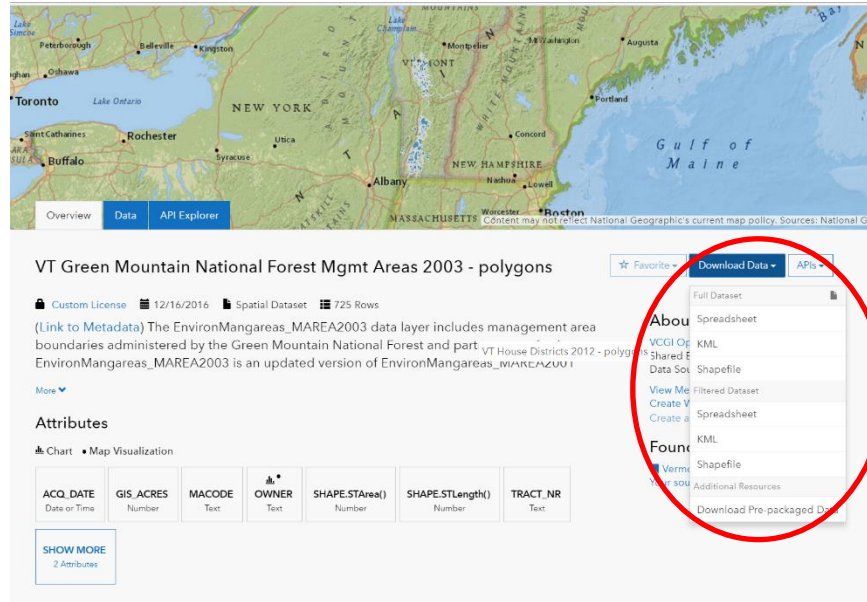
About: VCGI Open Data, Shared By: Opendata_VCGI, View Metadata, Create Webmap

Found In: Vermont Open Geodata Portal | Your source for geospatial data

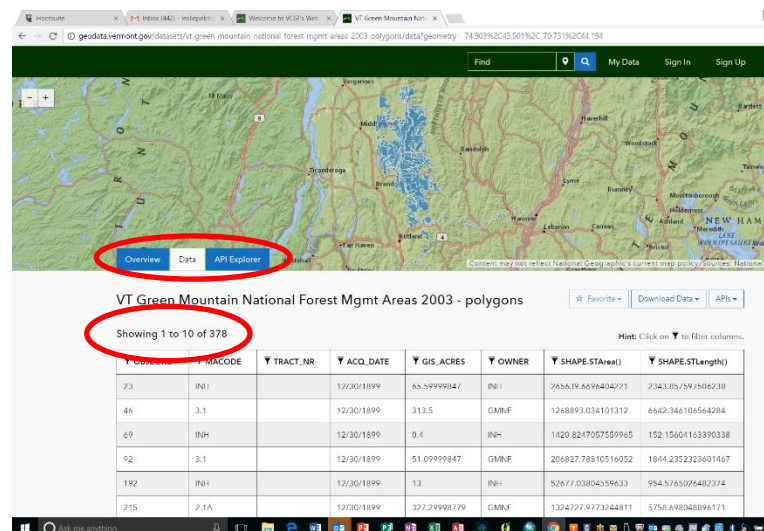
ACQ_DATE	GIS_ACRES	MACODE	OWNER	SHAPE.STArea()	SHAPE.STLength()	TRACT_NR
Date or Time	Number	Text	Text	Number	Number	Text

SHOW MORE 2 Attributes

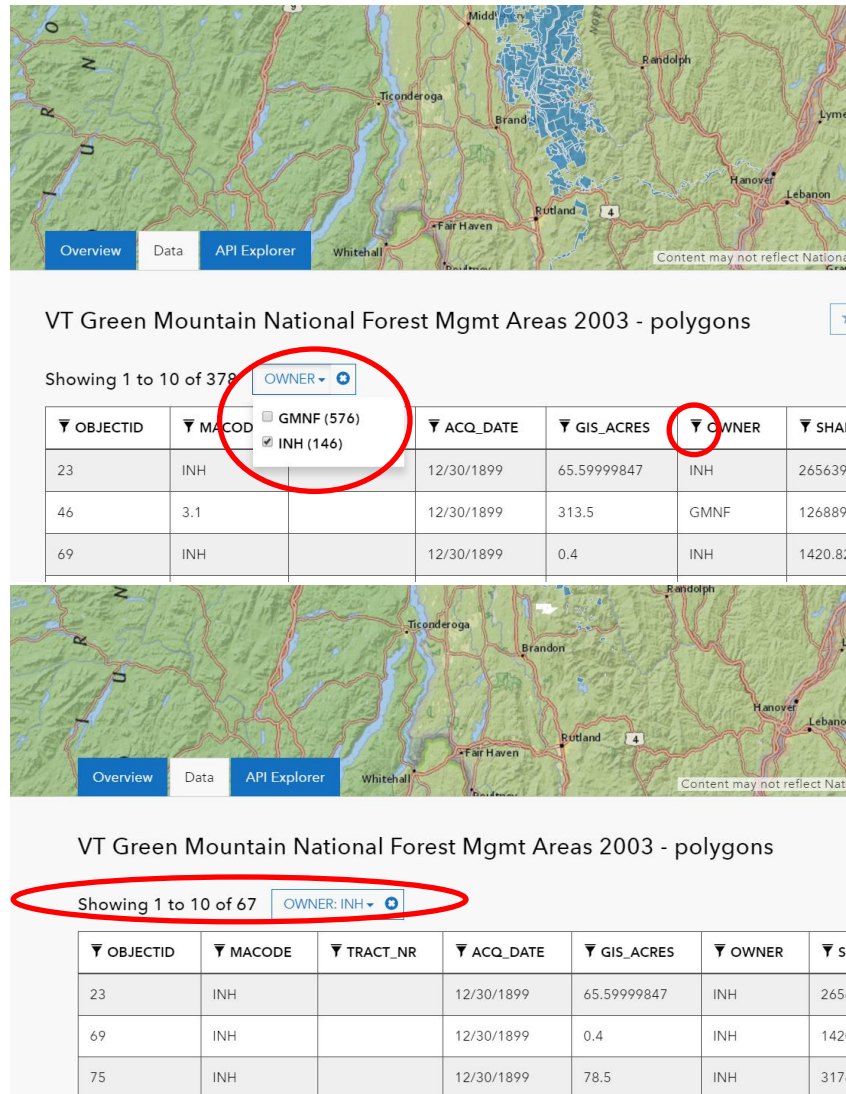
- Click on (Link to Metadata) in the description or “View Metadata” in the about section to view the metadata file.
- Close the metadata window when done.
- In the Search Results window click on the **Download Menu** in order to download the data layer. Note that you have format options as well as “filtered” options.



- Before we download the layer, let's take a look at how we can filter the layer and only download a subset of features.
- As in the image above, we can zoom in on the map that shows the features (plus and minus signs in upper left of map are zoom tools). Now, if we choose the filtered download option, we will only be download those features that occur within the chosen extent. See if you can zoom in on only the northern or southern section of the Green Mountain National Forest.
- If you click on the "Data" tab above the data layer description, you will see that the attribute list also reflects this filtering: by indicating the number of features being included "showing 1 to 10 of X" (try changing your zoom level).



In the data view, you can also filter by attribute value. Click on the filter symbol to the left of the attribute name, this will make the attribute name appear just below the name of the layer. Now click on the dropdown arrow to the right of the attribute name and indicate values that define what you would like to include. A date range, a parcel size, or owner type for instance. Hit enter to finalize the choice.



VT Green Mountain National Forest Mgmt Areas 2003 - polygons

Showing 1 to 10 of 378

OWNER

OBJECTID	MACOD	ACQ_DATE	GIS_ACRES	OWNER	SHAPE
23	INH	12/30/1899	65.59999847	INH	265639
46	3.1	12/30/1899	313.5	GMNF	126889
69	INH	12/30/1899	0.4	INH	1420.82

VT Green Mountain National Forest Mgmt Areas 2003 - polygons

Showing 1 to 10 of 67

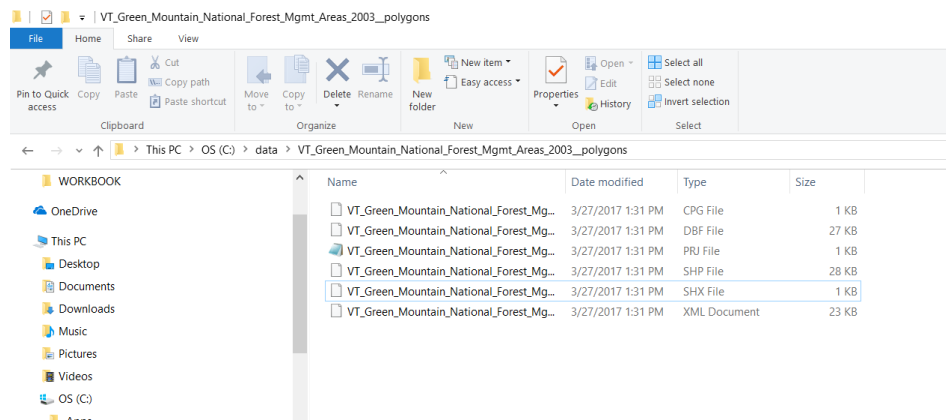
OWNER: INH

OBJECTID	MACODE	TRACT_NR	ACQ_DATE	GIS_ACRES	OWNER	SHAPE
23	INH		12/30/1899	65.59999847	INH	265639
69	INH		12/30/1899	0.4	INH	1420.82
75	INH		12/30/1899	78.5	INH	31760

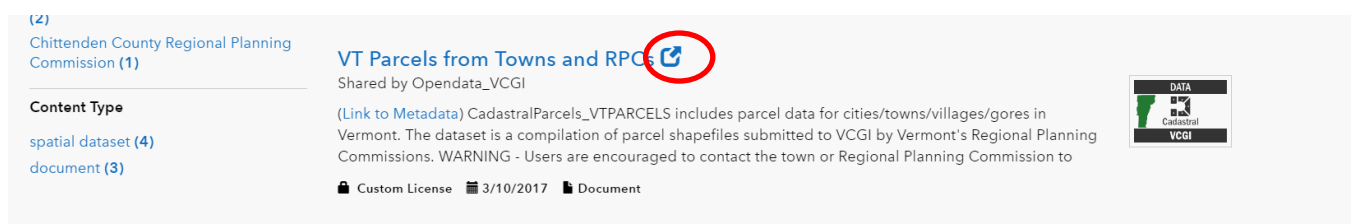
- Now, if I choose to download filtered data, I will only download these 67 parcels.
- When you do download the data you have found, you will be downloading a zipped file and, depending on your browser, it may automatically be saved in your downloads folder, or you may be given the choice to “Save As” (choose that!) so you can put it in a logical place.

- c) Using a file manager, navigate to where you have saved the file, right click and choose “extract all.” Make sure to save the unzipped data to an appropriate folder if necessary.

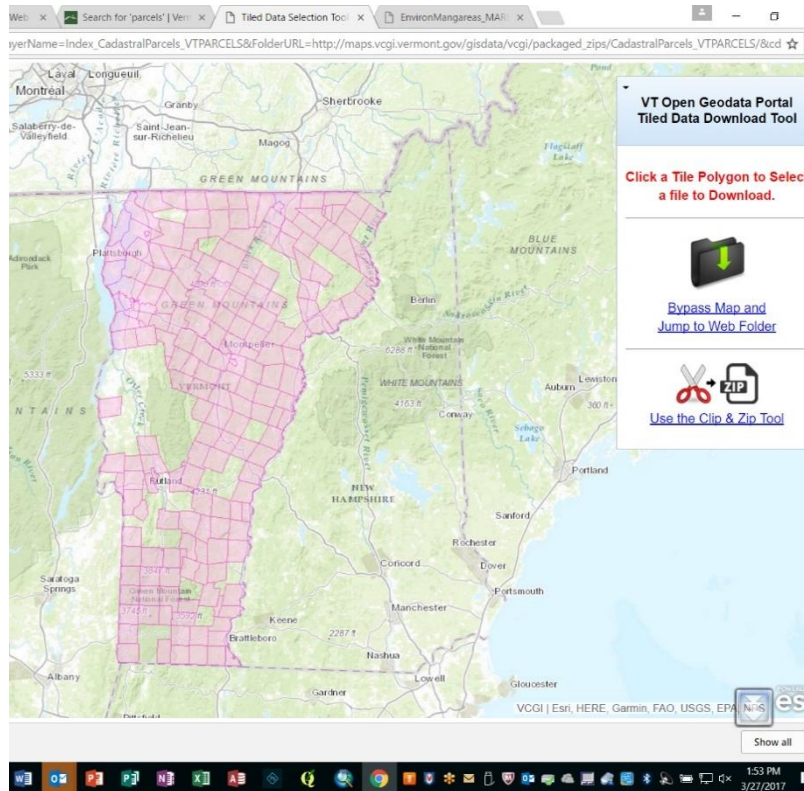
When extracted there will be a number of files that together are necessary for the ESRI Shapefile of The Green Mountain National Forest Management Areas.



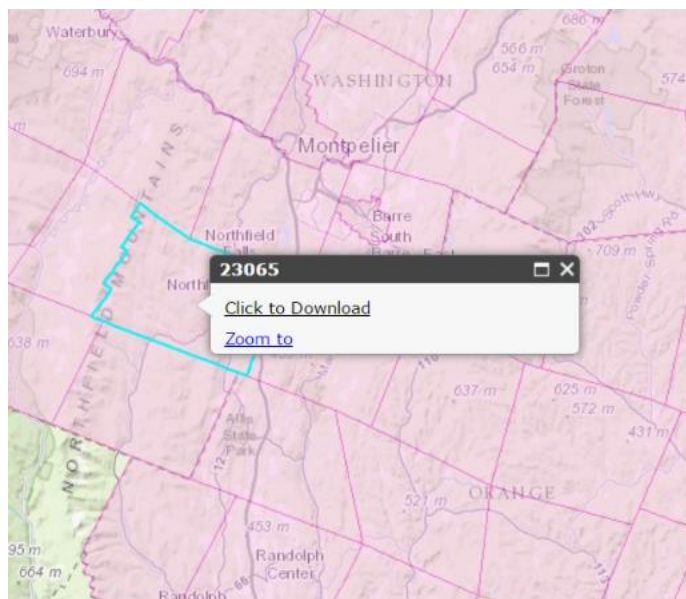
- Data that can be downloaded via external link will include ALL raster data (orthophotos, elevation data other than contours, scanned Topo maps) as well as some other layers like town parcel data. Applications and Web Services are also accessed via external link. How do you know that a search results item is actually an external link? It will have the following icon at the end of the title:



- Options for download may include:
 - Direct download - smaller and non-tiled data
 - Tiled Download – see example below
 - Bypass map download – see example below
 - Clip and Zip custom download – see example below
- Let’s try doing a keyword search on “town parcels.” Once you see “VT Parcels from Towns and RPCS” click on the title.

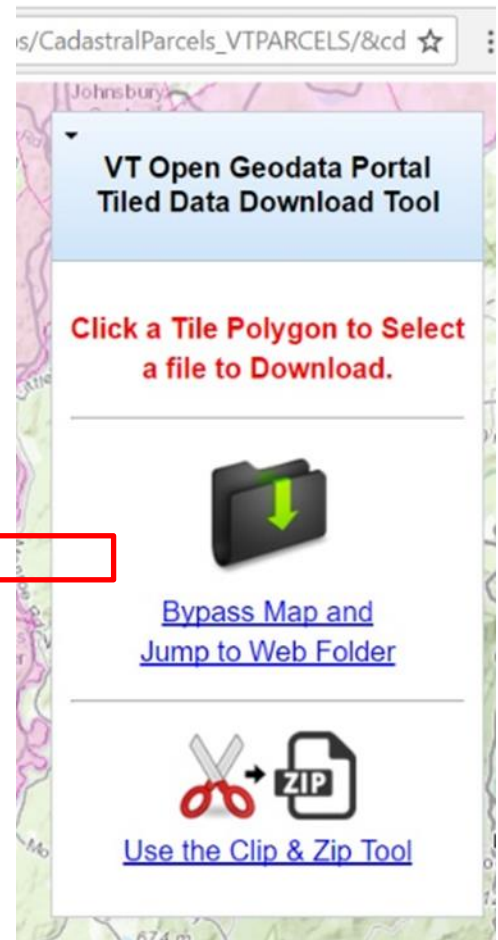


We will be taken to the Tiled Data Download tool at which we can click on a town and choose to download the parcel data for that town.

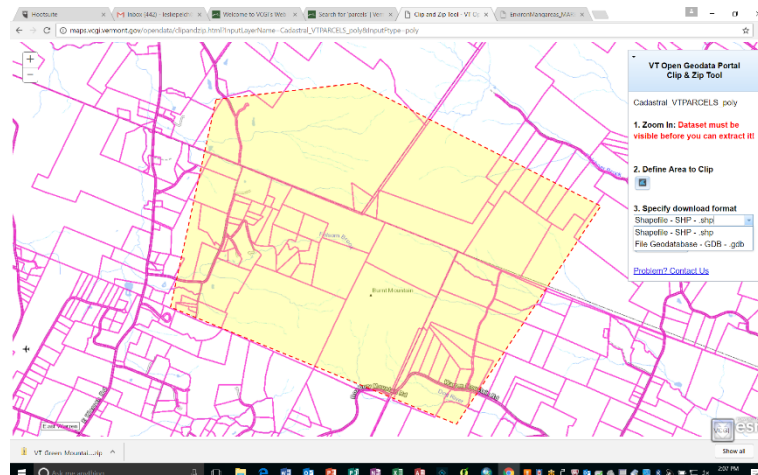


You have two other options for download, however!

If you click on “Bypass Map and Jump to Web Folder” you will view a listing of files (and names include name of town and year of data, which is helpful!).



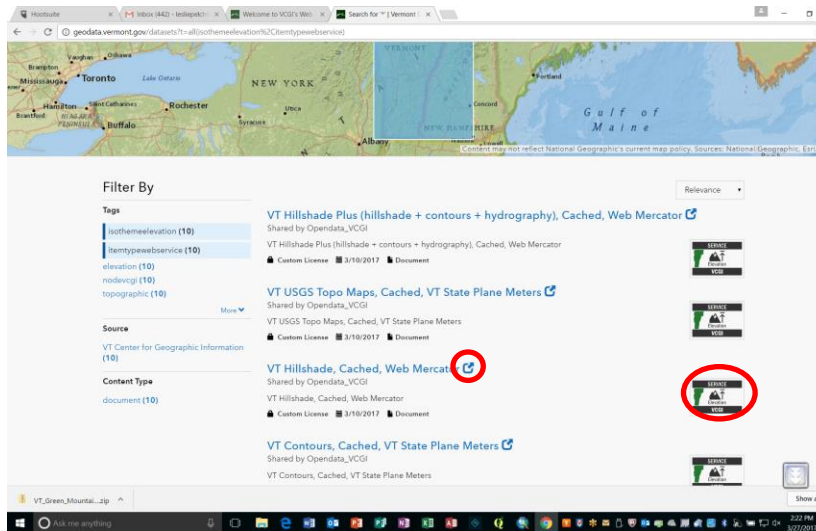
If you click on “Use the Clip and Zip Tool” you will have access to another way to select your area of interest. After zooming into the map enough to actually see the parcel boundaries, you can draw your area of interest (across towns lines if you like) and download in the format of your choice:



Web Services are another way to access the data provided by VCGI and its partners, but without downloading and saving the data onto your computer. You can find web services through any search method, but if you are specifically looking for services, you can use the Web Services search options below the Isothemes, Agency, Topic, and Other search options.

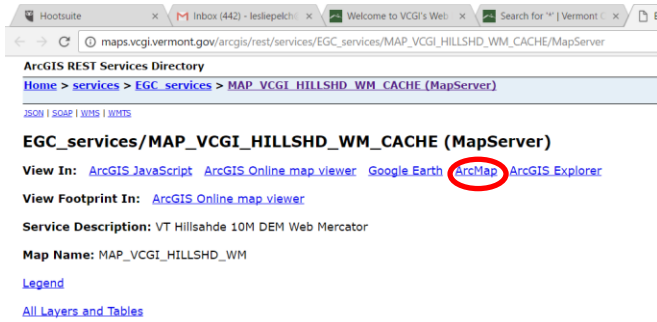


When you click on one of the options, you once again have a search results table, but this time all results are web services, which also means they are ALL external links.

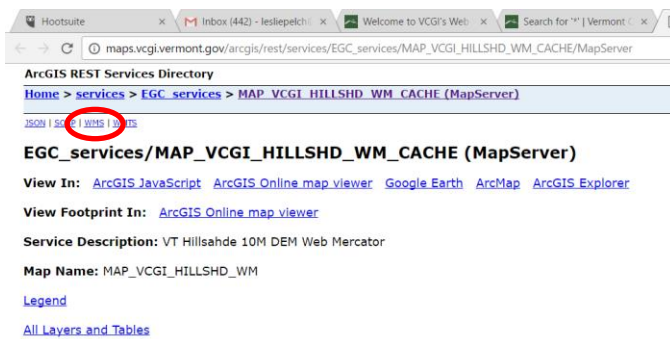


When you click on the title of a service, you will always be taken to a “Rest Endpoint Page”

If you are working in ArcGIS software, you can simply click on the “ArcMap” link and an ESRI “layer” file will download to your computer. This is simply a pointer, so if you double click on your open project will be connected to the service, or ArcGIS will open a new project that contains the service.



If you are working with an open source software like QGIS, you will right click on the WMS link and then copy that URL into the “Add WMS/WMTS” dialog box in QGIS.



Web Mapping Applications are online mapping tools, and when you come across them in the Open Geodata Portal, they are really just links to those apps. Try clicking on some to see examples:

Applications

Applications (Apps) provide simple access to information and tools for Vermonters and others looking to answer a wide range of questions. We recommend exploring the apps below.

Popular Apps: ANR Atlas | E911 Viewer | VT Interactive Map Viewer



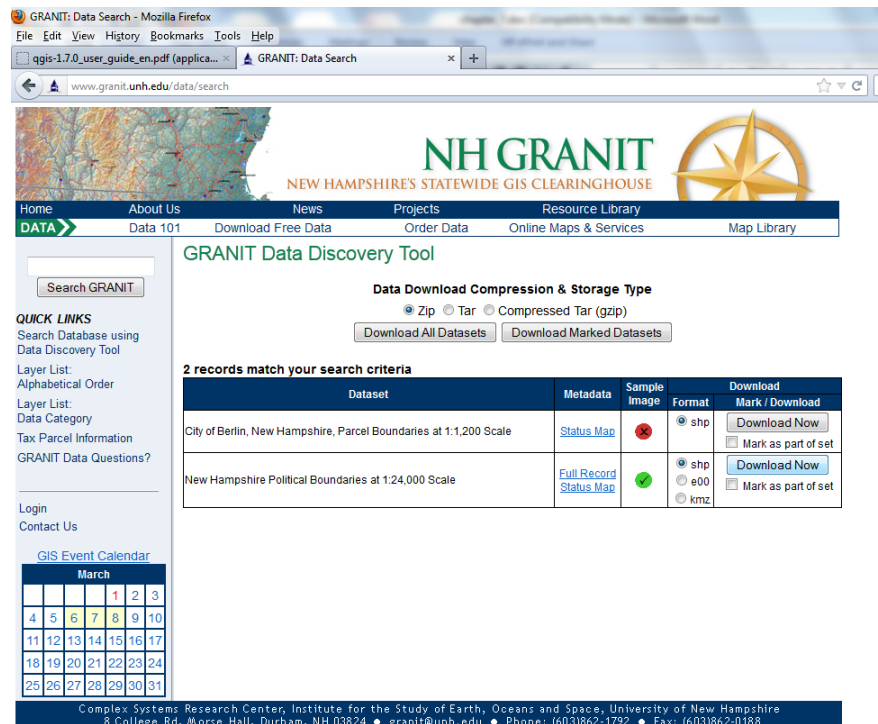
2. Data from Other Sources: NH Granit

- a) Go to <http://www.granit.unh.edu/>
- b) Click on "Download Free Data">>"Data Download"



- c) Click on " GRANIT Data Discovery Tool (DDT)"
- d) Next to the words "Search by theme keyword:" type "Town" in the box
- e) Click on the "Search" button

- f) Download the data called "New Hampshire Political Boundaries at 1:24,000 Scale" by clicking on the "Download Now" button (notice the various settings and options, but leave them set to their defaults)

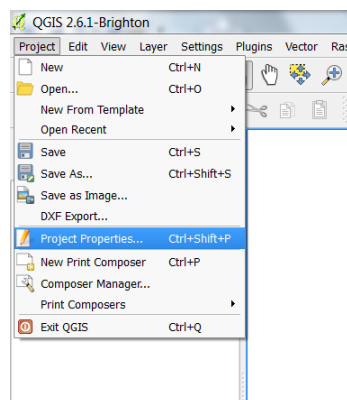


- g) Save the .zip file in an appropriate location and then extract it to access the two shapefiles within: pba.shp (line data) and pbp.shp (polygon data). Each of these shapefiles contains political boundaries including town, county, and state. Check the metadata (the .html file) for the codes that allow you to symbolize the data.

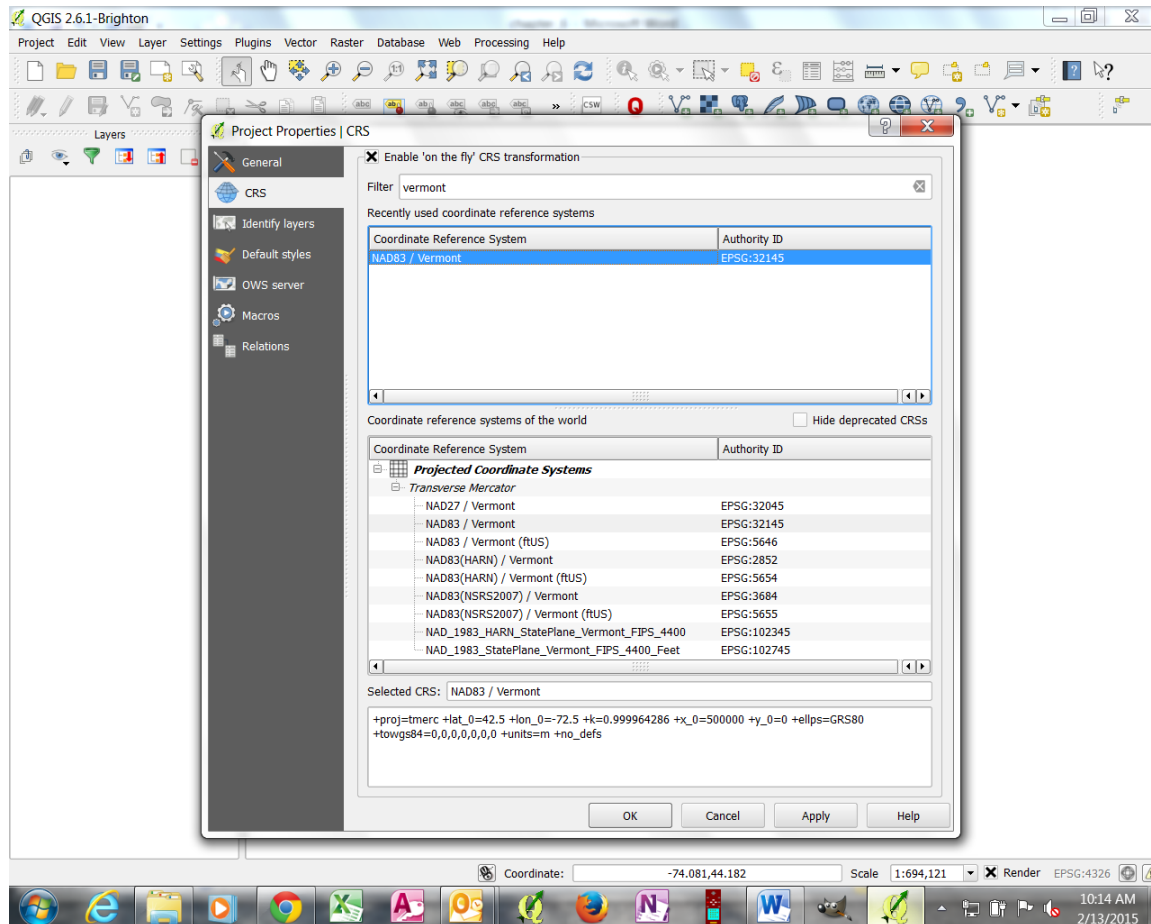
- h) Open QGIS



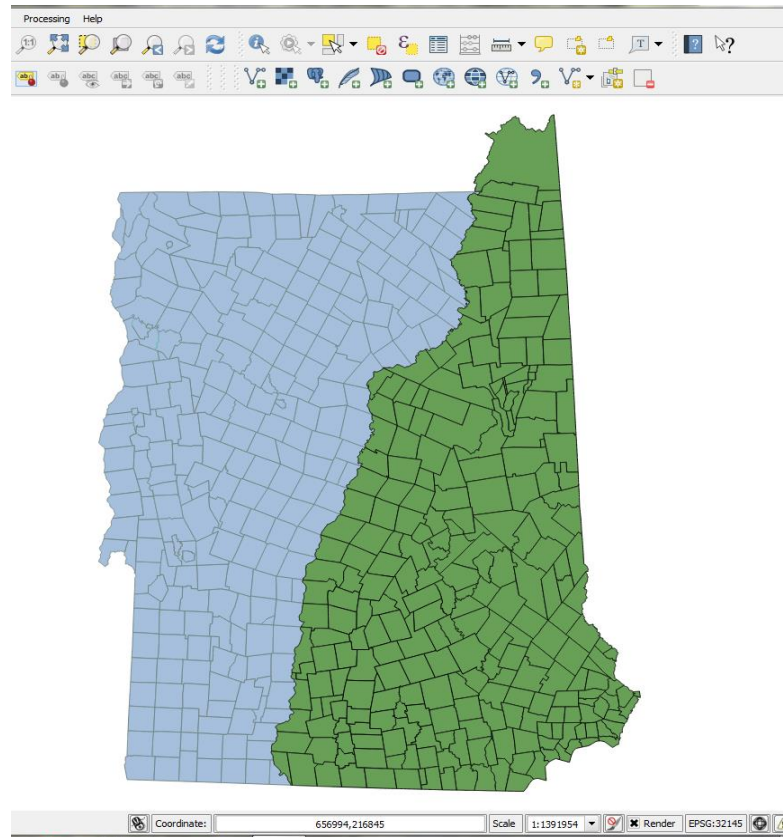
- i) Click on the Project menu and then Project Properties



- j) In the CRS tab, turn on “enable ‘on the fly’ CRS transformation”
- k) Type Vermont in the Filter
- l) Choose NAD83/Vermont and then click OK



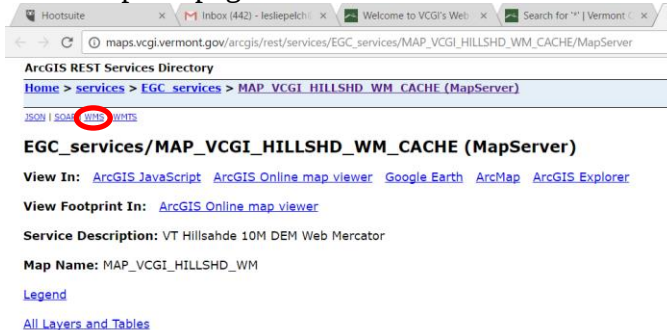
- m) One of the files associated with the shapefiles you downloaded from Granite is the .prj file (projection file) which basically says "this data is in NH State Plane Feet NAD 83 coordinate system" The software is able to do the mathematical transformation necessary to re-project the data into VT State Plane Meters. This is why the VT and NH Town boundaries will line up and scale properly.
- n) Setting the project to a particular projection and coordinate system also tells the software how to interpret measurements. If it knows the units of the coordinates are meters, it can calculate the distance and area measurements.



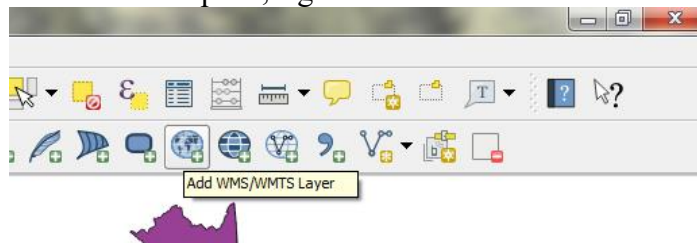
3. Accessing Data Via Web Map Service

VCGI, in collaboration with VT's Enterprise GIS Consortium (EGC), has developed a portfolio of "web services" which allow users to bring map layers and imagery into their GIS projects or map mashups without actually downloading any data. By connecting to the service of choice via the internet (a live internet connection is required in order for the images to appear in your GIS project) users can access a variety of resources. All available services can be found at the VT Open Geodata Portal, as described above. Here is more detail on how to connect.

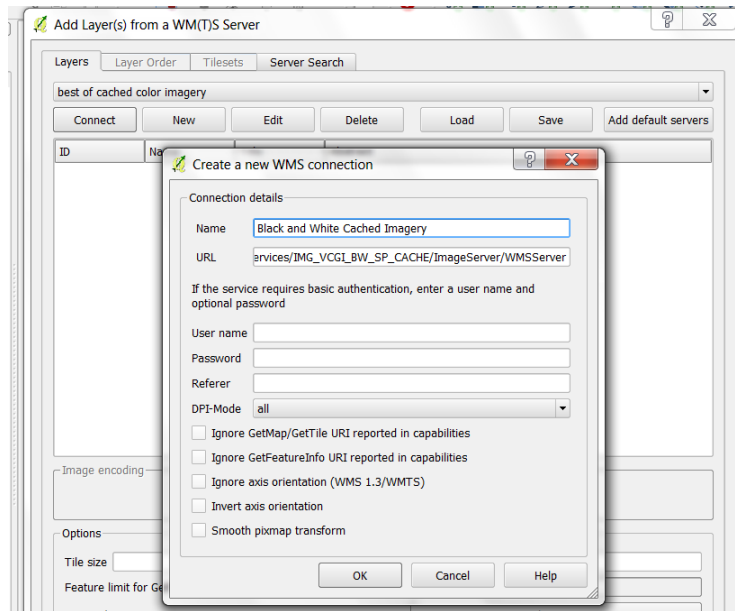
- a) Once you find the service that interests you, right click on the WMS link on the Rest Endpoint page.



- b) Choose “Copy Link Address” (Chrome) or “Copy Shortcut” (Internet Explorer) or “Copy Link” (Edge).
- c) Go back to QGIS and Click on the “Add WMS/WMTS Layer” button in QGIS or in the Browser pane, right click on WMS and click on “New Connection”

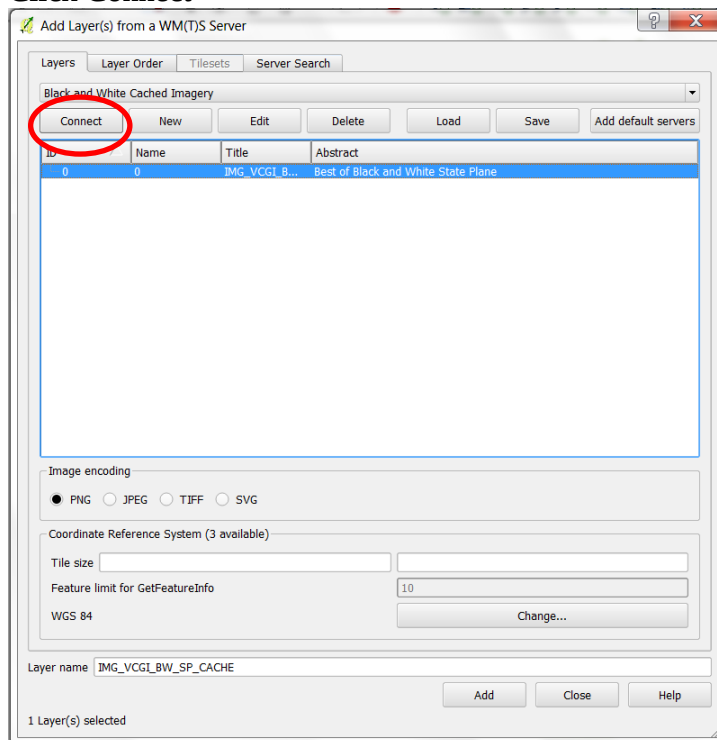


- d) Click on New in the “Add Layer(s) from a WM(T)S Server window”
- e) Paste the URL you copied into the appropriate box and type a name into the Name box in the Create a New WMS Connection Window



f) Click OK

g) Click Connect



h) Click on the service to highlight it

i) Type a name in the Layer Name box (or leave the default)

j) **Make sure that the Image encoding is set to PNG or JPG!** Leaving it set to TIF results in a washed out image.

k) Click Add then close the “Add Layer(s) from a WM(T)S Server” window.

NOTE: If you add the VCGI Basemap Service, you will be better off adding individual layers from it rather than trying to add the whole thing at once. For some reason, QGIS does not interpret such a complicated service well, and much of the information is lost.

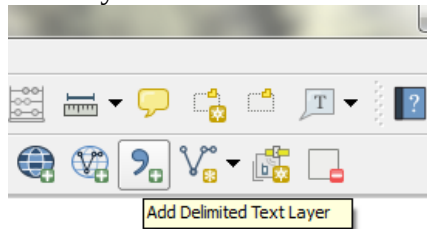
4. Data From a Text File or Spreadsheet Containing Coordinates

It is possible to use data from a delimited text file in a QGIS project if that data contains x, y coordinates of a known coordinate reference system (such as latitude longitude). If your data is stored in a spreadsheet, simply do a "save as" to create a .csv version of it (comma separated values).

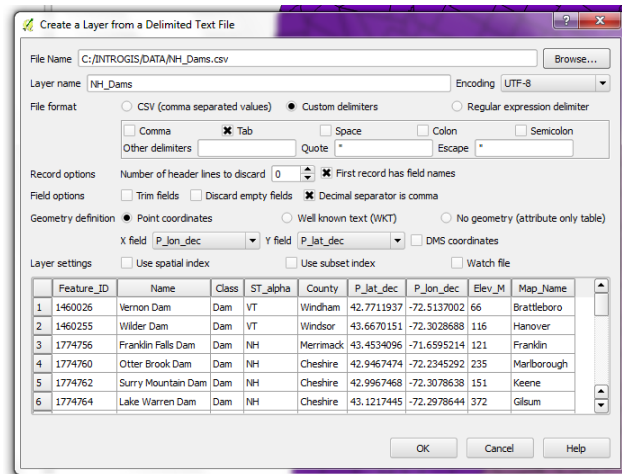
To view a delimited text file as layer, the text file must contain:

1. A delimited header row of field names. This must be the first line in the text file.
2. The header row must contain an X and Y field. These fields can have any name.
3. The x and y coordinates must be specified as a number. The coordinate system is not important, though you do need to know WHAT coordinate system it is.

We will use a function called "Add delimited text layer." You should see the following icon in your tool bar:

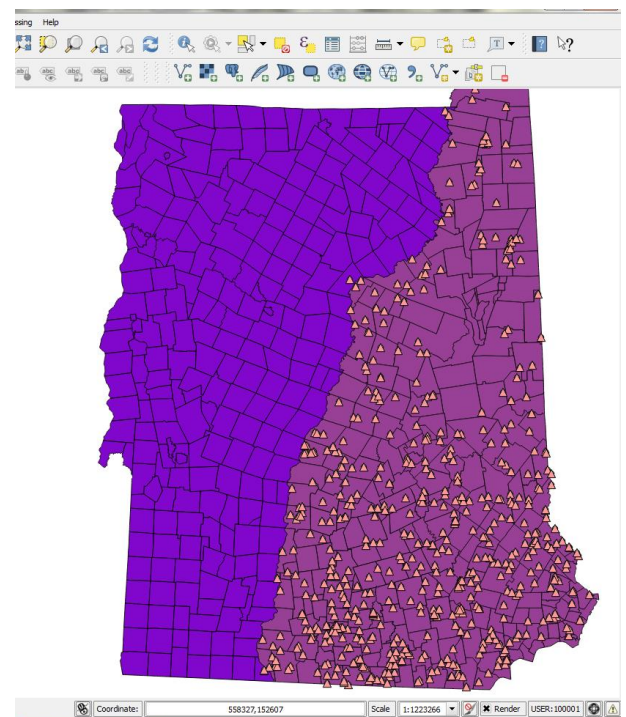


When you click on it a dialog box opens that allows you to indicate where your comma or tab delimited text file is located and which fields contain the latitude and longitude coordinates. You can use the file called "NH_DAMS.csv" found in your DATA directory or another file that your instructor may indicate.



Navigate to the location of the file, and make sure that all information in the dialog box is correct (is data delimited by tab or comma? Does the first row/record contain field names? Which fields contain the x/longitude info and y/latitude info?). The bottom window in the box will show what the data looks like based on the information you have provided, so you can make sure you have filled in the correct answers.

After you click on "OK" Quantum will ask you what Coordinate System the data is in. If you have latitude and longitude coordinates, choose from the "Geographic Coordinate Systems" and find "WGS 84." Click "OK" and the data should appear on your map in the correct area!



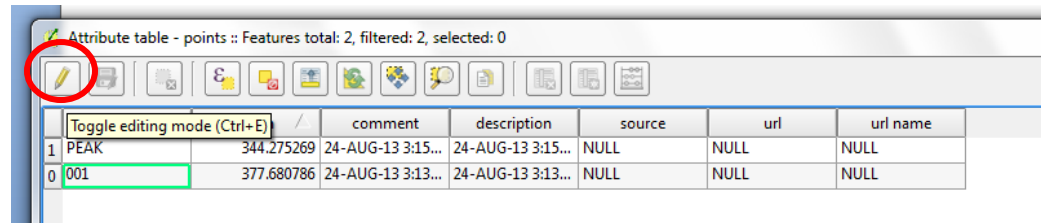
Editing, Joining, and Creating Data

1. Editing Data

We can edit our data if we need to add or change attributes or add or move features. We can only edit shapefiles, not .gpx or text file or joined data. You can always save your .gpx or text file or joined data out as a new shapefile, however, that you can then edit.

In order to edit the attributes of a data layer, we will highlight it, right click and then choose Open Attribute Table

- a) Click on the Toggle Editing icon (pencil) to turn on editing



- b) Once editing is turned on, you can click on any of the values in the attribute table and change them by double clicking and typing.
- c) Hold your cursor over the other icons at the top of the attribute table window to see what they do. Notice that you can add a new column and you can delete columns. You ARE changing the data when you perform these actions, so be careful!
- d) Click on the Field Calculator button to see that you can calculate values to populate new or existing fields. We'll do more with this function in the next section.
- e) When you turn editing off, you will be asked whether you want to save changes.

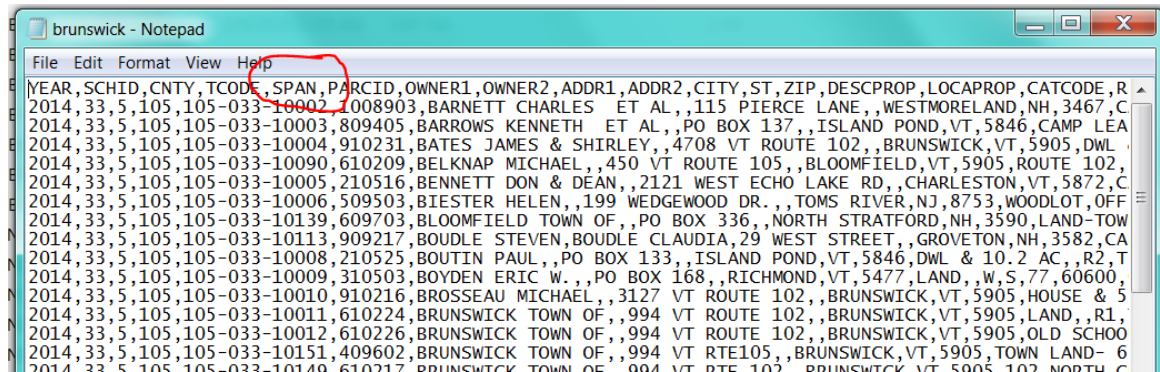
2. Joining Data from a Text File Via a Common Attribute

Another way to create geospatial data is to join tabular data to existing geospatial data. A great example of the usefulness of the join is when one has parcel boundary data and tabular grand list (assessment) data. We want to join that table to the existing parcel boundary data so that we can explore information like parcel size, value, value per acre, type of property (residential, commercial, etc.) and value reduction based on Use Value Appraisal status (Current Use).

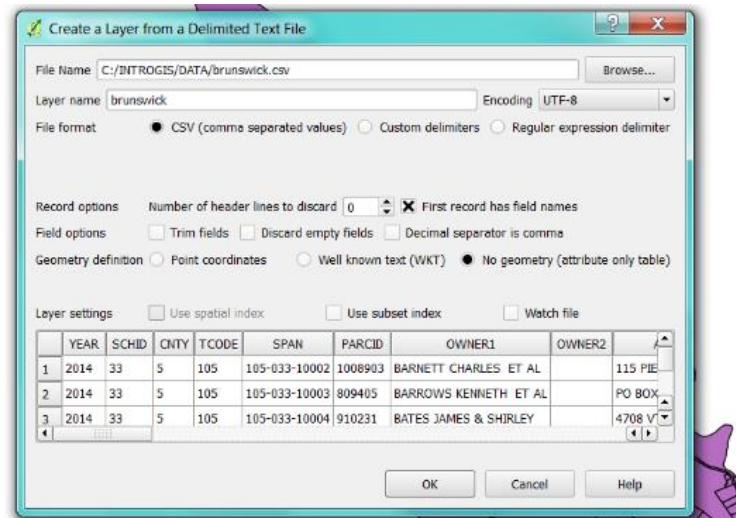
Before considering any table joins, the most vital aspect to consider is the JOIN FIELD. The join field must contain matching values (in the same format, such as STRING or INTEGER) in both the source table (grand list) and the target layer (parcel boundaries).

In our example, we have the SPAN or School Parcel Accounting Number – an identifier that is unique statewide and located in both the grand list and the parcel data for Brunswick.

- a. Let's start by looking at the two files we will join. Open the Brunswick.csv file in a text editor like Notepad. We can see that one of the columns is named "SPAN"

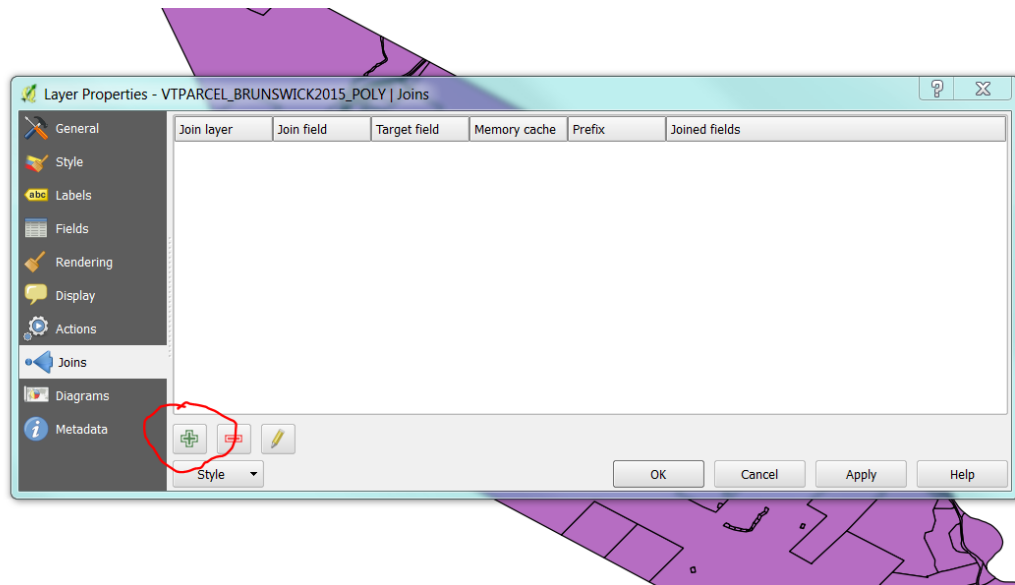


- b. Open or create a project containing the VT_BRUNSWICK2015_POLY parcel data. Take a look at the Attribute table to make sure it contains SPAN (the join attributes do not need to have the same name, but they do need to contain the same type of data and have some matching values!).
- c. Click on the Add Delimited Text Layer button and choose **brunswick.csv**, the Brunswick grand list.
- d. Make sure that you set the File Format to CSV (Comma Separated Values) and that you check the No Geometry (attribute only table) option

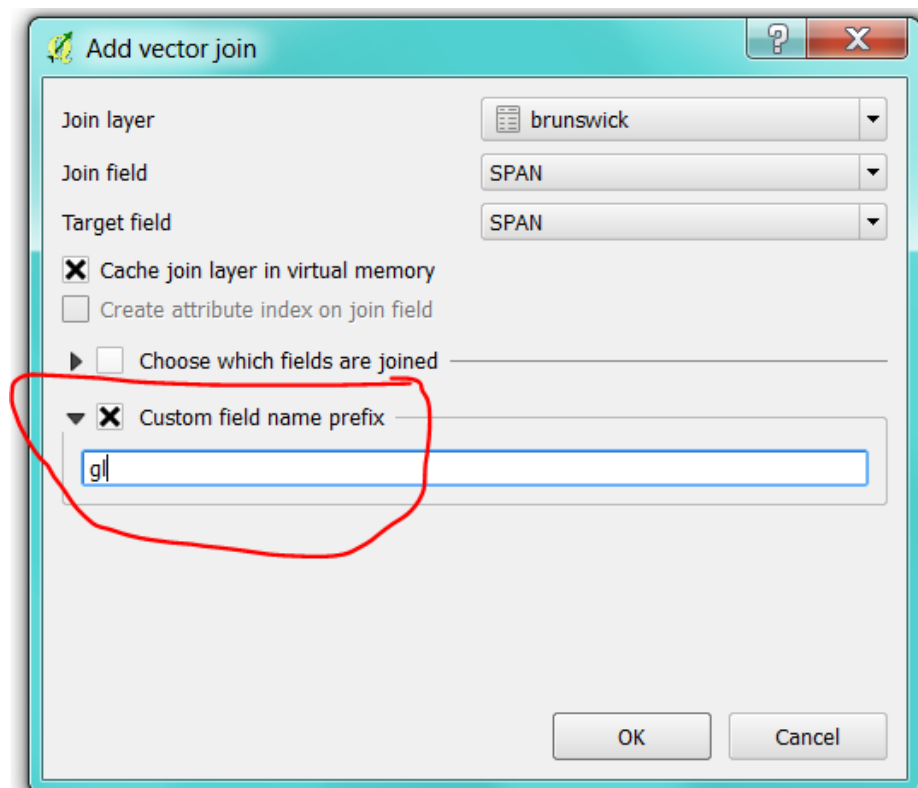


- e. Click OK, this will bring the grand list table into your project (You won't see anything change on the map, since this is simply tabular data)
- f. Right click on the Brunswick parcel boundaries layer in the layer list and choose Properties
- g. Choose the Joins section
- h. Click on the green Plus sign near the bottom of the dialogue box
- i. Choose the correct Join Layer (the CSV file you just brought in)

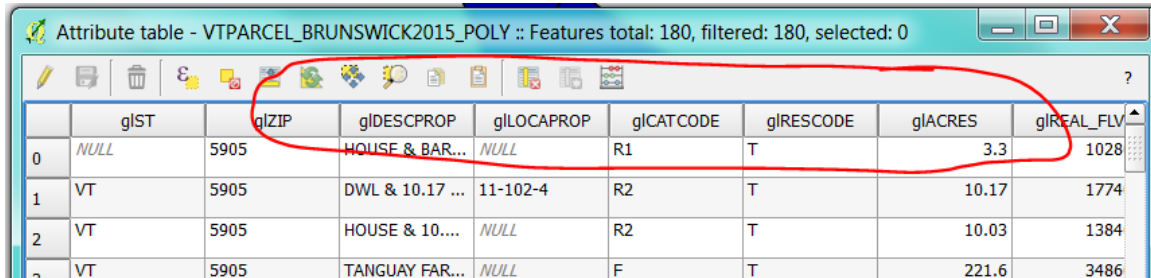
- j. Make sure the Join field and the Target field are correct and match (SPAN)



- k. Click to put a check in the box to create a Custom field name prefix and type in something that is only one or two letters like “gl” for grand list.
l. Click OK and OK again



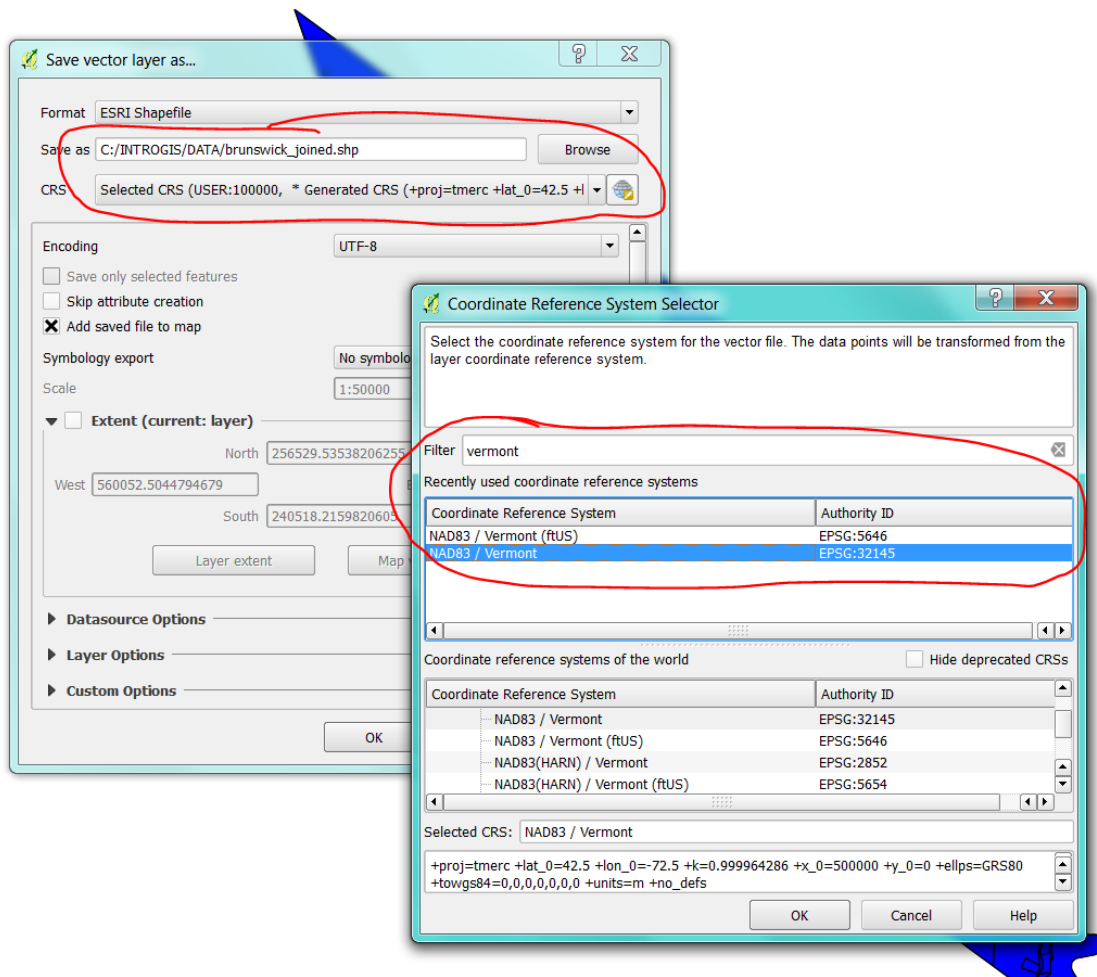
- m. Open the attribute table to see if the new set of attributes were successfully joined to your Town Boundary layer (look for the custom prefix you chose as well).



	glST	glZIP	glDESCPROP	glLOCAPROP	glCATCODE	glRESCODE	glACRES	glREAL_FLV
0	NULL	5905	HOUSE & BAR...	NULL	R1	T	3.3	1028
1	VT	5905	DWL & 10.17 ...	11-102-4	R2	T	10.17	1774
2	VT	5905	HOUSE & 10....	NULL	R2	T	10.03	1384
3	VT	5905	TANGUAY FAR...	NULL	F	T	221.6	3486

A few things to note about this joined data:

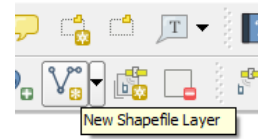
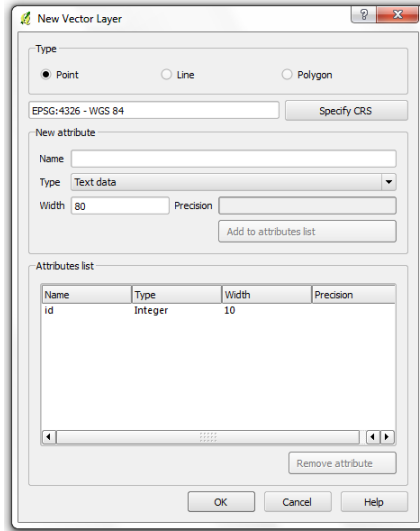
- The join only exists in this project. ***We have not edited or changed the table or the shapefile in any way.*** If you were to create a new project with this same shapefile, you would NOT automatically see the grand list information in the attribute table. When you save this project and open it again later you WILL see that the join persists.
- In order to save the join in a more permanent and transferable manner we can simply create a new shapefile using “Save as”
 - n. Right click on the name of the parcel shapefile in the list of layers
 - o. Choose “Save as”
 - p. Set the file type (to shapefile or perhaps GeoJson, an open source format) as well as the name and location where you would like to save the shapefile.



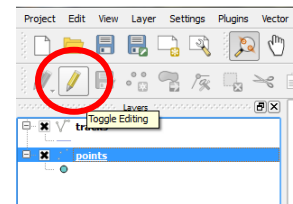
- q. Make sure to set the CRS as appropriate – we will make sure that any data we create is VT State Plane Meters, so we click on the icon to the right of the default CRS, type “vermont” (no quotes) in the filter, click on the Vermont State Plane Meters CRS (EPSG code 32145), and then click on OK.
- r. Click OK again and after a moment you will see that the new shapefile has been added to your project.
- s. Open the attribute table to confirm that the joined data is in the attribute table.

3. Creating New Vector Data

- a. Click on the New Shapefile Layer button



- b. Set the type of data you wish to create: Point, Line, or Polygon
- c. Specify the Coordinate Reference System
- d. Add attribute fields, choosing their specifications carefully (you might include attributes like name, type, size...whatever information you would want to include in an attribute table)
- e. Click OK when you are done creating your shapefile
- f. Save your new shapefile to an appropriate location
- g. Zoom to a level appropriate to the data you wish to create
- h. Click on the Toggle Editing button to turn editing on.
- i. Click on the Add Feature button next to the editing toggle and then start clicking on the map to add features.
- j. If you are creating point features you will be given the opportunity to fill in the attribute fields each time you click to add a point.
- k. If you are creating line or polygon features, you will left click to create vertices (points where your straight line can change direction) and you will right click to indicate you are done with that particular feature, and then you will be able to edit the attributes.
- l. Once you are done creating features, simply stop editing and say that you want to save the changes!



Online Resources:

The most important Vermont “other data source” and technical advice resource: Regional Planning Commissions: <http://www.vapda.org>. Other states have organizations similar to VCGI, and many federal agencies make GIS data available for download or via web map service. You will find links to their contact information as well as links to other GIS organizations here: <http://vcgi.vermont.gov/resources/othergis>