

# Exercise 4 – Focus on Data

**Objectives:** As the title suggests, this section requires you to get involved with the data side of GIS operations, in all its various forms. Accurate, stable data is the bedrock on which insightful analysis is produced. In this exercise we'll learn about...

- QGIS Data Management functionality
- Using the Plugins manager to access online earth imagery data (openLayers)
- Organising and preparing non-spatial table data and bringing it into QGIS

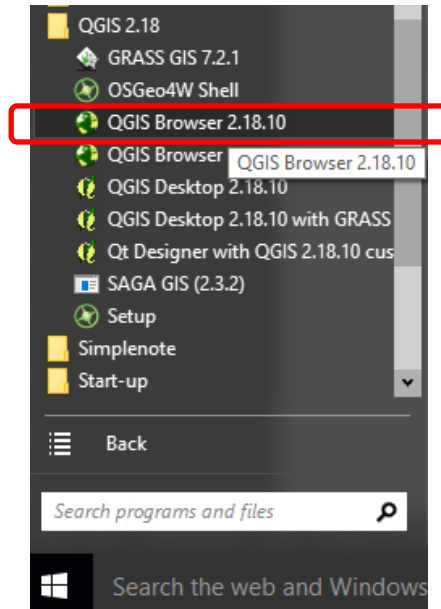
**Please ask if there is anything you don't understand!**

## Part 1 – Opening QGIS Browser

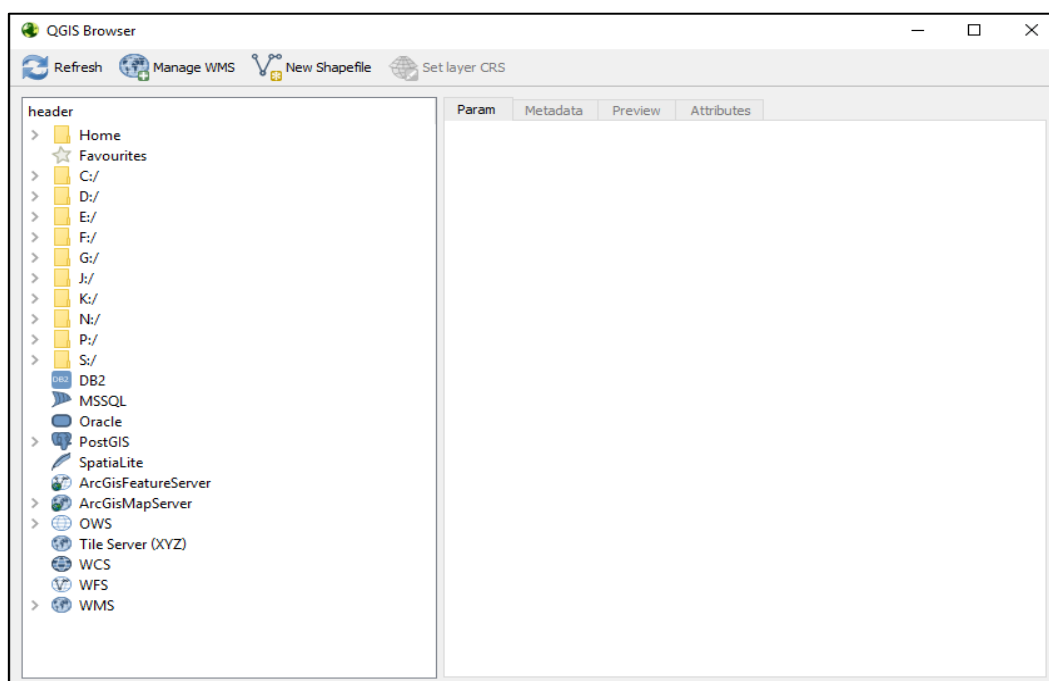
QGIS comes with a standalone application called QGIS Browser that is useful for managing and viewing your spatial data's metadata

- 1) Open QGIS Browser using the windows start menu

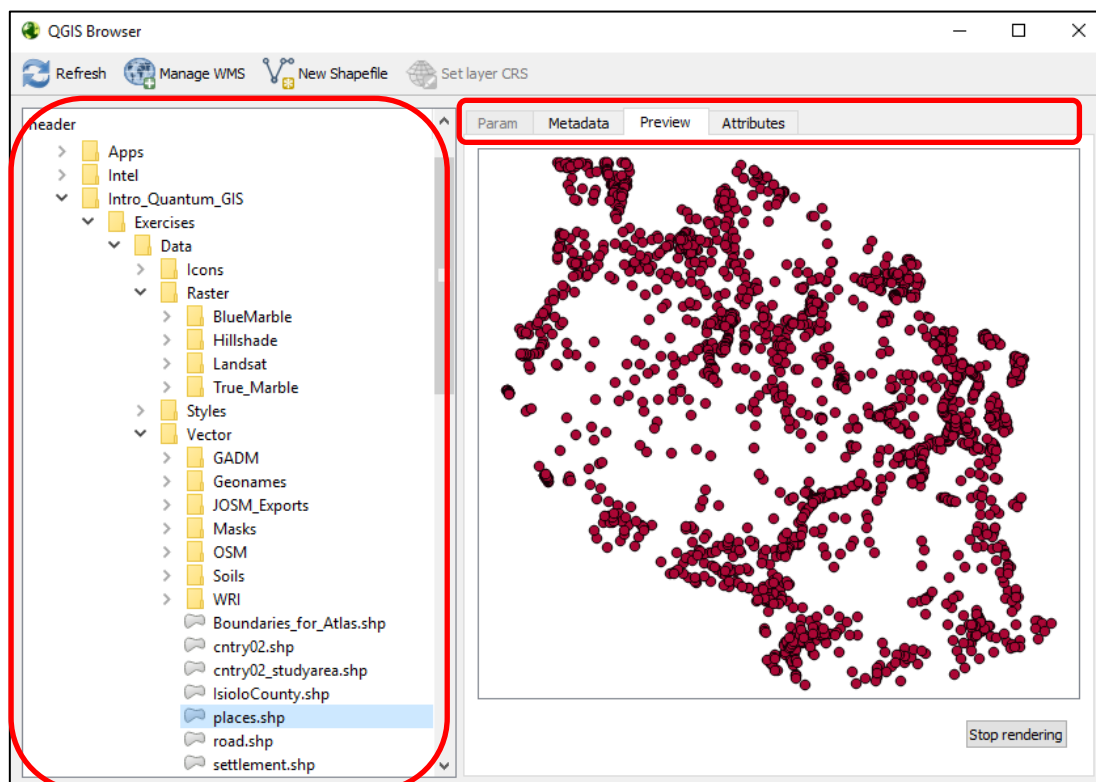
**Start > All Programs > QGIS 2.18 > QGIS Browser 2.18.10** Please note the name of the browser will vary according to your installed version of QGIS



- The browser should appear as shown below



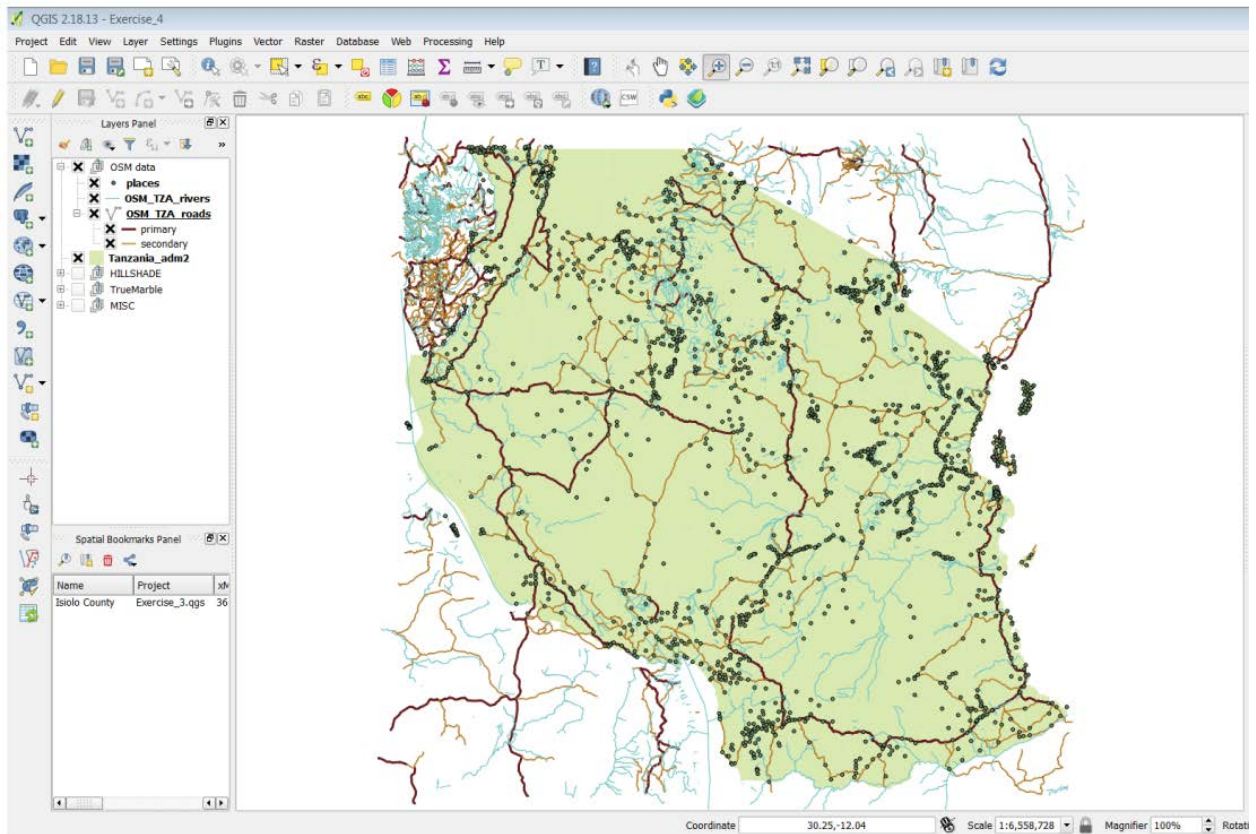
- Within the file explorer panel navigate to:  
**C:\Intro\_Quantum\_GIS\Exercises\Data\Vector**
- Select **places.shp** and then step through the 3 accessible tabs: **metadata**, **Preview** and **Attributes**; this is very useful outline information for datasets. QGIS Browser is particularly useful if you have a large volume of data to review.



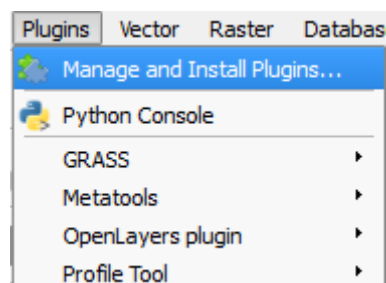
- Try using the browser to explore some of the other vector and raster datasets.

## Part 2 – The Open Layers Plugin

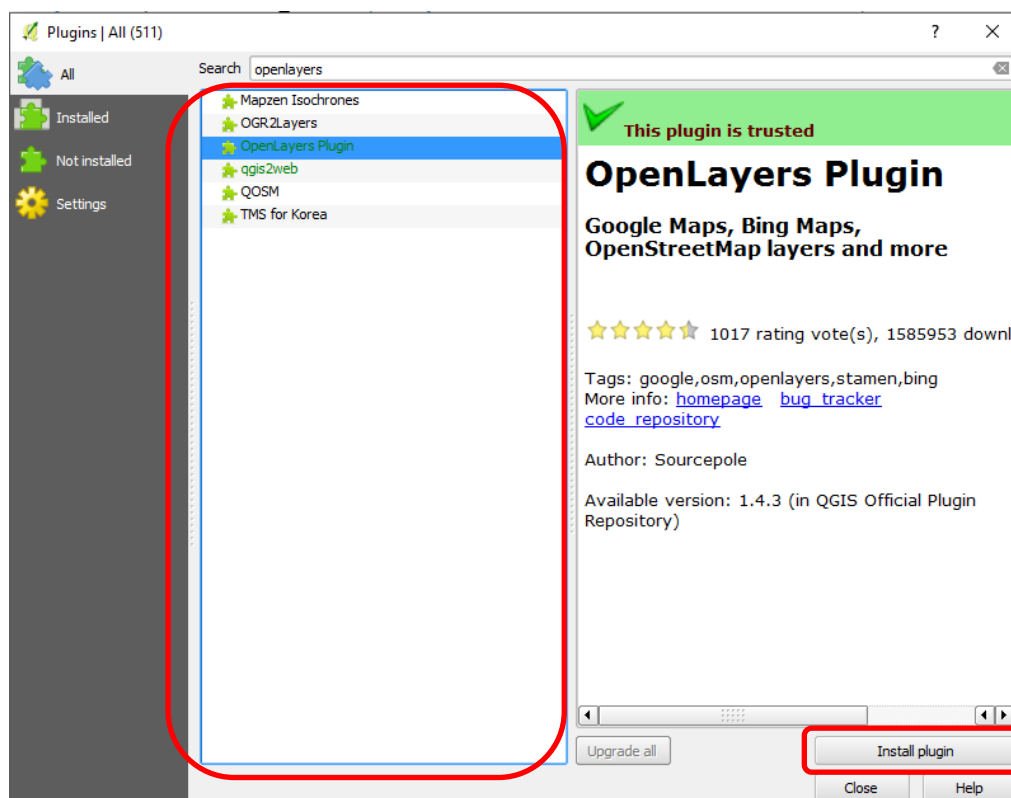
- Returning to QGIS desktop, open Exercise\_4.qgs from C:\Intro\_Quantum\_GIS\Exercises\.
- You are looking at Tanzania with some settlement, road and river data, sourced from Open Street Map.



- We will now add some background mapping to give the data some context. We will do that using the **OpenLayers** plugin
- From the **Plugins** drop-down menu click **Manage and Install Plugins**



- Type **OpenLayers** into the search bar

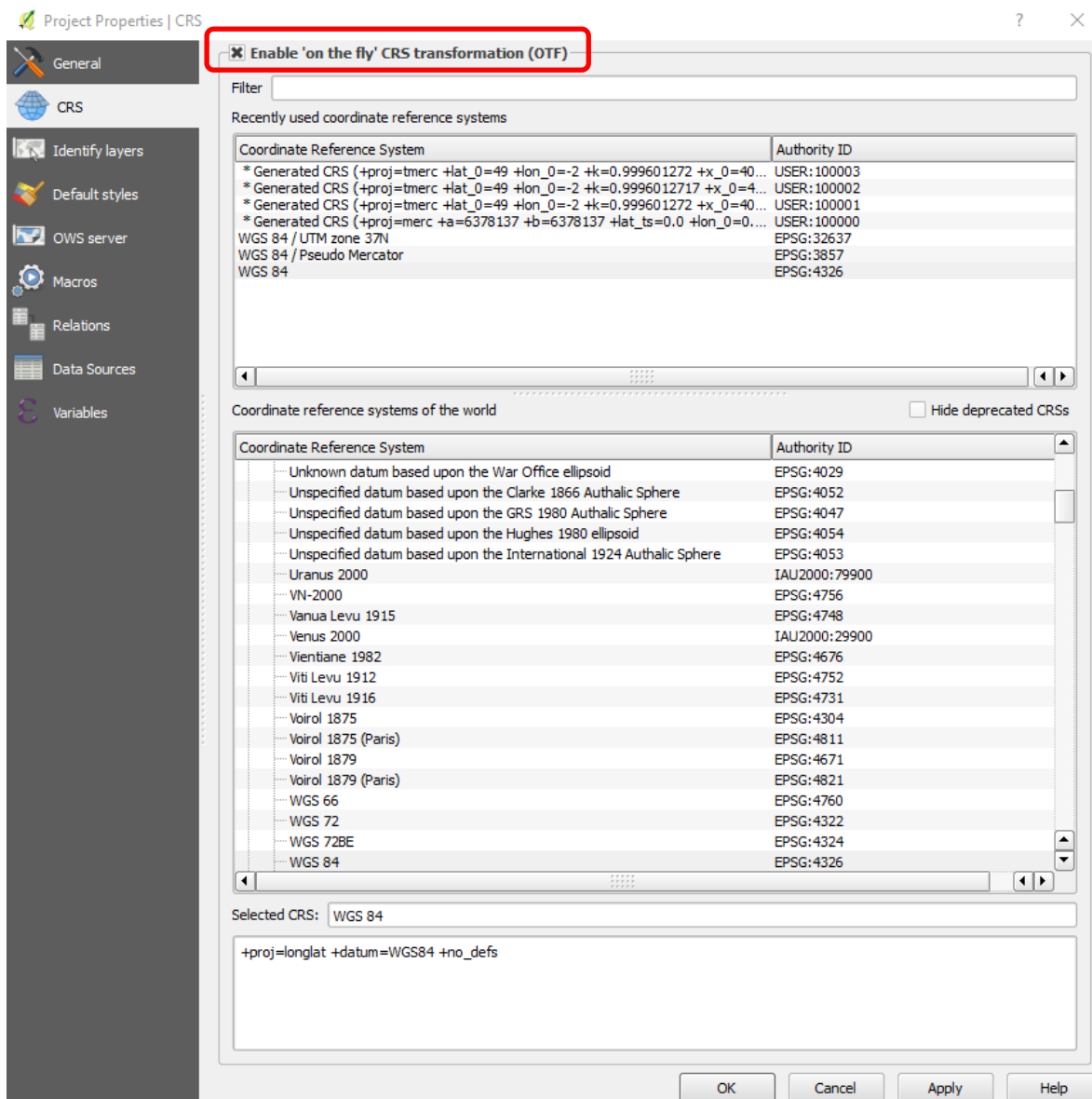


- Select the **OpenLayers Plugin** and then click install plugin
- QGIS will download and install the plugin
- Close the **Plugins** window

### Part 3 – Configuring coordinate reference systems to accurate display data

We must now ensure that 'on the fly' CRS transformation is enabled

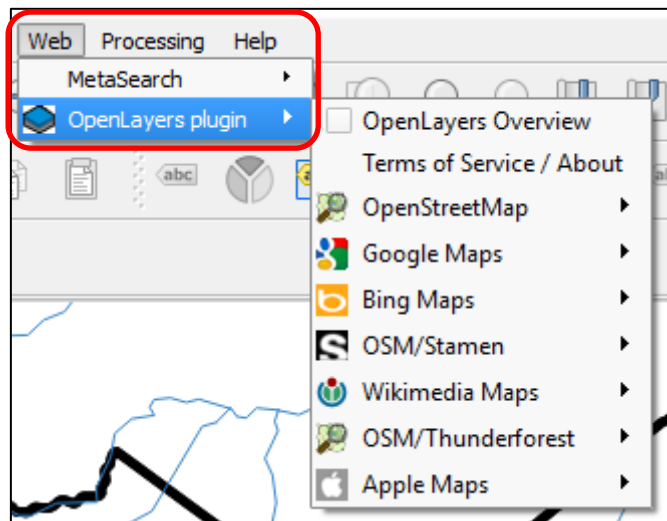
- From the **project** menu select **project properties**
- Click on the **CRS (Coordinate Reference System) tab** and make sure that **Enable 'on the fly' CRS transformation** is ticked.



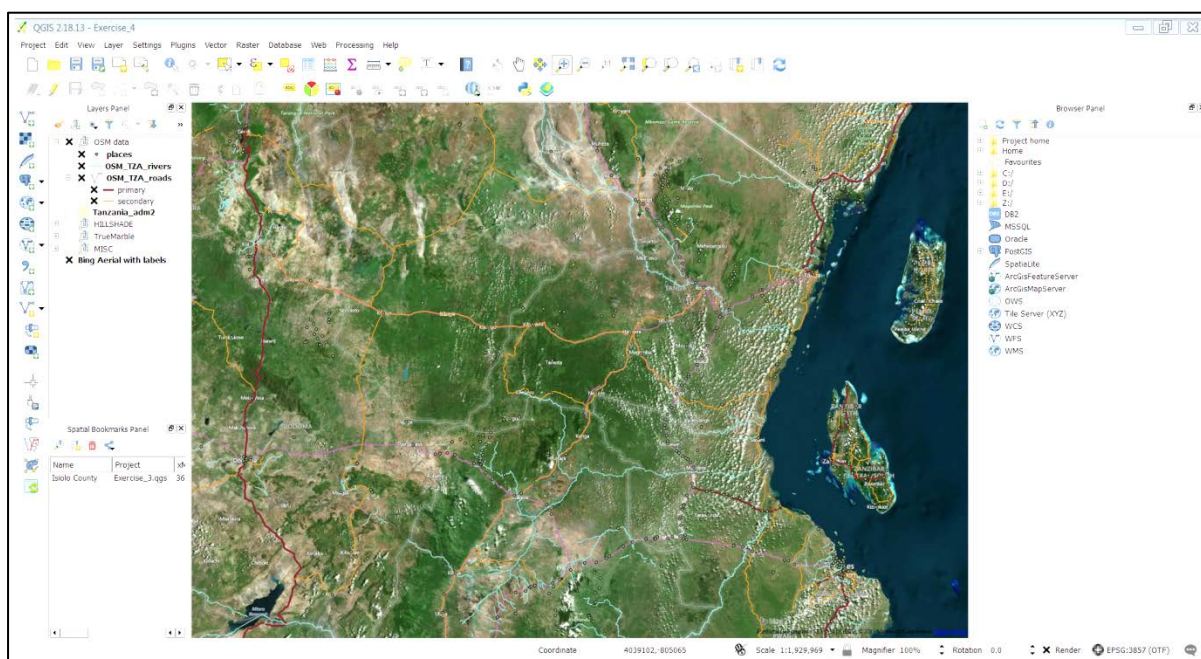
- This needs to be ticked because you are about to add some background mapping to your QGIS project which uses a different coordinate system to the data which is currently loaded.
- Press **OK**.

## Part 4 - Display OpenLayers background mapping.

- From the **web** menu expand the **OpenLayers Plugin**



- You can choose to display any of the available background maps; try *Google Hybrid* or *Bing Aerial with Labels*
- The imagery will be streamed into your project as a new layer, the speed of rendering will depend on your internet connection speed
- You may need to move the new layer around in the layers panel to ensure the drawing order is correct.
- Try panning and zooming around your map
- remember to turn layers on and off as you require; having too many background mapping layers turned on at once may affect performance and rendering speed.





## Challenge:

Return to the **Manage and Install Plugins** window (**Plugins** drop-down menu). If permissions for your PC enable you to do so, search for and Install **QuickOSM**; Once installed, go to the **Vector** drop-down menu and select **Quick OSM**. Try running a query to download some Open Street Map data.

Explore other plugins that are available and relevant to your interests.

## Part 5 – Data preparation: joining population data to boundaries


Both this section and the following section look at two types of raw table data which are commonly used in GIS analysis. Both data types require some preparation before GIS analysis can take place; we will now go through these processing steps.

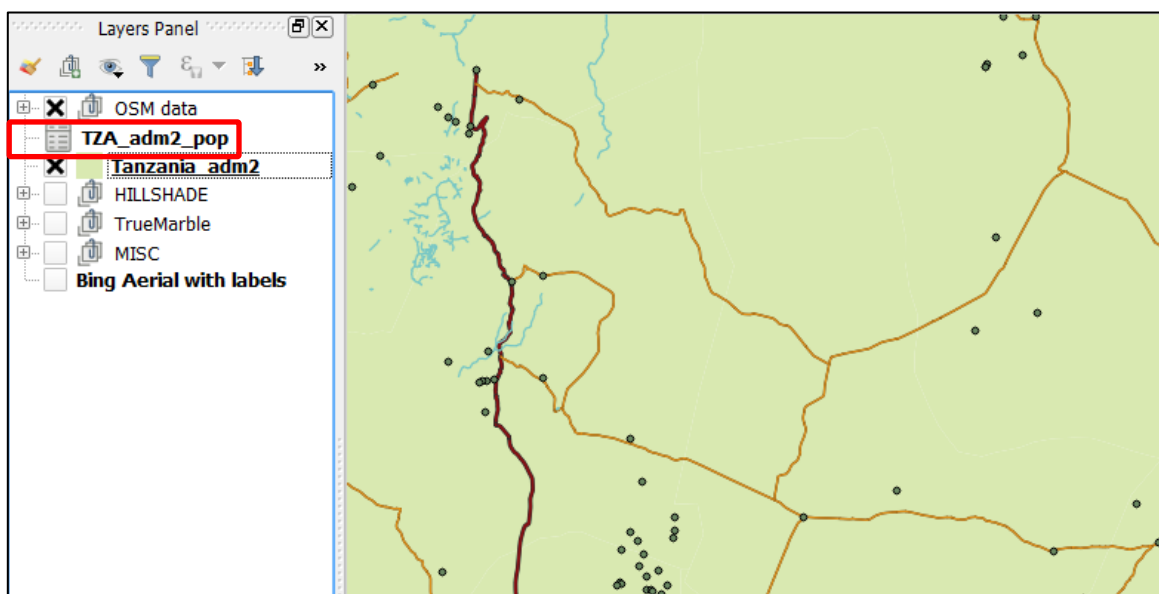
In this section we will run through the process of joining non-spatial population and demographic data for districts in Tanzania to corresponding spatial data features, in this case district boundaries for Tanzania. The data originates from the United States Census Bureau and is freely available from the DHS Spatial Data Repository (<https://spatialdata.dhsprogram.com/home/>).

1. In **Windows Explorer** Locate the following CSV data file and open it in **Microsoft Excel or Notepad**; then inspect the data:

`C:\Intro_Quantum_GIS\Exercises\Data\Tables\TZA_adm2_pop.csv`

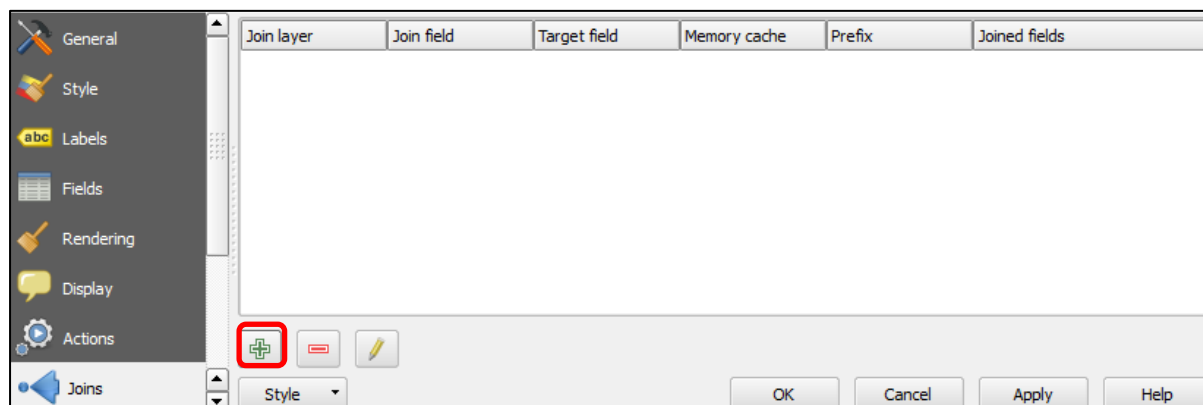
Now return to QGIS and locate the dataset *Tanzania\_adm2* in your **Layers Panel**. Right-click the layer and **Open Attribute Table**. Looking at this boundary data and the population data, can you identify a common field to both tables on which we could join them? They both have one field named **GEO-MATCH**. You will now join these two tables based on this field.

2. Bring the population data *TZA\_adm2\_pop.csv* into QGIS.
  - Locate the CSV file, via QGIS Browser or Windows Explorer, and drag it into the map window
  - Notice the alternative icon for table data (sometime known as *non-spatial data*) 

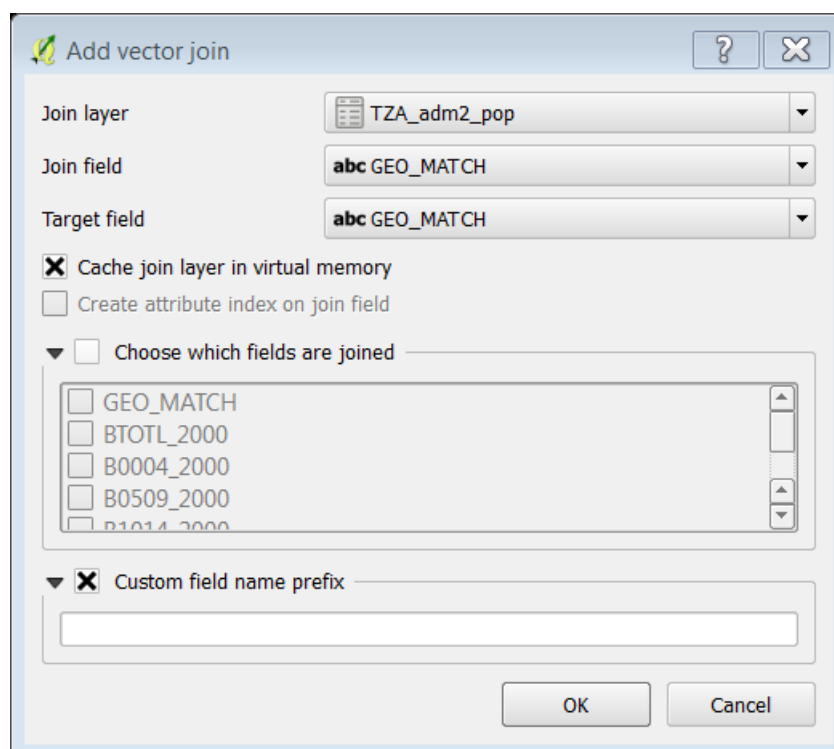


### 3. Configure the Join

- Right-click the spatial boundary dataset **Tanzania\_adm2** and click **Properties**; these are Tanzanian administrative boundary data from *United States Census Bureau*
- Click on the **Joins** tab.
- Click on the green plus sign

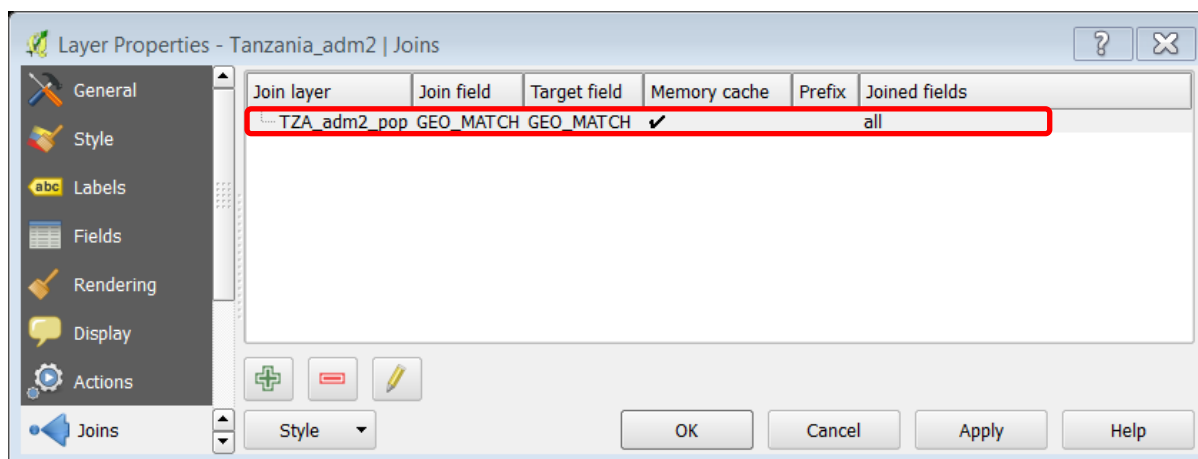


- A new window will appear.
- Select *TZA\_adm2\_pop* as the Join Layer.
- Select *GEO\_MATCH* as the Join and Target field
- Tick **Custom field name prefix** and delete all text (N.B. if this text is not deleted, each new joined field will have the prefix '*TZA\_adm2\_pop\_*' – this is a nuisance!)



- Press **OK**.
- You will now see that the Layer properties window shows you that *TZA\_adm2\_pop* has been joined to *Tanzania\_adm2*





- Press **OK**.

#### 4. View your join

- Open the attribute table for *Tanzania\_adm2* and ensure that the population data has successfully joined the boundary data
- This join is temporary and can be deleted at any time by clicking the 'red minus' in the Joins tab of the layer **Properties**.

#### 5. Save your join as a new shapefile dataset.

- Right click on *Tanzania\_adm2* and click **Save as**
- Save it as an **ESRI shapefile**
- Save it in this location: C:\Intro\_Quantum\_GIS\Exercises\Data\Output\_data
- Call it *Tanzania\_adm2\_pop*
- Tick **Add saved file to map**
- Leave all other settings as they are
- Click **Ok**

#### 6. View attributes for your new shapefile.

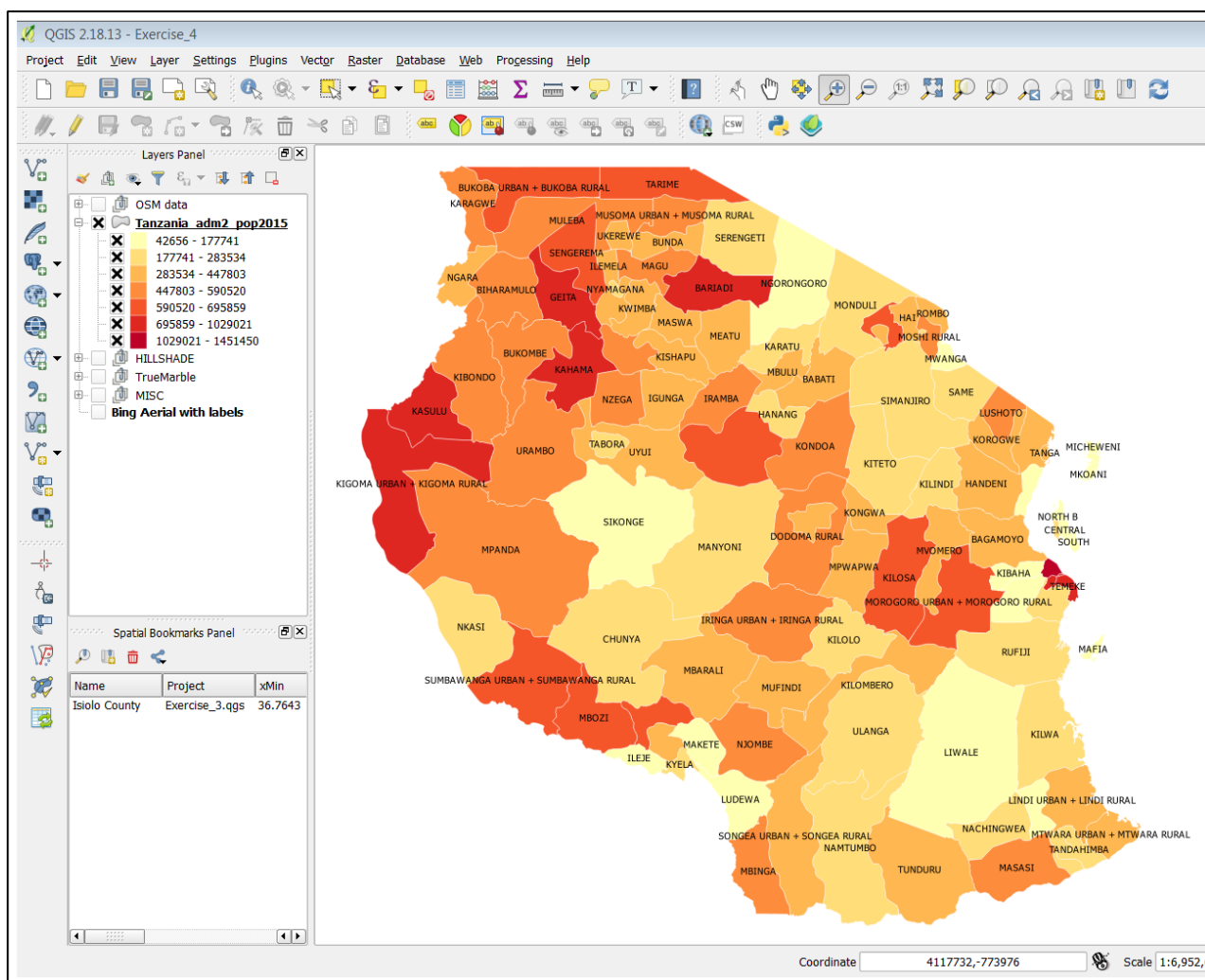
- Open the attribute table for this new layer. Notice it holds all of the population and demographic data from join you created.
- Open up the properties for this new layer and look at the **Joins** tab. You will see that there is no join. This means that you have now made a permanent dataset containing the attributes of both *Tanzania\_adm2* and *TZA\_adm2\_pop*.

#### 7. Remove the original pre-joined tables

- Right-click on *Tanzania\_adm2* and click **Remove**.
- Right-click on *TZA\_adm2\_pop* and click **Remove**.

### Challenge

Refer back to exercise 3 and style your data with a graduated symbology, to represent variable population counts in counties of Tanzania. You can choose to symbolise any of the population age group fields. Also, see if you can display labels showing the admin level 2 names. Your map may look something like the illustration below... (*Note: before attempting this, to simplify the map, untick the checkbox next to group layer **OSM***)



## Part 6 – Data preparation: Plotting X,Y coordinate locations from a text file.

We will now look at the second type of raw table data that is typically distributed for use in GIS analysis. You will be presented with a table containing the coordinate locations of Tanzanian Health Facilities. The data is freely available from the Humanitarian Data Exchange ([humdata.org](http://humdata.org)).

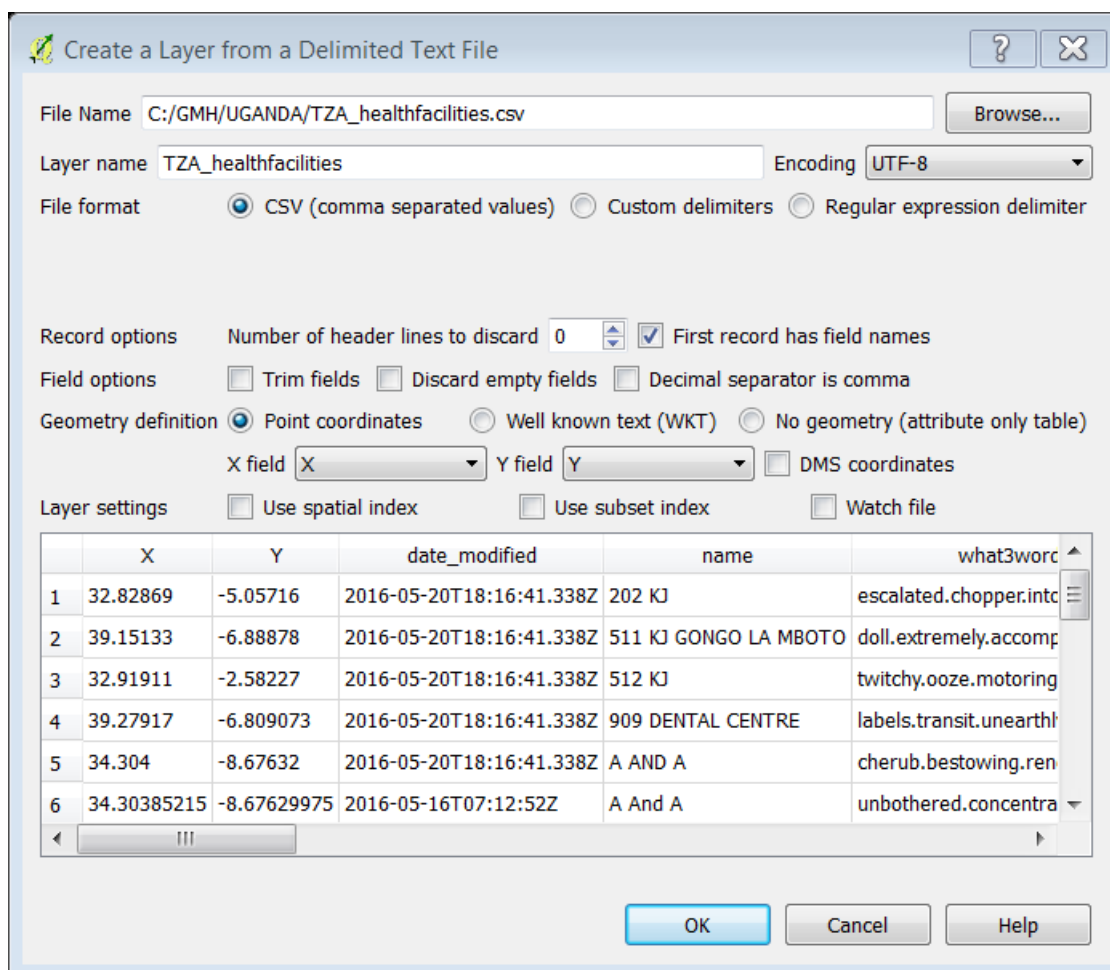
1. In **Windows Explorer** Locate the following CSV data file and open it in **Microsoft Notepad**; then inspect the data:

*C:\Intro\_Quantum\_GIS\Exercises\Data\Tables\TZA\_healthfacilities.csv*

- As in the previous section, the data format is *Comma Separated Values (.csv)*. It is important to understand the structure of this raw data as it is commonly used for sharing and distributing raw data. In Notepad, notice how each bit of data is separated by a comma. This type of format is known as delimited text format. Delimited text files typically use commas to structure the data, however other delimiter characters are also in use, such as space or tab. Provided that the delimiter is applied consistently, the data can then be formed into a table with rows and columns within a GIS or any other table-based software.
- Close Notepad.

## 2. Add this .csv file to QGIS

- Returning to the current QGIS project, start by turning off labels for the layer Tanzania\_adm2\_pop, to avoid confusion
- On the **Manage Layers** toolbar, click on the **Add Delimited Text Layer** button.
- Click on **Browse** and add:  
*C:\Intro\_Quantum\_GIS\Exercises\Data\Tables\TZA\_healthfacilities.csv*
- In the file format section, ensure **CSV** is selected
- Set the **X** and **Y** fields are correctly set
- The window should look like this:



**Create a Layer from a Delimited Text File**

File Name:

Layer name:  Encoding:

File format: ☒ CSV (comma separated values) ☐ Custom delimiters ☐ Regular expression delimiter

Record options: Number of header lines to discard:  ☒ First record has field names

Field options: ☐ Trim fields ☐ Discard empty fields ☐ Decimal separator is comma

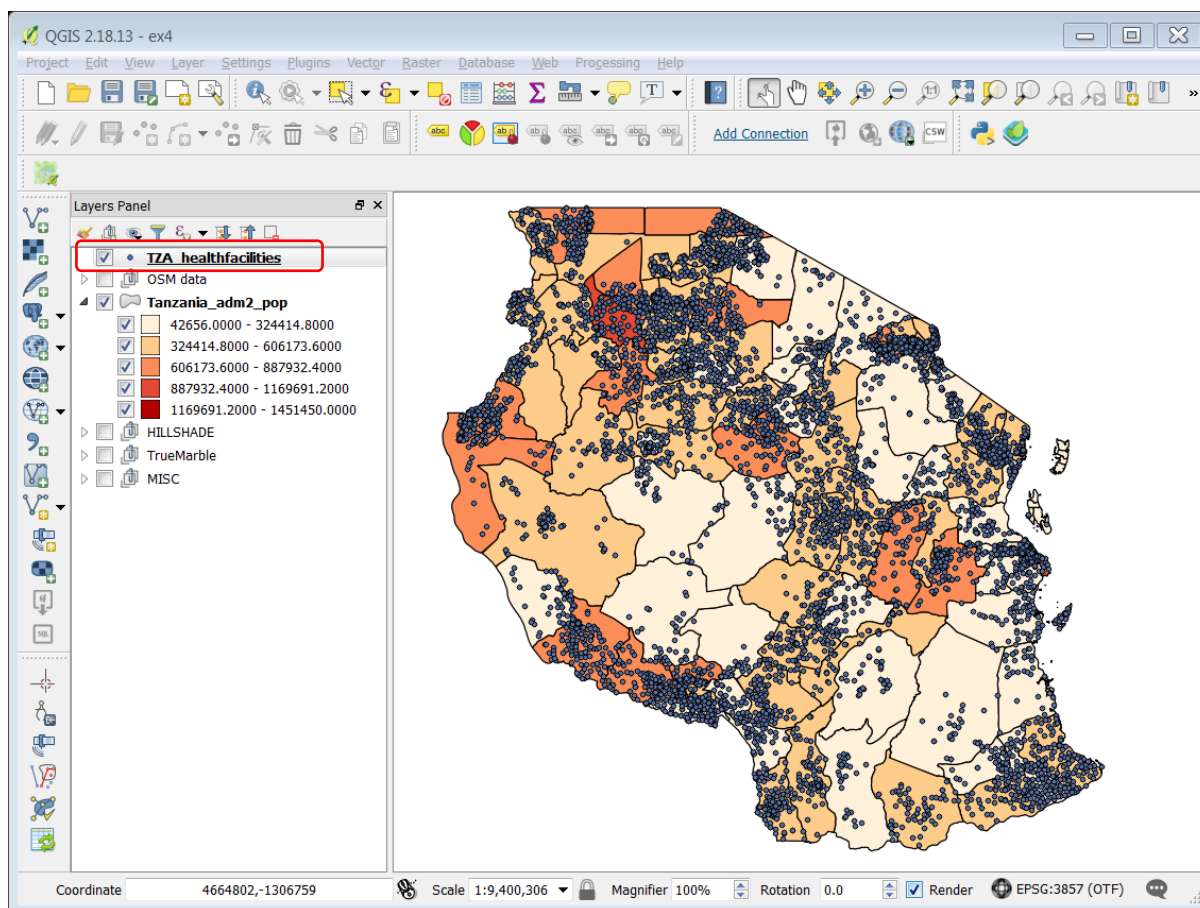
Geometry definition: ☒ Point coordinates ☐ Well known text (WKT) ☐ No geometry (attribute only table)

X field:  Y field:  ☐ DMS coordinates

Layer settings: ☐ Use spatial index ☐ Use subset index ☐ Watch file

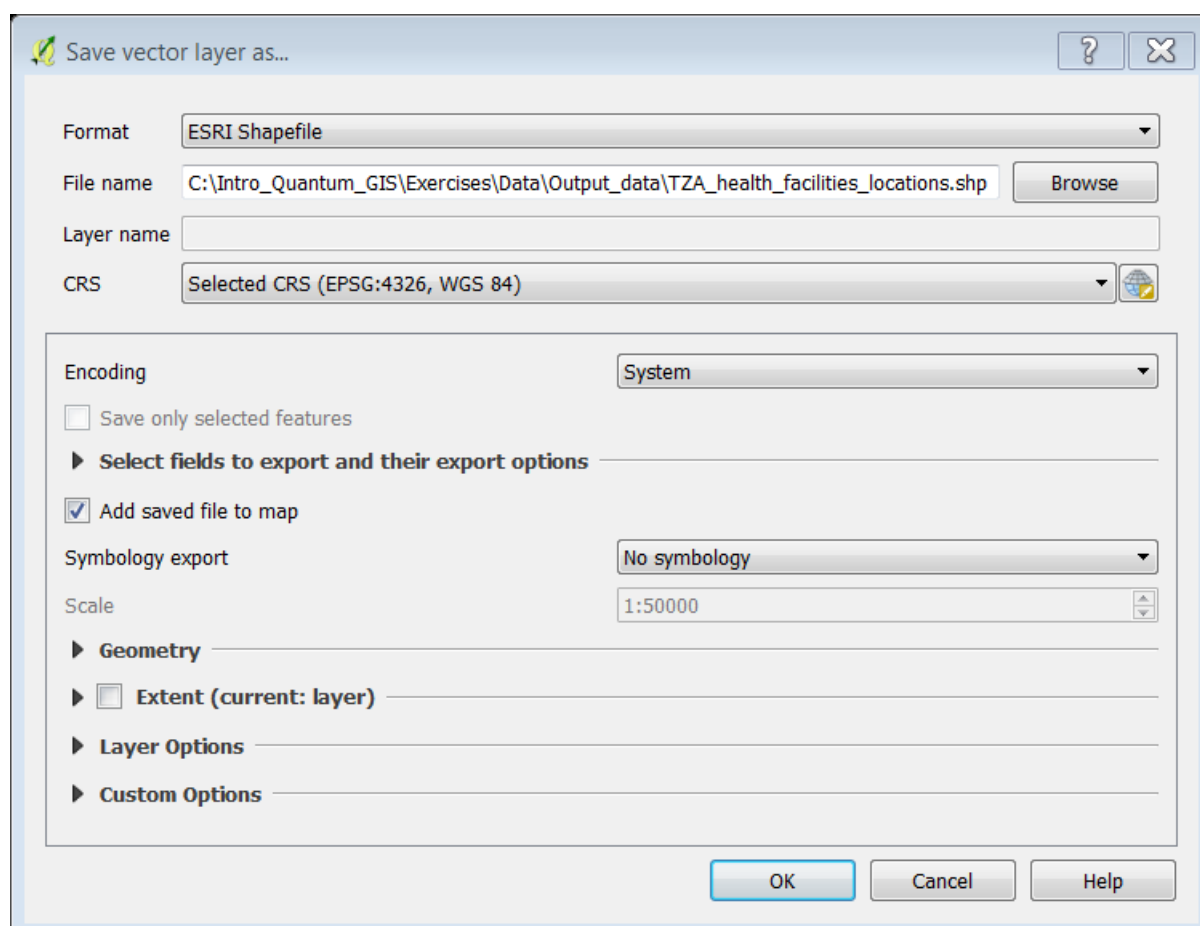
	X	Y	date_modified	name	what3word
1	32.82869	-5.05716	2016-05-20T18:16:41.338Z	202 KJ	escalated.chopper.intc
2	39.15133	-6.88878	2016-05-20T18:16:41.338Z	511 KJ GONGO LA MBOTO	doll.extremely.accomp
3	32.91911	-2.58227	2016-05-20T18:16:41.338Z	512 KJ	twitchy.ooze.motoring
4	39.27917	-6.809073	2016-05-20T18:16:41.338Z	909 DENTAL CENTRE	labels.transit.unearthl
5	34.304	-8.67632	2016-05-20T18:16:41.338Z	A AND A	cherub.bestowing.ren
6	34.30385215	-8.67629975	2016-05-16T07:12:52Z	A And A	unbothered.concentra

- You will notice that the table preview section at the bottom of the window changes if you change the options
- Press **OK**.
- If a window pops up asking you to select a Coordinate Reference System, select 'WGS 84 / EPSG 4326'. (N.B. use the **filter** input box to help you search)
- Press OK.
- You will notice that a number of points have now appeared on your map canvas and *TZA\_healthfacilities* point layer is visible in your Layers panel.



- This point layer is a temporary layer which is taking all of its data straight from the text file. It is not yet a permanent geometry layer.
- You can confirm this by right clicking on the layer in the Layers window and selecting **Properties**. If you look at the **Metadata** tab, the 'Source for this layer' is still pointing to the text file and has a line of code there to tell it to plot the points from scratch every time you open the QGIS project.
- This temporary plot of the locations has limitations for further analysis. It is really only intended for temporary visualisation and typically a GIS analyst would save the locations into a standalone spatial dataset, e.g. a shapefile; we will do this next.
- Now close the Layer Properties window.

3. Save your temporary table into a permanent spatial dataset.
  - Right click on your point layer in the Layer window and select **Save as**.
  - From the **Format** dropdown select **ESRI Shapefile**.
  - In the **Save as** dialogue click on **Browse** and browse to:  
*C:\Intro\_Quantum\_GIS\Exercises\Data\Output\_data*
  - Call it *TZA\_health\_facilities\_locations*
  - Tick **Add saved file to map**
  - Leave all other settings as they are
  - Click **Ok**



- You will see a temporary message telling you that the export has been completed.
  - Note that the new layer has appeared in your **Layers Panel** and use **Windows Explorer** to note the dataset location
4. You should now remove the temporary *TZA\_healthfacilities* layer from your QGIS project, as it is no longer required.
    - Right click on the *TZA\_healthfacilities* layer in the Layers window and click on **Remove**