Geographic Information for Vector Surveillance

Week 1 Day 1

Introduction to spatial data and QGIS

Learning objectives

- understand what is meant by 'GIS'
- become familiar with the different types of spatial data
- learn how to install QGIS and become familiar with its layout
- be able to change the size and colour of points and polygons within QGIS
- produce a simple map of entomological data within QGIS

Length : 1 day
Updated : September 2018

Introduction to week 1

The purpose of this course is to enable you to produce maps of your own vector control programme data, and demonstrate how these maps can be used to assist in disease vector control. The learning objectives are :

- To be able to produce a maps of disease vector surveillance data using the software QGIS
- To understand the data flow from field to paper to computer to map
- To be able to produce a map of your own disease vector data
- To understand the terminology associated with spatial data and mapping
- To be able to interpret maps to guide vector control activities
- To develop an enthusiasm to apply mapping approaches in your own work

Day 1 - 4	**Use example datasets**
Day 1	Introduction to spatial data and QGIS
Day 2	From paper forms to QGIS
Day 3	Generating spatial data
Day 4	Manipulating spatial data
Day 5	**Produce maps of your own data**

This course is predominantly practical to provide you with exposure to the QGIS software, and the general processes involved in producing maps of disease vector-related data. We will be delivering some material using a lecture format, however this manual contains all that is presented during these short lectures, in addition to providing you with a step-by-step guide of the practical exercises you will be completing. During the last day you will work with your own data where possible, and focus on generating maps that are specific to your own region. This will be self-guided, however the

instructors will be available to assist you. If you have any concerns about using your own data, please contact one of the instructors as early in the week as possible.

We hope you enjoy the course!

Introduction to GIS

A *Geographical Information System (GIS*) is a computer-based system for capturing, storing, visualising and analysing data that is associated with a geographical location on the Earth's surface. During this course you will be using GIS software to display maps of vector control programme data, which can then be examined and analysed to identify spatial patterns in disease risk. These patterns of risk can then be used to make the best possible decisions for where and how to do vector control.

Coordinate systems are used to assign a location on the Earth's surface to the data of interest. The most common *geographic coordinate system* that we will be using throughout this course is longitude and latitude (Figure 1.1). Lines of longitude run east and west of the prime meridian, which runs through Greenwich England, whereas lines of latitude run north and south of the equator. Both longitude and latitude are measured in degrees. At the equator a single degree represents a distance of approximately 111km. Longitude values range from -180° to +180°, with locations to the right (west) of the prime meridian having positive values, and location to the left (east) of the prime meridian having negative values. Latitude values range from -90° to 90°, with locations above (north) of the equator having positive values, and locations below (south) of the equator having negative values. You may sometimes find that instead of using positive and negative signs, the longitude coordinate are followed by or E and W representing east and west, and the latitude coordinate is followed by a N or S representing north and south. For example, 20° W is equivalent to -20° longitude, and 45°S is equivalent to -45° latitude. Coordinates of a location can be calculated using a Global Positioning System (GPS) which is a satellite-based navigation system. We will learn more about GPS on Day 3.

There are a number of ways of representing longitude and latitude in degrees, with the most common being decimal degrees (xx.xxxxxx), degrees, minutes and seconds (xx° xx' xx'') and degrees decimal minutes (xx° xx.xxxx'). The following example shows the same location in these three formats:

Decimal degrees	**e.g. 21.75555**
Degrees, minutes, seconds	e.g. 21° 45' 20''
Degrees, decimal minutes	e.g. 21° 45.333'

In this course we will primarily be using decimal degrees, and we will learn more about different types of coordinate systems on Day 3.

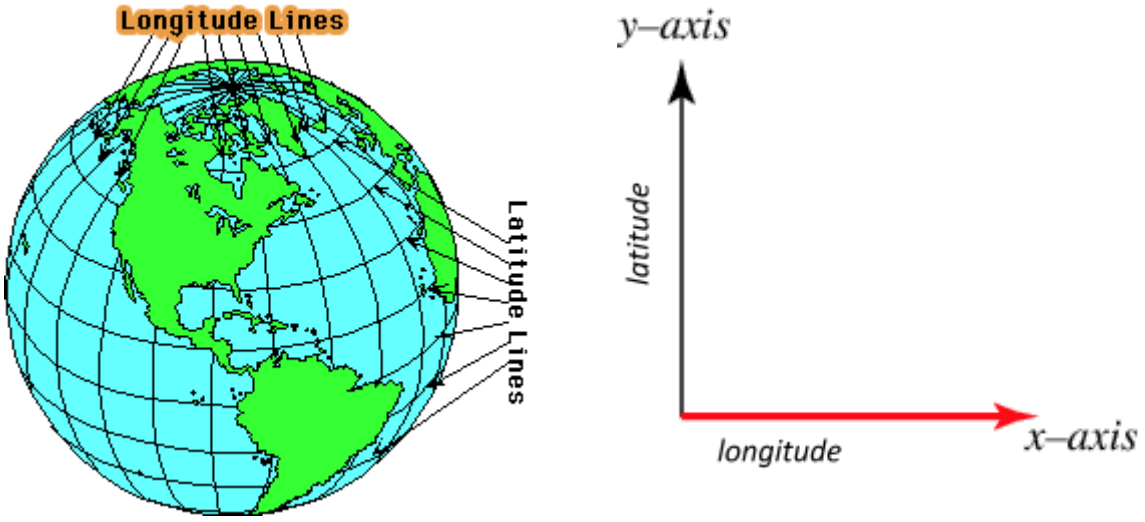


Figure 1.1: The lines on the globe represent longitude (lines that run north to south) and latitude (lines that run east to west). On a simple plot of coordinates, the x-axis displays the longitude values and the y-axis displays latitude values.

In this course we will be using QGIS, an open-source GIS software that is freely available to download from www.qgis.com. For this course, we will be using version 2.18.1 (Las Palmas) released in October 2016. QGIS is maintained by volunteers, and has a vast online community and support network which can be easily accessed if and when QGIS assistance is needed (https://www.qgis.org/en/site/forusers/support.html). There are a vast number of optional additional features that can be added to QGIS using 'plugins'. These are usually written by QGIS developers, and can be added into the software manually, with a few clicks after you have installed QGIS on your computer. We will be demonstrating a few plugins during this course.

Categories of spatial data

There are two categories of spatial data, namely **vector** data and **raster** data. The differences between these two types are described below. However, please note that throughout this course we will be using the word 'vector' to mean different things depending on the context. If you work in a disease control programme or public health you are likely to be most familiar with it in terms of diseases vectors and vector control activities – mosquitoes that carry and transmit malaria or Zika for example. Vector is also used in maths and navigation, as something which conveys direction and size. In GIS, vector data are data that represent locations in geographical space in the form of *points*, *lines* and *polygons*.

Vector data

Vector data are data that represent locations in geographical space in the form of *points*, *lines* and *polygons*. A shapefile is a common way of storing vector (point, line, polygon) data.

Points: Single points in geographical space. For example, point data may represent the exact location of different houses included in an entomological survey or general points of interest, or the points may represent larger geographical areas such the centre of a community.

Lines: Lines in geographical space. For example, line data may represent physical features such as rivers or roads, or routes travelled such as the route taken by a spraying truck, or the route walked by an individual. A single line can be stored as one object, or it may be split into smaller segments.

Polygons: Polygons are closed shapes within geographical space. For example, a polygon may represent a physical feature such as a pond, lake or park, or an administrative boundary such as a region or district. You can also create custom polygons relevant to your project such as survey blocks.

Vector (point, line, polygon) data are commonly stored as a **shapefile**. A shapefile, somewhat confusingly, actually consists of multiple files with the same initial name and different extensions after the '.' (e.g. a.shp, a.shx & a.dxf). These files contain different elements of the data, including the locations, plus variables (known as attributes) associated with each. All of the individual files which make up the shapefile must be saved in the same folder on your computer. Minimally, a shapefile consists of three files. These three files end in .shp, .shx, and .dbf . There may be other files that comprise your shapefile, with common optional extra files having the extensions .prj, cpg, and qpj. For example, in the training material distributed during this course, the shapefile for the locations of the surveyed houses (point data) is a set of six individual files named Example_houses_July with extensions .cpg, .dbf, .prj, .qpj, .qpj, .shp, .shx. You don't need to know what is in each file to be able to use them.

Raster data

Raster data represent something that is geographically continuous. The data are stored as a continuous grid of cells (or pixels), and each cell is assigned a value (single band raster), or possibly multiple values (multiband raster). The dimensions of the cell vary depending on the data that you are using, and is referred to as the 'spatial resolution' of the raster data. For example, we will be looking at gridded population estimate data that has a spatial resolution of 100m, meaning that each cell is 100m by 100m, and the value associated with the cell represents the estimated number of people living in the area covered by the cell.

Raster data are frequently used to represent information measured by remote sensing (usually satellites). For example, the images used by Google in their mapping products such as Google Earth are obtained from satellites. Sensors on satellites can also be used to derive other physical and environmental measures such as elevation, vegetation, land use, plus climate measure such as temperature and rainfall. Data from numerical models can also be represented by a raster. For example, heat maps are often used to represent disease transmission risk.

Raster data can be stored in the form of an image file with associated geographical information. For example, the population data we're using in the exercises consists of two files, namely the TIFF image file (.tif) plus an additional file with the same file name within which the geographical information is stored (.tif.aux).

Course datasets

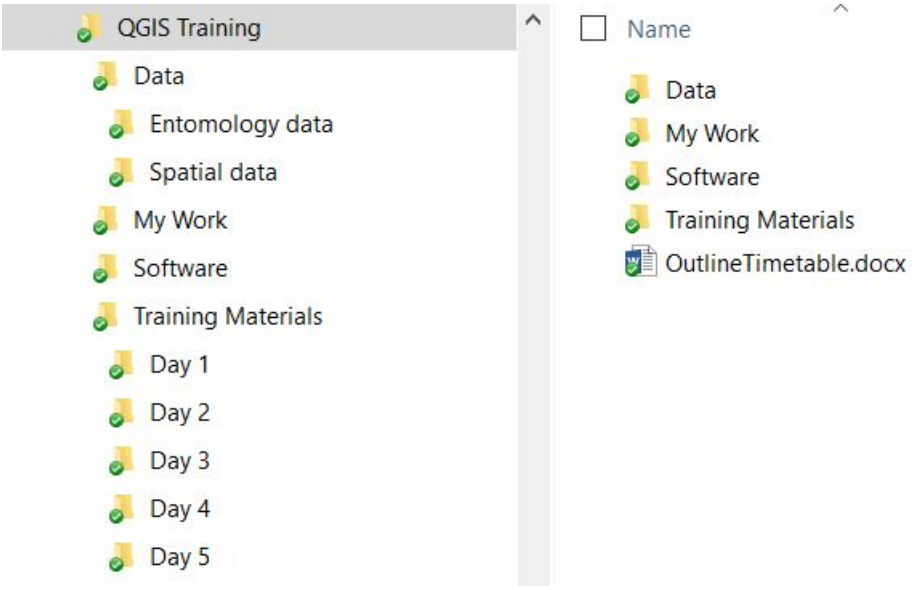
We will be providing you with datasets relevant to disease vector surveillance in order to demonstrate how to use QGIS. These datasets are listed in Table 1 below. These data can be found in the files on your pen drive.

Folder	Subfolder	File Name(s)	Description
Entomology data		household_containers.xlsx	An Excel file containing two sheets representing larvae/pupae survey data for July and September
		presence_aegypti_albopictus.csv	A CSV file containing all published locations where Aedes aegypti and Ae. albopictus have been found, as published in https://elifesciences.org/content/4/e08347
Spatial data	Lines	roads.shp and accompanying files	Shapefiles containing road data for Mexico city obtained from OpenStreetMap
		spray_truck.shp and accompanying files	Shapefiles containing spray truck route data for example area in Mexico city
	Points	houses_july.shp and accompanying files	Shapefiles containing locations of households in example larvae/pupae survey, plus accompanying household-level data
		populated_places.shp	Shapefile containing locations of populated places globally, as obtained from http://www.naturalearthdata.com
	Polygons	COL_adm0.shp and accompanying files	Shapefile containing the outline of Colombia, as obtained from www.gadm.org
		COL_adm1.shp and accompanying files	Shapefile containing the outlines of the departments of Colombia, as obtained from www.gadm.org
		COL_adm2.shp and accompanying files	Shapefile containing the outlines of the municipalities of Colombia, as obtained from www.gadm.org
		block.shp and accompanying files	Shapefile containing the locations of surveyed blocks in example larvae/pupae survey, plus accompanying block-level survey data

		MEX_adm0.shp and accompanying files	Shapefile containing the outline of Mexico, as obtained from www.gadm.org
		MEX_adm1.shp and accompanying files	Shapefile containing the outlines of the states of Mexico, as obtained from www.gadm.org
		MEX_adm2.shp and accompanying files	Shapefile containing the outlines of the municipalities of Mexico, as obtained from www.gadm.org
		ne_50m_admin_0_countries.shp and accompanying files	Shapefile containing the outlines of all of the countries of the world as obtained from http://www.naturalearthdata.com/

File structure

We have provided a folder called QGIS Training within this there are 4 main folders, Data, My Work, Software and Training Materials. Within the Training Materials folder there is one folder for each day of the training. The 'My Work' folder is empty and is a place for you to save files that you create yourself during the training.



It is important to pay attention to the file structure used throughout this course. QGIS does not store data within the software, but rather 'points to' the location where the file is stored on your computer. If, after creating a QGIS project (file extension .qgs), you later move these files to a different location you need to update this information within QGIS, or else the project will not display the data properly. Within this course we will be storing all spatial data within the folder

'Spatial data', and vector data will be stored according to the type of vector data i.e. subfolders 'Lines', 'Points', 'Polygons', whereas raster data will be stored in the 'Raster' subfolder. The non-spatial data for this course can be found in the file 'entomology', as they pertain to Excel and CSV files on mosquitoes, the disease vector in our example.

Example data

The data provided for use in the first part of this course are fictitious data that have been generated solely for demonstration purposes. The example is based principally around entomological data, with the addition of some epidemiological data on a vector-borne disease. We will introduce the data now to give context to the first maps we will look at and return to it later to look in more detail at data flow and manipulation.

Household entomological survey

The details of data collection forms tend to vary between control programs, unless they form part of a more widely coordinated effort and have been standardized or issued for a specific common purpose. The variables or kinds of variables required and general format of forms however, are often similar.

For our example, we are imagining that an entomological survey of households has taken place, at a total of 500 houses (10 houses in each of 50 blocks), at different time points (July and September), to monitor numbers of immature *Aedes aegypti*. Sometimes survey teams will enter data directly in the field, using a tablet or smartphone application, but often data is collected in hard-copy paper format. The data entry form below (Figure 1.2) is an example of what might be used to collect the raw data for the "household_survey_jul_sep.xlsx" dataset.

Demonstration data entry form

Aedes aegypti Surveillance: **Larvae and Pupae Survey**													
Name:				Date:				Municipality					
Unique ID		**GPS Coordinates**		**Containers**									
				Tanks		Tyres		Plants		Others <10 litres			
Block ID	Household ID	Longitude	Latitude	Number	No. Positive	Number	No. Positive	Number	No. Positive	Number	No. Positive		

Figure 1.2: An example of a data entry form used to collect data for monitoring immature Aedes aegypti.

Forms will often start with information that will be used to follow the progress of fieldwork, such as the name of the person collecting the data in the field. This need not be entered into a database unless it will be used in an electronic format for activity monitoring or quality checks for example.

Next in the example form, we have a row of information which is applicable to the whole form i.e. the survey date and the broader location details.

Below this, there are a series of rows, each representing one household. There is space for an intermediate level of location information – a Block identifier, which can be used to group the information later, as well as a column for household identification, and columns for the GPS coordinates of the household. The columns that follow represent a series of different containers to be checked for immature forms of *Aedes aegypti* mosquitoes. For each, the number of that type of container present are to be recorded, as well as the number that are positive – that is, the number with larvae or pupae present.

Lists of example data variables

We will be using the survey data both in its original granular form, at household level, and summarised by block. The variables that will be considered at each level are summarised below.

Household-level variables

Household-level variables can be found in the Excel spreadsheet "household_survey_jul_sep.xlsx". Most variables you will recognise as coming from the entomological survey, collected using the form as described above. Further fictitious data on disease case numbers on Zika virus disease have been added to the datasets to demonstrate the possibility and power of integrated data. Sometimes entomological surveys will be combined with prevalence surveys or other data collection. Programme managers may also have access to case numbers and their provenance, as well as entomological data. Where data has a common identifier of location the information can be displayed together on a map, which can be useful for situation analysis. You will see this household-level case data displayed in the day 1 example map that we will be looking at. A shortened name for each variable has also been listed in the below table; we will come back to this later.

Variable name – Used in Excel spreadsheet	**Short name – to be used in csv files**	**Description**
Block ID	blockid	Block ID, ranging from 1-50
Household ID	houseid	House ID within each block, ranging from 1-10
Unique ID	uniqueid	Unique household ID in the format BxxHxx
Number of tanks	num_tanks	Number of tanks found at the house
Number of positive tanks	pos_tanks	Number of tanks at the house where larvae/pupae were found
Number of tyres	num_tyres	Number of tyres found at the house
Number of positive tyres	pos_tyres	Number of tyres at the house where larvae/pupae were found
Number of others <10 litres	num_other	Number of other containers of less than ten litre capacity

Number of positive others <10 litres	pos_other	Number of other containers of less than ten litre capacity where larvae/pupae were found
Total number of containers	tot_contai	Total number of containers recorded at the house (variable to be calculated)
Total number of positive containers	tot_pos_co	Total number of larvae/pupae positive containers found at the house (variable to be calculated)
Larvae and/or pupae present	present	Indicator of larvae/pupae present (1= present, 0=absent) (variable to be calculated)
Date	date	Date of collection in DD/MM/YYYY format
Municipality	municipali	Municipality in which the household is situated
Cases	cases	Number of laboratory-confirmed Zika virus disease cases reported amongst those people residing in the surveyed house

Block-level variables

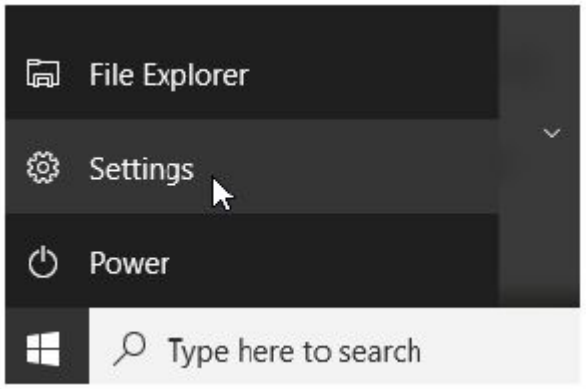
It is often useful to summarize data for reporting or to have an overview of the situation. In our example we want to look at block level indicators. We will be showing you how to derive this data from household-level data in Day 2.

Variable name – Used in Excel spreadsheet Pivot Table	Short name – used in csv files	Description
Block ID	blockid	Block ID, ranging from 1-50
Count of Household ID	num_houses	The total number of houses surveyed in each block
Sum of Total number of containers	num_contai	The total number of containers in each block
Sum of Total number of positive containers	num_pos_co	The total number of larvae/pupae positive containers in each block
Sum of Number of houses with larvae/pupae present	num_pos_ho	The total number of houses in the block with larvae/pupae present
House index (HI): percentage of houses infested with larvae and/or pupae	hi	Household index = number of positive houses[num_pos_ho]/total number of houses[num_houses]
Container index (CI): percentage of water-holding containers infested with larvae or pupae	ci	Container index = number of positive containers[num_pos_co]/number of containers[num_contai]
Breteau index (BI): number of positive containers per 100 houses inspected.	bi	Breteau index = 100*Number of positive containers[num_pos_co]/Number of houses[num_houses]

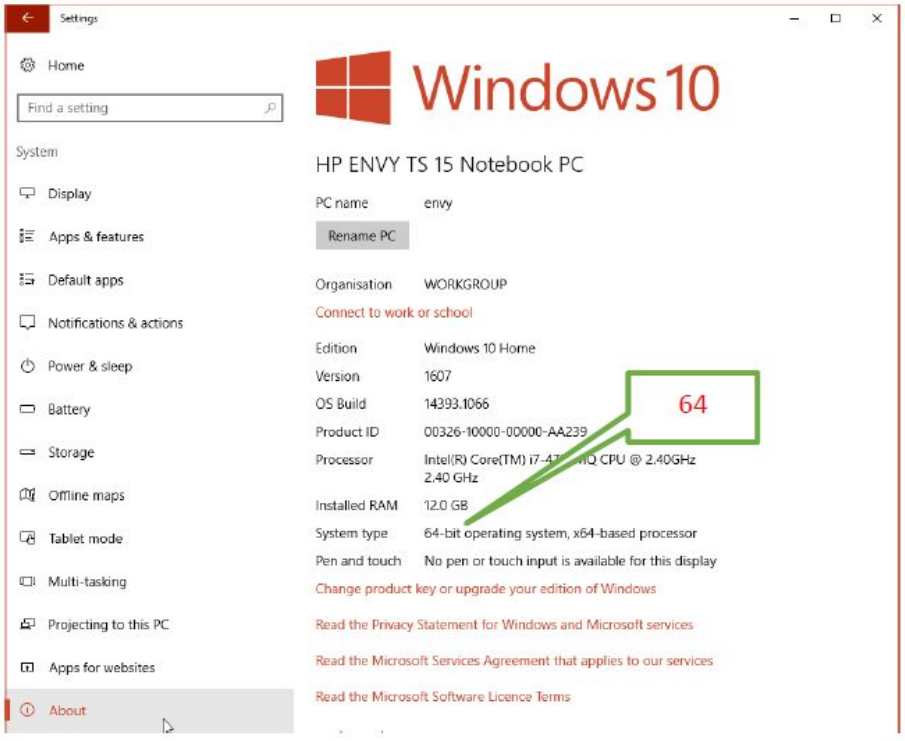
Install QGIS

The QGIS software is freely available to download from the internet. To save time, we have provided you with copies of the installation files on your USBs in the folder “software”.

There are three QGIS installation files provided. The file ending in “.dmg” is for installation on a Macintosh. The two ending in “.exe” are for PCs, one ends in “...x86.exe” and the other in “...x86_64.exe”. Which you choose will depend on whether the processor of your computer is “32-bit” or “64-bit”. If you are not sure which your computer has, you can verify this by clicking on the Windows icon in the bottom left-hand corner of your screen and searching for “system” to bring up the system information window, or looking for it manually within in Settings/System/About.



The system information screen will appear something like this:



Once you have confirmed which file you require, the steps for installation are detailed below.

Windows

To install QGIS,

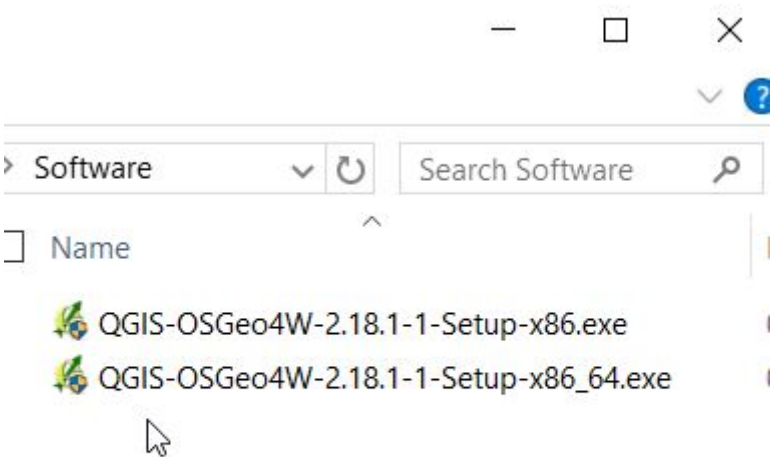
➔ double-click on one of the installation files provided in the USB in the folder “Software”.

Select :

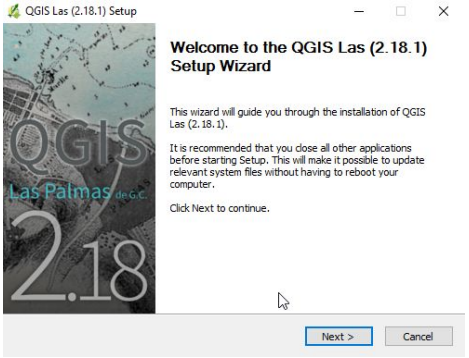
QGIS-OSGeo4W-2.18.1-1-Setup-x86_**64**.exe if your PC is **64-bits.**

OSGeo4W-2.18.1-1-Setup-x86.exe if your PC is 32-bits.

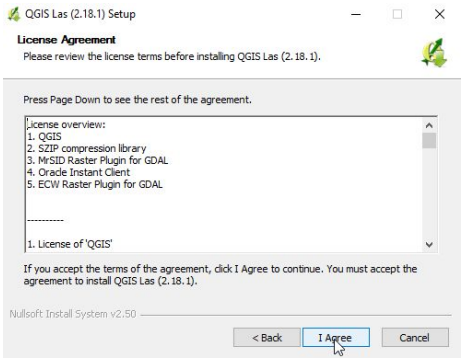
(See above guidance on how to check your PC system information)



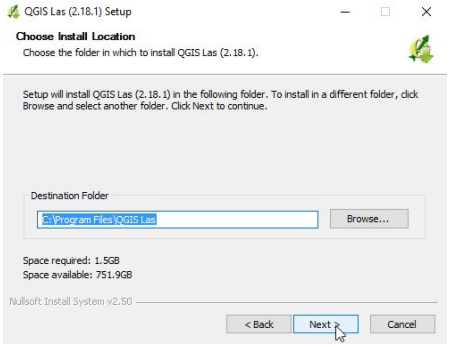
After accepting the first question window you should get a window like this :



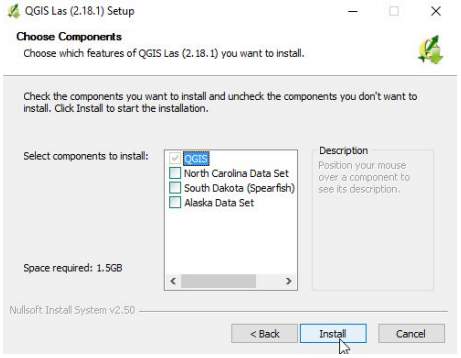
➔ Click 'I agree'



➔ Click 'Next' accepting the destination folder



➔ Click 'Install' :



QGIS can take many minutes to install after this step.

➔ click 'Finish'.

If you don't have access to the folders we have provided you can download the installation files from http://www.qgis.org/en/site/forusers/download.html, or search on google for "qgis download" to find. The download page currently looks like below :

Macintosh

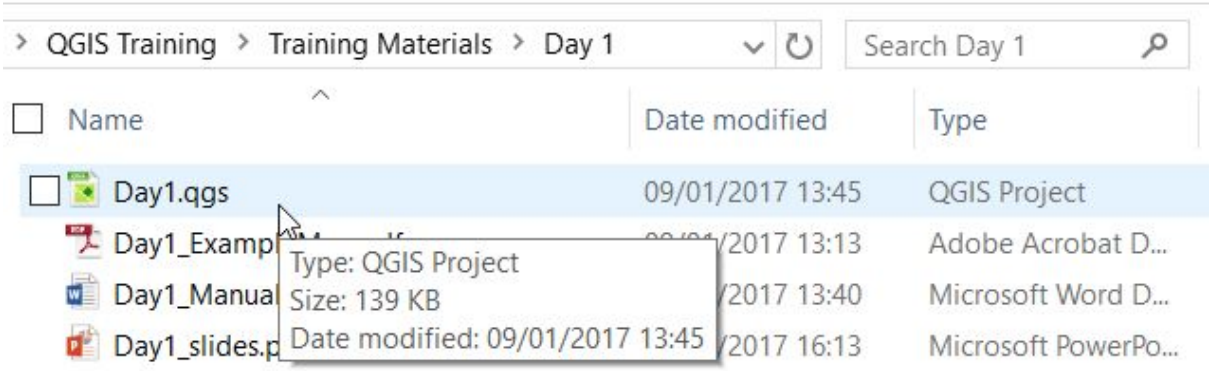
To install on Mac OS X follow the instructions from here:
http://drjill.net/install-qgis-on-a-mac-in-8-steps/

The instructions are in English, ask us and we can help you follow the 8 steps. It requires a single download from here: http://www.kyngchaos.com/software/qgis. We have included the Mac installation file on the USB drive (software\QGIS-2.18.2-1.dmg).

Exploring QGIS

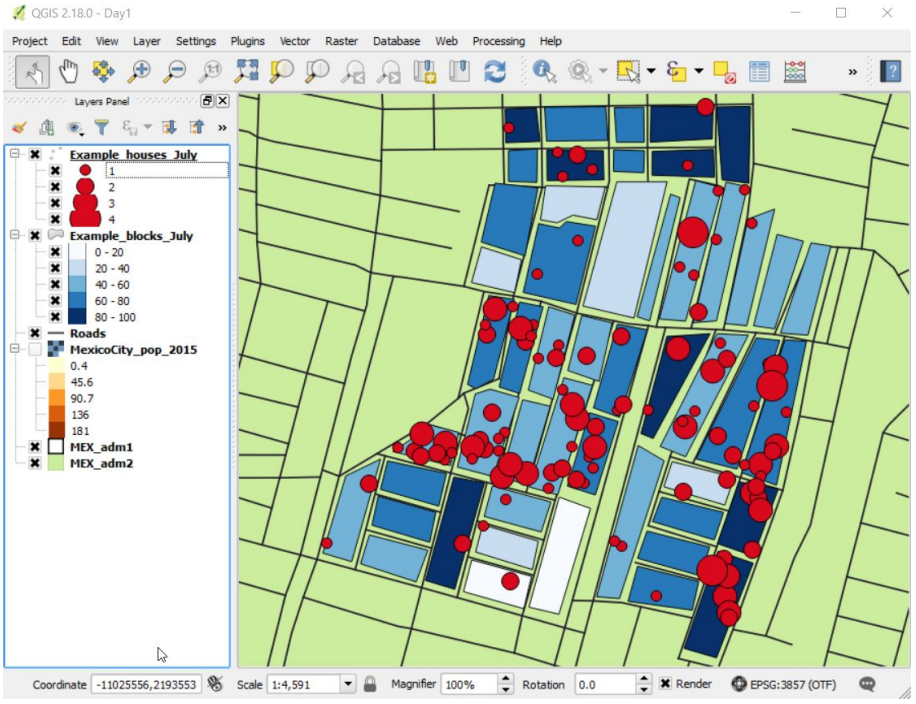
➔ Goto the folder 'training/day1'
➔ double click on the day1.qgs file.

(or if you have QGIS open already you can select Project, Open and find Day1.qgs).

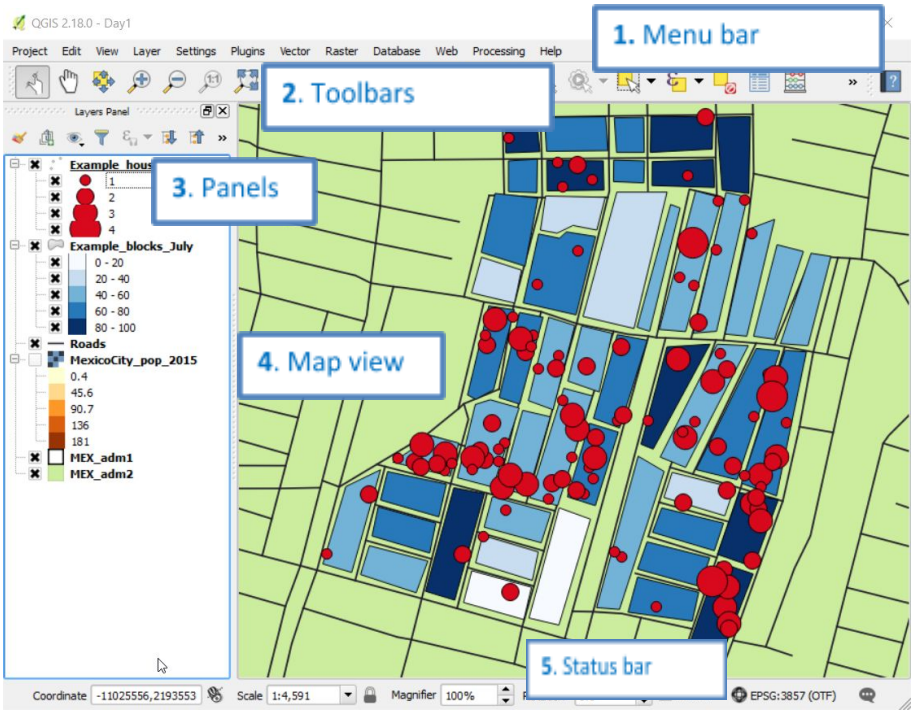


QGIS may take a while to start, during which time it will display this window :

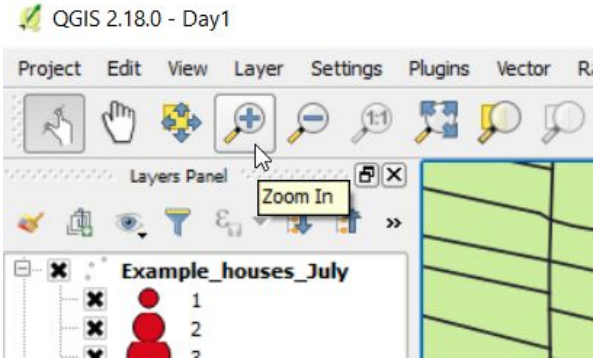
When the QGIS project opens it should look something like this :



Here the different areas of the QGIS User Interface are labelled :



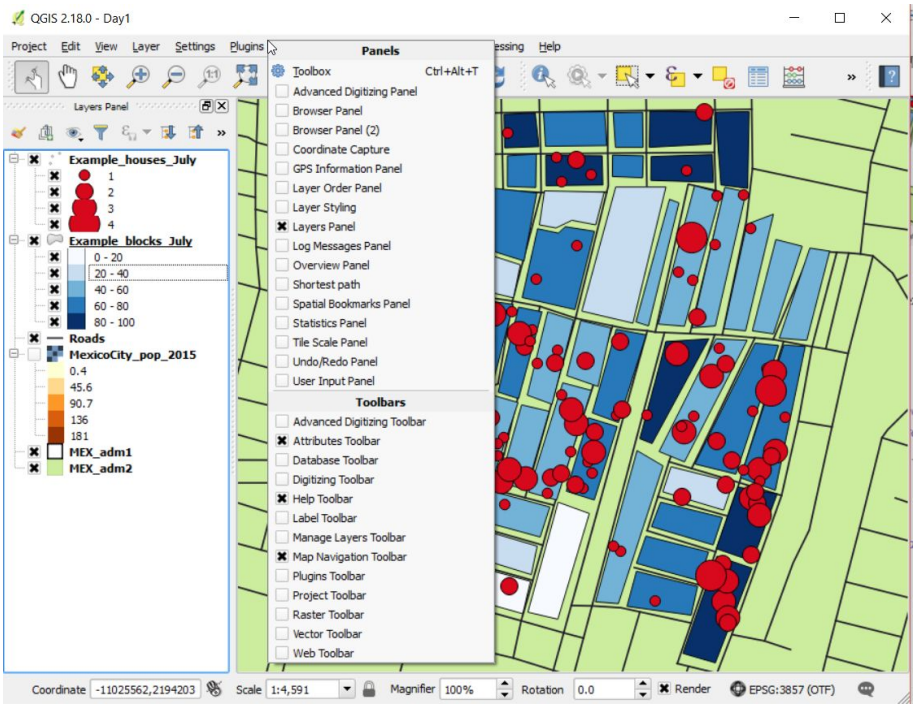
If you pause the cursor ('hover') over buttons in the toolbars a little message will appear telling you what that button does e.g. as seen for 'Zoom In' :



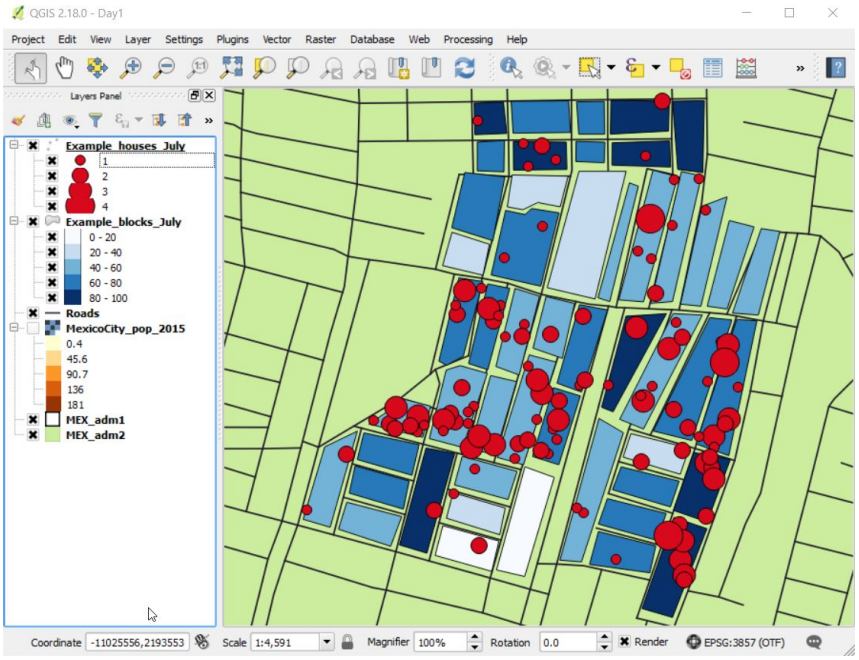
➔ Try 'hovering' over some of the buttons yourself. The hover message disappears after a while.

QGIS is a powerful piece of software and initially we will only need to use and understand a small part of it. We can turn off some components to focus on the basics to start with.

➔ Right click in Panels or Tool bars to be able to turn options on and off.
➔ Uncheck most options just to leave the 4 shown below (Layers Panel, Attributes Toolbar, Help Toolbar, Map Navigation Toolbar)

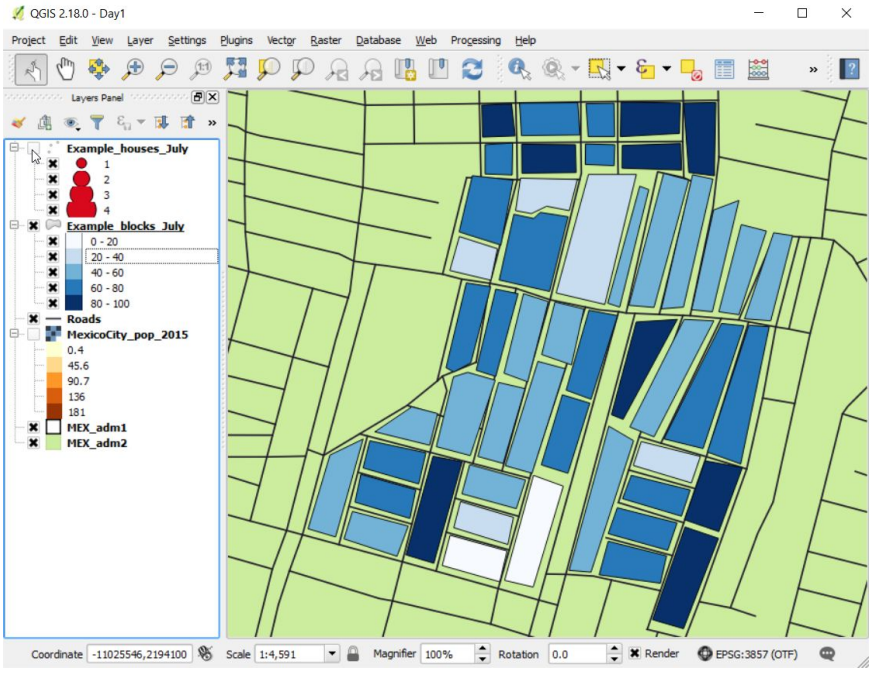


Data are displayed in the map as 'layers' that appear in the Layers Panel on the left. We will be using the Layers Panel a lot.



The checkboxes to the left of each layer name within the Layers Panel allow you to change what is displayed in the map view.

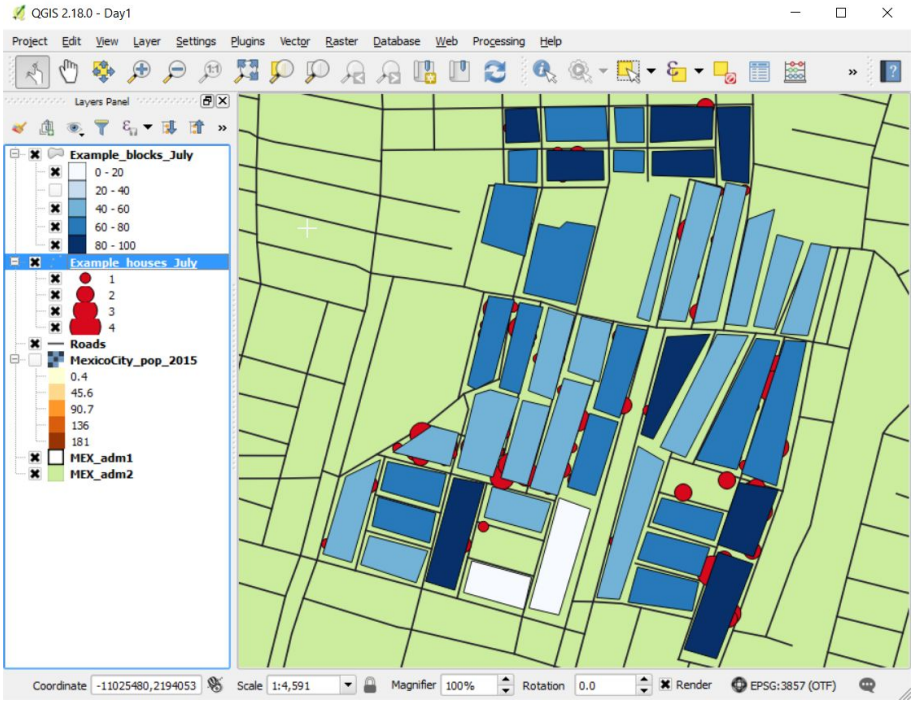
➔ Experiment with unchecking boxes and see how features no longer appear in the map view. Checking the box again makes them come back (they have not been deleted they are just not shown when they are unchecked).



The order of layers in the Layer Panel determines how they appear in the map view. The layers at the top are displayed last and those at the bottom are displayed first. Thus it is like the higher up layers are painted on top of the lower layers. This can mean that some layers get hidden by others. Later we will see that layers can be made partially transparent to allow them to be seen underneath others. You can change the order of layers by clicking and dragging them within the Layers Panel.

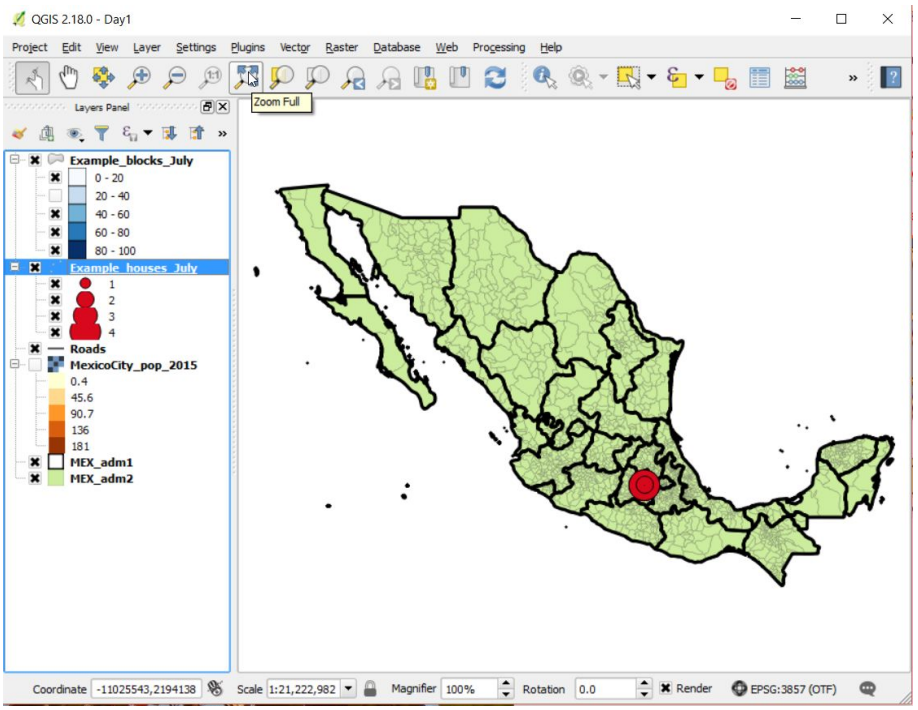
➔ In the Layers Panel Left click, hold and drag the blocks layer so that it is on top of the points layer

This is a little tricky, you need to let go of the mouse button in a certain place and if you don't the layer will not move. Try experimenting.



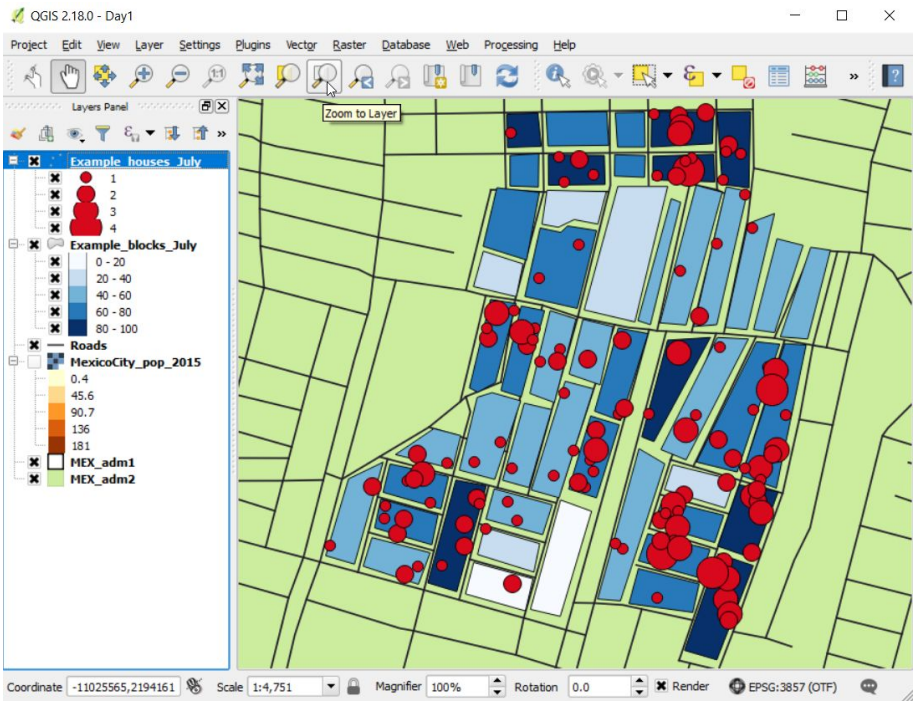
There are many ways to change the position and size of the area that is shown within the map view.

➔ use zoom buttons from the toolbar
➔ zoom in and out with the mouse wheel
➔ left click and drag to move the map



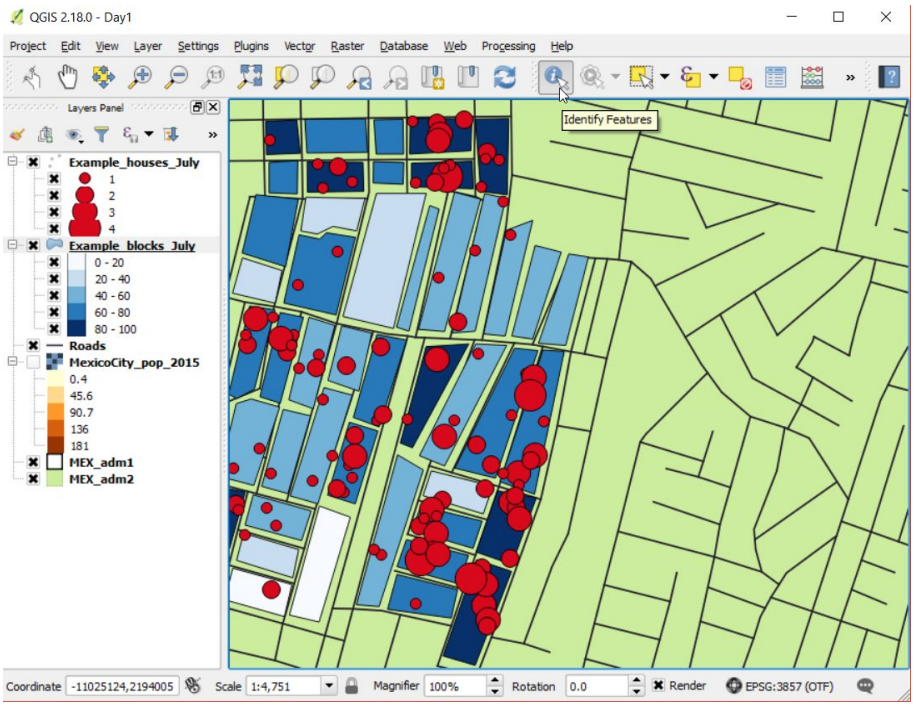
When you have a layer selected you can use the 'Zoom to layer' button to set the map view to show the whole of that layer :

- ➔ Left click a layer in the Layers Panel on the left
- ➔ Left click 'Zoom to layer' in the toolbar at the top to show the whole of that layer

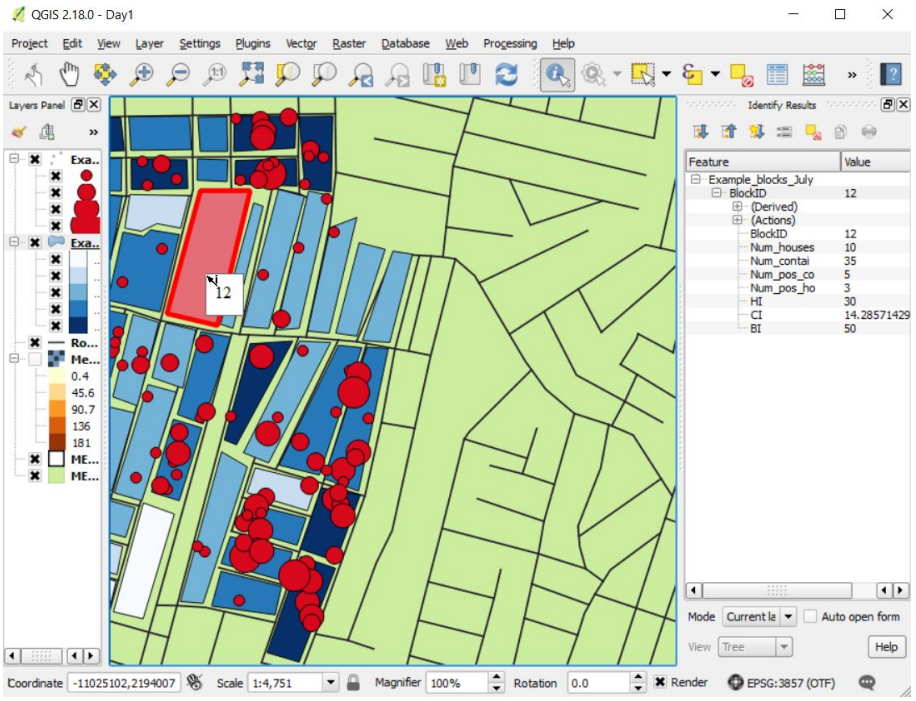


There are many ways to find out information about the map. One way is to use the 'Identify features' button to bring up information about map features stored in the currently selected layer :

- ➔ Left click the blocks layer in the Layers Panel to select it
- ➔ Left click the 'Identify Features' button in the toolbar
- ➔ Left click in one of the blue polygons on the map to select it



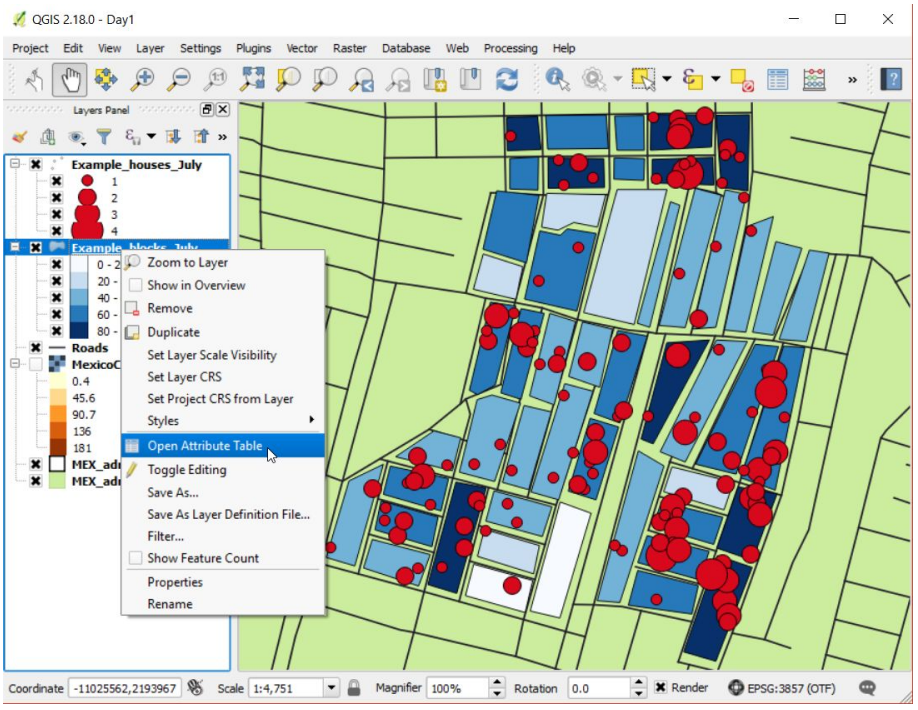
This should bring up an information panel on the right where you should be able to see e.g. the number of houses ('Num_houses') within each block.



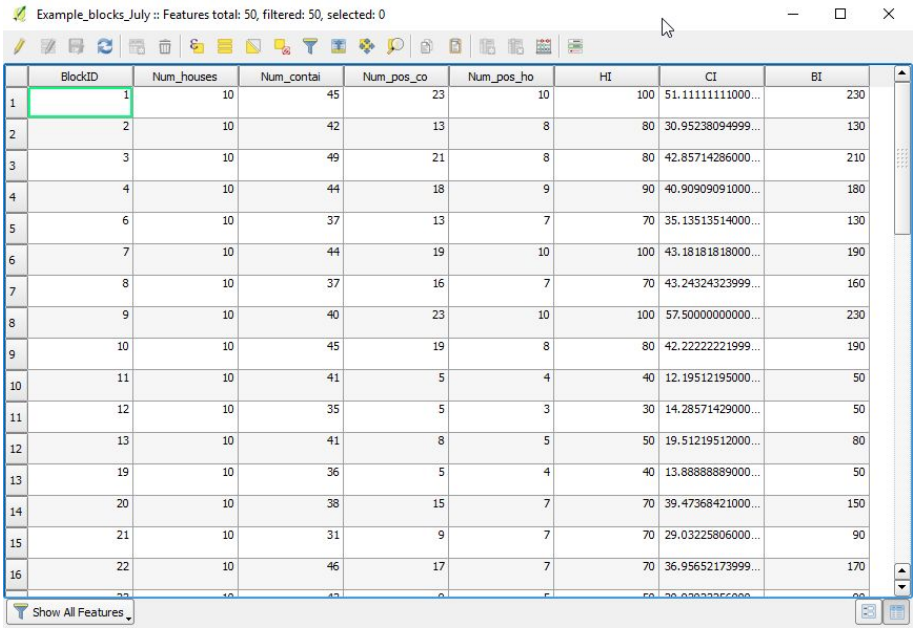
➔ Try selecting the 'Example_houses' layer and use the identify button to find out information about individual houses (the red circles).

This information about the layers is stored in 'Attribute tables' which are like spreadsheet tables. Attribute Tables are very useful, to view one :

➔ Right click on a layer in Layers Panel
➔ select ‘Open Attribute Table’



The attribute table for the blocks layer should look like below. Each row stores data about one feature In this example, a feature is a block hence there is one row per block. Have a look at the columns and values.



	BlockID	Num_houses	Num_contai	Num_pos_co	Num_pos_ho	HI	CI	BI
1	1	10	45	23	10	100	51.11111111000...	230
2	2	10	42	13	8	80	30.95238094999...	130
3	3	10	49	21	8	80	42.85714286000...	210
4	4	10	44	18	9	90	40.90909091000...	180
5	6	10	37	13	7	70	35.13513514000...	130
6	7	10	44	19	10	100	43.18181818000...	190
7	8	10	37	16	7	70	43.24324323999...	160
8	9	10	40	23	10	100	57.50000000000...	230
9	10	10	45	19	8	80	42.22222221999...	190
10	11	10	41	5	4	40	12.19512195000...	50
11	12	10	35	5	3	30	14.28571429000...	50
12	13	10	41	8	5	50	19.51219512000...	80
13	19	10	36	5	4	40	13.88888889000...	50
14	20	10	38	15	7	70	39.47368421000...	150
15	21	10	31	9	7	70	29.03225806000...	90
16	22	10	46	17	7	70	36.95652173999...	170

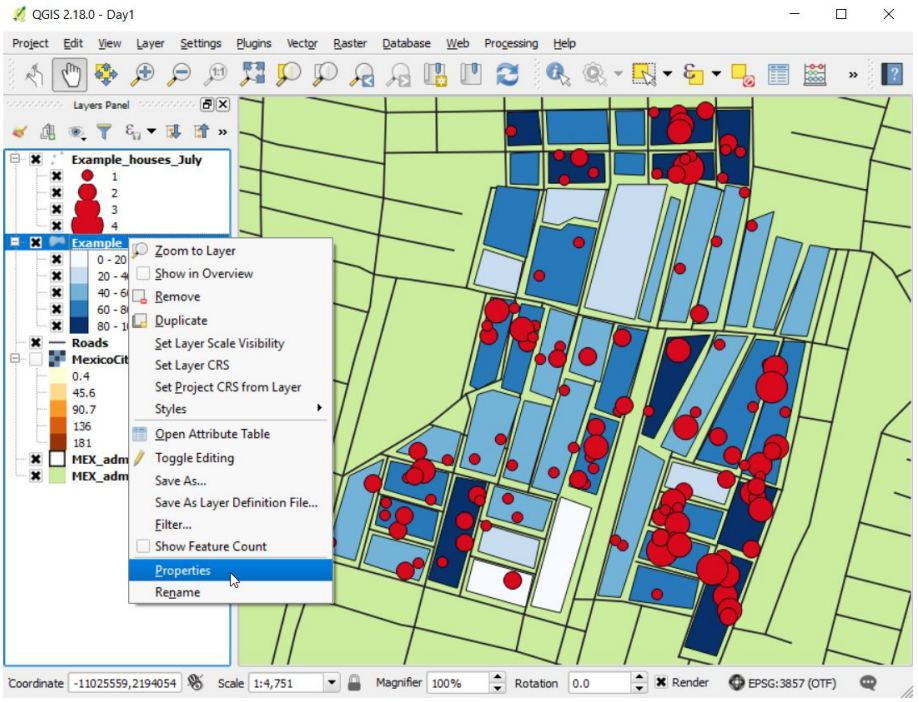
The column names give some indication to the data that are held in them. For the blocks data here is a brief summary of the columns :

blockID	a unique identifier for each block
num_houses	number of houses within the block
num_contai	number of containers
num_pos_co	number of positive containers (positive=containing larvae or pupae)
num_pos_ho	number of positive houses
HI	House Index, percentage of positive houses
CI	Container Index, percentage of positive houses
BI	Breteau Index, number of positive containers per 100 houses

- ➔ Close the attribute table by clicking the x in the top right corner.
- ➔ Open the attribute table for the houses (points) layer. Remember right click on the layer name in the Layer Panel then select 'Open Attribute Table'.

The houses data has more columns and you can use the lower scroll bar to view the columns to the right. In this table there is one row for each point (house). Look at the columns names and refer to section 1.3.1 to see what data they hold.

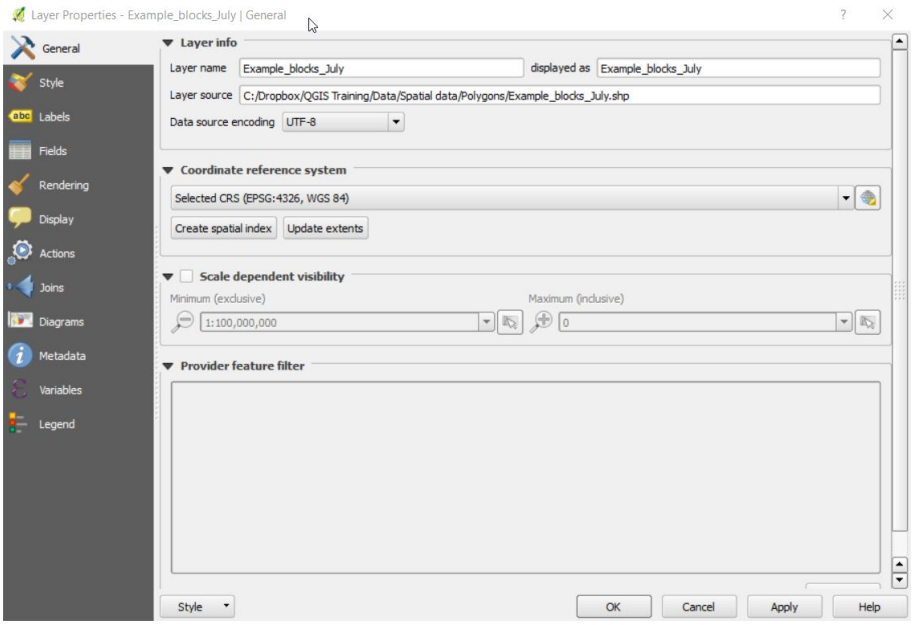
- ➔ To find out other information stored about a layer, right click on the layer and select 'Properties' :



This brings up the Layer properties window, it allows you to view, and later change, lots of information about each layer.

➔ For now just have a look at what is stored by clicking options on the left.

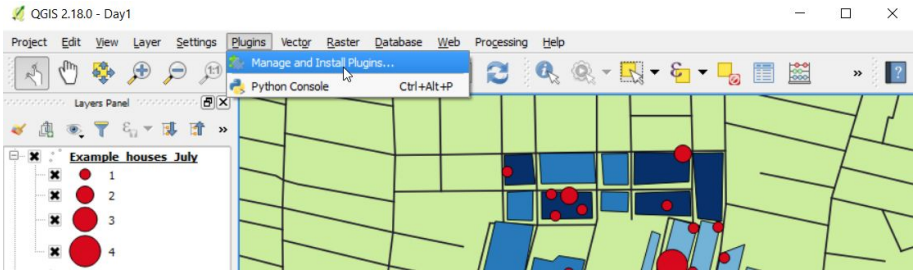
We will come back to Layer properties later.



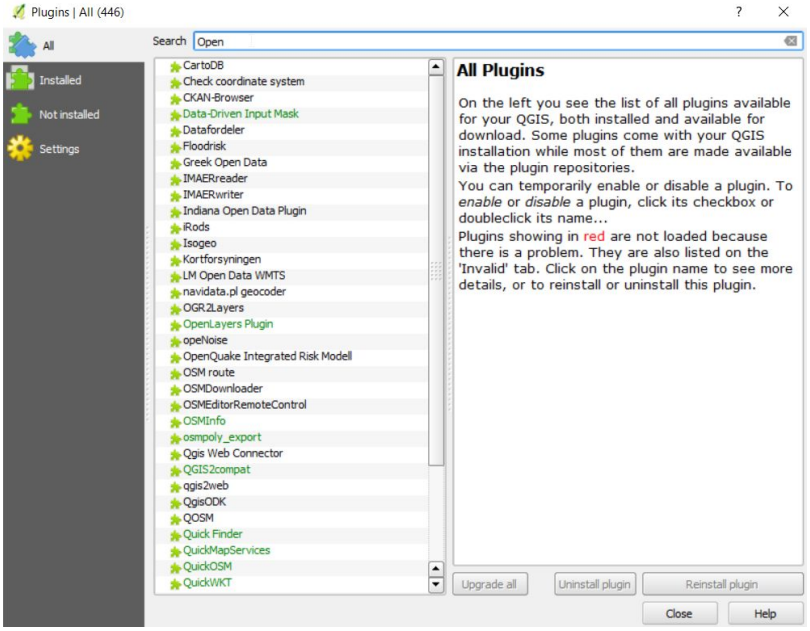
Adding basemaps to QGIS

We can add basemaps to our map but first we need to install a 'plugin' to do this. Plugins are optional features and functions created by the developers of QGIS, and other members of the GIS community. You only need to install a particular plugin once it will then be available for you to use in future QGIS sessions without having to repeat this process each time.

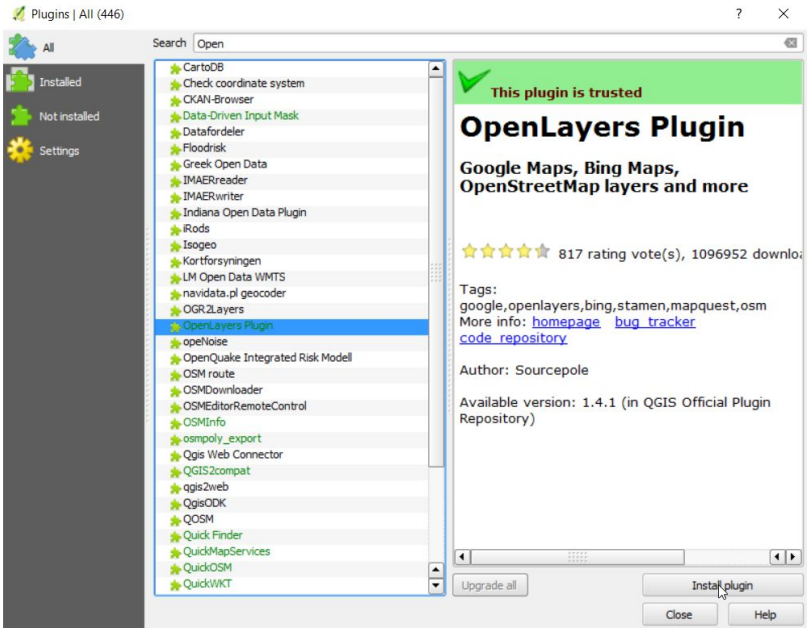
- ➔ Left click on Plugins in the menu bar
- ➔ Select 'Manage and install plugins'

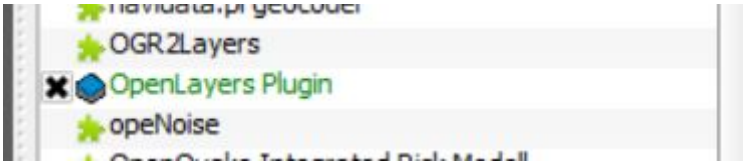


That should bring up this window :



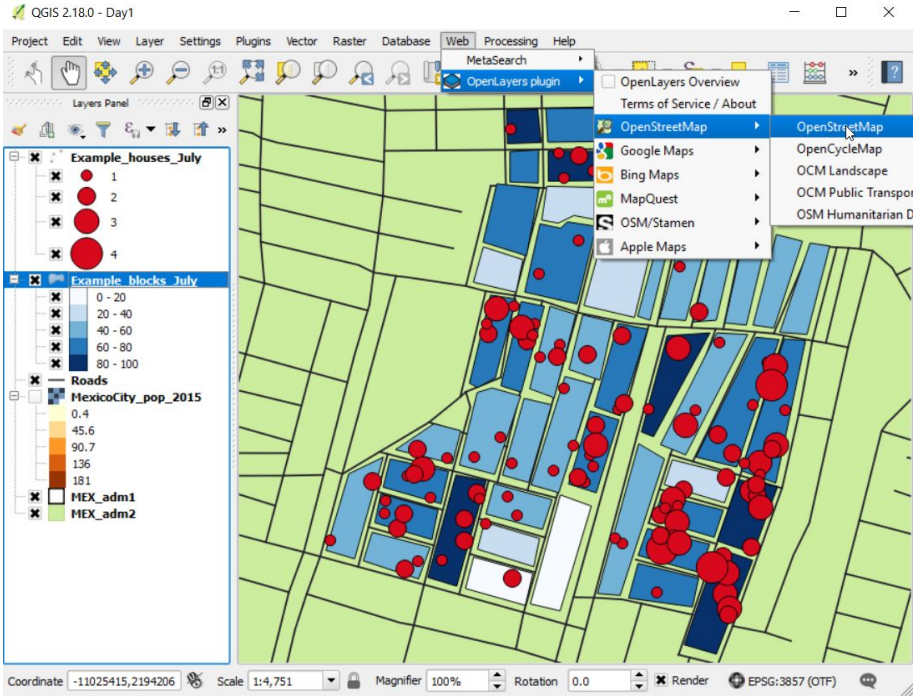
- ➔ Type open into the search box
- ➔ Left click on 'OpenLayers Plugin'.





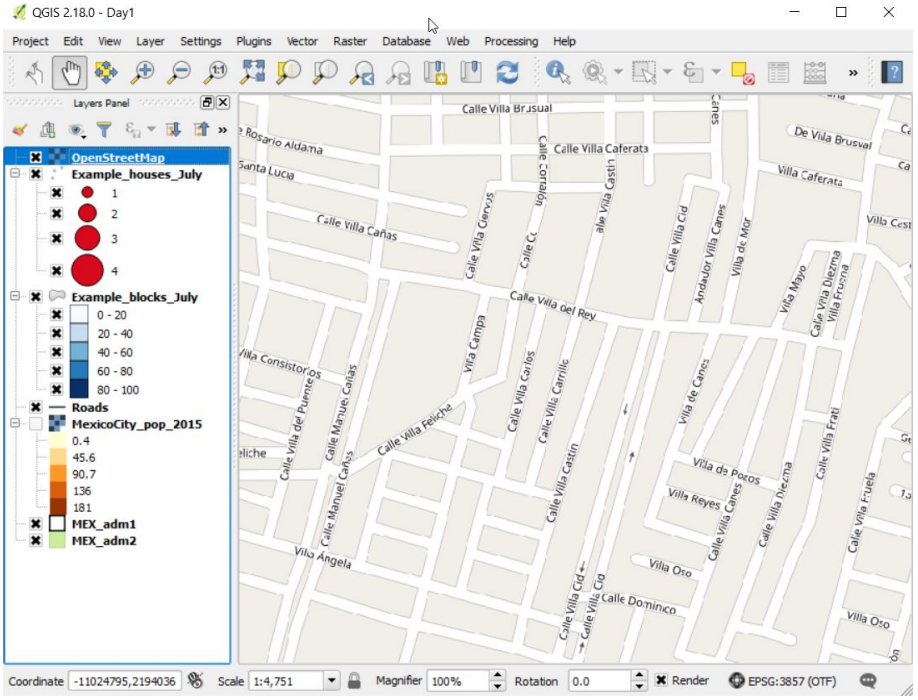
Once we have the OpenLayers plugin installed, we can use it from 'Web' in the Menu bar :

- ➔ Click on web in the menu bar
- ➔ select OpenStreetMap, OpenStreetMap



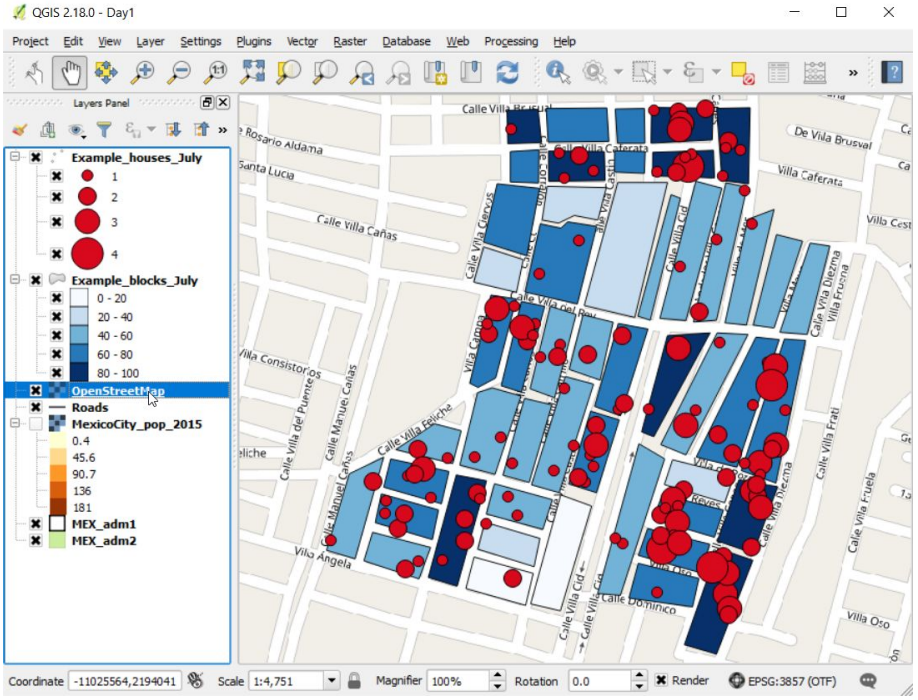
This should bring up a map of streets.

Q. Can you think why we can no longer see the data that we had before ? Can you think of a way of viewing our data on top of the street map ?



A. The 'OpenStreetMap' layer was added to the top of the layers list.

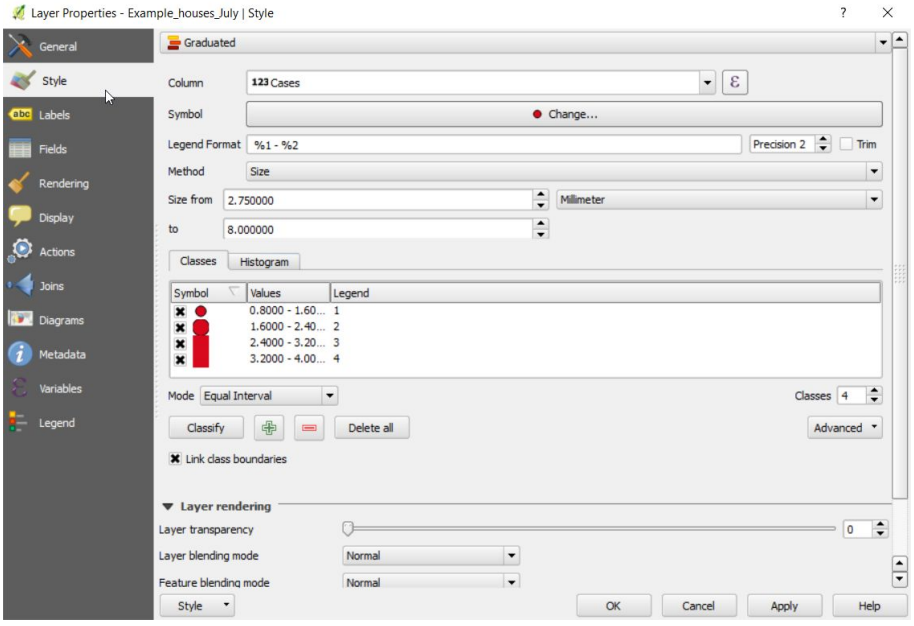
➔ Try clicking and dragging the layer to below the houses and blocks. You should then be able to see our data on top of the street map.



QGIS Symbology

Symbology determines how data are displayed on the map. For example, the current symbology we are using results in the locations of cases being presented on the map by a red circle, such that the size of the circle represents the number of cases identified. Symbology can be modified using Layer properties which we looked at earlier. Symbology options are different for point, line, polygon and raster data. We will start by looking at point data.

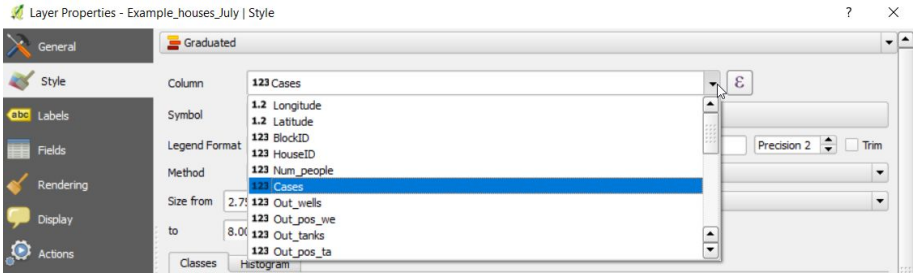
- ➔ Right click the houses layer to bring up the Properties window.
- ➔ Select the Style option on the left.



There are lots of Style options and we don't need to understand them all at first. Currently the house layer is set to present the variable ‘cases’ using the 'Graduated' style option, as seen at the top of the Style screen. The graduated option enables the size of the points to be determined by the values in one of the columns of the Attribute table, which initially is *cases*. Usually, we use increasing point sizes to represent increasing values.

➔ in the ‘Column’ choice box choose ‘Num_people’ to change the column that sets the appearance of the points

➔ click Apply or OK at the bottom of the window

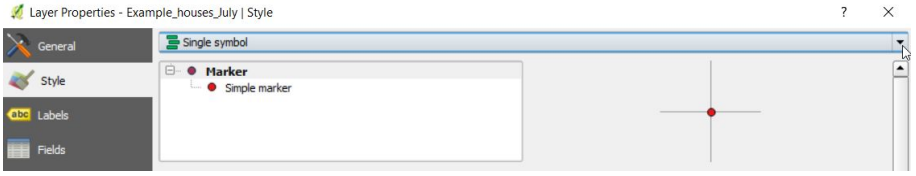


Changing Column to num_people means that the size of the points is now determined by the number of people in each house and the map will look different.

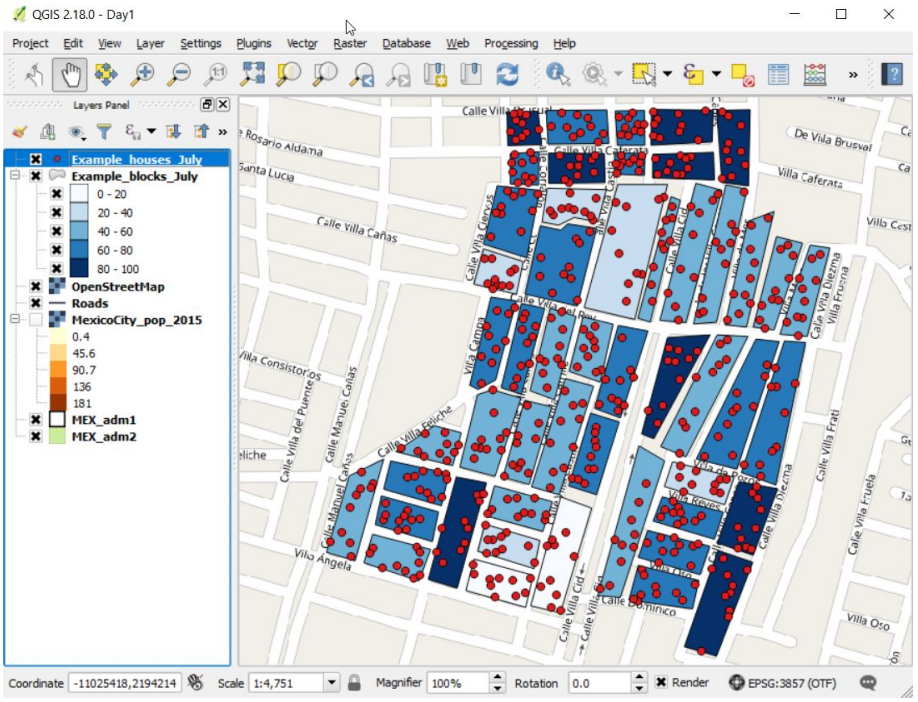
We can also have points of a constant size :

➔ change from the current ‘Graduated’ option at the top to ‘Single Symbol’

➔ click Apply or OK at the bottom of the window



Now the map should like the below with both the point symbols and legend changed :



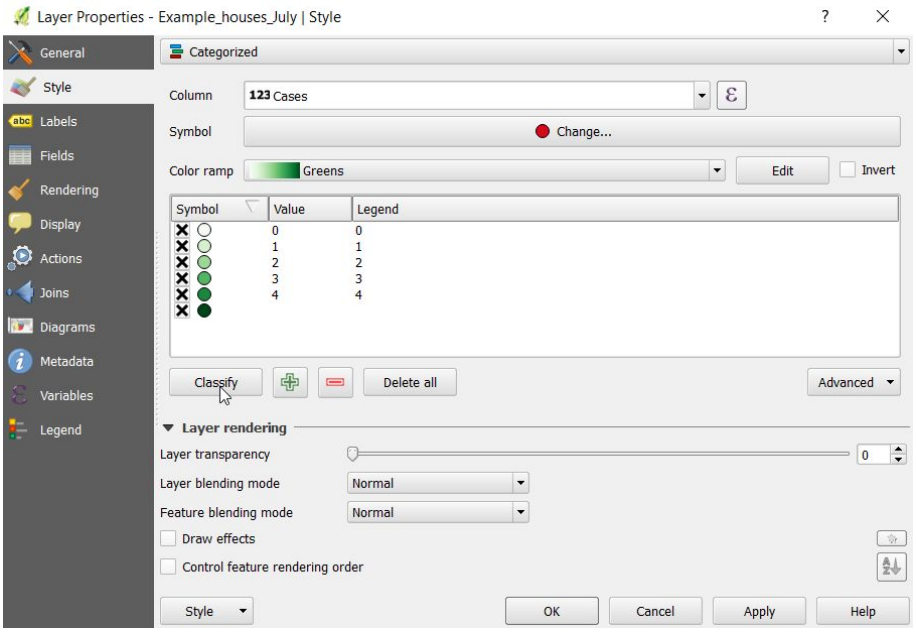
We can also use the color of points rather than their size to represent attribute data.

- ➔ Right click on the houses layer again to bring up the Properties, Style window
- ➔ Set the top option to 'Categorized'

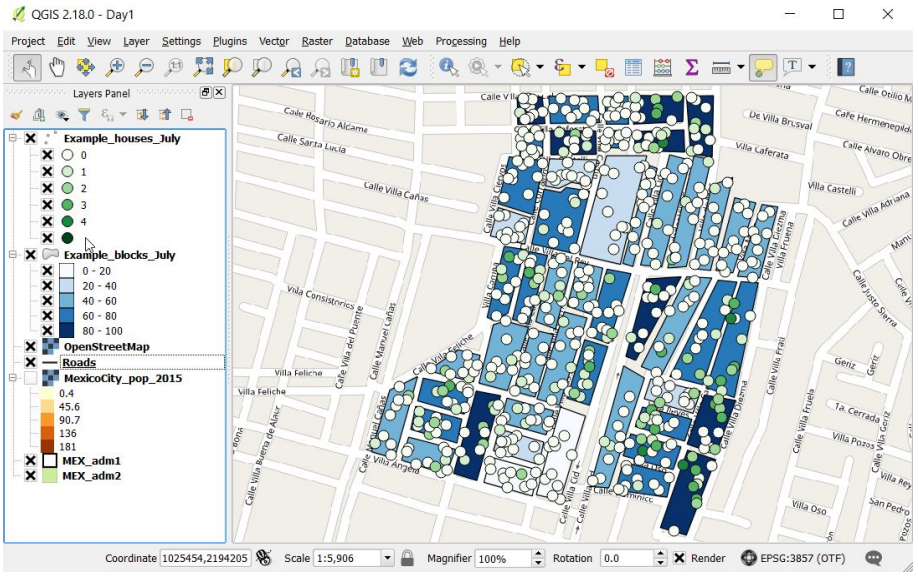
We can use categorized because the number of cases in a house is one of a category (i.e. 0,1,2,3,4) rather than being a continuous value between 0 & 4.

- ➔ Set 'Column' to 'Cases'
- ➔ Change Color ramp to Greens.
- ➔ Click 'Classify' and OK to make the changes.

The 'Classify' button calculates which colours to use for which points. You will see later that it is important to click this after making a change otherwise some points can end up with no colour.



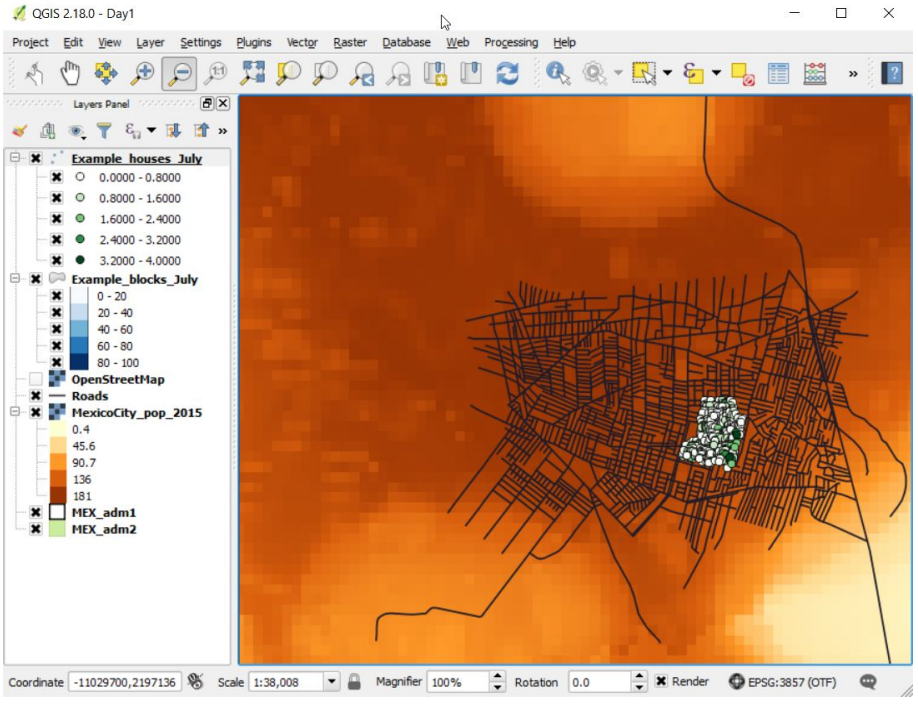
Now the map should appear as below with the points coloured dark green for high numbers of cases and light for lower numbers.



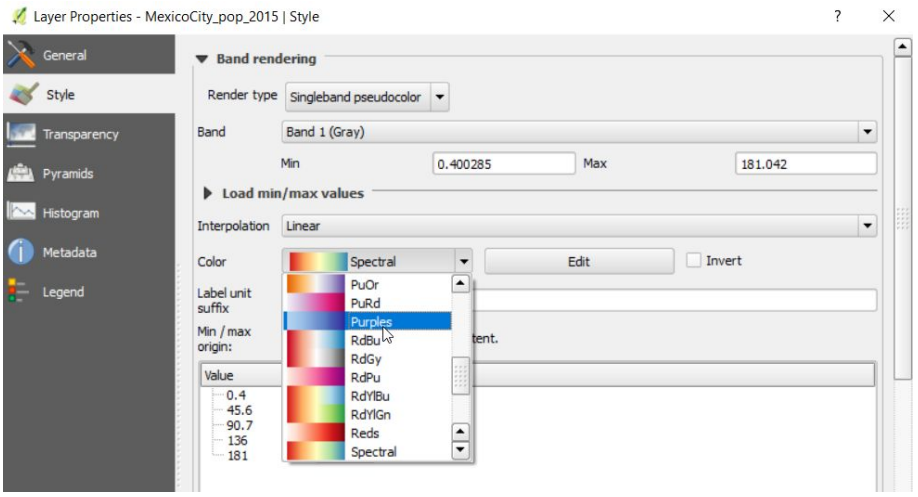
Similar options can be set for the blocks (polygon) layer. You can also control how the legend appears for each layer and we will come back to this later.

You can control the Style of raster (gridded) data too. We have raster data of estimated human population for Mexico city in the layer 'mexico_city_pop_per_hectare2015'. Select this layer in the layers panel and zoom out to see it. This layer is derived from data downloaded from the WorldPop project at : http://www.worldpop.org.uk/data/files/index.php?dataset=MEX-POP&action=group

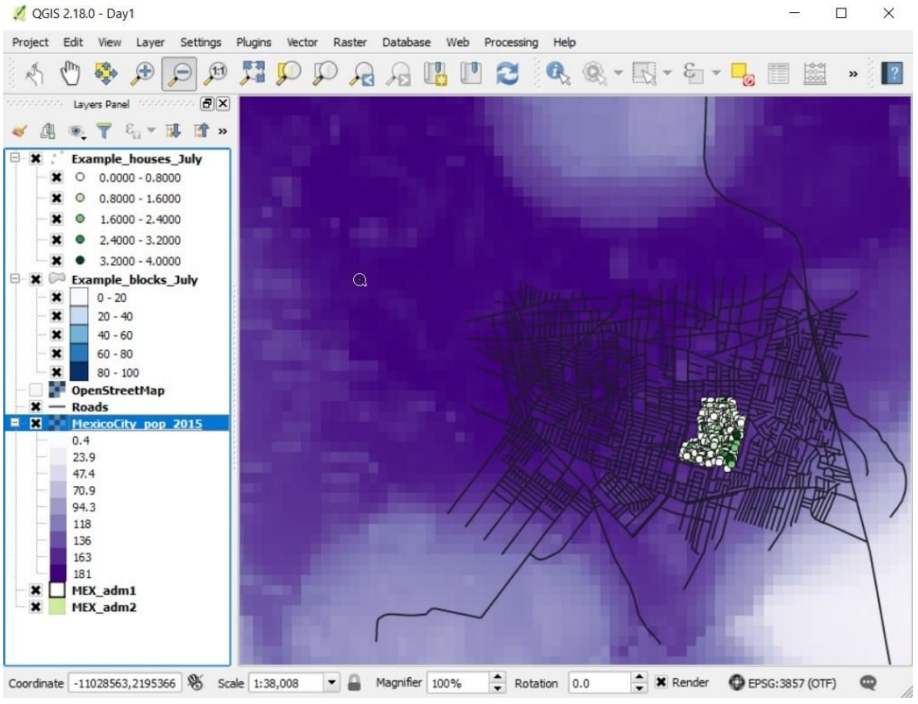
- ➔ In the layers panel : uncheck the OpenStreetMap layer
- ➔ click check box for MexicoCity_pop_2015
- ➔ zoom out on the map (either using mouse wheel or the zoom-out button)



➔ In the layers panel Right click on MexicoCity_pop_2015, select Properties
➔ select SStyle on the left
➔ Change 'Color' to *Purples* and press OK.



Now the map should look like this :



We will learn more about manipulating raster data in Day 4.

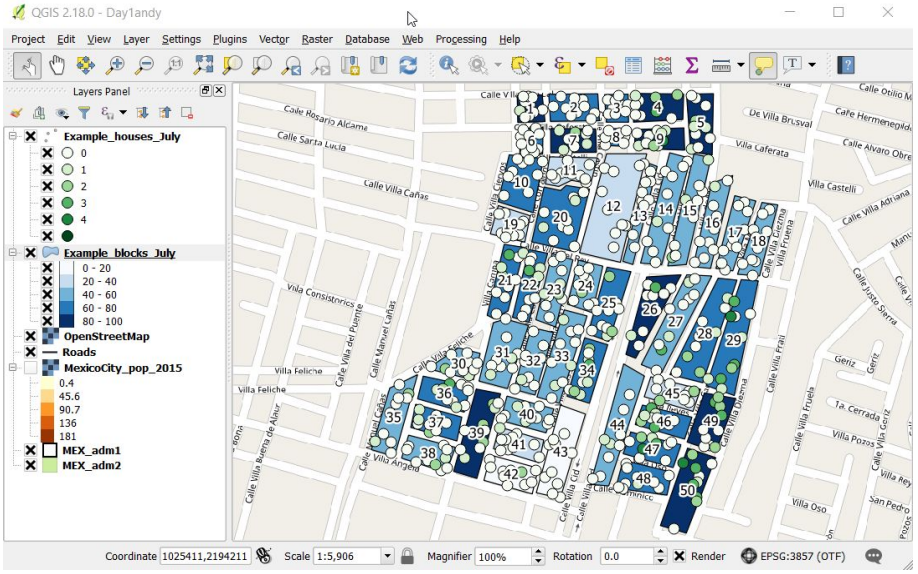
QGIS labels

Labels allow us to put text on the map associated with points, lines or polygons.

- ➔ in the layers panel uncheck the population raster layer
- ➔ zoom back to the extent of the blocks layer (right click, zoom to layer)
- ➔ right click on the blocks layer and select Properties.
- ➔ Select 'Labels' on the left
- ➔ AT the top change from ‘No labels’ to ‘Show labels for this layer’

- ➔ Set 'Label with' to 'blockID' and press OK at the bottom.

The blockID column in the Attribute table is used to label the blocks :



Now you have had a brief introduction to adding data to QGIS and changing how it is represented. We don't expect you to know everything and don't worry there will still be things that you don't understand. There are still parts of QGIS that we don't understand! However you now should know enough to start creating and modifying your own maps. Have a go at the Exercise below.

Exercise 1.1

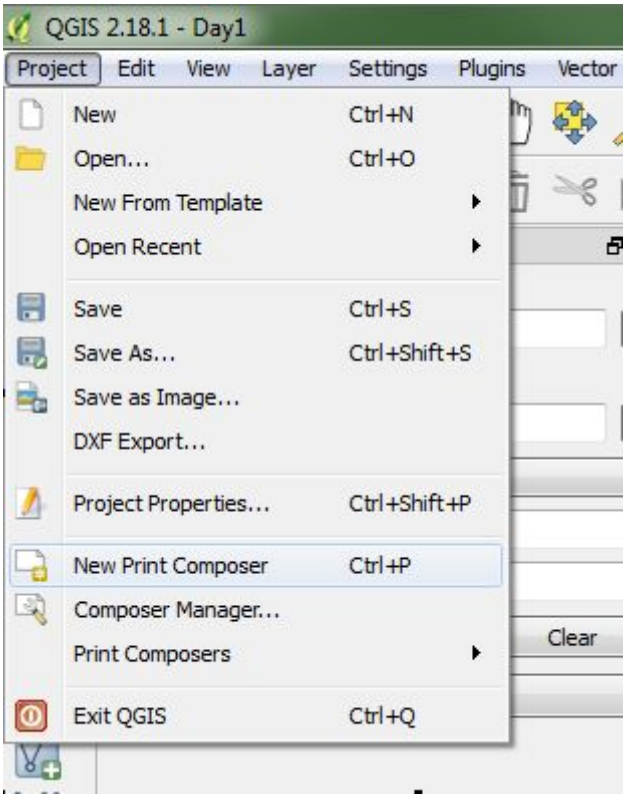
Categorise household level data according to whether or not any pupae or larvae were found in at least one container. Use the graduated colour ramp to colour the blocks according to Breteau Index. When you have finished save your project in your my_work folder, call it day1_myname. (Project, Save as, browse in your USB training_QGIS/my_work/projects_QGIS, type in a file name).

QGIS Print composer

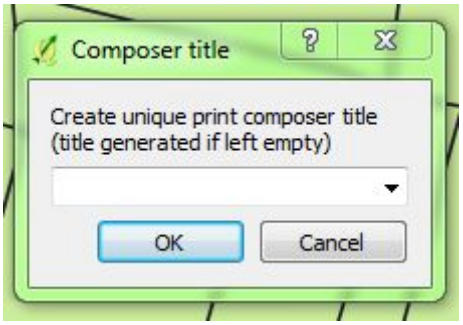
Print composer, as the name suggests, is the feature used to prepare and edit a map ready for printing or export in a format more suited to sharing (for example as an image or PDF). Scale bars, legends and north arrows should be added to your map to give it context and allow the information it displays to be more easily interpreted. Adjustments can be made to the size and position of each component until the desired layout is achieved.

To begin, you will need a map open to work with. For the purpose of this example, use the map created in the preceding exercise or load OpenStreetMaps as a base layer (ref section...).

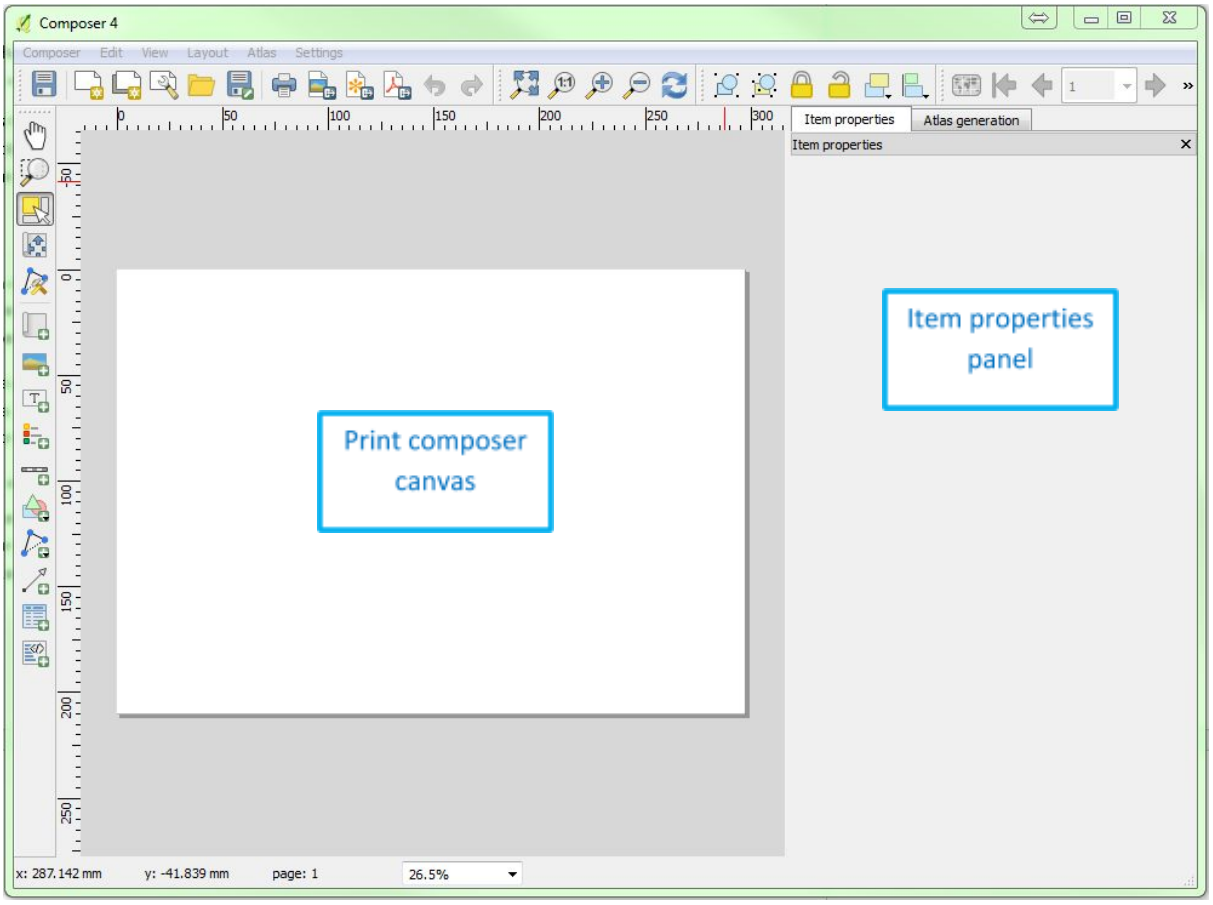
There are various ways to open the print composer, via the menu bar (Project, New Print Composer), using the regular shortcut keys for printing (Ctrl+P), or by clicking on the new print composer shortcut symbol on  the toolbar.



When you open new print composer QGIS will automatically ask you to name your file. Type the title of your map in the pop up box that appears, for example “Print Example”, and click on “OK”.



A new print composer window will then open that looks as follows below. This is your blank canvas to which you will add your map and any other required items, such as a scale and legend. It is also possible to include other elements in your layout like text, images and symbols if required. Note that the print composer window has its own menu options and tool bars, as well as multiple windows within for formatting and other functions. Basic navigation of the canvas is very similar to the layout of the QGIS project layout, with zoom control shortcuts in the tool bars or by simple scrolling of the mouse wheel, and further tools to select different elements or pan across. Hover the cursor over the different buttons to see what they do.



Adding a map to the print composer


This is the 'Add map' button. It is available in your tool bar on the left hand side or can be found via the menu under 'Layout'. If you move your mouse into the print composer canvas, you will see your cursor as the following symbol. . To add your map, hold down the left button of your mouse and drag a rectangle shape on your print canvas. When you release the mouse button your map will be displayed in the rectangle you have created.

Moving and resizing your map

 You can change the size and position of your added map using the “Select/Move Item” button. This will be the default mode after adding your map, but can also be found as an option in your tool bar, via the menu under “Layout”, or using the shortcut key “V”.

To change the position of your map on the page, use your left mouse button to drag and drop your map rectangle. Left-click on your map and hold down the mouse button as you use your mouse to move the map to the desired place on the canvas; release the button when the map is in position.

To make your map smaller or larger, move the cursor of your mouse to the corner of your map, when it changes from a single-headed to a double headed arrow () hold down your left mouse button and move your mouse to make the rectangle of your map smaller or larger. When you are happy with the size, release your mouse button.

 It is also possible to change the position of your map within its box. With this tool selected, you can drag and drop your map to reposition what is visible, or zoom in and out using your mouse scroller. Note that by holding down the Ctrl button when you scroll allows you to zoom by smaller increments.

 This tool is for zooming in and out on your whole print composer canvas. You can use the scroll of your mouse, or drag a rectangle to select the area you wish to zoom in on.

Adding a scale bar

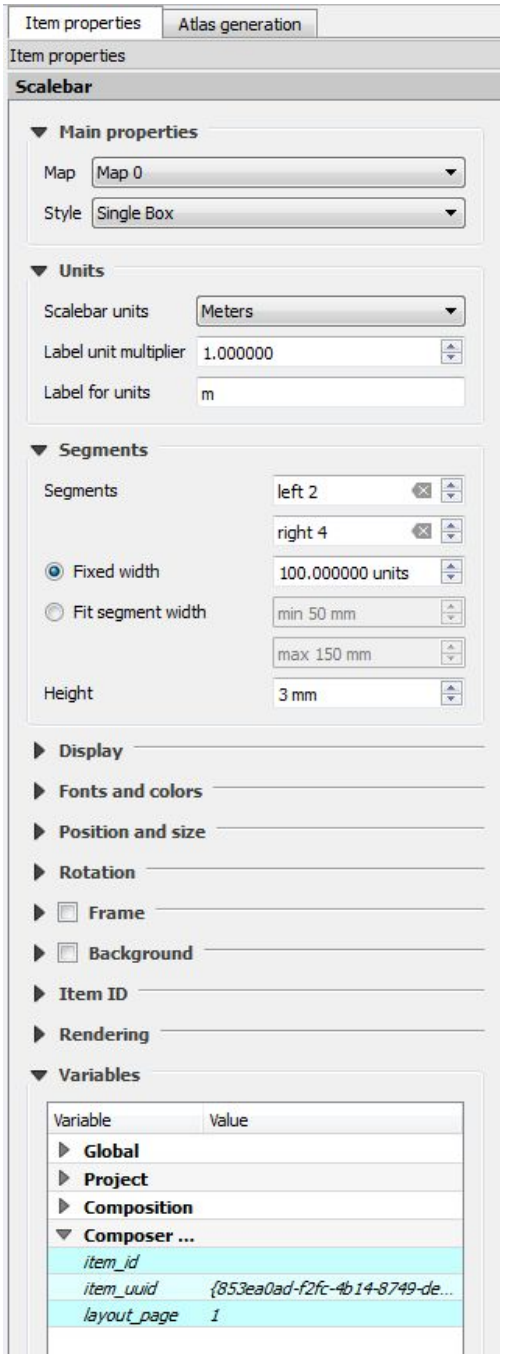
 This is the icon in the toolbar for adding a scale bar to your print canvas. It can also be found in the menu under “Layout”. To add a scale bar, simply click on the icon to select it, then click somewhere in your print canvas and a scale will appear.

The scale bar can be moved by dragging and dropping, as with the map. Use the “Select/Move item” tool and your mouse (clicking and holding the left mouse button). Notice guidelines will appear automatically to help you align the item with other objects on the canvas, such as the map.

To further edit the scale bar, the “item properties” tab should be used. This can be found in a panel on the right hand side of your screen. If it has been closed or hidden, the menu can be used to reopen it; go to “View”, “Panels”, then tick “Item properties”.

Within item properties for the scale bar, there are various subsections available for the customisation of your scale. Note that under “Main properties” there is an option “Map” where you can choose which map the scale bar is attached to (should you have more than one map present), and the option “Style” is also featured, which offers a drop down selection of available scale styles.

It is also possible to select which units will feature on your scale bar. Remember that the scale bar should provide useful information and context for the person reading your map, make sure you choose appropriately. Should the units be in meters, kilometres, nautical miles, feet, miles? Select the option that makes most sense for your map.

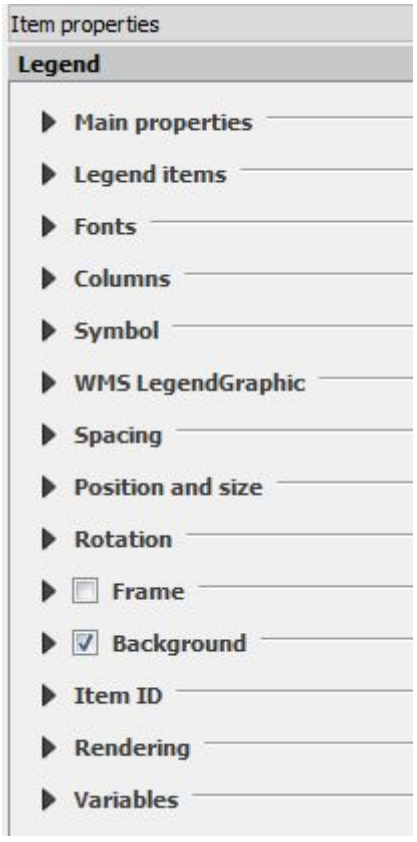


Adding a legend

A legend can provide important information to the map user on how to interpret the information presented. To add a legend to your map in the print composer, choose the “add new legend” icon from the toolbar, or find the “add legend” option under “Layout” in the menu bar. Left-clicking on this option and then clicking within the print canvas will provide you with the legend for your given map.

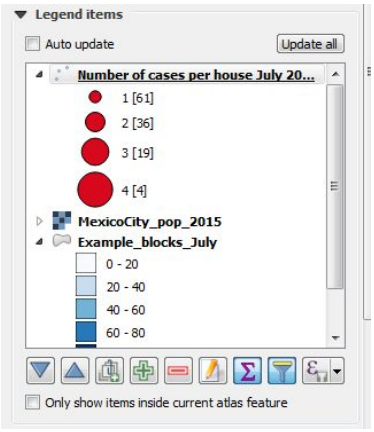
Editing a legend

Your legend should be meaningful to all those who might read your map. Make sure that the labels used are informative and will make sense to others. Units should be displayed where appropriate. Numbers should be formatted correctly. To do all this, the legend can be edited within the “item properties” panel. If you have just added the legend, this item will by default be selected in your canvas. If for whatever reason it is not, use the select/move item tool and click on the legend in the print canvas to select it. The item properties panel will show options for the selected feature only.

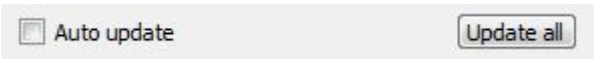


The legend panel in item properties offers various options for modification of your legend (see figure). Click on the triangles to expand each of the categories and reveal the various editing options they contain.

Expanding the “Legend items” category for example will reveal a list of layers and an editing toolbar. The items listed in this panel in the print composer reflect the layers in your main QGIS map canvas.



Changes made in the map canvas can be updated in the print composer by clicking on the “update all” button, or set to update automatically by ticking the “auto update” check box.



Alternatively, adjustments can be made within this panel in the print composer.

 The order of legend items can be changed using the up and down arrow buttons, or by dragging and dropping within the list itself.

 Layers can be added or removed using the plus and minus buttons.

 The text of each item can be edited using the button with a pencil and paper, or by directly double-clicking on the text in question. Use this to make sure your labels are meaningful. For example, what are displayed are your layer names, which by default are the names of those files. This may mean something to you, but may not be clear to others. It will also not reflect any manipulation you have made to how the information is displayed – what do your graduated colours represent? Are there units for the categories chosen? Changing the labels of your legend is a simple but effective way of making your map more informative.

The text of the legend title can edited within “Main properties”, “Title”.

Saving and exporting your map

 When you are happy with your map and its layout in the print composer, save your project using the icon in the toolbar, the option in the menu under “composer”, or the shortcut key “Ctrl”+”S”.

 If you have a printer installed and connected to your computer, you can simply print your map directly by using the print icon in the toolbar, or in the menu under “Composer”, or the shortcut key “Ctrl”+”P”.

 Alternatively, you can export your map as an image (PNG, BMP, TIF, JPEG),

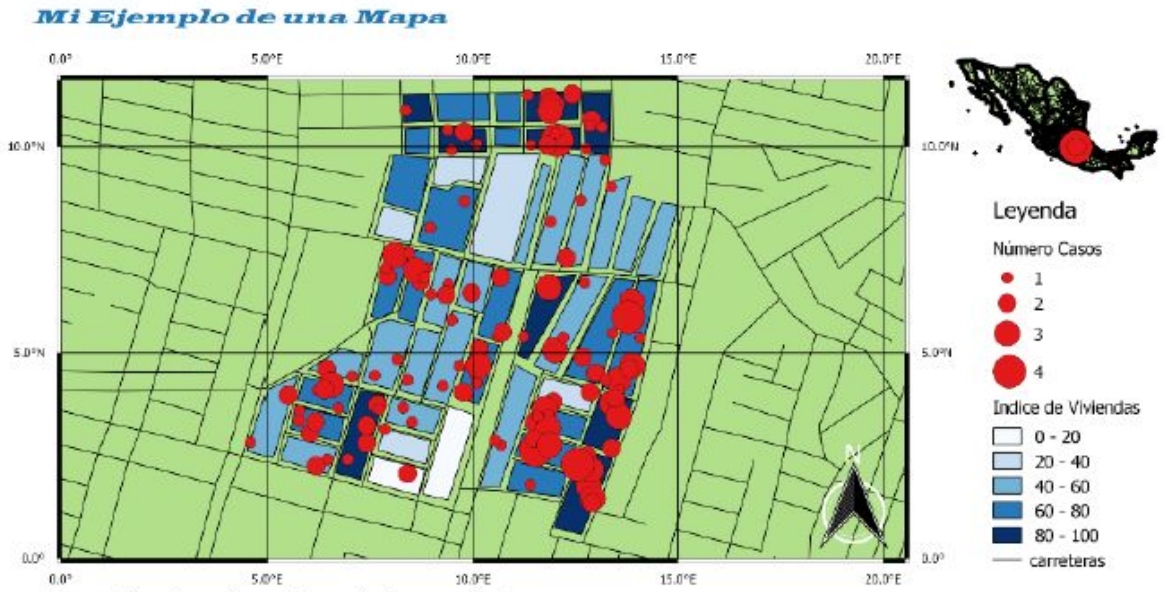
 or as a PDF.

These export options will open a “Save composition as” window, where you will be able to select your desired file type, a destination folder, and give it a name.

Exercise 1.2

Produce a map of the categorised points and graduated colour blocks from Exercise 1.1

Below is an example of some of the things that it is possible to include.



Id_manzana	tot_viv	tot_recip	tot_rec_po	tot_viv_po	iv	ir	ib
9	10	40	23	10	100	57.5	230
7	10	44	19	10	100	43.18181818	190
1	10	45	23	10	100	51.11111111	230
39	10	44	12	9	90	27.27272727	120
4	10	44	18	9	90	40.90909091	180
36	10	40	12	8	80	30	120
10	10	45	19	8	80	42.22222222	190
3	10	49	21	8	80	42.85714286	210
2	10	42	13	8	80	30.95238095	130

End of Day 1