Visual Communications

The art of showing complex data and messages effectively

Roberto Foa

foa@fas.harvard.edu

today

1. Principles of data presentation

2. A practical guide to GIS mapping

today

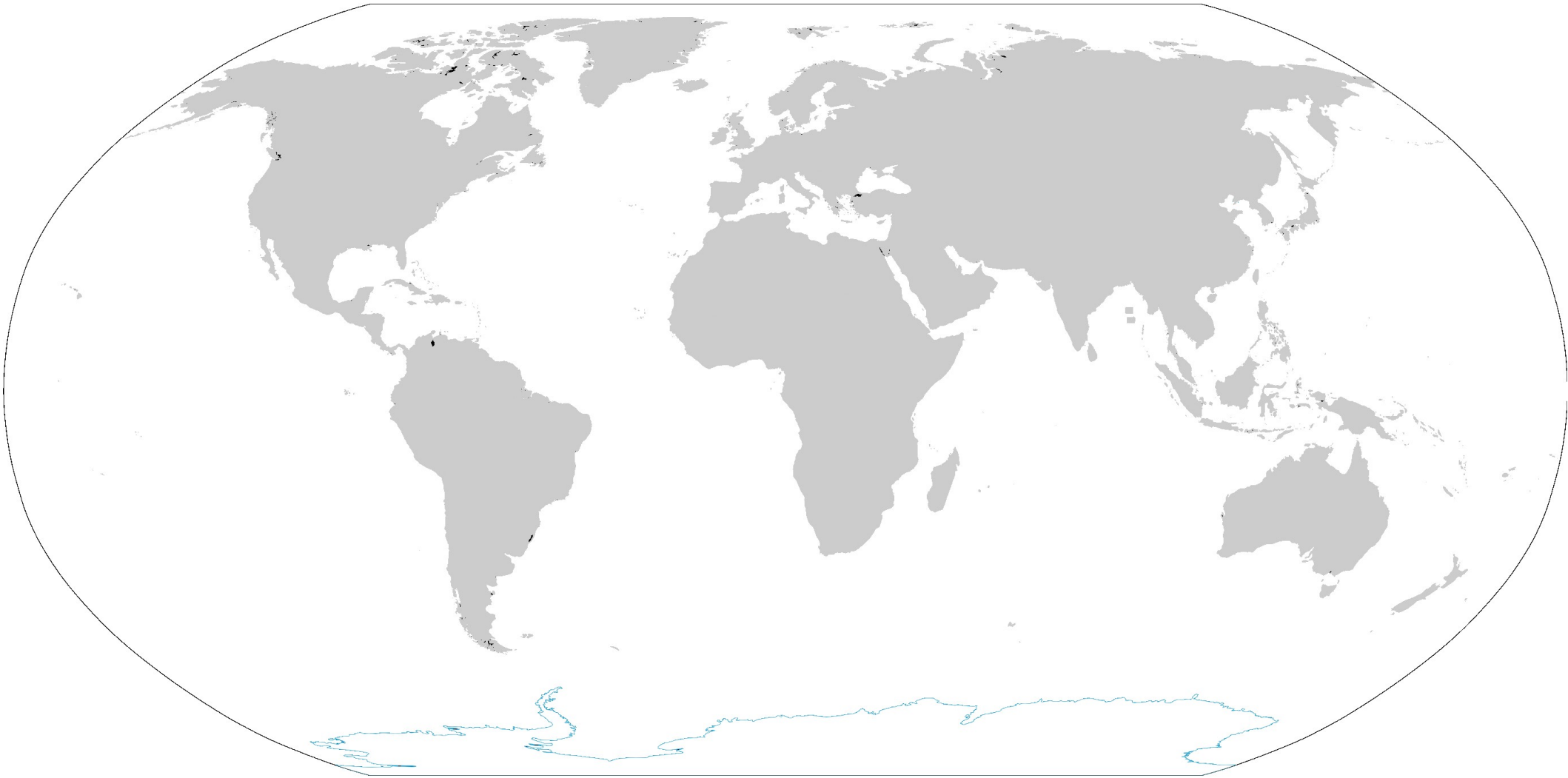
1. Principles of data presentation

2. A practical guide to GIS mapping

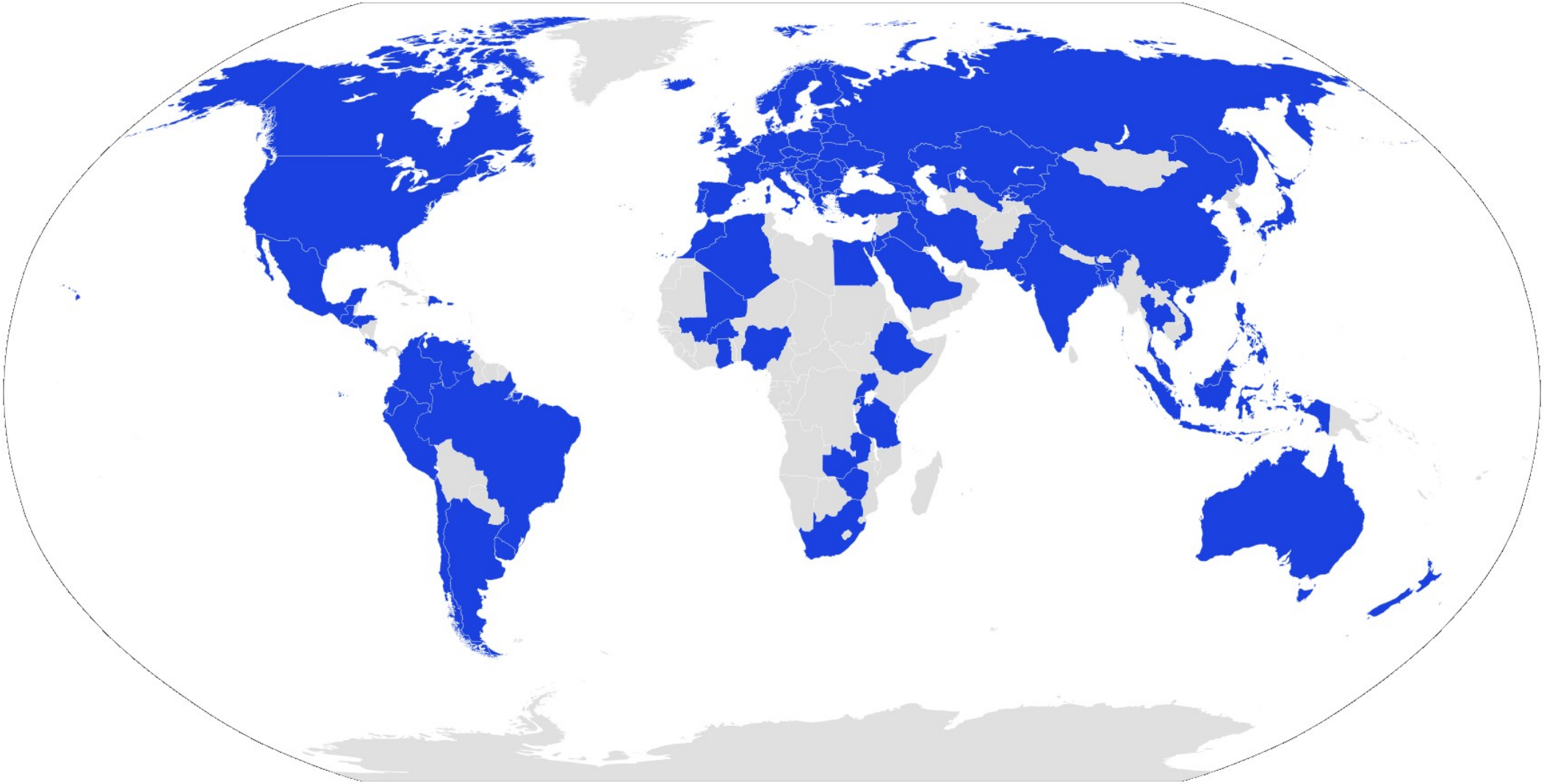
Automated Mapping

- One of the big breakthroughs in visual data representation is the rise of automated mapping.
- Mapping software (ARCGis, QGis) can fit data to a map quickly using shapefiles (.shp).
- About 5 years ago the more common method was to use a scalable vector (.svg) file and replace the tags for geographic spaces.
- This is also fairly easy.

A Very Simple Automated World Map

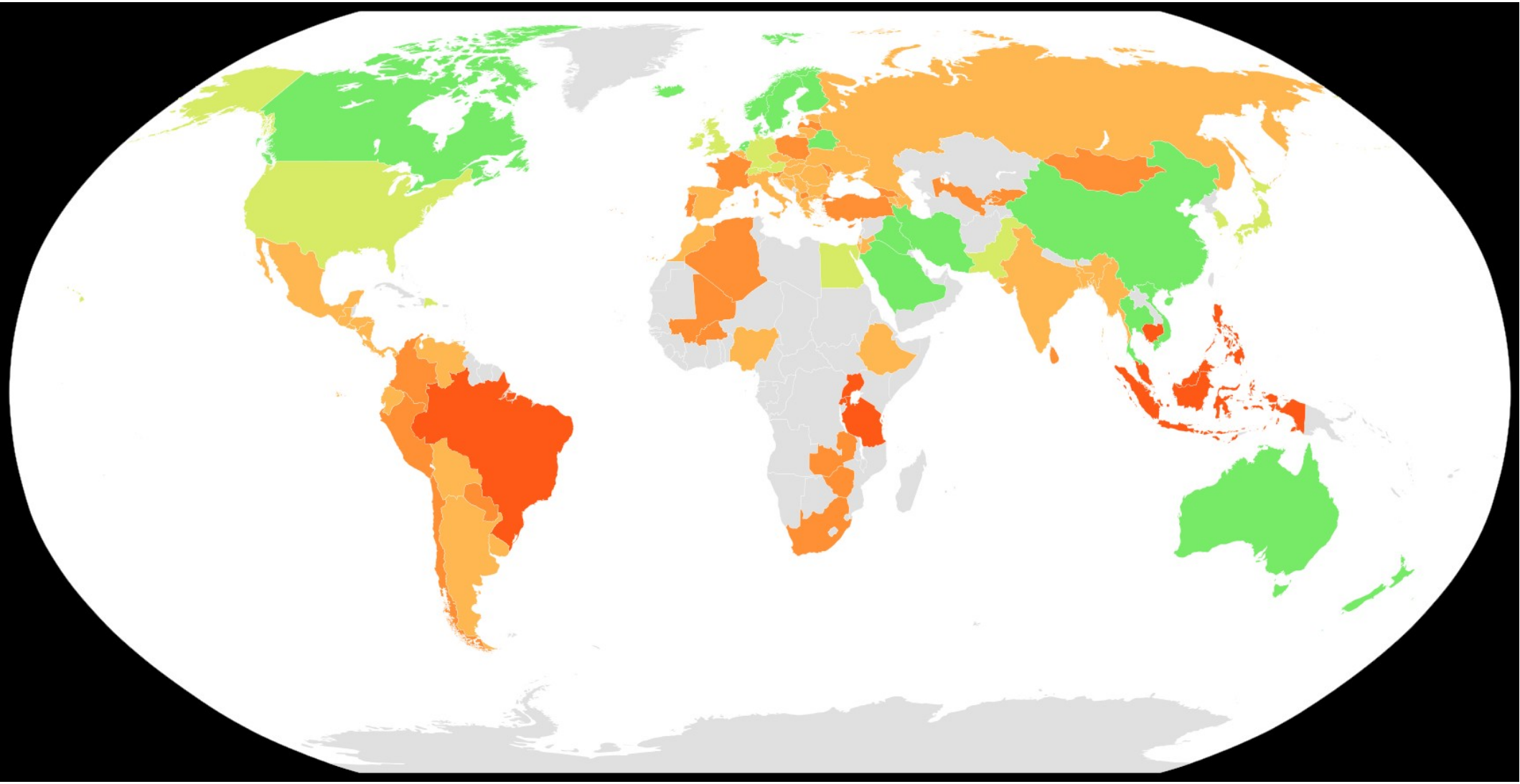


A Very Simple Automated World Map



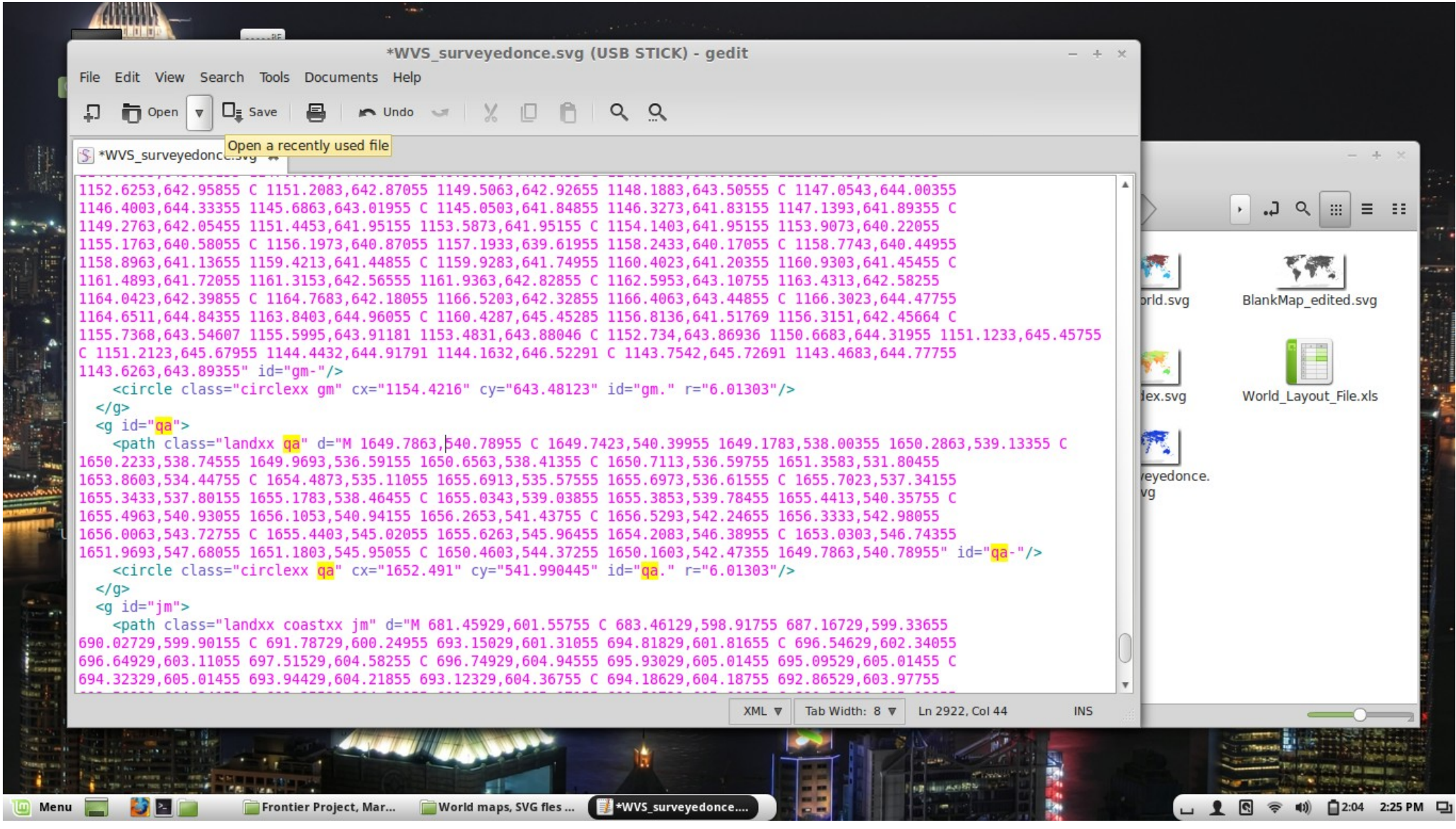
Countries surveyed once by the WVS

Mapping WVS Items

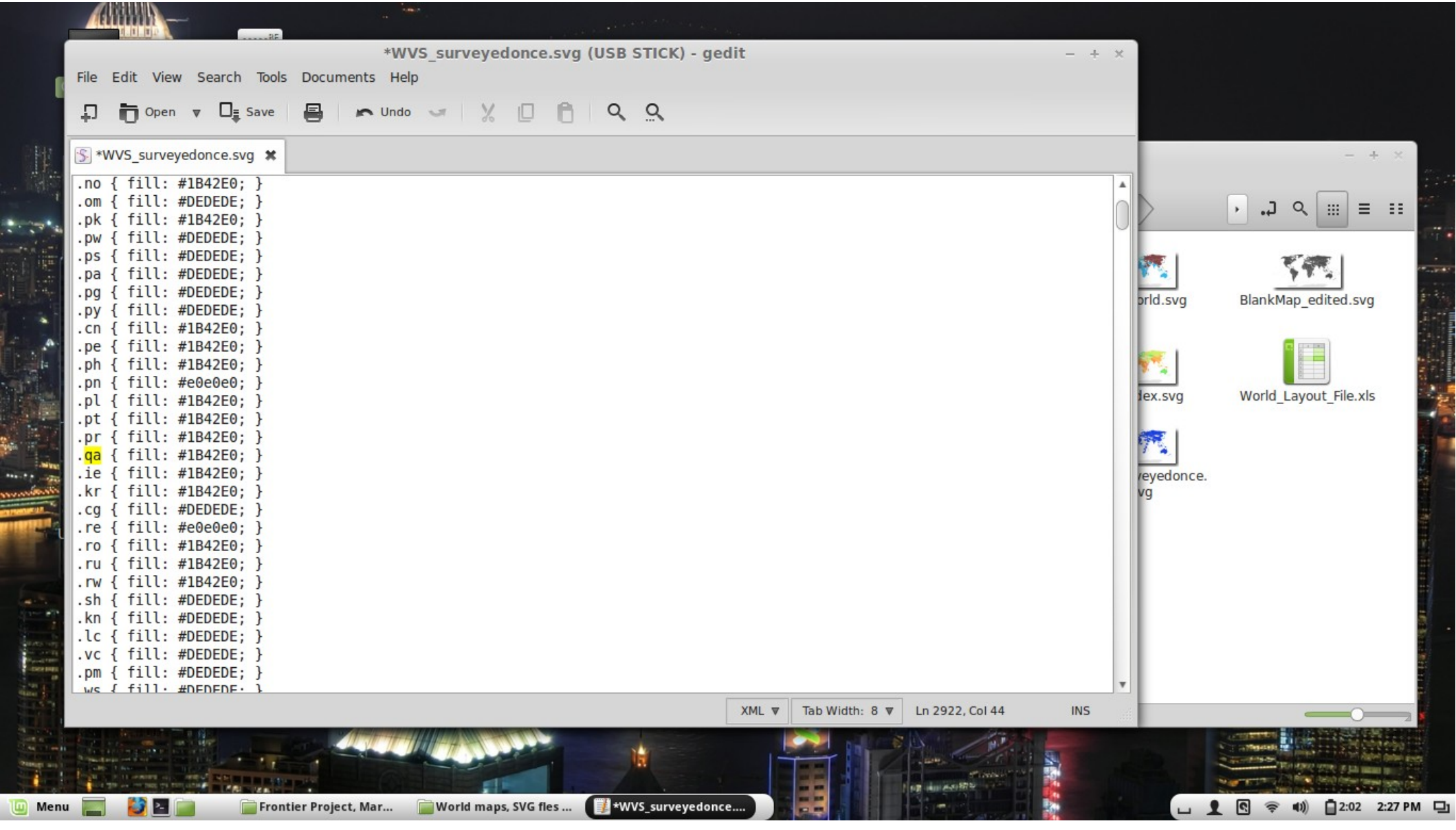


Social Trust Map of the World. Percentage Responding “People can generally be trusted”.
Source: World Values Survey (2000-5), Global Barometer surveys.

How to do it (1) – SVG Files



How to do it (1) – SVG Files



How to do it (2) – Use GIS Software

- A faster way is to use GIS software, such as ARCGIS (commercial) or Quantum GIS (free).

- These use shapefiles (.shp) to get the boundary outlines (like .svg files) but allow dynamic updating of the map, without having to toggle manually with the marker id's.

- Is also a lot more convenient for getting new maps (doesn't involve complex fiddling around with the raw data).

Getting going with QGIS

- Quantum GIS (QGIS) is a good program to get started with, as it has much of the functionality of ArcGIS but is also free to download.

- The software can be downloaded at the QGIS website (www.qgis.org).

- In this presentation I will assume you are using QGIS - though the steps for ArcGIS would be largely the same.

Getting going with QGIS

- The first thing you need is to get a 'shapefile' (extension .shp).

- A shapefile is simply a set of map coordinates and instructions of how to join them in order to make some basic shapes – such as the outlines of countries, or subnational country regions.

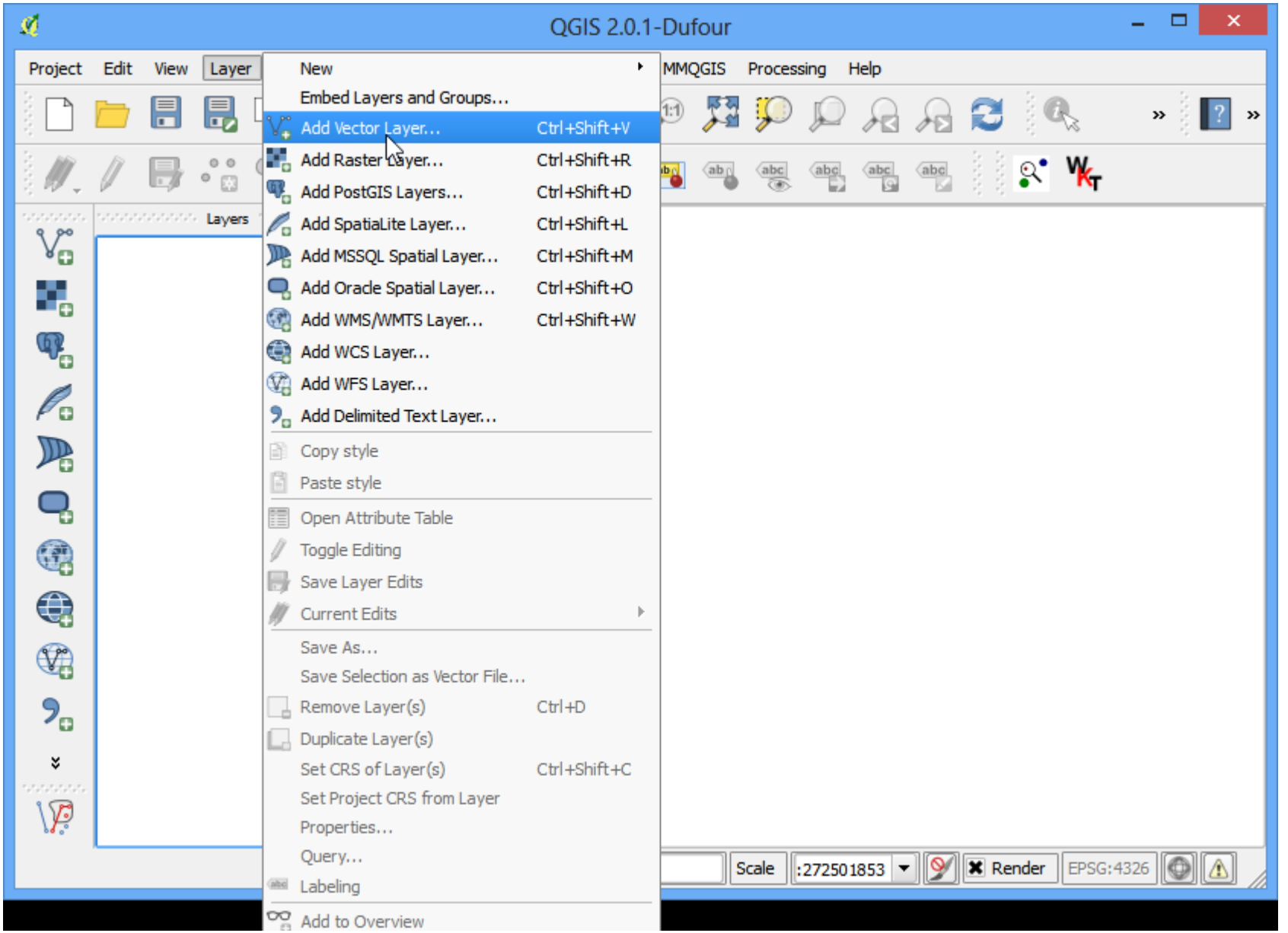
- Many common shapefiles are available for free online.

Getting going with QGIS

- Let's assume you have found a shapefile, and want to start using it for visualisation of your data.

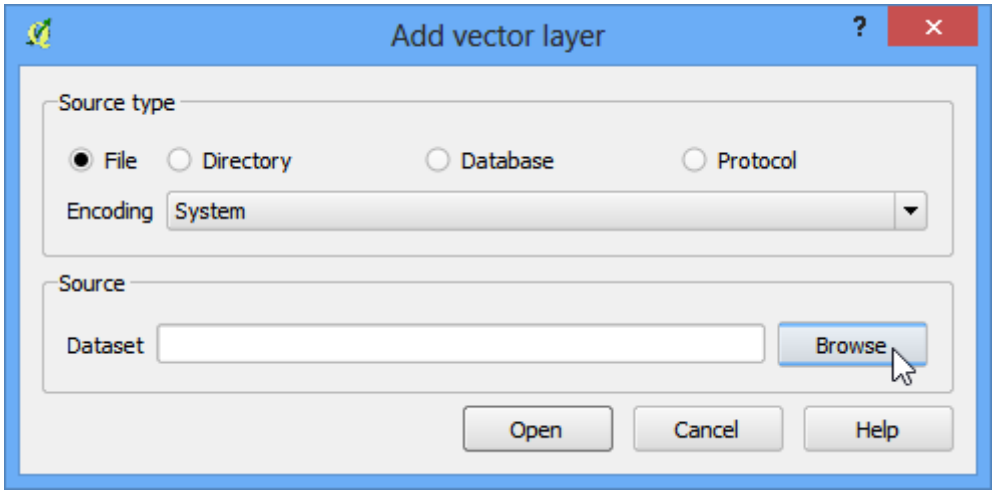
- The first thing is to add it as a vector layer – in other words, one level of the data/graphic file you'll be making.

Adding a vector layer

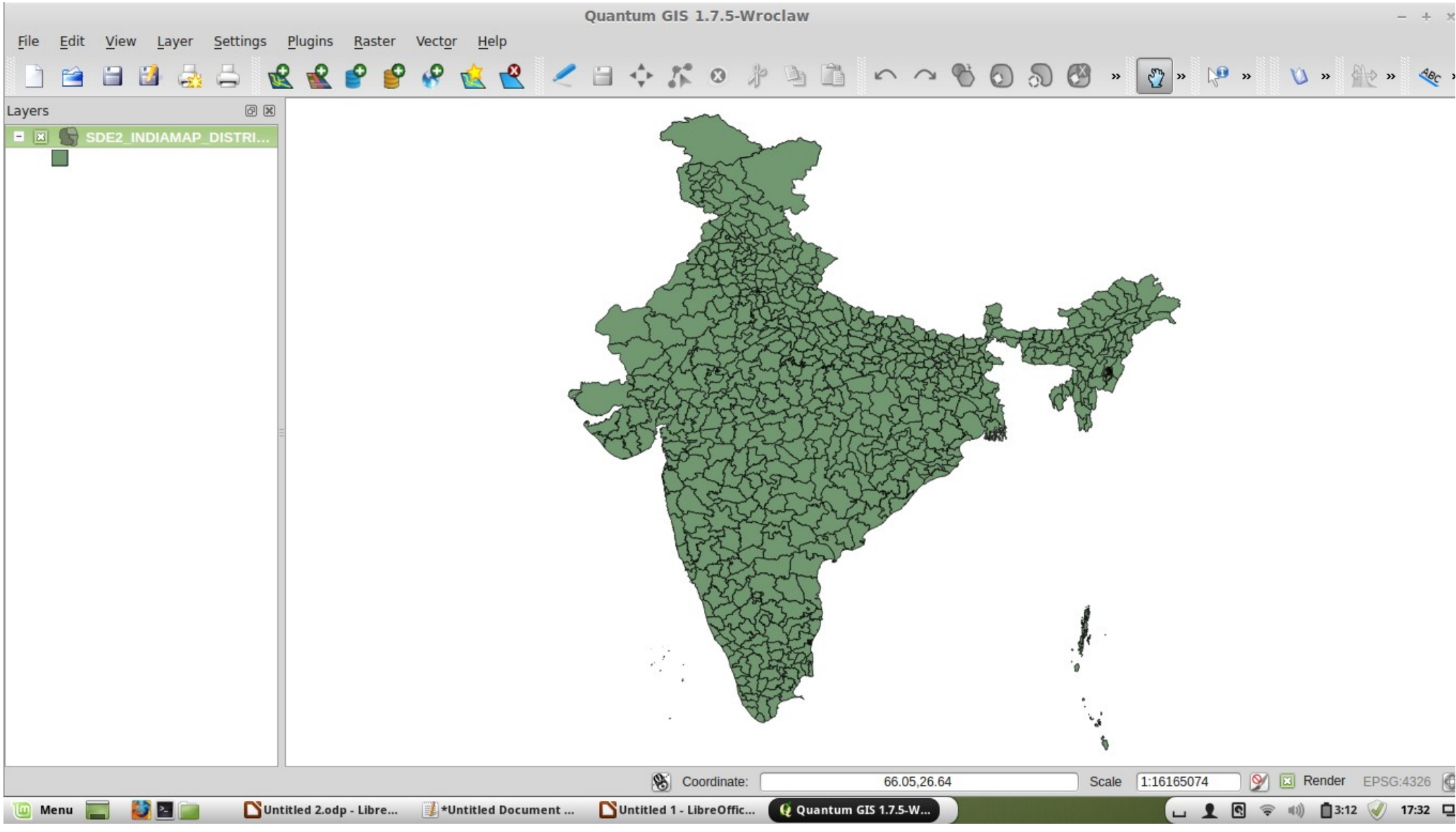


Getting going with QGIS

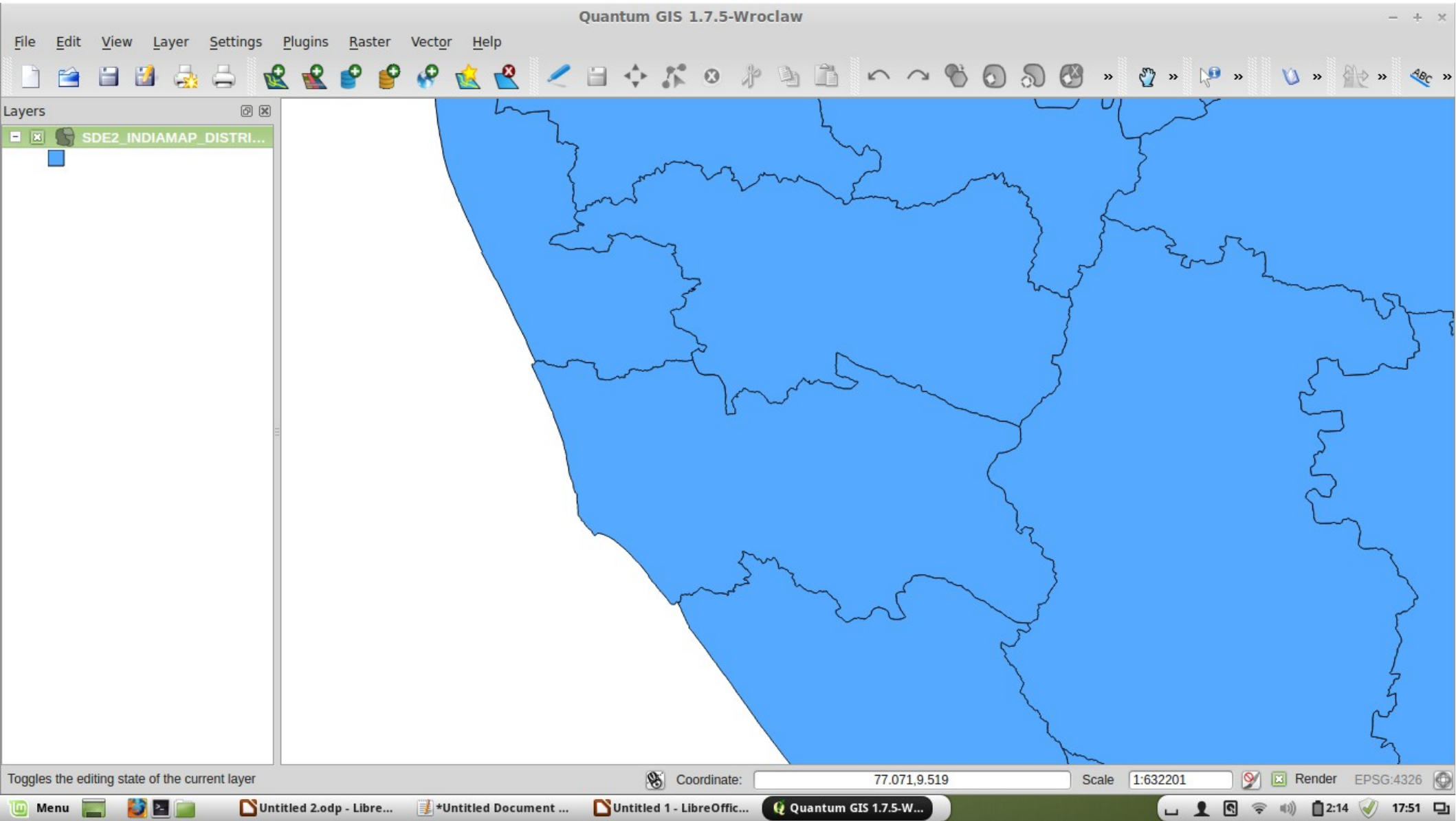
- This is how you load up the shapefile.



The shapefile, loaded



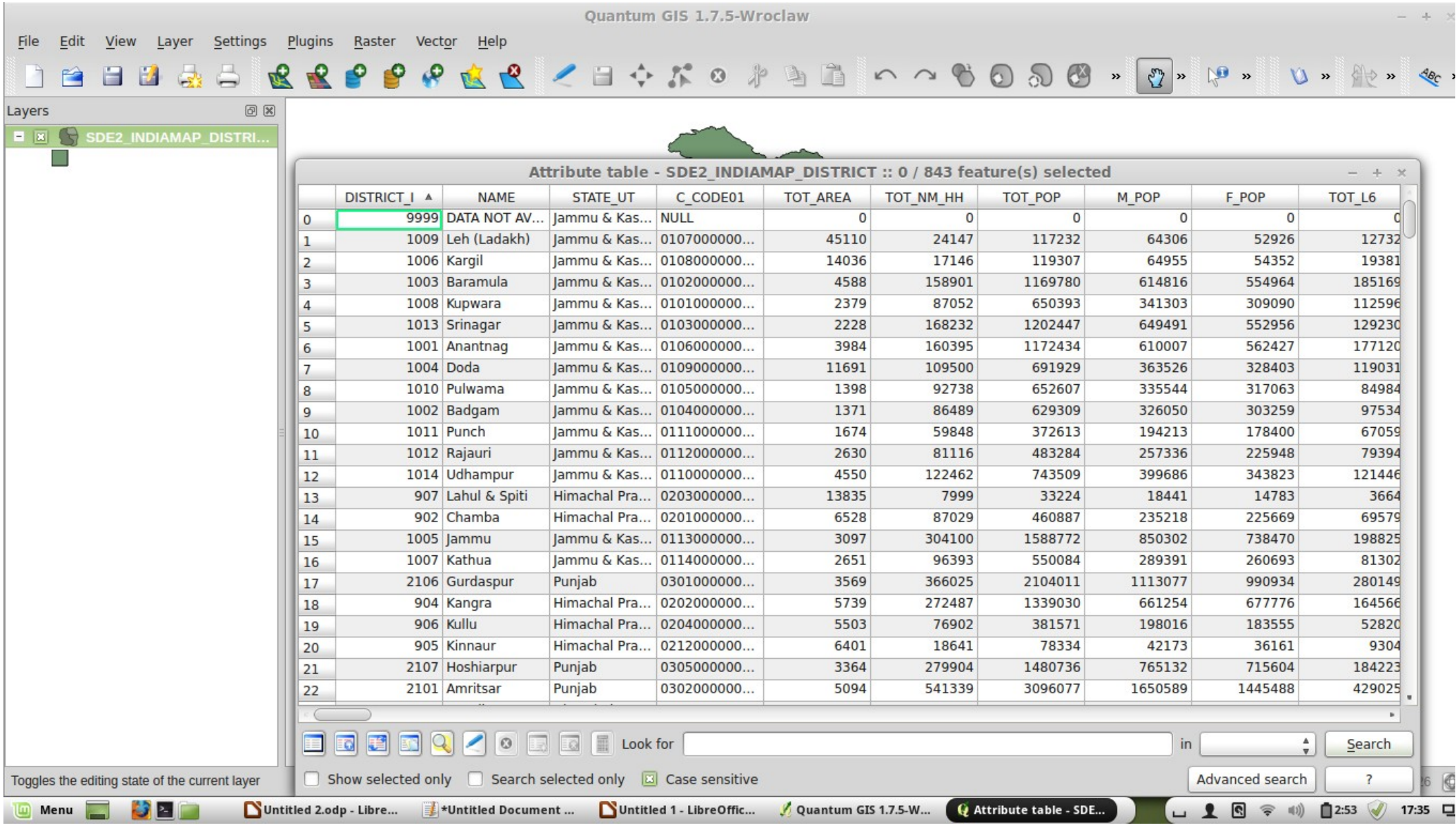
Zoom in, fairly detailed



Embedded data

- Often shapefiles already contain data embedded within them, that we can start using in order to generate maps.

- As this is a 2011 India census file, it contains a data table with a number of variables already.

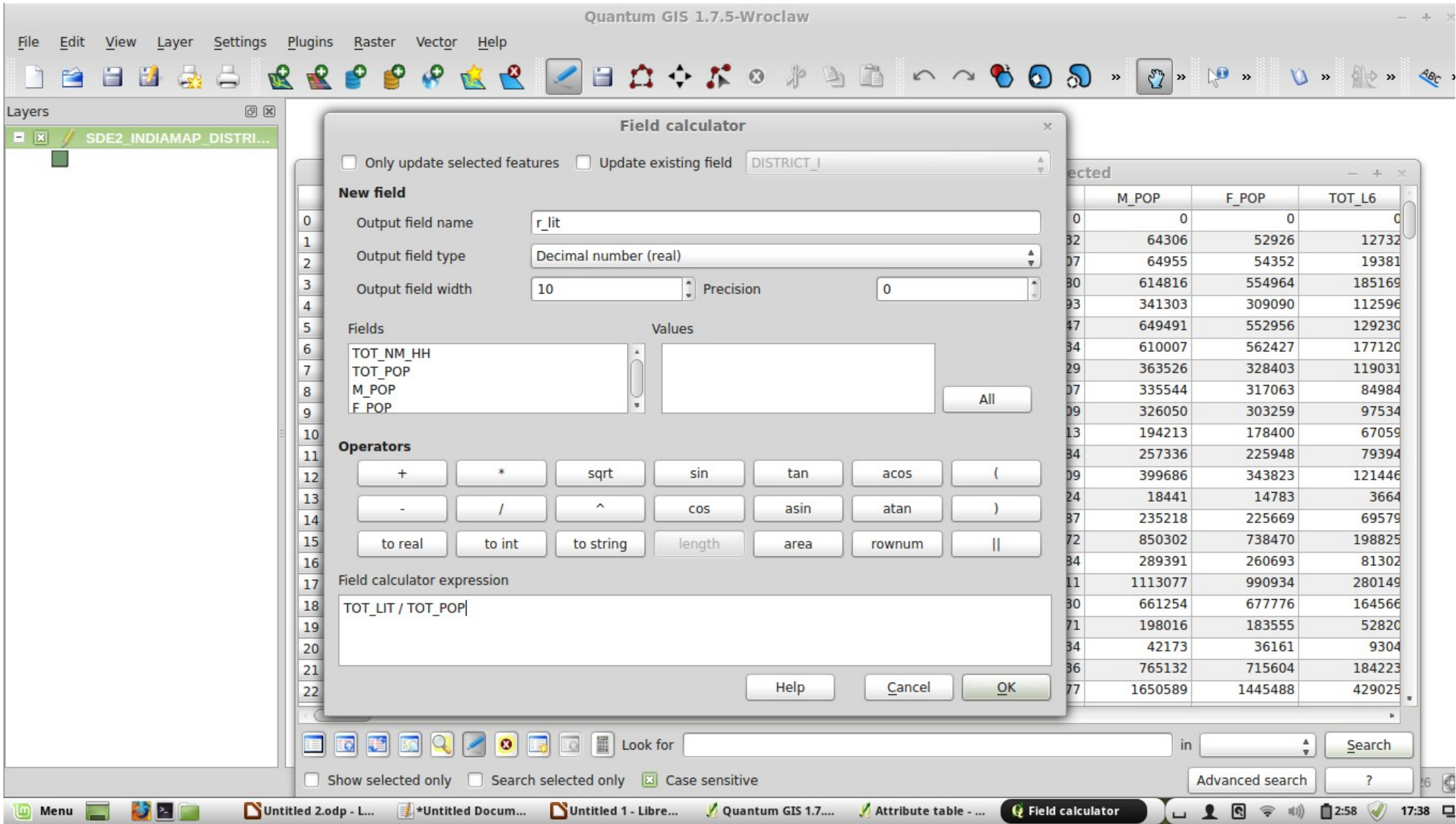


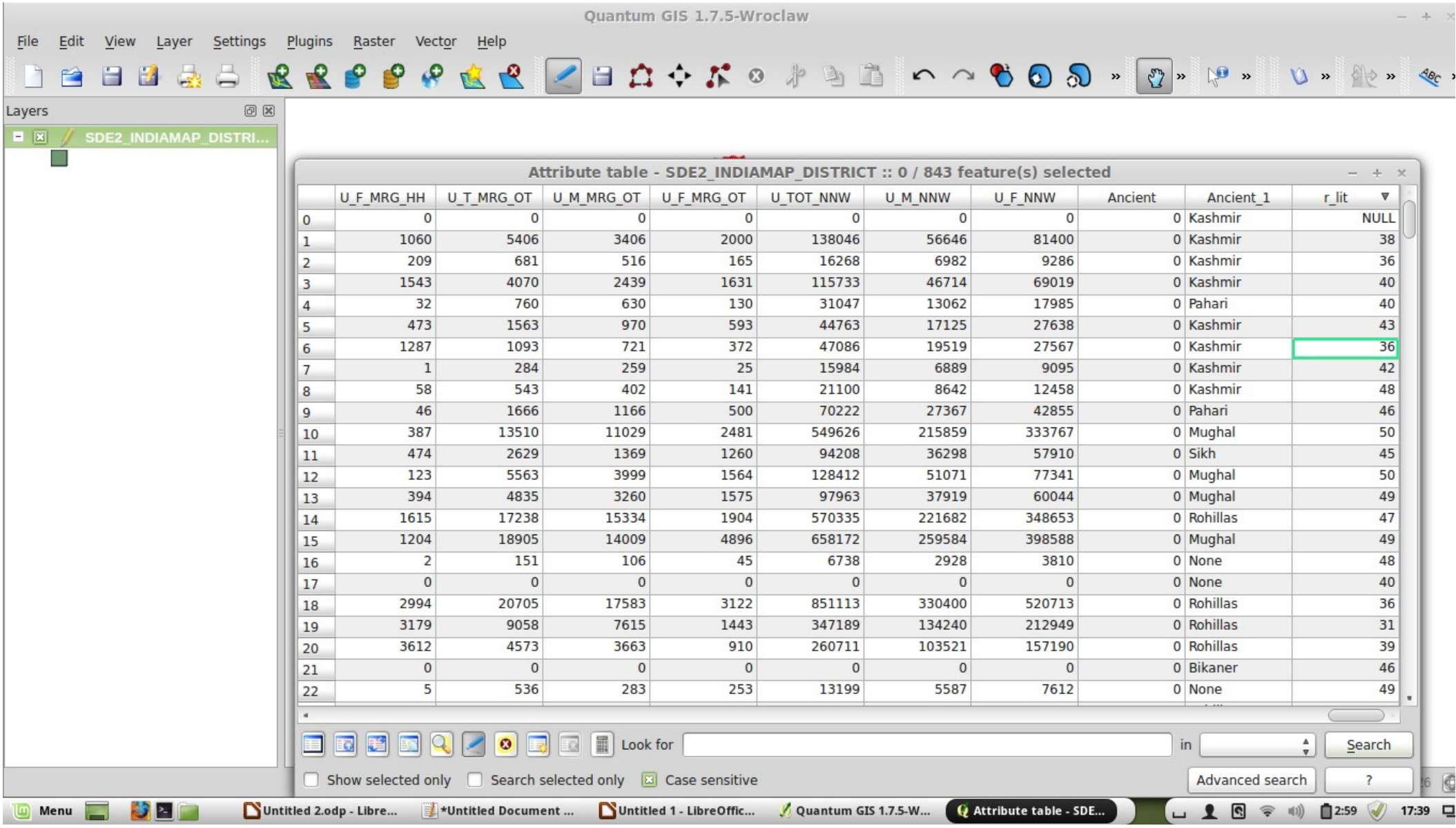
	DISTRICT_I	NAME	STATE_UT	C_CODE01	TOT_AREA	TOT_NM_HH	TOT_POP	M_POP	F_POP	TOT_L6
0	9999	DATA NOT AV...	Jammu & Kas...	NULL	0	0	0	0	0	0
1	1009	Leh (Ladakh)	Jammu & Kas...	0107000000...	45110	24147	117232	64306	52926	12732
2	1006	Kargil	Jammu & Kas...	0108000000...	14036	17146	119307	64955	54352	19381
3	1003	Baramula	Jammu & Kas...	0102000000...	4588	158901	1169780	614816	554964	185169
4	1008	Kupwara	Jammu & Kas...	0101000000...	2379	87052	650393	341303	309090	112596
5	1013	Srinagar	Jammu & Kas...	0103000000...	2228	168232	1202447	649491	552956	129230
6	1001	Anantnag	Jammu & Kas...	0106000000...	3984	160395	1172434	610007	562427	177120
7	1004	Doda	Jammu & Kas...	0109000000...	11691	109500	691929	363526	328403	119031
8	1010	Pulwama	Jammu & Kas...	0105000000...	1398	92738	652607	335544	317063	84984
9	1002	Badgam	Jammu & Kas...	0104000000...	1371	86489	629309	326050	303259	97534
10	1011	Punch	Jammu & Kas...	0111000000...	1674	59848	372613	194213	178400	67059
11	1012	Rajauri	Jammu & Kas...	0112000000...	2630	81116	483284	257336	225948	79394
12	1014	Udhampur	Jammu & Kas...	0110000000...	4550	122462	743509	399686	343823	121446
13	907	Lahul & Spiti	Himachal Pra...	0203000000...	13835	7999	33224	18441	14783	3664
14	902	Chamba	Himachal Pra...	0201000000...	6528	87029	460887	235218	225669	69579
15	1005	Jammu	Jammu & Kas...	0113000000...	3097	304100	1588772	850302	738470	198825
16	1007	Kathua	Jammu & Kas...	0114000000...	2651	96393	550084	289391	260693	81302
17	2106	Gurdaspur	Punjab	0301000000...	3569	366025	2104011	1113077	990934	280149
18	904	Kangra	Himachal Pra...	0202000000...	5739	272487	1339030	661254	677776	164566
19	906	Kullu	Himachal Pra...	0204000000...	5503	76902	381571	198016	183555	52820
20	905	Kinnaur	Himachal Pra...	0212000000...	6401	18641	78334	42173	36161	9304
21	2107	Hoshiarpur	Punjab	0305000000...	3364	279904	1480736	765132	715604	184223
22	2101	Amritsar	Punjab	0302000000...	5094	541339	3096077	1650589	1445488	429025

Generating variables internally

- We can find aggregates like total population, and total number of literates.
- This is a good opportunity to show how to manipulate variables from within QGIS.
- QGIS has a cumbersome interface called the field calculator. Alas it doesn't allow copy and paste from a spreadsheet, so this is how we have to alter variables.

Using the field calculator



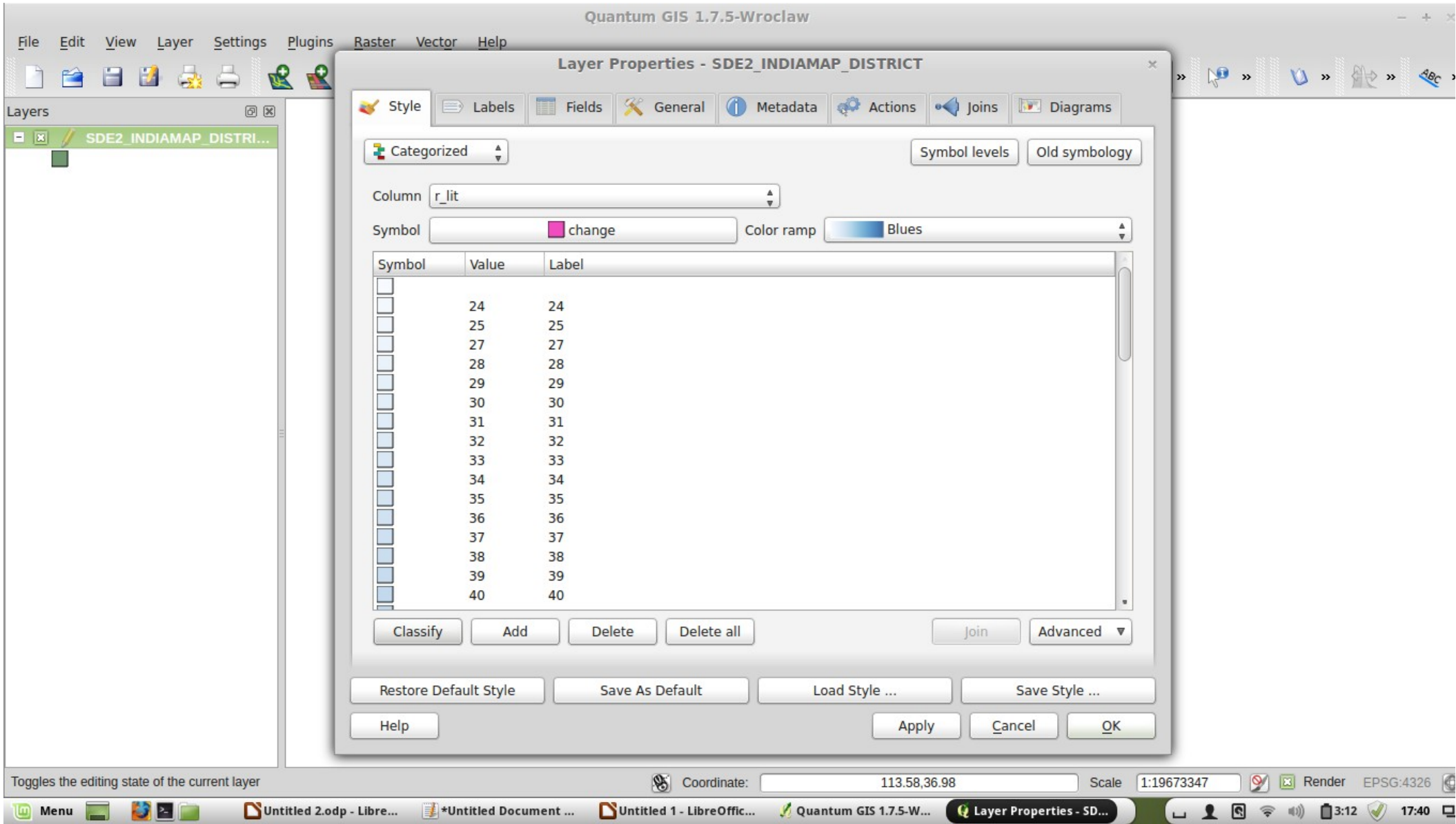


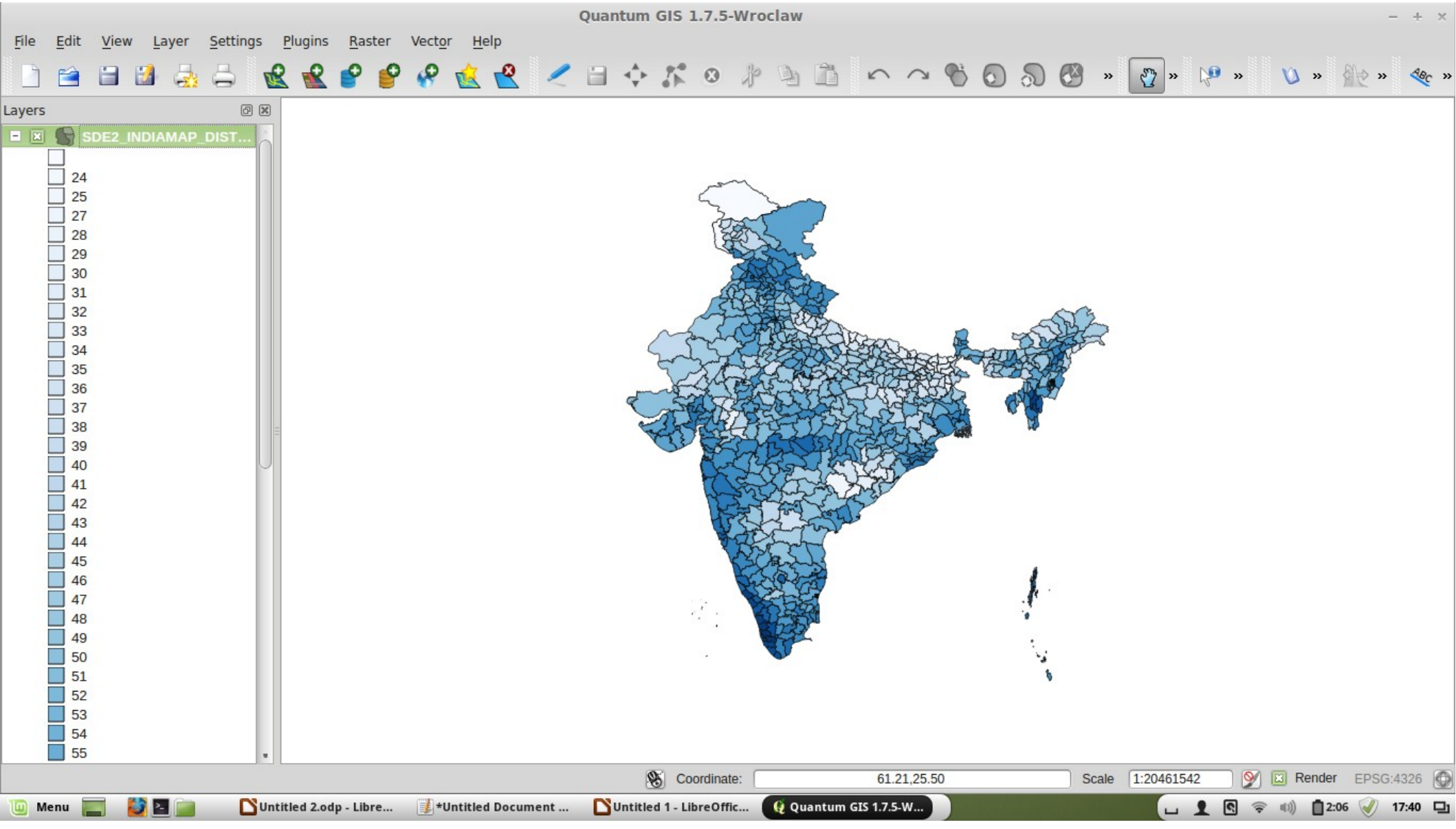
	U_F_MRG_HH	U_T_MRG_OT	U_M_MRG_OT	U_F_MRG_OT	U_TOT_NNW	U_M_NNW	U_F_NNW	Ancient	Ancient_1	r_lit
0	0	0	0	0	0	0	0	0	Kashmir	NULL
1	1060	5406	3406	2000	138046	56646	81400	0	Kashmir	38
2	209	681	516	165	16268	6982	9286	0	Kashmir	36
3	1543	4070	2439	1631	115733	46714	69019	0	Kashmir	40
4	32	760	630	130	31047	13062	17985	0	Pahari	40
5	473	1563	970	593	44763	17125	27638	0	Kashmir	43
6	1287	1093	721	372	47086	19519	27567	0	Kashmir	36
7	1	284	259	25	15984	6889	9095	0	Kashmir	42
8	58	543	402	141	21100	8642	12458	0	Kashmir	48
9	46	1666	1166	500	70222	27367	42855	0	Pahari	46
10	387	13510	11029	2481	549626	215859	333767	0	Mughal	50
11	474	2629	1369	1260	94208	36298	57910	0	Sikh	45
12	123	5563	3999	1564	128412	51071	77341	0	Mughal	50
13	394	4835	3260	1575	97963	37919	60044	0	Mughal	49
14	1615	17238	15334	1904	570335	221682	348653	0	Rohillas	47
15	1204	18905	14009	4896	658172	259584	398588	0	Mughal	49
16	2	151	106	45	6738	2928	3810	0	None	48
17	0	0	0	0	0	0	0	0	None	40
18	2994	20705	17583	3122	851113	330400	520713	0	Rohillas	36
19	3179	9058	7615	1443	347189	134240	212949	0	Rohillas	31
20	3612	4573	3663	910	260711	103521	157190	0	Rohillas	39
21	0	0	0	0	0	0	0	0	Bikaner	46
22	5	536	283	253	13199	5587	7612	0	None	49

Gradiating a map to some variable

- We now have a variable for the literacy rate, running from 0-100%.
- How to colour the map?
- For this we need to click on layer → properties.
- This will bring up the layer properties.

Setting a gradient variable




Quantum GIS 1.7.5-Wroclaw
File
Edit
View
Layer
Settings
Plugins
Raster
Vector
Help
Layers
SDE2_INDIAMAP_DIST...
24
25
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
Coordinate:
61.21,25.50
Scale
1:20461542
Render
EPSG:4326
Menu
Untitled 2.odp - Libre...
*Untitled Document ...
Untitled 1 - LibreOffic...
Quantum GIS 1.7.5-W...
2:06
17:40

Doing something with the map

- If we want to use the map (in a document or presentation) it can be exported using file → save as image.
- Note that it will take whatever is currently in the screen view.
- This can be useful when seeking to export only a portion of an image, but otherwise not.

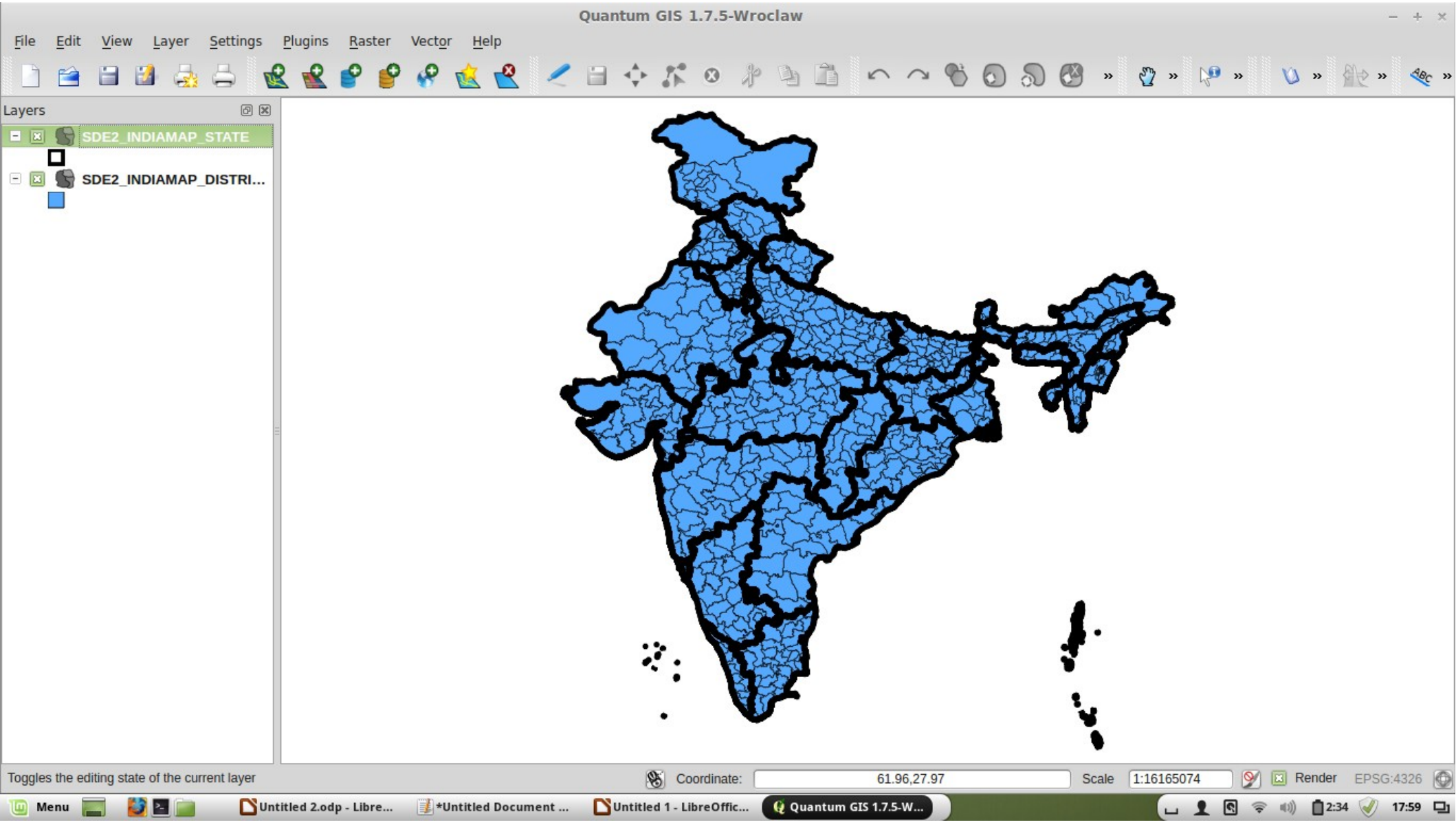
Multiple layers

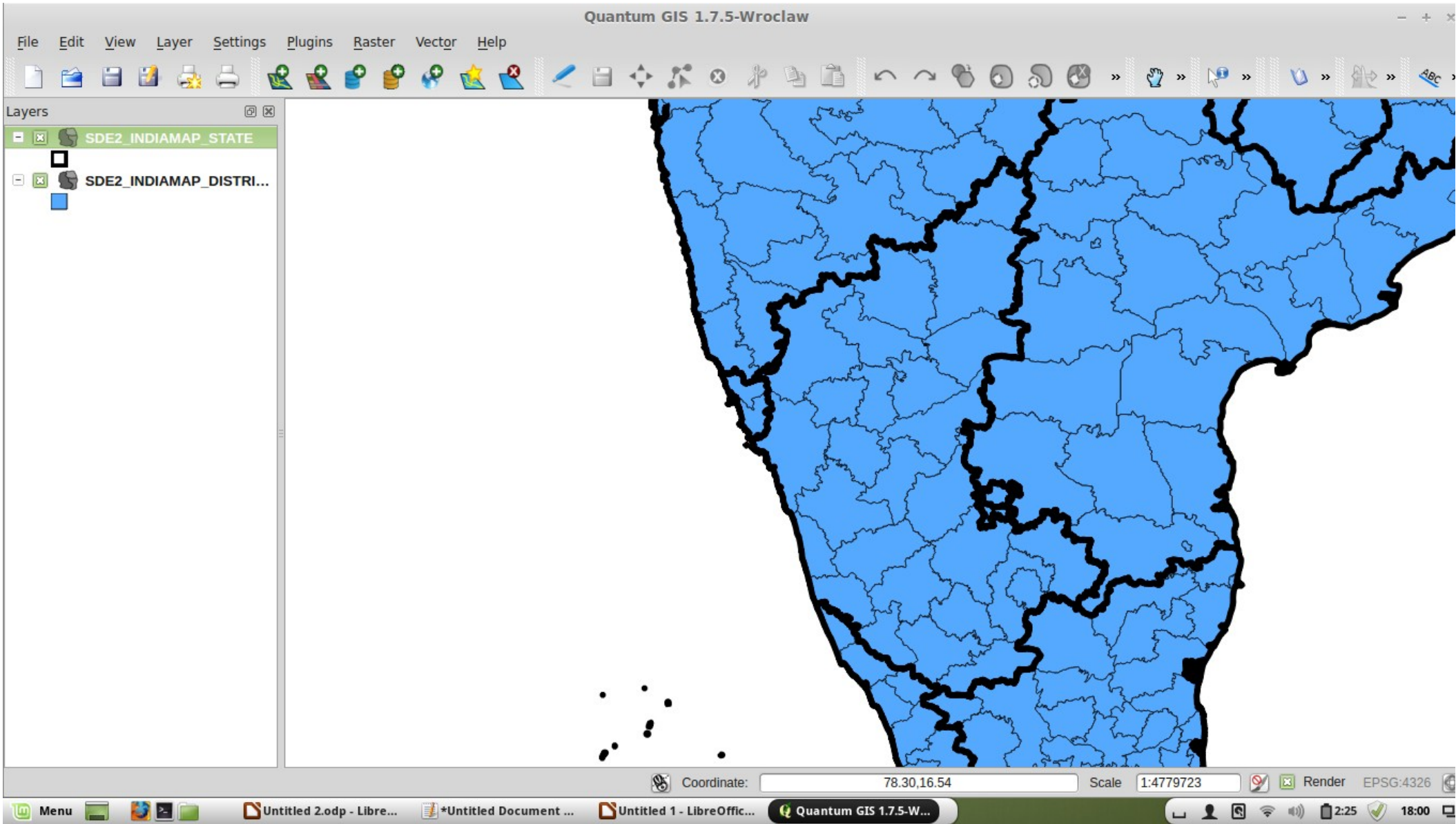
- Because the data is georeferenced, we can have multiple layers.
- An obvious extension is to have multiple layers of administrative boundary – e.g. a country map within a world map, or sub-district boundaries within a single district.
- This is really limited by the memory capacity of your computer (as can become quite resource intense)

Multiple layers

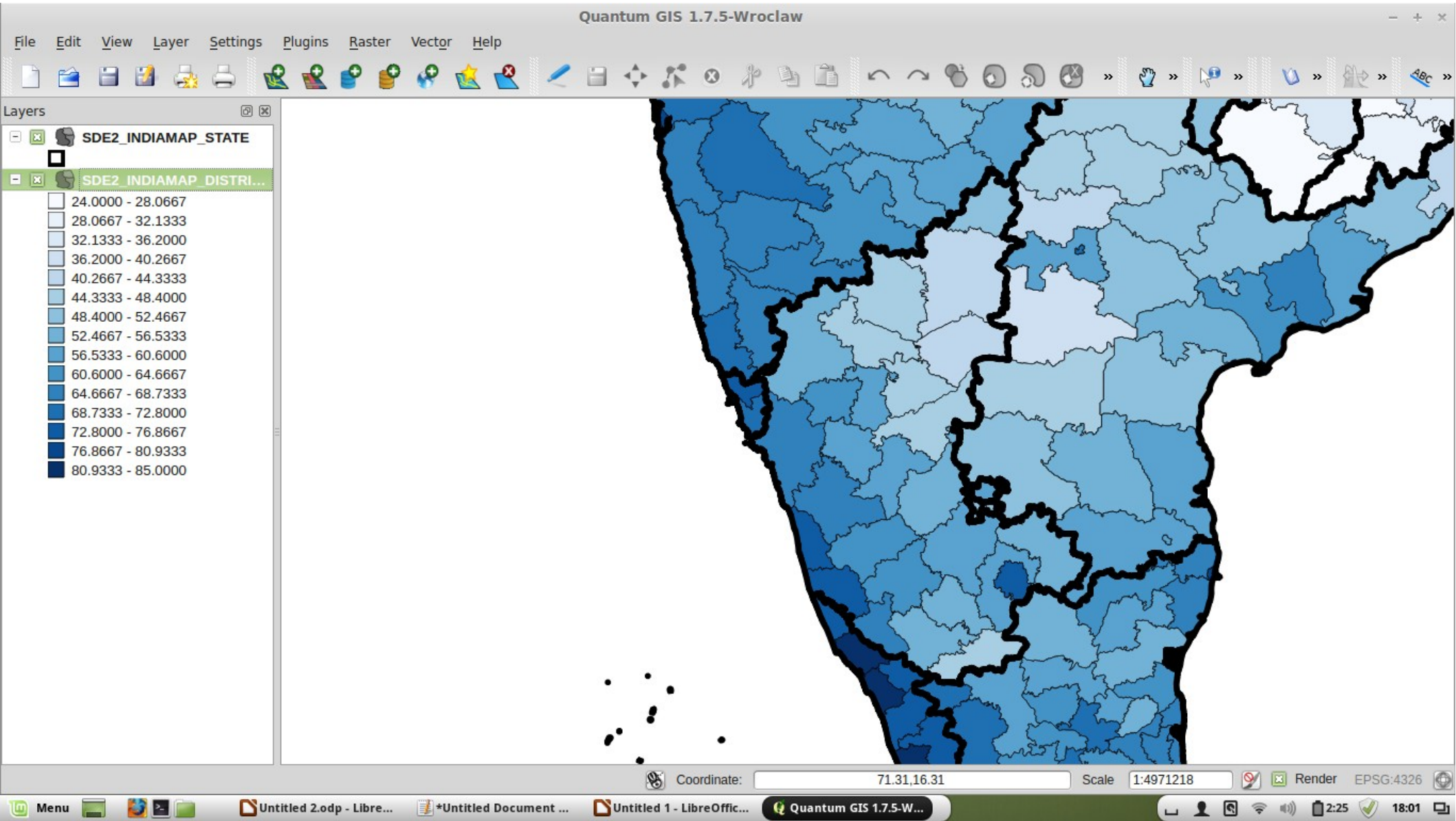
- Here I can add the federal state boundaries as well as the districts.
- This is done by downloading the federal state shapefile, and importing it to the workspace the same way that I imported the district boundaries file.

Two shapefiles together



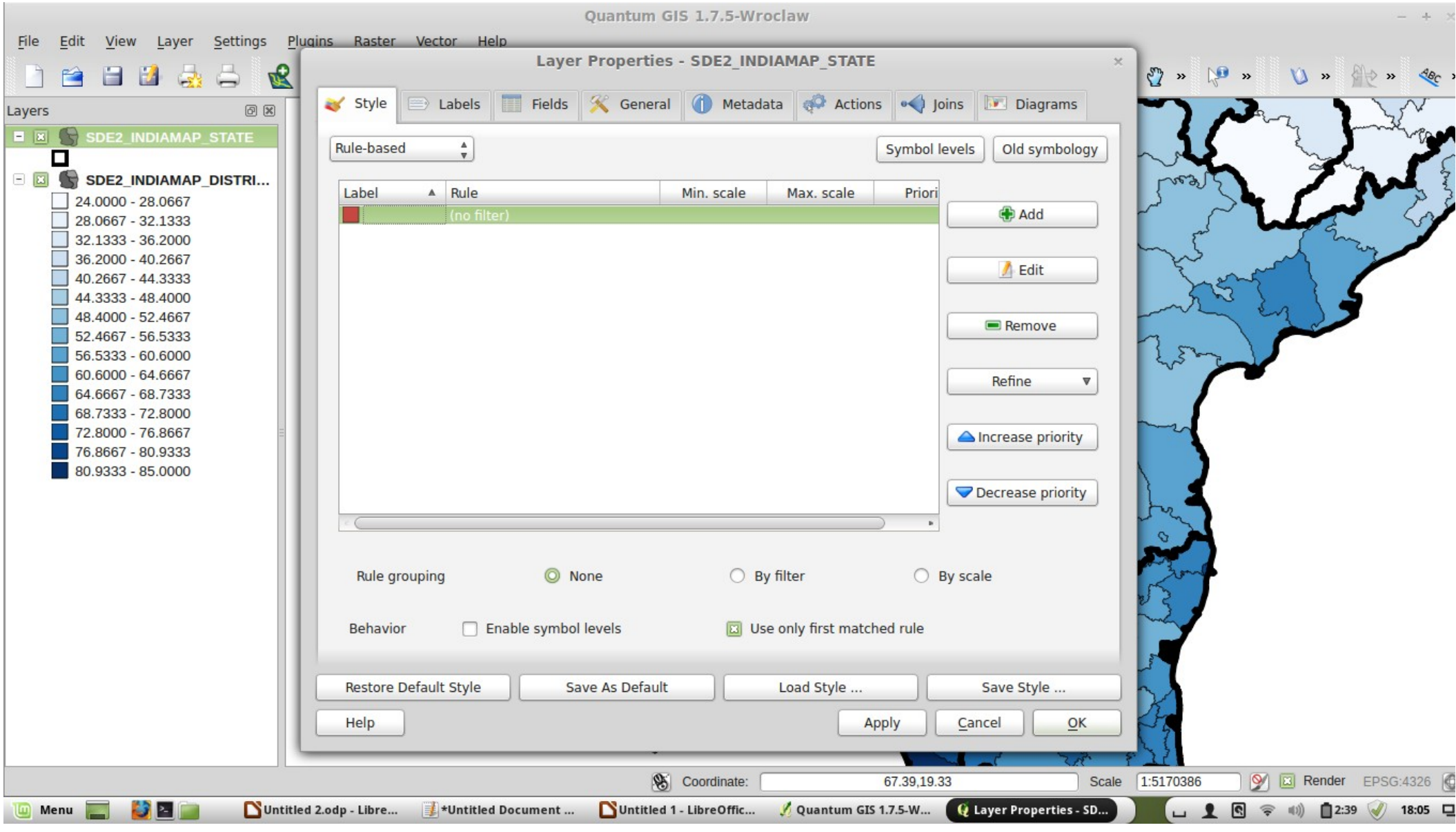

Quantum GIS 1.7.5-Wroclaw
File Edit View Layer Settings Plugins Raster Vector Help
Layers
SDE2_INDIAMAP_STATE
SDE2_INDIAMAP_DISTRI...
Coordinate: 78.30,16.54
Scale 1:4779723
Render
EPSG:4326
Menu
Untitled 2.odp - Libre...
*Untitled Document ...
Untitled 1 - LibreOffic...
Quantum GIS 1.7.5-W...
2:25
18:00

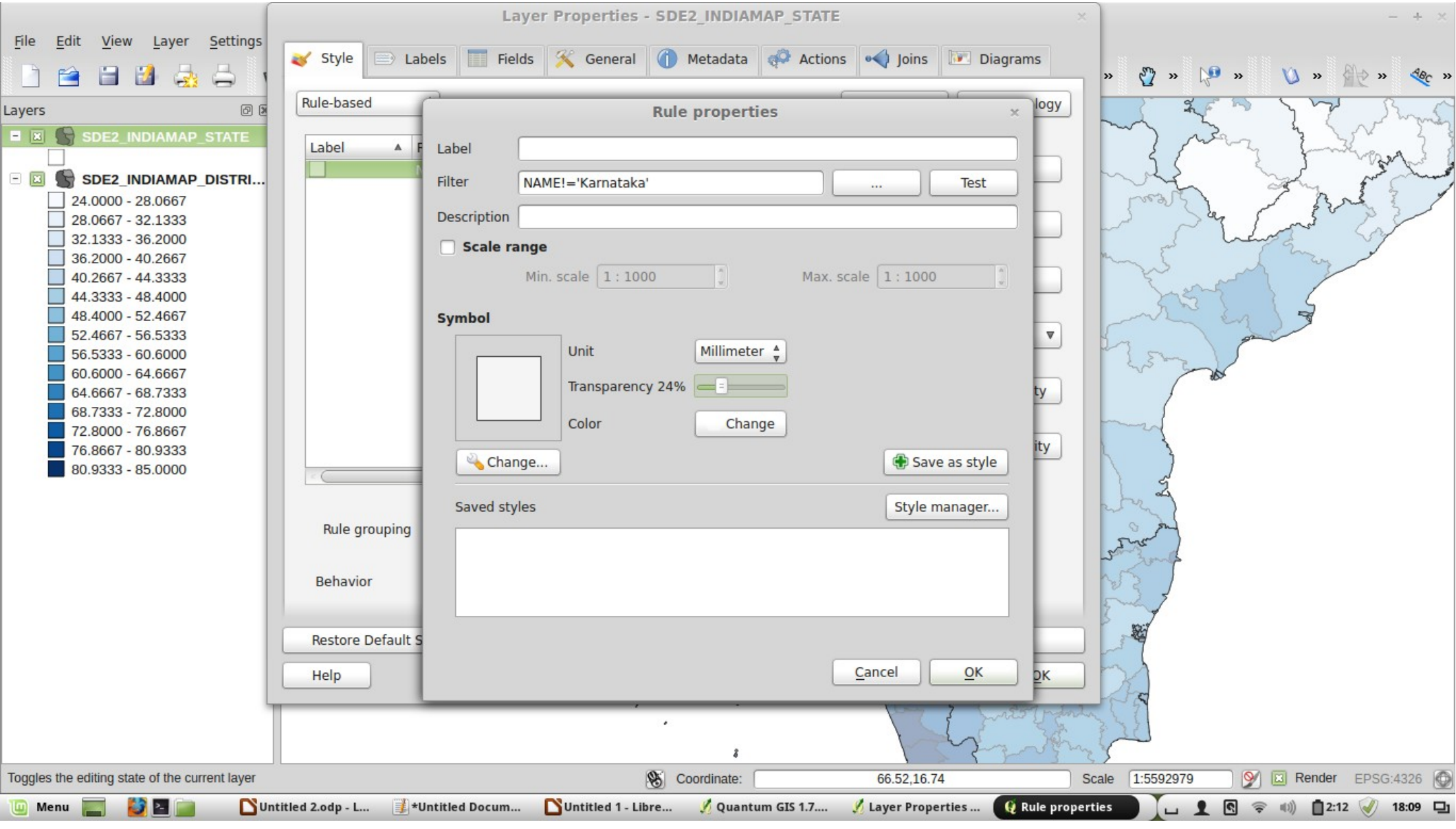
District gradation, state lines

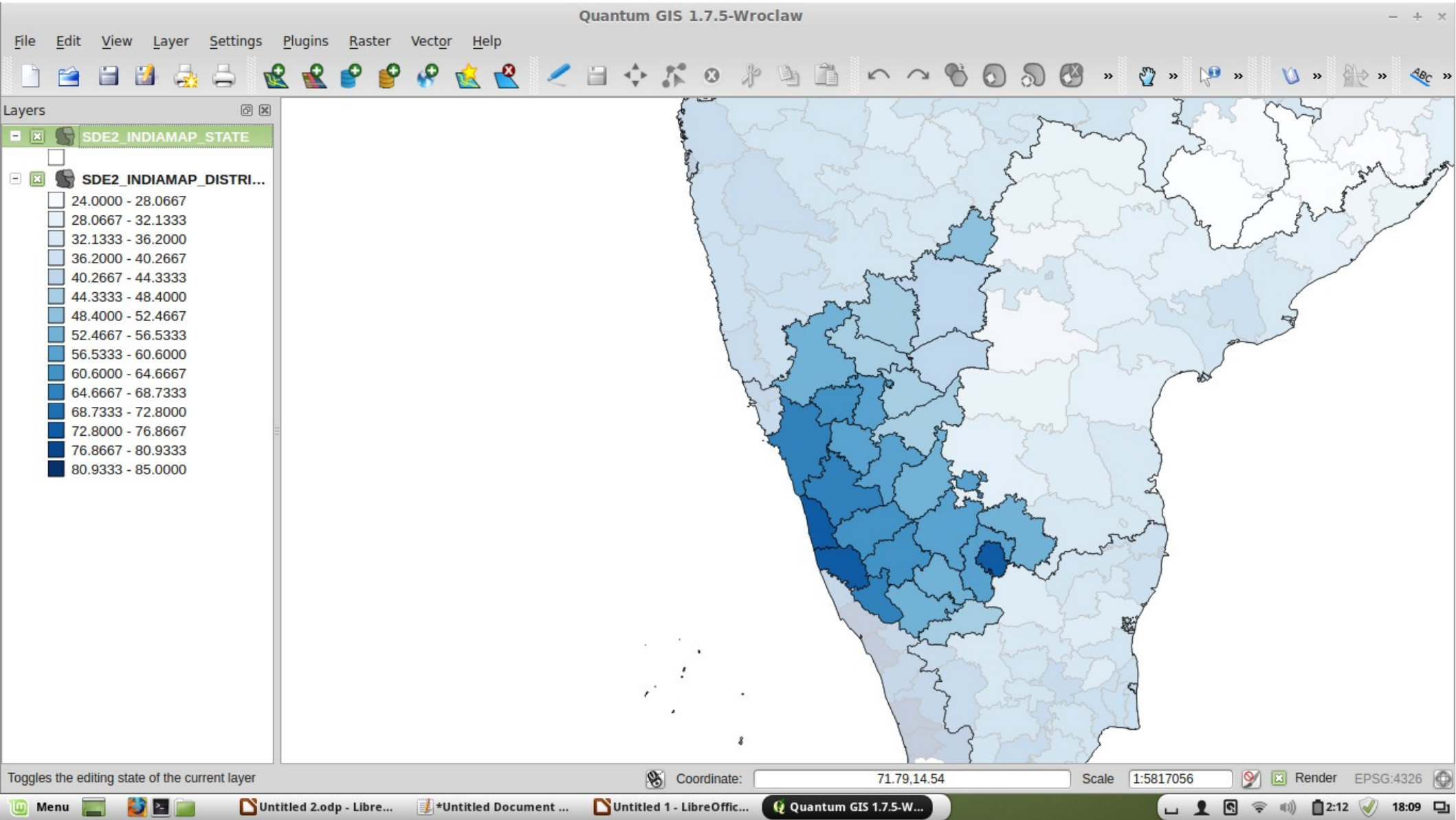


Multiple layers

- Note that the state level file also contains data, and we can alter its properties so as to manipulate the master image.
- For example, we might want to only show data from one state, and de-emphasise the others.
- To do this, we go to the properties window for the state-level file (on the left tab) and make the colouring 'rule-based'.
- We then make a rule to shade-out any entry that is not of a certain state.


Quantum GIS 1.7.5-Wroclaw
File Edit View Layer Settings Plugins Raster Vector Help
Layers
SDE2_INDIAMAP_STATE
SDE2_INDIAMAP_DISTRI...
24.0000 - 28.0667
28.0667 - 32.1333
32.1333 - 36.2000
36.2000 - 40.2667
40.2667 - 44.3333
44.3333 - 48.4000
48.4000 - 52.4667
52.4667 - 56.5333
56.5333 - 60.6000
60.6000 - 64.6667
64.6667 - 68.7333
68.7333 - 72.8000
72.8000 - 76.8667
76.8667 - 80.9333
80.9333 - 85.0000
Layer Properties - SDE2_INDIAMAP_STATE
Style Labels Fields General Metadata Actions Joins Diagrams
Rule-based
Symbol levels
Old symbology
Label Rule Min. scale Max. scale Priori
(no filter)
Add
Edit
Remove
Refine
Increase priority
Decrease priority
Rule grouping
None
By filter
By scale
Behavior
Enable symbol levels
Use only first matched rule
Restore Default Style
Save As Default
Load Style ...
Save Style ...
Help
Apply
Cancel
OK
Coordinate:
67.39,19.33
Scale
1:5170386
Render
EPSG:4326
Menu
Untitled 2.odp - Libre...
*Untitled Document ...
Untitled 1 - LibreOffic...
Quantum GIS 1.7.5-W...
Layer Properties - SD...
2:39
18:05


Layer Properties - SDE2_INDIAMAP_STATE
Style
Labels
Fields
General
Metadata
Actions
Joins
Diagrams
Rule-based
Rule properties
Label
Filter
NAME!='Karnataka'
Test
Description
Scale range
Min. scale
1 : 1000
Max. scale
1 : 1000
Symbol
Unit
Millimeter
Transparency 24%
Color
Change
Change...
Save as style
Saved styles
Style manager...
Rule grouping
Behavior
Restore Default S
Help
Cancel
OK
Layers
SDE2_INDIAMAP_STATE
SDE2_INDIAMAP_DISTRI...
24.0000 - 28.0667
28.0667 - 32.1333
32.1333 - 36.2000
36.2000 - 40.2667
40.2667 - 44.3333
44.3333 - 48.4000
48.4000 - 52.4667
52.4667 - 56.5333
56.5333 - 60.6000
60.6000 - 64.6667
64.6667 - 68.7333
68.7333 - 72.8000
72.8000 - 76.8667
76.8667 - 80.9333
80.9333 - 85.0000
Toggles the editing state of the current layer
Coordinate:
66.52,16.74
Scale
1:5592979
Render
EPSG:4326
Menu
Untitled 2.odp - L...
*Untitled Docum...
Untitled 1 - Libre...
Quantum GIS 1.7....
Layer Properties ...
Rule properties
2:12
18:09


Quantum GIS 1.7.5-Wroclaw
File Edit View Layer Settings Plugins Raster Vector Help
Layers
SDE2_INDIAMAP_STATE
SDE2_INDIAMAP_DISTRI...
24.0000 - 28.0667
28.0667 - 32.1333
32.1333 - 36.2000
36.2000 - 40.2667
40.2667 - 44.3333
44.3333 - 48.4000
48.4000 - 52.4667
52.4667 - 56.5333
56.5333 - 60.6000
60.6000 - 64.6667
64.6667 - 68.7333
68.7333 - 72.8000
72.8000 - 76.8667
76.8667 - 80.9333
80.9333 - 85.0000
Toggles the editing state of the current layer
Coordinate: 71.79,14.54
Scale 1:5817056
Render
EPSG:4326
Menu
Untitled 2.odp - Libre...
*Untitled Document ...
Untitled 1 - LibreOffic...
Quantum GIS 1.7.5-W...
2:12
18:09

Joining your own data

- Shapefiles typically have data already embedded. Sometimes this data is quite useful.

- The India file for example has lots of data already from the 2011 census.

- But this is quite limiting. what if we want to add our own data? I have some data on historical taxes across India that I have collected, and would like to add.

Joining your own data

- The method of inserting data into a shapefile is called 'joining'.

- Because we can have multiple layers in our workspace, we simply import our (csv) data table as a new layer, and then 'join' the data table to the existing shapefile.

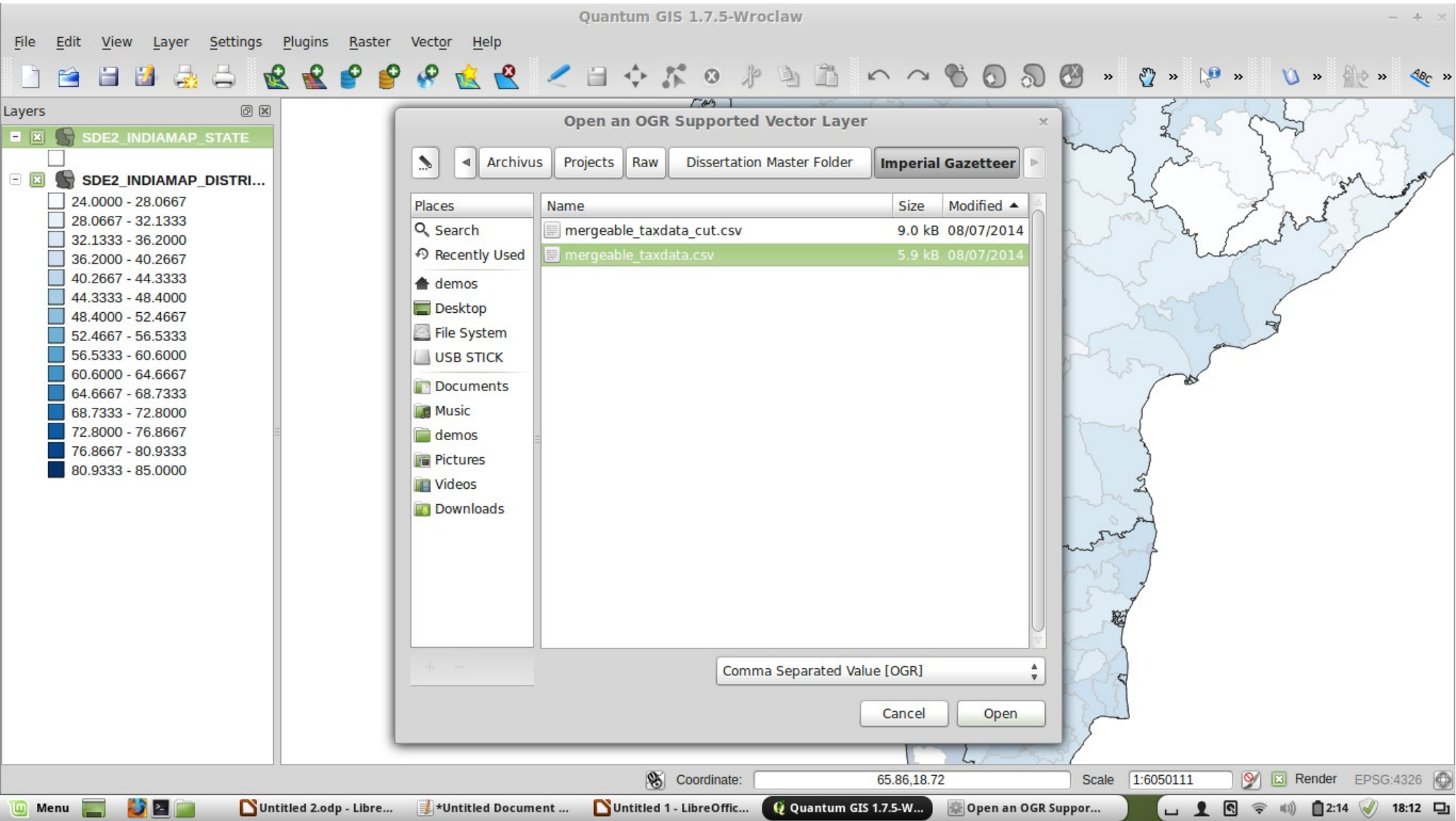
- They are matched on the basis of a shared variable – almost always the regional or country name/id.

Joining your own data

- Because QGIS demands use of the .csv format, first we need to get the data in a spreadsheet file (your statistics package may allow automatic export; otherwise you will have to copy and paste).

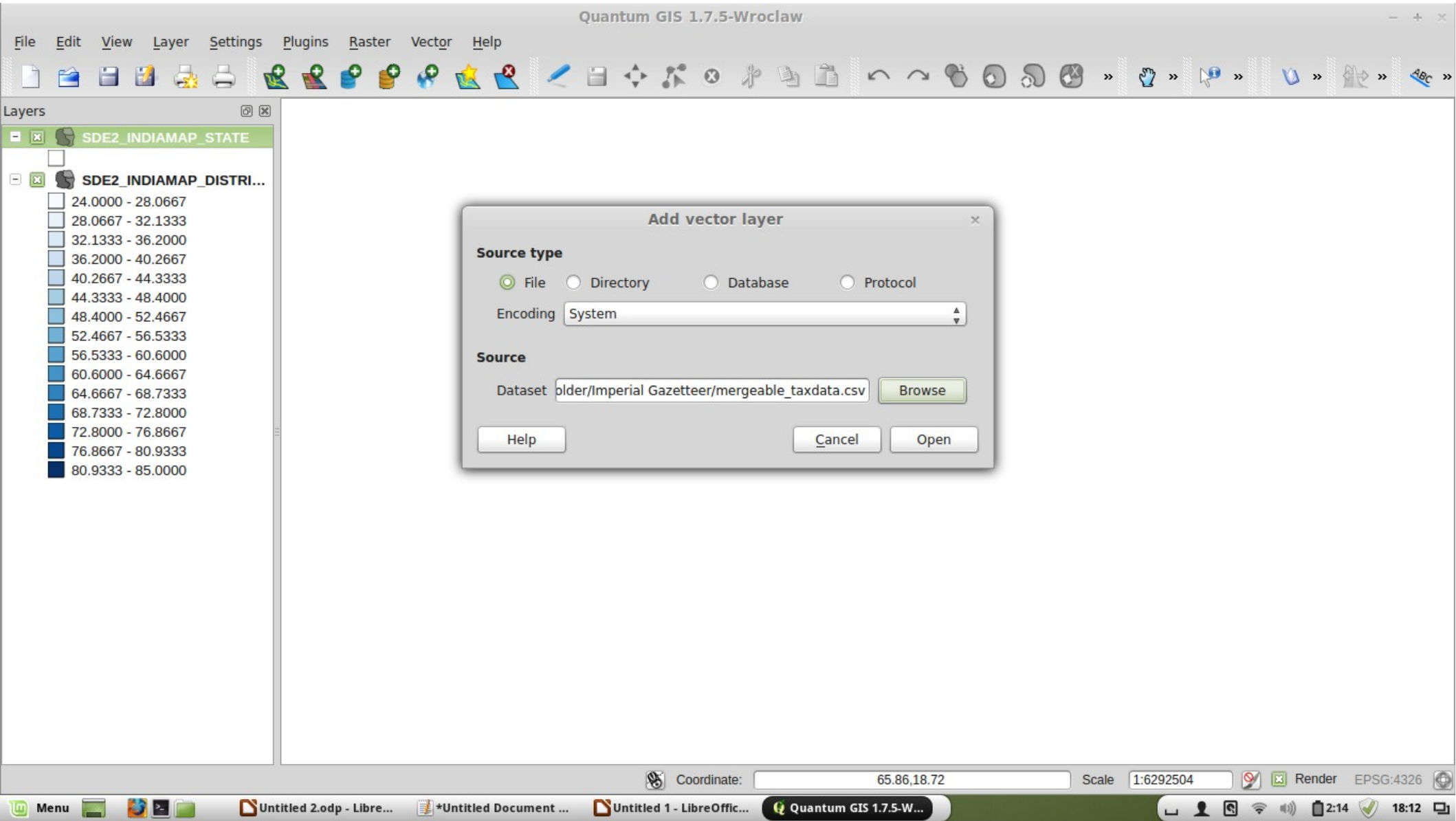
- **Make sure that the match variable does, in fact, match**. It can be either a string or a numerical id, but obviously any case than cannot be found in the corresponding dataset will simply be dropped.

Joining the data

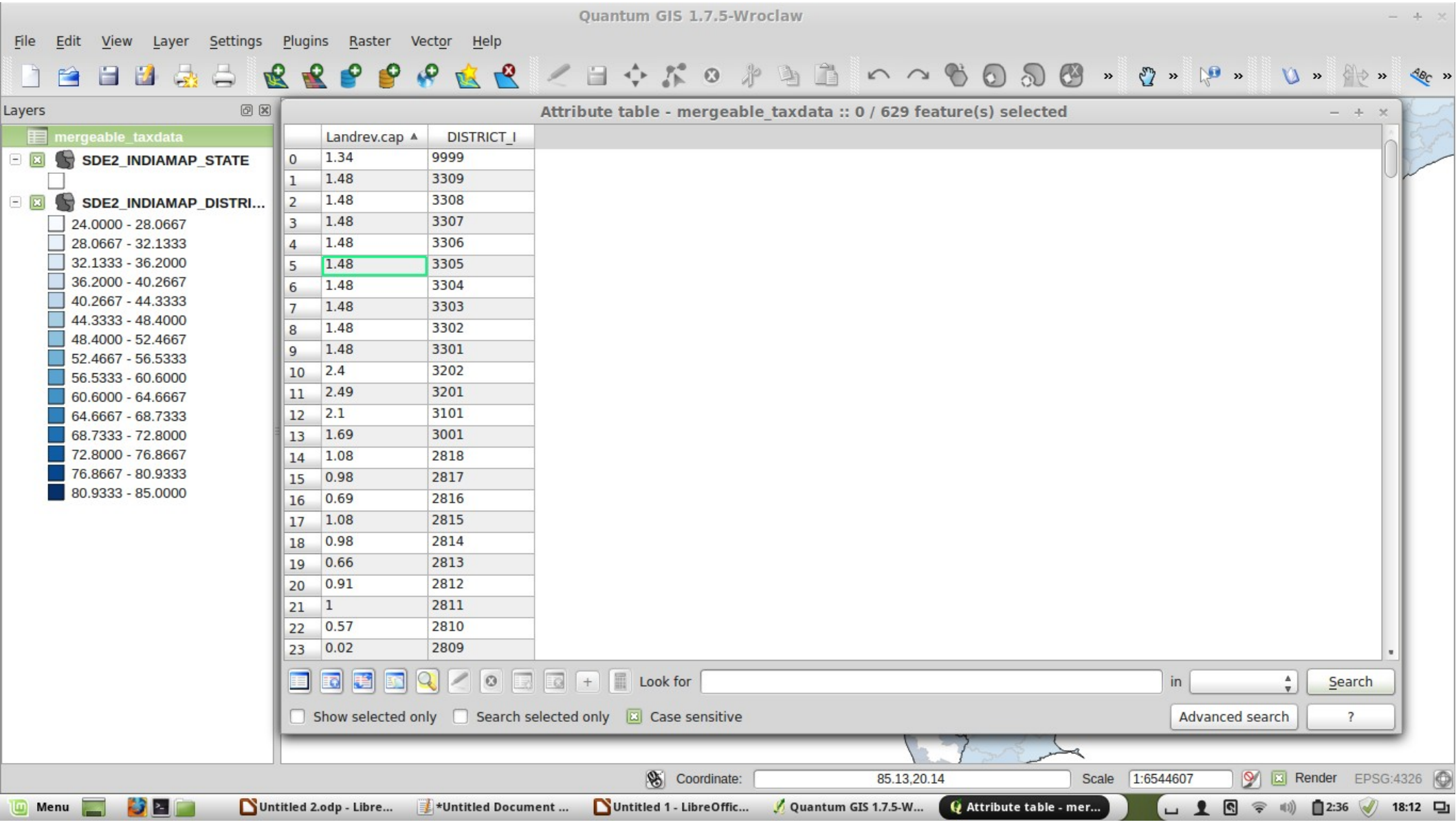


Joining your own data

- We just select the same option for adding a 'data layer' as we would for adding a new shape file.
- The data table will then appear alongside the map shapefiles in the right panel.
- QGIS basically treats spreadsheet as data tables that are not (yet) linked to any map coordinates or objects.

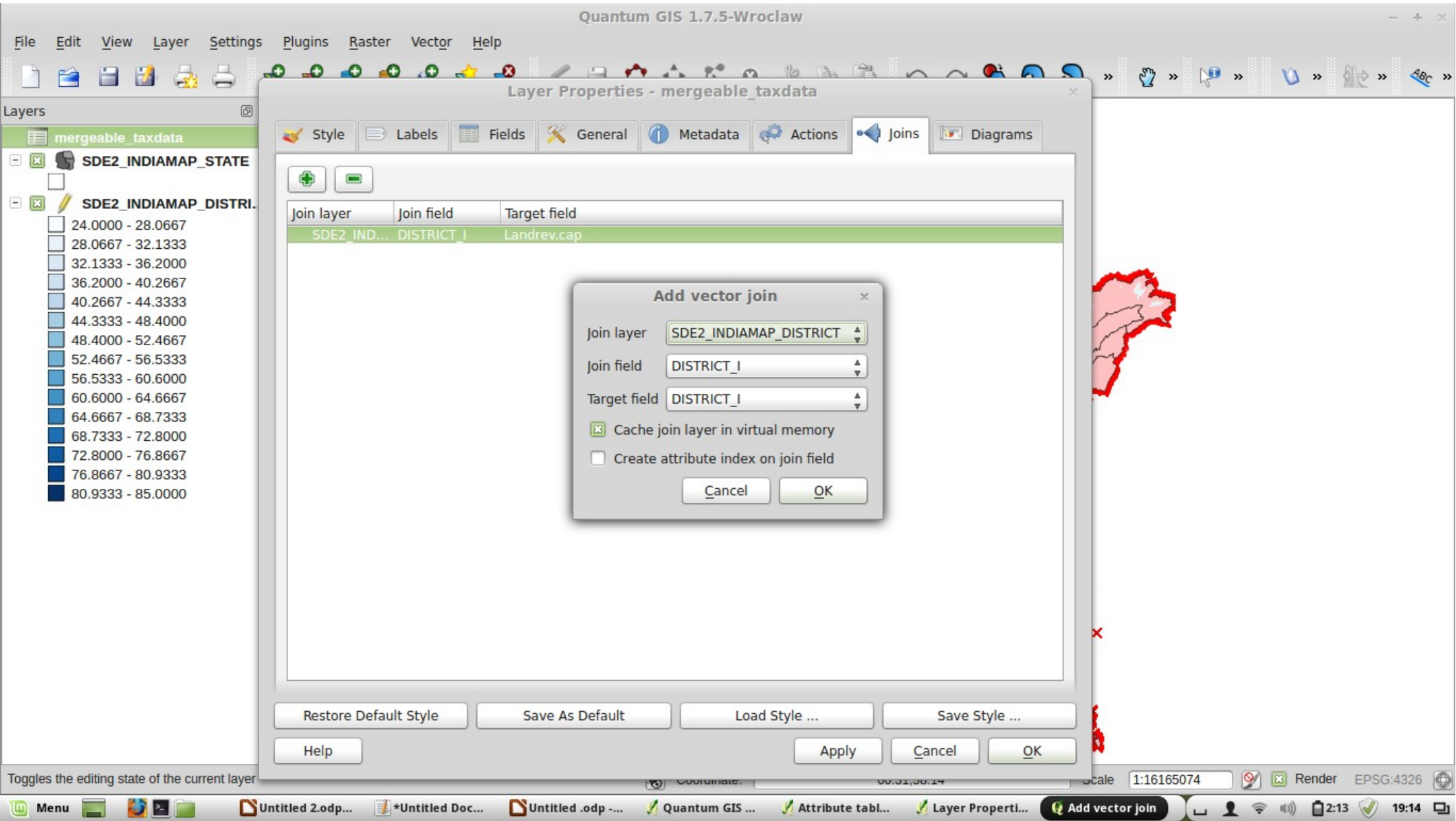

Quantum GIS 1.7.5-Wroclaw
File Edit View Layer Settings Plugins Raster Vector Help
Layers
SDE2_INDIAMAP_STATE
SDE2_INDIAMAP_DISTRI...
24.0000 - 28.0667
28.0667 - 32.1333
32.1333 - 36.2000
36.2000 - 40.2667
40.2667 - 44.3333
44.3333 - 48.4000
48.4000 - 52.4667
52.4667 - 56.5333
56.5333 - 60.6000
60.6000 - 64.6667
64.6667 - 68.7333
68.7333 - 72.8000
72.8000 - 76.8667
76.8667 - 80.9333
80.9333 - 85.0000
Add vector layer
Source type
File Directory Database Protocol
Encoding System
Source
Dataset older/Imperial Gazetteer/mergeable_taxdata.csv
Browse
Help Cancel Open
Coordinate: 65.86,18.72 Scale 1:6292504 Render EPSG:4326
Menu
Untitled 2.odp - Libre...
*Untitled Document ...
Untitled 1 - LibreOffic...
Quantum GIS 1.7.5-W...
2:14
18:12

And here is the data

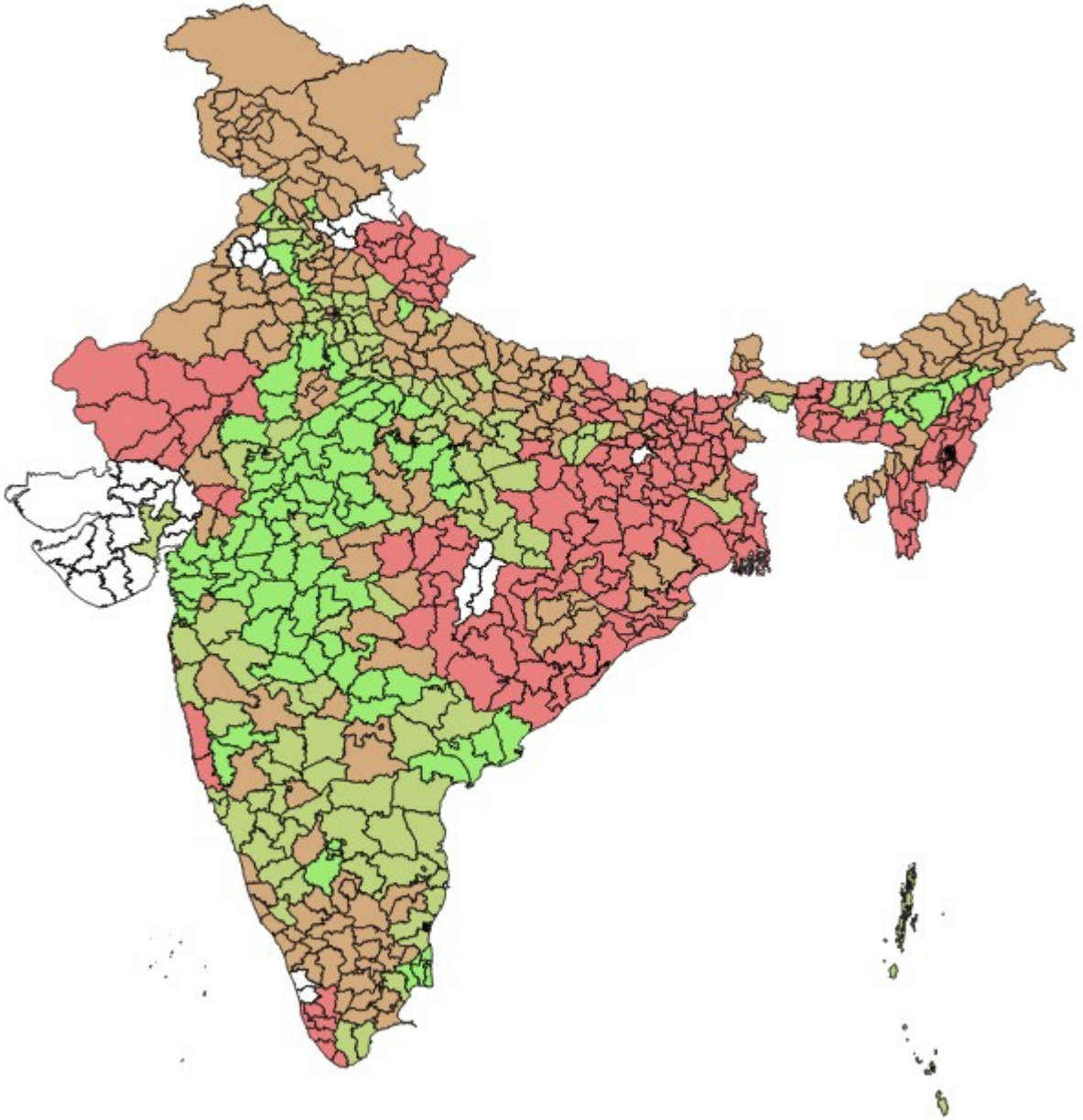


	Landrev.cap	DISTRICT_I
0	1.34	9999
1	1.48	3309
2	1.48	3308
3	1.48	3307
4	1.48	3306
5	1.48	3305
6	1.48	3304
7	1.48	3303
8	1.48	3302
9	1.48	3301
10	2.4	3202
11	2.49	3201
12	2.1	3101
13	1.69	3001
14	1.08	2818
15	0.98	2817
16	0.69	2816
17	1.08	2815
18	0.98	2814
19	0.66	2813
20	0.91	2812
21	1	2811
22	0.57	2810
23	0.02	2809

Joining the data to an existing file



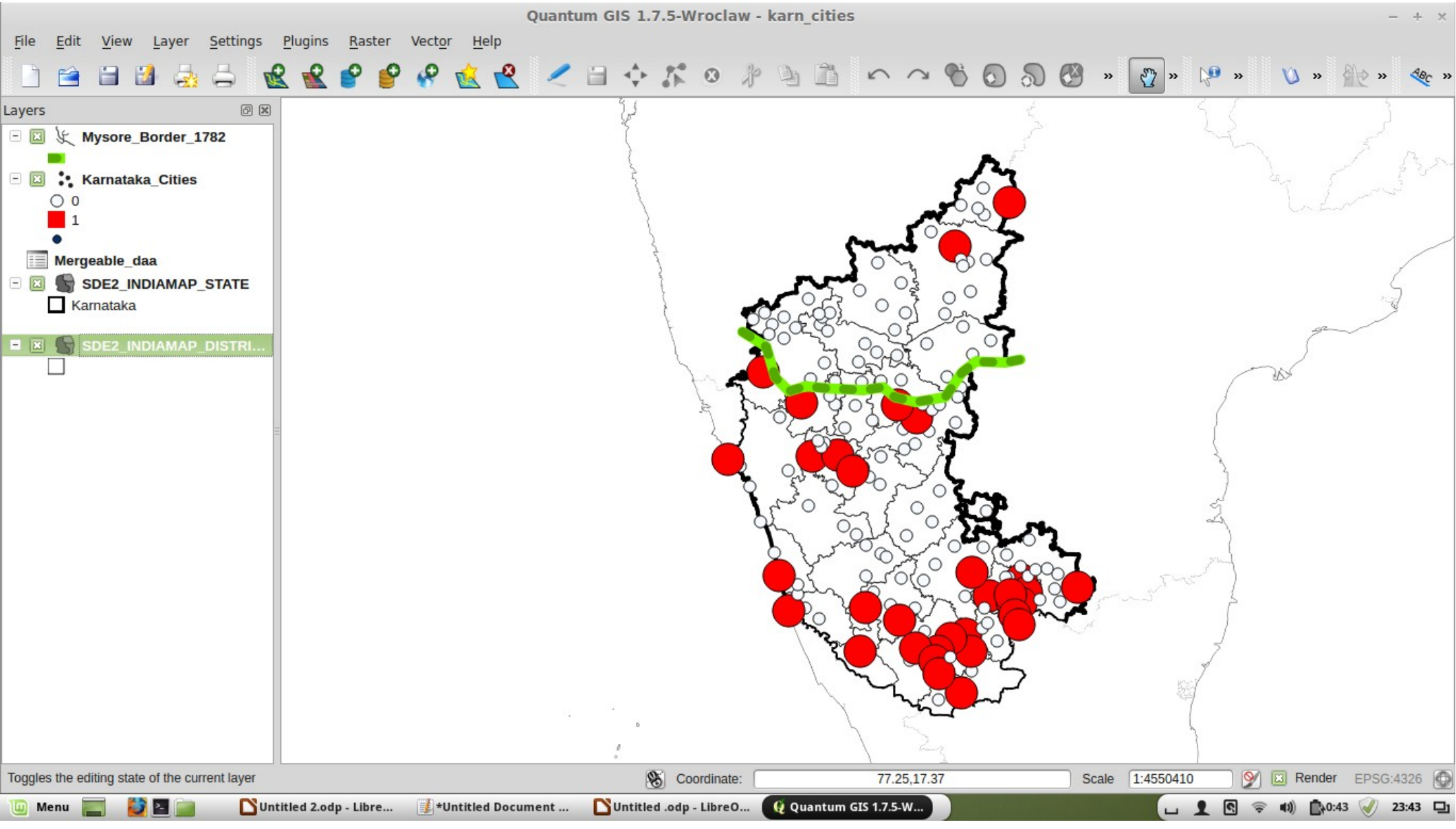
Mapping the new data



Other stuff

- We aren't restricted to adding shapefiles and gradiating our maps with new data.
- We could add a wide range of other things to a map, including (for example) contour lines, labels, features (lines), nodes (e.g. to represent cities, airports, etc)
- We might be able to import many of these. But we can also draw them ourselves.

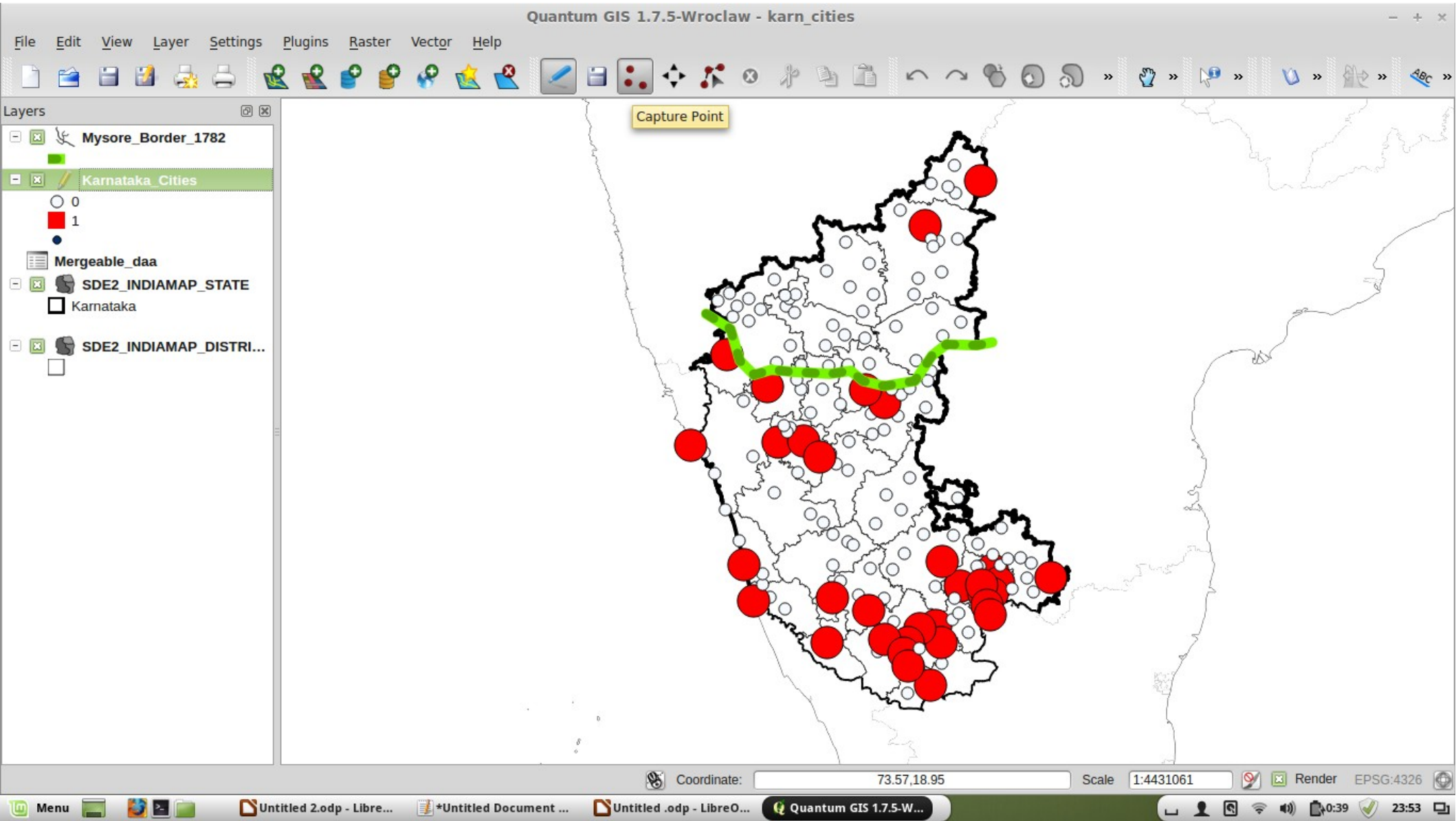
Adding nodes to a map



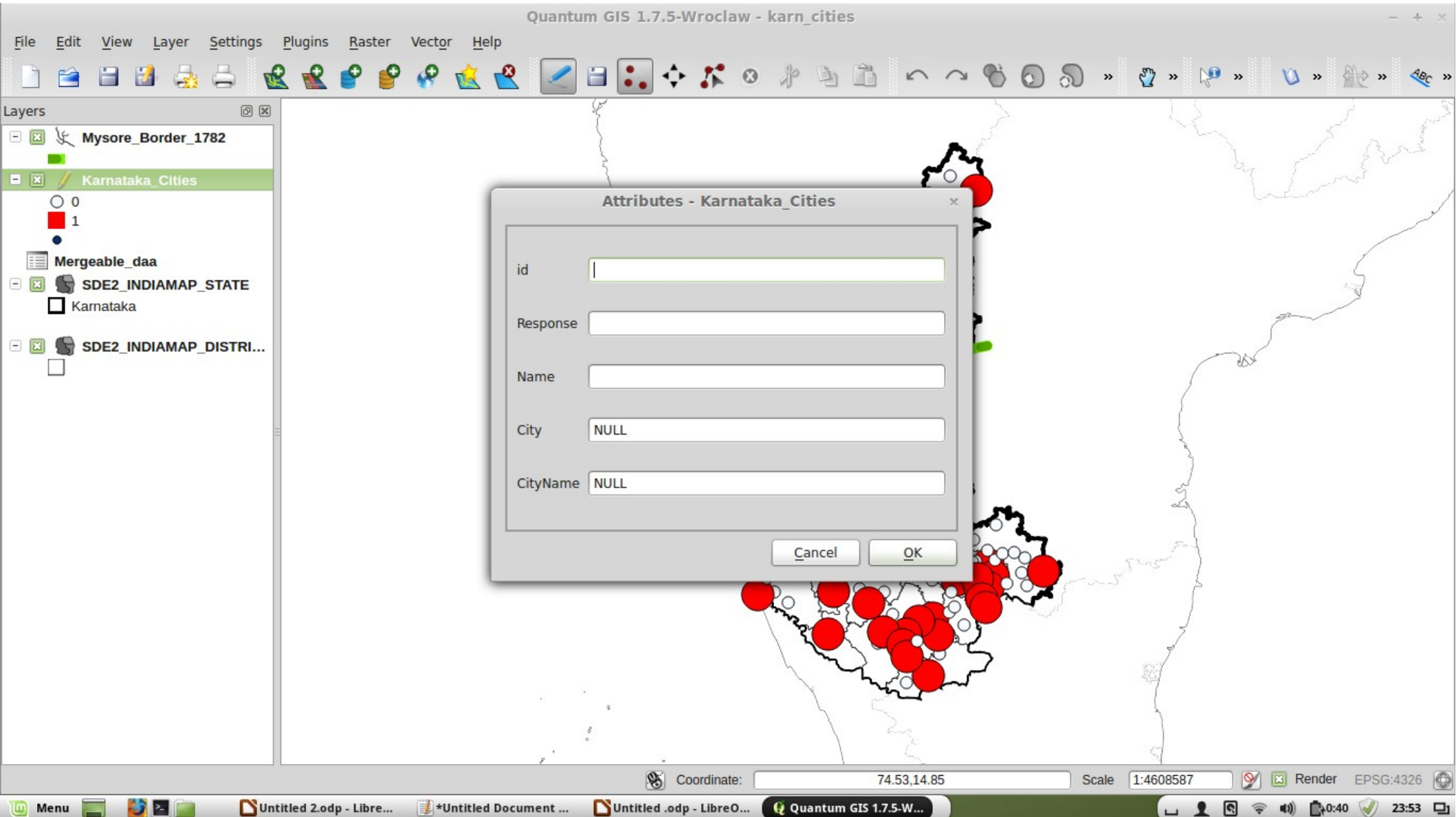
Other stuff

- Here there are two additional features – a set of nodes that represent major cities of Karnataka, and a line that represents the northernmost border of the Mysore Empire in 1782.
- Note that these are lain over the state and the district shapefiles (so there are here 4 layers of data)
- The nodes were simply drawn manually based on a graphic overlay.
- Easy to add / remove nodes and manipulate their appearance e.g. :

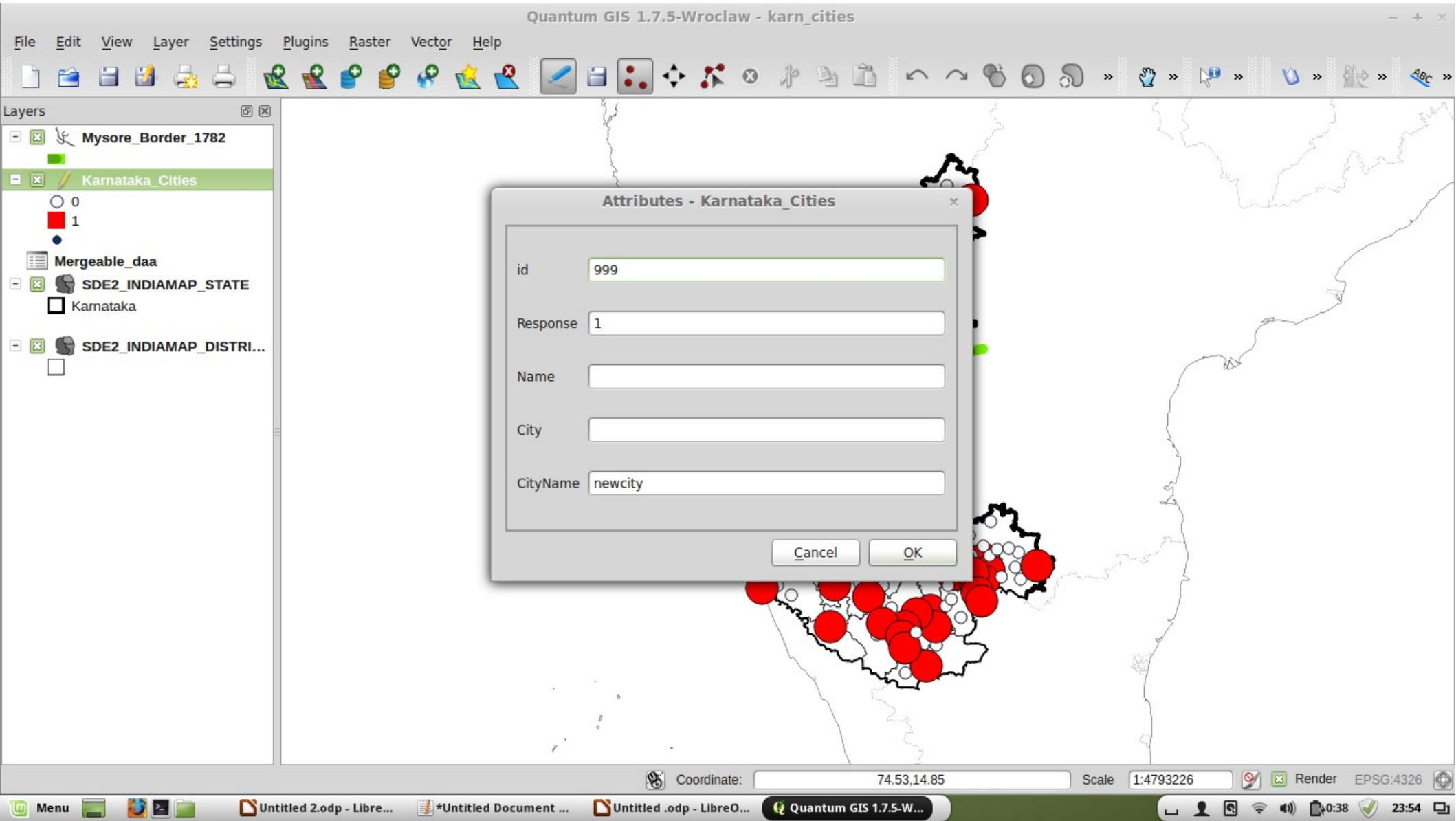
Capture a new point...



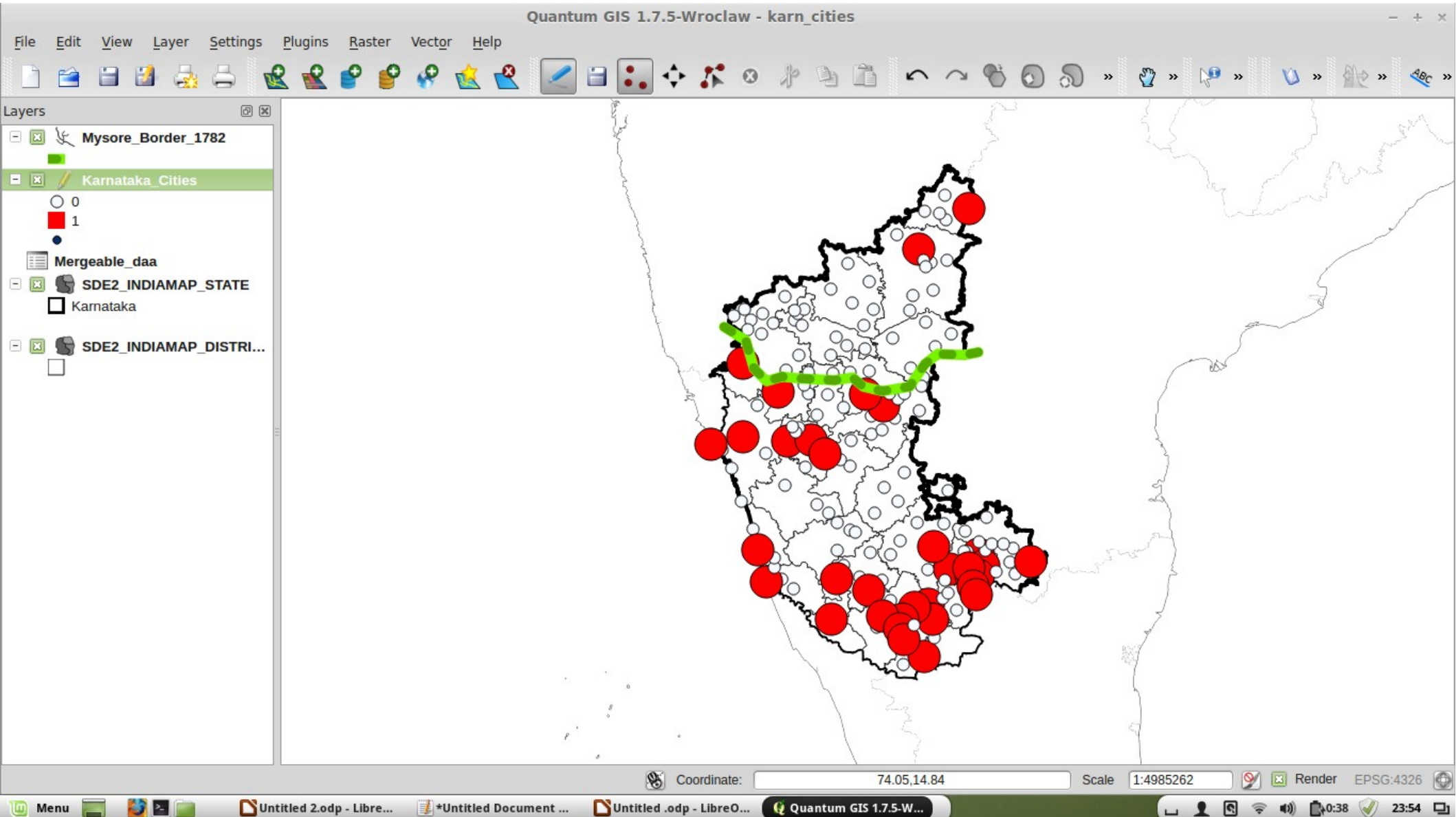
Give it some attributes



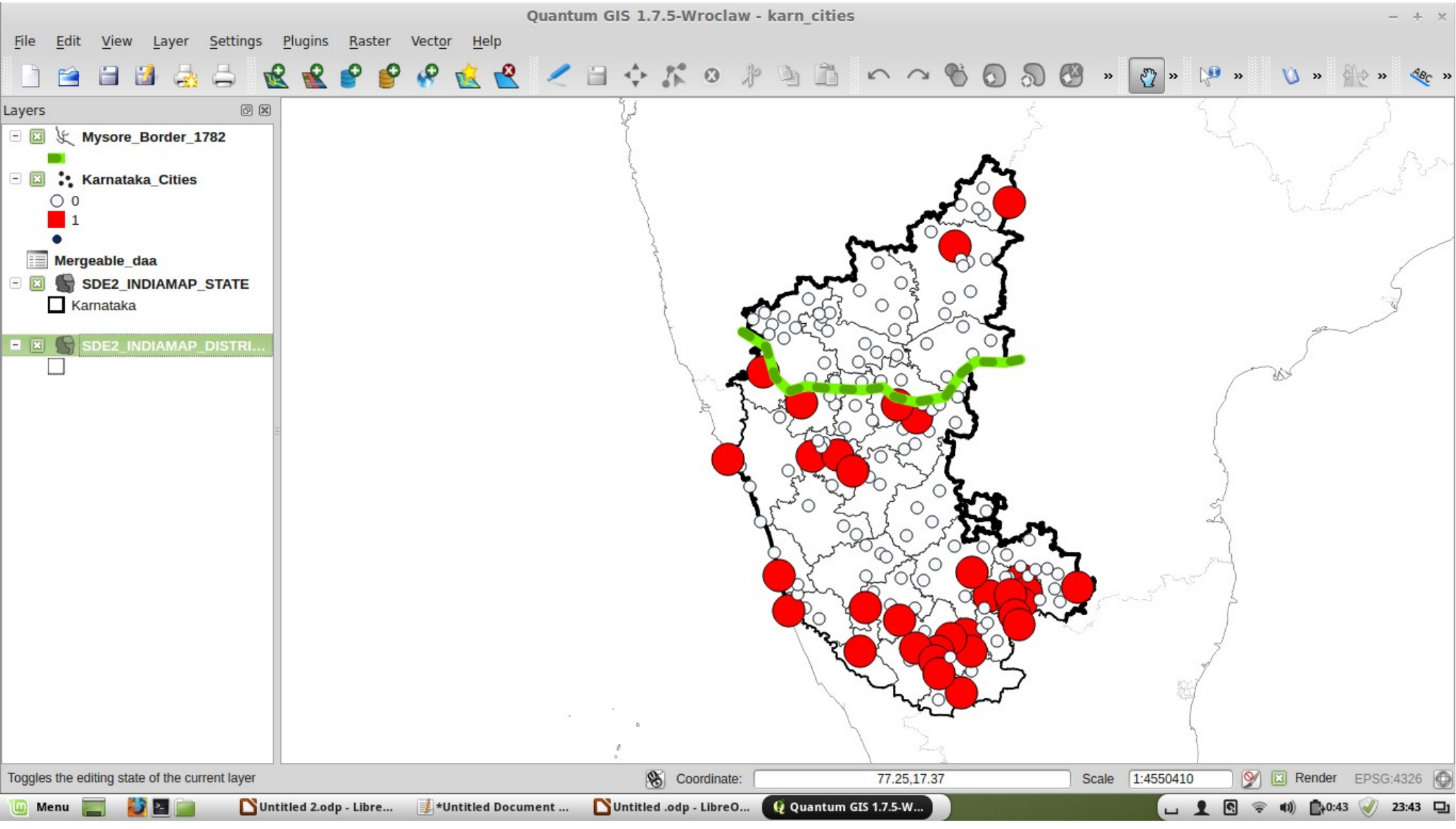
Give it some attributes



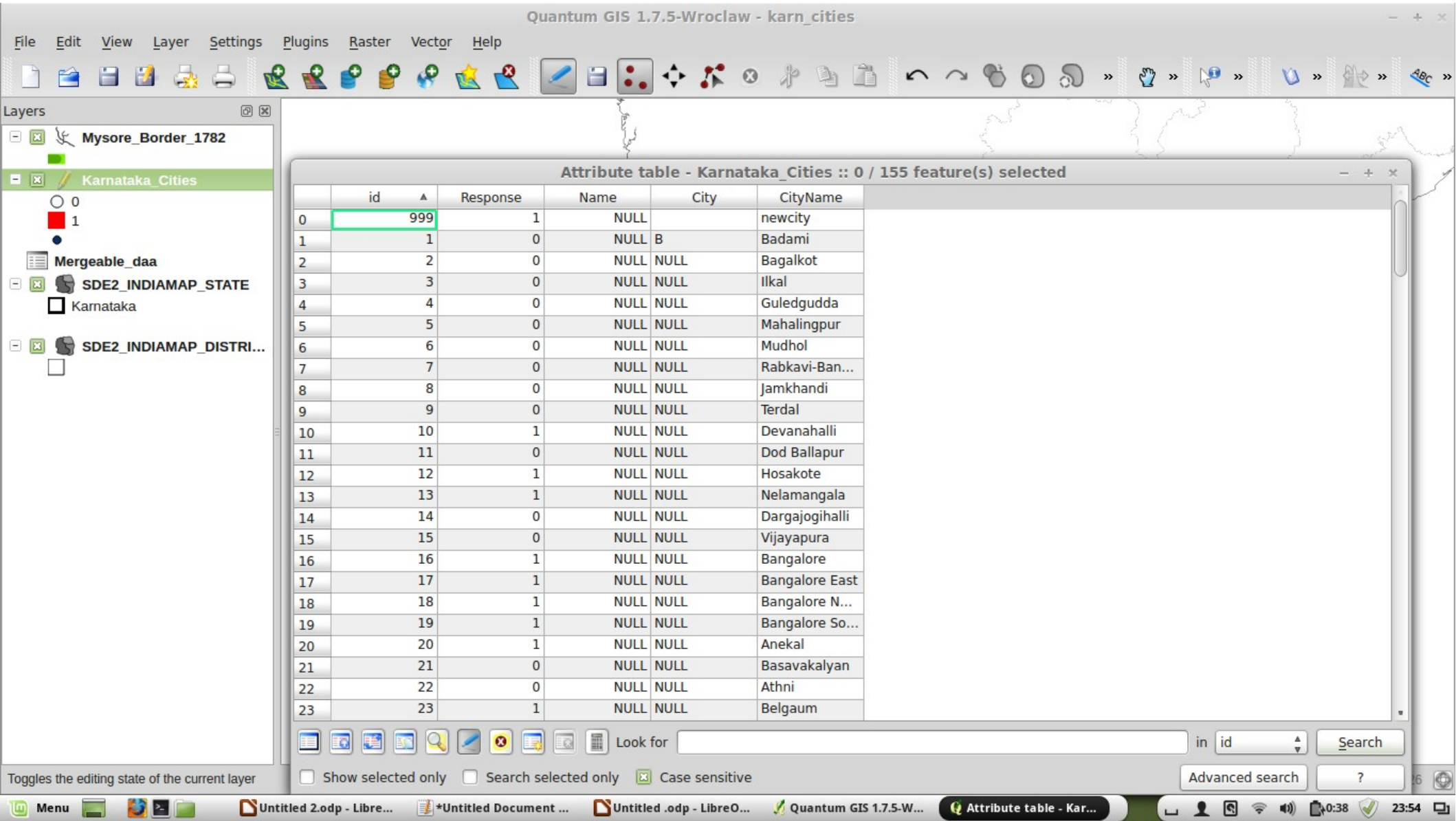
with new node



(before)


Quantum GIS 1.7.5-Wroclaw - karn_cities
File Edit View Layer Settings Plugins Raster Vector Help
Layers
Mysore_Border_1782
Karnataka_Cities
0
1
Mergeable_daa
SDE2_INDIAMAP_STATE
Karnataka
SDE2_INDIAMAP_DISTRI...
Toggles the editing state of the current layer
Coordinate:
77.25,17.37
Scale
1:4550410
Render
EPSG:4326
Menu
Untitled 2.odp - Libre...
*Untitled Document ...
Untitled .odp - LibreO...
Quantum GIS 1.7.5-W...
0:43
23:43

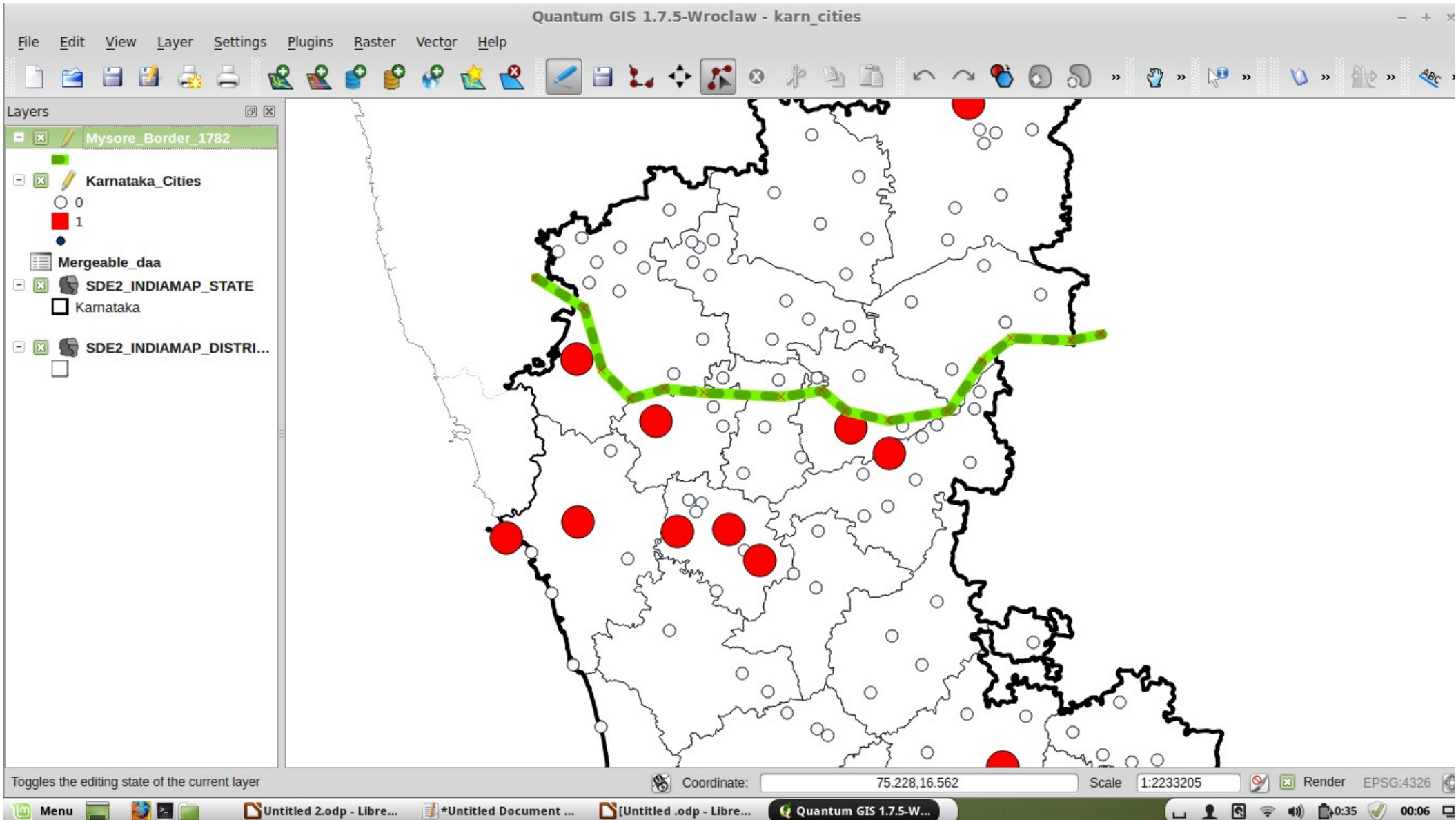
the node attributes



Customising nodes

- As with shapefile layers, we can fully customise the colour scheme. I've opted for a simple binary rule-based scheme whereby some of the nodes get coded as big red dots, and the others as small white ones, depending on the variable 'response'.
- We could do something more sophisticated. For example, having a variable for population and then sizing or colouring the node by the size of the city.
- Even better things are possible – e.g. having a georeference for each individual survey respondent, and then drawing 'heatmaps' based on answers to survey items.

The border


Quantum GIS 1.7.5-Wroclaw - karn_cities
File
Edit
View
Layer
Settings
Plugins
Raster
Vector
Help
Layers
Mysore_Border_1782
Karnataka_Cities
0
1
Mergeable_daa
SDE2_INDIAMAP_STATE
Karnataka
SDE2_INDIAMAP_DISTRI...
Toggles the editing state of the current layer
Coordinate:
75.228,16.562
Scale
1:2233205
Render
EPSG:4326
Menu
Untitled 2.odp - Libre...
*Untitled Document ...
[Untitled .odp - Libre...
Quantum GIS 1.7.5-W...
0:35
00:06

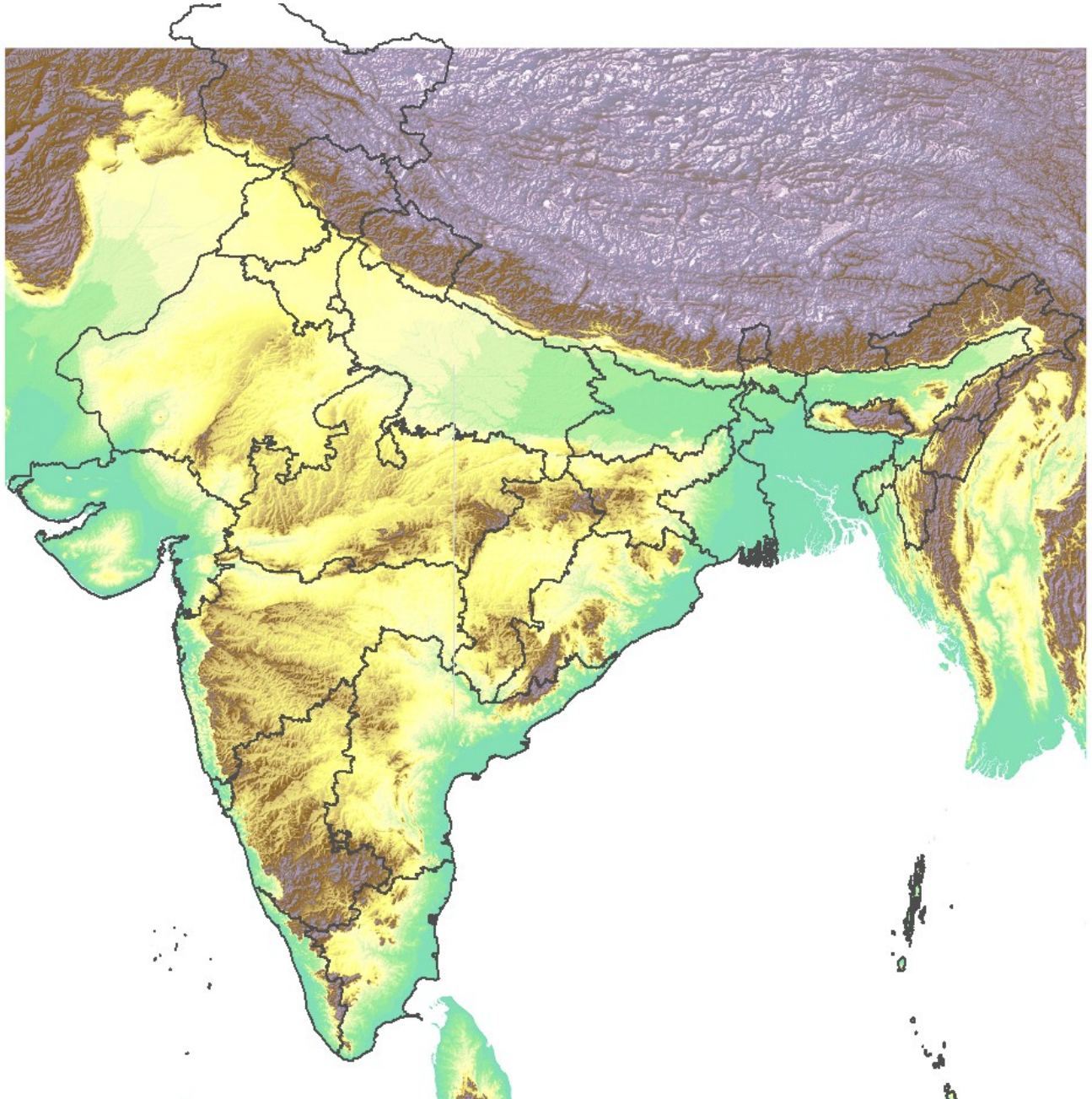
The border

- The border is just another layer, consisting a series of arbitrary (user-defined) nodes that are connected by a (slightly ugly) green line (following the line of a 1907 textbook map – Joppen's *Atlas of India*).
- As such, they are very easily manipulated – QGIS allows us to drag the points etc. similar to any graphics package would do.

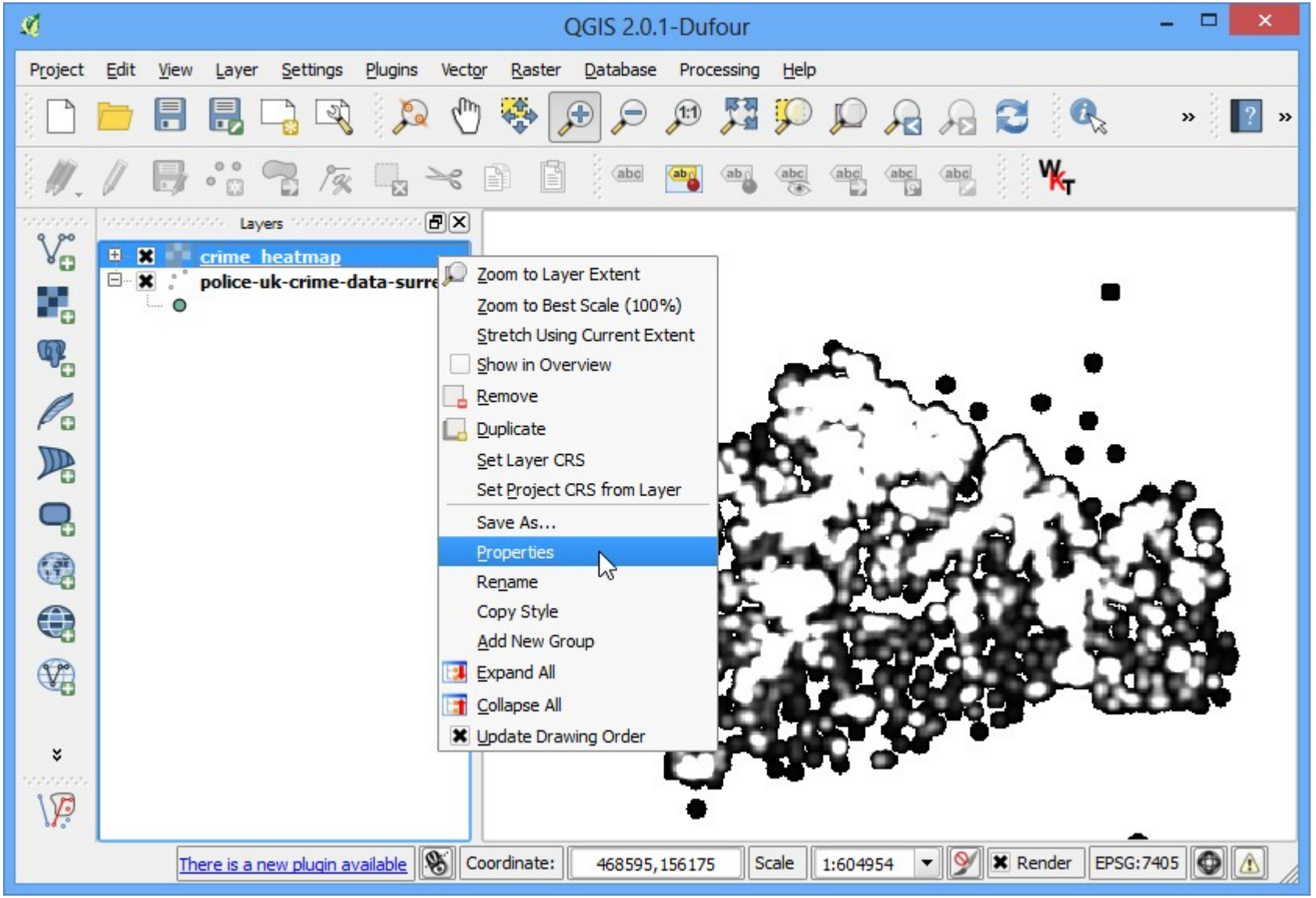
Other data

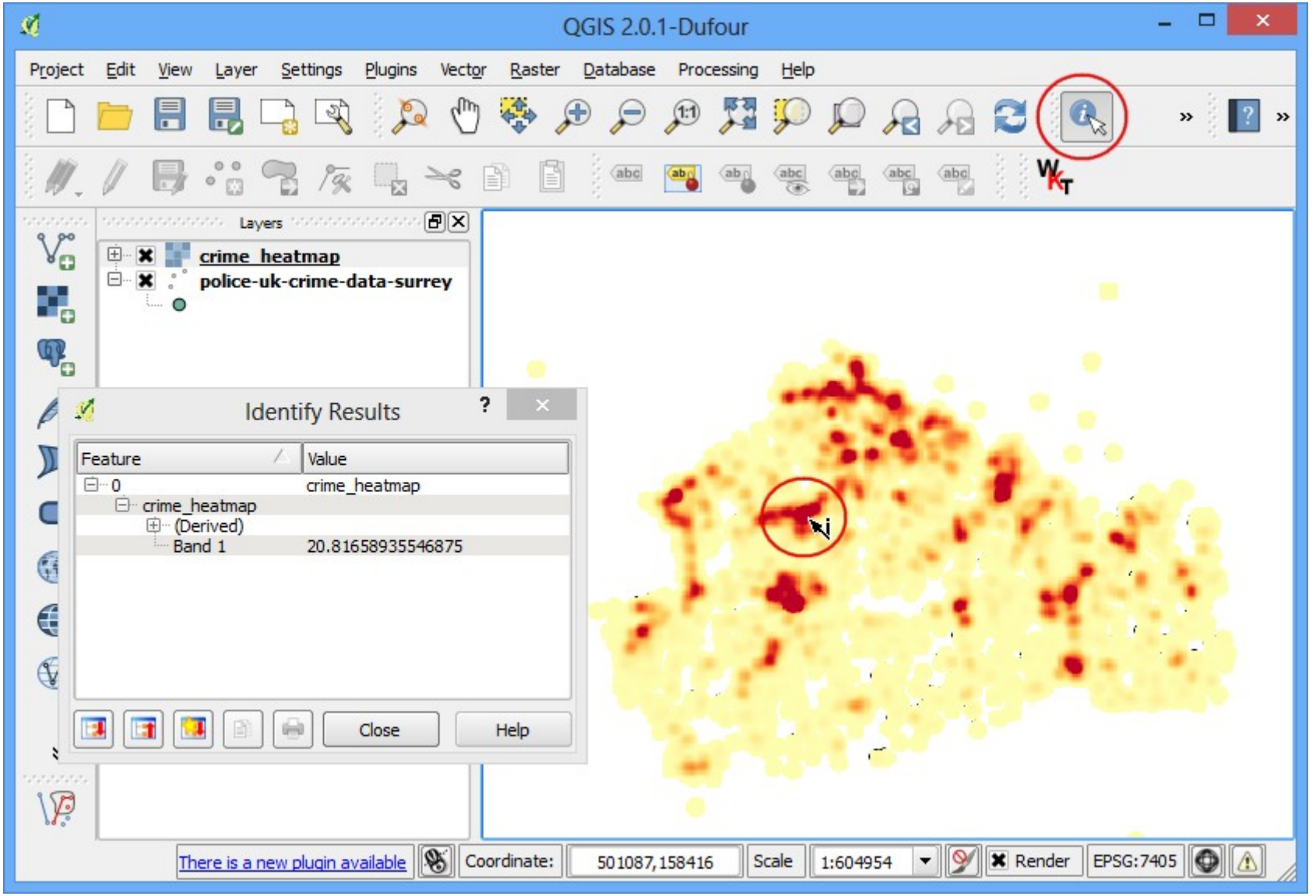
- QGIS also allows for the overlay of other kinds of data, for example contour maps.
- Also, we are not restricted to dealing with geographic maps: we could for example have neighbourhood maps, using data from openstreetmap.org, with nodes to show the location of survey respondents, or interviewees, or other subjects of interest.

E.g. overlaying a contour map layer

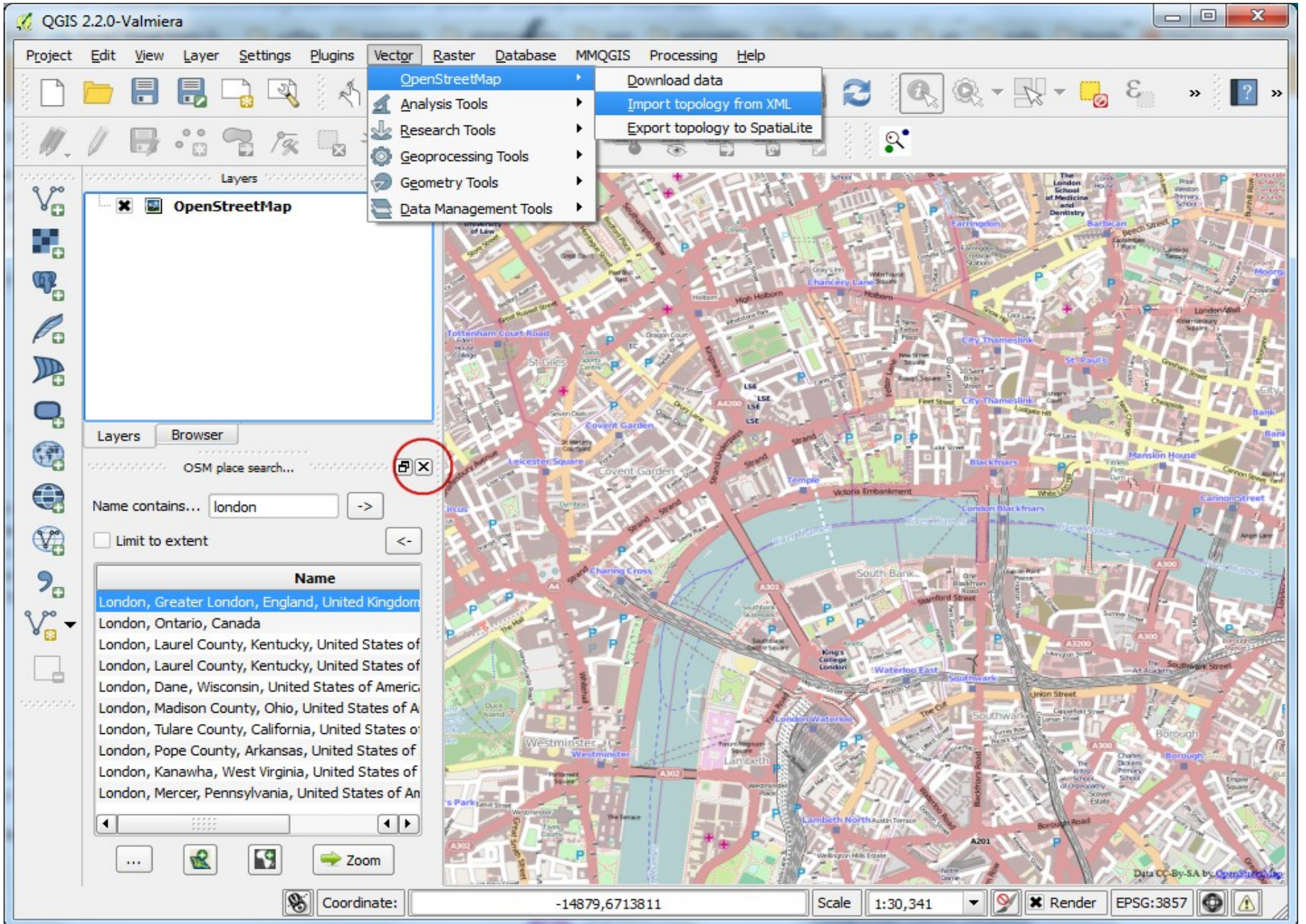


Example: what heatmaps look like

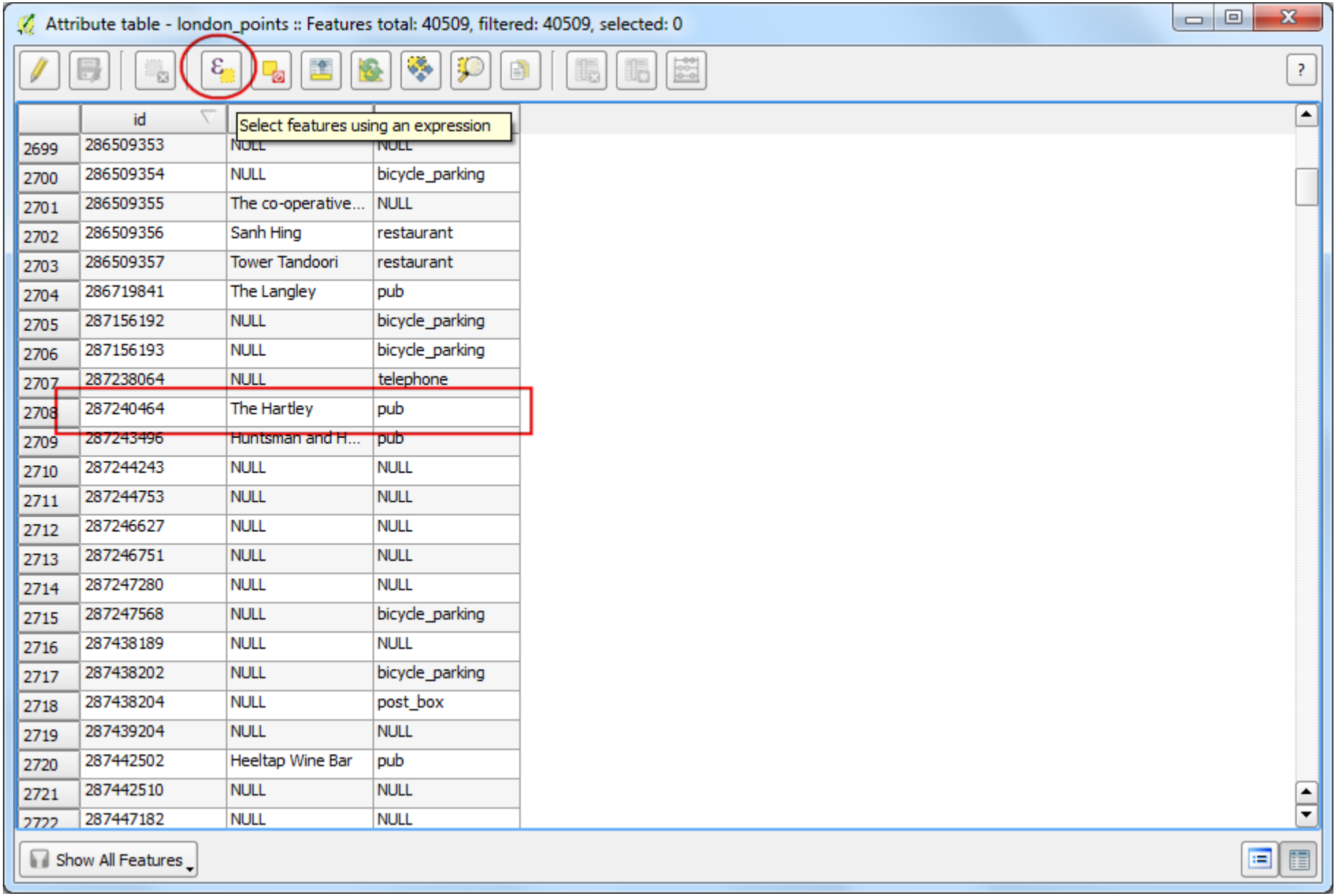



QGIS 2.0.1-Dufour
Project
Edit
View
Layer
Settings
Plugins
Vector
Raster
Database
Processing
Help
Layers
crime_heatmap
police-uk-crime-data-surrey
Identify Results
Feature
Value
0
crime_heatmap
crime_heatmap
(Derived)
Band 1
20.81658935546875
Close
Help
There is a new plugin available
Coordinate:
501087,158416
Scale
1:604954
Render
EPSG:7405

Or: using streetmap data

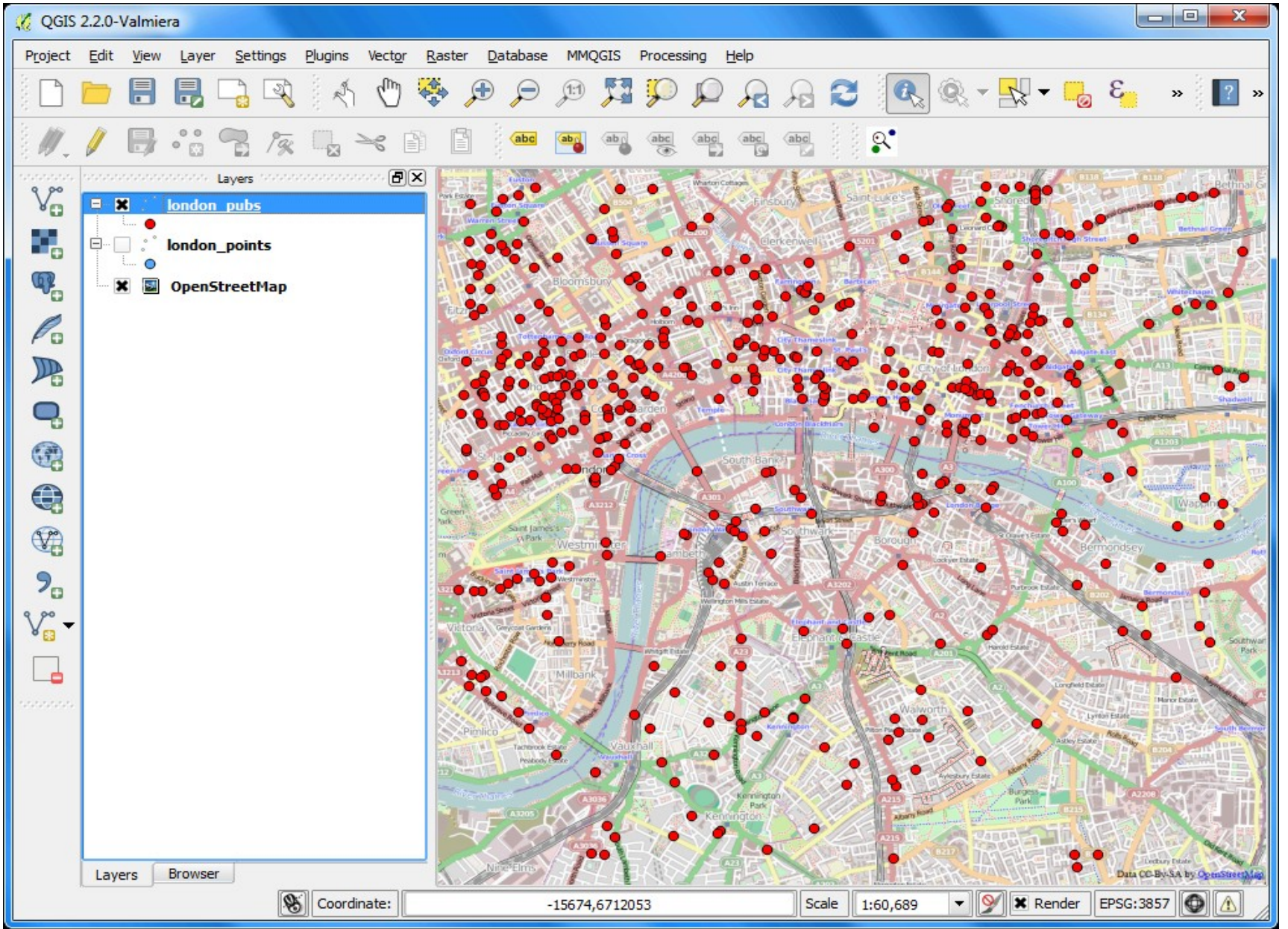


Underlying data – London landmarks



	id		
2699	286509353	NULL	NULL
2700	286509354	NULL	bicycle_parking
2701	286509355	The co-operative...	NULL
2702	286509356	Sanh Hing	restaurant
2703	286509357	Tower Tandoori	restaurant
2704	286719841	The Langley	pub
2705	287156192	NULL	bicycle_parking
2706	287156193	NULL	bicycle_parking
2707	287238064	NULL	telephone
2708	287240464	The Hartley	pub
2709	287243496	Huntsman and H...	pub
2710	287244243	NULL	NULL
2711	287244753	NULL	NULL
2712	287246627	NULL	NULL
2713	287246751	NULL	NULL
2714	287247280	NULL	NULL
2715	287247568	NULL	bicycle_parking
2716	287438189	NULL	NULL
2717	287438202	NULL	bicycle_parking
2718	287438204	NULL	post_box
2719	287439204	NULL	NULL
2720	287442502	Heeltap Wine Bar	pub
2721	287442510	NULL	NULL
2722	287447182	NULL	NULL

Highlighting the pubs in London



Summing up

- Geospatial mapping is a flexible way of visualising data in a visually appealing and intuitive way.
- It is easy to learn how to do.
- Is also becoming a new benchmark for the presentation of data: enables rapid analysis and descriptive statistics, and is aesthetically appealing.

This report was presented at the training methodological workshop
"Economic and Social Changes: values effects across Eurasia".

March 31 - April 6, 2015 – Turkey.

http://lcsr.hse.ru/en/seminar_m2015

Настоящий доклад был представлен на методологическом учебном семинаре
«Экономические и социальные изменения: оценка эффектов по всей Евразии».

31 марта – 6 апреля 2015 года – Турция.

http://lcsr.hse.ru/seminar_m2015