Spatial Data Visualization and Analytics

Learn techniques for working with all types of spatial datasets.

Ujaval Gandhi

Contents

This course is also offered as an in-person class. If you would like to attend one of my workshops, visit www.spatialthoughts.com/events to know details of upcoming classes. Please sign up for my mailing list to know when new sessions are scheduled.

Introduction

Spatial Data Model

geometry + properties

shape + table

geometry = coordinates + reference properties = type + data

A GeoJSON representation

```
{
     "type": "Feature",
     "geometry": {
       "type": "Point",
       "coordinates": [ 77.58270263671875, 12.963074139604124]
     },
     "properties": {
       "name": "Bengaluru"
     }
 }
```

Spatial Data Formats

Type	Non-Spatial Data	Spatial Data
Text	csv, json, xml	csv, geojson, gml, kml
Binary/Compressed	xlsx, zip	shapefile, geopackage
Images	tiff, jpg, png	geotiff, jpeg2000
Databases	SQLite, PostgreSQL, Oracle	Spatialite, PostGIS, Oracle Spatial

Spatial Data Types

Type	Sub Types
Vector	Point, Line (Curve), Polygon (Surface) MultiPoint, MultiLine, MultiPolyogn PointZ, LineZ, LineM, Multipatch Point Cloud
Raster	Grids Mesh Tiles

Introduction to QGIS

Plugins

Points

The simplest representation of spatial data can be done using a table. A place can be represented using a pair of coordinates - Latitude and Longitude - with other attribute information about the place. Many spatial data source come in this form. Excel sheets, CSV files, database tables etc.

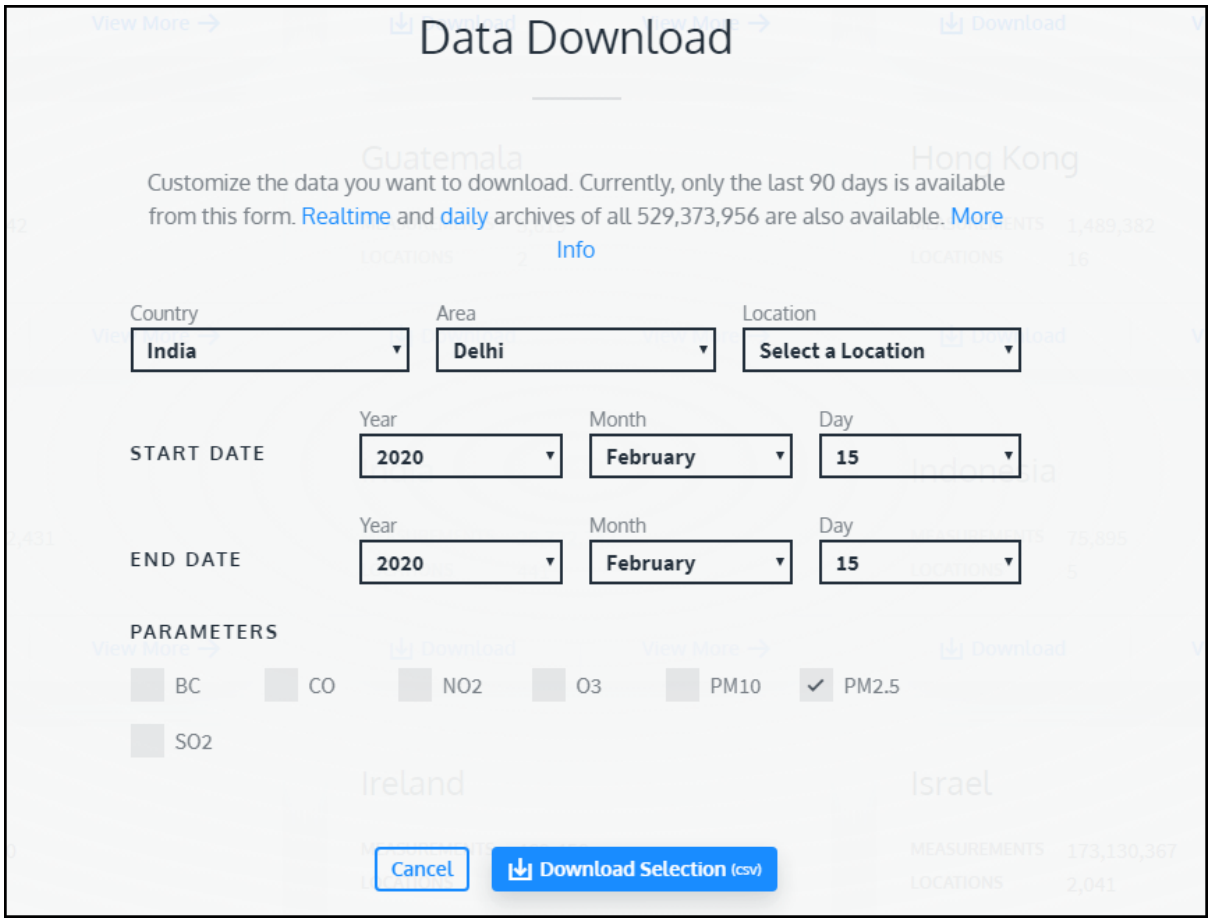
Exercise: Mapping Air Quality

Worsening air quality is a severe problem in many countries around the world. India - particularly - Delhi suffers from acute problems of high pollution levels.One of the first steps to better understand the problem, is to have continuous monitoring of air quality across the cities. Many organizations have stepped up and setup such sensors that collect air quality data and make it publicly available. OpenAQ is a platform that collects this data from all public sources and makes it available in an easy to use form.

> If you are interested in this topic, Urban Emissions has a lot of relevant information and datasets for India.

For this exercise, we will take the sensor data for PM2.5 concentrations 1 day and map it. The aim is to turn this tabular data info an informative spatial data visualization.

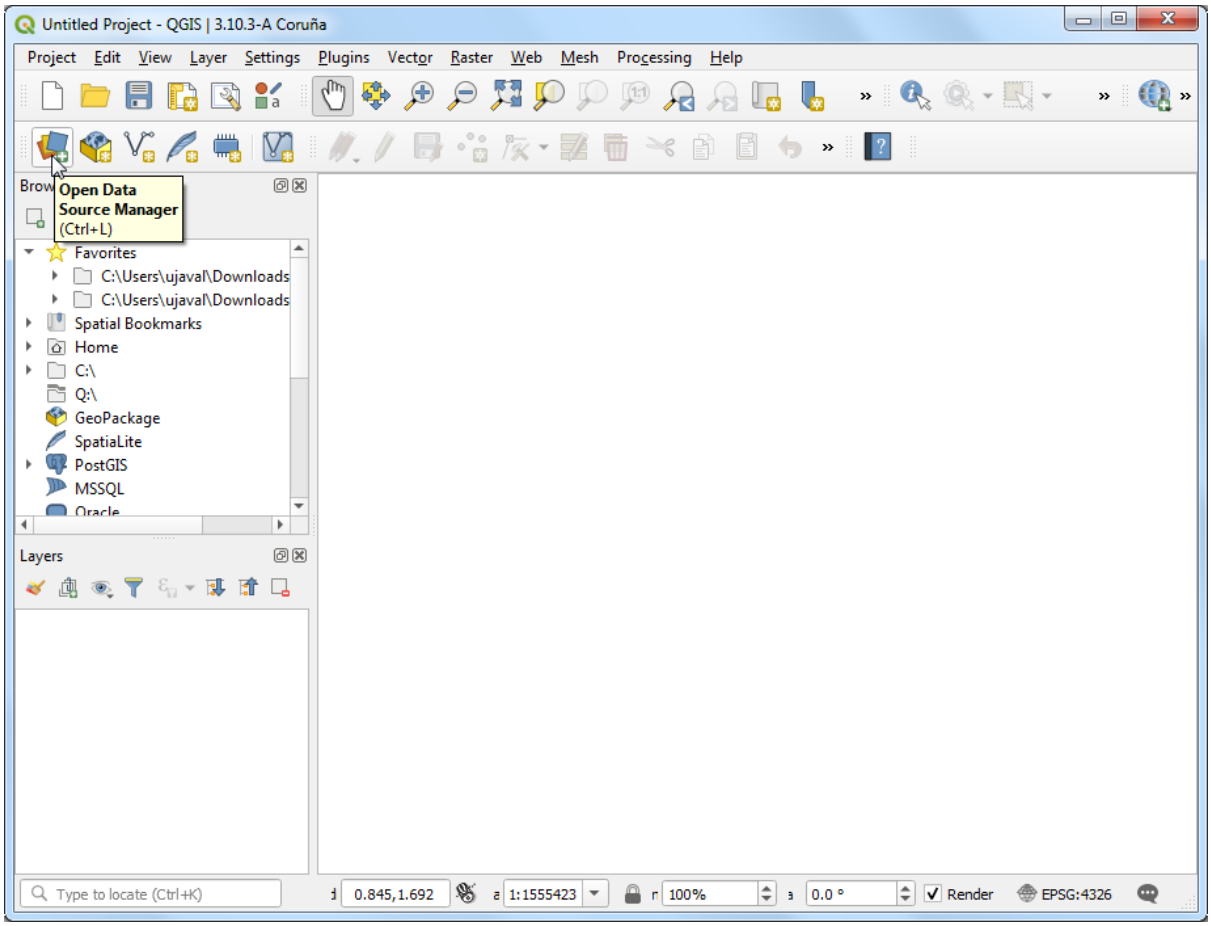
For this exercise, we are using daily average data for Delhi, India for February 15, 2020. This data was downloaded from OpenAQ Data Download



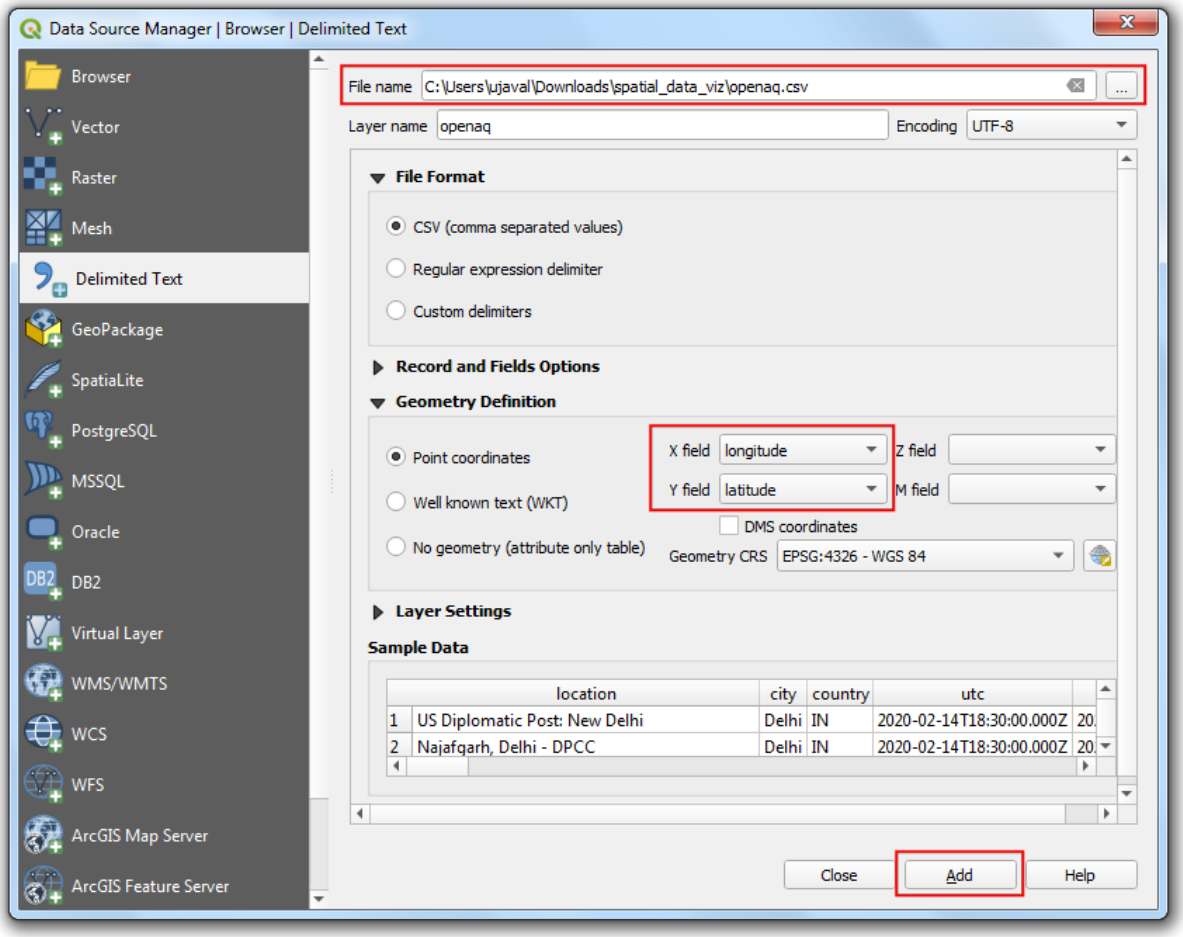
Open the `openaq.csv` file in a text editor and examine it. Each row of data contains data from 1 monitoring station. The `latitude` and `longitude` column contain the coordinates of the station and the `value` contains the daily average PM2.5 concentration

location	city	country	utc	local	parameter	value	unit	latitude	longitude	attribution
US Diplomatic Post: New Delhi	Delhi	IN	2020-02-14T18:30:00.000Z	2020-02-15T00:00:00+05:30	pm25	67	µg/m³	28.63576	77.22445	{"name":"EPA AirNow I
Najafgarh, Delhi - DPCC	Delhi	IN	2020-02-14T18:30:00.000Z	2020-02-15T00:00:00+05:30	pm25	195	µg/m³	28.570173	76.933762	{"name":"Central Pollu
Alipur, Delhi - DPCC	Delhi	IN	2020-02-14T18:30:00.000Z	2020-02-15T00:00:00+05:30	pm25	85	µg/m³	28.815329	77.15301	{"name":"Central Pollu
Mundka, Delhi - DPCC	Delhi	IN	2020-02-14T18:30:00.000Z	2020-02-15T00:00:00+05:30	pm25	68	µg/m³	28.684678	77.076574	{"name":"Central Pollu
Sri Aurobindo Marg, Delhi - DPCC	Delhi	IN	2020-02-14T18:30:00.000Z	2020-02-15T00:00:00+05:30	pm25	85	µg/m³	28.531346	77.190156	{"name":"Central Pollu
Narela, Delhi - DPCC	Delhi	IN	2020-02-14T18:30:00.000Z	2020-02-15T00:00:00+05:30	pm25	101	µg/m³	28.822836	77.101981	{"name":"Central Pollu
Vivek Vihar, Delhi - DPCC	Delhi	IN	2020-02-14T18:30:00.000Z	2020-02-15T00:00:00+05:30	pm25	84	µg/m³	28.672342	77.31526	{"name":"Central Pollu
Okhla Phase-2, Delhi - DPCC	Delhi	IN	2020-02-14T18:30:00.000Z	2020-02-15T00:00:00+05:30	pm25	102	µg/m³	28.530785	77.271255	{"name":"Central Pollu
Ashok Vihar, Delhi - DPCC	Delhi	IN	2020-02-14T18:30:00.000Z	2020-02-15T00:00:00+05:30	pm25	74	µg/m³	28.695381	77.181665	{"name":"Central Pollu
Dr. Karni Singh Shooting Range, Delhi - DPCC	Delhi	IN	2020-02-14T18:30:00.000Z	2020-02-15T00:00:00+05:30	pm25	68	µg/m³	28.498571	77.26484	{"name":"Central Pollu
Punjabi Bagh, Delhi - DPCC	Delhi	IN	2020-02-14T18:30:00.000Z	2020-02-15T00:00:00+05:30	pm25	63	µg/m³	28.674045	77.131023	{"name":"Central Pollu
R K Puram, Delhi - DPCC	Delhi	IN	2020-02-14T18:30:00.000Z	2020-02-15T00:00:00+05:30	pm25	72	µg/m³	28.563262	77.186937	{"name":"Central Pollu
Mandir Marg, Delhi - DPCC	Delhi	IN	2020-02-14T18:30:00.000Z	2020-02-15T00:00:00+05:30	pm25	70	µg/m³	28.636429	77.201067	{"name":"Central Pollu
Bawana, Delhi - DPCC	Delhi	IN	2020-02-14T18:30:00.000Z	2020-02-15T00:00:00+05:30	pm25	150	µg/m³	28.7762	77.051074	{"name":"Central Pollu
Sonia Vihar, Delhi - DPCC	Delhi	IN	2020-02-14T18:30:00.000Z	2020-02-15T00:00:00+05:30	pm25	77	µg/m³	28.710508	77.249485	{"name":"Central Pollu
Patparganj, Delhi - DPCC	Delhi	IN	2020-02-14T18:30:00.000Z	2020-02-15T00:00:00+05:30	pm25	52	µg/m³	28.623748	77.287205	{"name":"Central Pollu
Nehru Nagar, Delhi - DPCC	Delhi	IN	2020-02-14T18:30:00.000Z	2020-02-15T00:00:00+05:30	pm25	59	µg/m³	28.56789	77.250515	{"name":"Central Pollu
Jawaharlal Nehru Stadium, Delhi - DPCC	Delhi	IN	2020-02-14T18:30:00.000Z	2020-02-15T00:00:00+05:30	pm25	57	µg/m³	28.58028	77.233829	{"name":"Central Pollu
Jahangirpuri, Delhi - DPCC	Delhi	IN	2020-02-14T18:30:00.000Z	2020-02-15T00:00:00+05:30	pm25	106	µg/m³	28.73282	77.170633	{"name":"Central Pollu
Dwarka-Sector 8, Delhi - DPCC	Delhi	IN	2020-02-14T18:30:00.000Z	2020-02-15T00:00:00+05:30	pm25	103	µg/m³	28.5710274	77.0719006	{"name":"Central Pollu
NSIT Dwarka, Delhi - CPCB	Delhi	IN	2020-02-14T18:30:00.000Z	2020-02-15T00:00:00+05:30	pm25	118.51	µg/m³	28.60909	77.0325413	{"name":"Central Pollu
ITO, Delhi - CPCB	Delhi	IN	2020-02-14T18:30:00.000Z	2020-02-15T00:00:00+05:30	pm25	53	µg/m³	28.628624	77.24106	{"name":"Central Pollu
DTU, Delhi - CPCB	Delhi	IN	2020-02-14T18:30:00.000Z	2020-02-15T00:00:00+05:30	pm25	133.45	µg/m³	28.7500499	77.1112615	{"name":"Central Pollu
Shadipur, Delhi - CPCB	Delhi	IN	2020-02-14T18:30:00.000Z	2020-02-15T00:00:00+05:30	pm25	63.41	µg/m³	28.6514781	77.1473105	{"name":"Central Pollu
IHBAS, Dilshad Garden, Delhi - CPCB	Delhi	IN	2020-02-14T18:30:00.000Z	2020-02-15T00:00:00+05:30	pm25	52.4	µg/m³	28.6811736	77.3025234	{"name":"Central Pollu
North Campus, DU, Delhi - IMD	Delhi	IN	2020-02-14T18:30:00.000Z	2020-02-15T00:00:00+05:30	pm25	37.95	µg/m³	28.6573814	77.1585447	{"name":"Central Pollu

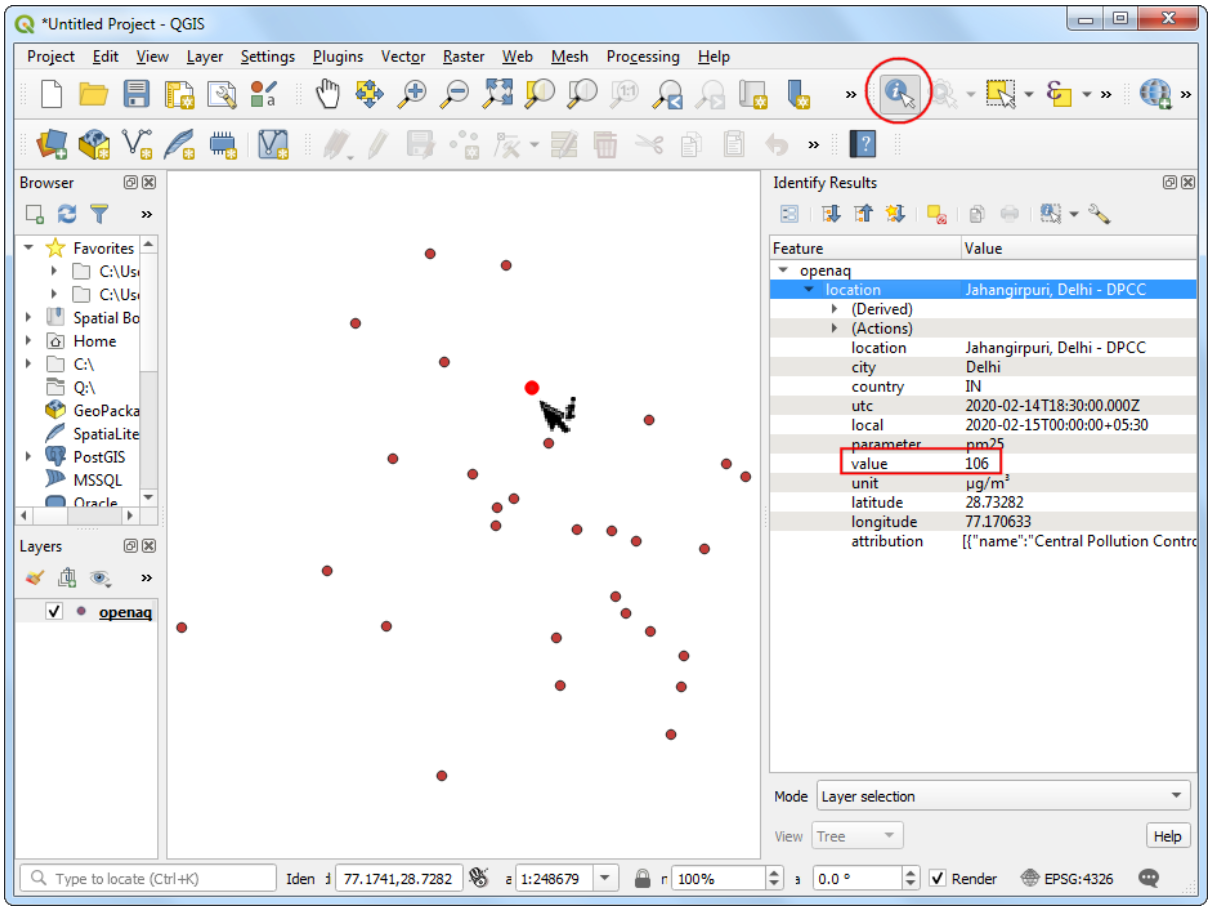
Tabular data in text files fall into the category of **Delimited Text** files, such as this can be imported in QGIS via *Data Source Manager*. Click *Open Data Source Manager* button.



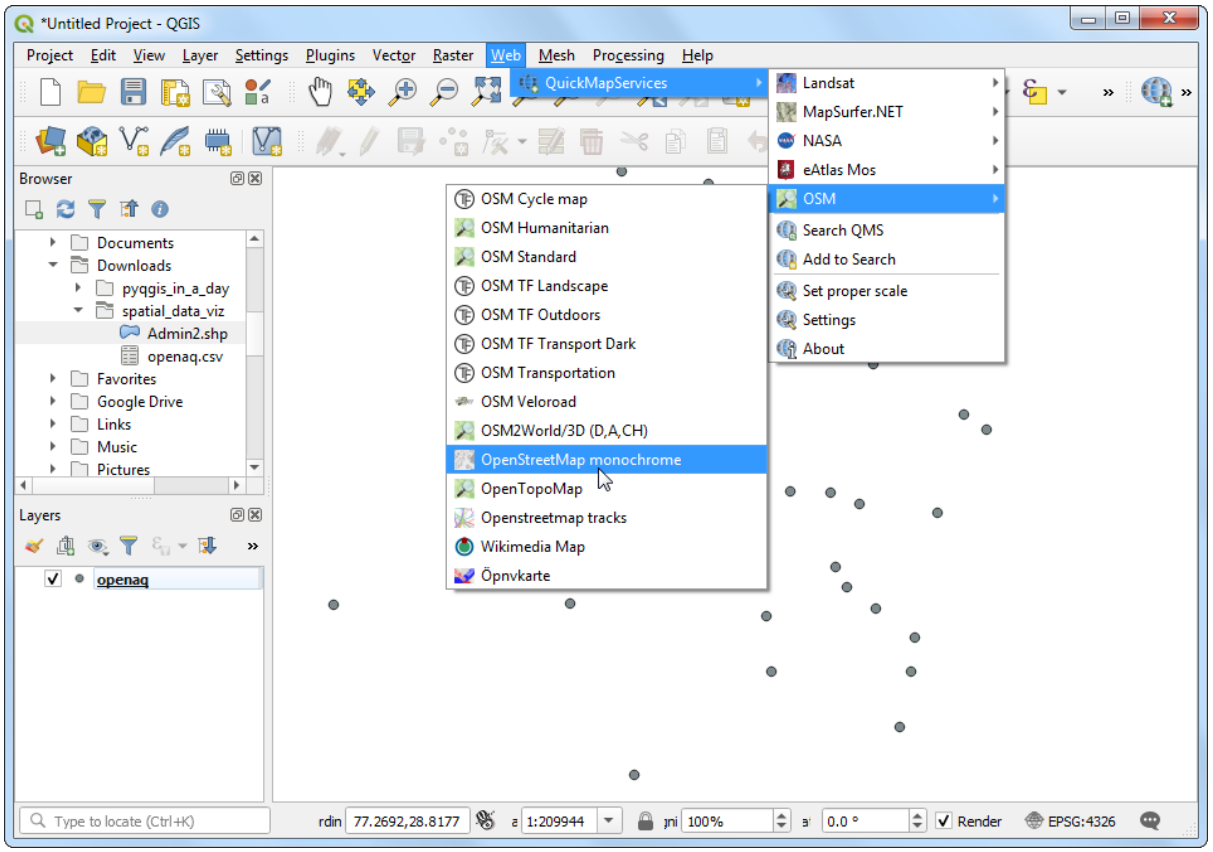
Browse to the `openaq.csv` file and open it. As we want to import this file as points, select *Point coordinates*. Choose `longitude` as *X Field* and `latitude` as *Y Field*. Click *Add*.



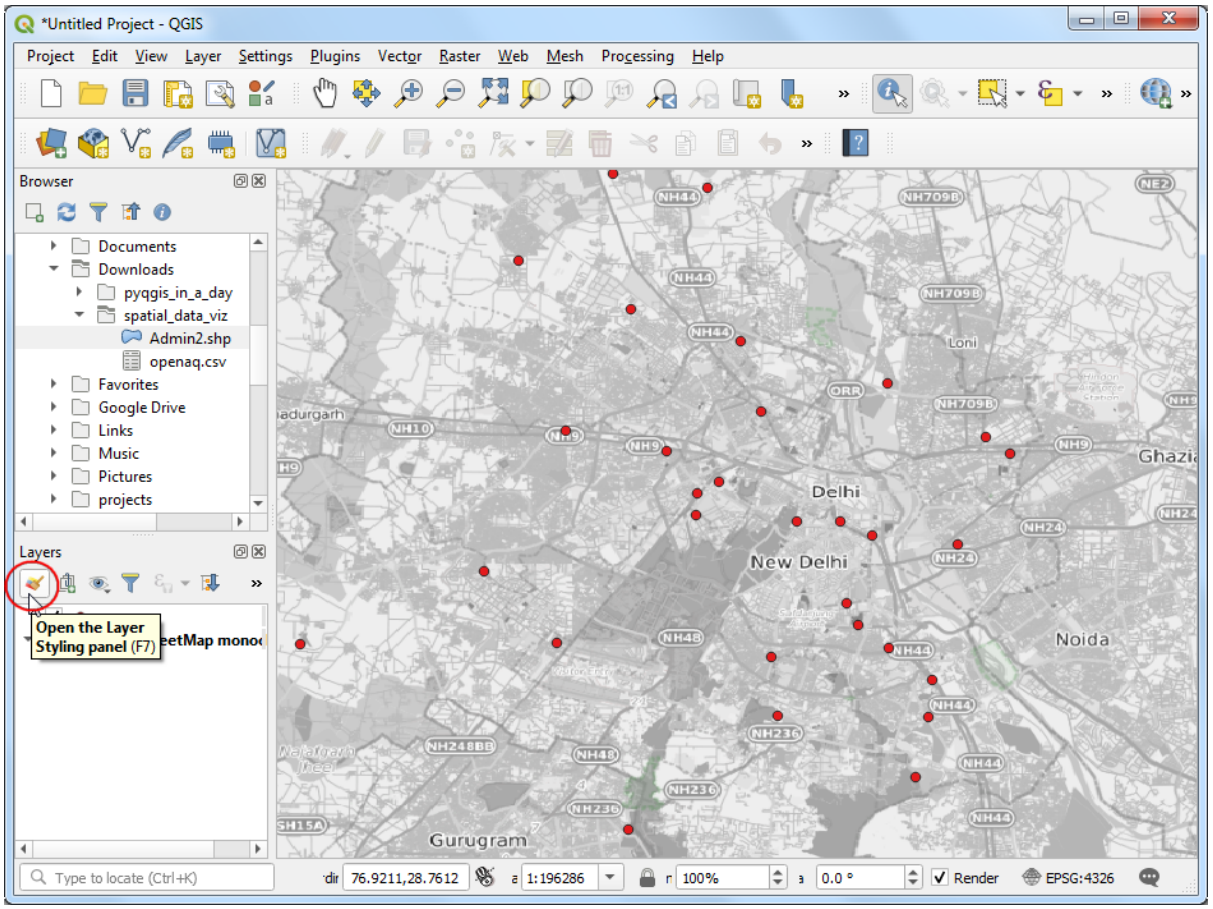
You will see the tabular data now loaded in QGIS canvas as a spatial data layer. Use the *Identify* button and click on any of the point. You will see the attribute data that is attached to each point.



Through we can see the point distribution is across the city of Delhi, we are missing the context on where is point is located. A base-map layer will help us understand this data better. The **QuickMapServices** plugin gives us ready access to many different types of base-maps. Go to **Web → QuickMapServices → OSM → OpenStreetMap monochrome** layer.

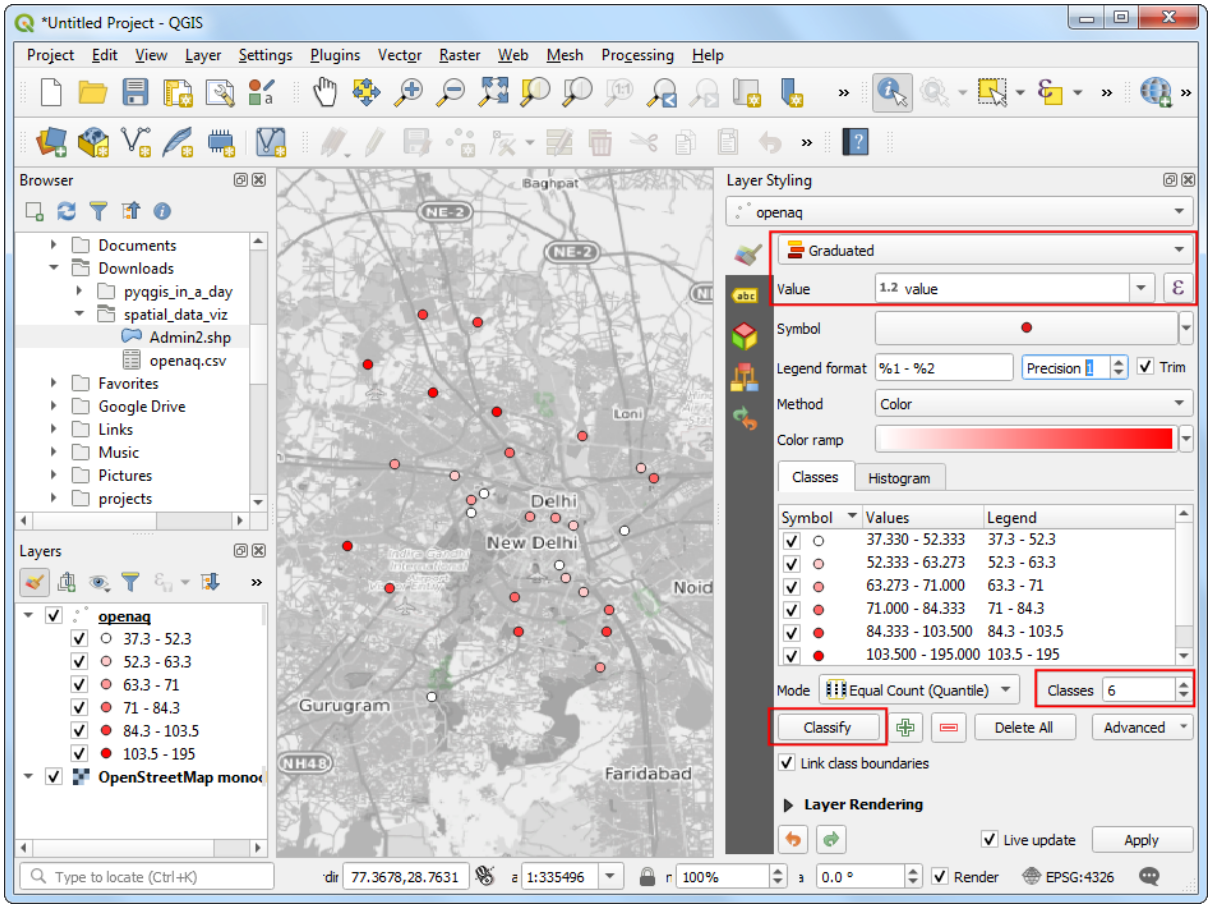


A new layer will get added to the *Layers* panel and the *Canvas*. Now you can see the points in the context of the city and surroundings. Let's style the point layer better now. Click *Open the Layer Styling Panel*.



We will color each point according to the observed PM2.5 value. Choose `Graduated`

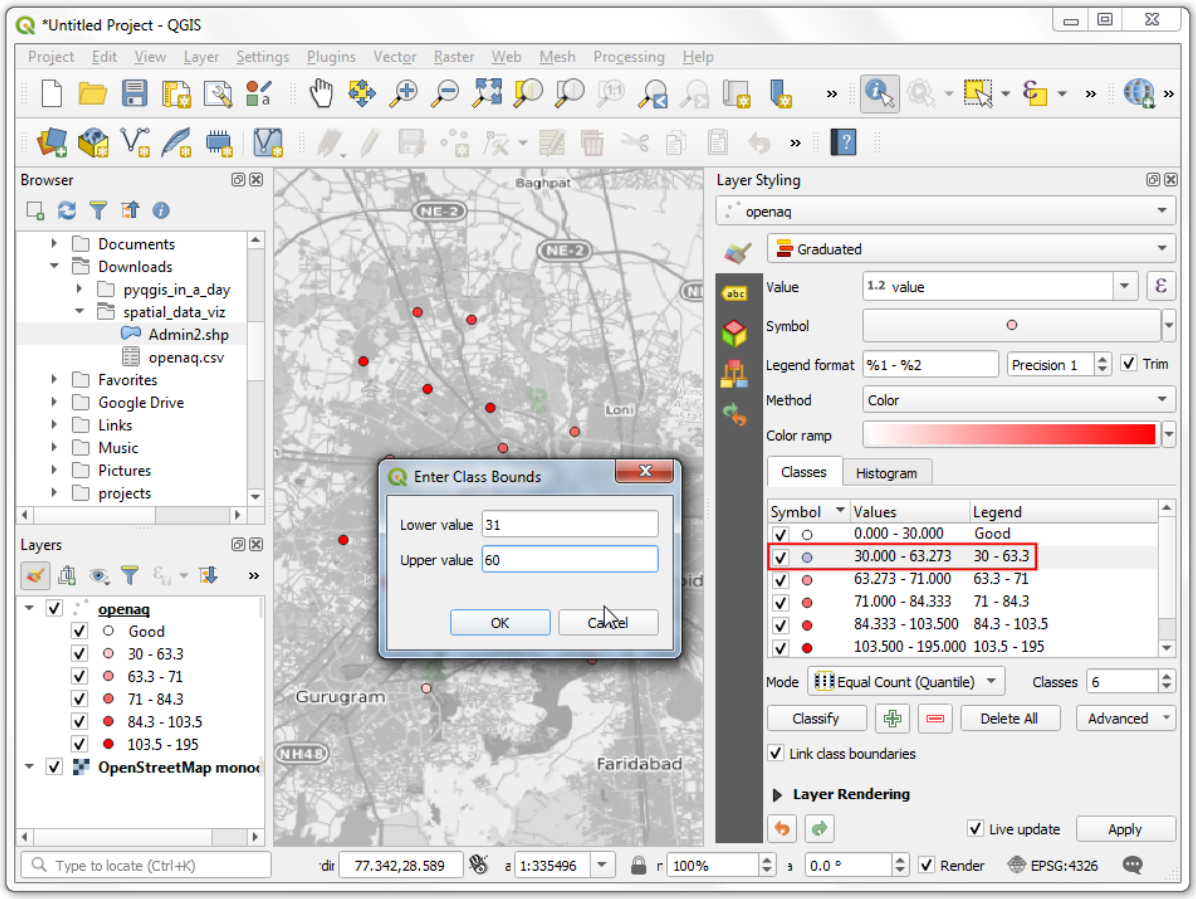
renderer and `value` as the *Value* column. Set the number of *Classes* to `6` and click *Classify*.



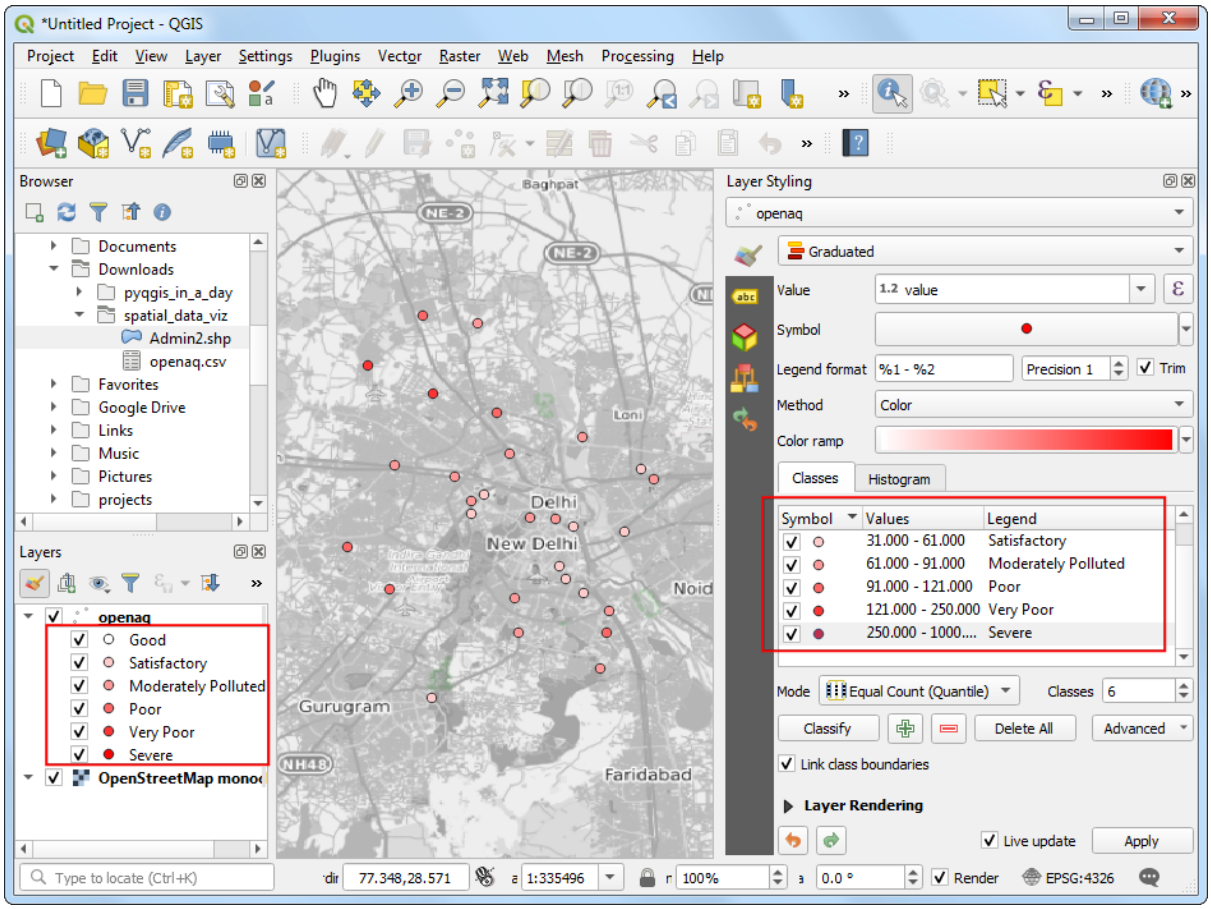
For the class ranges to have some meaning, we need to link them to the commonly used scale. India has adopted National Air Quality Index with the following definitions.

AQI Category, Pollutants and Health Breakpoints								
AQI Category (Range)	**PM_{10} 24-hr**	**$PM_{2.5}$ 24-hr**	**NO_2 24-hr**	**O_3 8-hr**	**CO 8-hr (mg/m^3)**	**SO_2 24-hr**	**NH_3 24-hr**	**Pb 24-hr**
Good (0-50)	**0-50**	**0-30**	**0-40**	**0-50**	**0-1.0**	**0-40**	**0-200**	**0-0.5**
Satisfactory (51-100)	**51-100**	**31-60**	**41-80**	**51-100**	**1.1-2.0**	**41-80**	**201-400**	**0.5 – 1.0**
Moderately polluted (101-200)	**101-250**	**61-90**	**81-180**	**101-168**	**2.1- 10**	**81-380**	**401-800**	**1.1-2.0**
Poor (201-300)	**251-350**	**91-120**	**181-280**	**169-208**	**10-17**	**381-800**	**801-1200**	**2.1-3.0**
Very poor (301-400)	**351-430**	**121-250**	**281-400**	**209-748***	**17-34**	**801-1600**	**1200-1800**	**3.1-3.5**
Severe (401-500)	**430 +**	**250+**	**400+**	**748+***	**34+**	**1600+**	**1800+**	**3.5+**

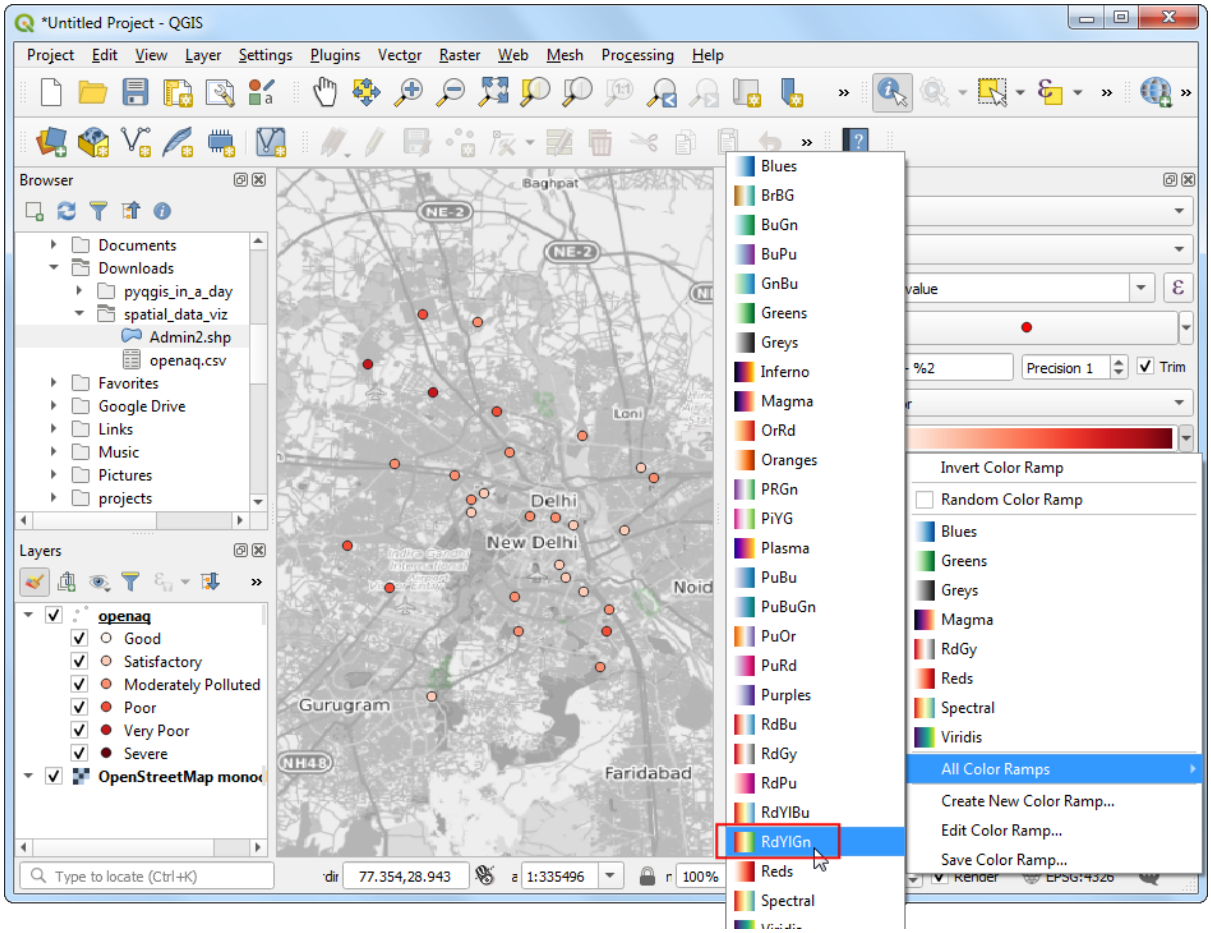
Let's adjust the class values to match those defined in the National Air Quality Index. We can also change the *Legend* labels to the human-readable category names. You can double-click each class range and edit it.



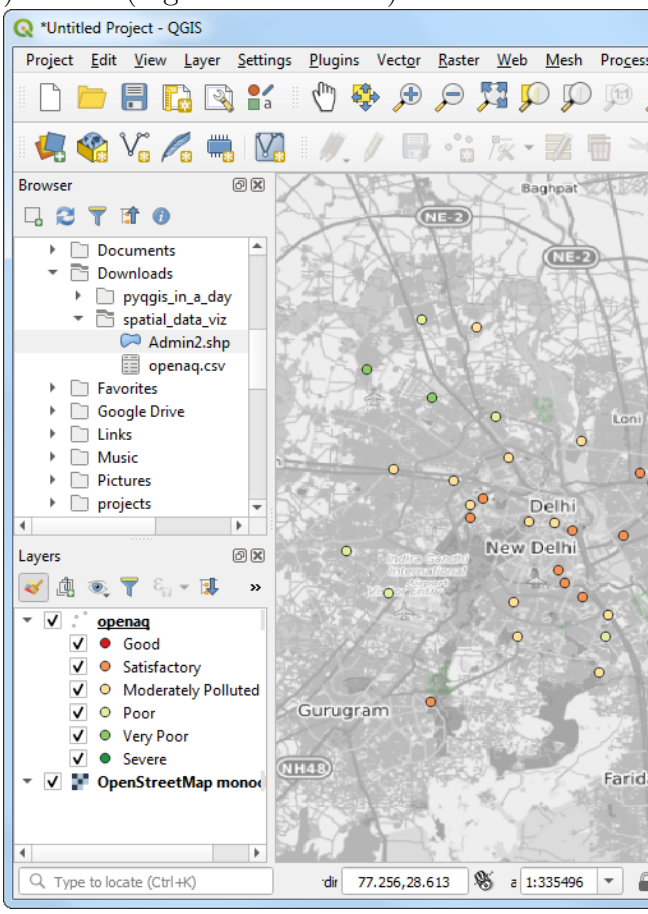
As you edit the categories, the map visualization will change accordingly. The layer legend will also show the legend labels now.

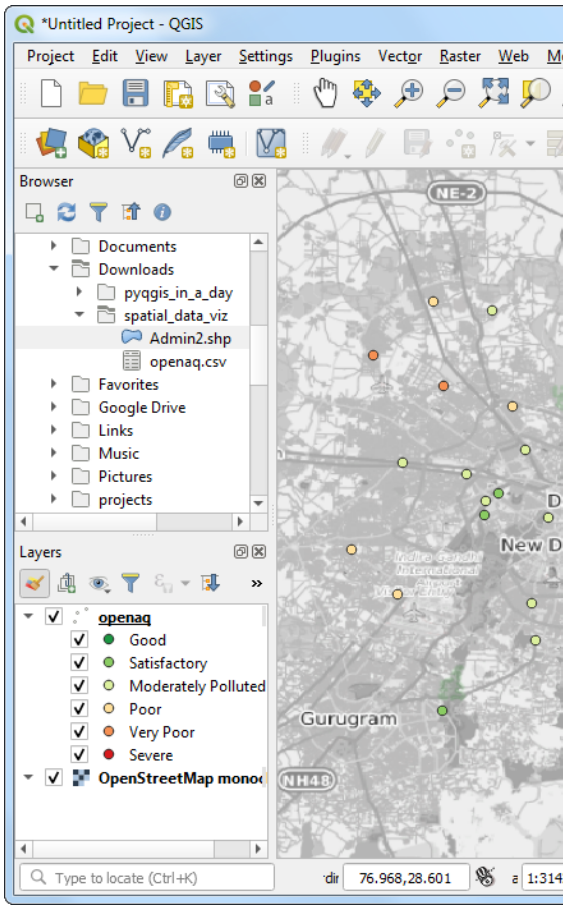


We can change the color also to match those defined in the index. Select the dropdown next to the color ramp and select the `RdYlGn` (Red-Yellow-Green) ramp.



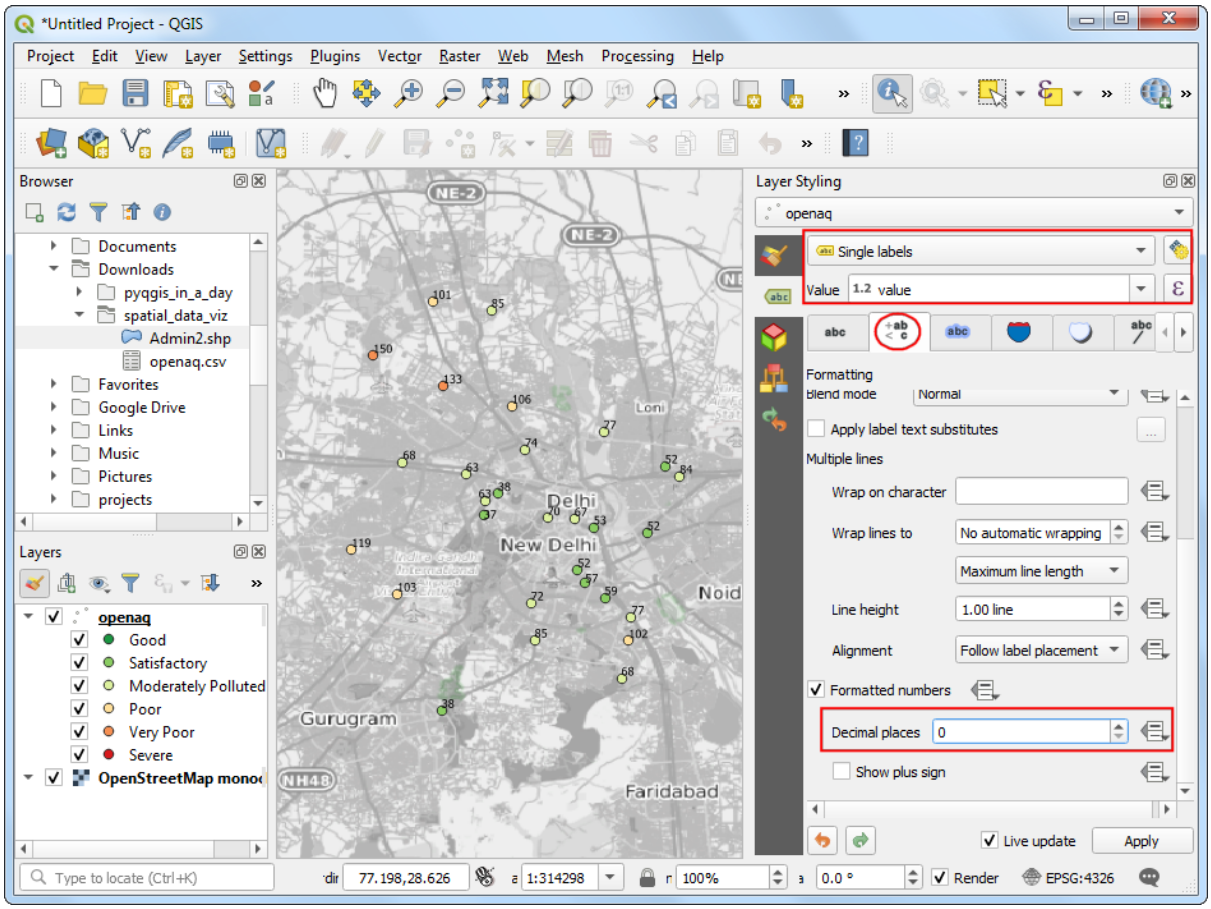
We want the ramp to go from Green (low PM2.5 values) to Red (high PM2.5 values) - so click the dropdown again and select *Invert Color Ramp*.



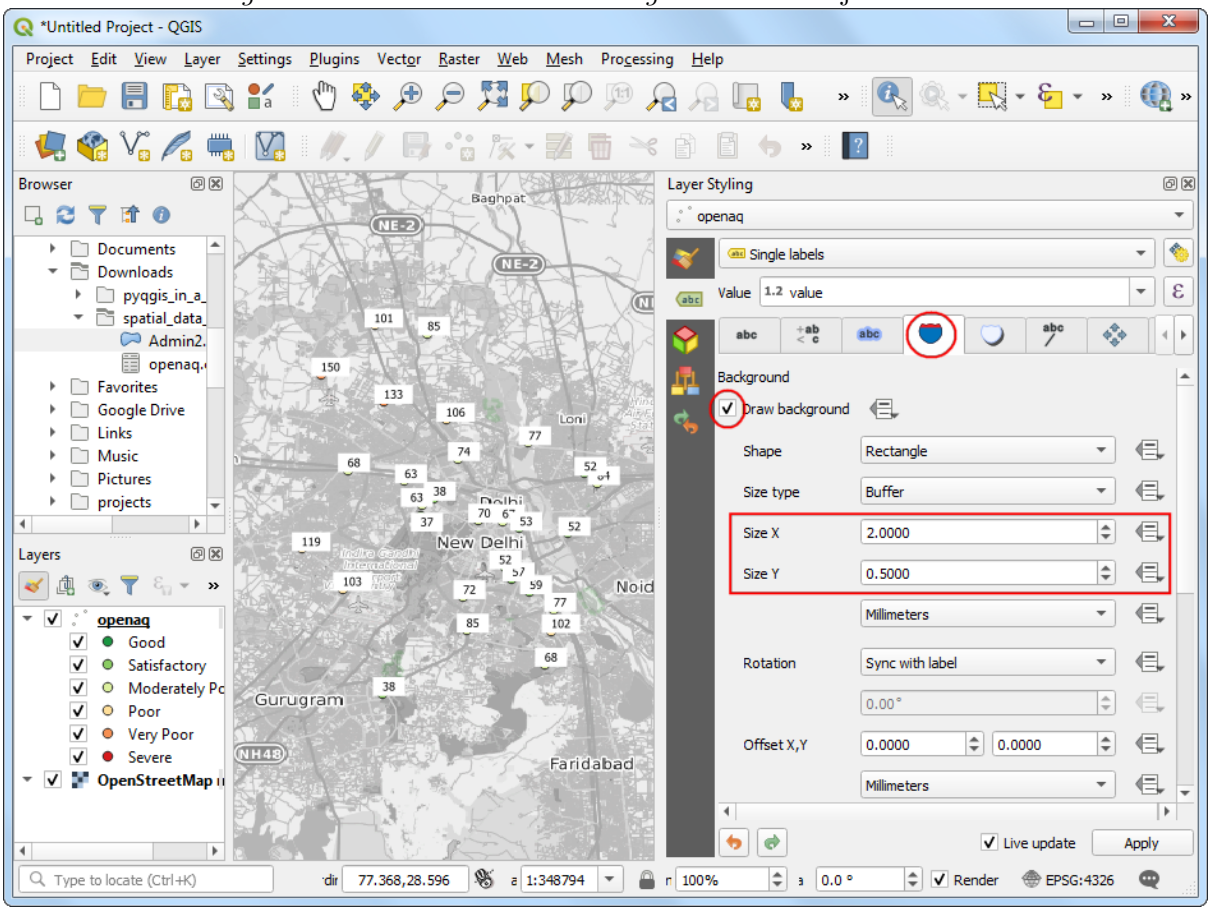


We will add labels to the points now. Switch to the *Labels* tab.

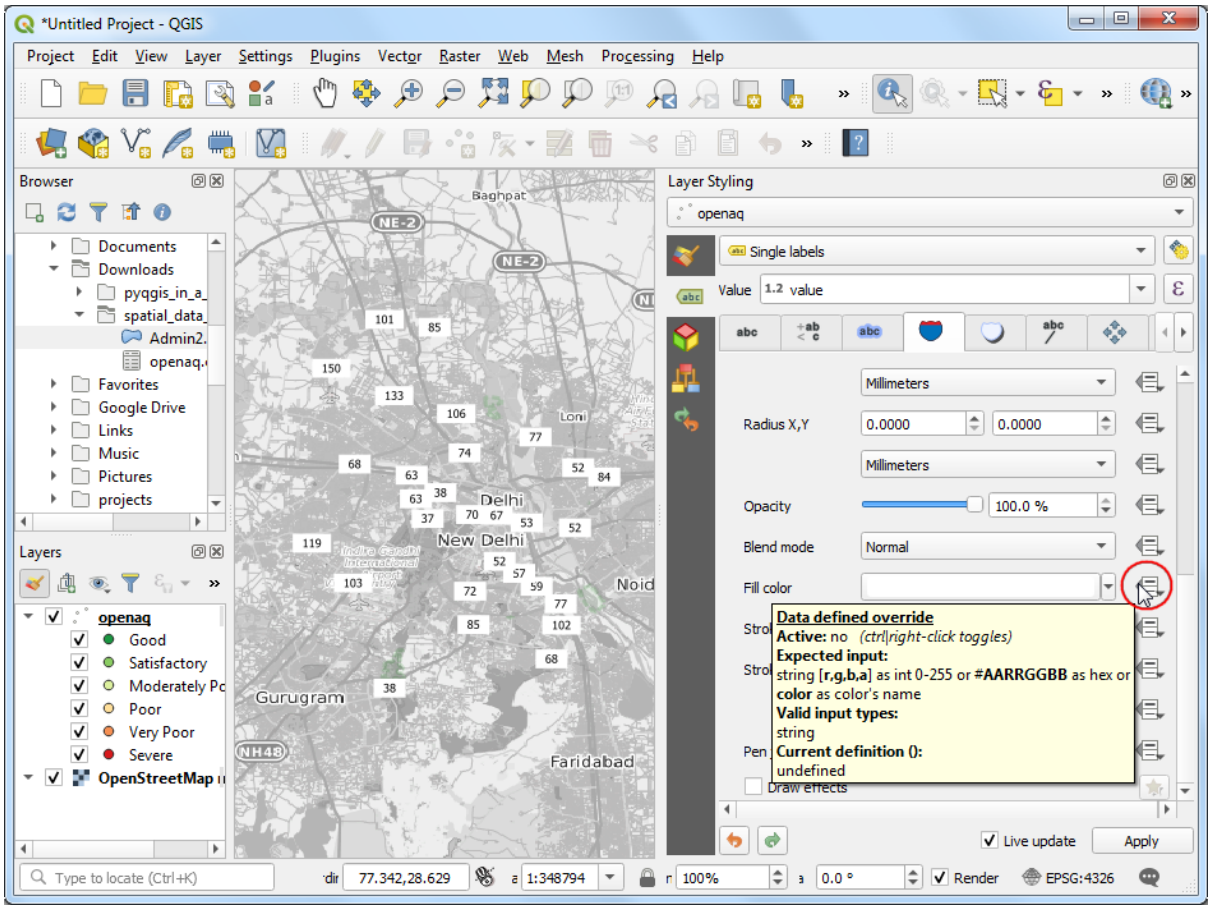
Choose `Single labels` and `value` as *Value*. Scroll down and check *Formatted numbers* and change the *Decimal places* to `0`.



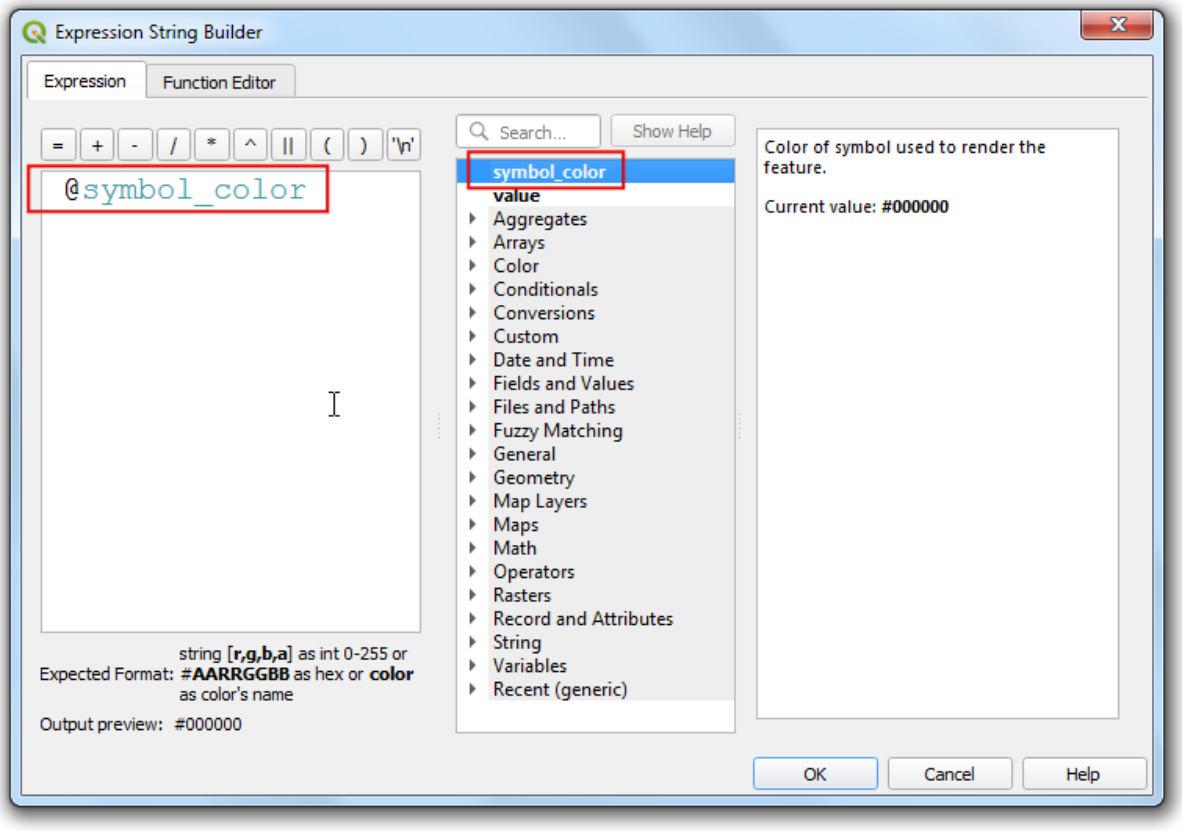
Next, switch to the *Background* tab. Check *Draw background* and adjust the *Size X* and *Size Y*.



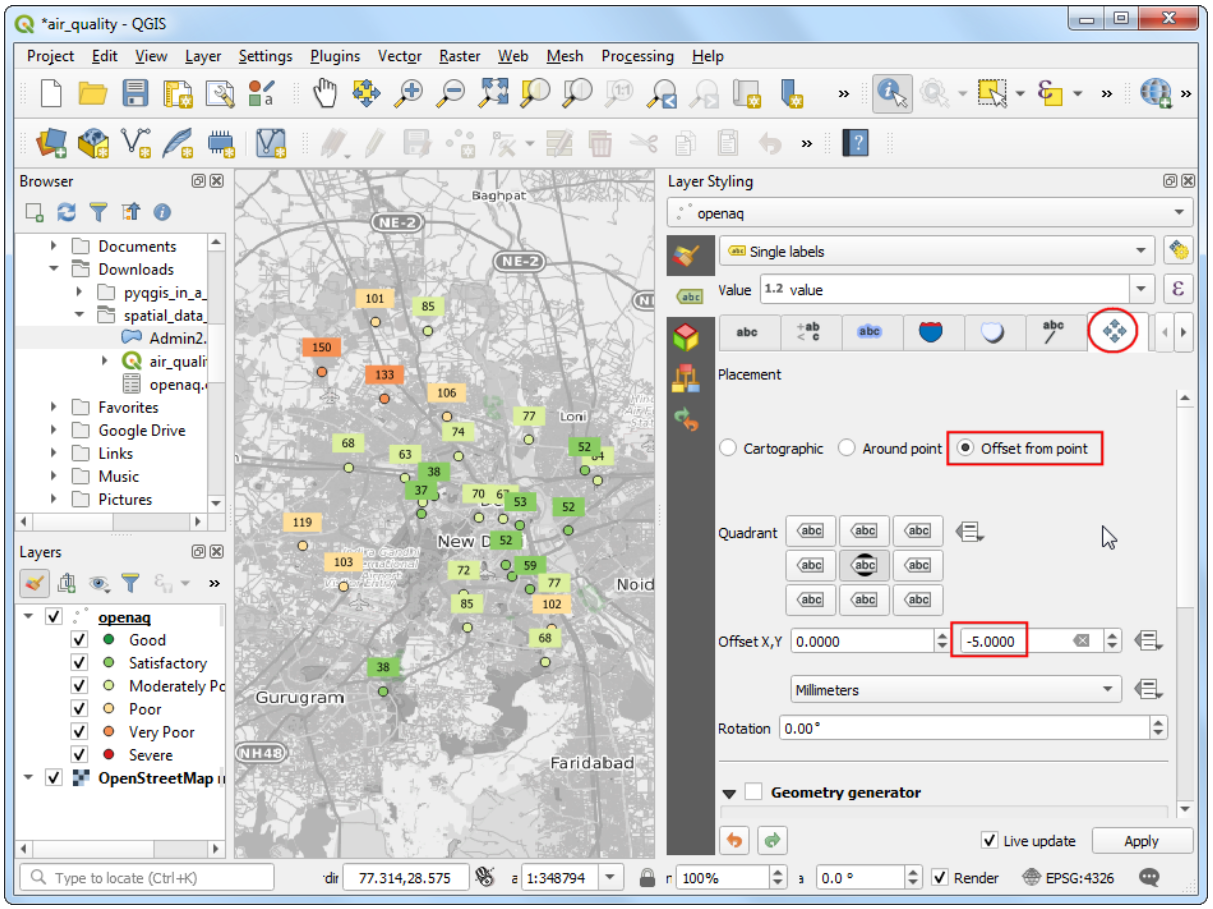
Scroll down to the *Fill color* section. We want the fill color for each box to match the color of the associated point. Click the *Data defined override* button and choose *Edit*.



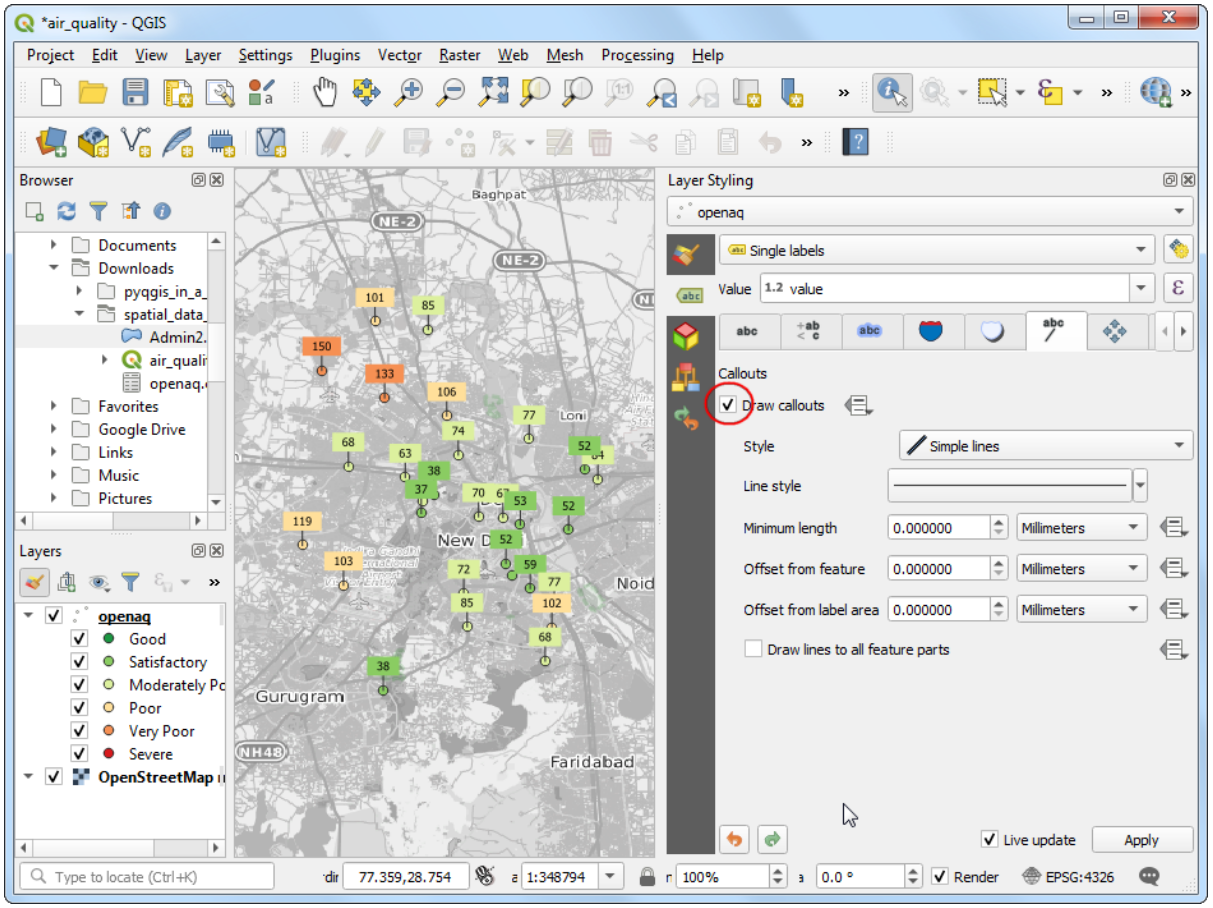
Double-click the `@symbol_color` variable to add it to the expression. Click OK.



You will see the map change and the label shields will have the color matching the category based on the values. We will move the labels slightly above so the points are also seen. Switch to the *Placement* tab and select *Offset from point*. Check the *Offset Y* to `-5`.

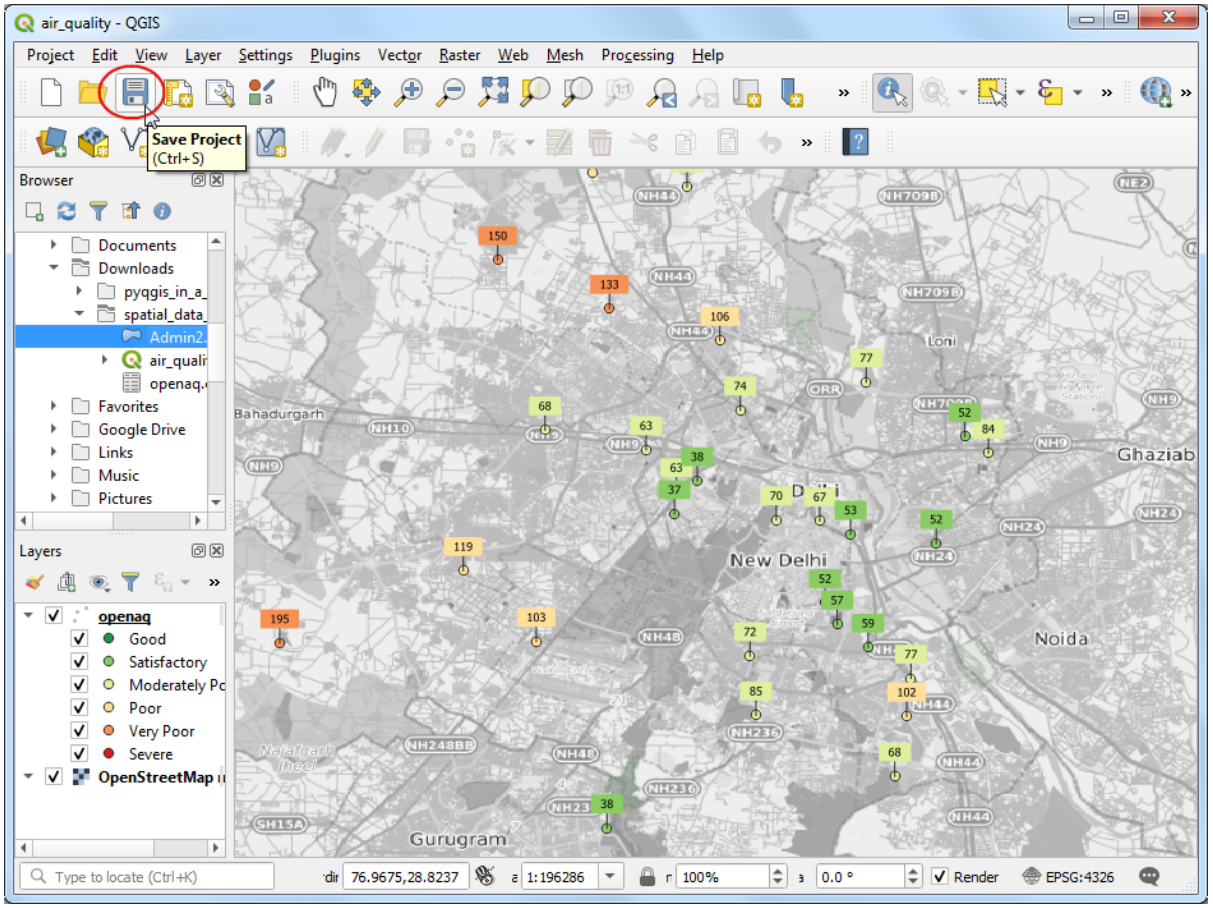


Lastly, go the *Callouts* tab and check *Draw callouts* to make it easy to see which labels belong to which point.

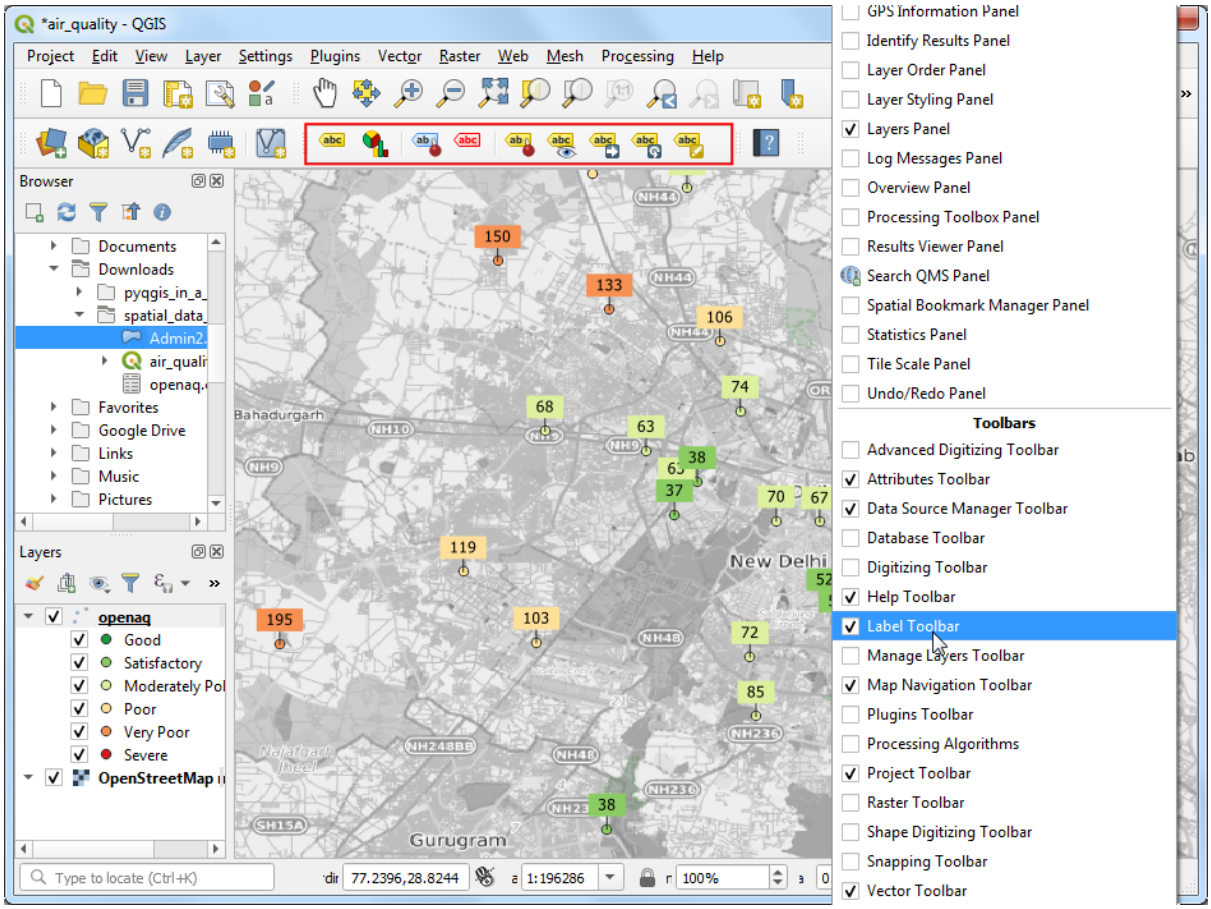


It is a good idea to save out work. Click the *Save Project* button and save it as

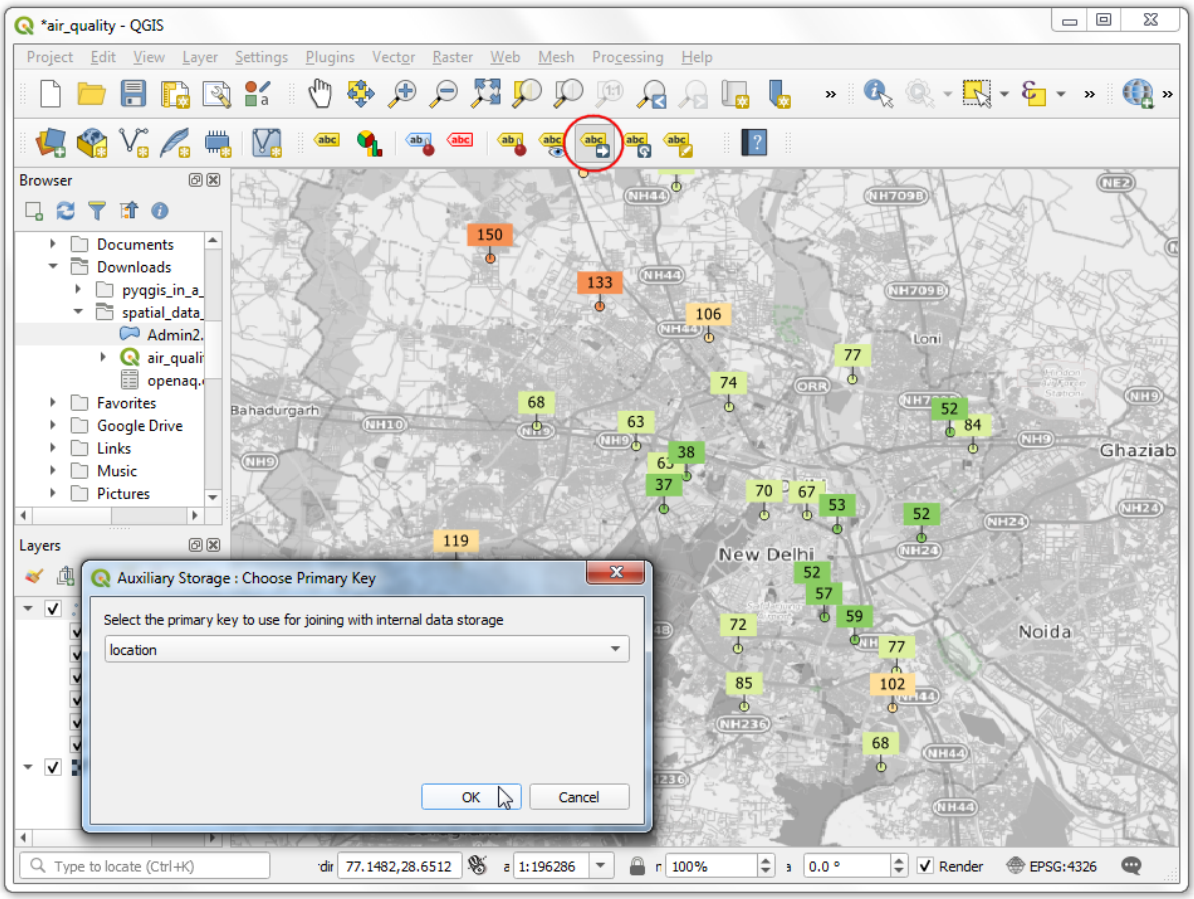
`air_quality.qgz`.



You can see most of the labels are legible, but some are too close and feel a bit cluttered. We can fix them by manually adjusting the placement. Right-click anywhere on the toolbar area and enable the *Label Toolbar*.

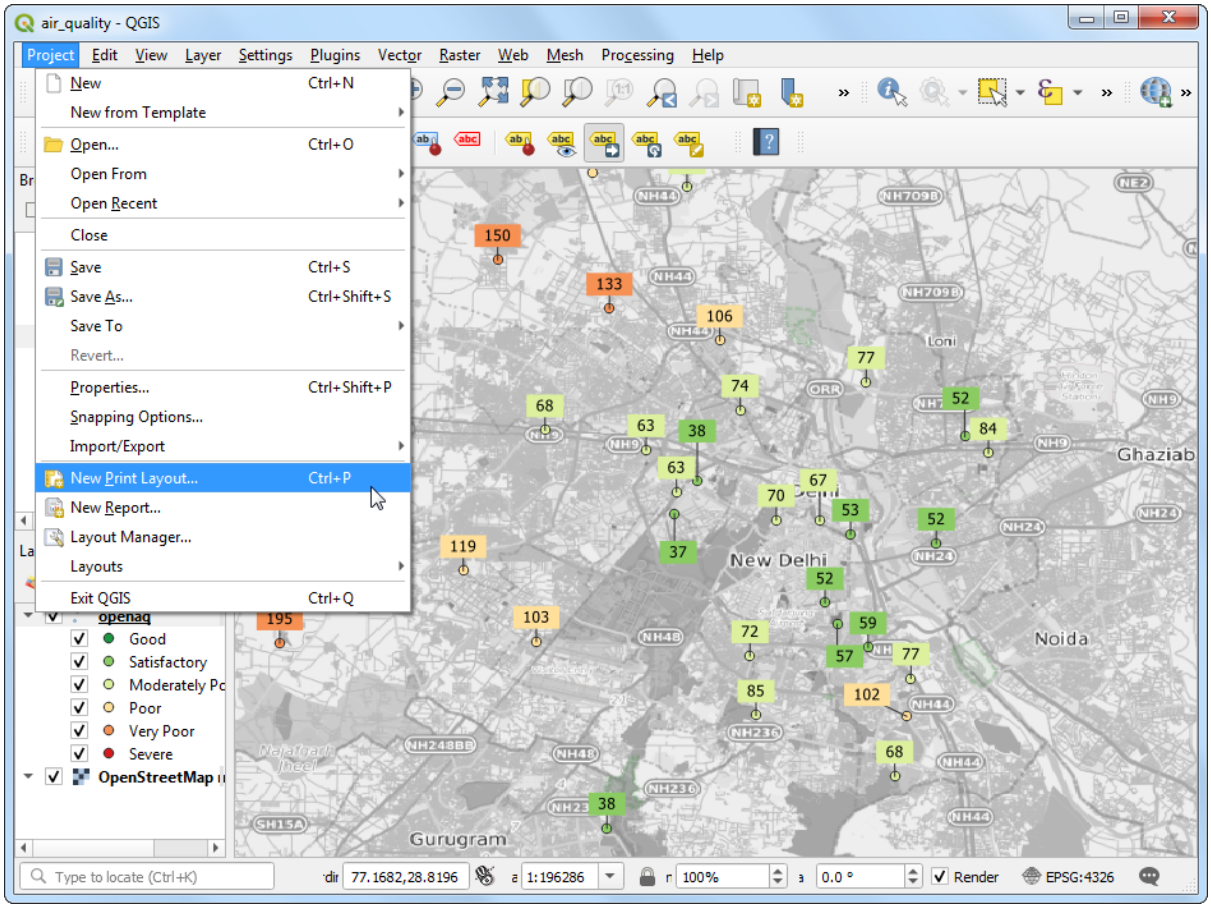


Click the *Move a Label* button. Click *OK* on the *Auxiliary Storage* prompt.

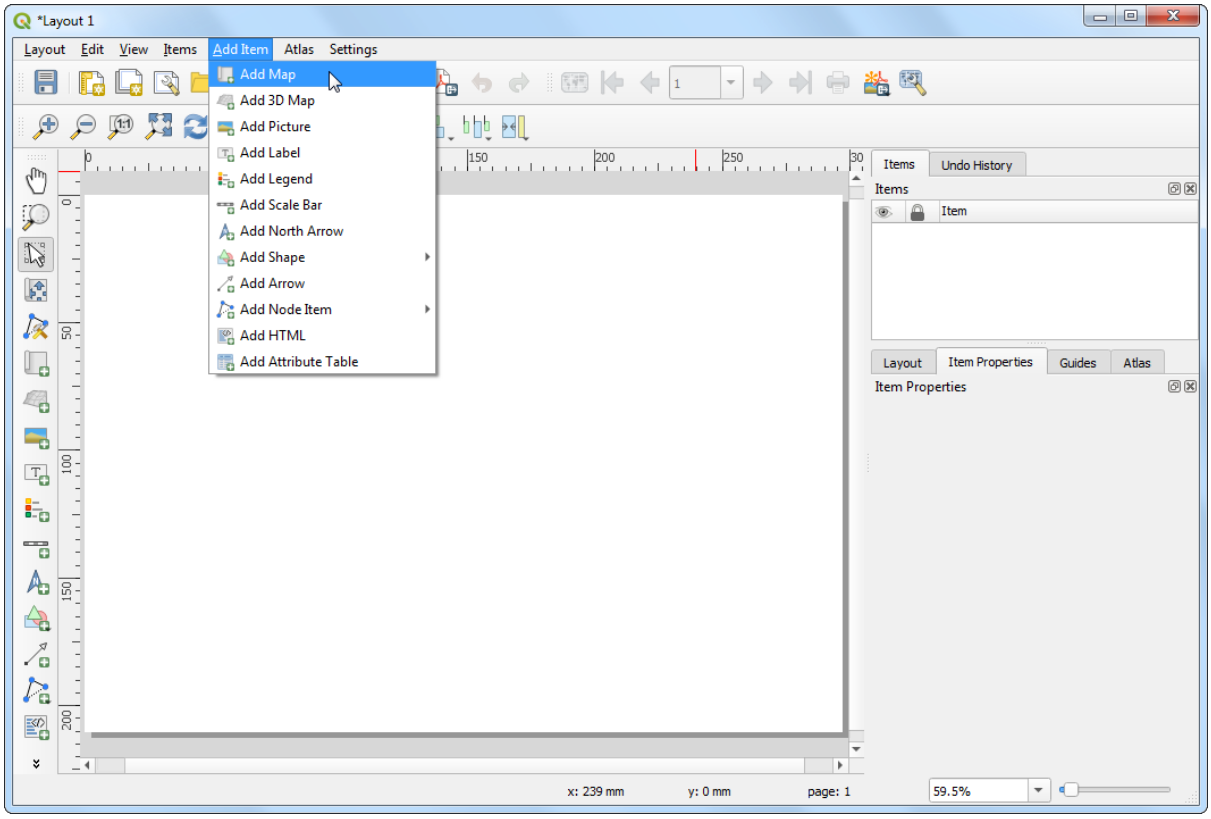


With the *Move a Label* button active, click on a any label to select it. Click at the place you want to move it to, and it will be placed there. Once you are satisfied, save the

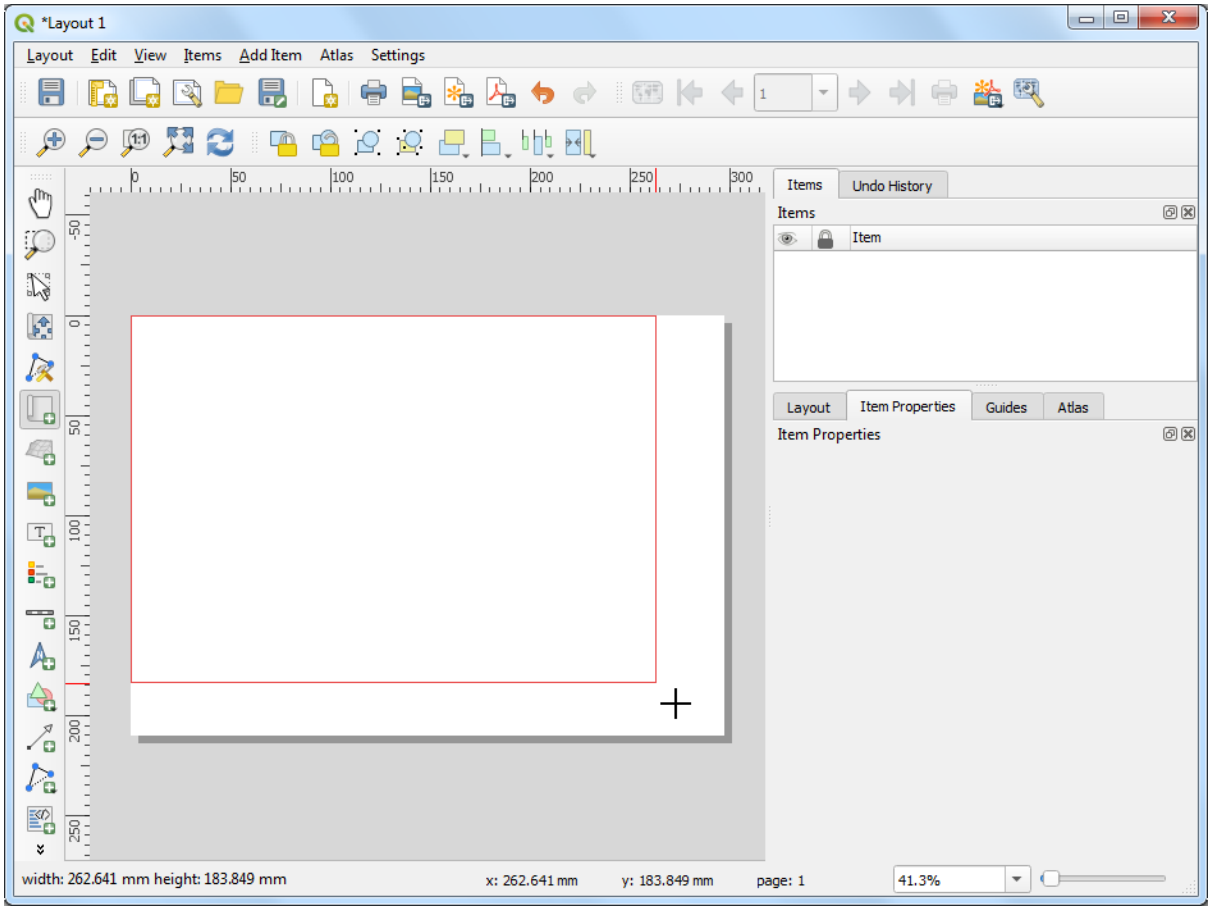
project. It is time to export our map. Go to **Project → New Print Layout...**. Leave the name empty and click *OK*.



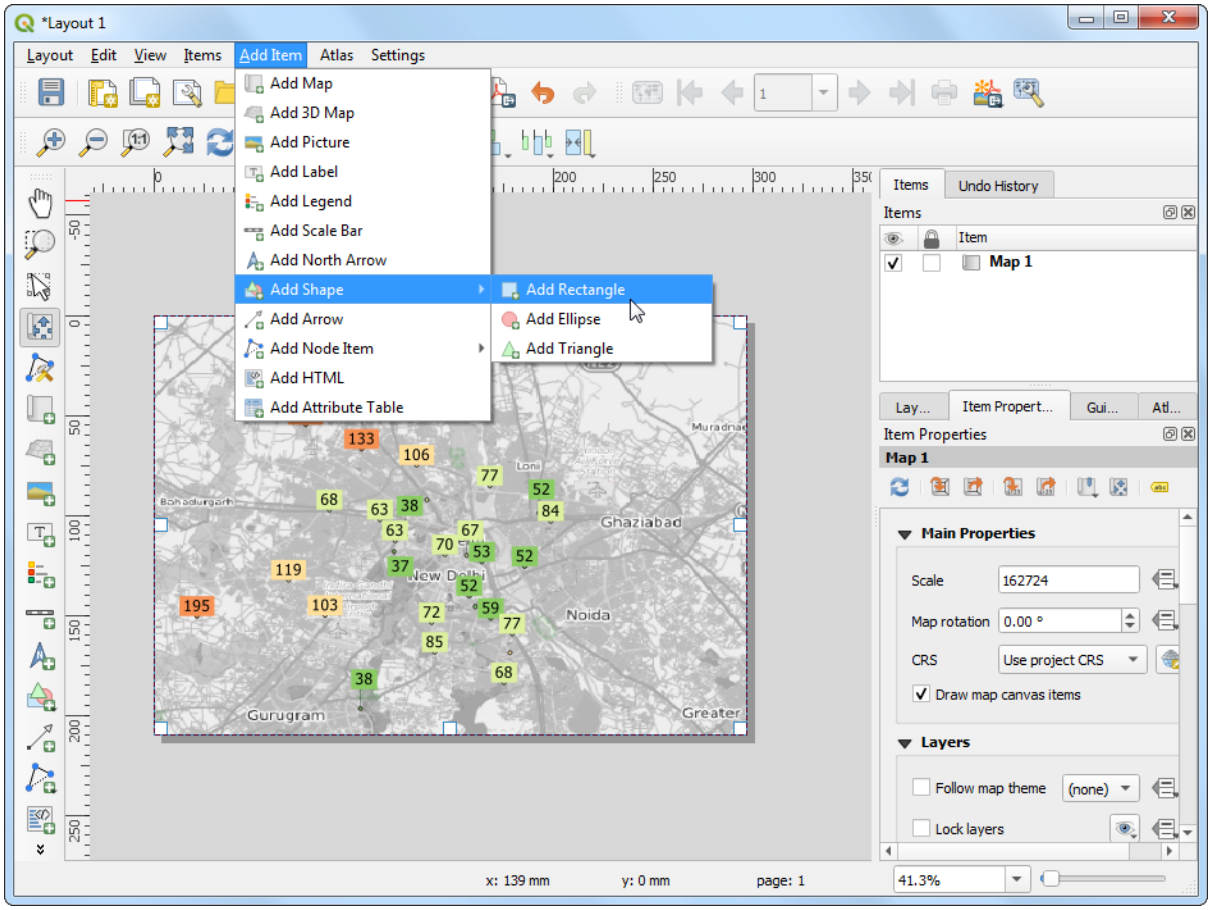
Print Layouts are a way to compose a static map with various elements such as labels, legends, north arrow, scale bar etc. Go to **Add Item → Add Map**.



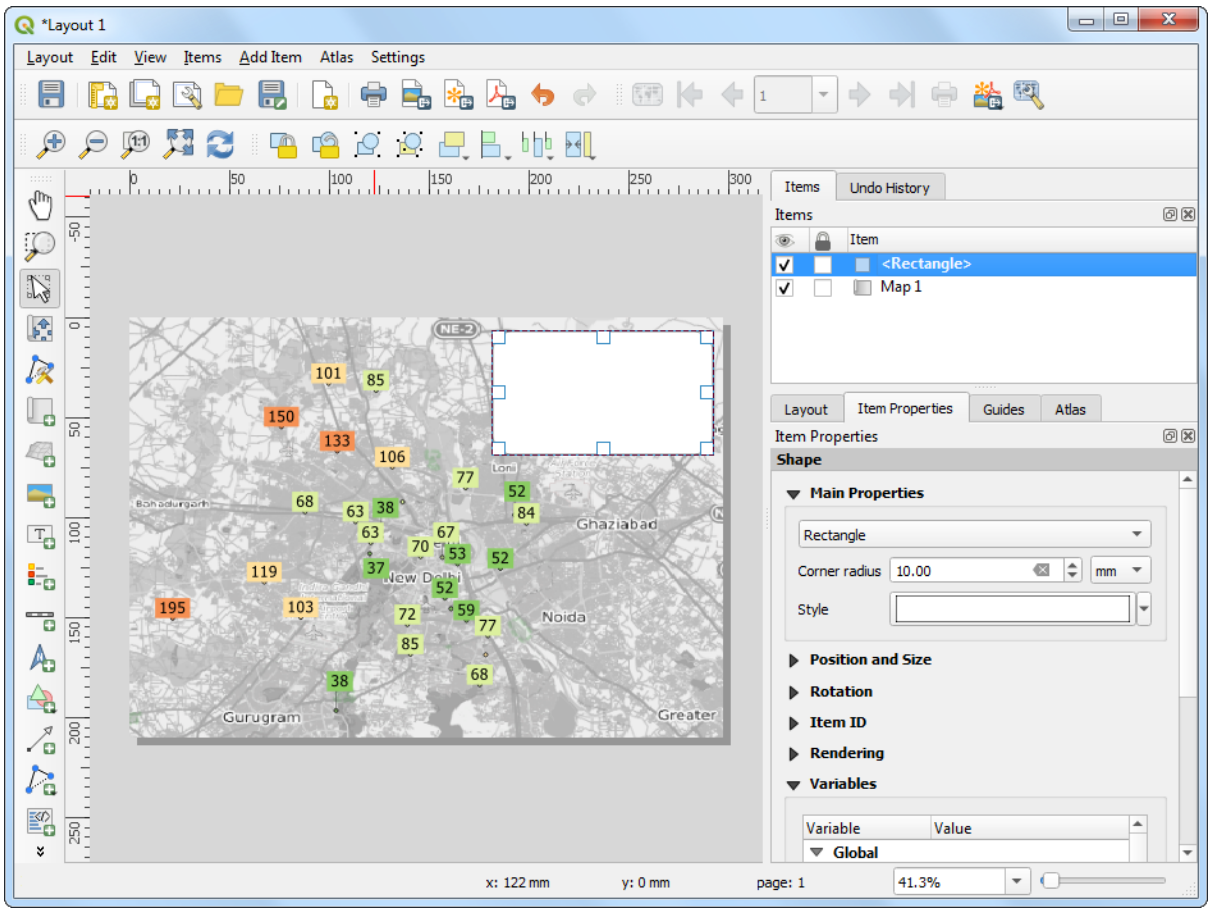
Drag a rectangle where you want the map to be rendered. Once you leg go of you your mouse button, the map from the main canvas will be loaded in the region.



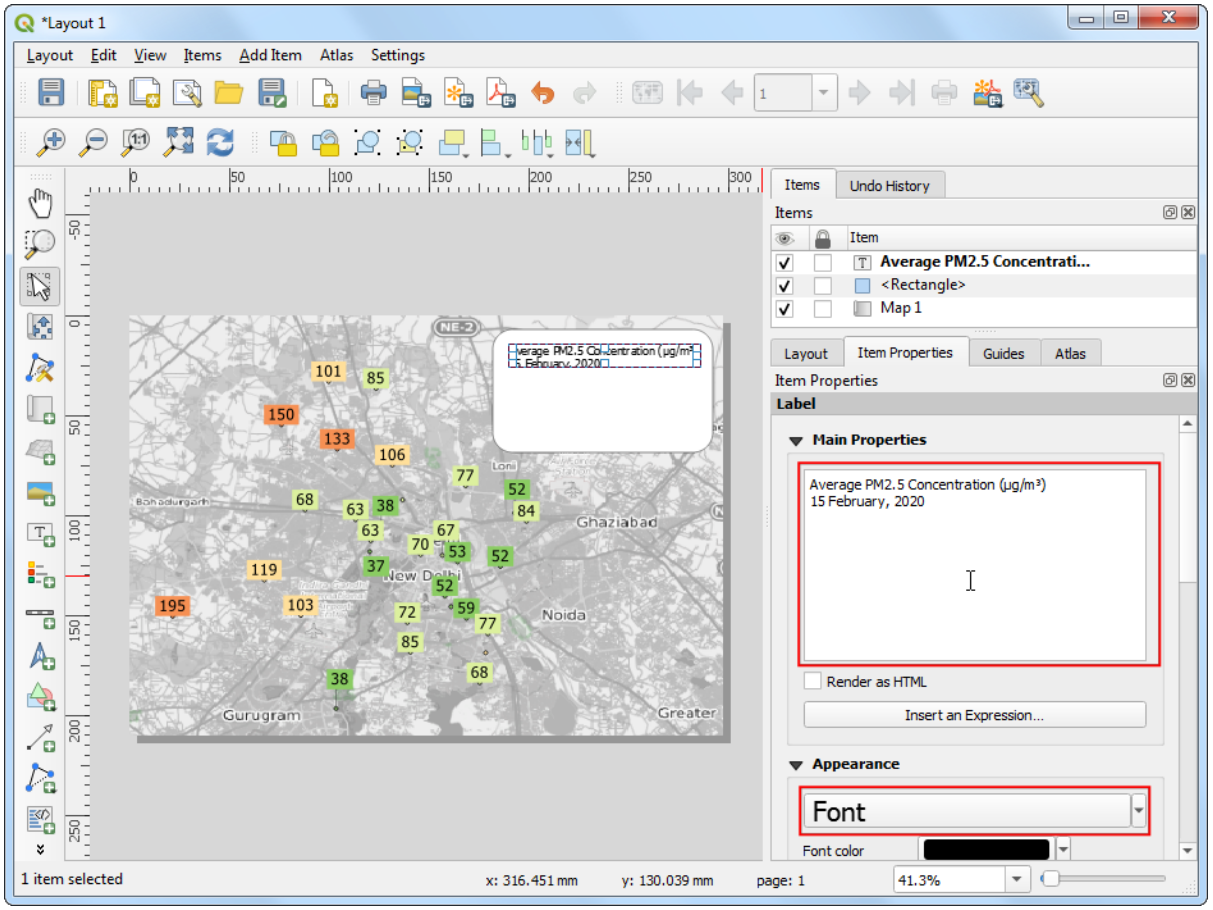
Next, we will add a *Rectangle* which will hold the tile, legend and attribution.



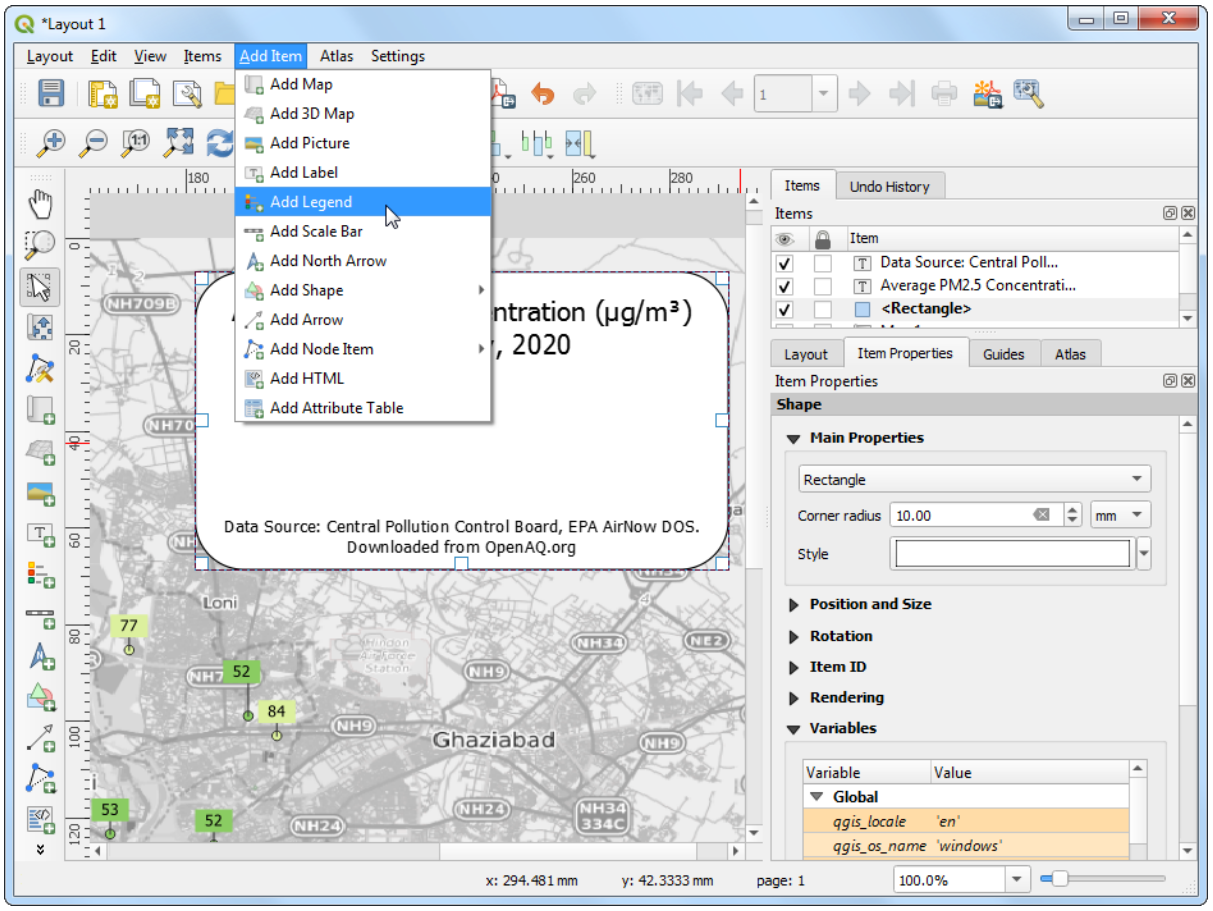
Place the rectangle on the top-right hand corner and change the *Corner radius* to `10`.



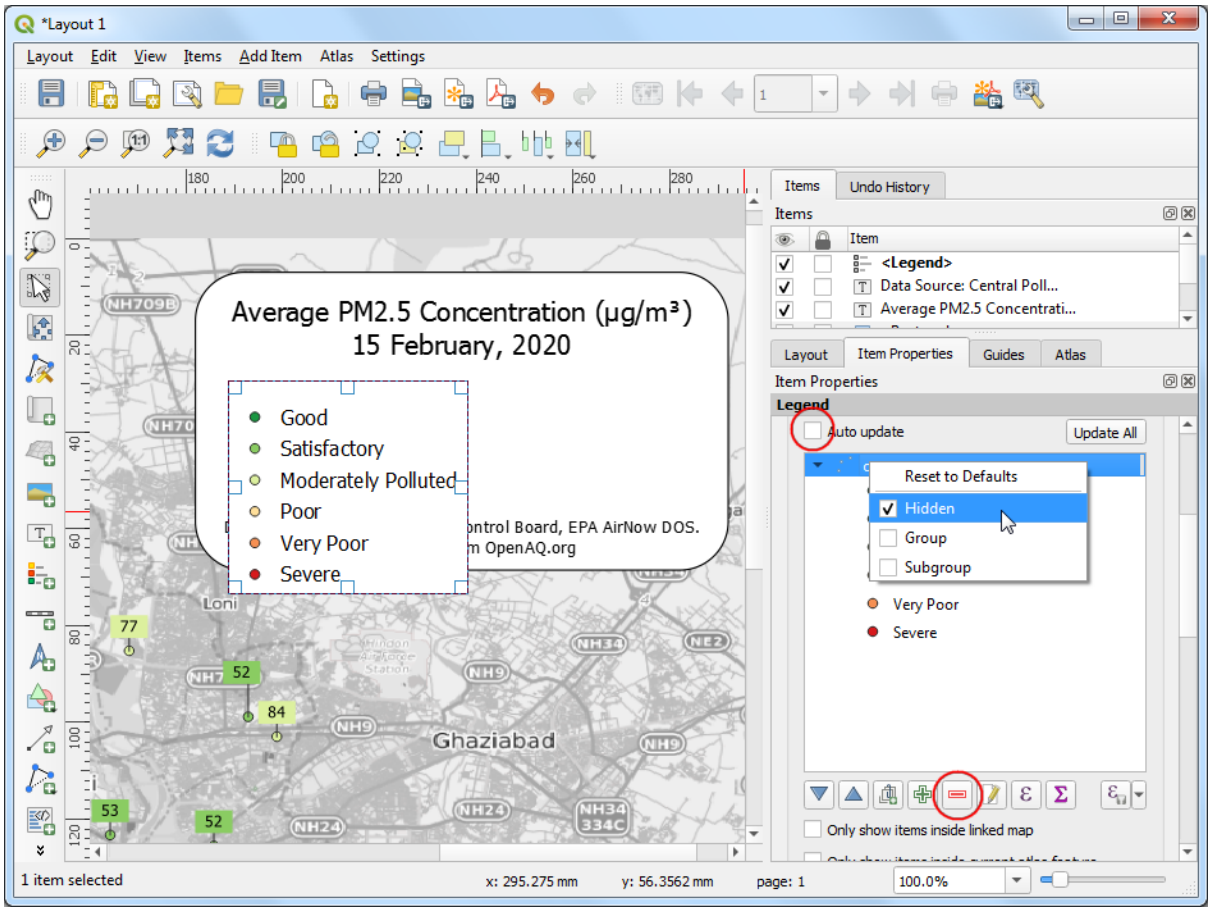
Next, add a *Label* with the title `Average PM2.5 Concentration (µg/m3)` and date `15 Bebruary, 2020`.



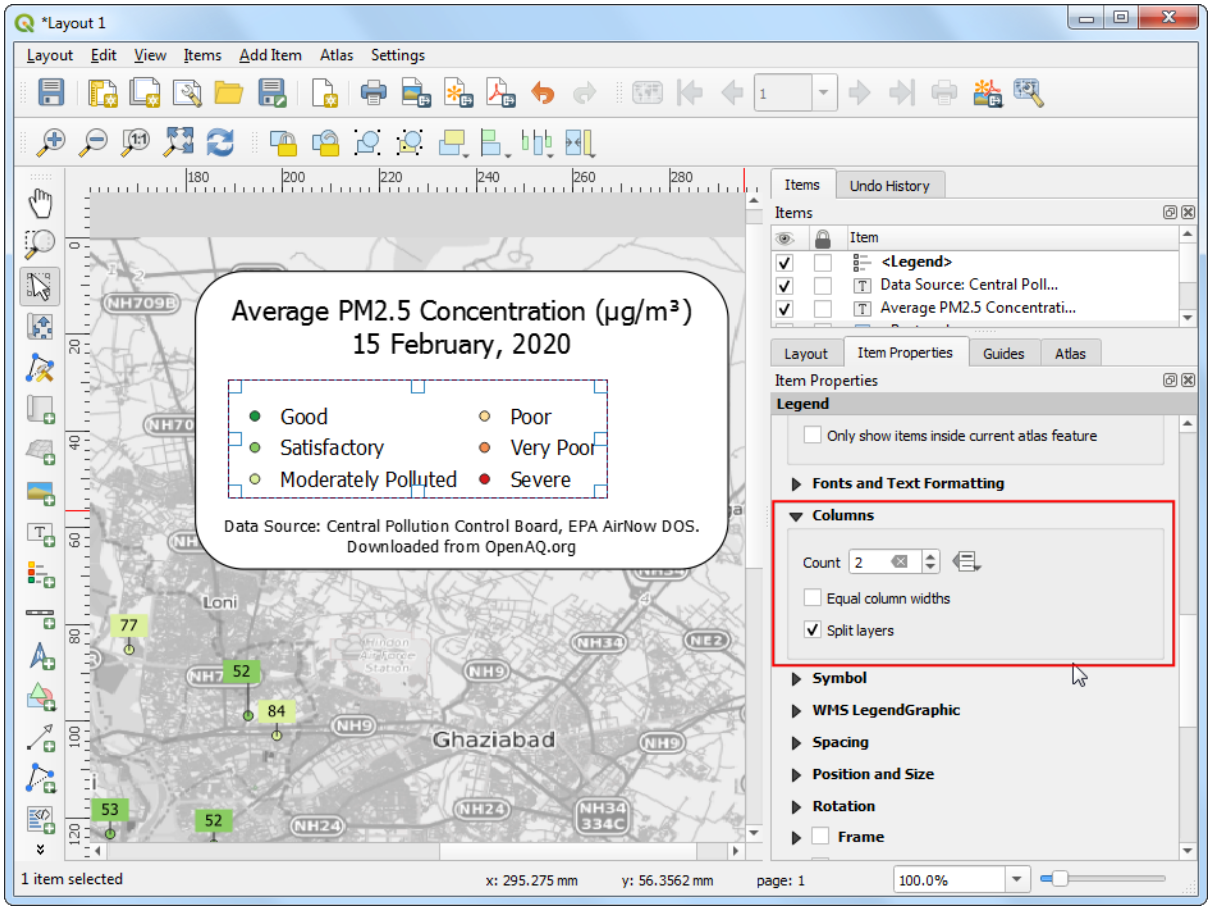
Add another *Label* with the attribution text `Data Source: Central Pollution Control Board, EPA AirNow DOS. Downloaded from OpenAQ.org`. Now we will add a legend, so our users know how to interpret various colors on the map. Go to **Add Item → Add Legend**



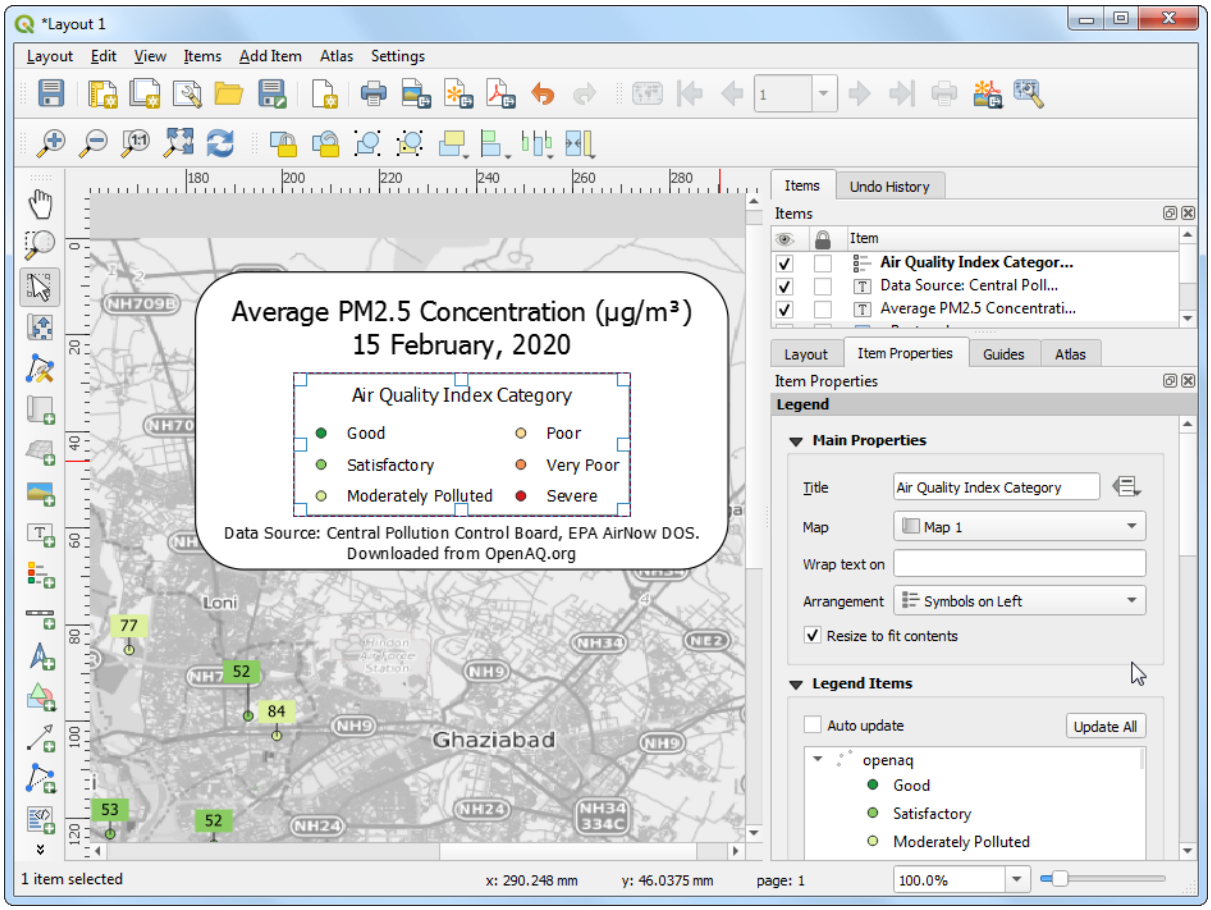
Drag a box to insert the legend. Go to the *Item Properties* tab and turn-off *Auto update*. Remove the `OpenStreetMap monochrome` layer and set the `openaq` group label to *Hidden*.



The default legend placement is vertical - in a single column. We can change it to be in multiple columns to make it span horizontal. Scroll down to the *Columns* section and change the *Count* to 2 and check *Split layers* button.

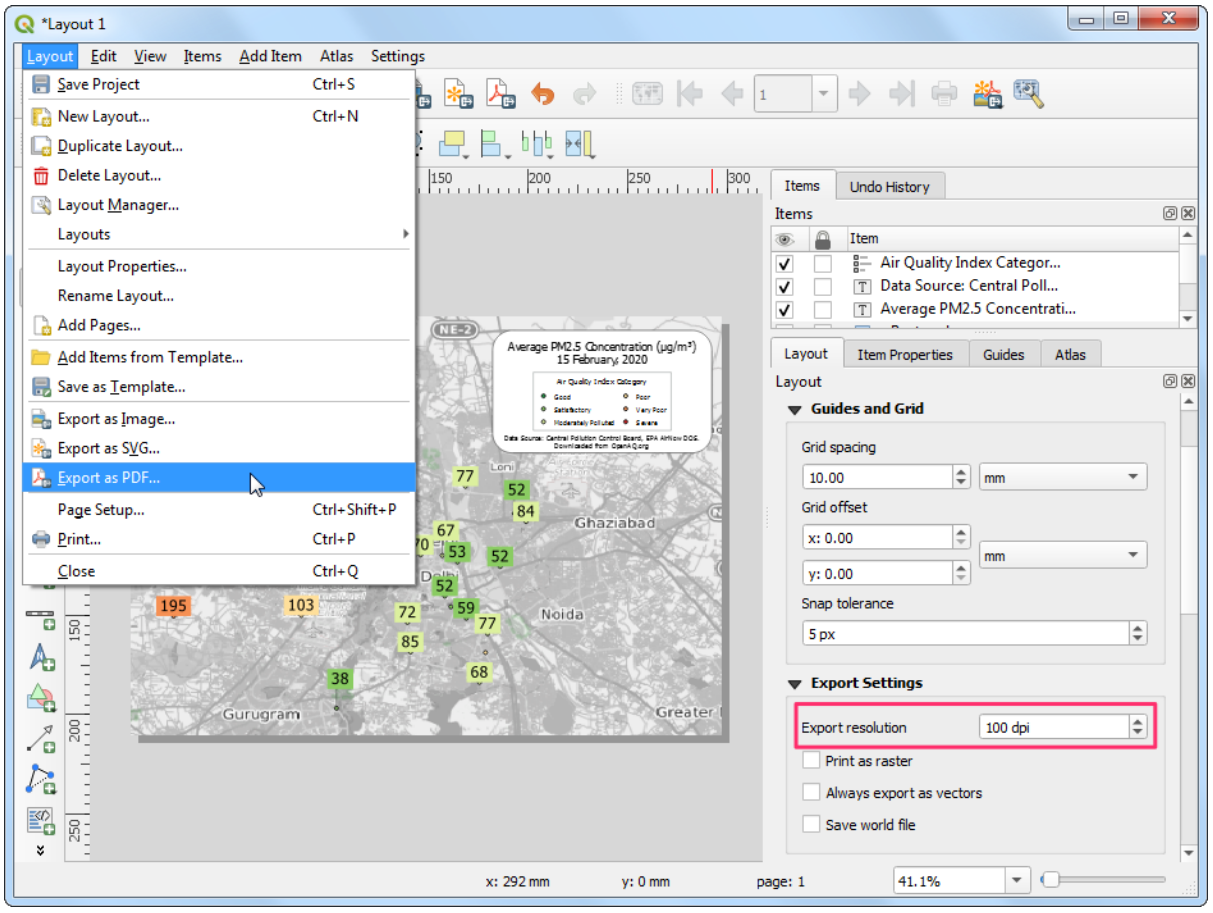


In the *Main Properties* section, enter a *Title* for the legend as `Air Quality Index Category`.

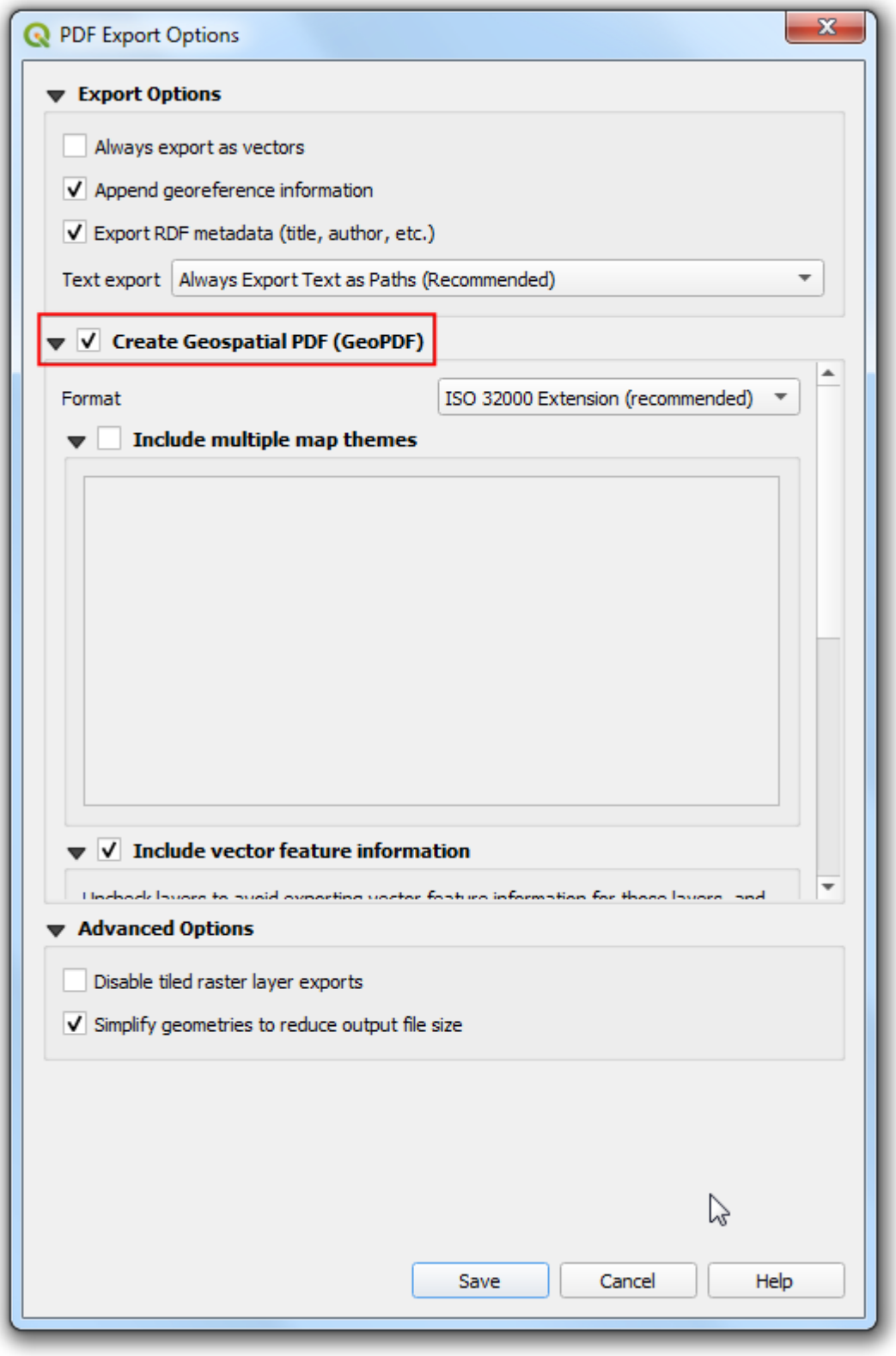


Once you are satisfied with the map, we can export the map. You can export the composition as an **Image** if you wanted to use it in a website, email, slideshow etc. You can also export it to as an **SVG** so it can be further edited in a graphics program such as InkScape. But the most common format for maps is still a PDF.

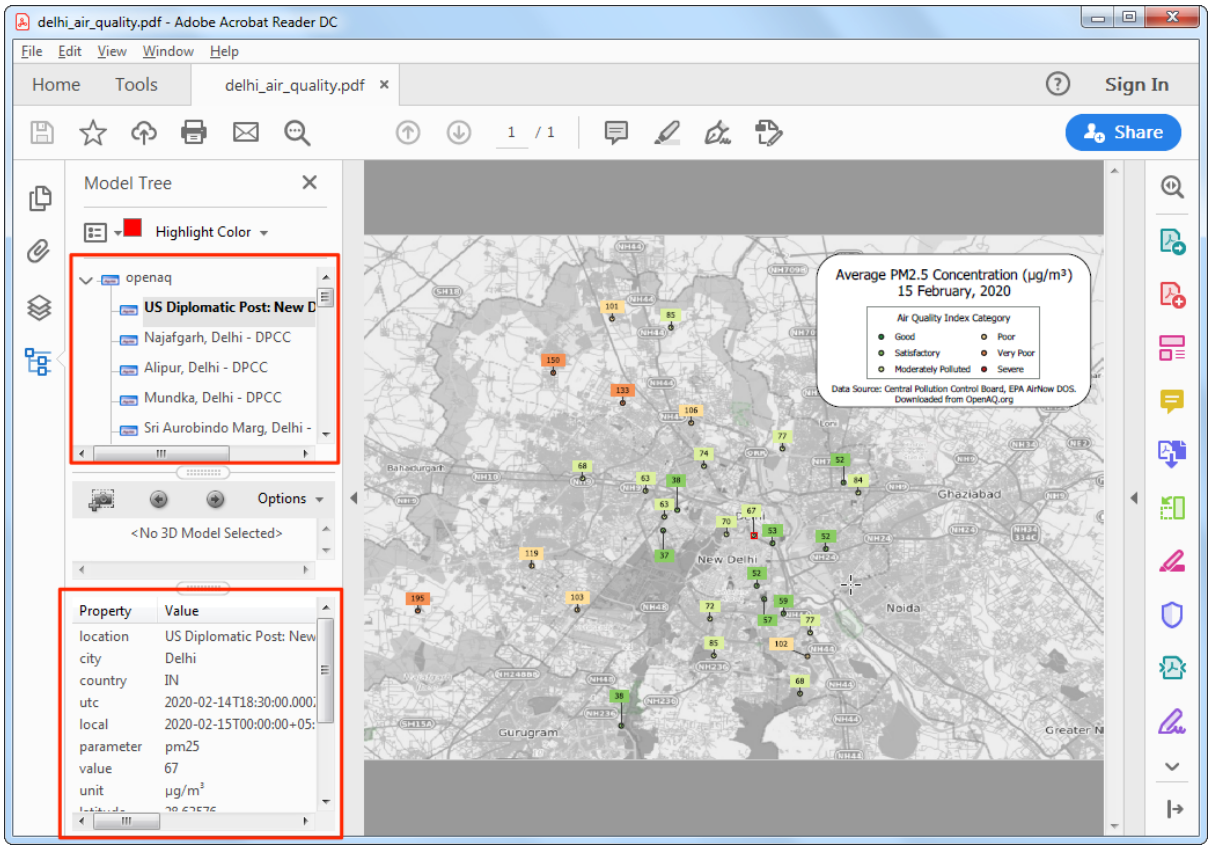
So let's export our map to a PDF. Before exporting, switch to the *Layout* tab and set the *Export resolution* to 100 dpi. Go to **Layout → Export as PDF**.



In the *PDF Export Options*, you can check *Create Geospatial PDF (GeoPDF)*. GeoPDF is an enhanced PDF format that is spatially aware and preserves layer and attribute information. Click *Save* and save the output as `delhi_air_quality.pdf`.



If you open the resulting PDF in a compatible reader such as Adobe Acrobat, you can toggle layer visibility, query attributes by features, measure distances and so on.



Data Credits

- Delhi PM2.5 concentrations. Downloaded from OpenAQ data download service. Source data from Central Pollution Control Board (CPCB) India and EPAAirNow DOC.

License