Fundamental Seminar
Analyzing Spatial Data with QGIS

Azarel Chamorro Obra
Salsabila Panji Arum

福田研究室
基礎ゼミナール

27 of April, 2018

Outline

- Processing GIS data
- Importing layers
- Setting CRS up
- Using plugins
- Aggregating data
- Attributes operation
- Filtering data
- Generating plots
- Exporting data

What is QGIS?

QGIS is a free geographic information system (GIS) software to view and analyze geospatial data.

The type of data QGIS processes is **geospatial/spatial data**.

Spatial data is the data equipped with particular **coordinate location**.

Source: www.arch.titech.ac.jp

東京工業大学 緑が丘1号館

〒152-8552 Tōkyō-to, Meguro-ku, Ookayama, 2 Chome−12−1 M1 東京工業大学緑が丘1号館

35.608755, 139.678955

(X,Y)

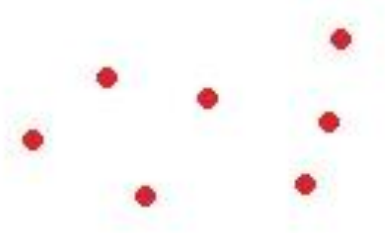
coordinate location

Benefits:

- See the **context of location** of our data in detail
- **Analyze the relationship** between our data and the surrounding environment

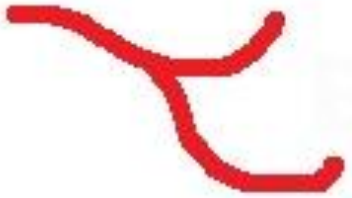
Types of Data

(1) Vector Data



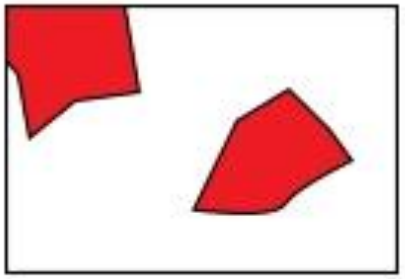
Point

Object is represented by a point



Line

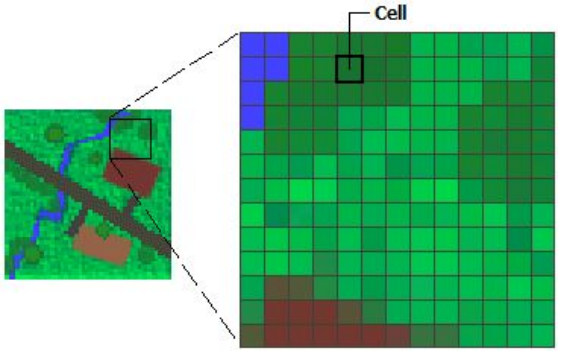
ex: railways, roads, river, etc.



Polygon

ex: buildings, prefecture, country

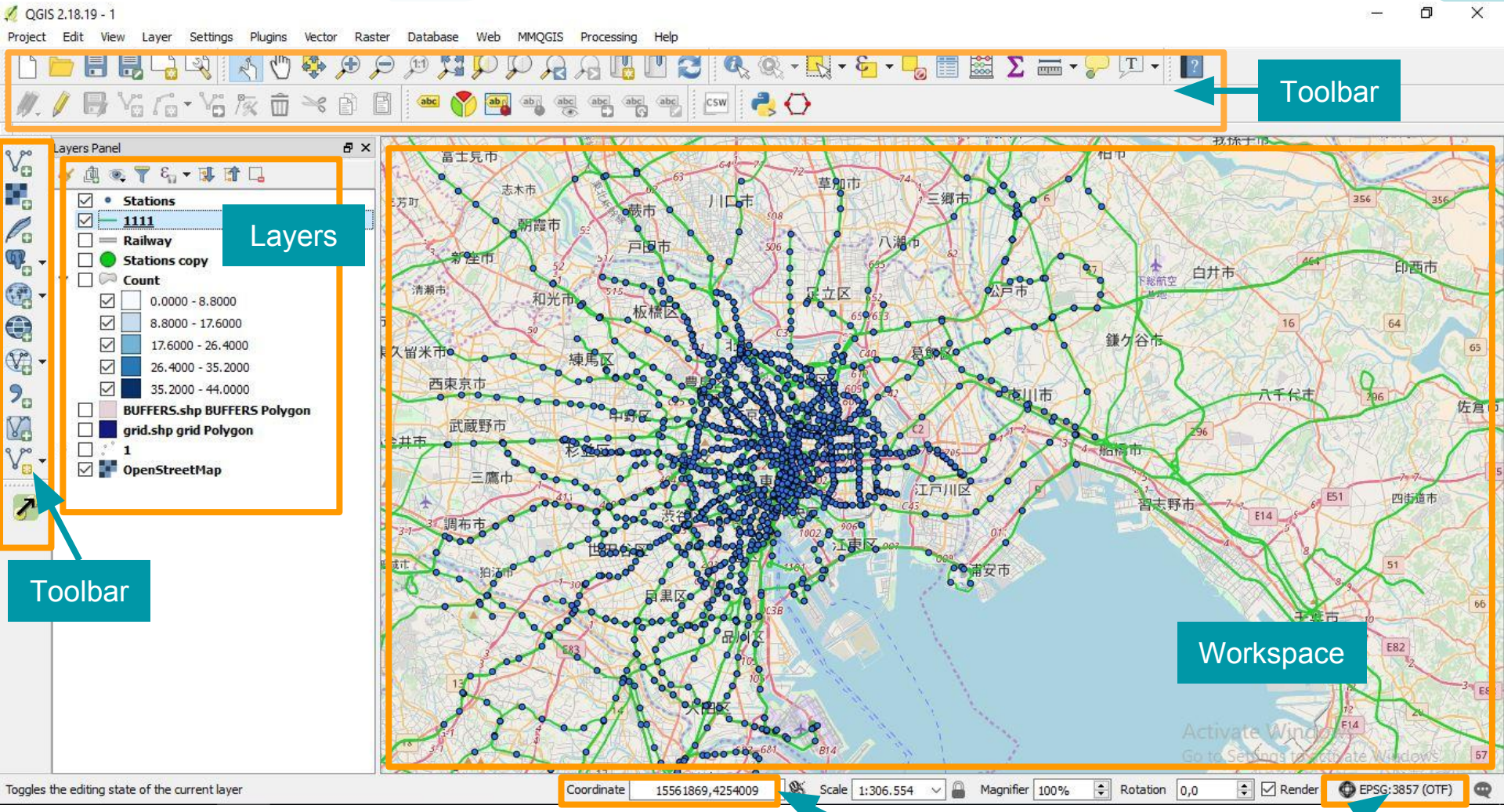
(2) Raster Data



Represents the world as surface divided by several grids of cells (rows and columns)

The smaller size of the cells we have, the higher accuracy of the data

User interface

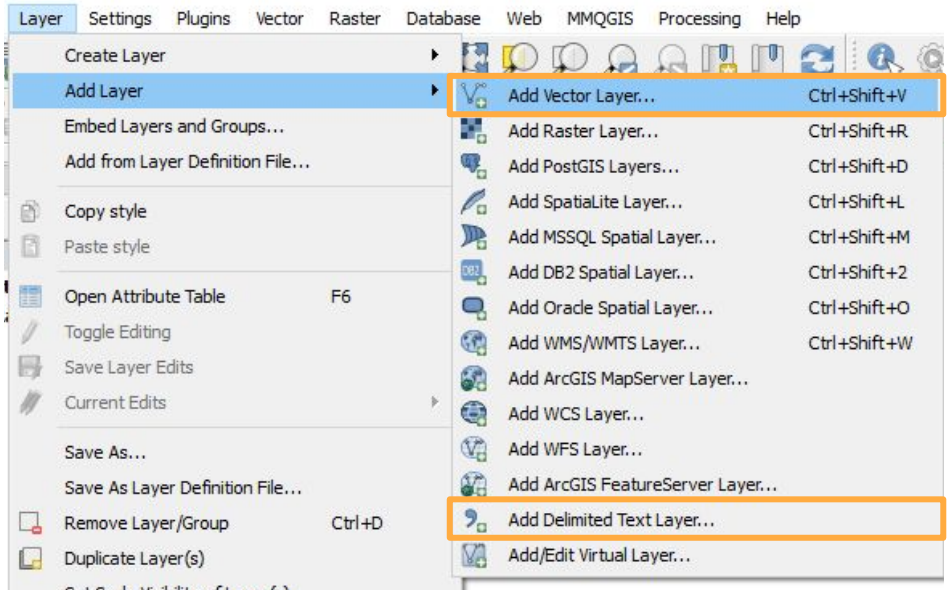


Importing Layers

Data from MLIT

Importing vector file in shapefile (.shp)

(1) Layer → Add layer → **Add Vector Layer**
(2) **Browse** the .shp file (Railway.shp)
(3) Set **encoding** to **Shift_JIS** for data with Japanese characters
(4) Click **Open**



Importing vector data from text file (.csv)

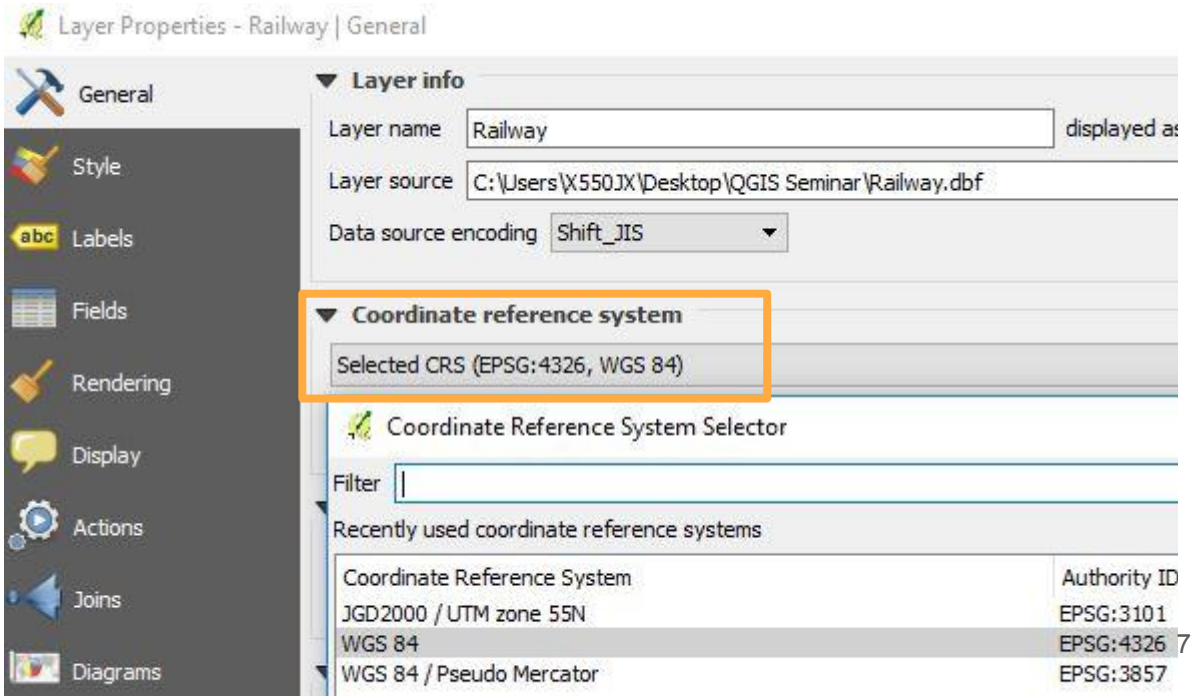
(1) Layer → Add layer → **Add Delimited Text Layer**
(2) **Browse** the .csv file (Stations.csv)
(3) Set **File Format** to **CSV**
(4) Choose the **encoding** (UTF-8)
(5) Set **Geometry Definition** as **Point Coordinates**
X field: x
Y field: y
(6) Click **Open**
(7) Set **CRS** (Coordinate Reference System) to **WGS 84**

Coordinate Reference System (CRS)

- We can specify the coordinate for each layer to best represent their location in the real world.
- Generally, *data source will specify which CRS shall be used for each data* (but we always have to double check). However, when there are no specification in the CRS, it is best to choose WGS 84 since it is compatible with most cases.

We can also change CRS after loading it to the worksheet by:

Right click on the Layer -> Properties -> General -> Coordinate Reference System

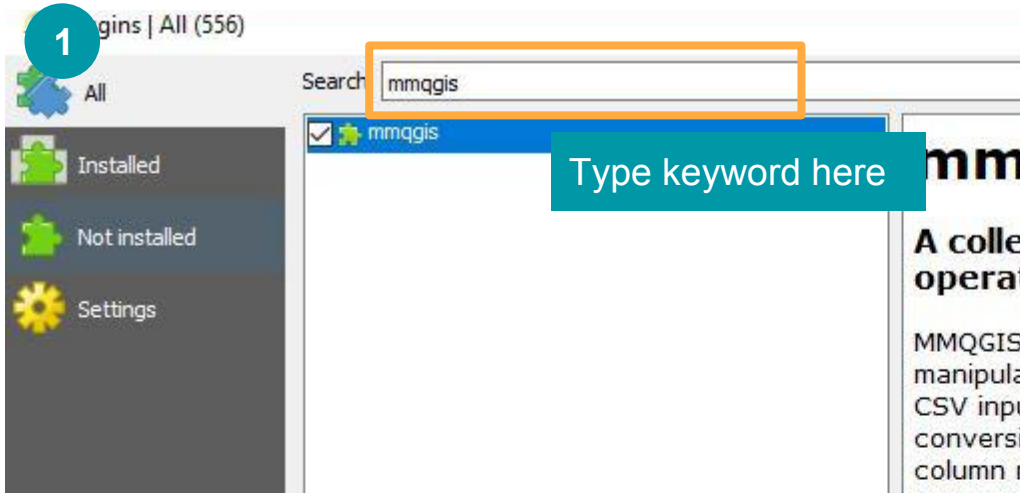


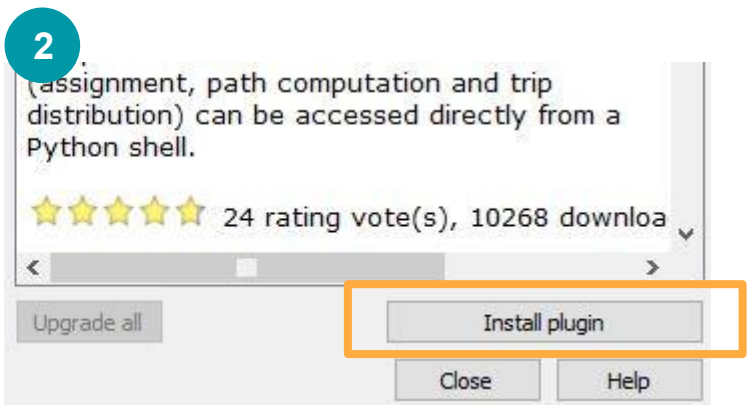
Installing Plugin

- Plugin is a tool for conducting data analysis that is not available on the QGIS interface
- Plugin is similar to ‘package’ in R and Python
- Since QGIS is mostly based on graphical user interface (GUI), most of the time, we do not have to use code.

Let’s install MMQGIS and OpenLayers

Plugins → Manage and Install Plugins → *Type on the keyword* → *Select* → Install Plugin

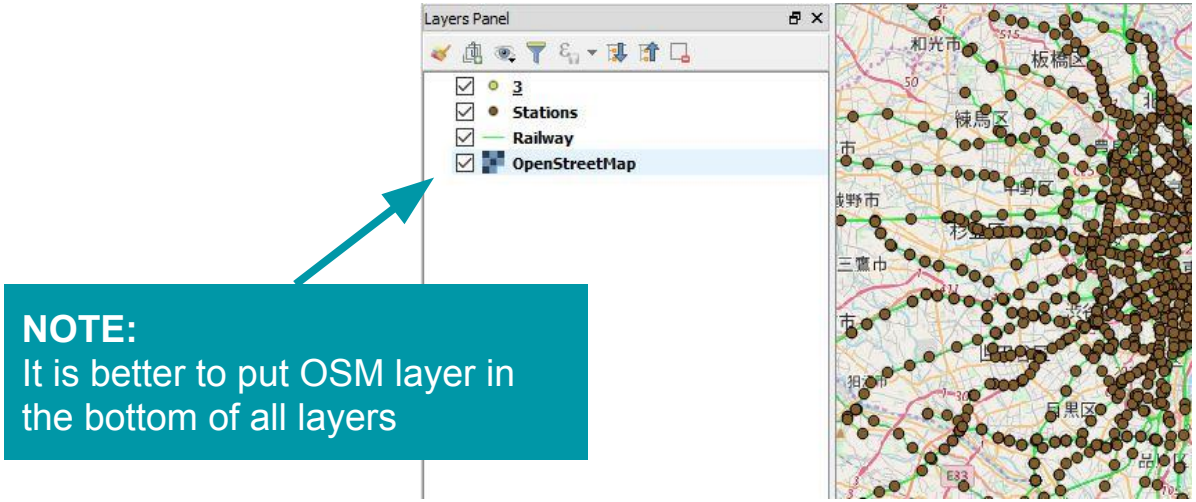




OpenStreetMap

Web → OpenLayers plugin → OpenStreetMap → OpenStreetMap

- OpenStreetMap may be added as a base layer in the working space.
- With OSM, we can visualize the location of our data better, since the map represents real world with roads, buildings, green space, water area (similar to Google Maps)
- It is possible to download data from OSM such as the road network

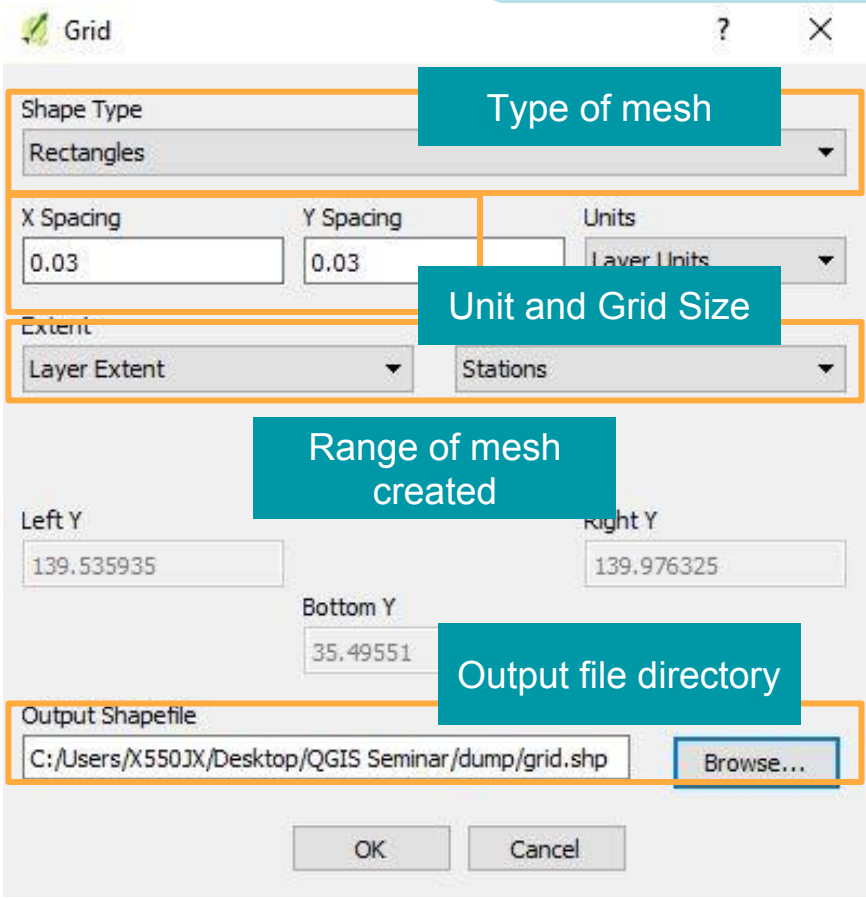


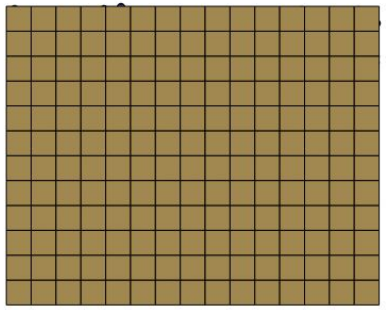
MMQGIS - Creating a Mesh

MMQGIS → Create → Create Grid Layer

- MMQGIS is a useful tool to do make mesh for aggregation. This time, we are going to make a mesh grid from aggregated ‘Stations’ data.

(1) Shape type: **rectangle** for mesh
(2) Specify **unit**: 0.01 and 0.01 in layer unit
(3) Extent: **layer extent**
(4) Layer: Stations
(5) **Browse** the directory of output file

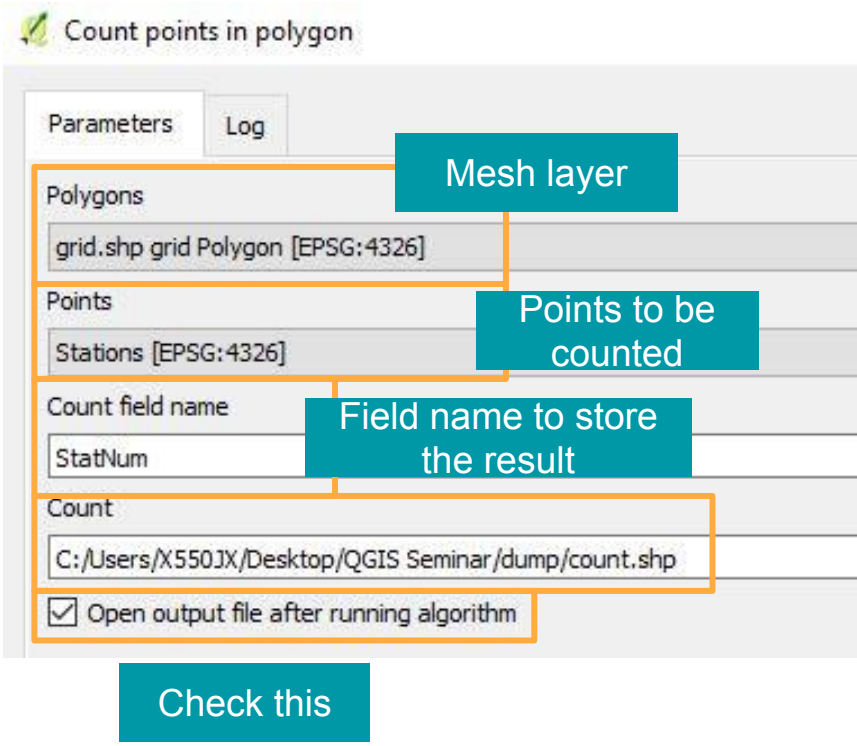




Data Aggregation

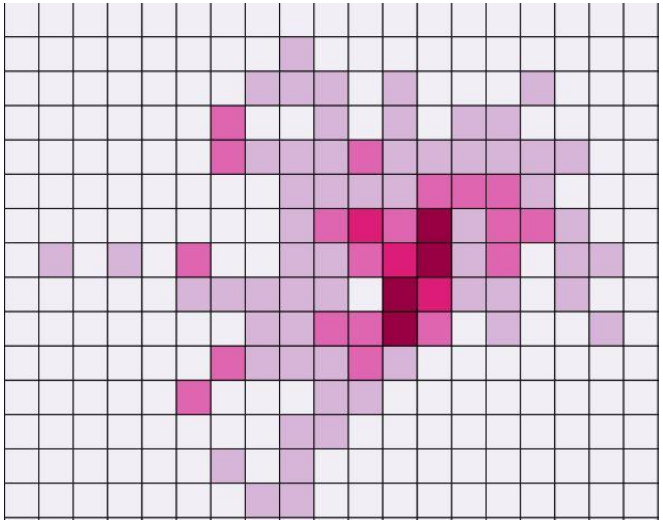
Now that we have the mesh and points of station, we can do a data aggregation analysis. For this exercise, we are going to aggregate the number of stations in every mesh.

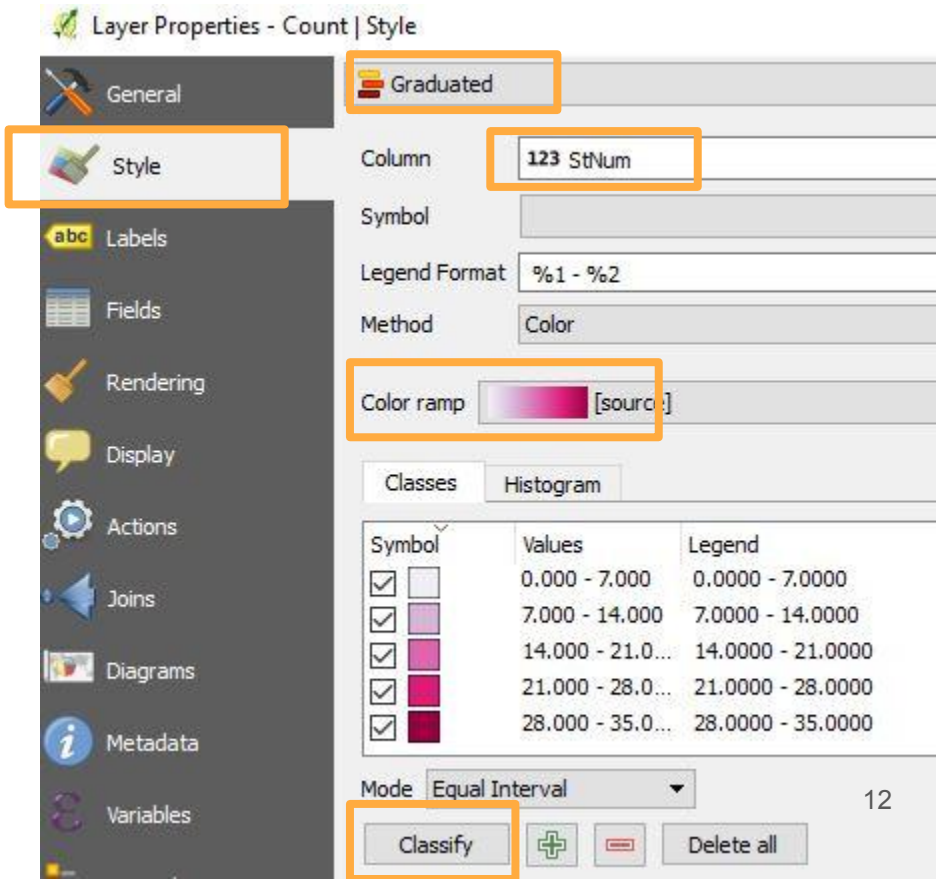
(1) **Vector → Analysis tools → count points in polygon**
(2) Set the **polygon layer** as the mesh layer we have,
point layer to be counted as 'Stations'
(3) **Name the count field** as **'StatNum'**
(4) **Browse** the directory for output file
(5) Click **Run**



Data visualisation

(1) **Right click** on our new mesh layer → **Properties**
(2) Open **Style**
(3) Change the **single symbol** to **graduated**
(4) Pick the **column name** as '**StatNum**'
(5) In the **Color Ramps**, choose your favorite color!
(6) Click on **Classify**
(7) Click **OK**

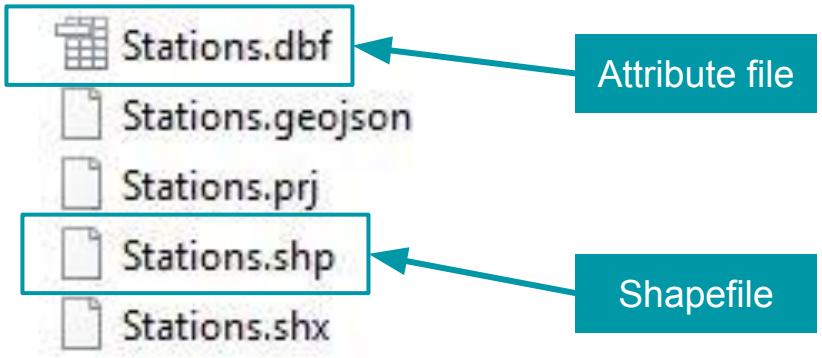




Analysing Attributes

Right click on the layer → Select "Open Attribute Table"

- Most of spatial data are equipped with "attributes''
- Spatial data sometimes comes with several other files. Attribute data is stored in .csv or database files (access, SQL….)
- Each row of the attribute represents the information of ONE point/line/polygon (of vector data) or cell (of raster data) in the selected layer



	N05_001	N05_002	N05_003	N05_004	N05_005b
1	1	東北新幹線	東日本旅客鉄道(	1982	2010
2	1	東北新幹線	東日本旅客鉄		02
3	1	東北新幹線	東日本旅客鉄道(...	1982	1991
4	1	東北新幹線	東日本旅客鉄道(...	1982	1985
5	1	東北新幹線	東日本旅客鉄道(...	1982	1982
6	1	上越新幹線	東日本旅客鉄道(...	1982	1982
7	1	北陸新幹線	東日本旅客鉄道(...	1997	1997
8	1	九州新幹線	九州旅客鉄道(旧...	2004	2011
9	1			004	2004
10	1			964	1964

Field name

Information of point #8

Attributes - Filtering

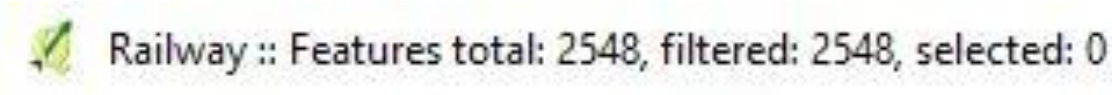
We can filter to select or remove attributes based on our conditions.

Select the **Railway** layer.

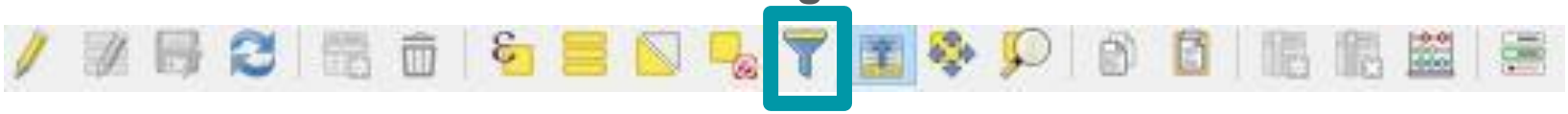
- In this example, we want to remove the old railway data which operation has been closed down.
- According to the data source, **existing railway is denoted by '9999'**. Therefore, we only want the value of '9999' in our data. This information is contained in the 'CLOSING' field of 'Railway' data attribute.

Attributes - Filtering

(1) Open attribute table
See on top of the window about the information of unfiltered data



(2) Click on the **select/filter features using form**



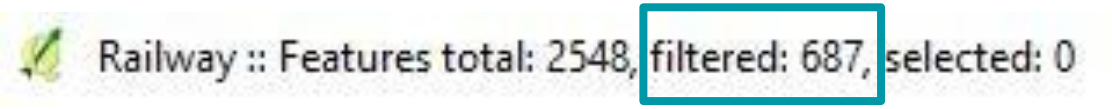
(3) In the '**CLOSING**', type **9999**

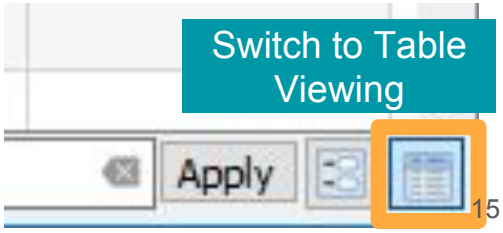
(4) Make sure the right option box is set to '**Contains**'

(5) On the bottom right, click '**Filter feature**'

(6) Again, on the bottom right, click '**Apply**' on the filter expression

(7) Your data has been successfully filtered!





Attributes - Remove Unfiltered Data

(1) Use **select/filter features using form** again in the attribute table

(2) In the '**CLOSING**', type **9999**

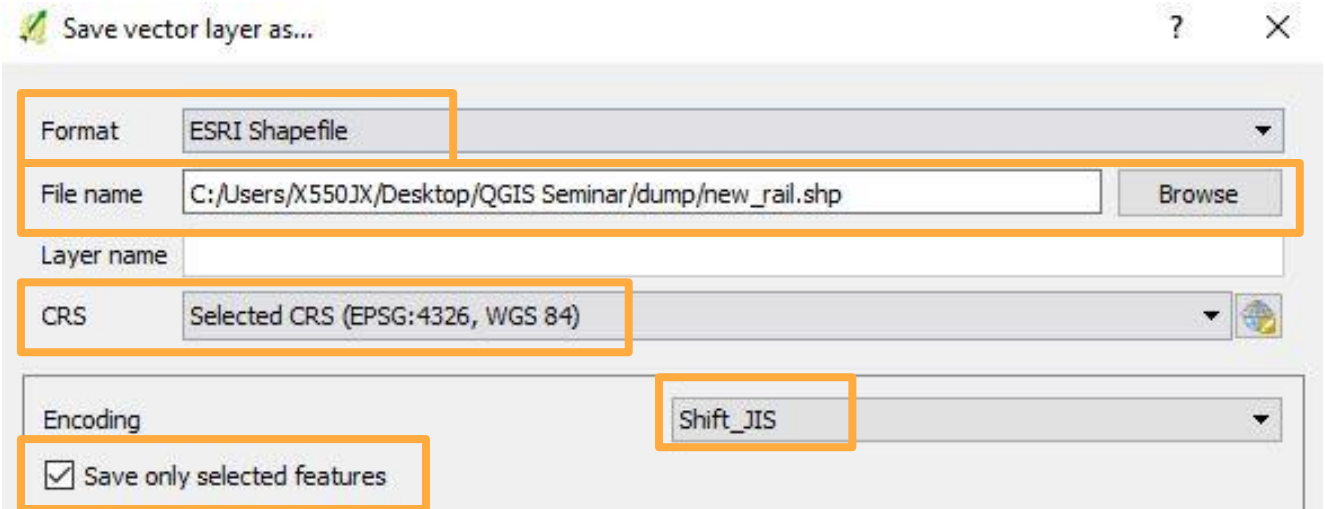
(3) Make sure the right option box is set to '**Contains**'

(4) On the bottom right, click '**SELECT feature**'

(5) Switch to table view, and go back to QGIS Workspace

(6) **Right click on the layer → save as**…

(7) Make sure the setting is right before clicking **OK**



Attributes - Adding New Field

We can add more variable to the attributes, either with our own conditions or based on another variable's condition.

(1) Open Attribute table of "Railway"

(2) Click on the **Open field calculator**

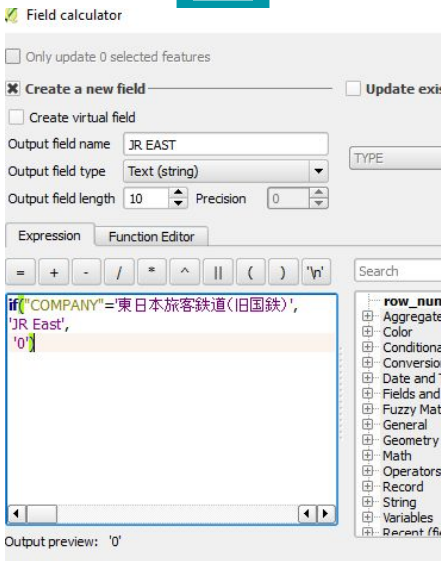
(3) Type the title of the new column on **Output Field Name**
We'll write down if a line belongs to JR EAST

(4) Specify the **output field type** as string and **length** as 10

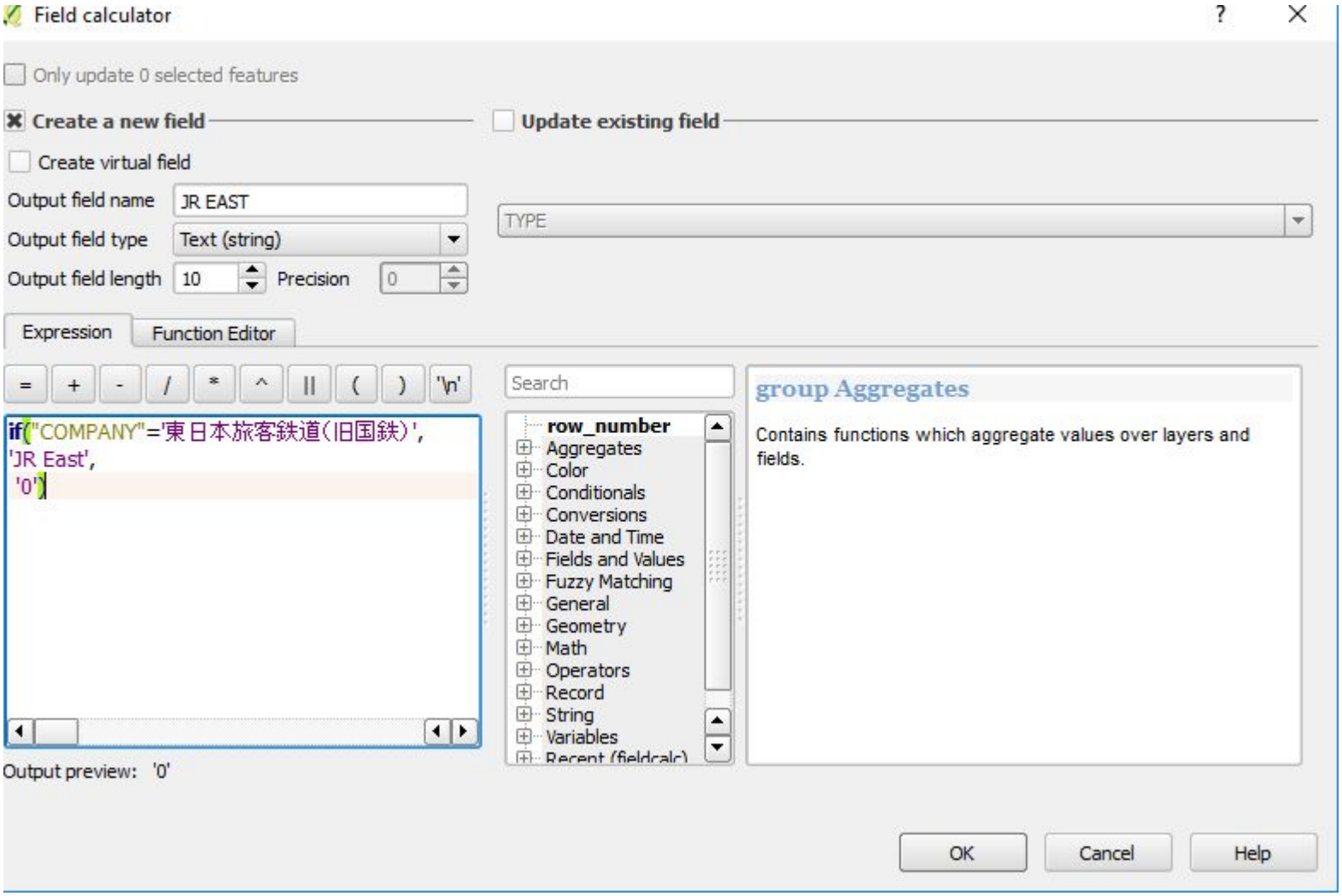
(5) Type:

if("COMPANY"='東日本旅客鉄道(旧国鉄)','JR East', '0')

(6) Click **OK**

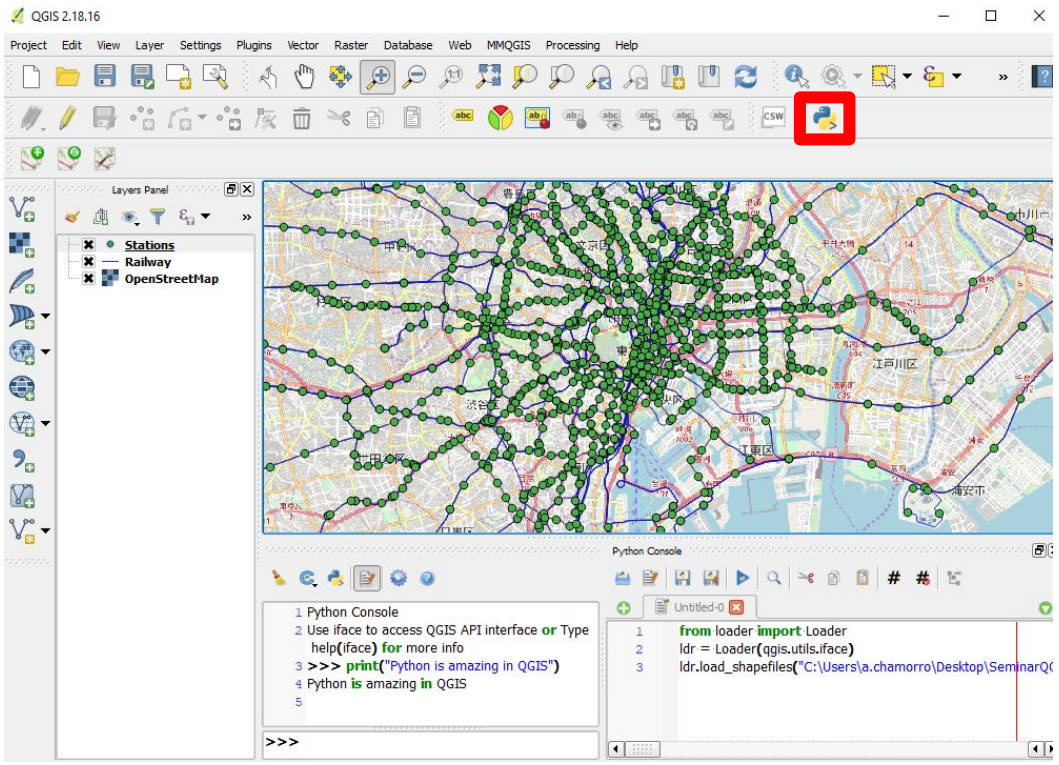


Attributes - Adding New Field



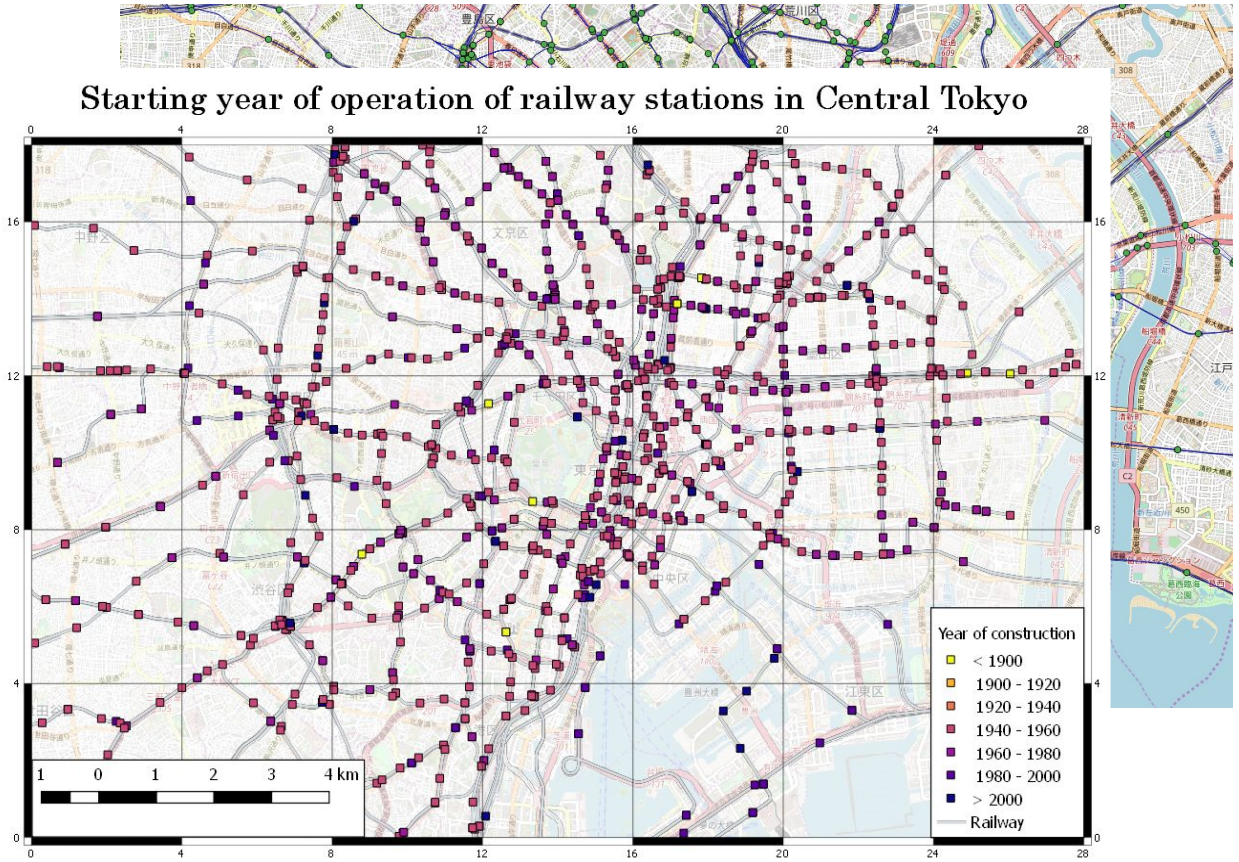
Automating QGIS

- Python scripts can be run in QGIS
- Useful when working with large sets of data
- To open Python console: Plugin→Python console
- For more information check QGIS documentation



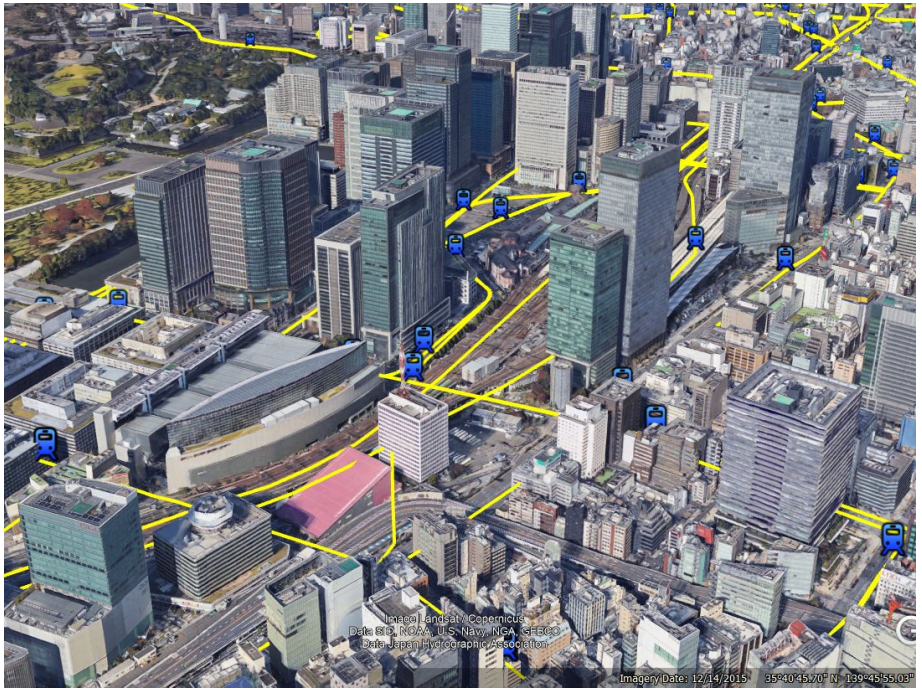
Printing out maps: print composer

- Plots are professional visualisations of the maps
- Useful when printing out showing results
- Plots are generated thro “print composer”:
  Project → New Print Comp
- You can add features su maps, legend, grid, etc
- ALWAYS draw the scale



Exporting maps to KML/KMZ

- Useful to send interactive layers to third parties (eg. client, reviewer)
- GoogleEarth and Maps allows people to see our GIS data
- Both software use .KML and .KMZ files
- To export a layer as .kml: “save as” -> Format =”Keyhole Markup Language”



Summary

- Today we learnt
  - Processing GIS data
  - Importing layers
  - Setting CRS up
  - Using plugins
  - Analysing and creating attributes
  - Filtering data
  - Generating plots
  - Exporting data

For more information check:

FINAL EXERCISE

- Create a layer of the decommissioned Japanese rail network by
- 15 periods:

  - Before 1951 (Imaoka)
  - 1951-1955 (Hirabayashi)
  - 1956-1960 (Ihoroi)
  - 1961-1965 (Suzuki)
  - 1966-1970 (Shiroma)
  - 1971-1975 (Kaneko)
  - 1976-1980 (Ogawa)
  - 1981-1985 (Muro)
  - 1986-1990 (Kawai)
  - 1991-1995 (Kita)
  - 1996-1999 (Koizumi)
  - After 2000 (Koike)
  - Still in operation (Nagasaki)

- Upload your layer to the GoogleMaps share folder:

https://drive.google.com/open?id=1vKSMpxqtDFE6vOWKzjQn-6QViSRohD6b&usp=sharing