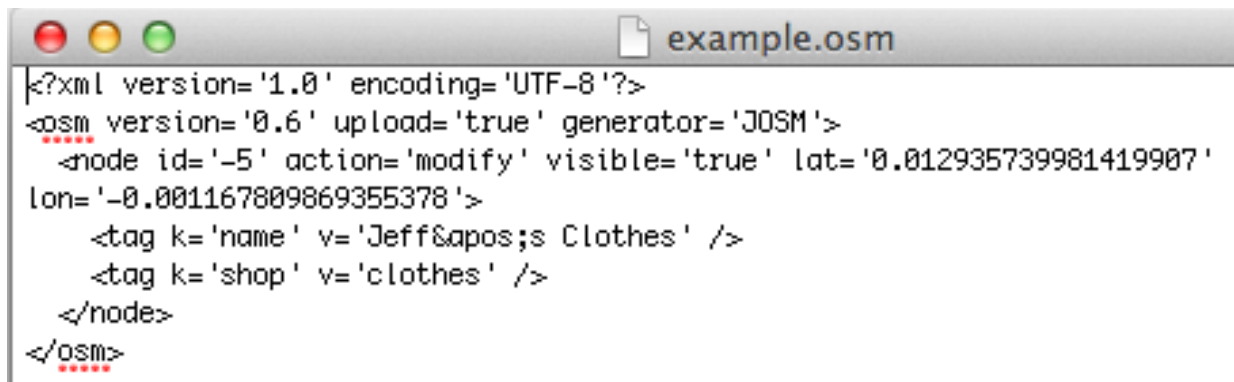# LearnOSM

File Formats

Reviewed 2016-09-05

Like any type of data, there are various ways of storing geographic data on a computer. It can be saved in a database, which is a specialized system for storing and retrieving data, and in fact there are database systems specifically designed for storing geographic data. It can also be stored in traditional computer files, though there are many different file formats for geographic data.

In this section we'll go through a few ways of storing geographic data, explain how they work and how they're typically used.

.OSM Files

The **.osm** file format is specific to OpenStreetMap. You won't come across it elsewhere. If you have ever downloaded data using JOSM and saved it as a file, you may have noticed that the file is saved with the extension **.osm**. If you are a GIS user, you may also have noticed that it is not easy to open these files using software such as QGIS.

So why is OSM data stored in a file format that nobody else uses? The answer is that many geographic data formats predate the modern internet era, and are designed for quick access and querying like one would query a database. OSM data, on the other hand is designed to be easily sent and received across the internet in a standard format. Hence, **.osm** files are coded in XML, and contain geographic data in a structured, ordered format. A simple **.osm** file would look like this if viewed in a text editor:

Acquiring data in **.osm** format is easy - in fact you do it every time that you download data in JOSM, but using these files for analysis and map design is not easy. Hence you are better off converting the data into another format, or getting it from a service that converts the data for you.

Raw OSM data is stored in **.osm** files usually, but you may also see files ending in **.bz2** and **.pbf**. These are essentially **.osm** files that have been compressed to save space, which can be extremely helpful when working with large data files.
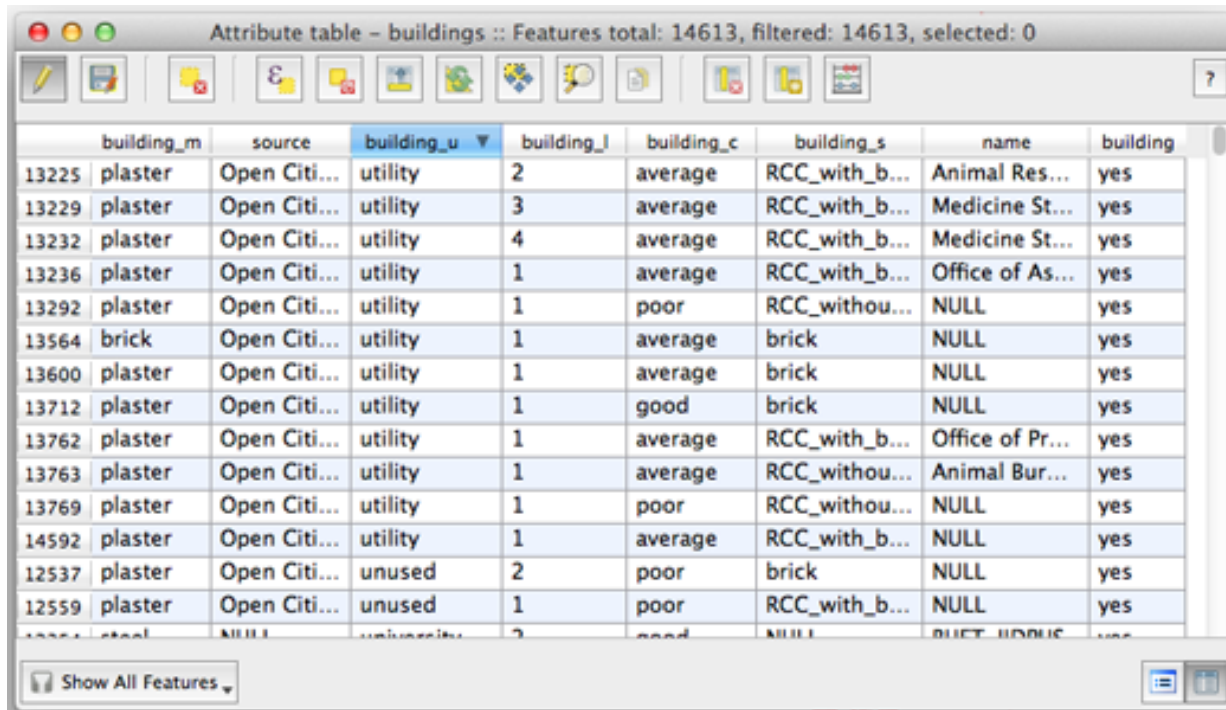
Shapefiles

The **shapefile** is a widely used format for storing vector geographic data. It was developed by ESRI, the company that makes ArcGIS, a popular suite of GIS applications.

Shapefiles are actually a collection of several different files. For example, a shapefile that contains building data might have files with the following extensions:

- buildings.**shp**
- buildings.**shx**
- buildings.**dbf**

Shapefiles will often have additional files too which contain other information.

A shapefile must be designated to hold only one type of feature (points, lines, or polygons), and each feature has it's attributes contained in a table. Unlike the OpenStreetMap system in which every object can have an unlimited number of tags, the attributes of features in a shapefile must fit into the shapefile's defined table structure, which might look something like this:

	building_m	source	building_u ▼	building_l	building_c	building_s	name	building
13225	plaster	Open Citi...	utility	2	average	RCC_with_b...	Animal Res...	yes
13229	plaster	Open Citi...	utility	3	average	RCC_with_b...	Medicine St...	yes
13232	plaster	Open Citi...	utility	4	average	RCC_with_b...	Medicine St...	yes
13236	plaster	Open Citi...	utility	1	average	RCC_with_b...	Office of As...	yes
13292	plaster	Open Citi...	utility	1	poor	RCC_withou...	NULL	yes
13564	brick	Open Citi...	utility	1	average	brick	NULL	yes
13600	plaster	Open Citi...	utility	1	average	brick	NULL	yes
13712	plaster	Open Citi...	utility	1	good	brick	NULL	yes
13762	plaster	Open Citi...	utility	1	average	RCC_with_b...	Office of Pr...	yes
13763	plaster	Open Citi...	utility	1	average	RCC_withou...	Animal Bur...	yes
13769	plaster	Open Citi...	utility	1	poor	RCC_withou...	NULL	yes
14592	plaster	Open Citi...	utility	1	average	RCC_with_b...	NULL	yes
12537	plaster	Open Citi...	unused	2	poor	brick	NULL	yes
12559	plaster	Open Citi...	unused	1	poor	RCC_with_b...	NULL	yes

Attribute table – buildings :: Features total: 14613, filtered: 14613, selected: 0

Show All Features

OpenStreetMap data can be converted into shapefiles. Various websites provide shapefiles converted from OSM data. These are discussed in the <u>next chapter</u>.

Databases

Many types of information are stored in database systems, which provide a logical way of organizing and accessing data. Geographic data is no different, although databases designed for geodata are specialized to handle the complex functions that querying geographic data requires.

OpenStreetMap data is often stored in a PostgreSQL database with PostGIS extensions. This type of database provides fast access to the data and can be used easily with Mapnik, a piece of software that creates the map tiles used in web slippy maps. There are several tools available for importing raw OSM data into a PostgreSQL database.

Another type of database is known as SQLite, which provides similar functionality as a PostgreSQL database, but is all stored in a single file and doesn't require database software to be running. These are a little more difficult to create yourself, but can be easier to work with for small sets of data.

Summary

In the following chapters we will see how you can download data in various formats from the internet, and how you can use various tools to manipulate the raw data on your own.

Was this chapter helpful? Let us know and help us improve the guides!

- learnosm@hotosm.org
- @learnOSM
- Hosted on Github

Official HOT OSM learning materials

Humanitarian
OpenStreetMap
Team