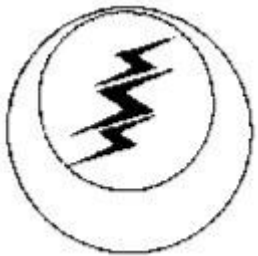
OSM Data Manipulation



Robert Savoye
Seneca Software & Solar, Inc.
rob@senecass.com
https://www.senecass.com

Resources We'll Use

- Download OSM data for Nepal
  http://download.geofabrik.de/asia/nepal-latest.osm.bz2
- Install Josm, osmconvert, osmfilter, gpsbabel
- Optionally install Qgis

Documentation

- http://wiki.openstreetmap.org/wiki/Osmfilter
- http://wiki.openstreetmap.org/wiki/Osmconvert
- https://www.gpsbabel.org

Primary Data Format Groups

- **OSM** (Open Street Map)
- **ODK** (Open Data Kit)
- **Output formats** (KML/KMZ, GPX, JSON, CSV)
- **Shapefiles**

Open Street Map Formats

- **osm** - Default XML based format
- **o5m** - High speed binary format
- **osc -** Change set for updating OSM data
- **pbf -** Very compressed binary format
- **poly -** Simple text file with a list of coordinates

Open Data Kit Formats

- **XLSForm**
  - Excel/LibreCalc form
  - http:/www.xlsform.org
- **XForm**
  - XLSForm converted to XForm for ODKCollect
  - Pyxform
  - http://build.opendatakit.org
- **ODK Instance**
  - XML output format from ODKCollect

Qgis can also export a layer to both XLSForm and XForm formats.

ODKCollect

- Android app
- Uses an XML format as a data input template, and a different output format.
  - Data input templates are built using the ODK builder or pyxform.
- Kathmandu Living Labs has derived versions that are more Nepal mapping task specific.
- Can use ODKAggregate to download forms and upload data files.
- Can also use ODKBriefcase to extract data using a laptop or desktop.
- Can also use the adb program, or mount the phone as a USB drive.

Shapefiles

- A binary format with arbitrary data fields that don’t map directly to OSM data fields.
- Commonly used by many public agencies.
- Can be read by Josm and Qgis.
- Often have many fields not relevant to OSM, so conversions need to be filtered.

Shapefile Sample

./shp2map.py -i ../MapData/Gilpin/RoadCenterlines/RoadCenterlines.shp --dump

Fields in: '../MapData/Gilpin/RoadCenterlines/RoadCenterlines.shp'

'DeletionFlag' 'centerline' 'fromleft' 'toleft' 'fromright' 'toright' 'fullname' 'altname' 'fedroute' 'fedrtetype' 'afedrte' 'afedrtetyp' 'stroute' 'strtetype' 'astrte' 'astrtetype' 'ctyroute' 'onewaydir' 'roadlevel' 'speedlimit' 'fromzelev' 'tozelev' 'inwater' 'roadclass' 'mtfcc' 'ownedby' 'maintby' 'munileft' 'muniright' 'coleft' 'coright' 'stateleft' 'stateright' 'zipleft' 'zipright' 'msagleft' 'msagright' 'esnleft' 'esnright' 'lastupdate' 'lasteditor' 'shape_Leng'

OSM Node

- A node represents a single location, and can be referenced by other application.

<node id="267680857" lat="28.1568909" lon="85.4147102" version="4" times
tamp="2010-11-11T22:56:01Z" changeset="6347737" uid="331348"
user="janedoe"/>

OSM Waypoint

<node id="632389857" lat="28.1522222" lon="85.3781199" version="5" timestamp="2014-02-24T23:02:05Z" changeset="20761187" uid="1772211" user="4rch">
<tag k="ele" v="1770"/>
<tag k="name" v="Landslide"/>
<tag k="name:np" v="pairo"/>
<tag k="tourism" v="guest_house"/>
<tag k="wpt_symbol" v="Lodging"/>
<tag k="wpt_description" v="Landslide lodge (1770m)"/>

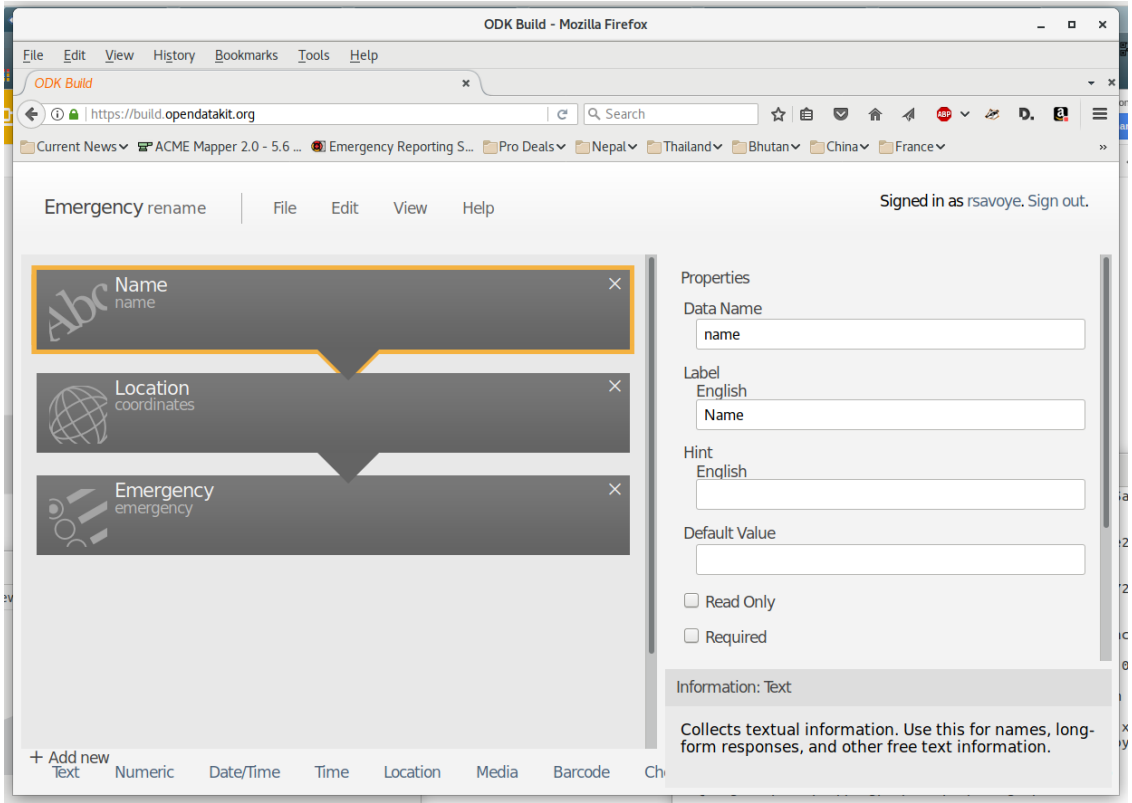
OSM Polygon or LineString

```
<way id="61427277" version="5" timestamp="2013-12-02T00:41:38Z"
changeset="19226127" uid="463033" user="synalik">
        <nd ref="632403048"/>
        <nd ref="766912443"/>
        <nd ref="766912342"/>
        <nd ref="766912416"/>
        <tag k="boundary" v="administrative"/>
```

The only difference between a Polygon and a LineString is Polygons are closed.

ODKBuild

- Web based design tool
- Creates XLSForm or XForm
- Requires OSM account



XLSForm

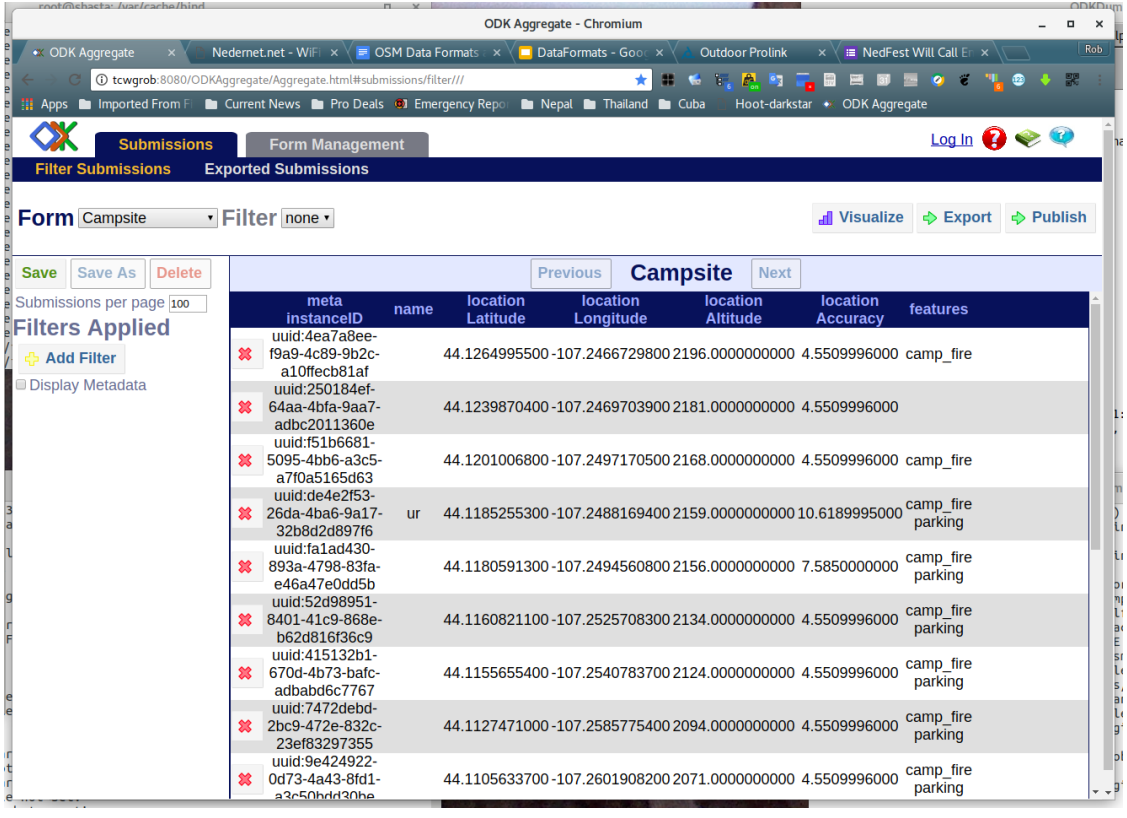
- Standard format used by LibreOffice, Google Sheets, and MS Office.
- Binary format
- Needs to be converted to XForm for use by ODKCollect

XForm

- XML format
- Can be edited using any text editor
- Template for ODKCollect
- Supports choices, multiple choices, labels, media files
- ODKAggregate can be used to store and download Xforms to ODKCollect based apps

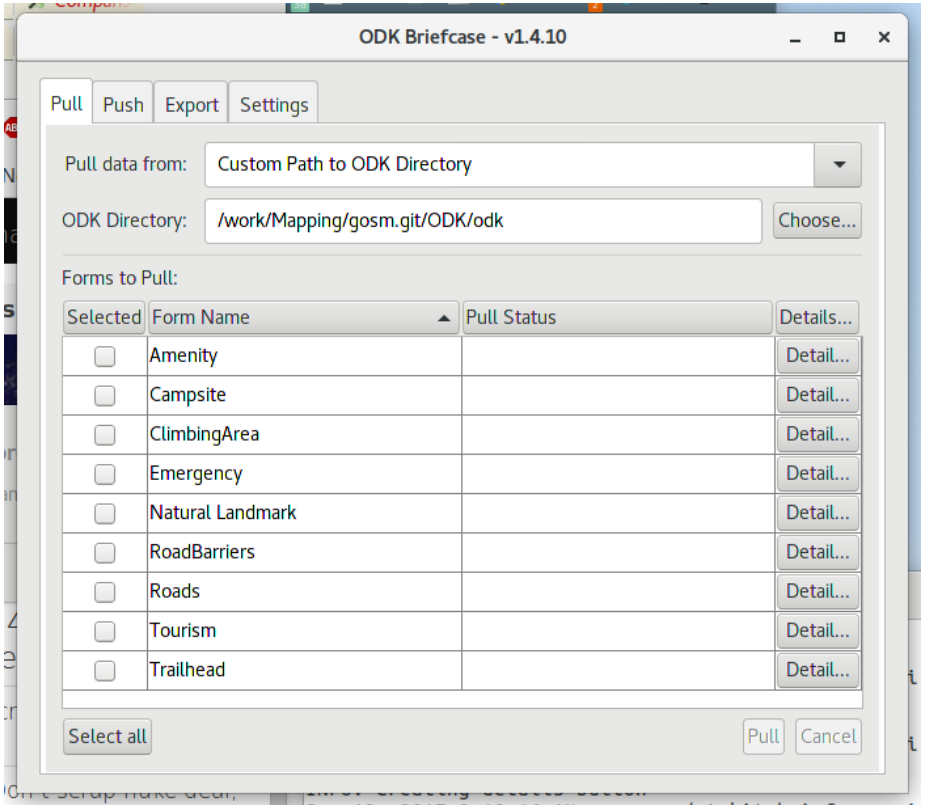
ODK Aggregate

- ODKCollect can download XForms for an input data template
- ODKCollect can upload data using a network connection
- Can export in data as KML, CSV, or JSON format
- Data is coupled to the version of the XForm



ODKBriefcase

- Reads data from mounted smartphone
- Reads data from local directory
- Exports CSV format
- Can also upload to ODKAggregate
- CVS file can be loaded into Josm



Desktop Tools

- **Qgis** - Powerful GIS editor
- **Josm** - Desktop OSM data editor
- **osmconvert** - Concerts between the various OSM formats and can also do filtering using polygons
- **osmosis** - Swiss army knife for OSM editing operations
- **osmfilter** - Can be used to filter OSM data based on tags

Android Tools

- **OSMAnd**
  - Primarily used for data viewing and navigation
- **Vespuci**
  - Full editor for OSM data
  - Can create and edit OSM data
  - Can download and upload OSM data
- **StreetComplete**
  - Used to find incomplete or wrong data

KML

- Created by Google for Maps and Earth
- Can be edited in a text editor

```
<Placemark>
    <name>Kharpani hot springs</name>
    <styleUrl>#HotSpring</styleUrl>
   <description>
      Osmid: 1959686217
   </description>
<Point><coordinates>83.958343088312887,28.360128796374827</coordinates>
</Point>
      </Placemark>
```

KMZ

- Binary zip format
- Contains a single KML file
- Also contains all custom icons

CSV

- Widely supported by many GIS applications
- Can be edited by any spreadsheet program
- Can be edited in a text editor

```
"name","latitude","longitude","altitude","accuracy","amenity"
"Porta potty 1","44.11682204","-107.25125132","2135.0","6.068","toilets"
"Porta potty 2","44.08421973","-107.30796647","1647.0","6.068","toilets"
"Sallys Cafe","44.03392162","-107.44992966","1347.0","7.585",""
"Ten Sleep Ranch","44.06935172","-107.34920497","1491.0","13.653","campground"
```

GPX

- Used by most GPS devices
- Widely supported by most all GIS applications
- Can be edited in a text editor

```
<wpt lat="28.360128796" lon="83.958343088">
    <name>Kharpani hot springs</name>
    <cmt>Osmid: 1959686217</cmt>
    <desc>Natural Hot Spring near Pokhara</desc>
  </wpt>
```

Filtering data`

- Most OSM data files are huge
- Data can be reduced by using bounding boxes or polygons using osmconvert or osmosis.
- Data can be reduced by searching for specific tags.
- Data can be reduced using Josm or QGIS

OSMConvert

- Can extract data using a bounding box or polygon
- Can produce an OSM change set by diffing OSM files
- Can upload OSM file into a database
- Can convert between OSM formats

OSMConvert Examples

osmconvert -B=test.poly tmp.osm --complete-ways -o=new.osm --drop-broken-refs

- **-B -** Input polygon
- **--complete-ways -** Include data that extends out of the polygon, like a road
- **--out-osm -** Output file in OSM format
- **--drop-broken-refs -** Drop refs outside of the boundary

OSMFilter

- Can filter based on tags
- Faster when using o5m file for input
- Produces OSM output file
- Can keep or delete based on the tags
- Can modify tags

OSMFilter Examples

osmfilter file.o5m --keep="surface=dirt highway=path bicycle=no" --out-osm

- **--keep -** Keep objects with the specified tags
- **--out-osm -** Output file in OSM format. Can also be osm or o5c

gpsbabel

- Can translate between many formats
- Used by Google Earth and Maps

gpsbabel -i kml -f AnnapurnaCA.kml -o gpx -F AnnapurnaCA.gpx

- **-i -** Input format
- **-f -** Input file
- **-o -** Output format
- **-F -** Output file

Examples

osmfilter file.o5m --keep="surface=dirt highway=path =footway bicycle=no" -o=out.osm

gpsbabel -i osm -f out.osm -o kml -F trails.kml

osmfilter file.o5m --keep=”historic=archaeological_site =building =ruins” -o=out.osm

gpsbabel -i osm -f out.osm -o gpx -F sites.gpx

Osmosis

- Can extract data using a bounding box or polygon
- Can produce an OSM change set by diffing OSM files
- Can upload OSM file into a database

Osmosis Examples

osmosis --read-xml file=colorado-latest.osm --write-apidb dbType="mysql" host="localhost:3306" validateSchemaVersion=no database="osm" user="rob" password="foobar"

osmosis --read-xml file=colorado.osm --tee 2 --bounding-polygon file=CO-polyfiles/Gilpin.poly -bounding-polygon file=CO-polyfiles/Boulder.poly --merge --write-xml file=bigger.osm

osmosis --read-xml file=colorado.osm --bounding-box top=38.96358 bottom=37.13613 left=-109.06578 right=-107.21458 --write-xml file=co-subset.osm

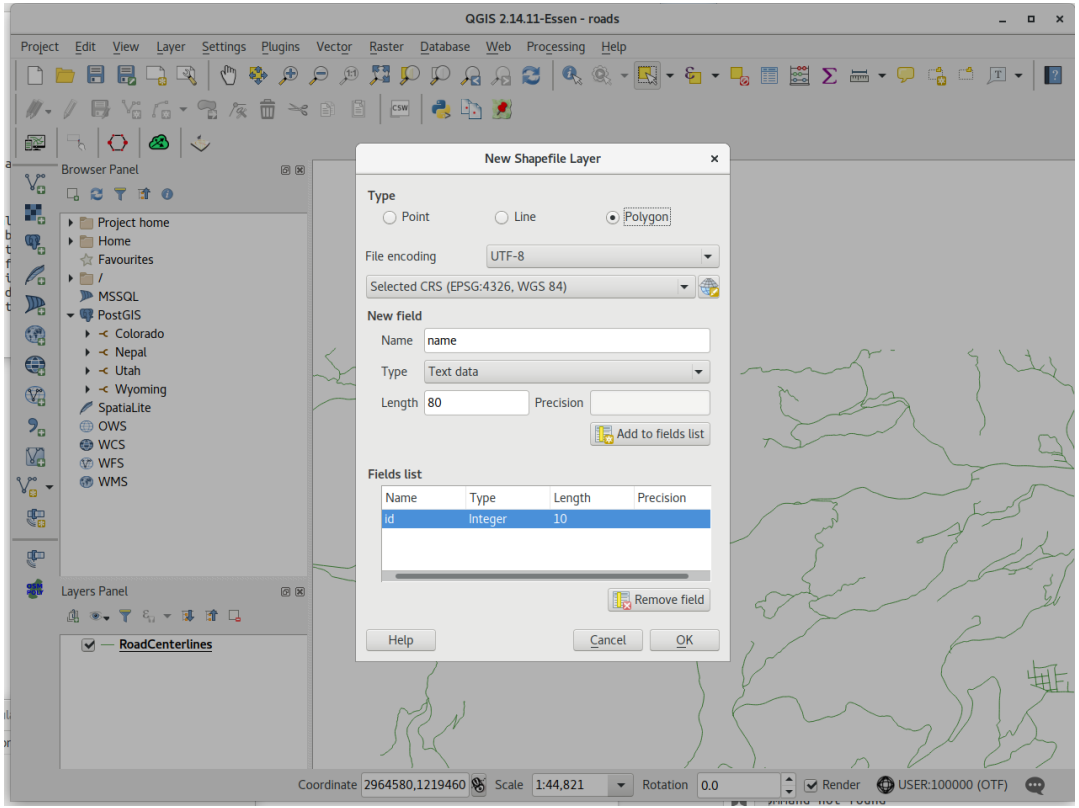
Online Conversion

- https://mygeodata.cloud/converter/
- http://www.gpsvisualizer.com
- https://www.gisconvert.com/
- https://geoconverter.hsr.ch/
- https://www.expertgps.com/convert.asp

Convenient but all have file size limitations. Most use gpsbabel to process the file.

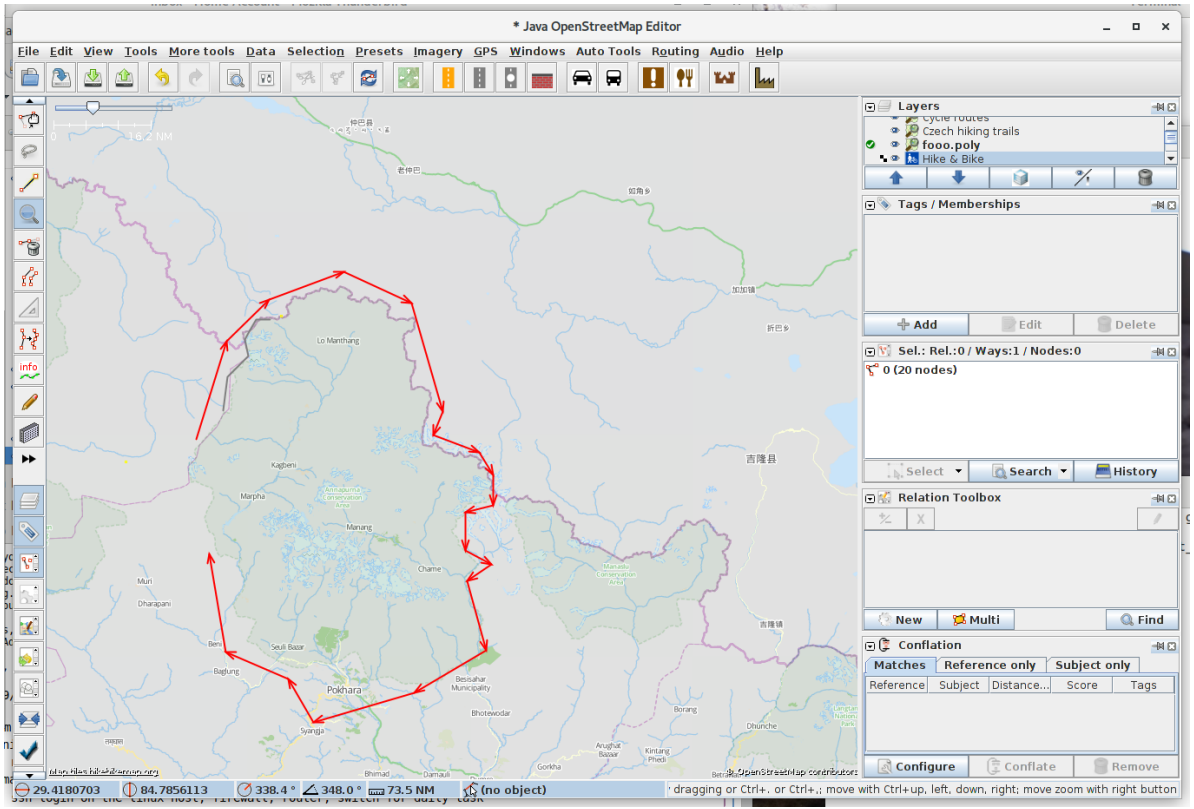
How To Make a Polygon With QGIS

- Create new Shapefile layer
- Select Polygon type
- Add name field
- Remove id field
- Click pencil icon to enable edit mode
- Click on Add Feature icon
- Draw polygon, right click to end
- Export OSM polygon



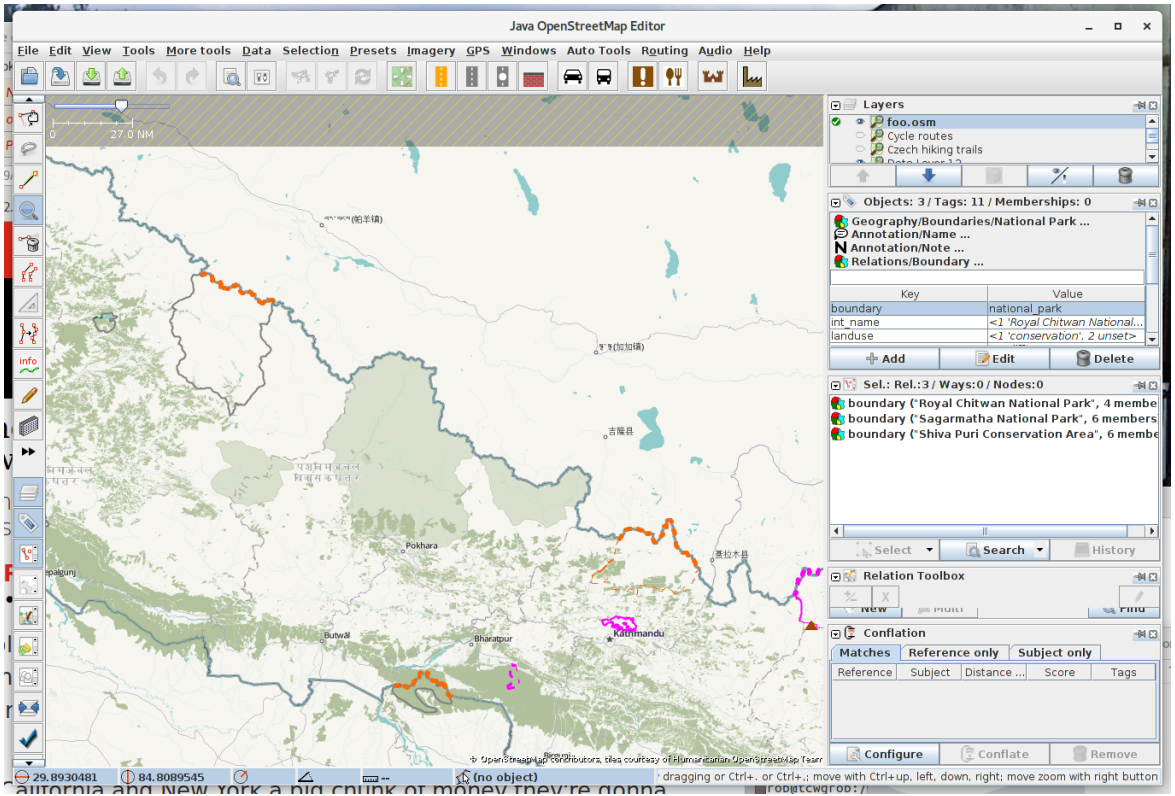
How To Make a Polygon With Josm

- Create New Layer
- Click on Draw Node
- Click to create nodes
- Last node most close polygon
- Save As “poly”



Extract Polygon with Josm

- You can extract an existing polygon
- Usually need to load an OSM file, data is too large for Overpass
- Select by searching presets for boundaries
- Paste into New Layer
- Save as poly

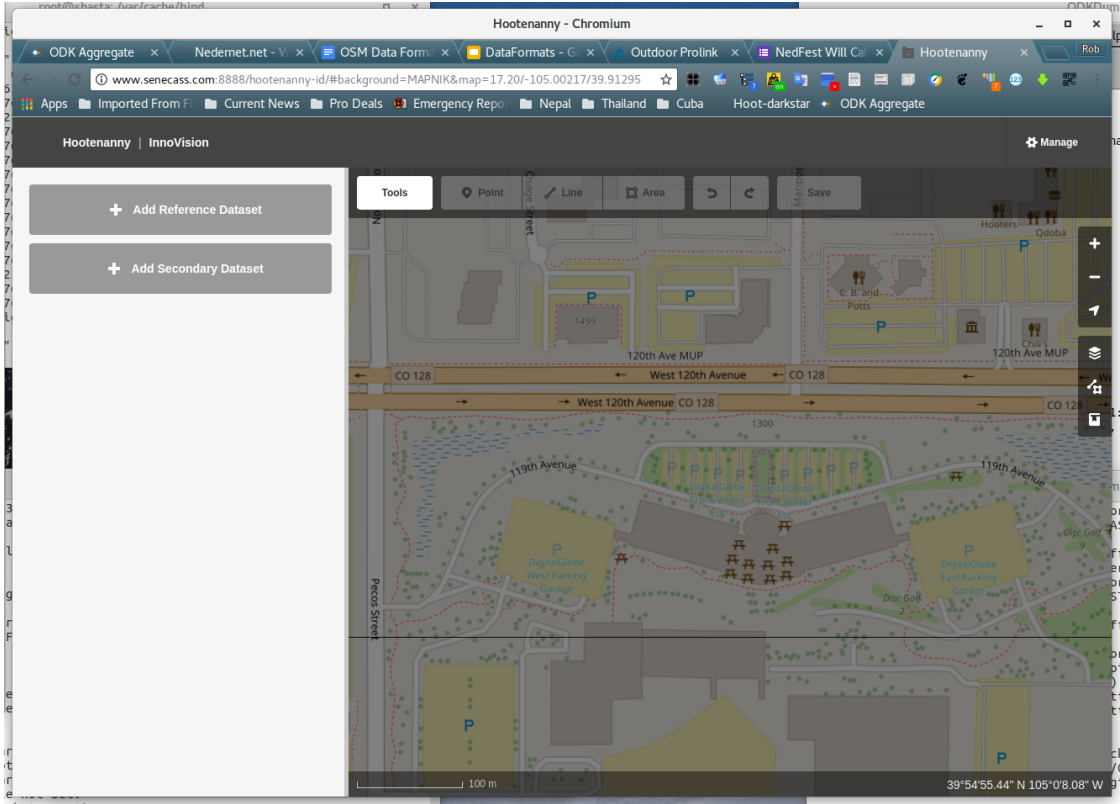


Conflation

- Osm_conflate
- JOSM conflate plugin
- Hootenanny
- Gconflate
- OpenJump with RoadMatcher plugin

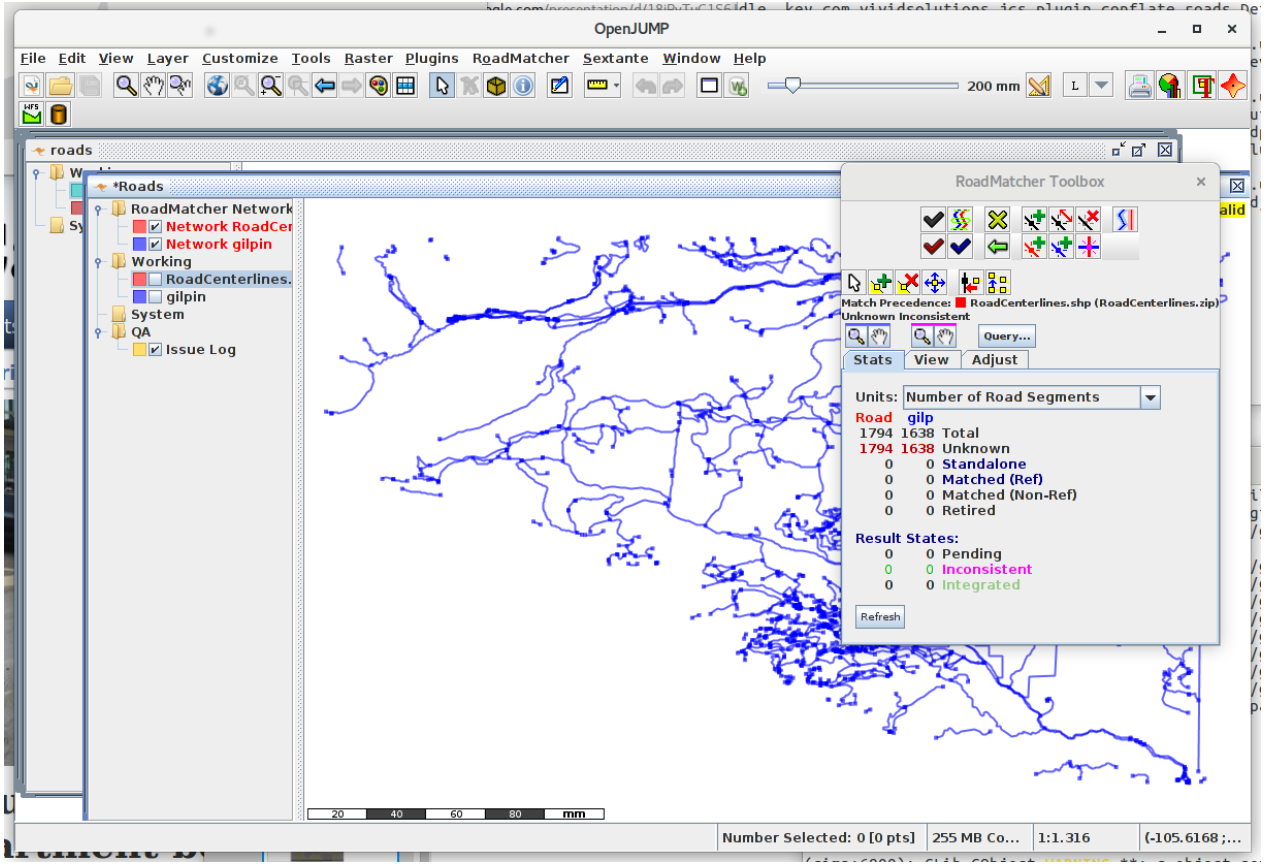
Hootenanny

- Only works online
- Can read Shapefiles
- Can read OSM file
- Flags conflicts with new tag in merged output file



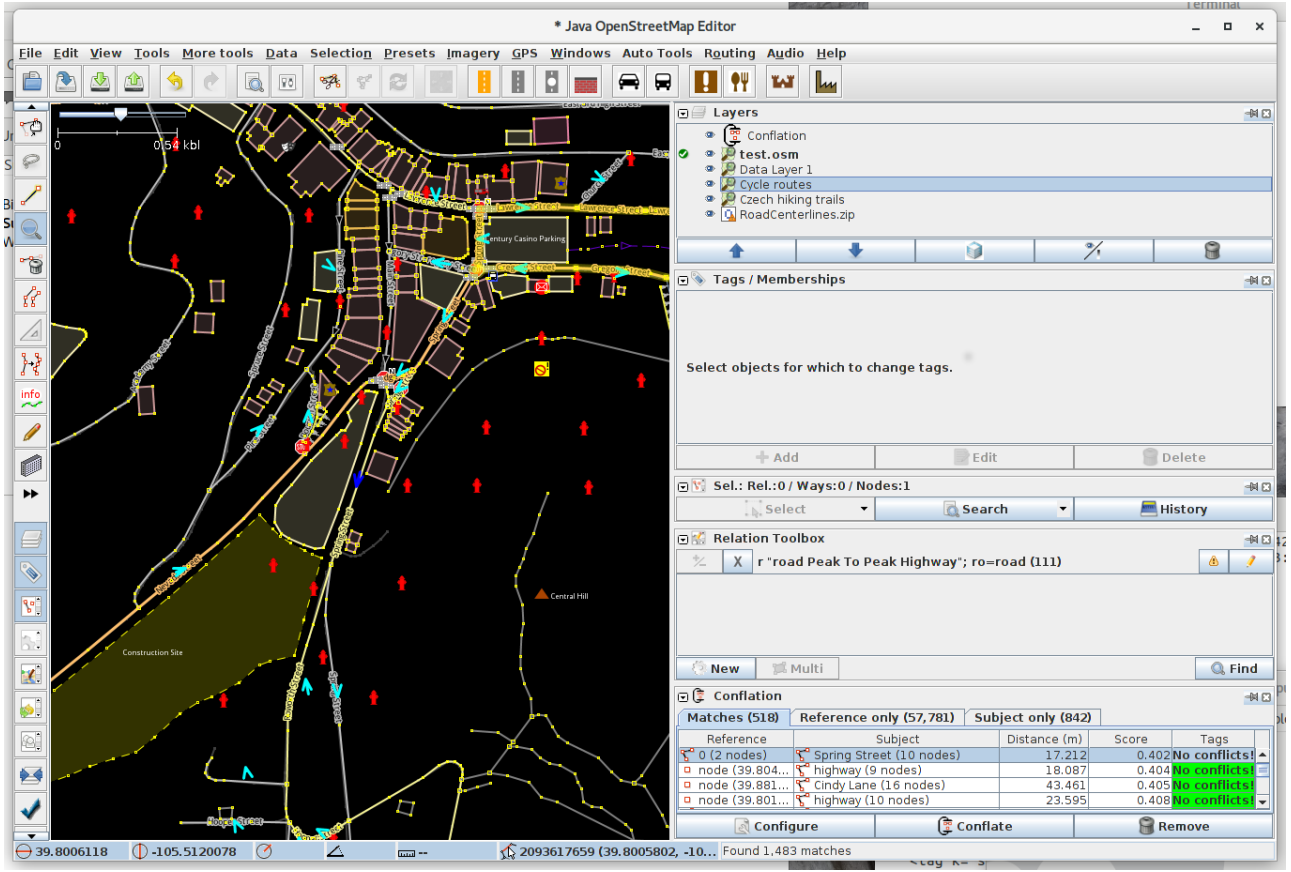
OpenJUMP With RoadMatcher

- Focuses on merging road data
- Can read Shapefile
- Can read OSM file
- Works offline



Josm Conflate Plugin

- Josm can also be used to conflate files
- Reference file can be Shapefile
- Each match or conflict has to be manually



Postgres Setup

createdb Nepal

psql Nepal -c ‘create extension hstore;’

psql Nepal -c ‘create extension postgis;’

psql Nepal -c ‘create extension dblink;’

Using Postgresql

- Data can be imported into the database using several applications

```
osm2pgsql -v --slim -C 1500 -d Nepal --number-processes 8 infile.osm --hstore
shp2pgsql -p in.zip > out.sql
ogr2ogr -f "PostgreSQL" "PG:Nepal=osm" infile.osm
```