

QGIS Application - Bug report #10427

.osm files have problems with iterator or number of features

2014-06-03 04:11 AM - Richard Duivenvoorde

Status:	Closed	
Priority:	Normal	
Assignee:		
Category:	Data Provider/OGR	
Affected QGIS version:	2.2.0	Regression?: No
Operating System:		Easy fix?: No
Pull Request or Patch supplied:		Resolution: end of life
Crashes QGIS or corrupts data:		Copied to github as #: 18841
Description		
<p>QGIS is loading .osm files fine, but there seems something wrong with the number of features QGIS thinks there is in the layer.</p> <p>If you open the osm-file which is in attached test.zip, and load those (small part of Amsterdam), you see a lot of points.</p> <p>But asking the number of features (for example in the properties/Metadata dialog) shows '-1'.</p> <p>Also python scripts which should work on for example a selected feature, fail because of null features.</p> <p>See also: #10000 (processing)</p> <p>Loading qml files created for this, crashes my QGIS (build master dd june 3 2014)</p> <p>QGIS is using the ogr provider for this, and if I use ogrinfo, it shows the number of features ok:</p> <p>ogrinfo lisdodde.osm -sql "SELECT COUNT FROM points"</p> <p>COUNT_* (Integer) = 6523</p>		

History

#1 - 2014-06-05 01:21 AM - Jukka Rahkonen

This relates / is the same as #10000

GDAL/ogrinfo does not report the feature count but it tells that it is unknown (-1). This is because it can be awfully slow to get the feature count from big osm files because all the data must be parsed first. For points it might be tolerable but resolving for example multipolygons is very heavy. You can try yourself

>ogrinfo germany.osm.pbf -sql "select count(*) from multipolygons"

Running this query is actually slower than converting the whole .pbf file into Spatialite with ogr2ogr. I canceled the query after 40 minutes. Probably some resource makes a bad bottleneck with large dataset. Getting an answer from finland-latest.osm.pbf took 2 and a half minutes which is also too long. Result was "COUNT_* (Integer) = 783693"

I guess that QGIS should be be made to load the data first and count the features once the job is done. Calculating feature count from big .osm or .pbf files just is not reasonable.

#2 - 2014-06-05 11:48 AM - Richard Duivenvoorde

Well, yes that is my point. And knowing the amount of features is not the most important missing. But the fact that it is not possible to iterate over the

features with Python/Processing is much more a problem.

#3 - 2014-06-28 07:45 AM - Jürgen Fischer

- Target version changed from Version 2.4 to Future Release - High Priority

#4 - 2016-10-07 03:43 AM - Even Rouault

OSM/PBF files are extremely GIS unfriendly and, unless very small, require conversion to another format to be used in practice. That said with the new Dataset::GetNextFeature() API added in GDAL 2.2dev per https://trac.osgeo.org/gdal/wiki/rfc66_randomlayerreadwrite things should be better, although there is no direct QGIS API match.

Not sure what to do with this ticket

#5 - 2017-05-01 01:08 AM - Giovanni Manghi

- Regression? set to No
- Easy fix? set to No

#6 - 2019-03-09 03:12 PM - Giovanni Manghi

- Status changed from Open to Closed
- Resolution set to end of life

End of life notice: QGIS 2.18 LTR

Source:
<http://blog.qgis.org/2019/03/09/end-of-life-notice-qgis-2-18-ltr/>

#7 - 2019-03-09 05:00 PM - Anita Graser

- Description updated

Information from provider (in layer properties) still shows "Feature count: unknown" in QGIS 3.7 but since #10000 is fixed, we can probably leave this one closed.

Files

test.zip	399 KB	2014-06-03	Richard Duivenvoorde
----------	--------	------------	----------------------