

QGIS Application - Bug report #19659

QGIS3: export to CSV with trailing comma

2018-08-21 03:06 AM - Tobias Wendorff

Status:	Closed	
Priority:	Normal	
Assignee:		
Category:	Data Provider/OGR	
Affected QGIS version:	3.3(master)	Regression?: No
Operating System:		Easy fix?: No
Pull Request or Patch supplied:		Resolution: up/downstream
Crashes QGIS or corrupts data:		Copied to github as #: 27484
Description When exporting only one attribute field to CSV from a multi-field dataset, the header gets a trailing comma [separator]. This makes trouble on some CSV parsers, which are interpreting the header. How to reproduce: 1. Load a dataset, which has more than one attribute fields. 2. Go to "Save vector layer as". 3. Unselect all but one fields and export to CSV. 4. Open the CSV with a text-editor and check the header. Expected: No trailing separator in the header line. This affects at least QGIS 3.2		

History

#1 - 2018-08-21 11:12 AM - Jürgen Fischer

- Category changed from Processing/Core to Data Provider/OGR

#2 - 2018-08-21 11:23 AM - Giovanni Manghi

- Status changed from Open to Feedback
- Operating System deleted (Microsoft Windows 7, 64-bit)

Does the same on 2.18?

#3 - 2018-08-27 01:14 PM - Andrea Giudiceandrea

I can reproduce the bug with QGIS 2.18.23.

#4 - 2018-08-27 01:14 PM - Andrea Giudiceandrea

I can reproduce the bug with QGIS 2.18.23.

#5 - 2018-08-27 06:16 PM - Andrea Giudiceandrea

QGIS relies on OGR CSV driver in order to save a layer as CSV file.

The OGR driver info page at https://www.gdal.org/drv_csv.html states that: "CSV files have one line for each feature (record) in the layer (table). The attribute field values are separated by commas. **At least two fields per line must be present.**"

It seems that the OGR CSV Layer class adds a "fake second blank field" if a single column csv file has to be written, thus the trailing comma, for "valid single column files" creation.

See:

<https://github.com/OSGeo/gdal/commit/92dcffccf8d91696e06b047f275a54a0b7ef0308>

<https://trac.osgeo.org/gdal/ticket/4824>

https://github.com/OSGeo/gdal/blob/master/gdal/ogr/ogrsf_frmts/csv/ogr_csv_layer.cpp#L2219-L2226

Tobias, which CSV parsers have troubles with trailing comma?

#6 - 2018-08-29 11:05 AM - Giovanni Manghi

Andrea Giudiceandrea wrote:

QGIS relies on OGR CSV driver in order to save a layer as CSV file.

*The OGR driver info page at https://www.gdal.org/drv_csv.html states that: "CSV files have one line for each feature (record) in the layer (table). The attribute field values are separated by commas. **At least two fields per line must be present.**"*

then we should close this, agree?

#7 - 2018-08-29 11:18 AM - Tobias Wendorff

Andrea Giudiceandrea wrote:

Tobias, which CSV parsers have troubles with trailing comma?

I've used a Python parser, which creates XLSX files. I need to go through my files to find it again. Actually, it tried to read the header after the first comma to assign a new field, but it was empty. That broke the process.

I've got tons of CSV files, which have one column/field only. Since CSV is a convention and not standard, this might be OGR specific.

I have to take care of this and remove the leading comma in post-processing. Can be closed for the moment.

#8 - 2018-08-29 11:27 AM - Giovanni Manghi

- Resolution set to up/downstream

- Status changed from Feedback to Closed