

QGIS Application - Bug report #14078

Incorrect loading of csv files

2016-01-08 08:17 AM - Paolo Cavallini

Status: Closed	
Priority: Normal	
Assignee: Sebastian Dietrich	
Category: Data Provider	
Affected QGIS version: 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 #: 22088
Description	
The first row of the attached file is not interpreted as column headings. Other, apparently similar, files are loaded fine.	
Related issues:	
Related to QGIS Application - Feature request # 14093: Prefer the QGIS-provid...	Open 2016-01-12

History

#1 - 2016-01-09 01:17 PM - Sebastian Dietrich

- File *QgisLoadCsvWithHeaders.png* added
- File *QgisCsvLayerFields.png* added

I can not reproduce using master (commit:8d105ff8f998bdf3426088d648cd75220b0b469c).

QgisLoadCsvWithHeaders.png

QgisCsvLayerFields.png

#2 - 2016-01-09 07:49 PM - Nyal Dawson

- Status changed from *Open* to *Feedback*

#3 - 2016-01-10 02:02 AM - Paolo Cavallini

- Status changed from *Feedback* to *Open*

Loading through the text delimited plugin is OK, but simply dragging it into the canvas shows the issue reported. Similar files load fine.

#4 - 2016-01-10 09:17 AM - Sebastian Dietrich

OK, now I could reproduce in master. Turns out this is a GDAL/OGR issue.

Note that files are loaded through GDAL/OGR when

- dropping them onto QGIS
- specifying them on the commandline
- double-clicking them on MAC

The meta-data will show OGR data provider.

This is different from opening a file using the *Create a Layer from a Delimited Text File* dialog where the provider will be Delimited text data provider.

The [GDAL/OGR .csv-driver](#) uses some logic to detect whether the first line is a header or not. In your case it finds some all numeric values in the first row (*D1 to D31*, *D* is considered a hex-digit) and treats it as a data record instead of as a header.

The easiest solution would be to put the column names within double quotes, at least the offending ones.

Bad:

```
D1,D2
10,20
```

Good:

```
"D1","D2"
10,20
```

Can you confirm the *similar files* that *load fine* have different, non-numeric column names?

#5 - 2016-01-10 11:39 PM - Paolo Cavallini

- *File test1.csv added*

The attached test1 loads fine.

Thanks for your feedback.

Perhaps worth opening a ticket on GDAL.

#6 - 2016-01-12 03:09 PM - Giovanni Manghi

- *Category changed from Data Provider/Delimited Text to Data Provider*

#7 - 2016-01-12 04:00 PM - Sebastian Dietrich

- *Status changed from Open to Closed*

- *Resolution set to up/downstream*

Paolo Cavallini wrote:

```
| The attached test1 loads fine.
| Thanks for your feedback.
| Perhaps worth opening a ticket on GDAL.
```

You're welcome.

You might consider opening a feature request to prefer the QGIS-providers over the OGR-provider for files that are supported by both, like CSVs.

#8 - 2016-01-12 04:07 PM - Sebastian Dietrich

- *Assignee set to Sebastian Dietrich*

#9 - 2016-01-12 11:02 PM - Paolo Cavallini

Done both, thanks.

#10 - 2016-01-13 01:21 AM - Even Rouault

Fixed in GDAL per <https://trac.osgeo.org/gdal/ticket/6305>. I'd note that the OGR CSV driver has gained editing capabilities in GDAL 2.1 dev (specifically to be used through QGIS), whereas AFAICS the QGIS CSV provider is read-only. It would be a shame that using the OGR CSV driver is no longer possible.

Files

test0.csv	223 Bytes	2016-01-08	Paolo Cavallini
QgisLoadCsvWithHeaders.png	34.6 KB	2016-01-09	Sebastian Dietrich
QgisCsvLayerFields.png	25.5 KB	2016-01-09	Sebastian Dietrich
test1.csv	517 Bytes	2016-01-10	Paolo Cavallini