QGIS Application - Bug report \#19322
Random selection within subsets throws exception
2018-07-03 04:02 PM - Dénes Szalontai

| Status: Closed |  |
| :---: | :---: |
| Priority: High |  |
| Assignee: |  |
| Category: Processing/QGIS |  |
| Affected QGIS version:3.2 | Regression?: Yes |
| Operating System: | Easy fix?: No |
| Pull Request or Patch supp lied: | Resolution: |
| Crashes QGIS or corrupls data: | Copied to github as \#: 27150 |

## Description

Random selection within subsets throws an exception if a subset has less element than the value set in 'Number/percentage of selected feature' field and Method is 'Number of selected features'

Error message:

Traceback (most recent call last):
File "C:/PROGRA~1/QGIS3~1.2/apps/qgis/./python/pluginslprocessinglalgslqgis\RandomSelectionWithinSubsets.py", line 141, in processAlgorithm
selran.extend(random.sample(subset, selValue))
File "C:\PROGRA~1\QGIS3~1.2lapps\Python36liblrandom.py", line 317, in sample
raise ValueError("Sample larger than population or is negative")
ValueError: Sample larger than population or is negative

It worked on the same data with QGIS 2.18.10 (but was extremely slow)

## Associated revisions

Revision 29207a16-2018-07-04 05:30 PM - Alexander Bruy
[processing] fix Random extract/select within subset when subset is
smaller than number of requested features (fix \#19322)

Revision 5ae24307-2018-07-10 03:15 AM - Alexander Bruy
[processing] fix Random extract/select within subset when subset is
smaller than number of requested features (fix \#19322)
(cherry-picked from 29207a16)

History
\#1-2018-07-03 06:26 PM - Giovanni Manghi

- Status changed from Open to Feedback

Did it worked as expected (with the described values) in previous releases?

## \#2-2018-07-03 06:58 PM - Dénes Szalontai

- Status changed from Feedback to Open

I checked it in QGIS 3.0.3 and the same issue occurs:

Traceback (most recent call last):
File "C:/PROGRA~1/QGIS3~1.0/apps/qgis/./python/plugins\processinglalgs\qgis\RandomSelectionWithinSubsets.py", line 137, in processAlgorithm
selran.extend(random.sample(subset, selValue))
File "C:\PROGRA~1\QGIS3~1.0\apps\Python36\lib\random.py", line 317, in sample
raise ValueError("Sample larger than population or is negative")
ValueError: Sample larger than population or is negative

It works well and produces the expected result in QGIS 2.18 .10 with the same input data and the same settings.

## \#3-2018-07-03 07:00 PM - Giovanni Manghi

- Regression? changed from No to Yes
- Priority changed from Normal to High


## \#4-2018-07-04 08:02 AM - Alexander Bruy

- Category changed from Python plugins to Processing/QGIS
- Status changed from Open to Feedback

Please provide test dataset and steps to reproduce the issue

## \#5-2018-07-04 02:05 PM - Dénes Szalontai

- File Sample.csv added
- Status changed from Feedback to Open


## Steps to reproduce:

1. Add the attached sample.csv layer with Layer $->$ Add Layer $->$ Add Delimited Text Layer (delimiter: semicolon, the first record has field names, EPSG: 4326)
2. Start the Random selection within subsets tool from Vector -> Research Tools
3. Select Sample as Input layer, SPEEDLIMIT as ID field, Number of selected features as Method and 5 as Number/percentage of selected features
4. Hit Run to start processing

Actual behaviour in QGIS 3.x:
An exception is thrown and the execution fails with the error message

Traceback (most recent call last):
File "C:/PROGRA~1/QGIS3~1.2/apps/qgis/./python/plugins\processinglalgs\qgis\RandomSelectionWithinSubsets.py", line 141, in processAlgorithm
selran.extend(random.sample(subset, selValue))
File "C:\PROGRA~1\QGIS3~1.2lapps\Python36\ib\random.py", line 317, in sample
raise ValueError("Sample larger than population or is negative")
ValueError: Sample larger than population or is negative

Actual behaviour in QGIS 2.18.10:
Process finishes successfully with the expected result

Note: The sample data contains the value 110 less than 5 times (only once) in SPEEDLIMIT column. If I set Number/percentage of selected features to 1 then process finishes successfully in QGIS 3.x as well but it fails with any other values.

Note 2: Random extract within subsets tool fails as well with this data in QGIS 3.x

## \#6-2018-07-04 06:53 PM - Alexander Bruy

- \% Done changed from 0 to 100
- Status changed from Open to Closed

Applied in changeset commit:qgis|29207a1617678eb3a3922ff3dd74c3b6b6b193a5.

## Files

| Random_selection_within_subsets.txt | 883 Bytes | $2018-07-03$ | Dénes Szalontai |
| :--- | ---: | ---: | :--- |
| Sample.csv | 5.97 KB | $2018-07-04$ | Dénes Szalontai |

