# Spatial Quality and Anomalies Diagnosis (SQUAD) Tool for QGIS Quick Start Guide

April 2018





## Spatial Quality and Anomalies Diagnosis (SQUAD) Tool for QGIS Quick Start Guide

April 2018

MEASURE Evaluation University of North Carolina at Chapel Hill 123 West Franklin Street, Suite 330 Chapel Hill, NC 27516 USA Phone: +1 919-445-9350 measure@unc.edu www.measureevaluation.org This publication was produced with the support of the United States Agency for International Development (USAID) under the terms of MEASURE Evaluation cooperative agreement AID-OAA-L-14-00004. MEASURE Evaluation is implemented by the Carolina Population Center, University of North Carolina at Chapel Hill in partnership with ICF International; John Snow, Inc.; Management Sciences for Health; Palladium; and Tulane University. Views expressed are not necessarily those of USAID or the United States government. TL-18-06 ISBN: 978-1-64232-021-3







### INTRODUCTION

QGIS is a free and open-source geographic information system (GIS) software. The Spatial Quality and Anomalies Diagnosis (SQUAD) tool is a QGIS plug-in that will assess data quality of large spatial data sets. Because it can be difficult to perform data quality checks on large data sets, the SQUAD tool rapidly and automatically looks for anomalies in the data that may indicate data quality issues.

The tool requires two data sets (in the form of shapefiles): one consisting of point locations and the other consisting of polygons that represent administrative units. The tool will review the point locations and identify one of six anomalies:

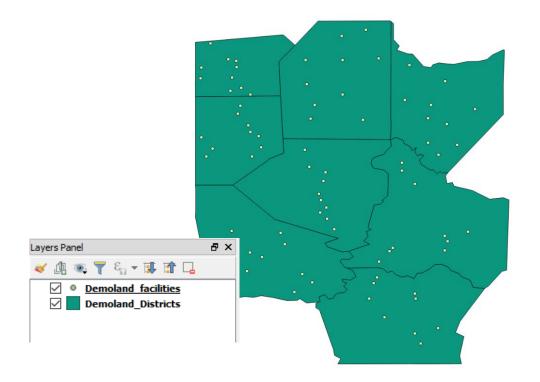
- 1. Missing coordinates
- 2. Truncated coordinates (lack of adequate precision)
- 3. Duplicate coordinates for distinct records
- 4. Duplicate key attributes (two identical names, but plotting in different locations)
- 5. Coordinate not located exactly where it would be expected (but falling within two kilometers of a border)
- 6. Coordinate not located anywhere near where expected

The quick start guide provides a general overview of the use of the tool with the sample data available for the tool. A more detailed overview is available on the MEASURE Evaluation web site, here: <u>https://www.measureevaluation.org/resources/tools/geographic-information-systems/squad-tool/spatial-quality-and-anomalies-diagnosis-squad-tool</u>.

NOTE: Screen shots are of QGIS version 2.18. While the interface of version 3.0 has a slightly different appearance, the steps are the same. The tool is available as a plug-in for both versions. QGIS is available as a free download at <u>http://qgis.org.</u>

#### Quick Start Guide for the SQUAD Plug-In for QGIS

- (1) Launch QGIS version 2.18 or 3.0 (search for and then double-click on the QGIS Desktop shortcut icon, or search for the app in the windows taskbar).
- (2) Click on **Project>New**. Click the **"Add Vector Layer"** button on the left-hand toolbar.
- (3) Browse to the "sample data" folder that came with these instructions. In the file type pop-up menu choose ESRI Shapefiles ESRI Shapefiles (\*.shp \*.SHP) and then select Demoland\_Districts.shp. Do the same for Demoland\_facilities.shp.
- (4) Make sure that the Demoland\_facilities layer appears above the Demoland\_Districts layer in the Layers Panel, and that both are visible on the screen:



- (5) Go to **Plugins>Manage and Install Plugins...** Click on the **All** tab, and in the search box **type** SQUAD.
  - a. For version 2.18, click on SQUAD tool.

- b. For version 3.0, click on SQUAD Tool v3.
- c. Click the "Install Plugin" button, then Close the dialog box.
- (6) Look for the SQUAD tool icon in the toolbar and click on it to activate the tool. (Alternatively, go to **Plugins>SQUAD>SQUAD plugin**.) This will start the tool. You should see a dialog box like this:

💋 SQUAD Tool		? ×
Parameters Log	Run as batch process	SQUAD
Site File Demoland_Districts [EPSG:4326]	× 2	This tool searches through a point file to identify data anomalies that could indicate data quality issues. Each of these new columns will be created in the original site file. A "1" indicates that the anomaly exists. It is possible for a site to have multiple anomalies identified.
Site Admin Unit Field Country Site Longitude Field		<ul> <li>Anomaly_1: Missing coordinate.</li> <li>Anomaly_2: Coordinate is truncated.</li> <li>Anomaly_3: The point has an identical coordinate as another point in the file.</li> <li>Anomaly_4: The point has an identical name as another point in the file.</li> <li>Anomaly 5: The point is outside of the expected administrative unit, but is within</li> </ul>
Shape_Area Site Latitude Field Shape_Area		<ul> <li>Anomaly_S. The point is obtained on the expected administrative unit, but is writing two kilometers of the expected unit's border.</li> <li>Anomaly_6. The point is outside of the expected administrative unit, and is greater than two kilometers of the expected unit's border.</li> </ul>
Site Name Field Country Site ID Field	<b></b>	
Country Administrative Units File	•	FROM THE AMERICAN PEOPLE PEPFAR MEASURE Evaluation
Demoland_Districts [EPSG:4326] Admin Unit Name Field Country	- 2 -	
Site Anomalies Output [Save to temporary file] Open output file after running algorithm		
	¥	< > > 0%
		Run Close

(7) Fill in the various parameters needed to check for anomalies in your data set. For the test data set included, Demoland, choose the following parameters:

Site File	
Demoland_facilities [EPSG:4326]	2
Site Admin Unit Field	
District	•
Site Longitude Field	
POINT_X	•
Site Latitude Field	
POINT_Y	•
Site Name Field	
NAME	•
Site ID Field	
UID	•
Administrative Units File	
Demoland_Districts [EPSG:4326]	2
Admin Unit Name Field	
Dist_name	•

(8) In the bottom box, under "Site Anomalies Output," click the Browse button:

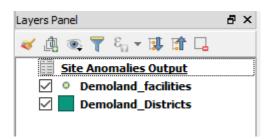
Site Anomalies Output	$\frown$
[Save to temporary file]	
Open output file after running algorithm	

... and choose "Save to file."

(9) In the resulting dialog box, you should choose to save the file as either a new Shapefile (SHP) or an Excel spreadsheet (XLSX). You may put it in the location of your choosing. NOTE: if you do not save to a file, a temporary output layer will be generated. You can then save this output by rightclicking on it in the Layers Panel and choosing "Save As." If you do not do either of these things, then when you quit QGIS, the information generated by the tool will be lost.

🦸 Save file		?	×
Look in:	C: \Users \bwilkes \Desktop \QGIS_SQUAD_test		:: =
S My Cor	mputer		
la bwilkes	s II		
File name:	Demoland_anomalies		Save
Files of type:	XLSX files (*.xlsx)	(	Cancel
Encoding:	System 🔻		

(10) Click "Save," then click on the "Run" button. The Log tab will appear and show you a progress bar while the tool is running. Your new output layer will display in the Layers Panel with the name "Site Anomalies Output." (You may double-click it to see and edit its name and source file, and source file location.)



(11) Right-click on the new layer and choose "Open Attribute Table" from the drop-down menu to view the contents of your newly-coded file. If you saved your output as a SHP file, you can select various locations in the table to highlight and view them on the map.

#### **MEASURE** Evaluation

University of North Carolina at Chapel Hill 123 West Franklin Street, Suite 330 Chapel Hill, NC 27516 USA Phone: +1 919-445-9350 measure@unc.edu

www.measureevaluation.org

This publication was produced with the support of the United States Agency for International Development (USAID) under the terms of MEASURE Evaluation cooperative agreement AID-OAA-L-14-00004. MEASURE Evaluation is implemented by the Carolina Population Center, University of North Carolina at Chapel Hill in partnership with ICF International; John Snow, Inc.; Management Sciences for Health; Palladium; and Tulane University. Views expressed are not necessarily those of USAID or the United States government. TL-18-06 ISBN: 978-1-64232-021-3





