

QGIS Application - Feature request #12607

Improved handling of calculated values

2015-04-20 02:32 AM - Dan Isaacs

Status:	Closed	Resolution: Copied to github as #: 20727
Priority:	Normal	
Assignee:		
Category:	Vectors	
Pull Request or Patch supplied:		
Easy fix?:	No	
Description		
<p>There is a major flaw in the Qgis system in dealing with calculated values. For example, we have a simple demographic of population densities consisting of polygons, their areas, their population, and a calculated attribute 'Population Density'(Population/Area).</p> <p>I say major flaw, because something fundamental seems to prevent the use of this calculation at every turn, making me think it's encoded deep within the way Qgis works. The options I've tried (and failed at) are;</p> <ol style="list-style-type: none">1. Use a simple calculated field with Field Calculator - Failed: If the boundary or population of one polygon changes the value is not updated (even if you click on 'update', which always baffled me).2. Use a virtual field so that the values are updated - Failed: No further use is then allowed, GroupStats does not recognise virtual fields, databases do not recognise virtual fields, spreadsheets, conversion to raster etc. all do not recognise the values in virtual fields.3. Use a spreadsheet to do the calculations and then join it to the polygons - Failed: Importing the data works, but then if any changes are made and saved, the calculated field reverts to a fixed integer. <p>Solutions to any of the methods above would work, but I see the simplest as making the update button work as expected. i.e. if you have one attribute calculated from the value of others, when you update, it should re-populate that column as if you had created it afresh, including any changes that might have been made to the attributes it is using in its calculation.</p>		

History

#1 - 2015-04-20 03:08 AM - Nathan Woodrow

2. Use a virtual field so that the values are updated - Failed: No further use is then allowed, GroupStats does not recognise virtual fields, databases do not recognise virtual fields, spreadsheets, conversion to raster etc. all do not recognise the values in virtual fields.

Virtual fields are only a new feature in 2.8 so many plugins might not yet support them correctly. Opening a ticket for that plugin is the best way to raise this issue.

#2 - 2015-04-26 03:08 PM - Giovanni Manghi

- Category set to Vectors
- Status changed from Open to Feedback

what "update" button do you refer to?

Generally speaking: I do not agree that this is a "major flow" in QGIS (or any other gis program that works this way) 1) as the capability of auto-update attributes after certain operations (like modifying a geometry) depends a lot on what datasources you are using (ex: postgis/spatialite and triggers). If you use for example shapefile you cannot do that, unless something else comes into play. 2) In the QGIS case this help is available as virtual fields and it is a quite recent feature. 3) I don't understand why do you think that joining a table to a vector would give you auto-updating attributes.

#3 - 2015-04-27 12:10 AM - Dan Isaacs

1) The 'update' button I'm referring to is actually the 'update all' button on the attribute table. If I have a column whose values are the result of the formula 'column1'*'column2', and I change a value in column1, I expect 'update all' to update *all* the attributes. I really can't see a circumstance in which anyone would want to keep the results of a calculation when the figures used in it have changed. Hence 'major flaw', the software is supposed to be about geographical analysis (i.e calculations). If they all become redundant the moment one piece of data changes, I do see that as a pretty major flaw.

2) I'm aware that virtual fields go some way to solving this problem, but again, the software is supposed to be about analysis, if one cannot do any further analysis on these fields because they cannot be exported or used by any other analytical engine, then they become far less meaningful. Also, if I have an ordinary (not virtual) column with an expression in it and I want it to update, I can delete the formula, re-write it and then press 'update all' and it works. I can't see why this process should be so difficult to automate without asking the user to re-write the formula manually.

3) I don't think that joining a table to a vector *should* give me auto-updating attributes. I'm suggesting it could, as a way around this problem. If Qgis could recognise and update the value of spreadsheet cells, even if they were the results of calculations, then you could harness all the processing power of a spreadsheet to a gis system.

Qgis is supposed to be about analysing geographical information. Information changes. Updating analysis as data changes should be at the heart of the software, not an afterthought.

#4 - 2015-04-27 12:31 AM - Giovanni Manghi

Dan Cuillin wrote:

1) The 'update' button I'm referring to is actually the 'update all' button on the attribute table. If I have a column whose values are the result of the formula 'column1'*'column2', and I change a value in column1, I expect 'update all' to update *all* the attributes. I really can't see a circumstance in which anyone would want to keep the results of a calculation when the figures used in it have changed. Hence 'major flaw', the software is supposed to be about geographical analysis (i.e calculations). If they all become redundant the moment one piece of data changes, I do see that as a pretty major flaw.

the "update all" button do indeed update all records, of a specific column (you choose on the left) and using a specific expression (you can write in the proper space or by opening the expression builder). If you are using data stored in a real database (postgis, spatialite, oracle, whatever) and you have a column that is computed using 'column1' and 'column2' and then 'column1' or 'column2' or both changes, then can have your column automatically updated with triggers. And this will work regardless of the software you are using to view/edit your data. If you use other formats, like shapefiles, you can't do that, and in this cases is the specific software you are using that must come to help. QGIS uses virtualfields. But virtualfields are a QGIS things, so the moment you export such datasources or use them in other software you of course loose such "help".

2) I'm aware that virtual fields go some way to solving this problem, but again, the software is supposed to be about analysis, if one cannot do any further analysis on these fields because they cannot be exported or used by any other analytical engine, then they become far less meaningful.

see above

Also, if I have an ordinary (not virtual) column with an expression in it and I want it to update, I can delete the formula, re-write it and then press 'update all' and it works. I can't see why this process should be so difficult to automate without asking the user to re-write the formula manually.

see above, it depends on the datasource, not the software.

3) I don't think that joining a table to a vector *should* give me auto-updating attributes. I'm suggesting it could, as a way around this problem. If Qgis could recognise and update the value of spreadsheet cells, even if they were the results of calculations, then you could harness all the processing power of a spreadsheet to a gis system.

I don't see the advantages over virtual fields here.

Updating analysis as data changes should be at the heart of the software, not an afterthought.

of course, but again, it depends a lot on the data sources.

#5 - 2015-04-27 02:41 AM - Dan Isaacs

My apologies if I have misinterpreted the purpose of this section, but I've not submitted a bug report, I've submitted a feature request. Your responses very helpfully and accurately reflect what *is* currently the case, what I'm talking about is what *should be* the case in order for the software to become more functional.

Qgis acts as a 'front-end' for many database types, if there is a way to create such auto-updating columns in other databases, then why not have a function within Qgis that does just that through the GUI?

Whatever is currently the case, there is a problem in that **auto-updating calculated expressions in shapefiles cannot be used in further analysis**. I think it is self-evident that it would be useful if they could. Therefore it is a legitimate feature request to ask that some means be found to do this. Making the 'update all' button act in a pre-programmed way as if it were re-writing the expression and updating the field would achieve this with very little change to the software. Whatever database one is using, Qgis acts as a front-end when it come to writing data to it and therefore it should be possible for it to write, and when necessary re-write the results of an expression in order to keep the data up to date.

#6 - 2015-04-27 03:03 AM - Nathan Woodrow

Maybe I'm missing something but the virtual fields already does all this by updating itself when something changes. What are the main issues with using virtual fields? If there are tools that are currently missing the virtual fields when listing the usable fields this will need to be addressed for each tool.

#7 - 2015-04-27 04:13 AM - Dan Isaacs

To use the example I gave in the original post, you have a simple set of polygons, each has an area value, a population value and a calculated density. You make the calculated density column and the area column virtual fields so that when the population or boundaries change it updates itself. You then need to produce some analysis. What is the average population density?, what is the most densely populated polygon under 5Ha?, what is the total population density? These are questions I think crop up quite frequently in geographical analysis.

GroupStats can't do it currently, but even if it could, it doesn't have the capacity that a spreadsheet has. Database reporting tools can do it, but the virtual figures aren't transferred to the databases. And what if you have (as I do) to produce such reports regularly? It is a simple matter to set up a spreadsheet template or database report to carry out the same process each time. This is not possible with tools like GroupStats even if we could get it to work with virtual columns.

Seeing as *if* this problem could be solved, someone will have to do some re-programming (I can only apologise that it can't be me). It seems to me that the most robust solution would be to get the auto-updating data into the databases (allowing endless data analysis) rather than get the makers of GroupStats (or similar) to understand virtual columns, which would only allow a very limited amount of analysis, and replicate functions that a spreadsheet can already do.

Its not as if this functionality is not possible, I can do it manually. I go into field calculator, tick 'update existing field', re-write the formula, click 'update all'. All I'm suggesting is that this process be automated so that the user doesn't have to remember and re-write what can sometimes be long complicated formulas. All the software needs to do is remember the last formula that was in that column.

#8 - 2015-04-27 12:12 PM - Giovanni Manghi

- Status changed from Feedback to Open

#9 - 2015-04-27 03:56 PM - Nathan Woodrow

Ok so the main issue is you need to have those fields in the database, say area, because you need to use the fields in other tools, but keep them updated in QGIS easily is that correct?

#10 - 2015-04-27 11:02 PM - Dan Isaacs

Yes, that's certainly the most robust solution I see, a simple automated way of keeping calculated Shapefile fields up to date in the actual .dbf files. If, instead of just looking for fresh formulae, the 'update all' process carried out a small program that re-wrote the last used formula in the column so that it triggers the database to re-calculate the values, based on the latest data.

#11 - 2016-08-18 02:39 PM - gcarrillo -

- Status changed from Open to Closed

I've just released a QGIS plugin that addresses exactly this issue. It's called AutoFields (<http://plugins.qgis.org/plugins/AutoFields/>).

You can even watch a video with the use case Dan has mentioned. An Area field and a Population Density field that are updated after feature creation or modification:

<https://vimeo.com/germap/autofields-area-population-density>