Best Practices, Tips and Altering Defaults in QGIS

View   Edit   Delete   Revisions   Clone   Translate

Signed in as scheeva   Account settings   Sign out
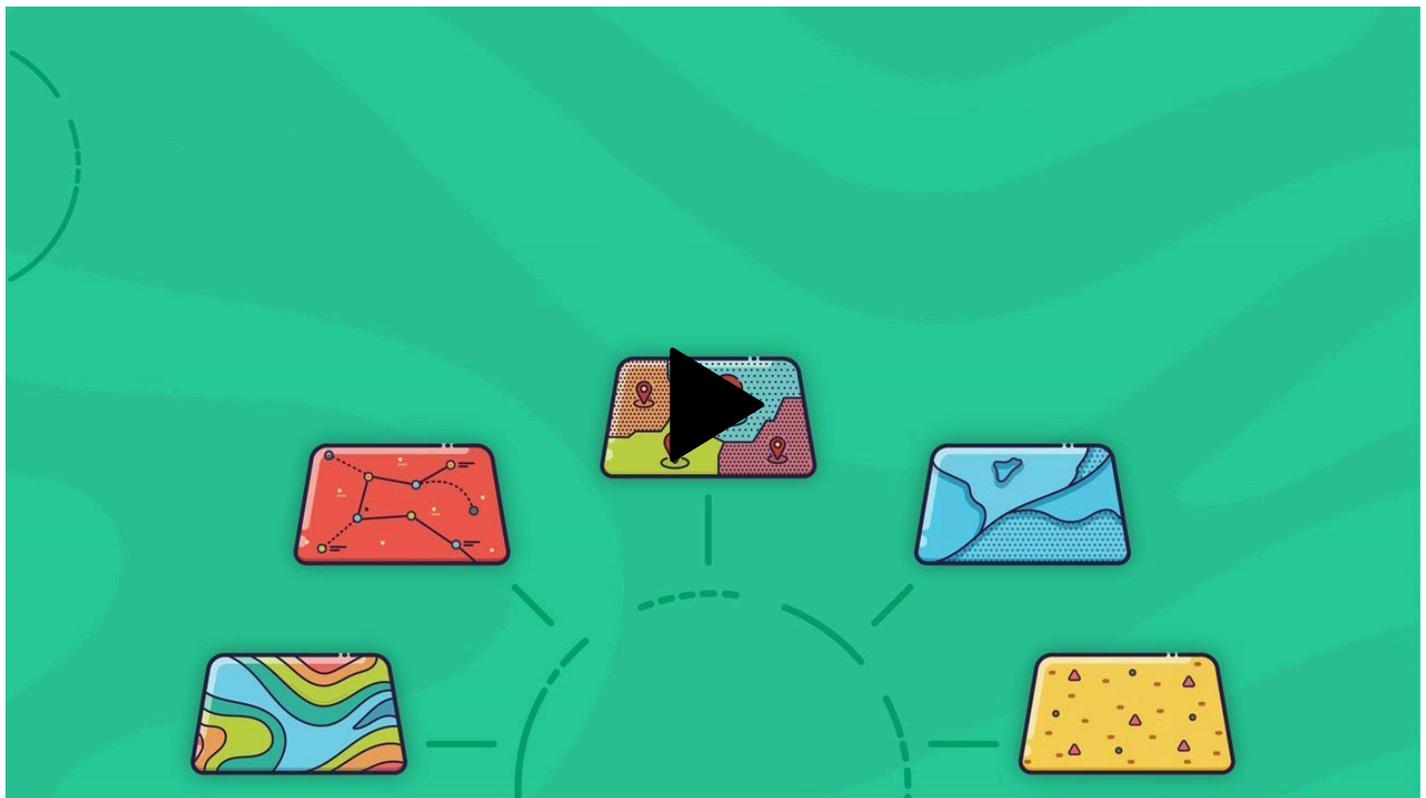
UUID Link: [uuid-link:node:a09da369-c531-4c5d-ad41-c601c420e447]

Catalogue number: <u>89200005</u>   **Issue number:** 2020011

Release date: November 19, 2020

QGIS Demo 11

▼ Best Practices, Tips and Altering Defaults in QGIS - Video transcript

(The Statistics Canada symbol and Canada wordmark appear on screen with the title: "Best Practices, Tips and Altering Defaults in QGIS")

Today we'll introduce some tips and best practices for using QGIS, covering topics such as file management, optimizing workflows and accessing, running and troubleshooting common issues with processing tools. We'll also briefly discuss changing program defaults, enabling some customization of the interface and treatment of data according to individual needs. These best practices and tips will help avoid frustration and facilitate processing, analyzing and sharing spatial data products and visualizations with others.

So the first tips concern file and directory management. Like many programs QGIS uses absolute file paths by default to link layers to project files. Therefore, if a directory or file is moved or renamed, the new path must be provided when reopening the project otherwise the

affected layers are discarded. So select layer and click Browse – navigating to the new directory or filename, and for a shapefile - select the .shp component of the layer.

So as noted, all spatial data should be in a common directory – here the Geospatial Data folder - using additional subdirectories and distinctive file-names for further organization. It is still best practice to avoid spaces and special characters in filenames or directories – as this can complicate saving or loading files. So substitute spaces with underscores or dashes as required. Finally, using GIS it's easy to rapidly create multiple files – so ensure to manage your directories judiciously.

The optimal file format for a dataset depends on the intended use. Shapefiles help quickly share layers with others for analysis, visualization and editing – while geodatabase and geopackage files enable layers of different geometry types to be stored in a single file; with the original layers locked from editing unlike Shapefiles – there are no limits on field name lengths. The Package Layers tool can be used to create a geopackage, where we could then combine the points from the Grain Elevators layer, lines from the road segments and polygons from the projected Manitoba census subdivisions. After saving to a permanent file, we could then load the layers like a file geodatabase. And there are a variety of other formats – such as KML for loading and displaying a vector layer in Google Earth. In general, use the Format drop-down in the Save Vector Layer As box to change to the desired file format. There are many sources detailing the applications, advantages and disadvantages of major formats, which can be consulted in determining the best format for your data.

And to improve rendering times for large vector datasets we can use the Create a Spatial Index tool from either the toolbox or the Source Tab of the Layer Properties Box. The raster equivalent is Build Overviews - creating coarser resolution versions of the input for rapid rendering at broader extents.

The next tips relate to GIS workflows. Ultimately, there are many ways to complete the same task in GIS. So the shortest workflow, in number of steps, intermediary outputs or processing times, which achieves the same result is the best workflow. Comparing these expressions – which produce the same selection - the second expression is better – as it avoids repeating the field name and operator for each attribute of interest. So apply these principles to your own workflows – whether it is the specific tools applied, the order in which they're implemented or, as just shown, the way that an expression or code is written.

So QGIS tools can be accessed from the menu-bar drop-downs or from the Processing Toolbox. Note there is some mutual exclusion in the available tools – such as the Check Geometries Core Plugin in the Vector drop-down or the additional GDAL, SAGA and GRASS tool set, as well as user-created models and processing scripts in the Toolbox. I find that the Toolbox is the fastest and easiest way to isolate available tools using the Search bar, as this will also return additional or alternative tools that may be relevant for your workflow. If needed, use the descriptions on the right-side of a tool to help parametrize it. Note that the parameters can vary depending upon the specific source of the tool. For example, the QGIS Slope tool has just two parameters, for the Digital Elevation Model and Z factor, while the GDAL Slope tool contains additional parameters such as expressing slope in percentages vs degrees. The appearance of tools can also vary according to the location that they're accessed from. So, for example,

opening the Select by Expression tool from the Toolbox is markedly different in it's appearance from that on the Attribute Toolbar – lacking the central drop-downs to help us construct our expressions.

The next item is on spatial properties. As noted, when using multiple layers in QGIS, the projection, datum and coordinate reference system should be uniform. Although QGIS re-projects layers on-the-fly for visualization to the Project Coordinate Reference System – established by the first loaded layer - it does not resolve these differing properties for processing and analysis. For spatial analysis, use a Projected Coordinate Reference System – tailoring the selected system to the required precision for your analytical needs.

Conversely, due to the potential effects on cell alignments and values, rasters should not be re-projected unless necessary such as for spatial analysis or integrating multiple rasters from different sources. In these cases, the alignment and resolution of cells should also match - which can be accomplished using the Align Rasters tool. Select the input layers, output file name and resampling method. The coarser resolution raster should be used as the Reference layer. And as we can see, the position of pixels compared against the original raster have been slightly shifted, but toggling on the aligned DEM we can see that their cells are aligned which we could then process and analyze further as required. Similarly, when sampling raster layers ensure that the minimum distance between points is greater than the resolution of cells to avoid violating assumptions of statistical independence.

The next tips concern running processing tools. Most tools can be run on single layers or Run as a Batch Process for multiple inputs. However, when run as a Batch Process – temporary layers and Selected Features Only are not available. The Multiple Selection box

can help rapidly select layers of interest, and where possible we can copy and paste parameters to reduce manual inputs. To store intermediary layers we can create a temporary directory which we can be deleted after processing – as I did to re-project layers to WGS UTM Zone 14, with the Scratch folder. Provided the layers are named with the desired filename we can just add a prefix and use the Autofill Settings, Fill with Parameter Values to automate the output filenames.

Alternatively, for vector processing we can enable the Edit in Place function in the Toolbox. This enables input layers to be modified without creating new layers. So we could re-project layers, or here take the AOI layer and Rotate Features by 180 degrees. We can use the Undo function to revert to the original inputs as needed. Another option is to create a process model, defining inputs and algorithms for repeated tasks, such as this one here which reprojects and clips a layer to a common coordinate reference system and extent. We could then double left-click it in the Toolbox to run it individually or as a batch process in standardizing the spatial properties and the extent of analysis. We'll cover the Process Modeler in a later demo.

So most QGIS tools are run in the Background – meaning that other tasks can be completed while processing tools are running. This is not necessarily applied to GRASS or SAGA tools. So be patient – even when the program appears frozen - often tools are still running and will complete given the required processing time. However, there is no auto-save in QGIS – so ensure to save edits to layers, visualizations and project files frequently, especially prior to running processing-intensive tools. And if QGIS crashes while using a processing tool, the Toolbox Icon may disappear from the Attribute Toolbar when the program is reopened. Since it's a core plugin, it can be reloaded from

the Manage and Install Plugins box, opened from the Plugins drop-down. We can then check the Processing box off and on again to have the icon reappear.

The Plugins are another key component of QGIS, integrating user-created functions. And they can be installed and updated directly from this window when connected to the internet or loaded from a compressed folder if downloaded from the Online Repository. Note that non-core plugins may rely on additional dependencies and can also become deprecated between QGIS versions – in which case they are listed in red.

Now let's quickly discuss editing defaults within QGIS. To do so, expand the Settings drop-down and select Options. Note that any changes made here apply to all project files, and require restarting the program to take effect.

Within the General Tab, we can alter the interface language - specifying the language and locale – here having selected Canadian French. As we can see this translates most aspects of the interface, including tools and outputs accordingly. Back in the General Tab, below are additional defaults on system prompts and project parameters. In the Coordinate Reference System tab we can change the default Coordinate Reference System. We'll leave it as WGS84, as this is the most widely used Geographic Coordinate Reference System. We can also alter how the coordinate reference system is established when loading layers – using either the Default, Prompting for each Layer or using the Project Coordinate Reference System.

In the Data Sources tab we can alter the behaviour and formatting of the attribute table. We can specify which features are shown, the default view as either form or table, and the defaults for copying the

table. So, the default here includes Well Known Text which are the coordinates for the geometries of each feature. And this enables tables to be processed and analyzed externally, and reloaded in a spatial file format. However, if no further analysis in GIS was required or the data could be rejoined via another means such as unique identifiers - we could switch to plain text, no geometry to reduce times in exporting the table.

Rendering provides information on the defaults for visualizing vector and raster layers, such as geometry simplification for vectors and default rendering styles for rasters. The next four tabs enable edits the selection and colours for other map interaction tools, pre-defined colours and scales, and parameters for feature delineations.

Within the Processing tab, we can select the default file formats for raster and vector layers, how to address invalid geometries in a vector – here leaving it in its default - as well as the displayed information when running tools and the default output folder. In the Menus drop-down, we can customize the tools listed in the menu-bar drop-downs and on the toolbars. So to add it to the menu-bar, copy the Menu Path syntax from a tool already added and paste it to a tool of interest. Then to add it to a toolbar simply provide an icon and check the "Add button in toolbar" box. So here I created a custom toolbar with Geoprocessing tools, including Extract by Location – using the Snipping tool to extract icons from the toolbox. The toolbar can then be accessed once QGIS is restarted – here being shown in the French interface.

The Project Properties box contains similar parameters - but are specific to the active project file. It can be opened by clicking on the Project Coordinate Reference System button in the bottom right

corner of the interface. Within the General tab, we can switch the Save Paths from Absolute to Relative for saving layers, which will reduce complications when sharing project files and directories with others. We can also specify default visualizations for different geometry types. And within the Relation tab we can establish layer relations, with the Referencing layer containing 'many' entries - such as the Census Subdivision layer - and the referenced layer containing one matching entry – here using the Census Division layer - and linking them by the census division identifier field.

Finally, let's discuss some common problems and resolutions for processing layers. Most resolutions link back to the best practices we've discussed. The first thing to do is to consult the Log tab for targeting your trouble-shooting initiatives. For example, if it returns Invalid Geometries – run the layers through a cleaning tool such as Fix Geometries - and then rerun through the tool of interest with the fixed output. If errors persist tools such as Check Validity and Topology Checker can help identify errors, which can then be resolved with more advanced cleaning tools such as v.clean and Check Geometries. There are also case-specific tools such as Delete Holes and Remove Null Geometries, which can be applied as required. Less favourable is altering the default settings for Invalid Filtering to Ignore - since it does not address underlying issues and may yield inconsistencies in the outputs and analysis.

If the Log tab indicates a layer or folder cannot be found, ensure once again there are no spaces or special characters in the directories, subdirectories or filenames.

Inconsistencies in projections of input layers can also produce failures. And the differences will be shown by the differing EPSG codes after the layer names – in which case simply re-project the layers to the same system. If a geoprocessing error is returned, this may indicate that layers may differ in their type – specifically as single or multi-part, which relates to the number of features and corresponding entries in the attribute table. In this case, simply use the Multipart to Single Part or Promote to Multi-part tools to ensure conformity between the layers.

Finally, similar issues can occur with tools that require conformity or have constraints on accepted field types or file formats of input layers.

If related to differing field types we can use the Refactor Fields tool to ensure that the field types are the same. Otherwise, differences in common fields between layers can cause Join Attributes by Field Values, Merge and other tools to fail. Within the tool we can specify the field types, and length and precision parameters. In addition to linking layers together, it can also be used to correctly attribute a field type based on its content – such as changing a string field type with numeric variables to integer or double for use in the field calculator, interpolation tools or applying a graduated symbology.

If pertaining to the accepted geometry types: there's a variety of geometry conversion tools to switch to the desired type. Some relevant tools include Buffer to generate polygons from lines or points, Polygons to Lines or Points to Path for Lines, and Centroids and Extract Vertices to extract points. Some layers may require additional formatting to convert successfully. And broadly, Polygonise and Rasterize tools can be used for converting between raster and vector formats.

If pertaining to the vector format: Use the Export – Save As box to change to the desired file format, such as enabling file geodatabase layers to be edited and processed.

Otherwise, use a comparable tool within the Processing Toolbox. And if substitutes also fail, this indicates that the issue likely lies with the input datasets. However, we can also troubleshoot online, exploring GIS forums and other online documentation. Seldom will you be the first to encounter an issue, and these particular resources are fantastic means to identify any issues or known bugs being reported, and ultimately resolve any issues you may encounter.

And finally we can explore and install plugins as substitutes to perform a task of interest.

So using these best practices will facilitate navigating, loading, editing and visualizing multiple geospatial datasets in QGIS. Apply these practices to minimize potential errors, frustrations or repeating processes when using QGIS. As with any program save edits to layers, symbology styles and the project file frequently to avoid information loss should the program close unexpectedly.

(The words: "For comments or questions about this video, GIS tools or other Statistics Canada products or services, please contact us: statcan.sisagrequestssrsrequetesag.statcan@canada.ca" appear on screen.)

(Canada wordmark appears.)