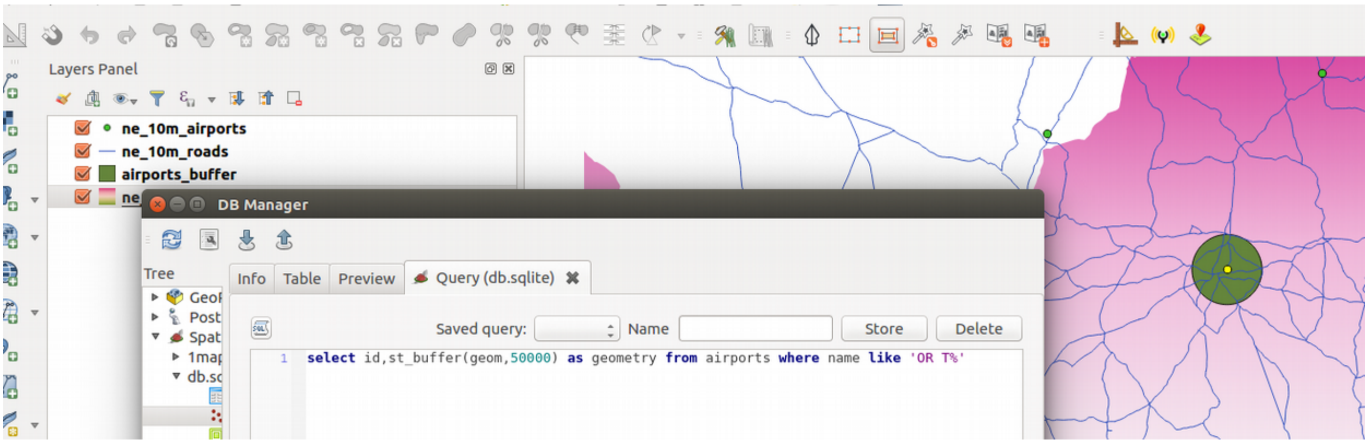
Section : 2. Data Providers

Module : 2.2. Databases (advanced)

Databases in Context

> *“Spatial databases are enhanced for storage and fast retrieval of data that defines a geometric space. The data is stored as coordinates, lines, points, polygon and topology”*

In this module, we will look at file-based databases and explore ways in which QGIS can interact with them. We will also learn about database syntax (SQL basics) and how databases can be utilized to do analysis ie geoprocessing. Databases can also store raster data. Storing raster data in a database has its own pros and cons. One distinct advantage is that analysis that combined vector data can be easily done whilst the main disadvantage is that rasters in nature tend to be big data and the upload process can take considerable time.



```
select id,st_buffer(geom,50000) as geometry from airports where name like 'OR T%'
```

You try:

Goal: To learn how to create and populate databases and run SQL queries

Problem: The government wants to expand OR Tambo international airport by a radius of 50km. Visualize the extent of the new airport.
Do all processing in the database.

- Load the listed layer in the table into QGIS.
- Use the algorithm `package layers` to create a geopackage database containing all the loaded layers.
- Navigate to the DB Manager.
- Create a connection for the geopackage you created above.
- Open the SQL window in DB Manager.
- Run the SQL `buffer layer expression`. Load the results into QGIS.
- Update the 'urban_precinct layer expression' by adding the `buffer layer expression` at the right location and run the SQL. Load the results into QGIS.

Can you modify the `urban_precinct layer expression` so that you can be able to calculate the expansion area ?

Name	Expectation

Layers from natural earth These layers are saved in the folder advanced-database which is downloaded as part of the zip file	ne_10m_airports, ne_10m_urban_areas, ne_10m_admin_0_countries
buffer layer Expression	select fid,st_buffer(geom,50000) geom from ne_10m_airports where name like 'OR T%'
urban_precinct layer expression	with buffer as (buffer layer expression) select a.fid,st_union(st_intersection(a.geom,b.geom)) as geom from buffer as a join ne_10m_urban_areas as b where st_intersects(a.geom,b.geom);

More about

Databases are very flexible in handling large amounts of data. They offer many advantages over other data sources and it is encouraged to use them as a data source. Spatial databases also store raster data. Spatial database can handle large volumes of data and can also be used for geoprocessing. The syntax that is used in the database is called SQL (Structured Query Language). This allows easy manipulation of the data and derivation of new products using the same data source.

GeoSpatial databases extend the standard set of functions you can use in your SQL queries to support 'spatial queries' - functions that specifically deal with spatial data. There is a very long list of spatial functions that can do everything from generating new geometries, cleaning geometries, calculating areas and distances, checking the relationship between geometries and so on.

You can create views in your database that save and replay queries. You can also add **query layers** in QGIS which are layers based on a specific SQL query and that can dynamically transform the data in your database as it is being loaded.

Check your knowledge:

1. A database is:
   a. *A storage device where GIS activities will take place*
   b. *A GIS database representing vector features.*
   c. *A type of data store that is used to store geographic data and non geographic*
2. In which instance would you prefer a spatial database:
   a. *When you have a a lot of CSV files or spreadsheets containing data for analysis*
   b. *When you need to do some advanced processing of the data.*
   c. *Generally, whenever I get hold of a data source I should load it into a spatial database.*
3. Why is it important to generate a spatial index for each layer loaded in the database:
   a. *To speed up rendering and fast retrieval of data when running SQL queries.*
   b. *To ensure that data is arranged in a chronological order using the primary key.*

Module video tutorial:

- https://youtu.be/7RvJ9Ej9kUs

Further reading:

- PrinterFriendly: http://www.bostongis.com/PrinterFriendly.aspx?content_name=spatialite_tut01
- Spatialite-3.0.0-BETA: https://www.gaia-gis.it/spatialite-3.0.0-BETA/spatialite-sql-3.0.0.html
- Spatialite: https://docs.qgis.org/3.16/en/docs/training_manual/databases/spatialite.html

Download the sample data for the lesson from http://changelog.qgis.org/media/images/lesson/worksheet/external_data/202daf4d3450d52a9241025ae5df5af7424c8606.zip.