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 spatialite databases and explore ways in which QGIS can interact with them. We will also learn about database syntax (SQL basics) and how to do simple geoprocessing.
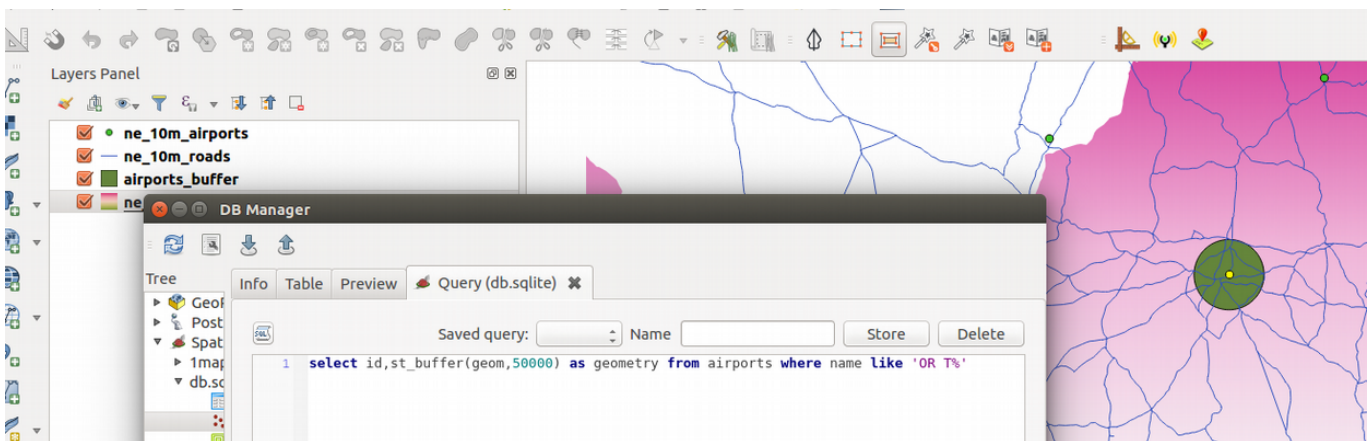
You try:

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.
- Create a new spatialite database.
- Filter the countries layer using to only show South Africa. Besides creating a filter what other methods can you use to identify South Africa ?.
- Perform a spatial selection to select urban areas and airports that are within South Africa.
  Hint Use Select by Location Algorithm
- Use the DB manager to load the filtered countries layer and the selected features from the airports and urban areas into the spatialite database.
  Hint: Use the import options specified and name the layers as countries, airports and urban_areas respectively.
- Open the SQL window in DB Manager.
- Create a table and call it **buffer** using the *buffer layer expression*. Load the resultant table in QGIS.
  What is the difference between a buffer done in the database and one done in QGIS ?.
- Create a table **urban_precinct** using the *urban_precinct layer expression* provided.
- Modify the *urban_precinct layer expression* to enable it to calculate the area affected by the expansion.
  What other methods can you use to calculate the area ?.

- Is it possible to combine the *buffer* and *urban_precinct* expressions to be a single expression. If possible write an expression that joins the two expressions into one ?.

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
Import Options	Toggle import selected, create spatial index, geom, reproject to target SRID 3857
buffer layer Expression	select id,st_buffer(geom,50000) as geometry from airports where name like 'OR T%'
urban_precinct layer expression	select a.id,st_union(st_intersection (a.geom,b.geom)) as geom from buffer as a join urban_areas as b where st_intersects (a.geom,b.geom)

ℹ 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.*

Further reading:

- Spatialite: http://docs.qgis.org/2.14/en/docs/training_manual/databases/spatialite.html
- Spatialite-3.0.0-BETA: https://www.gaia-gis.it/spatialite-3.0.0-BETA/spatialite-sql-3.0.0.html
- PrinterFriendly: http://www.bostongis.com/PrinterFriendly.aspx?content_name=spatialite_tut01

Download the sample data for the lesson from http://changelog.inasafe.org/media/images/lesson/worksheet/external_data/b8d0345e602396517b234e5e3a6c7f5b97a872df.zip.