**BI Analysis assignment**

Name: Tharanga Herath

Steps to installing the PostgressSQL Database with spatial data support and restore the give database backup for data analysis.

1. Installed PostgresSQL Database Version “**postgresql-9.5.3-1-windows-x64.exe**”.
2. Installed “**postgis**” 64bit version.
3. Create a new user Call **kademo**
4. Create a new database “bagdb” and assign the **kademo** user.
5. Create the extension **postgis**.
6. Restore the backup (**bag-amstelveen.backup**).

Note: refer to the below link for detail steps.

[**https://nlextract.readthedocs.io/en/latest/instructie-win-gui.html#postgresql-install**](https://nlextract.readthedocs.io/en/latest/instructie-win-gui.html#postgresql-install)

Step to install TileMill

1. Install the “**TileMill-v0.10.1-Setup.exe**” version
2. Follow the steps in setup. (Default setting will be sufficient)

**Prepare the environment for data analysis**

Loading new data set for the table “buildings” that is available in the new schema called “tilemill”

Data for the buildings table is extracted from the current schema “bagaveen”.

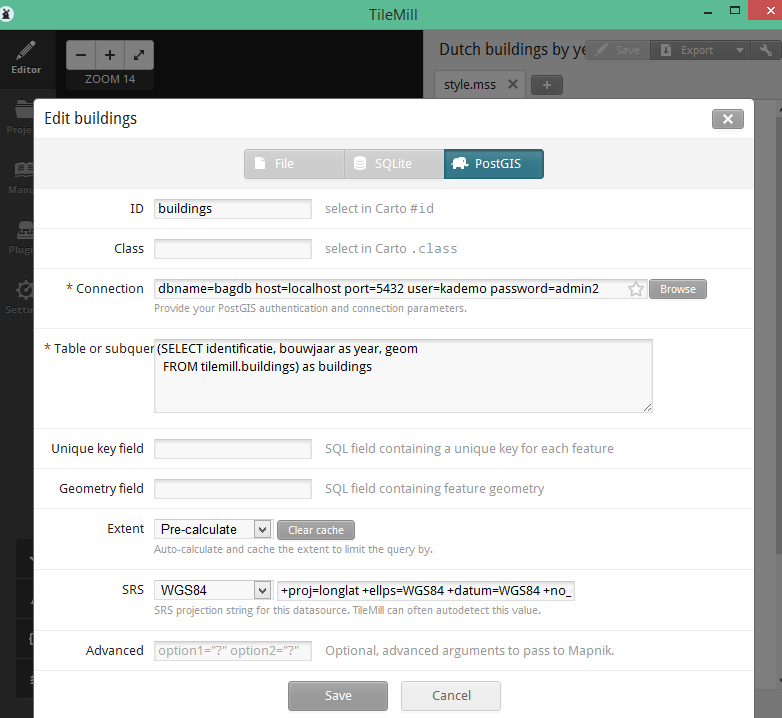
Later we are going to use the data in “buildings” in TileMill environment.

Script is available in the github scripts folder

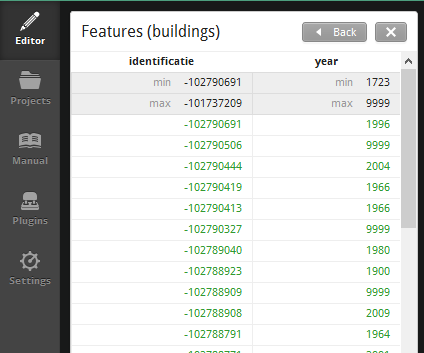
Script name : buildingdataextract.sql

**Setting up the TileMill project**

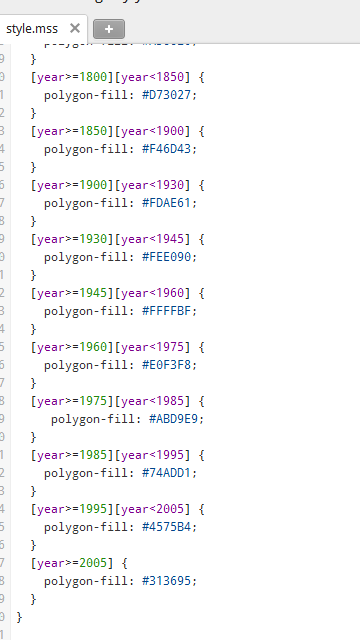
1. Create a project call “buldingmap”
2. Add a new layer call “buildings”
3. Add below information to pull the data from building table.

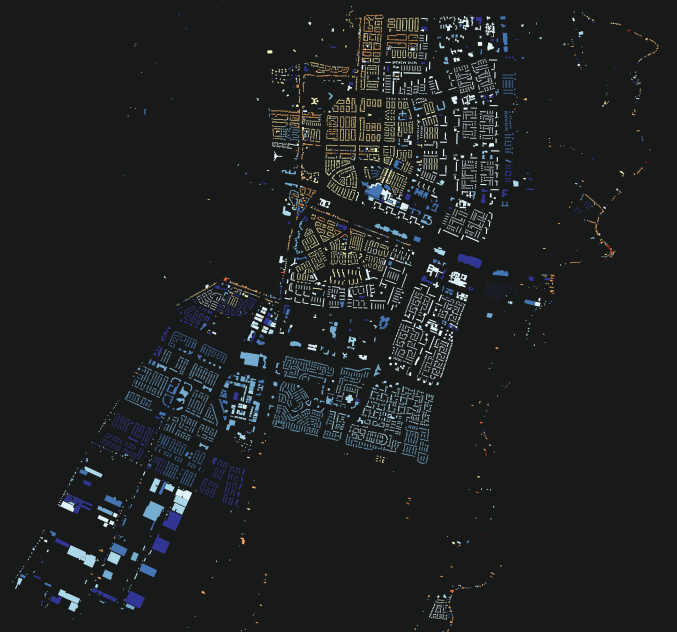


1. You can see the data in “features” option under layers.



1. Change the style.mms to plot the building map according to the year of construction.



1.  Building construction distribution in the plotted map.

Note:

TileMill project and the sql scripts are available in the “githhub” project

Conclusion

Above map show the building construction distribution over the past years in Netherlands.

Red and Orange areas shows the old building that are constructed before 1945 that is in the north-west side of the map. Looking at the area we can assume those building having historical value and won’t be any new development going to be initiated in this area due to archaeological value. To start any construction might need to take government approvals.

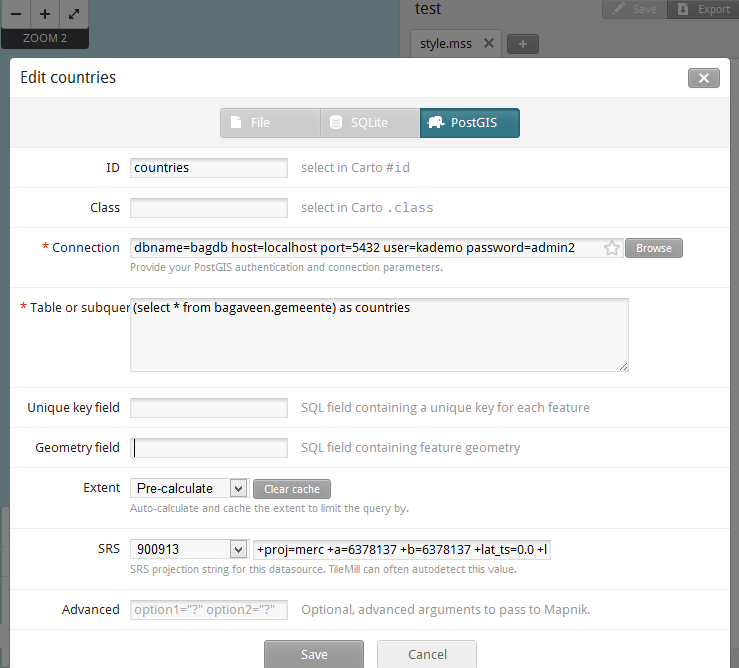
In the white area of the map contains buildings that are constructed between 1960 and 1985 and their might be an opportunity for start new projects. These building are bit old and may need to change or renovation. So people might sell these building and may move to developing areas that are mark in blue in the map. So these properties can be purchase and develop.

Arears mark in blue are the current developing areas, doing business in this territory might be challenging because most of the competitors are operating in this regions.

Areas indicating in black specially South-East side of the map we can’t see any construction. It might be a forest but it is useful to do a proper analysis about that area to identify any opportunities.

But before start any marking or strategic plan company need to do the SOWT analysis and a PEST analysis.

1. Make a connection to the postGIS



This is done just to check the connection.

dbname=bagdb host=localhost port=5432 user=kademo password=admin2

Then I went to the below link

<https://github.com/waagsociety/buildings>

Try to create the buildings table using below script

<script\_snip>

-- Aggregate mode function, to compute modal area and function

-- From: http://wiki.postgresql.org/wiki/Aggregate\_Mode

CREATE OR REPLACE FUNCTION \_final\_mode(anyarray)

RETURNS anyelement AS

$BODY$

SELECT a

FROM unnest($1) a

GROUP BY 1

ORDER BY COUNT(1) DESC, 1

LIMIT 1;

$BODY$

LANGUAGE 'sql' IMMUTABLE;

CREATE AGGREGATE mode(anyelement) (

SFUNC=array\_append, --Function to call for each row. Just builds the array

STYPE=anyarray,

FINALFUNC=\_final\_mode, --Function to call after everything has been added to array

INITCOND='{}' --Initialize an empty array when starting

);

CREATE SCHEMA tilemill;

CREATE TABLE tilemill.buildings AS

SELECT

p.identificatie::bigint, bouwjaar::int,

ST\_Transform(p.geovlak, 4326) AS geom,

round(mode(oppervlakteverblijfsobject)) AS oppervlakte,

mode(vg.gebruiksdoelverblijfsobject::text) AS gebruiksdoel

FROM verblijfsobjectactueelbestaand v

JOIN verblijfsobjectpandactueel vp

ON vp.identificatie = v.identificatie

JOIN pandactueelbestaand p

ON vp.gerelateerdpand = p.identificatie

JOIN verblijfsobjectgebruiksdoelactueel vg

ON v.identificatie = vg.identificatie

GROUP BY

p.identificatie, bouwjaar, p.geovlak;

CREATE INDEX buildings\_geom\_idx

ON tilemill.buildings

USING gist (geom);

</script\_snip>

And It throw the below error to me

/\*\*\*\*

ERROR: relation "verblijfsobjectactueelbestaand" does not exist LINE 29: FROM verblijfsobjectactueelbestaand v ^ \*\*\*\*\*\*\*\*\*\* Error \*\*\*\*\*\*\*\*\*\* ERROR: relation "verblijfsobjectactueelbestaand" does not exist SQL state: 42P01 Character: 879

\*\*\*\*\*\*\*\*\*\*\*\*/

I think the restore database and the script is referring to different schema appreciate if you can help me on this,