Creating and Loading a Spatial Database

1. Install psql and postGIS

- **Mac**

Postgres.app

- **Linux**

- PostgreSQL
- PostGIS

2. **Create a database and connect to it**

```
~> createdb --owner lucia --password nyc
[ createdb -0 lucia -W nyc ]


~> psql --dbname=nyc
[ Or \c nyc from psql prompt ]
[ \l to list all ]
```

3. **Enable spatial support**

```
=# CREATE EXTENSION postgis;

=# SELECT postgis_full_version();
```

4. Create sql load commands from shapefiles

```
~> mkdir sqlloads


~> shp2pgsql -I -s 26918 data/nyc_census_blocks.shp nyc_census_blocks
> sqlloads/nyc_census_blocks.sql


~> shp2pgsql -I -s 26918 data/nyc_neighborhoods.shp nyc_neighborhoods
> sqlloads/nyc_neighborhoods.sql


~> shp2pgsql -I -s 26918 data/nyc_streets.shp nyc_streets >
sqlloads/nyc_streets.sql


~> shp2pgsql -I -s 26918 data/nyc_subway_stations.shp
nyc_subway_stations > sqlloads/nyc_subway_stations.sql
```

Wait, what?

A shapefile is an vector data storage format for storing the location, shape, and attributes of geographic features.

It’s actually a collection of files:

- .shp → shape (geometry)
- .shx → index
- .dbf → attributes
- [.prj] → projection

Wait, what? (2)

SRID [Spatial Reference System Identifier] is a unique value that unambiguously specifies the projection and coordinates used to arrive to the given geometries

Two geometries with different SRIDs cannot be compared

SRID 4326 corresponds to the World Geodetic System → standard coordinate system for the Earth (in lat/lon) with with a standard spheroidal reference surface and origin at the Earth's center of mass. The error is believed to be less than 2cm

5. **Load spatial data into db**

```
~> psql -U lucia -d nyc -f sqlloads/nyc_census_blocks.sql

~> psql -U lucia -d nyc -f sqlloads/nyc_neighborhoods.sql

~> psql -U lucia -d nyc -f sqlloads/nyc_streets.sql

~> psql -U lucia -d nyc -f sqlloads/nyc_subway_stations.sql

~> psql -U lucia -d nyc -f data/nyc_census_sociodata.sql
```