**ETL project proposal**

**Extract:**

Data Source: New York open data source:

* Demographic data with borough (csv)
* Parking ticket (csv)
* Complaints (csv)
* Jobs (csv)
* Crime (csv)
* Pollution (csv)

Junction tables:

* Precincts-borough
* borough -zip

**Transformation:**

Data will be cleaned for:

* NAN,
* Unwanted data columns
* Normalization

**Load:**

There will be 6 tables and 2 junction tables. Data will be loaded in PostgreSQL and queries run on it.