

Angela Wang  
Data & Databases  
Final Project  
6/28/18

### Guided Project 4: NYC Housing Data

1. install the housing data database locally and Postgres
2. perform queries on the database to get a sense of what is in there, and develop a research question
3. write a one paragraph research question which explains what your central query or queries are, why you have decided on these queries, and how the results of these queries can be mapped.
4. find external sources that add information to the resulting table(s) that you have extracted from the database--these sources can be other data resources to scrape/join, as well as articles that might be included in your map, or research that has deepened your understanding of the topic.
5. submit the table(s) of your central queries, links to further research or data, and a written summary of how and why you plan to map the information you have discovered.

According to census data, the residents of Brooklyn's poorest tract, located in Brownsville, have a [median household income](#) of \$11,220. Just miles to the west, in DUMBO and Brooklyn Heights — often described as Brooklyn's priciest neighborhoods — residents have median household incomes [far exceeding](#) six figures.

I'm interested in how 311 complaints might differ across the most and least affluent neighborhoods in Brooklyn. For my project, I will examine complaints filed in two Brooklyn districts since 2014: Community Board 2 (Brooklyn Heights, DUMBO, Fulton Mall, Boerum Hill, Fort Greene, Brooklyn Navy Yard, Fulton Ferry, and Clinton Hill), Community Board 16 (Brownsville and Ocean Hill). I will look at complaint data from the Department of Housing Preservation and Development, which concerns conditions and maintenance issues in rental units, as well as complaint data from the Department of Buildings, which includes illegal construction work and building use. If time allows, I would also like to examine and map a rapidly gentrifying neighborhood: Community Board 4 (Bushwick).

I plan to create a map for each neighborhood (or one larger map with delineated districts), showing the the nature of the complaints in each district. Beyond revealing these patterns, I hope to be able to identify specific problematic buildings or areas.

#### Research sources:

NYCDB datasets:

- hpd\_complaints
- dob\_complaints
- pluto\_17v1

Detailed HPD complaint info from Open Data:

<https://data.cityofnewyork.us/Housing-Development/Complaint-Problems/a2nx-4u46>

<https://censusreporter.org/profiles/79500US3604002-nycbrooklyn-community-district-4bushwick-puma-ny/>

<http://ibo.nyc.ny.us/cgi-park/?p=740>

## **Central SQL Queries:**

### **HPD complaints in DUMBO:**

```
CREATE TABLE hpd_complaints_dumbo AS
SELECT hpd_complaints.complaintid, hpd_complaints.housenumber,
hpd_complaints.streetname, hpd_complaints.zip, hpd_complaints.apartment,
hpd_complaints.bbl, hpd_complaints.communityboard, hpd_complaintdetails.majorcategory,
hpd_complaintdetails.minorcategory, hpd_complaintdetails.code,
hpd_complaintdetails.statusdate,
hpd_complaintdetails.statusdescription, pluto_17v1.lat, pluto_17v1.lng
FROM hpd_complaints
LEFT JOIN pluto_17v1 ON hpd_complaints.bbl = pluto_17v1.bbl
JOIN hpd_complaintdetails ON hpd_complaints.complaintid = hpd_complaintdetails.complaintid
WHERE hpd_complaints.borough = 'BROOKLYN'
AND hpd_complaints.communityboard = 2;
```

*Rows in table: 18985*

### **HPD complaints in Brownsville:**

```
CREATE TABLE hpd_complaints_dumbo AS
SELECT hpd_complaints.complaintid, hpd_complaints.housenumber,
hpd_complaints.streetname, hpd_complaints.zip, hpd_complaints.apartment,
hpd_complaints.bbl, hpd_complaints.communityboard, hpd_complaintdetails.majorcategory,
hpd_complaintdetails.minorcategory, hpd_complaintdetails.code,
hpd_complaintdetails.statusdate,
hpd_complaintdetails.statusdescription, pluto_17v1.lat, pluto_17v1.lng
FROM hpd_complaints
LEFT JOIN pluto_17v1 ON hpd_complaints.bbl = pluto_17v1.bbl
JOIN hpd_complaintdetails ON hpd_complaints.complaintid = hpd_complaintdetails.complaintid
WHERE hpd_complaints.borough = 'BROOKLYN'
AND hpd_complaints.communityboard = 2;
```

*Rows in table: 40927*

#### **HPD complaints in Bushwick:**

```
CREATE TABLE hpd_complaints_bushwick AS
SELECT hpd_complaints.complaintid, hpd_complaints.housenumber,
hpd_complaints.streetname, hpd_complaints.zip, hpd_complaints.apartment,
hpd_complaints.bbl, hpd_complaints.communityboard, hpd_complaintdetails.majorcategory,
hpd_complaintdetails.minorcategory, hpd_complaintdetails.code,
hpd_complaintdetails.statusdate,
hpd_complaintdetails.statusdescription, pluto_17v1.lat, pluto_17v1.lng
FROM hpd_complaints
LEFT JOIN pluto_17v1 ON hpd_complaints.bbl = pluto_17v1.bbl
JOIN hpd_complaintdetails ON hpd_complaints.complaintid = hpd_complaintdetails.complaintid
WHERE hpd_complaints.borough = 'BROOKLYN'
AND hpd_complaints.communityboard = 4;
```

*Rows in table: 55310*

#### **Active DOB complaints in DUMBO:**

```
CREATE TABLE dob_complaints_dumbo AS
SELECT complaintcategory, housenumber, housestreet, bin, dispositioncode
FROM dob_complaints
WHERE communityboard = 302
AND status = 'ACTIVE'
GROUP BY complaintcategory, housenumber, housestreet, bin, dispositioncode
ORDER BY complaintcategory DESC;
```

Rows in table: 8061

#### **Active DOB complaints in Brownsville:**

```
CREATE TABLE dob_complaints_brownsville AS
SELECT complaintcategory, housenumber, housestreet, bin, dispositioncode
FROM dob_complaints
WHERE communityboard = 316
AND status = 'ACTIVE'
GROUP BY complaintcategory, housenumber, housestreet, bin, dispositioncode
ORDER BY complaintcategory DESC;
```

Rows in table: 4339

**Active DOB complaints in Bushwick:**

```
CREATE TABLE dob_complaints_bushwick AS
SELECT complaintcategory, housenumber, housestreet, bin, dispositioncode
FROM dob_complaints
WHERE communityboard = 304
AND status = 'ACTIVE'
GROUP BY complaintcategory, housenumber, housestreet, bin, dispositioncode
ORDER BY complaintcategory DESC;
```

Rows in table: 10413