# OpenStreetMap MongoDB Project

Seth Weidman

## Area chosen

I chose to analyze a strip consisting of about two thirds of the city of Chicago, including almost all the places I’ve roamed during the seven years I’ve lived here, excluding the far west and south sides.

## Problems encountered in the map

The only problem I fixed was the one fixed in the exercise: the fact that some street names are abbreviated, rather than fully written out. I fixed this by completing the “audit\_street\_types.py” code, attached.

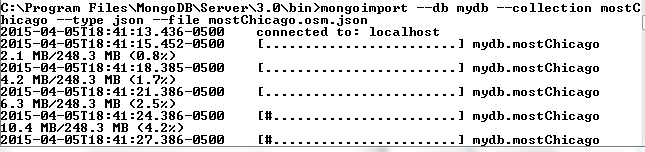
## Data overview

The file sizes were as follows:

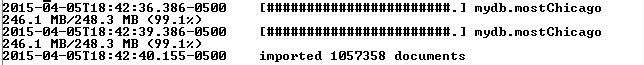
mostChicago.osm.json: 254 MB

mostChicago.osm: 228 MB

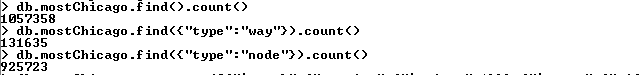
To import the data, from the command line, I used:



…



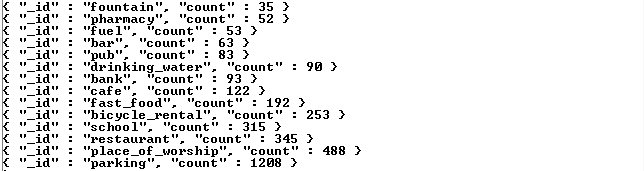
I ran some queries to determine the number of documents, then counted how many of each “type” there were:



I then ran a query to find the most popular “amenity” types in the city:



The ones that came out on top were:



I did a similar ranking to see which cuisines were most popular. Unsurprisingly, in Chicago, the setting for the famous “CHEEZBURGER CHEEZBURGER CHEEZBURGER” SNL sketch, burgers come out on top.



## 

## Conclusion

The data seems to be a fair representation of what the city of Chicago, at least the parts I’ve experienced, are actually like. OSM data has the potential to provide a ton of insights for cities; however, there should be better standards for how to classify different places on OSM – for example, what kind of spot qualifies as “parking” – any space on the street, or a space in a parking garage? With a consistent definition, (and clean data which would be a whole other step), cities could construct rankings of the availability of parking in different neighborhoods - or data wranglers could figure it out themselves!

Another, probably more minor, problem that better definitions for the “amenity” fields would solve is to define what qualifies as a “pub” vs. a “bar” e.g. does a bar have more features, such as a dance floor or pool tables? If you wanted to use OSM data to search for “bars” in a particular neighborhood, you might currently miss some perfectly good options that are just classified as pubs.

These inconsistent definitions prevent the OSM Chicago data from living up to its full potential.