Map Area:

Quincy, Massachusetts, United States

https://www.openstreetmap.org/export#map=12/42.2623/-71.0159

The reason that I chose this place is that I've always admired the founding fathers like Washington, Adams, Jefferson, and especially Jefferson. Seeing that there is no chance that I can get a dataset representing Monticello, instead I chose Quincy, the birthplace of the Adamses.

Problems encountered in the map
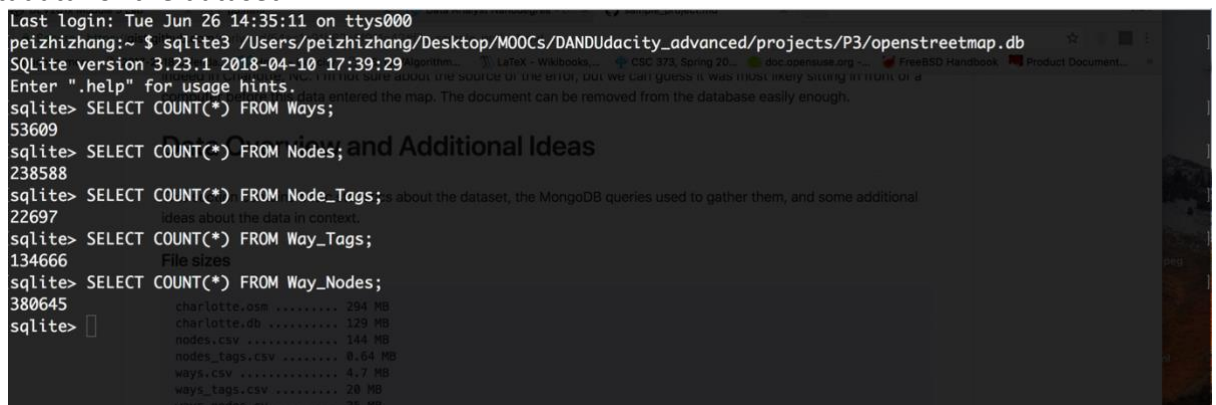
1. Invalid tag type and value

As further query shows, there are multiple tags with unmeaningful types and values, like for example, there is a tag with type 'noref' and value '4'. Without any gold standard data, there is no appropriate method to remedy this problem

Overview

1. Running ls -al on the files

```
drwxr-xr-x  18 peizhizhang  staff       576 Jun 26 16:53 .
drwxr-xr-x  12 peizhizhang  staff       384 Jun 12 17:58 ..
-rw-r--r--@  1 peizhizhang  staff      6148 Jun 25 16:27 .DS_Store
drwxr-xr-x   7 peizhizhang  staff       224 Jun 25 22:23 .idea
-rw-r--r--   1 peizhizhang  staff         0 Jun 12 15:52 README.md
drwxr-xr-x   3 peizhizhang  staff        96 Jun 22 18:17 __pycache__
-rw-r--r--@  1 peizhizhang  staff  64174633 Jun 21 15:38 data.osm
-rw-r--r--   1 peizhizhang  staff      6301 Jun 24 15:10 data.py
-rw-r--r--   1 peizhizhang  staff         0 Jun 26 16:53 file_sizes.txt
-rw-r--r--@  1 peizhizhang  staff  20156519 Jun 24 15:41 nodes.csv
-rw-r--r--@  1 peizhizhang  staff    879470 Jun 24 15:41 nodes_tags.csv
-rw-r--r--   1 peizhizhang  staff  33996800 Jun 24 20:51 openstreetmap.db
-rw-r--r--@  1 peizhizhang  staff     12149 Jun 25 22:30 p3.docx
-rw-r--r--   1 peizhizhang  staff      2313 Jun 22 16:54 schema.py
-rw-r--r--@  1 peizhizhang  staff   3254491 Jun 24 15:41 ways.csv
-rw-r--r--@  1 peizhizhang  staff   9101978 Jun 24 15:41 ways_nodes.csv
-rw-r--r--   1 peizhizhang  staff   4552938 Jun 24 15:41 ways_tags.csv
-rw-r--r--@  1 peizhizhang  staff       162 Jun 26 14:27 ~$p3.docx
```

2. metadata for the dataset

Top 10 contributing users

```
WesWeaver|114356
eugenebata|46990         UNT(*)
peace2|7532          e.user, COUNT(*
dchiles|6345         ELECT user FROM
woodpeck_fixbot|4746
chrmur|4268          num-1)  u;
Ann Khroshun|4121
choess|3370
dankpoet|2874
MSoslow|2848                   s
sqlite> SELECT SUM(count) as sum FROM (SELECT user, COUNT(*) as count FROM Nodes GROUP BY user ORDER BY count DESC LIMIT 10)
sub;
197450
sqlite> SELECT COUNT(*) FROM Nodes;
238588
```

Additional Ideas

1. More detailed info about the data may be required, like for restaurant nodes, land area of the restaurants can be vital.
2. After careful data exploration, I find that top ten contributing users make up about 82.8% of the dataset. Thus I suggest that some inciting measure must be undertaken in order for more users to contribute.