FILES AND CONTENT

1.sample\_file.py – The code to generate sample file from the original file.

2.tag\_count1.py—The code to generate counts of different types of tags in the xml file.

3.types\_of\_tags2.py—The code to generate tags with lowercolon,lower,problem characters

4.phone\_audit\_clean\_write.py—The code to audit,clean phonenumbers.The code is also used to write a new xml file with updated phone numbers.

5.post\_audit\_clean\_write.py-- The code to audit,clean postal codes.The code is also used to write a new xml file with updated postal codes.

6.street\_audit\_clean\_write.py-- The code to audit,clean postal codes.The code is also used to write a new xml file with updated postal codes.This file is the final xml file.

7.schema.py- schema for writing csv

8.mapparser.py—parses all elements and converts into tabular format for being written into csvs

9.data\_overview.py—gives file sizes

10.builiding\_databse.py—code to import all csv into database tables for sql querying

The files must be run in the above mentioned order.

London.pdf—ALL ANSWERS REGARDING THE PROJECT

References- References

Dataset details—details about the osm file.

please download the osm file from thelink in dataset details file and save it as 'london\_england.osm' .The python scripts must be run in the order which is specified in the readme file