Project 2 - Data Wrangle OpenStreetMaps Data - Additional Info

This document serves as a guide to the evaluator as to the location and meaning of the various project files:

There are two folders in the project:
- "**lesson6**" - This contains all the source code from the Lesson 6 problems in the course.
- "**p2**" - This folder contains the following elements:
  - **"project2_report.html"** - The report file for evaluation.
  - **"project2_references.pdf"** - File containing a list of reference material and websites used in the compilation of the report.
  - **"sample data"** - folder containing the a minimised JSON file selecting 1/12 records.
  - **Python source code**:
    - **"project2_constants.py"** - Contains all constants for use by other modules like "expected_street_types", etc/
    - **"project2_explore.py"** - The methods to parse and investigate the raw OSM XML file.
    - **"project2_process.py"** - used to generate the JSON file and employ the data-cleaning strategies identified during exploration.
    - **"project2_mongo_insert.py"** - used to parse the JSON file and insert the JSON objects into the MongoDB database.
    - **"project2_query_mongo.py"** - methods used to test and build queries using MongoDB's Aggregation framework.

The area of OpenStreetMap can be accessed through the following link:

http://www.openstreetmap.org/search?query=cape%20town#map=13/-33.9402/18.4614