



## PROJECT

## Wrangle OpenStreetMap Data

A part of the Data Analyst Nanodegree Program

## PROJECT REVIEW

## CODE REVIEW

## NOTES

SHARE YOUR ACCOMPLISHMENT!  

## Meets Specifications

## Code Functionality

Final project code functionality reflects the description in the project document.

The code reflects the description in the report document.

## Code Readability

Final project code follows an intuitive, easy-to-follow logical structure.

Final project code that is not intuitively readable is well-documented with comments.

The code is well formatted and appropriately commented. That make it easy to follow the analysis steps and identify a specific functional operation.

## Problems encountered in your map

Student response shows understanding of the process of auditing, and ways to correct or standardize the data, including dealing with problems specific to the location, e.g. related to language or traditional ways of formatting.

You did a good job standardized street type and names and postal code. It seems that you understand the process of auditing, and know ways to correct or standardize the data. Please consider expanding the "Problems encountered in your Map" further. You might want to include examples of other elements ( Email, phone numbers, URL, etc) that were not properly shaped or formatted and discuss ways in which they can be fixed programmatically. Were you able to find elements that are specific to Rio de Janeiro? Different language?

Some of the problems encountered during data audit are cleaned programmatically.

The problems that found in the dataset are corrected or cleaned programmatically.

## Overview of the data

The OSM XML file is at least 50 MB uncompressed.

Database queries are used to provide a statistical overview of the dataset, like:

- size of the file
- number of unique users
- number of nodes and ways
- number of chosen type of nodes, like cafes, shops etc.

Additional statistics not in the list above are computed. For SQL submissions some queries make use of more than one table.

The analysis makes use of different queries to present a meaningful overview about the dataset.

The submission document includes the database queries and statistics from above.

Other ideas about the dataset

Submission document includes one or more additional suggestions for improving the data or its analysis. The suggestions are backed up by at least one investigative query.

You bring up many other ideas for improvements here, that is excellence.

Submission document includes thoughtful discussion about the benefits as well as some anticipated problems in implementing the improvement.

The discussion about the other ideas includes the benefits as well as some anticipated problems in implementing the improvement.

Thoroughness and Succinctness of Submission

Submission document is long enough to thoroughly answer the questions asked without giving unnecessary detail. A good general guideline is that the question responses should take about 3-6 pages.

[↓ DOWNLOAD PROJECT](#)

[RETURN TO PATH](#)

[Student FAQ](#)