

Métodos e ferramentas

A base dos dados explorados neste projeto é disponibilizada pelo OpenStreetMap e foi extraída com a ajuda da ferramenta API Overpass com as coordenadas -44.3092,-20.0979,-43.6116,-19.7376 que representam a área de Belo Horizonte.

O arquivo foi exportado no formato OSM/XML e foi gerada uma amostra desta base com menos informação para facilitar e agilizar a leitura das informações na fase da auditoria. Script

executado: https://github.com/liebycardoso/OpenStreetMap/blob/master/Get_Sample_OSM.py

Antes de iniciar a análise dos dados executei o script Summary_OSM para identificar caracteres problemáticos, total de chaves e total de aparições. Script

executado: https://github.com/liebycardoso/OpenStreetMap/blob/master/Summary_OSM.py

A base (map_BH.osm) com todas as informações teve seu conteúdo analisado e os apontamentos e decisões tomadas estão documentadas na seção 2 deste documento. Após auditoria dos dados, e análise dos registros, um dicionário com os valores atualizados foi gerado no formato .JSON. Script

executado: https://github.com/liebycardoso/OpenStreetMap/blob/master/parse_duplicate_json.py. Script

auxiliar: https://github.com/liebycardoso/OpenStreetMap/blob/master/audit_map_BH.py

Todo o conteúdo do arquivo JSON foi importado através de linha de comando para a coleção mapbh no banco de dados MongoDB.

Foram executadas consultas aos dados importados e as novas inconsistências foram corrigidas conforme descrito neste relatório.