

Utilizando dados do IBGE para melhorar o Brasil no OpenStreetMap

Nelson A. de Oliveira

naoliv@gmail.com

Agenda

- Base Cartográfica Contínua do Brasil
- Faces de logradouros
 - Mapas do censo agropecuário / contagem da população
- Folhas topográficas
- Conclusão

Quem sou eu?

- Formação em computação
- Trabalho em uma universidade como administrador de sistemas
- Entusiasta de software livre há ~16 anos
- 5 anos editando no OpenStreetMap
 - Pato Donald bravo
 - Estatísticas de contribuição no OSM <http://hdyc.neis-one.org/?naoliv>
 - Onde contribuo <http://yosmhm.neis-one.org/?naoliv>
- "Reclamão" também encaixa na descrição, segundo o meu irmão

Base Cartográfica Contínua do Brasil

BC250

- Acessível através de
ftp://geoftp.ibge.gov.br/cartas_e_mapas/bases_cartograficas_continuas/bc250/versao2015/
(ou goo.gl/gm5LK8)
- Escala de 1:250000 (1 cm = 2.5 km)
- Abrange todo o Brasil
- Disponível em shapefiles (agrupados em algumas categorias) ou como um dump de postgis

BC250

Abrange:

- Unidades de proteção
- Unidades de conservação
- Terras indígenas
- Trechos rodoviários
- Pistas de pouso
- Massas d'água
- Ilhas
- Barragens
- Usinas de energia
- Localidades
- etc

Localidades do BC250

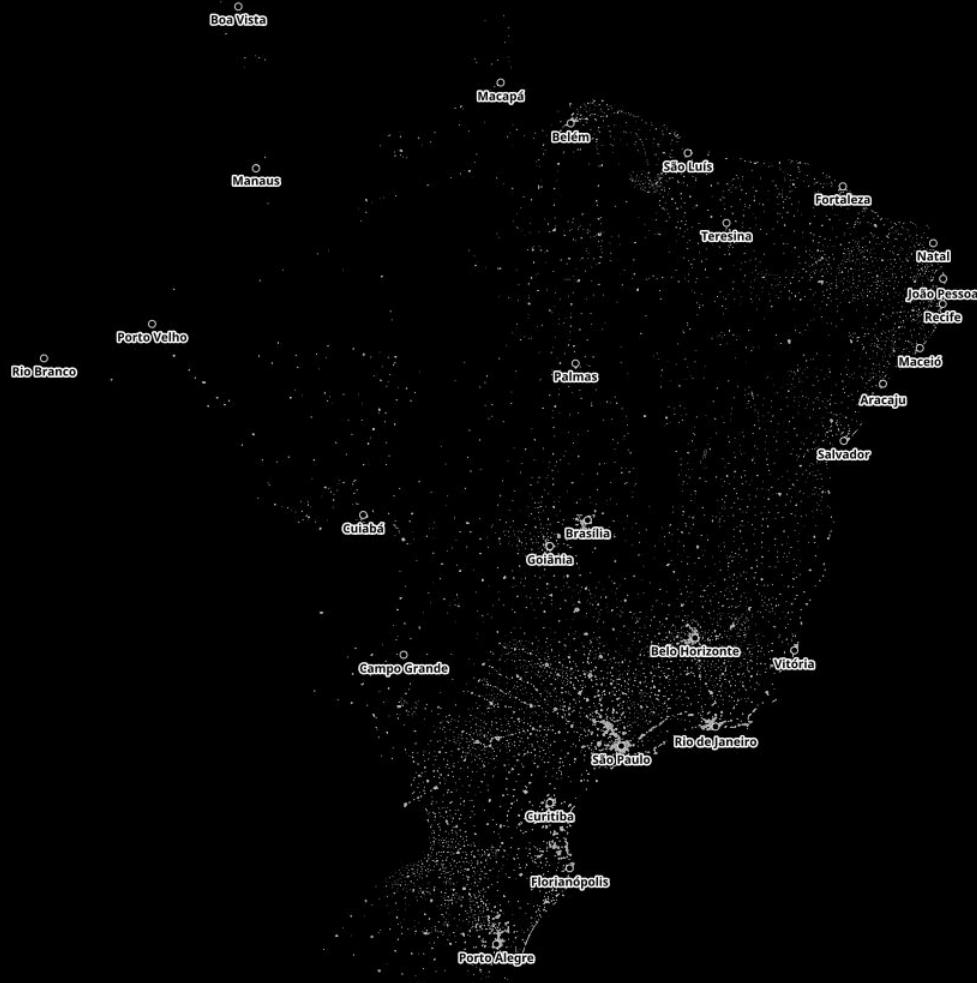
Possui:

- Capitais
- Cidades
- Vilas
- Aglomerados rurais isolados
- Habitações indígenas
- Áreas edificadas ("mancha urbana")

Visualizando

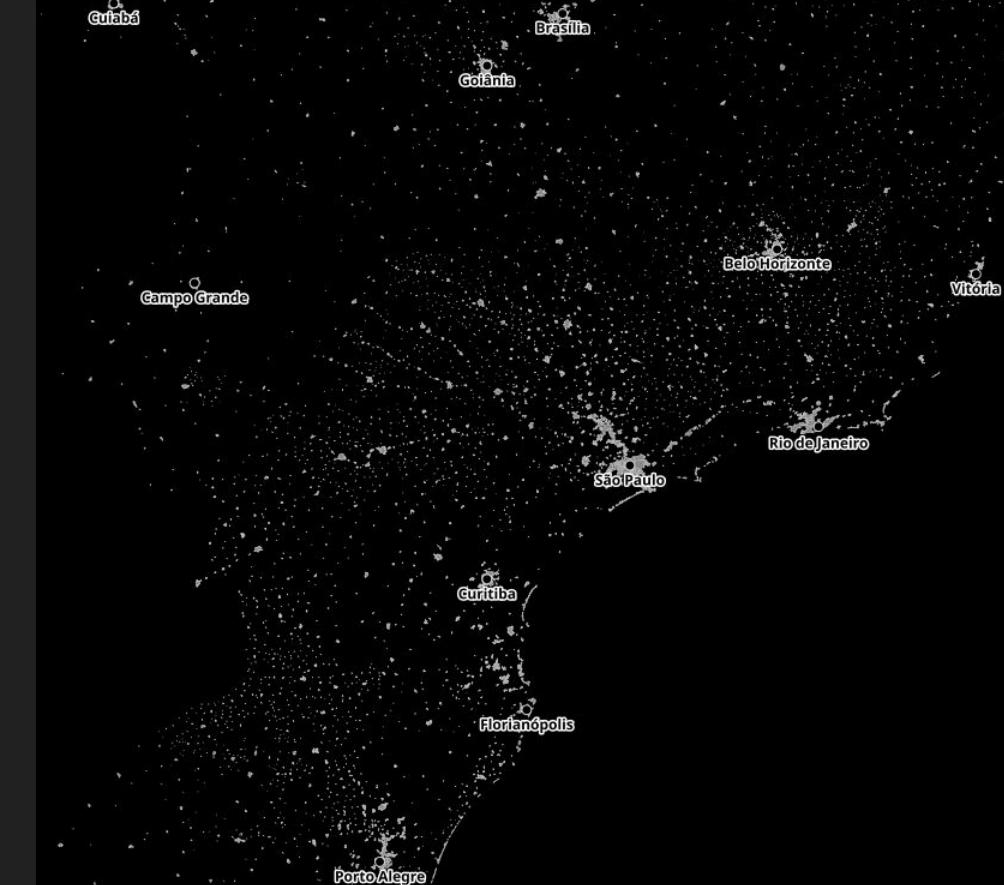
Os locais foram agrupados, classificados e inseridos em um banco de dados

Prototipada uma representação e depois disponibilizado como TMS (Tile Map Service)



Visualizando

Em níveis de zoom menores
(área de visualização maior)
apenas as capitais são
rotuladas



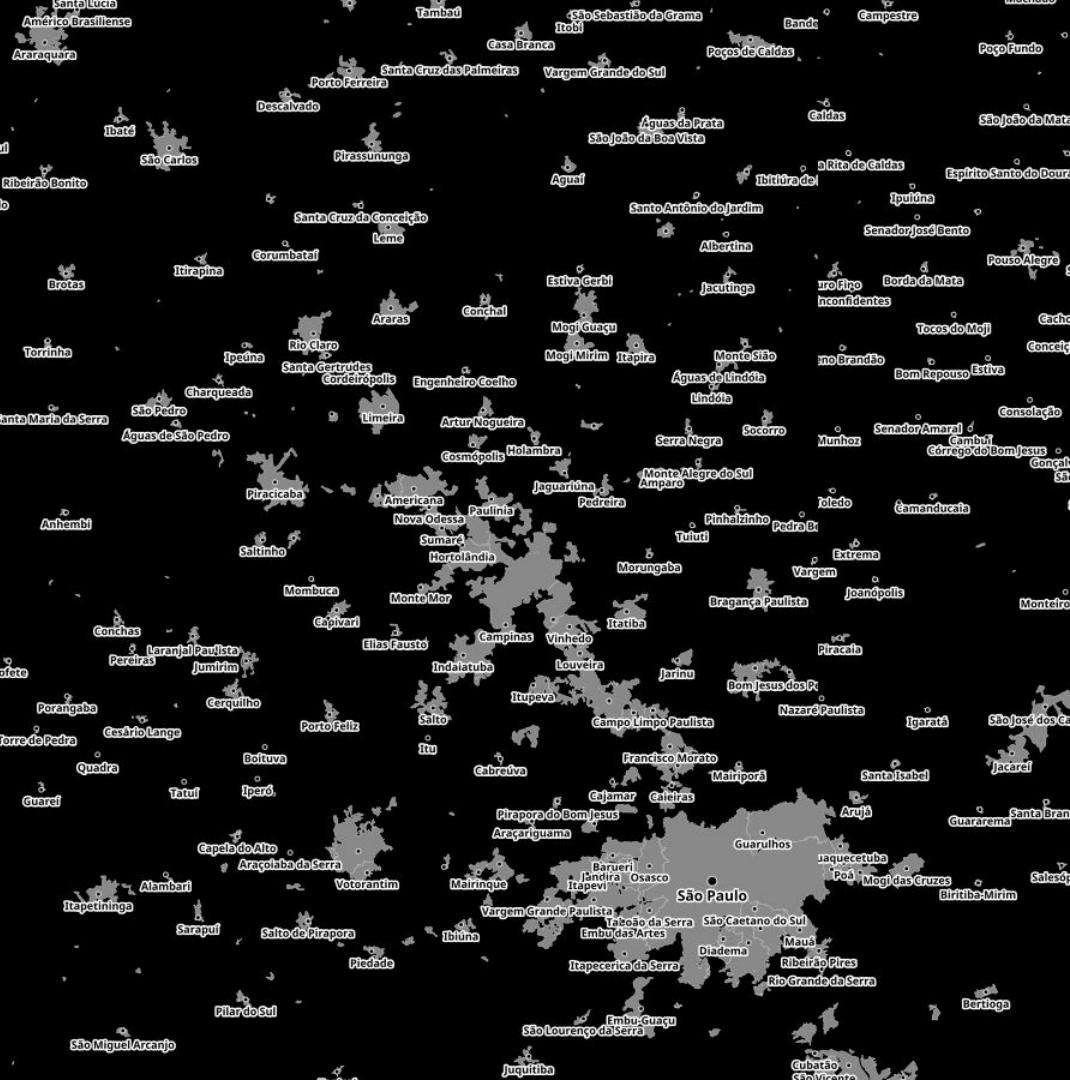
Visualizando

Manchas urbanas estão
presentes em todos os níveis
de zoom



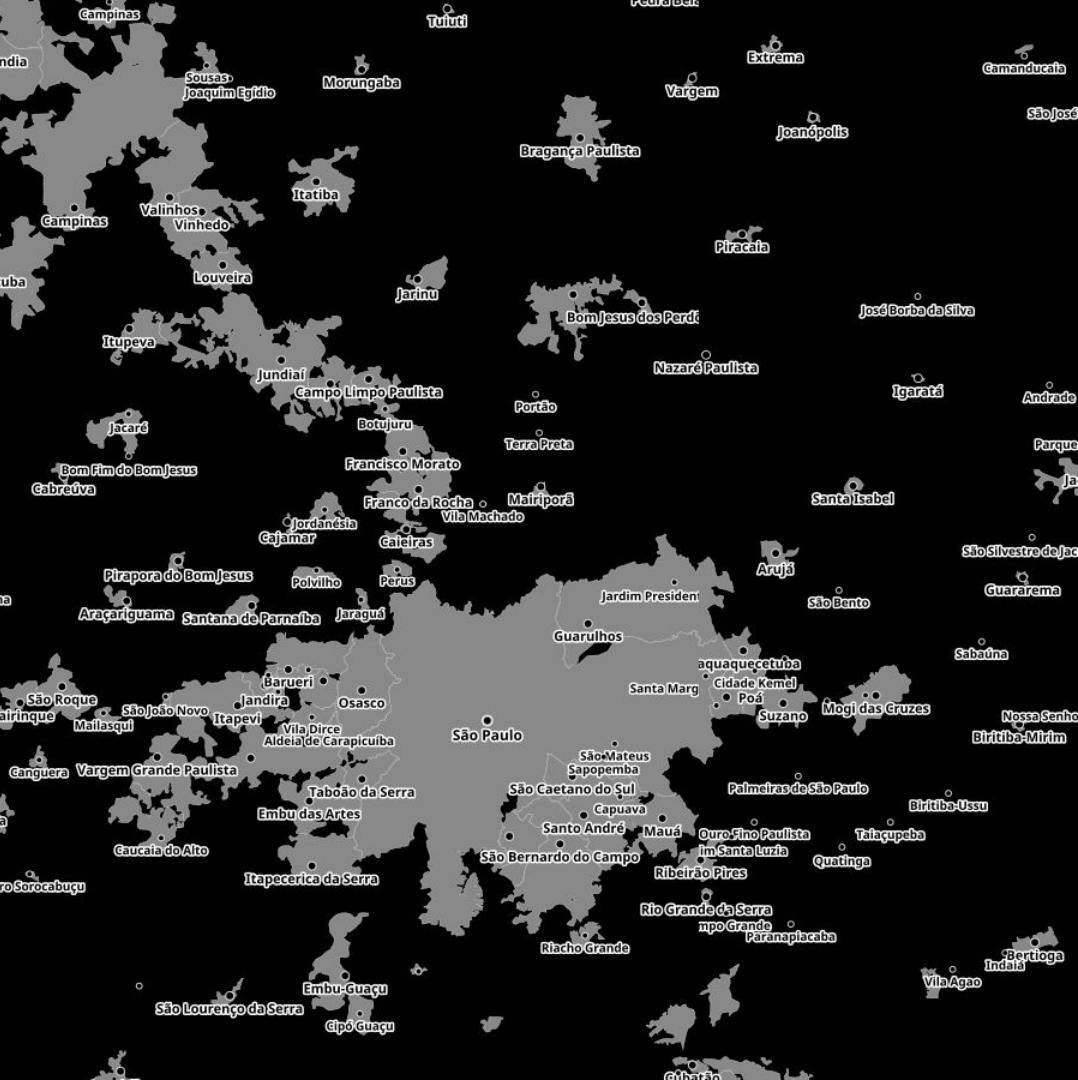
Visualizando

Rótulos de cidades a partir do nível de zoom 8



Visualizando

Vilas, povoados,
aglomerações rurais e áreas
indígenas aparecem
proporcionalmente com níveis
maiores de zoom



Visualizando



Visualizando

Identificadores dos locais a partir do nível de zoom 12

Muito útil para ter uma ideia aproximada do tipo de local



Visualizando

Razoavelmente bom com
várias camadas de fundo,
como o Bing



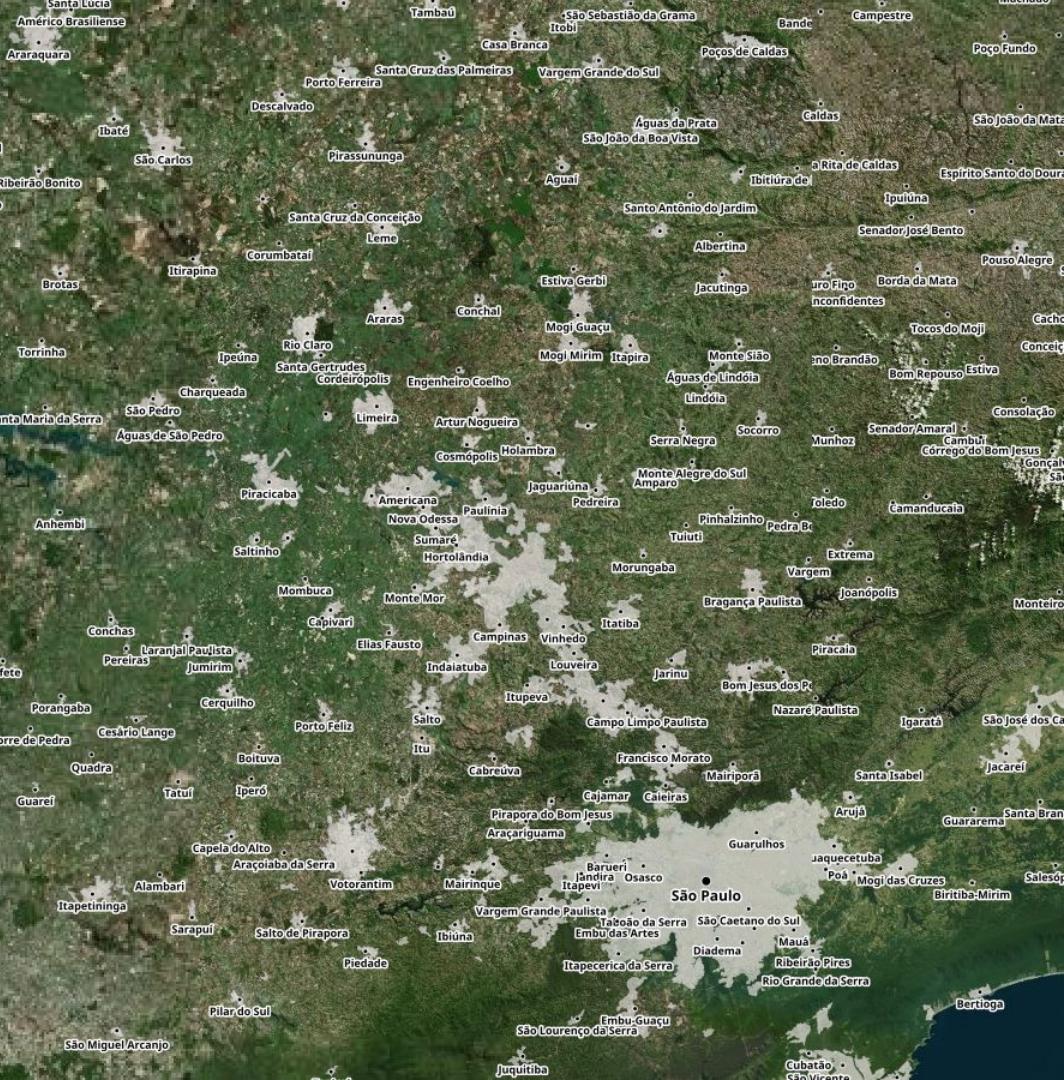
Visualizando



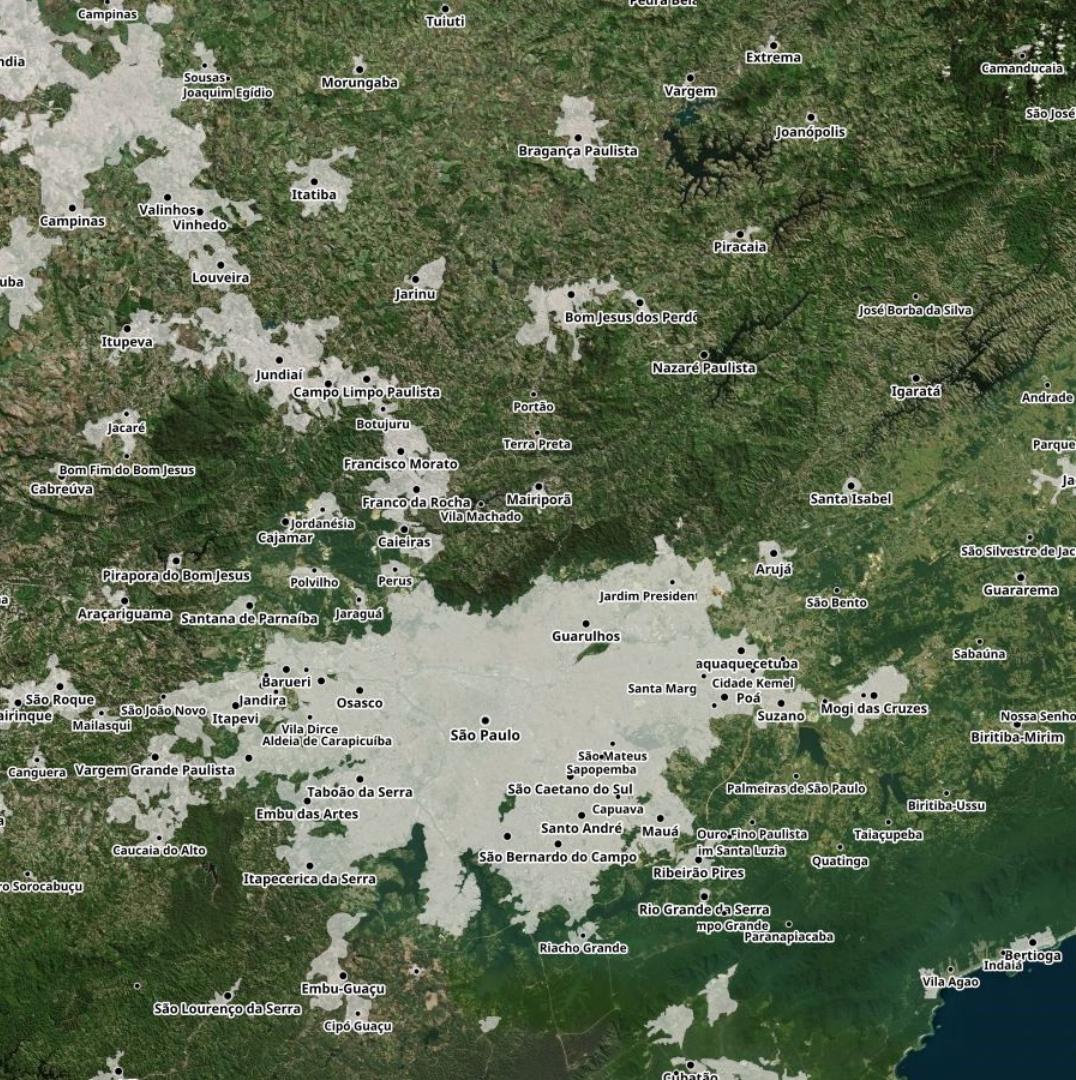
Visualizando



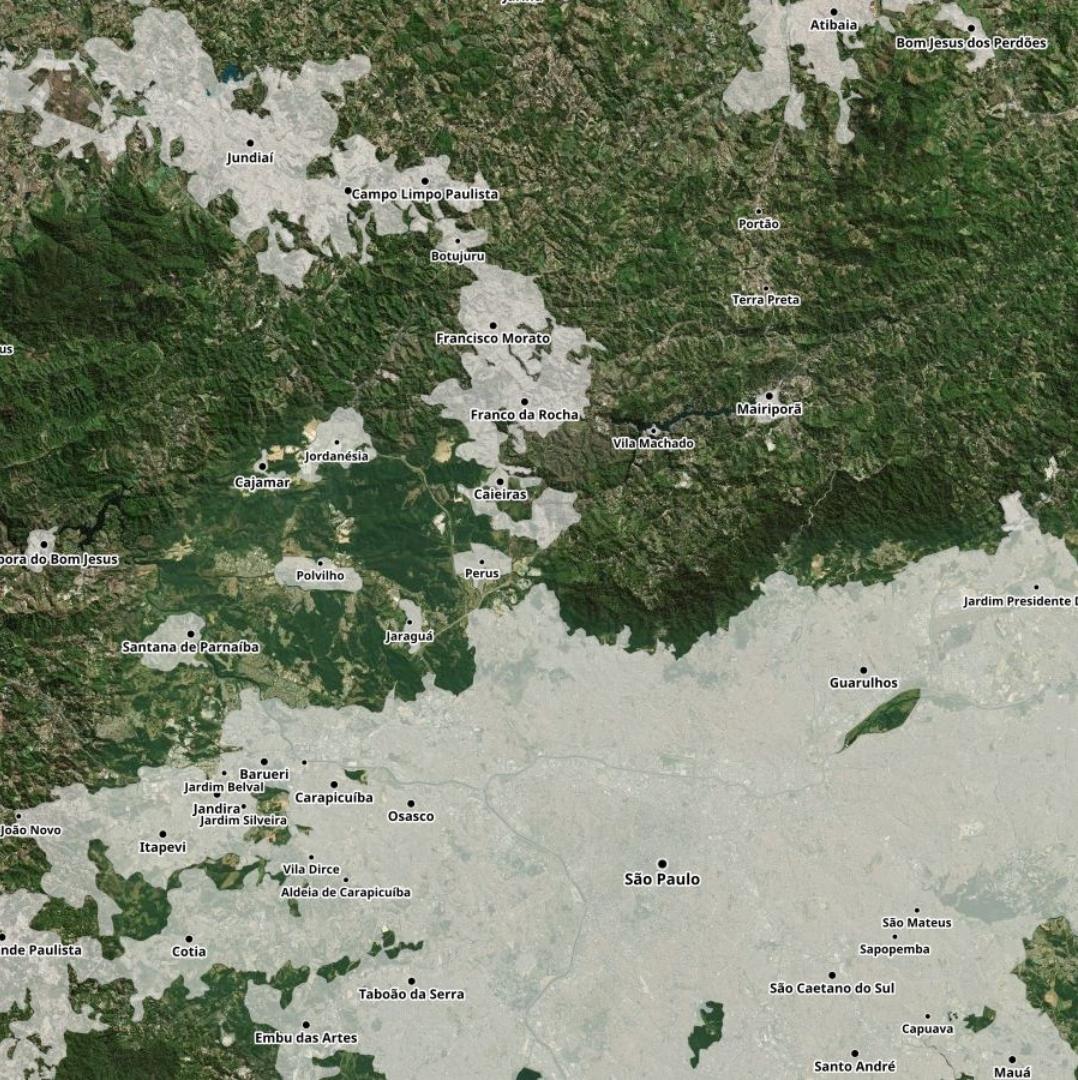
Visualizando



Visualizando



Visualizando

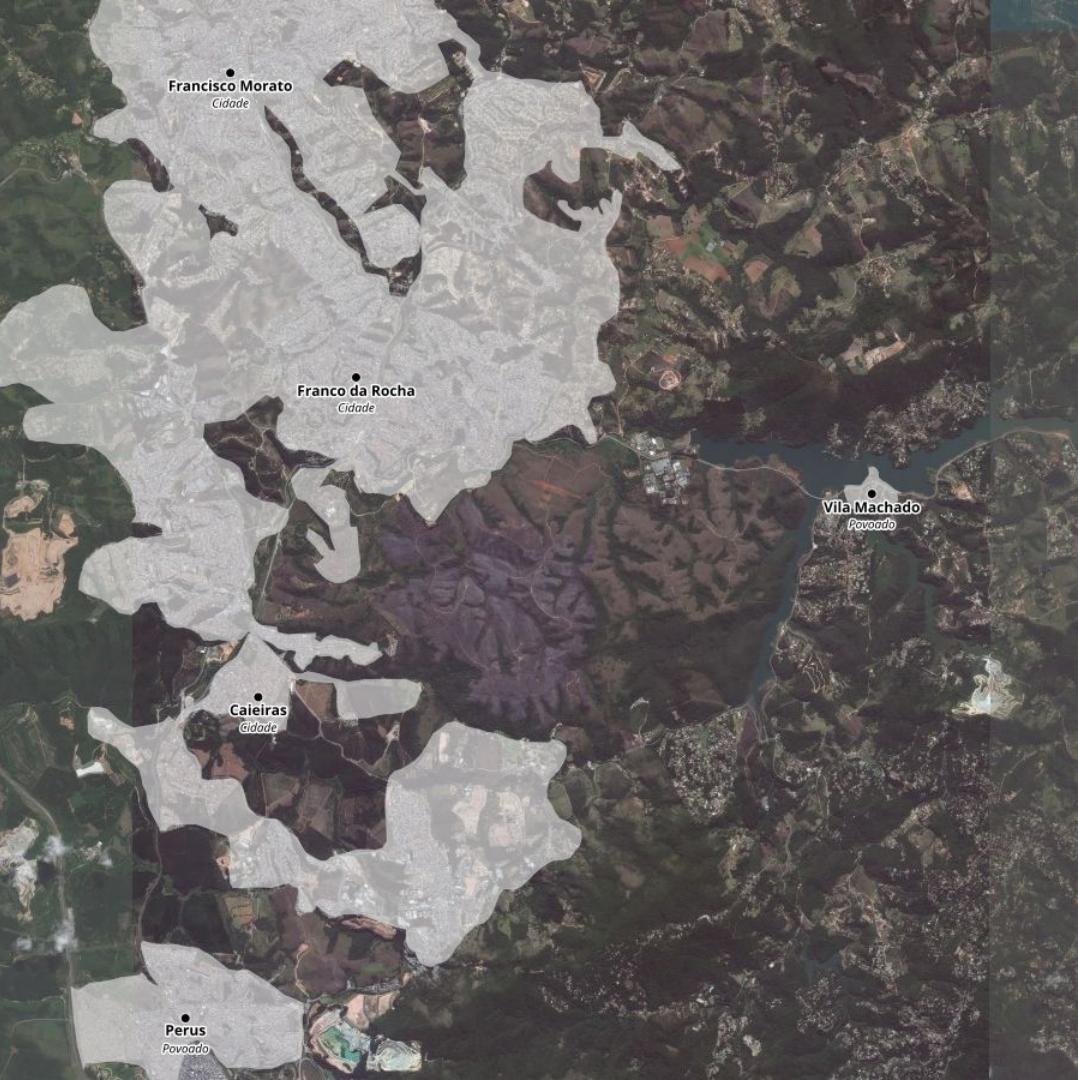


Visualizando



Visualizando

Leve transparência permite
visualizar a mancha e a
camada inferior



Uma maneira de utilizar a camada?

Passear no mapa

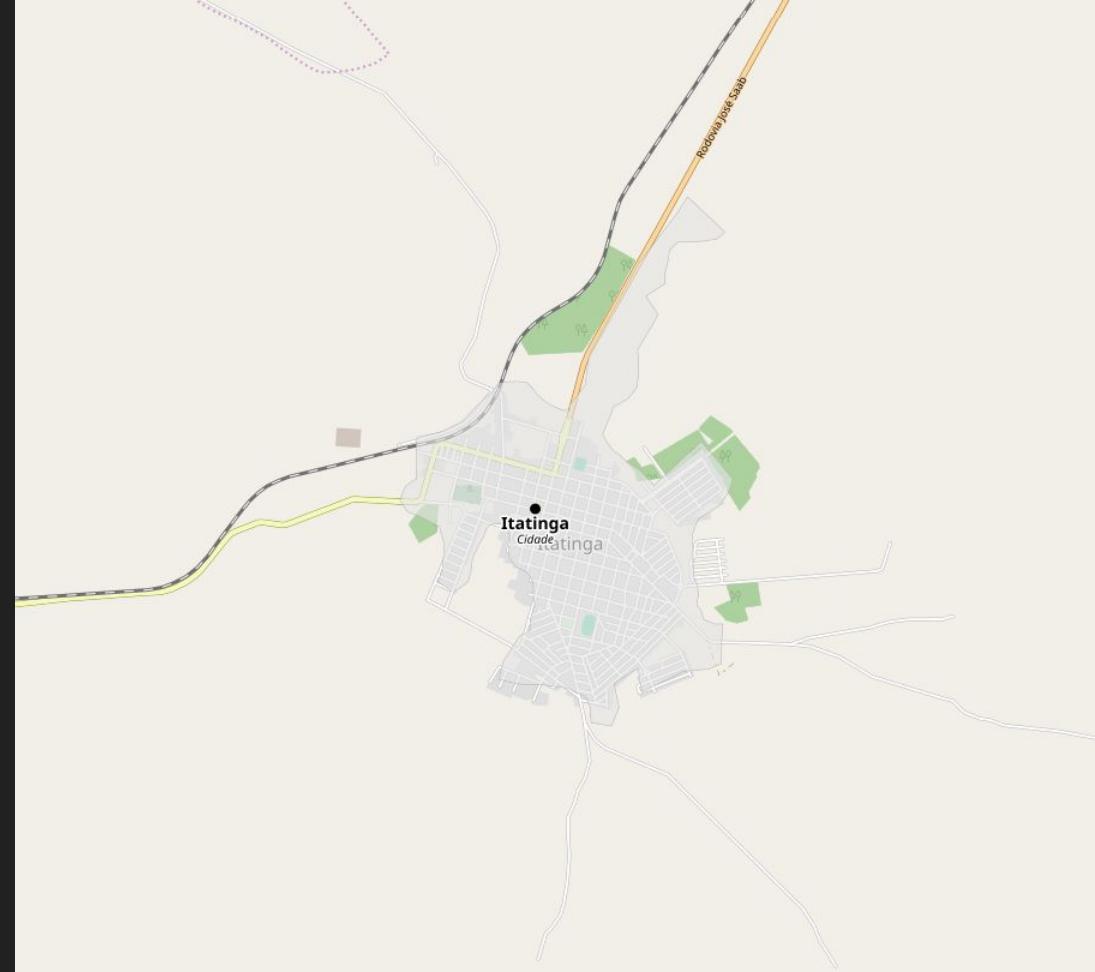
Áreas e pontos são relativamente bem definidos, permitindo verificar, de forma rápida os locais.

Basta habilitar uma camada do OSM com outra das localidades, e começar a "passear" no mapa.

Não requer conhecimentos técnicos ou habilidades específicas.

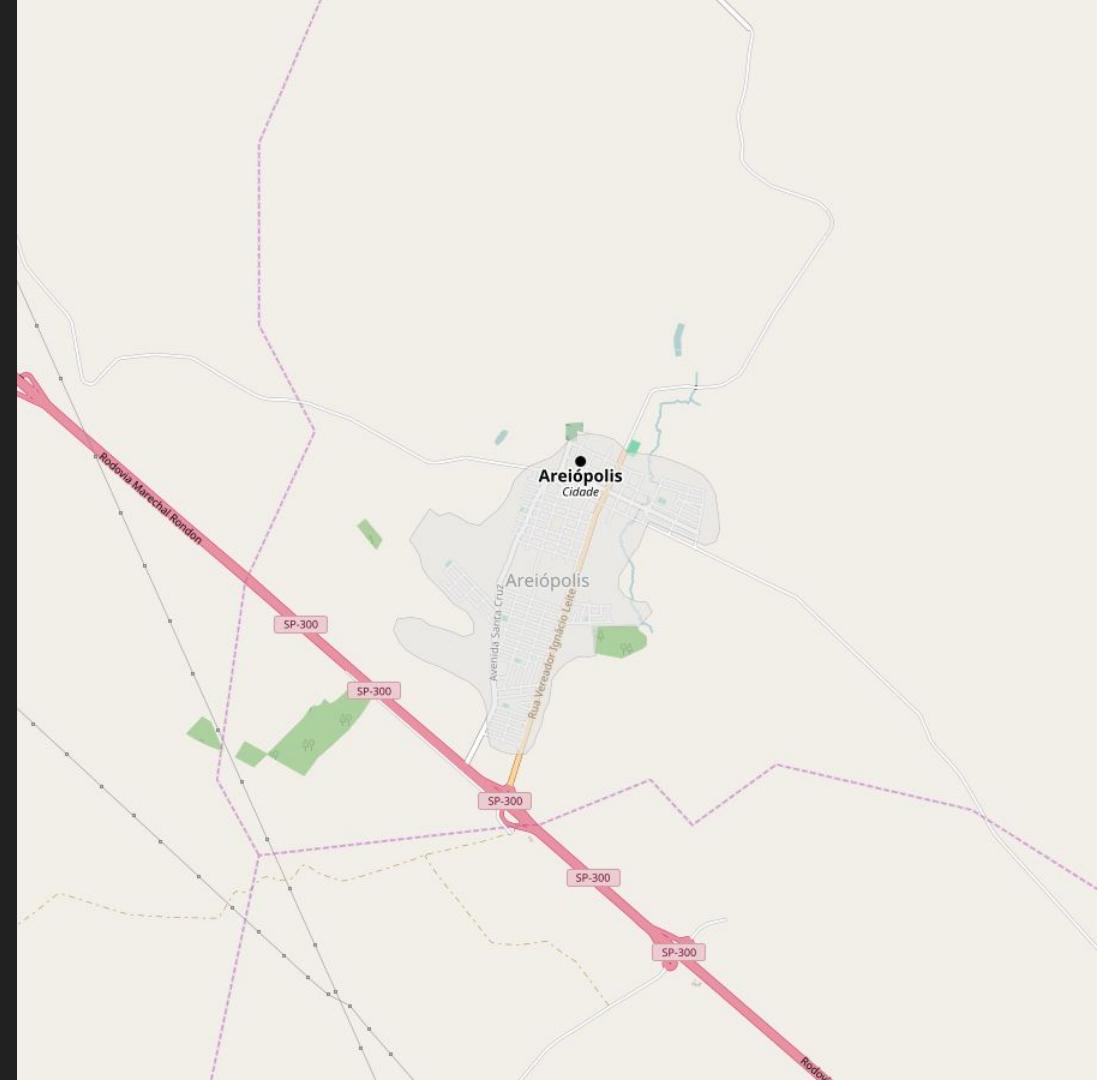
Passeando no mapa

Local possui nome e
aparentemente possui as ruas
mapeadas (correspondem de
forma aproximada com a área
urbana)



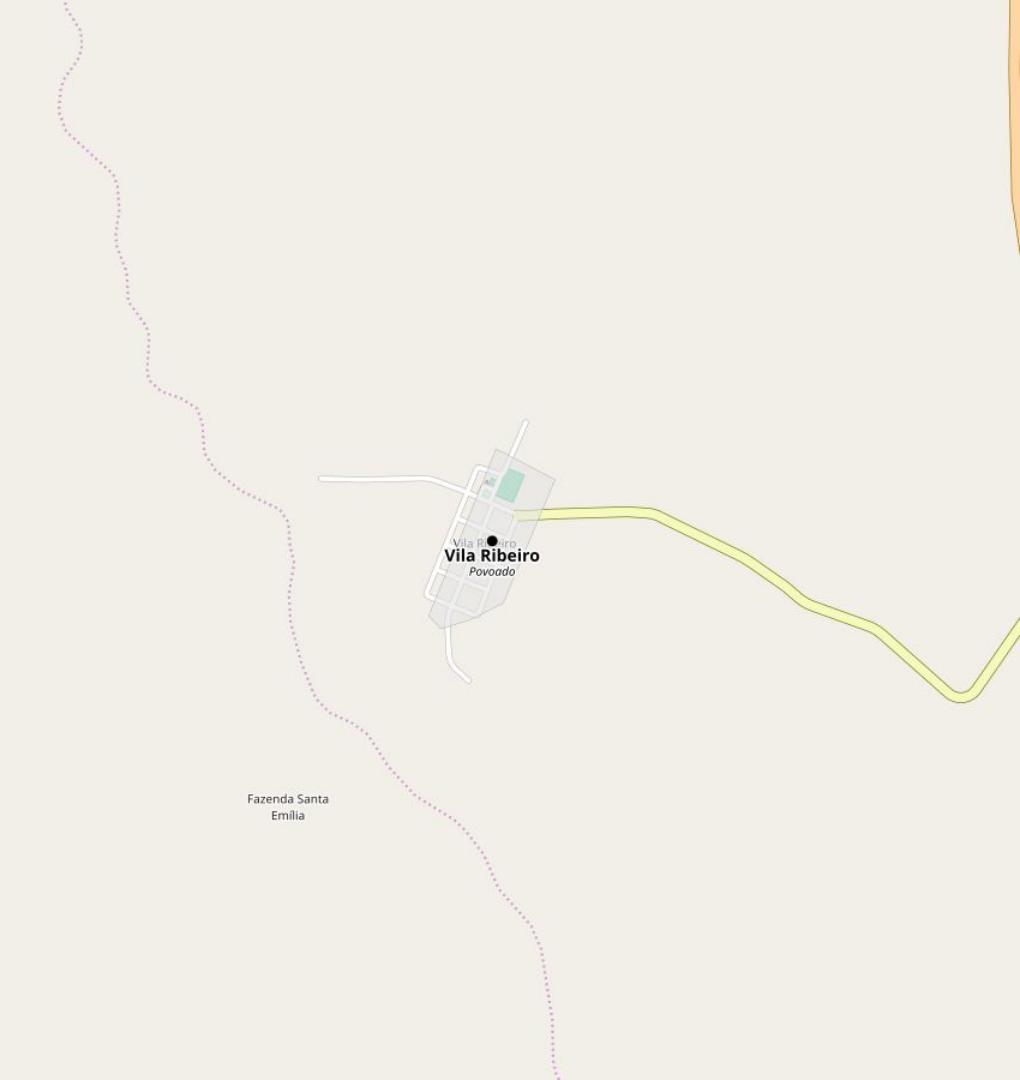
Passeando no mapa

Nó apresenta um leve deslocamento, mas o local também aparenta estar mapeado



Passeando no mapa

Nome e área mapeada se
assemelham aos dados do
IBGE

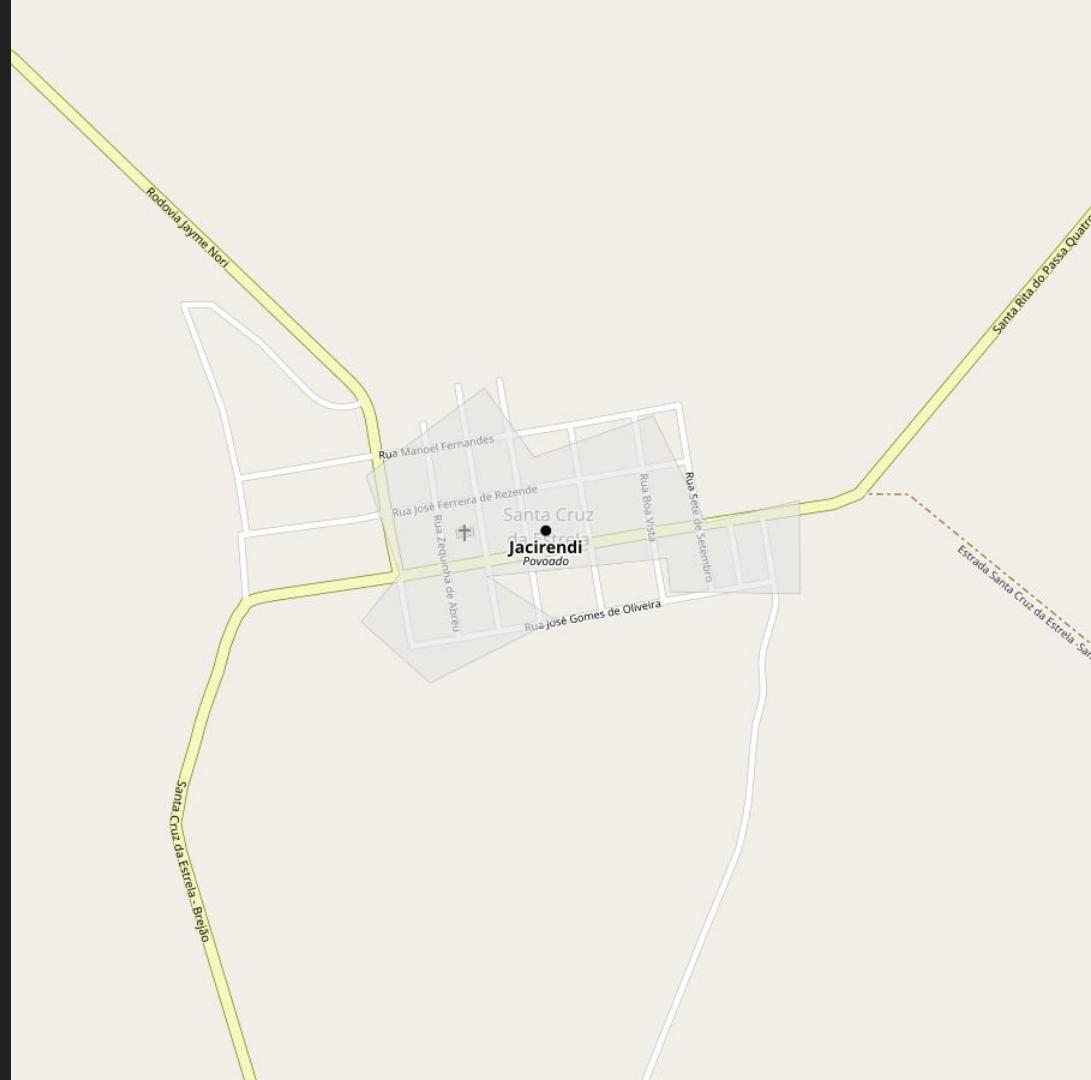


Passeando no mapa

Nome do local no OSM
diverge com o do IBGE: *Santa
Cruz da Estrela X Jacirendi*

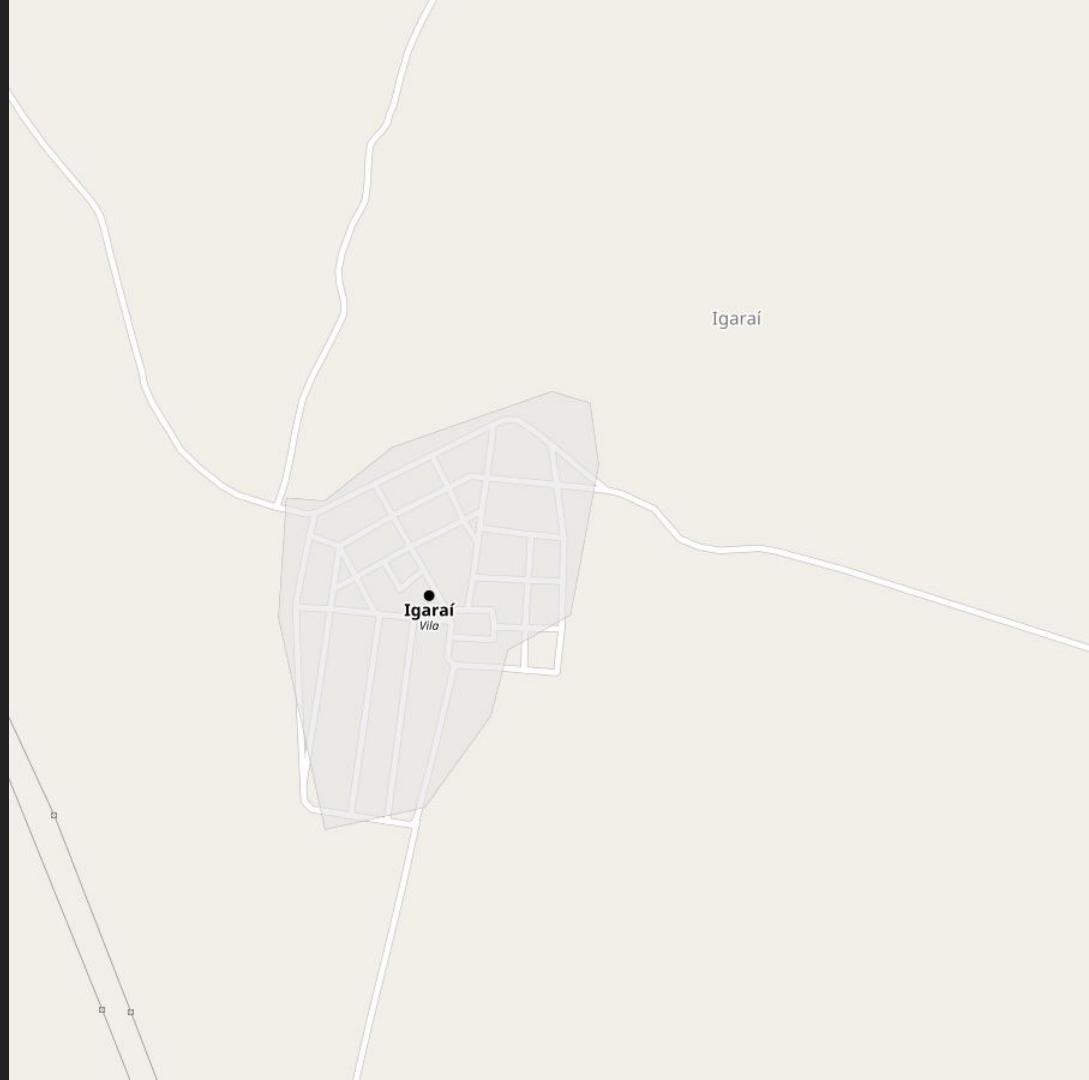
Verificar e corrigir.

?



Passeando no mapa

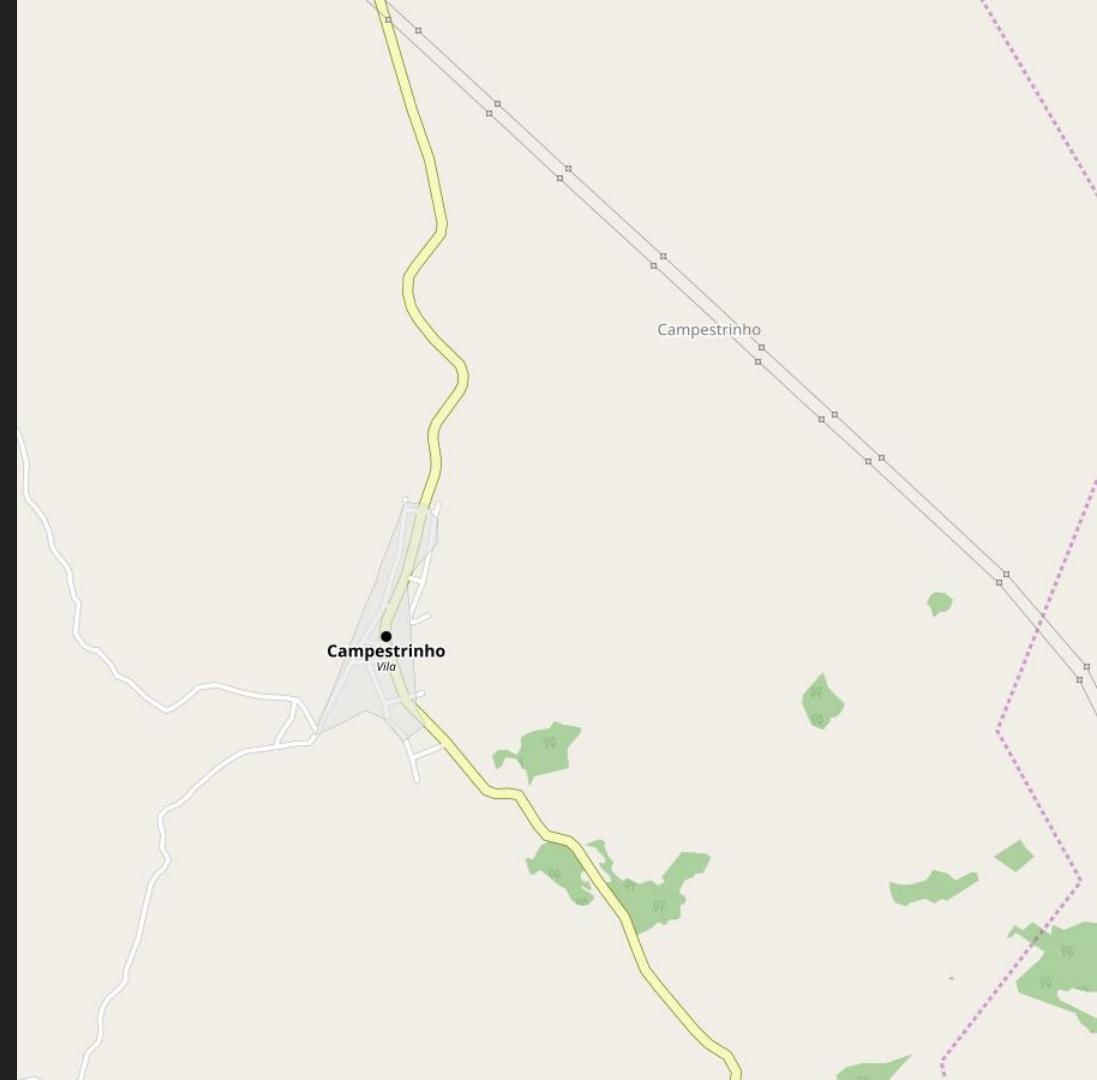
Local no OSM está deslocado (~770m). Precisa ajustar a posição do nó.



Passeando no mapa

Local no OSM está deslocado (~1.7km). Precisa ajustar a posição do nó.

Necessário tomar cuidado para não inserir um novo local.



Passeando no mapa

Possui área urbana demarcada no IBGE, mas não tem nada no OSM (não há ruas ou qualquer outro objeto dentro da mancha).



Passeando no mapa

Conferindo com o Bing é
possível ver que de fato existe
uma área incompleta no OSM.

Mapeie!



Passeando no mapa

Possui um ponto representando um povoado, mas não há nada no OSM (nem mesmo um nó representando o local).

●
Capoeirinha
Povoado

MG-170

Passeando no mapa

Verifique no Bing e adicione o local (o nó com *place*), junto com as ruas e estrada de ligação.



Outras formas de uso das localidades

É possível automatizar alguns testes para detectar ausências e controlar a qualidade dos locais no OSM. Por exemplo:

- Verificar se todos os locais estão representados no OSM
 - Buscar um nó no OSM, dentro de um raio determinado, que possua o mesmo nome que o ponto do IBGE; se não existir, possivelmente o local não está mapeado (ou foi removido)
- Detectar locais deslocados
 - Calcular a distância entre o nó no OSM e o ponto do IBGE; se estiver acima de determinado valor, gerar um aviso
- Detectar áreas ainda sem mapeamento
 - Utilizando as manchas urbanas, verificar se existem ruas no OSM que estejam dentro dessas áreas; caso não exista, muito provavelmente existe a necessidade de se mapear a localidade

Faces de logradouros

De onde obter o nome das ruas?

Alguns exemplos

- Leis
- Mapas/plantas da prefeitura
- Survey (anotar nome das placas)
- Perguntar para um desconhecido "Ow, qual o nome dessa rua?" (não é legal)
- Mapas do censo agropecuário 2006 / contagem da população 2007, do IBGE
- "É só habilitar a camada do setores urbanos do IBGE"

Censo agropecuário / contagem da população

Acessíveis através do diretório *mapas_de_setores_censitarios_urbanos* de cada estado, em

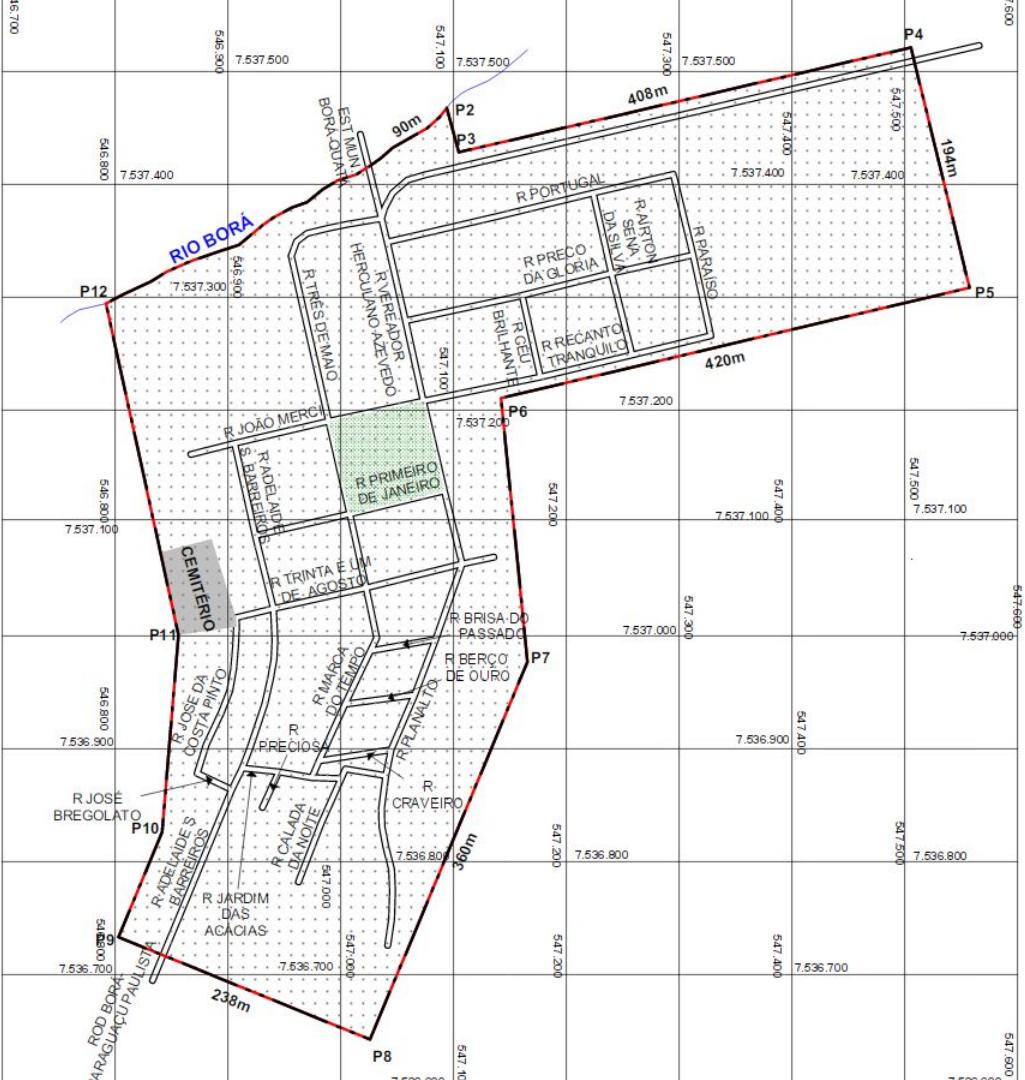
ftp://geoftp.ibge.gov.br/cartas_e_mapas/mapas_para_fins_de_levantamentos_estatisticos/contagem_da_populacao_e_censo_agropecuario_2007/mapas_urbanos_estatisticos_e_mapas_de_setores_censitarios_urbanos/ (ou goo.gl/3u2Wf6)

Apresentados como uma série de PDFs (quantidade proporcional ao tamanho das cidades, representando partes dos locais)

É possível ver nome de ruas (em alguns casos encontrados apenas nesses PDFs), nomes de locais, rios, posição de escolas, hospitais, órgãos públicos, etc

Exemplo

Mapa de Borá - SP, em
apenas um PDF



"É só habilitar a camada do IBGE"

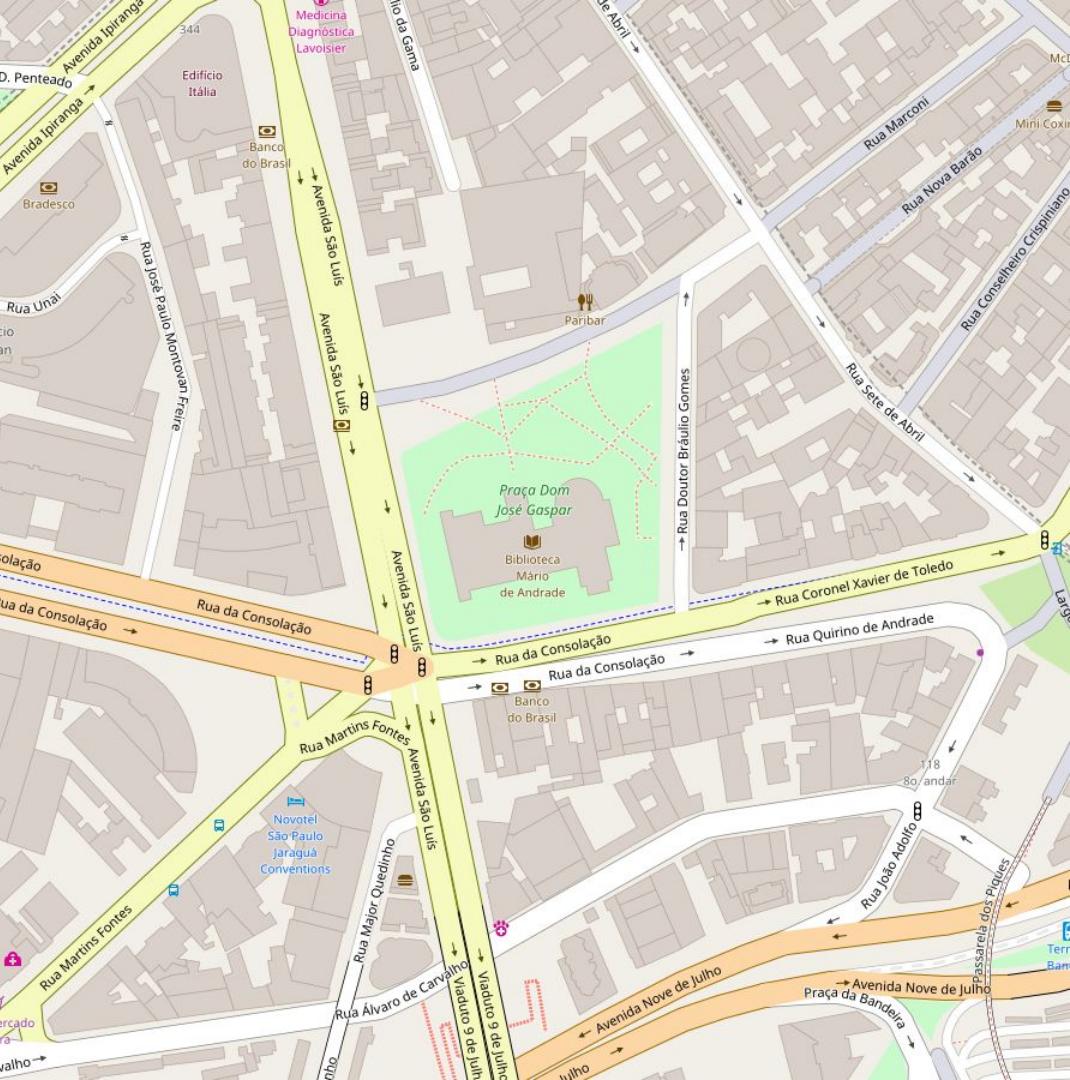
Excelente trabalho do Thiago Santos (tmpsantos)

Facilmente acessível no JOSM ou iD: basta clicar na camada "IBGE Mapa de Setores Urbanos"

É o método mais rápido e conhecido para se obter, de forma lícita, os nomes de ruas

Nomeando vias

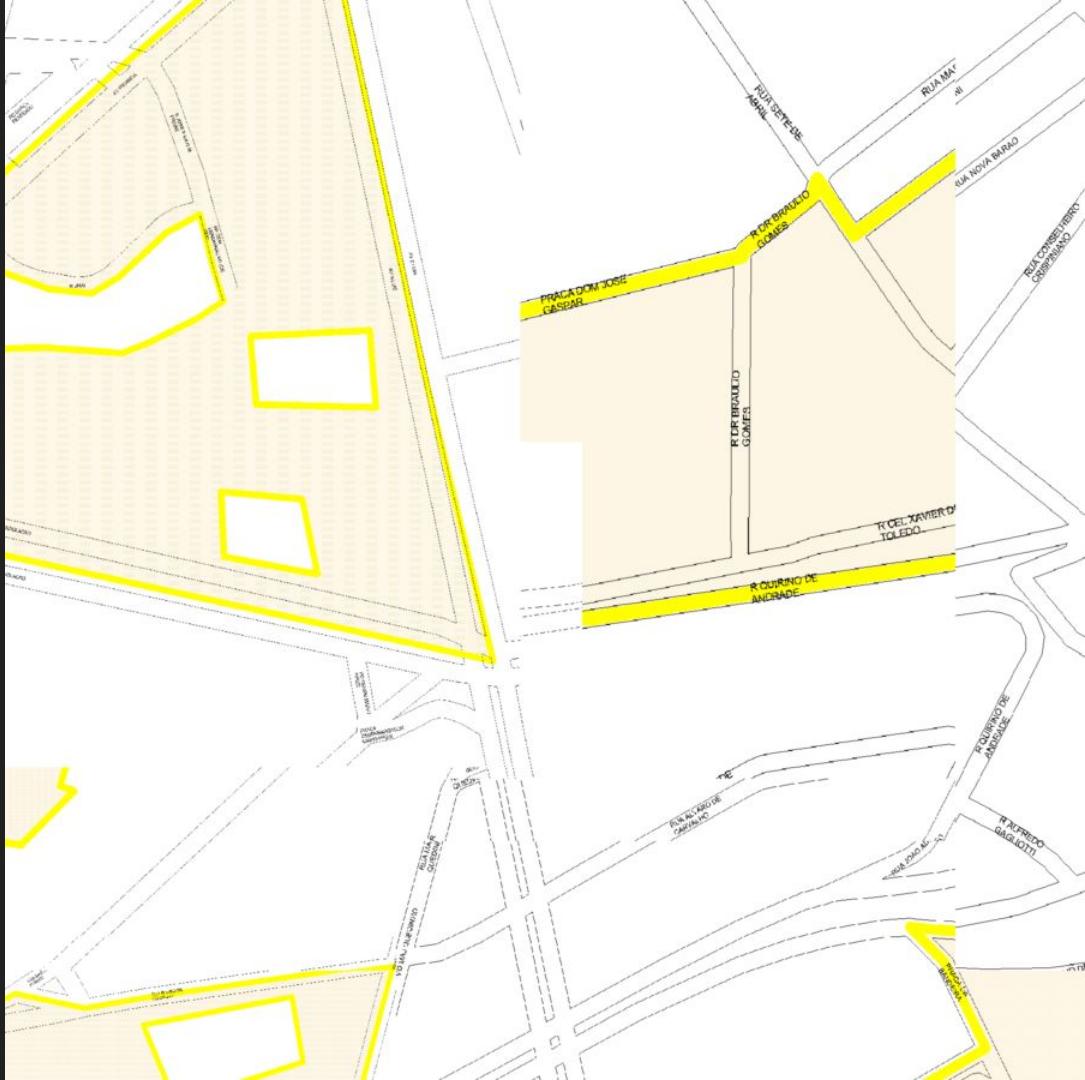
Se fôssemos nomear as ruas dessa região, o que usariamos?



Nomeando vias

Camada "IBGE Mapa de Setores Urbanos"

Há locais que apresentam falhas no encaixe dos PDFs, ausência de informações, qualidade baixa dos dados originais, etc



Há alguns
problemas em
algumas áreas

E se pudéssemos ter outra camada
para consultar os nomes?

Faces de logradouros

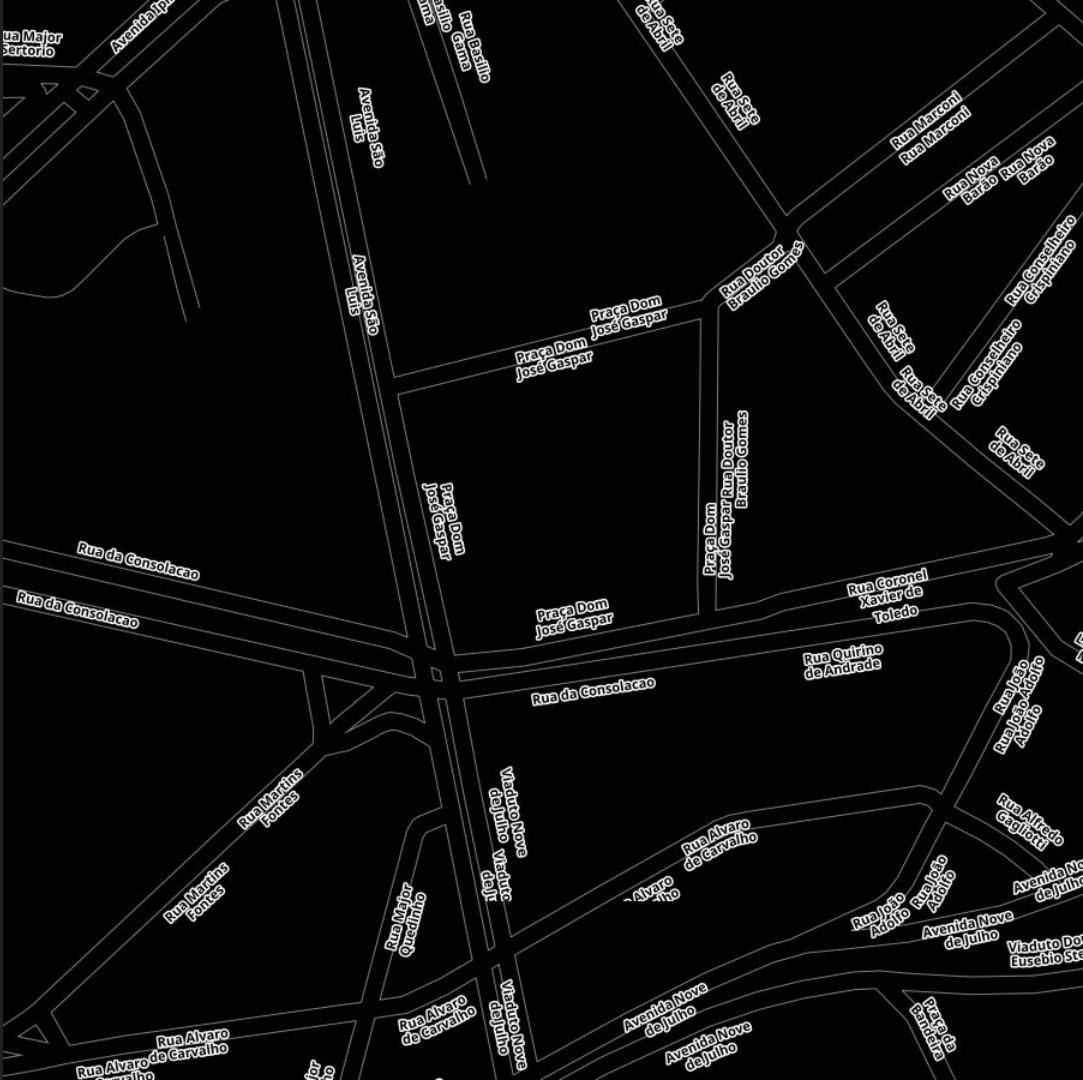
- Faces de Logradouros do Censo Demográfico 2010, disponibilizados em julho de 2016
- Acessíveis em ftp://geoftp.ibge.gov.br/recortes_para_fins_estatisticos/malha_de_setores_censitarios/censo_2010/base_de_faces_de_logradouros/ (ou goo.gl/HpzcrK)
- 10906 arquivos (1.9G no total, compactados)
- Cada arquivo possui informação das faces, setores e subdistritos de uma região
- Há 10810 shapefiles de interesse, totalizando 3.3G (descompactados)

Processo

- Foram extraídos apenas os shapefiles das faces
- Criada uma base contendo apenas um campo "nome", obtido pela concatenação dos campos *NM_TIPO_LO*, *NM_TITULO_* e *NM_NOME_LO*.
Exemplo:
 - *NM_TIPO_LO*: AVENIDA
 - *NM_TITULO_*: DUQUE
 - *NM_NOME_LO*: DE CAXIAS
- Removidos espaços duplicados e normalizado o nome
 - _ _ AVENIDA _ DUQUE _ _ _ DE _ CAXIAS → Avenida _ Duque _ de _ Caxias
- Alguns acentos foram corrigidos
 - Joao → João, Independencia → Independência, Barao → Barão, etc
- Logradouros como "Rua Sem Nome" e "Rua Projetada" ficaram sem nome

Resultado

Camada com cada face (lado da rua ou do quarteirão) nomeada



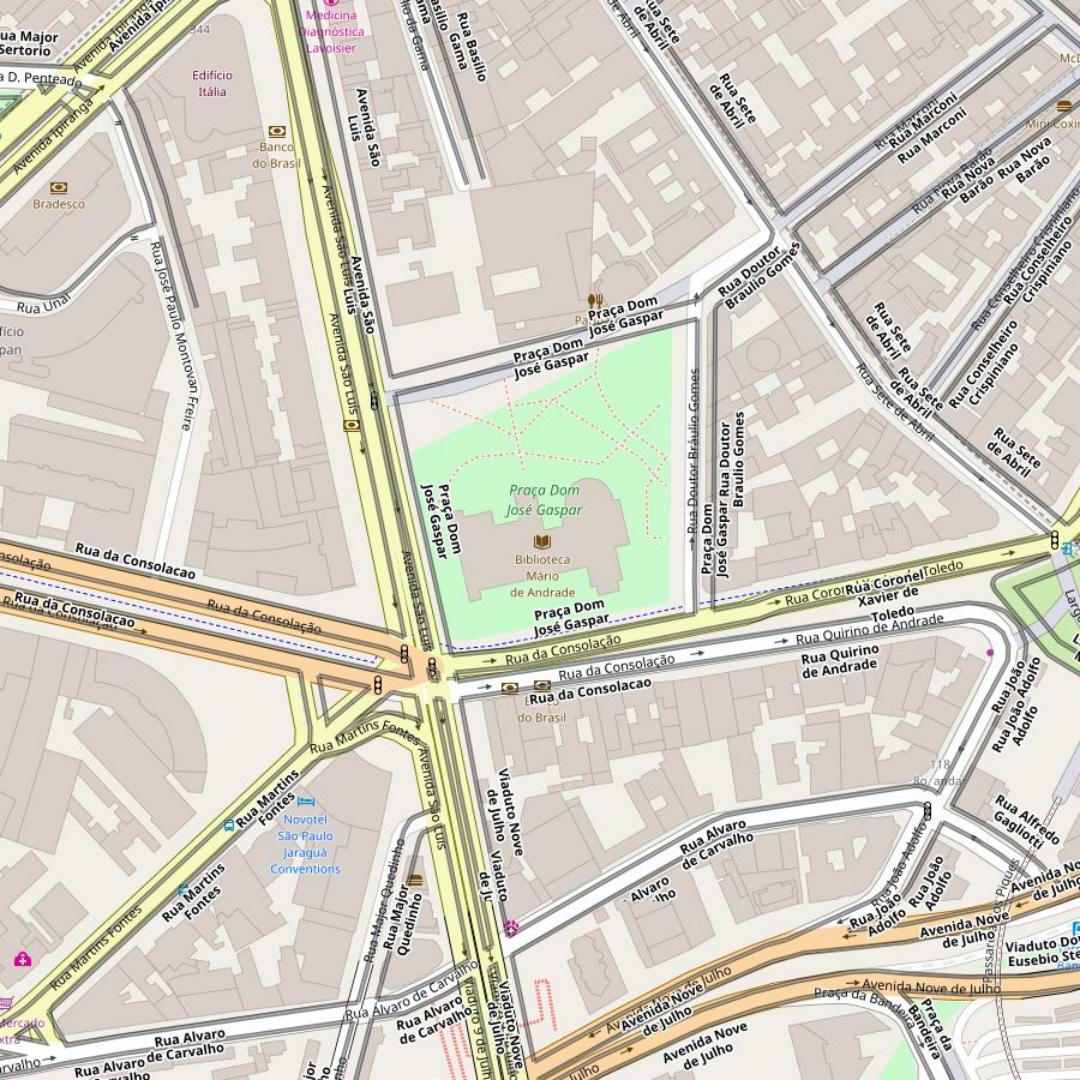
Resultado

Estilo prototipado para tentar ficar visível com Bing aerial imagery



Resultado

Com OpenStreetMap (Mapnik)



Resultado

Com Mapbox Satellite



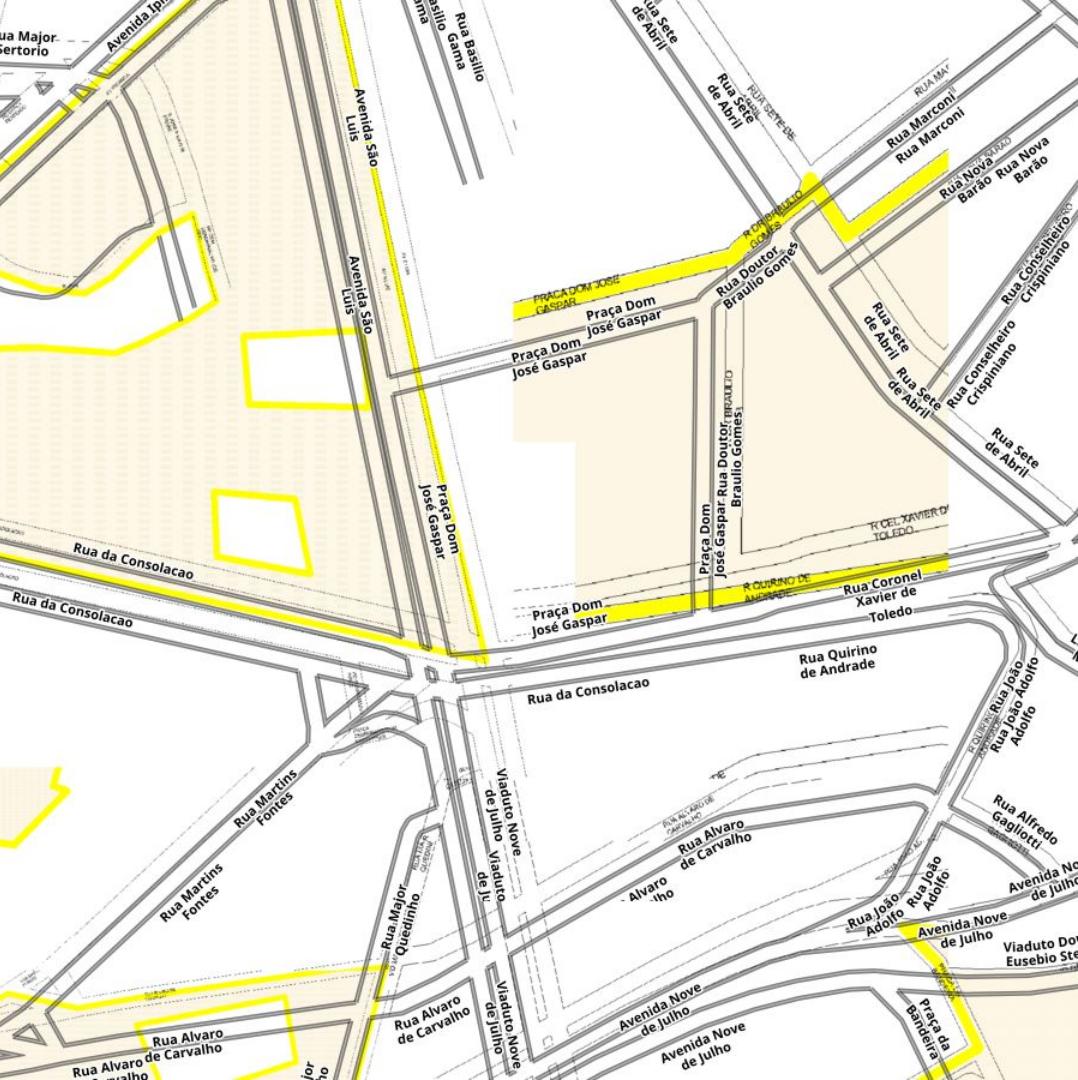
Resultado

Com Strava cycling and running heatmap



Resultado

Com IBGE Mapa de Setores
Urbanos



Resultado

Com dados do OSM no JOSM



Resultado

Em áreas densas com o uso de outros estilos, filtros ou visão *wireframe*



Exemplo

Uma área urbana qualquer



Exemplo

A camada atual do IBGE
possui algumas
inconsistências ou falta de
nomes no local



Exemplo

A camada de faces fornece as informações que buscamos



Problemas

Também existem
adversidades com as faces de
logradouros...

Há áreas que não possuem
dados



Problemas

Há cidades inteiras que não possuem nome, apesar de ter a geometria das faces



Problemas

Algumas ruas que possuem nome na camada "IBGE Mapa de Setores Urbanos"...



Problemas

... não estão nomeadas nas faces de logradouros.

As informações são complementares e não substitutas.



Folhas topográficas

Folhas topográficas

São as "cartas"

É possível comprá-las ou baixá-las gratuitamente em <http://loja.ibge.gov.br/>

Algumas foram vetorizadas e estão acessíveis em
ftp://geoftp.ibge.gov.br/cartas_e_mapas/folhas_topograficas/vetoriais/ (ou goo.gl/sVenEC)

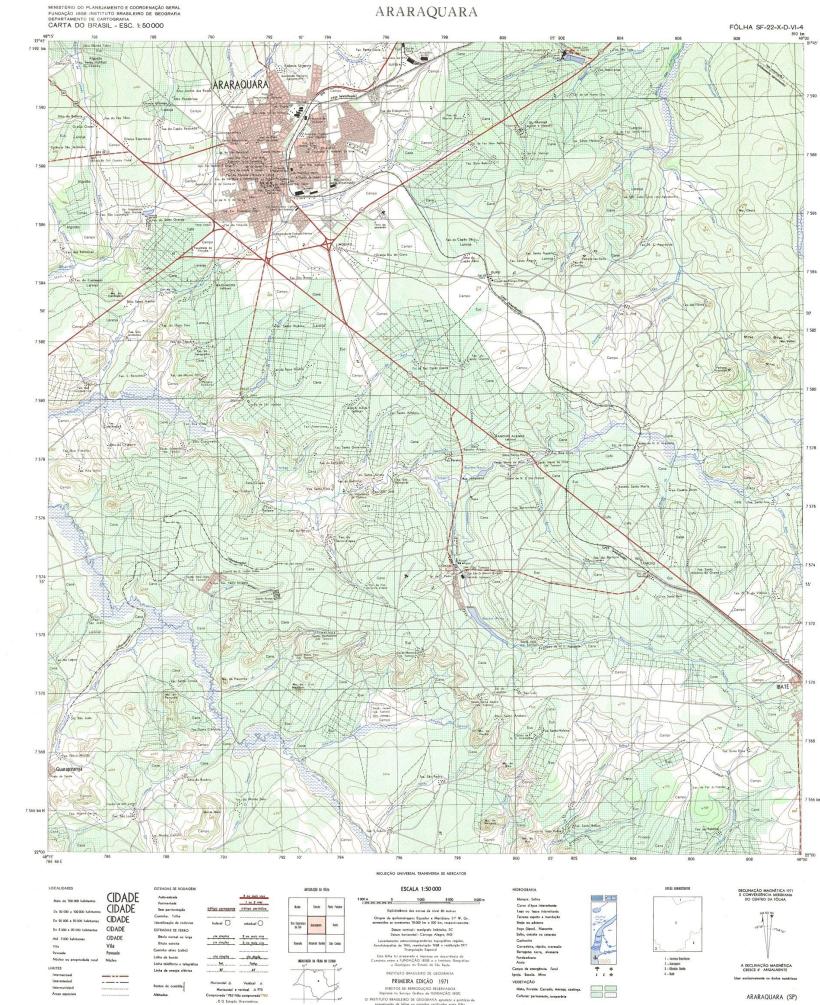
Há dados em várias escalas (desde 1:1000000 até 1:25000)

Folhas topográficas

Por exemplo, a folha de Araraquara, edição de 1971

Contém vários tipos de informações

Mesmo sendo "antiga", ainda há dados úteis



Folhas topográficas

As folhas vetorizadas são compostas de:

- Hidrografia
- Hipsografia
- Limite
- Localidade
- Obra e edificação
- Ponto de referência
- Sistema de transporte
- Sistema viário
- Vegetação

Folhas topográficas

As localidades vetorizadas em escala de 1:50000 são disponibilizadas como arquivos DGN

⚠️ Os arquivos DGN não possuem especificação da projeção dos dados

Além de ter cuidado para encontrar e utilizar a projeção correta, também é preciso prestar atenção na escala.

Folhas topográficas

Por exemplo, uma folha na projeção SAD69 UTM zona 22S (EPSG:5858)

```
+proj=utm +zone=22 +south +ellps=aust_SA +towgs84=-67.35,3.88,-38.22,0,0,0,0  
+units=m +no_defs
```

No caso desses DGN é necessário utilizar uma projeção personalizada, com unidades em quilômetros:

```
+proj=utm +zone=22 +south +ellps=aust_SA +towgs84=-67.35,3.88,-38.22,0,0,0,0  
+units=km +no_defs
```

Folhas topográficas

Dados estão disponíveis como pontos, linhas e polígonos

Possuem atributos de tipo (campo *Type*), nível (*Level*), cor (*ColorIndex*), peso (*Weight*), estilo (*Style*) e texto (*Text*)

Há tabelas que identificam cada tipo de objeto, utilizando uma combinação única dos valores dos campos

Os nomes dos objetos, geralmente, estão representados nos pontos

Folhas topográficas

Exemplo de hidrografia:

- Baliza: $Level = 55 + ColorIndex = 0$
- Rio Intermítente: $Level = 9 + ColorIndex = 79$
- Rio Permanente: $Level = 7 + ColorIndex = 77$

Dependendo do uso, pode-se simplesmente considerar os três tipos de objetos como "rio" (`waterway=river`).

Processo

Para localizar a zona UTM de cada DGN:

- Cada arquivo DGN é associado ao seu estado, utilizando uma tabela de municípios e estados
- O que sobra é verificado manualmente (podem conter erro de ortografia, por exemplo)
- Com todos os locais identificados, tenta-se localizá-los através do nominatim
- Os que não são localizados são verificados se existem no OSM, e manualmente adicionados quando necessário
- Com todos os nomes sendo encontrados no OSM, obtém-se a latitude e longitude e converte a localização em UTM

Resultados

Hidrografia

Rios, córregos, lagos,
represas



Resultados

Hipsografia

Curvas de nível, picos,
montanhas, pontos cotados



Resultados

Obra e edificação

Fazendas, sítios, indústrias,
aeroportos, prédios públicos



Resultados

Localidade

Cidades, bairros, distritos,
áreas rurais, manchas urbanas



Resultados

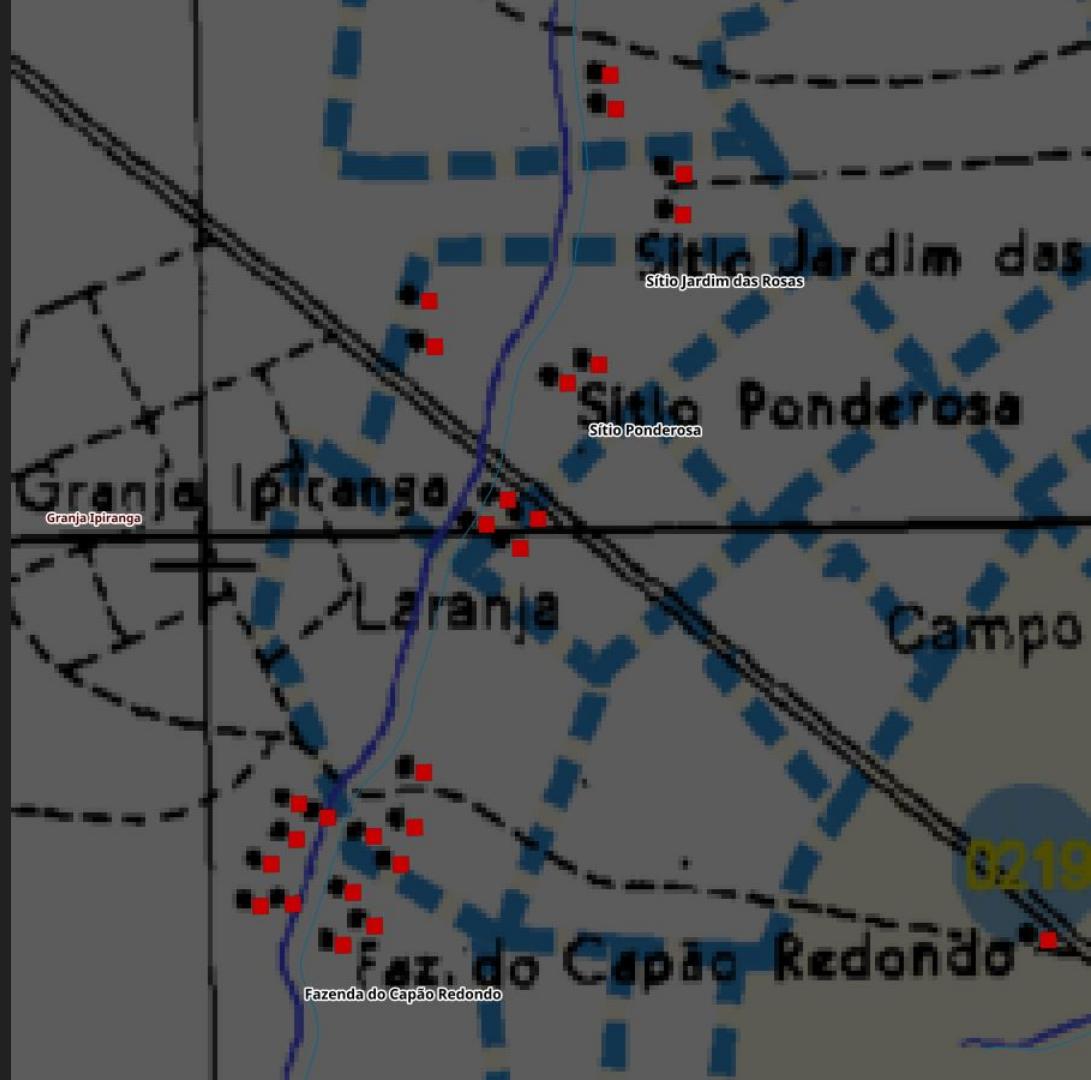
Sistema viário

Ruas, estradas rurais,
rodovias, estradas de ferro



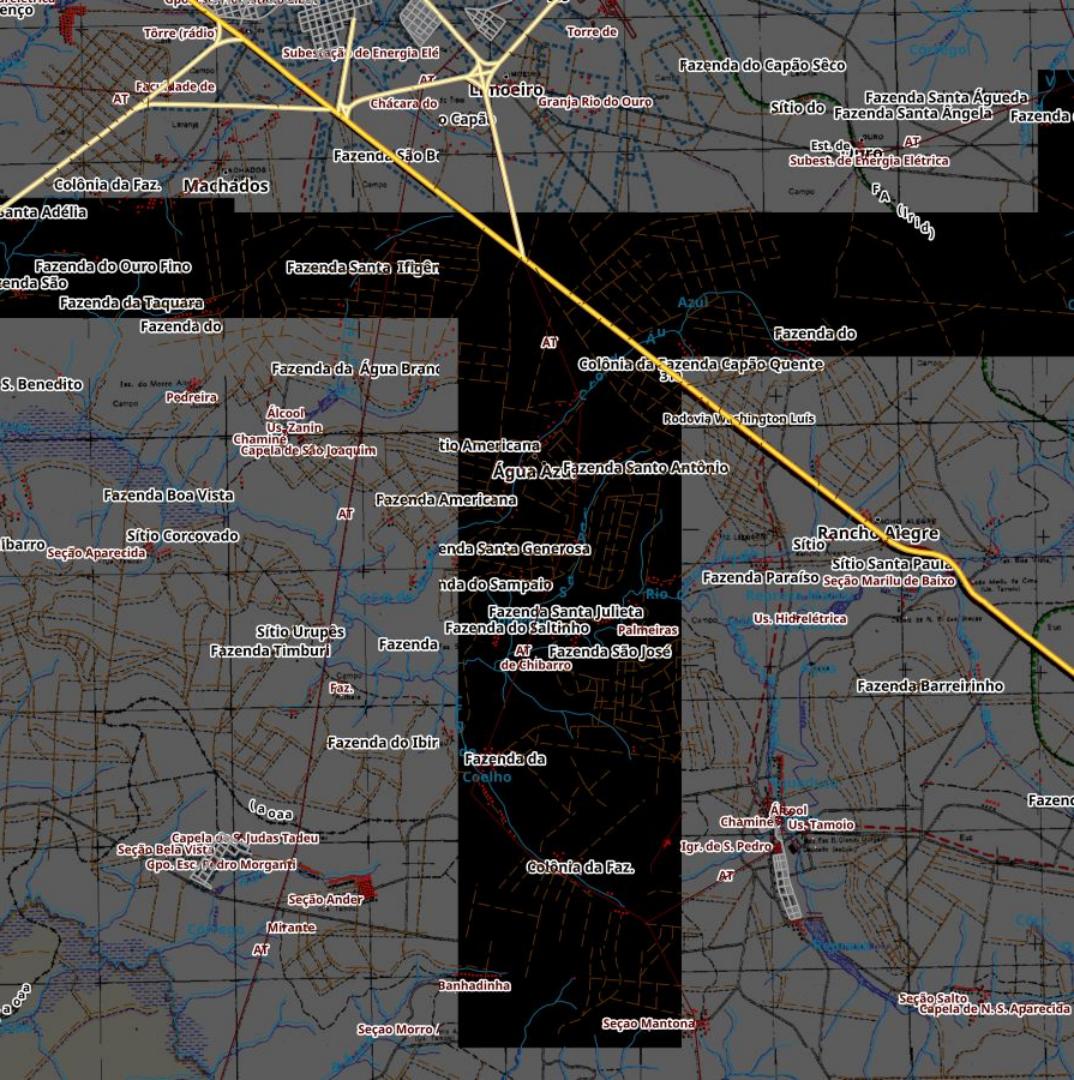
Resultados

Em muitos casos as informações são idênticas à camada "IBGE Mapa de Setores Rurais", já disponível no JOSM e iD



Resultados

Assim com as faces de logradouros, em vários momentos também complementa a camada já existente, quando esta não possui algumas informações



Usos das folhas topográficas

- Por serem vetorizadas, oferecem um leque muito grande de aplicações
- É possível prover várias camadas diferentes, cada uma com um tipo ou grupo de informação específico
- São mais complexas de serem trabalhadas, já que oferecem informações de vários temas
 - Mais trabalhoso para analisar os dados (precisa procurar parâmetros que identificam aquele tipo de informação)
 - Mais difícil para encontrar uma renderização que fique boa com todas as informações
- Permite os mesmos usos dos dados anteriores (verificar visualmente por problemas ou dados ausentes, por exemplo, ou através de processos automatizados)

Conclusão

Conclusão

- Há **muita** informação disponível no IBGE
- O OpenStreetMap pode se beneficiar bastante com esses elementos
- Apesar de haver alguns casos com dados de melhor qualidade, eles não substituem os recursos já existentes
- Com o devido cuidado e conhecimento de projeção, escala, etc, é possível utilizar os dados de maneiras bem flexíveis (para cálculos, comparações, controle de qualidade, complementação dos dados existentes no OSM, etc)
- O objetivo é tornar o Brasil o mais completo possível no OSM, mas com prudência: por exemplo, não é correto importar os dados sem a devida compreensão, detalhamento e discussão com a comunidade; precisamos também de qualidade, além de quantidade

Obrigado!