

Trabalho de TCC-2015 IME-USP

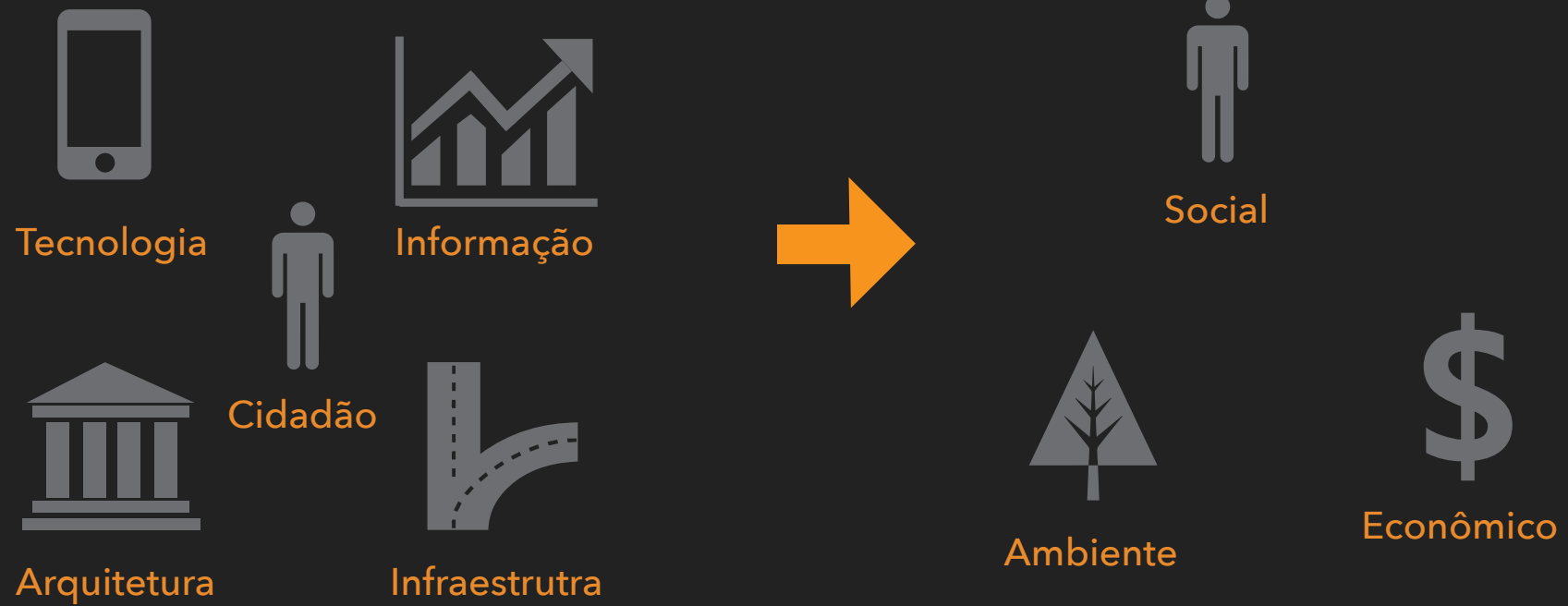
Orientando: André Kenji Yai

Orientador: Roberto Hirata Jr



ANÁLISE E VISUALIZAÇÃO DE DADOS DO TRANSPORTE DE SÃO PAULO

CIDADES INTELIGENTES



A blue bus is shown from a front-three-quarter view at night. The bus's headlights are on, and its interior lights are visible through the windshield. The background is dark and blurry, suggesting a city street at night.

Recentemente muitas
cidades estão
disponibilizando os
seus dados. Em 2013 a
SPTrans começou a
disponibilizar os
dados sobre as linhas
de ônibus.

OBJETIVO

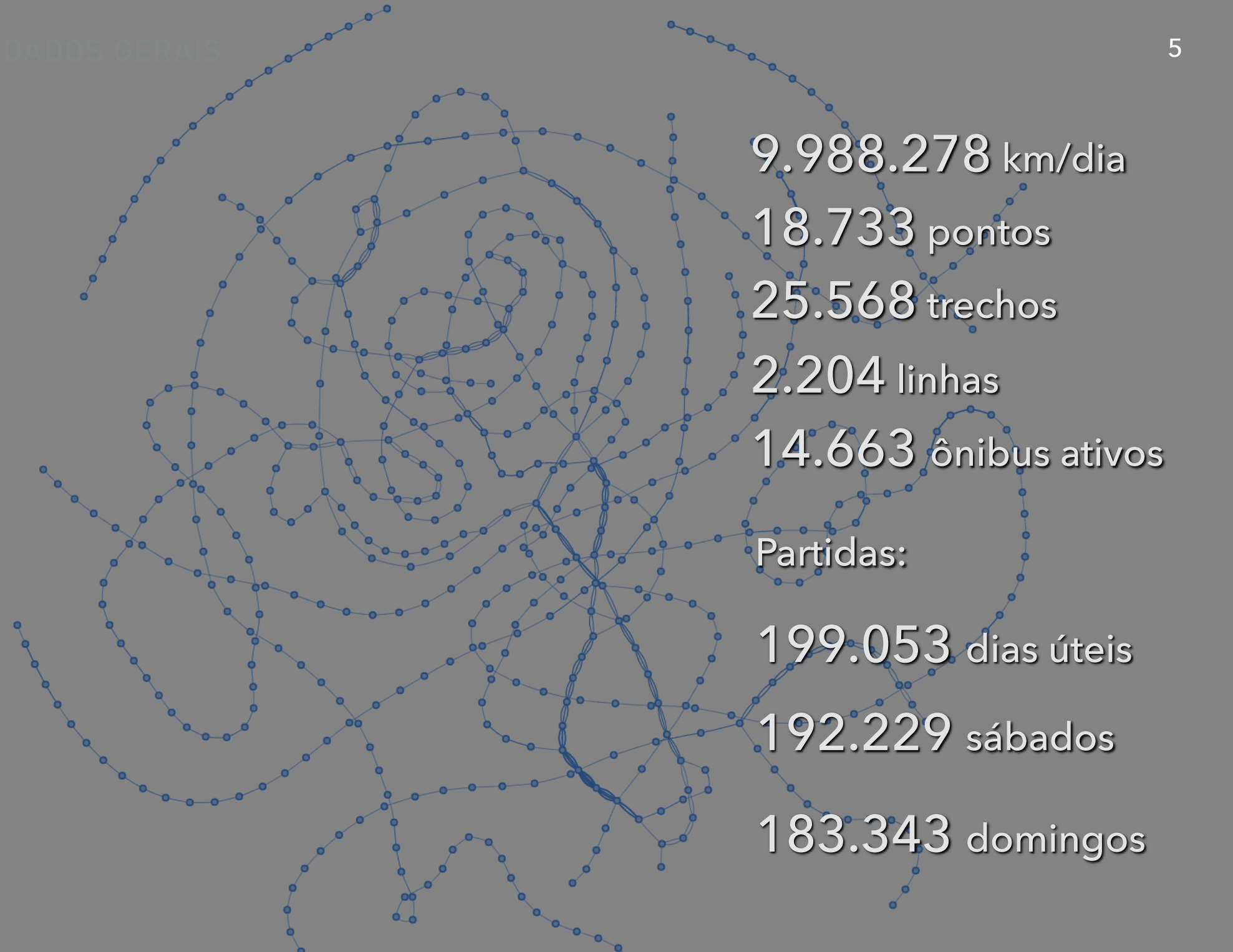
Este **projeto** tem como finalidade **explorar e analisar os dados** para **descobrir e visualizar** informações do sistema de **transporte por ônibus** da **cidade de São Paulo**.

DADOS

Neste trabalho utilizaremos 2 tipos de dados diferentes:

General Transit Feed Specification (GTFS): especificação que define um formato para troca de **informações de tabela** de horários do **transporte público** com as posições geográficas associadas.

Allocation Vehicle Location (AVL): equipamento embarcado no ônibus que **transmite automaticamente os dados de posição GPS do veículo a uma central** através de uma **conexão 3G**. Frequência de transmissão 45 segundos.



9.988.278 km/dia

18.733 pontos

25.568 trechos

2.204 linhas

14.663 ônibus ativos

Partidas:

199.053 dias úteis

192.229 sábados

183.343 domingos

FERRAMENTAS



METODOLOGIA

- ▶ explorar e analisar dados
- ▶ identificar anomalias
- ▶ criar e testar métricas para quantificar a qualidade dos dados e dos serviços.

EXPLORAÇÃO DE DADOS

Teve como o objetivo acostumar com a estruturação dos dados. A maioria destes dados foram feitas com os dados do GTFS. Na visualização destes gráficos temos as suas medias, medianas, desvios padrões.

Foram no total 7 métricas entre elas:

- Histograma mostra em um ponto de ônibus quantas linhas passam por ela. (fig1)
- histograma das quantidades de paradas por linha (fig2)
- histograma dos quilômetros das linhas
- histograma das distancias de cada entre ponto de onibus

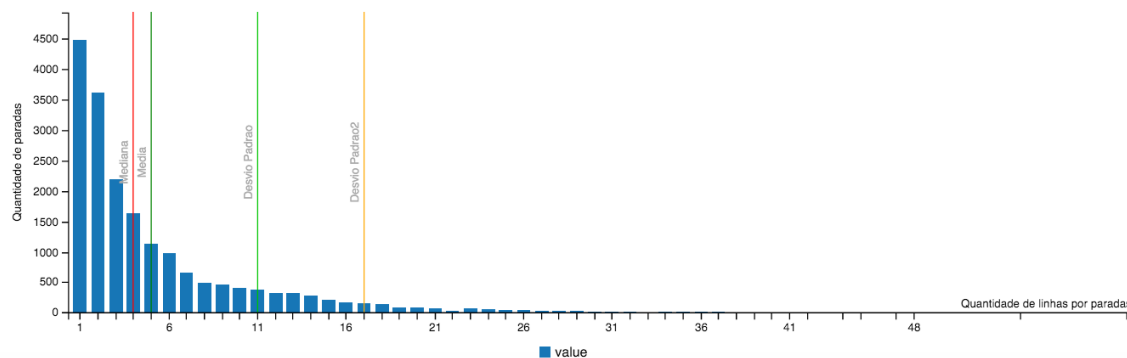


Fig1: Histograma mostra em um ponto de ônibus quantas linhas passam por ela.

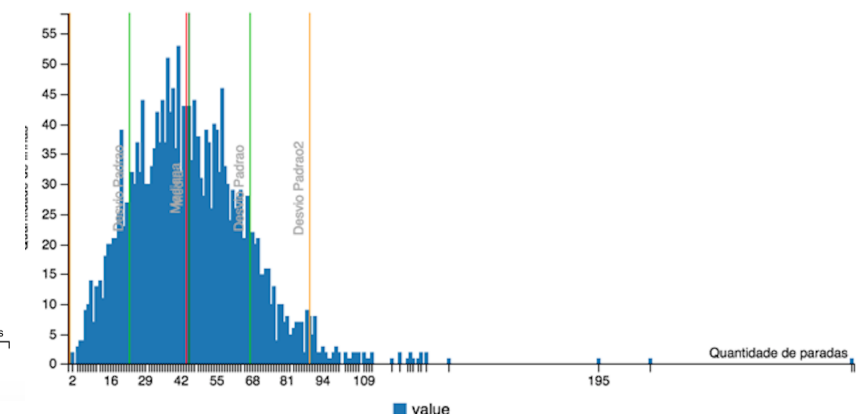


Fig1: histograma das quantidades de paradas por linha

RESULTADO

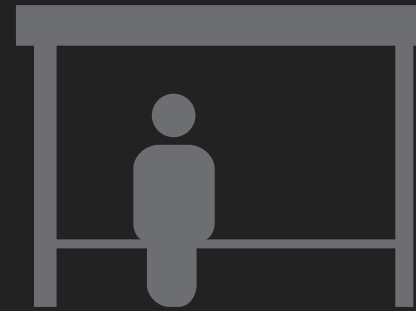


Média de linhas:

44 pontos de ônibus

15 km por linha

26.7% são linhas
circulares



Média em relação ao ponto de ônibus:

5 linhas por ponto de ônibus

350 m metros de distância entre os pontos
de ônibus

ANÁLISE DOS DADOS DO GTFS

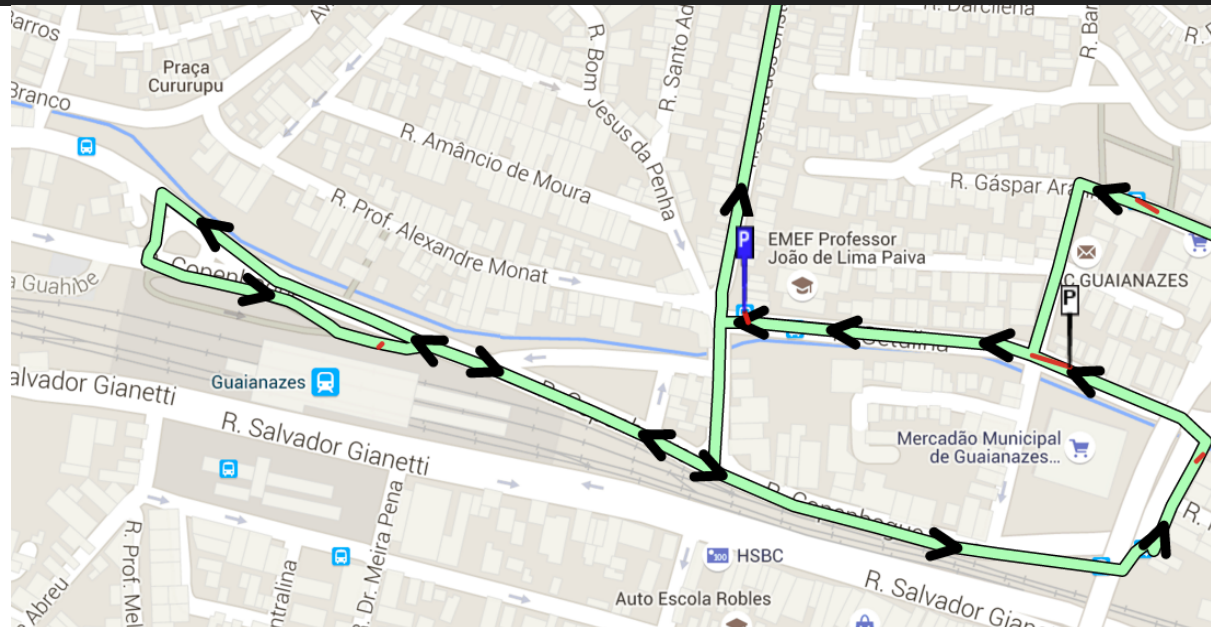
Nesta etapa foram feitas análises com o intuito de sabermos mais sobre o planejamento das linhas feitas pela SPTrans e detectar as suas anomalias.

No total foram **19 métricas**. Entre elas temos as métricas:

- ▶ Quantidade de pontos de ônibus duplicados
- ▶ Quantidade de pontos de ida e volta
- ▶ Quantidade de shapes
- ▶ Porcentagem de trechos compartilhados.

PONTOS DUPLICADOS

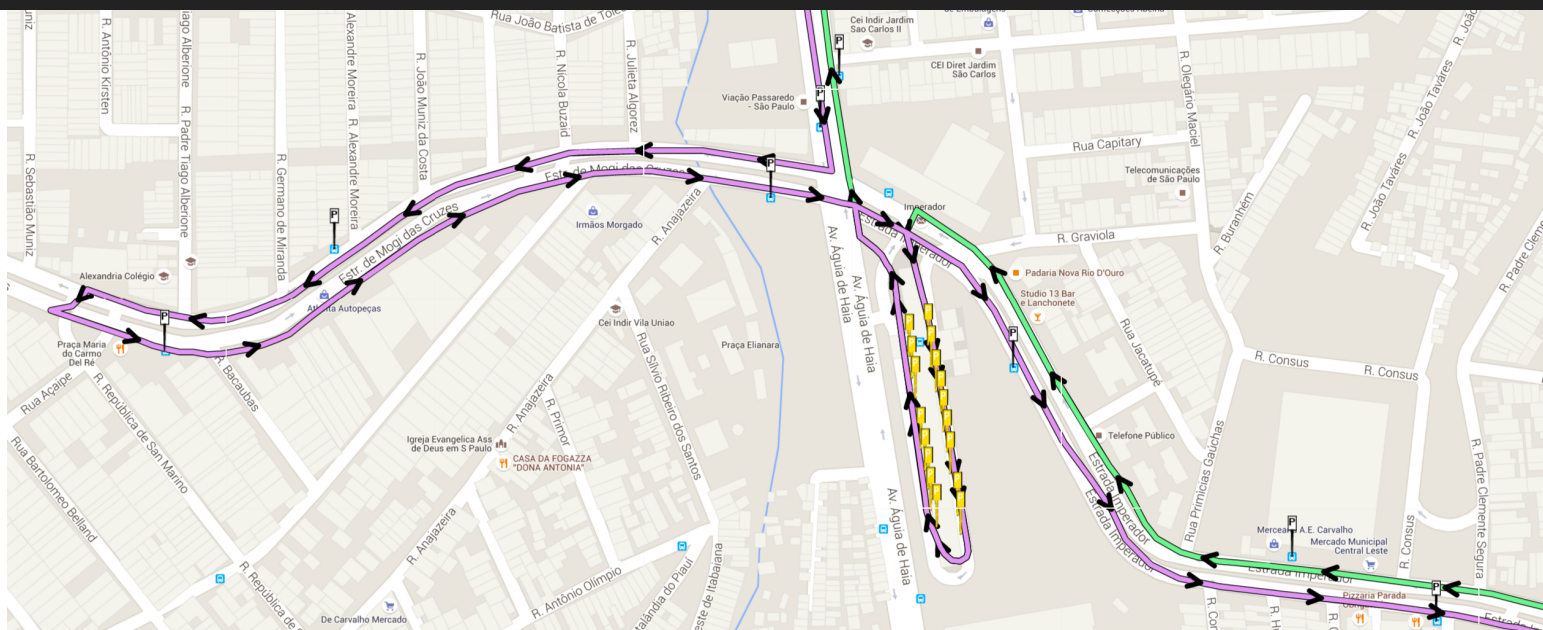
Ocorre se numa mesma linha um ponto é percorrido 2 vezes formando um loop. Isto será sinalizado pois trata-se de uma particularidade da linha, o qual pode confundir o passageiro.



Linha 2755-10 Metrô Penha

PONTOS IGUAIS NA IDA E NA VOLTÀ

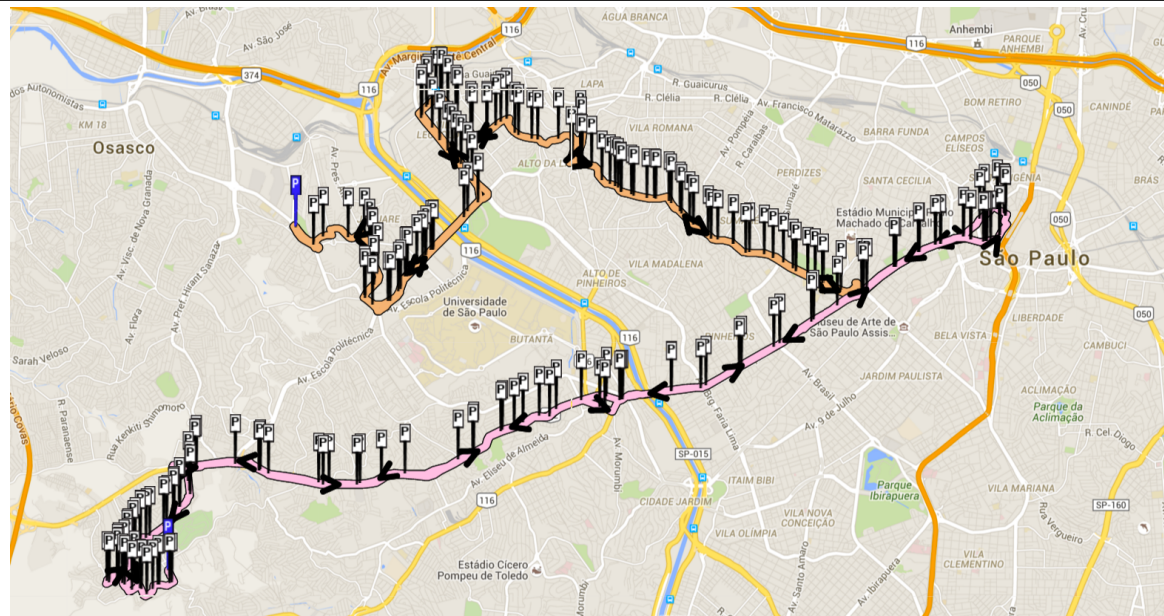
Ocorre se tanto no sentido TP-TS e TS-TP passamos por um mesmo ponto. Isto prejudica a identificação das viagens de um ônibus no AVL.



Linha: 2756-10-1 - Guaianazes (lilas) e 2756-10-1 - Metrô Patriarca (verde)

QUANTIDADE DE SHAPE

Duas linhas com o mesmo comprimento de linha, com o mesmo número de pontos mas com número diferente de mudanças de direção no trajeto podem ter uma alteração na sua velocidade média.

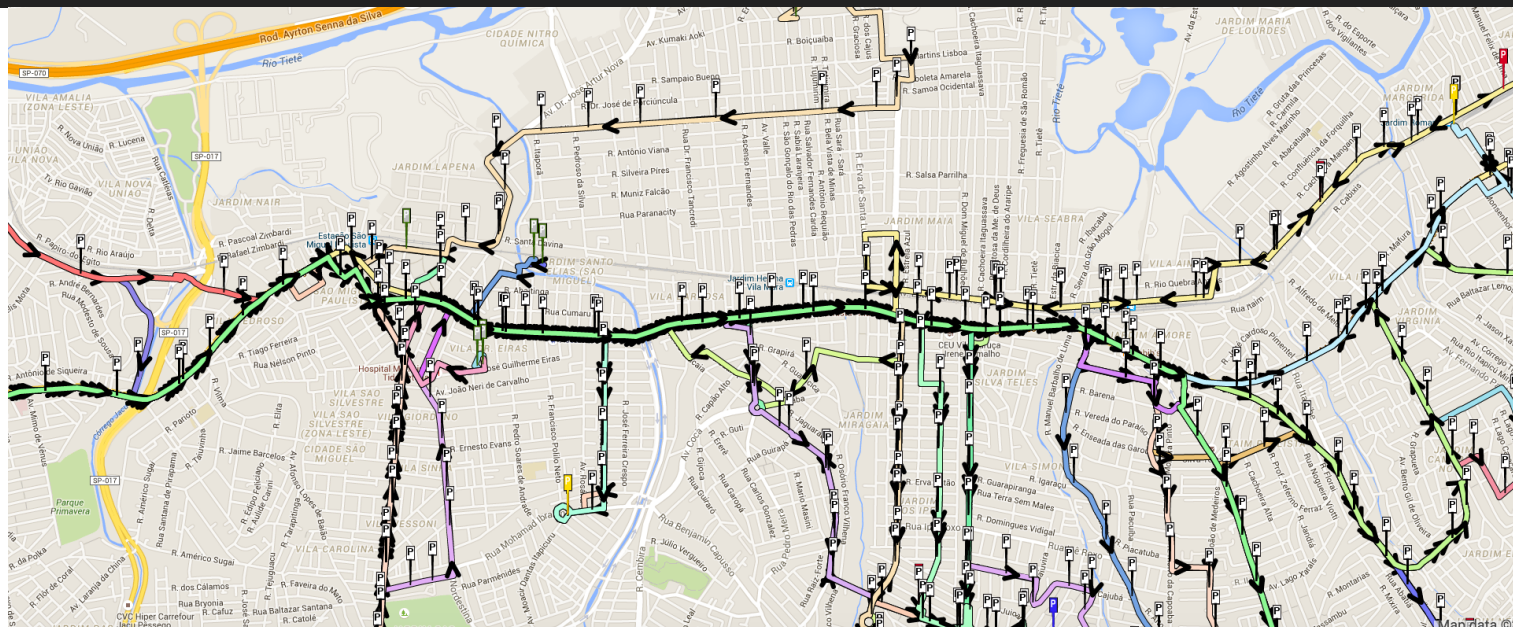


7903-10-0 - Praça Ramos de Azevedo (rosa) VM = 14.7 Km/h

7282-10-0 - Praça Ramos de Azevedo (laranja) VM = 10.9 Km/h

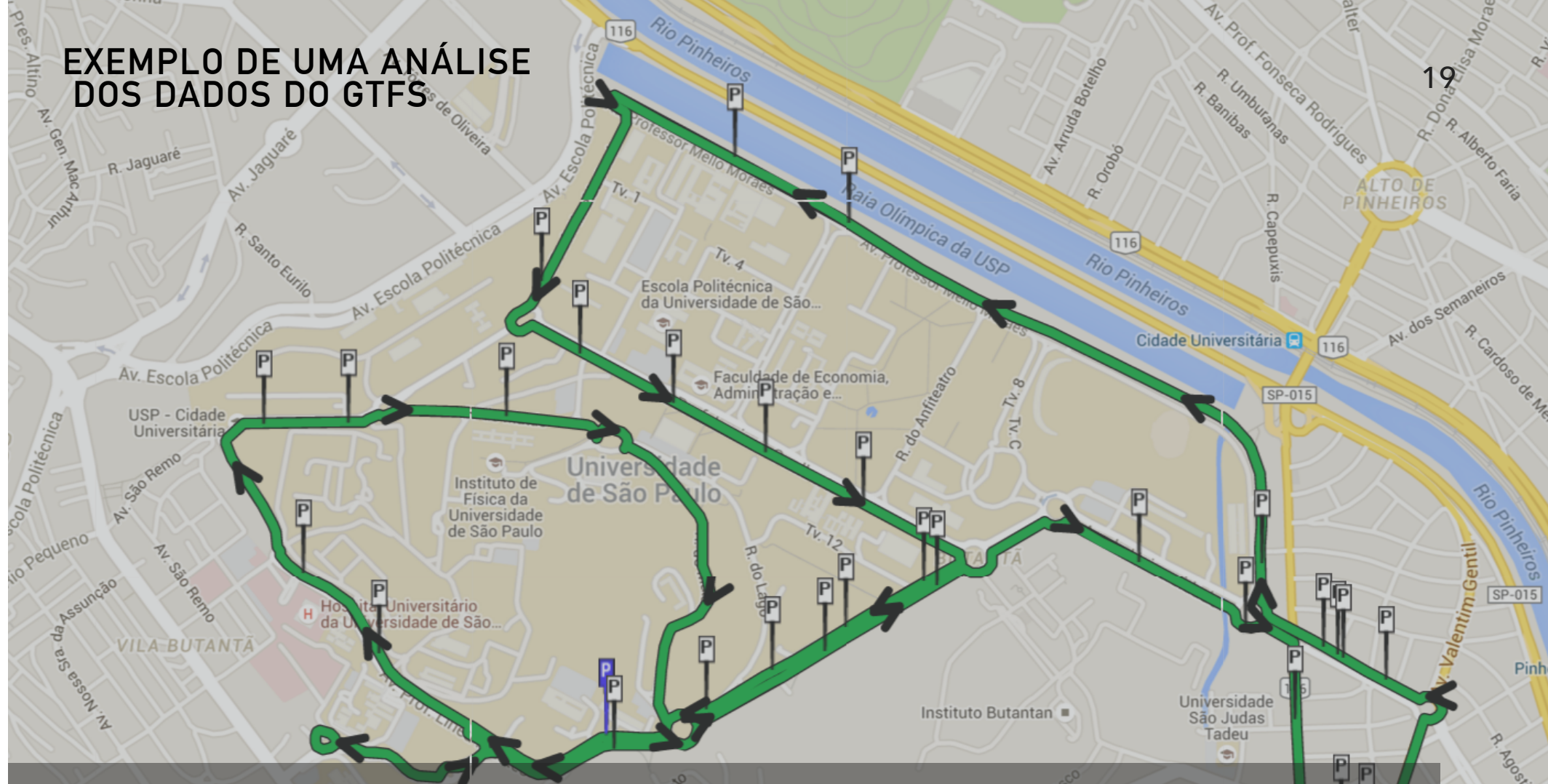
COMPARTILHAMENTO DE TRECHOS

Sabendo quantos trechos compartilhados tem uma linha pode-se verificar a sua flexibilidade para a cidade. Quanto maior o número de linhas que compartilham trechos, melhor será a fluidez na locomoção.



Trecho da marechal Tito (34 linhas que compartilham um determinado trecho)

EXEMPLO DE UMA ANÁLISE DOS DADOS DO GTFS



Linha 8022-10-0 : Cidade Universitária - Circular 2

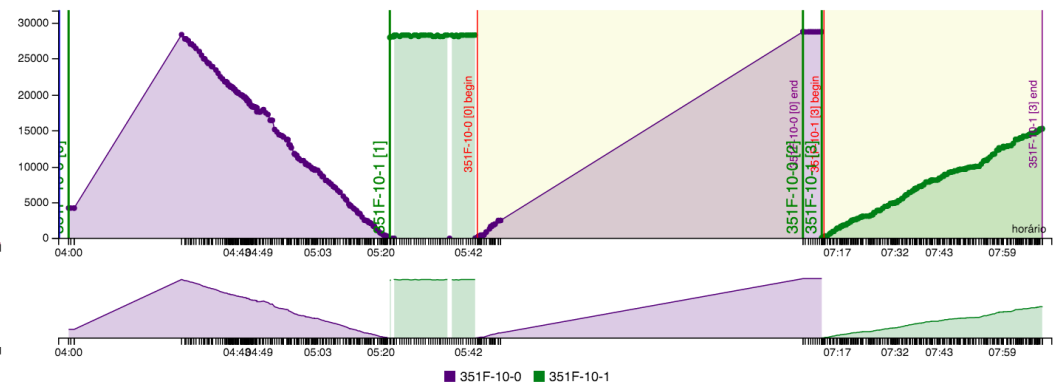
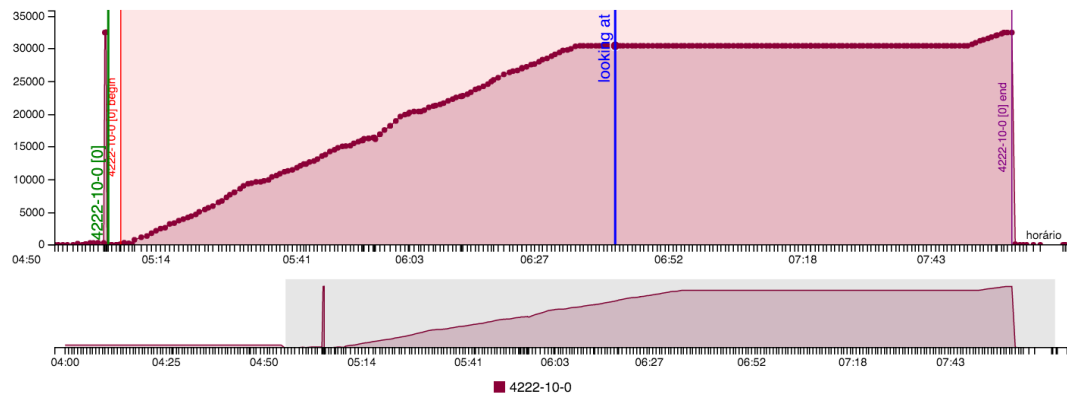
**Pontos Duplicados: 2 Pontos de Ida e Volta: 0 Quantidade de Shapes: 432
Compartilha trechos: 71% e Maior Compartilhamento: Vital Brasil (18 linhas)**

**Media: Pontos Duplicados: 184 linhas (total) Pontos de Ida e Volta: 520
linhas (total) Quantidade de Shapes: 330 Compartilha : 91% - 19 linhas**

ANALISE DOS DADOS DO AVL

Nesta etapa foram feitas análises sobre os dados emitidos pelos ônibus na cidade num determinado dia.

Estas análises tem como objetivo **identificar os detalhes das viagens**. Com isso é possível identificar os cumprimentos das viagens, um bom condutor e também os problemas da transmissão do AVL.



DEMO

BIBLIOGRAFIA

- ▶ Anthony M. Townsend: Smart Cities: Big Data, Civic Hackers, and the Quest for a New Utopia.
- ▶ SPTrans: <http://www.sptrans.com.br/desenvolvedores/>
- ▶ Banco de dados: <https://www.mongodb.org/>
- ▶ Plataforma Meteor: <https://www.meteor.com/>
- ▶ Google Feed Validator: <https://developers.google.com/transit/tools?hl=en>
- ▶ Visualização de Dados: www.c3js.org/

Obrigado!!! Dúvidas?

Apoio:



IME-USP

