

## Análise e visualização de dados do transporte público em São Paulo

### Introdução

Cidades inteligentes:



Sistema de ônibus em São Paulo (out/2014):

**9.988.278** Km/dia  
**18.733** pontos  
**25.568** arestas  
**2.204** linhas  
**14.663** ônibus ativos

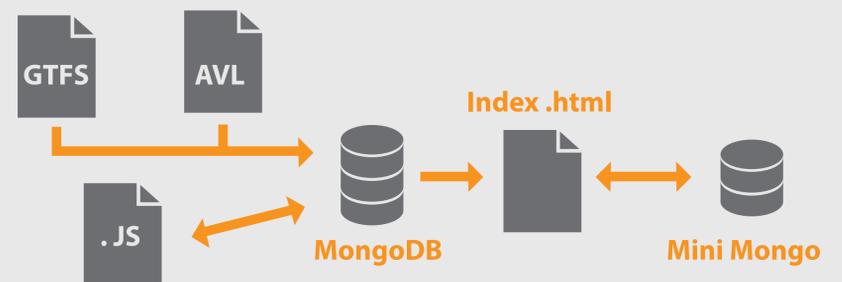
partidas:  
**199.053** dias úteis  
**192.229** sábados  
**183.343** domingos

Objetivo: explorar e analisar os dados para descobrir e visualizar informações do sistema de transporte por ônibus da cidade de São Paulo.

### Método

Usar técnicas computacionais e estatísticas para:

- (i) explorar e analisar dados
- (ii) identificar anomalias (situações fora da normalidade)
- (iii) criar e testar métricas para quantificar a qualidade dos dados.



### Resultados



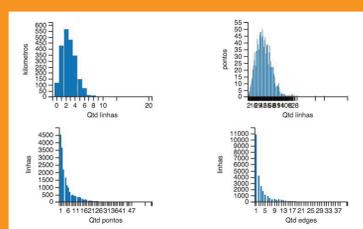
**GTFS (General Transit Feed Specification)** é uma especificação que define um formato para troca de informações de tabela de horários do transporte público com as posições geográficas associadas.

**Médias:**

**44** pontos por linha **26.7%** linhas circulares

**350m** de distância entre pontos **5** linhas por ponto

**15km** por linha **17h** operação linha



Foram criadas e testadas **19 métricas** sobre as informações disponíveis, entre elas:

- Quantidades de trechos com velocidade irreal
- Quantidade de pontos
- Quantidade de pontos duplicados
- Porcentagem da confiabilidade dos trechos

**Curiosidades:**

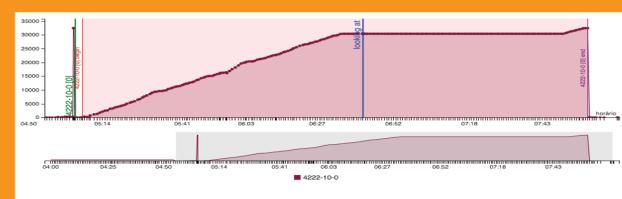
Na linha 372F-10-1 - Univ. São Judas Tadeu - tem 1km 5 pontos sai a cada 7 min

No ponto de onibus R. Rodovalho Júnior, 637 passa 60 linhas.



**AVL (Automatic Vehicle Location)** - 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

Nesta etapa foram analisados dados emitidos pelos ônibus num determinado dia (21/10/2014), com objetivo de identificar os detalhes das viagens.



Por meio de um gráfico de distância x tempo conseguimos identificar a qualidade da viagem. Na figura acima vimos uma viagem completa realizada.

A análise também é feita por meio da geolocalização do veículo em determinado momento



Na análise feita no dia 21/10/2014 temos que 45% dos dados do AVL são ruins.

\*Os gráficos referentes as informações coletadas estão na apresentação digital



### Referências:

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/>  
 Icones: <https://thenounproject.com/>

Google Feed Validator: <https://developers.google.com/transit/tools?hl=en>  
 Visualização de Dados: [www.c3js.org/](http://www.c3js.org/)

