

WORKSHOP: SUSTENTABILIDADE E INCLUSÃO SOCIAL NA GESTÃO DE  
TRANSPORTES PÚBLICOS: TÉCNICAS DE APOIO À DECISÃO

# DESCOBERTA DE CONHECIMENTO EM DADOS GEOESPACIAIS DE TRANSPORTES PÚBLICOS

14 DE DEZEMBRO DE 2021

**FCT**  
Fundação  
para a Ciência  
e a Tecnologia

**COMPETE**  
2020

**PORTUGAL**  
2020



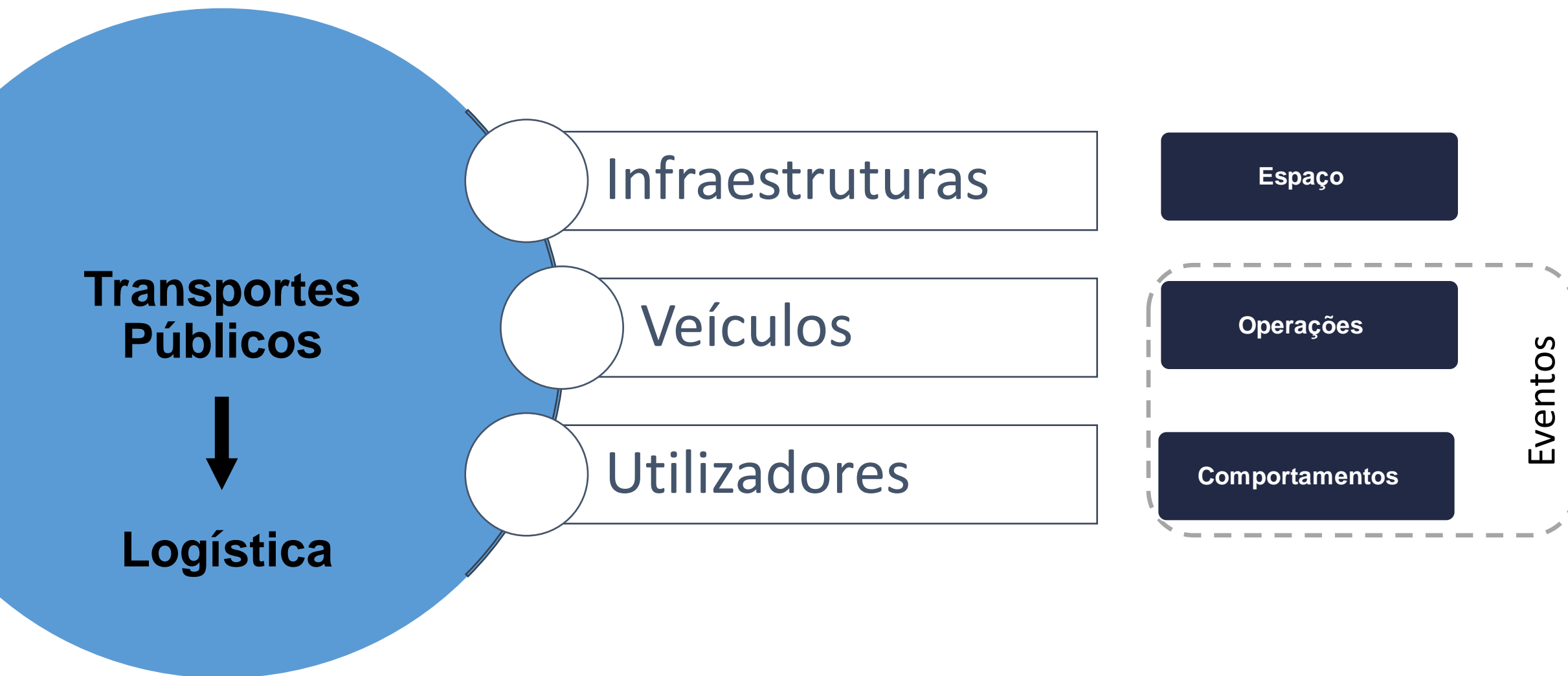
UNIÃO EUROPEIA  
Fundo Europeu  
de Desenvolvimento Regional

(PTDC/ECI-TRA/32053/2017 and POCI-01-0145-FEDER-032053)



JOEL RIBEIRO

# DESCOBERTA DE CONHECIMENTO EM DADOS GEOESPACIAIS



# ANÁLISE DE ACESSIBILIDADE: TRANSPORTES PÚBLICOS DA AMP

## Dados

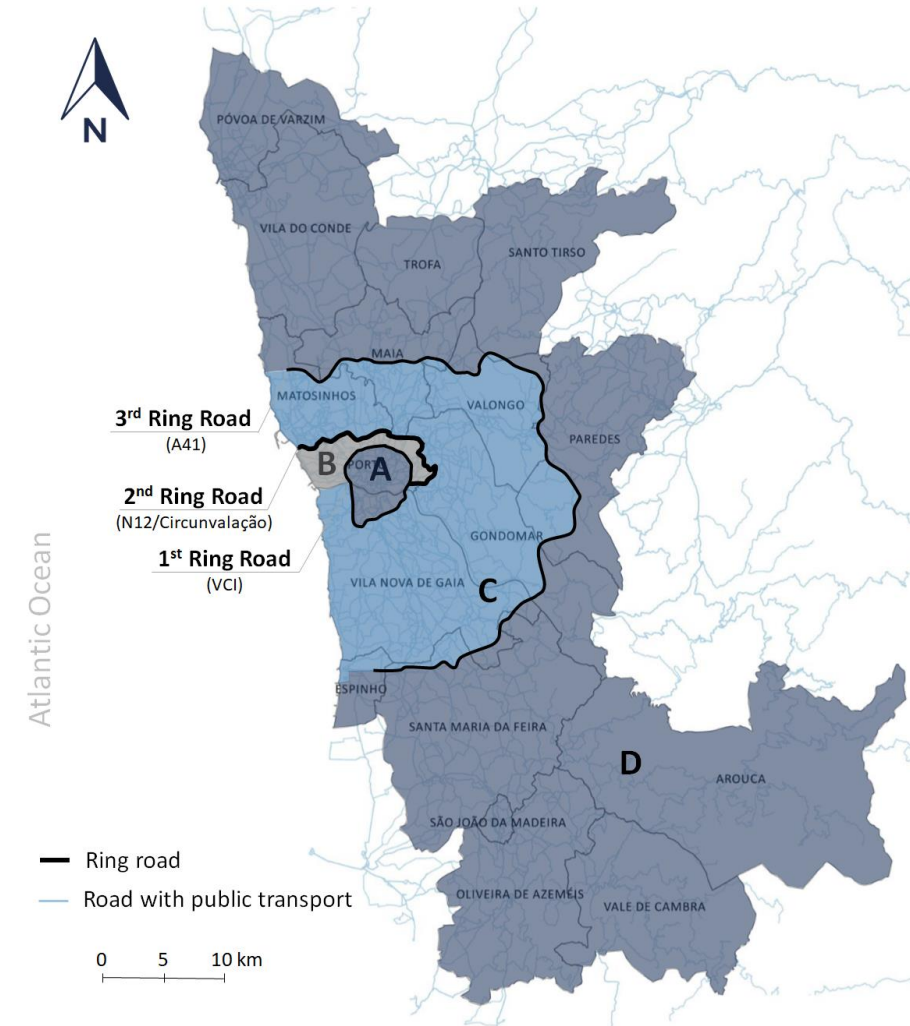
- Rede de transportes públicos
- Demografia
- Características antropométricas

## Método

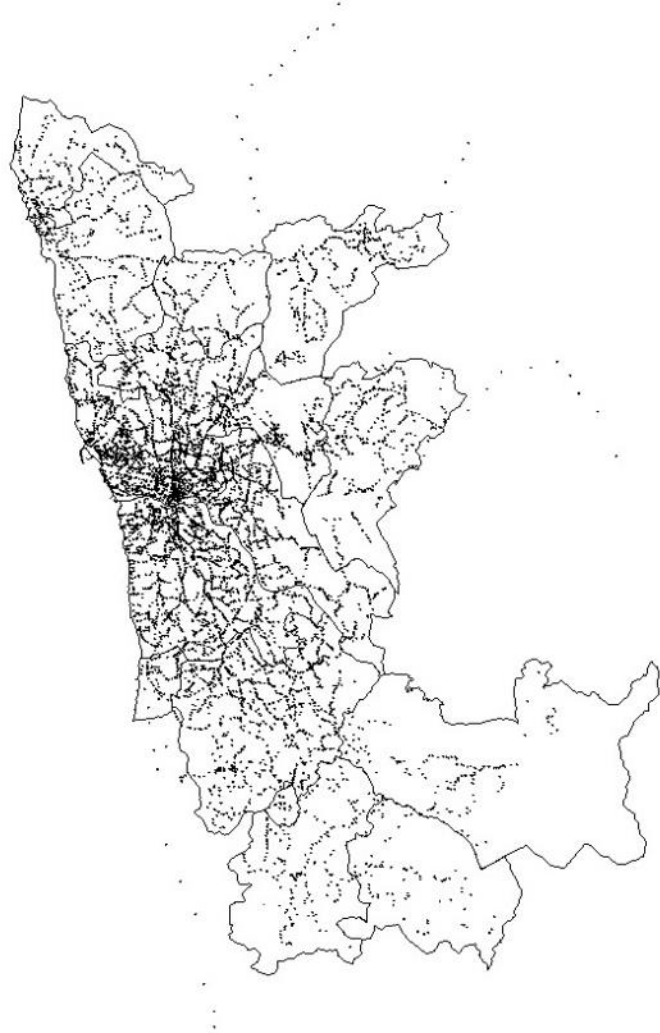
- Predisposição para caminhar até aos transportes públicos
- Capacidade para caminhar (6MWD)
- Cálculo de isolinhas
- Análise de cobertura

## Objetivos

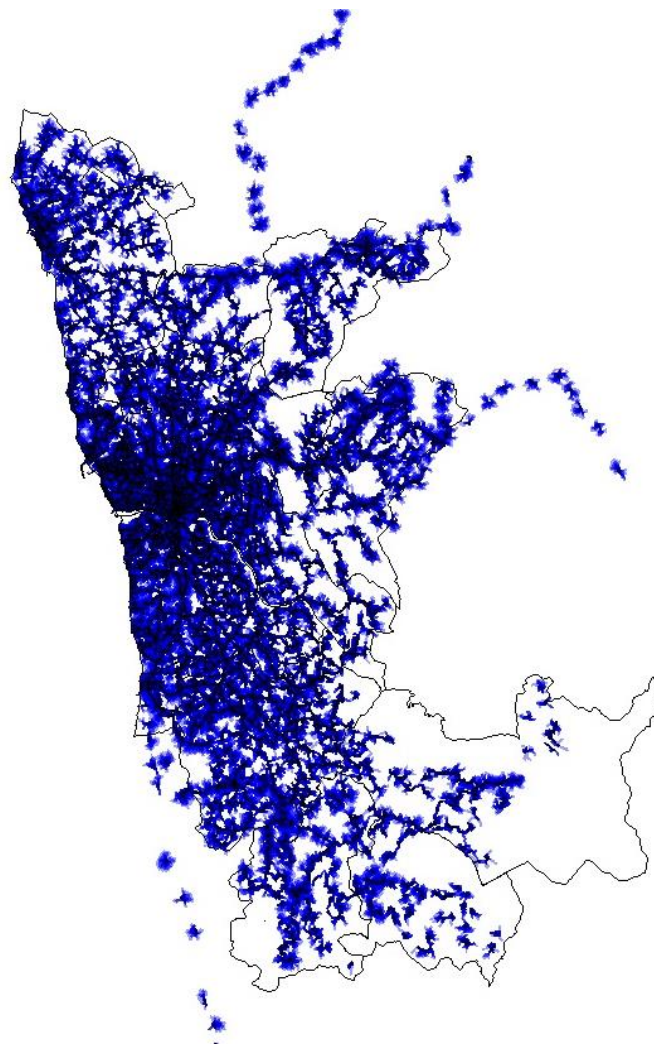
- Identificação de desigualdades (e.g. exclusão social)
- Quantificação da cobertura da rede de transportes
- Avaliação da qualidade de serviço



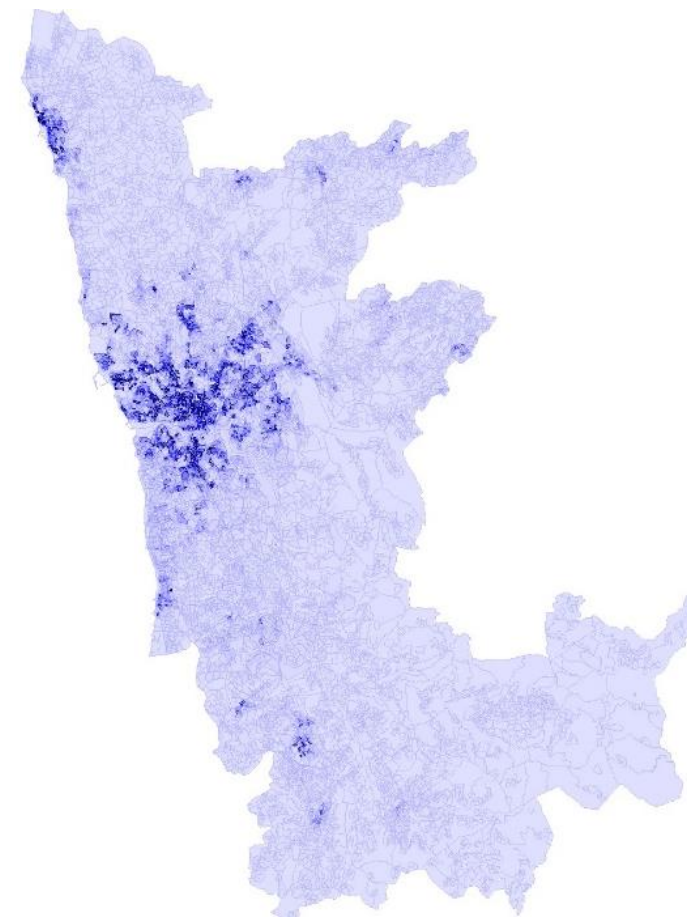
# ANÁLISE DE ACESSIBILIDADE: TRANSPORTES PÚBLICOS DA AMP



Estações + Paragens



Isolinhas



Demografia

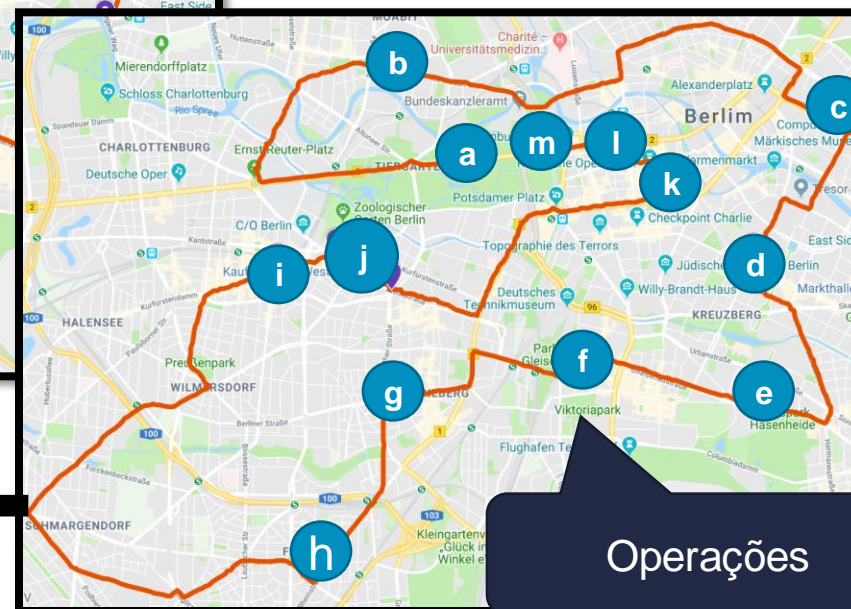
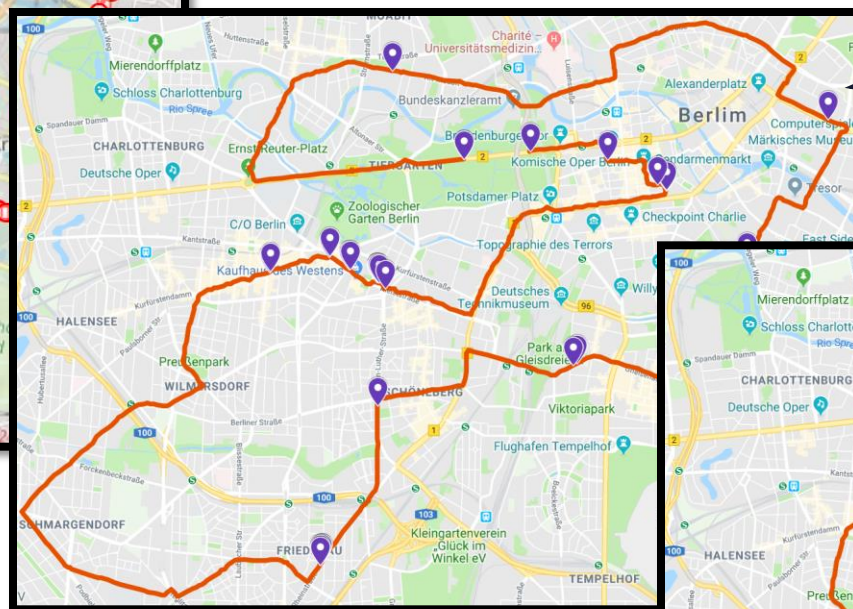
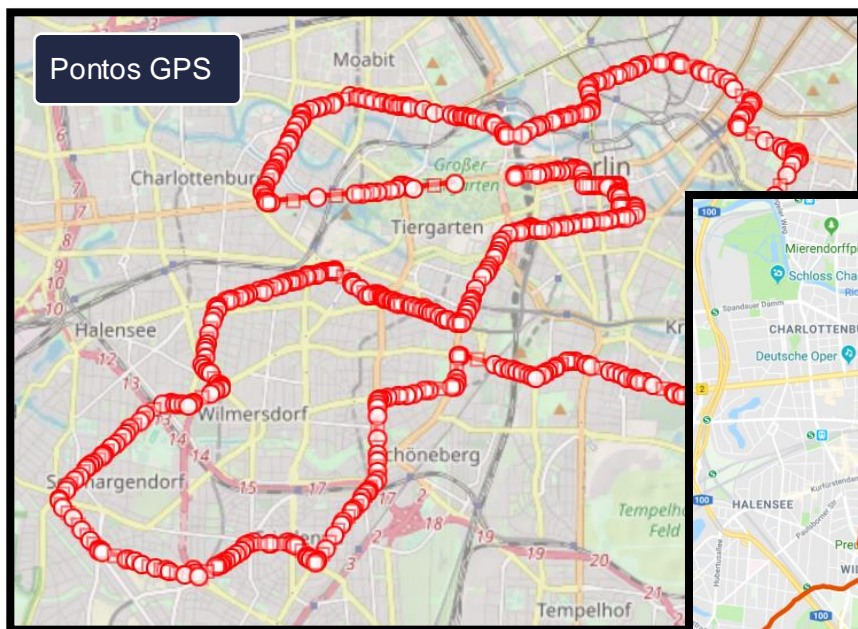
# ANÁLISE DE ACESSIBILIDADE: TRANSPORTES PÚBLICOS DA AMP

## População (%) sem acesso aos transportes públicos

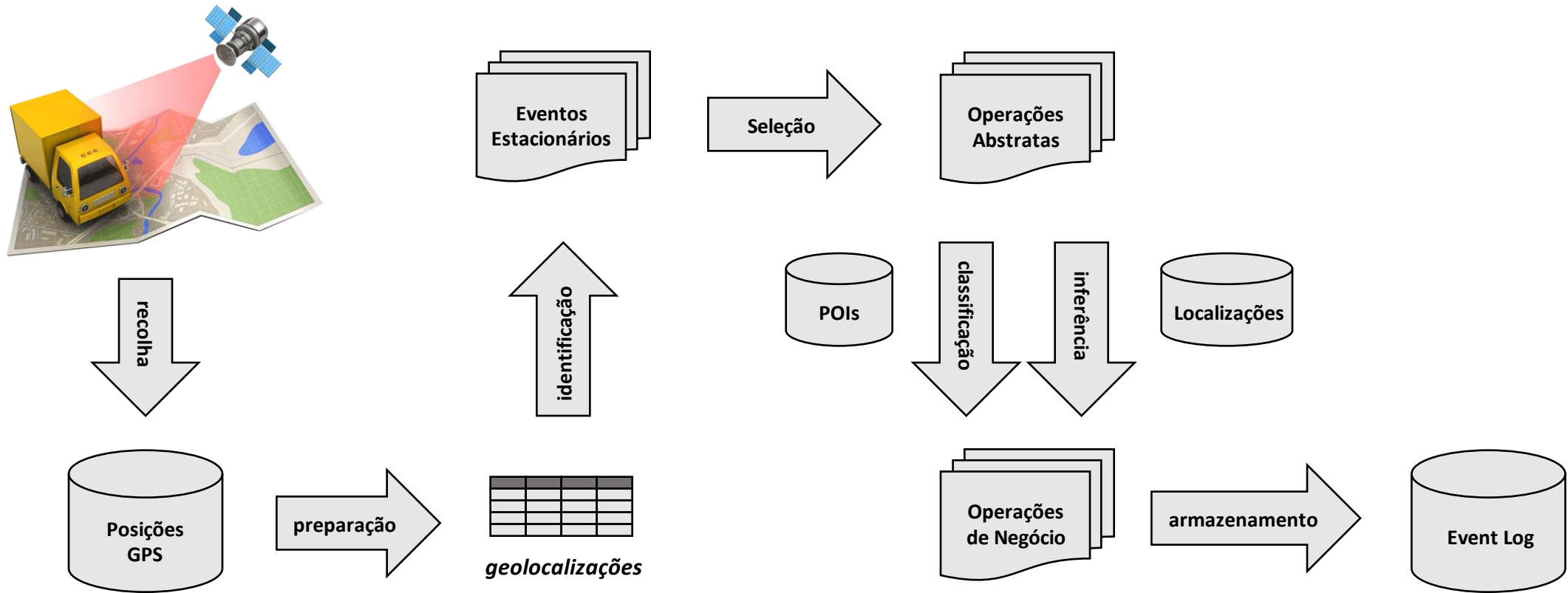
Area	Total	Resident Population (Males / Females)					
		5-9	10-13	14-19	20-24	25-64	65+
Zone A	0.5 (0.5 / 0.5)	0.7 / 0.7	0.5 / 0.5	0.7 / 0.4	0.4 / 0.5	0.5 / 0.5	0.5 / 0.6
Zone B	0.5 (0.4 / 0.6)	0.9 / 0.9	0.5 / 0.6	0.3 / 0.5	0.3 / 0.3	0.5 / 0.6	0.4 / 0.6
Zone C	5.6 (5.4 / 5.7)	6.6 / 6.5	5.2 / 5.5	4.7 / 5.3	4.8 / 5.3	5.4 / 5.7	5.5 / 6.1
Zone D	20.3 (19.7 / 20.9)	21.0 / 21.5	19.2 / 19.7	18.7 / 19.8	18.3 / 19.6	19.5 / 20.5	21.5 / 23.1
AMP	10.6 (10.4 / 10.8)	11.9 / 12.1	10.4 / 10.8	10.0 / 10.8	9.8 / 10.5	10.4 / 10.7	10.6 / 11.1

Resident population by age group and gender

# DESCOBERTA DE CONHECIMENTO EM DADOS GEOESPACIAIS: VEÍCULOS



# DETEÇÃO AUTOMÁTICA DE OPERAÇÕES COM VEÍCULOS



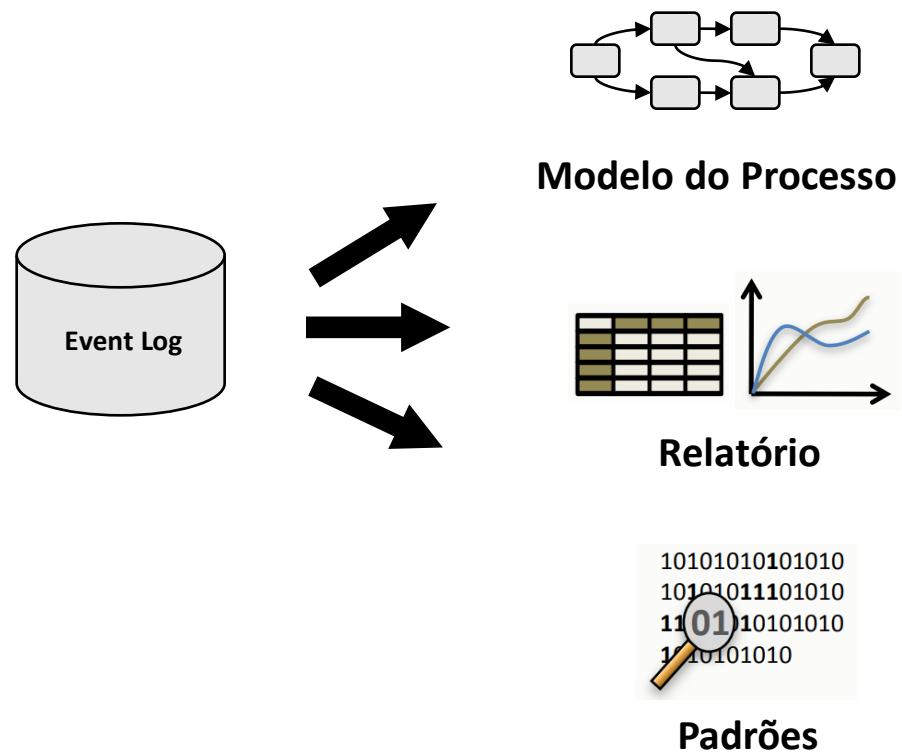
# EXTRAÇÃO DE CONHECIMENTO EM DADOS DE EVENTOS

## Process Mining

- Descoberta do Processo
- Verificação de Conformidade
- Análise de Performance

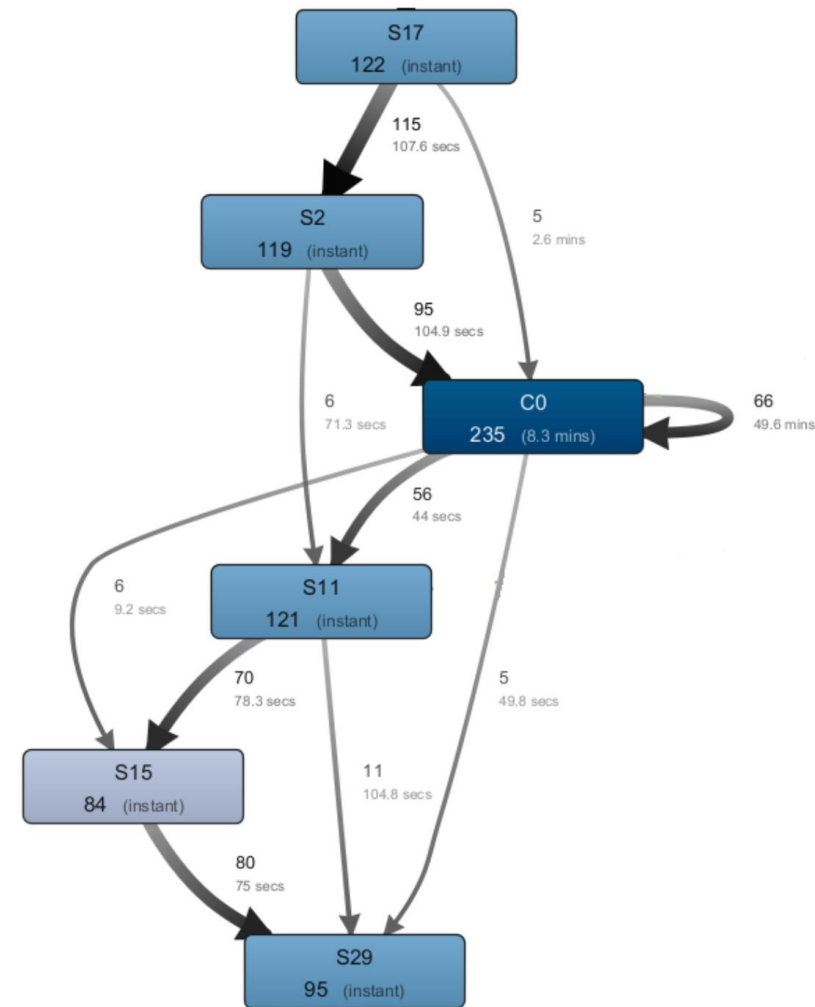
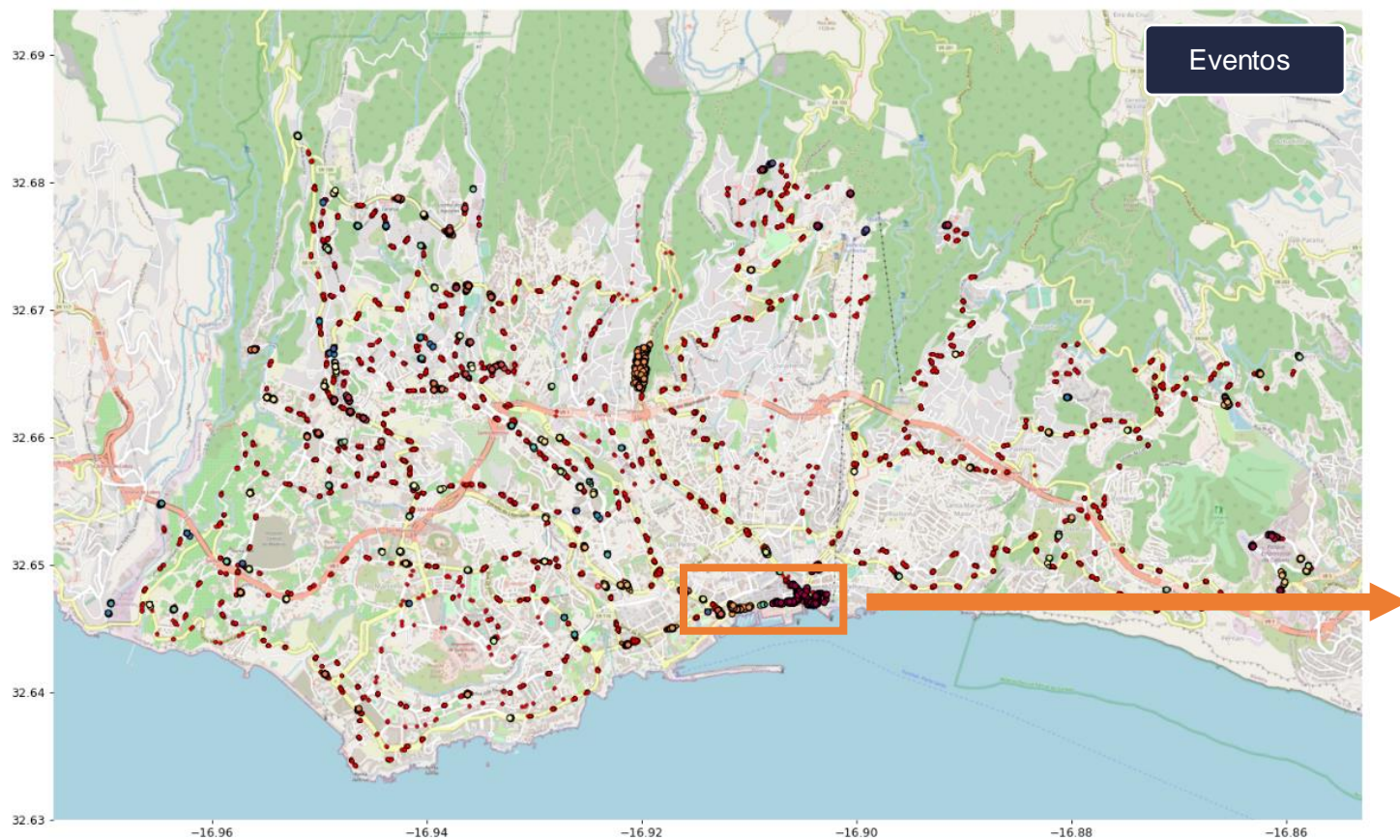
## Data Mining

- Descoberta de Padrões

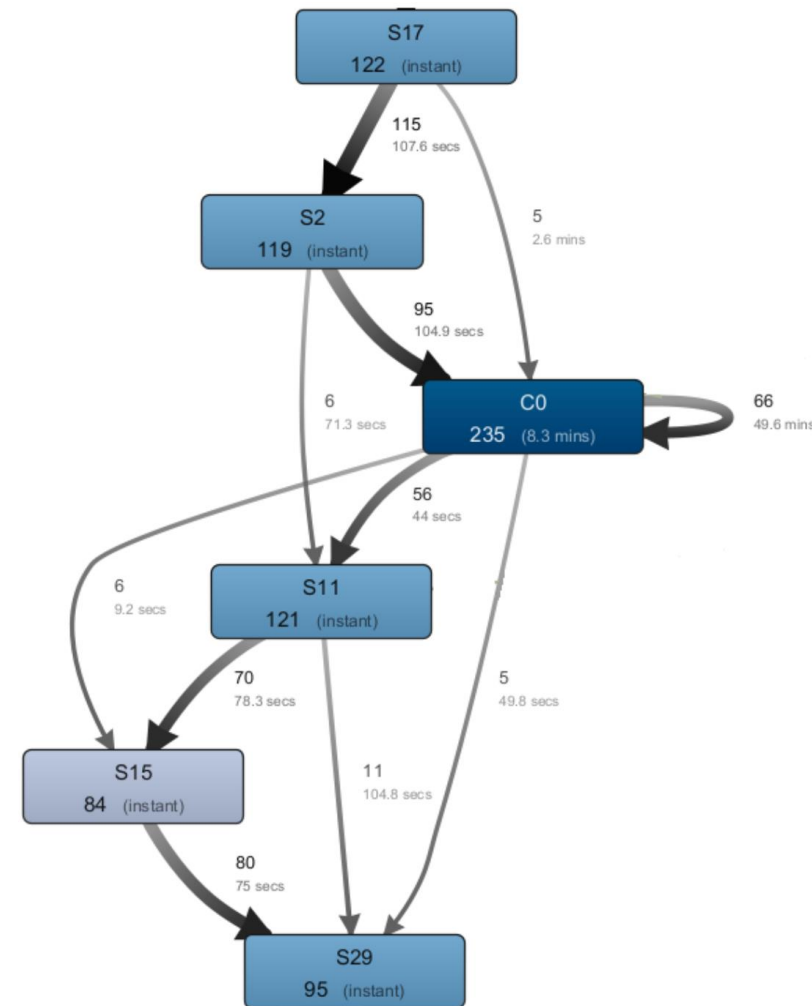




# DESCOBERTA DO PROCESSO: TRANSPORTES PÚBLICOS DO FUNCHAL



# DESCOBERTA DO PROCESSO: TRANSPORTES PÚBLICOS DO FUNCHAL



# VERIFICAÇÃO DE CONFORMIDADE: EMPRESA DE LOGÍSTICA

Work plan	State	Vehicle	Start time	Operation 1	Operation 2	Operation 3	Operation 4
Work plan 1	Active	V04	08:15	Location A 08:19 - 09:10	Location B <i>non-executed</i>	-	-
Work plan 2	Finished	V10	07:30	Location C 07:15 - 7:50	Location D 09:22 - 09:57	Location E <i>non-executed</i>	Location F 10:22 - 11:30
Work plan 3	Planned	V04	14:00	Location A <i>non-executed</i>	Location B <i>non-executed</i>	-	-
Work plan 4	Finished	V07	07:30	Location E 07:23 - 7:38	Location F 08:19 - 08:40	Location B 09:51 - 12:13	-

Indicator	Vehicle		
	V1	V2	V3
Geolocation entries	29328	26329	26335
in stationary event (average cadence)	15722 (53.6%) 92 s	17162 (65.2%) 102 s	21067 (80.0%) 102 s
Stationary events	589	1676	658
with known location in work plan	155 (26.3%) 210 (35.7%)	225 (13.4%) 804 (48.0%)	368 (55.9%) 431 (65.5%)
Work plans	18	21	56
fully fulfilled	17 (94.4%)	18 (85.7%)	52 (92.9%)
partially fulfilled	1 (5.6%)	2 (9.5%)	4 (7.1%)
Planned operations	36	67	113
with detected execution	35 (97.2%)	63 (94.0%)	108 (95.6%)

# ANÁLISE DE PERFORMANCE: EMPRESA DE LOGÍSTICA

Work plan	State	Vehicle	Start time	Operation 1	Operation 2	Operation 3	Operation 4
Work plan 1	Active	V04	08:15	Location A 08:19 - 09:10	Location B <i>non-executed</i>	-	-
Work plan 2	Finished	V10	07:30	Location C 07:15 - 7:50	Location D 09:22 - 09:57	Location E <i>non-executed</i>	Location F 10:22 - 11:30
Work plan 3	Planned	V04	14:00	Location A <i>non-executed</i>	Location B <i>non-executed</i>	-	-
Work plan 4	Finished	V07	07:30	Location E 07:23 - 7:38	Location F 08:19 - 08:40	Location B 09:51 - 12:13	-

Indicator	Vehicle		
	V1	V2	V3
Work plans			
Average throughput time	12:14:33	05:14:19	03:20:02
Average start time (executed vs planned)	-03:14:16	00:12:55	-00:52:28
Average delay	00:07:38	02:47:23	00:39:48
Started on time	16 (88.9%)	15 (78.9%)	41 (74.5%)
Planned operations			
Average execution time	04:48:18	01:01:44	01:18:23

# DESCOBERTA DE CONHECIMENTO EM DADOS GEOESPACIAIS: UTILIZADORES



- **Evento:** validação de título de transporte
- **Caso:** sequência de validações de um título ao longo de um determinado dia

# DESCOBERTA DE PADRÕES: TRANSPORTES PÚBLICOS DA AMP

## Data Mining

- Conjuntos frequentes

50.66% of the ticket validations are from 6:00 to 11:59 on weekdays

- Regras de associação

IF ticket validation from 6:00 to 8:59 on weekdays  
THEN ticket type is likely to be a monthly pass

- Padrões sequenciais

The sequence Trindade → not(Trindade) → Trindade happens in 4.32% of the daily journeys

# MULTIDIMENSIONAL SUBGROUP DISCOVERY ON EVENT LOGS

Senhora da Hora → not(Senhora da Hora) → Senhora da Hora

210000 validações

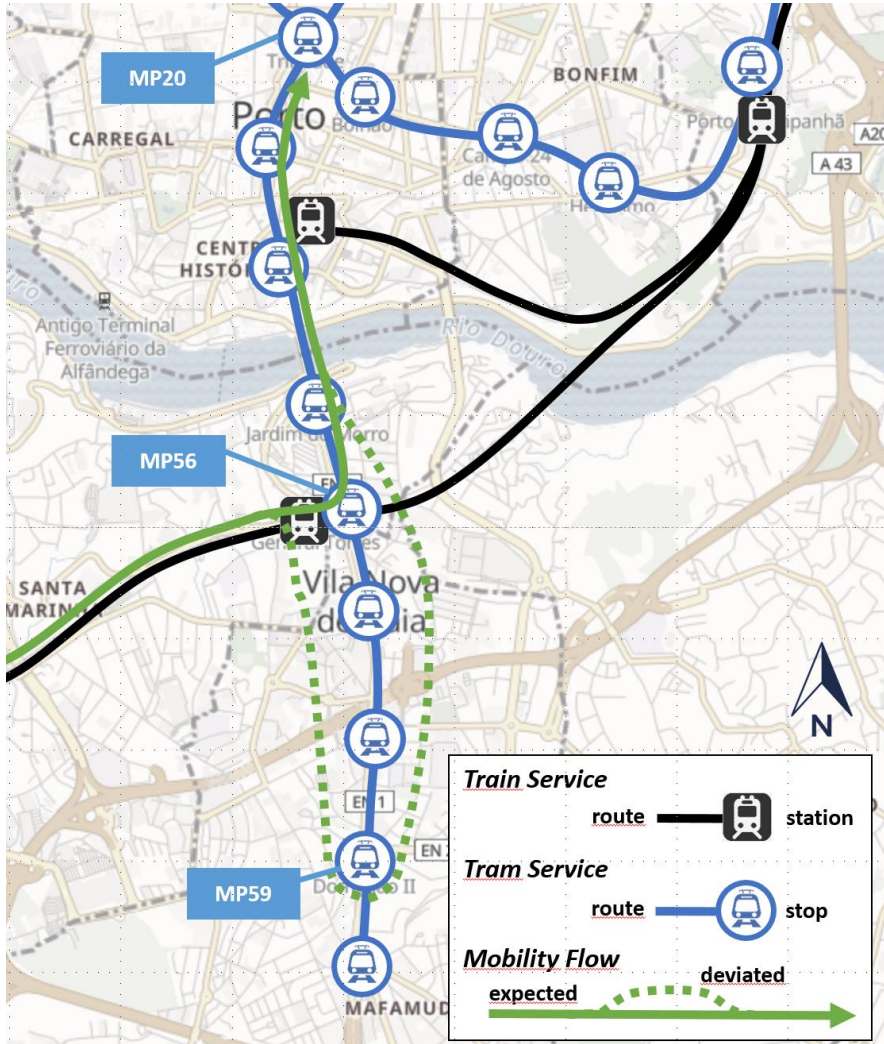
## 1. *Tipo de dia*

- **+2.7%** de validações nos **dias úteis**
- **-2.8%** de validações ao **fim de semana**

## 2. *Tipo de bilhete*

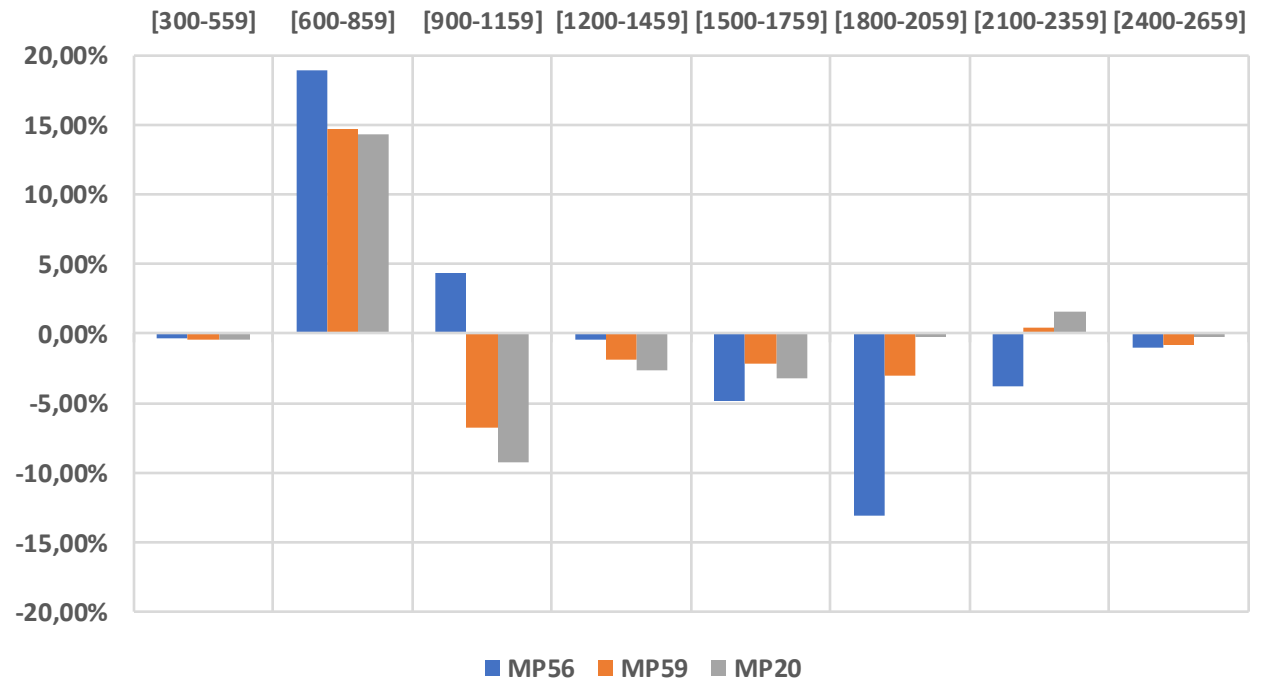
- **+10.8%** de validações com **titulo únicos**
- **-3.8%** de validações com **passes**

# ANÁLISE DE COMPORTAMENTOS: TRANSPORTES PÚBLICOS DA AMP



MP56 → MP59 → MP20

- *Hora da validação*





# SUMÁRIO

- **Descoberta do processo tal como está a ser executado**
- **Verificação de conformidade (realidade vs esperado)**
- **Deteção de pontos de gargalo e eventos não conformes no processo**
- **Quantificação de tempos de operação e atrasos**
- **Análise de performance**
- **Identificação de padrões e comportamentos**
- **Avaliação de qualidade de serviço**

# Referências

- Ribeiro, J.; Fontes, T.; Soares, C.; Borges, J.L. (2020), "Process discovery on geolocation data". Transportation Research Procedia 47: 139-146. <http://dx.doi.org/10.1016/j.trpro.2020.03.086>.
- Ribeiro, J.; Fontes, T.; Soares, C.; Borges, J.L. (2021), "Accessibility as an indicator to estimate social exclusion in public transport". Transportation Research Procedia 52 (2021): 740-747. <http://dx.doi.org/10.1016/j.trpro.2021.01.019>.
- Tavares, J.; Ribeiro, J.; Tânia F. (2021), "Discovery of Vehicle-Based Operations from Geolocation Data". 24th Euro Working Group on Transportation, Aveiro, 2021.
- Tavares, J. (2021), "Discovery of Transport Operations from Geolocation Data", (Advisor: J. Ribeiro, Co-advisor: T. Fontes). Dissertation presented for obtaining the MSc degree in Electrical and Computer Engineering, FEUP.
- Ribeiro, J.; Tavares, J.; Fontes, T. (2021), "Real-time detection of logistics vehicle-based geospatial operations". EAI INTSYS 2021.
- Ribeiro, J.; Fontes, T.; Soares, C.; Borges, J.L. (submitted) "Multidimensional Subgroup Discovery on Event Data". IEEE Transactions on Knowledge and Data Engineering.

Campus da Faculdade de  
Engenharia da Universidade  
do Porto

Rua Dr. Roberto Frias

4200-465 Porto, Portugal



T +351 222 094 000

[info@inesctec.pt](mailto:info@inesctec.pt)

[www.inesctec.pt](http://www.inesctec.pt)

