Icon

Description automatically generated

Francisca Amélia de Fernandes Barros

**Dashboards analíticos para a Indústria das Cidades Inteligentes**

Analytical Dashboards for the Smart Cities Industry

Icon

Description automatically generated

Francisca Amélia de Fernandes Barros

**Dashboards analíticos para a Indústria das Cidades Inteligentes**

Analytical Dashboards for the Smart Cities Industry

Dissertação de Mestrado

Mestrado Integrado em Engenharia e Gestão de Sistemas de Informação

Ano Letivo 2021/2022 – 5º Ano Curricular

Trabalho efetuado sob a orientação do

**Professor Manuel Filipe Vieira Torre dos Santos e Carlos Filipe Portela**

DIREITOS DE AUTOR E CONDIÇÕES DE UTILIZAÇÃO DO TRABALHO POR TERCEIROS

Este é um trabalho académico que pode ser utilizado por terceiros desde que respeitadas as regras e boas práticas internacionalmente aceites, no que concerne aos direitos de autor e direitos conexos.

Assim, o presente trabalho pode ser utilizado nos termos previstos na licença abaixo indicada.

Caso o utilizador necessite de permissão para poder fazer um uso do trabalho em condições não previstas no licenciamento indicado, deverá contactar o autor, através do RepositóriUM da Universidade do Minho.

https://licensebuttons.net/l/by-nc/3.0/88x31.png

**Atribuição-NãoComercial   
CC BY-NC**

https://creativecommons.org/licenses/by-nc/4.0/

AGRADECIMENTOS

Gostaria de agradecer ao Professor Carlos Filipe Portela, pelo estímulo e pela total disponibilidade que teve para me orientar nesta tese.

À minha querida mãe e querido namorado a minha enorme gratidão pela paciência e apoio inigualável, sem os quais não teria sido possível realizar este percurso académico.

Para vocês, Bella, Ninfo e Bugs, obrigada pela linda jornada que construímos e por estarem sempre presentes em todas as etapas mais difíceis e também por celebrarem comigo todas as vitórias. Um agradecimento especial à minha equipa de "Informática", Foguetinhos e Fitas pelo apoio e suporte. Obrigada, também ao pessoal das “Resis” pelos momentos de descontração que me proporcionaram e por estarem sempre presentes. Vocês, os 42, vou-vos levar sempre comigo no coração. Fizeram destes cinco anos académicos uma história bonita que jamais esquecerei.

Pretendo ainda deixar uma mensagem de agradecimento a todas as pessoas que, de alguma forma, contribuíram para que fosse possível evoluir ao longo desta viagem, particularmente, às minhas "nove mais uma" de Viana e ao pessoal do "Menina estás à Janela".

DECLARAÇÃO DE INTEGRIDADE

Declaro ter atuado com integridade na elaboração do presente trabalho académico e confirmo que não recorri à prática de plágio nem a qualquer forma de utilização indevida ou falsificação de informações ou resultados em nenhuma das etapas conducente à sua elaboração.

Mais declaro que conheço e que respeitei o Código de Conduta Ética da Universidade do Minho.

DASHBOARDS ANÁLITICOS PARA A INDÚSTRIA DAS SMART CITIES

RESUMO

A presente dissertação decorre no âmbito da Dissertação do 5º Ano do Mestrado Integrado em Engenharia e Gestão de Sistemas de Informação da Universidade do Minho. As cidades desempenham um papel crucial no desenvolvimento sustentável e na vida quotidiana da população. O trabalho proposto centra-se na questão de como, através de sensores e inteligência artificial, consegue-se melhorar a indústria das cidades. Problemas derivados do trânsito e do uso excessivo dos automóveis, provocam congestionamento, poluição, dificuldade em encontrar um parque de estacionamento e um maior número de acidentes. Deste modo, é esperado o desenvolvimento de um modelo descritivo-preditivo que, em tempo-real, contribua para a tomada de decisão a nível dos transportes públicos e/ou parques de estacionamento.

Inicialmente, foi contextualizada a era da indústria 4.0 e o que levou ao surgimento das cidades inteligentes. Assim, como, uma abordagem de diferentes áreas que evoluíram com as tecnologias emergentes desta revolução industrial. A importância de abordar áreas diferentes e conhecer o contexto de cada uma, leva a um mundo de conhecimento crucial para que a solução desenvolvida seja mais completa e eficiente. Sendo, a indústria das *Smart Cities*, uma indústria em constante evolução é importante acompanhar a progressão e explorar boas práticas já em vigor no mundo, mais concretamente, a nível de taxas de ocupação dos transportes públicos e de sistemas de gestão de parques de estacionamento.

Relativamente ao caso de estudo é apresentado um estudo analítico e um método de *Clustering* que teve como objetivo identificar quais os fatores que influenciam a tomada de decisão entre o uso de transportes públicos e automóveis, destacando seis variáveis como resultado do estudo. Assim como, uma exploração de diversas interfaces de programação de aplicativos (APIs), de forma, a garantir a recolha de todas as variáveis consideradas para o modelo inteligente. Seguidamente, e após uma análise de dados reais, foi desenvolvida uma *data warehouse*, um modelo OLAP e uma dashboard de visualização de dados para apoio à tomada de decisão. Deste modo, a solução desenvolvida é uma mais-valia para o processo de tomada de decisão e apresenta características fundamentais, como *drill-down* e *rollup* que permitem uma análise detalhada dos dados. Este documento apresenta assim a revisão de literatura e todo o trabalho desenvolvido.

**Palavras-chave:** Cidades Inteligentes, Indústria 4.0, Parques de Estacionamento e Transportes Públicos.

ANALYTICAL DASHBOARDS FOR THE SMART CITIES INDUSTRY

ABSTRACT

The present dissertation takes place within the 5th Year of the Integrated master’s in engineering and Management of Information Systems at the University of Minho. Cities play a crucial role in sustainable development and people's daily lives. The proposed work focuses on how it can improve the cities’ industry through sensors and artificial intelligence. Problems arising from traffic and excessive use of cars cause congestion, pollution, difficulty finding a car park and many accidents. Thus, it is expected to develop a descriptive-predictive model that, in real time, contributes to decision-making regarding public transport and car parks.

Initially, the era of Industry 4.0 was contextualized as what led to the emergence of smart cities. An approach to different areas that have evolved with the emerging technologies of this industrial revolution. The importance of approaching other sites and knowing the context of each one leads to a world of knowledge crucial for the developed solution to be complete and efficient. As the Smart Cities industry is in constant evolution, it is essential to follow the progression and explore promising practices already in place in the world, specifically in terms of occupancy rates of public transport and parking management systems.

Regarding the case study, analytical research and a Clustering method are presented, which aimed to identify which factors influence the decision-making between the use of public transport and cars, highlighting six variables as a result of the study. As well as an exploration of several application programming interfaces (APIs) to guarantee the collection of all variables considered for the intelligent model. Then, after an analysis of accurate data, a data warehouse, an OLAP model and a data visualization dashboard were developed to support decision-making. In this way, the developed solution is an asset to the decision-making process. It presents fundamental characteristics, such as drill-down and rollup, which allow a detailed data analysis. This document thus presents the literature review, and all the work developed.

**Keywords:** Car Parks, Industry 4.0, Public Transport and Smart Cities

ÍNDICE

[1. Introdução 1](#_Toc116652584)

[1.1. Contextualização do Projeto e Motivação 1](#_Toc116652585)

[1.2. Objetivos 2](#_Toc116652586)

[1.3. Estrutura do documento 3](#_Toc116652587)

[2. Revisão da Literatura 5](#_Toc116652588)

[2.1. Estratégia de Pesquisa 5](#_Toc116652589)

[2.2. Evolução da Indústria 5](#_Toc116652590)

[2.2.1. 1º Revolução Industrial 6](#_Toc116652591)

[2.2.2. 2º Revolução Industrial 7](#_Toc116652592)

[2.2.3. 3º Revolução Industrial 7](#_Toc116652593)

[2.2.4. 4º Revolução Industrial 7](#_Toc116652594)

[2.3. Pilares da Indústria 4.0 8](#_Toc116652595)

[2.4. A Indústria das Cidades Inteligentes 10](#_Toc116652596)

[2.4.1. Das Cidades tradicionais às Sustentáveis e Inteligentes 10](#_Toc116652597)

[2.4.2. Vantagens e Desvantagens das Cidades Inteligentes 12](#_Toc116652598)

[2.5. Projetos na Indústria 4.0 que impactam a decisão 13](#_Toc116652599)

[2.5.1. Área Têxtil 16](#_Toc116652600)

[2.5.2. Área Automóvel 20](#_Toc116652601)

[2.5.3. Cidades Inteligentes 23](#_Toc116652602)

[2.6. A indústria das Cidades 4.0 27](#_Toc116652603)

[2.6.1. Parques de Estacionamento 27](#_Toc116652604)

[2.6.2. Taxa de ocupação nos transportes públicos 30](#_Toc116652605)

[2.7. Visão Crítica do Estado da Arte 31](#_Toc116652606)

[3. Abordagem Metodológica, Materiais e Métodos 33](#_Toc116652607)

[3.1. Metodologias de Investigação 33](#_Toc116652608)

[3.1.1. Metodologia DSR 33](#_Toc116652609)

[3.1.2. Metodologia Caso de Estudo 36](#_Toc116652610)

[3.2. Metodologias Práticas 37](#_Toc116652611)

[3.2.1. Metodologia CRISP-DM 37](#_Toc116652612)

[3.2.2. Metodologia SCRUM 39](#_Toc116652613)

[3.3. Lista de Ferramentas 42](#_Toc116652614)

[4. Caso de Estudo 44](#_Toc116652615)

[4.1. Estudo das variáveis 44](#_Toc116652616)

[4.1.1. Critérios de apoio à tomada de decisão 44](#_Toc116652617)

[4.1.2. Descrição do Caso de Estudo 51](#_Toc116652618)

[4.1.3. Obtenção de dados 53](#_Toc116652619)

[4.1.4. Estudo do formulário 54](#_Toc116652620)

[4.1.5. Clustering 61](#_Toc116652621)

[4.2. Exploração de APIs 74](#_Toc116652622)

[4.2.1. Meteorologia e Eventos 74](#_Toc116652623)

[4.2.2. Localização 78](#_Toc116652624)

[4.3. Solução 82](#_Toc116652625)

[4.3.1. Arquitetura da Solução 82](#_Toc116652626)

[4.3.2. Processo de ETL 83](#_Toc116652627)

[4.3.3. Data Warehouse 91](#_Toc116652628)

[4.3.4. OLAP 93](#_Toc116652629)

[4.3.5. Dashboard Analítica 100](#_Toc116652630)

[5. Conclusão 113](#_Toc116652631)

[5.1. Considerações finais 113](#_Toc116652632)

[5.2. Trabalho Futuro 117](#_Toc116652633)

[5.3. Tabela de Riscos 117](#_Toc116652634)

[6. Apêndices 125](#_Toc116652635)

LISTA DE ABREVIATURAS

***APIs*** –Interface de Programação de Aplicativos

***APC*** – Contagem automática de passageiros

***BI*** – Inteligência empresarial

***CeNTI*** – Centro de Nanotecnologia e Materiais Técnicos, Funcionais e Inteligentes

***CFTV*** – Câmara de vigilância

***CLIP*** – Processo fotoquímico

***CMVNF*** –Câmara Municipal de Vila Nova de Famalicão

***DCBD*** – Descoberta de Conhecimento em Bases de Dados

***DLS*** – *Digital Light SynthesisTM*

***DS*** – Ciência de dados

***DSR***– *Design Science Research*

***DW*** – Armazém de dados

***ETL***– Extrair, transformar e carregar

***GPS SKM53*** – Sistema de Posicionamento Global SKM53

***IA*** – Inteligência Artificial

***IoT***– Internet das coisas

***KPI’S*** – Indicadores chaves de desempenho

***M2M***– Máquina com Máquina

***NoSQ*l** – Bancos de dados não relacionais

***RFID*** – Identificação por radiofrequência

***RSUs*** – Unidades de infraestrutura a beira da estrada

***SQL*** – Linguagem de Consulta Estruturada

***TPU & PLA*** – Tipos de plástico biodegradáveis

***VANETs*** – Redes *ad-hoc* veiculares

***V2X/ V2V/ V2I*** – Veículo para tudo; veículo para veículo, veículo para infraestrutura

***WSNS*** – Redes de sensores sem fio

LISTA DE FIGURAS

[Figura 1 - Evolução da Indústria. 6](file:///C:\Users\franciscabarros_iote\Documents\Work\Tese\Dissertação_FranciscaBarros_VersãoFinal2.docx#_Toc116652636)

[Figura 2 - Fluxograma das ações de entrada num parque de estacionamento. 29](file:///C:\Users\franciscabarros_iote\Documents\Work\Tese\Dissertação_FranciscaBarros_VersãoFinal2.docx#_Toc116652637)

[Figura 3 - Localização dos componentes eletrónicos nos autocarros. 30](file:///C:\Users\franciscabarros_iote\Documents\Work\Tese\Dissertação_FranciscaBarros_VersãoFinal2.docx#_Toc116652638)

[Figura 4 - Atividades da Metodologia DSR 33](file:///C:\Users\franciscabarros_iote\Documents\Work\Tese\Dissertação_FranciscaBarros_VersãoFinal2.docx#_Toc116652639)

[Figura 5 - Resultados Estatísticos das respostas do formulário. 55](file:///C:\Users\franciscabarros_iote\Documents\Work\Tese\Dissertação_FranciscaBarros_VersãoFinal2.docx#_Toc116652640)

[Figura 6 - Estatísticas (média, moda, mediana) das variáveis estudadas. 55](#_Toc116652641)

[Figura 7 - Dados Estatísticos relativamente à Idade. 56](file:///C:\Users\franciscabarros_iote\Documents\Work\Tese\Dissertação_FranciscaBarros_VersãoFinal2.docx#_Toc116652642)

[Figura 8 - Dados Estatísticos relativamente à Localização Geográfica. 56](file:///C:\Users\franciscabarros_iote\Documents\Work\Tese\Dissertação_FranciscaBarros_VersãoFinal2.docx#_Toc116652643)

[Figura 9 - Dados Estatísticos relativamente à Preferência de Transporte. 56](file:///C:\Users\franciscabarros_iote\Documents\Work\Tese\Dissertação_FranciscaBarros_VersãoFinal2.docx#_Toc116652644)

[Figura 10 - Nível de Correlação. 57](file:///C:\Users\franciscabarros_iote\Documents\Work\Tese\Dissertação_FranciscaBarros_VersãoFinal2.docx#_Toc116652645)

[Figura 11 - Resultados Estatísticos. 58](#_Toc116652646)

[Figura 12 - Resultados visuais do Clustering. 63](file:///C:\Users\franciscabarros_iote\Documents\Work\Tese\Dissertação_FranciscaBarros_VersãoFinal2.docx#_Toc116652647)

[Figura 13 - Arquitetura do projeto ioCity. 82](file:///C:\Users\franciscabarros_iote\Documents\Work\Tese\Dissertação_FranciscaBarros_VersãoFinal2.docx#_Toc116652648)

[Figura 14 - Arquitetura da Solução. 83](#_Toc116652649)

[Figura 15 - Data Warehouse. 92](file:///C:\Users\franciscabarros_iote\Documents\Work\Tese\Dissertação_FranciscaBarros_VersãoFinal2.docx#_Toc116652650)

[Figura 16 - Áreas com mais movimento anualmente. 94](file:///C:\Users\franciscabarros_iote\Documents\Work\Tese\Dissertação_FranciscaBarros_VersãoFinal2.docx#_Toc116652651)

[Figura 17 - Áreas com tráfego superior a dois mil movimentos por ano. 94](file:///C:\Users\franciscabarros_iote\Documents\Work\Tese\Dissertação_FranciscaBarros_VersãoFinal2.docx#_Toc116652652)

[Figura 18 - Parques mais movimentados mensalmente. 95](file:///C:\Users\franciscabarros_iote\Documents\Work\Tese\Dissertação_FranciscaBarros_VersãoFinal2.docx#_Toc116652653)

[Figura 19 - Parques com tráfego superior a mil e quinhentos movimentos por mês. 95](file:///C:\Users\franciscabarros_iote\Documents\Work\Tese\Dissertação_FranciscaBarros_VersãoFinal2.docx#_Toc116652654)

[Figura 20 - Dias mais movimentados. 96](file:///C:\Users\franciscabarros_iote\Documents\Work\Tese\Dissertação_FranciscaBarros_VersãoFinal2.docx#_Toc116652655)

[Figura 21 - Dias com tráfego superior a cento e cinquenta movimentos por dia. 96](file:///C:\Users\franciscabarros_iote\Documents\Work\Tese\Dissertação_FranciscaBarros_VersãoFinal2.docx#_Toc116652656)

[Figura 22 - Parques com mais capacidade. 97](file:///C:\Users\franciscabarros_iote\Documents\Work\Tese\Dissertação_FranciscaBarros_VersãoFinal2.docx#_Toc116652657)

[Figura 23 - Parques com capacidade superior a 300 lugares. 97](file:///C:\Users\franciscabarros_iote\Documents\Work\Tese\Dissertação_FranciscaBarros_VersãoFinal2.docx#_Toc116652658)

[Figura 24 - Taxa de ocupação anual. 98](file:///C:\Users\franciscabarros_iote\Documents\Work\Tese\Dissertação_FranciscaBarros_VersãoFinal2.docx#_Toc116652659)

[Figura 25 - Taxa de ocupação superior a 50%. 98](file:///C:\Users\franciscabarros_iote\Documents\Work\Tese\Dissertação_FranciscaBarros_VersãoFinal2.docx#_Toc116652660)

[Figura 26 - Horas com maior afluência. 98](file:///C:\Users\franciscabarros_iote\Documents\Work\Tese\Dissertação_FranciscaBarros_VersãoFinal2.docx#_Toc116652661)

[Figura 27 - Horas com tráfego superior a 200 movimentos por hora. 99](file:///C:\Users\franciscabarros_iote\Documents\Work\Tese\Dissertação_FranciscaBarros_VersãoFinal2.docx#_Toc116652662)

[Figura 28 - Passo 1 de utilização da dashboard. 103](#_Toc116652663)

[Figura 29 - Passo 2 de utilização da Dashboard. 104](#_Toc116652664)

[Figura 30 - Passo 3 de utilização da Dashboard. 104](#_Toc116652665)

[Figura 31 - Passo 4 de utilização da Dashboard. 105](#_Toc116652666)

[Figura 32 - Passo 5 de utilização da Dashboard. 105](#_Toc116652667)

[Figura 33 - Passo 6 de utilização da Dashboard. 106](#_Toc116652668)

[Figura 34 - Passo 7 de utilização da Dashboard. 106](#_Toc116652669)

[Figura 35 - Passo 8 de utilização da Dashboard. 107](#_Toc116652670)

[Figura 36 - Passo 9 de utilização da Dashboard. 108](#_Toc116652671)

[Figura 37 - Passo 10 de utilização da Dashboard. 108](#_Toc116652672)

[Figura 38 - Passo 11 de utilização da Dashboard. 109](#_Toc116652673)

[Figura 39 - Passo 12 de utilização da Dashboard. 109](#_Toc116652674)

[Figura 40 - Passo 13 de utilização da Dashboard. 110](#_Toc116652675)

[Figura 41 - Passo 14 de utilização da Dashboard 110](#_Toc116652676)

[Figura 42 - Diagrama de Gantt. 130](file:///C:\Users\franciscabarros_iote\Documents\Work\Tese\Dissertação_FranciscaBarros_VersãoFinal2.docx#_Toc116652677)

LISTA DE TABELAS

[Tabela 1 - Pilares da Indústria 4.0 8](#_Toc116652698)

[Tabela 2 - Projetos na Indústria 4.0 que impactam a decisão 14](#_Toc116652699)

[Tabela 3 - Fatores a considerar para melhorar a mobilidade urbana e Impactos 23](#_Toc116652700)

[Tabela 4 - Conceitos para desenvolver uma cidade digital 24](#_Toc116652701)

[Tabela 5 - Elementos de um novo modelo de estacionamentos 28](#_Toc116652702)

[Tabela 6 - Estratégias de pesquisa da metodologia Case Study 36](#_Toc116652703)

[Tabela 7 - Resumo de metodologias 40](#_Toc116652704)

[Tabela 8 - Mapeamento das Metodologias 40](#_Toc116652705)

[Tabela 9 - Lista de Ferramentas inerentes à execução do projeto 42](#_Toc116652706)

[Tabela 10 - Critérios relativos aos Transportes Públicos 46](#_Toc116652707)

[Tabela 11 - Critérios relativos aos Parques de Estacionamento 49](#_Toc116652708)

[Tabela 12 - Descrição do Caso de Estudo 51](#_Toc116652709)

[Tabela 13 - Grau de Importância dos critérios identificados 58](#_Toc116652710)

[Tabela 14 - Cenários Clustering 62](#_Toc116652711)

[Tabela 15 - Resultados do cenário A com o algoritmo K-Modes 64](#_Toc116652712)

[Tabela 16 - Resultados do cenário B com o algoritmo K-Modes 64](#_Toc116652713)

[Tabela 17 - Resultados do cenário C com o algoritmo K-Modes 64](#_Toc116652714)

[Tabela 18 - Perfis do Cenário A 66](#_Toc116652715)

[Tabela 19 - Perfis do Cenário B 68](#_Toc116652716)

[Tabela 20 - Perfis do Cenário C 70](#_Toc116652717)

[Tabela 21 - Exemplos de APIs de Metereologia 75](#_Toc116652718)

[Tabela 22 - Exemplos de APIs de Gestão de Eventos 77](#_Toc116652719)

[Tabela 23 - Exemplos de APIs do Google Maps 79](#_Toc116652720)

[Tabela 24 - Análise dos dados 84](#_Toc116652721)

[Tabela 25 - Estatísticas variáveis numéricas 87](#_Toc116652722)

[Tabela 26 - Estatísticas variáveis categóricas 88](#_Toc116652723)

[Tabela 27 - Indicadores e KPIs 89](#_Toc116652724)

[Tabela 28 - Cube.js vs Dashboard. 100](#_Toc116652725)

[Tabela 29 - Funcionalidades da solução 111](#_Toc116652726)

[Tabela 30 - Objetivos VS Resultados atingidos 115](#_Toc116652727)

[Tabela 31 - Tabela de Riscos 117](#_Toc116652728)

# Introdução

Neste capítulo é apresentada a contextualização do projeto e respetiva motivação para o desenvolvimento do mesmo. São identificados os objetivos, assim como, uma breve explicação de como o documento está organizado.

## Contextualização do Projeto e Motivação

Nos últimos anos as novas tecnologias inundaram e transformaram diversas áreas industriais, comerciais e sociais. Com o foco e os objetivos bem definidos estamos perante um aglomerar de oportunidades para alcançar o sucesso e aumentar a produtividade dos processos. Tudo isto é desencadeado com a quarta revolução industrial (Indústria 4.0) onde nasce uma iniciativa estratégica de alta tecnologia que alia uma estrutura onde os sistemas ciber-físicos através da Internet das coisas comunicam e fornecem informações em tempo real. Através da computação em nuvem estas chegam ao ser humano e providenciam informações que influenciam a tomada de decisões **(«Indústria 4.0», 2021).**

“O recente crescimento das tecnologias digitais está a permitir que as cidades sofram transformações para agilizar serviços inteligentes e oferecer novos produtos” **(Anthony Jnr, 2021)**. Nesta era, começou a ser adotado o conceito de cidades inteligentes com o propósito de melhorar a qualidade de vida dos cidadãos. Várias cidades começaram a usar sensores eletrônicos para recolher dados, trabalhá-los e usá-los para a tomada de decisão **(Anthony Jnr, 2021).**

Vários estudos na área da Indústria 4.0 foram realizados de forma a perceber as tendências e inovações ao nível da IOT e utilização da tecnologia que possam ser aplicados ao nível das *smart cities*. Deste modo, foi fundamental analisar boas práticas em outras áreas como têxtil e automóvel.

A preocupação de adotar “estratégias de redução de carros na cidade, e caminhar para alternativas de mobilidade”**(Aurigi et al., 2016)** com o objetivo de reduzir o congestionamento, o tempo, os acidentes e a poluição é um desafio comum a todas as grandes cidades europeias.

Este tema, cidades inteligentes, enquadra-se no projeto ioCity[[1]](#footnote-1) e é desenvolvido pela IOTECH (uma start-up que pretende resolver problemas sociais através do desenvolvimento de soluções inteligentes e inovadoras). Este projeto de investigação é apoiado pelos fundos públicos e visa ajudar a resolver este desafio, causado pelos transportes e pelo trânsito.

Deste modo, o ioCity propõe através da análise de um conjunto de dados (sensores, taxa de ocupação, tráfego, estacionamento, entre outros) ajudar o cliente a ter uma decisão mais informada e em tempo real a nível dos parques de estacionamento e transportes públicos. É esperado que, após a análise de diversos dados, seja modelado e desenvolvido um protótipo demonstrador de um sistema de apoio à tomada de decisão que possa ser útil para o cliente. Para isso será necessário analisar a taxa de ocupação de transportes públicos e a monitorização ativa de estacionamento de veículos dentro das cidades.

O interesse pela área de *Data Science* (DS) e pelas tecnologias envolventes é a principal motivação para a realização desta dissertação. “As cidades inteligentes são apoiadas pela digitalização e inovações tecnológicas que trazem benefícios ambientais, sociais e financeiros” **(Anthony Jnr, 2021)**, posto isto, a vontade de melhorar e inovar as cidades para que os cidadãos tenham uma melhor qualidade de vida é também uma motivação essencial. O projeto ioScience, desenvolvido pela ioTech,é também um contributo e motivação para a realização deste projeto onde o principal objetivo é a ciência de dados e inteligência empresarial como serviço.

Em suma, fazer a diferença e alcançar benefícios para as cidades através da tecnologia é o grande motor para a execução deste projeto.

## Objetivos

O apoio à decisão em tempo-real com base em dados gerados no momento é visto pelas organizações como um fator decisivo para o sucesso na tomada de uma decisão. No entanto, as organizações têm dificuldades em fazer uma análise desses dados em tempo real, devido à complexidade, quantidade e diversidade dos mesmos.

Com este projeto pretende-se ajudar na monitorização de locais de estacionamento, ou seja, aumentar a facilidade de encontrar um lugar de estacionamento disponível. Para isto, serão desenvolvidas *dashboards* e indicadores que podem ajudar à Indústria das Cidades.

No sentido de obter a resposta ideal ao problema, é apresentada a questão de investigação como objetivo principal: De que forma podemos melhorar a indústria das cidades através da utilização de sensores e Inteligência Artificial?

Para tal, é esperado a exploração de um conjunto de dados reais relativos à mobilidade urbana.

Deste modo, e para atingir o objetivo principal, desenvolver um sistema de apoio à tomada de decisão, foram definidos objetivos secundários que auxiliarão a dar resposta ao objetivo principal:

* Extrair, transformar e carregar os dados fornecidos (ETL);
* Criar uma *data warehouse* (DW);
* Interpretar o problema e perceção da realidade;
* Desenvolvimento dos *KPI’s;*
* Criar uma base de dados (SQL e noSQL);
* Recolher dados e interpretar os resultados obtidos (DCBD);
* Encontrar padrões entre os dados e retirar conclusões relevantes que sejam importantes para o apoio à tomada de decisão (*Data Science – DS).*

É previsto o desenvolvimento de um modelo descritivo-preditivo (inteligente) que, em tempo-real, contribua para uma decisão mais informada ao nível dos estacionamentos e/ou transporte.

Todo o trabalho decorreu com duas entidades parceiras da IOTech para o projeto ioCity, nomeadamente, o centro de nanotecnologia e materiais inteligentes (CeNTI) e câmara municipal de Vila Nova de Famalicão (CMVNF). A primeira organização referida, CeNTI, visa o desenvolvimento de sistemas de monitorização da taxa de ocupação de transportes públicos e sistemas de monitorização ativa de estacionamento de veículos dentro das cidades. O principal objetivo desta cooperação é a obtenção de dados reais, e, consequentemente, o desenvolvimento de um sistema de apoio à tomada de decisão com os dados recolhidos. Relativamente à CMVNF, o projeto ioCity, enquadra-se nos objetivos de desenvolvimento da Smart City de Vila Nova de Famalicão. Pelo que, o fornecimento de dados reais permitiu desenvolver uma solução baseada em dados verídicos e mais informada para os cidadãos.

## Estrutura do documento

O presente documento é constituído por 5 capítulos. Primeiramente o capítulo 1 (Introdução) onde é apresentada uma contextualização do projeto e a motivação para a realização do mesmo, assim como os objetivos, resultados esperados e a estrutura do documento.

Seguidamente, o capítulo 2 (Revisão da Literatura) onde é abordado a estratégia de pesquisa, o estado da arte e a visão crítica do estado da arte. No estado da arte são introduzidos conceitos chaves e essenciais para sustentar todo o trabalho desenvolvido na dissertação. Deste modo, esta secção esta dividida em subsecções onde temas como a Indústria 4.0 e Cidades Inteligentes são abordados de forma a explorar as diversas componentes da área.

Relativamente ao capítulo 3 (Abordagem metodológica, materiais e métodos) são exibidas as metodologias de investigação e práticas, assim como, a lista de ferramentas subjacentes à execução da dissertação.

No capítulo 4 (caso de estudo) é apresentado todo o trabalho realizado em três subsecções, nomeadamente, o estudo das variáveis a incorporar no modelo preditivo onde foram analisados diversos critérios de apoio à tomada de decisão, descrição detalhada do caso de estudo, forma de obtenção dos dados, estudo do formulário desenvolvido com os critérios que interferem com a decisão do método de transporte a escolher e o processo de *clustering* para identificar perfis e variáveis crucias na tomada de decisão.

Na subsecção Exploração de APIs de meteorologia, eventos e localização foram exploradas diversas soluções que assegurem a recolha dos dados a incorporar no modelo preditivo.

Por fim, é descrito detalhadamente todo o fluxo de desenvolvimento da solução. Inicialmente é apresentada a arquitetura da solução no âmbito do projeto, todo o processo de extração, transformação e carregamento dos dados para a criação de um *data warehouse* e o desenvolvimento de um modelo OLAP. O objetivo da presente dissertação é descrito expondo um caso prático de análise, ferramentas utilizadas para o desenvolvimento e funcionalidades inerentes à solução que permitem à mesma ser uma mais-valia no auxílio à tomada de decisão.

Por fim, o capítulo 5 (conclusão) onde é elaborado um remate final de todo o processo de escrita e desenvolvimento. Uma lista de riscos inerente ao desenvolvimento do projeto é apresentada, assim como, uma subsecção para o trabalho futuro.

# Revisão da Literatura

Neste capítulo são abordadas as estratégias de pesquisa para o desenvolvimento da revisão da literatura. Esta secção discute temas referentes à indústria 4.0 e cidades inteligentes são discutidos e, por último, é apresentada uma secção relativa à visão critica do estado da arte onde é exposta a opinião pessoal dos temas abordados no capítulo.

## Estratégia de Pesquisa

Este projeto de investigação tem por base a indústria das *smart cities*. Esta área tem vindo a ser bastante explorada e desenvolvida, pelo que, existem bastantes artigos que exploram várias questões sobre o tema.

Com o fim de identificar quais os artigos a selecionar foi realizada uma pesquisa geral, baseada em palavras-chaves pré-definidas, nomeadamente, “Indústria 4.0”, “*Smart Cities*”, “Impacto da Indústria 4.0 nas *Smart Cities*”, “Projetos inovadores na indústria” e “A indústria inteligente na mobilidade urbana – transportes públicos e parques de estacionamento”.

A estratégia de pesquisa baseou-se em processos manuais, onde através da visita a determinados sites foram encontradas informações relevantes para o tema. Assim como, através de plataformas auxiliares como *Google Scholar, Science Direct, Google, Web of Science, Scopus* e o repositório da Uminho onde uma variedade de artigos estão disponíveis. E ainda, através da lista de referências de determinado artigo, onde é possível encontrar referências de artigos relevantes para o desenrolar do projeto.

Consequentemente, foi elaborada uma seleção dos artigos através do nome, data de publicação, número de citações e do respetivo *abstract*, ou seja, foram eliminados artigos que não acrescentavam qualidade nem informação relevante, e, portanto, não foram considerados na revisão de literatura.

Numa fase de desenvolvimento e leitura, foram extraídas as informações essenciais para o desenvolvimento desta componente mais teórica. Deste modo, foi adotada uma sequência coerente na revisão de literatura, ou seja, tendo por base as palavras-chaves definidas inicialmente, foram definidas secções e subsecções com informações sintetizadas de todos os artigos que acrescentavam valor à dissertação.

## Evolução da Indústria

O setor da indústria nos últimos anos passou por várias tendências, evoluções e descobertas. Todo este crescimento levou a uma era de infinidade de melhorias tecnológicas que permitiu criar novas formas de suprimir as necessidades existentes e revolucionar de modo significativo as cadeias de valor **(Egreja, 2019)**.

Na Figura 1 é possível visualizar sucintamente as quatro revoluções indústrias e perceber os marcos significativos na indústria em cada era. De seguida é elaborado uma descrição das principais características de cada indústria.

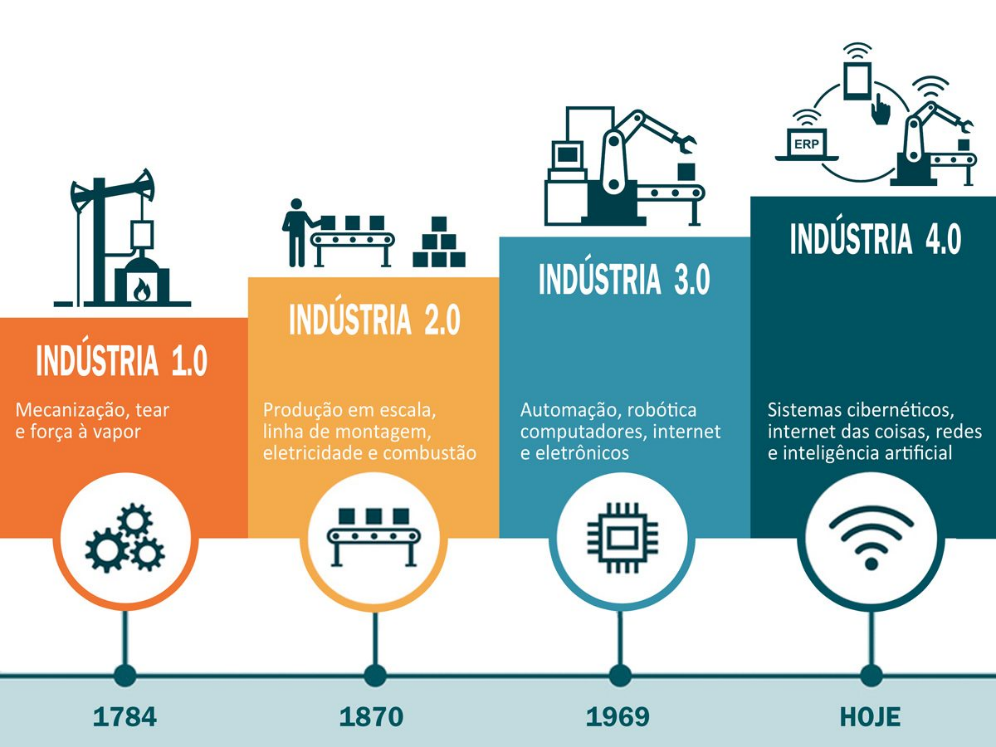


Figura 1 - Evolução da Indústria, retirado de (Egreja, 2019).

### 1º Revolução Industrial

Como podemos visualizar na imagem acima apresentada (Figura 1) a primeira revolução industrial foi gerada pela Revolução Comercial que ocorreu na Europa entre os séculos XV e meados do século XVIII onde o grande marco foi a passagem da manufatura para a maquinofatura. Por exemplo, a roca (instrumento manual de tecelagem) foi substituída pela máquina de fiar. Assim como, utilização das máquinas a vapor e a força hidráulica representaram uma grande revolução e um grande avanço na indústria do tecido, industrialização do ferro, navios a vapor, estradas de ferro, entre outras grandes mudanças que levaram a um aumento significativo da produção. Podemos afirmar que foi uma fase marcada pela transformação do ferro, carvão e energia a vapor em energia mecânica para o bom funcionamento das máquinas. Era marcada pelo surgimento de mecanização e da força hidráulica **(Fernandes, 2019; Neves et. al., 2022)**.

Nesta primeira revolução a manutenção era fundamentalmente corretiva, ou seja, após alguma falha a limpeza e a lubrificação dos equipamentos eram fulcrais para manter a produtividade **(Borlido, 2017)**.

### 2º Revolução Industrial

A segunda revolução industrial é marcada pela descoberta da energia elétrica e a produção em massa, em consequência da procura continua de otimização dos processos dos novos meios de produção e melhoria das técnicas. A substituição do ferro pelo aço teve um grande destaque, passando a ser utilizado nas ferrovias, na indústria naval e na fabricação de armamentos, por exemplo. A indústria química permitiu o desenvolvimento de remédios, fertilizantes, adubos, entre outros. A eletricidade passou a fazer parte do dia a dia da população assim como, permitiu criar e aprimorar diversos instrumentos de produção, uma vez que houve a substituição da energia a vapor pela energia elétrica.

É nesta fase que surge também a gasolina e o gás proveniente do uso do petróleo como fonte de energia (substituindo a utilização do carvão). Possibilitou o aprimoramento de técnicas, surgimentos de novas máquinas, introdução de novos meios de produção e automatização do trabalho que levaram a uma organização de produção industrial a menor custo e tempo. A segunda revolução industrial surge num período de pós-guerra onde se regista um grande aumento da procura o que, consequentemente, levou a um aumento da produtividade para conseguir corresponder á procura **(Fernandes, 2019; Neves et. al., 2022).** Deste modo, houve uma grande evolução no que diz respeito à manutenção. Começaram a ser registadas as manutenções feitas nos computadores e o agendamento e controlo da manutenção foi um dos passos mais significantes que começou a ser feito. Pelo que, passamos de uma manutenção corretiva para uma manutenção mais cautelosa e prudente, manutenção preventiva **(Borlido, 2017)**.

### 3º Revolução Industrial

Na década de 70 a terceira Revolução Industrial caracteriza-se pela computação e automação dos sistemas de informação. As máquinas começaram a ser programáveis o que originou uma progressiva automatização **(Neves et. al., 2022)**.

“Os principais marcos desse período podem ser vistos por meio dos aperfeiçoamentos e das inovações nas áreas de robótica, genética, telecomunicações, eletrônica, transporte e infraestrutura” **(Neves et. al., 2022)**.O sistema *Just In Time* nasce nesta época com a necessidade de reduzir ao máximo o tempo, consequentemente nasce o termo de manutenção preditiva **(Borlido, 2017)**.

### 4º Revolução Industrial

A Indústria 4.0 apresenta as mais recentes inovações tecnológicas relacionadas com a automação, controlo e tecnologias de informação, aplicadas aos processos de manufatura. Estas tecnologias emergiram com o foco da melhoria continua da eficiência e produtividade dos processos.

"A Quarta Revolução Industrial gera um mundo em que os sistemas de fabricação virtuais e físicos cooperam entre si de uma maneira flexível a nível global" **(Schwab Klaus, 2016).** Esta revolução trouxe inúmeras vantagens e é importante destacar que a obtenção de dados úteis, para o apoio à decisão, em tempo real revolucionou a indústria. Redução de custos e erros de produção, aumento da segurança e produções mais sustentáveis são alguns exemplos vantajosos da aplicação da indústria 4.0 **(Fernandes, 2019)**.

## Pilares da Indústria 4.0

Muitas indústrias estão a assistir à introdução das tecnologias que criam novas formas de suprimir as necessidades existentes e perturbam de modo significativo as cadeias de valor **(Fernandes, 2019)**. Deste modo, na Tabela 1 são expostos os nove pilares tecnológicos que emergiram com a indústria 4.0, nomeadamente, cibersegurança, realidade aumentada, robótica, simulações, digitalização 3d, sistemas integrados, computação em nuvem, internet das coisas (IoT) e big data.

Tabela 1 - Pilares da Indústria 4.0

| **Tecnologia** | **Descrição** |
| --- | --- |
| **Cibersegurança** | Com a quantidade de dados que emergem diariamente, é essencial que as empresas protejam os seus sistemas para prevenir possíveis vírus, ataques e crimes virtuais. Pelo que, é crucial a incorporação de sistemas de cibersegurança robustos para proteger os sistemas de informação **(Fernandes, 2019).** |
| **Realidade aumentada** | No artigo “Realidade aumentada na indústria: quais os ganhos e aplicações” **(Ferreira, 2020)** publicado a 17 de setembro de 2020, Ramon Ferreira afirma: “A realidade aumentada é uma tecnologia que realiza a integração do mundo virtual ao mundo real. Ela permite a sobreposição de objetos gerados por computador em um ambiente real, por meio de um dispositivo de visualização (smartphone, tablet ou óculos especiais). A partir disto, é possível interagir com estes elementos virtuais em diversas atividades”**.** |
| **Robótica** | Todas as tarefas de caracter repetitivo ficam à responsabilidade de robôs inteligentes, permitindo uma grande vantagem a nível de produção e tempo **(Iberdrola, 2022)**. |
| **Simulações** | Com a 4º revolução industrial, é exequível simular toda a cadeia de criação virtualmente e concludentemente obter dados para analise em tempo real. Com esta inovação, os processos podem ser testados e perceber até que ponto a mudança é exequível e garante uma otimização de recursos e uma melhor performance **(Altus, 2019)**. |
| **Digitalização 3d** | A digitalização 3d é o fenómeno que permite a criação de um modelo físico através de um modelo virtual **(Justus et al.,2015).** De forma mais eficiente e a um menor custo é possível ter um processo produtivo com menos falhas e mais qualidade. |
| **Sistemas Integrados** | A indústria 4.0 sugere uma melhor cooperação entre todos os elementos que façam parte da mesma organização. Com o intuito de garantir a eficiência organizacional propõe a integração de dados através num sistema ERP ou SAP (por exemplo) **(*Indústria 4.0*, 2021a).** |
| **Computação em nuvem (*Cloud computing*)** | A computação em nuvem disponibiliza serviços e armazenamento ilimitado, onde a informação e dos dados podem ser compartilhados entre diferentes sistemas. Através da internet “nuvem”, é possível aceder remotamente às informações desejadas **(«Computação em nuvem», 2021).** |
| **Internet das coisas (*Internet of Things*)** | O conceito de Internet das coisas (IoT) foi citado pela primeira vez por Kevin Ashton que descreveu um sistema “onde a Internet está conectada ao mundo físico através de sensores onipresentes”**(«Kevin Ashton», 2022).**  A IoT permite a interconexão digital de objetos físicos e ambientes, permitindo uma ampla rede capaz de agrupar e transmitir dados crucias para as organizações. Ou seja, com a conexão à internet é possível conectar um equipamento a outro (comunicação máquina com máquina – M2M). “Trata-se de uma ligação de todas as máquinas, dispositivos, sensores, câmaras e outros componentes que visam otimizar a vida das pessoas e empresas, além de melhorar as operações diárias e economizar tempo e dinheiro” **( Justus, et.al, 2015).** |
| **Big Data** | Com o surgimento da internet das coisas o volume de dados cresceu de forma exponencial. Deste modo, surge uma ferramenta capaz de tratar e analisar um grande conjunto de dados heterógenos. Volume, velocidade e variedade são três subdivisões de dados que o *Big Data* apresenta. É possível obter informações relevantes e crucias em tempo real que permitem às organizações criar indicadores e definir objetivos que possibilitam uma tomada de decisão mais assertiva **(«Indústria 4.0», 2021b).** |

## A Indústria das Cidades Inteligentes

Neste contexto, com a emersão de inúmeras tecnologias que surgiram com a quarta revolução industrial pode-se afirmar que estamos perante uma transformação digital. Este conceito é um processo que implica inovação e transformação através de tecnologias digitais para aumentar a produtividade e o valor de criação. Cada vez mais há uma tendência global de urbanização sustentada pelo digital, o desenvolvimento de cidades inteligentes é um movimento crescente que implementa tecnologia com o objetivo de melhorar a modernização das cidades criando benefícios a nível ambiental, social e financeiros **(Anthony Jnr, 2021).**

O conceito de *Smart City* surgiu na última década e pode ser considerado uma área urbana que integra e utiliza um vasto conjunto de métodos eletrónicos e sensores para captar dados. Com o objetivo de criar novas oportunidades que melhorem a qualidade de vida, este conceito “coordena e integra várias tecnologias que até então foram desenvolvidas separadamente umas das outras, mas têm sinergias claras em sua operação” **(Batty et al., 2012)**. Com vista à otimização da cidade são analisados uma diversidade de dados provenientes dos cidadãos, de edifícios (hospitais, escolas), dos transportes e mobilidade urbana, da utilização dos resíduos, entre outros dados, para que haja um crescimento e uma eficiência citadina com o fim de “integrar e agregar valor à prestação de serviços urbanos” **(Batty et al., 2012).**

### Das Cidades tradicionais às Sustentáveis e Inteligentes

As tecnologias como Inteligência artificial (IA), Internet das Coisas (IoT), computação em nuvem, big data, entre outras, aliadas aos processos quotidianos levaram ao surgimento de ecossistemas inteligentes. Permitem a capacidade de análise em tempo real e a obtenção rápida de dados, onde consequentemente, é possível processar e filtrar as informações relevantes e obter conhecimento essencial para a tomada de decisão.

Inicialmente, antes de se falar em cidades inteligentes, começaram a existir subsistemas inteligentes, por exemplo, infraestruturas inteligentes, transporte inteligente, ambiente inteligente, casas inteligentes e assim por diante. Focando nas casas inteligentes é possível perceber diversas dinâmicas que evoluíram nestes últimos tempos, salientado por exemplo, as máquinas de lavar, persianas automatizadas, fornos, entre outros.

Posto isto, todo o conjunto de subsistemas que foram emergindo originaram o ecossistema das cidades inteligentes.

“A ideia é usar tecnologias modernas para converter cada entidade de uma cidade convencional em um objeto autônomo que executa sua operação automaticamente sem qualquer ajuda externa substancial” **(Ahad et al., 2020).**

Sendo a sustentabilidade das cidades um dos objetivos da Organização das Nações Unidas (ONU) é fulcral que as cidades pra além de inteligentes se tornem sustentáveis **(Neto et al., 2016).**

Diversos fatores contribuem para que as cidades evoluam neste sentido, tal como, o uso de códigos reutilizáveis, matérias-primas ecológicas e recicláveis, consumos de energias e de recursos adequados e uma melhor gestão dos resíduos gerados pela computação.

“As cidades inteligentes devem integrar tecnologias e soluções de última geração para enfrentar os problemas de escassez de água limpa, esgotamento da qualidade do ar, diminuição das reservas de recursos naturais e desequilíbrio ecológico” **(Ahad et al., 2020).**

Segundo o artigo “ Descobrindo a história inicial do “Big Data” e da “Smart City” em Los Angeles” **(Vallianatos, 2015)** reflete sobre uma das ações visionarias primordiais na área das *smart cities* nos Estados Unidos na América em 1974. O *Community Analysis Bureau* utilizou bases de dados, análises de clusters e fotografias aéreas infravermelhas com o fim de produzir relatórios sobre a demografia e qualidade dos bairros para ajudar a evitar pragas e combater a pobreza.

Entre 1993 e 2000, a cidade europeia Amesterdão promoveu-se com o *De Digitale Stad* como“uma experiência inicial com redes cívicas e criação de comunidades virtuais” **(Aurigi et al., 2016)**. Inicialmente, houve um grande crescimento e adesão, no entanto o projeto não foi um sucesso, não tendo continuidade. Era um projeto mais focado na internet do que na própria cidade, levando a um desinteresse dos utilizadores **(Aurigi et al., 2016).**

Em 2009, Portugal desenvolveu uma Rede Portuguesa de cidades inteligentes, onde se propõe a analisar 25 municípios para adquirir uma maior eficiência energética. Sendo que em 2013, o conceito evoluiu e para além da análise do fator anteriormente mencionado, fatores como energia renováveis, gestão de água e resíduos foram incorporados com o intuito de retirar o melhor proveito dos mesmos **(*RPCI*, 2018).**

Em 2013, Bristol (cidade na Inglaterra) cria um evento inovador *Playable City.* Descrito como um jogo que visa promover o diálogo e o debate sobre a criação de experiências diferentes sob a cidade. O nível de participação da população foi tão alto que gerou uma quantidade de dados inesperada. Um ano depois com um investimento de 73 milhões, a cidade lança o *Bristol is Open* para desenvolver uma rede de alto desempenho de infraestruturas de mídia para o cento da cidade, que visa promover a cidade como inteligente **(Aurigi et al., 2016).**

Atualmente, em Barcelona, o evento *Smart City Expo World Congress* é um dos maiores na europa provendo a sustentabilidade aliada a tecnologia com o objetivo de alcançar maior eficiência e proveito da cidade, mostrando soluções inovadoras a nível mundial e impulsionando novos projetos nesta área. É um facto, que a evolução das *Smart Cities* está a evoluir exponencialmente e que são o futuro da sociedade, trazendo inúmeras vantagens competitivas **(*Expo World Congress*, 2022)**.

Em Portugal, o país tem manifestado a sua vontade de crescer nesta área, e apresenta o Portugal *Smart Cities Summit* onde empresas, universidades, start-ups, entre outras entidades, discutem e promovem o futuro das *smart cities.* Apostando na partilha de conhecimento entre o mundo académico e o científico com o objetivo comum de evoluir e fazer crescer o país de forma inteligente**(Passos, 2021)*.***

### Vantagens e Desvantagens das Cidades Inteligentes

As cidades inteligentes utilizam as recentes tecnologias para tornar a qualidade de vida e recursos mais eficientes apresentando inúmeras vantagens. No entanto, também existem fatores que podem não ser tão favoráveis **(PrimeStone, 2020).**

Nesta secção irá ser abordado, primeiramente as vantagens de uma Cidade Inteligente e seguidamente as desvantagens.

#### Vantagens

Segundo o artigo “Vantagens e Desvantagens das Cidades Inteligentes” **(PrimeStone, 2020)**, publicado no dia 2 de abril de 2020 no blog *PrimeStone* é possível enumerar alguns benefícios que melhoram o estilo de vida das populações, nomeadamente:

* A otimização dos serviços de transporte pois permite a localização de todos os automóveis e transportes públicos possibilitando uma gestão de tráfego mais eficiente **(PrimeStone, 2020)**;
* Um aumento da segurança que por sua vez leva a uma diminuição da criminalidade “Uma cidade inteligente é uma cidade mais segura” **(PrimeStone, 2020)**. Existem tecnologias que ajudam na segurança das cidades, como, reconhecimento facial, detetores de tiro, centros de crime conectados, sistemas de câmaras de vigilância, entre outros **(PrimeStone, 2020)**;
* A diminuição de desperdícios e de recursos naturais é outra das vantagens que o avanço tecnológico tem permitido, ou seja, efetuar uma gestão mais eficiente e controlada levando a uma redução de desperdícios como água, luz, entre outros **(PrimeStone, 2020)**;
* A diminuição da pegada ambiental é também possível com a implementação de *Smart Cities*, devido ao surgimento de ferramentas energéticas que ao utilizar fontes de energia renovável melhoram a qualidade do ar **(PrimeStone, 2020)**;
* Possibilidade de pontos específicos da cidade terem acesso a serviços de internet, permitindo a todos que nela circulam oportunidades de navegação **(PrimeStone, 2020);**
* Melhoria das infraestruturas, uma vez que, com as tecnologias que estão a emergir diariamente é possível efetuar uma análise das mesmas e identificar possíveis falhas **(PrimeStone, 2020)**;
* E, por fim, oportunidades de desenvolvimento económico pois uma cidade com estas características permitirá a tomada de decisões mais informadas e melhoria dos negócios, assim como, a atração de novos moradores e turistas **(PrimeStone, 2020)**.

#### Desvantagens

Contudo, e apesar das vantagens expostas na subsecção anterior (12) é possível enumerar também algumas desvantagens na adoção e implementação de Cidades Inteligentes, tais como:

* A privacidade pessoal torna-se bastante restrita pois com a adoção de todos os sistemas inteligentes, câmaras de segurança e técnicas de reconhecimento facial torna-se “mais difícil manter o anonimato” **(PrimeStone, 2020)**;
* Também é importante referir, que ao implementarmos uma *Smart City,* a cidade passa a ser profundamente tecnológica e, consequentemente, perdem autonomia na tomada de decisões caso seja exposta a falhas tecnológicas **(PrimeStone, 2020)**;
* A quantidade de dados e informações que as empresas, governos, entre outras entidades, tem acesso torna-se mais fácil proceder ao controlo e/ou manipulação da opinião pública **(PrimeStone, 2020)**;
* Por fim, mudança é sinonimo de resistência, deste modo, é importante que os cidadãos sejam informados das tecnologias em uso, para que as mesmas não sejam irrelevantes nas suas vidas diárias e não haja uma oposição da sua utilização **(Prasanna, 2022)**.

## Projetos na Indústria 4.0 que impactam a decisão

“A indústria 4.0 faculta transparência de ponta a ponta em tempo real, permitindo a verificação precoce das decisões de design na esfera da engenharia e as respostas mais flexíveis a perturbações no processo de produção” **(Kagermann, et.al, 2013)**.

Pelo que, permite a criação de maior valor através da incorporação de novas tecnologias e serviços. Deste modo, existem inúmeras áreas que criaram oportunidades de valor baseando-se nesta indústria **(Fernandes, 2019).**

Recolher informações e ideias das áreas que apresentam um maior crescimento e maior impacto na indústria 4.0 é uma mais-valia para as aplicar depois na área das cidades. Posto isto, na Tabela 2 são apresentadas uma síntese das três áreas abordadas, nomeadamente: área têxtil, área automóvel e cidades inteligentes com as respetivas tecnologias e casos de sucesso.

Tabela 2 - Projetos na Indústria 4.0 que impactam a decisão

|  |  |  |
| --- | --- | --- |
| **Área Têxtil** | | |
| **Tecnologia** | **Caso de Sucesso** | **Localização/ Empresa** |
| **IA**  **B**I  (Inteligência empresarial) | Capacidade de detetar erros de fabrico de malhas circulares na fase inicial de tricotagem. | Portugal (Porto) Smartex  **(*Smartex*, 2022)** |
| **DLS**  (*Digital Light SynthesisTM*)  **Impressão 3D** | Criação de uma sola intermediária para sapatilhas com rigor, precisão e adaptabilidade para o corredor. | Estados Unidos da América e Alemanha  *Adidas Running*  **(*4DFWD*, 2021)** |
| **Impressão 3D** | Criação de novas roupas através de um modelo digital. | Espanha (Bilbau)  Comme des Machines  ***(Aran et. al, 2020)*** |
| **Área Automóvel** | | |
| **Tecnologia** | **Caso de Sucesso** | **Localização/ Empresa** |
| **GPSSKM53**  (Sistema de Posicionamento Global)  **IoT** | Navegação em tempo real | ---------  **(Rahim et al., 2021)** |
| **Tecnologia** | **Caso de Sucesso** | **Localização/ Empresa** |
| **IoT**  **RSUs**  (unidades de infraestrutura a beira da estrada)  **V2V**  (veículo para veículo)  **V2I**  (veículo para infraestrutura)  **V2X**  (Veículo para tudo)  **VANETs**  (Redes ad-hoc veiculares) | Sistema de gestão de trânsito | ---------  **(Rahim et al., 2021)** |
| **IoT**  **5G** | Sistema de segurança e anti-roubo |
| **RFID**  (Identificação por radiofrequência)  **IoT** | Portagens com pagamento automático |
| **Cidades Inteligentes** | | |
| **Tecnologia** | **Caso de Sucesso** | **Localização/ Empresa** |
| **Sensores**  **IoT** | Otimizar as linhas dos transportes metropolitanos, monitorização dos estacionamentos e otimização da Mobilidade Urbana | Espanha (Barcelona)  **(Bria, 2018)** |
| **Big Data**  **IoT**  **CFTV inteligente**  **IA** | Owl Bus, monitorização dos estacionamentos, TOPIS (Operação de Transporte e Serviço de Informação), postes de iluminação LED e sistemas inteligentes de processamento de lixo. | Coreia do Sul (Seul)  **(Stevens, 2020)** |
| **5G**  **Fibra**  **Sensores** | Implementação de infraestruturas, otimização de transportes públicos, | Portugal (Aveiro)  **(Lab, 2021)** |
| **Sensores** | Recolher, processar, cruzar e analisar toda a informação do dia a dia dos famalicenses, para que, a gestão territorial a nível da mobilidade e gestão energética seja mais eficiente. | Portugal (Vila Nova de Famalicão)  **(CMVNF, 2019)** |

### Área Têxtil

A indústria têxtil e de vestuário enfrenta desafios como eficiência, sustentabilidade e resposta rápida aos requisitos dinâmicos dos clientes, bem como qualidade do produto e conformidade. A geração em que vivemos esta cada vez mais consciente da importância da sustentabilidade, pelo que, o ecossistema industrial e as tendências globais de negócios sustentáveis pressionam as indústrias a se tornarem mais sustentáveis, inovadoras e ágeis com o objetivo de conseguirem corresponder às expectativas dos consumidores. “Nada vai mudar enquanto mantiverem o seu sistema produtivo em modo linear (extrair, fabricar, descartar)” **(Marconi & Broega, 2021).**

Para conseguir alcançar estes objetivos a indústria têxtil cada vez mais se baseia em soluções tecnológicas da indústria 4.0. Mas a grande questão que se coloca é como é que as empresas de área têxtil podem melhorar os processos de criação de valor como o objetivo de manter a sustentabilidade através da integração de um sistema de negócio inteligente? Será que há barreiras e limitações associadas? Será linear afirmar que aplicação de um sistema deste género facilite de imediato o processo? Ou haverá barreiras implícitas que poderão não facilitar na execução? A necessidade de uma solução tecnológica deste nível deve-se à redução do tempo de entrega dos pedidos, ao aumento dos custos de matéria-prima (mão de obra e materiais), défice/excesso de stock e a vontade de reforçar a exportação e desenvolver uma solução mais sustentável **(Ahmad et al., 2020).** É importante também referir que a área têxtil é uma das indústrias mais poluentes do mundo, segundo o artigo “O impacto da produção e dos resíduos têxteis no ambiente”**(*Impacto da produção*, 2020)** com a última alteração feita em 12 de fevereiro de 2021, publicado no site do Parlamento Europeu, são apresentados vários fatores:

* Consumos de água muito elevados (devido à necessidade de terrenos para o cultivo de algodão e fibras);
* É responsável por 10% das emissões de gases com efeito de estufa;
* Poluição das águas (os produtos de tingimento e acabamento são responsáveis por 20% da poluição de água potável a nível mundial, assim como microfibras e mioplásticos provenientes da lavagem de matérias sintéticos);
* Elevados resíduos têxteis em aterros.

Segundo o artigo “Pode a indústria têxtil ser inovadora e sustentável? We Sustain diz que sim” **(ECO, 2020)**, publicado no dia 17 de julho de 2020 no blog Capital Verde a pandemia covid-19 veio acelerar todo este processo. “O consumo têxtil está em forte quebra, as fábricas estiveram vários meses fechadas, os níveis de exportações caíram brutalmente” **(Tavares, et.al, 2020).** Assim sendo, são apresentados vários casos de sucesso, que com a revolução industrial desenvolveram soluções tecnológicas capazes de contribuir e revolucionar a área têxtil.

Casos de Sucesso

* Portugal 🡪 Smartex (Porto)

A smartex, vencedora do prémio *Pitch* da *Web Summit* 2021, é uma start-up que promove uma solução capaz de detetar erros de fabrico de malhas circulares na fase de tricotagem. “Mais de 80% da origem dos defeitos têxteis está na fase de Malharia” **(ECO, 2021; *Smartex*, 2022).**

Na maior parte das vezes, estes defeitos são encontrados na fase final de produção quando a peça esta desenvolvida e pronta para ir para o mercado. Deste modo, e através da inteligência artificial foi desenvolvido um sistema de deteção de defeitos que pode ser facilmente conectado a teares circulares utilizando uma câmara industrial e um sistema de iluminação multiespectral que tem como objetivo filmar a malha e identificar o erro. Consequentemente, são recolhidas as imagens da malha, em tempo real, que posteriormente são examinadas por algoritmos através de *computer vision* num modelo de *machine learning*. A Smartex deteta e evita a produção com defeito **(*Smartex*, 2022).** Este software de monitorização de produção, adequado para análise de *Business Intelligence* permite que no momento exato em que é detetado um erro a produção pare. “Como os defeitos são detetados logo no início do processo, não há sequer necessidade de gastar matéria-prima, água (50.432 litros economizados), energia (gás natural para gerar calor e eletricidade, que emitem 684 kg de CO2 por cada rolo têxtil) e tempo no acabamento dos materiais têxteis” (**ECO, 2020);**

* Estados Unidos da América e Alemanha 🡪 Adidas *Running*

A adidas 4DFWD pretende proporcionar aos atletas uma experiência única com o desenvolvimento de uma sola idealizada onde o conforto, qualidade e adaptabilidade são os pontos chaves. Esta sola apresenta três vezes mais movimento que modelos anteriores e reduz em 15% a força que o corredor tem de fazer. “Proporcionando uma sensação de deslizamento única para nossos corredores” **(Sam, et.al, 2021).**

Para o desenvolvimento desta sola foi utilizada uma tecnologia que a marca já utiliza há quatro anos e com a qual já desenvolveu várias sapatilhas de sucesso, 4D *lattice midsole technology* em parceria com a Carbon. Deste modo, combinando os dados do atleta, tecnologia *Digital Light SynthesisTM (DLS) da Carbon* e impressão 3D conseguem produzir a sola intermediária da sapatilha com rigor e precisão, criando benefícios para os corredores **(*4DFWD*, 2021).**

Segundo o artigo “*Our Tecnology*”, **(«DLS 3D», 2022)**, a tecnologia DLS usa a projeção digital de luz, ótica permeável ao oxigênio e resinas líquidas programáveis ​​para fabricar produtos poliméricos finais duráveis ​​e de alto desempenho. Pelo que, através de um processo fotoquímico (CLIP) que tem a capacidade de, com luz ultravioleta (UV), transformar a resina plástica em partes sólidas, faz com que à medida que esta luz é projetada a peça se solidifique e a plataforma de construção suba. É importante referir que tal fenómeno acontece uma vez que é usada uma janela permeável ao oxigênio num reservatório de resina tratado por UV.

Em todo este processo existe uma zona denominada de “morta” característica por apresentar uma interface fina e líquida de resina não tratada entre a janela e a peça de impressão.  “A luz passa pela zona morta, curando a resina acima dela para formar uma parte sólida sem solidificar a parte na janela. (…) Os processos tradicionais de impressão 3D baseados em resina produzem peças fracas e quebradiças. O carbono supera isso incorporando uma segunda química programável ativada por calor em nossos materiais. Depois que uma peça é impressa em uma impressora de carbono, ela é cozida no forno. O calor desencadeia uma reação química secundária que faz com que os materiais se adaptem e se fortaleçam, assumindo características excecionalmente fortes. Isso produz peças de alta resolução com propriedades mecânicas de grau de engenharia”**(«DLS 3D», 2022).**

Em suma, com a utilização destas tecnologias a adidas consegue conceber uma sola de grande durabilidade, resistência e adaptabilidade para as 4DFWD;

* Espanha 🡪 Comme des machines (Bilbau)

Comme des machines é uma fábrica no setor têxtil, localizada em Bicais, Bilbau que utiliza uma das técnicas mais inovadoras da moda: Digitalização 3d.

A digitalização 3d permite a criação de objetos físicos a partir de um modelo digital. Esta tecnologia tem vindo a ser aperfeiçoada e trabalhada uma vez que contribui bastante para um mundo mais sustentável (desperdiçando menos materiais e produtos químicos), acelerando os processos de criação e inovação, minimizando os custos e possíveis falhas.

Segundo o artigo “*Comme des Machines – Future Perfect*” **(Aran et. al, 2020)**, publicado no dia 2 de junho de 2020 no blog Talento a Bordo, um dos criadores da fábrica, Aran Azkarate (2020), afirma que as marcas parceiras aumentaram até 30% de margem de lucro e que a digitalização 3d está a revolucionar a indústria da produção. Produzem roupas sustentáveis e customizáveis para marcas impactantes e bastante conhecidas em todo o mundo, como, Nike e Burberry **(Aran et. al, 2020).**

É também possível perceber a evolução que a digitalização 3d tem vindo a ter desde 2016 “Quando comecei, havia apenas amarelo, azul, vermelho e verde disponíveis, e fazer algo interessante com essas cores foi um verdadeiro desafio. Agora podemos escolher, por exemplo, entre muitos tipos diferentes de madeiras com tons lindos e sutis. Podemos até misturá-los, o que antes não era possível” **(Aran et. al, 2020).**

De entre vários materiais que a impressão 3d pode utilizar (como plástico TPU, PLA, silicone, entre outros) a *Comme des Machinnes* baseia-se em materiais sustentáveis como o PLA“um bio polímero feito com ingredientes naturais como cana-de-açúcar, café e batata, entre outros. É biodegradável e compostável” **(Aran et. al, 2020).**

A produção deste plástico utiliza bactérias para a produção de ácido lático por meio da fermentação de vegetais ricos em amido, utilizando fontes renováveis.Este plástico substitui o plástico convencional e é adaptável a várias indústrias como indústria alimentar e médica. Deste modo, no sistema, são criados os modelos com as medidas desejáveis. Através da projeção deste modelo tridimensional num software específico, os códigos são enviados para a impressora que produz o material desejado, a partir da injeção do plástico PLA por camadas. O PLA em camadas pode ser injetado por meio da mistura a vácuo /fusão /laser e/ou moldagem por injeção **(«Tecnologia 3D», 2020; Yu, 2018).**

Podemos afirmar, que os recursos humanos na área têxtil é uma dependência, e com a quarta revolução industrial através da robótica, BI, digitalização 3D, entre outras tecnologias inovadoras este fenómeno começou a ser contrariado. Começou a tornar-se mais sustentável, económico e prático. Adaptando-se à evolução da sociedade e à constante preocupação de preservar o planeta em que vivemos **(Ahmad et al., 2020).**

São apresentadas no site IAPMEI - Agendas Mobilizadoras para a Inovação Empresarial | Propostas para a Economia do Futuro, onde é possível constatar, através das *livestream*, que diversas start-ups em Portugal tencionam revolucionar a indústria têxtil portuguesa (não só em peças de roupa, mas também no calçado) assim como a criação de fábricas do futuro apostando essencialmente na tecnologia de digitalização 3d. Apesar das limitações subagentes à utilização destas tecnologias (custo, complexidade, competências técnicas, integração com novos sistemas, gestão de mudanças) a sustentabilidade gere a economia. Para tal, é necessário apostar em novas formas de confecionar os têxtis **(IAPMEI, 2021)**.

### Área Automóvel

A indústria automóvel é uma das mais diversificadas e influentes. Enfrenta desafios sustentáveis, mas também desafios para garantir a mobilidade urbana. É um facto, que o surgimento dos veículos é um marco fulcral para a mobilidade e uns dos grandes marcos na indústria.

Um dos primeiros carros foi idealizado por Nicolas Joseph Cugnot, no ano de 1700, com um peso quase de 10 toneladas e movido por uma espécie de cadeia a vapor, sendo o seu principal objetivo o transporte dos equipamentos de guerra. É de facto, uma evolução notória, e hoje em dia o automóvel é rapidamente associado a carros elétricos, inteligentes e híbridos **(Neon, sem data).**

Uma das maiores comodidades que o ser humano tem ao seu dispor é a locomoção, deste modo, desde que surgiu o primeiro automóvel este setor evolui no sentido de qualidade, conforto, inovação e sustentabilidade (mais precisamente a qualidade do ar que é um fator afetado diariamente com a emissão de gases e poeiras provenientes dos motores), regidos com o objetivo, de alcançar a eficiência.

Casos de Sucesso

Segundo o artigo “*Evolution of IoT-enabled connectivity and applications in automotive industry: A review*” **(Rahim et al., 2021)** publicado a 25 de agosto de 2021 a inteligência dos carros deve-se à IoT que permite a ligação automóvel a terminais do condutor (nomeadamente o ecrã do telemóvel). Para isto são usados sensores e softwares inteligentes. As redes de sensores sem fio (WSNS) são redes de vários nós que atuam como sensores que cooperam com outros nós e tem a capacidade de receber informações físicas e transferi-las para outros dispositivos.

Assim sendo, os condutores podem obter facilmente informações sobre o calor do motor, luz, calor do radiador, pressão dos pneus, movimento do veículo, km percorridos, entre outras. O mesmo artigo refere um conjunto de aplicações da tenologia IoT na indústria automóvel, como:

* Navegação em tempo real

“O IoT tem um grande impacto na garantia de uma instalação segura e inteligente. Tem características desejáveis para ajudar o condutor, seguindo mapas ao vivo (…) em qualquer lugar. Por outro lado, um modelo ativado por IoT utilizando o módulo GPS SKM53 e a fórmula Haversine para notificar a equipa de salvamento (…) como um acidente” **(Rahim et al., 2021)**.

Ou seja, o módulo GPS SKM53 obtém as coordenadas (latitude, longitude, hora, validade da posição e altitude) e a fórmula haversine providencia a distância em que se encontra de determinado ponto, obtendo a localização exata;

* Sistema de gestão de trânsito

“Como a quantidade de veículos na estrada está a crescer gradualmente, é possível gerir a temporização inteligente do semáforo com base na densidade de veículos com a ajuda da IoT, RSUs instalados localmente e uma câmara de vigilância. No caso de emergência, o sistema central pode ajudar a detetar a zona do acidente e informa a equipa de salvamento mais próxima. (…) Isto também ajuda o condutor a conhecer as condições da estrada (escorregadio, molhado e neve), situação do trânsito e informações acidentais na estrada” **(Rahim et al., 2021)**.

Isto é possível devido há existência de inúmeras redes de comunicação entre veículos para garantir a segurança rodoviária e eficiência no trânsito. Podemos destacar o V2V (partilha de informações de veículo para veículo), V2I (partilha de informação entre veículos e as unidades de infraestrutura à beira de estrada – RSUs - por acesso à Internet). O V2X (partilha de informação entre veículo e qualquer entidade que afete ou seja afetada pelo veículo) incorpora tanto o V2V e V2I, esta comunicação é possível através da IoT. Para além disto existem as VANETs, redes ad-hoc móveis veiculares, que permitem adaptar-se à mudança de posição do veículo e transformar cada veículo num nó (ponto final de comunicação) individual sem fio, com o fim de estabelecer uma rede temporária de grande alcance entre todos os nós de veículos, mantendo a segurança do tráfego automóvel **(Rahim et al., 2021)**;

* Sistema de segurança e antiroubo

“Beneficiando-se de um sistema auxiliado por IoT sistema antifurto, o proprietário de um veículo roubado pode rastrear rapidamente a sua localização exata com a ajuda do smartphone com acesso a internet. Essa tecnologia garante mais confiabilidade devido aos recursos poderem ser controlados remotamente pelo proprietário” **(Rahim et al., 2021).** Uma tecnologia de conectividade que pode tornar isto possível, seguro e mais eficiente é o 5G. Esta rede móvel apresenta uma largura de banda maior e, portanto, uma maior velocidade de transferência de dados (bits/s) o que é uma característica cada vez mais importante na indústria automóvel, assim como, conexões mais rápidas, com mais capacidade e menos consumo de energia comparativamente à 4G **(Rahim et al., 2021)**;

* Portagens com pagamento automático

A tecnologia RFID e a IoT permitem uma melhor gestão do trânsito, segurança e pagamento de portagens de forma fácil e eficaz. Basicamente, o sistema, no momento de passagem do veículo recolhe as informações físicas do mesmo antes de cobrar a taxa de pagamento através do acesso a um banco de dados **(Rahim et al., 2021).**

Pelo que, a tecnologia RFID é descrita “como uma tecnologia para identificar o objeto automaticamente e recolher informações com a ajuda de um leitor usando sinais eletromagnéticos de um adesivo chamado transcetor microchip” **(Rahim et al., 2021)** permitindo que os carros não tenham a necessidade de parar para efetuar um pagamento. Para além disto, esta tecnologia melhora “a utilização eficiente de matérias-primas, equipamentos e ferramentas” **(Rahim et al., 2021)** na indústria de fabricação de automóveis.

Estes são apenas alguns exemplos atuais que a indústria automóvel está a adotar. A inovação da tecnologia é a chave para uma economia competitiva, conduzindo a importantes ganhos de produtividade, permitindo que as empresas se direcionem para atividades de maior valor acrescentado. A IoT apresenta várias aplicações em diversos fatores, tais como, saúde, construção, energia e rede inteligente, monitoramento ambiental e cidades inteligentes **(Rahim et al., 2021).**

### Cidades Inteligentes

Os projetos de cidades inteligentes*,* nos dias de hoje, são na sua maioria projetos relacionados com a mobilidade, meio ambiente, governo, economia, pessoas e estilo de vida.

Segundo o artigo “Mobilidade Inteligente e Ambiente Inteligente nas cidades espanholas” **(Aletà et al., 2017**) publicado em 2017 afirma que a crescente concentração de pessoas, empresas e organizações nas cidades tem promovido a criatividade, inovação, diversidade e crescimento económico, para tal, as cidades devem apostar num processo de transformação que dê resposta a diferentes desafios que emergiram com o aumento da urbanização. É possível enumerar alguns, como o esgotamento dos recursos naturais, sustentabilidade ambiental, mobilidade, entre outros **(Aletà et al., 2017**).

Quando se fala em mobilidade há de imediato uma correlação com o trânsito e meios de transportes. Este estudo baseou-se em 62 cidade de Espanha e avaliou fatores de mobilidade, fatores ambientais e indicadores territoriais (Tabela 3).

Tabela 3 - Fatores a considerar para melhorar a mobilidade urbana e Impactos

| **Fatores a considerar** | **Impacto** |
| --- | --- |
| - Alternativas ao automóvel individual  - Aumento da eficiência das viagens através de uma gestão mais eficiente das ligações entre os diferentes tipos de transporte  - Gestão de controle inteligente para reduzir o congestionamento do tráfego  - Gestão da eficiência energética e hídrica | - Diminuição de custos econômicos  - Aumento da qualidade ambiental  - Redução de tempo |

O artigo referenciado “Mobilidade Inteligente e Ambiente Inteligente nas cidades espanholas” **(Aletà et al., 2017**) não apresenta soluções aos problemas, mas sim fatores que devem ser tomados em consideração para as cidades se inovem e se desenvolvam. Pretende também, informar os cidadãos dos avanços tecnológicos nos processos de urbanização.

Casos de Sucesso

* Barcelona, Espanha

Neste contexto, surge o caso de estudo da segunda maior cidade espanhola: Barcelona. Barcelona está se a tornar um modelo de sucesso de cidades inteligentes a nível global. Segundo o artigo “Caso de estudo de uma cidade inteligente – Barcelona” **(Wang, 2019)** publicado a 4 de agosto de 2019 demonstra como esta cidade tem vindo utilizar tecnologias como sensores e IoT nas áreas de ambiente urbano e mobilidade com o fim de desenvolver uma cidade digital. Para isto são descritos três conceitos fundamentais que são apresentados na Tabela 4, nomeadamente, transformação digital, inovação digital e capacitação digital.

Tabela 4 - Conceitos para desenvolver uma cidade digital

| **Conceito** | **Descrição** | **Impacto** |
| --- | --- | --- |
| **Transformação Digital** | Este conceito propõe uma transformação através das tecnologias emergentes, com o objetivo de aumentar o valor de criação. Mais detalhes podem ser consultados na secção anterior 2.4. | Segundo a “Barcelona Cidade Digital 2017-2020” o grande objetivo é ter uma enorme quantidade de dados para analisar. Deste modo, para alcançar o principal objetivo é necessário que a população tenha acesso a diferentes tipos de tecnologias digitais para que sejam implementados sensores e redes de fibra de alta velocidade. Assim "os dados tornam-se inteligentes e ágeis” **(Wang, 2019).** |
| **Inovação Digital** | Inovação Digital refere-se à capacidade que, neste caso, as cidades têm de se inovar (a nível de serviços e produtos baseados em tecnologia) | É expectável que surjam novos tipos de empreendedorismo baseados nas tecnologias. |
| **Capacitação Digital** | Envolver a tecnologia no dia-a-dia. | Com base no que foi dito anteriormente é esperado que estas tecnologias sejam inseridas no quotidiano da população. |

Através da leitura do artigo acima mencionado, é possível constatar diversas iniciativas que a cidade está a adotar. No decurso de otimizar as linhas dos transportes metropolitanos foram analisados uma série de dados que melhoram as rotas dos transportes públicos **(Wang, 2019)**.

Com o fim de reduzir a poluição foram adotados autocarros híbridos e as garagens onde os mesmo são guardados são mantidos por energia solar. Em diversas paragens há um ecrã onde é possível visualizar o tempo de espera para a chegada de um determinado autocarro **(Wang, 2019)**. A nível do estacionamento foram colocados “sensores eletromagnéticos sob o solo para monitorar se o estacionamento está ocupado e se um veículo está estacionado ilegalmente” **(Wang, 2019).**

Segundo o site “Urban Hub – Pessoas Dando Forma às cidades” no artigo Smart City 3.0 – pergunte a Barcelona sobre a próxima geração de cidades inteligentes **(Bria, 2018)** é apresentada uma iniciativa de mobilidade para a linha 9 do sistema de metro da cidade, onde foram desenvolvidos elevadores inteligentes que se adaptam as necessidades dos passageiros através da análise de dados em tempo real. “O movimento dos elevadores está otimizado para o uso dos passageiros e eles movem-se automaticamente para o andar da plataforma logo antes da chegada do metro. Isso agiliza a mobilidade dos passageiros, reduz as multidões e diminui o consumo de energia – para cerca de 30 milhões de passageiros por ano. Isso que é ser inteligente”**(Bria, 2018).**

Segundo o artigo “Infraestruturas de Mobilidade nas Cidades e Mudanças Climáticas: Uma análise através dos Super blocos de Barcelona” **(López et al., 2020)** publicado em 2020 Barcelona apostou num método de super blocos para mitigar problemas como o trânsito, emissões de gases e melhorar a saúde e qualidade de vida dos cidadãos. Estes super blocos caracterizam-se por “uma rede de blocos e estradas básicas formando um polígono, de aproximadamente 400m x 400m (…) representa um novo modelo de mobilidade que reestrutura a típica rede viária urbana” **(López et al., 2020).** Este conceito não é novo e muitos países já o adotaram. No entanto, no caso de Barcelona estes caracterizam-se por dois componentes: interiores e exteriores, sendo que cada um apresenta em médio 5000-6000 habitantes. Como referido no artigo, o interior é redirecionado para veículos motorizados e estacionamento (apenas acima do solo). Em casos excecionais, o interior pode ser usado para tráfico residencial, assim como serviços de emergência e para cargas/descargas.

Existem obstáculos inerentes, como por exemplo ”a necessidade de redesenhar a rede de transportes coletivos para que o tráfico de automóveis seja realmente reduzido nas cidades, a possível influência negativa no trânsito de entrada e saída da cidade, a falta de vantagens visíveis caso não sejam implantados em toda a cidade, o risco de gentrificação nas áreas com super blocos, oposição pública e oposição de alguns setores empresariais” **(López et al., 2020)**;

* Seoul, Coreia do Sul

“Seul: uma cidade inteligente de classe mundial” **(Stevens, 2020)**, Seul baseia-se nas novas tecnologias da Quarta Revolução Industrial e em Big Data. *Owl Bus* é um autocarro noturno que recolhe dados e consequentemente define os circuitos a realizar. Ou seja, através de chamadas de telemóvel os dados são analisados e de seguida são criadas rotas consoante a quantidade de chamadas realizadas por cada zona.

Para além disto, Seul procedeu à instalação de uma elevada quantidade de sensores de tecnologia IoT por toda a cidade, recolhendo diferentes tipos de dados, como o ruído, a vibração, raios ultravioleta, poeira fina, entre outros. A cidade, juntou a estes dados recolhidos dados do governo com o objetivo de os utilizar em diversas áreas, como, na análise de atividades de marketing, bem-estar, transporte, entre outras.

Uma particularidade é que estes dados são partilhados com qualquer cidadão, para que estes sugeriram novas formas de melhorarem a qualidade de vida. A cidade assenta num princípio em que acredita que existem três fatores fulcrais para ser uma cidade inteligente:“ infraestrutura de cidade inteligente, empresas inovadoras com tecnologias de ponta e cidadãos inteligentes” **(Stevens, 2020).** Como exemplo de melhorias, é possível destacar a utilização dos dados pelos sensores recolhidos sobre os lugares de estacionamento. O objetivo é que quando um lugar esteja livre seja indicado para que depois quem o queira ocupar tenha um acesso mais rápido e eficaz. Com todos estes dados “ Seul está criando novos serviços, como CFTV inteligente, táxis de IA e estacionamento compartilhado de IoT” **(Stevens, 2020)**. Como exemplo para as outras cidades que se querem tornar inteligentes e de forma a ajudá-las no seu crescimento Seul divulga soluções que devem ser consideradas “TOPIS (Operação de Transporte e Serviço de Informação), postes de iluminação LED incorporados em loT e sistemas inteligentes de processamento de lixo” **(Stevens, 2020)**;

* Aveiro, Portugal

*Urban Innovate Actions* é uma iniciativa da União Europeia que propõem soluções para enfrentar desafios urbanos. Aveiro, cidade portuguesa distinguida na área de *smart cities*, candidata-se a esta iniciativa com o objetivo de atuar em áreas de mobilidade, ambiente, energia e sustentabilidade. Como resultado desta candidatura, Aveiro recebeu recursos para ultrapassar desafios nas áreas desejadas, e por sua vez, implementar infraestruturas de 5G, fibra e sensores, beneficiando as empresas e centros de investigação. No que toca, à mobilidade foram utilizados sensores de geolocalização em transportes públicos com o fim de obter informações relativas aos hábitos dos cidadãos, identificar problemas associados aos transportes e consequentemente proceder a sua otimização. A nível da energia foram promovidos veículos elétricos assim como as estações de carregamento dos mesmos. Relativamente, ao meio ambiente com o objetivo de gerar um mapa ambiental da cidade com dados ativos para os cidadãos e/ou entidades de investigação foram colocados sensores fixos e moveis para recolher dados **(Lab, 2021)**;

* Vila Nova de Famalicão, Portugal

O projeto *b-smart[[2]](#footnote-2)*, é um projeto que visa o desenvolvimento de uma plataforma de inteligência urbana, que aglomera diversas ferramentas inteligentes presentes na gestão do município, e foi proposto no Portugal *Smart Cities* no ano de 2021. Este projeto visa criar uma cidade inovadora e simultaneamente inclusiva e moderna. Famalicão está em plena transformação apostando em diversas áreas, como por exemplo, melhor gestão energética e ambiental. A inovação a nível de infraestruturas e mobilidade também são objetivos a desenvolver. Para que a vila cresça na direção de uma cidade inteligente sustentável foi desenvolvida uma aplicação apelidada de centro de controlo, que processa, cruza e analisa toda a informação do dia a dia dos famalicenses, para que, a gestão territorial seja mais eficiente **(CMVNF, 2019)**.

## A indústria das Cidades 4.0

“Nas economias desenvolvidas, transporte e congestionamento estão a apresentar uma tendência crescente” **(Das, 2019).** Isto deve-se a uma crescente popularização e também ao desenvolvimento económico **(Das, 2019).** A mobilidade urbana está em mudança e os transportes públicos devem oferecer serviços mais informados e completos para os passageiros **(Dilax, sem data).** Assim como, a gestão de parques de estacionamento deve ser melhorada. Segundo o *USA Today* os motoristas gastam em média 17h por ano à procura de uma vaga de estacionamento **(Das, 2019).** Nesta secção irão ser abordado dois temas, nomeadamente, uma gestão mais eficaz para os parques de estacionamento e para os serviços de transporte público.

### Parques de Estacionamento

Segundo o artigo ”Uma nova gestão dos sistema de estacionamento, para cidades inteligentes, para economizar combustível, tempo e dinheiro” publicado em 2019 por Das  **(2019)** a dificuldade em encontrar um estacionamento livre está aumentar. Deste modo, a Inrix desenvolveu um estudo em várias cidades do EUA e concluiu que:

* Anualmente, 70 bilhões de dólares (equivalente a aproximadamente 62 bilhões de euros) são desperdiçados à procura de estacionamento;
* Relativamente ao tempo, são desperdiçadas 3,6 bilhões de horas;
* No que toca ao combustível 1,7 bilhões são desperdiçados.

Deste modo, e com a quantidade de recursos desperdiçados, é proposto um novo sistema de estacionamento, composto por dois módulos. É importante referir que a quantidade de componentes depende do local de implementação do sistema.

Na Tabela 5, são apresentados os elementos do sistema de Hardware e Software proposto por Das **(2019).**

Tabela 5 - Elementos de um novo modelo de estacionamentos

| **Módulo** | **Componente** | **Função** |
| --- | --- | --- |
| **Hardware** | Dispositivos *Raspberry Pi* | Principal Unidade de computação |
| Sensores de distância ultrassônico | Detetar se o carro está a chegar e comunicar com os dispositivos *Raspberry Pi* |
| Câmaras *Raspberry Pi* | Tirar foto à parte traseira do carro que está conectada ao *slot* da câmara do dispositivo *Raspberry Pi* |
| Placas de exibição | Um *display* por pista e um painel para cada vaga de estacionamento |
| *BreadBoards* e Fios | Conectar os componentes hardware com os dispositivos *Raspberry Pi* |
| **Software** | *Python* | Todo o código é escrito nesta linguagem |
| Base de dados em nuvem | Gestão de dados |
| *API Google Cloud Vision* | Extrair o número da matrícula |
| *Framework sendmail do Simple Mail Transfer Protocol* (SMTP) | Atribuições de estacionamento comunicadas |

De forma a perceber o funcionamento e interligação dos elementos do novo modelo de estacionamento é descrito todo o fluxo de forma textual e seguidamente apresentada uma imagem (**Error! Reference source not found.**) onde são apresentadas visualmente as ações de entrada no parque. A unidade Master *Raspberry Pi* (localizada na pista de entrada) tem um sensor de distância que está constantemente ativo a verificar se um carro se aproxima da entrada do estacionamento. Caso o sensor identifique o movimento de um carro, o mesmo envia um sinal para todas as 16 unidades *Raspberry Pi* que estão nessa faixa de entrada. Consequentemente, as 16 unidades acionam a câmara para captar a matrícula do carro. Através do código *python* este irá verificar se a imagem captada tem ou não a foto da matrícula através da *API Google Cloud Vision*. Em caso afirmativo, é enviada uma mensagem para a unidade *Master Raspberry Pi* com o número da matrícula identificado e o respetivo índice do *Raspberry Pi.* Um exemplo dado pelo artigo é “por exemplo, um *Raspberry Pi* pode comunicar (através de uma base de dados): Eu sou *Raspberry Pi* 4, e minha imagem encontrou o número da placa 6GWD345; outro *Raspberry Pi* pode comunicar: Eu sou Raspberry Pi 7, e minha imagem encontrou o número da placa Go-49ers” **(Das, 2019).** É de referir que dados ilegíveis serão ignorados pelo *Master Raspberry Pi* e todas as informações sobre a identificação do estacionamento e matrícula serão guardadas na base de dados.

Através da seguinte imagem (Figura 2), é descrita de forma mais objetiva e sequencial todo o fluxograma das ações de entrada.

Diagram, text, letter

Description automatically generated

Figura 2 - Fluxograma das ações de entrada num parque de estacionamento, adaptado de **(Das, 2019)**.

Concluindo, quando o veículo está estacionado no local, o *Raspberry Pi* altera a condição para lugar ocupado e, por sua vez, quando o mesmo sai do estacionamento a condição altera-se novamente para livre. Todas estas informações são comunicadas ao dispositivo master que guarda os dados na base de dados. Neste sistema, o carro, na entrada do parque, recebe um pager com o lugar que tem disponível para estacionar. Sendo que depois será confirmada através de todo o sistema se o carro está no devido lugar que lhe foi atribuído.

Uma outra solução, desenvolvida pela *All Traffic Solutions* visa partilhar com os utilizadores, através de uma plataforma, informações sobre a taxa de ocupação e lugares livres num determinado parque. **(«Connected Traffic Management Solutions», sem data).**

### Taxa de ocupação nos transportes públicos

Relativamente aos transportes públicos também é possível afirmar que uma das grandes dificuldades atualmente é saber quais os transportes que tem lugares vagos para um novo passageiro, qual a melhor rota e o tempo que o mesmo vai demorar a chegar **(Ctmucg, 2017).**

Deste modo, surge o “Milênio Bus” que visa resolver os problemas acima mencionados através de uma aplicação inteligente, que propõe em tempo real, a otimização de rotas e monitorização do número de passageiros em cada autocarro. Para isto baseiam-se em tecnologias como IoT, IA e visão computacional **(Ogando, sem data).** Sensores, componentes eletrónicos e GPS são instalados próximos da porta de entrada e saída do autocarro, através da Figura 3, a seta e o círculo preto demarcam os componentes referidos.



Figura 3 - Localização dos componentes eletrónicos nos autocarros, adaptado de **(Marcel Ogando, 2020)**.

A maior parte da população utiliza um smartphone, consequentemente, sempre que algum dispositivo passa pelo sensor é contabilizado. Assim sendo, é possível obter as informações de embarque e desembarque de cada pessoa e perceber as estações de paragem com maior aglomeração de pessoas **(Ogando et. al., 2021)**. O objetivo desta solução é “fornecer inteligência capaz de tornar o transporte coletivo mais eficiente e confortável nas cidades” **(Ogando, sem data).** No entanto, colocam-se as seguintes questões: E se uma pessoa possuir mais do que um smartphone? E se uma criança, que não possui dispositivo móvel, entra no autocarro? A contagem torna-se não fiável e “o volume de passageiros pode ser subestimado ou superestimado” **(Olivo et al., 2019).**

Para além do método baseado em dispositivos móveis, existem inúmeras técnicas que tem vindo a ser aplicadas para a contagem de passageiros, é possível enumerar e destacar algumas: venda de bilhetes/passes aos passageiros e consequentemente verificador do bilhete dentro do autocarro, assim como, a cobrança automática do bilhete. No entanto, estas técnicas não fornecem informações de rota nem de tempo, e apesar da obtenção do bilhete o passageiro não é obrigado a utilizá-lo **(Olivo et al., 2019).** Um outro sistema baseado no peso dos passageiros, que pode ser obtido através de sensores de carga no solo. Contudo, é apenas uma estimativa de volume de passageiros (contada pelo peso total de pessoas no transporte público) e não uma contagem exata e precisa. Assim como, “não oferecem dados sobre o fluxo de passageiros”**(Olivo et al., 2019).** Sistemas de tapetes (que são sensíveis a pressão e medem o volume total através dos passos das pessoas), tecnologia de infravermelhos (medido através de feixes de luz, ou seja, quando os mesmo são interrompidos é registada uma contagem) e sistemas de imagem de vídeo. Todas estas tecnologias apresentam várias desvantagens, porém, a precisão e a exatidão são as que mais se realçam neste tipo de sistemas **(Olivo et al., 2019).**

Posto isto, surge a DILAX que apresenta uma junção de algumas tecnologias que foram mencionadas anteriormente e garante “uma precisão de até 99%” **(Dilax, sem data)**. Este sistema apresenta dois componentes: o sensor que procede a contagem dos passageiros e a unidade que processa os dados, armazena e transmite. Este APC (contagem automática de passageiros) conta com informações da bilheteira, sensores capacitivos, GPS e tecnologia infravermelha. Por meio de uma plataforma online, em tempo real, consegue transmitir para os passageiros a localização do transporte público assim como o fluxo de passageiros e a taxa de ocupação de lugares **(Dilax, sem data).**

## Visão Crítica do Estado da Arte

A indústria 4.0 surge com a necessidade de melhorar os modelos de negócio, assim como, os sistemas de produção. Esta revolução trouxe inúmeras tecnologias que estão a permitir às cidades industrializarem-se e evoluírem. Assim como, a indústria têxtil e automóvel estão a crescer no sentido de sustentabilidade e inovação. É de facto importante referir que as tecnologias associadas a esta quarta revolução industrial trouxeram numerosos benefícios socias e económicos, tais como, o aumento da produtividade, redução de custos de produção, eficiência, automatização de processos repetitivos, entre outros **(Neves et. al., 2022)**.

Com esta transformação digital nasce o conceito de Cidades Inteligentes aliado a uma tendência global de urbanização sustentada pelo digital **(Anthony Jnr, 2021)**. É do senso comum, que as cidades, principalmente as metropolitanas, estão a cada dia que passa com maior número populacional, provocando um grande aumento do tráfego, número de acidentes e poluição **(Das, 2019)**. Neste sentido, através de uma análise da indústria das cidades 4.0, é percetível perceber que há várias soluções, a nível dos parques de estacionamento e transportes públicos, que aliadas seriam uma grande vantagem para as cidades. Das **(Das, 2019)** propõe um novo sistema de gestão de estacionamentos através da utilização de sensores, câmaras e microprocessadores. No entanto, uma das desvantagens que o sistema apresenta, diz respeito aquando da entrada do carro no parque, onde o mesmo tem de parar para receber um pager com a informação do número de lugar disponível para estacionar o veículo. Porém, esta condição não é de todo prática. Assim sendo, é possível adquirir uma solução mais viável como o desenvolvimento de uma aplicação de reserva de lugares num parque de estacionamento produzida por a *All Traffic Solutions* **(«Connected Traffic Management Solutions», sem data)**. Deste modo e juntando estas duas soluções é possível gerir de forma mais eficiente a gestão dos carros num parque de estacionamento. Onde vários fatores poderão diminuir com esta solução, nomeadamente, tempo, dinheiro e combustível. Relativamente à taxa de ocupação de transportes públicos, existem duas soluções relevantes, a *Milenio Bus* e a *Dilax*. A primeira solução referida, permite, em tempo real, a otimização de rotas e monitorização do número de passageiros em cada autocarro através da contabilização de cada smartphone que entra num transporte publico. Contudo, e como já foi referido, existem algumas desvantagens associadas, por exemplo uma criança pode não possuir nenhum smartphone e um adulto possuir dois smartphones **(Marcel Ogando, 2020)**. A segunda solução, *Dilax*, apresenta um sistema para contagem de passageiros através de GPS, tecnologia infravermelha, sensores e informação da bilheteira. Sendo um sistema tão completo tem a possibilidade de garantir até 99% de eficácia na contagem dos mesmos **(Dilax, sem data)**. Concluindo, esta solução da DILAX aliada à componente de otimização de rotas da *Milenio Bus* tornar-se-ia mais completa e informada para os passageiros, que poderiam contar com informações relevantes ao nível da lotação dos transportes públicos e das melhoras rotas a adquirir. Apesar das diversas soluções existentes nenhuma é 100% eficaz, apresentando falhas que impedem o sucesso e veracidade das mesmas. Este projeto de dissertação pretende ajuda o ioCity a colmatar algumas das limitações identificadas e promover uma solução mais eficaz. Em suma, a inovação da tecnologia é a chave para uma economia competitiva, levando à inovação e crescimento das cidades, permitindo um mundo mais sustentável e seguro.

# Abordagem Metodológica, Materiais e Métodos

Subjacente à realização da presente dissertação, existem metodologias, materiais e métodos que auxiliaram todo o processo de desenvolvimento.

## Metodologias de Investigação

Para o desenvolvimento da dissertação foram utilizadas duas metodologias, *Design Science Research* (DSR) e *Case Study*.

### Metodologia DSR

A metodologia DSR é importante na área de sistemas de informação, uma vez, que é uma metodologia orientada para a criação de artefactos. As técnicas e perspetivas que esta metodologia oferece para a pesquisa científica promove a qualidade de desempenho funcional dos sistemas de informação **(Cruz, 2011)**. Deste modo, existem seis etapas fundamentais, que podem ser visualizadas na Figura 4.

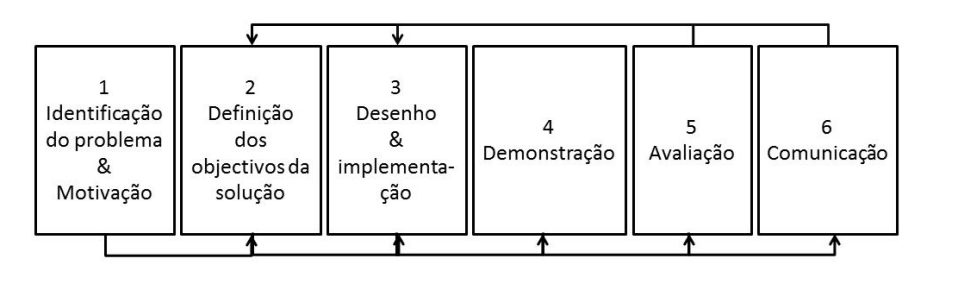


Figura 4 - Atividades da Metodologia DSR, retirado de **(Cruz, 2011)**.

* **Identificação do Problema e Motivação**

Inicialmente, é necessário identificar o problema específico de investigação. Consequentemente, o levantamento do estado da arte, para que o conhecimento na área auxilie na definição do problema e ajude a chegar à solução. Através da criação de artefactos, e do conhecimento adquirido, torna-se possível o progresso e a implementação de soluções baseadas em tecnologias para a resolução de problemas **(Cruz, 2011)**. Este projeto surge com a necessidade de alterar as alternativas de mobilidade, devido ao excesso congestionamento, poluição e acidentes. O problema e motivação foram previamente descritos na secção 1.1;

* **Definição dos objetivos da solução**

Após o desenvolvimento da primeira fase, identificação do problema e motivação, os objetivos devem ser definidos. É expectável que haja um conhecimento do estado do problema e das soluções existentes no mercado, para que o objetivo da solução seja inovador e eficiente **(Cruz, 2011)**. Neste trabalho, o principal objetivo é desenvolver um sistema de apoio à tomada de decisão, que em tempo real, contribua para uma decisão mais informada ao nível dos estacionamentos e/ou transporte (secção 1.2)

Para que este sistema, seja real e vá de encontro às expectativas dos cidadãos foi utilizada a metodologia *case study* nesta etapa. Esta metodologia de pesquisa tem como objetivo estudar o impacto que um sistema de informação tem num contexto social **(Antonio et al., 2018)**. Deste modo, foi desenvolvido um *google forms,* para identificar o grau de importância que cada critério de decisão, entre transportes públicos e parques de estacionamento, tem no quotidiano. Desta forma, todos os cidadãos, através da resposta ao formulário, previamente publicado nas redes sociais e enviado por email para todas as universidades do país, possam selecionar os fatores que priorizam na tomada de decisão. Este estudo tem também como objetivo ajudar a construir e otimizar o modelo inteligente a ser desenvolvido seguindo as necessidades e expectativas que a população e a cidade necessitam;

* **Desenho e implementação**

Numa terceira fase, propõem-se o desenvolvimento de um artefacto que contenha as funcionalidades subjacentes à solução, assim como, a sua definição e características, e a arquitetura do protótipo para que depois a construção do artefacto real seja coerente e o seu desenvolvimento simplificado. Este artefacto deve ter por base modelos, métodos, algoritmos e/ou técnicas **Cruz, 2011)**. Para a fase de desenho e implementação será seguida a metodologia CRISP-DM (descrita na subsecção 3.2.1).

Nesta fase, segundo a metodologia CRISP-DM existem seis etapas cruciais, as quais foram seguidas atentamente para o alcance de um trabalho exímio e útil no contexto. Inicialmente, e após a determinação dos objetivos, foi estudado todo o contexto subjacente ao projeto analisando diversas variáveis que podem interferir com a decisão entre a escolha do método de transporte a escolher. Para isto, um formulário foi desenvolvido (referido também na etapa anterior Definição dos objetivos da solução), assim como, análises estatísticas e um método de clustering foi aplicado aos resultados de forma a perceber quais as variáveis que os cidadãos priorizam na tomada de decisão. Na compreensão e preparação dos dados o processo de ETL e diversas análises foram executados de forma a compreender o tipo de dados, estrutura e indicadores/KPIs que podem ser estudados. Na etapa seguinte, modelação, todos os requisitos necessários foram desenvolvidos, nomeadamente, o desenvolvimento de um *data warehouse* para o armazenamento dos dados e um modelo OLAP para uma análise tridimensional sobre os mesmos. Por fim, a *dashboard* analítica foi criada e avaliada de forma a perceber se cumpria os requisitos necessários. Esta fase foi descrita detalhadamente na secção 4.1 e 4.3;

* **Demonstração**

Nesta etapa, é esperado, a demonstração de como o artefacto desenvolvido resolve o problema previamente identificado. Para que a prova do artefacto na resolução do problema seja eficaz e tenha valor é necessário a criação de um ambiente bem definido e apropriado, podendo envolver experiências, simulações, aplicações em casos de estudo, entre outros **(Cruz, 2011)**. A solução desenvolvida é escalável e os dados são indicativos do que é possível obter. Esta fase foi descrita detalhadamente na subsecção 4.3.5 de forma a compreender todo o trabalho elaborado e como a solução apresentada pode ser útil no contexto da tomada de decisão. Todas as funcionalidades são expostas, assim como a arquitetura da solução.

* **Avaliação**

A avaliação consiste na observação e medição da forma como o artefacto suporta a solução do problema, e se o mesmo vai de encontro aos requisitos exigidos e definidos inicialmente. A avaliação pode ser realizada através de diferentes métodos dependendo do problema em questão. Pode ser feita uma comparação das funcionalidades desenvolvidas com as pré-definidas no artefacto, ou, por exemplo, através de simulações medir quantitativamente a performance **(Cruz, 2011);**

* **Comunicação**

A comunicação, é a última fase da metodologia DSR, e assenta na comunicação e divulgação do problema e da sua relevância, assim como, a demonstração de todos artefactos e a sua utilidade, os resultados obtidos e a prova de eficiência da solução.

Através da Figura 4, onde estão representadas todas as atividades que devem ser executadas aquando da utilização da metodologia DSR, é possível perceber que o processo pode ser executado em sequência ou podem seguir uma ordem diferente. Na fase 5, avaliação, ou na última etapa, comunicação, pode haver um retrocedimento para a atividade 2 ou 3, dependendo dos resultados da avaliação, onde os objetivos possam ter de ser redefinidos **(Cruz, 2011)**.

Nesta fase três artigos foram escritos e publicados (apresentados na secção 6). Um artigo intitulado de “*Opinion clustering about mobility decisions - A practical case study*” onde toda investigação sobre quais as variáveis que influenciam a tomada de decisão foram apresentadas. Este artigo foi também apresentado na conferência internacional ARTIIS 2022[[3]](#footnote-3).

Um segundo artigo “*Intelligent dashboards for car parking flow monitoring*” onde o objetivo foi demonstrar a importância de um modelo OLAP na tomada de decisão e como todo o fluxo até a construção do modelo é fundamental para um modelo útil e eficaz. Por último, foi escrito um artigo denominado por “*An analysis of Smart Cities Initiatives in Portugal*” que demonstra o estado atual das *smart cities* em Portugal e o tipo de iniciativas que tem vindo a ser desenvolvidas. Nesta fase é também importante referir que toda a documentação subjacente à realização do projeto encontra-se no presente documento.

A metodologia DSR foi crucial para uma pesquisa intensa e organizada sobre todo o contexto subjacente a realização da presente dissertação, permitindo a aquisição de conhecimento essencial sobre todo o trabalho. Através do desenvolvimento de artefactos o entendimento do problema e da solução foram alcançados mantendo o foco de investigação e resultados válidos e coerentes de acordo com todo o projeto.

### Metodologia Caso de Estudo

Aquando da dificuldade de perceber um fenómeno estudado e o seu real contexto é utilizada a metodologia *case study* **(Antonio et al., 2018)***.* Existem características, que fazem desta metodologia, um método de pesquisa diferente dos já existentes, nomeadamente, a forma de elaboração da pergunta de pesquisa, procedimentos, recolha de dados, posterior análise e critérios de validação. É de realçar que a pergunta de pesquisa deve responder a “como” e “porque” que este estudo está a ser desenvolvido, tendo por base uma explicação contextual que descreva o sistema de informação em estudo, neste caso, transportes públicos e parques de estacionamento. Na Tabela 6, são apresentadas diferentes estratégias de pesquisa que auxiliam no desenvolvimento da metodologia em questão **(Antonio et al., 2018)**.

Tabela 6 - Estratégias de pesquisa da metodologia Case Study, retirado de (Antonio et al., 2018)

| **Estratégia** | **Forma da Questão da Pergunta** | **Exige controlo sobre eventos comportamentais** | **Focaliza acontecimentos contemporâneos** |
| --- | --- | --- | --- |
| **Experimento** | Como, porquê. | Sim | Sim |
| **Levantamento** | Quem, o que, quantos, quanto. | Não | Sim |
| **Análise de requisitos** | Quem, o que, quantos, quanto. | Não | Sim |
| **Pesquisa Histórica** | como, porque | Não | Não |
| **Estudo de caso** | como, porque | Não | Sim |

Algumas das características fundamentais do estudo de caso, são enumeradas na seguinte lista:

* O estudo deve ser desenvolvido num ambiente natural, baseado em várias fontes de evidências e dados recolhidos de diferentes meios e entidades **(Antonio et al., 2018)**;
* Com o fim de obter um bom resultado, o pesquisador deve ter uma atitude positiva perante os dados analisados **(Antonio et al., 2018)**;
* Proposições teóricas devem ser desenvolvidas para que a recolha e análise de dados seja objetiva e coerente **(Antonio et al., 2018)**.

Como referenciado anteriormente, na segunda etapa da metodologia *DSR,* o estudo de caso foi desenvolvido com o intuito de perceber que fatores entre o uso de veículos individuais e transportes públicos, os cidadãos priorizam. Este estudo pode ser realizado através de diferentes métodos (entrevistas, observação direta, análise de documentação e/ou análises mais complexas).

Através da metodologia *Case Study* toda a componente de pesquisa sobre a dificuldade de perceber o contexto real sobre quais as variáveis que os cidadãos priorizam na tomada de decisão foi necessário para um estudo completo, real e eficaz.

## Metodologias Práticas

Para as metodologias práticas, foram abordadas duas metodologias, CRISP-DM e SCRUM.

### Metodologia CRISP-DM

A metodologia CRISP-DM (*cross-industry process for data mining*) é apelidada como um processo padrão inter-indústrias para explorar dados e criar padrões. Descreve abordagens usadas por especialistas em mineração de dados para atacar problemas, oferecendo um modelo de orientação de projetos de *Data Mining*. De acordo com a metodologia mencionada anteriormente, este projeto é dividido em seis etapas distintas, organizadas de forma cíclica **(Rodrigues, 2020):**

* **Compreensão do Negócio**

Esta primeira etapa, tem como objetivo principal determinar as metas e objetivos para que o resultado seja de qualidade. É importante averiguar os recursos, restrições e perceber todo o contexto de negócio para que os objetivos projetados sejam relevantes e concretizáveis;

* **Compreensão dos dados**

Seguidamente, é importante recolher os dados, descrevê-los e averiguar a qualidade dos mesmos. Deve ser feita uma exploração de forma a perceber as características e informações presentes no conjunto de dados;

* **Preparação dos dados**

Inicialmente, deve ser feita uma análise e pré-seleção dos dados que se enquadram com as necessidades do projeto, e, posteriormente, verificar a qualidade dos mesmos. Pelo que, nesta fase, é efetuada uma limpeza, seleção e transformação, com o objetivo, de garantir a qualidade para uma posterior análise;

* **Modelação**

Na quarta fase são usadas técnicas de modelagem, assim como a configuração de parâmetros, usadas para diferentes cenários e testes com o fim de obter a otimização dos dados;

* **Avaliação**

Na avaliação são analisados ​​os impactos dos resultados dos modelos gerados a partir da etapa anterior e se os mesmos estão em coerência com os objetivos de negócio definidos previamente;

* **Implementação**

Por fim, a implementação, consiste num plano de desenvolvimento e implementação da solução num ambiente que seja disponibilizado ao utilizador. Dependendo dos requisitos do projeto, poderá ser apenas necessário a elaboração de um relatório final ou proceder-se efetivamente à implementação, monitorização e manutenção;

A metodologia CRISP-DM foi fundamental para a obtenção de uma solução de estrutura simples, de fácil compreensão e que vá de encontro aos objetivos e requisitos pretendidos.

### Metodologia SCRUM

Para a gestão do projeto foi utilizada a metodologia SCRUM, com o objetivo de melhorar a eficiência do projeto e reduzir desperdícios de tempo e recursos. Primeiramente, é necessário definir e perceber o papel a desempenhar no projeto, seguidamente listar as prioridades, definir e planear os sprints, e posteriormente, reuniões semanais e diárias **(Builder, 2020)**.

Existem três cargos possíveis a desempenhar nesta metodologia, nomeadamente:

* *Product Owner* (normalmente este papel é desempenhado por *stakeholders* que definem os objetivos a atingir e direciona a equipa no sentido de atingir os resultados esperados). No projeto ioCity este papel é desempenhado por Carlos Filipe Portela;
* *Scrum Master* (pessoa responsável por gerir o projeto e a equipa). No projeto ioCity este papel é desempenhado por Carlos Fernandes;
* Equipa de desenvolvimento (membros da equipa responsáveis por todo o desenvolvimento das tarefas definas para os sprints). No projeto ioCity este papel é desempenhado por três elementos, eu, Francisca Barros para a componente de *Data Science*, Beatriz Pereira responsável pelo front-end e João Oliveira responsável pelo back-end.

Os sprints são uma sequência de atividades com prazo definido, que devem ser desenvolvidos durante esse tempo. Deste modo, há uma reunião quinzenal para definir as tarefas para o sprint (*sprint planning*) onde o *project owner,* baseando-se no *product backlog* (lista com todas as funcionalidades desejadas do produto), divide as tarefas para o sprint em questão. Durante as duas semanas de trabalho, através de reuniões diárias, cada membro do projeto partilha as atividades a realizar para o dia, assim como, o que fez no dia anterior e respetivas dificuldades. Após o término do sprint, há uma reunião para validar todo o trabalho desenvolvido e consequentemente, uma nova reunião para o *sprint planning* **(Builder, 2020).**

Os sprints, neste projeto em concreto, têm uma duração de quinze dias e até ao término do projeto foram desenvolvidos 16 sprints. Como referido anteriormente, na presente secção, os sprints baseiam-se no *product backlog.* Consequentemente, as principais tarefas a concretizar, são: elaborar um estudo acerca dos fatores críticos que influenciam a tomada de decisão relativamente ao transporte a escolher, interpretação do problema e da realidade, extrair, carregar e transformar os dados para a criação da *data warehouse* (DW), desenvolvimento dos KPIs e do mecanismo OLAP. E por fim, retirar informações relevantes dos dados (*data science*) e construir um modelo inteligente de apoio á decisão. Todo o planeamento pode ser visualizado na secção 6 através do diagrama de *Gantt*.

Todo o planeamento delineado através da presente metodologia, proporcionou uma organização e ritmo de trabalho imprescindível para que a solução desenvolvida cumpra todos os requisitos.

Concluindo, esta metodologia proporciona uma maior qualidade, rigor e eficiência à gestão de projetos, potenciando a qualidade e agilidade enquanto equipa e desenvolvimento da solução em questão **(Builder, 2020)**.De forma sucinta e objetiva, na Tabela 7 - Resumo de metodologias é apresentado um resumo das quatro metodologias inerentes à execução do projeto que foram apresentadas na subsecção 3.1 e 3.2.

Tabela 7 - Resumo de metodologias

| **Metodologia** | **Tipo** | **Objetivo** |
| --- | --- | --- |
| **DSR** | Metodologia de Investigação | Criação de artefactos para o desenvolvimento do projeto |
| **Case Study** | Metodologia de Investigação | Analisar e definir um caso de estudo |
| **CRISP-DM** | Metodologia Prática | Conjunto de boas práticas para execução de um projeto de *data mining* |
| **SCRUM** | Metodologia Prática | Gestão de projetos |

Foi também elaborado um mapeamento das metodologias tendo por base as seis etapas fundamentais da metodologia DSR (descrita na secção 3.1.1). Ou seja, onde é que fases das diferentes metodologias se cruzam. Por exemplo, a etapa 2 da metodologia DSR (definição dos objetivos da solução) interfere com a metodologia Case Study e com a primeira e segunda fase da metodologia CRISP-DM. Todas as metodologias identificadas na Tabela 8 foram previamente descritas na secção 3.

Tabela 8 - Mapeamento das Metodologias

| **Metodologias**  **Etapas DSR** | **Case Study** | **CRISP-DM** | **SCRUM** |
| --- | --- | --- | --- |
| **Etapa 1**  **Identificação do Problema e Motivação** |  | X  (Fase: Compreensão do Negócio (1)) |  |
| **Etapa 2**  **Definição dos objetivos da solução** | X | X  (Fase: Compreensão do Negócio (1) e dos dados (2)) |  |
| **Etapa 3**  **Desenho e implementação** |  | X  (Fase: Compreensão dos dados (2), Preparação dos dados (3), Modelação (4) e Implementação (6)) | X |
| **Etapa 4**  **Demonstração** |  | X  (Fase: Avaliação (5) e Implementação (6)) | X |
| **Etapa 5**  **Avaliação** |  | X  (Fase: Avaliação (5) e Implementação (6)) | X |
| **Etapa 6**  **Comunicação** |  |  |  |

## Lista de Ferramentas

Para o desenvolvimento da presente dissertação, foram utilizadas diversas ferramentas para alcançar o objetivo pretendido. As ferramentas utilizadas, antes de serem escolhidas sofreram uma análise tendo em consideração as necessidades futuras. Desta forma, na Tabela 9, são apresentadas as ferramentas selecionadas para a execução do projeto.

Tabela 9 - Lista de Ferramentas inerentes à execução do projeto

| **Nome da Ferramenta** | **Tipo** | **Funcionalidade** |
| --- | --- | --- |
| ***Microsoft Office Excel*** | Folha de cálculo | Operações de suporte à  análise de dados. |
| ***Python*** | Linguagem de programação | Desenvolvimento do modelo de dados e protótipo. |
| ***MySQL Workbench*** | Ferramenta de design para base de dados | Criação e manutenção modelo de dados. |
| **Biblioteca Pandas** | Software de código aberto | Ferramenta para análise e manipulação de dados. |
| ***Past*** | Software estatístico | O *Paleontological Statistics* é um programa de software para análise de dados estatísticos. |
| ***Jupyter Notebook*** | Software de código aberto | Plataforma de computação interativa baseada na web, onde é possível desenvolver código aberto. |
| ***Cube.js*** | Software de código aberto | Plataforma interativa para construção de aplicações web analíticos. |
| ***Visual Studio Code*** | Software de código aberto | Editor de código fonte. |
| ***Vue.js*** | Linguagem de programação | Estrutura JavaScript *front-end* para o desenvolvimento da interface de visualização. |
| ***Vuetify*** | Biblioteca de interface gráfica para o utilizador | Framework que auxilia o desenvolvimento web, disponibilizando ferramentas de design úteis para o *front-end* da *dashboard.* |
| ***Apache ECharts*** | Biblioteca de visualização JavaScript de código aberto | Desenvolvimento de gráficos de visualização intuitivos, interativos e personalizáveis. |
| ***Google Maps*** | Plataforma de mapeamento web | Sistema de geolocalização através de mapas dinâmicos. |
| ***Node.js/Express.js*** | *Framework* de *back-end* para o desenvolvimento de APIs | Api *back-end* para definir as rotas dos dados via mysql. |

.

# Caso de Estudo

Neste capítulo é descrito todo o trabalho realizado em três subcapítulos. Inicialmente, foi elaborado um questionário (segundo a metodologia *Case Study*) de forma a abordar os critérios que influenciam a tomada de decisão da população entre o tipo de transporte a eleger (transportes públicos ou privados). Perante os critérios selecionados, foi feito um levantamento de APIs que asseguram a recolha de alguns dos dados necessários, sendo que a maior parte dos dados são assegurados pela Câmara Municipal de Vila Nova de Famalicão. Por fim, é apresentada a solução desenvolvida e todas as componentes necessárias para a realização da mesma, nomeadamente, criação de uma *Data Warehouse* e um modelo OLAP.

## Estudo das variáveis

Este estudo teve como principal objetivo perceber os fatores que são importantes no dia a dia da população e de que forma estes interferem com a mobilidade urbana, desta forma, o modelo a desenvolver vai de encontro aos fatores críticos selecionados pelos cidadãos. Após a publicação de um formulário nas redes socias e o envio de emails para as universidades do país como forma de divulgação, foi elaborado um levantamento das respostas, e, consequentemente, elaborada uma análise estatística para perceber quais os critérios que apresentavam maior impacto. Assim como, um estudo com a técnica de *clustering* foi desenvolvido para perceber os perfis associados e quais as variáveis com maior concentração de respostas de grau mais elevado.

### Critérios de apoio à tomada de decisão

A possibilidade de os cidadãos terem acesso a bens e serviços melhora a sua qualidade de vida. Existem vantagens e desvantagens perante o método de transporte eleito, contudo o uso de veículos individuais tem impactos bastante negativos no que diz respeito ao ambiente (poluição gasosa e sonora, consumo excessivo de combustível, entro outros) e também relativamente à mobilidade (aumento do congestionamento e sinistralidade rodoviária). Assim como, a utilização de transportes públicos limita os passageiros em termos de horários, flexibilidade e conforto **(Costa, 2008).**

Na área da mobilidade urbana o trânsito existente deve-se à grande utilização dos veículos pessoais e baixa utilização dos transportes públicos, assim como, a grande dificuldade de os condutores encontrarem um parque de estacionamento.

#### Transportes públicos

Existem várias perspetivas que podem ser tomadas em consideração dependendo da posição pela qual encaramos o problema. É possível pensar e agir como utilizadores, motoristas da viatura, operadores, autoridade, peões, comerciantes, entre outros **(Costa, 2008).**

A qualidade do serviço é um conceito subjetivo, dependendo da relação existente entre as expectativas e o serviço efetuado. Para que os transportes públicos cresçam na sua utilização é necessário compreender os diversos aspetos que tem um maior impacto na satisfação dos usufrutuários **(Sousa, 2020).** Visto que o projeto em questão, IoCity, é destinado à população irão ser descritas variáveis que influenciam a tomada de decisão dos mesmos.

Na Tabela 10 são apresentados os critérios de acordo com a norma EN 13816 **(Sousa, 2020)**, norma europeia que apresenta oito parâmetros que devem ser tidos em consideração para efetuar a avaliação da qualidade dos serviços dos transportes públicos, nomeadamente:

* **Disponibilidade:** Localização geográfica, fluxos de passageiros, horários e frequência dos transportes públicos;
* **Acessibilidade:** Custos inerentes e condições do meio de transporte com acessos especiais;
* **Informação:** Informações sobre os diversos meios de transporte;
* **Tempo:** Informações temporais significativas;
* **Atendimento ao utilizador:** Serviços que incrementam a qualidade e a relação entre o meio de transporte e o passageiro;
* **Conforto:** Características físicas que permitem um maior conforto para o passageiro;
* **Segurança:** Segurança dentro dos meios de transporte público assim como das vias e infraestruturas relacionadas;
* **Impacto Ambiental:** Impacto que este método de transporte tem no ambiente.

Nesta mesma tabela (Tabela 10) é apresentada para cada norma um conjunto de critérios que estão associados a uma definição, sendo que, todos os critérios foram definidos através da leitura de artigos seguindo a estratégia de pesquisa mencionada na secção 2.1.

Com o objetivo de identificar o grau de importância de cada critério foi desenvolvido um formulário, no *google forms*, para que todos os cidadãos selecionem os fatores que priorizam na tomada de decisão. O link do formulário encontra-se na secção 6.

O formulário foi realizado em março de 2022 e publicado nas redes socias para que qualquer cidadão possa ter voz ativa neste processo participativo.

Desta forma, consegue-se, rapidamente e de forma eficaz, obter informações pertinentes referentes aos critérios que apresentam uma maior importância. Na secção 4.1.4 encontra-se detalhado o estudo do formulário e respetivas conclusões.

Tabela 10 - Critérios relativos aos Transportes Públicos

| **Norma EN 13816** | **Critério** | **Definição** |
| --- | --- | --- |
| **Disponibilidade** | Distância média entre a localização geográfica do passageiro e o transporte | Refere-se à distância total média que o passageiro tem de percorrer até à paragem mais próxima. |
| Capacidade total do veículo | Capacidade do número total de passageiros que o transporte admite. |
| Percentagem de Lotação | Percentagem de ocupação do transporte. |
| Percentagem de Fluxo de passageiros | Fluxo de entrada e saída dos passageiros. |
| Número de transferências entre meios de transporte | Quantidade de mudanças que tem de ser efetuadas para chegar ao destino final. |
| Número de paragens intermediárias (viagens de longa distância) | Quantidade de paragens que o transporte irá realizar até ao destino desejado pelo passageiro. |
| **Acessibilidade** | Custo do bilhete | Custo médio que o passageiro tem no momento de compra do bilhete. |
| Custo de carga | Custo médio que o passageiro tem na aquisição de cargas extras. |
| Percentagem de versatilidade | Percentagem de Transportes adequados a passageiros com dificuldade motora, crianças e seniores. |
| **Informação** | Nº total dos diferentes tipos de transporte | Total de transportes públicos existentes. |
| Percentagem da frequência dos transportes | Percentagens da quantidade média que os transportes públicos efetuam transportes de passageiros |
| Percentagem de informações disponibilizadas | Percentagem de informações disponibilizadas para os utilizadores terem acesso aos dados da viagem. |
| **Tempo** | Tempo desde a localização geográfica do passageiro até à paragem | Refere-se ao tempo médio que um passageiro demora para  chegar à paragem. |
| Tempo de viagem (duração) | Tempo médio que um transporte demora a percorrer a linha. |
| Tempo de espera | O tempo médio de espera entre a chegada e a partida do transporte. |
| Tempo de transferência | Tempo médio que um passageiro demora a passar de uma linha para outra, isto é, tempo entre a saída de um transporte e entrada noutro. |
| Tempo total do percurso | Tempo médio que demora entre a origem da viagem e o seu destino. |
| Taxa de Pontualidade | Percentagem de transportes públicos que são pontuais nos seus itinerários. |
| **Atendimento ao Utilizador** | Percentagem de dispositivos eletrónicos | Percentagem de transportes públicos que em tempo real através de dispostitos eletrónicos possuem informações sobre horários, atrasos, compra de bilhetes, etc. |
| **Conforto** | Percentagem de Conforto físico | Percentagem relativa ao conforto do transporte público, como por exemplo, à temperatura ambiente do transporte, humidade, limpeza, nível de ruído, disponibilidade lugar a pé/sentado, etc.; |
| Percentagem de condições da Paragem | Se a paragem tem condições de espera como cobertura em caso de mau tempo, iluminação noturna, bancos, etc. |
| Percentagem tecnológica | Transporte que possui tomadas de carregamento, wi-fi , entre outros. |
| **Segurança** | Percentagem de acidentes | Percentagem de acidentes em transportes públicos anuais. |
| Percentagem de avarias | Percentagem de avarias em transportes públicos anuais. |
| Percentagens de roubo | Percentagem de roubos em transportes públicos anuais. |
| Percentagens de condições | Percentagem da qualidade da via onde os transportes públicos efetuam percursos. |
| **Impacto Ambiental** | Percentagens de veículos Híbridos | Percentagem anual de transportes públicos do tipo híbrido. |
| Percentagens de veículos a gasolina e gasóleo | Percentagem anual de transportes públicos do tipo gasolina e/ou gasóleo. |
| Percentagens de poluição | Percentagem de poluição dos transportes públicos: gases, nível de ruido, consumos energéticos, etc. |

#### Veículos particulares

O uso de transporte individual em detrimento do uso de transporte coletivo é fundamentado pelos cidadãos pela baixa qualidade dos transportes públicos.

O conforto, privacidade, flexibilidade são algumas das vantagens que levam os cidadãos a adotar veículos particulares. Porém, existem desvantagens associadas sendo uma delas o grande consumo de energia, sinistralidade, congestionamento e aumento da poluição **(Lima, 2018)**.

Os parques de estacionamentos são também uma desvantagem a ter em consideração, o “estacionamento afeta a facilidade de chegar aos destinos e, portanto, afeta a acessibilidade geral” (**Litman, 2021)**.

Existem problemas de oferta de estacionamento assim como problemas de gestão dos parques de estacionamento **(Litman, 2021)**, deste modo, irão ser descritas, na Tabela 11, variáveis que influenciam a tomada de decisão dos condutores na escolha dos parques de estacionamento.

Tal como na Tabela 10, na Tabela 11 são apresentados um conjunto de critérios que estão associados a uma definição, sendo que, todos os critérios foram definidos através da leitura de artigos seguindo a estratégia de pesquisa mencionada na subsecção 2.1.

Com o objetivo de identificar o grau de importância de cada critério foi desenvolvido um formulário, no *google forms*, para que todos os cidadãos selecionem os fatores que priorizam na tomada de decisão. O link do formulário encontra-se na secção 6.

O formulário foi realizado em março de 2022 e publicado nas redes socias para que qualquer cidadão possa ter voz ativa neste processo participativo. Desta forma, consegue-se, rapidamente e de forma eficaz, obter informações pertinentes referentes aos critérios que apresentam uma maior importância. Na subsecção 4.1.4 encontra-se detalhado o estudo do formulário e respetivas conclusões.

Tabela 11 - Critérios relativos aos Parques de Estacionamento

| **Critério** | **Definição** |
| --- | --- |
| **Custo do Combustível** | Custo médio que o condutor apresenta anualmente. |
| **Distância média entre o estacionamento do veículo e o local desejado** | Total da distância média que o condutor necessita de efetuar entre o local onde deixa o carro e o local pretendido. |
| **Taxa de disponibilidade dos parques de estacionamento** | Percentagem de disponibilidade que o parque de estacionamento tem para os veículos pessoais. |
| **Percentagem de parques públicos** | Percentagem de parques públicos que não acarretam custos para os proprietários dos carros. |
| **Preço médio do parque** | Custo de despesas ao nível do estacionamento no parque. |
| **Tempo total do percurso** | Tempo médio que demora entre a entrada no parque até encontrar um lugar de estacionamento. |
| **Tempo de procura de estacionamento** | Tempo médio que o condutor perde à procura de um estacionamento para o seu veículo. |
| **Percentagem de Pontualidade** | Percentagem média de trânsito no processo do estacionamento. |
| **Percentagem de acidentes** | Percentagem de acidentes em veículos individuais anuais. |
| **Percentagem de avarias** | Percentagem de avarias em veículos individuais anuais. |
| **Percentagem de roubo** | Percentagem de roubos em veículos individuais anuais. |
| **Percentagem de poluição** | Percentagem de poluição dos veículos individuais: gases, nível de ruido, consumos energéticos, etc. |
| **Percentagem de Conforto físico** | Percentagem de conforto relativamente à temperatura ambiente do parque, humidade, limpeza, nível de ruído, disponibilidade, etc. |
| **Percentagem de condições do parque de estacionamento** | Se o parque de estacionamento tem condições de solo (terra batida ou cimento), de edifício (parque ao ar livre ou fechado) e iluminação (escuro ou luminoso) |
| **Percentagem tecnológica** | Taxa de parques de estacionamento que possuem câmaras de vigilância em tempo real do lugar onde o carro está estacionado. |
| **Percentagem de pagamentos com sistemas eletrónicos** | Percentagem de parques que possuem sistema eletrónicos para efetuar pagamentos (cartões de crédito, débito, mbway, entre outros). |
| **Parques com lavagem do carro** | Parques que tem a possibilidade de lavagem interior e/ou exterior do veículo. |
| **Parques com carregamentos para carros elétricos** | Parques que possuem zonas de carregamento para carros híbridos e/ou elétricos. |
| **Percentagem de controlo de segurança** | Percentagem de parques que apresentam um sistema de segurança que permita a prevenção de roubos e assaltos. |
| **Localização geográfica** | Percentagem de parques que apresentam uma localização central e perto de edifícios empresariais e/ ou perto de transportes públicos. |

### Descrição do Caso de Estudo

Conforme descrito na subsecção 3.1.2 (onde as características referentes à metodologia *case study* são abordadas), serve a presente subsecção para mapear todo a componente teórica com o estudo desenvolvido no âmbito do projeto. Na seguinte tabela, encontra-se detalhado os requisitos da metodologia, onde as estratégias de experimento, levantamento, análise de requisitos, pesquisa histórica e estudo de caso são expostos. O caso de estudo em questão, é perceber que fatores os cidadãos priorizam na tomada de decisão entre o meio de transporte a escolher (Tabela 12).

Tabela 12 - Descrição do Caso de Estudo

| **Estratégia** | **Forma da Questão da Pergunta** | **Justificação** |
| --- | --- | --- |
| **Experimento** | Como, porquê. | Após a leitura de artigos relacionados com o tema de mobilidade urbana, foram encontrados diversos fatores que causam impacto na decisão entre o tipo de transporte a escolher. Após 49 critérios encontrados e a impossibilidade de todos entrarem no modelo inteligente a desenvolver torna-se crucial perceber quais os critérios que efetivamente tem maior impacto e quais os mais importantes para a população (secção 4.1.1). |
| **Levantamento** | Quem, o que, quantos, quanto. | De forma, a responder à justificação do experimento, torna-se necessário arranjar uma estratégia para priorizar esses fatores. Deste modo, o desenvolvimento de um formulário online foi a solução mais prática e eficaz de forma a abranger uma maior visibilidade a nível nacional. O formulário foi publicado nas redes socias, assim como, foram enviados emails para diversas universidades do país (secção 4.1.1). |
| **Análise de requisitos** | Quem, o que, quantos, quanto. | Após o término da data-limite do formulário, foram analisas todas as 211 respostas obtidas, através de análises estatísticas, de forma a perceber quais e quantos critérios os cidadãos tomam em consideração aquando da tomada de decisão (secção 4.1.4). |
| **Pesquisa Histórica** | Como, porquê. | Toda a pesquisa histórica, tanto da mobilidade urbana como dos critérios identificados, seguiram a estratégia de pesquisa mencionada na secção 2.1 de forma a obter os artigos mais relevantes para o caso de estudo em questão. |
| **Estudo de caso** | Como, porquê. | O estudo de caso, inicialmente, e através do formulário dirigiu-se a toda a população portuguesa (incluindo diversas cidades e ilhas de Portugal), contudo, o modelo a desenvolver irá incorporar apenas dados de Vila Nova de Famalicão. Após a análise dos resultados obtidos, foram identificados os critérios mais relevantes e consequentemente aqueles que irão incorporar o modelo a desenvolver (secção 4.1.4). |

### Obtenção de dados

Conforme referido anteriormente, Vila Nova de Famalicão é a cidade piloto do projeto *ioCity* enquadrando-se no desenvolvimento da *Smart City* que visa desenvolver uma cidade mais inclusiva e inovadora.

Deste modo, Famalicão assenta em seis eixos fundamentais (Pessoas, ambiente/energia, governança, mobilidade, modos de vida e economia) para desenvolver uma cidade inteligente que promova a diversidade, igualdade de oportunidades e participação cívica. Sendo que todas as ações a desenvolver devem ir de encontro à agenda urbana de 2030 definida pela ONU.

O projeto *b-smart*[[4]](#footnote-4), é um projeto que visa o desenvolvimento de uma plataforma de inteligência urbana, que aglomera diversas ferramentas inteligentes presentes na gestão do município, e foi proposto no Portugal *Smart Cities* no ano de 2021.

Posto isto, senda esta uma cidade com alguns projetos interessantes nesta área e tendo o objetivo de crescer, uma das formas de obtenção de dados, para os critérios acima mencionados (Tabela 10 e Tabela 11), baseia-se em dados da cidade fornecidos pela Câmara Municipal de Vila Nova de Famalicão (CMVNF) **(Passos, 2021).** Visto que, todos os fatores mencionados vão de encontro aos objetivos da cidade, o desenvolvimento da ioCity é uma mais-valia para a construção de um território mais inteligente e eficiente. Pelo que, um dos fornecedores de dados será a CMVNF.

Numa segunda fase, o projeto ioCity tem o objetivo de recolher os seus próprios dados. Dados relativos aos transportes, como por exemplo, a localização dos transportes públicos, taxa de ocupação, entre outros, e dados relativos aos parques de estacionamento, como, localização e ocupação. Outra fonte de dados é Google Maps, uma vez que, possui dados interessantes relativos às cidades e que podem ser úteis no âmbito do projeto.

### Estudo do formulário

Para determinar o grau de importância foi elaborado um questionário com os critérios pré-definidos para que cada cidadão votasse, numa escala de 1 a 5, o número que cada critério apresenta na sua tomada de decisão. Posto isto, e para avaliar os resultados obtidos de forma eficaz, foi utilizado o programa *Past*, programa de software para análise de dados estatísticos denominados por *Paleontological Statistics*.

Primeiramente, e após a exportação do Excel com os resultados, foi feita uma primeira análise global às respostas dadas pelos inquiridos, exceto aquelas que envolviam respostas com texto e intervalo de valores, nomeadamente, localização geográfica, idade e preferência de transporte.

A Figura 5 apresenta uma tabela com diversas estatísticas relacionadas com as respostas dadas por cada inquerido a cada pergunta, sendo que as perguntas são representadas entre P1 E P49. Na figura é possível observar o número total de respostas a cada pergunta (N), o valor mínimo (min) e máximo (max), a soma do valor das respostas (sum), a média (mean), o erro padrão (std. Error), a variância (variance), o desvio padrão (stand. dev), a mediana (median), o primeiro quartil (25 prcentil), terceiro quartil (75 prcentil) e o coeficiente de variância (coeff. var).

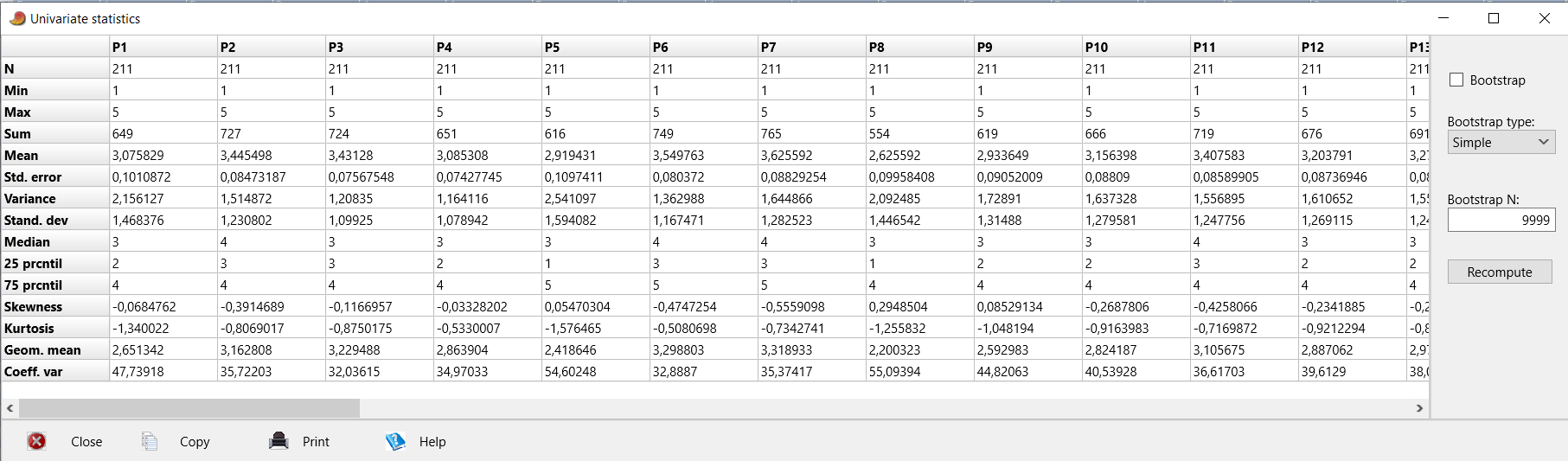


Figura 5 - Resultados Estatísticos das respostas do formulário.

É possível constatar que houve 211 inquiridos que participaram ativamente no formulário e como era de esperar a resposta mínima possível é de 1 e a máxima de 5. A média de respostas encontra-se por volta 3,170 valores e o desvio padrão 0,444, sendo apresentados valores bastante dispersos para cada pergunta.

Na Figura 6 é possível visualizar graficamente estatísticas, como a média, moda, mediana, das variáveis estudadas. É importante referir, que quanto mais próximo do zero for este valor mais coerência de respostas existe e consequentemente menos variância. Pelo que, é possível constatar que as pessoas dispersaram nas suas opiniões.

Chart, line chart

Description automatically generated

Figura 6 - Estatísticas (média, moda, mediana) das variáveis estudadas.

Relativamente à idade, na Figura 7, podemos visualizar os dados respetivos, concluindo que a maior parte da amostra se concentra entre os 18 e 24 anos de idade (63% da amostra total).

Graphical user interface, chart, application

Description automatically generated

Figura 7 - Dados Estatísticos relativamente à Idade.

Quanto à localização geográfica, podemos observar, na Figura 8, que a maior parte das respostas são de população do Norte, ocupando 61% da amostra total.

Graphical user interface, application

Description automatically generated

Figura 8 - Dados Estatísticos relativamente à Localização Geográfica.

Relativamente, à preferência entre transportes públicos ou privados, é preciso ter em consideração que a pergunta apenas foi colocada após o questionário ter sido lançado, tendo uma amostra de 76 pessoas (36% da amostra total). Contudo, averiguou-se que 71% prefere transporte individual (Figura 9).

Chart, pie chart

Description automatically generated

Figura 9 - Dados Estatísticos relativamente à Preferência de Transporte.

Seguidamente, foi feita uma análise para determinar o nível de correlação, nível de concordância de cada pergunta, através do coeficiente de correlação não-paramétrico *Kendall’s tau*. O coeficiente de correlação deve apresentar resultados num intervalo de [-1 e 1]. Caso o resultado seja mais próximo de 1 significa que existe um nível de concordância entre os inqueridos, e, portanto, uma variância menor.

Os resultados do coeficiente de correlação *Kendall’s tau* serão apresentados na imagem seguinte (Figura 10).

Table

Description automatically generated

Figura 10 - Nível de Correlação.

Posteriormente, foi elaborada a média das correlações obtidas para cada pergunta e, por sua vez, determinado o grau de importância de cada critério associado a pergunta. Sendo que, as correlações mais perto de 1 são as que apresentam maior nível de concordância, e, portanto, maior grau de importância.

Posto isto, o grau de importância 5 foi definido pelos critérios que apresentam média maior que 3 e um nível de correlação maior que 0,19. Na Tabela 13, é apresentada uma tabela com os critérios associados à média e à correlação obtida, e, posteriormente, o grau de importância.

Como era de esperar as respostas divergiram bastante, e portante, o nível de correlação é baixo. Apenas os critérios selecionados tem associado grau de importância 5, todos os outros não têm qualquer valor associado pois não cumprem os requisitos acima mencionados. Consequentemente, e depois de selecionados os critérios, foi identificado para cada um a estratégia de obtenção dos respetivos dados.

No gráfico a seguir é possível visualizar de forma mais objetiva a média de correlações e média de respostas por pergunta (Figura 11). Como podemos observar para o critério número 3 (percentagem de lotação) a média de correlações é 0,228 e média de respostas é 3,431.

Chart, line chart

Description automatically generated

Figura 11 - Resultados Estatísticos.

É de ressaltar que as perguntas percentagem de Lotação, fluxo de passageiros percentagens de parques públicos, tempo de procura de estacionamento, percentagem de pontualidade, percentagem tecnológica, percentagem de pagamentos com sistemas eletrónicos, percentagem de controlo de segurança e localização geográfica são as que apresentam um nível mais elevado, tendo média de correlação acima dos 0,19 e média acima de 3 valores (Tabela 13).

Tabela 13 - Grau de Importância dos critérios identificados

| **Nº** | **Critério** | **Média Correlação** | **Média do valor de respostas** | **Grau** |
| --- | --- | --- | --- | --- |
| **1** | Distância média entre a localização geográfica do passageiro e o transporte | 0,037 | 3,076 | - |
| **2** | Capacidade total do veículo | 0,187 | 3,445 | - |
| **3** | Percentagem de Lotação | 0,228 | 3,431 | 5 |
| **4** | Percentagem de Fluxo de passageiros | 0,229 | 3,085 | 5 |
| **5** | Número de transferências entre meios de transporte | 0,011 | 2,919 | - |
| **6** | Número de paragens intermediárias (viagens de longa distância) | 0,082 | 3,550 | - |
| **7** | Custo do bilhete | 0,024 | 3,626 | - |
| **8** | Custo de carga | 0,055 | 2,626 | - |
| **9** | Percentagem de versatilidade | 0,045 | 2,934 | - |
| **10** | N´º total dos diferentes tipos de transporte | 0,060 | 3,156 | - |
| **11** | Percentagem da frequência dos transportes | 0,066 | 3,408 | - |
| **12** | Percentagem de informações disponibilizadas | 0,087 | 3,204 | - |
| **13** | Tempo desde a localização geográfica do passageiro até à paragem | 0,075 | 3,275 | - |
| **14** | Tempo de viagem (duração) | 0,083 | 3,626 | - |
| **15** | Tempo de espera | 0,088 | 3,403 | - |
| **16** | Tempo de transferência | 0,088 | 3,095 | - |
| **17** | Tempo total do percurso | 0,096 | 3,668 | - |
| **18** | Taxa de Pontualidade | 0,089 | 3,777 | - |
| **19** | Percentagem de dispositivos eletrónicos | 0,109 | 3,028 | - |
| **20** | Percentagem de Conforto físico | 0,115 | 3,294 | - |
| **21** | Percentagem de condições da Paragem | 0,122 | 3,209 | - |
| **22** | Percentagem tecnológica | 0,122 | 2,308 | - |
| **23** | Percentagem de acidentes | 0,144 | 2,251 | - |
| **24** | Percentagem de avarias | 0,159 | 2,450 | - |
| **25** | Percentagens de roubo | 0,191 | 2,597 | - |
| **26** | Percentagens de condições | 0,140 | 3,171 | - |
| **27** | Percentagens de veículos Híbridos | 0,166 | 2,431 | - |
| **28** | Percentagens de veículos a gasolina e gasóleo | 0,053 | 3,332 | - |
| **29** | Percentagens de poluição | 0,121 | 3,441 | - |
| **30** | Custo do Combustível | 0,153 | 4,275 | - |
| **31** | Distância média entre o estacionamento do veículo e o local desejado | 0,165 | 3,360 | - |
| **32** | Taxa de disponibilidade dos parques de estacionamento | 0,161 | 3,389 | - |
| **33** | Percentagem de parques públicos | 0,208 | 3,270 | 5 |
| **34** | Preço médio do parque | 0,188 | 3,806 | - |
| **35** | Tempo total do percurso | 0,169 | 3,213 | - |
| **36** | Tempo de procura de estacionamento | 0,219 | 3,545 | 5 |
| **37** | Percentagem de Pontualidade | 0,223 | 3,265 | 5 |
| **38** | Percentagem de acidentes | 0,168 | 2,839 | - |
| **39** | Percentagem de avarias | 0,196 | 2,815 | - |
| **40** | Percentagem de roubo | 0,251 | 2,768 | - |
| **41** | Percentagem de poluição | 0,169 | 3,455 | - |
| **42** | Percentagem de Conforto físico | 0,123 | 3,735 | - |
| **43** | Percentagem de condições do parque de estacionamento | 0,165 | 3,379 | - |
| **44** | Percentagem tecnológica | 0,201 | 3,005 | 5 |
| **45** | Percentagem de pagamentos com sistemas eletrónicos | 0,220 | 3,261 | 5 |
| **46** | Parques com lavagem do carro | 0,032 | 2,066 | - |
| **47** | Parques com carregamentos para carros elétricos | 0,156 | 2,517 | - |
| **48** | Percentagem de controlo de segurança | 0,240 | 3,137 | 5 |
| **49** | Localização geográfica | 0,239 | 3,408 | 5 |

### Clustering

Foi desenvolvido um método de *clustering* com o objetivo de identificar os perfis associados à posição das pessoas em relação a tomada de decisão entre os transportes públicos e privados.

O clustering é uma técnica que tem como objetivo dividir os dados em diferentes grupos. Sendo que em cada grupo são identificados um conjunto de padrões idênticos nos dados. Existem vários algoritmos para realizar esta agregação **(Patil, 2021)**. Considerando a amostra e a quantidade de dados (resultados do estudo descrito na subsecção 4.1.4), uma das técnicas de agregação mais eficientes é o *K-Means*. No entanto, esta técnica só funciona para dados numéricos, uma vez que uma das medidas utilizadas para o agrupamento é a média. O método *K-Modes* alarga esta técnica utilizada para agrupar variáveis categóricas. Através de diferenças entre os pontos de dados, os modos são criados. Quanto menores forem as diferenças, mais semelhante será o conjunto de dados **(Aprilliant, 2021)**.

O primeiro passo é limpar os dados e obter um conjunto de Ks em várias divisões, e os dados estarão de acordo com a distância de cada ponto de dados até ao centroide. Deste modo, o valor de K refere-se ao número de centroides necessários perante o conjunto de dados, ou seja, um centroide é um protótipo para o agrupamento **(Alves, 2018)**. O método do cotovelo foi selecionado para determinar o número ótimo de clusters a partir dos dados. Calculando a soma das diferenças entre clusters (custo), foi definido o valor ideal de K. Para o efeito, os dados foram iterados, e gerados *clusters* para diferentes números de k, começando em um e terminando em nove. Depois, foi gerado um gráfico que traça o custo de cada valor de K. O método de Cotovelo apresenta o valor ideal de K no ponto em que o valor não diminui significativamente com a adição do valor de K. O passo seguinte foi gerar o objeto final de agrupamento e obter as etiquetas de agrupamento para todos os registos. Consequentemente, é essencial resumir todas as características dos agrupamentos e determinar os perfis de cada agrupamento **(Patil, 2021)**.

Em suma, através do desenvolvimento de *Clustering* foram determinados perfis, ou seja, identificados grupos idênticos de objetos dentro do mesmo grupo de dados. É um processo exploratório que auxilia a descoberta de relações entre os dados.

#### Cenários

Para proceder ao desenvolvimento desta prática foram selecionadas técnicas de modelização e utilizado o método de *K-Modes* uma vez que a amostra analisada apenas tem variáveis categóricas (descrito na subsecção 4.1.4).

Cenários são subconjuntos dos dados totais. Os dados obtidos têm duas áreas diferentes (transporte privado e público). Assim, foram criados três cenários para interpretar os resultados e investigar diferenças significativas (Tabela 14). O Cenário A foi estudado com o conjunto total de dados. O cenário B apenas com os dados relativos ao transporte individual. E o cenário C com as variáveis relativas ao transporte público. Consequentemente, o modelo foi construído com a biblioteca panda e posteriormente foram interpretados os resultados.

Tabela 14 - Cenários Clustering

|  |  |  |
| --- | --- | --- |
| **Cenários** | **Variáveis** | **Técnica** |
| **A** | Todas | K-Modes |
| **B** | Variáveis do Transporte Individual | K-Modes |
| **C** | Variáveis do Transporte Público | K-Modes |

#### Resultados

O algoritmo *K-Modes Clustering* identificou cinco perfis para os três cenários diferentes através do método *Elbow* que retorna o valor ideal de K no ponto em que o valor não diminui significativamente com a adição de K (subsecção 4.1.5). A Figura 12 mostra visualmente as diferenças entre cada cenário e agrupamento. Todos os cenários têm cinco cores, o que significa que têm cinco grupos diferentes identificados, e cada cor representa um agrupamento.

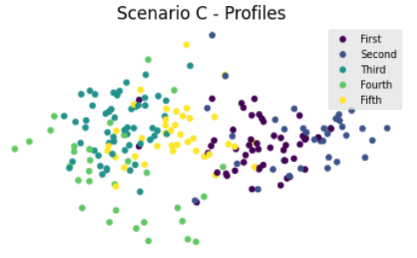
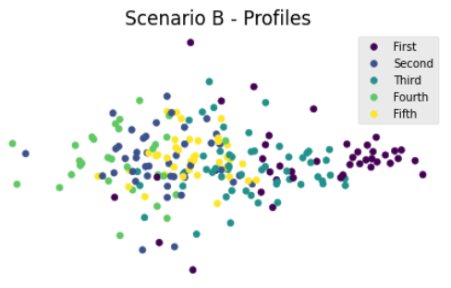
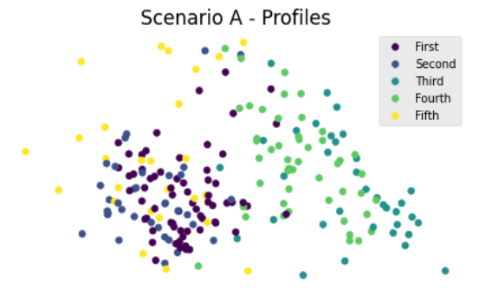


Figura 12 - Resultados visuais do Clustering.

Através dos cenários e modelos contruídos foram obtidos os seguintes resultados (Tabela 15, Tabela 16 e Tabela 17). Estes resultados, para cada cluster indicam um certo grau de importância. Para o cenário A, o cluster 1 tem 37 questões com um grau de importância 3. Quando o resultado é NaN, não há nenhuma questão com um grau de importância associado. Para o cenário B, o cluster 5 tem um valor NaN para o grau importante de 2 e apenas a variável custo do combustível tem um grau essencial de 5, e a percentagem de conforto físico tem um grau de 4. As restantes variáveis têm todas um grau de 3.

Tabela 15 - Resultados do cenário A com o algoritmo K-Modes

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Importância** | **Cluster 1** | **Cluster 2** | **Cluster 3** | **Cluster 4** | **Cluster 5** |
| **0** | 2 | 2 | 2 | 2 | 2 |
| **1** | 6 | 15 | 2 | 2 | 29 |
| **2** | 4 | 10 | NaN | 1 | 3 |
| **3** | 37 | 14 | 4 | 5 | 3 |
| **4** | 2 | 9 | NaN | 37 | 1 |
| **5** | 1 | 2 | 44 | 5 | 14 |

Tabela 16 - Resultados do cenário B com o algoritmo K-Modes

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Importância** | **Cluster 1** | **Cluster 2** | **Cluster 3** | **Cluster 4** | **Cluster 5** |
| **0** | 2 | 2 | 2 | 2 | 2 |
| **1** | 3 | 3 | 1 | 11 | 1 |
| **2** | NaN | 5 | 1 | 2 | NaN |
| **3** | NaN | 11 | 1 | 4 | 18 |
| **4** | NaN | 2 | 16 | 1 | 1 |
| **5** | 18 | NaN | 2 | 3 | 1 |

Tabela 17 - Resultados do cenário C com o algoritmo K-Modes

| **Importância** | **Cluster 1** | **Cluster 2** | **Cluster 3** | **Cluster 4** | **Cluster 5** |
| --- | --- | --- | --- | --- | --- |
| **0** | 2 | 2 | 2 | 2 | 2 |
| **1** | 1 | 1 | 9 | 18 | 2 |
| **2** | 4 | 1 | 4 | NaN | 3 |
| **3** | 4 | 1 | 13 | 1 | 22 |
| **4** | 17 | 2 | 4 | 1 | 3 |
| **5** | 4 | 25 | NaN | 10 | NaN |

##### Análise dos Resultados do Cenário A

Analisando os resultados do cluster (Tabela 15) é possível perceber de imediato que a variável idade, localização e preferência são constantes em cada cluster, não havendo alterações de valor. Sendo que a idade toma apenas valores de 1 (entre os 18-24), localização valores de 0 (Norte) e preferência valores de 0 também (NaN).

Este fenómeno deve-se ao facto de a distribuição de resultados ser assimétrica. Ou seja, de entre 211 respostas, 132 correspondem a pessoas de idade entre os 18-24, 128 da região norte e 135 são respostas sem valor para a variável preferência. Esta última variável foi adicionada ao formulário depois do mesmo ser criado e publicado para o público, por isso, apresenta 63,67 de valores nulos. Tanto a idade como a localização representam mais de metade da percentagem total da amostra, 63% e 61% respetivamente. Estes valores podem ser comprovados através das análises estatísticas feitas inicialmente (subsecção 4.1.4).

Concluindo, o impacto que estas representam no cluster é mais forte, e, portanto, mais significativo, não havendo dispersão dos valores e assumindo as características mais relevantes na análise. Relativamente, às outras faixas etárias, localização e preferências são padrões mais difíceis de encontrar diante da amostra disponível e, portanto, difícil de identificar classes associadas.

Depois de analisados os grupos encontrados pelo algoritmo *Kmodes Clustering*, foi possível identificar perfis da importância que cada inquirido dá a cada variável. Deste modo, são listados na Tabela 18 os perfis encontrados. Sendo que, para a região do norte e idade entre 18-24 anos foram identificados 5 perfis: Importância média, importância dispersa, alta importância, importância considerável e importância extrema.

É de realçar, que são discutidas 52 variáveis, contudo, na análise dos perfis irá ser assumido o valor de 49 variáveis devido ao facto de a idade, localização geográfica e preferência entre transportes públicos ou veículos privados serem constantes ao longo dos 5 perfis encontrados. É importante referir, que apesar da classificação de diversos perfis, em todos os clusters o custo de combustível apresenta um nível de importância 5.

Tabela 18 - Perfis do Cenário A

| **Perfil** | **Características Principais** | **Conclusão** |
| --- | --- | --- |
| **Cluster 1**  **Importância Média** | - Valor de importância 3 é predominante no cluster;  - Apenas a variável custo de combustível apresenta valor 5 de importância;  - Capacidade total e percentagem de ocupação são as duas variáveis que apresentam valor 4;  - Valores entre 1 e 2 são pouco significativos. | Estamos perante um cluster de impacto médio, ou seja, praticamente nenhuma variável (à exceção do custo de combustível) apresenta um grau de importância alto nem baixo. Sendo que todas as variáveis têm uma importância central. |
| **Cluster 2**  **Importância Dispersa** | - Valores de importância dispersos, maioritariamente, entre 1, 2 e 3;  - Variáveis percentagens de veículos a gasolina e gasóleo e custo de combustível apresentam valor 5 de importância.  - Existem 9 variáveis que tomam o valor de 4, sendo que as restantes 38 ocupam valores de 1,2 e 3. | De acordo, com os dados obtidos deste cluster, podemos afirmar que as variáveis apresentadas apresentam um grau de importância médio/baixo. Sendo que num intervalo de 49 variáveis apenas 2 apresentam um valor igual a 5, e 9 o valor de importância 4. |
| **Cluster 3**  **Alta importância** | - Respostas concentradas no valor 5, não existindo nem valores de importância 2 nem de 4;  - Apenas a variável parques com lavagem de carro apresenta um valor de importância um.  - De entre, todas as variáveis, só 4 é que apresentam valores de 3, nomeadamente, percentagem de fluxo de passageiros, percentagem de carros híbridos, percentagem de carros a gasolina e gasóleo e percentagem de tecnologia. | Perfil caracterizado por um alto extremo, onde 44/49 das variáveis estudadas apresentam um valor máximo de importância (5). |
| **Cluster 4**  **Importância considerável** | - Valor de importância 4 é predominante no cluster;  - Existem 5 variáveis que apresentam importância 5, nomeadamente, distância média entre a localização geográfica do passageiro e o transporte, custo do bilhete, percentagem de pontualidade, custo do combustível e preço do parque;  - Assim como, no grau de importância 3, também 5 variáveis tomam esse valor. | Concluindo, o perfil 4, é determinado por um grau médio/alto de importância, onde a maior parte das variáveis apresentam valor 4 (37/49). Grupo ponderado, no que toca, a respostas de valor máximo de importância. Apenas 5 ocupam esse valor. |
| **Cluster 5**  **Importância Extrema** | - Concentração de valores, variam entre o valor de 1 e valor de importância 5.  - De entre, o número total de variáveis analisadas, 29 das variáveis tomam o valor de 1 e 14 das variáveis tomam o valor de 5;  - Deste modo, apenas 9 variáveis apresentam um valor entre 2,3 e 4. | Perfil de extremos. Analisando, o cluster em questão, é possível observar que não existe meio termo. Ou a variável toma valores máximos ou mínimos. |

##### Análise dos Resultados do Cenário B

Nesta análise (Tabela 16), apenas foram consideradas as variáveis que tem impacto nos parques de estacionamento, sendo a amostra de 20 variáveis mais 3 variáveis gerais (idade, localização geográfica e preferência de transporte). Tal como, na análise geral, realizada anteriormente, o mesmo acontece neste cluster, onde as 3 variáveis gerais são constantes ao longo dos perfis. Isto deve-se à amostra de respostas ser mais incidente nos valores 18-24, norte e NaN correspondente á idade, localização e preferência, respetivamente. Uma explicação mais aprofundada deste fenómeno foi descrita anteriormente na subsecção Análise dos Resultados do Cenário A na página 65.

Deste modo, foram encontrados 5 perfis: Importância alta, Importância Central, Importância Considerável, Importância baixa e Importância Média (Tabela 19).

É de realçar, que a amostra de variáveis consideravas é de 20 e não de 23, uma vez que, a idade, localização geográfica e preferência entre transportes públicos ou veículos privados são constantes ao longo dos 5 perfis encontrados. O perfil 2 e 5 são bastante similares, apresentando uma amostra de variáveis, maioritariamente, com um impacto médio.

Tabela 19 - Perfis do Cenário B

| **Perfil** | **Características Principais** | **Conclusão** |
| --- | --- | --- |
| **Cluster 1**  **Importância alta** | - Valor de importância 5 é predominante no cluster;  - Não existindo valores de 2,3 e 4 associados a variáveis;  - Duas variáveis tomam valor de 1, nomeadamente, percentagem de lavagem dos veículos e percentagem de parques com carregamento para carros elétricos. | Perfil caracterizado por um alto extremo, onde 18/20 das variáveis estudadas apresentam um valor máximo de importância (5). |
| **Cluster 2**  **Importância Central** | - Respostas concentradas no valor de importância 3 (11/20), não existindo valores de importância 5 associados a alguma variável;  - Apenas as variáveis: preço do parque e custo do combustível apresentam um grau de importância 4 (2/20); | Estamos perante um cluster de impacto médio, ou seja, praticamente nenhuma variável (à exceção do custo de combustível e preço do parque) apresenta um grau de importância alto. Sendo que, mais de metade da amostra considerada apresenta uma importância central. |
| **Cluster 3**  **Importância Considerável** | - Valor de importância 4 é predominante no cluster;  - Existem 2 variáveis que apresentam importância 5, nomeadamente, custo do combustível e preço do parque;  - A variável percentagem de carregamentos para carros elétricos é a única que toma o valor de importância 3 e a percentagem de lavagem dos veículos a única que toma valores de 2. | Concluindo, o perfil 3, é determinado por um grau médio/alto de importância, onde a maior parte das variáveis apresentam valor 4 (16/20). Grupo ponderado, no que toca, a respostas de valor máximo de importância. Apenas 2 ocupam esse valor. |
| **Cluster 4**  **Importância baixa** | - Respostas concentradas no valor de importância 1;  - Apenas 4 variáveis tomam o valor de 3, 1 variável o valor de 1 e 3 variáveis o valor de 5;  - De entre, todas as variáveis, só 3 é que apresentam valores de 5, nomeadamente, custo do combustível, preço do parque e percentagem de conforto. | Perfil caracterizado por maioritariamente valores baixos, sendo que metade da amostra considera as variáveis com um grau de importância mínimo (valor de 1). Realce de três variáveis que tomam o valor de alta importância. |
| **Cluster 5**  **Importância Média** | - Bastante similar ao segundo cluster;  - Respostas concentradas no valor de importância 3, não existindo valores de importância 2 associados a alguma variável;  - Apenas a variável custo do combustível apresenta um grau de importância 5;  - A única variável que toma valores de 4 é a percentagem de conforto do parque. | Estamos perante um cluster de impacto médio, ou seja, nenhuma variável (à exceção do custo de combustível) apresenta um grau de importância alto (valor 5). Sendo que a maior parte das variáveis têm uma importância central (18/20). Relativamente ao cluster 2, é possível afirmar que apesar de a maior parte, em ambos os agrupamentos, serem de classe média, este cluster apresenta uma dispersão menor relativamente a outros valores de importância. |

##### Análise dos Resultados do Cenário C

Nesta análise (Tabela 17), apenas foram consideradas as variáveis que tem impacto nos transportes públicos, sendo a amostra de 29 variáveis mais 3 variáveis gerais (idade, localização geográfica e preferência de transporte). Tal como, na análise geral e na dos parques de estacionamento, realizada anteriormente, o mesmo acontece neste cluster, onde as 3 variáveis gerais são constantes ao longo dos perfis. Isto deve-se à amostra de respostas ser mais incidente nos valores 18-24, norte e NaN correspondente à idade, localização e preferência, respetivamente (figura 10). Uma explicação mais aprofundada deste fenómeno foi descrita na subsecção Análise dos Resultados do Cenário A na página 65.

Deste modo, foram encontrados 5 perfis: Importância alta, Importância Central, Importância Considerável, Importância baixa e Importância Média (Tabela 20).

É de realçar, que a amostra de variáveis consideravas é de 29 e não de 32, uma vez que, a idade, localização geográfica e preferência entre transportes públicos ou veículos privados são constantes ao longo dos 5 perfis encontrados.

Tabela 20 - Perfis do Cenário C

| **Perfil** | **Características Principais** | **Conclusão** |
| --- | --- | --- |
| **Cluster 1**  **Importância Considerável** | - Valor de importância 4 é predominante no cluster;  - Existem 4 variáveis que apresentam importância 5, nomeadamente, Distância média entre a localização geográfica do passageiro e o transporte, preço do bilhete, taxa de pontualidade e número de transferências entre meios de transporte;  - Tanto o grau de importância 2 e 3 tem 4 variáveis associadas. | Concluindo, o perfil 1, é determinado por um grau médio/alto de importância, onde a maior parte das variáveis apresentam valor 4 (17/29). Grupo ponderado, no que toca, a respostas de valor máximo de importância. Apenas 4 ocupam esse valor. |
| **Cluster 2**  **Importância alta** | - Valor de importância 5 é predominante no cluster;  - Duas variáveis tomam valor de 4, percentagem de carros híbridos e percentagens de veículos a gasolina e gasóleo  - Percentagem de Fluxo de passageiros tem um grau de importância de 2, sendo a única variável com este valor e percentagem de tecnologia a única com valor de 3. | Perfil caracterizado por um alto extremo, onde 25/29 das variáveis estudadas apresentam um valor máximo de importância (5). |
| **Cluster 3**  **Importância Moderada** | - Respostas concentradas no valor de importância 3 e 1, não existindo valores de importância 5 associados a alguma variável;  - Apenas as variáveis: capacidade total, Percentagem da frequência dos transportes, taxa de pontualidade e percentagens de veículos a gasolina e gasóleo apresentam um grau de importância 4 (4/29); | Perfil disperso e seletivo. Analisando, o cluster em questão, é possível observar que temos um valor de 13/29 com valores médios de importância e 8/29 de valor 1. Não existindo valores altos e apenas 4 variáveis com grau de importância 4. |
| **Cluster 4**  **Importância baixa** | - Respostas concentradas no valor de importância 1;  - Apenas 1 variável toma o valor de grau de importância de 3 e 4;  - Nenhuma variável toma o valor de 2;  - Na amostra 10 variáveis apresentam valores de 5, nomeadamente, capacidade total, percentagem de ocupação, número de transferências entre meios de transporte, preço do bilhete, tempo desde a localização geográfica do passageiro até à paragem, duração, tempo de espera, tempo de viagem, taxa de pontualidade e percentagens de veículos a gasolina e gasóleo | Perfil caracterizado por maioritariamente valores baixos, sendo que mais de metade da amostra considera as variáveis com um grau de importância mínimo, valor 1 (17/29). Apenas duas variáveis tomam valores intermédios de 3 e 4.  Com grau de importância 5 são destacadas 10 variáveis. Concluindo, perfil caracterizado por extremos. |
| **Cluster 5**  **Importância Média** | - Respostas concentradas no valor de importância 3, não existindo valores de importância 5 associados a alguma variável;  - Apenas 3 variáveis tem um grau de importância 4, tais como, capacidade total, percentagem de ocupação e número de transferências entre meios de transporte. | Estamos perante um cluster de impacto médio, ou seja, nenhuma variável apresenta um grau de importância alto (valor 5). Sendo que a maior parte das variáveis têm uma importância central (22/29). |

#### Conclusões da análise do Clustering

Inicialmente foi realizado ume estudo estatístico com o objetivo de conhecer em detalhe a amostra de dados. Posteriormente, através do algoritmo *kmodes Clustering* foram identificados perfis para três cenários distintos: uma análise global aos dados (cenário A), uma análise apenas dos dados relativos aos parques de estacionamento (cenário B) e, por fim, uma análise somente dos dados dos transportes públicos (cenário C).

Analisando os resultados do cluster é possível perceber de imediato, que para os três cenários estudados, a variável idade, localização e preferência são constantes em cada cluster de cada nível, não havendo alterações de valor. Sendo que a idade toma apenas valores de 1 (entre os 18-24), localização valores de 0 (Norte) e preferência valores de 0 também (NaN).

Este fenómeno deve-se ao facto de a distribuição de resultados ser assimétrica. Ou seja, de entre 211 respostas, 132 correspondem a pessoas de idade entre os 18-24, 128 da região norte e 135 são respostas sem valor para a variável preferência. Esta última variável foi adicionada ao formulário depois do mesmo ser criado e publicado para o público, por isso, apresenta 63,67 de valores nulos. Tanto a idade como a localização representam mais de metade da percentagem total da amostra, 63% e 61% respetivamente. Estes valores podem ser comprovados através das análises estatísticas feitas inicialmente (subsecção 4.1.4).

Concluindo, o impacto que estas representam no cluster é mais forte, e, portanto, mais significativo, não havendo dispersão dos valores e assumindo as características mais relevantes na análise. Relativamente, às outras faixas etárias, localização e preferências são padrões mais difíceis de encontrar diante da amostra disponível e, portanto, difícil de identificar classes associadas.

Depois de analisados os grupos encontrados pelo algoritmo *Kmodes Clustering*, foi possível identificar perfis da importância que cada inquirido dá a cada variável. Sendo que, para a região do norte e idade entre 18-24 anos foram identificados 5 perfis para cada nível.

É possível concluir que no cenário A, apesar da classificação de diversos perfis, em todos os *clusters* identificados, o custo de combustível apresenta um nível de importância 5.

No cenário B, as variáveis: preço médio do parque e custo do combustível, apresentam em todos os clusters, exceto no quinto, valores de alta importância (cluster 1,3 e 4 tomam valores de 5, e no cluster 2 tomam valores de 4). No *cluster* 5, o custo do combustível é a única variável com um grau de importância 5. Desta forma, é demonstrado o impacto e a importância que que as mesmas têm em grupos diferentes.

Para o cenário C as variáveis: capacidade total, percentagem de ocupação e número de transferências entre meios de transporte ocupam um grau de importância alto nos *cluster* 2,4 e 5. Sendo que, o número de transferências entre meios de transporte no cluster 1 também ocupa grau 5 de importância

Este trabalho demonstra, neste exemplo específico, quais as variáveis que os cidadãos priorizam na tomada de decisões entre o transporte público e o automóvel particular. A dispersão dos valores existentes prova que as pessoas têm a certeza de quais os critérios que consideram mais importantes e aqueles que priorizam na tomada de decisões. Foram identificados 5 *clusters* para três cenários realizados. É essencial mencionar que existem *clusters* extremos e medianos em todos os cenários disponíveis. Por exemplo, pessoas que dão prioridade a cada variável na tomada de decisão e outras pessoas que consideram não-variável importante. No entanto, alguns acreditam que todas as variáveis têm um grau de importância mediana. Outros cidadãos são mais críticos e votam apenas 5 (nível de significância mais elevado) para as variáveis centrais na tomada de decisão. São mais seletivos. É um grupo que seleciona vários níveis de importância para diferentes variáveis.

Para o cenário A, destaca-se o custo do combustível. O preço médio dos parques e o custo do combustível te destaque no cenário B. No cenário C, a capacidade total, a percentagem de ocupação e o número de transferências entre meios de transporte ocupam um elevado grau de importância.

Este estudo é um ponto de partida para a recolha de dados e, consequentemente, para responder à necessidade de otimizar e construir modelos inteligentes que satisfaçam as exigências da população e da cidade. O trabalho aqui descrito foi publicado em artigo e encontra-se na secção 6.

## Exploração de APIs

Nesta subsecção são apresentadas diversas aplicações que poderão ser úteis para a recolha de dados, que posteriormente, irão incorporar no modelo preditivo a desenvolver.

### Meteorologia e Eventos

A meteorologia e a gestão de eventos são fatores importantes que afetam a mobilidade urbana. “O clima afeta quase todas as áreas de nossas vidas – Condições climáticas como neblina, chuva e neve impactam o deslocamento para o trabalho e viagens” **(RapidAPI, 2022).** Assim como, a gestão de eventos é um ponto fulcral na vida de cada cidadão, e ter acesso a uma API que gere os eventos e nos alerte aquando da chegada de um é igualmente importante **(PURKAYASTHA, 2019).** Ou seja, se na tomada de decisão da escolha de um parque de estacionamento tivermos acesso a eventos que estão a ocorrer dentro de uma localização, entre 200m a 1 km de distância do parque de estacionamento, a probabilidade do parque estar praticamente lotado, e consequentemente, isso afetar a tomada de decisão é bastante grande. Assim como, o tempo (sol, chuva, tempestade) influencia a decisão de escolha entre um parque coberto ou um parque ao ar livre. Deste modo, na Tabela 21 e Tabela 22 são apresentados exemplos de APIs de meteorologia e de gestão de Eventos, respetivamente.

Tabela 21 - Exemplos de APIs de Metereologia

| **API Meteorologia** | | | | |
| --- | --- | --- | --- | --- |
| **API** | **Designação** | **Definição** | **Variáveis** | **Preço** |
| ***OpenWeatherMap***  **(owm, 2022)** | Previsão de tempo | Fornece a meteorologia atual, previsões, histórico, estações meteorológicas e alertas meteorológicos.  A API também fornece dados atuais e históricos de poluição do ar. | Dados meteorológicos atualizados;  Previsão do clima para os próximos 30 dias;  Dados de previsão diária;  Dados meteorológicos históricos;  Pesquisar dados meteorológicos; | Grátis |
| ***Weatherbit***  **(Weather, sem data)** | Previsão do tempo e alertas | Através das coordenadas de latitude e longitude, é possível obter dados de previsão de tempo retornados no formato JSON. | Previsão meteorológica dos próximos 5 dias;  Dados meteorológicos atuais de um local;  Previsão de 16 dias;  Previsão de 120 horas;  Alertas de mau tempo. | Paga |
| ***AccuWeather***  **(RapidAPI Staff, 2020)** | Condições climáticas e imagens | *AccuWeather* tem um conjunto de API que indicam a previsão, condições atuais, alertas e imagens. | Acesso à meteorologia do local selecionado nas últimas 24h e 6h;  Previsão meteorológica para o dia;  Retorna dados de condições atuais para as 50, 100 ou 150 principais cidades do mundo, com base na classificação.  Previsão diária por dia, hora e imagens;  Previsão meteorológica das cidades vizinhas à selecionada;  Previsão de ciclones (hora e localização) e lista de todos os ciclones ativos;  Alarme diário de meteorologia; | Paga |

Na próxima tabela (Tabela 22) são apresentadas as diversas API que foram exploradas no contexto de gestão de eventos.

Tabela 22 - Exemplos de APIs de Gestão de Eventos

| **API Eventos** | | | | |
| --- | --- | --- | --- | --- |
| **API** | **Designação** | **Definição** | **Variáveis** | **Preço** |
| ***Eventbrite***  **(*Eventbrite*, 2022; Shyam, 2019)** | Gestão de Eventos | Com esta API é possível criar e remover eventos, gerir classes, visualizar todos os eventos registados aplicação. | Criar e remover eventos;  Criar e remover classes e descontos de ingressos;  Criar perguntas sobre pedidos de ingressos;  Criar e atualizar locais de eventos;  Aceder a todos os eventos por localidade;  Criar relatórios e análises personalizadas. | Grátis  (No entanto limitada a 1000 chamadas por hora) |
| ***Meetup***  **(Shyam, 2019)** | Gestão de Eventos | Bastante similar à API *Eventbrite*,  contudo, a acessibilidade à aplicação é apenas para pessoas inscritas num grupo de interesse comum na app. | Criar eventos;  Localizar eventos;  Excluir eventos;  Atualizar eventos;  Listar participantes;  Listar eventos disponíveis numa localidade especifica. | Grátis |
| ***Events***  **(Shyam, 2019)** | Junção de plataformas de gestão de eventos (como Eventbrite, WordPress, Eventful, Facebook e Meetup). | Dispõem todos os eventos presentes nas plataformas parceiras, como, festivais, desporto, reuniões de negócios, exposições de arte e shows.  A API vem com métodos para pesquisar eventos usando um texto simples, nome de uma cidade ou código de um país. | Pesquisa e retorno de eventos através de uma palavra-chave, cidade ou código; | Grátis |
| ***Graph API***  **(Meta, 2022)** | Gestão de eventos do facebook | Dispõem de todos os eventos colocados no facebook | Pesquisa e retorno de eventos | Grátis |

A solução a utilizar está dependente dos modelos preditivos e variáveis a incluir, deste modo, ainda não foi selecionada nenhuma API de meteorologia e eventos.

### Localização

No sentido de desenvolver um sistema o mais completo possível, é importante incorporar uma API que disponibilize mapas estáticos e dinâmicos, rotas, acidentes, entre outros fatores que podem condicionar o transito e consequentemente influenciar a decisão de escolha no meio de transporte a escolher (Tabela 23).

Tabela 23 - Exemplos de APIs do Google Maps

| **APIs Google Maps** | | | | |
| --- | --- | --- | --- | --- |
| **API** | **Definição** | **Recursos** | **Serviços** | **Preço** |
| ***Maps JavaScript API*** | Criar mapas dinâmicos, interativos e personalizados, obter as localizações e experiências geoespaciais. | Diferentes tipos de mapa (estrada, satélite, híbrido, terreno e mapas personalizados);  Localização por texto;  Criar marcadores padrão ou exibir os pré-definidos;  Personalizar os controles de interface do utilizador que aparecem no mapa (zoom, escolha do tipo de mapa, girar, escala, *fullscreen*, atalhos, etc);  Escrever código que reaja às interações do utilizador e eventos do ciclo de vida (eventos da interface do usuário e/ou os eventos de erro, onde é possível manipular programaticamente);  Crie experiências 3D e 2D ricas no mapa base vetorial com o poder do WebGL;  Janelas de informações;  Possibilidade de desenhar formas no mapa com funções integradas;  Sobreposições personalizadas para exibir dados e imagens e sincronizar com o movimento do mapa;  Exiba GeoJSON e outros tipos de dados no mapa em vários formatos.  Mapas de calor (visualizar a densidade dos dados em pontos geográficos, por exemplo, densidade populacional, temperatura, tráfego, etc) | Rotas;  Distâncias e tempo de viagem para uma matriz de origens e destinos, através de diversas formas de transporte: caminhar, dirigir, andar de bicicleta, transporte público;  Elevação de locais na superfície da terra;  Geocodificação (Converte endereços ou Ids de lugares em coordenadas e vice-versa);  Zoom;  Steet View em 360 graus. | Pago  Contudo, as rovider tem direito a créditos: “As rovider verificadas do Google Cloud podem se inscrever para receber créditos da Plataforma Google Maps.  Os candidatos precisam ter permissão para aceitar o contrato de licença dos serviços principais do Google Maps e a política de créditos da Plataforma Google Maps em nome da organização.” |
| ***Local Context Library, Maps JavaScript API*** | Adicionar um mapa de contexto local a uma página web. O mapa de Contexto Local é um widget de IU simples e flexível. | Acesso aos principais locais de interesse próximos de um local;  Mapa interativo com um painel lateral composto por uma galeria de fotos;  Filtrar a lista de lugares com base no tipo, fornecer instruções de direção e personalizar a aparência visual do *widget*. | Rotas;  Locais de interesse;  Galeria de fotos. |
| ***Transit and Traffic layer in the Maps JavaScript API*** | Mostrar rotas de transportes públicos em um mapa; | Adicione informações de tráfego em tempo real;  Exibir a rede de transporte público de uma cidade; | Rotas de transportes públicos;  Camada de tráfego. |
| ***Places library in the Maps JavaScript API*** | Encontre empresas e outros lugares próximos e/ou pesquisa de lugares por tipo de lugar ou com uma *string* de consulta; | Pesquisa de lugares como estabelecimentos, parques de estacionamento, pontos de interesse, etc;  Preenchimento automático e sugestões de sítios pelas inicias escritas na barra de pesquisa; | Biblioteca de lugares;  Preenchimento automático. |

Desta forma, na Tabela 23 são analisadas as várias opções de APIs que o *google maps* disponibiliza **(*API Picker | Google Maps Platform*, sem data)**.

## Solução

Esta subsecção apresenta uma descrição pormenorizada da solução desenvolvida. De forma a compreender os requisitos e funcionalidades do modelo desenvolvido são apresentadas a arquitetura da solução, a *data warehouse*, o modelo OLAP e, por fim, a *dashboard* desenvolvida.

### Arquitetura da Solução

Como referido na secção 1.1 a presente dissertação insere-se no projeto ioCity que propõem através de uma análise de um conjunto de dados (sensores, taxa de ocupação, tráfego, estacionamentos, entre outros) ajudar o cliente a ter uma informação em tempo real e útil na decisão entre a escolha do método de transporte a escolher. Na Figura 13 é apresentada a arquitetura tecnológica do ioCity, as máquinas, bem como as linguagens e tecnologias a utilizar no desenvolvimento do projeto. A estrutura global do projeto está dividida em seis camadas:

* ***Acquisition*** – Aquisição dos dados por parte dos sensores dos parques, transportes e paragens;
* ***Adaptability*** – Modelo que permite adaptar a solução a outras realidades, tornando o sistema escalável;
* ***Data***– Armazenamento dos dados (*data lakes*) e organização dos mesmos (*data warehouse*);
* ***Knowledge***– Processamento dos dados e criação de modelos inteligentes;
* ***APIs*** – Servidor e mecanismo de integração das várias componentes do projeto;
* **Diagram

  Description automatically generated*Interface*** – Interface *web/mobile* para acesso por parte do utilizador.

Figura 13 - Arquitetura do projeto ioCity.

O contributo da solução apresentada entra na camada *Knowlegde* e *Interface*. Deste modo, na Figura 14, é apresentada a arquitetura e fluxo de todo o desenvolvimento até à solução pretendida: *dashboard* analítica. Relativamente à camada *Knowledge* a partir dos dados obtidos (*Data Sources*) foi iniciado o processo de ETL – Extração, Transformação e Carregamento (*Extract*, *transform* e *Load*) onde foi feita uma limpeza dos dados e consequentemente desenvolvida uma data warehouse para o armazenamento de todos os dados necessários à construção do modelo OLAP. Consequentemente o modelo OLAP foi criado com a ferramenta *cube.js*. Com a API do *cube.js* foi desenvolvida uma *dashboard* analítica que permite a visualização e criação de gráficos interativos que auxiliam o apoio à tomada de decisão. A solução é escalável e os dados são indicativos do que é possível obter.

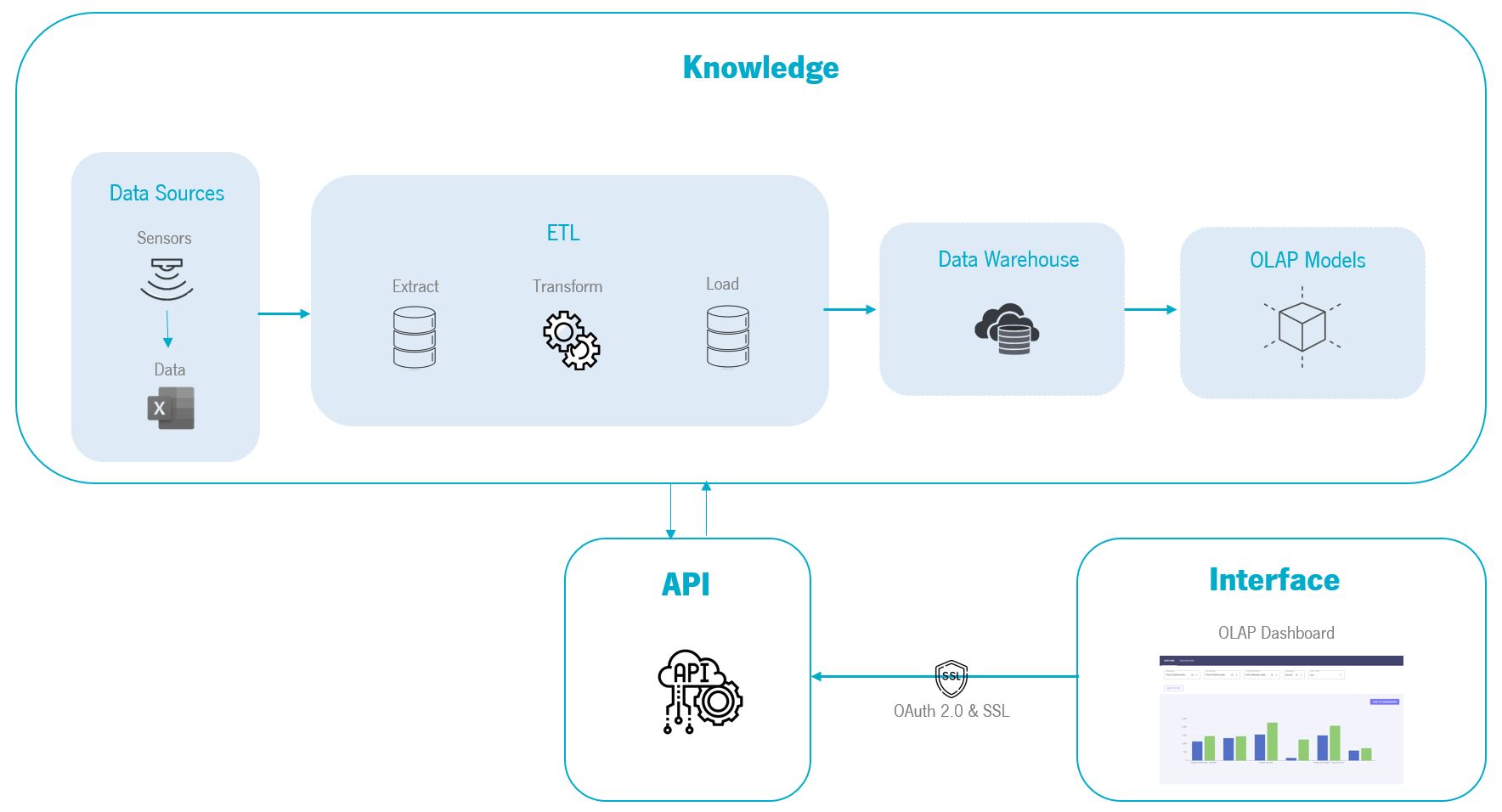


Figura 14 - Arquitetura da Solução.

### Processo de ETL

O processo de ETL (extração, transformação e carregamento) é fundamental para processar os dados e melhorar a performance da solução a desenvolver. Nesta fase, utilizando a linguagem de programação *python* e o *Microsoft Excel*, as colunas que não acrescentavam interesse perante os indicadores definidos foram eliminadas, erros de escrita foram corrigidos e análises estatísticas foram feitas.

#### Descrição dos dados

Após a recolha de dados pela Câmara Municipal de Vila Nova de Famalicão um conjunto de dados foi obtido e analisado de forma a compreender melhor os atributos existentes. Foram obtidos 27625 registos num período de dois meses, entre fevereiro e março. Na Tabela 24 é possível perceber os atributos obtidos, a descrição, o formato e os valores possíveis. Por exemplo, o primeiro atributo, *Parking lot id*, é o número que correspondente ao lote do parque, apresentado com números inteiros e tem valores de 1,2,3,4,5 e 7.

Tabela 24 - Análise dos dados

| **Atributos** | **Descrição** | **Formato** | **Valores Possíveis** |
| --- | --- | --- | --- |
| **id** | Número identificador de registo de dados | Inteiro | array([1, 27625]) |
| ***Parking\_lot\_id*** | O número correspondente ao lote do parque | Inteiro | 1,2,3,4,5 e 7 |
| ***Latitude*** | Coordenadas geográficas | Decimal | 41,40714;  41,40675;  41,40309;  41,41214;  41,4114;  41,4147 |
| ***Longitude*** | Coordenadas geográficas | Decimal | -8,51505;  -8,51413;  -8,51786;  -8,50915;  -8,50781;  -8,51672 |
| ***Name*** | Nome do parque | *String* | Estação Rodoviaria – Entrada;  Estação Rodoviária – CESPU;  Campo da Feira;  Parque da Cidade – Citeve;  Parque da Cidade – Casa Territorio;  Casa das Artes - Inf/coberto |
| ***Displat\_id*** | O número correspondente ao lote do parque | Inteiro | 1,2,3,4,5 e 7 |
| ***Provider\_type*** | Sensor proveniente dos dados | Inteiro | 1 |
| ***Provider\_id*** | O número correspondente ao lote do parque | Inteiro | 1,2,3,4,5 e 7 |
| ***parking\_cluster\_id*** | - | - | Vazio |
| ***parking\_cluster\_name*** | - | - | Vazio |
| ***parking\_cluster\_active*** | - | - | Vazio |
| ***Location*** | A rua onde o parque está localizado | *String* | Rua Henriques Nogueira;  Avenida José Manuel Marques;  Rua Fernando Mesquita;  Rua Farnando Mesquita;  Rua de São Vicente |
| ***Total\_capacity*** | A capacidade total de lugares que o parque tem | Inteiro | 79, 170, 618, 150, 113 e 98 |
| ***Observations*** | As zonas e secções que cada parque tem | *String* | “Zona1, Z1 E6-E | E5-S  Z2 E3-E | E4-S”;  “Zona2 Z2 E3-E | E4-S”;  “Feira Norte: Zona6 Z6 E5-E | E6-S Feira Sul: Zona6 Z6 E3-E | E4-S”;  “Zona3 Z3 E3-E | E4-S”;  “Zona4 Z4 E4-E | E3-S”;  Zona15 – Z15 E5-E | E6-S |
| ***Last Update*** | Hora e dia em que a recolha de dados foi feita | Data | Horas e dias do mês de fevereiro e março de 2022 |
| ***Total\_available\_places*** | Total de lugares disponíveis que o parque tem no momento em que os dados foram recolhidos | Inteiro | array([ 52., 160., 633., …, 3215., -97., 3220.]) |
| ***Total\_occupied*** | Número total de lugares ocupados no parque no momento em que os dados foram recolhidos | Inteiro | array([ 27., 10., -15., …, -3065., 715., -3070.]) |
| ***Total\_places*** | A capacidade total de lugares que o parque tem | Inteiro | 79, 170, 618, 150, 113 e 98 |
| ***Active*** | Se o parque está ou não ativo | *String* | Verdadeiro |
| ***Update\_date*** | Hora e data em que a recolha de dados terminou | Data | Horas e dias do mês de fevereiro e março de 2022 |
| ***no\_communication\_day*** | - | *String* | Falso |
| ***no\_communication\_hour*** | - | *String* | Falso |

#### Análise dos dados

Nesta subsecção são apresentadas duas tabelas, Tabela 25 com as variáveis numéricas e Tabela 26 com as variáveis categóricas, demonstrando os valores estatísticos dos atributos apresentados na subsecção “Descrição dos dados” na página 84.

Na Tabela 25 para cada atributo é possível observar o número total dados (Total), a média, o desvio padrão (std), o valor mínimo (Min), o primeiro quartil (25%), o segundo quartil (50%), o terceiro quartil (75%) e o valor máximo (Máx). Por exemplo, o atributo “*parking\_lot\_id*” tem o valor máximo de 7 e mínimo de 1 e é repetido 27625 vezes no dataset. Os valores apresentados como NaN significa que não tem qualquer valor numérico para análise.

Tabela 25 - Estatísticas variáveis numéricas

| **Atributos Númericos** | **Total** | **Média** | **Std** | **Min** | **25%** | **50%** | **75%** | **Máx** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| ***parking\_lot\_id*** | 27625 | 3.64 | 2.05 | 1 | 2 | 3 | 5 | 7 |
| ***displat\_id*** | 27625 | 3.64 | 2.05 | 1 | 2 | 3 | 5 | 7 |
| ***rovider\_type*** | 27625 | 1 | 0 | 1 | 1 | 1 | 1 | 1 |
| ***rovider\_id*** | 27625 | 3.64 | 2.05 | 1 | 2 | 3 | 5 | 7 |
| ***total\_capacity*** | 27625 | 208.95 | 193.91 | 79 | 98 | 113 | 170 | 618 |
| ***parking\_cluster\_id*** | 0.0 | NaN | NaN | NaN | NaN | NaN | NaN | NaN |
| ***parking\_cluster\_name*** | 0.0 | NaN | NaN | NaN | NaN | NaN | NaN | NaN |
| ***parking\_cluster\_active*** | 0.0 | NaN | NaN | NaN | NaN | NaN | NaN | NaN |
| ***total\_available\_places*** | 27625 | 313.91 | 562.70 | -109 | 38 | 94 | 318 | 3220 |
| ***total\_occupied\_places*** | 27625 | -104.96 | 582.41 | -3070 | 5 | 50 | 99 | 727 |
| ***total\_places*** | 27625 | 208.95 | 193.91 | 79 | 98 | 113 | 170 | 618 |

Na Tabela 26 para cada atributo é possível observar o número total de dados (Total), o número de valores únicos (Nº Valores Únicos), os dados que aparecem com mais frequência (Top) e o valor da frequência com que aparecem (Freq). Por exemplo, para o atributo “*name*” existem 6 valores únicos e o que aparece com mais frequência (4965 vezes) é o parque “Estação Rodoviaria – Entrada”.

Tabela 26 - Estatísticas variáveis categóricas

| **Atributos Categóricos** | **Total** | **Nº Valores Únicos** | **Top** | **Freq** |
| --- | --- | --- | --- | --- |
| ***name*** | 27625 | 6 | Estação Rodoviaria – Entrada | 4965 |
| ***location*** | 27625 | 4 | Rua Henriques Nogueira | 9930 |
| ***active*** | 27625 | 1 | verdadeiro | 27625 |
| ***observations*** | 27625 | 6 | Zona1, \nZ1 E6-E | E5-S\nZ2 E3-E | E4-S | 4965 |
| ***latitude*** | 27625 | 6 | 41.41470 | 4965 |
| ***longitude*** | 27625 | 6 | -8.50781 | 4965 |
| ***no\_communication\_day*** | 27625 | 1 | falso | 27625 |
| ***no\_communication\_hour*** | 27625 | 1 | falso | 27625 |

#### Indicadores e KPIs

Após a análise dos dados e de forma a ir de encontro aos objetivos foram definidos KPIs. Estes indicadores foram definidos com base no contexto e na análise de artigos relacionados. Um indicador é, por exemplo, as áreas com mais movimento e o respetivo KPI é as áreas com tráfego superior a dois mil movimentos por ano (Tabela 27).

Tabela 27 - Indicadores e KPIs

| **Número** | **Indicador** | **KPIs** |
| --- | --- | --- |
| **1** | Áreas com mais movimento | Com tráfego superior a dois mil movimentos por ano |
| **2** | Parques mais movimentados | Com tráfego superior a mil e quinhentos movimentos por mês |
| **3** | Dias mais movimentados | Com tráfego superior a cento e cinquenta movimentos por dia |
| **4** | Parques com maior capacidade | Capacidade superior a trezentos lugares |
| **5** | Taxa de ocupação | Taxa de ocupação superior a 50% |
| **6** | Horas com maior afluência | Com tráfego superior a duzentos movimentos por hora |

#### Seleção dos dados

Após uma análise detalhada do *dataset*, foi feita uma seleção das colunas tendo em conta os indicadores e KPIs definidos anteriormente (subsecção “Indicadores e KPIs” na página 89). Ou seja, as colunas que não são significativas foram excluídas.

Deste modo, as colunas eliminadas foram:

* *Id:* O número identificador de registo de dados não tem relevância ara as futuras análises;
* *Displat\_id:* Este atributo tem a mesma informação que o atributo “*parking\_lot\_id*”;
* *Provider\_type:* O atributo não tem relevância para as futuras análises;
* *Provider\_id:* Este atributo tem a mesma informação que o atributo “*parking\_lot\_id*”;
* *Active:* O atributo não tem relevância interesse para as futuras análises;
* *parking\_cluster\_id:* O atributo não tem valores, e, consequentemente, não tem relevância interesse para as futuras análises;
* *parking\_cluster\_name:* O atributo não tem valores, e, consequentemente, não tem relevância para as futuras análises;
* *parking\_cluster\_active:* O atributo não tem valores, e, consequentemente, não tem relevância para as futuras análises;
* *total\_places:* Este atributo tem a mesma informação que o atributo “*total\_capacity*”;
* *no\_communication\_day:* O atributo não tem relevância para as futuras análises;
* *no\_communication\_hour:* O atributo não tem relevância para as futuras análises;

#### Transformação e Carregamento dos dados

A nível da qualidade dos dados apenas um erro de escrita foi detetado e corrigido por meio do *Excel* (planilha que suporta as operações de análise de dados). A transformação que os dados sofreram foi:

* Uma das localizações estava escrita de duas maneiras diferentes, Rua Fernando Mesquita e Rua Farnando Mesquita. Pelo que através da função do *replace* o nome ficou alterado para Rua Fernando Mesquita uma vez que ambas as ruas se referenciam à mesma localização.

Para responder aos indicadores relacionados com o movimento, através da linguagem *python*, foi adicionada uma nova coluna “Dif” com a diferença entre os atributos *total\_available\_places* e *total\_occupied\_places.* Seguidamente foi criada uma função que sempre que os valores da diferença se alteram regista 1 numa nova coluna chamada “Mov”. Isto significa que, se os valores da diferença se alteram existe movimento no parque. Caso contrário, é registado na nova coluna o valor de 0.

### Data Warehouse

As *data warehouse* (DW) são repositórios centrais de dados de uma ou mais fontes que podem armazenar dados atuais e históricos num local único. Um modelo multidimensional é um tipo de modelação utilizada na conceção de *data warehouses*. Este tipo de modelação enfatiza a identificação, modelação e implementação de processos empresariais críticos. Um modelo multidimensional é composto por tabelas de factos e dimensões. As tabelas de factos contêm os indicadores de negócio necessários para as análises pretendidas, enquanto as dimensões descrevem os pontos **(OCI, 2022).**

#### Desenvolvimento do Data Warehouse

De acordo com os dados (Tabela 24), foi possível perceber que era necessário construir duas tabelas de factos, uma relativa aos movimentos e outra relativa aos dados de cada parque. O conjunto de dados fornece horas e dias de cada recolha efetuada, deste modo, foram construídas uma dimensão calendário (na figura apresentada como dim\_calendar) e outra dimensão relacionada com o tempo (na figura apresentada como dim\_time). O atributo latitude e longitude permitiu a criação de uma dimensão chamada localização (na figura apresentada como dim\_location). E o atributo com o nome da rua do parque permitiu a criação de uma dimensão local (na figura apresentada como dim\_local).

Após a identificação das tabelas de factos e dimensões, o modelo foi construído, e as possíveis ligações entre os elementos foram identificadas. Desta forma, é possível ver na Figura 15 que a tabela de factos de gestão de disponibilidade referente aos movimentos do parque (na figura apresentada como *facts\_table\_availabilityManegement*) está ligada à dimensão calendário e à dimensão tempo. E a tabela de factos relativa à gestão do parque (na figura apresentada como *facts\_table\_ParkingManengement*) está ligada à dimensão de localização e local.

A picture containing diagram

Description automatically generated

Figura 15 - Data Warehouse.

Nesta fase, foi criada a *data warehouse* com quatro dimensões: tempo, calendário, localização e local. E duas tabelas de facto: Gestão de disponibilidade e Gestão do parque.

Estas dimensões e tabelas de facto foram pensadas e estruturadas consoante os dados obtidos e posteriormente analisados. Ou seja, como tínhamos no *dataset* informações relativas ao tempo foi possível criar duas dimensões uma relativamente às horas e outra relativa aos dias. Na dimensão tempo tem como atributos o id, hora completa (*full-time*), hora (*hour*), minuto (*minute*), segundo (*second*) e se é am ou pm. Na dimensão calendário tem como atributos o id, a data completa (*date*), o dia da semana (*day of week*), do mês (*day of the month*) e do ano (*day of the year*), se é fim de semana ou dia da semana (*weekend*), mês (*month*) e número do mês (*month of year*), ano (*year*), assim como a data do dia anterior (*previous\_day*) e a data do dia seguinte (*next\_day*).

Através da informação do nome da rua proveniente do *dataset* foi possível concluir o concelho, distrito e cidade do parque, por isso foi criada a dimensão local.

Com as informações das coordenadas geográficas (latitude e longitude) foi criada a dimensão localização.

A tabela de facto *Parking Management* refere-se aos dados e informações únicas dos parques existentes. Ou seja, como todos os parques tem coordenadas geográficas e um local associado, existe uma ligação de um para muitos com as dimensões local e *location* para que a tabela de facto herde esses dados do parque. Assim como, tem atributos relativos ao lote, nome, observações e total de lugares do parque.

A tabela de facto *Availability Management* refere-se aos movimentos que cada parque tem. Neste sentido, a tabela esta ligada à dimensão tempo e calendário uma vez que cada movimento tem associado uma hora e uma data. Esta tabela de facto está também ligada à tabela de factos *parking management* com o objetivo de identificar e herdar o parque associado a cada movimento.

As colunas de interesse foram filtradas e, após as alterações necessárias, todos os dados selecionados foram carregados para cada dimensão da *data warehouse*. Todas as dimensões foram carregadas através do MySQL Workbench.

As dimensões tempo e calendário foram carregadas por scripts MySQL, enquanto as dimensões local e localização foram carregadas diretamente com os dados do excel.

As tabelas de factos foram carregadas usando a biblioteca do panda (um software de código aberto que permite a análise e manipulação de dados). Esta biblioteca foi utilizada pois era necessário carregar dados do Excel, mas também agregar dados das dimensões associadas. Esta foi a maneira mais viável de carregar todos os dados inseridos para corresponder uns aos outros.

### OLAP

Processamento Analítico Online (OLAP) é uma ferramenta que permite aos utilizadores analisar informação relevante de diferentes pontos de vista e em várias combinações. Através da ligação a uma base de dados, o sistema permite a visualização de gráficos estatísticos que possibilita a extração, consulta e recuperação de dados **(Onesmus Mbaabu, 2021)**.

#### Modelo

Para a construção do modelo OLAP foi utilizada a ferramenta cube.js. O Cube.js é um sistema de inteligência empresarial de código aberto que permite pré-calcular e pré-agregar dados, criando um cubo OLAP que torna a análise mais rápida. O cube.js permite a ligação a várias bases de dados, e consequentemente, o esquema é definido para cada dimensão e tabela de factos em vários cubos. Para as tabelas de factos, as chaves estrangeiras e as relações entre os outros cubos são definidas assim como os factos que queremos calcular. Através de pré-agregações, podemos guardar as combinações desejadas para uma análise mais rápida. O objetivo é obter *dashboards* analíticos baseados na DW implementada e obter uma visualização que satisfaçam as necessidades de informação para a gestão do negócio **(cube, 2022)**.

#### Resultados

As estatísticas a seguir apresentadas foram obtidas para cada indicador e KPIs no *playground* do cube.js. Estes indicadores e KPIs foram definidos previamente na subsecção “Indicadores e KPIs” na página 89. Os resultados são apenas indicativos e não devem ser entendidos como valores absolutos. O exemplo apresentado é facilmente replicado para outras cidades, sendo que, os resultados expostos dizem respeito ao projeto ioCity que ajudará a monitorar os fluxos nos parques. Esses dados podem ser cruzados com dados complementares, como clima e eventos, tornando o sistema ainda mais completo. O modelo é escalável e os dados são apenas indicativos do que é possível obter.

##### Indicador 1 – Áreas com mais movimento

Na Figura 16, é possível observar que a Rua Henriques Nogueira é a área de maior movimento (5385 movimentos, aproximadamente 35%). Em contrapartida, a Rua de São Vicente é o menos movimentado (1328 movimentos, cerca de 8,5%).

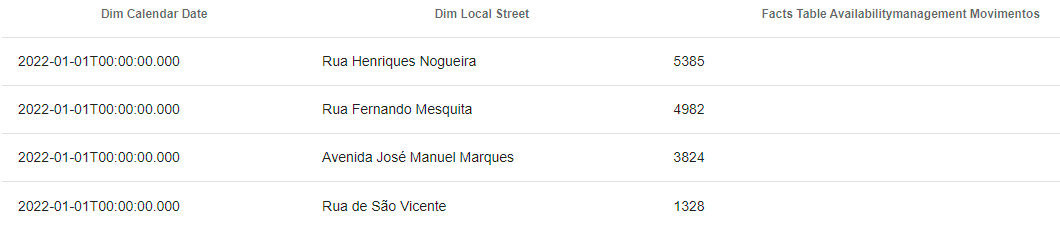


Figura 16 - Áreas com mais movimento anualmente.

Na Figura 17 é possível ver os locais com mais de dois mil movimentos por ano (aparecem apenas três áreas: Rua Henriques Nogueira, Rua Fernando Mesquita e Avenida José Manuel Marques).



Figura 17 - Áreas com tráfego superior a dois mil movimentos por ano.

##### Indicador 2 – Parques mais movimentados

Na Figura 18 é possível ver que o Parque Campo da Feira, tanto em fevereiro como em março, é o mais movimentado com aproximadamente 25% dos movimentos (1554 movimentos em fevereiro e 2270 movimentos em março).

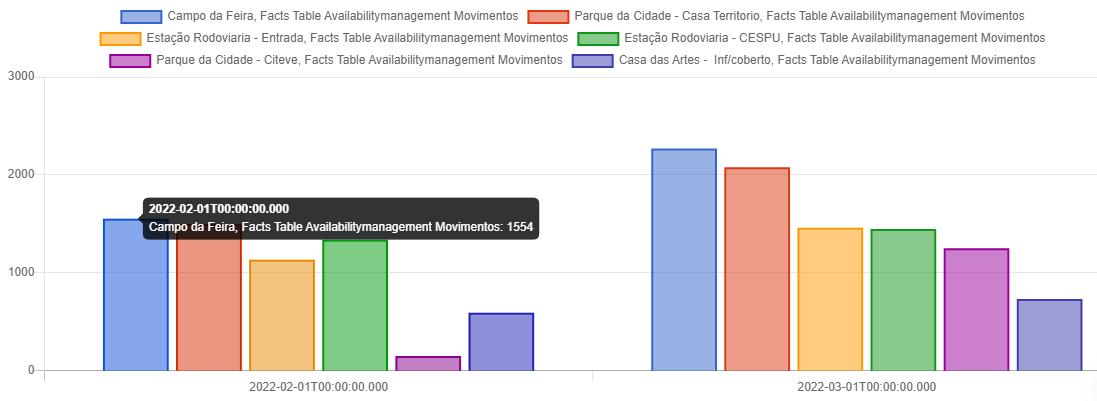


Figura 18 - Parques mais movimentados mensalmente.

Além disso, apenas o parque Campo da Feira em fevereiro teve mais mil e quinhentos movimentos (Figura 19).

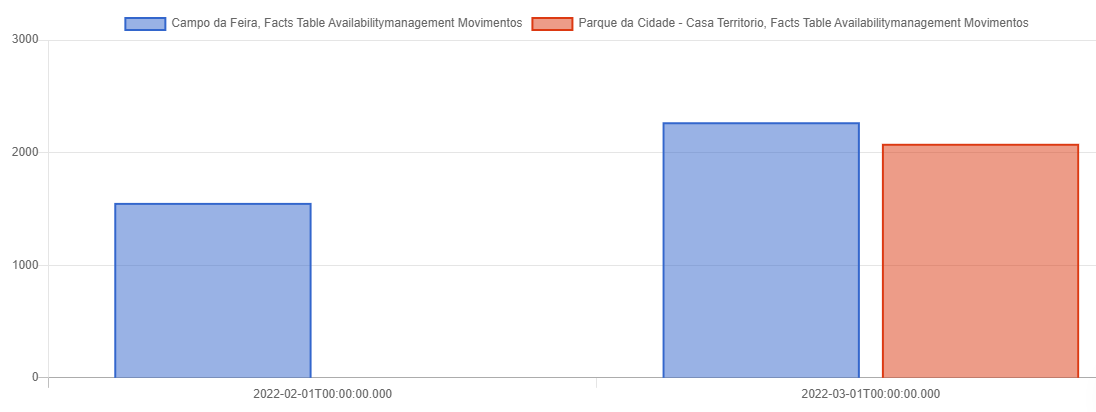


Figura 19 - Parques com tráfego superior a mil e quinhentos movimentos por mês.

##### Indicador 3 – Dias mais movimentados

Na Figura 20 é possível ver que o dia mais movimentado é 11 de março, com 535 movimentos. Além disso, o dia com o movimento menor é 21 de março, com 110 movimentos.

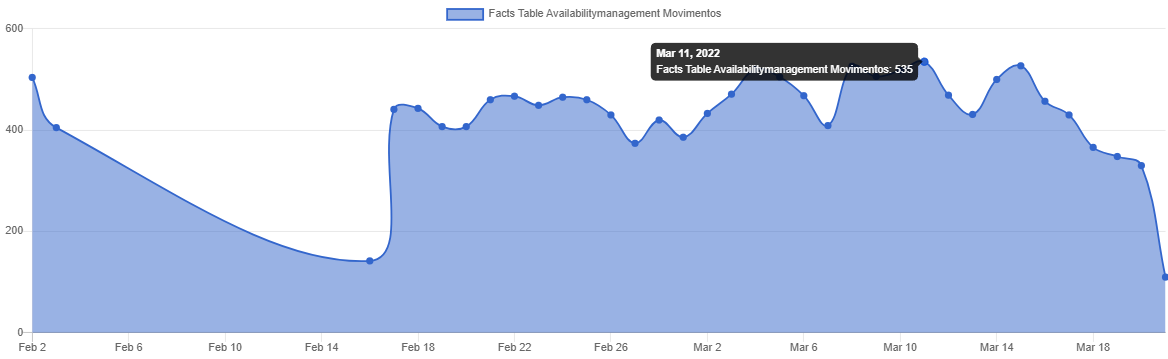


Figura 20 - Dias mais movimentados.

Na Figura 21 aparecem apenas os dias com mais de 150 movimentos. Como 16 de fevereiro e 21 de março não atendem ao requisito, eles não são mostrados no gráfico.

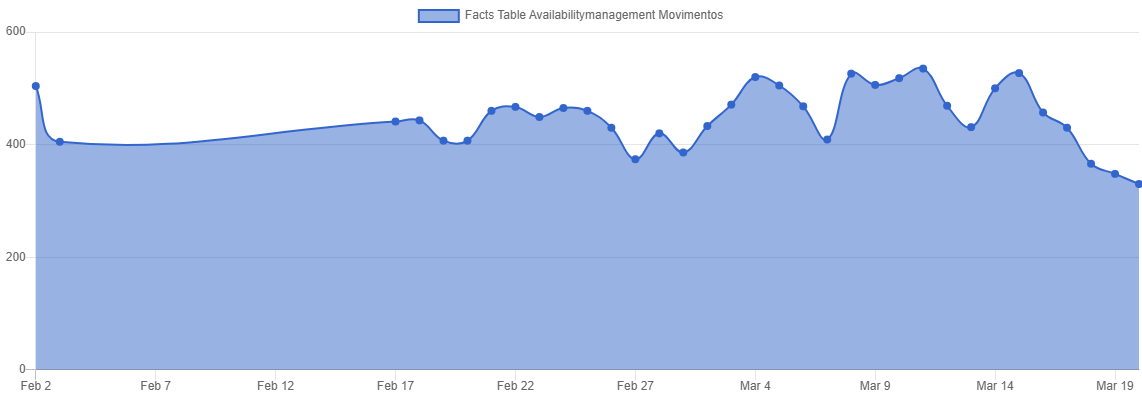


Figura 21 - Dias com tráfego superior a cento e cinquenta movimentos por dia.

##### Indicador 4 – Parques com maior capacidade

Na Figura 22 é possível ver o parque com maior lotação, neste caso, o Parque Campo da Feira, com 618 lugares (50,3% do total de lugares).

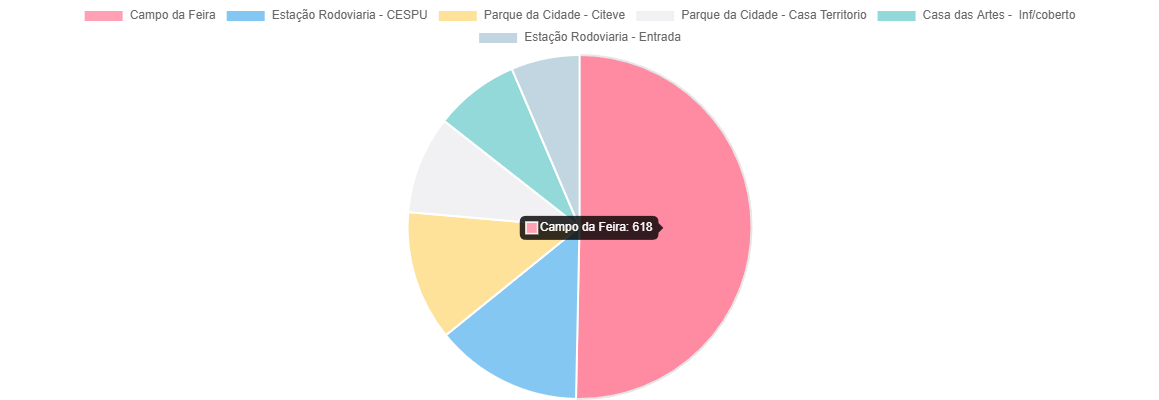


Figura 22 - Parques com mais capacidade.

Na Figura 23 é possível ver os parques em mais de trezentos lugares. O Parque Campo da Feira é o único que cumpre o requisito.

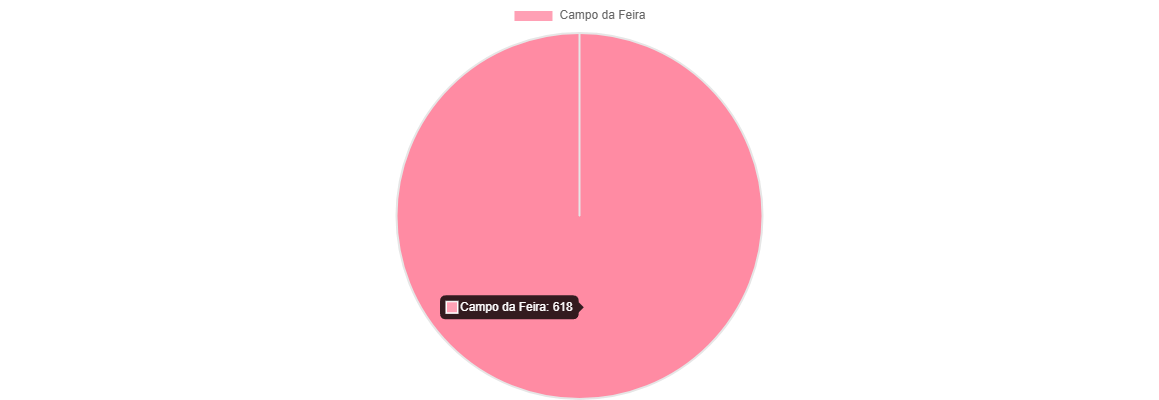


Figura 23 - Parques com capacidade superior a 300 lugares.

##### Indicador 5 – Taxa de ocupação

Na Figura 24 é possível observar que o parque com maior taxa de lotação anual é o parque Casa das Artes – Inf/coberto com 61,32%. Os parques Estação Rodoviária – Entrada e Parque da Cidade Citeve não aparecem na imagem porque não contêm dados válidos.

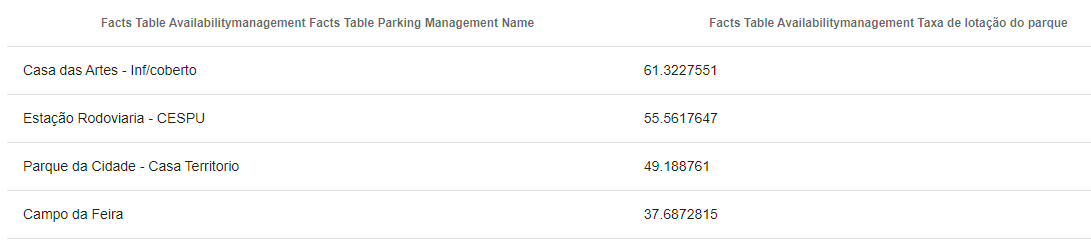


Figura 24 - Taxa de ocupação anual.

Na Figura 25 mostra os parques com mais de 50% de lotação (Parque Casa das Artes – Inf/coberto e Estação Rodoviária – CESPU).

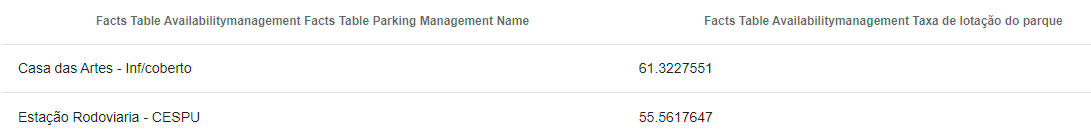


Figura 25 - Taxa de ocupação superior a 50%.

##### Indicador 6 – Horas com maior afluência

Na Figura 26 é possível visualizar as horas com mais movimento (dezoito horas com novecentos e quarenta e seis movimentos).

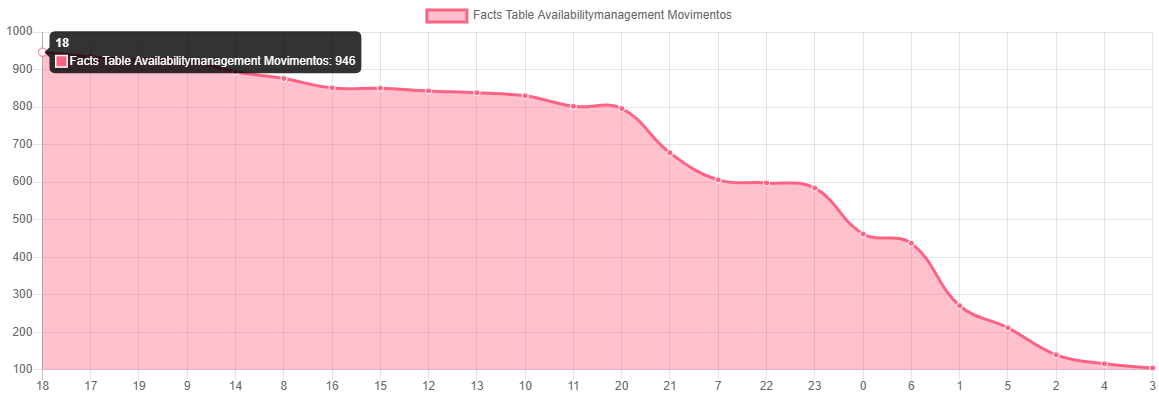


Figura 26 - Horas com maior afluência.

Na Figura 27 são mostradas apenas as horas com mais de duzentos movimentos.

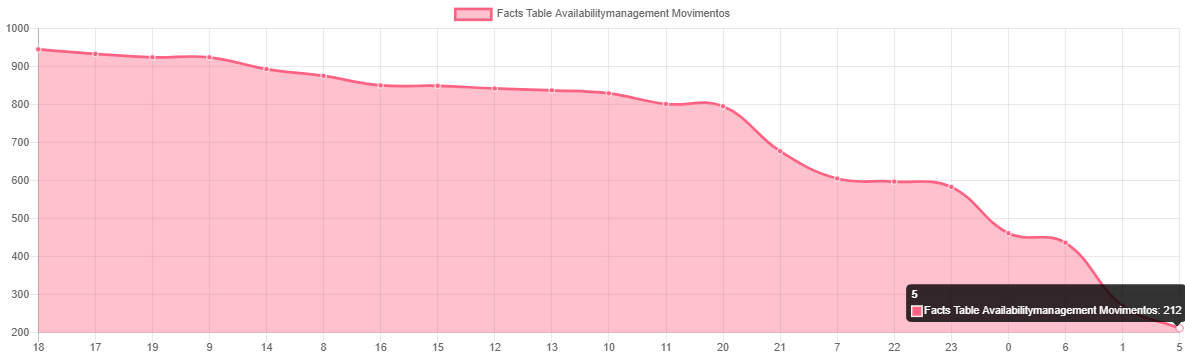


Figura 27 - Horas com tráfego superior a 200 movimentos por hora.

#### Análises dos resultados

Através do processo OLAP, é possível perceber diversas métricas e indicadores úteis na tomada de decisão. A análise do contexto atual permite uma melhor gestão do parque, criando mais valor e utilidade para quem usufrui e gere o estacionamento. Métricas como taxa de lotação, movimento e horas mais movimentadas podem ser analisadas. Por exemplo, supondo que há uma rua com muita procura em comparação com os lugares disponíveis existentes, a construção de um novo parque pode ser uma opção para criar mais condições de estacionamento e reduzir o fluxo de tráfego, permitindo uma melhor mobilidade.

Os KPIs são fundamentais para uma perceção da realidade do estacionamento dentro das cidades. Conhecer o contexto atual para futuras tomadas de decisões assertivas é crucial para o sucesso de uma organização. Após o desenvolvimento do sistema, foi possível mensurar os KPIs previamente definidos para este estudo de caso:

* O Parque Campo da Feira tem a maior capacidade com 618 lugares de estacionamento. Consequentemente, é o parque com maior movimento por ano (3824 movimentos por ano).
* Por outro lado, apesar de ser o parque de estacionamento com menos lugares (79 lugares), o parque de estacionamento Estação Rodoviária – Entrada é o quarto parque de estacionamento dos seis com mais tráfego. A rua mais movimentada é a Rua Henriques Nogueira (5385 movimentos) e nesta rua existem dois parques (Estação Rodoviária – CESPU e Estação Rodoviária – Entrada).
* É impossível concluir os resultados sobre a taxa de lotação porque os dados de origem estão errados e, portanto, os resultados estão incorretos.
* As horas de maior movimento são as horas de ponta: 8h às 9h, 17h às 19h e das 12h às 15h. Como estes são os horários de trabalho mais comuns (horários de entrada, hora de almoço e saída), é natural que tenham mais movimentos. Tanto os dias da semana como os dias de fim de semana apresentam movimento constante ao longo do ano.
* O sistema desenvolvido pode medir mais valores além dos KPIs, como as vagas disponíveis no parque por ano, mês e hora. Assim como as seções de cada parque têm. Por exemplo, o parque Estação Rodoviária – Entrada tem uma “Zona1, Z1 E6-E | E5-S Z2 E3-E | E4-S”.

O trabalho aqui descrito foi publicado em artigo e encontra-se na secção 6.

### Dashboard Analítica

Apesar da ferramenta cube.js ser bastante completa e permitir a modelação dos dados, assim como, um acesso rápido e eficaz dos mesmos, existem algumas adversidades que não permitem uma visão específica dos dados. Tais como processo de *drill-down*, *rollup*, vários tipos de gráfico e a integração em soluções de software como serviço (Saas), entre outras. Neste sentido, foi desenvolvida uma plataforma de visualização com a API do cube.js que apresenta características úteis e complementares a esta ferramenta de análise. Na Tabela 28 são apresentadas características das ferramentas.

Por exemplo, o processo de filtragem é uma característica comum tanto ao cube.js como à *dashboard* desenvolvida.

Tabela 28 - Cube.js vs Dashboard.

| **Características** | **Ferramenta** | |
| --- | --- | --- |
| ***Cube.js*** | ***Dashboard*** |
| **Definir estrutura dos dados** | X |  |
| **Pré-agregação dos dados** | X |  |
| **Possibilidade de conexão a várias *data warehouses*** | X |  |
| **Possibilidade de diversas combinações de dados** | X | X |
| **Processo de filtragem** | X | X |
| **Sugestão de dados no processo de filtragem** |  | X |
| **Diversos gráficos de visualização e mapa** |  | X |
| **Tabela auxiliar aos gráficos com dados em valor numérico e/ou percentagem** |  | X |
| ***Drill-Down*** |  | X |
| ***Rollup*** |  | X |
| **Página para guardar gráficos úteis para a tomada de decisão** |  | X |
| ***Labels* dos gráficos curtas e objetivas** |  | X |
| ***Guardar os gráficos para uso particular*** |  | X |

#### Ferramentas e linguagens utilizadas

Após uma pesquisa sobre quais as tecnologias mais adequadas ao desenvolvimento desta camada foram aplicadas as seguintes linguagens e ferramentas para desenvolver a solução.

Foi utilizado o vue.js para o desenvolvimento de *front-end*. O vue.js é uma *framework* de javascript para o desenvolvimento de interfaces web reativas que estende as linguagens de HTML e CSS para a construção de *front-end*. De forma a obter todo o modelo OLAP desenvolvido com a ferramenta cube.js foi utilizado a *query builder* da API do cube.js para que todos os dados fossem herdados. Uma biblioteca de interface gráfica para o utilizador, intitulada de Vuetify, foi utilizada para o desenvolvimento de um *front-end* mais apelativo e intuitivo. O Apache ECharts foi necessário aquando da criação dos gráficos de visualização dos dados para o utilizador e a API do googlemaps para a criação do gráfico mapa.

Uma *API backend* também foi incrementado para definir as rotas de dados via mysql. Para isto, foi criado uma *API* com *node.js/express.js* (*framework* de *back-end* para o desenvolvimento de *APIs* que facilita a criação de páginas Web com *node.js*) para obter os dados necessários. O node.js é um software de código aberto que permite a criação de aplicações javascript sem depender de um browser. A partir da ferramenta *Visual Studio Code* todo o código foi criado e desenvolvido para alcançar os objetivos do projeto.

#### Solução desenvolvida

A plataforma criada apresenta três páginas principais. Toda a solução desenvolvida vai de encontro aos objetivos definidos e as ferramentas e linguagens utilizadas foram descritas previamente na subsecção anterior “Ferramentas e linguagens utilizadas” na página 101. Deste modo, a solução irá ser apresentada segundo as páginas principais: *Explore*, *Dashboard* e KPIs onde todas as funcionalidades serão descritas respetivamente.

* **Explore**

Nesta página é possível explorar diversas combinações de dados entre factos, dimensões e dimensões temporais com granularidade através de diferentes tipos de gráficos. Os gráficos que permitem granularidade são os de linhas, área, barras e radar sendo que os gráficos circulares, *scatter* e *doughnut* não permitem este tipo de análise. É também possível adicionar filtros e guardar o gráfico na página da *Dashboard* associado a um nome. O processo de filtragem engloba todas as dimensões e factos presentes na *data warehouse* e diversos operadores para exploração dos dados. Nesta página é também possível a visualização de uma tabela auxiliar aos gráficos com os dados respetivos do gráfico criado, tanto em percentagem como em valor.

* **Dashboard**

Nesta página são apresentados todos os gráficos guardados. Os gráficos que permitem granularidade apresentam *drill-down* e *rollup* por tempo e por localização. Por defeito, ao clicar numa barra / linha/ área o processo de *drill-down* será apresentado e após a mensagem de término do processo é apresentada a possibilidade de voltar ao gráfico inicial ou de fazer *rollup*. O processo de filtragem também está disponível nesta página e é aplicado a todos os gráficos presentes, ou seja, caso o filtro se aplique a algum dos gráficos guardados o mesmo é aplicado, caso contrário, o gráfico continua igual. A possibilidade de eliminar algum gráfico também está disponível através do botão “*Delete*”.

* **KPIs**

São expostos os KPIs em diversos gráficos definidos na subsecção “Indicadores e KPIs” na página 89.

#### Exemplo prático de análise

De forma a complementar a descrição da subsecção “Solução desenvolvida” na página 102 um exemplo prático será apresentado. Neste exemplo será demonstrado como é possível fazer um gráfico com uma análise tridimensional associado a um processo de filtragem e *drill-down* temporal. Os dados apresentados foram expostos e analisados na subsecção 4.3.2.

1. Selecionar um facto, uma dimensão, uma dimensão tempo e um tipo de gráfico (Figura 28);

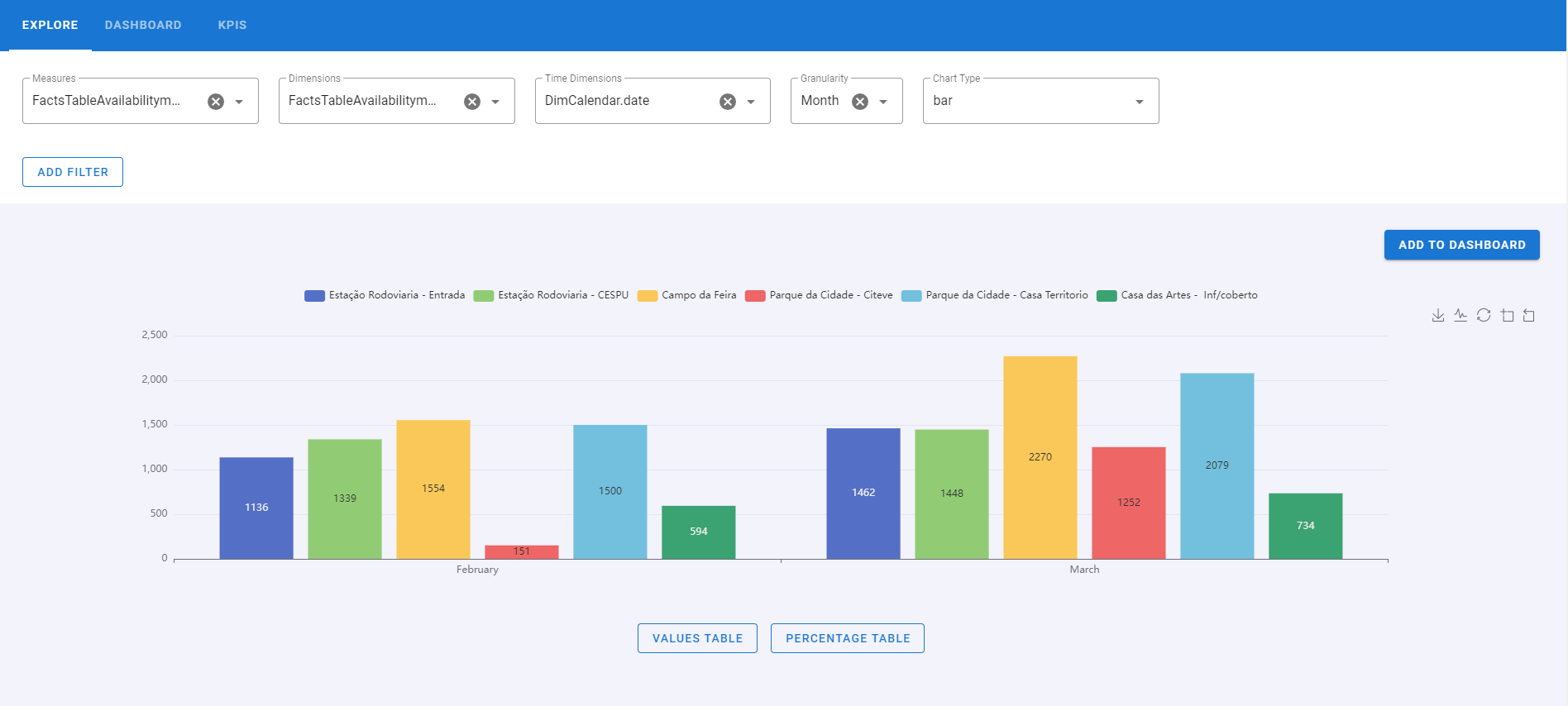


Figura 28 - Passo 1 de utilização da dashboard.

1. Após a criação do gráfico é possível visualizar uma tabela auxiliar ao gráfico tanto em valor (Figura 29) como em percentagem (Figura 30);

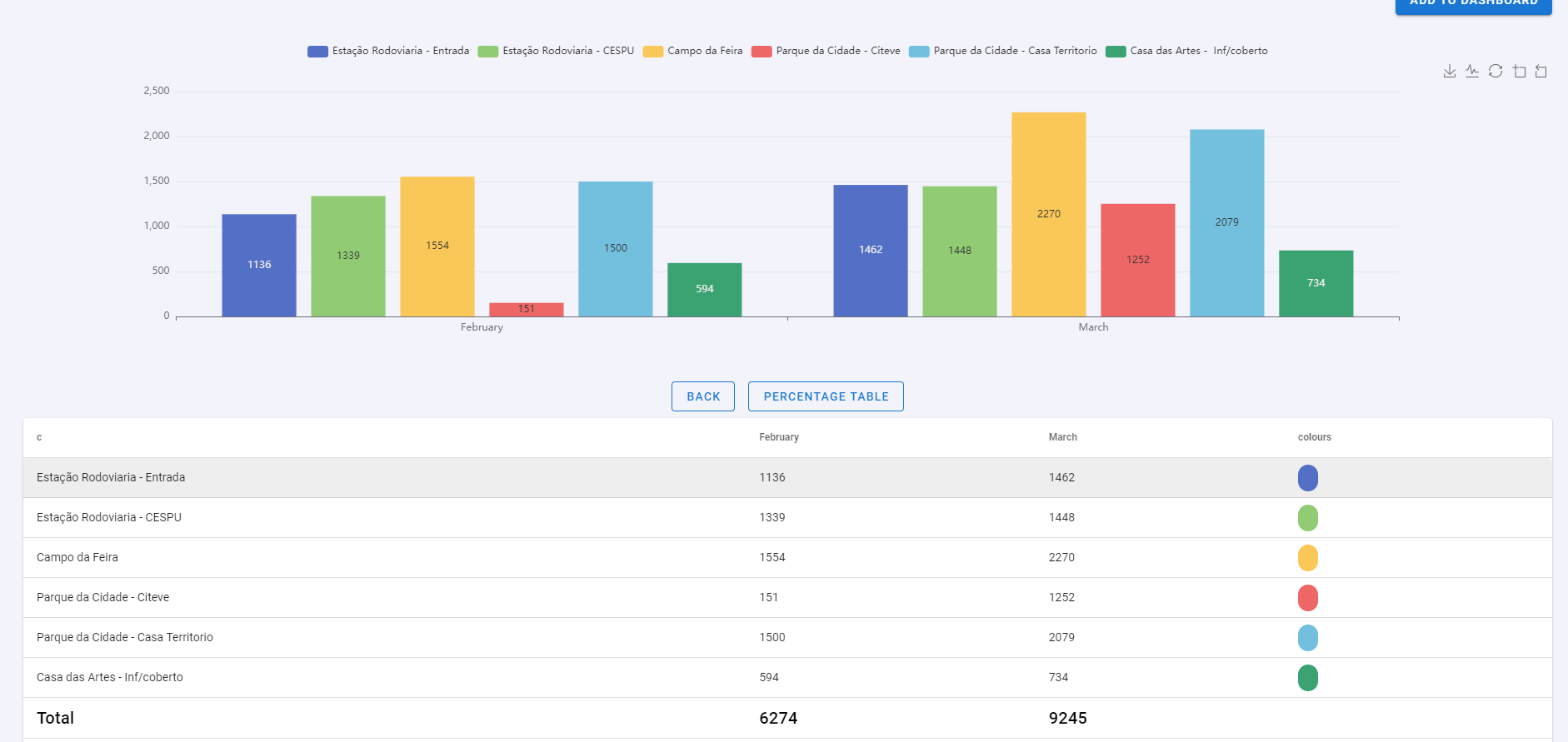


Figura 29 - Passo 2 de utilização da Dashboard.

Na próxima figura a tabela auxiliar apresenta os valores em percentagem (Figura 30);

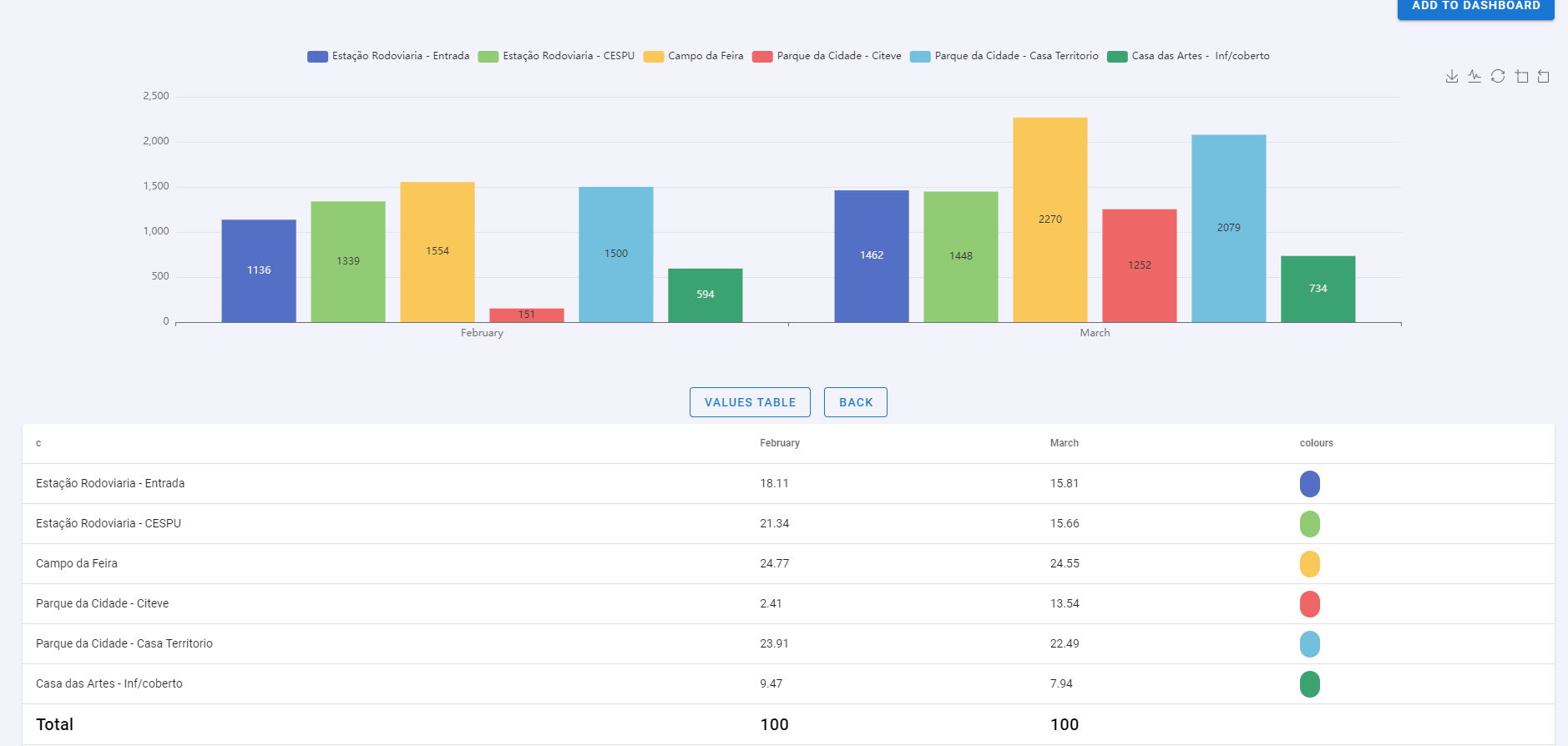


Figura 30 - Passo 3 de utilização da Dashboard.

1. Carregar no botão “*Add filter*”. Selecionar uma dimensão ou facto. Selecionar um operador (*equals*, *does note qual*, > , < , = , >= , <=). Inserir um valor caso tenha sido selecionado um facto ou selecionar um atributo caso tenha sido selecionado uma dimensão (Figura 31);

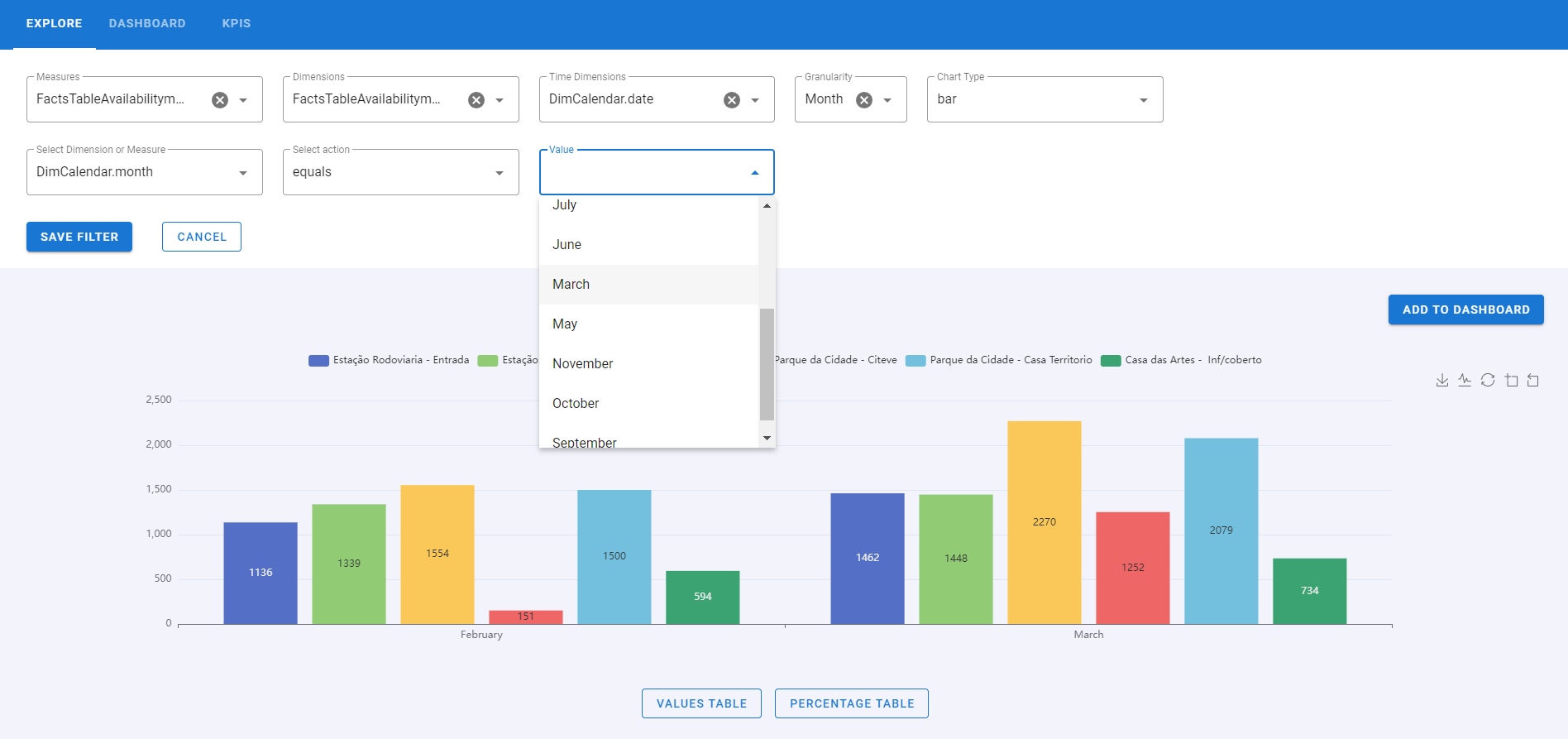


Figura 31 - Passo 4 de utilização da Dashboard.

1. Carregar no botão “*Save Filter*”. Podem ser inseridos vários filtros e posteriormente apagados (Figura 32);

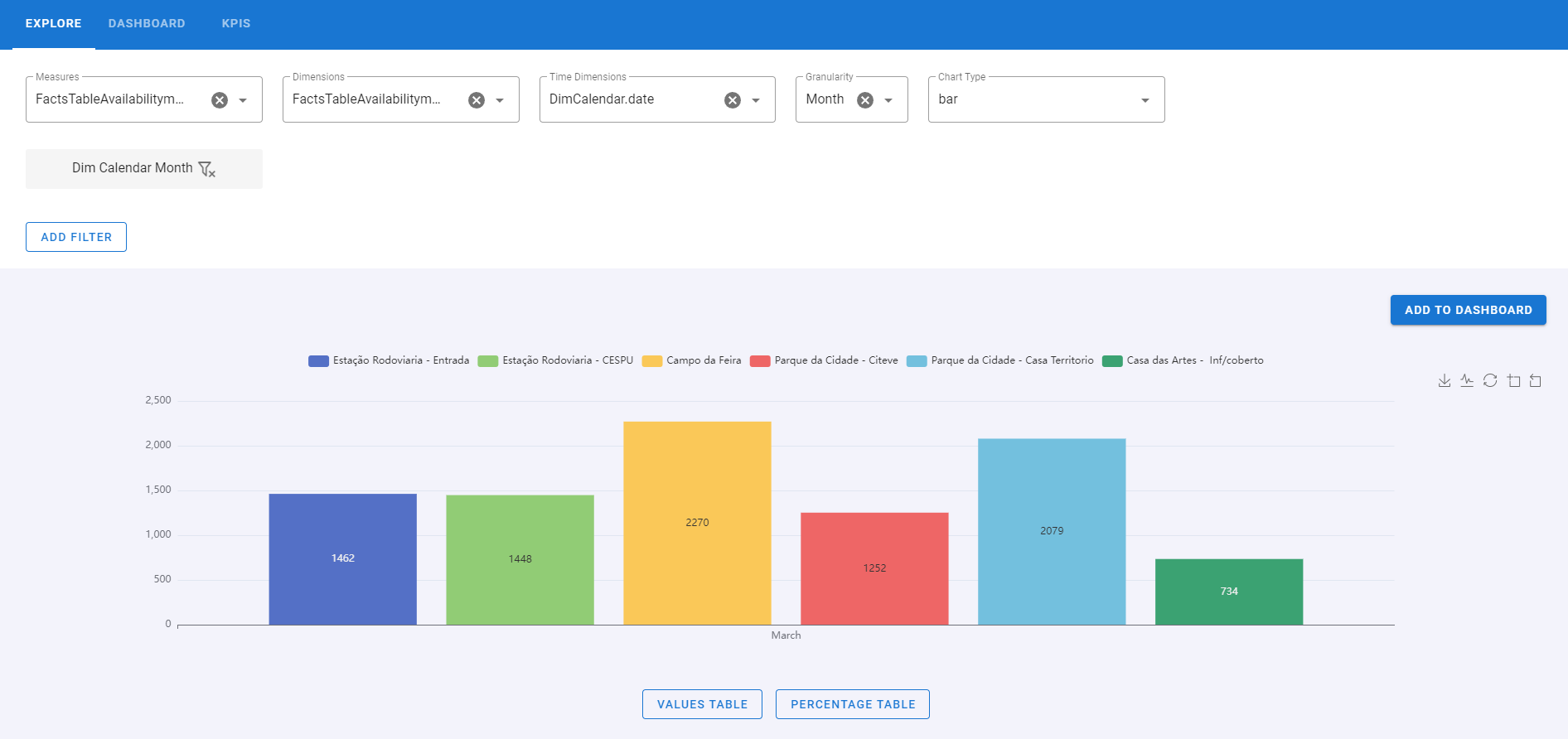


Figura 32 - Passo 5 de utilização da Dashboard.

1. Carregar no botão “*Add to Dashboard*”, dar um nome ao gráfico e carregar em “*Save*” (Figura 33).

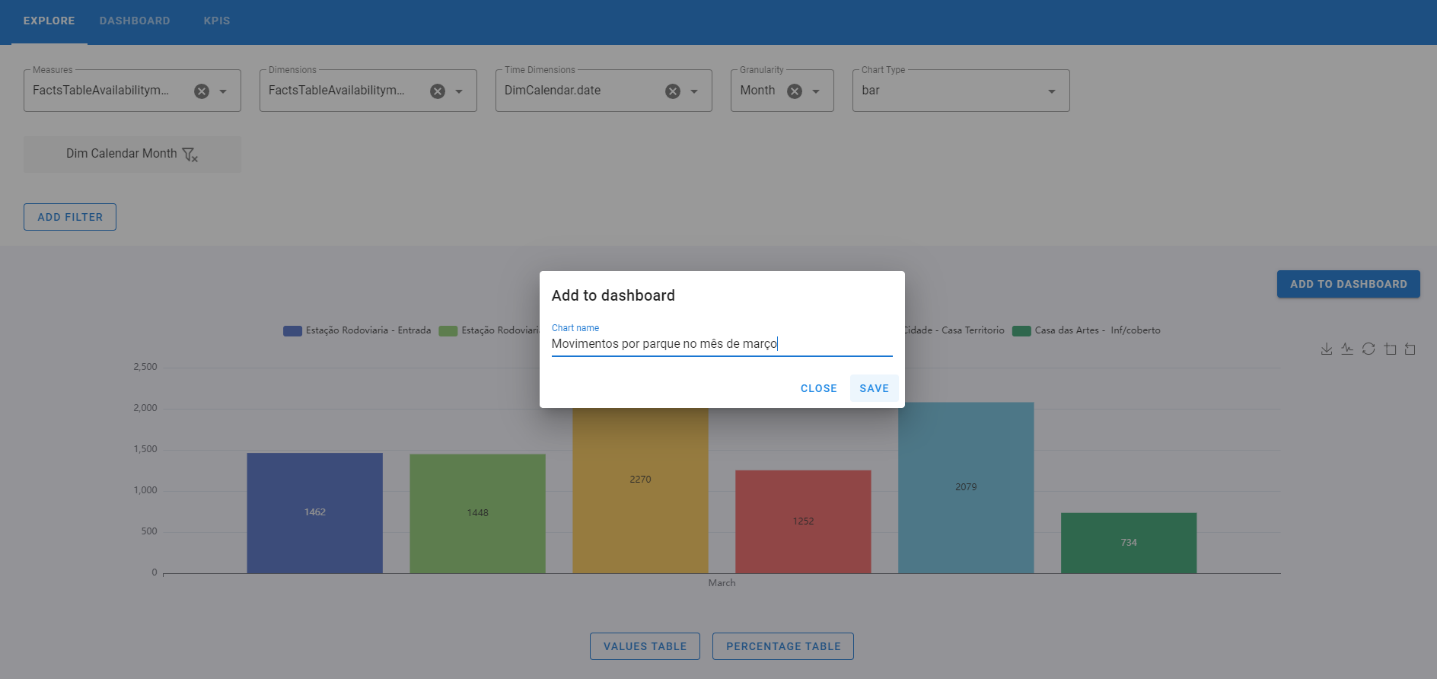


Figura 33 - Passo 6 de utilização da Dashboard.

1. O gráfico já se encontra na página da *Dashboard* e pode ser eliminado (Figura 34);

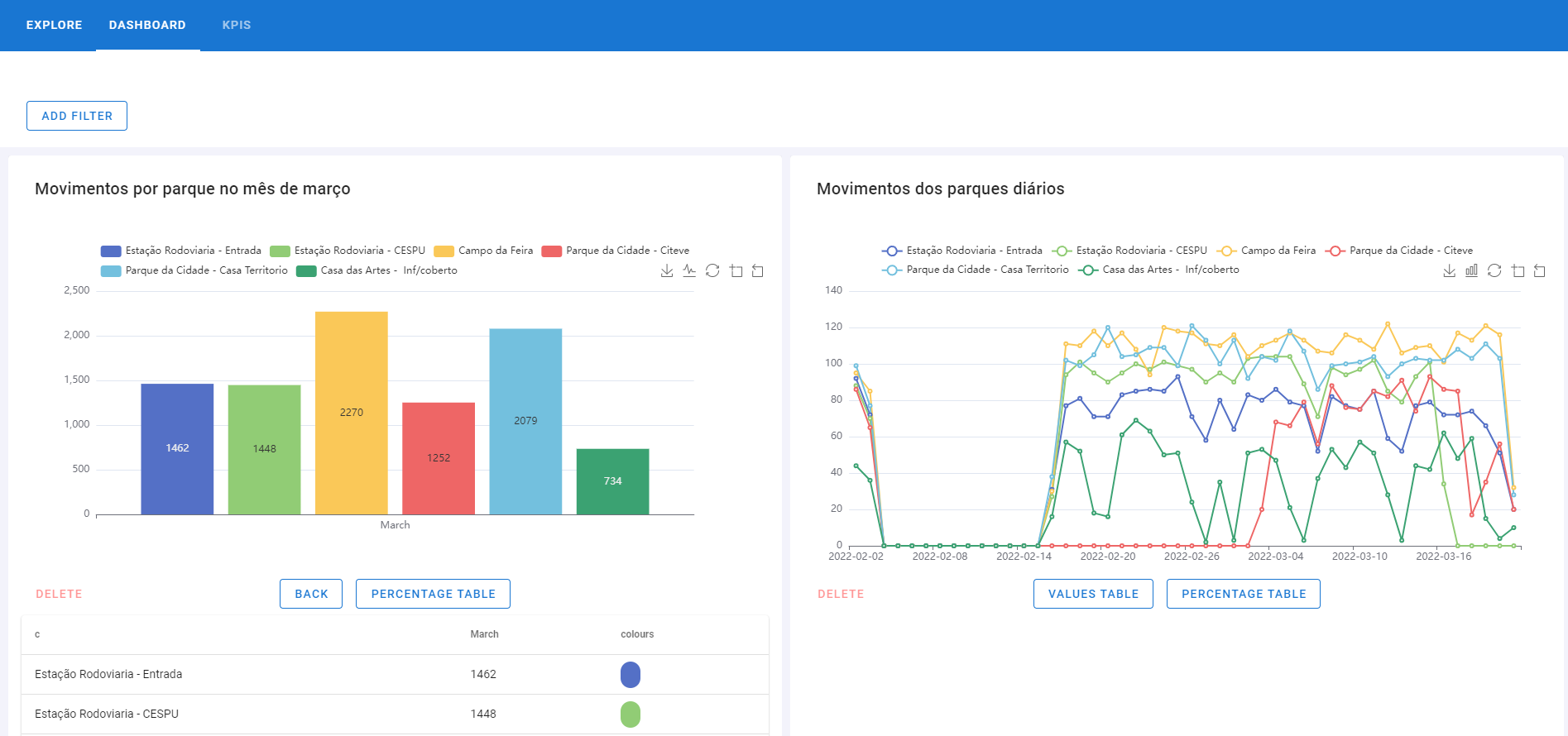


Figura 34 - Passo 7 de utilização da Dashboard.

1. Para fazer o processo de *drill-down* o passo 1 tem de ser rigorosamente cumprido. Neste processo basta carregar em cima da barra/linha/área e o gráfico irá descer de nível.

Durante o *drill-down* é exibido o grau onde o processo se encontra, ou seja, se o gráfico está por ano, quarter, mês, semanas, dias da semana ou horas e o atributo que foi anteriormente selecionado. Neste exemplo a barra do Campo da feira com 2270 movimentos foi selecionada (Figura 34) e consequentemente aparece o grau que foi selecionado e o respetivo atributo (*previous degree*): Mês de março (*Month - March*) e o grau em que o gráfico se encontra no momento (*current degree*): semana (week). Um novo gráfico por semana é exibido, sendo que, apenas as semanas 10, 11, 12 e 13 apresentam movimento (Figura 35).



Figura 35 - Passo 8 de utilização da Dashboard.

* + A semana 12 (do gráfico anteriormente mostrado Figura 35) foi selecionada. Um gráfico por dias é exibido. Neste caso apenas há movimentos nos dias 20 de março de 2022 e 21 de março de 2022. A barra correspondente ao dia 20 de março foi selecionada (Figura 36);

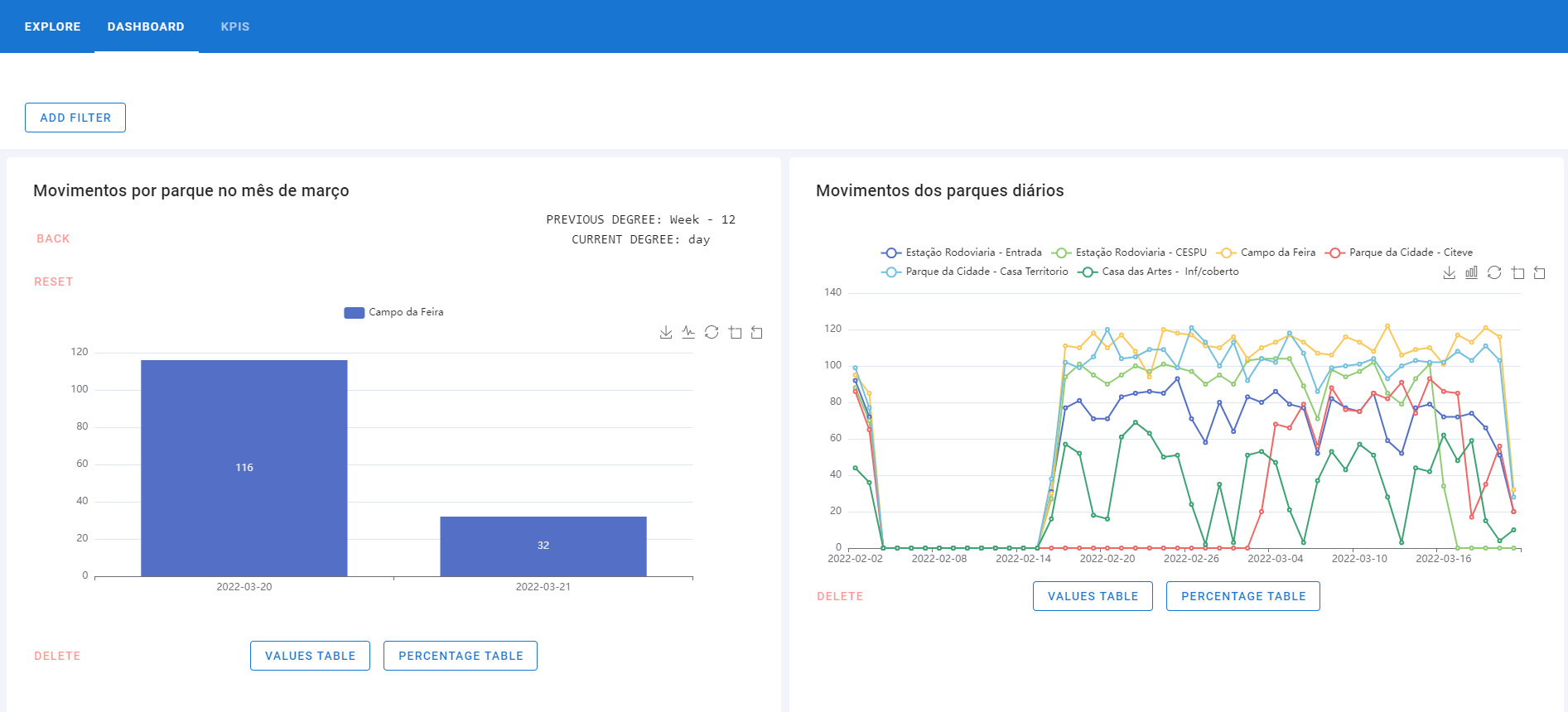


Figura 36 - Passo 9 de utilização da Dashboard.

* + As horas em que houve movimento nesse dia (2022-03-20) são expostas. Sendo as horas o atributo mais baixo do *drill-down* o processo termina (Figura 37);

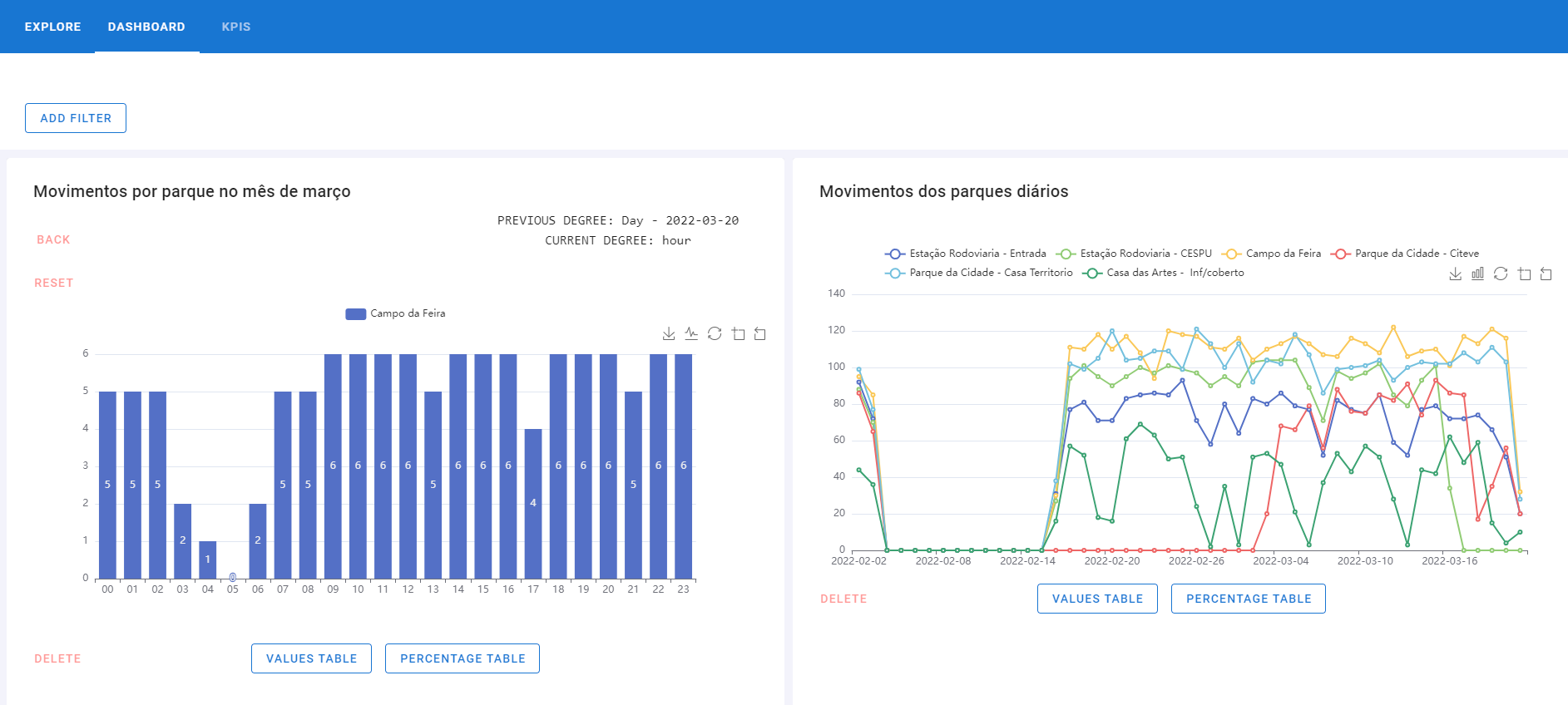


Figura 37 - Passo 10 de utilização da Dashboard.

1. Quando o *drill-down* acaba é exibida uma mensagem de término. A partir do momento que se efetua o *drill-down* um botão “*Reset*” é apresentado para voltar para o gráfico inicial. Também existe a possibilidade de fazer *rollup* (processo oposto do *drill-down*) através do botão “*Back*”. Ou seja, em vez de o grau descer (ano – quarter – mês – semanas – dias – horas) irá subir, pelo que, neste exemplo temporal o fluxo será: horas – dias – semanas – mês – quarter – ano. Sendo o ano o nível mais alto do processo (Figura 38).

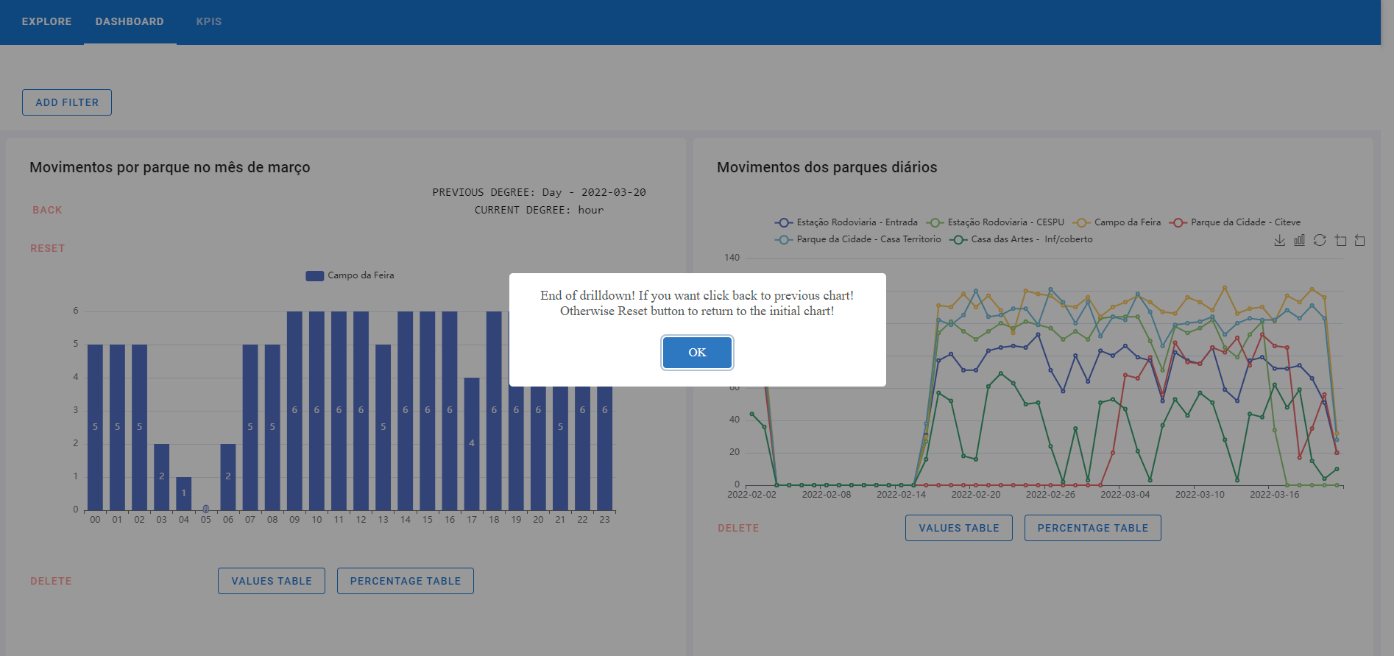


Figura 38 - Passo 11 de utilização da Dashboard.

1. Neste caso, se for pretendido voltar ao gráfico anterior apenas é necessário clicar no botão “Back” e o gráfico atualizará para o grau anterior, neste exemplo, para os dias da semana (Figura 39). Ou seja, após carregar no botão o gráfico é apresentado e as *labels* são atualizadas apresentando o grau anterior: as horas (*previous degree: hour*) e o grau atual: os dias (*current degree: day*);

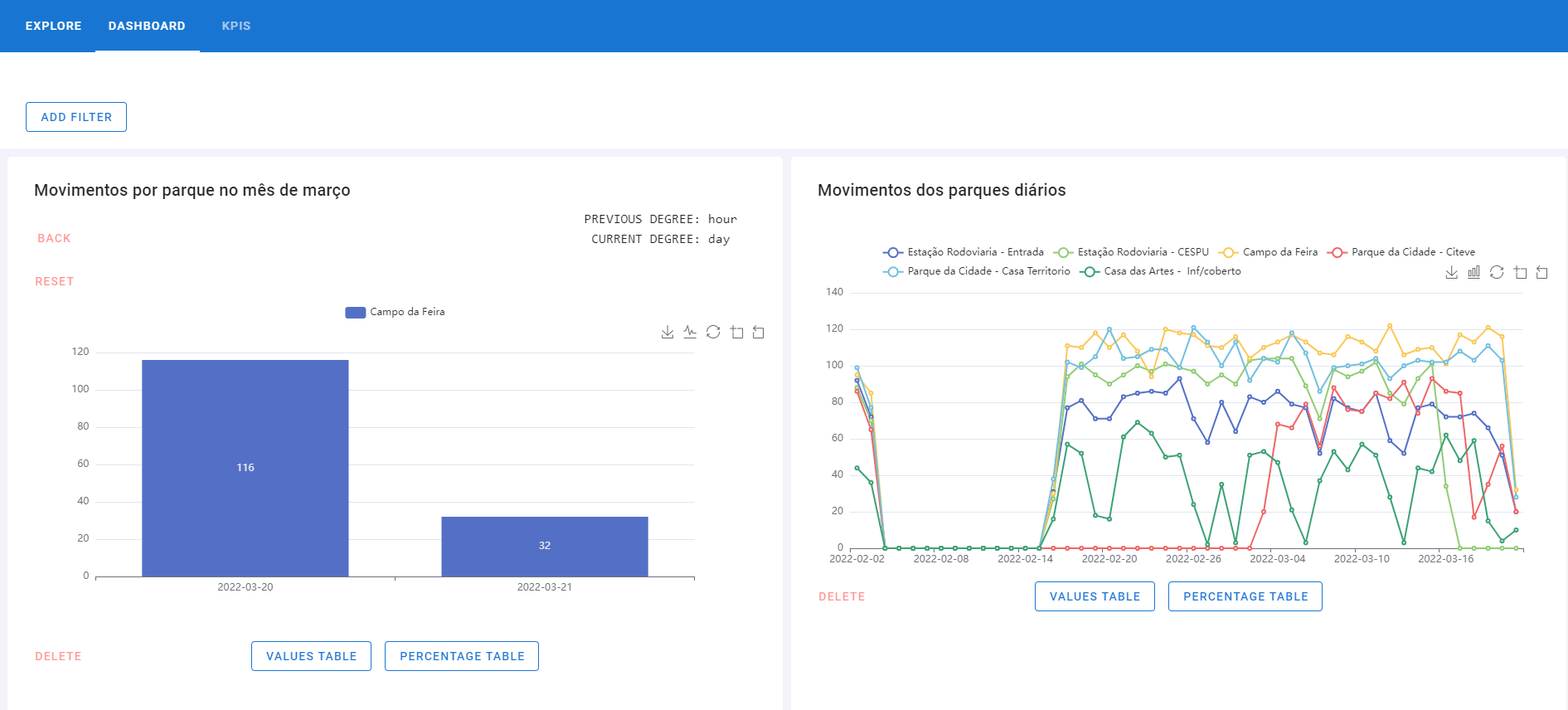


Figura 39 - Passo 12 de utilização da Dashboard.

1. Por fim, se pretendido, ao carregar no botão “*Reset* ” o gráfico original será apresentado (Figura 40). O processo de filtragem pode também ser incluído se assim for o desejado.

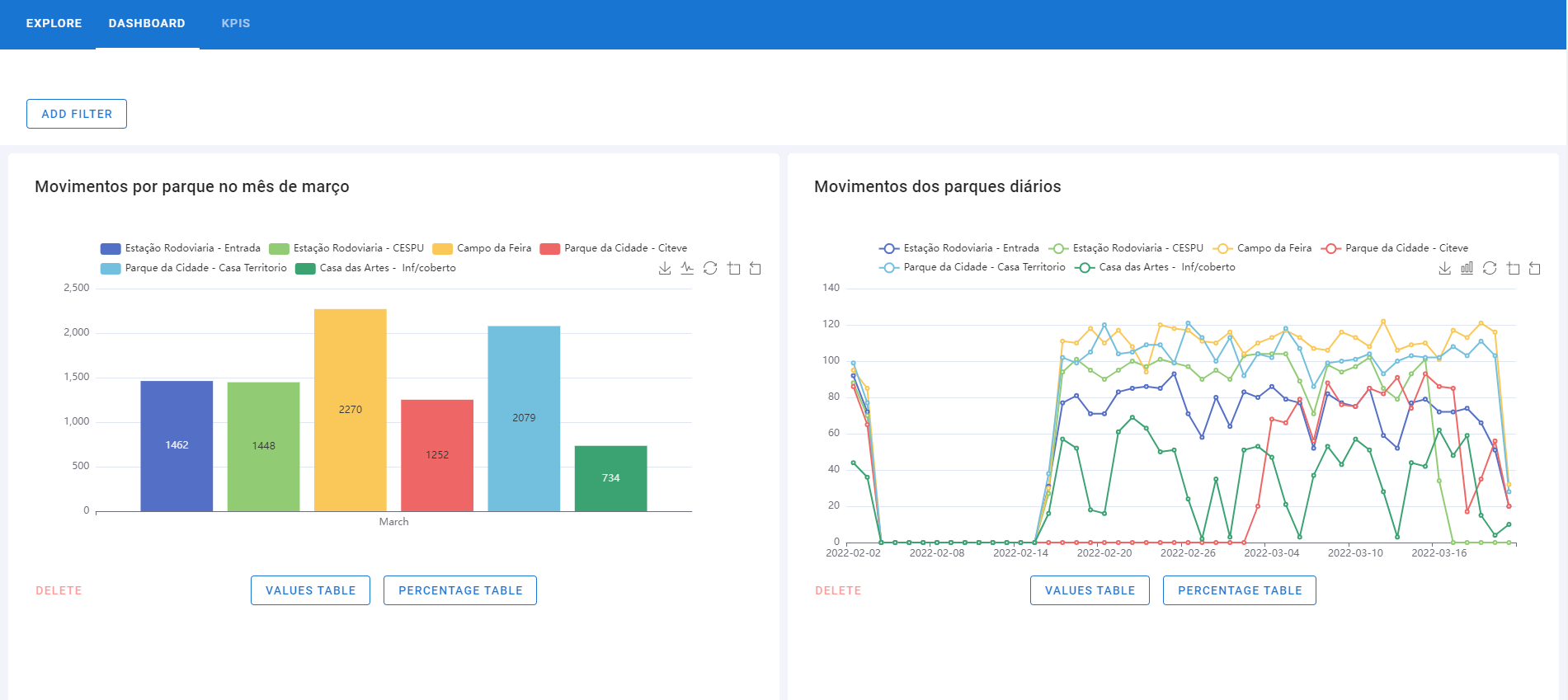


Figura 40 - Passo 13 de utilização da Dashboard.

1. Página com todos os KPIs apresentados (Figura 41).

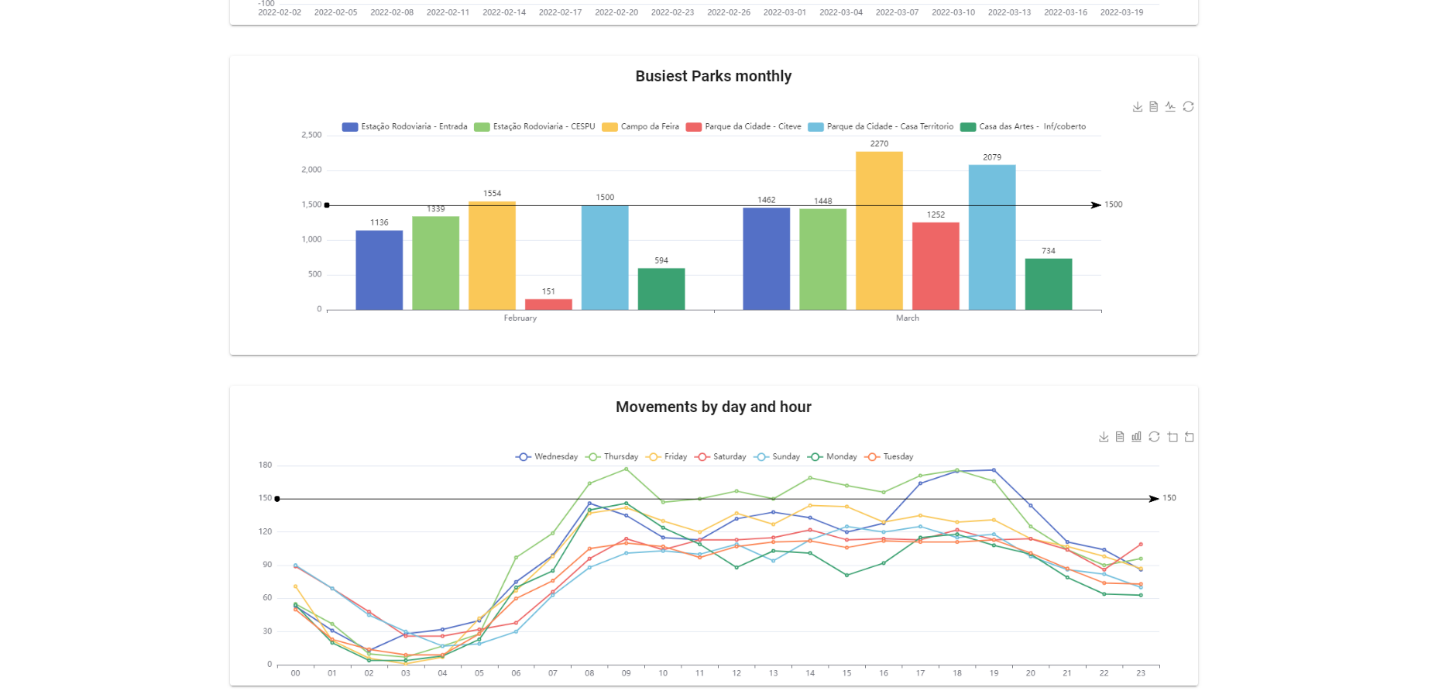


Figura 41 - Passo 14 de utilização da Dashboard

#### Especificação das funcionalidades da solução

De forma a perceber como a solução desenvolvida é capaz de responder ao problema e ao objetivo definido inicialmente é exposto na Tabela 29 as funcionalidades desenvolvidas. A *dashboard* analítica é altamente escalável e de fácil adaptabilidade a outro conjunto de dados permitindo uma análise tridimensional que auxilia no processo de tomada de decisão. Foram identificadas dez funcionalidades cruciais ao apoio à tomada de decisão, assim como, uma *data warehouse*, um modelo OLAP e seis indicadores foram construídos para a solução desenvolvida.

Tabela 29 - Funcionalidades da solução

| **Número** | **Funcionalidade** | **Descrição** |
| --- | --- | --- |
| **1** | **Combinar dados** | Esta funcionalidade permite ao utilizador criar gráficos até três dimensões. Selecionando um facto, uma dimensão e uma dimensão temporal. Ou seja, é possível uma análise tridimensional sobe múltiplas perspetivas. |
| **2** | **Adicionar filtro** | Após a criação do gráfico, o utilizador pode aplicar filtros sobe o gráfico e analisar e explorar o subconjunto de dados apresentado, retirando conclusões que podem auxiliar o processo de tomada de decisão. |
| **3** | **Sugerir dados no processo de filtragem** | Durante o processo de filtragem é possível escolher uma dimensão ou facto para o filtro, assim como, uma sugestão de operadores a utilizar. Por exemplo, maior, menor igual, entre outros. Por fim, caso o filtro seja por uma dimensão um conjunto de dados será sugerido ao utilizador. Caso o filtro seja por um facto o utilizador pode inserir o valor que pretender (visto que os factos são mensuráveis e o utilizador pode escolher o valor que pretender). |
| **4** | **Criar vários gráficos de visualização** | A solução desenvolvida permite ao utilizador criar gráficos de barras, linhas, área, radar, *scatter*, *doughnut* e circular. |
| **5** | **Visualizar dados em dois formatos (numéricos e percentagem)** | Todos os gráficos apresentam uma tabela auxiliar para uma melhor leitura dos dados. A tabela pode ser vista com os valores absolutos ou em percentagem. |
| ***6*** | **Efetuar *Drill-Down*** | O utilizador pode, à exceção dos gráficos circulares e *doughnuts* fazer o processo de drill-down carregando num ponto do gráfico. Este processo auxilia o utlizador a ver os dados com maior detalhe em vários níveis hierárquicos. |
| ***7*** | **Efetuar *Rollup*** | O processo de rollup é o processo oposto do drill-down e auxilia o utilizador a analisar os dados de uma forma mais global. |
| **8** | **Guardar gráficos para a tomada de decisão** | Todos os gráficos criados podem ser guardados na página da Dashboard para futura análise. |
| ***9*** | ***Labels* dos gráficos curtas e objetivas** | Através dos gráficos criados o utilizador consegue rapidamente perceber o contexto de cada um devido às legendas intuitivas e de rápida leitura. |
| **10** | **Descarregar gráficos** | Possibilidade de descarregar os gráficos para uso particular do utilizador. |

# Conclusão

Este capítulo apresenta as principais conclusões do desenvolvimento da solução e de todo o processo de pesquisa inerente. De forma a verificar os riscos e o que foi feito para minimizar o impacto dos mesmos no desenrolar do projeto é apresentada uma tabela de riscos, assim como, uma subsecção onde é exposto o trabalho futuro esperado a partir da solução atual.

## Considerações finais

A transformação digital provocada pela quarta revolução industrial levou ao conceito da indústria das Cidades Inteligentes. As cidades estão em constante evolução e desenvolvimento, no sentido de maior sustentabilidade, coesão e mobilidade, com o fim, de melhorar a vida dos cidadãos através de uma digitalização dos sistemas.

Este trabalho tem como foco a evolução da indústria 4.0, o impacto que a mesma teve em diversas áreas do quotidiano e como surgiram as *Smart Cities*. Sistemas relativos à mobilidade também foram explorados, no sentido, de encontrar soluções ideais para a resolução de problemas como o congestionamento e poluição.

Relativamente ao trabalho desenvolvido, foram explorados fatores críticos que influenciam a tomada de decisão ao nível de transportes públicos e parques de estacionamento, e consequentemente, desenvolvido um formulário online para definir o grau de importância que cada um tem no quotidiano. Análises estatísticas foram realizadas, com os resultados do inquérito, de forma perceber o grau de importância correspondente a cada um. Desta forma, foram destacadas nove variáveis (de quarenta e nove) que apresentam um maior impacto para a população, nomeadamente, a percentagem de lotação e fluxo dos passageiros num transporte público, percentagem de parques de estacionamento públicos, tempo de procura de estacionamento, percentagem de pontualidade, tecnológica, pagamentos com sistemas eletrónicos, controlo de segurança, e, por fim, localização geográfica.

De forma a complementar o estudo, e identificar perfis associados às pessoas na tomada de decisão foi desenvolvido um método de *clustering*. A dispersão dos valores existentes comprova que as pessoas têm certeza de quais os critérios consideram mais importantes e quais priorizam na tomada de decisão. De acordo com os cenários desenvolvidos, ocupam um alto grau de importância as variáveis: custo do combustível, o preço médio do parque, a capacidade total, a percentagem de ocupação e o número de transferências entre meios de transporte. Relativamente aos perfis encontrados, é possível afirmar que existem pessoas mais seletivas no que toca as variáveis com maior grau de importância, outras mais extremistas e outras medianas. Este estudo é um ponto de partida para a recolha de dados e, consequentemente, para responder à necessidade de otimizar e construir modelos inteligentes que respondam às exigências da população e da cidade.

A partir dos dados adquiridos foi criado uma *data warehouse* para armazenar os dados, identificados seis KPIs e por fim desenvolvido um modelo OLAP e *dashboards* para visualização dos indicadores. Todos as etapas até a construção do modelo OLAP são fundamentais para que o resultado final seja de qualidade e útil. Sem dados bem estruturados e sem entender o contexto dos mesmos o processo torna-se ineficiente, deste modo, foi feita uma análise dos dados e consequentemente criado uma *data warehouse* que permitisse identificar KPIs mensuráveis. Através do *cube.js* foi desenvolvido um modelo OLAP que permite a análise multidimensional de dados e torna a tomada de decisão mais eficiente e eficaz, desenvolvendo estratégias que vão ao encontro às necessidades. Apesar da ferramenta cube.js ser bastante útil existem aspetos que podem ser melhorados ao nível da visualização. Posto isto, foi desenvolvida uma *dashboard* com a api do cube.js que permite uma serie de características fundamentais numa análise OLAP, tais como, *drill-down*, *rollup*, filtragem dos dados, possibilidade de guardar todos os gráficos pertinentes numa secção a parte, integração em soluções de software como serviço (Saas), entre outras.

Foi desenvolvido também uma análise exploratória de diversas APIs com o objetivo de averiguar se alguns dos critérios identificados, que posteriormente darão origem às variáveis a integrar no sistema inteligente, são viáveis. A obtenção de dados foi também um tema explorado no presente documento, visto ser importante perceber se os dados obtidos serão reais e úteis no sistema a desenvolver.

Concluindo, todo o trabalho descrito responde à questão de investigação “De que forma podemos melhorar a indústria das cidades através da utilização de sensores e Inteligência Artificial?”. Ou seja, através dos modelos e da solução desenvolvida é possível analisar diversos fatores que os cidadãos priorizam e, consequentemente, analisar os dados (provenientes dos sensores) de forma a retirar conclusões importantes para a tomada da decisão. Com a solução desenvolvida é possível averiguar ao detalhe todos os dados recolhidos de forma objetiva e útil, fazendo com que a transformação digital nas cidades agregue muito mais valor e vá de encontro às necessidades dos habitantes.

Este trabalho apresenta um importante contributo para o projeto ioCity, assim como, para a patente associada ao projeto ioscience sendo agregado à solução pré-construída.

Todo o trabalho desenvolvido envolve diversas componentes como *data science*, desenvolvimento web e *data visualization*. A primeira componente envolve todo o tratamento e análise dos dados de forma a retirar informações importantes e identificar padrões para a tomada de decisão. Para o desenvolvimento da solução foi necessário o desenvolvimento web (*api back-end* para definir rotas dos dados via mysql) e toda a componente de *data visualization* de forma que toda a informação esteja disponível para o utilizador de forma clara e vantajosa.

Através das metodologias aplicadas foi possível desenvolver um trabalho de investigação completo, seguindo todas as normas e etapas necessárias e desenvolvendo sentido critico na investigação científica que auxiliou todo o processo de desenvolvimento, nomeadamente, o planeamento do projeto, estado da arte, estudo de casos práticos e todo o processo de dados até à construção da solução final.

Todo o trabalho apresentado exigiu bastante pesquisa, tanto no estado da arte como no próprio desenvolvimento, uma vez que o tema teve de ser explorado ao detalhe e até ao momento não existia qualquer tipo de familiarização com as linguagens de programação utilizadas. As adversidades foram ultrapassadas e a solução apresentada cumpre todos os objetivos e requisitos que foram previamente definidos atingindo a finalidade e respondendo à grande questão de investigação proposta.

Deste modo, na Tabela 30 é apresentado um mapeamento entre o que foi inicialmente idealizado e os resultados atingidos. Por exemplo, o objetivo principal (questão de investigação) foi atingido com o desenvolvimento de uma *dashboard* analítico altamente escalável e de fácil adaptabilidade a outro conjunto de dados permitindo uma análise tridimensional que auxilia no processo de tomada de decisão. Ou seja, a utilização de sensores permite através da utilização de técnicas de OLAP obter indicadores em tempo real (conforme observado com a realização desta dissertação na secção 4).

Tabela 30 - Objetivos VS Resultados atingidos

| **Objetivos** | **Resultados Atingidos** |
| --- | --- |
| Questão de investigação “De que forma podemos melhorar a indústria das cidades através da utilização de sensores e Inteligência Artificial?” | *Dashboard* Analítico |
| Interpretar o problema e perceção da realidade | Formulário e Clustering |
| Recolher dados e interpretar os resultados obtidos (DCBD) | Dados provenientes da CMVNV |
| Extrair, transformar e carregar os dados fornecidos (ETL) | Processo ETL |
| Criar uma base de dados (SQL e noSQL) | Base de dados |
| Criar uma *data warehouse* (DW) | Data Warehouse |
| Desenvolvimento dos *KPI’s* | Indicadores e KPIs |
| Encontrar padrões entre os dados e retirar conclusões relevantes que sejam importantes para o apoio à tomada de decisão (*Data Science – DS)* | Modelo OLAP |

De forma a quantificar todo o trabalho demonstrado é apresentada uma lista com os principais resultados atingidos:

* 1 formulário e 1 estudo a nível das variáveis que influenciam a tomada de decisão entre a escolha do método de transporte a escolher;
* 3 cenários e 1 método de *clustering* estudados;
* 1 processo de ETL;
* 1 base de dados;
* 1 *data warehouse*, 4 dimensões e 2 tabelas de facto;
* 6 indicadores e 6 KPIs;
* 1 modelo OLAP;
* 1 camada de *frontend* para operacionalização dos dados;
* 1 Dashboard Analítico
* 1 API *backend*;

Diversos fatores organizacionais incrementam a aprendizagem e conhecimento. A realização do estado da arte permitiu obter uma visão clarificada e objetiva de como o projeto a desenvolver pode ser útil e eficaz no dia-a-dia. Assim como, todo o trabalho desenvolvido é fundamental para a transformação digital das cidades seguir as prioridades da população e agregar valor à cidade, promovendo o bem-estar social e estratégias de desenvolvimento sustentável, social e econômico.

## Trabalho Futuro

Esta dissertação foi desenvolvida no âmbito de um projeto de I&D, pelo que, há um conjunto de tarefas ainda a realizar para que o projeto ioCity seja finalizado de acordo com os requisitos. A solução apresentada é um ponto de partida para que a mesma seja integrada no produto final do ioCity. O projeto é dividido em várias fases, sendo que nesta fase todos os objetivos foram atingidos com sucesso. Relativamente às fases futuras, melhorias a nível de *data visualization* serão efetuadas para que o produto final esteja apto para uma utilização mais intuitiva.

## Tabela de Riscos

As decisões inerentes ao desenvolvimento de um projeto estão sempre associadas a riscos. Desta forma, foi elaborada uma tabela com os principais riscos associados ao desenvolvimento da presente dissertação. O objetivo desta é verificar os riscos e o que foi feito para minimizar o impacto dos mesmos no desenrolar do projeto. Cada risco é associado a uma probabilidade e impacto (na escala de 1 a 5), e consequentemente a uma magnitude (resultante da multiplicação da probabilidade e do impacto).

Desta forma, perceber-se-á o impacto que cada um teve na execução do projeto (Tabela 31).

Tabela 31 - Tabela de Riscos

| **ID** | **Descrição** | **Probabilidade (P)**  **[1…5]**  **[** | **Impacto**  **(I)**  **[1…5]** | **Magnitude**  **(P X I)** | **Ações de correção** |
| --- | --- | --- | --- | --- | --- |
| **1** | Falta de experiência a lidar com as ferramentas | 3 | 5 | 15 | Estudo atempado das ferramentas que foram utilizadas e, no caso de persistência de dúvidas, foi pedida ajuda ao docente responsável. |
| **2** | Atraso no desenvolvimento do projeto | 2 | 4 | 8 | Elaboração de um bom planeamento do projeto, estabelecido desde o início as tarefas. |
| **3** | Objetivos de negócio e DM mal definidos. Necessidade de mudança dos requisitos do projeto | 3 | 5 | 15 | Reunião com o orientador para reestruturação do projeto, para que este fique de acordo com os novos requisitos. |
| **4** | Má qualidade dos dados | 2 | 4 | 8 | Análise dos dados pormenorizada e exaustiva para garantir a coerência e qualidade nas fases seguintes. |
| **5** | Perda de dados | 1 | 3 | 3 | Pretende-se mitigar a perda de dados com a manutenção de várias cópias de segurança dos mesmos. |

REFERÊNCIAS

1. 4DFWD: DATA-DRIVEN 3D PRINTED PERFORMANCE TECHNOLOGY DESIGNED TO MOVE YOU FORWARD. (2021, maio 5). Adidas News Site | Press Resources for All Brands, Sports and Innovations. https://news.adidas.com/running/4dfwd--data-driven-3d-printed-performance-technology-designed-to-move-you-forward/s/514baddb-1029-4686-abd5-5ee3985a304a
2. Ahad, M. A., Paiva, S., Tripathi, G., & Feroz, N. (2020). Enabling technologies and sustainable smart cities. Sustainable Cities and Society, 61, 102301. https://doi.org/10.1016/j.scs.2020.102301
3. Ahmad, S., Miskon, S., Alabdan, R., & Tlili, I. (2020). Towards Sustainable Textile and Apparel Industry: Exploring the Role of Business Intelligence Systems in the Era of Industry 4.0. Sustainability, 12(7), Art. 7. https://doi.org/10.3390/su12072632
4. Aletà, N. B., Alonso, C. M., & Ruiz, R. M. A. (2017). Smart Mobility and Smart Environment in the Spanish cities. Transportation Research Procedia, 24, 163–170. https://doi.org/10.1016/j.trpro.2017.05.084
5. Altus. (2019). Conheça os nove pilares da Indústria 4.0. https://www.altus.com.br/post/212/conheca-os-nove-pilares-da-industria-4-0-e-sua-relevancia-para-a-atividade-industrial
6. Alves, G. (2018, dezembro 13). Unsupervised learning with K-means. Neuronio. https://medium.com/neuronio/unsupervised-learning-with-k-means-3eaa0666eebf
7. Anthony Jnr, B. (2021). Managing digital transformation of smart cities through enterprise architecture – a review and research agenda. Enterprise Information Systems, 15(3), 299–331. https://doi.org/10.1080/17517575.2020.1812006
8. Antonio, N., Fornazin, M., & Araujo, R. (2018). Metodologia de Pesquisa de Estudo de Caso em Sistemas de Informação (pp. 41–67). https://doi.org/10.5753/sbc.7.2
9. API Picker | Google Maps Platform. (sem data). Google Developers. Obtido 18 de maio de 2022, de https://developers.google.com/maps/documentation/api-picker
10. Aprilliant, A. (2021, julho 21). The k-modes as Clustering Algorithm for Categorical Data Type. Geek Culture. https://medium.com/geekculture/the-k-modes-as-clustering-algorithm-for-categorical-data-type-bcde8f95efd7
11. Aran et. al. (2020, junho 2). Comme des Machines—Future Perfect. https://www.talentoabordo.com/en/design/comme-des-machines
12. Aurigi, A., Willis, K., & Melgaco, L. (2016). From «digital» to «smart»: Upgrading the city. Proceedings of the 3rd Conference on Media Architecture Biennale, 1–4. https://doi.org/10.1145/2946803.2946813
13. Batty, M., Axhausen, K. W., Giannotti, F., Pozdnoukhov, A., Bazzani, A., Wachowicz, M., Ouzounis, G., & Portugali, Y. (2012). Smart cities of the future. The European Physical Journal Special Topics, 214(1), 481–518. https://doi.org/10.1140/epjst/e2012-01703-3
14. Borlido, D. J. A. (2017). Indústria 4.0 – Aplicação a Sistemas de Manutenção. 77.
15. Bria, F. (2018). Smart City 3.0 – pergunte a Barcelona sobre a próxima geração de cidades inteligentes. URBAN HUB. https://www.urban-hub.com/pt-br/cities/barcelona-deixa-a-cidade-ainda-mais-inteligente/
16. Builder, P. (2020, abril 15). Entenda a metodologia Scrum e como utilizá-la na empresa. Project Builder. https://www.projectbuilder.com.br/blog/o-que-e-scrum/
17. CMVNF. (2019). Famalicão apresenta-se no Portugal Smart Cities Summit. https://www.famalicao.pt/famalicao-apresenta-se-no-portugal-smart-cities-summit
18. Computação em nuvem. (2021). Em Wikipédia, a enciclopédia livre. https://pt.wikipedia.org/w/index.php?title=Computa%C3%A7%C3%A3o\_em\_nuvem&oldid=61805595
19. Connected Traffic Management Solutions. (sem data). All Traffic Solutions. Obtido 22 de fevereiro de 2022, de https://www.alltrafficsolutions.com/our-solutions/
20. Costa, A. (2008). MANUAL DE PLANEAMENTO DAS ACESSIBILIDADES E DA GESTÃO VIÁRIA. Comissão de Coordenação e Desenvolvimento Regional do Norte (CCDR-N) Ministério do Ambiente e do Ordenamento do Território. https://www.imt-ip.pt/sites/IMTT/Portugues/Planeamento/DocumentosdeReferencia/Planeamento-Acessibilidades-Gestao-Viaria/Documents/13TransPub\_AF.pdf
21. Cruz, E. (2011). Design Science Research em Sistemas de Informação. unpublished.
22. Ctmucg. (2017). Sensores que contam passageiros em tempo real nas portas são instalados em ônibus da Metra e ideia recebe prêmio internacional – CTMU. http://ctmucg.com/sensores-que-contam-passageiros-em-tempo-real-nas-portas-sao-instalados-em-onibus-da-metra-e-ideia-recebe-premio-internacional/
23. cube. (2022). Cube—Headless BI for Building Data Applications. Cube.Dev. https://cube.dev/
24. Das, S. (2019). A Novel Parking Management System, for Smart Cities, to save Fuel, Time, and Money. 2019 IEEE 9th Annual Computing and Communication Workshop and Conference (CCWC), 0950–0954. https://doi.org/10.1109/CCWC.2019.8666537
25. Dilax. (sem data). Automatic Passenger Counting (APC): Sensors & Systems. Obtido 26 de fevereiro de 2022, de https://www.dilax.com/en/products/automatic-passenger-counting
26. DLS 3D. (2022). Carbon. https://www.carbon3d.com/carbon-dls-technology/
27. ECO. (2020, julho 17). Pode a indústria têxtil ser inovadora e sustentável? We Sustain diz que sim. ECO. https://eco.sapo.pt/2020/07/17/pode-a-industria-textil-ser-inovadora-e-sustentavel-we-sustain-diz-que-sim/
28. ECO. (2021, novembro 4). Smartex.ai de António Rocha vence prémio Pitch da Web Summit 2021. ECO. https://eco.sapo.pt/2021/11/04/smartex-ai-de-antonio-rocha-vence-premio-pitch-da-web-summit-2021/
29. Egreja, L. (2019, março 6). Por uma indústria mais moderna • Painel Logistico. Painel Logistico. https://www.painellogistico.com.br/﻿por-uma-industria-mais-moderna/
30. Eventbrite. (2022). Eventbrite. https://www.eventbrite.pt
31. Expo World Congress. (2022). Smart City Expo World Congress 2022. https://www.smartcityexpo.com/
32. Fernandes, S. A. S. (2019). Ferramentas de gestão e desenvolvimento de competências na transição para a Indústria 4.0: O caso de uma empresa na indústria automóvel. https://comum.rcaap.pt/handle/10400.26/29602
33. Ferreira. (2020, setembro 17). Realidade aumentada na indústria: Quais os ganhos e aplicações. ABii. https://abii.com.br/realidade-aumentada-na-industria-quais-os-ganhos-e-aplicacoes/
34. IAPMEI. (2021, dezembro 2). IAPMEI - Agendas Mobilizadoras para a Inovação Empresarial | Propostas para a Economia do Futuro. https://www.iapmei.pt/EVENTOS/Agendas-Mobilizadoras-para-a-Inovacao-Empresarial.aspx
35. Iberdrola, I. (2022). Indústria 4.0: Que tecnologias marcarão a Quarta Revolução Industrial? Iberdrola. https://www.iberdrola.com/inovacao/quarta-revolucao-industrial
36. Impacto da produção. (2020, dezembro 29). https://www.europarl.europa.eu/news/pt/headlines/society/20201208STO93327/o-impacto-da-producao-e-dos-residuos-texteis-no-ambiente-infografia
37. Indústria 4.0. (2021a). https://www.siembra.com.br/noticias/descubra-quais-sao-os-pilares-da-industria-40/
38. Indústria 4.0. (2021b). Em Wikipédia, a enciclopédia livre. https://pt.wikipedia.org/w/index.php?title=Ind%C3%BAstria\_4.0&oldid=62000439
39. Kevin Ashton. (2022). Em Wikipedia. https://en.wikipedia.org/w/index.php?title=Kevin\_Ashton&oldid=1065838499
40. Lab, O. (2021). Smart Cities em Portugal? Sim, claro. No Centro de Portugal. Observador. https://observador.pt/2020/10/02/smart-cities-em-portugal-sim-claro-no-centro-de-portugal/
41. Lima, F. (2018). INE, Mobilidade e funcionalidade do território nas Áreas Metropolitanas do Porto e de Lisboa 2017, Instituto Nacional de Estatística, Lisboa, 2018.
42. Litman, T. A. (2021). Parking Management. 32.
43. López, I., Ortega, J., & Pardo, M. (2020). Mobility Infrastructures in Cities and Climate Change: An Analysis Through the Superblocks in Barcelona. Atmosphere, 11(4), Art. 4. https://doi.org/10.3390/atmos11040410
44. Marcel Ogando. (2020, julho 13). Milênio Bus no canal Futura, na série inovação. https://www.youtube.com/watch?v=m-iV5Iaugt4
45. Marconi, bruna M., & Broega, A. C. (2021, novembro 24). ESTUDO DE ALTERNATIVAS SUSTENTÁVEIS PARA O SECTOR TÊXTIL PORTUGUÊS. Anais do VIII SDS 2021. Simpósio de Design Sustentável. https://doi.org/10.5380/8sds2021.art55
46. Meta. (2022). App Events API - Marketing API - Documentation. Facebook for Developers. https://developers.facebook.com/docs/marketing-api/app-event-api/
47. Neon. (sem data). A evolução dos carros: Da sua criação até os dias atuais—Neon Seguros. Obtido 11 de fevereiro de 2022, de https://www.neonseguros.com.br/curiosidades-automotivas/a-evolucao-dos-carros-da-sua-criacao-ate-os-dias-atuais/6427/
48. Neto, C., Masle-Farquhar, E., Ritz, I., Levy, I., Silva, P. D., Costa, N., & Garcia, R. (2016). O CONTEÚDO EDITORIAL DA REVISTA É DA TOTAL RESPONSABILIDADE DA PLATAFORMA PORTUGUESA DAS ONGD. 26.
49. Neves et. al. (2022). Revolução industrial: Causas, fases, consequências, resumo. Mundo Educação. https://mundoeducacao.uol.com.br/historiageral/revolucao-industrial-2.htm
50. OCI. (2022). What is a Data Warehouse? https://www.oracle.com/pt/database/what-is-a-data-warehouse/
51. Ogando et. al. (2021). Milênio Bus no canal Futura, na série inovação—YouTube. https://www.youtube.com/watch?v=m-iV5Iaugt4
52. Ogando, M. (sem data). Inteligência em Transporte Coletivo. Obtido 24 de fevereiro de 2022, de http://www.mileniobus.com.br/
53. Olivo, A., Maternini, G., & Barabino, B. (2019). Empirical Study on the Accuracy and Precision of Automatic Passenger Counting in European Bus Services. The Open Transportation Journal, 13(1). https://doi.org/10.2174/1874447801913010250
54. Onesmus Mbaabu. (2021). MOLAP vs ROLAP vs HOLAP in Online Analytical Processing (OLAP). Engineering Education (EngEd) Program | Section. https://www.section.io/engineering-education/molap-vs-rolap-vs-holap/
55. owm. (2022). OpenWeatherMap API - Free Weather Data (for Developers). https://rapidapi.com/community/api/open-weather-map
56. Passos, M. (2021, novembro 9). Famalicão in Portugal Smart Cities Summit. https://portal.vilanovadefamalicao.org/famalicao-in-portugal-smart-cities-summit
57. Patil, V. (2021, agosto 23). Clustering and profiling customers using k-Means. Analytics Vidhya. https://medium.com/analytics-vidhya/clustering-and-profiling-customers-using-k-means-9afa4277427
58. Prasanna. (2022, janeiro 16). Smart City Advantages and Disadvantages | List of All Benefits and Drawbacks of Smart Cities. A Plus Topper. https://www.aplustopper.com/smart-city-advantages-and-disadvantages/
59. PrimeStone. (2020, abril 2). ADVANTAGES AND DISADVANTAGES OF SMART CITIES. PrimeStone. https://primestone.com/en/advantages-and-disadvantages-of-smart-cities/
60. PURKAYASTHA. (2019, abril 9). Top 10 Best Events APIs: EventBrite, Meetup, Ticketmaster, and others. Rakuten RapidAPI Blog. https://blog.api.rakuten.net/top-10-best-events-apis/
61. Rahim, Md. A., Rahman, Md. A., Rahman, M. M., Asyhari, A. T., Bhuiyan, Md. Z. A., & Ramasamy, D. (2021). Evolution of IoT-enabled connectivity and applications in automotive industry: A review. Vehicular Communications, 27, 100285. https://doi.org/10.1016/j.vehcom.2020.100285
62. RapidAPI. (2022). Top 8 Best Free Weather APIs (2021) [25+ Reviewed]. The Last Call - RapidAPI Blog. https://rapidapi.com/blog/access-global-weather-data-with-these-weather-apis/
63. RapidAPI Staff. (2020, janeiro 7). Top 8 Best Free Weather APIs (2021) [25+ Reviewed]. The Last Call - RapidAPI Blog. https://rapidapi.com/blog/access-global-weather-data-with-these-weather-apis/
64. Rodrigues, I. (2020, fevereiro 20). CRISP-DM methodology leader in data mining and big data. Medium. https://towardsdatascience.com/crisp-dm-methodology-leader-in-data-mining-and-big-data-467efd3d3781
65. RPCI. (2018, julho 25). Fórum das Cidades. https://www.forumdascidades.pt/content/rede-portuguesa-de-cidades-inteligentes
66. Sara-Andreia-Silva-Fernandes.pdf. (sem data). Obtido 11 de fevereiro de 2022, de https://comum.rcaap.pt/bitstream/10400.26/29602/1/Sara-Andreia-Silva-Fernandes.pdf
67. Shyam. (2019, abril 9). Top 10 Best Events APIs: EventBrite, Meetup, Ticketmaster, and others. Rakuten RapidAPI Blog. https://blog.api.rakuten.net/top-10-best-events-apis/
68. Smartex. (2022). Smartex. https://www.smartex.ai/
69. Sousa, J. N. L. de. (2020). Especificação de requisitos para o desenvolvimento de um sistema de apoio à decisão para gestão de transportes públicos intermodais. https://repositorio-aberto.up.pt/handle/10216/127741
70. Stevens, C. L. (2020). Seoul: A World Class Smart City. https://www.cta.tech/Resources/i3-Magazine/i3-Issues/2020/January-February/Seoul-A-World-Class-Smart-City
71. Tecnologia 3D. (2020, junho 10). Tendências em Moda Íntima, Fitness e mais! | Blog da Zanotti Elásticos. https://zanotti.com.br/blog/tecnologia-3d-e-sua-aplicacao-na-industria-textil/
72. Vallianatos, M. (2015, junho 16). Uncovering the Early History of “Big Data” and the “Smart City” in Los Angeles. Boom California. https://boomcalifornia.org/2015/06/16/uncovering-the-early-history-of-big-data-and-the-smart-city-in-la/
73. Vantagens e Desvantagens. (2017). VANTAGENS E DESVANTAGENS de se usar o TRANSPORTE PRIVADO. http://transitoitumbiara.blogspot.com/2017/08/vantagens-e-desvantagens-de-se-usar-o\_19.html
74. Wang, T. (2019). Smart City Case Study.
75. Weather. (sem data). Historical Weather API, and Forecast API in one package! Obtido 18 de maio de 2022, de https://www.weatherbit.io/features
76. Yu, B. C. Y. (2018). FACULDADE DE TECNOLOGIA DE AMERICANA Curso Superior de Tecnologia em Produção Têxtil. 47.

# Apêndices

**Estudo de critérios de apoio à tomada de decisão [ioCity]**

A possibilidade de os cidadãos terem acesso a bens e serviços melhora a sua qualidade de vida. Irão ser expostos critérios decisivos para a tomada de decisão, ou seja, fatores que influenciam a escolha entre usar transportes públicos ou veículos individuais. Este questionário enquadra-se no projeto ioCity - A cidade não espera (NORTE-01-0247-FEDER-045397) que visa responder à questão de investigação: De que forma podemos melhorar a indústria das cidades através da utilização de sensores e Inteligência Artificial? Numa primeira secção irão ser abordados fatores relativos aos transportes públicos, e numa segunda parte fatores relativos a veículos individuais.

Link de acesso ao questionário desenvolvido.

<https://survey1.iocity.research.iotech.pt>

**Artigo “*Opinion clustering about mobility decisions - A practical case study*”**

Artigo que descreve toda a investigação sobre quais as variáveis que influenciam a tomada de decisão foram apresentadas. Este artigo foi também apresentado na conferência internacional ARTIIS 2022[[5]](#footnote-5).

Estado: Em publicação

Local de publicação: ARTIIS



**Artigo “*Intelligent dashboards for car parking flow monitoring*”**

Artigo que tem como objetivo demonstrar a importância de um modelo OLAP na tomada de decisão e como todo o fluxo até a construção do modelo é fundamental para um modelo útil e eficaz.

Estado: Em publicação

Local de publicação: ARTIIS

Text, letter

Description automatically generated

**Artigo “*An analysis of Smart Cities Initiatives in Portugal*”**

Artigo que demonstra o estado atual das *smart cities* em Portugal e o tipo de iniciativas que tem vindo a ser desenvolvidas.

Estado: Em publicação

Local de publicação: ARTIIS

Text

Description automatically generated

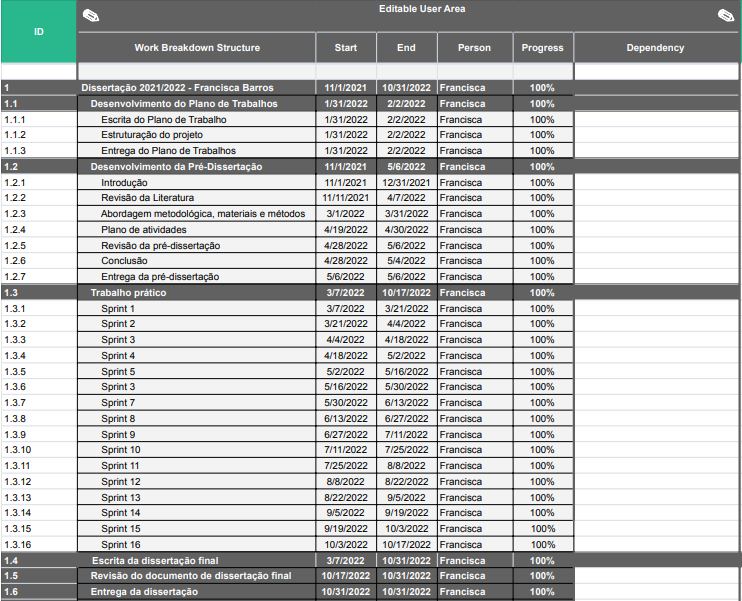
**Planeamento**

Este projeto terá a duração de aproximadamente 13 meses (outubro de 2021 até outubro de 2022).

Na Figura 41 é possível visualizar as tarefas subjacentes à realização da presente dissertação, assim como a data de início e fim das *milestones* e o progresso das mesmas. Existem 7 atividades principais, o desenvolvimento do plano de trabalhos, da pré-dissertação, do trabalho prático, e por fim, escrita, revisão, entrega e apresentação da dissertação.

A componente prática está dividida por sprints quinzenais (de acordo com a metodologia SCRUM) para realizar: ETL, DW, interpretação do problema, perceção da realidade e análise dos riscos, KPI’s, base de dados - SQL e noSQL, DCBD, *data science* e, por fim, revisão do material desenvolvido. Na imagem é possível visualizar o diagrama de *gantt* (Figura 41).

Figura 42 - Diagrama de Gantt.



1. ioCity (https://iocity.research.iotech.pt/) [↑](#footnote-ref-1)
2. b-smart (http://b-smart.famalicao.pt/) [↑](#footnote-ref-2)
3. https://www.artiis.org/ [↑](#footnote-ref-3)
4. *b-smart* (http://b-smart.famalicao.pt/) [↑](#footnote-ref-4)
5. https://www.artiis.org/ [↑](#footnote-ref-5)