



INSTITUTO
UNIVERSITÁRIO
DE LISBOA

Movimentação e Comportamento de pessoas nas zonas de diversão noturna de Lisboa com base em dados reais

Carolina Gabriel Albuquerque

Mestrado em Métodos Analíticos para a Gestão (Business Analytics)

Orientador:

Doutor João Carlos Amaro Ferreira, Professor Auxiliar (com agregação),
ISCTE-Instituto Universitário de Lisboa

Co-Orientador:

Professor Bruno Alexandre Francisco, Professor Assistente Convidado,
ISCTE-Instituto Universitário de Lisboa

Setembro, 2024

Departamento de Métodos Quantitativos para Gestão e Economia

**Movimentação e Comportamento de pessoas nas zonas de
diversão noturna de Lisboa com base em dados reais**

Carolina Gabriel Albuquerque

Mestrado em Métodos Analíticos para a Gestão (Business Analytics)

Orientador:

Doutor João Carlos Amaro Ferreira, Professor Auxiliar (com agregação),
ISCTE-Instituto Universitário de Lisboa

Co-Orientador:

Professor Bruno Alexandre Francisco, Professor Assistente Convidado,
ISCTE-Instituto Universitário de Lisboa

Setembro, 2024

Dedico esta dissertação aos meus pais, que estiveram sempre presentes em todas as etapas importantes da minha vida.

Agradecimentos

Quero agradecer a todas as pessoas que, de diferentes formas, me apoiaram ao longo deste percurso.

Aos meus amigos de Lisboa, que são uma constante na minha vida, proporcionando-me momentos de diversão e descontração e dando-me força de vontade para continuar a lutar pelos meus objetivos.

Aos meus amigos de Viseu, que me conhecem desde sempre e que acompanharam com orgulho as minhas conquistas. O vosso carinho e apoio foram essenciais para o meu desenvolvimento, tanto a nível pessoal como profissional.

À minha família, por todo o apoio incondicional ao longo da minha vida. Sempre me apoiaram e acreditaram em mim, permitindo-me chegar até aqui.

Por último, gostaria de agradecer aos meus orientadores que, com paciência e sabedoria, me guiaram no desenvolvimento desta dissertação. A vossa dedicação e apoio foram essenciais para o sucesso deste trabalho.

A todos, o meu sincero obrigada!

Resumo

Os espaços públicos dedicados à vida noturna têm um impacto relevante na dinâmica urbana, tornando crucial a sua gestão para atender aos interesses de utilizadores, residentes e comerciantes. Este estudo visa analisar a movimentação e o comportamento nas zonas de diversão noturna de Lisboa, utilizando dados anonimizados de dispositivos móveis fornecidos por uma operadora, incluindo dispositivos em *roaming* devido à relevância turística e impacto económico dos visitantes. É também interessante observar como estes movimentos humanos afetam o ambiente sonoro exterior, especificamente como os valores do parâmetro de nível sonoro muda ao longo do tempo.

A metodologia aplicada foi o CRISP-DM, guiando todo o processo de desenvolvimento de um *dashboard* em *Power BI*, e a seleção de indicadores-chave de desempenho (KPI) fundamentados na literatura e validados por especialistas. Como resultados, esta análise identifica padrões de movimento, avalia transferências entre zonas e relaciona esses movimentos com os níveis de ruído ambiental.

Por fim, a presente dissertação contribui para um sistema de apoio à decisão, proporcionando à LxDataLab uma ferramenta que facilita a análise e gestão de dados, permitindo uma compreensão mais profunda dos padrões de mobilidade urbana em áreas de lazer noturno.

Palavras-chave: Mobilidade urbana; Cidades Inteligentes; *Dashboard*; Análise do Comportamento; Vida Noturna; Lisboa

Classificação JEL: M10 e M15

Abstract

Public spaces dedicated to nightlife have a relevant impact on urban dynamics, making it crucial to manage them to meet the interests of users, residents and traders. This study aims to analyze movement and behavior in Lisbon's nightlife areas, using anonymized data from mobile devices provided by an operator, including *roaming* devices due to the tourist relevance and economic impact of visitors. It is also interesting to observe how these human movements affect the outdoor sound environment, specifically how the sound level parameter values change over time.

The methodology applied was CRISP-DM, which guided the entire process of developing a *dashboard* in *Power BI* and selecting key performance indicators (KPIs) based on the literature and validated by experts. As a result, this analysis identifies movement patterns, evaluates transfers between zones and relates these movements to environmental noise levels.

Finally, this dissertation contributes to a decision support system, providing LxDataLab with a tool that facilitates data analysis and management, enabling a deeper understanding of urban mobility patterns in night-time leisure areas.

Keywords: Urban Mobility; Smart Cities; Dashboard; Behavior Analysis; Nightlife; Lisbon

JEL Classification: M10 e M15

Índice

Agradecimentos.....	iii
Resumo.....	v
Abstract	vii
Índice.....	ix
Índice de Figuras	xi
Índice de Tabelas.....	xiii
Lista de abreviaturas, acrónimos e siglas	xv
1. Introdução	1
1.1. Tema e a sua delimitação.....	1
1.2. Problema e Questão de Investigação	2
1.3. Objetivos e Contributos.....	3
1.4. Abordagem metodológica.....	4
1.5. Estrutura e organização da dissertação	4
2. Revisão da Literatura	6
2.1. Mobilidade Urbana e Utilização de Dados de Dispositivos Móveis	6
2.1.2. Previsão de Posições e Regularidade Espaço-temporal.....	8
2.1.4. Turismo	8
2.2. <i>Dashboards</i>	9
2.2.1. Conceito de <i>dashboard</i>	9
2.2.2. Indicadores e métricas nos <i>dashboards</i>	10
2.2.3. Tipos de <i>dashboard</i>	11
2.2.4. Características de um <i>dashboard</i>	12
2.2.5. Tipos de gráficos mais adequados	13
2.2.6. Vantagens.....	14
2.2.7. Construção e avaliação de um <i>dashboard</i>	15
2.2.8. <i>Dashboards</i> de mobilidade humana.....	17
3. Metodologia	18
3.1. Compreensão do Negócio.....	18
3.2. Compreensão dos dados	19

3.3. Preparação dos dados.....	23
3.4. Modelação	27
3.5. Avaliação	30
3.6. Implementação.....	30
4. Resultados e discussão	33
4.1. <i>Dashboard</i> de monitorização da mobilidade urbana nas zonas de diversão noturna de Lisboa	33
4.1.1. Objetivo 1: Identificar as zonas mais movimentadas de Lisboa.....	33
4.1.2. Objetivo 2: Analisar o Comportamento da População em Eventos Festivos	37
4.1.3. Objetivo 3: Avaliar as Possíveis Transferências dos Utilizadores Entre as Zonas de Diversão Noturna	41
4.1.4. Objetivo 4: Relacionar os movimentos das pessoas com os níveis de ruído registados nos sensores ambientais nas áreas de estudo consideradas	43
4.2. Avaliação do <i>dashboard</i>	48
5. Conclusões	51
5.1. Contributos	53
5.2. Limitações	54
5.3. Recomendações e trabalho futuro.....	55
Referências Bibliográficas	57
Anexos.....	61

Índice de Figuras

Figura 2.1: Ordem de leitura de um <i>dashboard</i>	12
Figura 2.2: Etapas de construção do <i>dashboard</i>	16
Figura 3.1: Etapas da preparação de dados	23
Figura 3.2: <i>Grids</i> selecionadas para a análise	24
Figura 4.1: <i>Dashboard</i> para monitorizar áreas de diversão noturna mais movimentadas de Lisboa	33
Figura 4.2: Filtros disponíveis no <i>dashboard</i>	34
Figura 4.3: Indicadores de controlo (Nº Médio de Terminais por <i>Grid</i> , Nº de Terminais Portugueses e Nº de Terminais Estrangeiros)	34
Figura 4.4: Concentração de Terminais Portugueses e Estrangeiros	35
Figura 4.5: Concentração de Pessoas por Período do Dia.....	36
Figura 4.6: Densidade Populacional por Tipo de Dia	36
Figura 4.7: <i>Dashboard</i> de monitorização das Pessoas em Eventos Festivos.....	37
Figura 4.8: Concentração Populacional antes, durante e após as JMJ	38
Figura 4.9: Nº de Terminais por mês	39
Figura 4.10: Concentração de Pessoas durante as JMJ	40
Figura 4.11: Densidade Populacional por Tipo de Estação	40
Figura 4.12: <i>Dashboard</i> de monitorização de transferências de utilizadores	41
Figura 4.13: Transferência entre áreas de diversão noturna.....	42
Figura 4.14: Densidade Populacional por Período do Dia e por Freguesia	43
Figura 4.15: Localização dos sensores com limitadores de ruído.....	44
Figura 4.16: <i>Dashboard</i> de monitorização da concentração de pessoas e dos níveis de ruído	45
Figura 4.17: Níveis de Ruído por Freguesia	46
Figura 4.18: Relação da concentração de pessoas com os níveis de ruído registados	46
Figura 4.19: Níveis de Ruído por Hora e por Freguesia	47
Figura 4.20: Valor Máximo do Ruído antes, durante e após as JMJ.....	48
Figura 0.1: Tipos de <i>dashboard</i>	61
Figura 0.2: Gráfico com detalhe quando utilizador faz <i>mouse over</i>	66

Índice de Tabelas

Tabela 3.1: Descrição dos campos do Conjunto 1	20
Tabela 3.2: Descrição dos campos do Conjunto 2	21
Tabela 3.3: Descrição dos campos do Conjunto 3	22
Tabela 3.4: Descrição dos campos do Conjunto 4	22
Tabela 3.5: Elementos visuais desenvolvidos em Power BI.....	29
Tabela 3.6: Questionário de avaliação do <i>dashboard</i>	30
Tabela 4.1: Indicadores dos sensores de ruído	44
Tabela 4.2: Avaliação do <i>dashboard</i> por utilizadores.....	49
Tabela 0.1: Avaliação do <i>Dashboard</i> quanto ao <i>design</i>	61
Tabela 0.2: Avaliação do <i>Dashboard</i> quanto à análise	62
Tabela 0.3: Avaliação do <i>Dashboard</i> quanto à entrega	62
Tabela 0.4: Avaliação do <i>Dashboard</i> quanto à administração.....	63
Tabela 0.5: Avaliação do <i>Dashboard</i> quanto à infraestrutura	63
Tabela 0.6: Fórmula de cálculo (em linguagem DAX) para a criação da coluna "WeekendOrHoliday"	64
Tabela 0.7: Fórmula de cálculo (em linguagem DAX) para a criação da coluna "TipodeDia"	64
Tabela 0.8: Fórmula de cálculo (em linguagem DAX) para a criação da coluna "NTotal_Telemoveis".....	65
Tabela 0.12: Fórmulas de cálculo (em linguagem DAX) para criação das tabelas "Telemoveis_Por_Freguesia", "Tabela_Ranking", "Tabela_Periodos" e "Tabela_Horas"	65
Tabela 0.11: Fórmula de cálculo (em linguagem DAX) para a criação das métricas "Densidade_Media_DiaDaSemana", "Densidade_Media_FimDeSemana", "Densidade_Media_Feriado" e "Densidade_Media_Vespera_Feriado"	65
Tabela 0.10: Fórmula de cálculo (em linguagem DAX) para a criação da coluna "Tipo_Estacao".....	65
Tabela 0.9: Fórmula de cálculo (em linguagem DAX) para a criação das métricas "Densidade_Media_Verão" e "Densidade_Media_Outras_Estacoes"	65

Lista de abreviaturas, acrónimos e siglas

BPM - *Business Performance Management*

BSC - *Balanced Scorecard*

CML - Câmara Municipal de Lisboa

CRISP-DM - *CRoss-Industry Standard Process for Data Mining*

EIS - *Executive Information Systems*

GPI - *Global Peace Index*

GPS - *Global Positioning System*

IoT - *Internet of Things* (Internet das Coisas)

JMJ - Jornadas Mundiais da Juventude

KPI - *Key Performance Indicator*

WKT - *Well Known Text*

CAPÍTULO 1

1. Introdução

1.1. Tema e a sua delimitação

De acordo com o *Institute for Economics and Peace* (2024), está comprovado que Lisboa é uma das cidades mais seguras da Europa, apresentando elevados níveis de *Global Peace Index* (GPI)¹. Desta forma, torna-se uma mais-valia do ponto de vista turístico.

Tal como todos aqueles que habitam na cidade de Lisboa esperam de Lisboa tranquilidade, segurança e harmonia, também os turistas o fazem. A cidade lisboeta, cada vez mais, promove métodos de ordenamento do território, de forma a integrar adequadamente toda a população (Assembleia Municipal de Lisboa, 2018).

As políticas municipais de ordenamento do território, bem como os instrumentos de gestão que as executam, assentam fundamentalmente no planeamento da mobilidade. Como qualquer grande metrópole, o planeamento necessário para Lisboa é extremamente difícil devido à sua complexidade e à necessidade de recorrer a um conjunto de abordagens eficazes, informadas e rigorosas (Câmara Municipal de Lisboa, s.d.).

O conceito de mobilidade também se torna importante. A mobilidade humana, o movimento dos seres humanos no espaço e no tempo, reflete as características espaço-temporais do comportamento humano. Com a análise de *Big Data*, a investigação sobre mobilidade humana pode ser utilizada para facilitar o desenvolvimento de cidades inteligentes (Zhang et al., 2021).

A Câmara Municipal de Lisboa tem uma visão estratégica para a mobilidade de Lisboa até 2030 (Câmara Municipal de Lisboa, 2020). Lisboa assumiu o compromisso essencial de optar por uma mobilidade mais inteligente, que garanta a máxima acessibilidade com consequências mínimas. A cidade já tomou uma série de medidas para incentivar uma mobilidade mais sustentada, no entanto, precisa de uma estratégia integrada para garantir a coerência das opções escolhidas e criar uma visão clara (Câmara Municipal de Lisboa, s.d.).

¹ O *Global Peace Index* (GPI) é um índice que mede a paz em países e regiões ao redor do mundo. Ele é desenvolvido pelo *Institute for Economics and Peace* (IEP) e é a principal medida mundial de paz global.

Pretende-se que esta investigação crie uma abordagem baseada em dados que seja capaz de receber dados anonimizados do setor das telecomunicações e criar visualizações sobre a mobilidade dos cidadãos na cidade, utilizando as ferramentas analíticas adequadas, para extrair valor significativo dos dados. Pretende-se não só apresentar as visualizações, como também utilizar essas ferramentas para interpretar os dados de forma eficaz, identificando padrões e tendências na mobilidade dos cidadãos e na vida noturna da cidade.

O objetivo principal é desenvolver um sistema de apoio à decisão baseado em factos e dados reais e não em informações não sustentadas (Francisco et al., 2022). Desta forma, delineando um novo método de avaliação dos comportamentos, atividades e movimentos das pessoas para fins de planeamento e monitorização, o município poderá obter informações valiosas sobre os movimentos das pessoas nas áreas de diversão noturna.

Atualmente, com o aumento da utilização de dispositivos inteligentes, as operadoras móveis conseguem recolher grandes quantidades de dados relacionados à localização geográfica dos utilizadores, que podem ser extremamente úteis para análises de mobilidade urbana (Geoapify, 2023).

A análise desses dados pode fornecer informação valiosa sobre a regularidade espacial e temporal dos movimentos das pessoas, permitindo uma previsão mais precisa dos fluxos de mobilidade (Estes, 2016). Isto é especialmente relevante para otimizar serviços públicos e privados, como transporte e segurança, além de ajudar a direcionar ações estratégicas em áreas com grande concentração de pessoas.

Contudo, é essencial garantir que esses dados sejam tratados com cuidado, respeitando as regulamentações de privacidade e segurança, dado o potencial sensível da informação relacionada à localização (Geoapify, 2023; Estes, 2016).

1.2. Problema e Questão de Investigação

Os investigadores, devido a vários problemas, como ruído noturno, manutenção insuficiente, falta de segurança, são incentivados a estudar vários temas de forma a criar modelos precisos e eficazes para caracterizar a distribuição espaço-temporal dos indivíduos e da população como um todo, considerando os seres humanos uma fonte de informação devido aos seus dispositivos móveis e consequentes movimentos (Ali Haidery et al., 2020).

Sabe-se que os problemas nas zonas de maior concentração de estabelecimentos de diversão noturna, como o Bairro Alto, Santos e Cais do Sodré, continuam a existir. A população queixa-se do ruído, da limpeza urbana e da falta de segurança (Assembleia Municipal de Lisboa, 2018).

A análise dos dados disponíveis sobre a Internet das Coisas (IoT) em ambientes urbanos mostra que é possível enfrentar os desafios da urbanização, proporcionando uma nova qualidade de vida aos cidadãos, tornando o seu dia-a-dia mais conveniente e seguro (Leal et al., 2023).

Esta investigação analisa os padrões de deslocação e movimentos das pessoas em áreas de diversão noturna lisboeta, permitindo solucionar o facto de ser extremamente necessário monitorizar estas mesmas áreas. Os movimentos são obtidos através de dados (já anonimizados) de dispositivos móveis de uma operadora móvel portuguesa fornecidos à Câmara Municipal de Lisboa. Desta forma, os telemóveis tornam-se sensores úteis para recolher informações espaço-temporais sobre o comportamento e padrões de movimento humano. É também muito importante relacionar os movimentos de pessoas com os níveis de ruído registados nos sensores ambientais localizados na zona de estudo considerada.

Desta forma, este estudo visa responder à seguinte Questão de Investigação “Os dados anonimizados dos telemóveis fornecidos por operadoras móveis permitem identificar padrões de mobilidade?”.

Esta questão de investigação visa proporcionar um conhecimento aprofundado sobre a dinâmica da animação noturna nos espaços públicos ao ar livre em Lisboa, utilizando dados dos dispositivos móveis.

Para o planeamento urbano, gestão do turismo e o bem-estar geral da população, é imperativo conhecer os horários e os locais de diversão noturna, bem como a caracterização das variáveis relacionadas. A resposta a esta questão pode fornecer informações cruciais aos profissionais da administração local, aos planeadores urbanos e aos especialistas do turismo, permitindo uma administração mais eficiente das zonas de diversão noturna e uma experiência mais satisfatória para todos os envolvidos. Contribuirá também para o conhecimento científico da dinâmica das cidades à noite e constituirá um modelo para outras áreas urbanas que lidam com questões semelhantes.

1.3. Objetivos e Contributos

Tendo em conta a questão de investigação formulada anteriormente, o objetivo principal passa por identificar e estudar os padrões e comportamentos da população, tanto portuguesa quanto estrangeira, em áreas de lazer noturno. Para alcançar o objetivo geral são definidos objetivos específicos que incluem: (1) Identificar as áreas mais movimentadas de Lisboa; (2) Analisar o comportamento das pessoas durante eventos festivos, como as Jornadas Mundiais da Juventude

(JMJ); (3) Avaliar as possíveis transferências de utilizadores entre diferentes zonas de diversão noturna e (4) Relacionar os movimentos das pessoas com os níveis de ruído captados pelos sensores ambientais nas áreas estudadas. Este documento contribui para a transformação digital de Lisboa numa cidade inteligente, através de uma melhor compreensão dos padrões de mobilidade urbana (Leal et al., 2023).

1.4. Abordagem metodológica

De forma a resolver o problema do estudo e responder a todos os objetivos propostos, a metodologia adotada é o CRISP-DM (*Cross-Industry Standard Process for Data Mining*). Esta metodologia é constituída por seis fases, Compreensão do Negócio, sendo a segunda a Compreensão dos Dados, seguindo para Preparação dos Dados, a quarta etapa refere-se à Modelação e, as duas últimas fases, a Avaliação e Implementação (Chapman et al., 2000).

A escolha do CRISP-DM para esta investigação é vantajosa por vários motivos. Primeiramente, a sua flexibilidade permite adaptar-se a diferentes tipos de projetos, ajustando-se bem à análise de dados de mobilidade urbana (Chapman et al., 2000). Além disso, a sua estrutura clara e dividida em fases facilita o planeamento e a execução, especialmente em projetos complexos como a análise de dados de localização (Azevedo & Santos, 2008). Outra vantagem é o foco na compreensão do negócio desde o início, o que garante que o projeto esteja alinhado com os objetivos estratégicos, permitindo uma abordagem direcionada (Chapman et al., 2000). O carácter iterativo do CRISP-DM também é essencial, pois possibilita revisões constantes ao longo do processo, garantindo a precisão do modelo final. Finalmente, o CRISP-DM é amplamente aceite e utilizado, o que facilita a sua implementação (Wirth & Hipp, 2000).

O alcance dos objetivos estabelecidos é sustentado por uma abordagem que combina métodos quantitativos e qualitativos. Para identificar os indicadores de desempenho, é realizada uma revisão de literatura, complementada por entrevistas com especialistas da área. Na fase final, um questionário é aplicado aos potenciais utilizadores, com o intuito de avaliar o *dashboard* desenvolvido, focando-se na sua usabilidade e utilidade para os fins a que se propõe.

1.5. Estrutura e organização da dissertação

Esta dissertação é composta por cinco capítulos principais, incluindo esta introdução, onde se define a questão de investigação, a metodologia adotada e os objetivos que procuram responder ao problema identificado, o qual motiva o desenvolvimento deste estudo. O segundo capítulo, revisão de literatura, é dividido em três partes. Primeiro, são apresentados os conceitos

relevantes na área de estudo - Mobilidade Urbana e o Uso de Dados de Dispositivos Móveis- com foco na regularidade espaço-temporal e no turismo, para aprofundar o conhecimento e identificar as principais análises de mobilidade a serem realizadas.

Após essa introdução conceitual, realiza-se uma revisão de literatura com o propósito de identificar KPIs (*Key Performance Indicators*) para a construção do *dashboard* e fundamentar porquê e como deve ser construído um *dashboard*. Para encerrar o capítulo, faz-se a conexão entre o conceito de mobilidade urbana e os *dashboards*, abordando a implementação de um *dashboard* para monitorizar a concentração de pessoas.

O terceiro capítulo apresenta a metodologia adotada, o CRISP-DM, detalhando todas as etapas seguidas durante o processo. No penúltimo capítulo, resultados e discussão, são apresentados os resultados obtidos no estudo, ou seja, o *dashboard* construído com o objetivo de resolver o problema inicial, bem como os resultados da avaliação realizada. No final deste capítulo, os resultados são discutidos e são feitas recomendações para os especialistas.

Por fim, o último capítulo, a conclusão, traz as conclusões da investigação, destacando os contributos para o conhecimento científico e para os profissionais do setor em questão. As limitações encontradas ao longo da dissertação também são discutidas nesse capítulo, juntamente com recomendações para trabalhos futuros.

CAPÍTULO 2

2. Revisão da Literatura

A mobilidade urbana, definida como o movimento de pessoas e bens dentro de uma cidade (Ferrovia, s.d.), é um fenómeno complexo que tem implicações no planeamento urbano, na gestão de transportes e na sustentabilidade das cidades. A análise da mobilidade e conectividade em zonas de diversão noturna tem-se tornado um campo de estudo de considerável relevância nas ciências sociais e urbanas, particularmente com o surgimento de novas tecnologias de recolha e análise de dados (Fanticelli et al., 2022). Nesse contexto, o uso de dados anonimizados de operadoras móveis surgiu como uma ferramenta para a análise detalhada dos movimentos das pessoas. Esses dados permitem mapear, em tempo real, a localização, os movimentos e os comportamentos das pessoas, oferecendo informações sobre como estas interagem com o ambiente urbano (Calabrese et al., 2013; Wang et al., 2018).

Esta revisão de literatura reúne e analisa criticamente a informação proveniente de diversos estudos que utilizam dados de equipamentos móveis para explorar padrões de mobilidade e conectividade nas zonas de diversão noturna, relacionando com os níveis de ruído.

2.1. Mobilidade Urbana e Utilização de Dados de Dispositivos Móveis

A mobilidade urbana é um tema central em diversos estudos analisados. De acordo com Fanticelli et al. (2022), a análise de dados móveis pode revelar padrões típicos da vida metropolitana, apresentando uma visão detalhada das deslocações e transferências das pessoas dentro de uma cidade. Esta investigação representa a modelação da mobilidade baseada em dados para quantificar (no espaço e no tempo) os fluxos de frequência e de visita da população em diferentes zonas urbanas.

Segundo Sun et al. (2021), nos dias de hoje uma grande percentagem de pessoas utiliza frequentemente telemóveis. Os dados dos dispositivos móveis contêm várias informações que podem ser utilizadas para inferir quando e onde se encontra o utilizador. Em comparação com outras fontes de *big data* do movimento humano (por exemplo, redes sociais e dados de GPS), os dados dos telemóveis apresentam várias vantagens, como uma elevada cobertura da população, forte continuidade temporal e baixo custo de recolha.

Calabrese et al. (2013) foram pioneiros na utilização de dados de dispositivos móveis para perceber padrões individuais de mobilidade nas áreas urbanas. Neste estudo, demonstraram como os dados dos telemóveis podem ser utilizados para mapear deslocações diárias, como os trajetos de casa para o trabalho, proporcionando uma visão detalhada dos movimentos populacionais nas áreas metropolitanas. Os autores utilizaram *clusters* para identificar locais e trajetos significativos na mobilidade diária dos indivíduos, como as viagens entre casa e trabalho. Deste modo, permitiu a identificação de padrões recorrentes e a deteção de anomalias que poderiam indicar mudanças no comportamento de mobilidade, como eventos ou alterações nas infraestruturas urbanas. Esta investigação sugere que os dados de rastreio dos telemóveis representam um indicador da mobilidade individual e revelam um enorme potencial como fonte de dados alternativa e atualizável com maior frequência e como complemento dos inquéritos de viagem convencionais no estudo da mobilidade.

Outro estudo importante, conduzido por Kung et al. (2014), explorou padrões universais de deslocamento casa-trabalho. Os autores utilizaram dados de registos de chamadas telefónicas (*Call Detail Records* - CDRs) para explorar padrões de mobilidade entre casa e trabalho em diferentes cidades do mundo. A metodologia incluiu o processamento de grandes volumes de dados de localização associados a chamadas telefónicas, permitindo mapear os movimentos diários dos utilizadores entre localizações recorrentes, interpretadas como pontos de residência e trabalho. O estudo revelou que, apesar das diferenças culturais e infraestruturais, existem padrões de mobilidade comuns que podem ser observados em diversas cidades, sugerindo que certos comportamentos de deslocação são universais.

Versichele et al. (2012) adicionaram uma dimensão temporal à análise de mobilidade, utilizando dados de *Bluetooth* para gerir o movimento de pessoas durante eventos de massa na Bélgica. Refere-se a um estudo de caso do evento *Ghent Festivities* (1,5 milhões de visitantes durante 10 dias), cobrindo 22 locais na área de estudo com *scanners Bluetooth*, conseguindo extrair 152 487 trajetos gerados por 80 828 visitantes detetados. Este estudo é relevante para Lisboa, especialmente no contexto de áreas de diversão noturna, pois, apesar de se tratar de dados *Bluetooth*, oferece uma metodologia para capturar o comportamento de multidões em tempo real. A capacidade de rastrear os movimentos de grandes grupos de pessoas durante eventos pode ajudar as autoridades locais a gerenciar melhor o tráfego, melhorar a segurança pública e otimizar os serviços durante eventos de grande escala.

2.1.1. Regularidade Espaço-temporal

Segundo Jiang et al. (2013), os padrões de mobilidade humana são bastante regulares, o que permite o desenvolvimento de modelos preditivos eficazes. Esses modelos são utilizados para a otimização de rotas de transporte público, previsão da procura de serviços e planeamento de infraestrutura urbana.

A regularidade espaço-temporal é uma palavra-chave, pois os movimentos humanos estão frequentemente ligados à repetição de rotas e destinos ao longo do tempo. Para investigar a previsibilidade da dinâmica humana, Song et al. (2010) utilizaram as trajetórias dos telemóveis de milhões de utilizadores, recolhidas por empresas de telemóveis e tornadas anónimas para fins da investigação. A sua hipótese era que, dada a grande variedade de padrões de viagem que os diferentes utilizadores seguem, haveria também diferenças significativas entre a sua previsibilidade: Os utilizadores que viajam menos devem ser mais fáceis de prever do que aqueles que estão constantemente na estrada. Surpreendentemente, a previsibilidade foi de 93% em toda a base de utilizadores e, em geral, a previsibilidade dos indivíduos não desceu significativamente abaixo dos 80%.

Outro estudo importante de Lu et al. (2013) utilizou dados de chamadas telefónicas para analisar a regularidade espaço-temporal dos padrões de mobilidade em diferentes contextos urbanos. Foi possível identificar que os padrões de deslocamento variam significativamente de acordo com fatores como dia da semana, horário e localização, mas mantêm um alto grau de regularidade dentro desses contextos. Esse tipo de análise é importante para a presente dissertação, de forma a perceber como é que varia a concentração e o comportamento da população de acordo com os vários dias da semana, vários períodos do dia, entre outros.

A análise baseada na regularidade espaço-temporal também pode ajudar a mitigar problemas como a superlotação e a poluição sonora, promovendo um ambiente urbano mais seguro e agradável para os residentes e turistas (Park & Zhong, 2021).

2.1.2. Turismo

O turismo é um fator significativo que influencia a mobilidade em áreas urbanas, contribuindo para a complexidade dos padrões de movimento dentro das cidades. Estudos como os de Kovács et al. (2023) e Sun et al. (2021) utilizam dados de dispositivos móveis para identificar e analisar os padrões de mobilidade de turistas em destinos urbanos. Esses estudos revelam como os turistas se movem pelas cidades, permitindo a identificação de áreas de alta concentração turística e fornecendo insights valiosos sobre como a presença de turistas impacta a dinâmica das zonas de diversão noturna.

A investigação de Kovács et al. (2023) analisa o comportamento espaço-temporal dos visitantes internacionais, utilizando dados de telemóveis em três cidades da Hungria durante seis meses consecutivos. O objetivo deste estudo é identificar as origens dos visitantes, a duração da estadia, as concentrações sazonais e a mobilidade entre cidades. Os dados revelaram diferenças significativas no volume de negócios dos turistas, nos padrões de mobilidade e nos fluxos, indicando diferentes níveis de atividade e potencial turístico nas três cidades.

A pesquisa de Paananen & Minoia (2018) sobre a mobilidade de passageiros de cruzeiros em Helsínquia destaca a importância de considerar diferentes tipos de turistas ao estudar a mobilidade urbana. Os passageiros de cruzeiros, por exemplo, têm geralmente um comportamento de mobilidade distinto em comparação com outros turistas, pois tendem a explorar a cidade de forma concentrada e em grupos maiores, o que pode alterar significativamente os padrões de movimento. Isso pode resultar em picos de procura por transporte e serviços em horários específicos, exigindo uma gestão mais dinâmica e adaptativa da infraestrutura urbana.

Além disso, a sazonalidade do turismo pode afetar a mobilidade urbana. Durante a época alta, como o verão ou períodos de festas, o aumento no número de turistas pode levar a congestionamentos e maior pressão sobre os sistemas de transporte público. Por outro lado, na época baixa, a diminuição do fluxo turístico pode resultar numa menor utilização dos serviços, impactando a economia local. Portanto, a análise detalhada dos padrões de mobilidade pode ajudar as cidades a desenvolver estratégias para equilibrar a oferta e a procura de serviços ao longo do ano. (Brida & Zapata, 2010; Carić, 2016)

Desta forma, compreender os padrões de mobilidade dos turistas e a sua interação com os residentes locais é crucial para o desenvolvimento urbano sustentável. Planeadores urbanos e gestores de turismo podem utilizar essas informações para criar ambientes urbanos mais eficientes e acolhedores, promovendo uma experiência positiva tanto para os turistas quanto para os residentes.

2.2. Dashboards

2.2.1. Conceito de *dashboard*

Segundo Few (2006), em 1980 foram desenvolvidos os *Executive Information Systems* (EIS) que mostravam um conjunto de medidas financeiras através de uma interface simples. Em 1990, com a identificação de utilização de KPIs (*Key Performance Indicators*), o BSC (*Balanced*

Scorecard) e o BPM (*Business Performance Management*), os EIS evolui para o atual conhecido *dashboard*.

Few (2006) define *dashboard* como uma ferramenta visual que apresenta informações importantes e necessárias para atingir o(s) objetivo(s) proposto(s). Esta ferramenta é apresentada e consolidada num único ecrã, para que toda a informação possa ser monitorizada (Few, 2006).

Para Yigitbasioglu & Velcu (2012), um *dashboard* trata-se de um sistema de apoio à decisão baseado em dados, que fornece informações, num formato específico, ao decisor. Para Pauwels et al. (2009), um *dashboard* é um conjunto de indicadores-chave e métricas de desempenho interligados que refletem os interesses comuns a curto e a longo prazo a serem considerados por toda a organização. Para Nica et al. (2021), um *dashboard* é uma ferramenta de gestão de desempenho que analisa, seleciona e apresenta indicadores-chave de desempenho (KPIs) para identificar a situação da empresa e sintetizar a atividade desenvolvida. Oferece a possibilidade de monitorizar e melhorar o desempenho em tempo real, reduzindo o tempo de análise.

Apesar de existirem várias definições do conceito, todas têm em comum o facto de ser um sistema de apoio à decisão que apresenta informação relevante, que permita identificar ações a serem implementadas, com vista a atingir os objetivos estabelecidos.

2.2.2. Indicadores e métricas nos *dashboards*

Pauwels et al. (2009) definem indicadores-chave de desempenho, ou KPIs, como medições que refletem o desempenho de processos críticos para o sucesso organizacional. Segundo os autores, os elementos-chave de um *dashboard* incluem o resumo e a integração das principais métricas de desempenho com os fatores de desempenho subjacentes para comunicar o desempenho em toda a organização.

Nica et al. (2021) definem KPIs como indicadores financeiros e não financeiros utilizados para avaliar o sucesso de uma empresa na prossecução de objetivos a longo prazo previamente definidos. Os KPIs de alto nível centram-se no desempenho global da empresa, enquanto os KPIs de baixo nível consideram processos em departamentos como *marketing*, vendas e recursos humanos.

De acordo com Yigitbasioglu e Velcu (2012), um dos principais desafios no *design* de *dashboards* é selecionar adequadamente indicadores e métricas. Destacam que os *dashboards* eficazes são aqueles que incluem apenas os indicadores mais relevantes e evitam a sobrecarga de informações. Eles ressaltam a necessidade de uma escolha criteriosa de KPIs que consigam refletir diretamente os objetivos da organização e permitam uma ação corretiva imediata.

Os três objetivos dos KPIs passam por motivar a equipa, cumprir a utilização dos ativos dos processos organizacionais e alinhar com os objetivos da empresa, melhorando o desempenho e adotando as melhores práticas (Nica et al., 2021).

Para identificar as características de um KPI, na literatura é utilizada a regra SMART (Nica et al., 2021):

- Específico (S): o indicador é focado em objetivos de desempenho;
- Mensurável (M): pode ser apresentado quantitativamente;
- Atingível (A): os objetivos são razoáveis e exequíveis;
- Realista (R): o indicador é relevante para o trabalho realizado no projeto;
- Baseado no tempo (T): o indicador é mensurável num período.

2.2.3. Tipos de *dashboard*

Os *dashboards* podem ser classificados em diferentes tipos, dependendo de seu objetivo e do público-alvo. De acordo com Few (2006), Sarikaya et al. (2019) e Nica et al. (2021) os três principais tipos de *dashboard* são:

- *Dashboards* Operacionais: Permitem monitorizar atividades em tempo real, fornecendo uma visão geral das operações diárias (Nica et al., 2021). Estes *dashboards* devem ser concebidos de forma diferente dos que apoiam a tomada de decisões estratégicas ou a análise de dados. São úteis para detetar problemas ou oportunidades imediatamente, permitindo ações rápidas (Few, 2006). O *dashboard* operacional descreve o passado, o presente e futuro, sendo capaz de alertar para um valor problemático, sendo assim possível tomar uma ação imediata para corrigir o problema (Sarikaya et al., 2019);

- *Dashboards* Táticos: Destinados a utilizadores que precisam de tomar decisões de curto a médio prazo. Esses *dashboards* permitem monitorizar o progresso em relação a metas e objetivos específicos. Permite a análise dos processos em cada departamento (Nica et al., 2021). Os gestores departamentais utilizam a informação dos sistemas operacionais, permitindo que os analistas possam identificar problemas e, por consequência, proceder às suas correções, de

forma a assegurar os objetivos a curto e longo prazo, bem como responder às necessidades do negócio (Eckerson, 2010);

- *Dashboards* Estratégicos: Utilizados por gestores de alto nível para tomar decisões de longo prazo (Few, 2006). São utilizados para acompanhar os progressos no sentido do objetivo proposto (Nica et al., 2021). Os *dashboards* deste tipo centram-se em medidas de desempenho de alto nível, incluindo previsões para iluminar o caminho para o futuro (Few, 2006). Dado o objetivo de uma orientação estratégica a longo prazo, e não de reações imediatas a mudanças rápidas, estes *dashboards* não requerem dados em tempo real; em vez disso, é benéfico visualizações estáticas mensais, semanais ou diárias (Few, 2006)².

2.2.4. Características de um *dashboard*

A literatura descreve várias características fundamentais que tornam um *dashboard* eficaz, envolvendo aspetos de *design*, funcionalidade e adequação ao utilizador.

De acordo com Few (2006), a simplicidade é uma das principais características de um *dashboard* eficaz. Deve apresentar informações de forma clara e concisa, para que seja possível interpretar os dados com rapidez e precisão. As informações devem ser organizadas de forma que possam ser compreendidas "à primeira vista", permitindo a análise em tempo real sem sobrecarregar o utilizador com excesso de detalhes. Desta forma, é importante a posição das visualizações. O canto superior esquerdo deve conter as informações mais relevantes. A organização deve assumir um padrão em Z, isto é, com a informação mais importante a seguir este padrão de cima para baixo e da esquerda para a direita, tal como se mostra na Figura 2.1.

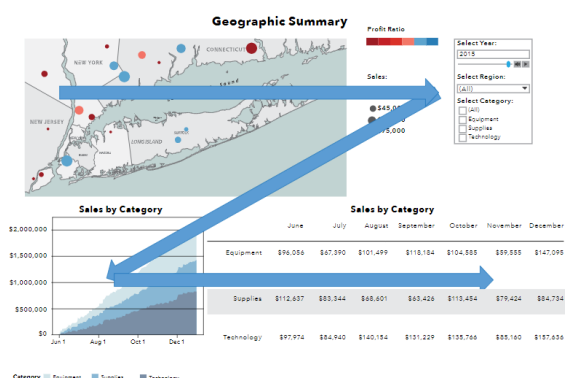


Figura 2.1: Ordem de leitura de um *dashboard*

Fonte: Firican (2017)

² No Anexo A estão resumidos os 3 tipos de *dashboards*. Resume qual o objetivo principal de cada tipo de *dashboard* e quem os utiliza.

Quanto às características funcionais, o *dashboard* deve permitir fazer *drill down* da informação (Few, 2006). A utilização de filtros é outra característica fundamental, que permite ao utilizador consultar a informação de forma condicionada (Firican, 2017; Few, 2006). Tornar o *dashboard* interativo permite que os utilizadores finais executem tarefas analíticas básicas e aprendam mais sobre os dados. Os visualizadores devem conseguir obter todas as informações necessárias no seu estado predefinido, mas permitir-lhes aprofundar uma vista fornece-lhes mais informações e respostas que podem ajudá-los a realizar o seu trabalho de forma mais eficiente (Firican, 2017).

A escolha da paleta de cores também é algo a ter em conta na construção de um *dashboard*. Firican (2017) aconselha o uso de paletes de *Brewer*, sempre que possível. Estas paletas são combinação de cores selecionadas compatíveis com os daltónicos e com impressões a preto e branco (as cores escuras permanecem mais escuras e as cores claras permanecem mais claras).

Yigitbasioglu e Velcu (2012) argumentam que as tabelas oferecem maior precisão na apresentação de informações em comparação com gráficos. No entanto, certos tipos de dados podem ser monitorizados de forma mais eficiente quando representados graficamente, uma vez que os gráficos simplificam a visualização ao reduzir o excesso de informações, algo que ocorre com frequência nas tabelas. Por isso, é sugerido que se dê preferência aos gráficos, a menos que a natureza dos dados torne a visualização em gráfico menos clara.

Para Firican (2017), a frequência de atualização dos dados deve ser definida com base nas necessidades e no propósito do *dashboard*, garantindo que as informações estejam sempre atualizadas. É essencial determinar com clareza a periodicidade da atualização e o período dos dados exibidos. Esse intervalo de tempo pode ser indicado no título, subtítulo, descrição ou até por meio de um filtro, quando apropriado. Outra recomendação do autor é incluir a data da última atualização e a fonte dos dados no cabeçalho ou rodapé.

Por último, o *dashboard* tem de ser acessível. O ideal é que seja construído na *web*, para que esteja disponível em qualquer lugar, desde que haja internet. Também é preciso ter cuidado com a segurança do *dashboard*, ou seja, se apresentar informações confidenciais é essencial restringir o acesso com um nome e uma palavra-passe (Firican, 2017)

2.2.5. Tipos de gráficos mais adequados

A escolha dos gráficos num *dashboard* é um fator crucial para garantir que as informações sejam transmitidas de forma clara, precisa e rápida.

Para Firican (2017), um gráfico de barras é uma das melhores e mais simples formas de visualizar uma ou mais séries de dados. Ao contrário de outros gráficos, um gráfico de barras permite analisar a diferença de volumes de alguma variável muito mais rapidamente. O autor recomenda a utilização do gráfico de barras para dados sobre várias séries relacionados com o tempo, produto, receita ou região.

Segundo Firican (2017), os gráficos de linhas são recomendados quando se pretende acompanhar as alterações num conjunto de dados dependentes durante um período de tempo.

De acordo com Few (2006), os gráficos de linhas são altamente eficazes para mostrar tendências ao longo do tempo. Eles permitem a visualização de dados contínuos, facilitando a identificação de padrões e mudanças numa série temporal. Quando se pretender apresentar a tendência num único conjunto de dados, é recomendável utilizar *sparklines*.

Quando se trata de um conjunto de dados geoespaciais, o gráfico do mapa é o recomendado (Firican, 2017). De acordo com Few (2006), mapas de calor geoespaciais são eficazes para mostrar a densidade de uma variável ou a intensidade de um evento em diferentes regiões. Estes gráficos utilizam uma paleta de cores para representar valores elevados ou baixos, facilitando a identificação rápida de "pontos quentes" (áreas com alta concentração de eventos ou valores) e "pontos frios" (áreas de baixa concentração).

Os gráficos de dispersão são recomendados para *dashboards* onde é necessário visualizar correlações ou relações entre duas variáveis quantitativas. No entanto, os gráficos de dispersão podem ser complexos para utilizadores que não estão familiarizados com esse tipo de visualização, sendo mais recomendados para *dashboards* destinados a especialistas que lidam com dados técnicos ou analíticos (Few, 2006).

Para todos os gráficos, Firican (2017) recomenda utilizar a mesma cor para a mesma série de dados. Se forem utilizadas várias cores em séries diferentes, uma legenda é essencial.

2.2.6. Vantagens

As vantagens dos *dashboards* incluem a monitorização em tempo real das alterações, centralização de todos os dados num único local para que possam ser acedidos por todos os utilizadores, previsão com base nas tendências atuais e identificação de tendências negativas (Nica et al., 2021).

Uma das maiores vantagens de um *dashboard* é que ele facilita a tomada de decisões informadas. Few (2006) argumenta que, ao apresentar informações complexas de maneira simplificada, os *dashboards* permitem que os utilizadores compreendam rapidamente os dados e tomem decisões com base em fatos. A visualização de KPIs e métricas essenciais de forma consolidada torna o processo de decisão mais eficiente.

Para Eckerson (2006), a razão de tantas organizações implementarem *dashboards* deve-se ao facto de estes oferecerem uma série de benefícios para todos os membros, desde gestores de topo a colaboradores. Os benefícios enunciados por Eckerson (2006) são os seguintes:

- Comunicar a estratégia corporativa em métricas, objetivos e iniciativas que são personalizadas para cada departamento ou membro organizacional. Os colaboradores, através do *dashboard*, obtêm uma visão clara dos objetivos estratégicos da organização e, consequentemente, das ações que precisam de tomar para os alcançar;
- Aperfeiçoar a estratégia consoante problemas internos ou tendências do setor de atividade;
- Os *dashboards* fornecem aos gestores de topo e colaboradores uma visão mais abrangente das operações diárias e do desempenho futuro, recolhendo dados relevantes de maneira oportuna e prevendo tendências com base no desempenho passado;
- Ao divulgar os dados de desempenho, os *dashboards* incentivam as equipas de diferentes departamentos a trabalharem juntas de forma mais coesa, além de promoverem o diálogo entre os líderes das equipas e os colaboradores sobre como melhorar o desempenho.

Concluindo, organizações que têm uma estratégia clara, uma cultura orientada por métricas e uma infraestrutura de informação forte podem fornecer sistemas de gestão de desempenho que causam um impacto dramático no resultado (Eckerson, 2006).

2.2.7. Construção e avaliação de um *dashboard*

De acordo com Nica et al. (2021), a construção de um *dashboard* é realizada em quatro etapas. No entanto, Smith (2013) inclui uma quinta etapa importante e, por isso, são consideradas cinco etapas para construir um *dashboard* (Figura 2.2).

01	02	03	04	05
Identificar problemas e definir objetivos	Determinar utilizadores finais e um plano	Selecionar os indicadores adequados	Elaborar o <i>dashboard</i>	Avaliar o <i>dashboard</i>
Selecionar os principais dados e KPIs a destacar no <i>dashboard</i>	Identificar a informação relevante para o utilizador final e que impactem o objetivo diretamente	Escolher os indicadores mais pertinentes para avaliar o resultado das ações	Utilizar formas que permitam uma boa visualização dos indicadores, para que a mensagem seja comunicada de forma eficaz	É necessário ter em conta um conjunto de critérios, de forma a responder às necessidades das partes interessadas

Figura 2.2: Etapas de construção do *dashboard*

Um *dashboard* deve comunicar as informações de forma clara e concisa. Isso inclui a utilização de gráficos apropriados, cores e disposição dos elementos visuais. O estudo de Few (2006) destaca a importância da simplicidade e clareza na apresentação dos dados, salientando que *dashboards* eficazes devem evitar o excesso de informações e focar-se apenas nas métricas mais relevantes.

Pauwels et al. (2009) reforçam que um *dashboard* deve ser construído de forma que diferentes níveis de utilizadores (desde gestores de topo até colaboradores) possam utilizá-lo facilmente. Se o *dashboard* for difícil de navegar ou interpretar, perde seu valor como ferramenta de apoio à decisão.

Os dados apresentados no *dashboard* devem ser relevantes para os objetivos do utilizador. Isso envolve a personalização do conteúdo exibido com base nas necessidades e preferências do público-alvo. De acordo com Sarikaya et al. (2018), a relevância do conteúdo e eficácia de um *dashboard* pode ser avaliada através de *feedbacks* dos utilizadores, que ajudam a identificar quais informações são mais úteis e quais podem ser descartadas. Questionários, entrevistas e grupos focais podem ser utilizados para obter opiniões sobre a usabilidade, a clareza e a relevância das informações apresentadas (Yigitbasioglu & Velcu, 2012).

Para Yigitbasioglu & Velcu (2012), os *dashboards* devem ser avaliados em função do *design* e da forma como o utilizador interage com eles para tomar decisões. Os autores sugerem que *dashboards* interativos são mais eficazes em contextos onde a análise de dados requer uma exploração detalhada e personalizada.

Medir o desempenho do *dashboard* em termos de tempo de resposta, velocidade de atualização de dados e estabilidade é muito importante para garantir uma boa experiência do utilizador, especialmente em contextos de monitorização em tempo real (Pauwels et al., 2009).

Para Eckerson (2006), para avaliar um *dashboard* devem ser considerados os critérios ao nível do *design*, da análise, da entrega, da administração e da infraestrutura³.

2.2.8. Dashboards de mobilidade humana

A utilização de *dashboards* para análise da mobilidade urbana tem-se mostrado uma ferramenta poderosa na gestão e planeamento das cidades inteligentes. *Dashboards* permitem monitorizar o tráfego e a movimentação urbana em tempo real, integrando dados de diversas fontes. O estudo de Prabhakar (2015) destaca os *dashboards* como uma ferramenta para análise de grandes volumes de dados de mobilidade, utilizando informações de transporte público e dispositivos móveis. A ferramenta não apenas monitoriza o tráfego, mas também prevê padrões de mobilidade, auxiliando na otimização de rotas e na eficiência do transporte urbano. Além da monitorização em tempo real, os *dashboards* são eficazes na análise de padrões históricos de deslocamento.

Além do tráfego de automóveis, os *dashboards* também são utilizados para analisar os padrões de mobilidade de pedestres, especialmente em áreas urbanas densas e durante eventos de grande escala. Estudos como o de Versichele et al. (2012) demonstram a utilização de *dashboards* para monitorizar o movimento de cidadãos e identificar áreas de alta concentração, o que é útil para o planeamento urbano e a segurança pública.

³ No Anexo B são apresentados os critérios de avaliação dos *dashboards* quanto ao *design*, análise, entrega, administração e infraestrutura.

CAPÍTULO 3

3. Metodologia

A metodologia a adotar para a concretização da dissertação é o CRISP-DM. Divide-se nas seguintes fases: compreensão do negócio, compreensão dos dados, preparação dos dados, modelação, avaliação e implementação.

A metodologia CRISP-DM envolve várias etapas que podem ser seguidas de maneira linear, mas também podem apresentar um fluxo não-linear, onde é necessário voltar a etapas anteriores conforme novas informações surgem.

Por exemplo, durante a fase de modelação, o analista pode descobrir que a preparação dos dados não foi adequada. Isso pode levá-lo a retornar à fase de preparação de dados para resolver problemas identificados, o que, por sua vez, pode impactar a modelação e as fases subsequentes.

De forma semelhante, durante a fase de avaliação, o analista pode perceber que a compreensão inicial do negócio não foi correta, o que significa que ele pode estar a tentar responder a uma pergunta inadequada. Nessa situação, o analista revisita a fase de compreensão do negócio para ajustar os objetivos e prosseguir com o processo com um propósito mais claro e bem definido (PSE, s.d.).

3.1. Compreensão do Negócio

Na fase de compreensão do negócio, é essencial a informação fornecida por uma operadora móvel, de forma a identificar e estudar padrões e comportamentos da população nas zonas de diversão noturna na cidade de Lisboa. A análise será conduzida utilizando as bases de dados fornecidas pela Câmara Municipal de Lisboa, referentes aos meses específicos.

Os objetivos do projeto incluem a otimização da experiência noturna, maximização do impacto económico do turismo noturno e a gestão eficiente do espaço público noturno. A fonte de dados é proveniente de uma operadora de telecomunicações, fornecida pela Câmara Municipal de Lisboa, com a necessidade de considerar potenciais problemas de qualidade dos dados ou questões de privacidade. A gestão da diversão noturna requer um equilíbrio delicado entre os interesses de moradores locais, utilizadores nacionais e turistas.

A implementação de IoT destaca-se como uma solução vantajosa, permitindo a monitorização em tempo real da movimentação de pessoas. Portanto, torna-se relevante identificar a movimentação das pessoas nas áreas de diversão noturna, identificando alturas e zonas específicas de maior movimento, como o caso da semana das JMJ (Jornadas Mundiais da Juventude), para que, por meios da IoT, seja possível alcançar uma gestão mais inteligente e equilibrada entre as partes envolvidas.

O objetivo geral passa por identificar e estudar padrões e comportamentos da população, tanto portuguesa quanto estrangeira, nas zonas de diversão noturna. Os objetivos específicos, já mencionados anteriormente, passam por (1) Identificar as zonas mais movimentadas de Lisboa; (2) Analisar o comportamento da população em eventos festivos, como as JMJ; (3) Avaliar as possíveis transferências dos utilizadores entre as zonas de diversão noturna; (4) Relacionar os movimentos das pessoas com os níveis de ruído registados nos sensores ambientais nas áreas de estudo consideradas. O objetivo analítico, de forma a ir de encontro ao objetivo de negócio, passa pelo desenvolvimento de um sistema de apoio à decisão baseado em dados (*dashboard*), recorrendo ao *Power BI*, para oferecer uma solução ao problema que foi encontrado – a necessidade de monitorização das áreas de diversão noturna de Lisboa.

Os critérios de sucesso incluem a qualidade dos resultados das análises, assim como a formulação de recomendações impactantes para a gestão urbana durante a noite em Lisboa. O projeto só será dado como concluído se a ferramenta de monitorização for avaliada, em termos de utilidade, qualidade e facilidade de utilização, com quatro (ou mais) pontos, numa escala de 0 (nada satisfeito) a 5 (extremamente satisfeito).

3.2. Compreensão dos dados

Em relação à compreensão dos dados, o objetivo principal é obter uma visão mais aprofundada dos dados disponíveis, perceber a sua estrutura, qualidade e conteúdo. Esta fase envolve explorar os dados, identificar padrões preliminares e definir os passos seguintes para a preparação dos mesmos.

As informações presentes no conjunto de dados foram devidamente anonimizadas pela operadora móvel de forma a cumprir com as leis e proteger a privacidade dos utilizadores. Assim, não é possível, de maneira nenhuma, efetuar qualquer análise específica de um utilizador individual. Não existe nada que possa ligar um utilizador específico a um evento, sendo possível apenas realizar análises baseadas em volumes de dados.

Foram disponibilizados 4 conjuntos de dados:

- Conjunto 1: Contém informação do número de telemóveis que entram, permanecem e saem por quadrícula de 200m x 200m num período de 15 minutos. Na Tabela 3.1 encontra-se o nome, a descrição e o tipo de dados das 24 variáveis presentes neste conjunto;

Tabela 3.1: Descrição dos campos do Conjunto 1

ID	Nome do campo	Descrição	Tipo de dados
1	grid_ID	Identificação da quadrícula da grelha	Númérico
2	Datetime	Data e hora do registo	Data/Hora
3	C1	N.º de terminais distintos na quadrícula	Númérico
4	C2	N.º de terminais distintos, em roaming na quadrícula	Númérico
5	C3	N.º de terminais distintos que permanecem na quadrícula	Númérico
6	C4	N.º de terminais distintos que permanecem na quadrícula, em roaming	Númérico
7	C5	N.º de entradas de terminais distintos na quadrícula	Númérico
8	C6	N.º de saídas de terminais distintos na quadrícula	Númérico
9	C7	N.º de entradas de terminais distintos, em roaming, na quadrícula	Númérico
10	C8	N.º de saídas de terminais distintos, em roaming, na quadrícula	Númérico
11	C9	N.º de terminais distintos com ligação de dados ativa, na quadrícula	Númérico
12	C10	N.º de terminais distintos com ligação de dados ativa, em roaming, na quadrícula	Númérico
13	C11	N.º de chamadas de voz com origem na quadrícula	Númérico
14	D1	Top 10 países de origem dos equipamentos terminais em roaming	Texto
15	E1	N.º de chamadas de voz que terminaram na quadrícula	Númérico
16	E2	Ritmo médio de downstream da quadrícula, na quadrícula	Númérico
17	E3	Ritmo médio de upstream da quadrícula, na quadrícula	Númérico
18	E4	Ritmo de pico de downstream da quadrícula, na quadrícula	Númérico
19	E5	Ritmo de pico de upstream da quadrícula, na quadrícula	Númérico
20	E6	Top 10 aplicações (texto com nomes separados por;)	Texto
21	E7	Duração da permanência mínima dentro quadrícula	Númérico
22	E8	Duração da permanência média dentro quadrícula	Númérico
23	E9	Duração da permanência máxima dentro quadrícula	Númérico
24	E10	N.º de dispositivos que efetuam partilha de ligação na quadrícula	Númérico

- Conjunto 2: Refere-se ao mapeamento das quadrículas da operadora móvel na cidade de Lisboa. É composta por 3742 registos. Este conjunto adicional inclui as coordenadas de cada grid, a(s) freguesia(s) em que a grid está localizada, o nome, a geometria e o WKT (Well-known text). Esses dados permitem a georreferenciação do conjunto de dados principal. As 14 variáveis estão descritas na Tabela 3.2;

Tabela 3.2: Descrição dos campos do Conjunto 2

ID	Nome do campo	Descrição	Tipo de campo
1	grelha_id	Número da grelha	Numérico
2	dicofre	Identificação da freguesia atribuída por entidades administrativas	Numérico
3	entity_id	Identificação da fonte de dados onde a informação foi gerada	texto
4	entity_type	Identificação da fonte de dados onde a informação foi gerada	texto
5	freguesia	Freguesia da grelha	texto
6	freguesias	Denominação anterior das freguesias	texto
7	grelha_x	Posição da grelha eixo x	Numérico
8	grelha_y	Posição da grelha eixo y	Numérico
9	latitude	Latitude	Numérico
10	longitude	Longitude	Numérico
11	nome	Designação do arruamento/zona	texto
12	objectid	ID do objeto na base de dados	Numérico
13	position	Grid no formato geométrico	Numérico
14	wkt	Grid no formato WKT	Numérico

- Conjunto 3: Diz respeito a dados da rede de sensores ambientais da cidade de Lisboa. Este conjunto de dados pertence aos dados da Rede de monitorização de parâmetros ambientais da cidade de Lisboa, composta por 80 estações, cada uma delas equipada com diferentes sensores de medição de parâmetros de qualidade do ar, ruído, estado do tempo e tráfego. As leituras são médias horárias obtidas tendo por base valores médios de períodos de 15 minutos. As 14 variáveis do conjunto de dados, juntamente com as suas descrições e tipos de variáveis, estão listadas na Tabela 3.3;

Tabela 3.3: Descrição dos campos do Conjunto 3

ID	Nome do campo	Descrição	Tipo de dados
1	DTM.UTC	Data e hora UTC	Data e hora
2	DTM.LOCAL	Data e hora local (UTC+1 em período de verão)	Data e hora
3	TEMÁTICA	Temática a que corresponde o parâmetro	Texto
4	COD.PARAMETRO	Código do Parâmetro	Texto
5	PARAMETRO	LAEQ	Texto
6	NR. ESTACAO	Número da estação	Número inteiro
7	CODE.SENSOR	Identificação do sensor : Temática (2 dígitos) Parâmetro (4 dígitos) Número da estação (4 dígitos), Ex: RULAEQ0001	Texto
8	LOCAL	Nome do local onde está instalada a estação	Texto
9	LATITUDE	Latitude da localização da estação	Numérico
10	LONGITUDE	Longitude da localização da estação	Numérico
11	UNIDADE	dB(A)	Texto
12	ETIQUETA.NIVEL	Etiqueta atribuída ao parâmetro: NORMAL: dB(A) 0 a 54, MODERADO: 55 a 59, ELEVADO: 60 a 64, MUITO ELEVADO: > 65 dB	Texto
13	COR.NIVEL	Cor atribuída ao nível: VERDE: dB(A) 0 a 54, AMARELO: 55 a 59, LARANJA: 60 a 64, VERMELHO: > 65 dB	Texto
14	VALOR	Valor observado	Numérico

- Conjunto 4: Contém a localização dos estabelecimentos com limitadores de ruído. Estes dados foram fornecidos pela DMAEVCE (Direção Municipal do Ambiente, Estrutura Verde, Clima e Energia) e contém informação da latitude e longitude dos estabelecimentos. O nome das variáveis, a descrição e o tipo estão descritos na Tabela 3.4. Este conjunto não será utilizado, uma vez que a informação da latitude e longitude dos estabelecimentos com limitadores de ruído já se encontra no Conjunto 3.

Tabela 3.4: Descrição dos campos do Conjunto 4

Nome do campo	Descrição	Tipo de campo
Latitude	Latitude do estabelecimento	Numérico
Longitude	Longitude do estabelecimento	Numérico

3.3. Preparação dos dados

A seguinte fase é a preparação dos dados. O foco é organizar, limpar e transformar os dados para que estejam prontos para a modelação.

Esta fase inclui várias etapas, representadas na Figura 3.1, seleção, limpeza, construção, integração e, por fim, formatação dos dados.



Figura 3.1: Etapas da preparação de dados

Na primeira etapa, na seleção de dados, em relação ao conjunto de dados da operadora móvel (Conjunto 1), serão utilizadas apenas 4 colunas: a *Grid_ID*, que é a chave única de identificação de cada *grid*; a *Datetime* que refere a data e hora de cada registo; a variável C3 que descreve utilizadores nacionais que permanecem em cada *grid* e a variável C4, por serem terminais em *roaming*, descreve o comportamento de permanência em cada *grid* por parte dos turistas. A variável C4 é utilizada visto que se pretende analisar a concentração e mobilidade não só dos utilizadores nacionais, mas também dos turistas, sendo necessário utilizar dados em *roaming*. De acordo com Uy (2021), entende-se por *roaming* de dados quando o dispositivo móvel se conecta a redes fora da sua rede doméstica. Foram selecionados todos os dados correspondentes ao ano de 2023, de forma a ter uma análise mais detalhada e precisa.

Em relação ao Conjunto 2, fornecido pela equipa da LxDataLab, também não foram utilizadas todas as variáveis. As variáveis utilizadas foram a *Grid_ID*, que identifica a *grid* e permite fazer ligação com o outro conjunto, a latitude, a longitude e o nome da zona. As variáveis *dicofre*, *entity_id* e *entity_type* foram eliminadas visto não serem úteis para a análise. Em relação à variável *freguesias* é uma variável que junta várias freguesias, sendo a variável *freguesia* suficiente e mais adequada para o estudo em questão. Neste conjunto de dados, foram fornecidas as *grids* em formato geográfico (*position*) e em formato *wkt*. Estas variáveis não são importantes porque serão utilizadas a longitude e latitude, que já fornecem as coordenadas das *grids*. Apenas foram analisadas as freguesias onde há mais concentração das pessoas no período noturno. As freguesias úteis para a análise, segundo o Tripadvisor (s.d.), são a Misericórdia, Estrela, Alcântara, Parque das Nações, Santo António, Santa Maria Maior e São Vicente, como se observar na Figura 3.2.

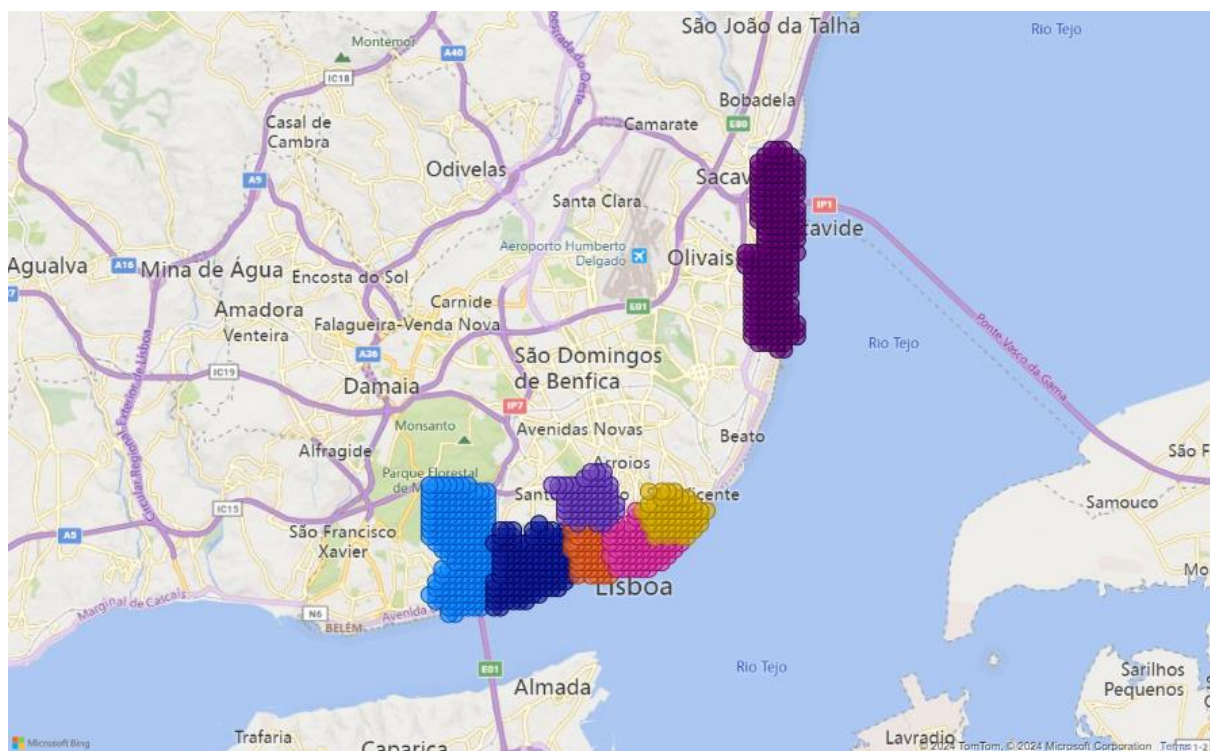


Figura 3.2: Freguesias seleccionadas para a análise

Relativamente ao Conjunto 3, que pertence à rede de monitorização de parâmetros ambientais da cidade de Lisboa, foram utilizadas todas as variáveis exceto a DTM.UTC que se refere à data e hora UTC, tendo-se optado por utilizar a data e hora local. A variável UNIDADE, TEMATICA, PARAMETRO e COD_PARAMETRO também foram excluídas porque contêm apenas um valor (a unidade de medida, a temática do parâmetro, o parâmetro e o código do parâmetro).

A variável “Grid_ID” tem como objetivo unir a informação através da base de dados fornecida pela operadora móvel com a base de dados geográfica, de modo a localizar os registos nas várias freguesias do concelho de Lisboa. Deste modo, a análise e a visualização dos dados em *Power BI* será mais enriquecida.

Na etapa de limpeza dos dados, com recurso a *Python*, analisaram-se os dados que tinham valores nulos ou duplicados. Não houve registos nulos e/ou duplicados em nenhuma das bases de dados, portanto, a limpeza não foi necessária.

Na fase de construção de dados, todas as colunas foram convertidas para o formato correto, como a coluna “Datetime” e procedeu-se à criação de novas variáveis, como a “Weekday”, “Date”, “Time”, “Holiday”, “PeriodoDia”. É crucial distinguir entre dias da semana, fins de semana e feriados, visto que o estudo se foca na concentração e mobilidade das pessoas e, por se focar também em zonas de diversão noturna, é importante diferenciar por período do dia. Só foi selecionado o período entre as 20h00 e as 06h00 (exclusive), o qual foi dividido nos seguintes subperíodos:

- Jantar: das 20h00 às 22h00 (exclusive);
- Noite: das 22h00 às 00h00 (exclusive);
- *Party*: das 00h00 às 04h00 (exclusive);
- *After Party*: das 04h00 às 06h00 (exclusive).

Foi também necessário criar uma coluna em *Power BI*, que permite diferenciar os dias em que os feriados são numa sexta-feira ou sábado, daqueles em que correspondem a um dia da semana. Desta forma, criou-se a coluna “WeekendOrHoliday”⁴ que divide os vários dias em três categorias:

- Fim de Semana: Quando for sexta-feira ou sábado (mesmo que seja feriado);
- Feriado: Quando for feriado em dias da semana (menos sexta-feira e sábado);
- Dia da Semana: Quando não for feriado e não for sexta-feira ou sábado.

Sendo esta uma análise de períodos noturnos em Lisboa, faz sentido considerar sextas-feiras e sábados como fim de semana, e domingos como dias de semana, pois a movimentação noturna é maior nas noites que antecedem dias sem obrigações laborais.

A variável “WeekendOrHoliday” classifica os dias como “Fim de Semana”, “Feriado” e “Dia da Semana”. No entanto, para a análise específica de períodos noturnos e vésperas de feriado, era necessário um nível adicional de detalhe que esta variável não fornecia. Desta forma, criou-se a variável “TipoDeDia”⁵.

A análise focada em períodos noturnos exige uma distinção clara entre diferentes tipos de dias durante a noite. Desta forma, existem quatro categorias:

- Vésperas de Feriado: São consideradas das 20h às 23h59 do dia anterior a um feriado, refletindo a antecipação das atividades noturnas em vista do feriado;

⁴ A forma de cálculo da coluna “WeekendOrHoliday” encontra-se no Anexo C

⁵ A nova variável “TipoDeDia” foi criada com a fórmula DAX que se encontra no Anexo C

- Feriado: São considerados especificamente entre 00h e 06h da manhã do dia do feriado, de forma a analisar a extensão das atividades noturnas que começam na noite anterior;
- Fim de Semana: Inclui sextas e sábados à noite;
- Dia da Semana: Qualquer dia que não seja feriado, véspera de feriado, ou fim de semana.

Foi também criada a coluna calculada "NTotal_Telemoveis" ⁶ com o objetivo de permitir a análise conjunta do número total de terminais, abrangendo tanto os terminais estrangeiros quanto os portugueses. A coluna faz o somatório, para cada *grid* ID, dos terminais móveis portugueses e dos estrangeiros. Desta forma, torna-se possível uma compreensão mais abrangente e precisa dos padrões de densidade populacional nas zonas de diversão noturna de Lisboa.

Como uma das análises passa por comparar a concentração de pessoas no verão com as outras estações, criou-se a coluna “Tipo_Estacao”⁷. Desta forma, caso os meses sejam junho, julho ou agosto, caracteriza-se como “Verão” e o resto dos meses como “Outro”.

Em termos de métricas, de forma a perceber, em média, se a concentração populacional varia por tipo de dia, criou-se as métricas “Densidade_Media_DiaDaSemana”, “Densidade_Media_FimDeSemana”, “Densidade_Media_Feriado” e “Densidade_Media_Vespera_Feriado”⁸.

Por fim, em relação às métricas, para entender como varia a densidade média de pessoas no verão em relação às restantes estações, criou-se a “Densidade_Media_Verao” e a “Densidade_Media_Outras_Estacoes”⁹

A tabela “Telemoveis_por_Freguesia” foi construída com o propósito de se compreender de forma precisa o número total de terminais móveis distribuídos pelas freguesias escolhidas. Foi também necessário criar a “Tabela_Ranking” de forma a ser possível ordenar as freguesias por densidade populacional, criando-se uma coluna de *ranking*. A “Tabela_Periodos” e a “Tabela_Horas” foram igualmente elaboradas com o intuito de possibilitar uma análise ordenada e mais precisa dos dados. Estas tabelas permitem a organização sequencial das horas, facilitando a compreensão da evolução ao longo do tempo¹⁰.

⁶ A fórmula da coluna “NTotal_Telemoveis” encontra-se no Anexo C

⁷ A fórmula da coluna “Tipo_Estacao” encontra-se no Anexo C

⁸ As fórmulas, em linguagem DAX, encontram-se no Anexo C

⁹ As fórmulas dessas métricas estão no Anexo C

¹⁰ As fórmulas de cálculo das três tabelas encontram-se no Anexo C

Com as primeiras etapas da preparação desenvolvidas, procede-se à integração dos dados através da construção do modelo de dados.

Por fim, realiza-se a formatação dos dados em Power BI. Por exemplo, uniformizar os formatos de data e hora nas várias tabelas, garantir a consistência das colunas em termos de letras maiúsculas e minúsculas, converter as variáveis para o seu formato correto, entre outros. Este processo assegura que todos os dados estão consistentes e prontos para a análise, eliminando variações que possam causar erros ou dificuldades na interpretação dos resultados.

3.4. Modelação

Neste momento, está tudo preparado para se passar para a fase de modelação. Com base nos dados recolhidos, o objetivo é construir um modelo descritivo da concentração de pessoas nas áreas de diversão noturna. Desta forma, é preciso selecionar as técnicas de modelação adequadas, técnicas de análise descritiva, calculando estatísticas descritivas como a média, utilizar vários gráficos, mapas de calor para perceber a variação da concentração, entre outros. Os gráficos criados serão capazes de serem interpretados pelo público-alvo através do *Power BI*, que é uma ferramenta de alto nível e acessível a todos os utilizadores de forma simples e prática.

Com a visualização dos *dashboards* e visualizações criadas pretende-se ter resposta para as seguintes questões, tendo em conta os objetivos propostos:

- *Objetivo 1: Identificar as zonas mais movimentadas de Lisboa*

- Qual é a distribuição de terminais portugueses e estrangeiros nas diferentes zonas de diversão noturna de Lisboa?
- Como varia a concentração de pessoas nas freguesias selecionadas ao longo da noite?
- Existem padrões específicos de densidade populacional em determinados dias da semana ou em feriados?

- *Objetivo 2: Analisar o Comportamento da População em Eventos Festivos*

- Como varia a densidade populacional nos vários meses do ano?
- Qual é o comportamento da população nas áreas de diversão noturna durante eventos específicos, como as Jornadas Mundiais da Juventude (JMJ)?
- A concentração varia mais durante o verão (junho, julho e agosto) do que no resto do ano?

Optou-se por analisar as Jornadas Mundiais da Juventude por se tratar de um evento de grande dimensão e relevância global, sendo a sua realização em Lisboa um acontecimento provavelmente único. Ao contrário de eventos como a *Web Summit* ou os Santos Populares, que são mais recorrentes na cidade, a JMJ oferece uma oportunidade singular para observar e compreender padrões de mobilidade e comportamento urbano associados a um evento de escala internacional. Esta escolha permite uma análise aprofundada de um evento extraordinário, cujos efeitos na cidade são de uma magnitude que dificilmente se repete, proporcionando informação valiosa sobre a mobilidade em cenários de elevada concentração populacional.

- *Objetivo 3: Avaliar as Possíveis Transferências dos Utilizadores Entre as Zonas de Diversão Noturna*

- Há picos de movimentação que indicam transferências significativas de utilizadores entre determinadas zonas de diversão noturna?
- Onde andam as pessoas na hora de jantar? E depois para onde se deslocam?

- *Objetivo 4: Relacionar os movimentos das pessoas com os níveis de ruído registados nos sensores ambientais nas áreas de estudo consideradas*

- Como os níveis de ruído variam em função da densidade populacional nas freguesias de Lisboa durante a noite?
- Existe uma correlação significativa entre a concentração de pessoas e os níveis de ruído em diferentes áreas de diversão noturna?
- Quais são as localizações dos estabelecimentos com limitadores de ruído e onde estão os principais focos de ruído?
- As JMJ impactaram os níveis de ruído nas áreas de diversão noturna lisboetas?

De forma a responder às questões, como já foi dito, foram desenvolvidas algumas métricas e indicadores (com base na literatura e entrevistas a especialistas da área) utilizados na construção de elementos visuais nos *dashboards*. Na Tabela 3.5 é apresentado de que forma os elementos visuais criados respondem às questões.

Tabela 3.5: Elementos visuais desenvolvidos em Power BI

KPI	Questão Respondida	Gráfico
Número Médio de Terminais	Qual é a distribuição de terminais portugueses e estrangeiros nas diferentes zonas de diversão noturna de Lisboa?	Cartão
Número de Terminais Portugueses	Qual é a distribuição de terminais portugueses e estrangeiros nas diferentes zonas de diversão noturna de Lisboa?	Cartão
Número de Terminais Estrangeiros	Qual é a distribuição de terminais portugueses e estrangeiros nas diferentes zonas de diversão noturna de Lisboa?	Cartão
Concentração de Terminais Portugueses	Como varia a concentração de pessoas nas diferentes freguesias ao longo da noite?	Mapa de Calor
Concentração de Terminais Estrangeiros	Como varia a concentração de pessoas nas diferentes freguesias ao longo da noite?	Mapa de Calor
Concentração de Pessoas por Período do Dia	Como varia a concentração de pessoas nas diferentes freguesias ao longo da noite?	Gráfico em anel
Densidade Populacional por Tipo de Dia	Existem padrões específicos de densidade populacional em determinados dias da semana ou em feriados?	Gráfico de Colunas Empilhadas
Nº Total Terminais por Mês	Como varia a densidade populacional nos vários meses do ano?	Gráfico de Colunas Empilhadas
Densidade Populacional por Tipo de Estação	A concentração varia mais durante o verão (junho, julho e agosto) do que no resto do ano?	Gráfico de Colunas Empilhadas
Concentração de pessoas nas semanas antes, durante e depois das JMJ	Qual é o comportamento da população nas áreas de diversão noturna durante eventos específicos, como as Jornadas Mundiais da Juventude (JMJ)?	Gráfico de Linhas
Concentração de pessoas por freguesia durante as JMJ	Qual é o comportamento da população nas áreas de diversão noturna durante eventos específicos, como as Jornadas Mundiais da Juventude (JMJ)?	Gráfico de Faixas
Transferências entre áreas de diversão noturna	Há picos de movimentação que indicam transferências significativas de utilizadores entre determinadas zonas de diversão noturna?	Gráfico de Linhas
Densidade Populacional por Período do Dia e por Freguesia	Onde andam as pessoas na hora de jantar? E depois para onde se deslocam?	Gráfico de Faixas
Níveis de Ruído por freguesia	Como os níveis de ruído variam em função da densidade populacional nas freguesias de Lisboa durante a noite?	Gráfico de Linhas
Relação da concentração de pessoas com os níveis de ruído registados	Existe uma correlação significativa entre a concentração de pessoas e os níveis de ruído em diferentes áreas de diversão noturna?	Gráfico de Dispersão
Níveis de Ruído por Hora e por Freguesia	Existe uma correlação significativa entre a concentração de pessoas e os níveis de ruído em diferentes áreas de diversão noturna?	Gráfico de Linhas e de Colunas Agrupadas
Localização dos estabelecimentos com limitadores de ruído	Quais são as localizações dos estabelecimentos com limitadores de ruído e onde estão os principais focos de ruído?	Mapa
Máximo do nível de ruído antes, durante e depois das JMJ	As JMJ impactaram os níveis de ruído nas áreas de diversão noturna lisboetas?	Gráfico de Linhas

Dessa maneira, verifica-se que os elementos visuais atendem às necessidades e expectativas delineadas na dissertação. Com isso, torna-se viável prosseguir para a penúltima fase, a avaliação do *dashboard*.

3.5. Avaliação

É importante avaliar o desempenho do modelo e perceber o impacto que poderá ter, no que toca à usabilidade e utilidade. Nesta fase de avaliação, desenvolve-se um questionário direcionado aos especialistas da CML e potenciais utilizadores do *dashboard*. As questões foram formuladas tendo em conta os critérios de avaliação presentes na literatura.

O questionário é composto por sete questões, seis delas com respostas em escala tipo *Likert*, de 1 a 5 pontos, e com uma questão com resposta de Sim/ Não. As perguntas que compõem o questionário estão apresentadas na Tabela 3.6.

Tal como definido na fase de “Compreensão do Negócio”, atinge-se o objetivo analítico quando os utilizadores potenciais do *dashboard* avaliarem a sua satisfação global com uma pontuação de 4 ou superior (numa escala de 1- Nada útil a 5- Extremamente útil).

Tabela 3.6: Questionário de avaliação do *dashboard*

Questões para avaliação do <i>dashboard</i>	
Q1	Como avalia a utilidade do <i>dashboard</i> para monitorização da concentração de pessoas nas zonas de diversão noturna de Lisboa? (numa escala de 1-Nada útil a 5-Extremamente útil)
Q2	Como avalia a facilidade de utilização do <i>dashboard</i> , numa escala de 1 (nada fácil) a 5 (extremamente fácil)?
Q3	Considera que os indicadores apresentados neste <i>dashboard</i> são adequados para monitorização da mobilidade urbana. Dê a sua resposta numa escala de 1 (nada adequado) a 5 (extremamente adequado)
Q4	Sugere acrescentar nova(s) característica(s) a este <i>dashboard</i> ?
Q5	Como avalia os gráficos utilizados para apresentar os diferentes tipos de indicadores? Responda numa escala de 1 (nada adequado) a 5 (extremamente adequado)
Q6	Como avalia o design visual do <i>dashboard</i> (uso de cores, gráficos, organização)? Dê a sua resposta numa escala de 1 (nada adequado) a 5 (extremamente adequado)
Q7	Qual é a sua satisfação geral com o <i>dashboard</i> , numa escala de 1 (nada útil) a 5 (extremamente útil)?

3.6. Implementação

A criação do modelo não é geralmente o final do projeto. Mesmo que o objetivo do modelo seja aumentar o conhecimento sobre os dados, o conhecimento adquirido terá de ser organizado e apresentado de forma que o cliente o possa utilizar efetivamente (Chapman et al., 2000).

Esta fase envolve transformar os *insights* obtidos em uma ferramenta prática e utilizável, como o *dashboard* desenvolvido. É essencial que os *dashboards* sejam projetados para incorporar indicadores e métricas relevantes que podem ser atualizados automaticamente com novos dados (Zhuang et al., 2022), permitindo que os utilizadores obtenham *insights* atualizados sem necessidade de intervenção humana constante.

A etapa de implementação envolve a elaboração do presente documento, bem como a sua apresentação, detalhando todas as tarefas realizadas. Além disso, esta etapa inclui conceder à Câmara Municipal de Lisboa o *dashboard* desenvolvido e perceber qual o contributo do mesmo para os profissionais da área e para a literatura.

CAPÍTULO 4

4. Resultados e discussão

Este capítulo apresenta o resultado após a fase de modelação. Apresenta um *dashboard* tático para monitorizar a concentração e mobilidade das pessoas nas áreas de diversão noturna de Lisboa.

Como já referido anteriormente, de forma a facilitar a compreensão e visualização por parte da CML, recorreu-se à ferramenta *Power BI*. Neste *dashboard* são apresentados os indicadores de maior importância, sendo possível aplicar filtros temporais e de localização. Neste capítulo ir-se-á proceder à apresentação do *dashboard* concluído e à explicação detalhada do mesmo, evidenciando-se a etapa de avaliação da metodologia CRISP-DM.

Para garantir uma melhor compreensão, optou-se por desenvolver um *dashboard* distinto para cada um dos quatro objetivos definidos. Essa abordagem permite uma apresentação mais clara e focada dos dados, facilitando a análise e interpretação das informações específicas de cada objetivo e atendendo de forma eficaz às necessidades dos utilizadores. No entanto, ao longo de toda a dissertação, refere-se ao *dashboard* no singular, pois foi criado um único modelo unificado a partir das fases de compreensão, preparação e modelação.

4.1. *Dashboard* de monitorização da mobilidade urbana nas zonas de diversão noturna de Lisboa

4.1.1. Objetivo 1: Identificar as zonas mais movimentadas de Lisboa

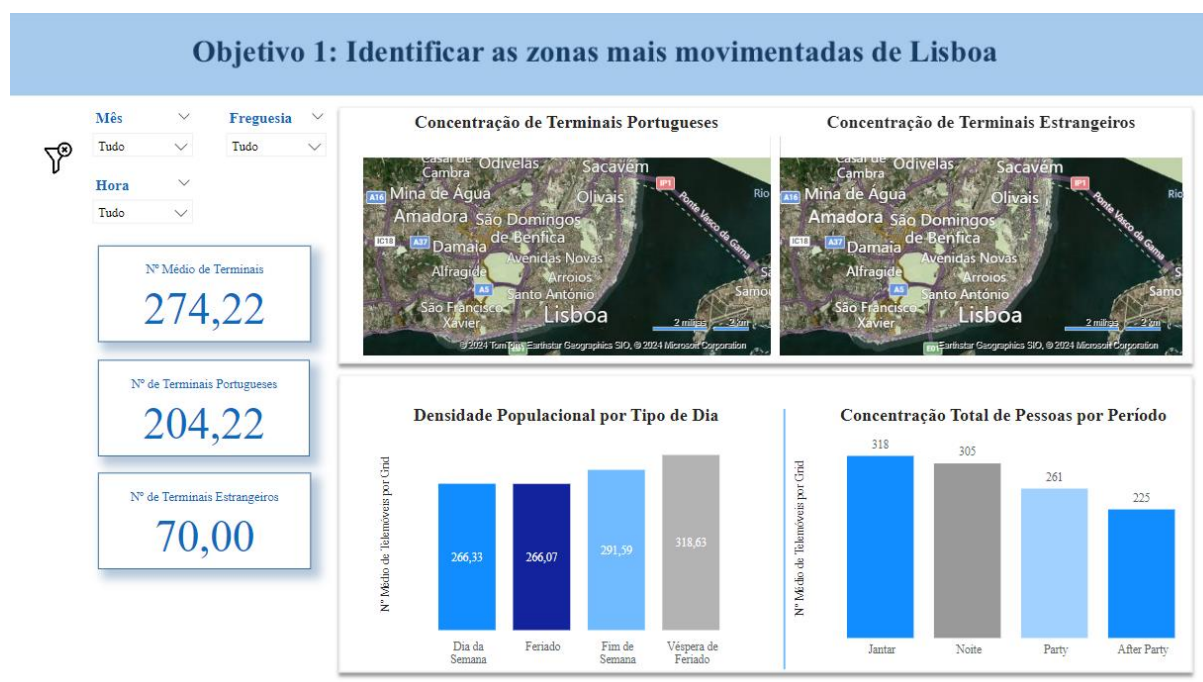


Figura 4.1: *Dashboard* para monitorizar áreas de diversão noturna mais movimentadas de Lisboa

Este *dashboard* (Figura 4.1) foi desenvolvido de forma a monitorizar as áreas de diversão noturna mais movimentadas de Lisboa. A seguir, será apresentada uma análise detalhada de cada componente do *dashboard*, que responde às questões de distribuição de terminais portugueses e estrangeiros, variações de concentração ao longo da noite, e padrões de densidade populacional em diferentes dias.

No canto superior esquerdo, são dispostos dois filtros essenciais: zona geográfica e período temporal (Figura 4.2). Estes filtros permitem aos utilizadores analisar e monitorizar a concentração de pessoas nas áreas de diversão noturna por freguesia e ajustar o intervalo de tempo para a análise, influenciando todos os elementos visuais do *dashboard* quando seleccionados. Se o utilizador seleccionar o ícone de remoção dos filtros, limpa todas as segmentações de dados.

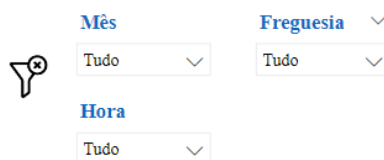


Figura 4.2: Filtros disponíveis no *dashboard*

Do lado esquerdo, são também apresentados indicadores considerados cruciais para o controlo e monitorização das zonas de diversão noturna (Figura 4.3). Foram desenvolvidos quatro indicadores:

- Número Médio de Terminais por *Grid* (274,22): indica a densidade média por *grid*, ajudando a identificar os *hotspots* de concentração populacional.
- Número de Terminais Portugueses (204,22);
- Número de Terminais Estrangeiros (70,00).



Figura 4.3: Indicadores de controlo (N° Médio de Terminais por *Grid*, N° de Terminais Portugueses e N° de Terminais Estrangeiros)

Os mapas de calor, localizados na parte superior direita do *dashboard*, visualizam a concentração de terminais portugueses e estrangeiros nas diferentes zonas de Lisboa. De acordo com a Figura 4.4, é possível perceber que as áreas com maior concentração, a azul-escuro, incluem Santa Catarina, Bairro Alto e Cais do Sodré. Estas são zonas populares pela vida noturna ativa, com muitos bares e restaurantes. As áreas com menor concentração, a azul-claro, são Alcântara, Santo António e São Vicente, que, embora movimentadas, têm uma densidade inferior de terminais portugueses. Em relação a terminais estrangeiros, como se também pode ver pela Figura 4.4, as áreas com maior concentração, são Baixa/Chiado, Cais do Sodré e Santa Catarina, refletindo o seu apelo turístico. As zonas com menor concentração são Alcântara, Santo António e São Vicente, atraindo menos turistas comparadas às zonas centrais. A análise dos mapas de calor revela que as zonas centrais como Santa Catarina, Bairro Alto, Cais do Sodré e Baixa/Chiado são os principais pontos de concentração noturna tanto para residentes quanto para turistas. Áreas como Alcântara, Santo António e São Vicente mostram uma menor densidade populacional noturna.



Figura 4.4: Concentração de Terminais Portugueses e Estrangeiros

Na parte inferior do *dashboard*, são apresentados gráficos que mostram a concentração total de pessoas por período do dia (“Jantar”, “Noite”, “Party” e “After party”) e a densidade populacional por tipo de dia (dia da semana, feriado, fim de semana e véspera de feriado), de forma a perceber como varia a concentração de pessoas.

O gráfico de barras do lado direito (Figura 4.5) apresenta a distribuição dos terminais móveis ao longo de diferentes períodos do dia. A análise revela que os períodos de "Jantar" e "Noite" apresentam as maiores concentrações de terminais móveis, sugerindo que estes são os horários de maior movimentação. Durante o jantar, as pessoas começam-se a reunir em restaurantes e bares, enquanto no período da noite, a concentração aumenta, refletindo o pico de atividades noturnas. A fase da festa ("Party") também tem uma alta concentração, indicando que muitas pessoas continuam nas áreas de diversão até tarde da madrugada. O período "After Party" mostra uma redução na concentração, mas ainda mantém uma quantidade significativa de pessoas, especialmente em locais que permanecem abertos até de manhã.

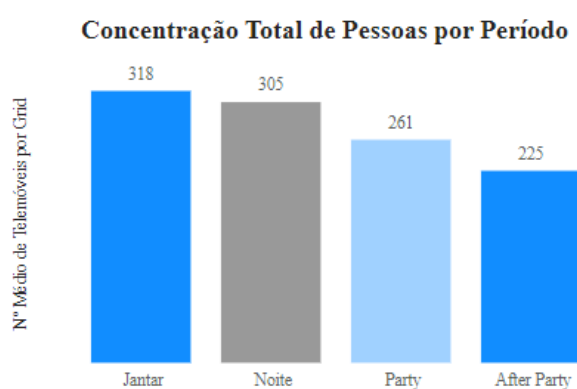


Figura 4.5: Concentração de Pessoas por Período do Dia

O gráfico de barras do lado esquerdo (Figura 4.6) compara a densidade populacional por tipo de dia. A análise mostra que a densidade populacional é maior nos fins de semana e vésperas de feriado, sugerindo que nesses dias há uma maior atividade noturna. Este comportamento deve-se ao facto de as pessoas geralmente terem mais tempo livre e disposição para sair e socializar durante esses dias, visto que, normalmente, não se trabalha no(s) dia(s) seguinte(s).

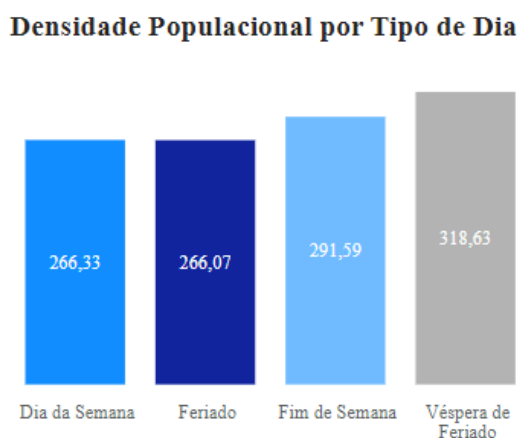


Figura 4.6: Densidade Populacional por Tipo de Dia

Quando o utilizador faz *mouse over* no gráfico da “Concentração Total de Pessoas por Período” consegue analisar com detalhe a concentração de pessoas nas horas correspondentes ao período “selecionado” ¹¹.

4.1.2. Objetivo 2: Analisar o Comportamento da População em Eventos Festivos

Para estudar este objetivo, foi construído um *dashboard* com quatro visualizações distintas.

Este *dashboard* (Figura 4.7) foi criado para responder às questões sobre a variação da densidade populacional nos vários meses do ano, se a concentração populacional é maior durante o verão (junho, julho e agosto) em comparação com o resto do ano e como varia a concentração de pessoas antes, durante e depois das JMJ. Os únicos gráficos que foi necessário implementar um filtro temporal foram os que abordam o tema das JMJ. No entanto, para cada um destes dois gráficos, foi aplicado um filtro temporal diferente, uma vez que um trata da densidade populacional antes, durante e após as JMJ e outro apenas analisa a densidade durante o evento. As JMJ decorreram de 1 a 6 de agosto de 2024 e, desta forma, o período de análise foi de 25 de julho de 2024 a 13 de agosto de 2024

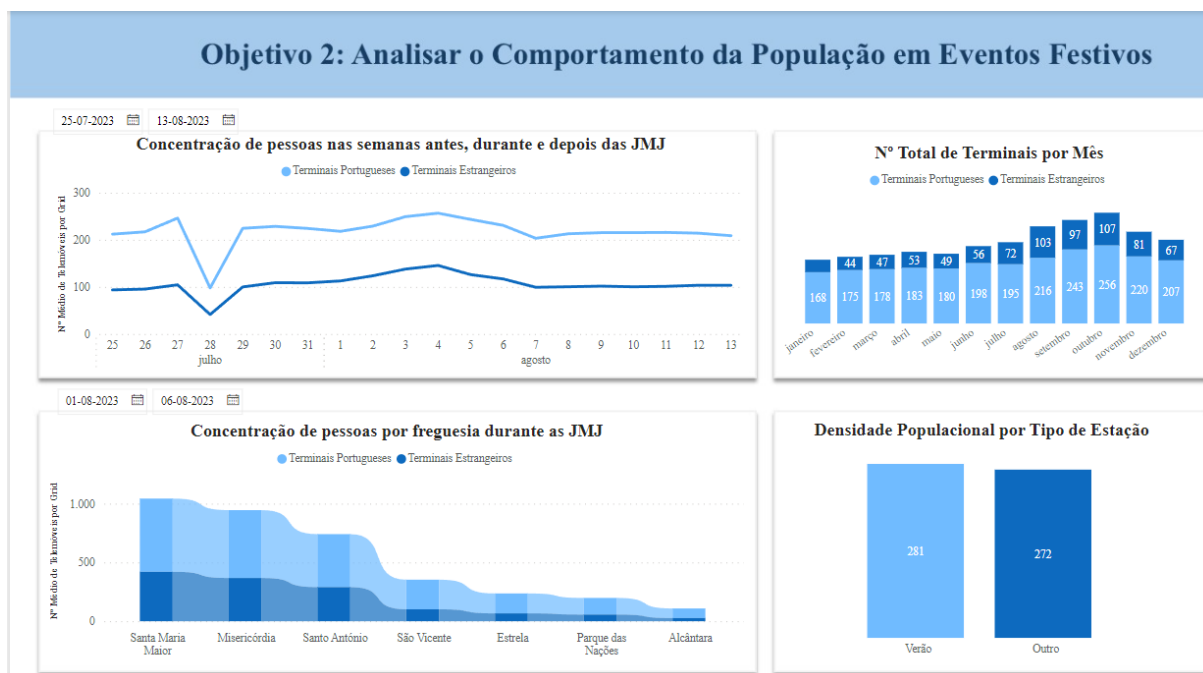


Figura 4.7: *Dashboard* de monitorização das Pessoas em Eventos Festivos

¹¹ Encontra-se no Anexo D como se visualiza o detalhe do gráfico quando o utilizador faz *mouse over*

O objetivo da visualização do lado superior esquerdo (Figura 4.8) é perceber se as JMJ foram acompanhadas por um aumento perceptível da densidade populacional. Ao analisar as três semanas, é possível obter informações sobre o impacto do evento no fluxo e na concentração de pessoas em Lisboa, ajudando a entender os efeitos mais amplos de um evento de tão grande escala na dinâmica da população local.

Cerca de três dias antes do evento, observa-se um aumento gradual na concentração de pessoas, sugerindo que tanto visitantes internacionais quanto residentes locais começaram a concentrar-se em Lisboa por antecipação ao evento. Durante o período das JMJ, a concentração de pessoas atinge seu pico, especialmente no dia 4 de agosto de 2024, refletindo a grande participação no evento. Após o dia 4 de agosto, observa-se uma diminuição na concentração de pessoas, mas é entre os dias 6 e 7 de agosto que ocorre uma descida mais acentuada, sugerindo que muitos participantes deixaram a cidade imediatamente após o encerramento do evento. Após o término das JMJ, a concentração de pessoas diminui gradualmente, retornando aos níveis próximos aos observados antes do evento. Este gráfico demonstra claramente que as JMJ tiveram um impacto significativo na densidade populacional de Lisboa, com um aumento notável durante o evento, seguido por uma normalização gradual após o encerramento do mesmo.

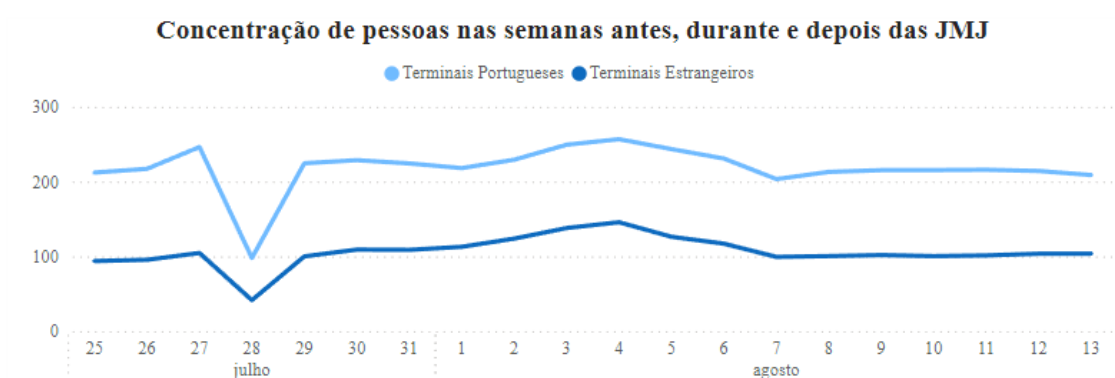


Figura 4.8: Concentração Populacional antes, durante e após as JMJ

O gráfico de barras à direita do lado superior (Figura 4.9) mostra o número médio de terminais móveis, divididos entre portugueses e estrangeiros, para cada mês do ano. De janeiro a maio, a densidade média é relativamente baixa, com cerca de 210 a 240 terminais, refletindo um período de menor atividade populacional. Em junho nota-se um aumento gradual, atingindo 254 terminais, indicando o início da temporada de verão e um aumento na movimentação. A partir do mês de julho, os valores são os mais altos, com um pico em outubro de 363 terminais. Nos meses posteriores a outubro, a densidade diminui novamente. Outubro é o mês com maior concentração de pessoas e maio destaca-se como o mês de menor concentração.

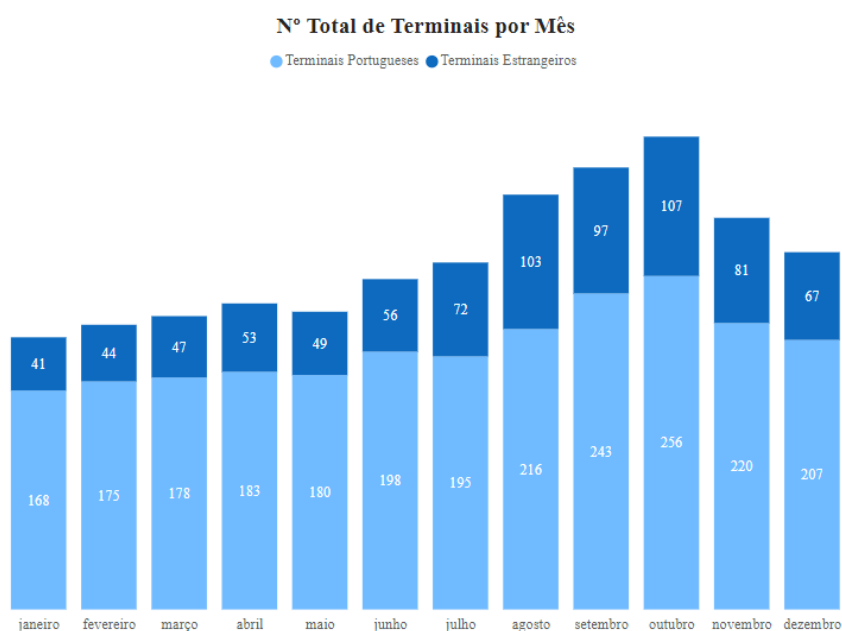


Figura 4.9: Nº de Terminais por mês

O gráfico “Concentração de pessoas por freguesia durante as JMJ” (Figura 4.10) desenvolveu-se com o intuito de perceber quais as freguesias com maior densidade populacional durante o grande evento das JMJ. Tal como os outros gráficos, apresenta a média de telemóveis por *grid* (a cada 15 minutos). Observa-se que a freguesia de Santa Maria Maior apresenta a maior concentração de pessoas, com um elevado número de terminais portugueses e estrangeiros. A freguesia do Parque das Nações (freguesia onde se realizaram as JMJ, mais propriamente no Parque Tejo) também não se destaca, revelando que não houve nenhum aumento considerável no número de pessoas que geralmente apresenta.

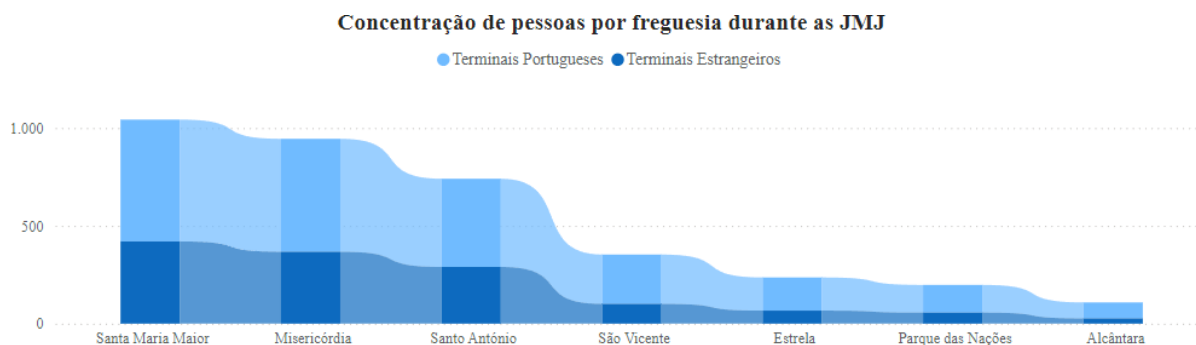


Figura 4.10: Concentração de Pessoas durante as JMJ

O gráfico "Densidade Populacional por Tipo de Estação" (Figura 4.11) compara a média da densidade populacional entre o verão e as outras estações do ano. Segundo os dados apresentados, a densidade populacional no verão é ligeiramente superior, com um valor de 281 pessoas por *grid*, em comparação com 272 pessoas no resto do ano. Embora o gráfico indique que o verão tem uma densidade populacional média maior, a diferença de apenas 9 pessoas não é particularmente significativa. Isso sugere que, embora haja um aumento na concentração populacional durante o verão, esta variação não é suficientemente marcante para indicar uma mudança drástica na dinâmica populacional desta estação em comparação com as restantes do ano.

Densidade Populacional por Tipo de Estação

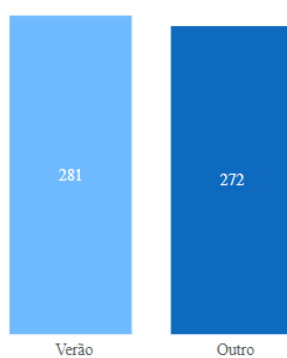


Figura 4.11: Densidade Populacional por Tipo de Estação

4.1.3. Objetivo 3: Avaliar as Possíveis Transferências dos Utilizadores Entre as Zonas de Diversão Noturna

Para responder ao objetivo 3, que consiste em avaliar as possíveis transferências de utilizadores entre as zonas de diversão noturna, foi desenvolvido outro *dashboard* (Figura 4.12).

O *dashboard* tem como finalidade analisar se há picos de movimentação que indiquem transferências significativas de pessoas entre as diferentes áreas de entretenimento noturno. Além disso, pretende entender os padrões de movimentação dos utilizadores, identificando onde eles se concentram no início da noite e para onde se deslocam após esse período.

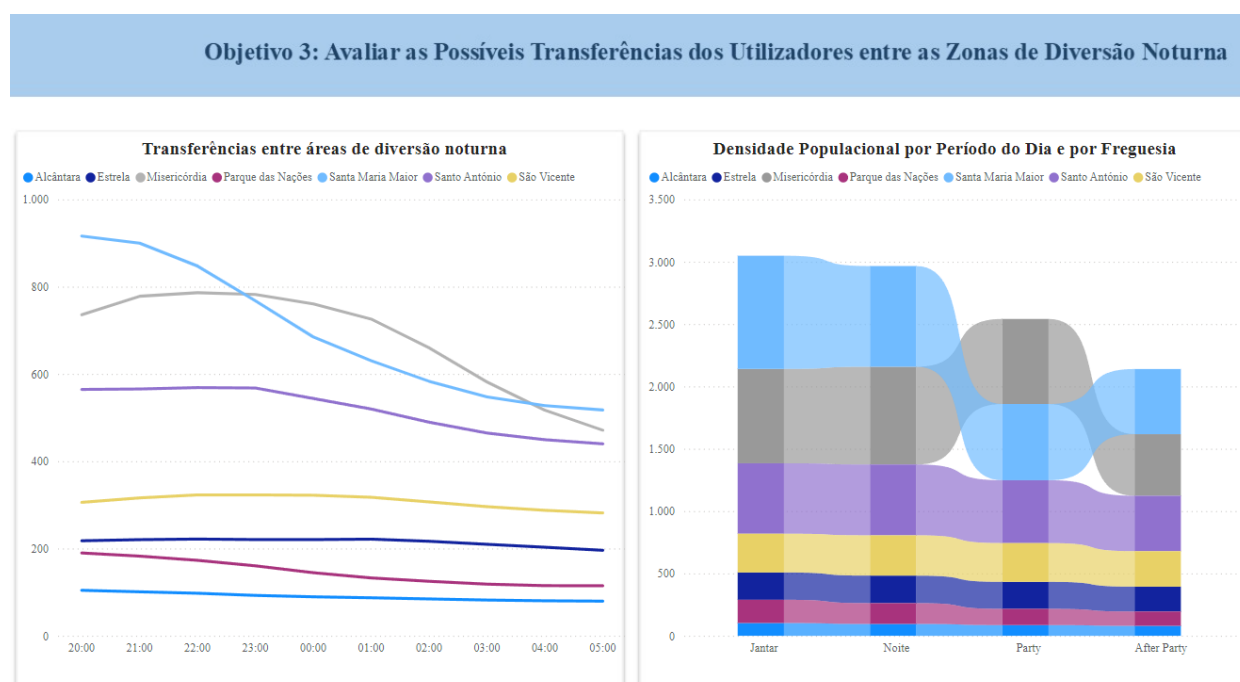


Figura 4.12: *Dashboard* de monitorização de transferências de utilizadores

O gráfico da esquerda (Figura 4.13) mostra as transferências de pessoas entre as diferentes freguesias de Lisboa ao longo da noite, desde as 20h00 às 05h00, representando como a concentração média de pessoas por *grid* (eixo Y) muda em cada freguesia com o passar do tempo (eixo X).

A freguesia de Santa Maria Maior destaca-se no início da noite com o maior número de pessoas, no entanto, essa concentração diminui gradualmente conforme a noite avança. A partir das 23h00, a freguesia da Misericórdia começa a aumentar a sua densidade populacional enquanto Santa Maria Maior diminui, sugerindo que Santa Maria Maior é popular para o Jantar e, após esse período, as pessoas deslocam-se para a Misericórdia.

Em geral, a maioria das freguesias mostra uma diminuição no número de pessoas ao longo da noite, sugerindo que as pessoas vão-se dispersando gradualmente.

Um destaque adicional é o comportamento da freguesia de Santa Maria Maior que, a partir das 04h00, a sua densidade volta a aumentar, sendo um sítio procurado no período “After Party”.

Este padrão sugere que as pessoas tendem a começar a noite concentradas em algumas áreas principais e vão-se dispersando ao longo da noite, mas sem grandes deslocamentos entre freguesias. A movimentação entre as áreas parece ser mais gradual e contínua, com Santa Maria Maior e Misericórdia sendo uns dos principais focos de concentração ao longo de toda a noite.

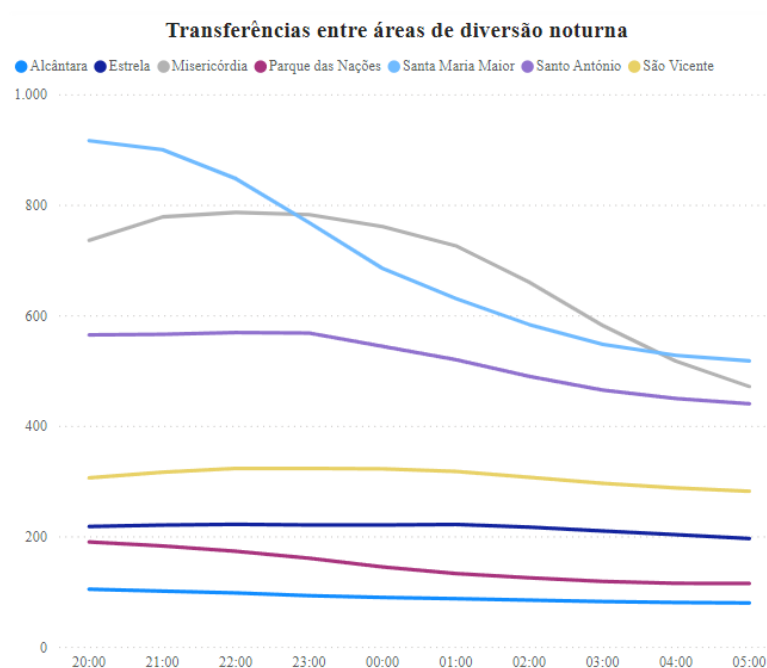


Figura 4.13: Transferência entre áreas de diversão noturna

O gráfico da "Densidade Populacional por Período do Dia e por Freguesia" (Figura 4.14) também permite analisar as transferências entre áreas ao longo da noite, embora de forma menos clara. Esta visualização foi especificamente criada para responder à pergunta "Onde andam as pessoas na hora de jantar? E depois, para onde se deslocam?".

Após análise do gráfico, confirma-se que Santa Maria Maior é a freguesia com a maior concentração de pessoas durante o período do “Jantar” e de “Noite”, reduzindo a sua concentração no período da “Party”

O gráfico mostra que, no período de "Jantar", as pessoas estão distribuídas de maneira relativamente uniforme entre várias freguesias, com destaque para o Santa Maria Maior, Misericórdia e Santo António. Como referido anteriormente, apenas se nota transferência de utilizadores da freguesia de Santa Maria Maior para a Misericórdia e só após o período da "Noite". Todas as outras freguesias têm um comportamento semelhante, sem grandes diferenças na densidade populacional ao longo da noite.

À medida que a noite avança para o período "Party", observa-se um menor pico de concentração. No período "After Party", a concentração populacional diminui em todas as freguesias, com Santa Maria Maior voltando a ser a mais frequentada.

Este gráfico não apenas confirma as áreas de maior concentração durante diferentes períodos da noite, mas também destaca como a movimentação entre freguesias após o período do "Jantar".

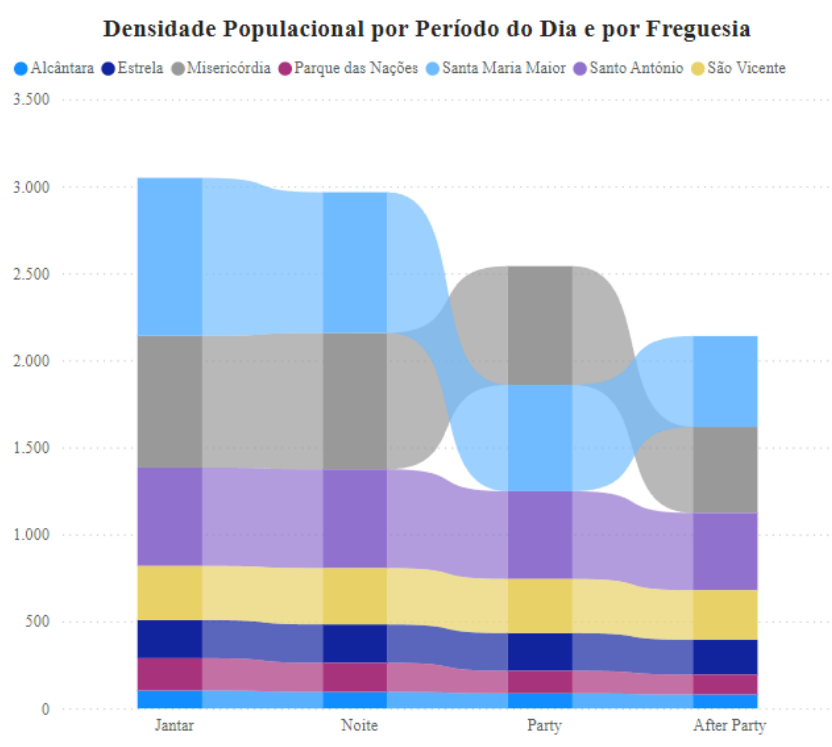


Figura 4.14: Densidade Populacional por Período do Dia e por Freguesia

4.1.4. Objetivo 4: Relacionar os movimentos das pessoas com os níveis de ruído registados nos sensores ambientais nas áreas de estudo consideradas

Para responder a este objetivo e analisar o ruído, foram utilizados dados da rede de sensores ambientais da cidade de Lisboa. As leituras são médias horárias obtidas tendo por base valores médios de períodos de 15 minutos.

Após uma análise, percebeu-se que existem 60 sensores com limitadores de ruído nas freguesias selecionadas e a sua localização apresenta-se na Figura 4.15.

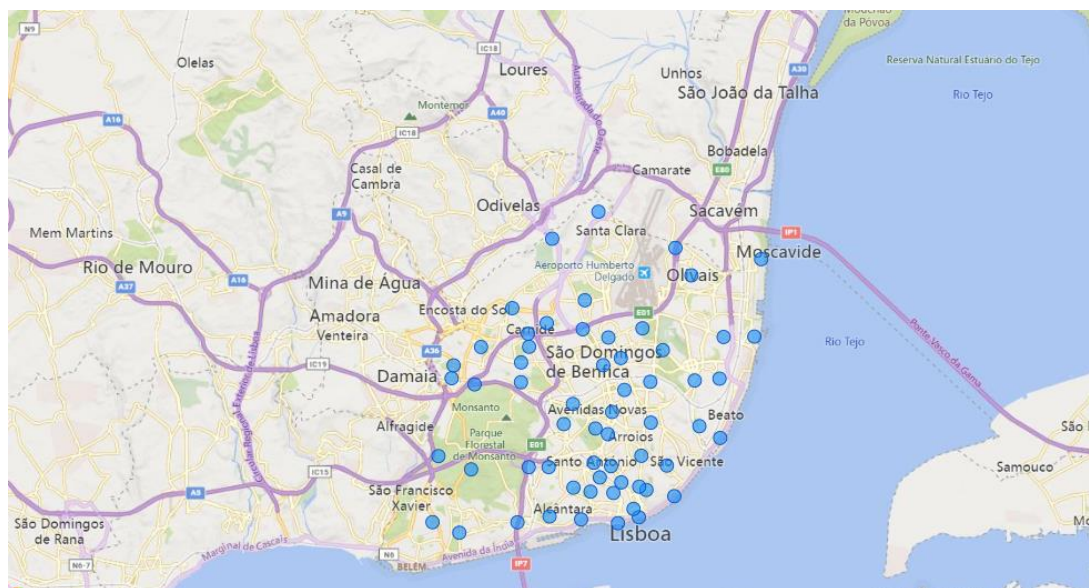


Figura 4.15: Localização dos sensores com limitadores de ruído

Existem quatro tipos de níveis de ruído: Normal (0 a 54 dB); Moderado (55 a 59 dB); Elevado (60 a 64 dB) e Muito Elevado (>65 dB).

De forma a obter uma primeira visão dos valores de ruído registados, a Tabela 4.1 foi elaborada como ponto de partida para uma breve análise exploratória dos dados, onde se analisam as seguintes medidas: contagem, mínimo, máximo, média e a proporção de cada nível de ruído face ao total.

Tabela 4.1: Indicadores dos sensores de ruído

Nível da Etiqueta	Quantidade	Mínimo	Máximo	Média	%
Normal	73 067	29	54	46,86	24%
Moderado	48 120	55	59	59,95	16%
Elevado	55 760	60	64	62,01	18%
Muito Elevado	129 396	65	99	70,61	42%
Total	306 343	29	99	59,86	100%

Após análise da tabela, percebe-se que 42% do total corresponde à etiqueta “Muito Elevado”, concluindo que Lisboa é uma cidade barulhenta. E será que existe uma relação entre a concentração de pessoas e os níveis de ruído em diferentes áreas de diversão noturna? Quais as horas e freguesias mais barulhentas? Qual o valor máximo antes, durante e após as JMJ? Para responder a estas questões foi desenvolvido outro *dashboard* (Figura 4.16).

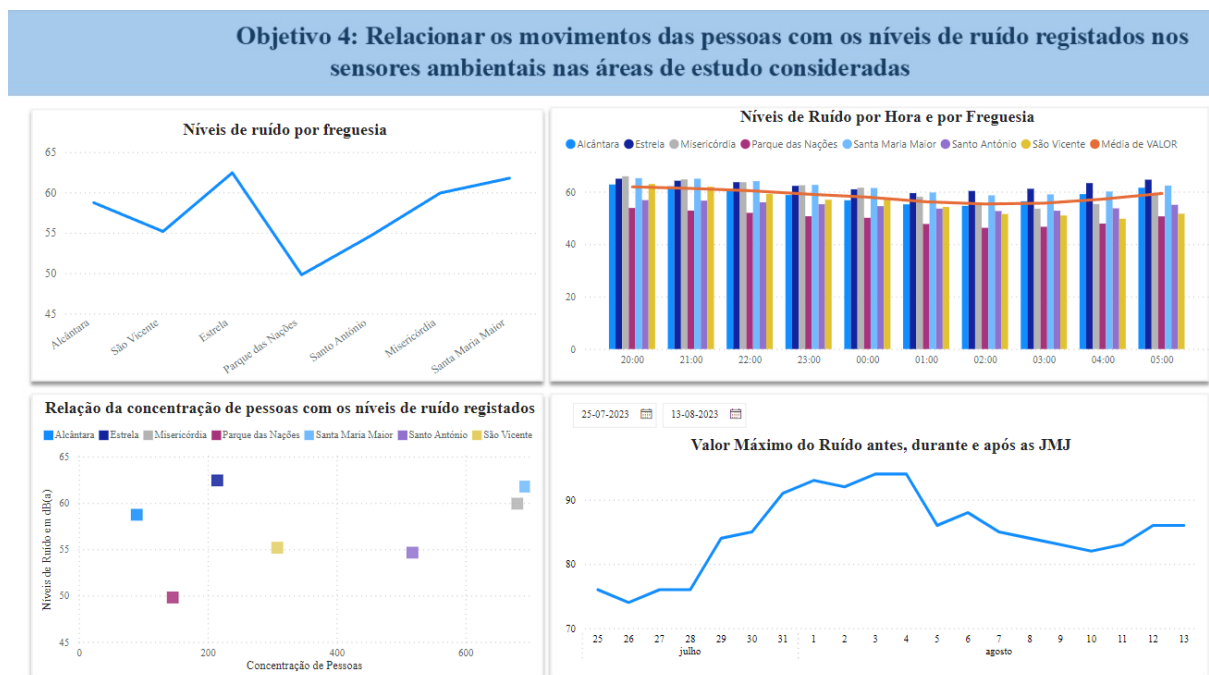


Figura 4.16: *Dashboard* de monitorização da concentração de pessoas e dos níveis de ruído

Foram elaborados dois gráficos com conclusões semelhantes, o “Níveis de Ruído por freguesia” e o “Relação de concentração de pessoas com os níveis de ruído registados”. O primeiro, apresentado na Figura 4.17, mostra o nível de ruído para da freguesia. É de notar que as freguesias estão ordenadas de forma ascendente, ou seja, da freguesia com menor concentração de pessoas (Alcântara) para aquela com maior concentração (Santa Maria Maior).

Este gráfico sugere que não há uma relação clara entre a concentração de pessoas e os níveis de ruído nas freguesias analisadas. Freguesias como Alcântara e São Vicente, que têm uma menor concentração de pessoas, apresentam níveis de ruído relativamente altos. Este facto contraria a expectativa de que menos pessoas significariam menos ruído.

A freguesia com valores de ruído mais baixos é o Parque das Nações, enquanto a estrela é a que apresenta valores mais elevados.

Este gráfico destaca que não existe uma correlação direta entre o número de pessoas presentes numa freguesia e os níveis de ruído registados, indicando que outros fatores, além da concentração populacional, influenciam significativamente os níveis de ruído nas diferentes áreas de Lisboa.



Figura 4.17: Níveis de Ruído por Freguesia

Em relação ao gráfico de dispersão (Figura 4.18), como dito anteriormente, as conclusões são semelhantes. Cada ponto no gráfico representa uma freguesia, identificada por uma cor específica, e posicionada de acordo com a sua concentração de pessoas (no eixo X) e o nível de ruído medido em decibéis (no eixo Y).

No geral, o gráfico de dispersão demonstra que a relação entre o número de pessoas numa área e o nível de ruído é complexa e não linear, indicando que a concentração de pessoas não é o único fator determinante para o ruído registado em diferentes freguesias de Lisboa.

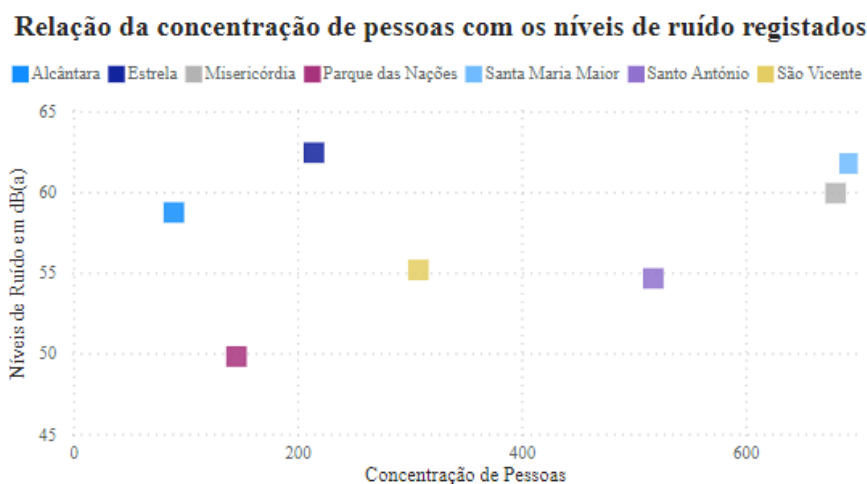


Figura 4.18: Relação da concentração de pessoas com os níveis de ruído registados

De modo realizar-se uma análise mais detalhada e perceber quais as horas e freguesias com mais barulho, desenvolveu-se o gráfico apresentado na Figura 4.19. Trata-se de um gráfico de linhas e colunas agrupadas onde a linha corresponde ao nível médio de ruído ao longo da noite e as colunas representam nível de ruído médio (eixo Y) de cada freguesia (dependendo da cor).

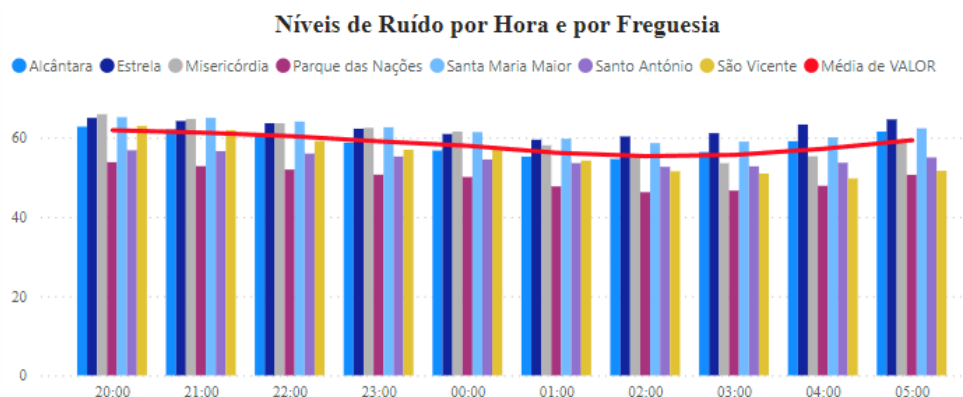


Figura 4.19: Níveis de Ruído por Hora e por Freguesia

Em relação às freguesias, as conclusões são semelhantes com as retiradas anteriormente. Interessante notar que, apesar das variações nos níveis de ruído ao longo da noite, todas as freguesias apresentam uma relativa consistência nos níveis de ruído, com as barras mantendo-se próximas umas das outras em cada intervalo de tempo.

Em relação às horas, o gráfico de linhas e colunas agrupadas apresenta a média horária dos níveis de ruído, representada pela linha vermelha.

É visível que, no início da noite, os níveis de ruído são relativamente elevados, com todas as freguesias apresentando valores semelhantes e a média de ruído situando-se em torno de 60 dB. Conforme a noite avança, especialmente a partir das 23h00, há uma tendência de queda nos níveis de ruído médio, com a linha vermelha descendo gradualmente até o ponto mais baixo, que ocorre por volta das 02h00.

Este declínio sugere que, durante a madrugada, há uma redução significativa nas atividades que contribuem para o ruído, possivelmente devido ao fecho de estabelecimentos ou à dispersão de pessoas. No entanto, após esse período de menor atividade, observa-se uma recuperação nos níveis de ruído a partir das 03h00 até as 05h00.

Em resumo, o gráfico demonstra que o nível de ruído médio diminui ao longo da noite, atingindo o ponto mais baixo às 02h00, antes de começar a subir novamente, refletindo a dinâmica das atividades noturnas nas diferentes freguesias de Lisboa. No entanto, é de notar que as variações são pouco acentuadas, apresentando um ruído médio por hora semelhante ao longo da noite.

Por fim, de forma a analisar se as JMJ impactaram os níveis de ruído nas áreas de diversão noturna lisboetas, construiu-se o gráfico da Figura 4.20. O período temporal, representado por um filtro, é desde 25 de julho de 2024 (uma semana antes das JMJ) até 13 de agosto de 2024 (uma semana após as JMJ).

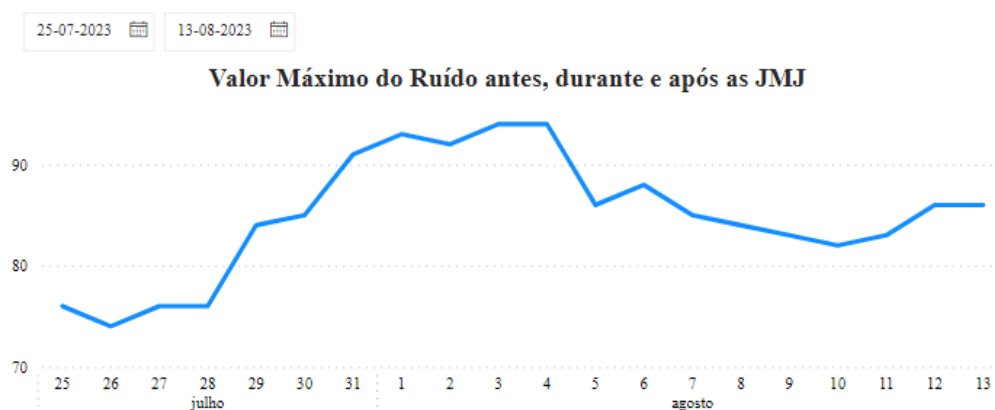


Figura 4.20: Valor Máximo do Ruído antes, durante e após as JMJ

Na semana que antecedeu as JMJ, os níveis de ruído (eixo Y) mantiveram-se relativamente estáveis e baixos. A partir do dia 29 de julho, os níveis de ruído começaram a aumentar, atingindo um pico a 4 de agosto.

Durante as JMJ, os níveis de ruído atingiram seus valores mais altos, refletindo o aumento de atividades e a presença de grandes multidões em Lisboa. O pico mais alto foi registado no dia 3 e 4 de agosto, indicando que esses foram os dias mais movimentados em termos de ruído. A partir do dia 4 de agosto, os níveis de ruído começaram a diminuir significativamente, acompanhando o encerramento gradual das atividades das JMJ.

Após o término das JMJ, houve uma queda nos níveis de ruído, sugerindo uma normalização da atividade na cidade, com o ruído retornando aos níveis habituais. A partir do dia 10 de agosto, notou-se uma ligeira recuperação nos níveis de ruído, embora sem alcançar os valores elevados observados durante as JMJ.

Resumindo, o gráfico demonstra que as JMJ tiveram um impacto significativo nos níveis de ruído em Lisboa, com picos substancialmente mais altos durante o evento, seguidos por uma redução acentuada após o seu término, refletindo a normalização da vida na cidade.

4.2. Avaliação do *dashboard*

De forma a avaliar o *dashboard* desenvolvido, aplicou-se o questionário apresentado na fase da “Avaliação” do CRISP-DM. Na Tabela 4.2 apresentam-se as questões e as respetivas respostas

de três potenciais utilizadores do *dashboard*. As questões foram enviadas para a equipa LxDataLab da CML e foram respondidas por três responsáveis do tema, sendo dois deles operacionais e um chefe de secção¹².

Tabela 4.2: Avaliação do *dashboard* por utilizadores

Questões para avaliação do <i>dashboard</i>	
Q1	Como avalia a utilidade do <i>dashboard</i> para monitorização da concentração de pessoas nas zonas de diversão noturna de Lisboa? (numa escala de 1-Nada útil a 5-Extremamente útil)
R1	5. Acho muito útil, especialmente para monitorizar a concentração nas zonas mais movimentadas e poder tomar medidas, caso necessário.
R2	4. O <i>dashboard</i> oferece uma visão geral útil, mas poderia incluir detalhe sobre mais eventos específicos, para além das JMJ.
R3	5. Muito útil para perceber a movimentação e as transferências entre zonas.
Q2	Como avalia a facilidade de utilização do <i>dashboard</i>, numa escala de 1 (nada fácil) a 5 (extremamente fácil)?
R1	5. <i>Layout</i> de fácil intuição, com facilidade de utilização dos filtros.
R2	5. Muito fácil de utilizar, a navegação é intuitiva e os dados estão bem organizados.
R3	5. Sou pouco experiente em Power Bi e, ainda assim, achei a navegação muito fácil.
Q3	Considera que os indicadores apresentados neste <i>dashboard</i> são adequados para monitorização da mobilidade urbana. Dê a sua resposta numa escala de 1 (nada adequado) a 5 (extremamente adequado)
R1	5. Os indicadores são relevantes para fornecer uma visão geral clara e útil da mobilidade, abrangendo aspetos essenciais para analisar a movimentação das pessoas.
R2	4. Acredito que os indicadores cobrem bem as principais necessidades de análise da monitorização da mobilidade, mas alguns detalhes adicionais poderiam ser incluídos
R3	5. Os indicadores apresentados são suficientes e adequados para avaliar a mobilidade e os padrões de comportamento de forma eficiente.
Q4	Sugere acrescentar nova(s) característica(s) a este <i>dashboard</i>?
R1	Não.
R2	Sim, talvez uma previsão do fluxo de pessoas em eventos futuros.
R3	Não, acho que o <i>dashboard</i> responde a todos os objetivos mencionados.
Q5	Como avalia os gráficos utilizados para apresentar os diferentes tipos de indicadores? Responda numa escala de 1 (nada adequado) a 5 (extremamente adequado)
R1	5. Os gráficos escolhidos estão adequados para a análise realizada.
R2	5. Os gráficos são adequados e de fácil visualização.
R3	5. Os gráficos são claros e fáceis de entender.
Q6	Como avalia o design visual do <i>dashboard</i> (uso de cores, gráficos, organização)? Dê a sua resposta numa escala de 1 (nada adequado) a 5 (extremamente adequado)
R1	5. As cores são bastante perceptíveis, permitindo identificar facilmente áreas de maior densidade populacional.
R2	4. Numa forma geral, está adequado.
R3	5. <i>Design</i> bem escolhido.
Q7	Qual é a sua satisfação geral com o <i>dashboard</i>, numa escala de 1 (nada útil) a 5 (extremamente útil)?
R1	5. Estou bastante satisfeito com o <i>dashboard</i> . Atende bem às minhas expectativas, pois oferece uma interface fácil de usar e informações relevantes para a monitorização da mobilidade urbana. A clareza dos gráficos e a forma como as informações são apresentadas facilitam muito a tomada de decisões rápidas.
R2	4. O <i>dashboard</i> é muito útil e bem desenhado, com dados apresentados de forma clara. No entanto, acredito que poderia ser melhorado com mais análises como relacionar com as condições climáticas, por exemplo. Ainda assim, é uma ferramenta valiosa para a monitorização de mobilidade.
R3	5. Estou extremamente satisfeito com o <i>dashboard</i> . Ele é intuitivo, proporciona <i>insights</i> valiosos e oferece uma visão clara dos padrões de mobilidade nas zonas de diversão noturna. A combinação de dados de mobilidade com outras métricas, como níveis de ruído, é uma grande mais-valia.

¹² De acordo com as leis de proteção de dados, os nomes não são divulgados neste contexto para garantir a segurança e a privacidade das informações pessoais.

Com as respostas dadas pelos potenciais utilizadores, conclui-se que o *dashboard* atingiu o critério de sucesso estabelecido, que exigia uma pontuação global de quatro ou superior. Em termos de utilidade, foi considerado muito útil para acompanhar a concentração de pessoas nos locais de diversão noturna de Lisboa. No entanto, houve recomendações esporádicas para incluir mais informações sobre eventos específicos. Os utilizadores que têm uma familiaridade mínima com o *Power BI* ou outras ferramentas semelhantes elogiaram a facilidade de utilização e de navegação dos filtros, atribuindo-lhe uma classificação unânime de 5 pontos.

Os KPIs apresentados foram considerados adequados para monitorizar a mobilidade urbana, no entanto houve uma sugestão para incluir mais detalhes em análises futuras. Quanto às melhorias sugeridas, a maioria dos utilizadores não propôs mudanças, com uma única sugestão que se foca na possibilidade de prever o movimento e concentração de pessoas.

Em suma, o *dashboard* atendeu as expectativas, sendo considerado uma ferramenta eficaz, útil e fácil de utilizar.

5. Conclusões

Existe uma grande diversidade de literatura sobre *dashboards*, mas relativamente à mobilidade urbana com utilização de dados de telemóveis é algo que ainda é pouco explorado cientificamente. Deste modo, esta investigação contribuiu para melhorar outros instrumentos de monitorização para este setor ou atividades semelhantes e reforçou a literatura.

Para alcançar o objetivo principal, foi construído um *dashboard* para cada um dos objetivos específicos delineados no início da investigação, utilizando a ferramenta *Power BI*. Cada *dashboard* foi desenvolvido tendo em conta os critérios apontados pela literatura.

O primeiro *dashboard* foi concebido para identificar as áreas mais movimentadas de Lisboa, permitindo monitorizar as zonas de maior concentração de pessoas nas áreas de diversão noturna. Desta forma, a visualização dos fluxos de movimento é facilitada, auxiliando os gestores urbanos a alocar recursos de maneira mais eficiente e a tomar decisões informadas sobre a gestão dessas zonas. Os gestores, através deste *dashboard*, conseguem também perceber os padrões específicos de densidade populacional em determinados dias da semana ou em feriados.

O segundo *dashboard* foi desenvolvido com o objetivo de analisar o comportamento das pessoas durante eventos festivos de grande escala, como as JMJ. Esta ferramenta permite comparar o comportamento da população em períodos de eventos com os padrões regulares de mobilidade, ou como a densidade populacional varia durante os vários meses do ano. Desta forma, oferece *insights* importantes sobre como grandes aglomerações impactam a mobilidade urbana.

Adicionalmente, foi criado um terceiro *dashboard* para avaliar as transferências de utilizadores entre diferentes zonas de diversão noturna e para onde se deslocam as pessoas ao longo da noite. Este *dashboard* permite analisar fluxos de pessoas entre várias freguesias como Santa Maria Maior ou Misericórdia, ajudando a CML a identificar padrões de movimento e a otimizar serviços urbanos, como transportes públicos e segurança.

Por fim, o quarto *dashboard* foi construído para correlacionar os movimentos de pessoas com os níveis de ruído captados por sensores ambientais. Deste modo, é possível ter uma visão abrangente da relação entre a concentração de pessoas em determinadas zonas e o impacto que isso tem nos níveis de ruído ao longo da noite, fornecendo uma ferramenta essencial para mitigar o impacto do ruído e melhorar a qualidade de vida dos residentes ou turistas.

Desde a compreensão inicial dos dados e do negócio, até à efetiva construção do modelo, a utilização da metodologia CRISP-DM no projeto foi vantajosa e permitiu uma abordagem estruturada e eficaz. Na fase de compreensão do negócio, foi definido o objetivo de negócio (identificar e estudar padrões e comportamentos da população, tanto portuguesa quanto estrangeira, nas zonas de diversão noturna) e o objetivo analítico (criar um *dashboard*). De seguida, passou-se para a compreensão dos dados, que permitiu ter uma visão mais aprofundada dos dados disponíveis, verificando a existência de erros e valores omissos. A seguinte fase, preparação dos dados, foi realizada em *Python* e em *Power BI*, renomeando colunas, eliminando alguns dados irrelevantes. Todas as etapas, desde a seleção dos dados à formatação dos mesmos, foram seguidas.

Com os dados já preparados e, com base na literatura e na entrevista com alguns especialistas, selecionaram-se as métricas importantes a incluir no *dashboard*. No passo seguinte, a modelação, procedeu-se então à construção do *dashboard*. Por fim, houve a necessidade de avaliar o *dashboard* quanto à utilidade e usabilidade. Foi possível perceber que, para diferentes utilizadores, esta ferramenta auxilia na monitorização da mobilidade urbana nos espaços noturnos.

Com os quatro objetivos específicos respondidos, o presente estudo conseguiu responder de forma eficaz à questão de investigação formulada - Os dados anonimizados dos telemóveis fornecidos por operadoras móveis permitem identificar padrões de mobilidade?

Baseado no trabalho e investigação realizada para esta dissertação, estou a realizar um artigo científico com o título “*Developing a Data-Driven Decision Support System for Managing Urban Nightlife: Insights into Mobility Patterns and Environmental Noise in Lisbon*” para a conferência EAI INTSYS 2024 – 8th EAI International Conference on Intelligent Transport Systems que irá ter lugar em Pisa, em Itália, nos dias 5 e 6 de dezembro de 2024.

Este trabalho oferece uma solução útil para a Câmara Municipal de Lisboa, permitindo monitorizar as zonas de diversão noturna de forma precisa e simples, respondendo a questões relacionadas com o comportamento da população e o impacto no ambiente sonoro. A abordagem adotada pode ser facilmente replicada noutras cidades que tenham acesso a dados de dispositivos móveis, o que torna esta solução aplicável em contextos urbanos semelhantes, para além da cidade de Lisboa. O estudo não só oferece uma solução prática para a gestão de espaços de lazer noturno, como também contribui para o conhecimento científico sobre a dinâmica lisboeta durante a noite. A replicabilidade desta solução em outros contextos urbanos demonstra o seu potencial e a relevância deste tipo de ferramentas para a gestão eficiente das cidades.

5.1. Contributos

Ao responder com sucesso à questão de investigação formulada, este estudo apresenta contributos a dois níveis: ao nível do conhecimento científico e ao nível profissional.

Contributos ao nível do conhecimento científico:

- O estudo proporciona um enriquecimento da literatura sobre mobilidade urbana e a utilização de dados de dispositivos móveis anonimizados como fonte de informação para o comportamento humano em espaços públicos, como espaços de diversão noturna. Ao explorar a monitorização de zonas de diversão noturna e a relação entre mobilidade e o ruído capturado por sensores, esta investigação acrescenta novas perspetivas sobre a utilização de *dashboards* e análise de dados no contexto da gestão urbana;
- A investigação identifica e valida KPIs relevantes (como os níveis de ruído por freguesia ou a densidade populacional por tipo de dia), combinando a literatura existente com o conhecimento de especialistas da área;
- O estudo torna-se mais um caso de sucesso da utilização da metodologia CRISP-DM, reforçando o valor desta metodologia. Com recurso a esta metodologia, foi resolvido um problema real, desenvolvendo um sistema de apoio à decisão baseado em *dashboards* que representam, em tempo real, a situação recolhida através dos dados;
- A investigação evidencia as potencialidades da ferramenta *Power BI* na criação de *dashboards* com utilidade e usabilidade, utilizando o *Power Query* e a linguagem DAX para criar métricas e indicadores com o objetivo de gerar visualizações que são interativas e fáceis de utilizar por parte dos gestores urbanos e outros profissionais envolvidos no planeamento e monitorização da cidade.

Contributos ao nível profissional:

- Ao disponibilizar um *dashboard* personalizado que permite monitorizar a movimentação e comportamento das pessoas nas zonas de diversão noturna lisboeta, este estudo oferece uma ferramenta prática para a Câmara Municipal de Lisboa, mais concretamente à equipa LxDataLab, e outros profissionais envolvidos na gestão urbana, conseguirem gerir essas áreas de forma mais eficiente;
- A solução apresentada é facilmente replicável noutras cidades, desde que tenham acesso a dados deste tipo. Desta forma, o sistema de apoio à decisão desenvolvido pode servir de modelo para outras áreas urbanas que enfrentam desafios semelhantes de mobilidade, segurança e poluição sonora;

- A investigação apresenta uma abordagem que combina a análise de dados de mobilidade fornecidos por operadoras móveis com os dados do nível de ruído ambiental, proporcionando aos planeadores urbanos e gestores uma ferramenta completa que permite gerir melhor as áreas de lazer noturno, respondendo às preocupações dos utilizadores dessas áreas, residentes e turistas;
- Este estudo contribui ainda para a transformação digital das cidades, apoiando o desenvolvimento de cidades inteligentes, onde os dados são utilizados para melhorar a qualidade de vida dos cidadãos, otimizar recursos e promover um planeamento urbano mais sustentável.

Concluindo, esta dissertação não só proporciona um exemplo prático da aplicação de *dashboards* para a gestão urbana, como também evidencia o potencial de combinar dados de mobilidade e ruído para uma compreensão mais profunda das dinâmicas de uma cidade à noite. Além disso, a abordagem utilizada pode servir como um modelo para outros estudos e para um futuro desenvolvimento de novas soluções tecnológicas em áreas urbanas.

5.2. Limitações

Apesar dos contributos significativos desta investigação já mencionados, também existiram algumas limitações que devem ser consideradas.

Uma das principais limitações está associada à acessibilidade e qualidade dos dados. Apesar dos dados anonimizados dos dispositivos móveis oferecerem informações úteis sobre a mobilidade nas zonas de diversão noturna, a sua precisão pode ser afetada pela conectividade da rede ou pelo número de dispositivos registados numa determinada região. Além disso, nem todos os padrões de mobilidade podem ser captados pela recolha de dados de dispositivos móveis, especialmente os das pessoas que não utilizam telemóveis ou que estão temporariamente desconectados da *internet*. Uma análise comportamental mais aprofundada é também dificultada pela falta de informações demográficas, como a idade e o género.

A disponibilidade dos sensores de ruído em todas as áreas em estudo foi outra limitação. A análise dos níveis de ruído foi baseada nos dados fornecidos pelos sensores já instalados nas freguesias selecionadas (cerca de 60), no entanto, estes dados podem não ter incluído todos os locais com maior densidade populacional, levando a possíveis lacunas na relação entre os níveis de ruído e a mobilidade. Além disso, o foco geográfico deste estudo foi limitado às freguesias com maior atividade noturna de Lisboa, o que pode limitar a generalização dos resultados a outras áreas urbanas com características distintas. A disponibilidade de dados semelhantes e o estado da infraestrutura local de monitorização urbana determinarão se este conceito pode ser replicado noutras cidades.

Outra limitação significativa foi o peso dos dados antes do procedimento de limpeza e preparação. Os dados fornecidos pela CML eram muito volumosos, o que dificultou o processo de transferência e manipulação inicial dos mesmos. O tempo de processamento foi afetado pelo tamanho dos ficheiros, especialmente no início, e foi necessário algum esforço para otimizar os dados e garantir que eram adequados para análise sem comprometer a sua integridade.

Por fim, o modelo de *dashboard* desenvolvido foi validado com base nas necessidades dos especialistas da Câmara Municipal de Lisboa, o que significa que os indicadores e métricas selecionados refletem as prioridades dessa cidade. Em outros contextos urbanos, os requisitos podem ser diferentes, exigindo ajustes no modelo proposto.

Apesar destes inconvenientes, os resultados são fiáveis e acrescentam bastante valor ao conhecimento científico e profissional. Abrem também oportunidades para novas investigações que podem aprofundar a análise destes problemas e alargar o estudo a outros contextos urbanos.

5.3. Recomendações e trabalho futuro

Com base nos resultados desta investigação, é possível propor algumas recomendações para melhorias futuras.

Uma das sugestões mais importantes é fornecer dados em tempo real às partes envolvidas na gestão das zonas de diversão noturna. Para as forças de segurança pública, serviços de emergência, transportes e outros serviços críticos, seria muito útil poder acompanhar em tempo real a densidade populacional nos locais mais movimentados. O acesso em tempo real a esta informação poderia evitar uma série de cenários perigosos ou problemáticos, como confrontos entre pessoas, crises médicas ou trânsito automóvel, melhorando os tempos de reação dos serviços e garantindo uma gestão mais eficaz das áreas metropolitanas. Seria possível tomar medidas imediatas, reduzindo os riscos e aumentando a segurança de todos.

A utilização de dados para o planeamento a médio e longo prazo é outra recomendação, particularmente no que diz respeito às exigências da indústria do turismo e dos seus setores relacionados (restaurantes, alojamento e transportes, por exemplo). O cruzamento e a análise contínua dos dados de mobilidade poderiam fornecer informações importantes sobre o crescimento do turismo em Lisboa, facilitando um planeamento mais eficiente dos recursos urbanos e promovendo o desenvolvimento de infraestruturas sustentáveis. As tendências futuras poderiam ser previstas através da utilização de técnicas de *Machine Learning*, fornecendo uma base de dados sólida para a criação de soluções inteligentes que se adaptem às mudanças na mobilidade urbana e no turismo.

Outra recomendação é o cruzamento de dados de mobilidade de dispositivos móveis com informações de cartões de transporte público, como o Viva ou Navegante. Isso possibilitaria uma análise mais aprofundada das entradas e saídas dos transportes públicos em relação ao movimento nas áreas de diversão noturna. Dessa forma, seria possível otimizar a rede de transportes, ajustar horários e garantir um serviço mais eficaz e adequado à procura. Com uma maior capacidade de processamento, seria possível analisar todos os dados mensais num único conjunto de dados e gerar gráficos mais dinâmicos, incluindo a análise de picos diários, durante o dia, tarde ou noite, tendo em conta o conjunto completo de dados.

Em suma, as recomendações para investigação futura incluem a incorporação de tecnologias em tempo real, a utilização de dados para planeamento estratégico, a aplicação de *Machine Learning* para implementar modelos preditivos e o cruzamento de diversas fontes de dados para obter uma compreensão mais abrangente da mobilidade urbana. Ao permitir uma administração urbana mais eficaz, inteligente e sustentável, estas iniciativas melhorariam consideravelmente o bem-estar da população.

Referências Bibliográficas

- Assembleia Municipal de Lisboa. (2018). *Segurança e Qualidade de Vida Nocturna na Cidade de Lisboa*. Assembleia Municipal de Lisboa.
<https://www.am-lisboa.pt/documentos/1525373122P2oQP9ja3Hp49JQ8.pdf>
- Azevedo, A., & Santos, M. F. (2008). KDD, semma and CRISP-DM: A parallel overview. *IADIS European Conference on Data Mining 2008, Amsterdam, the Netherlands, July 24-26, 2008. Proceedings*, 182–185.
- Brida, J. G., & Zapata, S. (2010). Economic Impacts of Cruise Tourism: The Case of Costa Rica. *Anatolia*, 21(2), 322–338. <https://doi.org/10.1080/13032917.2010.9687106>
- Calabrese, F., Diao, M., Di Lorenzo, G., Ferreira, J., & Ratti, C. (2013). Understanding individual mobility patterns from urban sensing data: A mobile phone trace example. *Transportation Research Part C: Emerging Technologies*, 26, 301–313.
<https://doi.org/10.1016/j.trc.2012.09.009>
- Câmara Municipal de Lisboa. (s.d.). *Desafio 74 - Movimentação de pessoas nas zonas de diversão noturna - Portal Dados Abertos*. Dados.cm-Lisboa.pt. <https://dados.cm-lisboa.pt/dataset/desafio-74-lxdatalab>
- Carić, H. (2016). Challenges and prospects of valuation – cruise ship pollution case. *Journal of Cleaner Production*, 111, 487–498. <https://doi.org/10.1016/j.jclepro.2015.01.033>
- Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C., & Wirth, R. (2000). CRISP-DM 1.0: Step-by-step data mining guide. SPSS.
- Eckerson, W. W. (2006). *Performance dashboards: measuring, monitoring, and managing your business*. (2ª ed.). John Wiley & Sons.
- Eckerson, W. W. (2010). *Performance Dashboards* (2ª ed.). John Wiley & Sons.
- Estes, B. (2016). Geolocation—The Risk and Benefits of a Trending Technology. *ISACA Journal*, 5. <https://www.isaca.org/resources/isaca-journal/issues/2016/volume-5/geolocationthe-risk-and-benefits-of-a-trending-technology>
- Fanticelli, H., Rabenjamina, S., Viana, A. C., Stanica, R., Santos, L., & Ziviani, A. (2022). Data-driven Mobility Analysis and Modeling: Typical and Confined Life of a Metropolitan Population. *ACM Transactions on Spatial Algorithms and Systems*, 8(3), 1–33.
<https://doi.org/10.1145/3517222>
- Ferrovial. (s.d.). *Urban mobility: what it is, types and projects*. Ferrovial.
<https://www.ferrovial.com/en/resources/mobility/>

- Few, S. (2006). *Information dashboard design: The Effective Visual Communication of Data* (1ª ed.). O'Reilly.
- Firican, G. (2017). Best Practices for Powerful Dashboards. *Business Intelligence Journal*, 22(2), 33–39.
- Francisco, B. A. M. (2022). *Analysis of the tourists behavior in Lisbon using data from a mobile operator* [Dissertação de mestrado, Iscte - Instituto Universitário de Lisboa. <http://hdl.handle.net/10071/27361>
- Geoapify. (2023, April 11). *What is Geolocation: How It Works and Its Many Uses - Geoapify*. Geoapify. <https://www.geoapify.com/what-is-geolocation/>
- González, M. C., Hidalgo, C. A., & Barabási, A.-L. (2008). Understanding individual human mobility patterns. *Nature*, 453(7196), 779–782. <https://doi.org/10.1038/nature06958>
- Institute for Economics and Peace. (2024). *Global Peace Index 2024*. <https://www.portugalglobal.pt/media/dzjlpsuk/gpi-2024-web.pdf>
- Jiang, S., Fiore, G. A., Yang, Y., Ferreira, J., Frazzoli, E., & González, M. C. (2013). A review of urban computing for mobile phone traces. *Proceedings of the 2nd ACM SIGKDD International Workshop on Urban Computing - UrbComp '13*, 2, 1–9. <https://doi.org/10.1145/2505821.2505828>
- Kovács, Z., Smith, M., Teleubay, Z., & Kovalcsik, T. (2023). Measuring visitor flows using mobile positioning data in three Hungarian second-tier cities. *International Journal of Tourism Cities*, 9(3), 656–674. <https://doi.org/10.1108/ijtc-03-2023-0049>
- Kung, K. S., Greco, K., Sobolevsky, S., & Ratti, C. (2014). Exploring Universal Patterns in Human Home-Work Commuting from Mobile Phone Data. *PLoS ONE*, 9(6), e96180. <https://doi.org/10.1371/journal.pone.0096180>
- Lu, X., Wetter, E., Bharti, N., Tatem, A. J., & Bengtsson, L. (2013). Approaching the Limit of Predictability in Human Mobility. *Scientific Reports*, 3(1). <https://doi.org/10.1038/srep02923>
- Marques, R. L. (2023). *Implementação de um dashboard para monitorização e controlo das atividades de manutenção numa empresa prestadora de serviços* [Dissertação]. <https://repositorio.iscte-iul.pt/handle/10071/30871>
- Medeiros, F. A. (2022). *Dashboard para monitorizar o Programa de Apoio à Restauração e Hotelaria* [Dissertação]. <https://repositorio.iscte-iul.pt/handle/10071/29414>
- Nica, I. C., CRĂCIUNESCU, S. L., ALEXANDRU, D., & IONESCU, Ș.-A. (2021). Using of KPIs and Dashboard in the Analysis of Carrefour Company's Performance Management.

- Journal of Organizational Management Studies*, 1–23.
<https://doi.org/10.5171/2021.852077>
- Paananen, K., & Minoia, P. (2018). Cruisers in the City of Helsinki: staging the mobility of cruise passengers. *Tourism Geographies*, 21(5), 801–821.
<https://doi.org/10.1080/14616688.2018.1490341>
- Park, S., & Zhong, R. R. (2021). Pattern Recognition of Travel Mobility in a City Destination: Application of Network Motif Analytics. *Journal of Travel Research*, 61(5), 1201–1216.
<https://doi.org/10.1177/00472875211024739>
- Pauwels, K., Ambler, T., Clark, B. H., LaPointe, P., Reibstein, D., Skiera, B., Wierenga, B., & Wiesel, T. (2009). Dashboards as a Service. *Journal of Service Research*, 12(2), 175–189.
<https://doi.org/10.1177/1094670509344213>
- Prabhakar, B. (2015). A Big Data Dashboard for Urban Mobility. *Procedia Computer Science*, 62, 7–8. <https://doi.org/10.1016/j.procs.2015.08.394>
- PSE. (s.d.). *Metodologia CRISP-DM - PSE*. <https://pse.pt/metodologia-crisp-dm/>
- Santos, B. (2023). *Analysis of the mobility of people and tourists in nightlife areas in the city of Lisbon* [Dissertação]. <https://repositorio.iscte-iul.pt/handle/10071/29160>
- Sanz, A. (2018). *Proposta de um dashboard para monitorizar falhas de energia numa rede elétrica inteligente* [Dissertação]. <https://repositorio.iscte-iul.pt/handle/10071/17580>
- Sarikaya, A., Correll, M., Bartram, L., Tory, M., & Fisher, D. (2019). What Do We Talk About When We Talk About Dashboards? *IEEE Transactions on Visualization and Computer Graphics*, 25(1), 682–692. <https://doi.org/10.1109/tvcg.2018.2864903>
- Smith, V. S. (2013). Data Dashboard as Evaluation and Research Communication Tool. *New Directions for Evaluation*, 2013(140), 21–45. <https://doi.org/10.1002/ev.20072>
- Song, C., Qu, Z., Blumm, N., & Barabasi, A.-L.. (2010). Limits of Predictability in Human Mobility. *Science*, 327(5968), 1018–1021. <https://doi.org/10.1126/science.1177170>
- Sun, H., Chen, Y., Lai, J., Yang, W., & Liu, X. (2021). Identifying Tourists and Locals by K-Means Clustering Method from Mobile Phone Signaling Data. *Journal of Transportation Engineering*, 147(10). <https://doi.org/10.1061/jtepbs.0000580>
- Tokola, H., Gröger, C., Järvenpää, E., & Niemi, E. (2016). Designing Manufacturing Dashboards on the Basis of a Key Performance Indicator Survey. *Procedia CIRP*, 57, 619–624. <https://doi.org/10.1016/j.procir.2016.11.107t>
- Tripadvisor. (s.d.). *Vida noturna em Lisboa*. Tripadvisor. https://www.tripadvisor.pt/Attractions-g189158-Activities-c20-Lisbon_Lisbon_District_Central_Portugal.html

- Uy, M. (2021, August 27). *What Is Data Roaming?* Lifewire. <https://www.lifewire.com/what-is-roaming-2377414>
- Versichele, M., Neutens, T., Delafontaine, M., & Van de Weghe, N. (2012). The use of Bluetooth for analysing spatiotemporal dynamics of human movement at mass events: A case study of the Ghent Festivities. *Applied Geography*, 32(2), 208–220. <https://doi.org/10.1016/j.apgeog.2011.05.011>
- Wang, Y., Correia, G. H. de A., van Arem, B., & Timmermans, H. J. P. (Harry). (2018). Understanding travellers' preferences for different types of trip destination based on mobile internet usage data. *Transportation Research Part C: Emerging Technologies*, 90, 247–259. <https://doi.org/10.1016/j.trc.2018.03.009>
- Wirth, R., & Hipp, J. (2000). *CRISP-DM: Towards a Standard Process Model for Data Mining*. <http://www.cs.unibo.it/~montesi/CBD/Beatriz/10.1.1.198.5133.pdf>
- Yigitbasioglu, O. M., & Velcu, O. (2012). A review of dashboards in performance management: Implications for design and research. *International Journal of Accounting Information Systems*, 13(1), 41–59. <https://doi.org/10.1016/j.accinf.2011.08.002>
- Zhang, A., Wang, A., Chan, E. H. W., Shi, W., Zhou, X., & Liu, Z. (2021). A Review of Human Mobility Research Based on Big Data and Its Implication for Smart City Development. *ISPRS International Journal of Geo-Information*, 10(1), 13. <https://doi.org/10.3390/ijgi10010013>
- Zhuang, M., Concannon, D., & Manley, E. (2022). A Framework for Evaluating Dashboards in Healthcare. *IEEE Transactions on Visualization and Computer Graphics*, 28(4), 1715–1731. <https://doi.org/10.1109/tvcg.2022.3147154>

Anexos

Anexo A – Tipo de *dashboards*

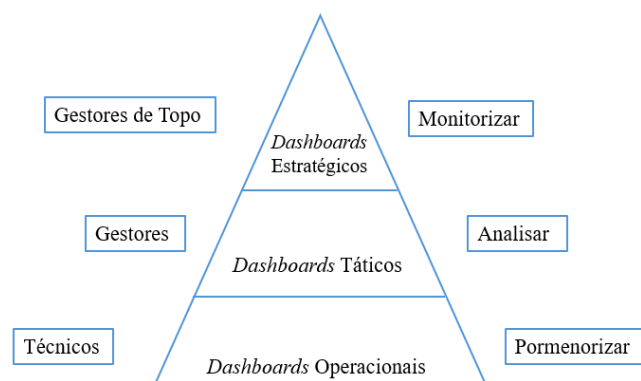


Figura 0.1: Tipos de *dashboard*

Anexo B – Critérios de avaliação dos *dashboards*

Tabela 0.1: Avaliação do *Dashboard* quanto ao *design*

Critérios
Simplificar o acesso dos utilizadores, centralizar a gestão e a administração dos dados, de forma a suportar milhares de utilizadores e evitar o download de grandes volumes de dados para computadores ligados a redes com baixa velocidade de ligação;
Definir objetivos, métricas, metas, limites, iniciativas e alertas; Associar objetivos, métricas, metas e iniciativas entre si;
Permitir que os utilizadores categorizem objetivos, métricas e iniciativas por diferentes perspetivas;
<i>Layouts</i> para fornecer várias formas de agrupar métricas relacionadas, scorecards e outros objetos no ecrã, como por exemplo, tabelas, colunas e <i>design</i> personalizado;
Permitir aos gestores de topo mapearem ligações entre métricas e estimarem e avaliarem o grau de correlação;
Personalizável, permitir aos utilizadores finais selecionarem métricas, alertas e outros objetos de listas autorizadas e organizarem o ecrã conforme as suas necessidades;
Facultar vários tipos de gráficos, símbolos e cores que permitam aos utilizadores avaliarem o estado do desempenho, as tendências e a variância de métricas críticas.

Fonte: Elaboração própria com base em Eckerson (2006)

Tabela 0.2: Avaliação do *Dashboard* quanto à análise

Crítérios
Organizar as informações em camada, ou seja, cada camada sucessiva fornece detalhes e perspectivas adicionais relativamente a métricas, processos ou eventos;
Associar cada gráfico, independentemente do seu tipo, uma tabela, em que o utilizador deve conseguir intercalar entre a apresentação da informação em gráfico e tabela;
Comparar os dados apresentados em gráficos e tabelas com as metas e os limites definidos;
Permitir aos utilizadores fazer <i>drill down</i> da informação sumariada até à informação de detalhe, através de um único clique;
<i>Drill across</i> permite aos utilizadores alternar as visualizações de uma métrica, alternando as dimensões (por exemplo, cliente, geografia), usando uma caixa de lista suspensa ou outro controlo gráfico;
<i>Drill through</i> é útil para obter detalhes de uma transação armazenada num sistema remoto, como, por exemplo, numa base de dados externa ou em relatórios <i>online</i> , criados noutras aplicações;
Relatórios interativos possibilitam aos utilizadores ordenar, filtrar, classificar, reagrupar ou formatar os dados, inserir e excluir colunas, modificar fórmulas;
Colocar pequenas instruções que ajudem os utilizadores menos experientes a analisar os dados e a navegar no dashboard, de forma a facilitar a chegada ao objetivo pretendido;
Permitir que os utilizadores definam e subscrevam novas visualizações de dados em tempo real;
Realizar análises com vários cenários distintos.

Fonte: Elaboração própria com base em Eckerson (2006)

Tabela 0.3: Avaliação do *Dashboard* quanto à entrega

Crítérios
Ter acesso a <i>scorecards</i> de diferentes níveis da organização;
Permitir que seja possível personalizar a forma de visualizar os dados na <i>internet</i> para utilização própria ou por terceiros;
Publicar e programar a informação em vários formatos (como <i>Excel</i> ou PDF) e em vários canais (<i>e-mail</i> , impressora e equipamentos via wireless);
Permitir que os utilizadores utilizem o <i>dashboard</i> sem estarem ligados à <i>internet</i> . Este deve ser portátil e o utilizador deve conseguir exportá-lo para <i>Excel</i> ou relatório;
Ser possível imprimir várias configurações do <i>dashboard</i> com quebras de página apropriadas e com títulos adequados;
Possibilitar que os utilizadores escrevam comentários nas métricas e que terceiros possam responder a esses comentários.

Fonte: Elaboração própria com base em Eckerson (2006)

Tabela 0.4: Avaliação do *Dashboard* quanto à administração

Critérios
Guardar definições e regras sobre métricas, dimensões, hierarquias, preferências e configurações do sistema para auditorias e possíveis pesquisas futuras;
Personalizar os ecrãs de acordo com as funções de cada colaborador, mostrando apenas as métricas, relatórios e dados que os utilizadores têm permissão;
Aumentar o nível de segurança nas bases de dados de forma a que os utilizadores com determinados perfis não tenham acesso a determinada informação;
Guardar cada alteração feita no sistema, por quem e quando, para fins de auditoria e controlo;
Impedir que utilizadores alterem dados e comentários inseridos manualmente, após uma determinada data, para evitar manipulações dos dados;
Ser possível configurar o <i>software</i> para funcionar com várias fontes de dados, desenhar modelos multidimensionais para análise, personalizar <i>layouts</i> , gerir a segurança e otimizar o desempenho do <i>software</i> ;
Resposta rápida de forma a que os programadores desenvolvam novas funcionalidades num espaço de dias ou semanas em vez de meses ou anos;
Permitir que os gestores de topo e líderes de equipas criem regras que acionem uma série de ações em resposta a uma condição excecional, como por exemplo, enviar diferentes tipos de alertas (<i>e-mail</i> , <i>sms</i> , ...) com base num problema identificado.

Fonte: Elaboração própria com base em Eckerson (2006)

Tabela 0.5: Avaliação do *Dashboard* quanto à infraestrutura

Critérios
Compatibilidade com o <i>hardware</i> , <i>software</i> , bases de dados e redes existentes;
Suportar interfaces, tecnologias e <i>frameworks</i> standards como, por exemplo, <i>web services</i> , XML, LDAP e arquitetura orientada a serviços;
Guardar os dados históricos do desempenho em <i>Data Marts</i> ou num <i>Data Warehouse</i> , guardar os dados atualizados num repositório de dados operacional e aceder aos dados em tempo real via <i>middleware</i> (EAI) ou através de <i>queries</i> de consulta à base de dados;
Integrar aplicações com aplicativos de terceiros, como portais, orçamentos, planeamentos, previsões, gestão de projetos e aplicações operacionais;
Ler todos os tipos de dados provenientes de qualquer sistema e guardá-los num servidor onde os dados possam ser limpos, transformados, reunidos e carregados num <i>dashboard</i>
Suportar vários tipos de visualização dos dados, normalmente entregues através de ferramentas OLAP, que armazenam os dados numa base de dados multidimensional;
Integrar o <i>dashboard</i> com os sistemas de segurança existentes na organização, como é o caso do LDAP;
Permitir que os programadores personalizem as funcionalidades do <i>dashboard</i> através de interfaces de programação de aplicações;
Tempos de resposta rápidos aos cliques do utilizador e ao pedido de dados, em segundos;
Independentemente do número de utilizadores ligados e do tipo de dados solicitados, o desempenho do <i>dashboard</i> não pode decrescer;
Garantir disponibilidade permanentemente mesmo quando novos dados estão a ser carregados ou atualizados.

Fonte: Elaboração própria com base em Eckerson (2006)

Anexo C – Fórmulas de cálculo (em linguagem DAX)

Tabela 0.6: Fórmula de cálculo (em linguagem DAX) para a criação da coluna "WeekendOrHoliday"

Fórmula de Cálculo
WeekendOrHoliday = SWITCH(TRUE(), '1semestre_limpo'[Weekday] IN {"Friday", "Saturday"}, "Fim de Semana", '1semestre_limpo'[Holiday] = 1 && NOT '1semestre_limpo'[Weekday] IN {"Friday", "Saturday"}, "Feriado", "Dia da Semana")

Tabela 0.7: Fórmula de cálculo (em linguagem DAX) para a criação da coluna "TipodeDia"

Fórmula de Cálculo
TipoDeDia = SWITCH(TRUE(), ano_limpo[WeekendOrHoliday] = "Fim de Semana", "Fim de Semana", ano_limpo[WeekendOrHoliday] = "Feriado" && TIME(HOUR(ano_limpo[time]), MINUTE(ano_limpo[time]), SECOND(ano_limpo[time])) >= TIME(0, 0, 0) && TIME(HOUR(ano_limpo[time]), MINUTE(ano_limpo[time]), SECOND(ano_limpo[time])) <= TIME(6, 0, 0), "Feriado", ano_limpo[WeekendOrHoliday] = "Dia da Semana" && LOOKUPVALUE(ano_limpo[WeekendOrHoliday], ano_limpo[date], ano_limpo[date] + 1) = "Feriado" && TIME(HOUR(ano_limpo[time]), MINUTE(ano_limpo[time]), SECOND(ano_limpo[time])) >= TIME(20, 0, 0) && TIME(HOUR(ano_limpo[time]), MINUTE(ano_limpo[time]), SECOND(ano_limpo[time])) <= TIME(23, 59, 59), "Véspera de Feriado", "Dia da Semana")

Tabela 0.8: Fórmula de cálculo (em linguagem DAX) para a criação da coluna "NTotal_Telemoveis"

Fórmula de Cálculo
NTotal_Telemoveis = ano_limpo[C3] + ano_limpo[C4]

Tabela 0.12: Fórmula de cálculo (em linguagem DAX) para a criação das métricas "Densidade_Media_Verao" e "Densidade_Media_Outras_Estacoes"

Fórmula de Cálculo
Densidade_Media_Verao = CALCULATE(AVERAGEX(FILTER(ano_limpo, MONTH(ano_limpo[Date]) IN {6, 7, 8}),ano_limpo[NTotal_Telemoveis]))
Densidade_Media_Outras_Estacoes = CALCULATE(AVERAGEX(FILTER(ano_limpo, NOT(MONTH(ano_limpo[Date]) IN {6, 7, 8})),ano_limpo[NTotal_Telemoveis]))

Tabela 0.11: Fórmula de cálculo (em linguagem DAX) para a criação da coluna "Tipo_Estacao"

Fórmula de Cálculo
Tipo_Estacao = SWITCH(TRUE(),MONTH(ano_limpo[Date]) IN {6, 7, 8}, "Verão","Outro")

Tabela 0.10: Fórmula de cálculo (em linguagem DAX) para a criação das métricas "Densidade_Media_DiaDaSemana", "Densidade_Media_FimDeSemana", "Densidade_Media_Feriado" e "Densidade_Media_Vespera_Feriado"

Fórmula de Cálculo
Densidade_Media_DiaDaSemana = CALCULATE(AVERAGEX(FILTER(ano_limpo, ano_limpo[TipoDeDia] = "Dia da Semana"), ano_limpo[NTotal_Telemoveis]))
Densidade_Media_FimDeSemana = CALCULATE(AVERAGEX(FILTER(ano_limpo, ano_limpo[TipoDeDia] = "Fim de Semana"), ano_limpo[NTotal_Telemoveis]))
Densidade_Media_Feriado = CALCULATE(AVERAGEX(FILTER(ano_limpo, ano_limpo[TipoDeDia] = "Feriado"), ano_limpo[NTotal_Telemoveis]))
Densidade_Media_Vespera_Feriado = CALCULATE(AVERAGEX(FILTER(ano_limpo, ano_limpo[TipoDeDia] = "Véspera de Feriado"), ano_limpo[NTotal_Telemoveis]))

Tabela 0.9: Fórmulas de cálculo (em linguagem DAX) para criação das tabelas "Telemoveis_Por_Freguesia", "Tabela_Ranking", "Tabela_Periodos" e "Tabela_Horas"

Fórmula de Cálculo
Telemoveis_Por_Freguesia = SUMMARIZE('ano_limpo', 'ano_limpo'[Freguesia], "TotalTelemoveis", SUM('ano_limpo'[NTotal_Telemoveis]))
Tabela_Ranking = SUMMARIZE(telemoveis_por_freguesia,telemoveis_por_freguesia[freguesia],"TotalTelemoveis", SUM(telemoveis_por_freguesia[TotalTelemoveis]))
Tabela_Periodos = DATATABLE("Periodo", STRING,"Indice", INTEGER, { {"Jantar", 1}, {"Noite", 2}, {"Party", 3}, {"After Party", 4} })
Tabela_Horas = ADDCOLUMNS(DISTINCT(SELECTCOLUMNS(ano_limpo, "HoraFormatada", ano_limpo[HoraFormatada])), "Indice_Hora", SWITCH(TRUE(),HOUR(TIMEVALUE([HoraFormatada])) >= 20, HOUR(TIMEVALUE([HoraFormatada])) - 20,HOUR(TIMEVALUE([HoraFormatada])) < 6, HOUR(TIMEVALUE([HoraFormatada])) + 4,HOUR(TIMEVALUE([HoraFormatada]))))

Anexo D – *Mouse Over* no gráfico

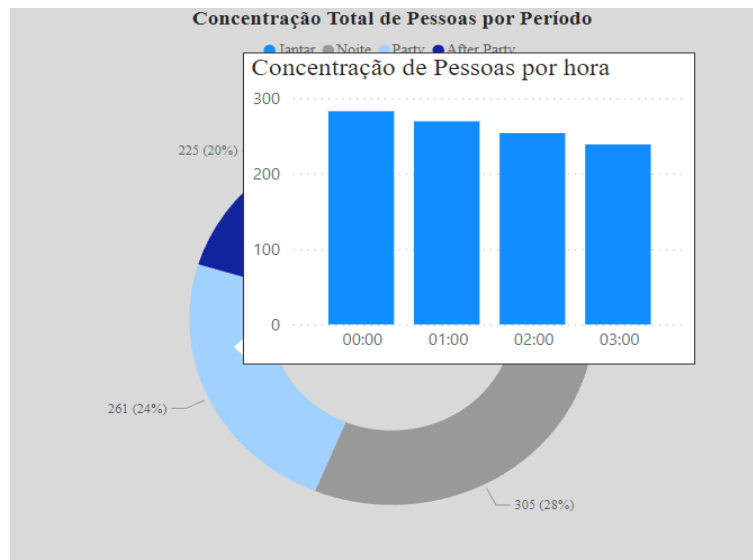


Figura 0.2: Gráfico com detalhe quando utilizador faz *mouse over*