

Data Science for Social Good Portugal



1

Projeto Revista CAIS





Índice

Introdução	3
Objetivo do Projeto	3
Análises	4
Sabia que 99% dos vendedores da CAIS têm alojamento?	5
Em detalhe para a cidade do Porto	7
Em detalhe para a cidade de Lisboa	8
.....	8
Sabia que 85% dos vendedores da CAIS têm mais de 40 anos de idade?	9
Em detalhe para a cidade do Porto	10
Em detalhe para a cidade de Lisboa	11
Sabia que o nível de escolaridade da maioria dos vendedores da CAIS se situa entre o primário e o básico?	12
Em detalhe para a cidade do Porto	13
Em detalhe para a cidade de Lisboa	14
Sabia que os vendedores da CAIS vivem com menos de 23% do custo de vida estimado das suas cidades?	15
Em detalhe para a cidade do Porto	17
Em detalhe para a cidade de Lisboa	18
Sabia que com a ajuda da CAIS esse valor aumenta para cerca de 45%?	19
Em detalhe para a cidade do Porto	20
Em detalhe para a cidade de Lisboa	21
Sabe onde é aplicado o valor obtido através da venda da Revista?	22
Em detalhe para a cidade do Porto	24
Em detalhe para a cidade de Lisboa	25
Sabe onde a Revista CAIS já foi vendida?	26
Fontes de dados	29
Limpeza dos dados	29
Análises incompletas	31
Dicionário de Dados	32
Os incríveis voluntários	35



Introdução

A Revista CAIS é a primeira criação da Associação CAIS e tem-se revelado uma importante ferramenta de intervenção social para a capacitação e integração de cidadãos em situação de pobreza e/ou exclusão social.

A venda da revista pretende ser uma resposta temporária que, para além de dar algum suporte financeiro, promove também a recuperação da autoestima do vendedor e a aquisição de uma série de competências sociais importantes no processo de transição para o regresso ao mercado de trabalho.

70% da receita das vendas revertem para os vendedores, que têm assim oportunidade de satisfazer algumas das suas necessidades básicas, como comer uma refeição, pagar um quarto, pagar contas, ou até apoiar algum membro da família em situação semelhante.

Objetivo do Projeto

Com a realização deste projeto pretende-se essencialmente atingir dois objetivos principais:

- Perceber e quantificar o impacto da Revista e comunicar esse impacto à sociedade, em jeito de campanha. Pretende-se não só contribuir para o aumento das vendas da Revista, mas acima de tudo promover a redução do estigma associado aos vendedores e impulsionar o envolvimento da comunidade com a missão da CAIS.
- Analisar padrões ao nível do perfil dos vendedores e/ou quantificar o seu sucesso na venda de Revistas, de forma a possibilitar a identificação atempada de casos mais problemáticos e poder apoiar a tomada de decisões estratégicas internas.



Análises

As análises foram divididas em 7 secções que achamos relevantes de acordo com os dados fornecidos. Em cada uma dessas secções é apresentada uma representação visual, uma tabela com os dados que originaram a representação visual e uma descrição textual dos resultados obtidos.

Na maior parte dessas secções incluímos também os resultados por cidade, de forma a que exista conteúdo disponível caso haja interesse em fazer-se alguma campanha de marketing ou sensibilização direcionado para uma região geográfica específica. Devido ao baixo número de vendedores por cidade, apenas foram apresentados os resultados para as cidades do Porto e Lisboa.

Sobre os dados utilizados para estas análise, é importante considerar que foram utilizados os dados disponíveis de 2016 a 2020. Isto traduz-se num total de 176 registos de vendedores.

Para a análise sobre os tipos de gastos foram apenas usados os dados mais recentes correspondentes aos primeiros quatro meses de 2020.

Note-se que em cada uma das análises um registo representa um vendedor que num determinado ano esteve ativamente a vender a Revista. Isto significa que, a título de exemplo, um vendedor que esteja ativo durante dois anos terá dois registos. O total de vendedores apresentado para vários anos corresponde assim à quantidade acumulada de registos.

Sabia que 99% dos vendedores da CAIS têm alojamento?

A figura 1 representa a distribuição dos vendedores da Revista CAIS por tipo de habitação, entre 2016 e 2020.



Figura 1 - Mapa em árvore com o valor acumulado de vendedores por situação habitacional (Âmbito geográfico: Nacional)

Podemos verificar que apenas uma pequena percentagem vive em habitação devoluta. A grande maioria dos vendedores necessita de pagar renda, o que significa uma despesa fixa que representa uma fatia importante do custo de vida total. Mais concretamente, 66% dos vendedores pagam renda (habitação alugada + quarto alugado) e apenas pouco mais de 0.5% vivem em situação de habitação devoluta.

Tabela 1 - Valor acumulado de vendedores por situação habitacional (Âmbito geográfico: Nacional)

	Quantidade de vendedores (Acumulado 5 anos)	% Sobre o total
Habitação Alugada	60	34.09%
Quarto Alugado	56	31.82%
Habitação Social	23	13.07%
Habitação Própria	9	5.11%
Informação não disponível	8	4.55%
Habitação Familiares	5	2.84%
Anexo	5	2.84%
Centro de Acolhimento	4	2.27%
Habitação Social (anexo)	3	1.70%
Instituição	1	0.57%
Habitação Devoluta	1	0.57%
Habitação Cedida	1	0.57%

Em detalhe para a cidade do Porto

Para a cidade do Porto a distribuição muda, mas a conclusão é idêntica: apenas 1 vendedor viveu em habitação devoluta, o que continua a representar uma percentagem muito baixa do total.

Percebemos também que na zona do Porto a percentagem de vendedores que alugaram casa ou quarto é acima do que vimos no cenário nacional, representando mais de 73% na zona do Porto.

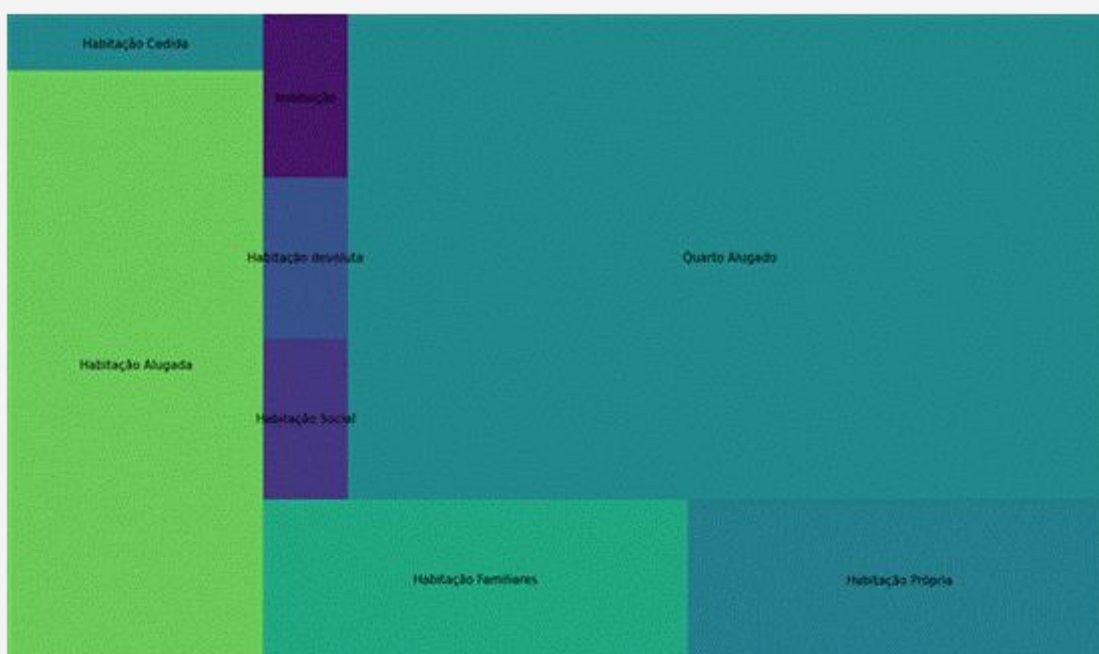


Figura 2 - Mapa em árvore com o valor acumulado de vendedores por situação habitacional (Âmbito geográfico: Porto)

Tabela 2 - Valor acumulado de vendedores por situação habitacional (Âmbito geográfico: Porto)

	Quantidade de vendedores (Acumulado 5 anos)	% Sobre o total
Quarto Alugado	27	51.92%
Habitação Alugada	11	21.15%
Habitação Própria	5	9.62%
Habitação Familiares	5	9.62%
Instituição	1	1.92%
Habitação devoluta	1	1.92%
Habitação Social	1	1.92%
Habitação Cedida	1	1.92%

Em detalhe para a cidade de Lisboa

Para a cidade de Lisboa não existe registo de vendedores em habitação devoluta. A situação mais comum continua aqui a ser o aluguer de habitação, seguido do aluguer de quarto.

A habitação social parece também ter mais relevância em Lisboa que no Porto, sendo importante referir que em Lisboa não existe o registo de casos de “Habitação Familiares” como na cidade do Porto.

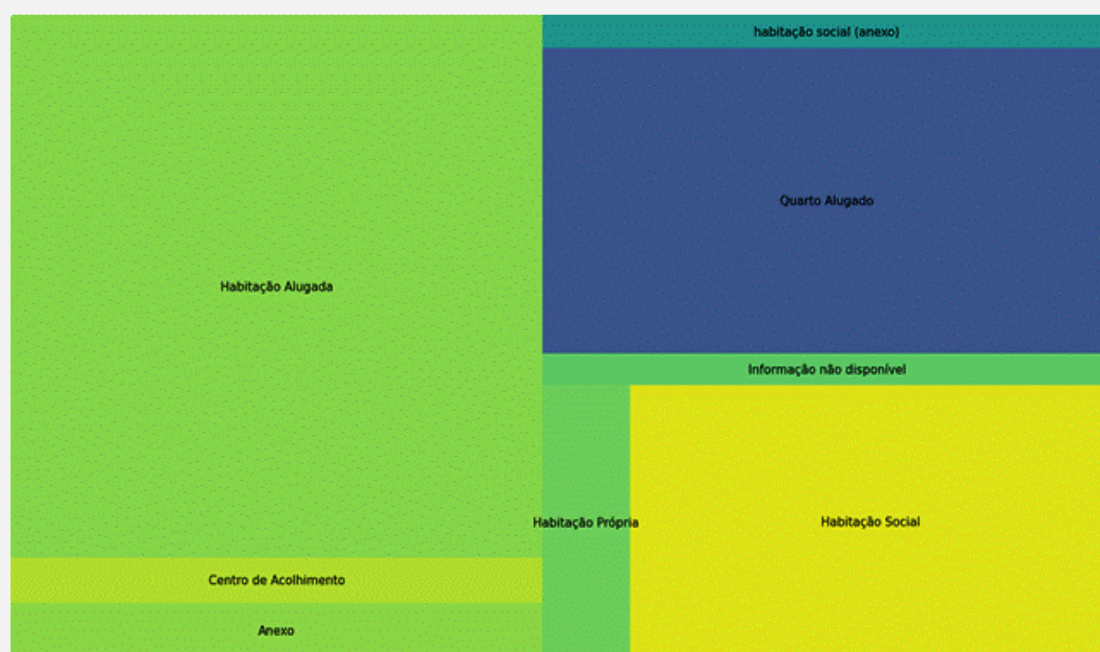


Figura 3 - Mapa em árvore com o valor acumulado de vendedores por situação habitacional (Âmbito geográfico: Lisboa)

Tabela 3 - Valor acumulado de vendedores por situação habitacional (Âmbito geográfico: Lisboa)

	Quantidade de vendedores (Acumulado 5 anos)	% Sobre o total
Habitação Alugada	48	40.68%
Quarto Alugado	29	24.58%
Habitação Social	22	18.64%
Anexo	5	4.24%
Habitação Própria	4	3.39%
Centro de Acolhimento	4	3.39%
Habitação social (anexo)	3	2.54%
Informação não disponível	3	2.54%

Sabia que 85% dos vendedores da CAIS têm mais de 40 anos de idade?

Para esta análise foram criadas classes de idades com um intervalo de 20 anos cada.

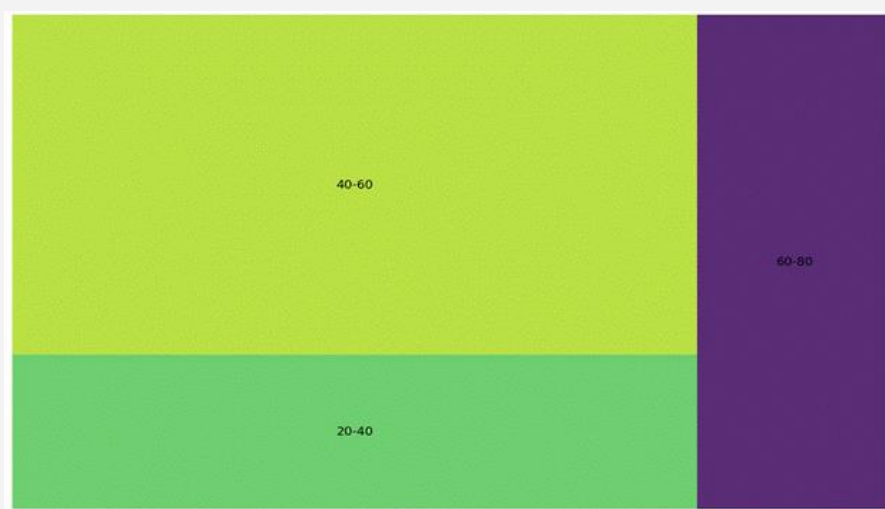


Figura 4 - Mapa em árvore com o valor acumulado de vendedores por intervalo de idades (Âmbito geográfico: Nacional)

Podemos constatar que a idade de mais de 56% dos vendedores se situa no intervalo entre os 40 e os 60 anos de idade. A idade dos restantes vendedores está distribuída pelo intervalo dos 60 aos 80 anos, com cerca de 30% dos vendedores com idade pertencente a este intervalo, e pelo intervalo dos 20 aos 40 anos, com 13% dos vendedores com idade pertencente a este intervalo.

Tendo em conta a última estatística feita pela PORDATA em 2018, que revela uma esperança média de vida em Portugal de 81 anos, e tendo em conta que mais de 85% dos vendedores têm mais de 40 anos, podemos concluir que mais de 85% dos vendedores da CAIS têm uma idade superior a metade da esperança média de vida.

Tabela 4 - Valor acumulado de vendedores por intervalo de idades (Âmbito geográfico: Nacional)

	Quantidade de vendedores (Acumulado 5 anos)	% Sobre o total
40-60	99	56.25%
60-80	53	30.11%
20-40	23	13.07%
Informação não disponível	1	0.57%

Em detalhe para a cidade do Porto

Na cidade do Porto o panorama é ligeiramente diferente. A grande maioria dos vendedores tem uma idade compreendida entre os 40 e os 60 anos de idade, intervalo que representa quase dois terços dos casos.

Além disso, a camada mais “jovem”, dos 20 aos 40 anos, deixa de ter representação.



Figura 5 - Mapa em árvore com o valor acumulado de vendedores por intervalo de idades (Âmbito geográfico: Porto)

Podemos confirmar também que, considerando apenas os vendedores da cidade do Porto cuja informação relativa à idade estava disponível, têm idade superior a 40 anos.

Tabela 5 - Valor acumulado de vendedores por intervalo de idades (Âmbito geográfico: Porto)

	Quantidade de vendedores (Acumulado 5 anos)	% Sobre o total
40-60	35	67.31%
60-80	16	30.77%
Informação não disponível	1	1.92%

Em detalhe para a cidade de Lisboa

No caso de Lisboa, embora o intervalo com maior registo de vendedores continue a ser o intervalo dos 20 aos 40 anos, ao contrário do que acontece na cidade do Porto, existem também registos de vendedores com idade entre os 20 e os 40 anos.

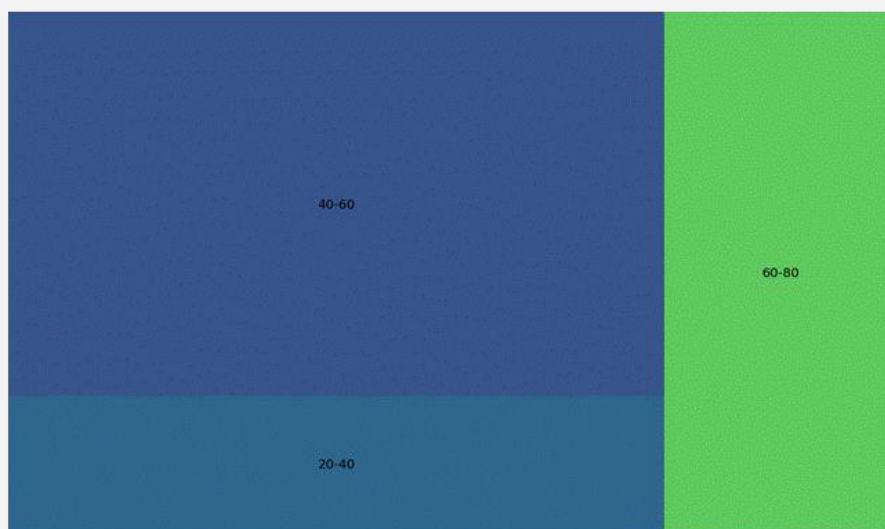


Figura 6 - Mapa em árvore com o valor acumulado de vendedores por intervalo de idades (Âmbito geográfico: Lisboa)

O intervalo mais jovem representa quase 20%, o que ainda é bastante significativo. Torna a média das idades em Lisboa de 51 anos de idade, enquanto esse valor é de 57 anos de idade para a cidade do Porto.

Tabela 6 - Valor acumulado de vendedores por intervalo de idades (Âmbito geográfico: Lisboa)

	Quantidade de vendedores (Acumulado 5 anos)	% Sobre o total
40-60	64	54.24%
60-80	31	26.27%
20-40	23	19.49%

Sabia que o nível de escolaridade da maioria dos vendedores da CAIS se situa entre o primário e o básico?

Apesar dos registos revelarem que a maior percentagem dos vendedores tem algum tipo de escolaridade, essa é apenas dividida em ensino básico e primário. O facto de o nível de escolaridade dos vendedores da CAIS ser baixo limita a oportunidade de estes conseguirem rendimentos estáveis.

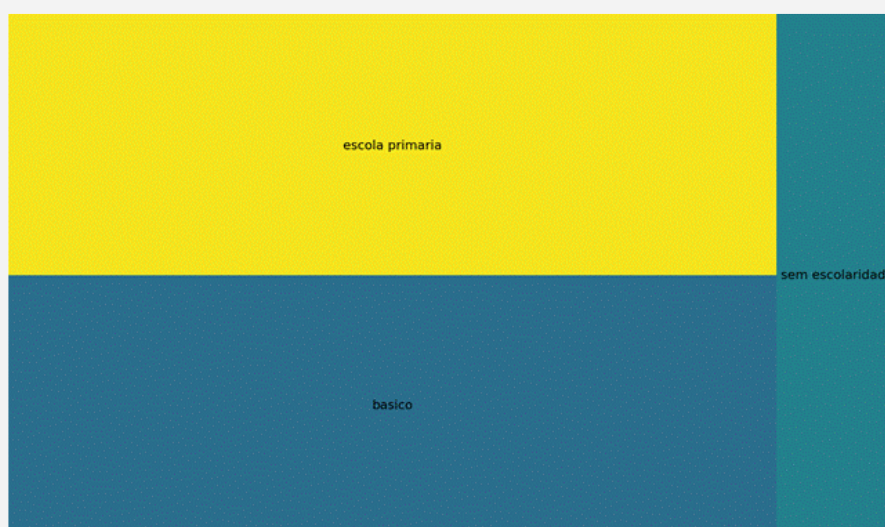


Figura 7 - Mapa em árvore com o valor acumulado de vendedores por grau de escolaridade (Âmbito geográfico: Nacional)

Tabela 7 - Valor acumulado de vendedores por grau de escolaridade (Âmbito geográfico: Nacional)

	Quantidade de vendedores (Acumulado 5 anos)	% Sobre o total
Escola Primária	76	43.18%
Ensino Básico	76	43.18%
Sem Escolaridade	24	13.64%

Em detalhe para a cidade do Porto

Na cidade do Porto todos os vendedores têm algum tipo de escolaridade, sendo que o nível de escolaridade se divide entre o ensino primário e o ensino básico.



Figura 8 - Mapa em árvore com o valor acumulado de vendedores por grau de escolaridade (Âmbito geográfico: Porto)

Podemos verificar que o Ensino Básico é o nível de escolaridade mais comum, com mais de 57% de vendedores na cidade do Porto com este nível de escolaridade.

Tabela 8 - Valor acumulado de vendedores por grau de escolaridade (Âmbito geográfico: Porto)

	Quantidade de vendedores (Acumulado 5 anos)	% Sobre o total
Ensino Básico	30	57.69%
Escola Primária	22	42.31%

Em detalhe para a cidade de Lisboa

No caso de Lisboa existe uma percentagem significativa de vendedores sem escolaridade, que corresponde a mais de 20% dos registos.

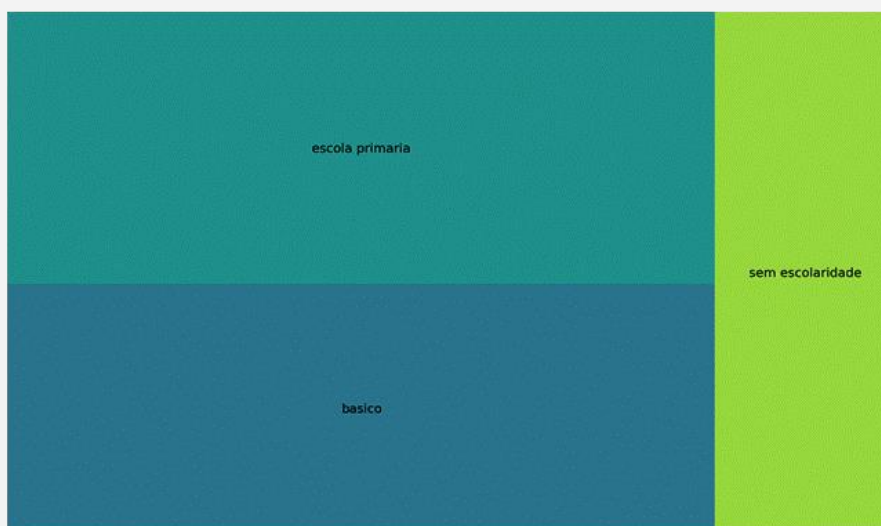


Figura 9 - Mapa em árvore com o valor acumulado de vendedores por grau de escolaridade (Âmbito geográfico: Lisboa)

Ao contrário do que acontece na cidade do Porto, em Lisboa a fatia maior é a de vendedores com o ensino primário. Isto revela que mais de 61% dos vendedores de Lisboa têm o nível de escolaridade primário, enquanto que na cidade do Porto a percentagem de vendedores com este nível de escolaridade é 42%.

Tabela 9 - Valor acumulado de vendedores por grau de escolaridade (Âmbito geográfico: Lisboa)

	Quantidade de vendedores (Acumulado 5 anos)	% Sobre o total
Escola Primária	49	41.53%
Ensino Básico	45	38.14%
Sem Escolaridade	24	20.34%

Sabia que os vendedores da CAIS vivem com menos de 23% do custo de vida estimado das suas cidades?

Para cada um dos vendedores foi calculado o custo de vida tendo em conta a sua situação habitacional, o número de dependentes e os custos de alimentação e valores de renda médios estimados para a sua cidade de habitação (fonte: NUMBEO - <https://www.numbeo.com/cost-of-living>).

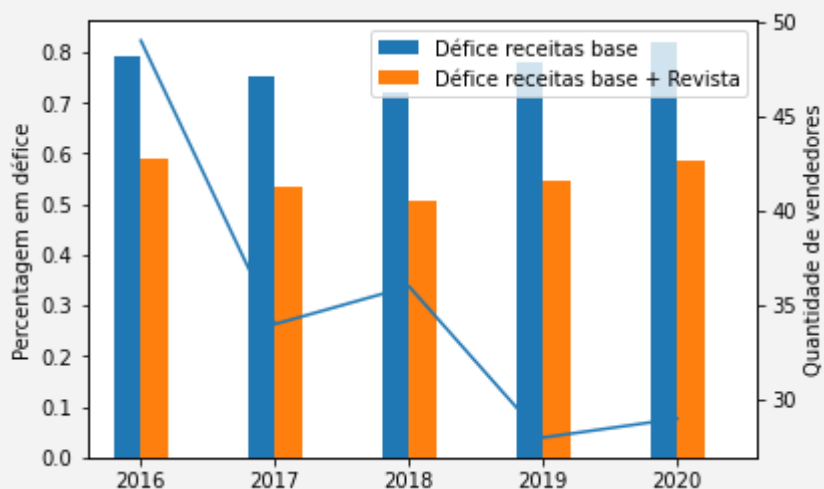


Figura 10 - Déficit das receitas dos vendedores, com e sem a venda da revista, para um custo de vida estimado (Âmbito geográfico: Nacional)

A linha da Imagem 10 representa a quantidade de vendedores, enquanto que as barras representam o déficit em relação ao custo de vida estimado para cada vendedor, com e sem a venda da Revista CAIS.

Com base na figura 10, podemos ver que o rendimento da venda da Revista não é a fonte principal de rendimentos para os vendedores, mas é antes um suporte adicional, para que estes possam ter uma vida mais digna. O valor dos rendimentos base foi estimado com base na informação oficial apresentada relativa a este tipo de ajudas por parte do Estado.

O que a figura nos mostra é a diferença percentual entre os rendimentos do vendedor em relação ao custo de vida estimado, que designamos por percentagem em déficit. Por exemplo, em 2016, sem a venda da Revista (barras azuis), os vendedores tinham rendimentos que correspondiam apenas a cerca de 20% do necessário para suportar o custo de vida estimado da sua cidade tendo em conta os fatores mencionados acima (situação habitacional, número de dependentes, etc.). Isto quer dizer que em 2016, em média os vendedores tinham um déficit de 80% em relação ao valor estimado do custo de vida. Com a venda da Revista esse déficit reduz para cerca de 60%.



Tabela 10 - Défice das receitas dos vendedores, com e sem a venda da revista, para um custo de vida estimado (Âmbito geográfico: Nacional)

	Défice sem a venda da Revista	Défice com a venda da Revista	Quantidade de Vendedores
2016	79.06%	58.84%	49
2017	75.09%	53.54%	34
2018	72.10%	50.61%	36
2019	77.87%	54.44%	28
2020	82.10%	58.64%	29

É interessante constatar que o défice base sem a venda da Revista atingiu um mínimo em 2018 e desde esse ano tem aumentado. Isto significa que o custo de vida está a aumentar mais do que os apoios/subsídios dados pelo Estado. Infelizmente, a venda da Revista não foi o suficiente para compensar este aumento do custo de vida, pois mesmo com a venda da Revista o défice tem aumentado a partir de 2018.

Em detalhe para a cidade do Porto

Na cidade do Porto, o défice entre o rendimento base dos vendedores e o custo de vida estimado é ligeiramente superior aos valores analisados para o panorama nacional. Contudo, neste caso o défice com e sem a venda da Revista tem-se mostrado mais estável. O número de vendedores também se tem mantido relativamente constante ao longo dos anos com um aumento ligeiro desde 2017, o que significa também um aumento da ajuda da Revista CAIS na cidade

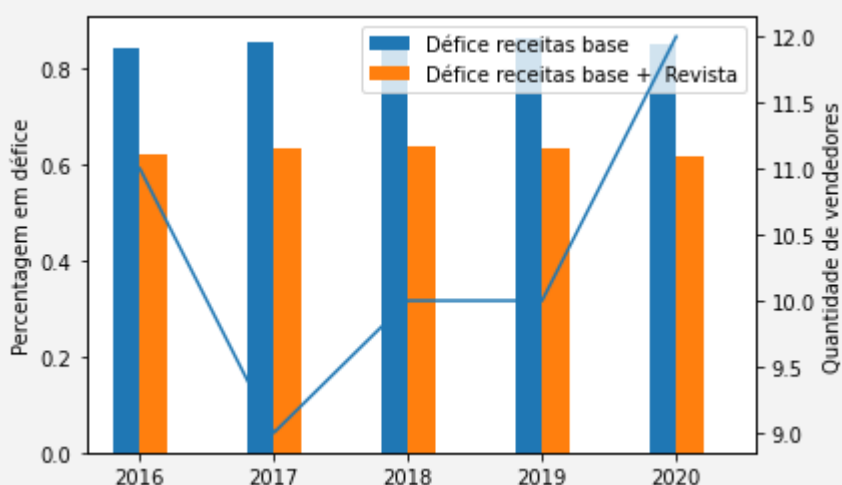


Figura 11 - Défice das receitas dos vendedores, com e sem a venda da revista, para um custo de vida estimado (Âmbito geográfico: Porto)

Tabela 11 - Défice das receitas dos vendedores, com e sem a venda da revista, para um custo de vida estimado (Âmbito geográfico: Porto)

	Défice sem a venda da Revista	Défice com a venda da Revista	Quantidade de Vendedores
2016	84.34%	62.30%	11
2017	85.44%	63.18%	9
2018	85.29%	63.73%	10
2019	86.44%	63.31%	10
2020	85.26%	61.60%	12

Em detalhe para a cidade de Lisboa

Em Lisboa, vemos um aumento significativo do défice ao longo dos últimos 2 anos, que contribui fortemente para o aumento mencionado anteriormente na análise do panorama nacional.

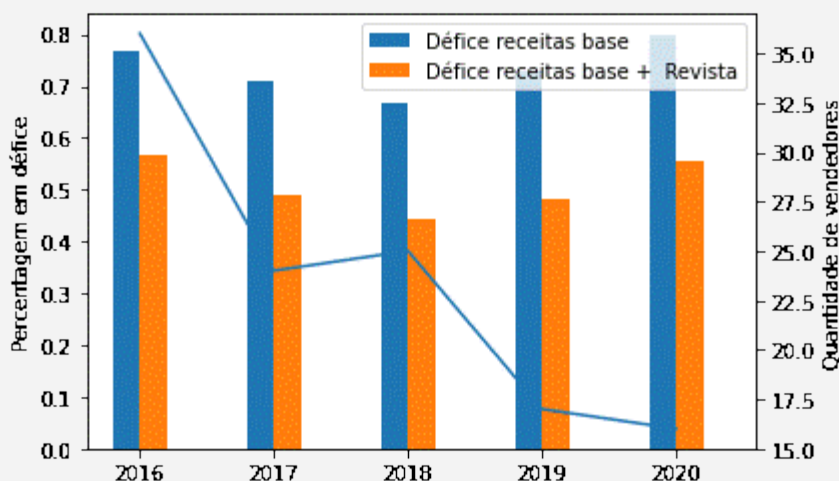


Figura 12 - Défice das receitas dos vendedores, com e sem a venda da revista, para um custo de vida estimado (Âmbito geográfico: Lisboa)

Com o suporte da Revista o aumento do défice não foi tão grande quanto seria sem o rendimento proveniente da venda da Revista. De facto, sem a revista temos um aumento do défice entre 2018 e 2020 de 13 pontos percentuais, sendo que com a revista o aumento é de 11 pontos percentuais. Ainda assim, a tendência tem sido o aumento do défice em relação ao custo de vida estimado, o que significa que na cidade de Lisboa em particular o custo de vida tem aumentado bastante nos últimos anos. Também nesta cidade a quantidade de vendedores diminuiu significativamente.

Tabela 12 - Défice das receitas dos vendedores, com e sem a venda da revista, para um custo de vida estimado (Âmbito geográfico: Lisboa)

	Défice sem a venda da Revista	Défice com a venda da Revista	Quantidade de Vendedores
2016	76.89%	56.59%	36
2017	71.08%	49.11%	24
2018	66.59%	44.46%	25
2019	72.82%	48.14%	17
2020	79.98%	55.53%	16

Sabia que com a ajuda da CAIS esse valor aumenta para cerca de 45%?

A venda da revista aumenta consideravelmente o rendimento do vendedor, mas ainda mantém os vendedores longe do custo de vida estimado.

O gráfico à direita quantifica quanto é que o rendimento extra que os vendedores receberam com a venda da Revista permite diminuir o défice em relação ao valor do custo de vida estimado. Podemos assim avaliar qual a percentagem de ajuda que a venda da Revista CAIS proporciona. Podemos ver que as vendas da Revista retiram uma média de aproximadamente 22% à diferença entre o vencimento do vendedor e o custo de vida aproximado.

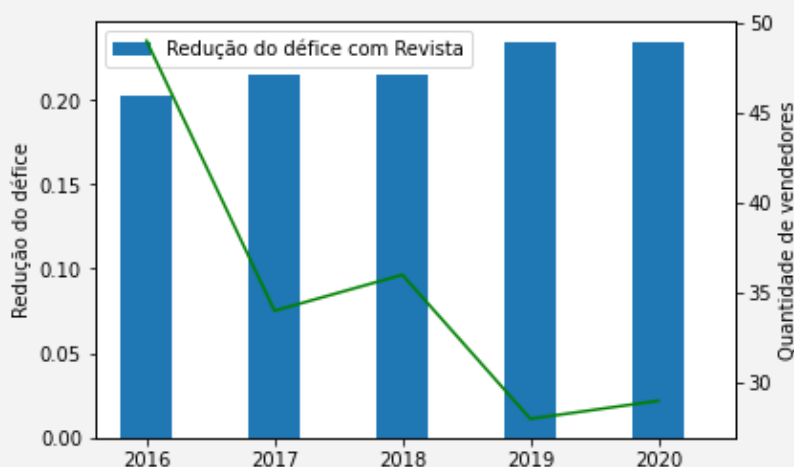


Figura 13 - Redução do défice para um custo de vida estimado com a ajuda da venda da Revista (Âmbito geográfico: Nacional)

O eixo à direita corresponde ao número de vendedores considerados para análise, representado pela linha verde.

Percebemos assim que a venda da Revista tem sido cada vez mais importante para ajudar a melhorar as condições de vida dos vendedores, o que se pode verificar pelo aumento da redução do défice entre 2016 e 2020.

Tabela 13 - Redução do défice para um custo de vida estimado com a ajuda da venda da Revista (Âmbito geográfico: Nacional)

	Redução do défice	Número de vendedores
2016	20.22%	49
2017	21.55%	34
2018	21.50%	36
2019	23.42%	28
2020	23.45%	29

Em detalhe para a cidade do Porto

Para a cidade do Porto, a média de redução do défice com a venda da Revista está um pouco acima da média nacional, principalmente nos primeiros anos em análise.

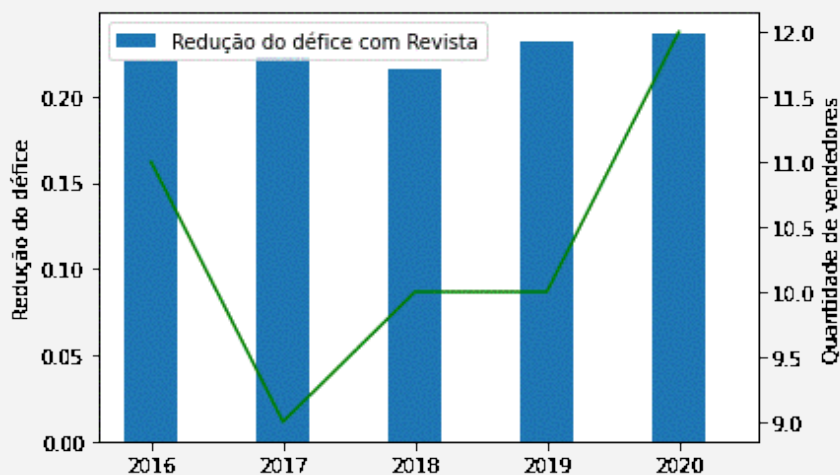


Figura 14 - Redução do défice para um custo de vida estimado com a ajuda da venda da Revista (Âmbito geográfico: Porto)

Podemos também verificar uma vez mais que o número de vendedores não tem seguido a quebra que vemos no panorama nacional.

Tabela 14 - Redução do défice para um custo de vida estimado com a ajuda da venda da Revista (Âmbito geográfico: Porto)

	Redução do défice de custo de vida	Número de vendedores
2016	22.04%	11
2017	22.26%	9
2018	21.56%	10
2019	23.13%	10
2020	23.66%	12

Em detalhe para a cidade de Lisboa

No caso de Lisboa vemos claramente que a redução do défice proveniente das receitas da Revista tem aumentado nos últimos anos.

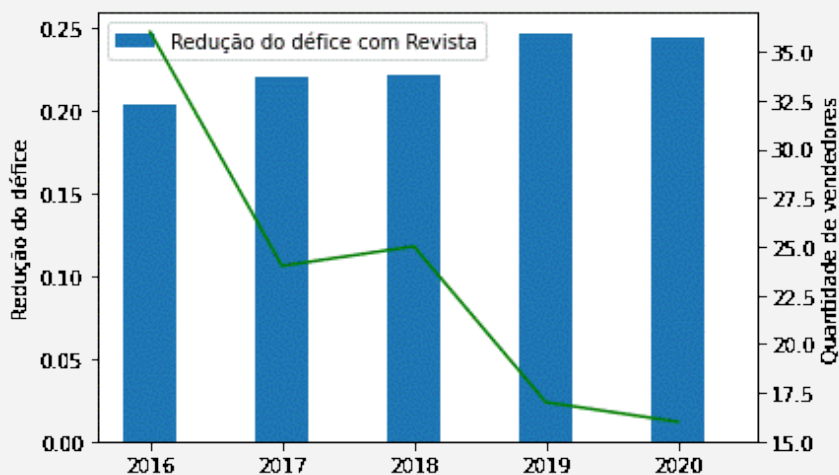


Figura 15 - Redução do défice para um custo de vida estimado com a ajuda da venda da Revista (Âmbito geográfico: Lisboa)

Podemos verificar também que o número de vendedores na cidade diminuiu significativamente, o que pode ser uma das causas do aumento da redução do défice, dado que um vendedor consegue vender mais Revistas por ter menos vendedores na cidade.

Nos últimos dois anos, os rendimentos obtidos com venda da Revista representaram cerca de 25% do custo de vida na cidade de Lisboa, o que é sem dúvida significativo para o aumento da qualidade de vida dos vendedores.

Tabela 15 - Redução do défice para um custo de vida estimado com a ajuda da venda da Revista (Âmbito geográfico: Lisboa)

	Redução do défice de custo de vida	Número de vendedores
2016	20.31%	36
2017	21.97%	24
2018	22.13%	25
2019	24.68%	17
2020	24.45%	16

Sabe onde é aplicado o valor obtido através da venda da Revista?

Para esta análise apenas foi possível utilizar dados dos vendedores do ano de 2020.

O valor obtido com a venda da Revista é utilizado maioritariamente em despesas com alimentação e habitação, seguidas de despesas pessoais. Seguem-se então despesas com medicação, transporte e ainda uma percentagem usa esse valor para ajudar a família.



Figura 16 - Mapa em árvore com a distribuição de categorias mencionadas como destino das receitas da Revista CAIS (Âmbito geográfico: Nacional)

Destacam-se as despesas com alimentação, com mais de 72% dos vendedores a mencionar usar receitas da venda da Revista para este fim.



Tabela 16 Distribuição de categorias mencionadas como destino das receitas da Revista CAIS (Âmbito geográfico: Nacional)

	Quantidade de vendedores (2020)	% dos vendedores
Alimentação	21	72.41%
Despesas pessoais	11	37.93%
Habitação	11	37.93%
Contas	7	24.14%
Medicação	5	17.24%
Passe de transporte	3	10.34%
Apoio a familiares	3	10.34%



Em detalhe para a cidade do Porto

Na cidade do Porto, podemos verificar que a grande parte dos vendedores usa o dinheiro obtido para despesas com alimentação, despesas pessoais, habitação e medicação.

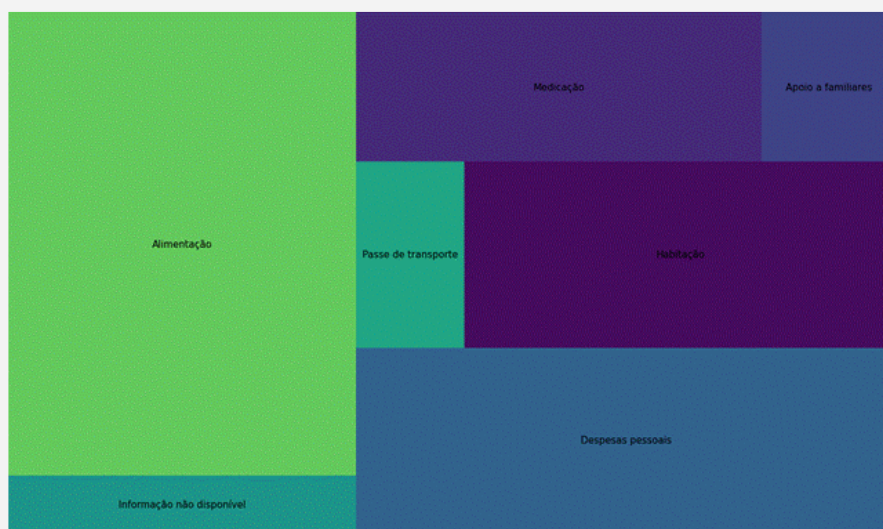


Figura 17 - Mapa em árvore com a distribuição de categorias mencionadas como destino das receitas da Revista CAIS (Âmbito geográfico: Porto)

Podemos verificar que cerca de um quarto dos vendedores menciona a medicação como destino das receitas.

Tabela 17 - Distribuição de categorias mencionadas como destino das receitas da Revista CAIS (Âmbito geográfico: Porto)

	Quantidade de vendedores (2020)	% dos vendedores
Alimentação	8	66.67%
Despesas pessoais	5	41.67%
Habitação	4	33.33%
Medicação	3	25.00%
Informação não disponível	1	8.33%
Passe de transporte	1	8.33%
Apoio a familiares	1	8.33%

Em detalhe para a cidade de Lisboa

No caso de Lisboa, para além das despesas com alimentação serem o destino mais comum das receitas, despesas com contas e habitação ganham mais relevância do que na cidade do Porto, o que pode ser reflexo do preço das rendas em Lisboa.



Figura 18 - Mapa em árvore com a distribuição de categorias mencionadas como destino das receitas da Revista CAIS (Âmbito geográfico: Lisboa)

A alimentação foi mencionada por 81% dos vendedores, o que é uma diferença relevante comparativamente aos valores nacionais.

Tabela 18 - Distribuição de categorias mencionadas como destino das receitas da Revista CAIS (Âmbito geográfico: Lisboa)

	Quantidade de vendedores (2020)	% dos vendedores
Alimentação	13	81.25%
Habitação	7	43.75%
Contas	7	43.75%
Despesas pessoais	5	31.25%
Medicação	2	12.50%
Passe de transporte	2	12.50%
Apoio a familiares	2	12.50%

Sabe onde a Revista CAIS já foi vendida?

Foi ainda possível analisar, através dos registos onde os vendedores se encontram na parte da manhã e da parte da tarde, as áreas onde a CAIS atua, isto é, os locais de venda dos vendedores da Revista CAIS. As zonas com maior intensidade são locais onde mais frequentemente os vendedores estiveram presentes.

Por exemplo para a cidade de Lisboa pode observar-se o seguinte:

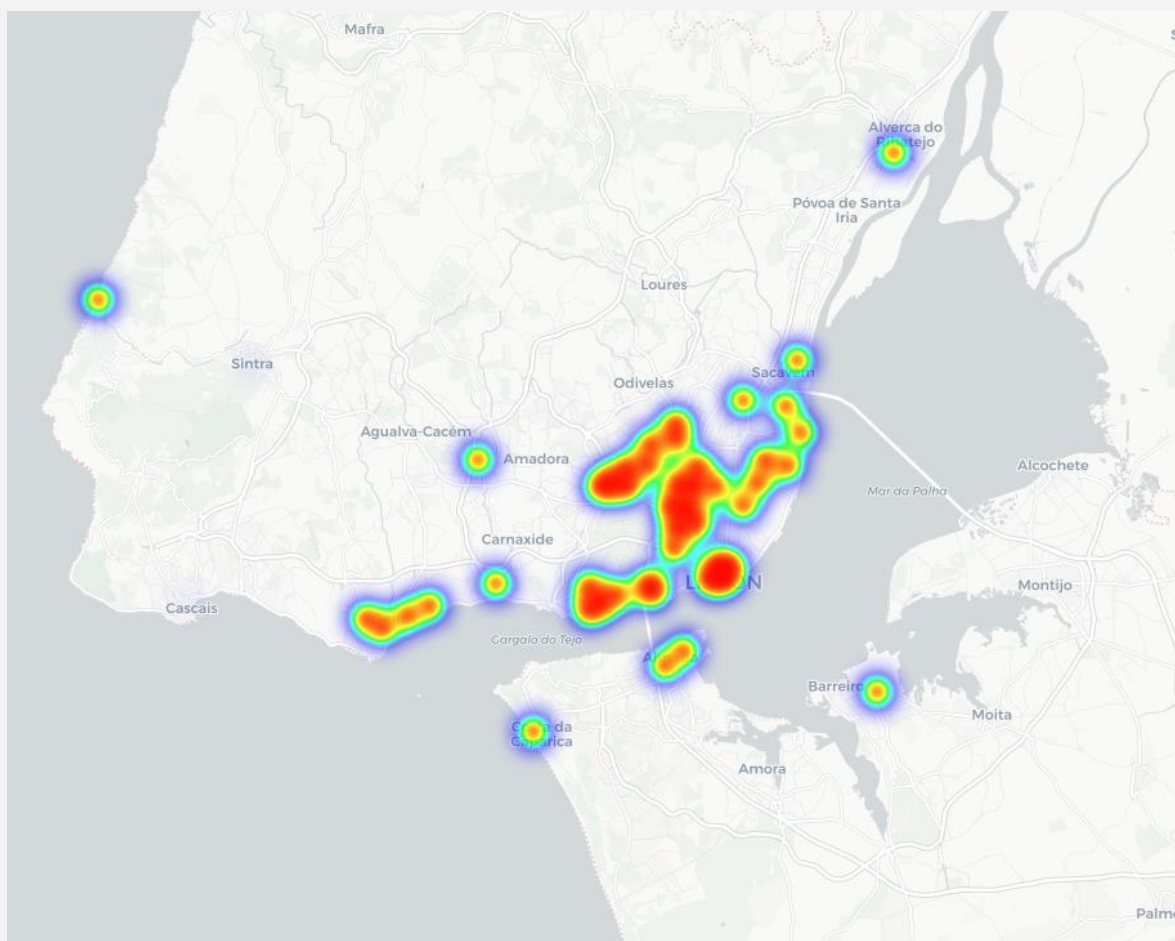


Figura 19 - Mapa de calor que representa a presença de vendedores da Revista CAIS em Lisboa

Consegue observar-se, a título de exemplo, que a estação de comboios de Santa Apolónia, a zona do Chiado e a zona do Parque Eduardo VII são as áreas com grande afluência de vendedores.

A mesma análise feita para a cidade do Porto leva à mesma conclusão: são as zonas mais populares que tendem a ter maior afluência de vendedores.

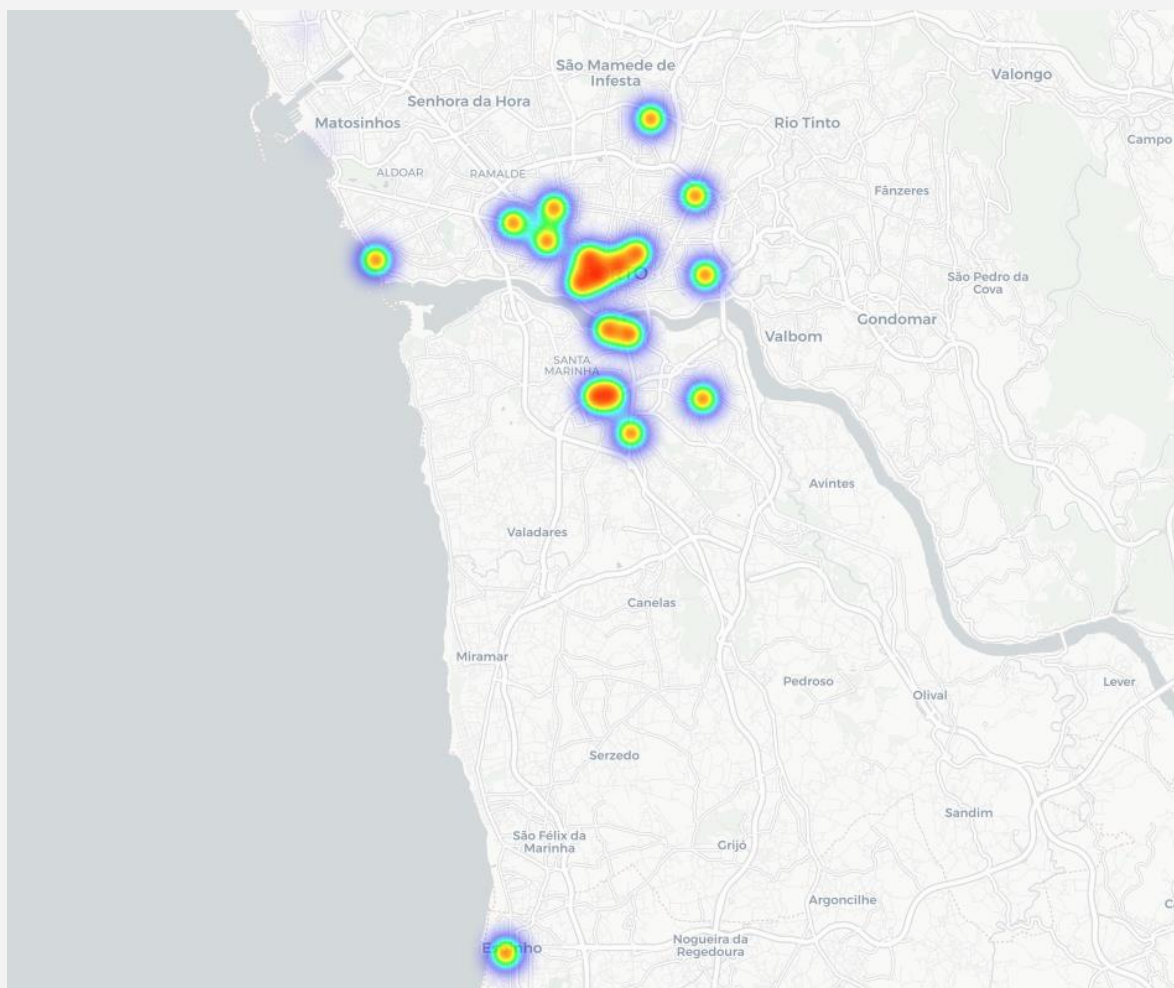


Figura 20 - Mapa de calor que representa a presença de vendedores da Revista CAIS no Porto

As zonas de maior afluência são, por ordem: Avenida Fernão Magalhães e Rua Júlio Dinis. Também é visível que a zona ribeirinha do lado de Gaia também é também um dos locais de venda prediletos.

Por fim, para Coimbra e com apenas 70 registos no total, a venda ocorreu apenas em dois locais.

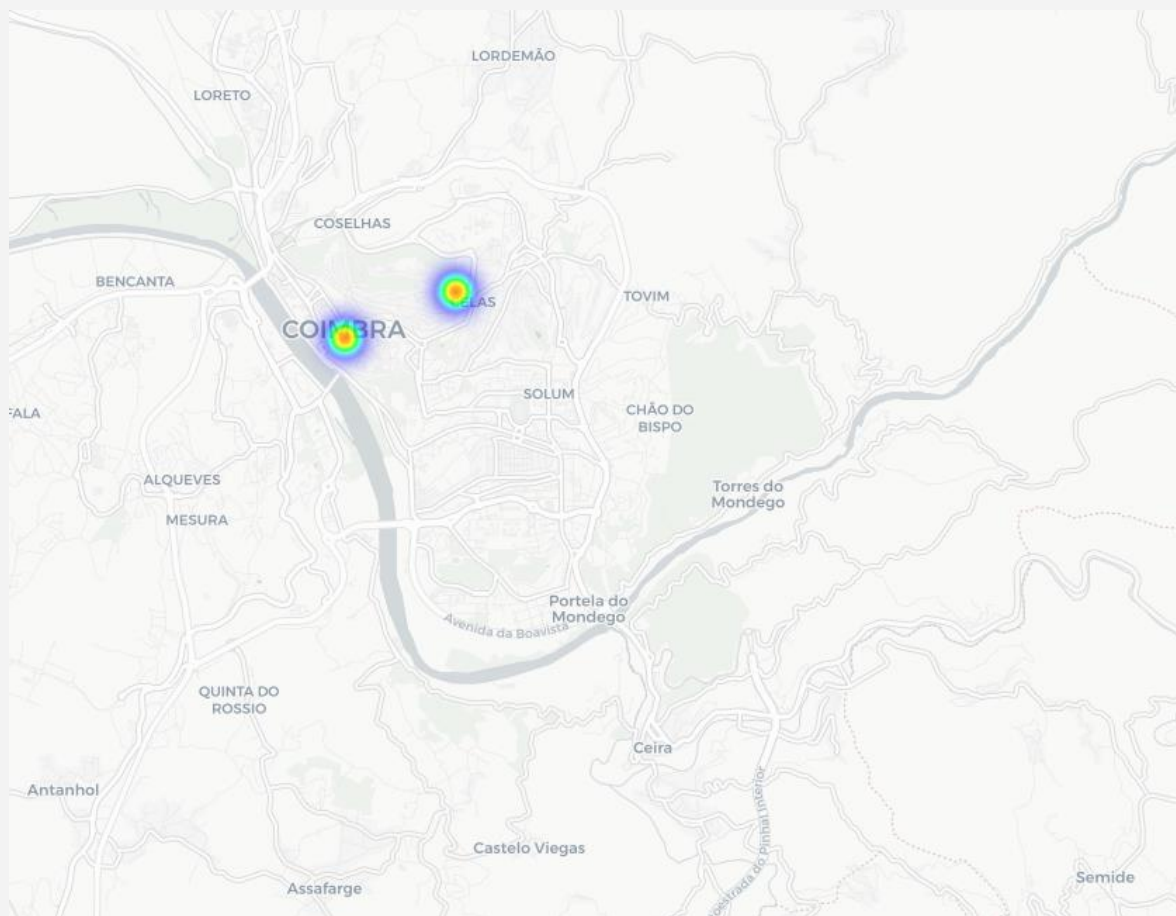


Figura 21 - Mapa de calor que representa a presença de vendedores da Revista CAIS em Coimbra

A Avenida Calouste Gulbenkian e a Rua Visconde da Luz são os dois locais onde há registo de ter havido vendas da Revista.

Devido à natureza dos dados ser extensa, disponibilizamos no nosso repositório as zonas com a respetiva contagem no seguinte link:

https://github.com/dssg-pt/project-cais/blob/master/data/geo_map_value_counts.csv

Fontes de dados

Existem notas a particularidades dos dados originais que foram usados nesta análise que têm que ser tomadas em consideração. Sabemos que é rara a fonte de dados que não tem erros ou imprecisões e no caso dos dados da Revista CAIS não foi diferente. Tentamos explicar nas secções seguintes as limitações, dificuldades e restrições encontradas que limitaram as análises feitas.

Limpeza dos dados

Inicialmente, os dados trabalhados vinham com bastantes problemas, tal como duplicação de colunas com informação diferente, inconsistência entre data de nascimento e idade, vendedores com idades impossíveis, etc.

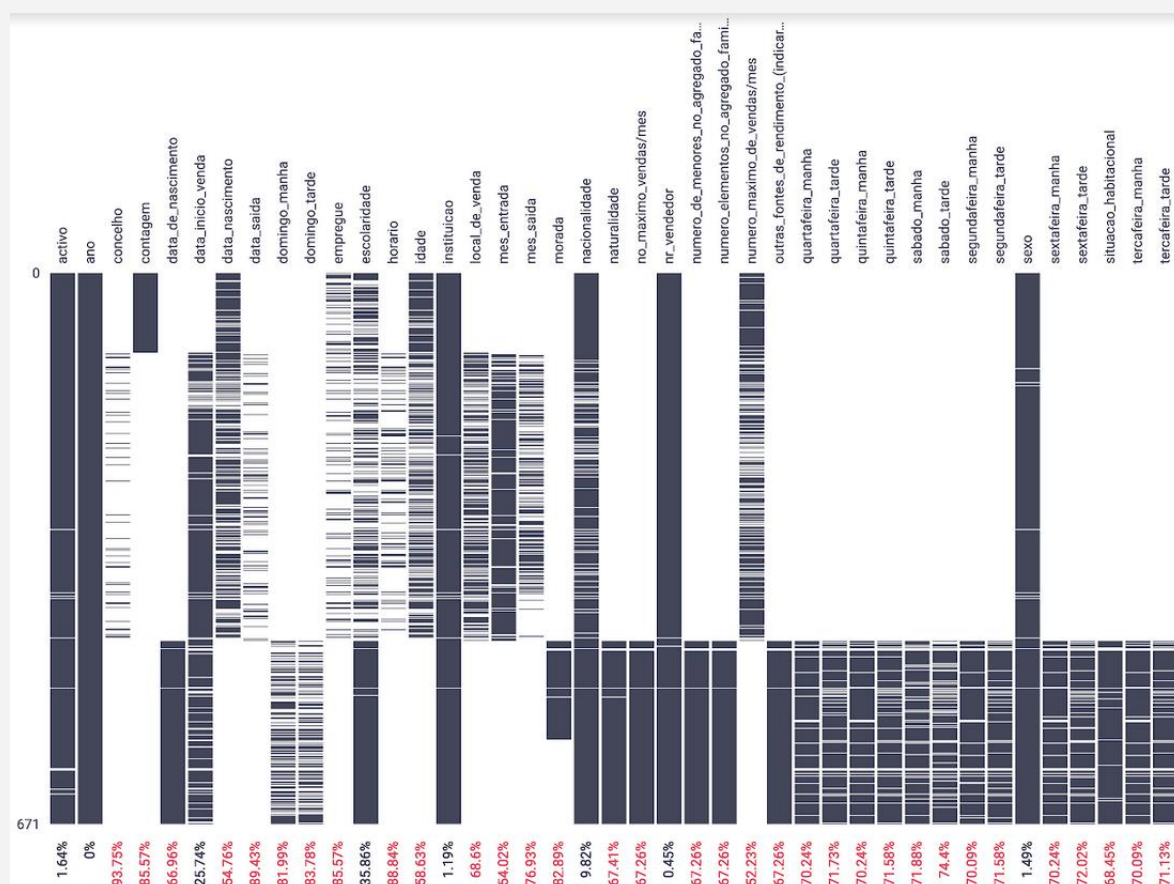


Figura 22 - Mapa de calor que representa o preenchimento do conjunto de dados original



No entanto, o maior problema foi a falta de informação em grande parte das colunas. A figura 22 representa a existência de dados ao longo do conjunto de dados disponibilizado. Cada linha é uma observação. Cada coluna é uma dimensão de dados. Cada linha nova está mais perto do ano atual. Áreas em branco são dados em falta. Zonas a cinzento escuro são áreas onde existem dados. Esta figura mostra de forma clara que os dados disponibilizados são muito incompletos até 2016. E só a partir de 2016 começa a haver um registo de dados mais significativo.

Mesmo assim, existem dados que ora estão presentes numa coluna ora estão noutra, pelo que dificulta o cruzamento de dados e/ou análise de correlações. Só os dados de 2016 começam a ter uma densidade mínima suficiente para tirar conclusões que pudessem ser mais relevantes.

No entanto com os dados originais foi possível criar colunas inferidas, como por exemplo:

- Idade: Obtida através da limpeza do campo idade em junção com data de nascimento.
- Ativo: Obtido através de texto livre
- Arquivado: Obtido através de texto livre
- Número de vendedor: uma aproximação inferida entre o número de colete, idade do vendedor, instituição, etc. Embora não seja muito preciso permite criar um identificador único de vendedor que permite identificar diferentes vendedores ao longo do tempo.
- Volume de Revistas vendidas por vendedor: através das Revistas destinadas a vender a uma percentagem do número de Revistas que costumam ser vendidas foi possível estimar um valor possivelmente mais real.
- Geolocalização: das vendas
- Utilizando o texto livre foi também possível criar categorias possíveis de agregar como por exemplo:
- Situação habitacional: Utilizado nas análises apresentadas.
- Razão de abandono: Não utilizado por falta de preenchimento
- Outras fontes de rendimento: Utilizado para estimar as receitas base sem a Revista
- Naturalidade: Não utilizado

Depois de todo o tratamento de inferências e após ignorar as colunas com um índice muito baixo de preenchimento, conseguimos obter um resultado mais satisfatório, principalmente para os últimos anos disponíveis acerca da venda da Revista. Nesse mesmo seguimento, para que as análises fossem representativas, focámo-nos nos dados disponíveis para os anos de 2016 a 2020.

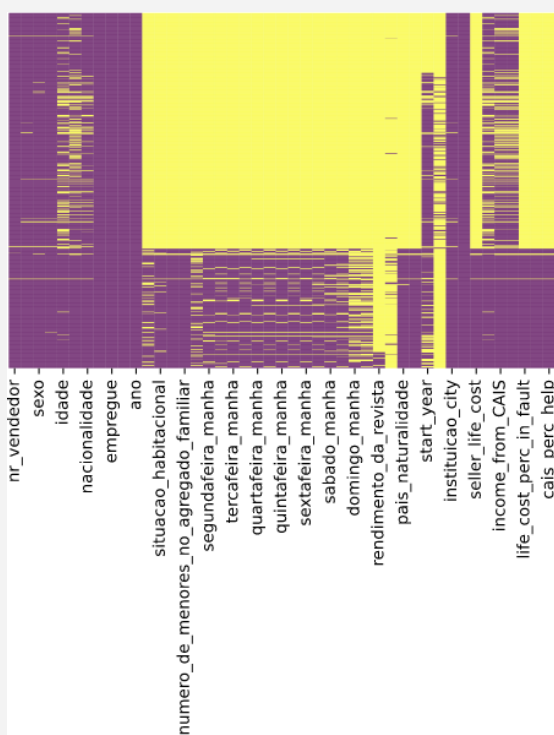


Figura 24 - Mapa de calor do preenchimento do dataset transformado total

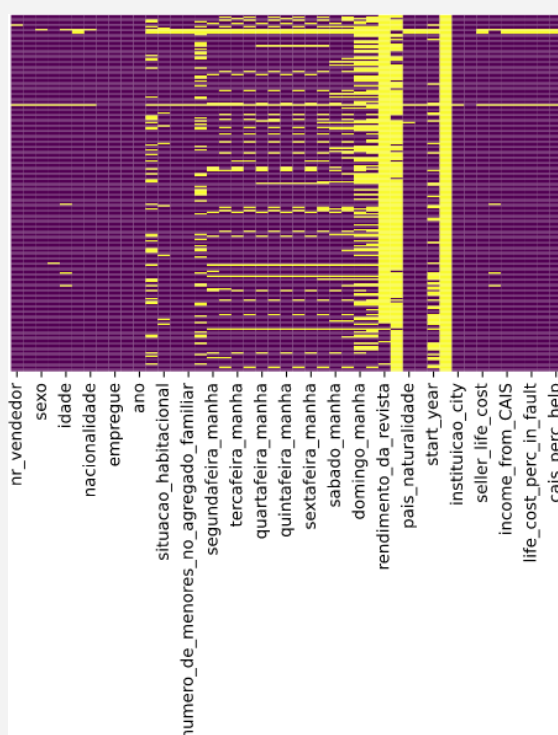


Figura 23 - Mapa de calor do preenchimento do dataset transformado de 2016 a 2020

Este conjunto de dados mais recentes é representado pela imagem mais à direita, onde vimos já uma falta de dados menos significativa.

Tal como foi referido anteriormente, todas as análises apresentadas foram feitas apenas para este intervalo, com a exceção da análise relativa ao tipo de gastos onde foi possível usar apenas os dados de 2020.

Foi com base nestes campos inferidos e calculados a partir dos originais que foi possível extrapolar, com os erros que lhes são inerentes, as conclusões feitas neste relatório. Estão elas sujeitas a erros, más interpretações, deduções erradas e atribuição de causalidade à correlação devido ao baixo número de vendedores nos anos analisados. Contudo, pelo menos intuitivamente parece-nos que apesar de todas as limitações, algumas das análises parecem fazer sentido e confirmam algumas das conclusões que eram já do conhecimento dos colaboradores da CAIS.

Análises incompletas

A análise dos dados disponibilizados possibilitou retirar conclusões importantes que vão de encontro ao primeiro objetivo estipulado: perceber e quantificar o impacto da Revista. Por outro lado, as análises necessárias à concretização do segundo objetivo



estipulado não foram possíveis pela dificuldade que a equipa teve em tirar conclusões e correlações devido à fraca qualidade dos dados e à baixa quantidade de vendedores nos anos entre 2016 e 2020. Por isso, propomos um conjunto de melhorias a ser adotadas pela CAIS, de forma a que no futuro se possa trabalhar no segundo objetivo proposto inicialmente e fazer análises potencialmente muito interessantes:

1. Criar um identificador único por vendedor. Cada vendedor deve ter um ID único (exemplo: número cartão de cidadão) usado para registar as atividades do vendedor. Assim um vendedor pode até ter coletes diferentes em dias diferentes, mas é possível identificar o mesmo vendedor em anos diferentes e acompanhar o seu progresso.
 - a. Com um identificador único passa a ser possível:
 - i. Acompanhar o progresso/história do vendedor ao longo do tempo
 - ii. Encontrar um padrão entre os vários vendedores
 - iii. Identificar/encontrar histórias de sucesso e insucesso da Revista CAIS (e otimizar a atuação)
 - iv. Encontrar o que varia entre a entrada de um vendedor na Revista CAIS e a sua saída
2. Usar nomes de rua mais estruturados em vez de nomes calão para os locais de venda dos vendedores (Exemplo: usar “avenida da liberdade, 98” em vez de “rotunda do rato”)

As propostas de melhoria não estão limitadas a estas, mas acreditamos que as propostas mencionadas são as essenciais para numa segunda iteração de análises voltar a identificar novas melhorias.

Dicionário de Dados

Coluna	Descrição	Tipo de dados
n._vendedor	Seller ID	Nominal
instituicao	Support Institution	Text
sexo	Gender	Categorical
activo	Active Member	Binary
contagem	(Unknown)	Binary



data_de_nascimento	Birthday	Date
idade	Age	Discrete
numero_maximo_de_vendas/mes	Target per month	Discrete
nacionalidade	Seller nationality(ies)	Categorical
escolaridade	Education	Categorical
empregue	Is employed	Binary
arquivado_cais_ou_so_revista	Is still an associated member	Categorical
ano	Activity Year	Discrete
data_inicio_venda	Date he started	Date
data_saida	Date he left	Date
mes_entrada	Month he started	Discrete
mes_saida	Month he left	Discrete
local_de_venda	Selling location	Text
concelho	Selling town(s)	Categorical
horario	Seller schedule	Categorical
arquivo	Is still an associated member	Categorical
naturalidade	Seller home city	Categorical
morada	Seller address	Text
situacao_habitacional	House situation	Categorical
numero_elementos_no_agregado_familiar	Number of household members	Discrete
numero_de_menores_no_agregado_familiar	Number of underage children	Discrete
outras_fontes_de_rendimento_(indicar_quais_e_quanto)	Other Seller income sources	Categorical
segunda-feira_manha	Sell Location on the Monday Morning	Text
segunda-feira_tarde	Sell Location on the Monday Afternoon	Text
terca-feira_manha	Sell Location on the Tuesday Morning	Text
terca-feira_tarde	Sell Location on the Tuesday Afternoon	Text
quarta-feira_manha	Sell Location on the Wednesday Morning	Text
quarta-feira_tarde	Sell Location on the Wednesday Afternoon	Text
quinta-feira_manha	Sell Location on the Thursday Morning	Text
quinta-feira_tarde	Sell Location on the Thursday Afternoon	Text





sextafeira_manha	Sell Location on the Friday Morning	Text
sextafeira_tarde	Sell Location on the Friday Afternoon	Text
sabado_manha	Sell Location on the Saturday Morning	Text
sabado_tarde	Sell Location on the Saturday Afternoon	Text
domingo_manha	Sell Location on the Sunday Morning	Text
domingo_tarde	Sell Location on the Sunday Afternoon	Text
rendimento_da_revista	In what the money is applied	Categorical





Os incríveis voluntários

Guilherme Cardoso

<https://www.linkedin.com/in/luminoso/>

João Brandão

<https://www.linkedin.com/in/jpabrandao/>

João Palmeiro

<https://www.linkedin.com/in/joaopalmeiro/>

Jorge Pérez

<https://www.linkedin.com/in/jlopezper/>

Pedro Moura

<https://www.linkedin.com/in/pedromanuelmoura>