



FACULDADE DE CIÊNCIAS DA UNIVERSIDADE DE LISBOA

VISUALIZAÇÃO DE DADOS - RELATÓRIO 2

PowerBI

António Rebelo · Clarisse Hetier · Tomás Quental

taught by
Professor MARIA BEATRIZ CARMO & ANA CLÁUDIO

6 de janeiro de 2021

Introdução

O seguinte relatório de utilização da ferramenta PowerBI tem como objetivo relatar a abordagem feita ao software bem como as dificuldades e críticas que novos utilizadores possam ter. O objetivo deste projeto foi de familiarização ao PowerBI ficando com uma ideia das técnicas de visualização que este permite realizar.

Os dados utilizados foram fornecidos pela Direção Geral de Estatística da Educação e Ciência (DGEEC) e são referentes ao ensino em escolas públicas portuguesas. Foi também utilizada a taxa de criminalidade em 2013 obtida através do portal do Instituto Nacional de Estatística. A taxa de natalidade obtida no ano 2001 e o poder de compra referente a 2015 foram também retirados do portal PORDATA [1, 2].

Como a partilha de workspaces do PowerBI estava desativada para as contas da FCUL, cada um dos membros do grupo fez os seus dashboards que publicámos com cada conta.

- <https://app.powerbi.com/groups/me/reports/bdc6c4a7-9694-46e1-8876-9b6f0d741c50?ctid=0bfa8500-b1f2-4566-baf1-6f59370893e7>
- <https://app.powerbi.com/groups/me/reports/1012654d-f277-4ded-84a4-93df4aceb6c1?ctid=0bfa8500-b1f2-4566-baf1-6f59370893e7>
- <https://app.powerbi.com/groups/me/dashboards/2bfa191c-70c4-42d6-87f1-528c3c18d6f8?ctid=0bfa8500-b1f2-4566-baf1-6f59370893e7>

Técnicas de visualização de dados oferecidas pelo PowerBI

Nesta seção expõem-se os tipos de visualização que o PowerBI disponibiliza, bem como as situações mais adequadas para o seu uso.

Gráficos de barras/colunas empilhadas, agrupadas e a 100%

Gráficos de colunas empilhadas são úteis quando a quantidade em interesse é a soma de duas ou mais variáveis, mas a proporção em que essa soma se dá também importa. Se a proporção é o que importa e o total nem tanto, então nesse caso é boa prática utilizar o gráfico de colunas a 100%. Se a quantidade de variáveis que participa na soma for grande (grande neste contexto é relativo), um gráfico de colunas agrupadas é mais útil para discernir as diferentes quantidades mais facilmente. Por último, quando a quantidade de pontos/colunas a mostrar é elevada, é melhor usar gráficos de barras e mostrar a informação na horizontal pois a comparação é mais fácil nesse caso. Ainda assim, a dificuldade de processar a informação quando há muitas colunas/barras é a principal desvantagem deste tipo de gráfico.

Gráficos de linhas

Os gráficos de linhas são muito versáteis e usam-se nas mais variadas situações, contudo eles adequam-se especialmente a mostrar evoluções periódicas ou outro tipo de tendências similares. A sua simplicidade é também o seu forte.

Gráficos de área e gráficos de áreas empilhadas

Este tipo de gráfico é mais adequado quando a magnitude é mais importante que os valores exatos, caso em que um gráfico de linhas seria preferível. De fato, as áreas coloridas tornam a percepção das magnitudes mais fáceis e portanto são ideias para fazer comparações.

Gráfico de friso

Este tipo de gráfico é similar ao gráfico de colunas empilhadas com a diferença que é ordenado e tem um maior foco na evolução do gráfico sendo assim especialmente adequado a mostrar evoluções de dados periódicos.

Gráfico da cascata

O gráfico de cascata também enfatiza a evolução, normalmente temporal de uma certa variável. Ao mostrar a variação a cada passo, normalmente codificada através de cores para diferenciar entre variações positivas e negativas, transmite a ideia de valores iniciais a se transformarem em valores finais e vice-versa. Geralmente este tipo de gráfico serve para visualizar a performance de dado indicador.

Gráfico em funil

Este tipo de gráfico é utilizado mais frequentemente para ilustrar as diferentes etapas de um processo sequencial. Contudo pode também ser usado numa forma similar ao gráfico de barras. Um gráfico de funil é útil principalmente em situações onde há mais de duas frações a comparar, caso contrário o gráfico de colunas empilhado ou gráfico circular darão melhores resultados.

Gráfico de dispersão

O gráfico de dispersão deve ser utilizado quando o pretendido é averiguar se existe uma correlação entre as variáveis.

Gráfico circular e gráfico em anel

Gráficos circulares devem ser usados para mostrar de uma forma mais geral uma fração, geralmente referente a uma dada categoria dentro de um todo. Quando a fração exata é importante outro tipo de gráfico deve ser usado. Um gráfico circular também pode ser usado para transmitir a ideia que um segmento do total é relativamente pequeno ou grande.

O gráfico em anel é similar ao gráfico circular, mas devido à maneira como o cérebro processa áreas e linhas de formas diferentes, o gráfico em anel geralmente transmite o seu conteúdo melhor do que um gráfico circular, tornando a sua leitura mais fácil. Contudo, também apresenta as suas desvantagens uma vez que a comparação com outros gráficos é difícil.

Treemap

O treemap adequa-se à visualização de dados hierárquicos, geralmente no contexto em que é útil saber a fração de um todo de muitas categorias e a comparação precisa entre as frações não é importante.

Mapas

A visualização de mapas é útil quando há uma relação geográfica com as variáveis em estudo.

Cartões

Quando um único número tem especial importância, o cartão permite dar-lhe o devido destaque.

Tabelas

A tabela é útil quando:

- A visualização é usada para consultar valores individuais
- É usada para comparar valores individuais, mas não séries inteiras de valores
- Valores precisos são necessários
- A informação quantitativa a ser comunicada envolve mais do que uma unidade de medição.

Segmentação de dados

A segmentação de dados é uma ferramenta poderosa de filtragem de dados que é muito versátil. É útil para mostrar/esconder níveis de uma dada hierarquia, para filtrar dados temporais, para selecionar uma dada métrica ou unidade dos dados. Há realmente uma grande gama de aplicações em que a segmentação de dados pode ser utilizada e ela torna a interação com o dashboard mais dinâmica e precisa.

Ferramentas analíticas

O PowerBI dispõe de ferramentas analíticas que permitem saber os principais influenciadores de uma dada variável e árvores de decomposição que permitem explorar os dados de uma forma interativa. Árvores de decomposição são também ótimas maneiras de representar dados hierárquicos.

Capacidades interativas do PowerBI

O PowerBI tem também capacidades interativas que permitem interagir entre os gráficos, filtrar categorias ou conjuntos de valores, ampliar gráficos e interagir com menus. Este conjunto de características torna o PowerBI muito poderoso e versátil.

Estudo do Poder de Compra

A hipótese em teste é a de que o poder de compra tem um impacto no desempenho escolar, uma vez que o acesso a ferramentas como explicadores e outro tipo de factores socioeconómicos aumentam com o poder de compra.

Primeiramente, um estudo foi feito para analisar a variável do poder de compra por município em si. Depois explorou-se a sua relação com o desempenho escolar e com a média de percentagem de alunos com auxílio social. Toda esta análise foi feita de forma a que possa ser escolhido o nível de granularidade geográfica que se pretende visualizar (Municípios, NUTS II ou NUTS III) recorrendo a um painel de segmentação de dados. O dashboard em questão é apresentado na Fig. 1

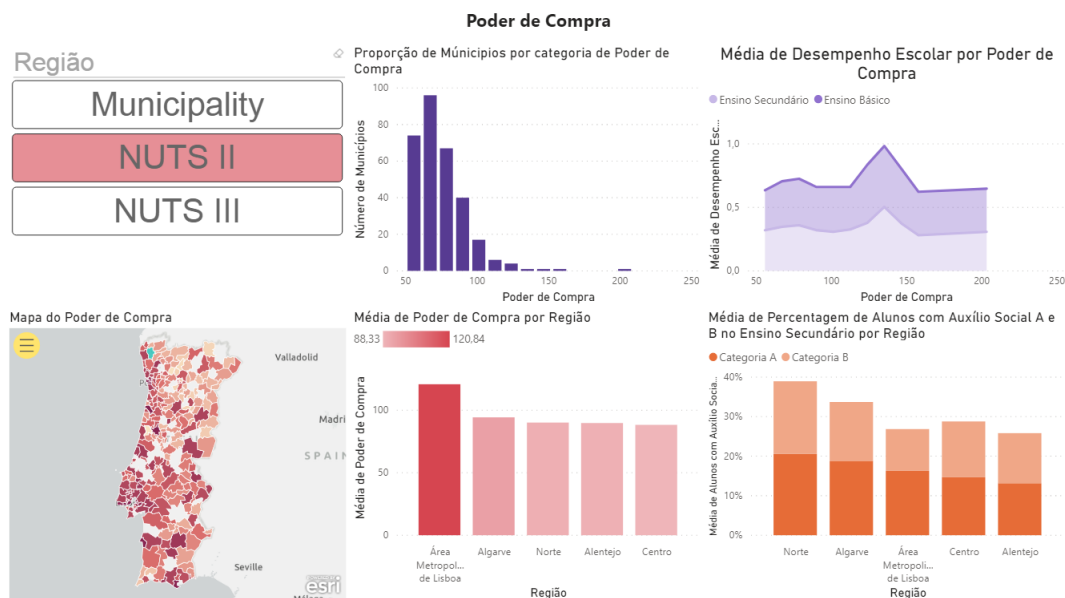


Figura 1: Dashboard relativo ao poder de compra.

Dada a gama de combinações e possibilidades de interação, analisar todas as combinações torna-se impraticável, então a análise aqui incidirá sobre o nível NUTS II. No entanto antes de seguir com a análise vale a pena demonstrar quais essas possibilidades.

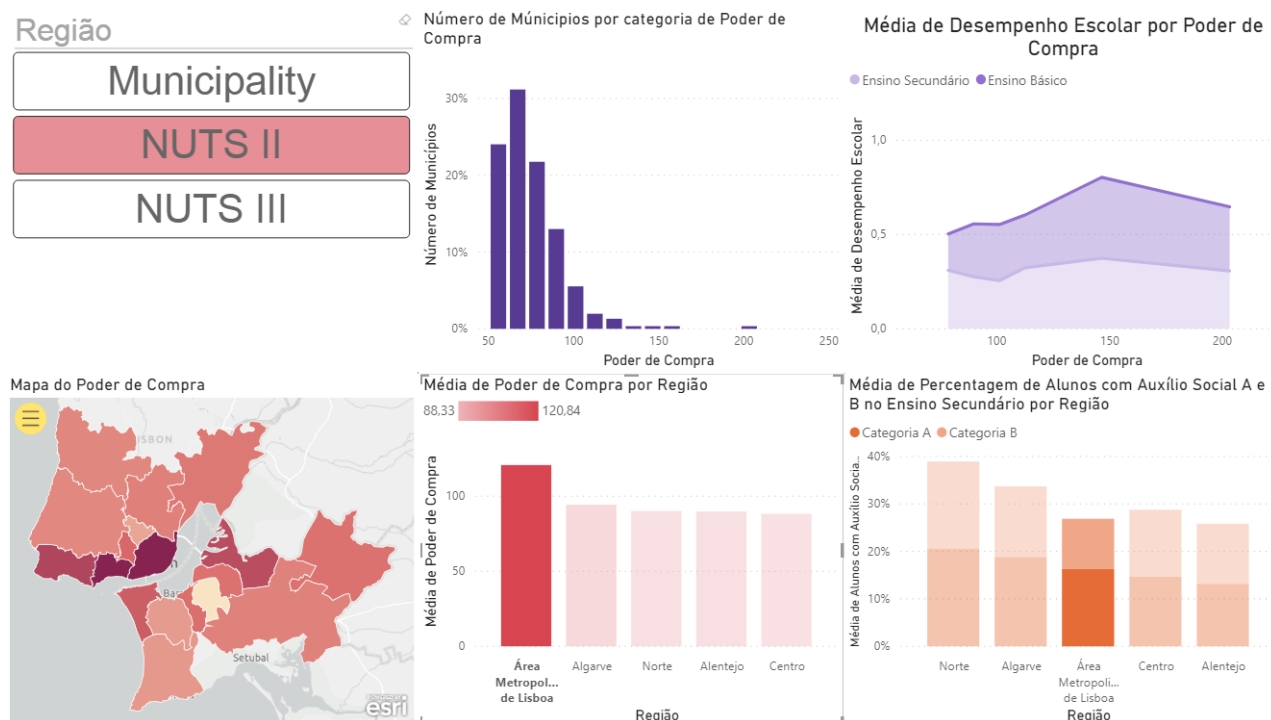


Figura 2: Dashboard exemplificando a seleção do nível NUTS II e a área metropolitana de Lisboa.

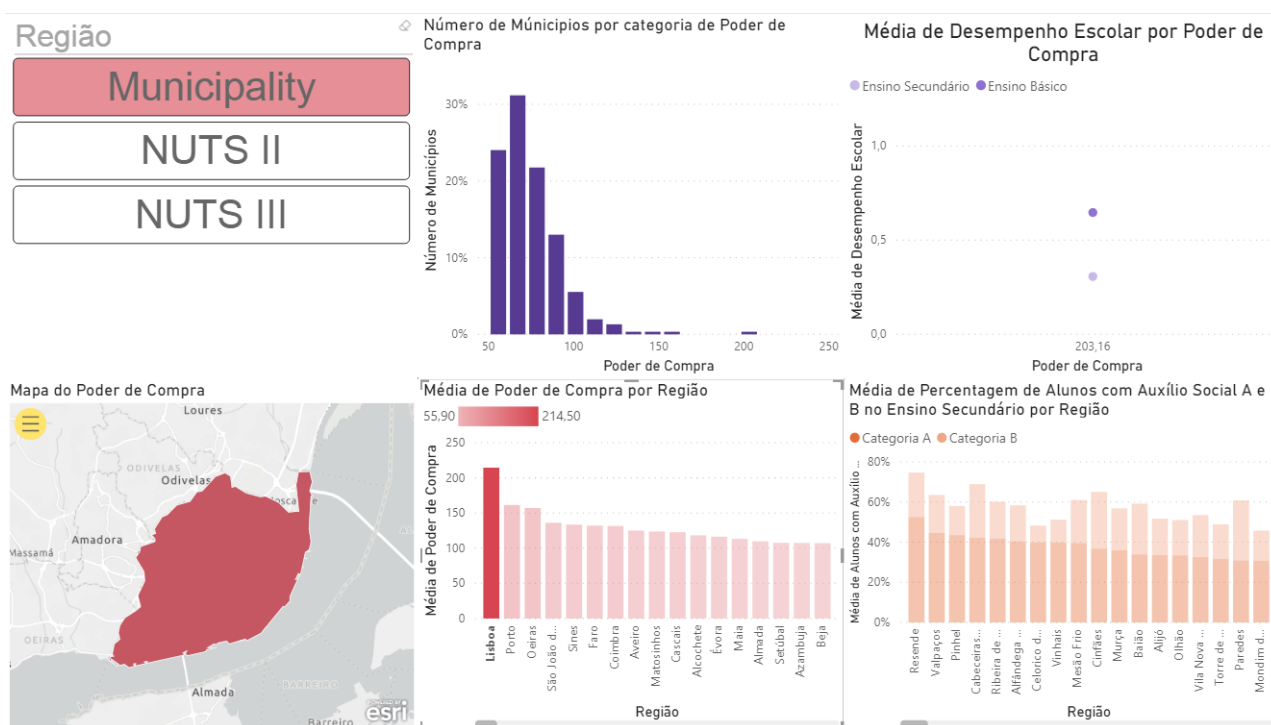
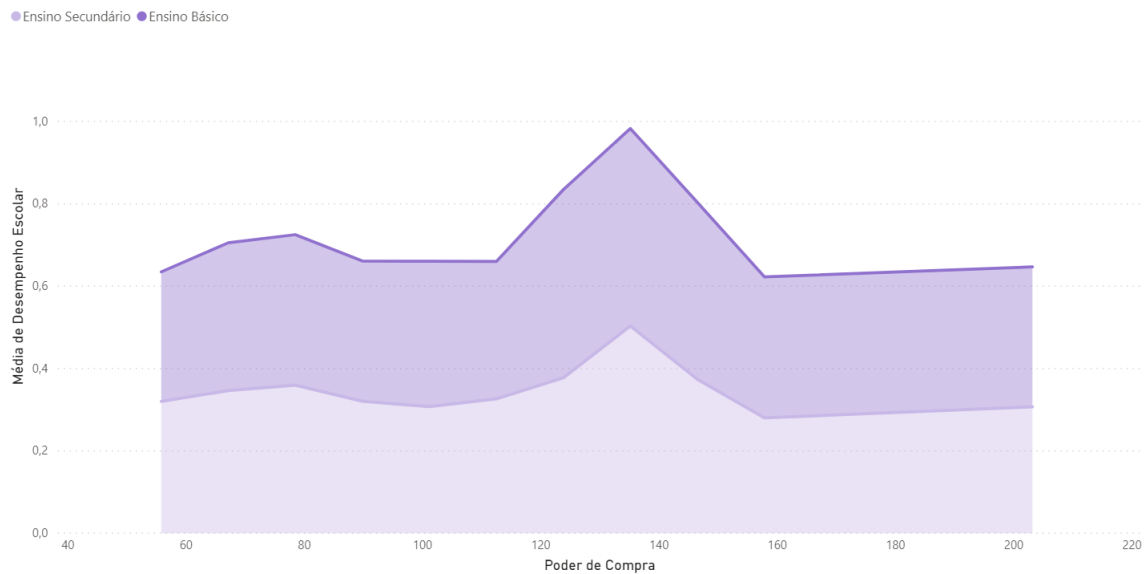


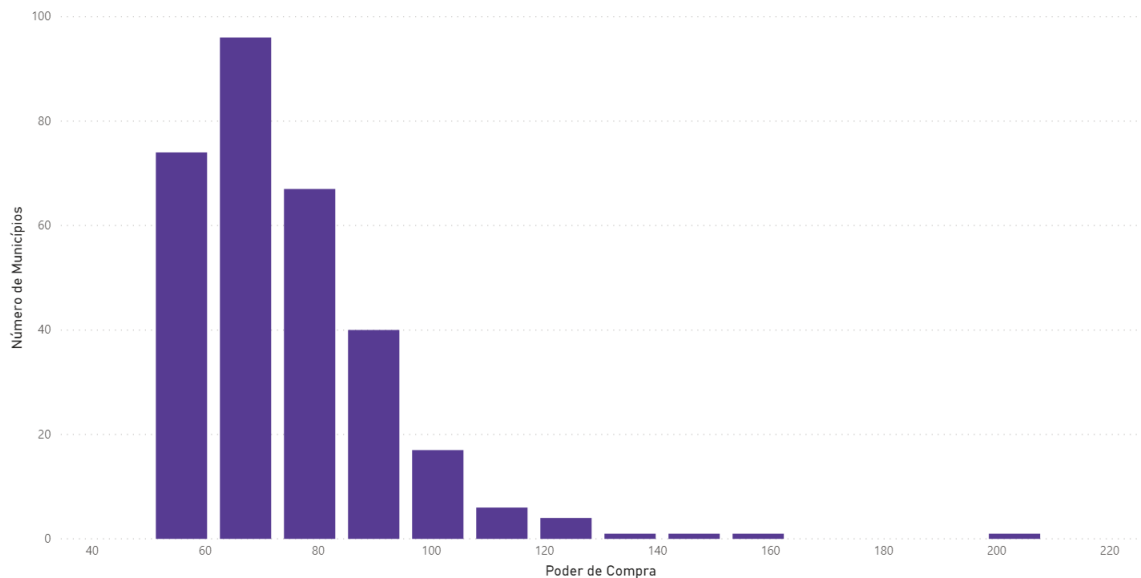
Figura 3: Dashboard exemplificando a seleção do nível "Municipality" e o município de Lisboa

Interpretação

Em primeiro lugar analisamos os resultados que não dependem da seleção da região, isto é, os gráficos a roxo.



(a) Média das médias municipais do desempenho Escolar por categoria de poder de compra para o ensino básico e secundário.



(b) Número de municípios por categoria de Poder de Compra. Esta distribuição influencia o gráfico (a).

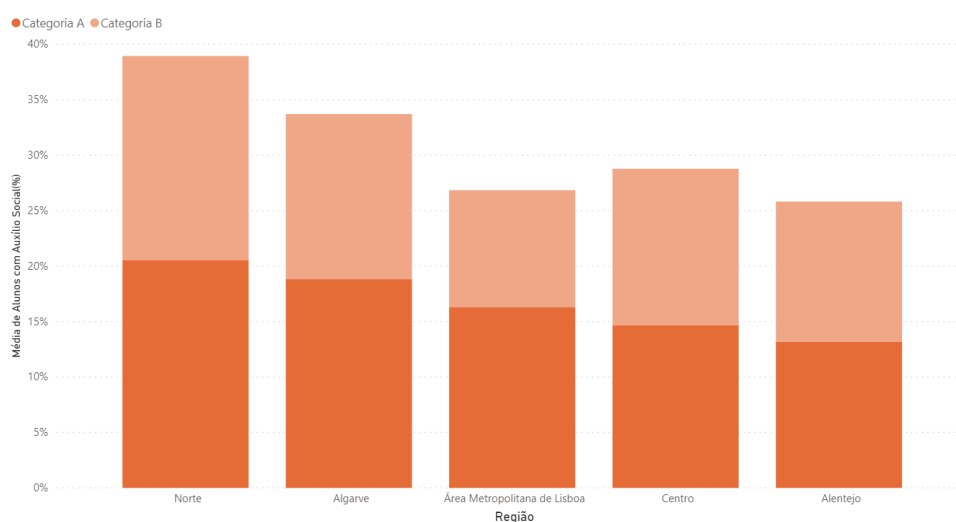
Figura 4: Relação entre poder de compra e aproveitamento escolar.

Na Fig. 4a vemos a relação entre o poder de compra e o desempenho escolar. Antes de mais, há que referir que a construção deste gráfico foi um desafio. O problema está relacionado com a estrutura dos dados. A informação relativa ao poder de compra é apenas referente a cada município e não a cada escola, como seria o ideal. O poder de compra é uma variável contínua. A fim de poder tornar possível a visualização, os municípios são agrupados por categoria de poder de compra. Assim, a média do desempenho escolar é calculada para cada município e depois dentro de cada categoria de poder de compra, outra média é calculada. Claro que isto faz com os dados apresentados sejam muito ruidosos porque há várias fontes de erro. Se dentro de um município houver escolas que têm realidades e poderes de compra associados completamente diferentes, a média vai cancelar essas diferenças e elas não ficarão visíveis no gráfico. É o caso do município de Lisboa, por exemplo, que tem o poder de compra mais alto do país e escolas que têm o melhor desempenho a nível nacional, mas que também

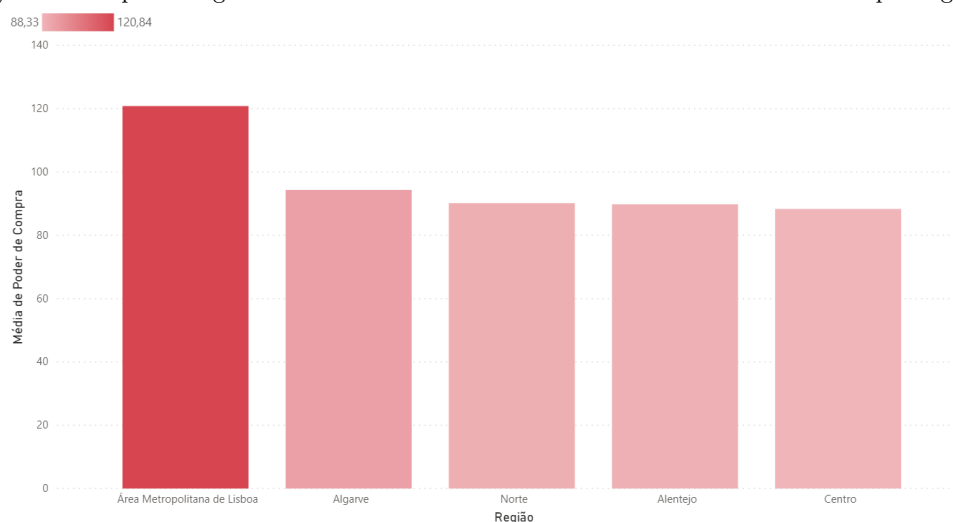
tem realidades mais pobres com desempenhos mais baixos.

Outra fonte de ruído vem da média feita ao nível da categoria de poder de compra, uma vez que os municípios muito populosos tendem a comportar-se como Lisboa. Ou seja, os erros propagam-se. Estes são tanto menores quanto maior o número de municípios usado para fazer a média. Por esta razão o gráfico 4b é apresentado. Observa-se que valores muito altos de poder de compra têm uma amostragem muito pouco significativa então poucas conclusões podem ser retiradas para essa parte do gráfico. Se no entanto, os municípios com poder de compra superior a 130 forem desconsiderados, é possível observar uma correlação positiva fraca entre o poder de compra e o desempenho escolar. Pelas razões acima referidas, esta não é uma conclusão fiável e um estudo mais aprofundado seria necessário para explorar a relação entre estas duas grandezas.

No mesmo dashboard apresenta-se a relação entre poder de compra e percentagem de alunos com auxílio social por região.



(a) Média de percentagem de alunos com auxílio social A e B no ensino secundário por região.



(b) Poder de compra por região.

Figura 5: Relação entre poder de compra, região e auxílio social.

Os gráficos da Fig. 5 de certa forma ilustram o ponto referido acima. Por exemplo, embora o Algarve tenha dos poderes de compra mais elevados, tem também uma das taxas de alunos com

auxílio social mais altas, evidenciando que existem realidades muito diferentes dentro desta região. Já relativamente ao Centro e ao Alentejo o contrário acontece: estas regiões têm baixo poder de compra comparativamente ao resto do país, mas taxa de auxílio social baixa, indicando que nestas regiões há menos desigualdade.

As regiões de NUTS II são muito abrangentes o que não é necessariamente mau, porque também há elações a tirar a este nível, contudo selecionando o nível "Municipality" é possível explorar estas relações com maior granularidade e aí o comportamento é diferente. Por exemplo, o município de Resende é o que tem mais alta taxa de alunos com auxílio social e tem o 6º poder de compra **mais baixo** de entre todos os municípios. Isto manifesta-se também no aproveitamento escolar como será apontado mais à frente. Aumentando a granularidade, é portanto possível retirar outro tipo de conclusões e diferenciar municípios mais e menos necessitados.

O único gráfico não mencionado até agora é o mapa do poder de compra por município.

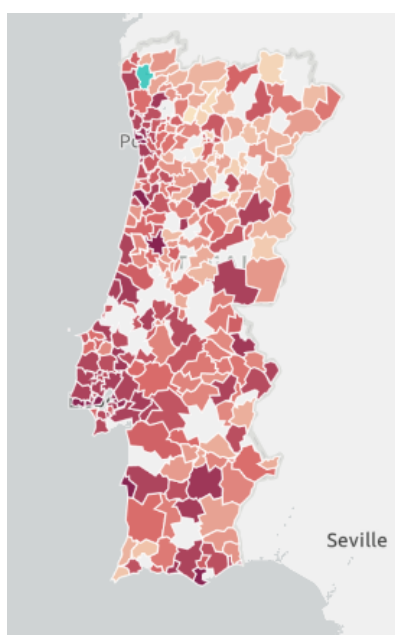


Figura 6: Mapa do poder de compra por município.

Este mapa serve mais como suporte aos restantes gráficos para que se possa ter uma ideia do poder de compra também a um nível geográfico. A sua execução foi no entanto um desafio. O PowerBI tenta encontrar as regiões associadas a cada município pelo seu nome, contudo nem sempre é bem sucedido. Como se pode ver no mapa, há vários municípios que não estão representados. Selecionando as subregiões de NUTS II verifica-se também que vários municípios estão representados no sítio errado. Para contornar este problema, a latitude e longitude de cada município foram inseridas, mas neste caso o problema seria que a representação pretendida não está disponível, uma vez que apenas círculos no lugar de cada município e não as suas fronteiras são renderizados.

Estudo da taxa de Natalidade

A taxa de Natalidade ou taxa de Natalidade bruta, definida como o número de nascimentos por mil habitantes referente ao ano de 2001 é explorada nesta secção. Apenas três relações são apresentadas: a sua distribuição geográfica, a sua relação com a taxa de ocupação das escolas e finalmente, com os anos de escolaridade da mãe.

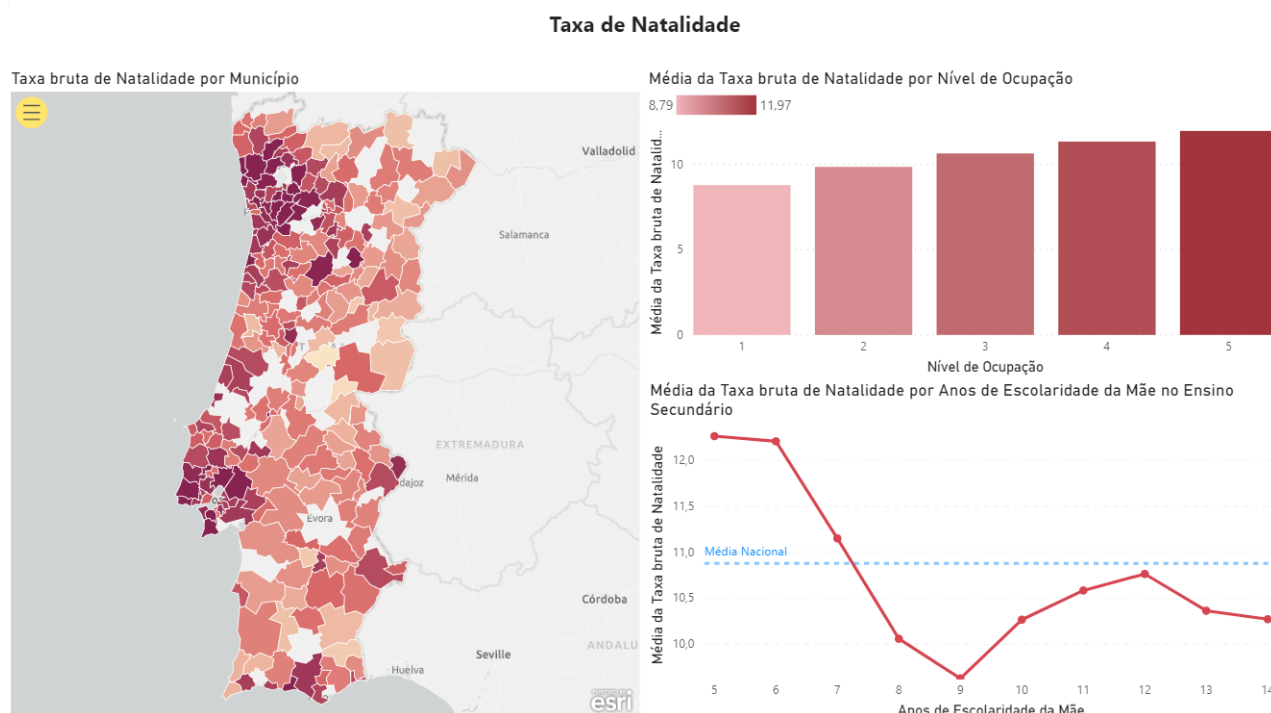


Figura 7: Dashboard relativo à taxa de Natalidade.

Observe-se primeiro a relação com a taxa de ocupação.

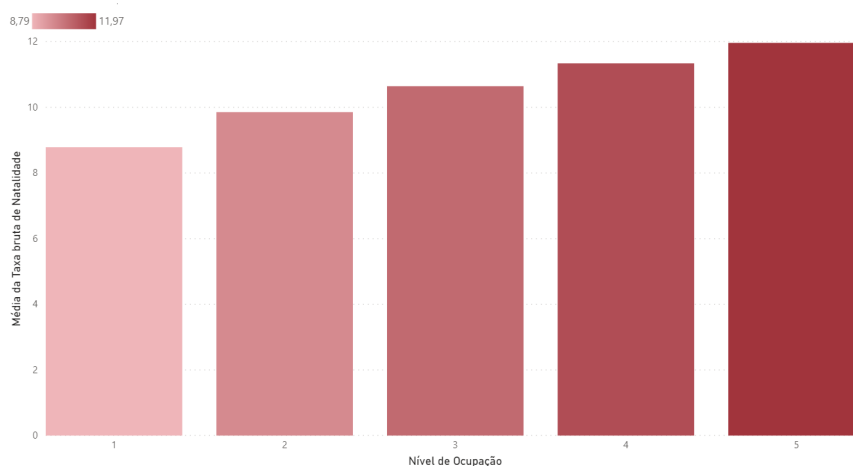


Figura 8: Média da Taxa bruta de Natalidade por Nível de Ocupação

Um gráfico de linha foi considerado para a visualização desta variável visto que seria mais compacto a nível de informação ao usar menos tinta, mas uma vez que aqui se tratam de níveis categóricos de ocupação, optou-se por um gráfico de colunas.

Existe uma relação linear clara entre o nível de ocupação e a média da taxa de natalidade. Note-se que embora os dados do nível de ocupação são referentes a 2015 a taxa de natalidade apresentada refere-se a 2001. É então natural que 14 anos depois, nos sítios onde a natalidade foi baixa o nível de ocupação também seja baixo.

De seguida estuda-se a relação da taxa de natalidade com a escolaridade maternal.

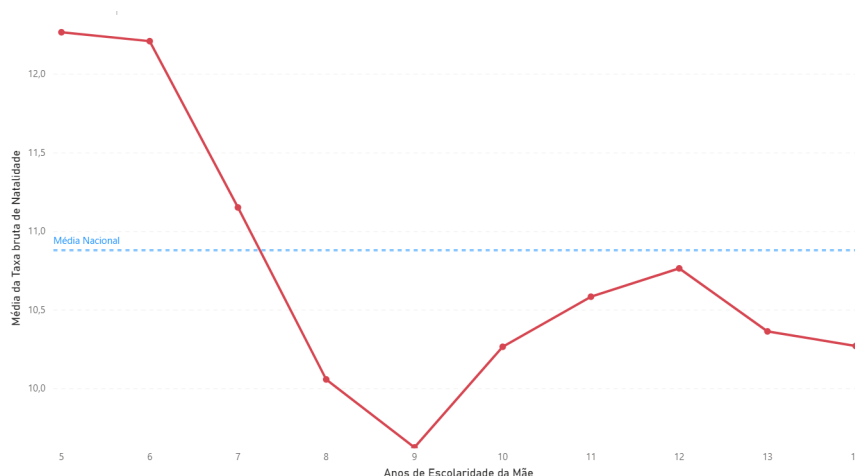


Figura 9: Média da Taxa bruta de Natalidade por anos de escolaridade da Mãe.

Antes da análise propriamente dita, vale afirmar que para esta análise, os mesmos problemas encontrados na seção do poder de compra se manifestam, uma vez que os dados estão estruturados da mesma forma. Note-se também que a escala da taxa de natalidade não começa em zero, uma vez que o ênfase está em como esta variável varia à volta da média nacional. Outra dificuldade que foi sentida não só neste gráfico, mas em todo o trabalho foi a de conseguir usar barras de erro. Se o PowerBI é capaz de fazer gráficos com base na média por categoria, poderia facilmente demonstrar uma opção para mostrar barras de erro correspondentes ao desvio padrão. Contudo, esse não é o caso e assim fica difícil avaliar a legitimidade de fazer a alteração de escala no eixo das ordenadas.

Ainda assim, verifica-se uma relação curiosa: municípios em que as mães têm níveis de escolaridade mais baixos têm taxas de natalidade elevadas. No nível de escolaridade intermédio é quando a taxa de natalidade é menor e depois volta a aumentar um pouco. A interpretação sociológica não é do foro dos nossos estudos, mas uma pesquisa superficial indica que a natalidade diminuir com o nível de educação materno é um resultado bem estabelecido na literatura. Talvez uma possível explicação para o vale do gráfico seja a de que as mães com mais anos de estudo têm empregos mais confortáveis, o que lhes dá a disponibilidade de ter mais filhos. De forma semelhante à conclusão relativa ao poder de compra, seria preciso um estudo mais aprofundado e um tratamento estatístico adequado que fogem ao escopo deste trabalho, para poder estar certo destes dados.

Finalmente, o mapa da taxa de natalidade.

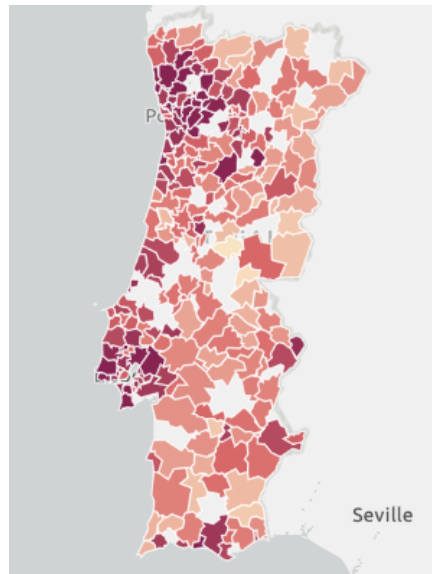


Figura 10: Mapa da Taxa bruta de Natalidade por Município

Observa-se que as zonas do interior e Alentejo têm natalidades mais baixas, como seria de esperar tendo em conta o processo de desertificação populacional que ocorre nestas áreas.

Estudo da taxa de desistência no Secundário

Com a variável "Dropout" tentou-se perceber em primeiro lugar qual era o panorama geral no país, verificando qual a percentagem média de desistência da escola no secundário em Portugal, refinado por NUTS II, NUTS III e Município. De seguida, criaram-se visualizações com duas variáveis que poderiam ter relação com a taxa de desistência: anos de escolaridade das mães e criminalidade. Obteve-se o seguinte dashboard:

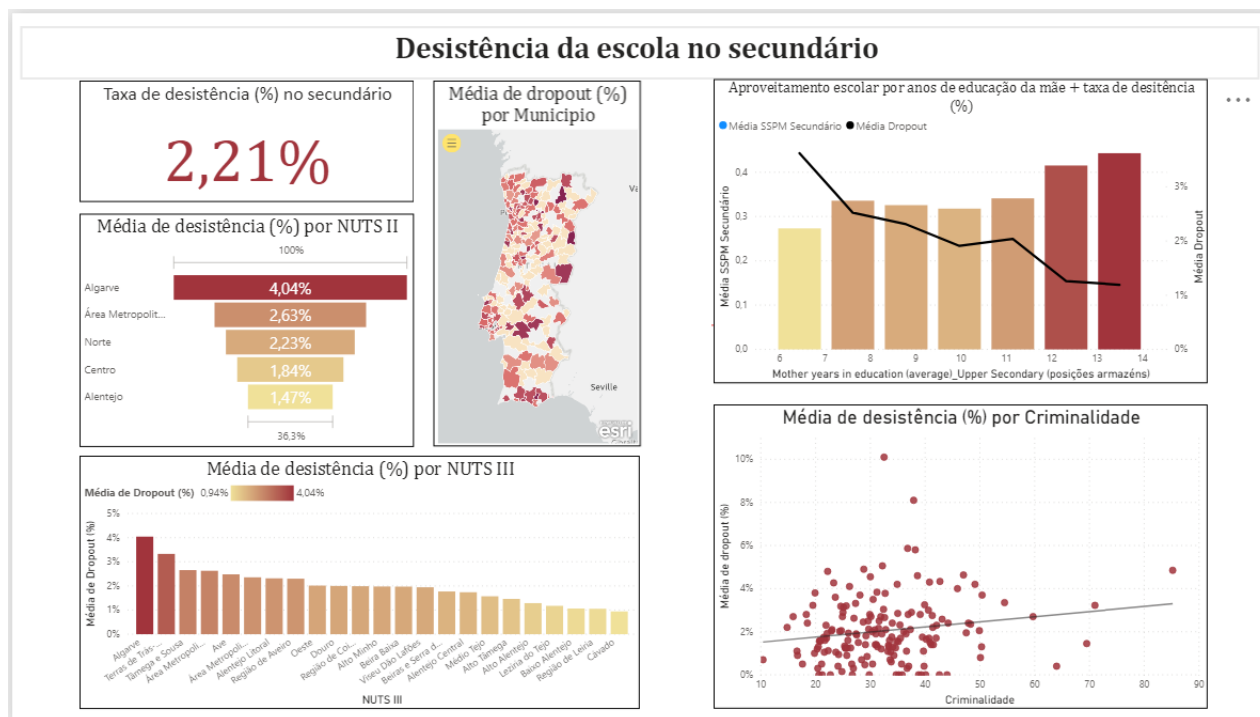


Figura 11: Dashboard relativo à desistência da escola no secundário

No lado esquerdo do dashboard encontram-se as taxas de desistência no país. Do lado direito, a relação da taxa de desistência com outras variáveis.

No canto superior esquerdo pode-se observar um cartão com uma percentagem referente à taxa de desistência global em Portugal. Por baixo, tem-se um esquema em funil com as taxas de desistência por NUTS II, com uma paleta de cores entre o vermelho escuro (valores mais altos) e amarelo (valores mais baixos). No fundo do lado esquerdo encontra-se um histograma que representa novamente a taxa de desistência mas desta vez por NUTS III com a mesma paleta de cores. No mapa de Portugal encontram-se os municípios coloridos com uma paleta de cores que representa, nos tons escuros de vermelho, maior taxa de desistência e, em tons de amarelo mais claro, a menor taxa de desistência.

Do lado direito vê-se no gráfico superior um histograma que representa o aproveitamento escolar dos alunos por anos de escolaridade das mães, com gráfico de linhas que representa a taxa de desistência por anos de escolaridade das mães. O gráfico inferior é um gráfico de dispersão da desistência por criminalidade com uma linha de tendência associada.

É possível com este dashboard focar-se apenas na NUTS II, NUTS III ou no município clicando sobre os valores nas visualizações correspondentes a essas representações. Estas observações são feitas nas páginas seguintes na Figura 14, Figura 15 e Figura 16, respetivamente.

Interpretação

Começa-se a interpretação dos gráficos do dashboard pelos gráficos representados à direita, que representam os gráficos das relações entre a taxa de desistência com as variáveis criminalidade, e aproveitamento escolar.

Na figura 12 está representado um gráfico de dispersão com um linha de tendência que tenta explicar a relação entre a Taxa de desistência da escola no secundário e a criminalidade. Pode-se observar uma ligeira tendência positiva, aumentando a criminalidade com a taxa de desistência da escola no secundário. Contudo, os pontos são muito dispersos não sendo possível tirar grandes conclusões.

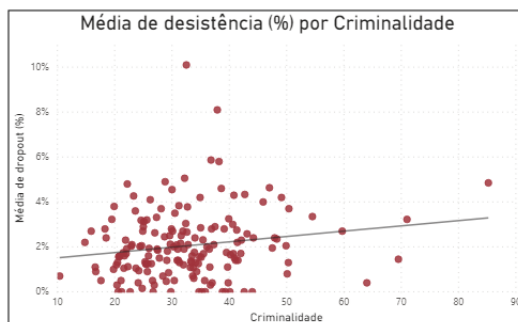


Figura 12: Gráfico de dispersão da Média de desistência (%) por criminalidade

De seguida, na figura 13 observa-se através do histograma que com os anos de escolaridade das mães, o nível de aproveitamento médio aumenta. Através da linha a preto vê-se que a taxa de desistência diminui consoante os anos de escolaridade das mães, faz sentido uma vez que o nível de aproveitamento é superior e que com um maior nível de educação das mães, vem todo um agregado de condições socioeconómicas favoráveis que acabam por impactar negativamente a desistência.

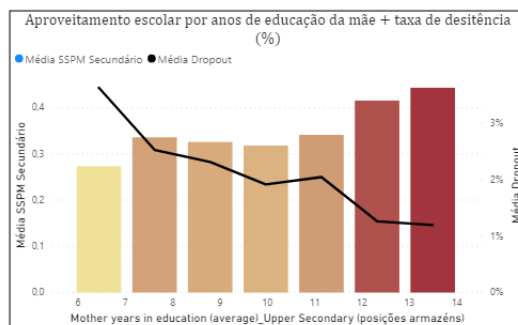


Figura 13: Gráfico de barras do aproveitamento escolar dos alunos por anos de escolaridade das mães, com gráfico de linhas que representa a taxa de desistência

Na figura 14 têm-se o exemplo do que acontece quando se interage no dashboard escolhendo na representação visual do funil a NUTS II referente ao Norte. Obtem-se uma taxa de desistência igual a 2.23%. No mapa aparecem os municípios do Norte e no Histograma das NUTS III, as NUTS III do Norte ficam em destaque.

Nos gráficos da direita também há alterações. Parece haver maior taxa de aproveitamento do que a média nacional, independentemente dos anos de escolaridade das mães. Consegue-se observar que a taxa de desistência é significativamente menor quantos mais anos de escolaridade têm as mães. No gráfico da média de desistência por criminalidade a tendência parece ser inversa à nacional, com uma ligeira diminuição da taxa de desistência com o aumento da criminalidade, mostrando mais uma vez de que poderá não haver relação entre as duas variáveis.

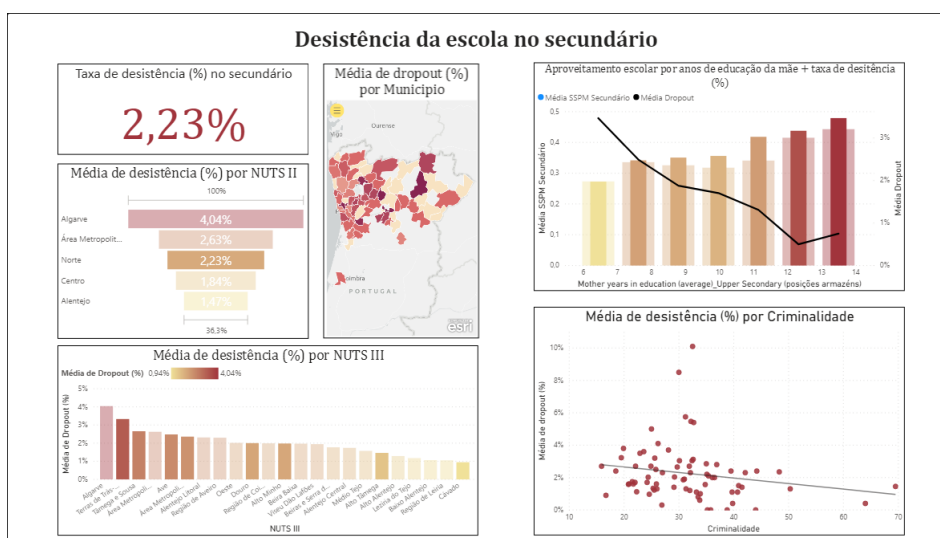


Figura 14: Dashboard relativo à desistência (%) da escola no secundário, com ênfase na NUTS II - NORTE

A figura 15 foca-se na NUTS III Terras de Trás-os-Montes. No funil da NUTS II conseguimos perceber que em Trás os Montes a média da taxa desistência é superior à do Norte.

Do lado direito, no histograma da direita, mais uma vez consegue-se ver que onde a o nível de aproveitamento é menor a taxa de desistência é maior. Mais uma vez, no gráfico de dispersão não parece haver relação entre a taxa de criminalidade e a taxa de desistência da escola no secundário.

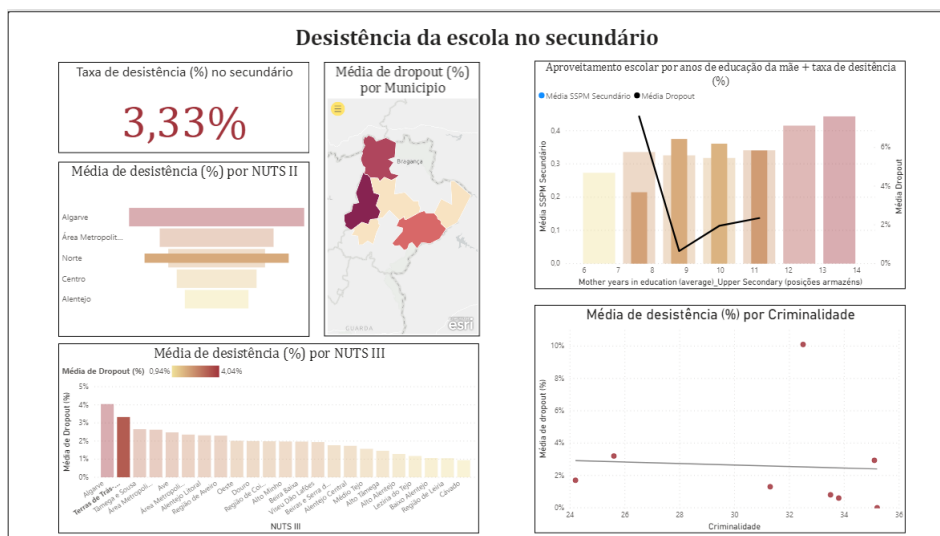


Figura 15: Dashboard relativo à desistência (%) da escola no secundário, com ênfase na NUTS III - Terras de Trás-os-Montes

Na figura 16, o exemplo escolhido foi o do município de Mirandela onde se observa uma taxa de desistência muito elevada de 10.10%, muito superior à média da sua NUTS II e III como conseguimos observar no gráfico do funil e no histograma abaixo.

À direita, nenhuma informação é possível retirar com apenas um ponto.

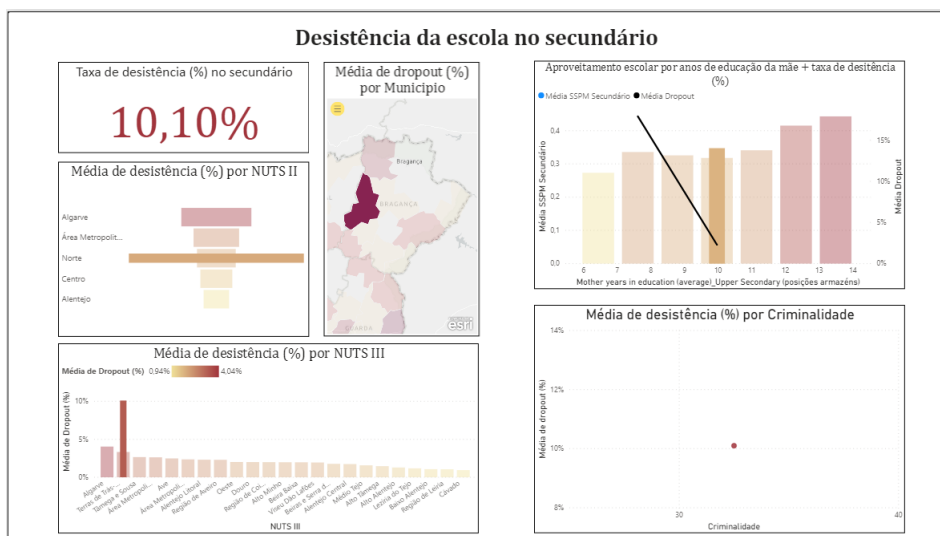


Figura 16: Dashboard relativo à desistência (%) da escola no secundário, com ênfase no município - Mirandela

Estudo sobre o aproveitamento Escolar dos alunos

Com este dashboard tem-se como objetivo mostrar dois tipos de visualização que o PowerBI nos permite fazer de modo automático dos principais influenciadores e árvores de decomposição de uma variável.

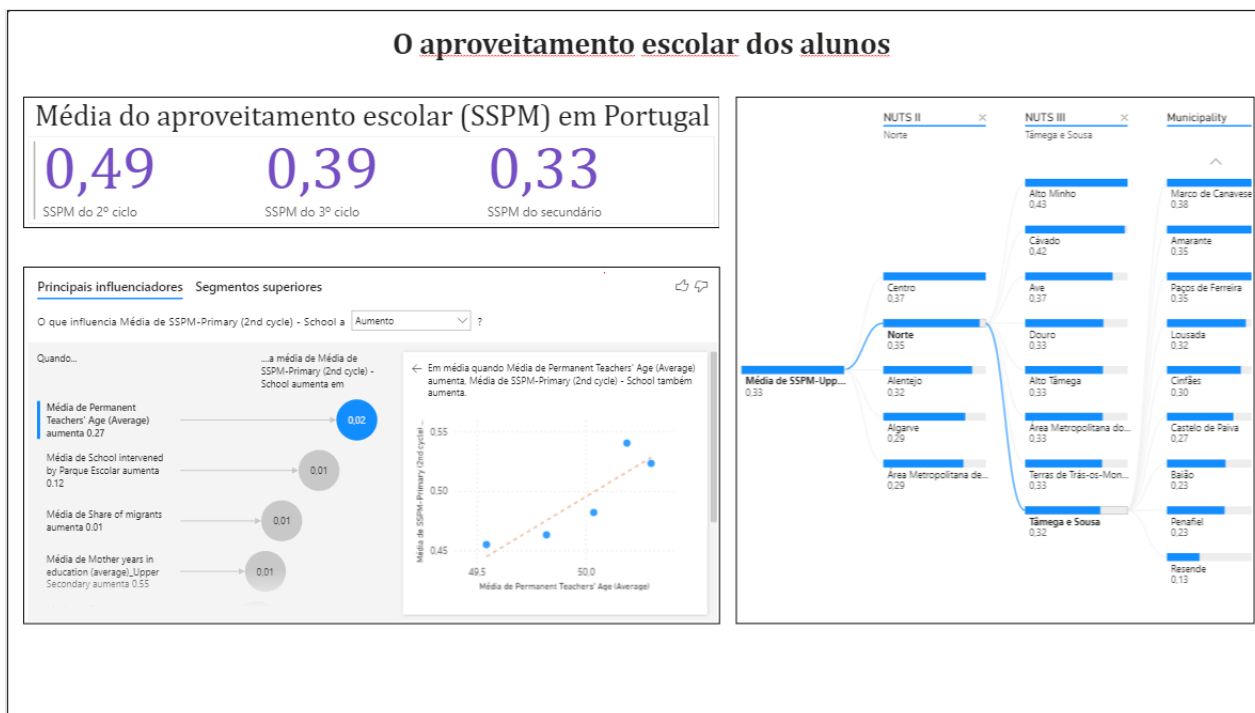


Figura 17: Dashboard relativo ao aproveitamento escolar dos alunos

No canto superior esquerdo colocou-se um cartão de linhas multiplas com a média do aproveitamento escolar no 2º ciclo do ensino básico, 3º ciclo do ensino básico e no secundário. No canto inferior esquerdo temos a ferramenta analítica dos principais influenciadores da variável SSPM (que representa o aproveitamento escolar). E do lado direito temos a ferramenta analítica que forma uma árvore de decomposição do aproveitamento escolar em Portugal, NUTS II, NUTS III e município.

Interpretação

Passa-se à interpretação das visualizações apresentadas anteriormente, iniciando-se pelo cartão de linhas múltiplas, onde se pode verificar que à medida que o nível escolaridade vai aumentando o aproveitamento diminui (Fig. 18).



Figura 18: Cartão com a média do aproveitamento escolar em Portugal, NUTS II, NUTS III e nos municípios

De seguida interpreta-se a visualização da funcionalidade analítica do PowerBI de verificar quais as variáveis que mais influenciam o aproveitamento escolar (Fig. 19). Foram escolhidas as seguintes variáveis teste: média de percentagem de imigrantes, média de professores permanentes, média da idade dos professores permanentes, média de escolas com intervenção da Parque Escola, média do número de estudantes e média de alunos com necessidades especiais. Com esta funcionalidade constatou-se que as variáveis mais influentes para o aproveitamento escolar possam ser a média da idade dos professores permanentes, a média do número de escolas com intervenção da Parque Escola e a média de percentagem de imigrantes.

Concentrando-se no exemplo da Fig. 19, vê-se que se forma automaticamente um gráfico de dispersão com uma linha de tendência que demonstra quanto maior for a média das idades dos professores permanentes mais alto é o nível de aproveitamento. Contudo, sendo crítico na análise observamos que a média das idades dos professores não varia entre 49.5 e 51, assim não é possível considerar este um fator relevante para o aproveitamento. Era necessário ter municípios com médias de idades dos professores muito diferentes para compreender se de facto a idade dos professores é relevante no aproveitamento.

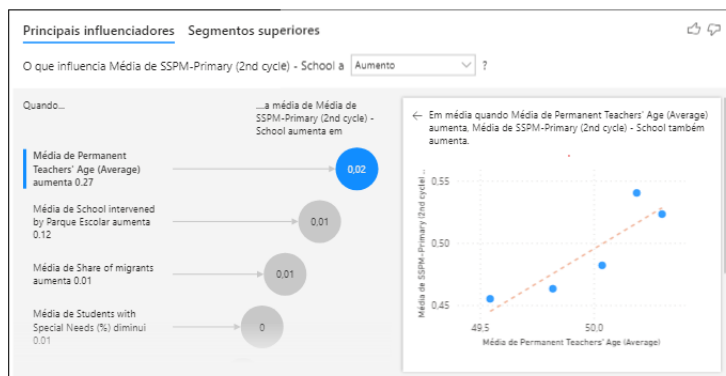


Figura 19: Variáveis influenciadoras do aproveitamento escolar no 2º ciclo

Na Fig. 20, observa-se a árvore de decomposição do aproveitamento escolar por região. Especial atenção para o Município de Resende que pertence a uma NUTS II (Norte) que tem alto aproveitamento, mas que é pior da sua região de NUTS III. Como já foi visto acima, este município tem baixo poder de compra.



Figura 20: Árvore de decomposição do aproveitamento escolar por região

Estudo do rácio aluno-professor no Portugal Urbano e Rural

Com as contagens do número de alunos e professores (permanentes e contratados), foi criada uma nova coluna com o rácio aluno-professor de cada estabelecimento de ensino. Foram filtrados os dados para que apenas se contassem escolas com 3º Ciclo e/ou Secundário. A tipologia TIPAU baseia-se em diversos fatores, começando pela densidade populacional de uma certa freguesia, produzindo três categorias de território: APU, área predominantemente urbana; AMU, área mediamente urbana; APR, área predominantemente rural. Comparou-se os dados da tipologias das áreas urbanas em que está inserida cada escola com os seus respectivos rácios aluno-professor (rácio A:P). Pretendia-se perceber se as áreas urbanas sofriam de números excessivos de alunos por cada professor, por terem maior densidade populacional.

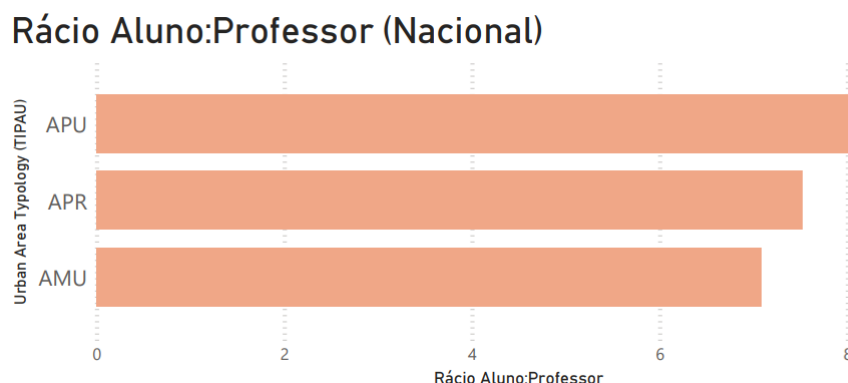


Figura 21: Rácio Aluno-Professor maior nas APU

Como era esperado, as áreas mais urbanizadas têm maior rácio A:P, mas as regiões rurais também acompanham esta tendência, enquanto que as zonas intermédias são as que beneficiam de maior número de professores por aluno, logo têm menor rácio A:P.

Com a ferramenta de análise de influência fornecida pelo Power BI, tentou-se descobrir quais as regiões (NUTS III) que maior impacto têm nestes resultados. Sem surpresa, as Áreas Metropolitanas de Lisboa e Porto têm maior influência, mas também o Alentejo Central, onde o rácio A:P das regiões rurais parece ser determinante para que estas sofram, nacionalmente, de um rácio superior às regiões mediamente urbanas.

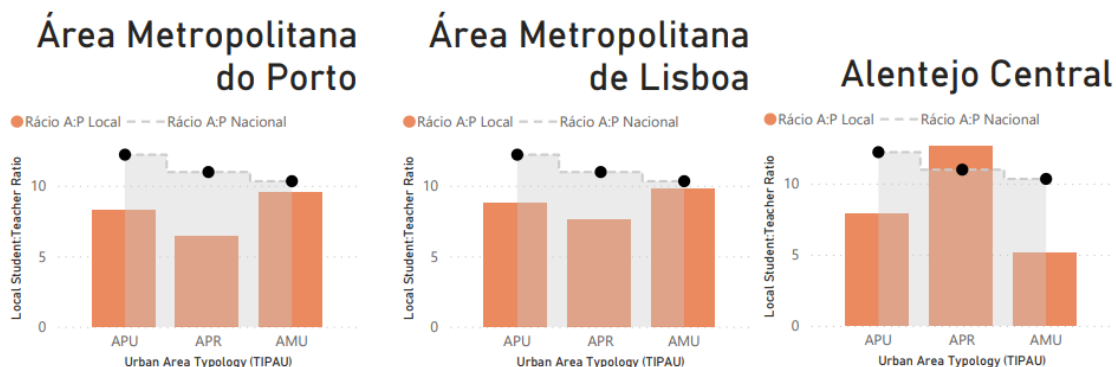


Figura 22: Comparadas as três regiões mais influentes com os valores nacionais

Como auxílio a estes gráficos, foram produzidos cartões e um mapa, interativos, finalizando este dashboard.

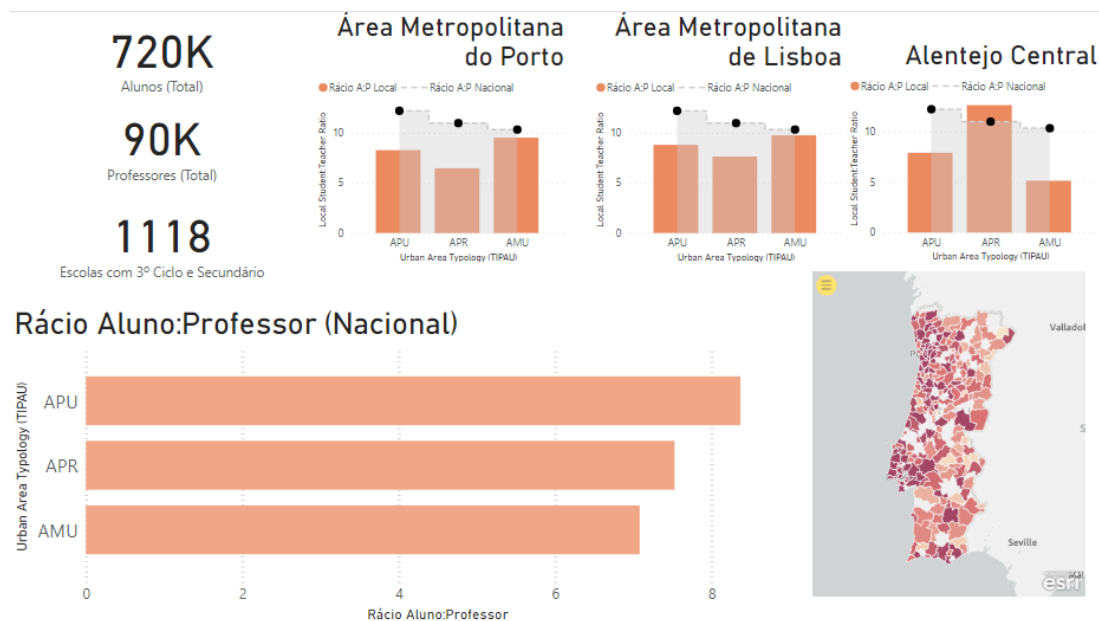


Figura 23: Dashboard Rácio A:P finalizado

Estudo do Envelhecimento do País e do Pessoal Docente

Foi encontrado, no sítio Pordata, o Índice de envelhecimento por município (número de idosos por cada jovem). Estes dados foram cruzados com as idades dos professores permanentes e contratados (escolas com 3º Ciclo e/ou Secundário). Para facilitar a visualização em dashboard foram apenas consideradas as médias por cada Região (NUTS III). Produziu-se o seguinte gráfico de combinação de linhas e barras.

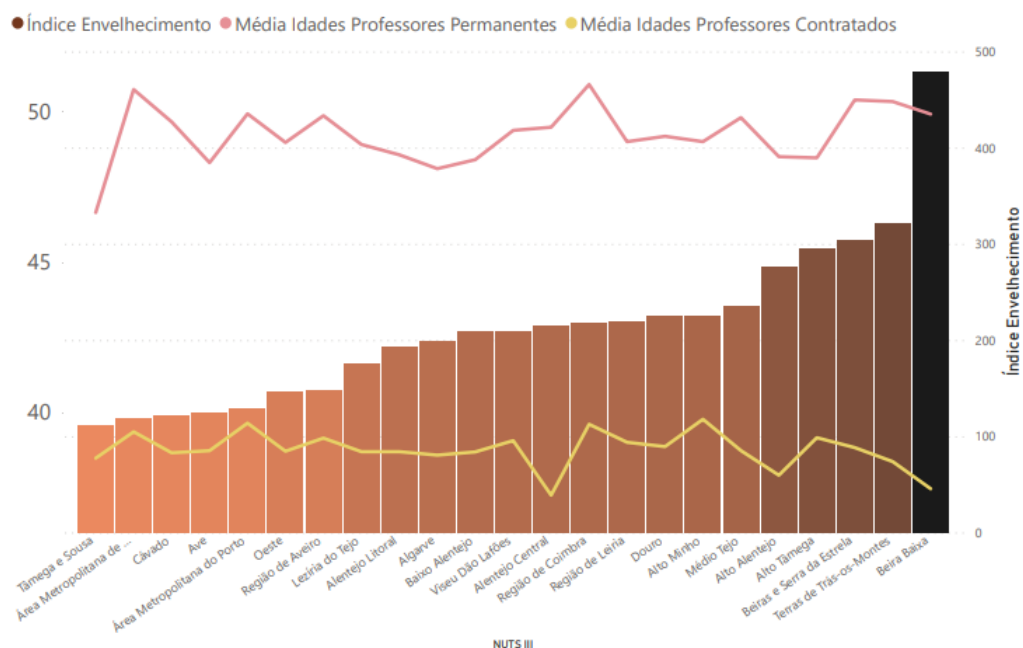


Figura 24: Barras com o Índice de Envelhecimento e as linhas com as idades do pessoal docente

As colunas foram coloridas, de acordo com o próprio Índice de Envelhecimento (IE), como auxílio visual. Percebe-se assim que as médias do IE são muito elevadas na Beira Baixa, mas as idades do pessoal docente não acompanham nem contrariam esta tendência. Não parece haver relação entre as componentes.

Foi produzido um mapa interativo para melhor perceber a relação geográfica destas regiões. O interior do país parece concentrar os IE mais elevados, de acordo com o esperado, no processo de desertificação.

Índice de Envelhecimento (Idosos por cada Jovem)

Um país envelhecido, especialmente no interior. Sem influência na distribuição de idades do pessoal docente.

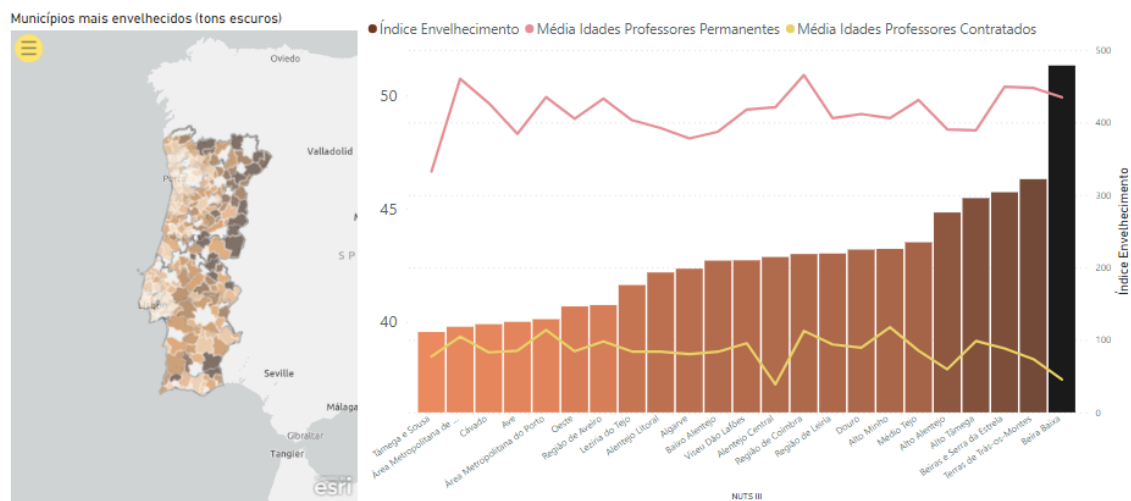


Figura 25: Dashboard Índice de Envelhecimento com Idade do Pessoal Docente

Estudo da Proporção de Professores Permanentes por Região

Tento saber-se como se relaciona a proporção de professores permanentes (em percentagem do total de professores), em cada região (NUTS III) e os respectivos municípios, se as maiores cidades concentram maior percentagem de professores do quadro nas suas zonas com maior densidade populacional ou não. Daqui foram isoladas várias Regiões de interesse, pela forma como diferem umas das outras.

● Professores Permanentes ● Professores Contratados

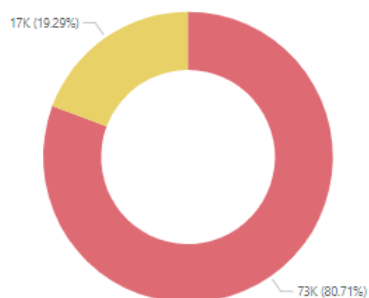


Figura 26: Gráfico de Rosca que permite saber a proporção e valor absoluto de professores do quadro a nível nacional

Para cada uma destas regiões e para os valores totais nacionais (NUTS III) foram produzidos bookmarks associados a botões que permitem uma melhor experiência para o utilizador, para que possa ver como cada região difere. As barras associadas a maiores centros urbanos (capitais ou cidades importantes) foram coloridas de forma contrastante, também como auxílio. Um mapa foi adicionado para que os valores de cada município sejam claros.

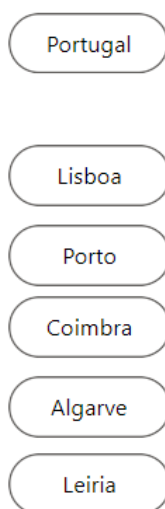


Figura 27: Botões de bookmark que permitem navegar entre regiões

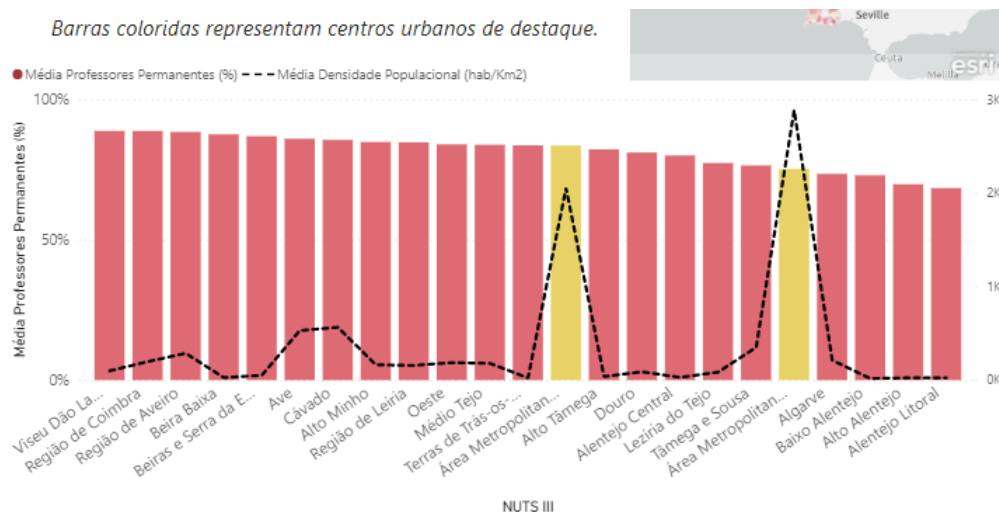


Figura 28: Gráfico barras e linhas entre proporção pessoal docente do quadro e densidade populacional, a nível nacional neste exemplo, com o botão "Portugal"selecionado

Assim, sabe-se que em regiões como a Área Metropolitana de Lisboa e do Porto, as zonas com maior densidade populacional tipicamente têm menos professores no quadro, sendo que em regiões como Região de Leiria e Algarve, a tendência é inversa.

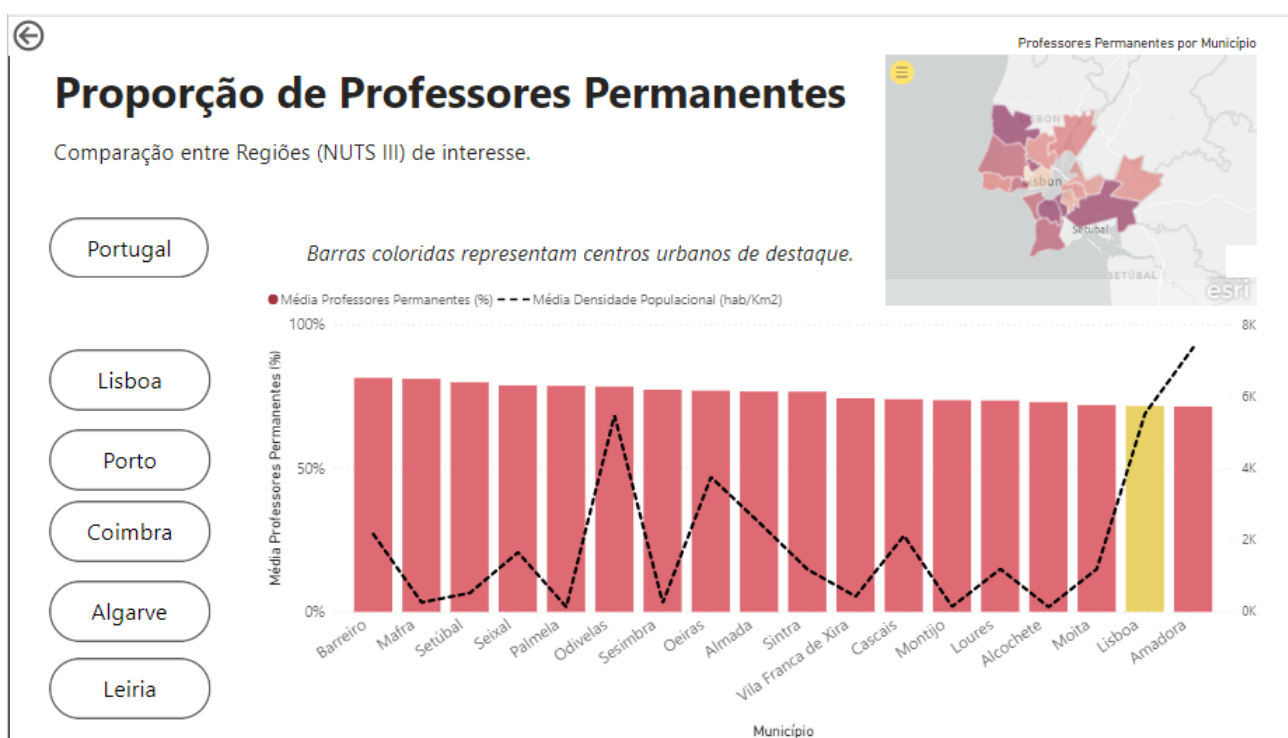


Figura 29: Dashboard final, com o botão "Lisboa"selecionado

Conclusão

Conclui-se este relatório classificando de 0 a 5 a ferramenta PowerBI em diferentes categorias:

Facilidade de utilização e de criação de Visualizações: 4. Neste ponto o powerBI é bastante *user friendly*, intuitivamente, através dos ícones apresentados não é difícil construir boas visualizações para

os dados.

Transformação dos dados: 3. Apesar de ser ótimo poder transformar os dados no mesmo software em que se criam as visualizações, houve algumas dificuldades no tratamento dos dados.

Facilidade em criar Dashboards: 4. A organização do dashboard por vezes poderia ser mais acessível.

Facilidade na partilha e interação de relatórios: 1, sem o PowerBI pro não foi possível partilhar os documentos e alterar os documentos uns dos outros.

Em conclusão, é uma ferramenta potente que quando bem utilizada facilita a produção de visualizações bonitas dos dados.

Dificuldades sentidas e Sugestões:

Para além de algumas dificuldades já referidas, as dificuldades sentidas foram essencialmente dificuldades de adaptação a um novo programa, mas que foram superadas em pouco tempo.

Achamos que uma ferramenta de trabalho em conjunto era essencial, a partilha tornou-se muito complicada ou mesmo impossível.

Outra sugestão é referente à visualização do mapa, alguns municípios ficaram em regiões erradas, uma forma de resolver isso seria poder arrastar os municípios que estão no sitio errado para o local certo.

Referências

- [1] PORDATA - Estatísticas, gráficos e indicadores de Municípios, Portugal e Europa. URL: <https://www.pordata.pt/>.
- [2] Portal do INE. URL: https://www.ine.pt/xportal/xmain?xpgid=ine_main&xpid=INE&xlang=pt.