



PROJETO DE VISUALIZAÇÃO E PREPARAÇÃO DE DADOS

Mestrado: Ciência de Dados e Engenharia

Unidade Curricular: Visualização e Preparação de Dados

Participantes:

- André Afonso - up2007002102@fe.up.pt
- Gabriel Copolecchia Carvalhal - up202103616@fe.up.pt

Existe oportunidade no meio do caos ?

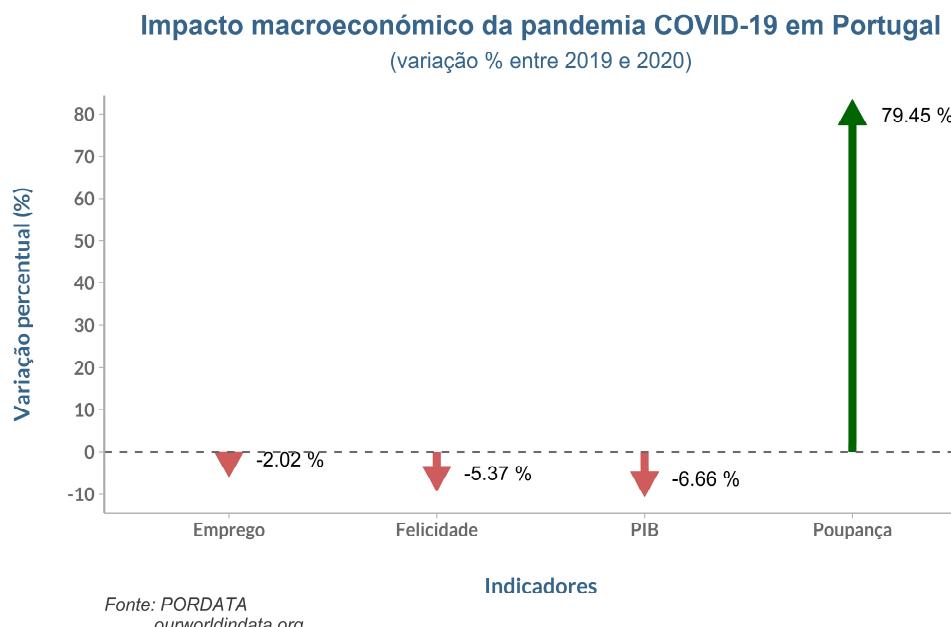
O mundo vive períodos atribulados. A volatilidade apodera-se dos mercados financeiros, e as pessoas sofrem as incertezas do mercado de trabalho, de cuidados de saúde e de bem-estar. Apesar disso, é importante manter um espírito positivo.

Neste contexto, e utilizando os mais diversos dados disponíveis em diferentes plataformas estatísticas, tentou-se encontrar oportunidades no meio do caos. Sendo este trabalho realizado por alunos do mestrado em ciência de dados, deseja-se perceber se, apesar da crise mundial que se instalou, as oportunidades de trabalho na área de tecnologia, em Portugal e na Europa, foram afetadas?

1. Qual o impacto macroeconómico da pandemia do Covid-19 em Portugal?

Antes de se conseguir entender as oportunidades no meio do caos, é preciso medir o impacto que a pandemia do covid-19 causou na vida das pessoas, em especial dos portugueses. Desta forma, optou-se por estudar a variação de indicadores económicos como o PIB (Produto Interno Bruto), o nível de felicidade, a quantidade de população empregada e o percentual que as famílias destinam à poupança.

Diante disso, utilizou-se três *datasets* provenientes da plataforma **PORDATA**, denominados "PIB.xlsx", "Pop_empregada.xlsx" e "Taxa_Poupanca_Familias.xlsx" que estão incluídos na pasta "*data*" deste projeto. Nos respectivos ficheiros o tratamento dos dados seguiu a mesma lógica, em que inicialmente leu-se os ficheiros em Excel, selecionando apenas as colunas de 'PT - Portugal' e 'Anos'. Em seguida, filtrou-se os anos de 2019 e 2020, adicionando-se uma coluna que especifica o indicador em análise e, por fim, converteu-se os dados de linha para coluna com a função "*spread*". O mesmo processo foi aplicado aos dados do indicador de felicidade, proveniente do website "*Our World in Data*". Por fim, as quatro tabelas foram consolidadas em um único *dataframe* (*Impacto*) e calculou-se a diferença percentual entre os dois anos para cada um dos indicadores. O gráfico resultante encontra-se abaixo:

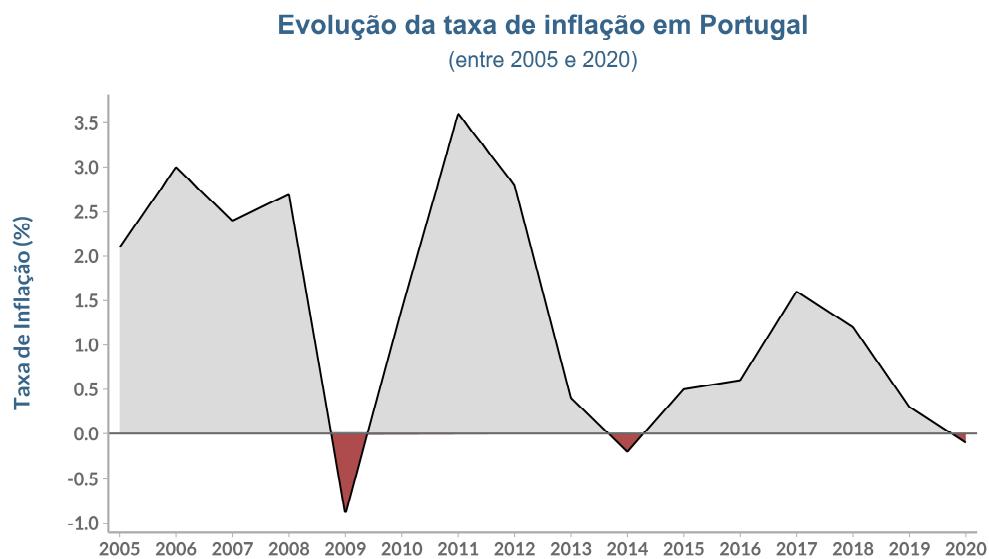


A partir da análise do gráfico, mede-se o impacto causado pela pandemia do covid-19, em que Portugal teve uma redução no número de pessoas empregadas, no índice de felicidade e no consumo de bens e serviços. Por outro lado, os portugueses passaram a destinar quase 80% a mais do seu orçamento à poupança. Neste cenário de queda de consumo e aumento de poupança, Portugal encontra-se em crise económica? Para responder esta pergunta, olhou-se para a evolução da taxa de inflação ao longo dos últimos anos.

2. Qual é a evolução da taxa de inflação em Portugal nos últimos anos?

Para responder a suposição, utilizou-se novamente um *dataset* da plataforma **PORDATA**, denominado "inflacao.xlsx". Neste documento, selecionou-se as colunas de 'Anos' e 'PT - Portugal', a partir do ano de 2005.

O desenvolvimento deste gráfico exigiu a criação de uma nova coluna, denominada 'status' com as categorias, "pos" e "neg" para cada taxa anual. Logo, permitiu-se colorir as áreas das taxas de inflação e deflação do gráfico abaixo.



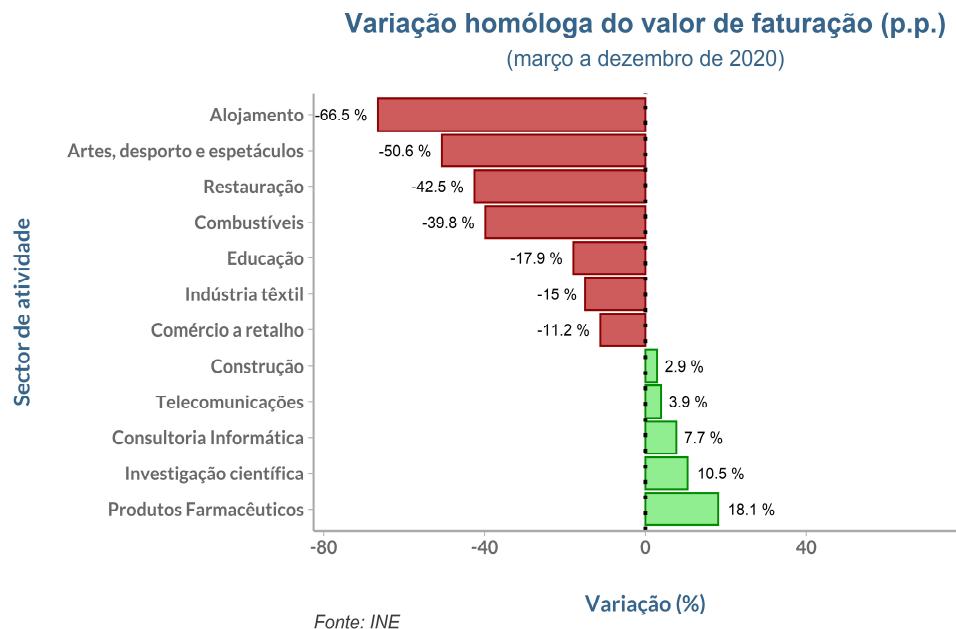
Fonte: PORDATA

O gráfico demonstra três momentos de deflação nos últimos quinze anos, em que os dois primeiros períodos ocorreram nos anos de 2009 e 2014, e que coincidem com recessões mundiais. O mesmo comportamento de deflação ocorreu em 2020, quando existe uma quebra no consumo, o que obrigou as empresas a reduzir seus preços para continuarem a vender. A causa deste fenômeno é uma determinada crise económica, que Portugal enfrentou e ainda lida com as consequências atualmente. Este fator acentuou-se ainda mais devido aos períodos de lockdown do ano de 2020, provocando uma nova queda de preços e causando uma retração da economia, vista na diminuição percentual do PIB. Entretanto, neste cenário de crise económica, todos os setores foram afetados da mesma forma?

3. Variação na faturação das empresas, por setor, no período da pandemia do Covid-19?

Neste contexto, estudou-se os dados partilhados pelo Instituto Nacional de Estatística (INE) com base na informação da plataforma e-fatura, da Autoridade Tributária e Aduaneira (AT), por meio do dataset “Destaque_E-fatura_PT.xlsx”. Em primeiro lugar, renomeou-se os setores em análise, a fim de obter uma visualização mais amigável. Subsequentemente, filtrou-se os setores de Indústria têxtil, Combustíveis, Produtos Farmacêuticos, Construção, Comércio a retalho, Alojamento, Restauração, Telecomunicação, Consultoria Informática, Investigação científica, Educação, Artes, desporto e espetáculos. Para finalizar a etapa de preparação dos dados, definiu-se uma nova coluna para categorizar valores positivos e negativos.

Optou-se por um gráfico de barras horizontais com o valor 0 centralizado, no intuito de simplificar a visão de faturação para cada segmento em Portugal entre os meses de março 2020 e dezembro 2020.



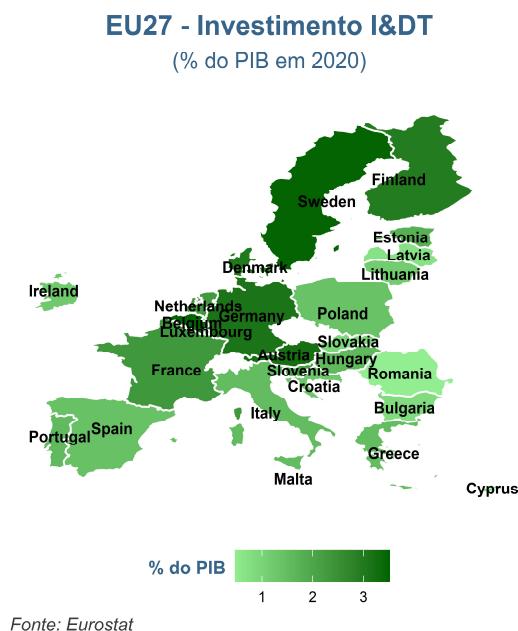
Diante disso, nota-se que no meio do caos social e económico, apenas os setores de Produtos Farmacêuticos, Investigação Científica, Consultoria Informática, Telecomunicações e Construção apresentaram aumento na faturação durante o primeiro ano de pandemia. Portanto, observam-se oportunidades nestes segmentos, que beneficiaram da atual conjuntura mundial.

Tendo em conta que os produtores deste estudo são atuais estudantes de engenharia e ciência de dados, decidiu-se detalhar as oportunidades de consultoria informática e investigação científica. Para isso, é importante perceber quais os países a nível europeu que atribuem uma maior parte dos seus recursos à investigação e ao desenvolvimento tecnológico.

4. Quais países da união europeia mais investem em investigação e desenvolvimento tecnológico?

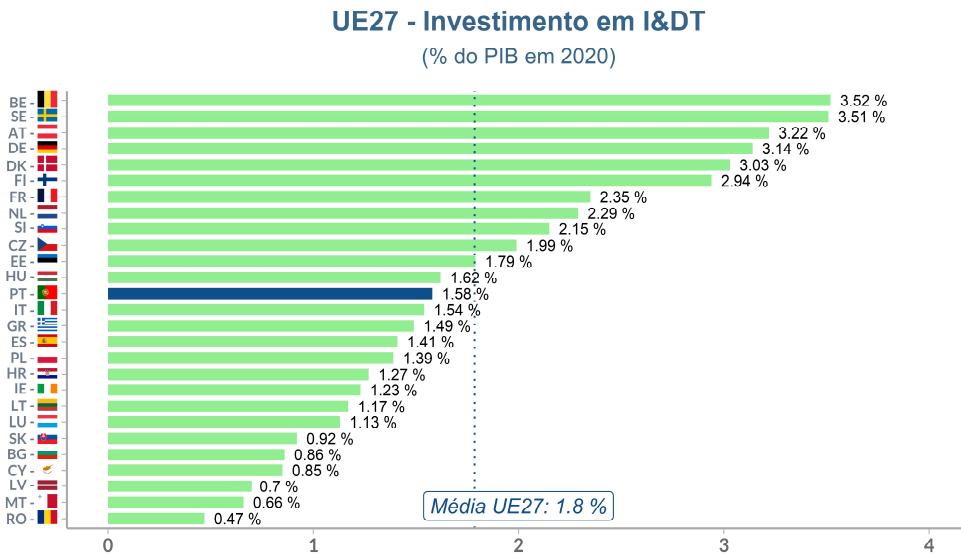
Identificou-se na plataforma Eurostat, o *dataset* “rd_gdp_eu27.xlsx”, que apresenta a percentagem em relação ao produto interno bruto (PIB) que cada país da união europeia investe em desenvolvimento tecnológico e investigação científica. A primeira opção de visualização proposta para este tipo de dados, foi a elaboração de um mapa em que os diferentes países fossem coloridos consoante o nível de investimento realizado.

Neste ficheiro, renomearam-se alguns países, a fim de se poder construir uma relação com a *library* “maps” do R. Finalmente, uniram-se os dois *dataframes* por forma a se obter os dados de latitude e longitude para cada um dos países da união europeia.



A partir de uma análise inicial, percebe-se que os países nórdicos destinam uma maior percentagem do PIB a investigação e desenvolvimento tecnológico (I&DT). Entretanto, a posição de Portugal não fica clara em relação ao resto dos países da união europeia. Logo, decidiu-se desenvolver um novo gráfico, a fim de facilitar a comparação entre os diversos países nestes setores.

Por meio do mesmo *dataset*, mas com a construção de um gráfico de barras, destacando os dados relativos a Portugal, identificou-se que o estado português destina menos do seu PIB a I&DT do que a média da união europeia. A análise do gráfico permite uma primeira conclusão semelhante à do mapa: os países do norte da europa são quem mais dedica parte do PIB ao investimento em investigação e desenvolvimento tecnológico. No entanto, com este gráfico de barras, é possível perceber que Portugal dedica uma maior percentagem do seu PIB ao investimento em ID&T do que países como Espanha, Itália ou Irlanda. No sentido contrário, países como Estónia, República Checa, ou Eslovénia, que sofreram grandes crises após o fim da União Soviética, investem a um ritmo superior do que a média europeia.



Fonte: Eurostat

No entanto, apesar de Portugal investir uma percentagem superior em I&DT do que, por exemplo Espanha, Itália, ou Irlanda, procurou-se perceber se este investimento se traduz num maior impacto em relação ao número de empregados e média salarial no segmento de informação, comunicação e tecnologia, no país.

5. Quais países europeus oferecem mais oportunidades e maiores salários aos profissionais do setor de tecnologia?

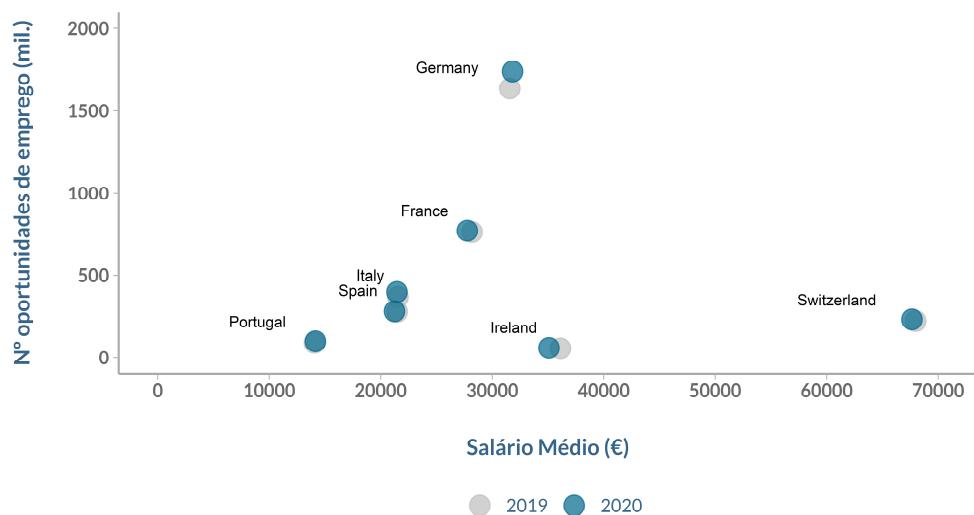
Novamente, recorreu-se à plataforma Eurostat para consultar dados relativos ao número de pessoas empregadas no setor tecnológico (*Human Resources in Science and Technology*), existente no ficheiro "hrst_st_ncat_page_spreadsheet.xlsx", e o valor do salário médio no país (*Net Earnings*), no documento "earn_nt_net_page_spreadsheet.xlsx".

O tratamento dado aos dois ficheiros foi semelhante: começar por selecionar apenas as colunas relativas ao nome do país, e aos valores do indicador para o ano de 2019 e 2020. Definimos ainda uma lista de países para ser possível efetuar uma comparação mais detalhada.

Ao compararmos com a figura anterior, em que se destaca o maior percentual português destinado ao I&DT do que por exemplo, Espanha, Itália ou Irlanda, percebe-se que este diferencial não se está a refletir em um maior número de oportunidades no setor em Portugal, apesar do salário médio ser bastante inferior, e possivelmente mais atrativo as empresas, quando comparado aos outros países. Portanto, Itália, Espanha e Irlanda alocam menos percentual do seu PIB no setor de investigação e tecnologia, mesmo assim oferecem condições mais atrativas aos profissionais.

UE27 - Oportunidades de emprego vs Salário em ITC

(2019 vs 2020)



Fonte: Eurostat

Atualmente, Portugal ainda tem um longo caminho a percorrer para poder ser um dos países mais atrativos aos profissionais do sector tecnológico. Apesar das condições de saúde, educação, e segurança pública estarem entre as melhores do mundo, os baixos salários continuam a ser o principal entrave de atração desta mão de obra especializada.

Independentemente de qual seja o melhor país para se trabalhar na área de investigação e tecnologia, todas estas nações citadas mantiveram ou aumentaram os postos de trabalho durante o primeiro ano de pandemia. Portanto, durante um período de turbulências, este setor aumentou a sua faturação, manteve postos de trabalho e proporcionou novas vagas, o que o caracteriza como uma boa oportunidade no meio do caos.

Datasets utilizados na elaboração deste relatório

1. PORDATA

PIB.xlsx

[https://www.pordata.pt/Portugal/Produto+Interno+Bruto+na+%c3%b3ptica+da+produ%c3%a7%c3%a3o+\(base+2016\)-2280-179832](https://www.pordata.pt/Portugal/Produto+Interno+Bruto+na+%c3%b3ptica+da+produ%c3%a7%c3%a3o+(base+2016)-2280-179832)

Taxa_Poupanca_Familias.xlsx

<https://www.pordata.pt/Portugal/Taxa+de+poupan%c3%a7a+das+fam%c3%adas-2340>

inflacao.xlsx

[https://www.pordata.pt/Europa/Taxa+de+Infla%c3%a7a%3a3o+\(Taxa+de+Varia%c3%a7a%3a3o+do+Pre%c3%a7o+Consumidor\)+total+e+por+tipo+de+bens+e+servi%c3%a7os-2656](https://www.pordata.pt/Europa/Taxa+de+Infla%c3%a7a%3a3o+(Taxa+de+Varia%c3%a7a%3a3o+do+Pre%c3%a7o+Consumidor)+total+e+por+tipo+de+bens+e+servi%c3%a7os-2656)

Pop_empregada.xlsx

<https://www.pordata.pt/Europa/Popula%c3%a7a%3a3o+empregada-1949>

2. INE

Destaque_E-fatura_PT.xlsx

https://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_indicadores&indOcorrCod=0008513&contexto=bd&selTab=tab2&xlang=PT

3. Eurostat

rd_gdp_eu27.xlsx

<https://ec.europa.eu/eurostat/web/science-technology-innovation/visualisations>

hrst_st_ncat_page_spreadsheet.xlsx

https://ec.europa.eu/eurostat/databrowser/view/hrst_st_ncat/default/table?lang=en

earn_nt_net_page_spreadsheet.xlsx

https://ec.europa.eu/eurostat/databrowser/view/earn_nt_net/default/table?lang=en

4. Our World in Data

felicidade.csv

<https://ourworldindata.org/happiness-and-life-satisfaction>