



UNIVERSIDADE
DE TRÁS-OS-MONTES
E ALTO DOURO

Escola de Ciências e Tecnologia – Departamento de Engenharias

Licenciatura em Engenharia Informática

Ano letivo 2021/2022

Introdução à Ciência dos Dados

Protocolo do Trabalho Experimental 1 – Enunciado B1

Trabalho Experimental 1 (conclusão até 29 de abril de 2022):

Utilize o seguinte Dataset de “**Energy (Archive)**” que pode ser descarregado em <https://github.com/owid/energy-data>. Neste link pode também encontrar informação detalhada sobre os diversos campos do Dataset.

Em função deste Dataset realize as seguintes tarefas em linguagem Python, que devem ficar devidamente documentadas no relatório deste Trabalho Experimental 1 a ser submetido no SIDE:

1. Carregue o ficheiro .csv para um DataFrame, e de seguida crie um novo DataFrame com apenas a informação dos países Estados Unidos, Canada, Brasil e México. Grave este novo DataFrame num novo ficheiro .csv.
2. A partir do novo DataFrame, faça um gráfico que lhe permita visualizar convenientemente a evolução da produção de eletricidade a partir do petróleo ('oil_electricity') nos referidos países.
3. Usando a biblioteca Matplotlib, crie um gráfico circular ('pie chart') que represente o número total de produção de eletricidade em função das diferentes fontes de produção ('biofuel_electricity', 'coal_electricity', 'fossil_electricity', 'gas_electricity', 'hydro_electricity', 'nuclear_electricity', 'oil_electricity') no ano 2010 para os Estados Unidos. Coloque as legendas adequadas.
4. Crie uma função que, dado o nome do país, apresente o ano em que houve mais consumo de energia nuclear ('nuclear_consumption'), bem como o respetivo valor.
5. Usando a biblioteca Seaborn, crie um gráfico de dispersão ('scatter plot') que permita visualizar o relacionamento entre o consumo de gás ('gas_consumption') e de energia nuclear ('nuclear_consumption'). Apresente também no gráfico uma regressão linear que relacione as duas variáveis. Explique convenientemente o seu significado.
6. Explore técnicas de Machine Learning que lhe permita fazer uma previsão de consumo de energia solar ('solar_consumption') no futuro. Documente bem as técnicas utilizadas e decisões tomadas.