

MODALIDADE:	Aprendizagem +	Não aplicável	
CURSO:	Técnico/a Especialista em Gestão de Informação e Ciência dos Dados		
UFCD:	Limpeza e transformação de dados em Python	CÓDIGO UFCD:	10808
FORMADOR/A:	Bruno Silva	DATA:	

Para este trabalho prático, vamos utilizar o ficheiro `nba.csv` (que deve estar na raiz da pasta do ficheiro onde está a trabalhar), onde o delimitador de valores das linhas é o símbolo `,` (vírgula).

Para trabalhar com a biblioteca Pandas, deve fazer a importação da biblioteca no início do ficheiro;

De seguida, vamos realizar os seguintes exercícios:

- **Exercício 1** – Crie a variável com o nome **data**, do qual vai ler o ficheiro com a função `read_csv()`, indicando a localização atual do ficheiro e ler o delimitador por vírgulas.
- **Exercício 2** – Mostre as informações gerais do ficheiro com a função `info()`;
- **Exercício 3** – Mostre os 10 primeiros registos do dataframe com a função `head()` e indique o número de registos a exibir;
- **Exercício 4** – Mostre os 5 últimos registos do dataframe com a função `tail()` e indique o número de registos a exibir;
- **Exercício 5** – Mostre o resumo dos dados do dataframe com a função `describe()`;
- **Exercício 6** – Mostre quantos registos existem no DataFrame com a função `len()`;
- **Exercício 7** – Mostre quantidade de linhas e colunas que existem no DataFrame com a função `shape()`;
- **Exercício 8** – Mostre o nome das colunas com a propriedade `columns`;
- **Exercício 9** – Altere o nome da coluna `Name` para `Nome` com a função `rename()`;
- **Exercício 10** – Mostre apenas o campo `College`;
- **Exercício 11** – Mostre as colunas `College` e `Team`;
- **Exercício 12** – Adicione nova coluna com o nome **Manager** com o valor “Sem info”;

- **Exercício 13** – Adicione nova coluna com o nome **Aumento_Salario** e calcule mais 10% de aumento em todos os salários (com base na coluna Salary);
- **Exercício 14** – Remova a coluna apenas a coluna Number com a função drop();
- **Exercício 15** – Remova as colunas Position e Age com a função drop();
- **Exercício 16** – Faça a filtragem do campo Weight para exibir valores maiores que 210;
- **Exercício 17** – Verifique se o ficheiro tem registos nulos com a função isnull() e faça a contagem com a função sum();
 - **Exercício 17.1** – Se tiver registos nulos deve:
 - Criar uma variável **dados_sem_nulos** e usar a função drop_na() para remover os dados nulos;
 - Volte a repetir o comando do exercício 20 e veja quantos registos existem na variável dados_sem_nulos;
- **Exercício 18** – Verifique se o ficheiro tem registos nulos com a função duplicated() e faça a contagem com a função sum();
 - **Exercício 18.1** – Se tiver registos duplicados deve:
 - Criar uma variável **dados_sem_nulos_duplicados** e usar a função drop_duplicates() para remover os dados duplicados;
 - Volte a repetir o comando do exercício 21 e veja quantos registos existem na variável dados_sem_nulos;
- **Exercício 19** – Grave a informação para um ficheiro excel (apenas com uma folha de cálculo) com o nome **“TP1_Simples.xlsx”**;
- **Exercício 20** – Grave a informação para um ficheiro excel, com o nome **“TP1_Combinado.xlsx”** e as seguintes folhas de cálculo:
 - **Criar** a folha de cálculo “Dados Originais” e coloque a informação da variável **data**;

- **Criar** a folha de cálculo “Dados sem Nulos” e coloque a informação da variável **dados_sem_nulos**;
- **Criar** a folha de cálculo “Dados sem erros” e coloque a informação da variável **dados_sem_nulos_duplicados**;