**04 - Operações com arrays**

Já conhecemos os arrays Numpy, aprendemos a criá-los de diferentes maneiras e testamos a sua performance em comparação às listas do Python. Agora aprenderemos a realizar operações aritméticas com esses arrays.

O conjunto de dados do nosso projeto possui informações de veículos, como ano de fabricação, quilometragem e valor. No início do curso fizemos um cálculo para descobrir a quilometragem média dos carros do conjunto, e para isso utilizamos a idade total e a quilometragem dos veículos.

Na seção **5.2 Operações aritméticas com arrays Numpy** do notebook deixamos duas células previamente preenchidas, uma delas com listas e outra com arrays Numpy, mas ambas com os mesmos valores. Começaremos trabalhando com as listas.

km = [44410., 5712., 37123., 0., 25757.]

anos = [2003, 1991, 1990, 2019, 2006]COPIAR CÓDIGO

Essas listas são ordenadas, ou seja, o primeiro valor de cada uma corresponde ao primeiro carro; o segundo, ao segundo carro; e assim sucessivamente. Queremos fazer um cálculo simples para descobrir a idade do veículo, algo que conseguiremos subtraindo o ano atual (2019 na gravação deste curso) pelo ano de fabricação de cada automóvel.

Criaremos entõa uma variável idade que receberá a subtração de 2019 pela lista anos.

idade = 2019 - anosCOPIAR CÓDIGO

A execução desse código nos retornará um erro, já que não é possível fazer a subtração de um inteiro por uma lista. Entretanto, isso é possível com os arrays Numpy.

km = np.array([44410., 5712., 37123., 0., 25757.])

anos = np.array([2003, 1991, 1990, 2019, 2006])COPIAR CÓDIGO

Ao subtrairmos 2019 pelo array anos, o retorno será outro array Numpy que poderá ser exibido na tela, este contendo as idades de cada um dos veículos.

idade = 2019 - anos

idadeCOPIAR CÓDIGO

array([16, 28, 29, 0, 13])

Também podemos realizar operações usando dois ou mais arrays. Por exemplo, se quisermos a quilometragem média (km\_media) podemos simplesmente dividir o array km pela variável idade que acabamos de criar.

km\_media = km / idadeCOPIAR CÓDIGO

Nesse caso teremos uma saída que indica um erro, já que nosso conjunto contém um valor 0, incorrendo em uma divisão por zero. Entretanto, o próprio Numpy preenche esse valor com nan (*not a number*) e continua a operação.

array([2775.625 , 204. , 1280.10344828, nan, 1981.30769231])

Perceba que não tivemos que utilizar nenhum tipo de laço ou criar outros arrays para realizarmos essas contas, que afinal são bem simples. Nesse momento estamos trabalhando apenas com uma dezena de dados, mas imagine se tivéssemos milhões? Levaria muito tempo para percorrermos todas as informações e realizarmos as operações usando laços, e isso custaria muito desempenho.

No vídeo anterior conhecemos os arrays de duas dimensões, e é possível realizar operações com eles. Para criarmos um array desse tipo, basta chamarmos a função np.array() passando como argumento uma lista contendo outros dois arrays. Nesse caso, criaremos o novo array dados usando como base os arrays km e anos que criamos anteriormente.

dados = np.array([km, anos])

dadosCOPIAR CÓDIGO

Como retorno, temos um array Numpy com duas linhas e 5 colunas (2, 5), que representando uma matriz.

array([[44410., 5712., 37123., 0., 25757.], [ 2003., 1991., 1990., 2019., 2006.]])

Isso é ilustrado pelo desenho abaixo, lembrando que a indexação se inicia no 0:



Agora queremos acessar esses dados de modo a calcularmos novamente a quilometragem média. Da mesma forma que fizemos com as listas, o índice 0 indica que estamos acessando o primeiro item - ou seja, o primeiro array/linha contendo 5 dados. Já com dados[1], acessaremos o segundo índice/linha. Ou seja, para obtermos km\_media, podemos simplesmente dividir dados[0], que representa a quilometragem, por 2019 - dados[1], que é a idade dos veículos.

km\_media = dados[0] / (2019 - dados[1])

km\_mediaCOPIAR CÓDIGO

Receberemos um aviso indicando que existe uma divisão por zero, mas a lista com os resultados será criada normalmente:

array([2775.625 , 204. , 1280.10344828, nan, 1981.30769231])

Aprenderemos mais sobre as seleções em arrays no próximo vídeo.