**04 - Criando DataFrames**

Vimos no último vídeo que é possível criar DataFrames e Series a partir de várias fontes (arquivos externos, listas, dicionários, etc).

Na aula que falamos sobre funções, desenvolvemos em nossos exercícios uma função que recebia um dicionário com um conjunto de informações sobre veículos, e calculava a quilometragem média anual de cada veículo. Esta função retornava o conteúdo do dicionário de *input* da função, incluindo as informações sobre a quilometragem média:

**In [1]:**

dados = {

'Crossfox': {'km': 35000, 'ano': 2005},

'DS5': {'km': 17000, 'ano': 2015},

'Fusca': {'km': 130000, 'ano': 1979},

'Jetta': {'km': 56000, 'ano': 2011},

'Passat': {'km': 62000, 'ano': 1999}

}COPIAR CÓDIGO

**In [2]:**

def km\_media(dataset, ano\_atual):

result = {}

for item in dataset.items():

media = item[1]['km'] / (ano\_atual - item[1]['ano'])

item[1].update({ 'km\_media': media })

result.update({ item[0]: item[1] })

return resultCOPIAR CÓDIGO

**In [3]:**

km\_media(dados, 2019)COPIAR CÓDIGO

**Out [3]:**

{'Crossfox': {'km': 35000, 'ano': 2005, 'km\_media': 2500.0},

'DS5': {'km': 17000, 'ano': 2015, 'km\_media': 4250.0},

'Fusca': {'km': 130000, 'ano': 1979, 'km\_media': 3250.0},

'Jetta': {'km': 56000, 'ano': 2011, 'km\_media': 7000.0},

'Passat': {'km': 62000, 'ano': 1999, 'km\_media': 3100.0}}COPIAR CÓDIGO

Assinale a alternativa que mostra a forma correta de se criar um DataFrame com o resultado obtido pela função acima. O DataFrame resultante deve ter a seguinte forma:

**Out [1]:**

|  | **km** | **ano** | **km\_media** |
| --- | --- | --- | --- |
| Crossfox | 35000.0 | 2005.0 | 2500.0 |
| DS5 | 17000.0 | 2015.0 | 4250.0 |
| Fusca | 130000.0 | 1979.0 | 3250.0 |
| Jetta | 56000.0 | 2011.0 | 7000.0 |
| Passat | 62000.0 | 1999.0 | 3100.0 |

Dica: Para resolver esta questão, será necessário relembrar de um recurso que aprendemos no curso anterior, quando falamos de arrays Numpy. Este recurso também pode ser aplicado a DataFrames do pandas:

***ndarray.T****: Retorna o array transposto, isto é, converte linhas em colunas e vice versa.*

import pandas as pd

carros = pd.DataFrame(km\_media(dados, 2019)).T

Alternativa correta! A propriedade T é uma forma de acessar o método transpose() do DataFrame.

 import pandas as pd

carros = pd.DataFrame(km\_media(dados, 2019))

Alternativa errada! Perfeito, só falta utilizar a dica do enunciado para obter a resposta correta.





import pandas as pd

carros = pd.DataFrame(km\_media(dados, 2019).T)

Alternativa errada! Dicionários não possuem o atributo T. Este atributo é do DataFrame.





import pandas as pd

carros = pd.Series(km\_media(dados, 2019)).T

Alternativa errada! O que o problema pede como resultado é um DataFrame e não uma Series.