**Python Para Data Science**

**Curso 1 – Primeiros Passos**

1. **Aula 1 – Função e Linguagem de Alto Nível:**
   1. Interpolação de str com f’{}’: Podemos fazer interpolação/concatenação entre str e texto em uma única linha usando essa técnica:

nome = 'Bruno'  
idade = 21  
print(f'O nome é {nome} e a idade é {idade} anos')

* + 1. Lembrando que podemos fazer essa formatação dessa forma:

nome = 'Bruno'  
idade = 21  
print('O nome é {} e a idade é {} anos'.format(nome, idade))

* + 1. O resultado final no console é o mesmo.
  1. Input(): Recebe uma entrada do usuário e o programa só continua sua execução após sua entrada:

def saudacao():  
 nome = input('Qual o seu nome?\nR: ')  
 idade = int(input('E a sua idade, qual é?\nR: '))  
 return print(f'Olá, {nome}! Você tem a mesma idade que eu: {idade} anos')  
  
saudacao()



* + 1. Lembrando que quando criamos uma função, precisamos executar ela no final do código para que rode, por isso o saudação() lá no fim.
  1. Linguagens de alto nível são aquelas que parecem muito com a linguagem humana. Diferente daquelas que parecem mais com as de máquinas.
     1. A principal vantagem é a facilidade de entendimento.
     2. A desvantagem é que o interpreter precisa traduzir o que escrevemos para linguagem de máquina, para que o computador execute.
     3. No caso das linguagens de baixo nível, elas já estão em linguagem de máquina, não necessitando do interpreter e tendo um desempenho melhor.
  2. O que aprendemos:
     1. Aprendemos o que são variáveis;
     2. Criamos nossa primeira função;
     3. Entendemos a diferença entre linguagens de alto e baixo nível.

1. **Aula 2 – Parâmetros, Condicionais e Conversão de Tipos:**
   1. Função com parâmetro: Passamos como parâmetro para função algo que ela tenha que usar dentro do seu código para executar, mas esse parâmetro é definido fora dela:

def saudacao\_com\_parametros(nome\_da\_pessoa):  
 return print(f'Bom dia, {nome\_da\_pessoa}!')  
  
nome = 'Luiza'  
saudacao\_com\_parametros(nome)

* 1. Podemos usar condições dentro de funções:

idade = 20  
  
def verifica\_se\_pode\_dirigir(idade\_usuario):  
 if idade\_usuario >= 18:  
 return print(f'Você pode dirigir!')  
 else:  
 raise ValueError('Você não tem idade para dirigir!')  
  
verifica\_se\_pode\_dirigir(idade)

* + 1. Optei por colocar um erro, mas poderia colocar somente uma mensagem negativa.
  1. Podemos converter tipos de str para int, int para str, float para str e assim por diante.
     1. Inputs sempre recebem valores em str, mesmo que o valor passado seja um número, caso estamos esperando receber um número e utilizar esse valor em uma verificação que precise de números e não str, podemos fazer isso:
        1. Código recebendo str:

def verifica\_se\_pode\_dirigir\_sem\_parametros():  
 idade = input('Qual a sua idade?\nR: ')  
 if idade >= 18:  
 return print(f'Você tem permissão dirigir!')  
 else:  
 raise ValueError('Você não tem idade para dirigir!')  
  
verifica\_se\_pode\_dirigir\_sem\_parametros()

* + - 1. Convertendo para int, part1:

def verifica\_se\_pode\_dirigir\_sem\_parametros():  
 idade = input('Qual a sua idade?\nR: ')  
 idade = int(idade)  
 if idade >= 18:  
 return print(f'Você tem permissão dirigir!')  
 else:  
 raise ValueError('Você não tem idade para dirigir!')  
  
verifica\_se\_pode\_dirigir\_sem\_parametros()

* + - 1. Convertendo para int, part2 e como prefiro fazer:

def verifica\_se\_pode\_dirigir\_sem\_parametros():  
 idade = int(input('Qual a sua idade?\nR: '))  
 if idade >= 18:  
 return print(f'Você tem permissão dirigir!')  
 else:  
 raise ValueError('Você não tem idade para dirigir!')  
  
verifica\_se\_pode\_dirigir\_sem\_parametros()

* + 1. A mesma lógica se aplica para converter em str e float. Basta colocar str(variável)/float(variável).
  1. Tipos de operadores:



* + 1. Apostila de python e orientação a objetos da caelum: <https://www.caelum.com.br/apostila-python-orientacao-objetos/declarando-e-usando-variaveis/#entrada-do-usurio>.
  1. O que aprendemos:
     1. Criamos uma função que recebe um parâmetro;
     2. Aprendemos a usar as condicionais if e else;
     3. Convertemos a string (str) digitando através do teclado em um valor inteiro (int).

1. **Aula 3 – Listas, Repetições e Tipos Booleanos:**
   1. Quando vemos o tipo das listas recebemos: “list”
   2. A contagem dos elementos de uma lista começa no 0, portanto, se vc quiser pegar o primeiro elemento de uma lista, print o índice 0, o segundo é o índice 1 e assim por diante:

idades = [18,20,22,35,47]  
  
print(idades)  
print(idades[0])  
print(idades[1])



* + 1. Podemos pedir para o índice printar do índice x até o y, fatiando nossa lista, mas lembre-se que o fatiamento é exclusivo, portanto, se pedirmos para printar do 0 até o 2, ele irá printar somente o 18 e 20, que estão nas posições 0 e 1, excluindo o 2:

idades = [18,20,22,35,47]  
  
print(idades)  
print(idades[0])  
print(idades[1])  
print(idades[0:2])



* + 1. Dessa forma, se quisermos printar até o terceiro elemento da lista, precisamos falar para ele printar de 0 até 3:

idades = [18,20,22,35,47]  
  
print(idades)  
print(idades[0])  
print(idades[1])  
print(idades[0:2])  
print(idades[0:3])



* + 1. Se quiser printar de um determinado elemento até o final, podemos passar somente o elemento antes do “:”, ignorando o próximo:

idades = [18,20,22,35,47]  
  
print(idades)  
print(idades[0])  
print(idades[1])  
print(idades[2:])



* + 1. Se quisermos pegar o último elemento de uma lista, basta colocar -1 no índice:

idades = [18,20,22,35,47]  
  
print(idades)  
print(idades[-1])



* + 1. A ordem regressiva continua, ou seja, se quiser o penúltimo item, coloque índice -2, -3 para o anti-penúltimo e assim por diante.
  1. Podemos criar loops e laços dentro de funções, facilitando nosso processo.
     1. Ao invés de criar uma função que recebe uma idade e verifica se a pessoa pode ou não dirigir para cada uma delas em um for, printando isso:

def verifica\_se\_pode\_dirigir(idade\_usuario):  
 if idade\_usuario >= 18:  
 print(f'{idade\_usuario} anos de idade, TEM permissão dirigir!')  
 else:  
 print(f'{idade\_usuario} anos de idade, NÃO tem idade para dirigir!')  
  
idades = [18,20,12,22,15,35,47,17]  
  
for idade in idades:  
 verifica\_se\_pode\_dirigir(idade)



* + 1. Podemos criar uma função que recebe uma lista de idade e já com esse for dentro fazendo a verificação para cada uma:

def verifica\_se\_pode\_dirigir\_loop\_(idades):  
 for idade in idades:  
 if idade >= 18:  
 print(f'{idade} anos de idade, TEM permissão dirigir!')  
 else:  
 print(f'{idade} anos de idade, NÃO tem idade para dirigir!')  
  
idades = [18,20,12,22,15,35,47,17]  
verifica\_se\_pode\_dirigir\_loop\_(idades)



* + 1. Diminuindo o código e deixando mais simples e prático.
  1. Podemos criar listas com valores booleanos:

permissoes = []  
idades = [18,14,22,15,47]  
  
def verifica\_se\_pode\_dirigir(idades, permissoes):  
 for idade in idades:  
 if idade >= 18:  
 permissoes.append(True)  
 else:  
 permissoes.append(False)  
 print(permissoes)  
  
verifica\_se\_pode\_dirigir(idades, permissoes)



* + 1. Utilizamos o append() para adicionar os valores na lista.
  1. O que aprendemos:
     1. Criamos nossa primeira lista (link com operações com listas: <https://www.alura.com.br/artigos/listas-no-python>) no Python;
     2. Aprendemos a utilizar a estrutura de repetição for;
     3. Descobrimos o tipo booleano.

1. **Aula 4 – O sistema de Imports:**
   1. Podemos colocar vários tipos numa mesma lista.
   2. Biblioteca Random: Randrange() aleatoriza um número de um começo a um fim. Exclusivo, portanto se quiser aleatorizar de 0 a 10, precisamos colocar até 11:

from random import randrange  
  
aleatorio = randrange(0, 11)  
print(aleatorio)

* 1. Podemos criar uma lista e ir acrescentando cada vez mais itens aleatórios dentro dela com um for e append:

from random import randrange  
  
notas\_matematicas = []  
  
for notas in range(8):  
 notas\_matematicas.append(randrange(0, 11))  
  
print(notas\_matematicas)



* 1. Essa biblioteca diz que gera números pseudo aleatórios pois se determinarmos uma seed para ela, a “aleatoriedade” dos números será sempre a mesma, deixando de ser igual.
     1. Ou seja, se escolhermos passar a seed 10 para nosso programa, os números do randrange que pedimos para ele gerar serão sempre os mesmos, portanto, aquelas notas matemáticas que colocamos seria os mesmos não importa quantas vezes rodarmos o programa.
  2. O que aprendemos:
     1. Vimos que é possível armazenar elementos de diferentes tipos em uma lista;
     2. Aprendemos como importar o método randrange para gerar números inteiros aleatórios entre 0 e 10.

1. **Aula 5 – Plotando com Matplotlib:**
   1. Importamos a matplotlib.pyplot como plt e aprendemos vários comandos:

x = list(range(1,9))

y = notas\_matematicas

plt.plot(x, y)

plt.title('Notas de matemática')

plt.xlabel('Provas')

plt.ylabel('Notas')

plt.show()

* + 1. Plt é a biblioteca, x e y são respectivamente linha e coluna.
    2. Plt.plot(x, y): basicamente manda fazer o gráfico com as informações das variáveis x e y, onde denominamos que x era uma lista que ia de 1 até 8 (exclusivo) pois é a quantidade de provas de matemática que tiveram, e y são as notas que o aluno tirou em todas elas, também composta por 8 valores com o código seguinte:

from random import randrange

notas\_matematicas = []

for notas in range(8):

  notas\_matematicas.append(randrange(0, 11))

* + 1. Como já tínhamos feito na aula anterior. Criamos uma lista das notas e fizemos um for para dar um append de um valor aleatório de 0 a 10, mimetizando as notas do aluno, em um range de 8 vezes.
    2. Após ter as informações dos valores de x e y, e mandar plotar o gráfico com elas, definimos que o título do gráfico seria ”Notas de Matemática", a etiqueta do eixo x “provas” e a etiqueta do eixo y “notas”, dando esse resultado:



* + 1. Para remover esse “Text” da parte superior usamos o .show(), dando esse resultado final:



* + 1. Podemos melhorar esse gráfico ainda mais, colocando umas bolinhas em cada ponto de cada nota, para fazer isso, incrementamos a linha do plot com mais um parâmetro:

x = list(range(1,9))

y = notas\_matematicas

plt.plot(x, y, marker='o')

# resto do Código…

* + 1. Nos levando a este resultado:



* 1. O que aprendemos:
     1. Aprendemos como importar a "Matplotlib" conforme a convenção como "plt";
     2. Vimos que é possível criar listas a partir do código "range";
     3. Aprendemos como criar gráficos, definindo título, rótulos e marcador, além de ocultar outros detalhes indesejáveis.

**Curso 2 – Introdução à Linguagem e Numpy:**

1. **Aula 1 – Ambiente do Cientista de Dados:**
   1. Todo o conteúdo dessa aula será feito no googlo colab. Link para acessar: <https://colab.research.google.com/drive/1vSPBCjnlEjD74x-vHtJ6Bfat4WyJ9dYy>.
   2. Estamos trabalhando com a biblioteca numpy.
      1. Importamos ela como np para simplificar o uso.
      2. Subimos arquivos txt com vários dados para o notebook para podermos utilizá-los e utilizamos a função loadtxt(‘nome\_do\_arquivo.txt’) da biblioteca, responsável por carregar um arquivo txt num Array numpy:

import numpy as np

km = np.loadtxt('carros-km.txt')

km



* + 1. No notebook não precisamos dar print(km). Só o fato de digitar o km no final da célula já faz a array ser impressa, como acima.
    2. Podemos passar um parâmetro na função loadtxt() dizendo o tipo queremos que essa Array seja:

anos = np.loadtxt('carros-anos.txt', dtype = int)

anos



* + 1. Como primeira operação, fizemos a km\_média de cada carro, basicamente dividindo o km pelo ano atual menos os anos de cada carro:

km\_media = km / (2019 - anos)



* + 1. Nesse caso acabei passando 2019 para ver o mesmo exemplo que foi dado na aula e ver essa mensagem.
    2. Na mensagem basicamente diz que não foi possível fazer o cálculo com alguns carros pois houveram carros que foram feitos em 2019, nesse caso, ele faz o tratamento de dados pra gente e devolve nan (not a number), já que 0 dividido por 0, não existe:



* + 1. Essa seria a km\_media de cada carro.
    2. O tipo de variável quando usamos o Array numpy é exatamente isso:



* + 1. Documentação da função loadtxt(): <https://numpy.org/doc/1.16/reference/generated/numpy.loadtxt.html>.
  1. O que aprendemos:
     1. Os ambientes de desenvolvimento para a linguagem Python;
     2. A carregar dados externos em arrays Numpy;
     3. A trabalhar de forma básica com arrays Numpy.

1. **Aula 2 – Características do Python:**
   1. Operadores matemáticos:
      1. Divisão com //: Devolve a parte inteira da divisão, como no caso de 10 // 3, devolverá apenas 3, ao invés do 3.333333... como em uma divisão normal.



* + 1. \*\*: operador de elevado, ou seja, 23 = 8. Esse é o operador que faz essa conta.



* + 1. %: devolve o resto de uma divisão, ou seja, 10%3 vai ser igual a 1, pois é o resto da divisão. Se fosse 10%2, o resto seria 0. Muito útil para verificação de números ímpares e pares.



* 1. Expressões matemáticas:
     1. Python entende expressões, identifica os operadores e executa a conta na ordem correta, independentemente da posição dos operadores:



* + 1. Para ter certeza que uma operação vai dar certo, podemos separar as contas:





* + 1. \_: Funciona como a memória da calculadora. É uma variável que guarda o último valor de uma operação matemática:





* 1. Formas de atribuição:
     1. =: atribuição direta;



* + 1. +=: soma o valor atual da variável mais o que vier depois



* + 1. -+: subtrai o valor atual da variável mais o que vier depois;



* 1. Atribuição múltipla:
     1. Podemos criar várias variáveis em uma única linha e atribuir os valores à elas respectivamente:

ano\_atual, ano\_fabricacao, km\_total = 2021, 2003, 44410.0



* + 1. Isso nos permite simplicar ainda mais um código. Ao invés de ficar assim:

ano\_atual = 2021

ano\_fabricacao = 2003

km\_total = 44410.0

km\_media = km\_total / (ano\_atual - ano\_fabricacao)

km\_media

* + 1. Ele fica assim:

ano\_atual, ano\_fabricacao, km\_total = 2021, 2003, 44410.0

km\_media = km\_total / (ano\_atual - ano\_fabricacao)

km\_media

* 1. Em str, ao colocar “ ” dentro de ‘ ’ python entende que é uma citação e as “” aparecem no texto final:



* + 1. O contrário também se aplica, se as de fora forem duplas as de dentro devem ser simples:



* + 1. Apas simples 3 vezes serve para escrever em várias linhas:



* + 1. \n: pula linha.
    2. None: indica que a variável é nula:



* + 1. Nem mostra nada quando excuta. Esse é o tipo dela:



* 1. Int(), float(), srt(): todos convertem uns aos outros para seu tipo. Se colocar uma variável dentro de () ela se torna respectivamente um inteiro, um float e uma string. O int() não arredonda, apenas pega a parte inteira, ignorando o pós virgula.