

Exercícios – 11/maio/2019

1. Acessar colab.research.google.com;
2. Se tiver conta no Google, entrar;
Quem não tiver, assinar para salvar seus trabalhos;
3. Trabalhar com o código abaixo:

```
from google.colab import files  
  
uploaded = files.upload()  
  
for fn in uploaded.keys():  
  
    print("C:\\Users\\julio.nievola\\Desktop\\CienciaDados11Maio2019".format(  
        name=fn, length=len(uploaded[fn])))
```

```
import pandas as pd  
  
dataset = pd.read_csv("iris.txt")  
  
dataset.describe()
```

4. Acessar <https://archive.ics.uci.edu/ml/datasets.php>;
5. Escolher três bases de dados e para cada uma delas realizar as seguintes ações:
 - a. Indicar o nome da base;
 - b. Que tipo de dado (ou problema) caracteriza esta base de dados;
 - c. Indicar quem forneceu esta base, caso haja indicações;
 - d. Quantos atributos (variáveis) possui;
 - e. Qual o tipo de tarefa a ser executada – classificação ou regressão;
 - f. Para cada uma das variáveis, indicar o que ela representa e seu tipo (nominal, real etc.);
 - g. Indicar se existem informações já publicadas sobre seu uso e, em caso positivo, quantas;
 - h. Quantos exemplos existem para cada classe, no caso de classificação, ou a faixa de valores da variável a ser prevista no caso de regressão.