



Faculdades Integradas de Bauru – FIB
Bacharelado em Ciências da Computação

Ciência de Dados I

Atividade Avaliada 1 – 1º Bimestre

Visualizando Dados

Objetivo

Consolidar os conceitos fundamentais de Ciência de Dados I desenvolvidos em sala de aula.

Metodologia

- Desenvolva seu trabalho **individualmente**.
- Enviar o arquivo .ipynb do notebook
- Postar o trabalho no Teams até às 23:59h do dia 14/03.
- Esta atividade terá peso 1,0 na média da disciplina.

Descrição

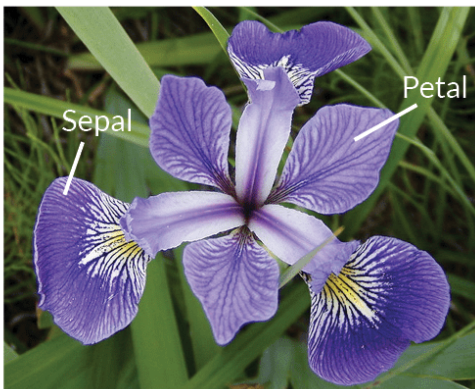
Para o desenvolvimento desta atividade vamos adotar o conhecido **dataset Iris** (<https://archive.ics.uci.edu/ml/datasets/iris>). Esse dataset consiste de 50 amostras de cada uma das seguintes três espécies de flores (Iris setosa, Iris virginica e Iris versicolor). Foram medidas em cada amostra, o comprimento e largura das sépalas e das pétalas (em cm).

O dataset está organizado da seguinte forma:

```
5.1,3.5,1.4,0.2,Iris-setosa  
7.0,3.2,4.7,1.4,Iris-versicolor  
5.8,2.7,5.1,1.9,Iris-virginica
```

Os valores separados por vírgula representam, respectivamente:

Comprimento da Sépala, Largura da Sépala, Comprimento da Pétala, Largura da Pétala, Tipo de flor



Iris Versicolor



Iris Setosa



Iris Virginica

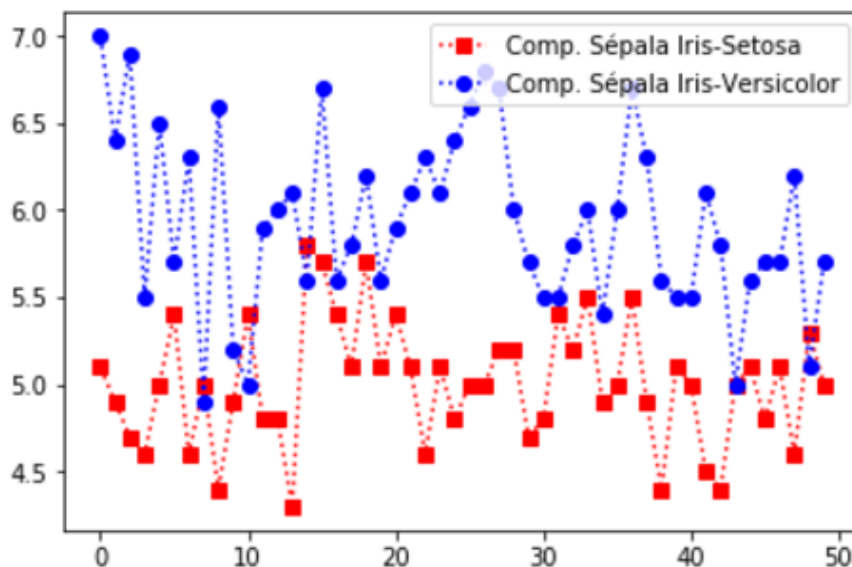
Considere as instruções a seguir para desenvolver o seu trabalho.

Instruções

- Baixe o dataset Iris do FIBOnline ou diretamente do site da Universidade da Califórnia em Irvine, endereço acima.
- Abra o Jupyter Notebook e crie um novo ambiente Python 3 (new -> Python 3)
- Instale o pacote Numpy em seu ambiente Python
 - `!pip install numpy`
- Importe a API Numpy
 - `import numpy as np`
- Carregue o dataset desconsiderando a coluna 5 sobre o tipo da flor
 - `data = np.genfromtxt('D:\Machine Learning\Datasets\iris.data.txt', delimiter=',', usecols=(0,1,2,3))`

Questões

1. Exiba (print) a primeira coluna (atributo) de todas as amostras
2. Faça o mesmo que acima considerando somente o tipo Versicolor
3. Considerando o que já estudamos, faça um gráfico como o abaixo:



4. Reconstrua o gráfico acima considerando as espécies Setosa e Virgínica
5. Faça um novo gráfico, semelhante ao acima, considerando os atributos comprimento da pétala entre as espécies Setosa e Virgínica.
6. Repita o gráfico acima para as espécies Versicolor e Virgínica.