

Roteiro Aula 2 Lab

1. Use numpy para criar uma matriz (dataset) composta por valores numéricos aleatórios com 200 linhas (instâncias) e 12 colunas (características). Apresente o dataset.
2. Converta o dataset do exercício 1 em um dataframe de pandas. As instâncias devem ser numeradas de 0 a 199 e as características nomeadas com letras de "a" a "l" (inclua "k" como coluna). Apresente o dataset no formato dataframe.
3. Mostre apenas as instâncias ímpares e as características "b", "e", "f" do dataset criado no exercício 2 ordenados crescentemente pelos valores da característica "f".
4. Apresente a soma dos valores, os valores mínimos e máximos, a média e o desvio padrão de cada coluna do dataset do exercício 3. Apresente também o intervalo de confiança de cada característica.
5. Faça um histograma com 7 faixas equidistantes de valores para a característica sepal_width da base iris.
6. Obtenha média, desvio padrão e intervalo de confiança de cada característica da base iris. Contraste os resultados com os obtidos fazendo um gráfico de barras com intervalo de confiança.
7. Faça um gráfico de dispersão apresentando os dados das características sepal_length e sepal_width diferenciando os exemplos de cada classe.
8. Faça um gráfico de caixas com dados de cada característica (coluna) do dataframe do exercício 2.
9. Apresente um gráfico com a projeção TSNE criado no exercício 1