

1ª Atividade – Apresentação em Sala de Aula.

1. Carregue o dataset do seu grupo
2. Quantas linhas e colunas tem seu dataset?
3. Plote o gráfico de dispersão com tamanho 12 x 8
4. Encontre todas medidas estatísticas do dataset (use describe())
5. Elimine a coluna: 'Unnamed: 0' do dataset
6. Verificar a existência de valores faltantes. Caso existam, completa-los com a **média** da respectiva coluna.
7. Verificar a existência de linhas (registros) duplicados e, caso existam, elimine-as.
8. Aplique o algoritmo K-Means ao dataset (Você deve decidir quantos cluster é a melhor escolha para seu dataset)
9. Obtenha todas as estatísticas descritivas para todos os clusters criados.
10. Mostrar coordenadas de todos os centroides criados.
11. Determine o valor da 'Silhueta' para a quantidade de cluster escolhida
12. Apresente e discuta os resultados com a classe.