### **1. Carregamento dos Dados**

* **Objetivo:** Carregar o dataset e entender suas variáveis.
* **Passo Executado:** Carregamos o dataset de rastreamento de membros de academia e visualizamos as primeiras linhas para entender a estrutura e os tipos de dados disponíveis.
* **Observação dos Dados:** O dataset inclui características demográficas (como idade e peso) e dados de desempenho (como BPM médio e porcentagem de gordura).

### **2. Preparação dos Dados**

* **Objetivo:** Selecionar variáveis relevantes para o agrupamento e padronizá-las para que estejam na mesma escala.
* **Passo Executado:** Selecionamos variáveis numéricas importantes (como idade, peso, altura, BPM médio, percentual de gordura e IMC) e aplicamos a padronização com StandardScaler, facilitando a performance dos algoritmos de agrupamento.

### **3. Definição e Teste dos Clusters**

* **Objetivo:** Testar diferentes números de clusters e calcular as métricas de validação.
* **Passo Executado:** Utilizamos o algoritmo **K-Means** com valores de KKK variando de 2 a 5 e registramos duas métricas de validação:
  + **Inércia (Soma das Distâncias Quadráticas):** Mede a coesão interna dos clusters, onde valores mais baixos indicam maior proximidade dos pontos aos centróides.
  + **Coeficiente de Silhueta:** Avalia a separação entre clusters, onde valores mais altos indicam uma maior distinção entre os grupos.

### **4. Identificação da Solução Ótima (Método do Cotovelo)**

* **Objetivo:** Determinar o número de clusters ideal utilizando a análise visual do gráfico de inércia.
* **Passo Executado:** Criamos um gráfico de inércia para diferentes valores de KKK de 1 a 10, buscando o ponto onde a curva começa a estabilizar (ponto de cotovelo), indicando que adicionar mais clusters não traz uma melhoria significativa na coesão dos grupos.

### **5. Conclusão sobre o Padrão de Clusters**

* **Padrão Identificado:** Com base no gráfico do método do cotovelo e nas métricas de inércia e silhueta, foi identificado que o valor de KKK entre 2 e 3 era um ponto estável, com uma boa separação entre clusters e uma coesão interna razoável.
* **Solução Final:** Optamos por K=3K = 3K=3, pois apresentou um equilíbrio entre baixa inércia e um coeficiente de silhueta ainda positivo, indicando que os clusters estavam bem definidos sem sobreposição excessiva.

### **Resumo da Conclusão**

A escolha de **3 clusters** foi considerada ideal devido ao equilíbrio entre coesão e separação dos grupos, refletindo características distintas entre os membros, como idade, condição física e intensidade do exercício, o que torna esses agrupamentos úteis para personalizar estratégias de treino ou saúde na academia.

Esse roteiro permite entender cada etapa do processo, desde o carregamento de dados até a escolha do modelo de agrupamento, focando em uma análise consistente para tomar decisões sobre a quantidade ideal de clusters.