Universidade de São Paulo ICMC/USP Bacharelado em Ciências de Dados Professor: Dr. Fabiano Rodrigues Coelho

1º Semestre de 2025Disciplina: 5500004

Instruções: Resolva os exercícios utilizando Google Colab (ou ferramentas similares) e envie a solução por e-mail em PDF. No PDF deve constar o nome do autor, bem como um link onde seja possível acessar o código on-line. O PDF deve ser o código executado, ou seja, imediatamente a cada bloco de código deve constar sua respectiva saída.

Trabalho 3: Data de Entrega 24/04/2025 às 23h59min

- 1. Implemente uma função chamada AED em Python, que receba uma lista de valores numéricos como parâmetro (Serie ou Data Frame). Em seguida, a função deve solicitar ao usuário que escolha o tipo de medida estatística que deseja calcular: medida de posição ou medida de dispersão. Após essa escolha, a função deve exibir uma lista de medidas disponíveis dentro da categoria selecionada e permitir que o usuário escolha uma delas. Por fim, a função deve calcular e exibir o resultado da medida escolhida.
- 2. Considere os dados disponíveis no site e-disciplinas do curso com o nome IDH e Gini, que tratam do IDH e do índice de Gini, das 27 Unidades da Federação do Brasil. Responda os itens
 - (a) Explique brevemente o que é e como calcular o IDH e o índice de Gini.
 - (b) Faça um histograma e um box-plot de cada uma das variáveis, descrevendo o comportamento das respectivas distribuições.
 - (c) Calcule as medidas resumo vistas em sala de aula e descreva o comportamento de cada uma delas.
 - (d) Calcule o coeficiente de correlação de Pearson entre o índice de Gini e o IDH. O que esse valor indica sobre a relação entre desigualdade de renda e desenvolvimento humano nas Unidades da Federação do Brasil?
 - (e) Identifique possíveis outliers em ambos os índices (Gini e IDH) utilizando o box-plot. Quais unidades da federação se destacam como outliers, e como isso pode ser interpretado?
 - (f) Identifique a unidade da federação com o maior IDH e a com o menor índice de Gini. Compare essas unidades da federação com outras da mesma região. O comportamento dessas UFs é consistente com o restante da região?
 - (g) Mapeie os índices de Gini e IDH em um mapa do Brasil, utilizando cores para representar diferentes intervalos de valores. O que você pode inferir sobre a distribuição geográfica dos índices de desenvolvimento e desigualdade?
 - (h) Agrupe as Unidades da Federação em quartis com base nos valores do IDH e compare as distribuições do índice de Gini para cada quartil. Existe alguma tendência observável à medida que o IDH aumenta?