Introdução à Ciência de Dados

usando Árvores

Enzo Benedetto Proença, 10418579

Gabriel Alves de Freitas Spinola Sucupira, 10418133

Lucas Pires de Camargo Sarai, 10418013

Vitor Alves Pereira, 10410862

Sumário

[Introdução 3](#_Toc182831436)

[Descrição do Dataset 4](#_Toc182831437)

[Modelagem nas Árvores 5](#_Toc182831438)

[Descrição da Implementação das Estruturas de Dados e as Operações Realizadas 6](#_Toc182831439)

[Perguntas Exploratórias 7](#_Toc182831440)

[Comparação de Desempenho das Árvores BST e AVL 8](#_Toc182831441)

[Conclusão 9](#_Toc182831442)

Introdução

**O que são?**

De acordo com a definição no site do Programa das Nações Unidas para o Desenvolvimento: “Os Objetivos de Desenvolvimento Sustentável (ODS), também conhecidos como Objetivos Globais, foram adotados pelos países-membros das Nações Unidas em 2015 como um apelo universal à ação para acabar com a pobreza, proteger o planeta e garantir que até 2030 todas as pessoas desfrutem de paz e prosperidade”.

Em suma, são objetivos variados definidos em reuniões entre os países membros para buscarem desenvolver e preservar o planeta.

**Objetivo escolhido**

Um desses objetivos é o 8: “Trabalho decente e crescimento econômico”. E nosso trabalho será focado nesse tópico, onde trabalhamos um banco de dados sobre a variação do PIB em todos os setores entre 2002 e 2024.

Descrição do Dataset

O *Dataset* “Tabelas PIB Mensal (08/24)” possui 5 planilhas, e optamos pela X: “”. Escolhemos essa por ser mais simples de se analisar, visto que não possui muitas variáveis envolvidas: trata-se do crescimento ou queda percentual do PIB do estado de São Paulo com base nos 3 principais setores econômicos: Agropecuária, Indústria e Serviços. Sendo assim, nossa análise nos permite aprofundar em um setor e analisarmos como fatos que ocorreram no período analisado afetaram a economia.

Esse *Dataset* está relacionado ao objetivo 8 pelo fato de o PIB ser usualmente o principal meio de análise de crescimento econômico em uma região, um estado ou em um país.

Modelagem nas Árvores

Em ambas as árvores, a chave que determina

Descrição da Implementação das Estruturas de Dados e as Operações Realizadas

Perguntas Exploratórias

• **Pergunta Exploratória 1**: *“Qual setor mais cresceu durante 2002 – 2010, em média? E qual teve o menor crescimento nesse mesmo período?*”

• **Pergunta Exploratória 2**: *“Qual setor mais cresceu durante 2010 – 20017, em média? E qual teve o menor crescimento nesse mesmo período?”*

• **Pergunta Exploratória 3**: *“Qual setor mais cresceu durante 2018 – 2019(pré-pandemia), em média? E qual teve o menor crescimento nesse mesmo período?”*

• **Pergunta Exploratória 4**: *“Qual setor mais cresceu durante 2020 – 2022(pandemia), em média? E qual teve o menor crescimento nesse mesmo período?”*

• **Pergunta Exploratória 5**: *“Qual setor mais cresceu durante 2023 – 2024(pós-pandemia), em média? E qual teve o menor crescimento nesse mesmo período?”*

Comparação de Desempenho das Árvores BST e AVL

Comparação do desempenho entre a BST e AVL, com gráficos e discussão sobre as diferenças observadas. Discutir sobre a eficácia de cada implementação de árvore para o tipo de análise realizada.

Conclusão

Conclusão do trabalho com uma discussão sobre as implicações dos resultados e as limitações do estudo.

Referências

Referências utilizadas