

# Análise Exploratória de Dados de Exploração de Floresta Plantada no Brasil

Isabelle Ramos de Azevedo Ferreira - 10365077

Universidade Presbiteriana Mackenzie (UPM)

**Resumo.** Este projeto realizou uma análise exploratória de dados sobre a exploração de floresta plantada no Brasil, utilizando um dataset contendo informações sobre detentores, imóveis rurais, empreendimentos, atividades de exploração, dados de inventário florestal e cronogramas de exploração. Através de técnicas de análise de dados, visualização e modelagem preditiva, foram identificados padrões, tendências e características importantes relacionadas à exploração de floresta plantada no país. Além disso, foi desenvolvido um modelo de regressão polinomial com regularização Ridge para prever o total de árvores cortadas nos próximos anos, fornecendo insights valiosos para o planejamento estratégico do setor.

## 1. Introdução

**Contextualização.** A exploração de floresta plantada desempenha um papel significativo na economia brasileira, fornecendo matéria-prima para diversos setores, como a indústria de papel e celulose, móveis e energia. Compreender a dinâmica e as características desse setor é fundamental para o planejamento estratégico, a tomada de decisões informadas e a formulação de políticas públicas adequadas.

**Justificativa.** A análise exploratória de dados sobre a exploração de floresta plantada no Brasil permite identificar padrões, tendências e insights relevantes que podem subsidiar ações e estratégias para o desenvolvimento sustentável do setor. Além disso, a aplicação de técnicas de modelagem preditiva possibilita a previsão do total de árvores cortadas nos próximos anos, fornecendo uma visão de longo prazo para o planejamento e a tomada de decisões.

**Objetivo.** O objetivo deste projeto é realizar uma análise exploratória abrangente dos dados de exploração de floresta plantada no Brasil, utilizando técnicas de análise de dados, visualização e modelagem preditiva para identificar padrões, tendências e características importantes do setor, bem como prever o total de árvores cortadas nos próximos anos.

**Opção do projeto.** A opção escolhida para este projeto é a análise exploratória de dados combinada com a modelagem preditiva, utilizando um dataset específico sobre exploração de floresta plantada no Brasil.

## **2. Descrição do Problema**

A exploração de floresta plantada no Brasil é uma atividade complexa que envolve diversos atores, como detentores de terras, empreendimentos, órgãos ambientais e profissionais da área. Compreender as características e a dinâmica desse setor é essencial para o planejamento e a tomada de decisões estratégicas. Além disso, a previsão do total de árvores cortadas nos próximos anos é fundamental para o planejamento de longo prazo e a sustentabilidade do setor. Este projeto busca aplicar técnicas de análise exploratória de dados e modelagem preditiva para extrair insights relevantes e fornecer previsões confiáveis.

## **3. Dataset**

O dataset utilizado neste projeto contém informações sobre a exploração de floresta plantada no Brasil, incluindo dados sobre detentores, imóveis rurais, empreendimentos, atividades de exploração, inventário florestal e cronogramas de exploração. Os dados foram obtidos de fontes confiáveis e foram anonimizados quando necessário para garantir a privacidade dos envolvidos.

Uma análise exploratória inicial foi realizada para compreender a estrutura e o conteúdo do dataset. Foram identificadas as colunas presentes, os tipos de dados e a presença de valores ausentes ou inconsistentes. Além disso, foram aplicadas técnicas de limpeza e preparação dos dados, como a remoção de duplicatas, o tratamento de valores ausentes e a formatação adequada das colunas.

O código Python utilizado para a análise exploratória, preparação dos dados e modelagem preditiva está disponível no notebook anexo a este relatório.

## **4. Metodologia**

A metodologia aplicada neste projeto consiste nas seguintes etapas:

**Coleta e preparação dos dados:** O dataset foi obtido de fontes confiáveis e passou por um processo de limpeza e preparação, conforme descrito anteriormente.

**Análise exploratória:** Utilizando técnicas de análise de dados e visualização, foram exploradas as características e os padrões presentes no dataset. Foram geradas estatísticas descritivas, gráficos e tabelas para obter insights relevantes sobre a exploração de floresta plantada no Brasil.

**Identificação de padrões e tendências:** Com base na análise exploratória, foram identificados padrões e tendências significativos no setor, como a distribuição geográfica dos empreendimentos, os principais detentores de terras, as espécies mais cultivadas e os volumes de exploração ao longo do tempo.

Modelagem preditiva: Foi desenvolvido um modelo de regressão polinomial com regularização Ridge para prever o total de árvores cortadas nos próximos anos. O modelo foi treinado utilizando os dados históricos disponíveis no dataset e avaliado quanto ao seu desempenho usando métricas apropriadas.

Interpretação dos resultados: Os resultados obtidos foram interpretados e discutidos, buscando compreender seu significado e implicações para o setor de exploração de floresta plantada no Brasil.

## 5. Resultados

Os resultados obtidos neste projeto incluem:

- Estatísticas descritivas sobre os detentores de terras, imóveis rurais, empreendimentos e atividades de exploração.
- Gráficos e visualizações que ilustram a distribuição geográfica dos empreendimentos, as principais espécies cultivadas e os volumes de exploração ao longo do tempo.
- Insights sobre padrões e tendências relevantes no setor, como a concentração de atividades em determinadas regiões, a predominância de certas espécies e a evolução dos volumes de exploração.
- Modelo preditivo de regressão polinomial com regularização Ridge para prever o total de árvores cortadas nos próximos anos, com os seguintes resultados:
  - Melhores parâmetros encontrados: {'poly\_\_degree': 2, 'ridge\_\_alpha': 1.0}
  - Erro Quadrático Médio no conjunto de teste: 24214309784.30
  - Abaixo está uma previsão realizada do total de árvores cortadas até 2050.
  - Ano    Total de Árvores
  - ----
  - 2021    23872.31
  - 2026    25534.56
  - 2031    27205.50
  - 2036    28885.13
  - 2041    30573.45
  - 2046    32270.47
  - 2050    33634.35

Todos os resultados detalhados estão disponíveis no notebook anexo e na organização pública no GitHub.

## 6. Conclusão

Este projeto realizou uma análise exploratória abrangente dos dados de exploração de floresta plantada no Brasil, utilizando técnicas de análise de dados, visualização e

modelagem preditiva. Os resultados obtidos fornecem insights valiosos sobre as características e a dinâmica do setor, contribuindo para uma melhor compreensão dos padrões e tendências existentes.

Além disso, o modelo preditivo desenvolvido demonstrou um bom desempenho na previsão do total de árvores cortadas nos próximos anos, fornecendo uma visão de longo prazo para o planejamento estratégico e a tomada de decisões no setor. As previsões indicam um crescimento contínuo no total de árvores cortadas até 2050, ressaltando a importância de práticas sustentáveis e do manejo adequado das florestas plantadas.

Os objetivos propostos foram alcançados, uma vez que foram identificados padrões relevantes, como a distribuição geográfica dos empreendimentos, as principais espécies cultivadas e a evolução dos volumes de exploração ao longo do tempo. Além disso, o modelo preditivo forneceu previsões confiáveis para o total de árvores cortadas nos próximos anos.

### **Endereço do Github**

SUBSTITUIR AQUI - Link para a organização pública no GitHub contendo os artefatos gerados, dataset e código-fonte }