

# Exploração de Dados (TJMA) — Preparação para Classe Usucapião (com fonte CNJ)

**Objetivo:** Documentar o notebook de preparação/exploração que identificou a classe processual USUCAPIÃO a partir do dataset geral, incluindo o passo a passo da obtenção do CSV no Painel Justiça em Números (CNJ).

## 1. Carregamento, Filtros e Amostra (resumo)

- Leitura de 'tjma-geral.csv' (cp1252; ';' on\_bad\_lines='skip') e normalização de colunas com utilitários de encoding/mojibake.
- Filtro da classe 'USUCAPIÃO' e exportação do recorte principal ('tjma-usucapiao.csv').
- Exploração por municípios, seleção de faixa balanceada (>10 e <120) e exportação da amostra ('tjma-usucapiao-amostra.csv').

## 2. Fonte dos Dados — Painel Justiça em Números (CNJ)

Utilizou-se o Painel Justiça em Números do CNJ, que consolida estatísticas oficiais do Poder Judiciário e oferece acesso a dados via API pública e painéis interativos. O Painel integra informações do DataJud desde 2020 e permite consultas detalhadas por município, classe e unidade judiciária.

### Passo a passo para obtenção do CSV (CNJ)

- Acessar o site do **Justiça em Números** (CNJ) e abrir o **Painel de Estatísticas**.
- Clicar em **Estatísticas** e aguardar o carregamento do dashboard (Power BI).
- No dashboard, acionar **Downloads** e selecionar **API Nova**.
- Na seção de filtro **Lista de Processos por Indicador**, configurar: Tribunal = **TJMA**; Indicador = **Julgados**; Grau = **1º Grau**; UF Município = **Todos**; Órgão Julgador = **Todos**.
- Efetuar o download/exportação dos dados e salvar como CSV para uso no notebook de exploração/preparação.

### Observações sobre o Painel/Api (CNJ)

- O Painel Justiça em Números é a principal fonte de estatísticas oficiais e disponibiliza uma **API pública** para download de bases com processos judiciais.
- Os dados são provenientes da **Base Nacional de Dados do Poder Judiciário (DataJud)**, instituída pela Resolução CNJ n. 331/2020, com atualização periódica.

## 3. Observações Finais

Recomenda-se registrar a **data/hora** da extração no relatório e versionar os recortes (CSV) com timestamp. Para integração ao pipeline (rotulagem/treino/fairness), mantenha chaves consistentes (nº do processo, município, vara/comarca).