

FIAP - FACULDADE DE INFORMÁTICA E ADMINISTRAÇÃO PAULISTA
DATA SCIENCE

Fábio Pereira de Lima – RM98803
Giovanna Cardoso Satorres – RM99944
Giullia Bianca Rocha Souza – RM552108
Gustavo Semenuk – RM550472
Mayara Prado Góes – RM98758

**Sprint 2: Entendimento dos Dados e Planejamento
de Análises para o Challenge MinSait**

Statistics for Data Science & Machine Learning
São Paulo
2023

Entendimento dos dados e planejamento de análises

” Levantar as fontes de dados preparados pela MINSAIT e gerar um relatório descrevendo quais dados são importantes e justificar o porquê.”

→ Matriz de Dados do Proagro

[Link Pesquisa ProAgro](#)

Formato: CSV, XLS, JSON

Essa fonte disponibiliza a quantidade e valor das Adesões e Sinistros por região, UF, produto, segmento, tipo de pessoa, idade e tipo de beneficiários. Aqui temos um histórico de valores do qual podemos avaliar as chances dos produtores rurais na obtenção de crédito, baseado nos dados também podemos indicar quais as atividades ao produtor rural ou cooperativa estão mais aptas a obter o crédito e assim esses possam concentrar esforços em atividades mais produtivas em sua região.

Análise preditiva: com esse histórico e cruzamentos das informações do beneficiário também poderemos gerar um gráfico com a projeção de adesão e sinistro específico por cultura do qual o produtor poderá avaliar quais as atividades com maior potencial de ganhos em sua região.

→ Conab - Custos de Produção

[Link Pesquisa Conab](#)

Formato: XLS

O CONAB, Companhia Nacional de Abastecimento, disponibiliza em seu portal uma grande quantidade de informações relacionadas a safra das principais atividades agrícolas, pecuárias e extrativista no Brasil, no portal é possível realizar o download de planilhas com as séries históricas de produção e detalhes como despesas de custeio, despesas pós-colheita, despesas financeiras, depreciação, custos fixos e demais variáveis envolvendo o setor agrícola.

Essa fonte de dados será de grande importância para entendermos toda a cadeia de produção das culturas escolhidas em nosso escopo: cana-de-açúcar, batata-doce, açaí, guaraná e mandioca. Com essa base de informação também poderemos aplicar os conhecimentos adquiridos na FIAP para gerar modelos de Análise de Dados e Machine Learning que sejam úteis aos produtores e assim possam organizar e otimizar seus trabalhos.