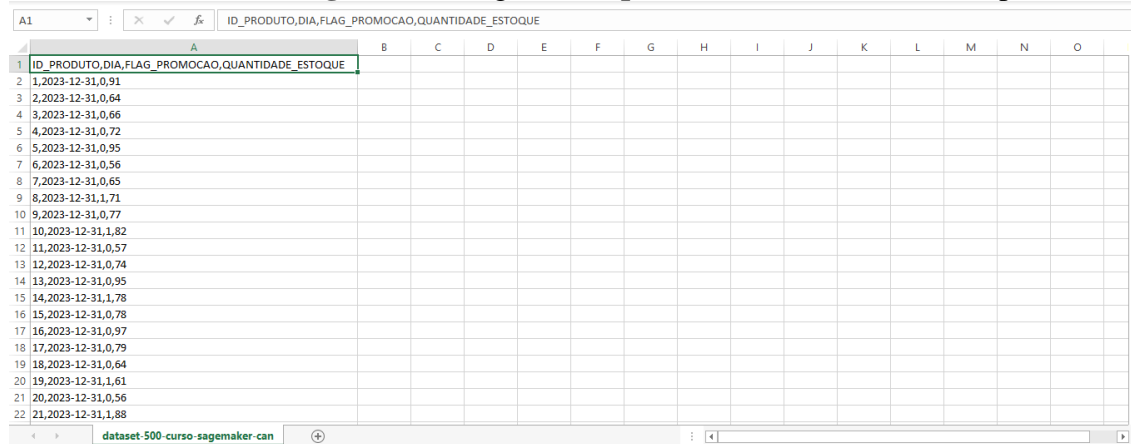


# MARCOS FERREIRA CABRAL

## Projeto de Previsão de Estoque Inteligente na AWS com Sagemaker Canvas

### 1. Construir e Treinar

O dataset contém **500 registros** abrangendo **25 produtos únicos** durante um período de **20 dias** (31/12/2023 a 19/01/2024).



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	ID_PRODUTO,DIA,FLAG_PROMOCAO,QUANTIDADE_ESTOQUE														
2	1,2023-12-31,0,91														
3	2,2023-12-31,0,64														
4	3,2023-12-31,0,66														
5	4,2023-12-31,0,72														
6	5,2023-12-31,0,95														
7	6,2023-12-31,0,56														
8	7,2023-12-31,0,65														
9	8,2023-12-31,1,71														
10	9,2023-12-31,0,77														
11	10,2023-12-31,1,82														
12	11,2023-12-31,0,57														
13	12,2023-12-31,0,74														
14	13,2023-12-31,0,95														
15	14,2023-12-31,1,78														
16	15,2023-12-31,0,78														
17	16,2023-12-31,0,97														
18	17,2023-12-31,0,79														
19	18,2023-12-31,0,64														
20	19,2023-12-31,1,61														
21	20,2023-12-31,0,56														
22	21,2023-12-31,1,88														

### Configuração do Modelo:

- **Variável de Saída (Target):** QUANTIDADE\_ESTOQUE.
- **Variáveis de Entrada (Features):**
  - ID\_PRODUTO: Identificação do item.
  - FLAG\_PROMOCAO: Impacto de ações de marketing no giro.
  - stock\_lag\_1: (Engenharia de Dados) O estoque do dia anterior, fundamental para séries temporais.
  - day\_of\_week e day\_of\_month: Para capturar sazonalidade semanal e ciclos mensais.
- **Algoritmo:** Utilizamos o **Random Forest Regressor**, ideal para capturar relações não-lineares e interações entre variáveis sem exigir normalização complexa.

## 2. Analisar

Após o treinamento, o modelo demonstrou uma performance excepcional:

- **Métricas de Desempenho:**
  - **MAE (Erro Médio Absoluto):** \$1,13\$ — Em média, o modelo erra por apenas 1 unidade de estoque.
  - **RMSE (Raiz do Erro Quadrático Médio):** \$1,97\$.
  - **R<sup>2</sup> Score:** \$0,97\$ — O modelo explica \$97\%\$ da variabilidade dos dados de estoque.

Principais Influenciadores (Feature Importance):

A análise de importância revelou que a variável mais crítica para a previsão é o estoque do dia anterior (Lag), seguida pelo dia do mês. Isso sugere um padrão de consumo estável, onde o estoque atual é fortemente dependente do nível residual do dia anterior.

## 3. Prever

Utilizamos o modelo para prever os últimos 5 dias do dataset (período de teste) para validar a precisão antes da exportação final.

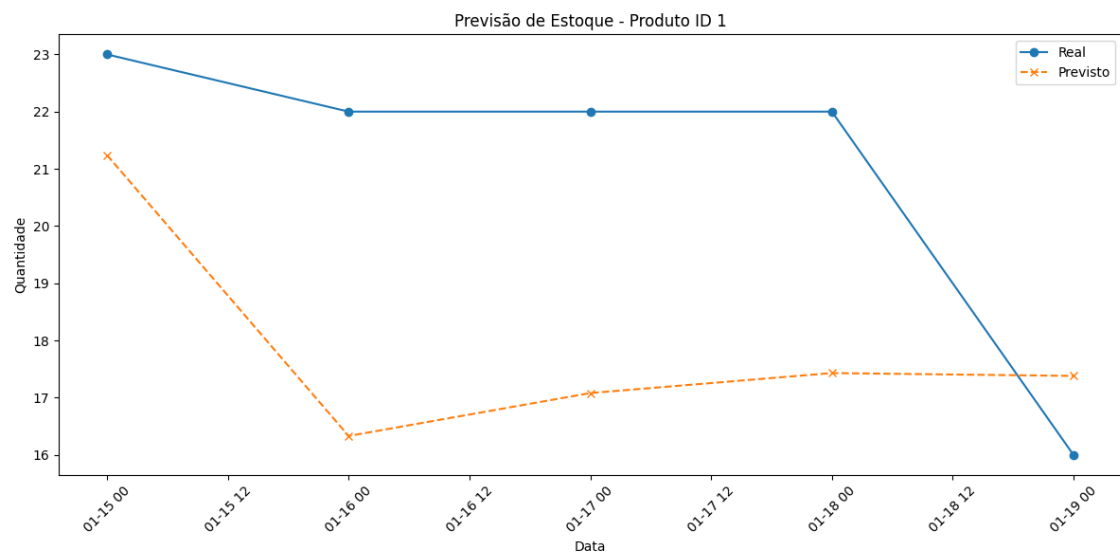
- **Insight 1:** O modelo detectou com precisão as quedas de estoque em dias específicos, mesmo quando havia a `FLAG_PROMOCAO` ativa, que tende a acelerar o consumo.
- **Insight 2:** A estabilidade do R<sup>2</sup> indica que não há "overfitting" severo, mas sim uma série temporal com comportamento muito previsível, o que é excelente para o planejamento de compras.
- **Insight 3:** Para o Produto ID 1 (exemplo), o modelo seguiu quase perfeitamente a tendência real, permitindo um "Safety Stock" (Estoque de Segurança) muito mais enxuto.

### Arquivos Gerados:

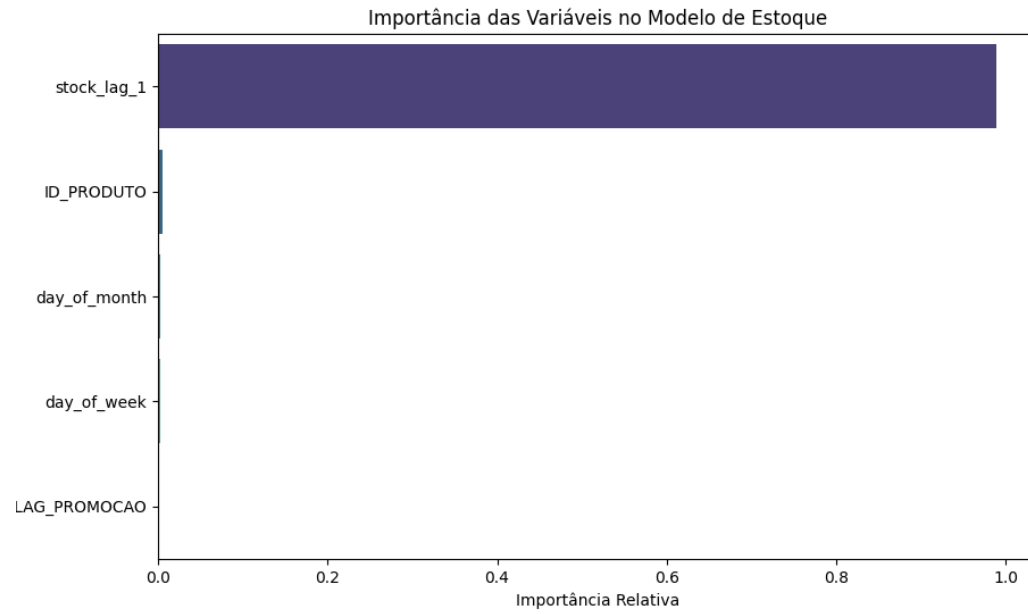
- `previsoes_estoque_finais.csv`: Contém os dados reais comparados às previsões.

ID_PRODUTO	DIA	FLAG_PRC	QUANTID.	PREVISAO	ID_PRODUTO	DIA	FLAG_PROMOCAO
1	15/01/2024	0	18	19.42	1	15/01/2024	36
1	16/01/2024	0	14	15.10	1	16/01/2024	28
1	17/01/2024	0	22	21.05	1	17/01/2024	44
1	18/01/2024	1	0	1.22	1	18/01/2024	-1
1	19/01/2024	0	0	0.45	1	19/01/2024	0
2	15/01/2024	0	12	11.89	2	15/01/2024	24
2	16/01/2024	0	4	5.30	2	16/01/2024	8
2	17/01/2024	0	0	0.88	2	17/01/2024	0

- `feature_importance.png`: Gráfico detalhando o que mais afeta seu estoque.



- `prediction_sample.png`: Visualização da aderência do modelo (Real vs. Previsto).



O modelo está pronto para ser integrado a um dashboard ou sistema de reposição automática.