

É possível prever
a demanda de
produtos
baseado em
dados
históricos?

Demand Forecasting

Forecasts for Product Demand - Make
Accurate Forecasts for Different
Products

Alunos:

Gilberto Filho

Wanderson Torres



AGENDA



O PROBLEMA

01



O DATASET

02



ANÁLISE EXPLORATÓRIA DOS DADOS

03



MODELOS ESTATÍSTICOS

04



**MODELOS DE MACHINE
LEARNING**

05



RESULTADOS

06



CONCLUSÃO

07

O Problema



PREVISÃO DE DEMANDA COM BASE EM DADOS HISTÓRICOS

A empresa está em período de reformulação de estratégias e de práticas de como fazer a previsão de demanda de seus produtos. Atualmente, os dados históricos do ano anterior servem como previsão de demanda para o ano seguinte..

Melhorar o processo de
previsão de demanda, pois o
método atual é ultrapassado
e gera perdas para a empresa.

A OPORTUNIDADE

O Dataset

Características:

- **Frequência de Demanda:** Diária;
- **Séries temporais intermitentes:** Sim, porém sem inclusão de zeros;
- **Informações de hierarquias de negócio** (código do produto, categorias de produto, centros de distribuição);
- **Variáveis exógenas:** Nenhuma.

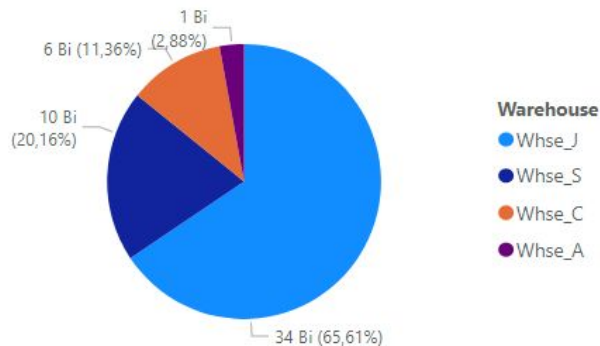


→ Tamanho : 1.048.575 linhas x 5 colunas

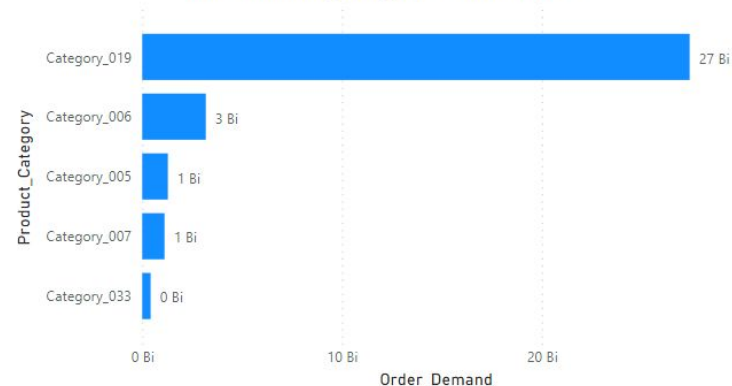
→ Categoria de Produtos: 33
Produtos: 2160
Centros de Distribuição: 4

Análise Exploratória dos Dados (EDA)

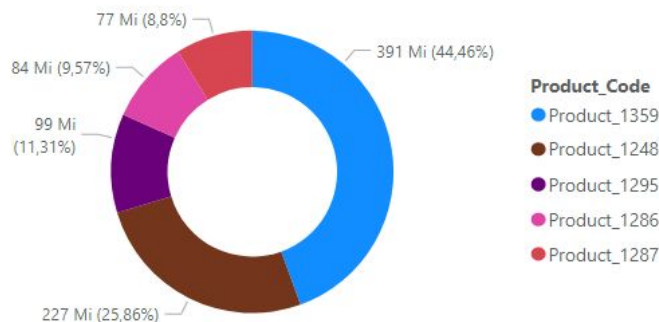
Order_Demand por Warehouse



TOP 5 Product_Category - Warehouse J



Top 5 Product_Code



Análise da Série Temporal (Produto 1359)

Gráfico Mensal

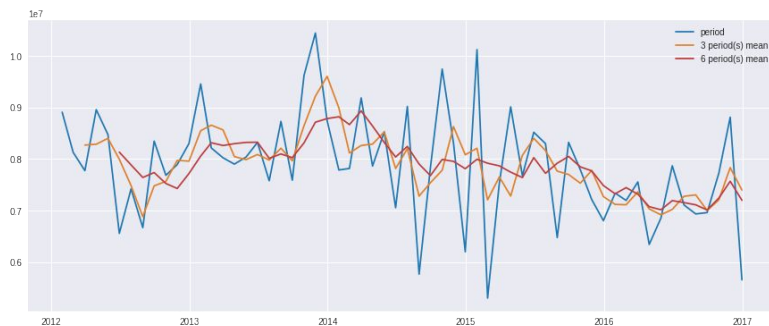


Gráfico Semanal

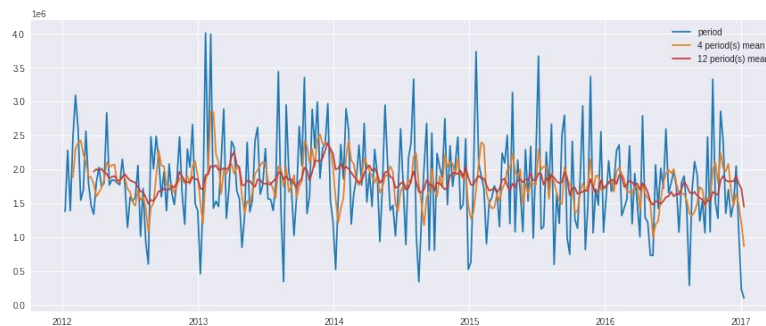
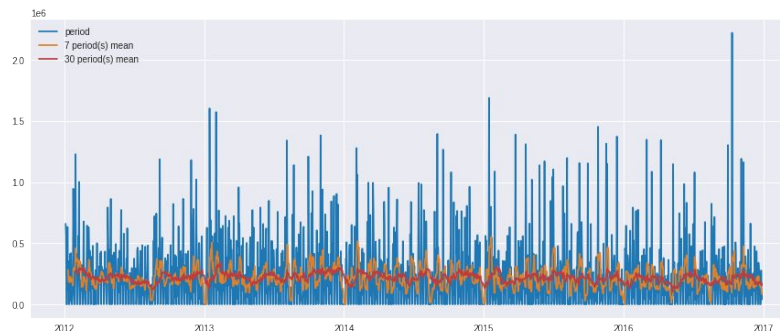


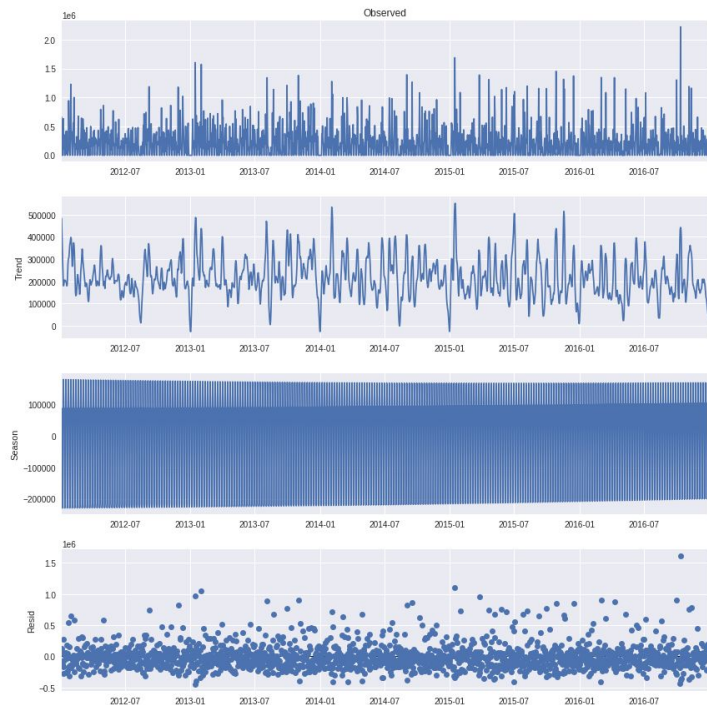
Gráfico Diário



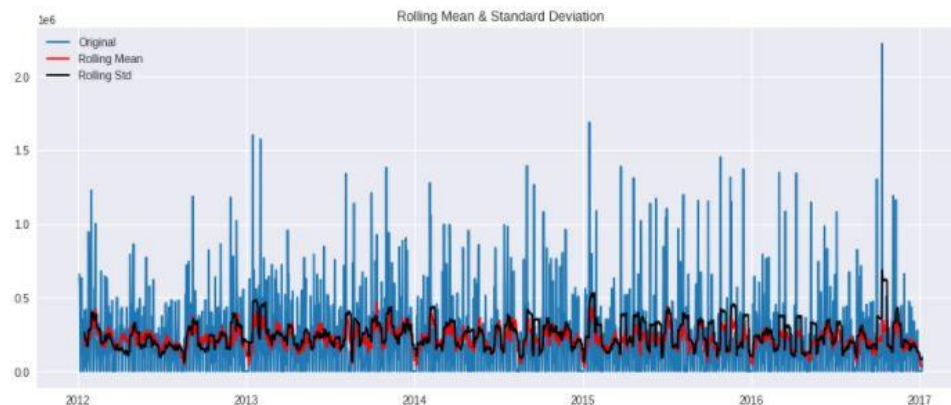
Análise da Série Temporal (Produto 1359)

Decomposição da Série Temporal (Diário)

Product 1359 - Warehouse J - Seasonal Decomposition



Teste de Estacionariedade (Diário)



Results of Dickey-Fuller Test:
 Test Statistic -1.040075e+01
 p-value 1.912366e-18
 #Lags Used 2.500000e+01
 Number of Observations Used 1.803000e+03
 Critical Value (1%) -3.433982e+00
 Critical Value (5%) -2.863144e+00
 Critical Value (10%) -2.567624e+00
 Strong evidence against the null hypothesis
 Reject the null hypothesis
 Data has no unit root and is stationary

Sazonalidade não identificada
 pela decomposição da série
 temporal do Produto 1359

Base Model

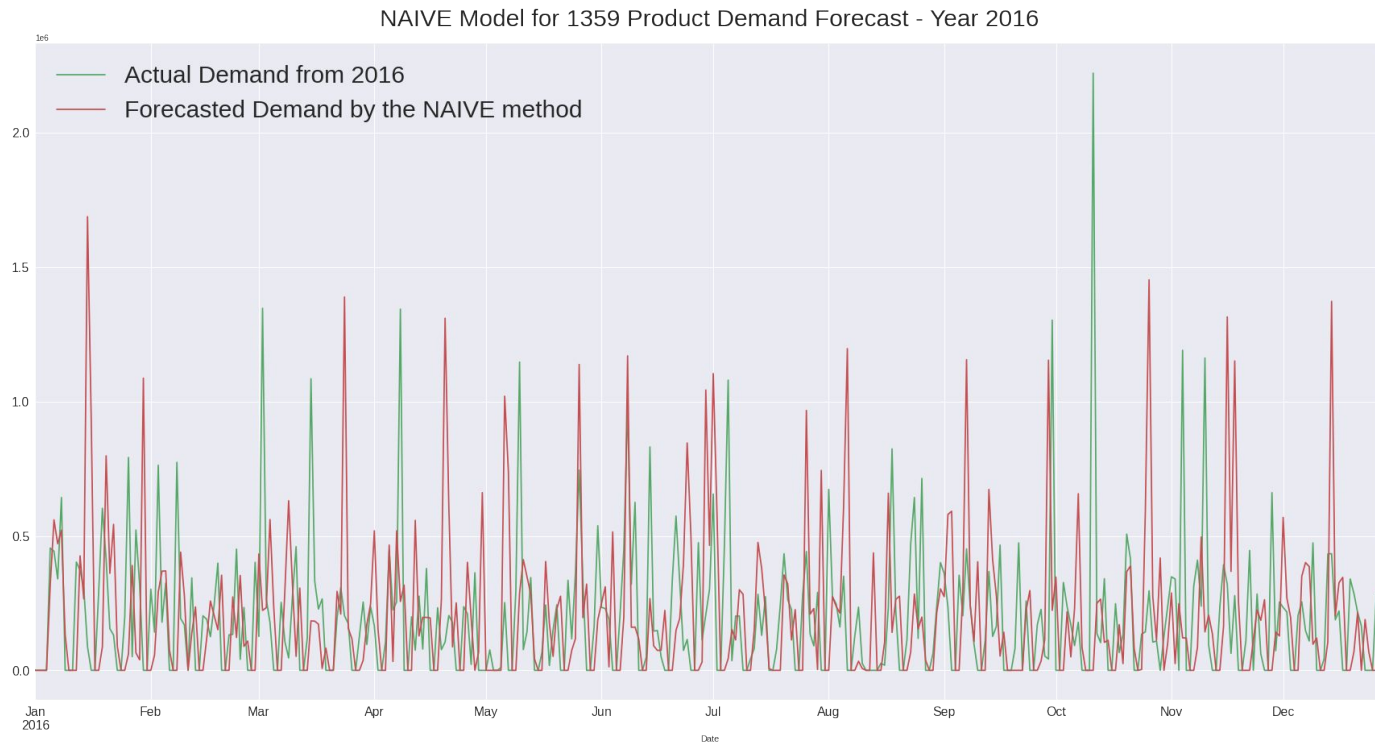
NAIVE

Previsão é o valor da última observação

MAE: 284522.63374

RMSE: 405853.3407

R2 SCORE: -1.42124



Modelos Clássicos

ARIMA

ORDEM (p, d, q): (3, 0, 2)

MAE: 156869.74047

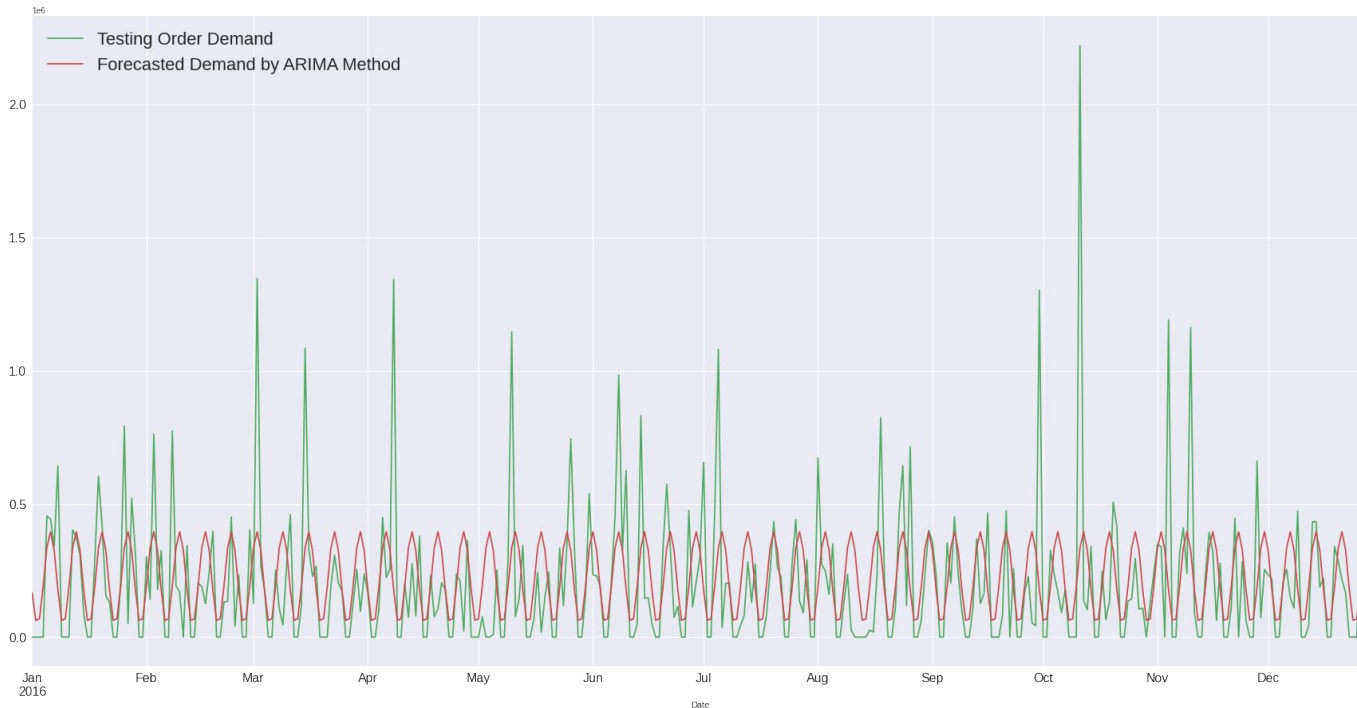
RMSE: 242612.48257

R2 SCORE: 0.14856

A estacionariedade da série foi identificada em teste, por isso de fato estamos diante do modelo ARMA (d=0)

Sazonalidade x SARIMAX

ARIMA Model for 1359 Product Demand Forecast - Year 2016



Modelagem ML

RANDOM FOREST

+ Interpretável/Explicável

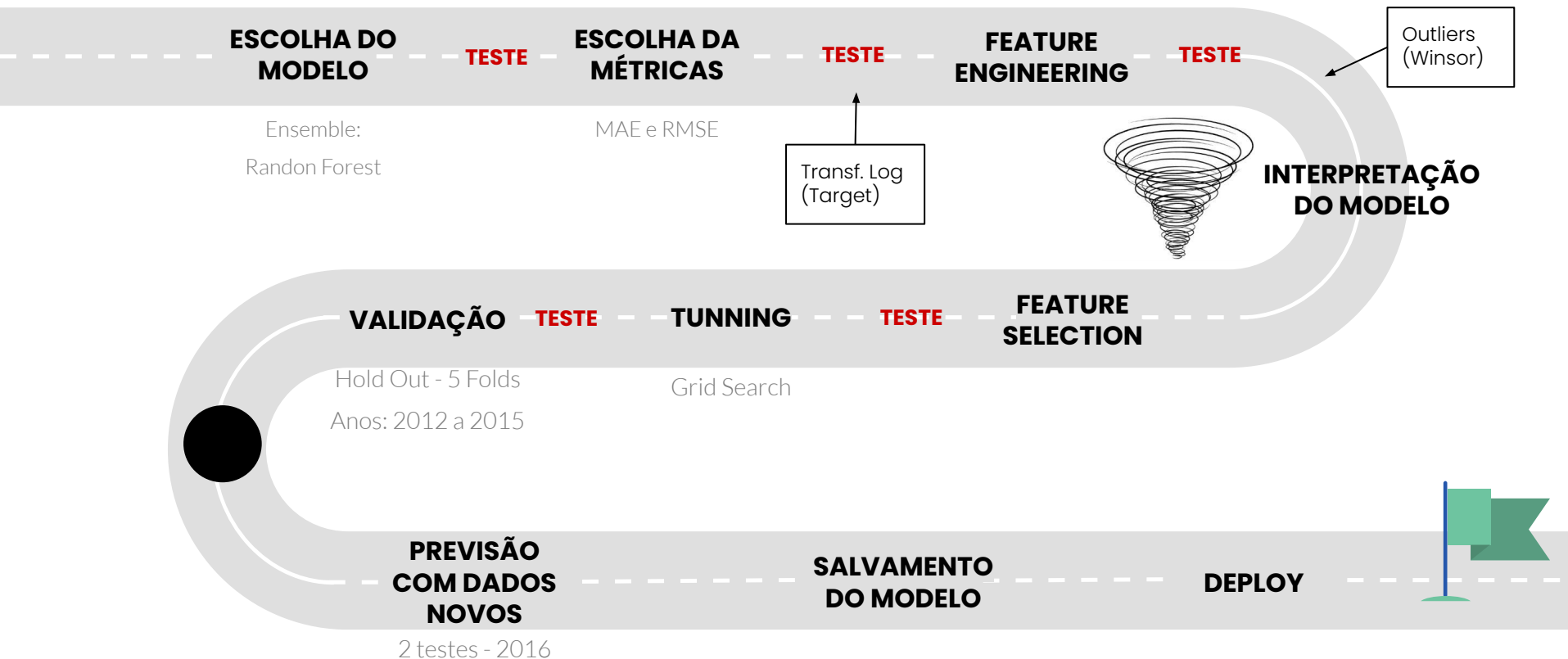
Trabalha melhor com dados escassos

Modelo mais leve (Consumo computacional)

Abordagem Diferente



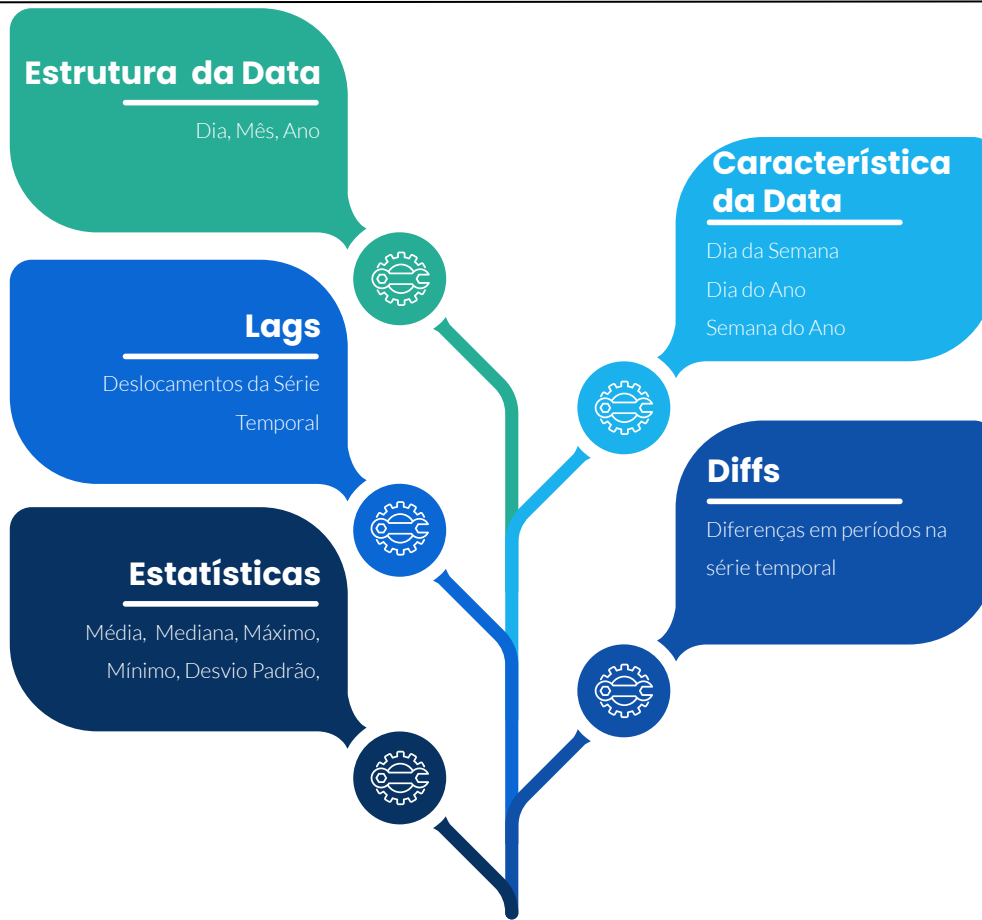
Modelagem ML



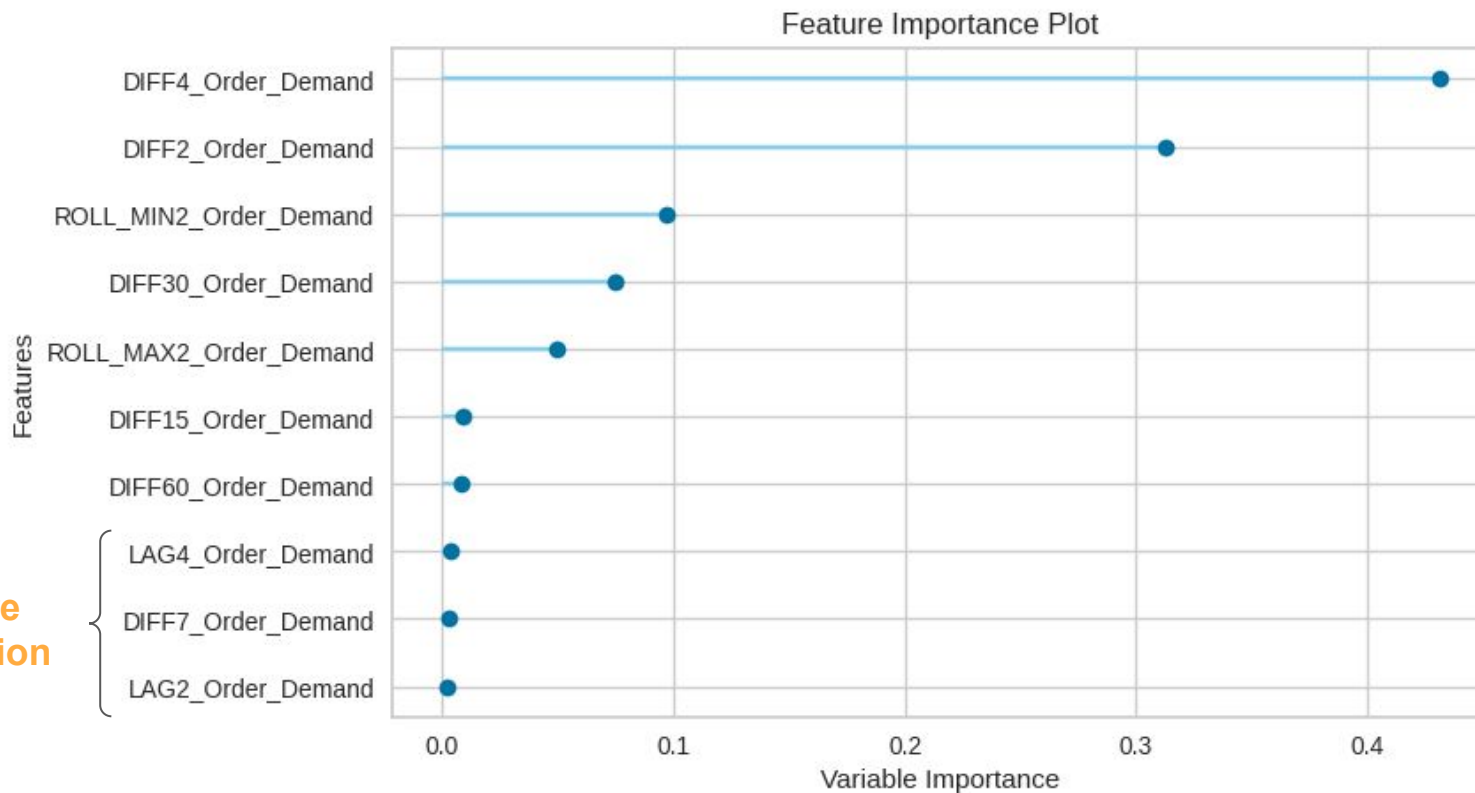
Feature Engineering



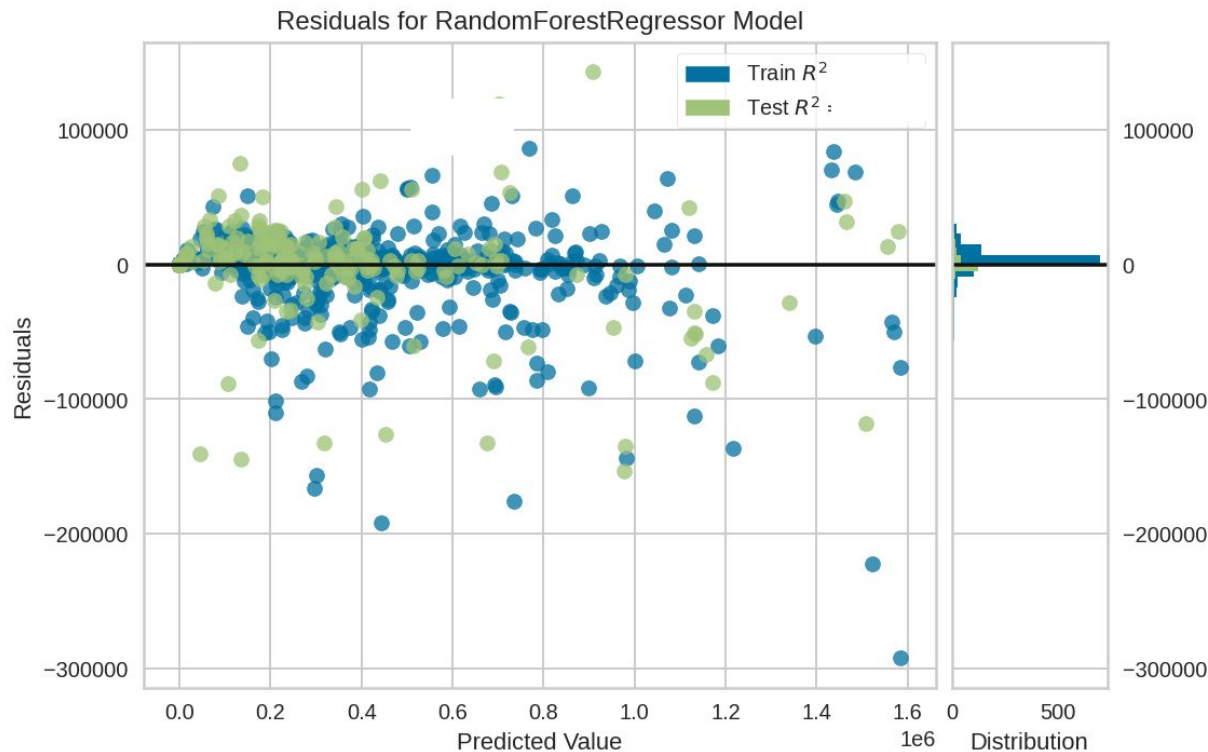
LAGS
[2, 4, 7, 15, 30, 60]



Interpretação do Modelo



Avaliação do Modelo

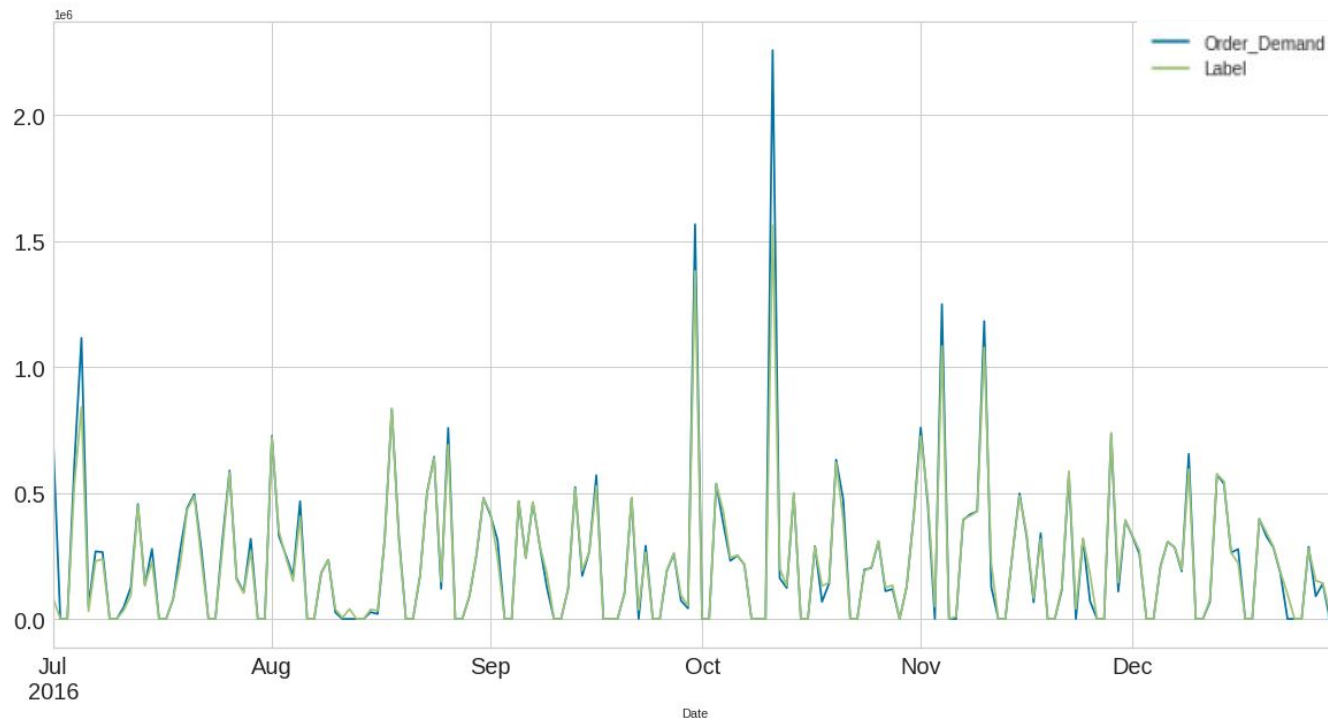


Model	MAE	MSE	RMSE
Random Forest Regressor	15527.1266	1.044605e+09	32320.3505

Resultados



Previsão de Demanda 2º Semestre 2016

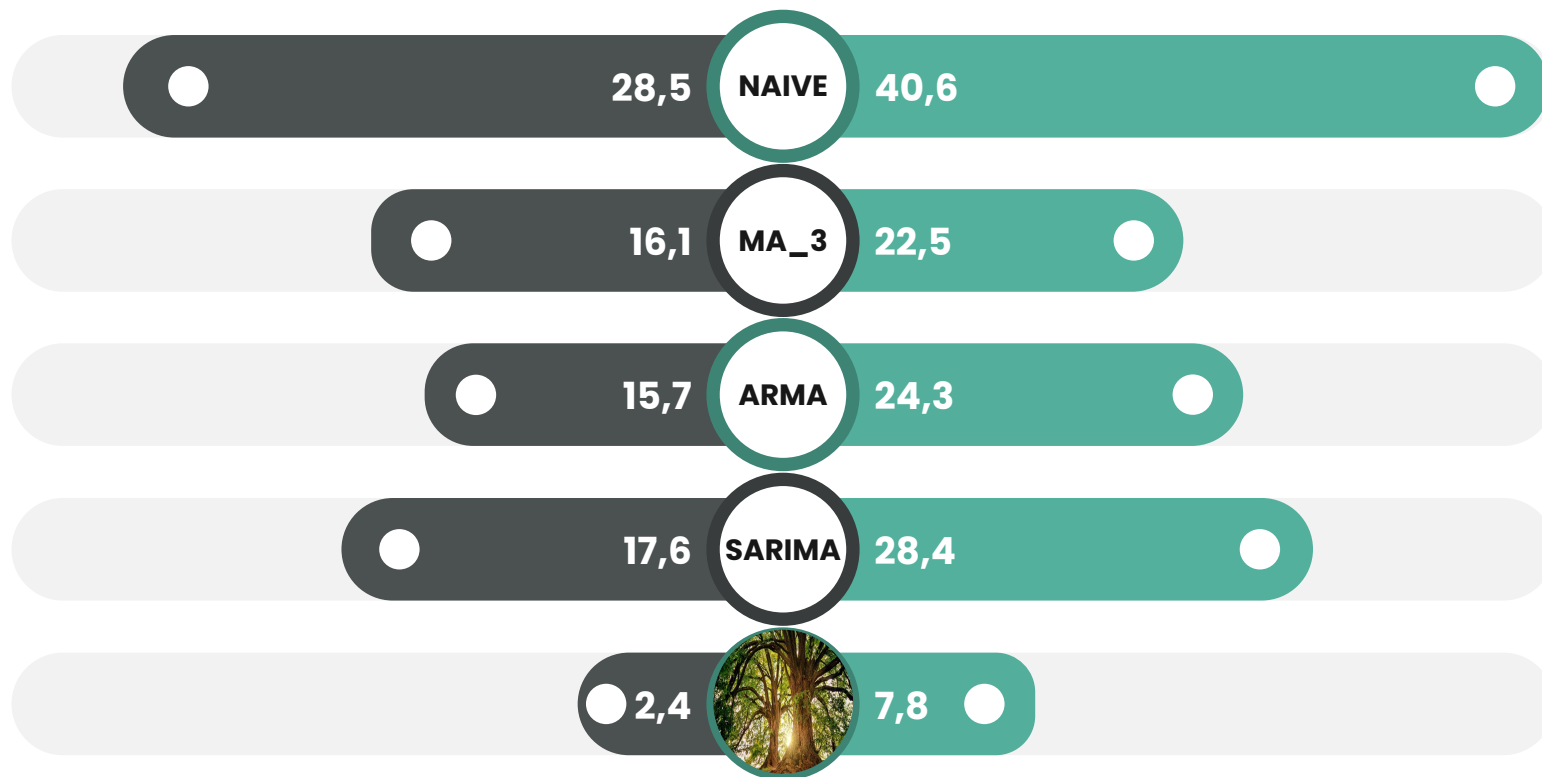


MAE: 24012.1581

RMSE: 78467.2199

R2 SCORE: 0.93

Erros dos Modelos



MAE (e5)

RMSE (e5)

Conclusão

Trabalhamos com:

- Data Cleaning, EDA, Séries Temporais, Modelos Estatísticos, Feature Engineering e Modelos de Machine Learning.
- **Principais Ferramentas:** Power BI, Statsmodel, Scikitlearn, Pandas, Numpy, Pycaret.

Dificuldades:

- Capacidade computacional

Trabalhos Futuros:

- Análises individuais de cada série temporal + Testes AutoML
- Cross-learning



Gilberto Filho

<https://www.linkedin.com/in/datagf/>

Wanderson Torres

<https://www.linkedin.com/in/wanderson-torres-31049522/>

