



Liga Competitiva de Ciências de Dados da USP

Constellation Challenge 2023

Leandro Coelho

IME - USP

ROTEIRO COM TUDO O QUE SERÁ APRESENTADO

- Descrição do Negócio
- Problema a ser resolvido
- Como é feito hoje
- Entendimento dos Dados
- Principais Insights Levantados
- Preparação dos Dados
- Justificar tomadas de decisões
- Modelagem
- Previsão
- Alcançamos o Sucesso?
- Próximos Passos

DESCRIÇÃO DO NEGOCIO

- Rede de supermercados do Equador
- Estamos presentes em 17 Estados
- Principalmente nas Capitais
- Majoritariamente em Pichincha

Estado	Qtd
Pichincha	19
Guayas	11
Santo Domingo de los Tsachilas	3
Azuay	3
Manabi	3
Cotopaxi	2
Tungurahua	2
Los Rios	2
El Oro	2
Chimborazo	1
Imbabura	1
Bolivar	1
Pastaza	1
Santa Elena	1
Loja	1
Esmeraldas	1



PROBLEMA DO NEGOCIO

- Prever a quantidade de produtos vendidos no próximo mês
- Regra de negocio
 - Preciso prever com um mês de antecedência
 - Preciso de tempo para faturar os pedidos e lidar com a logistica
- Conjuntos
 - Treino – 2013 a Maio 2017
 - Validação – Junho de 2017
 - Teste – Julho de 2017



Treino



Validação



Teste

BENCHMARK

- Como é feito hoje
 - Temos uma noção do que nossas lojas precisam, baseadas em um estudo antigo
- Diminuir o desperdício de alimentos relacionados ao excesso
- Satisfação do Cliente em ter o produto que precisa
- Queremos uma boa estimativa, o consumo exato está fora do nosso controle
- Estarmos preparados para Picos de Vendas*

Exploração dos Dados

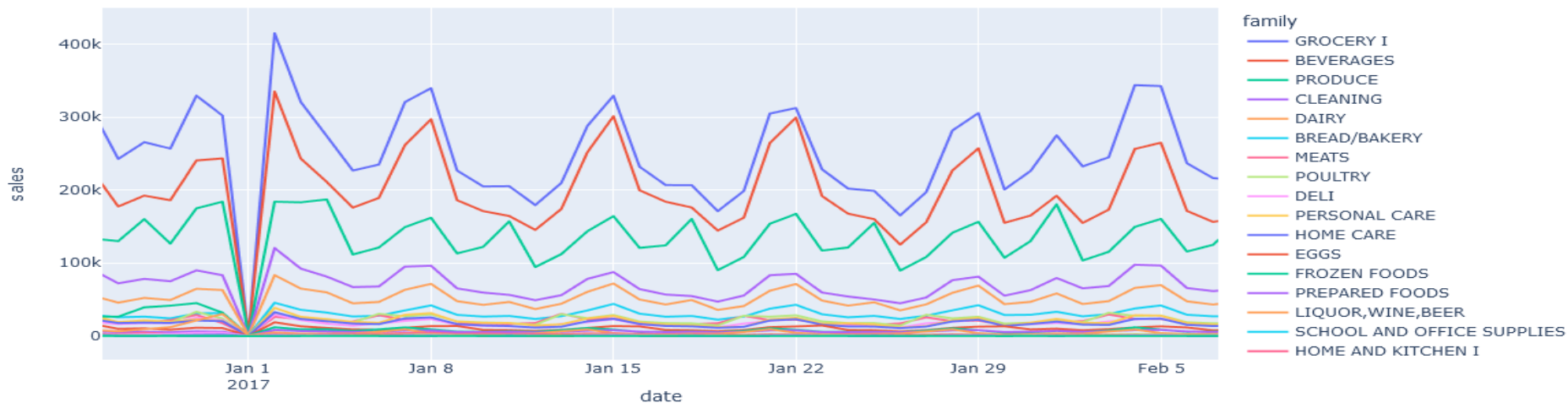


PRODUTOS MAIS VENDIDOS

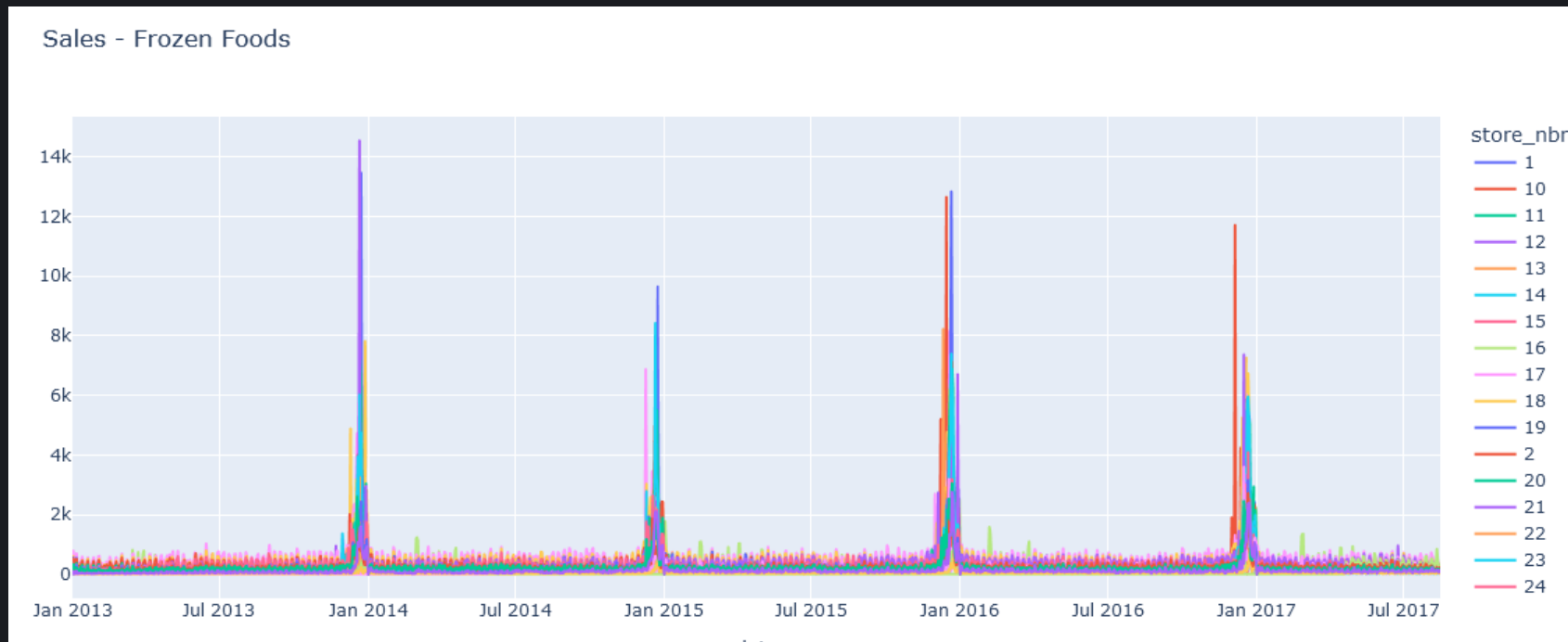
Media/dia

- Grocery I 204.000
- Beverages 128.000
- Produce 73.000
- Cleaning 58.000

Vendas por Família



DESTAQUE - FROZEN FOODS



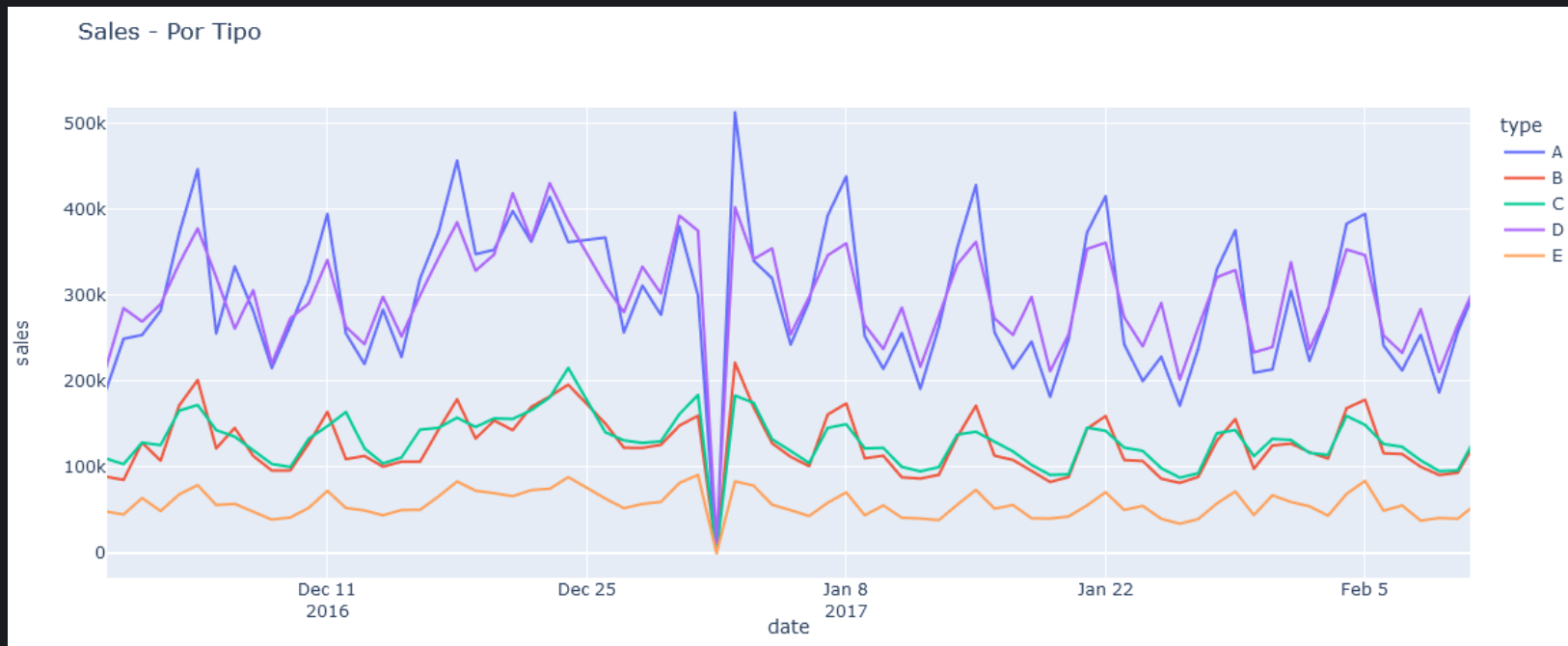
LOJAS QUE MAIS PRODUZEM

Loja	Madia/dia	Cidade
44	204.000	209.000
45	128.000	86.000
47	73.000	97.000
3	58.000	208.000

Tipo	Quantidade	Produção	Media/Dia(loja)
A	9	209.000	23.000
B	8	86.000	10.000
C	15	97.000	6.4000
D	18	208.000	11.000
E	4	35.000	8,75

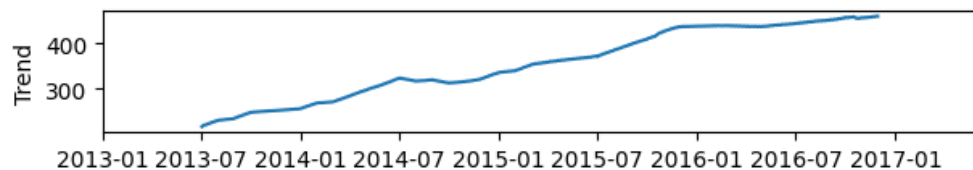
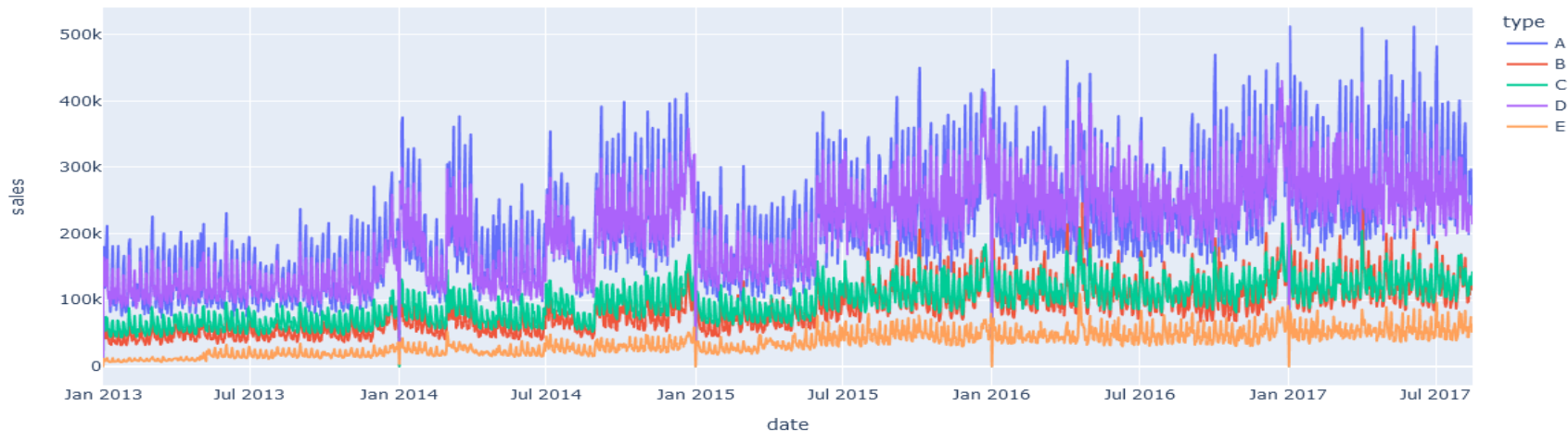
- Se fossemos priorizar algumas lojas, poderíamos segmentar as mais representativas

SAZONALIDADE – TIPO DE LOJA



TENDENCIA – POR TIPOS

Sales - Por Tipo

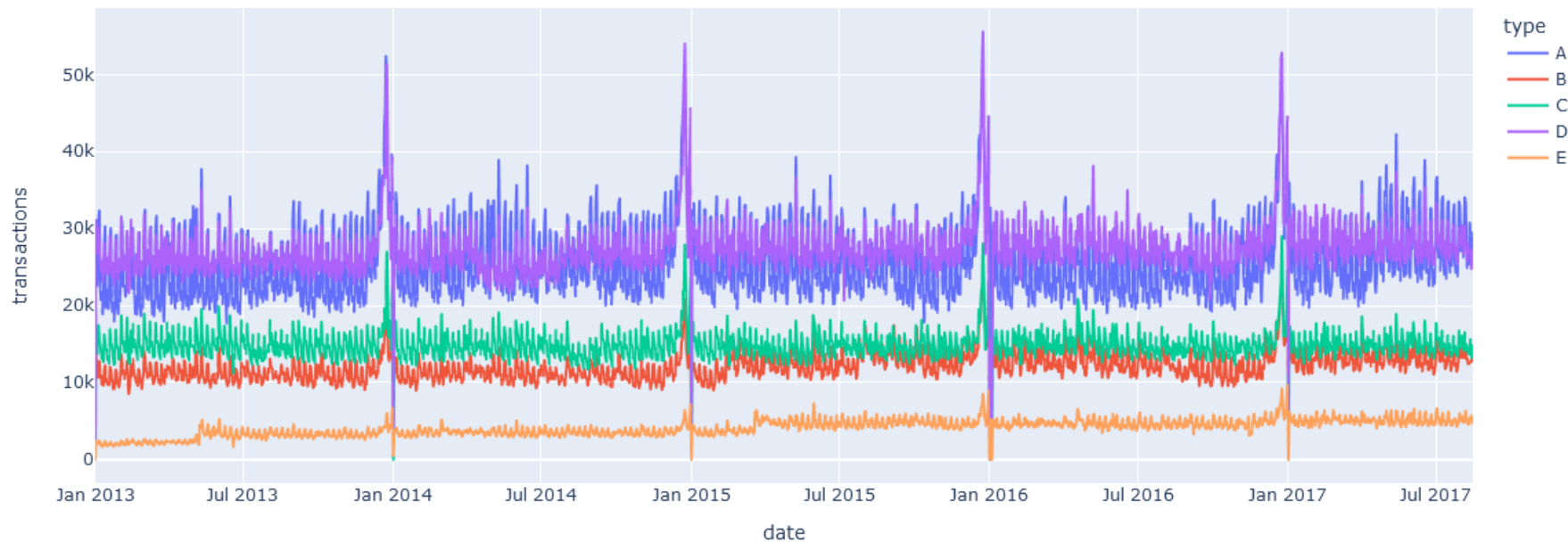


SAZONALIDADE - ANUAL

- Temos um aumento de vendas de fim de ano
- Devido a Festividades
- Puxado pelos produtos – FROZEN, Grocery
- Ano novo – Nenhuma Unidade Abre (Exceções)
- Início de ano são muitas contas – justificaria a diminuição de vendas em janeiro

SAZONALIDADE ANUAL – QUANTIDADES DE TRANSAÇÕES

Trasações por Cluster

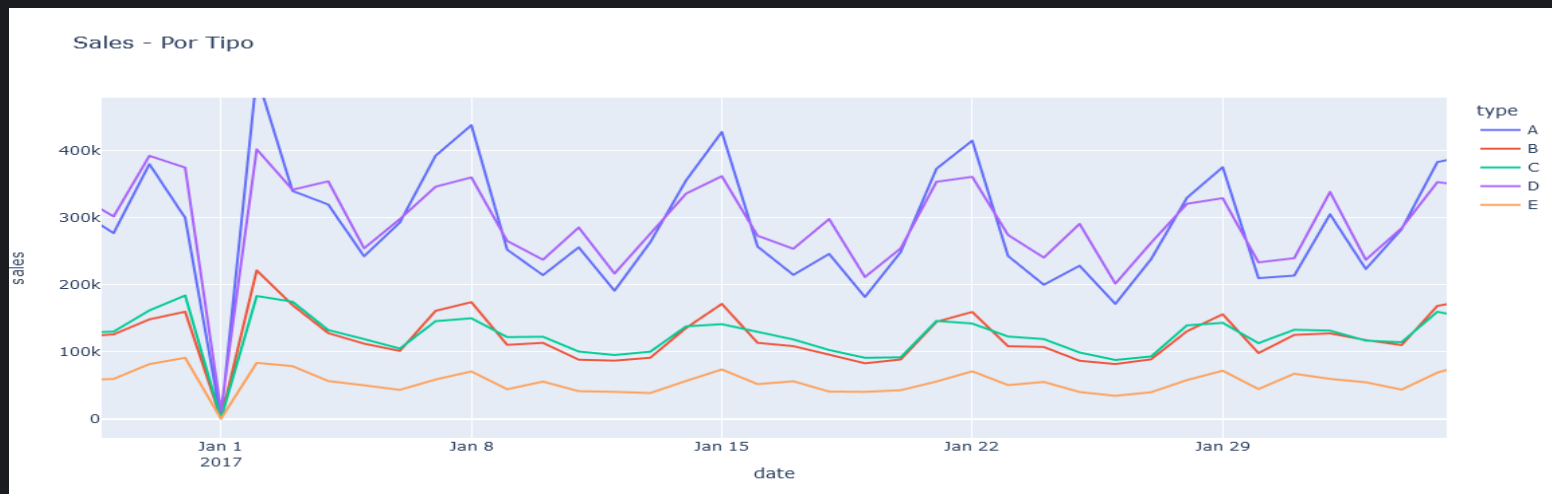


SAZONALIDADE MENSAL

- Início do mês tem mais vendas
- Recebimento de Salários
- Habito do consumidor

SAZONALIDADE SEMANAL

- Vendas quase dobram em finais de semana
- Dinâmicas dos Feriados
- Justificativa – Quando o trabalhador pode ir ao mercado
- Exemplo do brasil que mercados são cheios aos finais de semana



DINÂMICA DOS FERIADOS

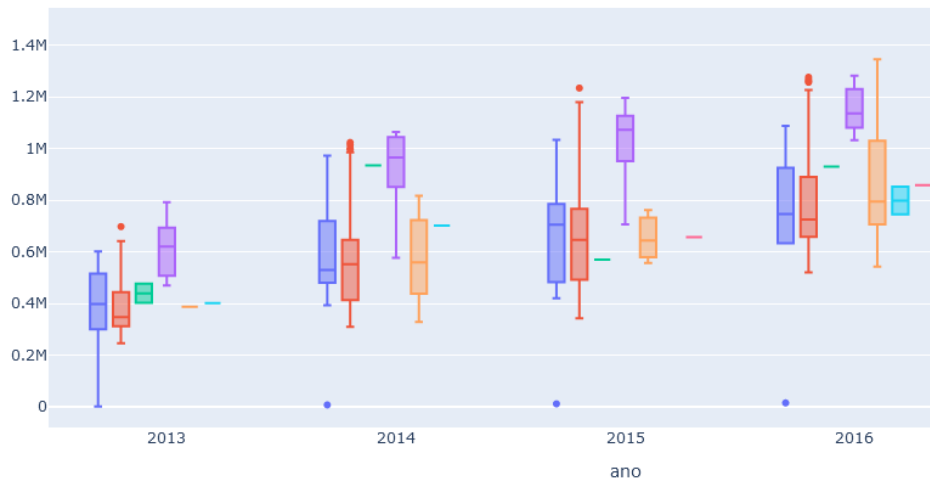
- Feriados
 - Eventos
 - Pontes e Compensações
 - Adicionais - Boxplot
-
- Obs: Nivel Nacional

FERIADOS - TRANSAÇÕES VS VENDAS

Transações - Feriados



Vendas - Feriados



type

- Holiday
- nan
- Work Day
- Additional
- Event
- Transfer
- Bridge

LOJAS FECHADAS

- Algumas lojas não foram inauguradas ou não venderam
- Períodos de lojas fechadas
- Obs: Algumas lojas não vendem alguns tipos de produtos

Modelagem



BENCHMARK

- Não Garanto a Estacionariedade de 'Lawn and Garden'
 - LAWN AND GARDEN | T-Statistic: -2.945 | Critical 1% value: -4.378 | P-Value: 0.322
- Modelo levemente ajustado
- Erro Quadrático Médio 229

PREPARAÇÃO DOS DADOS

- Conjunto de treino
- Conjunto de lojas – bom para lojas B e C
- Conjunto feriados – Informação importante

FEATURES

- Dia da semana
- Mês – específico de Dezembro
- Marcações de tipos de feriados
- Ano novo – marcação separada
- Fourier – mensal , freq. 4

ano_novo	dia_da_semana_Domingo	dia_da_semana_Quarta-feira	dia_da_semana_Quinta-feira	dia_da_semana_Segunda-feira
1	0	0	0	0
0	0	1	0	0
0	0	0	1	0
0	0	0	0	0
0	0	0	0	0
...
0	0	0	0	0
0	0	0	0	0
0	1	0	0	0
0	0	0	0	1
0	0	0	0	0

RESULTADOS

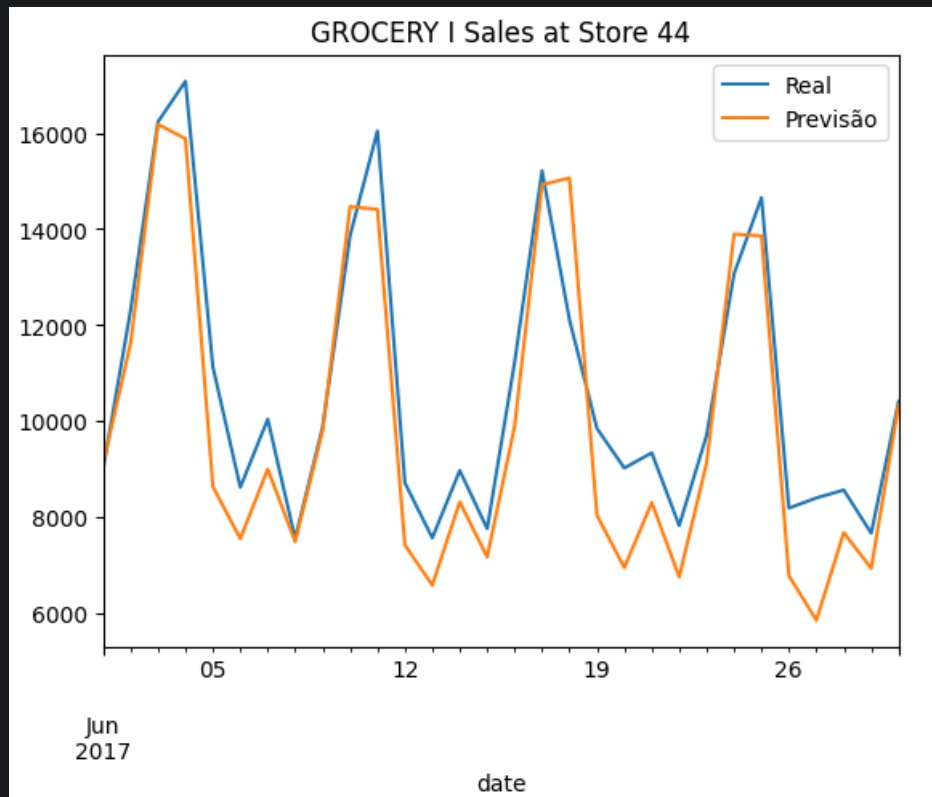
- Validação

- Erro Quadrático Médio 89

- Teste

- Erro Quadrático Médio 96

- Loja 44 – Grocery I



Próximos Passos



PRÓXIMOS PASSOS

- Encontrar grupos de produtos com sazonalidades diferentes
- Frozem Food é um exemplo de que é um bom caminho
- Inserir mais variáveis exógenas – preço do petróleo
- Testar outras abordagens e outros algoritmos
- Relações não lineares entre nossos dados e a variável a ser predita

OBRIGADO!!!

- Agradecimentos a Organização do BeeData e ao Leonardo
- Link do projeto
 - https://github.com/LeandroCoelhos/BeeData/tree/main/Desafio_Constellation
- Apresentação
- Redes Sociais
 - <https://www.linkedin.com/in/leandro-coelhos/>