

MBA
USP
ESALQ

Café e tecnologia: Seleção de métodos para projeção do preço da saca de café

Ramon Ferreira Cruz
Daniel Alvares Firmin

Introdução:

- Importante produto para economia brasileira
- Variação do preço influenciada por diversos fatores
- Forte valorização nos últimos anos



Objetivo:

Este trabalho se dedicou em entender quais fatores influenciam a cotação da saca de café, descrever o comportamento da série, por fim, propor diferentes modelos de natureza supervisionada para previsão de valores futuros.

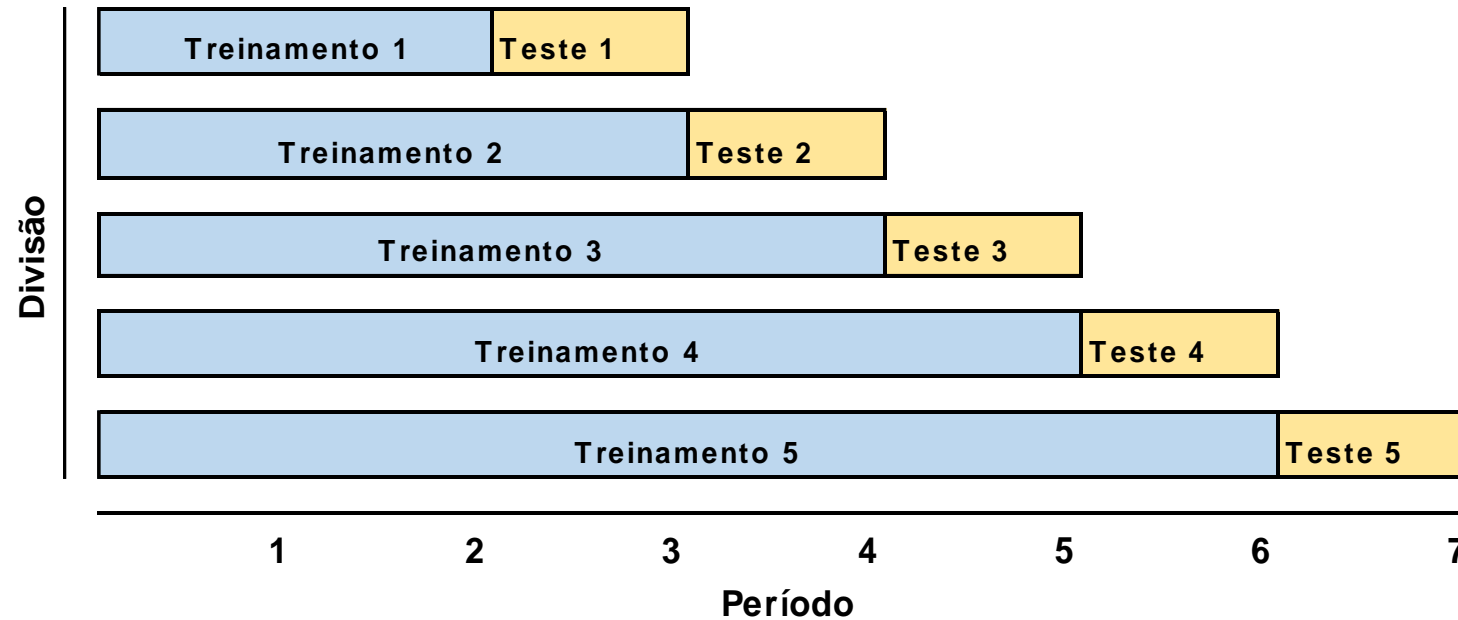


Material e Métodos:

- Coleta de dados
- Análise e descrição da série temporal
- Análise e descrição de cada variável exógena em relação a variável endógena
- Treinamento dos modelos
- Avaliação dos resultados
- Escolha do modelo

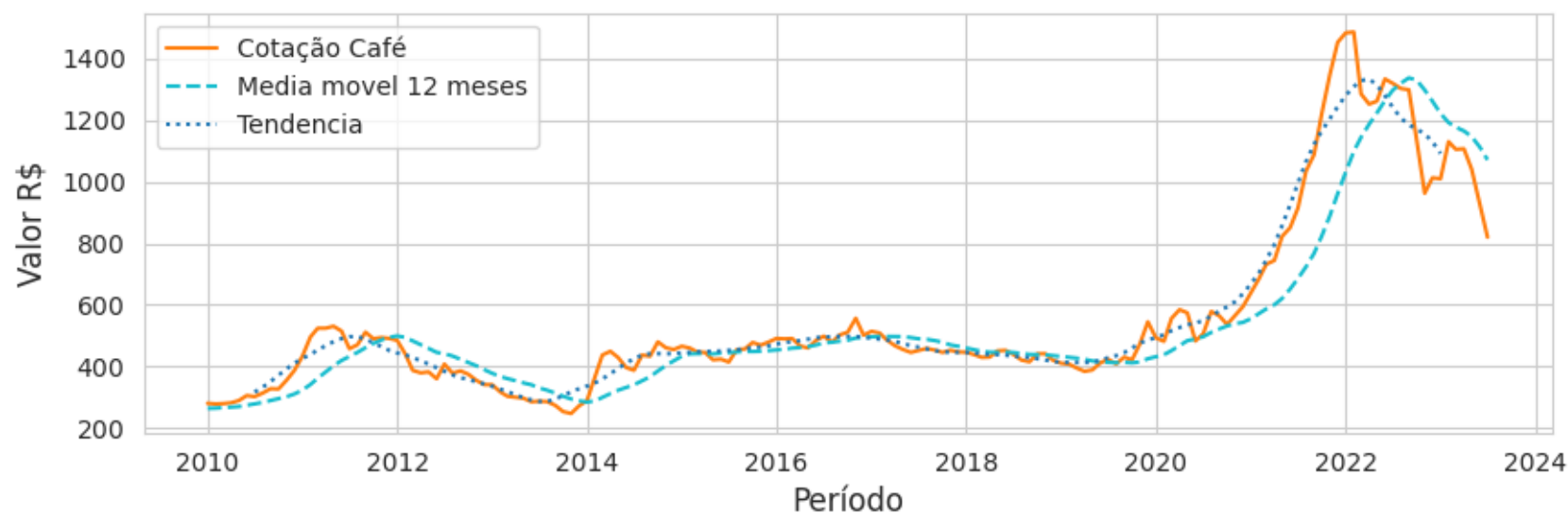
Material e Métodos:

Validação Cruzada



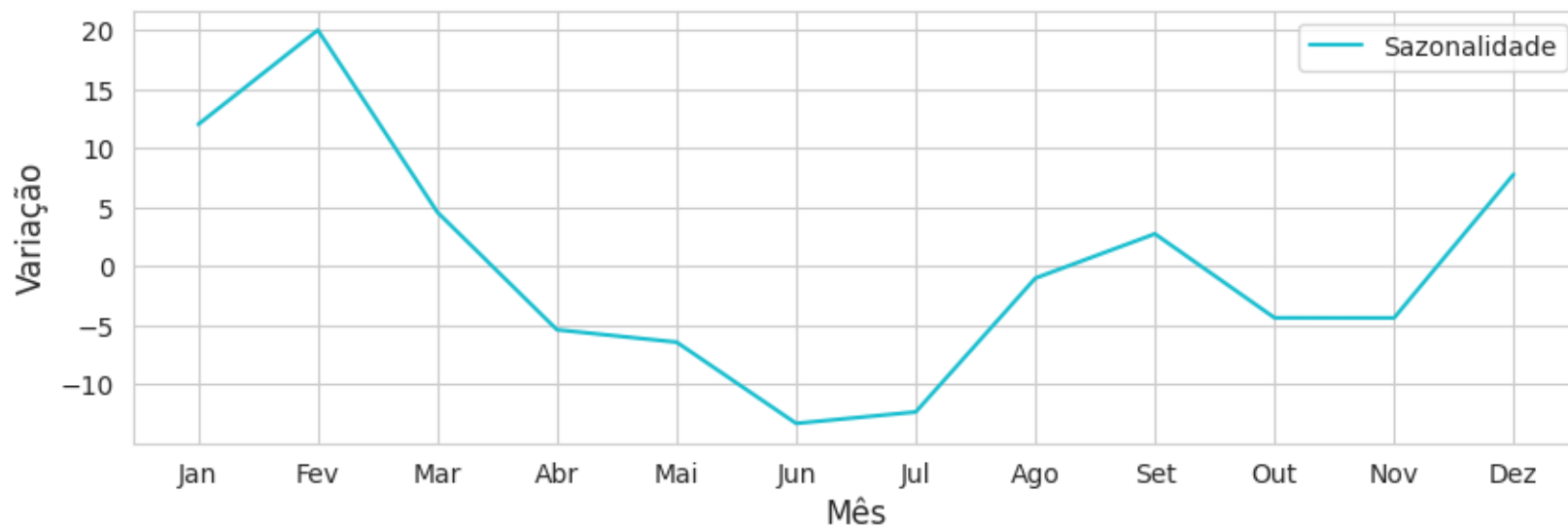
Resultados e Discussão:

Série temporal Café x tendência



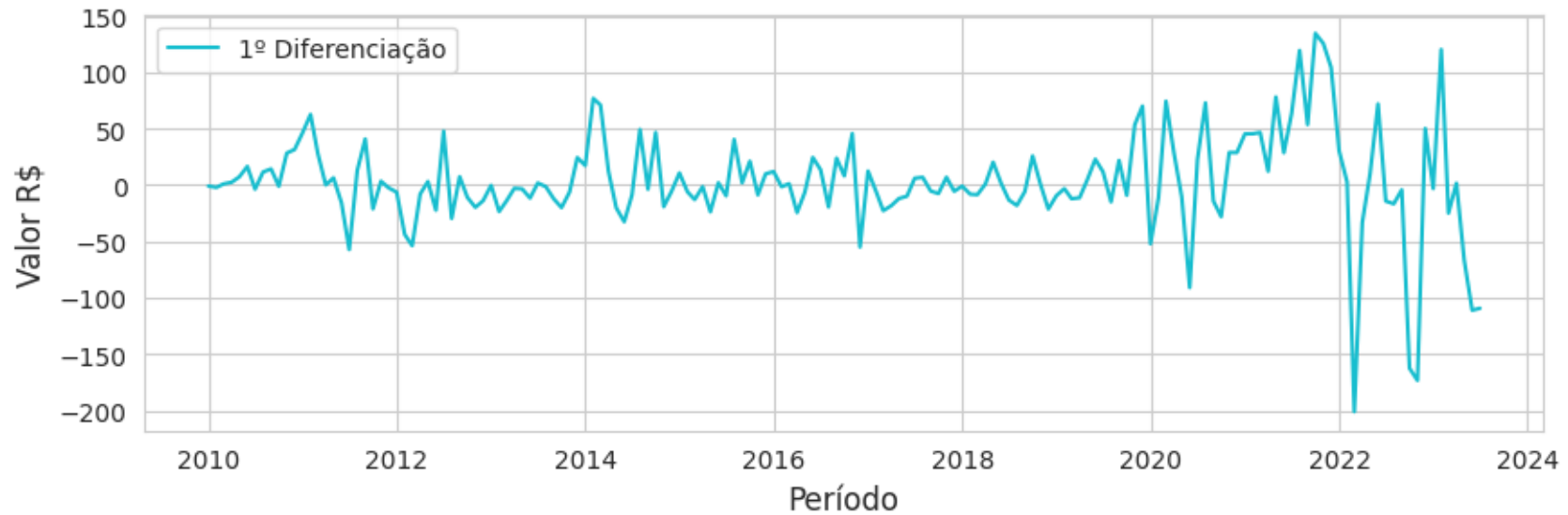
Resultados e Discussão:

Sazonalidade



Resultados e Discussão:

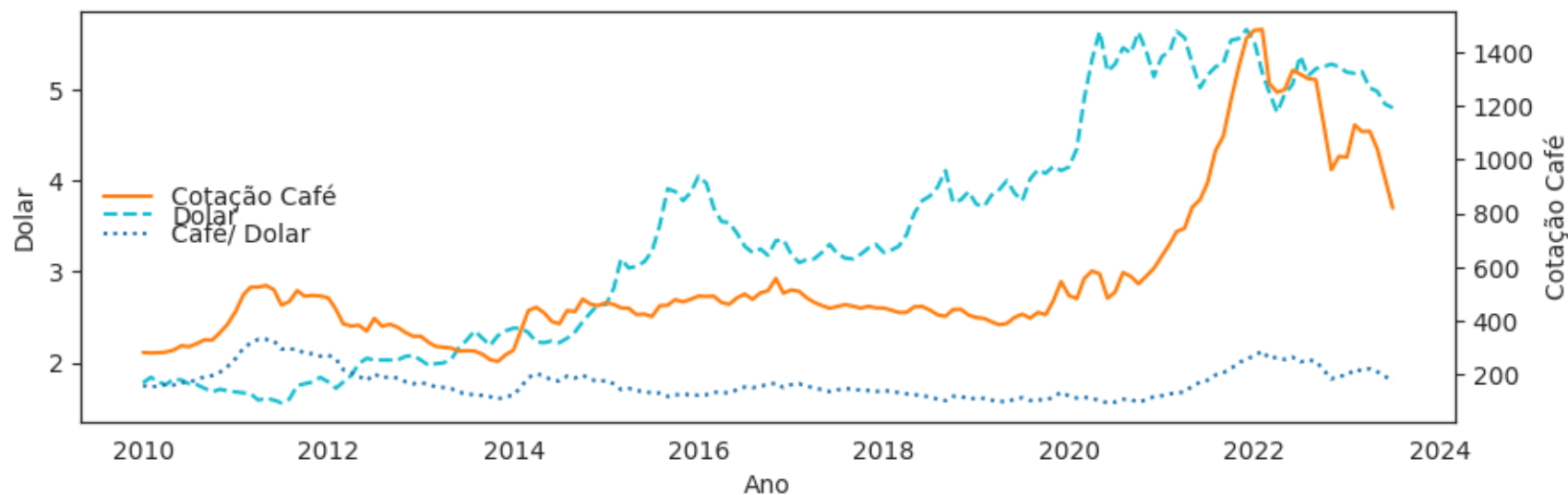
Estacionariedade



Teste estatístico de Dickey Fuller Aumentado

Resultados e Discussão:

Café x Dólar

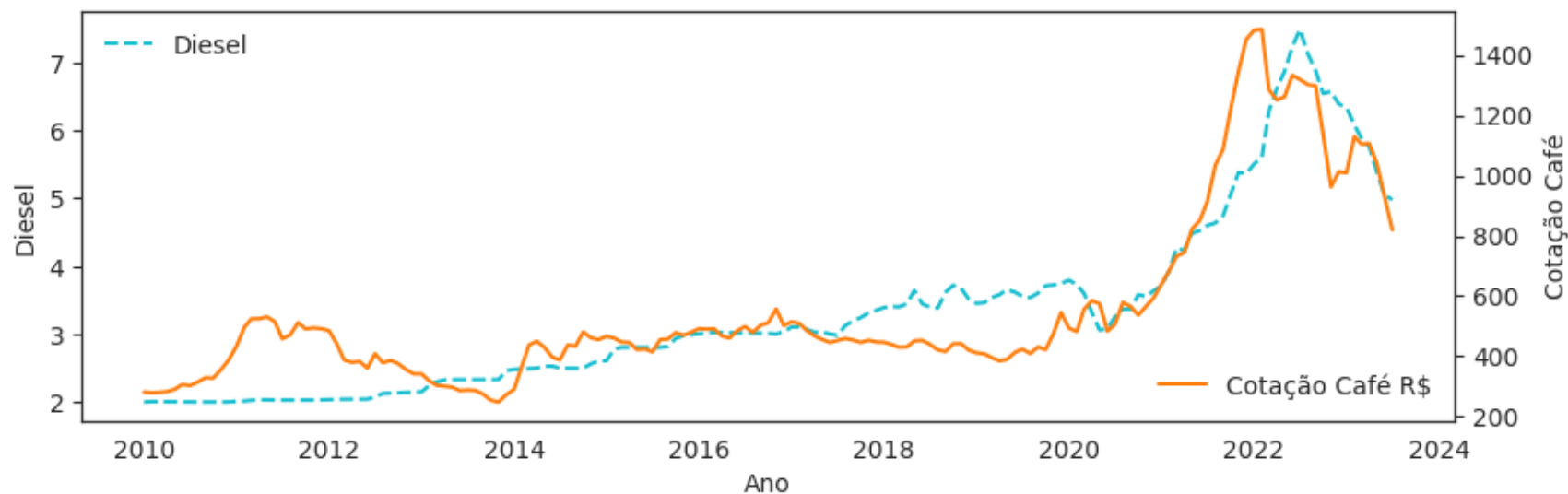


Correlação: 0.71

Teste de causalidade de Granger: 2 Atrasos

Resultados e Discussão:

Café x Diesel

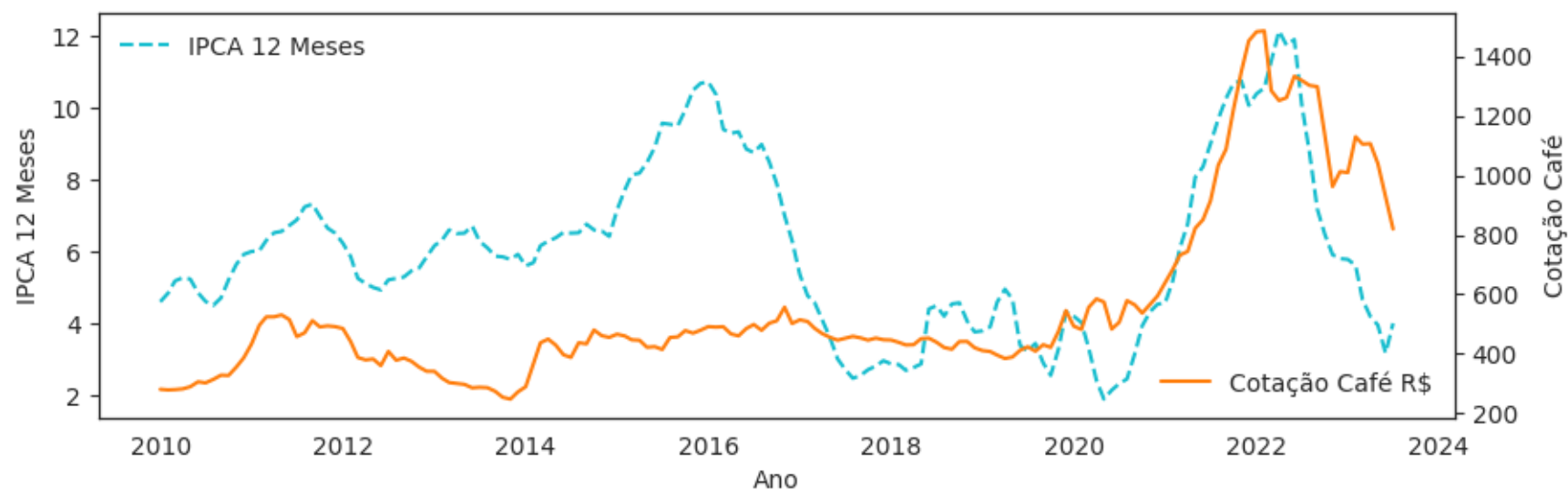


Correlação: 0.89

Teste de causalidade de Granger: 1 Atraso

Resultados e Discussão:

Café x IPCA 12 Meses

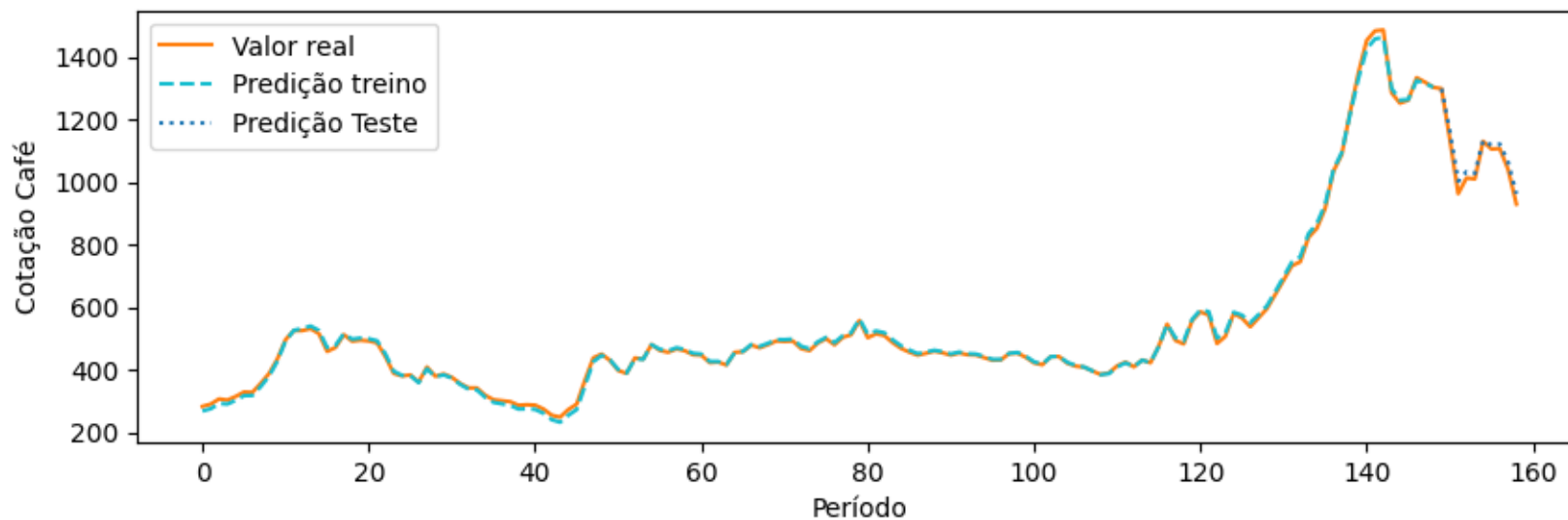


Correlação: 0.50

Teste de causalidade de Granger: 2 Atrasos

Resultados e Discussão:

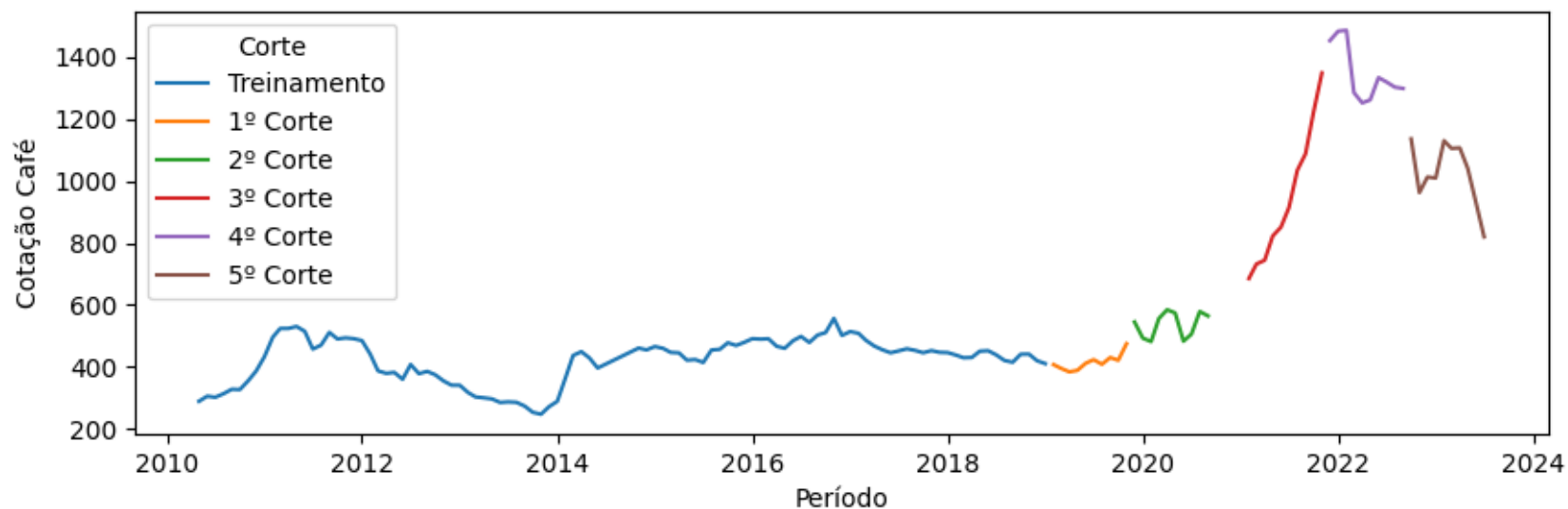
Redes neurais LSTM



Melhor ajuste parcial.

Resultados e Discussão:

Divisão Validação cruzada



Validar a capacidade de generalização dos modelos

Resultados e Discussão:

Métricas de resultado

Modelo	MAPE	MSE	MSLE	MAE	RMSE
Random Forest	3,92%	453,06	0,0025	16,47	21,29
Xgboost Arvore	5,19%	878,41	0,005	22,25	29,64
Xgboost Linear	3,42%	480,35	0,0025	14,81	21,92
LSTM	4,90%	1.344,21	0,0058	23,79	36,66
ARIMAX	7,19%	2.469,28	0,0098	36,04	49,69

Rodada 1

Resultados e Discussão:

Métricas de resultado

Modelo	MAPE	MSE	MSLE	MAE	RMSE
Random Forest	9,18%	3.734,73	0,0134	51,32	61,11
Xgboost Arvore	8,36%	3.272,95	0,0116	46,92	57,21
Xgboost Linear	8,79%	3.039,66	0,0112	46,82	55,13
LSTM	7,87%	2.336,30	0,0079	43,99	48,34
ARIMAX	7,02%	2.087,40	0,0069	38,84	45,69

Rodada 2

Resultados e Discussão:

Métricas de resultado

Modelo	MAPE	MSE	MSLE	MAE	RMSE
Random Forest	39,04%	201.456,53	0,3165	395,32	448,84
Xgboost Arvore	39,85%	207.855,16	0,3322	402,83	455,91
Xgboost Linear	7,04%	8.115,89	0,0074	73,1	90,09
LSTM	13,20%	16.962,26	0,0212	119,93	130,24
ARIMAX	6,65%	5.215,36	0,0057	64,63	72,22

Rodada 3

Resultados e Discussão:

Métricas de resultado

Modelo	MAPE	MSE	MSLE	MAE	RMSE
Random Forest	14,78%	47.969,76	0,0298	203,08	219,02
Xgboost Arvore	5,56%	7.444,78	0,0039	76,1	86,28
Xgboost Linear	11,83%	31.864,28	0,0159	155,2	178,51
LSTM	10,10%	24.947,48	0,0128	132,14	157,95
ARIMAX	3,64%	7.699,98	0,0039	47,8	87,75

Rodada 4

Resultados e Discussão:

Métricas de resultado

Modelo	MAPE	MSE	MSLE	MAE	RMSE
Random Forest	19,59%	45.408,06	0,0401	188,86	213,09
Xgboost Arvore	14,19%	24.388,61	0,022	139,63	156,17
Xgboost Linear	7,25%	7.084,04	0,0068	73,3	84,17
LSTM	2,80%	1.232,79	0,0012	28,47	35,11
ARIMAX	8,20%	10.465,74	0,0092	82,53	102,3

Rodada 5

Resultados e Discussão:

Métricas de resultado

Modelo	MAPE	MSE	MSLE	MAE	RMSE
Random Forest	17,31%	59.804,43	0,0805	171,01	192,67
Xgboost Arvore	14,63%	48.767,98	0,075	137,55	157,04
Xgboost Linear	7,67%	10.116,85	0,0088	72,65	85,96
ARIMAX	6,54%	5.587,55	0,0071	53,97	71,53
LSTM	7,77%	9.364,61	0,0098	69,67	81,66

Avaliação Final

Conclusão:

- Avaliação da série temporal
- Variáveis exógenas selecionadas influenciam a cotação da saca do café
- Melhor modelo: ARIMA
- Pesquisas Futuras
 - Investigação de outras variáveis
 - Implementação de outros modelos
 - Aprofundar na otimização de parâmetros

Obrigado!