

Rogério de Oliveira
Orlando Yesid Esparza Albarracin
Gustavo Rocha da Silva

Introdução às séries temporais

UMA ABORDAGEM
PRÁTICA EM PYTHON



Editora
Mackenzie

Introdução às séries temporais

UMA ABORDAGEM PRÁTICA EM PYTHON



50

UNIVERSIDADE PRESBITERIANA MACKENZIE

Reitor: Marco Tullio de Castro Vasconcelos

EDITORA MACKENZIE

Coordenador: John Sydenstricker-Neto

Conselho Editorial

Alexandre Nabil Ghobril

Ana Alexandra Caldas Osório

Cecília de Carvalho Castro e Silva

Gianpaolo Poggio Smanio

Gildásio Jesus Barbosa dos Reis

José Geraldo Simões Junior

José Luiz de Lima Filho

Maria Lucia Marcondes Carvalho Vasconcelos

Reinaldo Guerreiro

Rosângela Patriota Ramos

Walter Eustáquio Ribeiro

COLEÇÃO CONEXÃO INICIAL

Diretora: Rosângela Patriota Ramos

Rogério de Oliveira
Orlando Yesid Esparza Albarracin
Gustavo Rocha da Silva

Introdução às séries temporais

UMA ABORDAGEM PRÁTICA EM PYTHON

© 2024 Rogério de Oliveira, Orlando Yesid Esparza Albarracin e
Gustavo Rocha da Silva

Todos os direitos reservados à Editora Mackenzie.
Nenhuma parte desta publicação poderá ser reproduzida por qualquer meio ou forma
sem a prévia autorização da Editora Mackenzie.

Coordenação editorial: Jéssica Dametta
Preparação de texto: Jéssica Dametta
Revisão: Andréia Ferreira Cominetti
Diagramação: Pedro P. Videira Pancheri
Capa: Pedro P. Videira Pancheri e Isabelle Callegari Lopes
Estagiária editorial: Bruna Silva de Angelis

Dados Internacionais de Catalogação na Publicação (CIP)

O48i Oliveira, Rogério de.
Introdução às séries temporais: uma abordagem prática em
Python / Rogério de Oliveira, Orlando Yesid Esparza Albarracin,
Gustavo Rocha da Silva – São Paulo : Editora Mackenzie, 2024.
200 p. : il. ; 23 cm. – (Conexão Inicial ; 50).

Inclui referências bibliográficas, glossário e índice.
ISBN 978-65-264-0431-7

1. Python. 2. Linguagem de Programação. 3. Estruturas de
Dados. 4. Processamento de Dados. 5. Séries Temporais. I.
Albarracin, Orlando Yesid Esparza. II. Silva, Gustavo Rocha da.
III. Título. IV. Série.

CDD 005.133

Bibliotecária Responsável: Paola D'Amato - CRB 8/6271

EDITORA MACKENZIE
Rua da Consolação, 930
Edifício João Calvino, 6º andar
São Paulo – SP – CEP 01302-907
Tel.: (11) 2114-8774
editora@mackenzie.br
www.mackenzie.br/editora

Editora afiliada:



Sumário

Sobre os autores	7
Prefácio	9
Guia rápido	13
Funções deste livro – <code>tspplot()</code> , <code>tspdisplay()</code>	13
Trabalhando com datas no Python – <code>datetime</code> , <code>strptime</code> , <code>strptime</code> , <code>pd.to_datetime</code>	14
Strings para datas e datas para strings	14
Aritmética com datas	15
Datetime no Pandas	16
Validando o formato de datas	18
Aquisição dos dados – <code>DataFrame</code> , <code>pd.read_csv()</code>	19
Índices – <code>df.index</code> , <code>tspdisplay()</code>	19
Seleção dos dados – <code>df[<seleção de linhas>][<lista de colunas>]</code>	20
Gráficos – <code>plot()</code> <code>tspplot()</code>	21
Reamostragem – <code>.resample()</code>	23
Médias móveis – <code>rolling(n)</code>	24
Decomposição – <code>seasonal_decompose()</code> , <code>tspdecompose()</code>	25
Manipulações de séries temporais – <code>.diff()</code> <code>.shift()</code> <code>.cumsum()</code>	26
Introdução às séries temporais	29
Séries temporais	29
Decomposição de séries temporais	31
Construindo uma série temporal	34
Estacionariedade	37
Tendências e transformações	39
Sazonalidade	50
Resíduos	57
Intervalos de confiança	59
Resumo do capítulo	61
Modelos de suavização exponencial	63
Modelo 1 – suavização exponencial simples	68
Modelo 2 – suavização exponencial linear de Holt	73
Modelo 3 – suavização exponencial sazonal de Holt-Winters	77
Método multiplicativo de Holt-Winters	82
Resumo do capítulo	86

Modelos de predição ARMA e ARIMA	89
Conceitos importantes	90
Modelos autorregressivos	95
Modelos de médias móveis	98
Modelos autorregressivos e de médias móveis	100
Modelos autorregressivos integrados e de médias móveis	102
Resumo do capítulo	122
Modelos ARIMA sazonais e com variáveis exógenas	125
Modelos SARIMA	125
Modelos SARIMAX	140
Outros modelos com sazonalidade	151
Resumo do capítulo	152
Aprendizado de máquina	155
O aprendizado de máquina supervisionado para aproximação de funções	156
Aproximando uma função linear	156
Esquema geral do aprendizado supervisionado	158
Esquema geral do aprendizado supervisionado com o scikit-learn	159
Outros modelos não lineares	162
Outros modelos com o scikit-learn	163
Um modelo não linear: regressor k-vizinhos mais próximos	166
Outros modelos regressores	167
Empregando regressores para séries temporais	169
Série de emissões de CO ₂	169
Bike Sharing: machine learning	172
Modelo ARIMA	173
Modelo de aprendizado de máquina	174
Modelos neurais	178
Redes neurais	179
Modelo multivariado simples, $(x_t^a, x_t^b, \dots) \rightarrow x_{t+1}^a$	181
Deep learning	183
Resumo do capítulo	184
Conclusão	185
Referências	187
Bibliografia comentada	189
Glossário	195
Índice	197

Sobre os autores

ROGÉRIO DE OLIVEIRA é doutor em Ciências Aplicadas pela Escola Politécnica da Universidade de São Paulo (Poli-USP), mestre em Engenharia Elétrica pela Universidade Presbiteriana Mackenzie (UPM), graduado em Matemática (IME-USP) e em Ciências Econômicas (FEA-USP), com Especialização em Neurociências e Psicologia Aplicada pela UPM. É professor das disciplinas de Python, Ciências de Dados, Inteligência Artificial, e Probabilidade e Estatística Aplicadas nos cursos de graduação da UPM e do Instituto Mauá de Tecnologia (IMT), e professor colaborador do Programa de Pós-Graduação em Computação Aplicada da UPM. Pela Coleção Conexão Inicial, é também autor do livro *Visualização de dados com Python*.

ORLANDO YESID ESPARZA ALBARRACIN é doutor e mestre em Estatística pelo Instituto de Matemática e Estatística (IME-USP) e graduado em Matemática pela Universidad Industrial de Santander (Colômbia). É professor das disciplinas de Estatística Aplicada e Análise de Dados nos cursos da Escola de Engenharia da UPM.

GUSTAVO ROCHA DA SILVA é doutor e mestre em Astrofísica pelo Instituto de Astronomia, Geofísica e Ciências Atmosféricas (IAG-USP), possui pós-doutorado pela Universidade de Princeton (Estados Unidos) e é bacharel em Física pela USP. Atualmente, é gerente na área de Modelagem e Data Analytics para prevenção à fraude. É professor das disciplinas de Análise e Ciência de Dados da Escola de Engenharia da UPM.

Prefácio

Introdução às séries temporais: uma abordagem prática em Python se destina a você que está interessado em conhecer e aplicar métodos de análise e previsão de séries temporais em campos como Engenharia (para previsão de demandas de produtos), Comércio (previsão de preços), Finanças (séries econômicas e de ações), Ciências (séries meteorológicas e outros fenômenos temporais) e até mesmo Ciências Sociais e da Saúde (séries de casos de Covid-19 e evolução de indicadores sociais).

Este livro foi elaborado como um material de apoio para disciplinas de cursos de Engenharia e Administração, mas pode ser empregado total ou parcialmente em cursos de diferentes áreas. Sendo seu objetivo principalmente didático, adotamos a forma dialógica no texto e, muitas vezes, vamos nos referir diretamente ao leitor.

Ao longo do texto não só explicamos e discutimos os principais conceitos de séries temporais, como também trazemos implementações práticas de análise e previsões para séries de dados, que julgamos ser de interesse para um público diverso. Os exemplos vão de séries de dados da bolsa de valores e do aquecimento global até dados de demanda por aluguel de bicicletas.

Para a implementação dos modelos, empregamos o Python e suas bibliotecas, mas buscamos construir um texto que possibilitasse a leitura também pelo leitor não interessado na codificação, ou mesmo não familiarizado com Python e outras linguagens de programação. A ideia é que você possa entender e acompanhar os programas, isto é, os modelos e o que eles *produzem*, deixando os detalhes de codificação para uma segunda leitura ou para quando tiver a necessidade de aplicar esses modelos. Procuramos, assim, elaborar um texto útil, tanto para programadores quanto para não programadores.

De qualquer modo, para quem quiser se iniciar nesse campo da programação e desenvolver seus próprios modelos, fornecemos desde conceitos básicos (e você poderá encontrar no *site* do livro suplementos digitais de Python, Pandas e matplotlib) até tópicos avançados, além de todo o código deste livro.

No decorrer do livro, você encontrará os modelos tradicionais empregados na maior parte dos livros-texto de análise de séries temporais, como os modelos AR, ARMA, ARIMA, SARIMAX etc., e também um capítulo dedicado ao tratamento de séries temporais com modelos de aprendizado de máquina. Há uma lacuna desses modelos em textos mais introdutórios, e buscamos preenchê-la na medida em que o uso desses modelos vem crescendo rapidamente e possibilita a solução de uma grande quantidade de problemas com grande volume de dados, complexidade e multidimensionalidade.

Todo o texto pré-edição e os códigos deste livro foram desenvolvidos e executados em Python no ambiente Google Colab e podem ser encontrados no TSP Book Site em <https://github.com/Introducao-Series-Temporais-em-Python/Book>. Lá você também encontrará atualizações do código, do texto e materiais complementares, bem como as bases de dados empregadas.



TSP BOOK SITE

Códigos completos. Muitos códigos ao longo do texto são inseridos apenas de modo parcial, apresentando os pontos essenciais para a compreensão da tarefa ou conceito e omitindo detalhes menos importantes, como os de edição e/ou visualização dos dados. Os códigos completos, incluindo a produção de figuras do texto, podem ser acessados diretamente no TSP Book Site.

Assim, para uma leitura mais fluida do texto e sem prejuízo para o seu entendimento, um código como este:

```

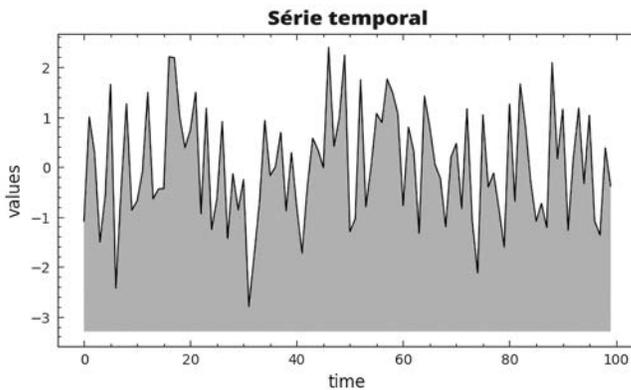
import numpy as np
from matplotlib import pyplot as plt

np.random.seed(123)
TS = np.random.randn(100)
t = [t for t in range(len(TS))]

plt.plot(TS, '-')
plt.fill_between(t, TS, TS.min()-0.5, where=(TS > TS.min()-
0.5),alpha=0.3)

plt.title("Série Temporal")
plt.xlabel('time')
plt.ylabel('values')
plt.show()

```



poderá aparecer simplesmente como:

```

TS = np.random.randn(100)
t = [t for t in range(len(TS))]

plt.plot(TS, '-')

[...] some comments

```

Você certamente preferirá acessar o código para execução do que reescrevê-lo a partir da edição impressa. O *site* também traz outros suplementos digitais, como seções de referências adicionais para cada capítulo e um capítulo extra sobre problemas encontrados no tratamento de séries de dados reais e outros modelos.

INTRODUÇÃO ÀS SÉRIES TEMPORAIS UMA ABORDAGEM PRÁTICA EM PYTHON

COMPREENDER A ANÁLISE E A PREVISÃO DE SÉRIES TEMPORAIS É CADA vez mais essencial para a tomada de decisões e para a antecipação de tendências. *Introdução às séries temporais: uma abordagem prática em Python* é um guia introdutório para aqueles que desejam dominar esses métodos nos mais variados campos, desde a Engenharia e Finanças até as Ciências Sociais e da Saúde.

Elaborado para ser um instrumento de apoio para os estudantes de Engenharia, Ciências de Dados e de Computação, entre outros, este livro adota uma abordagem didática e acessível. Os conceitos teóricos são apresentados com implementações práticas em Python, tornando o conteúdo bastante claro tanto para programadores quanto para não programadores.

Os exemplos práticos de análise e previsão de séries temporais abrangem uma ampla gama de conjuntos de dados, desde séries econômicas até dados de temperatura global e demanda por aluguel de bicicletas. Além dos modelos tradicionais, também são explorados os modelos de aprendizado de máquina, preenchendo uma lacuna importante nos textos introdutórios sobre o assunto.

Como material complementar, o leitor tem acesso a um *site* em que são disponibilizados exercícios, dados e códigos, que passam por atualizações contínuas. Aproveite essa oportunidade para aprender a análise de séries temporais em Python na prática.

Prof. Dr. Diógenes de Souza Bido

*Professor do Programa de Pós-Graduação em Administração de
Empresas da Universidade Presbiteriana Mackenzie*

