

E-Book

Do Zero à Análise de Dados Econômicos e Financeiros com Python

Um guia mão na massa para iniciar no mercado



Índice

| | |
|---|----|
| Bem vindo(a)! | 10 |
| Parte I - O ciclo da análise de dados | 11 |
| 1. Introdução | 12 |
| 2. O roteiro | 12 |
| 2.1 Objetivo | 14 |
| 2.2 Dados | 14 |
| 2.3 Exploração | 15 |
| 2.4 Modelagem | 16 |
| 2.5 Validação | 17 |
| 2.6 Implantação | 18 |
| Parte II - Programação: Python e Colab | 19 |
| 3. Por que usar o Python? | 20 |
| 3.1 O que usar para implementar o ciclo? | 20 |
| 4. Google Colab e Olá, Mundo! | 21 |
| 4.1 Ambiente de programação | 21 |
| 4.2 Google Colab | 22 |
| 4.3 Olá, Mundo! | 23 |
| 5. Bibliotecas | 24 |
| 5.1 O que são bibliotecas? | 24 |
| 5.2 Como instalar bibliotecas? | 24 |
| 5.3 Entendendo as mensagens | 24 |
| 6. Introdução ao Pandas | 25 |
| 6.1 Bibliotecas para análise de dados | 25 |
| 6.2 Pandas | 26 |
| 7. Criando scripts | 26 |
| 7.1 Regras, boas práticas e estilo | 27 |
| 7.2 Código versus comentários | 27 |



| | |
|--|-----------|
| 7.3 Estrutura de código | 27 |
| 8. Lidando com objetos | 28 |
| 8.1 Definição de objetos | 28 |
| 8.1.1 Regras de nomes | 29 |
| 8.2 Cópia e modificação | 29 |
| 8.3 Tipos de objetos | 30 |
| 9. Estruturas de dados | 31 |
| 9.1 Hierarquia de tipos | 31 |
| 9.2 Tamanho de objetos | 32 |
| 9.3 Informações de tabelas | 32 |
| 10. Operadores | 33 |
| 10.1 Operadores de comparação | 33 |
| 10.2 Operadores lógicos | 33 |
| 10.3 Operadores de aritméticos | 34 |
| 10.4 Operadores diversos | 34 |
| Parte III - Exploração de dados | 35 |
| 11. Tipos de variáveis | 36 |
| 11.1 O que é uma variável quantitativa? | 36 |
| 11.2 O que é uma variável qualitativa? | 37 |
| 11.3 Exemplo prático | 37 |
| 11.4 Conclusão | 39 |
| 12. Processo Gerador dos Dados | 39 |
| 12.1 O que é o PGD? | 40 |
| 12.2 Como identificar o PGD? | 40 |
| 12.3 Conclusão | 42 |
| 13. Estatísticas descritivas | 42 |
| 13.1 O que são estatísticas descritivas de dados? | 42 |
| 13.1.1 Dados de séries temporais | 43 |
| 13.1.2 Dados de corte transversal | 43 |
| 13.2 Fórmulas e cálculos de estatísticas descritivas | 43 |
| 13.2.1 Média | 44 |



| | |
|--|----|
| 13.2.2 Mediana | 44 |
| 13.2.3 Mínimo | 44 |
| 13.2.4 Máximo | 44 |
| 13.2.5 Percentis | 44 |
| 13.2.6 Frequência | 44 |
| 13.2.7 Desvio Padrão | 44 |
| 13.2.8 Variância | 44 |
| 13.3 Como gerar um sumário de estatísticas descritivas? | 44 |
| 13.4 Conclusão | 46 |
| 13.5 Referências | 46 |
| 14. Gráfico de Dispersão | 46 |
| 14.1 Análise de dados possíveis com gráficos de dispersão | 47 |
| 14.2 O que é um gráfico de dispersão? | 47 |
| 14.3 Como gerar gráficos de dispersão? | 48 |
| 14.4 Conclusão | 49 |
| 14.5 Referências | 49 |
| 15. Gráfico de Histograma | 49 |
| 15.1 Análise de dados possíveis com gráficos de histograma | 50 |
| 15.2 O que é um gráfico de histograma? | 50 |
| 15.3 Como gerar gráficos de histograma? | 51 |
| 15.4 Conclusão | 52 |
| 16. Gráfico de Boxplot | 52 |
| 16.1 O que é visualização de dados? | 52 |
| 16.2 Análise de dados possíveis com gráficos de boxplot | 52 |
| 16.3 O que é um gráfico de boxplot? | 53 |
| 16.4 Como gerar gráficos de boxplot? | 54 |
| 16.5 Conclusão | 55 |
| 16.6 Referências | 55 |
| 17. Gráfico de Linha | 55 |
| 17.1 O que é visualização de dados? | 55 |
| 17.2 Análise de dados possíveis com gráficos de linha | 56 |



| | |
|--|-----------|
| 17.3 O que é um gráfico de linha? | 56 |
| 17.4 Como gerar gráficos de linha? | 57 |
| 17.5 Conclusão | 58 |
| 18. Análise de Correlação | 58 |
| 18.1 O que é correlação? | 59 |
| 18.2 Prós e contras: quando usar e quando não usar correlação? | 59 |
| 18.3 Como calcular a correlação entre duas variáveis? | 60 |
| 18.4 Como interpretar o valor da correlação? | 60 |
| 18.5 Exemplo prático: correlação entre variáveis macro-financeiras | 61 |
| 18.6 Conclusão | 62 |
| 18.7 Referências..... | 62 |
| 19. Análise de Autocorrelação | 62 |
| 19.1 O que é autocorrelação? | 63 |
| 19.2 Quando usar usar a autocorrelação? | 63 |
| 19.3 Como calcular a autocorrelação entre duas variáveis? | 64 |
| 19.4 Como interpretar o valor da autocorrelação? | 65 |
| 19.5 Exemplo prático: correlação entre variáveis macro-financeiras | 66 |
| 19.6 Tendência e Sazonalidade | 67 |
| 19.7 Conclusão | 68 |
| 19.8 Referências | 68 |
| 20. Análise de Estacionariedade | 68 |
| 20.1 Processo Estocástico e Séries Temporais | 68 |
| 20.2 Componentes de uma série temporal | 68 |
| 20.2.1 Componentes | 69 |
| 20.2.1.1 Tendência | 69 |
| 20.2.1.2 Sazonalidade | 69 |
| 20.2.1.3 Ciclo e o Termo de erro | 69 |
| 20.2.1.4 Pontos importantes: | 69 |
| 20.2.1.5 Decomposição | 70 |
| 20.3 Tipos de processos | 70 |
| 20.3.1 Processos Estacionários | 71 |



| | |
|--|----|
| 20.3.2 Processos Não Estacionários | 71 |
| 20.4 Exemplos de Processos | 71 |
| 20.4.1 Ruído Branco | 71 |
| 20.4.2 Passeio Aleatório (Random Walk) | 72 |
| 20.4.3 Passeio Aleatório com drift | 72 |
| 20.5 Estacionariedade | 73 |
| 20.5.1 O que é uma Série Temporal Estacionária? | 73 |
| 20.5.2 Como verificar a estacionariedade? | 74 |
| 20.5.2.1 Abordagem Intuitiva | 74 |
| 20.5.2.2 Abordagem estatística | 75 |
| 20.5.2.3 Teste de Dickey-Fuller Aumentado | 75 |
| 20.5.2.4 Teste de Kwiatkowski-Phillips-Schmidt-Shin | 75 |
| 20.6 Teste de Estacionariedade com o Python | 76 |
| 20.7 Conclusão | 77 |
| 20.8 Referências | 77 |
| 21. Análise de Séries Temporais | 77 |
| 21.0.1 Gráfico de linha | 78 |
| 21.0.2 Gráfico de sazonalidade | 79 |
| 21.0.3 Gráfico de decomposição: sazonalidade, tendência, ruído | 81 |
| 21.0.4 Gráfico de FAC e FACP | 82 |
| 21.1 Conclusão | 83 |
| 21.2 Referências | 83 |
| Parte IV - Modelagem e Previsão | 84 |
| 22. Fluxo de trabalho | 85 |
| 22.1 Modelagem preditiva em 6 etapas | 85 |
| 22.1.1 Organizar os dados | 86 |
| 22.1.2 Visualizar os dados | 86 |
| 22.1.3 Especificar o modelo | 86 |
| 22.1.4 Estimar o modelo | 86 |
| 22.1.5 Avaliar a performance | 87 |
| 22.1.6 Produzir previsões | 87 |



| | |
|---|-----|
| 22.2 Conclusão | 87 |
| 22.3 Referências | 87 |
| 23. Exemplo aplicado | 87 |
| 23.1 Organizar os dados | 88 |
| 23.2 Visualizar os dados | 88 |
| 23.3 Especificar o modelo | 90 |
| 23.4 Estimar o modelo | 90 |
| 23.5 Avaliar a performance | 90 |
| 23.6 Produzir previsões | 92 |
| 23.7 Conclusão | 93 |
| 23.8 Referências | 93 |
| Parte V - Projeto: previsão do IPCA | 94 |
| 24. Modelos SARIMA | 95 |
| 24.1 Modelos auto-regressivos (AR) | 95 |
| 24.2 Modelos de média móvel (MA) | 97 |
| 24.3 Modelos auto-regressivos integrados de média móvel (ARIMA) | 98 |
| 24.4 Modelos sazonais auto-regressivos integrados de média móvel (SARIMA) | 99 |
| 24.5 Conclusão | 100 |
| 24.6 Referências | 100 |
| 25. Previsão do IPCA | 100 |
| 25.1 Exemplo de modelo AR(p) | 100 |
| 25.2 Exemplo de modelo MA(q) | 103 |
| 25.3 Exemplo de modelo ARIMA(p,d,q) | 106 |
| 25.4 Exemplo de modelo ARIMA(p,d,q)(P,D,Q)m | 108 |
| 25.5 Conclusão | 111 |
| 25.6 Referências | 111 |
| 26. Conclusão | 112 |
| 26.1 Como continuar aprendendo? | 112 |

