

Analytics e Inteligência Artificial

Aula 14

Projeto Parte 3 - Regressão





BUSINESS SCHOOL

Graduação, pós-graduação, MBA, Pós- MBA, Mestrado Profissional, Curso In *Company* e EAD



CONSULTING

Consultoria personalizada que oferece soluções baseadas em seu problema de negócio



RESEARCH

Atualização dos conhecimentos e do material didático oferecidos nas atividades de ensino



Líder em Educação Executiva, referência de ensino nos cursos de graduação, pós-graduação e MBA, tendo excelência nos programas de educação. Uma das principais **escolas de negócio do mundo**, possuindo convênios internacionais com Universidades nos EUA, Europa e Ásia. +8.000 **projetos de consultorias** em organizações públicas e privadas.



Único curso de graduação em administração a receber as notas máximas



A primeira escola brasileira a ser finalista da maior competição de MBA do mundo



Única *Business School* brasileira a figurar no *ranking* LATAM



Signatária do Pacto Global da ONU



Membro fundador da ANAMBA - Associação Nacional MBAs



Credenciada pela AMBA - Association of MBAs



Credenciada ao Executive MBA Council



Filiada a AACSB - Association to Advance Collegiate Schools of Business



Filiada a EFMD - European Foundation for Management Development



Referência em cursos de MBA nas principais mídias de circulação

O **Laboratório de Análise de Dados** – LABDATA é um Centro de Excelência que atua nas áreas de ensino, pesquisa e consultoria em análise de informação utilizando técnicas de **Big Data, Analytics** e **Inteligência Artificial**.



O LABDATA é um dos pioneiros no lançamento dos cursos de **Big Data** e **Analytics** no Brasil
Os diretores foram professores de grandes especialistas do mercado
+10 anos de atuação
+1000 alunos formados

Docentes

- Sólida formação acadêmica: doutores e mestres em sua maioria
- Larga experiência de mercado na resolução de **cases**
- Participação em Congressos Nacionais e Internacionais
- Professor assistente que acompanha o aluno durante todo o curso

Estrutura

- 100% das aulas realizadas em laboratórios
- Computadores para uso individual durante as aulas
- 5 laboratórios de alta qualidade (investimento +R\$2MM)
- 2 Unidades próximas a estação de metrô (com estacionamento)





Profª Dra.
**Alessandra
Montini**

Diretora do LABDATA-FIA, apaixonada por dados e pela arte de lecionar. Têm muito orgulho de ter criado na FIA cinco laboratórios para as aulas de Big Data e inteligência Artificial. Possui mais de 20 anos de trajetória nas áreas de Data Mining, Big Data, Inteligência Artificial e Analytics. Cientista de dados com carreira realizada na Universidade de São Paulo. Graduada e mestra em estatística aplicada pelo IME-USP e doutora pela FEA-USP. Com muita dedicação chegou ao cargo de professora e pesquisadora na FEA-USP, ganhou mais de 30 prêmios de excelência acadêmica pela FEA-USP e mais de 30 prêmios de excelência acadêmica como professora dos cursos de MBA da FIA. Orienta alunos de mestrado e de doutorado na FEA-USP. Membro do Conselho Curador da FIA, Coordenadora de Grupos de Pesquisa no CNPQ, Parecerista da FAPESP e Colunista de grandes Portais de Tecnologia.

 [linkedin.com/in/alessandramontini/](https://www.linkedin.com/in/alessandramontini/)



Prof. Dr.
Adolpho Walter Canton

Diretor do LABDATA-FIA. Consultor em Projetos de *Analytics, Big Data* e Inteligência Artificial. Professor FEA – USP. PhD em Estatística Aplicada pela *University of North Carolina at Chapel Hill*, Estados Unidos.



Currículo – Prof. João Nogueira

FORMAÇÃO ACADÊMICA | EXPERIÊNCIA PROFISSIONAL

5

- **(2019-Presente)** – Professor nos cursos de Extensão, Pós e MBA em Big Data e Data Mining na Fundação Instituto de Administração (FIA) – www.fia.com.br
- **(2018-Presente)** – Cientista de Dados na Via Varejo – <https://viavarejo.com.br>
- **(2016-Presente)** – Doutorando em Física Computacional e Estatística pelo Departamento de Física na Universidade Federal do Ceará – <https://fisica.ufc.br>
- **(2014-2016)** – Mestre em Física da Matéria Condensada pelo Departamento de Física na Universidade Federal do Ceará - <https://fisica.ufc.br>
- **(2012-2013)** – Estudante Intercambista na Universidade de Coimbra – Portugal – <https://www.uc.pt>
- **(2010-2014)** – Bacharel em Física pela Universidade Federal do Ceará – <http://www.ufc.br>
- **Contatos:**
 - E-mail: joaonogueira@fisica.ufc.br



Conteúdo Programático da Disciplina – Projeto de Inteligência Artificial



Data	Horário	Tema
09/03/2021	19:00	Aula 1 - Introdução ao Ambiente de Desenvolvimento
11/03/2021	19:00	Aula 2 – Revisão de Python
16/03/2021	19:00	Aula 3 – Manipulação de Dados
18/03/2021	19:00	Aula 4 – Análise Exploratória de Dados
23/03/2021	19:00	Aula 5 – Projeto da disciplina – Parte 1 – Análise Exploratória de Dados
25/03/2021	19:00	Aula 6 - Introdução, Motivação e Framework de Machine Learning
06/04/2021	19:00	Aula 7 - Analytical Base Table
08/04/2021	19:00	Aula 8 - Aprendizagem Supervisionada - Classificação
13/04/2021	19:00	Aula 9 - Aprendizagem Supervisionada - Classificação
15/04/2021	19:00	Aula 10 - Aprendizagem Supervisionada - Classificação
20/04/2021	19:00	Aula 11 - Projeto da disciplina - Parte 2 - Machine Learning - Classificação
22/04/2021	19:00	Aula 12 - Projeto da disciplina - Parte 2 - Machine Learning - Classificação
27/04/2021	19:00	Aula 13 - Aprendizagem Supervisionada - Regressão
29/04/2021	19:00	Aula 14 - Aprendizagem Supervisionada - Regressão
04/05/2021	19:00	Aula 15 - Projeto da disciplina - Parte 3 - Machine Learning - Regressão
06/05/2021	19:00	Aula 16 - Aprendizagem Não-Supervisionada
11/05/2021	19:00	Aula 17 - Aprendizagem Não-Supervisionada
13/05/2021	19:00	Aula 18 - Projeto da disciplina - Parte 4 - Machine Learning - Clusterização
18/05/2021	19:00	Aula 19 - AutoML
20/05/2021	19:00	Aula 20 – Demonstração de Deploy de Machine Learning

Conteúdo da Aula

- 1. Introdução
- 2. Aprendizagem Supervisionada - Regressão
- 3. Exercícios
- 4. Entregas e Prazos



Material das aulas

- Iremos utilizar o Google Colab para desenvolver os códigos durante as aulas.
- Acesse <https://bit.ly/tutorial-colab-projeto> para realizar o tutorial de utilização do Google Colab.



1. Introdução



1. Introdução

BASES DE DADOS

- Essa etapa do projeto da disciplina iremos aplicar os conceitos sobre Aprendizagem Supervisionada para problemas de regressão.



Icons made by [Prosymbols](https://www.flaticon.com) from www.flaticon.com

@2020 LABDATA FIA. Copyright all rights reserved.



2. Aprendizagem Supervisionada - Regressão



2. Aprendizagem Supervisionada - Classificação

BASES DE DADOS



ORDERS

- **ORDERS** - Informações sobre os pedidos realizados.

Icon made by [Srip](https://www.flaticon.com) from www.flaticon.com

@2020 LABDATA FIA. Copyright all rights reserved.



2. Aprendizagem Supervisionada - Regressão

PROBLEMA

- O departamento de Vendas do iFood fez uma solicitação para o time de Ciência de Dados para criar um modelo de **Machine Learning** que faça a previsão da receita por dia. O modelo será executado diariamente, fornecendo uma estimativa do quanto será vendido no dia seguinte.

Fonte: <https://bit.ly/3h7FcC9>



2. Aprendizagem Supervisionada - Regressão

PROBLEMA


O objetivo desse modelo é auxiliar na otimização de provisionamento e investimento de recursos tanto em termos de marketing (para os dias que se estima uma venda baixa) quanto em termos de infraestrutura (para os dias que se estima uma venda muita alta).

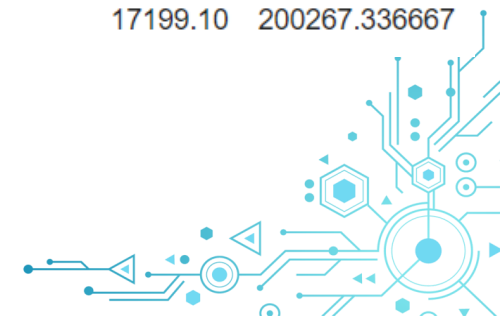


2. Aprendizagem Supervisionada - Regressão

O QUE IREMOS APRENDER?

Série Temporal como um problema de **Aprendizagem Supervisionada de Regressão**

	order_date	receita		order_date	receita	mes	dia	weekday	dayofyear	receita_atual	receita_anterior	diff_receitas	media_movel		
0	2019-06-01	209061.01		0	2019-06-01	209061.01	6	1	5	152	NaN	NaN	NaN	NaN	
1	2019-06-02	213995.65		1	2019-06-02	213995.65	6	2	6	153	209061.01	NaN	NaN	NaN	NaN
2	2019-06-03	108973.96		2	2019-06-03	108973.96	6	3	0	154	213995.65	209061.01	4934.64	NaN	NaN
3	2019-06-04	106973.23		3	2019-06-04	106973.23	6	4	1	155	108973.96	213995.65	-105021.69	177343.540000	NaN
4	2019-06-05	122957.05		4	2019-06-05	122957.05	6	5	2	156	106973.23	108973.96	-2000.73	143314.280000	NaN
5	2019-06-06	130615.30		5	2019-06-06	130615.30	6	6	3	157	122957.05	106973.23	15983.82	112968.080000	NaN
6	2019-06-07	185314.25		6	2019-06-07	185314.25	6	7	4	158	130615.30	122957.05	7658.25	120181.860000	NaN
7	2019-06-08	199144.33		7	2019-06-08	199144.33	6	8	5	159	185314.25	130615.30	54698.95	146295.533333	NaN
8	2019-06-09	216343.43		8	2019-06-09	216343.43	6	9	6	160	199144.33	185314.25	13830.08	171691.293333	NaN
9	2019-06-10	107707.36		9	2019-06-10	107707.36	6	10	0	161	216343.43	199144.33	17199.10	200267.336667	NaN



2. Aprendizagem Supervisionada - Regressão

O QUE IREMOS APRENDER?

- Realizar a **criação da ABT**
- Aplicar diferentes algoritmos de classificação
- Encontrar a melhor combinação de hiperparâmetros
- Entender os resultados obtidos
- Entender as **métricas** para os problemas de **regressão**



4. Exercícios



4. Exercícios

BASES DE DADOS

- Utilize o notebook abaixo para realizar os exercícios.



Abra o arquivo **"aula14-projeto-parte3-aprendizagem-supervisionada-regressão.ipynb"**



5. Entregas e Prazos



5. Entregas e Prazos

BASES DE DADOS

- A entrega deverá ser submetida no portal do aluno, com o upload de apenas um notebook no formato `.ipynb`, com o código executado
- Recomendo também enviar uma cópia do notebook para o email do monitor: denner.lopes86@faculdadefia.edu.br
- No campo de **Assunto** deve ser exatamente:
 - Entrega – TURMA 2 BIG DATA EAD – Projeto Parte 3 - Regressão **(obrigatório)**
- No **Corpo** do e-mail deve ser adicionado o seu **nome completo**.
- **Prazo: 02/03 até às 23:59**
- O projeto final terá um total de **100 pontos**.
- A entrega dos projetos acontecerá em 4 partes e cada parte vale **25 pontos**

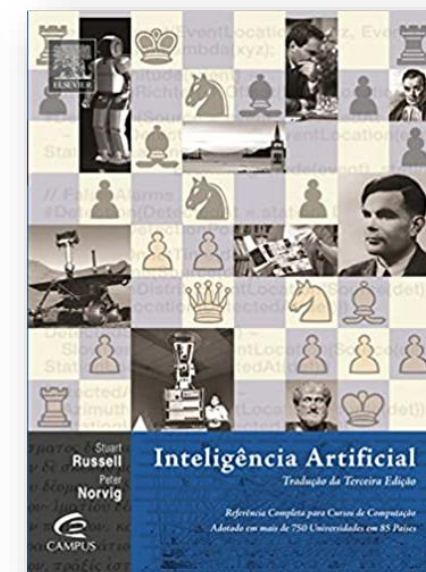
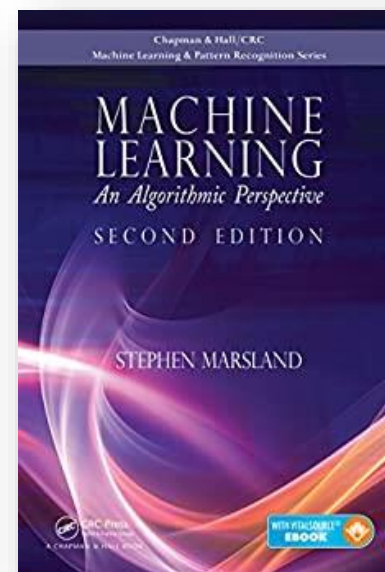
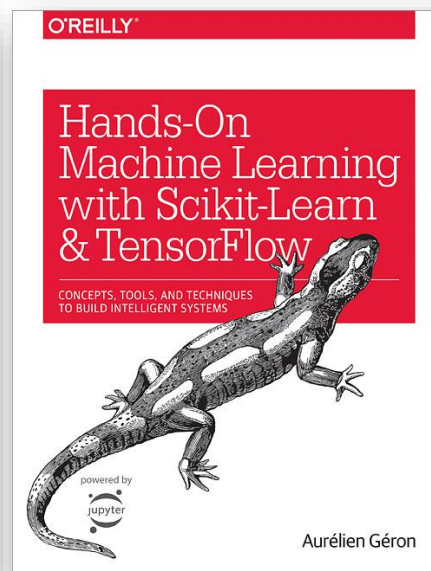
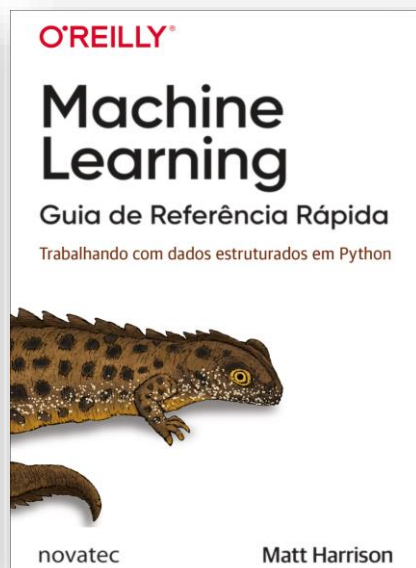
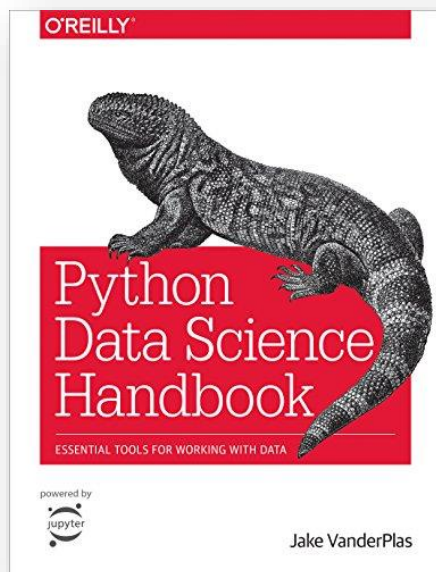


Referências Bibliográficas



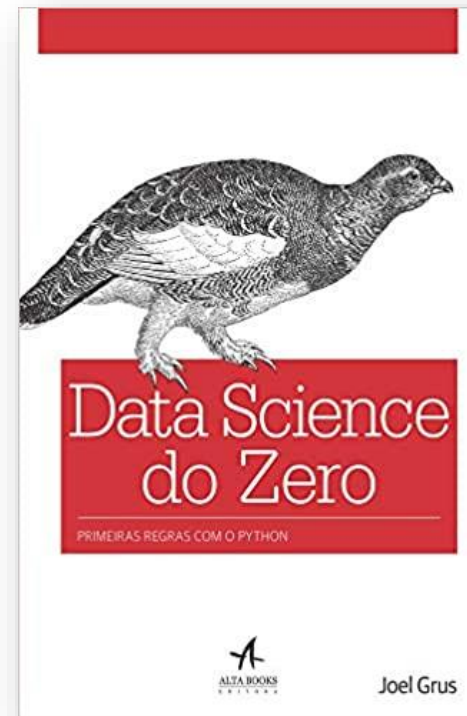
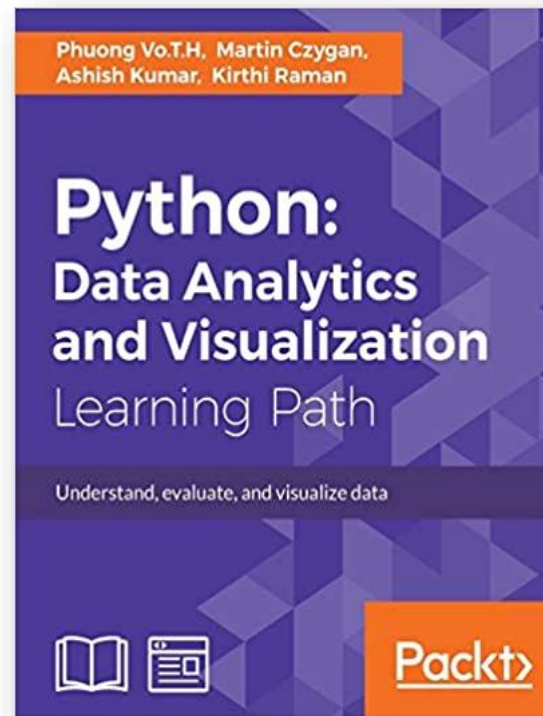
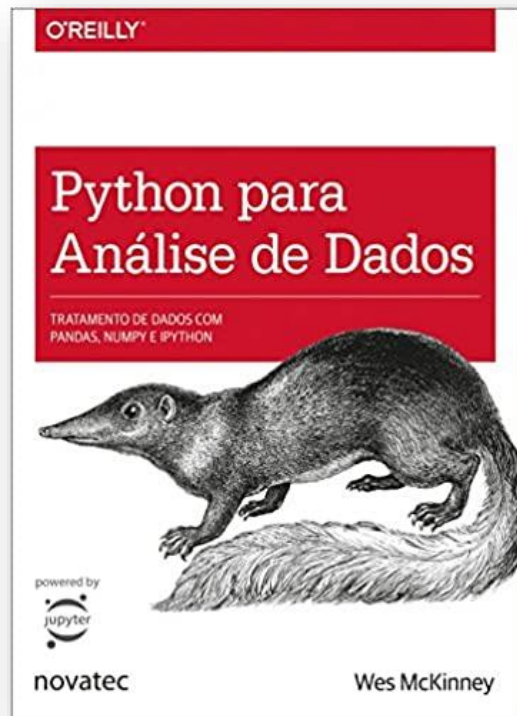
Referências Bibliográficas

LIVROS



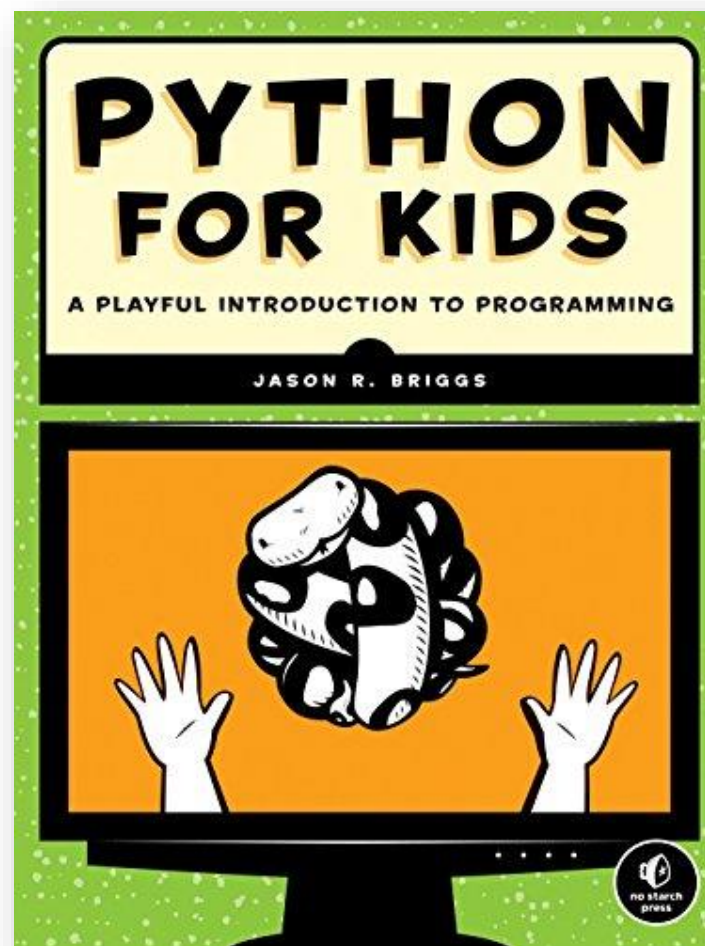
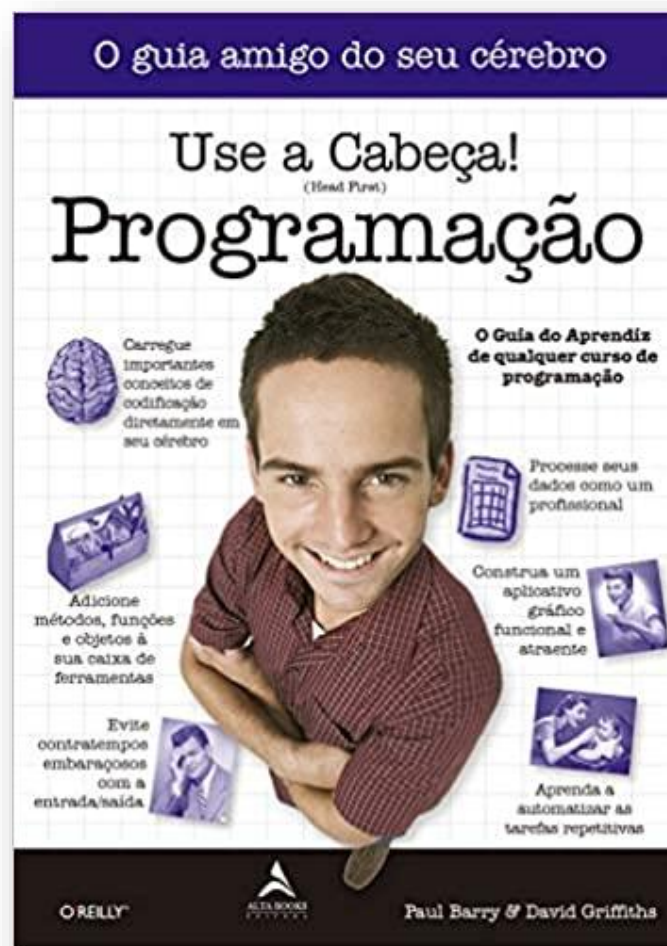
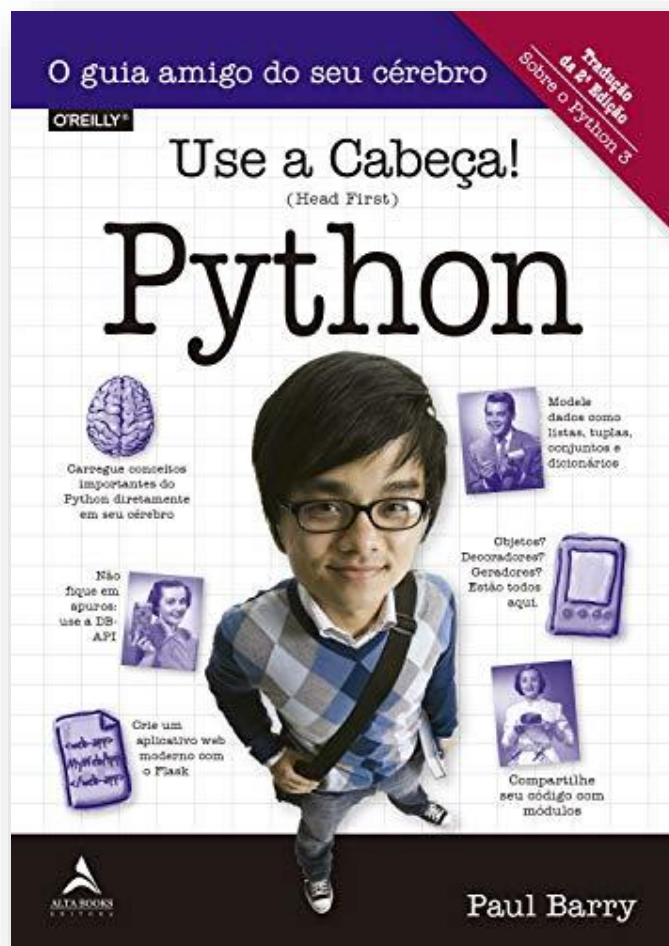
Referências Bibliográficas

LIVROS



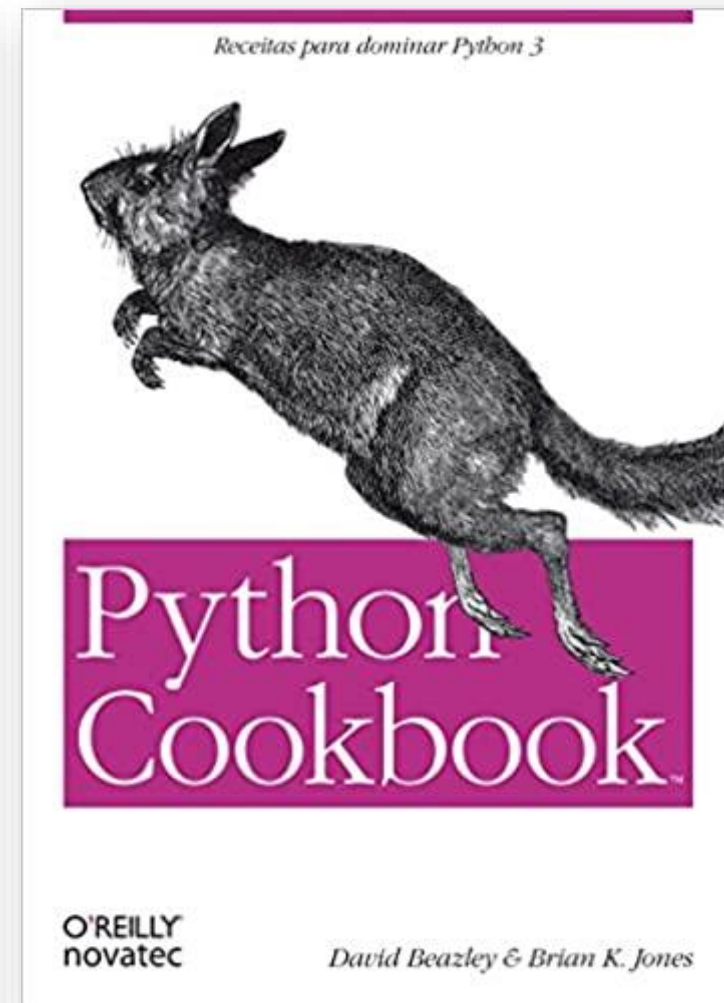
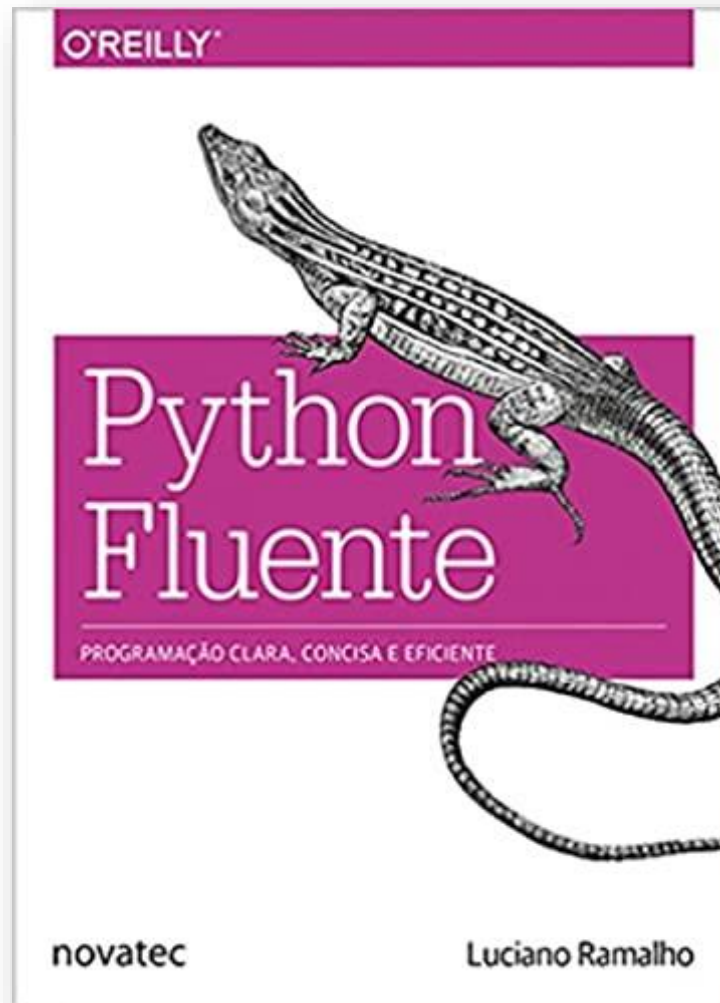
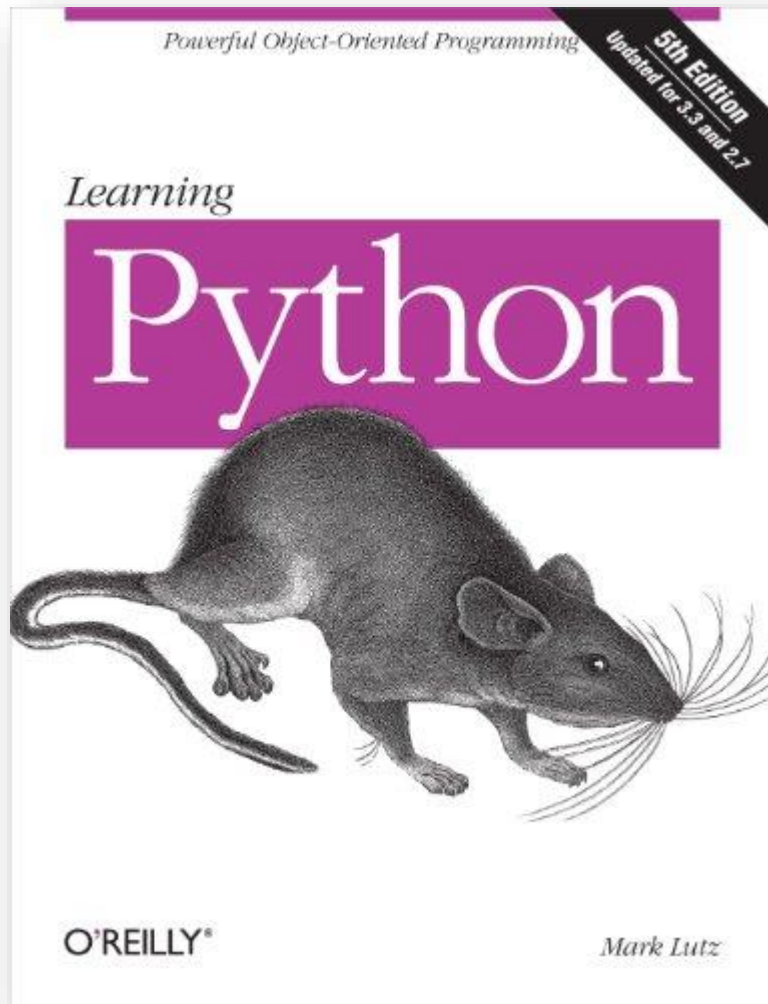
Referências Bibliográficas

LIVROS



Referências Bibliográficas

LIVROS



Referências Bibliográficas

LINKS, ÍCONES, IMAGENS

- As referências de links utilizados podem ser visualizados em <http://urls.dinomagri.com/refs>
- Tutoriais disponíveis no site oficial do Pandas - <http://pandas.pydata.org/pandas-docs/stable/>
- Livro de receitas disponíveis no site oficial do Pandas - <http://pandas.pydata.org/pandas-docs/stable/cookbook.html>
- As imagens foram Icon made by [Srip](#), [Pixel perfect](#), [Eucalyp](#) e [Prettycons](#) from www.flaticon.com

