

DATA MINING





Coordenação:

**Profa. Dra. Alessandra de
Ávila Montini**

Tema da Aula: Python para Machine Learning

Prof. Luís Moneda

Outubro – 2019

Currículo

- Bacharel em economia, USP
- Engenheiro da computação, USP
- Mestrando em ciência da computação, USP
- Data Scientist desde 2015
- Kaggle: competições e datasets
- Data Scientist no Nubank desde 2017
- Um dos organizadores do Meet-up de ML do Nubank

Contatos

- [Twitter](#)
- [Linkedin](#)
- [Site](#)
- [Kaggle](#)
- [Github](#)
- [Meetup de Machine Learning do Nubank](#)

Conteúdo da Aula

- Introdução ao Python
- Python para Data Science

Introdução ao Python

- ▼ [1 Revisão das estruturas básicas](#)
 - [1.1 Strings](#)
 - [1.2 Números](#)
 - ▼ [1.3 Listas](#)
 - [1.3.1 Acessando os elementos de uma lista](#)
 - [1.4 Dicionário](#)
 - [1.5 Booleanos](#)
 - [1.6 Operações matemáticas](#)
 - [1.7 Manipulação de strings](#)
 - [1.8 Operadores lógicos](#)
 - [1.9 Comparação](#)
 - [1.10 Condicional](#)
 - [1.11 Loops](#)
 - [1.12 Funções](#)
 - [1.13 Identação](#)
- ▼ [2 Elementos intermediários](#)
 - [2.1 List comprehension](#)
 - [2.2 PEP8](#)
 - [2.3 Ser "Pythonico"](#)
- ▼ [3 Elementos avançados](#)
 - [3.1 Testes unitários](#)
 - [3.2 Funções com multi argumentos](#)
 - [3.3 Tipagem](#)
 - [3.4 Multi processamento](#)

Python para Data Science

- ▼ 1 Numpy
 - 1.1 Vetores
 - 1.2 Operações com vetores
 - 1.3 Matrizes
- ▼ 2 Pandas
 - 2.1 DataFrame
 - 2.2 Manipulando dados em um DataFrame
 - 2.3 Carregando dados externos
 - ▼ 2.4 Análise Exploratória de Dados
 - 2.4.1 Tipos dos dados
 - 2.4.2 Dados faltantes
 - 2.4.3 Dados ao longo do tempo
 - 2.4.4 Investigando características dos dados
 - 2.5 Visualização de dados
- ▼ 3 Scikit-learn
 - 3.0.1 Regressão
 - 3.0.2 Classificação

Continuamos no Jupyter Notebook!



Referências Bibliográficas

- <https://towardsdatascience.com/how-to-be-pythonic-and-why-you-should-care-188d63a5037e>
- https://share.cocalc.com/share/741ad81231a9db8d8f83bf312458c606ddae7b1d/01_IntroStatisticsWithPython_InClass/01_IntroStatisticsWithPython.ipynb?viewer=share