

Programação Python para Ciência de Dados

Luis Gustavo Nonato

Assistentes:

Bruno Gomes Coelho

Gustavo Sutter Pessurno de Carvalho

Tobias Veiga

Conteúdo

Módulo 1:

- Introdução
- Controle de versões
- Ambiente de Programação: python 3, jupyter notebook, pacotes e github
- Python I: tipos de dados, sequências e operações, estruturas de controle e repetição
- Prática: construindo e operando listas e strings

Conteúdo

Módulo 2:

- Python II: referências, funções, funções avançadas e comprehension
- Python III: I/O e noções básicas de classes e objetos
- Prática dos conteúdos estudados: construindo um pequeno pacote para manipular arquivos texto.

Conteúdo

Módulo 3:

- Numpy I: array, slicing, fancy index, copy and view
- Pandas I: dataframes, series, index, Pandas I/O (csv, json, excell)
- Prática dos conteúdos estudados: Processando e extraindo informações de arquivos csv

Conteúdo

Módulo 4:

- Numpy II e Matplotlib: operações com array, broadcasting, construção de gráficos usuais
- Pandas II: limpeza de dados e estatísticas básicas
- Prática dos conteúdos estudados: manipulando dados do ensino médio

Conteúdo

Módulo 5:

- Scikit learn: introdução a mecanismos de clustering, regressão e classificação
- Web Crawler: request e funcionalidades associadas
- Prática dos conteúdos estudados: clusterizando e predizendo o desempenho de escolas do ensino médio

Dias das Aulas

- 21/09 - 9h as 13h20
- 28/09 - 9h as 13h20
- 05/10 - 9h as 13h20
- 09/11 - 9h as 13h20
- 23/11 - 9h as 13h20

Wifi

Username: cursopython

Password: #pyth0n#

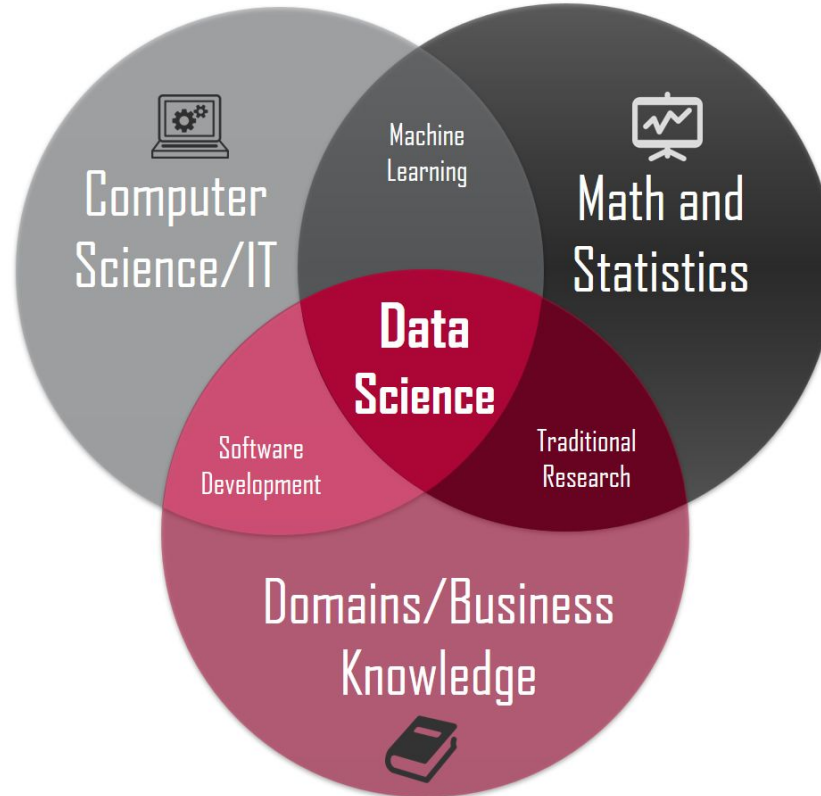
Material do Curso

<http://tiny.cc/nfvycz>

Instalação Python

- <https://www.anaconda.com/>
- <https://colab.research.google.com/notebooks/welcome.ipynb>

What is Data Science



Data Science Pipeline



Question
Hypothesis ??

