



## Python para Ciência de Dados

Aula 01

SENAI CETIQT

SENAI



## Currículo: Profa. Edileusa de Estefani do Prado

### FORMAÇÃO:

- ☐ Mestranda em Métodos Quantitativos - FEA/USP - 2019-2022.
- ☐ MBA em Analytics - Big Data - FIA - 2015/2016 - Turma 1.
- ☐ Disciplina de Banco de dados avançados - UNICAMP - 2005.
- ☐ Pós-Graduação em Consultoria em Internet - FASP - 2000/2001.
- ☐ Graduação em Ciências da Computação - UNESP - 1989/1993.

### QUALIFICAÇÕES:

- ☐ Experiência profissional de 26 anos no mercado, tendo atuado como Data Scientist, DW DBA, BI, ETL, AD e Analista de Sistemas em empresas: Banco Itaú, Equifax do Brasil, Banco ABN Amro, IBM, Serasa Experian, Unisys Brasil, e outras.
- ☐ Professora dos cursos de MBA e Pós-Graduação em Analytics e Big Data pela FIA.
- ☐ Professora do curso de Pós-Graduação em Big Data pelo Senac - SJRP - SP.
- ☐ Data Scientist Consultant

### CONTATO:

- ☐ E-mail : [edileusa.estefani@gmail.com](mailto:edileusa.estefani@gmail.com)
- ☐ LinkedIn : <https://www.linkedin.com/in/edileusa-est%C3%A9fani-prado-027a957/>

## Agenda

1. Objetivo
2. Introdução - Abordagem Gerencial
3. Instalação do Python
4. Referências Bibliográficas



## Objetivo



- ☐ Apresentar a metodologia de ensino utilizada
- ☐ Definir o que é a linguagem Python
- ☐ Comentar sobre o surgimento da linguagem Python até os dias de hoje
- ☐ Mostrar as principais características técnicas da linguagem e porque Python é considerada tão especial pela comunidade de desenvolvedores
- ☐ Comparar a utilização do Python com outras linguagens no mercado
- ☐ Elencar Vantagens e Desvantagens

## Objetivo

- ☐ Comentar sobre a Instalação do Python, o uso do Anaconda e do Google Colab.
- ☐ Comentar as referências bibliográficas utilizadas com sugestão de links e livros tanto em Português como em Inglês, direcionando para o auto estudo e aprofundamento.



## Introdução – Abordagem Gerencial

### Curiosidades

O termo python em Inglês refere-se a uma espécie de cobra grande e não venenosa. Anacondas são um bom exemplo de cobras da espécie python.

### O que é a linguagem Python?





## Introdução – Abordagem Gerencial

### O que é a linguagem Python?

“Python é uma linguagem de programação de alto nível, amplamente utilizada para programação de propósito geral, criada por Guido Von Rossum (foto ao lado) e liberada inicialmente em 20 de fevereiro de 1991.”

Foi desenvolvida pela [Python Software Foundation](https://www.python.org/).

<http://www.python.org>

<http://www.python.org.br>

Fonte - [https://en.wikipedia.org/wiki/Python\\_\(programming\\_language\)](https://en.wikipedia.org/wiki/Python_(programming_language))



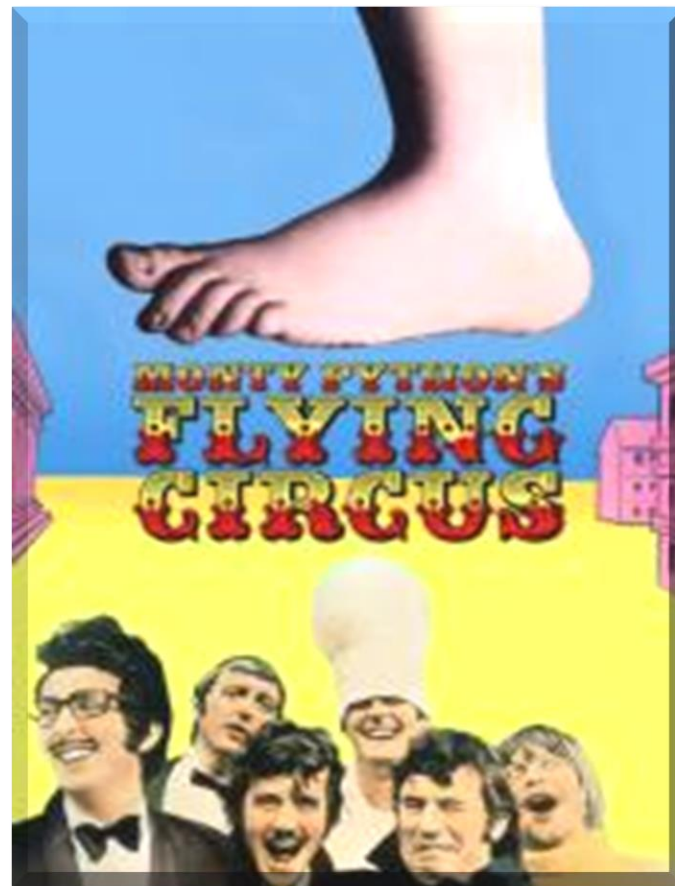
O Holandês Guido Rossum é o pai do Python

## Introdução – Abordagem Gerencial

### Um pouco de história sobre Python...

#### Curiosidades

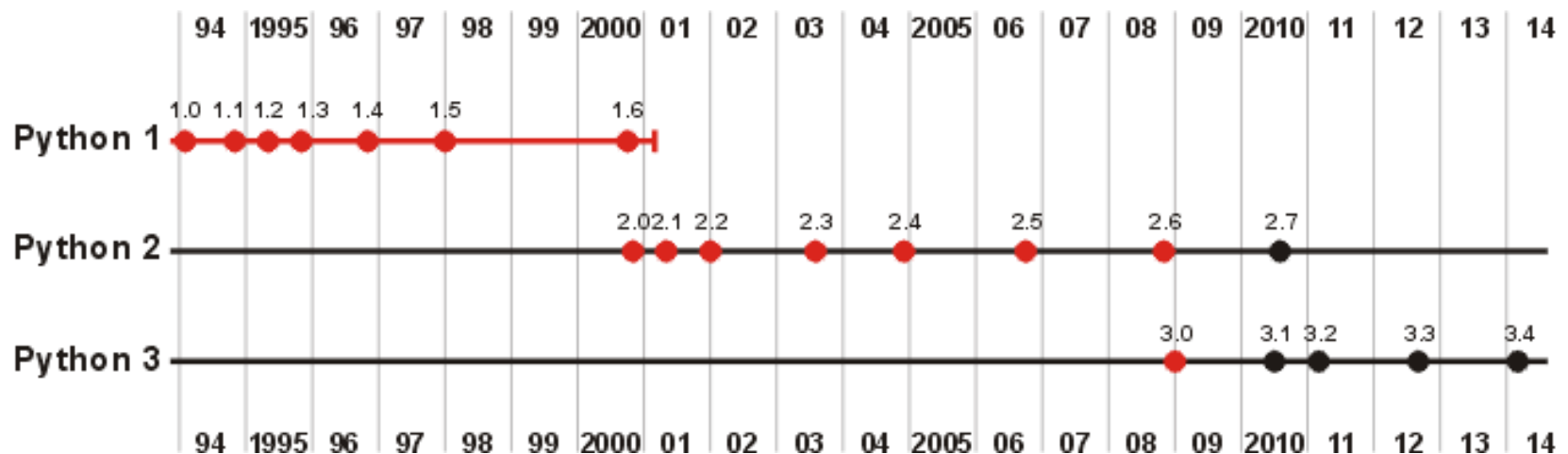
Apesar do uso de uma cobra no logo, a inspiração inicial do criador da linguagem Python foi de dar o mesmo nome que um grupo de **comédia britânico Monty Python's**.





# Abordagem Gerencial

## Desenvolvimento da linguagem Python na linha do tempo (timeline)

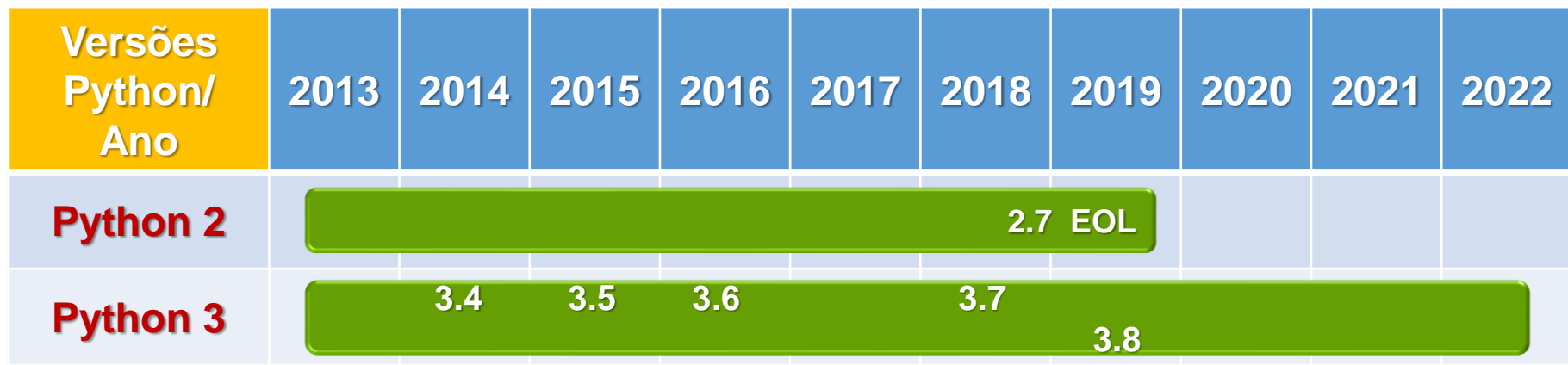


Fonte:

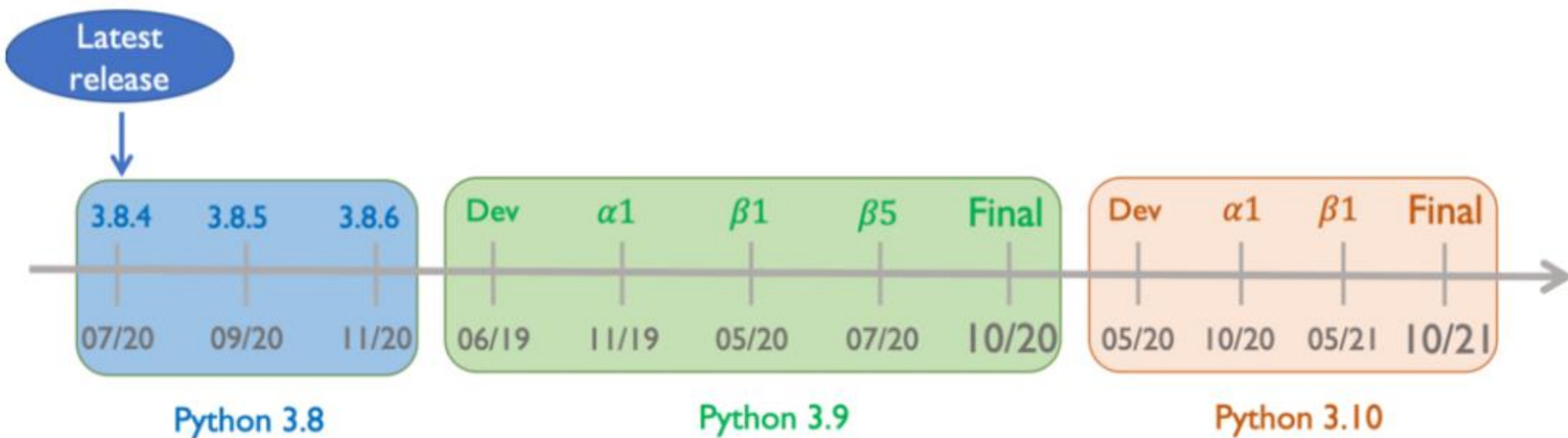
[https://es.wikipedia.org/wiki/Historia\\_de\\_Python](https://es.wikipedia.org/wiki/Historia_de_Python)

## Abordagem Gerencial

### Desenvolvimento da linguagem Python na linha do tempo (timeline)



## Desenvolvimento da linguagem Python na linha do tempo (timeline)



*Edileusa Estefani Prado*

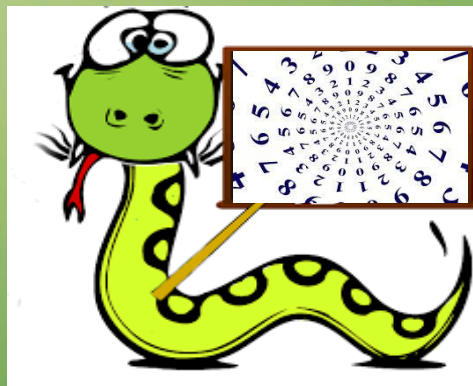
## Abordagem Gerencial

**Quão Python é utilizada  
pela comunidade de  
desenvolvedores?**



## Abordagem Gerencial

# Vamos aos números!



Editora Stefani Prada

## Abordagem Gerencial

### Python X Outras ferramentas

#### Python ultrapassa R no top 2018 Kdnuggets rankings

#### Python ultrapassa R

Python (2017) tinha mais de 50% de participação e aumentou sua participação para 66%. A participação de R diminuiu e caiu para menos de 50%.

#### RapidMiner surge

O RapidMiner, surge como a principal plataforma de Data Science aumentou sua participação p/ cerca de 50%.

#### KDnuggets Analytics, Data Science, Machine Learning Software Poll, 2016-2018

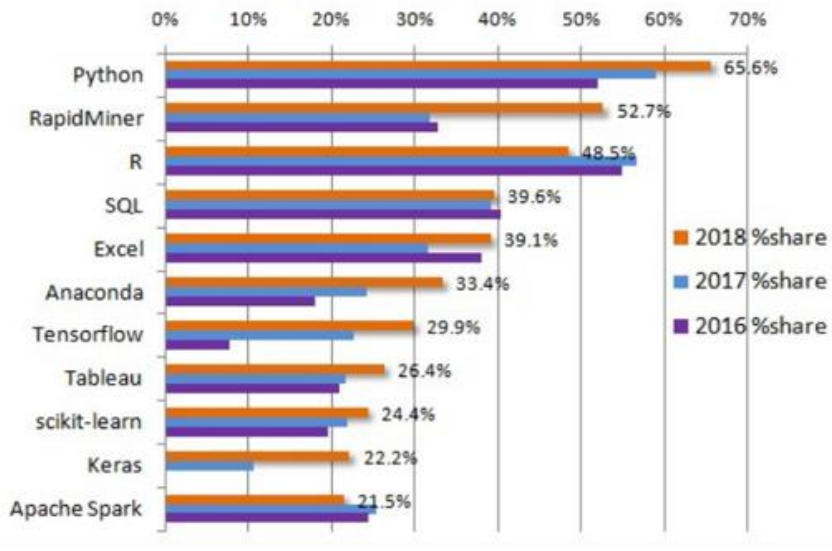


Fig 1: Análise do KDnuggets Analytics / Data Science 2018: principais ferramentas em 2018 e sua participação nas pesquisas 2016-7



## Abordagem Gerencial

### Python X Outras ferramentas

#### The 2018 Top Programming Languages from IEEE

Observa-se um declínio de R (uma linguagem específica para ciência de dados) em # 7. Python (uma linguagem de uso geral com muitos aplicativos de ciência de dados) permanece no 1º lugar.

Motivo: um contínuo desenvolvimento de aplicativos de inteligência de máquina, impulsionado pelo crescimento de iniciativas de “Big Data”.



# Abordagem Gerencial

## Python X Outras ferramentas

"Parece que há uma consolidação acontecendo no mercado de programação em ciências de dados. O Python se tornou o grande vencedor. Uma possível razão para isso é que a programação na ciência de dados está encontrando seu caminho da universidade para a indústria, onde Python é mais aceito"

May 2019	May 2018	Change	Programming Language	Ratings	Change
1	1		Java	16.005%	-0.38%
2	2		C	14.243%	+0.24%
3	3		C++	8.095%	+0.43%
4	4		Python	7.830%	+2.64%
5	6	▲	Visual Basic .NET	5.193%	+1.07%
6	5	▼	C#	3.984%	-0.42%
7	8	▲	JavaScript	2.690%	-0.23%
8	9	▲	SQL	2.555%	+0.57%
9	7	▼	PHP	2.489%	-0.83%
10	13	▲	Assembly language	1.816%	+0.82%
11	15	▲	Objective-C	1.626%	+0.69%
12	12		Delphi/Object Pascal	1.406%	+0.39%
13	18	▲	Perl	1.394%	+0.48%
14	16	▲	MATLAB	1.366%	+0.44%
15	10	▼	Ruby	1.343%	+0.16%
16	17	▲	Visual Basic	1.317%	+0.40%
17	91	▲	Groovy	1.173%	+1.06%
18	19	▲	Swift	1.150%	+0.24%

## Abordagem Gerencial

### Principais características técnicas

- ❑ Fácil de ler, escrever, depurar e expandir. Código limpo.
- ❑ Para escrever “Hello, world!” (Olá, mundo!), use apenas a função print seguido de aspas duplas ou simples.



```
print("Hello, world!")
```

# Abordagem Gerencial

## Principais características técnicas

- ❑ Programação de **alto nível**;
- ❑ Programação de **propósito geral**;
- ❑ **Linguagem interpretada** (*cada linha do código fonte é executada*);
- ❑ **Sem ponto e vírgula** no final da linha de código;
- ❑ **Código legível** (usa o recuo de espaços em branco **(4 espaços)** nos códigos fontes para delimitar um bloco de código ao invés de usar chaves ou palavras chaves.
  - ✓ **Não usa {}.**
  - ✓ **Não usa begin end;**
  - ✓ **Indentação**



## Abordagem Gerencial

### Principais características técnicas

Indentação em Python é muito importante.

Se não respeitada, dá erro.



indentation

```
*Python 3.3.4: Blockly-Demo.py - C:/Python33/Blockly-Demo.py
File Edit Format Run Options Windows Help

import weather
import math

def convert(TheStream):
    global temp, celsius, Cstream
    for temp in TheStream:
        celsius = round((temp - 32) / 1.8)
        Cstream.append(celsius)

Cstream = []
Fstream = weather.get_forecasts('Blacksburg, VA')
print(Fstream)
convert(Fstream)
print(Cstream)
```

Editeusa Stefani Prado

Ln: 17 Col: 0

# Abordagem Gerencial

## Principais características técnicas

- ❑ Usa menos linhas de código do que C ou Java.

Comparação de código Python com outras linguagens como C e Java

## Comparison with other languages

C Program	Java Program	Python Program
<pre>main() {     printf("hello, world\n"); }</pre>	<pre>class myfirstjavaprogram {     public static void main(String args[])     {         System.out.println( "Hello World!");     } }</pre>	<pre>print ("Hello World!!")</pre>





## Abordagem Gerencial

### Principais características técnicas

- ❑ Software livre e código aberto;
- ❑ Case sensitive, ou seja, uma variável denominada NOME é diferente de nome.
- ❑ Suporta múltiplos paradigmas de programação, entre os principais:
  - ✓ orientação a objetos e
  - ✓ funcional;
- ❑ Possui ampla e abrangente bibliotecas (libraries) e bibliotecas *open source*, tais como:
  - ✓ Pandas
  - ✓ NumPy
  - ✓ MatPlotLib



## Abordagem Gerencial

### Principais características técnicas

- ❑ Python executa em diversas plataformas: Linux, MAC, Windows, Android, Unix etc., 21 no total.

*Edileusa Estefani Prado*

### ...Onde Roda...

Linux – Windows – Mac – Unix – Android – iOS –  
Symbian – JVM - .NET



## Abordagem Gerencial



**7 REASONS**  
**Why You Should**  
**LEARN PYTHON**  
**Right Now**

# Abordagem Gerencial



## Porque Python é uma linguagem tão especial?

- 1) Destinada a produção de projetos empresariais de substancial importância;
- 2) Usada praticamente p/ qualquer implementação, muito versátil.
- 3) Linguagem de propósito geral, simples de ler, depurar, expandir e escrever.
- 4) Verdadeiramente universal (usado p/ atender às diversas necessidades de desenvolvimento), tais como:
  - ✓ para operações do sistema;
  - ✓ desenvolvimento web e mobile;
  - ✓ ferramentas de servidor e administrativas;
  - ✓ Implantação;
  - ✓ modelagem científica, jogos e muito mais.

## Abordagem Gerencial



### Porque Python é uma linguagem tão especial?

- 5) Google - usa a linguagem para uma série de aplicativos (Grandes players como patrocinadores). Eles ainda possuem um portal de desenvolvedores dedicado a Python , com aulas gratuitas, vídeos de conferências e muito mais.
- 6) O aumento no uso do framework Django para desenvolvimento web e um declínio na popularidade do PHP também contribuiu para o sucesso de Python.

## Abordagem Gerencial

### Porque Python é uma linguagem tão especial?



7) Python tem uma comunidade substancial no mundo, muito ativa e que oferece suporte a linguagem de forma ágil.

- ✓ Linguagens com pouca documentação e suporte fraco tendem ao fracasso.
- ✓ **Python é o contrário**, tem documentação abundante e gratuita, tutoriais, livros, cursos, etc., além de ter um ativo suporte da comunidade que o desenvolve, onde no caso de um problema, você consegue ajuda em tempo hábil.
- ✓ **Desenvolver aplicações com Python oferece menos riscos às empresas**, pois contam com um ativo suporte da comunidade, documentação abundante, muitos cases de sucesso.
- ✓ **As aplicações terão longevidade**, uma vez que as empresas podem contar com o constante desenvolvimento, aperfeiçoamento e suporte à linguagem por parte da comunidade.



## Abordagem Gerencial

**Como Python está  
sendo percebida pelo  
mercado mundial?**



*Edileusa Stefani Prado*

## Abordagem Gerencial

Como Python está sendo percebida pelo mercado mundial?

COMPANIES USING PYTHON

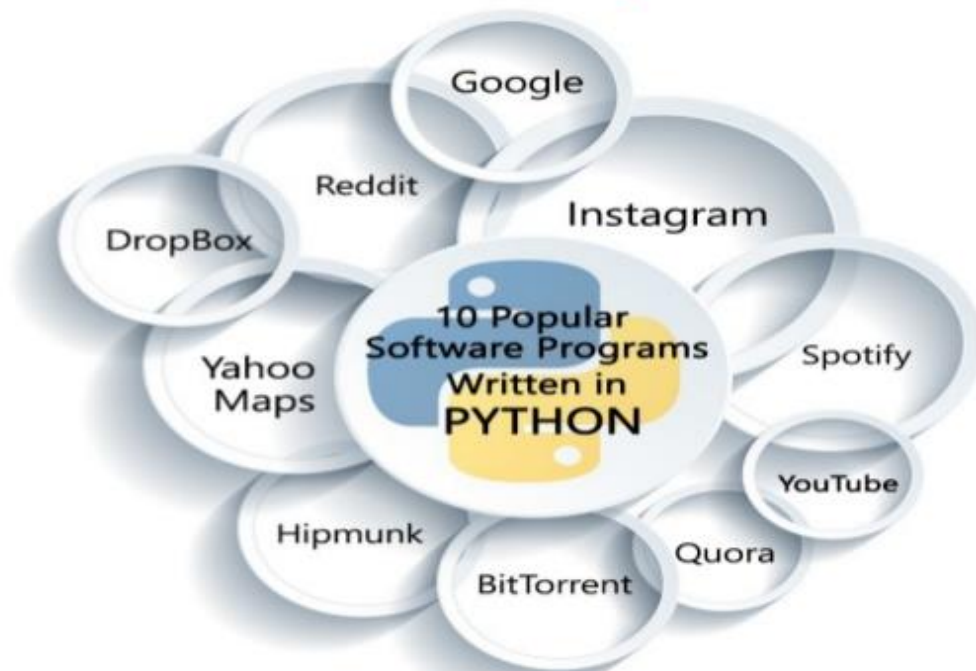


## Abordagem Gerencial

Como Python está  
sendo percebida pelo  
mercado mundial?

*Edileusa Stefani Prado*

## 10 Popular Software Programs Written in Python



## Quão Python está sendo rentável aos profissionais

[https://www.glassdoor.com.br/Sal%C3%A1rios/desenvolvedor-python-sal%C3%A1rio-SRCH\\_KO0,20\\_IP3.htm](https://www.glassdoor.com.br/Sal%C3%A1rios/desenvolvedor-python-sal%C3%A1rio-SRCH_KO0,20_IP3.htm)

glassdoor

Vagas

Avaliações de empresa

Salários

Entrevistas



Entrar

Escrever avaliação

Para Empresas

Salários



Localização

Busca

## Salários por empresa

Ordenar: Popularidade | Salários publicados | Salário

Filtrar por localização

Pesquisar por uma empresa específica

Busca

ou

Busca

Empresa

Faixa salarial base em (BRL)

Baixo

Alto

**Desenvolvedor Python Pleno:**  
**mensal**

Snowman Labs

1 salário de funcionário ou estimativa

Cerca de  
R\$ 4 mil - R\$ 5 mil

R\$ 4 mil

R\$ 5 mil

**Desenvolvedor Python Sênior:**  
**prestador de serviços mensal**

TheAuberginePanda

1 salário de funcionário ou estimativa

Cerca de  
R\$ 9 mil - R\$ 10 mil

R\$ 9 mil

R\$ 10 mil

**Desenvolvedor Python Sênior:**  
**mensal**Trade Force Serviços EM Tecnologia  
SA

1 salário de funcionário ou estimativa

Cerca de  
R\$ 6 mil - R\$ 7 mil

R\$ 6 mil

R\$ 7 mil

**Desenvolvedor Python Sênior:**  
**prestador de serviços mensal**

Nexus Edge

1 salário de funcionário ou estimativa

Cerca de  
R\$ 8 mil - R\$ 8 mil

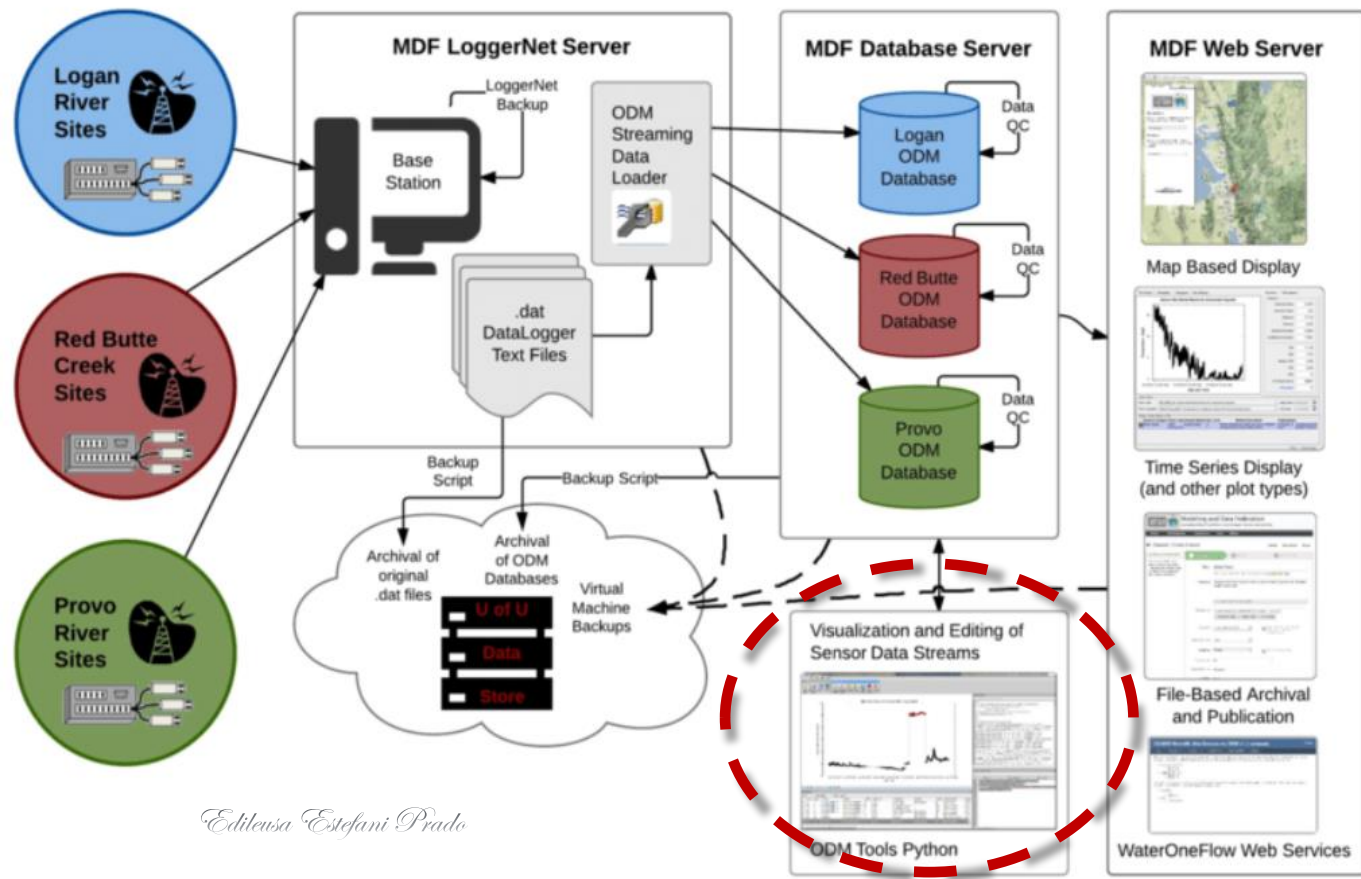
R\$ 8 mil

R\$ 8 mil

## Arquitetura em Python

Gerenciamento e compartilhamento automatizado de dados dentro de uma Rede de Sensores Heterogêneos de Grande Escala

[https://www.researchgate.net/publication/263928853\\_Automating\\_Data\\_Management\\_and\\_Sharing\\_within\\_a\\_Large-Scale\\_Heterogeneous\\_Sensor\\_Network](https://www.researchgate.net/publication/263928853_Automating_Data_Management_and_Sharing_within_a_Large-Scale_Heterogeneous_Sensor_Network)

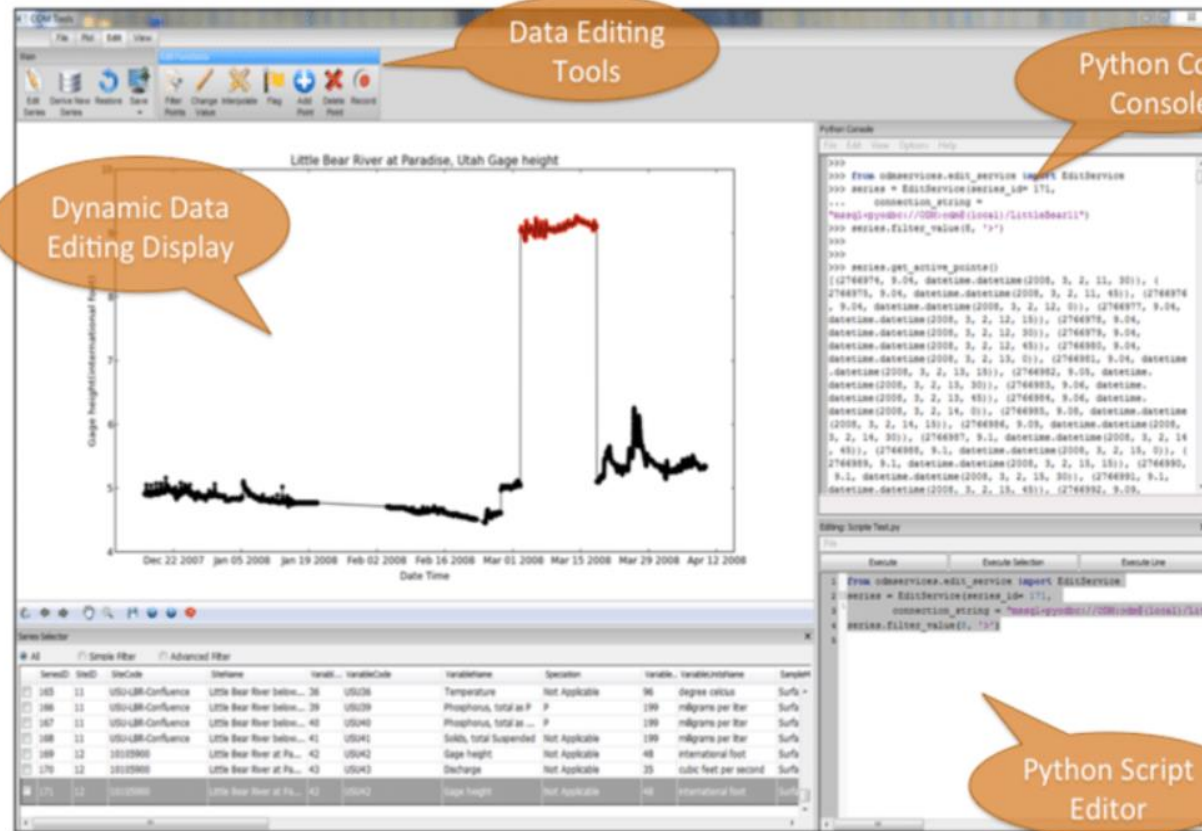


*Edileusa Estefani Prado*

## Arquitetura em Python

Gerenciamento e compartilhamento automatizado de dados dentro de uma Rede de Sensores Heterogêneos de Grande Escala

[https://www.researchgate.net/publication/263928853\\_Automating\\_Data\\_Management\\_and\\_Sharing\\_within\\_a\\_Large-Scale\\_Heterogeneous\\_Sensor\\_Network](https://www.researchgate.net/publication/263928853_Automating_Data_Management_and_Sharing_within_a_Large-Scale_Heterogeneous_Sensor_Network)



Captura de tela da interface gráfica do usuário do ODM Tools Python mostrando o editor de scripts Python e o console Python.

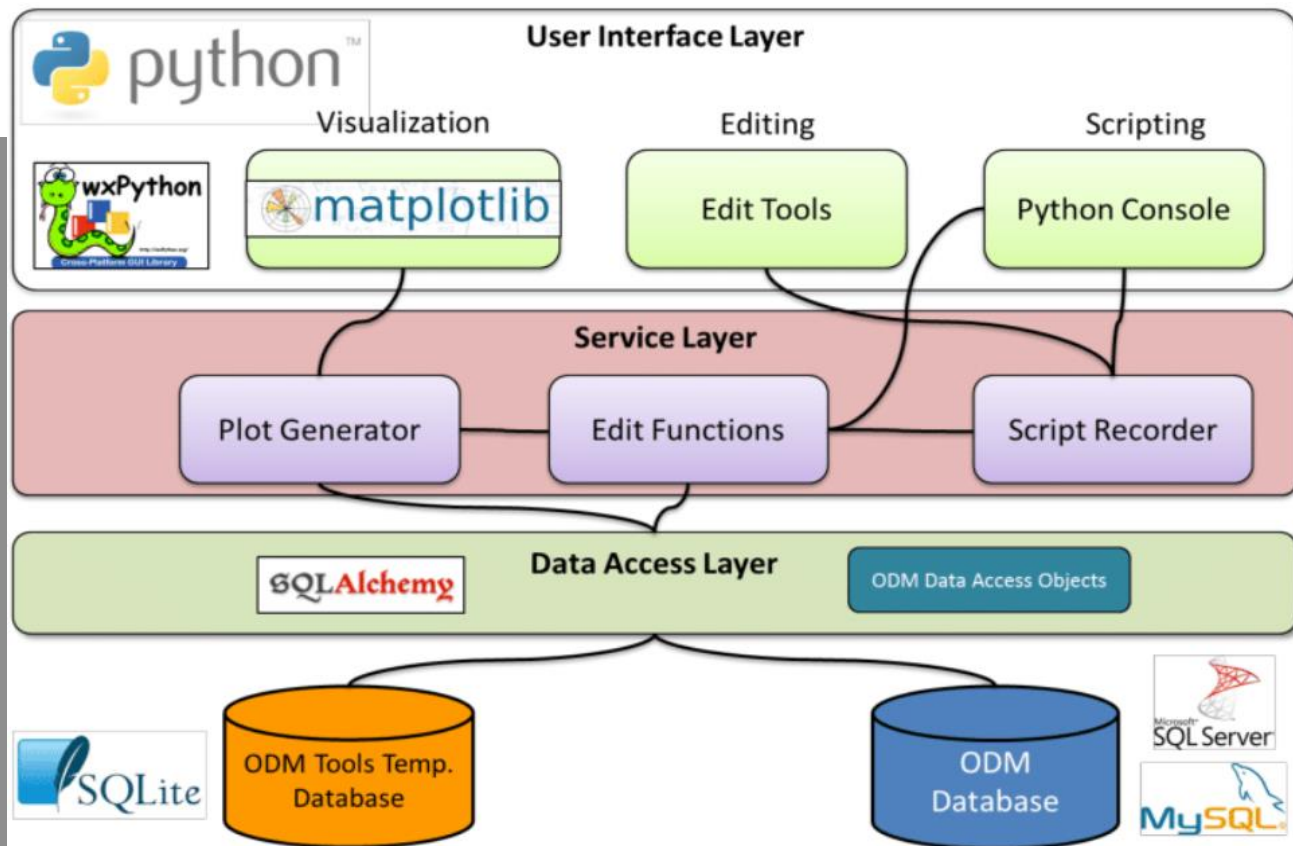
*Editeusa Estefani Prado*



## Arquitetura em Python

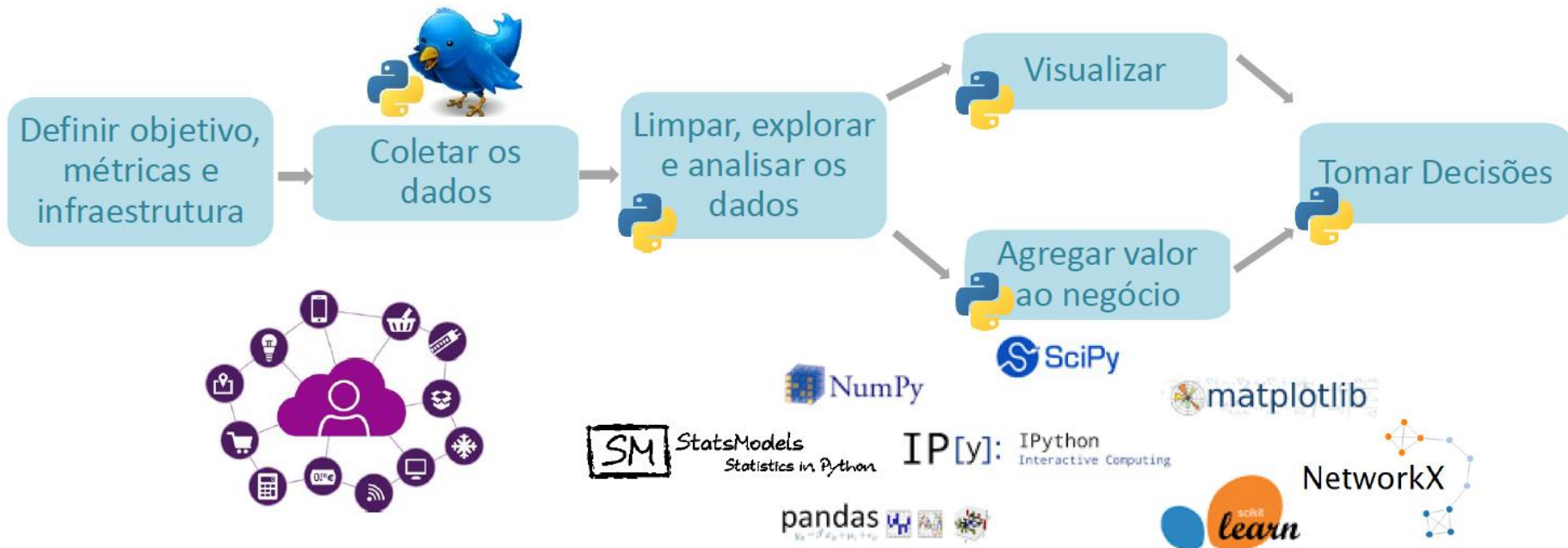
Gerenciamento e compartilhamento automatizado de dados dentro de uma Rede de Sensores Heterogêneos de Grande Escala

[https://www.researchgate.net/publication/263928853\\_Automating\\_Data\\_Management\\_and\\_Sharing\\_within\\_a\\_Large-Scale\\_Heterogeneous\\_Sensor\\_Network](https://www.researchgate.net/publication/263928853_Automating_Data_Management_and_Sharing_within_a_Large-Scale_Heterogeneous_Sensor_Network)



Arquitetura do software Python do ODM Tools.

## Por que Python para Big Data?



## Referências Bibliográficas

### Sites:

- <https://developers.google.com/edu/python/>
- <https://www.kdnuggets.com/2017/07/6-reasons-python-suddenly-super-popular.html>
- <http://letzgrou.net/blog/creating-ai-using-python/>
- <https://dzone.com/articles/which-are-the-popular-languages-for-data-science>

### Livros:

- Curso Intensivo de Python, Novatec, Eric Matthes
- Introdução à Programação com Python, Novatec, Nilo Ney Coutinho Menezes, 2ª Revisão.
- Pense em Python, Novatec, Allen B. Downey
- Learn to Program with Python, Apress, Irb Kalb
- Data Science Projects with Python: A case study approach to gaining valuable insights from real data with machine learning, 2nd Edition, Packt Publishing, Stephen Klosterman
- Projetos de ciência de dados com Python: Abordagem de estudo de caso para a criação de projetos de ciência de dados bem-sucedidos usando Python, pandas e scikit-learn, Novatec, Stephen Klosterman