

Loyanne Repolho

Healthdata
Bootcamp
November 10th-13th

PYTHON DJANGO

Organizers



Supporters

Google Cloud



Sobre mim

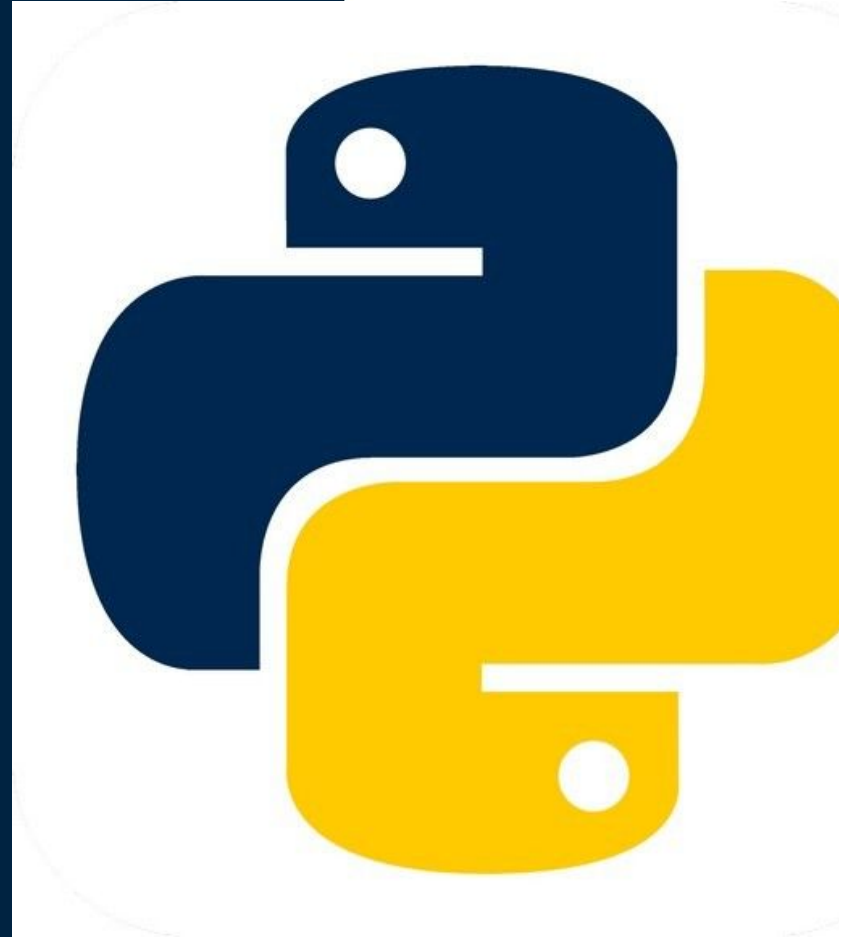
Loyanne, mulher, 31 anos, brasileira, manauara com orgulho, é médica com especialização em Ciência da Computação e Desenvolvimento de Software pela National College of Ireland, onde desenvolveu uma habilidade para Alexa voltada para pacientes com Hipertensão Arterial. Com uma mente inquieta e disruptiva, ela é uma entusiasta em Inteligência Artificial, Ciência de Dados, Aprendizado de Máquina, Design Thinking e tudo o que envolve a área de saúde e tecnologia. Sempre que possível, ela trabalha na comunidade com contribuições de código aberto e incentiva ativamente outras mulheres a se tornarem ainda mais interessadas em tecnologia.



Loyanne Repolho | 01

Python é para todos

- Python é uma linguagem de programação de alto nível, interpretada, script, imperativa, orientada a objetos, funcional, de tipagem dinâmica e forte.
- Open source
- Multi paradigmas, extensível, muitas bibliotecas
- Crianças e Adultos
- Profissionais de TI e outras profissões, como médico, administrador, engenharia, contador, farmacêutico, cientistas, etc.
- Desenvolvedores e para o pessoal de Infra
- Estudantes nível fundamental e universitários.



Loyanne Repolho | 02

Python é usado em quais áreas?

- Desenvolvimento Web
- Inteligência Artificial
- Computação Gráfica
- Enquadramento de Testes
- Big Data
- Scripting e Automação
- Ciência de Dados

Loyanne Repolho | 03

Python em Nuvem

Loyanne Repolho | 03

Python em Nuvem

Quando os conjuntos de dados são grandes, os usuários da nuvem podem fazer spool em grandes tarefas em hardware alugado capaz de fazer todo o trabalho muito mais rápido.

Agora você pode apertar o botão, criar dúzias de instâncias de nuvem carregadas com toneladas de memória e assistir seu código falhar em alguns minutos, economizando tempo e, não raro, também dinheiro.

Existem perigos também. O maior é a preocupação amorfa com a privacidade.

Algumas análises de dados envolvem informações pessoais de pessoas que confiaram em você para protegê-las. Estamos acostumados às questões de segurança envolvidas no bloqueio de dados em um disco rígido em seu laboratório. É difícil saber exatamente o que está acontecendo na nuvem.

Tanto que hoje temos a HIPAA nos EUA que é responsável pela regulação desses dados.

Loyanne Repolho | 03

A cada dia a complexidade para fazer Data Science. Isso é impulsionado por três fatores principais:

1. Aumento na geração de dados
2. Baixo custo de armazenamento de dados
3. Poder computacional barato

Loyanne Repolho | 03

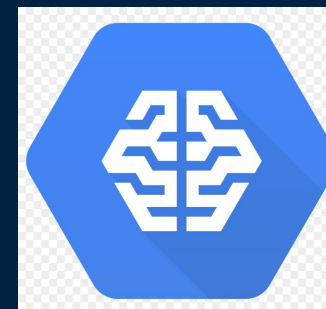
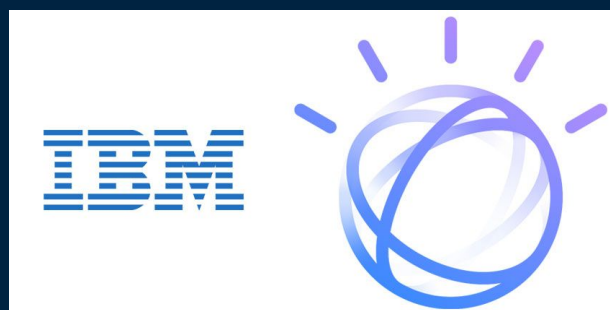
Qual o motivo de rodar na nuvem?

1. Necessidade de executar soluções que ganham escala
2. Custo
3. Colaboração
4. Compartilhamento
5. Amplo ecossistema para implantar sistemas de aprendizado de máquina
6. Construção rápida de protótipos

Loyanne Repolho | 03

Opções para rodar Data Science na nuvem:

- 1. Amazon Web Services (AWS)
- 2. Machine Learning Azure
- 3. Google Cloud Machine Learning Engine
- 4. IBM Watson Studio



Loyanne Repolho | 02

Desafios na execução de Data Science na nuvem:

- 1. Relutância em compartilhar dados com terceiros
- 2. Necessidade de fazer upload/download de grandes quantidades de dados

Loyanne Repolho | 02

Django para Data Science

É um framework para desenvolvimento web utilizando o Python, antes disso precisamos entender sobre o Model Template View(MTV). O MVT é a arquitetura utilizada pelo Django.

django

Loyanne Repolho | 02

Dicas de escalabilidade e eficiência

- Qualidade e Limpeza de Dados
- Armazenamento de Dados
- Evitar memory leaking

Loyanne Repolho | 03

Pandas

Series X Dataframe

É como um array unidimensional, uma lista de valores. Toda Series possui um índice, o index, que dá rótulos a cada elemento da lista.

```
['Doe, John', 54, 23.5, 0, 1, 115]
```

Já um DataFrame é uma estrutura bidimensional de dados, como uma planilha.

```
[
    ['Doe, Jane', F, 54, 23.5, 0, 1],
    ['Doe, John', M, 56, 34.2, 0, 0],
    ['Doe, Jenny', F, 11, 77.0, 1, 1],
    ['Doe, Jake', M, 17, 28.3, 0, 1],
    ...
]
```



Loyanne Repolho | 04

Series

	apples
0	3
1	2
2	0
3	1

+

Series

	oranges
0	0
1	3
2	7
3	2

=

DataFrame

	apples	oranges
0	3	0
1	2	3
2	0	7
3	1	2

Loyanne Repolho | 05

```
height_in_meters = (dataframe['Height'] / 25.4 ) / 1000  
weight_in_kg = dataframe['Weight'] / 2.2  
dataframe['BMI'] = weight_in_kg / (height_in_meters ** 2)
```

```
us_cdi_data['YearsDelta'] = us_cdi_data['YearEnd'] - us_cdi_data['YearStart']
```

Loyanne Repolho | 04

A computação em nuvem deve continuar ganhando espaço por conta dos benefícios que oferece, e é apenas uma questão de tempo até que alguns destes serviços se tornem o padrão (se já não forem). Espero que você ache estes serviços úteis e que te ajudem quando você precisar deles.

Organizers



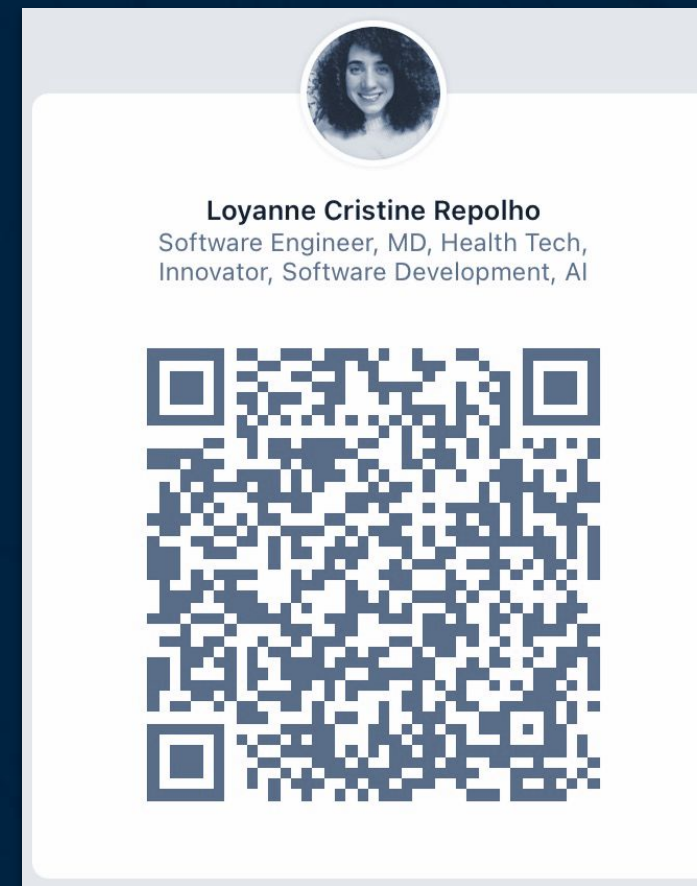
Supporters



Obrigada!

Loyanne Cristine da Costa Repolho
<https://loyannec.github.io/>
loyanne.cristine@gmail.com

Social Media: loyannec



Organizers



Supporters

