



## ABERTURA

**DR. RENATO A. CORRÊA DOS SANTOS  
Biólogo UBIRATAN BATISTA**

**E COMISSÃO ORGANIZADORA**

**V WORKSHOP DE PYTHON  
PARA DADOS BIOLÓGICOS**

**26 DE SETEMBRO DE 2022**



**Bioinformática  
Uma imensidão de  
dados!  
Perl**

Dr. rer. nat. Diego M. Riaño Pachón

2013

# Motivação



<https://www.lsnogueira.com.br/projeto/centro-nacional-de-pesquisa-em-energia-e-materiais-cnpeem/>



MINISTÉRIO DA  
CIÊNCIA, TECNOLOGIA,  
INOVAÇÕES E COMUNICAÇÕES

GOVERNO  
FEDERAL



# Motivação

2013, **2014,**

**Bioinformática**  
**Uma imensidate de**  
**dados!**  
**Perl**

Dr. rer. nat. Diego M. Riaño Pachón



Raniere Gaia Silva





Bioinformática  
Uma imensidão de  
dados!  
Perl

Dr. rer. nat. Diego M. Riaño Pachón



Raniere Gaia Silva

# Motivação

2013, 2014, 2015, **2016**,



“An Introduction to  
Programming for  
Bioscientists: A  
Python-Based Primer”

Berk Ekmekci , Charles E.  
McAnany , Cameron Mura



Dr. Cameron Mura

Introduction to  
coding, via Python,  
with concrete  
examples and  
exercises





Dr. rer. nat. Diego M. Riaño Pachón



Raniere Gaia Silva

Bioinformática  
Uma imensidão de  
dados!  
Perl

# Motivação

2013, 2014, 2015, 2016, **2017**



“An Introduction to  
Programming for  
Bioscientists: A  
Python-Based Primer”

Berk Ekmekci , Charles E.  
McAnany , **Cameron Mura**



Dr. Cameron Mura



Dra. Flávia V. Winck





# I Workshop de Python para Dados Biológicos

2013, 2014, 2015, 2016, **2017**



Raniere Gaia Silva



Dr. rer. nat. Diego M. Riaño Pachón



Dra. Flávia V. Winck



- Alisson Hayasi da Costa
- Bruno Ferreira de Souza
- Ana C. Martins Ciconelle
- Franciele Grego Esteves
- Renato Augusto Corrêa dos Santos
- Leonardo Martins-Silva



# Breve história do workshop

- Início em 2017 (Instituto de Química da USP - São Paulo)
- Organizado por estudantes (graduação e pós-graduação)
- Apoiado por docentes e pesquisadores
- Língua portuguesa
- Público-alvo
  - Pesquisadores na área de biologia, com pouco ou nenhum conhecimento de programação (dando também oportunidade a graduandos - mesmo aqueles com pouco resultado de pesquisa)
- Ensino de conceitos introdutórios de programação, usando dados biológicos da “vida real”
- Python
  - Linguagem com sintaxe de leitura “fácil”
  - Amplamente usada para automatizar tarefas e na bioinformática
  - Popular entre os organizadores



# I Workshop de Python para Dados Biológicos (2017)





# II Workshop de Python para Dados Biológicos (2018)





# III Workshop de Python para Dados Biológicos (2020)



III Workshop de Python para Dados Biológicos

8 a 11 de Setembro de 2020

Aplicação de Python na análise de dados biológicos!



IV Workshop de Python para Dados Biológicos também foi online

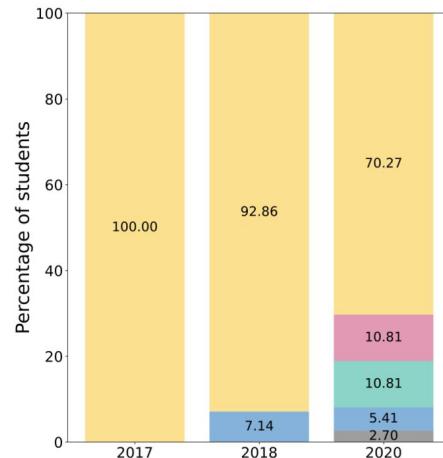


# Impacto do workshop ao longo do tempo

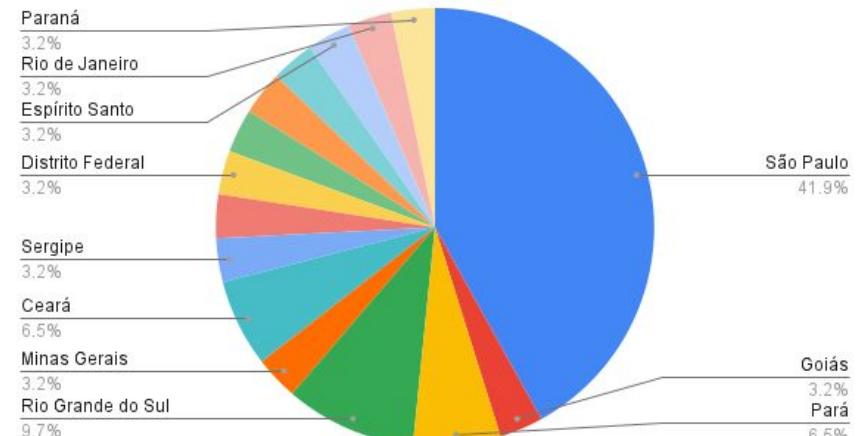
- Aproveitamento de habilidades transferíveis entre os participantes
- Desenvolvimento de um ambiente de pesquisa
  - Leitura e discussão de artigos (Twitter & Instagram; aplicação prática)
  - Escrita de projeto ao comitê de ética e de artigo científico
  - Habilidades no ensino entre os participantes
  - Incentivo à ciência em pensamento computacional e reprodutibilidade



# Inclusão de pessoas em outras regiões



Count of Região do país



Zuvanov et al. (2021)



# Palestras ao longo dos anos

Exemplificar a aplicação de programação em análises biológicas mais complexas

Edition	Talks	
	Keynote speakers	Title
2017	Profa. Dr. Flávia Vischi Winck	Biological problems and computational solutions
	Prof. Dr. Diego Mauricio Riaño Pachon	
2018	Dr. Anally Ribeiro da Silva Menegasso	Applications of Proteomic Analysis
	-	Corporate scientist. How to navigate in this world?
	Prof. Dr. Renato Vicentini dos Santos	Systems Biology and Genetic Regulation Networks
	Dr. Marcelo Falsarella Carazzolle	Applications of Bioinformatics in Biotechnology
	-	Automatization in Biotech
	Prof. Dr. Ricardo Cerri	A Framework using Machine Learning to identify and classify transposable elements
	-	Using Python for linear regression from Molecular Markers data (SNPs)
2020	Prof. Dr. Diego Mauricio Riaño Pachon	Python application in biological data: the example of the "PloidyNGS" tool
	Dr. Andréa T. Thomaz	Python in evolutionary studies: from simulations to data analysis
	Prof. Dr. Rommel Ramos	Evaluation metrics of genome assembly with Python

# Pesquisa atual (2022)

## 1<sup>a</sup> autoria compartilhada

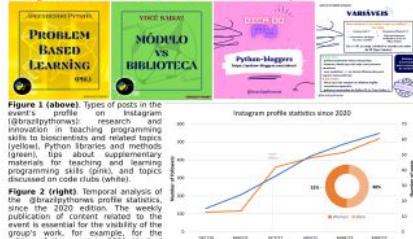


- Redes sociais
- Construção colaborativa do evento e comitê de ética
- Desenvolvimento de Code Club de programação em Python

**Experiences of the Brazilian Python Workshop for Biological Data in 2022**  
Gustavo Schiavone **Crestana**<sup>1</sup>, Michelli Inácio Gonçalves **Funnicelli**<sup>2</sup>, Ubiratan da Silva **Baptista**<sup>3</sup>, Balissa Meto de **Sousa**<sup>4</sup>, Pedro Henrique **Narciso Ferreira**<sup>5</sup>, Matheus **Scacetti**<sup>6</sup>, Leonardo André Mumbach da **Silva**<sup>7</sup>, Víncius **Franceschini-Santos**<sup>8</sup>, Cláudia Barros **Monteiro-Vitorello**<sup>9</sup>, Larissa Graciano **Braga**<sup>10</sup>, Diego Maurício **Riano-Pachón**<sup>11</sup>, Renato Augusto Corrêa dos **Santos**<sup>12</sup>

<sup>1</sup>ESALQ, University of São Paulo (USP), <sup>2</sup>School of Agricultural and Veterinary Sciences, São Paulo State University (UNESP), <sup>3</sup>Federal University of Ouro Preto (UFOP), <sup>4</sup>Federal University of Pará (UFPa), <sup>5</sup>Federal University of São Carlos (UFSCar), <sup>6</sup>University of Campinas (UNICAMP), <sup>7</sup>University of Vale do Rio dos Sinos (UNISINOS), <sup>8</sup>Faculty of Philosophy, Sciences and Letters of Ribeirão Preto, University of São Paulo (USP), <sup>9</sup>CENA, University of São Paulo (USP)  
Correspondence: renata.santos@gmail.com (R.A.C.S), gustave.crestana@usp.br (G.S.C), michelli.funnicelli@unesp.br (M.I.G.F), ubir1di@gmail.com (U.S.B)

### Using social media to increase our social impact



**Figure 1 (above):** Types of posts in the event's profile on Instagram (@brasiliapython) research and teaching programming skills to biologists and related topics (e.g., Python modules, code snippets, tips about supplementary materials for teaching and learning programming concepts and topics discussed on code clubs (white)).

### Encouraging collaborative construction of research project using Python

To conduct the research in 2021, the participants were invited to participate in the research through forms. Those who agreed to participate answered feedback questionnaires with objective and qualitative questions at the beginning and end of the workshop. In order to characterize the target audience in general aspects of the participants, as well as to identify difficulties, positive points and suggestions for future events, data was collected and analyzed by members of the organizing committee (manuscript in preparation). For the analysis of the anonymized data, fundamental and basic Python programming concepts were employed by the current organizing committee to generate didactic graphical figures, depending on the type of question asked: the type of data obtained and the Python library used (Table 1).

As a new feature for the next edition of our workshop, scheduled to take place in September of 2022, selected participants will be able to make voluntary donations during the course of the event. The money will be donated in full to a Brazilian educational entity committed to teaching programming skills, to increase diversity and promote the inclusion of minority groups in science.

**Table 1:** Graphic output generated by the analysis of collected data from questionnaires submitted by participants of the 2021 workshop edition.

Question Type	Data Type	Used Library	Graphic Output
Objective	Nominal	Basic Functions/ Matplotlib	Donut/ Bar Chart
Discursive	Qualitative Ordinal		Stacked Bar Char

### Organizing code clubs to engage and empower the team

Given our objective of promoting basic programming skills for biologists, the code clubs meetings organized in 2022 as a means to empower them to engage the members of the organizing committee and give the members an experience that make them comfortable with teaching during the workshop. Topics covered in code clubs comprise the basic programming concepts and tasks, using the Python language (Table 2). Considering that organizers have distinct backgrounds and knowledge about programming computers, and come from different institutions, the meetings also contribute to equality.

**Table 2:** Topics implemented for code club and their relationship with the content of the edition described by Zuvanov et al. (2021) (\*).

Topics	Title	Description	Social Media	*Previous workshop content
1 and 2	Variable and operators <sup>1,2</sup>	Arithmetic, assignment comparison and logical operators	147/128	Box 1/First day
3	Strings <sup>1,2</sup>	Indexing, methods and functions applied to strings	257	Box 2/Fifth day
4	Sequences <sup>1,2</sup>	Tuples, lists, ranges	186	Box 2/Fifth day
5	Control Flow Statements <sup>1,2</sup>	If, elif, and else	144	Box 2/Fifth day
6	Sets <sup>1,2</sup>	Structure, methods and functions applied to Sets	169	Box 2/Fifth day
7	Control Structures Repetition <sup>1,2</sup>	for and while	-	Box 1/Second day
8	Dictionaries <sup>1,2</sup>	Structure, methods and functions applied to Dictionaries	-	Box 1/Second day
9	Functions and Recursion <sup>1,2</sup>	Function definition, structure	-	-
10	Numpy <sup>2</sup>	Arrays structure, vector and matrix operations	-	-
11, 12	Pandas <sup>2</sup>	Series, Dataframe, and data manipulation	-	Box 1/Second and Third day
14	Data cleaning <sup>2</sup>	Cleaning and treatment of biological data sets	-	Box 1/Third day
15	Matplotlib <sup>2</sup>	Data visualization	-	Box 1/Third day
26	Biopython <sup>1</sup>	Manipulating sequence data	-	Box 2/Fourth day

The metric of interaction (number of views) on Instagram was obtained as a measure of effectiveness and scale of social impact of the workshop. The main goal is to increase our social media. Code club meetings occur biweekly over the year and started on the third week of January, 2022, and is planned to end in September (close to the workshop date). Numbers in column "Title" indicate whether topics were covered in code clubs before submission of ISMB 2022 poster (1), or are planned to be covered before (2) and after the V edition of code clubs (Zuvanov et al. 2021).

### References

Zuvanov, Lúcia, et al. "The experience of teaching introductory programming skills to biologists in Brazil." PLoS computational biology 17.11 (2021): e1009534.



# *Press Release e Apoio em Visibilidade*



RedeComCiencia · 23 de mai. · 5 min para ler

## Workshop gratuito ensina programação para biocientistas

Por Larissa Graciano Braga e Renato Augusto Corrêa dos Santos

*Iniciativa do Workshop de Python para Dados Biológicos partiu do biólogo Renato A. Corrêa dos Santos, membro da RedeComCiência*

O aumento exponencial na quantidade de dados biológicos, associado à diversidade e complexidade característica desse tipo de dado requer cada vez mais o uso de computação para organização e análise. O trabalho de exploração e de extração de informações é essencial e pode ser otimizado com uso de uma linguagem de programação. Atualmente, a linguagem Python é uma das mais utilizadas no mundo para essa finalidade.

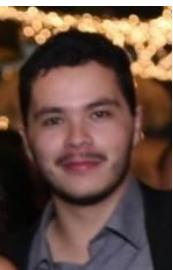


- **RedeComCiência**
- **Fundação de Estudos Agrários Luiz de Queiroz**

# Comissão Organizadora em 2022



# Docentes e pós-doutorandos



# Conselheiro





# Objetivos do Workshop

Contribuir para desenvolvimento de atividades básicas envolvendo aquisição, limpeza e manipulação de dados ensinando o básico de programação em Python



V Workshop de Python para Dados Biológicos || 26 a 30 de setembro de 2022

# Como?

Mostrando como aplicar habilidades básicas de programação em tarefas que biocientistas lidam diariamente usando Python.



# O que esperar do workshop esta semana?

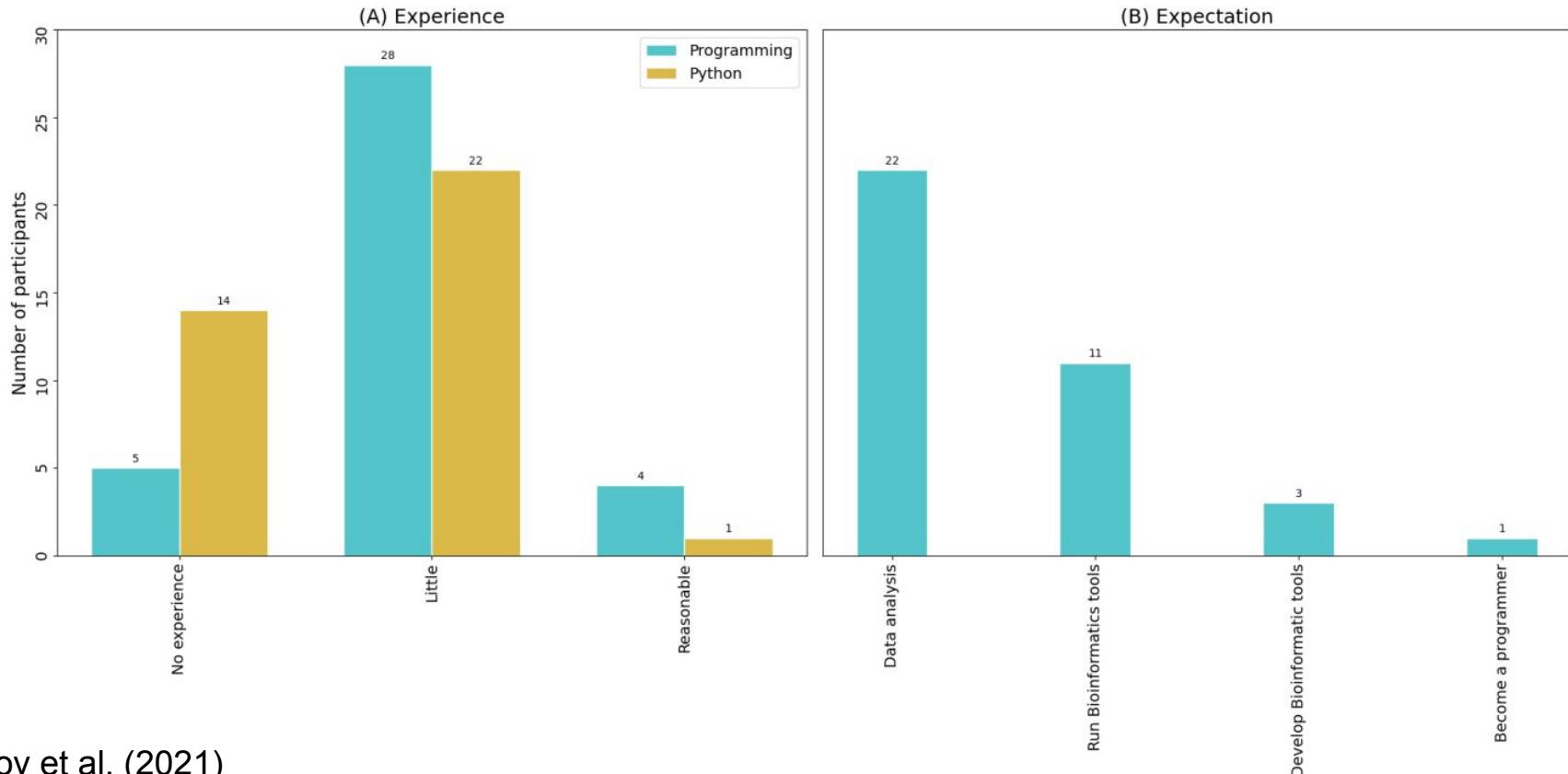
Introdução a algumas **habilidades de programação com aplicações em dados biológicos usando a linguagem Python**

**Público esperado:** *biocientistas com pouco ou nenhum conhecimento de ciência da computação, com pouca ou nenhuma experiência em Python.* Serão introduzidos *aspectos básicos da programação (sintaxe), mostrando aplicações em análise de dados biológicos.*

Consideramos o material como **introdutório**.



# Experiência prévia / Expectativas em 2020





# O que NÃO esperar do workshop esta semana?

- Conhecimento aprofundado em ciência da computação ou em biologia
- Aprendizado de novos conceitos em estatística ou matemática



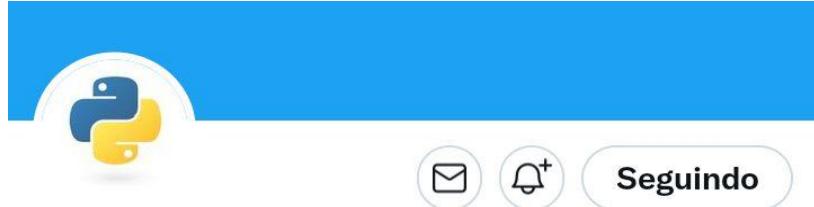
# Estratégias aplicadas neste workshop

- Quais problemas nós temos no laboratório que poderiam ser resolvidos com programação?
- União de pessoas (comissão organizadora) com diferentes “backgrounds”
- Introdução do problema biológico
- Apresentação da parte computacional de forma prática e aplicada
- De preferência totalmente prático, com apresentações de conceitos
- Com palestrantes para dar um gostinho de “como ir além” com programação



## Pesquisa associada

- Comitê de ética (todos devem ter preenchido formulário)
  - Análises e gráficos feitos pela comissão organizadora usando Python :-)
- Rede sociais (literatura e interação com comunidade brasileira e internacional)
  - Área de constante mudança e avanços
  - Desenvolvimento de ferramentas com atualizações quase que diariamente
- Code clubs
  - Encontros quinzenais
  - Discussão de temas biológicos
  - Exercícios aplicados
- Novo manuscrito em desenvolvimento liderado pelos alunos



## Brazilian Python Workshop for Biological Data

@BrazilPythonWS Segue você

Brazilian Python Workshop for Biological Data

Entrou em julho de 2020

212 Seguindo 318 Seguidores

brazilpythonws

Enviar mensagem Seguir ...

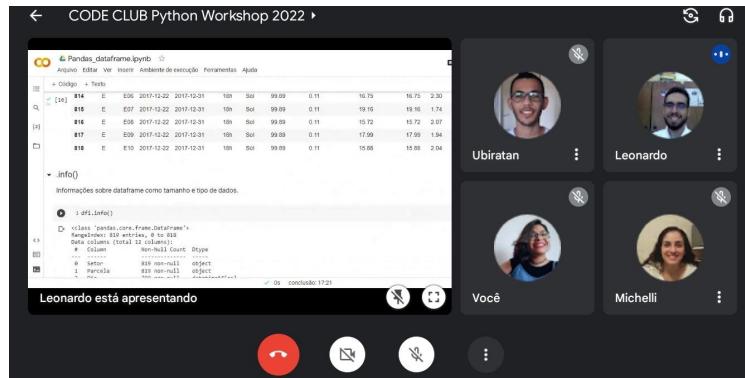
77 publicações 630 seguidores 234 seguindo

Wshop Python Dados Biológicos  
APLICAÇÃO DE PYTHON PARA A ANÁLISE DE DADOS BIOLÓGICOS

WORKSHOP ANUAL  
Link da edição deste ano (2022) ↗  
wpdb2022.netlify.app

ORGANIZA... IV EDIÇÃO III EDIÇÃO OBJETIVO

PUBICAÇÕES REELS MARCADOS





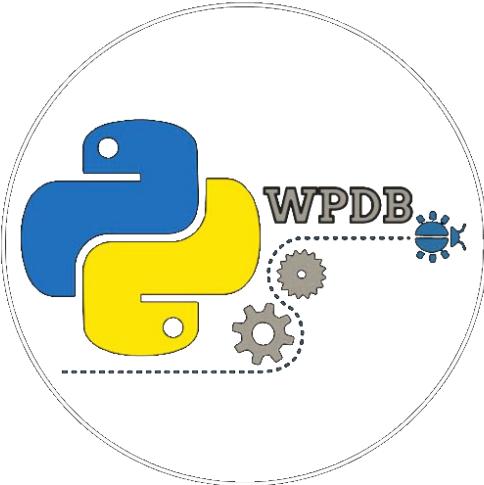
# Flash talks

- Critérios para produção das apresentações de *Flash Talks*
  - Instigar ideias para resolução de problemas usando Python
  - 10 minutos: breve apresentação do projeto (2 min), foco nos dados e perguntas aos dados!
- Comissão científica
  - Seleção de projetos
  - Avaliação e formulação de perguntas (caso os outros participantes não tenham)

# VI Workshop de Python para Dados Biológicos (2023)



- Também será *online*, talvez semipresencial.
- Também permitirá participação de pessoas de todo o país!
- Diferente do evento deste ano, PRETENDEMOS abrir para participantes externos (ouvintes)



@brazilpythonws



<https://github.com/brazilpythonws>

## Realização



ESALQ



## Apoio



AlfaHelix  
ESCOLA DE COMPUTAÇÃO  
[WWW.ALFAHELIX.COM.BR](http://WWW.ALFAHELIX.COM.BR)