

Experiences of the Brazilian Python Workshop for Biological Data in 2022

Gustavo Schiavone **Crestana**¹, Michelli Inácio Gonçalves **Funnicelli**², Ubiratan da Silva **Batista**³, Raissa Melo de **Sousa**⁴, Pedro Henrique Narciso **Ferreira**⁵, Matheus **Scaketti**⁶, Leonardo André Mumbach da **Silva**⁷, Vinícius **Franceschini-Santos**⁸, Claudia Barros **Monteiro-Vitorello**¹, Larissa Graciano **Braga**², Diego Mauricio **Riaño-Pachón**⁹, Renato Augusto Corrêa dos **Santos**⁹

¹ESALQ, University of São Paulo (USP), ²School of Agricultural and Veterinarian Sciences, São Paulo State University (UNESP), ³Federal University of Ouro Preto (UFOP), ⁴Federal University of Pará (UFPA), ⁵Federal University of São Carlos (UFSCar), ⁶University of Campinas (UNICAMP), ⁷University of Vale do Rio dos Sinos (UNISINOS), ⁸Faculty of Philosophy, Sciences and Letters of Ribeirao Preto, University of São Paulo (USP), ⁹CENA, University of São Paulo (USP)

Correspondence: renatoacsantos@gmail.com (R.A.C.S), gustavo.crestana@usp.br (G.S.C), michelli.funnicelli@unesp.br (M.I.G.F), ubira1pdi@gmail.com (U.S.B)

Using social media to increase our social impact

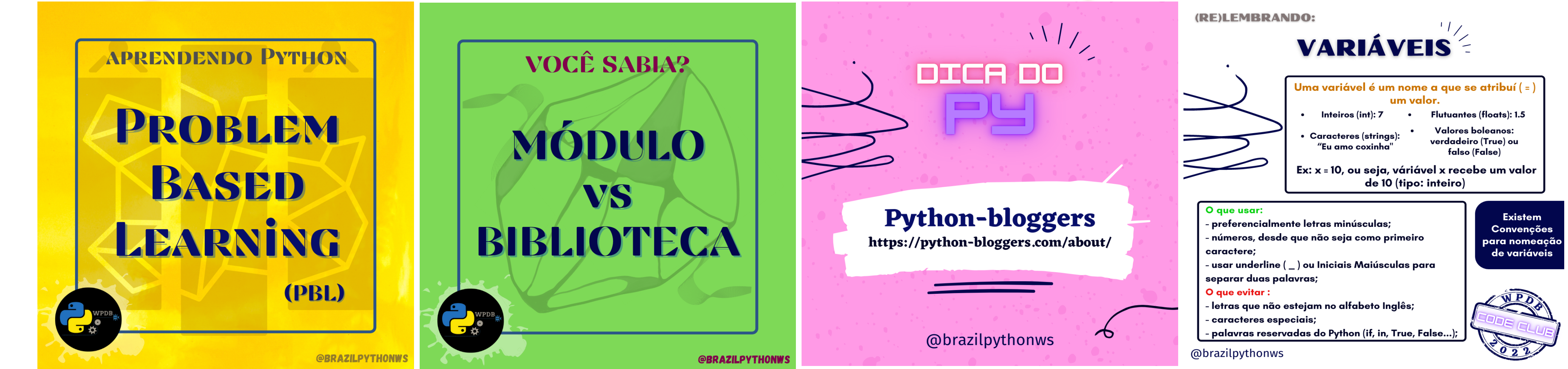
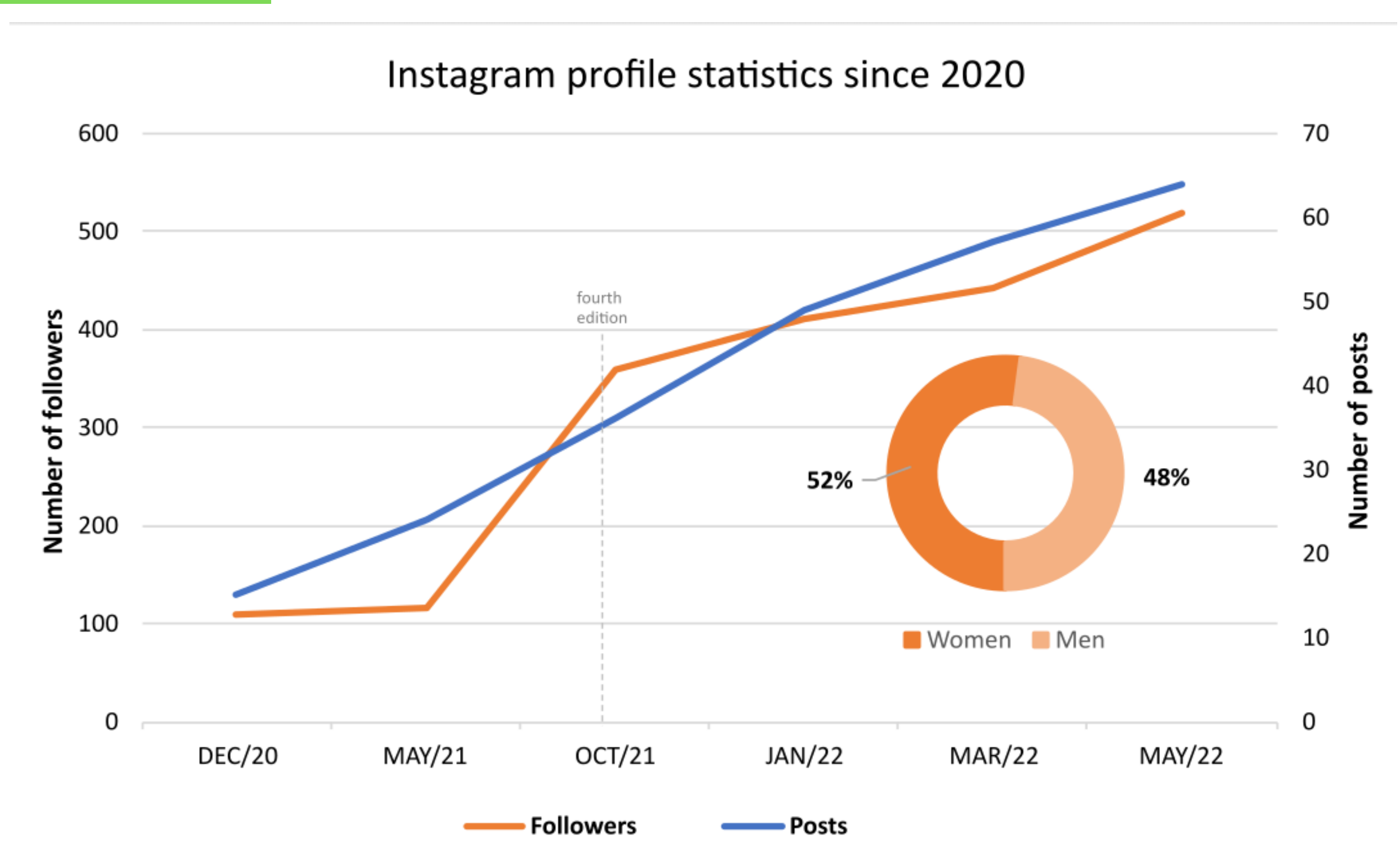


Figure 1 (above). Types of posts in the event's profile on Instagram (@brazilpythonws): research and innovation in teaching programming skills to bioscientists and related topics (yellow), Python libraries and methods (green), tips about supplementary materials for teaching and learning programming skills (pink), and topics discussed on code clubs (white).

Figure 2 (right). Temporal analysis of the @brazilpythonws profile statistics, since the 2020 edition. The weekly publication of content related to the event is essential for the visibility of the group's work, for example, for the edition of the event in 2021 (dashed line in gray).



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 discursive questions at the beginning and end of the event. In order to characterize the target audience and general aspects of the course, 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 to participants of the 2021 workshop edition.

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

Organizing code clubs to engage and empower the team

Given our objective of promoting basic programming skills for bioscientists, the code club meetings emerged in 2022 as a means to ameliorate these skills among 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 equalize.

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

Fortn.	Title	Description	Social Media	*Previous workshop content
1 and 2	Variable and operators ^{1,2}	Arithmetic, assignment comparison and logical operators	147/138	Box 1/First day
3	Strings ^{1,2}	Indexing, methods and functions applied to strings	257	Box 1/First day
4	Sequences ^{1,2}	Tuples, lists, ranges	186	Box 1/First day
5	Control Flow, Statements ^{1,2}	if, elif, and else	144	Box 1/First day
6	Sets ^{1,2}	Structure, methods and functions applied to Sets	169	Box 1/First day
7	Control Structures Repetition ²	for and while	-	Box 1/Second day
8	Dictionaries ²	Structure, methods and functions applied to Dictionaries	-	Box 1/Second day
9	Functions and Recursion ²	Function definition, structure	-	-
10	Numpy ²	Arrays structure, vector and matrix operations	-	-
11, 12 and 13	Pandas ²	Series, Dataframe, and data manipulation	-	Box 1/ Second and Third day
14	Data cleaning ²	Cleaning and treatment of biological data sets	-	Box 1/Third day
15	Matplotlib ²	Data visualization	-	Box 1/Third day
16	Biopython ³	Manipulating sequence data	-	Box 1/Fourth day

The reach of publications (number of views) on Instagram was obtained as a measure of diffusion and increase of social impact of the workshop, which is indicated in column '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 club before submission of ISMB 2022 poster (1), or are planned to be covered before (2) and after the V edition of the workshop (2022) (3).

References

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

