Teaching portfolio[1]
Daniel H. Chitwood (he/him)

Plants&Python, a series of lessons in coding, plant biology, computation, and bioinformatics

Plants&Python, una serie de lecciones de programación, biología vegetal, cómputo, y bioinformática

I have contributed to the development of the curriculum *Plants&Python*. I have led its development and wrote lessons that teach coding learning objectives in Python. The second part of the curriculum that teaches command line was written by Dr. Alejandra Rougon Cardoso, at Universidad Nacional Autónoma de México (UNAM) La Escuela Nacional de Estudios Superiores (ENES) in León, Guanajuato. The third part of the curriculum was developed by Dr. Robert VanBuren (Dept. Horticulture, Michigan State University), teaching students bioinformatics.

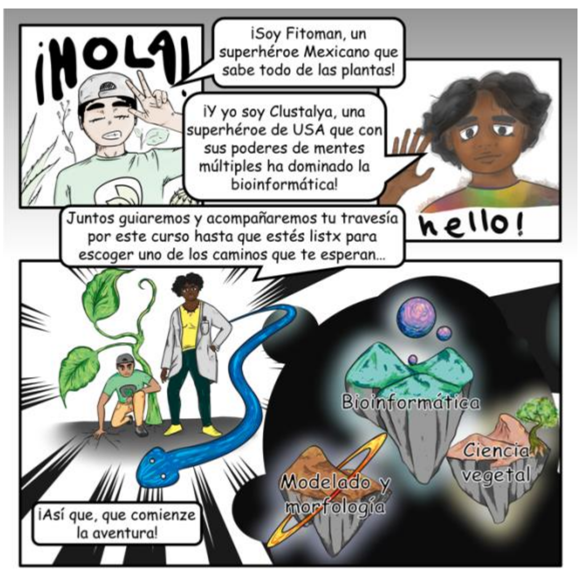


Illustration by Eddy Mendoza Galindo

Students have contributed Spanish translations of the Jupyter notebook lessons and video tutorials and provided illustrations of coding and bioinformatic "heros" that accompany students through the lessons and provide inspiration. *Plants&Python* has been published in *The Plant Cell*. Student co-authors are indicated in bold in the citation below:

VanBuren R, Rougon-Cardoso A, Amézquita EJ, **Coss-Navarette EL**, **Espinosa-Jaime A**, **Gonzalez-Iturbe A**, **Luckie-Duque AC**, **Mendoza-Galindo E**, Pardo J, **Rodríguez-Guerrero G**, **Rosiles-Loeza PY**, **Vásquez-Cruz M**, **Fernandez-Valverde SL**, Hernández-Hernández T, Palande S, Chitwood DH (2022) Plants & Python, a series of lessons in coding, plant biology, computation, and bioinformatics / una serie de lecciones de programación, biología vegetal, cómputo, y bioinformática. *Plant Cell.* 34(7):e1

All the materials for the course are publicly available in English and Spanish through a Jupyter Book:
https://plantsandpython.github.io/PlantsAndPython

[1] For transparency into the tenure process for others, a copy of my teaching portfolio can be found at the following link: https://github.com/DanChitwood/promotion_materials/blob/main/teaching_portfolio.pdf

Plants&Python structure

- Video tutorial lessons, closed-captioned and with translated subtitles on YouTube, that students watch before class. Code and transcripts are also available as Jupyter notebooks.
    - Installing Anaconda and using Jupyter notebooks: https://youtu.be/FrDYpLVuTkQ
    - Introduction to Plants and Python: https://youtu.be/wXAf_XSNeZ4
    - Video 1: Variables, Functions, Lists, & Indexing, https://youtu.be/4XIllJVnT4Y
    - Video 2: Slicing and data types, https://youtu.be/CIlwbP_QM1w
    - Video 3: Visualizing data using matplotlib, https://youtu.be/PJ1dvAgOAj0
    - Video 4: Loops and the Fibonacci sequence, https://youtu.be/Cvi3dByz9SE
    - Video 5; Pandas, dataframes, and seaborn, https://youtu.be/jEQRU55x0e4
- Practice Jupyter notebooks that students complete before class.
- Group activities in-class, again conducted using Jupyter notebooks.
- The lessons that I wrote are labelled as the Preface and Lessons 1-4 in the Jupyter Book.
- There are no traditional assessments. The course assumes that the gate-keeping and shaming tactics traditionally used by data science curriculums have negative long-term impacts on learning.
- Rather, the materials are presented as freely available for students to learn to advance *their intended impacts*. For biology students who don't know how to code, learning objectives are selected to get them coding in Python as soon as possible. For students of computational and data science, the examples involve mathematical phenomena inherent to plant biology and data that forces them to deal with "real world" science and development empirical intuition.
- The course emphasizes peer-to-peer learning both across disciplines and across cultures, bringing together plant & computational scientists in the USA and México. See pedagogical research, below.
- The second half of the HRT841/UNAM course is a class project that applies what students have learned and results in a class manuscript, with students as co-authors. The following publication and pre-print are the class projects from 2019 and 2020, respectively, with student co-authors in bold. The 2021 and 2022 class projects are currently in preparation.

    - **Bryson AE***, **Wilson Brown M***, **Mullins J***, **Dong W***, **Bahmani K**, **Bornowski N**, **Chiu C**, **Engelgau P**, **Gettings B**, **Gomezcano F**, **Gregory LM**, **Haber AC**, **Hoh D**, **Jennings EE**, **Ji Z**, **Kaur P**, Kenchanmane Raju SK, **Long Y**, **Lotreck SG**, **Mathieu DT**, **Ranaweera T**, **Ritter EJ**, **Sadohara R**, **Shrote RZ**, **Smith KE**, **Teresi SJ**, **Venegas J**, **Wang H**, **Wilson ML**, **Tarrant AR**, Frank MH, Migicovsky Z, Kumar J, VanBuren R, Londo JP, Chitwood DH (2020) Composite modeling of leaf shape across shoots discriminates *Vitis* species better than individual leaves. *Applications in Plant Sciences.* 8(12): 11404

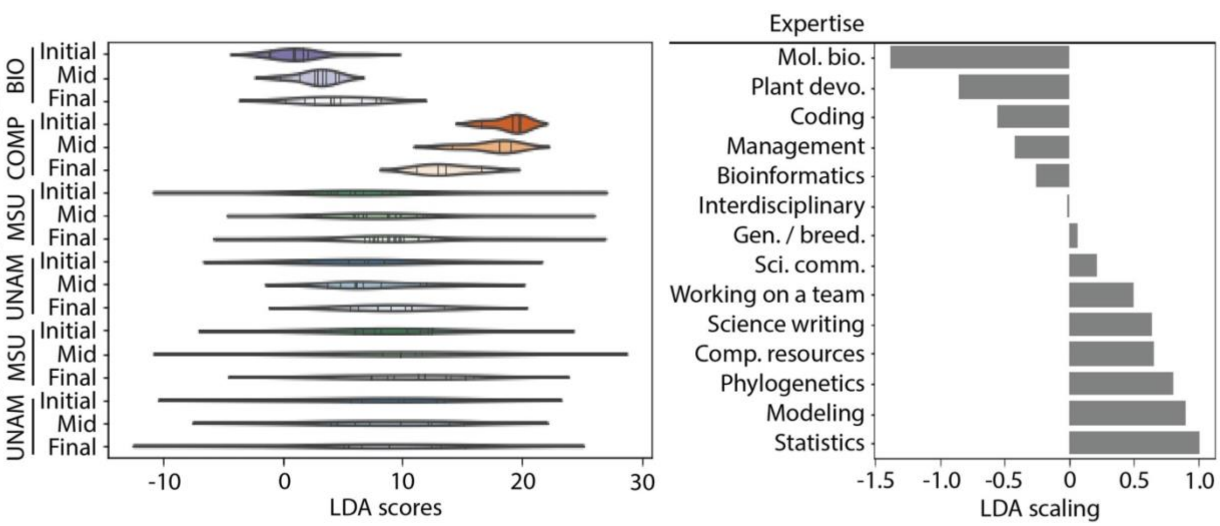
    - Palande S*, **Kaste JAM***, **Roberts MDS***, **Segura Aba K***, **Claucherty C**, **Dacon J**, **Doko R**, **Jayakody TB**, **Jeffrey HR**, **Kelly N**, **Manousidaki A**, **Parks HM**, **Roggenkamp EM**, **Schumacher AM**, **Yang J**, Percival S, **Pardo J**, Husbands AY, Krishnan A, Montgomery BL, Munch E, Thompson AM, Rougon-Cardoso A, Chitwood DH, VanBuren R (2022) The topological shape of gene expression across the evolution of flowering plants. *bioRxiv*.

    - I have recently extended the idea of including students as co-authors to CMSE201, where I offered students in 2021 the chance to participate in research for their class project, which resulted in the following publication:

Migicovsky Z, Quigley MY, Mullins J, **Ali T**, Swift JF, **Agasaveeran AR**, **Dougherty JD**, **Grant BM**, **Korkmaz I**, **Malpeddi MR**, **McNichol EL**, **Sharp AW**, Harris JL, Hopkins DR, Jordan LM, Kwasniewski MT, Striegler RK, Dowtin AL, Stotts S, Cousins P, Chitwood DH (2022) X-ray imaging of 30 year old wine grape wood reveals cumulative impacts of rootstocks on scion secondary growth and Ravaz index. *Horticulture Research.* uhac226

Interdisciplinary/intercultural pedagogical research

Plants&Python was originally conceived as an interdisciplinary class, as part of the NSF Research Traineeship (NRT) program Integrated training Model in Plant And Compu-Tational Sciences (IMAPCTS), to bring biologists who have never coded and data scientists who have no prior experience in plant biology together. In 2020, all students were from MSU, members of plant biology or computational departments. At the beginning, middle, and end of the semester, I gave students a survey to self-report their expertise in a variety of topics related to the learning objectives of the course. Building a Linear Discriminant Analysis (LDA) model, to find the self-reported expertise values that most separate biology vs. computational students in 2020, yielded predictable results (below). In 2020, biology students (purple) were strongly separated by the model from computational students (orange), reporting predictable expertise (see LDA scaling values, right). As the semester progressed, biology students became more "computational" and computational students more "biological", reflecting the interdisciplinary goals of the class.

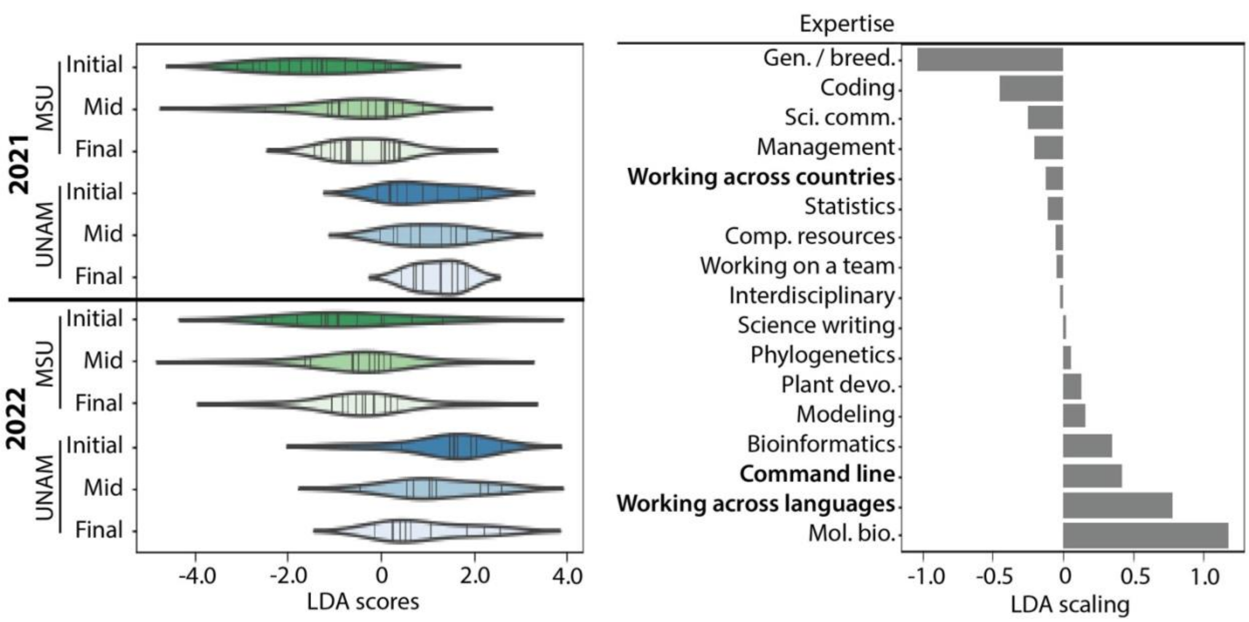


In 2021 and 2022, UNAM students joined the class with MSU students. The UNAM system is vast, and the majors of the students could not neatly be divided into "biology" and "computational" groups. Still, biological and computational expertise was more or less equally divided between MSU and UNAM students, as seen when the LDA predicting MSU biology and computational students in 2020 is applied to MSU and UNAM students in 2021 and 2022 (above).

Although we could not group UNAM students into "biology" or "computation", the presence of MSU and UNAM students together in the class—US and México, predominately English speakers and bilingual English and Spanish speakers—embodied the original interdisciplinary premise of the curriculum, but now across nationalities, languages, and cultures. We wondered how these demographic differences would translate into scientific expertise, and if students in each group would influence each other scientifically. We gave students the same survey at the beginning, middle, and end of the semester, but added three additional questions: expertise in “Working across countries”, “Working across languages”, and “Command line”. The first two questions were added because of the new international context of the class, and the last because of the lessons on command line added by Dr. Alejandra Rougon Cardoso.

An LDA model was built from 2021 and 2022 data to find the self-reported expertise that best separates MSU and UNAM student responses (see below). MSU and UNAM students reproducibly separated in 2021 and 2022. Reflecting the similar distribution of biological and computational expertise between the groups (see above figure), MSU students reported more expertise in genetics and breeding while UNAM students more expertise in molecular biology; MSU students reported more expertise in coding while UNAM students reported more expertise in command line (see LDA scaling values, below). These similar—but differing technical skills—within biology and computation domain areas separating the student groups show that each group has equivalent expertise *between* domains but differs in the exact skills *within* each domain. One of the biggest differences between the groups was that UNAM students are strongly differentiated by their ability to “[work] across languages”, whereas “working across countries” only slightly distinguishes MSU students.

Do students from different countries in the same classroom influence each other, similar to how students across disciplines did at MSU in 2020? In 2021, MSU students became more “UNAM-like” (and UNAM students became even more UNAM-like). In 2022, the influence was more mutual, with each group of students coming to report similar expertise more closely by the end of the semester (see below). We will continue monitoring these results in the future, but they suggest that culture is at least as important (if not more) than discipline in defining the self-reported expertise of students, and that students can influence each other across both discipline and culture.



Classes that I teach

HRT841: *Foundation in Plant and Computational Science* (3 credits), part of the NSF Research Traineeship (NRT) program Integrated training Model in Plant And Compu-Tational Sciences (IMAPCTS) and **UNAM Temas selectos: *bioinformática y minería de datos con python***, as part of the CANR Global Scholars program that helped support the writing of the *Plants&Python* curriculum.

- Fall 2019, 26 MSU students
- Fall 2020, 16 MSU students
- Fall 2021, 26 MSU students, 33 UNAM students
- Fall 2022, 18 MSU students, 15 UNAM students

Plant Genomics @ MSU Research Experiences for Undergraduates (REU), where *Plants&Python* is taught as a primer for coding to students participating in genomics and bioinformatics projects during the summer.

- Summer 2021, 20 US students, 9 Mexican students
- Summer 2022, 18 US students

CMSE 201: Introduction to Computational Modeling, Michigan State University, 4 credits, undergraduate. Computational modeling using a wide variety of application examples. Algorithmic thinking, dataset manipulation, model building, data visualization, and numerical methods all implemented as programs in Python. No prior coding experience required.

- Spring 2019, 30 students
- Fall 2021, 34 students

HRT 892: Seminar in plant phenomics, Michigan State University, 1 credit, graduate. Students choose and present papers in plant phenomics research and contribute to planning the Corteva-PBGB-IRT IMPACTS Plant Phenomics Symposium.

- Fall 2019, 18 students

PLB 2410: Introductory Plant Diversity and Evolution, Cornell University, Guest Lectures: a tutorial to ImageJ/FIJI is provided to students (https://docs.google.com/document/d/13xRE1uzQWn-bb44_tcofwtEs_kVNKelOKJJWpWaA85I/edit?usp=sharing) so that they can explore plant anatomy and morphology using a variety of X-ray Computed Tomography (CT) scans (https://drive.google.com/drive/folders/1-_tX1vQTIBPZ5NtbgT3n07XI4w-RAQNJ?usp=share_link). A second lecture introduces students to Topological Data Analysis (TDA) methods to analyze X-ray CT data. ***Fall 2020***, ***Fall 2021***, ***Fall 2022***.

HRT 812: Laboratory Research Techniques, Michigan State University. Guest lectures: Students produced their own X-ray CT scan reconstructions of plants, explored and visualized data, and made a presentation of their results to the class. 2 credits. ***Fall 2018***, ***Fall 2021***

Student evaluation

The inspiration for *Plants&Python* comes from CMSE201: *Introduction to Computational Modeling*. This innovative undergraduate class uses a flipped classroom approach, Jupyter notebooks, and a student project to effectively teach Python-based coding and modeling to undergraduates at MSU with no assumed prior experience. Another positive aspect of the class is that it is taught in sections, and the instructors, which include teaching professionals, meet weekly to discuss student feedback and to update the curriculum for next year if needed. However, I do not have control over the content or most of the delivery of the course. The only way I could implement changes in response to the student evaluations below was to develop a new course, which resulted in *Plants&Python*. In 2021 my Student

Instructional Rating Survey (SIRS) scores for CMSE201 were mostly positive (1.65 to 2.2 for the seven composite factors, with 1 being “good” and 5 being “bad”). Although the class and the Learning Assistant and I received positive written comments overall, the students offered constructive feedback about how to improve the CMSE201 experience:

- For someone who has never coded before, I thought the pace of the content moved a bit fast and the topics were very challenging at times. I wish rather than pre-class assignments, we had video lectures with PowerPoint notes about how to structure the code so we could look back at it. Overall, Dr. Chitwood was a nice guy and a fair grader and Meghan was a very good LA.
- I have never programmed and this course is advertised to people like me that it does not require any past knowledge. But since the start, I felt I was behind. The necessity to google things up is a HUGE part of this class, or at least for me it was, and I did not like this. I would recommend that more examples and applications for the code are represented in the class in order to gain prevent a part of the google-ing for students like me. Another problem I spotted throughout the class was that things sometimes were very badly phrased and that made the course frustrating at these times. I would suggest a bit more simpler and logical approach to some explanations. Other than that I enjoyed the course very much!

I took these comments to heart when designing *Plants&Python*, which is truly meant for biologists who have never coded. Rather than include too many learning objectives, the class goes at a slower pace than CMSE201, focuses more on active learning, and intends to expose students to the possibilities that coding offers in their scientific and professional lives that they can reference back to in course materials when the need arises. The SIRS scores for *Plants&Python* in 2022 were improved compared to my CMSE201 scores in 2021 (1.00 to 1.85). Some of the written comments from MSU students in the class include:

- Great class, nice instructors, active classmates. I enjoyed and have started working on learning more about coding in the field that I am interested, inspired by this class.
- This was an excellent course, the interdisciplinary approach was highly successful in part due to the enthusiasm and consistency of the instructors. A good model for teaching and practicing critical research methods.
- All instructors were very enthusiastic about teaching the material. The course provided the opportunity to work with group members from diverse scientific backgrounds. I would recommend the course to students interested in taking an introductory course in Python programming through plant-related topics.

As part of a write-up for the publication of *Plants&Python* in *The Plant Cell*[2], the student translators gave quotes about their experiences, which I am copying below:

- **Eddy Mendoza-Galindo** (Laboratory of Agrigenomic Sciences, Universidad Nacional Autónoma de México, ENES-León, León, Guanajuato, México): “Plants are amazing, and so is their science. As biologists, big data and programming are sometimes scary. Plants and Python is a unique way to learn computational plant science. Learning was never that easy. Going through the lessons is an exciting journey. It was made for everyone: it doesn’t matter which academic background you come from, which part of the world, or what language you speak. We all are invited to admire plant beauty and magnificence from our computers!”

[2] https://plantae.org/new-teaching-tool-plants-and-python/

- **Marilyn Vásquez-Cruz** (Laboratorio Nacional de Genómica para la Biodiversidad (Langebio)-Centro de Investigación y de Estudios Avanzados (Cinvestav), Irapuato, Guanajuato, México: “Plants & Python is great material for those who want to explore this programming language. It does not matter if you are a beginner or have advanced knowledge on this topic. The exercises are designed using plant data as examples, which makes the lessons more and more interesting as you go along. You will be able to solve mathematical problems until you draw some beautiful graphs”

The Plants & Python authors. From left to right, top to bottom: Alejandra Rougon-Cardoso, Robert VanBuren, Erik J. Amézquita, Evelia L. Coss-Navarrete, Eddy Mendoza-Galindo, Aarón Espinosa-Jaime, Jeremy Pardo, Guillermo Rodríguez-Guerrero, Pablo Y. Rosiles-Loeza, Marilyn Vásquez-Cruz, Selene L. Fernandez-Valverde, Sourabh Palande, and Daniel H. Chitwood

- **Erik J. Amézquita** (Michigan State University, East Lansing, Michigan, USA): “This material can be very insightful for the non-biology crowd, even for those who already feel comfortable with coding. As a mathematician by training, I was very surprised about the computational richness within plants. It may sound naive, but I thought plants were simply growing there. Turns out that, dare I say it, plant biology is way more difficult than math and computer science. Evolution, development, multiple omics, oh my!”
- **Aarón Espinosa-Jaime** (Laboratory of Agrigenomic Sciences, Universidad Nacional Autónoma de México, ENES-León, León, Guanajuato, México): “This is just the beginning. You just need a sparkle of imagination and inspiration to start your own plant omic project.”
- **Pablo Y. Rosiles-Loeza** (Laboratorio Nacional de Genómica para la Biodiversidad (Langebio)-Centro de Investigación y de Estudios Avanzados (Cinvestav), Irapuato, Guanajuato, México): “An incredible effort to bring the foundations of Python data analysis to a wider audience.”

Syllabus

A copy of the syllabus from the fall 2022 MSU HRT841: *Foundation in Plants and Computational Science* and UNAM *Temas selectos: bioinformática y minería de datos con python* class is attached after this document.

Syllabus

Michigan State University HRT 841: Foundations in Computational Plant Sciences

UNAM Temas selectos: Bioinformática y minería de datos con python

Fall 2022, Tuesdays & Thursdays 10:20-11:40 am,

In-person PLB151 Michigan State University,

also offered virtually to those at UNAM or any student who prefers virtual

3 credit hours

Instructors: Dr. Dan Chitwood (MSU), Dr. Alejandra Rougon (UNAM León), Dr. Robert VanBuren (MSU)

Emails: bobvanburen@gmail.com; dhchitwood@gmail.com; alerougon@gmail.com

Office hours: Students are highly encouraged to schedule individual one-on-one and small group meetings with the instructors. Please email instructors to set-up an appointment. Additionally, the instructors will try to be available after the normally scheduled class period if needed.

Class materials: Available as a Jupyter Book at https://plantsandpython.github.io/PlantsAndPython

IMPORTANT NOTE: We are still in a pandemic(s). Your health and well-being always take precedence. Although Michigan State University (MSU) classes in person, this class is already being offered virtually, as students in both the USA and México are enrolled. If you are at MSU, you can choose to take the class in person, virtually, or both. We ask though that you notify the instructors of your plans using the poll that will be sent so that we can plan appropriately (for example, we need to know how many students are participating virtually for limits that we need to adjust for on zoom and Discord, the online platforms we will be using). Similarly, if you need to skip a class or need deadlines extended because of the exceptional times we are living in, we understand. We ask that you simply communicate what you need to the instructors (you don't need to provide a reason). This course is meant to be a learning experience for you and you know best what you need to further your intended goals and career. Please be in communication as needed with the instructors.

Course description:

This course will bring together plant biologists and computational/data scientists to address grand challenges in plant biology. It also brings together students from both the USA and México, at Michigan State University (MSU) and Universidad Nacional Autónoma de México (UNAM). This project-based course will teach fundamental concepts in plant science through a computational lens. The first half of the course will consist of a mix of lectures and thought experiments and hands-on activities in Python implemented in Jupyter notebooks. No prior coding experience is required. Students will learn basic programming concepts implemented using cutting edge datasets and computational approaches. In the second half of the course, students will download previously published data and conduct a bioinformatics class project implementing the computational and data science skills they have learned. Results from the project will be written up as a manuscript, the goal of which is to submit to a research journal with students as authors. Foundation in Computational Plant Sciences is open to graduate students in any program and no prerequisite experience in coding or computational biology is required. Foundation in Computational Plant Sciences is the first course in the NSF-IMPACTS series training program at MSU and it is offered virtually through UNAM in México as *Temas selectos: Bioinformática y minería de datos con python*. This course will be taught using a flipped classroom approach, meaning that lectures will be available online through YouTube video tutorials and Jupyter notebooks and there will be group activities to explore the learning concepts in class the following day.

Course goals:

The course will synthesize concepts from computational and data science with plant biology principles. The goals for this course are for students to:

(1) acquire foundational knowledge and skill sets in plant biology and data science

(2) formulate hypothesis-driven questions stemming from grand challenges at the interface of plant biology and data sciences

(3) synthesize the acquired skills to solve questions in collaborative teams

(4) communicate persuasively across disciplines and cultures with peers and the public

By the end of the course, students will be able to identify and create strategies to address current problems in plant biology using cutting edge computational tools and approaches. Through interactive lectures and 'thought experiments' students will learn to communicate across disciplines and cultures and to explore emergent biological processes.

Course Format:

This course is hybrid, offered both in person at Michigan State University in the USA, virtually through UNAM in México, and also available virtually to any MSU student that wishes to participate as such. We will be using zoom and Discord for lectures and small group meetings. Discord allows a record of chats and links to resources for each lecture to be recorded and accessible at later dates. Breakout rooms can be created, and students and instructors can freely move between them and small groups can spontaneously video chat if they want to. An invitation link will be sent to you by the instructors to join the Discord server for the class. We will use zoom at the beginning of class and then break out into groups on Discord. Using zoom and Discord does not preclude the use of other messaging and video chat platforms, especially if small groups or individuals prefer to meet using other platforms outside of class. The platform used is flexible, and if the class prefers a different one, it can be used as well.

This course is uniquely international. It is being offered to both students in the USA and México through MSU and UNAM. The course is in English. However, many of the course materials have been translated into Spanish (see https://plantsandpython.github.io/PlantsAndPython). If it helps, please use the Spanish lessons, either to help you learn the material and in learning English related to coding and plant science; or, for English speakers to learn about coding and plant science in a different language or to help their Spanish speaking skills. Agriculture and plant science in North America involves multiple countries and both English and Spanish speakers. This course is meant to be **inter-disciplinary**, between the computational and plant sciences; the course is also meant to be **inter-cultural**, synthesizing and bringing together the cultures and languages that we use in agriculture and the plant sciences in North America. **Be respectful of the cultures and languages in this course and learn from each other, both across scientific disciplines and unique cultural perspectives.**

The first half of the semester will be lectures. The first set of lessons by Dr. Dan Chitwood will introduce foundational coding concepts in Python. The second set of lectures by Dr. Alejandra Rougon will introduce students to the command line. The third set of lectures by Dr. Bob VanBuren will introduce students to bioinformatics concepts. Generally, there will be ***pre-class lessons*** and video tutorials for you to go over at your own pace at home the night before. There will also be a pre-class ***practice assignment*** for you to complete. These assignments are meant to check your understanding of the material. They are only graded for being completed and turned-in, not for being "correct". Please use the ***practice assignments*** to check your mastery of the material and ask the instructors and peers in class questions of any problems you ran into the next day. Finally, there will be ***in-class activities****.* Groups will always consist of both USA and México students and be changed periodically: so get to know your classmates during group work! **Because this is a hybrid class, it is important that if you are in class that you are using your Discord audio/video so that all students can fully participate.** The in-class activities are meant to fully explore the learning objectives that you learn the night before. This is group work, and we stress peer-to-peer learning. However, we ask that you complete and turn in your own in-class activity. We will go over assignments and lessons in class and on the Discord channel.

The second part of the course will be a class project. It will likely involve measuring and comparing the shapes of a large number of empirically determined and predicted protein structures. We hope that the class project leads to a class manuscript with students as co-authors. More about this later!

Please give the instructors feedback on your preferences about online platforms and class structure: we are here for you!

Assessment and grading:

Graded assignments include a midterm and end of the semester project as well as weekly attendance, participation, and completion of the Jupyter notebooks. The breakdown is as follows:

50% Attendance, participation, and completion of the pre-class lesson and in-class activity Jupyter notebooks and weekly summaries

12.5% Attendance

12.5% Participation

25% Assignments: pre-class lesson, in-class activity Jupyter notebooks and weekly summaries

25% Midterm project

25% Final project

Attendance/Participation: Because this course is hybrid with students participating both in-person and virtually, it is vital that you attend classes and arrive on time. Attendance will be taken promptly when the class starts at 10:20am ET on Tuesday and Thursday and it is part of your grade (because of the large number of students, we will use an automatic attendance program on Discord for virtual students). The class format is flipped, meaning that through Jupyter notebooks and online videos, lectures are viewed at home. In class, the instructors will begin with a small lecture and review, but most of the class will be in small groups performing the in-class activity. Groups will report back at the end of class and present results. Participation is graded by contributing to group activities and group reports at the end of class.

Assignments: For the first half of the class, Python coding and bioinformatics will be taught using Jupyter notebooks. There is a pre-class lesson and in-class activity Jupyter notebook for each lecture. You are expected to read through and complete each pre-class lesson notebook before the class. If there are video tutorials in the notebook, you are expected to watch them before coming to class (approximately 30-40 minutes total). There are in-class activities for each class as well. At the end of class, after reporting group results, each individual is expected to turn in a notebook. You will be graded by turning in the notebook and completing it, yourself. Grading is solely based on an honest attempt by a student to complete the notebook with their group. Students are expected to heavily comment on their code and the instructors will go over each notebook and provide feedback. Although working in a group, you should be completing code as you understand it from the group discussion. In the second half of the course, students will be working in groups on the class project. Students are expected to 1) give weekly one slide summaries of progress to the class for feedback and discussion and 2) turn in weekly Jupyter notebooks with their coding progress. Grading for assignments in the second half of the course is based on participating in the above activities.

Midterm/Final project: You will be working on a class project this semester where you put to use the coding and computational skills you learn during the first part of the course. Hopefully, the project will result in a research manuscript authored by students (see the manuscript from the 2019 class: **https://doi.org/10.1002/aps3.11404**). The project will have two parts: 1) a plan, composed by students with specific aims, objectives, and deadlines for completing analyses and work and 2) the final product, a draft manuscript with results written up and the conclusions presented, as well as documented code and data for others to reproduce the work. The manuscript will be submitted to a research journal the semester after the course. The midterm will be a presentation and submission of the project plan for the rest of the semester and the final the presentation and submission of the project results.

This year, the class project will analyze and compare the similarity of thousands (millions?) of protein structures. These structures are either empirically determined or predicted, through new algorithms such as AlphaFold. In using a field of mathematics, Topological Data Analysis, that measures the structure of data, we will compare the overall similarity of each structure to every other, allowing questions about protein structure prediction and evolutionary changes in protein structure to be asked. Through discussion, students this year will have a role in framing questions, hypotheses, and new directions they wish to take the analysis. We will discuss more about the project in class!

Communication: This class is taking place under unprecedented circumstances of a continuing pandemic. The above grading policies on attendance, participation, and assignments will be strictly enforced. However, because of the exceptional circumstances, without question students only need to communicate to the instructors that they will miss an online class period or need an extension on turning in an assignment and the request will be granted. The class is flexible and accommodating to students, but it ***requires communication with the instructors BEFORE the student will miss a class or turn in an assignment late*** (of course, for truly exceptional circumstances, you can communicate your absence after the fact, but we ask for you to tell us before class if possible)***.***

Grade Point Assignment (Grading Scale)

The table below describes the relationships between letter grades, percent, and performance. The first column describes the letter grade. The second column describes the percentage associated with that letter grade. The third column describes the performance represented by that letter grade and percentage.

Grade Point **(US/México)**	***Percentage***	***Performance***
4.0 / 10	90 to 100%	Excellent Work
3.5 / 9.5	80 to 89%	Above average
3.0 / 9	70 to 79%	Good Work
2.5 / 8.5	60 to 69%	Mostly Good Work
2.0 / 8	50 to 59%	Average work
1.5 / 7	40 to 49%	Below average work
1.0 / 6	30 to 39%	Poor work
0 / 5	0 to 29%	Failing work

Course Policies

Diversity Equity and Inclusiveness

Diversity, Equity and Inclusion are important, interdependent components of everyday life in the College of Agriculture and Natural Resources (CANR) and are critical to our pursuit of academic excellence. Our aim is to foster a culture where every member of CANR feels valued, supported and inspired to achieve individual and common goals with an uncommon will. This includes providing opportunity and access for all people across differences of race, age, color, ethnicity, gender, sexual orientation, gender identity, gender expression, religion, national origin, migratory status, disability / abilities, political affiliation, veteran status and socioeconomic background. (See the full CANR statement: https://www.canr.msu.edu/news/canr-statement-on-diversity-equity-and-inclusion)

Commit to Integrity: Academic Honesty

Article 2.3.3 of the Academic Freedom Report states that "The student shares with the faculty the responsibility for maintaining the integrity of scholarship, grades, and professional standards." In addition, the (insert name of unit offering course) adheres to the policies on academic honesty as specified in General Student Regulations 1.0, Protection of Scholarship and Grades; the all-University Policy on Integrity of Scholarship and Grades; and Ordinance 17.00, Examinations. (See Spartan Life: Student Handbook and Resource Guide and/or the MSU Web site: www.msu.edu.)

Therefore, unless authorized by your instructor, you are expected to complete all course assignments, including homework, lab work, quizzes, tests and exams, without assistance from any source. You are expected to develop original work for this course; therefore, you may not submit course work you completed for another course to satisfy the requirements for this course. Also, you are not authorized to use the www.allmsu.com Web site to complete any course work in this course. Students who violate MSU academic integrity rules may receive a penalty grade, including a failing grade on the assignment or in the course. Contact your instructor if you are unsure about the appropriateness of your course work. (See also the Academic Integrity webpage.)

Inform Your Instructor of Any Accommodations Needed

From the Resource Center for Persons with Disabilities (RCPD): Michigan State University is committed to providing equal opportunity for participation in all programs, services and activities. Requests for accommodations by persons with disabilities may be made by contacting the Resource Center for Persons with Disabilities at 517-884-RCPD or on the web at rcpd.msu.edu. Once your eligibility for an accommodation has been determined, you will be issued a Verified Individual Services Accommodation ("VISA") form. Please present this form to me at the start of the term and/or two weeks prior to the accommodation date (test, project, etc.). Requests received after this date may not be honored.

Participation and Engagement

During all classes, the instructor expects students to be fully engaged and prepared to discuss reading assignments. Students are encouraged to ask questions of the instructor, guest speakers, and their peers.

Active participation includes, but is not limited to, the following behaviors:

1. Asking and answering questions of the instructors, peers, or guest speakers
2. Bringing forth new ideas, information, or perspectives to academic conversations
3. Discussing your readings and reflections with instructors and peers
4. Meeting with the instructors to discuss your interests, assignments, or project
5. Questioning information presented and discussed
6. Participating in small group discussions and activities
7. Assuming responsibility for personal behavior and learning

While working on group projects, students should be mindful of other students in their group; therefore, it is important for all participants to exercise:

- Respect for themselves, each other
- Openness and a positive attitude toward new ideas and other's ideas
- Flexibility and tolerance of ambiguity
- Good communications amongst themselves.

General college and university policies

All other general college and university policies applicable to this course are available at https://www.canr.msu.edu/academics/courses/policies . Please review these policies.
Topics covered in these general policies include:

- Students with disabilities, Resource Center for Persons with Disabilities (RCPD) and accommodations
- Student rights under the family educational rights and privacy act (FERPA)
  - Student release authorization form
- Religious holiday policies
- Grief absence policies
- Students in distress policies
- MSU student athlete policies
- Course add-drop policies
- Honors options
- Course Management system policies
- Final exam policy and attendance
- Grade dispute policies
- Academic honesty and integrity, plagiarism, and disciplinary procedures
- Disruptive behavior
- Harassment and discrimination policies
- RVSM University reporting protocols
- Limits to confidentiality
- Social media policy
- Web accessibility policies
- MSU Code of Teaching Responsibility
- SIRS
- Commercialization of lecture notes

· University Learning Goals

Detailed course schedule:

Day	Topic	Instructor
thurs, 1 sept	Introduction	
mon, 9 sept	US Labor Day	
tues, 6 sept	1. Variables, Lists, & Indexing	Dr. Daniel Chitwood
thurs, 8 sept	2. Visualizing Data with Matplotlib	Dr. Daniel Chitwood
tues, 13 sept	3. Loops & the Golden Angle	Dr. Daniel Chitwood
thurs, 15 sept	4. Data Analysis with Pandas	Dr. Daniel Chitwood
fri, 16 sept	México Día de la Independencia	
tues, 20 sept	5. UNIX Command Line	Dr. Alejandra Rougon Cardoso
thurs, 22 sept	6. UNIX Commands for Data Mining	Dr. Alejandra Rougon Cardoso
tues, 27 sept	7. Scripting with Bash	Dr. Alejandra Rougon Cardoso
thurs, 29 sept	8. Genomics and Biopython	Dr. Robert VanBuren
tues, 4 oct	9. Comparative Genomics	Dr. Robert VanBuren
thurs, 6 oct	10. RNASeq Analysis	Dr. Robert VanBuren
tues, 11 oct	11. Variant Discovery	Dr. Robert VanBuren
thurs, 13 oct	Guest seminar: AlphaFold and protein structure	Dr. Annabel Romero
tues, 18 oct	Guest seminar: Topological Data Analysis	Dr. Sarah Percival
thurs, 20 oct	Guest seminar: Deep learning methods	Sarah McGuire, PhD student
tues, 25 oct	No class: MSU break day	
thurs, 27 oct	Class project	
tues, 1 nov	Class project	
thurs, 3 nov	Class project	
tues, 8 nov	Class project	
thurs, 10 nov	Class project	
fri, 11 nov	US Veterans Day	
tues, 15 nov	Class project	
thurs, 17 nov	Class project	
mon, 21 nov	México Día de la Revolución	
tues, 22 nov	Class project	
thurs, 24 nov	No class: US Thanksgiving	

tues, 29 nov	Class project	
thurs, 1 dec	Class project	
tues, 6 dec	Class project	
thurs, 8 dec	Group presentations	