

Introduction to Coding

Starts at 9:30am CT

• • •

Instructors: Amanda Farah, Maria Hernandez, Katie Dixon

Teaching Assistants: Salman Yousaf, Shrikanth Subramanian

[Powerpoint will be added to git as a pdf after class]

Section One: Introductions and Goals

...

Maria
[she/her]

Instructor and TA intros

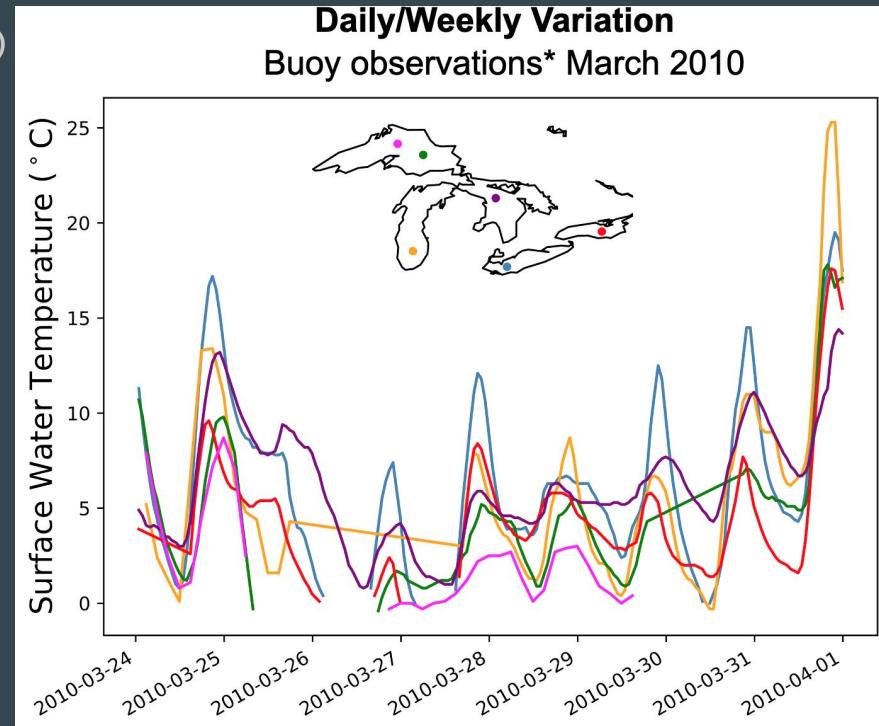
• • •

We are all students, so we are also learning!

Maria 4th year Geophysical Sciences (PSD)

Research: Biological oceanography (but in lakes)

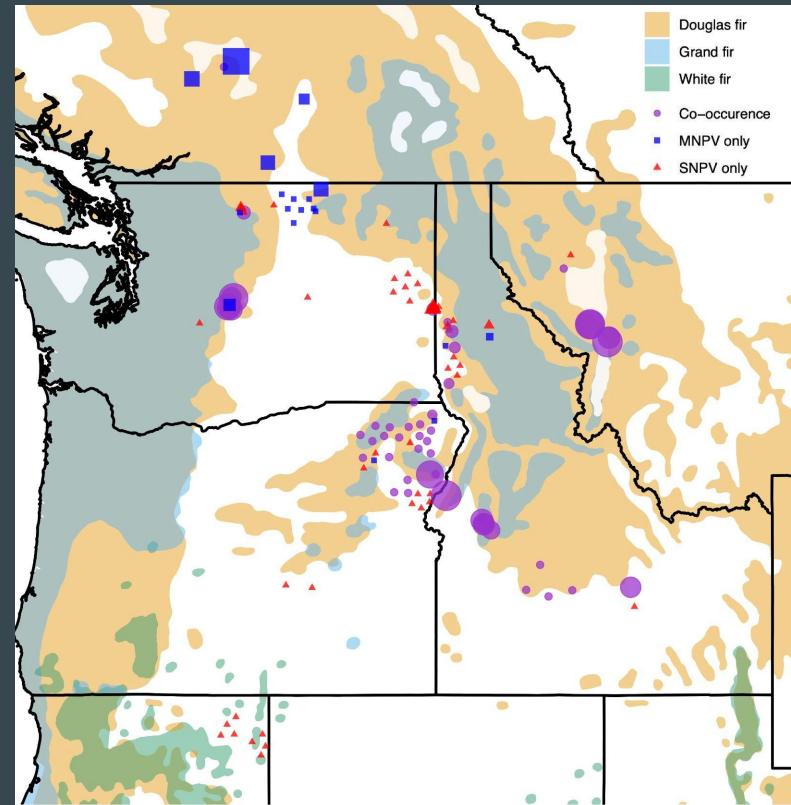
Fun fact: I have sampled all five great lakes!



Katie 6th year Ecology and Evolution (BSD)

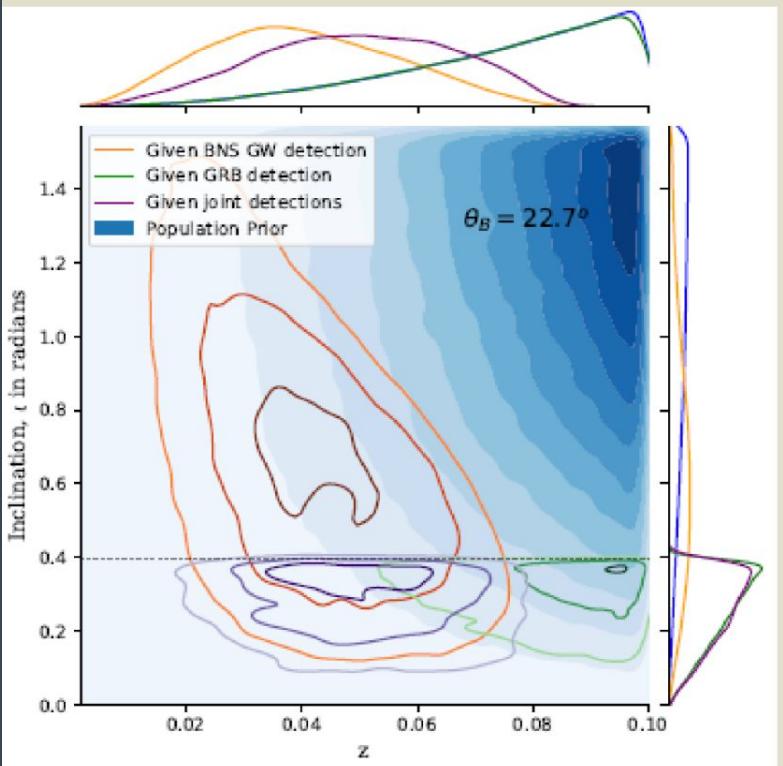
Research: Wildlife disease ecology

Fun fact: I have my advanced scuba certification



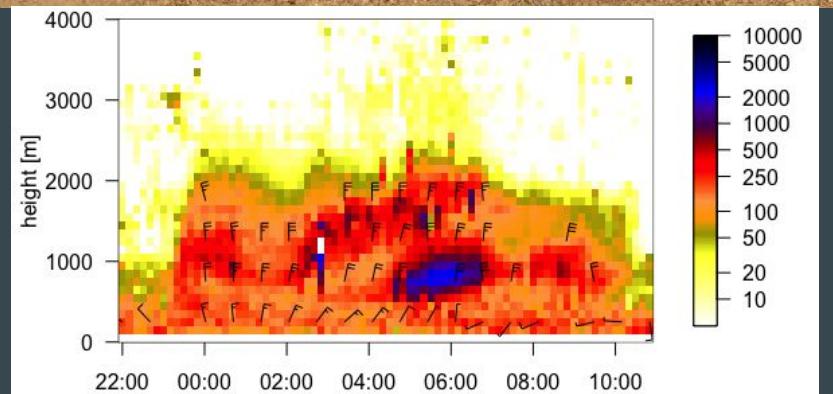
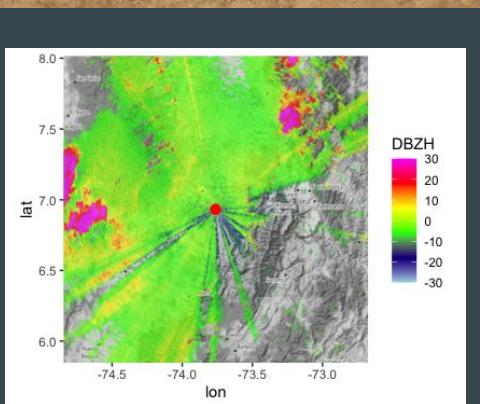
Research Focus: Gravitational Wave Astrophysics

Fun Fact : I'm a new parent to many plant babies



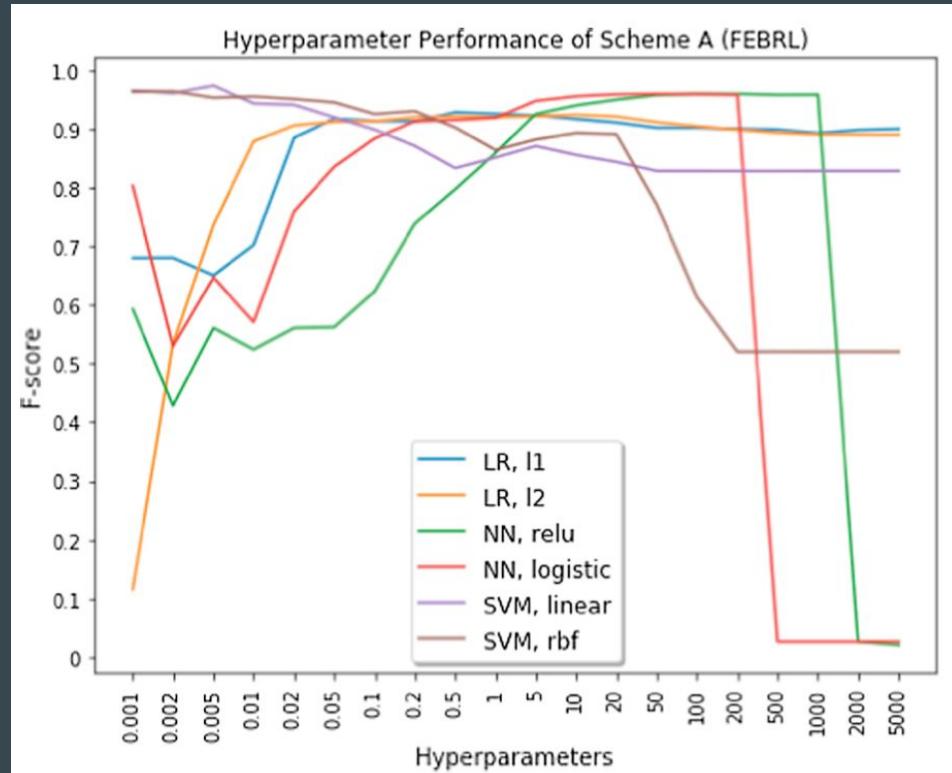
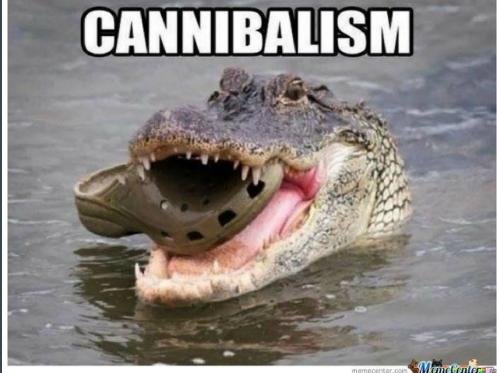
Research focus |
ecological plasticity
in tropical climates

Fun fact |
album drop
in October!



Salman 1st Year MScA (Data Science) Student (PSD)

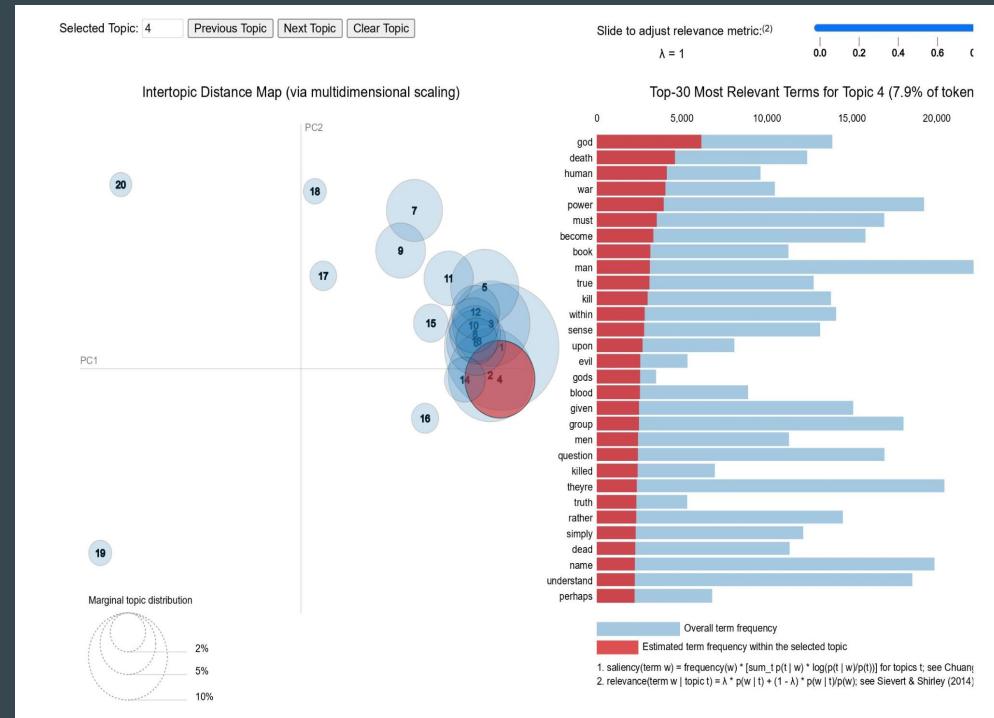
- Data Science / Analytics
- Fun Fact: I love animals and nature!



Shrikanth 2nd year Masters Program in CS (PSD)

Software Engineering/Data Science

Fun fact: I speak 4 $\frac{1}{2}$ languages.



Intro Poll - [Link](#)

Poll Results - [Link](#)

Class intro

• • •

Quick Background poll

Goals:

Students will be able to:



1. Understand coding basics like loops, datatypes, and operators
2. Comfortable with basic python libraries such as numpy, matplotlib, and pandas
3. Comfortable loading in, manipulating, and plotting 2D datasets
4. Comfortable approaching and learning new code

Together we will create a welcoming environment

- Everyone here is learning... to code or to teach so be patient and respectful
- We are students and we chose content that we think will be helpful to you.
- We cannot, in 13 days teach you all of python! But we will teach you how to learn more programming on your own!
- We appreciate any feedback that will improve this course so please reach out with constructive suggestions. **[we will send a short poll at the end of each class asking for feedback]**



Roles

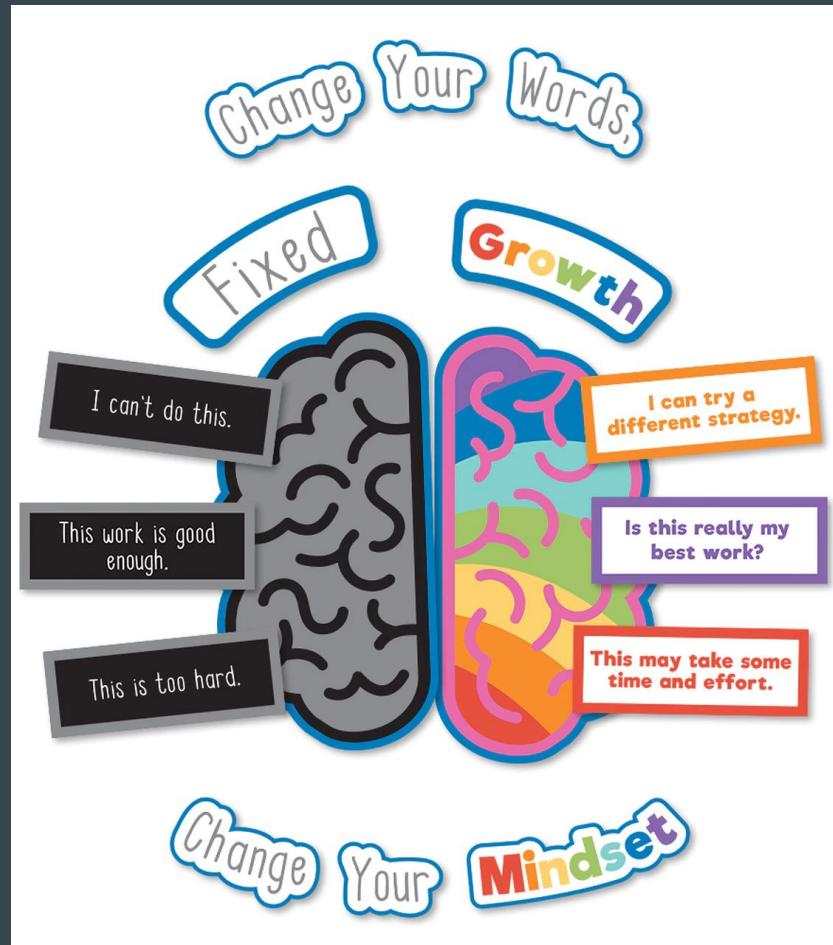
- Instructor:
 - Teach
- Facilitator:
 - Monitor class and slack to make sure everyone is following
 - Work with TAs
 - Gives instructor feedback
 - Launch polls on slack
 - Record lessons
- TAs:
 - Answer questions in person/slack
 - will add an emoji to a question on slack as a note that they are working on an answer
 - Communicate with facilitator
- Students:
 - Follow instructor
 - Ask questions in person/slack
 - Complete homework

Covid19 Protocol

- We **strongly request** that folks in person wear a mask because we will be in the same room for three weeks and we don't want anyone to get sick
- 15-minute coffee breaks will happen outside of the classroom
- If you are **SICK** (symptoms or positive test) please join via zoom

Growth Mindset

- You don't know how to code! Yet!! Having a growth mindset is key!
- Ask questions, we cannot help you if you don't ask for help
- Please use the slack channel and follow the protocols we have in place, we have tried our best to create a system that will give everyone an equal opportunity to ask for help



Course logistics

- Daily sessions 9:30-12:30 pm CT
 - Instructor will go over lessons
 - Lessons will include some problems that we will do in class
 - We will take a few breaks
 - You will have a skill practice homework (short problems to practice what you learned)
 - The instructor may not get through all the code they prepared for the day. It is in your best interest to complete the code on your own and visit office hours with questions.
- Syllabus tour
- We expect that you will attend all classes and complete the skills practice
- Ask TAs questions
 - Send screenshots of your errors

BREAK TIME

Section Two: Programming and Jupyter Notebook

...

Katie

What is programming and why?

Programming is instructing your computer to do various tasks

Why learn to program?

Speed and energy

Computers are much faster and don't need to sleep (generally)

Can automate tasks

Can connect to high performance computers to be even faster

Can create amazing visuals to explain your data analysis

Example of data analysis- Pittsburgh parking meters

1. Extracted the raw data from a large .csv file

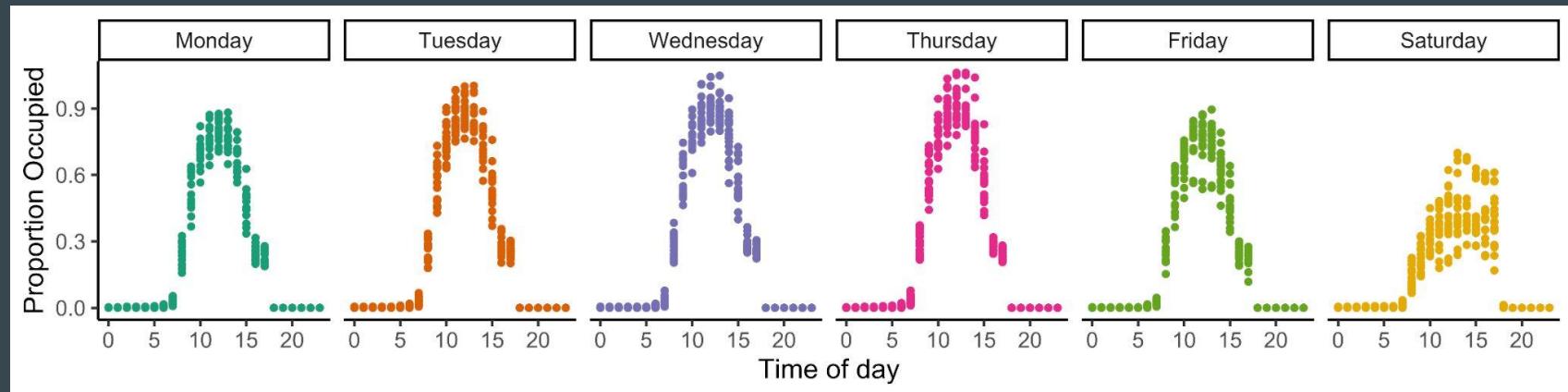
Example of data analysis- Pittsburgh parking meters

1. Extracted the raw data from a large .csv file
2. Reformatted the data, removed data we didn't need, added features

zone_id	occupied	date_adj	time_adj	month	day	year	wday	hour
421 - NorthSide	856	2018-04-02	2018-04-02 13:00:00	Apr	2	2018	Monday	13
421 - NorthSide	839	2018-04-02	2018-04-02 12:00:00	Apr	2	2018	Monday	12
421 - NorthSide	830	2018-04-02	2018-04-02 12:30:00	Apr	2	2018	Monday	12
421 - NorthSide	825	2018-04-02	2018-04-02 13:30:00	Apr	2	2018	Monday	13
421 - NorthSide	805	2018-04-02	2018-04-02 14:00:00	Apr	2	2018	Monday	14
421 - NorthSide	786	2018-04-02	2018-04-02 11:30:00	Apr	2	2018	Monday	11
421 - NorthSide	777	2018-07-11	2018-07-11 12:30:00	Jul	11	2018	Wednesday	12
421 - NorthSide	774	2018-06-07	2018-06-07 13:00:00	Jun	7	2018	Thursday	13
415 - SS & SSW	771	2018-03-24	2018-03-24 15:30:00	Mar	24	2018	Saturday	15
421 - NorthSide	770	2018-08-04	2018-08-04 13:30:00	Aug	4	2018	Saturday	13
415 - SS & SSW	770	2018-10-06	2018-10-06 13:00:00	Oct	6	2018	Saturday	13

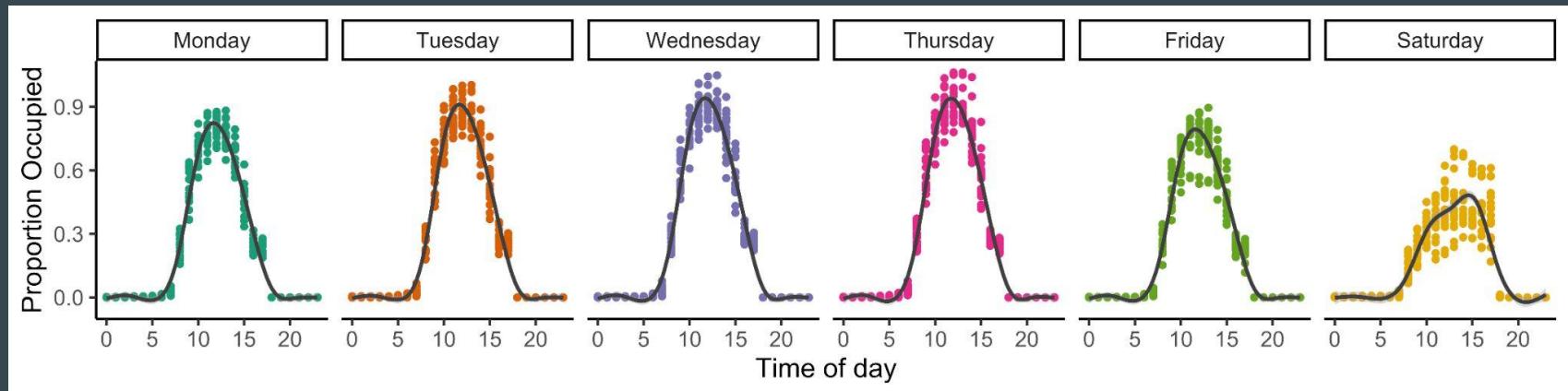
Example of data analysis- Pittsburgh parking meters

1. Extracted the raw data from a large .csv file
2. Reformatted the data, removed data we didn't need, added features
3. Did some exploratory plotting to see how parking rates vary



Example of data analysis- Pittsburgh parking meters

1. Extracted the raw data from a large .csv file
2. Reformatted the data, removed data we didn't need, added features
3. Did some exploratory plotting to see how parking rates vary
4. Used a GAM to fit the time component, quick estimation of parking rates



Computations you can't/ don't want to do by hand

Simulating a SEIR disease model with demography

Impossible/ nearly impossible by hand

Almost instantaneous with a ordinary differential equation (ODE) solver

$$\frac{dS}{dt} = A - \mu S + \frac{\beta SI}{N}$$

$$\frac{dE}{dt} = \frac{\beta SI}{N} - (\mu + \delta)E$$

$$\frac{dI}{dt} = \delta E - (\mu + \gamma)I$$

$$\frac{dR}{dt} = \gamma I - \mu R$$

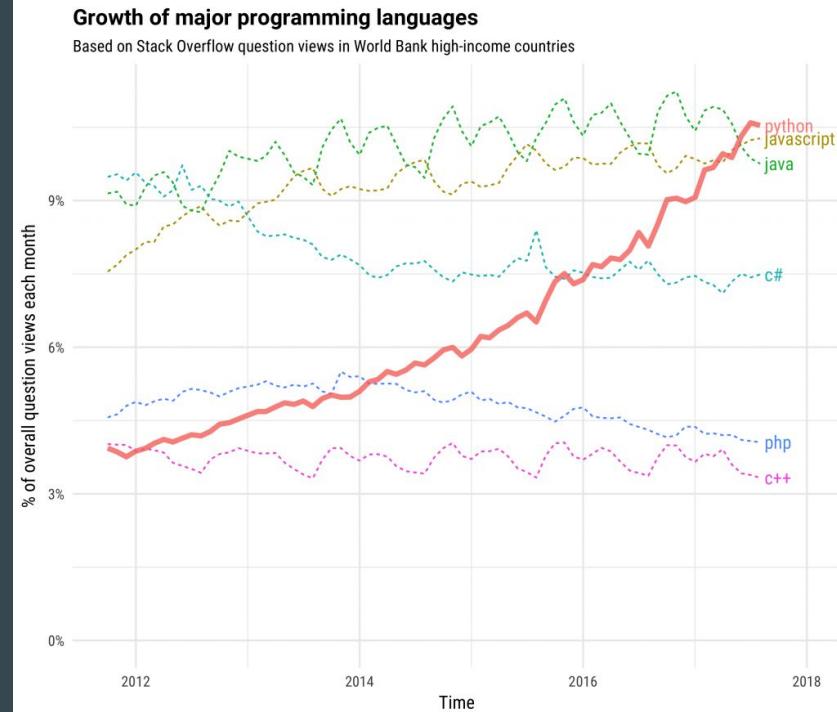
Why Python?

General purpose, **open source** coding language

Very intuitive programming language, simple to learn

Many libraries to do what you want to do

Many, many users so there's lots of documentation and help resources

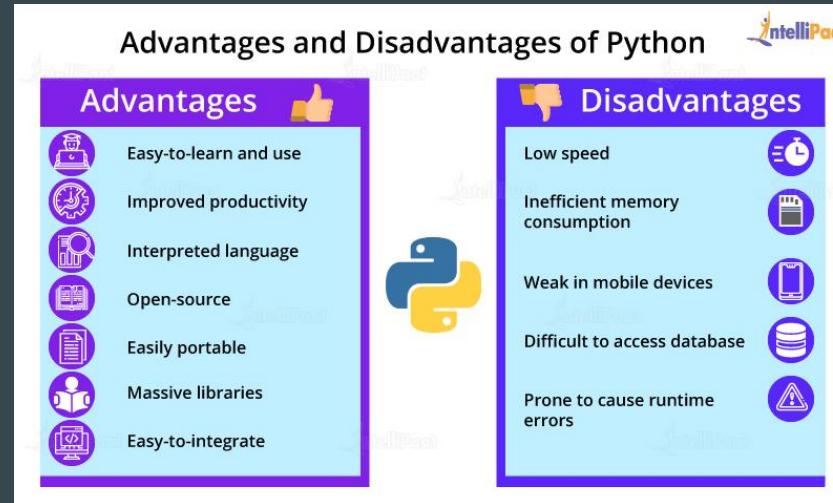


Why Python?

The best data scientists know the strengths and weaknesses of different languages

Code can be translated between languages relatively easily

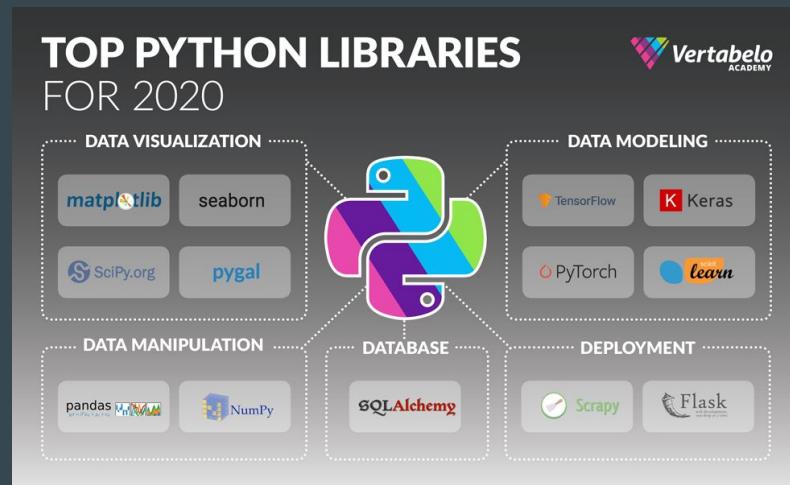
Once you pick up one, the time spent learning a new one goes down



What is a library?

Python Libraries are a set of useful functions that eliminate the need for writing code from scratch.

There are over 137,000 python libraries as of today and they play a vital role in developing machine learning, data science, data visualization, image and data manipulation applications, etc.



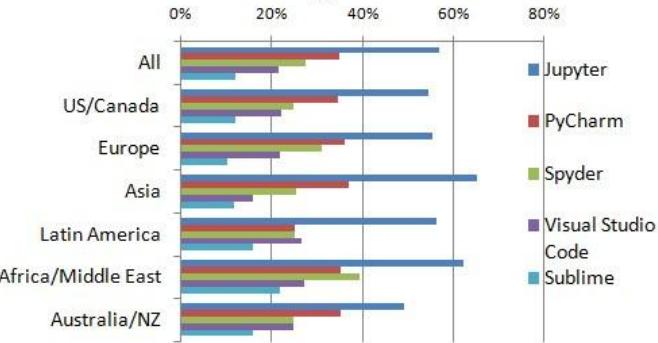
Different environments for Python

Integrated Development Environment (IDE)

Can help you debug

Can be hard to choose which one is the best one

Most Popular Python IDE/Editors by Region



IDLE



SPYDER



PyCharm



eric



Atom



jupyter



Anaconda



PyDev



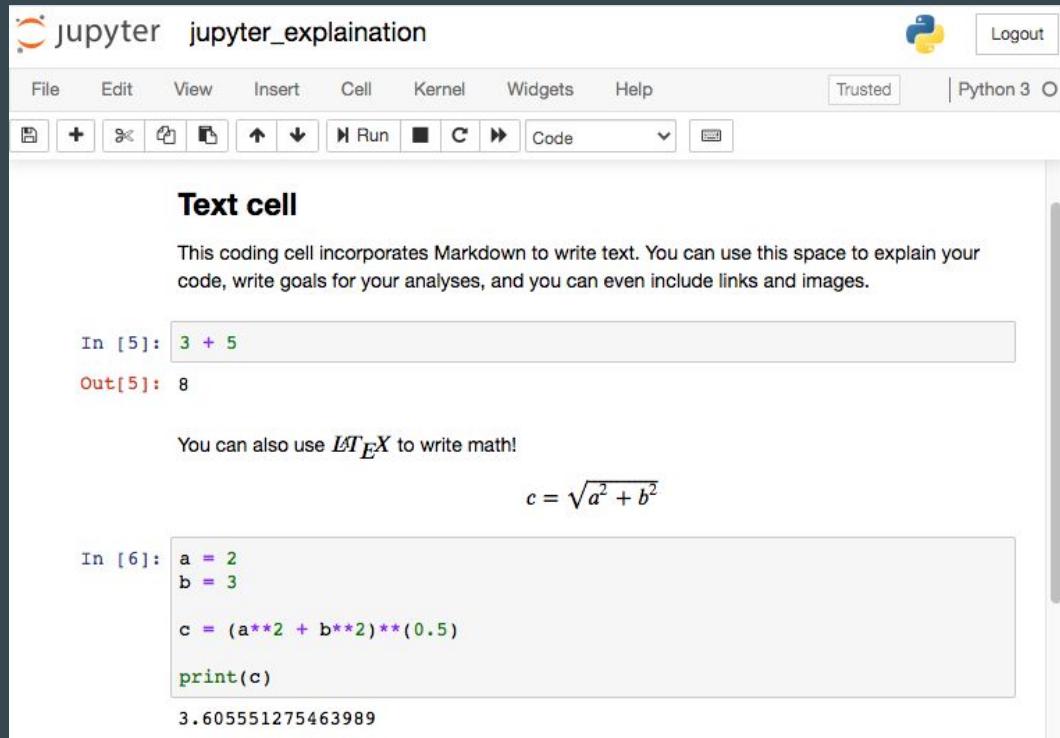
Thonny
Python IDE for beginners

Jupyter Notebooks

Incorporates Markdown to write text

Can use other languages using a kernel (R, Julia, SQL, etc.)

Readable, **reproducible** interface



The screenshot shows a Jupyter Notebook interface with the title "jupyter_explaination". The top menu bar includes File, Edit, View, Insert, Cell, Kernel, Widgets, Help, a Trusted button, and Python 3. The toolbar below the menu has icons for file operations like Open, Save, and Run, along with a Code dropdown.

Text cell:
This coding cell incorporates Markdown to write text. You can use this space to explain your code, write goals for your analyses, and you can even include links and images.

In [5]: `3 + 5`
Out[5]: `8`

You can also use *LaTeX* to write math!

$$c = \sqrt{a^2 + b^2}$$

In [6]:

```
a = 2  
b = 3  
  
c = (a**2 + b**2)**(0.5)  
  
print(c)
```

3.605551275463989

Move to jupyter notebook

Resources

Although we tried, we can't cover everything

Expert coders are still learning something new all the time

Error messages: sometimes helpful! Read them and learn from the mistakes

Expert coders make errors all the time

I still make the same mistakes as I did 9 years ago, now I'm faster at fixing

In this course: your TAs are here to help!



Questions for us?

Raise your hand and we will call on you

Note: we will not be able to do this other days. In the future, ask in the slack!

BREAK TIME

Section Three: Setup for the Course

...

Amanda