

---

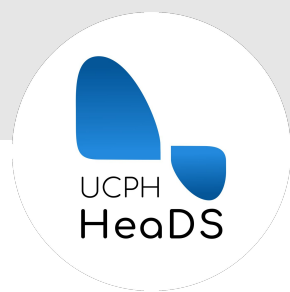
# Python Tsunami

Part 1: Intro to Python by HeaDS

---

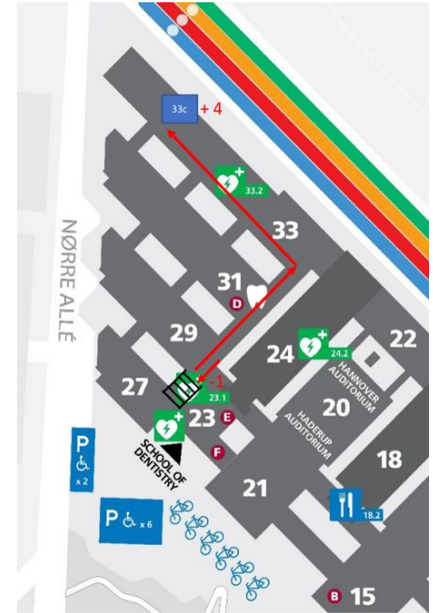
# Who are we?

# Center for Health Data Science (HeaDS)



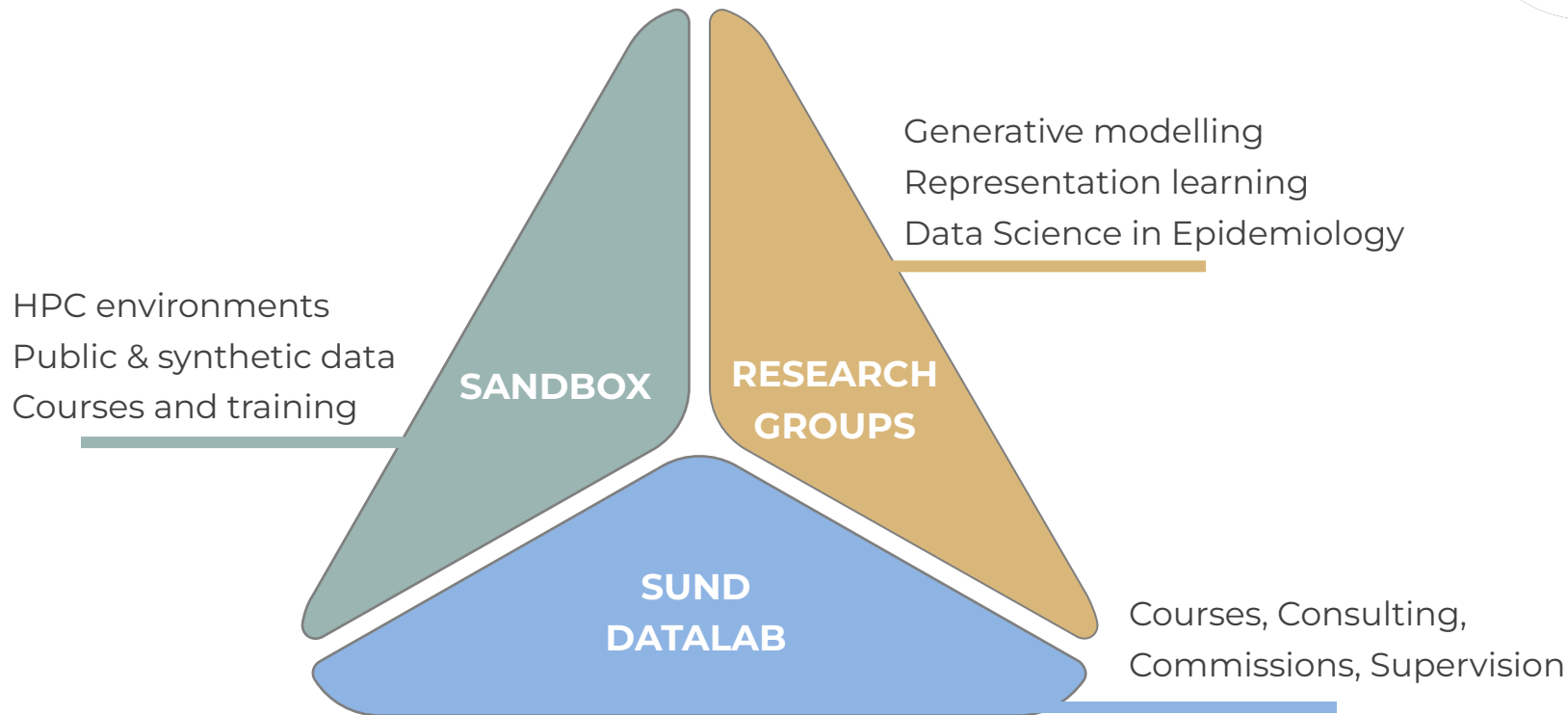
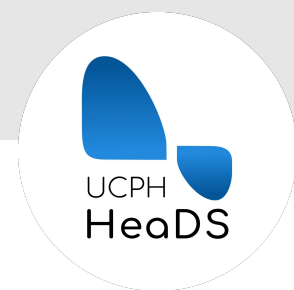
The mission of the Center is to strengthen health data science within the Faculty:

- Active and visible hub for Health Data Science
- Providing data science support for researchers at SUND
- Courses, workshops and training environments to improve data science skills
- Support a network of researchers and educators



<https://heads.ku.dk>

# CENTER STRUCTURE



---

# About this course

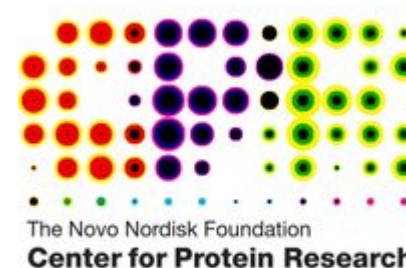
# About this course

---



Originally developed at the Center for Protein Research (CPR) by:

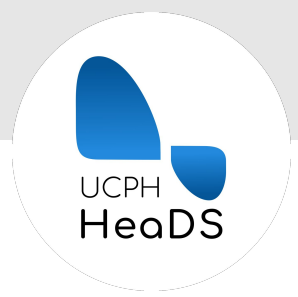
- Alberto Santos Delgado (University of Oxford)
- Henry Webel (NNF CPR)
- Annelaura Bach Nielsen (NNF CPR)
- Rita Colaço (PRI)



We say thank you for the course material which we have adapted.

# About this course

---



## Your teachers:

Thilde Terkelsen (HeaDS)

Rita Colaço (PRI)

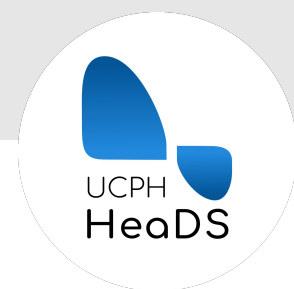
Valentina Sora (HeaDS)

Henrike Zschach (HeaDS)

Inigo Prada Luengo (HeaDS)



# About this course



Starting time	Day 1 (Hannover Aud.)	Day 2 (Holst Aud.)	Day 3 (Holst Aud.)
8:30	Morning coffee (optional)		
8:45	Motivation	Pandas: Series and dataframes, import and examine data, Renaming index/column	Yesterday Questions + Recap Quiz
9:05	Variables & Data types		
9:45	Coffee break		Coffee break
10	Iterables I: Lists		Coffee break
10:15			
10:30		Pandas: Indexing and Selecting Data, Summary functions	
11:30	Iterables II: sets, dicts, tuples		
12:00	Lunch		
13:00	Booleans, operators & conditions	Pandas: Modifying data	Virtual environments and local python installations
13:30			
14:00			Coffee break
14:15	Coffee break	Pandas: GroupBy Operations, Sorting and Imputation	
14:30	Loops		
15:30			Pandas Q + A
16:00	- END -		



---

# Python

# Why Python?

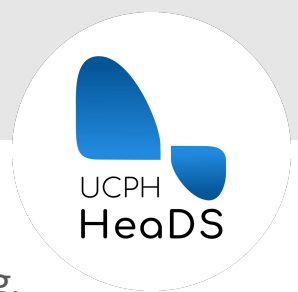
---




Python is a great programming language for both beginners and advanced programmers:

- Easy to grasp, close to natural language
- Many learning resources available
- Large community (i.e. stackoverflow for questions)
- Libraries
- Can do very advanced things like neural networks

# Where to get help



Online communities such as stackoverflow are an important tool in programming.



[All](#) [Images](#) [Videos](#) [News](#) [Maps](#) [Settings](#)

☐ Denmark ☐ Safe search: moderate ☐ Any time

[https://stackoverflow.com > questions > 28256761 > select-pandas-rows-by-excluding-index-num...](https://stackoverflow.com/questions/28256761/select-pandas-rows-by-excluding-index-num...)

**select pandas rows by excluding index number - Stack Overflow**

I'm looking to slice a **Pandas** dataframe by using index numbers. I have a list/core index with the index numbers that i do NOT need, shown below **pandas**. Stack Overflow ... Use a list of values to **select rows** from a **Pandas** dataframe. 2015. Delete a column from a **Pandas** DataFrame. 1290. How to drop **rows** of **Pandas** DataFrame whose value in a certain ...

**select pandas rows by excluding index number**

<http://stackoverflow.com/questions/28256761/ddg#28256...>

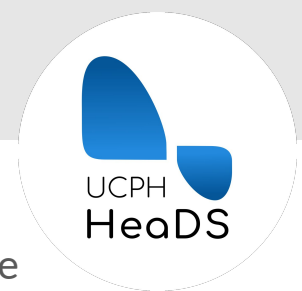
Not sure if that's what you are looking for, posting this as an answer, cause it's too long for a comment:

You can also ask chatGPT, but remember to verify the answer!

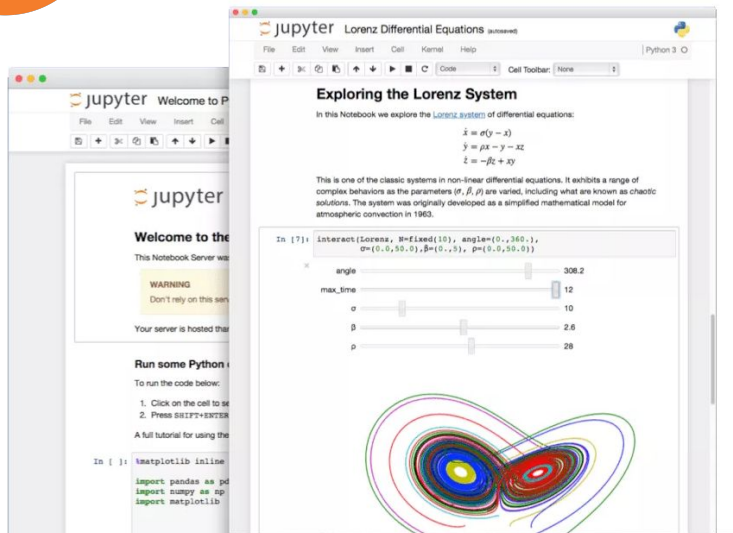
---

# Python environments

# Jupyter notebook



The Jupyter Notebook is an **open-source application** to create and share documents that contain code, equations, visualizations and text (markdown).



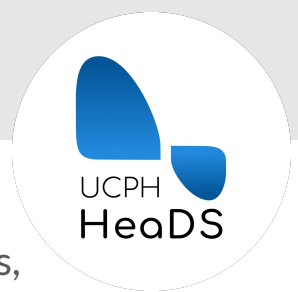
- Browser-based development environment for creating, running and sharing Python code
- Combine code with text and output
- Runs from your **local installation**. I.e., you need Python and the libraries you want to use installed on your computer



Google Colab is a Jupyter Notebook hosted on Google's servers, not your own machine. It still runs in your browser.

- tool to write, execute and share Python code through the browser
- requires no setup to use and provides free access to computing resources on Google's servers including GPUs
- is connected to a Google account and data and notebooks can be accessed through Google Drive.

We'll use Colab during the course.



---

# Using libraries/packages

Python has many libraries, also called packages, that other programmers have developed. Find and **use** them!

Well-maintained libraries generally are:

- Tested
- Optimized
- Documented

There is no need to reinvent the wheel. During this course we will use:

- Pandas (all the data analysis!)
- Math (basic math)
- Plotly express (visualization)



If you are running Python from a local installation, you need to have libraries **installed** before you can use them.

On Google Colab you can generally just import, they are already installed.

- Import the math library:

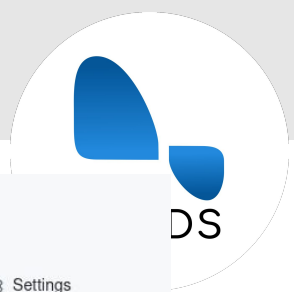
```
import math
```

- Now I can use functions from that library, i.e. calculating the logarithm or square root:

```
math.log(3)
```

```
math.sqrt(4)
```

# Course material



Center-for-Health-Data-Science / PythonTsunami Public  
forked from pythontsunami/teaching

<> Code Issues 10 Pull requests Discussions Actions Projects Wiki Security Insights Settings

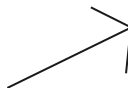
You can find the course material here:

<https://github.com/Center-for-Health-Data-Science/PythonTsunami>

spring2022 16 branches 0 tags  
Go to file Add file Code

This branch is 111 commits ahead of pythontsunami/teaching:heads. Contribute Fetch upstream

hezscha Add files via upload 078c67a 2 days ago 376 commits	
Conditionals	Add files via upload 6 days ago
Exercise	Add files via upload 2 days ago
Functions	Add files via upload 2 days ago
Introduction_and_tools	Add files via upload 7 days ago
Iterables	Add files via upload 2 days ago
Loops	Add files via upload 2 days ago
Pandas	Minor changes to pandas examples 6 days ago
Recap	Add files via upload 9 days ago
Variables_data_types	Add files via upload 6 days ago



# Course material



spring2022 PythonTsunami / Variables\_data\_types /

Go to file Add file ...



This branch is 111 commits ahead of pythontsunami/teaching:heads. Contribute Fetch upstream

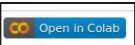
hezscha Add files via upload ab1c127 6 days ago History

..

README.md	Minor fixes	7 days ago
variables.ipynb	Update Colab link within the notebook	6 days ago
variables_solutions.ipynb	Add files via upload	6 days ago

README.md



notebook	content
<a href="#">variables.ipynb</a> 	Variables and data types



# Course material



Remember to **save a copy** to your own google drive so you can save your notes and exercises!



# Short Introduction

---



Take the next 5 mins to introduce yourself at your table:

- Name
- Position
- Unit
- What you do (very briefly!)