

Learning outcomes

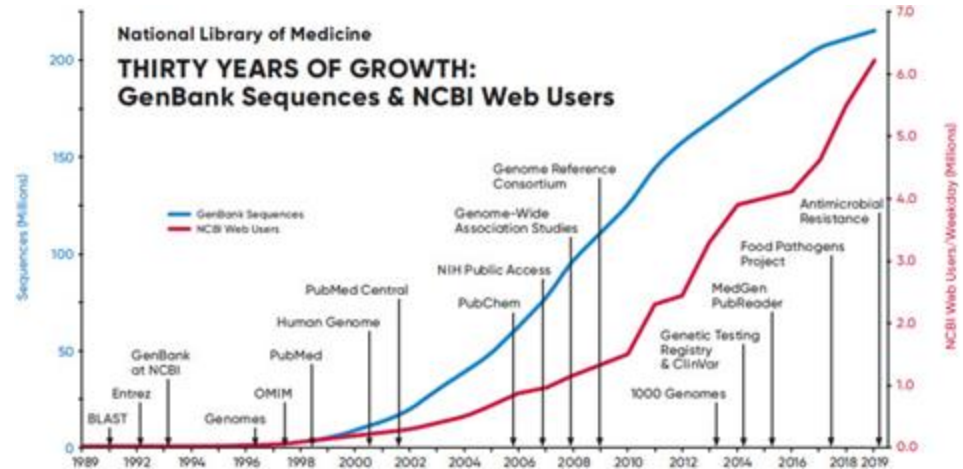
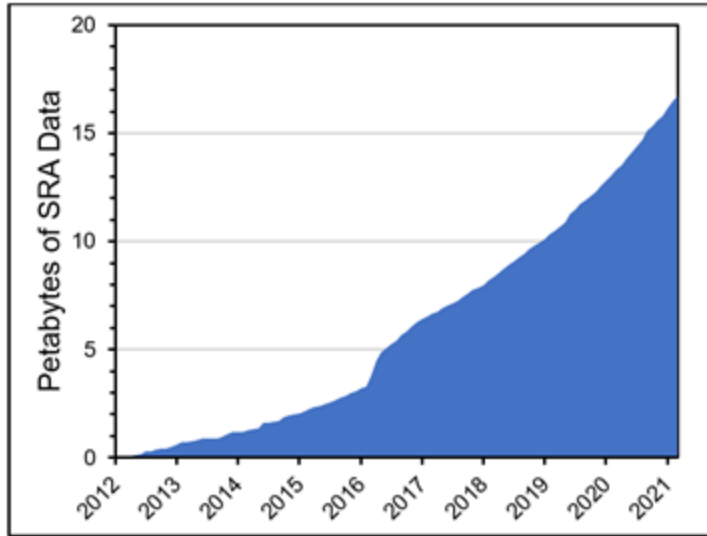
By the end of this workshop, you will have tools to select, gather and prepare data for analysis, and will be familiar with statistical methods to analyze the data.

You will learn the basis of:

- **Analytical methods** for processing, describing and visualizing complex biomedical data, with the help of AI;
- **Terminology** for statistical tools and machine learning in biomedical discovery.

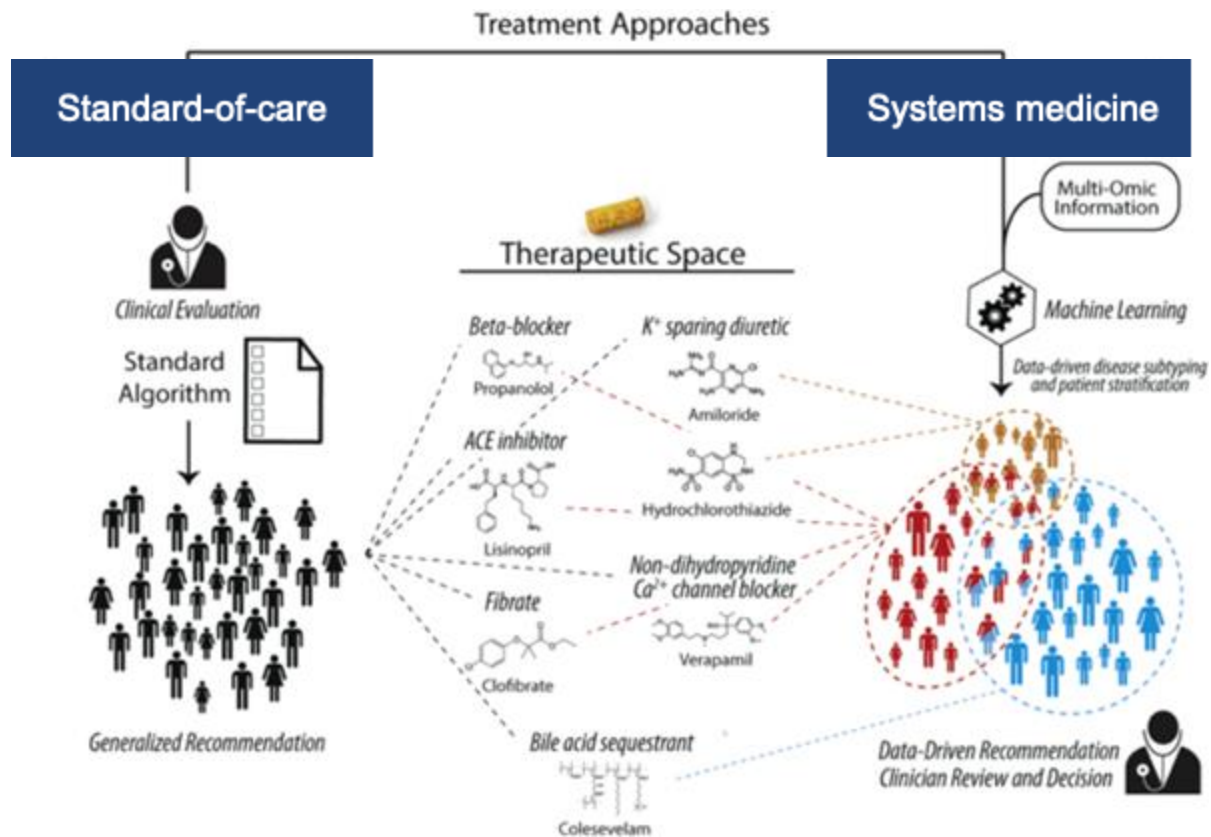
From data analysis to meaningful biological and clinical insights

Technological advances = more data = more insights

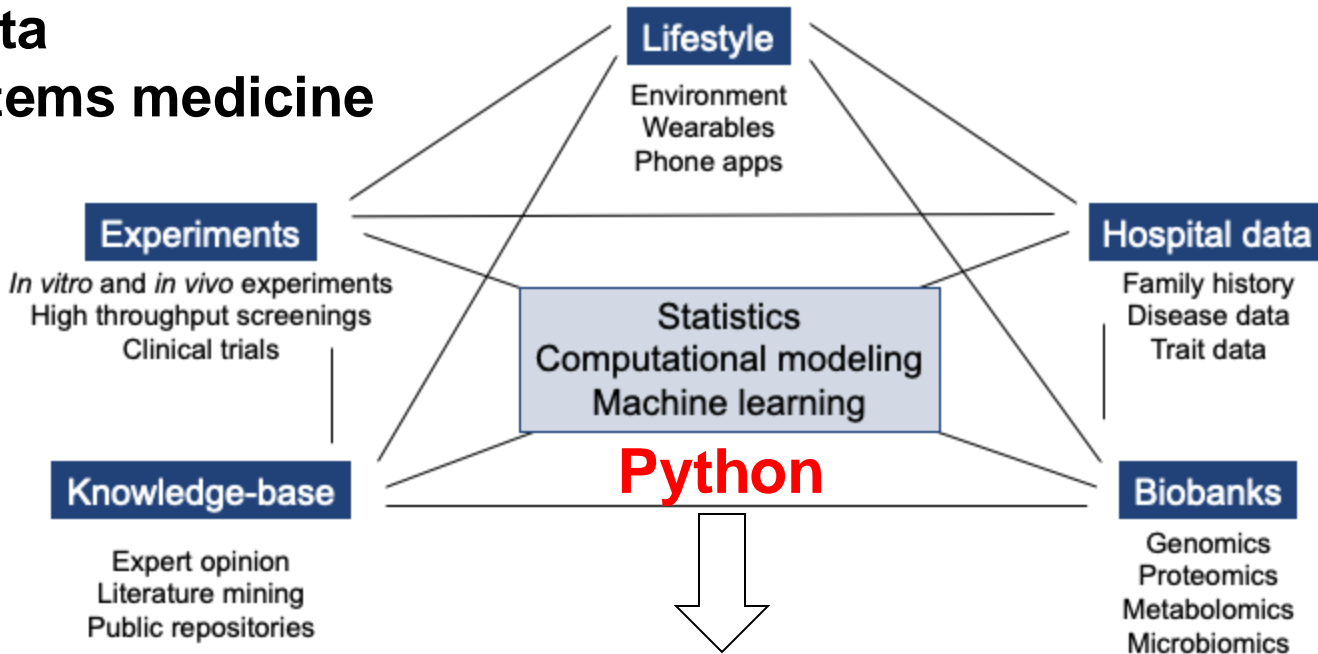


Problem: how to deal with all this data? **Data science (and this workshop!)**

Why big data analysis: standard-of-care vs systems medicine



Big data in systems medicine



Greater understanding of onset, progression of disease

Understanding of molecular mechanisms

Disease prevention

Early diagnostic

Personalized treatments (more effective treatments, less side effects)



pythonTM

Introduction to Python

What is Python?

- Python is a high-level, interpreted, interactive and object-oriented programming language used for general-purpose software engineering.
- It was designed initially by Guido Van Rossum in 1991 and mainly developed for emphasis on code readability and easy implementation → Easy to learn!
- It is an open-source and one of the most popular programming language for data science (with a strong community of users in machine learning, artificial intelligence, data modeling, etc).

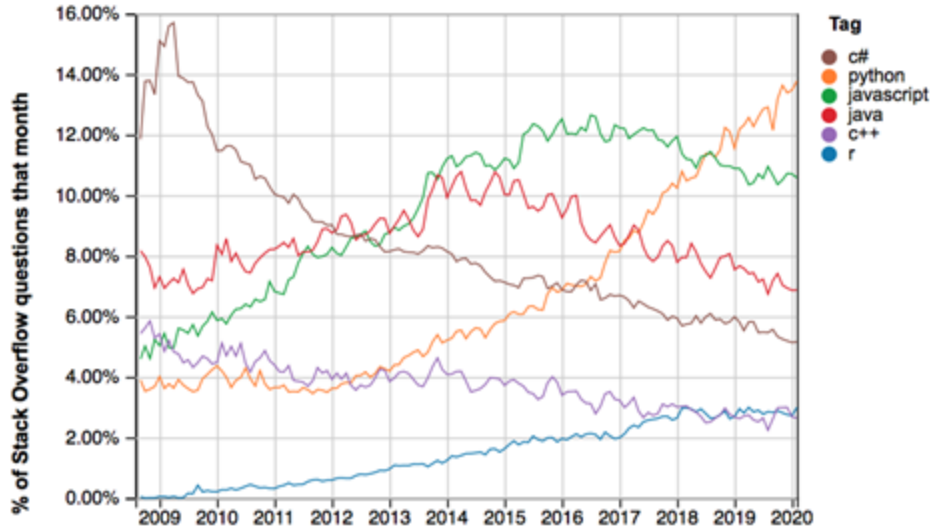
Shall you code with the strength of many men, Sir Knight!



Why is Python?

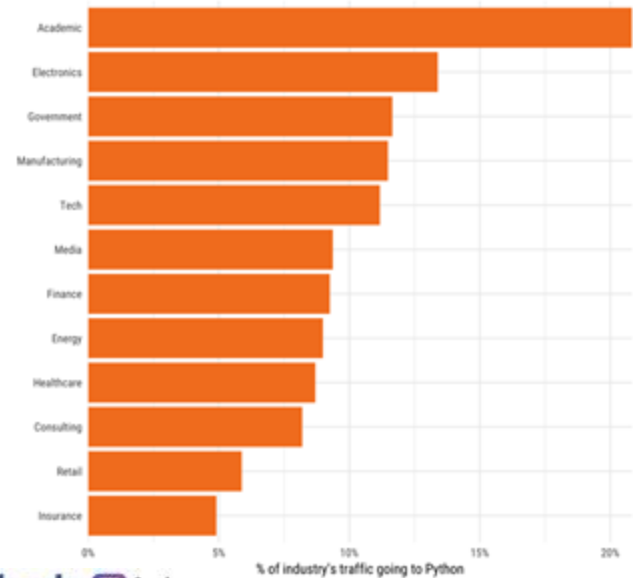
- **Versatile:**
 - Python is used in Data Mining, Data Science, AI, Machine Learning, Web Development, Web Frameworks, Embedded Systems, Graphic Design applications, Gaming, Network development, Product development, Rapid Application Development, Testing, Automation Scripting...
 - It offers great flexibility and has extensive open-source libraries for multiple applications: Pandas, numpy, scikit-learn, tensorflow, etc.
 - Useful in every industry (healthcare, financial services, marketing, education, etc)
- **Easy:** High readability, simple syntax. Easier to maintain and more efficiently-written alternative to languages that perform similar functionalities like C, R, and Java.
- **Community-based Support:** StackOverflow has answers to most of your questions already.
- **Scalable:** great for quick prototyping, and major projects
- **Fastest growing programming language:** it can keep up with technologies and needs
- **Now you can use AI to generate code in Python!**

Python popularity vs other programming languages and by industry



Visits to Python by industry

Based on visits to Stack Overflow questions from the US/UK in January-August 2017. The denominator in each is the total traffic from that industry.



MAJOR COMPANIES THAT USE



python



Introduction to Colaboratory Notebooks



What is a Notebook

The Notebook is a web based interactive computational environment which provides a unique combination of code, shell environment and text. Three components:

The notebook web application: An interactive web application for writing and running code interactively and authoring notebook documents.

Kernels: Separate processes started by the notebook web application that runs user's code in a given language and returns output back to the notebook web application.

Notebook documents: Self-contained documents that contain a representation of all the contents visible in the notebook web application, including inputs and outputs of the computations, narrative text, equations, images, and rich media representations of objects. Each notebook document has its own kernel.

Introduction to Colaboratory Notebooks



Different Cell Types?

Notebooks consist of a linear sequence of cells. There are four basic cell types:

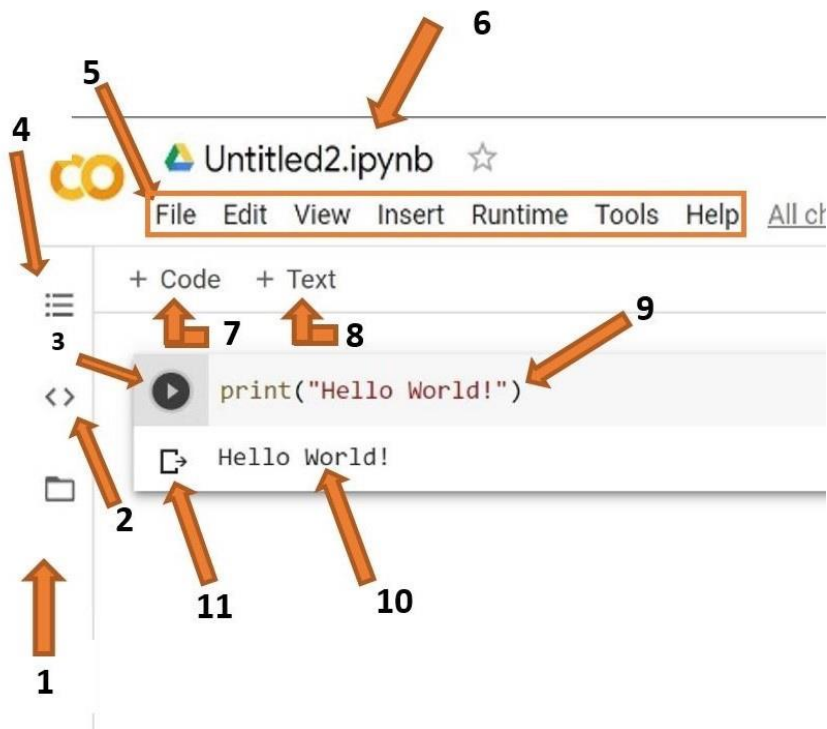
Code cells: Input and output of live code that is run in the kernel

Markdown cells: Narrative text with embedded LaTeX equations

Heading cells: 6 levels of hierarchical organization and formatting

Raw cells: Unformatted text that is included, without modification, when notebooks are converted to different formats using nbconvert These cell types can be viewed by clicking Cell -> Cell Type in menu bar

The anatomy of a Notebook



- 1.Files:** Here you will be able to upload datasets and other files from both your computer and Google Drive
- 2.Code Snippets:** Here you will be able to find prewritten snippets of code for different functionalities like adding new libraries or referencing one cell from another.
- 3.Run Cell:** This is the run button. Clicking this will run any code that is inserted in the cell beside it. You can use the shortcut shift+enter to run the current cell and exit to a new one.
- 4.Table of Contents:** Here you will be able to create and traverse different sections inside of your notebook. Sections allow you to organize your code and improve readability.
- 5.Menu Bar:** Like in any other application, this menu bar can be used to manipulate the entire file or add new files. Look over the different tabs and familiarize yourself with the different options. In particular, make sure you know how to upload or open a notebook and download the notebook (all of these options are under "File").
- 6.File Name:** This is the name of your file. You can click on it to change the name. Do not edit the extension (.ipynb) while editing the file name as this might make your file unopenable.
- 7.Insert Code Cell:** This button will add a code cell below the cell you currently have selected.
- 8.Insert Text Cell:** This button will add a text cell below the cell you currently have selected.
- 9.Cell:** This is the cell. This is where you can write your code or add text depending on the type of cell it is.
- 10.Output:** This is the output of your code, including any errors, will be shown.
- 11.Clear Output:** This button will remove the output.

Where will be the material for this workshop?

systemsmedicine.univr.it → Open Science

https://github.com/MaximilianNuber/ICOCIMS_2024_python_workshop

When you need help:

<https://stackoverflow.com/>