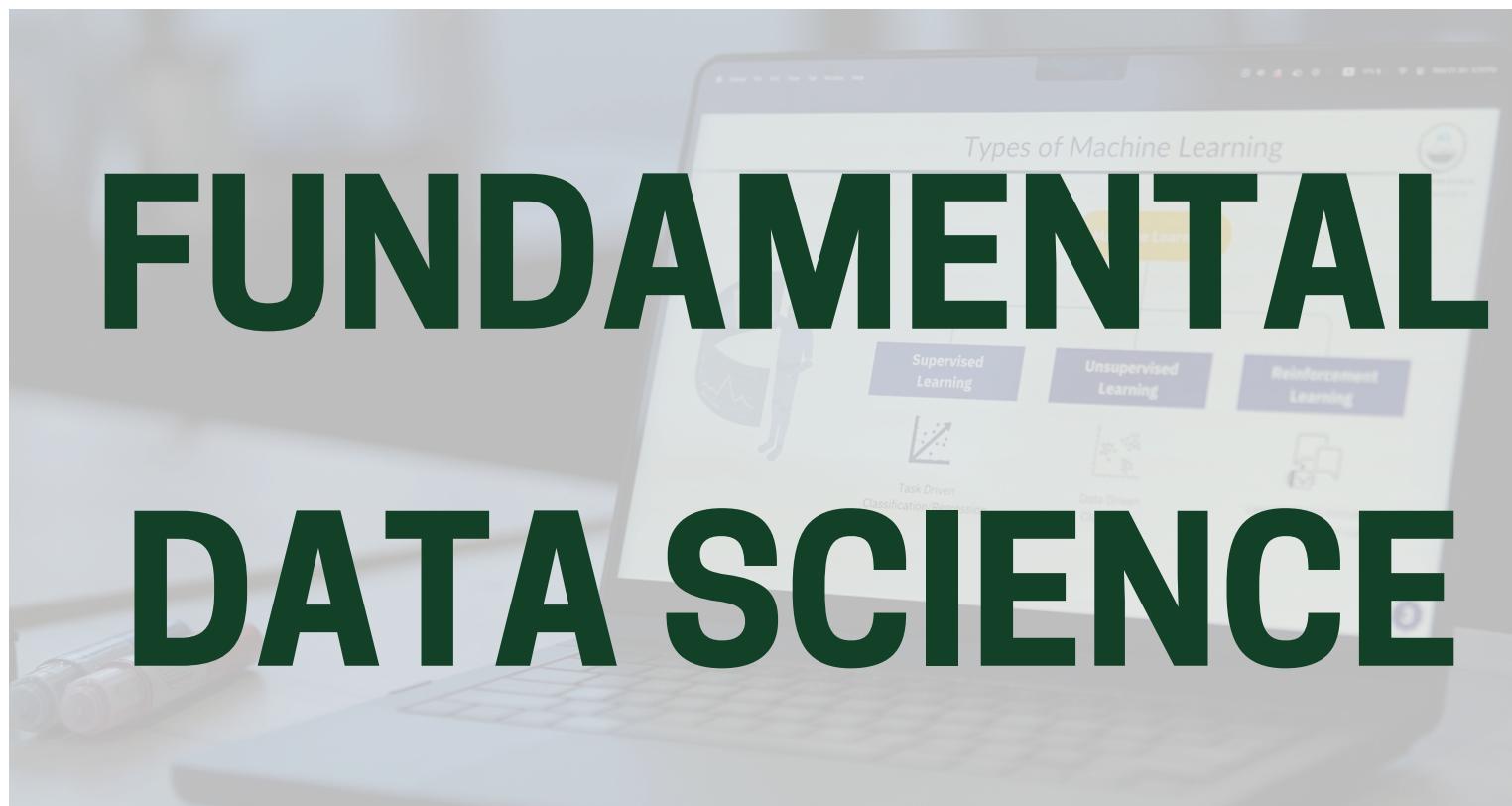




ත්‍රිජාත්‍යාල සීල් රුහුණ්

SUNRISE INSTITUTE



EMBARKING ON A JOURNEY
INTO DATA SCIENCE

YA MANON



You can have data without information but you
cannot have information without data.

-Daniel Keys Maran

COURSE STRUCTURE



This is a **project-based** course for students looking for a practical, hands-on, and highly engaging approach to learning, Data Science, Python essentials for data analysis.

Additional resources include:

- ☀ **Source Code Link**: to serve as a helpful reference when you're offline or on the go
- ☀ **Quizzes & Assignments** to test and reinforce key concepts, with step-by-step solutions
- ☀ **Interactive demos** to keep you engaged and apply your skills throughout the course

SETTING EXPECTATIONS



Who this is for:

- Analysts or BI professionals looking to learn **Python for data analysis** and **Data Science**
- Students looking to learn the most popular open-source analytics tool (free)
- Anyone who wants to understand the core fundamentals of the Python language and syntax



Who this is NOT for:

- Experienced Python programmers or advanced users
- Students looking to learn Python for Software or Web Development
- Anyone who would rather copy and paste code or run packages without building the foundational skills

“Data is the new oil”

British mathematician
Clive Humby, 2006



INTRO TO DATA SCIENCE



In this section we'll **introduce the field of data science**, discuss how it compares to other data fields, and walk through each phase of the data science workflow

TOPICS WE'LL COVER:

What is Data Science

What is Artificial Intelligence AI?

Data Science vs. AI

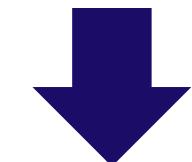
Data Science vs. Data Analytics

GOALS FOR THIS SECTION:

- Compare data science and machine learning with other common data analytics fields
- Introduce supervised and unsupervised learning, and examples of each technique
- Review the machine learning landscape and commonly used algorithms
- Discuss essential skills, and review each phase of the data science workflow

What is Data Science?

What's Data Science?



- 1 **The study of data to extract meaningful insights for business.**
- 2 **The process of creating physical models to simulate real-world events without using any data.**
- 3 **A branch of engineering focused solely on designing hardware components for data storage systems.**
- 4 **The study of astronomical phenomena using telescopes and computational tools.**

What is Data?

Data is a collection of information gathered by observations, measurements, research or analysis.

Data Types

Structures data

Dimension (rows, columns),
csv, excel , ..

Example: phone number, zip codes,
identity number, gender,..

Unstructures data

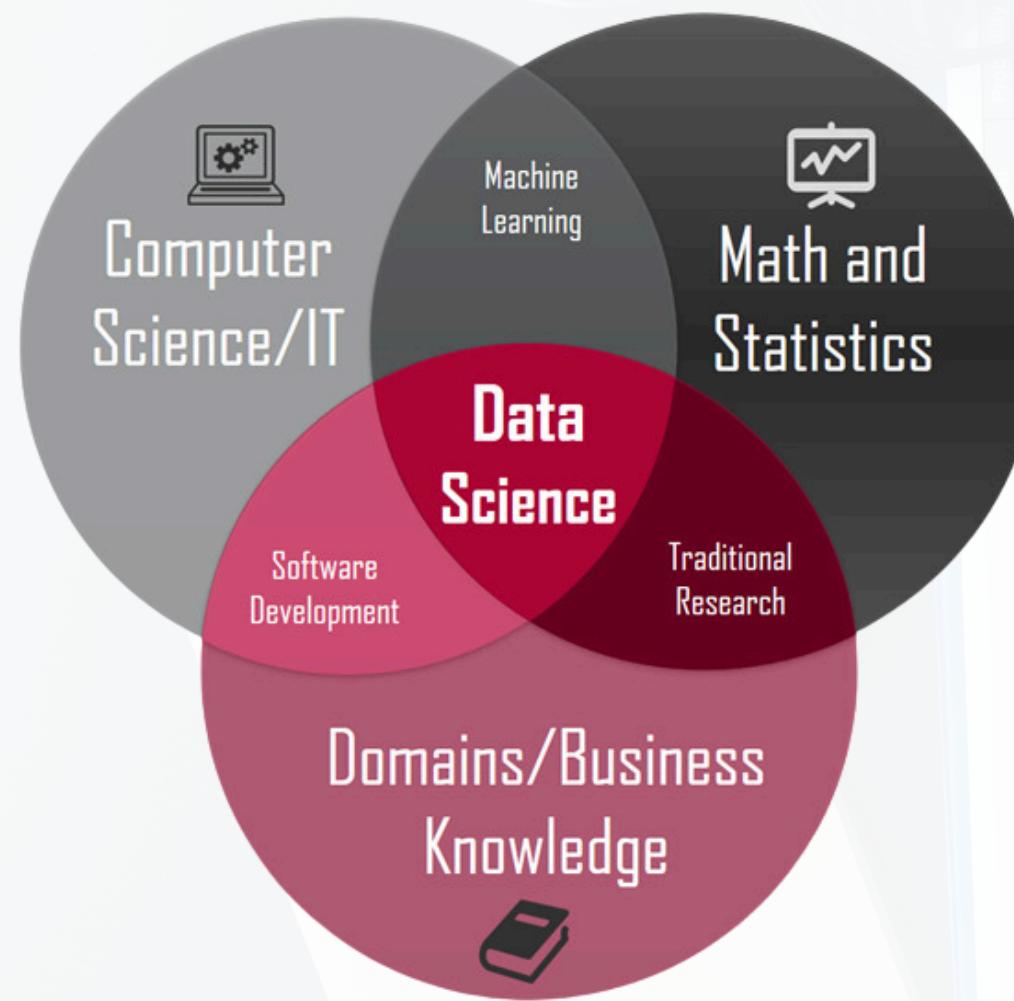
Everything else (no rows,
columns)...

Example: emails, social media post, video,
sensor, data...

Data Science

1

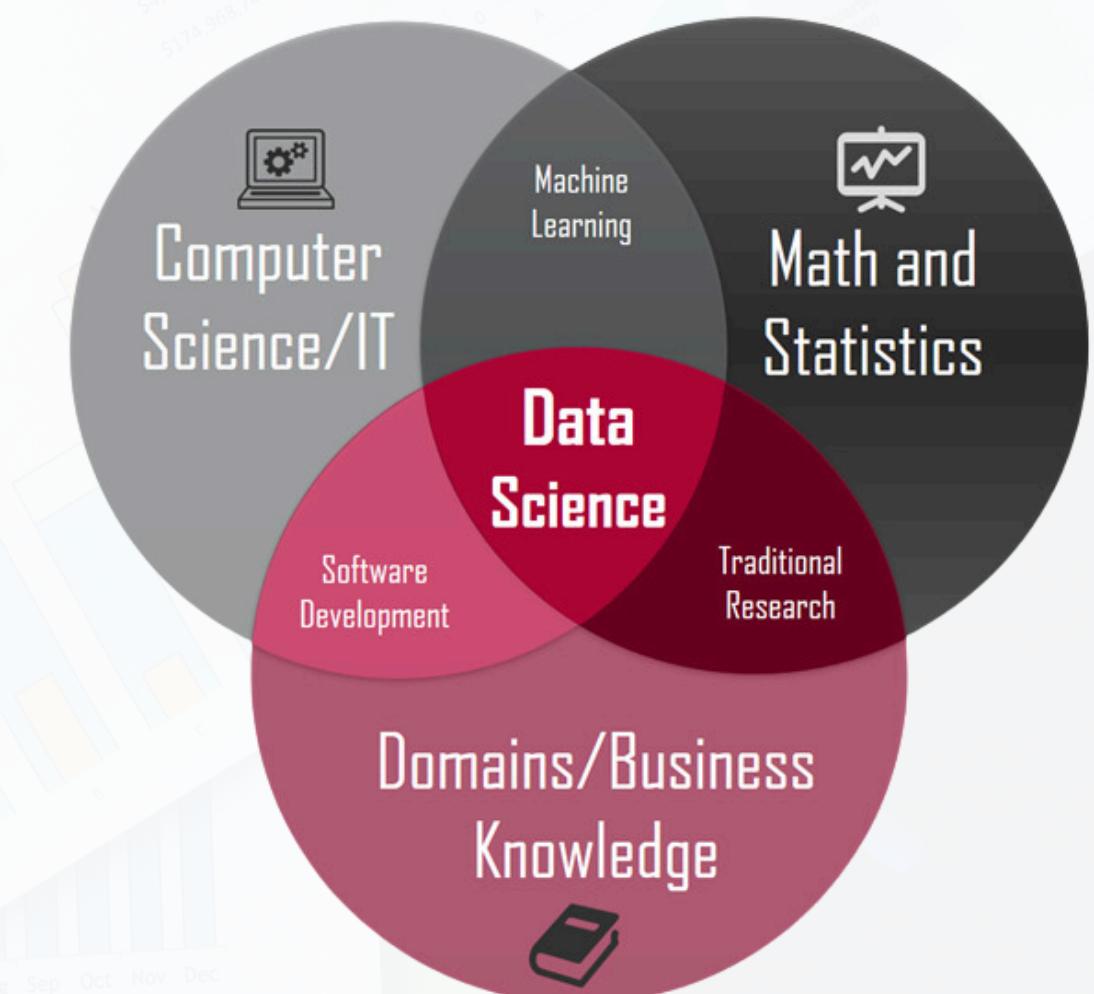
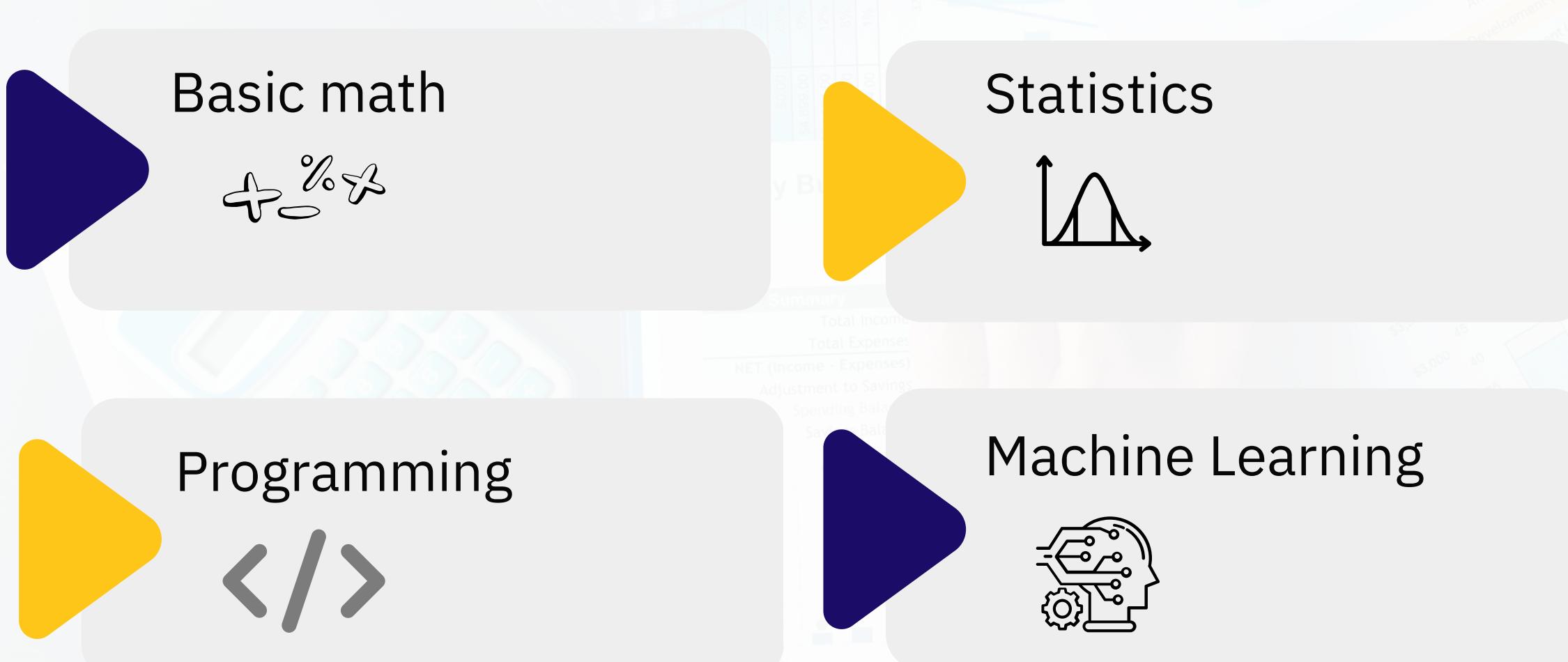
The study of data to extract meaningful insights for business.



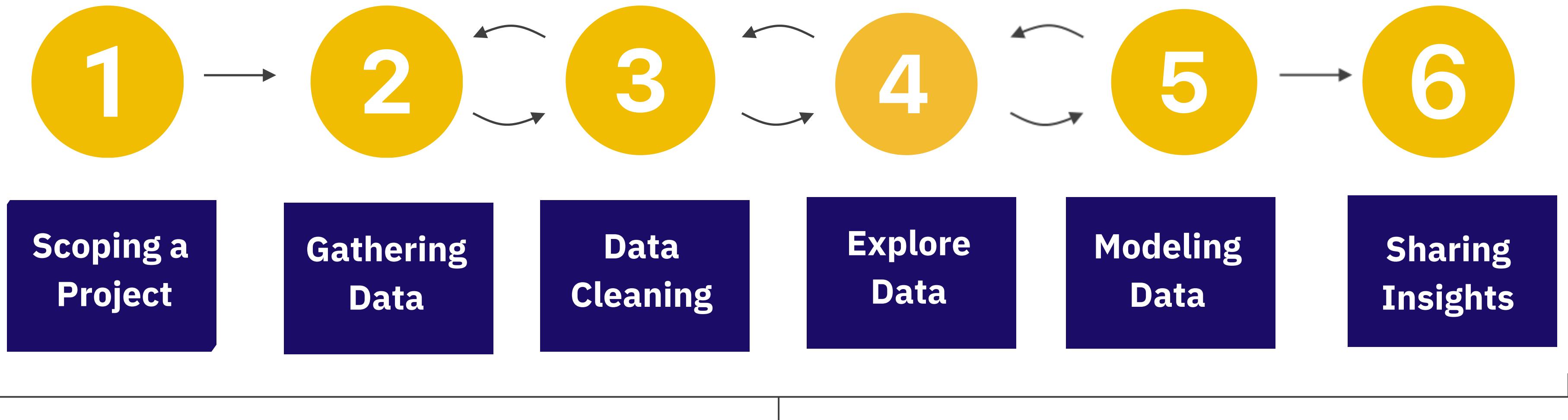
Data science is about using data to make smart decisions.

Data Science

Data science combines **math** and **statistics**, specialized **programming**, **advanced analytics**, **a** and **machine learning** with specific subject matter expertise to uncover actionable insights hidden in an organization's data.



Data Science Workflow



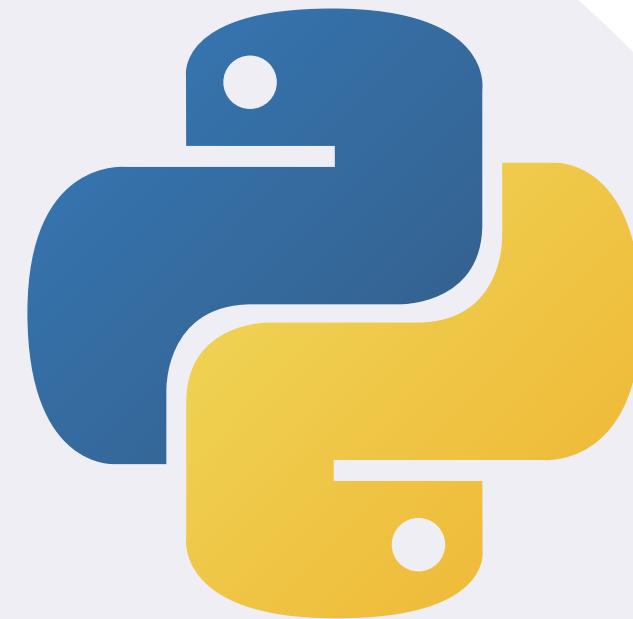
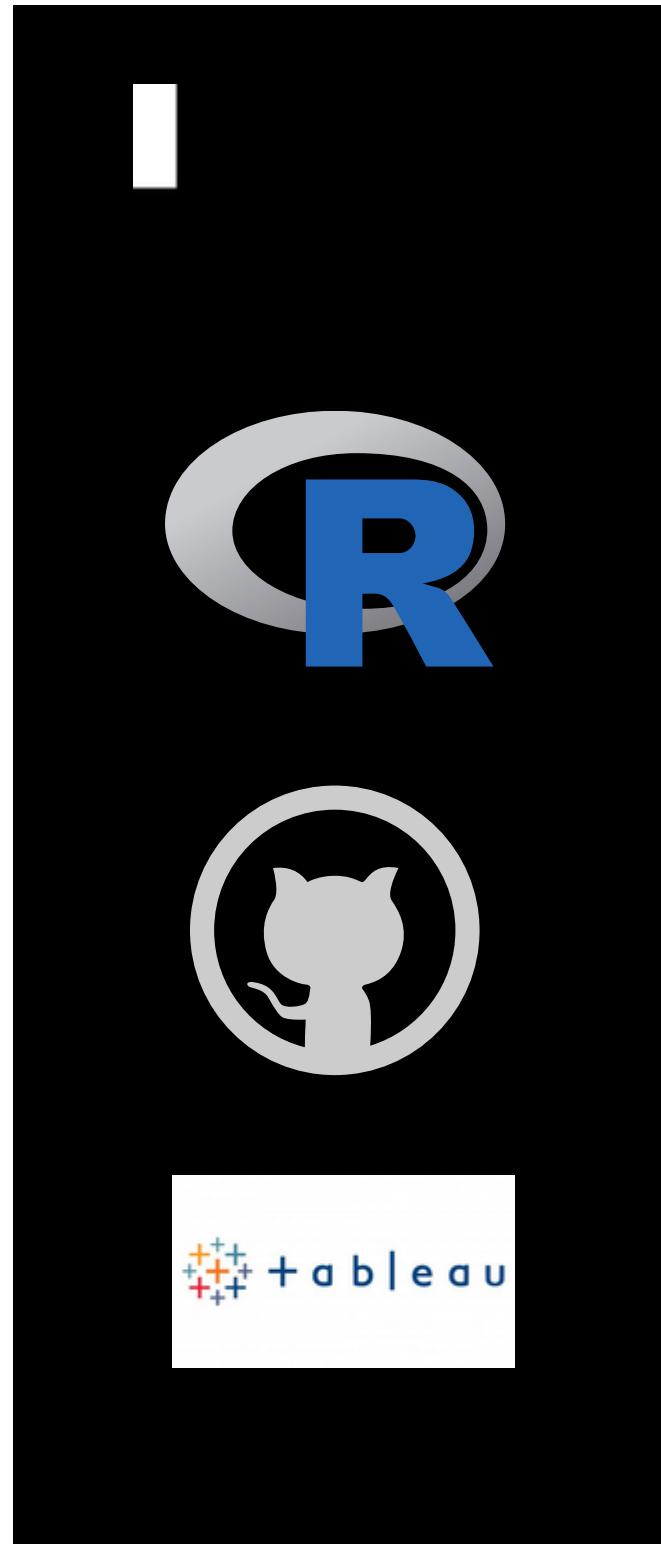
This is not a linear process! You'll likely go back to further gather, clean and explore your data

A semi-transparent background image of a woman with dark, curly hair. She is resting her chin on her right hand, looking thoughtfully upwards and to the side. She is wearing a yellow, textured sweater.

Question?

?

Data Science tools



ផ្សេងៗ python ដែលជា
បកបន្ទាន់



MEET PYTHON

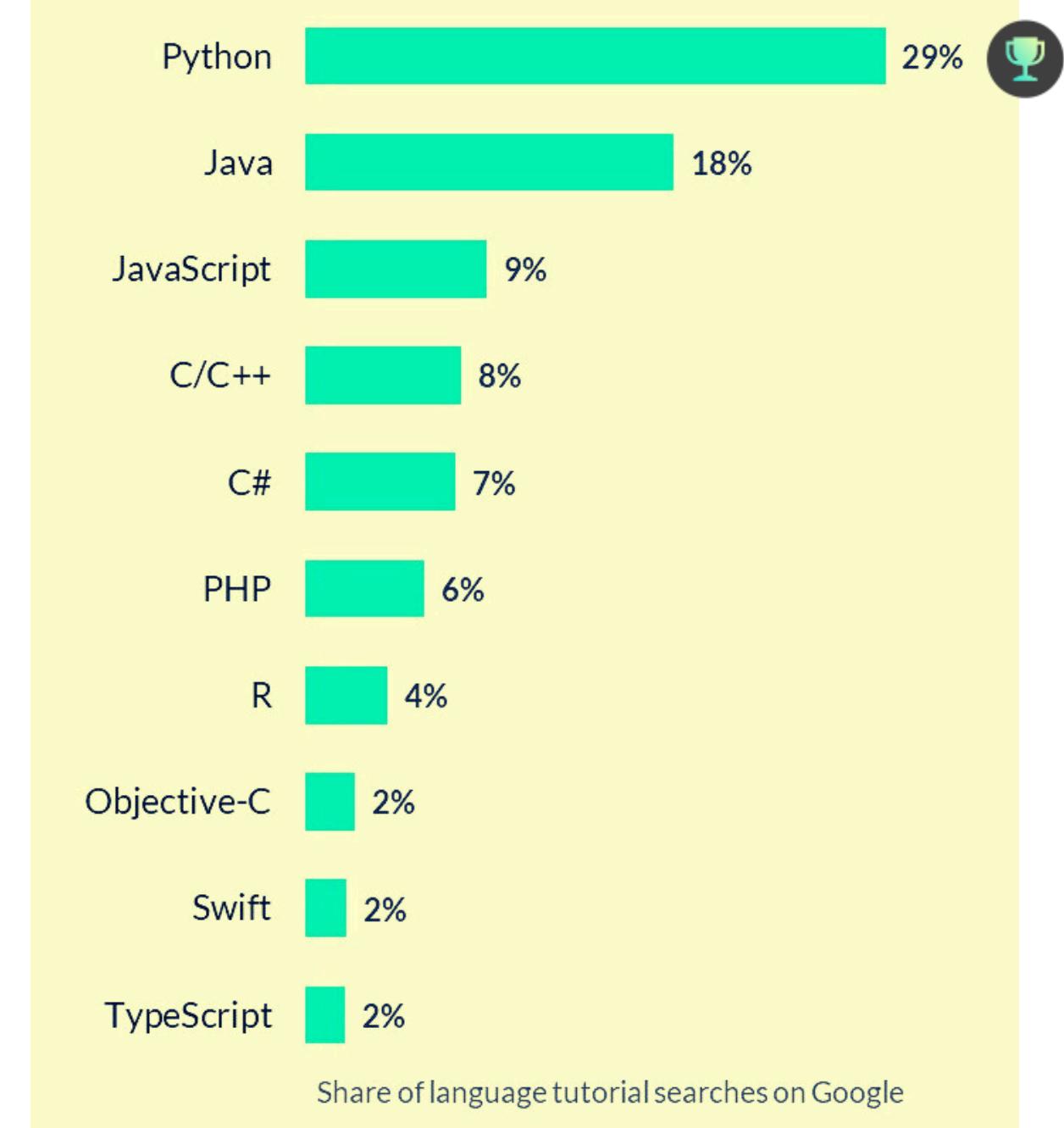


Python is a free, **open-source programming language** that is both powerful and easy to learn.

It has become one of the most popular languages in the world due to its accessibility, flexibility, ease of use, and wide range of applications, including:

- **Data Analytics**
- **Software Development**
- **Machine Learning**
- **Web Scraping**
- **Game Development And more!**

Popularity of Programming Language (PYPL):



Why Does Programming Play Such a Big Role?

-  Data Manipulation and Cleaning
-  Statistical and Mathematical Modeling
-  Building Machine Learning Models
-  Data Visualization
-  Scalability and Automation
-  Collaboration and Reproducibility

Installation (Mac)

1) Go to anaconda.com/products/distribution and click 

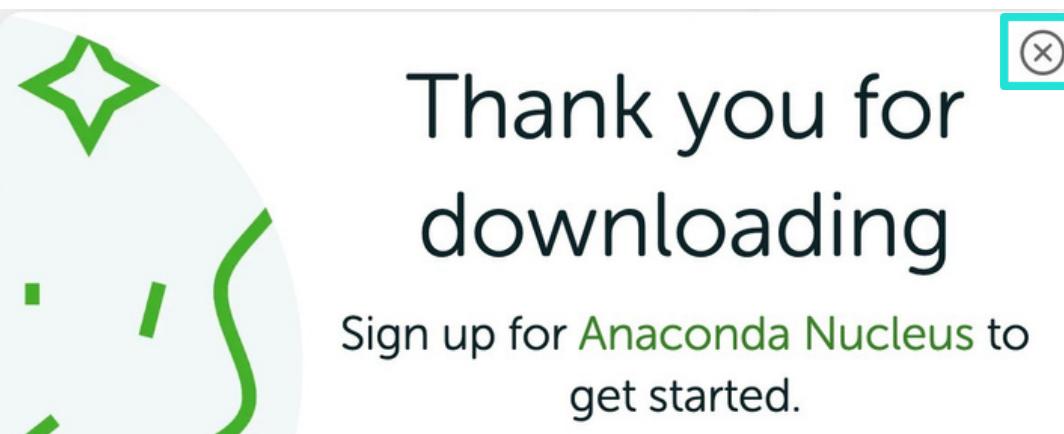
Individual Edition is now

ANACONDA DISTRIBUTION

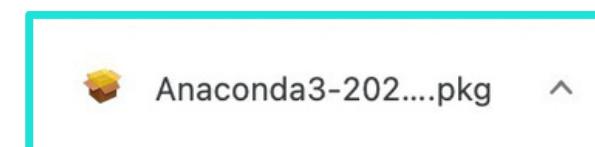
The world's most popular open-source Python distribution platform



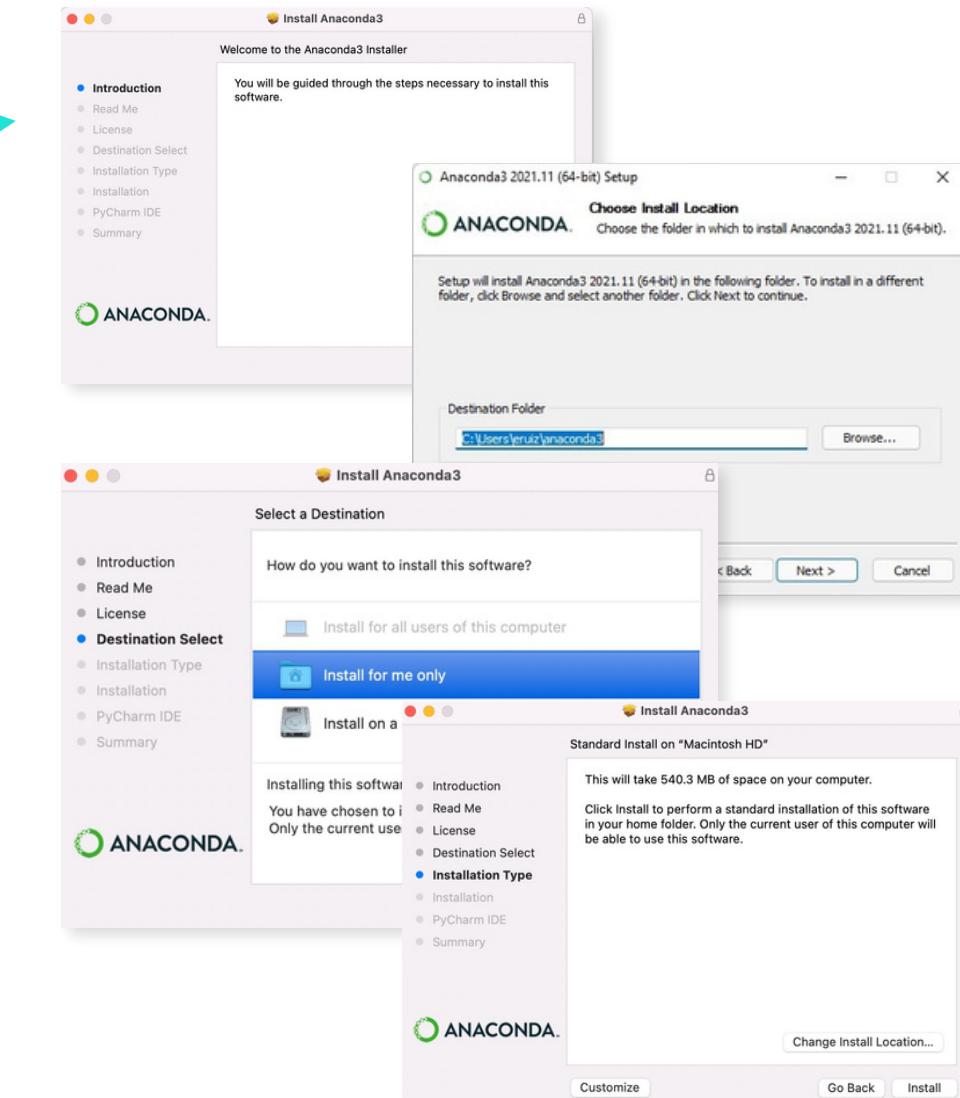
2) Click X on the Anaconda Nucleus pop-up (no need to launch)



3) Launch the downloaded Anaconda **pkg** file



4) Follow the **installation steps** (default settings are OK)



Installation (PC)

1) Go to anaconda.com/products/distribution and click

Download 

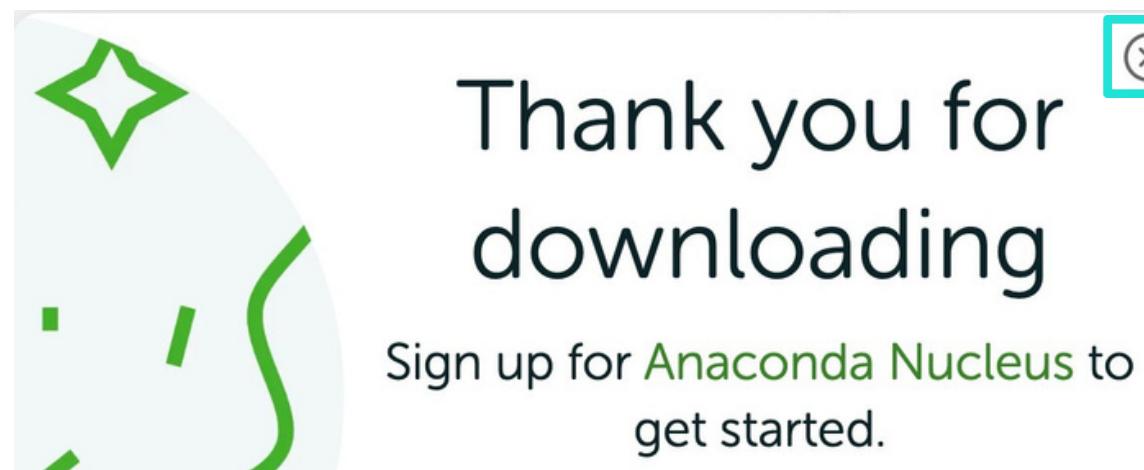
Individual Edition is now

ANACONDA DISTRIBUTION

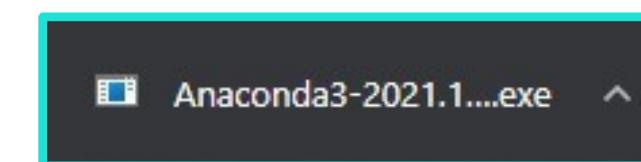
The world's most popular open-source Python distribution platform



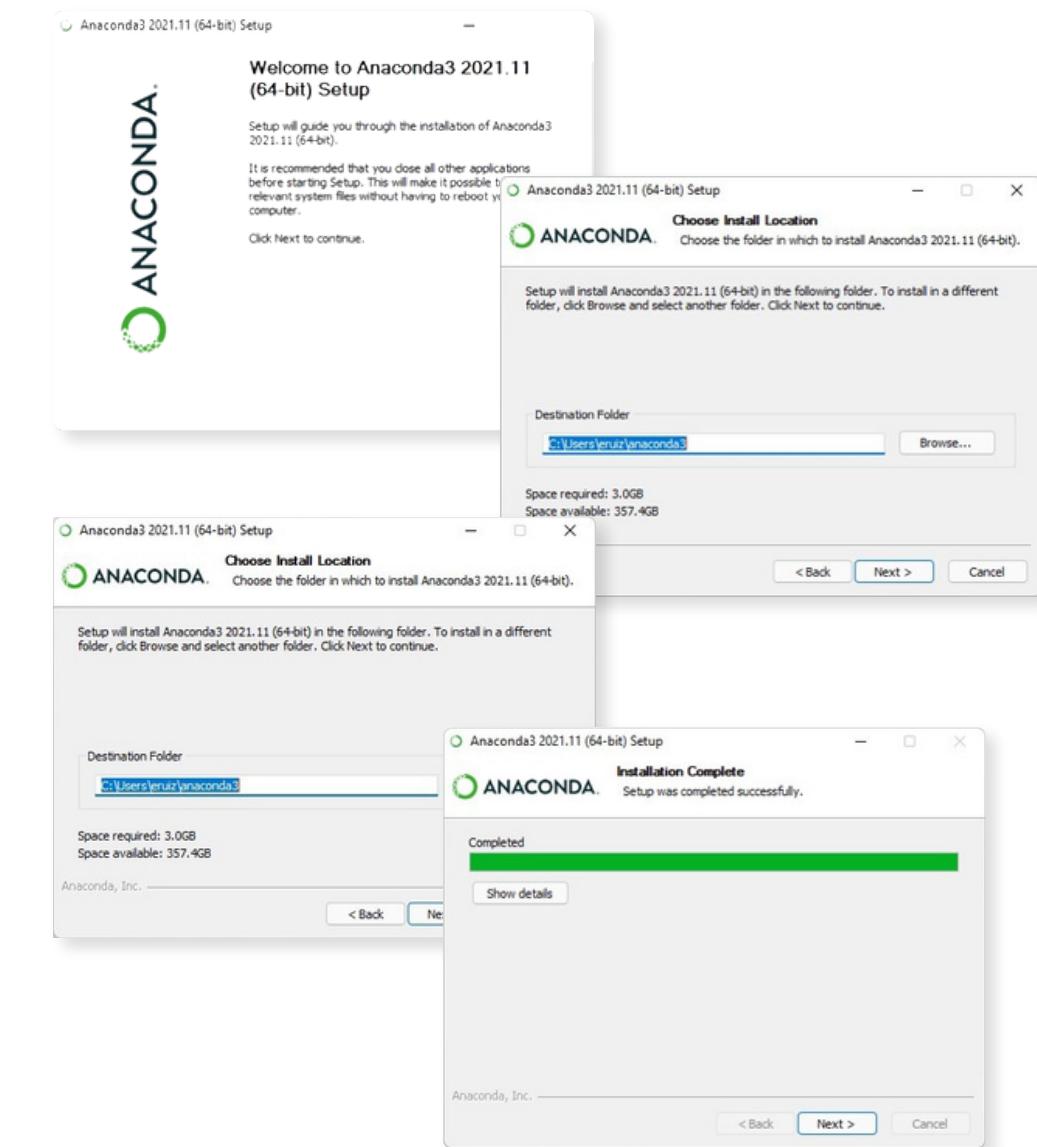
2) Click X on the Anaconda Nucleus pop-up
(no need to launch)



3) Launch the downloaded Anaconda **exe** file



4) Follow the **installation steps**
(default settings are OK)



Why Data Science

Data science is a vital part of every business

You can minor in data science

Job opportunities in all kinds of businesses

Work with diverse teams

Be a part of innovation and progress

Be part of a growing community

Data science jobs are in high demand

Data scientists earn good money

Job security in the field of data science

Use the power of data for social change

Data Science vs. Data Analytics

Data Analytics

What happened?

- Descriptive Analytics
- Data Analysis

Data Science

What's going to happen?

- Predictive Analytics