



WELCOME  
TO

# Giới thiệu



**Cao Tuấn Anh**

**Giảng viên**

anh.cao@codegym.vn



**Nguyễn Hồng Ngọc**

**Giáo vụ**

ngoc.nguyen2@codegym.vn

# Data Science

## Lập trình căn bản, OOP

Biến, kiểu dữ liệu, toán tử.  
Điều kiện, lặp  
Hàm module  
OOP, Tuples, Dictionary, Set



## Phân tích dữ liệu với Python

python căn bản



Dữ liệu, thống kê, phân tích các loại dữ liệu,  
trực quan hoá dữ liệu

File, Data, Database, CSV  
Numpy, Pandas, Matplotlib  
Data Visualization

## Machine Learning xử lý dữ liệu

Mô hình phân loại  
Hồi quy



Machine Learning

## Tools MS Azure ML Studio/Tableau



MS Azure ML Studio  
Tableau

# Khoá học: Python Căn bản

# Nội dung

1. Giới thiệu khóa học python
2. Định hướng học tập
3. Hướng dẫn sử dụng hệ thống
4. Cài đặt Jupyter Notebook
5. Hiểu được khái niệm về Data Science
6. Python với Data Science
7. Chạy ứng dụng đầu tiên với Python

# Khóa học Python

## Mục đích

- Cung cấp cho học viên những kiến thức nền tảng và kỹ năng cần thiết về lập trình Python, xây dựng ứng dụng game, ứng dụng tính toán sử dụng kiến thức lập trình căn bản, tạo nền tảng vững chắc khi chuyển sang khóa học chuyên sâu về data science như phân tích dữ liệu, học máy

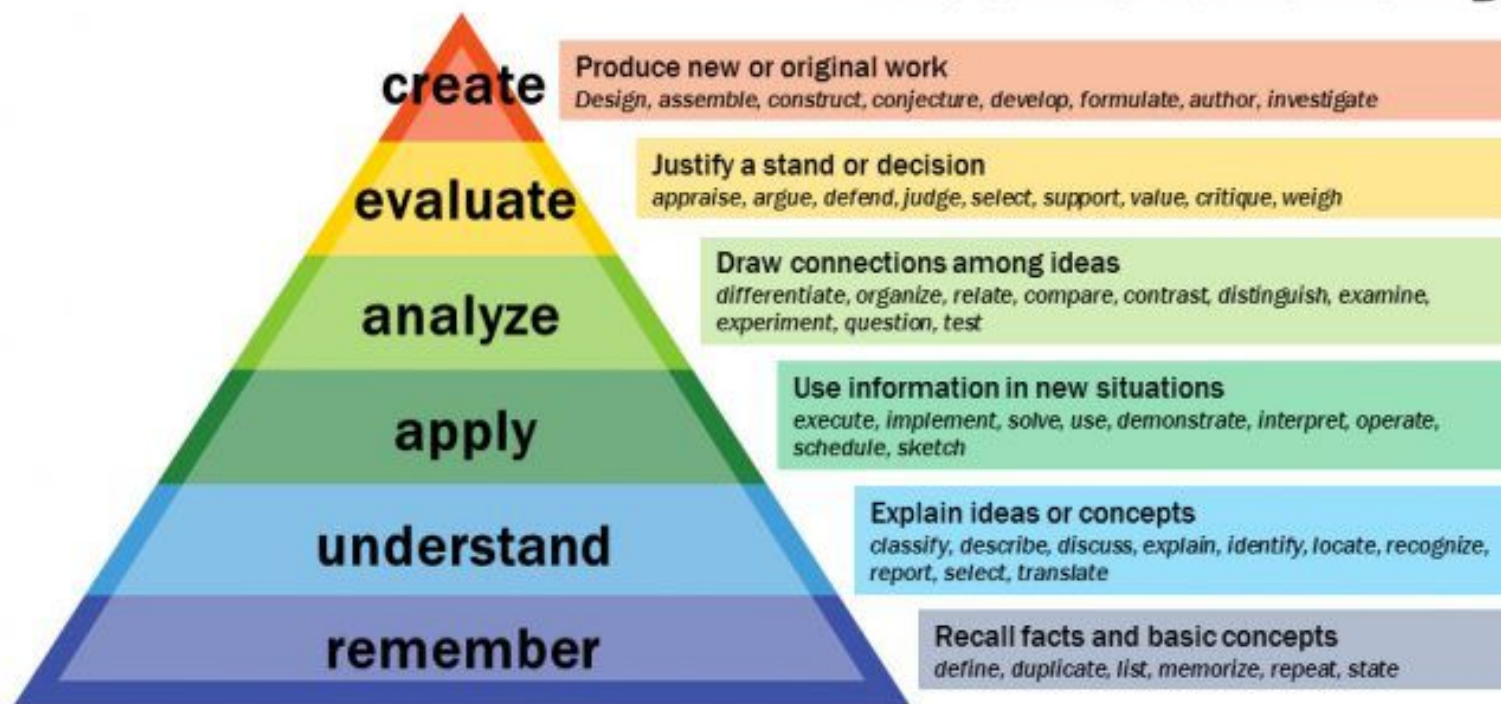
## Các kiến thức trọng tâm

- Biến, kiểu dữ liệu và toán tử
- Cấu trúc điều khiển: Điều kiện, lặp
- Hàm
- OOP
- Thư viện

# Định hướng học tập

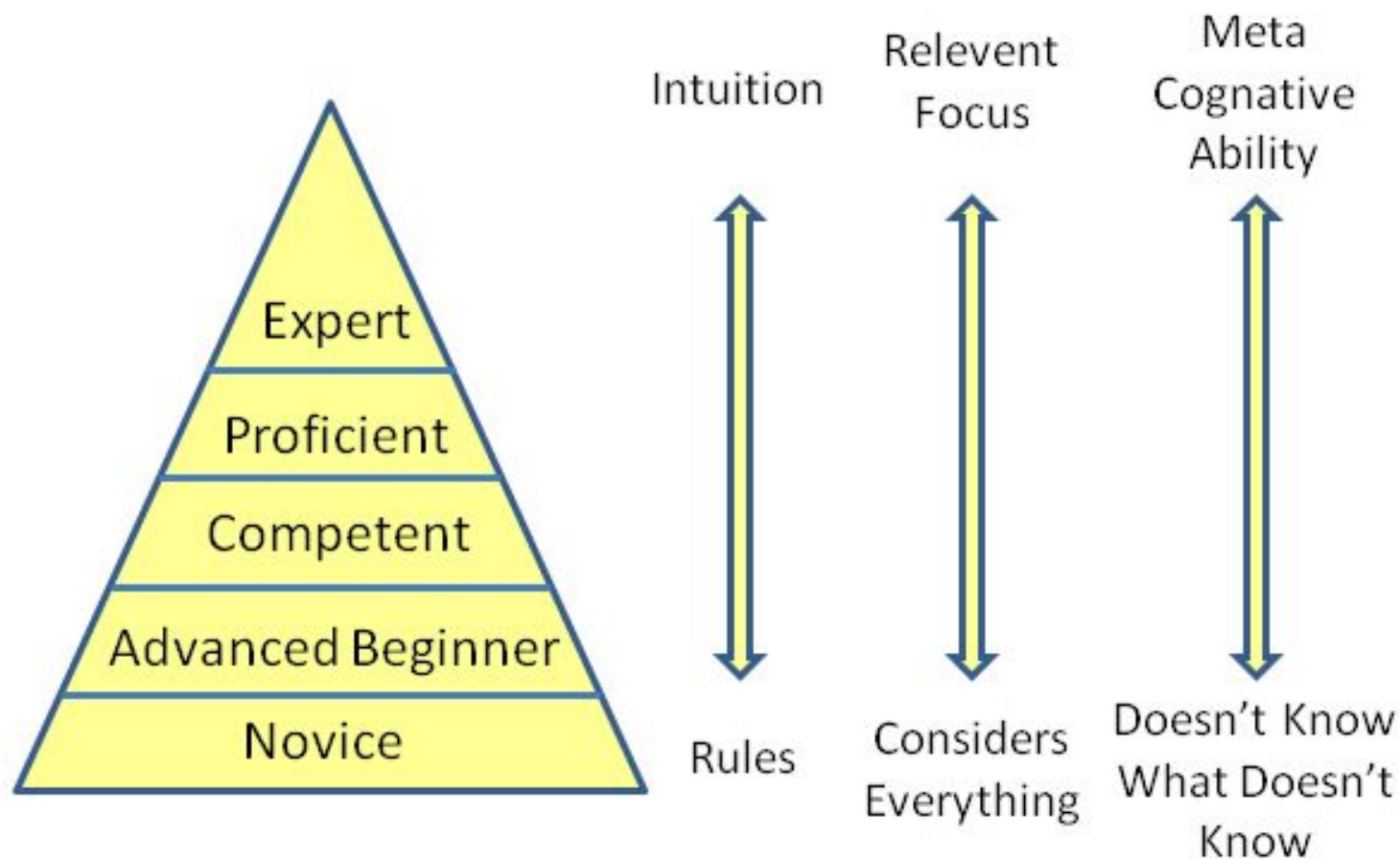
# Phải biết mình ở đâu: Kiến thức

## Bloom's Taxonomy

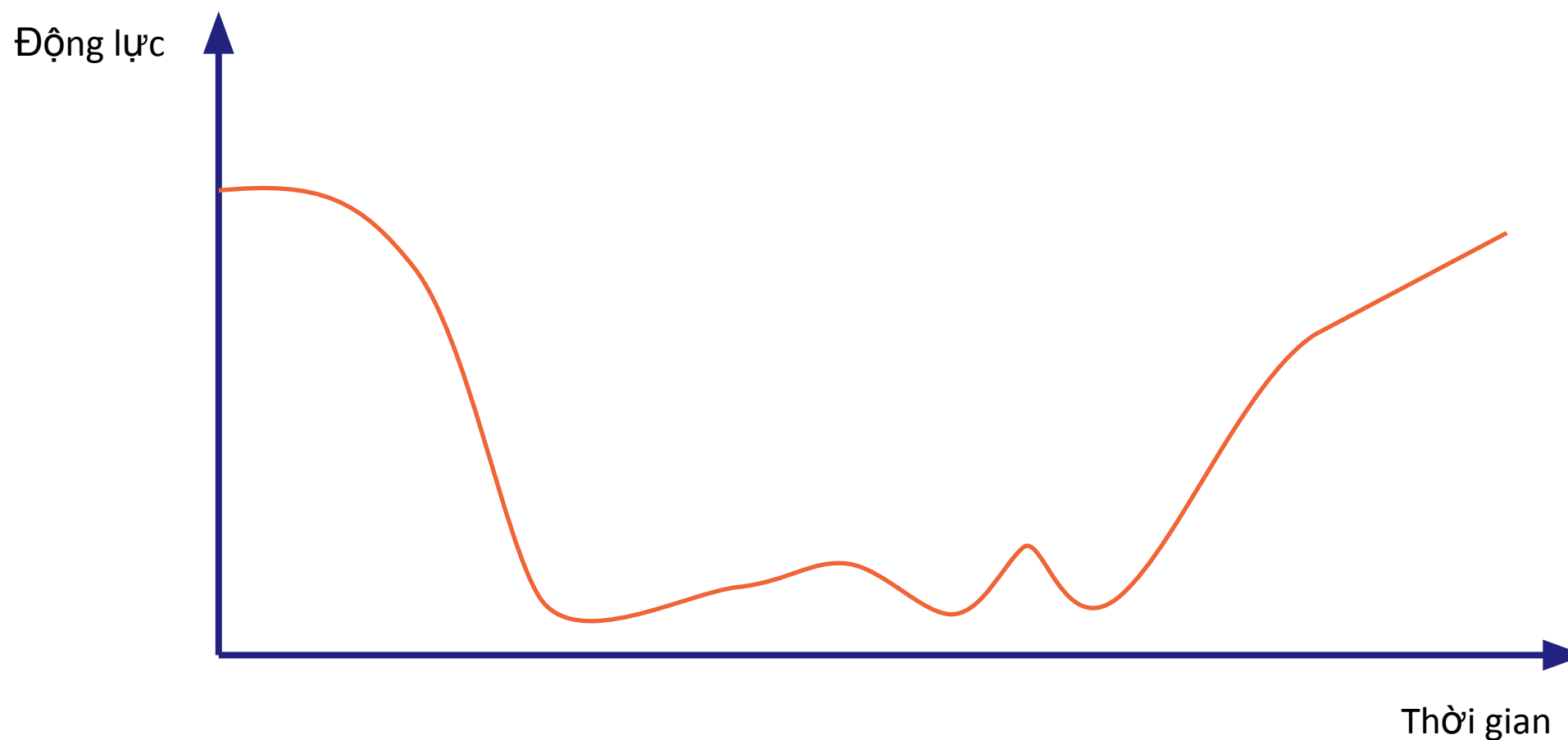




# Phải biết mình ở đâu: Kỹ năng



# Động lực học tập lên xuống theo giai đoạn



# Làm sao để có động lực học tập

- Xác định mục tiêu học tập
- Nhận thức năng lực hiện tại của mình
- Kế hoạch học tập rõ ràng
- Tạo các thắng lợi ngắn hạn
- Nhìn thấy được kết quả học tập theo tuần, ngày
- Tăng cường hỏi, giao tiếp
- Không sợ sai

# Hướng dẫn sử dụng hệ thống

# Hệ thống học tập

- Learn
  - Khóa học Python căn bản trên learn
  - Nhóm lớp
  - Nhóm cộng đồng
- [Codegym.Bob](#): Luyện tập thuật toán

# Hướng dẫn cài đặt Jupyter Notebook

# Jupyter Notebook

Cài Anaconda:

Truy cập vào trang và tải phiên bản phù hợp:

<https://www.anaconda.com/products/individual>

Mở jupyter Notebook trên anaconda

# Khái niệm Data Science



# Data Science

- Data Science là thu thập, khai thác và phân tích các giá trị từ khối dữ liệu khổng lồ theo cấu trúc.

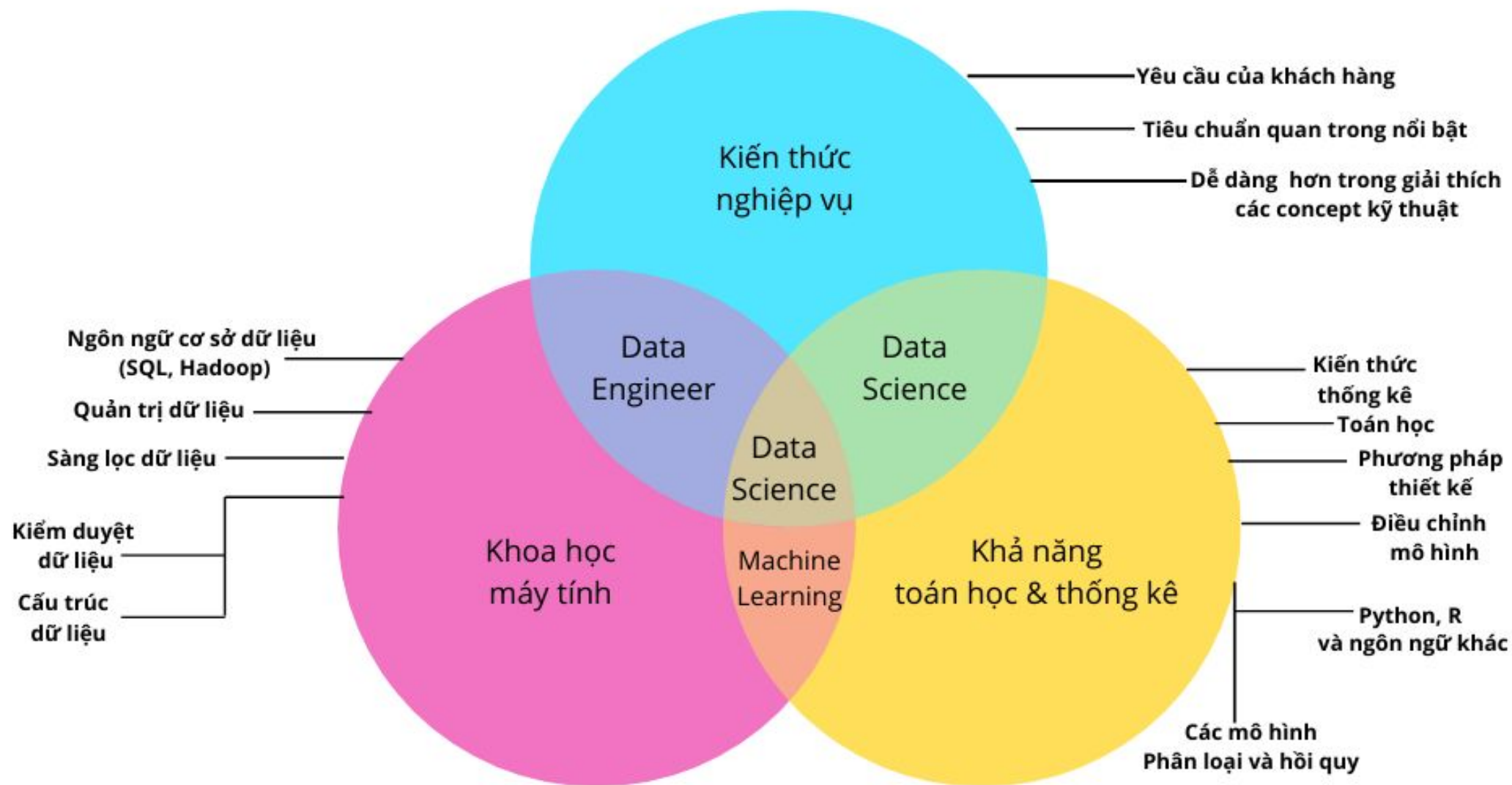
Data Science =

Advanced Mathematics (Statistics, Linear Algebra, Calculus)

+ Programming (Python, Database , DSA)

+ Kiến thức về lĩnh vực kinh tế xã hội cụ thể

- Data Science giúp phân tích, đưa ra các xu hướng, mô hình phát triển và dự báo tương lai.

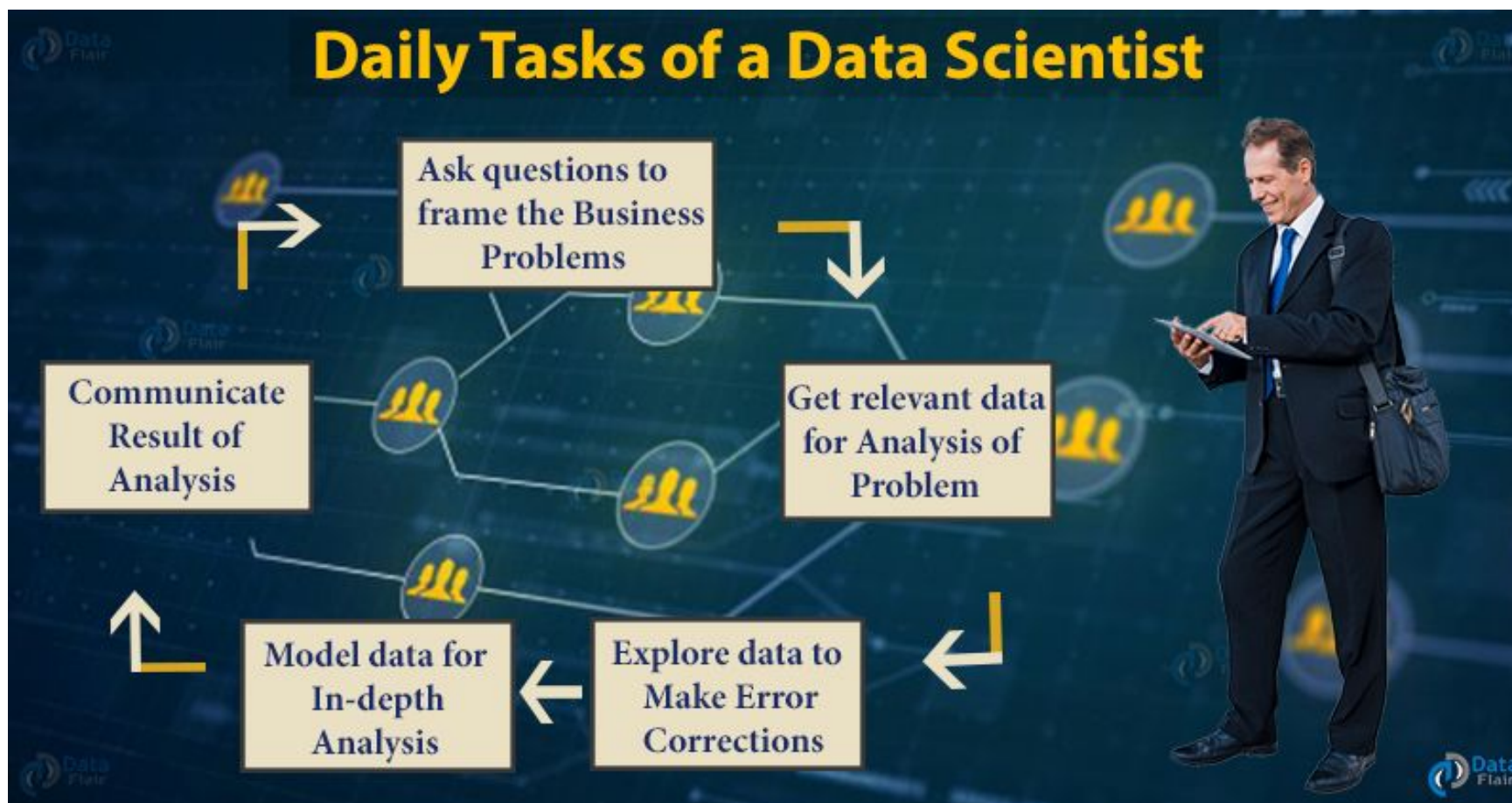


# Data Scientist

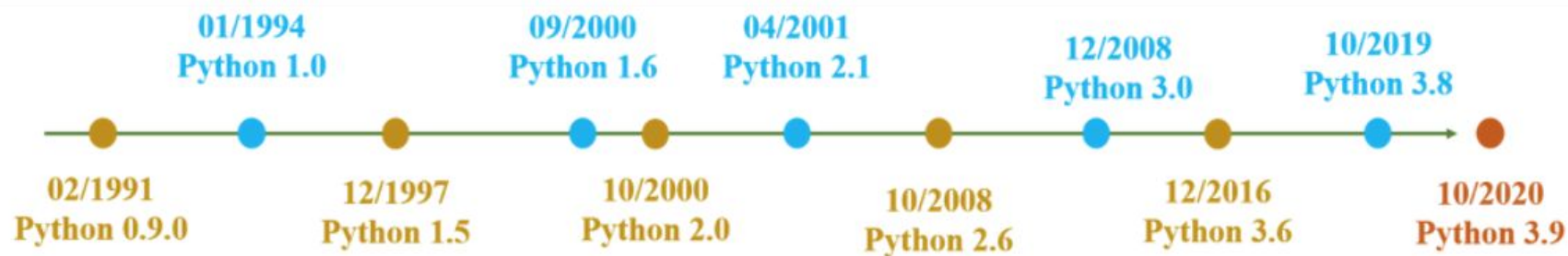
Data Scientist - Nhà khoa học dữ liệu là những người tạo ra giá trị từ dữ liệu với nhiệm vụ chính là:

- Thu thập, xử lý dữ liệu để tìm ra những giá trị ẩn chứa trong nó
- Chuyển hóa những giá trị đó thành các giải pháp/hành động từ đó tạo ra các giá trị và lợi ích thực cho tổ chức.

# Luồng công việc trong data science



# Python trong Data Science



Ý tưởng từ 1980s



Bắt đầu cài đặt  
từ 12/1989



Được đặt tên theo  
nhóm hài Monty Python



Guido van Rossum



228,855 python packages  
(PyPI)

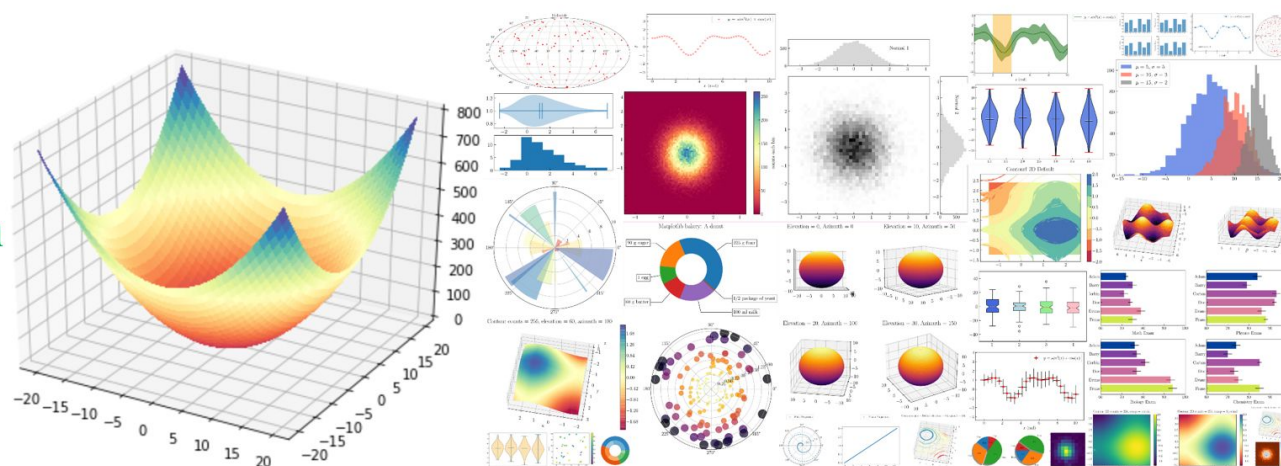


Hỗ trợ rất mạnh cho  
Data Science và Machine Learning

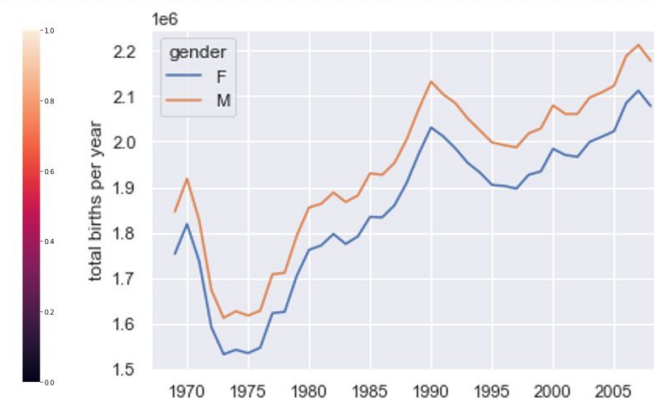


# Thư viện nổi tiếng python

## Visualization

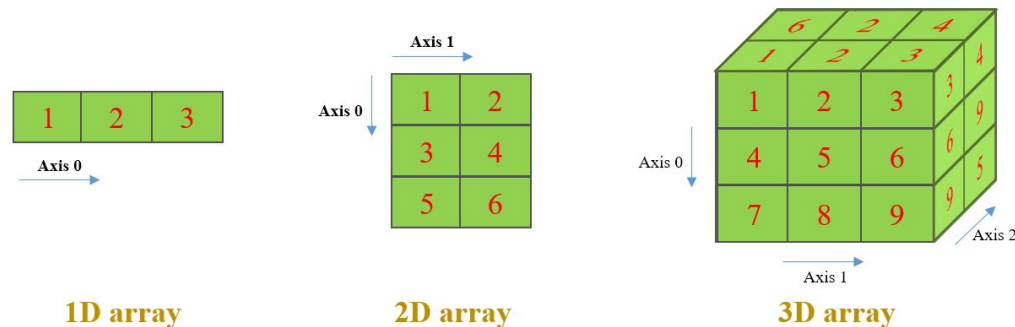


## Pandas

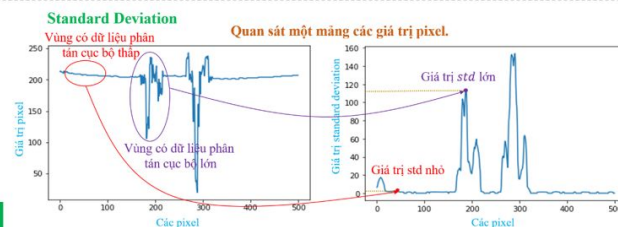


# Thư viện nổi tiếng python

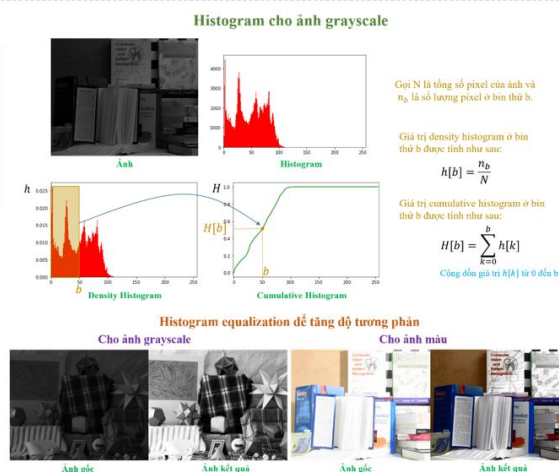
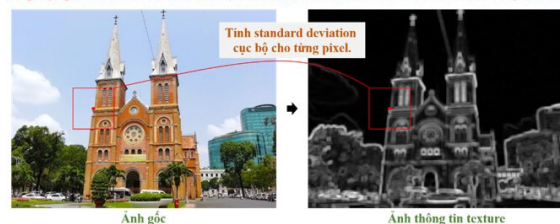
**Numpy**



**Math and applications (1)**



Ứng dụng tính chất của variance (~standard deviation) để tìm texture cho một hình



# Demo

Chạy chương trình Python đầu tiên với Jupyter Notebook



# Tổng kết

Học cách sử dụng hệ thống

Hiểu về Data Science

Cài đặt Jupyter Notebook

Chạy ứng dụng Python đầu tiên