

Introduction

Janis Keuper



Introduction to ML



INSTITUTE FOR MACHINE
LEARNING AND ANALYTICS



Introduction

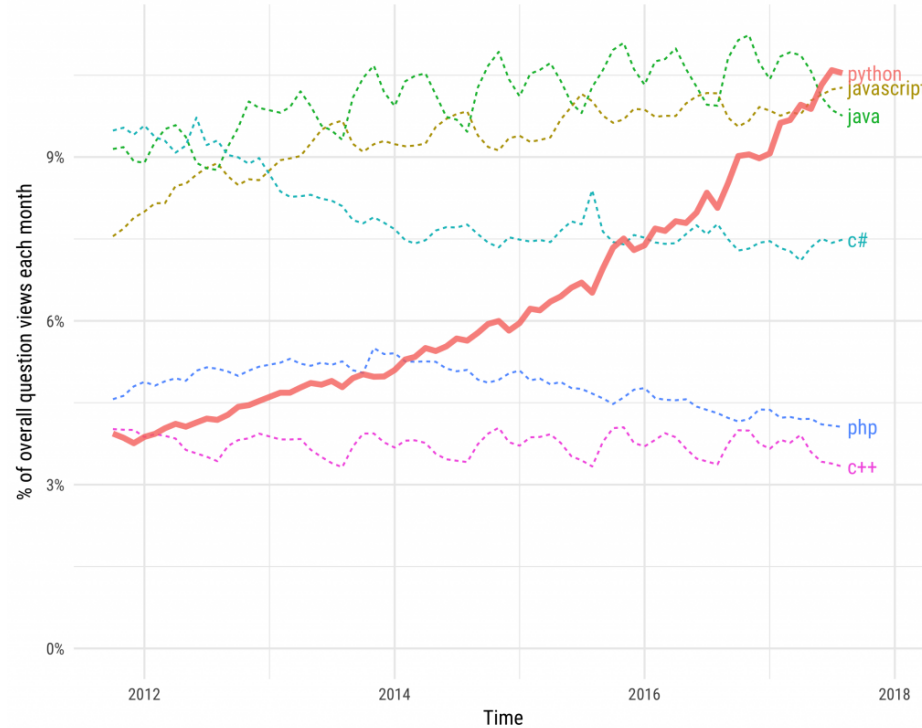


- We will use Python (more exactly tools with python interface) for this course
- Python is the most dominant language used in Data Science, Data Analytics and Machine Learning today
- Alternatives: R and Julia



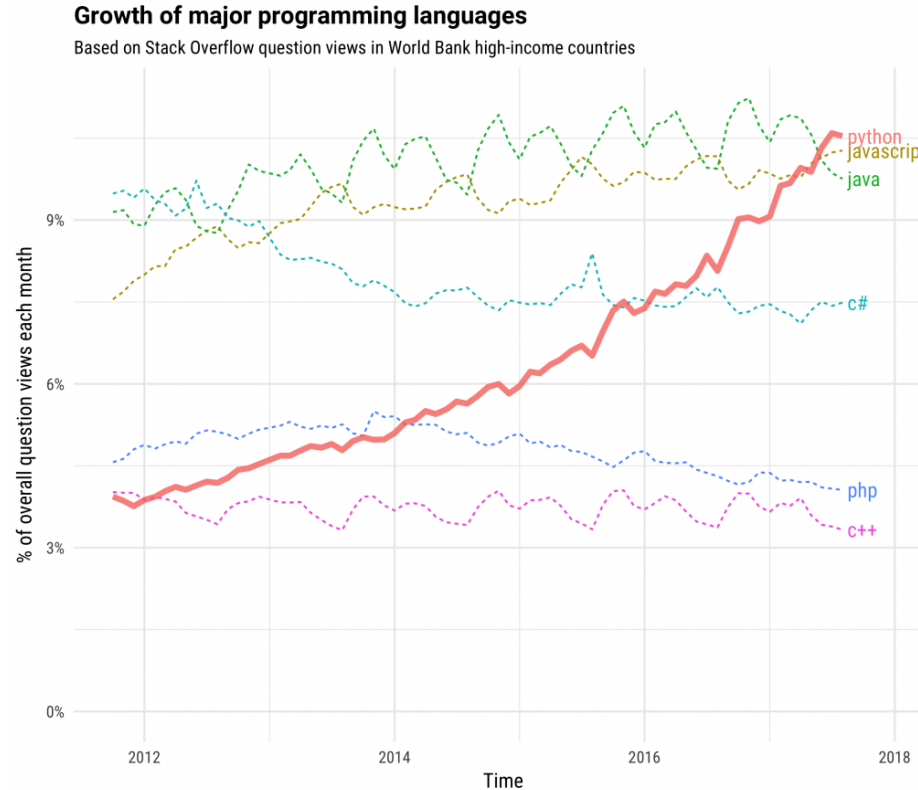
Growth of major programming languages

Based on Stack Overflow question views in World Bank high-income countries



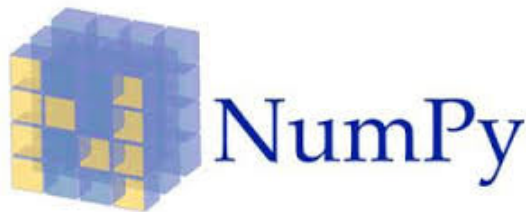
[2]

- Triggered by ML + Data Science
- Python pushes convergence of technologies:
 - HPC
 - Cloud
 - Big-Data



[2]

- Very low entry barrier (easy to learn)
- Quite universal in programming approaches
- Easy to interface existing high performance libs
- Huge Community
- **Libraries !!!**
- **Python as „glue code“ for rapid development and markup**





<https://www.anaconda.com/distribution/>

- Easy installation on Linux/Mac/Win
 - See read.me in Moodle
 - User space install (no admin needed)
- Conda Python package management
 - Simple install/update/sharing of software
- All open source!



NOTE:

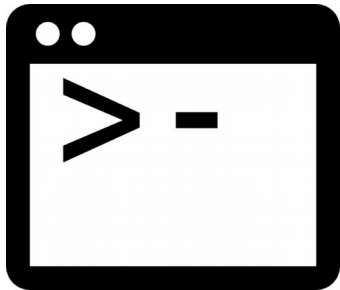
There are two versions of Python:

Python2 and Python3

Support for Python2 will end 31.12.19

→ get started with Python3!





Interactive Python Introduction

→ Start lab session now!