

Data Science School

Programing

박두진

데사스 프로그래밍 과정의 목표

데이터 분석에 필요한 프로그래밍 실력 쌓기 !!!

데이터 분석을 위해 저와 배우는 것들

- 데이터 분석을 위한 파이썬 배우기
- 데이터 저장 및 가공을 위한 **SQL** 배우기
- 데이터 수집을 위한 웹 서비스 배우기
- 클라우드 서버를 활용하여 데이터 수집하기

데이터 분석을 위해 저와 배우는 것들

Day 1 - 5 (20h)

- 파이썬 문법과 데이터 분석을 위한 기본 패키지
- Numpy, Pandas

Day 6 - 11 (24h)

- 서버의 사용방법과 데이터 베이스 구축 및 활용
- AWS EC2, Linux, Mysql, Mongodb, SQLAlchemy

Day 12 - 17 (24h)

- 웹에서의 데이터 수집
- HTML, CSS-Selector, Xpath, requests, selenium, scrapy

Day 18 - 19 (8h)

- 서버를 이용한 웹 데이터 크롤링 자동화
- Terraform, AWS Lambda, AWS Cloud Watch

과정을 잘 배우고 나면...

AWS 클라우드 서버를 이용하여...

수집한 데이터를 DB에 저장하여...

DB에 저장된 데이터를 파이썬을 이용하여 분석하고...

분석한 데이터를 시각화 하고

AI 모델을 만들어 서비스화 할 수 있습니다.

왜 파이썬인가?

- 프로그램 언어를 배우고 싶은데 어떤 언어를 배우면 될까요?
 - C, C++, C#, Java, Javascript, Python, Ruby, C#, Go, Rust, Scala, Perl, Obj-C, PHP, R
- 외국어를 배우고 싶은데 어떤 언어를 배우면 될까요?
- 언어는 커뮤니케이션의 도구
- 컴퓨터와 커뮤니케이션을 하면 컴퓨터에게 일을 시킬수 있다.
- 중국에 가서 중국사람과 일하고 싶은데 어떤 언어를 배우면 될까요?
- 컴퓨터로 데이터분석을 하고 싶은데 어떤 프로그래밍 언어를 배우면 될까요?
 - 파이썬을 배우세요. 왜?

왜 파이썬인가?

- 데이터 분석을 위한 빠르고 강력한 라이브러리를 가지고 있다.
 - Numpy, Pandas, Matplotlib, scikit-learn, statsmodels
- 다른언어에 비해 문법이 쉽고 간결하다.
- 파이썬을 배우면 할수 있는것들이 많다.
 - 데이터 분석, 웹서비스 개발, 해킹, GUI 프로그래밍, 타 언어와 결합등등
 - <https://github.com/vinta/awesome-python>

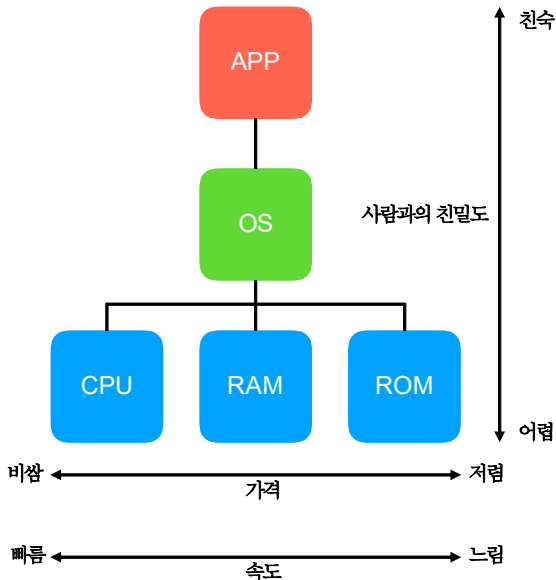
파이썬은 어떤 언어인가?

- Python is an interpreted high-level programming language for general-purpose programming. ([wiki](#))
- 파이썬의 특징
 - 플랫폼에 독립적
 - 인터프리터 언어
 - 객체지향적
 - 동적타이핑

-> 컴퓨터에 대해 먼저 알아보자 !

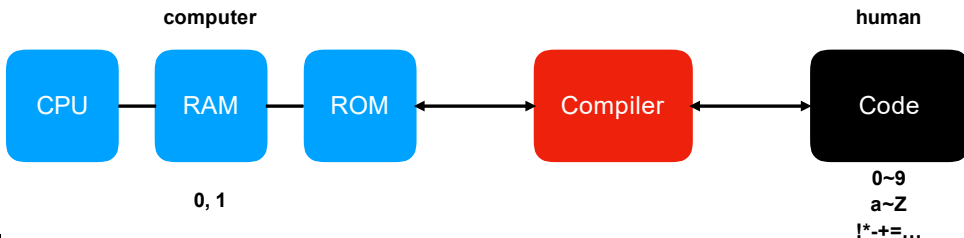
컴퓨터를 이해해 보자!

- 계산을 수행하는 기계
- 컴퓨터를 구성하는 기본 요소
 - CPU (Central Processing Unit)
 - RAM (Random Access Memory)
 - ROM (Read-Only Memory)
- 운영체제 (OS : Operating System)
 - Kernel
 - 하드웨어를 컨트롤하는 소프트웨어 (운영체제의 핵심)
 - CPU, RAM, ROM 자원을 사용하는 방법을 정의
- 어플리케이션 (App : Application)
 - OS 기반 응용 프로그램
 - 대부분의 프로그래밍의 영역



컴퓨터에서 프로그램의 동작원리는?

- CPU, RAM, ROM은 0과 1밖에 모름
- 프로그램 언어는 숫자와 알파벳과 특수기호를 사용
- 각 언어의 **Compiler**가 코드를 0과 1의 이진수로 변환
- 컴파일러 언어
 - 모든 코드를 컴파일링 후에 컴퓨터에서 처리 -> 처리속도가 빠르지만 프로그램 실행을 위해 컴파일링 시간을 기다려야함
- 인터프리터 언어
 - 한줄씩 코드를 컴파일링 하면서 컴퓨터에서 처리 -> 처리속도가 느리지만 컴파일링 시간 없이 바로바로 프로그램 실행
- 코드를 잘 만들면 컴퓨터가 효율적으로 일할수 있다!



파이썬은 어떤 언어인가?

- 플랫폼에 독립적
 - 어떠한 종류의 OS에도 같은 문법을 사용
- 인터프리터 언어
 - 한줄씩 컴파일링 하면서 코드를 수행
 - 처리속도가 느린데 감히 데이터 분석을?
- 객체지향적
 - 실제세계를 모델링하여 공통적인 기능을 묶어서 개발하는 방식
 - 추상화, 캡슐화, 상속, 다형성의 특징을 가짐
 - 반대의 개념은 절차지향
- 동적타이핑
 - 변수 선언시 데이터 타입을 지정해 주지 않아도 데이터에 따라서 자동으로 타이핑

파이썬의 종류는?

- Cpython
 - C로 만들어진 파이썬
 - 우리가 코딩하는 부분은 인터프리터지만 안에서는 다 컴파일러언어인 C로 동작되어 속도가 빠름 -> 데이터 분석 가능
- Jython
 - Java로 만들어진 파이썬
- IronPython
 - C#으로 만들어진 파이썬
- Pypy
 - Python 으로 만들어진 파이썬
 - Cpython 보다 빠르게 수행되는것을 목표로 함 (<http://pypy.org/>)

프로그래밍을 한다는것은?

컴퓨터와의 효율적인 커뮤니케이션으로

자신이 생각하는 목적을 컴퓨터가 잘 처리할수 있도록 하는것!

프로그래밍을 배운다는것은?

프로그래밍 언어는 언어다.

나이가 많다고 해서 말을 잘 하지는 않는다.

프로그래밍을 오래 했다고 해서 프로그래밍을 잘하지 않는다.

많은 관심과 애정 그리고 연습만이 프로그래밍을 잘하는 방법이다.