

정치학과 파이썬

엄기홍(경북대학교)

2018년 10월 5일 건국대학교 특강

개요

- 파이썬 소개
- 파이썬 실습
 - 데이터 수집
 - 데이터 분석
- Q&A

정치학과 프로그래밍

- 오늘날의 정치학은 경험적 · 실시간 분석이 필요한 시대
- 소규모 데이터(여론조사, 집합자료 등)를 위한 분석도구
- 빅데이터(SNS, 회의록 등)를 위한 분석 도구
 - 컴퓨터 프로그래밍, 특히 파이썬

컴퓨터 프로그래밍이란?

다른 언어를 사용하는 외국인과 영어로 소통하듯이
컴퓨터가 알아들을 수 있는 **언어**를 통해
의사소통 하는 행위

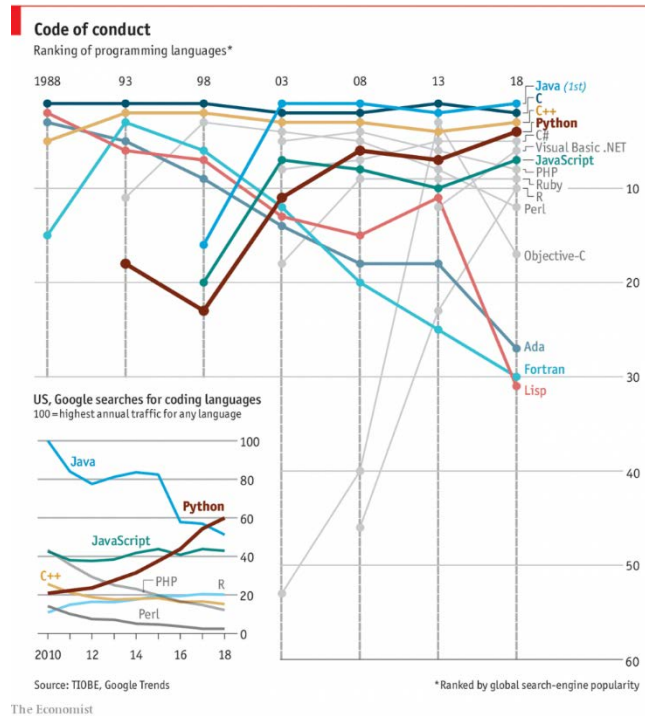


파이썬은?

Dec 2017	Dec 2016	Change	Programming Language	Ratings	Change
1	1		Java	13.268%	-4.59%
2	2		C	10.158%	+1.43%
3	3		C++	4.717%	-0.62%
4	4		Python	3.777%	-0.46%
5	6	▲	C#	2.822%	-0.35%
6	8	▲	JavaScript	2.474%	-0.39%
7	5	▼	Visual Basic .NET	2.471%	-0.83%
8	17	▲	R	1.906%	+0.08%
9	7	▼	PHP	1.590%	-1.33%
10	18	▲	MATLAB	1.569%	-0.25%
11	13	▲	Swift	1.566%	-0.57%
12	11	▼	Objective-C	1.497%	-0.83%
13	9	▼	Assembly language	1.471%	-1.07%
14	10	▼	Perl	1.437%	-0.90%
15	12	▼	Ruby	1.424%	-0.72%
16	15	▼	Delphi/Object Pascal	1.395%	-0.55%
17	16	▼	Go	1.387%	-0.55%
18	25	▲	Scratch	1.374%	+0.19%
19	20	▲	PL/SQL	1.368%	-0.13%
20	14	▼	Visual Basic	1.347%	-0.62%

출처: <https://www.tiobe.com/tiobe-index/>(검색일: 2017년 12월 28일)

“Python is becoming the world’s most popular coding language.” by The Economist



출처: <https://amp.economist.com/graphic-detail/2018/07/26/python-is-becoming-the-worlds-most-popular-coding-language/> (검색일: 2018년 10월 3일)

파이썬은?

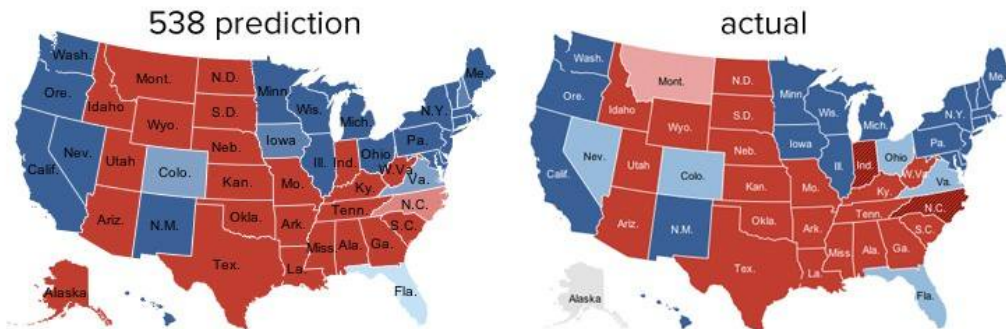


- 객체 지향형 프로그래밍 언어
- 직관적이고 아주 쉬운 문법
- 플랫폼 독립
- 인터프리터 언어
- 오픈소스(Open Source)
- 다양한 라이브러리 & 커뮤니티 활성화

파이썬의 창시자 "귀도 반 로섬"

출처: https://ko.wikipedia.org/wiki/귀도_반_로섬

정치와 프로그래밍 - 미국 대선



Nate Silver 선거예측

출처: <https://mashable.com/2012/11/07/nate-silver-wins/#Tad57oJTauqK>



오바마 재선 - 마이크로 타겟팅

데이터 유형

Structured data	vs.	Unstructured data	vs.	Semi-structured data
<ul style="list-style-type: none">• 열과 행으로 정형화된 데이터• 관계형 데이터		<ul style="list-style-type: none">• 텍스트 자료• 이미지, 오디오, 비디오와 같은 다중매체 자료		<ul style="list-style-type: none">• XML• HTML• JSON

연구과정

연구주제, 기존 문헌 검토, 이론 및 가설

데이터 수집

- 웹 스크래핑
- API

데이터 분석

- 워드클라우드, 의미망분석
- Classification, Clustering, Topic Modeling

결론 및 함의

Jupyter를 활용한 실습

- Off line 실습
 - 경북대 사회과학기초자료연구소에서 다운로드 (<http://www.basicdata.kr/seminar/>)
 - Off line 실습용 파일
- On line 실습
 - 크롬을 열고, 다음 주소를 입력

Jupyter를 활용한 실습(계속)

- <http://155.230.193.142:8000>

Username	Password
student02	student02
student03	student03
.....
.....
student25	student25