

# 파이썬과 데이터 분석

성공회대학교

IT융합자율학부 인공지능전공

홍성준 ([sjhong@skhu.ac.kr](mailto:sjhong@skhu.ac.kr))

## 데이터 분석과 파이썬

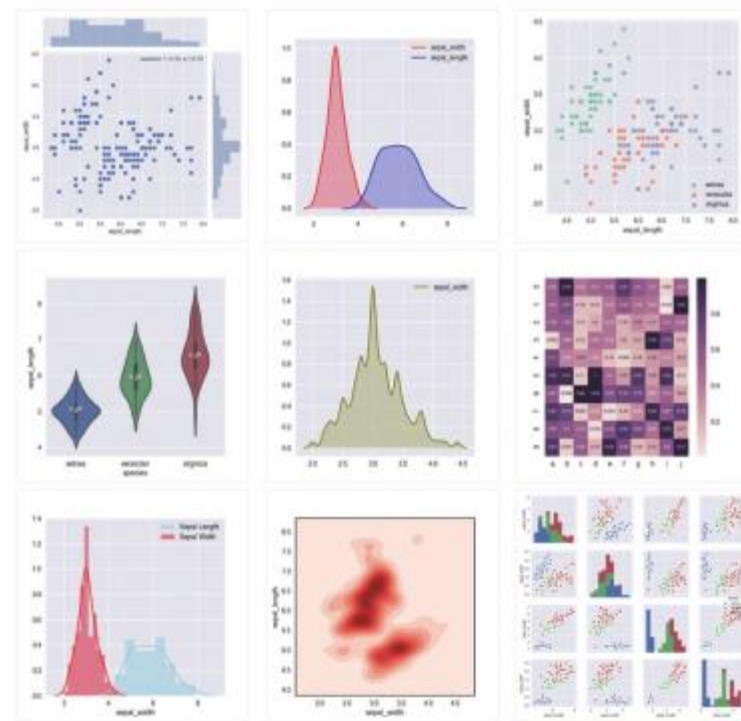
- 파이썬은 데이터를 분석하는데 많이 사용되는 프로그래밍 언어
- 기업, 학계, 언론 등 다양한 분야에서 데이터 분석가들이 파이썬을 사용

# 파이썬의 특징

- 오픈 소스 프로젝트
  - 네덜란드 프로그래머 귀도 반 로섬<sup>Guido van Rossum</sup>이 개발
  - 무료로 사용할 수 있는 오픈 소스로 소스 코드를 공개하고, 파이썬으로 응용 소프트웨어를 개발하거나 상업화 할 수 있도록 허용
- 데이터 분석의 대중화
  - 데이터 분석 기술의 장벽을 낮추는 역할
  - 오픈 소스 생태계 안에서 새로운 분석 기법이 패키지 형태로 배포
  - 작은 기업이나 개인 사용자도 데이터 분석 기술을 손쉽게 활용케 함

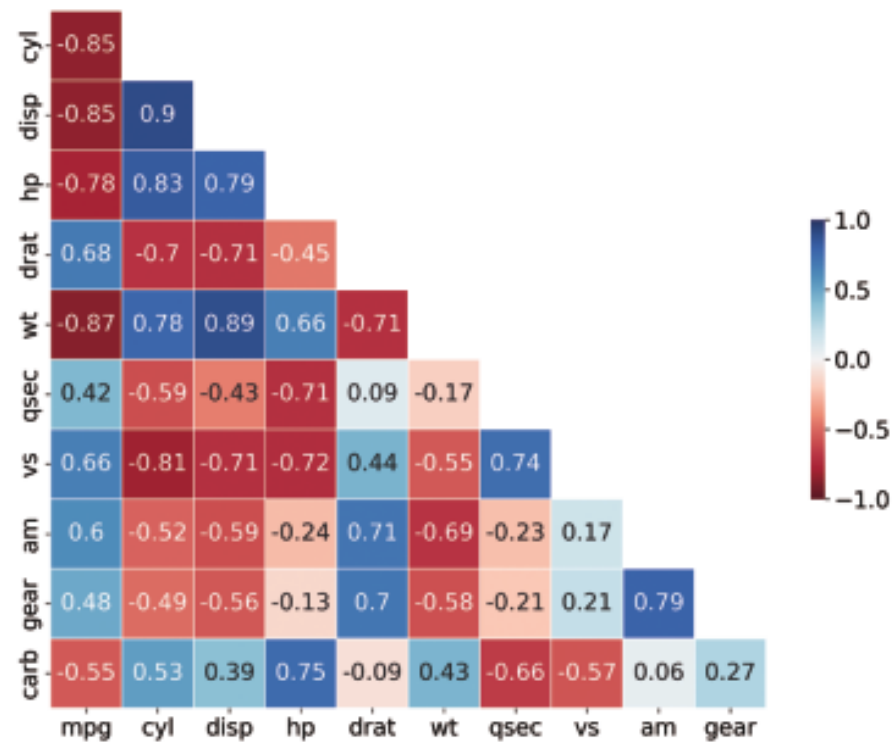
# 파이썬의 특징

- 다양한 교육 콘텐츠
  - 사용자가 많은 만큼, 책, 온라인 강의, 온라인 문서 등 파이썬을 다룬 다양한 교육 콘텐츠가 개발
- 손쉬운 시각화 패키지가 제공
  - 데이터 분석에서부터 그래프를 만드는 일까지 하나의 도구로 손쉽게 효율적으로 작업



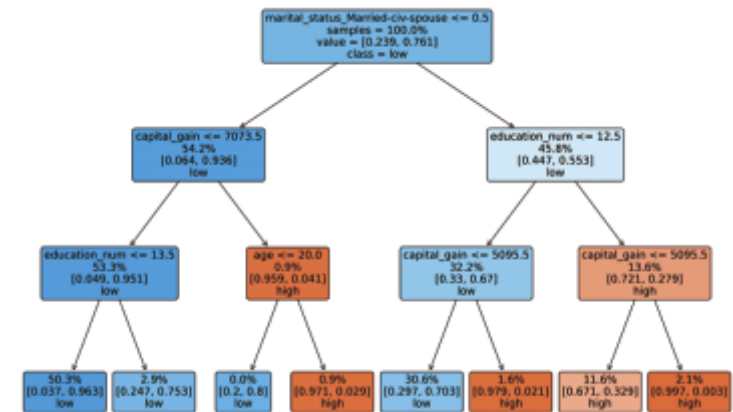
## 파이썬으로 할 수 있는 작업

- 통계 분석
  - 데이터의 특징을 살펴보는 기초 통계 분석부터  
가설 검정 등 고급 통계 분석에 이르기까지  
다양한 통계 분석 기법에 활용



# 파이썬으로 할 수 있는 작업

- 머신 러닝 machine learning
  - 많은 데이터를 이용해 특정 변수를 예측하는 예측 모델을 만드는 기법
  - Decision Tree, Random Forest, Neural Networks 등 최신 머신러닝 알고리즘을 쉽게 구현할 수 있음



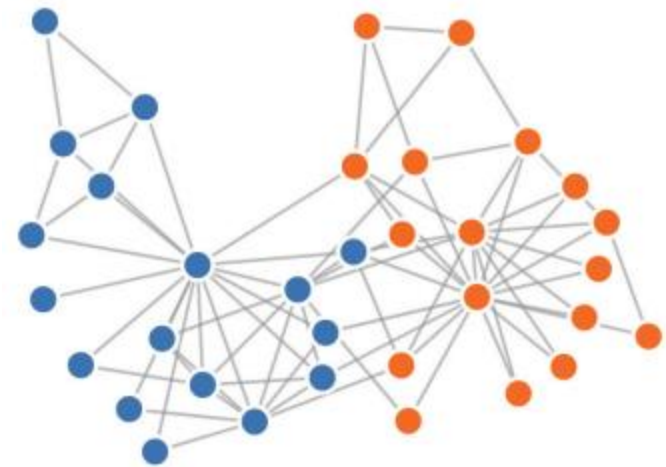
## 파이썬으로 할 수 있는 작업

- 텍스트 마이닝<sup>text mining</sup>
  - 문장에 자주 등장하는 단어를 찾거나, 어떤 감정을 표현하는 단어가 자주 사용되는지 등 문자로 이루어진 데이터를 분석



## 파이썬으로 할 수 있는 작업

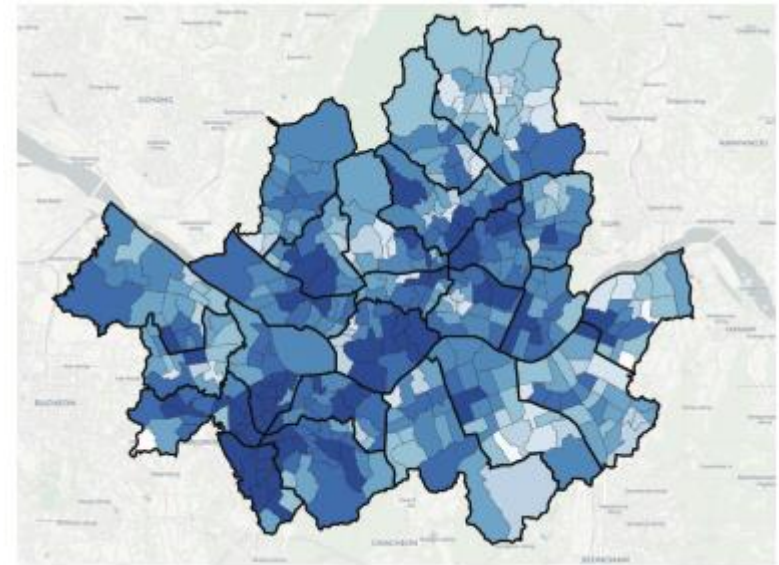
- 소셜 네트워크 분석
  - 트위터, 페이스북과 같은 소셜 네트워크 서비스에서 사람들이 어떤 관계를 형성하고 있는지, 어떤 경로로 메시지가 퍼져 나가는지 등을 분석





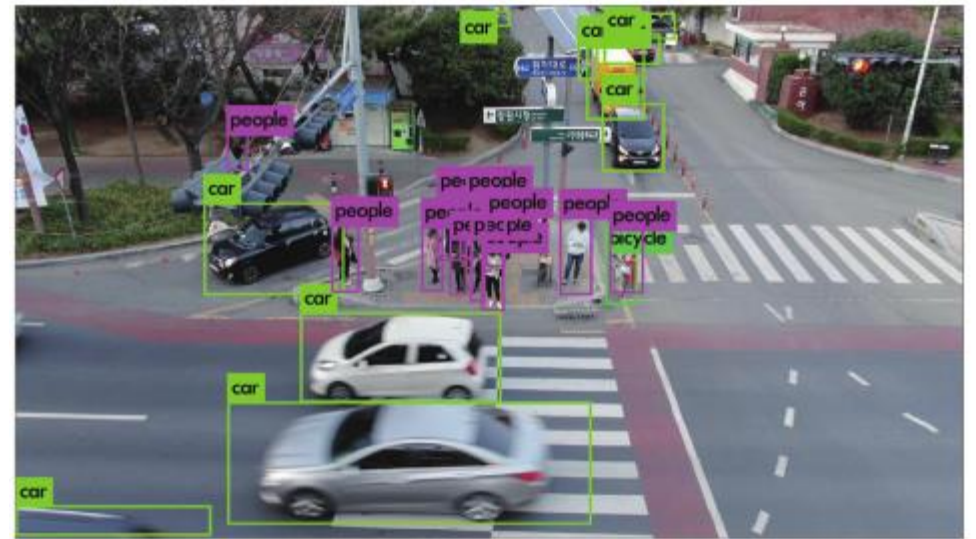
## 파이썬으로 할 수 있는 작업

- 지도 시각화
  - 위도, 경도 등의 지리 정보와 지역 통계를 활용하여 국가별 GDP, 시군구 별 범죄율 등 지역별 특징을 지도로 표현



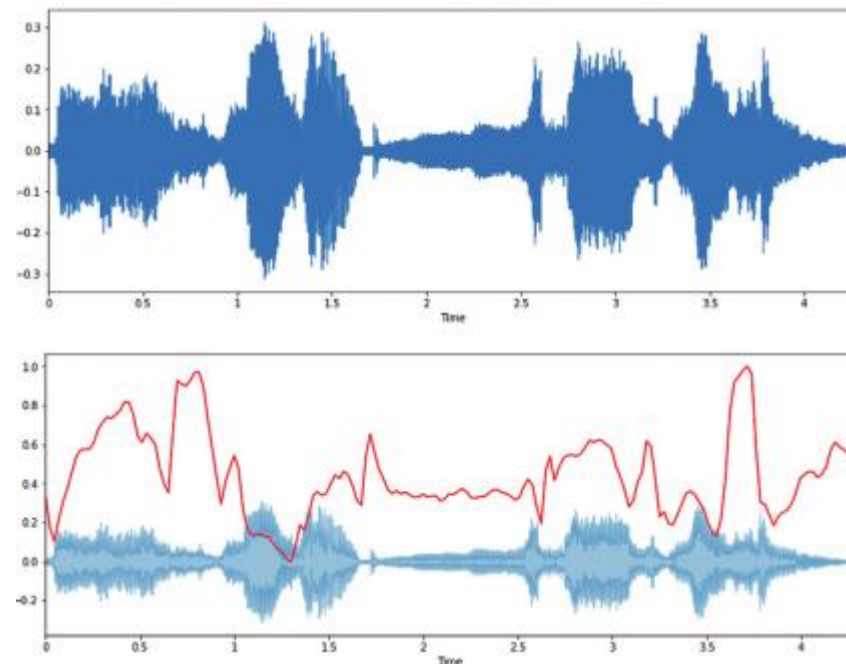
# 파이썬으로 할 수 있는 작업

- 이미지 분석
  - 사진, 그림 등의 이미지에서 속성을 추출하여 데이터로 변환 후 이미지 분류, 객체 검출, 객체 분할 등 영상 인식 알고리즘을 개발하는데 활용



## 파이썬으로 할 수 있는 작업

- 사운드 분석
  - 소리 데이터에서 음량, 진폭 등의 속성을 추출해 데이터로 변환하고 시각화하며, 추출한 데이터를 이용해 음성 인식 등의 알고리즘을 개발



# 파이썬으로 할 수 있는 작업

- 응용 소프트웨어 개발
  - 파이썬은 범용 프로그래밍 언어로 Django, Flask, FastAPI 등의 프레임워크를 활용하여 유튜브, 인스타그램, 넷플릭스와 같은 서비스를 운용하는데 사용



## 캐글Kaggle

- 데이터 분석 대회가 열리는 온라인 플랫폼
  - 기업이 당면한 문제와 데이터를 사이트에 올리면 개인 참가자들이 알고리즘을 개발해 제출하는 방식으로 진행
- 캐글의 설문조사에 따르면 캐글 이용자가 가장 많이 사용하는 프로그래밍 언어가 파이썬

