

# 한발 더 나아가기

성공회대학교

IT융합자율학부 인공지능전공

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

## 집중할 방향 정하기

- 사람마다 데이터에 대한 정의를 다르게 내림
- 스스로 데이터 분석가라 하는 사람들도 서로 다른 일을 하기도 함

## 데이터 분석 관련 분야

- 데이터 분석
- 데이터 엔지니어링
- 데이터 시각화
- 웹 애널리스트

## 데이터 분석data analytics

- 데이터를 이용해 현상을 해석하거나, 변수 간의 관계를 검증하거나, 미래를 예측하는 예측 모델을 만드는 등 데이터 분석에 초점을 맞춘 분야
- 파이썬을 비롯한 다양한 데이터 분석 도구를 다룰 수 있어야 하고 통계 분석 및 머신러닝 등 데이터 분석 관련 지식이 필요
- 데이터 분석가data analyst

## 데이터 엔지니어링data engineering

- 대량의 데이터가 지속적으로 발생할 때 데이터를 효율적으로 저장하고 추출할 수 있도록 데이터베이스를 구축하고 관리하는 분야
- 하둡Hadoop, 스팍Spark과 같은 분산 처리 시스템이나 AWS, GCP 같은 클라우드 컴퓨팅 서비스를 잘 다룰 수 있어야 함
- 데이터 엔지니어data engineer

## 데이터 시각화<sup>data visualization</sup>

- 데이터를 쉽게 이해할 수 있도록 이미지로 표현하는 분야
- 데이터 분석에서는 분석을 통해 요약된 데이터를 시각화하는 반면, 데이터 시각화는 원자료 자체를 시각화 하는데 중점을 두고 인터랙티브 그래프를 웹 페이지에 구현하기도 함
- D3.js, 태플로<sup>Tableau</sup>와 같은 데이터 시각화 도구를 다룰 수 있어야 함

## 웹 애널리틱스web analytics

- 인터넷 사용자의 웹 서비스 이용 행태를 중점적으로 분석하는 분야
- 웹 로그web log를 분석해 사용자가 언제 많이 접속하는지, 어떤 주제의 콘텐츠를 많이 공유하는지, 어떤 링크를 통해 방문하는지 등 웹 서비스 운영에 도움이 되는 정보를 도출
- 구글 애널리틱스Google Analytics 같은 웹 로그 분석 전문 도구를 다룰 줄 알아야 함

# 효율적으로 데이터 분석 능력 키우는 방법

- 분석 도구에 익숙해지기
  - 데이터 분석 관련 파이썬 패키지를 찾아 튜토리얼 진행하기
- MOOC 온라인 강의를 활용하여 학습
  - Coursera.org 등에서 python data analysis 등으로 검색해서 심화 과정 공부
- 공공데이터 활용하기
  - 공공데이터포털: <http://data.go.kr>
  - 서울 열린데이터 광장: <http://data.seoul.go.kr>
  - 한국복지패널: <http://koweps.re.kr:442/main.do>
- 자기만의 데이터 분석 프로젝트 진행하기
- 온라인 데이터 분석 대회 캐글<sup>kaggle</sup>에 공개된 데이터 분석 보고서 리뷰



# 재미있는 데이터 분석 관련 사이트

- 사이트
  - Kdnuggets <http://kdnuggets.com>
- 유튜브
  - 데이터 홀릭 <https://www.youtube.com/@dataholic4/videos>
- 페이스북 그룹
  - 한국 파이썬 사용자 모임 <http://fb.com/groups/pythonkorea>
  - 통계 마당 <https://fb.com/groups/statground>
  - 싸이그래머 <https://fb.com/groups/psygrammer>