# Foundations of Data Science

# Capstone Project Proposal

## 프로젝트 명 ( Project Title)

**학과별 취업 현황 분석 및 시각화**

## 팀원 (Team Members)

* 박철, 컴퓨터 공학부 데이터공학과
* 김숙, 컴퓨터 공학부 데이터공학과
* 홍길동, 행정학과

## 프로젝트의 개요 (Abstract)

저희 프로젝트는 취업에 관련된 다양한 데이터를 비교 분석하고 시각화 하는 프로젝트입니다. 저희 조는 현재 모두 대학교 2학년에 재학하고 있습니다. 때문에 2~3년 사이에 졸업을 하여 취직을 해야 하기에 취직을 할 때 가장 중요한 점이 무엇일까 하고 생각해봤습니다. 그 결과, 저희와 같은 학생들이 관심을 가질 만한 급여, 만족도, 취업률에 관한 데이터를 다뤄 보기로 결정했습니다.

그래서 2018년~ 2019년 동안의 전국 모든 대학교의 학과들, 그리고 그 졸업생들의 종합되어 있는 급여, 만족도, 취업률 데이터를 찾아서 시각화를 해보려고 합니다. 이러한 데이터를 찾아본 후 저희는 임금 수준에 따른 만족도와 학과 별 취업률 현황 등의 정보가 궁금했습니다. 그래서 만족도 데이터, 취업률 데이터, 임금 데이터 이 세가지 데이터를 비교 분석하여 그 관련성에 대해 중점적으로 알아보고자 합니다. 이 데이터들에 대한 분석과 시각화가 잘 된다면, 대학생들이나 청소년들이 직업을 결정하는 데에 조금이라도 더 도움이 될 것이라고 생각합니다.

## 목표 (Goals)

* 2018년과 2019년의 학과 별 임금 데이터를 비교한 것을 토대로 시각화
* 임금이 가장 높은 학과 10개와 낮은 학과 10개에 대한 시각화
* 임금 수준 별 만족도 수준 비교 분석을 통한 관계성 파악
* 취업률이 가장 높은 학과 10개의 취업률 데이터 시각화

## 데이터 확보와 분석에 사용될 기술

* <http://datakorea.datastore.or.kr/> 에서 검색 키워드에 학과들을 넣어서 학과 별 임금에 대한 데이터를 확보하려고 합니다.
* 분석에 사용될 기술 (어떤 방법으로 데이터 분석을 할 것인가 설명)
* Csv파일을 R studio에 불러와서 modelling 하고 visualization하려고 합니다.

## 주별 계획 (Timeline and Plans)

* 1주차:
  + 데이터 수집 및 분석
  + 데이터 분석 및 시각화
  + 데이터 모델링
* 2주차:
  + 모델링 결과 분석 및 선정
  + 시각화 테마와 색상 선정
  + 시각화 완성 및 리포트 작성
* 3주차:
  + 발표 준비 (슬라이드)