Foundations of Data Science with Capstone for SMU

# Data Project Guidelines

8번의 강의 후 4일 동안에 팀별 (2인 1팀) 또는 개인 실습프로젝트를 진행하게 됩니다. 실습프로젝트는 데이터프로젝트를 정하고 아래와 같은 task를 진행하여야 합니다.

1. 데이터 프로젝트 정하고 데이터 찾기 (Research)
   1. <https://www.kaggle.com/datasets>
2. 제안서 작성하기 (Communication): 페이지2에 첨부된 제안서 템플릿을 사용하여 작성.
3. 데이터 모으기 (Data Collection)
4. 데이터 정리하기 (Data Wrangling)
5. 데이터 시각화 (Data Visualization)와 데이터 분석 (Analysis)을 통하여 결과 도출하기
6. 결과를 리포트로 만들기 (Documenting) – data\_project\_report.Rmd 화일을 사용하여 리포트 작성. 이 리포트를 바탕으로 slides 작성.
7. 결과를 10분 정도의 프레젠테이션으로 만들고 발표하기 (Communication). 가능하다면 리포트의 중요한 내용을 R Presentation화일로 슬라이드 형식으로 만들어서 RPubs에 출판 예정.

Data Project Proposal

## 프로젝트 명 (Project Title)

## 팀원 (Team Members)

## 프로젝트의 개요 (Abstract)

* 이 프로젝트가 어떤 data project인지에 대해 설명합니다.
* 구체적으로 데이터를 통해 알아보고 싶은 질문이나 가정에 대해 기술합니다.
* 어떤 분석을 할 것인지 계획합니다.
* 분석 결과가 어떤 영향을 미칠 수 있는지, 어떤 의미를 가지는지에 대해 기술합니다.

## 목표 (Goals)

* 어떤 결과물을 만들어 낼 계획인지 설명합니다.

## 데이터 확보와 분석 방법

* 데이터 확보 (어디서, 어떻게, 어떤 데이터를 확보할 것인가 설명합니다.)
* 분석에 사용될 기술 (어떤 방법으로 데이터 분석을 할 것인가 설명합니다.)

(참고: 프로젝트 제안서는 단순히 무엇을 하겠다고 마음먹는 단계에서 쓰는 것이 아닙니다. 먼저 원하는 가설이나 질문이 무엇인지 알았으면, 그 가설에 대한 검증, 질문에 대한 답을 할 수 있는 분석결과를 도출하는 것이 데이터 프로젝트의 목표입니다. 그래서 제안서는 그 목표를 위해 필요한 제반 사항들이 있고 목표를 주어진 시간 안에 만들어 낼 수 있는지 확실하게 조사를 한 상태에서 써야합니다. 말하자면 여러분의 제안서는 다음의 질문들에 답할 수 있어야 합니다.

1. 데이터를 통해서 어떤 것을 알아보고 싶은가?
2. 필요한 데이터가 있는가?
3. 그 데이터는 결과물을 만들기에 충분한 정보를 가지고 있는가?
4. 결과물은 만들어 질 수 있는가?
5. 결과물이 정해진 시간 안에 만들어 질 수 있는가?
6. 만들어진 결과물을 프로젝트의 목적에 맞게 사용할 수 있는가?