

단국대학교 2020 동계 빅데이터 강의 계획서

교육과정명	파이썬을 이용한 공공빅데이터 분석 및 시각화 표현
교육일정	2020년 12월 28(월) ~ 30일, 2021년 1월4일~5일, 9:00~19:00 (점심시간 1시간 제외 1일 9시간)
교육시수	총 45시간 (5일)
교육목표	파이썬 라이브러리를 이용한 공공빅데이터 분석과 시각화 표현을 학습한다. 시각화 플랫폼인 Power BI와 스플링크를 이용해 빅데이터 플랫폼을 학습한다.
교육내용	<p>Day 1 – 빅데이터 분석과 파이썬 환경 설정 및 기초</p> <ul style="list-style-type: none"> ① 비정형 데이터 분석(실습) <ul style="list-style-type: none"> • R을 이용한 비정형 데이터 분석 실습 • R문법을 이용한 네이버 실시간 영화순위 가져오기 ② 파이썬 환경 설정 및 구축 ③ 파이썬의 기초1 <ul style="list-style-type: none"> • 변수 • 제어구조 • 함수 • 자료구조 ④ 파이썬의 기초2 <ul style="list-style-type: none"> • 자료형(data type) • 제어문(if, for) • 파일입출력(read, write) • 예외처리 <p>Day 2 – 빅데이터와 데이터 시각화 기초</p> <ul style="list-style-type: none"> ① 빅데이터 시각화의 개념 ② 빅데이터 시각화 접근법 ③ 기상청 공공데이터 살펴보기 ④ 공공빅데이터의 CSV파일 분석환경 설정하기 ⑤ 서울의 기온데이터 분석하기 ⑥ 파이썬 코드 구현하기 ⑦ 공공데이터 분석하기 <ul style="list-style-type: none"> • 기온 데이터 분석 • 2019년 여름중 가장 더웠던 날은 언제일까? • 내생일 기온 변화를 그래프로 그리기 ⑧ 실시간 데이터 분석 <ul style="list-style-type: none"> • 서울시 부동산 예측 분석 <p>Day 3– 인구 공공데이터와 대중교통 데이터로 보는 세상</p> <ul style="list-style-type: none"> ① 우린 동네 인구구조 시각화하기 ② 인구구조를 다양한 형태로 시각화 하기 ③ 우리동네 인구 구조를 파이 차트로 나타내기 <ul style="list-style-type: none"> • 제주도에 여성비율이 정말 더 높을까? 차트로 확인하기

	<ul style="list-style-type: none"> • 혈액형으로 보는 남녀 비율은? • 우리동네 연령대 성별 비율은? <p>④ 대중교통 데이터시각화 하기</p> <ul style="list-style-type: none"> • 지하철 유/무임별 이용현황 차트로 알아보기 • 지하철 시간대별 이용현황 데이터로 알아보기 • 출근시간대 사람들이 가장 많이 타고 내리는 역은 어디일까? • 시간대별로 사람들이 가장 많이 타고 내리는 역은 어디일까? <p>Day 4- 시각화 플랫폼 Power BI</p> <ol style="list-style-type: none"> ① power BI desktop 시작하기 ② 데이터 가져오기와 테이블 다루기 ③ 기본 쿼리 작업 ④ 데이터 모델링 ⑤ 데이터 시각화 <ul style="list-style-type: none"> • 다양한 시각화 유형 활용하기 ⑥ power BI 를 이용한 공공데이터 분석1 <ul style="list-style-type: none"> • 혼밥족이 살고 싶은 곳은 어디? • 서울의 미세먼지 탈탈 털어보자 <p>Day 5- 시각화 인텔리전스 플랫폼 스플링크</p> <ol style="list-style-type: none"> ① powr BI를 활용한 공공데이터 분석2 <ul style="list-style-type: none"> • 직장 그만두고 카페나 차려볼까? • 날짜별 창업 및 폐업 추이 분석 • 카페 규모와 생존 기간 관계 분석 • 유명 브랜드 카페 분석 • 카페 과밀지역 분석 ② 스플링크 엔터프라이즈 설치 및 설정 ③ 스플링크 플랫폼을 이용한 실시간 로그 시각화
<p>교재</p>	<p>강의자료 별도 제공</p>