

XML 이론(데이터 분석) 및 실습

강윤희

yhkang@bu.ac.kr, yunh.kang@gmail.com

차례

- 빅데이터시대 준비하기
- 수업목표
- 교재 및 부교재
- 주차별 수업내용
- 수강 이유
- 수업 진행 및 평가

주차별 수업 내용

- 수업내용
 - 시작에서 적용까지

파이썬 언어 이해	데이터 준비를 위한 자료 다루기	데이터 분석 프로젝트
자료형 및 제어문	CSV JSON XML	공공데이터 살펴보기
함수 및 예외처리		텀 프로젝트 계획서 작성하기
클래스와 객체 모듈	오픈 API HTTP 및 REST	프로젝트 목표 정의 데이터 수집 데이터 처리 결과 저장 및 시각화
	자료처리 패키지: numpy, pandas, matplotlib	

주차별 수업 내용

주차	수업주제 및 내용	주차 목표
1	수업 목표 및 수업 내용 소개	파이썬 환경 구축
2	파이썬 자료형 - 숫자, 문자열자료형 - 컬렉션 자료형	<ul style="list-style-type: none">• IDLE 사용 자료를 변수에 저장한 후 확인함• 공공데이터자료 소개 후 프로젝트 주제 선정 준비
3	파이썬 제어문 - 조건문 - 반복문	활용 자료 탐색 및 이해하기
4	파이썬 함수 및 예외처리 - 파이썬 함수 정의 및 호출 - 예외처리	활용 자료 형식 및 특징이해하기
5	파이썬 클래스 및 객체 활용 - 클래스와 객체 - 모듈	텀프로젝트 계획서 작성
6	파이썬 자료 활용 - CSV 개요 및 활용하기	오픈 API 소개
7	파이썬 자료 활용 - JSON 개요 및 활용하기	HTTP 및 REST 이해하기 (공공데이터 접근: JSON)

주차별 수업 내용

주차	수업주제 및 내용	주차 목표
8	파이썬 자료 활용 - XML 개요 및 활용하기	HTTP 및 REST 이해하기 (공공데이터 접근: XML)
9	빅데이터 자료처리 - numpy 패키지	숫자데이터 활용 통계처리
10	빅데이터 자료처리 - pandas 패키지	접근 자료를 데이터프레임으로 구성
11	빅데이터 자료처리 - matplotlib 패키지	데이터 시간화
12	웹 자료크롤링(crawling) 및 스크래핑(scraping)	HTTP 요청/응답으로 HTML 문서 다운로드 및 자료추출
13	지도데이터 처리	데이터 융합 및 데이터 저장
14	팀별 프로젝트 결과 발표	팀별 발표

수강 이유는 ?

- 다양한 데이터 활용에 관심 있음 ?
 - 나만의 공공데이터 활용 과정을 수행함
- 문제 해결을 위한 프로그램을 익힘
 - 응용프로그램에서 자료에서 가치를 발견함
- 다양한 패키지를 사용하여 데이터 분석 방법을 익힘
 - 개발자(데이터 엔지니어/데이터 과학자)에게 있어서는 필수적인 도구를 익힘
- 데이터는 다양한 환경에서 생성되고 여러 목적으로 사용됨
 - 데이터는 모바일, 웹, 기업 내 사용하고 있음

Q&A

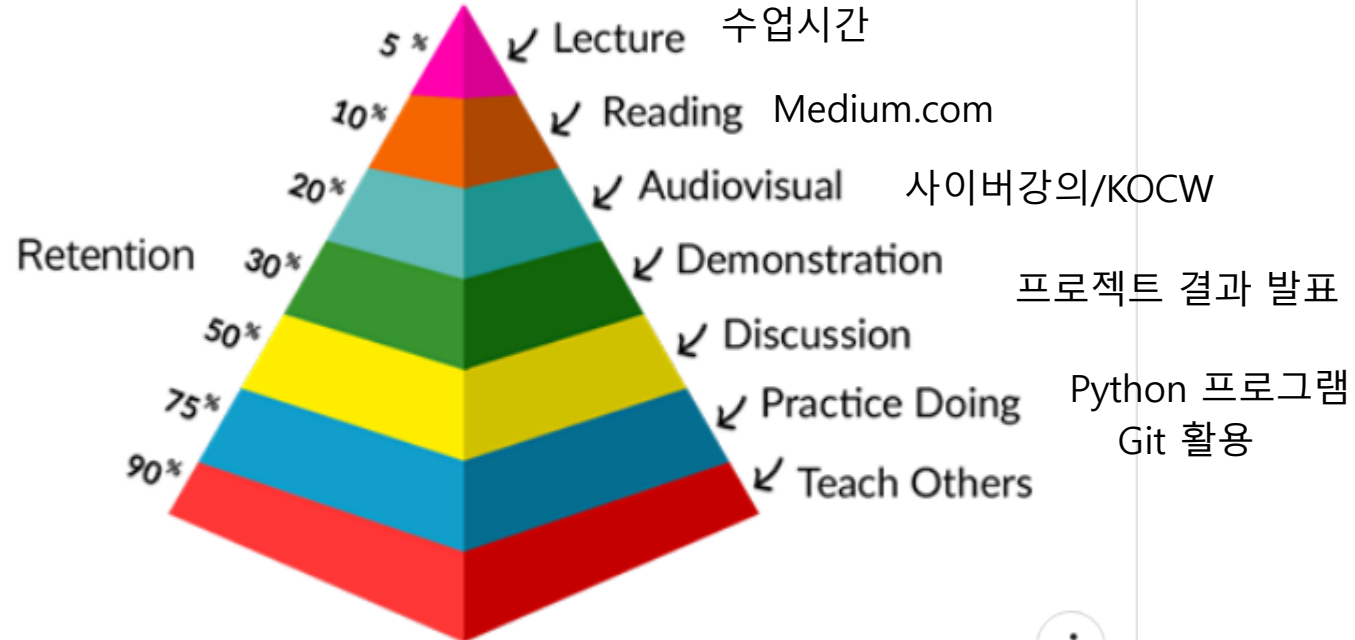
- 수업 시작 전에 게시판 참조
 - 사이버 게시판
- 개인적인 질문
 - E-mail: YunH.KangATgmail.com
 - Offline meeting : 본부동 1128호
- 수업진행 관련
 - 수업 내용에 질문이 있는 경우 수업 중 문제 해결
 - 리포트 관련 문의는 사이버 게시판 이용 문제 해결
 - 리포트 관련 사이버 게시판의 경우 접근
 - 자신이 단계별로 수행한 내용을 적고, 해결이 필요한 부분을 명확히 기술함
 - Simple and Clear

수업진행



Henry Choi

4 hrs · 👥



MEDIUM.COM

Explanation Effect: Why You Should Always Teach What You Learn

학점 부여 평가방법 및 기준

- 성적 (100점)
 - 출석 : 20 %
 - ❖ 수업 이후 필수적으로 확인하여야함
 - 중간 : 30 %
 - ❖ Python 문법에서 자료활용
 - 기말 : 30%
 - ❖ Python 패키지 활용
- Term Project : 20 %
 - ❖ 수행 결과가 있어야 하며, 보고서가 있는 경우 인정됨
 - ❖ 주어진 문제를 정의하고 관련 프로그램 작성

학점 부여 평가방법 및 기준

- 학점 기준

- A 학점 : 주어진 문제 해결을 위해 자료 처리(분석) 기술 적용을 이해하고, 해당 문제를 프로그램을 통해 해결할 수 있음
- B 학점 : 주어진 문제 해결을 위해 자료 처리(분석) 기술 적용을 이해하고 일부 문제를 프로그램을 통해 해결할 수 있음(미흡)
- C 학점 :: 주어진 문제 해결을 위한 자료 처리(분석) 기술 적용을 이해하지만 해당 문제를 프로그램을 통해 해결하지 못함
- D 학점 : 주어진 문제 해결을 위한 자료 처리(분석) 기술 적용 이해가 미흡하고 해당 문제를 프로그램을 통해 해결하지 못함

❖ 프로그램을 작성하지 못한 경우 C/D 학점 부여

질문 및 응답

혁신은 울타리 밖에서 일어난다
데이터의 활용을 다양한 부분에서 찾아봄