

2023학년도 1학기

데이터분석과 시각화

기말고사 대비 요점정리 및 미완성 프로그램 설명

기말고사 안내



1. Lecture Note 학습평가 : 20문제

2. 나머지 : 10문제 단답형

- 문제해결 프로그램을 완성한 후
- 완성된 프로그램의 코드를 보고 시험 응시

기말고사 안내

시험 범위	9~14주차
주차별 시험 요약	
9주차	데이터 분석 프로세스에서의 모델링 정의 데이터 분석 방법론별 정의와 차이점 데이터 분석 기법 중 회귀분석과 탐색적 자료 분석의 정의
10주차	용어의 정의 (확률, 추정, 가설검정, 표본, 기술통계, 표준편차, 평균, 분산)
11주차	단변량, 다변량의 정의와 변수의 구분 단변량 데이터 분석 전 확인사항 다변량 데이터 시각화 방법 차원축소의 정의
12주차	Alberto Cairo의 Visualization Wheel 에서 Abstraction - Figuration (추상-형상) 의미 시각화 절차에서의 데이터 전처리 중요성 시각화 오류 또는 오해의 소지가 있는 시각화 사례와 집단화 법칙의 정의
13주차	정형 데이터 정의와 이미지 데이터의 구성 분석의 목적
14주차	텍스트 데이터 분석과 워드클라우드, 텍스트 데이터 분석이 어려운 이유

문제해결 프로그램

[문제상황]

경상남도 밀양시 지역의 주택가격을 예측하는 모델을 만들기 위하여 데이터를 분석하고자 한다.

모델생성은 설명 가능한 회귀분석 방법을 이용할 예정이며, 분석과정 중간에 시각화를 진행하려고 한다.

이러한 과정을 수행하는 분석 프로그램을 완성하세요.

완성된 프로그램을 이용하여 시험문제에 답하세요.

- 미완성 문제해결 프로그램 다운로드 : 학습자료실

[문제예시]

▼ [문제 - 1]

전체 프로그램이 문제없이 실행되기 위해 필요한 필수 패키지를 import 하세요.

import 하는 패키지의 갯수 제한은 없지만, 필수가 아닌 패키지를 import하는 경우 하나 당 -0.1점이 감점됩니다.

```
[ ] import ?  
import ?  
import ?  
...  
...  
...  
import ?
```