

기출복원_05회_문제

※ 작업형3은 5회까지는 시행되지 않아 별도 제공되지 않으며, 모의고사 3회분이 추후 제공 될 예정입니다.

작업형1

01. 아래는 공공데이터 포털에서 제공하는 폐기물관리법」에 따른 지방자치단체별생활쓰레기 및 음식물쓰레기 종량제 봉투 가격에 대한 정보와 관련된 데이터(전국_종량제봉투_가격표준.csv)의 일부이다. 전국의 2L 음식물쓰레기 규격봉투의 평균 가격을 계산하여라.(단, 평균 계산 시 가격이 0원인 경우는 제외하고 결과는 소숫점을 버리고 정수로 출력할 것)

02. 주어진 bmi 데이터셋(bmi.csv)를 통해 BMI를 계산한 후 정상 체중에 속하는 인원과 과체중에 속하는 인원의 차이에 대한 절대값을 구하여라.(단, 출력시 정수로 출력할 것) 여기서 BMI는 몸무게/키의 제곱(몸무게의 단위는 kg, 키의 단위는 m)으로 계산되며, 아래와 같은 기준으로 분류된다.

- 18.5 미만이면 저체중
- 18.5 이상이고 23.0미만이면 정상체중
- 23.0 이상이고 25.0미만이면 과체중
- 25.0 이상이면 비만

03. 다음은 학교알리미에서 제공하는 공개용 데이터의 일부(2022_부산초등학교_전출입학생현황.csv)로 부산광역시 교육청 산하의 초등학교의 전출입 현황이다. 순전입학생의 수가 가장 많은 순전입학교의 전체학생수를 구하라. 여기서 순전입학생의 수는 총 전입학생의 수 - 총 전출학생의 수로 계산되며, 순전입학교는 총 전입학생의 수가 총 전출학생의 수보다 많은 학교를 의미한다.

작업형2

01. 아래는 중고 포드(Ford) 자동차의 가격 예측을 위한 데이터의 일부이다.

(가) 제공 데이터 목록

- ① Ford_Price_train.csv : 13,470대의 중고 자동차의 가격과 이와 관련된 주요 컬럼 정보 데이터 (학습용), csv 형식의 파일
- ② Ford_Price_test.csv : 4,490대의 중고 자동차의 가격에 관련된 주요 컬럼 정보 데이터 (평가용), csv 형식의 파일

(나) 데이터 형식 및 내용

컬럼	설명
price	중고 자동차의 가격(단위:달러)
model	브랜드명
year	생산연도
transmission	변속기(Automatic, Manual, Semi-Auto)
mileage	주행 마일 수
fuel_Type	연료 타입(Petrol, Diesel, Hybrid)
Annual Tax	연간 세금
mpg	마일/갤런
engineSize	자동차 엔진 크기

13,470대에 대한 학습용 데이터(Ford_Price_train.csv)를 이용하여 가격 예측 모델을 만든 후 이를 평가용 데이터(Ford_Price_test.csv)에 적용하여 얻은 4,490대 예측 가격을 다음과 같은 형식의 CSV 파일로 생성하시오.(제출한 모델의 성능은 RMSE 평가지표에 따라 채점)

<제출형식>

변수	설명
pred	중고 자동차의 가격(단위:달러)에 대한 예측값

<유의사항>

- 성능이 우수한 예측모형을 구축하기 위해서는 적절한 데이터 전처리, Feature Engineering, 회귀알고리즘 사용, 초매개변수 최적화, 모형 앙상블 등이 수반되어야 한다.
- result.csv 파일이 만들어지도록 코드를 제출한다.
- 제출한 모델의 성능은 RMSE 평가지표에 따라 채점한다.