

## 2025 단 · 쪽 · 키 빅데이터 분석기사 파이썬 작업형 정오표

도서 내의 오류 관련 내용은 홈페이지 내의 <자료실·빅분기> 코너를 통해 알려드립니다. 앞으로 믿을 수 있는 도서를 만들기 위해 더욱 노력하겠습니다.

| 정오표 최종수정일 | 2025년 04월 011일                                                                                                                                                                                                                                                                                      | 현재 판쇄                                                                                                                                                                                                                                                                                                                                                        | 1판 1쇄 |
|-----------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------|
| 위치        | 현재 내용                                                                                                                                                                                                                                                                                               | 바로잡기                                                                                                                                                                                                                                                                                                                                                         | 비고    |
| 174       | ans=round(boston_sort_80['CRIM'].mean(), 3)<br>정답 5.759                                                                                                                                                                                                                                             | ans=round(boston_sort_80['CRIM'].mean(), 2)<br>정답 5.76                                                                                                                                                                                                                                                                                                       |       |
| 188       | 기출문제 09<br>df['date_added']=pd.to_datetime(df['date_added'],<br>format='%Y-%m-%d',errors='coerce')                                                                                                                                                                                                  | df['date_added']=pd.to_datetime(df['date_added'],<br>errors='coerce')<br>• format='%Y-%m-%d' 옵션을 사용했을 때 오류가 발생하는 이유는 데이터의 실제 날짜 형식과 format이 맞지 않기 때문입니다.<br>• format='%Y-%m-%d' 제외하고 실행해주세요<br>• 경고일 뿐 오류는 아님 → 코드 실행은 정상적으로 되고, 결과도 정확함                                                                                                                     |       |
| 212       | 20. 국가별 맥주 소비량 데이터이다.                                                                                                                                                                                                                                                                               | 20. 국가별 맥주 소비량 데이터이다.<br>import pandas as pd<br>df=pd.read_csv('c:/data/8_type1_1.csv')                                                                                                                                                                                                                                                                      |       |
| 215       | df=pd.read_csv("c:/data/8_type1_2/.csv")                                                                                                                                                                                                                                                            | df=pd.read_csv("c:/data/8_type1_2/.csv",encoding='cp949')                                                                                                                                                                                                                                                                                                    |       |
| 286       | 기출문제 06<br>from sklearn.preprocessing import OneHotEncoder                                                                                                                                                                                                                                          | from sklearn.preprocessing import LabelEncoder                                                                                                                                                                                                                                                                                                               |       |
| 289       | X = train.drop(columns=['Performance Index'])<br>Y = train['Performance Index']<br>le = LabelEncoder()<br>train['Extracurricular Activities']=<br>le.fit_transform(train['Extracurricular Activities'])<br>test['Extracurricular Activities'] =<br>le.transform(test['Extracurricular Activities']) | le = LabelEncoder()<br>train['Extracurricular Activities']=<br>le.fit_transform(train['Extracurricular Activities'])<br>test['Extracurricular Activities'] =<br>le.transform(test['Extracurricular Activities'])<br>X = train.drop(columns=['Performance Index'])<br>Y = train['Performance Index']                                                          |       |
| 290       | 7) 모형평가<br>y_test=test_encoded["Performance Index"]<br>y_pred = model.predict(X_test)<br>rmse=np.sqrt(mean_squared_error(y_test, y_pred))<br>print(rmse)                                                                                                                                            | from sklearn.model_selection import train_test_split<br>X_train, X_valid, Y_train, Y_valid = train_test_split(X, Y,<br>test_size=0.2, random_state=42)<br>model = RandomForestRegressor(random_state=42)<br>model.fit(X_train, Y_train)<br>Y_pred_valid = model.predict(X_valid)<br>rmse = np.sqrt(mean_squared_error(Y_valid, Y_pred_valid))<br>print(rmse) |       |