공동교육과정 빅데이터분석

**데이터 시각화 수행평가**

|  |  |  |  |
| --- | --- | --- | --- |
| **이름 :** | 박준 | **희망진로:** | 게임 개발자 |
| ▶주제 | | | |
| 최근 20년간 제주도의 월 평균 기온차 | | | |
| ▶선정 이유 | | | |
| 월이 지날때마다 날이 추워지거나 더워지는 걸 체감할 수 있는데 이것을 데이터로 분석하여 언제 가장 추워지고 가장 더워지는지 언제가 가장 기온차이가 컸는지와 같은 정보를 수치화하고 이를 시각화하면 재미있을 것 같아서 | | | |
| ▶수집한 데이터에 대한 설명(사이트 url이나 출처 반드시 작성할 것) | | | |
| - 설명 : 해당 지역의 날짜와 그에 따른 평균기온, 최저기온과 최고기온 등의 정보가 있다.  - 컬럼 : '날짜', '지점', '평균기온', '최저기온', '최고기온’    - 데이터 수 :  - 출처 : <https://data.kma.go.kr/stcs/grnd/grndTaList.do?pgmNo=70>  코랩 주소 : https://colab.research.google.com/drive/1SfGPswbhH5o3FM2B6qJ9EI3XQYohte5W?usp=sharing | | | |
| ▶데이터 분석을 위한 과정(알고리즘) | | | |
| 1. 기상자료개방포털\_기후 데이터 사이트에 접속한 후 제주도의 1970년부터 현재까지의 기후를 csv파일로 다운로드 받는다.  2. 데이터 분석의 편의를 의해 csv파일에서 쓸모없는 부분을 삭제한다 .  3. pandas 활용을 위해 csv 파일을 pandas를 이용하여 데이터프레임으로 바꾼 후 칼럼명을 바꾼다.  4. 반복문과 조건문을 이용하여 2020년부터 현재까지의 월 평균 기온을 담은 “monthValue”라는 이름의 리스트 변수를 선언 후 다시 반복문을 사용하여 각 월의 기온차를 “1월기온 - 12월기온, 2월기온 - 1월기온, ……, 12월 기온 - 11월 기온” 의 순서로 담은 “월교차” 라는 이름의 리스트 변수를 선언한다.  5. matplotlib.pyplot 라이브러리와 위에서 만든 리스트를 사용하여 월 평균 기온차를 막대그래프와 선 그래프로 시각화한다.  6. 리스트에 절댓값을 취한 후 5의 과정을 반복한다. | | | |
| ▶데이터 시각화 그래프(캡처) | | | |
| 1. 월 평균 기온차에 절댓값을 취하기 전      1. 월 평균 기온차에 절댓값을 취한 후 | | | |
| ▶데이터 분석으로 알 수 있는 결과 | | | |
| 제주도에서 온도가 가장 많이 떨어지는 때는 평균적으로 5.41°C 가 떨어지는 10월과 11월 사이이고 가장 많이 오르는 달은 4.44°C 가 오르는 6월에서 7월 사이이다. 그리고 연교차가 절댓값으로 가장큰 때는 달은 5.41°C 가 떨어지는 10월과 11월이고 가장 적게 차이가 나는 때는 0.81°C 가 오르는 1월과 2월 사이이다. | | | |