**[오픈소스개발프로젝트 2021F]**

**출석 과제**

주의사항: 본문 글자 크기, 자간, 장평 등 서식 변경 금지

작성자: 홍진우 학번: 2018037044

|  |
| --- |
| 해당 주차의 동영상을 시청하고 내용을 요약하시오. (공백포함 600자 이상) |
| 챕터3 마지막 부분에서 데이터를 추출하여 Print에 조건식을 작성하여 true/false형으로 만든다음 true형만 추출하는 방법이 있다. 직접 true/false리스트를 만들어서 적용할 수도 있다. 벡터와 스칼라를 이용해서 사칙연산을 하는 법을 배웠다. 챕터4에서는 추출된 데이터들을 이용해서 시각화하는 것을 배웠다. 단순 표로 나열된 수치로만 확인하면 문제가 발생할 수 있기 때문에 4개의 그룹으로 쪼개서 수치화 하는 것이 필요하다. 파이썬의 matplotlib라이브러리를 이용하여 그래프를 그릴 수 있다. plot함수를 이용하여 x,y값을 받아서 그래프를 그리고 o인자를 사용하여 점으로 표현할 수 있다. Set\_title, suptitle함수를 사용하여 각 그래프의 제목, 전체 그래프의 제목을 만들 수 있다. 히스토그램은 구간별 값들의 분포를 보여준다. 구간 범위를 변경하기 위해 bin인자를 사용할 수 있고 tips셋을 이용하여 데이터를 불러와서 히스토그램 그래프를 그린다. scatter함수를 이용하여 막대로 된 그래프를 점들의 집합으로 표현한다. 박스그래프는 최솟값, q1(<25),q2(<50),q3(<75),최댓값이 필요하며 데이터들의 분포를 비교하기가 수월하다. 흔히 알고 있는 주식그래프와 형태가 유사하다. 2차원 박스그래프에 변수가 3개이면 그냥 그래프로 표현하기가 불가능하다. 연속형 변수들은 일반적으로 그래프를 그리고 이산형 변수들은 색상,점의 크기 등으로 표현을 해서 해결하는 방법이 있다. |
| 동영상 내용에서 질문 사항이 있는 경우 작성하시오. |
| 1)  2) |