

| 데이터 마이닝 및 실습 - 5주차 실습



명지대학교
MYONGJI UNIVERSITY

과제 제출 유의사항

- 이번주 과제는 . **.ipynb** 제출 바랍니다.
- 파일에 모든 과제가 다 들어갈 수 있도록 한개 파일로(comment 로 구분 필요)
- 각 그림별 5점

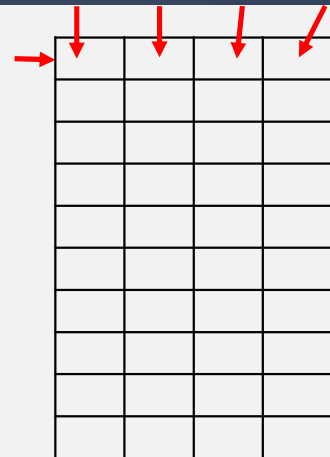
문제1

아래는 1~100 까지의 랜덤 정수 값을 10x4 개를 만들어 numpy 타입의 score 변수에 넣는 문장이다. (랜덤값이므로 각자마다 다른 결과가 나올 수 있음)

```
import numpy as np
Score = np.random.randint(100, size=(10, 4))
```

첫번째 학생

국어영어수학과학



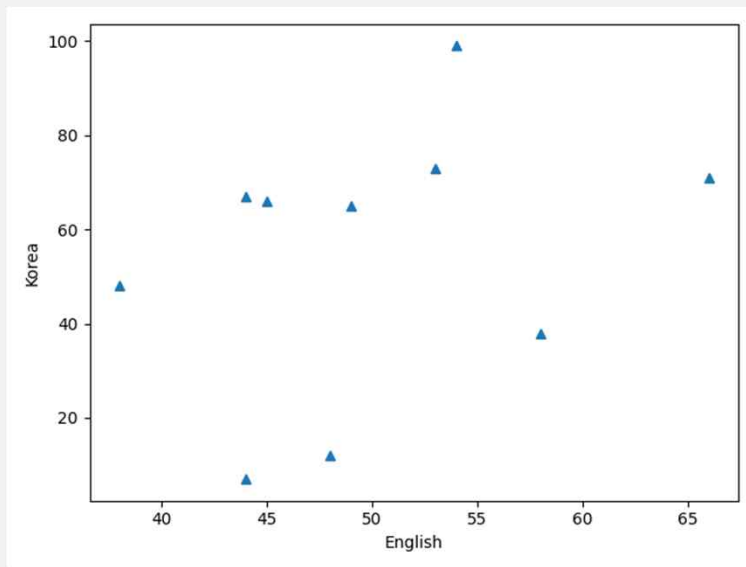
열번째 학생

위의 문장을 이용하여 함수로 구분하여 다음을 풀어라. 점수는 배열은 메인 함수에 두고 다 매개 변수로 넘기고 각 함수에서 print 하시오

Score 변수는 10 명의 학생의 국어(첫번째 열), 영어(두번째열), 수학(세번째열), 과학(네번째열)이라고 할 경우,

그림1,2

1. X 축은 영어 y 축은 국어로 흩뿌리기 하여라



2. 다음과 같이 각과목별 히스토그램을 그리되 4개의 작은 그림이 한 화면에 나타나도록 하여라

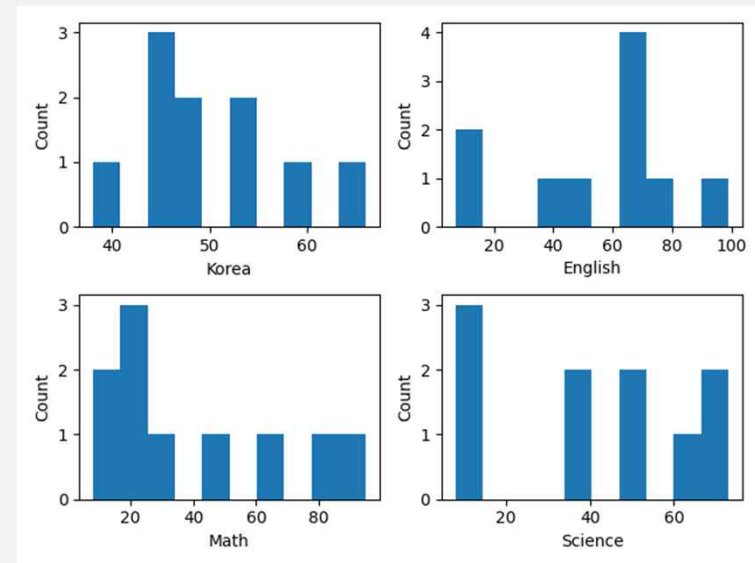
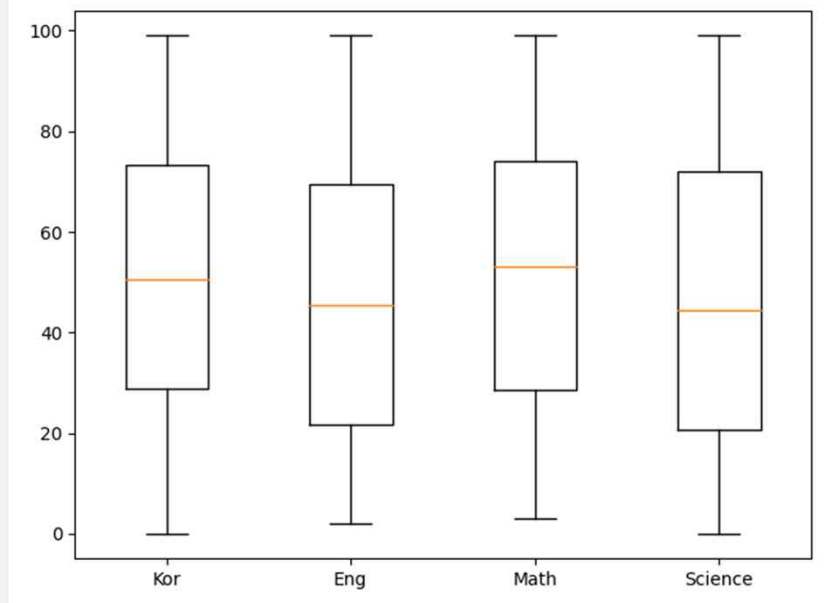
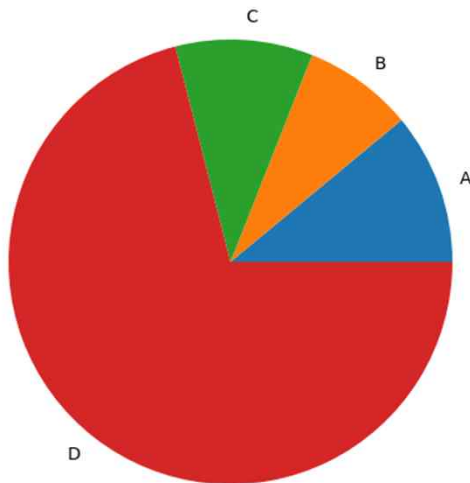


그림3,4

3. plt.pie(리스트 타입의 데이터)는 파이 차트를 나타낸다.
- 국어 점수가 90 이상 A
 - 국어 점수가 80 이상 90미만. B
 - 국어 점수가 70 이상 80미만. C
 - 국어 점수가 70 미만. D
- 라고 할 경우 아래와 같은 그림을 그리시오
4. 다음과 같이 과목의 최소 최대 및 평균값을 나타낼 수 있는 boxplot 을 사용하여 plt.boxplot (데이터) 는 데이터의 각 열별로 최소, 최대 값별로 다음과 같이 그림이 그진다. Xlabel 도 나타나도록 하여라.

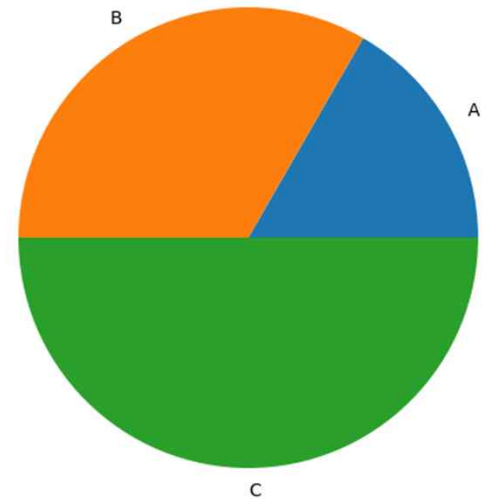


3번 파이차트 참고

3번 참고

```
arrlist = [ 10,20,30]  
label=['A','B','C']  
plt.pie(arrlist,labels=label)  
plt.show()
```

Arrlist 는 총합이 60 이기때문에
10/60, 20/60. 30/60 만큼의 비율로 그려지고
Labels는 [10,20,30]에 해당 하는 만큼의 문자열
리스트가 그대로 대응된다.



4번 참고

```
arrlist = np.array([[ 10,20,30],[15,22,33],[25,21,36]])  
plt.boxplot(arrlist)  
plt.show()
```

- 1열의 값은 10, 15, 25 이기때문에 1번째 boxplot은 평균 20과 최소 최대를 나타낸다.
- 2열의 값은 20, 22, 21 이기때문에 2번째 boxplot은 평균 21과 최소 최대를 나타낸다.
- 3열의 값은 30, 33, 36 이기때문에 3번째 boxplot은 평균 33과 최소 최대를 나타낸다.

