

데이터 마이닝 시각화 및 실습 - 6주차 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)
- 그림 그리기 5점
- 데이터 만들기 5점
- **Numpy pandas의 함수를 이용하고 for와 같은 반복문은 사용하지 않기**

6주차 실습

옆 코드는 iris 데이터를 읽어 오는 코드이다.

붓꽃은 꽃잎의 길이 넓이, 꽃받침의 길이 넓이로 3가지로 구분 가능하다.

```
from sklearn.datasets import load_iris
import matplotlib.pyplot as plt
import pandas as pd
import numpy as np
```

```
iris = load_iris()
```



6주차 실습

Iris 데이터를 살펴보면

`iris.target` 은 붓꽃의 종류에 따라 0,1,2로 구분된다.

```
array([0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,  
       0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,  
       0, 0, 0, 0, 0, 0, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,  
       1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,  
       1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2,  
       2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2,  
       2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2])
```

iris.data 은 꽃잎 길이, 꽃잎 넓이, 꽃받침 길이, 꽃받침 넓이의 값으로 총 붓꽃의 개수는 150개 이다

6주차 실습

1. 다음과 같은 `pandas.DataFrame` 데이터를 만들고, 처음 5개의 예를 보여주어라.

- 아래는 처음 5개의 예를 보여준다.

	sepal length (cm)	sepal width (cm)	petal length (cm)	petal width (cm)	target	targetName
0	5.1	3.5	1.4	0.2	0.0	setosa
1	4.9	3.0	1.4	0.2	0.0	setosa
2	4.7	3.2	1.3	0.2	0.0	setosa
3	4.6	3.1	1.5	0.2	0.0	setosa
4	5.0	3.6	1.4	0.2	0.0	setosa

6주차 실습

2.1번에서 만든 pandas데이터를 이용하여 아래 그림을 그려라

- 다음과 같이 꽃잎의 길이 vs 꽃잎의 넓이 의 그림과 꽃받침의 길이 vs 꽃받침의 넓이 의 그림을 각각 붓꽃의 클래스에 따라 다른색으로 구분하여 그림을 그리고, X,Y축의 의미와 각 색의 라벨 값도 적어라

