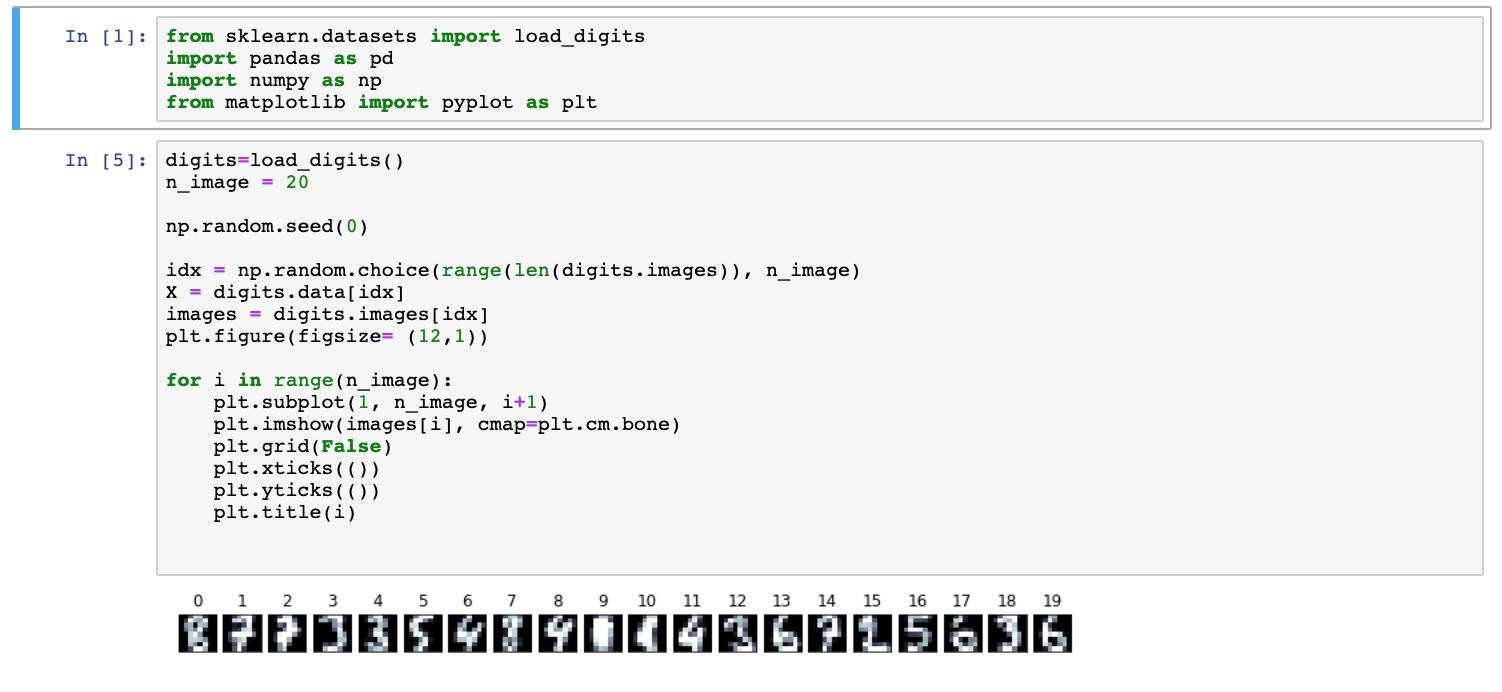
머신러닝과 딥러닝

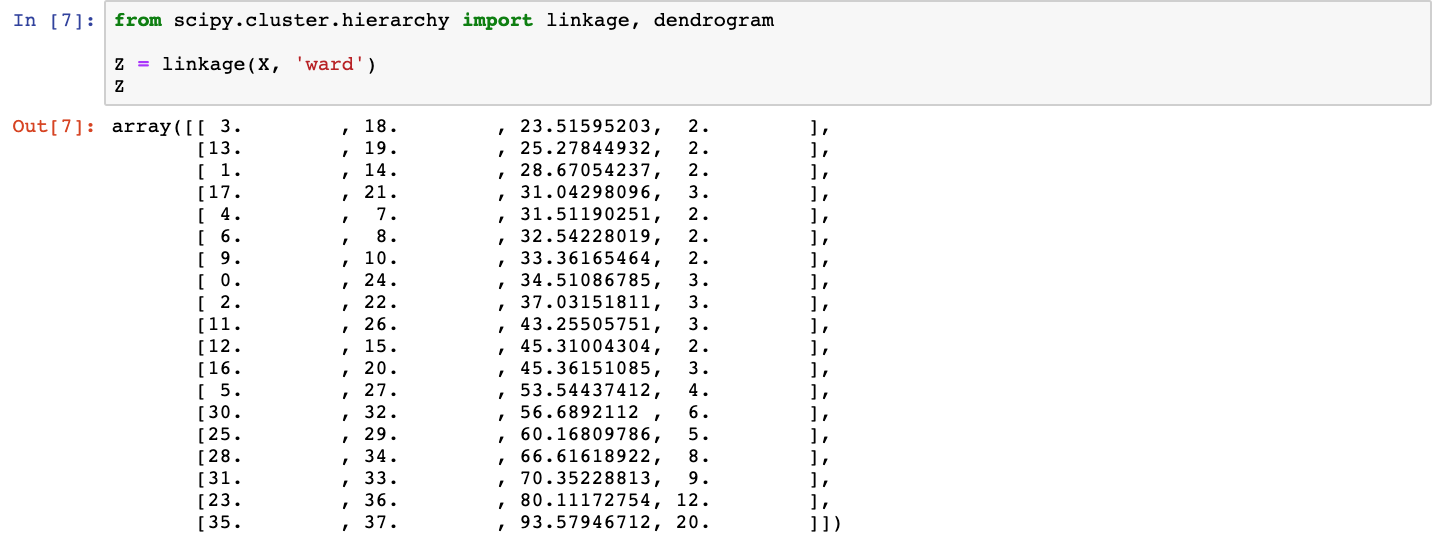
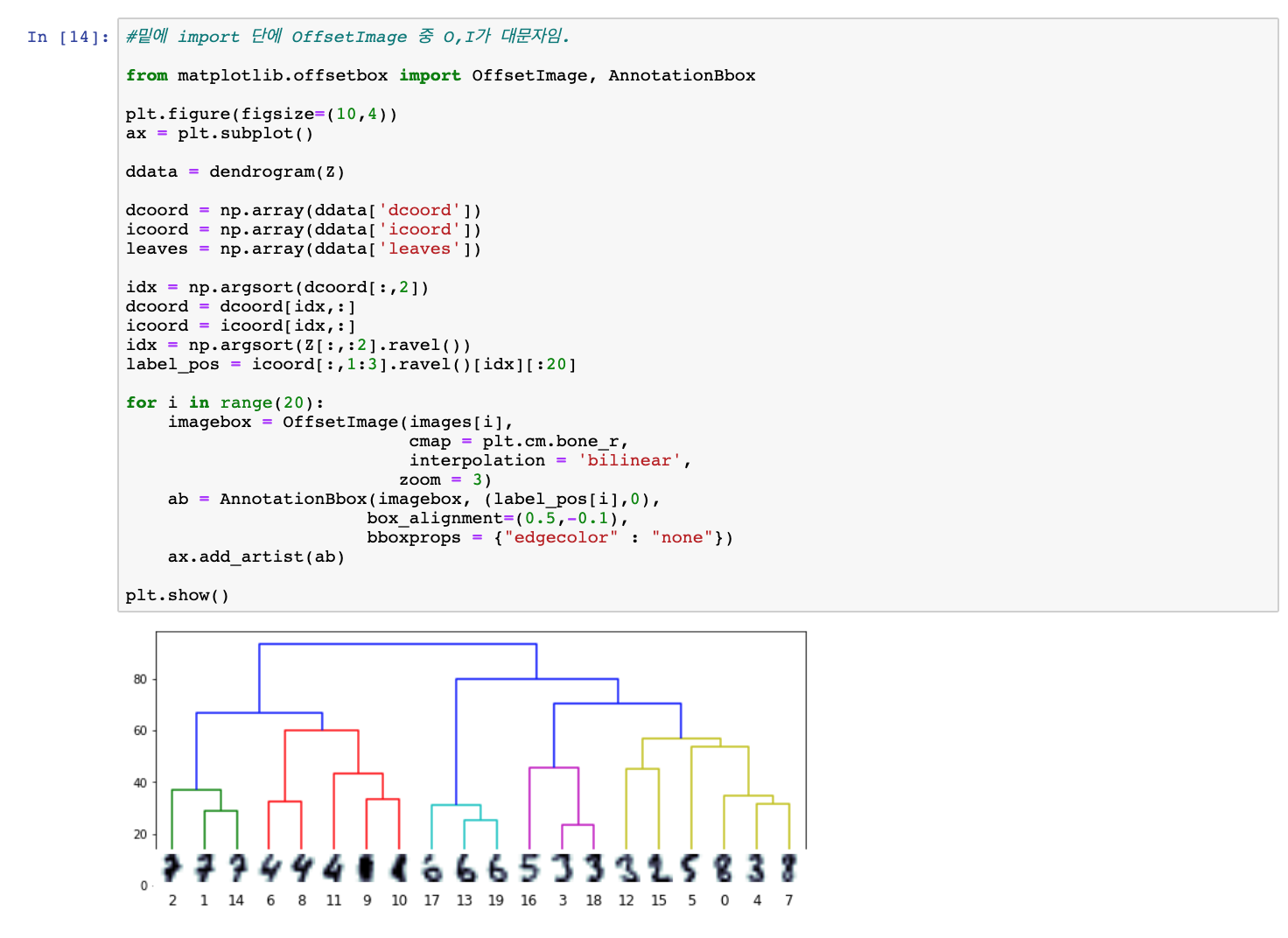
Report9

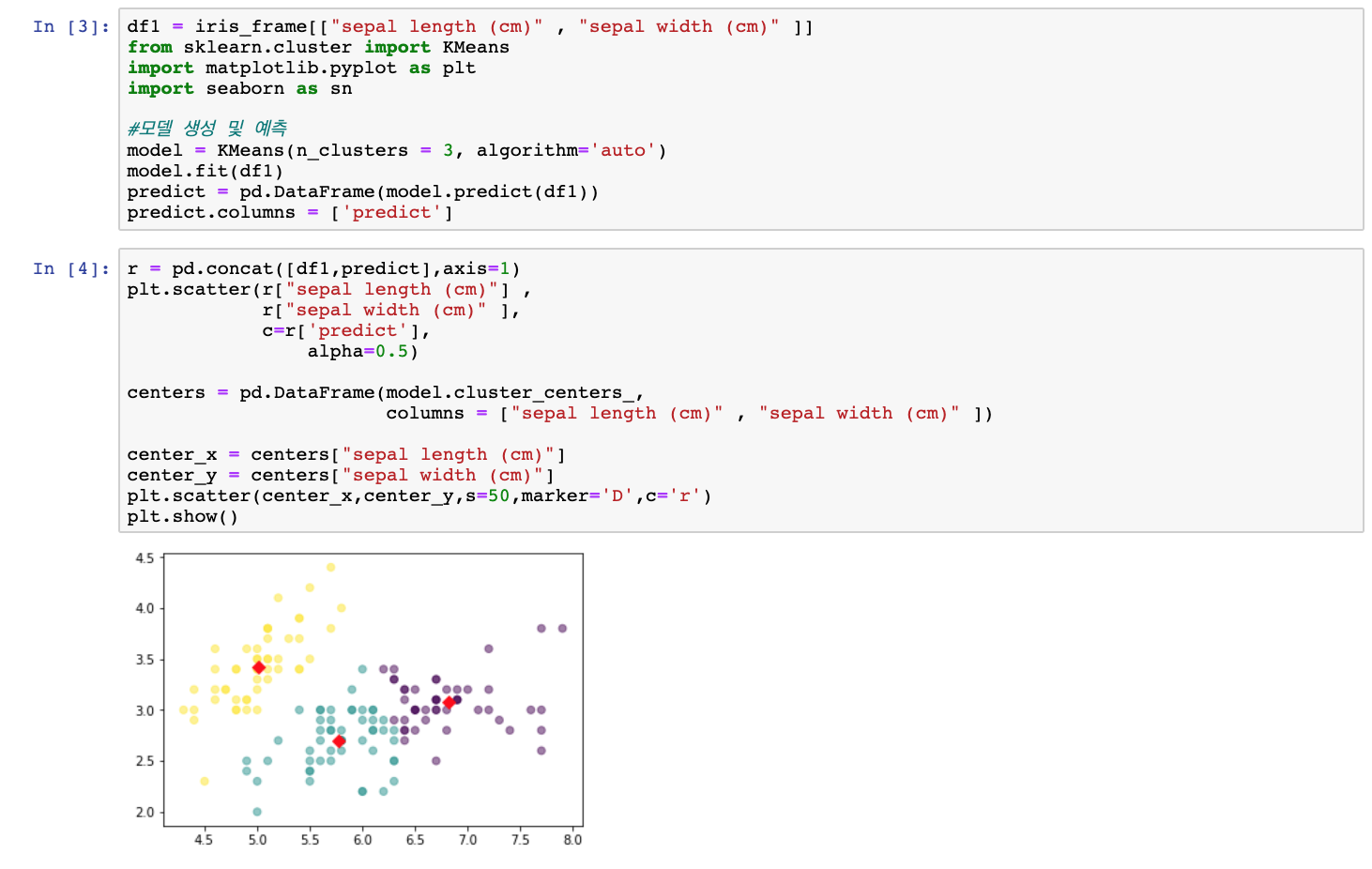
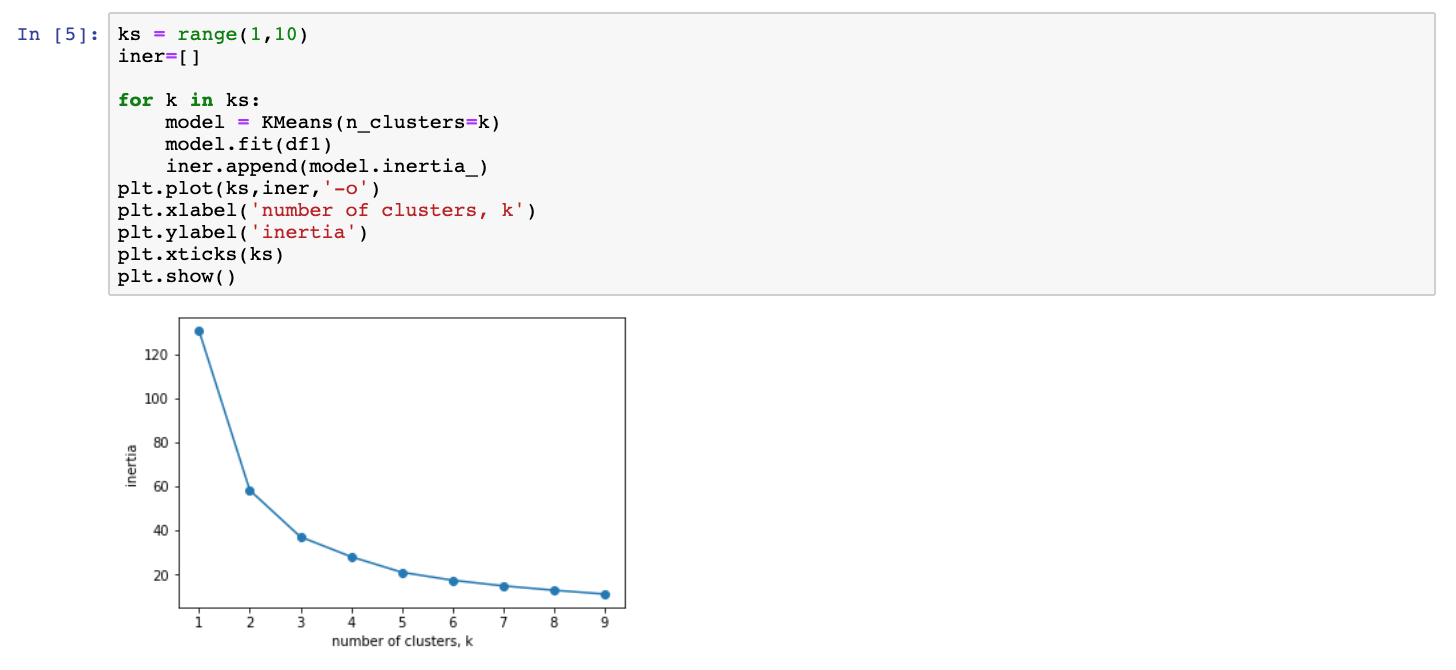
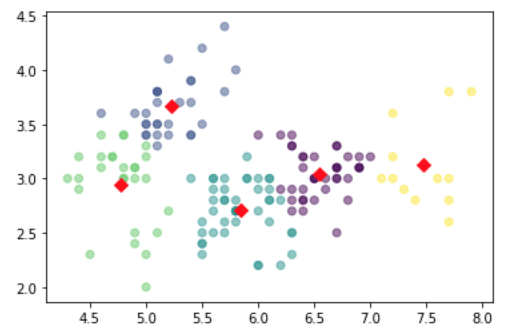
소프트웨어학과

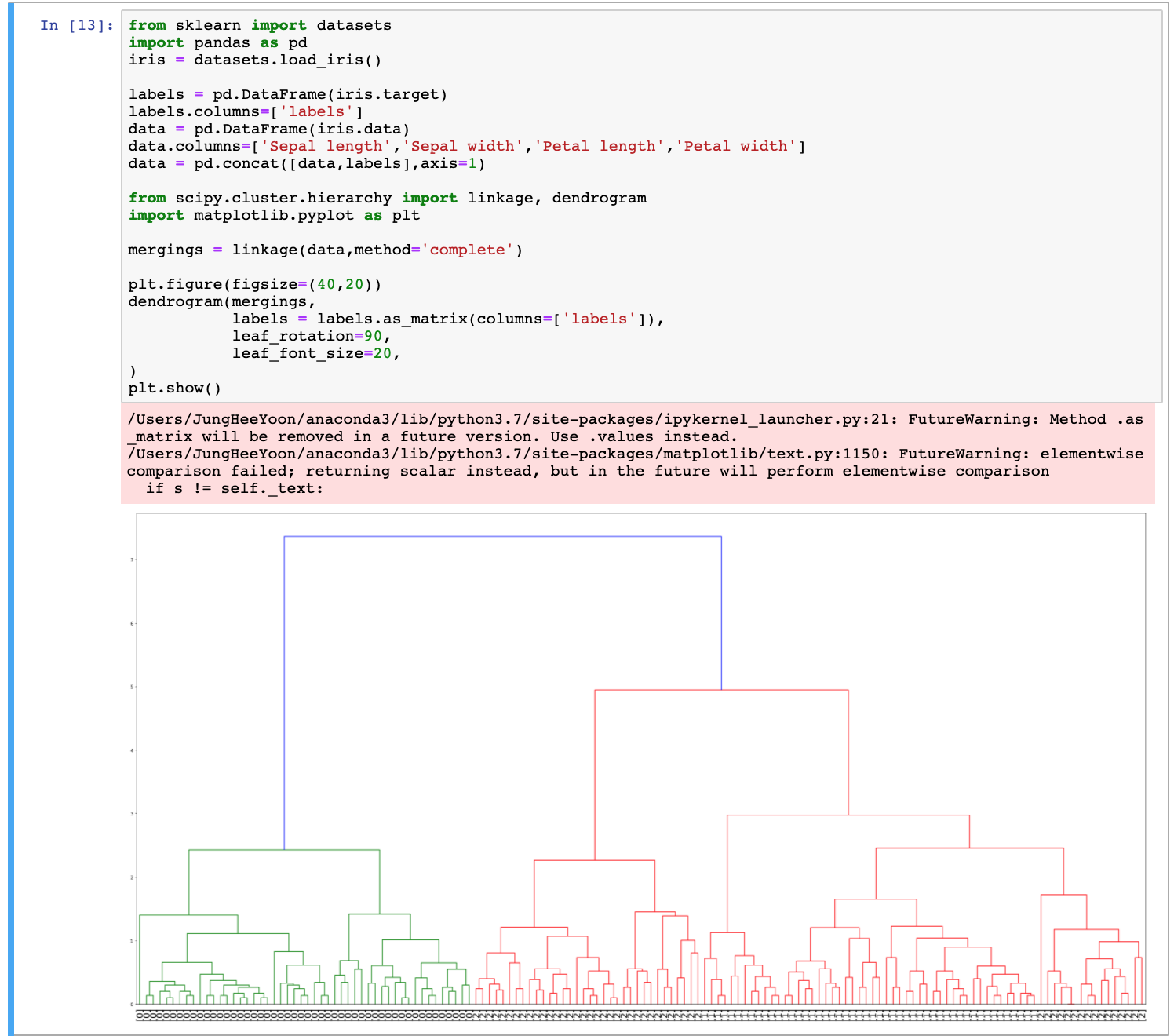
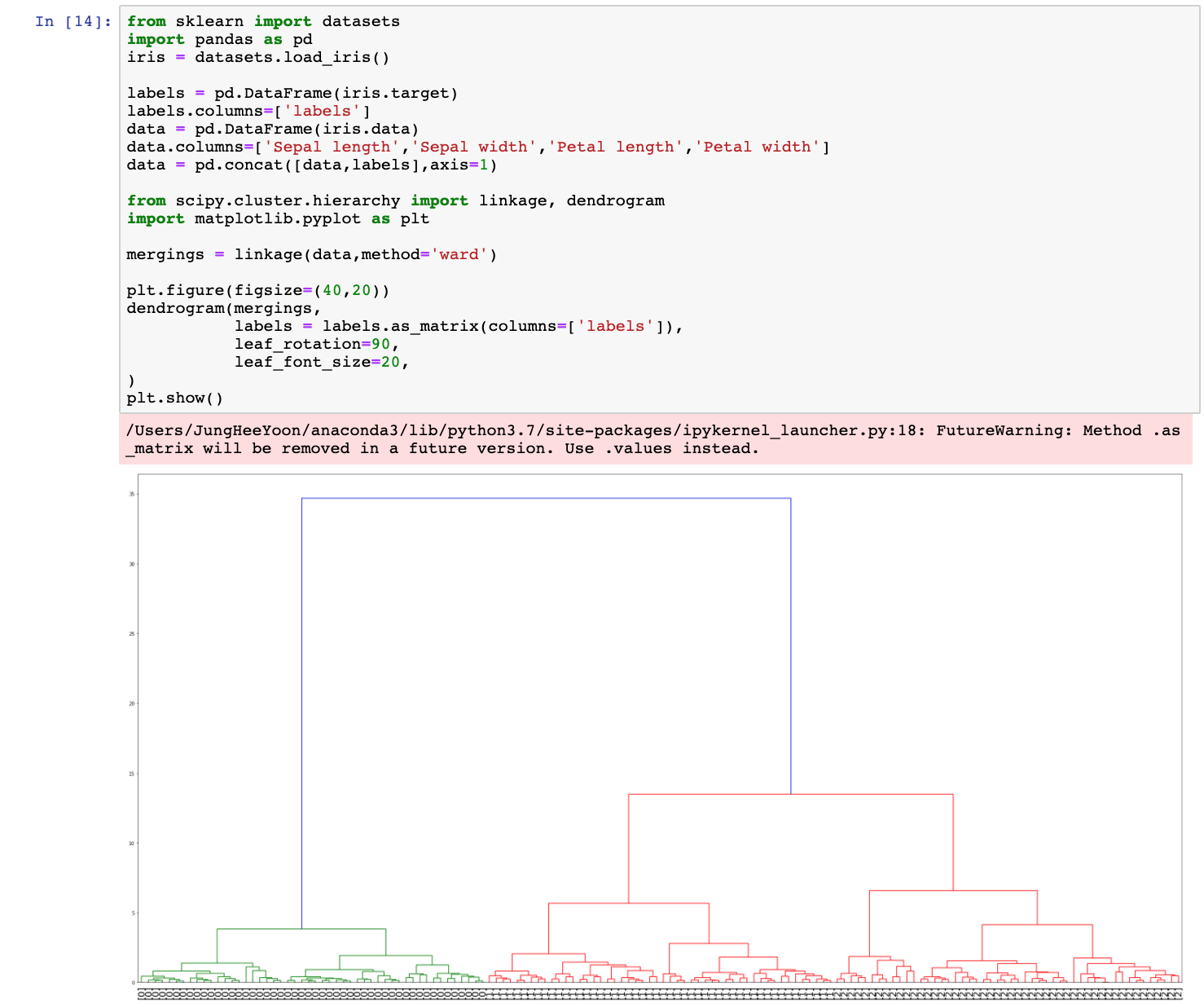
2016312568 정희윤

1. digit데이터\_클러스터링  
   load\_digits 데이터셋에서 랜덤으로 20개의 데이터를 가져온다.

linkage와 dendrogram을 이용하여 위와 같이 군집화 시킬 수 있다.



1. IRIS데이터\_클러스터링  
   IRIS 데이터를 KMeans를 사용하여 학습을 시킨다. 위는 클러스터의 갯수가 3개인 경우이다.  
   위는 클러스터 수에 따른 inertia 변수의 변화를 나타낸 그래프이다. 이를 통하여 적절한 클러스터의 갯수를 알 수 있는데 그래프의 기울기의 절대값이 작아질수록 의미가 없어진다는 의미이다. 이를 잘 판단하여 클러스터 수를 정해야 한다. 위 경우에는 클러스터 수가 5 일때 가장 적절할 것같다.  
   아래 그림은 클러스터 수가 5일때의 학습결과이다.

아래는 Iris dataset으로 hierarchical clustering한 결과이다.   
 다음은 linkage method = complete로 했을경우이다.  
  
  
  
다음은 linkage method = word로 했을경우이다.