DBSCAN Report

2016024893 오성준

1. **Summary of your algorithm**
2. **Detailed description of your codes (for each function)**
3. **Instructions for compiling your source codes at TA's computer**

**1) requirements.txt**

이번 과제부터 python 가상 환경에서 진행하여 **requirements.txt** 파일을 첨부해서 함께 git lab에 올립니다. 보고서 3번 내용 자체를 지금에서야 이해하게 된 점에 대해 죄송합니다. 동시에 보고서에 사진도 첨부합니다.

텍스트이(가) 표시된 사진

자동 생성된 설명

**2) 실행 명령어**

파일의 실행은 명세와 똑같이 구현했습니다.

텍스트이(가) 표시된 사진

자동 생성된 설명

**3) 채점**

채점 시스템이 잘 작동하지 않아서 친구에게 부탁해서 채점을 해보았습니다. 잘 채점이 된 모습입니다.

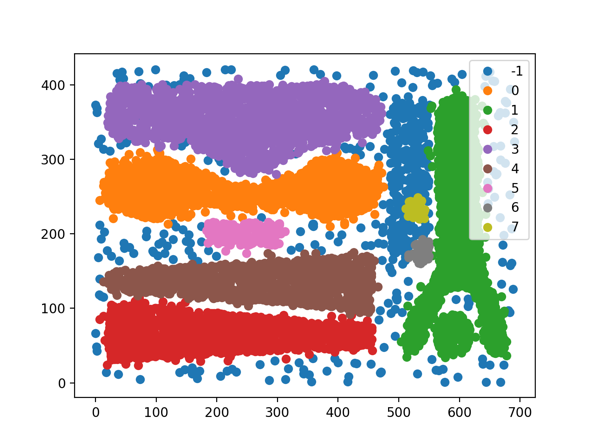
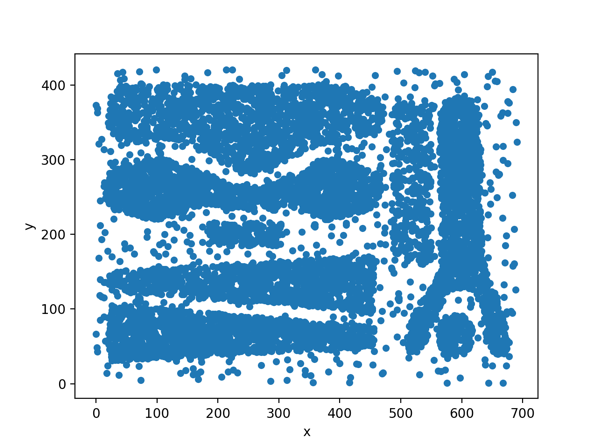
텍스트이(가) 표시된 사진

자동 생성된 설명

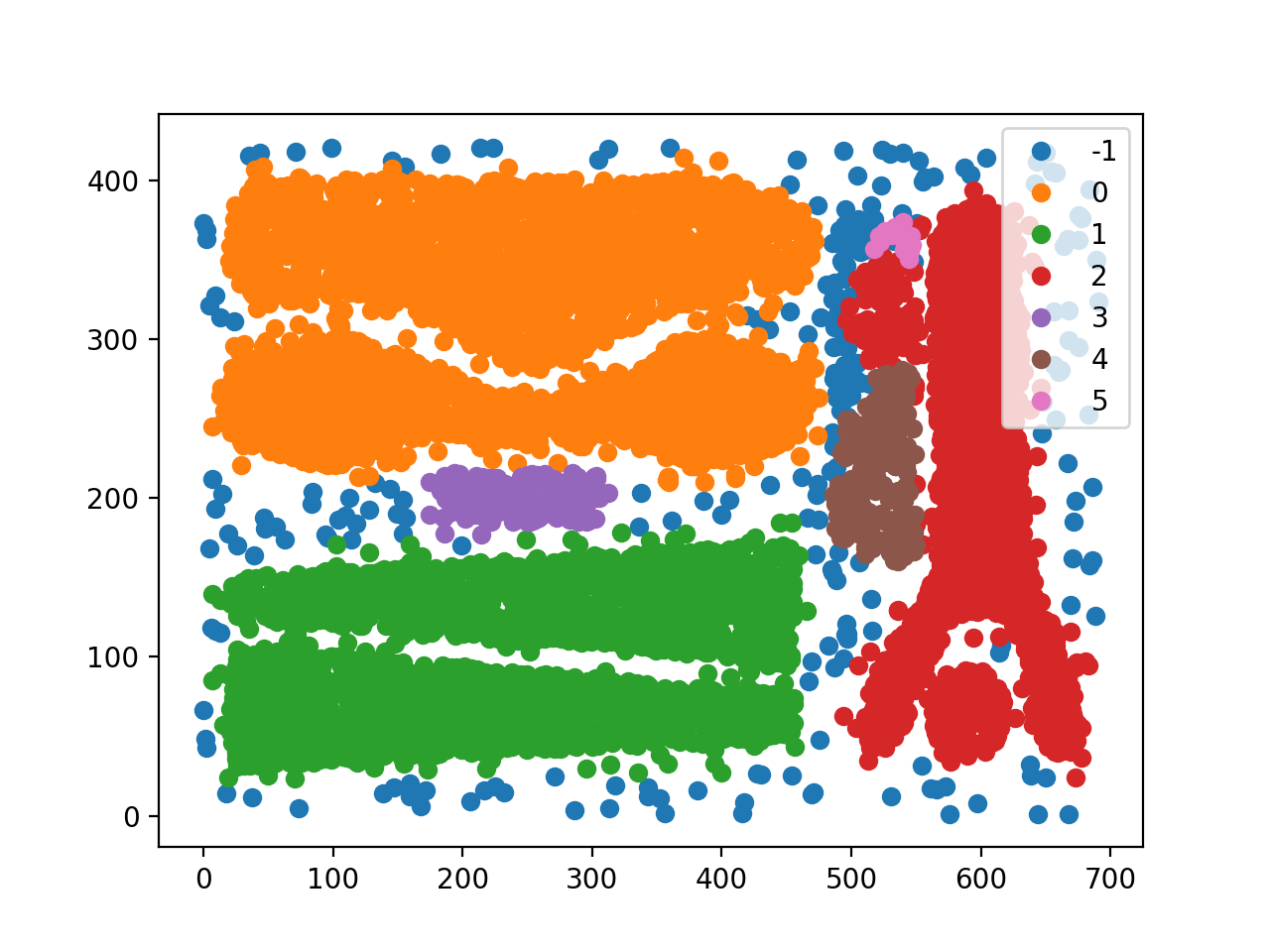
1. **Any other specification of your implementation and testing**

* 이번 과제를 진행하면서 data들을 시각화 하고 싶다는 생각이 들어서 **show.py**에 matplotlib을 이용해서 scatter로 점들을 시각화 했습니다. 동시에 clustering 이후에 색깔을 붙여서 다시 시각화 해서 전 후를 비교할 수 있도록 했습니다. 아래에 결과물을 첨부합니다.

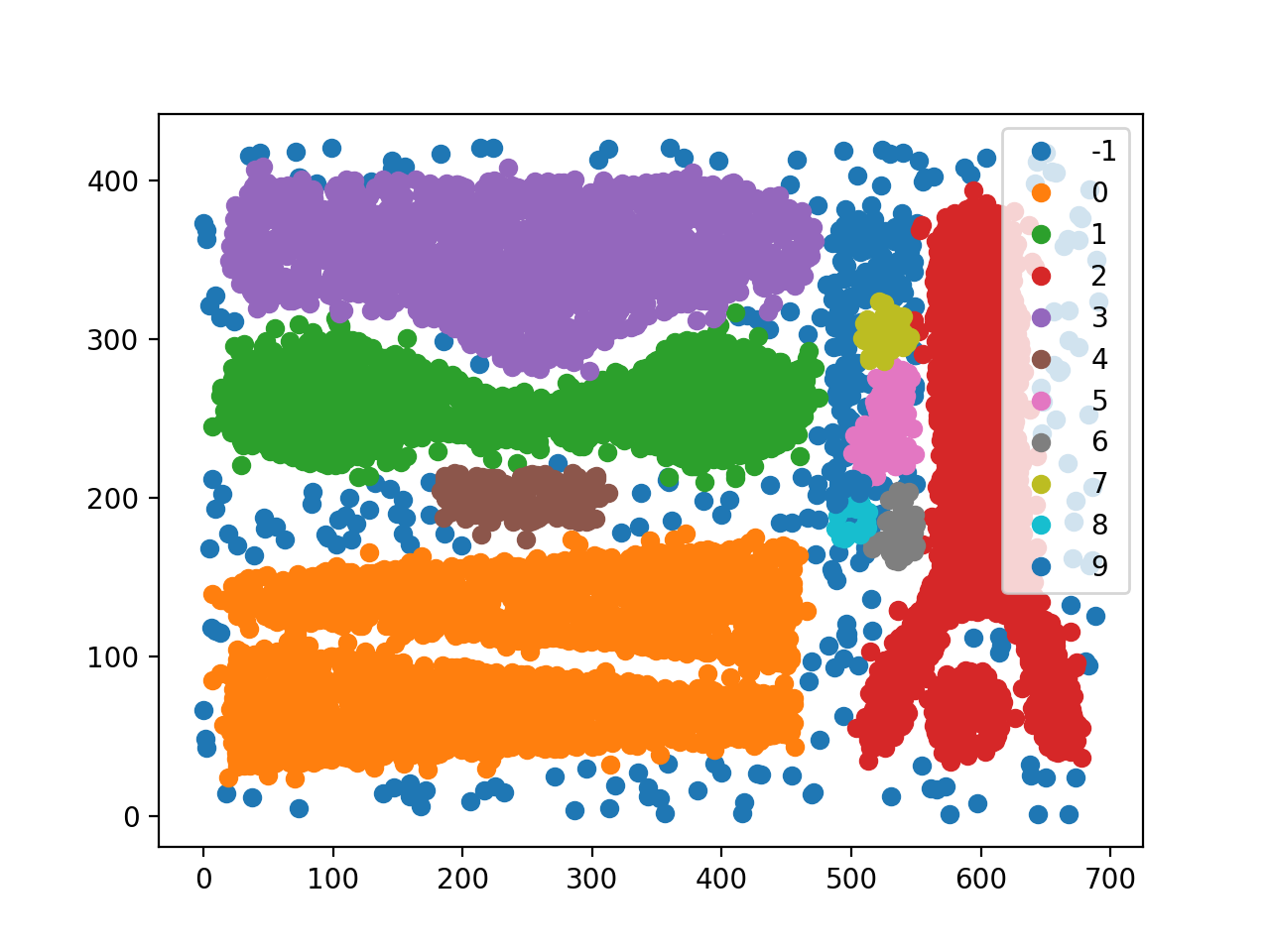
1. **Input1.txt**



* 아래로 긴 직사각형 모양의 cluster가 제대로 잡히지 않았는데, 우선 결과값은 대략 99점으로 제대로 나오긴 했습니다.
* Eps 값을 늘리거나 min points 값을 줄이면 cluster에 포함되곤 했지만, 단독으로 존재하기 보단, 주변의 cluster와 함께 묶이는 상황이 나타나곤 했습니다.
  + Eps 값이 17, min points 값이 22일 때,

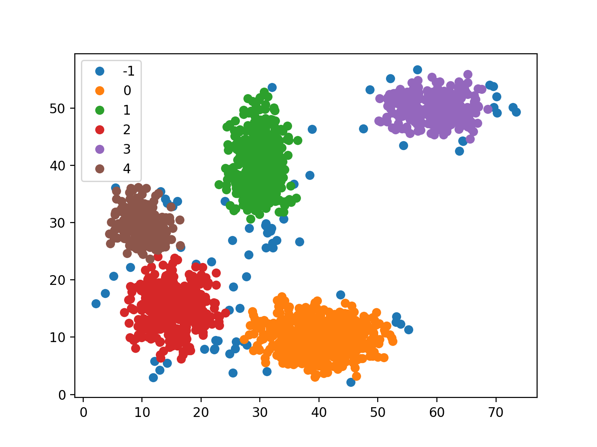
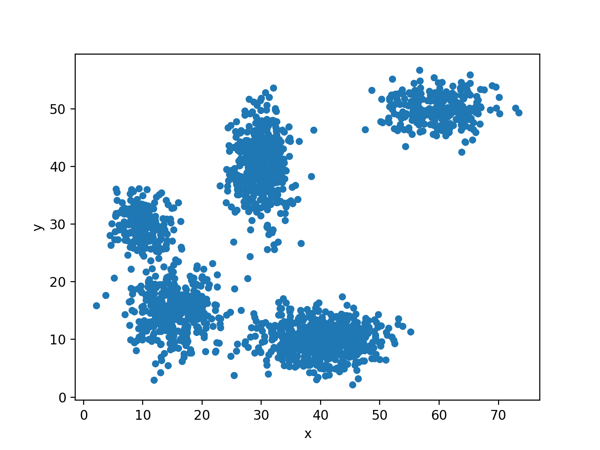


* + - Cluster의 수가 전체적으로 줄었고, 더 많은 점이 cluster에 포함된 모습이지만 결과가 만족스럽진 못하다고 생각합니다. 그 이유는 구분되어야 할 cluster가 합쳐졌기 때문입니다.
  + Eps 값이 15, min points 값이 20일 때,

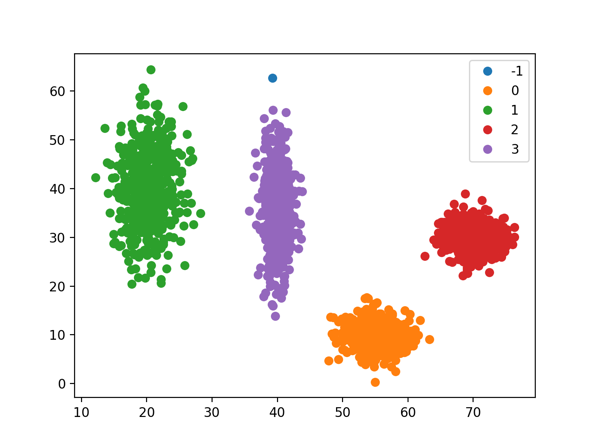
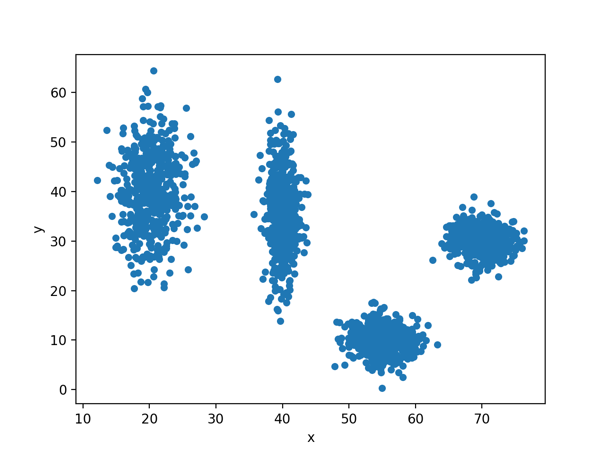


* + - 확실히, 아래로 긴 직사각형 부분에서 clustering되는 부분이 많아졌지만, 그럼에도 불구하고 각자가 쪼개진 모습입니다. 그리고 이번에는 주황색에 해당하는 부분이 합쳐졌기 때문에 마찬가지로 지금 시점이 이득이라고 볼 수 없을 겁니다.
* 이번 시도를 통해서 DBScan이 확실히 parameter에 민감하다는 것과, 시각적으로 완벽히 cluster라고 할 수 있는 것들이 제대로 clustering 되지 않을 수 있다는 것을 알 수 있었습니다.

1. **Input2.txt**



1. **Input3.txt**



* Input2와 input3에 관해선 특별한 점이 발견되지 않고 잘 clustering 됬다고 생각해서 다른 실험을 해보진 않았습니다.
* 전반적인 data의 분포와 결과를 분석할 수 있고, 여러가지 실험을 해볼 수 있어서 좋았습니다. Show.py의 구현 부분은 다음과 같습니다.

1. **이전 과제에 대해서**

* 앞서 말한 것처럼 저는 지금에서야 보고서 3번 문항에 대해서 알게 되었고, 모든 과제를 가상 환경에서 다시 실행했습니다. 그리고 현재의 requirements.txt 환경에서 잘 작동하는 것을 확인했습니다. 이전 과제들도 채점하실 때 참고 하셨으면 좋겠습니다.
* 2번째 과제에 대해서 말씀드릴 건 제가 구현한 dt.py 코드를 보면서 보고서를 쓰던 도중에 한글로 주석을 달아놓고 그대로 내는 바람에 아마도 dt.py가 잘 실행이 되지 않았을 수도 있을 것 같습니다.

# -\*- coding: utf-8 -\*-

이 주석을 가장 맨 위에 붙여놓으면 잘 작동하는 것을 확인했습니다. 참고해서 채점해 주셨으면 합니다.

* 여러모로 이런 부분들을 미리 확인하지 못하고 과제를 제출한 뒤에 뒤늦게 말하는 점 죄송합니다. 이번 과제와 recommender 과제는 주의해서 보고서를 잘 작성하도록 하겠습니다.