dbscan

# 라이브러리 링크

<https://github.com/scikit-learn/scikit-learn/blob/7e1e6d09b/sklearn/cluster/_dbscan.py#L164>

# 기초 설명

DBSCAN(Density-Based Spatial clustering of applications with noise)은 대표적인 밀도 방식의 클러스터링 기법이다. 군집 간의 거리를 이용하여 클러스터링을 하는 대표적인 알고리즘은 K-means와는 달리, 밀도를 호라용한 클러스터링은 점이 모여있는 영역을 합쳐나가는 방식이다.

# 버전 정보

* Numpy == 1.19.4 (pip install numpy)
* Matplotlib == 3.1.3 (pip install matplotlib)
* Pandas == 1.0.1 (pip install pandas)
* Sklearn == 0.22.1 (pip install sklearn)

# 데이터셋 설명 및 출처

* 쇼핑몰 고객 데이터 셋 (고객 수입과 소비 점수에 대한 정보)

<https://github.com/Daniel695/datasets>

# 코드 설명

* 위 링크에서 쇼핑몰 고객 데이터 셋을 다운받은 후 코드 파일과 같은 공간에 데이터 셋을 둔다.
* 주어진 고객 데이터 셋의 수입과 소비 점수를 사용하여 DBSCAN을 통해 clustering하는 코드.

# 검증 방법

* x