

# 머신러닝 - 군집기초

---

# 목차

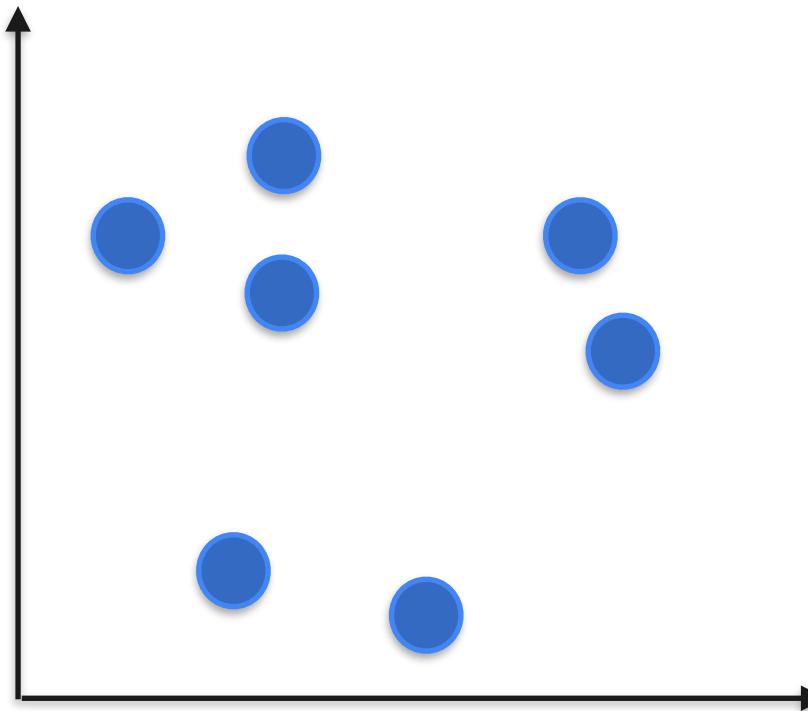
1 K-means 이해하기

2 모델만들기

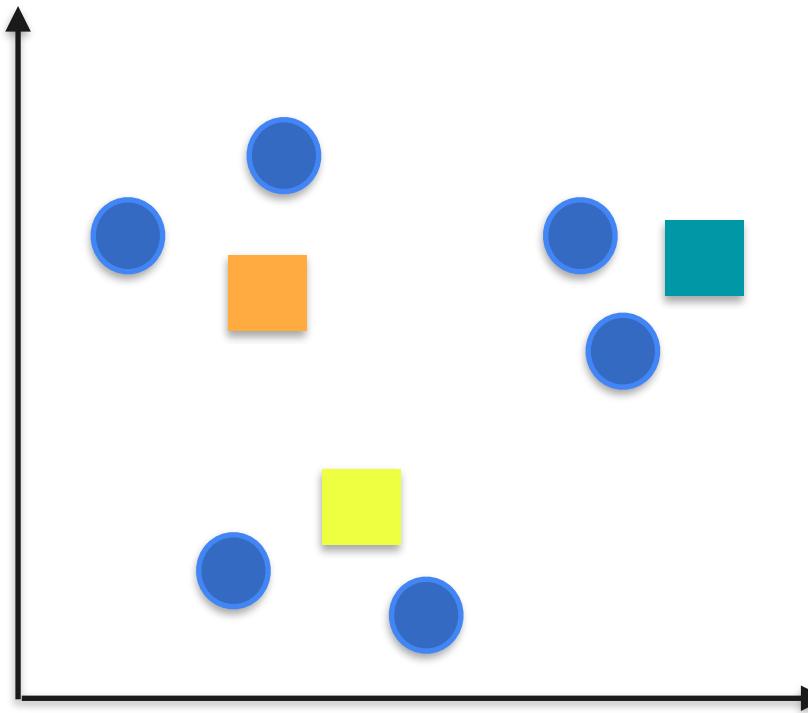
# K-means

- 비지도 학습 클러스터링 기법
  - 클러스터 개수( $k$ ) 사전 지정 필요
  - 랜덤하게  $k$ 개 centroid 초기화
  - 데이터 포인트 -> 가장 가까운 centroid에 할당
  - Centroid 업데이트 (할당된 데이터 중심점)
  - 반복하여 centroid 수렴까지
  - 거리 기반 클러스터링 (유clidean 거리 등)
  - 레이블 없는 데이터에 적용 가능 (사전 지식 불필요)
- 
- 장점: 간단, 직관적, 계산 비용 낮음 (대용량 데이터 적합)
  - 단점: 초기값 영향 받음, 구형 클러스터 가정

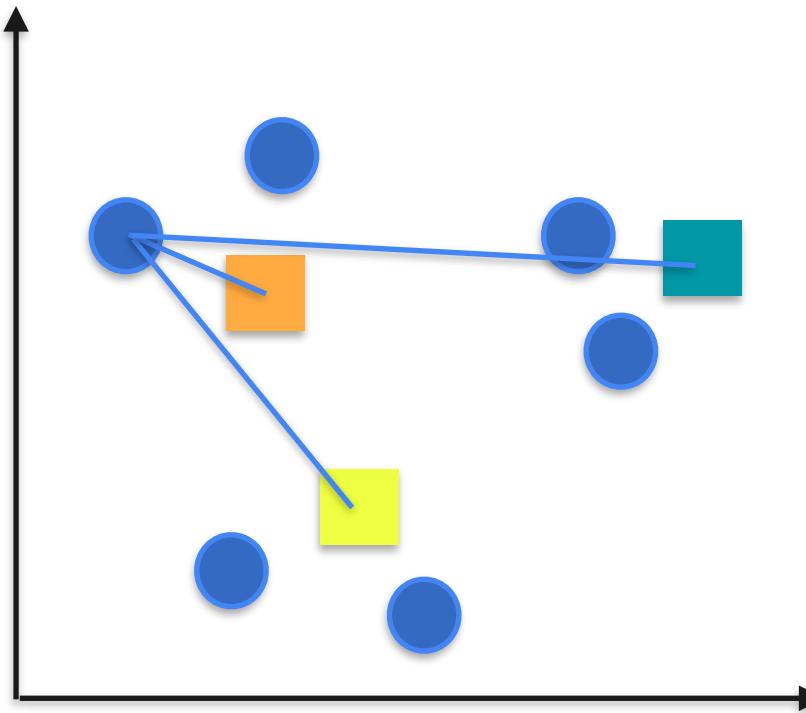
## 1. 데이터에서 군집의 개수(K) 설정하기



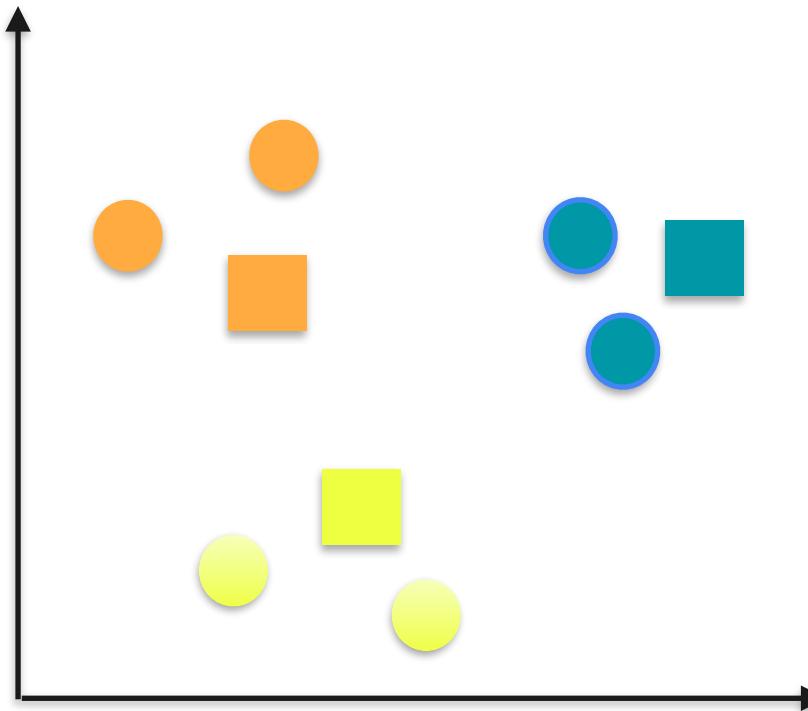
## 2. 초기 중심점 랜덤 설정하기



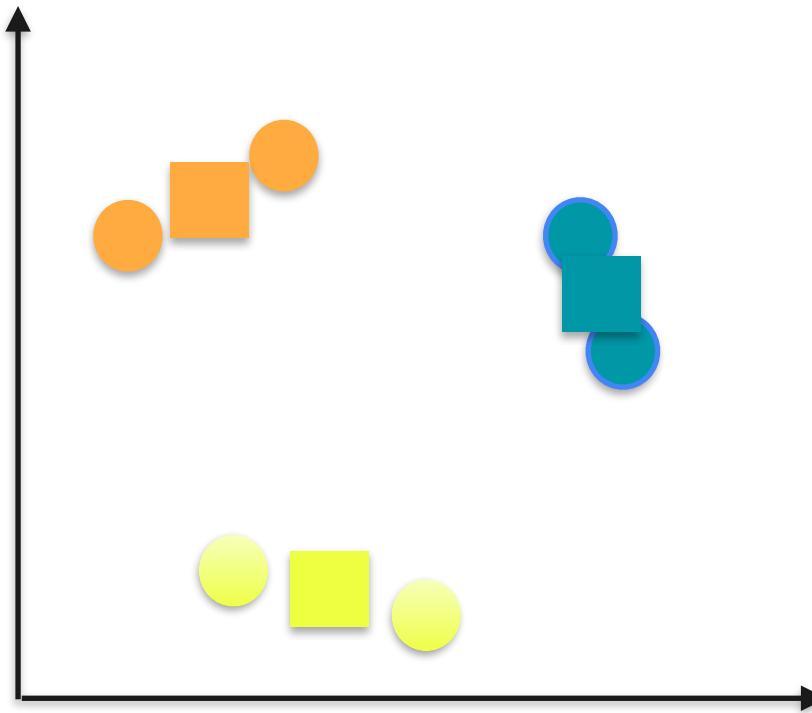
### 3. 데이터를 군집에 할당하기



### 3. 데이터를 군집에 할당하기



#### 4. 중심점 재설정



## 5. 데이터를 군집에 재할당

