

---

# Feature Selection

데이터 분석 스터디 2주차

**1. Correlation**

**2. Mutual Information**

**3. Shap value**

**4. Permutation Importance**

## Feature Selection

데이터의 특성 중에서 **가장 중요하게** 생각되는 변수를 선택하는 것  
Goal : 모델 훈련 과정에서 **가장 유용한** 특성을 선택하는 것

### 장점

- 차원의 저주를 줄임
- 모델의 성능을 높임
- 과적합 방지
- 데이터 분석에서 **중요하지 않은** 특성 제거

## 차원의 저주

차원이 증가하면서 **문제공간이 지수적으로 커지는** 현상

- 용량이 커진다 → **필요한 메모리 큼**
- 데이터보다 변수 수가 많아짐 → **모델의 성능 저하**

차원이 커질 수록 공간이 많이 필요하면서 사용할 수 있는 **정보량은 상대적으로 작아지는** 현상

