# 데이터 타입 조정

각자의 데이터를 가지고 활용해 보려면 몇가지 문제에 직면 하게 됩니다. 그 문제들을 미리 예상해서 그 문제들을 해결할 수 있는 방법을 알아보겠습니다. 원핫인코딩을 할때, 변수의 데이터 타입때문에 발생하는 문제가 있고, 그문제에 해당하는 내용이

* 변수(컬럼) 타입 확인: 데이터.dtypes
* 변수를 범주형으로 변경:
  + 데이터['컬럼에'].astype('category')
* 변수를 수치형으로 변경:
  + 데이터['컬럼명'].astype('int')
  + 데이터['컬럼명'].astype('float')

이만큼 이고요. 그다음에 데이터 안에 내용 중에 N/A값 때문에 발생하는 문제가 있습니다.

* NA 값의 처리
  + NA 갯수 체크: 데이터.isna().sum()
  + NA 값 채우기: 데이터['칼럼명'].fillna(특정숫자)

두가지만 여러분이 대응할 수 있으면 대부분의 표형식의 데이터를 읽어 들여서 딥러닝 모델에 넣어서 학습시키는데 문제가 없을 것입니다.

# 

## 