**Pandas Basics**

**1) 개요**

★pandas의 장점

1. 차원에 대한 장점

2. numpy기반으로 만들어짐. 데이터 분석 도구도 포함

-> 기본적인 numpy함수는 다 지원함

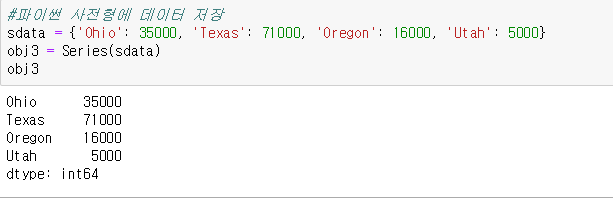
PANDAS=Panel+Data

**2) Series**

-일련의 객체를 담을 수 있는 1차원 배열 같은 자료 구조

-색인이라고 하는 배열의 데이터에 연관된 이름을 가짐

- 사전형 데이터 형태를 지원함

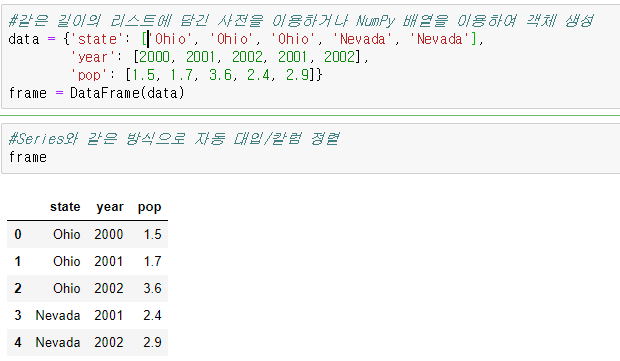


**3) DataFraame**

- 여러 개의 데이터 타입을 column별로 담은 데이터!

- 각 column을 feature, variable 등으로 볼 수 있다.

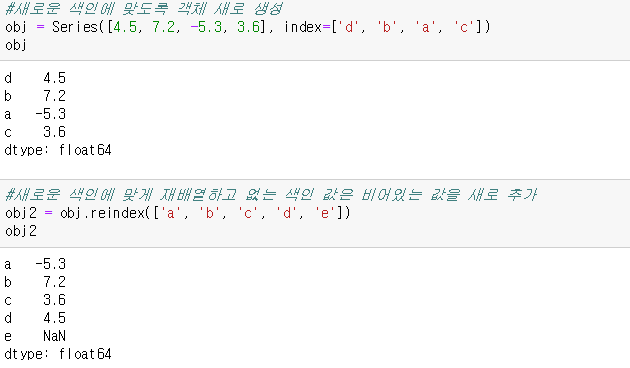
- 데이터베이스도 이와 마찬가지의 형태를 보인다. (row = record, col = feature)



\*\*\*Index나 숫자를 통해 접근가능한 Series 혹은 R과 다르게 pandas. DataFrame에선 column 이름/속성형식으로만 접근할 수 있다

**5) 주요함수**

- Re-indexing

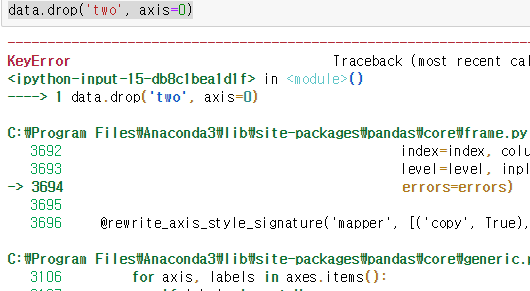


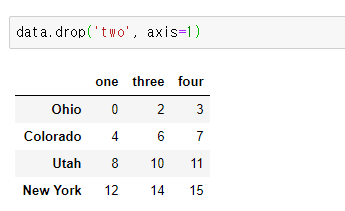
**-drop**



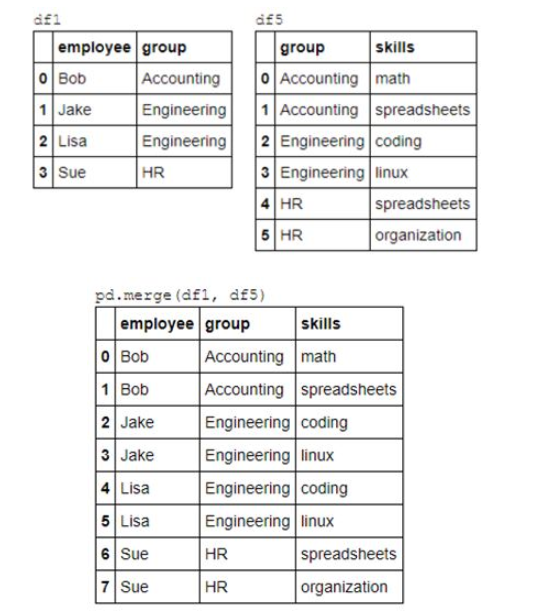
axis값은 default로 0 (행을 지우는 경우에는 axis를 입력하지 않아도 됨)

반면 column을 지우는 경우, axis값을 안 넣으면….

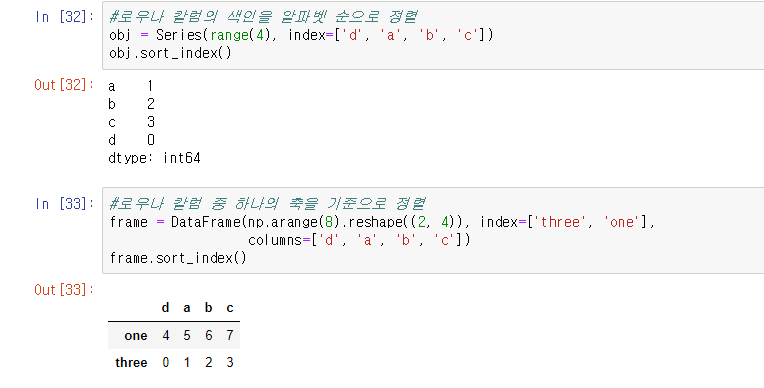




**-merge**

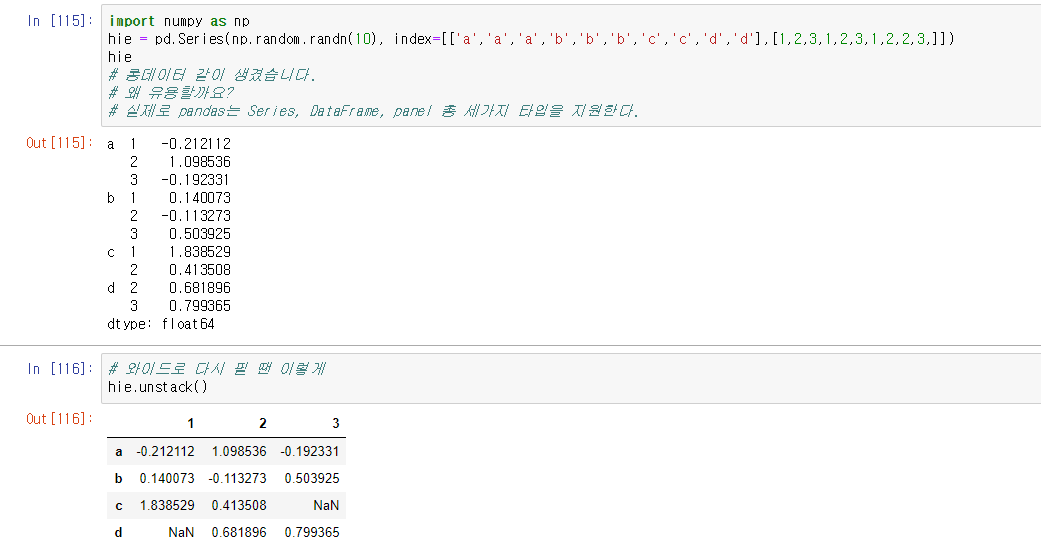


**-sort**



6) 주요특징

**-Indexing**



**-다른 데이터 포맷간의 계산**

e.g) series와 dataframe 간의 계산

