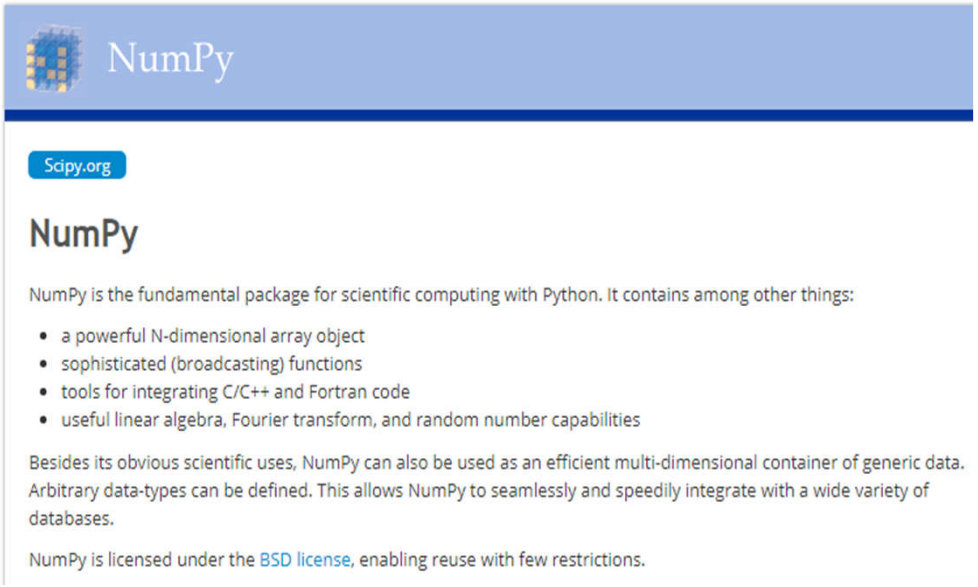
데이터 분석 패키지

파이썬 주요 데이터분석 패키지들 소개

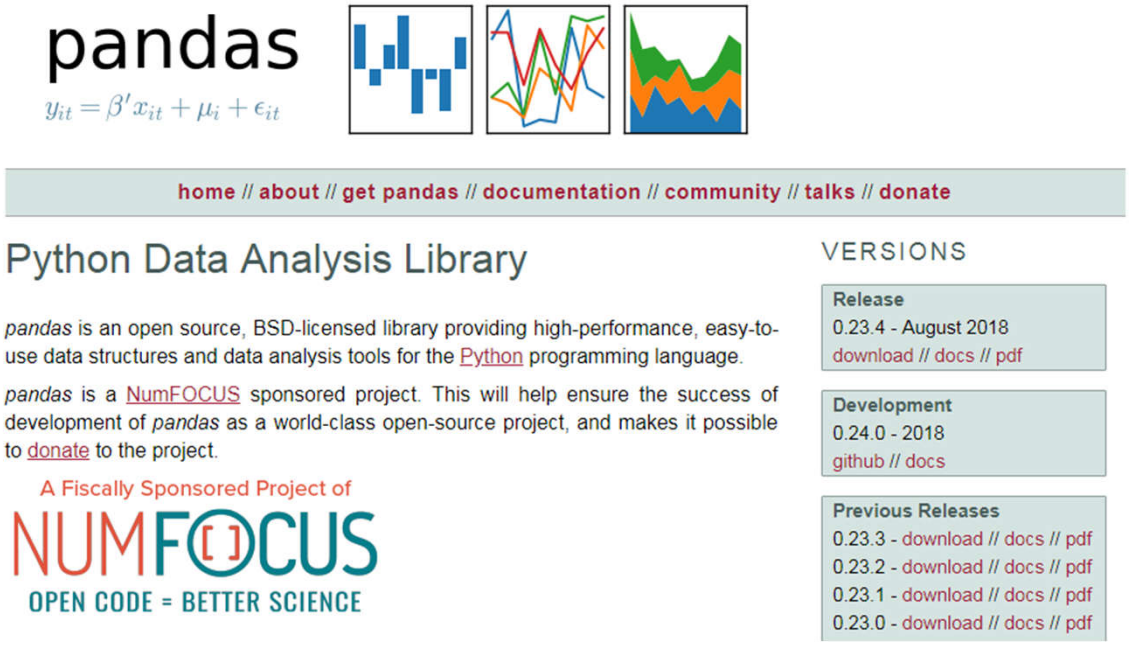
주요 패키지 소개

- Numpy
  - http://www.numpy.org/
  - C로 구현된 고성능 수치계산을 위한 라이브러리.
  - 고성능 과학연산을 위한 패키지로 데이터분석, 머신러닝등에 필수로 사용된다
  - 벡터, 행렬 연산관련 다양한 기능을 제공.
  - Pandas, matplotlib, scikit-learn의 기반이 되는 패키지.



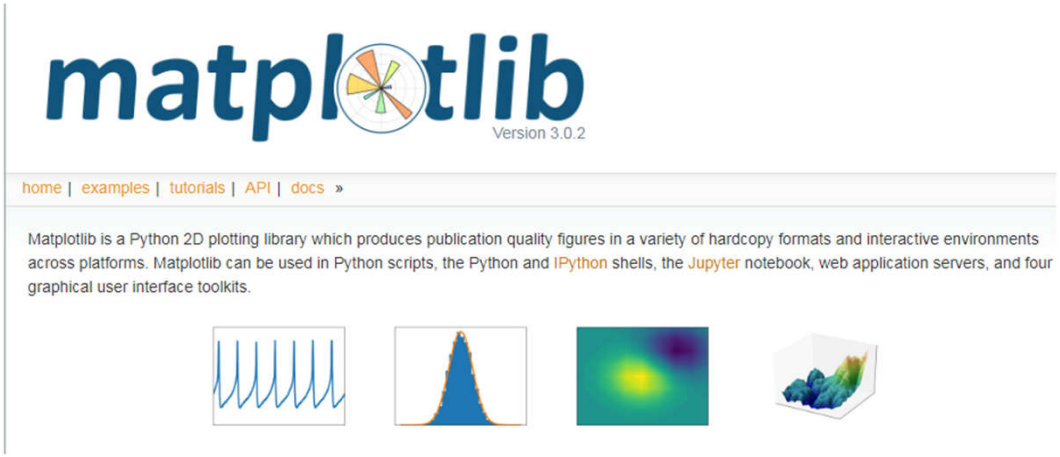
주요 패키지 소개

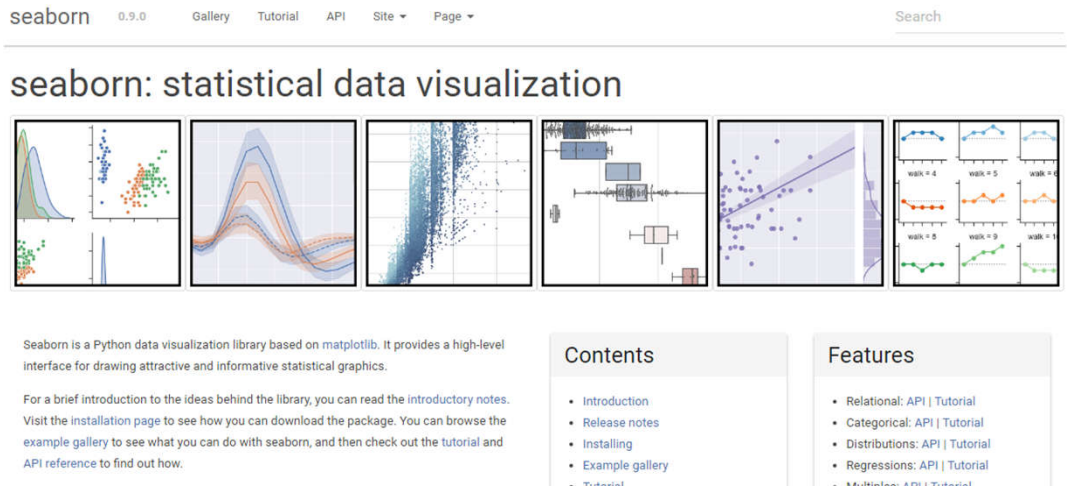
- Pandas
  - 데이터 분석시 가장 많이 다루게 되는 표 형태의 데이터셋을 쉽게 다루는 기능을 제공하는 패키지.
  - csv나 excel파일을 읽어 분석하는데 사용하는 다양한 기능을 제공.
  - https://pandas.pydata.org



주요 패키지 소개 – 시각화 패키지

- matplotlib
  - 파이썬 시각화의 기본이 되는 패키지
  - https://matplotlib.org
- Seaborn
  - matplotlib를 기반으로 하는 시각화 패키지
  - matplotlib보다 좀더 쉽게 그래프를 그릴 수 있도록 한다.
  - https://seaborn.pydata.org





주요 패키지 소개

- scikit-learn
  - 머신러닝(기계학습) 관련 기능을 제공하는 패키지.
  - https://scikit-learn.org