

[데이터 시각화]

1. 시각화: 원본 데이터나 분석된 결과 데이터를 그래프로 표현하는 것

2. **matplotlib** cf) 유용한사이트: https://matplotlib.org/stable/api/pyplot_summary.html

: 파이썬에서 가장 많이 사용하는 시각화 라이브러리로, 판다스에서도 내부적으로 사용한다.

(1) 패키지 임포트 및 기본 설정

```
import matplotlib.pyplot as plt # import
%matplotlib inline # 주피터 노트북이 실행되는 브라우저에서 시각화
%config InlineBackend.figure_format='retina' # 그래프의 해상도를 높임 (기본값: 'png')
```

(2) 그래프 객체

- Figure 객체: Matplotlib 에서 그래프를 그리기 위한 객체 / **plt.figure()**

(3) 그래프 영역 나누기

- 1) **subplot()** : 서브플롯 추가
- 2) **subplots()** : 서브플롯 집합 추가

(4) 그래프 그리는 함수들

plot()	선(lines)과 점(marker)으로 주어진 x, y 값을 표시
scatter()	산점도 그래프 plt.scatter(x= , y=)
bar() / barh()	수직 막대그래프 / 수평 막대그래프
axvline() / axhline()	수평선 / 수직선
hist()	히스토그램(도수분포표)
boxplot() / violinplot()	박스플롯 / 바이올린 플롯
fill() / fill_between()	그래프 영역을 색으로 채우기

(5) 그래프 커스터마이징

- 1) **set(linestyle= , linewidth=)**: 선 스타일 / 선의 굵기 지정
- 2) **text() / annotate()**: 텍스트 삽입 / 텍스트를 삽입하여 화살표로 위치 표시
- 3) 축과 눈금

set_xlim() / set_ylim	x 축/ y 축의 범위 지정
set_xlabel() / set_ylabel()	x 축/ y 축의 제목 지정
set_xticks() / set_yticks()	x 축/ y 축의 눈금 위치를 리스트 형식으로 지정
set_xticklabels() / set_yticklabels()	x 축/ y 축의 각 눈금의 이름을 리스트 형식으로 지정
(축 공유) twinx() / twiny()	x 축 공유, y 축 오른쪽에 표시 / y 축 공유, x 축 위쪽에 표시

- 4) **legend()**: 범례 표시