

Matplotlib을 이용한 데이터 시각화

송실대학교
베어드교양대학
강의선 교수
iami86@ssu.ac.kr

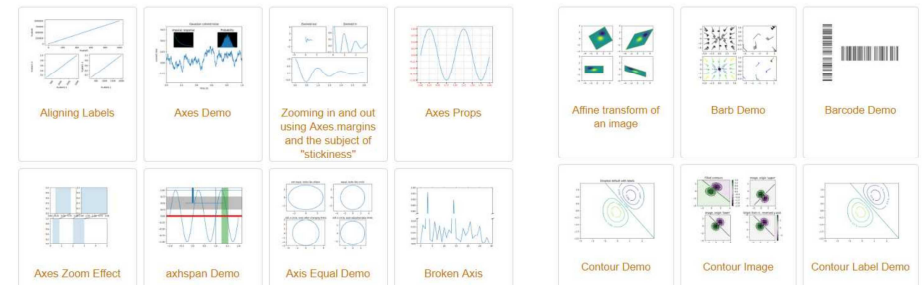
matplotlib으로 간단한 차트 그려보기

matplotlib이란?

- matplotlib(mat + plot + lib)
 - 데이터를 차트나 Plot등 다양한 형태로 그려주는 데이터 시각화 패키지
 - 설치 : `pip install matplotlib`

<https://matplotlib.org/>

matplotlib
version 3.4.2



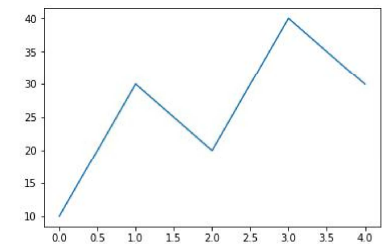
2

간단한 Plot 차트 그리기

```
In [1]: import matplotlib.pyplot as plt
```

```
In [2]: values = [10, 30, 20, 40, 30]
```

```
In [3]: plt.plot(values)
```



4

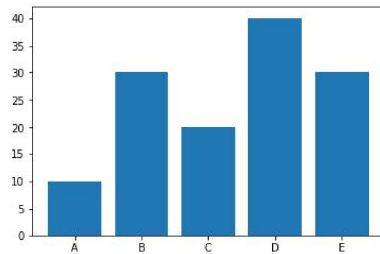
간단한 Bar 차트 그리기

```
In [1]: import matplotlib.pyplot as plt
```

```
In [5]: items = ['A', 'B', 'C', 'D', 'E']  
values = [10, 30, 20, 40, 30]
```

```
In [6]: plt.bar(items, values)
```

```
Out [6]: <BarContainer object of 5 artists>
```



5개의 상품의 판매량을 비교한다고 해보자.

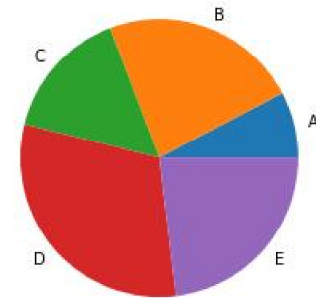
두 개의 리스트를 생성한 후

bar 함수에 입력해주면 된다.

5

간단한 Pie Chart 만들기

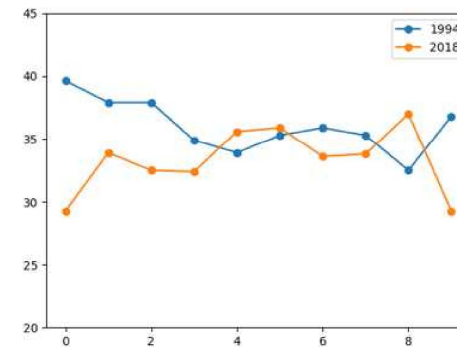
```
1 import matplotlib.pyplot as plt  
2  
3 items = ['A', 'B', 'C', 'D', 'E']  
4 values = [10, 30, 20, 40, 30]  
5  
6 plt.pie(values, labels=items)  
7 plt.show()
```



6

단계별로 옵션 배우기

주제: 최고기온 비교하기



8

공공데이터 가져오기

http://www.weather.go.kr/weather/climate/past_tendays.jsp

국내기후자료 | 과거자료

일별자료 | 요소별자료 | 순별자료

지점 | 서울(유) | 선택 | 년도 | 1994 | 선택 | 월 | 8 | 선택 | 시기 | 상순 | 선택

108 서울(유) / 1994년 8월 상순

요소		08월 01일	08월 02일	08월 03일	08월 04일	08월 05일	08월 06일	08월 07일	08월 08일	08월 09일	08월 10일	순계	순평균
평균기온 (℃)	관측	26.3	28.7	28.9	29.1	30.2	30.5	29.7	29.8	30.4	26.3	289.9	29.0
	평년	26.7	26.7	26.6	26.6	26.5	26.5	26.6	26.5	26.4	26.5	265.6	26.6
최고기온 (℃)	관측	29.3	33.9	32.5	32.4	35.6	35.9	33.6	33.8	37.0	29.3	333.3	33.3
	평년	30.6	30.6	30.6	30.6	30.5	30.6	30.7	30.6	30.5	30.6	305.9	30.6

9

Step1. 기본 차트 생성하기

```
import matplotlib.pyplot as plt

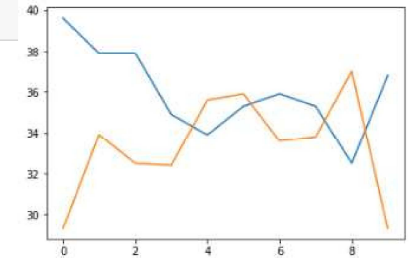
temp1994 = [39.6, 37.9, 37.9, 34.9, 33.9, 35.3, 35.9, 35.3, 32.5, 36.8]
temp2018 = [29.3, 33.9, 32.5, 32.4, 35.6, 35.9, 33.6, 33.8, 37.0, 29.3]

plt.plot( temp1994 )
plt.plot( temp2018 )

plt.show( )
```

입력이 귀찮은 경우 데이터를 3~4개만 우선 입력 테스트

한 번에 실행이 안되는 경우 (텍스트만 표시)
한 번 더 실행한다. (Shift + Enter)



10

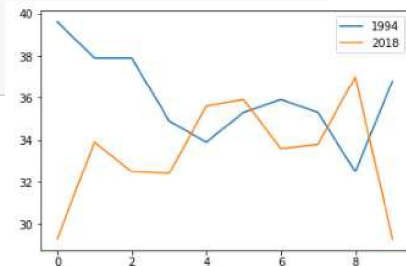
Step2. 범례 표시하기

```
import matplotlib.pyplot as plt

temp1994 = [39.6, 37.9, 37.9, 34.9, 33.9, 35.3, 35.9, 35.3, 32.5, 36.8]
temp2018 = [29.3, 33.9, 32.5, 32.4, 35.6, 35.9, 33.6, 33.8, 37.0, 29.3]

plt.plot( temp1994, label='1994' )
plt.plot( temp2018, label='2018' )
plt.legend()

plt.show( )
```



11

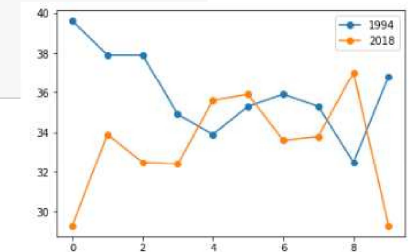
Step3. 마커 적용하기

```
import matplotlib.pyplot as plt

temp1994 = [39.6, 37.9, 37.9, 34.9, 33.9, 35.3, 35.9, 35.3, 32.5, 36.8]
temp2018 = [29.3, 33.9, 32.5, 32.4, 35.6, 35.9, 33.6, 33.8, 37.0, 29.3]

plt.plot( temp1994, marker='o', label='1994' )
plt.plot( temp2018, marker='o', label='2018' )
plt.legend()

plt.show( )
```



12

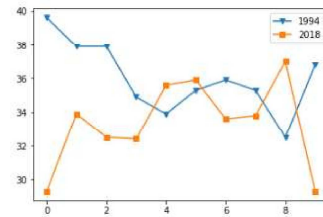
Step3. 마커 적용하기

여러 가지 마커

marker	사용
o	o
v	v
^	^
s	s
+	+
*	*

```
marker='v', label='1994')
marker='s', label='2018')
```

```
label='1994', marker='v' )
label='2018', marker='s' )
```



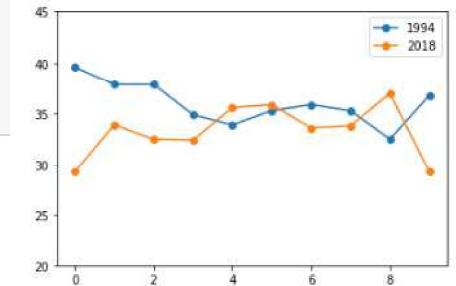
13

Step4. Y축 범위 변경하기 : y-axis limit

```
import matplotlib.pyplot as plt
temp1994 = [39.6, 37.9, 37.9, 34.9, 33.9, 35.3, 35.9, 35.3, 32.5, 36.8]
temp2018 = [29.3, 33.9, 32.5, 32.4, 35.6, 35.9, 33.6, 33.8, 37.0, 29.3]
plt.plot( temp1994, label='1994', marker='o' )
plt.plot( temp2018, label='2018', marker='o' )
```

```
plt.ylim(20, 45)
```

```
plt.legend()
plt.show()
```



14

Step5. X축 범위 변경하기 : x-axis limit

```
import matplotlib.pyplot as plt
temp1994 = [39.6, 37.9, 37.9, 34.9, 33.9, 35.3, 35.9, 35.3, 32.5, 36.8]
temp2018 = [29.3, 33.9, 32.5, 32.4, 35.6, 35.9, 33.6, 33.8, 37.0, 29.3]
plt.plot( temp1994, label='1994', marker='o' )
plt.plot( temp2018, label='2018', marker='o' )
```

```
plt.ylim(20, 45)
plt.xlim(0, 9) # 인덱스 0부터, 인덱스 9까지 표시
```

```
plt.legend()
plt.show()
```

Xlim 함수는 현재 불필요하므로 코드 주석 처리

15

Step6. 제목(Title) 표시하기

```
import matplotlib.pyplot as plt
temp1994 = [39.6, 37.9, 37.9, 34.9, 33.9, 35.3, 35.9, 35.3, 32.5, 36.8]
temp2018 = [29.3, 33.9, 32.5, 32.4, 35.6, 35.9, 33.6, 33.8, 37.0, 29.3]
plt.plot( temp1994, label='1994', marker='o' )
plt.plot( temp2018, label='2018', marker='o' )
plt.legend()
```

```
plt.ylim(20, 45)
plt.xlim(2, 8)
```

```
plt.title('temperature')
```

```
plt.title('temperature')
```

```
plt.title('최고 기온 비교')
```

```
plt.show()
```

위의 제목을 한글 제목 '최고 기온 비교' 로 변경해보자. (에러 발생)

16

Step7. 한글 제목(Title)

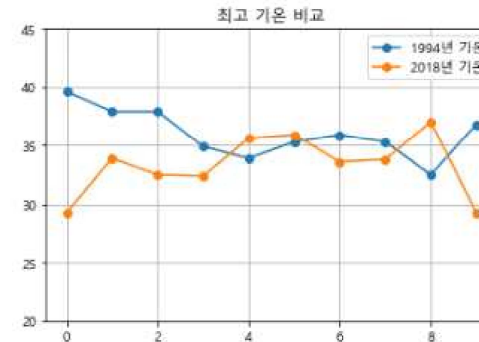
```
import matplotlib.pyplot as plt
temp1994 = [39.6, 37.9, 37.9, 34.9, 33.9, 35.3, 35.9, 35.3, 32.5, 36.8]
temp2018 = [29.3, 33.9, 32.5, 32.4, 35.6, 35.9, 33.6, 33.8, 37.0, 29.3]
plt.plot( temp1994, label='1994년 기온', marker='o' )
plt.plot( temp2018, label='2018년 기온', marker='o' )
plt.legend()
plt.ylim(20, 45)
```

```
plt.rc('font', family='Malgun Gothic')
plt.title('최고 기온 비교')
plt.show( )
```



박스로 표시될 경우 다시 한번 실행

Step8. 격자 표시하기



```
plt.grid()
```

```
plt.show()
```

```
plt.grid(True)
```

Step 9. 축 제목 표시하기

```
In [65]: import matplotlib.pyplot as plt
temp1994 = [39.6, 37.9, 37.9, 34.9, 33.9, 35.3, 35.9, 35.3, 32.5, 36.8]
temp2018 = [29.3, 33.9, 32.5, 32.4, 35.6, 35.9, 33.6, 33.8, 37.0, 29.3]
plt.plot( temp1994, label='1994년 기온', marker='o' )
plt.plot( temp2018, label='2018년 기온', marker='o' )
plt.legend()
plt.ylim(20, 45)
plt.rc('font', family='Malgun Gothic')
plt.title('최고 기온 비교')
plt.grid(True)
```

```
plt.xlabel('날짜 변화')
plt.ylabel('최고 기온')
```

```
plt.show()
```

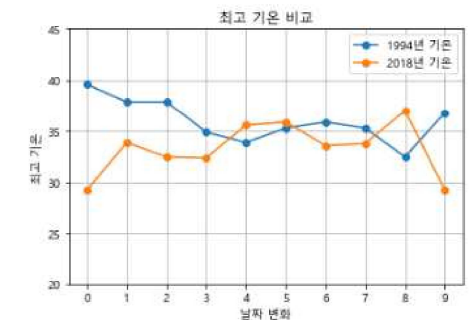
Step 10. x축 값의 표시 단위 : xticks 함수

```
plt.xlabel('날짜 변화')
plt.ylabel('최고 기온')
plt.xticks( np.arange(0, 10, 1))
plt.show()
```

아래와 같이 수정해보자. (1 에서 0.5로 변경)

```
plt.xticks( np.arange(0, 10, 0.5))
```

```
import matplotlib.pyplot as plt
import numpy as np
```

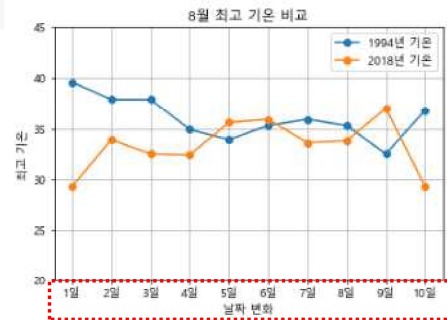


Step 11. x축 값의 표시 단위

```
plt.xlabel('날짜 변화')
plt.ylabel('최고 기온')

# plt.xticks( np.arange(0, 10, 0.5))
plt.xticks( np.arange(10), ('1일', '2일', '3일', '4일', '5일', '6일', '7일', '8일', '9일', '10일') )

plt.show()
```

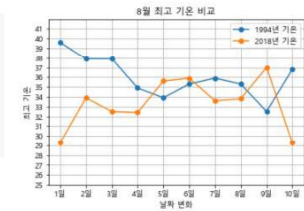


21

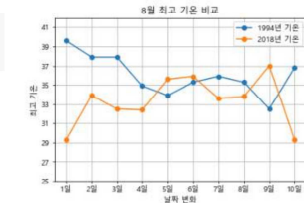
Step 12. y축 값의 표시 단위 : yticks 함수

```
plt.ylim(25, 42)
plt.yticks( np.arange(25, 42, 1))

plt.show()
```



```
plt.yticks( np.arange(25, 42, 2))
```



22

수고하셨습니다.