

day2

데이터의 시각화(matplotlib)

2019 경북고등학교

#1. matplotlib 그래프 알아보기

#2. matplotlib을 이용하여 다양한 그래프 그려보기

#3. [project] 우리나라 지역별 인구밀도 scatter graph로 나타내기
[project] 매년 내 생일의 서울 기온 변화 선 그래프로 나타내기

#1. matplotlib 그래프 알아보기



<https://matplotlib.org/>

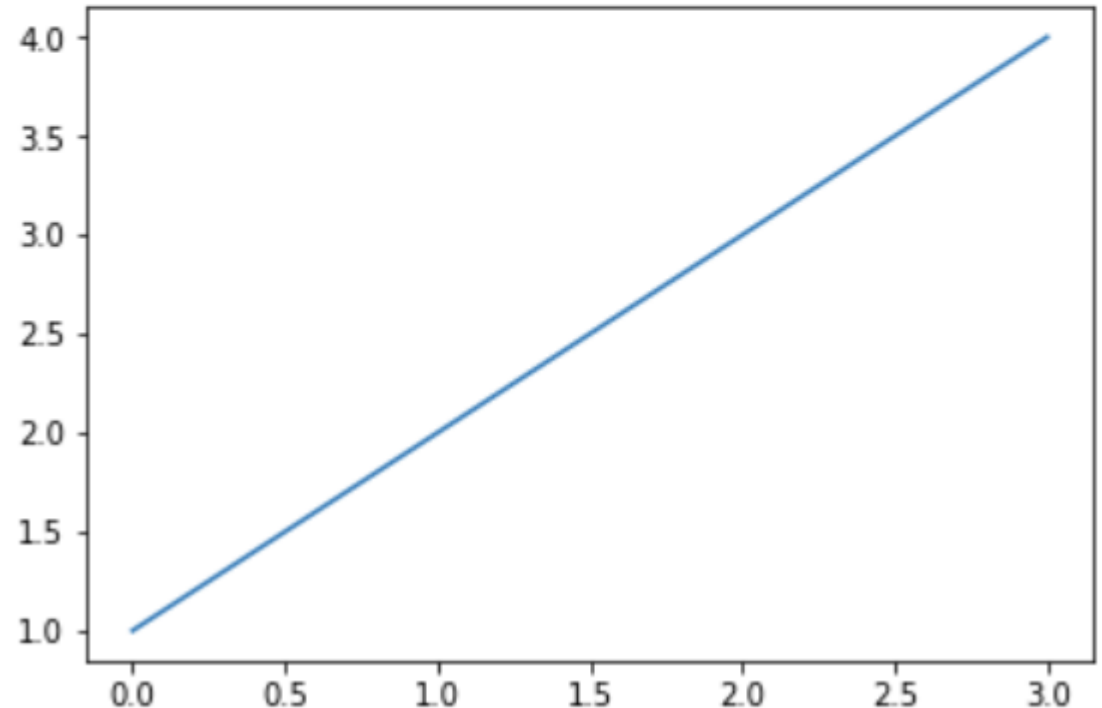
- python에서 그래프를 그리기 위한 라이브러리
- matplotlib.pyplot : matlab 스타일의 그래프를 그리기 위한 서브 라이브러리
- import matplotlib.pyplot as plt : plt(별칭)으로 matplotlib.pyplot을 импорт 함.

[선(plot) 그래프 그리기]

`plt.plot([x,],y[,fmt])`

1. 그래프 라이브러리 импорт
2. 그래프에 그릴 데이터 결정
3. 그래프 그리기
4. 그래프 보여주기

```
import matplotlib.pyplot as plt  
data=[1,2,3,4]  
plt.plot(data)  
plt.show()
```

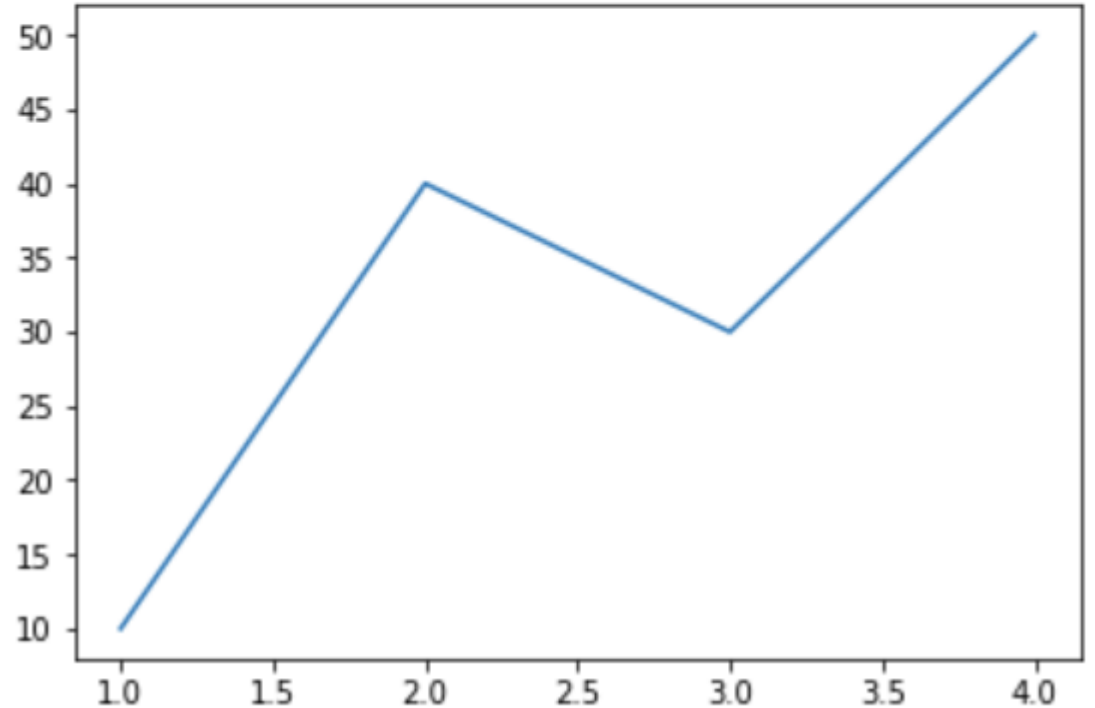


[선(plot) 그래프 그리기]

`plt.plot([x,],y[,fmt])`

1. 그래프 라이브러리 импорт
2. 그래프에 그릴 데이터 결정
 - 1) x축 데이터
 - 2) y축 데이터
3. 그래프 그리기
4. 그래프 보여주기

```
import matplotlib.pyplot as plt  
xdata=[1,2,3,4]  
ydata=[10,40,30,50]  
plt.plot(xdata,ydata)  
plt.show()
```



[그래프 꾸미기]

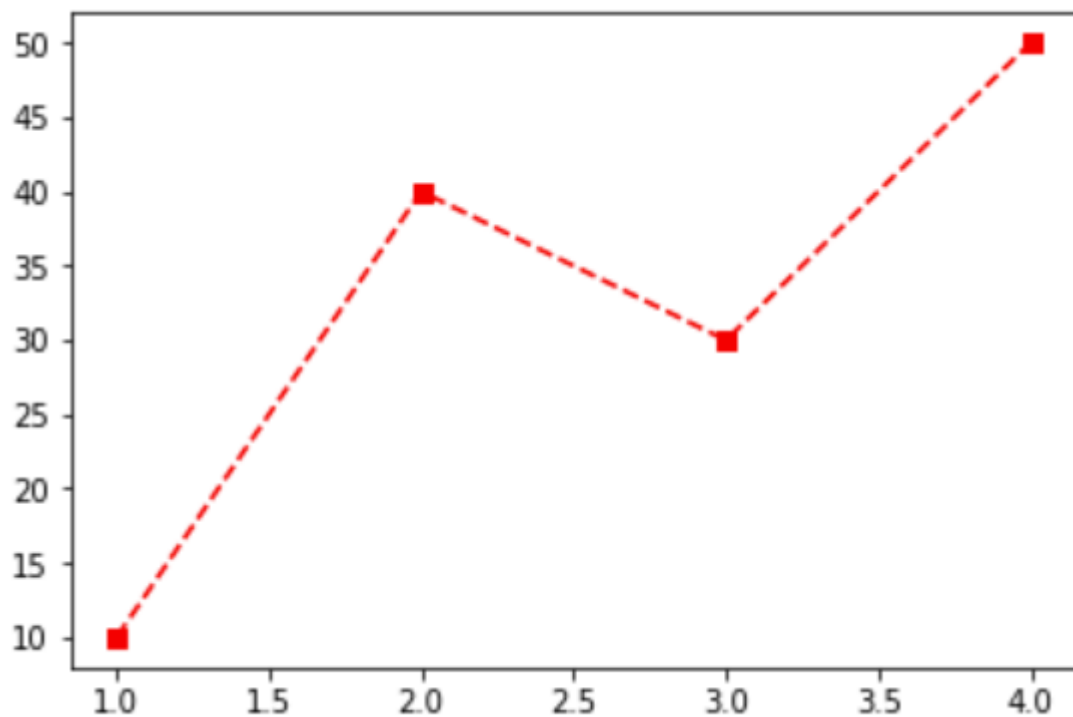
컬러 약어	컬러
b	blue
g	green
r	red
c	cyan
m	magenta
y	yellow
k	black
w	white

선스타일약어	선 스타일
-	실선(solid line)
--	파선(dashed line)
:	점선(dotted line)
-.	파선,점선 혼합(dash-dot line)

마커약어	마커
o	원모양
^,v,<,>	삼각형
s	사각형
*	별모양
+	플러스 모양
d	다이아몬드

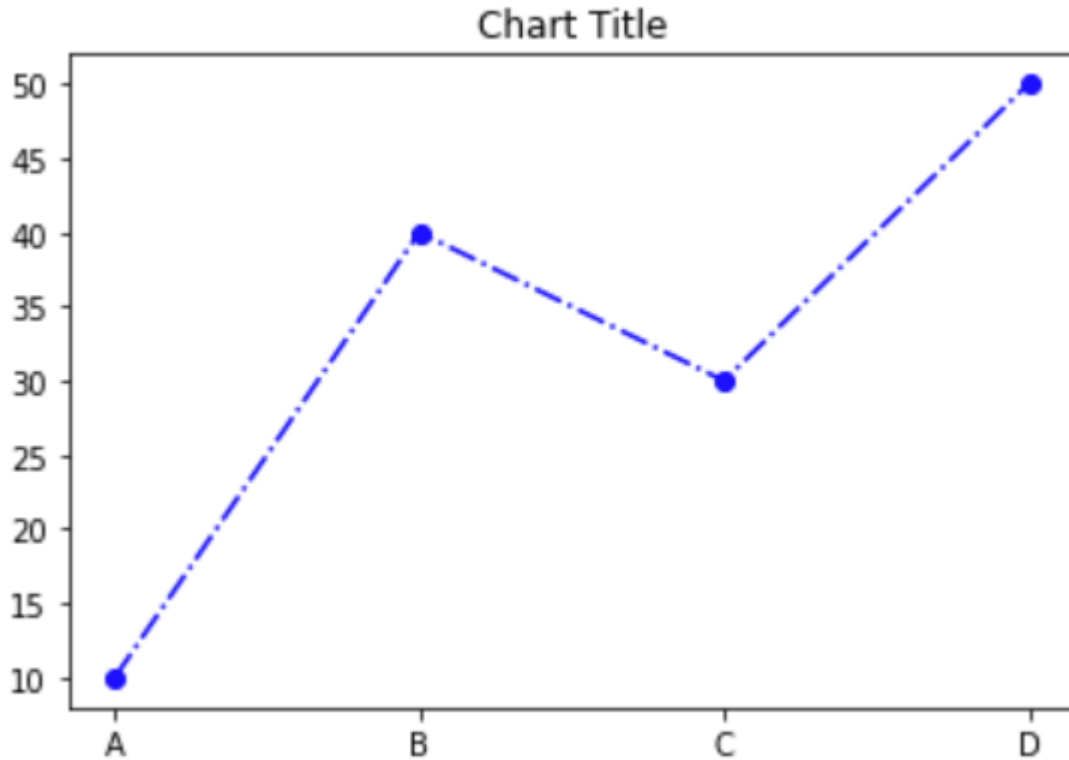
[그래프 꾸미기]

```
import matplotlib.pyplot as plt  
plt.plot([1,2,3,4],[10,40,30,50], 'rs--')  
plt.show()
```



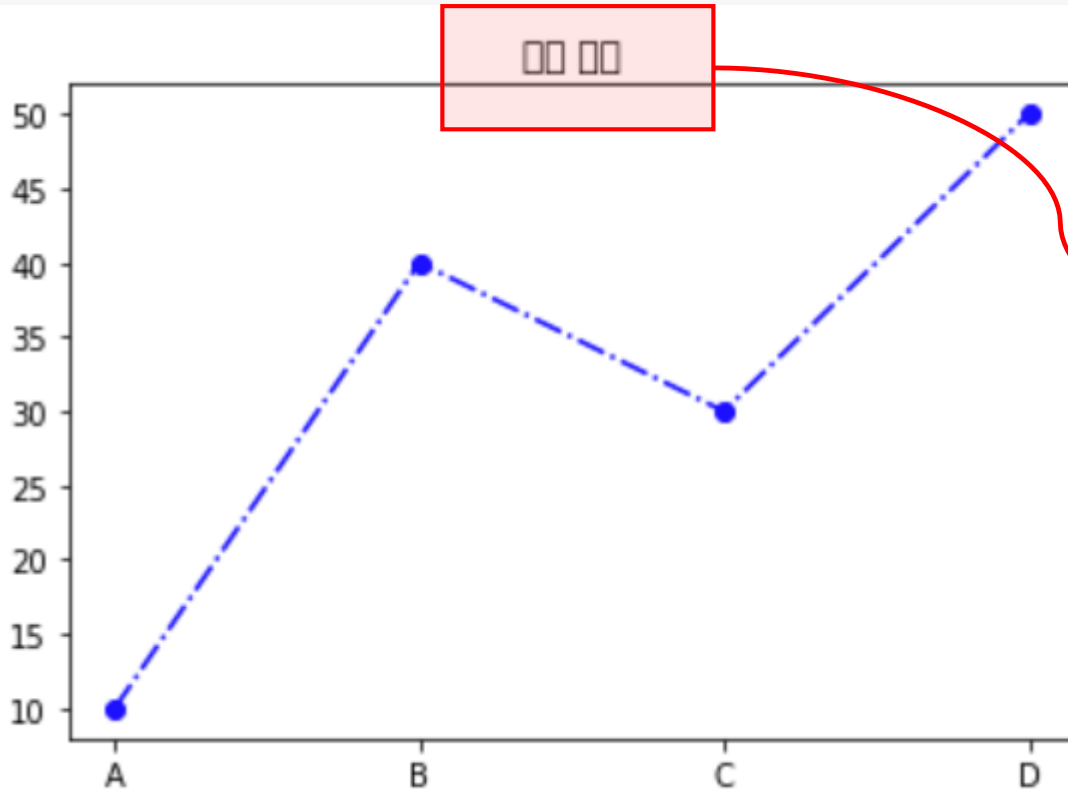
[그래프 제목 넣기]

```
import matplotlib.pyplot as plt  
plt.plot(['A', 'B', 'C', 'D'], [10, 40, 30, 50], 'ob-.')  
plt.title('Chart Title')  
plt.show()
```



[그래프 제목 넣기 - 한글 사용]

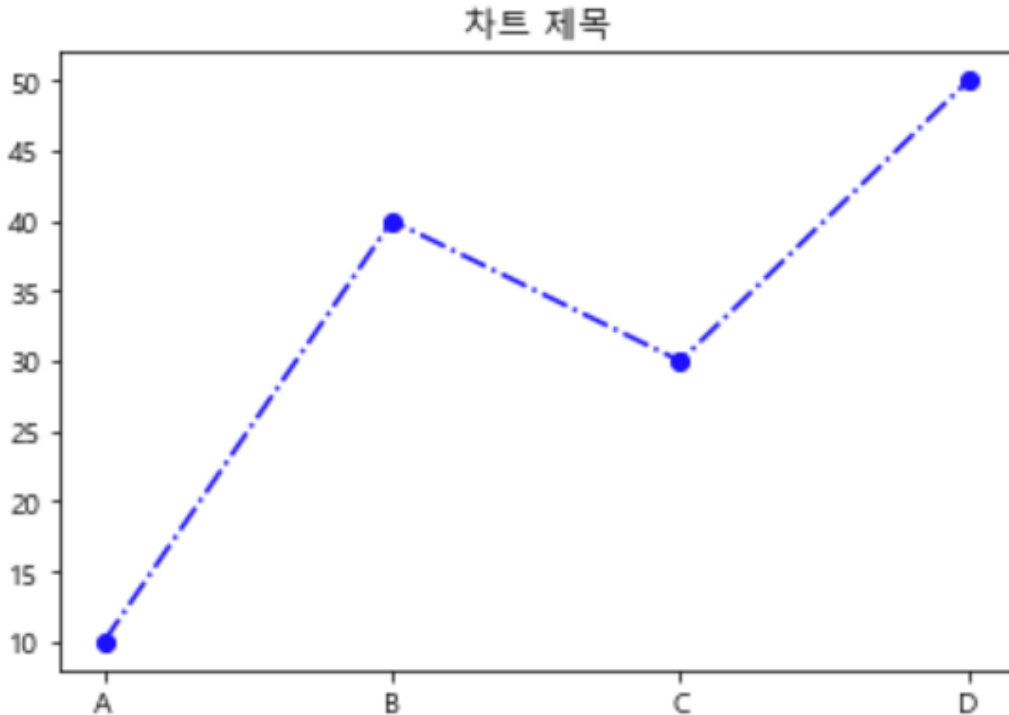
```
import matplotlib.pyplot as plt
plt.plot(['A', 'B', 'C', 'D'], [10, 40, 30, 50], 'ob-.')
plt.title('차트 제목')
plt.show()
```



→ 한글이 제대로 처리 되지 않음

[그래프 제목 넣기 - 한글 사용]

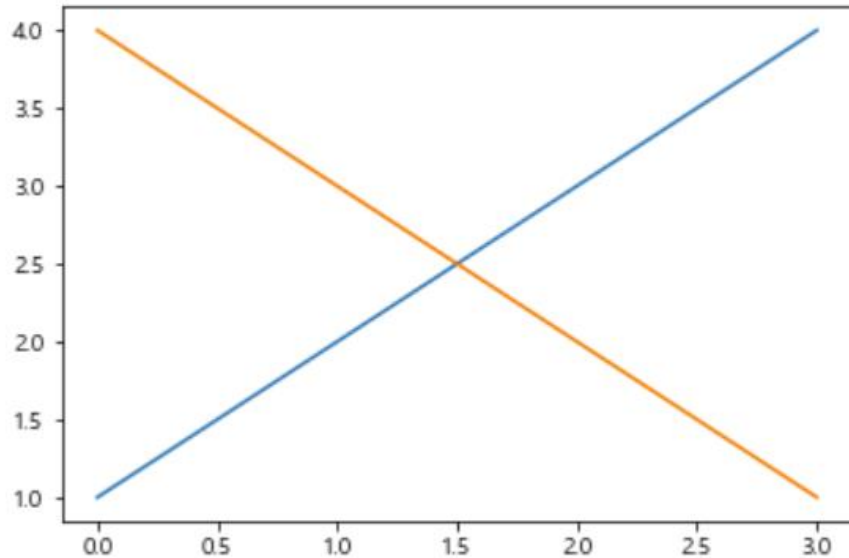
```
import matplotlib.pyplot as plt
plt.plot(['A', 'B', 'C', 'D'], [10, 40, 30, 50], 'ob-.')
plt.rcParams['font.family'] = 'Malgun Gothic' # 한글처리
plt.title('차트 제목')
plt.show()
```



[여러 개의 선 그래프]

- 여러 개의 plot 함수 사용으로 하나의 그림에 겹쳐서 그릴 수 있다

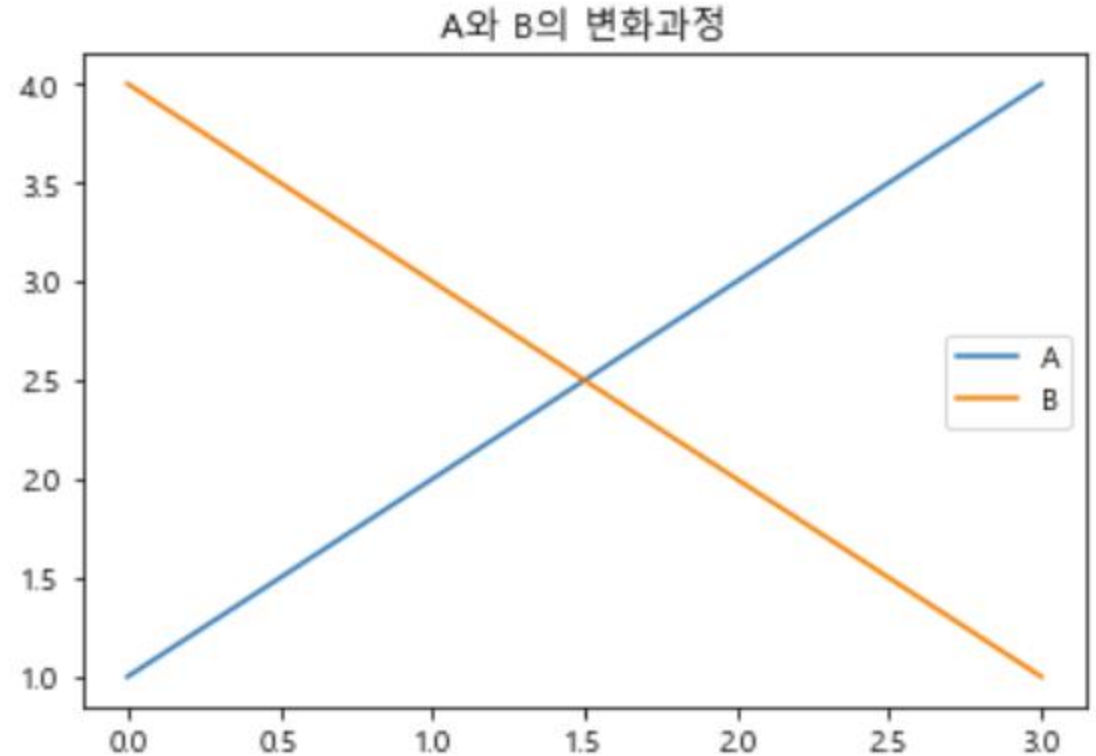
```
import matplotlib.pyplot as plt  
plt.plot([1,2,3,4])  
plt.plot([4,3,2,1])  
  
plt.show()
```



[여러 개의 선 그래프- 범례 표시하기]

- 레이블(label)을 설정한 후 범례(legend)를 그림

```
import matplotlib.pyplot as plt
plt.plot([1,2,3,4],label='A')
plt.plot([4,3,2,1],label='B')
plt.legend()
plt.rcParams['font.family']='Malgun Gothic'
plt.title('A와 B의 변화과정')
plt.show()
```



[참고-그래프의 각 부분의 명칭]

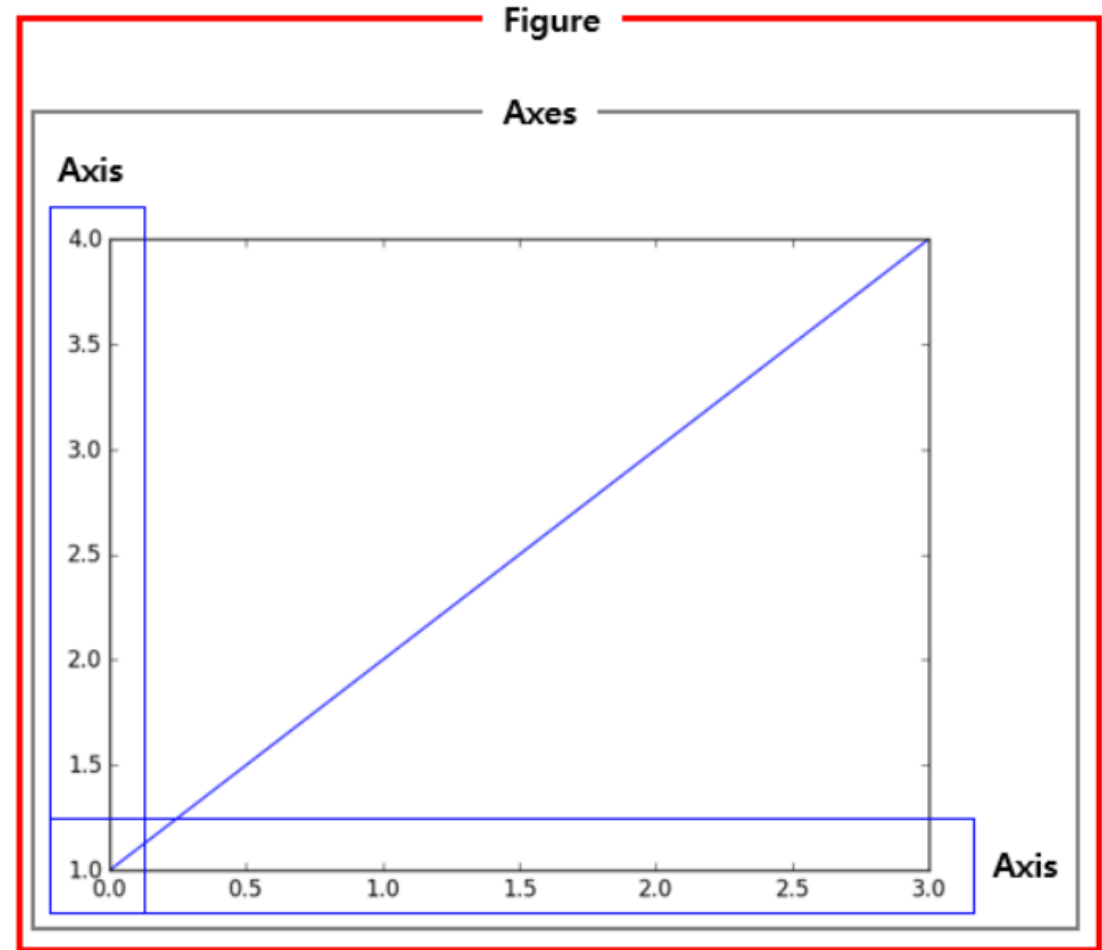
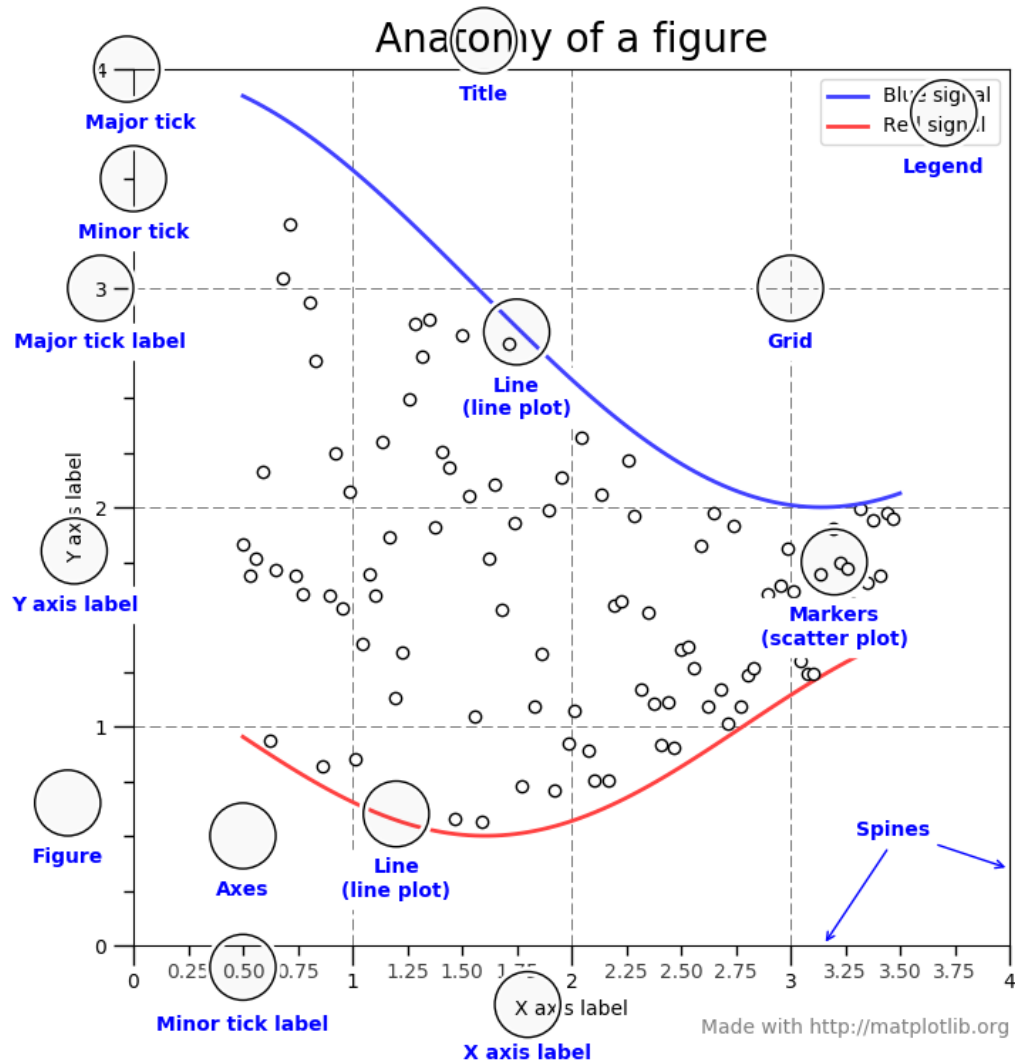


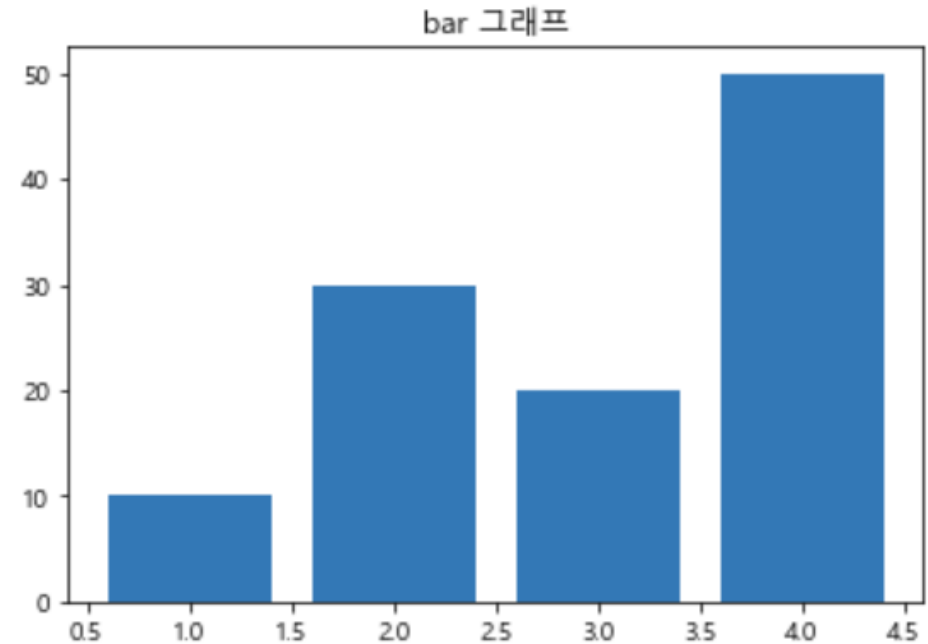
그림 15.10 Figure의 구성 <https://wikidocs.net/4763>

#2. matplotlib을 이용하여 다양한 그래프 그려보기

[막대(bar) 그래프 그리기 : bar(수직방향)]

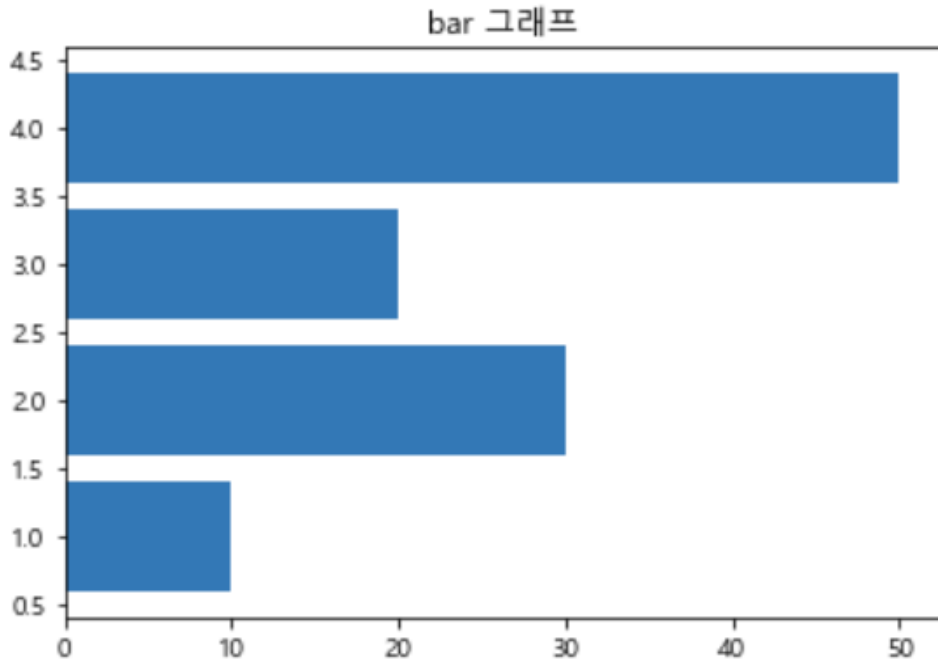
`plt.bar(x,height,[width=width_f, color=colors, tick_label=tick_labels, align='center'(기본) 혹은 'edge', label=labels])`

```
import matplotlib.pyplot as plt
plt.rcParams['font.family']='Malgun Gothic'
plt.title('bar 그래프')
plt.bar([1,2,3,4],[10,30,20,50])
plt.show()
```



[막대(bar) 그래프 그리기 : barh(수평방향)]

```
import matplotlib.pyplot as plt
plt.rcParams['font.family']='Malgun Gothic'
plt.title('bar 그래프')
plt.barh([1,2,3,4],[10,30,20,50])
plt.show()
```



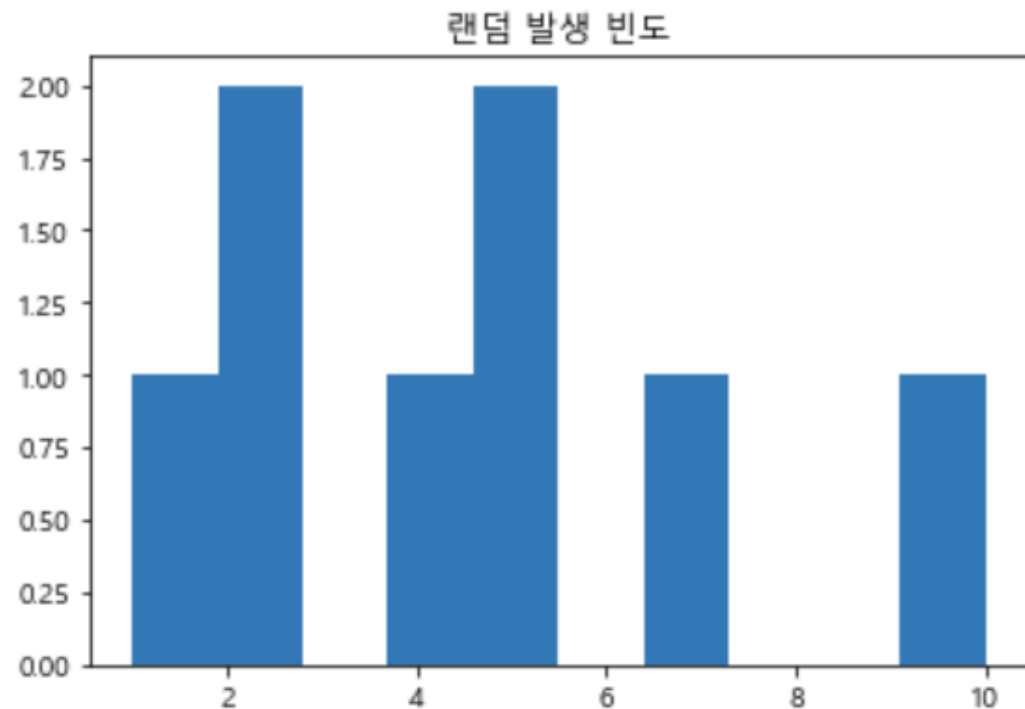
[히스토그램(histogram)]

데이터를 정해진 간격으로 나눈 후 그 간격 안에 들어간 데이터의 개수를 막대로 표시한 그래프

plt.hist(x,[,bins=bins_n 혹은 'auto']

```
import matplotlib.pyplot as plt
data=[]
import random
for i in range(8) :
    data.append(random.randint(1,10))
print(data)
plt.rcParams['font.family']='Malgun Gothic'
plt.title('랜덤 발생 빈도')
plt.hist(data)
plt.show()
```

[5, 7, 2, 2, 5, 4, 1, 10]

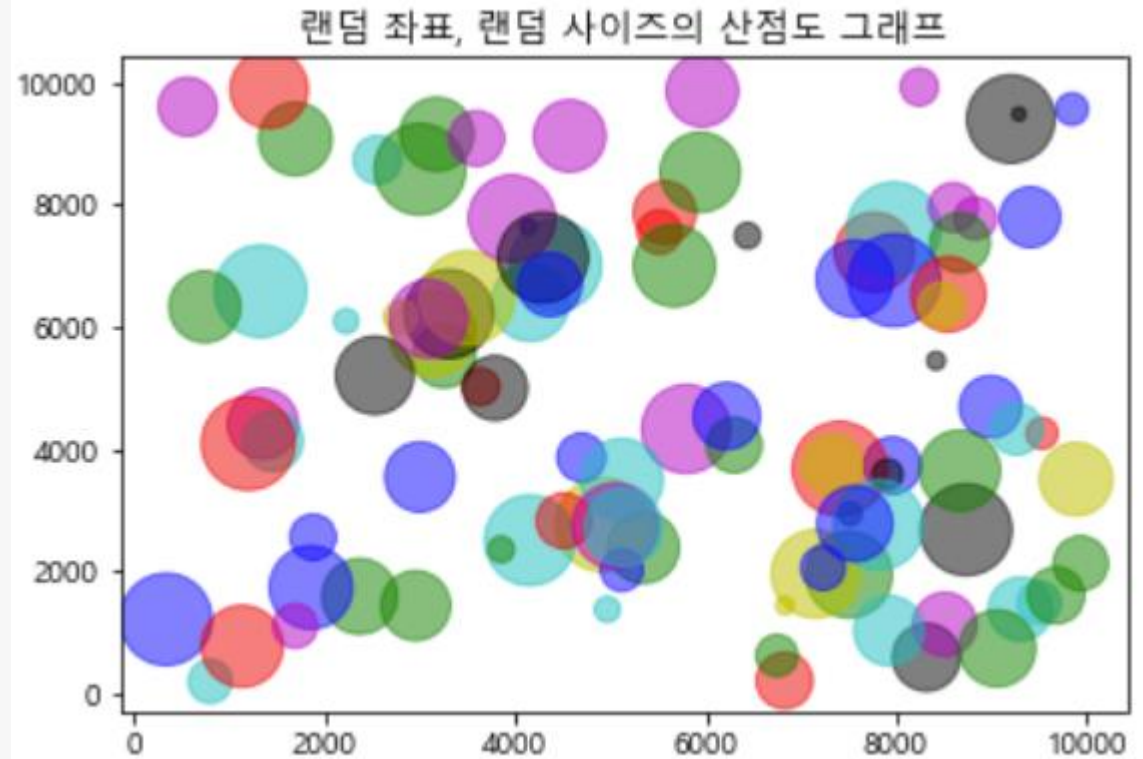


[산점도(scatter)]

```
plt.scatter(x,y[,s=size_n, c=colors, marker='marker_string', alpha=alpha_f])
```

```
import matplotlib.pyplot as plt
x=[]
y=[]
size=[]
c=[]
colors=['r','g','b','c','m','k','y']
plt.title('랜덤 좌표, 랜덤 사이즈의 산점도 그래프')
import random
for i in range(100) :
    x.append(random.randint(1,10000))
    y.append(random.randint(1,10000))
    size.append(random.randint(1,1000))
    c.append(random.choice(colors))

plt.scatter(x,y,s=size,c=c,alpha=0.5)
plt.show()
```



[파이 그래프]

`plt.pie(x,[,labels=label_seq,autopct='비율표시형식(ex :%0.1f)',
shadow=False(기본) 혹은 True, explode=explode_seq,
counterclock=True(기본) 혹은 False, startangle=각도(기본은0)])`

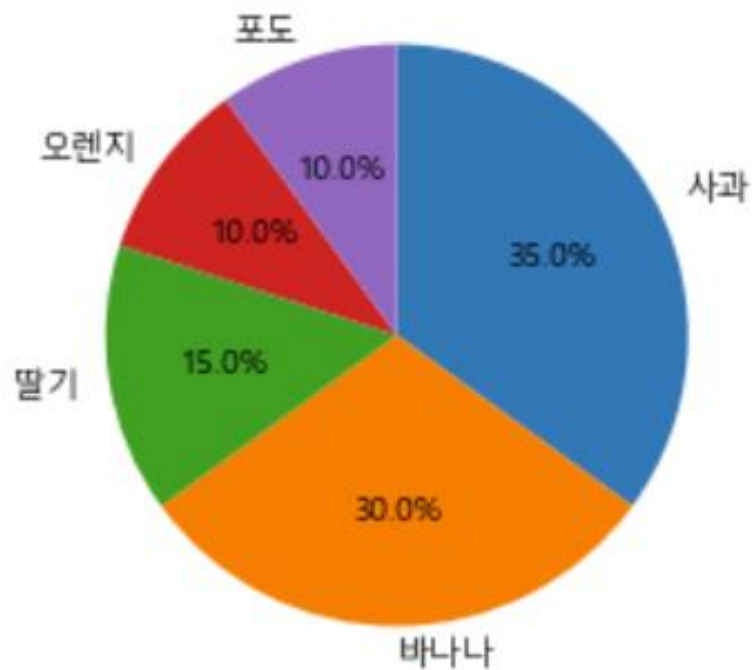
```
import matplotlib.pyplot as plt
fruit=['사과','바나나','딸기','오렌지','포도']
result=[7,6,3,2,2]
plt.rcParams['font.family']='Malgun Gothic'
plt.title('좋아하는 선택 과일 결과')
plt.pie(result)
plt.show()
```

좋아하는 선택 과일 결과

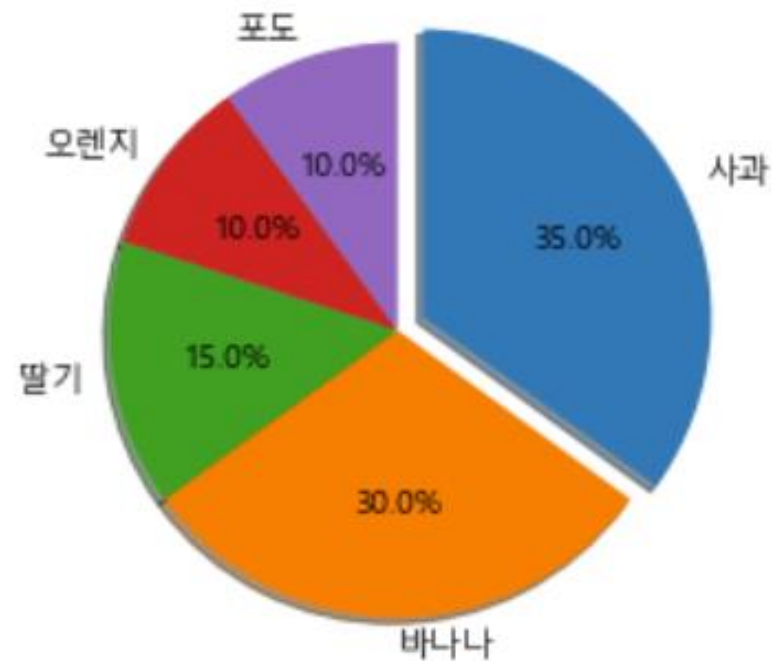


[파이 그래프]

좋아하는 선택 과일 결과

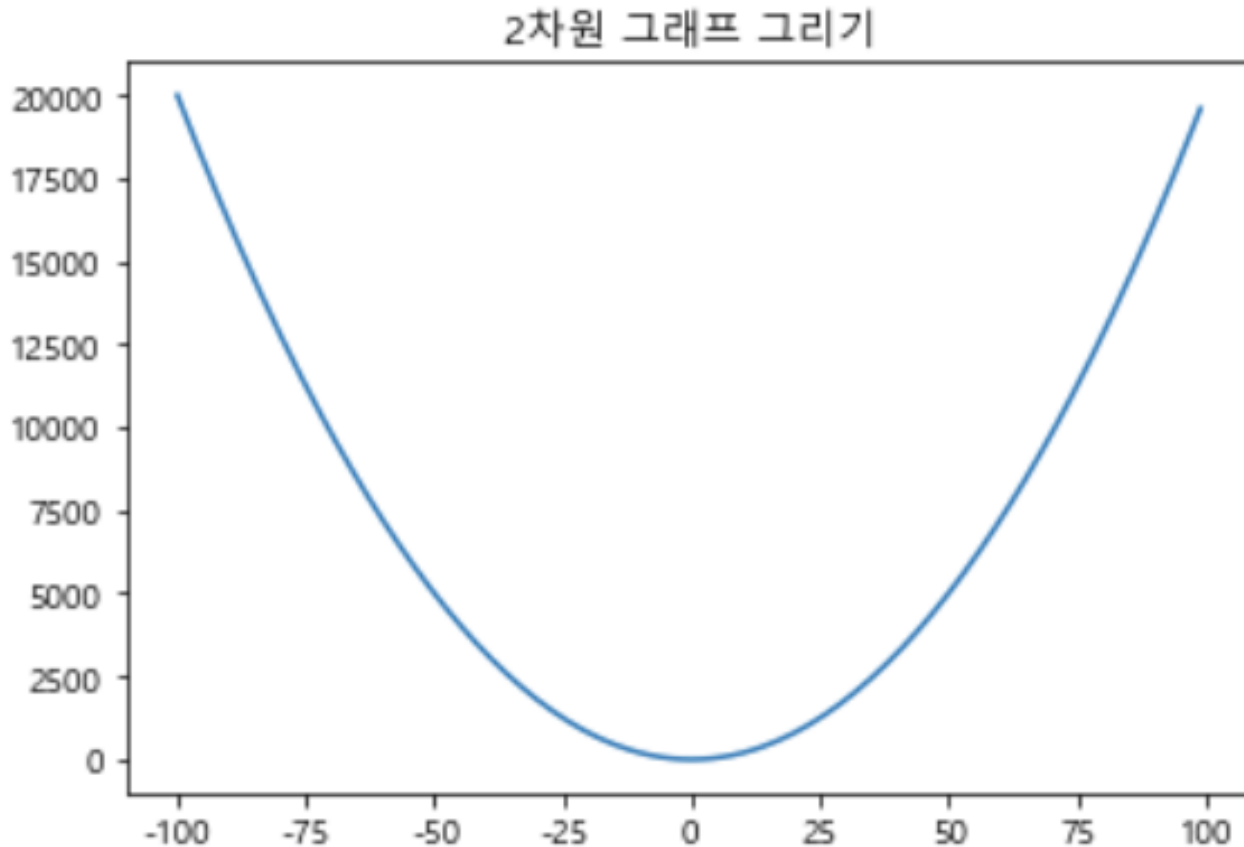


좋아하는 선택 과일 결과



[실습1]

$y = 2x^2$ 그래프를 그려볼까요?



x축 데이터 : -100~100까지의 숫자

y축 데이터 : x축 데이터에 따라 계산하기

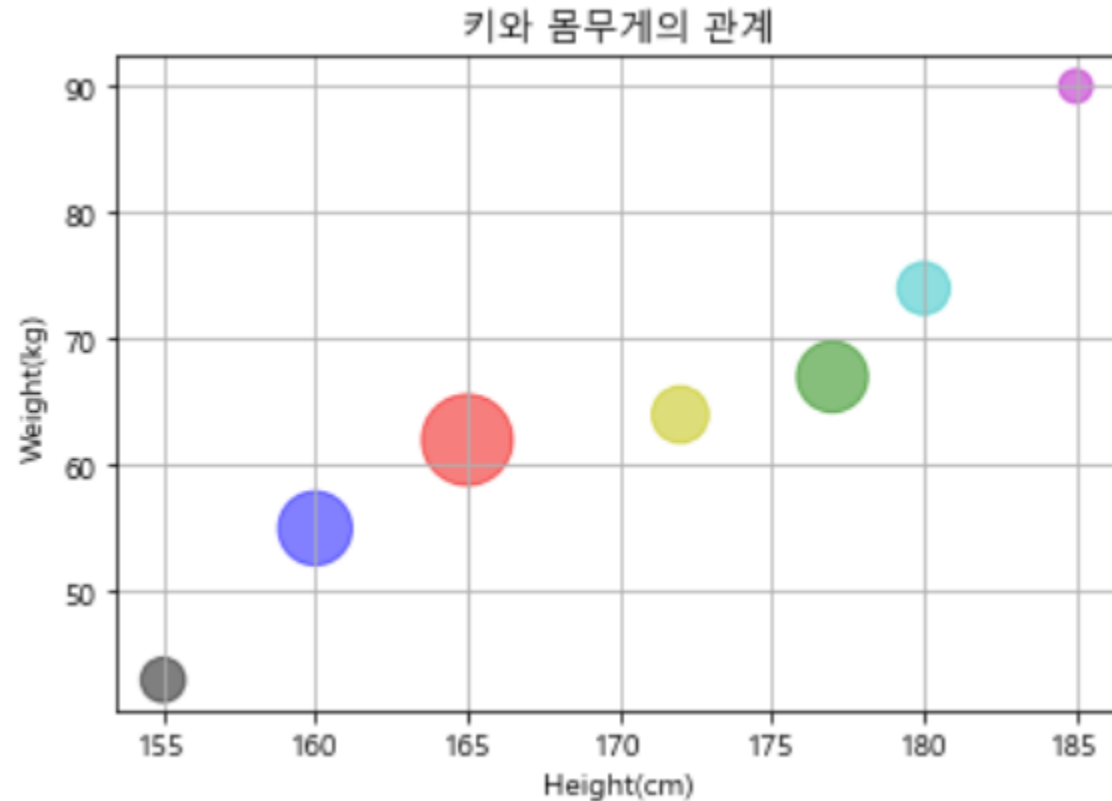
[실습1]

$y = 2x^2$ 그래프를 그려볼까요?

```
import matplotlib.pyplot as plt
plt.rcParams['font.family']='Malgun Gothic' #한글 처리
plt.rcParams['axes.unicode_minus']=False #마이너스 기호(-) 문제 해결
x=[]
y=[]
for i in range(-100,100) :
    x.append(i)
    y.append(2*i**2)
plt.title('2차원 그래프 그리기')
plt.plot(x,y)
plt.show()
```

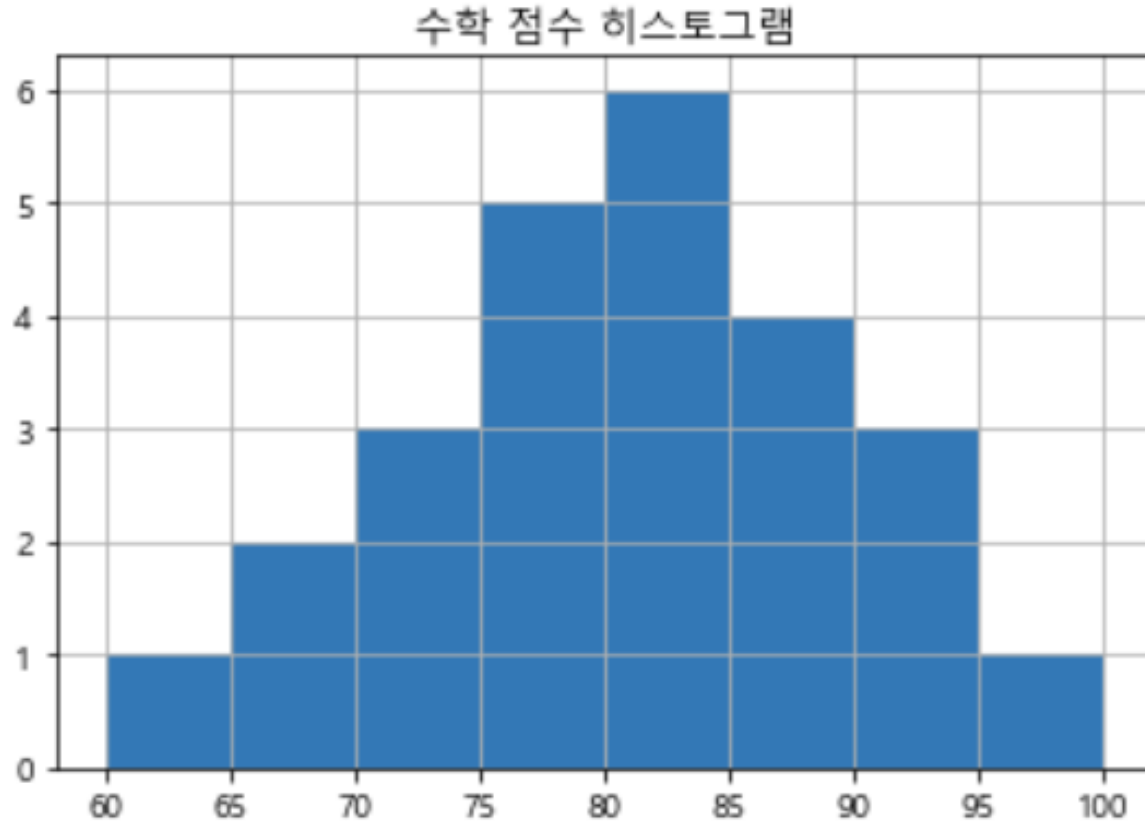
[실습2]

키와 몸무게의 관계를 나타내는 산점도 그래프 그리기
(자료 : 'data\height_weight.csv')



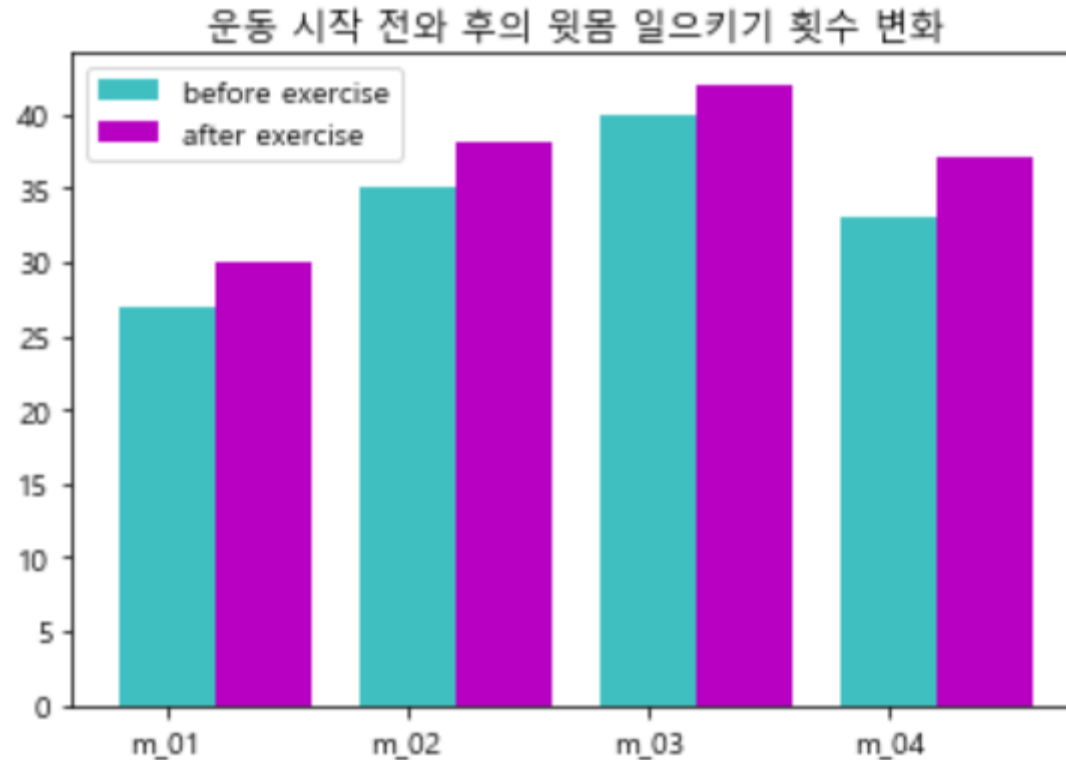
[실습3]

학생 25명에 대한 수학점수 히스토그램 그리기
(자료 : 'data\math_score.csv')



[실습4]

운동 시작 전과 후의 뒷몸 일으키기 횟수 변화 막대 그래프 그리기
(자료 : 'data/exercise_result.csv')



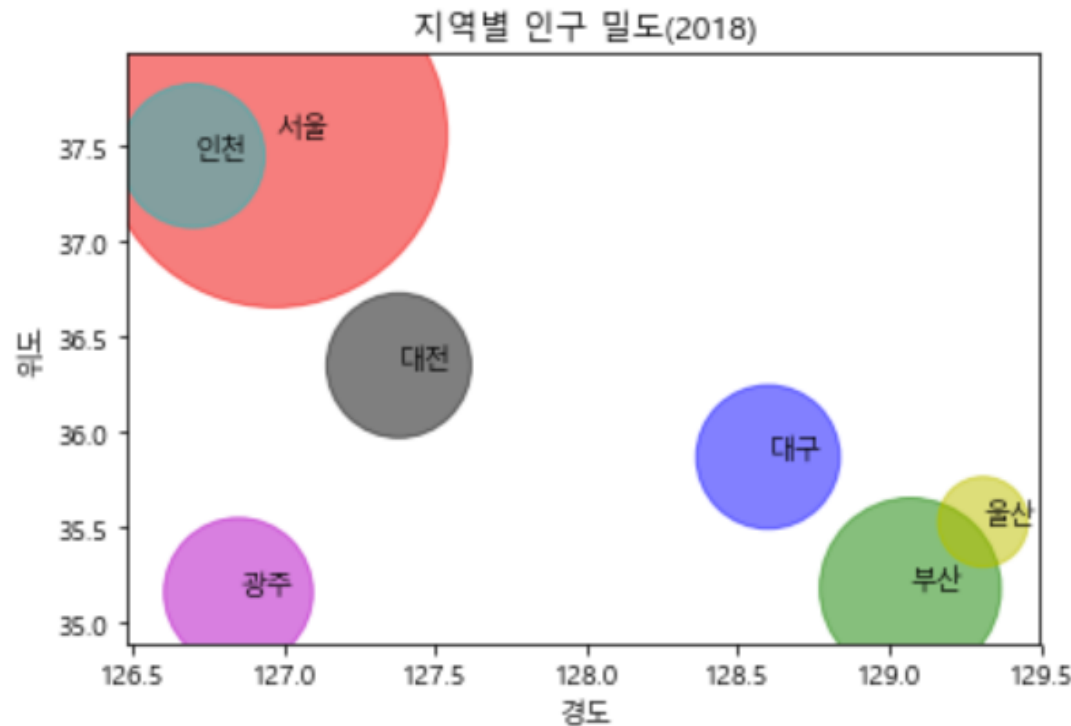
#3. [project]

#3-1.우리나라 지역별 인구밀도를 scatter graph로 나타내기

#3-2.매년 내 생일의 기온 변화 선 그래프 그리기

#3. [project1]

우리나라 지역별 인구밀도를 scatter graph로 나타내기
(자료 : 'data/pop_density.csv')



#3. [project2]

매년 내 생일의 기온 변화 선 그래프 그리기
(자료 : 'data/seoul_temp.csv')

