/* elice */

파이썬으로 시작하는 데이터 분석

Matplotlib 데이터 시각화

임원균 선생님

목차

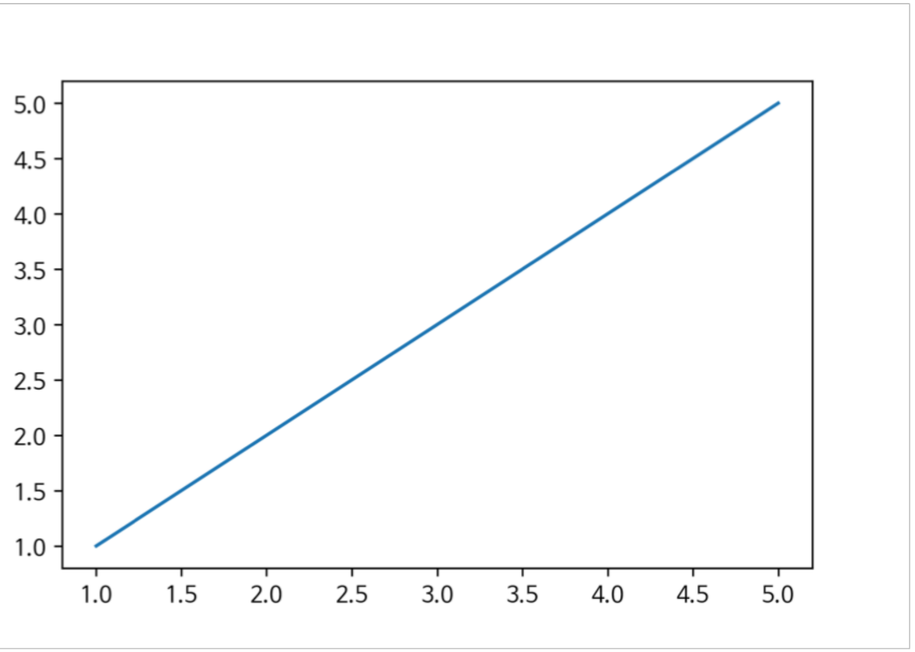
Matplotlib 소개

Matplotlib

파이썬에서 데이터를 그래프나 차트로
시각화할 수 있는 라이브러리

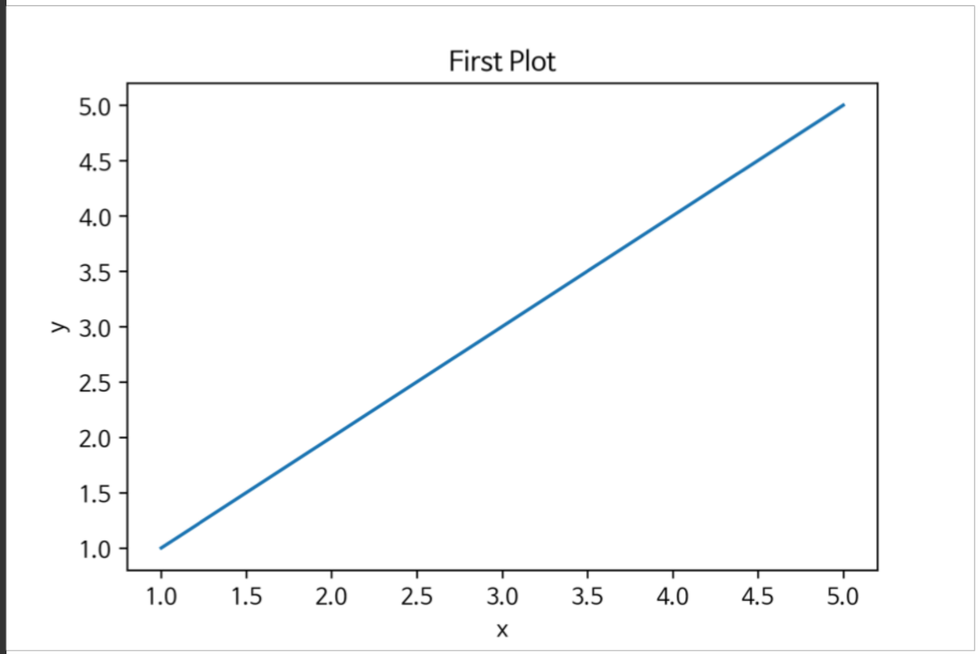
그래프 그려보기

```
x = [1, 2, 3, 4, 5]
y = [1, 2, 3, 4, 5]
plt.plot(x, y)
```



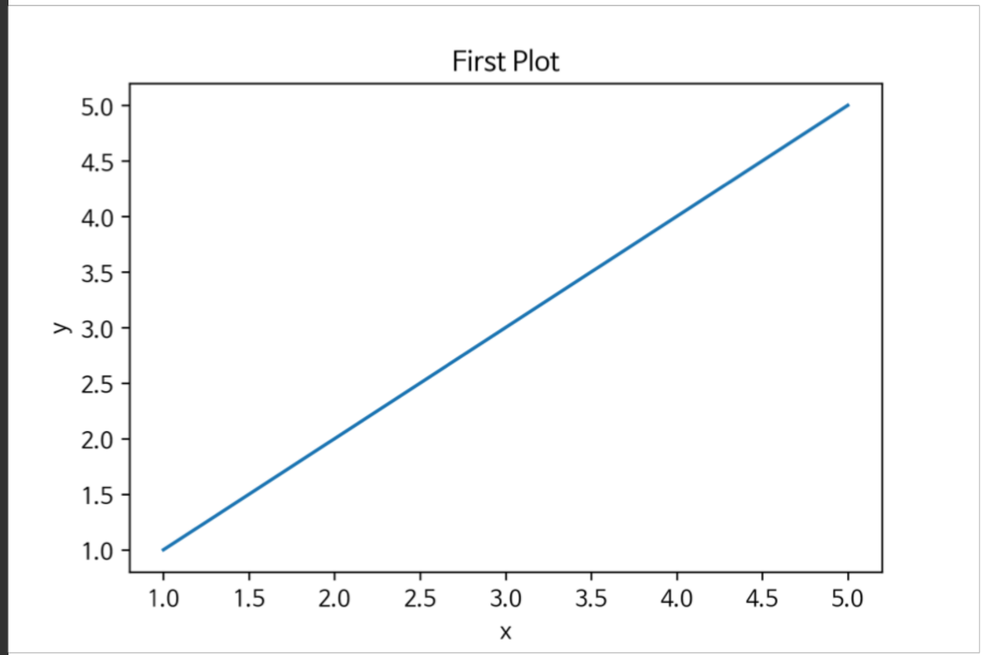
그래프 그려보기

```
x = [1, 2, 3, 4, 5]
y = [1, 2, 3, 4, 5]
plt.plot(x, y)
plt.title("First Plot")
plt.xlabel("x")
plt.ylabel("y")
```

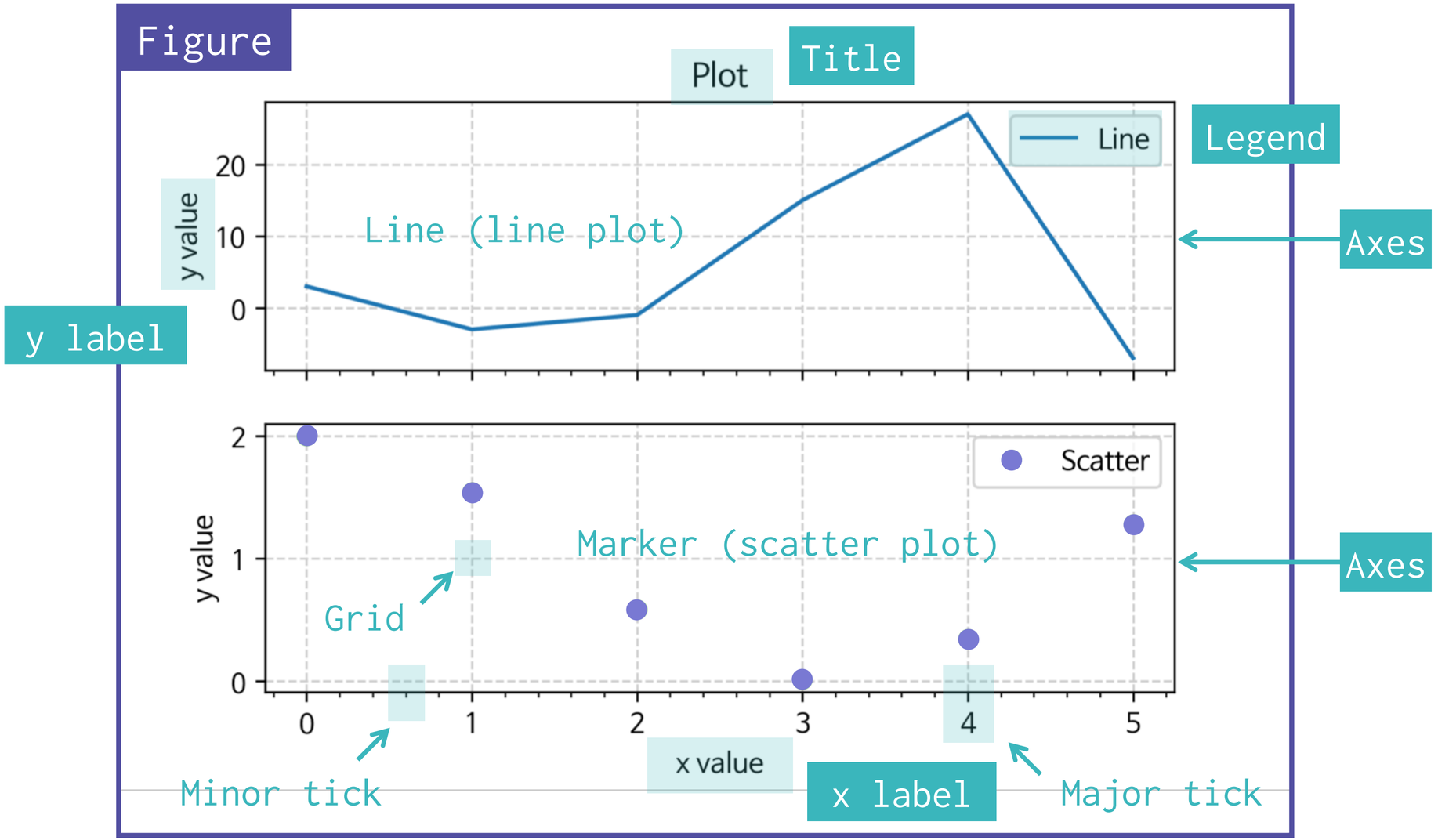


그래프 그려보기

```
x = [1, 2, 3, 4, 5]
y = [1, 2, 3, 4, 5]
fig, ax = plt.subplots()
ax.plot(x, y)
ax.set_title("First Plot")
ax.set_xlabel("x")
ax.set_ylabel("y")
```

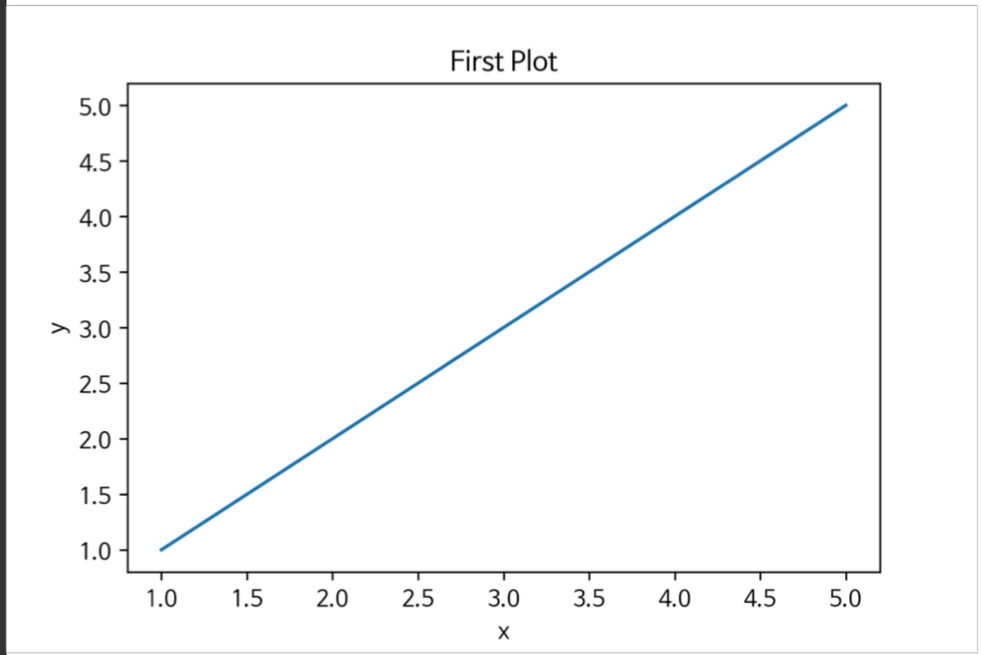


Matplotlib 구조



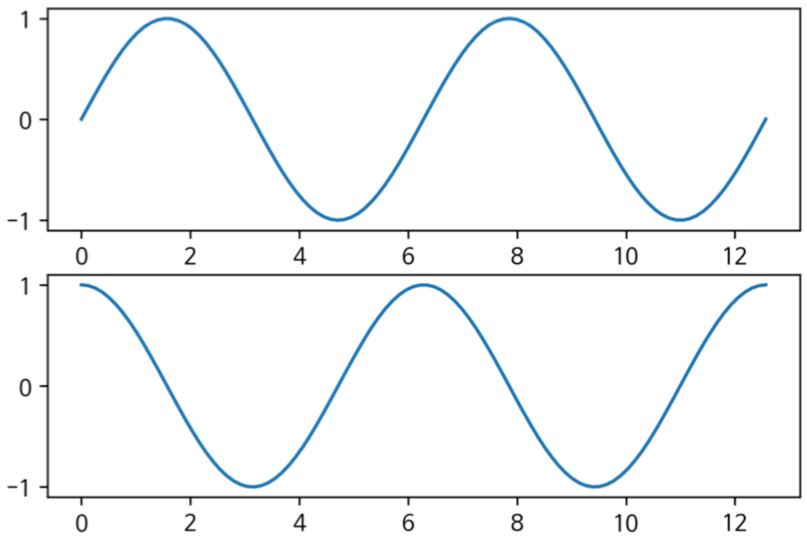
저장하기

```
x = [1, 2, 3, 4, 5]
y = [1, 2, 3, 4, 5]
fig, ax = plt.subplots()
ax.plot(x, y)
ax.set_title("First Plot")
ax.set_xlabel("x")
ax.set_ylabel("y")
fig.set_dip(300)
fig.savefig(”first_plot.png”)
```



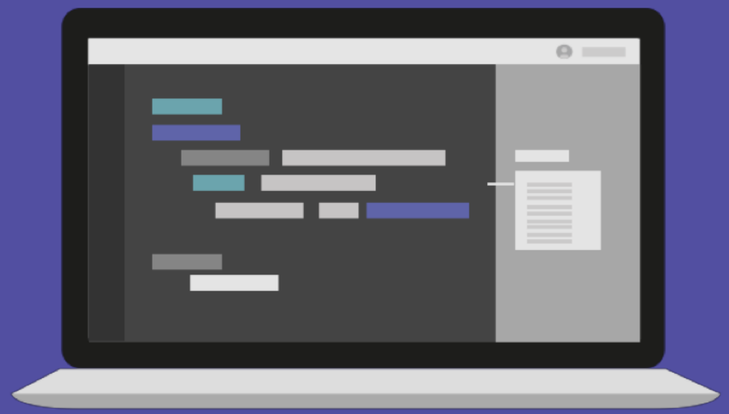
여러개 그래프 그리기

```
x = np.linspace(0, np.pi*4, 100)
fig, axes = plt.subplots(2, 1)
axes[0].plot(x, np.sin(x))
axes[1].plot(x, np.cos(x))
```



[실습1]

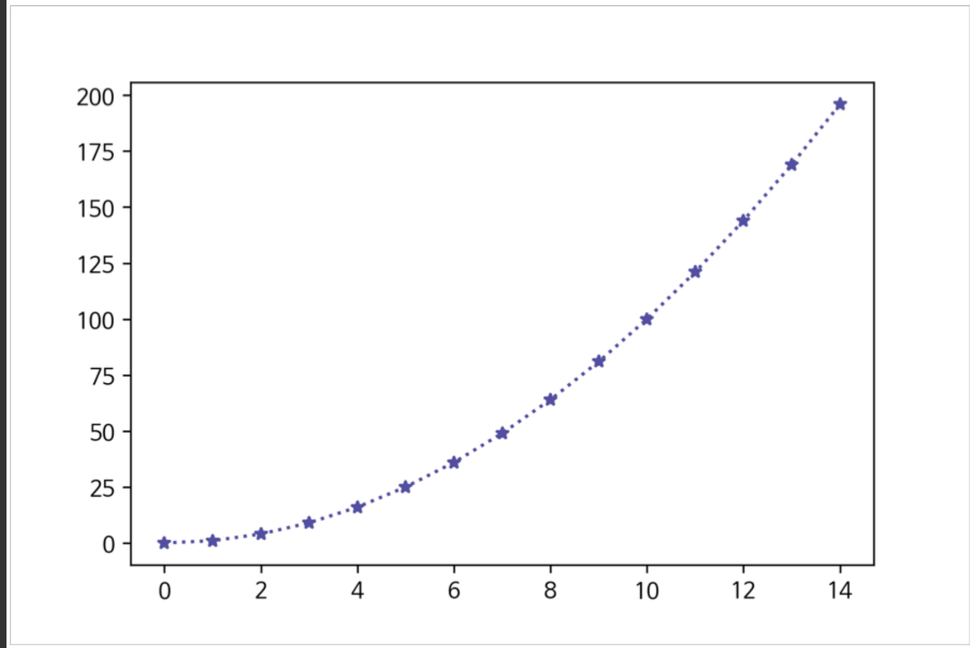
Matplotlib 기초



Matplotlib 그래프

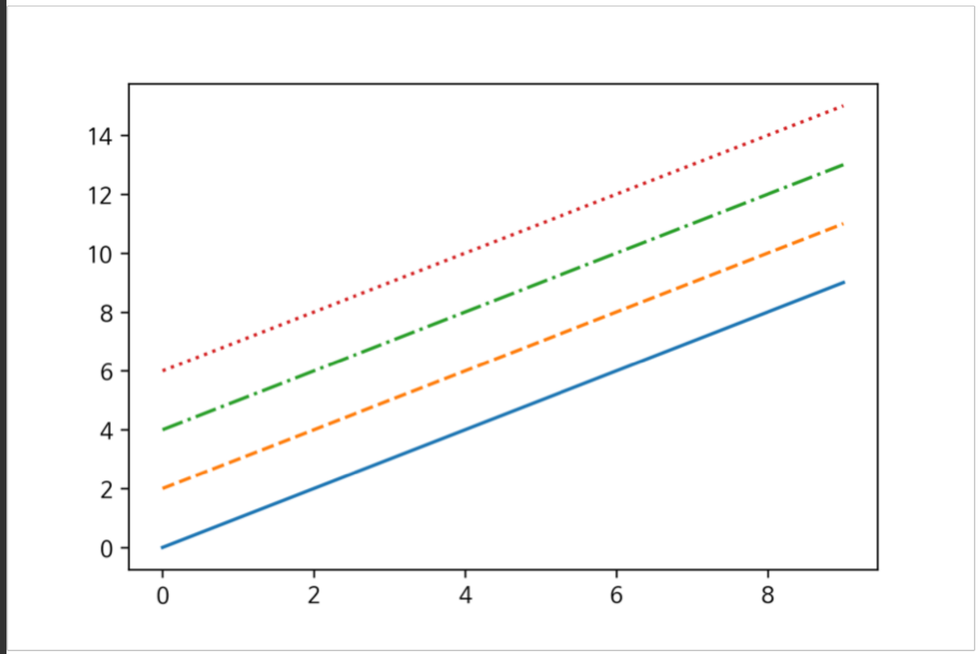
Line plot

```
fig, ax = plt.subplots()
x = np.arange(15)
y = x ** 2
ax.plot(
    x, y,
    linestyle=":",
    marker="*",
    color="#524FA1"
)
```



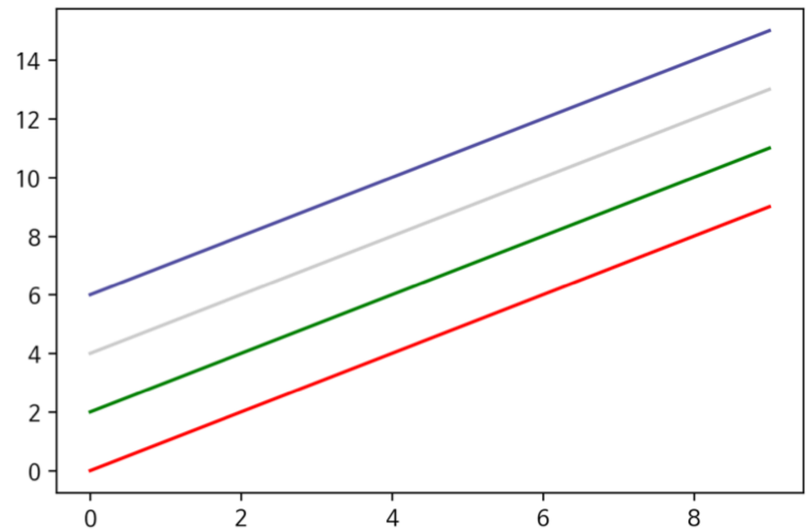
Line style

```
x = np.arange(10)
fig, ax = plt.subplots()
ax.plot(x, x, linestyle="-")
# solid
ax.plot(x, x+2, linestyle="--")
# dashed
ax.plot(x, x+4, linestyle="-.")
# dashdot
ax.plot(x, x+6, linestyle=":")
# dotted
```



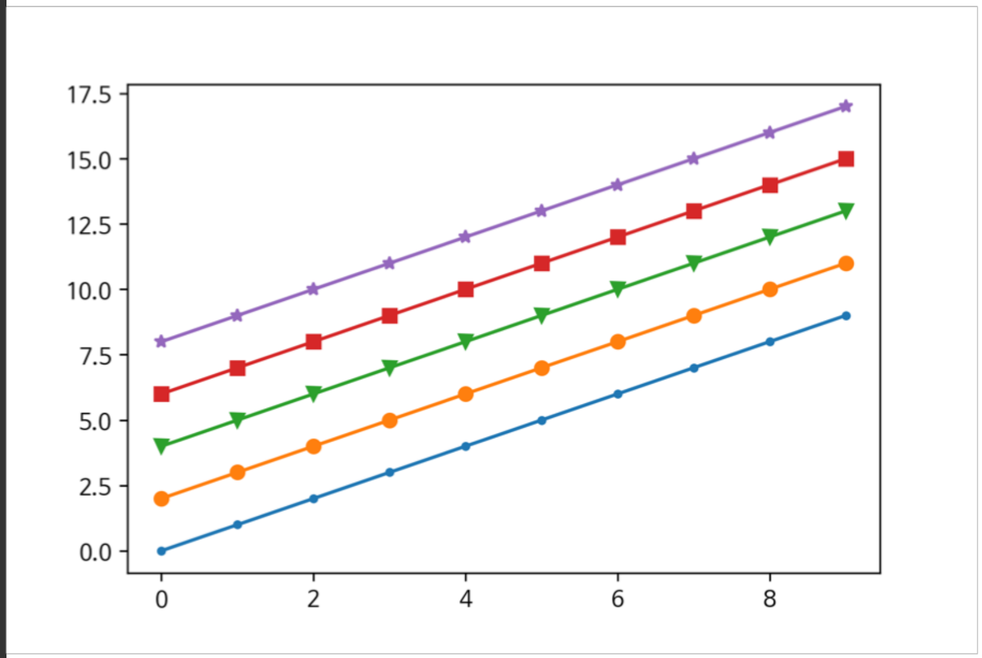
Color

```
x = np.arange(10)
fig, ax = plt.subplots()
ax.plot(x, x, color="r")
ax.plot(x, x+2, color="green")
ax.plot(x, x+4, color='0.8')
ax.plot(x, x+6, color="#524FA1")
```



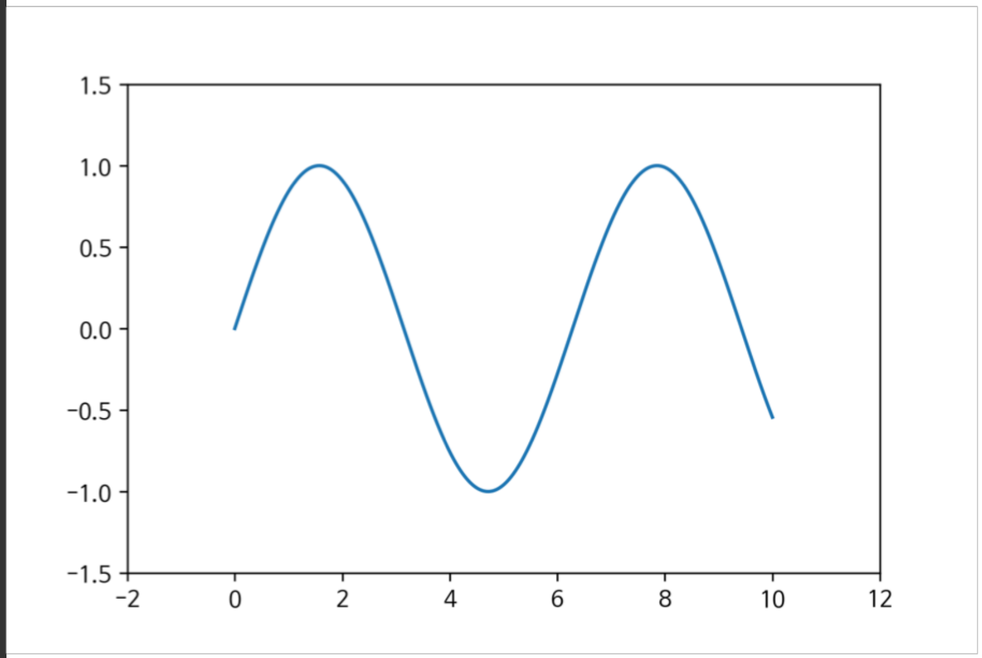
Marker

```
x = np.arange(10)
fig, ax = plt.subplots()
ax.plot(x, x, marker=".")
ax.plot(x, x+2, marker="o")
ax.plot(x, x+4, marker='v')
ax.plot(x, x+6, marker="s")
ax.plot(x, x+8, marker="*")
```



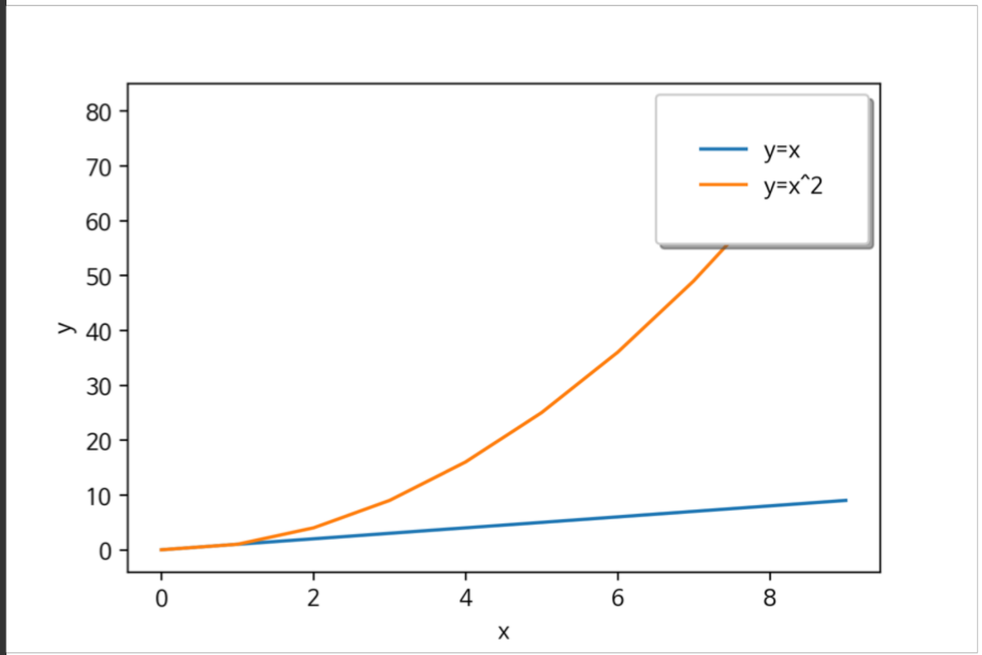
축 경계 조정하기

```
x = np.linspace(0, 10, 1000)
fig, ax = plt.subplots()
ax.plot(x, np.sin(x))
ax.set_xlim(-2, 12)
ax.set_ylim(-1.5, 1.5)
```



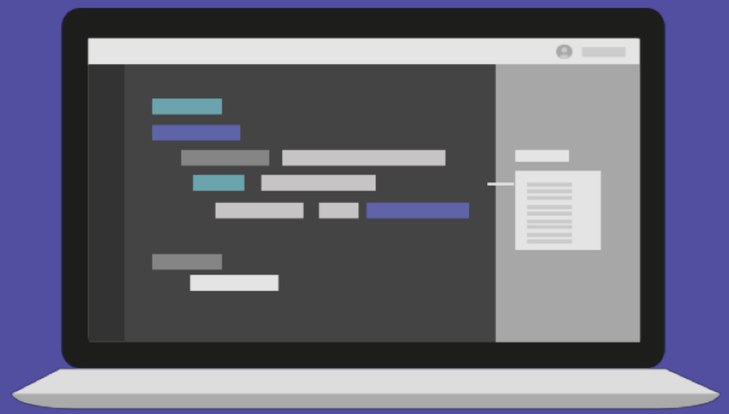
범례

```
fig, ax = plt.subplots()
ax.plot(x, x, label='y=x')
ax.plot(x, x**2, label='y=x^2')
ax.set_xlabel("x")
ax.set_ylabel("y")
ax.legend(
    loc='upper right',
    shadow=True,
    fancybox=True,
    borderpad=2
)
```



[실습2]

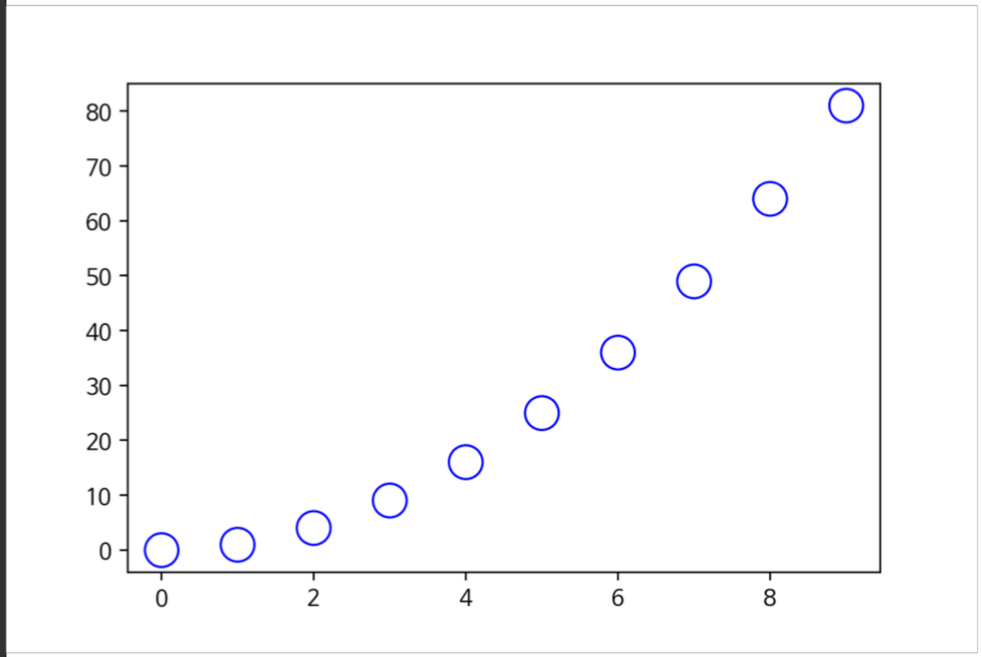
Line Graph



Scatter

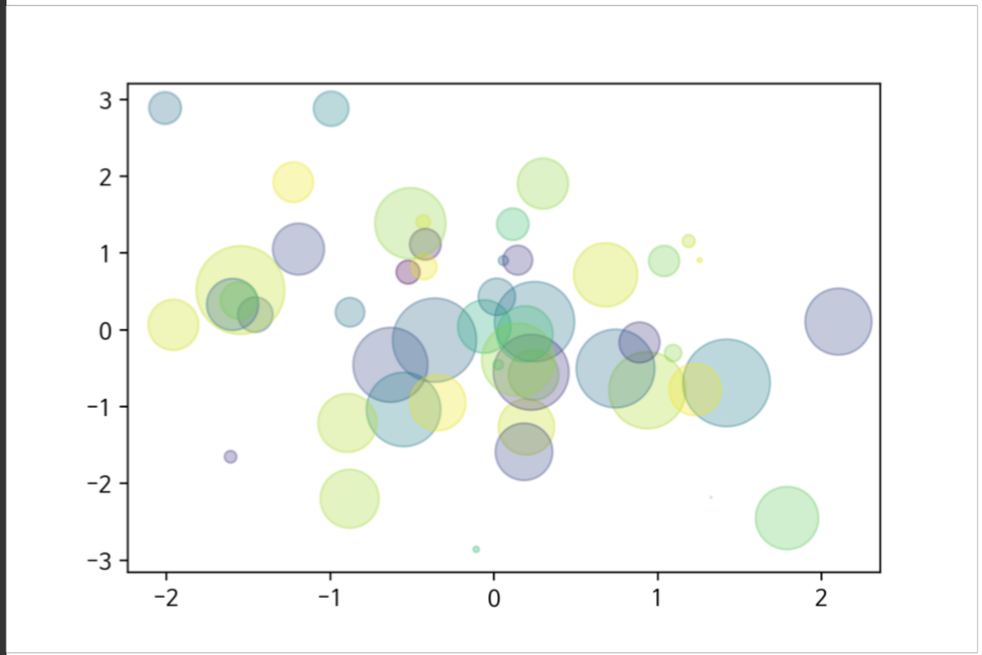
Scatter

```
fig, ax = plt.subplots()
x = np.arange(10)
ax.plot(
    x, x**2, "o",
    markersize=15,
    markerfacecolor='white',
    markeredgecolor="blue"
)
```



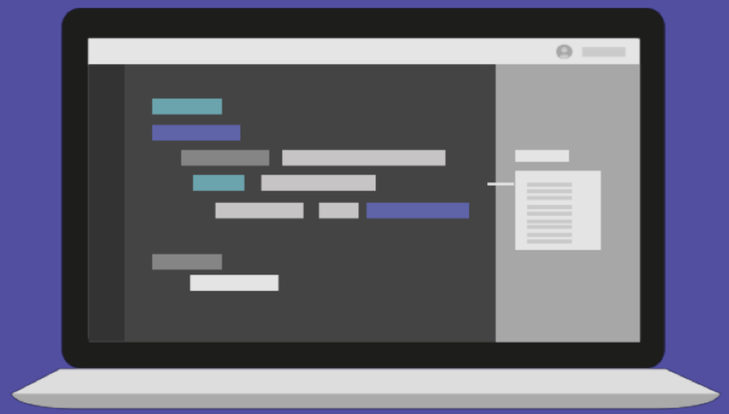
Scatter

```
  fig, ax = plt.subplots()
  x = np.random.randn(50)
  y = np.random.randn(50)
  colors = np.random.randint(0, 100,
50)
  sizes = 500 * np.pi *
np.random.rand(50) ** 2
  ax.scatter(
      x, y, c=colors, s=sizes,
alpha=0.3
  )
```



[실습3]

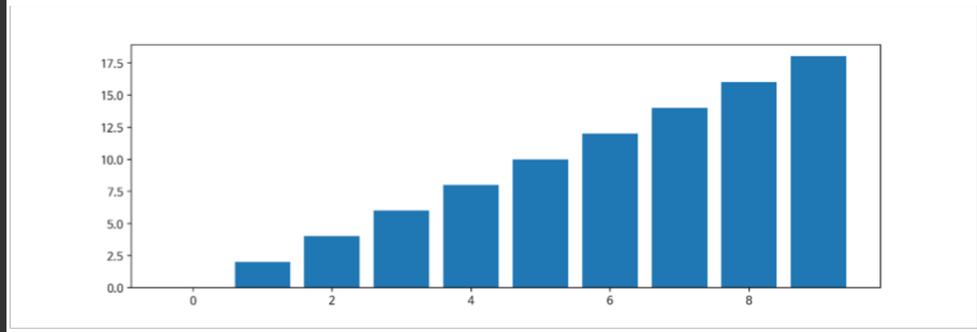
Scatter



Bar & Histogram

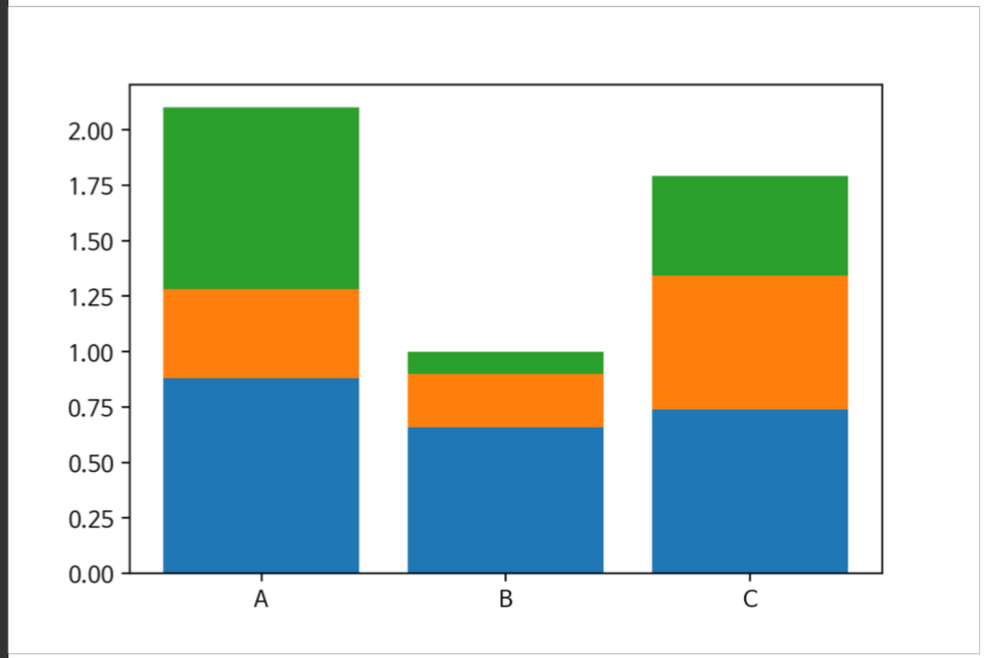
Bar plot

```
  # bar
  x = np.arange(10)
  fig, ax =
plt.subplots(figsize=(12, 4))
  ax.bar(x, x*2)
```



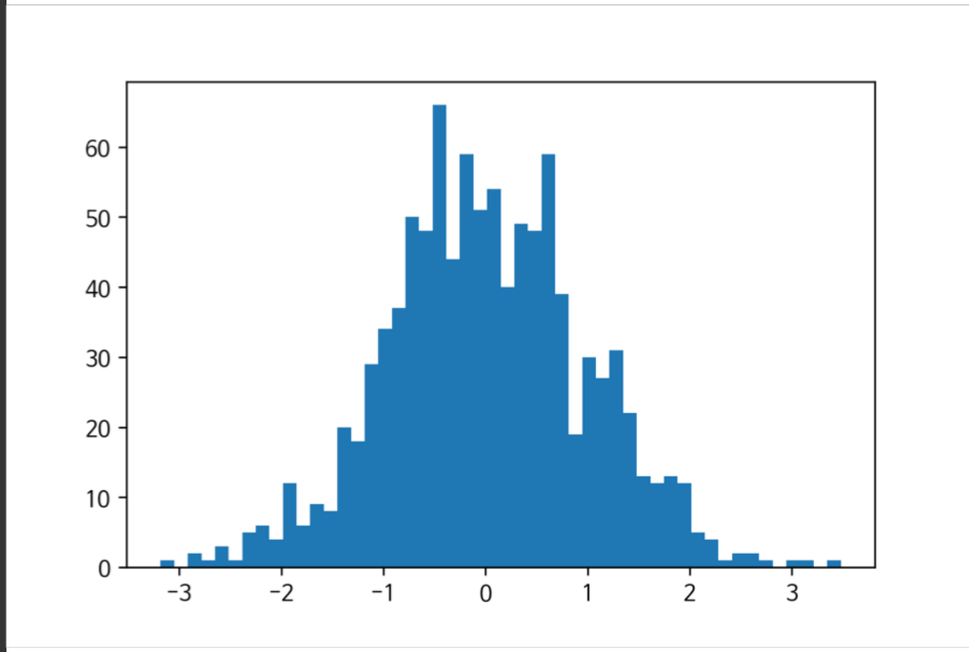
Bar plot

```
x = np.random.rand(3)
y = np.random.rand(3)
z = np.random.rand(3)
data = [x, y, z]

fig, ax = plt.subplots()
x_ax = np.arange(3)
for i in x_ax:
    ax.bar(x_ax, data[i],
    bottom=np.sum(data[:i], axis=0))
ax.set_xticks(x_ax)
ax.set_xticklabels(["A", "B", "C"])
```



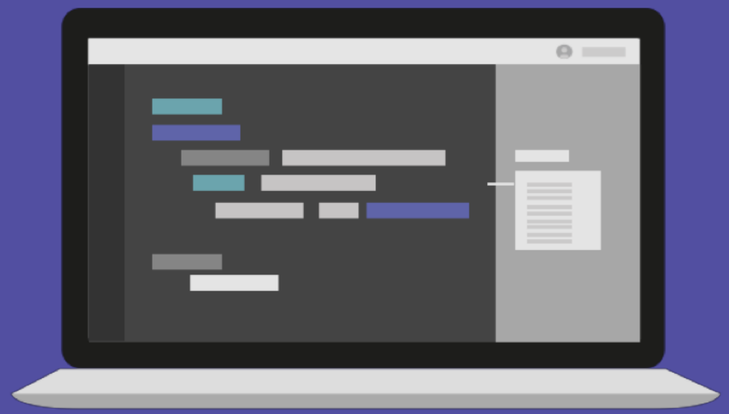
Histogram

```
fig, ax = plt.subplots()

data = np.random.randn(1000)

ax.hist(data, bins=50)
```



[실습4]

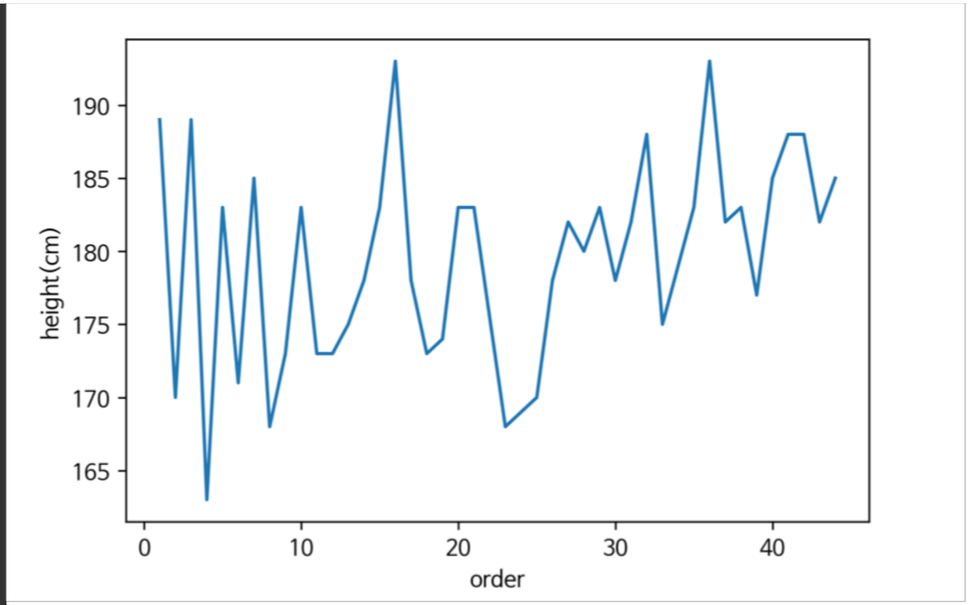
Bar & Histogram



Matplotlib with Pandas

Matplotlib with pandas

```
df = pd.read_csv(
"./president_heights.csv")
fig, ax = plt.subplots()
ax.plot(df["order"],
df["height(cm)"], label="height")
ax.set_xlabel("order")
ax.set_ylabel("height(cm)")
```



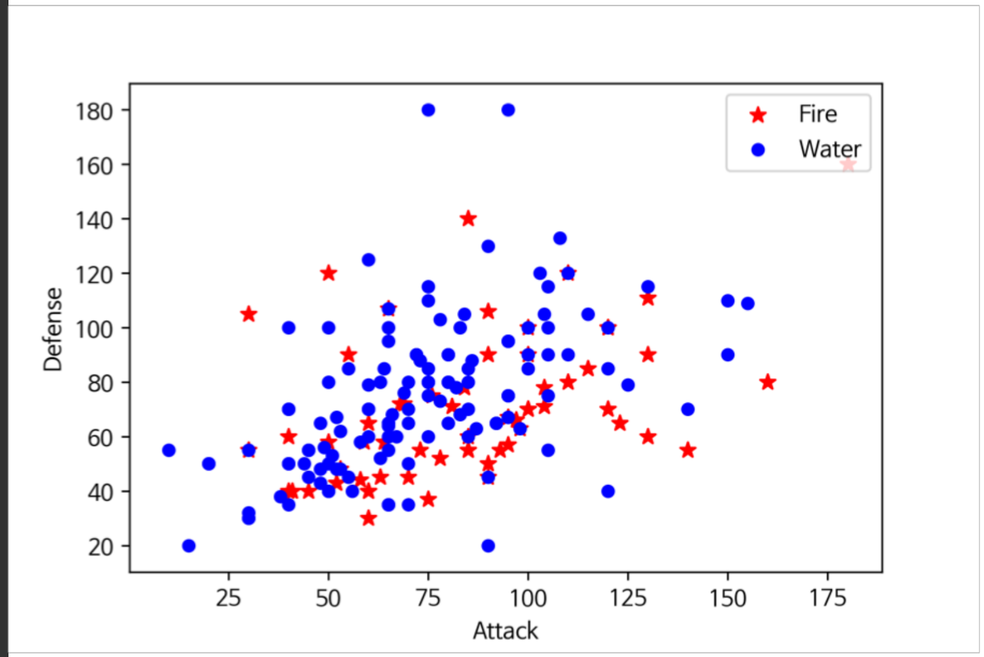
	order	name	height(cm)
0	1	George Washington	189
1	2	John Adams	170
2	3	Thomas Jefferson	189
3	4	James Madison	163
4	5	James Monroe	183

Matplotlib with pandas

	#	Name	Type 1	Type 2	Total	HP	Attack	Defense	Sp. Atk	Sp. Def	Speed	Generation	Legendary
0	1	Bulbasaur	Grass	Poison	318	45	49	49	65	65	45	1	False
1	2	Ivysaur	Grass	Poison	405	60	62	63	80	80	60	1	False
2	3	Venusaur	Grass	Poison	525	80	82	83	100	100	80	1	False
3	3	VenusaurMega Venusaur	Grass	Poison	625	80	100	123	122	120	80	1	False
4	4	Charmander	Fire	NaN	309	39	52	43	60	50	65	1	False
5	5	Charmeleon	Fire	NaN	405	58	64	58	80	65	80	1	False
6	6	Charizard	Fire	Flying	534	78	84	78	109	85	100	1	False
7	6	CharizardMega Charizard X	Fire	Dragon	634	78	130	111	130	85	100	1	False

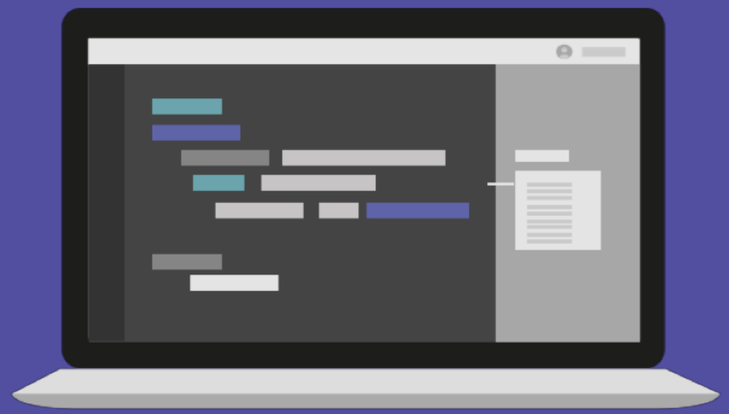
Matplotlib with pandas

```
    fire = df[
          (df['Type 1']=='Fire') | ((df['Type
2'])=="Fire")
    ]
    water = df[
          (df['Type 1']=='Water') | ((df['Type
2'])=="Water")
    ]
    fig, ax = plt.subplots()
    ax.scatter(fire['Attack'], fire['Defense'],
          color='R', label='Fire', marker="*", s=50)
    ax.scatter(water['Attack'], water['Defense'],
          color='B', label="Water", s=25)
    ax.set_xlabel("Attack")
    ax.set_ylabel("Defense")
    ax.legend(loc="upper right")
```



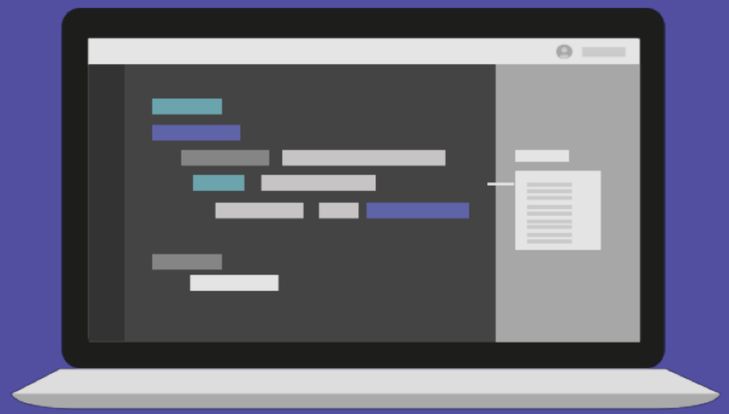
[실습5]

Matplotlib with Pandas



[실습6]

토끼와 거북이 경주 결과 시각화



/* elice */

문의 및 연락처

academy.elice.io

contact@elice.io

facebook.com/elice.io

medium.com/elice