



CHAPTER

08

데이터 시각화 Data Visualization

컴퓨터소프트웨어공학과
김대영

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from datetime import datetime
```

```
def main():
```

```
    data = np.arange(10)
    print(data)
    plt.plot(data)
    plt.show()
```

```
    fig = plt.figure()
    ax1 = fig.add_subplot(2,2,1)
    ax2 = fig.add_subplot(2,2,2)
    ax3 = fig.add_subplot(2,2,3)
    plt.plot(np.random.randn(50).cumsum(), 'k--')
    ax1.hist(np.random.randn(100), bins=20, color='k',
alpha=0.3)
    ax2.scatter(np.arange(30), np.arange(30) + 3 *
np.random.randn(30))
    plt.show()
```

```
    fig, axes = plt.subplots(2,2, sharex=True,
sharey=True)
    for i in range(2):
        for j in range(2):
            axes[i, j].hist(np.random.randn(500),
bins=50, color='k', alpha=0.5)
    plt.subplots_adjust(wspace=0, hspace=0)
    plt.show()
```

```
    data = np.arange(10)
    plt.plot(data, 'g--')
    plt.show()
    plt.plot(data, linestyle='--', color='r')
    plt.show()
```

```
    data = np.random.randn(50).cumsum()
    plt.plot(data, 'ko--')
    plt.show()
    plt.plot(data, color='b',
linestyle='dashed', marker='o')
    plt.show()
```

```
    data = np.random.randn(30).cumsum()
    plt.plot(data, 'k--', label='Default')
    plt.plot(data, 'k-', drawstyle='steps-post',
label='steps-post')
    plt.legend(loc='best')
    plt.show()
```

```
    fig = plt.figure()
    ax = fig.add_subplot(1,1,1)
    ax.plot(np.random.randn(1000).cumsum())
    #plt.show()
```

```
    ticks = ax.set_xticks([0,250,500,750,1000])
    labels = ax.set_xticklabels(['one', 'two', 'three',
'four', 'five'], rotation=30, fontsize='small')
    ax.set_title('My matplotlib plot')
    ax.set_xlabel('Stages')
    plt.show()
```

```
fig = plt.figure()
ax = fig.add_subplot(1,1,1)
ax.plot(np.random.randn(1000).cumsum(), 'k',
label='one')
ax.plot(np.random.randn(1000).cumsum(), 'g--',
label='two')
ax.plot(np.random.randn(1000).cumsum(), 'r.',
label='three')
ax.legend(loc='best')
plt.show()

fig = plt.figure()
ax = fig.add_subplot(1,1,1)
data = pd.read_csv('LEC/examples/spx.csv',
index_col=0, parse_dates=True)
spx = data['SPX']
spx.plot(ax=ax, style='k-')
crisis_data = [
    (datetime(2007,10,11), 'Peak of bull market'),
    (datetime(2008,3,12), 'Bear Stearns Fails'),
    (datetime(2008,9,15), 'Lehman Bankruptcy')
]
for date, label in crisis_data:
    ax.annotate(label, xy=(date, spx.asof(date)+75),
xytext=(date, spx.asof(date)+225), arrowprops=
dict(facecolor='black', headwidth=4, width=2,
headlength=4), horizontalalignment='left',
verticalalignment='top')

ax.set_xlim(['1/1/2007', '1/1/2011'])
ax.set_ylim([600, 1800])
ax.set_title('Important dates in the 2008-2009
financial crisis')
```

```
ax.set_xlabel('DATE')
plt.show()

fig = plt.figure()
ax = fig.add_subplot(1,1,1)

rect = plt.Rectangle((0.2, 0.75), 0.4, 0.15,
color='k', alpha=0.3)
circ = plt.Circle((0.7, 0.2), 0.15, color='b',
alpha=0.3)
pgon = plt.Polygon([[0.15, 0.15], [0.35, 0.4], [0.2,
0.6]], color='g', alpha=0.5)

ax.add_patch(rect)
ax.add_patch(circ)
ax.add_patch(pgon)
plt.show()

font_options = {'family' : 'monospace',
'weight' : 'bold', 'size' : 16}
plt.rc('font', **font_options)

fig = plt.figure()
ax = fig.add_subplot(1,1,1)
ax.plot(np.random.randn(100).cumsum(), 'k-',
label='one')
#plt.savefig('./LEC/figure.png', dpi=400,
bbox_inches='tight')
plt.show()

if __name__ == '__main__': # main()
    main()
```