Reshaping&Pivoting

August 29, 2022

[1]:
```
import pandas as pd
import numpy as np
```

[ ]:
```
data = pd.DataFrame(np.arange(6).reshape((2, 3)),
        index=pd.Index(['Ohio', 'Colorado'], name='state'),
        columns=pd.Index(['one', 'two', 'three'],
        name='number'))
```

[ ]:
```
data
```

[ ]:
```
result = data.stack()
```

[ ]:
```
result
```

[ ]:
```
result.unstack()
```

[ ]:
```
result.unstack(1)
```

[ ]:
```
result.unstack(0)
```

[ ]:
```
result.unstack('state')
```

[ ]:
```
s1 = pd.Series([0, 1, 2, 3], index=['a', 'b', 'c', 'd'])
```

[ ]:
```
s2 = pd.Series([4, 5, 6], index=['c', 'd', 'e'])
```

[ ]:
```
data2 = pd.concat([s1, s2], keys=['one', 'two'])
```

[ ]:
```
s1
```

[ ]:
```
s2
```

[ ]:
```
data2
```

[ ]:
```
data2.unstack()
```

[ ]:
```
data2.unstack().stack()
```

[ ]:
```
data2.unstack().stack(dropna=False)
```

1 Plotting & Visualization

[2]:
```
import matplotlib.pyplot as plt
```

[ ]:
```
plt.plot()
```

[ ]:
```
plt.plot([2,4,6,8, 10])
```

[ ]:
```
plt.plot([2,4,2,6,8,4])
```

[ ]:
```
salaries=[55000,65000,72000,90000,115000,150000]
ages = [20,25,30,32,40,45]
plt.plot(ages, salaries)
```

[ ]:
```
plt.plot(ages, salaries)
plt.show()
```

[ ]:
```
import numpy as np
nums = np.arange(5)
```

[ ]:
```
plt.plot(nums,nums)
plt.plot(nums, nums*nums)
plt.plot(nums, nums**3)
```

[ ]:
```
plt.plot(np.random.randn(50).cumsum(), 'k--')
```

[ ]:
```
np.random.randn(50).cumsum()
```

[ ]:
```
import pandas as pd
import numpy as np
fig = plt.figure()
ax1 = fig.add_subplot(2, 2, 1)
ax2 = fig.add_subplot(2, 2, 2)
ax3 = fig.add_subplot(2, 2, 3)
ax1.hist(np.random.randn(100), bins=20, color='k', alpha=0.4)
ax2.scatter(np.arange(30), np.arange(30) + 3 * np.random.randn(30))
plt.plot(np.random.randn(50).cumsum(), 'k--')
```

[ ]:
```
fig, axes = plt.subplots(2, 3)
```

[ ]:
```
import matplotlib.pyplot as plt
import numpy as np

#plot 1:
x = np.array([0, 1, 2, 3])
```

```
y = np.array([3, 8, 1, 10])

plt.subplot(1, 2, 1)
plt.plot(x,y)

#plot 2:
x = np.array([0, 1, 2, 3])
y = np.array([10, 20, 30, 40])

plt.subplot(1, 2, 2)
plt.plot(x,y)

plt.show()
```

[ ]:
```
#plot 1:
x = np.array([0, 1, 2, 3])
y = np.array([3, 8, 1, 10])

plt.subplot(2, 1, 1)
plt.plot(x,y)

#plot 2:
x = np.array([0, 1, 2, 3])
y = np.array([10, 20, 30, 40])

plt.subplot(2, 1, 2)
plt.plot(x,y)

plt.show()
```

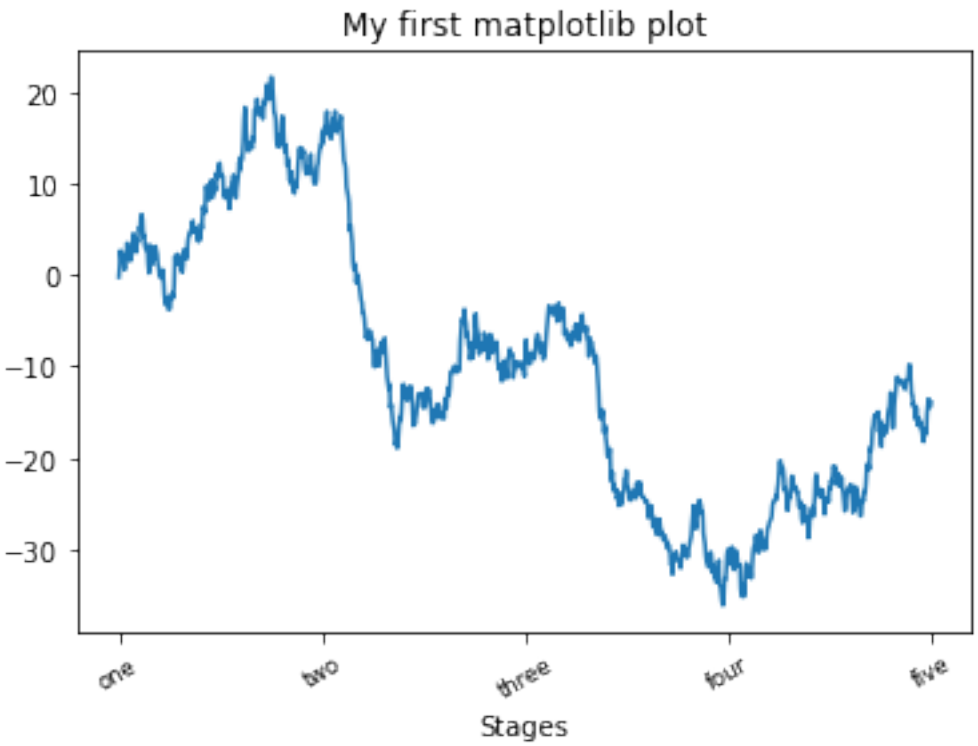
[ ]:
```
x = np.array([0, 1, 2, 3])
y = np.array([3, 8, 1, 10])

plt.subplot(2, 3, 1)
plt.plot(x,y)

x = np.array([0, 1, 2, 3])
y = np.array([10, 20, 30, 40])

plt.subplot(2, 3, 2)
plt.plot(x,y)

x = np.array([0, 1, 2, 3])
y = np.array([3, 8, 1, 10])

plt.subplot(2, 3, 3)
plt.plot(x,y)
```

```
x = np.array([0, 1, 2, 3])
y = np.array([10, 20, 30, 40])

plt.subplot(2, 3, 4)
plt.plot(x,y)

x = np.array([0, 1, 2, 3])
y = np.array([3, 8, 1, 10])

plt.subplot(2, 3, 5)
plt.plot(x,y)

x = np.array([0, 1, 2, 3])
y = np.array([10, 20, 30, 40])

plt.subplot(2, 3, 6)
plt.plot(x,y)
```

[ ]:
```
x = np.array([5,7,8,7,2,17,2,9,4,11,12,9,6])
y = np.array([99,86,87,88,111,86,103,87,94,78,77,85,86])
plt.scatter(x, y)

#day two, the age and speed of 15 cars:
x = np.array([2,2,8,1,15,8,12,9,7,3,11,4,7,14,12])
y = np.array([100,105,84,105,90,99,90,95,94,100,79,112,91,80,85])
plt.scatter(x, y)

plt.show()
```

[3]:
```
import matplotlib.pyplot as plt
import numpy as np
fig = plt.figure()
ax = fig.add_subplot(1, 1, 1)
ax.set_xticks([0, 250, 500, 750, 1000])
ax.set_title('My first matplotlib plot')
ax.set_xlabel('Stages')
ax.set_xticklabels(['one', 'two', 'three', 'four', 'five'], rotation=30,␣
 ↪fontsize='small')
ax.plot(np.random.randn(1000).cumsum())
```

[3]:
```
[<matplotlib.lines.Line2D at 0x24c9907a880>]
```



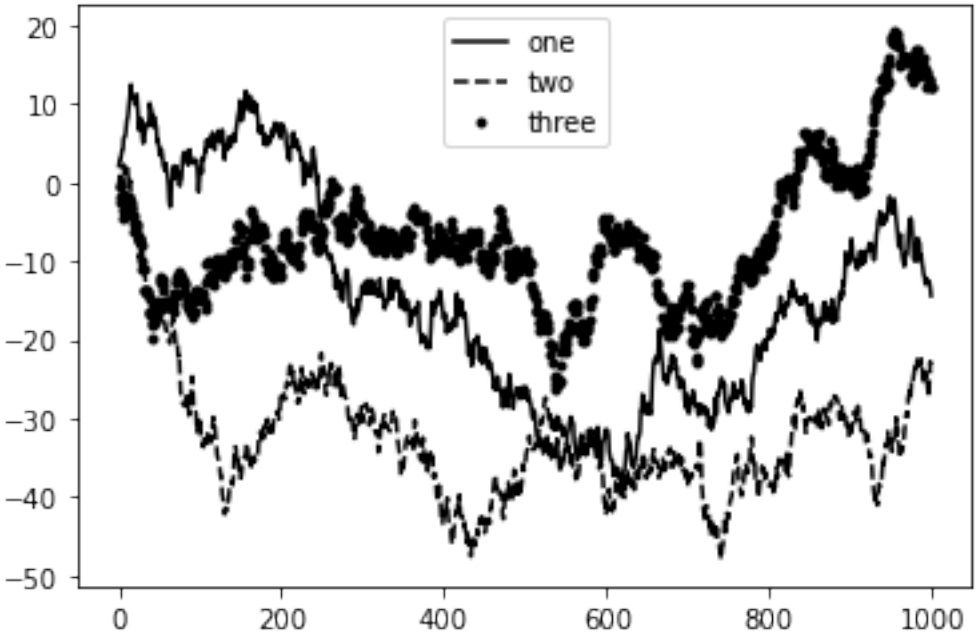
[ ]:
```
ax.set_xticks([0, 250, 500, 750, 1000])
```

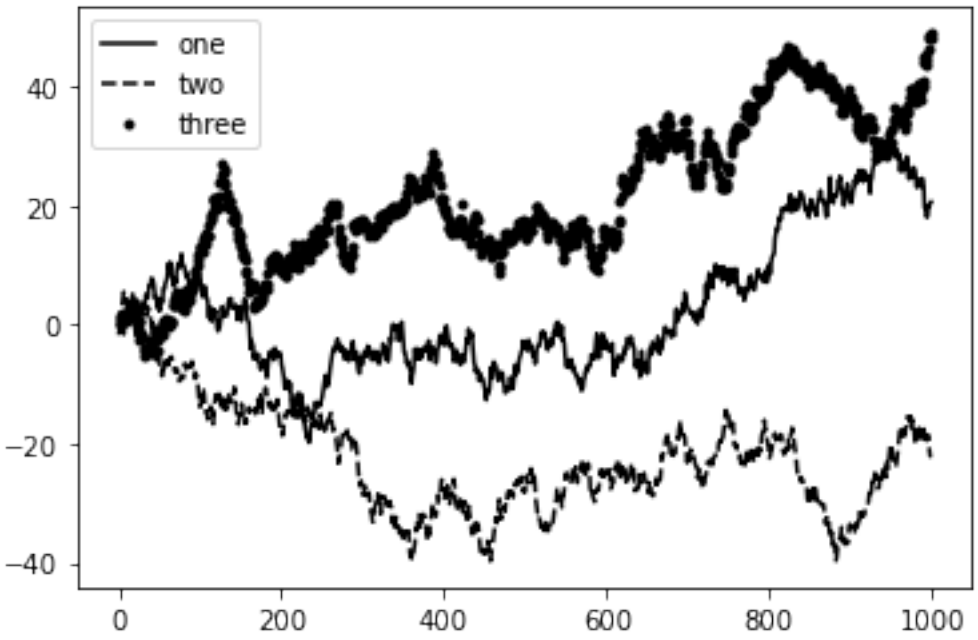
[ ]:
```
labels = ax.set_xticklabels(['one', 'two', 'three', 'four', 'five'],␣
 ↪rotation=30, fontsize='small')
```

[4]:
```
from numpy.random import randn
fig = plt.figure();
ax = fig.add_subplot(1, 1, 1)
ax.plot(randn(1000).cumsum(), 'k', label='one')

ax.plot(randn(1000).cumsum(), 'k--', label='two')

ax.plot(randn(1000).cumsum(), 'k.', label='three')

ax.legend(loc='best')

fig = plt.figure();

ax = fig.add_subplot(1, 1, 1)
ax.plot(randn(1000).cumsum(), 'k', label='one')

ax.plot(randn(1000).cumsum(), 'k--', label='two')
```

```
ax.plot(randn(1000).cumsum(), 'k.', label='three')

ax.legend(loc='best')
```

[4]: <matplotlib.legend.Legend at 0x24c99831f70>





[5]:
```
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.savefig('figpath.png', dpi=400, bbox_inches='tight')
```