

데이터 시각화 (2024)

데이터과학부 정진명

(jmjung@suwon.ac.kr, 글로벌경상관 918호)

4 주차

bar

bar

```
matplotlib.pyplot.bar(x, height, width=0.8, bottom=None, *, align='center', data=None, **kwargs)
```

[\[source\]](#)

Make a bar plot.

The bars are positioned at *x* with the given *alignment*. Their dimensions are given by *width* and *height*. The vertical baseline is *bottom* (default 0).

Each of *x*, *height*, *width*, and *bottom* may either be a scalar applying to all bars, or it may be a sequence of length N providing a separate value for each bar.

Parameters:

x : sequence of scalars

The x coordinates of the bars. See also *align* for the alignment of the bars to the coordinates.

height : scalar or sequence of scalars

The height(s) of the bars.

width : scalar or array-like, optional

The width(s) of the bars (default: 0.8).

bottom : scalar or array-like, optional

The y coordinate(s) of the bars bases (default: 0).

align : {'center', 'edge'}, optional, default: 'center'

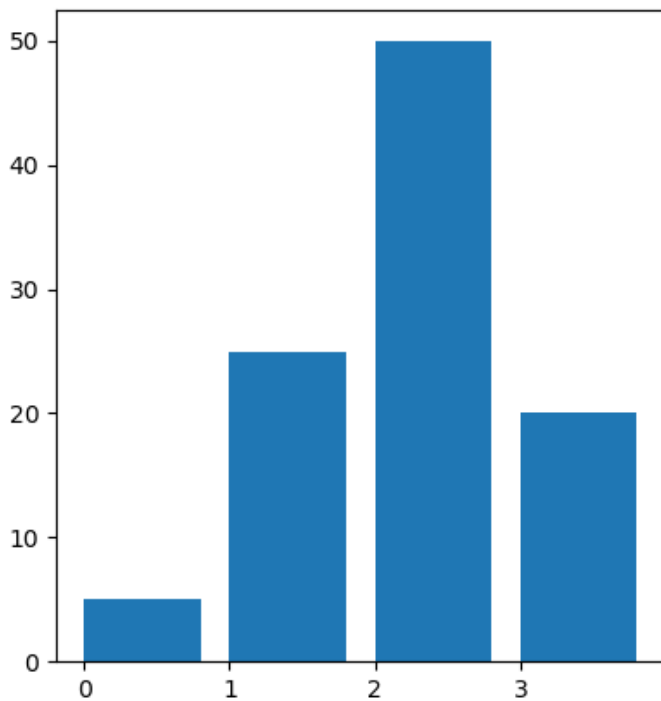
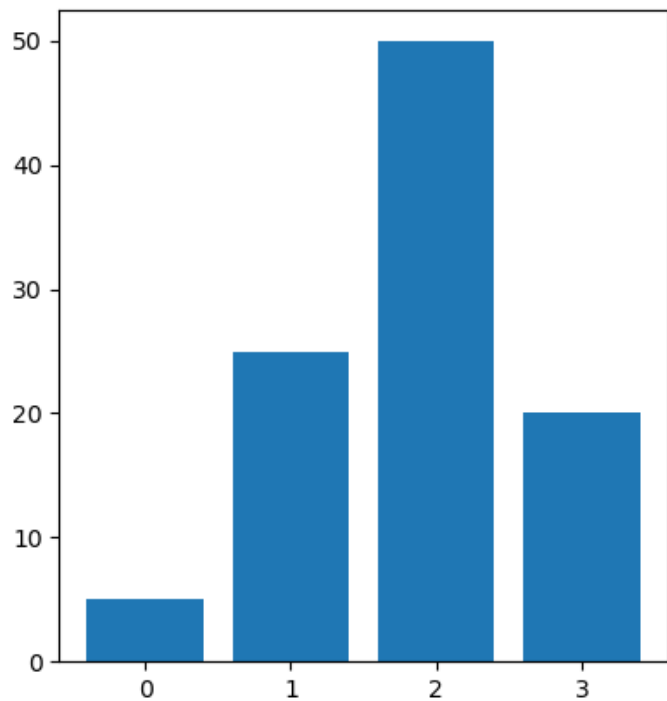
bar 예제 1

```
fig=plt.figure(figsize=(10,5), dpi=100)
axs=fig.subplots(1,2)

data = np.array([5,25,50,20])
axs[0].bar(range(len(data)),data)
axs[1].bar(range(len(data)),data, align='edge')
```

X

height



bar: various width

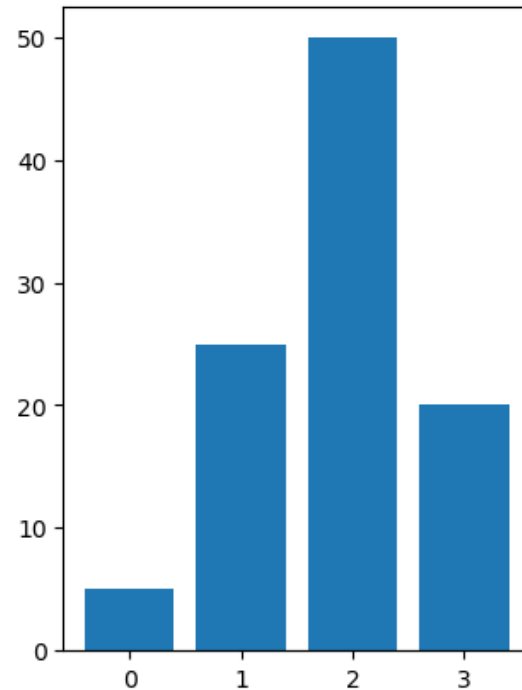
```
fig=plt.figure(figsize=(12,5), dpi=100)
axs=fig.subplots(1,3)
```

```
def bar_width(ax,data,w):
```

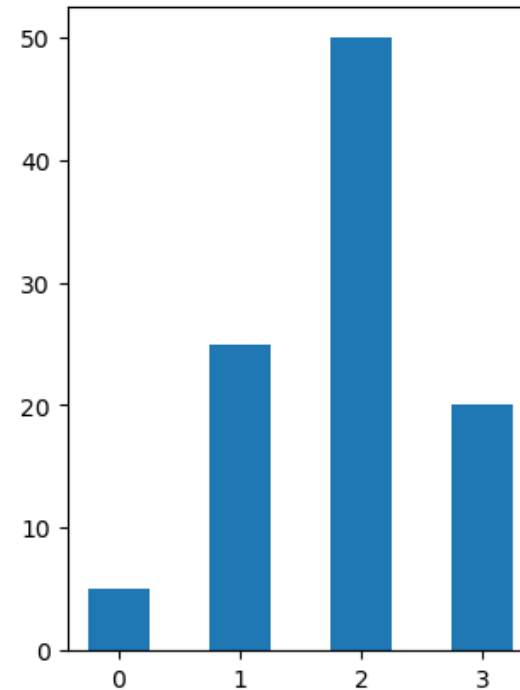
```
    ax.bar(range(len(data)),data, width=w)
    ax.set_title('width: {}'.format(w))
```

```
data = np.array([5,25,50,20])
bar_width(axs[0], data, w=0.8) # default width: 0.8
bar_width(axs[1], data, w=0.5)
bar_width(axs[2], data, w=1)
```

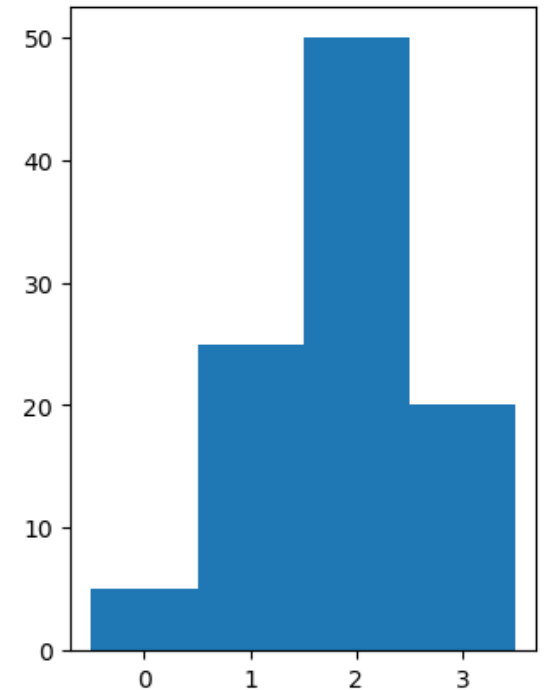
width: 0.8



width: 0.5



width: 1



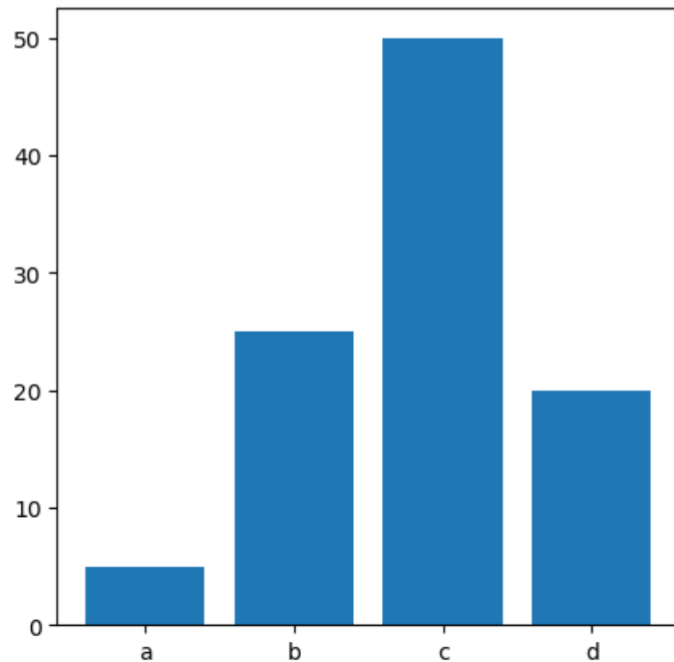
string X & series bar plot

3.3 string X

```
fig=plt.figure(figsize=(5,5), dpi=100)
ax=fig.subplots()

data = np.array([5,25,50,20])
X=list('abcd')
ax.bar(X,data)
```

<BarContainer object of 4 artists>



3.4 Series bar plot

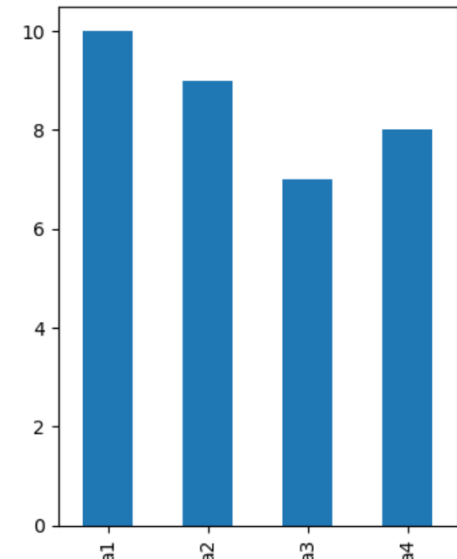
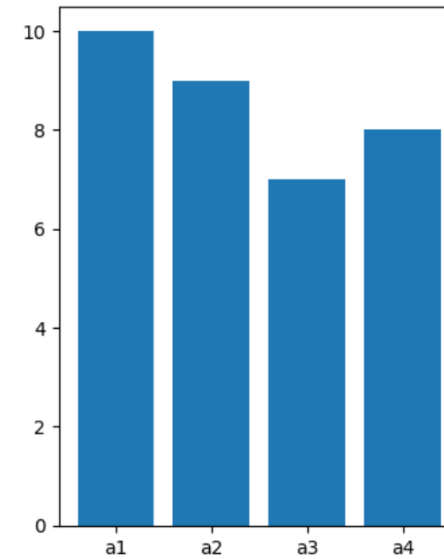
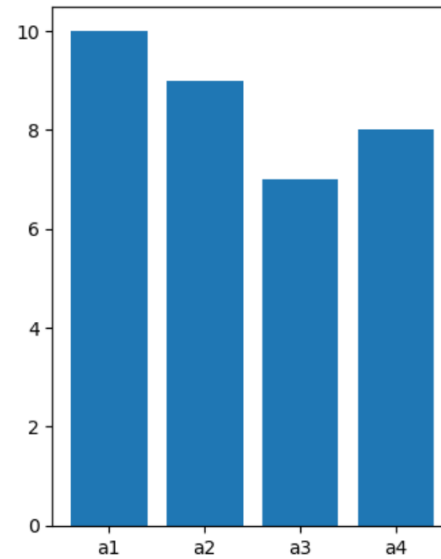
```
fig=plt.figure(figsize=(13,5), dpi=100)
axs=fig.subplots(1,3)

sr1=pd.Series(data=[10,9,7,8], index=['a1','a2','a3','a4'])
axs[0].bar(sr1.index, sr1)
axs[1].bar(sr1.index, sr1.values)
sr1.plot.bar(ax=axs[2])
```

<BarContainer object of 4 artists>

<BarContainer object of 4 artists>

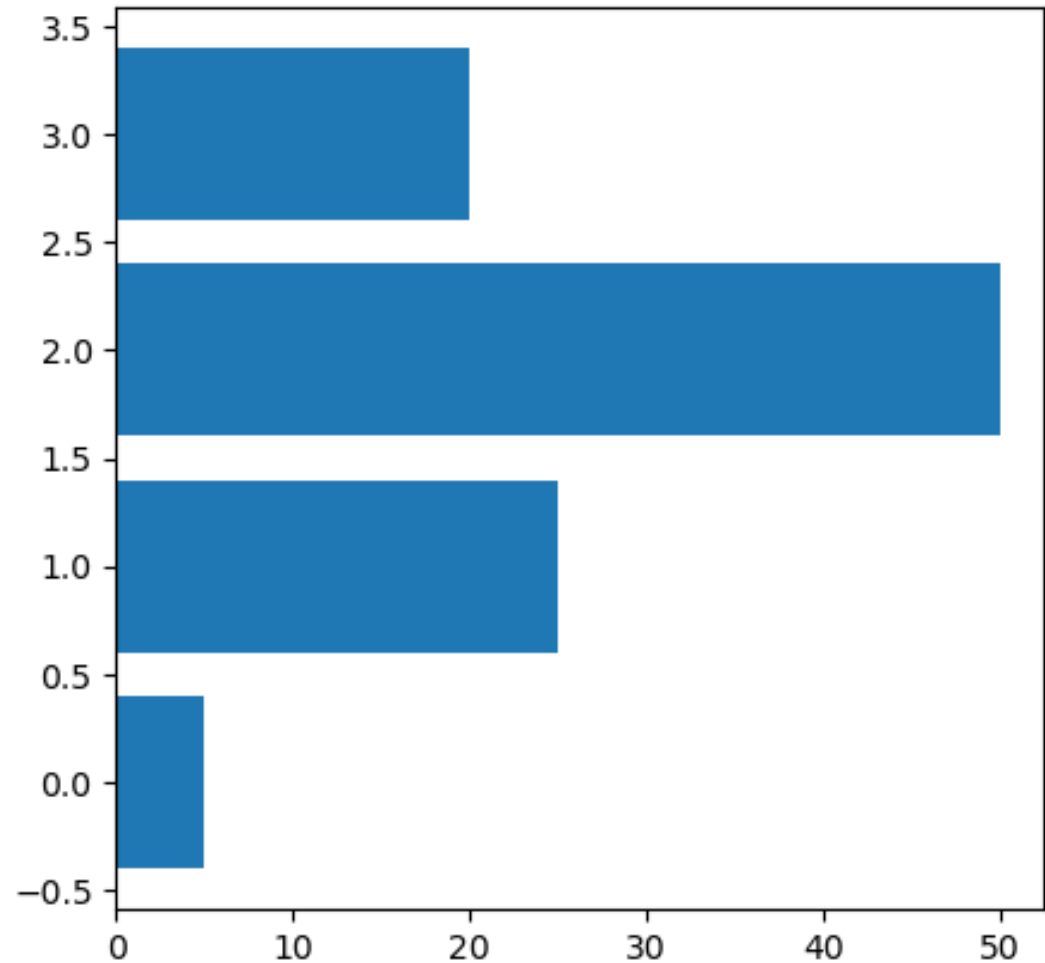
<Axes: >



수평막대 (horizontal bar: barh)

```
fig=plt.figure(figsize=(5,5), dpi=100)
ax=fig.subplots()

data = np.array([5,25,50,20])
ax.barh(range(len(data)),data)
```



다중막대차트 그리기

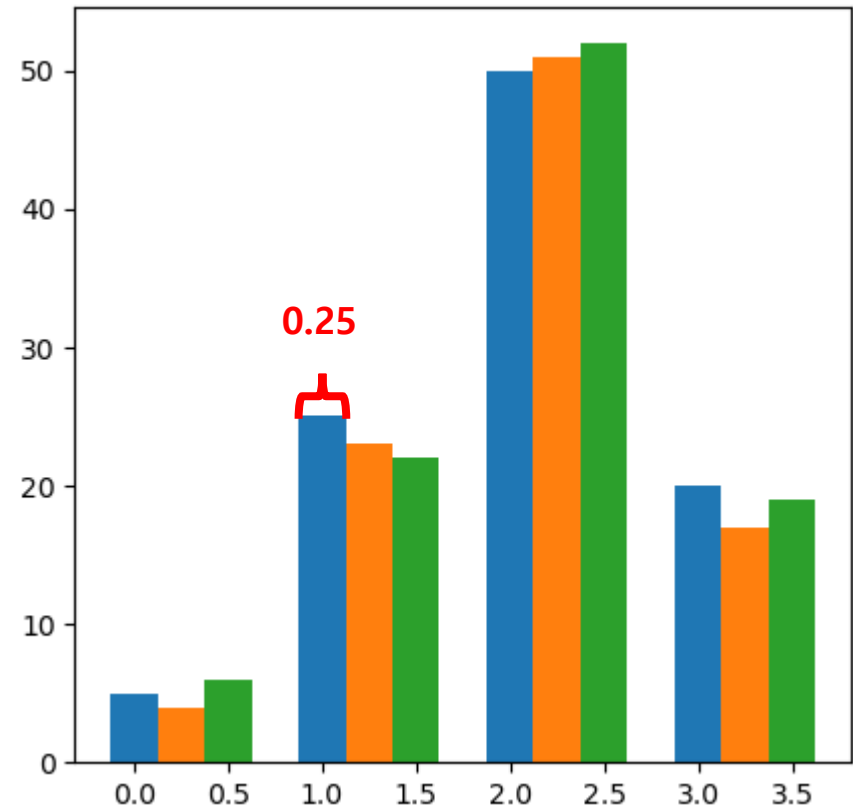
```
fig=plt.figure(figsize=(5,5), dpi=100)
ax=fig.subplots()

data=np.array([[5,25,50,20],
               [4,23,51,17],
               [6,22,52,19]])

X=np.arange(4)

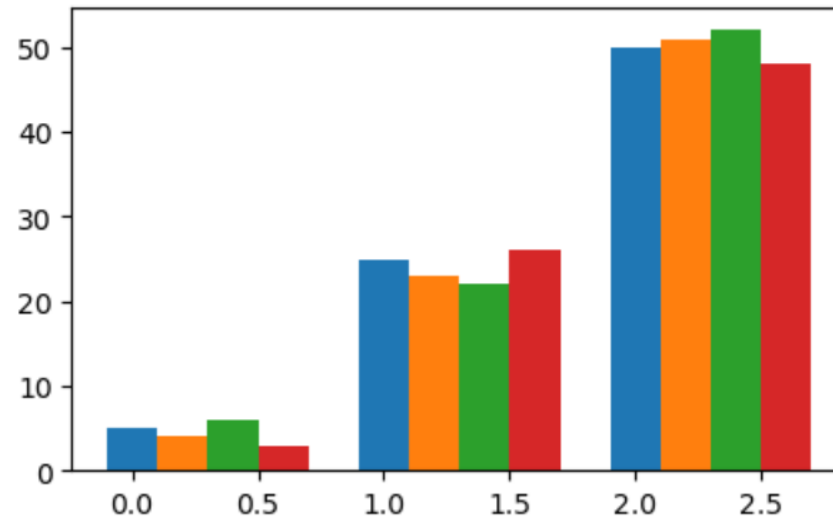
ax.bar(X+0.00, data[0], width=0.25)
ax.bar(X+0.25, data[1], width=0.25)
ax.bar(X+0.50, data[2], width=0.25)
```

직접 color 지정하지 않으면 미리 정해진 color 사용



실습 1

- for loop을 사용하여 아래와 같은 bar graph를 그리시오 (실습 1)



dataframe에서 특정 column, row selection

```
data=pd.read_table('dat_bar.txt',sep='\t', index_col=0)  
data
```

```
data.loc[:, 'iphone']  
data['iphone']  
data.iloc[:,0]
```

특정 column 선택
(column 이름 또는 위치)

```
data.loc['jan',:]  
data.loc['jan']  
data.iloc[0,:]  
data.iloc[0]
```

특정 row(index) 선택
(index 이름 또는 위치)

jan 76
feb 81
mar 84
apr 91
may 90
jun 88
jul 81
aug 73
sep 71
oct 63
nov 61
dec 57
Name: iphone, dtype: int64

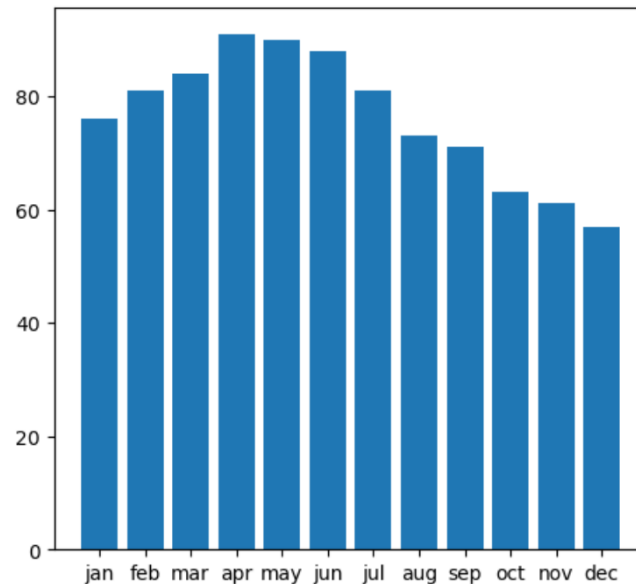
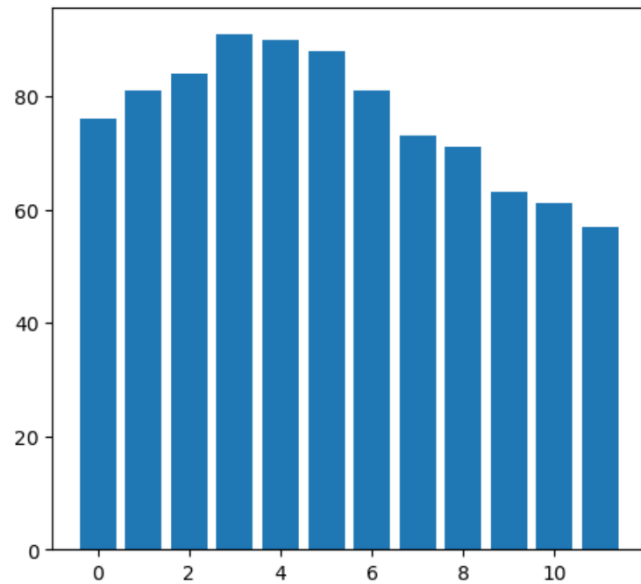
iphone 76
galaxy 54
Name: jan, dtype: int64

| | iphone | galaxy |
|-----|--------|--------|
| jan | 76 | 54 |
| feb | 81 | 47 |
| mar | 84 | 42 |
| apr | 91 | 35 |
| may | 90 | 36 |
| jun | 88 | 57 |
| jul | 81 | 65 |
| aug | 73 | 75 |
| sep | 71 | 80 |
| oct | 63 | 82 |
| nov | 61 | 85 |
| dec | 57 | 90 |

파일(dat_bar.txt)로부터 bar 그리기

```
fig=plt.figure(figsize=(12,5), dpi=100)
ax1, ax2=fig.subplots(1,2)

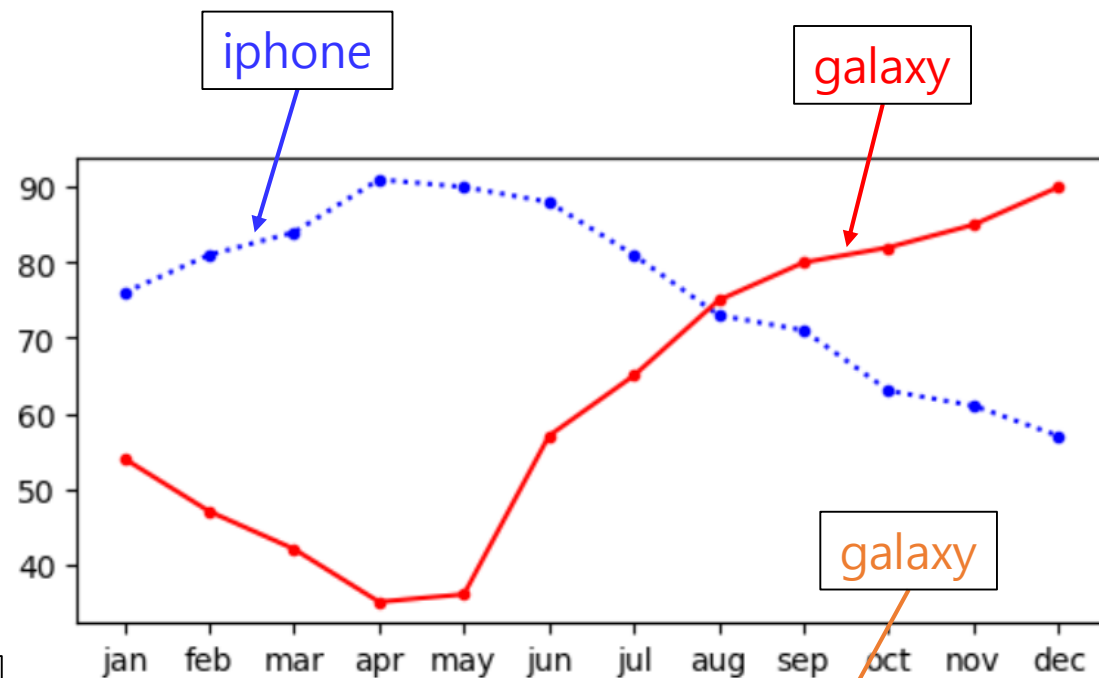
data=pd.read_table('data/dat_bar.txt',sep='\t', index_col=0)
data
ax1.bar(range(len(data)), data['iphone'])
ax2.bar(data['iphone'].index, data['iphone'])
```



| | iphone | galaxy |
|-----|--------|--------|
| jan | 76 | 54 |
| feb | 81 | 47 |
| mar | 84 | 42 |
| apr | 91 | 35 |
| may | 90 | 36 |
| jun | 88 | 57 |
| jul | 81 | 65 |
| aug | 73 | 75 |
| sep | 71 | 80 |
| oct | 63 | 82 |
| nov | 61 | 85 |
| dec | 57 | 90 |

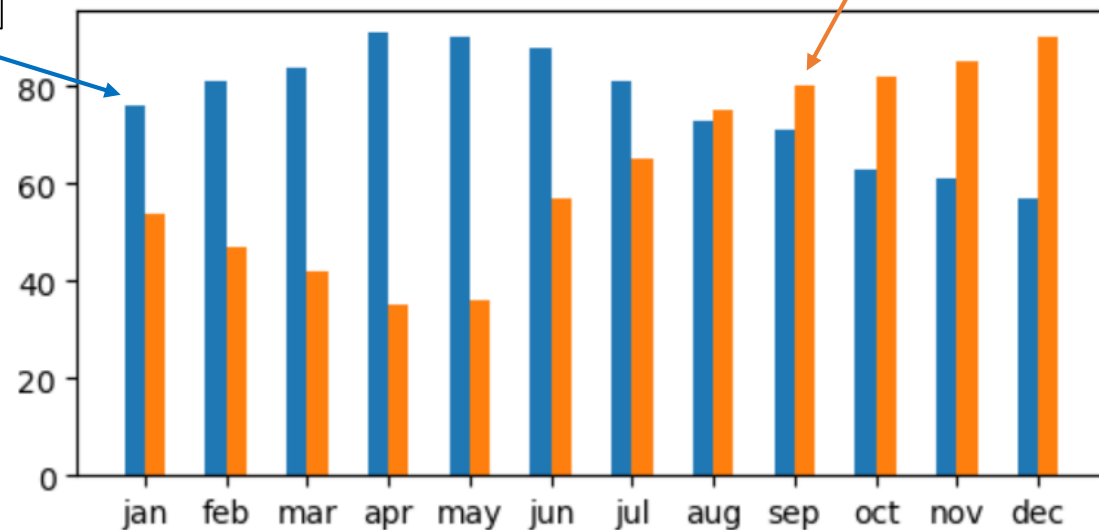
실습 2

- dat_bar.txt 파일을 읽어서 아래와 같은 그림을 그리시오



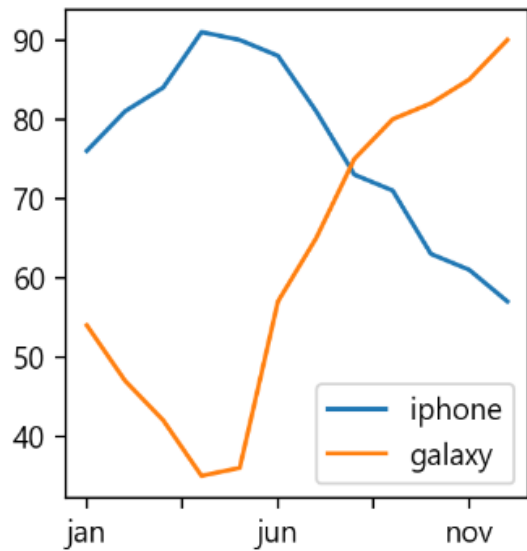
iphone

```
_ax2.set_xticks(ticks=X+0.125, labels=data.index)
```

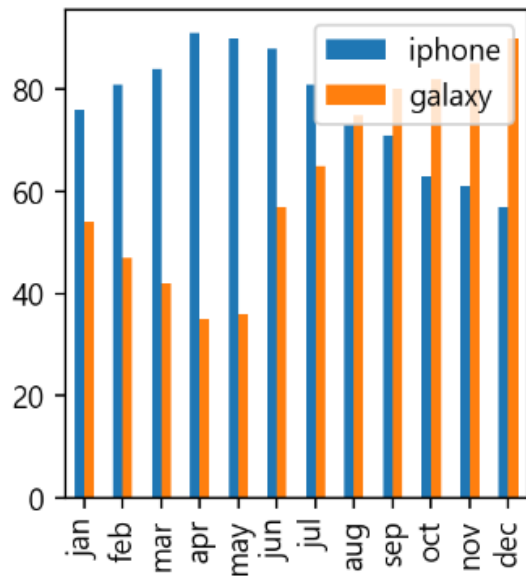


실습 3

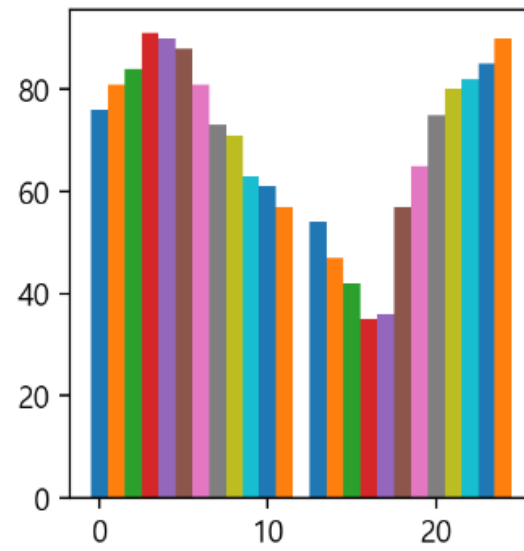
- 아래의 dat_bar.txt를 읽어서 오른쪽과 같은 bar plot을 그리시오



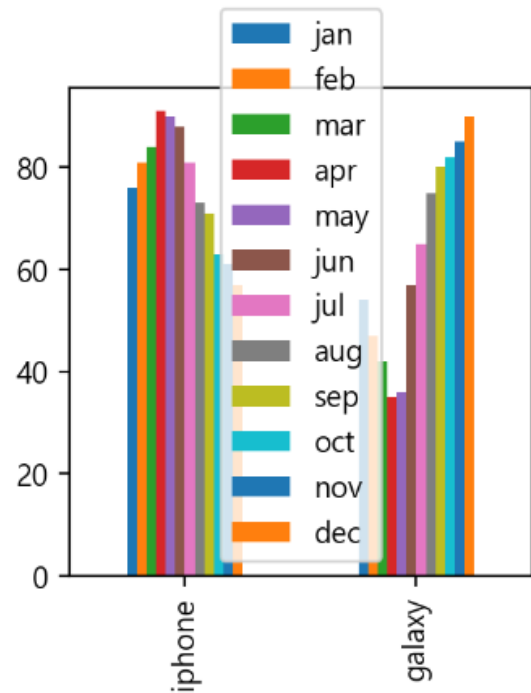
pandas.plot.line



pandas.plot.bar



axe.bar
(loc으로 각 행 접근)



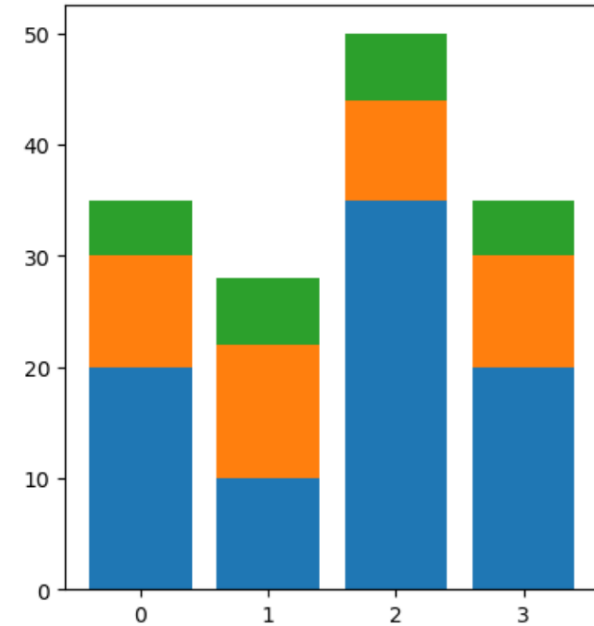
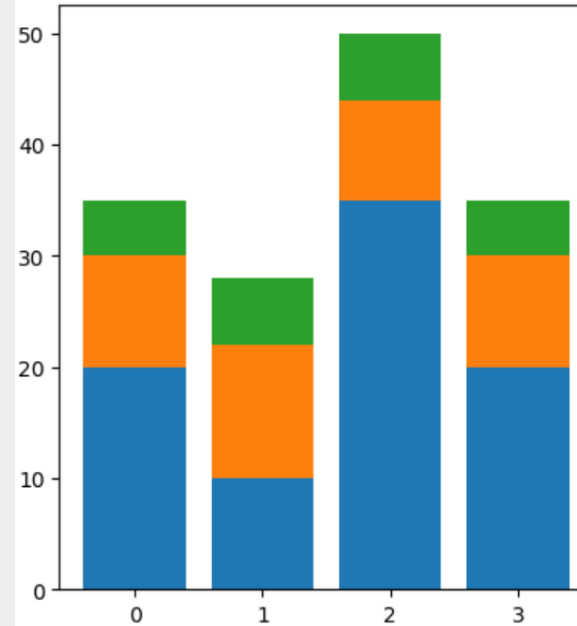
pandas.plot.bar

분할 막대차트 그리기

```
fig=plt.figure(figsize=(10,5), dpi=100)
ax1,ax2=fig.subplots(1,2)
|
data=np.array([[20,10,35,20],
               [10,12,9,10],
               [5,6,6,5]])
X=np.arange(4)

_=ax1.bar(X,data[0])
_=ax1.bar(X,data[1],bottom=data[0])
_=ax1.bar(X,data[2],bottom=data[0]+data[1])
      A+B: numpy element-wise 덧셈

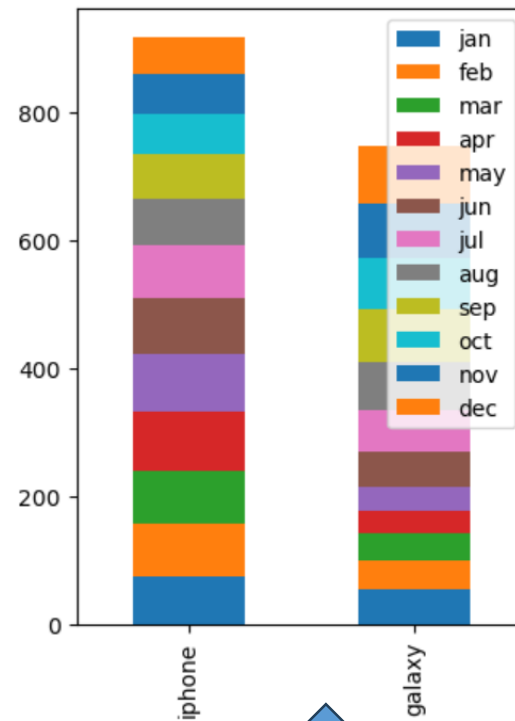
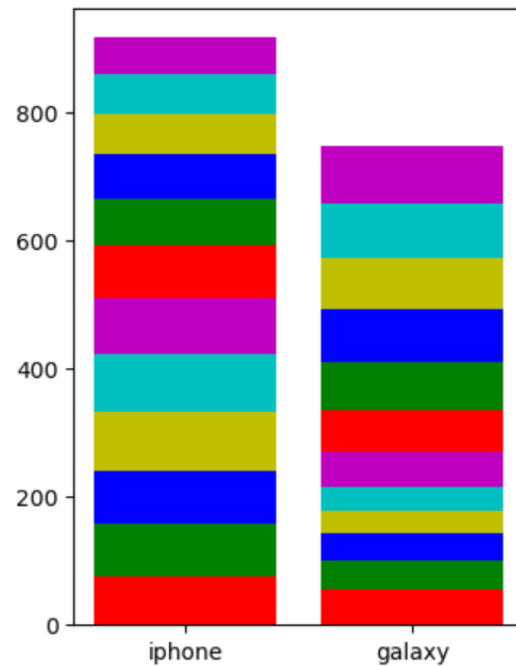
for i in range(len(data)):
    _=ax2.bar(X, data[i], bottom=np.sum(data[:i], axis=0))
```



실습 4

- 아래의 dat_bar.txt를 읽어서 오른쪽과 같은 분할 막대차트를 그리시오

| | iphone | galaxy |
|-----|--------|--------|
| jan | 76 | 54 |
| feb | 81 | 47 |
| mar | 84 | 42 |
| apr | 91 | 35 |
| may | 90 | 36 |
| jun | 88 | 57 |
| jul | 81 | 65 |
| aug | 73 | 75 |
| sep | 71 | 80 |
| oct | 63 | 82 |
| nov | 61 | 85 |
| dec | 57 | 90 |



Color_list의 12개의 값을 jan ~ dec 까지 순차적으로 사용하시오

dataframe.plot.bar

Q & A

Thank you