

[Open in app](#)

Published in Python Everywhere -from Beginner to Advanced



Sean Yeh

[Follow](#)

May 11, 2021 · 6 min read



Save



Matplotlib資料視覺化-2 柱狀圖與直方圖



Hengchun, Taiwan ,photo by Sean Yeh

在資料視覺化時，常用的圖表類型有折線圖、散佈圖、柱狀圖、直方圖與圓餅圖等。其中有兩種圖看起來很類似，卻不相同。我們就先看看這兩種圖形：柱狀圖與直方圖。

Bar chart與Histogram chart不同

柱狀圖（又稱長條圖），它可以用來呈現多筆離散的資料，資料與資料之間彼此不相關，也沒有一定的順序關係。柱狀圖（bar chart）有一個長相看似相近的兄弟，叫做直方圖（histogram），兩者在資料表達上，是有差異的，大家要特別注意，可不要搞混了。

柱狀圖(Bar graph)

柱狀圖（又稱長條圖）描述的是確切的類別，並說明正類別各組資料的大小。由於柱狀圖的類



1



[Open in app](#)

直方圖histogram

直方圖所描述的數據資料是連續性的 (continuous datas)，也就是說在X軸各個區間之間是有順序性的。比如說年齡從0歲一直到100歲，或者是金額從0元一直到100000元，身高、體重等等都屬於這種。

直方圖的資料間隔 (bin) 可以依照需要隨意的調整，間隔調整後並不會因此破壞原先的數據。調整間隔的目的只是讓我們更容易瀏覽。以年齡為例，我們可以以五歲為一個級距，也可以以十歲為一個級距甚至於使用20歲為一個級距。可以想像為將資料分裝置不同大小的量杯裡面，我們可以依照需要選擇量杯的大小。

柱狀圖 (bar chart) 的繪製方式

繪製柱狀圖的方式基本上與前面的折線圖類似，透過bar的方式可以繪製柱狀圖。語法如下：

```
plt.bar(x,y)
```

只要傳入x、y軸的資料，就可以繪製出柱狀圖。其中x、y軸的資料，通常為1D 陣列。

例如：我們透過下面的x、y資料繪製出柱狀圖。

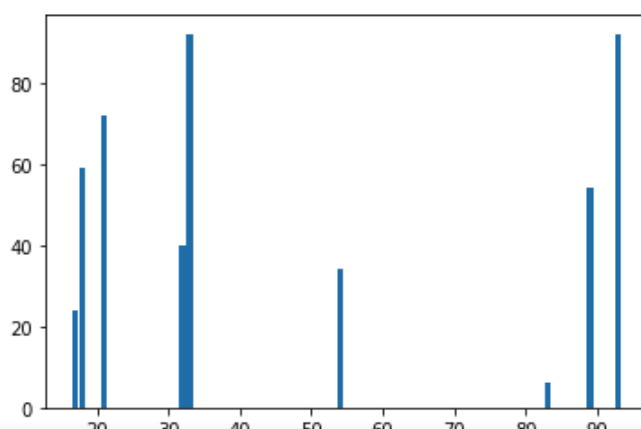
```
import matplotlib.pyplot as plt

x = np.random.randint(1000, size=10)

y = np.random.randint(1000, size=10)

plt.bar(x,y)

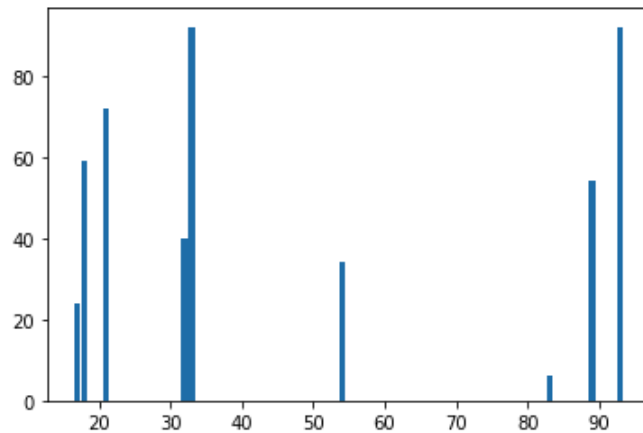
plt.show()
```



[Open in app](#)

```
plt.bar(x, y, tick_label=['A', 'B', 'C', 'D', 'E', 'F', 'G', 'H', 'I', 'J'])
```

設定後結果如下：



tick_label的值為串列list，這個串列裡面含有標籤的名稱。我們也可以把它寫在外面。

```
label = ['A', 'B', 'C', 'D', 'E', 'F', 'G', 'H', 'I', 'J']  
  
plt.bar(x, y, tick_label=label)
```

直方圖 (histogram) 的繪製方式

直方圖用來呈現連續資料間的次數，可以看出資料的分佈情況。使用Matplotlib繪製時語法如下：

```
plt.hist(x, bins)
```

我們可以在hist()裡面放入兩個參數x與bins。其中，x為我們想要畫出的資料。bins為橫軸的組距。若指定為auto則表示自動決定。

第一個例子

如下面例子，我們用numpy隨機取出0到100的數字10個，放入x變數中。並使用hist()製作直方圖，組距設為10。

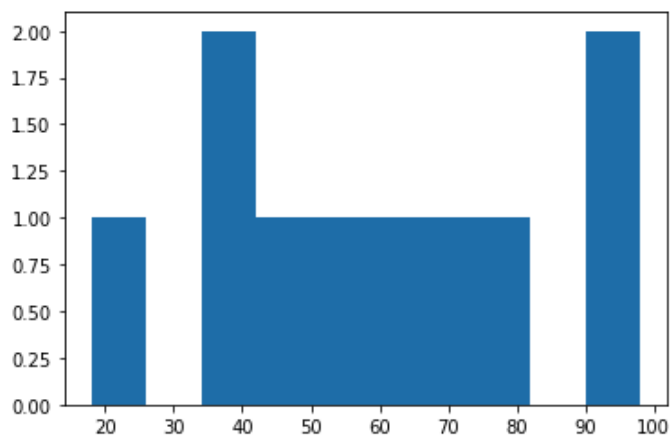
```
import matplotlib.pyplot as plt
```



[Open in app](#)

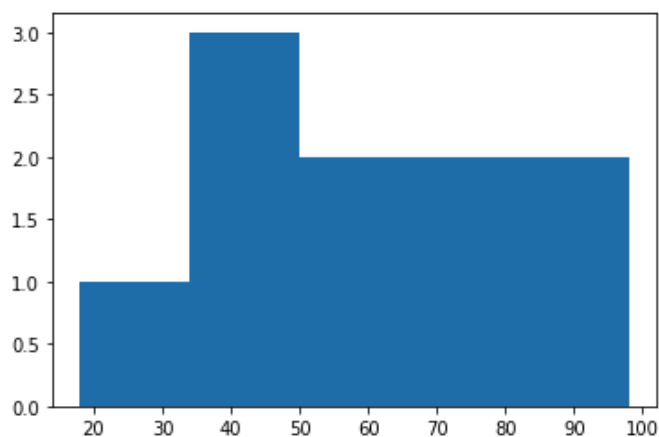
結果顯示：

```
(array([1., 0., 2., 1., 1., 1., 1., 1., 0., 2.]),  
 array([18., 26., 34., 42., 50., 58., 66., 74., 82., 90., 98.]),  
 <BarContainer object of 10 artists>)
```



使用auto的話，呈現的結果如下：

```
plt.hist(x, bins='auto')
```



第二個例子

在下面的例子，我們用random製作一個具備10,000個亂數的z值：

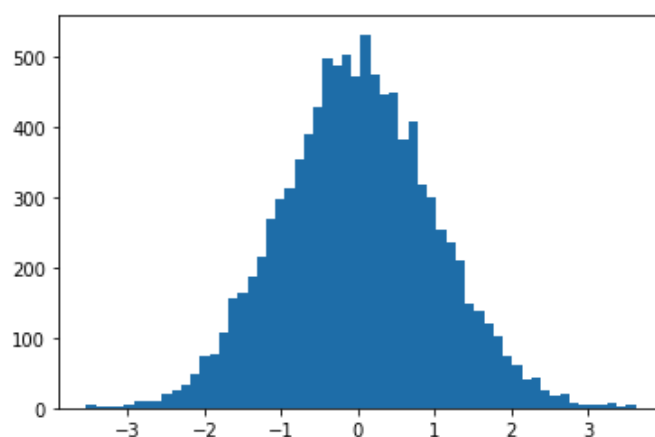
```
z = np.random.randn(10000)
```

先使用自動組距畫出直方圖。



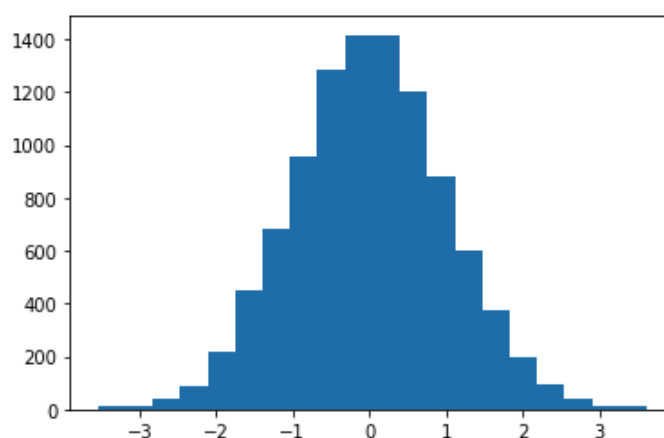
[Open in app](#)

結果呈現吊鐘型分佈，也就是統計上的常態分佈（normal distribution）。左右兩邊對稱。



如果換成20為組距，結果如下：

```
plt.hist(z, bins=20)
```

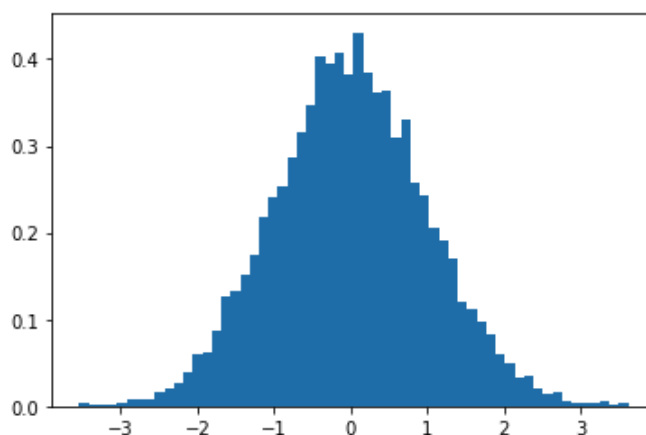


density 機率密度

```
plt.hist(z, bins='auto', density=True)
```

圖中 Y 軸的數字為機率，其中0.4是指40%的意思。



[Open in app](#)

結語

上一篇我們說明了各種基礎的繪圖概念，在這篇裡面我們進一步討論了柱狀圖與直方圖的差異點。由於大家常常會使用到這兩種圖形，因此必須要了解兩者間的差異性，才可以適當的選擇圖形表現方式。

延伸閱讀

Matplotlib 讓資料視覺化-1

在Python的世界中，處理資料視覺化的工具很多，其中Matplotlib可以說是一切的基礎。Matplotlib是個Python套件，在Python繪圖領域中被廣泛使用。

medium.com

Get an email whenever Sean Yeh publishes.

Subscribe

Emails will be sent to jfjan2005@gmail.com.

[Not you?](#)



