

# Data Visualization with Matplotlib - Exercises

จงทำตามคำสั่งต่อไปนี้ด้วย data ที่กำหนดให้ต่อไปนี้

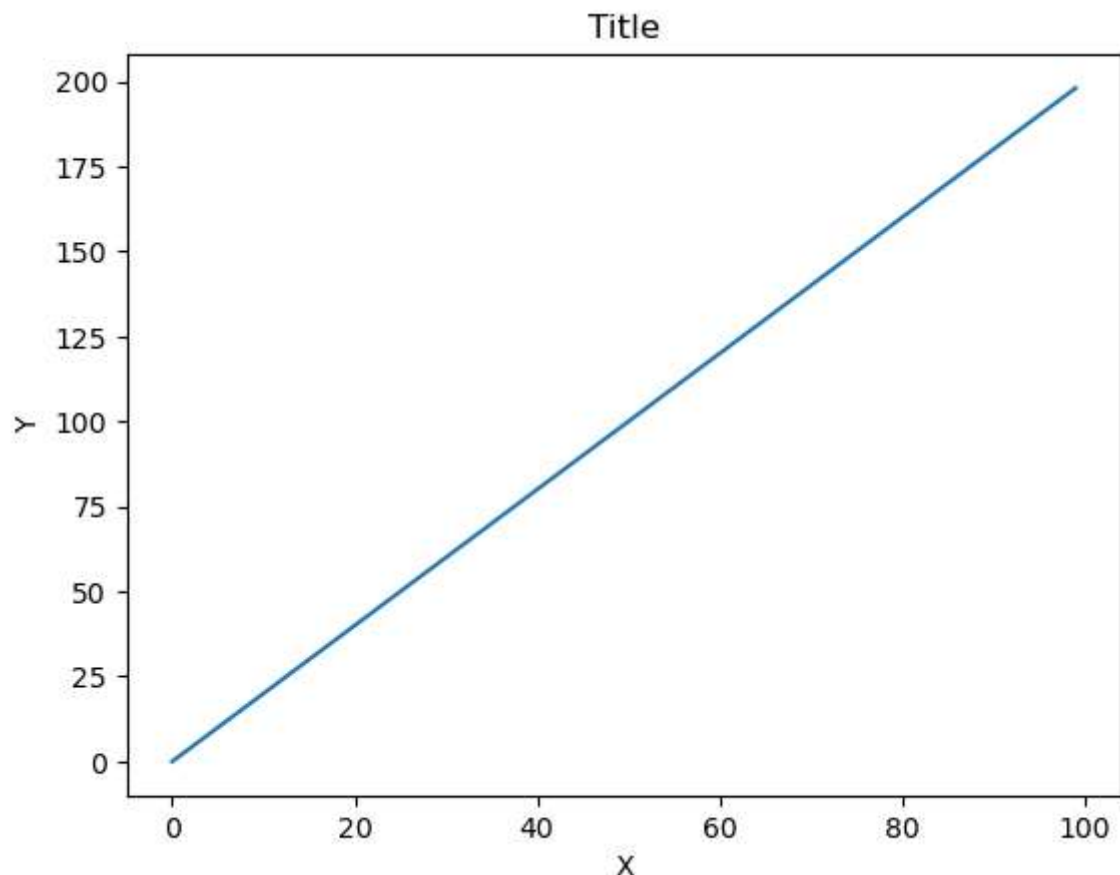
## Data

```
In [39]: import matplotlib.pyplot as plt
import numpy as np
import pandas as pd
x = np.arange(0,100)
y = x*2
z = x**2
df = pd.read_csv('Superstore.csv',encoding = 'iso-8859-1')
```

## Exercise 1

```
In [40]: plt.plot(x,y)
plt.title('Title')
plt.xlabel('X')
plt.ylabel('Y')
```

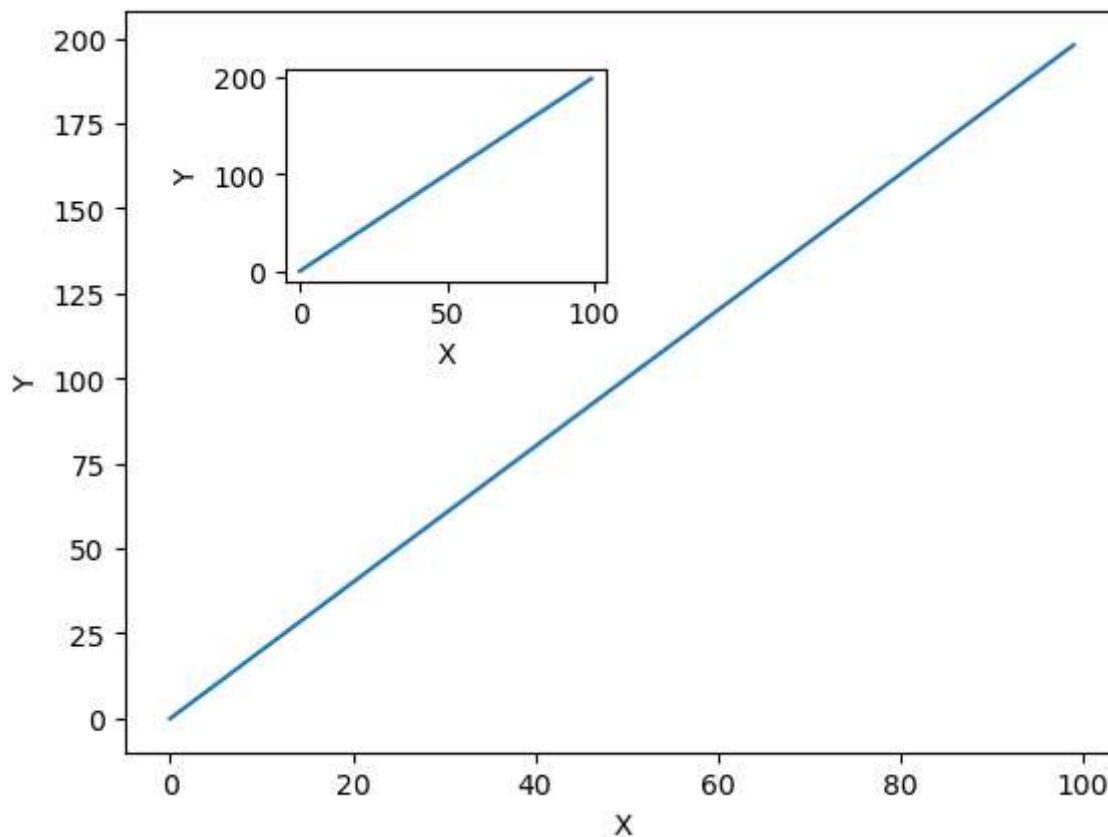
Out[40]: Text(0, 0.5, 'Y')



## Exercise 2

```
In [41]: fig = plt.figure()

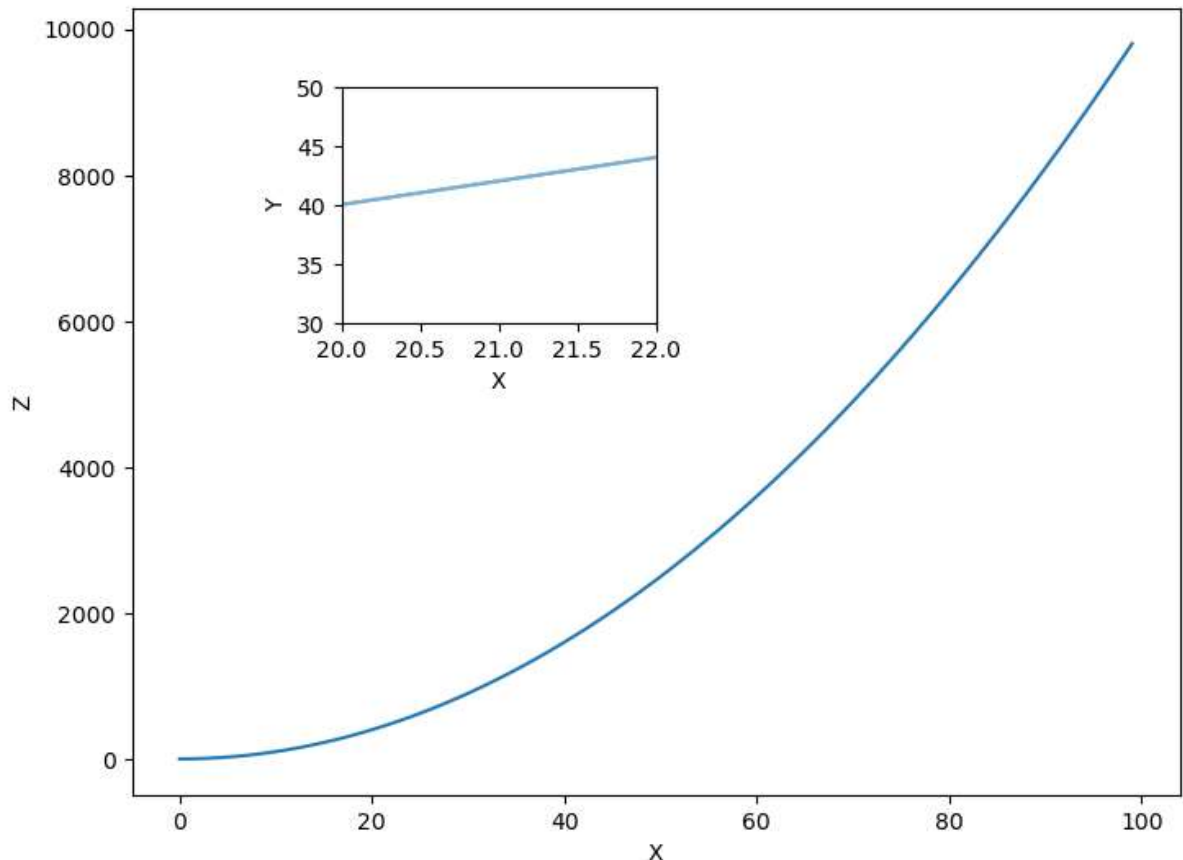
plt.plot(x,y)
plt.xlabel('X')
plt.ylabel('Y')
axes1 = plt.axes([.25,.6,.25,.22])
axes1 = plt.xlabel('X')
axes1 = plt.ylabel('Y')
axes1 = plt.plot(x,y)
```



## Exercise 3

ใช้ arrays x, y และ z เพื่อทำการ plot บนแกนที่สร้างจากข้อที่แล้ว (Notice อย่าลืมกำหนด x - limits และ y - limits )

```
In [52]: fig = plt.figure()
axes1 = fig.add_axes([0,0,1,1])
axes1.plot(x,z)
axes1 = plt.xlabel('X')
axes1 = plt.ylabel('Z')
axes2 = fig.add_axes([0.2,0.6,0.3,0.3])
axes2.plot(x,y,y,color = '#7AAED2')
axes2.set_ylim(30,50)
axes2.set_xlim(20,22)
axes2 = plt.xlabel('X')
axes2 = plt.ylabel('Y')
```

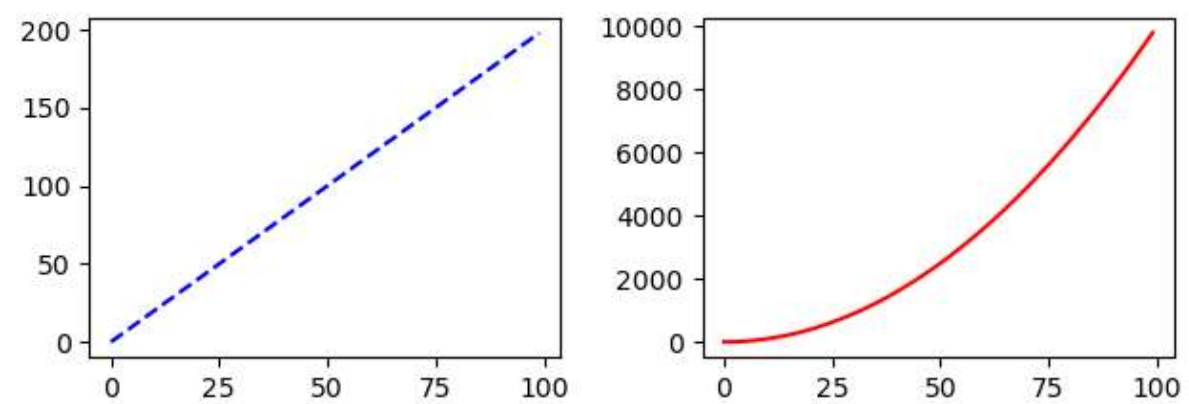


## Exercise 4

จงใช้คำสั่ง `plt.subplots(nrows=1, ncols=2)`

จากนั้นให้ทำการ plot (x,y) และ plot (x,z) บนแกน axes และให้ใช้งานคำสั่ง `linewidth` and `style` เพื่อตกแต่งเส้นของกราฟ

```
In [62]: fig,axes = plt.subplots(nrows=1, ncols=2)
axes[0].plot(x,y,'--b')
axes[1].plot(x,z,'r')
fig.tight_layout();
fig.set_figheight(2);
```



Exercise 5

```
In [63]: df.head()
```

Out[63]:

	Order ID	Customer Name	Segment	Day	Month	Year	Ship Mode	City	State	Category	C
0	CA-2016-152156	Claire Gute	Consumer	8	11	2016	Second Class	Henderson	Kentucky	Furniture	Bo
1	CA-2016-152156	Claire Gute	Consumer	8	11	2016	Second Class	Henderson	Kentucky	Furniture	
2	CA-2016-138688	Darrin Van Huff	Corporate	12	6	2016	Second Class	Los Angeles	California	Office Supplies	
3	US-2015-108966	Sean O'Donnell	Consumer	11	10	2015	Standard Class	Fort Lauderdale	Florida	Furniture	
4	US-2015-108966	Sean O'Donnell	Consumer	11	10	2015	Standard Class	Fort Lauderdale	Florida	Office Supplies	

```
In [64]: df.info()
```

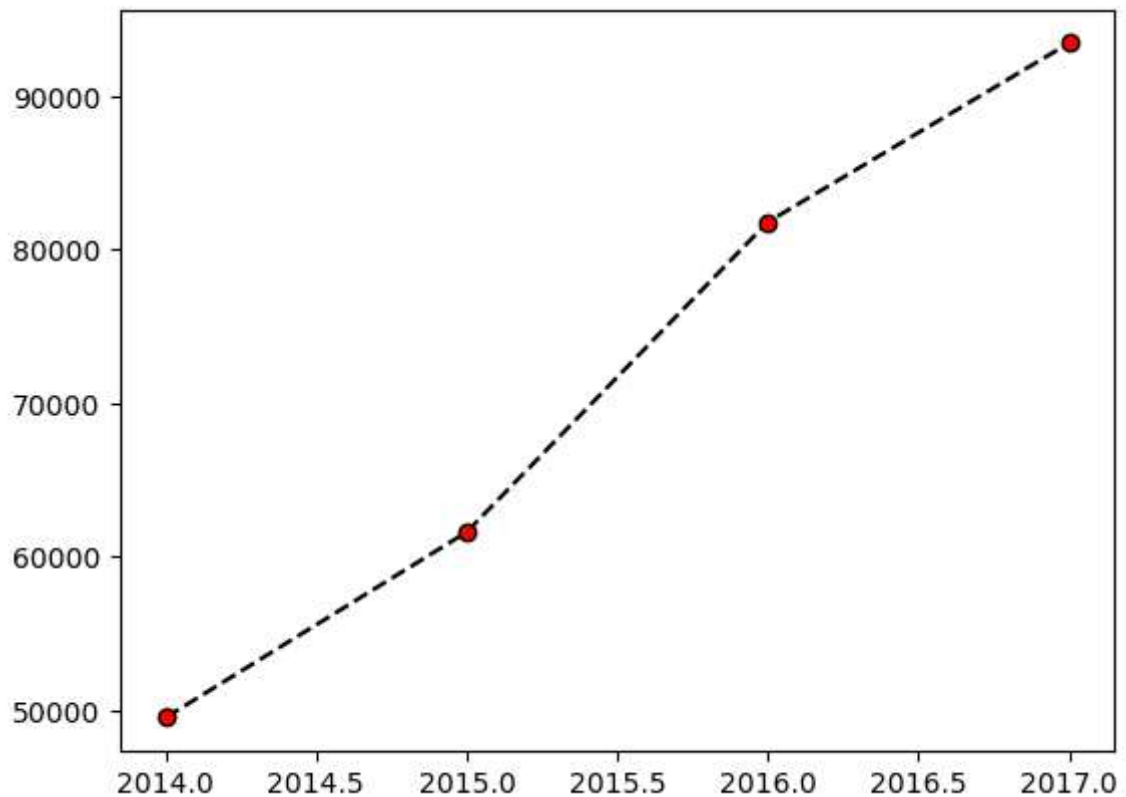
```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 9994 entries, 0 to 9993
Data columns (total 16 columns):
 #   Column              Non-Null Count  Dtype  
---  -
 0   Order ID            9994 non-null   object 
 1   Customer Name       9994 non-null   object 
 2   Segment             9994 non-null   object 
 3   Day                 9994 non-null   int64  
 4   Month               9994 non-null   int64  
 5   Year                9994 non-null   int64  
 6   Ship Mode           9994 non-null   object 
 7   City                9994 non-null   object 
 8   State               9994 non-null   object 
 9   Category            9994 non-null   object 
10   Sub-Category        9994 non-null   object 
11   Product Name        9994 non-null   object 
12   Sales               9994 non-null   float64 
13   Quantity            9994 non-null   int64  
14   Discount            9994 non-null   float64 
15   Profit              9994 non-null   float64 
dtypes: float64(3), int64(4), object(9)
memory usage: 1.2+ MB
```

## จงแสดงกราฟรายได้ของทุกปี

```
In [45]: df1 = df.groupby('Year')['Profit'].sum()
a = df1.index
b = df1
```

```
In [83]: plt.plot(a,b,'--k',marker='o',mfc='red')
```

```
Out[83]: [<matplotlib.lines.Line2D at 0x267baa37790>]
```



โค้ดต่อไปนี้จะใช้ทั้งสองข้อสุดท้าย

```
In [84]: df['Category'].unique()
```

```
Out[84]: array(['Furniture', 'Office Supplies', 'Technology'], dtype=object)
```

```
In [85]: df[ df['Category'] == 'Furniture' ].groupby('Year').sum()['Profit']
```

```
Out[85]: Year
2014    5457.7255
2015    3015.2029
2016    6959.9531
2017    3018.3913
Name: Profit, dtype: float64
```

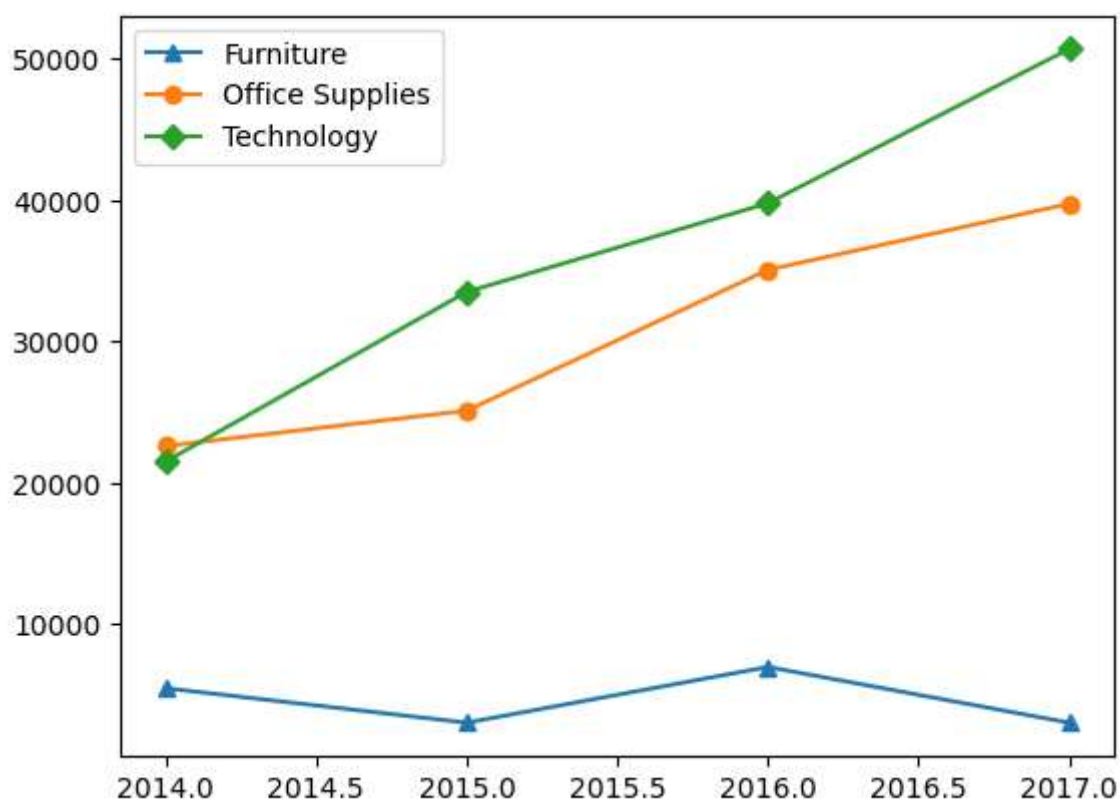
### Dictionary of Category

```
In [86]: arr_df = {}
for i in range(0,df['Category'].nunique()) :
    arr_df[df['Category'].unique()[i]] = df[ df['Category'] == df['Category'].u
```

## จงแสดงกราฟรายได้ของแต่ละ Category ในแต่ละปีในกราฟเดียว

```
In [102]: x = arr_df['Furniture'].index
y = arr_df['Furniture']
x1 = arr_df['Office Supplies'].index
y1 = arr_df['Office Supplies']
x2 = arr_df['Technology'].index
y2 = arr_df['Technology']
plt.plot(x,y,marker='^',label='Furniture')
plt.plot(x1,y1,marker='o',label='Office Supplies')
plt.plot(x2,y2,marker='D',label='Technology')
plt.legend(loc='best')
```

Out[102]: <matplotlib.legend.Legend at 0x267bc10e9d0>



## จแสดงกราฟรายได้ของแต่ละ Category ในแต่ละปี แบบแยกกราฟ

```
In [106]: fig = plt.figure()
axes1 = fig.add_axes([0.5,1.2,1,1])
axes1.plot(arr_df["Technology"].index,arr_df["Technology"],"s:g",label="Technol")
axes1.set_title("Technology")
axes1.set_xticks(np.arange(2014,2018,1))

axes2 = fig.add_axes([0,0,1,1])
axes2.plot(arr_df["Furniture"].index,arr_df["Furniture"],"^-b",label="Furnitur")
axes2.set_title("Furniture")
axes2.set_xticks(np.arange(2014,2018,1))

axes3 = fig.add_axes([1.1,0,1,1])
axes3.plot(arr_df["Office Supplies"].index,arr_df["Office Supplies"],"s-.r",label="Office Supplies")
axes3.set_title("Office Supplies")
axes3.set_xticks(np.arange(2014,2018,1))
```

```
Out[106]: [<matplotlib.axis.XTick at 0x267bdb42010>,
<matplotlib.axis.XTick at 0x267bdb40890>,
<matplotlib.axis.XTick at 0x267bd637d90>,
<matplotlib.axis.XTick at 0x267bdb6e6d0>]
```

