Data Visualization with Matplotlib - Exercises 2

จงทำตามคำสั่งต่อไปนี้ด้วย data ที่กำหนดให้ต่อไปนี้

In [2]:
```
import matplotlib.pyplot as plt
import numpy as np
import pandas as pd
```

อ่านไฟล์ Superstore.csv

In [3]:
```
df = pd.read_csv('Superstore.csv',encoding = 'iso-8859-1')
```

In [4]:
```
df.head()
```

Out[4]:

	Order ID	Customer Name	Segment	Day	Month	Year	Ship Mode	City	State	Category	
0	CA-2016-152156	Claire Gute	Consumer	8	11	2016	Second Class	Henderson	Kentucky	Furniture	Be
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 [5]:
```
df.info()
```

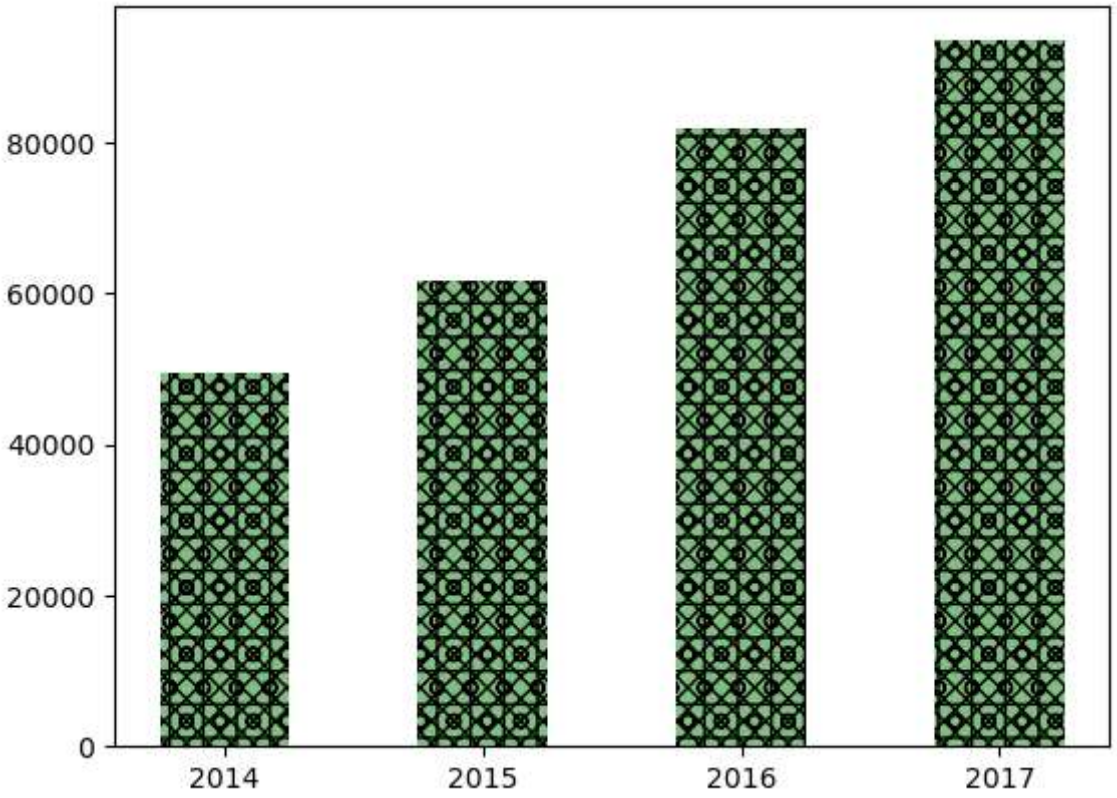
```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 9994 entries, 0 to 9993
Data columns (total 16 columns):
 #   Column         Non-Null Count  Dtype  
---  ------         --------------  -----  
 0   Order ID       9994 non-null   object 
 1   Customer Name  9994 non-null   object 
 2   Segment        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
```

Exercise 1

จงวาดกราฟแท่งแสดงรายได้ของปี 2014 - 2017 และตกแต่งให้สวยงาม

In [26]:
```
df1 = df.groupby('Year')['Profit'].sum()
x = df1.index
y = df1
plt.bar(x,y,color='g',alpha=0.5,hatch = 'xx\/-|o',width=0.5)
plt.xticks([2014,2015,2016,2017])
```

Out[26]:
```
([<matplotlib.axis.XTick at 0x186c98274d0>,
  <matplotlib.axis.XTick at 0x186c9833150>,
  <matplotlib.axis.XTick at 0x186c98f47d0>,
  <matplotlib.axis.XTick at 0x186c9876b50>],
 [Text(2014, 0, '2014'),
  Text(2015, 0, '2015'),
  Text(2016, 0, '2016'),
  Text(2017, 0, '2017')])
```



Exercise 2

จงวาดกราฟแท่งแสดงรายได้ของปี 2014 - 2017 ในกราฟเดียวแยกตามหมวดหมู่ พร้อมตกแต่งให้สวยงาม

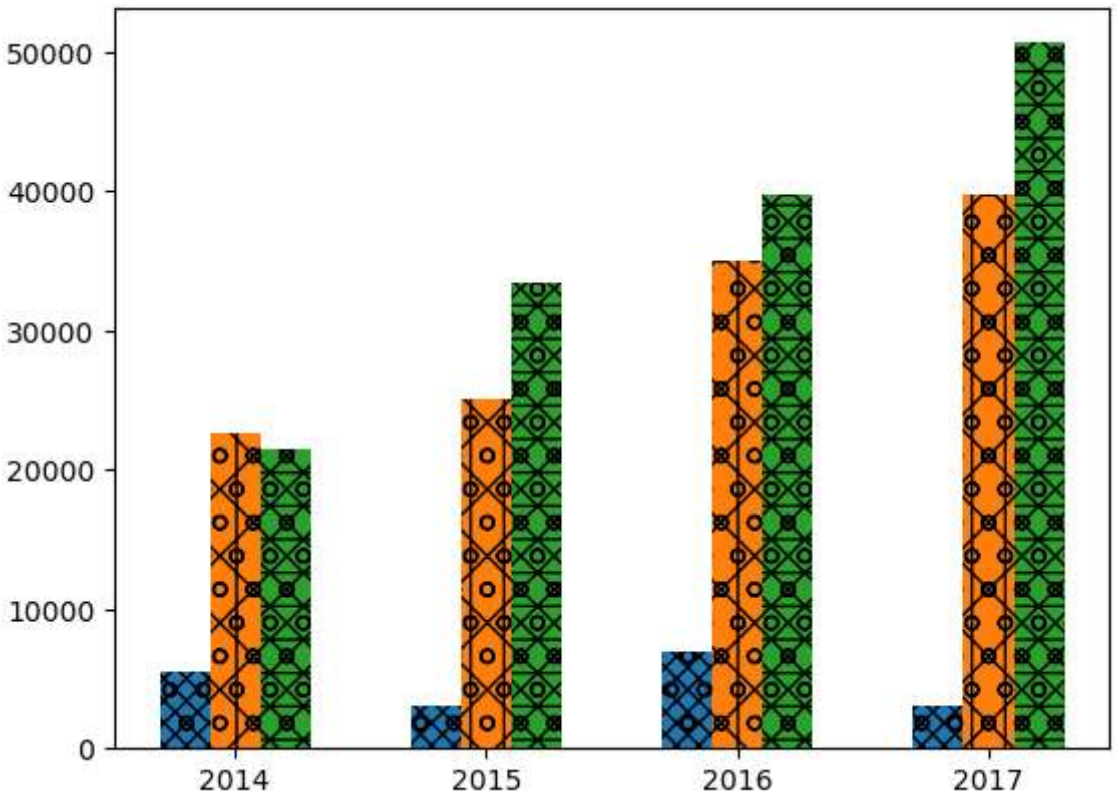
In [32]:
```
arr_df = {}
for i in range(0,df['Category'].nunique()) :
    arr_df[df['Category'].unique()[i]] = df[ df['Category'] == df['Category'].
```

In [42]:
```
x1 = arr_df["Furniture"].index
x2 = arr_df["Office Supplies"].index
x3 = arr_df["Technology"].index

y1 = arr_df["Furniture"]
y2 = arr_df["Office Supplies"]
y3 = arr_df["Technology"]

bw = 0.2
plt.bar(x,y1,width=bw,label="Furniture",hatch = 'xx\/o')
plt.bar(x+bw,y2,width=bw,label="Office Supplies",hatch = '\/|o')
plt.bar(x+bw*2,y3,width=bw,label="Technology",hatch = 'xx-o')
plt.xticks(x+bw,['2014','2015','2016','2017'])
```

Out[42]:
```
([<matplotlib.axis.XTick at 0x186cfe8b110>,
  <matplotlib.axis.XTick at 0x186d01a7690>,
  <matplotlib.axis.XTick at 0x186d01bb810>,
  <matplotlib.axis.XTick at 0x186d1204350>],
 [Text(2014.2, 0, '2014'),
  Text(2015.2, 0, '2015'),
  Text(2016.2, 0, '2016'),
  Text(2017.2, 0, '2017')])
```



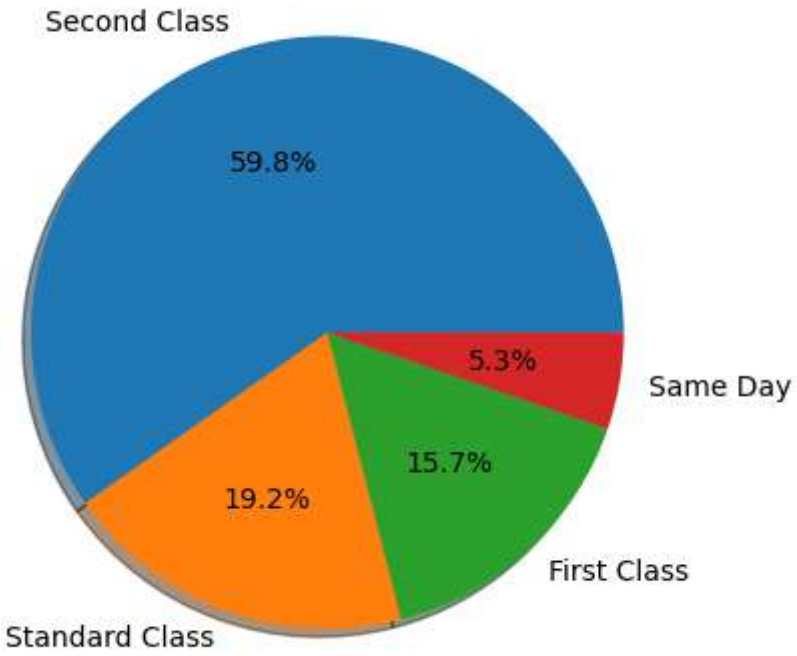
Exercise 3

จงวาดกราฟวงกลม แสดงเปอร์เซ็นต์การขนส่งแต่ละแบบ (Ship Mode) พร้อมตกแต่งให้สวยงาม

In [68]:
```
data = df.groupby('Order ID')['Ship Mode'].unique().value_counts()
lb = df['Ship Mode'].unique()

plt.pie(data,labels=lb,shadow = True,autopct="%.1f%%")
# plt.show
```

Out[68]:
```
([<matplotlib.patches.Wedge at 0x186d29f5e90>,
  <matplotlib.patches.Wedge at 0x186d29f7990>,
  <matplotlib.patches.Wedge at 0x186d29c1690>,
  <matplotlib.patches.Wedge at 0x186d29c3610>],
 [Text(-0.3324299954073045, 1.048565829194095, 'Second Class'),
  Text(-0.37942596150215924, -1.0324901644752662, 'Standard Class'),
  Text(0.7466110434072194, -0.8078192556892806, 'First Class'),
  Text(1.0849555427808886, -0.1813049094454643, 'Same Day')],
 [Text(-0.1813254520403479, 0.5719449977422335, '59.8%'),
  Text(-0.2069596153648141, -0.5631764533501451, '19.2%'),
  Text(0.4072423873130287, -0.4406286849214257, '15.7%'),
  Text(0.5917939324259391, -0.09889358697025323, '5.3%')])
```



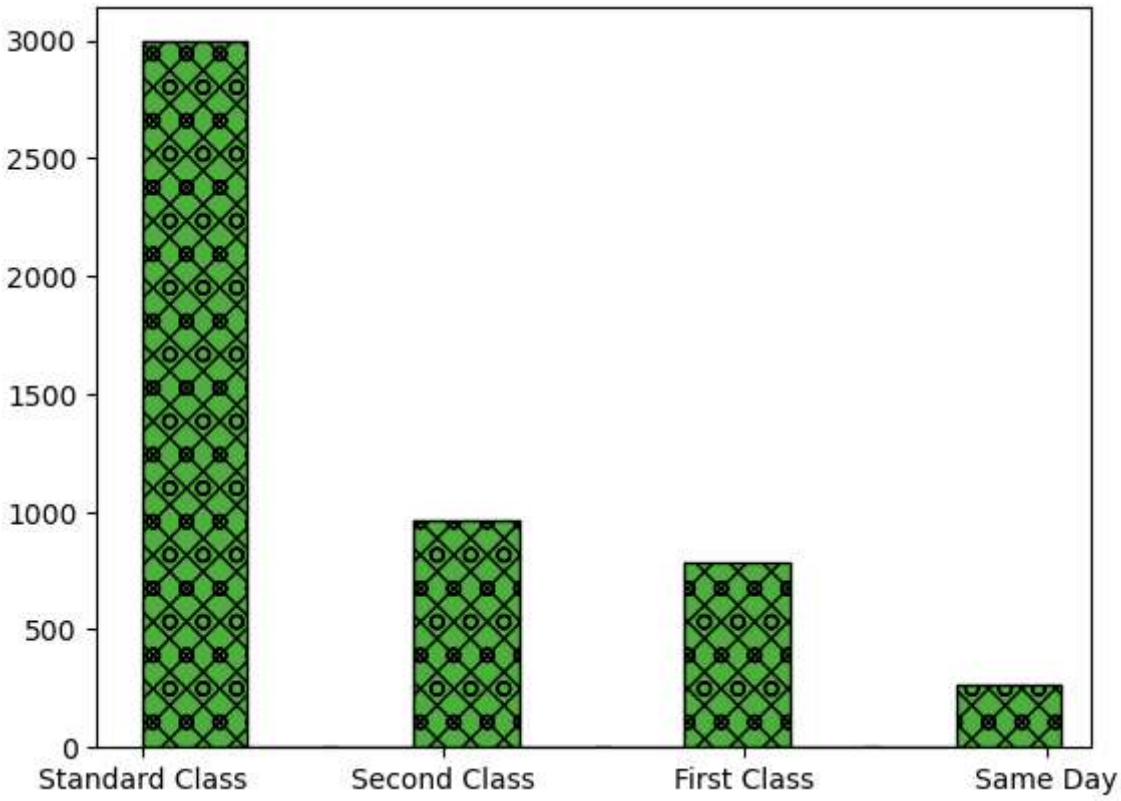
Exercise 4

จงวาดกราฟความถี่ แสดงจำนวนการขนส่งแต่ละแบบ (Ship Mode) พร้อมตกแต่งให้สวยงาม

In [112]:
```
orderIdShipMode = pd.DataFrame({"Order ID":df["Order ID"],"Ship Mode":df["Ship
collapsed_df = orderIdShipMode.groupby("Order ID").first().reset_index()

plt.hist(collapsed_df["Ship Mode"],ec='k',width=0.35,color='#4FAF3E',hatch = '
```

Out[112]:
```
(array([2994.,    0.,    0.,  964.,    0.,    0.,  787.,    0.,    0.,
         264.]),
 array([0. , 0.3, 0.6, 0.9, 1.2, 1.5, 1.8, 2.1, 2.4, 2.7, 3. ]),
 <BarContainer object of 10 artists>)
```



Exercise 5

จงวาดกราฟจุด(Scatter) แสดงราคาขายกับกำไรที่ได้ (Sales , Profit) พร้อมตกแต่งให้สวยงาม

Type *Markdown* and LaTeX: α^2

In [8]:
```
Sales = df['Sales']
Profit = df['Profit']
plt.figure (figsize =  [6,4])
plt.scatter(Sales,Profit,marker='o',c='#319697',s=25,ec ='k',lw = 0.8)
```