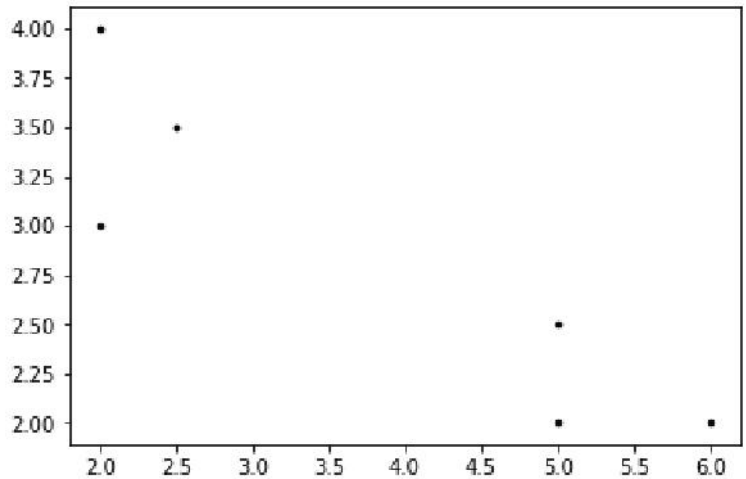
K mean Clustering

Mohammad Shakirul Islam

ID: 151-15-311

In [26]:
```
#import libraries
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
```

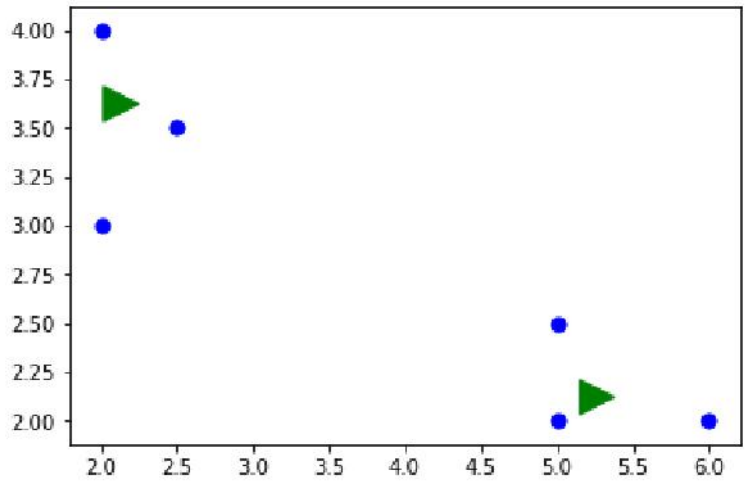
In [27]:
```
def dis(x1,y1,x2,y2):
    a=np.sqrt((x1-x2)**2+(y1-y2)**2) #Distance
    return a

x1=[2.0,2.0,5.0,6.0,5.0,2.5] #Hypothetical data
y1=[4.0,3.0,2.0,2.0,2.5,3.5]

x=np.array(x1)
y=np.array(y1)

#rint(x)
#rint(y)
c1=np.array([2.0,4.0])
c2=np.array([5.0,2.0])

c1xx=[2.0]
c1yy=[4.0]

c2xx=[5.0]
c2yy=[2.0]

plt.scatter(x, y, c='black', s=7)
plt.show()
```



In [28]:
```
ite=int(len(x))

for i in range(0,ite):
    a=dis(c1[0],c1[1],x[i],y[i])
    b=dis(c2[0],c2[1],x[i],y[i])

    if(a<b):
        c1xx.append(x[i])
        c1yy.append(y[i])

        c1x=np.array(c1xx)
        c1y=np.array(c1yy)

        c1[0]=(c1x.sum())/(len(c1x))
        c1[1]=(c1y.sum())/(len(c1y))
    else:
        c2xx.append(x[i])
        c2yy.append(y[i])

        c2x=np.array(c2xx)
        c2y=np.array(c2yy)

        c2[0]=(c2x.sum())/(len(c2x))
        c2[1]=(c2y.sum())/(len(c2y))

    print(c1[0],"",c1[1])
    print(c2[0],"",c2[1])
    print()


x1.append(c1[0])
x1.append(c2[0])

y1.append(c1[1])
y1.append(c2[1])

fx=[c1[0],c2[0]]
fy=[c1[1],c2[1]]

ffx=np.array(fx)
ffy=np.array(fy)

plt.scatter(x,y,c='b',s=50)
plt.scatter(ffx,ffy,marker='>',c='g',s=300)
plt.show()
```

```
2.0  4.0
5.0  2.0

2.0  3.6666666666666665
5.0  2.0

2.0  3.6666666666666665
5.0  2.0

2.0  3.6666666666666665
```

```
5.333333333333333  2.0

2.0  3.6666666666666665
5.25  2.125

2.125  3.625
5.25  2.125
```



In [ ]: