

In [83]:

```
import pandas as pd
import matplotlib.pyplot as plt
import numpy as np
```

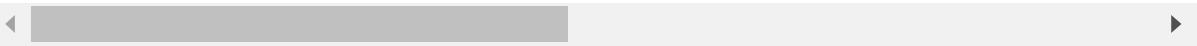
In [5]:

```
data=pd.read_csv('E:\\bigmart_data.csv')
data
```

Out[5]:

	Item_Identifier	Item_Weight	Item_Fat_Content	Item_Visibility	Item_Type	Item_MRP	Outl
0	FDA15	9.300	Low Fat	0.016047	Dairy	249.8092	
1	DRC01	5.920	Regular	0.019278	Soft Drinks	48.2692	
2	FDN15	17.500	Low Fat	0.016760	Meat	141.6180	
3	FDX07	19.200	Regular	0.000000	Fruits and Vegetables	182.0950	
4	NCD19	8.930	Low Fat	0.000000	Household	53.8614	
...	
8518	FDF22	6.865	Low Fat	0.056783	Snack Foods	214.5218	
8519	FDS36	8.380	Regular	0.046982	Baking Goods	108.1570	
8520	NCJ29	10.600	Low Fat	0.035186	Health and Hygiene	85.1224	
8521	FDN46	7.210	Regular	0.145221	Snack Foods	103.1332	
8522	DRG01	14.800	Low Fat	0.044878	Soft Drinks	75.4670	

8523 rows × 12 columns



In [11]:

```
data.isnull().sum()
```

Out[11]:

```
Item_Identifier      0
Item_Weight          1463
Item_Fat_Content     0
Item_Visibility     0
Item_Type            0
Item_MRP             0
Outlet_Identifier    0
Outlet_Establishment_Year  0
Outlet_Size          2410
Outlet_Location_Type  0
Outlet_Type          0
Item_Outlet_Sales    0
dtype: int64
```

In [13]:

```
data.shape
```

Out[13]:

```
(8523, 12)
```

In [16]:

```
data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 8523 entries, 0 to 8522
Data columns (total 12 columns):
#   Column                                Non-Null Count  Dtype
---  -
0   Item_Identifier                       8523 non-null  object
1   Item_Weight                           7060 non-null  float64
2   Item_Fat_Content                      8523 non-null  object
3   Item_Visibility                      8523 non-null  float64
4   Item_Type                            8523 non-null  object
5   Item_MRP                             8523 non-null  float64
6   Outlet_Identifier                     8523 non-null  object
7   Outlet_Establishment_Year            8523 non-null  int64
8   Outlet_Size                          6113 non-null  object
9   Outlet_Location_Type                 8523 non-null  object
10  Outlet_Type                          8523 non-null  object
11  Item_Outlet_Sales                    8523 non-null  float64
dtypes: float64(4), int64(1), object(7)
memory usage: 566.0+ KB
```

In [17]:

```
data.describe()
```

Out[17]:

	Item_Weight	Item_Visibility	Item_MRP	Outlet_Establishment_Year	Item_Outlet_Sales
count	7060.000000	8523.000000	8523.000000	8523.000000	8523.000000
mean	12.857645	0.066132	140.992782	1997.831867	2181.288914
std	4.643456	0.051598	62.275067	8.371760	1706.499616
min	4.555000	0.000000	31.290000	1985.000000	33.290000
25%	8.773750	0.026989	93.826500	1987.000000	834.247400
50%	12.600000	0.053931	143.012800	1999.000000	1794.331000
75%	16.850000	0.094585	185.643700	2004.000000	3101.296400
max	21.350000	0.328391	266.888400	2009.000000	13086.964800

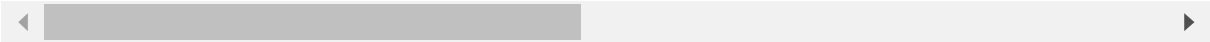
In [19]:

```
data.fillna(0)
```

Out[19]:

	Item_Identifier	Item_Weight	Item_Fat_Content	Item_Visibility	Item_Type	Item_MRP	Outl
0	FDA15	9.300	Low Fat	0.016047	Dairy	249.8092	
1	DRC01	5.920	Regular	0.019278	Soft Drinks	48.2692	
2	FDN15	17.500	Low Fat	0.016760	Meat	141.6180	
3	FDX07	19.200	Regular	0.000000	Fruits and Vegetables	182.0950	
4	NCD19	8.930	Low Fat	0.000000	Household	53.8614	
...	
8518	FDF22	6.865	Low Fat	0.056783	Snack Foods	214.5218	
8519	FDS36	8.380	Regular	0.046982	Baking Goods	108.1570	
8520	NCJ29	10.600	Low Fat	0.035186	Health and Hygiene	85.1224	
8521	FDN46	7.210	Regular	0.145221	Snack Foods	103.1332	
8522	DRG01	14.800	Low Fat	0.044878	Soft Drinks	75.4670	

8523 rows × 12 columns

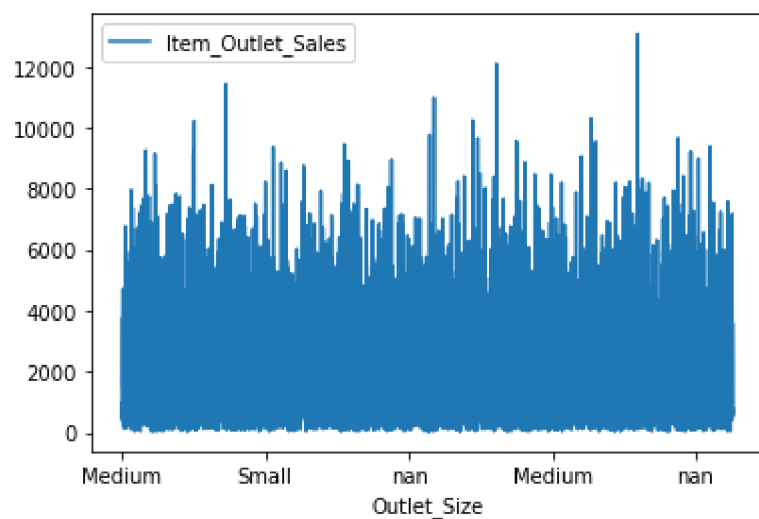


In [26]:

```
data.plot('Outlet_Size', 'Item_Outlet_Sales')
```

Out[26]:

<AxesSubplot:xlabel='Outlet_Size'>

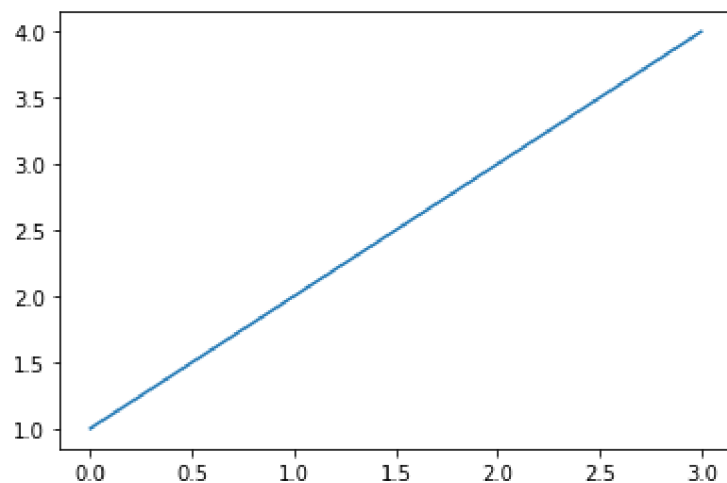


In [31]:

```
mlt.plot([1,2,3,4])
```

Out[31]:

[<matplotlib.lines.Line2D at 0x6772100>]

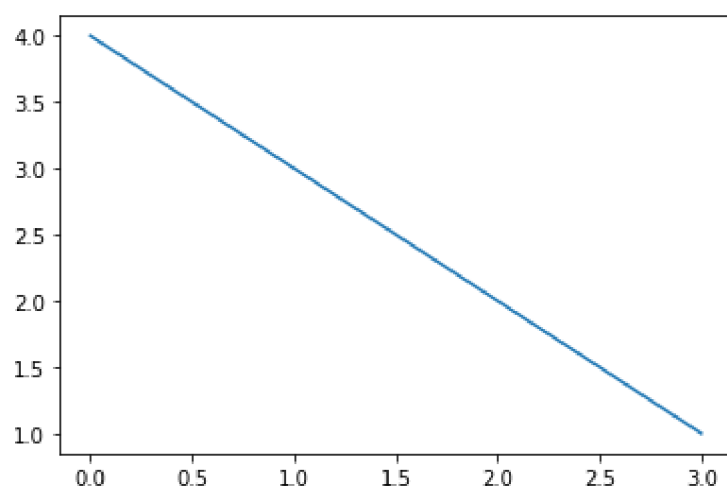


In [32]:

```
plt.plot([4,3,2,1])
```

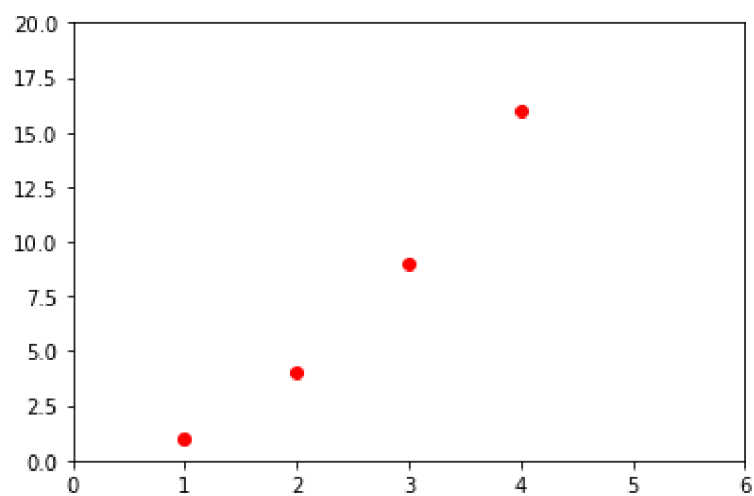
Out[32]:

[<matplotlib.lines.Line2D at 0x63c7ac0>]



In [36]:

```
plt.plot([1, 2, 3, 4], [1, 4, 9, 16], 'ro')  
plt.axis([0, 6, 0, 20])  
plt.show()
```

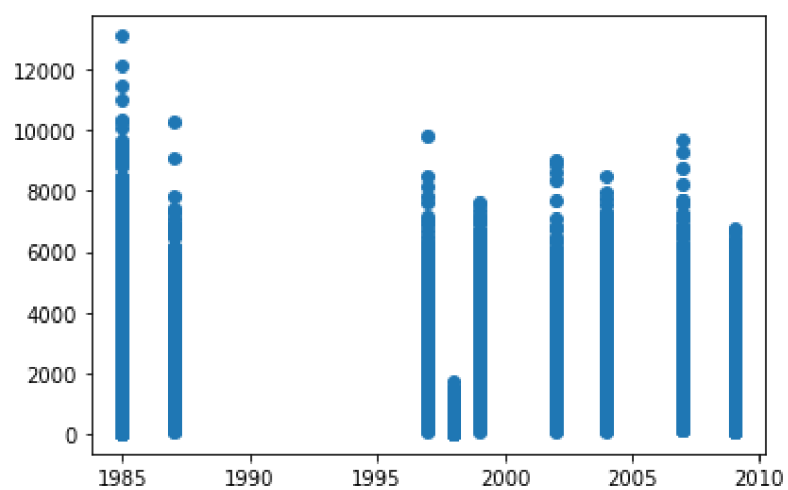


In [41]:

```
plt.scatter(data['Outlet_Establishment_Year'], data['Item_Outlet_Sales'])
```

Out[41]:

<matplotlib.collections.PathCollection at 0x66dc970>



In [42]:

```
data
```

Out[42]:

	Item_Identifier	Item_Weight	Item_Fat_Content	Item_Visibility	Item_Type	Item_MRP	Outlet
0	FDA15	9.300	Low Fat	0.016047	Dairy	249.8092	
1	DRC01	5.920	Regular	0.019278	Soft Drinks	48.2692	
2	FDN15	17.500	Low Fat	0.016760	Meat	141.6180	
3	FDX07	19.200	Regular	0.000000	Fruits and Vegetables	182.0950	
4	NCD19	8.930	Low Fat	0.000000	Household	53.8614	
...
8518	FDF22	6.865	Low Fat	0.056783	Snack Foods	214.5218	
8519	FDS36	8.380	Regular	0.046982	Baking Goods	108.1570	
8520	NCJ29	10.600	Low Fat	0.035186	Health and Hygiene	85.1224	
8521	FDN46	7.210	Regular	0.145221	Snack Foods	103.1332	
8522	DRG01	14.800	Low Fat	0.044878	Soft Drinks	75.4670	

8523 rows × 12 columns

In [46]:

```
data['Outlet_Location_Type'].isnull().sum()
```

Out[46]:

0

In [47]:

```
def outlet_loc(x):
    if x=='Tier 1':
        data['Outlet_Location_Type']=1
    elif x=='Tier 2':
        data['Outlet_Location_Type']=2
    else:
        data['Outlet_Location_Type']=3
```

In [52]:

```
data['Outlet_Location_Type'].apply(lambda x:outlet_loc(x))
```

Out[52]:

```
0      None
1      None
2      None
3      None
4      None
...
8518   None
8519   None
8520   None
8521   None
8522   None
```

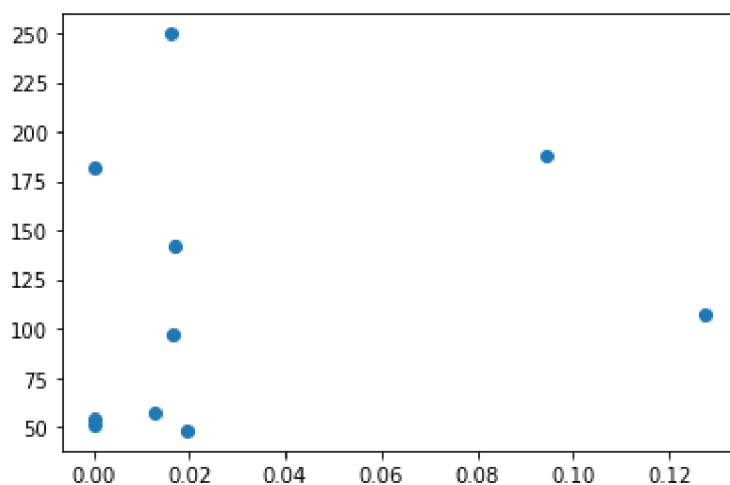
Name: Outlet_Location_Type, Length: 8523, dtype: object

In [56]:

```
mlt.scatter(data['Item_Visibility'].head(10),data['Item_MRP'].head(10))
```

Out[56]:

<matplotlib.collections.PathCollection at 0x7d6e1d8>

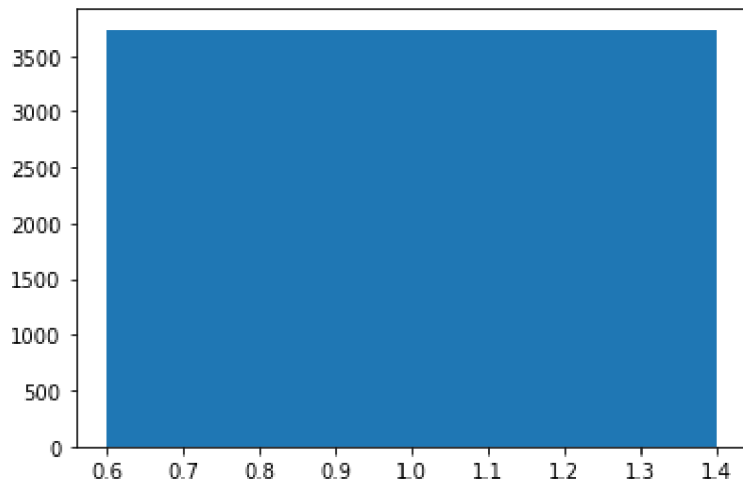


In [59]:

```
mlt.bar(data['Outlet_Location_Type'].head(3),data['Item_Outlet_Sales'].head(3))
```

Out[59]:

<BarContainer object of 3 artists>

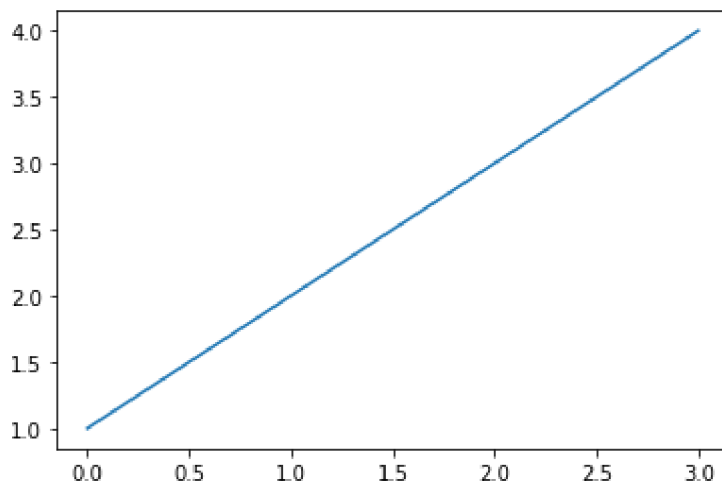


In [5]:

```
data=[1,2,3,4]  
data1=[3,4,1,2]  
mlt.plot(data)
```

Out[5]:

[<matplotlib.lines.Line2D at 0x65f4a48>]

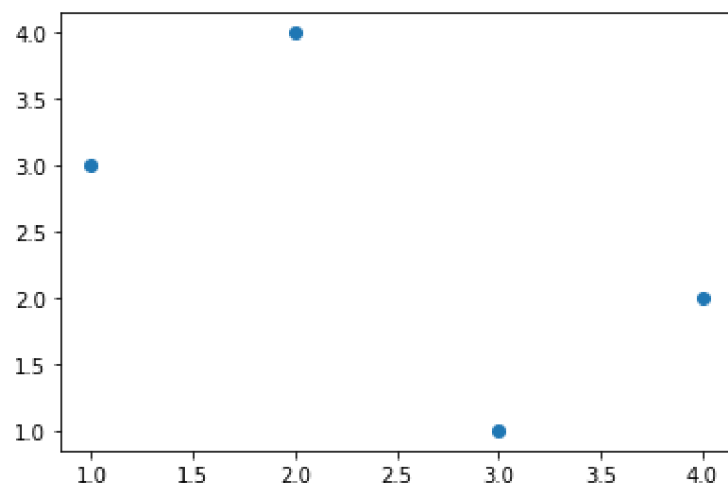


In [6]:

```
mlt.scatter(data,data1)
```

Out[6]:

<matplotlib.collections.PathCollection at 0x66325b0>

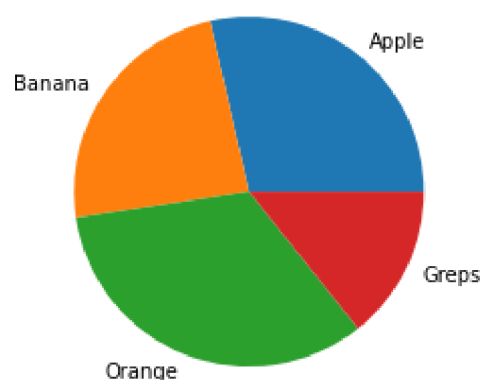


In [28]:

```
data2=[60,50,70,30]
data3=['Apple', 'Banana', 'Orange', 'Greps']
plt.pie(data2,labels=data3)
```

Out[28]:

```
([<matplotlib.patches.Wedge at 0xcaf4358>,
 <matplotlib.patches.Wedge at 0xcaf46e8>,
 <matplotlib.patches.Wedge at 0xca83388>,
 <matplotlib.patches.Wedge at 0xca83c88>],
 [Text(0.6858387475358717, 0.8600146582346299, 'Apple'),
 Text(-0.9088627028569363, 0.6196519889063413, 'Banana'),
 Text(-0.4018750081066444, -1.0239611700935172, 'Orange'),
 Text(0.991065834488076, -0.4772719473323923, 'Greps')])
```



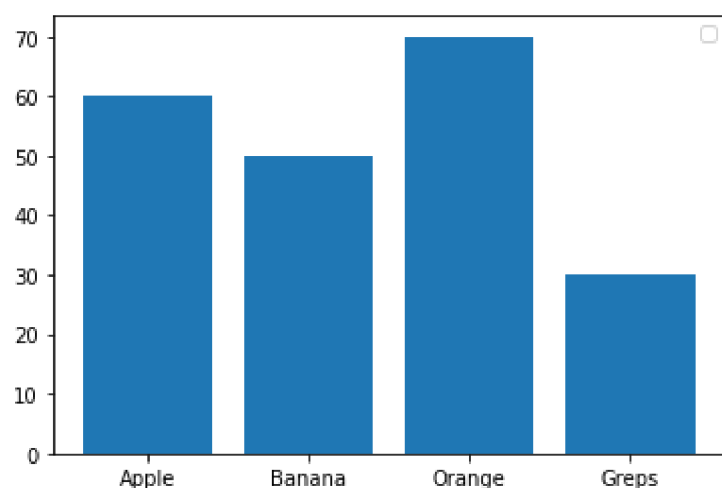
In [135]:

```
plt.bar(data3,data2)
plt.legend()
```

No handles with labels found to put in legend.

Out[135]:

<matplotlib.legend.Legend at 0xfa4b970>

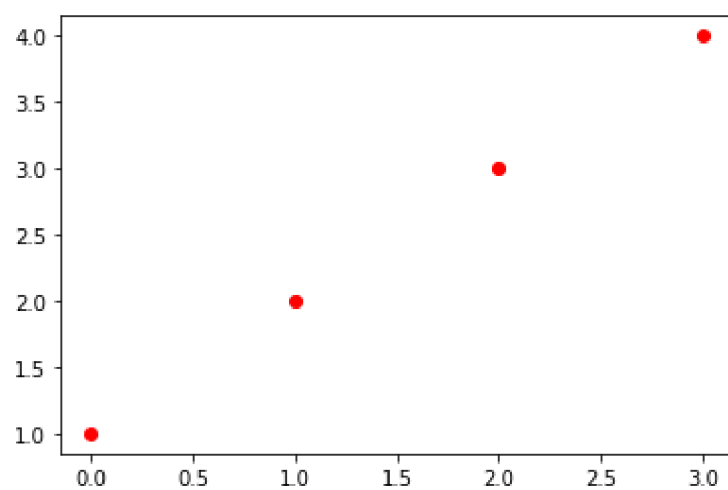


In [19]:

```
mlt.plot(data, 'ro')
```

Out[19]:

[<matplotlib.lines.Line2D at 0xcac9bc8>]



In [20]:

```
data1
```

Out[20]:

[3, 4, 1, 2]

In [21]:

```
data
```

Out[21]:

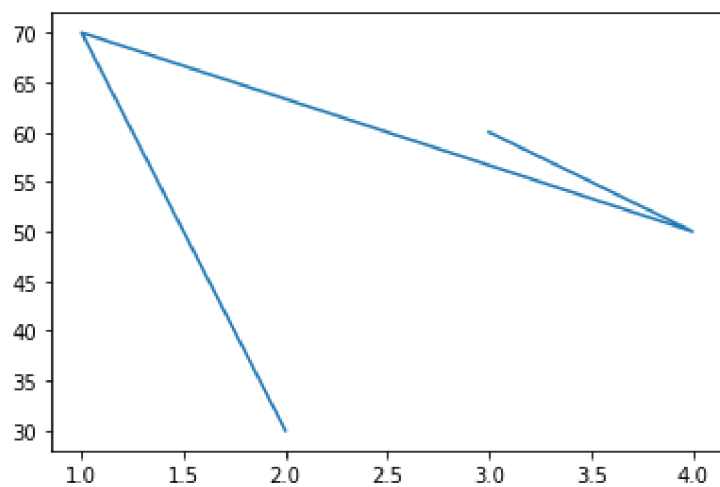
[1, 2, 3, 4]

In [32]:

```
plt.plot(data1,data2)
```

Out[32]:

[<matplotlib.lines.Line2D at 0xcbecef8>]

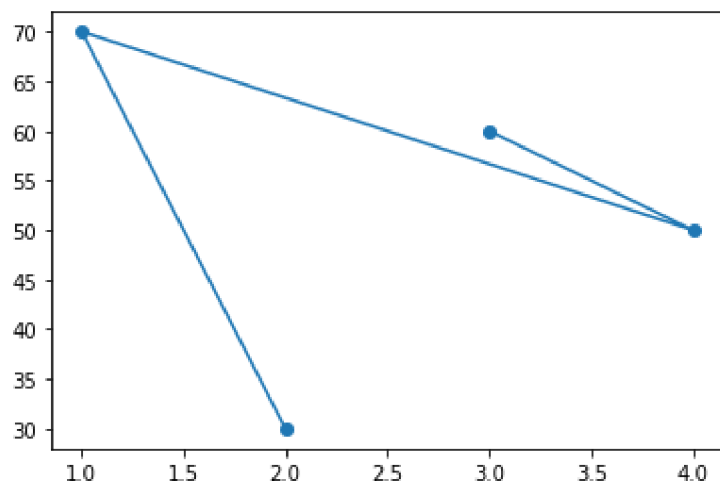


In [33]:

```
plt.plot(data1,data2,marker='o')
```

Out[33]:

[<matplotlib.lines.Line2D at 0xcc45e38>]

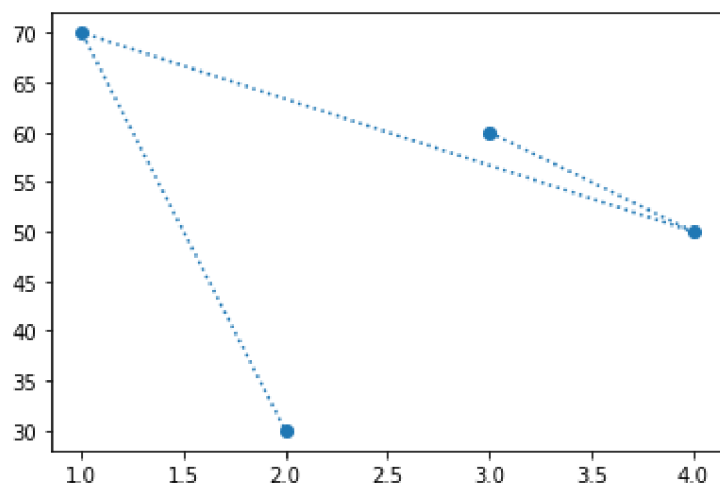


In [35]:

```
plt.plot(data1,data2,marker='o',linestyle='dotted')
```

Out[35]:

[<matplotlib.lines.Line2D at 0xcd2a658>]

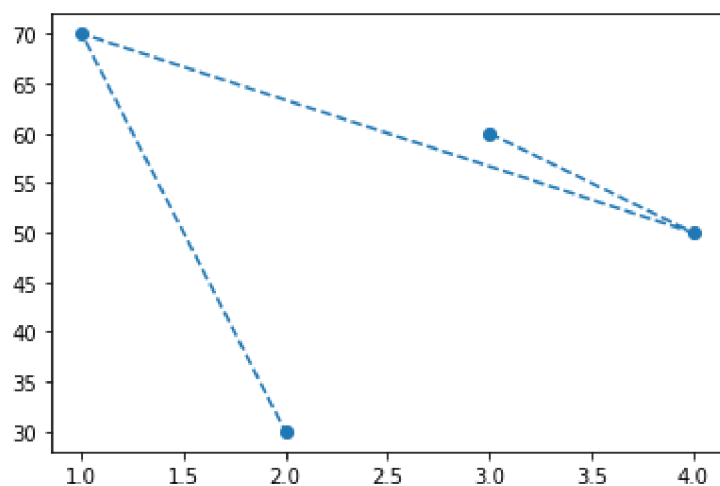


In [36]:

```
plt.plot(data1,data2,marker='o',linestyle='dashed')
```

Out[36]:

[<matplotlib.lines.Line2D at 0xcd65478>]

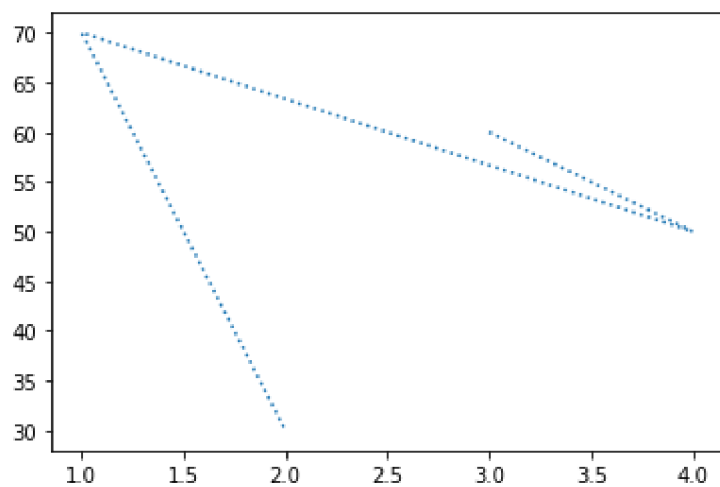


In [39]:

```
plt.plot(data1,data2,ls=':')#indicates doted
```

Out[39]:

[<matplotlib.lines.Line2D at 0xcdde088>]

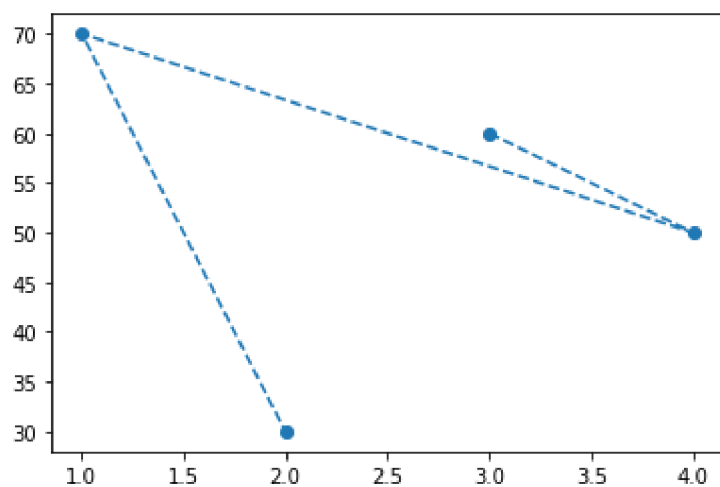


In [40]:

```
plt.plot(data1,data2,marker='o',ls='--')
```

Out[40]:

[<matplotlib.lines.Line2D at 0xce09e80>]

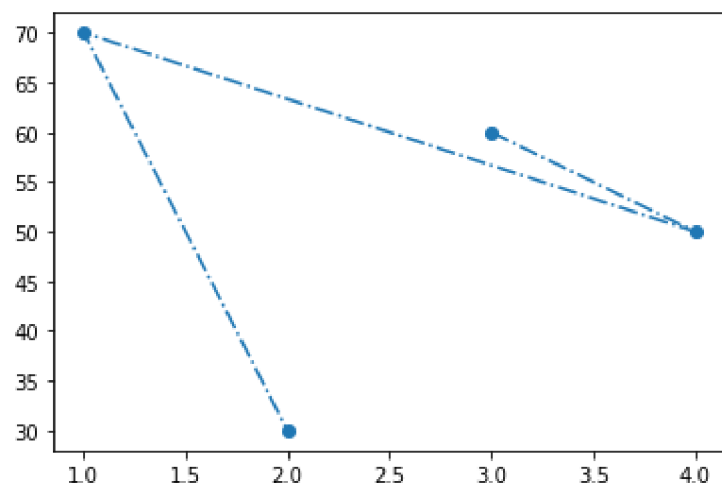


In [41]:

```
plt.plot(data1,data2,marker='o',ls='-.')
```

Out[41]:

[<matplotlib.lines.Line2D at 0xcc6e568>]

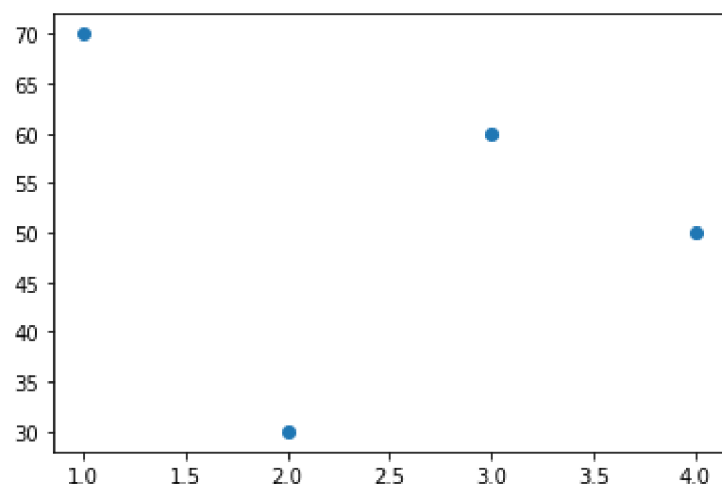


In [42]:

```
plt.plot(data1,data2,marker='o',ls='')
```

Out[42]:

[<matplotlib.lines.Line2D at 0xcd01f0>]

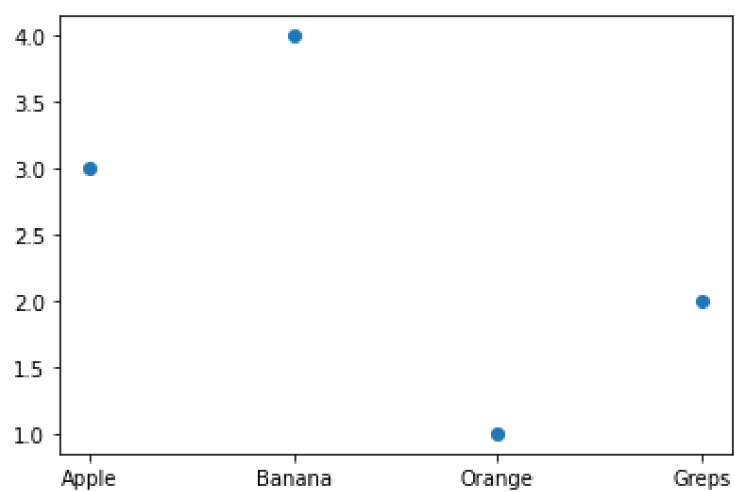


In [47]:

```
mlt.scatter(data3,data1,label=['fruits','quantity'])
```

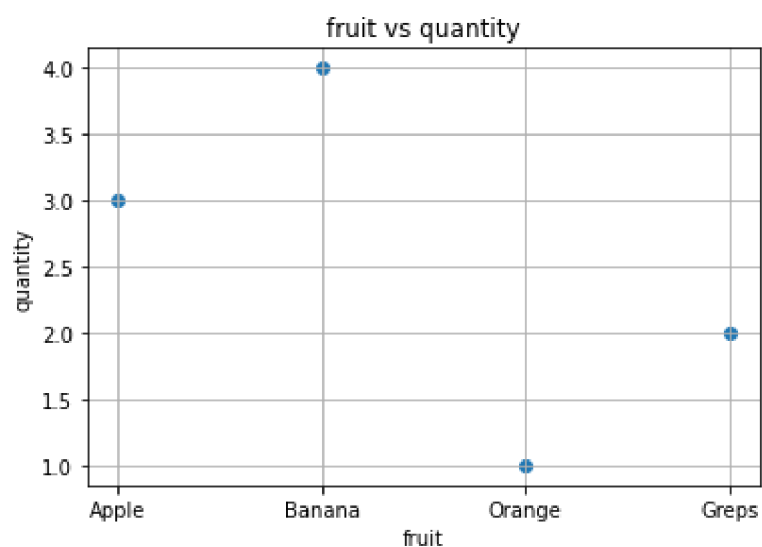
Out[47]:

<matplotlib.collections.PathCollection at 0xdf15700>



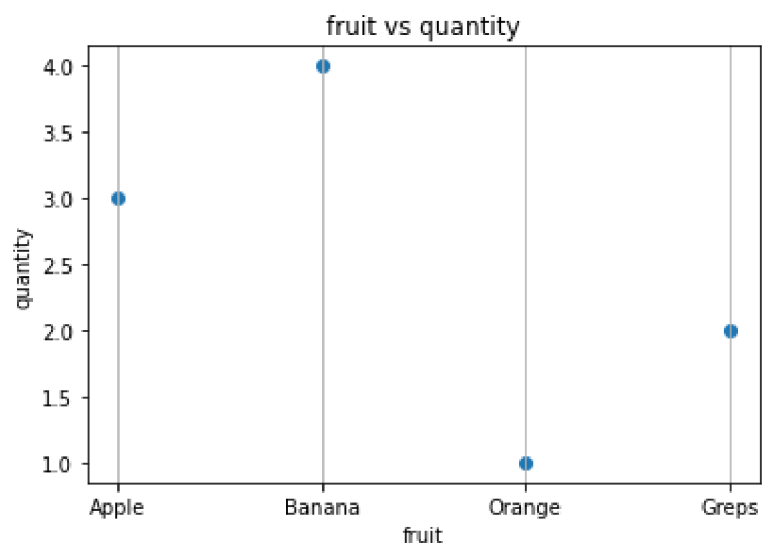
In [51]:

```
mlt.xlabel('fruit')  
mlt.ylabel('quantity')  
mlt.scatter(data3,data1)  
mlt.title('fruit vs quantity')  
mlt.grid()
```



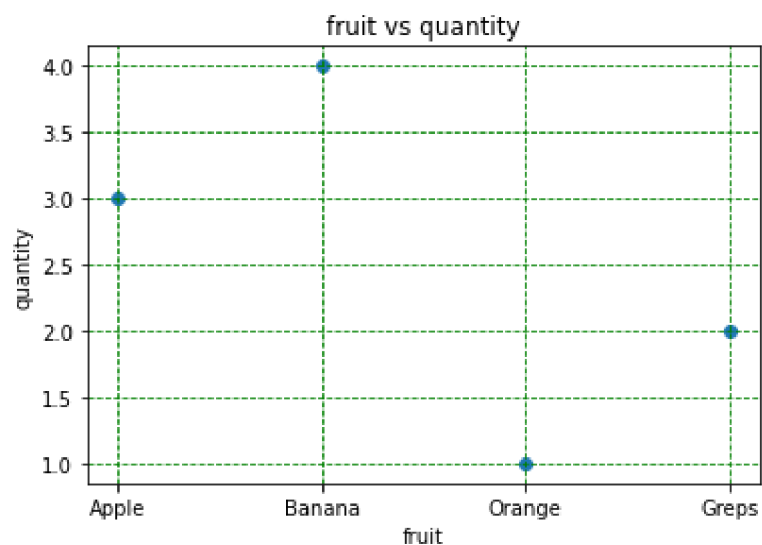
In [53]:

```
mlt.xlabel('fruit')
mlt.ylabel('quantity')
mlt.scatter(data3,data1)
mlt.title('fruit vs quantity')
mlt.grid(axis='x')
```



In [65]:

```
mlt.xlabel('fruit')
mlt.ylabel('quantity')
mlt.scatter(data3,data1)
mlt.title('fruit vs quantity')
mlt.grid(color='Green',linestyle='--')
```

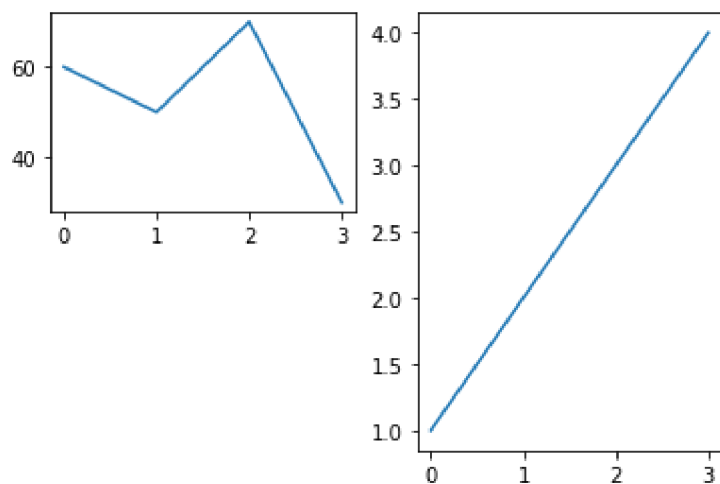


In [78]:

```
datax=[1,2,3]
plt.subplot(1,2,1)
plt.plot(data1)
plt.subplot(1,2,2)
plt.plot(data)
plt.subplot(2,2,1)
plt.plot(data2)
```

Out[78]:

[<matplotlib.lines.Line2D at 0xe147388>]



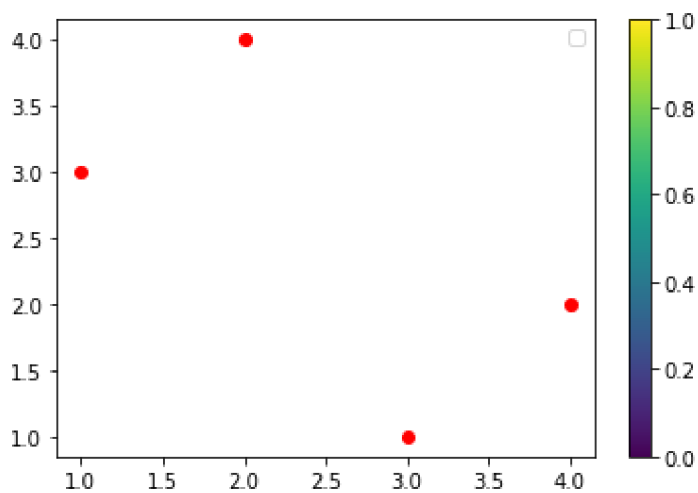
In [134]:

```
color=np.array(['red','black','blue','green','voilet'])
mlt.scatter(data1,data,c='red',cmap='viridis')
mlt.colorbar()
mlt.legend()
```

No handles with labels found to put in legend.

Out[134]:

<matplotlib.legend.Legend at 0xfa1adf0>



In [93]:

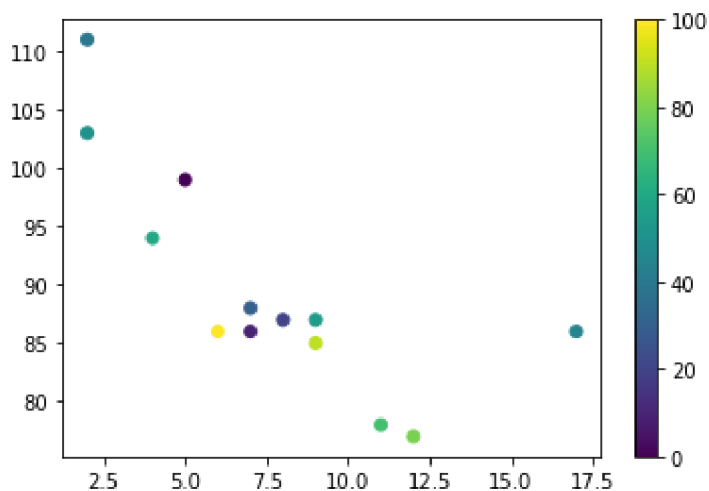
```
import matplotlib.pyplot as plt
import numpy as np

x = np.array([5,7,8,7,2,17,2,9,4,11,12,9,6])
y = np.array([99,86,87,88,111,86,103,87,94,78,77,85,86])
colors = np.array([0, 10, 20, 30, 40, 45, 50, 55, 60, 70, 80, 90, 100])

plt.scatter(x, y, c=colors, cmap='viridis')

plt.colorbar()

plt.show()
```



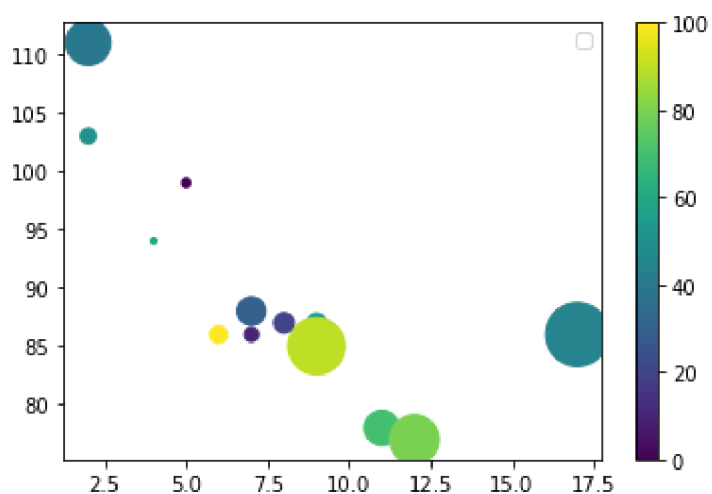
In [133]:

```
import matplotlib.pyplot as plt
import numpy as np

x = np.array([5,7,8,7,2,17,2,9,4,11,12,9,6])
y = np.array([99,86,87,88,111,86,103,87,94,78,77,85,86])
colors = np.array([0, 10, 20, 30, 40, 45, 50, 55, 60, 70, 80, 90, 100])
sizes = np.array([20,50,100,200,500,1000,60,90,10,300,600,800,75])
plt.scatter(x, y, c=colors, cmap='viridis',s=sizes)

plt.colorbar()
plt.legend()
plt.show()
```

No handles with labels found to put in legend.

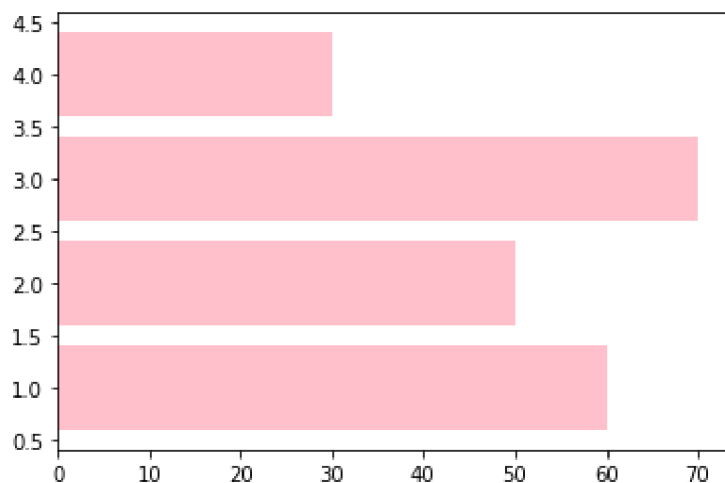


In [102]:

```
mlt.barh(data,data2,color='pink')
```

Out[102]:

<BarContainer object of 4 artists>

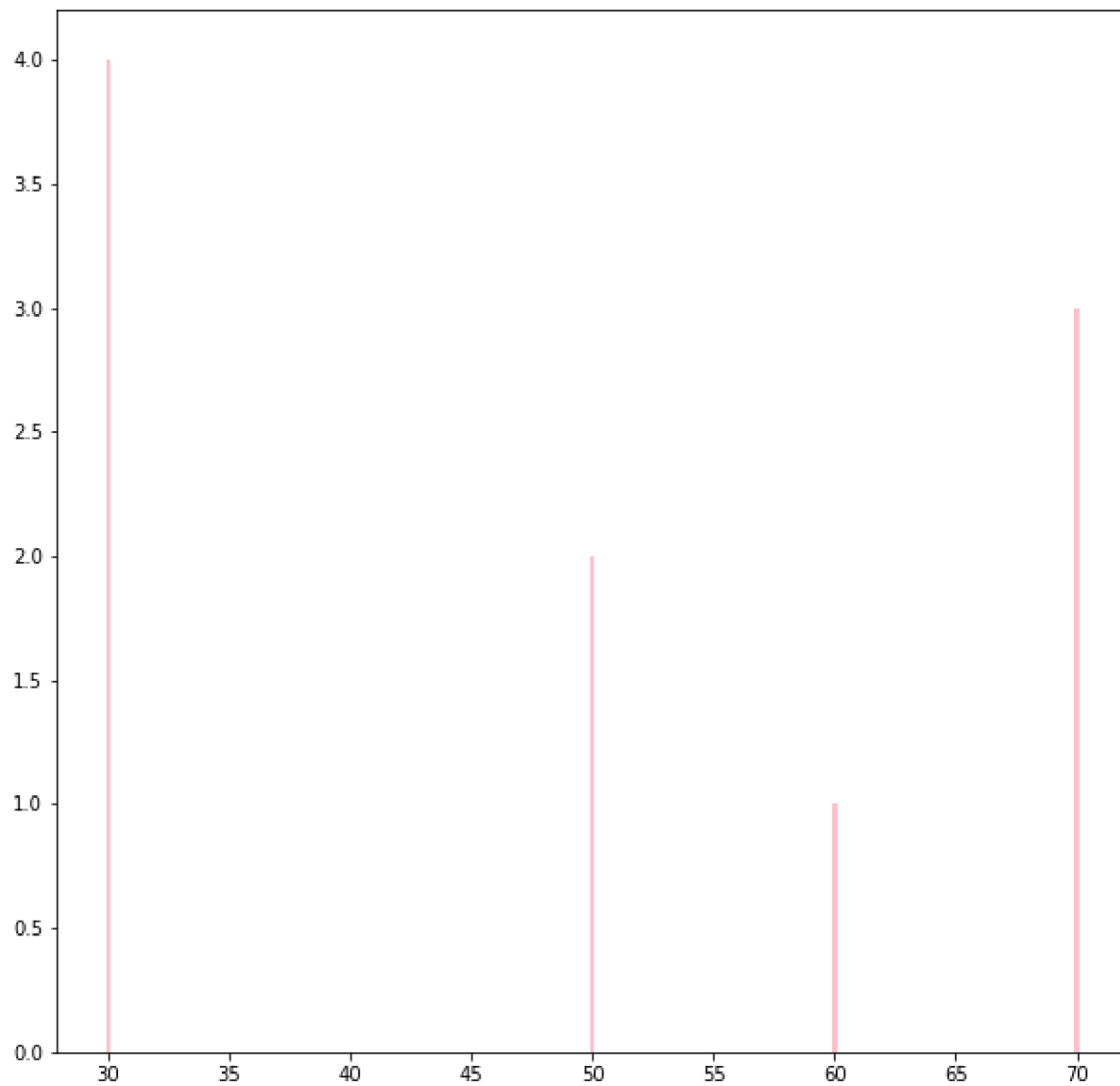


In [108]:

```
plt.figure(figsize=[10,10])  
plt.bar(data2,data,color='Pink',width=0.20)
```

Out[108]:

<BarContainer object of 4 artists>



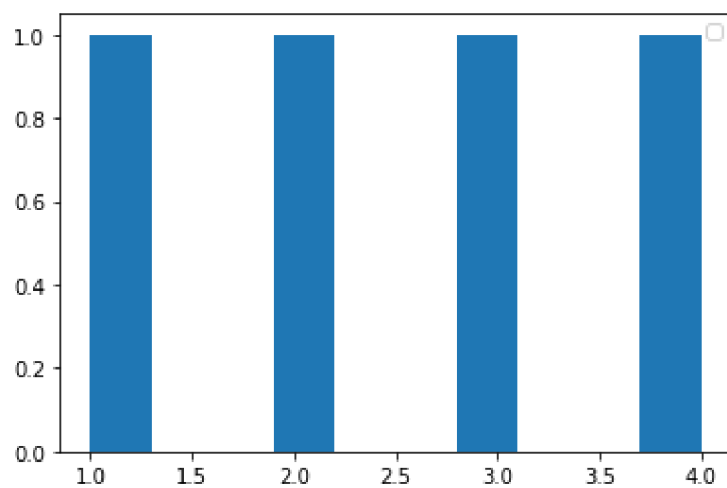
In [131]:

```
mlt.hist(data1)  
mlt.legend()
```

No handles with labels found to put in legend.

Out[131]:

<matplotlib.legend.Legend at 0xe5e8e98>



In [113]:

```
data
```

Out[113]:

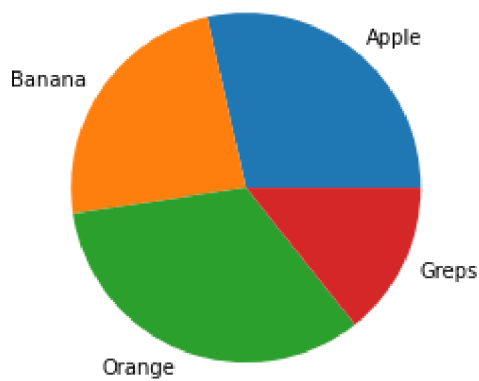
[1, 2, 3, 4]

In [121]:

```
data2=[60,50,70,30]  
data3=['Apple','Banana','Orange','Greps']  
plt.pie(data2,labels=data3)
```

Out[121]:

```
([<matplotlib.patches.Wedge at 0xe67faa8>,  
 <matplotlib.patches.Wedge at 0xe67fd18>,  
 <matplotlib.patches.Wedge at 0xe67ff28>,  
 <matplotlib.patches.Wedge at 0xe68e160>],  
 [Text(0.6858387475358717, 0.8600146582346299, 'Apple'),  
  Text(-0.9088627028569363, 0.6196519889063413, 'Banana'),  
  Text(-0.4018750081066444, -1.0239611700935172, 'Orange'),  
  Text(0.991065834488076, -0.4772719473323923, 'Greps')])
```



In [130]:

```
data2=[60,50,70,30]  
data3=['Apple','Banana','Orange','Greps']  
plt.pie(data2,labels=data3,explode=[0.2,1,0.3,0])  
plt.legend()
```

Out[130]:

<matplotlib.legend.Legend at 0xe3b5ac0>

