In [2]:

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

```
C:\Users\student\AppData\Roaming\Python\Python37\site-packages\pandas\comp
at\_optional.py:138: UserWarning: Pandas requires version '2.7.0' or newer
of 'numexpr' (version '2.6.8' currently installed).
  warnings.warn(msg, UserWarning)
```

In [3]:

```
df=pd.read_csv("sales_data_sample (1).csv")
df
```

Out[3]:

	ORDERNUMBER	QUANTITYORDERED	PRICEEACH	ORDERLINENUMBER	SALES	ORD
0	10107	30	95.70	2	2871.00	
1	10121	34	81.35	5	2765.90	5/7/
2	10134	41	94.74	2	3884.34	7/1/
3	10145	45	83.26	6	3746.70	
4	10159	49	100.00	14	5205.27	1
...	...	...	...	...	...	
2818	10350	20	100.00	15	2244.40	
2819	10373	29	100.00	1	3978.51	
2820	10386	43	100.00	4	5417.57	3/1/
2821	10397	34	62.24	1	2116.16	
2822	10414	47	65.52	9	3079.44	5/6/

2823 rows × 16 columns

In [4]:

```
df.columns
```

Out[4]:

```
Index(['ORDERNUMBER', 'QUANTITYORDERED', 'PRICEEACH', 'ORDERLINENUMBER',
       'SALES', 'ORDERDATE', 'STATUS', 'QTR_ID', 'MONTH_ID', 'YEAR_ID',
       'PRODUCTLINE', 'MSRP', 'PRODUCTCODE', 'CUSTOMERNAME', 'COUNTRY',
       'DEALSIZE'],
      dtype='object')
```

In [7]:

```
newdf=df.groupby('COUNTRY')
country=df['COUNTRY'].unique()
sum(newdf.get_group('USA')['SALES'])
```

Out[7]:

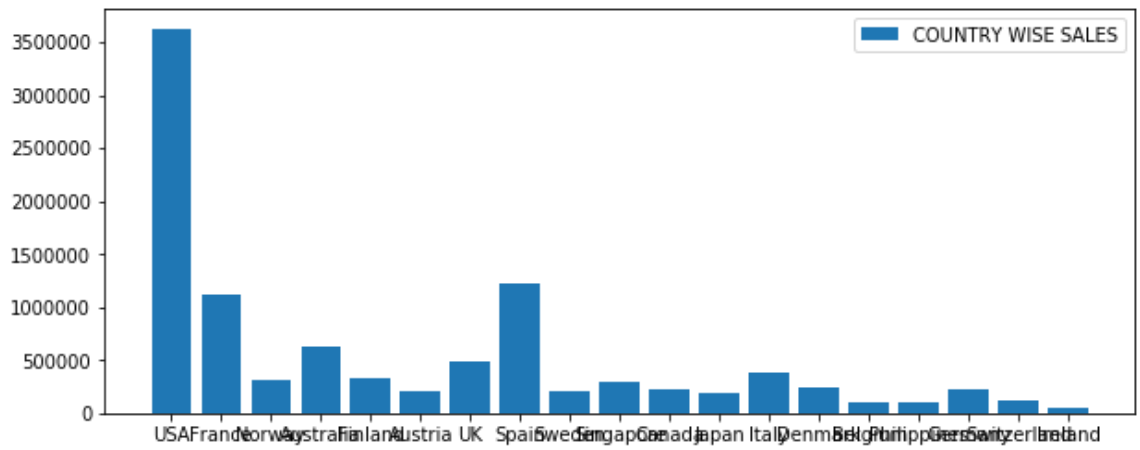
```
3627982.83
```

In [11]:

```
newdf=df.groupby('COUNTRY')
country=df['COUNTRY'].unique()
sales=[]
for cname in country:
    sales.append(sum(newdf.get_group(cname)['SALES']))

f=plt.figure()
f.set_figwidth(30)
f.set_figwidth(10)
newdf=df.groupby('COUNTRY')
country=df['COUNTRY'].unique()
sales=[]
for cname in country :
    sales.append(sum(newdf.get_group(cname)['SALES']))
f=plt.figure()
f.set_figwidth(30)
f.set_figwidth(10)

font1={'family':'serif','color':'blue','size':20}
font2={'family':'serif','color':'darkred','size':15}
plt.bar(country,sales,label="COUNTRY WISE SALES")
plt.legend(loc="best")
```

Out[11]:

```
<matplotlib.legend.Legend at 0x21a4ea75470>

<Figure size 720x288 with 0 Axes>
```

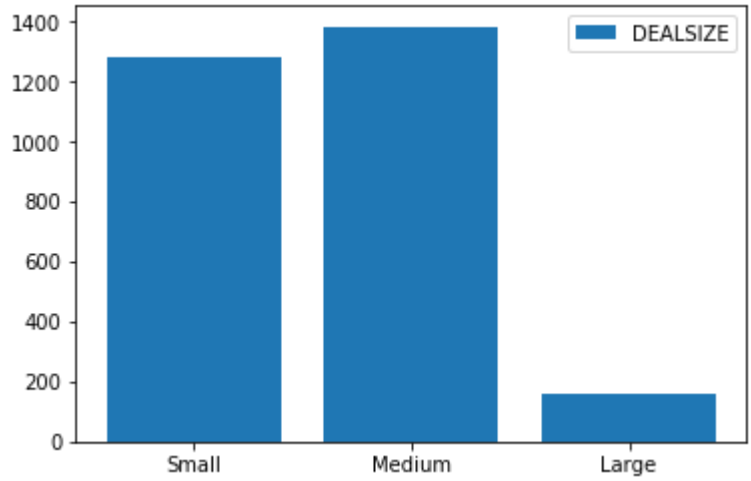


In [14]:

```
dsize=df['DEALSIZE'].unique()
deal=[]
newdf=df.groupby('DEALSIZE')
for dname in dsize:
    deal.append(newdf.get_group(dname)['DEALSIZE'].count())

plt.bar(df['DEALSIZE'].unique(),deal, label="DEALSIZE")
plt.legend(loc="best")
```

Out[14]:

```
<matplotlib.legend.Legend at 0x21a4e7d0358>
```



In [15]:

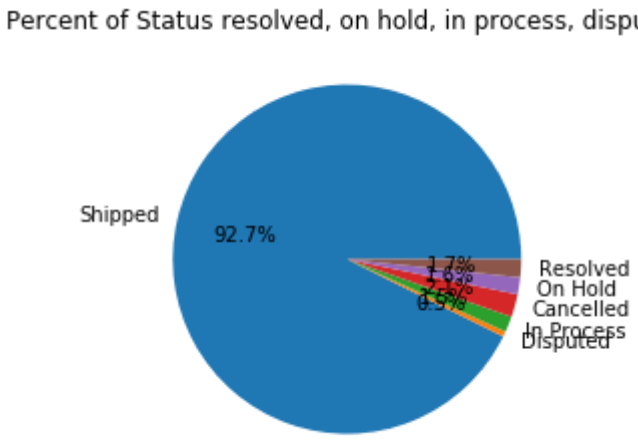
```
newdf=df.groupby('STATUS')
tot=df['STATUS'].count()
status=df['STATUS'].unique()
percent=[]
for sname in status:
    percent.append(newdf.get_group(sname)['STATUS'].count()*100/tot)

```

In [16]:

```
plt.pie(percent,labels=status,autopct='%1.1f%%')
plt.title('Percent of Status resolved, on hold, in process, disputed')
```

Out[16]:

```
Text(0.5, 1.0, 'Percent of Status resolved, on hold, in process, dispute
d')
```

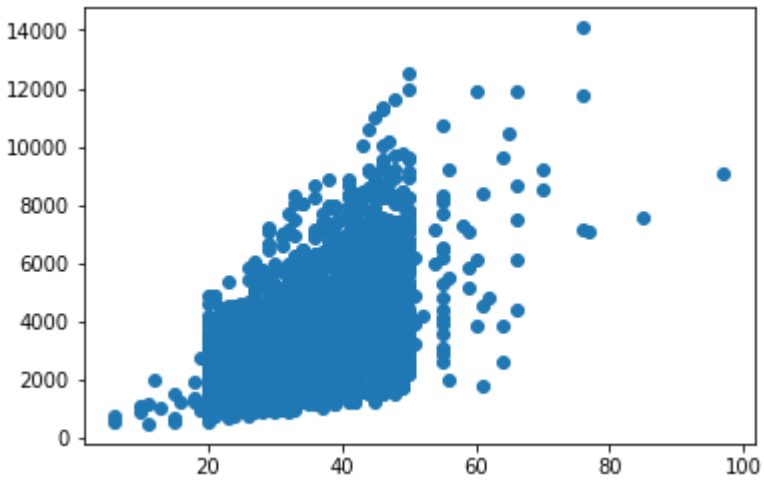


In [17]:

```
plt.scatter(df['QUANTITYORDERED'],df['SALES'])
```

Out[17]:

```
<matplotlib.collections.PathCollection at 0x21a51ca4ef0>
```



In [ ]:

```

```