In [9]:
```
import matplotlib.pyplot as plt
import pandas as pd
import numpy as np
```

In [10]:
```
df=pd.read_csv("sales_data_sample.csv")
df
```

2815	10315	40	55.69	5	2227.60	10
2816	10327	37	86.74	4	3209.38	11
2817	10337	42	97.16	5	4080.72	11
2818	10350	20	100.00	15	2244.40	1
2819	10373	29	100.00	1	3978.51	1
2820	10386	43	100.00	4	5417.57	3/1/2
2821	10397	34	62.24	1	2116.16	3
2822	10414	47	65.52	9	3079.44	5/6/2

In [11]:
```
df.columns
```

Out[11]:
```
Index(['ORDERNUMBER', 'QUANTITYORDERED', 'PRICEEACH', 'ORDERLINENUMBER',
       'SALES', 'ORDERDATE', 'STATUS', 'QTR_ID', 'MONTH_ID', 'YEAR_ID',
       'PRODUCTLINE', 'MSRP', 'PRODUCTCODE', 'CUSTOMERNAME', 'COUNTRY',
       'DEALSIZE'],
      dtype='object')
```

In [12]:
```
newdf=df.groupby('COUNTRY')
country=df['COUNTRY'].unique()
sum(newdf.get_group('USA')['SALES'])
```

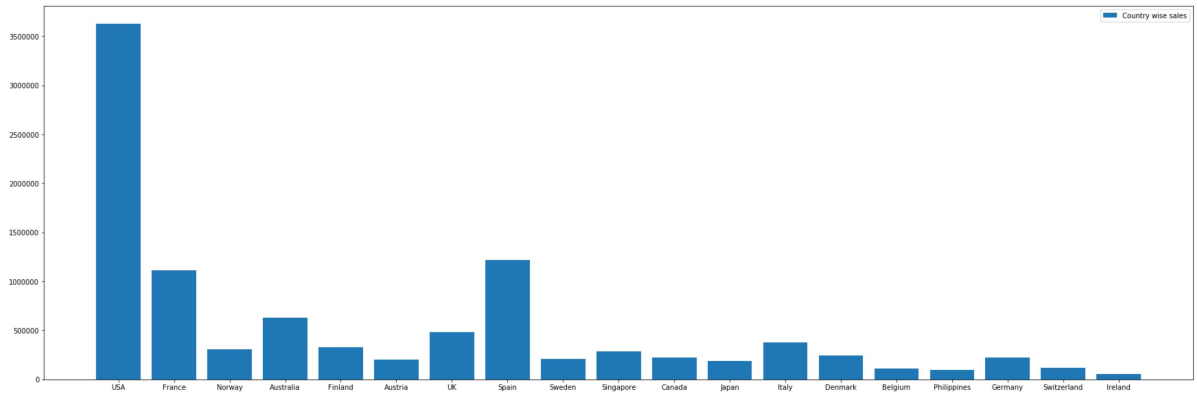
Out[12]:
```
3627982.83
```

In [15]:
```
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_figheight(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")
```

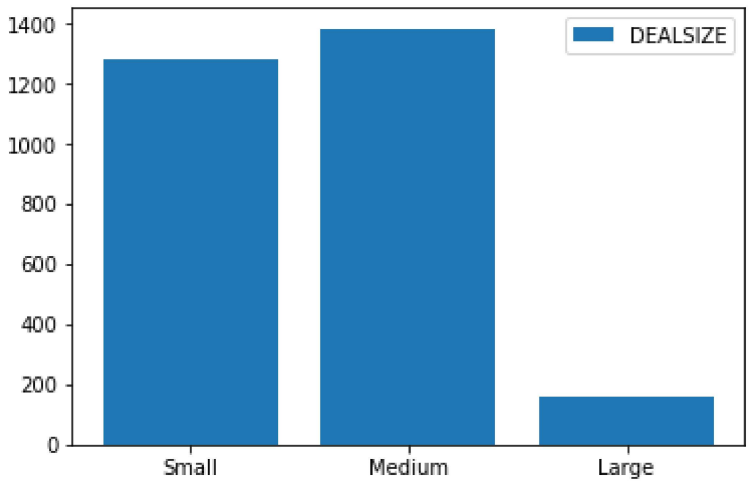
Out[15]: <matplotlib.legend.Legend at 0x2a1fea69358>



In [16]:
```
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")
```

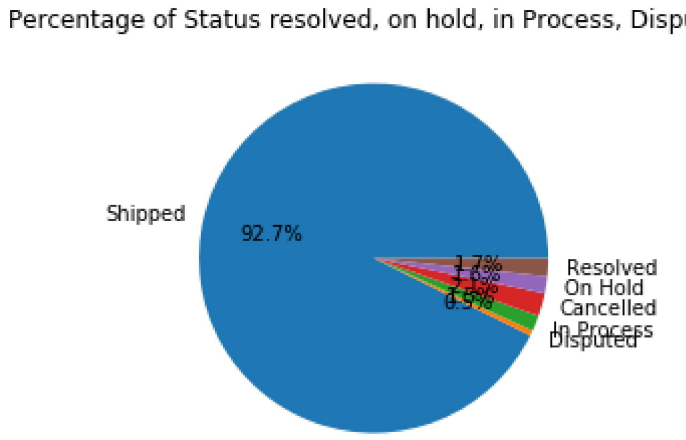
Out[16]: <matplotlib.legend.Legend at 0x2a1fea45eb8>



In [17]:
```
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)
```

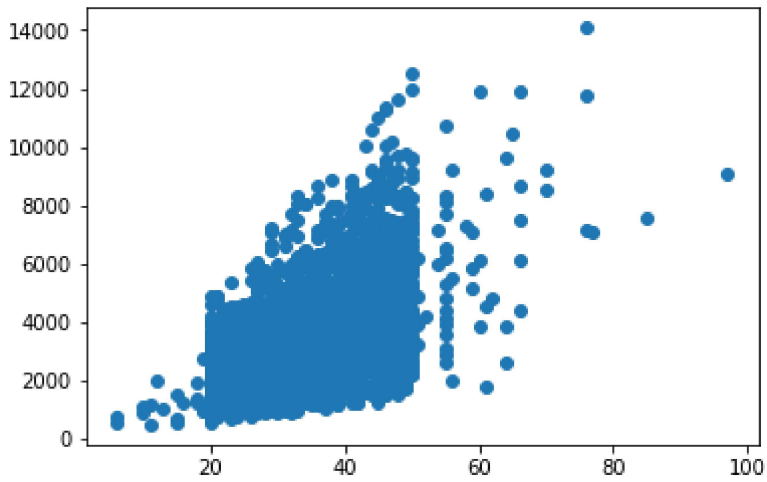
In [19]:
```
plt.pie(percent, labels=status,autopct='%1.1f%%')
plt.title('Percentage of Status resolved, on hold, in Process, Disputed')
```

Out[19]:
```
Text(0.5, 1.0, 'Percentage of Status resolved, on hold, in Process, Dispute
d')
```



In [18]:
```
plt.scatter(df['QUANTITYORDERED'],df['SALES'])
```

Out[18]:
```
<matplotlib.collections.PathCollection at 0x2a180072a90>
```



In [ ]: