```
Name: Sayali Jadhav.
```

```
PRN: 202201070086
```

```
 Roll no:732
```

In [11]:
```
import matplotlib.pyplot as plt
import pandas as pd
import numpy as np

```

In [12]:
```
df=pd.read_csv("sales_data_sample.csv")
df
```

Out[12]:

	ORDERNUMBER	QUANTITYORDERED	PRICEEACH	ORDERLINENUMBER	SALES	ORDER
0	10107	30	95.70	2	2871.00	2/24
1	10121	34	81.35	5	2765.90	5/7/200
2	10134	41	94.74	2	3884.34	7/1/200
3	10145	45	83.26	6	3746.70	8/25
4	10159	49	100.00	14	5205.27	10/10
...	...	...	...	...	...	
2818	10350	20	100.00	15	2244.40	12/2
2819	10373	29	100.00	1	3978.51	1/31
2820	10386	43	100.00	4	5417.57	3/1/200
2821	10397	34	62.24	1	2116.16	3/28
2822	10414	47	65.52	9	3079.44	5/6/200

2823 rows × 16 columns

In [13]:
```
df.columns
```

Out[13]:
```
Index(['ORDERNUMBER', 'QUANTITYORDERED', 'PRICEEACH', 'ORDERLINENUMBER',
       'SALES', 'ORDERDATE', 'STATUS', 'QTR_ID', 'MONTH_ID', 'YEAR_ID',
       'PRODUCTLINE', 'MSRP', 'PRODUCTCODE', 'CUSTOMERNAME', 'COUNTRY',
       'DEALSIZE'],
      dtype='object')
```

In [14]:
```
newdf=df.groupby('COUNTRY')
country=df['COUNTRY'].unique()
sum(newdf.get_group('USA')['SALES'])
```

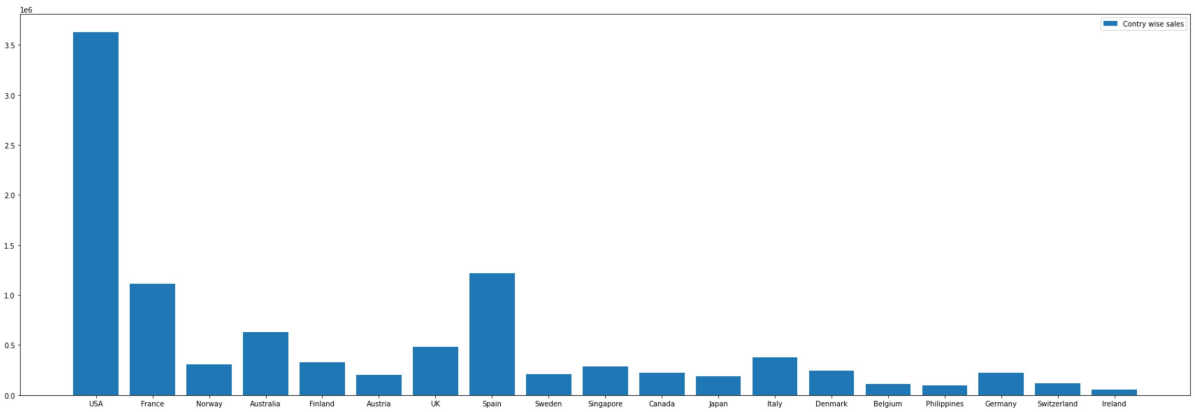
Out[14]: 3627982.83

In [15]:
```
#1 identfy contry wise sales
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}
font1 = {'family':'serif','color':'darked','size':15}
plt.bar(country, sales, label="Contry wise sales")
plt.legend(loc="best")
```

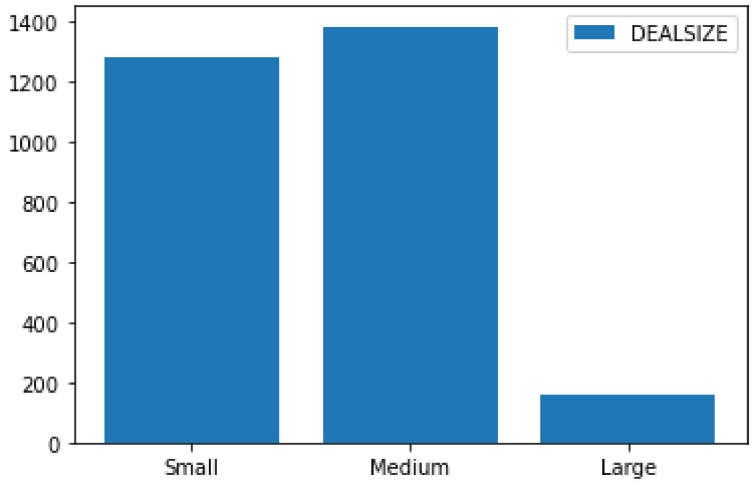
Out[15]: <matplotlib.legend.Legend at 0x23e0680dac0>



In [16]:
```
# 2 identifythe most common DEALSIZE
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")
```

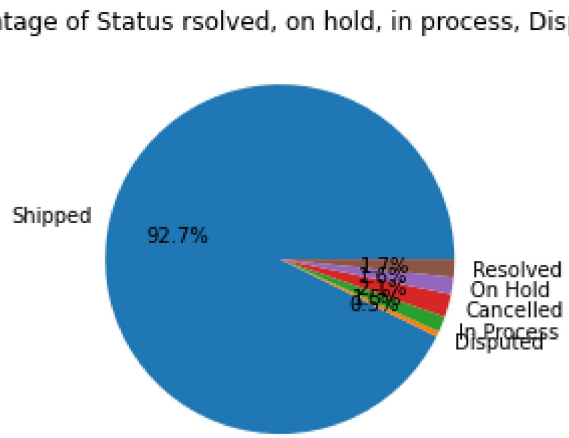
Out[16]: <matplotlib.legend.Legend at 0x23e0680d1c0>



In [17]:
```
#3 find percentage of status resolved, on hold, in process, disputed
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)
```

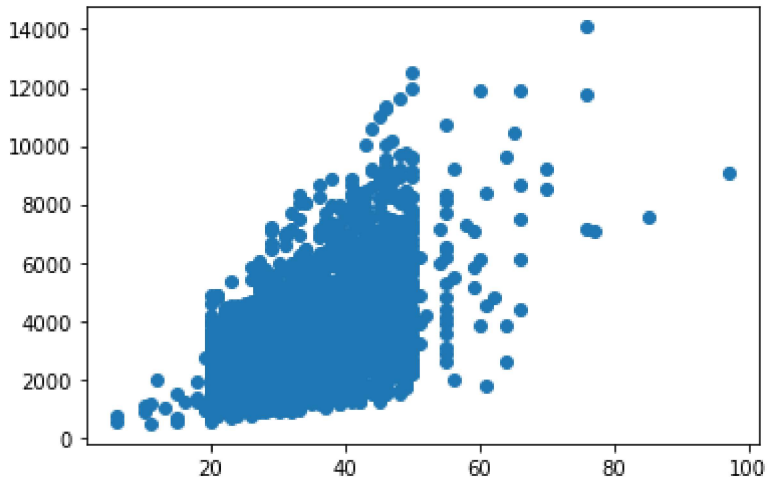
In [18]:
```
plt.pie(percent, labels=status, autopct='%1.1f%%')
plt.title('Percentage of Status rsolved, on hold, in process, Disputed')
```

Out[18]: Text(0.5, 1.0, 'Percentage of Status rsolved, on hold, in process, Disputed')



In [19]:
```
#4 identyfy relationship between Quantity order and sales
plt.scatter(df['QUANTITYORDERED'],df['SALES'])
```

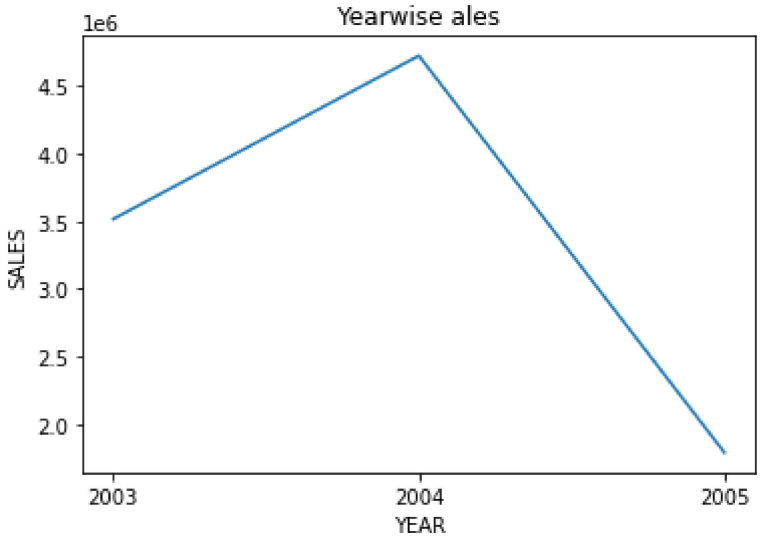
Out[19]: <matplotlib.collections.PathCollection at 0x23e069f4100>



In [20]:
```
#5 identfy year wise sales
year=df['YEAR_ID'].unique()
sales=[]
for yr in year:
    sales.append(sum(newdf.get_group(yr)['SALES']))

plt.plot(year.astype(str),sales)
plt.xlabel('YEAR')
plt.ylabel('SALES')
plt.title('Yearwise ales')
```

Out[20]: Text(0.5, 1.0, 'Yearwise ales')

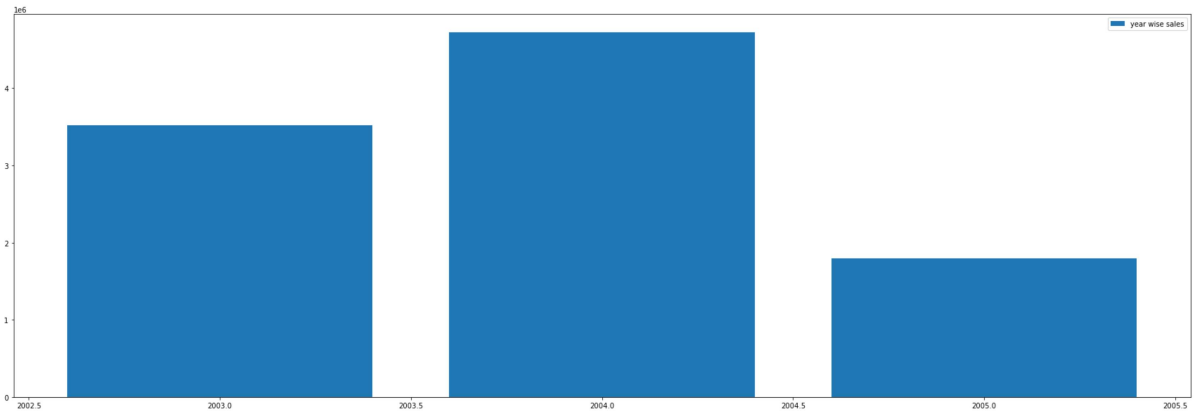


In [22]:
```
#5identfy year wise sales
year=df['YEAR_ID'].unique()
sales=[]
for yr in year:
    sales.append(sum(newdf.get_group(yr)['SALES']))


f= plt.figure()
f.set_figwidth(30)
f.set_figheight(10)

font1 = {'family':'serif','color':'green','size':20}
font1 = {'family':'serif','color':'darked','size':15}
plt.bar(year, sales, label="year wise sales")
plt.legend(loc="best")
```

Out[22]: <matplotlib.legend.Legend at 0x23e06941730>

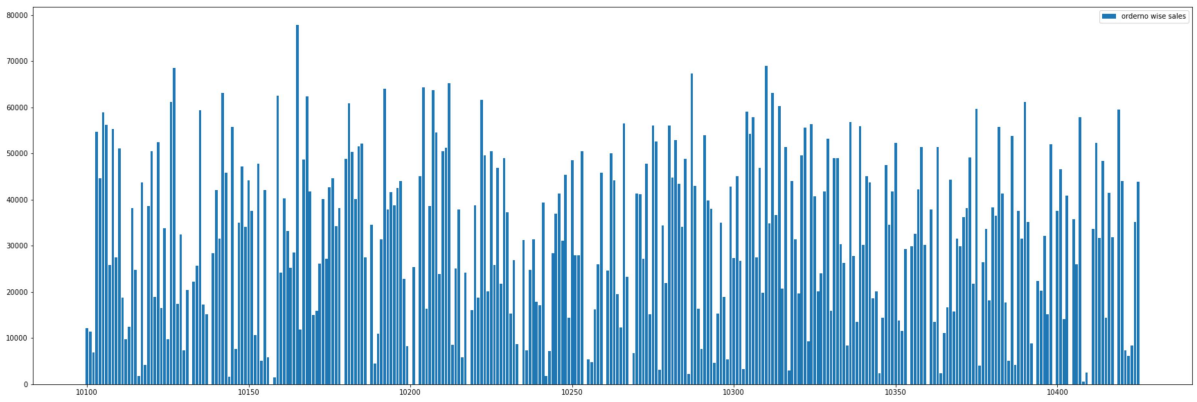


In [29]:
```
#6 identfy ORDERNUMBER wise sales
orderno=df['ORDERNUMBER'].unique()
sales=[]
for on in orderno:
    sales.append(sum(newdf.get_group(on)['SALES']))


f= plt.figure()
f.set_figwidth(30)
f.set_figheight(10)

font1 = {'family':'serif','color':'blue','size':20}
font1 = {'family':'serif','color':'darked','size':15}
plt.bar(orderno, sales, label="orderno wise sales")
plt.legend(loc="best")
```

Out[29]: <matplotlib.legend.Legend at 0x23e0790c790>



In [ ]:
```
1
```