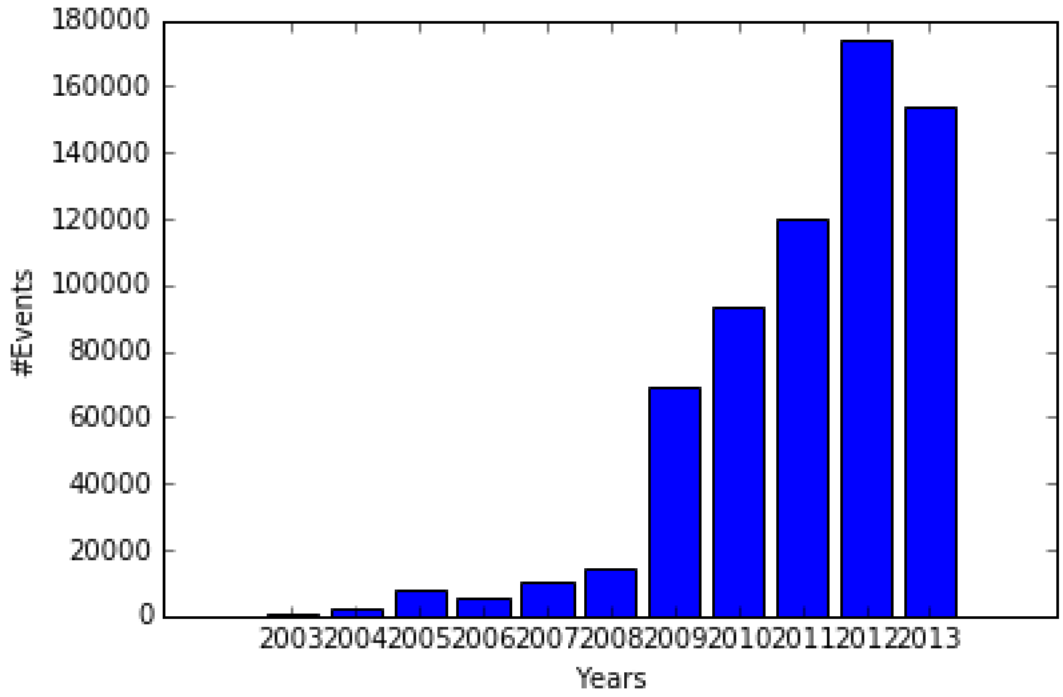
Q2 Plot of Years vs Number of events



Code

```
#closed_tstamp 3889 null val
#created_tstamp 169286 null
#start_tstamp 204946 null
#confirmed_tstamp 227843 null

from pandas import read_csv
import matplotlib.pyplot as plt
import numpy as np

#df = read_csv("events_train_holdout.tsv",sep="\t", error_bad_lines=False)

df['year'] = df.closed_tstamp.str[:4]
grouped=df.groupby('year')
xval = ['2003','2004','2005','2006','2007','2008','2009','2010','2011','2012','2013'
yval = [419,2501,7693,5300,10630,14259,68909,93181,120376,174405,154082]
plt.bar(np.arange(len(xval)),yval,align='center',alpha=2.0)
plt.xticks(np.arange(len(xval)),xval)

plt.ylabel('#Events')
plt.xlabel('Years')

plt.show()
```

```
In [161]: grouped.size()
Out[161]:
year
2003       419
2004      2501
2005      7693
2006      5300
2007     10630
2008     14259
2009     68909
2010     93181
2011    120376
2012    174405
2013    154082
acci      1966
devi       168
dist         4
obst        49
road        37
spec         4
traf        13
dtype: int64
```