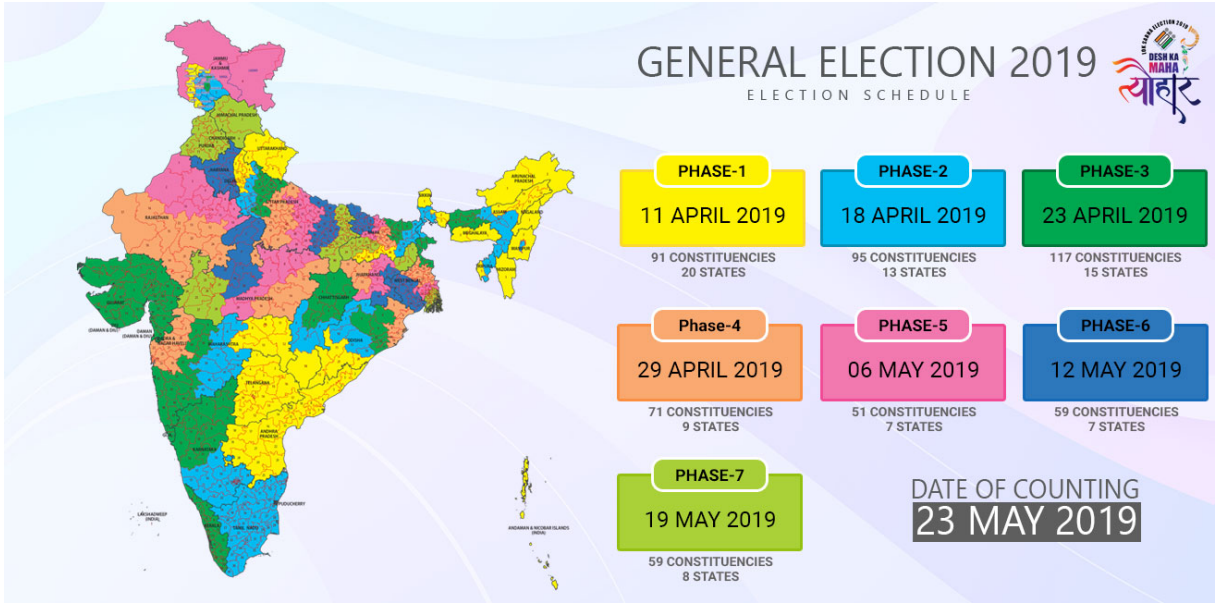
GENERAL ELECTION 2019

In [1]:
```
from IPython.display import display, HTML
display(HTML('''<img src="https://eci.gov.in/img/ge14-schedule1.jpg">'''))
```



DESCRIPTION

With over 600 Million voters voting for 8500+ candidates across 543 constituencies, the general elections in the world's largest democracy are a potential goldmine of data. While there are existing separate datasets about the votes each candidate received and the personal information

of each candidate, there was no comprehensive dataset that included both these information. Thus, this dataset will provide more usability than most existing datasets in this domain.

IMPORTING NECESSARY LIBRARIES FOR DATA VISUALIZATION

In [2]:
```
import pandas as pd
import numpy as np
from matplotlib import pyplot as plt
import seaborn as sms
```

IMPORTIN DATASET FOR VISUALIZATION

In [3]:
```
File=pd.read_csv("project.csv")
```

CHECKING TOP 10

In [4]:
```
File.head(10)
```

Out[4]:

	STATE	CONSTITUENCY	NAME	WINNER	PARTY	SYMBOL	GENDER	CRIMINA
0	Telangana	ADILABAD	SOYAM BAPU RAO	1	BJP	Lotus	MALE	
1	Telangana	ADILABAD	Godam Nagesh	0	TRS	Car	MALE	

	STATE	CONSTITUENCY	NAME	WINNER	PARTY	SYMBOL	GENDER	CRIMINA
2	Telangana	ADILABAD	RATHOD RAMESH	0	INC	Hand	MALE	
3	Telangana	ADILABAD	NOTA	0	NOTA	NaN	NaN	
4	Uttar Pradesh	AGRA	Satyapal Singh Baghel	1	BJP	Lotus	MALE	
5	Uttar Pradesh	AGRA	Manoj Kumar Soni	0	BSP	Elephant	MALE	
6	Uttar Pradesh	AGRA	Preeta Harit	0	INC	Hand	FEMALE	
7	Maharashtra	AHMADNAGAR	Dr. SUJAY RADHAKRISHNA VIKHEPATIL	1	BJP	Lotus	MALE	
8	Maharashtra	AHMADNAGAR	SANGRAM ARUNKAKA JAGTAP	0	NCP	Clock	MALE	
9	Maharashtra	AHMADNAGAR	SUDHAKAR LAXMAN AVHAD	0	VBA	Cup & Saucer	MALE	

Describing Our data set

In [5]:
```
File.describe()
```

Out[5]:

	WINNER	CRIMINAL\nCASES	AGE	GENERAL\nVOTES	POSTAL\nVOTES	TOTAL
count	2263.000000	2018.000000	2018.000000	2.263000e+03	2263.000000	2.26
mean	0.238179	1.453915	52.273538	2.615991e+05	990.710561	2.62
std	0.426064	7.636973	11.869373	2.549906e+05	1602.839174	2.55
min	0.000000	0.000000	25.000000	1.339000e+03	0.000000	1.34
25%	0.000000	0.000000	43.250000	2.103450e+04	57.000000	2.11
50%	0.000000	0.000000	52.000000	1.539340e+05	316.000000	1.54
75%	0.000000	1.000000	61.000000	4.858040e+05	1385.000000	4.87
max	1.000000	240.000000	86.000000	1.066824e+06	19367.000000	1.06

Checking last 5 data set

In [6]:
```
File.tail(5)
```

Out[6]:

	STATE	CONSTITUENCY	NAME	WINNER	PARTY	SYMBOL	GENDER	CRIMINAL\n
2258	Maharashtra	YAVATMAL-WASHIM	Anil Jayram Rathod	0	IND	SHIP	MALE	
2259	Telangana	ZAHIRABAD	B.B.PATIL	1	TRS	Car	MALE	
2260	Telangana	ZAHIRABAD	MADAN MOHAN RAO	0	INC	Hand	MALE	

	STATE	CONSTITUENCY	NAME	WINNER	PARTY	SYMBOL	GENDER	CRIMINAL\n
2261	Telangana	ZAHIRABAD	BANALA LAXMA REDDY	0	BJP	Lotus	MALE	
2262	Telangana	ZAHIRABAD	NOTA	0	NOTA	NaN	NaN	

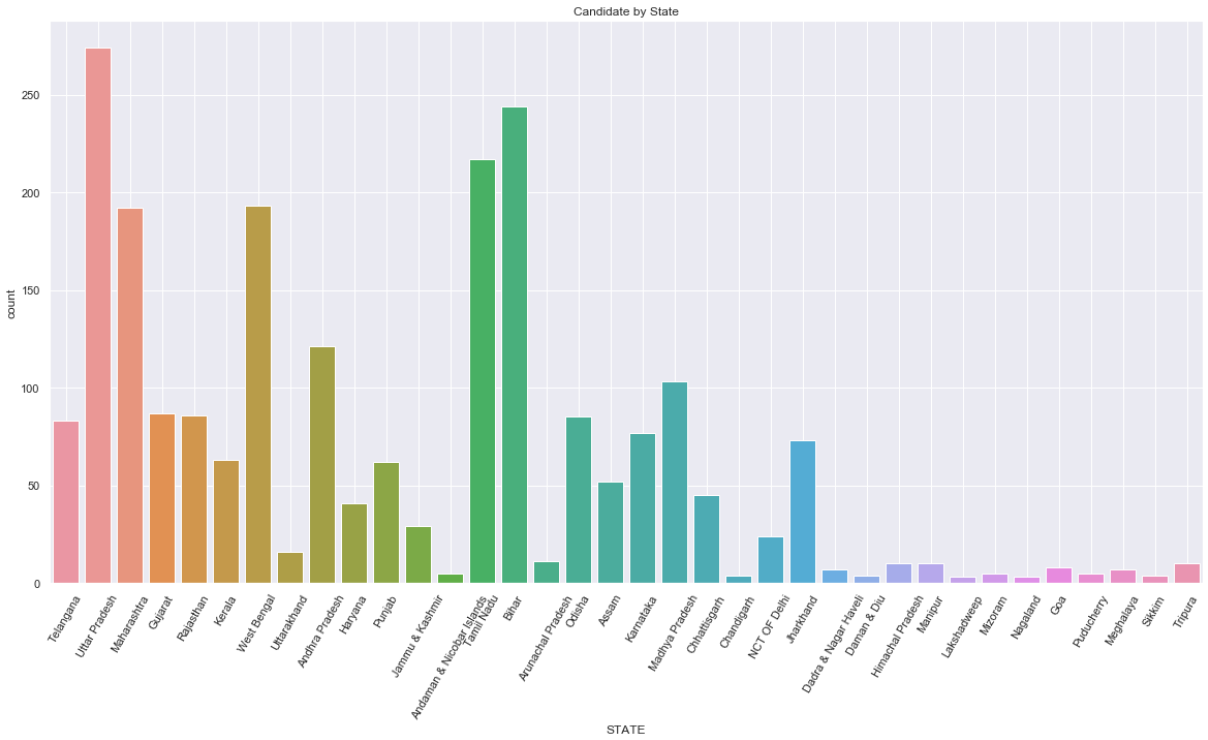
Number of Rows and Column

In [7]:
```
print('Number of rows and columns :',File.shape)
```

```
Number of rows and columns : (2263, 19)
```

Graph-STATE VS CANDIDATE

In [8]:
```
plt.figure(figsize=(20,10))
sns.set(style="dark")
a=sns.countplot(x="STATE",data=File)
plt.grid()
plt.title('Candidate by State')
a=a.set_xticklabels(a.get_xticklabels(),rotation=60)
```



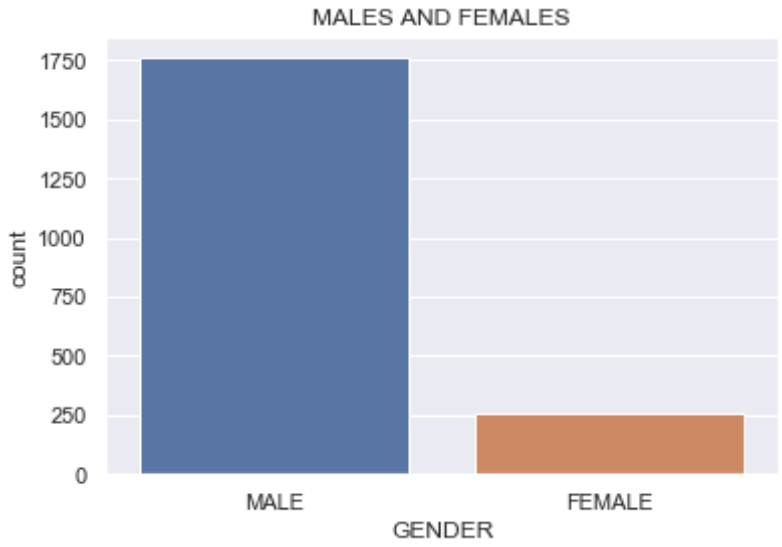
OBSERVATION FROM ABOVE GRAPH

->Above graph tell us Uttar Pradesh has highest number of Candidates. ->The second largest candidate is in Bihar. ->Lakshadweeb has lowest number of Candidates.

MALES VS FEMALES IN ELECTION

In [16]:
```
plt.title("MALES AND FEMALES")
plt.grid()
sms.countplot(x ='GENDER', data = File)
```

Out[16]:
```
<matplotlib.axes._subplots.AxesSubplot at 0x1d711827d88>
```

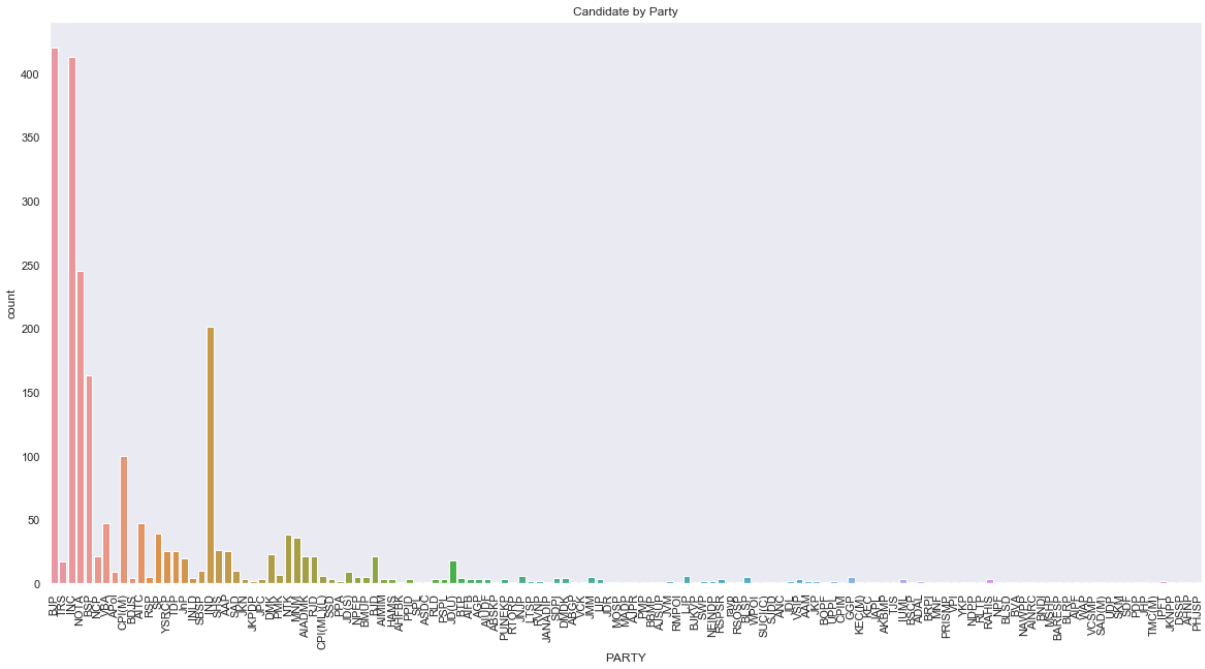


OBSERVATION FROM ABOVE GRAPH

->We can say that Females in Election is much much lesser than Males.

Parties in the election

In [10]:
```
plt.figure(figsize=(20,10))
a=sms.countplot(x='PARTY', data=File)
plt.title('Candidate by Party')
a=a.set_xticklabels(a.get_xticklabels(),rotation=90)
```

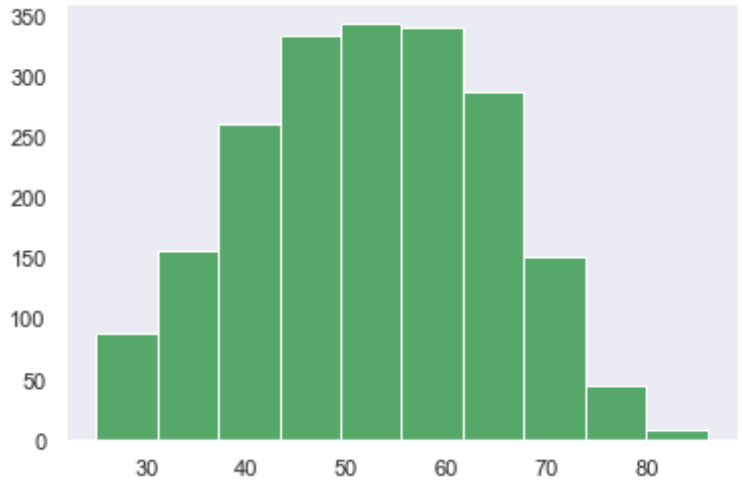


OBSERVATIONS:

->BJP has highest candidates ->congress is second highest

In [11]:
```
plt.hist(File["AGE"],color="g");
```

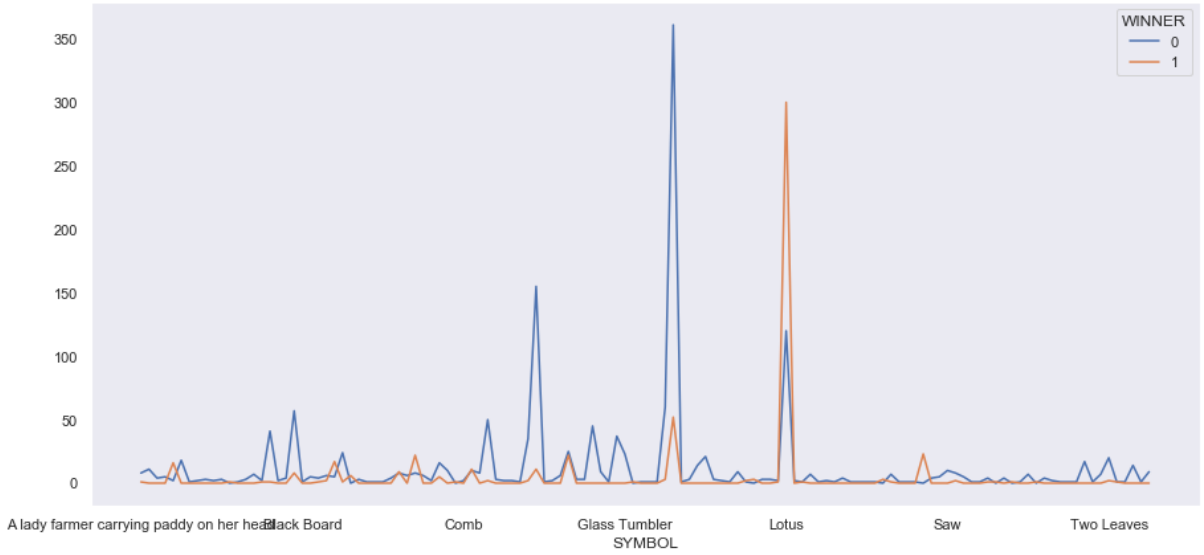
```
C:\Users\deepa\anaconda3\lib\site-packages\numpy\lib\histograms.py:839:
RuntimeWarning: invalid value encountered in greater_equal
  keep = (tmp_a >= first_edge)
C:\Users\deepa\anaconda3\lib\site-packages\numpy\lib\histograms.py:840:
RuntimeWarning: invalid value encountered in less_equal
  keep &= (tmp_a <= last_edge)
```



GRAPH- SYMBOL VS WINNER

In [12]:
```
pd.crosstab(File['SYMBOL'],File['WINNER']).plot(figsize=(15,7))
```

Out[12]:
```
<matplotlib.axes._subplots.AxesSubplot at 0x1d712093488>
```

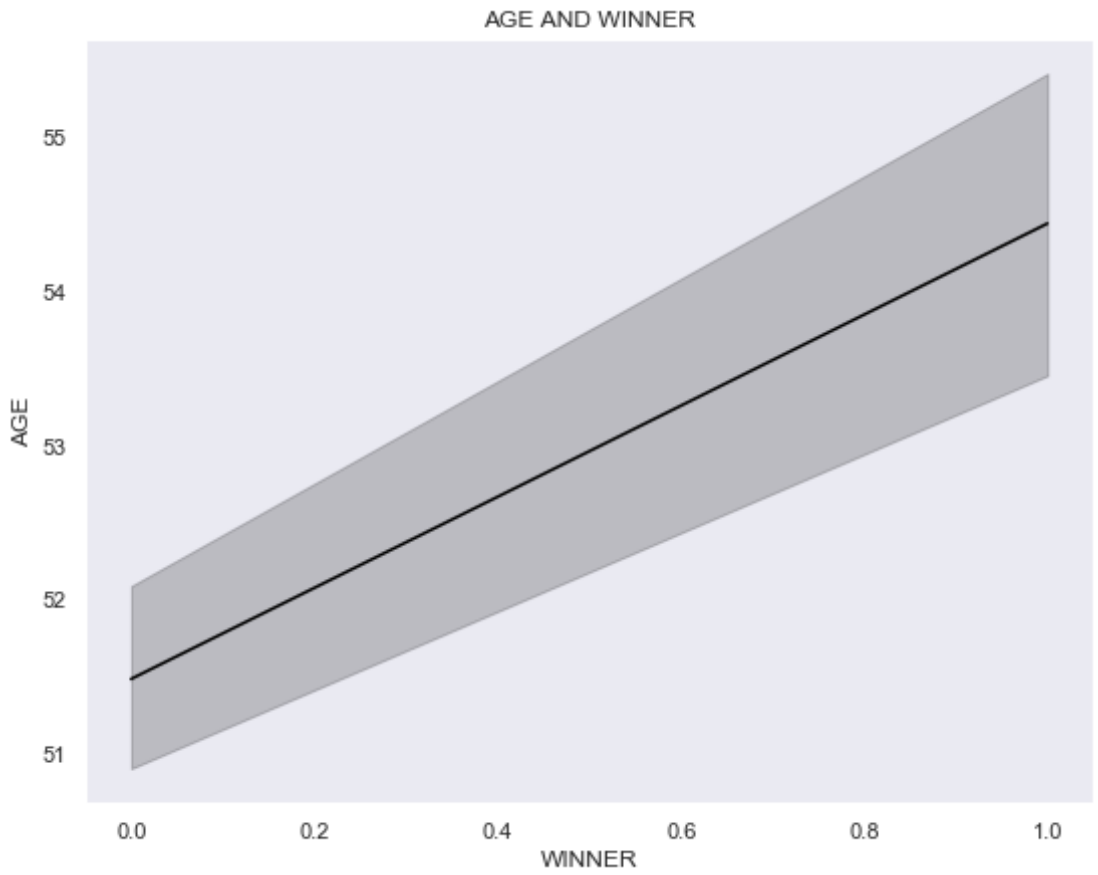


Observation

->Glass tumber party member has highest losers ->Lotus party member has highest no of candiates

GRAPH AGE VS WINNERS

In [17]:
```
plt.figure(figsize=(9,7))
sms.lineplot(x=File['WINNER'],y=File['AGE'],color="black");
plt.title("AGE AND WINNER");
plt.xlabel("WINNER");
plt.ylabel("AGE");
```



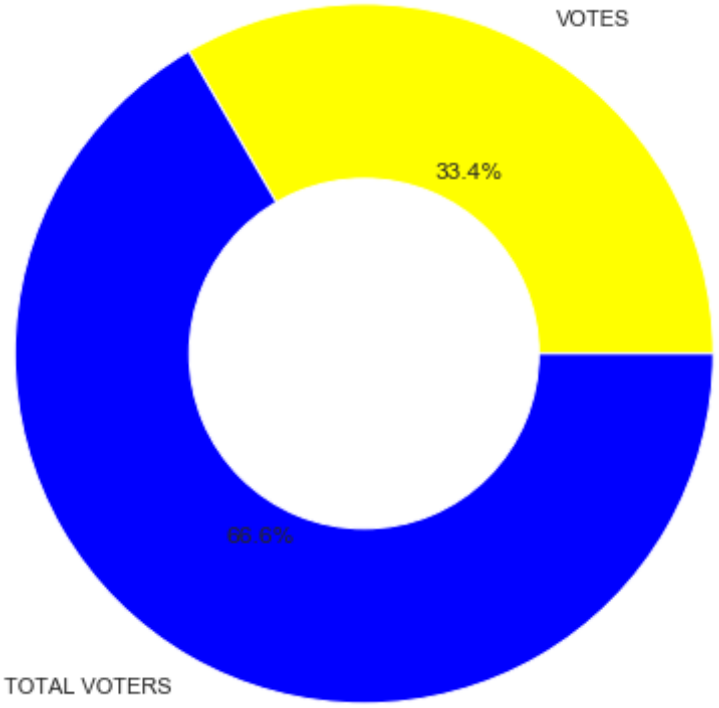
OBERVATIONS

->This shows that winners volume is less at the age of 52 and the winner values is high at the age of 55.

GRAPH BETWEEN TOTAL VOTES VS VOTES

This graph will show the No of votes vs votes.

In [14]:
```
a= File['TOTAL\nVOTES'].sum()
b=a+File['GENERAL\nVOTES'].sum()
mains=["VOTES","TOTAL VOTERS"]
quantity=[a,b]
plt.pie(quantity,labels=mains,radius=2,autopct="%0.1f%%",colors=["yello
w","blue"])
plt.pie([b],colors=["white"],radius=1)
plt.show()
```



OBESERVATIONS

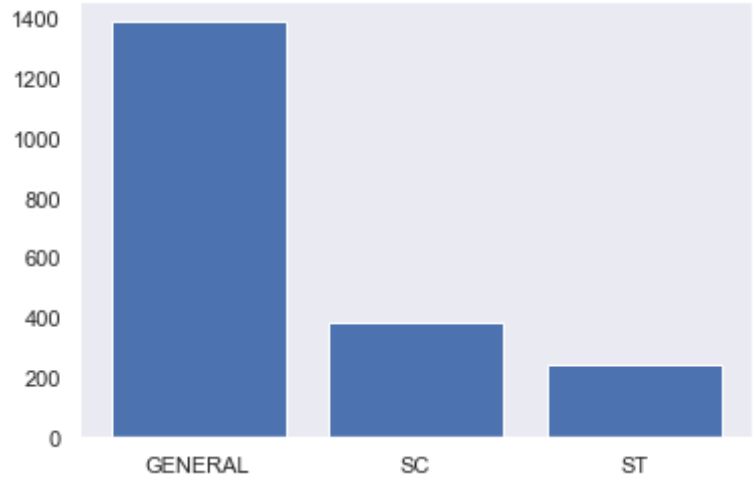
This graph show the Total number of votes vs Total votes.

GRAPH SHOWING CATEGORIES

This graph will show the no of different category from the column "CATEGORY" thats is General,SC,ST

In [15]:
```
b=File.groupby('CATEGORY').size().reset_index(name='NUMBER')
data = {'GENERAL':1392, 'SC':383, 'ST':243}
keys = data.keys()
values =data.values()
plt.bar(keys, values)
```

Out[15]:
```
<BarContainer object of 3 artists>
```



OBERVATIONS

->The highest no of candidates is of General Category ->The second Highest no Candidates is of SC and the last is of ST