Analysis of legislative assembly election 2017

Let's analyse the important factors which played a major role in candidate success.

As a first step, read the csv file which has the details of candidates.

In [4]:
```
import pandas as pd

file =pd.read_csv('/home/arvind/Desktop/GL_prj_4/LA_2017.csv')
print (file.head())
```

```
  ST_CODE ST_NAME  MONTH  YEAR  DIST_NAME  AC_NO  AC_NAME AC_TYPE  \
0     S05     Goa      3  2017  North Goa      1  Mandrem     GEN
1     S05     Goa      3  2017  North Goa      1  Mandrem     GEN
2     S05     Goa      3  2017  North Goa      1  Mandrem     GEN
3     S05     Goa      3  2017  North Goa      1  Mandrem     GEN
4     S05     Goa      3  2017  North Goa      1  Mandrem     GEN

                                CAND_NAME CAND_SEX CAND_CATEGORY  CAND_AGE  \
0                 DAYANAND RAGHUNATH SOPTE        M           GEN      53.0
1                       LAXMIKANT PARSEKAR        M           GEN      60.0
2                    SHRIDHAR LADU MANJREKAR        M           GEN      69.0
3  DEVENDRA KRISHNAJI PRABHU PARSEKAR DESAI        M           GEN      53.0
4                         None of the Above      NaN           NaN       NaN

  PARTYABBRE  TOTALVALIDVOTESPOLLED  POSITION
0        INC                  16490         1
1        BJP                   9371         2
2        MAG                    678         3
3       AAAP                    620         4
4       NOTA                    415         5
```

Finding number of male and female candidates.

In [10]:
```
import numpy as np
import pandas as pd
%matplotlib inline
import matplotlib.pyplot as plt
import seaborn as sns

file =pd.read_csv('/home/arvind/Desktop/GL_prj_4/LA_2017.csv')
#print (file.head())

#candidate gender distribution
candidate_sex = file["CAND_SEX"].value_counts()
candidate_sex
```

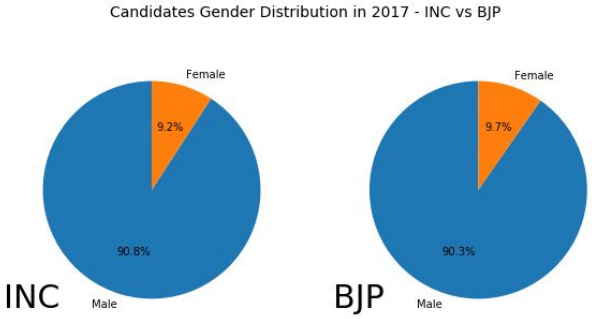
Out[10]:
```
M    6576
F     663
0       4
Name: CAND_SEX, dtype: int64
```

Candidate gender distribution between INC and BJP

In [16]:
```
plt.figure(figsize=(10,5))
plt.subplot(1,2,1)
plt.pie(file[(file["PARTYABBRE"]=='INC')]['CAND_SEX'].value_counts(), labels=['Male','Female'],autopct='%1.1f%%', s
tartangle=90)

fig = plt.gcf()
fig.suptitle("Candidates Gender Distribution in 2017 - INC vs BJP", fontsize=14)
ax = fig.gca()
label = ax.annotate("INC", xy=(-1.1,-1), fontsize=30, ha="center",va="center")
ax.axis('off')
ax.set_aspect('equal')
ax.autoscale_view()

plt.subplot(1,2,2)
plt.pie(file[(file["PARTYABBRE"]=='BJP')]['CAND_SEX'].value_counts(), labels=['Male','Female'],autopct='%1.1f%%', s
tartangle=90)
fig = plt.gcf()
ax = fig.gca()
label = ax.annotate("BJP", xy=(-1.1,-1), fontsize=30, ha="center",va="center")
ax.axis('off')
ax.set_aspect('equal')
ax.autoscale_view()
plt.show();
```

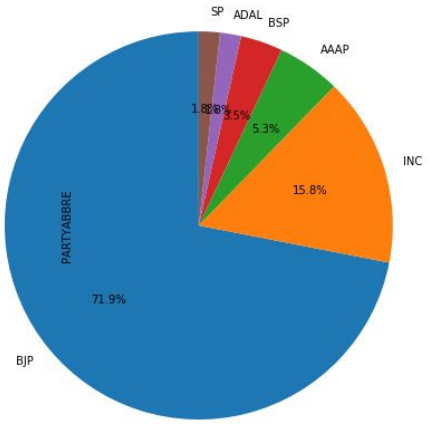


Number of women winners

In [17]:
```
df_womenwinners = file[(file['POSITION']==1)&(file["CAND_SEX"]=="F")]

ax = df_womenwinners["PARTYABBRE"].value_counts().plot(kind="pie",radius=2,autopct='%1.1f%%', startangle=90)
x = df_womenwinners["PARTYABBRE"].value_counts()
x
```

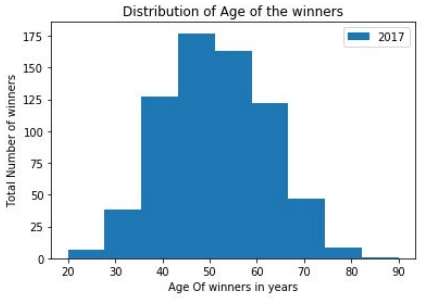
Out[17]:
```
BJP     41
INC      9
AAAP     3
BSP      2
ADAL     1
SP       1
Name: PARTYABBRE, dtype: int64
```



Age of winners

```
In [21]: Age=file[(file.POSITION==1) & (file.YEAR==2017)]['CAND_AGE'].tolist()
         bins = np.linspace(20, 90, 10)
         plt.hist([Age], bins, label=['2017'])

         plt.legend(loc='upper right')
         plt.xlabel('Age Of winners in years')
         plt.ylabel('Total Number of winners')
         plt.title('Distribution of Age of the winners')
         plt.show()
```



Party Wise winners

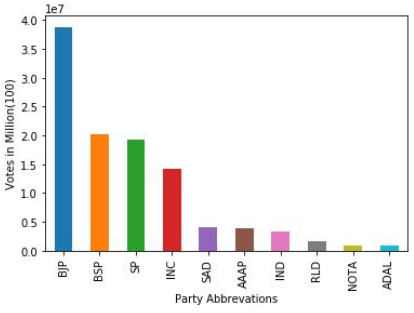
```
In [22]: df_winners = file[file['POSITION']==1]
         DF = df_winners['PARTYABBRE'].value_counts().head(10)
         DF
```

```
Out[22]: BJP     406
         INC     140
         SP       47
         AAAP     20
         BSP      19
         SAD      15
         ADAL      9
         IND       9
         SBSP      4
         NPEP      4
         Name: PARTYABBRE, dtype: int64
```

TOtal valid valid polled

```
In [23]: votespartywise = file.groupby('PARTYABBRE')['TOTALVALIDVOTESPOLLED'].sum()
         x09 = votespartywise.sort_values(ascending=False)[:10].plot(kind="bar")
         x09.set_xlabel('Party Abbrevations')
         x09.set_ylabel('Votes in Million(100)')
         votespartywise.sort_values(ascending=False)[:10]
```

```
Out[23]: PARTYABBRE
         BJP     38800694
         BSP     20101331
         SP      19158491
         INC     14196036
         SAD      4027773
         AAAP     3915496
         IND      3277542
         RLD      1557844
         NOTA      948827
         ADAL      851336
         Name: TOTALVALIDVOTESPOLLED, dtype: int64
```



Poll Percentage of electors:

In [34]:
```
electorfile=pd.read_csv('/home/arvind/Desktop/GL_prj_4/electors_2017.csv')

pollper = electorfile.groupby("ST_NAME").mean()
LS09 = pollper[['POLLPERCENTAGE']].round(1).sort_values('POLLPERCENTAGE',ascending=False)
ax1 =LS09['POLLPERCENTAGE'].plot(kind='bar',figsize=(20, 15))
for p in ax1.patches:
    ax1.annotate(format(p.get_height()), (p.get_x()+0.1, p.get_height()+2),fontsize=12)
```