In [1]:
```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
```

Read the bond purchaser file, make the unique alphanumeric code by combining the 'Prefix' and 'Bond Number' fields. Delete the fields that we are not interested. Rename the 'Denomination' column as 'Amount paid'.

In [2]:
```
pdf = pd.read_csv('purchasers.csv')
pdf['ANcode'] = pdf['Prefix'] + pdf['Bond Number'].astype(str)  # make the alphanumeric code
pdf = pdf.set_index('ANcode')

# Remove columns that we are not using
pdf = pdf.drop(columns =['Prefix', 'Bond Number', 'Reference No (URN)', 'Date of Expiry'], axis = 'columns')
pdf = pdf.drop(columns =['Issue Branch Code', 'Issue Teller', 'Journal Date', 'Sr No', 'Status' ], axis = '

pdf.rename(columns = {'Denominations':'Amount paid'}, inplace = True)   # rename 'Denominations' to 'Amount

pdf['Amount paid'] = pdf['Amount paid'].apply(lambda x:float(x.replace(',','')))  # comma to dot
pdf['Amount paid'] = pdf['Amount paid'].apply(lambda x:x * 1e-7)                  # convert to crores
ptotal = pdf['Amount paid'].sum()
print('Number of Bonds purchased = %d (Rs. %10.4f Crores)'%(len(pdf), ptotal))
```

```
Number of Bonds purchased = 18871 (Rs. 12155.5132 Crores)
```

Read the bond receiver file, make the unique alphanumeric code by combining the 'Prefix' and 'Bond Number' fields. Delete the fields that we are not interested. Rename the 'Denomination' column as 'Amount received'.

In [3]:
```
rdf=pd.read_csv('receivers.csv')
rdf['ANcode'] = rdf['Prefix'] + rdf['Bond Number'].astype(str)
rdf = rdf.set_index('ANcode')

rdf = rdf.drop(columns = ['Prefix', 'Bond Number', 'Pay Branch Code', 'Pay Teller'], axis = 'columns')
rdf = rdf.drop(columns = ['Account no. of Political Party', 'Sr No'], axis = 'columns')

rdf.rename(columns = {'Denominations':'Amount received'}, inplace = True)

rdf['Amount received'] = rdf['Amount received'].apply(lambda x:float(x.replace(',','')))
rdf['Amount received'] = rdf['Amount received'].apply(lambda x:x * 1e-7)  # convert to crores
rtotal = rdf['Amount received'].sum()
print('Number of Bonds encashed = %d (Rs. %10.4f Crores)'%(len(rdf), rtotal))
print('Source unknown %d bonds (Rs. %10.4f)'%(len(rdf)-len(pdf), rtotal - ptotal))
```

```
Number of Bonds encashed = 20421 (Rs. 12769.0893 Crores)
Source unknown 1550 bonds (Rs.   613.5761)
```

match the bond purchaser and receiver data

Combine the the purchaser and receiver Dataframes. Result saved to 'combined.csv'.

In [4]:
```
combined = pd.concat([pdf, rdf], axis=1, sort = True)
combined['DateTime'] = pd.to_datetime(rdf['Date of Encashment'])
combined.to_csv('combined.csv')
```

Donations from different types of companies

pharma, infrastructure, construction, media etc.

Pharma companies

In [5]:
```
pharma =combined[combined['Name of the Purchaser'].str.contains('pharma', na = False, case = False)]
pharma.to_csv('pharma.csv')
print(pharma['Amount received'].sum())
#pharma.groupby(by='Name of the Purchaser')['Amount received'].sum().nlargest(15)
pharma.groupby(by='Name of the Political Party')['Amount received'].sum().nlargest(15)
```

```
348.2
```

Out[5]:
```
Name of the Political Party
BHARATIYA JANATA PARTY                                        238.95
BHARAT RASHTRA SAMITHI                                         55.00
PRESIDENT, ALL INDIA CONGRESS COMMITTEE                        18.25
TELUGU DESAM PARTY                                             16.50
SIKKIM KRANTIKARI MORCHA                                        7.00
JANASENA PARTY                                                  5.00
ADYAKSHA SAMAJVADI PARTY                                        3.00
YSR  CONGRESS PARTY  (YUVAJANA SRAMIKA RYTHU CONGRESS PARTY)    3.00
AAM AADMI PARTY                                                 1.00
SIKKIM DEMOCRATIC FRONT                                         0.50
Name: Amount received, dtype: float64
```

Hotel

In [6]:
```
hotel =combined[combined['Name of the Purchaser'].str.contains('hotel', na = False, case = False)]
hotel.to_csv('pharma.csv')
print(hotel['Amount received'].sum())
#hotel.groupby(by='Name of the Purchaser')['Amount received'].sum().nlargest(15)
hotel.groupby(by='Name of the Political Party')['Amount received'].sum().nlargest(15)
```

```
1370.75
```

Out[6]:
```
Name of the Political Party
ALL INDIA TRINAMOOL CONGRESS                                    542.00
DRAVIDA MUNNETRA KAZHAGAM (DMK)                                 503.00
YSR  CONGRESS PARTY  (YUVAJANA SRAMIKA RYTHU CONGRESS PARTY)    154.00
BHARATIYA JANATA PARTY                                          105.00
PRESIDENT, ALL INDIA CONGRESS COMMITTEE                          50.75
SIKKIM KRANTIKARI MORCHA                                         11.00
SIKKIM DEMOCRATIC FRONT                                           5.00
Name: Amount received, dtype: float64
```

Media companies

In [7]:
```
media =combined[combined['Name of the Purchaser'].str.contains('media', na = False, case = False)]
media.to_csv('pharma.csv')
print(media['Amount received'].sum())
#media.groupby(by='Name of the Purchaser')['Amount received'].sum().nlargest(15)
media.groupby(by='Name of the Political Party')['Amount received'].sum().nlargest(15)
```

```
42.0
```

Out[7]:
```
Name of the Political Party
BHARATIYA JANATA PARTY    42.0
Name: Amount received, dtype: float64
```

Energy Companies

In [8]:
```
ener =combined[combined['Name of the Purchaser'].str.contains('energy', na = False, case = False)]
ener.groupby(by='Name of the Political Party')['Amount received'].sum().nlargest(15)
```

Out[8]:
```
Name of the Political Party
ALL INDIA TRINAMOOL CONGRESS                                  281.0
BHARATIYA JANATA PARTY                                        115.0
BHARAT RASHTRA SAMITHI                                         19.0
PRESIDENT, ALL INDIA CONGRESS COMMITTEE                        18.0
YSR  CONGRESS PARTY  (YUVAJANA SRAMIKA RYTHU CONGRESS PARTY)   16.0
Name: Amount received, dtype: float64
```

Infrastructure companies

In [9]:
```
infra =combined[combined['Name of the Purchaser'].str.contains('infra', na = False, case = False)]
print(infra['Amount received'].sum())
infra.groupby(by='Name of the Political Party')['Amount received'].sum().nlargest(15)
```

```
1654.51
```

Out[9]:
```
Name of the Political Party
BHARATIYA JANATA PARTY                                        984.40
BHARAT RASHTRA SAMITHI                                        247.84
PRESIDENT, ALL INDIA CONGRESS COMMITTEE                       117.64
ALL INDIA TRINAMOOL CONGRESS                                  111.75
DRAVIDA MUNNETRA KAZHAGAM (DMK)                                85.00
YSR  CONGRESS PARTY  (YUVAJANA SRAMIKA RYTHU CONGRESS PARTY)   37.00
TELUGU DESAM PARTY                                             34.88
BIHAR PRADESH JANTA DAL(UNITED)                                11.00
ADYAKSHA SAMAJVADI PARTY                                       10.00
JANATA DAL ( SECULAR )                                          5.00
JANASENA PARTY                                                  4.00
SHIVSENA                                                        3.00
SHIROMANI AKALI DAL                                             1.50
NATIONALIST CONGRESS PARTY MAHARASHTRA PRADESH                  1.00
JHARKHAND MUKTI MORCHA                                          0.50
Name: Amount received, dtype: float64
```

Construction companies

In [10]:
```
constru =combined[combined['Name of the Purchaser'].str.contains('constru', na = False, case = False)]
print(constru['Amount received'].sum())
constru.groupby(by='Name of the Political Party')['Amount received'].sum().nlargest(15)
```

```
258.82
```

Out[10]:
```
Name of the Political Party
BHARATIYA JANATA PARTY                          100.23
SHIVSENA                                         85.00
BHARAT RASHTRA SAMITHI                           54.09
JANATA DAL ( SECULAR )                           10.00
PRESIDENT, ALL INDIA CONGRESS COMMITTEE           6.50
NATIONALIST CONGRESS PARTY MAHARASHTRA PRADESH    2.00
AAM AADMI PARTY                                   1.00
Name: Amount received, dtype: float64
```

Engineering Companies

In [11]:
```
engi =combined[combined['Name of the Purchaser'].str.contains('engineer', na = False, case = False)]
print(engi['Amount received'].sum())
engi.groupby(by='Name of the Political Party')['Amount received'].sum().nlargest(15)
```

```
1119.45
```

Out[11]:
```
Name of the Political Party
BHARATIYA JANATA PARTY                                    718.75
BHARAT RASHTRA SAMITHI                                    202.70
DRAVIDA MUNNETRA KAZHAGAM (DMK)                            85.00
YSR  CONGRESS PARTY  (YUVAJANA SRAMIKA RYTHU CONGRESS PARTY)    37.00
TELUGU DESAM PARTY                                         28.00
PRESIDENT, ALL INDIA CONGRESS COMMITTEE                    24.00
BIHAR PRADESH JANTA DAL(UNITED)                            10.00
BIJU JANATA DAL                                             5.00
JANATA DAL ( SECULAR )                                      5.00
JANASENA PARTY                                              4.00
Name: Amount received, dtype: float64
```

Mining Companies

In [12]:
```
mining =combined[combined['Name of the Purchaser'].str.contains('mining', na = False, case = False)]
print(mining['Amount received'].sum())
mining.groupby(by='Name of the Political Party')['Amount received'].sum().nlargest(15)
```

```
234.5
```

Out[12]:
```
Name of the Political Party
BIJU JANATA DAL                             174.5
BHARATIYA JANATA PARTY                       52.5
JHARKHAND MUKTI MORCHA                        3.0
ALL INDIA TRINAMOOL CONGRESS                  2.5
PRESIDENT, ALL INDIA CONGRESS COMMITTEE       2.0
Name: Amount received, dtype: float64
```

Metals

In [13]:
```
metal =combined[combined['Name of the Purchaser'].str.contains('metal', na = False, case = False)]
print(metal['Amount received'].sum())
metal.groupby(by='Name of the Political Party')['Amount received'].sum().nlargest(15)
```

```
81.99999999999997
```

Out[13]:
```
Name of the Political Party
BIJU JANATA DAL                             45.0
BHARATIYA JANATA PARTY                      16.5
PRESIDENT, ALL INDIA CONGRESS COMMITTEE     14.5
ALL INDIA TRINAMOOL CONGRESS                 5.5
AAM AADMI PARTY                              0.5
Name: Amount received, dtype: float64
```

Cement

In [14]:
```
cem =combined[combined['Name of the Purchaser'].str.contains('cement', na = False, case = False)]
print(cem['Amount received'].sum())
#cem.groupby(by='Name of the Purchaser')['Amount received'].sum().nlargest(15)
cem.groupby(by='Name of the Political Party')['Amount received'].sum().nlargest(15)
```

```
233.5
```

Out[14]:
```
Name of the Political Party
BHARATIYA JANATA PARTY                                          77.5
BIJU JANATA DAL                                                 56.5
PRESIDENT, ALL INDIA CONGRESS COMMITTEE                         39.5
YSR  CONGRESS PARTY  (YUVAJANA SRAMIKA RYTHU CONGRESS PARTY)    24.5
ALL INDIA TRINAMOOL CONGRESS                                    20.5
TELUGU DESAM PARTY                                               5.5
DRAVIDA MUNNETRA KAZHAGAM (DMK)                                  4.0
SHIVSENA                                                         3.0
BIHAR PRADESH JANTA DAL(UNITED)                                  1.0
RASHTRIYA JANTA DAL                                              1.0
BHARAT RASHTRA SAMITHI                                           0.5
Name: Amount received, dtype: float64
```

Pharma Companies facing investigation

Hetero labs, Torrent pharma, zydus healthcare, glenmark, cipla, ipca labs, intas pharma

In [15]:
```
het =combined[combined['Name of the Purchaser'].str.contains('hetero', na = False, case = False)]

het.groupby(by='Name of the Political Party')['Amount received'].sum().nlargest(15)
```

Out[15]:
```
Name of the Political Party
BHARAT RASHTRA SAMITHI    50.0
BHARATIYA JANATA PARTY    10.0
Name: Amount received, dtype: float64
```

In [16]:
```
tor =combined[combined['Name of the Purchaser'].str.contains('torrent pha', na = False, case = False)]

tor.groupby(by='Name of the Political Party')['Amount received'].sum().nlargest(15)
```

Out[16]:
```
Name of the Political Party
BHARATIYA JANATA PARTY                      61.0
SIKKIM KRANTIKARI MORCHA                     7.0
PRESIDENT, ALL INDIA CONGRESS COMMITTEE      5.0
ADYAKSHA SAMAJVADI PARTY                     3.0
AAM AADMI PARTY                              1.0
SIKKIM DEMOCRATIC FRONT                      0.5
Name: Amount received, dtype: float64
```

In [17]:
```
zyd =combined[combined['Name of the Purchaser'].str.contains('zydus health', na = False, case = False)]

zyd.groupby(by='Name of the Political Party')['Amount received'].sum().nlargest(15)
```

Out[17]:
```
Name of the Political Party
BHARATIYA JANATA PARTY                      18.0
SIKKIM KRANTIKARI MORCHA                     8.0
PRESIDENT, ALL INDIA CONGRESS COMMITTEE      3.0
Name: Amount received, dtype: float64
```

In [18]:
```
glen =combined[combined['Name of the Purchaser'].str.contains('glenmark', na = False, case = False)]

glen.groupby(by='Name of the Political Party')['Amount received'].sum().nlargest(15)
```

Out[18]:
```
Name of the Political Party
BHARATIYA JANATA PARTY    9.75
Name: Amount received, dtype: float64
```

In [19]:
```
glen =combined[combined['Name of the Purchaser'].str.contains('glenmark', na = False, case = False)]

glen.groupby(by='Name of the Political Party')['Amount received'].sum().nlargest(15)
```

Out[19]:
```
Name of the Political Party
BHARATIYA JANATA PARTY    9.75
Name: Amount received, dtype: float64
```

In [20]:
```
cip =combined[combined['Name of the Purchaser'].str.contains('cipla', na = False, case = False)]

cip.groupby(by='Name of the Political Party')['Amount received'].sum().nlargest(15)
```

Out[20]:
```
Name of the Political Party
BHARATIYA JANATA PARTY                      37.0
PRESIDENT, ALL INDIA CONGRESS COMMITTEE      2.2
Name: Amount received, dtype: float64
```

In [21]:
```
ipca =combined[combined['Name of the Purchaser'].str.contains('ipca', na = False, case = False)]
ipca.groupby(by='Name of the Political Party')['Amount received'].sum().nlargest(15)
```

Out[21]:
```
Name of the Political Party
BHARATIYA JANATA PARTY      10.0
SIKKIM KRANTIKARI MORCHA     3.5
Name: Amount received, dtype: float64
```

In [22]:
```
intas =combined[combined['Name of the Purchaser'].str.contains('intas', na = False, case = False)]

intas.groupby(by='Name of the Political Party')['Amount received'].sum().nlargest(15)
```

Out[22]:
```
Name of the Political Party
BHARATIYA JANATA PARTY    20.0
Name: Amount received, dtype: float64
```

In [ ]:
```

```