In [5]:
```
import requests
import pandas as pd
from bs4 import BeautifulSoup

url = 'https://results.eci.gov.in/'
response = requests.get(url)
soup = BeautifulSoup(response.content, 'html.parser')

# Extract specific data
data = soup.find_all('div', class_='data-class')
for item in data:
    print(item.text)
```

In [6]:
```
import requests
import pandas as pd
from bs4 import BeautifulSoup
url="https://results.eci.gov.in/"
r=requests.get(url)
```

In [10]:
```
soup=BeautifulSoup(r.text,"lxml")
table=soup.find("table",class_="body > main > div.container-fluid > section > div > div > div:nth-child(1)")

print(table)
```

```
None
```

In [12]:
```
import requests
import pandas as pd
from bs4 import BeautifulSoup
url="https://results.eci.gov.in/"
r=requests.get(url)

soup=BeautifulSoup(r.text,"lxml")
table=soup.find("table",class_="body > main > div.container-fluid > section > div > div > div:nth-child(1)")

print(table)



headers=table.find_all("th")
print(headers)
```

```
None
---------------------------------------------------------------------------
AttributeError                            Traceback (most recent call last)
Cell In[12], line 14
      8 table=soup.find("table",class_="body > main > div.container-fluid > section > div > div > div:nth-child
(1)")
     10 print(table)
---> 14 headers=table.find_all("th")
     15 print(headers)

AttributeError: 'NoneType' object has no attribute 'find_all'
```

In [13]:
```
pip install requests beautifulsoup4 pandas
```

```
Requirement already satisfied: requests in c:\users\mahadevan\anaconda3\lib\site-packages (2.28.1)
Requirement already satisfied: beautifulsoup4 in c:\users\mahadevan\anaconda3\lib\site-packages (4.11.1)
Requirement already satisfied: pandas in c:\users\mahadevan\anaconda3\lib\site-packages (1.5.3)
Requirement already satisfied: idna<4,>=2.5 in c:\users\mahadevan\anaconda3\lib\site-packages (from requests) (
3.4)
Requirement already satisfied: certifi>=2017.4.17 in c:\users\mahadevan\anaconda3\lib\site-packages (from reque
sts) (2022.12.7)
Requirement already satisfied: charset-normalizer<3,>=2 in c:\users\mahadevan\anaconda3\lib\site-packages (from
requests) (2.0.4)
Requirement already satisfied: urllib3<1.27,>=1.21.1 in c:\users\mahadevan\anaconda3\lib\site-packages (from re
quests) (1.26.14)
Requirement already satisfied: soupsieve>1.2 in c:\users\mahadevan\anaconda3\lib\site-packages (from beautifuls
oup4) (2.3.2.post1)
Requirement already satisfied: pytz>=2020.1 in c:\users\mahadevan\anaconda3\lib\site-packages (from pandas) (20
22.7)
Requirement already satisfied: python-dateutil>=2.8.1 in c:\users\mahadevan\anaconda3\lib\site-packages (from p
andas) (2.8.2)
Requirement already satisfied: numpy>=1.21.0 in c:\users\mahadevan\anaconda3\lib\site-packages (from pandas) (1
.23.5)
Requirement already satisfied: six>=1.5 in c:\users\mahadevan\anaconda3\lib\site-packages (from python-dateutil
>=2.8.1->pandas) (1.16.0)
Note: you may need to restart the kernel to use updated packages.
```

In [15]:
```
import requests
from bs4 import BeautifulSoup
import pandas as pd

url = 'https://results.eci.gov.in/'
response = requests.get(url)
soup = BeautifulSoup(response.content, 'html.parser')
```

```
# Update this to target the specific table
table = soup.find('table', {'class': 'table-class-name'})  # Replace with actual class or id

if table:
    # Extract table headers
    headers = [th.text.strip() for th in table.find_all('th')]

    # Extract table rows
    rows = []
    for tr in table.find_all('tr')[1:]:
        cells = tr.find_all('td')
        row = [cell.text.strip() for cell in cells]
        rows.append(row)

    # Create DataFrame
    df = pd.DataFrame(rows, columns=headers)
    print(df)
    df.to_csv('election_results.csv', index=False)
else:
    print("Table not found")
```

```
Table not found
```

In [18]:
```
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
import pandas as pd

try:
    df = pd.read_csv("assembly.csv")
    print(df.head())
except FileNotFoundError:
    print("File not found. Please check the file path.")
except pd.errors.EmptyDataError:
    print("The file is empty.")
except pd.errors.ParserError:
    print("Error parsing the file.")
except Exception as e:
    print(f"An error occurred: {e}")
```

```
An error occurred: 'utf-8' codec can't decode byte 0x93 in position 36679: invalid start byte
```

In [19]:
```
import pandas as pd

# Try different encodings
try:
    df = pd.read_csv("assembly.csv", encoding='latin1')  # or 'ISO-8859-1'
    print(df.head())
except FileNotFoundError:
    print("File not found. Please check the file path.")
except pd.errors.EmptyDataError:
    print("The file is empty.")
except pd.errors.ParserError:
    print("Error parsing the file.")
except Exception as e:
    print(f"An error occurred: {e}")
```

```
          ST_NAME  YEAR  AC_NO  #      AC_NAME AC_TYPE                 NAME  \
0  Andhra Pradesh  1955      1  1  ICHCHAPURAM     GEN      UPPADA RANGABABU
1  Andhra Pradesh  1955      1  2  ICHCHAPURAM     GEN      HARIHARA PATNAIK
2  Andhra Pradesh  1955      1  3  ICHCHAPURAM     GEN       PUDI LOKANADHAM
3  Andhra Pradesh  1955      1  4  ICHCHAPURAM     GEN  KALLA BALARAMA SWAMY
4  Andhra Pradesh  1955      1  5  ICHCHAPURAM     GEN    BANIKINKARA SHARMA

  SEX  AGE CATEGORY PARTY    VOTES
0   M  NaN      NaN   KLP  14565.0
1   M  NaN      NaN   IND   7408.0
2   M  NaN      NaN   IND   6508.0
3   M  NaN      NaN   IND   3002.0
4   M  NaN      NaN   IND    682.0
C:\Users\Mahadevan\AppData\Local\Temp\ipykernel_12408\752085104.py:5: DtypeWarning: Columns (1,5,7,9) have mixe
d types. Specify dtype option on import or set low_memory=False.
  df = pd.read_csv("assembly.csv", encoding='latin1')  # or 'ISO-8859-1'
```

In [20]:
```
df
```

Out[20]:

	ST_NAME	YEAR	AC_NO	#	AC_NAME	AC_TYPE	NAME	SEX	AGE	CATEGORY	PARTY	VOTES
0	Andhra Pradesh	1955	1	1	ICHCHAPURAM	GEN	UPPADA RANGABABU	M	NaN	NaN	KLP	14565.0
1	Andhra Pradesh	1955	1	2	ICHCHAPURAM	GEN	HARIHARA PATNAIK	M	NaN	NaN	IND	7408.0
2	Andhra Pradesh	1955	1	3	ICHCHAPURAM	GEN	PUDI LOKANADHAM	M	NaN	NaN	IND	6508.0
3	Andhra Pradesh	1955	1	4	ICHCHAPURAM	GEN	KALLA BALARAMA SWAMY	M	NaN	NaN	IND	3002.0
4	Andhra Pradesh	1955	1	5	ICHCHAPURAM	GEN	BANIKINKARA SHARMA	M	NaN	NaN	IND	682.0
...	...	...	...	...	...	...	...	...	...	...	...	...
398697	Rajasthan	2013	75	12	WEIR	NaN	BALSWAROOP	NaN	NaN	NaN	RVP	178.0
398698	Rajasthan	2013	75	13	WEIR	NaN	PUSHKAR JATAV	NaN	NaN	NaN	IPGP	152.0
398699	Rajasthan	2013	75	14	WEIR	NaN	KAMLESH	NaN	NaN	NaN	BBP	142.0
398700	Rajasthan	2013	75	15	WEIR	NaN	UDAY SINGH	NaN	NaN	NaN	IVD	120.0
398701	Rajasthan	2013	75	5	WEIR	NaN	NONE OF THE ABOVE	NaN	NaN	NaN	NOTA	1472.0

398702 rows × 12 columns

In [21]:
```
df.columns
```

Out[21]:
```
Index(['ST_NAME', 'YEAR', 'AC_NO', '#', 'AC_NAME', 'AC_TYPE', 'NAME', 'SEX',
       'AGE', 'CATEGORY', 'PARTY', 'VOTES'],
      dtype='object')
```

In [23]:
```
df.dtypes
```

Out[23]:
```
ST_NAME      object
YEAR         object
AC_NO         int64
#             int64
AC_NAME      object
AC_TYPE      object
NAME         object
SEX          object
AGE         float64
CATEGORY     object
PARTY        object
VOTES       float64
dtype: object
```

In [24]:
```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 398702 entries, 0 to 398701
Data columns (total 12 columns):
 #   Column    Non-Null Count   Dtype
---  ------    --------------   -----
 0   ST_NAME   398702 non-null  object
 1   YEAR      398702 non-null  object
 2   AC_NO     398702 non-null  int64
 3   #         398702 non-null  int64
 4   AC_NAME   398702 non-null  object
 5   AC_TYPE   391465 non-null  object
 6   NAME      398702 non-null  object
 7   SEX       391465 non-null  object
 8   AGE       78578 non-null   float64
 9   CATEGORY  78579 non-null   object
 10  PARTY     398702 non-null  object
 11  VOTES     398452 non-null  float64
dtypes: float64(2), int64(2), object(8)
memory usage: 36.5+ MB
```

In [25]:
```
df.describe()
```

Out[25]:

	AC_NO	#	AGE	VOTES
count	398702.000000	398702.000000	78578.000000	398452.000000
mean	126.174494	8.008864	44.342780	8881.917478
std	97.603249	28.950246	11.106754	15293.814044
min	1.000000	1.000000	21.000000	0.000000
25%	46.000000	2.000000	36.000000	360.000000
50%	103.000000	5.000000	43.000000	1447.000000
75%	191.000000	9.000000	52.000000	10604.000000
max	430.000000	946.000000	93.000000	584098.000000

In [26]:
```
df.fillna("0", inplace=True)
```

In [27]:
```
df.drop_duplicates(inplace=True)
df
```

Out[27]:

	ST_NAME	YEAR	AC_NO	#	AC_NAME	AC_TYPE	NAME	SEX	AGE	CATEGORY	PARTY	VOTES
0	Andhra Pradesh	1955	1	1	ICHCHAPURAM	GEN	UPPADA RANGABABU	M	0	0	KLP	14565.0
1	Andhra Pradesh	1955	1	2	ICHCHAPURAM	GEN	HARIHARA PATNAIK	M	0	0	IND	7408.0
2	Andhra Pradesh	1955	1	3	ICHCHAPURAM	GEN	PUDI LOKANADHAM	M	0	0	IND	6508.0
3	Andhra Pradesh	1955	1	4	ICHCHAPURAM	GEN	KALLA BALARAMA SWAMY	M	0	0	IND	3002.0
4	Andhra Pradesh	1955	1	5	ICHCHAPURAM	GEN	BANIKINKARA SHARMA	M	0	0	IND	682.0
...	...	...	...	...	...	...	...	...	...	...	...	...
398697	Rajasthan	2013	75	12	WEIR	0	BALSWAROOP	0	0	0	RVP	178.0
398698	Rajasthan	2013	75	13	WEIR	0	PUSHKAR JATAV	0	0	0	IPGP	152.0
398699	Rajasthan	2013	75	14	WEIR	0	KAMLESH	0	0	0	BBP	142.0
398700	Rajasthan	2013	75	15	WEIR	0	UDAY SINGH	0	0	0	IVD	120.0
398701	Rajasthan	2013	75	5	WEIR	0	NONE OF THE ABOVE	0	0	0	NOTA	1472.0

398652 rows × 12 columns

In [28]:
```
df
```

Out[28]:

	ST_NAME	YEAR	AC_NO	#	AC_NAME	AC_TYPE	NAME	SEX	AGE	CATEGORY	PARTY	VOTES
0	Andhra Pradesh	1955	1	1	ICHCHAPURAM	GEN	UPPADA RANGABABU	M	0	0	KLP	14565.0
1	Andhra Pradesh	1955	1	2	ICHCHAPURAM	GEN	HARIHARA PATNAIK	M	0	0	IND	7408.0
2	Andhra Pradesh	1955	1	3	ICHCHAPURAM	GEN	PUDI LOKANADHAM	M	0	0	IND	6508.0
3	Andhra Pradesh	1955	1	4	ICHCHAPURAM	GEN	KALLA BALARAMA SWAMY	M	0	0	IND	3002.0
4	Andhra Pradesh	1955	1	5	ICHCHAPURAM	GEN	BANIKINKARA SHARMA	M	0	0	IND	682.0
...	...	...	...	...	...	...	...	...	...	...	...	...
398697	Rajasthan	2013	75	12	WEIR	0	BALSWAROOP	0	0	0	RVP	178.0
398698	Rajasthan	2013	75	13	WEIR	0	PUSHKAR JATAV	0	0	0	IPGP	152.0
398699	Rajasthan	2013	75	14	WEIR	0	KAMLESH	0	0	0	BBP	142.0
398700	Rajasthan	2013	75	15	WEIR	0	UDAY SINGH	0	0	0	IVD	120.0
398701	Rajasthan	2013	75	5	WEIR	0	NONE OF THE ABOVE	0	0	0	NOTA	1472.0

398652 rows × 12 columns

In [30]:
```
df.PARTY
```

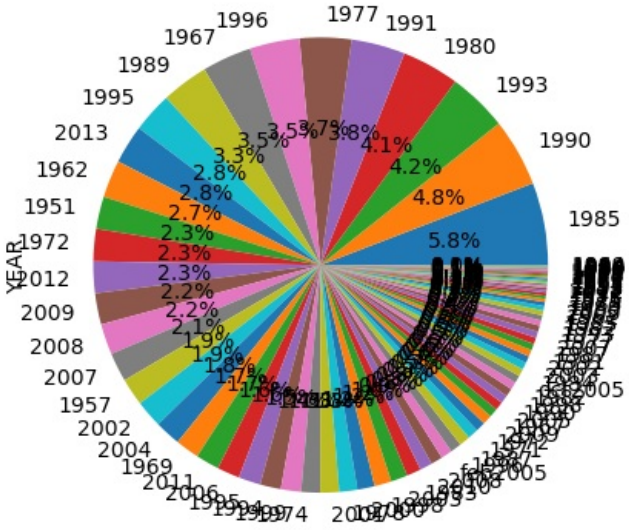
Out[30]:
```
0          KLP
1          IND
2          IND
3          IND
4          IND
          ...
398697     RVP
398698    IPGP
398699     BBP
398700     IVD
398701    NOTA
Name: PARTY, Length: 398652, dtype: object
```

In [31]:
```
df.PARTY.unique()
```

Out[31]:
```
array(['KLP', 'IND', 'CPI', ..., 'BSuD', 'TP', 'RBD'], dtype=object)
```

In [39]:
```
df["YEAR"].value_counts().plot.pie(autopct="%1.1f%%")
```

Out[39]:
```
<Axes: ylabel='YEAR'>
```



In [41]:
```
df.info
```

Out[41]:
```
<bound method DataFrame.info of                ST_NAME  YEAR  AC_NO   #      AC_NAME AC_TYPE  \
0       Andhra Pradesh  1955      1   1  ICHCHAPURAM     GEN
1       Andhra Pradesh  1955      1   2  ICHCHAPURAM     GEN
2       Andhra Pradesh  1955      1   3  ICHCHAPURAM     GEN
3       Andhra Pradesh  1955      1   4  ICHCHAPURAM     GEN
4       Andhra Pradesh  1955      1   5  ICHCHAPURAM     GEN
...                ...   ...    ...  ..          ...     ...
398697       Rajasthan  2013     75  12         WEIR       0
398698       Rajasthan  2013     75  13         WEIR       0
398699       Rajasthan  2013     75  14         WEIR       0
398700       Rajasthan  2013     75  15         WEIR       0
398701       Rajasthan  2013     75   5         WEIR       0

                     NAME SEX AGE CATEGORY PARTY    VOTES
0           UPPADA RANGABABU   M   0        0   KLP  14565.0
1           HARIHARA PATNAIK   M   0        0   IND   7408.0
2            PUDI LOKANADHAM   M   0        0   IND   6508.0
3       KALLA BALARAMA SWAMY   M   0        0   IND   3002.0
4         BANIKINKARA SHARMA   M   0        0   IND    682.0
...                      ...  ..  ..      ...   ...      ...
398697            BALSWAROOP   0   0        0   RVP    178.0
398698         PUSHKAR JATAV   0   0        0  IPGP    152.0
398699               KAMLESH   0   0        0   BBP    142.0
398700            UDAY SINGH   0   0        0   IVD    120.0
398701     NONE OF THE ABOVE   0   0        0  NOTA   1472.0

[398652 rows x 12 columns]>
```

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]: