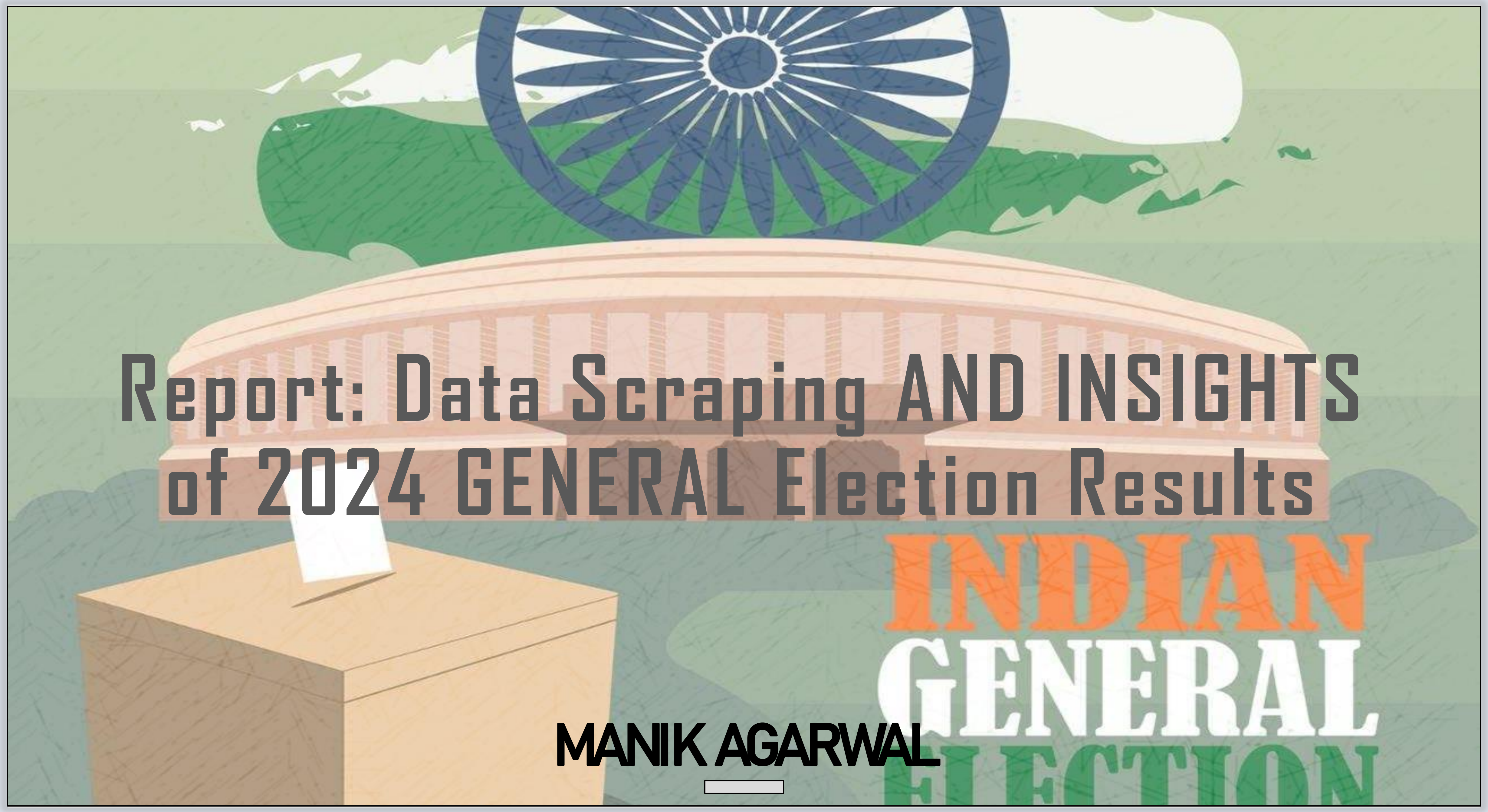
Report: Data Scraping AND INSIGHTS of 2024 GENERAL Election Results



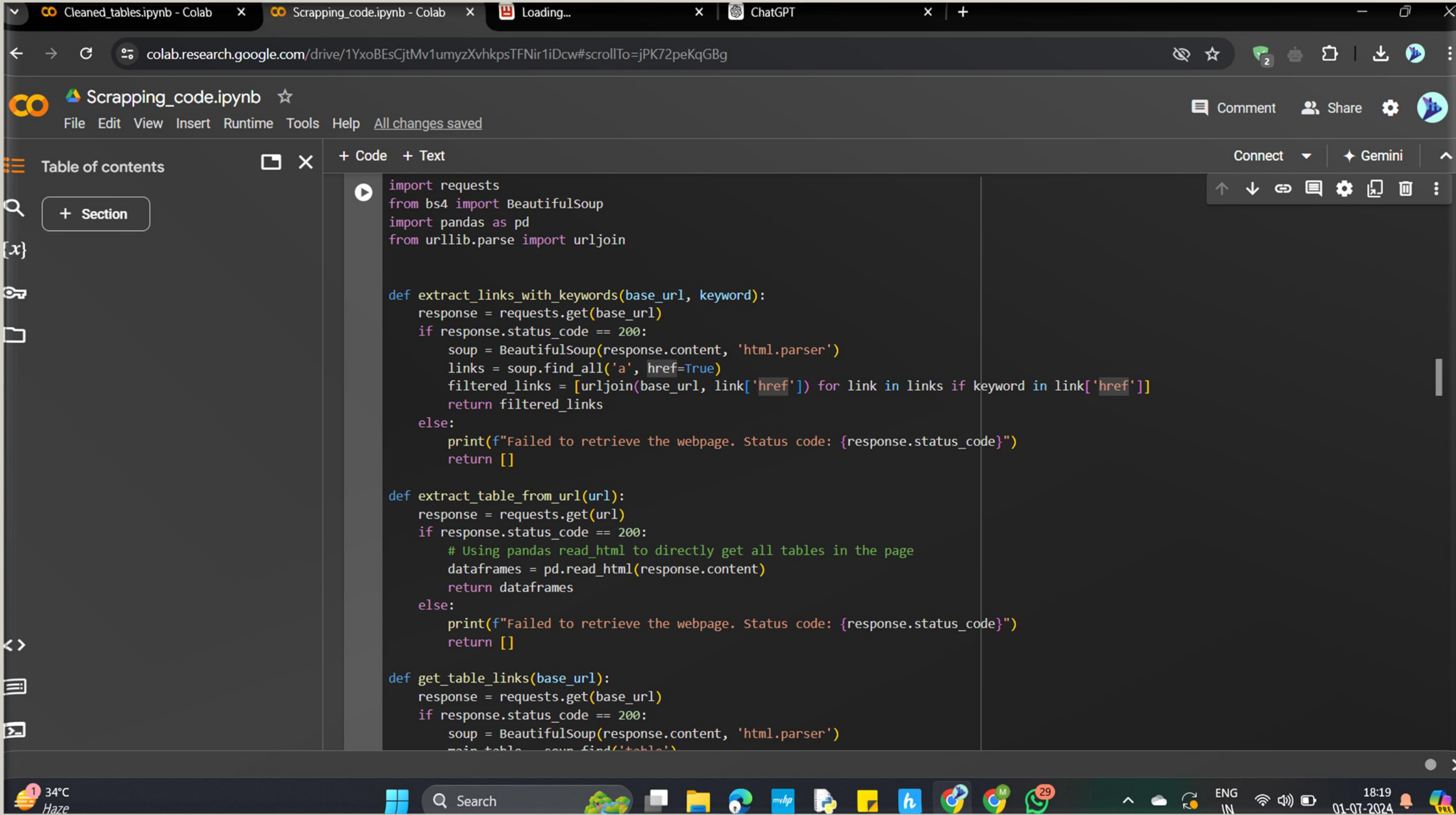
MANIK AGARWAL

Technical Details:

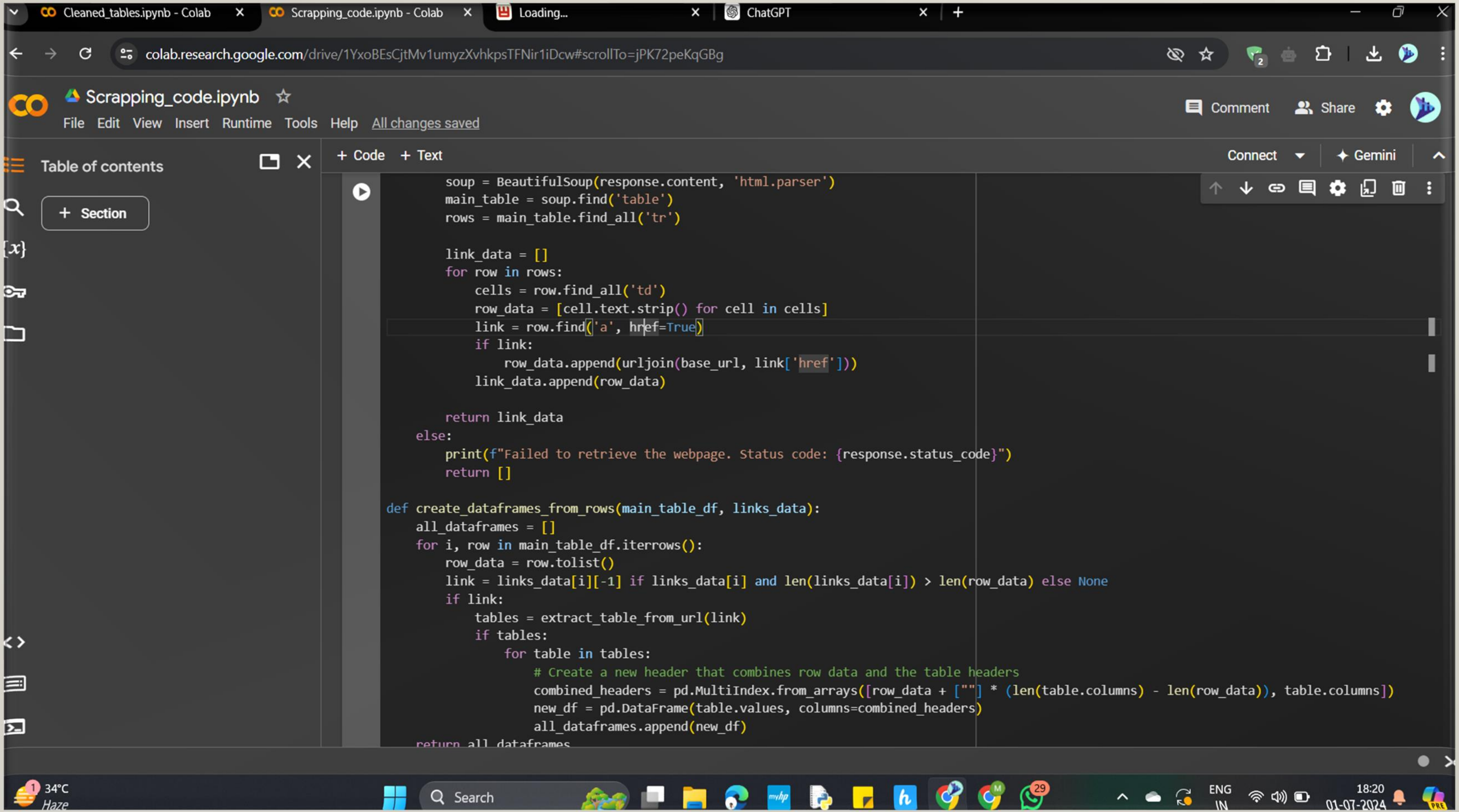
Utilized Python libraries, BeautifulSoup and Requests, to scrape data from the official Election Commission of India website (https://results.eci.gov.in), which contains the election results for India in 2024.
The aim was to extract comprehensive details such as seats won, seats lost, total seats, leading seats, state-wise results, and member-wise results.
Using the Requests library, I sent HTTP requests to fetch the HTML content of the webpage. BeautifulSoup was then employed to parse this HTML content and extract relevant data. The data extraction process involved identifying specific HTML tags and classes that contained the necessary information about the election results.
I focused on several key elements:

- **Seats Win/Lose/Total**: Extracted the number of seats each party won, lost, and their total tally.
- **Leading Seats**: Identified the number of seats where each party was leading.
- **State-wise Results**: Gathered data on the election results from each state.
- **Member-wise Results**: Compiled detailed results for individual members, including their winning or losing status.

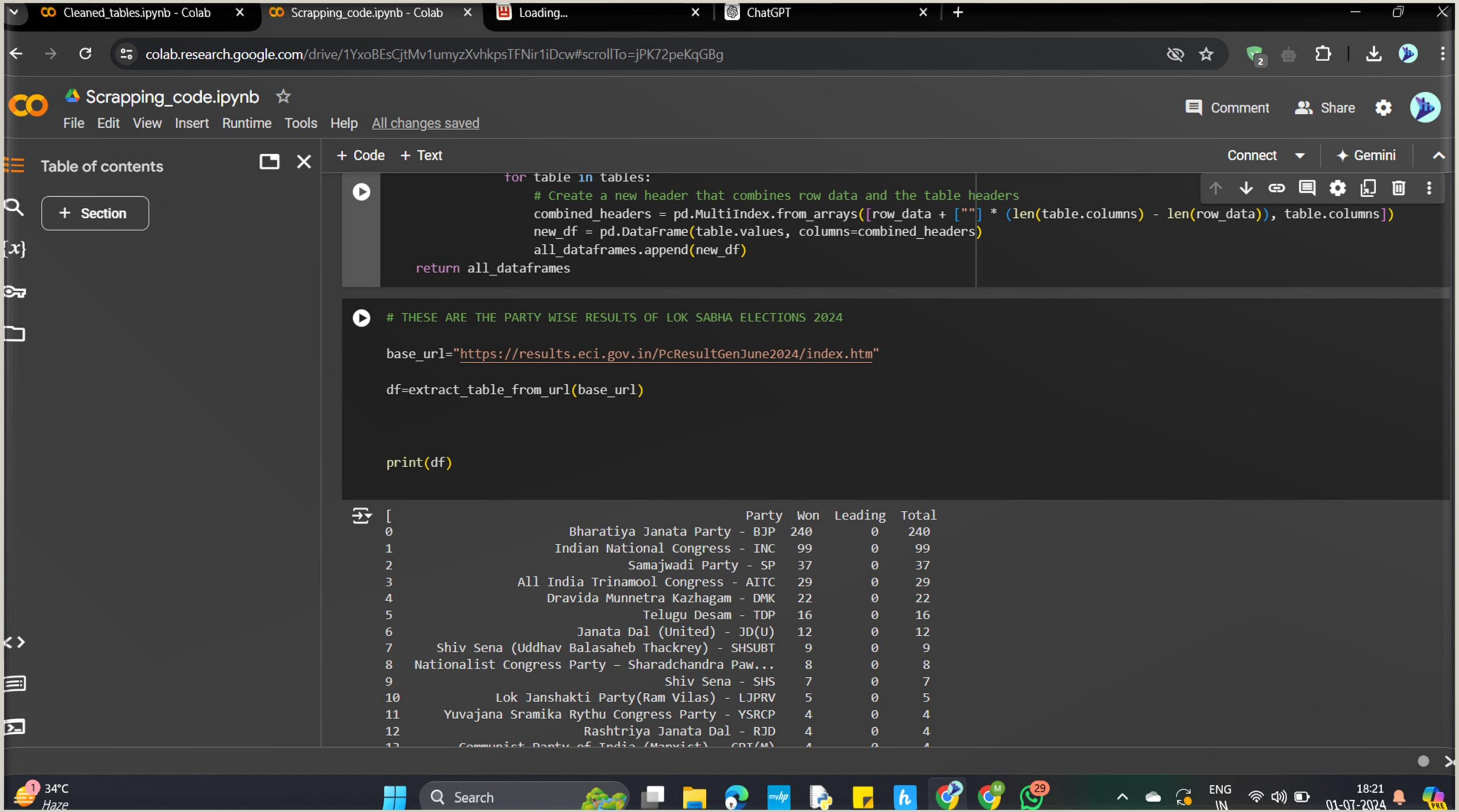
The scraped data was then organized into a structured format, enabling further analysis and visualization. This approach ensured accurate and up-to-date information was collected efficiently, providing valuable insights into the 2024 election outcomes.


Cleaned_tables.ipynb - Colab
Scrapping_code.ipynb - Colab
Loading...
ChatGPT
colab.research.google.com/drive/1YxoBEsCjtMv1umyzXvhkpsTFNir1iDcw#scrollTo=jPK72peKqGBg
Scrapping_code.ipynb
File Edit View Insert Runtime Tools Help All changes saved
Comment
Share
Table of contents
+ Section
+ Code + Text
Connect
Gemini
34°C
Haze
Search
ENG
IN
18:19
01-07-2024

```
import requests
from bs4 import BeautifulSoup
import pandas as pd
from urllib.parse import urljoin


def extract_links_with_keywords(base_url, keyword):
    response = requests.get(base_url)
    if response.status_code == 200:
        soup = BeautifulSoup(response.content, 'html.parser')
        links = soup.find_all('a', href=True)
        filtered_links = [urljoin(base_url, link['href']) for link in links if keyword in link['href']]
        return filtered_links
    else:
        print(f"Failed to retrieve the webpage. Status code: {response.status_code}")
        return []

def extract_table_from_url(url):
    response = requests.get(url)
    if response.status_code == 200:
        # Using pandas read_html to directly get all tables in the page
        dataframes = pd.read_html(response.content)
        return dataframes
    else:
        print(f"Failed to retrieve the webpage. Status code: {response.status_code}")
        return []

def get_table_links(base_url):
    response = requests.get(base_url)
    if response.status_code == 200:
        soup = BeautifulSoup(response.content, 'html.parser')
        main_table = soup.find('table')
```


Cleaned_tables.ipynb - Colab
Scrapping_code.ipynb - Colab
Loading...
ChatGPT
colab.research.google.com/drive/1YxoBEsCjtMv1umyzXvhkpsTFNir1iDcw#scrollTo=jPK72peKqGBg
Scrapping_code.ipynb
File Edit View Insert Runtime Tools Help All changes saved
Comment
Share
Table of contents
Section
+ Code + Text
Connect
Gemini
34°C
Haze
Search
ENG
IN
18:20
01-07-2024

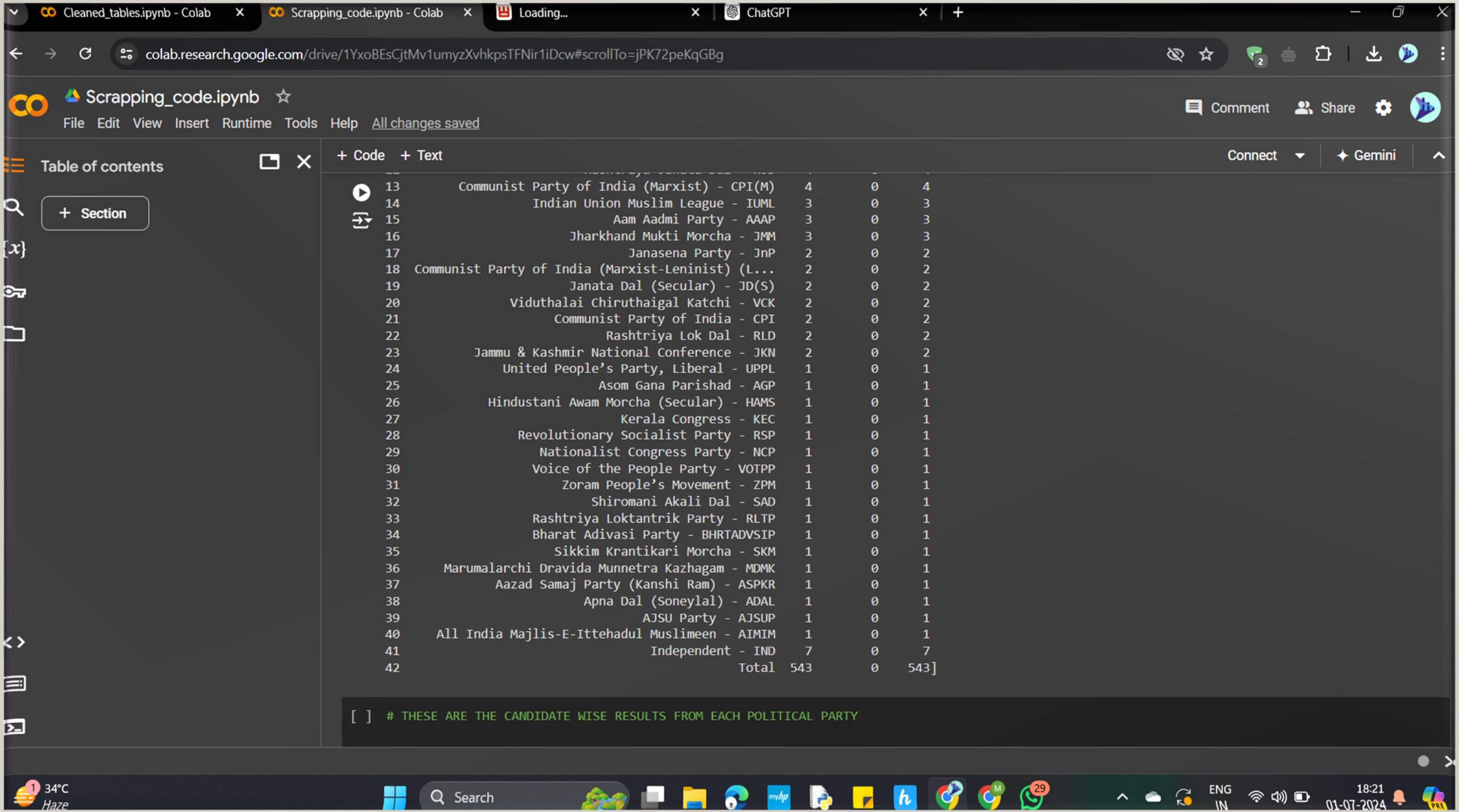
```
        soup = BeautifulSoup(response.content, 'html.parser')
        main_table = soup.find('table')
        rows = main_table.find_all('tr')

        link_data = []
        for row in rows:
            cells = row.find_all('td')
            row_data = [cell.text.strip() for cell in cells]
            link = row.find('a', href=True)
            if link:
                row_data.append(urljoin(base_url, link['href']))
            link_data.append(row_data)

        return link_data
    else:
        print(f"Failed to retrieve the webpage. Status code: {response.status_code}")
        return []

def create_dataframes_from_rows(main_table_df, links_data):
    all_dataframes = []
    for i, row in main_table_df.iterrows():
        row_data = row.tolist()
        link = links_data[i][-1] if links_data[i] and len(links_data[i]) > len(row_data) else None
        if link:
            tables = extract_table_from_url(link)
            if tables:
                for table in tables:
                    # Create a new header that combines row data and the table headers
                    combined_headers = pd.MultiIndex.from_arrays([row_data + [""] * (len(table.columns) - len(row_data)), table.columns])
                    new_df = pd.DataFrame(table.values, columns=combined_headers)
                    all_dataframes.append(new_df)
    return all_dataframes
```



```
                for table in tables:
                    # Create a new header that combines row data and the table headers
                    combined_headers = pd.MultiIndex.from_arrays([row_data + [""] * (len(table.columns) - len(row_data)), table.columns])
                    new_df = pd.DataFrame(table.values, columns=combined_headers)
                    all_dataframes.append(new_df)
    return all_dataframes
```

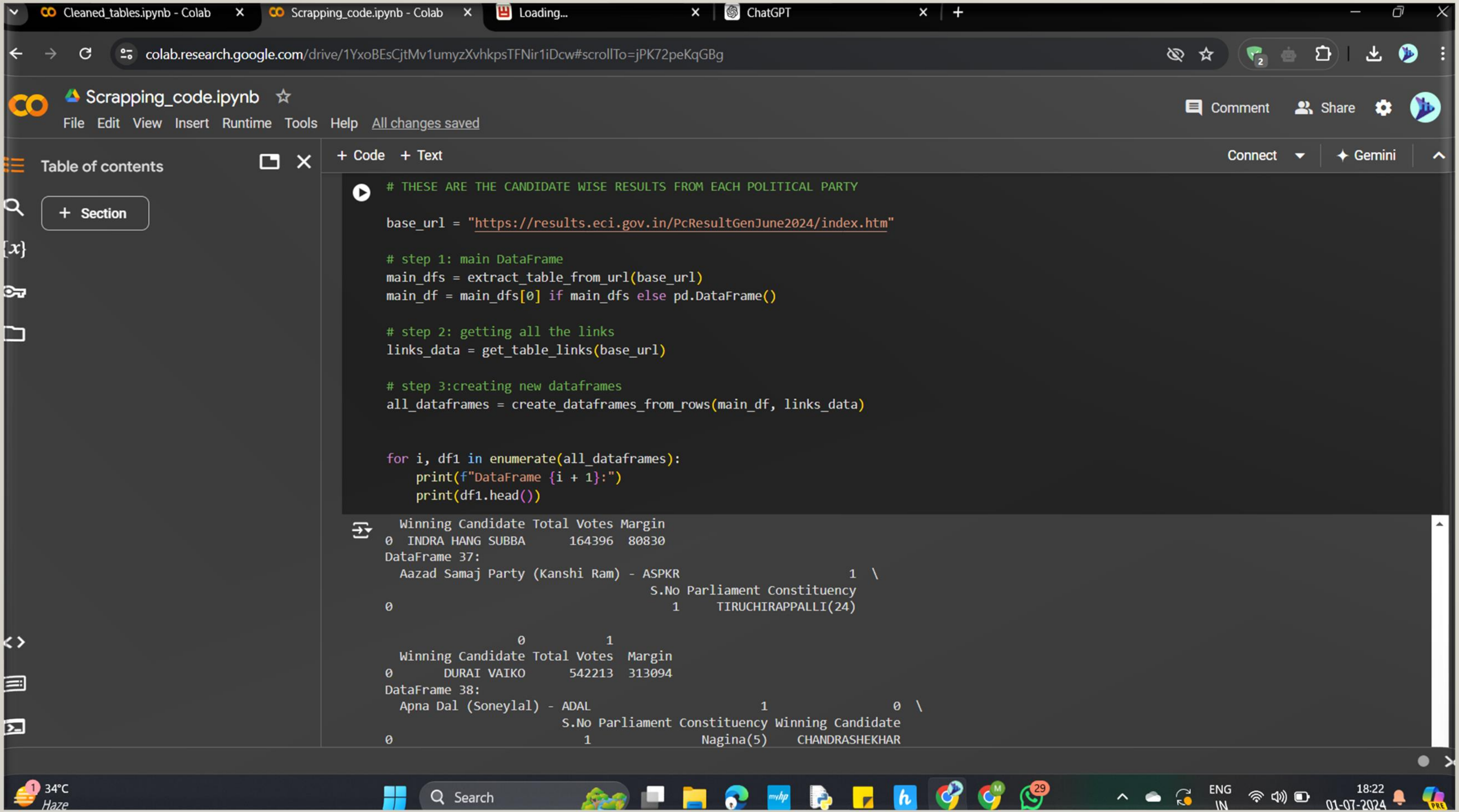
```
# THESE ARE THE PARTY WISE RESULTS OF LOK SABHA ELECTIONS 2024

base_url="https://results.eci.gov.in/PcResultGenJune2024/index.htm"

df=extract_table_from_url(base_url)



print(df)
```

```
[                                              Party  Won  Leading  Total
0                        Bharatiya Janata Party - BJP  240        0    240
1                      Indian National Congress - INC   99        0     99
2                                Samajwadi Party - SP   37        0     37
3                 All India Trinamool Congress - AITC   29        0     29
4                     Dravida Munnetra Kazhagam - DMK   22        0     22
5                                 Telugu Desam - TDP   16        0     16
6                         Janata Dal (United) - JD(U)   12        0     12
7         Shiv Sena (Uddhav Balasaheb Thackrey) - SHSUBT    9        0      9
8    Nationalist Congress Party – Sharadchandra Paw...    8        0      8
9                                   Shiv Sena - SHS    7        0      7
10                 Lok Janshakti Party(Ram Vilas) - LJPRV    5        0      5
11          Yuvajana Sramika Rythu Congress Party - YSRCP    4        0      4
12                          Rashtriya Janata Dal - RJD    4        0      4
13             Communist Party of India (Marxist) - CPI(M)    4        0      4
```



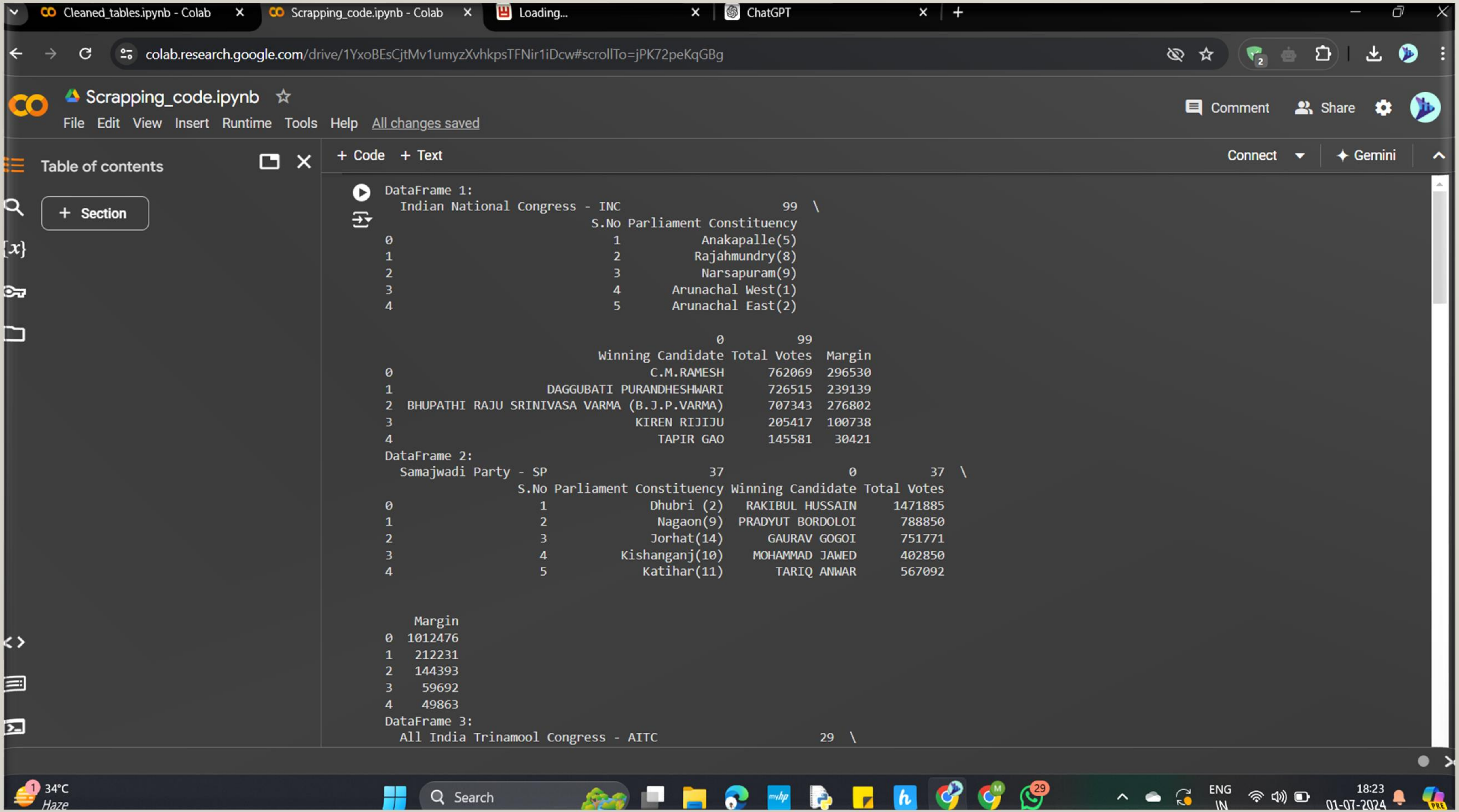
```
13          Communist Party of India (Marxist) - CPI(M)    4      0      4
14                Indian Union Muslim League - IUML    3      0      3
15                        Aam Aadmi Party - AAAP    3      0      3
16                   Jharkhand Mukti Morcha - JMM    3      0      3
17                         Janasena Party - JnP    2      0      2
18  Communist Party of India (Marxist-Leninist) (L...    2      0      2
19                   Janata Dal (Secular) - JD(S)    2      0      2
20             Viduthalai Chiruthaigal Katchi - VCK    2      0      2
21                 Communist Party of India - CPI    2      0      2
22                     Rashtriya Lok Dal - RLD    2      0      2
23         Jammu & Kashmir National Conference - JKN    2      0      2
24             United People’s Party, Liberal - UPPL    1      0      1
25                      Asom Gana Parishad - AGP    1      0      1
26          Hindustani Awam Morcha (Secular) - HAMS    1      0      1
27                        Kerala Congress - KEC    1      0      1
28              Revolutionary Socialist Party - RSP    1      0      1
29                 Nationalist Congress Party - NCP    1      0      1
30               Voice of the People Party - VOTPP    1      0      1
31                  Zoram People’s Movement - ZPM    1      0      1
32                      Shiromani Akali Dal - SAD    1      0      1
33               Rashtriya Loktantrik Party - RLTP    1      0      1
34             Bharat Adivasi Party - BHRTADVSIP    1      0      1
35                  Sikkim Krantikari Morcha - SKM    1      0      1
36     Marumalarchi Dravida Munnetra Kazhagam - MDMK    1      0      1
37            Aazad Samaj Party (Kanshi Ram) - ASPKR    1      0      1
38                     Apna Dal (Soneylal) - ADAL    1      0      1
39                          AJSU Party - AJSUP    1      0      1
40      All India Majlis-E-Ittehadul Muslimeen - AIMIM    1      0      1
41                          Independent - IND    7      0      7
42                                      Total  543      0    543]
```

```
[ ]  # THESE ARE THE CANDIDATE WISE RESULTS FROM EACH POLITICAL PARTY
```

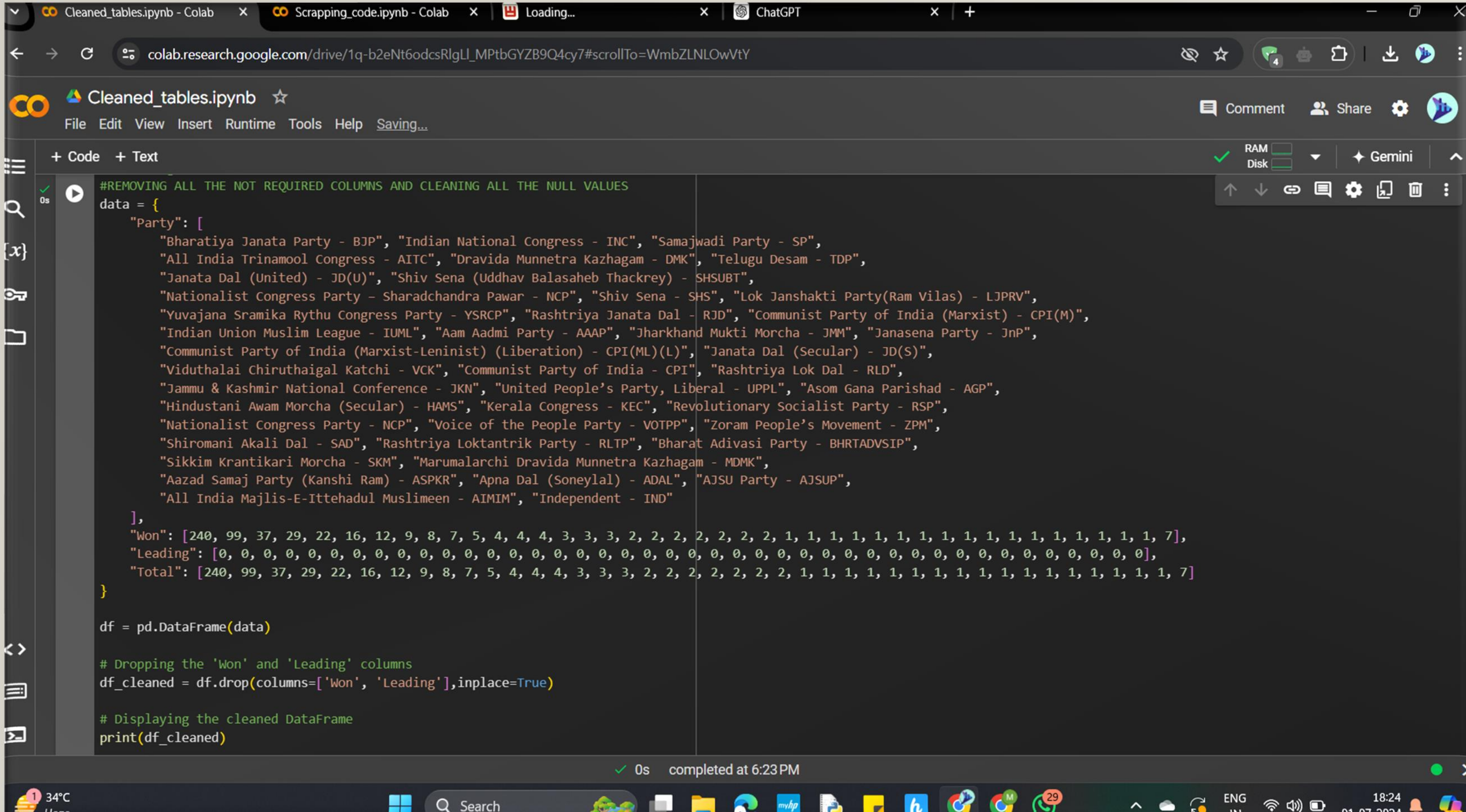

Cleaned_tables.ipynb - Colab
Scrapping_code.ipynb - Colab
Loading...
ChatGPT
colab.research.google.com/drive/1YxoBEsCjtMv1umyzXvhkpsTFNir1iDcw#scrollTo=jPK72peKqGBg
Scrapping_code.ipynb
File Edit View Insert Runtime Tools Help All changes saved
Comment
Share
Table of contents
+ Section
+ Code + Text
Connect
Gemini
34°C
Haze
Search
ENG
IN
18:22
01-07-2024

```python
# THESE ARE THE CANDIDATE WISE RESULTS FROM EACH POLITICAL PARTY

base_url = "https://results.eci.gov.in/PcResultGenJune2024/index.htm"

# step 1: main DataFrame
main_dfs = extract_table_from_url(base_url)
main_df = main_dfs[0] if main_dfs else pd.DataFrame()

# step 2: getting all the links
links_data = get_table_links(base_url)

# step 3:creating new dataframes
all_dataframes = create_dataframes_from_rows(main_df, links_data)


for i, df1 in enumerate(all_dataframes):
    print(f"DataFrame {i + 1}:")
    print(df1.head())
```

```
  Winning Candidate Total Votes Margin
0  INDRA HANG SUBBA      164396  80830
DataFrame 37:
  Aazad Samaj Party (Kanshi Ram) - ASPKR                       1  \
                                    S.No Parliament Constituency
0                                      1      TIRUCHIRAPPALLI(24)


                  0           1
  Winning Candidate Total Votes  Margin
0       DURAI VAIKO      542213  313094
DataFrame 38:
  Apna Dal (Soneylal) - ADAL                       1                 0  \
                        S.No Parliament Constituency Winning Candidate
0                          1               Nagina(5)    CHANDRASHEKHAR
```



```
DataFrame 1:
  Indian National Congress - INC                      99  \
                            S.No Parliament Constituency
0                              1           Anakapalle(5)
1                              2          Rajahmundry(8)
2                              3           Narsapuram(9)
3                              4       Arunachal West(1)
4                              5       Arunachal East(2)

                                             0          99
                             Winning Candidate Total Votes  Margin
0                                   C.M.RAMESH      762069  296530
1                     DAGGUBATI PURANDHESHWARI      726515  239139
2  BHUPATHI RAJU SRINIVASA VARMA (B.J.P.VARMA)      707343  276802
3                                 KIREN RIJIJU      205417  100738
4                                    TAPIR GAO      145581   30421
DataFrame 2:
  Samajwadi Party - SP                      37                 0          37  \
                  S.No Parliament Constituency Winning Candidate Total Votes
0                    1              Dhubri (2)   RAKIBUL HUSSAIN     1471885
1                    2               Nagaon(9)  PRADYUT BORDOLOI      788850
2                    3              Jorhat(14)      GAURAV GOGOI      751771
3                    4          Kishanganj(10)    MOHAMMAD JAWED      402850
4                    5             Katihar(11)       TARIQ ANWAR      567092


    Margin
0  1012476
1   212231
2   144393
3    59692
4    49863
DataFrame 3:
  All India Trinamool Congress - AITC                      29  \
```



Cleaned_tables.ipynb - Colab
Scrapping_code.ipynb - Colab
Loading...
ChatGPT
colab.research.google.com/drive/1q-b2eNt6odcsRlgLl_MPtbGYZB9Q4cy7#scrollTo=WmbZLNLOwVtY
Cleaned_tables.ipynb
File Edit View Insert Runtime Tools Help Saving...
Comment
Share
+ Code + Text
RAM
Disk
Gemini
0s

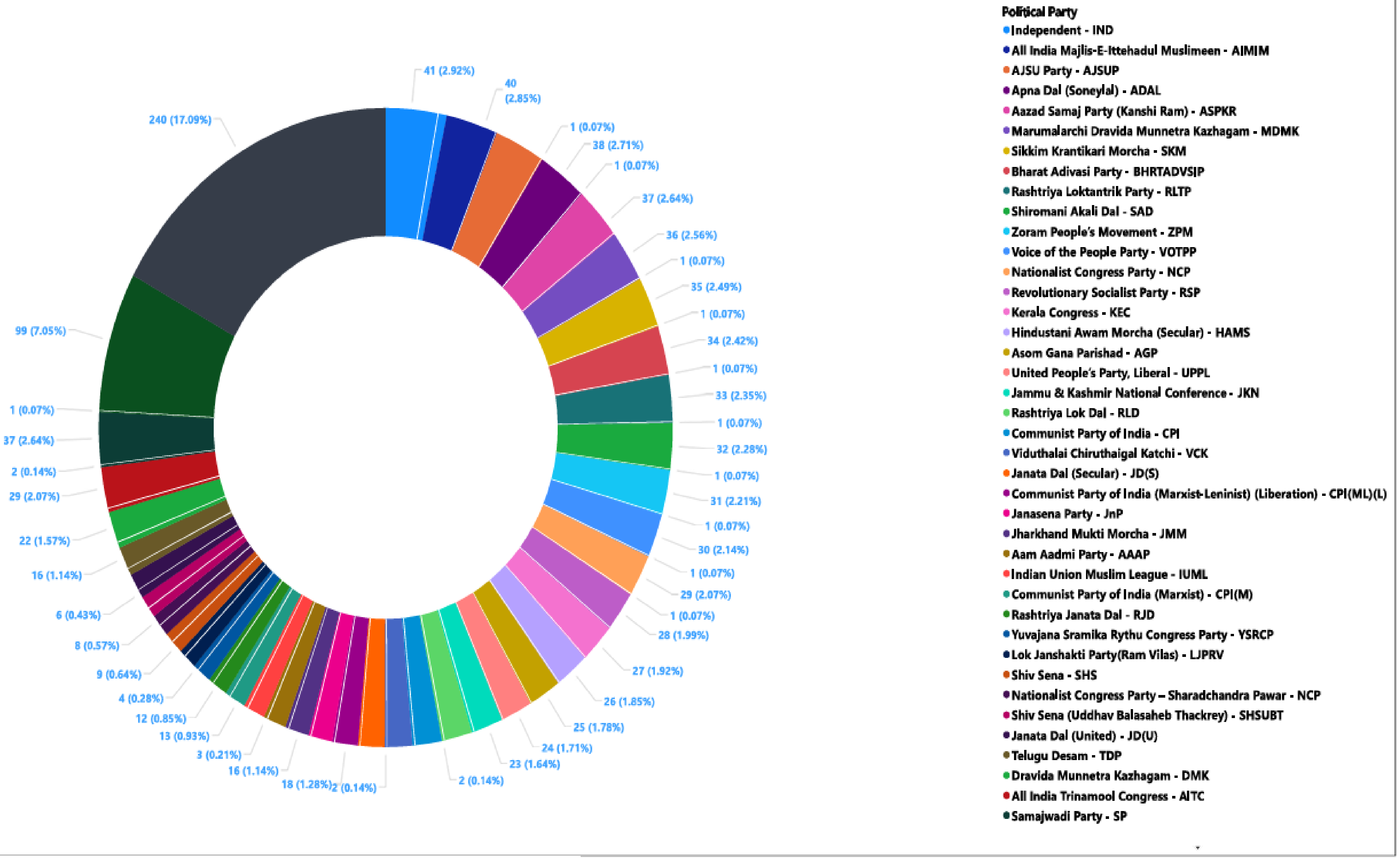
```python
#REMOVING ALL THE NOT REQUIRED COLUMNS AND CLEANING ALL THE NULL VALUES
data = {
    "Party": [
        "Bharatiya Janata Party - BJP", "Indian National Congress - INC", "Samajwadi Party - SP",
        "All India Trinamool Congress - AITC", "Dravida Munnetra Kazhagam - DMK", "Telugu Desam - TDP",
        "Janata Dal (United) - JD(U)", "Shiv Sena (Uddhav Balasaheb Thackrey) - SHSUBT",
        "Nationalist Congress Party – Sharadchandra Pawar - NCP", "Shiv Sena - SHS", "Lok Janshakti Party(Ram Vilas) - LJPRV",
        "Yuvajana Sramika Rythu Congress Party - YSRCP", "Rashtriya Janata Dal - RJD", "Communist Party of India (Marxist) - CPI(M)",
        "Indian Union Muslim League - IUML", "Aam Aadmi Party - AAAP", "Jharkhand Mukti Morcha - JMM", "Janasena Party - JnP",
        "Communist Party of India (Marxist-Leninist) (Liberation) - CPI(ML)(L)", "Janata Dal (Secular) - JD(S)",
        "Viduthalai Chiruthaigal Katchi - VCK", "Communist Party of India - CPI", "Rashtriya Lok Dal - RLD",
        "Jammu & Kashmir National Conference - JKN", "United People’s Party, Liberal - UPPL", "Asom Gana Parishad - AGP",
        "Hindustani Awam Morcha (Secular) - HAMS", "Kerala Congress - KEC", "Revolutionary Socialist Party - RSP",
        "Nationalist Congress Party - NCP", "Voice of the People Party - VOTPP", "Zoram People’s Movement - ZPM",
        "Shiromani Akali Dal - SAD", "Rashtriya Loktantrik Party - RLTP", "Bharat Adivasi Party - BHRTADVSIP",
        "Sikkim Krantikari Morcha - SKM", "Marumalarchi Dravida Munnetra Kazhagam - MDMK",
        "Aazad Samaj Party (Kanshi Ram) - ASPKR", "Apna Dal (Soneylal) - ADAL", "AJSU Party - AJSUP",
        "All India Majlis-E-Ittehadul Muslimeen - AIMIM", "Independent - IND"
    ],
    "Won": [240, 99, 37, 29, 22, 16, 12, 9, 8, 7, 5, 4, 4, 4, 3, 3, 3, 2, 2, 2, 2, 2, 2, 2, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 7],
    "Leading": [0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0],
    "Total": [240, 99, 37, 29, 22, 16, 12, 9, 8, 7, 5, 4, 4, 4, 3, 3, 3, 2, 2, 2, 2, 2, 2, 2, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 7]
}

df = pd.DataFrame(data)

# Dropping the 'Won' and 'Leading' columns
df_cleaned = df.drop(columns=['Won', 'Leading'],inplace=True)

# Displaying the cleaned DataFrame
print(df_cleaned)
```

✓ 0s completed at 6:23 PM

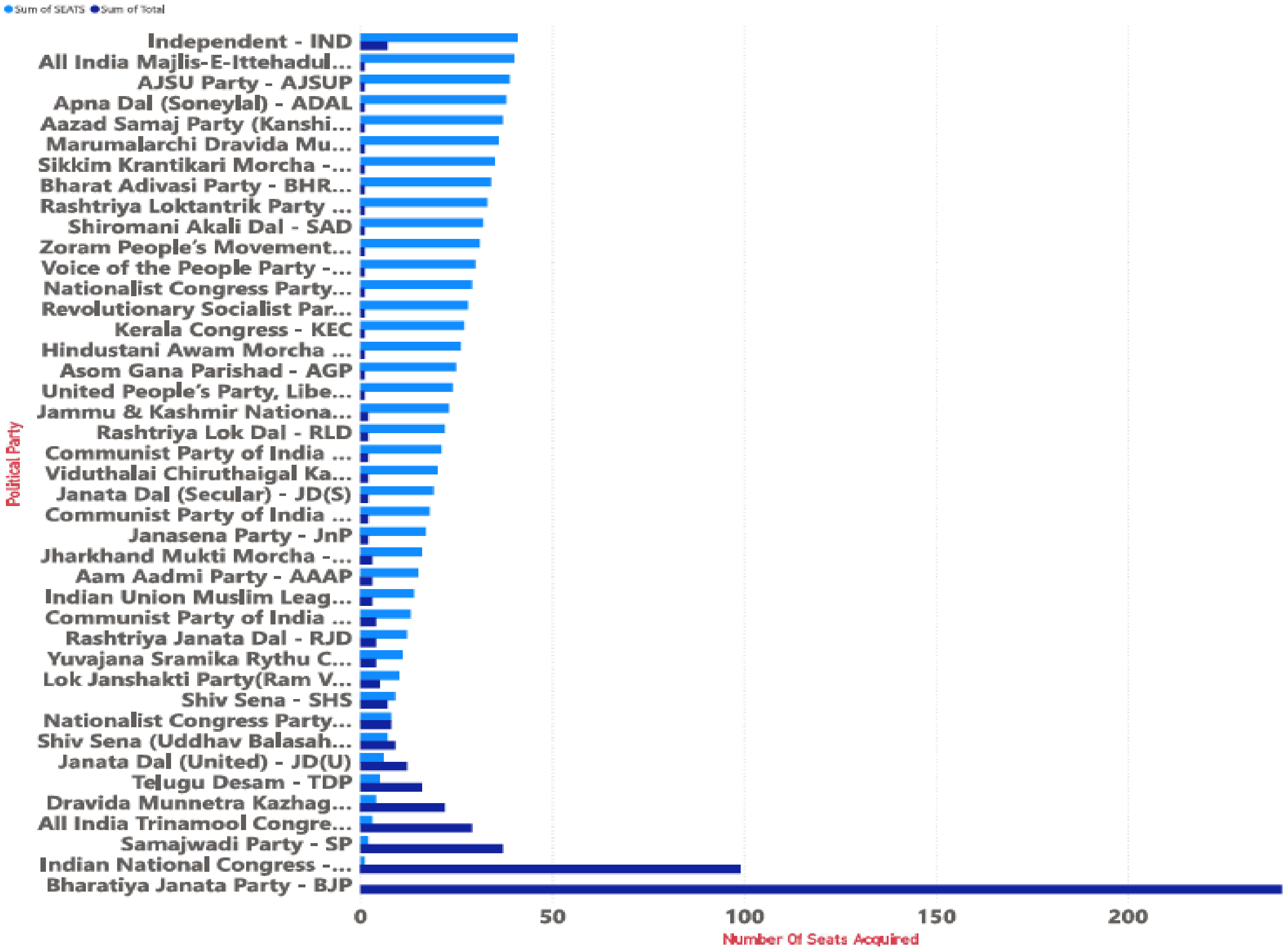
Report: Insights Generation on India 2024 Election Results Using Power BI

- For analyzing the India 2024 election results, I scraped data from the Election Commission of India's official website (https://results.eci.gov.in) using Python's BeautifulSoup and Requests libraries. The goal was to extract detailed information on seats won, lost, total seats, leading seats, state-wise results, and member-wise results.
- After collecting the raw data, I used Python's Pandas and NumPy libraries to clean and preprocess it. This involved handling missing values, correcting data formats, and organizing the data into structured dataframes. These steps were crucial in ensuring the accuracy and reliability of the subsequent analysis.
- With the cleaned data, I imported it into Power BI for visualization and insights generation. Key insights included:
- **Party Performance**: Visualization of seats won, lost, and total for each party, highlighting their overall performance.
- **Leading Trends**: Analysis of leading seats to identify potential winners in various regions.
- **State-wise Analysis**: Comparative analysis of election results across different states to identify regional strongholds and battlegrounds.
- **Member-wise Results**: Detailed insights into the performance of individual candidates, including win/loss status and vote margins.
- Power BI's robust visualization tools enabled me to create interactive dashboards, making it easier to interpret the election results and uncover trends and patterns. This comprehensive approach provided valuable insights into the 2024 election outcomes, facilitating informed decision-making and strategic planning.

party wise parliamentary seats



Sum of SEATS and Sum of Total by Political Party



THANK YOU