```
!pip install selenium
!pip install pandas
!pip install openpyxl
!apt-get update
!apt install -y chromium-chromedriver
!cp /usr/lib/chromium-browser/chromedriver /usr/bin
```

```
Unpacking libudev1:amd64 (249.11-0ubuntu3.12) over (249.11-0ubuntu3.10) ...
Setting up libudev1:amd64 (249.11-0ubuntu3.12) ...
Selecting previously unselected package udev.
(Reading database ... 122133 files and directories currently installed.)
Preparing to unpack .../udev_249.11-0ubuntu3.12_amd64.deb ...
Unpacking udev (249.11-0ubuntu3.12) ...
Selecting previously unselected package libfuse3-3:amd64.
Preparing to unpack .../libfuse3-3_3.10.5-1build1_amd64.deb ...
Unpacking libfuse3-3:amd64 (3.10.5-1build1) ...
Selecting previously unselected package snapd.
Preparing to unpack .../snapd_2.63+22.04_amd64.deb ...
Unpacking snapd (2.63+22.04) ...
Setting up apparmor (3.0.4-2ubuntu2.3) ...
Created symlink /etc/systemd/system/sysinit.target.wants/apparmor.service → /lib/systemd/system/apparmor.service.
Setting up liblzo2-2:amd64 (2.10-2build3) ...
Setting up squashfs-tools (1:4.5-3build1) ...
Setting up udev (249.11-0ubuntu3.12) ...
invoke-rc.d: could not determine current runlevel
invoke-rc.d: policy-rc.d denied execution of start.
Setting up libfuse3-3:amd64 (3.10.5-1build1) ...
Setting up snapd (2.63+22.04) ...
Created symlink /etc/systemd/system/multi-user.target.wants/snapd.apparmor.service → /lib/systemd/system/snapd.apparmor.service.
Created symlink /etc/systemd/system/multi-user.target.wants/snapd.autoimport.service → /lib/systemd/system/snapd.autoimport.service.
Created symlink /etc/systemd/system/multi-user.target.wants/snapd.core-fixup.service → /lib/systemd/system/snapd.core-fixup.service.
Created symlink /etc/systemd/system/multi-user.target.wants/snapd.recovery-chooser-trigger.service → /lib/systemd/system/snapd.recover
Created symlink /etc/systemd/system/multi-user.target.wants/snapd.seeded.service → /lib/systemd/system/snapd.seeded.service.
Created symlink /etc/systemd/system/cloud-final.service.wants/snapd.seeded.service → /lib/systemd/system/snapd.seeded.service.
Unit /lib/systemd/system/snapd.seeded.service is added as a dependency to a non-existent unit cloud-final.service.
Created symlink /etc/systemd/system/multi-user.target.wants/snapd.service → /lib/systemd/system/snapd.service.
Created symlink /etc/systemd/system/timers.target.wants/snapd.snap-repair.timer → /lib/systemd/system/snapd.snap-repair.timer.
Created symlink /etc/systemd/system/sockets.target.wants/snapd.socket → /lib/systemd/system/snapd.socket.
Created symlink /etc/systemd/system/final.target.wants/snapd.system-shutdown.service → /lib/systemd/system/snapd.system-shutdown.servi
Selecting previously unselected package chromium-browser.
(Reading database ... 122363 files and directories currently installed.)
Preparing to unpack .../chromium-browser_1%3a85.0.4183.83-0ubuntu2.22.04.1_amd64.deb ...
=> Installing the chromium snap
==> Checking connectivity with the snap store
===> System doesn't have a working snapd, skipping
Unpacking chromium-browser (1:85.0.4183.83-0ubuntu2.22.04.1) ...
Selecting previously unselected package chromium-chromedriver.
Preparing to unpack .../chromium-chromedriver_1%3a85.0.4183.83-0ubuntu2.22.04.1_amd64.deb ...
Unpacking chromium-chromedriver (1:85.0.4183.83-0ubuntu2.22.04.1) ...
Selecting previously unselected package systemd-hwe-hwdb.
Preparing to unpack .../systemd-hwe-hwdb_249.11.5_all.deb ...
Unpacking systemd-hwe-hwdb (249.11.5) ...
Setting up systemd-hwe-hwdb (249.11.5) ...
Setting up chromium-browser (1:85.0.4183.83-0ubuntu2.22.04.1) ...
update-alternatives: using /usr/bin/chromium-browser to provide /usr/bin/x-www-browser (x-www-browser) in auto mode
update-alternatives: using /usr/bin/chromium-browser to provide /usr/bin/gnome-www-browser (gnome-www-browser) in auto mode
Setting up chromium-chromedriver (1:85.0.4183.83-0ubuntu2.22.04.1) ...
Processing triggers for udev (249.11-0ubuntu3.12) ...
Processing triggers for hicolor-icon-theme (0.17-2) ...
Processing triggers for libc-bin (2.35-0ubuntu3.4) ...
/sbin/ldconfig.real: /usr/local/lib/libtbbbind_2_5.so.3 is not a symbolic link

/sbin/ldconfig.real: /usr/local/lib/libtbbbind_2_0.so.3 is not a symbolic link
```

```
!pip install webdriver-manager
```

```
Collecting webdriver-manager
  Downloading webdriver_manager-4.0.1-py2.py3-none-any.whl (27 kB)
Requirement already satisfied: requests in /usr/local/lib/python3.10/dist-packages (from webdriver-manager) (2.31.0)
Collecting python-dotenv (from webdriver-manager)
  Downloading python_dotenv-1.0.1-py3-none-any.whl (19 kB)
Requirement already satisfied: packaging in /usr/local/lib/python3.10/dist-packages (from webdriver-manager) (24.1)
Requirement already satisfied: charset-normalizer<4,>=2 in /usr/local/lib/python3.10/dist-packages (from requests->webdriver-manager) (3
Requirement already satisfied: idna<4,>=2.5 in /usr/local/lib/python3.10/dist-packages (from requests->webdriver-manager) (3.7)
Requirement already satisfied: urllib3<3,>=1.21.1 in /usr/local/lib/python3.10/dist-packages (from requests->webdriver-manager) (2.0.7)
Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.10/dist-packages (from requests->webdriver-manager) (2024.6.
Installing collected packages: python-dotenv, webdriver-manager
Successfully installed python-dotenv-1.0.1 webdriver-manager-4.0.1
```

```
!apt install -y chromium-chromedriver
!cp /usr/lib/chromium-browser/chromedriver /usr/bin/chromedriver
```

```
Reading package lists... Done
Building dependency tree... Done
Reading state information... Done
chromium-chromedriver is already the newest version (1:85.0.4183.83-0ubuntu2.22.04.1).
0 upgraded, 0 newly installed, 0 to remove and 45 not upgraded.
cp: '/usr/lib/chromium-browser/chromedriver' and '/usr/bin/chromedriver' are the same file
```

testing the code for scraping

Double-click (or enter) to edit

```python
import requests
from bs4 import BeautifulSoup
import pandas as pd


def scrape_data(url):
    try:
        # Fetch the webpage content
        response = requests.get(url)
        response.raise_for_status()
        soup = BeautifulSoup(response.content, 'html.parser')
        tables = soup.find_all('table')

        # Extract data from tables
        table_data = []
        for table in tables:
            table_rows = table.find_all('tr')
            for row in table_rows:
                cols = row.find_all(['th', 'td'])
                cols = [ele.text.strip() for ele in cols]
                table_data.append(cols)
        df = pd.DataFrame(table_data)

        df = df.dropna(how='all').dropna(axis=1, how='all')

        return df
    except Exception as e:
        print(f"Error scraping {url}: {e}")
        return pd.DataFrame()

# Combined list of States and Union Territories with URLs for Election Commission party-wise results pages
state_urls = {
    'INDIA': 'https://results.eci.gov.in/PcResultGenJune2024/index.htm',
    'Andaman and Nicobar Islands': 'https://results.eci.gov.in/PcResultGenJune2024/partywiseresult-U01.htm',
    'Andhra Pradesh': 'https://results.eci.gov.in/PcResultGenJune2024/partywiseresult-S01.htm',
    'Arunachal Pradesh': 'https://results.eci.gov.in/PcResultGenJune2024/partywiseresult-S02.htm',
    'Assam': 'https://results.eci.gov.in/PcResultGenJune2024/partywiseresult-S03.htm',
    'Bihar': 'https://results.eci.gov.in/PcResultGenJune2024/partywiseresult-S04.htm',
    'Chandigarh': 'https://results.eci.gov.in/PcResultGenJune2024/partywiseresult-U02.htm',
    'Chhattisgarh': 'https://results.eci.gov.in/PcResultGenJune2024/partywiseresult-S26.htm',
    'Dadra and Nagar Haveli and Daman and Diu': 'https://results.eci.gov.in/PcResultGenJune2024/partywiseresult-U03.htm',
    'Goa': 'https://results.eci.gov.in/PcResultGenJune2024/partywiseresult-S05.htm',
    'Gujarat': 'https://results.eci.gov.in/PcResultGenJune2024/partywiseresult-S06.htm',
    'Haryana': 'https://results.eci.gov.in/PcResultGenJune2024/partywiseresult-S07.htm',
    'Himachal Pradesh': 'https://results.eci.gov.in/PcResultGenJune2024/partywiseresult-S08.htm',
    'Jammu and Kashmir': 'https://results.eci.gov.in/PcResultGenJune2024/partywiseresult-U08.htm',
    'Jharkhand': 'https://results.eci.gov.in/PcResultGenJune2024/partywiseresult-S27.htm',
    'Karnataka': 'https://results.eci.gov.in/PcResultGenJune2024/partywiseresult-S10.htm',
    'Kerala': 'https://results.eci.gov.in/PcResultGenJune2024/partywiseresult-S11.htm',
    'Lakshadweep': 'https://results.eci.gov.in/PcResultGenJune2024/partywiseresult-U09.htm',
    'Madhya Pradesh': 'https://results.eci.gov.in/PcResultGenJune2024/partywiseresult-S12.htm',
    'Maharashtra': 'https://results.eci.gov.in/PcResultGenJune2024/partywiseresult-S13.htm',
    'Manipur': 'https://results.eci.gov.in/PcResultGenJune2024/partywiseresult-S14.htm',
    'Meghalaya': 'https://results.eci.gov.in/PcResultGenJune2024/partywiseresult-S15.htm',
    'Mizoram': 'https://results.eci.gov.in/PcResultGenJune2024/partywiseresult-S16.htm',
    'Nagaland': 'https://results.eci.gov.in/PcResultGenJune2024/partywiseresult-S17.htm',
    'Odisha': 'https://results.eci.gov.in/PcResultGenJune2024/partywiseresult-S18.htm',
    'Puducherry': 'https://results.eci.gov.in/PcResultGenJune2024/partywiseresult-U10.htm',
    'Punjab': 'https://results.eci.gov.in/PcResultGenJune2024/partywiseresult-S19.htm',
```

```
        'Rajasthan': 'https://results.eci.gov.in/PcResultGenJune2024/partywiseresult-S20.htm',
        'Sikkim': 'https://results.eci.gov.in/PcResultGenJune2024/partywiseresult-S21.htm',
        'Tamil Nadu': 'https://results.eci.gov.in/PcResultGenJune2024/partywiseresult-S22.htm',
        'Telangana': 'https://results.eci.gov.in/PcResultGenJune2024/partywiseresult-S23.htm',
        'Tripura': 'https://results.eci.gov.in/PcResultGenJune2024/partywiseresult-S24.htm',
        'Uttar Pradesh': 'https://results.eci.gov.in/PcResultGenJune2024/partywiseresult-S25.htm',
        'Uttarakhand': 'https://results.eci.gov.in/PcResultGenJune2024/partywiseresult-S28.htm',
        'West Bengal': 'https://results.eci.gov.in/PcResultGenJune2024/partywiseresult-S29.htm'
}



all_state_results = []

# Iterate over each state and scrape data
for state_name, state_url in state_urls.items():
    state_df = scrape_data(state_url)
    if not state_df.empty:
        state_df['State'] = state_name
        all_state_results.append(state_df)

if all_state_results:
    # Concatenate all DataFrames into a single DataFrame
    final_df = pd.concat(all_state_results, ignore_index=True)
    combined_output_file = 'results_of_elections_including_states.xlsx'
    final_df.to_excel(combined_output_file, index=False)
    print(final_df.head())
else:
    print("No data scraped. Check URLs or script for errors.")
```

```
Error scraping https://results.eci.gov.in/PcResultGenJune2024/partywiseresult-U10.htm: 404 Client Error: Not Found for url: https://resu
                                    0    1        2      3  State
0                               Party  Won  Leading  Total  INDIA
1          Bharatiya Janata Party - BJP  240        0    240  INDIA
2         Indian National Congress - INC   99        0     99  INDIA
3                  Samajwadi Party - SP   37        0     37  INDIA
4  All India Trinamool Congress - AITC   29        0     29  INDIA
```