PROJECT REPORT

Job Preference Data Collection System Using Selenium Automation

Business Efficiency Architect (BEA)

Prepared by: Krishna Priya V
Date: June 2025
Institution: Panimalar Engineering College Chennai City Campus
Course: B.E. Computer Science and Engineering

S.No	Section Title	Page No.
1	Introduction	1
2	Objective	1
3	Tools & Technologies Used	1
4	System Requirements	2
5	Project Workflow	2
6	HTML Form Design	2
7	Python Automation Script	3
8	Sample Data (CSV Output)	4
9	Data Visualization (Apple Numbers)	4
10	Relevance to Business Efficiency Architect (BEA)	5
11	Conclusion	5
12	Future Scope	5
13	Attachments	6

---

1. Introduction

Understanding individual preferences is key to efficient workforce planning. This project showcases a real-time application that collects user job role preferences using an HTML form and automates data extraction through Python's Selenium WebDriver. The data is stored in CSV format and visualized using Apple Numbers. The solution reflects the automation and optimization skills crucial for a Business Efficiency Architect (BEA).

---

2. Objective

The goal of this project is to:

- Create an interactive form to collect job role preferences.
- Use Selenium to automate the process of reading data from the form.
- Save the responses into a structured CSV file.
- Allow users to visualize results using Apple Numbers.

---

3. Tools & Technologies Used

Technology	Description
HTML	Frontend form for data collection
Python	Programming logic and data handling
Selenium	Web automation
Pandas	Data formatting and CSV export
ChromeDriver	Interface for browser automation
Apple Numbers	Visual representation (charts)

---

4. System Requirements

- macOS operating system
- Python 3.x installed
- Chrome browser with corresponding ChromeDriver
- Required Python packages: selenium, pandas

---

5. Project Workflow

1. Users fill out a local HTML form containing personal data and job preference.
2. Form data is dynamically added to a visible HTML table (responseTable).
3. A Python script (with Selenium) waits for at least 4 form submissions.
4. Data from the table is scraped using Selenium.
5. Data is processed with Pandas and exported as a CSV file.
6. CSV file is opened in Apple Numbers to create visual charts.

---

6. HTML Form Design

The form contains:

- Input for name
- Dropdown for age group
- Radio buttons for student/professional status
- Job role selection: BEA, Cloud Architect, Cybersecurity, etc.
- Submit button
- JavaScript code to dynamically update the table below the form with each new entry

---

7. Python Automation Script

- Launches a local browser window with the form using Selenium
- Waits 30+ seconds for user to complete entries
- Identifies the response table by its ID
- Extracts table content row by row
- Saves the information into job_preferences.csv
- Closes the browser after saving

Python Code:

```
from selenium import webdriver
from selenium.webdriver.chrome.service import Service
from selenium.webdriver.common.by import By
import pandas as pd
import time
import os

# Setup correct ChromeDriver path
driver_path = "./chromedriver"
service = Service(driver_path)

# Launch Chrome using Service
driver = webdriver.Chrome(service=service)

# Load the local HTML form
html_path = f"file://{os.getcwd()}/job_form.html"
driver.get(html_path)

# Wait for manual form submissions
print("Please fill the form manually now...")
time.sleep(60)  # wait for at least 4 entries

# Scrape the response table
table = driver.find_element(By.ID, "responseTable")
rows = table.find_elements(By.TAG_NAME, "tr")[1:]

data = []
for row in rows:
    cols = row.find_elements(By.TAG_NAME, "td")
    data.append([col.text for col in cols])

# Save to CSV
df = pd.DataFrame(data, columns=["Name", "Age Group", "Status", "Job Role"])
df.to_csv("job_preferences.csv", index=False)
print("✅ Data saved to job_preferences.csv")

driver.quit()
```

---

8. Sample Data (CSV Output)

Name	Age Group	Status	Job Role
Krishna Priya	Under 18	Student	Business Efficient Architect (BEA)
Mirdhula	Under 18	Student	Business Efficient Architect (BEA)
Navya Shree	Under 18	Student	Application Developer
Pooja	Under 18	Student	Cloud Architect

9.Data Visualization (Apple Numbers)

Importing the `job_preferences.csv` file into **Apple Numbers** is a crucial step toward understanding user preferences in a visual format. Although charts have not yet been created, the structured data collected using Selenium.

1. **CSV Export**: The Python automation script successfully generated a CSV file containing user responses.

2. **Import into Apple Numbers**:

    - Open Apple Numbers.
    - Choose “File” → “Open” and select `job_preferences.csv`.
    - Data appears in tabular format with columns for Name, Age Group, Status, and Job Role.

---

10. Relevance to Business Efficiency Architect (BEA)

This project demonstrates BEA responsibilities such as:

- Workflow automation
- Data collection optimization
- Structured data handling
- Real-time insights through visual representation

---

11. Conclusion

This project successfully demonstrates how automation can simplify data collection and improve operational efficiency — core objectives of the Business Efficiency Architect (BEA) role. By integrating a custom HTML form with Selenium automation and Python data handling, we created a seamless system to collect, process, and visualize user preferences for job roles.

The solution reflects real-world applications such as employee interest surveys, feedback systems, and resource planning. It emphasizes the importance of workflow automation, structured data analysis, and meaningful presentation — all essential skills for a BEA.

This hands-on project also shows the ability to work with automation tools, data processing libraries, and real-time interfaces, aligning with industry expectations for BEA professionals.

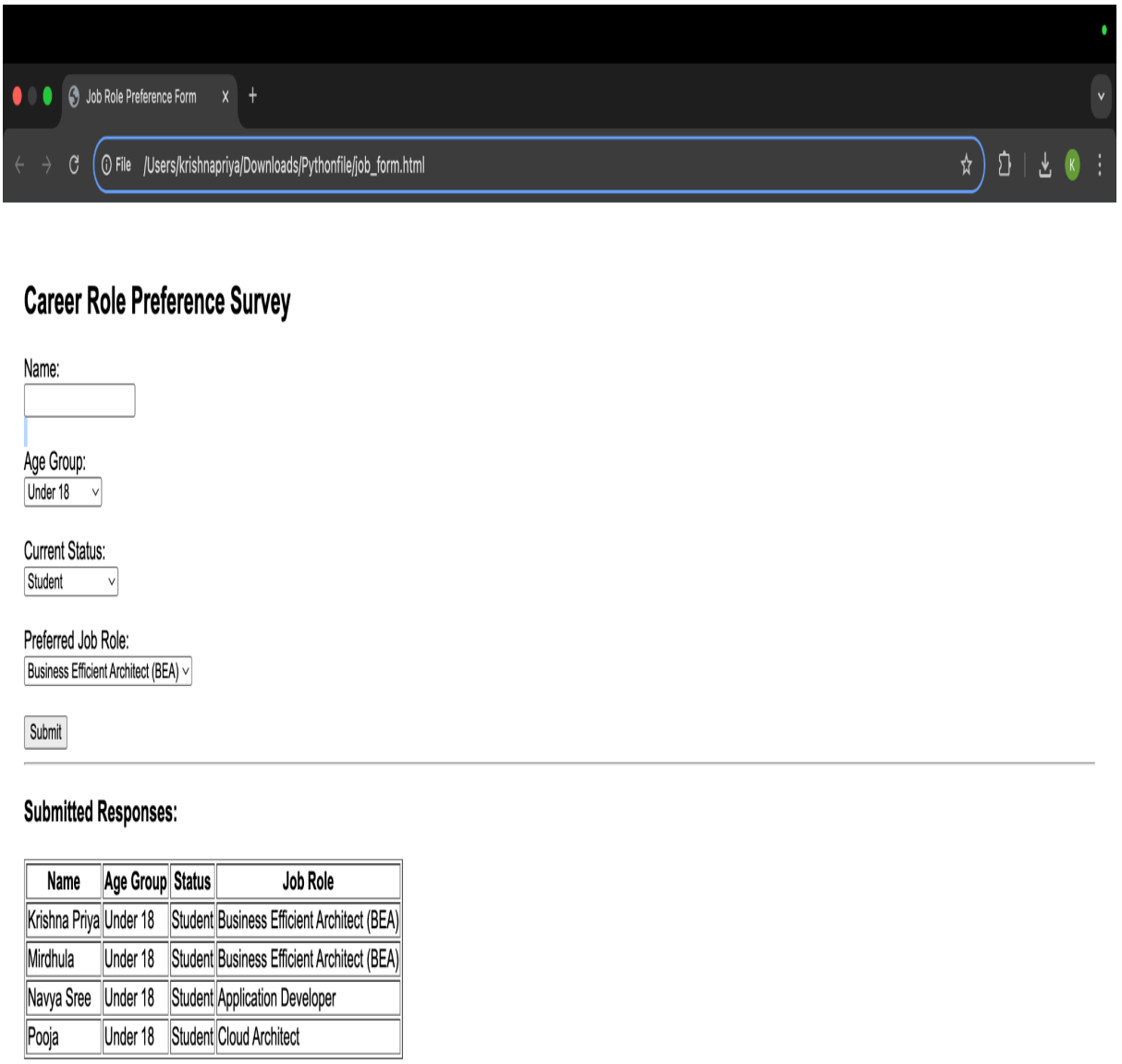
---

12. Future Scope

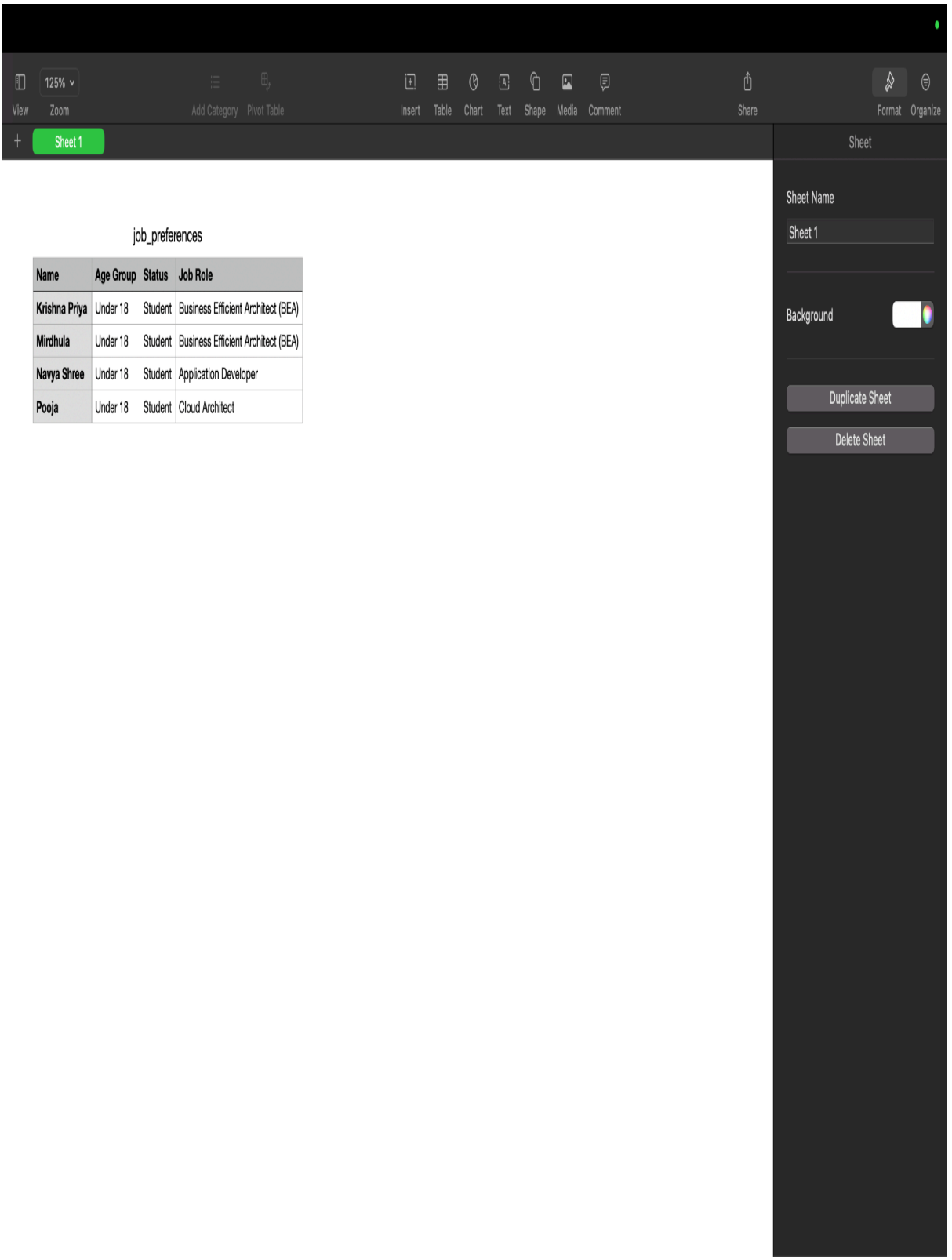
- Host the form online and collect submissions via a database
- Integrate with cloud services like Google Sheets
- Add user login functionality
- Create automated reports and charts using Python libraries

Attachments

- Screenshot of form



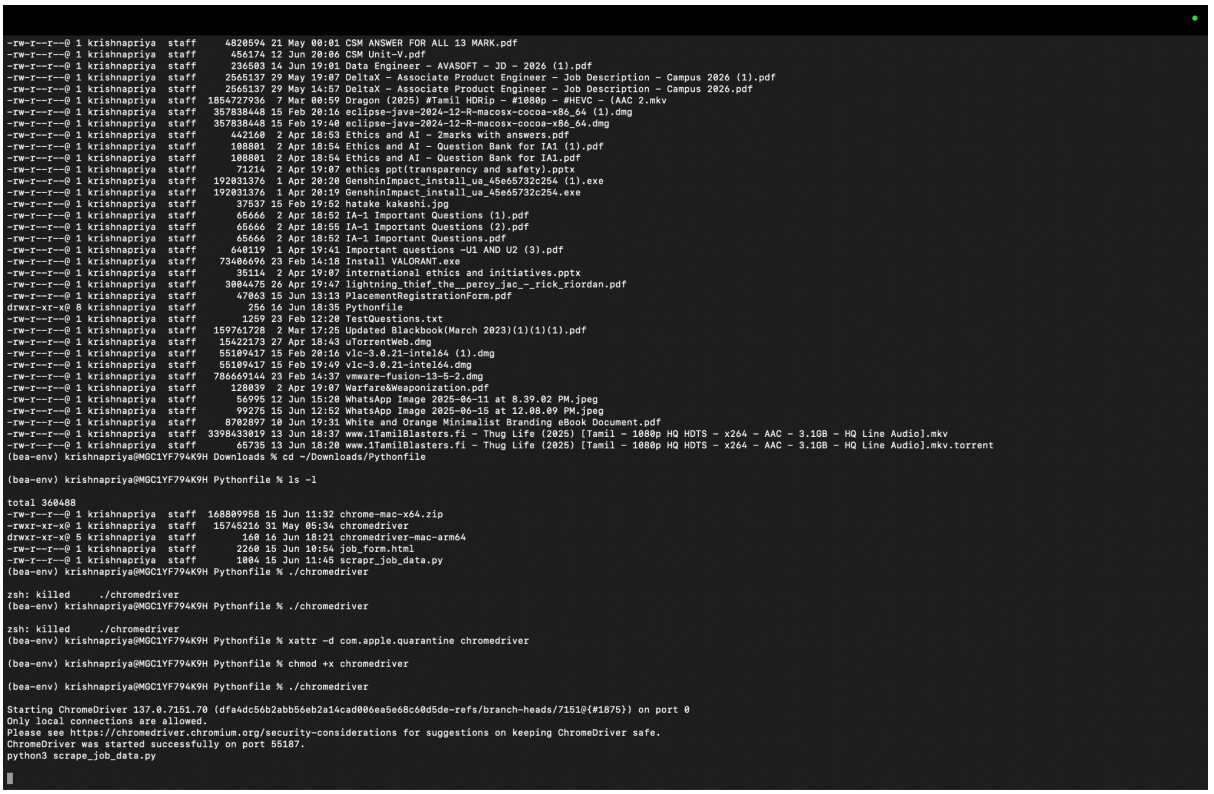
- Sample chart from Apple Numbers



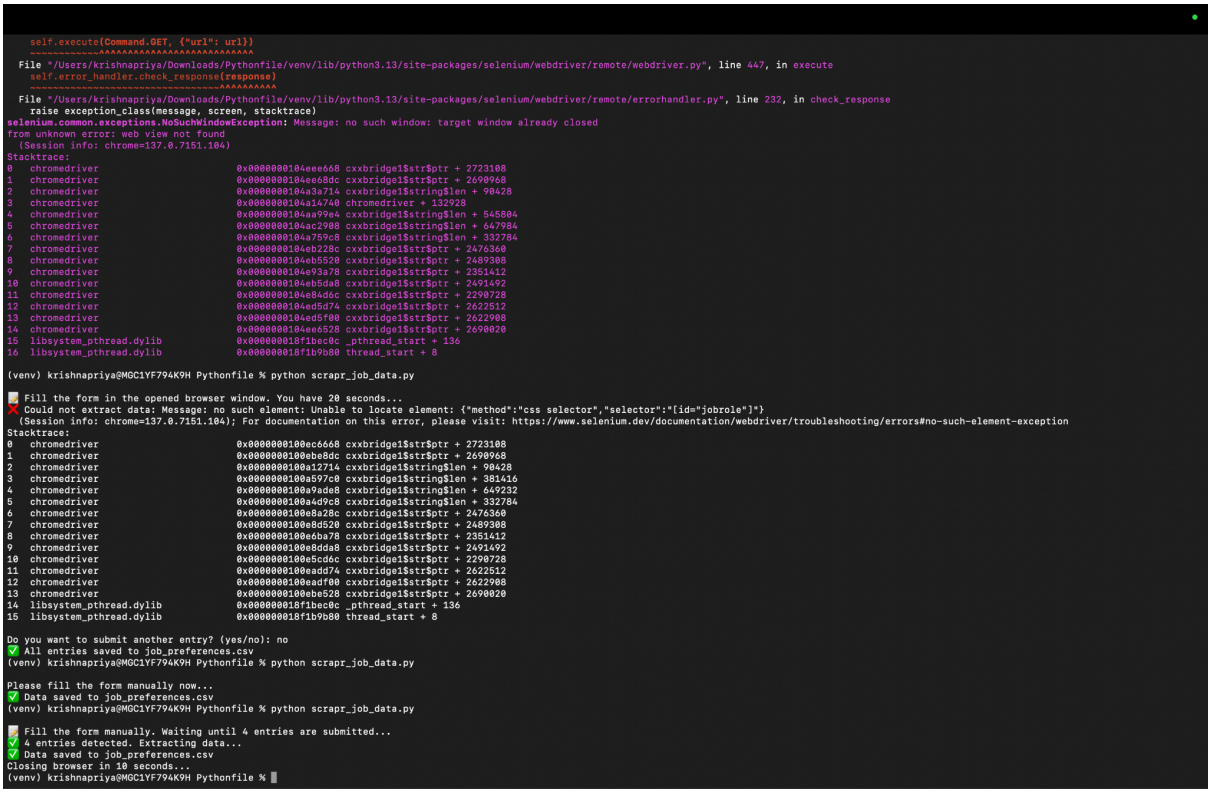
job_preferences

Name	Age Group	Status	Job Role
Krishna Priya	Under 18	Student	Business Efficient Architect (BEA)
Mirdhula	Under 18	Student	Business Efficient Architect (BEA)
Navya Shree	Under 18	Student	Application Developer
Pooja	Under 18	Student	Cloud Architect

- Execution Output & Terminal Logs



```
-rw-r--r--@ 1 krishnapriya  staff     4820594 21 May 00:01 CSM ANSWER FOR ALL 13 MARK.pdf
-rw-r--r--@ 1 krishnapriya  staff      456174 12 Jun 20:06 CSM Unit-V.pdf
-rw-r--r--@ 1 krishnapriya  staff      236503 14 Jun 19:01 Data Engineer - AVASOFT - JD - 2026 (1).pdf
-rw-r--r--@ 1 krishnapriya  staff     2565137 29 May 19:07 DeltaX - Associate Product Engineer - Job Description - Campus 2026 (1).pdf
-rw-r--r--@ 1 krishnapriya  staff     2565137 29 May 14:57 DeltaX - Associate Product Engineer - Job Description - Campus 2026.pdf
-rw-r--r--@ 1 krishnapriya  staff  1854727936  7 Mar 00:59 Dragon (2025) #Tamil HDRip - #1080p - #HEVC - (AAC 2.mkv
-rw-r--r--@ 1 krishnapriya  staff   357838448 15 Feb 20:16 eclipse-java-2024-12-R-macosx-cocoa-x86_64 (1).dmg
-rw-r--r--@ 1 krishnapriya  staff   357838448 15 Feb 19:40 eclipse-java-2024-12-R-macosx-cocoa-x86_64.dmg
-rw-r--r--@ 1 krishnapriya  staff      442160  2 Apr 18:53 Ethics and AI - 2marks with answers.pdf
-rw-r--r--@ 1 krishnapriya  staff      108801  2 Apr 18:54 Ethics and AI - Question Bank for IA1 (1).pdf
-rw-r--r--@ 1 krishnapriya  staff      108801  2 Apr 18:54 Ethics and AI - Question Bank for IA1.pdf
-rw-r--r--@ 1 krishnapriya  staff       71214  2 Apr 19:07 ethics ppt(transparency and safety).pptx
-rw-r--r--@ 1 krishnapriya  staff   192031376  1 Apr 20:20 GenshinImpact_install_ua_45e65732c254 (1).exe
-rw-r--r--@ 1 krishnapriya  staff   192031376  1 Apr 20:19 GenshinImpact_install_ua_45e65732c254.exe
-rw-r--r--@ 1 krishnapriya  staff       37537 15 Feb 19:52 hatake kakashi.jpg
-rw-r--r--@ 1 krishnapriya  staff       65666  2 Apr 18:52 IA-1 Important Questions (1).pdf
-rw-r--r--@ 1 krishnapriya  staff       65666  2 Apr 18:55 IA-1 Important Questions (2).pdf
-rw-r--r--@ 1 krishnapriya  staff       65666  2 Apr 18:52 IA-1 Important Questions.pdf
-rw-r--r--@ 1 krishnapriya  staff      640119  1 Apr 19:41 Important questions -U1 AND U2 (3).pdf
-rw-r--r--@ 1 krishnapriya  staff   734066696 23 Feb 14:18 Install VALORANT.exe
-rw-r--r--@ 1 krishnapriya  staff       35114  2 Apr 19:07 international ethics and initiatives.pptx
-rw-r--r--@ 1 krishnapriya  staff     3004475 26 Apr 19:47 lightning_thief_the__percy_jac_-_rick_riordan.pdf
-rw-r--r--@ 1 krishnapriya  staff       47063 15 Jun 13:13 PlacementRegistrationForm.pdf
drwxr-xr-x@ 8 krishnapriya  staff         256 16 Jun 18:35 Pythonfile
-rw-r--r--@ 1 krishnapriya  staff        1259 23 Feb 12:20 TestQuestions.txt
-rw-r--r--@ 1 krishnapriya  staff   159761728  2 Mar 17:25 Updated Blackbook(March 2023)(1)(1)(1).pdf
-rw-r--r--@ 1 krishnapriya  staff    15422173 27 Apr 18:43 uTorrentWeb.dmg
-rw-r--r--@ 1 krishnapriya  staff    55109417 15 Feb 20:16 vlc-3.0.21-intel64 (1).dmg
-rw-r--r--@ 1 krishnapriya  staff    55109417 15 Feb 19:49 vlc-3.0.21-intel64.dmg
-rw-r--r--@ 1 krishnapriya  staff   786669144 23 Feb 14:37 vmware-fusion-13-5-2.dmg
-rw-r--r--@ 1 krishnapriya  staff      128039  2 Apr 19:07 Warfare&Weaponization.pdf
-rw-r--r--@ 1 krishnapriya  staff       56995 12 Jun 15:20 WhatsApp Image 2025-06-11 at 8.39.02 PM.jpeg
-rw-r--r--@ 1 krishnapriya  staff       99275 15 Jun 12:52 WhatsApp Image 2025-06-15 at 12.08.09 PM.jpeg
-rw-r--r--@ 1 krishnapriya  staff     8702897 10 Jun 19:31 White and Orange Minimalist Branding eBook Document.pdf
-rw-r--r--@ 1 krishnapriya  staff  3398433019 13 Jun 18:37 www.1TamilBlasters.fi - Thug Life (2025) [Tamil - 1080p HQ HDTS - x264 - AAC - 3.1GB - HQ Line Audio].mkv
-rw-r--r--@ 1 krishnapriya  staff       65735 13 Jun 18:20 www.1TamilBlasters.fi - Thug Life (2025) [Tamil - 1080p HQ HDTS - x264 - AAC - 3.1GB - HQ Line Audio].mkv.torrent
(bea-env) krishnapriya@MGC1YF794K9H Downloads % cd ~/Downloads/Pythonfile

(bea-env) krishnapriya@MGC1YF794K9H Pythonfile % ls -l

total 360488
-rw-r--r--@ 1 krishnapriya  staff  168809958 15 Jun 11:32 chrome-mac-x64.zip
-rwxr-xr-x@ 1 krishnapriya  staff   15745216 31 May 05:34 chromedriver
drwxr-xr-x@ 5 krishnapriya  staff        160 16 Jun 18:21 chromedriver-mac-arm64
-rw-r--r--@ 1 krishnapriya  staff       2260 15 Jun 10:54 job_form.html
-rw-r--r--@ 1 krishnapriya  staff       1804 15 Jun 11:45 scrapr_job_data.py
(bea-env) krishnapriya@MGC1YF794K9H Pythonfile % ./chromedriver

zsh: killed     ./chromedriver
(bea-env) krishnapriya@MGC1YF794K9H Pythonfile % ./chromedriver

zsh: killed     ./chromedriver
(bea-env) krishnapriya@MGC1YF794K9H Pythonfile % xattr -d com.apple.quarantine chromedriver

(bea-env) krishnapriya@MGC1YF794K9H Pythonfile % chmod +x chromedriver

(bea-env) krishnapriya@MGC1YF794K9H Pythonfile % ./chromedriver

Starting ChromeDriver 137.0.7151.70 (dfa4dc56b2abb56eb2a14cad006ea5e68c60d5de-refs/branch-heads/7151@{#1875}) on port 0
Only local connections are allowed.
Please see https://chromedriver.chromium.org/security-considerations for suggestions on keeping ChromeDriver safe.
ChromeDriver was started successfully on port 55187.
python3 scrape_job_data.py
```



```
    self.execute(Command.GET, {"url": url})
    ~~~~~~~~~~~~^^^^^^^^^^^^^^^^^^^^^^^^^^^^
  File "/Users/krishnapriya/Downloads/Pythonfile/venv/lib/python3.13/site-packages/selenium/webdriver/remote/webdriver.py", line 447, in execute
    self.error_handler.check_response(response)
    ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~^^^^^^^^^^
  File "/Users/krishnapriya/Downloads/Pythonfile/venv/lib/python3.13/site-packages/selenium/webdriver/remote/errorhandler.py", line 232, in check_response
    raise exception_class(message, screen, stacktrace)
selenium.common.exceptions.NoSuchWindowException: Message: no such window: target window already closed
from unknown error: web view not found
  (Session info: chrome=137.0.7151.104)
Stacktrace:
0   chromedriver                        0x0000000104eee668 cxxbridge1$str$ptr + 2723108
1   chromedriver                        0x0000000104ee68dc cxxbridge1$str$ptr + 2690968
2   chromedriver                        0x0000000104a3a714 cxxbridge1$string$len + 90428
3   chromedriver                        0x0000000104a14740 chromedriver + 132928
4   chromedriver                        0x0000000104a99e4 cxxbridge1$string$len + 545804
5   chromedriver                        0x0000000104ac2908 cxxbridge1$string$len + 647984
6   chromedriver                        0x0000000104a759c8 cxxbridge1$string$len + 332784
7   chromedriver                        0x0000000104eb228c cxxbridge1$str$ptr + 2476360
8   chromedriver                        0x0000000104eb5520 cxxbridge1$str$ptr + 2489308
9   chromedriver                        0x0000000104e93a78 cxxbridge1$str$ptr + 2351412
10  chromedriver                        0x0000000104eb5da8 cxxbridge1$str$ptr + 2491492
11  chromedriver                        0x0000000104e84d6c cxxbridge1$str$ptr + 2290728
12  chromedriver                        0x0000000104ed5d74 cxxbridge1$str$ptr + 2622512
13  chromedriver                        0x0000000104ed5f00 cxxbridge1$str$ptr + 2622908
14  chromedriver                        0x0000000104ee6528 cxxbridge1$str$ptr + 2690020
15  libsystem_pthread.dylib             0x00000018f1bec0c _pthread_start + 136
16  libsystem_pthread.dylib             0x00000018f1b9b80 thread_start + 8

(venv) krishnapriya@MGC1YF794K9H Pythonfile % python scrapr_job_data.py

📝 Fill the form in the opened browser window. You have 20 seconds...
❌ Could not extract data: Message: no such element: Unable to locate element: {"method":"css selector","selector":"[id="jobrole"]"}
  (Session info: chrome=137.0.7151.104); For documentation on this error, please visit: https://www.selenium.dev/documentation/webdriver/troubleshooting/errors#no-such-element-exception
Stacktrace:
0   chromedriver                        0x0000000100ec6668 cxxbridge1$str$ptr + 2723108
1   chromedriver                        0x0000000100ebe8dc cxxbridge1$str$ptr + 2690968
2   chromedriver                        0x0000000100a12714 cxxbridge1$string$len + 90428
3   chromedriver                        0x0000000100a597c0 cxxbridge1$string$len + 381416
4   chromedriver                        0x0000000100a9ade8 cxxbridge1$string$len + 649232
5   chromedriver                        0x0000000100a4d9c8 cxxbridge1$string$len + 332784
6   chromedriver                        0x0000000100e8a28c cxxbridge1$str$ptr + 2476360
7   chromedriver                        0x0000000100e8d520 cxxbridge1$str$ptr + 2489308
8   chromedriver                        0x0000000100e6ba78 cxxbridge1$str$ptr + 2351412
9   chromedriver                        0x0000000100e8dda8 cxxbridge1$str$ptr + 2491492
10  chromedriver                        0x0000000100e5cd6c cxxbridge1$str$ptr + 2290728
11  chromedriver                        0x0000000100eadd74 cxxbridge1$str$ptr + 2622512
12  chromedriver                        0x0000000100eadf00 cxxbridge1$str$ptr + 2622908
13  chromedriver                        0x0000000100ebe528 cxxbridge1$str$ptr + 2690020
14  libsystem_pthread.dylib             0x00000018f1bec0c _pthread_start + 136
15  libsystem_pthread.dylib             0x00000018f1b9b80 thread_start + 8

Do you want to submit another entry? (yes/no): no
✅ All entries saved to job_preferences.csv
(venv) krishnapriya@MGC1YF794K9H Pythonfile % python scrapr_job_data.py

Please fill the form manually now...
✅ Data saved to job_preferences.csv
(venv) krishnapriya@MGC1YF794K9H Pythonfile % python scrapr_job_data.py

📝 Fill the form manually. Waiting until 4 entries are submitted...
✅ 4 entries detected. Extracting data...
✅ Data saved to job_preferences.csv
Closing browser in 10 seconds...
(venv) krishnapriya@MGC1YF794K9H Pythonfile %
```