

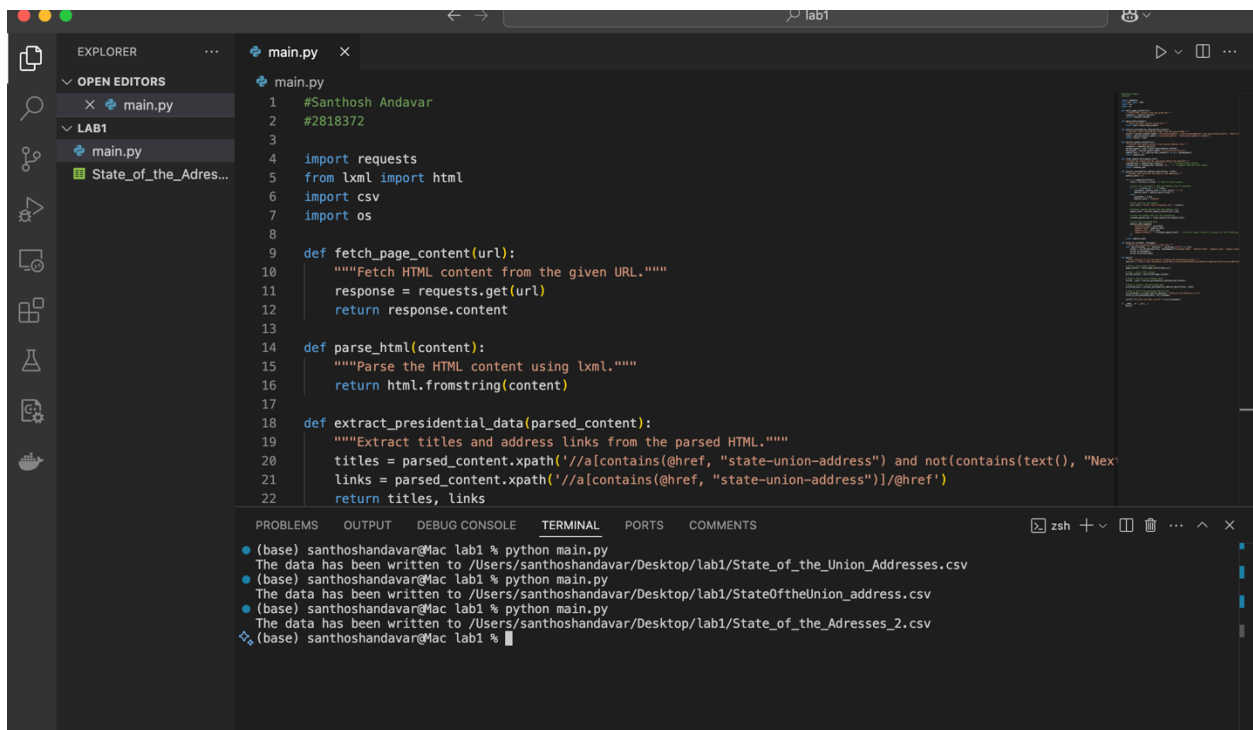
Santhosh Andavar

2818372

CIS 492

Lab1

Step 1: The data has been exported into a CSV file.



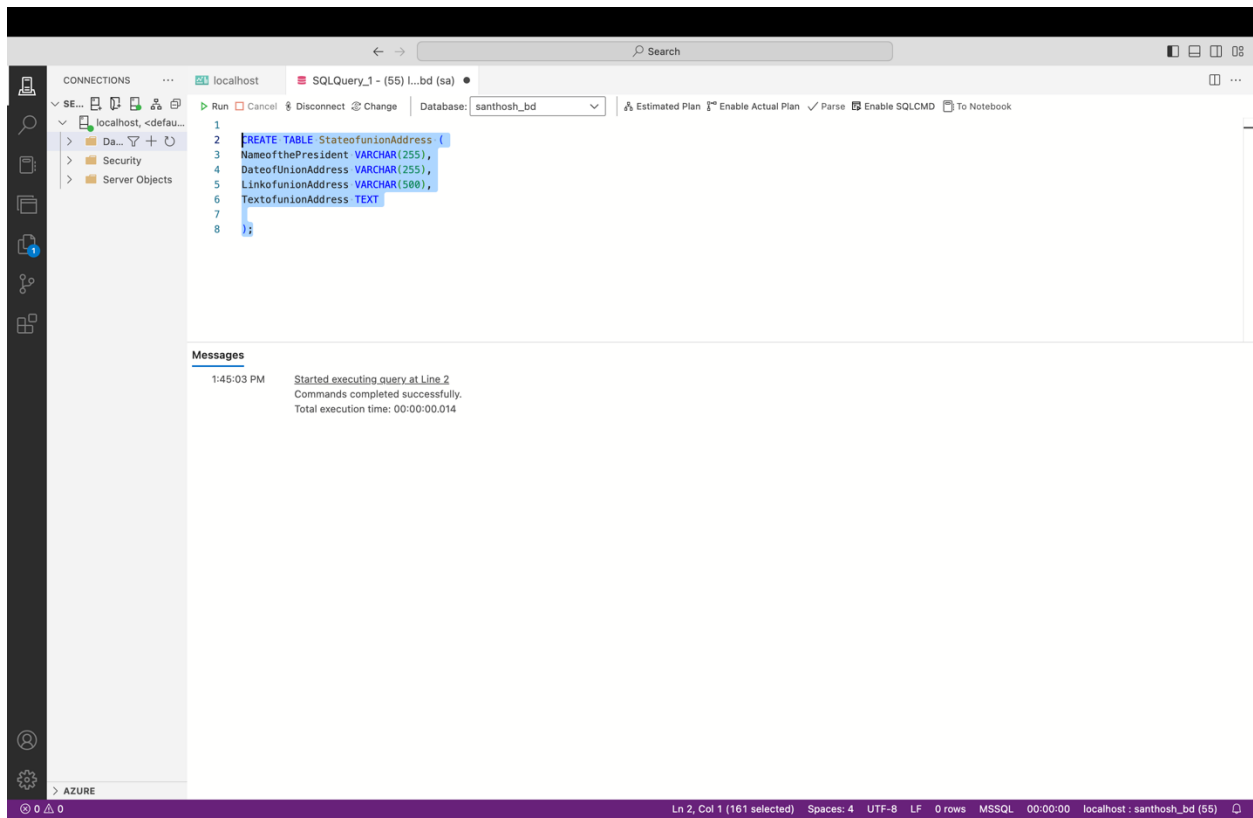
The screenshot shows a Visual Studio Code editor window with a Python script named `main.py` open. The script defines three functions: `fetch_page_content`, `parse_html`, and `extract_presidential_data`. The `extract_presidential_data` function uses `lxml` to parse HTML content and extract titles and links. The terminal at the bottom shows the execution of the script, which successfully writes data to two CSV files: `State_of_the_Union_Addresses.csv` and `StateOftheUnion_address.csv`.

```
1 #Santhosh Andavar
2 #2818372
3
4 import requests
5 from lxml import html
6 import csv
7 import os
8
9 def fetch_page_content(url):
10     """Fetch HTML content from the given URL."""
11     response = requests.get(url)
12     return response.content
13
14 def parse_html(content):
15     """Parse the HTML content using lxml."""
16     return html.fromstring(content)
17
18 def extract_presidential_data(parsed_content):
19     """Extract titles and address links from the parsed HTML."""
20     titles = parsed_content.xpath('//a[contains(@href, "state-union-address") and not(contains(text(), "Next"))]')
21     links = parsed_content.xpath('//a[contains(@href, "state-union-address")]/@href')
22     return titles, links
```

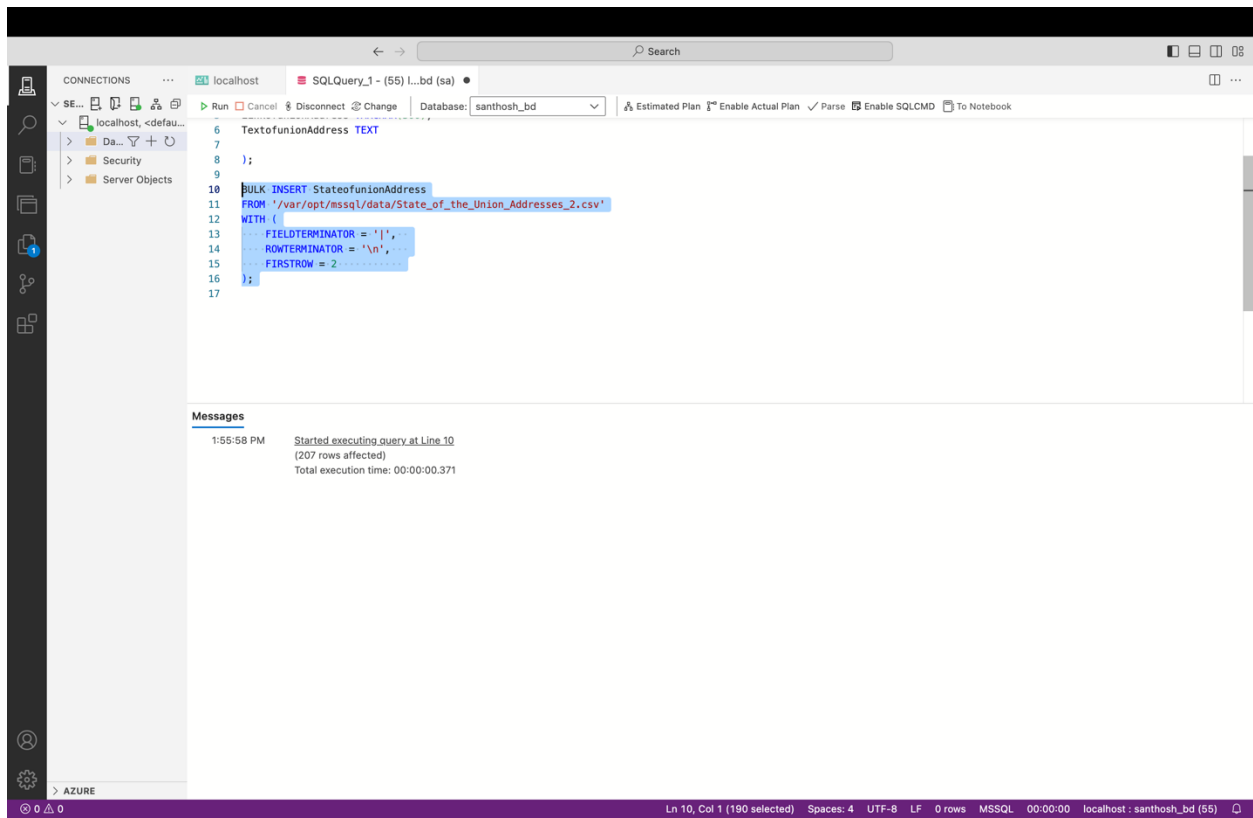
Terminal Output:

```
(base) santhoshandavar@Mac lab1 % python main.py
The data has been written to /Users/santhoshandavar/Desktop/lab1/State_of_the_Union_Addresses.csv
(base) santhoshandavar@Mac lab1 % python main.py
The data has been written to /Users/santhoshandavar/Desktop/lab1/StateOftheUnion_address.csv
(base) santhoshandavar@Mac lab1 % python main.py
The data has been written to /Users/santhoshandavar/Desktop/lab1/State_of_the_Addresses_2.csv
(base) santhoshandavar@Mac lab1 %
```

Step 2: The Extracted file



Step 4: The data from the CSV file has been inserted into the database using a bulk insert method.



Step 5: Final Output

SQLQuery_1 - (55) |...bd (sa)

```
6 TextofunionAddress TEXT
7
8 );
9
10 BULK INSERT StateofunionAddress
11 FROM '/var/opt/mssql/data/State_of_the_Union_Addresses_2.csv'
12 WITH (
13     FIELDTERMINATOR = '|',
14     ROWTERMINATOR = '\n',
15     FIRSTROW = 2
16 );
17
18
19 SELECT* from StateofunionAddress;
```

Results Messages

	NameofthePresident	DateofUnionAddress	LinkofunionAddress	TextofunionAddress
1	George Washington	January 8, 1790	https://www.infoplease.com/primary-sources/governmen...	"" Fellow-Citizens of the Senate and House of Repre...
2	George Washington	December 8, 1790	https://www.infoplease.com/primary-sources/governmen...	"" Fellow-Citizens of the Senate and House of Repre...
3	George Washington	October 25, 1791	https://www.infoplease.com/primary-sources/governmen...	"" Fellow-Citizens of the Senate and House of Repre...
4	George Washington	November 6, 1792	https://www.infoplease.com/primary-sources/governmen...	"" Fellow-Citizens of the Senate and House of Repre...
5	George Washington	December 3, 1793	https://www.infoplease.com/primary-sources/governmen...	"" Fellow-Citizens of the Senate and House of Repre...
6	George Washington	November 19, 1794	https://www.infoplease.com/primary-sources/governmen...	"" Fellow-Citizens of the Senate and House of Repre...
7	George Washington	December 8, 1795	https://www.infoplease.com/primary-sources/governmen...	"" Fellow-Citizens of the Senate and House of Repre...
8	George Washington	December 7, 1796	https://www.infoplease.com/primary-sources/governmen...	"" Fellow-Citizens of the Senate and House of Repre...
9	John Adams	November 22, 1797	https://www.infoplease.com/primary-sources/governmen...	"" Fellow-Citizens of the Senate and House of Repre...
10	John Adams	December 8, 1798	https://www.infoplease.com/primary-sources/governmen...	"" Gentlemen of the Senate and Gentlemen of the Hou...
11	John Adams	December 3, 1799	https://www.infoplease.com/primary-sources/governmen...	"" Gentlemen of the Senate and Gentlemen of the Hou...
12	John Adams	November 11, 1800	https://www.infoplease.com/primary-sources/governmen...	"" Gentlemen of the Senate and Gentlemen of the Hou...
13	Thomas Jefferson	December 8, 1801	https://www.infoplease.com/primary-sources/governmen...	"" Gentlemen of the Senate and Gentlemen of the Hou...
14	Thomas Jefferson	December 15, 1802	https://www.infoplease.com/primary-sources/governmen...	"" Fellow Citizens of the Senate and House of Repre...
15	Thomas Jefferson	October 17, 1803	https://www.infoplease.com/primary-sources/governmen...	"" To the Senate and House of Representatives: Wh...
16	Thomas Jefferson	November 8, 1804	https://www.infoplease.com/primary-sources/governmen...	"" To The Senate and House of Representatives of th...
17	Thomas Jefferson	December 2, 1806	https://www.infoplease.com/primary-sources/governmen...	"" The Senate and House of Representatives of the U...
18	Thomas Jefferson	October 27, 1807	https://www.infoplease.com/primary-sources/governmen...	"" The Senate and House of Representatives of the U...
19	Thomas Jefferson	November 8, 1808	https://www.infoplease.com/primary-sources/governmen...	"" The Senate and House of Representatives of the U...

Ln 19, Col 1 (33 selected) Spaces: 4 UTF-8 LF 207 rows MSSQL 00:00:00 localhost: santhosh_bd (55)

I have successfully extracted data from the Union Address website, saved it in a CSV file, and imported it into a database.