RED BUS

PROJECT

DOCUMENT

# INDEX

* Objective
* Web Scraping using Selenium
* SQL Database
* Stream lit Application

Objective:

To Develop a web scraper to extract the bus details, schedules, and relevant information from the RedBus website for multiple states throughout the nation. Store the extracted data in an SQL database and visualize it using a Stream lit app.

Web Scraping:

Approach :

* Use selenium-python to automate web browsing and data extraction from RedBus website.
* Handle dynamic content loading , pagination, and potential Pop-ups.

Steps:

1. Initialize web driver: Open the browser and handle the Browser like maximizing the browser and navigate to the RedBus website
2. Load Web Page: Load the specific URL for the target state, Handling any loading delays.
3. Scrape Bus Routes: Identify and extract all bus route links and name on the page, managing pagination to capture all routes.
4. Scrape Bus Details: navigate to each bus route link and extract detailed information about available buses, such as name, type, departing time, duration, reaching time, star rating, price, and seat availability.
5. Handling Errors: Using try-exception helps to implement the error handling for the element not found or loading failures, logging errors and continuing the scraping process.

SQL Database:

Approach:

* Use Python’s MySQL Database to store the scraped data

Steps:

1. Installation: Install MySQL database and create a work space and connect to python using connector.
2. Database Setup: create a database and create a table structure to store the scraped data.
3. Data Insertion: Insert the scraped data into the SQL database, ensuring data integrity and handling the duplicates or errors.

|  |  |
| --- | --- |
| Column names | Data types |
| Id | Int PRIMARY KEY AUTO INCREMENT |
| Route Name | TEXT |
| Route Link | TEXT |
| Bus Name | TEXT |
| Bus Type | TEXT |
| Departing Time | VARCHAR(100) |
| Duration | TEXT |
| Reaching Time | VARCHAR(100) |
| Star Rating | FLOAT |
| Price | FLOAT |
| Seat Availability | INT |

Stream lit Application:

Approach:

* Develop a Stream lit app to query and visualize the data from the SQL Database.

Steps:

1. Database Connection: Establish a Database connection with MySQL Database.
2. Query Data: Fetch the data from the database to be displayed in the app.
3. Filtering: Use stream lit components to filter the scraped data accordingly the user interest.

***Stream lit App Features:***

* *Display a table of buses and schedules by the users search.*
* *Using widgets, radio buttons, select box and search box user will easily see the buses information.*

THANK YOU