Redbus Data Scraping and Filtering with Streamlit Application

A Comprehensive Solution for Transportation Data Management

> By Harikrishnan G

Introduction

Efficient data management is crucial in the transportation industry. The "Redbus Data Scraping and Filtering with Streamlit Application" automates the collection, analysis, and visualization of bus travel data. By leveraging web scraping and interactive dashboards, this project provides real-time insights into bus schedules, prices, and seat availability, significantly enhancing decision-making and operational efficiency.

Problem Statement

Problem: Manual data collection is time-consuming and prone to errors, making it difficult to manage bus schedules, prices, and seat availability effectively.

Challenges: Discuss challenges like data inconsistency, lack of real-time information, and difficulty in visualizing complex datasets.

Goal: Automate data collection and provide tools for analysis and visualization.

Solution Overview

Data Collection: Using Selenium for web scraping.

Data Storage: Storing data in a SQL database.

Data Visualization: Using Streamlit to create an interactive user interface.

Outcome: Streamlined data management and enhanced decision-making capabilities.

Technology Stack

Web Scraping: Selenium for automated data extraction from Redbus.

Database: MySQL for structured data storage and querying.

Backend: Python for data processing and backend logic.

Frontend: Streamlit for creating an interactive dashboard.

Tools/libraries: Pandas, MySQL Connector, time, selenium & streamlit

Web Scraping with Selenium

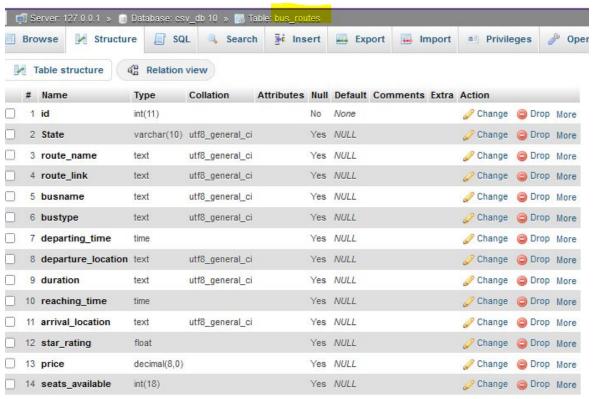
Explanation: Describe how Selenium is used to navigate Redbus, extract data, and handle dynamic content.

Features:

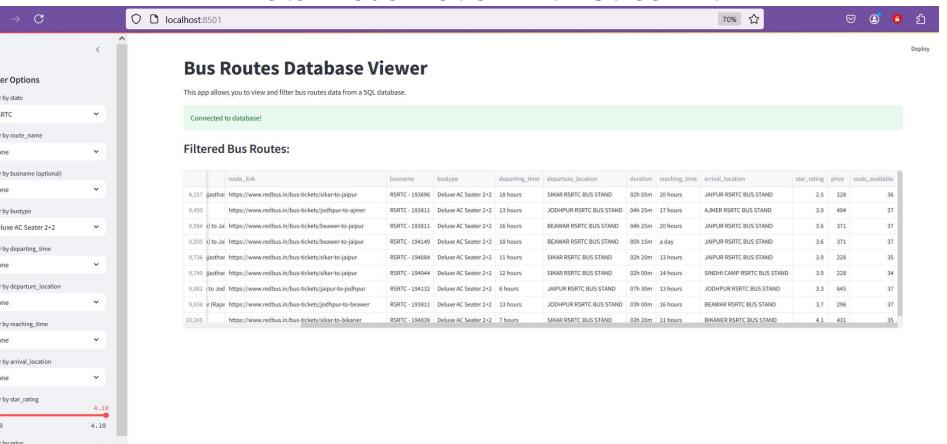
Automation of repetitive tasks.

Collection of complex data like bus schedules, prices, and seat availability.

Data Storage with MySQL



Data Visualization with Streamlit



Conclusion

This project can significantly improve operational efficiency, strategic planning, and overall user experience in the transportation sector.