Redbus Data Scraping with Selenium & Dynamic Filtering using Streamlit

Project Overview

This project aims to scrape bus data from Redbus, store it in a SQL database, and provide an interactive filtering interface using Streamlit. It focuses on automating the collection of travel data, analyzing it, and providing an interactive web application to filter and visualize the data.

Skills and Technologies Used

- Web Scraping: Selenium, Python
- Data Storage: SQL
- Data Analysis & Filtering: Streamlit, SQL queries
- Data Visualization: Plotly, Streamlit

Problem Statement

The goal of this project is to provide a comprehensive solution for collecting, analyzing, and visualizing bus travel data from the Redbus platform. By scraping real-time data, users can interactively filter bus routes, departure times, prices, seat availability, star ratings, and more.

Business Use Cases

This solution can be used for:

- **Travel Aggregators**: Provide real-time bus schedules and seat availability for users.
- **Market Research**: Analyze travel patterns and market preferences.
- **Customer Service**: Offer customized travel options based on filtered data.
- **Competitor Analysis**: Compare pricing and bus service quality with competitors.

Approach

Data Scraping

- Use Selenium to automate the extraction of bus-related data such as routes, schedules, prices, and seat availability from the Redbus website.

Data Storage

- Store the scraped data in a SQL database for structured storage and easy querying.

Streamlit Application

- Develop a Streamlit application to display and filter the scraped data.
- Implement various interactive filters, including bus type, route, price range, star rating, and seat availability.

Data Set

- **Source**: The data is scraped from the Redbus website.
- **Link**: Redbus
- **Required Fields**: Bus routes, bus name, bus type (Sleeper/Seater), departure time, duration,

Setup Instructions

Set up the SQL database:

Create the necessary database and tables.

Update the database connection in the file.

Run the scraping script:python sql.ipynb

Run the Streamlit app: streamlit run app.py

Open your browser and go to http://localhost:8509/ to interact with the application.

My Project

Here is a screenshot of my project

Deploy

Navigate

- Home
- Select Bus

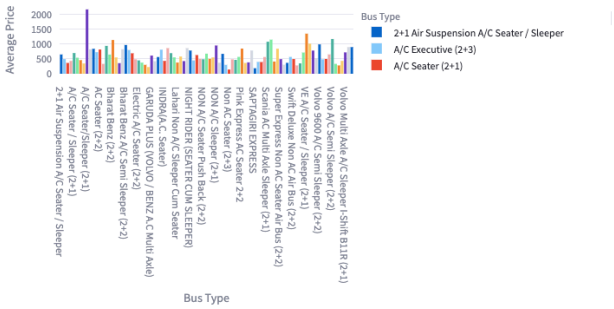
Redbus Data Explorer

Explore bus routes, prices, ratings, and seat availability!

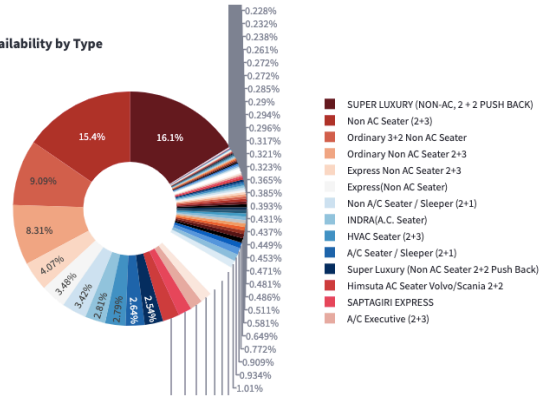
Find your perfect bus with customized filters and visual insights.

📊 Overview of Bus Data

Average Price by Bus Type



Bus Availability by Type



Deploy

Navigate

- Home
- Select Bus

Bus Filters

Select Route

Vijayawada to Hyderabad

Select Bus Type

SUPER LUXURY (NON-AC, 2 + 2 P…

Price Range

52 — 5000

52 — 5000

Star Rating

1 — 5

1 — 5

Show only buses with available seats

Sort By

Price: Low to High

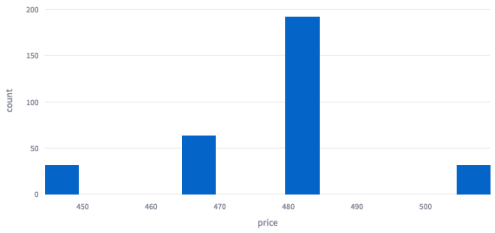
Found 320 buses matching your criteria

	route_name	route_link	busname
17,143	Vijayawada to Hyderabad	https://www.redbus.in/bus-tickets/vijayawada-to-hyderabad	APSRTC - 4052
19,201	Vijayawada to Hyderabad	https://www.redbus.in/bus-tickets/vijayawada-to-hyderabad	APSRTC - 4052
2,061	Vijayawada to Hyderabad	https://www.redbus.in/bus-tickets/vijayawada-to-hyderabad	APSRTC - 4052
18,857	Vijayawada to Hyderabad	https://www.redbus.in/bus-tickets/vijayawada-to-hyderabad	APSRTC - 4052
17,487	Vijayawada to Hyderabad	https://www.redbus.in/bus-tickets/vijayawada-to-hyderabad	APSRTC - 4052
3,431	Vijayawada to Hyderabad	https://www.redbus.in/bus-tickets/vijayawada-to-hyderabad	APSRTC - 4052
15,773	Vijayawada to Hyderabad	https://www.redbus.in/bus-tickets/vijayawada-to-hyderabad	APSRTC - 4052
15,429	Vijayawada to Hyderabad	https://www.redbus.in/bus-tickets/vijayawada-to-hyderabad	APSRTC - 4052
14,059	Vijayawada to Hyderabad	https://www.redbus.in/bus-tickets/vijayawada-to-hyderabad	APSRTC - 4052
3,775	Vijayawada to Hyderabad	https://www.redbus.in/bus-tickets/vijayawada-to-hyderabad	APSRTC - 4052

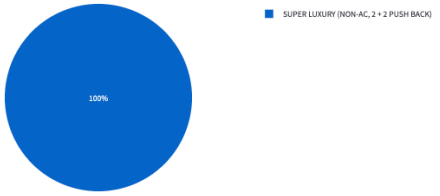
Download Data as CSV

Visual Insights from Filtered Data

Price Distribution



Bus Type Distribution



Price vs Star Rating