

RedBus Route Scraper Documentation:

Overview

This script scrapes bus route information from RedBus.in for 10 different state transport Organizations and stores the data in a MySQL database. It uses Selenium for web scraping And PyMySQL for database operations.

Code Explanation:

This code divided into two files, Redbus.ipynb and streamlitt.py. redbus file further divided into two code snippets. The first snippet scrape all the route names and route URL for ten states. It will be stored in a MYSQL database in states table. The second snippet fetch route URL from state table and scrape bus details for all the ten states. This will be stored in bus details table.

Streamlitt.py will be used for filter the data from bus details and displayed. Screenshoots attached below.

Database Configuration:

Host: "localhost"

User: "gokuld"

Password: "Guvi"

Database: "redbus"

Database Schema:

States Table

```
CREATE TABLE states (  
state_id INT AUTO_INCREMENT PRIMARY KEY,  
state_name VARCHAR(255) UNIQUE  
)
```

```
mysql> desc states;  
+-----+-----+-----+-----+-----+-----+  
| Field      | Type          | Null | Key | Default | Extra          |  
+-----+-----+-----+-----+-----+-----+  
| state_id   | int           | NO   | PRI | NULL    | auto_increment |  
| state_name | varchar(255)  | YES  | UNI | NULL    |                |  
+-----+-----+-----+-----+-----+-----+  
2 rows in set (0.02 sec)
```

```
mysql> select * from states;
```

state_id	state_name
10	Assam State Transport Corporation (ASTC)
7	Chandigarh Transport Undertaking (CTU)
9	HRTC
3	KAAC TRANSPORT
2	KSRTC (Kerala)
4	Meghalaya Transport Corporation(MTC)
6	PEPSU (Punjab)
5	RSRTC
1	Sikkim Nationalised Transport (SNT)
8	TGSRTC

```
10 rows in set (0.01 sec)
```

Routes Table:

```
CREATE TABLE routes (
route_id INT AUTO_INCREMENT PRIMARY KEY,
state_id INT,
route_name VARCHAR(255),
route_link TEXT,
UNIQUE(state_id, route_name),
FOREIGN KEY (state_id) REFERENCES states(state_id) ON DELETE CASCADE)
```

```
mysql> desc routes;
```

Field	Type	Null	Key	Default	Extra
route_id	int	NO	PRI	NULL	auto_increment
state_id	int	YES	MUL	NULL	
route_name	varchar(255)	YES		NULL	
route_link	text	YES		NULL	

```
4 rows in set (0.01 sec)
```

```
mysql> select * from routes;
```

route_id	state_id	route_name	route_link
1	1	Siliguri to Gangtok	https://www.redbus.in/bus-tickets/siliguri-to-gangtok
2	1	Gangtok to Siliguri	https://www.redbus.in/bus-tickets/gangtok-to-siliguri
3	1	Siliguri to Rangpo	https://www.redbus.in/bus-tickets/siliguri-to-rangpo
4	1	Rangpo to Siliguri	https://www.redbus.in/bus-tickets/rangpo-to-siliguri
5	1	Siliguri to Singtham (Sikkim)	https://www.redbus.in/bus-tickets/siliguri-to-singtham-sikkim
6	1	Pelling (Sikkim) to Siliguri	https://www.redbus.in/bus-tickets/pelling-sikkim-to-siliguri
7	1	Siliguri to Pelling (Sikkim)	https://www.redbus.in/bus-tickets/siliguri-to-pelling-sikkim
8	1	Singtham (Sikkim) to Siliguri	https://www.redbus.in/bus-tickets/singtham-sikkim-to-siliguri
9	1	Siliguri to Ravangla (Sikkim)	https://www.redbus.in/bus-tickets/siliguri-to-ravangla-sikkim
10	1	Siliguri to Jorethang (Sikkim)	https://www.redbus.in/bus-tickets/siliguri-to-jorethang-sikkim
11	1	Ravangla (Sikkim) to Siliguri	https://www.redbus.in/bus-tickets/ravangla-sikkim-to-siliguri
12	1	Namchi (Sikkim) to Siliguri	https://www.redbus.in/bus-tickets/namchi-sikkim-to-siliguri
13	1	Siliguri to Namchi (Sikkim)	https://www.redbus.in/bus-tickets/siliguri-to-namchi-sikkim
14	1	Jorethang (Sikkim) to Siliguri	https://www.redbus.in/bus-tickets/jorethang-sikkim-to-siliguri
15	1	Siliguri to Melli (Sikkim)	https://www.redbus.in/bus-tickets/siliguri-to-melli-sikkim
16	2	Bangalore to Kozhikode	https://www.redbus.in/bus-tickets/bangalore-to-kozhikode
17	2	Kozhikode to Ernakulam	https://www.redbus.in/bus-tickets/kozhikode-to-ernakulam
18	2	Kozhikode to Bangalore	https://www.redbus.in/bus-tickets/kozhikode-to-bangalore
19	2	Ernakulam to Kozhikode	https://www.redbus.in/bus-tickets/ernakulam-to-kozhikode
20	2	Mysore to Kozhikode	https://www.redbus.in/bus-tickets/mysore-to-kozhikode
21	2	Kozhikode to Mysore	https://www.redbus.in/bus-tickets/kozhikode-to-mysore
22	2	Kozhikode to Thrissur	https://www.redbus.in/bus-tickets/kozhikode-to-thrissur
23	2	Kozhikode to Thiruvananthapuram	https://www.redbus.in/bus-tickets/kozhikode-to-thiruvananthapuram
24	2	Kozhikode to Kottayam	https://www.redbus.in/bus-tickets/kozhikode-to-kottayam
25	2	Thrissur to Kozhikode	https://www.redbus.in/bus-tickets/thrissur-to-kozhikode
26	2	Bangalore to Malpetta (Kerala)	https://www.redbus.in/bus-tickets/bangalore-to-kalpetta
27	2	Bangalore to Mannur (Kerala)	https://www.redbus.in/bus-tickets/bangalore-to-kannur
28	2	Malpetta (Kerala) to Bangalore	https://www.redbus.in/bus-tickets/kalpetta-to-bangalore
29	2	Kottayam to Kozhikode	https://www.redbus.in/bus-tickets/kottayam-to-kozhikode
30	2	Thiruvananthapuram to Kozhikode	https://www.redbus.in/bus-tickets/thiruvananthapuram-to-kozhikode
31	2	Mannur (Kerala) to Bangalore	https://www.redbus.in/bus-tickets/kannur-to-bangalore
32	2	Kozhikode to Aluva	https://www.redbus.in/bus-tickets/kozhikode-to-aluva
33	3	Diphu to Guwahati	https://www.redbus.in/bus-tickets/diphu-to-guwahati
34	3	Bokolia (Assam) to Guwahati	https://www.redbus.in/bus-tickets/bokolia-assam-to-guwahati
35	3	Guwahati to Diphu	https://www.redbus.in/bus-tickets/guwahati-to-diphu

Bus Details table

```
CREATE TABLE bus_details (
  id INT AUTO_INCREMENT PRIMARY KEY,
  route_id INT,
  bus_name VARCHAR(255),
  bus_type VARCHAR(255),
  departing_time VARCHAR(50),
  duration VARCHAR(50),
  reaching_time VARCHAR(50),
  star_rating DECIMAL(3,2),
  price DECIMAL(10,2),
  seat_availability INT,
  FOREIGN KEY (route_id) REFERENCES routes(route_id) ON DELETE CASCADE)
```

```
mysql> desc bus_details;
```

Field	Type	Null	Key	Default	Extra
id	int	NO	PRI	NULL	auto_increment
route_id	int	YES	MUL	NULL	
bus_name	varchar(255)	YES		NULL	
bus_type	varchar(255)	YES		NULL	
departing_time	varchar(50)	YES		NULL	
duration	varchar(50)	YES		NULL	
reaching_time	varchar(50)	YES		NULL	
star_rating	decimal(3,2)	YES		NULL	
price	decimal(10,2)	YES		NULL	
seat_availability	int	YES		NULL	

10 rows in set (0.01 sec)

```
mysql> select * from bus_details;
```

id	route_id	bus_name	bus_type	departing_time	duration	reaching_time
1	1	Sikkim Nationalised Transport (SNT) - 171242	A/C Seater (2+1)	18:30	05h 00m	15:30
2	1	Sikkim Nationalised Transport (SNT) - 171242	A/C Seater (2+1)	18:30	05h 00m	15:30
3	1	Sikkim Nationalised Transport (SNT) - 171242	A/C Seater (2+1)	18:30	05h 00m	15:30
4	1	Sikkim Nationalised Transport (SNT) - 171242	A/C Seater (2+1)	18:30	05h 00m	15:30
5	1	Sikkim Nationalised Transport (SNT) - 171242	A/C Seater (2+1)	18:30	05h 00m	15:30
6	1	Sikkim Nationalised Transport (SNT) - 171242	A/C Seater (2+1)	18:30	05h 00m	15:30
7	1	Sikkim Nationalised Transport (SNT) - 171242	A/C Seater (2+1)	18:30	05h 00m	15:30
8	1	Sikkim Nationalised Transport (SNT) - 171242	A/C Seater (2+1)	18:30	05h 00m	15:30
9	1	Sikkim Nationalised Transport (SNT) - 171242	A/C Seater (2+1)	18:30	05h 00m	15:30
10	1	Sikkim Nationalised Transport (SNT) - 171242	A/C Seater (2+1)	18:30	05h 00m	15:30
11	1	Sikkim Nationalised Transport (SNT) - 171242	A/C Seater (2+1)	18:30	05h 00m	15:30
12	1	Sikkim Nationalised Transport (SNT) - 171242	A/C Seater (2+1)	18:30	05h 00m	15:30
13	1	Sikkim Nationalised Transport (SNT) - 171242	A/C Seater (2+1)	18:30	05h 00m	15:30
14	1	Sikkim Nationalised Transport (SNT) - 171242	A/C Seater (2+1)	18:30	05h 00m	15:30
15	2	Sikkim Nationalised Transport (SNT) - 178859	Non AC Seater (2+3)	18:25	07h 05m	17:30

State Transport Organizations Covered:

- Sikkim Nationalised Transport (SNT)
- KSRTC (Kerala)
- KAAC TRANSPORT
- Meghalaya Transport Corporation (MTC)
- RSRTC
- PEPSU (Punjab)
- Chandigarh Transport Undertaking (CTU)
- TGSRTC
- HRTC
- Assam State Transport Corporation (ASTC)

Data Fields Collected:

- Route Name
- Route Link
- Bus name
- Bus type
- Departure time
- Duration
- Arrival time
- Star rating (if available)
- Price (in INR)
- Seat availability

Streamlit Screenshot:

