# IT VEDANT INSTITUTE, THANE.

# MASTER IN DATA SCIENCE & ANALYTICS WITH ARTIFICIAL INTELLIGENCE


**PROJECT FOR PYTHON MODULE**

**ON**

## List Of Mumbai Suburban Railway Stations

**BY**

Madhur Ramesh Dhanawade

**UNDER THE GUIDENCE OF**

**Mr. Sameer Warsolkar**

**Academic year: 2024-2025**

# Acknowledgment

I would like to express my heartfelt gratitude to everyone who supported me during this web scraping project on Mumbai Suburban Railway stations. First and foremost, I thank my mentors and instructors for their invaluable guidance and encouragement throughout the process. A special thanks to my peers for their collaborative spirit and insights, which helped enhance my understanding. I also appreciate the open-source communities and online forums for providing the resources and knowledge essential for this project. Furthermore, I acknowledge the developers of the libraries and tools used in this project, which simplified the web scraping process. Lastly, I extend my appreciation to the Mumbai Suburban Railway authorities for maintaining accessible data that contributed to the success of this endeavour. Your support has been instrumental in the completion of this project.

Name:- Madhur Ramesh Dhanawade

Module:-Python for Data Science Module

Institute:-IT Vedant

# Abstraction

This project aims to create a comprehensive list of Mumbai Suburban Railway stations through web scraping techniques. Utilizing Python libraries such as Beautiful Soup and Requests, we extracted data from relevant railway websites. The focus was on gathering essential information, including station names, codes, and locations.

The collected data is structured into a CSV format for easy accessibility and analysis. By automating the data retrieval process, we ensure up-to-date information is readily available for users. This project demonstrates the effectiveness of web scraping in collecting large datasets efficiently. Furthermore, it highlights the importance of accessible public transport information in urban planning and commuter navigation. The final dataset serves as a valuable resource for developers, researchers, and commuters alike.

**Keywords:**

Web Scraping
Python
Beautiful Soup
Data Extraction
E-commerce

# TABLE OF CONTENTS

# INTRODUCTION

Mumbai, known as the financial capital of India, boasts one of the most extensive suburban railway networks in the world. The Mumbai Suburban Railway plays a crucial role in the daily commute of millions, connecting various neighbourhoods and facilitating economic activities. However, accessing updated and comprehensive information about the railway stations can be challenging. This project aims to address that gap by employing web scraping techniques to compile a detailed list of Mumbai Suburban Railway stations.

Utilizing tools such as Python's Beautiful Soup and Requests library, we will extract key data, including station names, codes, and locations. This effort not only enhances data accessibility for commuters but also serves as a valuable resource for developers and researchers interested in transportation systems. Through this project, we demonstrate the power of automation in gathering real-time data to support informed decision-making in urban transit.

# List Of Mumbai Suburban Railway Stations Description.

This project involves the development of a web scraping tool to gather information about the stations in the Mumbai Suburban Railway network. Using Python, we employ libraries such as Beautiful Soup and Requests to extract relevant data from official railway websites. The primary focus is on collecting essential details, including station names, codes, locations, and operational information.

The scraped data is processed and organized into a structured format, such as CSV, for easy access and analysis. This initiative not only aims to provide commuters with up-to-date station information but also serves as a foundation for further research in transportation analytics. Additionally, the project showcases the potential of web scraping as a valuable technique for data collection in the digital age. Ultimately, it contributes to enhancing public transportation awareness and accessibility in Mumbai.

## Outline:

From this Website we are going to grab the following details.

- Railway Stations Names
- Railway Stations Names (Marathi)
- Railway line (Central,Westren,Harber)
- Fast_Train_Halt (Yes / No)
- Long-distance (Indicates that trains other than local trains halt at this station. These trains are usually part of the Indian Railways network.)

# METHODOLOGY

**Tools and Technologies Used:**

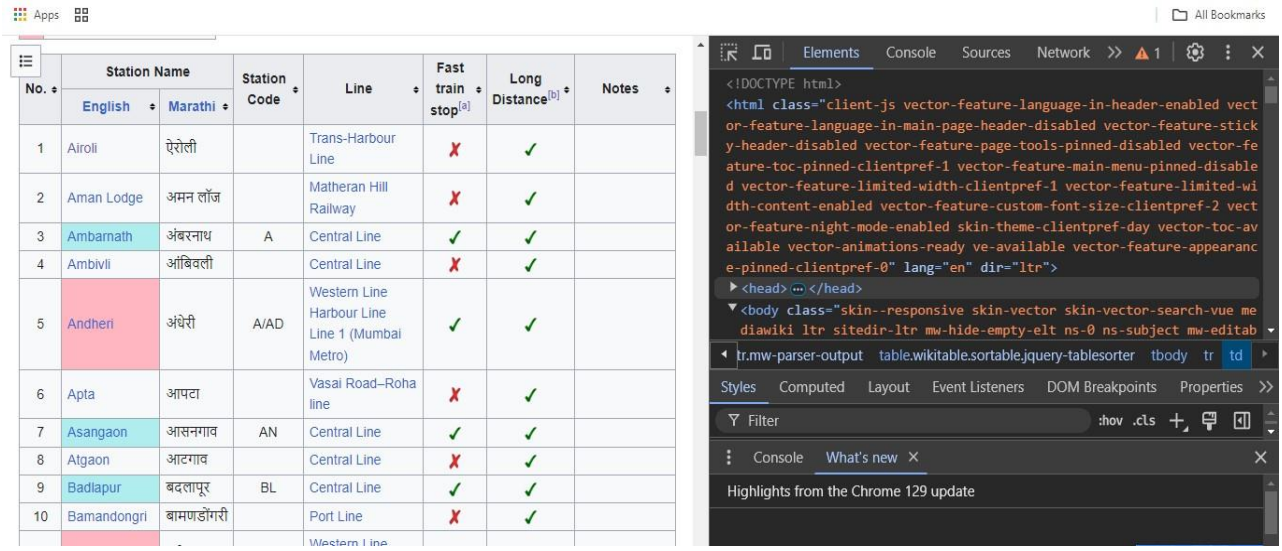
- Python: Programming language
- BeautifulSoup: Python library used for web scraping purposes to extract data from HTML and XML files
- Requests: Requests is a Python library used for sending HTTP requests to websites and web services.
- Pandas: it provides data structures and functions needed to work with structured data, particularly data that is stored in tables or time series.

**Steps Involved in Web Scraping:**

- Import the libraries need to access the website
- Accessing the website.
- Inspecting the HTML structure of List Of Mumbai Suburban Railway Stations Wikipedia pages.
- Extracting product details (RailwayStation Names,Railway Line,Fast Train Halt,Long Distance).
- Storing the extracted data in a structured format (e.g., CSV or DataFrame).
- Extract the data for the analysis.

# CODE EXPLANATION

## Step 1: Choose the Website and Webpage URL
## Inspect the Website



| No. | Station Name | | Station Code | Line | Fast train stop[a] | Long Distance[b] | Notes |
|---|---|---|---|---|---|---|---|
| | English | Marathi | | | | | |
| 1 | Airoli | ऐरोली | | Trans-Harbour Line | ✗ | ✓ | |
| 2 | Aman Lodge | अमन लॉज | | Matheran Hill Railway | ✗ | ✓ | |
| 3 | Ambarnath | अंबरनाथ | A | Central Line | ✓ | ✓ | |
| 4 | Ambivli | आंबिवली | | Central Line | ✗ | ✓ | |
| 5 | Andheri | अंधेरी | A/AD | Western Line Harbour Line Line 1 (Mumbai Metro) | ✓ | ✓ | |
| 6 | Apta | आपटा | | Vasai Road–Roha line | ✗ | ✓ | |
| 7 | Asangaon | आसनगाव | AN | Central Line | ✓ | ✓ | |
| 8 | Atgaon | आटगाव | | Central Line | ✗ | ✓ | |
| 9 | Badlapur | बदलापूर | BL | Central Line | ✓ | ✓ | |
| 10 | Bamandongri | बामणडोंगरी | | Port Line | ✗ | ✓ | |

## Step 2: Installing or importing the essential libraries for scrapping process.

- requests
- BeautifulSoup
- pandas

```
[2]: import requests
     from bs4 import BeautifulSoup
     import pandas as pd
```

## Step 3: using requests library, requests the website for providing data

```
[2]: url = "https://en.wikipedia.org/wiki/List_of_Mumbai_Suburban_Railway_stations"
     response = requests.get(url)
     response

[2]: <Response [200]>
```

## Step 4: using BeautifulSoup library import the HTML content of the website

```
[3]: soup = BeautifulSoup(response.content, "html.parser")
     soup

[3]: <!DOCTYPE html>

     <html class="client-nojs vector-feature-language-in-header-enabled vector-feature-language-in-main-page-header-disabled vector-feature-sticky-header-di
     sabled vector-feature-page-tools-pinned-disabled vector-feature-toc-pinned-clientpref-1 vector-feature-main-menu-pinned-disabled vector-feature-limited
     -width-clientpref-1 vector-feature-limited-width-content-enabled vector-feature-custom-font-size-clientpref-1 vector-feature-appearance-pinned-clientpr
     ef-1 vector-feature-night-mode-enabled skin-theme-clientpref-day vector-toc-available" dir="ltr" lang="en">
     <head>
     <meta charset="utf-8"/>
     <title>List of Mumbai Suburban Railway stations - Wikipedia</title>
     <script>(function(){var className="client-js vector-feature-language-in-header-enabled vector-feature-language-in-main-page-header-disabled vector-feat
     ure-sticky-header-disabled vector-feature-page-tools-pinned-disabled vector-feature-toc-pinned-clientpref-1 vector-feature-main-menu-pinned-disabled ve
     ctor-feature-limited-width-clientpref-1 vector-feature-limited-width-content-enabled vector-feature-custom-font-size-clientpref-1 vector-feature-appear
     ance-pinned-clientpref-1 vector-feature-night-mode-enabled skin-theme-clientpref-day vector-toc-available";var cookie=document.cookie.match(/(?:^|; )en
     wikimwclientpreferences=([^;]+)/);if(cookie){cookie[1].split('%2C').forEach(function(pref){className=className.replace(new RegExp('(^| )'+pref.replace
     (/-clientpref-\w+$|[^\w-]+/g,'')+'-clientpref-\\w+( |$)'),'$1'+pref+'$2');});}document.documentElement.className=className;}());RLCONF={"wgBreakFrame
     s":false,"wgSeparatorTransformTable":["",""],"wgDigitTransformTable":["",""],"wgDefaultDateFormat":"dmy",
     "wgMonthNames":["","January","February","March","April","May","June","July","August","September","October","November","December"],"wgRequestId":"2089c9
     e0-41d3-4e84-af80-ebd0a07404e3","wgCanonicalNamespace":"","wgCanonicalSpecialPageName":false,"wgNamespaceNumber":0,"wgPageName":"List of Mumbai Suburba
```

## Step 5: Find the “td” elements.

```
[8]: td_elements = soup.find_all("td")
     td_elements    # td_elements

[8]: [<td style="background:#afeeee;">†
      </td>,
      <td>Terminal station
      </td>,
      <td style="background:#ff9;">*
      </td>,
      <td>Transfer station (excluding transfer to <a href="/wiki/Indian_Railways" title="Indian Railways">Indian Railways</a>)
      </td>,
      <td style="background:#ffb6c1;">†*
      </td>,
      <td>Terminal and transfer station to other lines
      </td>,
      <td style="text-align:center;">1</td>,
      <td><a href="/wiki/Airoli_railway_station" title="Airoli railway station">Airoli</a></td>,
      <td>ऐरोली</td>,
      <td>
      </td>,
```

.

## Step 6: Extract text from each td element.

```
[10]: # Extract text from each td element
      list_general=[]
      for td in td_elements:
          text = td.get_text()
          list_general.append(text)
      list_general        # list_general
```

```
[10]: ['†\n',
       'Terminal station\n',
       '*\n',
       'Transfer station (excluding transfer to Indian Railways)\n',
       '†*\n',
       'Terminal and transfer station to other lines\n',
       '1',
       'Airoli',
       'ऐरोली',
       '\n',
       'Trans-Harbour Line',
       'N',
       'Y',
       '\n',
       '2',
       'Aman Lodge',
       'अमन लॉज',
```

## Step 7: Search first row text,last row text and search there index.

```
[11]: search_first_row_text="Airoli"
      search_last_row_text="Targhar"
      print(search_first_row_text)
      print(search_last_row_text)
```

```
Airoli
Targhar
```

```
[12]: search_index_first_row=list_general.index(search_first_row_text)
      search_index_last_row=list_general.index(search_last_row_text)

      print("The index of the'Airoli'in above list is: ",search_index_first_row )
      print("The index of the'Targhar' in the above list is: ",search_index_last_row)
```

```
The index of the'Airoli'in above list is:  7
The index of the'Targhar' in the above list is:  1199
```

## Step 8: Accessing the Railway Stations Names.

```
[28]: name=[]
      for i in range(1,len(new_list),8):
          name.append(new_list[i])
      print(name)
```

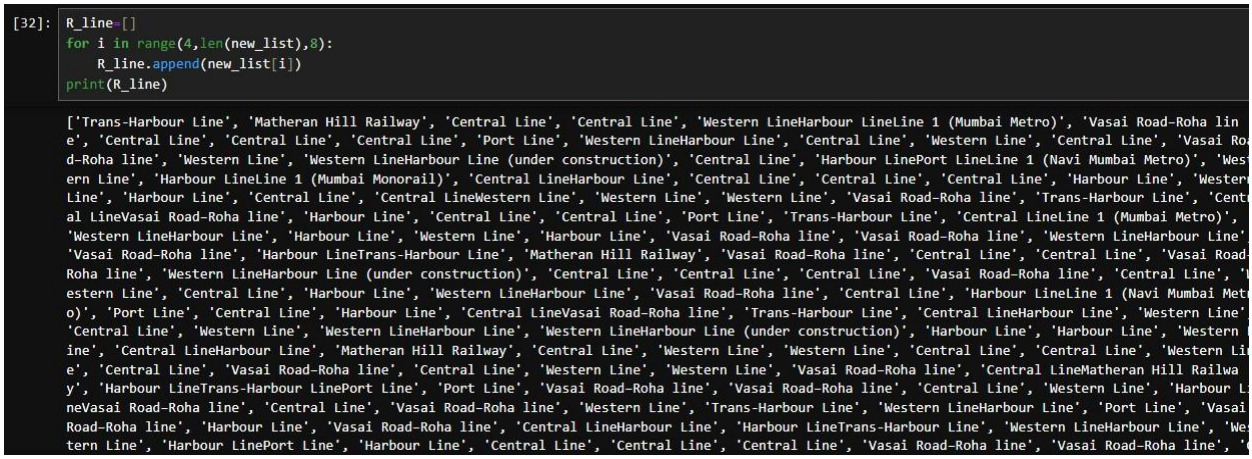
```
['Airoli', 'Aman Lodge', 'Ambarnath', 'Ambivli', 'Andheri', 'Apta', 'Asangaon', 'Atgaon', 'Badlapur', 'Bamandongri', 'Bandra', 'Bhandup', 'Bhayandar', 'B
hivpuri Road', 'Bhiwandi Road', 'Boisar', 'Borivali', 'Byculla', 'CBD Belapur', 'Charni Road', 'Chembur', 'Chhatrapati Shivaji Maharaj Terminus', 'Chikha
le', 'Chinchpokli', 'Chouk', 'Chunabhatti', 'Churchgate', 'Cotton Green', 'Currey Road', 'Dadar', 'Dahanu Road', 'Dahisar', 'Dativali', 'Digha Gaon', 'Di
va Junction', 'Dockyard Road', 'Dolavli', 'Dombivli', 'Dronagiri', 'Ghansoli', 'Ghatkopar', 'Goregaon', 'Govandi', 'Grant Road', 'Guru Tegh Bahadur Naga
r', 'Hamrapur', 'Jite', 'Jogeshwari', 'Juchandra', 'Juinagar', 'Jummapatti', 'Kalamboli', 'Kalwa', 'Kalyan Junction', 'Kaman Road', 'Kandivli', 'Kanjur M
arg', 'Karjat', 'Kasara', 'Kasu', 'Kelavli', 'Kelve Road', 'Khadavli', 'Khandeshwar', 'Khar Road', 'Kharbav', 'Khardi', 'Kharghar', 'Kharkopar', 'Khopol
i', "King's Circle", 'Kopar', 'Koparkhairane', 'Kurla', 'Lower Parel', 'Lowjee', 'Mahalaxmi', 'Mahim Junction', 'Malad', 'Mankhurd', 'Mansarovar', 'Marin
e Lines', 'Masjid', 'Matheran', 'Matunga', 'Matunga Road', 'Mira Road', 'Mohope', 'Mulund', 'Mumbai Central', 'Mumbra', 'Nagothane', 'Nahur', 'Naigaon',
'Nallasopara', 'Navde Road', 'Neral Junction', 'Nerul', 'Nhava Sheva', 'Nidi', 'Nilaje', 'Palasdari', 'Palghar', 'Panvel', 'Parel', 'Pen', 'Prabhadevi',
'Rabale', 'Ram Mandir', 'Ranjanpada', 'Rasayani', 'Reay Road', 'Roha', 'Sandhurst Road', 'Sanpada', 'Santacruz', 'Saphale', 'Seawoods-Darave', 'Sewri',
'Shahad', 'Shelu', 'Sion', 'Somatne', 'Taloje Panchnand', 'Thakurli', 'Thane', 'Thansit', 'Tilak Nagar', 'Titwala', 'Turbhe', 'Ulhasnagar', 'Umbermali',
'Umroli', 'Uran', 'Vaitarna', 'Vangani', 'Vangaon', 'Vasai Road', 'Vashi', 'Vasind', 'Vidyavihar', 'Vikhroli', 'Vile Parle', 'Virar', 'Vithalwadi', 'Wada
la Road', 'Water Pipe', 'Gavan', 'Sagar Sangam', 'Targhar']
```

## Step 9: Accessing the Railway Stations Names In Marathi.

```
[28]: name=[]
      for i in range(1,len(new_list),8):
          name.append(new_list[i])
      print(name)
```

```
['Airoli', 'Aman Lodge', 'Ambarnath', 'Ambivli', 'Andheri', 'Apta', 'Asangaon', 'Atgaon', 'Badlapur', 'Bamandongri', 'Bandra', 'Bhandup', 'Bhayandar', 'B
hivpuri Road', 'Bhiwandi Road', 'Boisar', 'Borivali', 'Byculla', 'CBD Belapur', 'Charni Road', 'Chembur', 'Chhatrapati Shivaji Maharaj Terminus', 'Chikha
le', 'Chinchpokli', 'Chouk', 'Chunabhatti', 'Churchgate', 'Cotton Green', 'Currey Road', 'Dadar', 'Dahanu Road', 'Dahisar', 'Dativali', 'Digha Gaon', 'Di
va Junction', 'Dockyard Road', 'Dolavli', 'Dombivli', 'Dronagiri', 'Ghansoli', 'Ghatkopar', 'Goregaon', 'Govandi', 'Grant Road', 'Guru Tegh Bahadur Naga
r', 'Hamrapur', 'Jite', 'Jogeshwari', 'Juchandra', 'Juinagar', 'Jummapatti', 'Kalamboli', 'Kalwa', 'Kalyan Junction', 'Kaman Road', 'Kandivli', 'Kanjur M
arg', 'Karjat', 'Kasara', 'Kasu', 'Kelavli', 'Kelve Road', 'Khadavli', 'Khandeshwar', 'Khar Road', 'Kharbav', 'Khardi', 'Kharghar', 'Kharkopar', 'Khopol
i', "King's Circle", 'Kopar', 'Koparkhairane', 'Kurla', 'Lower Parel', 'Lowjee', 'Mahalaxmi', 'Mahim Junction', 'Malad', 'Mankhurd', 'Mansarovar', 'Marin
e Lines', 'Masjid', 'Matheran', 'Matunga', 'Matunga Road', 'Mira Road', 'Mohope', 'Mulund', 'Mumbai Central', 'Mumbra', 'Nagothane', 'Nahur', 'Naigaon',
'Nallasopara', 'Navde Road', 'Neral Junction', 'Nerul', 'Nhava Sheva', 'Nidi', 'Nilaje', 'Palasdari', 'Palghar', 'Panvel', 'Parel', 'Pen', 'Prabhadevi',
'Rabale', 'Ram Mandir', 'Ranjanpada', 'Rasayani', 'Reay Road', 'Roha', 'Sandhurst Road', 'Sanpada', 'Santacruz', 'Saphale', 'Seawoods-Darave', 'Sewri',
'Shahad', 'Shelu', 'Sion', 'Somatne', 'Taloje Panchnand', 'Thakurli', 'Thane', 'Thansit', 'Tilak Nagar', 'Titwala', 'Turbhe', 'Ulhasnagar', 'Umbermali',
'Umroli', 'Uran', 'Vaitarna', 'Vangani', 'Vangaon', 'Vasai Road', 'Vashi', 'Vasind', 'Vidyavihar', 'Vikhroli', 'Vile Parle', 'Virar', 'Vithalwadi', 'Wada
la Road', 'Water Pipe', 'Gavan', 'Sagar Sangam', 'Targhar']
```

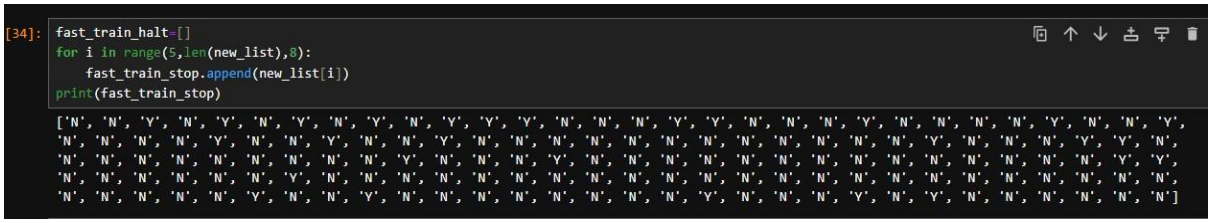
## Step 10: Accessing the Railway Station Line.



```
[32]: R_line=[]
      for i in range(4,len(new_list),8):
          R_line.append(new_list[i])
      print(R_line)
```

```
['Trans-Harbour Line', 'Matheran Hill Railway', 'Central Line', 'Central Line', 'Western LineHarbour LineLine 1 (Mumbai Metro)', 'Vasai Road-Roha lin
e', 'Central Line', 'Central Line', 'Central Line', 'Port Line', 'Western LineHarbour Line', 'Central Line', 'Western Line', 'Central Line', 'Vasai Ro
d-Roha line', 'Western Line', 'Western LineHarbour Line (under construction)', 'Central Line', 'Harbour LinePort LineLine 1 (Navi Mumbai Metro)', 'Wes
ern Line', 'Harbour LineLine 1 (Mumbai Monorail)', 'Central LineHarbour Line', 'Central Line', 'Central Line', 'Central Line', 'Harbour Line', 'Wester
Line', 'Harbour Line', 'Central Line', 'Central LineWestern Line', 'Western Line', 'Western Line', 'Vasai Road-Roha line', 'Trans-Harbour Line', 'Cent
al LineVasai Road-Roha line', 'Harbour Line', 'Central Line', 'Central Line', 'Port Line', 'Trans-Harbour Line', 'Central LineLine 1 (Mumbai Metro)',
'Western LineHarbour Line', 'Harbour Line', 'Western Line', 'Harbour Line', 'Vasai Road-Roha line', 'Vasai Road-Roha line', 'Western LineHarbour Line'
'Vasai Road-Roha line', 'Harbour LineTrans-Harbour Line', 'Matheran Hill Railway', 'Vasai Road-Roha line', 'Central Line', 'Central Line', 'Vasai Road
Roha line', 'Western LineHarbour Line (under construction)', 'Central Line', 'Central Line', 'Central Line', 'Vasai Road-Roha line', 'Central Line', '
estern Line', 'Central Line', 'Harbour Line', 'Western LineHarbour Line', 'Vasai Road-Roha line', 'Central Line', 'Harbour LineLine 1 (Navi Mumbai Met
o)', 'Port Line', 'Central Line', 'Harbour Line', 'Central LineVasai Road-Roha line', 'Trans-Harbour Line', 'Central LineHarbour Line', 'Western Line'
'Central Line', 'Western Line', 'Western LineHarbour Line', 'Western LineHarbour Line (under construction)', 'Harbour Line', 'Harbour Line', 'Western
ine', 'Central LineHarbour Line', 'Matheran Hill Railway', 'Central Line', 'Western Line', 'Western Line', 'Central Line', 'Central Line', 'Western Li
e', 'Central Line', 'Vasai Road-Roha line', 'Central Line', 'Western Line', 'Western Line', 'Vasai Road-Roha line', 'Central LineMatheran Hill Railwa
y', 'Harbour LineTrans-Harbour LinePort Line', 'Port Line', 'Vasai Road-Roha line', 'Vasai Road-Roha line', 'Central Line', 'Western Line', 'Harbour L
neVasai Road-Roha line', 'Central Line', 'Vasai Road-Roha line', 'Western Line', 'Trans-Harbour Line', 'Western LineHarbour Line', 'Port Line', 'Vasai
Road-Roha line', 'Harbour Line', 'Vasai Road-Roha line', 'Central LineHarbour Line', 'Harbour LineTrans-Harbour Line', 'Western LineHarbour Line', 'We
tern Line', 'Harbour LinePort Line', 'Harbour Line', 'Central Line', 'Central Line', 'Central Line', 'Vasai Road-Roha line', 'Vasai Road-Roha line', '
```

## Step 11: Accessing the Fast Train Halt (Y/N).



```
[34]: fast_train_halt=[]
      for i in range(5,len(new_list),8):
          fast_train_stop.append(new_list[i])
      print(fast_train_stop)
```

```
['N', 'N', 'Y', 'N', 'Y', 'N', 'Y', 'N', 'Y', 'N', 'Y', 'Y', 'Y', 'N', 'N', 'N', 'Y', 'Y', 'N', 'N', 'N', 'Y', 'N', 'N', 'N', 'N', 'Y', 'N', 'N', 'Y',
'N', 'N', 'N', 'N', 'Y', 'N', 'N', 'Y', 'N', 'N', 'Y', 'N', 'N', 'N', 'N', 'N', 'N', 'N', 'N', 'N', 'N', 'Y', 'N', 'N', 'N', 'Y', 'Y', 'N',
'N', 'N', 'N', 'N', 'N', 'N', 'N', 'N', 'N', 'Y', 'N', 'N', 'N', 'Y', 'N', 'N', 'N', 'N', 'N', 'N', 'N', 'N', 'N', 'N', 'N', 'N', 'N', 'N', 'Y', 'Y',
'N', 'N', 'N', 'N', 'N', 'N', 'Y', 'N', 'N', 'N', 'N', 'N', 'N', 'N', 'N', 'N', 'N', 'N', 'N', 'N', 'N', 'N', 'N', 'N', 'N', 'N', 'N', 'N', 'N', 'N',
'N', 'N', 'N', 'N', 'N', 'Y', 'N', 'N', 'Y', 'N', 'N', 'N', 'N', 'N', 'N', 'N', 'N', 'Y', 'N', 'N', 'N', 'Y', 'N', 'Y', 'N', 'N', 'N', 'N', 'N']
```

## Step 12: Accessing the Long Distance Halt. (Y/N).

```
41]: long_distance=[]
     for i in range(6,len(new_list),8):
         long_distance.append(new_list[i])
     print(long_distance)
```

```
['Y', 'Y', 'Y', 'Y', 'Y', 'Y', 'Y', 'Y', 'Y', 'N', 'N', 'Y', 'Y', 'Y', 'Y', 'N', 'N', 'Y', 'N', 'N', 'N', 'Y', 'N', 'Y', 'N', 'N', 'N', 'N', 'Y',
 'Y', 'N', 'Y', 'Y', 'Y', 'N', 'Y', 'Y', 'Y', 'Y', 'N', 'N', 'N', 'N', 'N', 'Y', 'Y', 'N', 'Y', 'Y', 'Y', 'Y', 'Y', 'Y', 'N', 'N', 'Y', 'Y', 'Y',
 'Y', 'Y', 'Y', 'Y', 'N', 'Y', 'Y', 'Y', 'Y', 'Y', 'N', 'Y', 'Y', 'N', 'N', 'Y', 'N', 'N', 'N', 'N', 'Y', 'N', 'N', 'Y', 'N', 'N', 'Y', 'Y', 'N', 'Y',
 'Y', 'Y', 'N', 'Y', 'Y', 'Y', 'Y', 'Y', 'Y', 'Y', 'Y', 'Y', 'N', 'Y', 'N', 'Y', 'N', 'Y', 'N', 'Y', 'N', 'Y', 'N', 'Y', 'Y', 'N', 'Y',
 'Y', 'N', 'Y', 'Y', 'Y', 'Y', 'Y', 'N', 'Y', 'Y', 'Y', 'Y', 'Y', 'Y', 'Y', 'Y', 'Y', 'Y', 'Y', 'Y', 'Y', 'Y', 'N', 'N', 'N', 'Y', 'Y', 'N', 'Y', 'Y', 'Y', 'Y']
```

## Step 13: using pandas library make a data frame of the dictionary you have created.

```
[43]: import pandas as pd

      df=pd.DataFrame({"Railway_station_Name":name,"Name_in_marathi":name_marathi,"Railway_line":R_line,"Fast_Train_Halt":fast_train_stop,"Long_Distance":long_
      df
```

[43]:

| | Railway_station_Name | Name_in_marathi | Railway_line | Fast_Train_Halt | Long_Distance |
|---|---|---|---|---|---|
| 0 | Airoli | ऐरोली | Trans-Harbour Line | N | Y |
| 1 | Aman Lodge | अमन लॉज | Matheran Hill Railway | N | Y |
| 2 | Ambarnath | अंबरनाथ | Central Line | Y | Y |
| 3 | Ambivli | आंबिवली | Central Line | N | Y |
| 4 | Andheri | अंधेरी | Western LineHarbour LineLine 1 (Mumbai Metro) | Y | Y |
| ... | ... | ... | ... | ... | ... |
| 145 | Wadala Road | वडाळा रोड | Harbour LineLine 1 (Mumbai Monorail) | N | N |
| 146 | Water Pipe | वॉटर पाईप | Matheran Hill Railway | N | Y |
| 147 | Gavan | गव्हाण | Port Line | N | Y |
| 148 | Sagar Sangam | सागरसंगम | Port Line | N | Y |

## Step 14: save the file to csv mode in your computer and read that file.

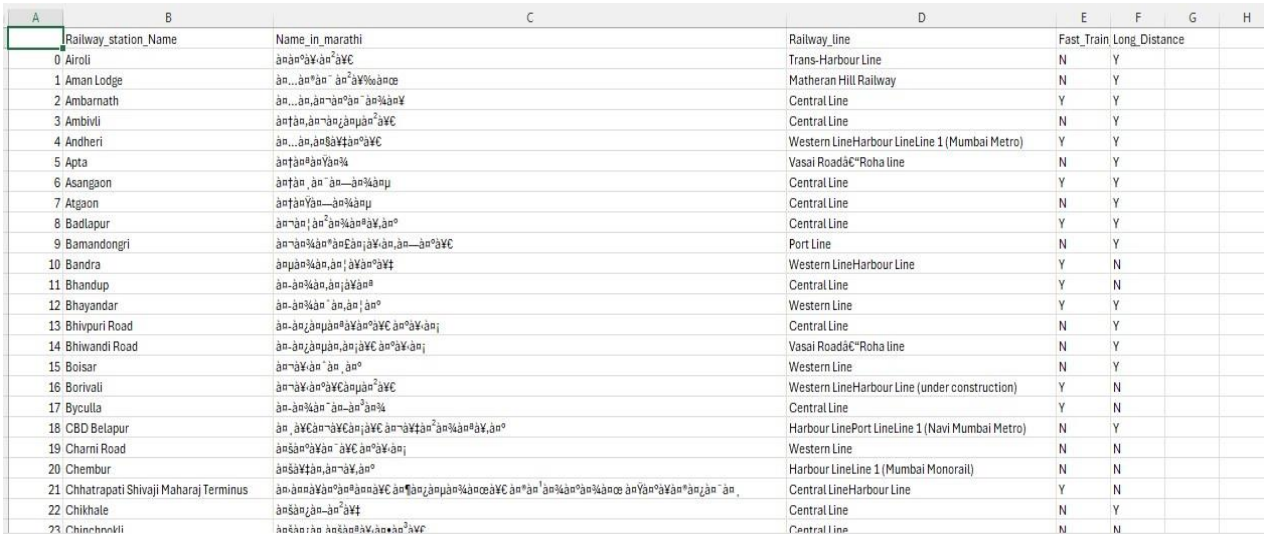
```
[47]: df.to_csv("List of Mumbai Suburban Railway stations.csv")

[48]: pwd()

[48]: 'C:\\Users\\HP\\Desktop\\IT vedant\\python\\Web scrapping'

[49]: read=pd.read_csv("List of Mumbai Suburban Railway stations.csv")
      read
```

[49]:

| | Unnamed: 0 | Railway_station_Name | Name_in_marathi | Railway_line | Fast_Train_Halt | Long_Distance |
|---|---|---|---|---|---|---|
| 0 | 0 | Airoli | ऐरोली | Trans-Harbour Line | N | Y |
| 1 | 1 | Aman Lodge | अमन लॉज | Matheran Hill Railway | N | Y |
| 2 | 2 | Ambarnath | अंबरनाथ | Central Line | Y | Y |
| 3 | 3 | Ambivli | आंबिवली | Central Line | N | Y |
| 4 | 4 | Andheri | अंधेरी | Western LineHarbour LineLine 1 (Mumbai Metro) | Y | Y |
| ... | ... | ... | ... | ... | ... | ... |
| 145 | 145 | Wadala Road | वडाळा रोड | Harbour LineLine 1 (Mumbai Monorail) | N | N |
| 146 | 146 | Water Pipe | वॉटर पाईप | Matheran Hill Railway | N | Y |

# RESULT AND ANALYSIS.

**check the file by opening from the computer into MS excel.**



| | Railway_station_Name | Name_in_marathi | Railway_line | Fast_Train | Long_Distance |
|---|---|---|---|---|---|
| 0 | Airoli | à¤à¤°à¥‹à¤²à¥€ | Trans-Harbour Line | N | Y |
| 1 | Aman Lodge | à¤…à¤®à¤¨ à¤²à¥‰à¤œ | Matheran Hill Railway | N | Y |
| 2 | Ambarnath | à¤…à¤¨à¤¾à¤¥ à¤¸à¤¥ | Central Line | Y | Y |
| 3 | Ambivli | à¤†à¤‚à¤¬à¤¿à¤µà¤²à¥€ | Central Line | N | Y |
| 4 | Andheri | à¤…à¤‚à§à¥‡à¤°à¥€ | Western LineHarbour LineLine 1 (Mumbai Metro) | Y | Y |
| 5 | Apta | à¤†à¤ªà¥Yà¤¾ | Vasai Roadâ€“Roha line | N | Y |
| 6 | Asangaon | à¤†à¤¸à¤—à¤¾à¤µ | Central Line | Y | Y |
| 7 | Atgaon | à¤†à¤Yà¤—à¤¾à¤µ | Central Line | N | Y |
| 8 | Badlapur | à¤¬à¤¦à¤²à¤¾à¤ªà¥‚à¤° | Central Line | Y | Y |
| 9 | Bamandongri | à¤¬à¤¾à¤®à¤£à¤¡à¥‹à¤‚à¤—à¤°à¥€ | Port Line | N | Y |
| 10 | Bandra | à¤µà¤¾à¤‚à¦à¥à¤°à¥‡ | Western LineHarbour Line | Y | N |
| 11 | Bhandup | à¤-à¤¾à¤‚à¤¡à¥à¤ª | Central Line | Y | N |
| 12 | Bhayandar | à¤-à¤¾à¤¯à¤‚à¦à¤° | Western Line | Y | Y |
| 13 | Bhivpuri Road | à¤-à¤¿à¤µà¤ªà¥à¤°à¥€ à¤°à¥-à¤¡ | Central Line | N | Y |
| 14 | Bhiwandi Road | à¤-à¤¿à¤µà¤‚à¤¡à¥€ à¤°à¥-à¤¡ | Vasai Roadâ€“Roha line | N | Y |
| 15 | Boisar | à¤¬à¥-à¤¯à¤¸à¤° | Western Line | N | Y |
| 16 | Borivali | à¤¬à¥-à¤°à¥€à¤µà¤²à¥€ | Western LineHarbour Line (under construction) | Y | N |
| 17 | Byculla | à¤-à¤¾à¤¯à¤–à¤³à¤¾ | Central Line | Y | N |
| 18 | CBD Belapur | à¤¸à¥€à¤¬à¥€à¤¡à¥€ à¤¬à¥‡à¤²à¤²à¤¾à¤ªà¥‚à¤° | Harbour LinePort LineLine 1 (Navi Mumbai Metro) | N | Y |
| 19 | Charni Road | à¤šà¤°à¥à¤£à¥€ à¤°à¥-à¤¡ | Western Line | N | N |
| 20 | Chembur | à¤šà¥‡à¤‚à¤¬à¥‚à¤° | Harbour LineLine 1 (Mumbai Monorail) | N | N |
| 21 | Chhatrapati Shivaji Maharaj Terminus | à¤›à¤¤à¥à¤°à¤ªà¤¤à¥€ à¤¶à¤¿à¤µà¤¾à¤œà¥€ à¤®à¤¹à¤¾à¤°à¤¾à¤œ à¤Ÿà¤°à¥à¤®à¤¿à¤¨à¤¸ | Central LineHarbour Line | Y | N |
| 22 | Chikhale | à¤šà¤¿à¤–à¤²à¥‡ | Central Line | N | Y |
| 23 | Chinchpokli | à¤šà¤¿à¤‚à¤šà¤ªà¥‹à¤•à¤³à¥€ | Central Line | N | N |

# CONCLUSION

The web scraping project successfully compiled a comprehensive list of Mumbai Suburban Railway stations, showcasing the effectiveness of automated data collection methods. By utilizing Python and relevant libraries, we were able to extract and organize crucial information that enhances accessibility for commuters and researchers alike. The analysis of the collected data provided insights into the network's structure and operational dynamics, highlighting the importance of efficient public transportation systems in urban environments. This project not only serves as a valuable resource but also emphasizes the potential of web scraping in addressing data gaps in real-time. Future enhancements could include integrating live updates on station facilities and service disruptions, further enriching the dataset. Overall, this initiative contributes to a deeper understanding of Mumbai’s transportation landscape and supports informed decision-making in urban planning.