**Configure:**

|  |
| --- |
| # import module |
|  |

|  |
| --- |
| import requests |
|  |

|  |
| --- |
| from bs4 import BeautifulSoup |
|  |

|  |
| --- |
| import pandas as pd |
|  |

|  |
| --- |
|  |
|  |

|  |
| --- |
| # user define function |
|  |

|  |
| --- |
| # Scrape the data |
|  |

|  |
| --- |
| def getdata(url): |
|  |

|  |
| --- |
| r = requests.get(url) |
|  |

|  |
| --- |
| return r.text |
|  |

|  |
| --- |
|  |
|  |

|  |
| --- |
| # input by geek |
|  |

|  |
| --- |
| train\_name = "03391-rajgir-new-delhi-clone-special-rgd-to-ndls" |
|  |

|  |
| --- |
|  |
|  |

|  |
| --- |
| # url |
|  |

|  |
| --- |
| url = "https://www.railyatri.in/live-train-status/"+train\_name |
|  |

|  |
| --- |
|  |
|  |

|  |
| --- |
| # pass the url |
|  |

|  |
| --- |
| # into getdata function |
|  |

|  |
| --- |
| htmldata = getdata(url) |
|  |

|  |
| --- |
| soup = BeautifulSoup(htmldata, 'html.parser') |
|  |

|  |
| --- |
|  |
|  |

|  |
| --- |
| # traverse the live status from |
|  |

|  |
| --- |
| # this Html code |
|  |

|  |
| --- |
| data = [] |
|  |

|  |
| --- |
| for item in soup.find\_all('script', type="application/ld+json"): |
|  |

|  |
| --- |
| data.append(item.get\_text()) |
|  |

|  |
| --- |
|  |
|  |

|  |
| --- |
| # convert into dataframe |
|  |

|  |
| --- |
| df = pd.read\_json(data[2]) |
|  |

|  |
| --- |
|  |
|  |

|  |
| --- |
| # display this column of |
|  |

|  |
| --- |
| # dataframe |
|  |

|  |
| --- |
| print(df["mainEntity"][0]['name']) |
|  |

print(df["mainEntity"][0]['acceptedAnswer']['text'])