**Registered Projects**

To scrape the first 6 projects under the "Registered Projects" heading from the provided URL and extract the fields GSTIN No, PAN No, Name, and Permanent Address, you can use the following **Python program:**

import requests

from bs4 import BeautifulSoup

import pandas as pd

from requests.packages.urllib3.exceptions import InsecureRequestWarning

requests.packages.urllib3.disable\_warnings(InsecureRequestWarning)

def get\_project\_links(base\_url):

response = requests.get(base\_url, verify=False)

soup = BeautifulSoup(response.text, 'html.parser')

print(soup.prettify())

table = soup.find('table', {'id': 'example1'})

if not table:

print("Table not found")

return []

links = []

for row in table.find\_all('tr')[1:7]:

cols = row.find\_all('td')

if cols:

link = cols[0].find('a')['href']

links.append(f"{base\_url}/{link}")

return links

def extract\_project\_details(link):

response = requests.get(link, verify=False)

soup = BeautifulSoup(response.text, 'html.parser')

details = {}

details['GSTIN No'] = soup.find(text='GSTIN No:').find\_next('td').text.strip()

details['PAN No'] = soup.find(text='PAN No:').find\_next('td').text.strip()

details['Name'] = soup.find(text='Name of Promoter:').find\_next('td').text.strip()

details['Permanent Address'] = soup.find(text='Permanent Address:').find\_next('td').text.strip()

return details

def scrape\_hprera(base\_url):

project\_links = get\_project\_links(base\_url)

all\_projects = []

for link in project\_links:

details = extract\_project\_details(link)

all\_projects.append(details)

df = pd.DataFrame(all\_projects)

return df

if \_name\_ == "\_main\_":

base\_url = "https://hprera.nic.in/PublicDashboard"

df = scrape\_hprera(base\_url)

print(df)

df.to\_csv("projects.csv", index=False)

**Pip Install requirements:**

pip install requests beautifulsoup4 pandas