

# Title: Zone and Region-wise Data Extraction from FCI Website using Selenium

A Web Scraping Project Presented by: Punit Upadhye



# Objective

- Extract structured data from the Food Corporation of India (FCI) website
- Automate the process using Selenium and BeautifulSoup
- Save data for all zone and region combinations into a structured DataFrame

# Tools and Libraries

- Selenium (for automation)
- BeautifulSoup (for HTML parsing)
- Pandas (for data storage and analysis)
- Chrome WebDriver

# Workflow Overview

- Launch FCI website
- Click "Translate" to switch to English
- Loop through zone-region combinations
- Extract data using BeautifulSoup
- Store in Pandas DataFrame

# Setup

- Import libraries
- Set up Chrome WebDriver path
- Launch browser and open FCI website
- Click the translate button to enable English

# Zone-Region Loop

- Defined all zone-region pairs manually
- Select zone using dropdown
- Wait for region options to load
- Select region



# Data Extraction

- Parse HTML with BeautifulSoup
- Extract:
  - Titles (e.g., "Total Capacity")
  - Values (e.g., "4,69,494 MT")
- Create dictionary with zone, region, titles, and values

# Building DataFrame

- Append all dictionaries to a list
- Convert list to Pandas DataFrame
- Easy to read, analyze, and export



- Run Selected Cell Shift+Enter
- Run Selected Cell and Insert Below Alt+Enter
- Run Selected Cell and Do not Advance Ctrl+Enter
- Run Selected Text or Current Line in Console
- Run All Above Selected Cell
- Run Selected Cell and All Below
- Render All Markdown Cells
- Run All Cells
- Restart Kernel and Run All Cells...
- Cell Type

## 1. Introduction

### Select

### Beautiful

### Panda

### time

ation (e.g., selecting dropdowns, clicking buttons).

to extract specific data.

format (DataFrame).

the website loads properly before the next step.

```
[ ]: from selenium import webdriver
from selenium.webdriver.chrome.service import Service
from selenium.webdriver.chrome.options import Options
from selenium.webdriver.common.by import By
from selenium.webdriver.support.ui import Select
import time
from bs4 import BeautifulSoup
import pandas as pd
```

## 2. Setting up the Chrome Driver

driver\_path: path to your downloaded ChromeDriver executable.

Service and Options are used to configure and initialize the Chrome browser in automation mode.

```
[ ]: driver_path = "C:/Users/Abus/Data Science/tech-scanning/chromedriver.exe"
```

# Conclusion

- Demonstrated automation and data extraction
- Delivered clean, structured government data
- Ready for further analysis and reporting

Thank you!

