**Day 1**

**Session 1: Introduction to Web Scraping**

* Introduction and course overview
* Understanding the basics of web scraping
* Benefits and ethical considerations of web scraping
* Demo: Collecting data from a simple website using each tool (Excel, Google Sheets, R, and Python)

**Session 2: Excel and Google Sheets for Web Scraping**

* Advantages and limitations of using spreadsheets for web scraping
* Using Excel's Web Query feature to import data from the web
* Demonstrating data filtering and basic data cleaning in Excel
* Google Sheets' IMPORTXML, IMPORTDATA, and IMPORTHTML functions for web scraping
* Hands-on exercise: Scraping a basic website and structuring data in Google Sheets

**Session 3: Setting up R and Python for web scraping**

* Setting up the environment and installing necessary packages/add-ons (R and Python)

**Day 2**

**Session 4: Web Scraping with R**

* Overview of R and its capabilities for web scraping
* Using rvest package for simple HTML web scraping
* Reading and parsing HTML content from websites
* Handling dynamic websites using RSelenium package
* Storing scraped data (e.g., CSV)
* Hands-on exercise: Scraping a dynamic website and extracting specific information

**Session 5: Web Scraping with Python**

* Web scraping with Pandas
* Python's powerful libraries for web scraping (e.g., BeautifulSoup, Requests, Selenium)
* Sending HTTP requests and parsing responses
* Handling dynamic websites
* Scraping data from APIs
* Storing scraped data (e.g., CSV)
* Hands-on exercise: Scraping a dynamic website and extracting specific information

**Session 6: Data harvesting via API’s (e.g. IMF, World Bank, data.gov.in etc.)**

* Introduce open data formats (CSV, JSON, XML)
* Demonstrate how to interact with APIs using R
* Demonstrate how to interact with APIs using Python

**Introduction to Open Data Formats (15 minutes)**

* Brief overview of CSV, JSON, and XML formats
* Differences and use cases for each format
* Example data in each format

**Using APIs in R (60 minutes)**

* Introduction to APIs and their purpose
* Setting up an environment for API access (API keys, authentication)
* Example: Accessing World Bank Data using the WDI package
  + Install and load the WDI package
  + Retrieve a list of available indicators
  + Fetch data for a specific indicator and country
  + Handling API responses (JSON/XML)
  + Data manipulation and visualization

**Break (10 minutes)**

**Using APIs in Python (60 minutes)**

* Setting up an environment for API access in Python
* Example: Accessing data.gov.in API
  + Import necessary libraries (requests, pandas)
  + Construct the API request URL
  + Send the request and handle the response
  + Parse the JSON data
  + Data manipulation and analysis using pandas

**Recap and Q&A (15 minutes)**

* Summarize key points from the session
* Address any questions or concerns from participants

**Resources:**

* Sample CSV, JSON, and XML data files
* R script with World Bank API example
* Python script with data.gov.in API example
* API documentation for World Bank and data.gov.in

**Day 3**

**Session 7: AI-Assisted Document Analysis**

* Introduction to AI in document analysis and its potential benefits.
* Demonstration of AI tools (Copilot, ChatGPT, Gemini, Perplexity etc.)
* Hands-on exercise: Analyzing a sample document using AI tools and extracting key information.
* Ethical considerations and potential biases in AI document analysis.

**Session 8: AI-Powered Writing**

* AI's role in enhancing report writing efficiency and consistency.
* Demonstration of AI tools (Copilot, ChatGPT, Gemini, Perplexity etc.)
* Best practices for AI-assisted report writing, including prompt engineering.
* Hands-on exercise: Generating a draft report using AI tools and refining the output.
* Q&A session to address challenges and concerns.