**Project Title: Web Scraping Public Supporter Data from TeamWater.org**

**Objective**

Extract publicly available supporter information from <https://teamwater.org/donations> and deliver a structured CSV file containing all relevant fields.

**Scope of Work**

1. **Data Extraction**
   * Scrape supporter details from the donations page.
   * Capture the following fields if available:
     + Name
     + Donation amount
     + Donation date
     + Location
     + Message/note
   * Ensure no duplication of supporter entries.
2. **Data Processing**
   * Clean extracted data (trim whitespace, handle missing fields, standardize formats).
   * Organize data into a structured table.
3. **Data Output**
   * Deliver a CSV file in the provided sample format.
   * Ensure the CSV has consistent headers and rows.
4. **Script Development**
   * Write a web scraping script (Python recommended, using BeautifulSoup, Selenium, or similar).
   * Script should:
     + Be reusable for future scraping of the same page.
     + Respect site rate limits and avoid unnecessary load.
     + Handle pagination or dynamic loading if applicable.
5. **Documentation**
   * Provide brief usage instructions:
     + Dependencies (e.g., Python version, required libraries).
     + How to run the script.
     + Expected output location (CSV file).

**Deliverables**

1. CSV file containing all scraped supporter data.
2. Web scraping script.
3. README/usage instructions for running the script.

**Requirements**

* Strong experience with web scraping (Python + BeautifulSoup/Selenium preferred).
* Responsible approach: must respect site rate limits and avoid disruption.
* Clear communication and ability to deliver within 2–3 days.

**Budget & Timeline**

* Budget: $50 (fixed).
* Completed: 2–3 days from project start.

**Project Breakdown:**

<https://www.notion.so/TeamWater-Supporter-Scraper-25c442b7e6c6800daab8d5254e7f7f76>

1. Setup environment:

# Navigate to your project folder

cd TeamWater\_Supporter\_Scraper

# Create a virtual environment named 'venv'

python -m venv venv

# Activate the environment

# On Windows:

venv\Scripts\activate

# On Mac/Linux:

source venv/bin/activate

**Project Repo:**

<https://github.com/danaremu/data_engineering_projects>