

Web Scraper Challenge (Medium)

- Introduction
- Problem Statement
 - Website Details
 - JSON File structure
- Requirements
- Expectations
- Setup

Introduction

Hello **Challenger**, howdy☺..

We have been hearing a lot about you. let us give you a challenge that suits your talent.

As you know nowadays it's no secret that data is what matters most. and we at **@GDGAlgiers Corp** are not an exception to this rule, we do care a lot about data, but what we also care about, is when other companies provide public **API's** to consume their data.

Unfortunately that's not the case with every website out there. happy for you it's exactly what we need you for today.

Your challenge if you are brave enough to take it, is to create a web scraper, that's given a website it should extract all the required information from it.

Problem Statement

We have been wanting to ship a new product to the market. a machine learning powered **Recipes mobile application**. that uses **AI** to process food recipes and their ingredients in correlation with market prices & ingredients availability.

We want to provide our app users (mostly restaurants & chiefs) with intelligent indicators and measures to help them chose the best menu for the day.

You might be asking where does the Web Scraper fits in all of this. Well happy to answer you that we did find the perfect dataset for our needs. but unfortunately it's not available through a public **API** or **Database**, thus we need to scrape it directly from a website that's behind an auth wall.

Your task is to figure out a way to extract all the required information from the page, process it and save it into a **JSON** file for further processing by our data analysts team.

You will find bellow the website address and the required credentials to login, as well as the **JSON** file format.

Website details

- **URL:** `https://devfest22-recipes-heaven.gdgalgiers.com`
- **Username:** `gdg_algiers`
- **Password:** `devfest2022`

JSON file structure

```
{
  "recipes": [
    {
      "id": 1,
      "name": "Homemade Cinnamon Rolls",
      "slug": "homemade-cinnamon-rolls",
      "keywords": "bake sale",
      "description": "Nothing says Sunday morning like a warm",
      "thumbnail_url": "https://img.buzzfeed.com/video-api-prod/assets/9d589367531e4c12a4937e30e521c86",
      "cook_time": 151,
      "ratings_negative": 475,
      "ratings_positive": 17340,
      "score": 98,
    },
    ...
  ]
}
```

Requirements

1. Solution must be written in **Javascript** or **Typescript**.
2. Solution must not use a 3rd party SaaS service.
3. Solution must use the provided starter code.
4. Solution can use any **JS/TS** library.

Expectations

- This challenge should take around **2 to 3** hours to complete.
- Your web scraper needs to gather all required information.
- The scraped dataset must include all recipes available in the website.

Evaluation

This challenge has maximum points of **10** for implementing all the required features.

Setup

To get you started with this challenge, we have provided you with a starting point. so you need first to copy to your local environment the `starter` folder you see next to this challenge. and install all the dependencies:

```
cd starter-folder
npm install
```

Good luck and happy hunting buddy.

↑ back to top