Web Scraping Test Description

Objective:

The objective of this test is to assess your ability to scrape product data from a website effectively. You will be required to extract information foe products from the website. The extracted data should include the product name, brand/manufacturer, price, and EAN code.

Instructions:

1. Website:

https://shop.rewe.de/

Or

https://www.citygross.se/matvaror

2. Requirements:

- Extract data for as many products as you can, ideally all.
- For each product, gather the following information:
 - Product name
 - o Brand/Manufacturer
 - Price
 - o EAN/GTIN (European Article Number) code This one is mandatory
- Tools: You are free to use any web scraping tools or libraries of your choice (e.g., Beautiful Soup, Scrapy, Selenium, etc.). Ensure that the chosen tools are capable of extracting the required information accurately.

3. Submission:

Once you have completed the scraping process, submit your code on GITHUB along with a brief
explanation of your approach. The code should be well-commented and organized for clarity.

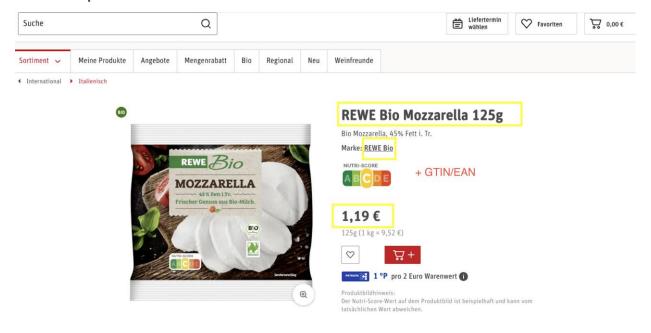
4. Assessment Criteria:

- Accuracy of the extracted data.
- Efficiency of the scraping process.
- Quality of code structure and documentation.

5. Screenshots:

 To understand what fields, we need for the scrapes, we provided two screenshots with highlighted information that is needed. EANs/GTINs are not visible on the website but they can be found somewhere inside the structure of the website. **Screenshot 1**: This screenshot illustrates the layout of the product page for Rewe.de website. In top right corner are the categories.

Shop.rewe.de



Screenshot 2: This screenshot displays the layout of the product page for Citygross.se website. Use this screenshot as a reference to locate the necessary information for extraction.

CITYGROSS.SE:

