



KYC HUB

Data Extraction Assignment

Submission Date:	Exactly 24 Hours from receiving the email
Total Points for Assignment:	30 points

Instructions:

Hello Candidate, Below you will find one data extraction assignment that you are to complete using the **Python-Selenium framework**. The data extracted from the source needs to be mapped to a JSON file in a structured manner. (An example structure has been provided below for your reference).

```
[
  {
    "name": "John Trevor Roche Baines",
    "address": "c/o Isle of Man Prison, Jurby, Isle of Man",
    "dob": "19 Dec 1939",
    "periodOfDisqualification": "15 Years 0 Months 0 Days",
    "datesOfDisqualification": "From 15 Jul 2010 To 15 Jul 2025"
  },
  {
    "name": "Fenella Jane Carter",
    "address": "30 North Quay, Douglas, Isle of Man IM1 4LD",
    "dob": "6 August 1984",
    "periodOfDisqualification": "6 Years 0 Months 0 Days",
    "datesOfDisqualification": "From 5 July 2018 to 5 July 2024"
  }
]
```

You may refer to online materials to understand how to use Python Selenium. **Do not copy and paste code from online sources.** You will have to include the following in your final submission.

1. **Python code** files for the extraction (**only .py files**)
2. Extracted data in a **JSON File**
3. A document briefly explaining your approach to this assignment.

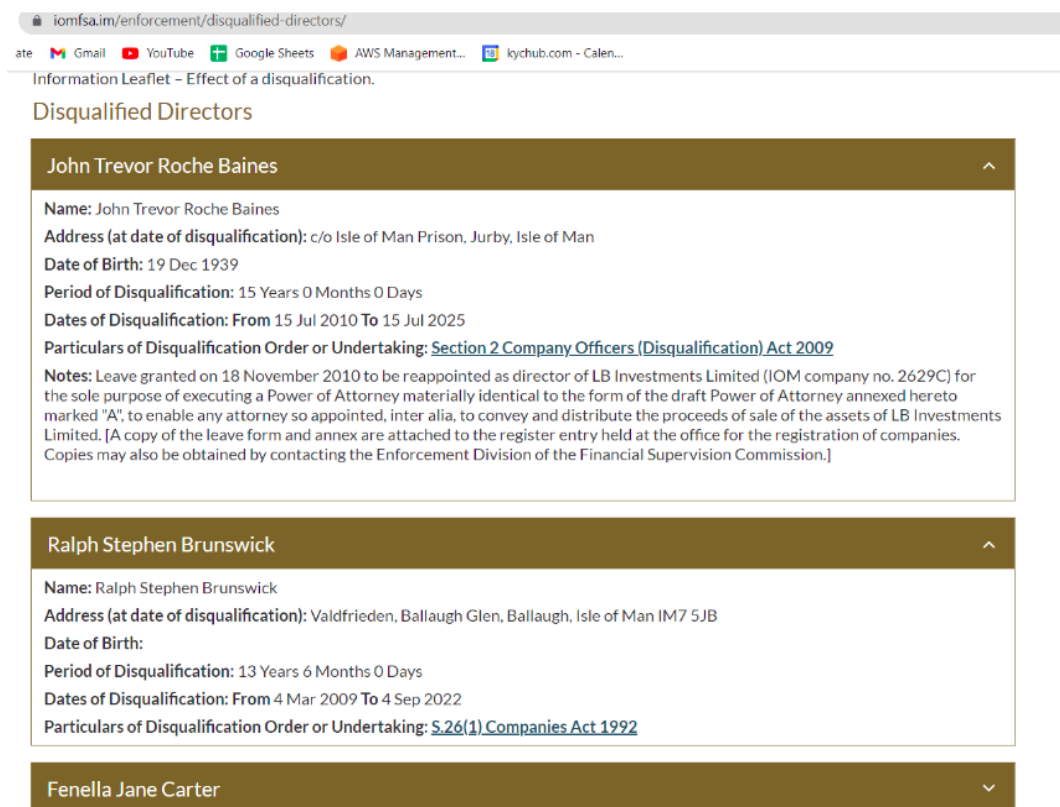
You are required to submit the assignment within 24 Hours of receiving this document.

Assignment

Please find the data source in the following link;

Link: <https://www.iomfsa.im/enforcement/disqualified-directors/>.

You are required to extract this data using the **Python Selenium framework**. Please provide the code in .py file format only, an explanation document, and also a structured JSON file in your final submission. Explore the entire website and extract all relevant information. Please ensure that no duplicate data is present in your JSON.



Sample image of the website

Evaluation Criteria:

(Total Points: 30)

1. **Working Code:** 0 to 12 points
2. **Structured JSON File:** 0 to 12 points
3. **Explanation Document:** 0 to 6 points
4. **Late Submission:** -6 to -10 points
5. **Copied Code:** -15 points
6. **Hardcoded extraction:** -10 points

☆All the best☆
