

Problem statement : Web scraping challenge

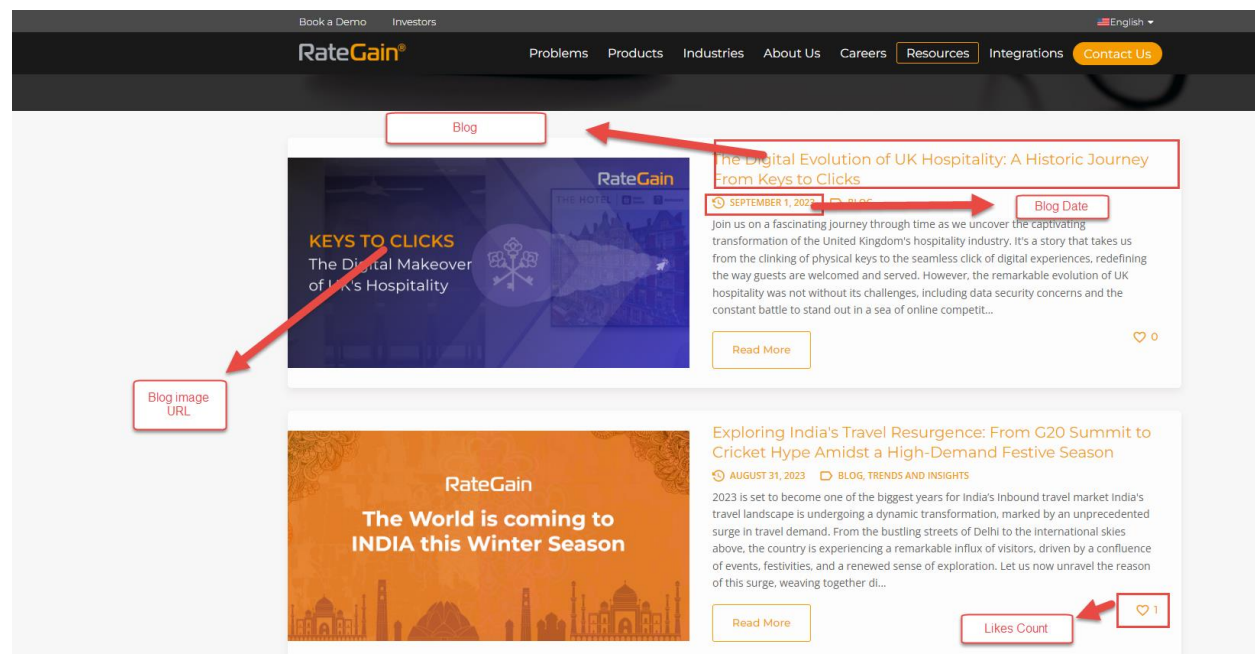
Overview

This hackathon invites participants to showcase their coding skills and creativity to solve a real-world problem: extracting valuable data from web pages.

In this challenge, Develop a program capable of extracting specific information from web pages. The target URL for this exercise is "<https://rategain.com/blog>".

Task

Your journey begins with a visit to "<https://rategain.com/blog>." On this webpage, you will find numerous blog posts. Your application must identify and collect data from these posts, which are highlighted in red.



This task is not limited to a single page. You must navigate through various pages to gather data comprehensively. Your application should extract the following details:

- **Blog title:** Capture the titles of the blog posts.
- **Blog date:** Retrieve the publication dates of each blog post.
- **Blog image URL:** Extract the URLs of the images associated with the blogs.
- **Blog likes count:** Record the number of likes each blog post has received.

- **Data management:** The data you gather should be organized and saved efficiently. The preferred format for storage is either Excel or CSV. Ensure that the extracted data is accurately structured for easy analysis.

Resources

- [Requests](#)
- [Selenium](#)
- [Import.io](#)
- [Legal and Ethical Issues of Web Scraping](#)
- [Working with Excel Files in Python](#)