**1. A developer is assigned a task to scrape 1 lakh website pages from a directory site, while scrapping he is facing such hcaptcha, which are placed to stop people from scrapping As a project Coordinator suggest ways to solve this problem**

|  |
| --- |
| * First thing This captcha is to prevent too many requests in a small amount of time so, I would suggest adding a line of code to give time gaps between scraping. |
|  |
| * I would suggest looking for an Alternative source of the same data, such as some API, or raw Data in the form of XML or JSON. |
|  |
| * Proxy Servers: we can utilize proxy servers to route the scraping request through different IP addresses. Proxies help hide real IPs and provide a pool of IP addresses for scraping tasks. |
|  |
| * Headless Browsers: Headless Browsers can render and interact with web pages, making it harder for websites to detect automated scraping. |

**2. Our client has around 10k LinkedIn people profiles, he wants to know the estimated income range of these profiles. Suggest ways on how to do this?**

|  |
| --- |
| * We can analyze Job Titles and Industries - Job titles correlate with specific salary ranges. We can use online LinkedIn profiles / online resources to estimate average income. |
|  |
| * Location-Based Estimates - Salaries vary based on regions. |
|  |
| * Education and Experience - Research on average salary earned based on level of education and experience. |
|  |
| * Company Size and Reputation - We should also consider the size and reputation of the company. |
|  |
| * We can also use LinkedIn Premium Feature / Third Party Data Analysis Tools if we are provided with the same budgets. |

**3. We have a list of 1L company names, need to find LinkedIn company links of these profiles, how to go about this?**

|  |
| --- |
| * You can automate the process by utilizing web scraping techniques like Beautiful Soup or other Python libraries. |
|  |
| * We can also use API to fetch the LinkedIn Profiles and save them in the database. We can use SQL and write queries to do this. |

**4. How to identify list of companies whose tech stack is built on Python. Give names of 5 companies if possible, by your suggested approach**

|  |
| --- |
| * The first thing I will do is go for GitHub, Bitbucket, and Gitlab, many companies post their open-source projects on these platforms, so I would look for Python files. |
|  |
| * Technology Blogs and Articles - companies share their tech stack on blogs and articles. |
|  |
| * Tech Communities and Conferences: Participate in Python-related tech communities and conferences. Companies actively participating or sponsoring these events are likely to be using Python in their projects. |
|  |
| * Job Sites - We can also look for the Python Keyword in the Job listing by the Companies. Websites like LinkedIn, Indeed, or company career pages can be searched for job postings that mention Python as a required skill. |
|  |
| * Amazon, Google, Dropbox, Spotify, Uber – These are the 5 companies whose tech stack is built on Python. |

**5. Need to find an API, through which we can send LinkedIn messages to other LinkedIn users**

|  |
| --- |
| * Sending a direct message to a person using the LinkedIn API requires obtaining the appropriate API permissions and making an API request to the LinkedIn API endpoint. The LinkedIn API provides the /v2/messages endpoint for sending messages. |
|  |
| * Obtain API permissions: To use the LinkedIn API, we need to obtain the appropriate API permissions. This can be done by creating a LinkedIn Developer Account and registering the application. |
| * Authenticate the API request. |
| * Send the API request: Once you have obtained the access token, we can send a POST request to the /v2/messages endpoint. |