Internship Project – 1

Problem Statement.: We need to scrap the data from the website : https://github.com/topics

The vendor for this project is an ed-tech company which is looking out to extract the best resources available on the open source environment for various topics.

1) Python.
2) Data Visualization
3) react
4) javascript
5) Flask
6) GraphQL
7) Java
8) JSON
9) Machine Learning
10) PostgresSQL

We need to extract the information via webscrapping for the above topics anc create individual excel sheets for each and every topic.

The excel sheet should consist the data in the following pattern.

Sr No	Repository Name	User Name	Stars	URL

We need a maximum of top 200 repository data for each and every topic in a separate excel sheet.

Once all the excel sheet is ready we shall then share our insight with our client for further actions.