

KARTIK BHATIA
9990643661
kartikbhatia2705@gmail.com

GitHub repo: <https://github.com/kartik27b/anakin-webscraping>

Q1) For this question the solution to get the latitude and longitudes is to find the api calls and extract data from them.

I found that there are 2 api calls which are useful.

First is the <https://portal.grab.com/foodweb/v2/search> API which is being hit to gather restaurant data. But this request does not work on its own and requires a X-Recaptcha-Token header to work.

This captcha token head can be extracted from another API call which is <https://www.google.com/recaptcha/api2/reload> which gets the token in its response body

So I used these 2 api's to extract the data from the grab website.

Data link: https://drive.google.com/file/d/1Qs5YAnNAF0-F7Sgjp-jcKutsvM-Ko5AW/view?usp=share_link

Code: <https://github.com/kartik27b/anakin-webscraping/blob/main/grab.py>

Q2) For this question I have first extracted the super categories then for each super category I extracted the category then for each category I extracted the subcategory and link to final page where we have to extract products. So using the final page link I extracted 10 products and stored them to json file

Spreadsheet link: https://docs.google.com/spreadsheets/d/1cEnMAzPXk9g_YJWiFh8ZpJu-UgBtLwyiFz9P_YartKg/edit?usp=share_link

JSON Data link: https://drive.google.com/file/d/1F3pcuzYox8useYCy5t7KRmjWaN8HM2gP/view?usp=share_link

Code: <https://github.com/kartik27b/anakin-webscraping/blob/main/bigbasket.py>

Q3)

We can do this with the help of Wireshark

1. Install WireShark

2. now we have to create Android virtual device(AVD) so we will download Android SDK from official site.
3. Android SDK come with emulator for testing after setting up Android SDK, create a Android virtual Device(AVD) on which we will install app
4. after creating virtual device install app using adb command `adb install milkbasket.apk`
5. now we can start capturing the packets so I will suggest to close other application on your computer which are using network so our captured packets would be more relevant.
6. now start wireshark with root access
7. select interface which you want to capture and click start to start capturing.
8. now start using that app so packets will transfer to and fro and wireshark will capture it.
9. The next step is to filter the packets based on http and see the info carefully for listed packets you will see lot of important details, API keys, cookies and also find the API used for fetching the products data.
10. We can copy this API endpoint and use in our code to fetch data