

# Anakin Grab Assignment

Viral Barodia | Jr. Software Engineering Intern @Growexx

- Approach

The scraper is designed following Object-Oriented Programming (OOP) principles and optimized for handling high load using multithreading. I have used the following address to scrape the locality's website:

Chong Boon Dental Surgery - Block 456 Ang Mo Kio Avenue 10,  
#01-1574, Singapore, 560456

The problem faced during the implementation

- Problem 1: Handling dynamic content loading

The Grab website dynamically loads restaurant data as the user scrolls. This means that not all restaurant data is available immediately when the page loads.

Solution/Approach: To solve this issue, we implemented a scrolling mechanism in the `load_all_restaurants` method. This method scrolls to the bottom of the page in a loop, waits for new content to load, and repeats until no new content is loaded. This ensures that all restaurant data is loaded before we start scraping.

Steps to execute the code locally

1. Extract the code files and ready them for running in your local system
2. Update the ``proxy`` variable in the `main.py` with the proxy you're using
3. Run the code by simply writing ``python main.py`` in the command line of the project directory
4. Wait for the program to execute
5. Once the code running is complete, we can see the `restaurants.ndjson` for results.

- Room for improvement: Usage of a proxy/VPN to avoid 403 error.