# ĐẠI HỌC QUỐC GIA THÀNH PHỐ HÒ CHÍ MINH TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN

# THU THẬP VÀ TIỀN XỬ LÝ DỮ LIỆU BÀI TẬP 3

# Thu thập dữ liệu tọa độ các đỉnh của đa giác tạo thành bản đồ TP.HCM

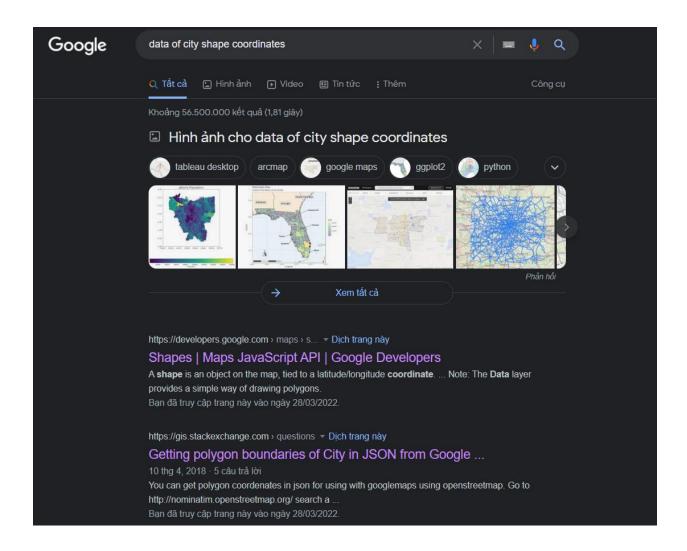
Họ và tên: Lưu Quang Tiến Hoàng

MSSV: 20521342

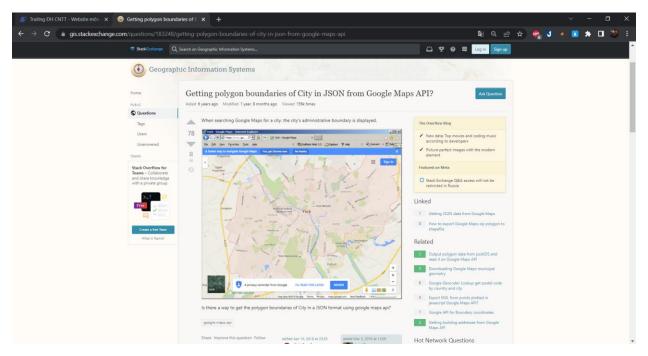
Lóp: DS103.M21

# Tìm kiếm nguồn dữ liệu:

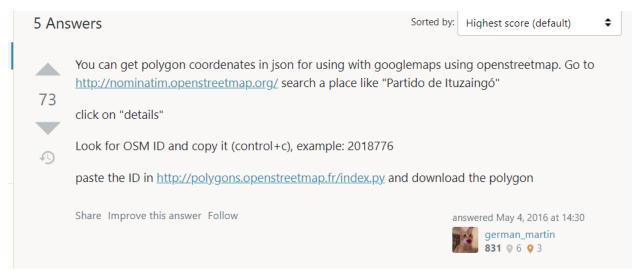
Đầu tiên ta sử dụng Google để tìm kiếm về dữ liệu của đa giác tạo nên bản đồ Tp.HCM. Từ khóa tìm kiếm là "Data of city shape coordinates".



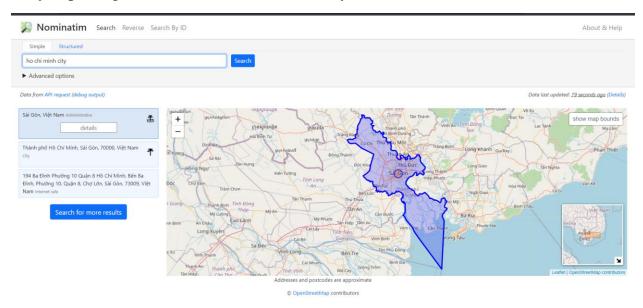
Tham khảo thì thấy kết quả tìm kiếm số 2 phù hợp với yêu cầu của bài nhất nên ta chọn.



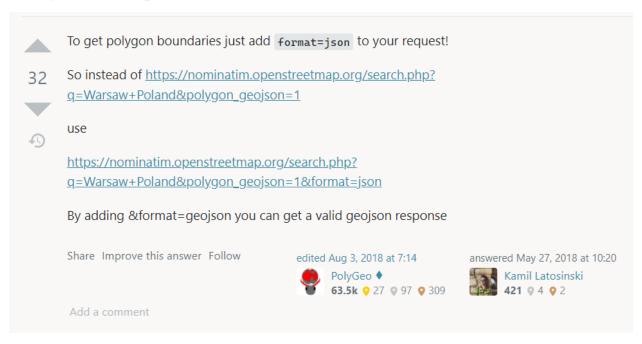
Sau khi đọc và tham khảo thì thấy bình luận dưới đây là lượt bình luận có tính xác thực và chính xác cao nhất nên ta làm theo hướng dẫn.



Truy cập trang và tìm kiếm "ho chi minh city" và click detail.



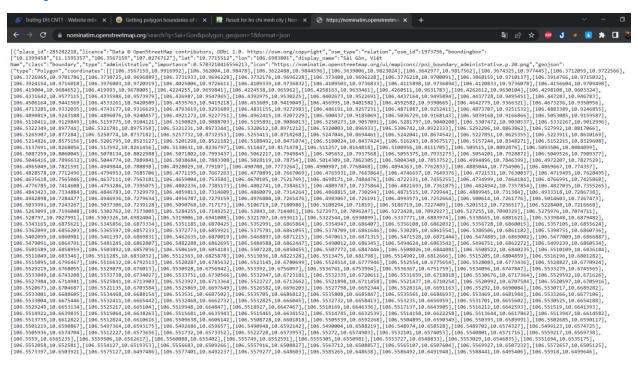
Theo nguồn của trang <a href="https://gis.stackexchange.com/questions/183248/getting-polygon boundaries-of-city-in-json-from-google-maps-api">https://gis.stackexchange.com/questions/183248/getting-polygon boundaries-of-city-in-json-from-google-maps-api</a> ở trên thì ta tìm được hướng dẫn thu thập dữ liệu tọa độ.



Chúng ta thêm vào đường dẫn URL "&polygon\_geojson=1&format=json" thì ta sẽ được trang web trả về file json chứa kinh độ và vĩ độ của đa giác ranh giới của TP.HCM.

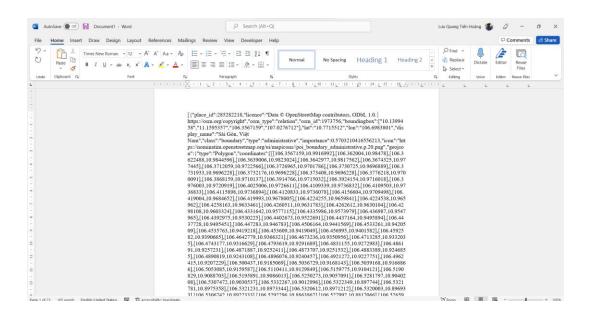
#### **URL**:

https://nominatim.openstreetmap.org/search?q=Sai+Gon&polygon\_geojson=1&format=json

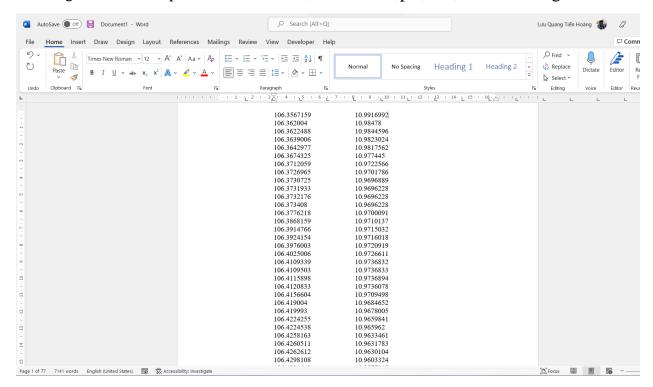


## Cách xử lý:

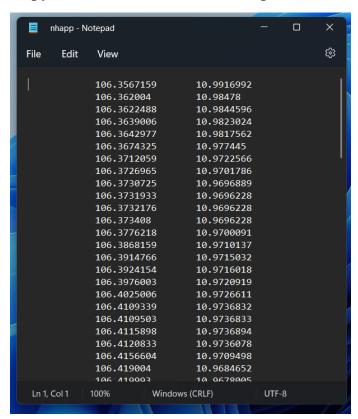
Sử dụng các phần mềm như Word, Excel, Notepad...để xử lý dữ liệu. Ban đầu chúng ta copy tất cả vào file Word.



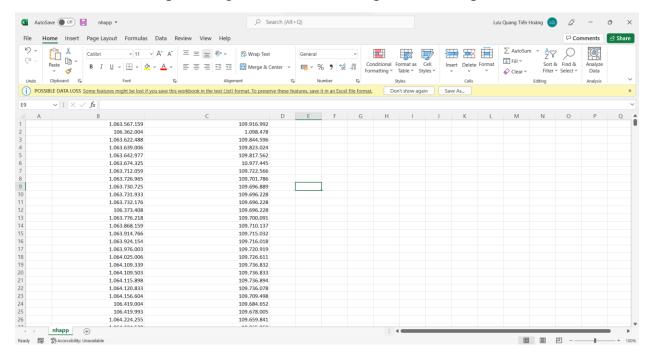
# Sử dụng Find and Replace để đưa dữ liệu về mỗi cặp tọa độ thành 1 hàng.



### Copy toàn bộ lưu lại vào file notepad.



Sau đó, mở file notepad bằng Excel và cuối cùng lưu lại kết quả.



Hoàn tất bài làm.