

E1:

Crawl_Data - BeautifulSoup:

Khó khăn:

- + Tìm hiểu và sử dụng cấu trúc HTML
- + Khi sử dụng các thư viện liên quan (cài đặt thư viện, các bug khi code khó khăn khi fix,...)
- + Bị lỗi đường dẫn, chặn truy cập
- + Khi phân tích HTML với lượng thông tin lớn, mã hóa dữ liệu
- + Dữ liệu có thể chứa các ký tự đặc biệt hoặc mã hóa không đúng, gây lỗi khi xử lý.
- + Xử lý dữ liệu bị thiếu

E2: