Họ và tên: Lê Thị Bảo Yến

MSSV: 0850080057

Lớp: 08\_ĐH\_CNPM

BUỔI 3

NGUYÊN CỨU TẠO DỮ LIỆU TỰ ĐỘNG

1. Giả thiết

Xây dựng một hệ thống website quản lý vật tư y tế, cần số lượng dữ liệu lớn để kiểm tra chất lượng hệ thống, việc tạo tay gây tốn thời gian. Nghiên cứu tạo dữ liệu tự động bằng Python.

1. Thư viện

Sử dụng các thư viên sau:

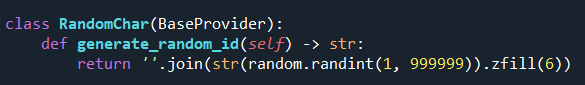
* Thư viện Faker để tạo dữ liệu giả.
* Thư viện Random để lấy giá trị ngẫu nhiên.
* Module Workbook của thư viện openpyxl để tạo và xử lý file Excel.
* Thư viện os dùng để thao tác với thư mục lấy địa chỉ lưu file.
* Module date của thư viện Datetime để xử lý thời gian của ngày sinh.
* Callable, Any để định nghĩa kiểu dữ liệu đầu vào là Faker và trả về là List bất kỳ.
* BaseProvider để tùy chỉnh các giá trị giả mạo

1. Thực nghiệm

Tạo dữ liệu giả cho bảng Nhân Viên gồm các thuộc tính sau: ['ID', 'Ten', 'CCCD', 'GioiTinh','NgaySinh','QueQuan','DiaChi','SDT','Mail','ChiNhanh'].

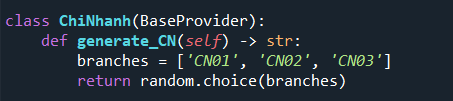
* 1. Tạo Random cho ID

Tạo ngẫu nhiên chuỗi với độ dài là 6 kí tự số.



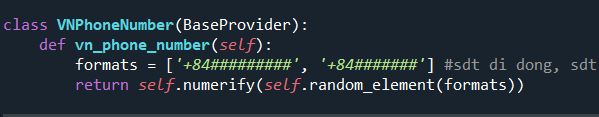
* 1. Tùy chỉnh giá trị ChiNhanh

Gỉa sử hệ thống quản lý có 3 ChiNhanh là CN01, CN02 và CN03, tùy chỉnh sẽ lấy một trong 3 lựa chọn.



* 1. Định dạng kiểu điện thoại vùng VN

Với số điện thoại di động là +84######### và +84####### tương ứng với số điện thoại bàn ở Việt Nam. Tạo số điện thoại giả sẽ random số ở kí tự #.

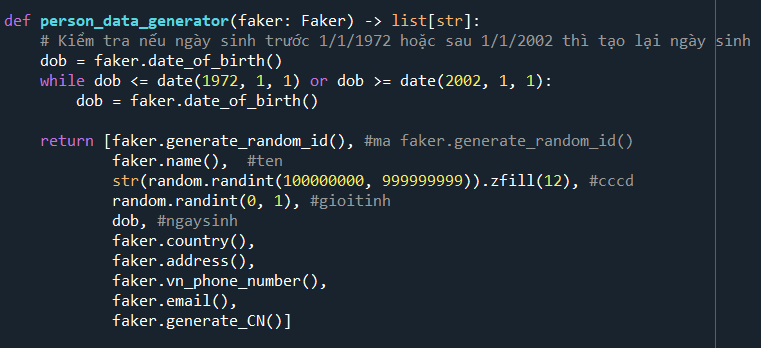


* 1. Xây dựng hàm tạo dữ liệu nhân viên sử dụng Faker trả về List kiểu string.

Trong đó có kiểm tra ngày sinh của nhân viên, nếu ngày sinh trước 1/1/1972 hoặc sau 1/1/2002 thì tạo lại ngày sinh, có thể tùy chỉnh lại.

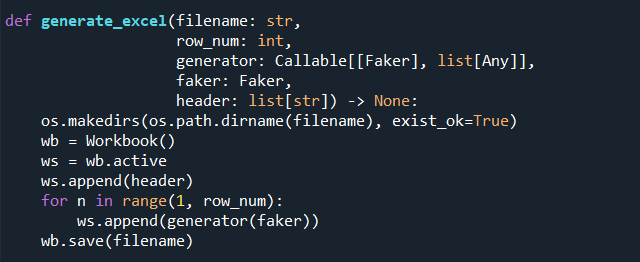
Sau đó là lệnh Return trả về với các dòng lần lượt là:

* Tạo ID sử dụng hàm generate\_random\_id()
* Tên nhân viên
* CCCD với 12 số
* Random giới tính kiểu 0-1 để phù hợp với dữ liệu kiểu bit trong database, với 0 là nam và 1 là nữ.
* Lấy ngày sinh đã tạo tự động và phù hợp điều kiện ở trên
* Tạo Quê quán
* Tạo địa chỉ
* Tạo số điện thoại sử dụng hàm vn\_phone\_number()
* Tạo email
* Tạo Chi Nhánh làm việc với hàm generate\_CN()



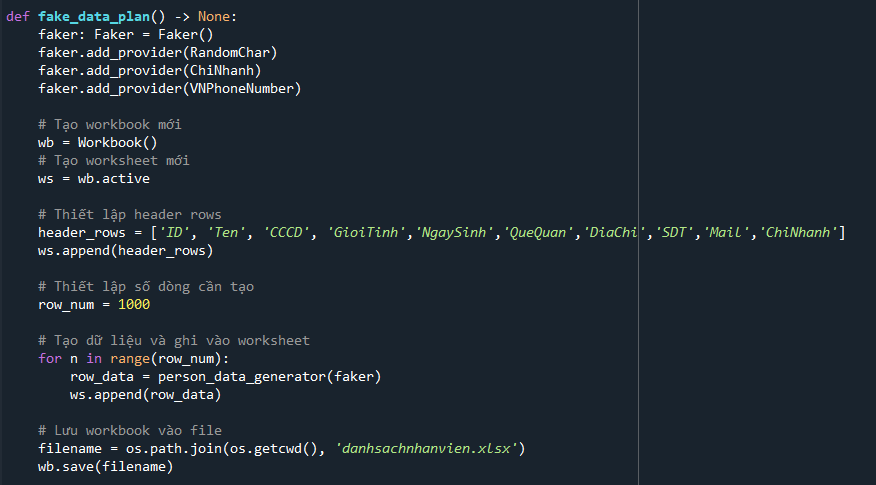
* 1. Xây dựng hàm tạo file Excel

Hàm bao gồm khai báo filename, tổng số record muốn tạo tự động, hàm với đối số là Faker trả về kiểu dữ liệu List, là header của file. Khởi tạo đường dẫn lưu file và tạo workbook với các dòng dữ liệu thêm sau vào cho tới hết số record.

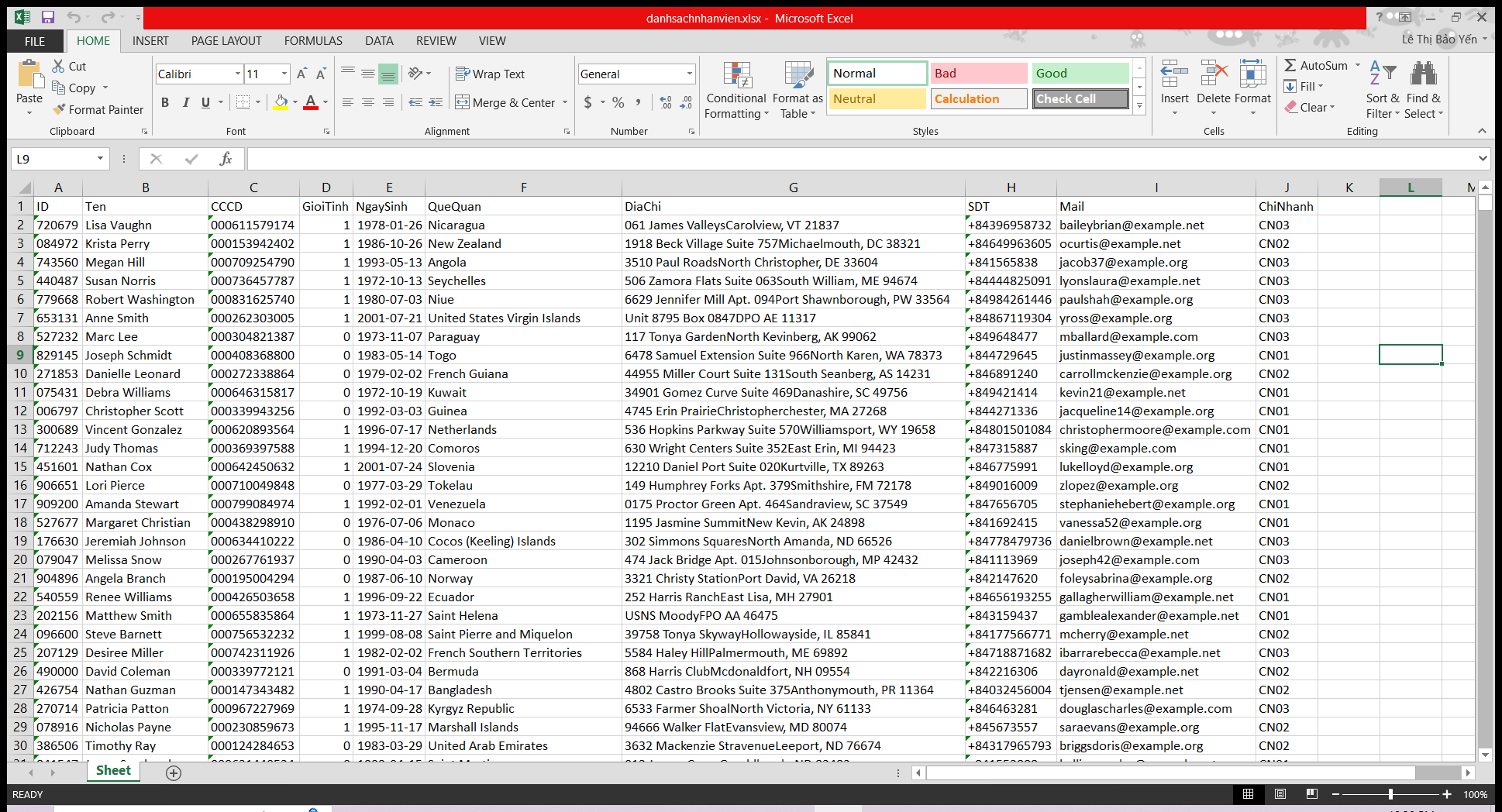


* 1. Xây dựng hàm sinh dữ liệu giá

Khởi tạo Faker, Add Provider các Class chứa hàm generate đã tùy chỉnh. Khởi tạo workbook và worksheet, thiết lâp header và số record. Sau đó tạo dữ liệu bằng hàm person\_data\_generator() đã tạo trước đó, cuối cùng lưu file excel với tên là danhsachnhanvien tại thư mục gốc.



* 1. Kết quả



1. Kết luận

Hiện tại chỉ xuất dữ liệu về dạng excel, chưa xuất về dạng Json, có thể sử dụng tạm trang web <https://products.aspose.app/cells/vi/conversion/excel-to-json> để chuyển đổi sang Json.

Kết quả sau khi chuyển về Json

