**Báo cáo kết quả nội dung thực hành Week 4**

**Data Understanding and Visualization**

Họ và Tên: Kiều Thế Vinh – nhóm 3

**Các nội dung thực hành được:**

* Tạo một DataFrame, lọc các dữ liệu cần thiết với pandas
* Trực quan hoá bằng matplotlib

**Những công việc đã làm:**

* Đọc dữ liệu file dạng json
* Lấy ra date object từ string có chứa object
* Lấy ra chủ đề của bài báo từ url của bài
* Loại các cột không cần thiết và các hàng có tác giả là null ( do từ 2014 trở về trước css selector của trang cho tác giả khác nhau)
* Tạo 3 figure để trực quan hoá số bài theo tác giả, theo chủ đề, theo năm

**Những việc chưa làm được:**

* Số bài của mỗi tác giả qua các năm
* Chủ đề mà mỗi tác giả viết nhiều nhất