

Project tracker

Task	Details	Assigned to	Estimated start	Estimated finish	Notes
1 [Data collection] Purposes, objectives, sources details	Trình bày trong file "Data_collection.ipynb". Mục thông tin dataset trên Kaggle có link Github về việc scrape dataset	Châu Anh	9/28/2024	10/5/2024	
2 [Data collection] Kiểm tra và chỉnh sửa markdown. Load dataset	Download dataset từ Kaggle, đọc file	Kim Anh	10/5/2024	10/6/2024	
3 [Data exploration] File 'worldometer.csv'	Dòng, cột, ý nghĩa. Kiểm tra trùng lặp, missing values, các giá trị thống kê, phân bố numerical feature và categorical feature. Lưu cleaned data	Dịu Huyền	10/7/2024	10/12/2024	
4 [Data exploration] File 'clean_complete.csv'	Dòng, cột, ý nghĩa. Kiểm tra trùng lặp, missing values, các giá trị thống kê, phân bố numerical feature và categorical feature. Lưu cleaned data	Kim Anh	10/7/2024	10/12/2024	
5 [Data exploration] File 'day_wise.csv'	Dòng, cột, ý nghĩa. Kiểm tra trùng lặp, missing values, các giá trị thống kê, phân bố numerical feature và categorical feature. Lưu cleaned data	Châu Anh	10/7/2024	10/12/2024	Xóa file này vì không cần data ở file này do có thể nội suy từ file clean_complete.csv
6 [Data exploration] Kiểm tra lại việc kiểm tra các giá trị bất thường. Sửa lại việc xử lý giá trị Region và WHO Region tại file worldometer	Thiếu giá trị ở cột Region và WHO Region (Drop hoặc Fill)	Kim Anh	10/12/2024	10/13/2024	Thiếu giá trị do đó là du thuyền Diamond Princess. Xử lý bằng cách fill giá trị ở cột Region là Cruise, những địa điểm không có WHO Region thì fill là No-WHORegion
7 Bổ sung Mục lục cho các file notebook	Tạo mục lục, dùng html để có thể link vị trí bằng id	Châu Anh	10/14/2024	10/16/2024	
8 [Data collection + Data exploration] Kiểm tra lại và chỉnh sửa markdown nếu có. Kiểm tra có thiếu nội dung nào đã yêu cầu hay không	Chỉnh sửa lại outlines nếu cần. Thêm các nội dung còn thiếu, xem xét xử lý dataset hợp lý chưa	Dịu Huyền	10/14/2024	10/16/2024	
9 Bổ sung phần license vào Data_collection	Tác giả là ai, có cho phép không? Có thể đặt trong mục Data source detail (nhớ in đậm đề mục)	Châu Anh	10/14/2024	10/18/2024	
10 Sửa mục Data source details trong Data_collection	Chia nội dung trong đó thành 2 sub đề rõ ý thành 3 phần: Source (link), License, Collecting method (Cách thức tác giả lấy được data)	Châu Anh	10/14/2024	10/18/2024	
11 [Data analysis] Visualization	Tổng quan về dữ liệu: Correlation matrix, Tổng các ca Xác nhận, Tử vong, Hồi phục, Top 10 các nước có số ca mắc cao nhất và thấp nhất	Dịu Huyền	10/23/2024	10/31/2024	
12 [Data analysis] Đặt câu hỏi 1	Câu hỏi là gì? Tại sao phải trả lời câu hỏi? Cách để có được câu trả lời? Preprocess nếu cần để trả lời câu hỏi (như bên Intro)	Châu Anh	11/1/2024	11/13/2024	

Task	Details	Assigned to	Estimated start	Estimated finish	Notes
13 [Data analysis] Đặt câu hỏi 2	Câu hỏi là gì? Tại sao phải trả lời câu hỏi? Cách để có được câu trả lời? Preprocess nếu cần để trả lời câu hỏi (như bên Intro)	Kim Anh	11/1/2024	11/13/2024	
14 [Data analysis] Đặt câu hỏi 3	Câu hỏi là gì? Tại sao phải trả lời câu hỏi? Cách để có được câu trả lời? Preprocess nếu cần để trả lời câu hỏi (như bên Intro)	Dju Huyền	11/1/2024	11/13/2024	
15 Kiểm tra lại câu hỏi		Kim Anh	11/14/2024	11/17/2024	
16 Reflection	Khó khăn phải đối mặt trong đồ án (mỗi thành viên). Bài học (mỗi thành viên). Nếu còn thời gian sẽ làm gì?	All	11/18/2024	11/22/2024	
17 References	Bỏ nguồn vào	All	9/28/2024	12/27/2024	
18 [Data analysis] Trả lời câu hỏi 1	Preprocess lại nếu cần. Markdown đầy đủ cho các bước	Châu Anh	11/25/2024	12/6/2024	
19 [Data analysis] Trả lời câu hỏi 2	Preprocess lại nếu cần. Markdown đầy đủ cho các bước	Kim Anh	11/25/2024	12/6/2024	
20 [Data analysis] Trả lời câu hỏi 3	Preprocess lại nếu cần. Markdown đầy đủ cho các bước	Dju Huyền	11/25/2024	12/6/2024	
21 [Data analysis] Kiểm tra lại		Kim Anh	12/9/2024	12/13/2024	
22 Chia line bài thuyết trình và thống nhất nội dung thuyết trình		All	12/14/2024	12/26/2024	