**CỘNG HÒA XÃ HỘI CHỦ NGHĨA VIỆT NAM**

**Độc lập - Tự do- Hạnh phúc**

**BÁO CÁO HỌC TẬP NHÓM**

*Kính gửi*: Giảng viên bộ môn Lập Trình Python nâng cao- thầy Cao Diệp Thắng

*Tên nhóm*: nhóm 3 Lớp: DHKL18A3HN

*Thành viên nhóm:*

* Đỗ Thị Vân Anh- 24174600128
* Đinh Diệu Linh- 24174600172
* Nguyễn Thùy Linh- 24174600134
* Hoàng Thị Mai- 24174600153
* Hoàng Nguyễn Khánh Vy- 24174600156

*Tên chủ đề* : **CASE STUDY 13 – DỮ LIỆU TIÊM CHỦNG & SỨC KHỎE CỘNG ĐỒNG**

*Thời gian hoàn thành:*trước 0 giờ ngày 17/12/2025.

**I. Giới thiệu chủ đề:**

Dữ liệu tiêm chủng trong cộng đồng thường được thu thập từ nhiều nguồn như trạm y tế,

bệnh viện, chiến dịch tiêm chủng lưu động. Điều này dẫn đến dữ liệu không đồng nhất, dễ

phát sinh lỗi về mã số bệnh nhân, hình thức ghi ngày tháng, cách viết tên vaccine hoặc mức

độ phản ứng sau tiêm. Case study này giúp rèn luyện kỹ năng khảo sát – làm sạch

– tổng hợp – phân tích dữ liệu y tế thực tế theo đúng quy trình chuẩn.

III. Phương pháp xử lý và phân tích dữ liệu

**Dữ liệu thô**

- patient\_info.csv

- vaccination\_record.csv

- reaction\_log.csv

**Truy vấn & thống kê mô**

**Đọc và làm sạch dữ liệu**

**Pivot Table + Stack/Unstack**

**GroupBy & Tổng hợp**

**Merge dữ liệu**

**II. Phân công nhiệm vụ:**

|  |  |  |  |
| --- | --- | --- | --- |
| **Tên thành viên** | **Nhiệm vụ** | **Kĩ thuật sử dụng** | **Hoàn thành** |
| Đỗ Thị Vân Anh | **Task 5 – Pivot Table + Stack/Unstack:**  - Tạo pivot tổng số liều tiêm theo nhóm tuổi ×vaccine\_type.  - Tạo pivot mức độ phản ứng trung bình theo vaccine\_type × age\_group.  - Thực hành stack/unstack  - Tạo pivot đa cấp để thống kê số bệnh nhân, số lần tiêm và tỷ lệ có phản ứng | pivot\_table(), "stack(), unstack() | x |
| Đinh Diệu Linh | **Task 4 – Merge dữ liệu:**  - Merge patient\_info với vaccination\_record theo patient\_id.  - Merge thêm reaction\_log  - Phát hiện dữ liệu bị thiếu  - Tạo bảng full\_outer\_merge để liệt kê tất cả bệnh nhân và phân tích nguyên nhân lệch dữ liệu. | merge() với inner/left, | x |
| Nguyễn Thùy Linh | **Task 1 – Đọc & làm sạch dữ liệu**  - Khảo sát cấu trúc 3 bảng  - Chuẩn hoá mã và nhãn quan trọng.  - Xử lý dữ liệu thiếu và giá trị bất thường.  - Kiểm tra tính hợp lệ logic  - Đảm bảo khả năng merge  - Xuất các file sạch | read\_csv(), to\_csv(), dropna(), fillna(), astype(),  to\_datetime(), str.strip(),  str.title() | x |
| Hoàng Thị Mai | **Task 3 – GroupBy & Tổng hợp:**  - Tổng số liều tiêm theo nhóm tuổi và loại vaccine.  - Trung bình số liều tiêm theo giới tính.  - Tỷ lệ bệnh nhân có phản ứng sau tiêm theo vaccine\_type.  - Mức độ phản ứng trung bình theo nhóm tuổi. | groupby(), agg() | x |
| Hoàng Nguyễn Khánh Vy | **Task 2 – Truy vấn & thống kê mô tả:**  - Đếm số bệnh nhân theo nhóm tuổi.  - Tìm bệnh nhân đã tiêm từ 2 loại vaccine trở lên.  - Thống kê số lần tiêm theo loại vaccine.  - Liệt kê các bản ghi có dose\_number bất thường | loc, iloc, isin | x |

#### **III. Kết quả :**

|  |  |  |
| --- | --- | --- |
| **Chỉ số** | **Facebook** | **Twitter** |
| Bình luận tích cực | Cao | Trung bình |
| Bình luận tiêu cực | Trung bình | Cao |
| Tính ổn định theo thời gian | Ổn định | Biến động mạnh |

### **IV. Nhận xét & Đánh giá:**

* Facebook thể hiện xu hướng người dùng phản hồi tích cực hơn.
* Twitter nhạy với sự kiện, tạo nhiều bình luận tiêu cực/nhạy cảm.

V. Kết luận

*Nhóm 3 xin chân thành cảm ơn!*

**NGƯỜI BÁO CÁO**

*(Ký và ghi rõ họ tên)*

Vân Anh

Đỗ Thị Vân Anh