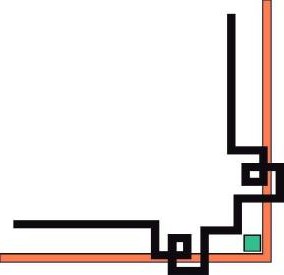
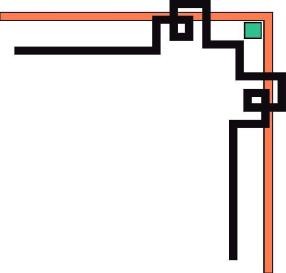
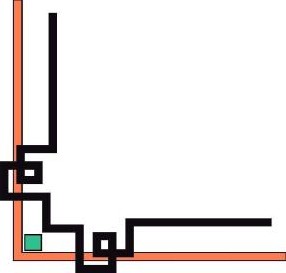
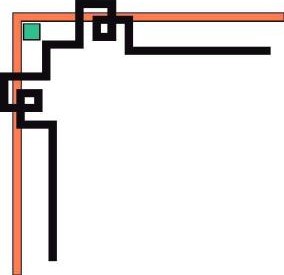
**BỘ GIÁO DỤC VÀ ĐÀO TẠO**



**TRƯỜNG ĐẠI HỌC SƯ PHẠM KỸ THUẬT TP.HCM**

**KHOA CƠ KHÍ ĐỘNG LỰC**

**MÔN HỌC: NHẬP MÔN LẬP TRÌNH PYTHON**

 **\*\*\*\*\*\*🙡🕮🙣\*\*\*\*\*\***

**BÁO CÁO CUỐI KÌ**

**ĐỀ TÀI: MATH-ASSETMENT-CRUD-ANALYSIS**

**GVHD: TRẦN QUANG KHẢI**

|  |  |  |
| --- | --- | --- |
| **SVTH:** | **PHAN KHÁNH DUY** | **24154020** |
|  | **NGUYỄN DUY HÀO** | **24154032** |
|  | **NGUYỄN DUY HẢO** | **24154034** |
|  | **HỒ PHẠM SAO KHUÊ** | **24154056** |
|  | **NGÔ TRUNG HOÀNG ĐỨC** | **24154028** |

**Tp. Hồ Chí Minh – 12/2024**



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

**Độc lập – Tự do – Hạnh phúc**

**Tp. Hồ Chí Minh, Tháng 12, Năm 2024**

**DANH SÁCH THÀNH VIÊN THAM GIA VIẾT BÁO CÁO**

**HỌC KỲ I NĂM HỌC 2024 – 2025**

**Lớp học phần:241INPY131685\_08**

**Giảng viên hướng dẫn: Trần Quang Khải**

**Tên đề tài:** “*Math-Assetment-Crud-Analysis”*

|  |  |  |  |
| --- | --- | --- | --- |
| **STT** | **Họ và tên sinh viên** | **Nội dung thực hiện** | **Tỉ lệ** |
| 1 | Phan Khánh Duy | Code xây dựng giao diện | 100% |
| 2 | Ngô Trung Hoàng Đức | Code xây dựng giao diện, miêu tả dữ liệu | 100% |
| 3 | Nguyễn Duy Hào | Code xây dựng chức năng | 100% |
| 4 | Nguyễn Duy Hảo | Code xây dựng đồ thị | 100% |
| 5 | Hồ Phạm Sao Khuê | Tổng hợp và hoàn thiện code, xây dựng báo cáo phần mở đầu | 100% |

***Nhận xét của giáo viên***

........................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................ ..............................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................................

*Tp. Hồ Chí Minh – Tháng 12 năm 2024*

**MỤC LỤC**

[PHẦN 1: MỞ ĐẦU 1](#_Toc12608)

[1.1 Mục tiêu 1](#_Toc24288)

[1.2 Đối tượng và phạm vi nghiên cứu 1](#_Toc8118)

[PHẦN 2: NỘI DUNG 2](#_Toc21890)

[CHƯƠNG 1: THÔNG TIN DỮ LIỆU 2](#_Toc2793)

[1.1 Nguồn gốc 2](#_Toc26432)

[1.2 Công dụng 2](#_Toc6597)

[1.3 Hạn chế 2](#_Toc13328)

[CHƯƠNG 2: MÔ TẢ DỮ LIỆU 4](#_Toc30289)

[2.1 Các tập dữ liệu 4](#_Toc18455)

[CHƯƠNG 3: PHÂN TÍCH DỮ LIỆU 5](#_Toc25945)

[3.1 Khởi tạo dữ liệu 5](#_Toc326)

[3.1.1 Tkinter 5](#_Toc4115)

[3.1.2 Pandas: 5](#_Toc17117)

[3.1.3 Os: 5](#_Toc14365)

[3.1.4 Matplotlib: 5](#_Toc28151)

[CHƯƠNG 4: CÁC HÀM CHỨC NĂNG 6](#_Toc4173)

[4.1 Tải dữ liệu-load data 6](#_Toc3983)

[4.2 Hiển thị-display data 6](#_Toc2347)

[4.3 Tự động điển thông tin-auto fill 7](#_Toc30626)

[4.4 Thêm-add data 8](#_Toc5644)

[4.5 Xóa-delete data 9](#_Toc19026)

[4.6 Tìm kiếm-Search data 10](#_Toc22562)

[4.7 Cập nhật- update 11](#_Toc25403)

[4.8 Sắp xếp -sort 12](#_Toc18180)

[CHƯƠNG 5: BIỂU ĐỒ HỖ TRỢ 13](#_Toc25690)

[5.1 Biểu đồ cột chồng 13](#_Toc27256)

[5.2 Biểu đồ tròn 14](#_Toc21806)

[5.3 Biểu đồ diện tích (Area Chart) 15](#_Toc6021)

[5.4 Hàm render\_chart 16](#_Toc7178)

[5.5 Hàm clear\_chart\_area 17](#_Toc29458)

[CHƯƠNG 6 :PHẦN KẾT LUẬN 18](#_Toc10914)

[6.1 Về kiến thức: 18](#_Toc20138)

[6.2 Về sản phẩm: 18](#_Toc16476)

[6.3 Ưu và nhược điểm 19](#_Toc22546)

[6.4 Hướng phát triển. 19](#_Toc30714)

[6.5 Tổng kết 20](#_Toc6868)

[TÀI LIỆU THAM KHẢO 21](#_Toc17284)

**DANH SÁCH HÌNH**

[Hình 2.1: data 4](#_Toc185931753)

[*Hình 2:Hình ảnh code đọc dữ liệu* 6](#_Toc185931754)

[Hình 3: Hình ảnh code hiển thị data 7](#_Toc185931755)

[*Hình 4: Hình ảnh code đổ dữ liệu lên ô nhập* 8](#_Toc185931756)

[*Hình 5: Hình ảnh code chức năng thêm dữ liệu* 9](#_Toc185931757)

[*Hình 6: Hình ảnh code chức năng xóa dữ liệu* 9](#_Toc185931758)

[*Hình 7:Hình ảnh code chức năng tìm kiếm dữ liệu* 10](#_Toc185931759)

[*Hình 8:Hình ảnh code chức năng tìm kiếm dữ liệu* 11](#_Toc185931760)

[*Hình 9: Hình ảnh code hỗ trợ sắp xếp dữ liệu* 12](#_Toc185931761)

[*Hình 10: Hình ảnh code biểu đồ cột chồng* 13](#_Toc185931762)

[*Hình 11: Hình ảnh code biểu đồ tròn* 15](#_Toc185931763)

[*Hình 12: hình ảnh code biểu đồ diện tích* 16](#_Toc185931764)

[*Hình 13: Hình ảnh code hỗ trợ hiển thị biểu đồ* 17](#_Toc185931765)

[*Hình 14: Hình ảnh code hỗ trợ xóa widget con* 17](#_Toc185931766)

# PHẦN 1: MỞ ĐẦU

## 1.1 Mục tiêu

Đánh giá tối ưu hóa phương pháp giảng dạy toán học , để nghiên cứu và phân tích khả năng học toán của sinh viên đại học. Dữ liệu này hỗ trợ phát triển các phương pháp giảng dạy toán hiệu quả, cải thiện chiến lược học tập dựa trên đánh giá và phản hồi từ sinh viên. Ngoài ra, nó có thể được sử dụng trong các nghiên cứu trí tuệ nhân tạo để phân loại, hồi quy và phân cụm trong giáo dục toán học là mục tiêu của dữ liệu cũng như đề tài “Mathematics learning in higher education” nhóm em.

## 1.2 Đối tượng và phạm vi nghiên cứu

**Đối tượng nghiên cứu:**

* Tập thông tin của sinh viên: Bao gồm Id sinh viên , quốc gia .
* Tập thông tin về thông tin khảo sát: Bao gồm Id câu hỏi , loại câu trả lời (đúng/sai) , mức độ câu hỏi (cơ bản/nâng cao), chủ đề toán học , chủ đề con .

**Phạm vi nghiên cứu:**

* Phân tích dữ liệu : khai thác và hiển thị được dữ liệu từ tệp tin.
* Xây dựng giao diện : Thực hiện được cái thao tác cơ bản trên dữ liệu đồng thời xây dựng biểu đồ trực quan trên dữ liệu gốc .

# PHẦN 2: NỘI DUNG

## CHƯƠNG 1: THÔNG TIN DỮ LIỆU

### Nguồn gốc

MathE được phát triển và triển khai bởi một nhóm gồm bảy đối tác tổ chức từ năm quốc gia châu Âu: Học viện Bách khoa Bragança (Bồ Đào Nha), Học viện Công nghệ Limerick (Ireland), Đại học Genova, Pixel (Ý), Đại học Công nghệ Kaunas (Litva), Đại học Kỹ thuật Iasi (Romania) và EuroED (Romania) , Dữ liệu được thu thập từ tháng 2 năm 2019 đến tháng 12 năm 2023.

### Công dụng

Đánh giá học sinh: Học sinh có thể tự đánh giá kiến thức của mình và giáo viên có thể tổ chức bài kiểm tra trực tuyến về các chủ đề Toán đã chọn mục tiêu của phần tự đánh giá nhu cầu là cung cấp cho học viên bài đánh giá đào tạo để kiểm tra xem một chủ đề nào đó mà học viên đã đăng ký đã biết và hiểu chưa.

Là nơi gặp gỡ ảo nơi người dùng có thể chia sẻ kinh nghiệm giảng dạy và học tập. Theo cách này, một mạng lưới thực hành giảng dạy và học tập vững chắc sẽ được tạo ra, cung cấp một nơi mà giáo viên và học sinh có thể chia sẻ kinh nghiệm,kiến thức và thông tin của họ.

### Hạn chế

Một số câu hỏi đã được trả lời nhiều lần, trong khi một số khác chỉ được trả lời ít hoặc chỉ một lần.

Số lượng câu trả lời cho những câu hỏi cơ bản cao hơn nhiều so với số lượng câu trả lời cho những câu hỏi nâng cao.

Bộ dữ liệu không cân bằng theo quốc gia: phần lớn câu trả lời được cung cấp bởi sinh viên đến từ Bồ Đào Nha, Litva và Ý.

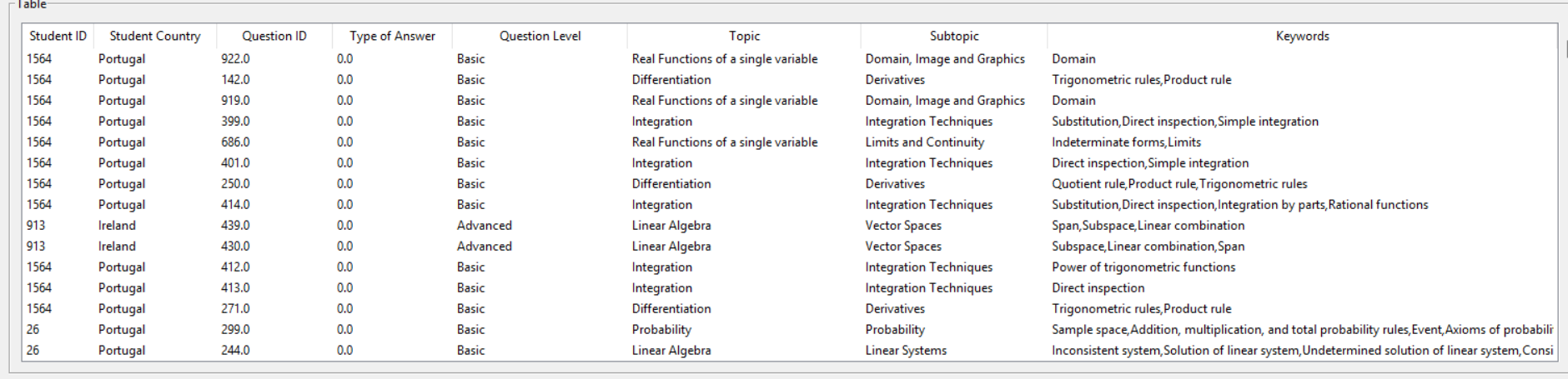
Giới hạn trong phân tích: Dữ liệu không cung cấp thông tin về các yếu tố ngoài toán học (như yếu tố tâm lý hoặc xã hội) ảnh hưởng đến kết quả học tập.

Dữ liệu không đầy đủ: Một số sinh viên có thể thiếu thông tin hoặc không tham gia đủ bài kiểm tra.

## CHƯƠNG 2: MÔ TẢ DỮ LIỆU

### 2.1 Các tập dữ liệu

|  |  |
| --- | --- |
| Thông tin | Mô tả |
| Tên tập dữ liệu | Dataset.csv |
| Kích thước | 8 cột và 9546 dòng |
| Loại dữ liệu | Số nguyên, số thực, danh mục , chuỗi ký tự |



Hình 2.1: data

Bộ dữ liệu có 9546 câu trả lời cho các câu hỏi trong các chủ đề Toán học được giảng dạy trong giáo dục đại học. Tệp có tám tính năng, được đặt tên là: Mã sinh viên, Quốc gia của sinh viên, Mã câu hỏi, Loại câu trả lời (đúng hoặc sai), Cấp độ câu hỏi (cơ bản hoặc nâng cao), Chủ đề toán học, Chủ đề toán học phụ và Từ khóa câu hỏi. Cấp độ câu hỏi được liên kết với giáo sư đã gửi câu hỏi.

## CHƯƠNG 3: PHÂN TÍCH DỮ LIỆU

### 3.1 Khởi tạo dữ liệu

### 3.1.1 Tkinter

**Chức năng:** Thư viện giao diện đồ họa tiêu chuẩn tích hợp sẵn trong Python. Nó được sử dụng để tạo các ứng dụng GUI (Graphical User Interface).

**Cụ thể:** tk: Tạo cửa sổ chính và quản lý các widget (nút, nhãn, hộp văn bản, v.v.).

**messagebox:** Hiển thị các hộp thoại thông báo, xác nhận, hoặc cảnh báo cho người dùng.

### 3.1.2 Pandas:

**Chức năng:** Thư viện mạnh mẽ để xử lý và phân tích dữ liệu dạng bảng

**Cụ thể:** Đọc và ghi dữ liệu từ các tệp CSV, Excel, hoặc SQL.

### 3.1.3 Os:

**Chức năng:** Thư viện tiêu chuẩn của Python để tương tác với hệ thống tệp và thư mục.

**Cụ thể:** Truy cập tệp, kiểm tra sự tồn tại của tệp, tạo hoặc xoá thư mục, v.v.

### 3.1.4 Matplotlib:

**Chức năng:** Thư viện trong python được sử dụng để vẽ và trực quan hóa dữ liệu

**Cụ thể:** Dữ liệu sẽ được xử lý (ví dụ, bằng pandas) và được hiển thị dưới dạng biểu đồ trong ứng dụng.

Sử dụng FigureCanvasTkAgg, biểu đồ sẽ được hiển thị trong một cửa sổ Tkinter

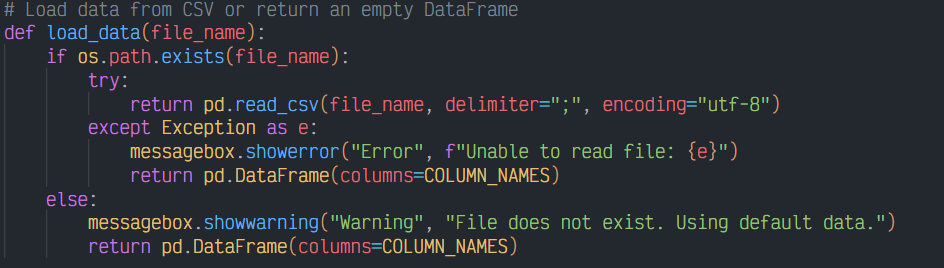
## CHƯƠNG 4: CÁC HÀM CHỨC NĂNG

### 4.1 Tải dữ liệu-load data

**Mô tả:** Hàm này sẽ tải dữ liệu từ tệp CSV và trả về một DataFrame của pandas. Nếu tệp không tồn tại hoặc không thể đọc được, hàm sẽ hiển thị thông báo lỗi hoặc cảnh báo và trả về một DataFrame trống với các cột mặc định.

**Chức năng:** Kiểm tra sự tồn tại của tệp.

* Nếu tệp tồn tại, cố gắng đọc dữ liệu từ tệp và trả về một DataFrame.
* Nếu không thể đọc dữ liệu (ví dụ: lỗi tệp), thông báo lỗi sẽ được hiển thị.



*Hình 2:Hình ảnh code đọc dữ liệu*

**Giải thích:**

* os.path.exists(file\_name): Kiểm tra xem tệp có tồn tại không.
* pd.read\_csv(): Đọc tệp CSV thành DataFrame.
* Nếu có lỗi khi đọc tệp, một thông báo lỗi được hiển thị và một DataFrame trống được trả về.

### 4.2 Hiển thị-display data

**Mô tả:** Hàm này hiển thị dữ liệu từ DataFrame trong widget Treeview. Dữ liệu có thể được sắp xếp theo cột chỉ định. Sau khi tải dữ liệu, hàm xóa tất cả các mục hiện có trong Treeview và chèn lại dữ liệu đã được sắp xếp.

**Chức năng:** Lấy dữ liệu từ tệp CSV thông qua load\_data().

* Sắp xếp dữ liệu theo cột được chỉ định (sort\_by).
* Xóa tất cả các mục trong Treeview và sau đó thêm dữ liệu đã sắp xếp vào.

A computer screen shot of text

Description automatically generated

Hình 3: Hình ảnh code hiển thị data

**Giải thích:**

* data.sort\_values(): Sắp xếp dữ liệu theo cột chỉ định.
* tree.delete(\*tree.get\_children()): Xóa tất cả các mục trong Treeview.
* tree.insert(): Thêm các hàng dữ liệu vào Treeview.

### 4.3 Tự động điển thông tin-auto fill

**Mô tả:** Hàm này được gọi khi người dùng chọn một mục trong Treeview. Nó sẽ tự động điền thông tin từ dòng đã chọn vào các ô nhập liệu (input fields).

**Chức năng:** Khi người dùng chọn một mục trong Treeview, hàm này sẽ điền các giá trị của dòng đó vào các trường nhập liệu.

A screen shot of a computer code

Description automatically generated

*Hình 4: Hình ảnh code đổ dữ liệu lên ô nhập*

**Giải thích:**

* tree.selection(): Lấy các mục được chọn trong Treeview.
* tree.item(selected\_items[0], "values"): Lấy các giá trị của mục được chọn.
* Dòng input\_fields[i].delete(0, "end") và input\_fields[i].insert(0, value) sẽ xóa và điền lại dữ liệu vào các ô nhập liệu tương ứng.

### 4.4 Thêm-add data

**Mô tả:** Hàm này sẽ lấy dữ liệu từ các ô nhập liệu và thêm một dòng mới vào DataFrame. Sau đó, nó lưu lại dữ liệu vào tệp CSV và làm mới dữ liệu hiển thị trong Treeview.

**Chức năng:** Thu thập dữ liệu từ các ô nhập liệu.

* Thêm dữ liệu vào DataFrame.
* Lưu lại dữ liệu vào tệp CSV.
* Cập nhật Treeview để hiển thị dữ liệu mới.

A computer code with colorful text

Description automatically generated

*Hình 5: Hình ảnh code chức năng thêm dữ liệu*

**Giải thích:**

* new\_data = [entry.get() for entry in input\_fields]: Lấy dữ liệu từ các ô nhập liệu.
* data.loc[len(data)] = new\_data: Thêm dòng mới vào DataFrame.
* data.to\_csv(): Lưu dữ liệu vào tệp CSV.

### 4.5 Xóa-delete data

**Mô tả:** Hàm này xóa các dòng dữ liệu được chọn trong Treeview. Sau khi xóa, dữ liệu sẽ được lưu lại vào tệp CSV và cập nhật lại Treeview.

**Chức năng:** Kiểm tra xem có mục nào được chọn trong Treeview.

* Xóa các mục đã chọn từ DataFrame.
* Lưu lại tệp CSV và cập nhật Treeview.

A screen shot of a computer code

Description automatically generated

*Hình 6: Hình ảnh code chức năng xóa dữ liệu*

**Giải thích:**

* tree.selection(): Lấy các mục được chọn trong Treeview.
* data.drop(): Xóa các dòng dữ liệu tương ứng với các chỉ số đã chọn.
* data.to\_csv(): Lưu lại dữ liệu sau khi xóa.

### 4.6 Tìm kiếm-Search data

**Mô tả:** Hàm này tìm kiếm dữ liệu dựa trên các giá trị trong các ô nhập liệu. Nó lọc dữ liệu và hiển thị các kết quả phù hợp trong Treeview.

**Chức năng:** Hiển thị các kết quả phù hợp trong Treeview.

* Lấy giá trị từ các ô nhập liệu và lọc dữ liệu trong DataFrame theo các giá trị tìm kiếm.

A screen shot of a computer code

Description automatically generated

*Hình 7:Hình ảnh code chức năng tìm kiếm dữ liệu*

**Giải thích:**

* search\_values: Lấy các giá trị tìm kiếm từ các ô nhập liệu.
* filtered\_data = filtered\_data[filtered\_data[COLUMN\_NAMES[i]].astype(str).str.contains(value, case=False, na=False)]: Lọc dữ liệu theo các giá trị tìm kiếm.
* tree.delete(\*tree.get\_children()): Xóa tất cả các mục trong Treeview.
* tree.insert(): Thêm các kết quả tìm kiếm vào Treeview.

### 4.7 Cập nhật- update

**Mô tả:** Hàm này xử lý việc cập nhật dữ liệu trong Treeview sau khi người dùng chỉnh sửa các trường nhập liệu và chọn một mục trong Treeview để cập nhật.

**Chức năng:** Kiểm tra xem người dùng có chọn mục nào trong Treeview hay không.

* Nếu có, lấy dữ liệu từ các trường nhập liệu và cập nhật dữ liệu tương ứng trong DataFrame.
* Lưu lại dữ liệu đã cập nhật vào tệp CSV và làm mới dữ liệu hiển thị trong Treeview.

A screen shot of a computer code

Description automatically generated

*Hình 8:Hình ảnh code chức năng tìm kiếm dữ liệu*

**Giải thích:**

* tree.selection(): Lấy các mục được chọn trong Treeview.
* updated\_data = [field.get() for field in input\_fields]: Lấy dữ liệu từ các trường nhập liệu.
* data.iloc[index] = updated\_data: Cập nhật dòng dữ liệu trong DataFrame tại chỉ số index.
* data.to\_csv(FILE\_NAME, sep=";", index=False, encoding="utf-8"): Lưu lại dữ liệu đã cập nhật vào tệp CSV.
* display\_data(): Cập nhật lại dữ liệu trong Treeview sau khi cập nhật.

### 4.8 Sắp xếp -sort

**Mô tả:**Hàm này sắp xếp dữ liệu trong Treeview theo cột được chỉ định và thay đổi thứ tự sắp xếp (tăng dần hoặc giảm dần).

**Chức năng:** Thay đổi thứ tự sắp xếp của cột mỗi khi người dùng nhấn vào tiêu đề cột.

*A screen shot of a computer code

Description automatically generated*

*Hình 9: Hình ảnh code hỗ trợ sắp xếp dữ liệu*

**Giải thích:**

* ascending\_order: Biến toàn cục lưu trữ thứ tự sắp xếp của từng cột.
* display\_data(): Hàm này được gọi lại để làm mới Treeview theo thứ tự mới của cột.

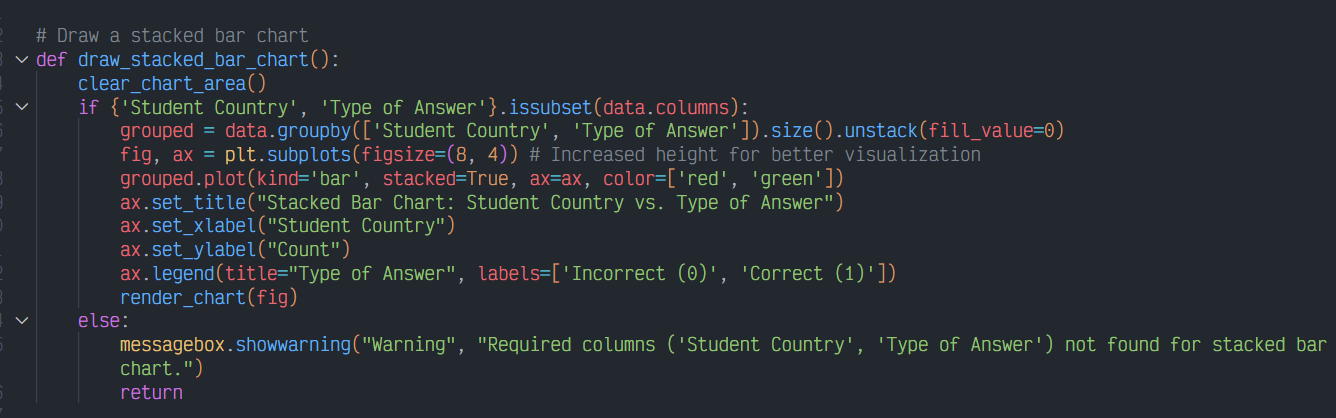
## CHƯƠNG 5: BIỂU ĐỒ HỖ TRỢ

### 5.1 Biểu đồ cột chồng

**Mô tả:** Hàm này vẽ một biểu đồ thanh chồng (stacked bar chart) để so sánh số lượng các loại câu trả lời (Correct/Incorrect) cho mỗi quốc gia học sinh (Student Country).

**Chức năng:**

* Kiểm tra cột: Hàm kiểm tra xem các cột 'Student Country' và 'Type of Answer' có tồn tại trong DataFrame data hay không. Nếu không, sẽ hiển thị thông báo cảnh báo.
* Nhóm và tính toán dữ liệu: Dữ liệu được nhóm theo hai cột 'Student Country' và 'Type of Answer'. Sau đó, hàm sử dụng unstack() để chuyển đổi dữ liệu thành dạng có thể vẽ biểu đồ thanh chồng, điền giá trị 0 cho các tổ hợp không có dữ liệu.
* Vẽ biểu đồ: Biểu đồ thanh chồng được vẽ với hai màu: red cho câu trả lời sai và green cho câu trả lời đúng.
* Cài đặt biểu đồ: Các tiêu đề cho trục X, trục Y, và tên của các legend được thiết lập.
* Hiển thị biểu đồ: Sau khi vẽ, biểu đồ sẽ được hiển thị thông qua hàm render\_chart().



*Hình 10: Hình ảnh code biểu đồ cột chồng*

**Giải thích:**

* clear\_chart\_area(): Trước khi vẽ biểu đồ mới, hàm này sẽ xóa các biểu đồ cũ đã được vẽ trong khu vực hiển thị.
* groupby(): Nhóm dữ liệu theo 'Student Country' và 'Type of Answer', sau đó sử dụng .size() để tính số lượng mỗi nhóm.
* unstack(): Chuyển đổi dữ liệu từ dạng dài (long) thành dạng rộng (wide), mỗi giá trị của 'Type of Answer' trở thành một cột riêng biệt.
* plot(kind='bar', stacked=True): Vẽ biểu đồ thanh chồng.
* render\_chart(fig): Vẽ biểu đồ lên giao diện.

### 5.2 Biểu đồ tròn

**Mô tả:**Hàm này vẽ một biểu đồ tròn để hiển thị tỷ lệ phần trăm của các mức độ câu hỏi khác nhau (Question Level) trong dữ liệu.

**Chức năng:**

* Kiểm tra cột:Hàm kiểm tra xem cột 'Question Level' có tồn tại trong DataFrame data hay không. Nếu không, sẽ hiển thị thông báo cảnh báo cho người dùng.
* Tính toán dữ liệu:Dữ liệu trong cột 'Question Level' được tính tần suất xuất hiện bằng value\_counts().
* Vẽ biểu đồ:Biểu đồ tròn được vẽ với tỷ lệ phần trăm (autopct='%1.1f%%') hiển thị trên từng phần của biểu đồ. Góc bắt đầu của biểu đồ được đặt ở 90 độ để sắp xếp dễ nhìn.
* Cài đặt biểu đồ:Tiêu đề biểu đồ được thiết lập là "Pie Chart: Question Levels" để mô tả nội dung.
* Hiển thị biểu đồ:Sau khi hoàn thành, biểu đồ được hiển thị thông qua hàm render\_chart()

A computer code on a black background

Description automatically generated

*Hình 11: Hình ảnh code biểu đồ tròn*

**Giải thích:**

* clear\_chart\_area():Xóa các biểu đồ cũ trong khu vực hiển thị trước khi vẽ biểu đồ mới.
* value\_counts():Đếm số lượng mỗi giá trị duy nhất trong cột 'Question Level'.
* ax.pie():Vẽ biểu đồ tròn và thêm nhãn (labels) cùng tỷ lệ phần trăm (autopct).
* render\_chart(fig):Hiển thị biểu đồ tròn trên giao diện ứng dụng.
* ax.pie(): Vẽ biểu đồ tròn và thêm nhãn (labels) cùng tỷ lệ phần trăm (autopct).
* render\_chart(fig): Hiển thị biểu đồ tròn trên giao diện ứng dụng.

### 5.3 Biểu đồ diện tích (Area Chart)

**Mô tả:**Hàm này vẽ một biểu đồ diện tích để biểu diễn số lượng câu hỏi thuộc các chủ đề (Topic) trong dữ liệu.

**Chức năng:**

* Kiểm tra cột: Hàm kiểm tra xem cột 'Topic' có tồn tại trong DataFrame data hay không. Nếu không, sẽ hiển thị thông báo cảnh báo.
* Tính toán dữ liệu: Dữ liệu trong cột 'Topic' được tính tần suất xuất hiện bằng value\_counts().
* Vẽ biểu đồ: Biểu đồ diện tích được vẽ bằng cách sử dụng fill\_between() để tô màu dưới đường biểu diễn.
* Cài đặt biểu đồ: Tiêu đề biểu đồ được đặt là "Area Chart: Topics", trục X hiển thị các chủ đề, và trục Y hiển thị số lượng.
* Hiển thị biểu đồ: Sau khi hoàn thành, biểu đồ được hiển thị trên giao diện thông qua hàm render\_chart().

A screen shot of a computer code

Description automatically generated

*Hình 12: hình ảnh code biểu đồ diện tích*

**Giải thích:**

* clear\_chart\_area():Xóa các biểu đồ cũ trong khu vực hiển thị trước khi vẽ biểu đồ mới.
* value\_counts():Đếm số lượng mỗi giá trị duy nhất trong cột 'Topic'.
* ax.fill\_between():Tô màu dưới đường biểu diễn để tạo biểu đồ diện tích.
* ax.plot():Vẽ đường biểu diễn trên biểu đồ để làm nổi bật xu hướng.
* render\_chart(fig):Hiển thị biểu đồ diện tích trên giao diện ứng dụng.

### 5.4 Hàm render\_chart

**Mô tả:** Hiển thị biểu đồ matplotlib trong giao diện Tkinter

**Chức năng:**

* Xóa nội dung hiện có trong khung hiển thị trước khi vẽ biểu đồ mới.
* Tạo FigureCanvasTkAgg để kết nối biểu đồ matplotlib với Tkinter.
* Hiển thị biểu đồ trong khung.

A screen shot of a computer code

Description automatically generated

*Hình 13: Hình ảnh code hỗ trợ hiển thị biểu đồ*

**Giải thích :**

* fig (matplotlib.figure.Figure): Đối tượng biểu đồ matplotlib cần hiển thị.
* chart\_display\_frame (tkinter.Frame): Khung giao diện để hiển thị biểu đồ.

### 5.5 Hàm clear\_chart\_area

**Mô tả:** Làm sạch khung hiển thị biểu đồ, xóa tất cả các widget con trong khung.

**Chức năng:** Duyệt qua tất cả các widget con trong khung và hủy chúng.A screenshot of a computer program

Description automatically generated

*Hình 14: Hình ảnh code hỗ trợ xóa widget con*

**Giải thích:** chart\_display\_frame (tkinter.Frame): Khung cần được làm sạch.

## CHƯƠNG 6 :PHẦN KẾT LUẬN

### 6.1 Về kiến thức:

**Dựa trên tập dữ liệu:**

* Phân tích dữ liệu lớn: Code cung cấp khả năng thao tác và hiển thị trên tập dữ liệu CSV có cấu trúc. Qua đó, giúp hiểu được các thành phần quan trọng của dữ liệu, mối quan hệ giữa các yếu tố như "Student Country", "Type of Answer", "Question Level",...
* Phát hiện lỗi: Khi xử lý các dữ liệu lớn, việc kiểm tra cẩn thận từng bước như đọc file, lọc dữ liệu và hiển thị giúp người dùng nhận biết các lỗi phổ biến trong dữ liệu hoặc trong quá trình nhập liệu, từ đó áp dụng biện pháp sửa lỗi.
* Tạo trực quan hóa: Hỗ trợ việc tạo các biểu đồ như stacked bar chart, pie chart, và area chart. Điều này cải thiện khả năng nhìn nhận tổng quan các xu hướng dữ liệu và các điểm bất thường.
* Tối ưu hiệu năng: Khi làm việc với tập dữ liệu lớn, tốc độ xử lý hiện tại phụ thuộc vào pandas và matplotlib. Hệ thống vẫn chưa tối ưu để xử lý nhanh với tập dữ liệu cực lớn, đặc biệt trong các khâu tìm kiếm, sắp xếp, và cập nhật.

### 6.2 Về sản phẩm:

* Quản lý CRUD: Ứng dụng cung cấp tính năng thêm, sửa, xóa, và tìm kiếm trực quan giúp quản lý thông tin dữ liệu dễ dàng hơn.
* Trực quan hóa dễ hiểu: Các biểu đồ minh họa được tích hợp trực tiếp trong ứng dụng. Điều này giúp người dùng không cần chuyển đổi qua công cụ khác để phân tích trực quan dữ liệu.
* Khả năng sắp xếp: Hỗ trợ sắp xếp cột theo ý muốn, thuận tiện khi so sánh hoặc phân tích nhanh các giá trị.

### 6.3 Ưu và nhược điểm

**Ưu điểm:**

Ứng dụng toàn diện:

* Hỗ trợ đầy đủ các tính năng quản lý dữ liệu (CRUD).
* Hiển thị và trực quan hóa trực tiếp trong ứng dụng mà không cần phần mềm ngoài.

Tính dễ sử dụng:

* Giao diện đồ họa trực quan (Tkinter).
* Tích hợp các thông báo lỗi hoặc cảnh báo giúp người dùng dễ dàng xác định các bước cần điều chỉnh.

**Nhược điểm:**

Hiệu năng:

* Khả năng xử lý tập dữ liệu lớn vẫn còn hạn chế, gây chậm khi số lượng dòng tăng đáng kể.

Chưa có kiểm tra đầu vào:

* Việc nhập liệu thủ công có thể gây lỗi (như thiếu hoặc sai định dạng dữ liệu), ảnh hưởng đến quá trình xử lý và trực quan hóa.

### 6.4 Hướng phát triển.

Tối ưu hiệu năng: Ứng dụng các thư viện khác để tối ưu hiệu năng

Tăng tính linh hoạt: Thêm tính năng import dữ liệu từ nhiều định dạng khác như Excel, JSON, hoặc kết nối cơ sở dữ liệu (MySQL, SQLite).

Kiểm tra và chuẩn hóa dữ liệu: Tích hợp thêm các bước kiểm tra đầu vào hoặc gợi ý tự động để tránh nhập sai dữ liệu.

Ứng dụng vào thực tế: Sản phẩm có thể được phát triển thành một công cụ phân tích giáo dục, giúp quản lý điểm số hoặc phân tích xu hướng học tập của học sinh.

### 6.5 Tổng kết

Code hiện tại là một nền tảng tốt cho việc quản lý và phân tích dữ liệu cơ bản. Tuy nhiên, cần cải thiện về hiệu năng và tính linh hoạt để đáp ứng tốt hơn cho các yêu cầu thực tế phức tạp hơn.

## TÀI LIỆU THAM KHẢO

Nguồn data: [Dataset for Assessing Mathematics Learning in Higher Education - UCI Machine Learning Repository](https://archive.ics.uci.edu/dataset/1031/dataset+for+assessing+mathematics+learning+in+higher+education)

Git : Saokhue8326/math-assetment-crud-analysis: Đồ án Python: Phân tích và xử lý dữ liệu