Đại học Quốc gia Thành phố Hồ Chí Minh Trường Đại học Bách Khoa Khoa Khoa học và Kỹ thuật Máy tính



LẬP TRÌNH NÂNG CAO (CO2039)

Bài Tập Lớn

XÂY DỰNG APP TRA CỨU SAO KÊ

GVHD: Lê Đình Thuận

Lớp-Nhóm: L01-503

SVTH: Võ Trần Phi Phong

MSSV: 2420006



Contents

1	Nội Dung Đề Tài	2
2	Phương Pháp Tiếp Cận 2.1 Ngôn Ngữ	2 2 2
3	Cấu trúc dữ liệu	3
4	Thuật toán	4
5	Thiết lập môi trường	5
6	Demo ứng dụng	5

1 Nội Dung Đề Tài

Xây dựng một web/app tra cứu dữ liệu sao kê của MTTQ VN với khoảng 200.000 bản ghi. Ứng dụng cần có giao diện (web/app/CLI) có thể tra cứu thông tin sao kê theo: số tiền, tên người gửi, nội dung.

2 Phương Pháp Tiếp Cận

2.1 Ngôn Ngữ

Python rất phù hợp để tìm kiếm dữ liệu trong tệp CSV có khoảng 200.000 bản ghi. Mặc dù 200.000 bản ghi là một lượng dữ liệu đáng kể, nhưng các thư viện của Python có thể xử lý một cách hiệu quả.

- 1. Pandas: là một thư viện tuyệt vời để làm việc với các tập dữ liệu lớn. Nó có thể đọc nhanh các tệp CSV vào DataFrame, cho phép thao tác và tìm kiếm dữ liệu nhanh chóng. Pandas có thể xử lý dữ liệu với hàng triệu hàng và các tối ưu hóa như sử dụng dtype có thể cải thiện hiệu suất.
- 2. CSV module: là một thư viện built-in của Python, làm việc với dữ liệu CSV thô mà không cần tải toàn bộ tập dữ liệu vào bộ nhớ bằng cách đọc dữ liệu theo từng dòng. Phương pháp này sử dụng ít bộ nhớ hơn nhưng có thể chậm hơn nếu tìm kiếm nhiều điều kiện hoặc thực hiện các truy vấn phức tạp.
- 3. SQLite: để có chức năng tìm kiếm nâng cao hơn và tối ưu hóa, tải dữ liệu CSV vào cơ sở dữ liêu SQLite. SQLite nhe và cho phép truy vấn SQL nhanh.

2.2 Thư Viện

CSV module là một thư viện Python tích hợp được sử dụng để đọc và ghi tệp CSV.

• csv.DictReader: đọc tệp CSV và trả về mỗi hàng dưới dạng dictionary, trong đó key là tên cột và value là giá trị hàng. Điều này hữu ích khi xử lý dữ liệu CSV một cách linh động.

OS module cung cấp một cách thức để tương tác với hệ điều hành, chẳng hạn như đọc đường dẫn file, thư mục và các chức năng khác ở cấp độ hệ điều hành.

• os.path.dirname và os.path.abspath: tìm đường dẫn tuyệt đối của thư mục hiện tại và phục vụ tệp index.html một cách chính xác.

Flask là một web framework nhẹ cho Python được sử dụng để tạo các ứng dụng web. Trong ứng dụng này, Flask được sử dụng để tạo các API endpoint, xử lý các HTTP request, xử lý dữ liệu từ file CSV và trả kết quả cho người dùng. Flask đóng vai trò như một backbone của máy chủ web, quản lý các yêu cầu từ người dùng và trả về các phản hồi phù hợp.

- Flask: core class, được sử dụng để khởi tạo ứng dụng
- request: xử lý dữ liệu HTTP request (ví du: tham số truy vấn)



- jsonify: chuyển đổi các đối tượng Python (như dictionary) thành kiểu dữ liệu JSON
- send_from_directory: phục vụ các tệp tĩnh (static), trong trường hợp này là index .html

3 Cấu trúc dữ liệu

Một file CSV (Comma Separated Values) là tệp văn bản chứa dữ liệu theo định dạng có cấu trúc, trong đó mỗi giá trị được phân cách bằng dấu phẩy. Mỗi dòng biểu diễn một bản ghi và mỗi giá trị trong dòng tương ứng với một trường trong bản ghi.

Ví dụ data từ file CSV: date_time, trans_no, credit, debit, detail là các header.

```
date_time, trans_no, credit, debit, detail
03/09/2024_5216.65140,9,200000,0,120167.030924.100642.NGUYEN
HOAI NAM ung ho
03/09/2024_5017.43849,13,9000,0,888828.030924.121121.NGUYEN
THI KIEU OANH chuyen tien
04/09/2024_5240.80637,55,500000,0,MBVCB.6944619812.chung tay
gop suc .CT tu 0351000801710 TRAN DANH VU toi
0011001932418 MAT TRAN TO QUOC VN - BAN CUU TRO TW
```

Đối tượng csv.DictReader là một iterator đọc file CSV và ánh xạ từng hàng vào một dictionary trong đó key là tên cột và value là dữ liệu tương ứng trong hàng đó. Ngoài ra, đối tương fieldnames trả về danh sách tên cột xuất hiện ở hàng đầu tiên của file CSV.

Dữ liệu được đọc trong Python:

```
{
  "date_time": "03/09/2024_5216.65140",
  "trans_no": "9",
  "credit": "200000",
  "debit": "0",
  "detail": "120167.030924.100642.NGUYEN HOAI NAM ung ho"
},
  "date_time": "03/09/2024_5017.43849",
  "trans_no": "13",
  "credit": "9000",
  "debit": "0",
  "detail": "888828.030924.121121.NGUYEN THI KIEU OANH chuyen
      tien"
},
  "date_time": "04/09/2024_5240.80637",
  "trans_no": "55",
  "credit": "500000",
  "debit": "0",
  "detail": "MBVCB.6944619812.chung tay gop suc .CT tu
     0351000801710 TRAN DANH VU toi 0011001932418 MAT TRAN TO
      QUOC VN - BAN CUU TRO TW"
}
```



Các key biểu thị tên cột là điều kiện tìm kiếm (như credit, detail) và các giá trị là các giá trị cần tìm kiếm. Kết quả tìm kiếm được lưu trữ trong danh sách dictionary. Mỗi dictionary đại diện cho một hàng từ file CSV phù hợp với điều kiện tìm kiếm.

Flask cung cấp jsonify để chuyển đổi các cấu trúc dictionary sang định dạng JSON.

4 Thuật toán

Đọc dữ liệu CSV: csv.DictReader đọc file CSV và lưu trữ từng hàng dưới dạng dictionary với tên cột là key và giá trị ô là value. Các hàng được lặp lại và mỗi hàng được kiểm tra để khớp với các điều kiện.

Lọc tìm kiếm: ứng dụng lọc dữ liệu dựa trên các điều kiện tìm kiếm bằng cách sử dụng phương pháp **linear search**. Đối với mỗi hàng trong CSV, ứng dụng sẽ kiểm tra xem hàng đó có đáp ứng tất cả các điều kiện do người dùng chỉ định hay không. Các bước thực hiện:

- Lặp lại qua từng hàng của tệp CSV.
- Đối với mỗi điều kiện trong search_conditions, kiểm tra xem hàng đó có đáp ứng điều kiện không.
 - Đối với các chuỗi khớp chính xác (detail, trans_no), hãy kiểm tra xem giá trị của cột có chứa thuật ngữ tìm kiếm không.
 - Đối với các phạm vi (credit_min, credit_max), kiểm tra xem giá trị có nằm trong phạm vi đã chỉ định không.
 - Nếu hàng khớp với tất cả các điều kiện, hàng đó sẽ được thêm vào danh sách kết quả.

```
with open(file_path, mode="r", newline="") as file:
    reader = csv.DictReader(file)
    results = []

for row in reader:
    match = True
    for column_name, search_term in search_conditions.items()
    :
    if column_name == "credit_min":
        if float(row["credit"]) < search_term:
            match = False
            break
    elif column_name == "credit_max":
        if float(row["credit"]) > search_term:
            match = False
            break
        break
```



```
else:
    if search_term not in row[column_name].lower():
        match = False
        break
    if match:
        results.append(row)
```

Time Complexity:

- Đọc file CSV: time complexity là O(n), trong đó n là số hàng trong file CSV.
- Tìm kiếm: quá trình tìm kiếm bao gồm việc lặp lại từng hàng và kiểm tra từng điều kiện tìm kiếm. Time complexity trong trường hợp xấu nhất là $O(n \times m)$, trong đó n là số hàng và m là số điều kiện trong search_conditions.

5 Thiết lập môi trường

Trong Python, môi trường ảo (virtual environment) là một thư mục độc lập chứa cài đặt Python và các packages hoặc dependencies. Nó cho phép các nhà phát triển quản lý các dependencies cho các dự án khác nhau một cách riêng biệt, đảm bảo rằng mỗi dự án có môi trường riêng với các phiên bản thư viện cụ thể mà nó cần. Sự cô lập này (isolation) giúp tránh xung đột giữa các packages hoặc dependencies giữa các dự án khác nhau.

Trong thư mục chứa dự án, tiến hành mở Terminal:

- 1. Cài đặt môi trường ảo (venv): \$ sudo apt install python3-venv
- 2. Tao một môi trường ảo mới có tên flask_env: \$ python3 -m venv flask_env
- 3. Kích hoat: \$ source flask_env/bin/activate
- 4. Cài đặt Flask trong môi trường ảo: (flask_env)\$ pip install Flask
- 5. Chay ứng dụng Python: (flask_env)\$ python3 app.py
- 6. Để thoát khỏi môi trường ảo: (flask_env)\$ deactivate

6 Demo ứng dụng

Ủng dụng có các chức năng cơ bản như tìm kiếm thông tin theo tên người gửi, số tiền gửi, ngày gửi, và sắp xếp giá trị tăng dần/giảm dần theo cột.

"date_time"	Transaction Number
09/09/2024	Enter transaction number
Credit Amount	Debit Amount
Enter a range (e.g., 100-100 or 100-500)	Enter debit amount
Transaction Detail	
Enter transaction detail	
Search	

credit	debit	detail	trans_no	"date_time"
55555	0	PARTNER.DIRECT_DEBITS_VCB.MSE.6637 7041397.20240908.66377041397-0988549441_ Mong moi dieu tot dep a	687	09/09/2024_5161.90230
200000	0	020097042209082241272024U3UA760455.7997 7.224128.TRAN THI HANH TRINH ho tro ba con khac phuc bao YAGI	688	09/09/2024_5389.79977
2000000	0	MBVCB.6979460863.PHAM THI LIEN chuyen tien.CT tu 0351000813059 PHAM THI LIEN toi 0011001932418 MAT TRAN TO QUOC VN - BAN CUU TRO TW	689	09/09/2024_5242.92630
200000	0	MBVCB.6979453135.Ung ho khac phuc hau qua bao so 3 Yagi.CT tu 1024149794 KHUONG NGOC BICH toi 0011001932418 MAT TRAN TO QUOC VN - BAN CUU TRO TW	690	09/09/2024_5241.92651
100000	0	304088.080924.224230.HA THI HAI chuyen tien ung ho khac phuc bao lu	691	09/09/2024_5017.66459
100000	0	MBVCB.6979454212.PHAM XUAN TRUNG ung ho khac phuc bao so 3.CT tu 1019885185 PHAM XUAN TRUNG toi 0011001932418 MAT TRAN TO QUOC VN - BAN CUU TRO TW	692	09/09/2024_5244.92702
2000000	0	622852.080924.224307.HA THU HIEN UNG HO KHAC PHUC HAU QUA BAO SO 3 FT24253830067707	693	09/09/2024_5219.67051
500000	0	020097042209082243282024JU0W958288.8196 3.224329.DO THI Y PHUONG xin duoc gui de cung moi nguoi khac phuc sau bao yagi	694	09/09/2024_5189.81963
100000	0	623580.080924.224331.Cuu tro bao so 3 FT24253997709200	695	09/09/2024 5017.68139

Hình 1: Tìm kiếm tất cả giao dịch trong ngày 09/09/2024



"date_time"	Transaction Number
Enter value for "date_time"	Enter transaction number
Credit Amount	Debit Amount
Enter a range (e.g., 100-100 or 100-500)	Enter debit amount
Transaction Detail	
NGUYEN THI THU TRANG	
Search	

credit	debit	detail	trans_no	"date_time"
2000	0	913031.090924.164004.NGUYEN THI THU TRANG chuyen tien	2418	09/09/2024_5216.56265
1000000	0	722887.090924.185123.NGUYEN THI THU TRANG chuyen tien ung ho khac phuc bao so 3 yogi	3178	09/09/2024_5218.72592
1000000	0	020097042209091941052024A8TO654923.7947 1.194058.Nguyen Thi Thu Trang ung ho khac phuc hau qua bao Yagi	3841	09/09/2024_5388.79471
200000	0	0200970405090920245820246J9F078865.87929. 202458.Vietcombank:0011001932418:nguyen thi thu trang ung ho khac phuc sau bao	4559	09/09/2024_5388.87929
2000000	0	328751.090924.211759.NGUYEN THI THU TRANG ung ho cuu tro bao lu Yaki FT24253823407525	5613	09/09/2024_5215.62154
100000	0	575660.090924.224958.NGUYEN THI THU TRANG xin quyen gop chia se dong hanh dot bao Yagi va lu lut mong moi nguoi duoc binh an	6974	10/09/2024_5214.61263
100000	0	371515.090924.231015.Nguyen Thi Thu Trang ung ho khac phuc bao Yagi	7694	10/09/2024_5212.87026
200000	0	626412.090924.232612.NGUYEN THI THU TRANG ung ho dong bao Viet Nam FT24254828980711	8206	10/09/2024_5219.3670
1000000	0	670731.100924.000432.NGUYEN THI THU TRANG chuyen ung ho khac phuc sau bao FT24254430017396	9489	10/09/2024_5216.36675
200000	0	020097042209100016572024KBD5373221.2911. 001659.NGUYEN THI THU TRANG ung ho dong	9905	10/09/2024_5189.2911

Hình 2: Tìm kiếm theo tên người gửi NGUYEN THI THU TRANG



"date_time"	Transaction Number
Enter value for "date_time"	Enter transaction number
Credit Amount	Debit Amount
Enter a range (e.g., 100-100 or 100-500)	Enter debit amount
Transaction Detail	
NGUYEN THI THU TRANG	
Search	

credit	debit	detail	trans_no	"date_time"
100000000	0	MBVCB.6990563464.NGUYEN THI THU TRANG chuyen tien.CT tu 0081001211057 NGUYEN THI THU TRANG toi 0011001932418 MAT TRAN TO QUOC VN - BAN CUU TRO TW	28084	10/09/2024_5240.14689
50000000	0	MBVCB.6994047786.Gia dinh BS Duc Megadom ung ho dong bao vung lu phia Bac.CT tu 0021000409461 NGUYEN THI THU TRANG toi 0011001932418 MAT TRAN TO QUOC VN - BAN CUU TRO TW	88544	10/09/2024_5244.26741
30000000	0	MBVCB.6997698076. Ung ho dong bao lu lut.CT tu 0491000019713 NGUYEN THI THU TRANG toi 0011001932418 MAT TRAN TO QUOC VN - BAN CUU TRO TW	157033	10/09/2024_5245.44871
10000000	0	MBVCB.6989924341.NGUYEN THI THU TRANG BAC NINH UNG HO DONG BAO CHIU A.H CUA CON BAO SO 3.CT tu 0351001168368 NGUYEN THI THU TRANG toi 0011001932418 MAT TRAN TO QUOC VN - BAN CUU TRO TW	22458	10/09/2024_5242.94711
6000000	0	278327.100924.130659.NGUYEN THI THU TRANG uh khac phuc hau qua bao Yagi FT24254078537055	54749	10/09/2024_5214.79616
2000000	0	328751.090924.211759.NGUYEN THI THU TRANG ung ho cuu tro bao lu Yaki FT24253823407525	5613	09/09/2024_5215.62154
2000000	0	961357.100924.080000.NGUYEN THI THU TRANG Chuyen tien	16898	10/09/2024_5017.40056
2000000	0	MBVCB.6992430751.ung ho cuu tro bao Yagi.CT tu 0201000590007 NGUYEN THI THU TRANG toi 0011001932418 MAT TRAN TO QUOC VN - BAN CUU TRO TW	60946	10/09/2024_5243.73927
2000000	0	MBVCB.6997811566.ung ho bao lu.CT tu 0851000011014 NGUYEN THI THU TRANG toi 0011001932418 MAT TRAN TO QUOC VN - BAN CUU TRO TW	160468	10/09/2024_5245.48120
1000000	0	722887.090924.185123.NGUYEN THI THU TRANG chuyen tien ung ho khac phuc bao so 3 yogi	3178	09/09/2024_5218.72592
1000000	0	020097042209091941052024A8TO654923.7947 1.194058.Nguyen Thi Thu Trang ung ho khac phuc hau qua bao Yagi	3841	09/09/2024_5388.79471

Hình 3: Sắp xếp giảm dần theo cột credit

		Search CSV D	ata		
"date_tir	ne"	Transaction	Number		
Enter v	/alue for	"date_time" Enter trans	saction number		
Credit A	mount	Debit Amou	nt		
50000-	-50000	Enter debi	t amount		
Transacti	ion Deta	ail			
NGUYE	EN THI T	THU TRANG			
10 recor	debit			trans_no	"date_time"
50000	0	771005.100924.150724.NGUYEN THI THU TRANG Chuyen tien		76851	10/09/2024_5217.22533
50000	0	820652.100924.151906.NGUYEN THI THU TRANG chuyen FT2425468080184	45	78836	10/09/2024_5220.71861
50000	0	MBVCB.6993910987.NGUYEN THI THU TRANG UNG HO DONG BAO.CT tu 1 THU TRANG toi 0011001932418 MAT TRAN TO QUOC VN - BAN CUU TRO		86526	10/09/2024_5241.22391
50000	0	327182.100924.172042.NGUYEN THI THU TRANG chuyen khoan		99538	10/09/2024_5017.40473
50000	0	MBVCB.6995675445.NGUYEN THI THU TRANG chuyen tien.CT tu 07310009. TRANG toi 0011001932418 MAT TRAN TO QUOC VN - BAN CUU TRO TW	28251 NGUYEN THI THU	108968	10/09/2024_5239.79987

Hình 4: Tìm kiếm theo tên người gửi NGUYEN THI THU TRANG với số tiền 50000

 $020097042209101952472024 \text{K4RM444172.6963} \ 3.195248. \text{NGUYEN THI THU TRANG chuyen tien} \\$

MBVCB.6997255888.NGUYEN THI THU TRANG chuyen tien.CT tu 1039488032 NGUYEN THI THU

02009704220910202447202456TW517634.7742 3.202448.NGUYEN THI THU TRANG chuyen tien

668703.100924.213024.NGUYEN THI THU TRANG ung ho dong bao sau bao yagi FT24255556078905

332350.100924.201226.NGUYEN THI THU TRANG ung ho khac phuc bao FT24254764025466

TRANG toi 0011001932418 MAT TRAN TO QUOC VN - BAN CUU TRO TW

50000 0

50000 0

50000 0

50000 0

50000 0

135900

142421

143421

146764

171893

10/09/2024_5390.69633

10/09/2024_5218.96892

10/09/2024_5240.30354

10/09/2024_5189.77423

10/09/2024_5216.14257



"date_time"	Transaction Number
Enter value for "date_time"	Enter transaction number
Credit Amount	Debit Amount
50000-100000	Enter debit amount
Transaction Detail	
NGUYEN THI THU TRANG	
Search	

credit	debit	detail	trans_no	"date_time"
100000	0	575660.090924.224958.NGUYEN THI THU TRANG xin quyen gop chia se dong hanh dot bao Yagi va lu lut mong moi nguoi duoc binh an	6974	10/09/2024_5214.61263
100000	0	371515.090924.231015.Nguyen Thi Thu Trang ung ho khac phuc bao Yagi	7694	10/09/2024_5212.87026
100000	0	MBVCB.6988439781.Ung ho khac phuc hau qua bao Yagi.CT tu 1021525599 NGUYEN THI THU TRANG toi 0011001932418 MAT TRAN TO QUOC VN - BAN CUU TRO TW	10520	10/09/2024_5242.53189
100000	0	MBVCB.6988484506.YTT ung ho khac phuc hau qua Bao so 3 Yagi.CT tu 9938829536 NGUYEN THI THU TRANG toi 0011001932418 MAT TRAN TO QUOC VN - BAN CUU TRO TW	11517	10/09/2024_5240.54325
100000	0	MBVCB.6988484978.YNY ung ho khac phuc hau qua Bao so 3 Yagi.CT tu 9938829536 NGUYEN THI THU TRANG toi 0011001932418 MAT TRAN TO QUOC VN - BAN CUU TRO TW	11542	10/09/2024_5244.54349
100000	0	MBVCB.6988536976. Mong moi nguoi deu binh an .CT tu 1042208218 NGUYEN THI THU TRANG toi 0011001932418 MAT TRAN TO QUOC VN - BAN CUU TRO TW	12505	10/09/2024_5240.55404
100000	0	MBVCB.6988567005.YTT ung ho khac phuc hau qua Bao Yagi.CT tu 0181003437838 NGUYEN THI THU TRANG toi 0011001932418 MAT TRAN TO QUOC VN - BAN CUU TRO TW	13124	10/09/2024_5241.56177
100000	0	MBVCB.6988626611.NGUYEN THI THU TRANG chuyen tien ung ho dong bao lu lut mien bac.CT tu 1018518134 NGUYEN THI THU TRANG toi 0011001932418 MAT TRAN TO QUOC VN - BAN CUU TRO TW	13818	10/09/2024_5078.38818
75000	0	020097042209100818092024LHV9692002.24343 .081810.NGUYEN THI THU TRANG chuyen tien	17791	10/09/2024_5189.24343
100000	0	$020097042209100827452024l52B819713.38660.\ 082737. NGUYEN\ THI\ THU\ TRANG\ chuyen\ tien\ ung\ ho$	18321	10/09/2024_5189.38660
70000	0	MBVCB.6991191760.NGUYEN THI THU TRANG chuyen tien.CT tu 1036219816 NGUYEN THI THU TRANG toi 0011001932418 MAT TRAN TO QUOC VN - BAN CUU TRO TW	37212	10/09/2024_5239.34197
100000	0	986413.100924.115609.NGUYEN THI THU TRANG Chuyen tien	41406	10/09/2024_5218.8916
100000	0	MBVCB.6992180586.NGUYEN THI THU TRANG chuyen tien ung ho dong bao cac tinh bi thiet hai.CT tu 1032282623 NGUYEN THI THU TRANG toi 0011001932418 MAT TRAN TO QUOC VN - BAN CUU TRO TW	55478	10/09/2024_5244.66035
100000	0	MBVCB.6992242464.chung tay dong gop cung dong bao to quoc chut nho.CT tu 0381000415593 NGUYEN THI THU TRANG toi 0011001932418 MAT TRAN TO QUOC VN - BAN CUU TRO TW	57000	10/09/2024_5243.68002
100000	0	324506.100924.131951.NGUYEN THI THU TRANG chuyen tien ung ho dong bao bi bao Yagi FT24254279870312	57147	10/09/2024_5214.21104
100000	0	MBVCB.6992388660. ung ho phong chong bao lu mien bac.CT tu 1013115862 NGUYEN THI THU TRANG toi 0011001932418 MAT TRAN TO QUOC VN - BAN CUU TRO TW	59884	10/09/2024_5244.72430
100000	0	020097042209101336442024IX8W363178.2067 3.133645.NGUYEN THI THU TRANG chuyen tien	60029	10/09/2024_5387.20673
100000	0	020097042209101431422024UD5G271468.5053 2.143123.NGUYEN THI THU TRANG chuyen tien	70640	10/09/2024_5390.50532
50000	0	771005.100924.150724.NGUYEN THI THU TRANG Chuyen tien	76851	10/09/2024_5217.22533
50000	0	820652.100924.151906.NGUYEN THI THU TRANG chuven FT24254680801845	78836	10/09/2024 5220.71861

Hình 5: Tìm kiếm theo tên người gửi NGUYEN THI THU TRANG với số tiền từ 50000 đến 100000



"date_tim	ie"		Transaction Number		
Enter va	alue for '	date_time"	4559		
Credit An	nount		Debit Amount		
Enter a	range (e	g., 100-100 or 100-500)	Enter debit amount		
Transactio	on Detai				
NGUYE	N THI TH	U TRANG			
Search	s) for used				
1 record(s) iouna				
1 record(detail		trans_no	"date_time"

Hình 6: Tìm kiếm theo tên người gửi NGUYEN THI THU TRANG với mã số $4559\,$