

ĐẠI HỌC QUỐC GIA TP HCM
TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN
KHOA CÔNG NGHỆ THÔNG TIN



Báo cáo Bài tập thực hành
LAB-02: LÀM VIỆC VÀ TRỰC QUAN HÓA DỮ LIỆU
CHUỖI THỜI GIAN

CHỦ ĐỀ: Phân tích giá cổ phiếu của top 5 ngân hàng có giá trị thương hiệu lớn nhất Việt Nam

Môn học: Trực quan hóa dữ liệu

Nhóm thực hiện:

Nhóm 23

Giảng viên hướng dẫn:

Lê Nhựt Nam

Ngày 4 tháng 6 năm 2024

LỜI CẢM ƠN

Chúng em xin chân thành cảm ơn các thầy cô đã tận tình hướng dẫn và nhận xét, góp ý phù hợp để nhóm chúng em có thể hoàn thành đồ án này.

Trân trọng.

TP. Hồ Chí Minh, ngày 04 tháng 06 năm 2024

NHÓM SINH VIÊN THỰC HIỆN

Đại diện

Lê Thị Minh Phương

Mục lục

1	Chủ đề	3
2	Thông tin nhóm	3
3	Yêu cầu bài tập	3
4	Sơ lược pha "Khám phá các mối quan hệ trong dữ liệu"	4
5	Phân công công việc	5
6	Timeline công việc	6
7	Mức độ hoàn thành	6
7.1	Mức độ hoàn thành tổng thể của các yêu cầu	6
7.2	Mức độ hoàn thành của các thành viên	7
8	Tự đánh giá theo các yêu chí của bài tập	7
	Tài liệu	8

1 Chủ đề

Tên chủ đề: Phân tích giá cổ phiếu của top 6 ngân hàng có giá trị thương hiệu lớn nhất Việt Nam

Nguồn dữ liệu: [Investing.com](https://www.investing.com)

2 Thông tin nhóm

Tên nhóm: Nhóm 23

Danh sách thành viên

MSSV	Họ và tên
20120454	Lê Công Đất
21120279	Lê Trần Minh Khuê
21120290	Hoàng Trung Nam
21120296	Lê Trần Như Ngọc
21120533	Lê Thị Minh Phương

Bảng 1: Danh sách thành viên nhóm 23

Nhóm trưởng: Lê Thị Minh Phương

3 Yêu cầu bài tập

Trong bài tập thực hành này, sinh viên tìm hiểu mối quan hệ giữa các trường dữ liệu thực tế thông qua việc trực quan hóa dữ liệu. Sinh viên thực hiện tìm kiếm dữ liệu đã được công khai trên trang [Data World](#) về một chủ đề mà nhóm sinh viên quan tâm, hứng thú. Dữ liệu phải được cấu trúc hóa thành một bảng gồm ít nhất 5 trường dữ liệu và 1000 dòng.

Giai đoạn	Yêu cầu
A. Thu thập dữ liệu	<ul style="list-style-type: none"> - Nêu ngữ cảnh và động cơ lựa chọn thực hiện việc thu thập dữ liệu. - Nêu chủ đề và nguồn của dữ liệu 100. - Kiểm tra quyền sử dụng và các giấy phép liên quan. - Trình bày cách thức và phương pháp thu thập dữ liệu.
B. Khám phá dữ liệu	<ul style="list-style-type: none"> - Cho biết ý nghĩa của các dòng và các cột. - Cho biết kiểu dữ liệu, giá trị phân bố của các cột. - Trình bày các bước tiền xử lý dữ liệu
C. Khám phá mối quan hệ trong dữ liệu	<ul style="list-style-type: none"> - Chọn và thảo luận về các trường dữ liệu cần thể hiện trực quan. - Lựa chọn biểu đồ phù hợp với tính chất của dữ liệu và giải thích lý do. - Thể hiện mối quan hệ từ đơn giản đến phức tạp bằng cách sử dụng nhiều loại biểu đồ. - Sử dụng màu sắc và kỹ thuật trình bày để làm cho dữ liệu trở nên dễ hiểu hơn. - Thể hiện dữ liệu biến đổi theo thời gian và rút ra nhận định. - Sử dụng các kỹ thuật trình bày dữ liệu để minh họa mối quan hệ.

Bảng 2: Yêu cầu trong quá trình làm việc với dữ liệu

4 Sơ lược pha "Khám phá các mối quan hệ trong dữ liệu"

Các câu hỏi về các trường dữ liệu và quan hệ giữa chúng do nhóm đặt ra:

1. Có sự tương quan (có ý nghĩa về thống kê) giữa giá cổ phiếu và khối lượng giao dịch đối với từng ngân hàng; trong tổng thể hay không?
2. So sánh và nhận xét về giá cổ phiếu và khối lượng giao dịch theo thời gian giữa các ngân hàng. Nên chọn đầu tư cổ phiếu nào?
3. Đánh giá biến động về giá theo thời gian dựa vào giá trung bình của các ngân hàng theo thời gian?
4. Trung bình Change (%) của các ngân hàng theo từng năm nói lên điều gì?
5. Phân tích sự biến động giá cổ phiếu của VPBank, Agribank, VietinBank, Techcombank vào năm 2021 và 2022?

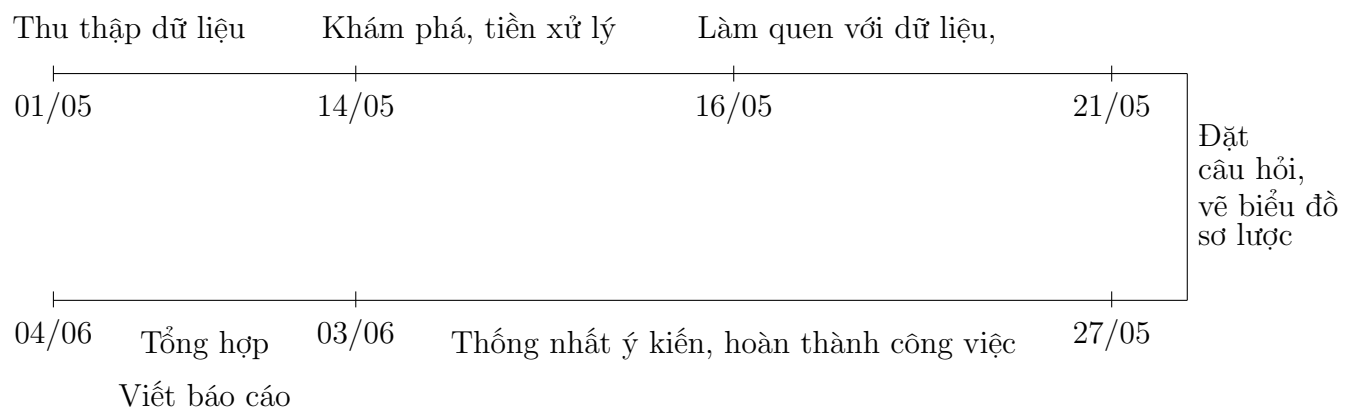
6. Liệu có sự tương quan nào giữa thông tin cổ phiếu với giá trị thương hiệu của top 5 ngân hàng có giá trị thương hiệu cao nhất trong giai đoạn 2021-2023?
7. Moving Average
 - Moving Average là gì? Ứng dụng của Moving Average trong dữ liệu thời gian nói chung, và dữ liệu cổ phiếu nói riêng? Các đường Moving Average phổ biến.
 - Xác định tín hiệu giao dịch (mua/ bán) bằng Moving Average.
8. Kiểm tra tính dừng của các time series về giá cổ phiếu.
 - Sử dụng các phương pháp đánh giá: quan sát biểu đồ trực quan time series, đánh giá thống kê mô tả với chunking, đánh giá xấp xỉ phân phối chuẩn bằng biểu đồ histogram và normal test, kiểm định Augmented Dickey-Fuller Test.
9. Tìm tính chu kỳ trong các time series về giá cổ phiếu của từng ngân hàng.
 - Sử dụng thuật toán Fast Fourier Transform.
10. Thực hiện các phép biến đổi để đưa về time series thỏa mãn tính dừng.
 - Sử dụng Additive Decomposition, Multiplicative Decomposition, Log Transformation, 1st and 2nd Order Differencing, Seasonal Differencing.
 - Kiểm tra tính dừng của time series sau khi biến đổi.
11. Xây dựng mô hình SARIMA và mô hình LSTM để dự đoán giá cổ phiếu.
 - Auto Arima, LSTM.

5 Phân công công việc

Công việc	Đất	Ngọc	Nam	Khuê	Phương
Thu thập dữ liệu			X		
Khám phá dữ liệu				X	
Khám phá mối quan hệ trong dữ liệu	X	X	X	X	X
Câu hỏi	1,2	3,4	5	6,7	8,9,10,11

Bảng 3: Bảng phân công công việc

6 Timeline công việc



7 Mức độ hoàn thành

7.1 Mức độ hoàn thành tổng thể của các yêu cầu

Yêu cầu	Hoàn thành
A. Thu thập dữ liệu: <ul style="list-style-type: none">- Nêu ngữ cảnh và động cơ lựa chọn thực hiện việc thu thập dữ liệu- Nêu chủ đề và nguồn của dữ liệu- Kiểm tra quyền của dữ liệu- Trình bày cách thức và phương pháp thu thập dữ liệu	100%
B. Tiền xử lý và khám phá dữ liệu: <ul style="list-style-type: none">- Cho biết ý nghĩa của các dòng và các cột- Cho biết kiểu dữ liệu, giá trị phân bố của các cột- Trình bày các tiền xử lý dữ liệu	100%
C. Khám phá các mối quan hệ trong dữ liệu: <ul style="list-style-type: none">- Chọn trường dữ liệu và biểu đồ thể hiện quan hệ- Xác định mối quan hệ giữa các trường dữ liệu- Vẽ biểu đồ thể hiện các mối quan hệ- Sử dụng màu sắc để thể hiện dữ liệu- Thể hiện dữ liệu biến đổi theo thời gian	100%

Bảng 4: Yêu cầu trong quá trình làm việc với dữ liệu

7.2 Mức độ hoàn thành của các thành viên

Họ và tên	Mức độ hoàn thành
Lê Công Đất	100 %
Lê Trần Minh Khuê	100%
Hoàng Trung Nam	100%
Lê Trần Như Ngọc	100%
Lê Thị Minh Phương	100%

Bảng 5: Bảng mức độ hoàn thành công việc của các thành viên

8 Tự đánh giá theo các yêu chí của bài tập

Tiêu chí	Hoàn thành
Thu thập và tiền xử lý dữ liệu (5%)	5%
Chọn lựa, giải thích, trực quan các trường và các mối quan hệ giữa chúng (50%)	50%
Rút ra ý nghĩa hợp lý sau mỗi dữ liệu được trực quan (20%)	20%
Xem xét trên nhiều quan hệ, nhiều góc nhìn khác nhau (10%)	10%
Báo cáo trình bày bố cục và định dạng hợp lý, rõ ràng (15%)	15%
Có những phân tích, trực quan hóa bằng những biểu đồ mới lạ và rút ra những thông tin hữu ích. Sử dụng mô hình học máy cơ bản (5%)	5%
Hiểu và giải thích được tổng quan mã nguồn (5%)	5%
Tổng (110%)	110%

Bảng 6: Các yêu cầu và mức độ hoàn thành của bài tập

Tài liệu

- [1] 11 cách để giao dịch với đường MA (Moving Average). [Kakata.vn](#). Truy cập lần cuối: 27/05/2024.
- [2] AgriBank Reports Fourth Quarter 2021 and Year-End Financial Results. [PR Newswire](#). Truy cập lần cuối: 27/05/2024.
- [3] AgriBank Securities JSC (AGR) - Historical Data. [Investing.com](#). Truy cập lần cuối: 27/05/2024.
- [4] AGRISECO INSIGHT REPORT. [Kinh tế Chứng khoán](#). Truy cập lần cuối: 27/05/2024.
- [5] Báo cáo thường niên về các thương hiệu giá trị nhất và mạnh nhất Việt Nam (năm 2022). [BrandDirectory.com](#). Truy cập lần cuối: 27/05/2024.
- [6] Báo cáo thường niên về các thương hiệu giá trị nhất và mạnh nhất Việt Nam (năm 2023). [BrandDirectory.com](#). Truy cập lần cuối: 27/05/2024.
- [7] Chính sách tiền tệ. [Ngân hàng Nhà nước Việt Nam](#). Truy cập lần cuối: 27/05/2024.
- [8] Chứng khoán Agribank (AGR) lãi gấp đôi sau soát xét, đạt 312 tỷ đồng. [Vietstock](#). Truy cập lần cuối: 27/05/2024.
- [9] Chứng khoán Agribank (AGR): Lãi ròng quý 4 tăng gấp 5 lần cùng kỳ, LNST cả năm 2021 xấp xỉ 400 tỷ đồng, vượt 23% kế hoạch. [CafeF.vn](#). Truy cập lần cuối: 27/05/2024.
- [10] Cổ phiếu ngành Ngân hàng: Khả quan để đầu tư? [Ngân hàng Nhà nước Việt Nam](#). Truy cập lần cuối: 27/05/2024.
- [11] Cổ phiếu 'vua' hạ nhiệt, ngân hàng chùn bước lên sàn. [Báo Đầu tư](#). Truy cập lần cuối: 27/05/2024.
- [12] Joint Stock Commercial Bank for Foreign Trade of Vietnam (VCB) - Historical Data. [Investing.com](#). Truy cập lần cuối: 27/05/2024.
- [13] Joint Stock Commercial Bank for Investment and Development of Vietnam (BID) - Historical Data. [Investing.com](#). Truy cập lần cuối: 27/05/2024.
- [14] KB Security - Company report: Vietinbank (CTG). [KB Securities](#). Truy cập lần cuối: 27/05/2024.
- [15] Lợi nhuận quý 4/2022 của VPBank sụt giảm đáng kể. [CafeLand.vn](#). Truy cập lần cuối: 27/05/2024.

- [16] Lợi nhuận Techcombank giảm hơn 20% quý 4/2022 xuống thấp nhất 9 quý. [CafeBiz.vn](#). Truy cập lần cuối: 27/05/2024.
- [17] Microsoft Stock Price Prediction - PARAMARTHA SENGUPTA. [Kaggle](#). Truy cập lần cuối: 27/05/2024.
- [18] Top 500 Ngân hàng có giá trị thương hiệu tốt nhất thế giới (năm 2022). [BrandDirectory.com](#). Truy cập lần cuối: 27/05/2024.
- [19] Top 500 Ngân hàng có giá trị thương hiệu tốt nhất thế giới (năm 2023). [BrandDirectory.com](#). Truy cập lần cuối: 27/05/2024.
- [20] Vietnam Prosperity Joint Stock Commercial Bank (VPB) - Historical Data. [Investing.com](#). Truy cập lần cuối: 27/05/2024.
- [21] Vietnam Technological And Commercial Joint Stock Bank (TCB) - Historical Data. [Investing.com](#). Truy cập lần cuối: 27/05/2024.
- [22] VPB - Sự trở lại của cổ phiếu dẫn đầu. [Thời báo Tài chính Việt Nam](#). Truy cập lần cuối: 27/05/2024.
- [23] Jason Fernando. Moving Average (MA): Purpose, Uses, Formula, and Examples. [Investopedia](#). Truy cập lần cuối: 27/05/2024.
- [24] FilippoMB. Python Time Series Handbook. [FilippoMB.github.io](#). Truy cập lần cuối: 27/05/2024.