

BÁO CÁO NHÓM 10 CASE 1

Đề tài: Phân tích biến động giá vàng TP.HCM – Hà Nội trong 6 tháng

Lớp/Nhóm: DHKL18A1HN– Nhóm 10

Thành viên:

| | |
|------------------|-------------|
| Dương Văn Quỳnh | 24174600058 |
| Hà Giang Nam | 24174600014 |
| Nguyễn Mạnh Tiến | 24174600177 |
| Nguyễn Vũ An | 24174600025 |
| Nguyễn Đức Trung | 24174600049 |

Ngày nộp: 25/11/2025

I. Đặt vấn đề

Công ty tài chính muốn theo dõi biến động giá vàng giữa hai khu vực TP.HCM và Hà Nội trong suốt 6 tháng để nhận diện mức chênh lệch giá theo thời gian. Dữ liệu gồm hai tệp CSV: gold_price_hcm.csv và gold_price_hanoi.csv, mỗi tệp có dạng: YYYY-MM-DD, price.

II. Mục tiêu

- Vận dụng NumPy để xử lý dữ liệu dạng mảng 1 chiều hoặc 2 chiều.
- Thực hành các thao tác slicing, indexing, reshaping, stacking.
- Áp dụng boolean mask và fancy indexing.
- Sử dụng broadcasting để tính toán chênh lệch giữa các mảng.
- Phân tích thống kê cơ bản: mean, std, max, min, argmax, argmin.
- Phân tích biến động theo thời gian bằng np.diff.
- Tạo bảng tổng hợp và xuất file CSV.
- Vẽ biểu đồ minh họa biến động giá.

III. Phương pháp

1. Đọc dữ liệu bằng np.genfromtxt, bỏ header.
2. Tách cột ngày và giá, convert sang datetime64[D].

3. Slicing 7 ngày đầu, giữa, cuối; boolean mask, fancy indexing.
4. Broadcasting tính chênh lệch, np.diff phân tích biến động.
5. Tính thống kê cơ bản: mean, std, max, min, argmax, argmin.
6. Tạo bảng tổng hợp và lưu CSV ket_qua_case_1.csv.
7. Vẽ biểu đồ line minh họa giá TP.HCM, Hà Nội và chênh lệch.

IV. Kết quả

Các kết quả phân tích dữ liệu đã lưu trong ket_qua_case_1.csv và thể hiện trực quan trong biểu đồ line và bar.

V. Nhận xét

Giá vàng TP.HCM thường cao hơn Hà Nội nhưng chênh lệch không lớn. Biến động giá ổn định, top-5 ngày giá cao nhất tập trung vào tháng cuối cùng. Việc sử dụng NumPy giúp xử lý dữ liệu nhanh và chính xác.

VI. Nhật ký đóng góp nhóm

| Thành viên | Mã số | Nhiệm vụ cụ thể |
|------------------|-------------|--|
| Dương Văn Quỳnh | 24174600058 | Tổng hợp & báo cáo + code chính: viết báo cáo nhóm (pdf), tổng hợp dữ liệu từ các thành viên, gộp code NumPy, đọc file CSV, tách cột ngày và giá, convert sang datetime64[D], tạo CSV tổng hợp, vẽ biểu đồ line/bar minh họa. |
| Hà Giang Nam | 24174600014 | Xử lý dữ liệu cơ bản: slicing 7 ngày đầu/giữa/cuối, kiểm tra dtype & shape. |
| Nguyễn Mạnh Tiến | 24174600177 | Top-5 ngày giá cao nhất: tìm top-5 ngày giá HCM & HN, chuẩn bị bảng nhận xét cho báo cáo. |
| Nguyễn Vũ An | 24174600025 | Biến động ngày liền kề: tính np.diff cho HCM & HN, phân tích biến động tăng/giảm, nhận xét biến động dữ liệu. |
| Nguyễn Đức Trung | 24174600049 | Thống kê cơ bản & CSV chuẩn bị: tính mean, std, max, min, argmax, argmin; hỗ trợ chuẩn bị bảng dữ liệu lưu CSV. |

VII. Kết luận

Mục tiêu đề ra đã hoàn thành: xử lý NumPy, phân tích thống kê, tạo CSV, vẽ biểu đồ. Kết quả có thể sử dụng cho báo cáo nội bộ công ty hoặc phân tích các chỉ số tài chính.