**BÁO CÁO THỰC HÀNH TUẦN 05**

Dataset: **Starter: Bitcoin time series Data Till df26fa3c-8**

1. **plotPerColumnDistribution(df1, 10, 5)**

* Không có cột nào phù hợp để vẽ biểu đồ phân phối.

1. **plotCorrelationMatrix(df1, 8)**

* Ma trận tương quan trên biểu diễn mối quan hệ tuyến tính giữa các biến số trong tập dữ liệu Bitcoin. Hệ số tương quan nằm trong khoảng từ -1 đến 1:
* **1 hoặc -1**: Hai biến có tương quan hoàn hảo (cùng hoặc ngược chiều).
* **0**: Không có mối quan hệ tuyến tính giữa hai biến.
* Nhận xét chính từ ma trận tương quan

*Mối quan hệ rất mạnh giữa các chỉ số giá (Price, Open, High, Low)*

* Price, Open, High, và Low có hệ số tương quan gần bằng **1**.
* Điều này hợp lý vì trong một ngày, giá mở cửa, giá cao nhất, giá thấp nhất, và giá đóng cửa của Bitcoin thường thay đổi theo xu hướng chung.

*Date có rất ít tương quan với các chỉ số giá*

* Cột Date có hệ số tương quan rất thấp với Price, Open, High, và Low (khoảng -0.04 đến -0.06).
* Điều này có nghĩa là **ngày giao dịch không ảnh hưởng mạnh đến giá trị Bitcoin** nếu xét theo mối quan hệ tuyến tính.

*Change % không có mối quan hệ mạnh với giá trị giá Bitcoin*

* Change % có hệ số tương quan rất nhỏ với Price, Open, High, Low (dưới 0.1).
* Điều này có nghĩa là **tỷ lệ thay đổi giá không bị ảnh hưởng quá mạnh bởi giá trị thực của Bitcoin**.
* Đây là một điểm quan trọng vì nó cho thấy **biến động giá Bitcoin không bị ảnh hưởng đơn thuần bởi mức giá hiện tại** mà có thể bị tác động bởi các yếu tố khác như tin tức thị trường hoặc sự kiện   
  kinh tế.
* Kết luận
* Giá Bitcoin (Price, Open, High, Low) có mối quan hệ **rất chặt chẽ** với nhau.
* **Ngày giao dịch (Date) không có ảnh hưởng tuyến tính mạnh** đến giá Bitcoin.
* **Tỷ lệ thay đổi giá (Change %) không có mối quan hệ mạnh với giá trị thực của Bitcoin**, điều này có thể gợi ý rằng các yếu tố phi tuyến tính như tâm lý thị trường, tin tức, hoặc khối lượng giao dịch có thể đóng vai trò quan trọng hơn.
* Gợi ý tiếp theo:
* **Phân tích phi tuyến tính** như hồi quy phi tuyến hoặc mô hình machine learning để hiểu rõ hơn về sự biến động của giá Bitcoin.
* **Kiểm tra tương quan với các yếu tố khác** như khối lượng giao dịch (Vol.) hoặc dữ liệu bên ngoài như biến động thị trường chứng khoán hoặc tin tức kinh tế.

1. **plotScatterMatrix(df1, 12, 10)**

* Mối quan hệ tuyến tính mạnh giữa các chỉ số giá (Price, Open, High, Low)
* Các biến Price, Open, High, Low có sự phân tán **gần như theo đường chéo 45 độ**.
* Hệ số tương quan gần bằng **1** → Chúng có quan hệ tuyến tính rất mạnh, điều này hợp lý vì giá Bitcoin trong ngày thường không thay đổi đột ngột giữa các mức giá này.

*Cột Date không có mối quan hệ tuyến tính với giá Bitcoin*

* Các điểm trong biểu đồ giữa Date và các chỉ số giá (Price, Open, High, Low) **phân bố rời rạc, không có xu hướng rõ ràng**.
* Hệ số tương quan rất nhỏ (~0) → **Ngày giao dịch không ảnh hưởng trực tiếp đến giá trị Bitcoin theo mô hình tuyến tính**.
* Tuy nhiên, điều này không có nghĩa là thời gian không có tác động, có thể cần kiểm tra theo chu kỳ hoặc bằng mô hình phi tuyến tính.

*Tỷ lệ thay đổi giá (Change %) không có tương quan đáng kể với giá Bitcoin*

* Các biểu đồ scatter giữa Change % và các chỉ số giá **phân tán rất nhiều** mà không có hình dạng rõ ràng.
* Hệ số tương quan nhỏ (dưới 0.1) → Điều này chứng tỏ **biến động giá không bị ảnh hưởng mạnh bởi giá trị thực của Bitcoin**.
* Change % có thể bị tác động bởi các yếu tố phi tuyến tính khác như tin tức thị trường, khối lượng giao dịch, hoặc tâm lý nhà đầu tư.

*Phân phối dữ liệu*

* Các biểu đồ mật độ trên đường chéo cho thấy:
  + Date có phân phối khá đồng đều, phản ánh khoảng thời gian giao dịch rộng.
  + Price, Open, High, Low có **phân phối lệch phải**, cho thấy sự tăng trưởng mạnh của Bitcoin qua thời gian.
  + Change % có phân phối **gần như chuẩn**, với phần lớn giá trị dao động quanh 0%, phản ánh sự biến động hàng ngày của Bitcoin.
* Kết luận
* Giá Bitcoin có quan hệ tuyến tính rất chặt chẽ giữa các chỉ số (Price, Open, High, Low).
* Ngày giao dịch (Date) không có tác động tuyến tính mạnh đến giá Bitcoin.
* Tỷ lệ thay đổi giá (Change %) không có tương quan đáng kể với giá trị Bitcoin, gợi ý rằng biến động giá có thể chịu tác động bởi các yếu tố khác như thị trường hoặc tin tức.
* Gợi ý tiếp theo:
* **Phân tích phi tuyến tính** để kiểm tra tác động của thời gian (Date) đến giá Bitcoin.
* **Kiểm tra mối quan hệ giữa khối lượng giao dịch (Vol.) và biến động giá (Change %)**, vì có thể có ảnh hưởng từ tâm lý nhà đầu tư.