**CHƯƠNG 1: GIỚI THIỆU**

**1.1 ĐẶT VẤN ĐỀ**

Tại hầu hết các quốc gia trên thế giới, thị trường chứng khoán là một phần  
không thể thiếu trong nền kinh tế. Nó chính là “lá phổi” của nền kinh tế. Bên  
cạnh các kênh thu hút vốn đầu tư từ vay vốn từ ngân hàng, vay vốn từ nhà nước  
theo các chương trình hỗ trợ… thì thị trường chứng khoá cũng là một kênh thu  
hút vốn đầu tư lớn cho các doanh nghiệp. Bởi thế giá cổ phiếu, chỉ số sàn chứng  
khoán là một trong những thông tin quan trọng mà những nhà đầu tư, nhà quản lí  
hết sức quan tâm. Chính vì thế việc dự đoán, dự báo được giá, các chỉ số chứng  
khoán là hết sức quan trọng. Nó giúp cho những nhà quản lí, đầu tư có các quyết  
định đúng đắn. Qua đó, không chỉ góp phần vào phát triển thị trường chứng  
khoán mà còn góp phần phát triển nền kinh tế của quốc gia.

Ra đời vào đầu năm 2000, thị trường chứng khoán Việt Nam đã trở thành một kênh đầu tư hết sức hấp dẫn đối với các nhà đầu tư, từ các tổ chức đầu tư chuyên  
nghiệp cho đến các nhà đầu tư cá nhân nghiệp dư nhỏ lẻ. Tuy nhiên, bên cạnh  
mức sinh lợi cao, đây cũng là hoạt động luôn tồn tại nhiều rủi ro tiềm ẩn bởi nhà  
đầu tư không phải lúc nào cũng dự đoán được chính xác xu hướng của giá cổ  
phiếu trong tương lai. Do đó, việc dự đoán dự báo chính xác sự biến động giá  
của cổ phiếu để có một sách lược nhằm phục vụ cho công việc kinh doanh của các cá nhân, tổ chức hay hoạch định chiến lược của một quốc gia đã thu hút rất  
nhiều sự quan tâm của các nhà kinh tế lượng tài chính trong và ngoài nước.

Tại thị trường chứng khoán Việt Nam, giá của một loại cổ phiếu thuộc một lĩnh vực kinh doanh bất kì có sự thay đổi có thể kéo theo sự thay đổi của hệ thống  
các cổ phiếu khác không chỉ trong cùng một lĩnh vực mà còn ở các lĩnh vực  
khác. Vì vậy việc dự đoán dự báo sự tăng giảm của giá cổ phiếu không chỉ phản  
ánh sự phát triển của công ty mà còn giúp các nhà đầu tư, nhà quản lí nhận biết  
chiều hướng thay đổi để qua đó ra quyết định chính xác hơn.

Trong khuôn khổ đề tài, tôi đề xuất sử dụng mô hình ARIMA và phương  
pháp Box - Jenkins để dự báo giá cổ phiếu của ngân hàng thương mại cổ phần Ngoại Thương Việt Nam (VCB) trong ngắn hạn căn cứ vào chuỗi dữ liệu quá khứ. George Box và Gwilym Jenkins (1976) đã nghiên cứu mô hình ARIMA và tên  
của họ thường được dùng để gọi tên các quá trình ARIMA tổng quát, áp dụng  
vào việc phân tích và dự báo các chuỗi thời gian. Phương pháp Box-Jenkins với  
bốn bước: nhận dạng mô hình thử nghiệm; ước lượng; kiểm định bằng chẩn  
đoán; và dự báo.

**1.2 GIỚI THIỆU VỀ VCB**

Ngân hàng Ngoại thương Việt Nam trước đây, nay là Ngân hàng TMCP Ngoại thương Việt Nam (Vietcombank) được thành lập và chính thức đi vào hoạt động ngày 01/4/1963 với tổ chức tiền thân là Cục Ngoại hối (trực thuộc Ngân hàng Nhà nước Việt Nam).

Là ngân hàng thương mại nhà nước đầu tiên được Chính phủ lựa chọn thực hiện thí điểm cổ phần hoá, Vietcombank chính thức hoạt động với tư cách là một ngân hàng thương mại cổ phần vào ngày 02/06/2008 sau khi thực hiện thành công kế hoạch cổ phần hóa thông qua việc phát hành cổ phiếu lần đầu ra công chúng. Ngày 30/6/2009, cổ phiếu Vietcombank (mã chứng khoán VCB) chính thức được niêm yết tại Sở Giao dịch chứng khoán TP.HCM.

Trải qua gần 60 năm xây dựng và trưởng thành, Vietcombank đã có những đóng góp quan trọng cho sự ổn định và phát triển của kinh tế đất nước, phát huy tốt vai trò của một ngân hàng đối ngoại chủ lực, phục vụ hiệu quả cho phát triển kinh tế trong nước, đồng thời tạo những ảnh hưởng quan trọng đối với cộng đồng tài chính khu vực và toàn cầu.

Từ một ngân hàng chuyên doanh phục vụ kinh tế đối ngoại, Vietcombank ngày nay đã trở thành một ngân hàng đa năng, hoạt động đa lĩnh vực, cung cấp cho khách hàng đầy đủ các dịch vụ tài chính hàng đầu trong lĩnh vực thương mại quốc tế; trong các hoạt động truyền thống như kinh doanh vốn, huy động vốn, tín dụng, tài trợ dự án…cũng như mảng dịch vụ ngân hàng hiện đại: kinh doanh ngoại tệ và các công vụ phái sinh, dịch vụ thẻ, ngân hàng điện tử…

Sở hữu hạ tầng kỹ thuật ngân hàng hiện đại với việc chuyển đổi thành công hệ thống ngân hàng lõi (Core Banking) vào đầu năm 2020, Vietcombank có nhiều lợi thế trong ứng dụng công nghệ tiên tiến vào xử lý tự động các dịch vụ ngân hàng, phát triển các sản phẩm, dịch vụ ngân hàng điện tử dựa trên nền tảng công nghệ cao. Không gian giao dịch công nghệ số cùng các dịch vụ ngân hàng số đa tiện ích cung cấp cho khách hàng cá nhân, doanh nghiệp như: VCB Digibank, VCB Pay, VCB - iB@nking, VCB CashUp, VCB DigiBiz... đã, đang và sẽ tiếp tục thu hút đông đảo khách hàng bằng sự tiện lợi, nhanh chóng, an toàn, hiệu quả, tạo thói quen thanh toán không dùng tiền mặt cho đông đảo khách hàng.

**1.3 MÔ HÌNH ARIMA**

Mô hình sử dụng dữ liệu chuỗi thời gian, xem giá trị trong quá khứ của một biến  
số cụ thể là một chỉ tiêu tốt phản ánh giá trị trong tương lai của nó, cụ thể, cho  
Yt là giá trị của biến số tại thời điểm t với Yt = f(Yt-1, Yt-2, ..., Y0, t).  
Mục đích của phân tích là để thấy rõ một số mối quan hệ giữa các giá trị Yt được  
quan sát đến nay để cho phép chúng ta dự báo giá trị Yt trong tương lai. Phương  
pháp này đặc biệt hữu ích cho việc dự báo trong ngắn hạn.

Mô hình tự hồi quy p - AR(p): trong mô hình tự hồi qui quá trình phụ thuộc vào  
tổng trọng số của các giá trị quá khứ và số hạng nhiễu ngẫu nhiên:

Yt = φ1Yt-1 + φ2Yt-2 + ...+φpYt-p +δ +εt

Mô hình trung bình trượt q – MA(q): trong mô hình trung bình trượt, quá trình  
được mô tả hoàn toàn bằng tổng trọng số của các ngẫu nhiên hiện hành có độ trễ:

Yt = μ +εt −θ1εt-1 −θ2εt-2 −...−θqεt-q

Mô Hình Hồi Quy Kết Hợp Trung Bình Trượt - ARMA(p,q):

Yt = φ1Yt-1 + φ2Yt-2 + ...+φpYt-p +δ +εt − θ1εt-1 −θ2εt-2 −...−θqεt-q

Tuy nhiên trong bài nghiên cứu này , thì mô hình chạy ở đây tôi xin dùng hàm có sẵn để có thể dự đoán xu hướng giá trong bài này

**CHƯƠNG 2: DỮ LIỆU VÀ XỬ LÝ DỮ LIỆU**

**2.1 DỮ LIỆU:**

- Nguồn cập nhật số liệu là từ một thư viện có sẵn trong Python mang tên FinanceDataReader. Đây là một thư viện cho phép người xài có thể trích xuất giá Open, Close, High, Low, Change và Volume của cổ phiếu bất kì

- Số liệu được lấy từ ngày 1/7/2009 tới ngày 20/6/2022.

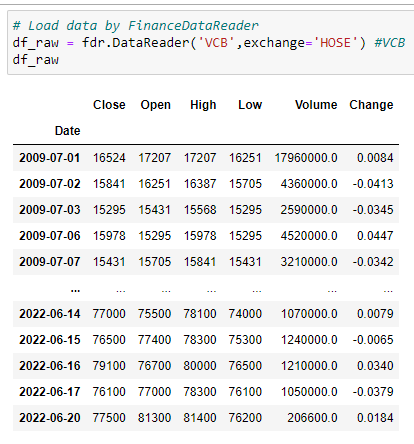
- Tôi chọn chuỗi số liệu và công ty cho bài dự đoán lần này là vì:

+ Công ty ngân hàng thương mại cổ phần Ngoại Thương Việt Nam (VCB) là một trong các công ty có giá cổ phiếu thuộc loại cao. Được nhiều nhà đầu tư quan tâm.

+ Trong thời gian gần đây lĩnh vực ngân hàng đang trở thành một trong những vấn đề chính không chỉ các nhà đầu tư để ý mà còn được nhà nước hết sức quan tâm. Không chỉ bởi nó là một lĩnh vực nhiều tiềm năng mà còn ở chỗ đó là, vào thời điểm hiện tại, tín dụng vi mô số cũng như công nghệ tài chính là một bộ phận chủ yếu trong phát triển kinh tế quốc gia.

+ Bởi thế dự đoán, dự báo được giá cổ phiếu VCB vào thời điểm hiện tại không chỉ giúp các nhà đầu tư có quyết định chính xác mà còn cho ta thấy cái sự phát triển, rủi ro trong lĩnh vực ngân hàng tài chính này.

- Sau đây là code và kết quả:



**2.2 XỬ LÝ DỮ LIỆU:**

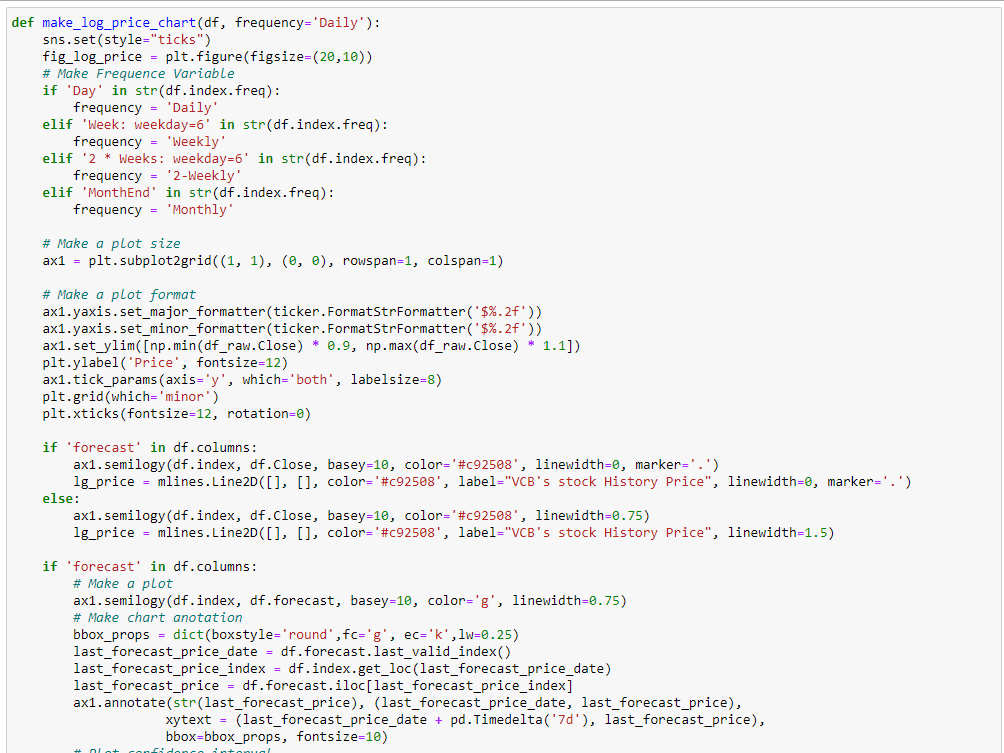
- Trong bài nghiên cứu này, tôi sẽ dự đoán giá hay xu hướng giá của cổ phiếu VCB theo ngày, tuần, 2 tuần và tháng qua đó tại đây tôi sẽ xử lý data này chia nó thành 4 dataframe theo ngày, tuần, 2 tuần và tháng. Đây là code của tôi.

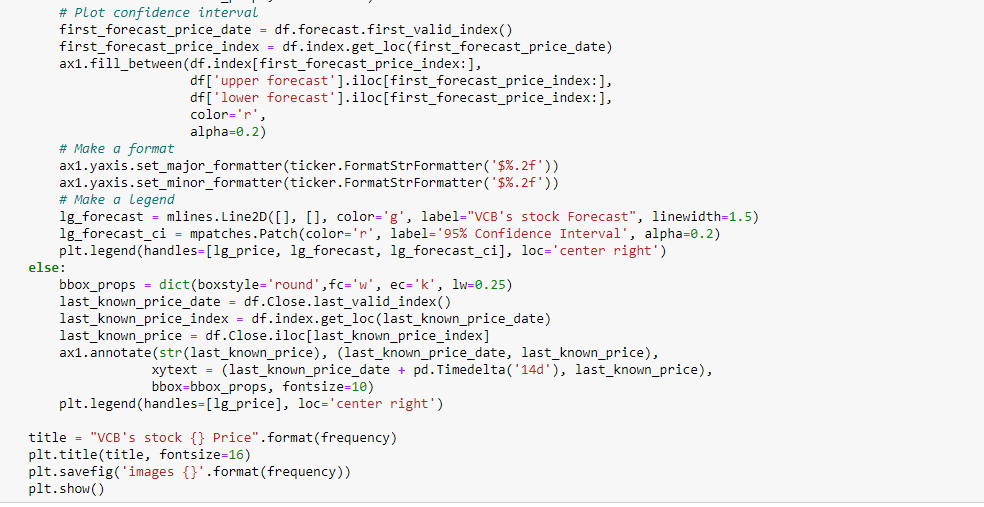


Tại, đây tôi dùng hàm Resample() và Mean(), Copy() để có thể lấy mẫu theo ý muốn ban đầu của tôi đề ra. Ngoài ra, tôi cũng chỉ cần 2 cột là Close và Volume để có thể dự đoán trong bài nghiên cứu này. bên cạnh đó việc tôi def thêm một hàm tên New\_Sort() hàm này có tác dụng là dựa trên cột dữ liệu được Resample theo ngày thì những ngày có dữ liệu NA ( Có thể là dữ liệu của những ngày thứ 7, Chủ Nhật hoặc ngày nghỉ của thị trường chứng khoán Việt Nam qua các năm thì sẽ không có dữ liệu) thì cột giá đóng cửa sẽ bằng ngày trước và volume = 0.

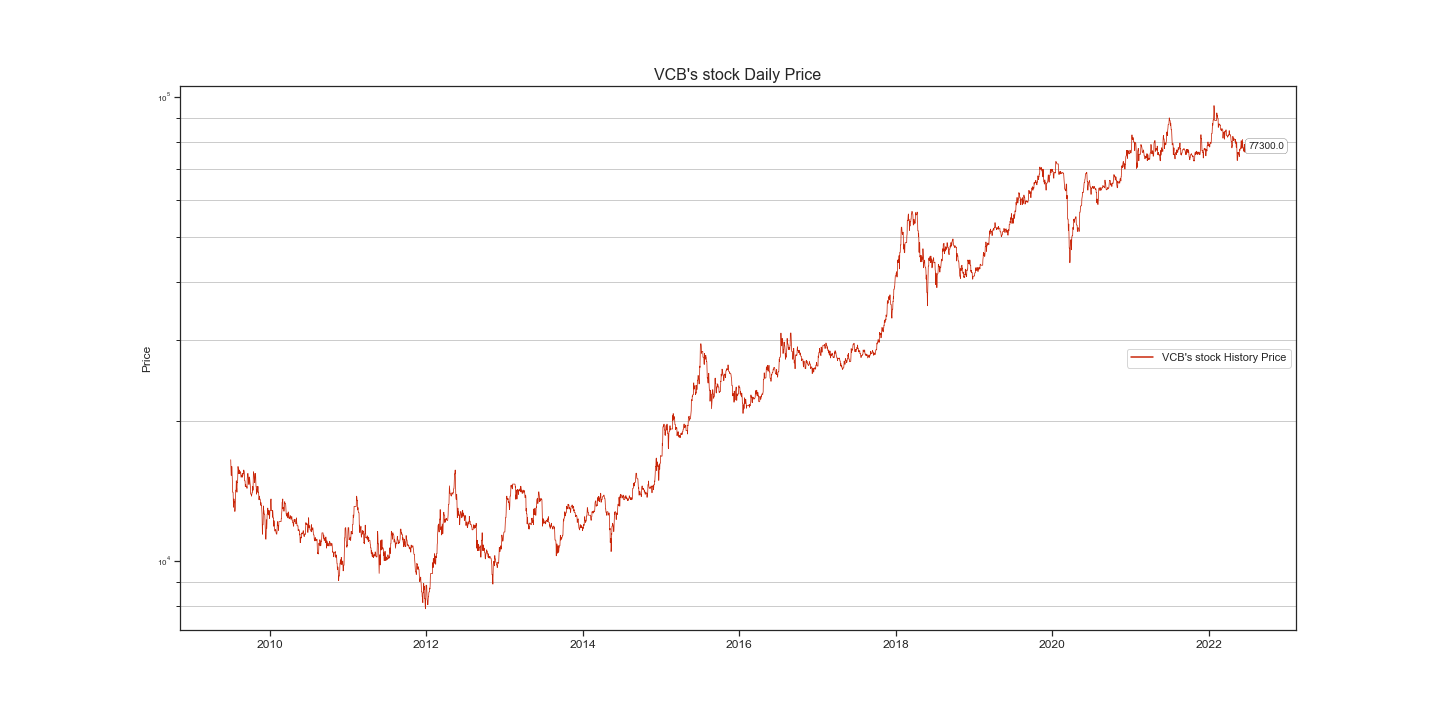
**CHƯƠNG 3: LỊCH SỬ GIÁ CỦA VCB**

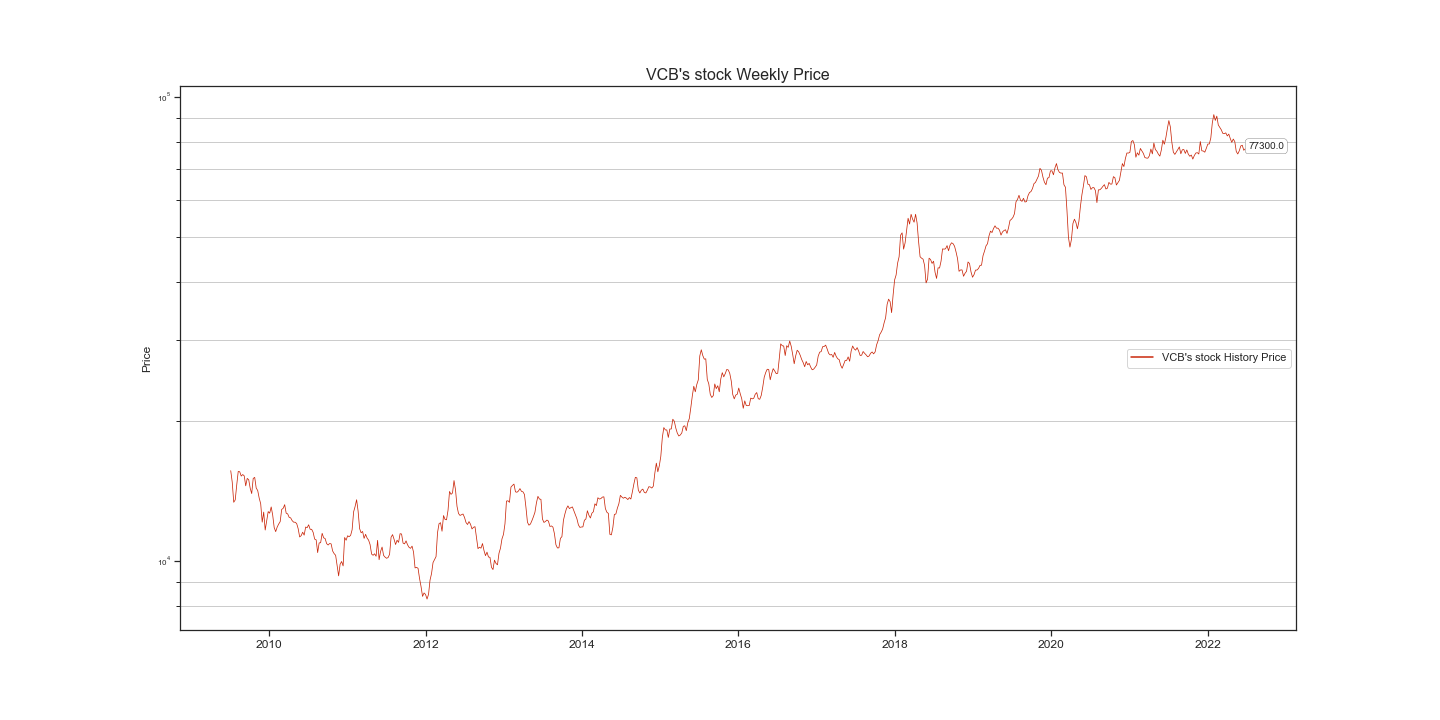
Để có cái nhìn tổng quan nhất về 4 dataframe vừa tạo ở mục 2.2 thì tại đây, tôi đã def ra một function vừa dành cho mục dự đoán mà vừa dành cho mục vẽ đồ thị. Đây là code

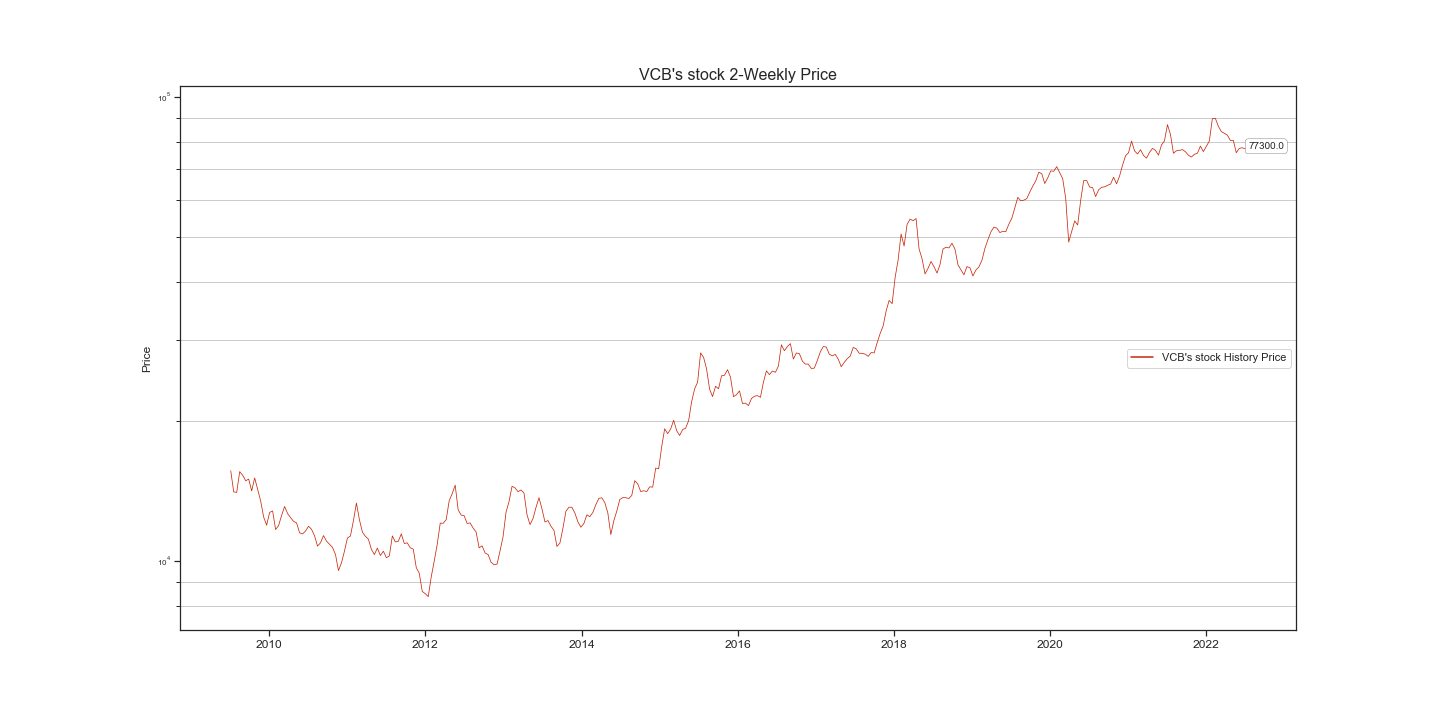


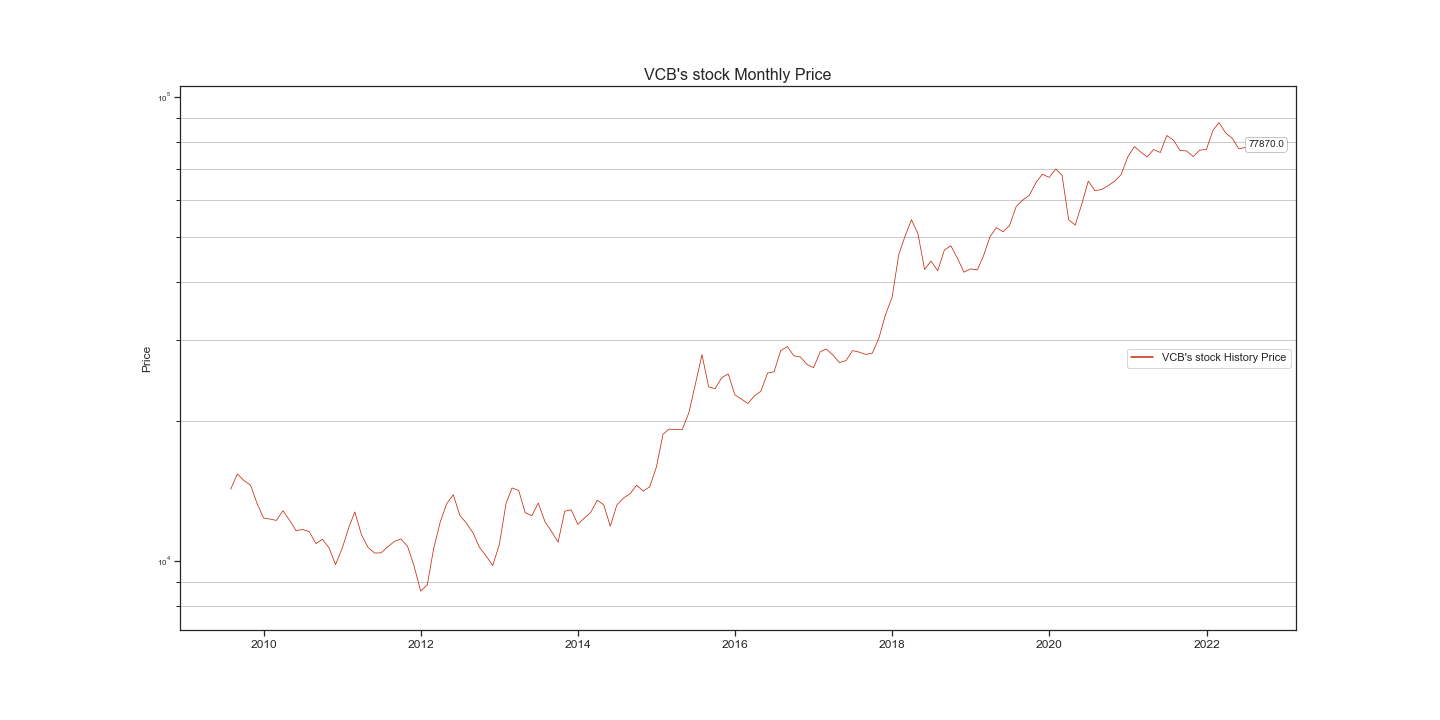


Đây là kết quả các biểu đồ mà tôi thu được từ 4 cột dataframe vừa tạo ra





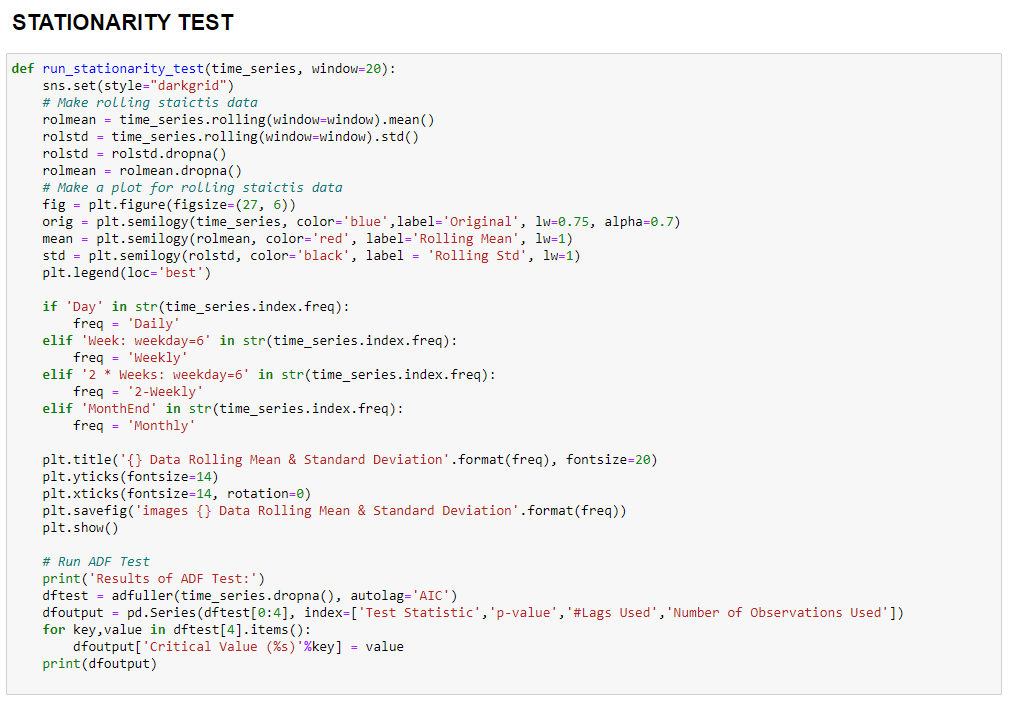




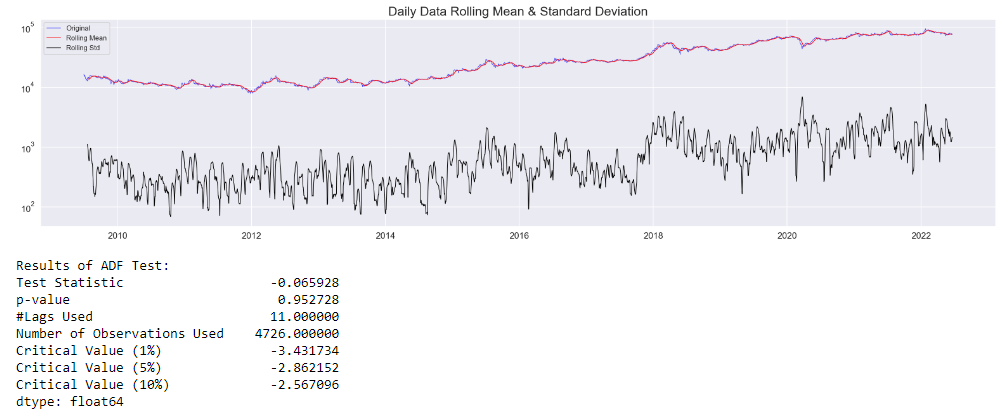
Nhìn chung, trong xuyên suốt lịch sử của VCB ta thấy rõ được sự biến động mạnh mẽ của ngân hàng cũng như phát triển vượt bậc , khi qua các năm thì giờ giá trị của cổ phiếu VCB đã tăng mạnh hơn nhiều so với lúc trước điều này là một tin rất đáng mừng vì điều này chứng mình rằng công ty này đang ngày càng phát triển mạnh mẽ dù có biến động như nào thì công ty vẫn phát triển tốt

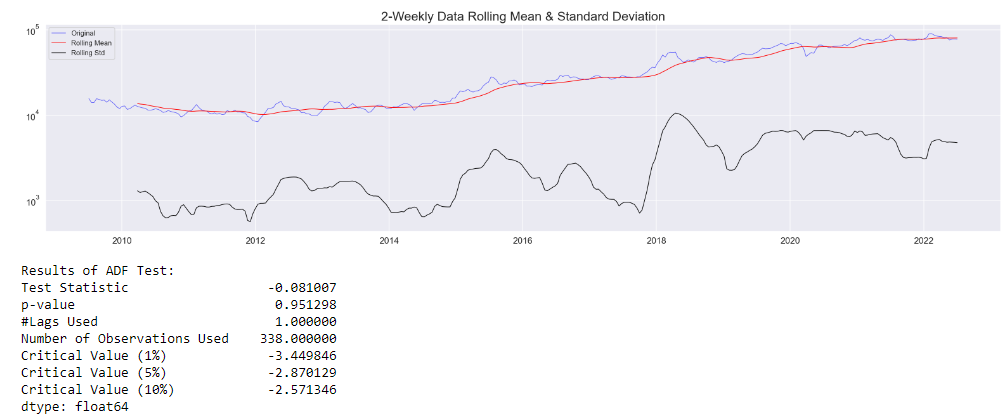
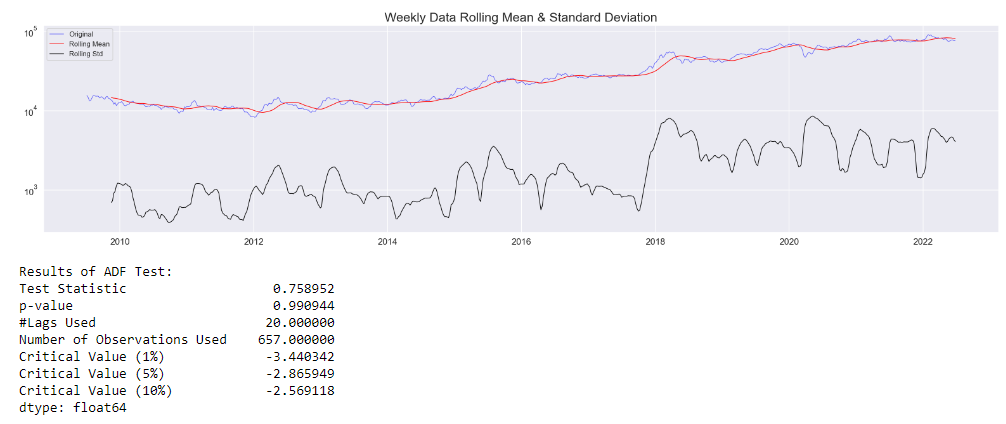
CHƯƠNG 4: KIỂM ĐỊNH TÍNH DỪNG, BOX COX TRANSFORMATION VÀ SỰ TƯƠNG QUAN

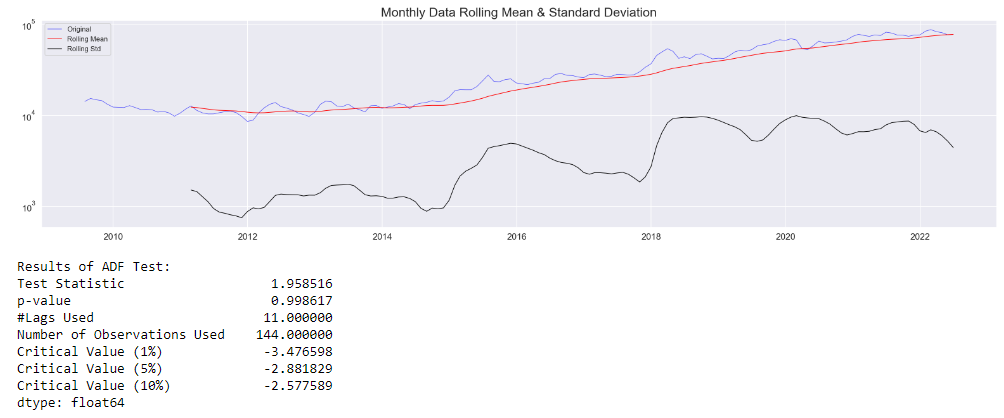
Để có thể dự đoán được xu hướng giá thì đầu tiên ta phải kiểm định xem các bộ dữ liệu này có mang tính dừng hay không, và đây là code mà tôi xây dựng nên để kiểm tra



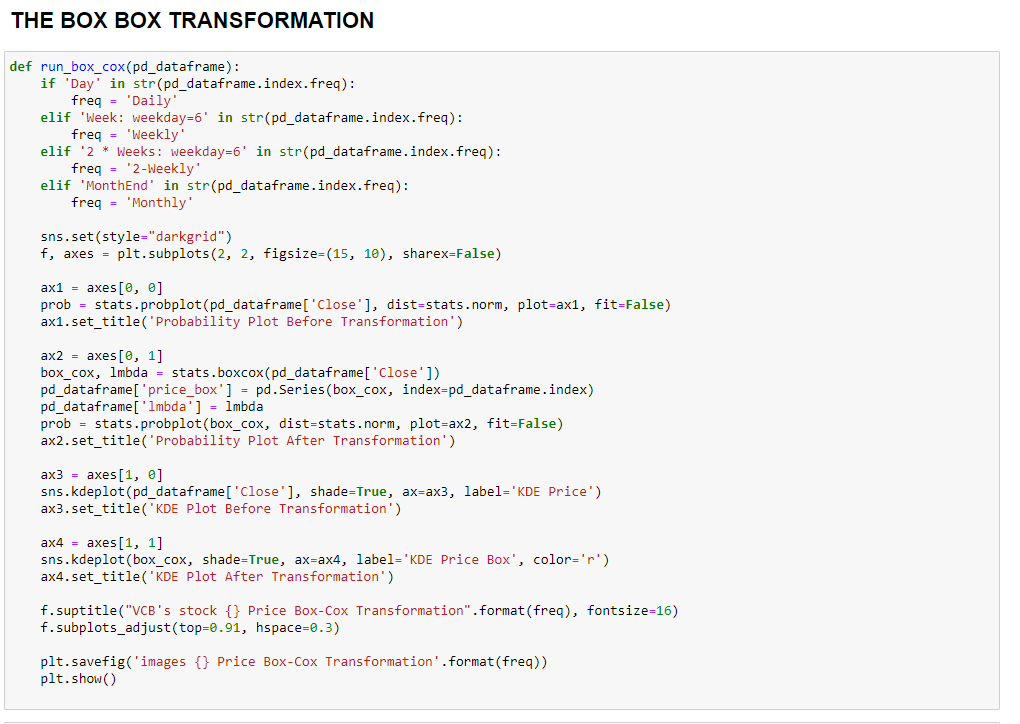
Đây là kết quả



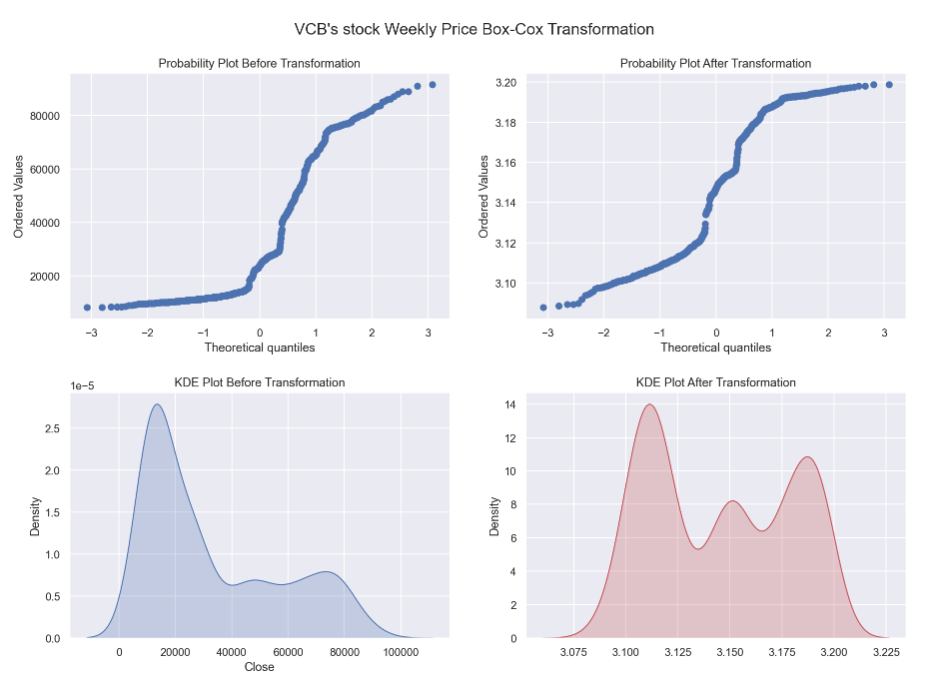
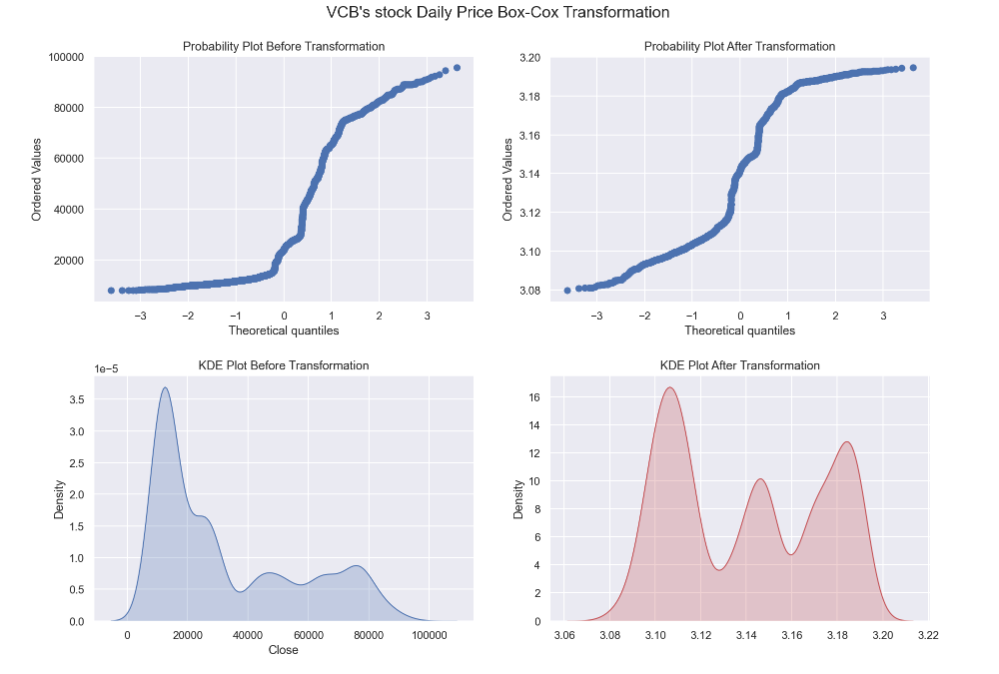


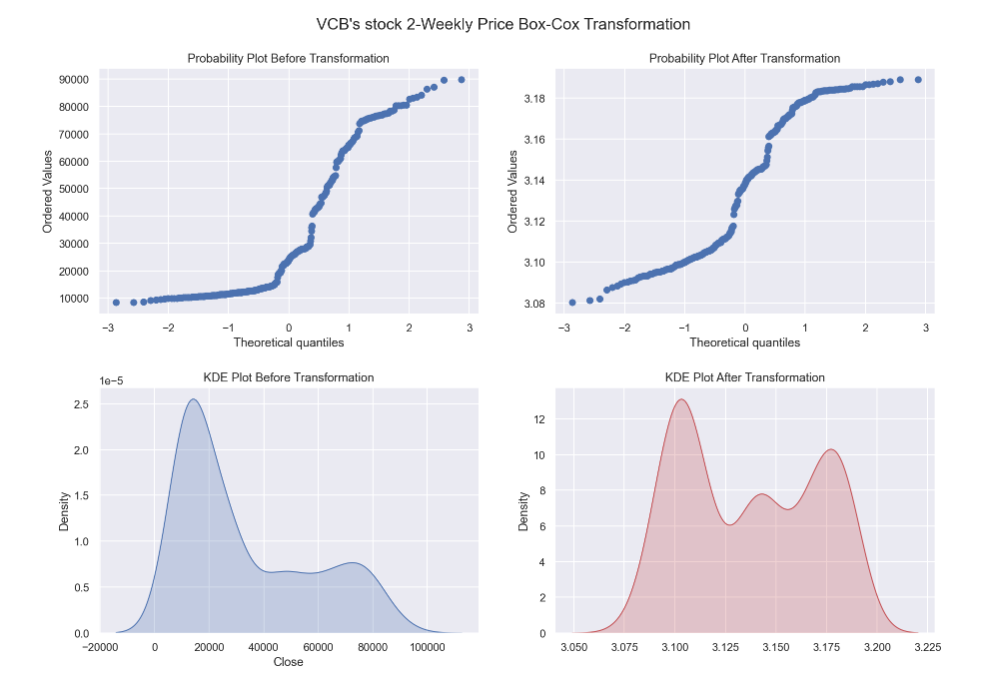


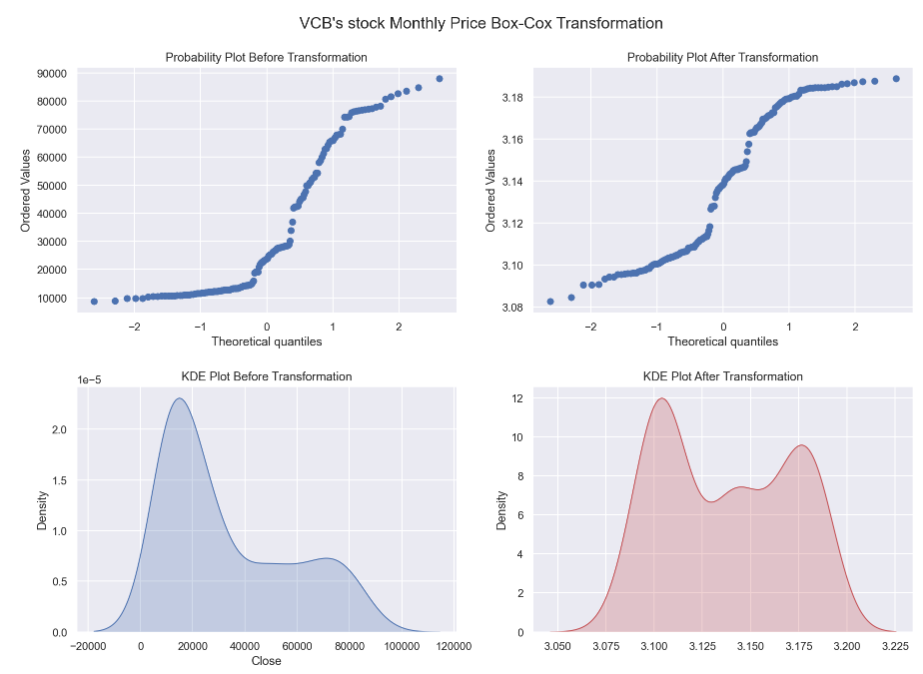
Nhìn 4 kết quả trên ta thấy , các p đều > 0.05 qua đó gián tiếp cho thấy chúng ta chưa thể kết luận rằng bộ dữ liệu này không mang tính dừng mà có thể có hoặc không nhưng vì p xấp xỉ = 1 nên có thể nói bộ dữ liệu rất có thể mang tính dừng. Do đó tôi dùng một phương pháp mang tên box\_cox transformation để chuyển hoá dữ liệu . Đây là code



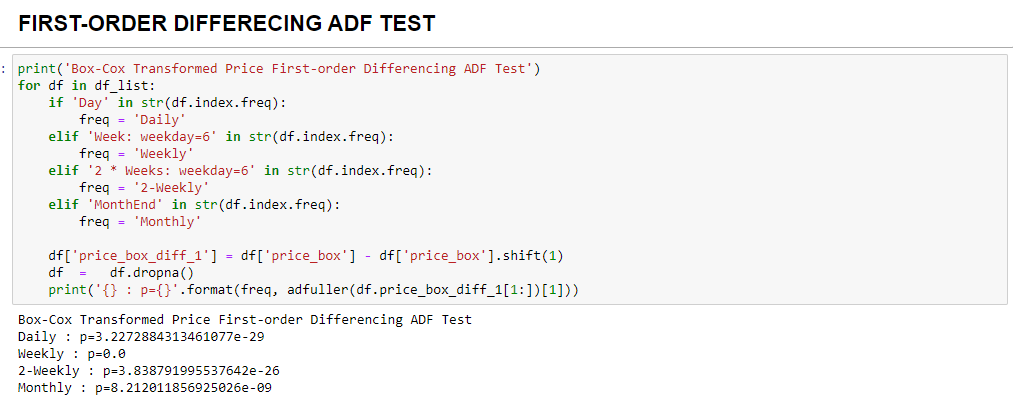
Đây là kết quả



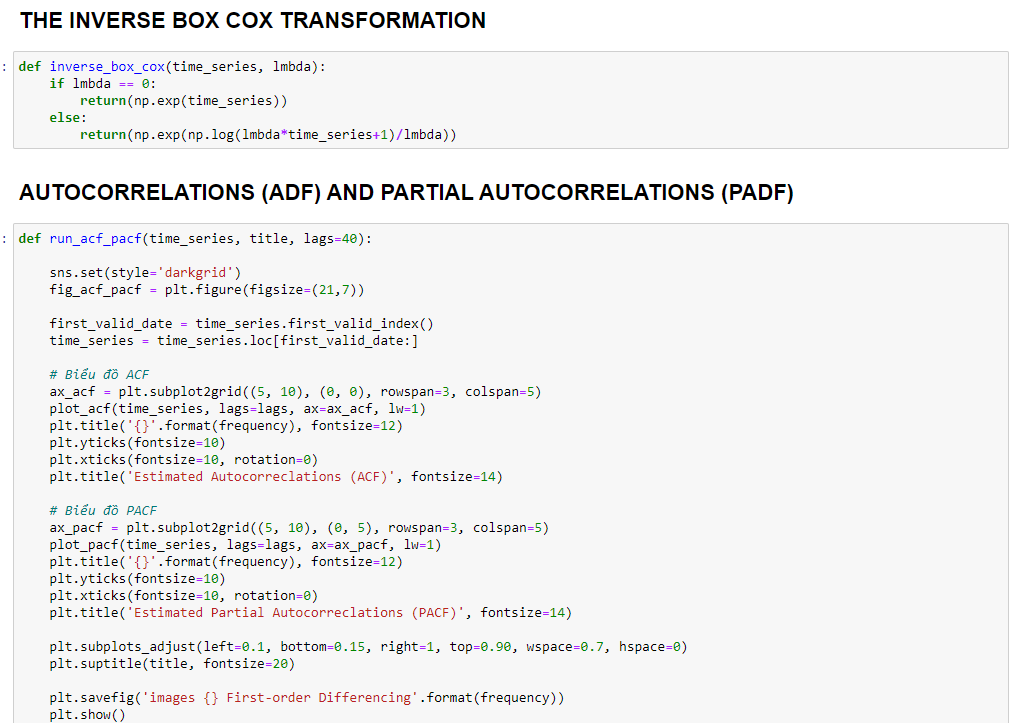




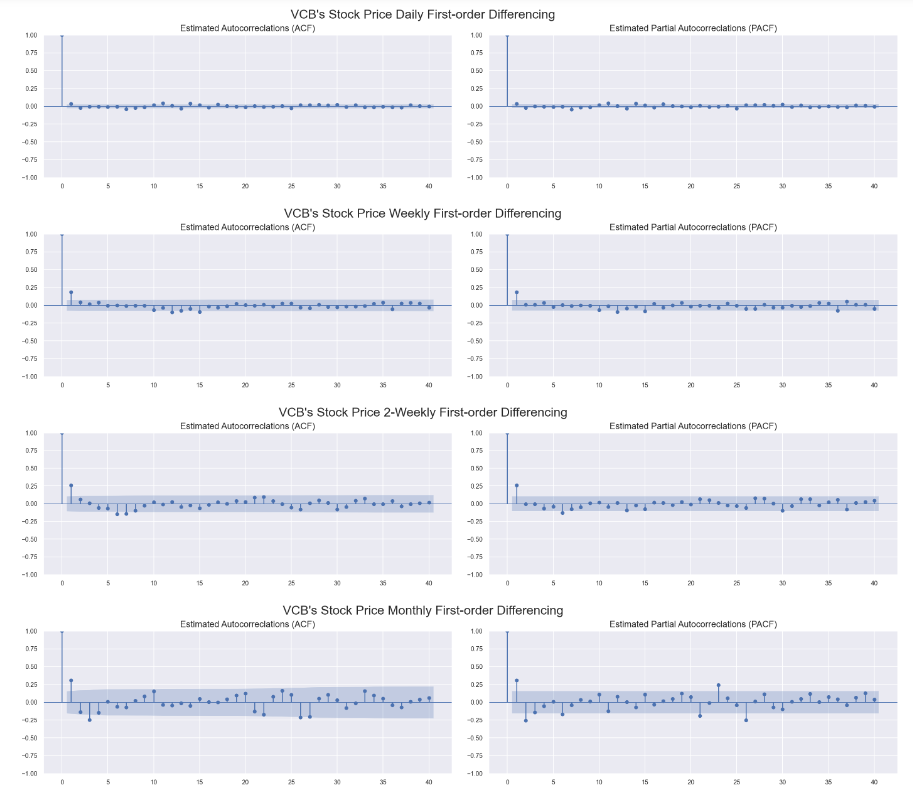
Nhìn chung kết quả sau khi box cox nhìn khá khả quan nên ngay từ lần box cox đầu tiên tôi sẽ cho kiểm định lại.



Sau khi phân sai bằng box cox bậc 1 tôi để cho thấy được dữ liệu này không còn tính dừng nữa. Tiếp theo tui sẽ kiểm định sự tự tương quan của các biến này bằng các code sau



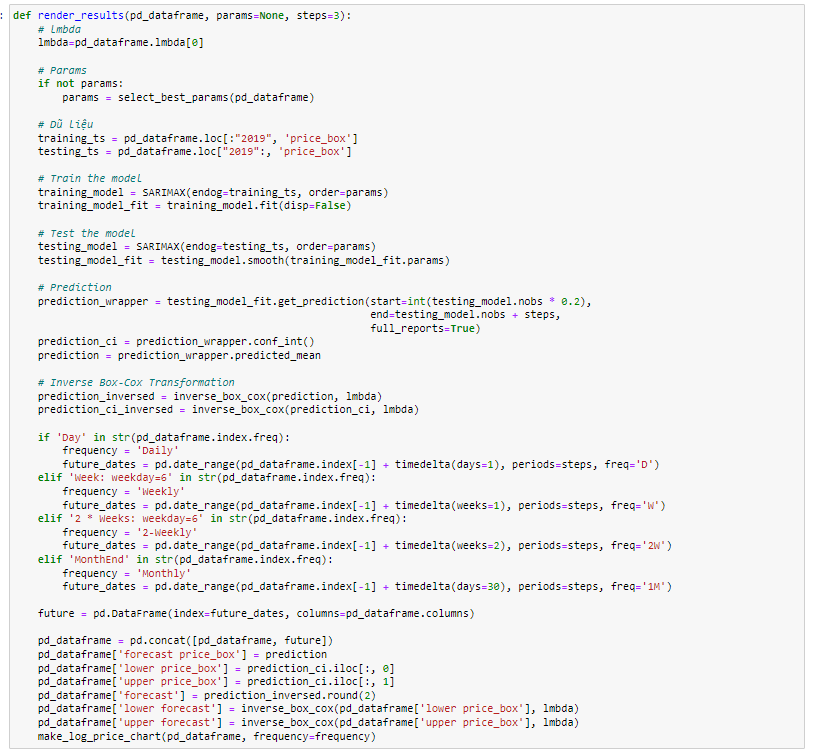
Đây là kết quả tôi thu được:



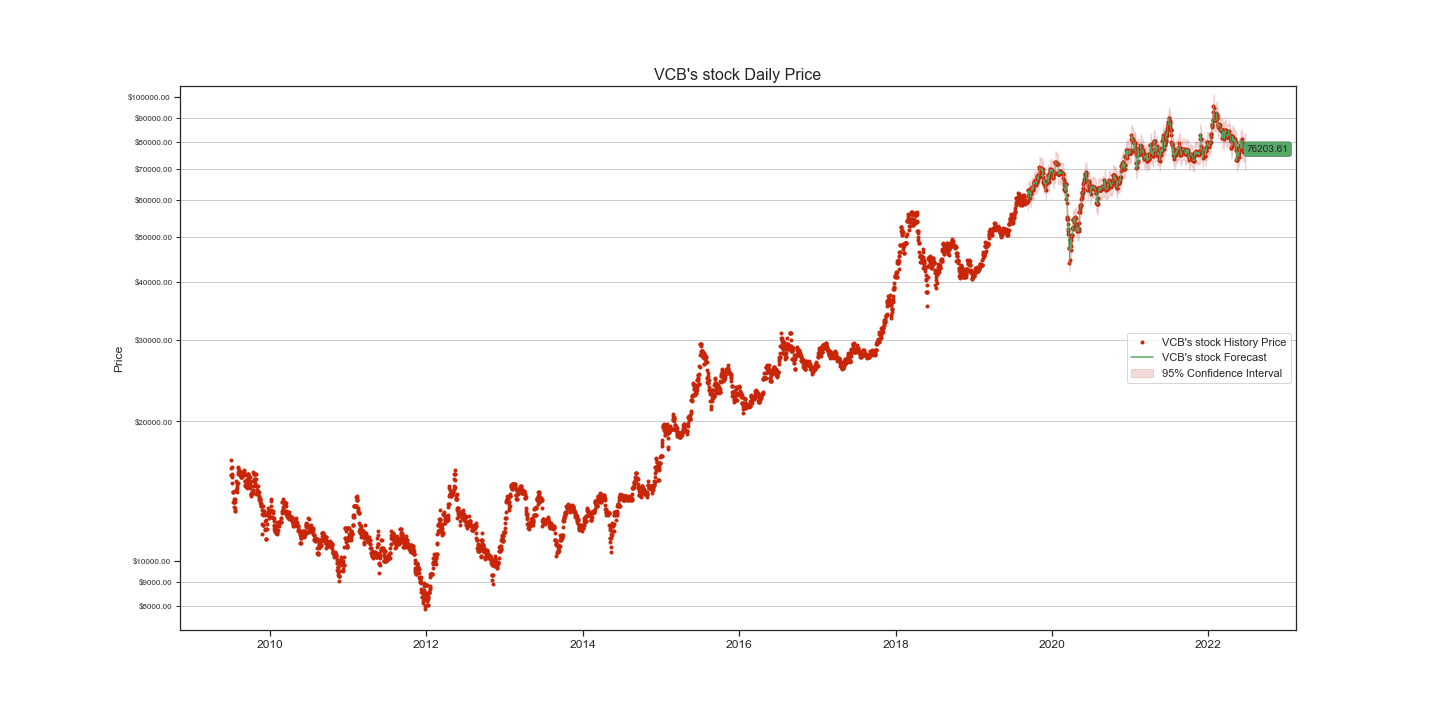
Nhìn chung kết quả tự tương quan thu được cũng khá tốt, vì thế tôi đã tiếp tục tiến đến bước cuối cùng chính là lựa chọn params và cho kết quả dự đoán dựa trên params đã chọn.

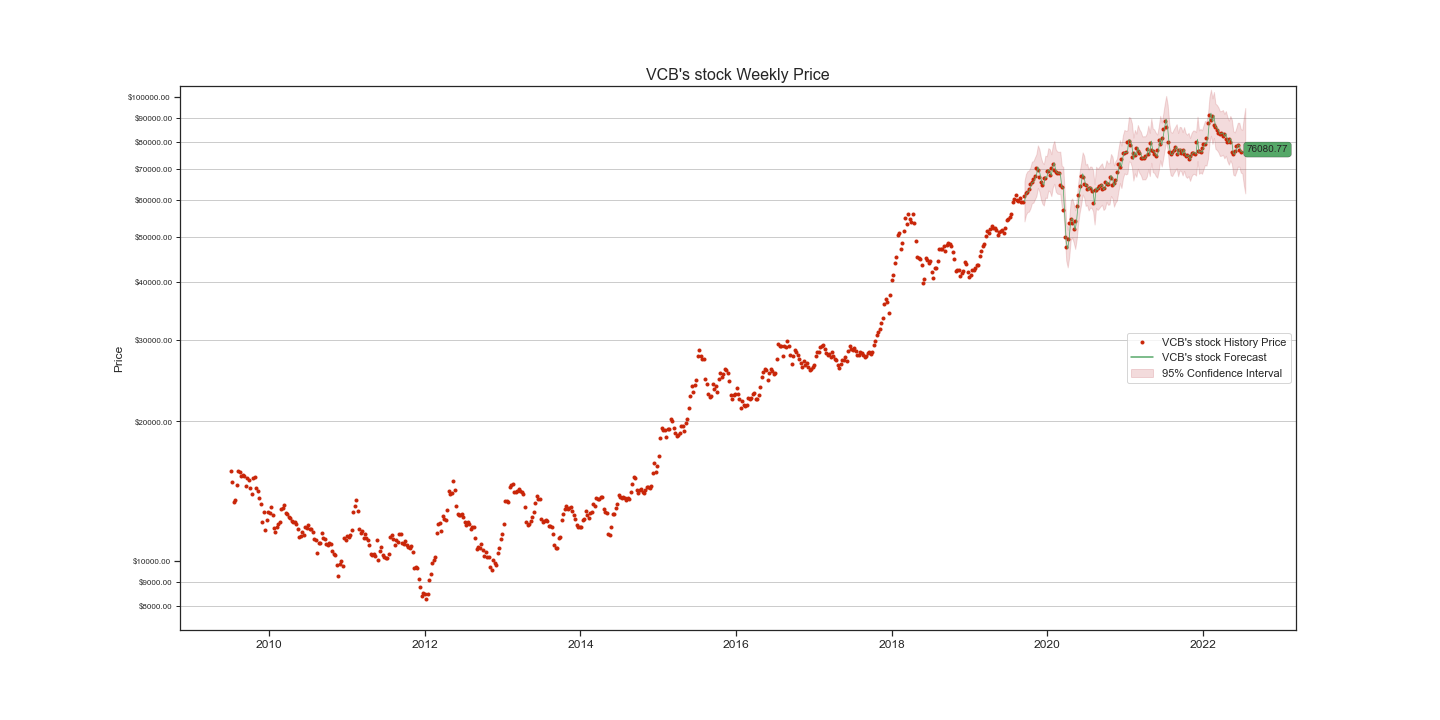
Kết quả chạy mô hình học là từ 2019 trở về trước và mô hình để chạy test là 2019 trở về sau. Đây là code lựa chọn mô hình và vẽ ra kết quả dự đoán

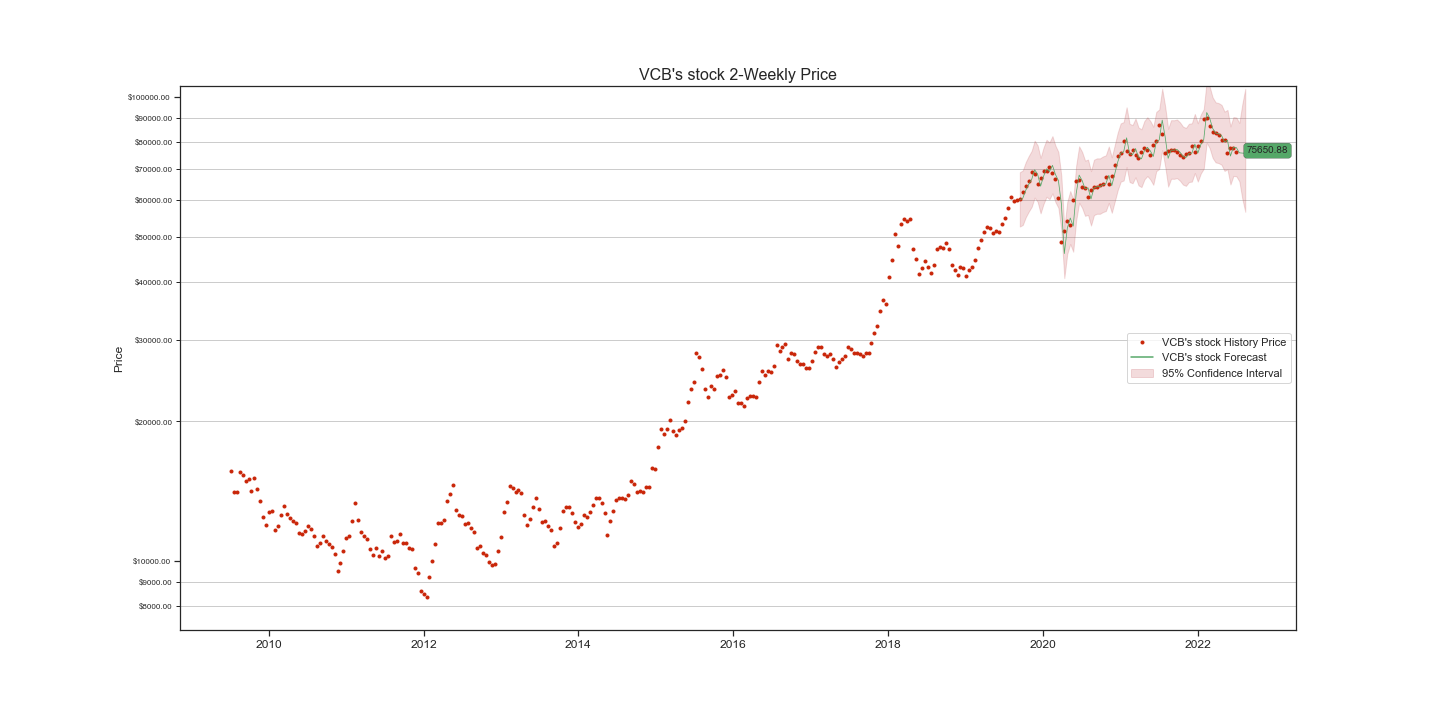
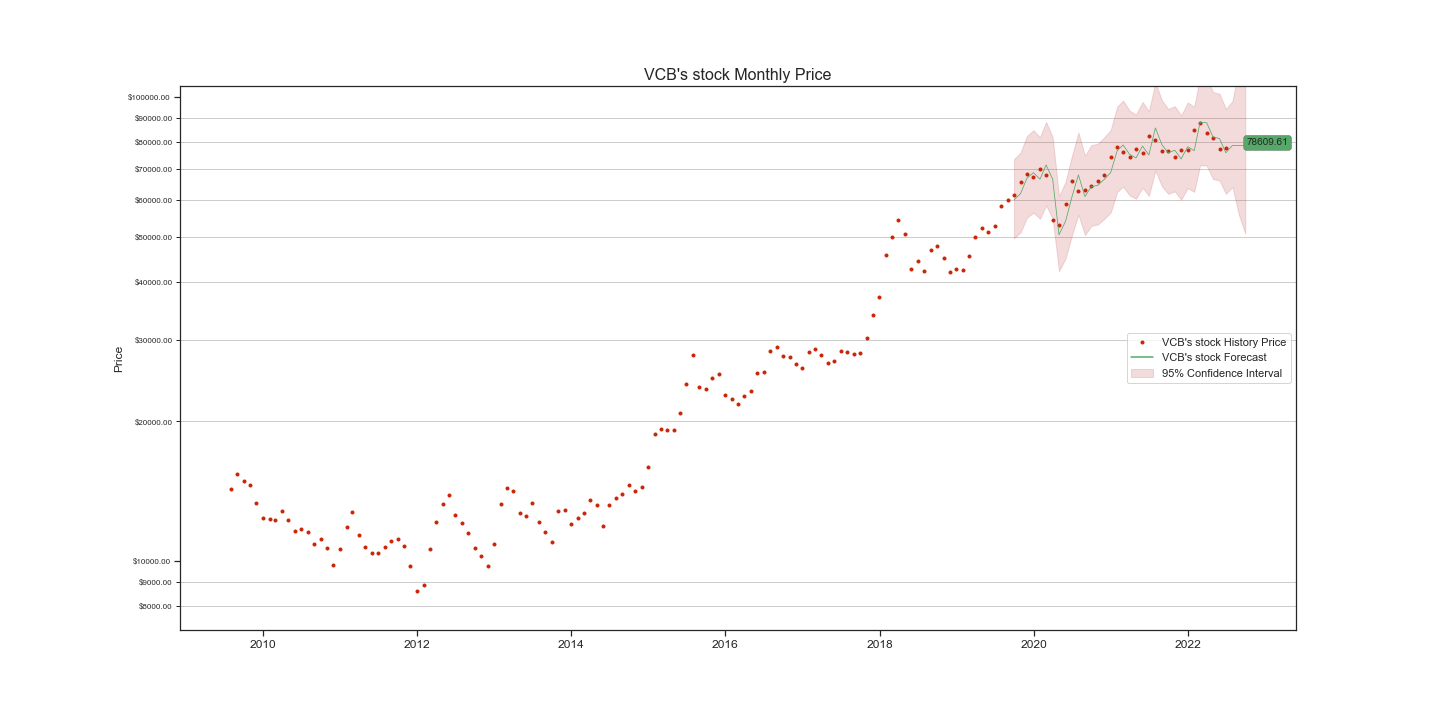




Và đây là kết quả







Nhìn chung kết quả dự đoán của mô hình này khá khả quan hơn mong đợi qua đó cho thấy mô hình khá phù hợp cho việc dự báo lần này