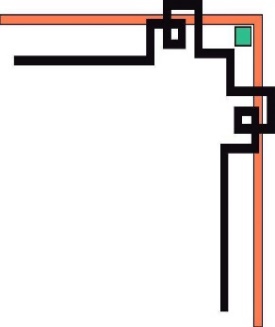
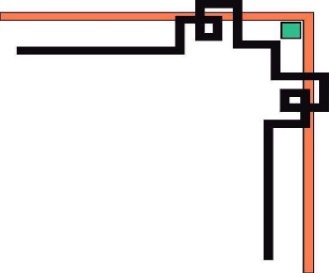
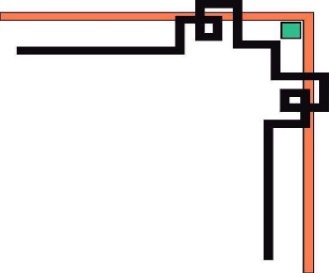
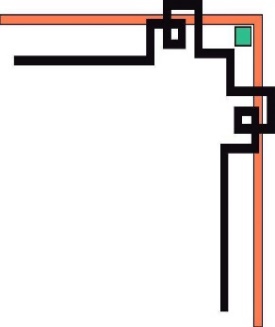
**TRƯỜNG ĐẠI HỌC SƯ PHẠM KỸ THUẬT TP. HCM**



**KHOA ĐÀO TẠO CHẤT LƯỢNG CAO**

**🕯✡🕮🕮✡🕯**



**MÔN : MACHINE LEARNING**

BÁO CÁO

Đề Tài: Ứng dụng thuật toán máy học trong việc phân tích,dự đoán mua bán hàng hóa trong siêu thị

**GVHD: Thầy Vũ Quang Huy**

**Sinh viên thực hiện:**

Vương Văn Hậu – 15110044

Nguyễn Đức Thịnh – 15110132

Hồ Chí Minh tháng 12 năm 2018

**Mục Lục**

[1. Giới thiệu. 4](#_Toc531638480)

[2. Abstract 5](#_Toc531638481)

[3. Mục Tiêu 6](#_Toc531638482)

[4. Các bước thực hiện 6](#_Toc531638483)

[5. Flow chart 7](#_Toc531638484)

[6. Block chart 7](#_Toc531638485)

[7. Methodolory. 8](#_Toc531638486)

[8. Implementation. 13](#_Toc531638487)

[9. Test. 18](#_Toc531638488)

[10. Result and Accuracy. 19](#_Toc531638489)

**BẢNG PHÂN CHIA CÔNG VIỆC**

|  |  |
| --- | --- |
| **Tasks** | **Assignment** |
| Documents for project and Definition | **Thịnh** |
| Determine purpose of forecast | **Hậu** |
| Collect Data set | **Thịnh, Hậu** |
| Choose and Analysis (Data exploration,algorithms) | **Thịnh, Hậu** |
| Analysis Techniques | **Hậu** |
| Methodology(Data Preprocessing) | **Thịnh, Hậu** |
| Implement | **Thịnh, Hậu** |
| Check forecast accuracy | **Thịnh, Hậu** |

# 1. Giới thiệu.

Trong môi trường thị trường hỗn loạn hiện tại, dự báo khối lượng cầu và

phát triển tương lai có vai trò không thể thay thế trong việc quản lý mọi bán lẻ

kinh doanh. Dự báo nhu cầu đáng tin cậy là điều kiện tiên quyết cho hiệu quả và hiệu quả sự hài lòng của khách hàng. Nó có thể sử dụng hiệu quả các nguồn lực và giảm chi phí liên quan đến dư thừa hoặc thiếu hàng hóa trong chuỗi cung ứng, vì có liên quan. Các chất lượng của quá trình dự báo nhu cầu trong lĩnh vực bán lẻ được liên kết chặt chẽ với lựa chọn các phương pháp dự báo phù hợp dựa trên thông tin liên quan từ các nguồn thông tin có sẵn cho các nhà bán lẻ và đồng thời có thể chuyển thành kết quả dự báo ảnh hưởng của tất cả các yếu tố bên trong và bên ngoài đáng kể , ảnh hưởng đến nhu cầu của khách hàng trong tương lai trên thị trường B2C.

Trong thị trường B2C, nhu cầu chủ yếu chịu ảnh hưởng của xã hội, văn hóa, tâm lý

và các yếu tố nhân cách. Mỗi khách hàng đều có ưu tiên riêng của mình, nhưng cũng có lý do để mua một số hàng hóa. Điểm quan trọng là do đó tính cách và lối sống của khách hàng.

Khi lựa chọn thực phẩm, khách hàng cũng xem xét một số yếu tố vốn có trong

thức ăn đặc biệt. Thông thường, họ đánh giá tính mới mẻ, tự nhiên và chi phí của

thức ăn . Nhu cầu về thực phẩm bị ảnh hưởng bởi nhiều yếu tố. Bên cạnh chất lượng hàng hóa và hành vi thường xuyên của người mua và giá của hàng hóa, nó cũng đề cập đến các hình thức xúc tiến bán hàng khác nhau trong kinh doanh bán lẻ, đề xuất từ ​​các khách hàng khác, lòng trung thành của thương hiệu, v.v.

Các nhà bán lẻ có thể chọn từ nhiều phương pháp dự báo khác nhau. Sử dụng định tính phương pháp dự báo (ý kiến ​​của nhà bán lẻ, ước tính chuyên gia, nghiên cứu tiếp thị) không phải là điển hình để dự đoán nhu cầu đối với một loạt hàng hóa trên thị trường tiêu dùng, bởi vì nó rất tốn thời gian và tốn kém . Do đó, các nhà bán lẻ chủ yếu làm việc với phân tích chuỗi thời gian của doanh số bán hàng trong quá khứ, cho phép chúng tương đối dễ dàng để xác định biến động theo mùa về doanh thu cũng như thay đổi xu hướng trong toàn bộ sản phẩm vòng đời. Để định lượng tác động của giá sản phẩm và khuyến mại trên mô hình yêu cầu, hồi quy hoặc kỹ thuật khai thác dữ liệu được sử dụng, dựa trên trên cơ sở dữ liệu lớn về thông tin về hành vi của khách hàng dưới dạng điểm dữ liệu bán hàng (dữ liệu POS). Nó là một cơ sở dữ liệu về doanh số bán hàng trong quá khứ trong bối cảnh thực hiện

các hình thức khuyến mãi và kết hợp khác nhau của chúng, mức độ nhạy cảm về giá của người tiêu dùng, hiệu ứng ăn thịt người, ảnh hưởng của thời tiết, vv Dữ liệu POS đại diện cho nguồn thông tin có giá trị về nhu cầu, vì dữ liệu về bán hàng trong thương mại bán lẻ không bị bóp méo bởi các hiệu ứng liên quan đến chuỗi và cũng có sẵn từ bán lẻ .Chất lượng của quá trình dự báo chủ yếu được phản ánh trong tính chính xác của dự báo được sản xuất, là một trong những nguyên nhân chính gây ra hiệu ứng bullwhip trong chuỗi cung ứng. Dự báo không chính xác của các nhà bán lẻ tạo ra các đơn đặt hàng không đầy đủ cho các liên kết tiếp tục xuống chuỗi cung ứng và gây ra nhiều khó khăn trong việc phối hợp các luồng nguyên liệu trong chuỗi cung ứng.

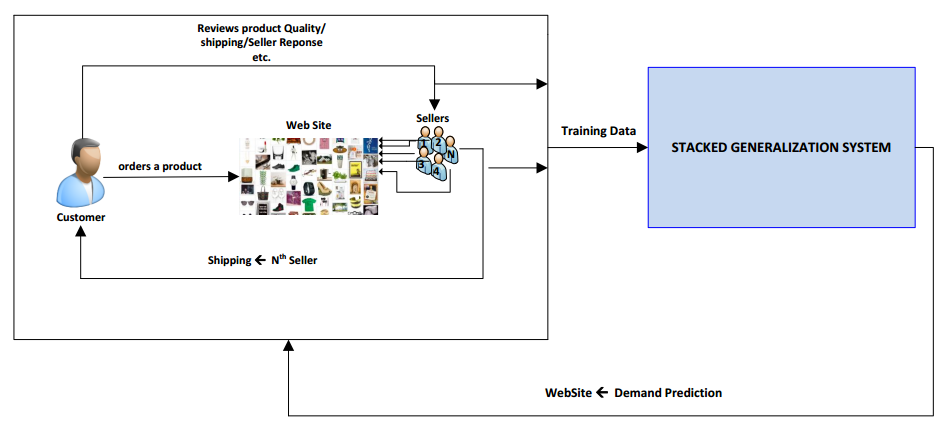
# 2. Abstract

Dự báo nhu cầu đóng một vai trò quan trọng trong việc quản lý các quy trình hậu cần trong bán lẻ. Chất lượng của các dự báo kết quả có ảnh hưởng trực tiếp đến hiệu quả của các quy trình không chỉ trong các cửa hàng siêu thị bán lẻ, mà còn trong toàn bộ chuỗi cung ứng. Các cửa hàng siêu thị bán lẻ trên thực tế liên hệ hàng ngày với khách hàng cuối cùng trong chuỗi cung ứng và do đó họ phần lớn tham gia vào dự báo trong việc áp dụng hợp tác các chiến lược trong quản lý chuỗi cung ứng (ví dụ: Đáp ứng nhanh, Người tiêu dùng hiệu quả Đáp ứng, lập kế hoạch hợp tác, dự báo và bổ sung).

Mức dự báo nhu cầu trong các cửa hàng siêu thị bán lẻ tại Việt Nam là chủ đề của nghiên cứu định lượng. Mục tiêu chính của nghiên cứu là tìm hiểu phương pháp nào được sử dụng trong dự báo nhu cầu, những yếu tố ảnh hưởng đến nhu cầu thực phẩm được xem xét và những nguồn thông tin cửa hàng bán lẻ sử dụng để căn cứ dự báo trên. Khi phân tích dữ liệu thu được thông qua các yêu cầu cá nhân tại 75 cửa hàng siêu thị bán lẻ, các phương pháp thống kê mô tả và suy luận được sử dụng.Nghiên cứu cho thấy rằng các nhà bán lẻ phụ thuộc vào nhu cầu dự báo chủ yếu dựa trên chính họ trực giác và kinh nghiệm trong bán lẻ, và do đó các phương pháp dự báo định tính là áp dụng thường xuyên nhất. Mặc dù các nhà bán lẻ xem xét nhiều yếu tố ảnh hưởng đến nhu cầu, nghiên cứu đã xác định sự vắng mặt của các phương pháp thống kê có thể thực hiện hiệu quả ảnh hưởng của các yếu tố vào các dự đoán cuối cùng. Tỷ lệ áp dụng định lượng thấp các phương pháp dự báo trong các cửa hàng bán lẻ và sự sẵn có của một cơ sở đầu vào nhỏ thông tin về nhu cầu của khách hàng cuối có thể có tác động tiêu cực đến thực hiện các phương pháp hiện đại trong quản lý chuỗi cung ứng.

# 3. Mục Tiêu

Dự báo không dùng máy tính thường không tốt bằng dự đoán máy tính kết hợp thông tin chi tiết về chất lượng. Dự báo hiếm khi (nếu có) chính xác… một trong những mục tiêu dự báo là sai lầm “ít nhất”. Dự báo chính xác hơn nếu được phát triển ở nhóm chứ không phải cấp mục. Độ chính xác của dự báo lớn nhất cho các khoảng thời gian gần hơn với thời gian dự báo được phát triển.Vì vậy mục tiêu chính là có thể dự đoán ít sai nhất về sản phẩm sẽ được tiêu thụ trong tương lai để tránh việc tồn kho gây lỗ.



# 4. Các bước thực hiện

**Bước 1:** Xác định mục đích dự báo

**Bước 2:** Thiết lập một đường thời gian

**Bước 3:** Chọn kỹ thuật dự báo

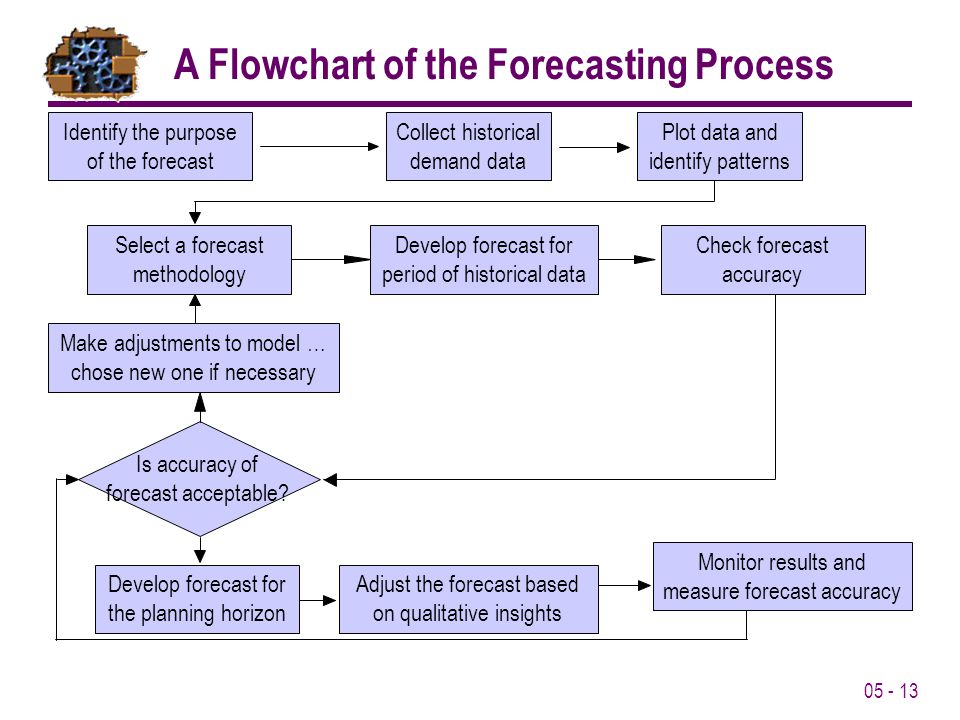
**Bước 4:** Thu thập và phân tích dữ liệu

**Bước 5:** Chuẩn bị dự báo

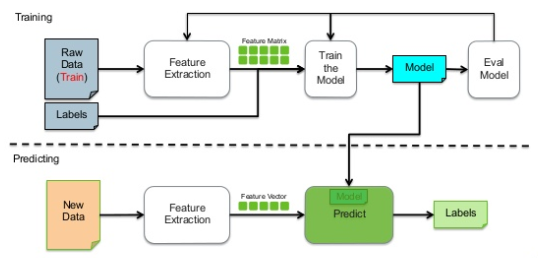
**Bước 6:** Theo dõi dự báo “Dự báo”. Xác định mục đích của dự báo (lập kế hoạch phạm vi dài, giới thiệu sản phẩm mới, các hoạt động hàng ngày).

Thiết lập một đường thời gian mà dự báo sẽ bao gồm (hãy nhớ rằng độ chính xác dự báo sẽ giảm dần theo thời gian). Thu thập dữ liệu nhu cầu lịch sử và chọn phương pháp luận (thường là kết hợp thống kê / toán học, máy tính và hiểu biết về chất lượng). Chuẩn bị dự báo. Theo dõi dự báo

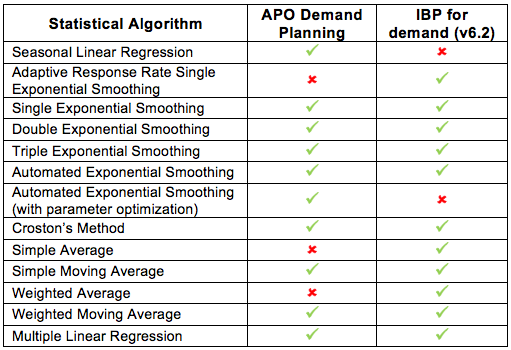
# 5. Flow chart



# 6. Block chart



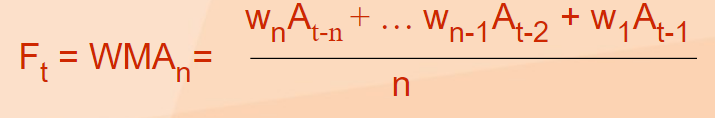
# 7. Methodolory.



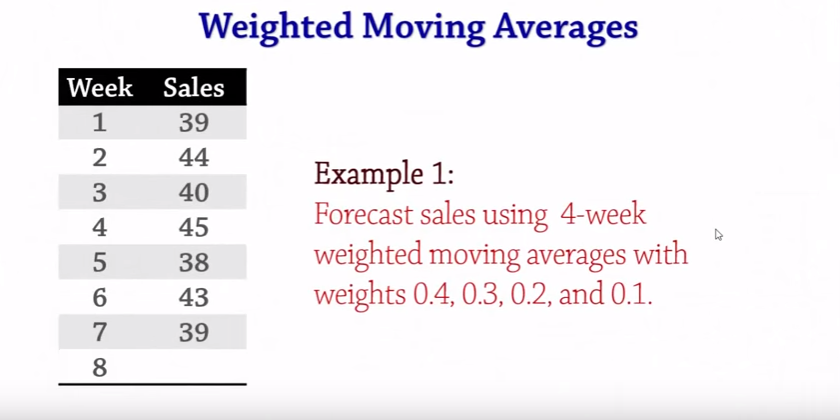
Có rất nhiều thuật toán có thể áp dụng cho bài toán dự báo nhu cầu. Trong báo giới thiệu này, nhóm áp dụng thuật toán Weighted Moving Averages.

Thuật toán Weighted Moving Averages (WMA): Các giá trị gần đây hơn sẽ được cho trọng số cao hơn.

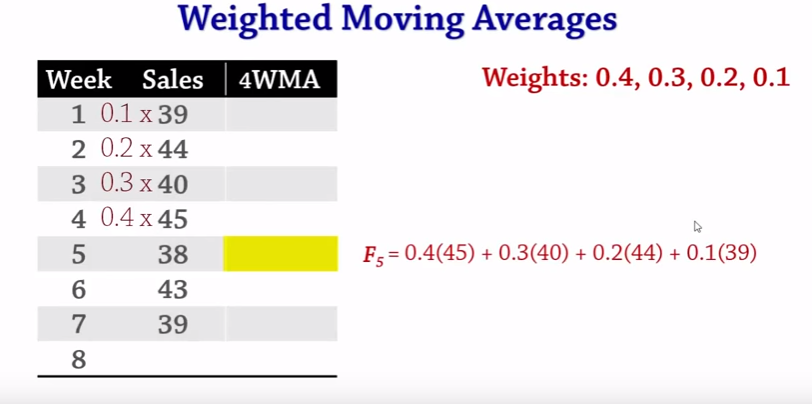
WMA được tính theo công thức:

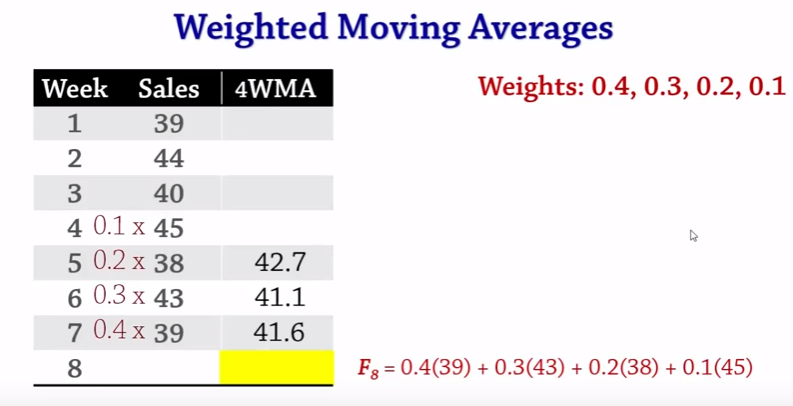


Example:

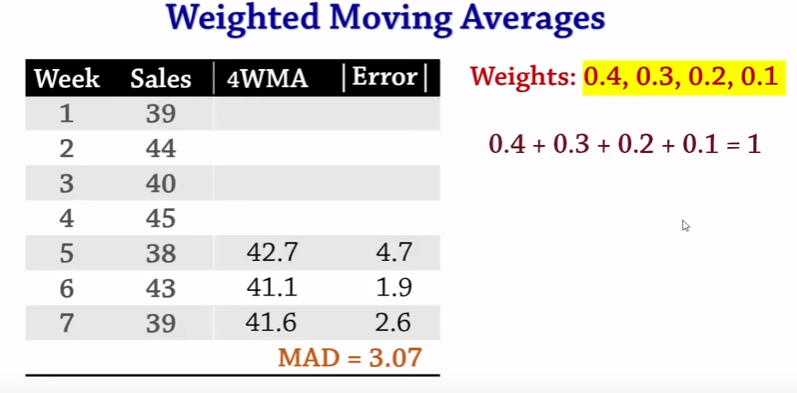


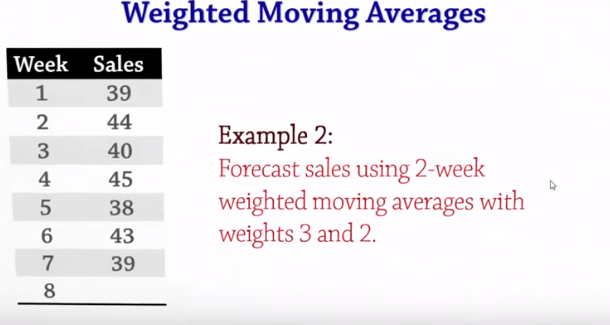
Chọn trọng số và tính WMA cho các tháng.

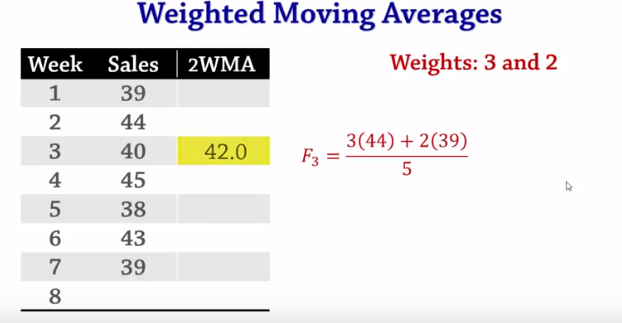




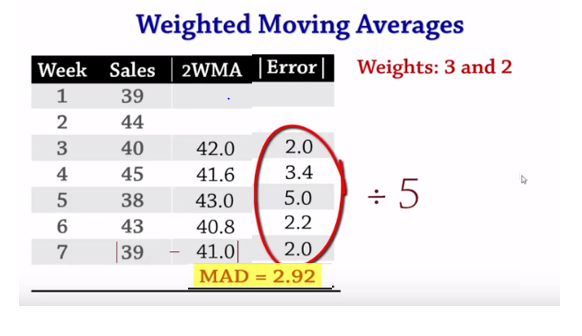
Sau đó tính sai số so với thực tế.



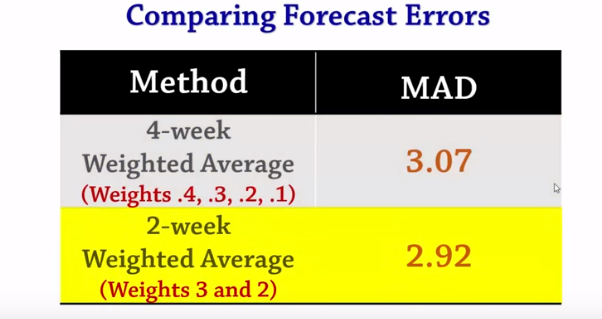


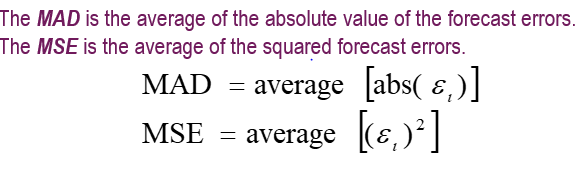


Tính trung bình cộng các sai số.



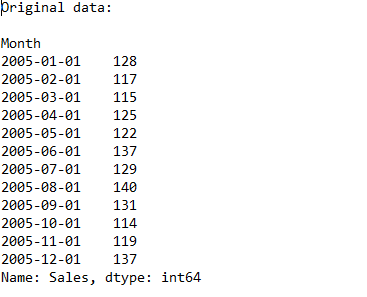
Sau khi dự báo chúng ta sẽ xét xem mức độ sai lệt so với thực tế bằng cách trung bình cộng các sai số tính được và chọn được trọng số tối ưu nhất để làm trọng số dự đoán cho bài toán.



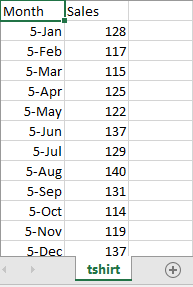


# 8. Implementation.

Thu thập dữ liệu bán hàng về 1 sản phẩm đó là áo tay dài

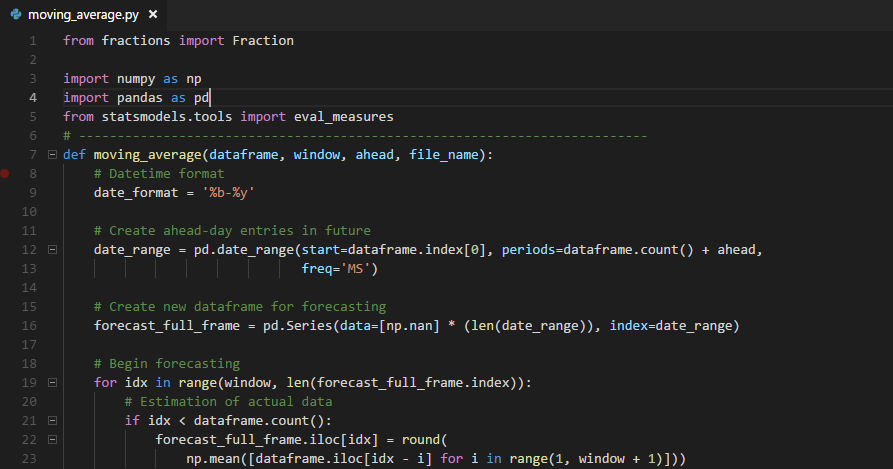


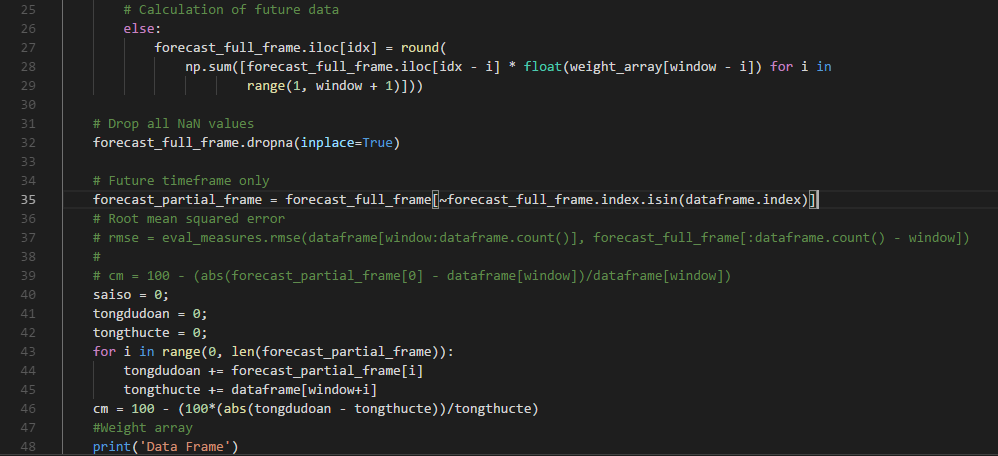
Sau khi đã thu tập dữ liệu ta sẽ làm cho dữ liệu trở nên dễ dàng sử dụng hơn trong thuật toán Weighted Moving Averages để có thể dự đoán được chính xác hơn về năng suất bán cho tháng tiếp theo và sau đó nữa.

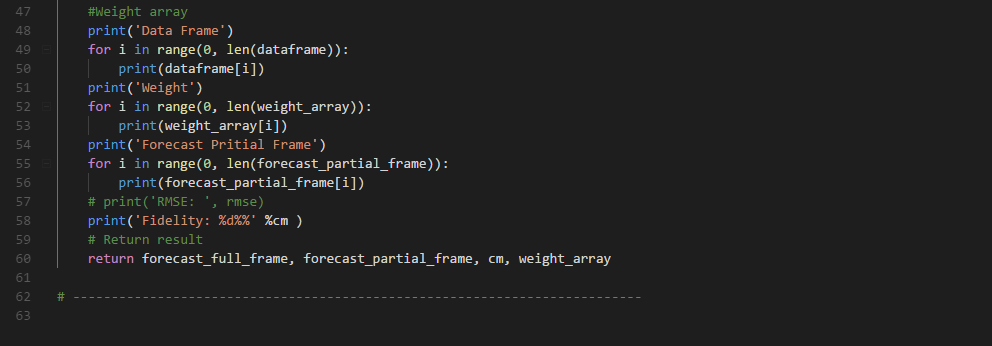


Sau khi đã có dữ liệu tối ưu thì ta sẽ thực hiên việc coding và áp dụng dữ liệu trên

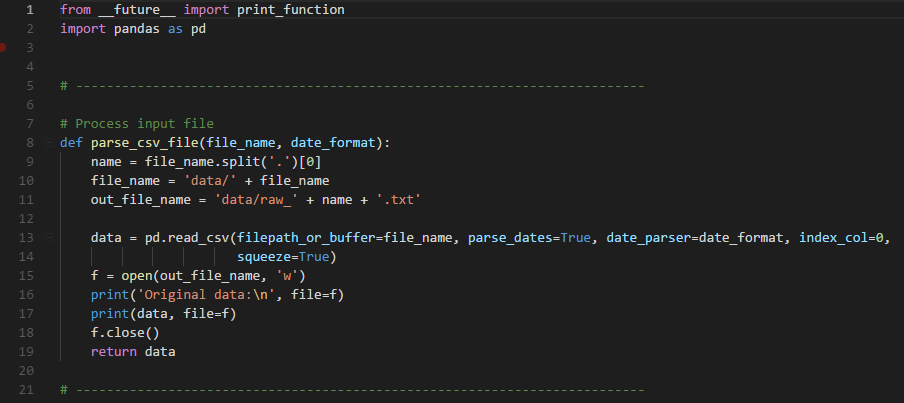
**Bước 1:** Ta sẽ tạo lớp “moving\_average.py” sẽ là lớp (model) chứa thuật toán để dự đoán.



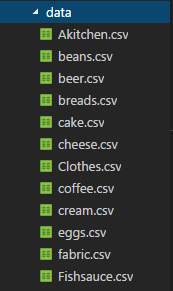




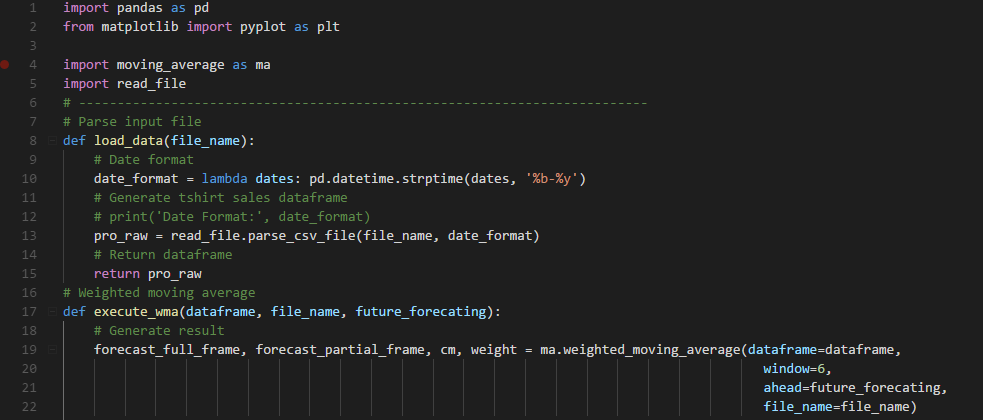
**Bước 2:** Ta cần phải có 1 lớp để đọc file data đầu vào “read\_file.py”.

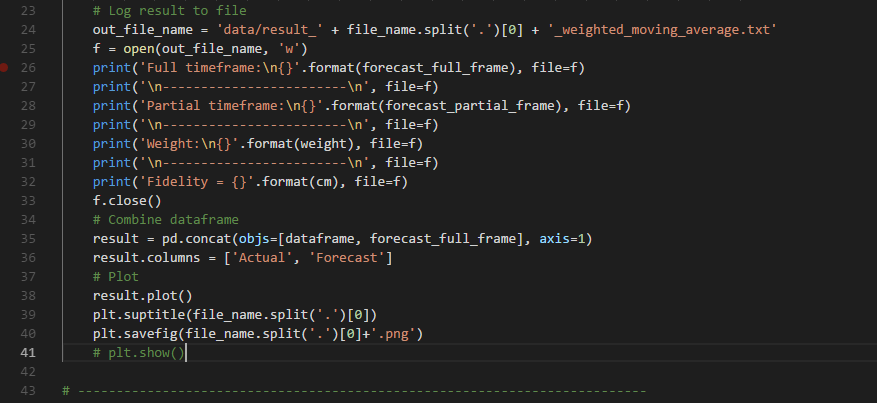


File đầu vào sẽ được đọc dưới dạng .csv

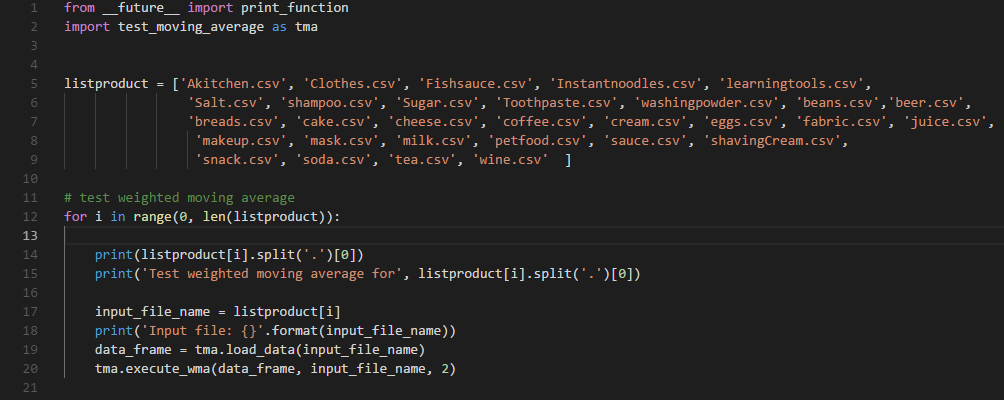


**Bước 3:** Tiếp theo là lớp test để kiểm tra kết quả sau khi training và đưa ra dự đoán cho sản phẩm đó trong tương lai gần. “test\_moving\_average.py”



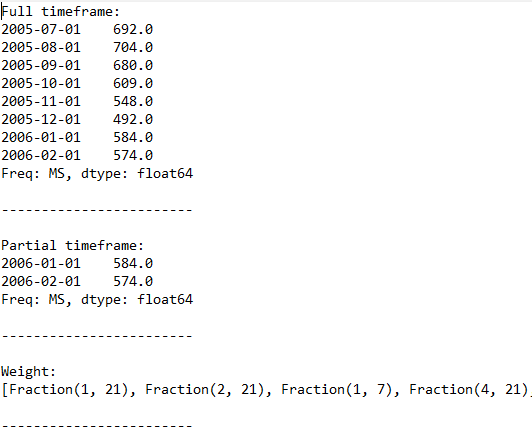


**Bước 4:** Ta sẽ tạo lớp “main.py” để chạy thuật toán dự đoán vừa viết trên.

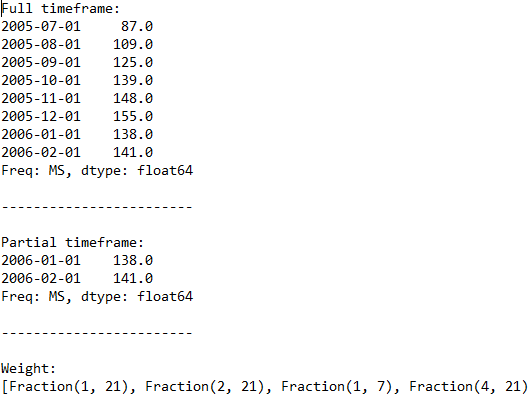


# 9. Test.

Đây là file test đầu ra cho sản phẩm soda của cửa hàng bán lẻ



Đây là file test đầu ra cho sản phẩm petfood của cửa hàng bán lẻ



# 10. Result and Accuracy.

Sau khi chạy main.py thì phần đồ thị sẽ được biểu diễn cho thấy được kết quả dự đoán của thuật toán.

