

Họ và tên: Nguyễn Tuấn Khang

MSV: 2151264665

- Kịch bản chia data:

+ Dữ liệu được đọc từ tệp CSV 'Gia SMP va SMPcap 2021(Giá thị trường SMP).csv'.

+ Chọn cột 6 và 7 để phân tích.

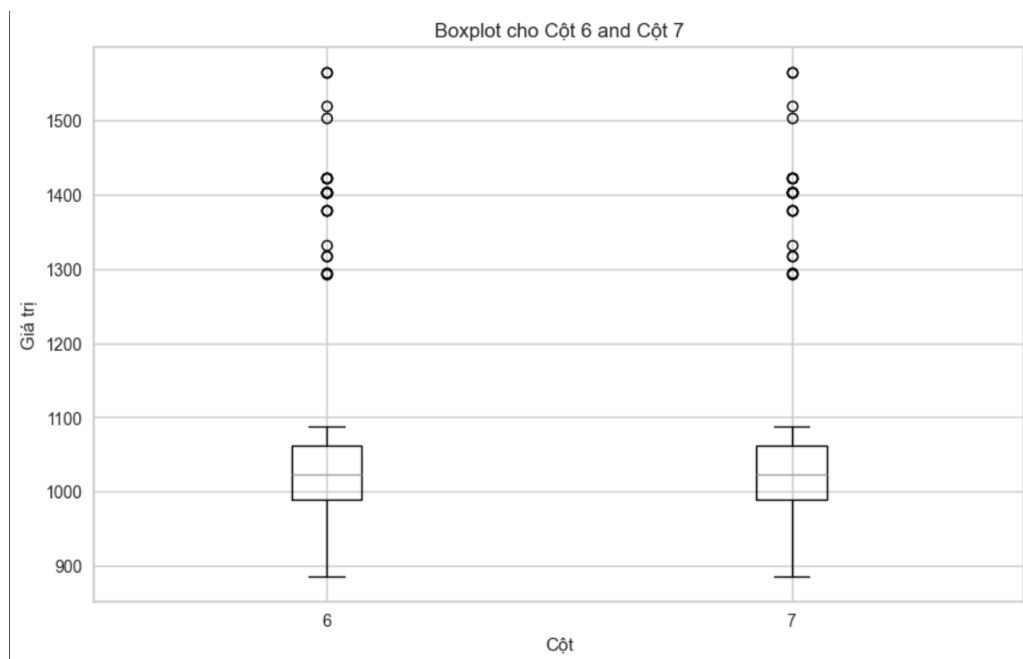
+ Trực quan hóa dữ liệu bằng các biểu đồ Plotbox, Scatterplot, Line Graph và Bar graph

+ Sử dụng Gaussian Mixture với số components là 7.

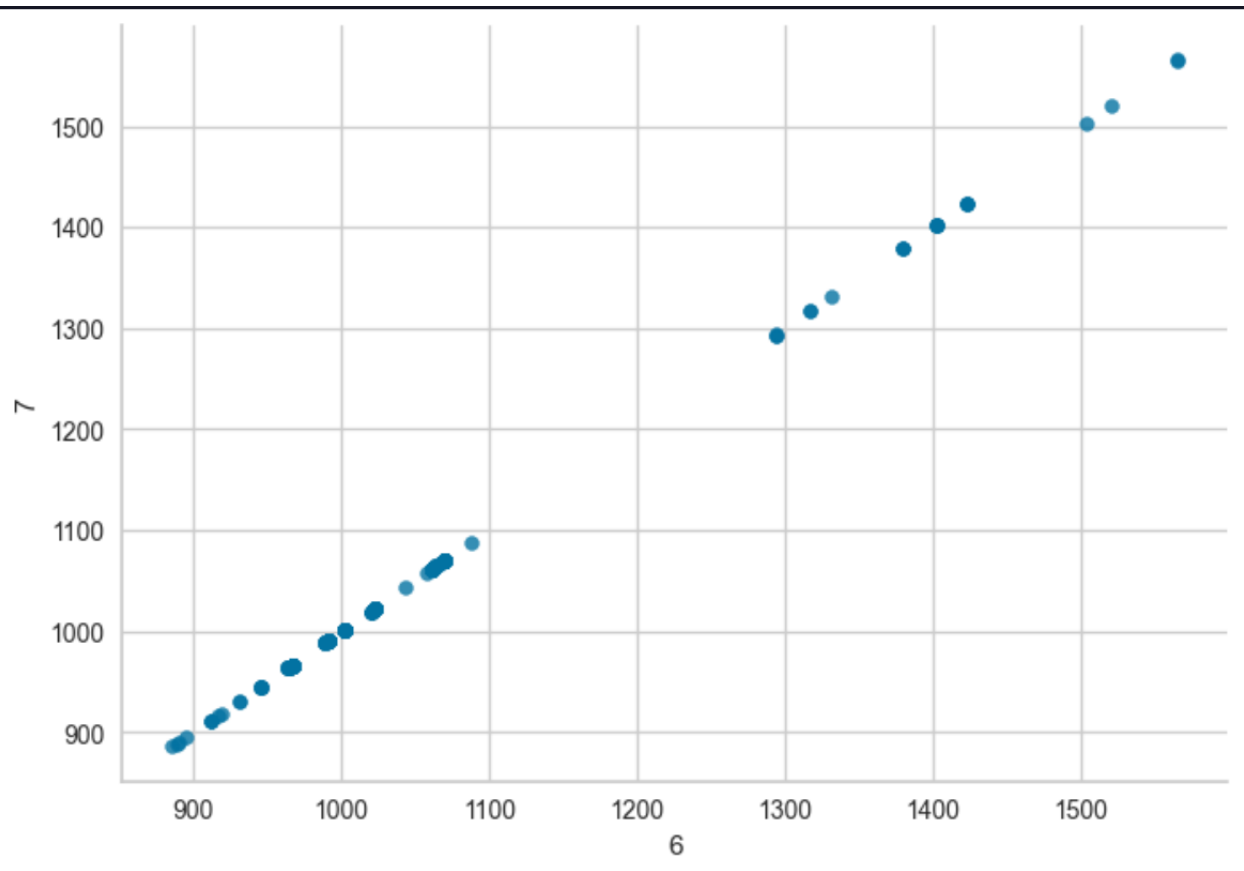
+ Sử dụng Sarima dự đoán.

- Ảnh trực quan hóa:

+ Plotbox:

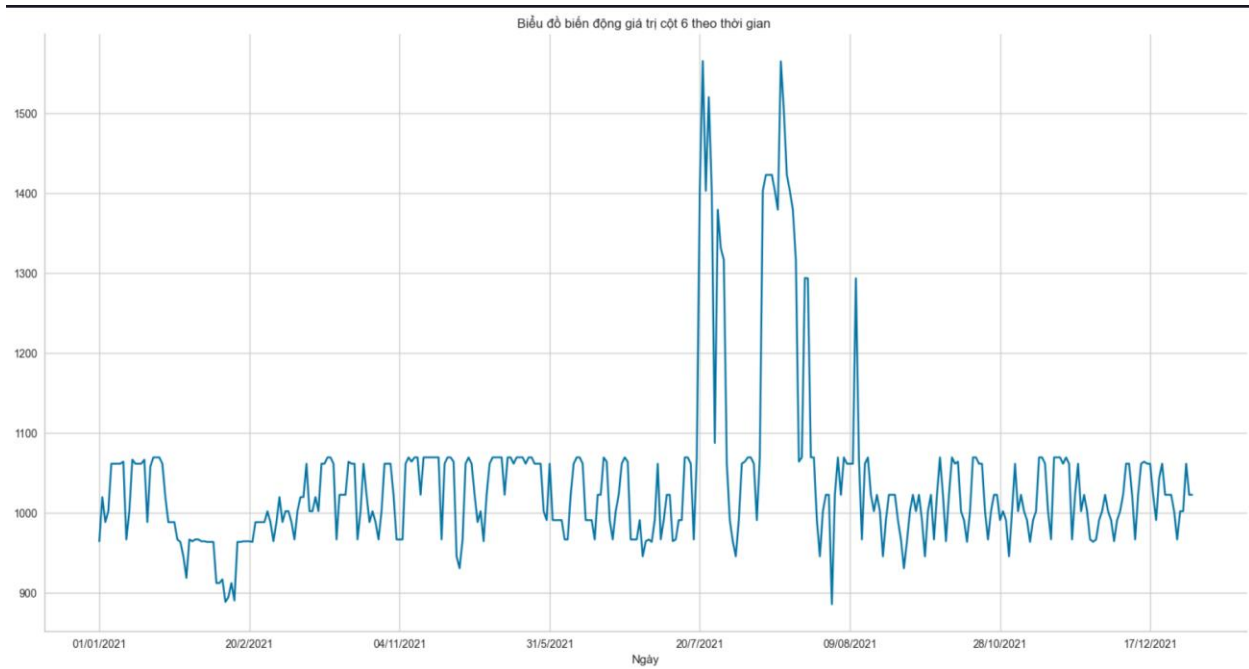


+ Scatterplot:

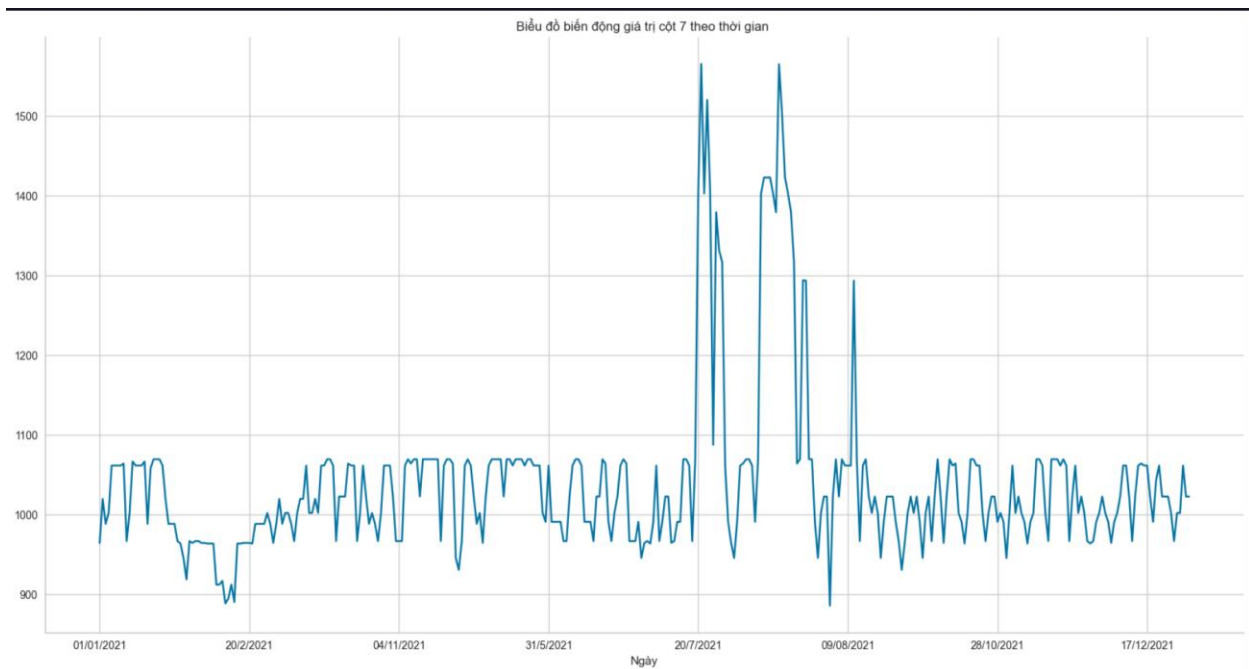


+ Line Graph:

- Cột 6

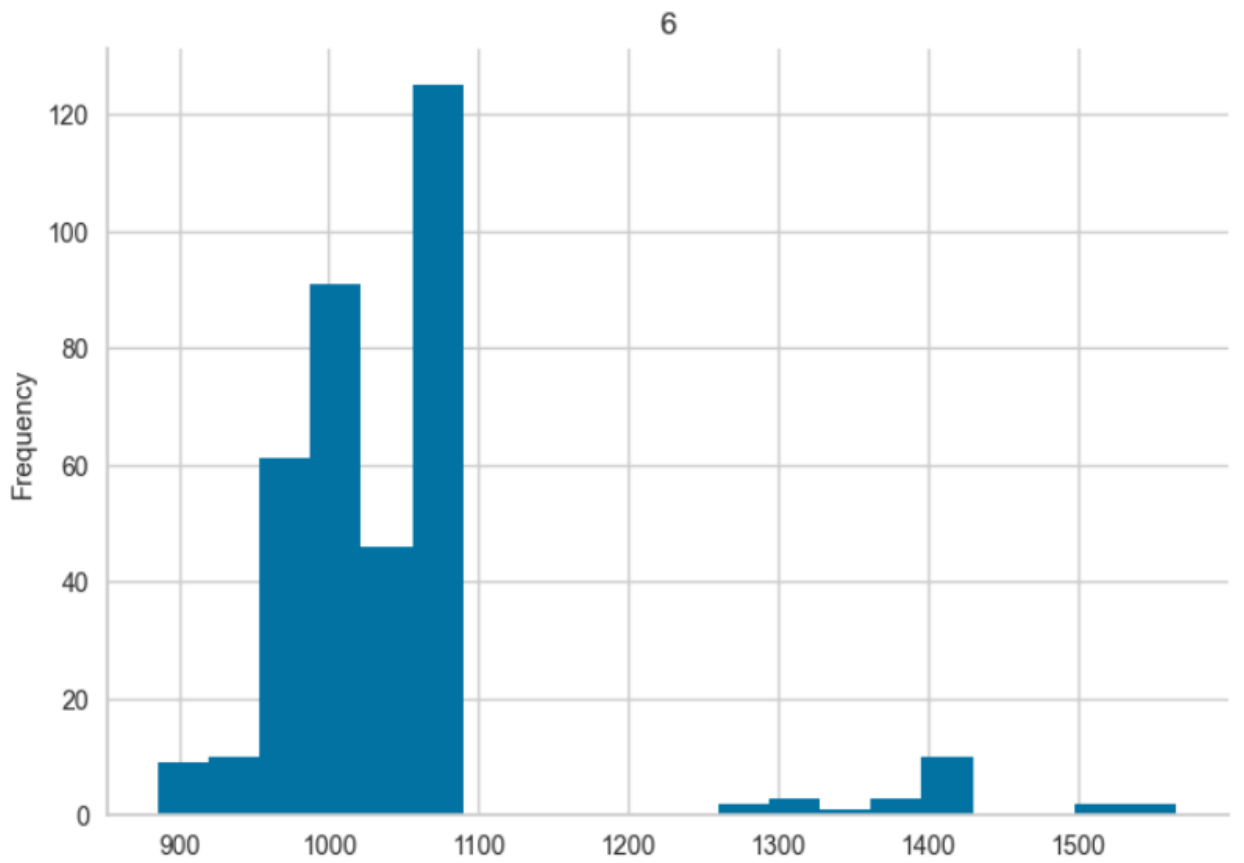


- Cột 7:

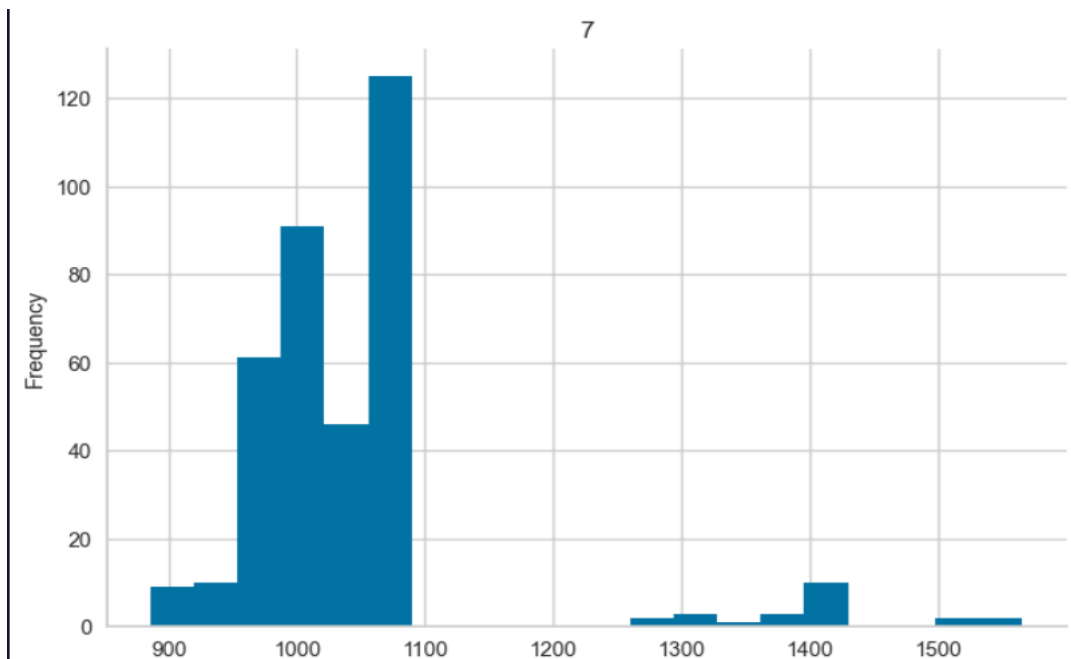


+ Bar Graph:

- Cột 6:



- Cột 7:



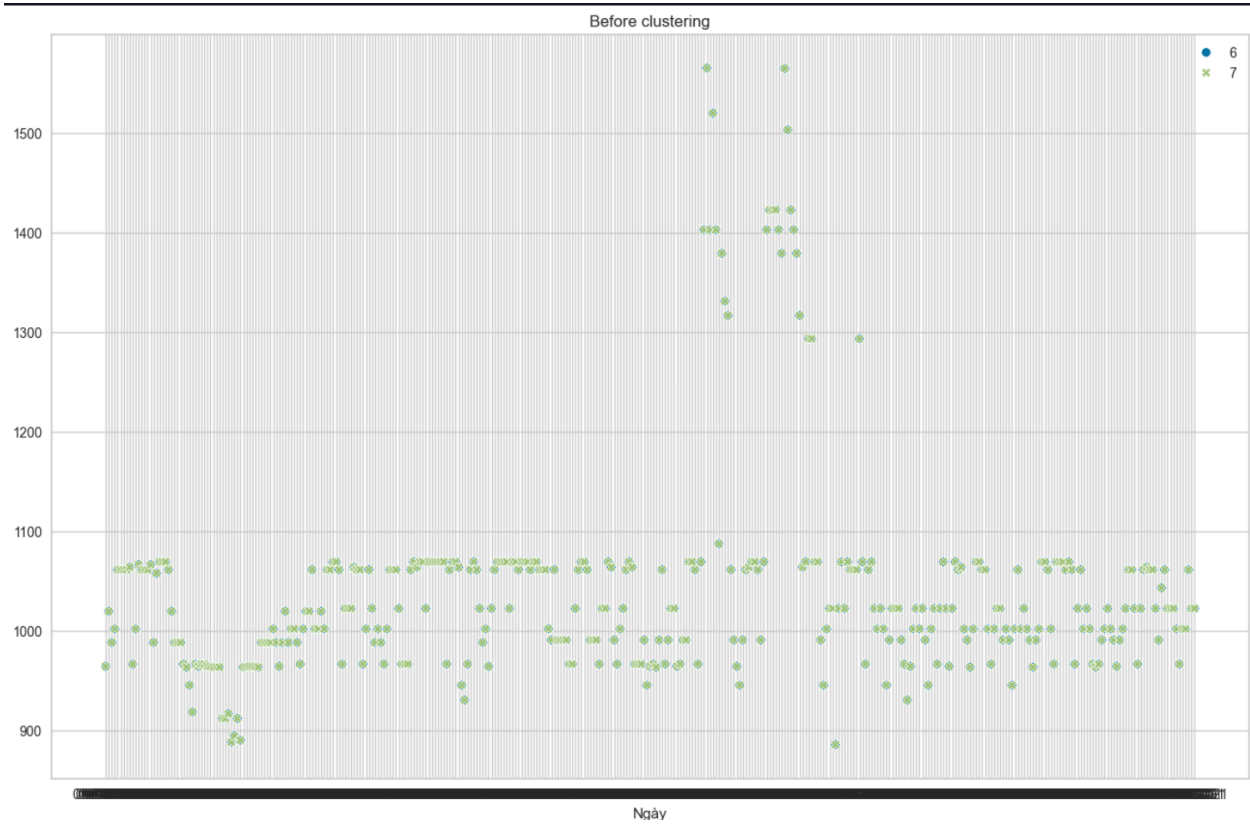
- Anh Training:

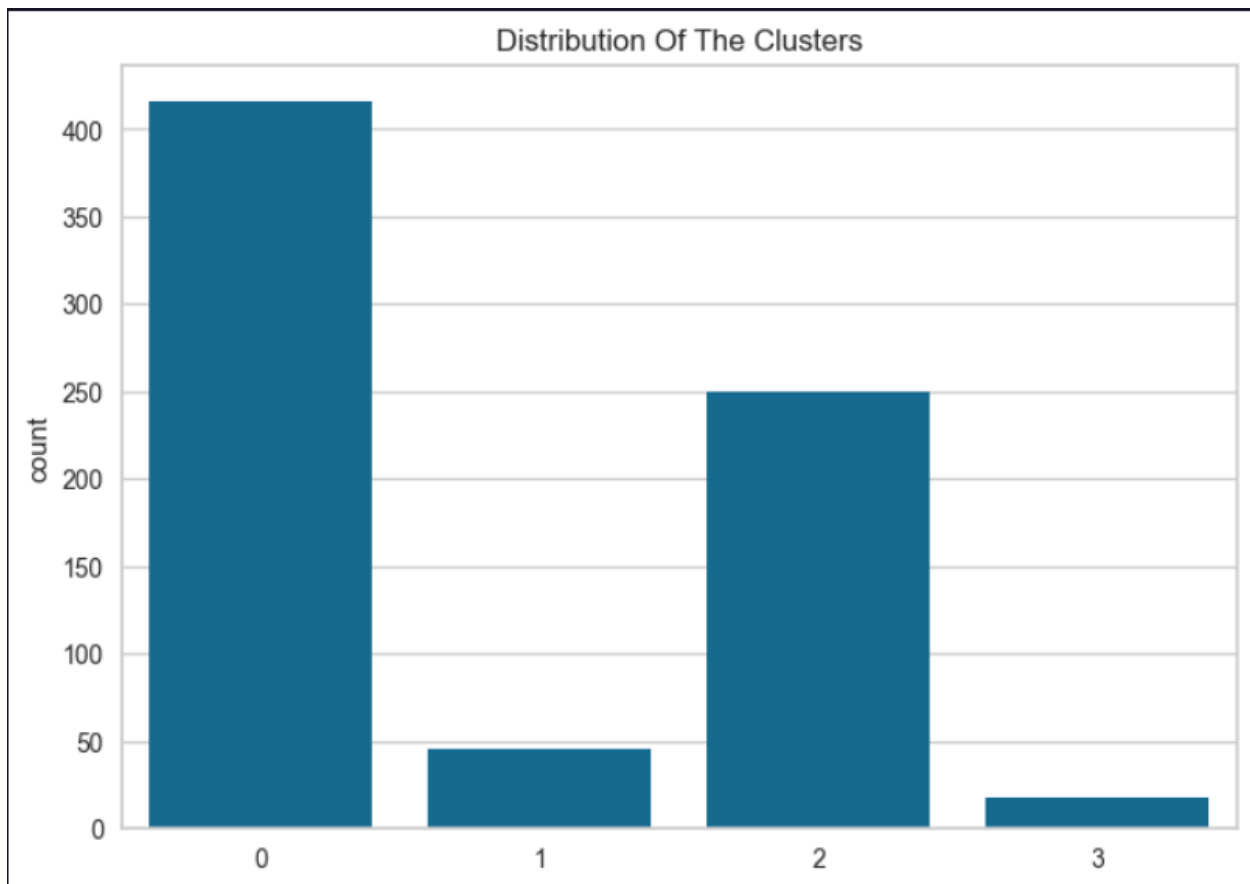
+ BayesianGaussianMixture

```
[200]    valid_0's multi_logloss: 1.71072e-05
[400]    valid_0's multi_logloss: 1.71072e-05
[200]    valid_0's multi_logloss: 2.9727e-06
[400]    valid_0's multi_logloss: 2.9727e-06
[200]    valid_0's multi_logloss: 2.96634e-06
[400]    valid_0's multi_logloss: 2.96634e-06
[200]    valid_0's multi_logloss: 2.80756e-06
[400]    valid_0's multi_logloss: 2.80756e-06
[200]    valid_0's multi_logloss: 3.11811e-06
[400]    valid_0's multi_logloss: 3.11811e-06
[200]    valid_0's multi_logloss: 3.31111e-06
[400]    valid_0's multi_logloss: 3.31111e-06
[200]    valid_0's multi_logloss: 5.75321e-06
[400]    valid_0's multi_logloss: 5.75321e-06
[200]    valid_0's multi_logloss: 3.26431e-06
[400]    valid_0's multi_logloss: 3.26431e-06
[200]    valid_0's multi_logloss: 3.13843e-06
[400]    valid_0's multi_logloss: 3.13843e-06
[200]    valid_0's multi_logloss: 3.13837e-06
[400]    valid_0's multi_logloss: 3.13837e-06
[200]    valid_0's multi_logloss: 4.06637e-06
[400]    valid_0's multi_logloss: 4.06637e-06
```

- Ảnh kết quả:

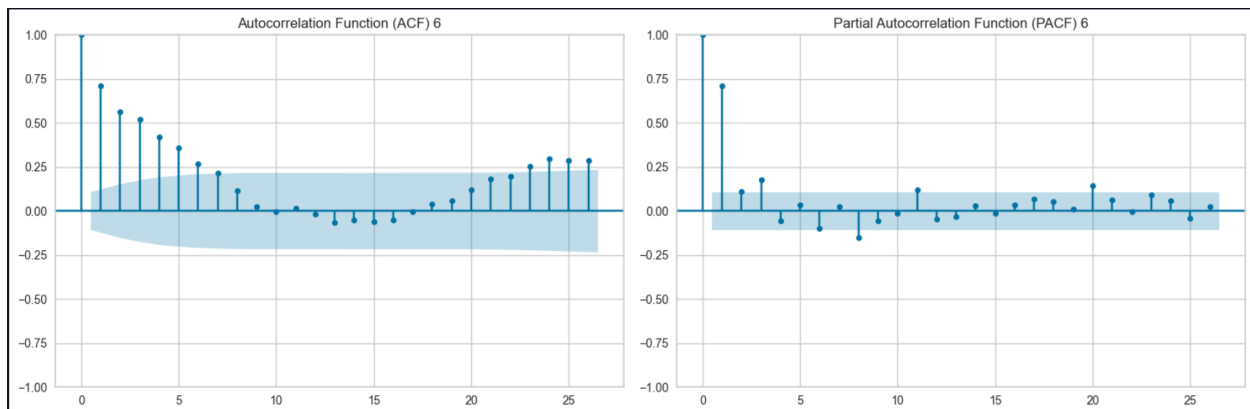
+ BayesianGaussianMixture:



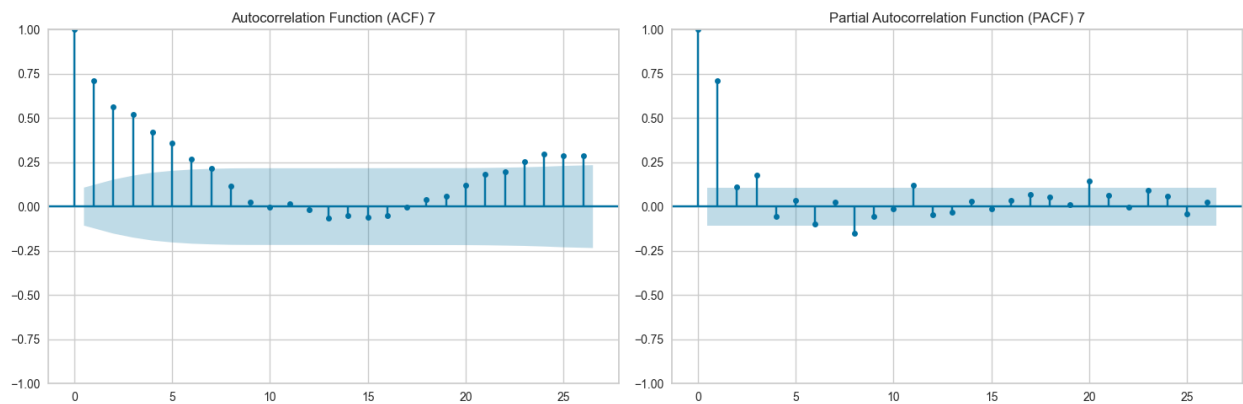


+ SARIMA:

Cột 6:

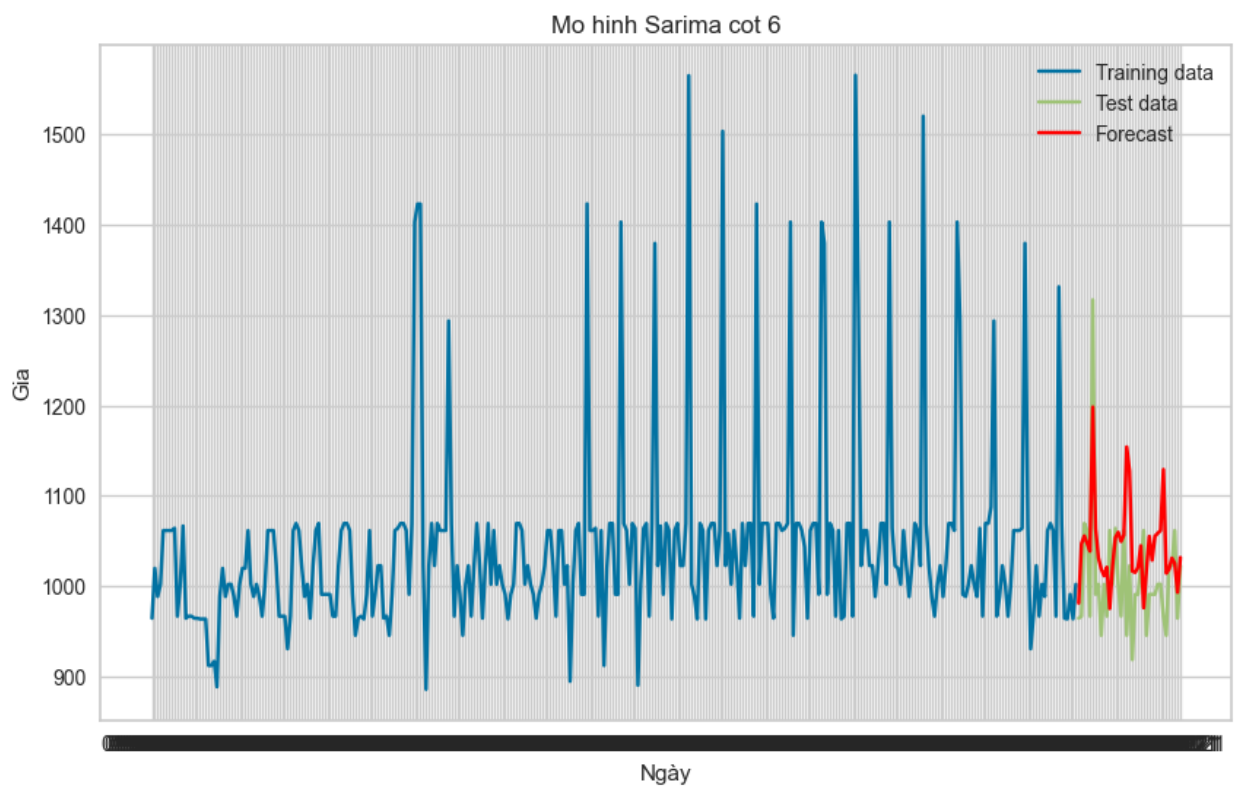


Cột 7:

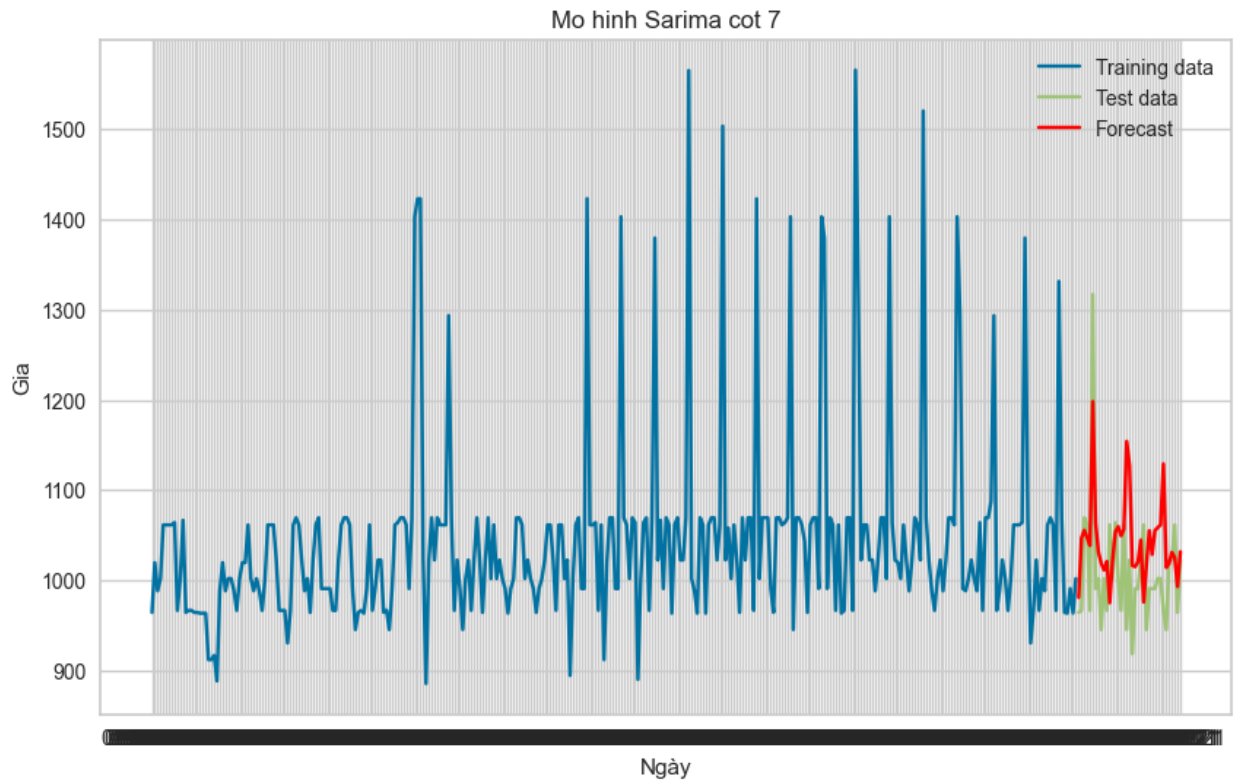


+ Dự đoán:

- Cột 6:



- Cột 7:



Link Github: <https://github.com/tu4nkhag/phan-tich-chuoi-thoi-gian/blob/main/TH2.ipynb>