Time Series Analysis to predict Medical Properties Trust stock

Veronika Titarchuk

May 2024

1 Introduction

Proficiency in forecasting future stock performance is critical for financial institutions, analysts, and investors alike in the world of financial markets. The development of sophisticated analytical methods like time series analysis has given scholars and industry professionals stronger instruments for understanding complex patterns found in stock market data.

Medical Properties Trust (MPT), a real estate investment trust (REIT) specializing in healthcare facilities, stands as a compelling subject for such analysis. MPT presents a compelling case study for examining the accuracy of time series analysis in predicting stock performance because of its concentration on owning and leasing healthcare properties in the US and other nations.

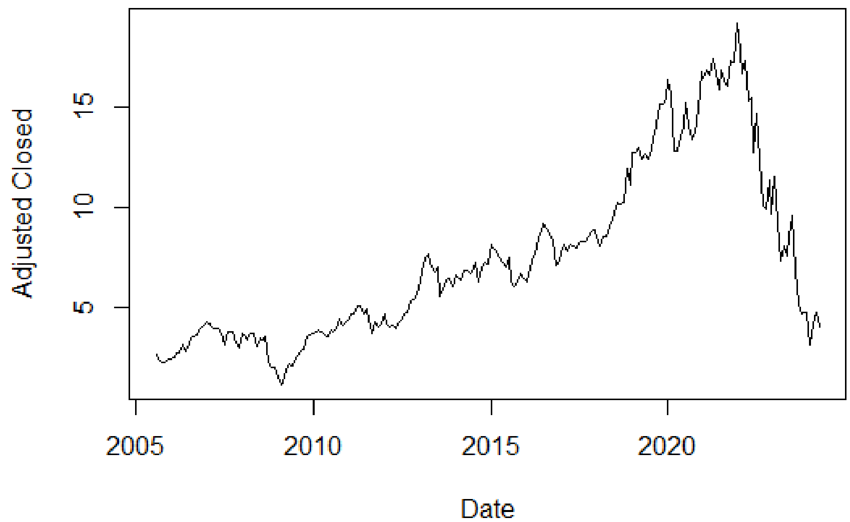
The corporation boasts a portfolio comprising 439 properties and encompassing 43,000 licensed beds, with operational presence extending across nine nations, namely Colombia, Finland, Germany, Italy, Portugal, Spain, Switzerland, the United Kingdom, and the United States. Of particular note, Medical Properties Trust (MPT) holds the distinction of ranking as the second-largest non-governmental proprietor of hospitals globally, rendering it a compelling subject for comprehensive stock analysis.

In order to gain valuable insights and make forecasts, this study attempts to explore the area of time series analysis applied to MPT's stock. We aim to build strong models that can forecast the stock prices of MPT and spot possible trends by utilizing historical stock data, market indicators, and pertinent macroeconomic factors.

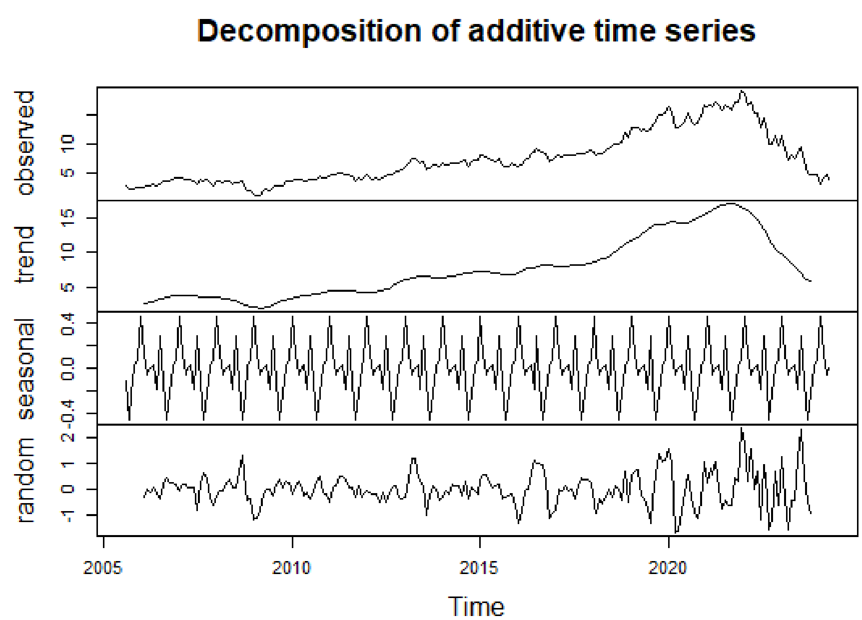
2 Exploratory Data Analysis

In order to understand what we should do with the data, it is a good idea to start with plotting a series.

Time series plot of mpw:



As we can see, the series has a trend. It could have some seasonality but in order to see it, we should decompose the data and plot it. So, let's do it.



From the decomposed plot, we can definitely the upwards trend at first, and then downward trend after the start of 2023. After making some research, I found out that the stock fell due to the inflation rates, since all the real estate companies are severely dependable on it. Let's keep this in mind.

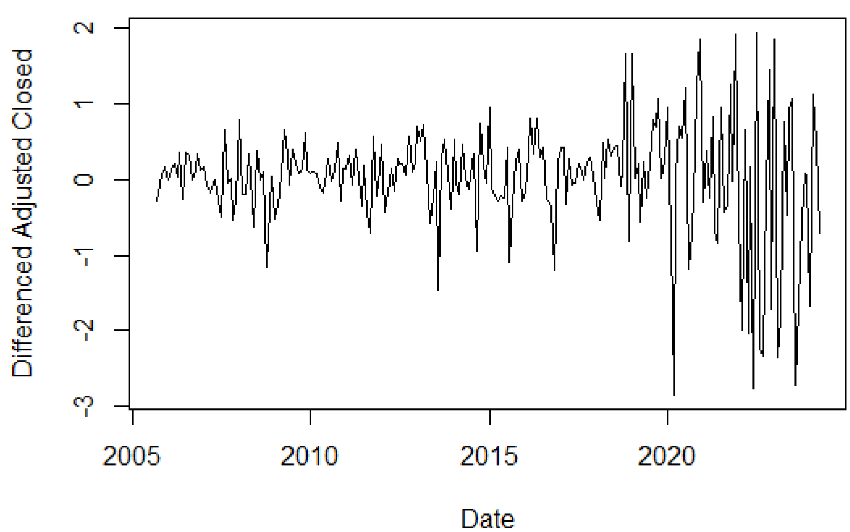
Another conclusion that we can make from the plot is that the data has seasonal pattern, which is important to note now, because it will play a big role while choosing our model later.

Random plot shows us that there are not that many irregular fluctations, but definitely we could have some problems with heteroscedasticity, which ruins our assumption of white noise residuals.

3 Stationarity

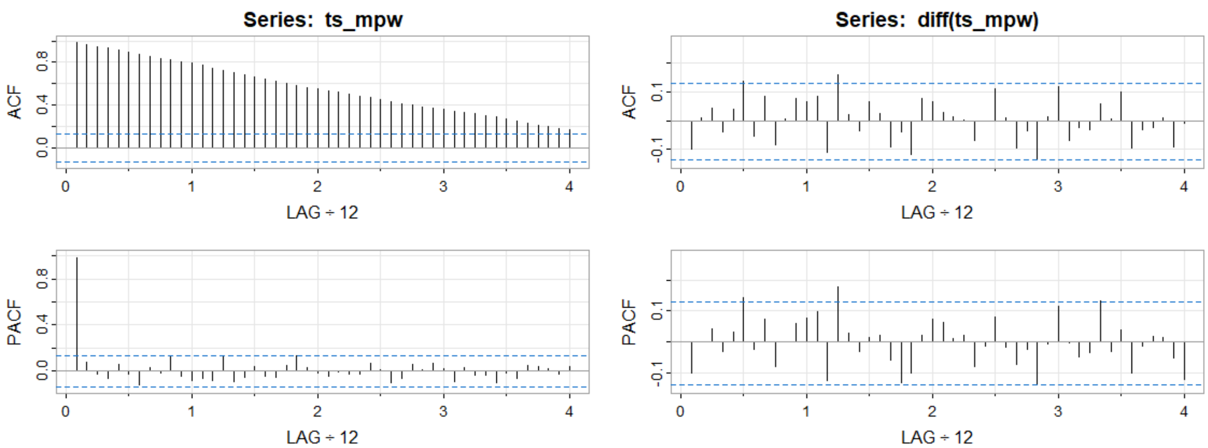
In order to successfully continue working with the data, we need to transform the data to reach one of our main assumptions - stationarity. The main method to do it is to differenciate the data. This will also help us get rid of the trend part.

Differenced time series plot of mpw:



From the differenced plot we see that we got rid of the trend and reached stationarity, but there are definitely some issues with constant variance assumption, which is important to note, since we are violating the main time series assumption. Now we can make a conclusion that we need to work with GARCH model along with the SARIMA to create a reasonable model.

Let's look at the ACF and PACF plots of the original time series and differenced one to make a conclusion that we reached stationarity.



We make the conclusion that we reached the stationarity, but let's use Augmented Dickey-Fuller Test to make a final decision. The null hypothesis H_0 : The series is not stationary, while the alternative hypothesis H_1 : The series is stationary.

```
        Augmented Dickey-Fuller Test

data:  ts_mpw
Dickey-Fuller = -0.9046, Lag order = 6, p-value = 0.9511
alternative hypothesis: stationary
```

```
        Augmented Dickey-Fuller Test

data:  diff(ts_mpw)
Dickey-Fuller = -5.0238, Lag order = 6, p-value = 0.01
alternative hypothesis: stationary
```

We see that in the original time series, we cannot reject the null hypothesis, since the p-value is 0.9511, while in the differenced time series the p-values is less than 0.01, so we reject the null hypothesis and conclude that we reached stationarity.

4 Model selection

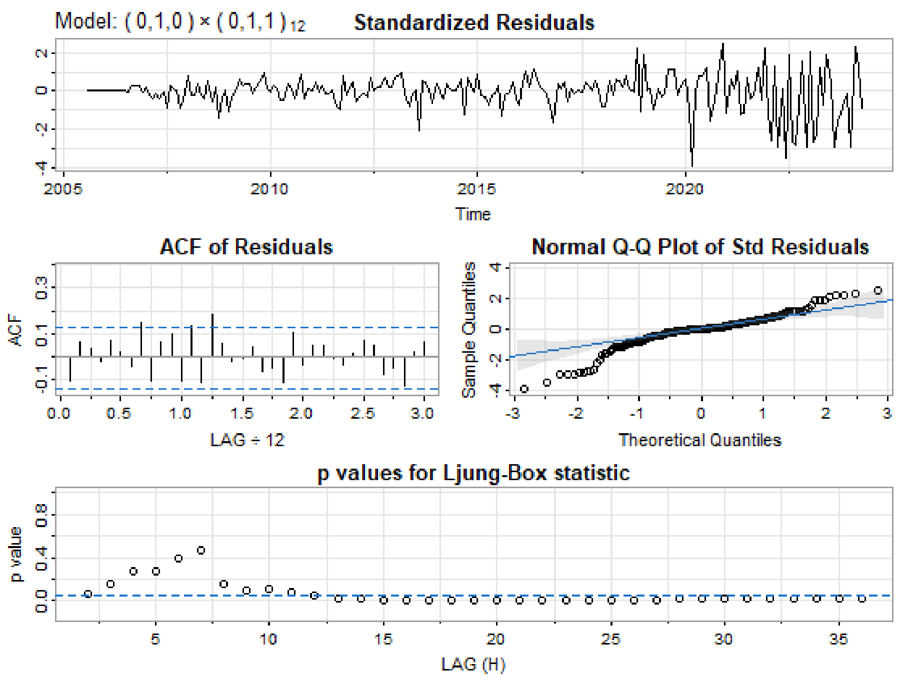
The best way for me to start choosing SARIMA model is to run *auto.arima*() function since it gives the idea where we should start. The result of such action were not really promising, since the seasonal part that we could see in the decomposed plot is not selected.

```
Series: ts_mpw
ARIMA(0,1,0)

sigma^2 = 0.5648:  log likelihood = -253.85
AIC=509.7   AICc=509.72   BIC=513.11
```

We see that the model provided is just differenced once model. Let's note that AIC of the model is 509.7 for future reference. In order to find any other model I decided to use implemented in the class function *get.best.arima*(), which gave me the following result:

```
[[1]]
[1] 492.6933

[[2]]

Call:
arima(x = x.ts, order = c(p, d, q), seasonal = list(order = c(P, D, Q), frequency(x.ts)),
    method = "CSS")

Coefficients:
         sma1
      -0.8862
s.e.   0.0404

sigma^2 estimated as 0.5804:  part log likelihood = -243.14

[[3]]
[1] 0 1 0 0 1 1
```

The AIC of the model is provided in [[1]] section, which is 492.6933 and it is better than the previous model, but it also includes seasonal part, that we can see in [[3]] part of the output. So, we will continue with the model: nonseasonal part: no AR and MA, but differenced once; seasonal part: no AR, MA of order 1 and differenced once.



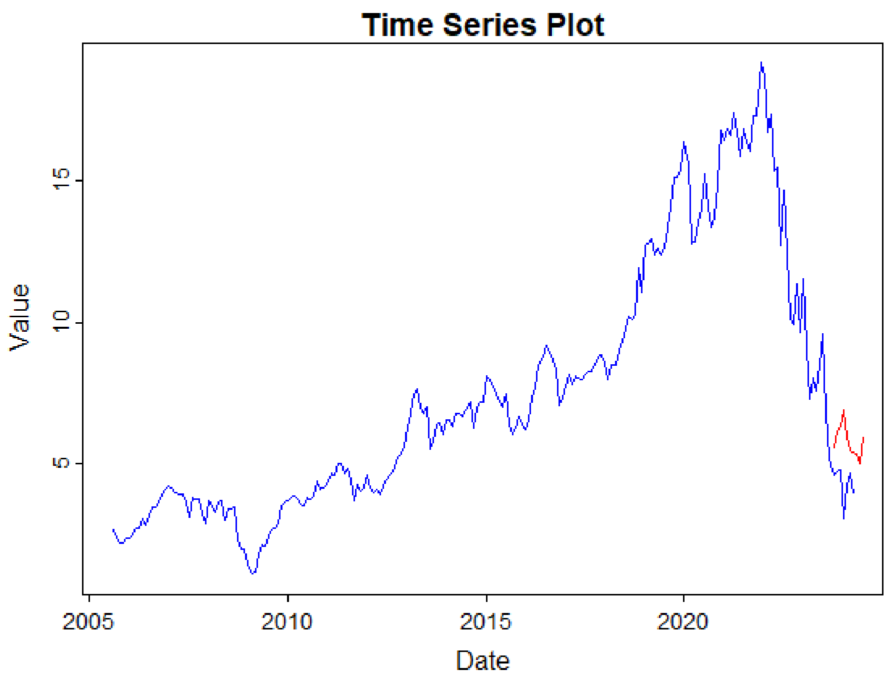
From the plots we still can see the issues with the volatility of the data and in order to work with it, I will also add GARCH model to he data so the overall model will also include explanation of the heteroscedasticity.

```
#garch
garch_spec <- ugarchspec(variance.model = list(model = "sGARCH", garchOrder = c(1, 1)),
                         mean.model = list(armaOrder = c(0, 0), include.mean = FALSE),
                         distribution.model = "std")
garch_fit <- ugarchfit(data = sarima_residuals, spec = garch_spec)
```

Now we can combine SARIMA and GARCH together to create a good prediction model.

5 Validation of the model

In order to see whether we are going in the correct direction, I am going to create a train time series and plot it together with the actual data, the results are given below.

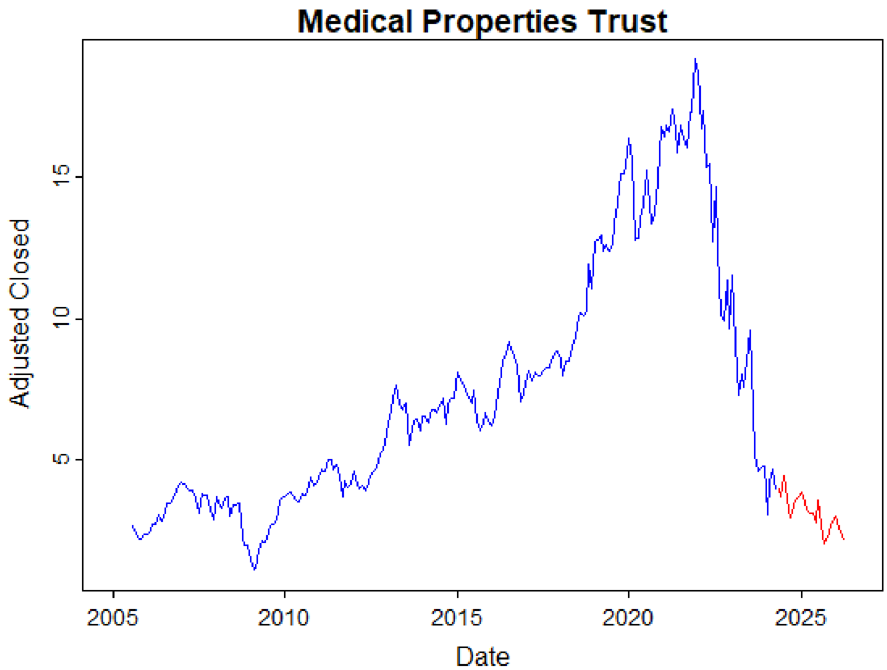


The red line shows the prediction that my model made while the blue line is the original data. We see that the prediction has the same shape as the real data, but it does not go that low as the real data. That can be explained by the fact of the inflation rates that we have already discussed before. In order to create an ideal model, we should also add the column of the rates into the data and work with VAR model. But for now, we will continue working with the model that we created now.

6 Predictions

```
#Forecast arima and garch

garch_forecast <- ugarchforecast(garch_fit, n.ahead = 24)
sarima_forecast <- sarima.for(ts_mpw, 24, 0,1,0,0,1,1,12)

combined_forecast <- garch_forecast@forecast$seriesFor + sarima_forecast$pred
print(combined_forecast)
```

We are trying to predict the data 2 years ahead, which we can see in the code in the number 24, which means 24 months. The results of the prediction is provided below.



We see that the model overall makes a pretty reasonable predictions. So overall for now I would not recommend investing into this stock. But we can't make those conclusions, since we have a chance to create a better model if we include inflation rates data.

7 Conclusion

During this project I learned how to use different techniques on real life stock data as well as understand what factors can influence the outcome of the predictions. I learned how to make the data stationary and choose the best model to work with. Combination of 2 different models gave me much better results than working with each of those separately. In the future work, I am planning to include the inflation rates column so that I can also explain the abrupt ups and downs of the data. Based on the predictions provided by my model it is not the best idea to invest money right now. But who knows how the inflation will play role on the stock.