Task1 a)Discuss why the US unemployment rate, the US inflation rate, the federal funds rate and the growth rate of the S&P 500 could affect Bitcoin growth.

Bitcoin is a digital currency that has seen a extreme rise in popularity over the last few years. Its value has been the subject of much speculation, and many people are interested in understanding the factors that can influence its growth.

One of the key factors that can impact the growth of Bitcoin is the economic situation in the United States. Specifically, the US unemployment rate, inflation rate, federal funds rate, and growth rate of the S&P 500 can all have an impact on Bitcoin's value.

When the unemployment rate is high, there is typically less economic activity, which can result in reduced consumer spending and a decreased demand for Bitcoin. On the other hand, when the unemployment rate is low, there is often more economic activity, leading to increased consumer spending and a greater demand for Bitcoin.

Similarly, high inflation rates can lead to decreased purchasing power for consumers and businesses, resulting in a lower demand for Bitcoin. However, Bitcoin has been considered a hedge against inflation, and its value may increase during periods of high inflation.

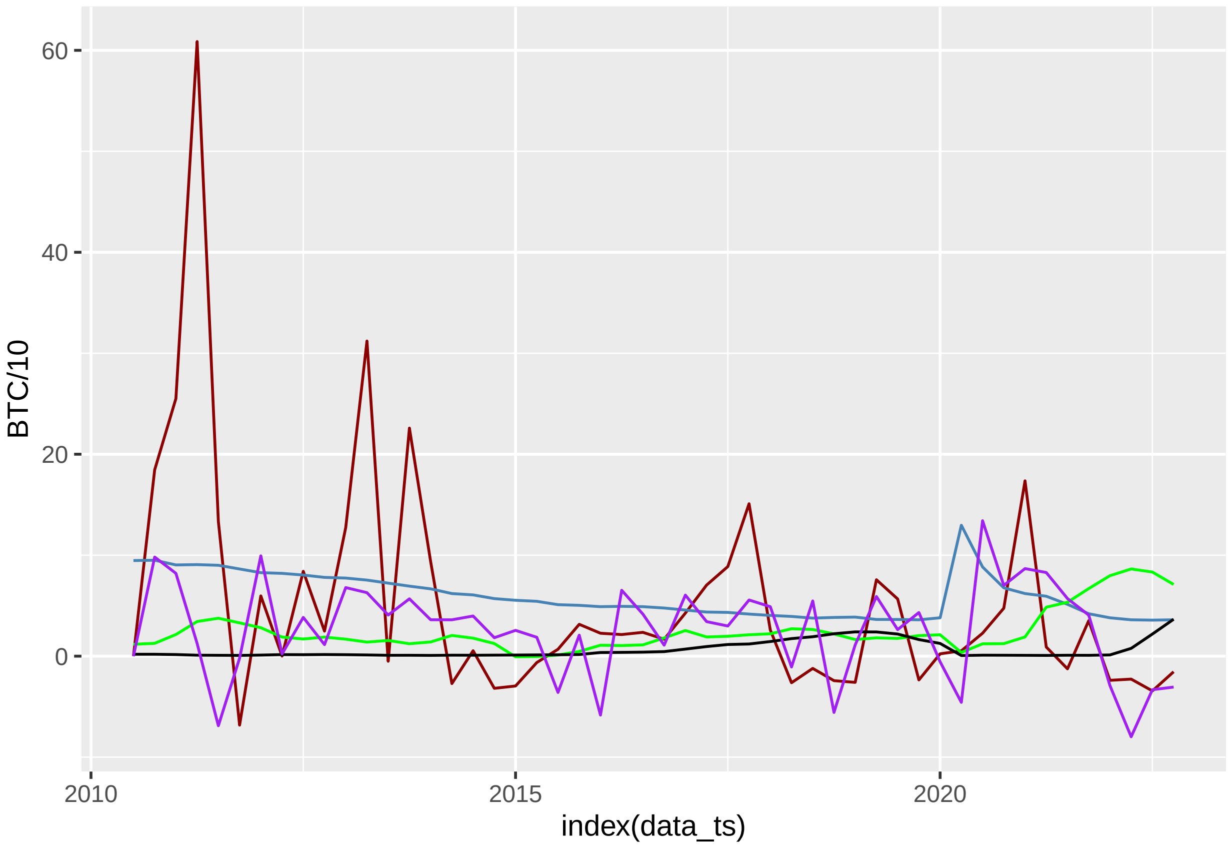
Changes in the federal funds rate can also impact the value of the US dollar, which can in turn impact the value of Bitcoin. A decrease in the federal funds rate can weaken the US dollar, leading to an increase in the price of Bitcoin. Conversely, an increase in the federal funds rate can strengthen the US dollar and lead to a decrease in the price of Bitcoin.

Finally, the growth rate of the S&P 500, which is a benchmark for the performance of the US stock market, can also impact Bitcoin's value. During times of economic stability, investors may be less likely to invest in alternative assets like Bitcoin. However, during times of economic uncertainty or market volatility, investors may turn to Bitcoin as a safe haven asset, leading to an increase in its value.

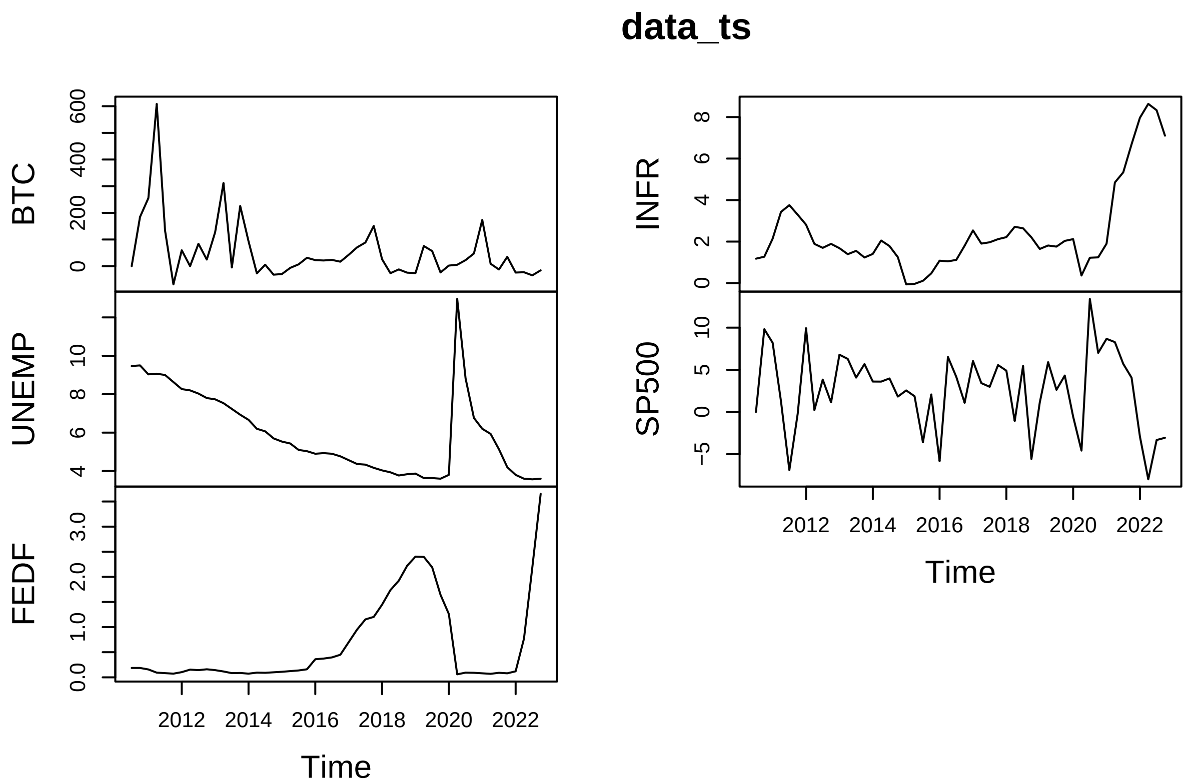
Overall, it's important to note that Bitcoin is a relatively new and volatile asset, and its value can be influenced by a range of unpredictable factors beyond just these economic indicators. However, understanding how the economic situation in the US can impact Bitcoin's growth is an important part of understanding this digital currency and its potential for the future

b)Generate a plot that shows Bitcoin's growth rate and the four potential drivers (at quarterly frequency) over the longest time period available.

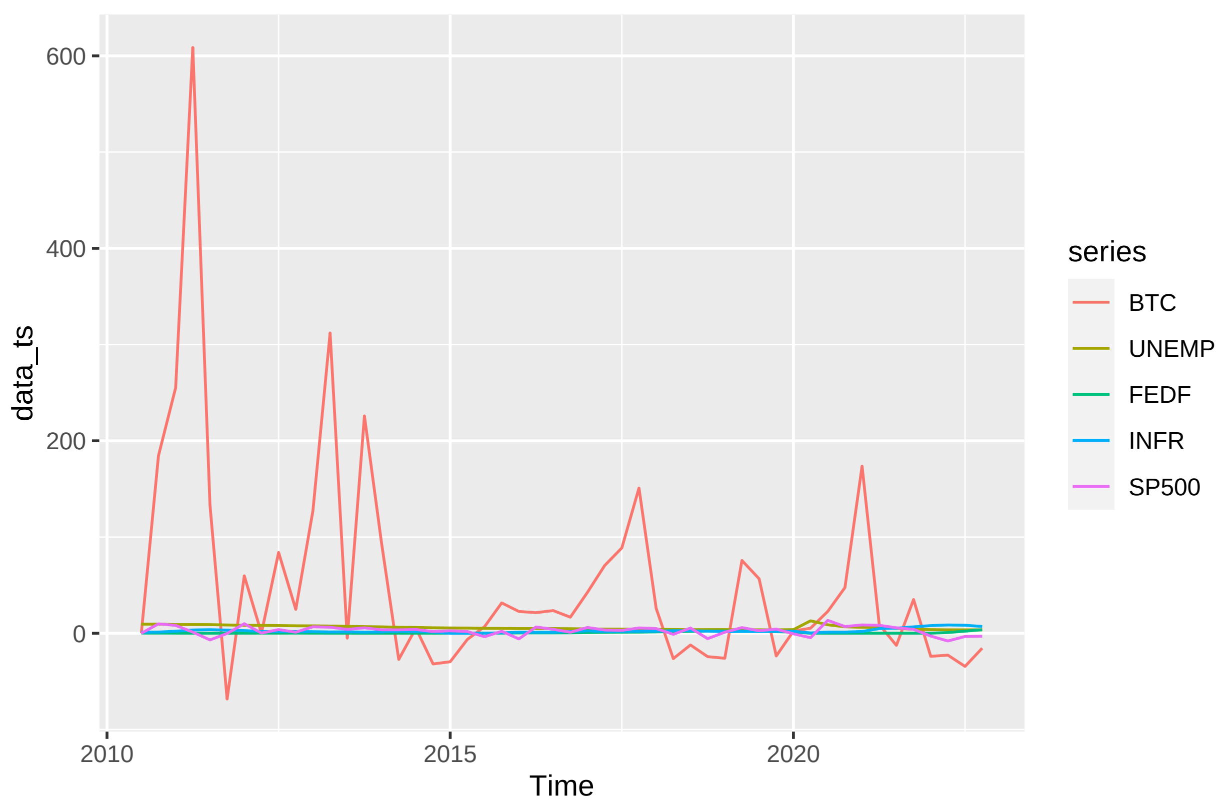
```
data <- read_csv("~/Desktop/CaseStudies/dataset_QD.csv")
data <- data_frame(data$BTC_gr,data$UNEMP,data$FEDF,data$INFR,data$SP500_gr)
colnames(data) <- c("BTC","UNEMP","FEDF","INFR","SP500")
data_ts <- ts(data,start=c(2010,3),frequency = 4)
ggplot(data, aes(x=index(data_ts))) +
  geom_line(aes(y = BTC/10), color = "darkred") +
  geom_line(aes(y = UNEMP), color="steelblue") +
  geom_line(aes(y = INFR), color="green")+
  geom_line(aes(y = FEDF), color="black") +
  geom_line(aes(y = SP500), color="purple")
```

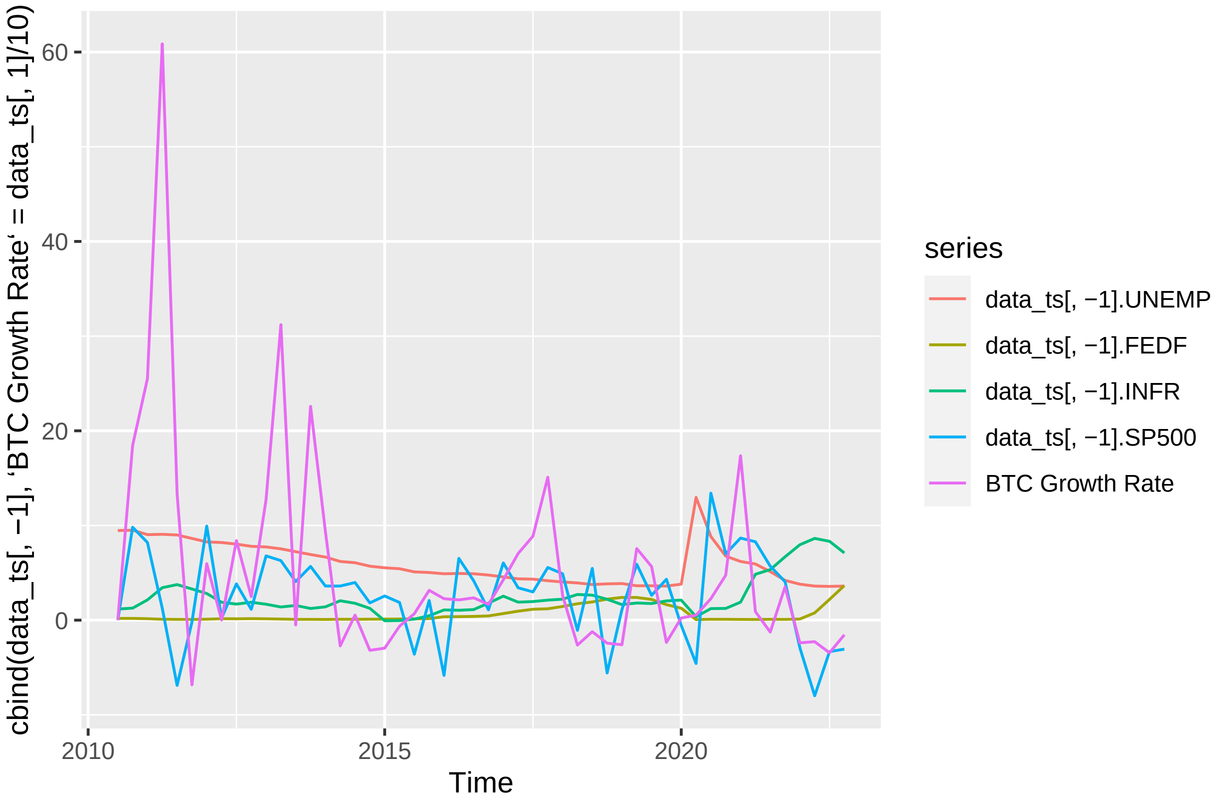



```
#or
plot(data_ts)
```



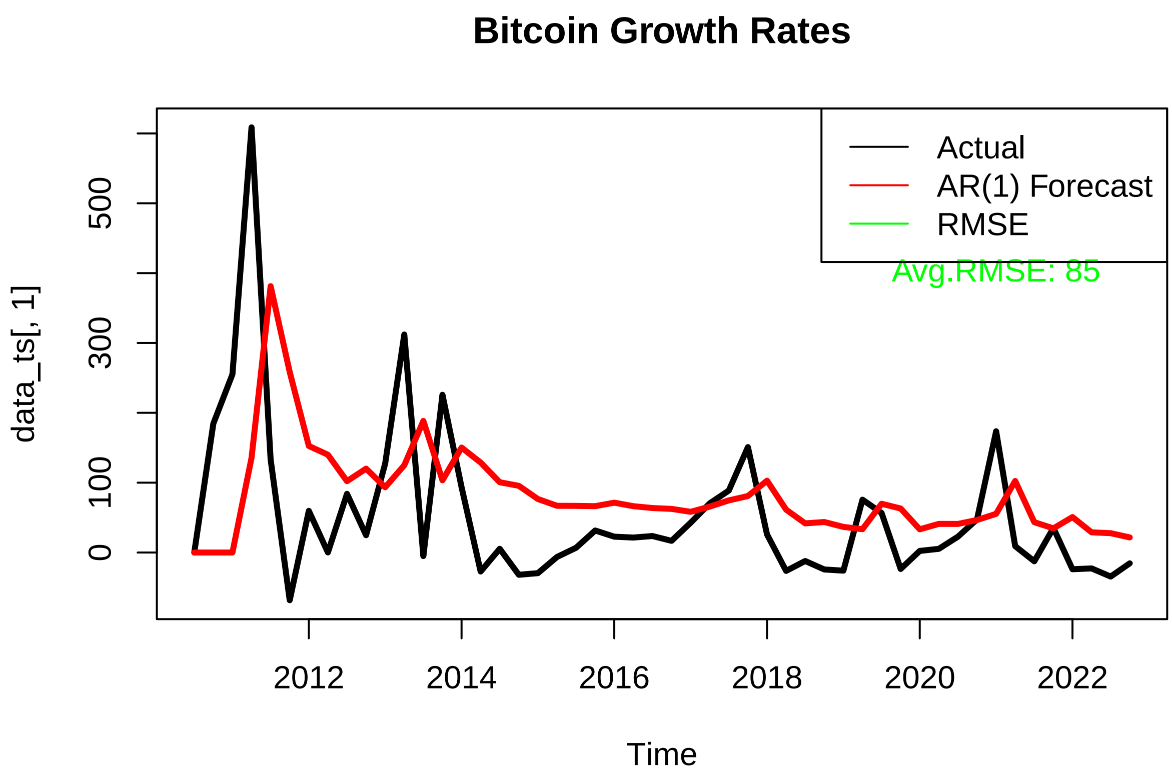

```
#or

autoplot(data_ts)
```



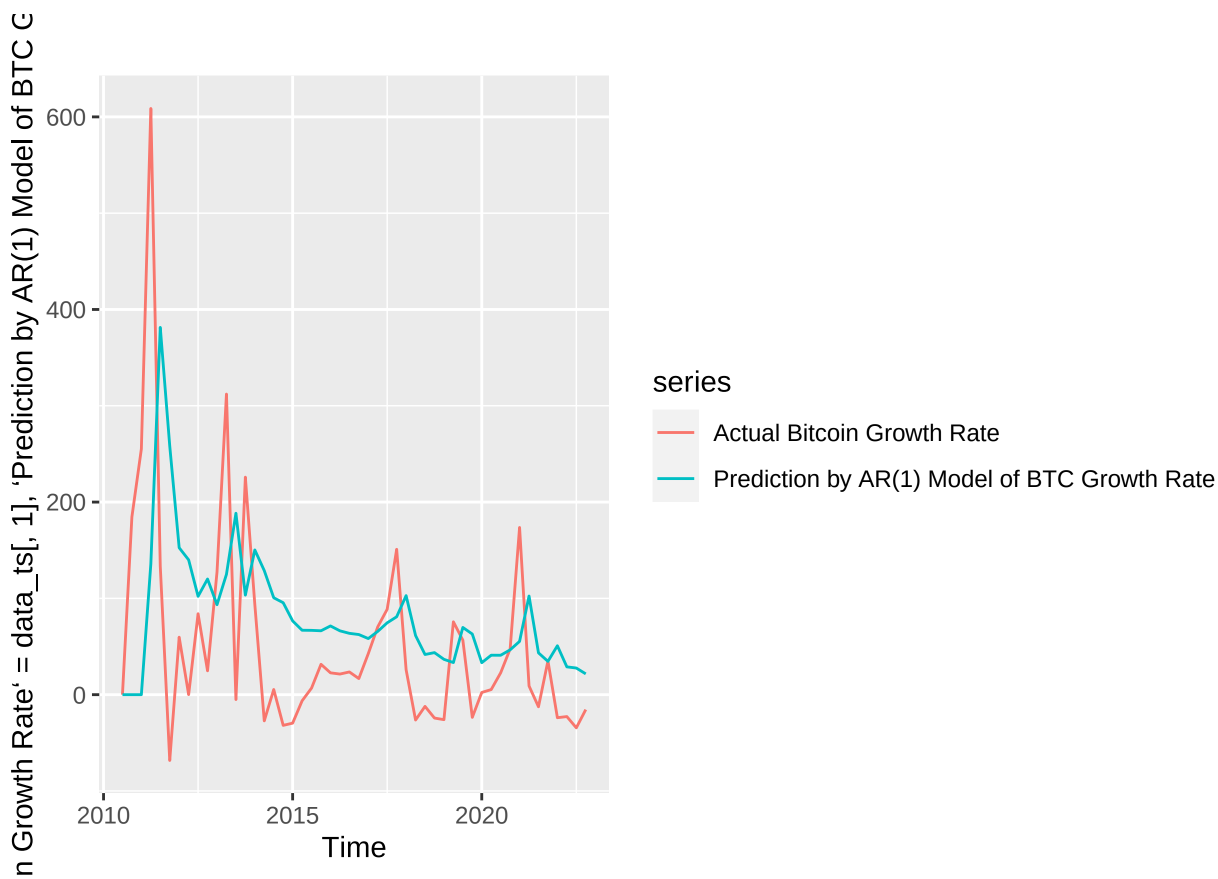


c)-Create one-quarter-ahead forecasts of Bitcoin growth for this time period using an autoregressive model of order one (i. e., an AR(1) model). Illustrate the forecasts together with the actual growth rates in one plot. Calculate the root mean squared forecasting error over the whole period.2 Here and in the following you can set all required starting values equal to zero.

```
ar1_fcast <- c(0,0)

rmse <- c(0,0)
for (i in 2:49) {
  ar1 <- arima(data_ts[,1][1:i], order = c(1,0,0))
  ar1_fcast[i+1] <- forecast(ar1, h = 1)$mean[1]
  rmse[i+1] <- sqrt(mean((ar1_fcast[i+1] - data_ts[,1][i+1])^2))
}
for_ts <- ts(ar1_fcast,start=c(2010,3),frequency = 4)
plot(data_ts[,1], main = "Bitcoin Growth Rates",lwd=3)
lines(for_ts, col = "red",lwd=3)
#lines(ts(rmse,start=c(2010,3),frequency = 4), col = "green")
text(x = c(2021,2), y = 400, col="green",labels = paste0("Avg.RMSE: ", round(mean(rmse), 1)))
legend("topright", c("Actual", "AR(1) Forecast", "RMSE"), col=c("black", "red", "green"), lty=1)
```

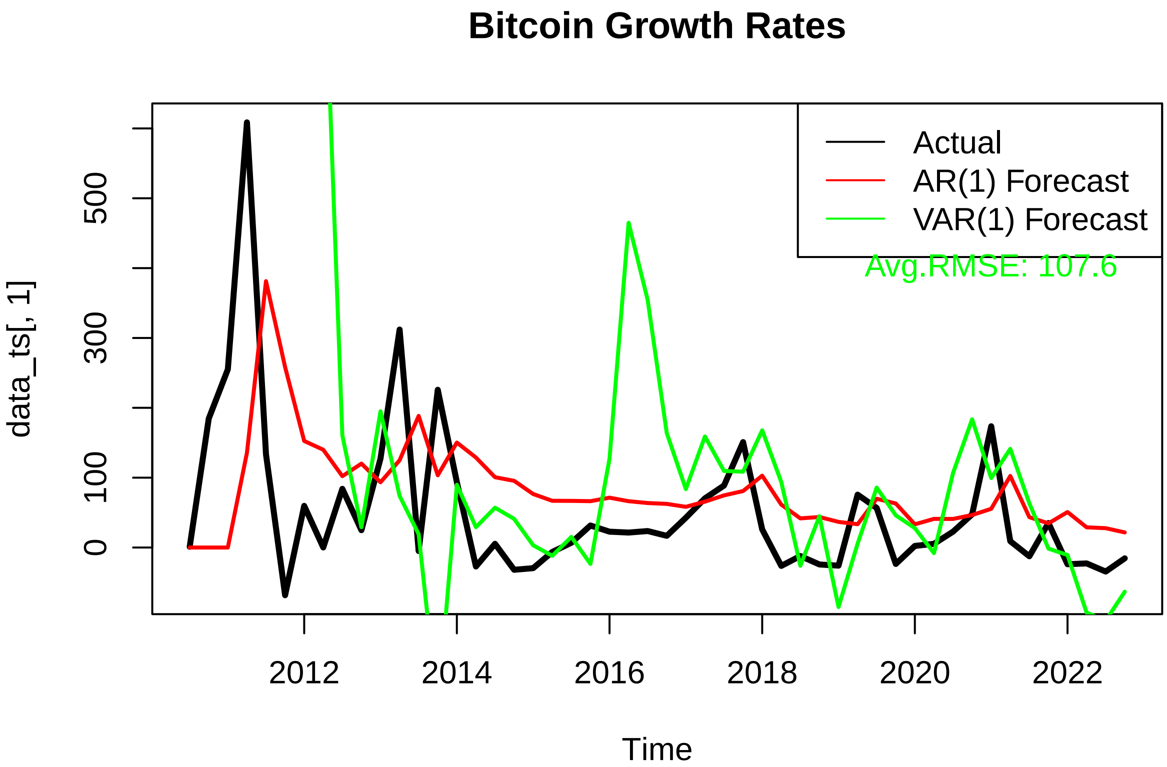



```
autoplot(cbind("Actual Bitcoin Growth Rate"=data_ts[,1],"Prediction by AR(1) Model of BTC Growth Rate" =
```

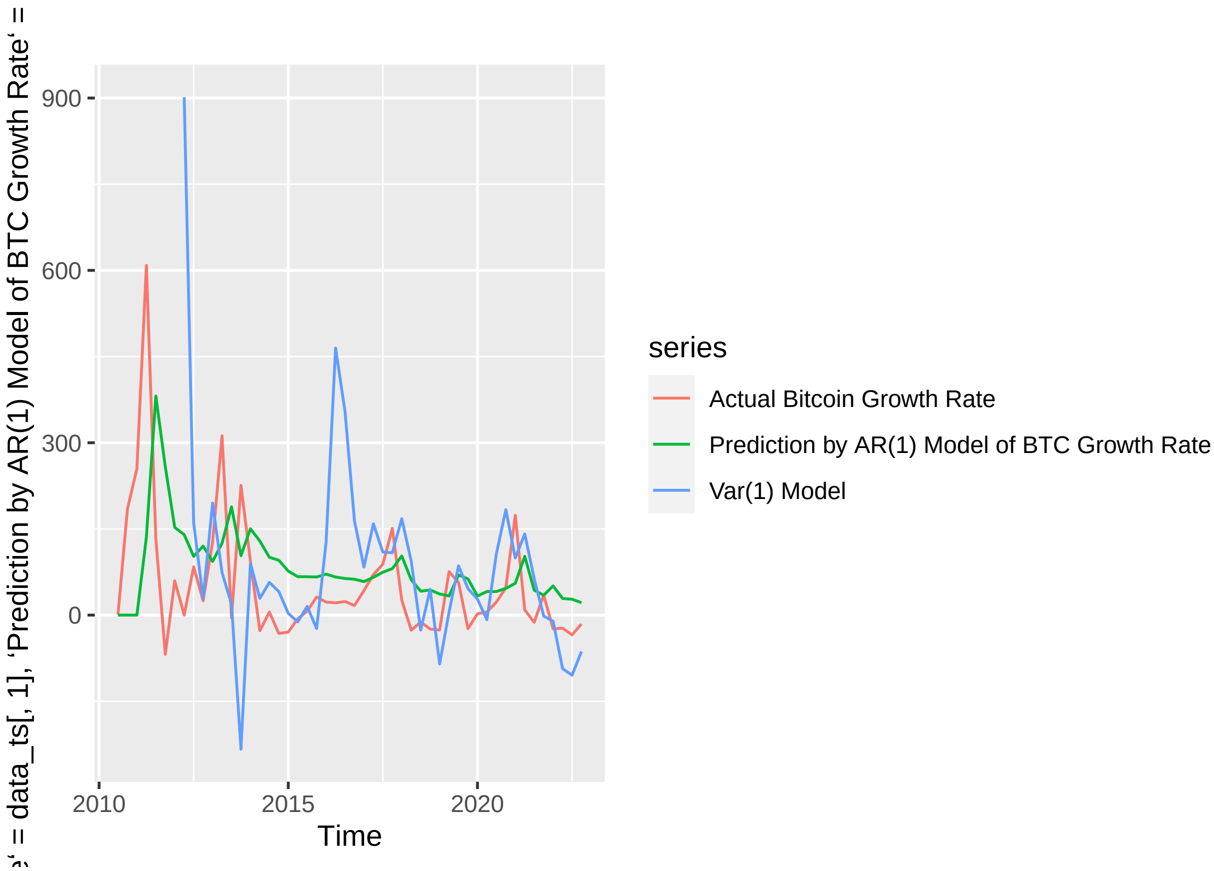


d)Create one-quarter-ahead forecasts of Bitcoin growth for the same period using a vector autoregressive model of order one (i. e., a VAR(1) model) containing the four additional variables. Illustrate the forecasts together with the actual growth rates and the forecasts based on the AR(1) model in one plot, potentially extending the plot in (c). Calculate the root mean squared forecasting error over the whole period.

```
var1_fcast <- matrix(NA, nrow = nrow(data_ts), ncol = 1)
rmse_var1 <- rep(NA,times=7)
data_ts[1,1] <- 1
data_ts[1,2] <- 1
# Loop over all observations, starting from the second observation
for (i in 7:49) {

  # Subset the data up to the previous observation
  data_sub <- data_ts[1:i,]

  # Estimate a VAR(1) model using the previous data
  var1_fit <- VAR(data_sub, p = 1,season=NULL,type="const",exogen = NULL)

  # Forecast the next observation using the VAR(1) model
  var1_fcast[i+1] <- predict(var1_fit, n.ahead = 1,ci=0.95)$fcst$BTC[1]
  #na.fill(var1_fcast[i+1],0)
  rmse_var1[i+1] <- sqrt(mean((var1_fcast[i+1] - data_ts[,1][i+1])^2))
}

#var1_fcast[4:7]<- c(0,0,0,0)
#rmse_var1[2:7]<- c(0,0,0,0,0,0)
plot(data_ts[,1], main = "Bitcoin Growth Rates",lwd=3)
lines(ts(ar1_fcast,start=c(2010,3),frequency = 4), col = "red",lwd=2)
lines(ts(var1_fcast,start=c(2010,3),frequency = 4), col = "green",lwd=2)
legend("topright", c("Actual", "AR(1) Forecast", "VAR(1) Forecast"), col=c("black", "red", "green"), lty
text(x = c(2021,2), y = 400, col="green",labels = paste0("Avg.RMSE: ", round(mean(rmse_var1[8:50]), 1)))
```




```
var1_ts <- ts(var1_fcast,start=c(2010,3),frequency = 4)

autoplot(cbind("Actual Bitcoin Growth Rate"=data_ts[,1],"Prediction by AR(1) Model of BTC Growth Rate" =
```



e)

Granger causality is a statistical concept that measures the causal relationship between two time series variables. The concept was introduced by Clive Granger in 1969 and is widely used in econometrics to assess the direction of causality between economic variables. Granger causality actually means that, on average over the sample period, including lags of one variable helps reduce the squared error in predicting another variable. It does not imply causality, only precedence.

In a VAR(1) model, which is a Vector Autoregressive model of order one, the current value of each variable is modeled as a linear function of the past values of all the variables in the system. Granger causality in this context is tested by examining whether the inclusion of lagged values of one variable improves the forecasting ability of another variable.

Kilian and Lütkepohl (2017) discuss the concept of Granger causality in Chapters 2.5 and 7. In Chapter 2.5, they explain that a variable X Granger causes another variable Y if the past values of X contain information that helps to predict the current value of Y, beyond what can be predicted using only the past values of Y. In other words, X Granger causes Y if the inclusion of past values of X in the forecasting model of Y improves the accuracy of the forecasts.

In Chapter 7, Kilian and Lütkepohl extend the Granger causality concept to a multivariate setting and show how to test for Granger causality between multiple variables. In this context, Granger causality between two variables X and Y is said to exist if the past values of X help to predict the current value of Y, beyond what can be predicted using the past values of Y and all other variables in the system.

For example, we could test whether the past values of the unemployment rate help to predict the current value of Bitcoin growth, beyond what can be predicted using only the past values of Bitcoin growth. If we find evidence of Granger causality between unemployment rate and Bitcoin growth, we would conclude that unemployment rate Granger causes Bitcoin growth. We could repeat this analysis for each of the other variables of interest to identify which variables Granger cause Bitcoin growth.

```r
var_model_UNEMP <- VAR(data_ts[,-c(3,4,5)], p=1, type = "const")
var_model_FEDF <- VAR(data_ts[,-c(2,4,5)], p=1, type = "const")
var_model_INFR <- VAR(data_ts[,-c(2,3,5)], p=1, type = "const")
var_model_SP500 <- VAR(data_ts[,-c(2,3,4)], p=1, type = "const")


granger_test_UNM<- causality(var_model_UNEMP, cause = "UNEMP")$Granger$method
granger_test_FEDF <- causality(var_model_FEDF, cause = "FEDF")$Granger$method
granger_test_INFR <- causality(var_model_INFR, cause = "INFR")$Granger$method
granger_test_SP500 <- causality(var_model_SP500, cause = "SP500")$Granger$method




print(granger_test_UNM)
```

```
## [1] "Granger causality H0: UNEMP do not Granger-cause BTC"
```

```r
print(granger_test_FEDF)
```

```
## [1] "Granger causality H0: FEDF do not Granger-cause BTC"
```

```r
print(granger_test_INFR)
```

```
## [1] "Granger causality H0: INFR do not Granger-cause BTC"
```

```r
print(granger_test_SP500)
```

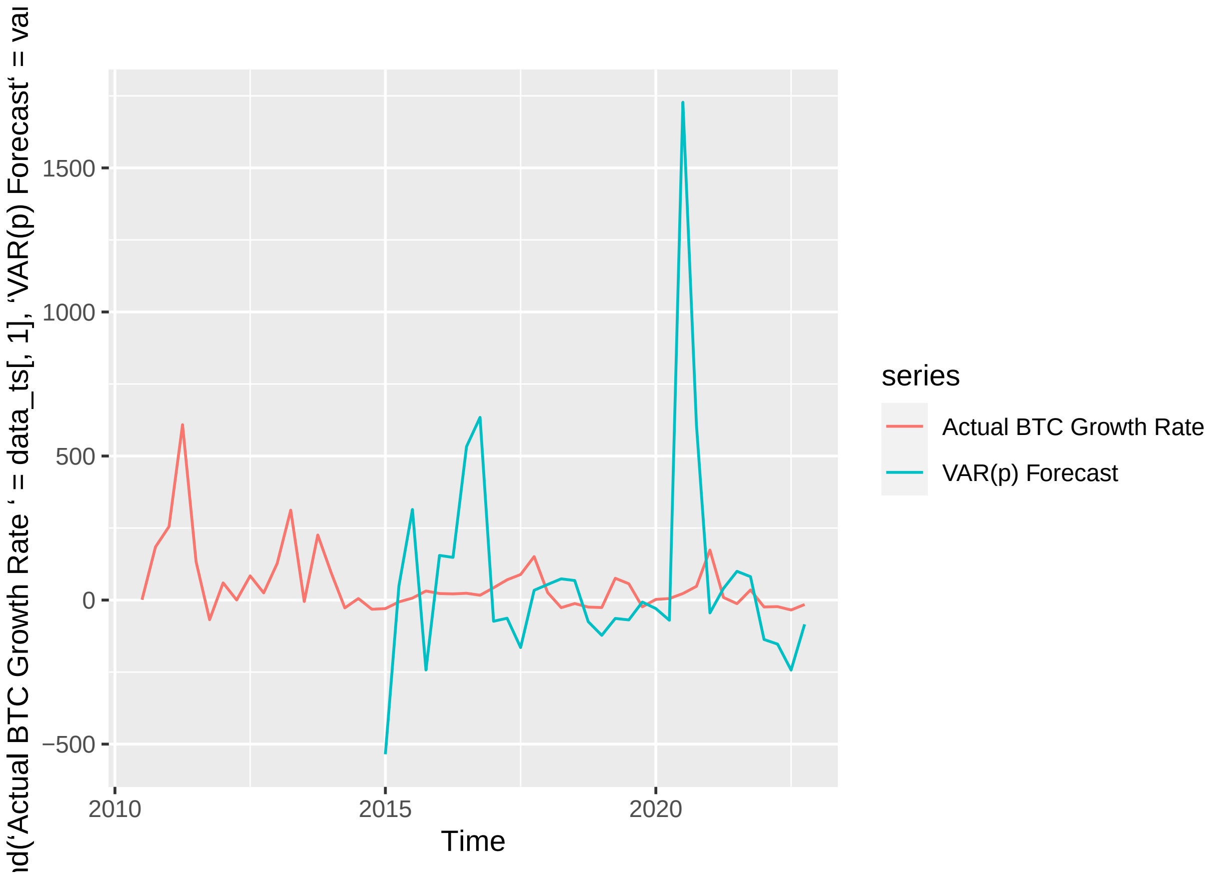
```
## [1] "Granger causality H0: SP500 do not Granger-cause BTC"
```

```r
#All together cause= BTC
#var_model <- VAR(data_ts, p=1, type = "const",season=NULL,exog=NULL)
#grng_btc <- causality(var_model, cause = "BTC")$Granger$method
#print(grng_btc)

#summary(var_model)


#Var Decomp
#plot(fevd(var_model,n.head=1))
```

f) Akaike's Information Criterion (AIC) is a statistical measure that provides a means of selecting the appropriate lag order for a Vector Autoregressive (VAR) model. The lower the AIC, the better the model.To determine the appropriate lag order for a VAR model, one can estimate VAR models with different lag orders and choose the one that has the lowest AIC value.

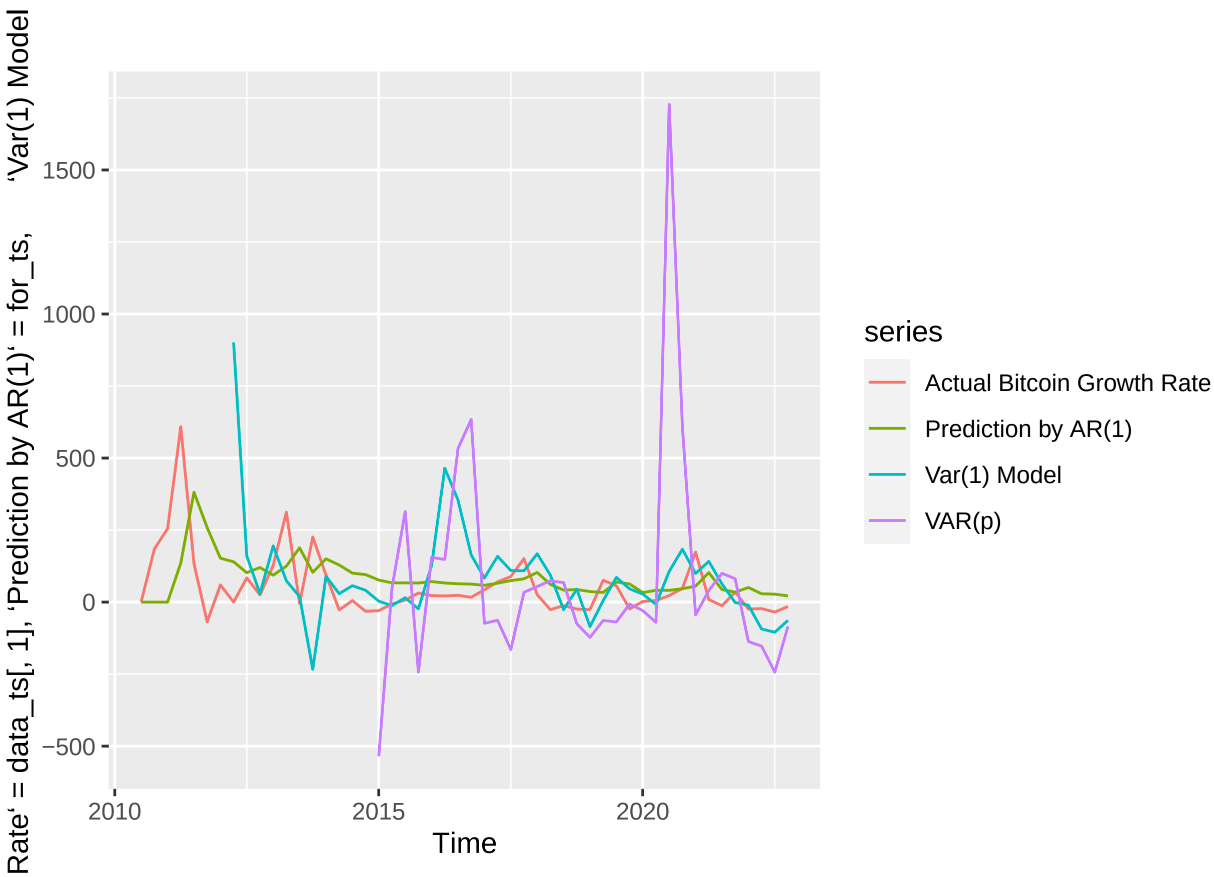
```r
fcst <- rep(NA,times=16)
#Optimal lag by AIC
var_select <- VARselect(data_ts,lag.max = 3,type = "const")$selection["AIC(n)"][1][[1]]
```

```r
#which is 3
rmse_varp <- rep(NA,times=16)

for (i in 16:49) {

var_model <- VAR(data_ts[1:i,],p=var_select,type="none")
fcst[i+1] <- predict(var_model, n.ahead = 1,ci=0.95)$fcst[[1]][,1]
rmse_varp[i+1] <- sqrt(mean(fcst[i+1] - data_ts[,1][i+1])^2)
}
varp_ts <- ts(fcst,start=c(2010,3),frequency = 4)

autoplot(cbind("Actual BTC Growth Rate "=data_ts[,1],"VAR(p) Forecast"=varp_ts))
```



OR

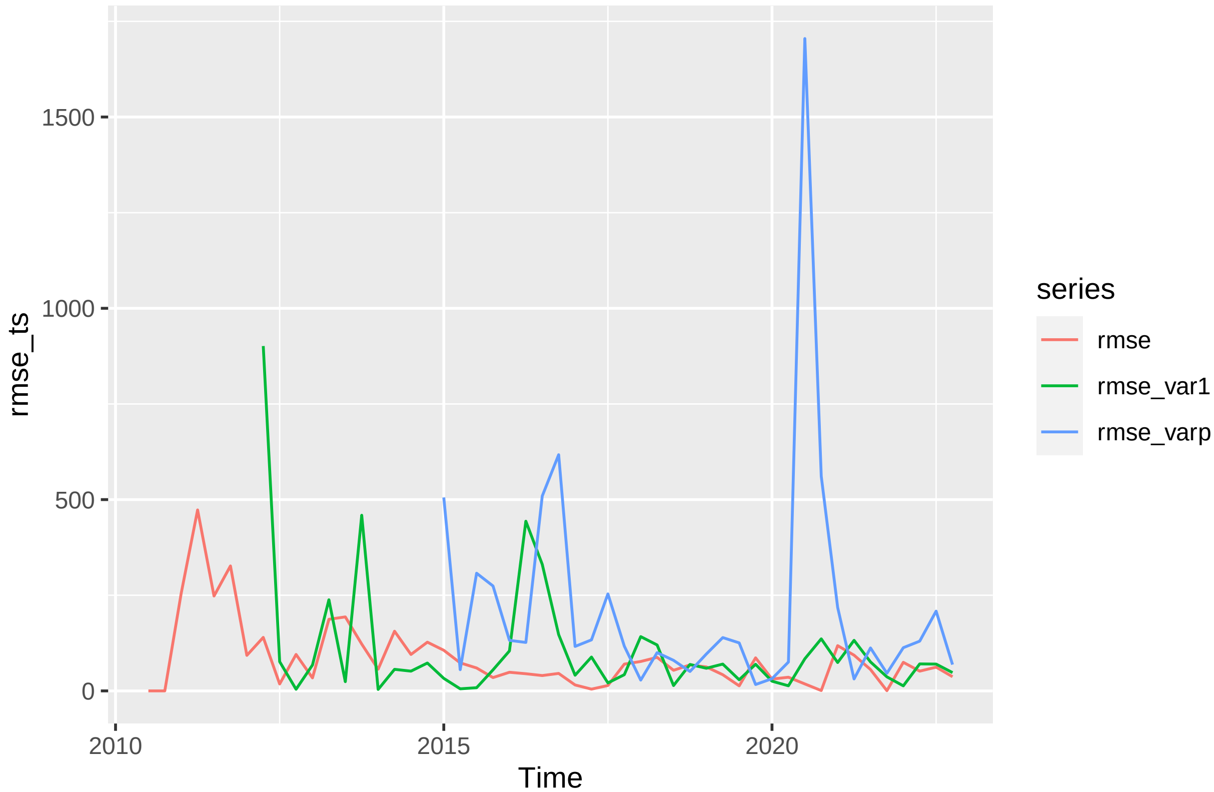
```r
varp_ts <- ts(fcst,start=c(2010,3),frequency = 4)
autoplot(cbind("Actual Bitcoin Growth Rate"=data_ts[,1],"Prediction by AR(1)" =for_ts,"Var(1) Model" = v
```




```
rmse_df <- data_frame(rmse,rmse_var1,rmse_varp)
rmse_ts <- ts(rmse_df,start = c(2010,3),frequency = 4)

mean_of_rmse <- 0
mean_of_rmse[1] <- sum(rmse[3:50])/length(3:50)
mean_of_rmse[2] <-sum(rmse_var1[8:50])/length(8:50)
mean_of_rmse[3] <- sum(rmse_varp[19:50])/length(17:50)
mean_of_rmse
```

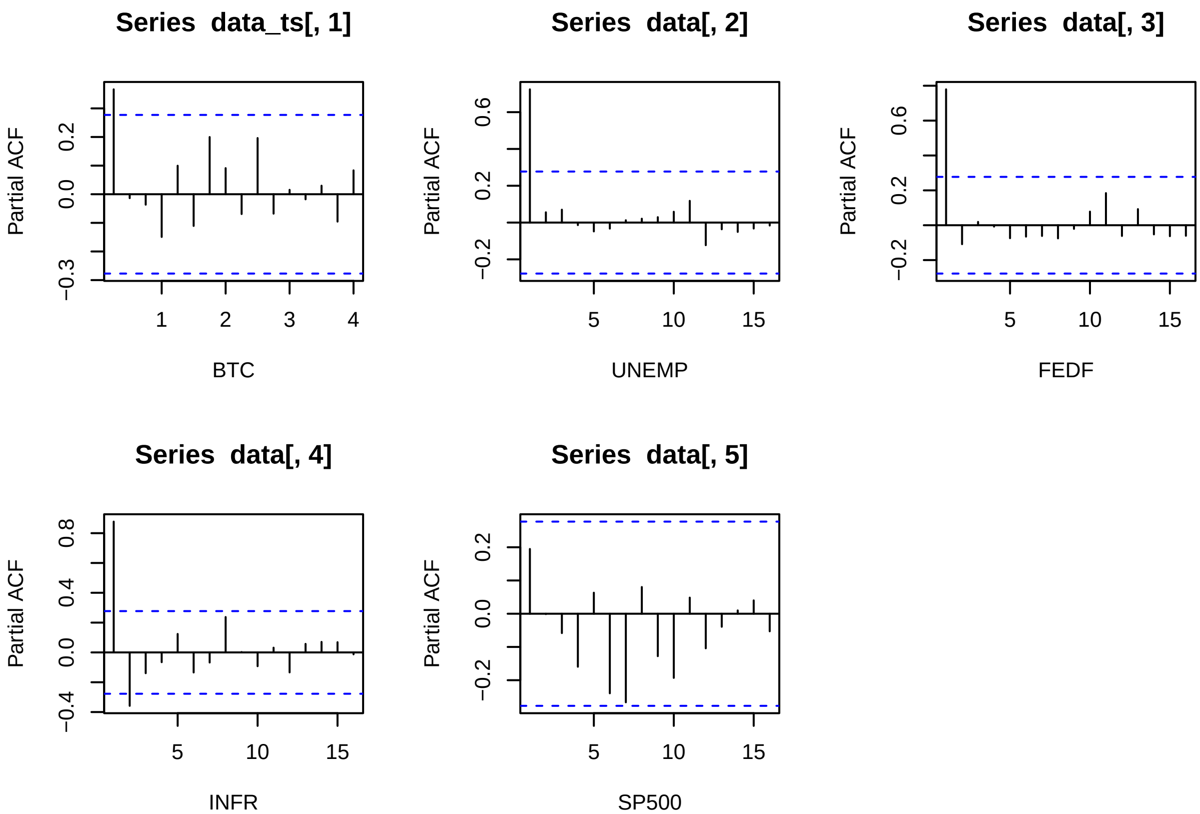
```
## [1]  88.54307 107.64069 208.44828
```

```
autoplot(rmse_ts)
```



when we check the pacg of the each variablesi n our dataset

```
par(mfrow=c(2,3))
pacf(data_ts[,1],xlab="BTC")
pacf(data[,2],xlab="UNEMP")
pacf(data[,3],xlab="FEDF")
pacf(data[,4],xlab="INFR")
pacf(data[,5],xlab="SP500")
```



f) Does it suffice to merely fit an AR(1) model? In which periods is it particularly advantageous to exploit information in other variables? Are there any periods in which the forecasts fit poorly to the actual observations? Are these periods potentially related to "economic events"? Is there any structure in the data not captured by the methods used? How could you account for these features?

When considering different models for predicting Bitcoin (BTC) growth rate, we find that the AR(1) model outperforms others. The reason for this is that the AR(1) model only uses the first lag of BTC and doesn't include additional variables. For short-term forecasts with a strong autocorrelation structure, fitting an AR(1) model may be sufficient. However, for quarterly data, the correlation structure is weak, resulting in poor predictions.

To improve the model, we may incorporate additional variables that share common trends or patterns with BTC or have known causal relationships with BTC. For example, including the macroeconomic variables such as inflation rate, unemployment rate, and federal funds rate affect BTC prices. However, we tested for Granger causality and did not find any causal effect from other variables to BTC.

Periods in which the forecasts fit poorly to the actual observations may indicate that the model is missing important features or that the data has changed in a way that was not captured by the model. These periods could potentially be related to economic events, such as the COVID-19 pandemic in 2019 to 2022. The pandemic had a significant impact on the global economy, leading to an increase in unemployment rates and a decline in stock prices, which in turn affected BTC prices where investors saw BTC as an alternative investment the during the uncertain market situation.

Incorporating such structural breaks and outliers in the model can be achieved through the use of robust estimation techniques or dummy variables that capture the effect of specific events. These techniques help account for the deviations from the normal pattern of the data.

In conclusion, while an AR(1) model may be sufficient for short-term forecasts with a strong autocorrelation structure, incorporating additional variables or using more complex models may be necessary for longer-term forecasts or when the correlation structure is weak. Furthermore, accounting for structural breaks and outliers through robust estimation techniques or dummy variables can help improve the accuracy of the model.

Task 2