Final Exam

Aniketh Moparthi

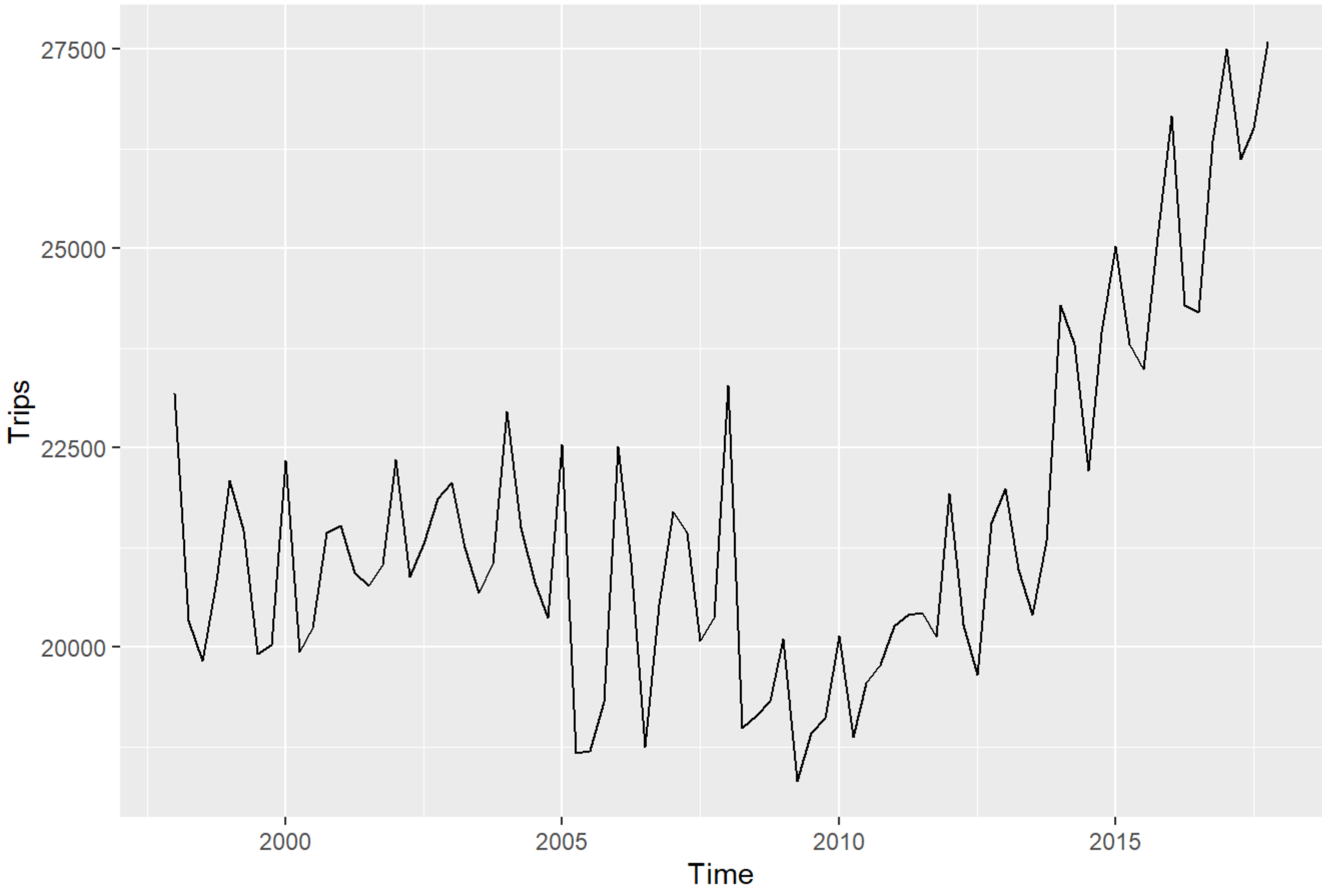
05-08-2023

1. Compute the total domestic overnight trips across Australia from the tourism dataset.

```
setwd("D:/Business Analytics/Business forecasting/Final Exam")

tourism <- read_excel("tourism.xlsx")

tourism <- tourism %>%
  mutate(Quarter = yearquarter(Quarter)) %>%
  as_tsibble(
    index = Quarter,
    key = c(Region, State, Purpose)
  )

tourism <- tourism %>%
  summarise(Trips = sum(Trips))

holiday <- tourism
```

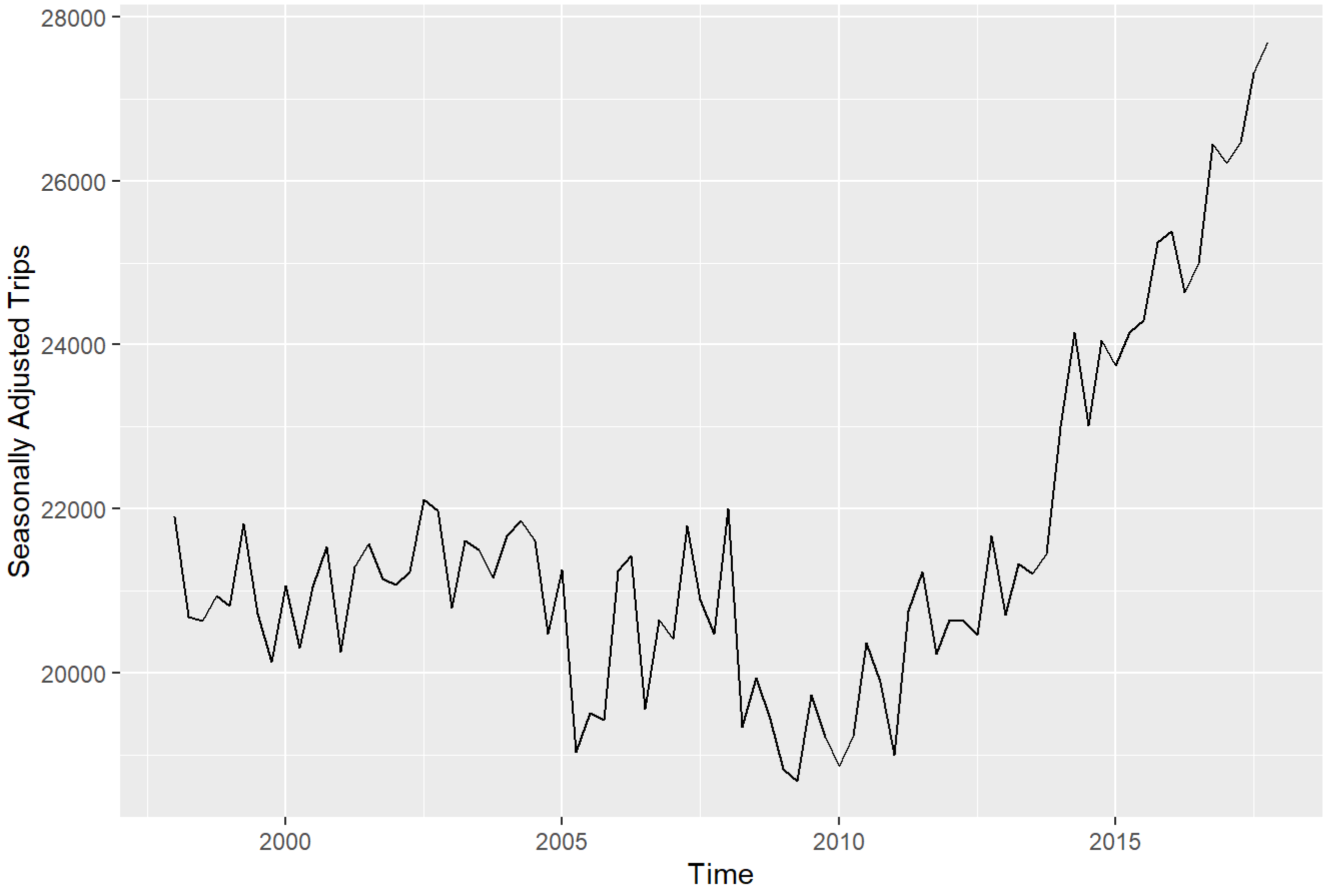
a. Plot the data and describe the main features of the series.

```
holiday_ts <- as.ts(holiday)
autoplot(holiday_ts, ts.geom = "line", ylab = "Trips")
```



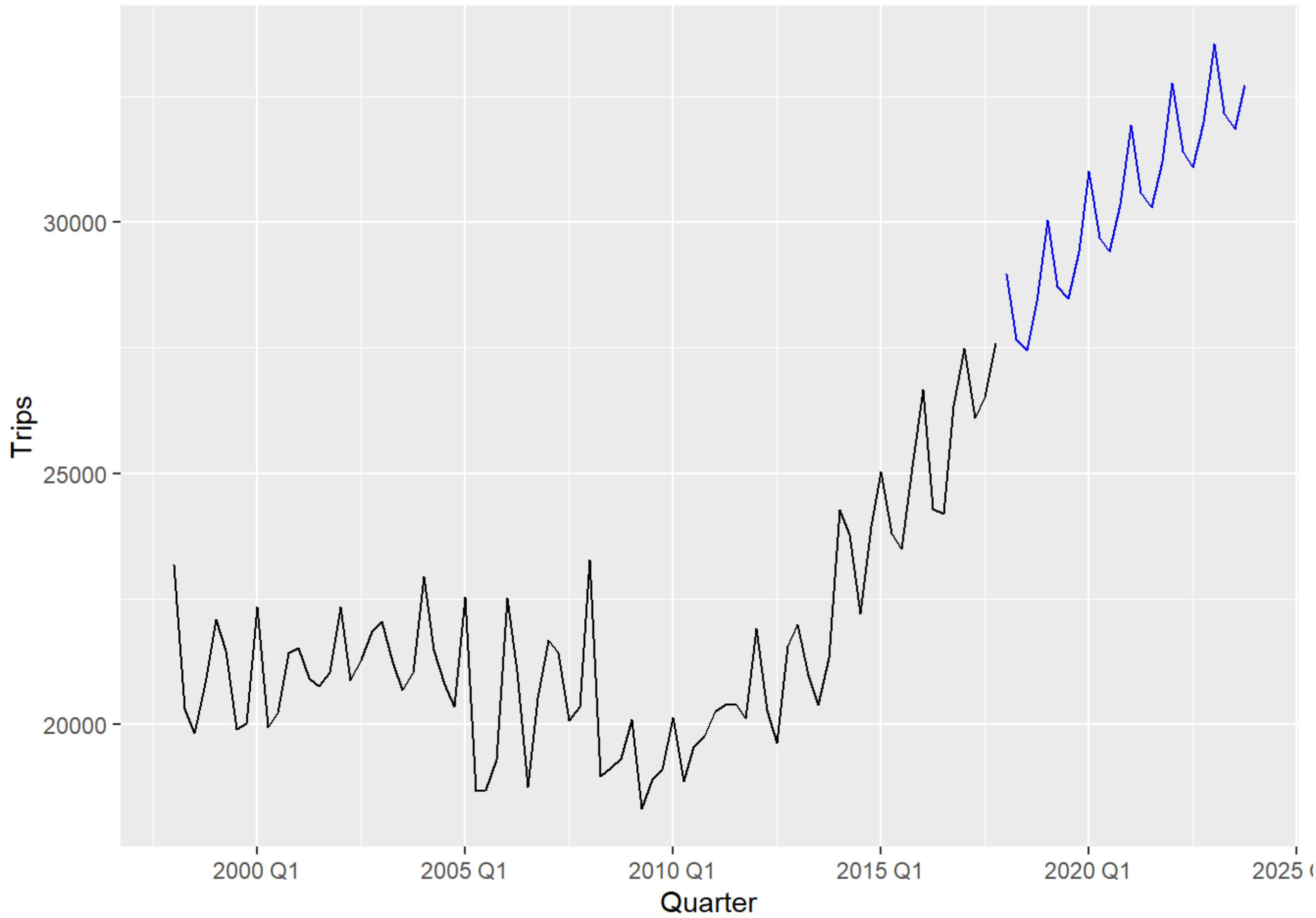
b. Decompose the series using STL and obtain the seasonally adjusted data

```
holiday_stl <- stl(holiday_ts, s.window = "periodic")
holiday_sa <- seasadj(holiday_stl)
autoplot(holiday_sa, ylab = "Seasonally Adjusted Trips")
```



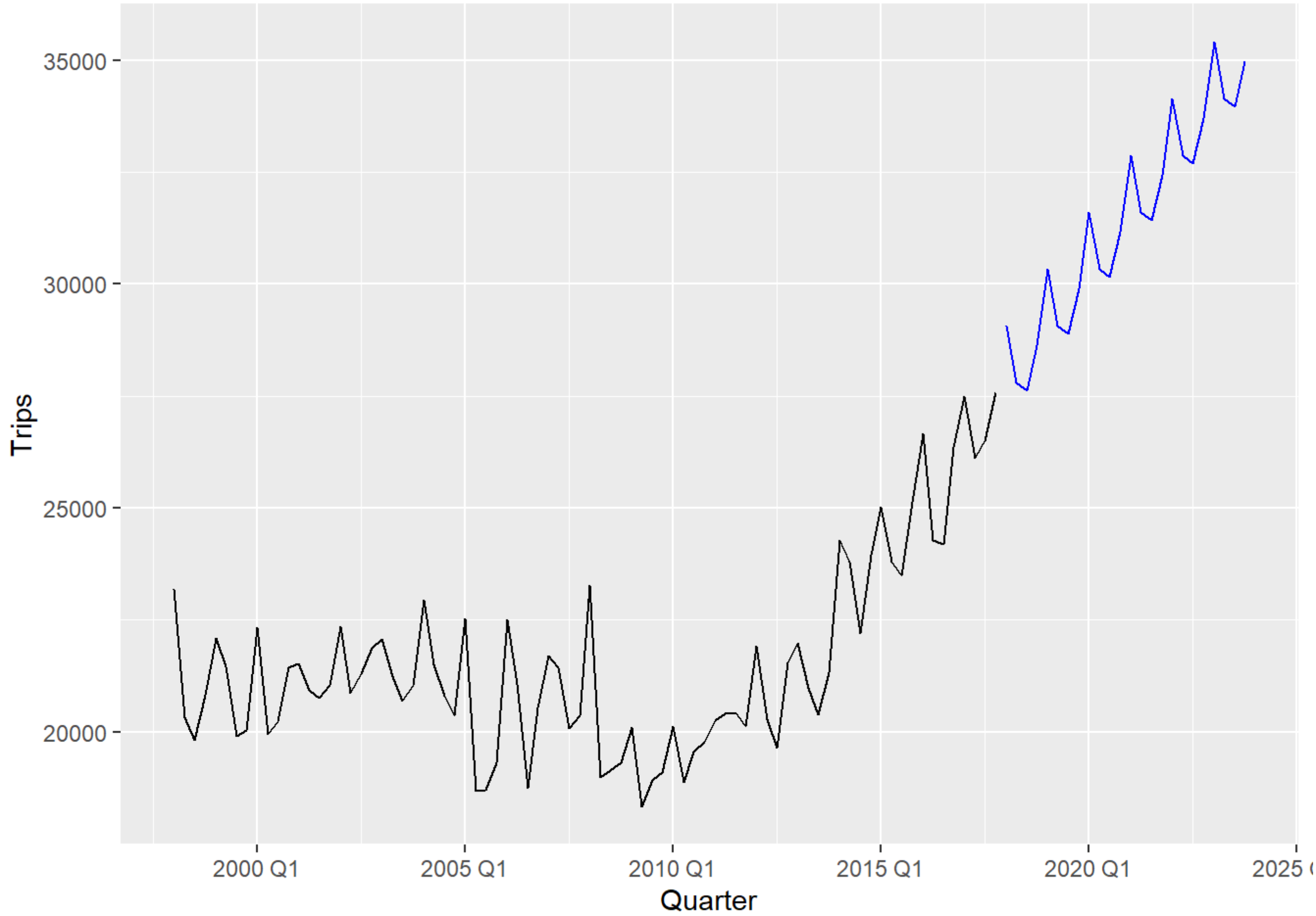
c. Forecast the next two years of the series using an additive damped trend method applied to the seasonally adjusted data. (This can be specified using decomposition_model().)

```
library(forecast)
fit <- holiday %>%
  model(
    "additive damped trend" = ETS(Trips ~ error("A") + trend("Ad") + season("A")))

fc <- fit %>% forecast(h = 24)
fc %>%
  autoplot(holiday, level = NULL)
```



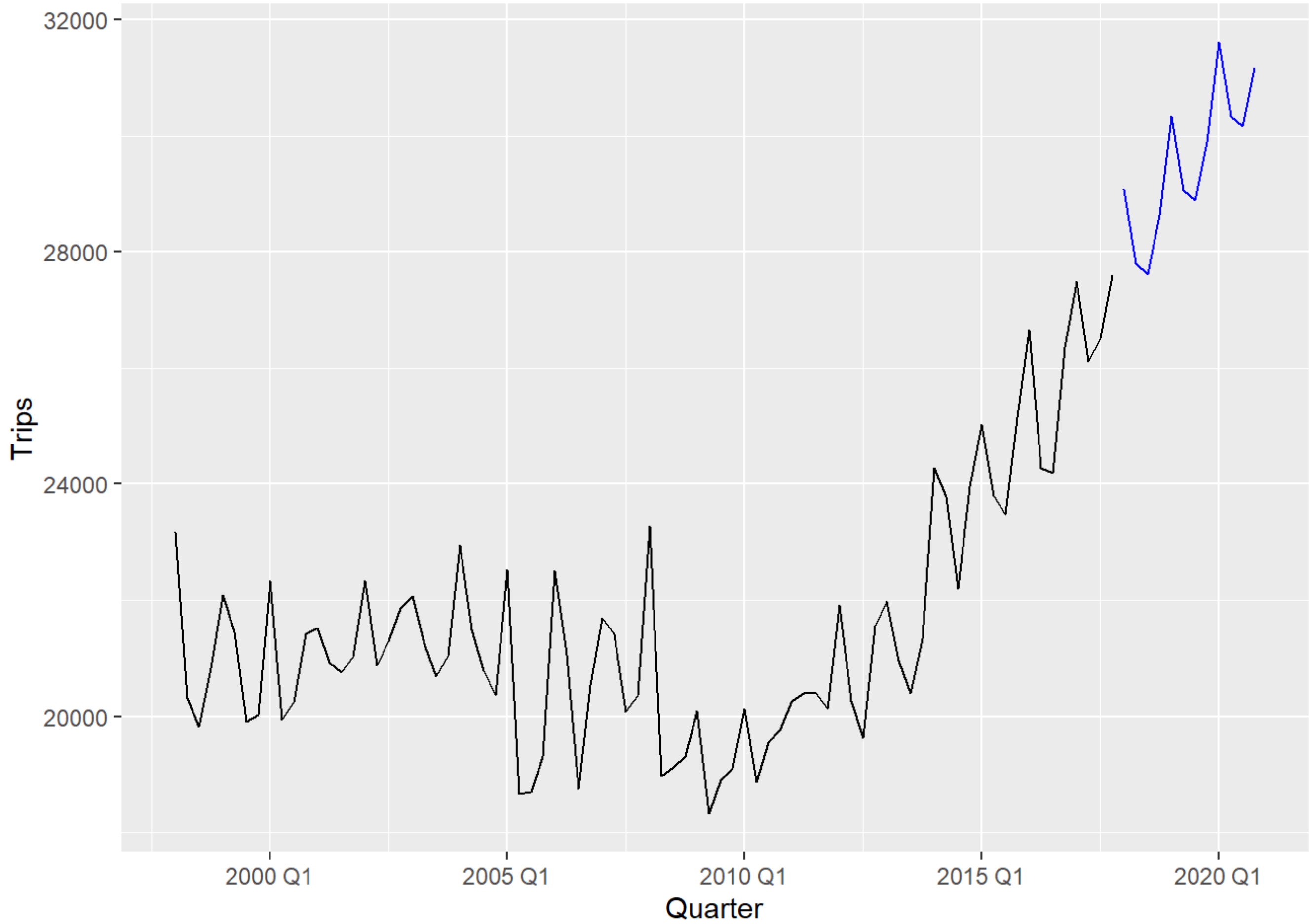
d. Forecast the next two years of the series using an appropriate model for Holt's linear method applied to the seasonally adjusted data (as before but without damped trend).

```
fit_holt <- holiday %>%
  model(
    "Holt's linear method" = ETS(Trips ~ error("A") + trend("A") + season("A"))
  )

fc_holt <- fit_holt %>% forecast(h = 24)
fc_holt %>%
  autoplot(holiday, level = NULL)
```



e. Now use ETS() to choose a seasonal model for the data.

```
fit_ets <- holiday %>%
  model(ETS(Trips))

fc_ets <- fit_ets %>% forecast(h = 12)

fc_ets %>% autoplot(holiday, level = NULL)
```



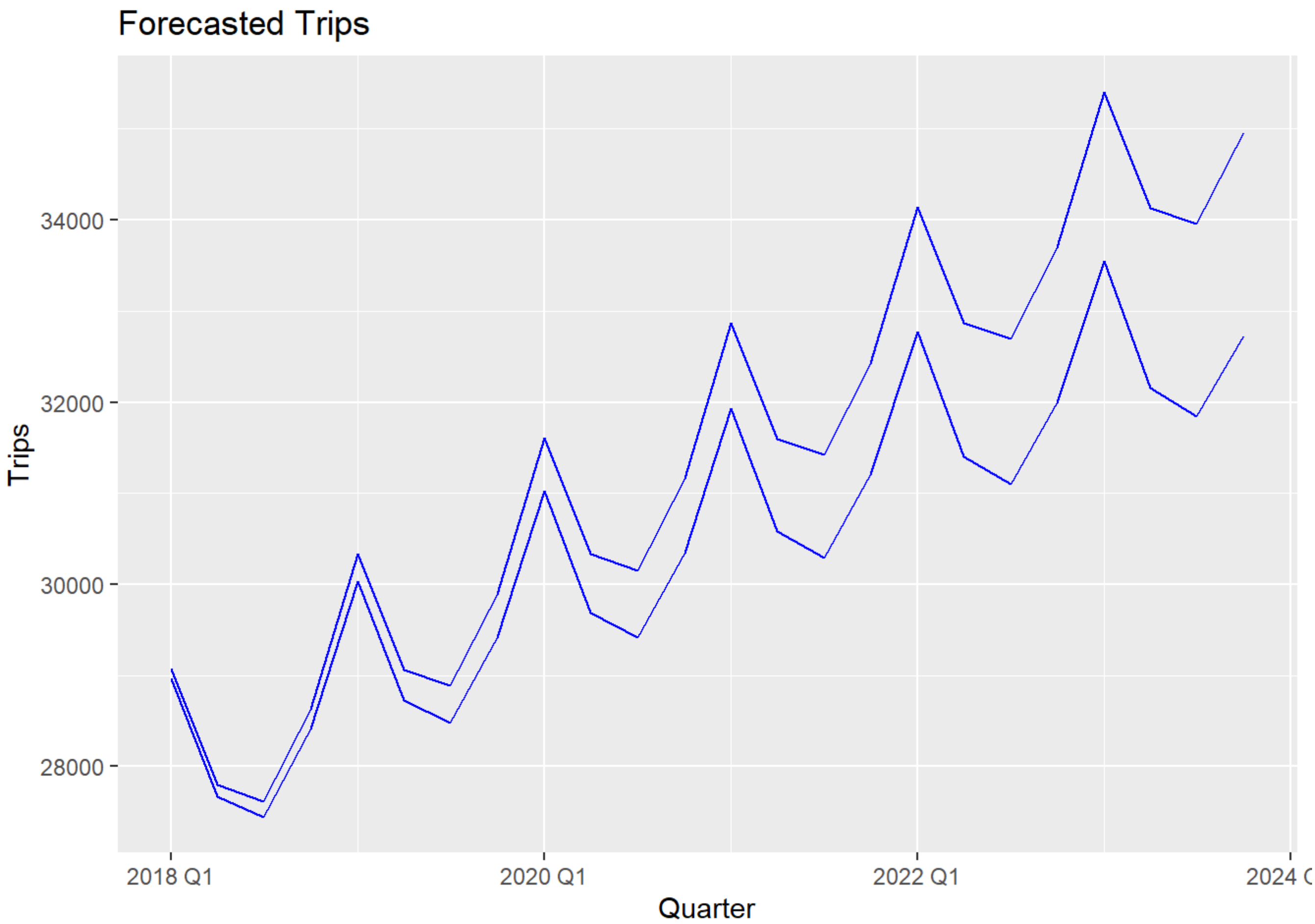
f. Compare the RMSE of the ETS model with the RMSE of the models you obtained using STL decompositions. Which gives the better in-sample fits?

```
hol_day <- holiday %>%
  model(
    "Holt Linear Method" = ETS(Trips ~ error("A") +  trend("Ad") + season("A")),
    "Holt Multiplicative" = ETS(Trips ~ error("A") +  trend("Ad") + season("N")),
    "ETS" = ETS(Trips))
accuracy(hol_day)
```

```
## # A tibble: 3 × 10
##   .model              .type      ME  RMSE   MAE   MPE  MAPE  MASE RMSSE     ACF1
##   <chr>               <chr>   <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>    <dbl>
## 1 Holt Linear Method  Traini…  103.  795.  606. 0.357  2.86 0.638 0.654  4.34e-4
## 2 Holt Multiplicative Traini…  135. 1215. 1000. 0.354  4.66 1.05  1.00   9.53e-2
## 3 ETS                 Traini…  105.  794.  604. 0.379  2.86 0.636 0.653 -1.51e-3
```

g. Compare the forecasts from the three approaches? Which seems most reasonable?

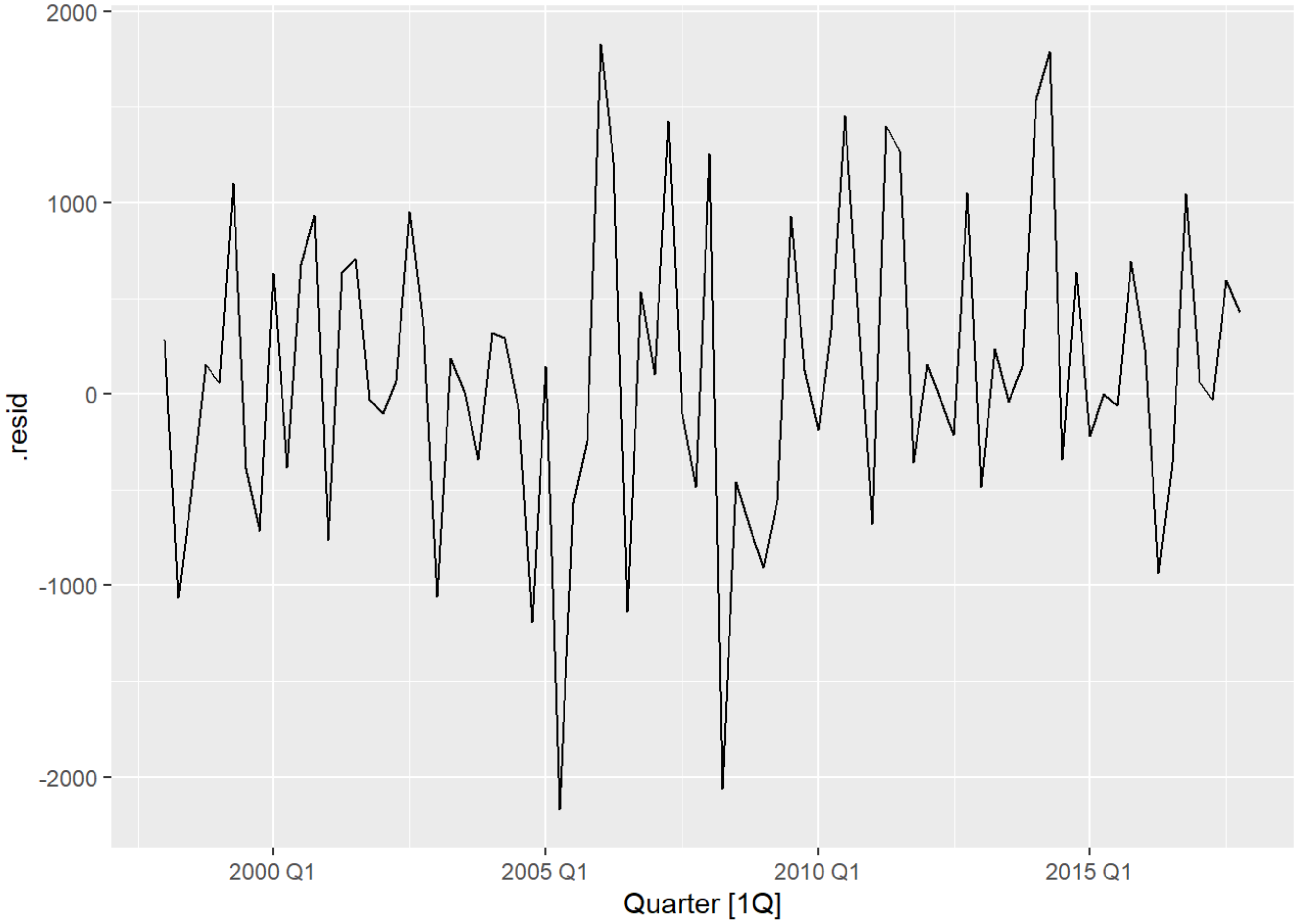
```
autoplot(fc, level = NULL) +
  autolayer(fc_holt, level = NULL) +
  autolayer(fc_ets, level = NULL) +
  xlab("Quarter") + ylab("Trips") +
  ggtitle("Forecasted Trips")
```



h. Check the residuals of your preferred model. Is it white noise? Explain briefly.

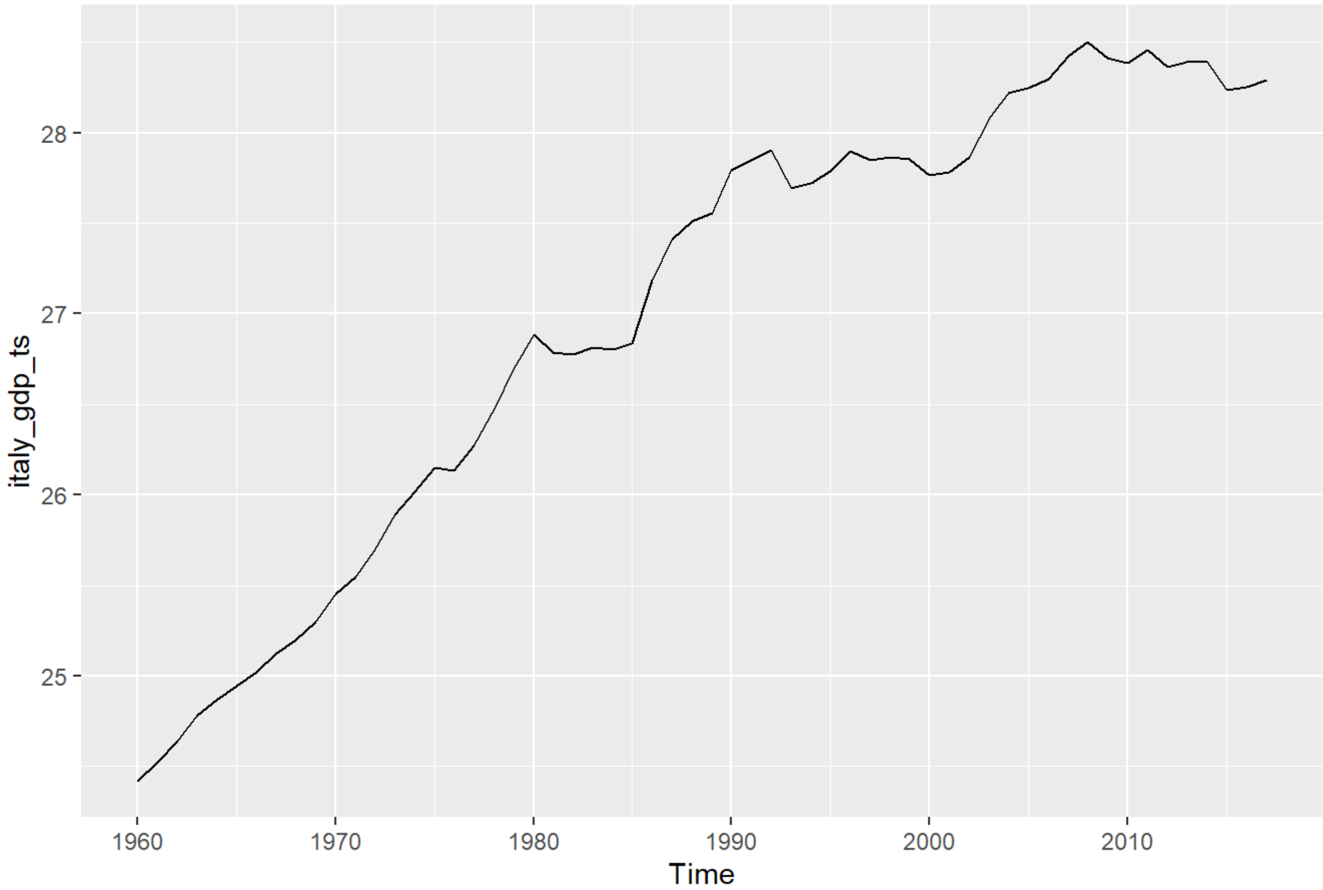
```
resid_ets <- resid(fit_ets)

autoplot(resid_ets, ylab = "Residuals")
```

```
## Plot variable not specified, automatically selected `.vars = .resid`
```



2. Choose a random European country's GDP(from global_economy):

```
set.seed(123)
italy_gdp <- global_economy %>%
  filter(Country == 'Italy')

italy_gdp_ts <- ts(log(italy_gdp$GDP), start = c(italy_gdp$Year[1], 1), frequency = 1)
autoplot(italy_gdp_ts)
```



a. Show the Box-Cox transformation is required and take the logarithm of the series.

- Based on the lambda value it indicates box cox transformation is required

```
lambda <- BoxCox.lambda(italy_gdp_ts)
lambda
```

```
## [1] -0.232903
```

b. fit a suitable ARIMA model to the transformed data using ARIMA();

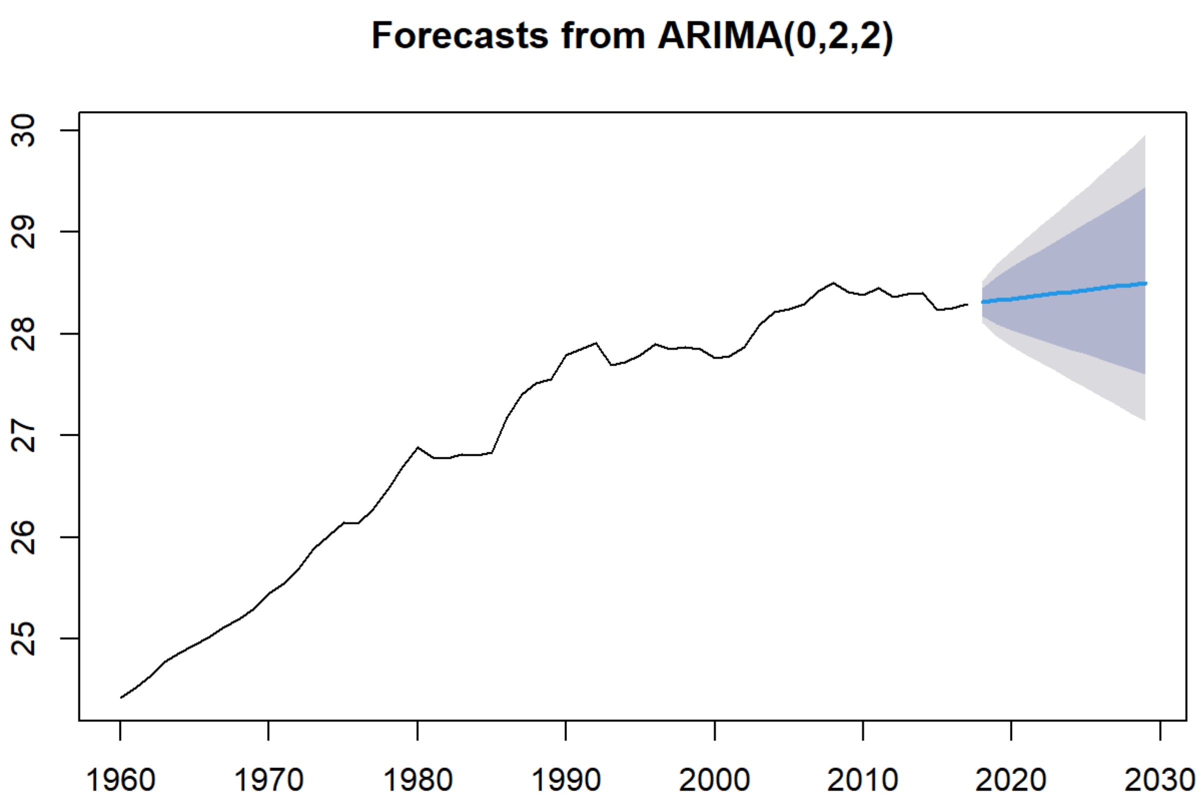
```
arima_model <- auto.arima(italy_gdp_ts, lambda = lambda)
summary(arima_model)
```

```
## Series: italy_gdp_ts
## ARIMA(0,2,2)
## Box Cox transformation: lambda= -0.232903
##
## Coefficients:
##           ma1      ma2
##       -0.6229  -0.2565
## s.e.   0.1201   0.1137
##
## sigma^2 = 2.915e-06:  log likelihood = 279.63
## AIC=-553.25   AICc=-552.79   BIC=-547.18
##
## Training set error measures:
##                       ME       RMSE         MAE         MPE      MAPE      MASE
## Training set -0.01714565 0.09724452 0.06849885 -0.06251957 0.2519853 0.6934574
##                     ACF1
## Training set -0.05461743
```

c. produce forecasts of 1 year of your fitted model. Do the forecasts look reasonable?

- Yes the forecasts looks reasonable

```
arima_forecast <- forecast::forecast(arima_model, h = 12)
plot(arima_forecast)
```



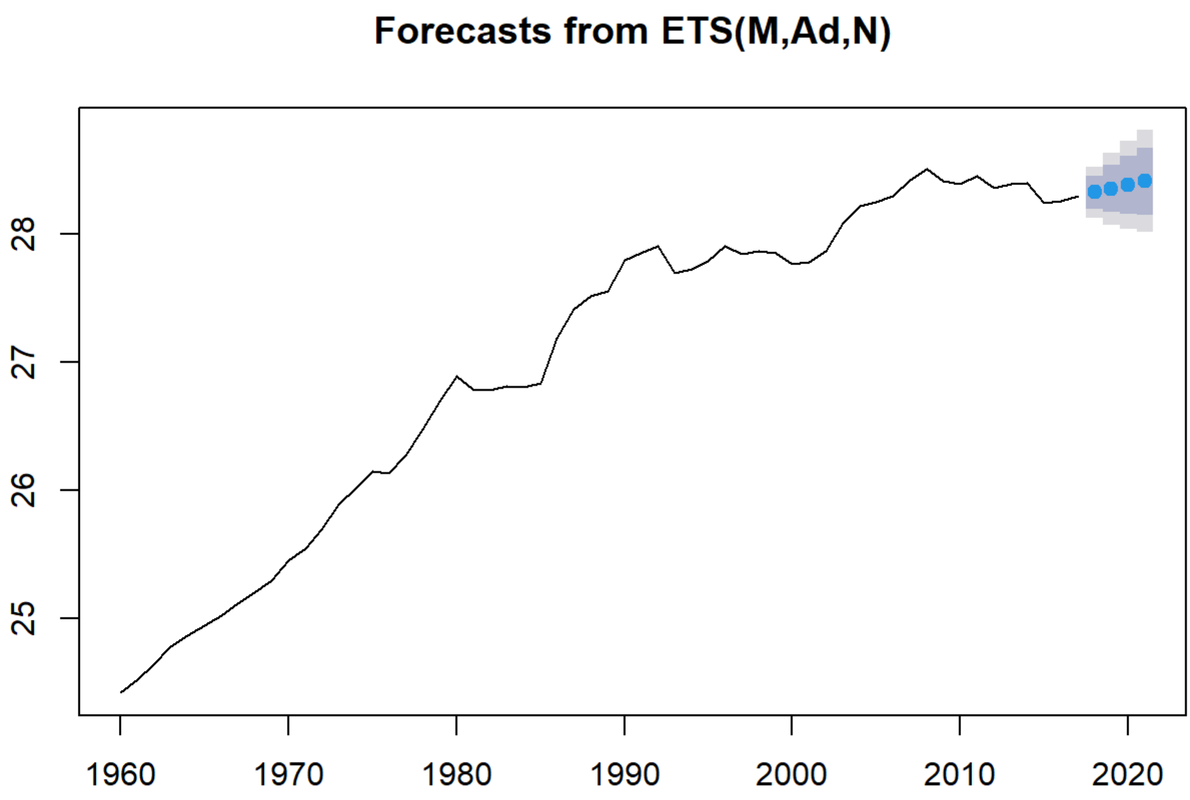
d. Fit an ETS model.

```
fit_ets <- ets(italy_gdp_ts)
summary(fit_ets)
```

```
## ETS(M,Ad,N)
##
## Call:
##  ets(y = italy_gdp_ts)
##
##   Smoothing parameters:
##     alpha = 0.9999
##     beta  = 1e-04
##     phi   = 0.9738
##
##   Initial states:
##     l = 24.2968
##     b = 0.1461
##
##   sigma:  0.0036
##
##       AIC      AICc       BIC
## -28.96007 -27.31302 -16.59742
##
## Training set error measures:
##                        ME      RMSE       MAE         MPE      MAPE      MASE
## Training set -0.004714582 0.0936841 0.0682602 -0.01740154 0.2508017 0.6910413
##                   ACF1
## Training set 0.251261
```

e. By using ETS model, forecast 1 year ahead.

```
ets_forecast <- forecast::forecast(fit_ets, h = 4)
plot(ets_forecast)
```



f. Compare the forecasted models of ARIMA and ETS to choose which one should be the best model.

```
accuracy_arima <- forecast::accuracy(arima_forecast)
accuracy_ets <- forecast::accuracy(ets_forecast)
cat("ARIMA Accuracy measures:\n")
```

```
## ARIMA Accuracy measures:
```

```
print(accuracy_arima)
```

```
##                       ME       RMSE        MAE         MPE      MAPE      MASE
## Training set -0.01714565 0.09724452 0.06849885 -0.06251957 0.2519853 0.6934574
##                     ACF1
## Training set -0.05461743
```

```
cat("ETS Accuracy measures:\n")
```

```
## ETS Accuracy measures:
```

```
print(accuracy_ets)
```

```
##                         ME       RMSE        MAE         MPE      MAPE      MASE
## Training set -0.004714582 0.0936841 0.0682602 -0.01740154 0.2508017 0.6910413
##                   ACF1
## Training set 0.251261
```

The best model is the one that has the lowest values for the error measures such as Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), Mean Absolute Percentage Error (MAPE), etc. Based on these ARIMA model is the best model.

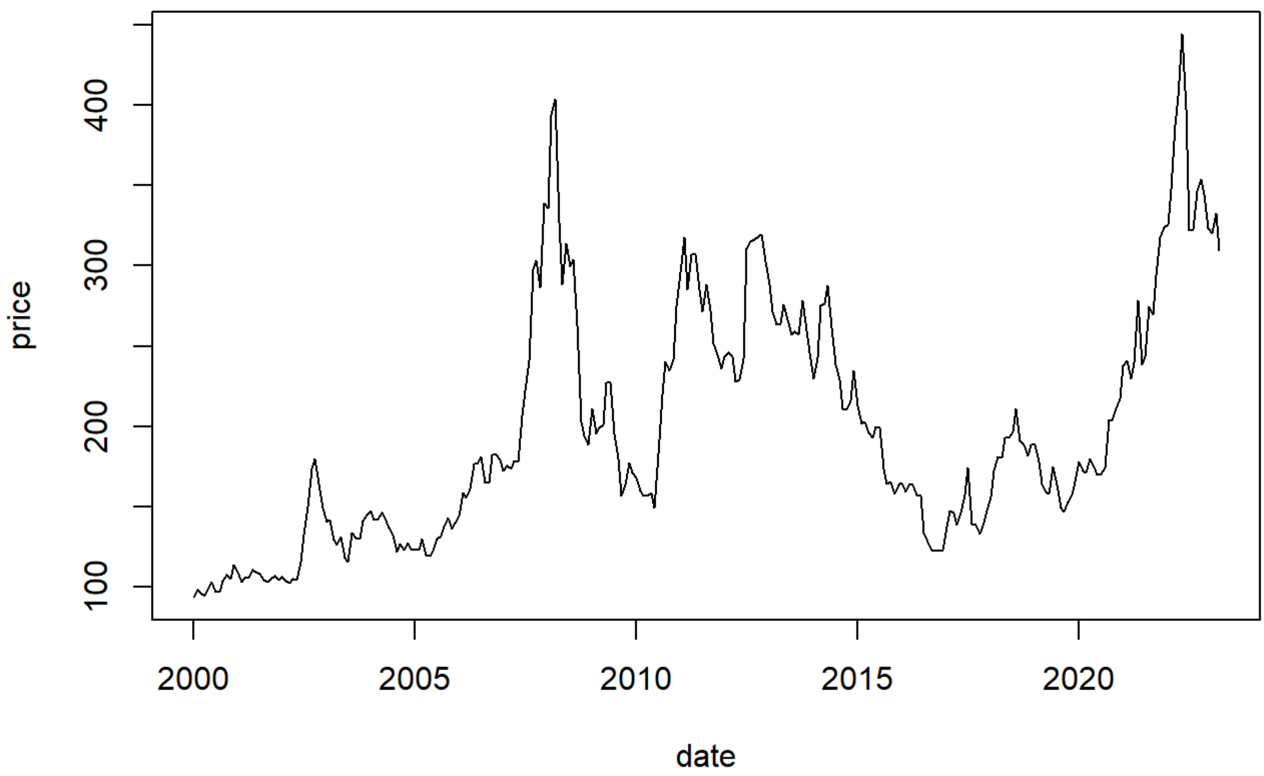
3. Get your own data:

a. Choose a series from any source.

```
wheat_prices <- read_excel("D:/Business Analytics/Business forecasting/Final Exam/Wheat Price
s.xls")
```

b. Plot the graph of the data and try to identify an appropriate ARIMA model.

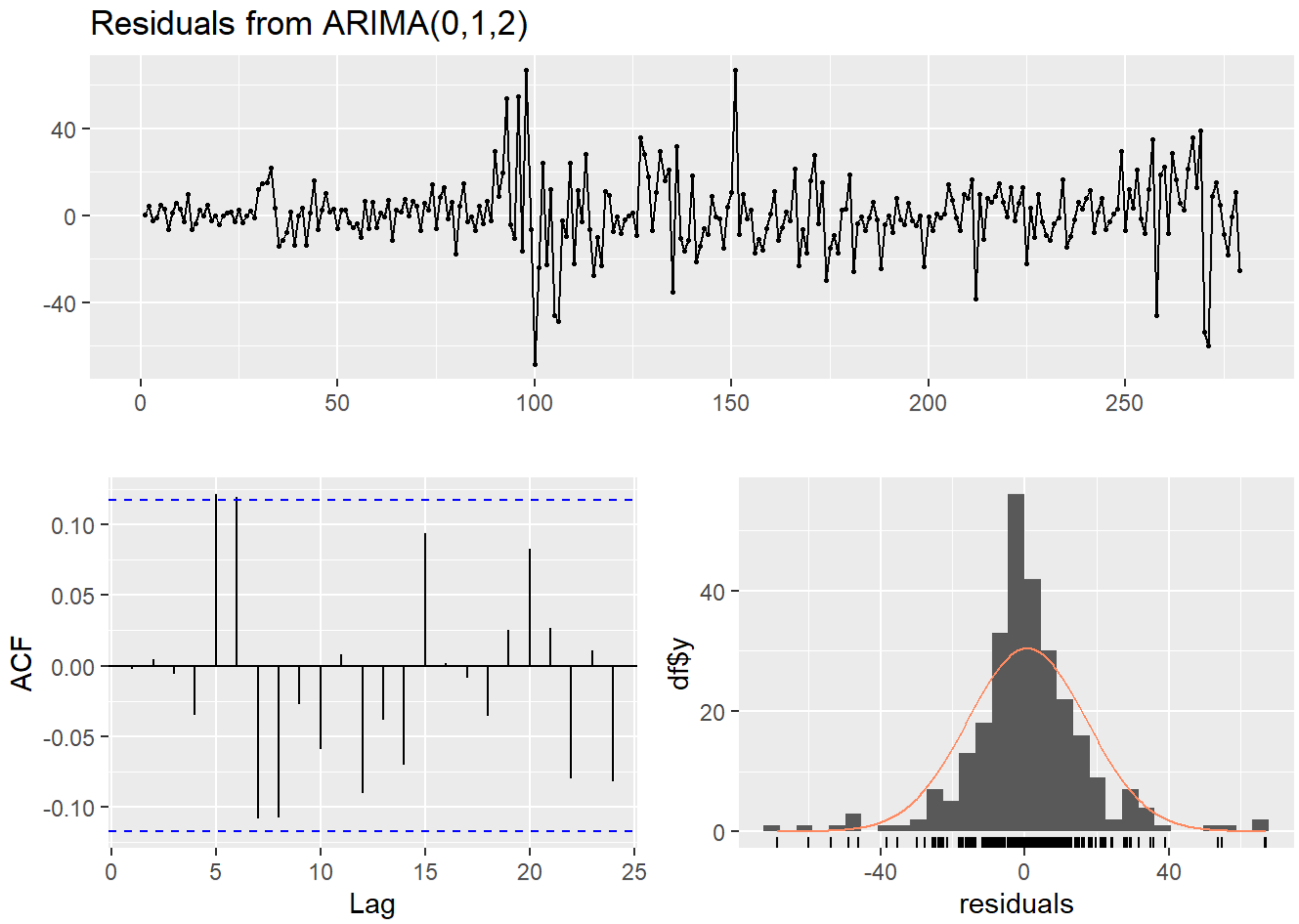
```
plot(wheat_prices$date, wheat_prices$price, type="l", xlab="date", ylab="price")
```




```
arima_model <- auto.arima(wheat_prices$price)
```

c. Do residual diagnostic checking of your ARIMA model. Are the residuals white noise?

```
checkresiduals(arima_model)
```




```
##
##  Ljung-Box test
##
## data:  Residuals from ARIMA(0,1,2)
## Q* = 16.595, df = 8, p-value = 0.03461
##
## Model df: 2.   Total lags used: 10
```

d. Use your chosen ARIMA model to forecast the next four years.

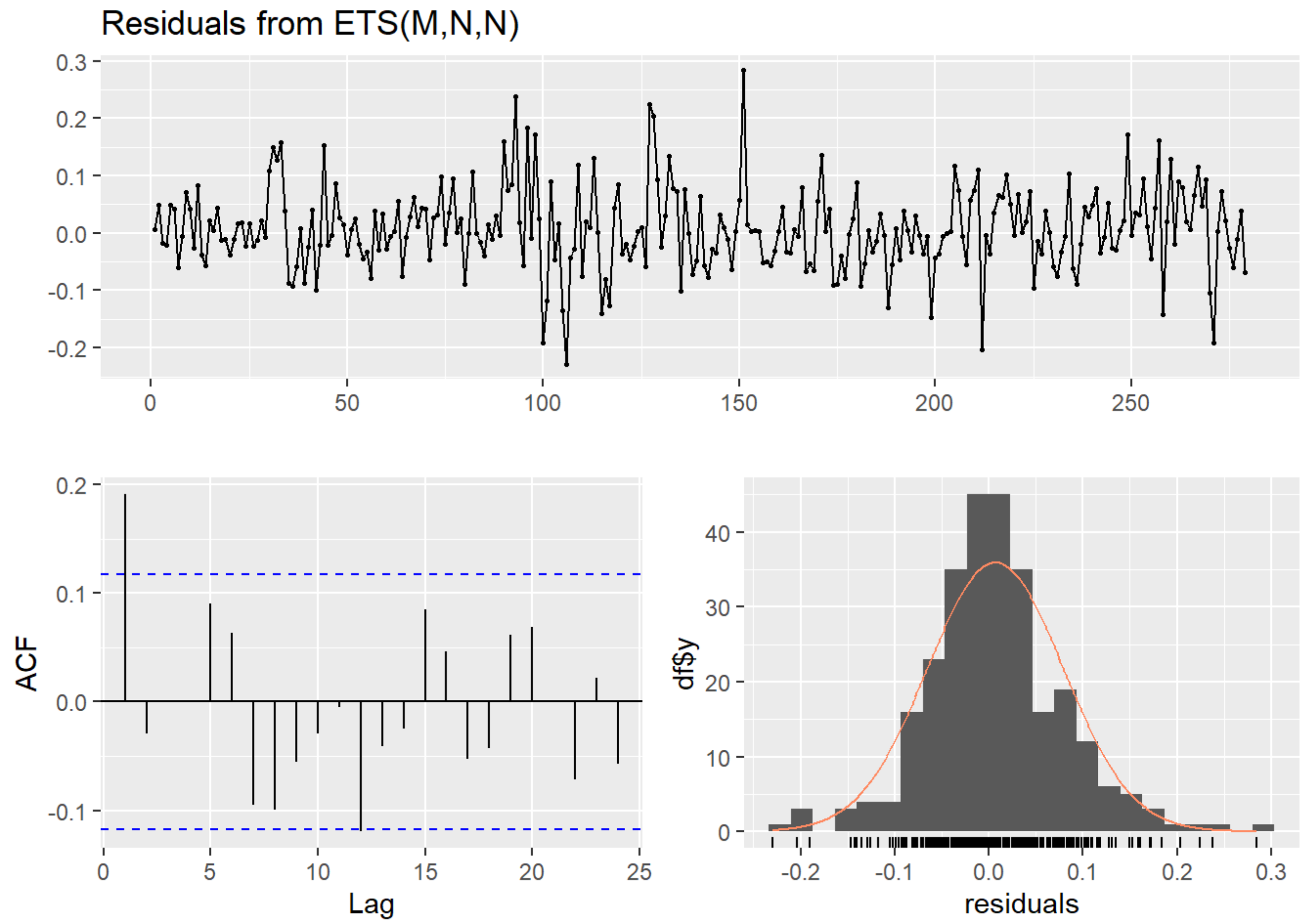
```
arima_forecast<- forecast::forecast(arima_model, h=48)
```

e. Now try to identify an appropriate ETS model.

```
ets_model <- ets(wheat_prices$price)
```

f. Do residual diagnostic checking of your ETS model. Are the residuals white noise?

```
checkresiduals(ets_model)
```



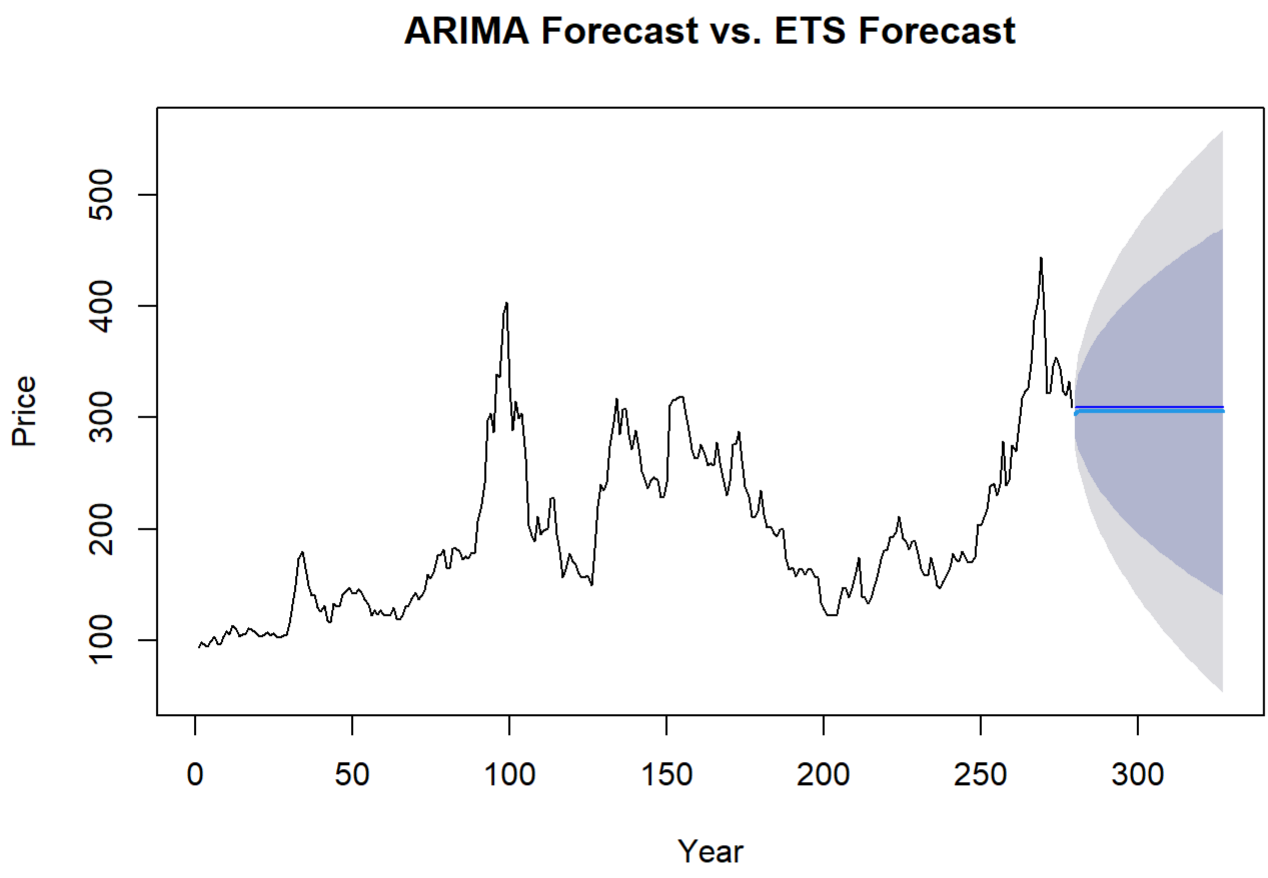

```
##
##  Ljung-Box test
##
## data:  Residuals from ETS(M,N,N)
## Q* = 20.582, df = 8, p-value = 0.008346
##
## Model df: 2.   Total lags used: 10
```

g. Use your chosen ETS model to forecast the next four years.

```
ets_forecast <- forecast::forecast(ets_model, h=48)
```

h. Which of the two models do you prefer?

```
plot(arima_forecast, xlab="Year", ylab="Price", main="ARIMA Forecast vs. ETS Forecast")
lines(ets_forecast$mean, col="blue")
```




```
forecast::accuracy(arima_forecast)
```

```
##                     ME     RMSE      MAE       MPE    MAPE      MASE
## Training set 0.6787217 16.52482 11.15713 0.1626657 5.26815 0.9943604
##                      ACF1
## Training set -0.002539645
```

```
forecast::accuracy(ets_forecast)
```

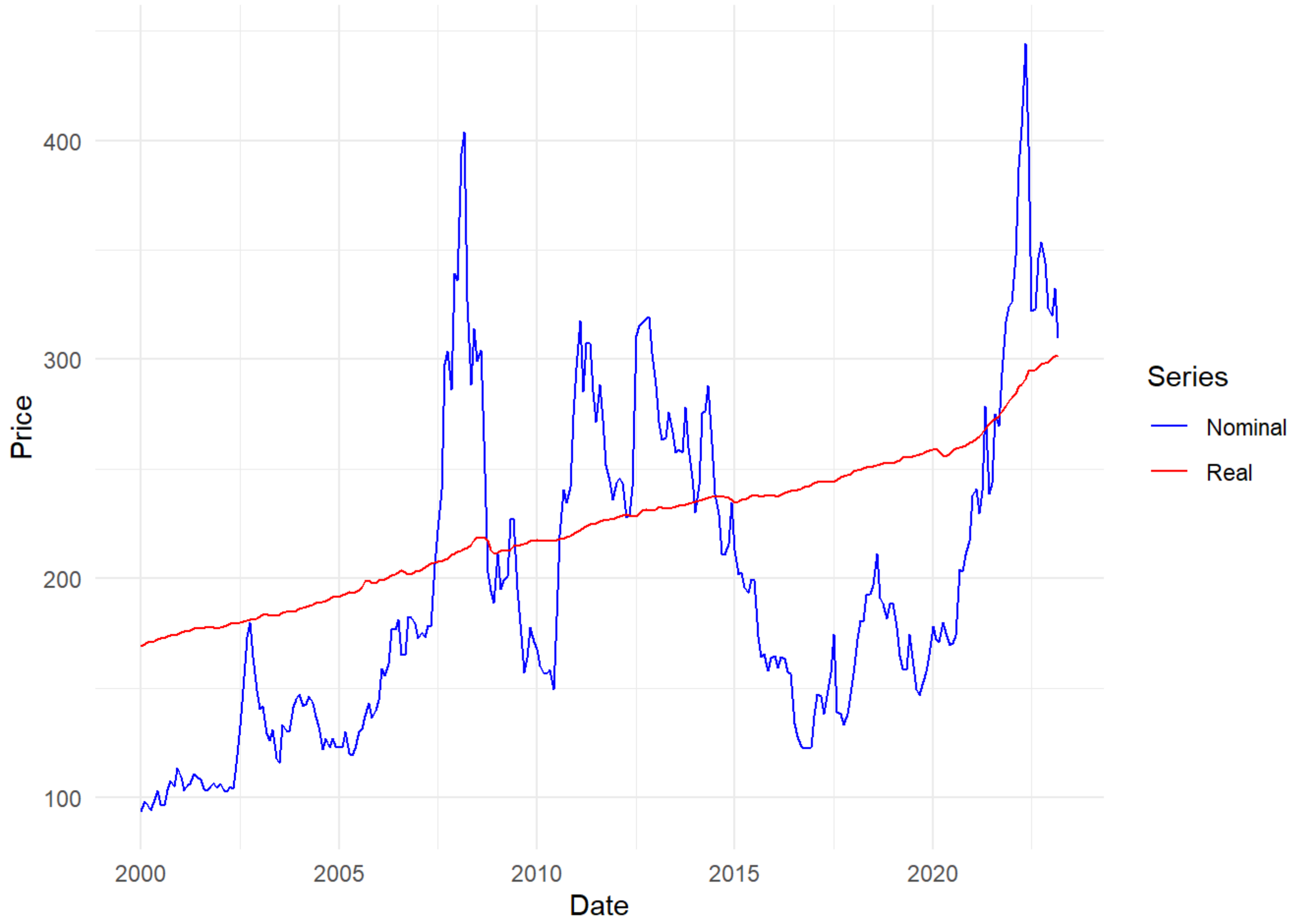
```
##                     ME     RMSE     MAE       MPE    MAPE      MASE      ACF1
## Training set 0.7754415 16.98876 11.1823 0.1732938 5.27173 0.9966033 0.1848067
```

4. Get Consumer Price Index(CPI) from Quandl

a. Use CPI to make the real value of the series you chose in question #3. (Hint: ex: the series in question#3 divided by CPI. They need to be on the same length)

```
wheat_cpi <- read_excel("D:/Business Analytics/Business forecasting/Final Exam/CPI - Wheat.xl
s")

wheat_cpi$real_price <- wheat_prices$price / wheat_cpi$cpi
```

```
ggplot() +
  geom_line(data = wheat_prices, aes(x = date, y = price, color = "Nominal")) +
  geom_line(data = wheat_cpi, aes(x = date, y = cpi, color = "Real")) +
  labs(x = "Date", y = "Price", color = "Series") +
  scale_color_manual(values = c("Nominal" = "blue", "Real" = "red")) +
  theme_minimal()
```



5. By using pdfetch or quantmod packages, obtain any stock's price from Dow Jones 30(different than APPLE that I did in class notes) and Dow Jones Industrial Average (DJIA)

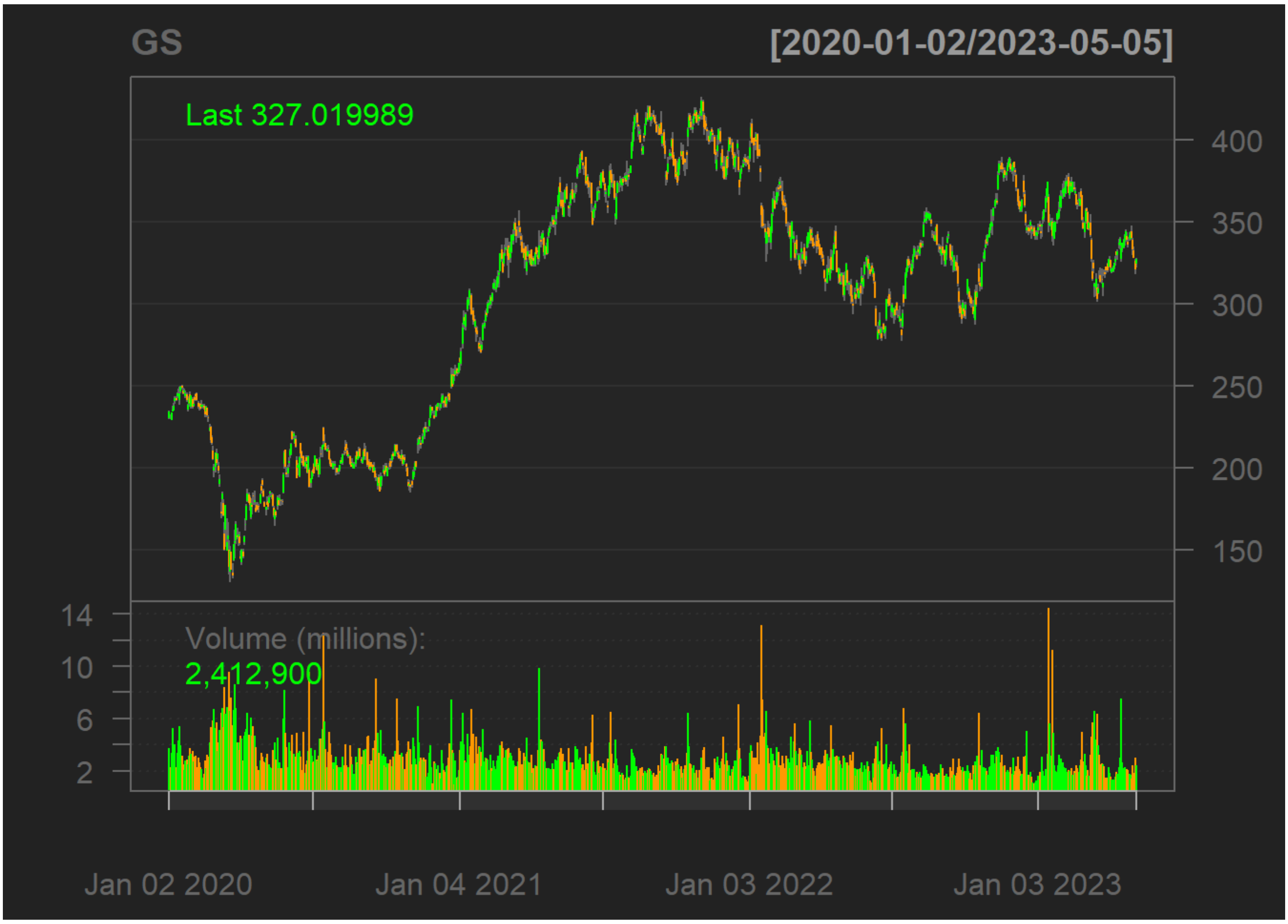
a. Plot each series. What do you observe from the graphs?

- Graph shows the price of GS and DJIA both increased over a time period, But magnitude and volatility are different for each series.

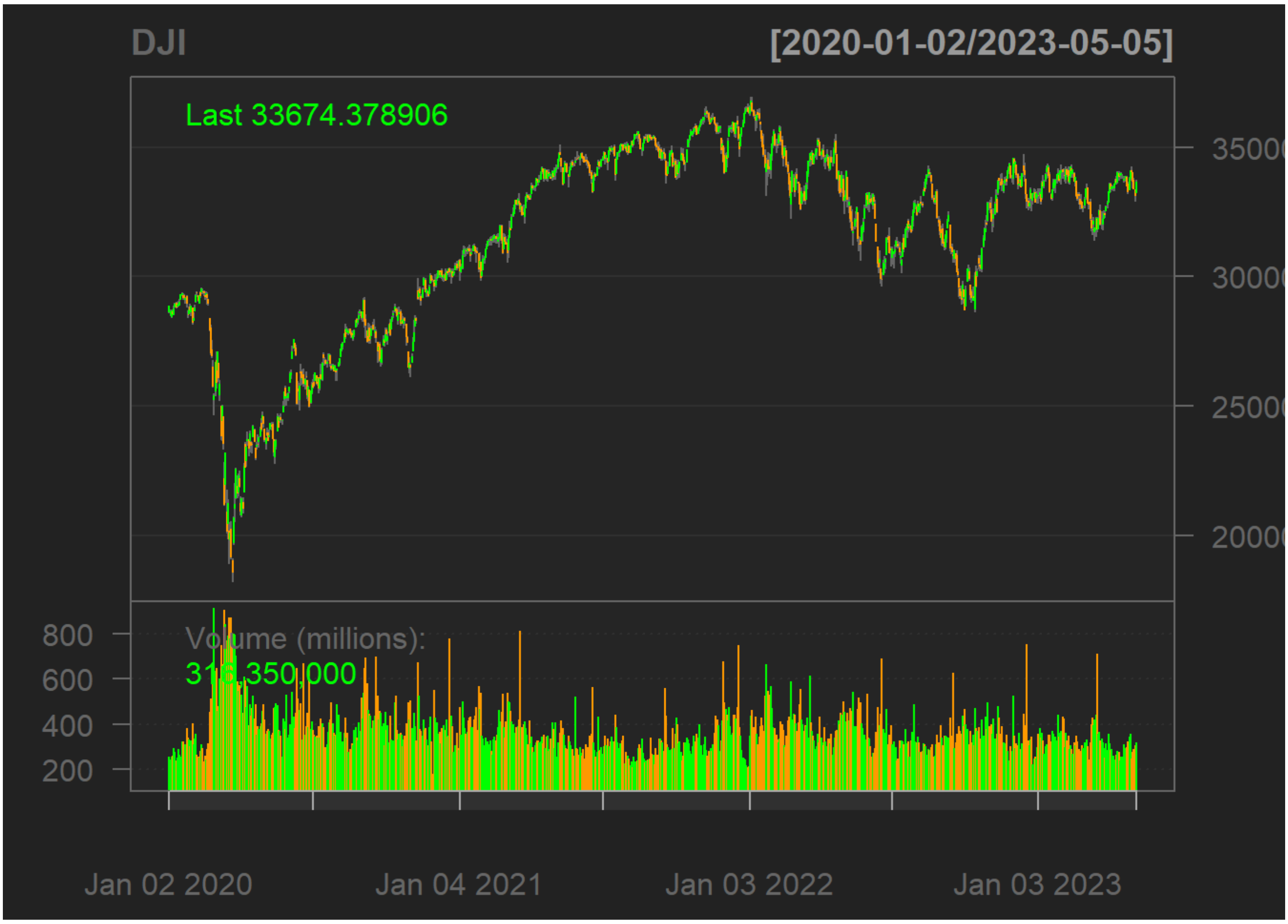
```
getSymbols(c("GS", "^DJI"), from = "2020-01-01", to = Sys.Date())
```

```
## [1] "GS"   "^DJI"
```

```
chartSeries(GS)
```




```
chartSeries(DJI)
```



b. Calculate the return of the stock price you chose.

```
GS_returns <- dailyReturn(GS)
```

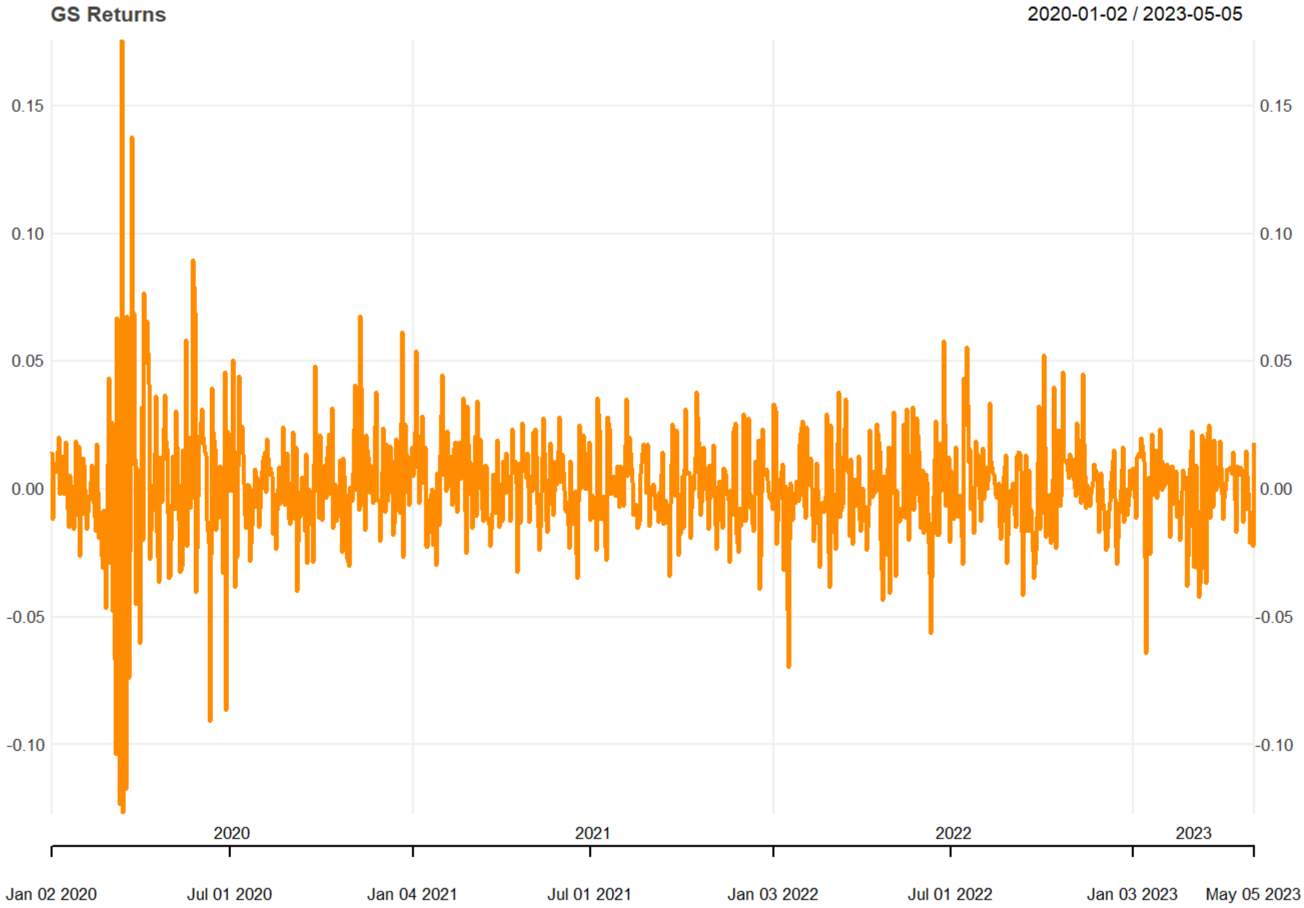
c. Calculate the return of DJIA

```
DJI_returns <- dailyReturn(DJI)
```

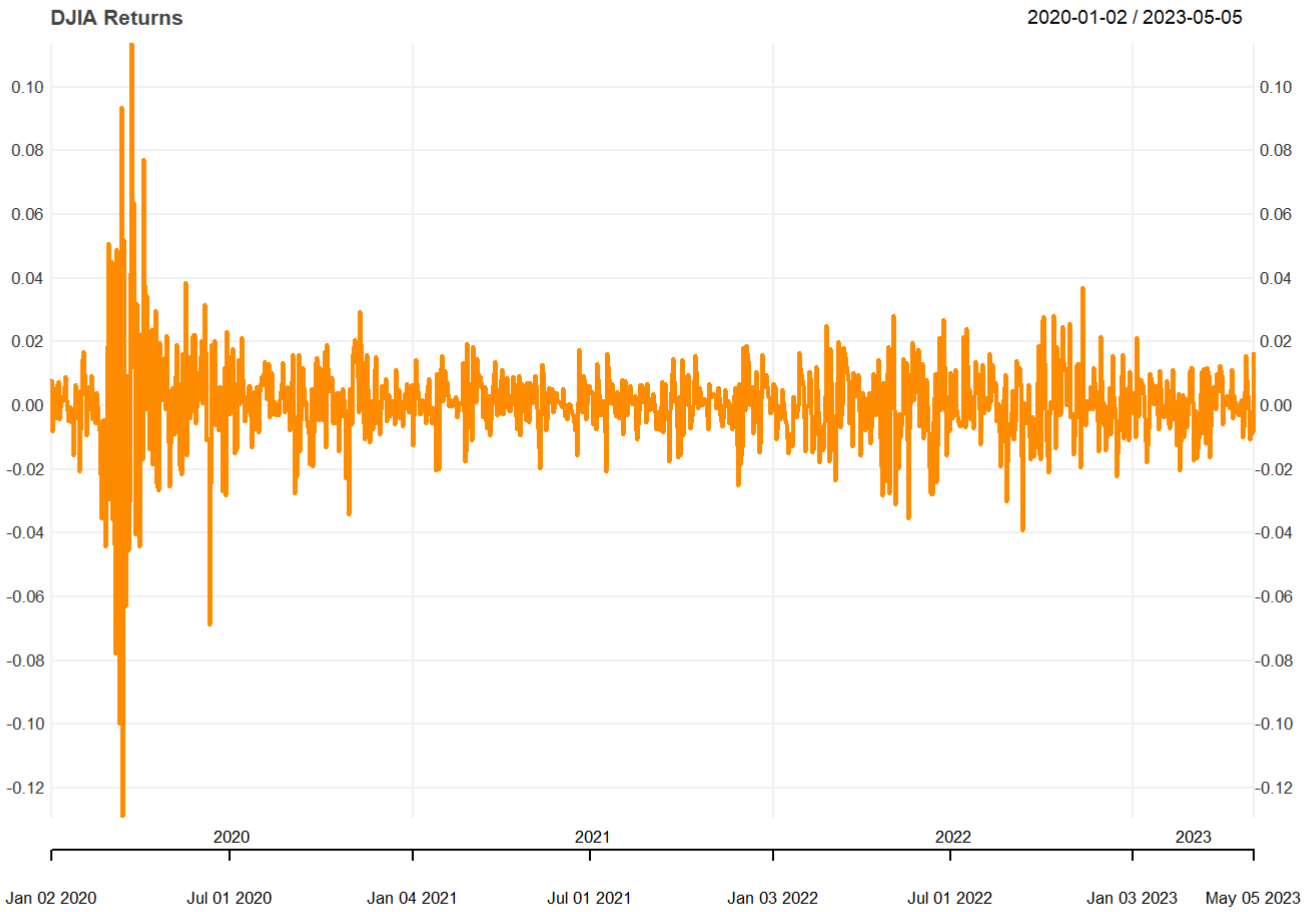
d. Plot the return of each series. What do you observe?

- From the graphs, we can observe that the returns of both GS and DJIA were not too volatile, but GS had higher fluctuations than DJIA.

```
chart_Series(GS_returns, name = "GS Returns")
```




```
chart_Series(DJI_returns, name = "DJIA Returns")
```



e. What is the correlation coefficient between the stock price and DJI. Interpret the result. Is there a strong (or weak) positive (or negative) correlation between them?

```
cor(GS, DJI)
```

```
##               DJI.Open   DJI.High    DJI.Low  DJI.Close DJI.Volume DJI.Adjusted
## GS.Open      0.9509351  0.9527283  0.9485643  0.9496181 -0.4176854    0.9496181
## GS.High      0.9484563  0.9513550  0.9463820  0.9483758 -0.4098737    0.9483758
## GS.Low       0.9516116  0.9537378  0.9506171  0.9516914 -0.4301794    0.9516914
## GS.Close     0.9488922  0.9521943  0.9481993  0.9510714 -0.4209202    0.9510714
## GS.Volume   -0.3437632 -0.3383632 -0.3581810 -0.3501161  0.5675515   -0.3501161
## GS.Adjusted  0.9456075  0.9493179  0.9442833  0.9476726 -0.4187037    0.9476726
```

f. Estimate CAPM regression equation (without interest rates).

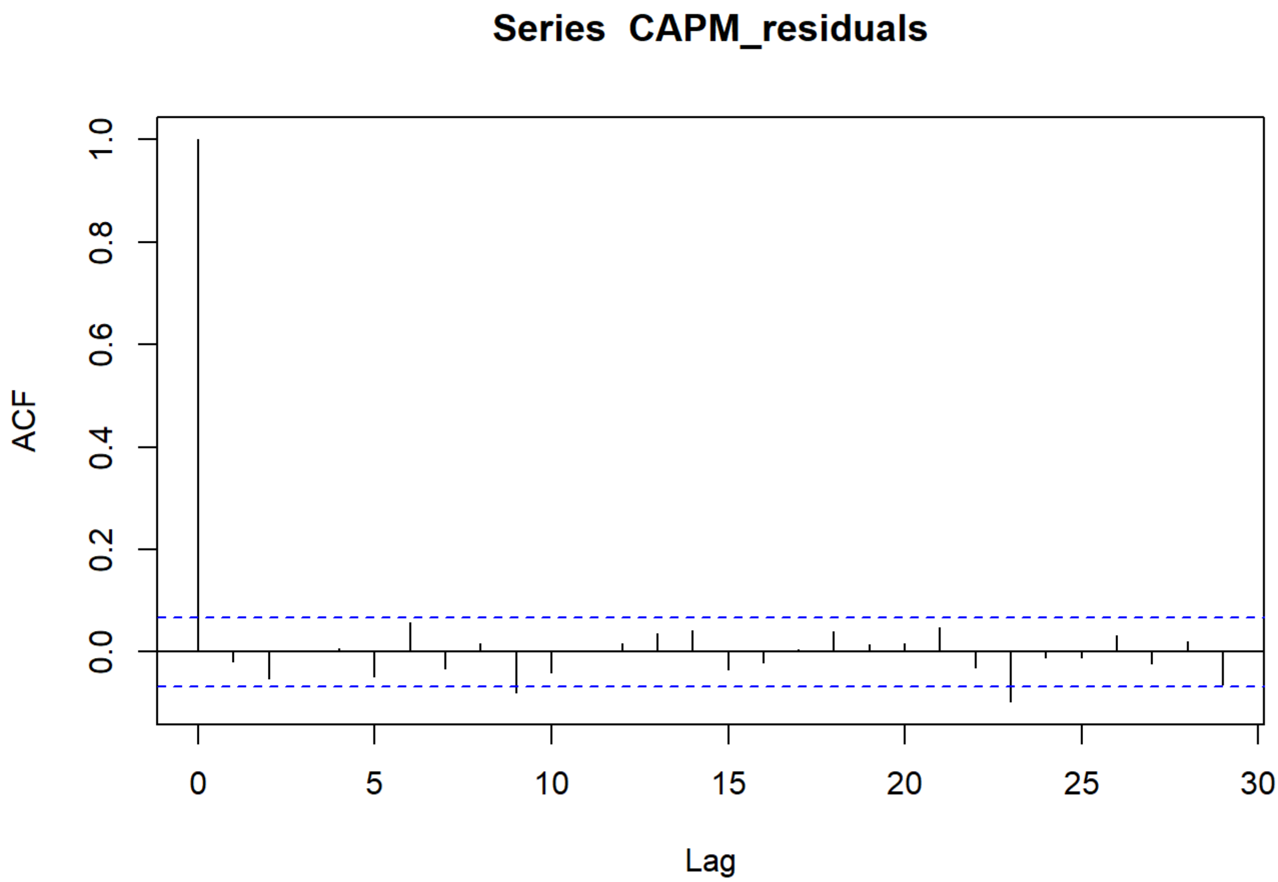
```
CAPM <- lm(GS_returns ~ DJI_returns)
summary(CAPM)
```

```
##
## Call:
## lm(formula = GS_returns ~ DJI_returns)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.051310 -0.007125 -0.000687  0.007541  0.062265
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.0002917  0.0004475   0.652    0.515
## DJI_returns 1.2505642  0.0293305  42.637   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.01298 on 840 degrees of freedom
## Multiple R-squared:  0.684,	Adjusted R-squared:  0.6836
## F-statistic:  1818 on 1 and 840 DF,  p-value: < 2.2e-16
```

g. What is the beta of the stock. Does it tell you is it riskier than market or not?

With a beta of 1.2505642, the stock is estimated to have a slightly higher volatility or risk compared to the overall market represented by the DJIA. However, it is important to note that beta alone does not provide a complete measure of a stock's risk. Other factors, such as the stock's specific characteristics, sector, and company fundamentals, should also be considered when assessing its risk profile.

h. Check the residuals diagnosis. Is it white noise. Explain what is white noise.


Series CAPM_residuals
ACF
0.0
0.2
0.4
0.6
0.8
1.0
0
5
10
15
20
25
30
Lag