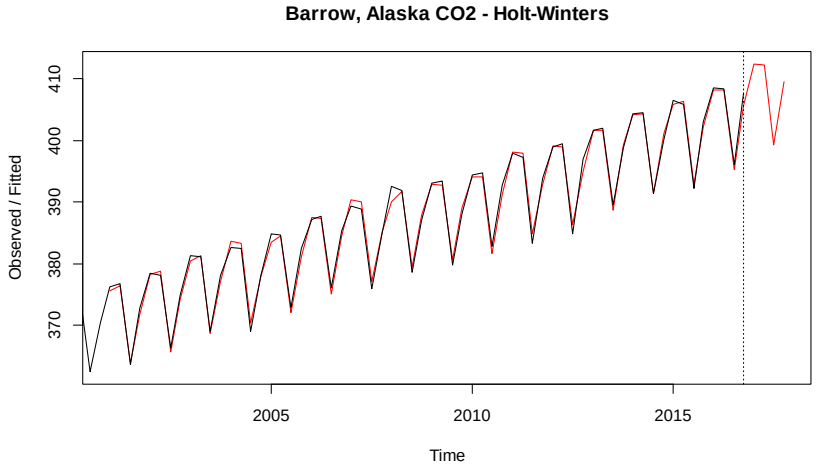
STATS 326
Applied Time Series
ASSIGNMENT TWO
ANSWER GUIDE

Question One:

```
> HW.fit = HoltWinters(red.CO2.ts)
> HW.fit
Holt-Winters exponential smoothing with trend and additive seasonal
component.

Call:
HoltWinters(x = red.CO2.ts)

Smoothing parameters:
 alpha: 0.5732177
 beta : 0.03867122
 gamma: 0.4140589

Coefficients:
          [,1]
a  406.4502058
b    0.5868941
s1   5.3662780
s2   4.6109173
s3  -8.8837987
s4   0.7093982

> HW.pred = predict(HW.fit,n.ahead=4)
> HW.pred
         Qtr1     Qtr2     Qtr3     Qtr4
2017 412.4034 412.2349 399.3271 409.5072

> actual
2017.1 2017.2 2017.3 2017.4
413.75 412.42 398.47 408.44

> HW.RMSEP = sqrt(1/4*sum((actual-HW.pred)^2))
> HW.RMSEP
[1] 0.9645109
```

```
> plot(HW.fit,HW.pred,main="Barrow, Alaska CO2 - Holt-Winters")
```



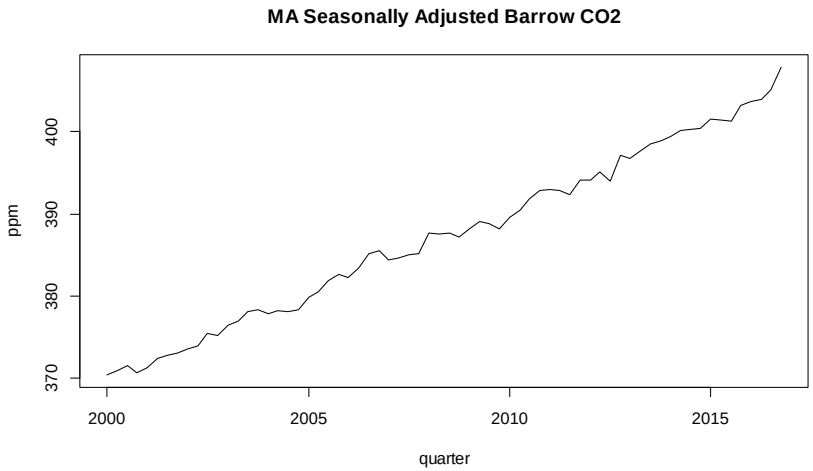
The plot of the Holt-Winters model shows an excellent fit with the model (red line) being very close to the actual observations (black line).

The predictions are below the 2017 actual values for Quarters 1 & 2 and above the actual values for Quarters 3 & 4.

The RMSEP shows that, on average, the prediction error is 0.96 ppm.

Question Two:

```
> plot(MA.CO2.ts,main="MA Seasonally Adjusted Barrow
  CO2",xlab="quarter",ylab="ppm")
```



The plot of the moving average seasonally adjusted series shows a reasonably linear trend with Quarter 4 of 2016 being slightly higher than expected, given the trend.

```
> MA.Barrow$figure
[1]  4.9286328  4.3242578 -9.0356641 -0.2172266
```

The seasonal estimates for Quarter 1 & Quarter 2 are very similar positive values (above the trend) while Quarter 3 and Quarter 4 are below the trend. Quarter 1 is the highest (4.93) while Quarter 3 is the lowest (-9.04).

```
> summary(MA.Barrow.fit2)

Call:
lm(formula = MA.CO2.ts[-1] ~ Time[-1] + MA.CO2.ts[-68])

Residuals:
    Min      1Q  Median      3Q     Max
-1.9475 -0.4709 -0.0720  0.4565  2.5262

Coefficients:
               Estimate Std. Error t value Pr(>|t|)
(Intercept)    198.50565   46.50963   4.268 6.65e-05 ***
Time[-1]         0.28589    0.06611   4.324 5.46e-05 ***
MA.CO2.ts[-68]   0.46263    0.12615   3.667    5e-04 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.7696 on 64 degrees of freedom
Multiple R-squared:  0.9946,    Adjusted R-squared:  0.9944
F-statistic:  5897 on 2 and 64 DF,  p-value: < 2.2e-16
```

```
> t.69.ma.pred = MA.Barrow.fit2$coef[1]+MA.Barrow.fit2$coef[2]*69+
MA.Barrow.fit2$coef[3]*MA.CO2.ts[68]
> t.69.ma.pred
(Intercept)
   406.9338
> t.69.pred = t.69.ma.pred+MA.Barrow$figure[1]
> t.69.pred
(Intercept)
   411.8624
> t.70.ma.pred = MA.Barrow.fit2$coef[1]+MA.Barrow.fit2$coef[2]*70+
  MA.Barrow.fit2$coef[3]*t.69.ma.pred
> t.70.ma.pred
(Intercept)
   406.7786
> t.70.pred = t.70.ma.pred+MA.Barrow$figure[2]
> t.70.pred
(Intercept)
   411.1028
> t.71.ma.pred = MA.Barrow.fit2$coef[1]+MA.Barrow.fit2$coef[2]*71+
  MA.Barrow.fit2$coef[3]*t.70.ma.pred
> t.71.ma.pred
(Intercept)
   406.9927
> t.71.pred = t.71.ma.pred+MA.Barrow$figure[3]
> t.71.pred
(Intercept)
    397.957
> t.72.ma.pred = MA.Barrow.fit2$coef[1]+MA.Barrow.fit2$coef[2]*72+
  MA.Barrow.fit2$coef[3]*t.71.ma.pred
> t.72.ma.pred
(Intercept)
   407.3776
> t.72.pred = t.72.ma.pred+MA.Barrow$figure[4]
> t.72.pred
(Intercept)
   407.1604

> MA.pred = c(t.69.pred,t.70.pred,t.71.pred,t.72.pred)
> names(MA.pred) = c("2017.1","2017.2","2017.3","2017.4")
> MA.pred
  2017.1   2017.2   2017.3   2017.4
411.8624 411.1028 397.9570 407.1604

> RMSEP.MA.Barrow = sqrt(1/4*sum((actual-MA.pred)^2))
> RMSEP.MA.Barrow
[1] 1.341502
```

Question Three:

The seasonal estimates show that the CO2 concentration is above the overall trend for the first 2 quarters with Quarter 1 being the highest (4.93) and below the overall trend in the last 2 quarters with Quarter 3 being the lowest (-9.04).

The plot of the seasonally adjusted series shows an increasing reasonably linear trend.The last value, Quarter 4, 2016 is higher than expected given the overall trend.

The final model included a trend term and a lagged response to take care of autocorerelation detected in the Residual Series.

For the final model, the Residual Series appears to be random scatter about 0 with a large positive residual for the last observation. The plot of the autocorrelation function of the residuals shows no significant lags. The residuals appear to be normally distributed (Shapiro-Wilk *P-value* = 0.16) with slight right skew (min = -1.95, max = 2.53). The assumptions appear to be satisfied.

We have strong evidence against the hypothesis that the coefficient associated with the Time variable is 0 (*P-value* = 0.0000546). We have very strong evidence against the hypothesis of no autocorrelation (*P-value* = 0.0005).

The *F*-statistic provides extremely strong evidence against the hypothesis that none of the variables are related to the seasonally adjusted CO2 concentration (*P-value* ≈ 0). The Multiple R^2 is almost 1 (0.9946) indicating that nearly all the variation in the seasonally adjusted CO2 concentration is explained by the model.

The Residual Standard Error is 0.77 ppm so prediction intervals should be reasonably narrow. The model predictions can be relied on as the assumptions appear to be satisfied. The RMSEP for the 2017 predictions was 1.34 ppm which was smaller than that for the Seasonal Trend Lowess model (1.41 ppm). Our predictions for 2017 were:

Quarter 1: 411.86 ppm
Quarter 2: 411.10 ppm
Quarter 3: 397.96 ppm
Quarter 4: 407.16 ppm

Question Four:

```
> HW.Full.fit = HoltWinters(full.CO2.ts)
> HW.Full.fit
Holt-Winters exponential smoothing with trend and additive seasonal
component.

Call:
HoltWinters(x = full.CO2.ts)

Smoothing parameters:
 alpha: 0.6655127
 beta : 0.02583596
 gamma: 0.5504329

Coefficients:
          [,1]
a  407.8598399
b    0.5251393
s1   5.6488674
s2   4.5145080
s3  -9.0761435
s4   0.6730279

> HW.Full.pred = predict(HW.Full.fit,n.ahead=4)
> HW.Full.pred
         Qtr1     Qtr2     Qtr3     Qtr4
2018 414.0338 413.4246 400.3591 410.6334
```

The Holt-Winters estimates for the Full timeframe model are very similar to those from Reduced timeframe model.

Since the model follows the observations closely, the predictions should be reasonably reliable.

The prediction intervals are between 3.5 and 5.6 ppm in width.