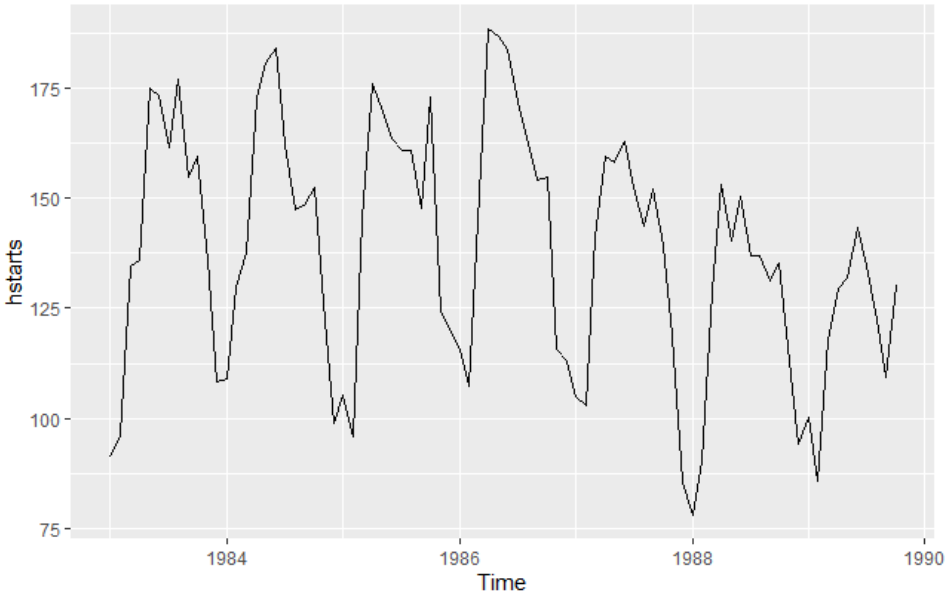
Module 2 Peer Assignment Sample Answers

Q1-A

head(hstarts,10)

```
      Jan   Feb   Mar   Apr   May   Jun   Jul   Aug   Sep   Oct
1983 91.3  96.3 134.6 135.8 174.9 173.2 161.6 176.8 154.9 159.3
```

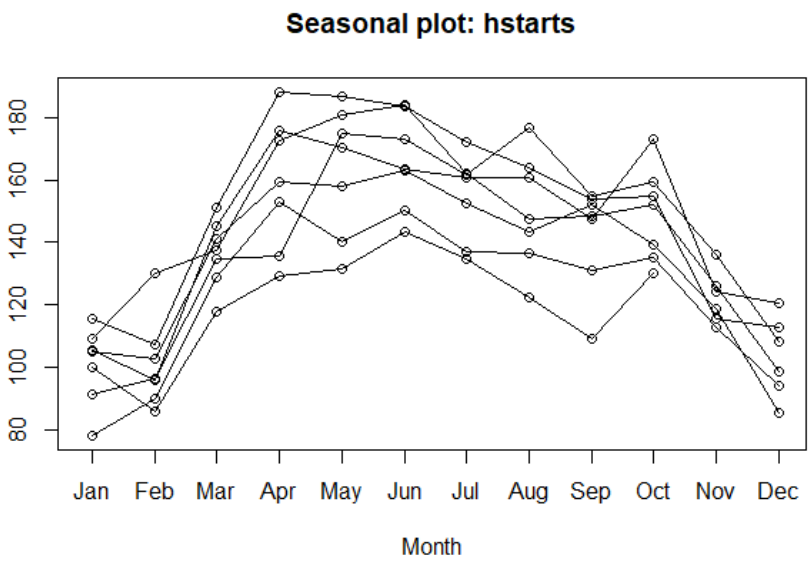
Q1-B



Q1-C

The y-axis is the number of housing starts. The x-axis is the time period in years. There do not seem to be a trend. (use your judgement here.) There does seem to be a seasonal pattern. Housing starts seem to rise. During the summer months and drop off at the end of the year.

Q1-D



Q1-E

The housing starts are on the y-axis. The time period in months is on the x-axis. There does not seem to be a downward or upward trend. (use your judgement here.) The number of housing starts seems to rise during April through September. The number of housing starts seen to drop in the fall months.

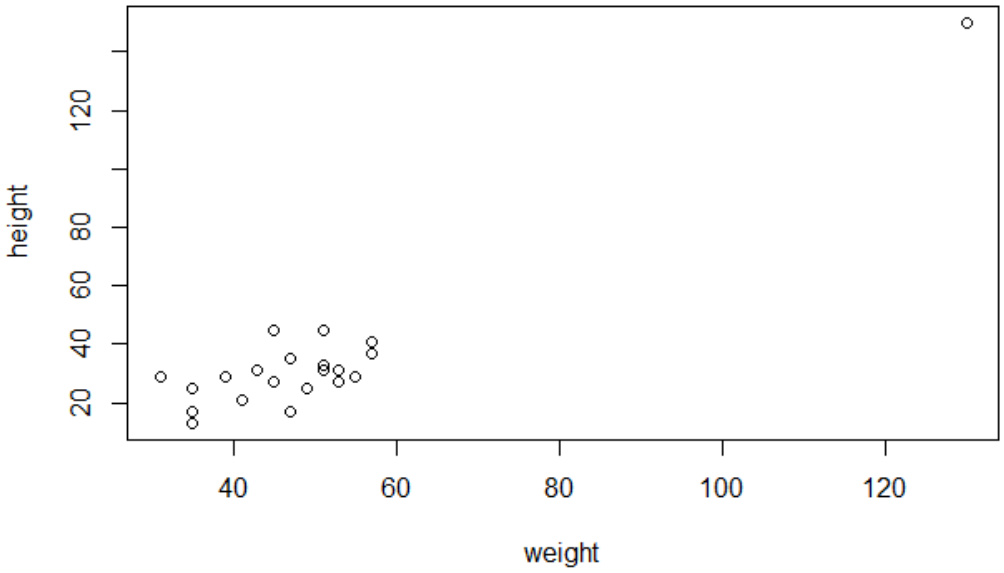
Q2-A

```
summary(height)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  13.00   25.00   29.00   35.14   35.00  150.00
summary(weight)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
     31      41      47      50      53     130
```

Q2-B

The range of the height is 137 with a minimum of 13 and a maximum of 150. The median height is 29 and the mean height is 35.14. The range of the weight is 99 with a minimum of 31 and a maximum of 130. The median weight is 47 and the mean height is 50.

Q2-C



Q2-D

The y-axis represents the height and the x-axis represents the weight. All put one data point is clustered at the bottom left corner. There seems to be an outlier in the top right corner. Inspecting the actual data, the outlier has a value of 130 weight and 150 height.

Q2-E

```
fit <- lm(height~weight)
summary(fit)

Call:
lm(formula = height ~ weight)

Residuals:
    Min      1Q  Median      3Q     Max
-14.219  -7.298  -2.372   8.243  18.706

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -30.2495     5.8203  -5.197 5.13e-05 ***
weight        1.3078     0.1085  12.051 2.41e-10 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 9.646 on 19 degrees of freedom
Multiple R-squared:  0.8843,   Adjusted R-squared:  0.8782
F-statistic: 145.2 on 1 and 19 DF,  p-value: 2.412e-10
```

Q2-F

- Coefficient 1.3078
- The coefficient is significant at the .001 level

- Y will change by 1.3078 units when there is a one unit change in X
- Adj. R-squared .8782. This can be interpreted as 88% of hte variance in Y can be explained by X

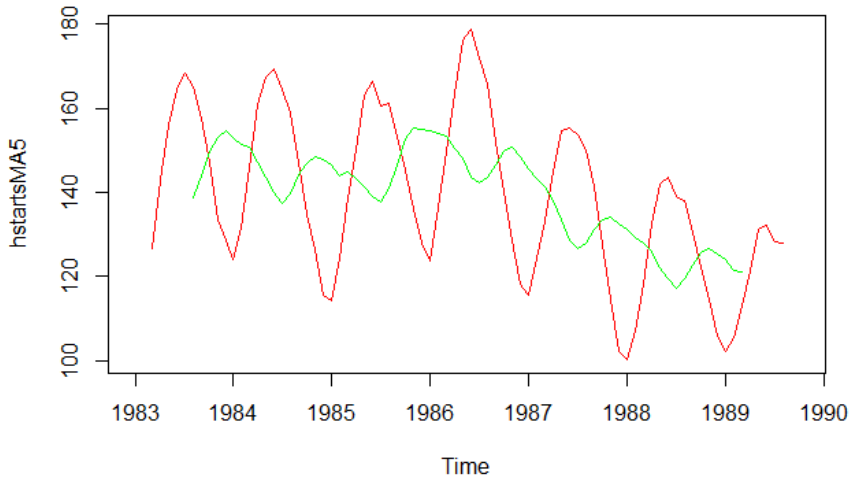
Q3-A

hstartsMA5 <- ma(hstarts,5)

Q3-B

hstartsMA15 <- ma(hstarts,15)

Q3-C



The MA15 model is smoother. The curve is dampened and does not exhibit the extremes of the seasonal cycle.

Q4-A

hstartsSES5 <- ses(hstarts,h=5)

Q4-B

```
> accuracy(hstartsSES5)
                     ME      RMSE       MAE         MPE     MAPE      MASE
ACF1
Training set 0.4756625 19.46083 14.66371 -0.6897685 11.0748 1.137856 0.159
6446
```

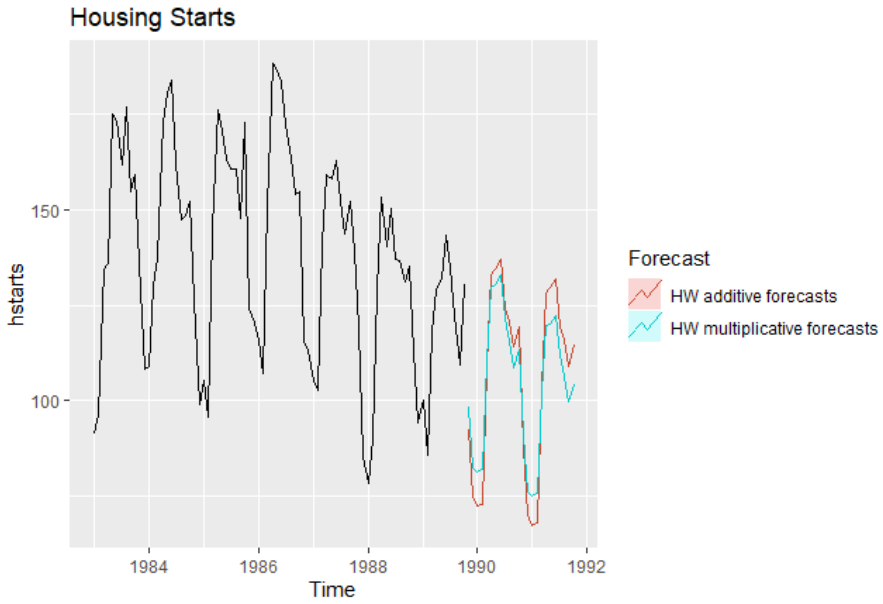
Q4-C

hstartsHWa <- hw(hstarts,seasonal = "additive")

Q4-D

hstartsHWm <- hw(hstarts,seasonal = "multiplicative")

Q4-E



Both models are good possibilities for forecasts as they both exhibit the cyclical cycles as well the downward trend from 1987 onward. However, the multiplicative mode also seems to dampen the cycles. A visual inspection of the previous 3-4 cycles seem to confirm the dampening of the cycles.