Financial Time Series Homework2

Name: Venkata Avinash

Question 1: In this problem we will perform multiple regression on the Boston housing data. The data contains 506 records with 14 variables. The variable medv is the response variable. To assess the data use library(MASS) data(Boston)

(a) First perform a multiple regression with all the variables, what can you say about the significance of the variables based on only the p-values. Next use the" step" function to perform backward selection using (1) the AIC criteria and (2) the BIC criteria then compare the results. (By default, the step function in R performs variable selection based on AIC criteria. Read the documentation to find out how to do the selection using BIC criteria.)

```r
library(MASS)
fix(Boston)
summary(Boston)
```

```
##       crim                zn             indus            chas
##  Min.   : 0.00632   Min.   :  0.00   Min.   : 0.46   Min.   :0.00000
##  1st Qu.: 0.08204   1st Qu.:  0.00   1st Qu.: 5.19   1st Qu.:0.00000
##  Median : 0.25651   Median :  0.00   Median : 9.69   Median :0.00000
##  Mean   : 3.61352   Mean   : 11.36   Mean   :11.14   Mean   :0.06917
##  3rd Qu.: 3.67708   3rd Qu.: 12.50   3rd Qu.:18.10   3rd Qu.:0.00000
##  Max.   :88.97620   Max.   :100.00   Max.   :27.74   Max.   :1.00000
##       nox               rm             age              dis
##  Min.   :0.3850   Min.   :3.561   Min.   :  2.90   Min.   : 1.130
##  1st Qu.:0.4490   1st Qu.:5.886   1st Qu.: 45.02   1st Qu.: 2.100
##  Median :0.5380   Median :6.208   Median : 77.50   Median : 3.207
##  Mean   :0.5547   Mean   :6.285   Mean   : 68.57   Mean   : 3.795
##  3rd Qu.:0.6240   3rd Qu.:6.623   3rd Qu.: 94.08   3rd Qu.: 5.188
##  Max.   :0.8710   Max.   :8.780   Max.   :100.00   Max.   :12.127
##       rad              tax           ptratio          black
##  Min.   : 1.000   Min.   :187.0   Min.   :12.60   Min.   :  0.32
##  1st Qu.: 4.000   1st Qu.:279.0   1st Qu.:17.40   1st Qu.:375.38
##  Median : 5.000   Median :330.0   Median :19.05   Median :391.44
##  Mean   : 9.549   Mean   :408.2   Mean   :18.46   Mean   :356.67
##  3rd Qu.:24.000   3rd Qu.:666.0   3rd Qu.:20.20   3rd Qu.:396.23
##  Max.   :24.000   Max.   :711.0   Max.   :22.00   Max.   :396.90
##      lstat            medv
##  Min.   : 1.73   Min.   : 5.00
##  1st Qu.: 6.95   1st Qu.:17.02
##  Median :11.36   Median :21.20
##  Mean   :12.65   Mean   :22.53
##  3rd Qu.:16.95   3rd Qu.:25.00
##  Max.   :37.97   Max.   :50.00
```

```r
fit=lm(medv~.,data=Boston)
summary(fit)
```

```
##
## Call:
## lm(formula = medv ~ ., data = Boston)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -15.595  -2.730  -0.518   1.777  26.199
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  3.646e+01  5.103e+00   7.144 3.28e-12 ***
## crim        -1.080e-01  3.286e-02  -3.287 0.001087 **
## zn           4.642e-02  1.373e-02   3.382 0.000778 ***
## indus        2.056e-02  6.150e-02   0.334 0.738288
## chas         2.687e+00  8.616e-01   3.118 0.001925 **
## nox         -1.777e+01  3.820e+00  -4.651 4.25e-06 ***
## rm           3.810e+00  4.179e-01   9.116  < 2e-16 ***
## age          6.922e-04  1.321e-02   0.052 0.958229
## dis         -1.476e+00  1.995e-01  -7.398 6.01e-13 ***
## rad          3.060e-01  6.635e-02   4.613 5.07e-06 ***
## tax         -1.233e-02  3.760e-03  -3.280 0.001112 **
## ptratio     -9.527e-01  1.308e-01  -7.283 1.31e-12 ***
## black        9.312e-03  2.686e-03   3.467 0.000573 ***
## lstat       -5.248e-01  5.072e-02 -10.347  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.745 on 492 degrees of freedom
## Multiple R-squared:  0.7406, Adjusted R-squared:  0.7338
## F-statistic: 108.1 on 13 and 492 DF,  p-value: < 2.2e-16
```

AIC criteria

```r
model.aic.backward <- step(fit, data=Boston,direction="backward", trace = 1,k =2)
```

```
## Start:  AIC=1589.64
## medv ~ crim + zn + indus + chas + nox + rm + age + dis + rad +
##     tax + ptratio + black + lstat
##
##           Df Sum of Sq   RSS    AIC
## - age      1      0.06 11079 1587.7
## - indus    1      2.52 11081 1587.8
## <none>                 11079 1589.6
## - chas     1    218.97 11298 1597.5
## - tax      1    242.26 11321 1598.6
## - crim     1    243.22 11322 1598.6
## - zn       1    257.49 11336 1599.3
## - black    1    270.63 11349 1599.8
## - rad      1    479.15 11558 1609.1
## - nox      1    487.16 11566 1609.4
```

```
## - ptratio  1   1194.23 12273 1639.4
## - dis      1   1232.41 12311 1641.0
## - rm       1   1871.32 12950 1666.6
## - lstat    1   2410.84 13490 1687.3
##
## Step:  AIC=1587.65
## medv ~ crim + zn + indus + chas + nox + rm + dis + rad + tax +
##     ptratio + black + lstat
##
##           Df Sum of Sq   RSS    AIC
## - indus    1      2.52 11081 1585.8
## <none>                 11079 1587.7
## - chas     1    219.91 11299 1595.6
## - tax      1    242.24 11321 1596.6
## - crim     1    243.20 11322 1596.6
## - zn       1    260.32 11339 1597.4
## - black    1    272.26 11351 1597.9
## - rad      1    481.09 11560 1607.2
## - nox      1    520.87 11600 1608.9
## - ptratio  1   1200.23 12279 1637.7
## - dis      1   1352.26 12431 1643.9
## - rm       1   1959.55 13038 1668.0
## - lstat    1   2718.88 13798 1696.7
##
## Step:  AIC=1585.76
## medv ~ crim + zn + chas + nox + rm + dis + rad + tax + ptratio +
##     black + lstat
##
##           Df Sum of Sq   RSS    AIC
## <none>                 11081 1585.8
## - chas     1    227.21 11309 1594.0
## - crim     1    245.37 11327 1594.8
## - zn       1    257.82 11339 1595.4
## - black    1    270.82 11352 1596.0
## - tax      1    273.62 11355 1596.1
## - rad      1    500.92 11582 1606.1
## - nox      1    541.91 11623 1607.9
## - ptratio  1   1206.45 12288 1636.0
## - dis      1   1448.94 12530 1645.9
## - rm       1   1963.66 13045 1666.3
## - lstat    1   2723.48 13805 1695.0
```

BIC model for model selection

```r
model.bic.backward <- step(fit, data=Boston, direction="backward", trace = 1, k=log(nrow(Boston)))
```

```
## Start:  AIC=1648.81
## medv ~ crim + zn + indus + chas + nox + rm + age + dis + rad +
##     tax + ptratio + black + lstat
##
##           Df Sum of Sq   RSS    AIC
## - age      1      0.06 11079 1642.6
## - indus    1      2.52 11081 1642.7
```

```
## <none>                 11079 1648.8
## - chas     1    218.97 11298 1652.5
## - tax      1    242.26 11321 1653.5
## - crim     1    243.22 11322 1653.6
## - zn       1    257.49 11336 1654.2
## - black    1    270.63 11349 1654.8
## - rad      1    479.15 11558 1664.0
## - nox      1    487.16 11566 1664.4
## - ptratio  1   1194.23 12273 1694.4
## - dis      1   1232.41 12311 1696.0
## - rm       1   1871.32 12950 1721.6
## - lstat    1   2410.84 13490 1742.2
##
## Step:  AIC=1642.59
## medv ~ crim + zn + indus + chas + nox + rm + dis + rad + tax +
##     ptratio + black + lstat
##
##           Df Sum of Sq   RSS    AIC
## - indus    1      2.52 11081 1636.5
## <none>                 11079 1642.6
## - chas     1    219.91 11299 1646.3
## - tax      1    242.24 11321 1647.3
## - crim     1    243.20 11322 1647.3
## - zn       1    260.32 11339 1648.1
## - black    1    272.26 11351 1648.7
## - rad      1    481.09 11560 1657.9
## - nox      1    520.87 11600 1659.6
## - ptratio  1   1200.23 12279 1688.4
## - dis      1   1352.26 12431 1694.6
## - rm       1   1959.55 13038 1718.8
## - lstat    1   2718.88 13798 1747.4
##
## Step:  AIC=1636.48
## medv ~ crim + zn + chas + nox + rm + dis + rad + tax + ptratio +
##     black + lstat
##
##           Df Sum of Sq   RSS    AIC
## <none>                 11081 1636.5
## - chas     1    227.21 11309 1640.5
## - crim     1    245.37 11327 1641.3
## - zn       1    257.82 11339 1641.9
## - black    1    270.82 11352 1642.5
## - tax      1    273.62 11355 1642.6
## - rad      1    500.92 11582 1652.6
## - nox      1    541.91 11623 1654.4
## - ptratio  1   1206.45 12288 1682.5
## - dis      1   1448.94 12530 1692.4
## - rm       1   1963.66 13045 1712.8
## - lstat    1   2723.48 13805 1741.5
```

(b) Now make a histogram of the response variable (use hist()) to see if it is skewed. Using log(medv) as the response variable, perform the stepwise selection as previously using both AIC and BIC criteria. Compare with the previous results in terms of selected variables and adjusted R2 .

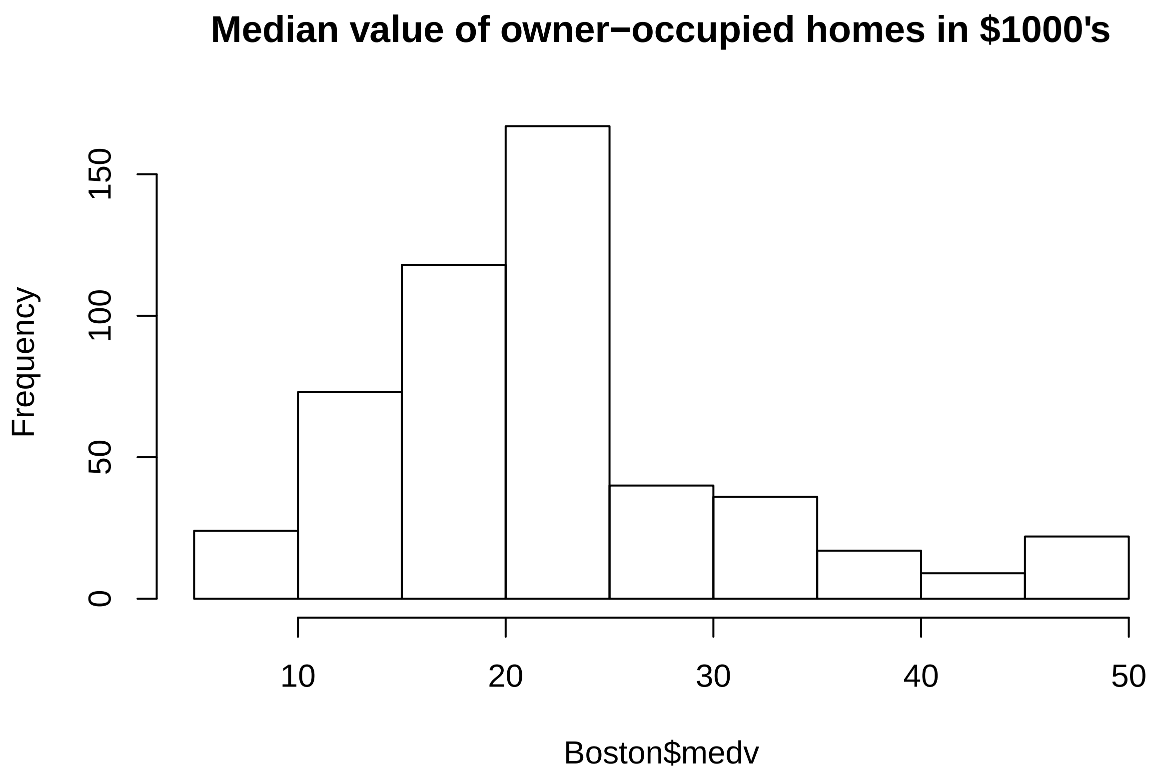
```
AIC(fit)
```

```
## [1] 3027.609
```

```
BIC(fit)
```

```
## [1] 3091.007
```

```
hist(Boston$medv, main= "Median value of owner-occupied homes in $1000's")
```



log(medv) as the response variable to model the dataset

```
fit2 <- lm(log(medv) ~ . , data = Boston)
summary(fit2)
```

```
##
## Call:
## lm(formula = log(medv) ~ ., data = Boston)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.73361 -0.09747 -0.01657  0.09629  0.86435
##
## Coefficients:
```

```
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)  4.1020423  0.2042726  20.081  < 2e-16 ***
## crim        -0.0102715  0.0013155  -7.808 3.52e-14 ***
## zn           0.0011725  0.0005495   2.134 0.033349 *
## indus        0.0024668  0.0024614   1.002 0.316755
## chas         0.1008876  0.0344859   2.925 0.003598 **
## nox         -0.7783993  0.1528902  -5.091 5.07e-07 ***
## rm           0.0908331  0.0167280   5.430 8.87e-08 ***
## age          0.0002106  0.0005287   0.398 0.690567
## dis         -0.0490873  0.0079834  -6.149 1.62e-09 ***
## rad          0.0142673  0.0026556   5.373 1.20e-07 ***
## tax         -0.0006258  0.0001505  -4.157 3.80e-05 ***
## ptratio     -0.0382715  0.0052365  -7.309 1.10e-12 ***
## black        0.0004136  0.0001075   3.847 0.000135 ***
## lstat       -0.0290355  0.0020299 -14.304  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1899 on 492 degrees of freedom
## Multiple R-squared:  0.7896, Adjusted R-squared:  0.7841
## F-statistic: 142.1 on 13 and 492 DF,  p-value: < 2.2e-16
```

Using Backward Selection AIC to model the data using log(medv) as the response

```
model.aic.backward1 <-step(fit2, data=Boston, direction="backward",trace =1,k =2)
```

```
## Start:  AIC=-1667.19
## log(medv) ~ crim + zn + indus + chas + nox + rm + age + dis +
##     rad + tax + ptratio + black + lstat
##
##           Df Sum of Sq    RSS     AIC
## - age      1    0.0057 17.755 -1669.0
## - indus    1    0.0362 17.786 -1668.2
## <none>                 17.749 -1667.2
## - zn       1    0.1643 17.914 -1664.5
## - chas     1    0.3088 18.058 -1660.5
## - black    1    0.5339 18.283 -1654.2
## - tax      1    0.6235 18.373 -1651.7
## - nox      1    0.9351 18.684 -1643.2
## - rad      1    1.0413 18.791 -1640.3
## - rm       1    1.0637 18.813 -1639.7
## - dis      1    1.3639 19.113 -1631.7
## - ptratio  1    1.9270 19.676 -1617.0
## - crim     1    2.1995 19.949 -1610.1
## - lstat    1    7.3809 25.130 -1493.2
##
## Step:  AIC=-1669.03
## log(medv) ~ crim + zn + indus + chas + nox + rm + dis + rad +
##     tax + ptratio + black + lstat
##
##           Df Sum of Sq    RSS     AIC
## - indus    1    0.0363 17.791 -1670.0
```

```
## <none>                 17.755 -1669.0
## - zn       1    0.1593 17.914 -1666.5
## - chas     1    0.3138 18.069 -1662.2
## - black    1    0.5431 18.298 -1655.8
## - tax      1    0.6205 18.376 -1653.7
## - nox      1    0.9645 18.720 -1644.3
## - rad      1    1.0356 18.791 -1642.3
## - rm       1    1.1452 18.900 -1639.4
## - dis      1    1.5471 19.302 -1628.8
## - ptratio  1    1.9224 19.677 -1619.0
## - crim     1    2.1988 19.954 -1612.0
## - lstat    1    8.1949 25.950 -1479.0
##
## Step:  AIC=-1670
## log(medv) ~ crim + zn + chas + nox + rm + dis + rad + tax + ptratio +
##     black + lstat
##
##           Df Sum of Sq    RSS     AIC
## <none>                 17.791 -1670.0
## - zn       1    0.1451 17.936 -1667.9
## - chas     1    0.3399 18.131 -1662.4
## - black    1    0.5344 18.326 -1657.0
## - tax      1    0.6139 18.405 -1654.8
## - nox      1    0.9350 18.726 -1646.1
## - rad      1    1.0088 18.800 -1644.1
## - rm       1    1.1171 18.909 -1641.2
## - dis      1    1.7385 19.530 -1624.8
## - ptratio  1    1.8862 19.678 -1621.0
## - crim     1    2.2229 20.014 -1612.4
## - lstat    1    8.1604 25.952 -1481.0
```

Using Backward Selection BIC model with log(medv) as the response

```r
model.bic.backward <- step(fit2, data=Boston, direction="backward", trace = 1,    k=log(nrow(Boston)))
```

```
## Start:  AIC=-1608.02
## log(medv) ~ crim + zn + indus + chas + nox + rm + age + dis +
##     rad + tax + ptratio + black + lstat
##
##           Df Sum of Sq    RSS     AIC
## - age      1    0.0057 17.755 -1614.1
## - indus    1    0.0362 17.786 -1613.2
## - zn       1    0.1643 17.914 -1609.6
## <none>                 17.749 -1608.0
## - chas     1    0.3088 18.058 -1605.5
## - black    1    0.5339 18.283 -1599.2
## - tax      1    0.6235 18.373 -1596.8
## - nox      1    0.9351 18.684 -1588.3
## - rad      1    1.0413 18.791 -1585.4
## - rm       1    1.0637 18.813 -1584.8
## - dis      1    1.3639 19.113 -1576.8
## - ptratio  1    1.9270 19.676 -1562.1
## - crim     1    2.1995 19.949 -1555.1
```

```
## - lstat    1    7.3809 25.130 -1438.3
##
## Step:  AIC=-1614.09
## log(medv) ~ crim + zn + indus + chas + nox + rm + dis + rad +
##     tax + ptratio + black + lstat
##
##           Df Sum of Sq    RSS     AIC
## - indus    1    0.0363 17.791 -1619.3
## - zn       1    0.1593 17.914 -1615.8
## <none>                 17.755 -1614.1
## - chas     1    0.3138 18.069 -1611.5
## - black    1    0.5431 18.298 -1605.1
## - tax      1    0.6205 18.376 -1602.9
## - nox      1    0.9645 18.720 -1593.5
## - rad      1    1.0356 18.791 -1591.6
## - rm       1    1.1452 18.900 -1588.7
## - dis      1    1.5471 19.302 -1578.0
## - ptratio  1    1.9224 19.677 -1568.3
## - crim     1    2.1988 19.954 -1561.2
## - lstat    1    8.1949 25.950 -1428.3
##
## Step:  AIC=-1619.28
## log(medv) ~ crim + zn + chas + nox + rm + dis + rad + tax + ptratio +
##     black + lstat
##
##           Df Sum of Sq    RSS     AIC
## - zn       1    0.1451 17.936 -1621.4
## <none>                 17.791 -1619.3
## - chas     1    0.3399 18.131 -1615.9
## - black    1    0.5344 18.326 -1610.5
## - tax      1    0.6139 18.405 -1608.3
## - nox      1    0.9350 18.726 -1599.6
## - rad      1    1.0088 18.800 -1597.6
## - rm       1    1.1171 18.909 -1594.7
## - dis      1    1.7385 19.530 -1578.3
## - ptratio  1    1.8862 19.678 -1574.5
## - crim     1    2.2229 20.014 -1565.9
## - lstat    1    8.1604 25.952 -1434.5
##
## Step:  AIC=-1621.4
## log(medv) ~ crim + chas + nox + rm + dis + rad + tax + ptratio +
##     black + lstat
##
##           Df Sum of Sq    RSS     AIC
## <none>                 17.936 -1621.4
## - chas     1    0.3388 18.275 -1618.2
## - tax      1    0.5229 18.459 -1613.1
## - black    1    0.5386 18.475 -1612.7
## - rad      1    0.9601 18.897 -1601.2
## - nox      1    1.0250 18.961 -1599.5
## - rm       1    1.2650 19.201 -1593.1
## - dis      1    1.6967 19.633 -1581.9
## - crim     1    2.1377 20.074 -1570.7
## - ptratio  1    2.5632 20.500 -1560.0
```

```
## - lstat    1    8.1516 26.088 -1438.1
```

```r
AIC(fit2)
```

```
## [1] -229.2284
```

```r
BIC(fit2)
```

```
## [1] -165.8304
```

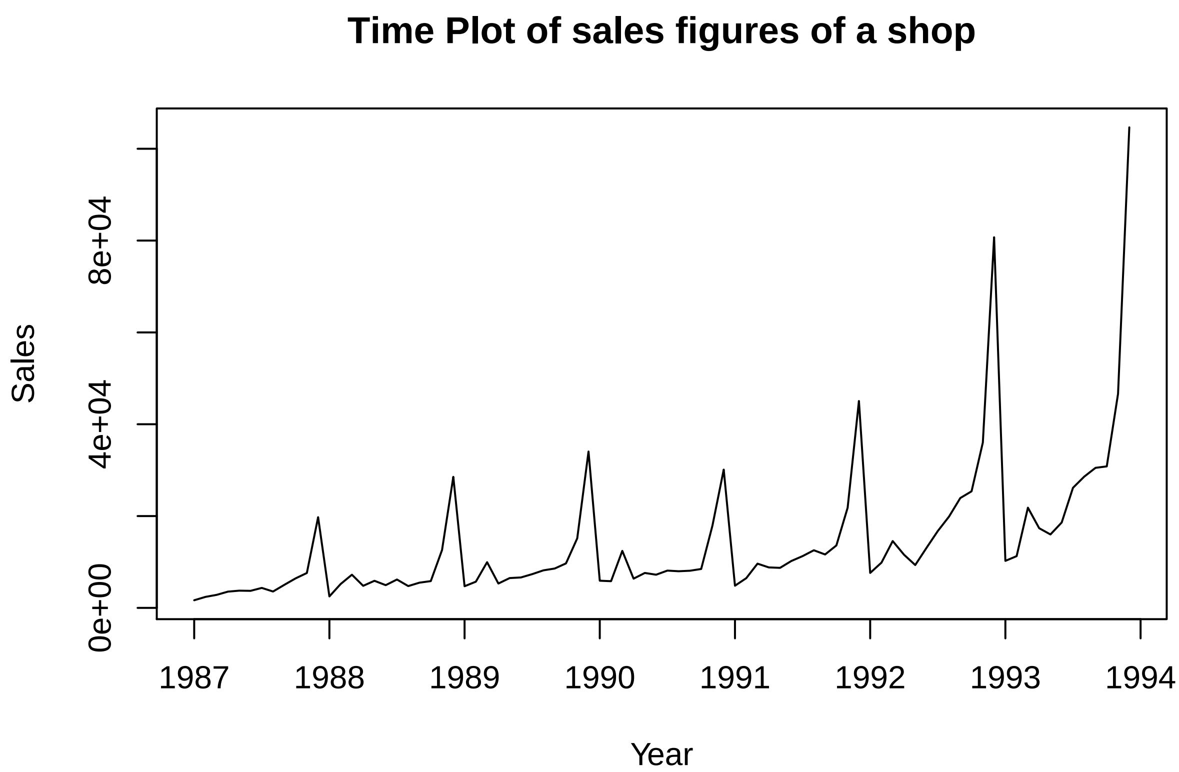
Question 2. The data set fancy (you need to library the fpp package to get the dataset) concerns the monthly sales figures of a shop which opened in January 1987 and sells gifts, souvenirs, and novelties. The sales volume varies with the seasonal population of tourists.

(a) Produce a time plot of the data and describe the patterns in the graph. Identify any unusual or unexpected fluctuations in the time series.

```r
library(fpp)
```

```
## Loading required package: forecast

## Warning in as.POSIXlt.POSIXct(Sys.time()): unknown timezone 'zone/tz/2018c.
## 1.0/zoneinfo/America/New_York'

## Loading required package: fma

##
## Attaching package: 'fma'

## The following objects are masked from 'package:MASS':
##
##     cement, housing, petrol

## Loading required package: expsmooth

## Loading required package: lmtest

## Loading required package: zoo

## Warning: package 'zoo' was built under R version 3.4.3

##
## Attaching package: 'zoo'

## The following objects are masked from 'package:base':
##
##     as.Date, as.Date.numeric

## Loading required package: tseries

## Warning: package 'tseries' was built under R version 3.4.3
```

```r
plot.ts(fancy, main = "Time Plot of sales figures of a shop",xlab = "Year",ylab="Sales")
```



(b) Use R function tslm to fit a regression model to the logarithms of these sales data with a linear trend and seasonal component.

```r
fit_fancy <- tslm(log(fancy) ~ trend + season)
summary(fit_fancy)
```

```
##
## Call:
## tslm(formula = log(fancy) ~ trend + season)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.41644 -0.12619  0.00608  0.11389  0.38567
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 7.6058604  0.0768740  98.939  < 2e-16 ***
## trend       0.0223930  0.0008448  26.508  < 2e-16 ***
## season2     0.2510437  0.0993278   2.527 0.013718 *
## season3     0.6952066  0.0993386   6.998 1.18e-09 ***
## season4     0.3829341  0.0993565   3.854 0.000252 ***
## season5     0.4079944  0.0993817   4.105 0.000106 ***
## season6     0.4469625  0.0994140   4.496 2.63e-05 ***
## season7     0.6082156  0.0994534   6.116 4.69e-08 ***
## season8     0.5853524  0.0995001   5.883 1.21e-07 ***
## season9     0.6663446  0.0995538   6.693 4.27e-09 ***
## season10    0.7440336  0.0996148   7.469 1.61e-10 ***
## season11    1.2030164  0.0996828  12.068  < 2e-16 ***
```

```
## season12    1.9581366  0.0997579  19.629  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1858 on 71 degrees of freedom
## Multiple R-squared:  0.9527, Adjusted R-squared:  0.9447
## F-statistic: 119.1 on 12 and 71 DF,  p-value: < 2.2e-16
```

(c) Use multiple regression with trend variable and seasonal dummy variables to redo the regression as shown in the lecture example. Check to see that you obtain the same results as tslm.

```r
fit_fancy2 <- lm(log(fancy) ~., data=fancy)
summary(fit_fancy2)
```

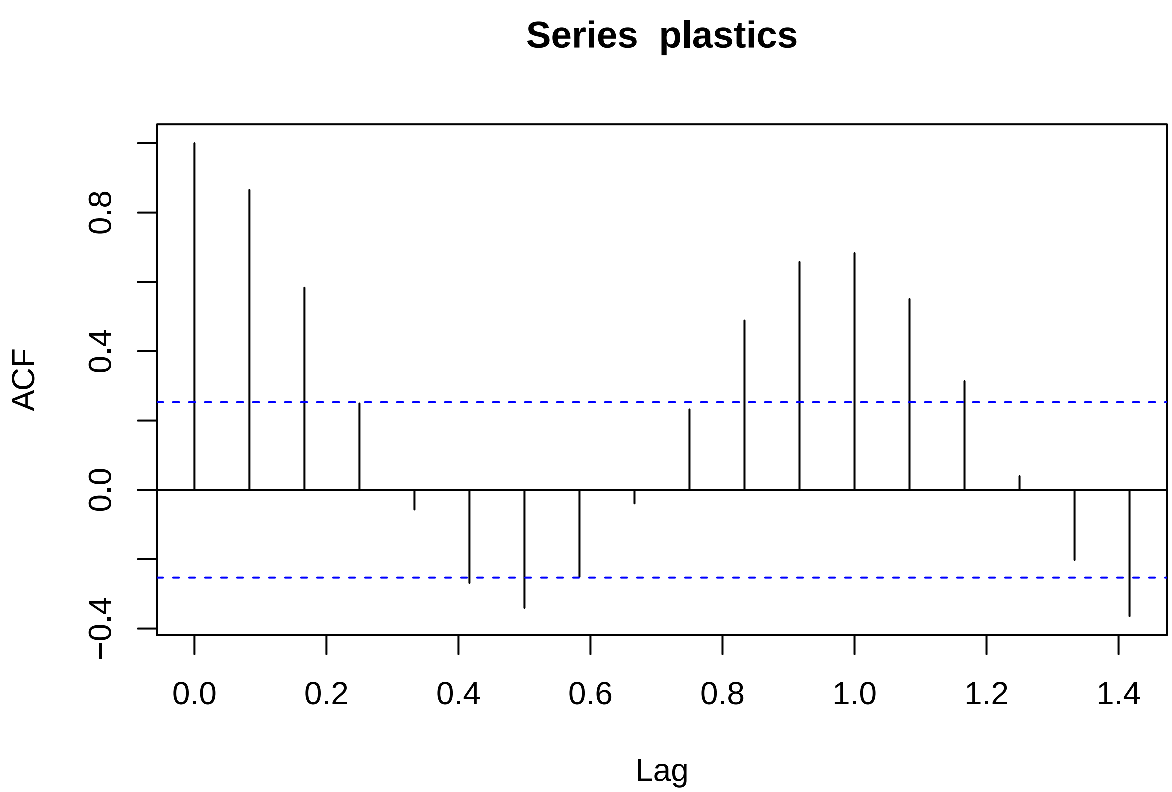
```
##
## Call:
## lm(formula = log(fancy) ~ ., data = fancy)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.53705 -0.19087  0.07914  0.33173  0.45046
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 8.606e+00  6.094e-02  141.21   <2e-16 ***
## x           4.290e-05  2.873e-06   14.93   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4122 on 82 degrees of freedom
## Multiple R-squared:  0.7311, Adjusted R-squared:  0.7278
## F-statistic: 222.9 on 1 and 82 DF,  p-value: < 2.2e-16
```

Question 3. The data set plastics (you need to library the fpp package to get the dataset) represents the monthly sales (in thousands) of product A for a plastics manufacturer for years 1 through 5 (data set plastics). You need to load fpp package to have this dataset.

(a) Plot the time series of sales of product A. Can you identify seasonal fluctuations or a trend?

```r
library(fpp)
plot(plastics, main = "Monthly sales of product A(in thousands)")
```


```
acf(plastics)
```

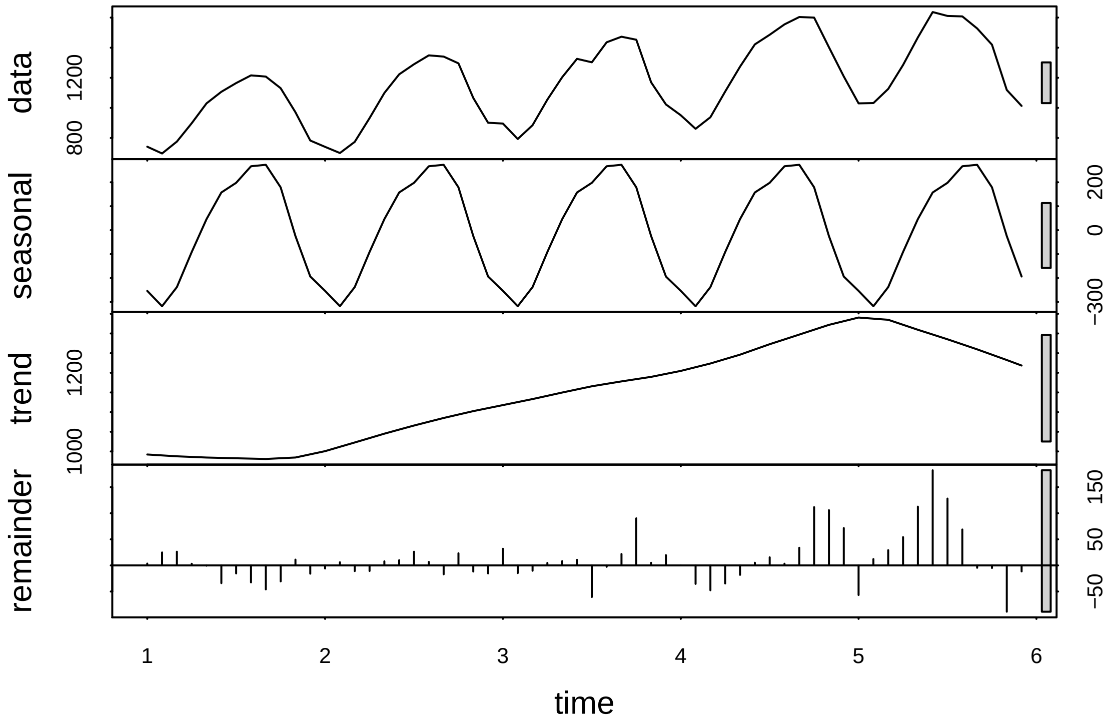


##(b)

Perform a classical additive decomposition using" stl" function. Plot out the decomposition. Try four combinations: (1) s.window="periodic", (2) s.window="periodic", t.window=5, (3) s.window="periodic", t.window=50 and (4) s.window=5, t.window=50. Explain the differences you see in the plots.

Case 1: s.window="periodic"

```
fit_plastic1 <- stl(plastics, s.window="periodic", robust=TRUE )
plot(fit_plastic1)
```



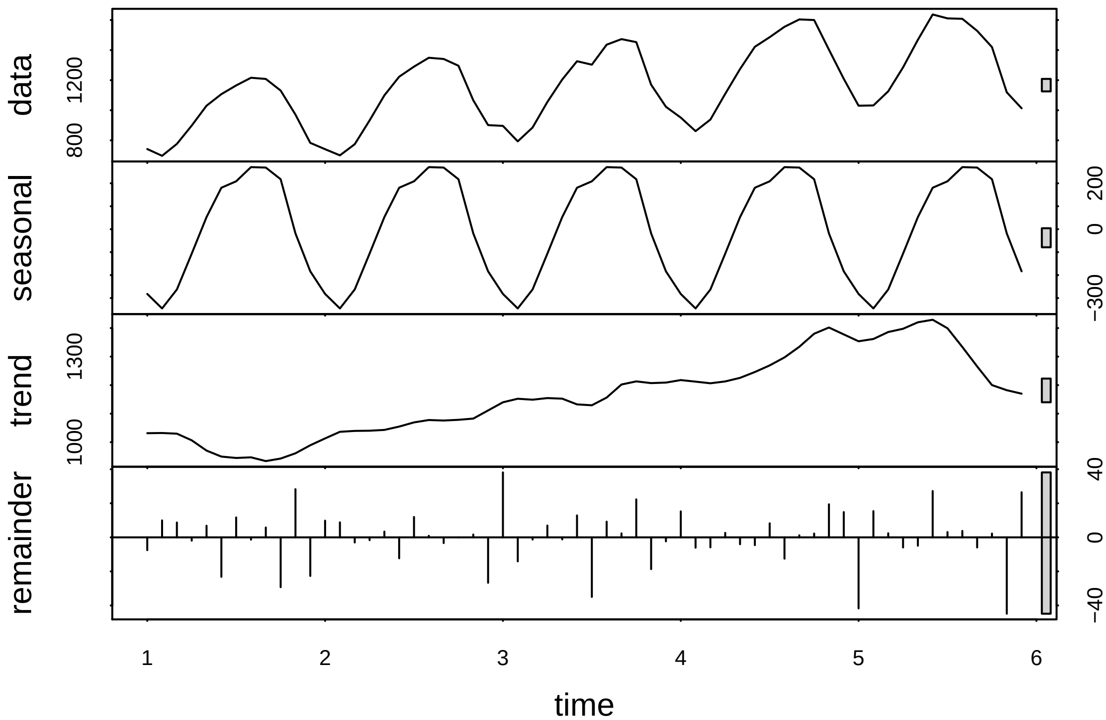

```
summary(fit_plastic1)
```

```
##  Call:
##  stl(x = plastics, s.window = "periodic", robust = TRUE)
##
##  Time.series components:
##     seasonal              trend               remainder
##  Min.   :-317.8402   Min.   : 980.8963   Min.   :-88.67139
##  1st Qu.:-204.8979   1st Qu.:1031.2002   1st Qu.:-14.83808
##  Median :  10.5747   Median :1161.7531   Median :  4.19361
##  Mean   :   0.0000   Mean   :1151.6866   Mean   : 10.68007
##  3rd Qu.: 183.5807   3rd Qu.:1259.4471   3rd Qu.: 23.71173
##  Max.   : 272.9791   Max.   :1340.5082   Max.   :182.28778
##  IQR:
##      STL.seasonal STL.trend STL.remainder data
##      388.48       228.25     38.55         414.75
##    %  93.7          55.0       9.3          100.0
##
##  Weights:
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0000  0.7722  0.9452  0.7726  0.9873  1.0000
##
##  Other components: List of 5
##  $ win  : Named num [1:3] 601 19 13
##  $ deg  : Named int [1:3] 0 1 1
##  $ jump : Named num [1:3] 61 2 2
```

```
##   $ inner: int 1
##   $ outer: int 15
```

Case 2: s.window="periodic", t.window=5

```
fit_plastic2 <- stl(plastics,  s.window="periodic", t.window=5)
plot(fit_plastic2)
```



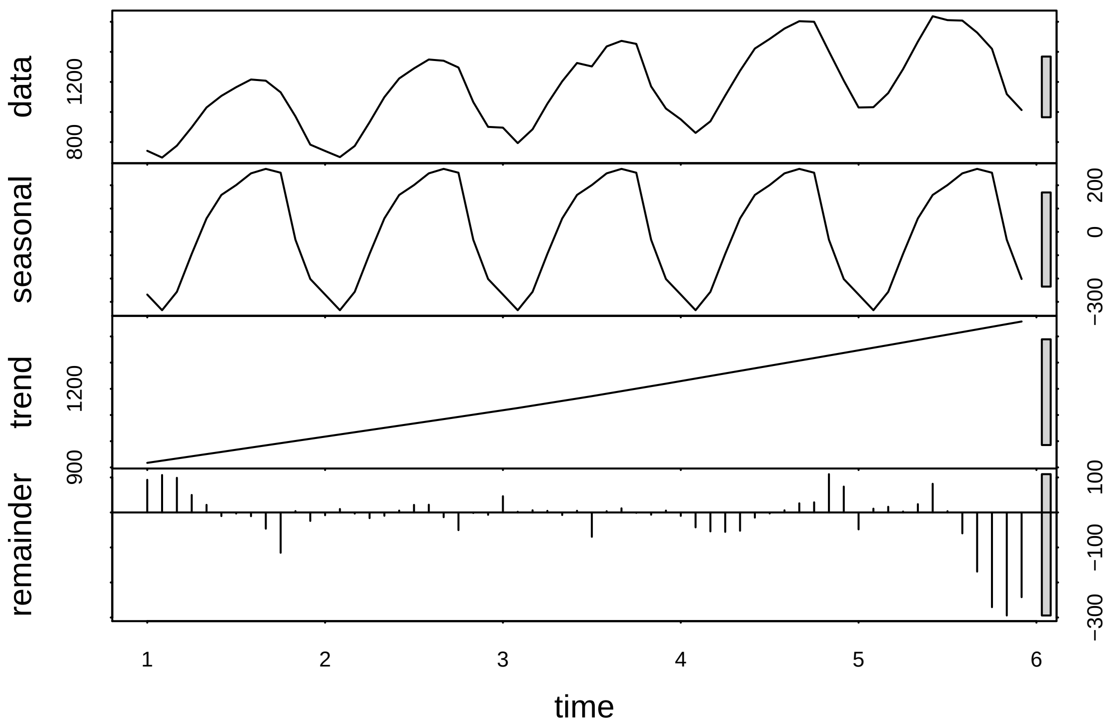

```
summary(fit_plastic2)
```

```
##   Call:
##   stl(x = plastics, s.window = "periodic", t.window = 5)
##
##   Time.series components:
##      seasonal             trend              remainder
##   Min.    :-345.1684   Min.    : 933.8063   Min.    :-44.84399
##   1st Qu.:-203.3249   1st Qu.:1040.0924   1st Qu.: -5.90772
##   Median :  17.1211   Median :1153.6267   Median :  1.42814
##   Mean   :   0.0000   Mean   :1162.3395   Mean   :  0.02721
##   3rd Qu.: 210.8602   3rd Qu.:1250.8643   3rd Qu.:  8.91246
##   Max.   : 270.4102   Max.   :1429.1903   Max.   : 38.14019
##   IQR:
##       STL.seasonal STL.trend STL.remainder data
##       414.19       210.77    14.82         414.75
##     %  99.9         50.8      3.6          100.0
##
##   Weights: all == 1
##
##   Other components: List of 5
##   $ win  : Named num [1:3] 601 5 13
```

```
##  $ deg  : Named int [1:3] 0 1 1
##  $ jump : Named num [1:3] 61 1 2
##  $ inner: int 2
##  $ outer: int 0
```

Case 3: s.window="periodic", t.window=50,

```
fit_plastic3 <-  stl(plastics,  s.window="periodic", t.window=50, robust=TRUE)
plot(fit_plastic3)
```



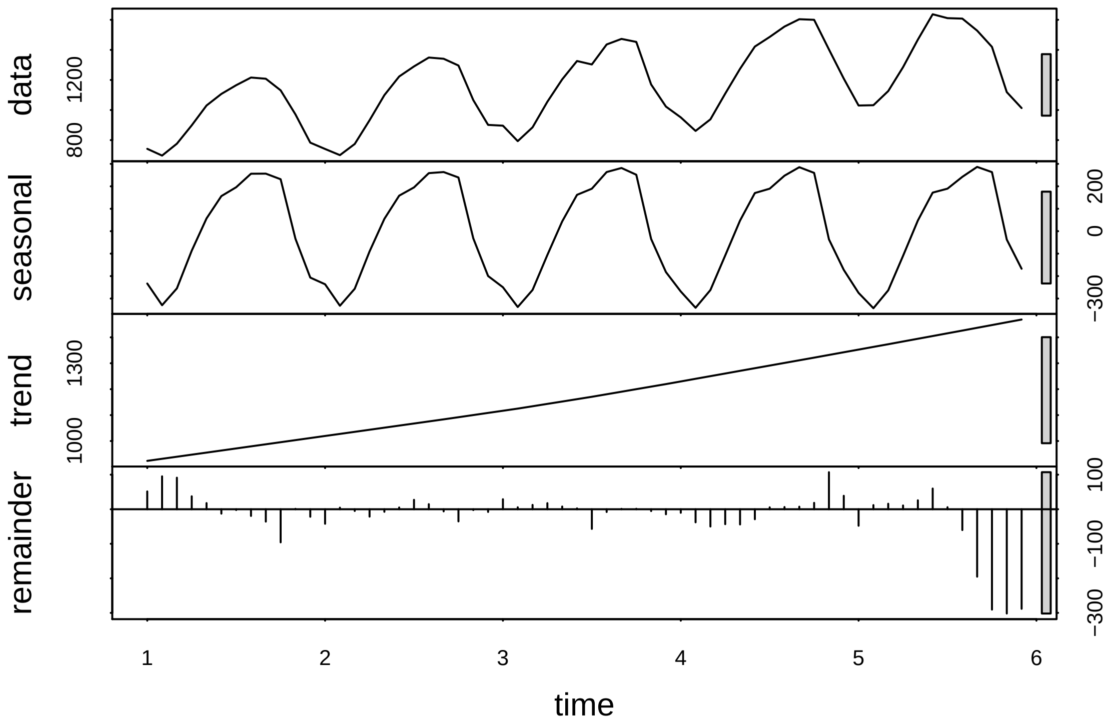

```
summary(fit_plastic3)
```

```
##  Call:
##  stl(x = plastics, s.window = "periodic", t.window = 50, robust = TRUE)
##
##  Time.series components:
##     seasonal             trend            remainder
##  Min.   :-335.7097   Min.   : 917.409   Min.   :-294.21654
##  1st Qu.:-215.7585   1st Qu.:1040.479   1st Qu.: -18.18613
##  Median :  11.9464   Median :1167.370   Median :  -1.39773
##  Mean   :   0.0000   Mean   :1175.977   Mean   : -13.61019
##  3rd Qu.: 213.2684   3rd Qu.:1310.075   3rd Qu.:  12.92090
##  Max.   : 270.1984   Max.   :1456.967   Max.   : 109.42339
##  IQR:
##      STL.seasonal STL.trend STL.remainder data
##      429.03       269.60     31.11         414.75
##    % 103.4         65.0       7.5          100.0
##
##  Weights:
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
```

```
## 0.0000 0.4998 0.9452 0.7128 0.9932 1.0000
##
## Other components: List of 5
## $ win  : Named num [1:3] 601 50 13
## $ deg  : Named int [1:3] 0 1 1
## $ jump : Named num [1:3] 61 5 2
## $ inner: int 1
## $ outer: int 15
```

Case 4: s.window=5, t.window=50

```
fit_plastic4 <-  stl(plastics,  s.window=5, t.window=50, robust=TRUE)
plot(fit_plastic4)
```



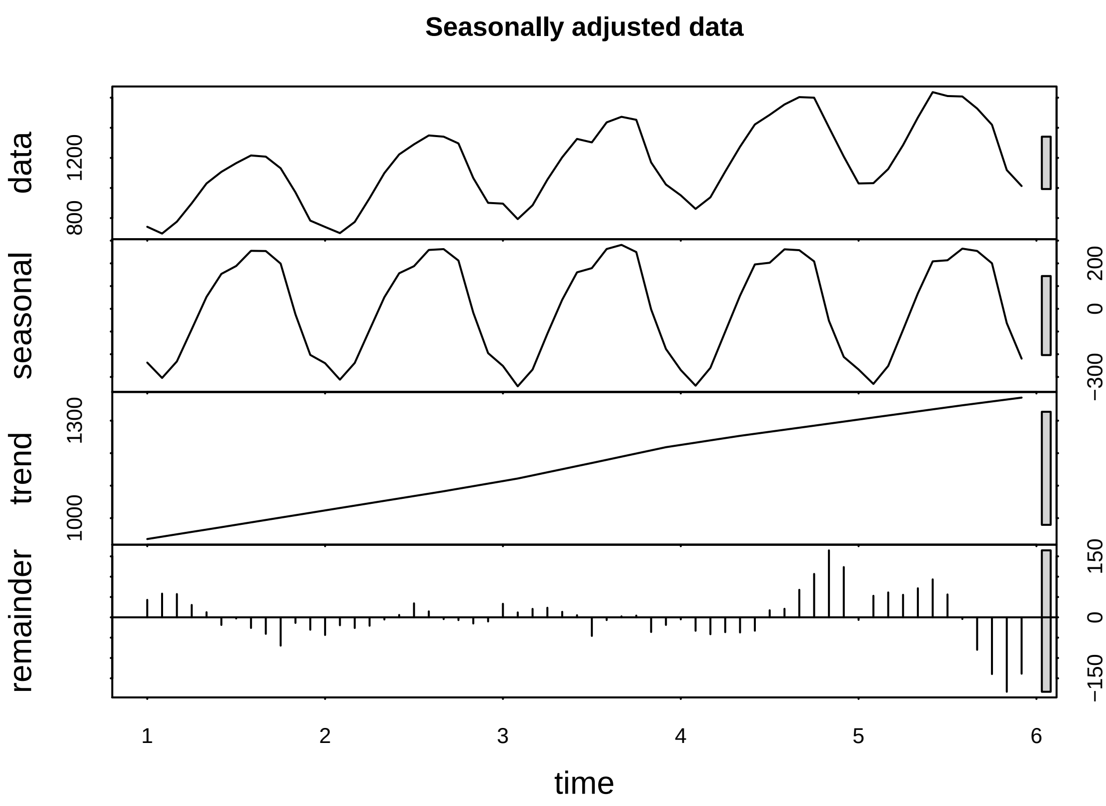

```
summary(fit_plastic4)
```

```
## Call:
## stl(x = plastics, s.window = 5, t.window = 50, robust = TRUE)
##
## Time.series components:
##    seasonal             trend               remainder
## Min.   :-343.1171   Min.   : 923.4691   Min.   :-301.81367
## 1st Qu.:-213.2872   1st Qu.:1041.4928   1st Qu.: -30.50249
## Median :   5.9745   Median :1166.0301   Median :  -0.54606
## Mean   :   1.1227   Mean   :1178.8929   Mean   : -17.64900
## 3rd Qu.: 204.7696   3rd Qu.:1314.0207   3rd Qu.:  13.16553
## Max.   : 286.5406   Max.   :1468.6791   Max.   : 107.32375
## IQR:
##     STL.seasonal STL.trend STL.remainder data
##     418.06       272.53     43.67         414.75
```

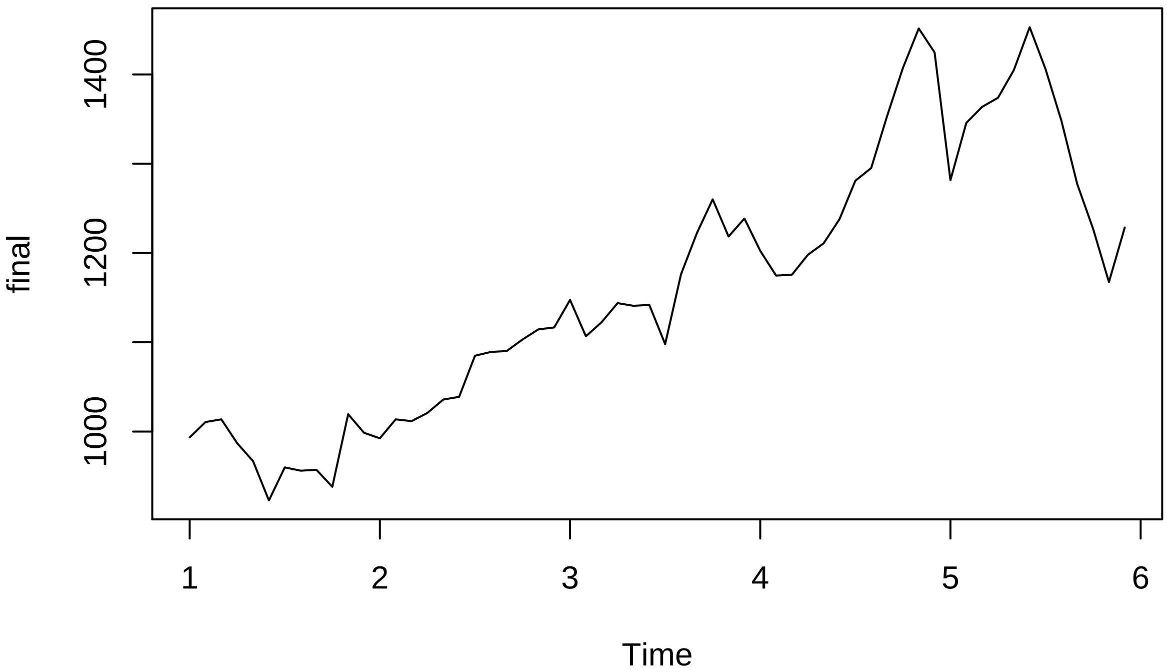
```
##    % 100.8         65.7      10.5         100.0
##
##  Weights:
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0000  0.7053  0.9452  0.7725  0.9914  0.9998
##
##  Other components: List of 5
##  $ win  : Named num [1:3] 5 50 13
##  $ deg  : Named int [1:3] 0 1 1
##  $ jump : Named num [1:3] 1 5 2
##  $ inner: int 1
##  $ outer: int 15
```

(c) Compute and plot the seasonally adjusted data. You need to do this only for case of s.window="periodic",t.window=50.

```
plot(stl(plastics,s.window=5,t.window=50), main ="Seasonally adjusted data")
```




```
bbb=stl(plastics,s.window="periodic",t.window=50)
final=plastics-bbb$time.series[,1]
plot(final)
```



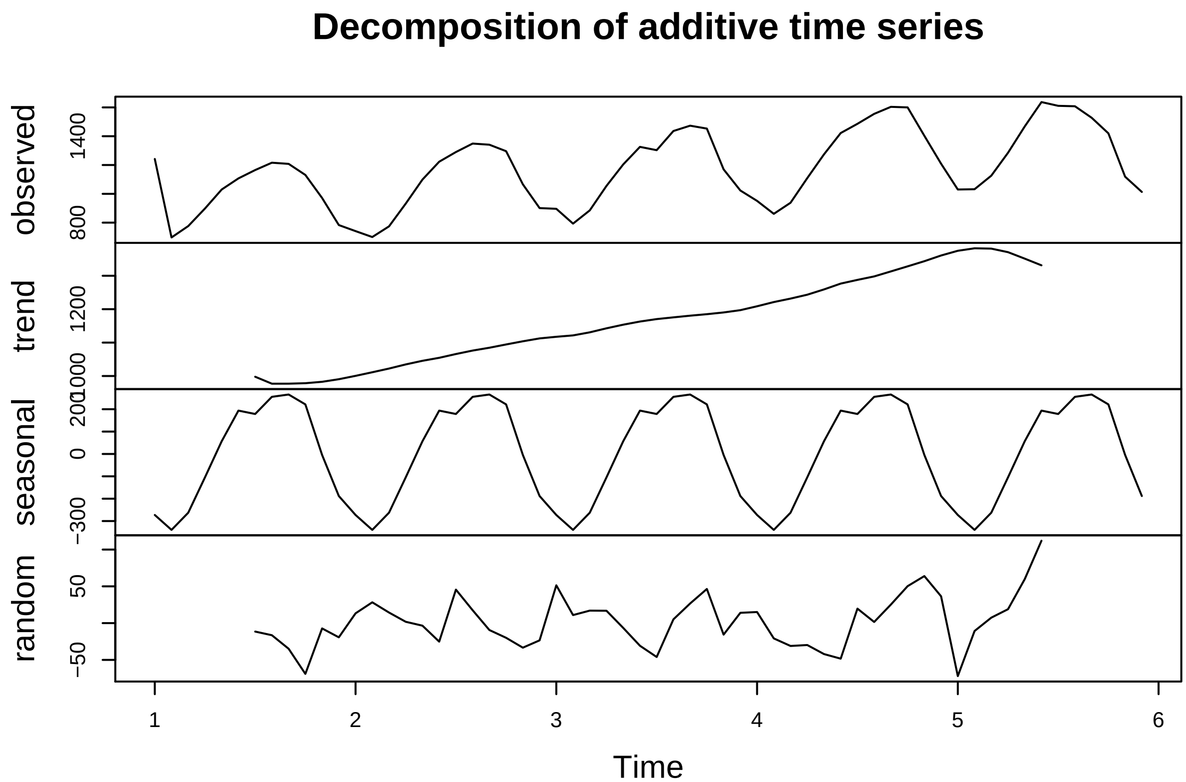
(d) Change one observation to be an outlier (pick one data point and add 500 to its value. For instance, if you picked July of the third year, the current value is 1303, then the modified value will be 1803) and recompute the seasonally adjusted data. What is the effect of the outlier. Again, you need to do this only for case of s.window="periodic",t.window=50. Does it make any difference if the outlier is near the end rather than in the middle of the time series? Try it out. 2

let 1st point be the outlier. The value of the 1st point is 742. After adding 500, we get 1242.

```
plastics[1]
```

```
## [1] 742
```

```
outlier1 <- plastics
outlier1[1] = outlier1[1] + 500

fit_plastic5 <- decompose(outlier1, type = 'additive')

plot(fit_plastic5)
```




```r
plastics[45]
```

```
## [1] 1604
```

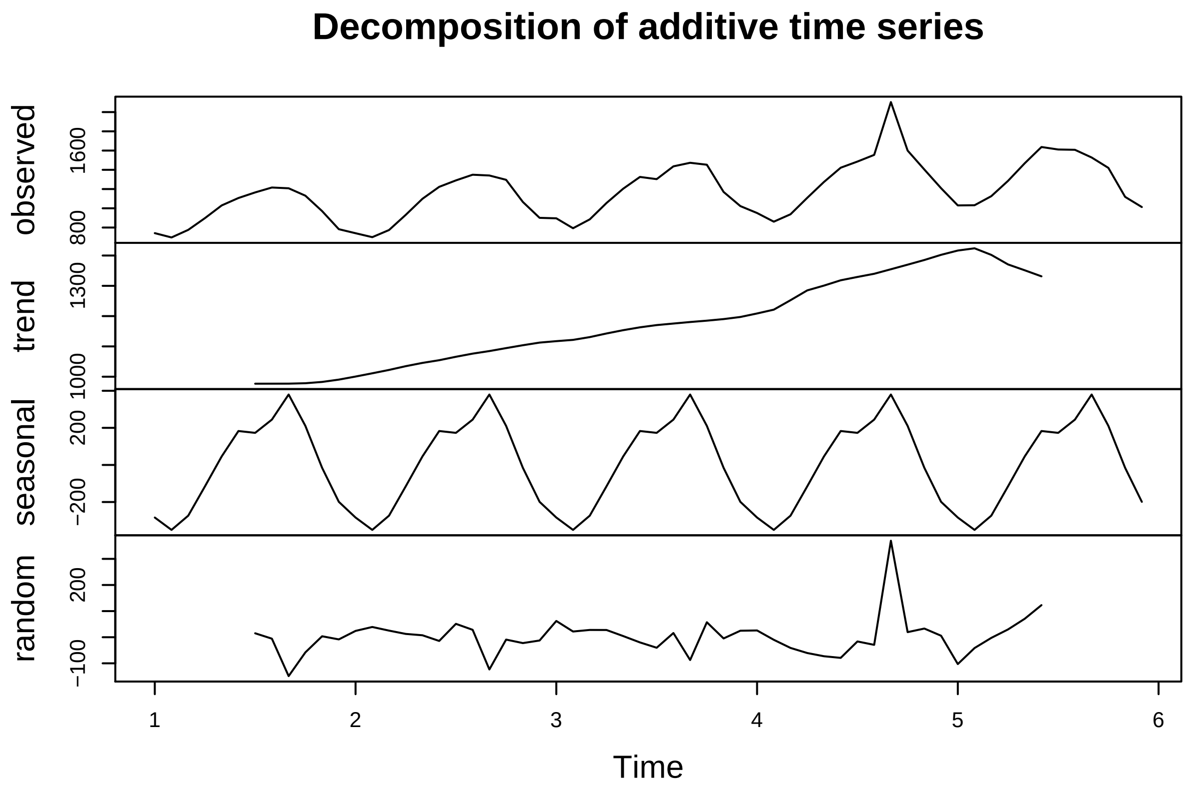
```r
outlier2 <- plastics
outlier2[45] = outlier1[45] + 500
outlier2
```

```
##    Jan  Feb  Mar  Apr  May  Jun  Jul  Aug  Sep  Oct  Nov  Dec
## 1  742  697  776  898 1030 1107 1165 1216 1208 1131  971  783
## 2  741  700  774  932 1099 1223 1290 1349 1341 1296 1066  901
## 3  896  793  885 1055 1204 1326 1303 1436 1473 1453 1170 1023
## 4  951  861  938 1109 1274 1422 1486 1555 2104 1600 1403 1209
## 5 1030 1032 1126 1285 1468 1637 1611 1608 1528 1420 1119 1013
```

```r
outlier2[45]
```

```
## [1] 2104
```

```r
fit_plastic6 <- decompose(outlier2, type = 'additive')
plot(fit_plastic6)
```


Decomposition of additive time series
observed
800
1600
trend
1000
1300
seasonal
−200
200
random
−100
200
1
2
3
4
5
6
Time