Marketing Mix Models

Abhilash K R

2 June 2017

Implementation of MMM in R:

The data we consider here is the dataset used in Chpater 9 of the book 'Cutting-Edge Marketing Analytics: Real World Cases and Data Sets for Hands On Learning'. The data set is available to download from www.dmanalytics.org.

Importing data

```
library(xlsx)
```

```
## Loading required package: rJava
```

```
## Loading required package: xlsxjars
```

```
data = read.xlsx('MMM.xlsx',sheetIndex = 1)
head(data)
```

```
##   BrandName Brand.ID Year Absolut Aristocrat Barton Belvedere Burnett
## 1   Absolut       15 1995       1          0      0         0       0
## 2   Absolut       15 1996       1          0      0         0       0
## 3   Absolut       15 1997       1          0      0         0       0
## 4   Absolut       15 1998       1          0      0         0       0
## 5   Absolut       15 1999       1          0      0         0       0
## 6   Absolut       15 2000       1          0      0         0       0
##   Chopin Crystal.Palac Finlandia Fleischmann.s Fris Gilbey.s Gordon.s
## 1      0             0         0             0    0        0        0
## 2      0             0         0             0    0        0        0
## 3      0             0         0             0    0        0        0
## 4      0             0         0             0    0        0        0
## 5      0             0         0             0    0        0        0
## 6      0             0         0             0    0        0        0
##   Grey.Goose Kamchatka Ketel.One Level McCormick Polar.Ice Popov Pravda
## 1          0         0         0     0         0         0     0      0
## 2          0         0         0     0         0         0     0      0
## 3          0         0         0     0         0         0     0      0
## 4          0         0         0     0         0         0     0      0
## 5          0         0         0     0         0         0     0      0
## 6          0         0         0     0         0         0     0      0
##   Seagram.s Skol Sky Smirnoff Stolicnaya Tanqueray Three.Olives TotalSales
## 1         0    0   0        0          0         0            0       3000
## 2         0    0   0        0          0         0            0       3340
## 3         0    0   0        0          0         0            0       3440
## 4         0    0   0        0          0         0            0       3630
## 5         0    0   0        0          0         0            0       4050
## 6         0    0   0        0          0         0            0       4605
##   LagTotalSales X2LagTotalSales  LnSales LnLSales Ln2Lsales    LnDiff
## 1          2905            2790 8.006368 7.974189  7.933797 0.0321789
## 2          3000            2905 8.113726 8.006368  7.974189 0.1073580
## 3          3340            3000 8.143227 8.113726  8.006368 0.0295010
## 4          3440            3340 8.196988 8.143227  8.113726 0.0537615
```

```
## 5          3630            3440 8.306472 8.196988  8.143227 0.1094837
## 6          4050            3630 8.434898 8.306472  8.196988 0.1284266
##         diff IfDom DollarSales PriceRerUnit LagPrice  LnPrice LnLPrice
## 1 0.03270224     0      352040     117.3467 109.2943 4.765132 4.694044
## 2 0.11333333     0      370072     110.8000 117.3467 4.707727 4.765133
## 3 0.02994012     0      381152     110.8000 110.8000 4.707727 4.707727
## 4 0.05523256     0      419737     115.6300 110.8000 4.750396 4.707727
## 5 0.11570248     0      468302     115.6301 115.6300 4.750397 4.750395
## 6 0.13703704     0      560025     121.6124 115.6301 4.800839 4.750396
##       Mag  News Outdoor Broad   Print    LnMag   LnNews    LnOut  LnBroad
## 1 25713.7 291.4   332.7   1.0 26005.1  0.00000 0.000000 0.000010 0.000010
## 2 25658.0 177.6   625.0   1.0 25835.6 10.15261 5.325933 6.384520 0.000010
## 3 27013.9 201.7   515.7   1.0 27215.6 10.20411 5.306781 6.245525 0.000010
## 4 27617.0 308.2  1420.4 161.9 27925.2 10.22619 5.730749 7.258694 5.086979
## 5 30605.3 258.1  1416.8  70.3 30863.4 10.32893 5.553347 7.256156 4.252772
## 6 33971.3 533.4  1336.7 143.1 34504.7 10.43327 6.279272 7.197959 4.963544
##    LnPrint Tier1 Tier2 TotalMinusSales LagTotalMinusSales TierSales
## 1 10.16605     1     0           60840              60903      1080
## 2 10.15951     1     0           60532              60840      1255
## 3 10.21155     1     0           60464              60532      1532
## 4 10.23728     1     0           60306              60464      1730
## 5 10.33733     1     0           59918              60306      2118
## 6 10.44885     1     0           59395              59918      2549
##   OutsideTierSales LagTierSales LagOutsideTierSales Firstintro Marketshare
## 1            59760         1035               59868          0   0.1339226
## 2            59277         1080               59760          0   0.1420068
## 3            58932         1255               59277          0   0.1467577
## 4            58576         1532               58932          0   0.1461059
## 5            57800         1730               58576          0   0.1612839
## 6            56846         2118               57800          0   0.1733680
##   LagMktshare YearID total.ad NA. NA..1 NA..2 NA..3 NA..4 NA..5 NA..6
## 1   0.1242409      3  52343.9  NA    NA    NA    NA    NA    NA    NA
## 2   0.1339226      4  52297.2  NA    NA    NA    NA    NA    NA    NA
## 3   0.1420068      5  54947.9  NA    NA    NA    NA    NA    NA    NA
## 4   0.1467577      6  57432.7  NA    NA    NA    NA    NA    NA    NA
## 5   0.1461059      7  63213.9  NA    NA    NA    NA    NA    NA    NA
## 6   0.1612839      8  70489.2  NA    NA    NA    NA    NA    NA    NA
```

```
length(unique(data$BrandName))
```

```
## [1] 27
```

```
attach(data)
```

This data set includes a total of 27 brands of vodka manufacturing companies. For MMM, let's choose a brand and analyse the impact of Price on sales. For instance, let's choose 'Finlandia' as the brand for our analysis.

```
Finlandia = subset(data,BrandName == 'Finlandia')
dim(Finlandia)
```

```
## [1] 13 74
```

```
print(Finlandia)
```

```
##    BrandName Brand.ID Year Absolut Aristocrat Barton Belvedere Burnett
## 78 Finlandia       21 1995       0          0      0         0       0
```

```
## 79 Finlandia       21 1996       0          0      0         0       0
## 80 Finlandia       21 1997       0          0      0         0       0
## 81 Finlandia       21 1998       0          0      0         0       0
## 82 Finlandia       21 1999       0          0      0         0       0
## 83 Finlandia       21 2000       0          0      0         0       0
## 84 Finlandia       21 2001       0          0      0         0       0
## 85 Finlandia       21 2002       0          0      0         0       0
## 86 Finlandia       21 2003       0          0      0         0       0
## 87 Finlandia       21 2004       0          0      0         0       0
## 88 Finlandia       21 2005       0          0      0         0       0
## 89 Finlandia       21 2006       0          0      0         0       0
## 90 Finlandia       21 2007       0          0      0         0       0
##    Chopin Crystal.Palac Finlandia Fleischmann.s Fris Gilbey.s Gordon.s
## 78      0             0         1             0    0        0        0
## 79      0             0         1             0    0        0        0
## 80      0             0         1             0    0        0        0
## 81      0             0         1             0    0        0        0
## 82      0             0         1             0    0        0        0
## 83      0             0         1             0    0        0        0
## 84      0             0         1             0    0        0        0
## 85      0             0         1             0    0        0        0
## 86      0             0         1             0    0        0        0
## 87      0             0         1             0    0        0        0
## 88      0             0         1             0    0        0        0
## 89      0             0         1             0    0        0        0
## 90      0             0         1             0    0        0        0
##    Grey.Goose Kamchatka Ketel.One Level McCormick Polar.Ice Popov Pravda
## 78          0         0         0     0         0         0     0      0
## 79          0         0         0     0         0         0     0      0
## 80          0         0         0     0         0         0     0      0
## 81          0         0         0     0         0         0     0      0
## 82          0         0         0     0         0         0     0      0
## 83          0         0         0     0         0         0     0      0
## 84          0         0         0     0         0         0     0      0
## 85          0         0         0     0         0         0     0      0
## 86          0         0         0     0         0         0     0      0
## 87          0         0         0     0         0         0     0      0
## 88          0         0         0     0         0         0     0      0
## 89          0         0         0     0         0         0     0      0
## 90          0         0         0     0         0         0     0      0
##    Seagram.s Skol Sky Smirnoff Stolicnaya Tanqueray Three.Olives
## 78         0    0   0        0          0         0            0
## 79         0    0   0        0          0         0            0
## 80         0    0   0        0          0         0            0
## 81         0    0   0        0          0         0            0
## 82         0    0   0        0          0         0            0
## 83         0    0   0        0          0         0            0
## 84         0    0   0        0          0         0            0
## 85         0    0   0        0          0         0            0
## 86         0    0   0        0          0         0            0
## 87         0    0   0        0          0         0            0
## 88         0    0   0        0          0         0            0
## 89         0    0   0        0          0         0            0
## 90         0    0   0        0          0         0            0
```

```
##    TotalSales LagTotalSales X2LagTotalSales  LnSales LnLSales Ln2Lsales
## 78        240           285             250 5.480639 5.652489  5.521461
## 79        220           240             285 5.393628 5.480639  5.652489
## 80        183           220             240 5.209486 5.393628  5.480639
## 81        248           183             220 5.513429 5.209486  5.393628
## 82        252           248             183 5.529429 5.513429  5.209486
## 83        275           252             248 5.616771 5.529429  5.513429
## 84        342           275             252 5.834811 5.616771  5.529429
## 85        319           342             275 5.765191 5.834811  5.616771
## 86        327           319             342 5.789960 5.765191  5.834811
## 87        357           327             319 5.877736 5.789960  5.765191
## 88        365           357             327 5.899898 5.877736  5.789960
## 89        381           365             357 5.942800 5.899898  5.877736
## 90        362           381             365 5.891644 5.942800  5.899898
##        LnDiff        diff IfDom DollarSales PriceRerUnit LagPrice  LnPrice
## 78 -0.1718502 -0.15789474     0     35170.0     146.5417 117.1930 4.987310
## 79 -0.0870109 -0.08333333     0     24449.0     111.1318 146.5417 4.710717
## 80 -0.1841421 -0.16818182     0     26043.5     142.3142 111.1318 4.958037
## 81  0.3039432  0.35519126     0     28024.0     113.0000 142.3142 4.727388
## 82  0.0159998  0.01612903     0     28476.0     113.0000 113.0000 4.727388
## 83  0.0873423  0.09126984     0     34998.0     127.2655 113.0000 4.846275
## 84  0.2180400  0.24363636     0     43128.0     126.1053 127.2655 4.837117
## 85 -0.0696201 -0.06725146     0     38216.0     119.7994 126.1053 4.785819
## 86  0.0247688  0.02507837     0     39175.0     119.8012 119.7994 4.785834
## 87  0.0877762  0.09174312     0     46626.0     130.6050 119.8012 4.872178
## 88  0.0221620  0.02240896     0     48234.0     132.1479 130.6050 4.883922
## 89  0.0429020  0.04383562     0     50041.0     131.3412 132.1479 4.877799
## 90 -0.0511560 -0.04986877     0     49073.0     135.5608 131.3412 4.909420
##    LnLPrice    Mag  News Outdoor Broad  Print    LnMag     LnNews
## 78 4.763822  234.0   1.0     1.0   1.0  235.0 0.000000  0.0000000
## 79 4.987310    1.0   1.0     1.0   1.0    2.0 0.000000  0.0000000
## 80 4.710717 2412.2   0.4     1.8   1.0 2412.6 7.788294 -0.9162907
## 81 4.958037 8156.4  90.2   575.6 325.9 8246.6 9.006558  4.5020290
## 82 4.727388 5092.0 404.7   158.0 272.8 5496.7 8.535426  6.0031460
## 83 4.727388 1372.1 682.0   301.2   5.8 2054.1 7.224098  6.5250300
## 84 4.846275  389.2 396.0    54.8   1.0  785.2 5.964093  5.9814140
## 85 4.837117  692.1 112.8    22.0   1.0  804.9 6.539731  4.7256160
## 86 4.785819  756.9   1.0     1.0   1.0  757.9 6.629231  0.0000000
## 87 4.785834  167.3 188.5    49.3   1.0  355.8 5.119789  5.2390980
## 88 4.872178 1457.7 219.7    92.3   1.0 1677.4 7.284615  5.3922630
## 89 4.883922 5926.7   1.0   201.2 201.2 5927.7 8.687222  0.0000000
## 90 4.877799 7350.7   1.0   298.6 119.7 7351.7 8.902551  0.0000000
##        LnOut  LnBroad  LnPrint Tier1 Tier2 TotalMinusSales
## 78 0.0000100 0.000010 5.459586     0     1           63600
## 79 0.0000100 0.000010 0.000010     0     1           63652
## 80 0.5877866 0.000010 7.788460     0     1           63721
## 81 6.3554130 5.786591 9.017556     0     1           63688
## 82 5.0625950 5.608739 8.611903     0     1           63716
## 83 5.7077750 1.757858 7.627593     0     1           63725
## 84 4.0036900 0.000010 6.665938     0     1           63690
## 85 3.0910430 0.000010 6.690718     0     1           63745
## 86 0.0000100 0.000010 6.630551     0     1           63769
## 87 3.8979240 0.000010 5.874369     0     1           63771
## 88 4.5250440 0.000010 7.425000     0     1           63795
```

```
## 89 5.3042990 5.304299 8.687391     0     1           63811
## 90 5.6991050 4.784988 8.902687     0     1           63862
##    LagTotalMinusSales TierSales OutsideTierSales LagTierSales
## 78              63523      6106            57494         6152
## 79              63600      6696            56956         6106
## 80              63652      6407            57314         6696
## 81              63721      7652            56036         6407
## 82              63688      7208            56508         7652
## 83              63716      7267            56458         7208
## 84              63725      8103            55587         7267
## 85              63690      9175            54570         8103
## 86              63745     10716            53053         9175
## 87              63769     11476            52295        10716
## 88              63771     12830            50965        11476
## 89              63795     14005            49806        12830
## 90              63811     15504            48358        14005
##    LagOutsideTierSales Firstintro Marketshare LagMktshare YearID total.ad
## 78               57371          0   0.0107138   0.0121889      3    472.0
## 79               57494          0   0.0093537   0.0107138      4      6.0
## 80               56956          0   0.0078072   0.0093537      5   4828.0
## 81               57314          0   0.0099819   0.0078072      6  17394.7
## 82               56036          0   0.0100354   0.0099819      7  11424.2
## 83               56508          0   0.0103531   0.0100354      8   4415.2
## 84               56458          0   0.0123044   0.0103531      9   1626.2
## 85               55587          0   0.0107230   0.0123044     10   1632.8
## 86               54570          0   0.0101455   0.0107230     11   1517.8
## 87               53053          0   0.0103370   0.0101455     12    761.9
## 88               52295          0   0.0100335   0.0103370     13   3448.1
## 89               50965          0   0.0097785   0.0100335     14  12257.8
## 90               49806          0   0.0087090   0.0097785     15  15121.7
##    NA. NA..1 NA..2 NA..3 NA..4 NA..5 NA..6
## 78  NA    NA    NA    NA    NA    NA    NA
## 79  NA    NA    NA    NA    NA    NA    NA
## 80  NA    NA    NA    NA    NA    NA    NA
## 81  NA    NA    NA    NA    NA    NA    NA
## 82  NA    NA    NA    NA    NA    NA    NA
## 83  NA    NA    NA    NA    NA    NA    NA
## 84  NA    NA    NA    NA    NA    NA    NA
## 85  NA    NA    NA    NA    NA    NA    NA
## 86  NA    NA    NA    NA    NA    NA    NA
## 87  NA    NA    NA    NA    NA    NA    NA
## 88  NA    NA    NA    NA    NA    NA    NA
## 89  NA    NA    NA    NA    NA    NA    NA
## 90  NA    NA    NA    NA    NA    NA    NA
```

There are a lot of columns not of our interest. Let's filter them out.

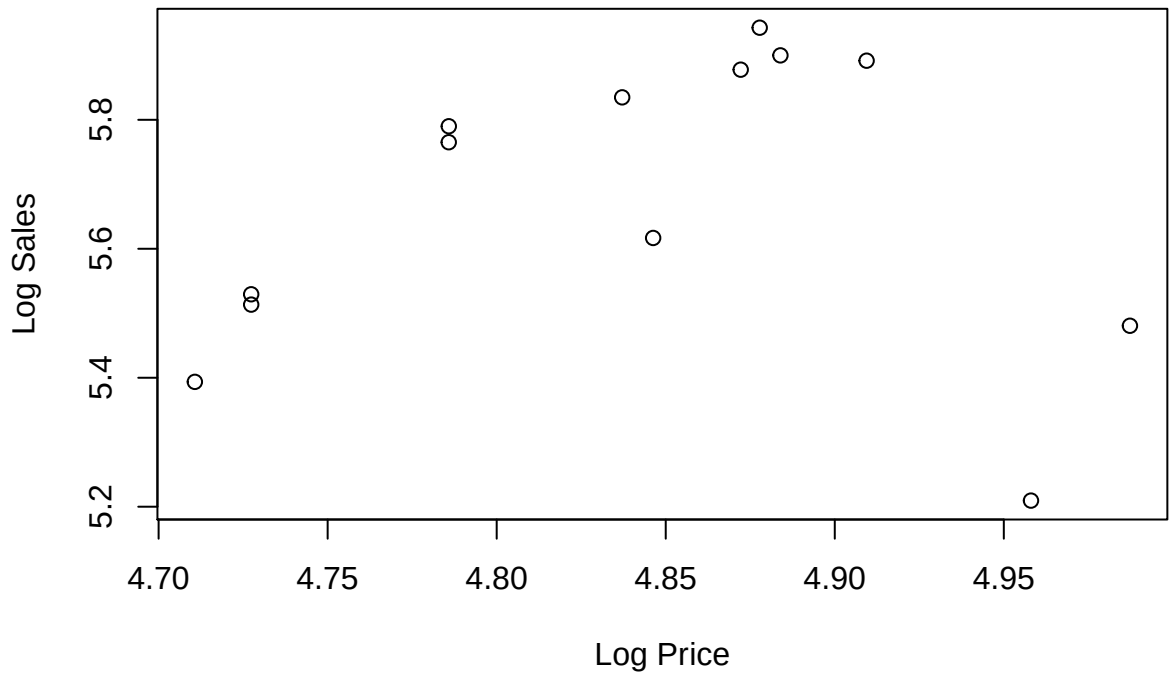
```
Pr_Finlandia = data.frame(Finlandia$LnSales,Finlandia$LnPrice)
names(Pr_Finlandia) = c('LnSales','LnPrice')
Pr_Finlandia
```

```
##     LnSales  LnPrice
## 1  5.480639 4.987310
## 2  5.393628 4.710717
## 3  5.209486 4.958037
```

```
## 4  5.513429 4.727388
## 5  5.529429 4.727388
## 6  5.616771 4.846275
## 7  5.834811 4.837117
## 8  5.765191 4.785819
## 9  5.789960 4.785834
## 10 5.877736 4.872178
## 11 5.899898 4.883922
## 12 5.942800 4.877799
## 13 5.891644 4.909420
```

In order to find the impact of Price on sales,let's build a regression model on the dataset. First let's plot some graphs to understand the data points.

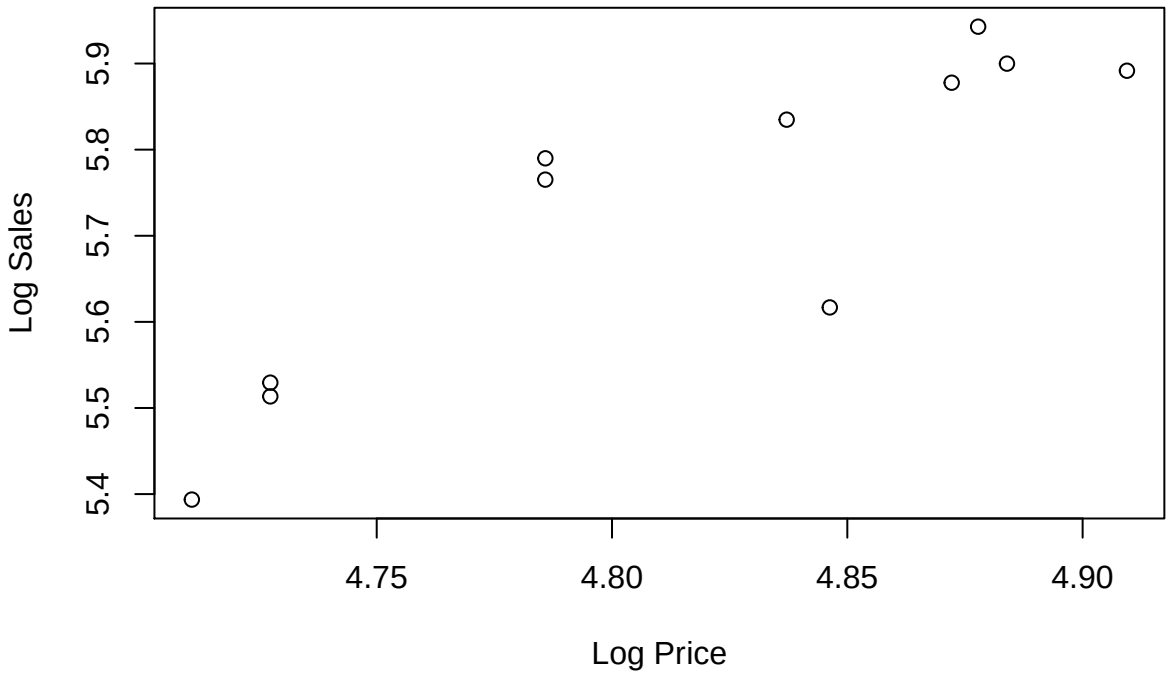
```
plot( x = Pr_Finlandia$LnPrice,y = Pr_Finlandia$LnSales,xlab = 'Log Price',ylab = 'Log Sales')
```



We can infer that price variation is almost linear in nature with sales with outliers.

We can remove the outliers and therefore fit out regression line.

```
Pr_Finlandia = subset(Pr_Finlandia,LnPrice < 4.95)
plot( x = Pr_Finlandia$LnPrice,y = Pr_Finlandia$LnSales,xlab = 'Log Price',ylab = 'Log Sales')
```



Linear Regression :

```
result = lm(formula = 'LnSales ~ LnPrice',data = Pr_Finlandia)
summary(result)
```

```
##
## Call:
## lm(formula = "LnSales ~ LnPrice", data = Pr_Finlandia)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.188187 -0.037886  0.007768  0.057953  0.124954
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -5.4167     2.0426  -2.652 0.026394 *
## LnPrice       2.3155     0.4242   5.459 0.000401 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.09511 on 9 degrees of freedom
## Multiple R-squared:  0.768,  Adjusted R-squared:  0.7423
## F-statistic:  29.8 on 1 and 9 DF,  p-value: 0.000401
```

Here, the p-value of price is almost zero that indicates that price is significant indicator of sales. Here, the value of R-squared is 0.768 i.e the price variable indicates nearly 77% of the data points. The co-efficient of price tells that for every unit increase in price, there is 2.31 times increase in sales. Let's add more variables to the regression and see what happens with the R-squared.

```
Ad_Finlandia = subset(Finlandia,select = c('LnSales','LnMag','LnNews','LnOut','LnBroad','LnPrint','LnPr
dim(Ad_Finlandia)
```

```
## [1] 13  7
```

```
Ad_Finlandia
```

```
##     LnSales    LnMag     LnNews     LnOut  LnBroad  LnPrint  LnPrice
## 78 5.480639 0.000000  0.0000000 0.0000100 0.000010 5.459586 4.987310
## 79 5.393628 0.000000  0.0000000 0.0000100 0.000010 0.000010 4.710717
## 80 5.209486 7.788294 -0.9162907 0.5877866 0.000010 7.788460 4.958037
## 81 5.513429 9.006558  4.5020290 6.3554130 5.786591 9.017556 4.727388
## 82 5.529429 8.535426  6.0031460 5.0625950 5.608739 8.611903 4.727388
## 83 5.616771 7.224098  6.5250300 5.7077750 1.757858 7.627593 4.846275
## 84 5.834811 5.964093  5.9814140 4.0036900 0.000010 6.665938 4.837117
## 85 5.765191 6.539731  4.7256160 3.0910430 0.000010 6.690718 4.785819
## 86 5.789960 6.629231  0.0000000 0.0000100 0.000010 6.630551 4.785834
## 87 5.877736 5.119789  5.2390980 3.8979240 0.000010 5.874369 4.872178
## 88 5.899898 7.284615  5.3922630 4.5250440 0.000010 7.425000 4.883922
## 89 5.942800 8.687222  0.0000000 5.3042990 5.304299 8.687391 4.877799
## 90 5.891644 8.902551  0.0000000 5.6991050 4.784988 8.902687 4.909420
```

```
result_ad = lm(formula ='LnSales ~ LnMag + LnNews + LnOut + LnBroad + LnPrint + LnPrice',data=Ad_Finlan
summary(result_ad)
```

```
##
## Call:
## lm(formula = "LnSales ~ LnMag + LnNews + LnOut + LnBroad + LnPrint + LnPrice",
##     data = Ad_Finlandia)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.28787 -0.06936  0.03921  0.13734  0.19446
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 16.30608    9.98348   1.633   0.1535
## LnMag       -0.07663    0.08489  -0.903   0.4015
## LnNews      -0.10075    0.06977  -1.444   0.1988
## LnOut        0.21095    0.10703   1.971   0.0962 .
## LnBroad     -0.16092    0.09485  -1.697   0.1407
## LnPrint      0.12320    0.12571   0.980   0.3649
## LnPrice     -2.30179    2.12122  -1.085   0.3195
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.231 on 6 degrees of freedom
## Multiple R-squared:  0.5011, Adjusted R-squared:  0.002126
## F-statistic: 1.004 on 6 and 6 DF,  p-value: 0.498
```

The R-squared value has decreased. Here the p-values of some variables are high which can be accounted due to interaction effect and some other factors. Let's try out the interation effect method between variables broad and print

```
result_inter = lm('LnSales ~ LnMag + LnNews + LnOut * LnBroad + LnPrint + LnPrice',data=Ad_Finlandia)
summary(result_inter)
```

```
##
## Call:
## lm(formula = "LnSales ~ LnMag + LnNews + LnOut * LnBroad + LnPrint + LnPrice",
##     data = Ad_Finlandia)
```

```
##
## Residuals:
##        78        79        80        81        82        83        84
## -0.016749 -0.036166 -0.075965  0.019294  0.017785 -0.165698  0.019895
##        85        86        87        88        89        90
## -0.141204  0.103224  0.207687  0.051581  0.003465  0.012852
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)   30.47222    7.98562   3.816  0.01243 *
## LnMag         -0.18276    0.06526  -2.800  0.03798 *
## LnNews        -0.20077    0.05600  -3.585  0.01579 *
## LnOut          0.36721    0.08646   4.247  0.00811 **
## LnBroad        0.28660    0.15967   1.795  0.13261
## LnPrint        0.28170    0.09688   2.908  0.03349 *
## LnPrice       -5.31606    1.69756  -3.132  0.02591 *
## LnOut:LnBroad -0.10413    0.03428  -3.038  0.02882 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.15 on 5 degrees of freedom
## Multiple R-squared:  0.8247, Adjusted R-squared:  0.5792
## F-statistic:  3.36 on 7 and 5 DF,  p-value: 0.1004
```

Boom! The results here are surprised. The presence of significant interaction indicates that the effect of one predictor variable on the response variable is different at different values of the other predictor variable. Therefore,the individial effects of the predictors also include the effects of the co-efficients of the interaction terms.

The R-squared value has also increased indicating the model can explain most of the data points now.