MIS40970 Assignment 1 Orange Juice Sales

Contents

1. Load the oj.csv data into R.

```
oj <- read.csv(file="E:/Smurfit/Semester 2/DM/Week 2/oj.csv",head=TRUE,sep=",")
summary(oj)
```

```
##      store                brand           week          logmove
##  Min.   :  2.00   dominicks  :9649   Min.   : 40.0   Min.   : 4.159
##  1st Qu.: 53.00   minute.maid:9649   1st Qu.: 70.0   1st Qu.: 8.490
##  Median : 86.00   tropicana  :9649   Median :101.0   Median : 9.034
##  Mean   : 80.88                      Mean   :100.5   Mean   : 9.168
##  3rd Qu.:111.00                      3rd Qu.:130.0   3rd Qu.: 9.765
##  Max.   :137.00                      Max.   :160.0   Max.   :13.482
##       feat            price           AGE60              EDUC
##  Min.   :0.0000   Min.   :0.520   Min.   :0.05805   Min.   :0.04955
##  1st Qu.:0.0000   1st Qu.:1.790   1st Qu.:0.12210   1st Qu.:0.14598
##  Median :0.0000   Median :2.170   Median :0.17065   Median :0.22939
##  Mean   :0.2373   Mean   :2.282   Mean   :0.17313   Mean   :0.22522
##  3rd Qu.:0.0000   3rd Qu.:2.730   3rd Qu.:0.21395   3rd Qu.:0.28439
##  Max.   :1.0000   Max.   :3.870   Max.   :0.30740   Max.   :0.52836
##      ETHNIC            INCOME          HHLARGE           WORKWOM
##  Min.   :0.02425   Min.   : 9.867   Min.   :0.01351   Min.   :0.2445
##  1st Qu.:0.04191   1st Qu.:10.456   1st Qu.:0.09794   1st Qu.:0.3126
##  Median :0.07466   Median :10.635   Median :0.11122   Median :0.3556
##  Mean   :0.15556   Mean   :10.617   Mean   :0.11560   Mean   :0.3592
##  3rd Qu.:0.18776   3rd Qu.:10.797   3rd Qu.:0.13517   3rd Qu.:0.4023
##  Max.   :0.99569   Max.   :11.236   Max.   :0.21635   Max.   :0.4723
##     HVAL150            SSTRDIST          SSTRVOL          CPDIST5
##  Min.   :0.002509   Min.   : 0.1321   Min.   :0.4000   Min.   :0.7725
##  1st Qu.:0.123486   1st Qu.: 2.7670   1st Qu.:0.7273   1st Qu.:1.6262
##  Median :0.346154   Median : 4.6507   Median :1.1154   Median :1.9634
##  Mean   :0.343766   Mean   : 5.0973   Mean   :1.2073   Mean   :2.1204
##  3rd Qu.:0.528313   3rd Qu.: 6.6506   3rd Qu.:1.5385   3rd Qu.:2.5337
##  Max.   :0.916700   Max.   :17.8560   Max.   :2.5714   Max.   :4.1079
##     CPWVOL5
##  Min.   :0.09456
##  1st Qu.:0.27167
##  Median :0.38323
##  Mean   :0.43891
##  3rd Qu.:0.56024
##  Max.   :1.14337
```

2. How many records and how many attributes are in the orange juice dataset?
3. What is the mean, standard deviation and range of the price of orange juice?
4. What is the median of the 'log of number of units sold' (logmove)?

```
dim(oj)
```

```
## [1] 28947    17
```

```r
names(oj)
```

```
##  [1] "store"    "brand"    "week"     "logmove"  "feat"     "price"
##  [7] "AGE60"    "EDUC"     "ETHNIC"   "INCOME"   "HHLARGE"  "WORKWOM"
## [13] "HVAL150"  "SSTRDIST" "SSTRVOL"  "CPDIST5"  "CPWVOL5"
```

```r
mean(oj$price)
```

```
## [1] 2.282488
```

```r
sd(oj$price)
```

```
## [1] 0.6480007
```

```r
range(oj$price)
```

```
## [1] 0.52 3.87
```

```r
median(oj$logmove)
```

```
## [1] 9.03408
```

5. What are the names of the 3 orange juice brands? (Hint: brand is of type factor, and a factor is comprised of a number of . . . <insert function name here!. . .)

```r
unique(oj$brand)
```

```
## [1] tropicana   minute.maid dominicks
## Levels: dominicks minute.maid tropicana
```

```r
juicebrand<- factor(oj$brand)
table(juicebrand)
```

```
## juicebrand
##   dominicks minute.maid   tropicana
##        9649        9649        9649
```

6. Create a histogram of prices for each brand of orange juice.

```r
hist(oj$price[oj$brand=='tropicana'],main=" Price of Tropicana Orange Juice",xlab="Price",breaks =50, c
```

```
hist(oj$price[oj$brand=='minute.maid'],main=" Price of Mindute Maid Orange Juice",xlab="Price",breaks =
```

```
hist(oj$price[oj$brand=='dominicks'],main=" Price of Dominicks Orange Juice",xlab="Price",breaks =50, c

"Alternative Method Using ggplot2"
```

```
## [1] "Alternative Method Using ggplot2"
```

```
library(ggplot2)
```

```
ggplot(oj, aes(x= oj$price, fill = brand)) + geom_histogram(binwidth=1,position = "dodge")+ ylab("Count
  xlab("Orange Juice Price") +
  ggtitle("Prices of Orange Juice and Count of Each Type")
```

7. Generate a boxplot, which includes a separate plot for the prices of each brand (Hint: the boxplot should look something like the example below)

```
boxplot(oj$price~oj$brand, data=oj, notch=TRUE,
  col=(c("gold","darkgreen")),main="Boxplot of Prices for each Orange Juince brand", xlab="Price and Bra
```

8. What does the boxplot tell us about the relative prices of each brand? (Hint: answer this one by interpreting what you observe in the boxplot, i.e., no coding required!).

9. Generate a scatterplot of the logmove compared to price, and color the points according to their brand (e.g., "red" for brand 1, "green" for brand 2 etc).

```
library(scatterplot3d)
attach(oj)
pairs(~logmove+price, pch=21,main="Scatter plot of Log of the number of units Sold vs Price of Orange Ju
```

```
plot(logmove,price, xlab="Units Sold",ylab="Price",pch=21,col=topo.colors(3)[unclass(brand)])
legend("bottomleft", inset=0.01, title="Brands",c("Tropicana","Minute Maid","Dominicks"), fill=topo.col
```

```r
library(plotly)
```

```
##
## Attaching package: 'plotly'

## The following object is masked from 'package:ggplot2':
##
##     last_plot

## The following object is masked from 'package:stats':
##
##     filter

## The following object is masked from 'package:graphics':
##
##     layout
```

```r
Sys.setenv("plotly_username"="deepak91g")
Sys.setenv("plotly_api_key"="Gb3tMDu2UzMuGdXsMB9n")

p <- plot_ly(data = oj, x = ~logmove, y = ~price, color = ~brand, colors = "Set1")

# Create a shareable link to your chart
# Set up API credentials: https://plot.ly/r/getting-started
chart_link = plotly_POST(p, filename="colorbrewer")
```

```
## No trace type specified:
##   Based on info supplied, a 'scatter' trace seems appropriate.
##   Read more about this trace type -> https://plot.ly/r/reference/#scatter
```

```
## No scatter mode specifed:
##   Setting the mode to markers
##   Read more about this attribute -> https://plot.ly/r/reference/#scatter-mode

## No encoding supplied: defaulting to UTF-8.

## Success! Modified your plotly here -> https://plot.ly/~deepak91g/1
```

```
chart_link
```

10. Based on what you observe in the scatterplot, what can we say about the price of each brand of chilled orange juice and the volume of sales?
11. Calculate the mean price of orange juice sold each week, and create a line plot of this timeseries. (Hint: use tapply(), and an example plot is provided below.)

Note: Using tapply() method here

```
xnames <- tapply(oj$price,oj$week,FUN=mean,na.rm=TRUE)
plot(tapply(oj$price,oj$week,FUN=mean,na.rm=TRUE),xaxt="n",pch=16,type="l",col="green",xlab = "Week",yla
```

NOTE: Answer to question 11, using aggreagate method

```
meanweeklyprice <- aggregate(oj[ ,6], list(oj$week), mean)
plot(meanweeklyprice$Group.1,meanweeklyprice$x,type="l",col="green",xlab = "Week",ylab="meanweeklyprice
```

12. Extract the mean weekly price of orange juice sold each week according to each brand (Hint: use tapply() with a list for the INDEX)

```
#Solution :1 using aggregate()
mean_weekly_brand_price <- aggregate(oj[ ,6], list(Brand=oj$brand,Week=oj$week), mean)
mean_weekly_brand_price
```

```
##           Brand Week        x
## 1     dominicks   40 1.590000
## 2   minute.maid   40 2.890000
## 3     tropicana   40 3.509315
## 4     dominicks   41 2.447015
## 5   minute.maid   41 1.990000
## 6     tropicana   41 3.514627
## 7     dominicks   42 2.461940
## 8   minute.maid   42 2.911343
## 9     tropicana   42 3.536567
## 10    dominicks   43 2.452597
## 11  minute.maid   43 2.892597
## 12    tropicana   43 3.517792
## 13    dominicks   44 2.463506
## 14  minute.maid   44 2.590000
## 15    tropicana   44 3.534156
## 16    dominicks   45 2.467838
## 17  minute.maid   45 2.908649
## 18    tropicana   45 3.541486
## 19    dominicks   46 2.471053
```

```
## 20  minute.maid   46 2.913158
## 21    tropicana   46 3.547500
## 22    dominicks   47 2.086800
## 23  minute.maid   47 2.390000
## 24    tropicana   47 3.534000
## 25    dominicks   48 2.086883
## 26  minute.maid   48 2.907013
## 27    tropicana   48 3.540000
## 28    dominicks   49 2.086842
## 29  minute.maid   49 1.990000
## 30    tropicana   49 3.545000
## 31    dominicks   50 2.083924
## 32  minute.maid   50 2.911392
## 33    tropicana   50 3.539114
## 34    dominicks   51 1.889000
## 35  minute.maid   51 2.921625
## 36    tropicana   51 3.552375
## 37    dominicks   52 1.888987
## 38  minute.maid   52 2.913797
## 39    tropicana   52 3.243544
## 40    dominicks   53 1.888987
## 41  minute.maid   53 2.192025
## 42    tropicana   53 3.243544
## 43    dominicks   54 1.790000
## 44  minute.maid   54 2.908765
## 45    tropicana   54 3.245062
## 46    dominicks   55 1.779610
## 47  minute.maid   55 2.909481
## 48    tropicana   55 3.542857
## 49    dominicks   56 1.790000
## 50  minute.maid   56 2.909114
## 51    tropicana   56 3.169114
## 52    dominicks   57 2.469000
## 53  minute.maid   57 1.990000
## 54    tropicana   57 3.170375
## 55    dominicks   58 1.202469
## 56  minute.maid   58 2.657160
## 57    tropicana   58 3.250494
## 58    dominicks   59 1.590000
## 59  minute.maid   59 2.912375
## 60    tropicana   59 3.432375
## 61    dominicks   60 1.590000
## 62  minute.maid   60 1.979367
## 63    tropicana   60 3.431772
## 64    dominicks   61 1.290000
## 65  minute.maid   61 2.914304
## 66    tropicana   61 3.431519
## 67    dominicks   62 0.990000
## 68  minute.maid   62 1.990000
## 69    tropicana   62 3.547778
## 70    dominicks   63 1.590000
## 71  minute.maid   63 1.990000
## 72    tropicana   63 2.981951
## 73    dominicks   64 2.414938
```

```
## 74  minute.maid   64 2.913951
## 75    tropicana   64 2.980123
## 76    dominicks   65 2.219877
## 77  minute.maid   65 1.990000
## 78    tropicana   65 3.284321
## 79    dominicks   66 1.390000
## 80  minute.maid   66 1.990000
## 81    tropicana   66 3.281852
## 82    dominicks   67 2.217073
## 83  minute.maid   67 1.990000
## 84    tropicana   67 2.390000
## 85    dominicks   68 2.229737
## 86  minute.maid   68 1.990000
## 87    tropicana   68 2.390000
## 88    dominicks   69 2.226962
## 89  minute.maid   69 1.990000
## 90    tropicana   69 3.247975
## 91    dominicks   70 0.990000
## 92  minute.maid   70 2.401500
## 93    tropicana   70 2.590000
## 94    dominicks   71 2.230488
## 95  minute.maid   71 2.401707
## 96    tropicana   71 2.590000
## 97    dominicks   72 2.230988
## 98  minute.maid   72 1.988889
## 99    tropicana   72 3.276914
## 100   dominicks   73 1.490000
## 101 minute.maid   73 2.386829
## 102   tropicana   73 3.275854
## 103   dominicks   74 1.490000
## 104 minute.maid   74 2.318313
## 105   tropicana   74 2.490000
## 106   dominicks   75 2.242099
## 107 minute.maid   75 2.323210
## 108   tropicana   75 3.278642
## 109   dominicks   76 2.237654
## 110 minute.maid   76 2.318642
## 111   tropicana   76 3.275432
## 112   dominicks   77 1.938889
## 113 minute.maid   77 1.989012
## 114   tropicana   77 3.278642
## 115   dominicks   78 1.037683
## 116 minute.maid   78 2.386829
## 117   tropicana   78 3.270244
## 118   dominicks   79 2.248701
## 119 minute.maid   79 2.393377
## 120   tropicana   79 3.278052
## 121   dominicks   80 2.227632
## 122 minute.maid   80 1.725000
## 123   tropicana   80 3.241579
## 124   dominicks   81 1.590000
## 125 minute.maid   81 1.723086
## 126   tropicana   81 1.603086
## 127   dominicks   82 1.590000
```

```
## 128 minute.maid   82 2.387949
## 129   tropicana   82 3.274359
## 130   dominicks   83 1.590000
## 131 minute.maid   83 2.048000
## 132   tropicana   83 1.987500
## 133   dominicks   84 0.990000
## 134 minute.maid   84 2.048519
## 135   tropicana   84 3.275432
## 136   dominicks   85 2.237073
## 137 minute.maid   85 2.049024
## 138   tropicana   85 1.990000
## 139   dominicks   86 2.233250
## 140 minute.maid   86 2.049750
## 141   tropicana   86 1.990000
## 142   dominicks   87 2.245325
## 143 minute.maid   87 1.390000
## 144   tropicana   87 3.282468
## 145   dominicks   88 2.248025
## 146 minute.maid   88 2.206420
## 147   tropicana   88 2.290000
## 148   dominicks   89 1.590000
## 149 minute.maid   89 2.200533
## 150   tropicana   89 3.279600
## 151   dominicks   90 1.590000
## 152 minute.maid   90 2.204487
## 153   tropicana   90 3.286410
## 154   dominicks   91 2.114000
## 155 minute.maid   91 2.214933
## 156   tropicana   91 1.990000
## 157   dominicks   92 1.380854
## 158 minute.maid   92 1.660854
## 159   tropicana   92 1.990000
## 160   dominicks   93 1.255802
## 161 minute.maid   93 2.144074
## 162   tropicana   93 3.300494
## 163   dominicks   94 1.814810
## 164 minute.maid   94 1.688481
## 165   tropicana   94 3.143544
## 166   dominicks   95 1.820519
## 167 minute.maid   95 1.701429
## 168   tropicana   95 1.990000
## 169   dominicks   96 1.590000
## 170 minute.maid   96 2.223462
## 171   tropicana   96 1.990000
## 172   dominicks   97 1.595921
## 173 minute.maid   97 2.211974
## 174   tropicana   97 2.038026
## 175   dominicks   98 1.753671
## 176 minute.maid   98 2.217089
## 177   tropicana   98 3.189241
## 178   dominicks   99 0.990000
## 179 minute.maid   99 2.210390
## 180   tropicana   99 2.189870
## 181   dominicks  100 1.026265
```

```
## 182 minute.maid  100 2.218193
## 183   tropicana  100 2.229157
## 184   dominicks  101 1.486709
## 185 minute.maid  101 2.082152
## 186   tropicana  101 3.314810
## 187   dominicks  102 1.476829
## 188 minute.maid  102 1.286341
## 189   tropicana  102 3.312683
## 190   dominicks  103 1.389870
## 191 minute.maid  103 2.214156
## 192   tropicana  103 3.315195
## 193   dominicks  104 1.389875
## 194 minute.maid  104 2.195750
## 195   tropicana  104 1.988750
## 196   dominicks  105 1.403049
## 197 minute.maid  105 1.974634
## 198   tropicana  105 3.237805
## 199   dominicks  106 1.583750
## 200 minute.maid  106 1.960000
## 201   tropicana  106 2.556750
## 202   dominicks  107 1.583614
## 203 minute.maid  107 1.868193
## 204   tropicana  107 3.271205
## 205   dominicks  108 1.567273
## 206 minute.maid  108 1.790000
## 207   tropicana  108 3.012078
## 208   dominicks  109 1.545823
## 209 minute.maid  109 1.820127
## 210   tropicana  109 2.956329
## 211   dominicks  110 1.571625
## 212 minute.maid  110 1.690000
## 213   tropicana  110 2.960875
## 214   dominicks  111 1.677683
## 215 minute.maid  111 2.205732
## 216   tropicana  111 2.957439
## 217   dominicks  112 1.649880
## 218 minute.maid  112 1.489759
## 219   tropicana  112 2.959036
## 220   dominicks  113 1.617927
## 221 minute.maid  113 2.205854
## 222   tropicana  113 2.966951
## 223   dominicks  114 1.630854
## 224 minute.maid  114 2.157805
## 225   tropicana  114 3.021463
## 226   dominicks  115 1.687654
## 227 minute.maid  115 1.974321
## 228   tropicana  115 1.990000
## 229   dominicks  116 1.548293
## 230 minute.maid  116 1.971098
## 231   tropicana  116 2.866463
## 232   dominicks  117 1.576463
## 233 minute.maid  117 1.984756
## 234   tropicana  117 3.025854
## 235   dominicks  118 1.588795
```

```
## 236 minute.maid  118 2.218795
## 237   tropicana  118 3.024940
## 238   dominicks  119 1.590244
## 239 minute.maid  119 2.175244
## 240   tropicana  119 2.900610
## 241   dominicks  120 1.617470
## 242 minute.maid  120 1.974819
## 243   tropicana  120 2.384699
## 244   dominicks  121 1.605732
## 245 minute.maid  121 1.972561
## 246   tropicana  121 2.362927
## 247   dominicks  122 1.625556
## 248 minute.maid  122 2.026420
## 249   tropicana  122 2.273580
## 250   dominicks  123 1.643614
## 251 minute.maid  123 2.462651
## 252   tropicana  123 2.337711
## 253   dominicks  124 1.541205
## 254 minute.maid  124 2.455422
## 255   tropicana  124 2.994699
## 256   dominicks  125 1.466829
## 257 minute.maid  125 2.478049
## 258   tropicana  125 2.994878
## 259   dominicks  126 1.491707
## 260 minute.maid  126 2.484268
## 261   tropicana  126 2.992073
## 262   dominicks  127 1.522892
## 263 minute.maid  127 1.982651
## 264   tropicana  127 1.990000
## 265   dominicks  128 1.557308
## 266 minute.maid  128 1.971026
## 267   tropicana  128 2.990769
## 268   dominicks  129 1.533659
## 269 minute.maid  129 2.018171
## 270   tropicana  129 1.982683
## 271   dominicks  130 1.505732
## 272 minute.maid  130 2.469512
## 273   tropicana  130 1.803902
## 274   dominicks  131 1.530120
## 275 minute.maid  131 1.988916
## 276   tropicana  131 1.788916
## 277   dominicks  132 1.583735
## 278 minute.maid  132 2.016024
## 279   tropicana  132 1.860120
## 280   dominicks  133 1.603659
## 281 minute.maid  133 2.413415
## 282   tropicana  133 2.902195
## 283   dominicks  134 1.688675
## 284 minute.maid  134 1.978916
## 285   tropicana  134 2.500000
## 286   dominicks  135 1.528293
## 287 minute.maid  135 2.033171
## 288   tropicana  135 2.534756
## 289   dominicks  136 1.290000
```

```
## 290 minute.maid  136 2.600976
## 291   tropicana  136 2.987561
## 292   dominicks  137 1.329398
## 293 minute.maid  137 1.960000
## 294   tropicana  137 2.832530
## 295   dominicks  138 1.860964
## 296 minute.maid  138 2.619036
## 297   tropicana  138 2.240120
## 298   dominicks  139 1.870123
## 299 minute.maid  139 2.618025
## 300   tropicana  139 2.991235
## 301   dominicks  140 1.965244
## 302 minute.maid  140 2.605732
## 303   tropicana  140 2.992439
## 304   dominicks  141 1.689880
## 305 minute.maid  141 1.960120
## 306   tropicana  141 2.991205
## 307   dominicks  142 1.689615
## 308 minute.maid  142 2.622436
## 309   tropicana  142 2.989103
## 310   dominicks  143 1.701538
## 311 minute.maid  143 2.686154
## 312   tropicana  143 2.490000
## 313   dominicks  144 1.956543
## 314 minute.maid  144 2.678889
## 315   tropicana  144 2.381481
## 316   dominicks  145 1.957662
## 317 minute.maid  145 2.661429
## 318   tropicana  145 1.990000
## 319   dominicks  146 1.910000
## 320 minute.maid  146 2.545750
## 321   tropicana  146 2.977250
## 322   dominicks  147 1.852099
## 323 minute.maid  147 2.010247
## 324   tropicana  147 2.673210
## 325   dominicks  148 1.327051
## 326 minute.maid  148 2.242692
## 327   tropicana  148 2.983590
## 328   dominicks  149 1.885309
## 329 minute.maid  149 2.680123
## 330   tropicana  149 2.990617
## 331   dominicks  150 1.693250
## 332 minute.maid  150 2.278750
## 333   tropicana  150 2.986375
## 334   dominicks  151 1.483544
## 335 minute.maid  151 2.677089
## 336   tropicana  151 2.997342
## 337   dominicks  152 1.869747
## 338 minute.maid  152 2.619620
## 339   tropicana  152 2.908987
## 340   dominicks  153 1.933247
## 341 minute.maid  153 1.988701
## 342   tropicana  153 3.009481
## 343   dominicks  154 1.587051
```

```
## 344 minute.maid  154 1.966923
## 345   tropicana  154 2.774615
## 346   dominicks  155 1.273590
## 347 minute.maid  155 1.689615
## 348   tropicana  155 2.892051
## 349   dominicks  156 1.959610
## 350 minute.maid  156 1.716234
## 351   tropicana  156 2.490000
## 352   dominicks  157 1.975316
## 353 minute.maid  157 1.966709
## 354   tropicana  157 2.535823
## 355   dominicks  158 1.860000
## 356 minute.maid  158 2.622875
## 357   tropicana  158 3.000625
## 358   dominicks  159 1.489877
## 359 minute.maid  159 2.513704
## 360   tropicana  159 2.783210
## 361   dominicks  160 1.880000
## 362 minute.maid  160 2.182500
## 363   tropicana  160 2.973250
```

```r
#Solution :2 using tapply()
weekly_brand_price<-tapply(oj$price,list(Week=oj$week,oj$brand),FUN=mean,na.rm=TRUE)
weekly_brand_price
```

```
##
## Week  dominicks minute.maid tropicana
##   40   1.590000    2.890000  3.509315
##   41   2.447015    1.990000  3.514627
##   42   2.461940    2.911343  3.536567
##   43   2.452597    2.892597  3.517792
##   44   2.463506    2.590000  3.534156
##   45   2.467838    2.908649  3.541486
##   46   2.471053    2.913158  3.547500
##   47   2.086800    2.390000  3.534000
##   48   2.086883    2.907013  3.540000
##   49   2.086842    1.990000  3.545000
##   50   2.083924    2.911392  3.539114
##   51   1.889000    2.921625  3.552375
##   52   1.888987    2.913797  3.243544
##   53   1.888987    2.192025  3.243544
##   54   1.790000    2.908765  3.245062
##   55   1.779610    2.909481  3.542857
##   56   1.790000    2.909114  3.169114
##   57   2.469000    1.990000  3.170375
##   58   1.202469    2.657160  3.250494
##   59   1.590000    2.912375  3.432375
##   60   1.590000    1.979367  3.431772
##   61   1.290000    2.914304  3.431519
##   62   0.990000    1.990000  3.547778
##   63   1.590000    1.990000  2.981951
##   64   2.414938    2.913951  2.980123
##   65   2.219877    1.990000  3.284321
##   66   1.390000    1.990000  3.281852
##   67   2.217073    1.990000  2.390000
```

```
##   68   2.229737    1.990000  2.390000
##   69   2.226962    1.990000  3.247975
##   70   0.990000    2.401500  2.590000
##   71   2.230488    2.401707  2.590000
##   72   2.230988    1.988889  3.276914
##   73   1.490000    2.386829  3.275854
##   74   1.490000    2.318313  2.490000
##   75   2.242099    2.323210  3.278642
##   76   2.237654    2.318642  3.275432
##   77   1.938889    1.989012  3.278642
##   78   1.037683    2.386829  3.270244
##   79   2.248701    2.393377  3.278052
##   80   2.227632    1.725000  3.241579
##   81   1.590000    1.723086  1.603086
##   82   1.590000    2.387949  3.274359
##   83   1.590000    2.048000  1.987500
##   84   0.990000    2.048519  3.275432
##   85   2.237073    2.049024  1.990000
##   86   2.233250    2.049750  1.990000
##   87   2.245325    1.390000  3.282468
##   88   2.248025    2.206420  2.290000
##   89   1.590000    2.200533  3.279600
##   90   1.590000    2.204487  3.286410
##   91   2.114000    2.214933  1.990000
##   92   1.380854    1.660854  1.990000
##   93   1.255802    2.144074  3.300494
##   94   1.814810    1.688481  3.143544
##   95   1.820519    1.701429  1.990000
##   96   1.590000    2.223462  1.990000
##   97   1.595921    2.211974  2.038026
##   98   1.753671    2.217089  3.189241
##   99   0.990000    2.210390  2.189870
##   100  1.026265    2.218193  2.229157
##   101  1.486709    2.082152  3.314810
##   102  1.476829    1.286341  3.312683
##   103  1.389870    2.214156  3.315195
##   104  1.389875    2.195750  1.988750
##   105  1.403049    1.974634  3.237805
##   106  1.583750    1.960000  2.556750
##   107  1.583614    1.868193  3.271205
##   108  1.567273    1.790000  3.012078
##   109  1.545823    1.820127  2.956329
##   110  1.571625    1.690000  2.960875
##   111  1.677683    2.205732  2.957439
##   112  1.649880    1.489759  2.959036
##   113  1.617927    2.205854  2.966951
##   114  1.630854    2.157805  3.021463
##   115  1.687654    1.974321  1.990000
##   116  1.548293    1.971098  2.866463
##   117  1.576463    1.984756  3.025854
##   118  1.588795    2.218795  3.024940
##   119  1.590244    2.175244  2.900610
##   120  1.617470    1.974819  2.384699
##   121  1.605732    1.972561  2.362927
```

```
##   122  1.625556    2.026420  2.273580
##   123  1.643614    2.462651  2.337711
##   124  1.541205    2.455422  2.994699
##   125  1.466829    2.478049  2.994878
##   126  1.491707    2.484268  2.992073
##   127  1.522892    1.982651  1.990000
##   128  1.557308    1.971026  2.990769
##   129  1.533659    2.018171  1.982683
##   130  1.505732    2.469512  1.803902
##   131  1.530120    1.988916  1.788916
##   132  1.583735    2.016024  1.860120
##   133  1.603659    2.413415  2.902195
##   134  1.688675    1.978916  2.500000
##   135  1.528293    2.033171  2.534756
##   136  1.290000    2.600976  2.987561
##   137  1.329398    1.960000  2.832530
##   138  1.860964    2.619036  2.240120
##   139  1.870123    2.618025  2.991235
##   140  1.965244    2.605732  2.992439
##   141  1.689880    1.960120  2.991205
##   142  1.689615    2.622436  2.989103
##   143  1.701538    2.686154  2.490000
##   144  1.956543    2.678889  2.381481
##   145  1.957662    2.661429  1.990000
##   146  1.910000    2.545750  2.977250
##   147  1.852099    2.010247  2.673210
##   148  1.327051    2.242692  2.983590
##   149  1.885309    2.680123  2.990617
##   150  1.693250    2.278750  2.986375
##   151  1.483544    2.677089  2.997342
##   152  1.869747    2.619620  2.908987
##   153  1.933247    1.988701  3.009481
##   154  1.587051    1.966923  2.774615
##   155  1.273590    1.689615  2.892051
##   156  1.959610    1.716234  2.490000
##   157  1.975316    1.966709  2.535823
##   158  1.860000    2.622875  3.000625
##   159  1.489877    2.513704  2.783210
##   160  1.880000    2.182500  2.973250
```

13. Create a plot which compares the mean weekly price of orange juice for all brands versus each individual brand.

```
weekly_brand_price<-tapply(oj$price,list(Week=oj$week,oj$brand),FUN=mean,na.rm=TRUE)
allbrand <- tapply(oj$price,list(Week=oj$week),FUN=mean,na.rm=TRUE)
df <- cbind(allbrand,weekly_brand_price)
write.table(df, file = 'q13.csv',sep = ",",row.names = T)
library(readr)
q13 <- read_csv("E:/Smurfit/Semester 2/DM/R/Data-Mining/q13.csv",
    col_names = FALSE, skip = 1)
```

```
## Parsed with column specification:
## cols(
##   X1 = col_integer(),
##   X2 = col_double(),
```

```
##   X3 = col_double(),
##   X4 = col_double(),
##   X5 = col_double()
## )
```

```
ggplot(q13, aes(q13$X1,xlab="Week")) +
  geom_line(aes(y = q13$X2, colour = "All Brand")) +
  geom_line(aes(y = q13$X3, colour = "Dominicks")) +
  geom_line(aes(y = q13$X4, colour = "Minute Maid")) +
  geom_line(aes(y = q13$X5, colour = "Tropicana")) + xlab("Week") +ylab("Price")+
  ggtitle("Weekly price of Orange Juice for All Brands vs Individual Brand")
```

```
individual_brand <- aggregate(oj[ ,"price"], list(Price=oj$price,Week= oj$week,Brand= oj$brand), FUN=mea
head(individual_brand)
```

```
##   Price Week     Brand    x
## 1  1.59   40 dominicks 1.59
## 2  1.85   41 dominicks 1.85
## 3  2.29   41 dominicks 2.29
## 4  2.49   41 dominicks 2.49
## 5  2.69   41 dominicks 2.69
## 6  1.85   42 dominicks 1.85
```

```
all_brand <- aggregate(oj[ ,"price"], list(Week=oj$week), mean)
head(all_brand)
```

```
##   Week        x
## 1   40 2.663105
```

```
## 2   41 2.650547
## 3   42 2.969950
## 4   43 2.954329
## 5   44 2.862554
## 6   45 2.972658
```

```r
typeof(individual_brand)
```

```
## [1] "list"
```

14.

```r
adcap <- factor(oj$feat)
tapply(oj$logmove,adcap,FUN=mean,na.rm=TRUE)
```

```
##         0         1
##  8.862963 10.148049
```