**[Test Information](https://blackboard.valpo.edu/webapps/assessment/take/launch.jsp?course_assessment_id=_16300_1&course_id=_13703_1&content_id=_440740_1&step=null" \o "Hide)**

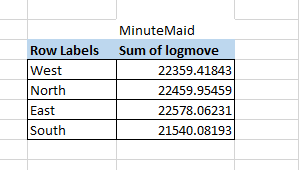
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Description | You are the data analyst on the regional management team for the orange juice company **shown in your question**.  Your juice is being sold in 83 stores in the greater Chicago area and is competing with two other brands.  Store numbers are organized as follows:   * 2-53:  West * 54-86:  North * 88-111:  East * 112-137:  South   Your sales manager has reviewed your analysis from the previous exercise, but isn't convinced yet that she can see how the sales per region are really different from one another; it is your job this week to investigate whether they are.  You will he using the same[dataset with weekly sales and demographic information](http://www.biz.uiowa.edu/faculty/jledolter/DataMining/oj.csv).  The attributes in your dataset are:   |  |  | | --- | --- | | store | Store number | | brand | OJ Brand | | week | Week in which advertisements are run | | logmove | Log units sold, normalized | | feat | Feature advertisement | | price | Retail price | | AGE60 | Percentage of the population that is aged 60 or older | | EDUC | Percentage of the population that has a college degree | | ETHNIC | Percentage of the population that is African-American or Hispanic | | INCOME | Median income | | HHLARGE | Percentage of households with 5 or more persons | | WORKWOM | Percentage of women with full-time jobs | | HVAL150 | Percentage of households worth more than $150,000 | | SSTRDIST | Distance to the nearest warehouse store (e.g. Walmart) | | SSTRVOL | Ratio of sales of this store to the nearest warehouse store (e.g. Walmart) | | CPDIST5 | Average distance in miles to the nearest 5 supermarkets | | CPWVOL5 | Ratio of sales of this store to the average of the nearest five stores | |

You work for the **Minute Maid** brand.  Your state sales manager is looking for more answers.

Here are the questions she needs you to answer:

1. Given all the regional sales for your brand, are the sales per region really that different from each other? (3 points)

Using Microsoft Excel, generated the sales ratio of minute maid of all the regions like from north, east, south and west region. In order to identify the sales in all region, used pivot table.



Yes, as per the pivot table, it identifies that sales in all region varies as per the value , but how ever the value is having is not having more difference ,but in south value is compared to be less than compared to be very less than compared to other regions. Taking internally all the factors which is internally dependent on the class attribute it defines that ratio of sales varies.

R Code :  
  
Goteti<- read\_excel("~/minute.xlsx")

library(nutshell)

library(lattice)

str(minute)

summary(minute)

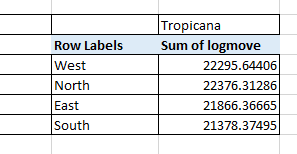
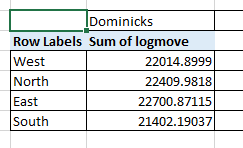
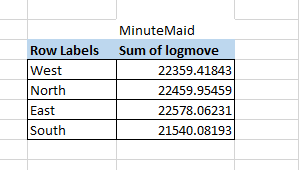
Output:

|  |
| --- |
| > Goteti<- read\_excel("~/minute.xlsx")  > library(nutshell)  Loading required package: nutshell.bbdb  Loading required package: nutshell.audioscrobbler  > library(lattice)  > str(minute)  Classes ‘tbl\_df’, ‘tbl’ and 'data.frame': 9649 obs. of 17 variables:  $ store : num 2 2 2 2 2 2 2 2 2 2 ...  $ brand : chr "minute.maid" "minute.maid" "minute.maid" "minute.maid" ...  $ week : num 40 46 47 48 50 51 52 53 54 57 ...  $ logmove : num 8.41 8.45 9.91 7.9 8.01 ...  $ feat : num 0 0 1 0 0 0 0 1 0 1 ...  $ price : num 3.17 3.17 2.39 3.17 3.17 3.17 3.17 2.19 3.17 1.99 ...  $ AGE60 : num 0.233 0.233 0.233 0.233 0.233 ...  $ EDUC : num 0.249 0.249 0.249 0.249 0.249 ...  $ ETHNIC : num 0.114 0.114 0.114 0.114 0.114 ...  $ INCOME : num 10.6 10.6 10.6 10.6 10.6 ...  $ HHLARGE : num 0.104 0.104 0.104 0.104 0.104 ...  $ WORKWOM : num 0.304 0.304 0.304 0.304 0.304 ...  $ HVAL150 : num 0.464 0.464 0.464 0.464 0.464 ...  $ SSTRDIST: num 2.11 2.11 2.11 2.11 2.11 ...  $ SSTRVOL : num 1.14 1.14 1.14 1.14 1.14 ...  $ CPDIST5 : num 1.93 1.93 1.93 1.93 1.93 ...  $ CPWVOL5 : num 0.377 0.377 0.377 0.377 0.377 ...  > summary(minute)  store brand week logmove feat  Min. : 2.00 Length:9649 Min. : 40.0 Min. : 5.768 Min. :0.0000  1st Qu.: 53.00 Class :character 1st Qu.: 70.0 1st Qu.: 8.476 1st Qu.:0.0000  Median : 86.00 Mode :character Median :101.0 Median : 9.026 Median :0.0000  Mean : 80.88 Mean :100.5 Mean : 9.217 Mean :0.2885  3rd Qu.:111.00 3rd Qu.:130.0 3rd Qu.: 9.829 3rd Qu.:1.0000  Max. :137.00 Max. :160.0 Max. :13.290 Max. :1.0000  price AGE60 EDUC ETHNIC INCOME  Min. :0.880 Min. :0.05805 Min. :0.04955 Min. :0.02425 Min. : 9.867  1st Qu.:1.990 1st Qu.:0.12210 1st Qu.:0.14598 1st Qu.:0.04191 1st Qu.:10.456  Median :2.170 Median :0.17065 Median :0.22939 Median :0.07466 Median :10.635  Mean :2.241 Mean :0.17313 Mean :0.22522 Mean :0.15556 Mean :10.617  3rd Qu.:2.490 3rd Qu.:0.21395 3rd Qu.:0.28439 3rd Qu.:0.18776 3rd Qu.:10.797  Max. :3.170 Max. :0.30740 Max. :0.52836 Max. :0.99569 Max. :11.236  HHLARGE WORKWOM HVAL150 SSTRDIST SSTRVOL  Min. :0.01351 Min. :0.2445 Min. :0.002509 Min. : 0.1321 Min. :0.4000  1st Qu.:0.09794 1st Qu.:0.3126 1st Qu.:0.123486 1st Qu.: 2.7670 1st Qu.:0.7273  Median :0.11122 Median :0.3556 Median :0.346154 Median : 4.6507 Median :1.1154  Mean :0.11560 Mean :0.3592 Mean :0.343766 Mean : 5.0973 Mean :1.2073  3rd Qu.:0.13517 3rd Qu.:0.4023 3rd Qu.:0.528313 3rd Qu.: 6.6506 3rd Qu.:1.5385  Max. :0.21635 Max. :0.4723 Max. :0.916700 Max. :17.8560 Max. :2.5714  CPDIST5 CPWVOL5  Min. :0.7725 Min. :0.09456  1st Qu.:1.6262 1st Qu.:0.27167  Median :1.9634 Median :0.38323  Mean :2.1204 Mean :0.43891  3rd Qu.:2.5337 3rd Qu.:0.56024  Max. :4.1079 Max. :1.14337 |
|  |
| |  | | --- | | > | |

Here the mean value of the internally dependent value would manipulate our and general sales manager ideology that using common sense we get if price is high and population is high, we need to have high price but value differs in this scenario, where mean and plot value varies as on the plot requires on the region where high sales varies as per the categorization.

1. Given all brand sales in all regions, does region really matter?  Make sure to provide a clear explanation that your state sales manager will understand with recommendations as to how to use the information in order to increase sales. (3 points)

Attaching the screenshot of all the regions of sales, and categorization the region as per the brand.

  
  
  
As per the images above displayed , we can get a categorical ideology about how the brand values varies with sales, and how the values of interdependent with class value.

Elaborating the description, treating LOGMOVE as a class attribute, we identify how to increase the sales in south region, as we got to a idea with the screenshot that the sales value in south region, and has to increase the sales as region value is dependent on internal values, and yes for that idealizing the example of correlation with the exact attribute of age and logmove we would resolve with positive relationship, which means it identifies that yes region really matters when compared with different brands as well.

R Code:

cor(Goteti$logmove,Goteti$AGE60,use="complete.obs",method=c("pearson"))

Output:  
> cor(Goteti$logmove,Goteti$AGE60,use="complete.obs",method=c("pearson"))

[1] 0.07975037

As per the region , we think south region sales are very less, and for that cor relation of keeping store far away from the people would not help in making great sales, it helps to have making with higher sales when sales of ratio of nearest super market would be having with in south with age it relates with positive co- relation so we suggest to have a nearest super market, and decrease the price to ensure people visit the store. In terms of publicity, we can focus on the age, super market near by with lesser price, so it would help in increasing the sales.

R Code :   
  
Goteti<- read\_excel("~/minute.xlsx")

library(nutshell)

library(lattice)

str(minute)

summary(minute)

cor(minute)

cor(Goteti$logmove,Goteti$AGE60,use="complete.obs",method=c("pearson"))

cor(Goteti$logmove,Goteti$brand,use="complete.obs",method=c("pearson"))

plot(logmove~SSTRDIST,data=minute)

boxplot(Goteti$logmove)

plot(SSTRVOL~SSTRDIST,data=minute)

plot(CPDIST5~SSTRDIST,data=minute)

plot(CPWVOL5~SSTRDIST,data=minute)

is.null(Goteti$logmove)

is.character(Goteti$logmove)

is.numeric(Goteti$logmove)

is.factor(Goteti$logmove)

is.numeric(Goteti)

Goteti$brand=as.numeric(Goteti$brand)

is.numeric(Goteti$brand)

summary(Goteti$brand)

barchart(Goteti$logmove)

m1=lm(brand~.,data=minute)

summary(m1)

m2=lm(SSTRDIST~.,data=minute)

summary(m2)

m3=lm(SSTRVOL~.,data=minute)

summary(m3)

m4=lm(CPDIST5~.,data=minute)

summary(m4)

m5=lm(CPWVOL5~.,data=minute)

summary(m5)

cor(minute)  
  
  
  
Output:

|  |
| --- |
| > Goteti<- read\_excel("~/minute.xlsx")  > library(nutshell)  > library(lattice)  > str(minute)  Classes ‘tbl\_df’, ‘tbl’ and 'data.frame': 9649 obs. of 17 variables:  $ store : num 2 2 2 2 2 2 2 2 2 2 ...  $ brand : chr "minute.maid" "minute.maid" "minute.maid" "minute.maid" ...  $ week : num 40 46 47 48 50 51 52 53 54 57 ...  $ logmove : num 8.41 8.45 9.91 7.9 8.01 ...  $ feat : num 0 0 1 0 0 0 0 1 0 1 ...  $ price : num 3.17 3.17 2.39 3.17 3.17 3.17 3.17 2.19 3.17 1.99 ...  $ AGE60 : num 0.233 0.233 0.233 0.233 0.233 ...  $ EDUC : num 0.249 0.249 0.249 0.249 0.249 ...  $ ETHNIC : num 0.114 0.114 0.114 0.114 0.114 ...  $ INCOME : num 10.6 10.6 10.6 10.6 10.6 ...  $ HHLARGE : num 0.104 0.104 0.104 0.104 0.104 ...  $ WORKWOM : num 0.304 0.304 0.304 0.304 0.304 ...  $ HVAL150 : num 0.464 0.464 0.464 0.464 0.464 ...  $ SSTRDIST: num 2.11 2.11 2.11 2.11 2.11 ...  $ SSTRVOL : num 1.14 1.14 1.14 1.14 1.14 ...  $ CPDIST5 : num 1.93 1.93 1.93 1.93 1.93 ...  $ CPWVOL5 : num 0.377 0.377 0.377 0.377 0.377 ...  > summary(minute)  store brand week logmove feat  Min. : 2.00 Length:9649 Min. : 40.0 Min. : 5.768 Min. :0.0000  1st Qu.: 53.00 Class :character 1st Qu.: 70.0 1st Qu.: 8.476 1st Qu.:0.0000  Median : 86.00 Mode :character Median :101.0 Median : 9.026 Median :0.0000  Mean : 80.88 Mean :100.5 Mean : 9.217 Mean :0.2885  3rd Qu.:111.00 3rd Qu.:130.0 3rd Qu.: 9.829 3rd Qu.:1.0000  Max. :137.00 Max. :160.0 Max. :13.290 Max. :1.0000  price AGE60 EDUC ETHNIC INCOME  Min. :0.880 Min. :0.05805 Min. :0.04955 Min. :0.02425 Min. : 9.867  1st Qu.:1.990 1st Qu.:0.12210 1st Qu.:0.14598 1st Qu.:0.04191 1st Qu.:10.456  Median :2.170 Median :0.17065 Median :0.22939 Median :0.07466 Median :10.635  Mean :2.241 Mean :0.17313 Mean :0.22522 Mean :0.15556 Mean :10.617  3rd Qu.:2.490 3rd Qu.:0.21395 3rd Qu.:0.28439 3rd Qu.:0.18776 3rd Qu.:10.797  Max. :3.170 Max. :0.30740 Max. :0.52836 Max. :0.99569 Max. :11.236  HHLARGE WORKWOM HVAL150 SSTRDIST SSTRVOL  Min. :0.01351 Min. :0.2445 Min. :0.002509 Min. : 0.1321 Min. :0.4000  1st Qu.:0.09794 1st Qu.:0.3126 1st Qu.:0.123486 1st Qu.: 2.7670 1st Qu.:0.7273  Median :0.11122 Median :0.3556 Median :0.346154 Median : 4.6507 Median :1.1154  Mean :0.11560 Mean :0.3592 Mean :0.343766 Mean : 5.0973 Mean :1.2073  3rd Qu.:0.13517 3rd Qu.:0.4023 3rd Qu.:0.528313 3rd Qu.: 6.6506 3rd Qu.:1.5385  Max. :0.21635 Max. :0.4723 Max. :0.916700 Max. :17.8560 Max. :2.5714  CPDIST5 CPWVOL5  Min. :0.7725 Min. :0.09456  1st Qu.:1.6262 1st Qu.:0.27167  Median :1.9634 Median :0.38323  Mean :2.1204 Mean :0.43891  3rd Qu.:2.5337 3rd Qu.:0.56024  Max. :4.1079 Max. :1.14337  > cor(minute)  Error in cor(minute) : 'x' must be numeric  > cor(Goteti$logmove,Goteti$AGE60,use="complete.obs",method=c("pearson"))  [1] 0.07975037  > cor(Goteti$logmove,Goteti$brand,use="complete.obs",method=c("pearson"))  Error in cor(Goteti$logmove, Goteti$brand, use = "complete.obs", method = c("pearson")) :  'y' must be numeric  > plot(logmove~SSTRDIST,data=minute)  > boxplot(Goteti$logmove)  > plot(SSTRVOL~SSTRDIST,data=minute)  > plot(CPDIST5~SSTRDIST,data=minute)  > plot(CPWVOL5~SSTRDIST,data=minute)  > is.null(Goteti$logmove)  [1] FALSE  > is.character(Goteti$logmove)  [1] FALSE  > is.numeric(Goteti$logmove)  [1] TRUE  > is.factor(Goteti$logmove)  [1] FALSE  > is.numeric(Goteti)  [1] FALSE  > Goteti$brand=as.numeric(Goteti$brand)  Warning message:  NAs introduced by coercion  > is.numeric(Goteti$brand)  [1] TRUE  > summary(Goteti$brand)  Min. 1st Qu. Median Mean 3rd Qu. Max. NA's  NA NA NA NaN NA NA 9649  > barchart(Goteti$logmove)  > m1=lm(brand~.,data=minute)  Error in lm.fit(x, y, offset = offset, singular.ok = singular.ok, ...) :  NA/NaN/Inf in 'y'  In addition: Warning message:  In model.response(mf, "numeric") : NAs introduced by coercion  > summary(m1)  Call:  lm(formula = GPM ~ ., data = FuelEff)  Residuals:  Min 1Q Median 3Q Max  -0.33622 -0.13499 -0.00486 0.08701 0.48325  Coefficients:  Estimate Std. Error t value Pr(>|t|)  (Intercept) 4.9864020 1.1309678 4.409 0.000123 \*\*\*  MPG -0.1151678 0.0160110 -7.193 5.27e-08 \*\*\*  WT 0.0860549 0.2949136 0.292 0.772450  DIS -0.0003056 0.0017785 -0.172 0.864705  NC 0.2152198 0.0819830 2.625 0.013497 \*  HP 0.0076517 0.0047066 1.626 0.114471  ACC 0.0136109 0.0282887 0.481 0.633905  ET -0.4770695 0.1774648 -2.688 0.011608 \*  ---  Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1  Residual standard error: 0.1928 on 30 degrees of freedom  Multiple R-squared: 0.9775, Adjusted R-squared: 0.9722  F-statistic: 185.8 on 7 and 30 DF, p-value: < 2.2e-16  > m2=lm(SSTRDIST~.,data=minute)  Error in `contrasts<-`(`\*tmp\*`, value = contr.funs[1 + isOF[nn]]) :  contrasts can be applied only to factors with 2 or more levels  > summary(m2)  Call:  lm(formula = GPM ~ WT, data = FuelEff[train, ])  Residuals:  Min 1Q Median 3Q Max  -0.88189 -0.31582 -0.01301 0.21347 1.09266  Coefficients:  Estimate Std. Error t value Pr(>|t|)  (Intercept) -0.1704 0.3100 -0.55 0.586  WT 1.5795 0.1069 14.77 <2e-16 \*\*\*  ---  Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1  Residual standard error: 0.4302 on 35 degrees of freedom  Multiple R-squared: 0.8618, Adjusted R-squared: 0.8578  F-statistic: 218.2 on 1 and 35 DF, p-value: < 2.2e-16  > m3=lm(SSTRVOL~.,data=minute)  Error in `contrasts<-`(`\*tmp\*`, value = contr.funs[1 + isOF[nn]]) :  contrasts can be applied only to factors with 2 or more levels  > summary(m3)  Error in summary(m3) : object 'm3' not found  > m4=lm(CPDIST5~.,data=minute)  Error in `contrasts<-`(`\*tmp\*`, value = contr.funs[1 + isOF[nn]]) :  contrasts can be applied only to factors with 2 or more levels  > summary(m4)  Error in summary(m4) : object 'm4' not found  > m5=lm(CPWVOL5~.,data=minute)  Error in `contrasts<-`(`\*tmp\*`, value = contr.funs[1 + isOF[nn]]) :  contrasts can be applied only to factors with 2 or more levels  > summary(m5)  Error in summary(m5) : object 'm5' not found |
|  |
| |  | | --- | | > | |

3.In what region do you see the largest potential for sales increase?   Explain the thinking steps and show the R code you used to determine this.  Make sure to provide an answer to your question and recommendations for sales strategies that your state sales manager will understand. (4 points)

Please include screenshots of your R code and its outcomes.  Remember that the results of your work and the sales strategy recommendations must be interpreted in language that the state sales manager can understand.

Largest potential of sales increase are in South Region for minute maid Brand. It is identified with R and Excel, so get more clear idea , is the value of class and internal attributes are they internally related or not.   
  
  
R Code:   
  
Goteti<- read\_excel("~/minute.xlsx")

library(nutshell)

library(lattice)

str(minute)

summary(minute)

cor(minute)

cor(Goteti$logmove,Goteti$AGE60,use="complete.obs",method=c("pearson"))

cor(Goteti$logmove,Goteti$brand,use="complete.obs",method=c("pearson"))

plot(logmove~SSTRDIST,data=minute)

boxplot(Goteti$logmove)

plot(SSTRVOL~SSTRDIST,data=minute)

plot(CPDIST5~SSTRDIST,data=minute)

plot(CPWVOL5~SSTRDIST,data=minute)

is.null(Goteti$logmove)

is.character(Goteti$logmove)

is.numeric(Goteti$logmove)

is.factor(Goteti$logmove)

is.numeric(Goteti)

Goteti$brand=as.numeric(Goteti$brand)

is.numeric(Goteti$brand)

summary(Goteti$brand)

barchart(Goteti$logmove)

m1=lm(brand~.,data=minute)

summary(m1)

m2=lm(SSTRDIST~.,data=minute)

summary(m2)

m3=lm(SSTRVOL~.,data=minute)

summary(m3)

m4=lm(CPDIST5~.,data=minute)

summary(m4)

m5=lm(CPWVOL5~.,data=minute)

summary(m5)

cor(minute)

n=length(Goteti$brand)

diff=dim(n)

percdiff=dim(n)

for (k in 1:n) {

train1=c(1:n)

train=train1[train1!=k]

m1=lm(brand~.,data=Goteti[train,])

pred=predict(m1,newdat=Goteti[-train,])

obs=Goteti$brand[-train]

diff[k]=obs-pred

percdiff[k]=abs(diff[k])/obs

}

n1=length(Goteti$logmove)

diff=dim(n1)

percdiff=dim(n1)

for (k in 1:n) {

train1=c(1:n)

train=train1[train1!=k]

m2=lm(logmove~.,data=Goteti[train,])

pred=predict(m2,newdat=Goteti[-train,])

obs=Goteti$brand[-train]

diff[k]=obs-pred

percdiff[k]=abs(diff[k])/obs

}

filled.contour(volcano, color=terrain.colors, asp=1, plot.axes=contour(volcano, add=T))

persp(volcano, theta=25, phi=30, expand=0.5, col="blue")

parallelplot(~Goteti[1:4] | logmove, data=Goteti)

median(Goteti$logmove,na.rm=TRUE)

cor(Goteti)

me=mean(Goteti)

rmse=sqrt(mean(Goteti\*\*2))

mape=100\*(mean(percGoteti))

me # mean error

rmse # root mean square error

mape

data1=data[minus-2]

z<-lm(logmove~.,data=Goteti)

summary(z)  
  
  
Output:

|  |
| --- |
| > z=lm(Goteti)  Error in stats::model.frame(formula = Goteti, drop.unused.levels = TRUE) :  object 'Goteti' not found  > summary(z)  Error in summary(z) : object 'z' not found  > z=lm(logmove~.,data=Goteti)  Error in is.data.frame(data) : object 'Goteti' not found  > summary(z)  Error in summary(z) : object 'z' not found  > z<-lm(logmove~.,data=Goteti)  Error in is.data.frame(data) : object 'Goteti' not found  > Goteti<- read\_excel("~/minute.xlsx")  Error: could not find function "read\_excel"  > Goteti<- read\_excel("~/minute.xlsx")  Error: could not find function "read\_excel"  > library(readxl)  > minute <- read\_excel("~/minute.xlsx")  > View(minute)  > Goteti<- read\_excel("~/minute.xlsx")  > z<-lm(logmove~.,data=Goteti)  Error in `contrasts<-`(`\*tmp\*`, value = contr.funs[1 + isOF[nn]]) :  contrasts can be applied only to factors with 2 or more levels  > z<-lm(logmove~.,data=Goteti)  Error in `contrasts<-`(`\*tmp\*`, value = contr.funs[1 + isOF[nn]]) :  contrasts can be applied only to factors with 2 or more levels  > z=lm(logmove~.,data=Goteti)  Error in `contrasts<-`(`\*tmp\*`, value = contr.funs[1 + isOF[nn]]) :  contrasts can be applied only to factors with 2 or more levels  > data2<-Goteti[,-2]  > z=lm(logmove~.,data=data2)  > summary(z)  Call:  lm(formula = logmove ~ ., data = data2)  Residuals:  Min 1Q Median 3Q Max  -2.61853 -0.36032 -0.01696 0.32358 2.91046  Coefficients:  Estimate Std. Error t value Pr(>|t|)  (Intercept) 12.7429558 0.5051817 25.224 < 2e-16 \*\*\*  store 0.0006338 0.0001919 3.303 0.000961 \*\*\*  week -0.0035076 0.0001731 -20.261 < 2e-16 \*\*\*  feat 1.1179707 0.0138326 80.822 < 2e-16 \*\*\*  price -1.1348268 0.0160704 -70.616 < 2e-16 \*\*\*  AGE60 2.0619217 0.1942354 10.616 < 2e-16 \*\*\*  EDUC 0.6167503 0.1551242 3.976 7.06e-05 \*\*\*  ETHNIC 0.5989527 0.0571027 10.489 < 2e-16 \*\*\*  INCOME -0.1352242 0.0508801 -2.658 0.007881 \*\*  HHLARGE -0.5017965 0.3521707 -1.425 0.154228  WORKWOM -0.2672190 0.2239255 -1.193 0.232766  HVAL150 0.5565707 0.0638732 8.714 < 2e-16 \*\*\*  SSTRDIST -0.0078095 0.0022572 -3.460 0.000543 \*\*\*  SSTRVOL -0.0146395 0.0149633 -0.978 0.327920  CPDIST5 0.0361835 0.0096180 3.762 0.000170 \*\*\*  CPWVOL5 -0.4871137 0.0402429 -12.104 < 2e-16 \*\*\*  ---  Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1  Residual standard error: 0.5773 on 9633 degrees of freedom  Multiple R-squared: 0.6572, Adjusted R-squared: 0.6567  F-statistic: 1231 on 15 and 9633 DF, p-value: < 2.2e-16  > Goteti<- read\_excel("~/minute.xlsx")  > library(nutshell)  Loading required package: nutshell.bbdb  Loading required package: nutshell.audioscrobbler  > library(lattice)  > str(minute)  Classes ‘tbl\_df’, ‘tbl’ and 'data.frame': 9649 obs. of 17 variables:  $ store : num 2 2 2 2 2 2 2 2 2 2 ...  $ brand : chr "minute.maid" "minute.maid" "minute.maid" "minute.maid" ...  $ week : num 40 46 47 48 50 51 52 53 54 57 ...  $ logmove : num 8.41 8.45 9.91 7.9 8.01 ...  $ feat : num 0 0 1 0 0 0 0 1 0 1 ...  $ price : num 3.17 3.17 2.39 3.17 3.17 3.17 3.17 2.19 3.17 1.99 ...  $ AGE60 : num 0.233 0.233 0.233 0.233 0.233 ...  $ EDUC : num 0.249 0.249 0.249 0.249 0.249 ...  $ ETHNIC : num 0.114 0.114 0.114 0.114 0.114 ...  $ INCOME : num 10.6 10.6 10.6 10.6 10.6 ...  $ HHLARGE : num 0.104 0.104 0.104 0.104 0.104 ...  $ WORKWOM : num 0.304 0.304 0.304 0.304 0.304 ...  $ HVAL150 : num 0.464 0.464 0.464 0.464 0.464 ...  $ SSTRDIST: num 2.11 2.11 2.11 2.11 2.11 ...  $ SSTRVOL : num 1.14 1.14 1.14 1.14 1.14 ...  $ CPDIST5 : num 1.93 1.93 1.93 1.93 1.93 ...  $ CPWVOL5 : num 0.377 0.377 0.377 0.377 0.377 ...  > summary(minute)  store brand week logmove feat  Min. : 2.00 Length:9649 Min. : 40.0 Min. : 5.768 Min. :0.0000  1st Qu.: 53.00 Class :character 1st Qu.: 70.0 1st Qu.: 8.476 1st Qu.:0.0000  Median : 86.00 Mode :character Median :101.0 Median : 9.026 Median :0.0000  Mean : 80.88 Mean :100.5 Mean : 9.217 Mean :0.2885  3rd Qu.:111.00 3rd Qu.:130.0 3rd Qu.: 9.829 3rd Qu.:1.0000  Max. :137.00 Max. :160.0 Max. :13.290 Max. :1.0000  price AGE60 EDUC ETHNIC INCOME  Min. :0.880 Min. :0.05805 Min. :0.04955 Min. :0.02425 Min. : 9.867  1st Qu.:1.990 1st Qu.:0.12210 1st Qu.:0.14598 1st Qu.:0.04191 1st Qu.:10.456  Median :2.170 Median :0.17065 Median :0.22939 Median :0.07466 Median :10.635  Mean :2.241 Mean :0.17313 Mean :0.22522 Mean :0.15556 Mean :10.617  3rd Qu.:2.490 3rd Qu.:0.21395 3rd Qu.:0.28439 3rd Qu.:0.18776 3rd Qu.:10.797  Max. :3.170 Max. :0.30740 Max. :0.52836 Max. :0.99569 Max. :11.236  HHLARGE WORKWOM HVAL150 SSTRDIST SSTRVOL  Min. :0.01351 Min. :0.2445 Min. :0.002509 Min. : 0.1321 Min. :0.4000  1st Qu.:0.09794 1st Qu.:0.3126 1st Qu.:0.123486 1st Qu.: 2.7670 1st Qu.:0.7273  Median :0.11122 Median :0.3556 Median :0.346154 Median : 4.6507 Median :1.1154  Mean :0.11560 Mean :0.3592 Mean :0.343766 Mean : 5.0973 Mean :1.2073  3rd Qu.:0.13517 3rd Qu.:0.4023 3rd Qu.:0.528313 3rd Qu.: 6.6506 3rd Qu.:1.5385  Max. :0.21635 Max. :0.4723 Max. :0.916700 Max. :17.8560 Max. :2.5714  CPDIST5 CPWVOL5  Min. :0.7725 Min. :0.09456  1st Qu.:1.6262 1st Qu.:0.27167  Median :1.9634 Median :0.38323  Mean :2.1204 Mean :0.43891  3rd Qu.:2.5337 3rd Qu.:0.56024  Max. :4.1079 Max. :1.14337  > cor(Goteti$logmove,Goteti$AGE60,use="complete.obs",method=c("pearson"))  [1] 0.07975037  > cor(Goteti$logmove,Goteti$brand,use="complete.obs",method=c("pearson"))  Error in cor(Goteti$logmove, Goteti$brand, use = "complete.obs", method = c("pearson")) :  'y' must be numeric  > Goteti<- read\_excel("~/minute.xlsx")  > library(nutshell)  > library(lattice)  > str(minute)  Classes ‘tbl\_df’, ‘tbl’ and 'data.frame': 9649 obs. of 17 variables:  $ store : num 2 2 2 2 2 2 2 2 2 2 ...  $ brand : chr "minute.maid" "minute.maid" "minute.maid" "minute.maid" ...  $ week : num 40 46 47 48 50 51 52 53 54 57 ...  $ logmove : num 8.41 8.45 9.91 7.9 8.01 ...  $ feat : num 0 0 1 0 0 0 0 1 0 1 ...  $ price : num 3.17 3.17 2.39 3.17 3.17 3.17 3.17 2.19 3.17 1.99 ...  $ AGE60 : num 0.233 0.233 0.233 0.233 0.233 ...  $ EDUC : num 0.249 0.249 0.249 0.249 0.249 ...  $ ETHNIC : num 0.114 0.114 0.114 0.114 0.114 ...  $ INCOME : num 10.6 10.6 10.6 10.6 10.6 ...  $ HHLARGE : num 0.104 0.104 0.104 0.104 0.104 ...  $ WORKWOM : num 0.304 0.304 0.304 0.304 0.304 ...  $ HVAL150 : num 0.464 0.464 0.464 0.464 0.464 ...  $ SSTRDIST: num 2.11 2.11 2.11 2.11 2.11 ...  $ SSTRVOL : num 1.14 1.14 1.14 1.14 1.14 ...  $ CPDIST5 : num 1.93 1.93 1.93 1.93 1.93 ...  $ CPWVOL5 : num 0.377 0.377 0.377 0.377 0.377 ...  > summary(minute)  store brand week logmove feat  Min. : 2.00 Length:9649 Min. : 40.0 Min. : 5.768 Min. :0.0000  1st Qu.: 53.00 Class :character 1st Qu.: 70.0 1st Qu.: 8.476 1st Qu.:0.0000  Median : 86.00 Mode :character Median :101.0 Median : 9.026 Median :0.0000  Mean : 80.88 Mean :100.5 Mean : 9.217 Mean :0.2885  3rd Qu.:111.00 3rd Qu.:130.0 3rd Qu.: 9.829 3rd Qu.:1.0000  Max. :137.00 Max. :160.0 Max. :13.290 Max. :1.0000  price AGE60 EDUC ETHNIC INCOME  Min. :0.880 Min. :0.05805 Min. :0.04955 Min. :0.02425 Min. : 9.867  1st Qu.:1.990 1st Qu.:0.12210 1st Qu.:0.14598 1st Qu.:0.04191 1st Qu.:10.456  Median :2.170 Median :0.17065 Median :0.22939 Median :0.07466 Median :10.635  Mean :2.241 Mean :0.17313 Mean :0.22522 Mean :0.15556 Mean :10.617  3rd Qu.:2.490 3rd Qu.:0.21395 3rd Qu.:0.28439 3rd Qu.:0.18776 3rd Qu.:10.797  Max. :3.170 Max. :0.30740 Max. :0.52836 Max. :0.99569 Max. :11.236  HHLARGE WORKWOM HVAL150 SSTRDIST SSTRVOL  Min. :0.01351 Min. :0.2445 Min. :0.002509 Min. : 0.1321 Min. :0.4000  1st Qu.:0.09794 1st Qu.:0.3126 1st Qu.:0.123486 1st Qu.: 2.7670 1st Qu.:0.7273  Median :0.11122 Median :0.3556 Median :0.346154 Median : 4.6507 Median :1.1154  Mean :0.11560 Mean :0.3592 Mean :0.343766 Mean : 5.0973 Mean :1.2073  3rd Qu.:0.13517 3rd Qu.:0.4023 3rd Qu.:0.528313 3rd Qu.: 6.6506 3rd Qu.:1.5385  Max. :0.21635 Max. :0.4723 Max. :0.916700 Max. :17.8560 Max. :2.5714  CPDIST5 CPWVOL5  Min. :0.7725 Min. :0.09456  1st Qu.:1.6262 1st Qu.:0.27167  Median :1.9634 Median :0.38323  Mean :2.1204 Mean :0.43891  3rd Qu.:2.5337 3rd Qu.:0.56024  Max. :4.1079 Max. :1.14337  > cor(minute)  Error in cor(minute) : 'x' must be numeric  > cor(Goteti$logmove,Goteti$AGE60,use="complete.obs",method=c("pearson"))  [1] 0.07975037  > cor(Goteti$logmove,Goteti$brand,use="complete.obs",method=c("pearson"))  Error in cor(Goteti$logmove, Goteti$brand, use = "complete.obs", method = c("pearson")) :  'y' must be numeric  > plot(logmove~SSTRDIST,data=minute)  > boxplot(Goteti$logmove)  > plot(SSTRVOL~SSTRDIST,data=minute)  > plot(CPDIST5~SSTRDIST,data=minute)  > plot(CPWVOL5~SSTRDIST,data=minute)  > is.null(Goteti$logmove)  [1] FALSE  > is.character(Goteti$logmove)  [1] FALSE  > is.numeric(Goteti$logmove)  [1] TRUE  > is.factor(Goteti$logmove)  [1] FALSE  > is.numeric(Goteti)  [1] FALSE  > Goteti$brand=as.numeric(Goteti$brand)  Warning message:  NAs introduced by coercion  > is.numeric(Goteti$brand)  [1] TRUE  > summary(Goteti$brand)  Min. 1st Qu. Median Mean 3rd Qu. Max. NA's  NA NA NA NaN NA NA 9649  > barchart(Goteti$logmove)  > m1=lm(brand~.,data=minute)  Error in lm.fit(x, y, offset = offset, singular.ok = singular.ok, ...) :  NA/NaN/Inf in 'y'  In addition: Warning message:  In model.response(mf, "numeric") : NAs introduced by coercion  > summary(m1)  Call:  lm(formula = GPM ~ ., data = FuelEff)  Residuals:  Min 1Q Median 3Q Max  -0.33622 -0.13499 -0.00486 0.08701 0.48325  Coefficients:  Estimate Std. Error t value Pr(>|t|)  (Intercept) 4.9864020 1.1309678 4.409 0.000123 \*\*\*  MPG -0.1151678 0.0160110 -7.193 5.27e-08 \*\*\*  WT 0.0860549 0.2949136 0.292 0.772450  DIS -0.0003056 0.0017785 -0.172 0.864705  NC 0.2152198 0.0819830 2.625 0.013497 \*  HP 0.0076517 0.0047066 1.626 0.114471  ACC 0.0136109 0.0282887 0.481 0.633905  ET -0.4770695 0.1774648 -2.688 0.011608 \*  ---  Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1  Residual standard error: 0.1928 on 30 degrees of freedom  Multiple R-squared: 0.9775, Adjusted R-squared: 0.9722  F-statistic: 185.8 on 7 and 30 DF, p-value: < 2.2e-16  > m2=lm(SSTRDIST~.,data=minute)  Error in `contrasts<-`(`\*tmp\*`, value = contr.funs[1 + isOF[nn]]) :  contrasts can be applied only to factors with 2 or more levels  > summary(m2)  Call:  lm(formula = GPM ~ WT, data = FuelEff[train, ])  Residuals:  Min 1Q Median 3Q Max  -0.88189 -0.31582 -0.01301 0.21347 1.09266  Coefficients:  Estimate Std. Error t value Pr(>|t|)  (Intercept) -0.1704 0.3100 -0.55 0.586  WT 1.5795 0.1069 14.77 <2e-16 \*\*\*  ---  Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1  Residual standard error: 0.4302 on 35 degrees of freedom  Multiple R-squared: 0.8618, Adjusted R-squared: 0.8578  F-statistic: 218.2 on 1 and 35 DF, p-value: < 2.2e-16  > m3=lm(SSTRVOL~.,data=minute)  Error in `contrasts<-`(`\*tmp\*`, value = contr.funs[1 + isOF[nn]]) :  contrasts can be applied only to factors with 2 or more levels  > summary(m3)  Error in summary(m3) : object 'm3' not found  > m4=lm(CPDIST5~.,data=minute)  Error in `contrasts<-`(`\*tmp\*`, value = contr.funs[1 + isOF[nn]]) :  contrasts can be applied only to factors with 2 or more levels  > summary(m4)  Error in summary(m4) : object 'm4' not found  > m5=lm(CPWVOL5~.,data=minute)  Error in `contrasts<-`(`\*tmp\*`, value = contr.funs[1 + isOF[nn]]) :  contrasts can be applied only to factors with 2 or more levels  > summary(m5)  Error in summary(m5) : object 'm5' not found  > Goteti<- read\_excel("~/minute.xlsx")  > library(nutshell)  > library(lattice)  > str(minute)  Classes ‘tbl\_df’, ‘tbl’ and 'data.frame': 9649 obs. of 17 variables:  $ store : num 2 2 2 2 2 2 2 2 2 2 ...  $ brand : chr "minute.maid" "minute.maid" "minute.maid" "minute.maid" ...  $ week : num 40 46 47 48 50 51 52 53 54 57 ...  $ logmove : num 8.41 8.45 9.91 7.9 8.01 ...  $ feat : num 0 0 1 0 0 0 0 1 0 1 ...  $ price : num 3.17 3.17 2.39 3.17 3.17 3.17 3.17 2.19 3.17 1.99 ...  $ AGE60 : num 0.233 0.233 0.233 0.233 0.233 ...  $ EDUC : num 0.249 0.249 0.249 0.249 0.249 ...  $ ETHNIC : num 0.114 0.114 0.114 0.114 0.114 ...  $ INCOME : num 10.6 10.6 10.6 10.6 10.6 ...  $ HHLARGE : num 0.104 0.104 0.104 0.104 0.104 ...  $ WORKWOM : num 0.304 0.304 0.304 0.304 0.304 ...  $ HVAL150 : num 0.464 0.464 0.464 0.464 0.464 ...  $ SSTRDIST: num 2.11 2.11 2.11 2.11 2.11 ...  $ SSTRVOL : num 1.14 1.14 1.14 1.14 1.14 ...  $ CPDIST5 : num 1.93 1.93 1.93 1.93 1.93 ...  $ CPWVOL5 : num 0.377 0.377 0.377 0.377 0.377 ...  > summary(minute)  store brand week logmove feat  Min. : 2.00 Length:9649 Min. : 40.0 Min. : 5.768 Min. :0.0000  1st Qu.: 53.00 Class :character 1st Qu.: 70.0 1st Qu.: 8.476 1st Qu.:0.0000  Median : 86.00 Mode :character Median :101.0 Median : 9.026 Median :0.0000  Mean : 80.88 Mean :100.5 Mean : 9.217 Mean :0.2885  3rd Qu.:111.00 3rd Qu.:130.0 3rd Qu.: 9.829 3rd Qu.:1.0000  Max. :137.00 Max. :160.0 Max. :13.290 Max. :1.0000  price AGE60 EDUC ETHNIC INCOME  Min. :0.880 Min. :0.05805 Min. :0.04955 Min. :0.02425 Min. : 9.867  1st Qu.:1.990 1st Qu.:0.12210 1st Qu.:0.14598 1st Qu.:0.04191 1st Qu.:10.456  Median :2.170 Median :0.17065 Median :0.22939 Median :0.07466 Median :10.635  Mean :2.241 Mean :0.17313 Mean :0.22522 Mean :0.15556 Mean :10.617  3rd Qu.:2.490 3rd Qu.:0.21395 3rd Qu.:0.28439 3rd Qu.:0.18776 3rd Qu.:10.797  Max. :3.170 Max. :0.30740 Max. :0.52836 Max. :0.99569 Max. :11.236  HHLARGE WORKWOM HVAL150 SSTRDIST SSTRVOL  Min. :0.01351 Min. :0.2445 Min. :0.002509 Min. : 0.1321 Min. :0.4000  1st Qu.:0.09794 1st Qu.:0.3126 1st Qu.:0.123486 1st Qu.: 2.7670 1st Qu.:0.7273  Median :0.11122 Median :0.3556 Median :0.346154 Median : 4.6507 Median :1.1154  Mean :0.11560 Mean :0.3592 Mean :0.343766 Mean : 5.0973 Mean :1.2073  3rd Qu.:0.13517 3rd Qu.:0.4023 3rd Qu.:0.528313 3rd Qu.: 6.6506 3rd Qu.:1.5385  Max. :0.21635 Max. :0.4723 Max. :0.916700 Max. :17.8560 Max. :2.5714  CPDIST5 CPWVOL5  Min. :0.7725 Min. :0.09456  1st Qu.:1.6262 1st Qu.:0.27167  Median :1.9634 Median :0.38323  Mean :2.1204 Mean :0.43891  3rd Qu.:2.5337 3rd Qu.:0.56024  Max. :4.1079 Max. :1.14337  > cor(minute)  Error in cor(minute) : 'x' must be numeric  > cor(Goteti$logmove,Goteti$AGE60,use="complete.obs",method=c("pearson"))  [1] 0.07975037  > cor(Goteti$logmove,Goteti$brand,use="complete.obs",method=c("pearson"))  Error in cor(Goteti$logmove, Goteti$brand, use = "complete.obs", method = c("pearson")) :  'y' must be numeric  > plot(logmove~SSTRDIST,data=minute)  > boxplot(Goteti$logmove)  > plot(SSTRVOL~SSTRDIST,data=minute)  > plot(CPDIST5~SSTRDIST,data=minute)  > plot(CPWVOL5~SSTRDIST,data=minute)  > is.null(Goteti$logmove)  [1] FALSE  > is.character(Goteti$logmove)  [1] FALSE  > is.numeric(Goteti$logmove)  [1] TRUE  > is.factor(Goteti$logmove)  [1] FALSE  > is.numeric(Goteti)  [1] FALSE  > Goteti$brand=as.numeric(Goteti$brand)  Warning message:  NAs introduced by coercion  > is.numeric(Goteti$brand)  [1] TRUE  > summary(Goteti$brand)  Min. 1st Qu. Median Mean 3rd Qu. Max. NA's  NA NA NA NaN NA NA 9649  > barchart(Goteti$logmove)  > m1=lm(brand~.,data=minute)  Error in lm.fit(x, y, offset = offset, singular.ok = singular.ok, ...) :  NA/NaN/Inf in 'y'  In addition: Warning message:  In model.response(mf, "numeric") : NAs introduced by coercion  > summary(m1)  Call:  lm(formula = GPM ~ ., data = FuelEff)  Residuals:  Min 1Q Median 3Q Max  -0.33622 -0.13499 -0.00486 0.08701 0.48325  Coefficients:  Estimate Std. Error t value Pr(>|t|)  (Intercept) 4.9864020 1.1309678 4.409 0.000123 \*\*\*  MPG -0.1151678 0.0160110 -7.193 5.27e-08 \*\*\*  WT 0.0860549 0.2949136 0.292 0.772450  DIS -0.0003056 0.0017785 -0.172 0.864705  NC 0.2152198 0.0819830 2.625 0.013497 \*  HP 0.0076517 0.0047066 1.626 0.114471  ACC 0.0136109 0.0282887 0.481 0.633905  ET -0.4770695 0.1774648 -2.688 0.011608 \*  ---  Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1  Residual standard error: 0.1928 on 30 degrees of freedom  Multiple R-squared: 0.9775, Adjusted R-squared: 0.9722  F-statistic: 185.8 on 7 and 30 DF, p-value: < 2.2e-16  > m2=lm(SSTRDIST~.,data=minute)  Error in `contrasts<-`(`\*tmp\*`, value = contr.funs[1 + isOF[nn]]) :  contrasts can be applied only to factors with 2 or more levels  > summary(m2)  Call:  lm(formula = GPM ~ WT, data = FuelEff[train, ])  Residuals:  Min 1Q Median 3Q Max  -0.88189 -0.31582 -0.01301 0.21347 1.09266  Coefficients:  Estimate Std. Error t value Pr(>|t|)  (Intercept) -0.1704 0.3100 -0.55 0.586  WT 1.5795 0.1069 14.77 <2e-16 \*\*\*  ---  Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1  Residual standard error: 0.4302 on 35 degrees of freedom  Multiple R-squared: 0.8618, Adjusted R-squared: 0.8578  F-statistic: 218.2 on 1 and 35 DF, p-value: < 2.2e-16  > m3=lm(SSTRVOL~.,data=minute)  Error in `contrasts<-`(`\*tmp\*`, value = contr.funs[1 + isOF[nn]]) :  contrasts can be applied only to factors with 2 or more levels  > summary(m3)  Error in summary(m3) : object 'm3' not found  > me=mean(Goteti)  Warning message:  In mean.default(Goteti) : argument is not numeric or logical: returning NA  > rmse=sqrt(mean(Goteti\*\*2))  > mape=100\*(mean(percGoteti))  Error in mean(percGoteti) : object 'percGoteti' not found  > me # mean error  [1] NA  > rmse # root mean square error  [1] NA  > mape  [1] NA  > data1=data[minus-2]  Error: object 'minus' not found  > z<-lm(logmove~.,data=Goteti)  Error in lm.fit(x, y, offset = offset, singular.ok = singular.ok, ...) :  0 (non-NA) cases  > summary(z)  Call:  lm(formula = logmove ~ ., data = data2)  Residuals:  Min 1Q Median 3Q Max  -2.61853 -0.36032 -0.01696 0.32358 2.91046  Coefficients:  Estimate Std. Error t value Pr(>|t|)  (Intercept) 12.7429558 0.5051817 25.224 < 2e-16 \*\*\*  store 0.0006338 0.0001919 3.303 0.000961 \*\*\*  week -0.0035076 0.0001731 -20.261 < 2e-16 \*\*\*  feat 1.1179707 0.0138326 80.822 < 2e-16 \*\*\*  price -1.1348268 0.0160704 -70.616 < 2e-16 \*\*\*  AGE60 2.0619217 0.1942354 10.616 < 2e-16 \*\*\*  EDUC 0.6167503 0.1551242 3.976 7.06e-05 \*\*\*  ETHNIC 0.5989527 0.0571027 10.489 < 2e-16 \*\*\*  INCOME -0.1352242 0.0508801 -2.658 0.007881 \*\*  HHLARGE -0.5017965 0.3521707 -1.425 0.154228  WORKWOM -0.2672190 0.2239255 -1.193 0.232766  HVAL150 0.5565707 0.0638732 8.714 < 2e-16 \*\*\*  SSTRDIST -0.0078095 0.0022572 -3.460 0.000543 \*\*\*  SSTRVOL -0.0146395 0.0149633 -0.978 0.327920  CPDIST5 0.0361835 0.0096180 3.762 0.000170 \*\*\*  CPWVOL5 -0.4871137 0.0402429 -12.104 < 2e-16 \*\*\*  ---  Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1  Residual standard error: 0.5773 on 9633 degrees of freedom  Multiple R-squared: 0.6572, Adjusted R-squared: 0.6567  F-statistic: 1231 on 15 and 9633 DF, p-value: < 2.2e-16  >  > |
|  |
| |  | | --- | | > | |

By the above point we identify that , sales in south region are very less than compared to other regions, so we would recommend the sales manager to decrease the price, as the graph and regression identifies that price is inversely proportional to log move which would make the sales in the region, and population and store near by would become inter related so increase sales in the south region.  
  
Now it the choice of sales manager to take a step to increase with the idea we had given on this.

1. Increase the sales in south region, by decreasing price, and bringing the store near to the people, and can increase publicity.
2. Or else, he can focus on other region where the sales are already high, so taking some simple step can help to increase more way having region of store, and make more publicity about the product and the store.