Balance Prediction Linear Regression

Niranjan

21 December 2018

Loading the required packages

Importing the case file and examining it

```
setwd("I:/DA/To Submit/Linear Regression Balance Prediction/")
cust_data <- fread("Cust_Database.csv")
str(cust_data)
```

```
## Classes 'data.table' and 'data.frame':   5000 obs. of  132 variables:
##  $ custid           : chr  "3964-QJWTRG-NPN" "0648-AIPJSP-UVM" "5195-TLUDJE-HVO" "4459-VLPQUH-3OL" .
##  $ region           : num  1 5 3 4 2 4 2 3 2 2 ...
##  $ townsize         : num  2 5 4 3 2 4 5 4 3 2 ...
##  $ gender           : num  1 0 1 0 0 0 1 1 1 0 ...
##  $ age              : num  20 22 67 23 26 64 52 44 66 47 ...
##  $ agecat           : num  2 2 6 2 3 5 5 4 6 4 ...
##  $ birthmonth       : chr  "September" "May" "June" "May" ...
##  $ ed               : num  15 17 14 16 16 17 14 16 12 11 ...
##  $ edcat            : num  3 4 2 3 3 4 2 3 2 1 ...
##  $ jobcat           : num  1 2 2 2 2 3 1 1 1 6 ...
##  $ union            : num  1 0 0 0 0 0 0 0 0 0 ...
##  $ employ           : num  0 0 16 0 1 22 10 11 15 19 ...
##  $ empcat           : num  1 1 5 1 1 5 3 4 4 5 ...
##  $ retire           : num  0 0 0 0 0 0 0 0 1 0 ...
##  $ income           : num  31 15 35 20 23 107 77 97 16 84 ...
##  $ lninc            : num  3.43 2.71 3.56 3 3.14 4.67 4.34 4.57 2.77 4.43 ...
##  $ inccat           : num  2 1 2 1 1 4 4 4 1 4 ...
##  $ debtinc          : num  11.1 18.6 9.9 5.7 1.7 5.6 1.9 14.4 2.6 4.1 ...
##  $ creddebt         : num  1.2 1.22 0.93 0.02 0.21 1.06 0.5 5.95 0.1 1.77 ...
##  $ lncreddebt       : num  0.18 0.2 -0.07 -3.78 -1.54 0.06 -0.69 1.78 -2.28 0.57 ...
##  $ othdebt          : num  2.24 1.57 2.54 1.12 0.18 4.93 0.96 8.02 0.31 1.67 ...
##  $ lnothdebt        : num  0.81 0.45 0.93 0.11 -1.74 1.6 -0.04 2.08 -1.16 0.52 ...
##  $ default          : num  1 1 0 1 0 0 0 0 0 0 ...
##  $ jobsat           : num  1 1 4 2 1 2 2 5 2 4 ...
##  $ marital          : num  0 0 1 1 1 0 0 1 0 0 ...
##  $ spoused          : num  -1 -1 13 18 13 -1 -1 15 -1 -1 ...
##  $ spousedcat       : num  -1 -1 2 4 2 -1 -1 3 -1 -1 ...
##  $ reside           : num  3 2 3 5 4 1 1 2 1 2 ...
##  $ pets             : num  0 6 3 0 0 11 2 10 1 1 ...
##  $ pets_cats        : num  0 0 2 0 0 1 0 0 1 1 ...
##  $ pets_dogs        : num  0 0 1 0 0 1 2 2 0 0 ...
##  $ pets_birds       : num  0 0 0 0 0 0 0 0 0 0 ...
##  $ pets_reptiles    : num  0 0 0 0 0 0 0 0 0 0 ...
##  $ pets_small       : num  0 0 0 0 0 2 0 0 0 0 ...
##  $ pets_saltfish    : num  0 0 0 0 0 0 0 0 0 0 ...
##  $ pets_freshfish   : num  0 6 0 0 0 7 0 8 0 0 ...
##  $ homeown          : num  0 1 1 1 0 1 0 1 1 1 ...
##  $ hometype         : num  2 3 1 3 2 1 3 3 1 1 ...
##  $ address          : num  0 2 30 3 3 31 21 20 21 19 ...
```

```
## $ addresscat       : num  1 1 5 2 2 5 4 4 4 4 ...
## $ cars             : num  2 2 3 3 1 0 2 1 1 4 ...
## $ carown           : num  1 1 1 1 0 -1 1 1 1 1 ...
## $ cartype          : num  0 1 1 1 1 -1 0 0 1 0 ...
## $ carvalue         : num  14.3 6.8 18.8 8.7 10.6 -1 25.6 55.5 8.6 41 ...
## $ carcatvalue      : num  1 1 1 1 1 -1 2 3 1 3 ...
## $ carbought        : num  0 0 0 0 0 -1 1 0 0 1 ...
## $ carbuy           : num  0 0 1 1 1 0 0 1 0 0 ...
## $ commute          : num  8 1 4 1 6 8 4 5 4 1 ...
## $ commutecat       : num  4 1 3 1 3 4 3 3 3 1 ...
## $ commutetime      : num  22 29 24 38 32 23 32 31 25 29 ...
## $ commutecar       : num  0 1 1 1 0 0 1 0 1 1 ...
## $ commutemotorcycle: num  1 0 0 0 0 0 0 0 1 1 ...
## $ commutecarpool   : num  1 0 1 0 0 0 0 0 0 0 ...
## $ commutebus       : num  0 1 1 0 0 0 1 0 1 0 ...
## $ commuterail      : num  0 0 1 0 0 1 0 1 1 0 ...
## $ commutepublic    : num  0 0 0 0 1 0 0 0 0 0 ...
## $ commutebike      : num  0 1 0 0 0 0 1 0 0 0 ...
## $ commutewalk      : num  1 0 0 0 1 1 1 0 0 0 ...
## $ commutenonmotor  : num  0 1 0 0 0 0 0 0 0 0 ...
## $ telecommute      : num  0 1 0 0 0 0 0 0 0 0 ...
## $ reason           : num  9 9 2 9 9 9 9 2 9 9 ...
## $ polview          : num  6 4 5 3 4 4 4 6 3 3 ...
## $ polparty         : num  1 1 1 0 0 0 0 0 1 1 ...
## $ polcontrib       : num  0 0 0 0 0 0 0 0 1 1 ...
## $ vote             : num  1 0 0 0 0 0 0 1 1 1 ...
## $ card             : num  3 2 2 2 4 2 5 1 3 5 ...
## $ cardtype         : num  1 4 1 1 2 4 3 1 4 4 ...
## $ cardbenefit      : num  1 1 4 4 1 1 1 4 2 1 ...
## $ cardfee          : num  0 0 0 0 0 1 1 0 0 0 ...
## $ cardtenure       : num  2 4 35 5 8 18 3 25 26 2 ...
## $ cardtenurecat    : num  2 2 5 2 3 5 2 5 5 2 ...
## $ card2            : num  5 4 4 3 1 3 2 3 2 2 ...
## $ card2type        : num  3 1 1 2 3 3 3 1 4 2 ...
## $ card2benefit     : num  1 3 3 4 2 2 4 4 3 1 ...
## $ card2fee         : num  0 0 0 0 0 1 0 0 0 0 ...
## $ card2tenure      : num  3 4 25 5 9 9 2 20 17 2 ...
## $ card2tenurecat   : num  2 2 5 2 3 3 2 5 5 2 ...
## $ carditems        : num  5 5 9 17 8 11 20 6 12 11 ...
## $ cardspent        : num  81.7 42.6 184.2 341 255.1 ...
## $ card2items       : num  4 2 7 1 7 0 5 7 6 5 ...
## $ card2spent       : num  67.8 34.9 175.8 18.4 252.7 ...
## $ active           : num  0 1 0 1 1 0 0 0 1 1 ...
## $ bfast            : num  3 1 3 1 3 3 2 3 3 3 ...
## $ tenure           : num  5 39 65 36 21 28 15 46 53 3 ...
## $ churn            : num  1 0 0 0 0 0 0 0 0 1 ...
## $ longmon          : num  6.5 8.9 28.4 6 3.05 ...
## $ lnlongmon        : num  1.87 2.19 3.35 1.79 1.12 2.09 1.34 2.7 3.05 1.29 ...
## $ longten          : num  34.4 330.6 1858.3 199.4 74.1 ...
## $ lnlongten        : num  3.54 5.8 7.53 5.3 4.31 5.58 3.8 6.42 6.98 3 ...
## $ tollfree         : num  1 0 0 0 1 1 0 1 0 0 ...
## $ tollmon          : num  29 0 0 0 16.5 ...
## $ lntollmon        : num  3.37 NA NA NA 2.8 3.29 NA 3.18 NA NA ...
## $ tollten          : num  161 0 0 0 388 ...
```

```
##  $ lntollten        : num  5.08 NA NA NA 5.96 6.59 NA 7.01 NA NA ...
##  $ equip            : num  1 1 0 0 0 1 0 0 0 0 ...
##  $ equipmon         : num  29.5 54.9 0 0 0 ...
##  $ lnequipmon       : num  3.38 4 NA NA NA 3.57 NA NA NA NA ...
##  $ equipten         : num  126 1975 0 0 0 ...
##  $ lnequipten       : num  4.84 7.59 NA NA NA 6.88 NA NA NA NA ...
##   [list output truncated]
##  - attr(*, ".internal.selfref")=<externalptr>
```

```
cust_ <- cust_data
```

Making necessary changes to variable types

```
cust_$region <- factor(cust_data$region,levels = c(1,2,3,4,5),labels = c("zone 1","zone 2","zone 3","zo
cust_$townsize <- factor(cust_$townsize, levels = c(1,2,3,4,5),labels = c("> 250,000","50,000-249,999",
cust_$gender <- factor(cust_$gender, levels = c(0,1),labels = c("Male","Female"))
cust_$agecat <- factor(cust_$agecat, levels = c(1,2,3,4,5,6,9), labels = c("<18","18-24","25-34","35-49
cust_$birthmonth <- factor(cust_$birthmonth, levels = c("April","August","December","February","January
cust_$edcat <- ordered(cust_$edcat, levels = c(1,2,3,4,5), labels = c("Did not complete high school","H
cust_$jobcat <- factor(cust_$jobcat, levels= c(1,2,3,4,5,6), labels = c("Managerial and Professional","
cust_$jobcat <- factor(cust_$union, levels = c(0,1),labels = c("No","Yes"))
cust_$employ <- as.factor(cust_$employ)
cust_$empcat <- factor(cust_$empcat, levels = c(1,2,3,4,5), labels = c("Less than 2","2 to 5","6 to 10"
cust_$retire <- factor(cust_$retire, levels = c(0,1), labels = c("No","Yes"))
cust_$inccat <- ordered(cust_$inccat, levels = c(1,2,3,4,5), labels = c("Under $25","$25 - $49","$50 - 
cust_$default <- factor(cust_$default, levels = c(0,1), labels = c("No","Yes"))
cust_$jobsat <- ordered(cust_$jobsat, levels = c(1,2,3,4,5), labels = c("Highly dissatisfied","Somewhat
cust_$marital <- factor(cust_$marital, levels = c(0,1), labels = c("Unmarried","Married"))
cust_$spousedcat <- factor(cust_$spousedcat, levels = c(-1,1,2,3,4,5),labels = c("Not married","Did not
cust_$homeown <- factor(cust_$homeown, levels = c(0,1), labels = c("Rent","Own"))
cust_$hometype <- factor(cust_$hometype, levels = c(1,2,3,4), labels = c("Single-family","Multiple-Fami
cust_$address <- as.factor(cust_$address)
cust_$addresscat <-  ordered(cust_$addresscat, levels = c(1,2,3,4,5), labels = c("Less than 3","4 to 7"
cust_$cars <- as.ordered(cust_$cars)
cust_$carown <- factor(cust_$carown, levels = c(-1,0,1), labels = c("N/A","Lease","Own"))
cust_$cartype <- factor(cust_$cartype, levels = c(-1,0,1), labels = c("N/A","Domestic","Import"))
cust_$carcatvalue <- ordered(cust_$carcatvalue, levels = c(-1,1,2,3), labels = c("N/A","Economy","Stand
cust_$carbought <-factor(cust_$carbought, levels= c(-1,0,1), labels = c("N/A","No","Yes"))
cust_$carbuy <- factor(cust_$carbuy, levels = c(0,1), labels = c("No","Yes"))
cust_$commute <- factor(cust_$commute, levels = c(1,2,3,4,5,6,7,8,9,10), labels = c("Car","Motorcycle",
cust_$commutecat <- factor(cust_$commutecat, levels = c(1,2,3,4,5), labels = c("Single occupancy","Mult
cust_$commutecar <- factor(cust_$commutecar, levels = c(0,1), labels = c("No","Yes"))
cust_$commutemotorcycle <- factor(cust_$commutemotorcycle, levels = c(0,1), labels = c("No","Yes"))
cust_$commutecarpool<- factor(cust_$commutecarpool, levels = c(0,1), labels = c("No","Yes"))
cust_$commutebus <- factor(cust_$commutebus, levels = c(0,1), labels = c("No","Yes"))
cust_$commuterail<- factor(cust_$commuterail, levels = c(0,1), labels = c("No","Yes"))
cust_$commutepublic <- factor(cust_$commutepublic, levels = c(0,1), labels = c("No","Yes"))
cust_$commutebike <- factor(cust_$commutebike, levels = c(0,1), labels = c("No","Yes"))
cust_$commutewalk <- factor(cust_$commutewalk, levels = c(0,1), labels = c("No","Yes"))
cust_$commutenonmotor <- factor(cust_$commutenonmotor, levels = c(0,1), labels = c("No","Yes"))
cust_$telecommute <- factor(cust_$telecommute, levels = c(0,1), labels = c("No","Yes"))
cust_$reason <- factor(cust_$reason, levels = c(1,2,3,4,8,9), labels = c("Prices","Convenience","Servic
cust_$polview <- factor(cust_$polview, levels = c(1,2,3,4,5,6,7), labels = c("Extremely liberal","Liber
cust_$polparty <- factor(cust_$polparty, levels = c(0,1), labels = c("No","Yes"))
```

```
cust_$polcontrib <- factor(cust_$polcontrib, levels = c(0,1), labels = c("No","Yes"))
cust_$vote <- factor(cust_$vote, levels = c(0,1), labels = c("No","Yes"))
cust_$card <- factor(cust_$card, levels = c(1,2,3,4,5), labels = c("American Express","Visa","Mastercard
cust_$cardtype <- factor(cust_$cardtype, levels = c(1,2,3,4), labels = c("None","Gold","Platinum","Other
cust_$cardbenefit <- factor(cust_$cardbenefit, levels = c(1,2,3,4), labels = c("None","Cash back","Airli
cust_$cardfee <- factor(cust_$cardfee, levels = c(0,1), labels = c("No","Yes"))
cust_$cardtenure <- ordered(cust_$cardtenure)
cust_$cardtenurecat <- ordered(cust_$cardtenure, levels= c(1,2,3,4,5), labels= c("Less than 2","2 to 5"
cust_$card2 <- factor(cust_$card2, levels = c(1,2,3,4,5), labels = c("American Express","Visa","Masterca
cust_$card2type <- factor(cust_$card2type, levels = c(1,2,3,4), labels = c("None","Gold","Platinum","Oth
cust_$card2benefit <- factor(cust_$card2benefit, levels = c(1,2,3,4), labels = c("None","Cash back","Ai
cust_$card2fee <- factor(cust_$card2fee, levels = c(0,1), labels = c("No","Yes"))
cust_$card2tenure <- ordered(cust_$card2tenure)
cust_$card2tenurecat <- ordered(cust_$card2tenure, levels= c(1,2,3,4,5), labels= c("Less than 2","2 to 5
cust_$active <- factor(cust_$active, levels = c(0,1), labels = c("No","Yes"))
cust_$bfast <- factor(cust_$bfast, levels = c(1,2,3), labels = c("Energy bar","Oatmeal","Cereal"))
cust_$churn <- factor(cust_$churn, levels = c(0,1), labels = c("No","Yes"))
cust_$tollfree <- factor(cust_$tollfree, levels = c(0,1), labels = c("No","Yes"))
cust_$equip <- factor(cust_$equip, levels = c(0,1), labels = c("No","Yes"))
cust_$callcard <- factor(cust_$callcard, levels = c(0,1), labels = c("No","Yes"))
cust_$wireless <- factor(cust_$wireless, levels = c(0,1), labels = c("No","Yes"))
cust_$multline <- factor(cust_$multline, levels = c(0,1), labels = c("No","Yes"))
cust_$voice <- factor(cust_$voice, levels = c(0,1), labels = c("No","Yes"))
cust_$pager <- factor(cust_$pager, levels = c(0,1), labels = c("No","Yes"))
cust_$internet <- factor(cust_$internet, levels = c(0,1,2,3,4), labels = c("None", "Dia-up","DSL","Cable
cust_$callid <- factor(cust_$callid, levels = c(0,1), labels = c("No","Yes"))
cust_$callwait <- factor(cust_$callwait, levels = c(0,1), labels = c("No","Yes"))
cust_$forward <- factor(cust_$forward, levels = c(0,1), labels = c("No","Yes"))
cust_$confer <- factor(cust_$confer, levels = c(0,1), labels = c("No","Yes"))
cust_$ebill <- factor(cust_$ebill, levels = c(0,1), labels = c("No","Yes"))
cust_$owntv <- factor(cust_$owntv, levels = c(0,1), labels = c("No","Yes"))
cust_$ownvcr <- factor(cust_$ownvcr, levels = c(0,1), labels = c("No","Yes"))
cust_$owndvd <- factor(cust_$owndvd, levels = c(0,1), labels = c("No","Yes"))
cust_$owncd <- factor(cust_$owncd, levels = c(0,1), labels = c("No","Yes"))
cust_$ownpda <- factor(cust_$ownpda, levels = c(0,1), labels = c("No","Yes"))
cust_$ownpc <- factor(cust_$ownpc, levels = c(0,1), labels = c("No","Yes"))
cust_$ownipod <- factor(cust_$ownipod, levels = c(0,1), labels = c("No","Yes"))
cust_$owngame <- factor(cust_$owngame, levels = c(0,1), labels = c("No","Yes"))
cust_$ownfax <- factor(cust_$ownfax, levels = c(0,1), labels = c("No","Yes"))
cust_$news <- factor(cust_$news, levels = c(0,1), labels = c("No","Yes"))
cust_$response_01 <- factor(cust_$response_01, levels = c(0,1), labels = c("No","Yes"))
cust_$response_02 <- factor(cust_$response_02, levels = c(0,1), labels = c("No","Yes"))
cust_$response_03 <- factor(cust_$response_03, levels = c(0,1), labels = c("No","Yes"))
```

A user defined function to see about the data

```
about_nums <- function(x){

  n = length(x)
  nmiss = sum(is.na(x))
  nmiss_pct = (mean(is.na(x)))*100
  sum = sum(x, na.rm=T)
  mean = mean(x, na.rm=T)
```

```
  median = quantile(x, p=0.5, na.rm=T)
  std = sd(x, na.rm=T)
  var = var(x, na.rm=T)
  range = max(x, na.rm=T)-min(x, na.rm=T)
  pctl = quantile(x, p=c(0, 0.01, 0.05,0.1,0.25,0.5, 0.75,0.9,0.95,0.99,1), na.rm=T)
  return(c(N=n, Nmiss =nmiss, Nmiss_pct = nmiss_pct, sum=sum, avg=mean, meidan=median, std=std, var=var

}

about_cats <- function(x){

  n = length(x)
  nmiss = sum(is.na(x))
  nmiss_pct = (mean(is.na(x)))*100
  return(c(N=n, Nmiss =nmiss, Nmiss_pct = nmiss_pct))

}
```

Taking numerics and non-numerics as two subsets

```
numerics <- names(cust_)[sapply(cust_, is.numeric)]
# cust_nums <- cust_[,..numerics,] ------> creates subset of numeric vars

non_nums <- as.data.frame(names(cust_))
nums <- as.data.frame(numerics)
rest <- non_nums[!non_nums$`names(cust_)`%in%nums$numerics,]
rest <- as.vector(rest)
# non_nums <- cust_[,..rest,] -------> creates subset of non-nums
```

Examining through the variables

Numeric

```
numbs <- as.data.frame(t(sapply(cust_[,..numerics,], about_nums)))
kep <- row.names(numbs[numbs$Nmiss_pct<50,]) # -------- removing variables with missing more than 50% m
numerics_cust <- cust_[,..kep]
# For the rest of the variables with missing, the mean value is inserted
miss_treat_num = function(x){
  x[is.na(x)] = mean(x,na.rm=T) # replace missings with mean
  return(x)
}
numerics_cust <- data.table(apply(numerics_cust,2, FUN = miss_treat_num))
numbs <- as.data.frame(t(sapply(numerics_cust, about_nums)))
fwrite(numbs,"num_outliers.csv")
```

Capping outliers in the data

Capping with 95 percentile

```
outlier_treat_95 <- function(x){
  UC1 = quantile(x, p=0.95,na.rm=T)
  LC1 = quantile(x, p=0.05,na.rm=T)
```

```
  x=ifelse(x>UC1, UC1, x)
  x=ifelse(x<LC1, LC1, x)
  return(x)
}

cap95 <- numerics_cust[,c(1,21)]
cap95 <- data.table(apply(cap95,2, FUN = outlier_treat_95))
```

Capping with 99 percentile value

```
outlier_treat_99 <- function(x){
  UC1 = quantile(x, p=0.99,na.rm=T)
  LC1 = quantile(x, p=0.01,na.rm=T)

  x=ifelse(x>UC1, UC1, x)
  x=ifelse(x<LC1, LC1, x)
  return(x)
}

cap99 <- numerics_cust[,c(-1,-21)]
cap99 <- data.table(apply(cap99,2, FUN = outlier_treat_99))

numerics_cust <- cbind(cap95,cap99)
```

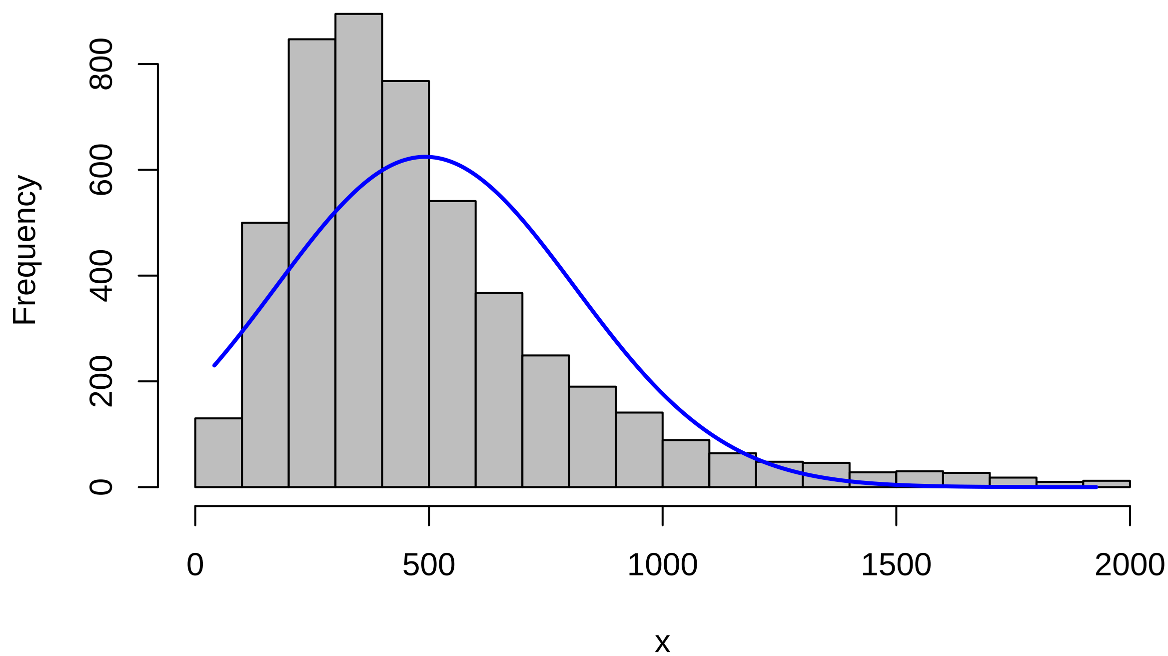
Categorical

```
cats <- as.data.frame(t(sapply(cust_[,-..numerics,], about_cats)))

cat_cust <- cust_[,-..numerics,]
cat_cust$cardtenurecat <- NULL
cat_cust$card2tenurecat <- NULL

cust_final <- cbind(numerics_cust,cat_cust)
cust_final$total_spent <- cust_final$cardspent+cust_final$card2spent
```

Checking regression assumptions

Normality of Y variable

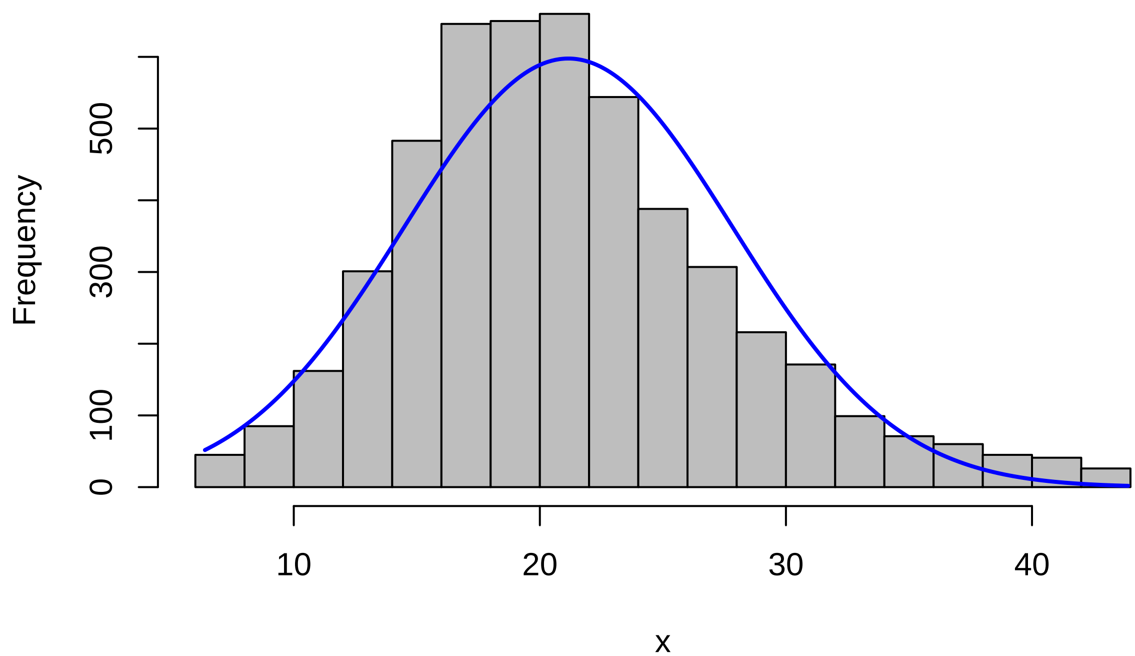
```
plotNormalHistogram(cust_final$total_spent)
```




```
shapiro.test(cust_final$total_spent)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  cust_final$total_spent
## W = 0.87136, p-value < 0.00000000000000022
```

```
# This is not normal hence it is transformed
cust_final$total_spent_sqrt <-  sqrt(cust_final$total_spent)
plotNormalHistogram(cust_final$total_spent_sqrt)
```




```
cust_final$total_spent_sqrt <-  sqrt(cust_final$total_spent)
```

This shows a more normal distribution, thus this will be used as Y

Checking if Y variables and X Variables are correlated

```
numerics_cust <- cbind(numerics_cust,total_spent_sqrt=cust_final$total_spent_sqrt,total_spent = cust_fi
cormat <- as.data.frame(cor(numerics_cust))
fwrite(cormat,"cormat.csv",row.names = T)
remove_ <- c("age","union","debtinc","spoused","reside","pets","pets_cats","pets_dogs","pets_birds",
             "pets_reptiles","pets_small","pets_saltfish","pets_freshfish","commutetime","tenure","long
             "lnlongmon","longten","lnlongten","tollmon","equipmon","equipten","cardmon","lncardmon",
             "cardten","lncardten","wiremon","hourstv")
# removing all numeric variables that are nor correlated with total_spent
cust_final <- cust_final[,-..remove_,]
# removing variables with high inter-item correlation with 0.6 as limit
remove_ <- c("carvalue","income","creddebt","lnothdebt","card2items")
cust_final <- cust_final[,-..remove_,]
```

Checking the relationship of categorical X variables with Y variable

```
Anov <- function(x){
summary(aov(cust_final$total_spent~x))
}
```

```r
cat_cust$custid <- NULL
apply(cat_cust, 2, FUN = Anov)
```

```
## $region
##                Df    Sum Sq Mean Sq F value Pr(>F)
## x               4   1054636  263659   2.588  0.035 *
## Residuals    4995 508790498  101860
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $townsize
##                Df    Sum Sq Mean Sq F value Pr(>F)
## x               4    136780   34195   0.335  0.855
## Residuals    4993 509622803  102067
## 2 observations deleted due to missingness
##
## $gender
##                Df    Sum Sq Mean Sq F value      Pr(>F)
## x               1   3068393 3068393   30.26 0.0000000396 ***
## Residuals    4998 506776741  101396
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $agecat
##                Df    Sum Sq Mean Sq F value              Pr(>F)
## x               4  14551607 3637902   36.69 <0.0000000000000002 ***
## Residuals    4995 495293527   99158
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $birthmonth
##                Df    Sum Sq Mean Sq F value Pr(>F)
## x              11    439019   39911   0.391   0.96
## Residuals    4988 509406116  102126
##
## $edcat
##                Df    Sum Sq Mean Sq F value        Pr(>F)
## x               4   5050828 1262707   12.49 0.000000000409 ***
## Residuals    4995 504794307  101060
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $jobcat
##                Df    Sum Sq Mean Sq F value Pr(>F)
## x               1     70766   70766   0.694  0.405
## Residuals    4998 509774368  101996
##
## $employ
##                Df    Sum Sq Mean Sq F value       Pr(>F)
## x              51  13336411  261498   2.606 0.00000000446 ***
## Residuals    4948 496508723  100345
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## $empcat
##              Df    Sum Sq Mean Sq F value            Pr(>F)
## x             4   7045584 1761396    17.5 0.0000000000000286 ***
## Residuals  4995 502799550  100661
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $retire
##              Df    Sum Sq  Mean Sq F value              Pr(>F)
## x             1  14674020 14674020   148.1 <0.0000000000000002 ***
## Residuals  4998 495171114    99074
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $inccat
##              Df    Sum Sq  Mean Sq F value              Pr(>F)
## x             4  76130691 19032673   219.2 <0.0000000000000002 ***
## Residuals  4995 433714443    86830
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $default
##              Df    Sum Sq Mean Sq F value Pr(>F)
## x             1     38194   38194   0.374  0.541
## Residuals  4998 509806941  102002
##
## $jobsat
##              Df    Sum Sq Mean Sq F value      Pr(>F)
## x             4   3596431  899108   8.871 0.000000391 ***
## Residuals  4995 506248704  101351
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $marital
##              Df    Sum Sq Mean Sq F value Pr(>F)
## x             1     94361   94361   0.925  0.336
## Residuals  4998 509750773  101991
##
## $spousedcat
##              Df    Sum Sq Mean Sq F value Pr(>F)
## x             5   1528394  305679   3.003 0.0104 *
## Residuals  4994 508316740  101785
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $homeown
##              Df    Sum Sq Mean Sq F value      Pr(>F)
## x             1   2673048 2673048   26.34 0.000000297 ***
## Residuals  4998 507172086  101475
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $hometype
##              Df    Sum Sq Mean Sq F value   Pr(>F)
```

```
## x              3   2057899  685966   6.749 0.000154 ***
## Residuals   4996 507787235  101639
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $address
##               Df    Sum Sq Mean Sq F value    Pr(>F)
## x             56  10884689  194369   1.926 0.0000453 ***
## Residuals   4943 498960445  100943
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $addresscat
##               Df    Sum Sq Mean Sq F value             Pr(>F)
## x              4   6843762 1710940   16.99 0.0000000000000759 ***
## Residuals   4995 503001373  100701
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $cars
##               Df    Sum Sq Mean Sq F value Pr(>F)
## x              8    677094   84637    0.83  0.576
## Residuals   4991 509168041  102017
##
## $carown
##               Df    Sum Sq Mean Sq F value              Pr(>F)
## x              2  10107155 5053578   50.53 <0.0000000000000002 ***
## Residuals   4997 499737979  100008
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $cartype
##               Df    Sum Sq Mean Sq F value Pr(>F)
## x              2    391449  195724    1.92  0.147
## Residuals   4997 509453686  101952
##
## $carcatvalue
##               Df    Sum Sq  Mean Sq F value              Pr(>F)
## x              3  57143454 19047818   210.2 <0.0000000000000002 ***
## Residuals   4996 452701680    90613
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $carbought
##               Df    Sum Sq Mean Sq F value Pr(>F)
## x              2    354037  177018   1.736  0.176
## Residuals   4997 509491098  101959
##
## $carbuy
##               Df    Sum Sq Mean Sq F value Pr(>F)
## x              1    117182  117182   1.149  0.284
## Residuals   4998 509727953  101986
##
## $commute
```

```
##              Df    Sum Sq Mean Sq F value Pr(>F)
## x             9    949548  105505   1.035  0.409
## Residuals  4990 508895587  101983
##
## $commutecat
##              Df    Sum Sq Mean Sq F value Pr(>F)
## x             4    303267   75817   0.743  0.562
## Residuals  4995 509541868  102010
##
## $commutecar
##              Df    Sum Sq Mean Sq F value Pr(>F)
## x             1     41137   41137   0.403  0.525
## Residuals  4998 509803998  102002
##
## $commutemotorcycle
##              Df    Sum Sq Mean Sq F value Pr(>F)
## x             1    230742  230742   2.263  0.133
## Residuals  4998 509614393  101964
##
## $commutecarpool
##              Df    Sum Sq Mean Sq F value Pr(>F)
## x             1     18970   18970   0.186  0.666
## Residuals  4998 509826165  102006
##
## $commutebus
##              Df    Sum Sq Mean Sq F value Pr(>F)
## x             1     63297   63297   0.621  0.431
## Residuals  4998 509781837  101997
##
## $commuterail
##              Df    Sum Sq Mean Sq F value Pr(>F)
## x             1     22512   22512   0.221  0.639
## Residuals  4998 509822622  102005
##
## $commutepublic
##              Df    Sum Sq Mean Sq F value Pr(>F)
## x             1        30      30       0  0.986
## Residuals  4998 509845105  102010
##
## $commutebike
##              Df    Sum Sq Mean Sq F value Pr(>F)
## x             1    477132  477132   4.682 0.0305 *
## Residuals  4998 509368002  101914
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $commutewalk
##              Df    Sum Sq Mean Sq F value Pr(>F)
## x             1     10699   10699   0.105  0.746
## Residuals  4998 509834435  102008
##
## $commutenonmotor
##              Df    Sum Sq Mean Sq F value Pr(>F)
## x             1     69678   69678   0.683  0.409
```

```
## Residuals   4998 509775457  101996
##
## $telecommute
##               Df    Sum Sq Mean Sq F value Pr(>F)
## x              1     34691   34691    0.34   0.56
## Residuals   4998 509810444  102003
##
## $reason
##               Df    Sum Sq Mean Sq F value              Pr(>F)
## x              4  14817637 3704409   37.38 <0.0000000000000002 ***
## Residuals   4995 495027497   99105
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $polview
##               Df    Sum Sq Mean Sq F value Pr(>F)
## x              6   1209432  201572   1.979 0.0651 .
## Residuals   4993 508635703  101870
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $polparty
##               Df    Sum Sq Mean Sq F value Pr(>F)
## x              1     41327   41327   0.405  0.524
## Residuals   4998 509803807  102002
##
## $polcontrib
##               Df    Sum Sq Mean Sq F value Pr(>F)
## x              1    494343  494343   4.851 0.0277 *
## Residuals   4998 509350791  101911
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $vote
##               Df    Sum Sq Mean Sq F value     Pr(>F)
## x              1   2170590 2170590   21.37 0.00000388 ***
## Residuals   4998 507674545  101576
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $card
##               Df    Sum Sq  Mean Sq F value              Pr(>F)
## x              4  72491796 18122949     207 <0.0000000000000002 ***
## Residuals   4995 437353338    87558
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $cardtype
##               Df    Sum Sq Mean Sq F value Pr(>F)
## x              3    129747   43249   0.424  0.736
## Residuals   4996 509715387  102025
##
## $cardbenefit
##               Df    Sum Sq Mean Sq F value Pr(>F)
```

```
## x              3    546128  182043   1.786  0.148
## Residuals   4996 509299006  101941
##
## $cardfee
##               Df    Sum Sq Mean Sq F value Pr(>F)
## x              1     13579   13579   0.133  0.715
## Residuals   4998 509831555  102007
##
## $cardtenure
##               Df    Sum Sq Mean Sq F value   Pr(>F)
## x             40   7780859  194521   1.921 0.000443 ***
## Residuals   4959 502064275  101243
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $card2
##               Df    Sum Sq Mean Sq F value              Pr(>F)
## x              4  15378634 3844659   38.84 <0.0000000000000002 ***
## Residuals   4995 494466500   98992
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $card2type
##               Df    Sum Sq Mean Sq F value Pr(>F)
## x              3     62545   20848   0.204  0.893
## Residuals   4996 509782589  102038
##
## $card2benefit
##               Df    Sum Sq Mean Sq F value Pr(>F)
## x              3    242830   80943   0.794  0.497
## Residuals   4996 509602304  102002
##
## $card2fee
##               Df    Sum Sq Mean Sq F value Pr(>F)
## x              1    498034  498034   4.887 0.0271 *
## Residuals   4998 509347100  101910
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $card2tenure
##               Df    Sum Sq Mean Sq F value    Pr(>F)
## x             30   7616630  253888   2.512 0.0000101 ***
## Residuals   4969 502228504  101072
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $active
##               Df    Sum Sq Mean Sq F value Pr(>F)
## x              1     51818   51818   0.508  0.476
## Residuals   4998 509793316  101999
##
## $bfast
##               Df    Sum Sq Mean Sq F value Pr(>F)
## x              2    551504  275752   2.706 0.0669 .
```

```
## Residuals   4997 509293631  101920
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $churn
##               Df    Sum Sq Mean Sq F value Pr(>F)
## x              1     50479   50479   0.495  0.482
## Residuals   4998 509794656  102000
##
## $tollfree
##               Df    Sum Sq Mean Sq F value    Pr(>F)
## x              1   1730013 1730013   17.02 0.0000376 ***
## Residuals   4998 508115121  101664
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $equip
##               Df    Sum Sq Mean Sq F value   Pr(>F)
## x              1   1347603 1347603   13.25 0.000276 ***
## Residuals   4998 508497531  101740
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $callcard
##               Df    Sum Sq Mean Sq F value Pr(>F)
## x              1    436712  436712   4.285 0.0385 *
## Residuals   4998 509408422  101922
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $wireless
##               Df    Sum Sq Mean Sq F value     Pr(>F)
## x              1   2219718 2219718   21.86 0.00000302 ***
## Residuals   4998 507625416  101566
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $multline
##               Df    Sum Sq Mean Sq F value   Pr(>F)
## x              1   1200261 1200261   11.79 0.000599 ***
## Residuals   4998 508644873  101770
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $voice
##               Df    Sum Sq Mean Sq F value Pr(>F)
## x              1    583189  583189   5.724 0.0168 *
## Residuals   4998 509261945  101893
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $pager
##               Df    Sum Sq Mean Sq F value   Pr(>F)
## x              1   1540683 1540683   15.15 0.000101 ***
```

```
## Residuals   4998 508304451  101702
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $internet
##               Df    Sum Sq Mean Sq F value   Pr(>F)
## x              4   2945817  736454   7.257 0.000008 ***
## Residuals   4995 506899317  101481
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $callid
##               Df    Sum Sq Mean Sq F value    Pr(>F)
## x              1   1559884 1559884   15.34 0.0000911 ***
## Residuals   4998 508285250  101698
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $callwait
##               Df    Sum Sq Mean Sq F value    Pr(>F)
## x              1   1707810 1707810    16.8 0.0000422 ***
## Residuals   4998 508137324  101668
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $forward
##               Df    Sum Sq Mean Sq F value   Pr(>F)
## x              1   1490594 1490594   14.65 0.000131 ***
## Residuals   4998 508354540  101712
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $confer
##               Df    Sum Sq Mean Sq F value    Pr(>F)
## x              1   1732684 1732684   17.04 0.0000371 ***
## Residuals   4998 508112451  101663
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $ebill
##               Df    Sum Sq Mean Sq F value Pr(>F)
## x              1    453338  453338   4.448  0.035 *
## Residuals   4998 509391796  101919
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $owntv
##               Df    Sum Sq Mean Sq F value      Pr(>F)
## x              1   2594300 2594300   25.56 0.000000444 ***
## Residuals   4998 507250834  101491
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $ownvcr
```

```
##               Df    Sum Sq Mean Sq F value              Pr(>F)
## x              1   9206679 9206679   91.91 <0.0000000000000002 ***
## Residuals   4998 500638455  100168
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $owndvd
##               Df    Sum Sq Mean Sq F value              Pr(>F)
## x              1   9843802 9843802    98.4 <0.0000000000000002 ***
## Residuals   4998 500001333  100040
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $owncd
##               Df    Sum Sq Mean Sq F value              Pr(>F)
## x              1   7543624 7543624   75.06 <0.0000000000000002 ***
## Residuals   4998 502301510  100501
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $ownpda
##               Df    Sum Sq Mean Sq F value      Pr(>F)
## x              1   2514947 2514947   24.78 0.000000666 ***
## Residuals   4998 507330187  101507
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $ownpc
##               Df    Sum Sq Mean Sq F value   Pr(>F)
## x              1   1311225 1311225   12.89 0.000334 ***
## Residuals   4998 508533909  101747
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $ownipod
##               Df    Sum Sq Mean Sq F value Pr(>F)
## x              1    494879  494879   4.856 0.0276 *
## Residuals   4998 509350255  101911
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $owngame
##               Df    Sum Sq Mean Sq F value Pr(>F)
## x              1    362131  362131   3.552 0.0595 .
## Residuals   4998 509483003  101937
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $ownfax
##               Df    Sum Sq Mean Sq F value    Pr(>F)
## x              1   1772441 1772441   17.44 0.0000302 ***
## Residuals   4998 508072693  101655
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
## $news
##                Df    Sum Sq Mean Sq F value  Pr(>F)
## x               1    841780  841780   8.266 0.00406 **
## Residuals    4998 509003354  101841
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $response_01
##                Df    Sum Sq Mean Sq F value Pr(>F)
## x               1     28651   28651   0.281  0.596
## Residuals    4998 509816483  102004
##
## $response_02
##                Df    Sum Sq Mean Sq F value Pr(>F)
## x               1    258071  258071   2.531  0.112
## Residuals    4998 509587063  101958
##
## $response_03
##                Df    Sum Sq Mean Sq F value   Pr(>F)
## x               1   1358424 1358424   13.35 0.000261 ***
## Residuals    4998 508486711  101738
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# Categories that are insignificant are removed except for those that needed to be kept in relation to
remove_ <- c("townsize","birthmonth","jobcat","default","marital","cars","carbought","carbuy","commute"
             "commutecar","commutemotorcycle","commutecarpool","commutebus","commuterail","commutepubli
             "commutewalk","commutenonmotor","telecommute","polparty","response_01","response_02")
cust_final <- cust_final[,-..remove_,]
```

Creating dummy variables and preparing data for building the model

```
cust_final$custid <- NULL
ordered <- names(cust_final)[sapply(cust_final, is.ordered)]
cust_ordered <- cust_final[,..ordered,]
cust_final <- cust_final[,-..ordered,]
factor <- names(cust_final)[sapply(cust_final, is.factor)]
dd <- dummy_cols(cust_final, remove_first_dummy = TRUE)
cust_final <- dd[,-..factor,]
cust_final <- cbind(cust_final,cust_ordered)
```

Building the Regression model

Taking samples

```
set.seed(654)
samp <-  sample(1:nrow(cust_final),size = floor(nrow(cust_final)*0.7))
Dev <- cust_final[samp,]
Test <- cust_final[-samp,]

fit <- lm(total_spent_sqrt~.,data = Dev)
summary(fit)
```

```
##
## Call:
## lm(formula = total_spent_sqrt ~ ., data = Dev)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -6.4751 -0.5459  0.2084  0.7768  2.5887
##
## Coefficients: (11 not defined because of singularities)
##                                            Estimate  Std. Error t value
## (Intercept)                              6.91333713  0.69649351   9.926
## ed                                      -0.00140350  0.02745284  -0.051
## lninc                                    0.69831765  0.12968129   5.385
## lncreddebt                               0.02072706  0.02219918   0.934
## othdebt                                 -0.00435045  0.00732124  -0.594
## carditems                                0.21583534  0.00815976  26.451
## cardspent                                0.01727199  0.00016413 105.236
## card2spent                               0.02166340  0.00021350 101.466
## tollten                                  0.00002357  0.00004415   0.534
## wireten                                  0.00004175  0.00004369   0.955
## total_spent                                      NA          NA      NA
## `region_zone 5`                         -0.01199707  0.06720150  -0.179
## `region_zone 3`                         -0.05912040  0.06687519  -0.884
## `region_zone 4`                         -0.03102811  0.06730258  -0.461
## `region_zone 2`                         -0.00462108  0.06736190  -0.069
## gender_Male                              0.01607822  0.04265714   0.377
## `agecat_>65`                             0.30952765  0.18588498   1.665
## `agecat_25-34`                           0.01098643  0.13450969   0.082
## `agecat_50-64`                           0.20159468  0.17112258   1.178
## `agecat_35-49`                           0.22167723  0.15815212   1.402
## employ_16                                0.13608612  0.19637449   0.693
## employ_1                                -0.07360934  0.10751250  -0.685
## employ_22                               -0.21872293  0.27169424  -0.805
## employ_10                               -0.24871889  0.16408343  -1.516
## employ_11                               -0.19072643  0.16163823  -1.180
## employ_15                               -0.15733291  0.18356144  -0.857
## employ_19                               -0.01923199  0.20499162  -0.094
## employ_8                                -0.32437281  0.15598327  -2.080
## employ_4                                -0.19700026  0.13344517  -1.476
## employ_12                               -0.11113669  0.17881030  -0.622
## employ_3                                -0.11442906  0.12859232  -0.890
## employ_27                               -0.09195162  0.28106075  -0.327
## employ_31                               -0.06843752  0.28908829  -0.237
## employ_24                               -0.40904652  0.27423092  -1.492
## employ_29                               -0.02915824  0.30029734  -0.097
## employ_2                                -0.18443642  0.11999283  -1.537
## employ_18                               -0.30479053  0.21514811  -1.417
## employ_5                                -0.31111477  0.13953403  -2.230
## employ_28                               -0.32166996  0.32738624  -0.983
## employ_26                               -0.28970764  0.25924798  -1.117
## employ_23                               -0.05837279  0.23783918  -0.245
## employ_13                                0.08633709  0.17891518   0.483
## employ_17                               -0.12162279  0.19590012  -0.621
## employ_30                               -0.26851161  0.28863370  -0.930
```

```
## employ_14                                 -0.32357015  0.19616578  -1.649
## employ_6                                  -0.09682807  0.14317633  -0.676
## employ_36                                 -0.09748749  0.39349958  -0.248
## employ_7                                  -0.02732906  0.15270504  -0.179
## employ_9                                  -0.31214508  0.15948150  -1.957
## employ_21                                  0.11025111  0.23443307   0.470
## employ_20                                 -0.32042492  0.20989408  -1.527
## employ_34                                 -0.18393779  0.30485298  -0.603
## employ_35                                 -0.35755690  0.41717182  -0.857
## employ_33                                  0.25730651  0.34459653   0.747
## employ_25                                  0.20487864  0.27201180   0.753
## employ_39                                  0.11810635  0.40181513   0.294
## employ_45                                 -1.49377124  0.93903219  -1.591
## employ_32                                 -0.09263997  0.31905724  -0.290
## employ_40                                  0.29390447  0.64758249   0.454
## employ_37                                 -0.29185086  0.43319791  -0.674
## employ_48                                  0.12053092  1.25252903   0.096
## employ_38                                 -0.10225568  0.38256365  -0.267
## employ_46                                 -0.13090572  1.28211748  -0.102
## employ_43                                  0.11452598  0.74317357   0.154
## employ_47                                  0.15455844  1.26495485   0.122
## employ_41                                  2.03681933  1.26722011   1.607
## employ_42                                  0.58745218  0.58633973   1.002
## employ_49                                 -0.04278096  1.26978033  -0.034
## employ_44                                  0.64427027  0.59314132   1.086
## employ_52                                 -0.77639107  1.32660798  -0.585
## employ_51                                  1.12080188  1.26486671   0.886
## `empcat_More than 15`                              NA          NA      NA
## `empcat_6 to 10`                                   NA          NA      NA
## `empcat_11 to 15`                                  NA          NA      NA
## `empcat_2 to 5`                                    NA          NA      NA
## retire_Yes                                -0.13076625  0.11158786  -1.172
## `spousedcat_High school degree`           -0.01542996  0.06318898  -0.244
## `spousedcat_College degree`                0.08686225  0.08385085   1.036
## `spousedcat_Some college`                  0.07511438  0.07425164   1.012
## `spousedcat_Did not complete high school` -0.05906807  0.07220504  -0.818
## `spousedcat_Post-undergraduate degree`    -0.06612115  0.15717449  -0.421
## homeown_Own                               -0.08221209  0.06575953  -1.250
## `hometype_Condominium/Townhouse`           0.04096802  0.08096860   0.506
## `hometype_Single-family`                   0.08338028  0.06948214   1.200
## `hometype_Mobile Home`                     0.07957723  0.10888896   0.731
## address_2                                  0.10200118  0.17936558   0.569
## address_30                                 0.16729462  0.26494662   0.631
## address_3                                 -0.14223948  0.19713677  -0.722
## address_31                                -0.24830995  0.26641323  -0.932
## address_21                                -0.34830969  0.24134252  -1.443
## address_20                                -0.42550647  0.24837483  -1.713
## address_19                                -0.19324641  0.25062460  -0.771
## address_14                                 0.09134708  0.23159017   0.394
## address_5                                 -0.22844201  0.21136591  -1.081
## address_9                                  0.06770719  0.22992286   0.294
## address_32                                 0.14293497  0.27380554   0.522
## address_29                                -0.33736660  0.26791205  -1.259
## address_8                                 -0.18724233  0.22138681  -0.846
```

```
## address_12                              -0.25785390  0.22447382  -1.149
## address_10                              -0.15827347  0.22914394  -0.691
## address_38                              -0.14594241  0.31043023  -0.470
## address_4                               -0.04048651  0.20223667  -0.200
## address_57                               0.41442638  0.76477420   0.542
## address_16                              -0.03713520  0.23887844  -0.155
## address_27                              -0.30492604  0.25956912  -1.175
## address_6                                0.12146786  0.21844693   0.556
## address_35                              -0.06375698  0.30499429  -0.209
## address_25                              -0.26161880  0.25492123  -1.026
## address_36                              -0.42981706  0.31262594  -1.375
## address_22                               0.00805815  0.25450330   0.032
## address_24                              -0.11488722  0.25929307  -0.443
## address_13                              -0.18287390  0.23661088  -0.773
## address_11                              -0.00851662  0.23406302  -0.036
## address_1                               -0.08382493  0.17514142  -0.479
## address_23                              -0.06951789  0.25419416  -0.273
## address_26                               0.12203288  0.27172620   0.449
## address_39                              -0.41895961  0.31825961  -1.316
## address_42                               0.34579034  0.38807231   0.891
## address_17                              -0.00903240  0.24400929  -0.037
## address_34                               0.03681511  0.30505709   0.121
## address_18                              -0.03020161  0.24268533  -0.124
## address_28                               0.04003076  0.26504997   0.151
## address_40                              -0.36953494  0.33217137  -1.112
## address_45                              -0.35331731  0.39269128  -0.900
## address_15                              -0.10664044  0.23782480  -0.448
## address_46                              -0.43767120  0.45348584  -0.965
## address_50                               0.79501763  0.91317294   0.871
## address_33                              -0.35059583  0.29636300  -1.183
## address_37                              -0.45253587  0.32541230  -1.391
## address_7                               -0.27515982  0.22255630  -1.236
## address_43                               0.01450470  0.38374119   0.038
## address_41                              -0.05757248  0.34204777  -0.168
## address_44                              -0.39233230  0.37629755  -1.043
## address_47                               0.09336105  0.42687314   0.219
## address_51                               0.44166962  0.62417786   0.708
## address_49                               0.45763670  0.59995320   0.763
## address_55                               0.44387576  0.69368452   0.640
## address_48                              -0.05259738  0.39598119  -0.133
## address_52                              -1.10564463  0.76616394  -1.443
## address_53                               0.58210288  0.76559199   0.760
## address_54                              -0.07285986  1.28192304  -0.057
## carown_Lease                            -0.07298849  0.06474099  -1.127
## `carown_N/A`                            -0.75590974  0.22731822  -3.325
## cartype_Import                           0.02069739  0.04462388   0.464
## `cartype_N/A`                                     NA          NA      NA
## commutebike_Yes                          0.09442859  0.06459386   1.462
## reason_Convenience                       0.22115889  0.08461657   2.614
## reason_Prices                           -0.18314144  0.07469532  -2.452
## reason_Service                           0.18941597  0.20182169   0.939
## reason_Other                            -0.20651114  0.14882712  -1.388
## polview_Moderate                        -0.06134514  0.06516523  -0.941
## `polview_Slightly conservative`         -0.04271156  0.07328199  -0.583
```

```
## `polview_Slightly liberal`            -0.11497170  0.07929323  -1.450
## polview_Liberal                       -0.01679273  0.08002080  -0.210
## `polview_Extremely liberal`           -0.10684640  0.13100881  -0.816
## `polview_Extremely conservative`      -0.54539936  0.17409627  -3.133
## polcontrib_Yes                         0.00490912  0.05153765   0.095
## vote_No                                0.02312499  0.04401293   0.525
## card_Visa                              0.04281189  0.06673973   0.641
## card_Discover                         -0.02120147  0.06688167  -0.317
## card_Other                             0.03232410  0.11108033   0.291
## `card_American Express`                0.61018542  0.07343112   8.310
## cardtype_Other                         0.14258207  0.05994568   2.379
## cardtype_Gold                          0.09949933  0.05975772   1.665
## cardtype_Platinum                      0.03601710  0.05940127   0.606
## cardbenefit_Other                      0.08819895  0.05991422   1.472
## `cardbenefit_Cash back`                0.06935353  0.06027168   1.151
## `cardbenefit_Airline miles`            0.05666919  0.05920487   0.957
## cardfee_Yes                           -0.02844970  0.05399748  -0.527
## card2_Discover                        -0.11174355  0.09464254  -1.181
## card2_Mastercard                      -0.19365728  0.09284366  -2.086
## `card2_American Express`               0.40454440  0.10079660   4.013
## card2_Visa                            -0.13583978  0.09396779  -1.446
## card2type_None                         0.08767290  0.06034070   1.453
## card2type_Gold                        -0.00814959  0.05948013  -0.137
## card2type_Other                        0.02368243  0.05864442   0.404
## `card2benefit_Airline miles`          -0.05450746  0.05981524  -0.911
## card2benefit_Other                     0.03465276  0.05954244   0.582
## `card2benefit_Cash back`               0.02241328  0.06004380   0.373
## card2fee_Yes                           0.10532395  0.05403990   1.949
## active_Yes                            -0.02244068  0.04487629  -0.500
## `bfast_Energy bar`                     0.00546388  0.05294530   0.103
## bfast_Oatmeal                         -0.06069318  0.05975190  -1.016
## churn_No                              -0.08509279  0.05643230  -1.508
## tollfree_No                            0.06783040  0.07622704   0.890
## equip_No                               0.06246459  0.06397208   0.976
## callcard_No                           -0.09853112  0.06021930  -1.636
## wireless_Yes                           0.12479668  0.09056195   1.378
## multline_No                            0.00403647  0.05392225   0.075
## voice_No                              -0.07948692  0.06819764  -1.166
## pager_No                               0.12125585  0.07255543   1.671
## internet_Other                         0.06784565  0.10028063   0.677
## internet_DSL                          -0.06830071  0.09266118  -0.737
## `internet_Cable modem`                -0.06493017  0.09257026  -0.701
## `internet_Dia-up`                      0.00064387  0.08209999   0.008
## callid_Yes                             0.03385622  0.06279174   0.539
## callwait_No                            0.09715885  0.06275970   1.548
## forward_No                             0.04596418  0.06231682   0.738
## confer_No                             -0.04351673  0.06432521  -0.677
## ebill_Yes                              0.02141591  0.06051085   0.354
## owntv_No                               0.29579768  0.19697955   1.502
## ownvcr_No                              0.07223014  0.09586350   0.753
## owndvd_No                              0.01647664  0.09470029   0.174
## owncd_Yes                              0.12216660  0.10475481   1.166
## ownpda_Yes                            -0.10468998  0.06848494  -1.529
## ownpc_Yes                             -0.03496883  0.07034934  -0.497
```

```
## ownipod_No                          0.00363586  0.05120243   0.071
## owngame_No                         -0.05761714  0.05188234  -1.111
## ownfax_Yes                          0.03154471  0.07252440   0.435
## news_Yes                           -0.03418979  0.05204889  -0.657
## response_03_Yes                     0.13358594  0.06920118   1.930
## edcat.L                            -0.16044235  0.22662837  -0.708
## edcat.Q                             0.06369961  0.06120442   1.041
## edcat.C                            -0.01599009  0.04854237  -0.329
## edcat^4                             0.04055652  0.04574734   0.887
## inccat.L                           -0.61319701  0.21453182  -2.858
## inccat.Q                           -0.36598625  0.07391583  -4.951
## inccat.C                           -0.05329064  0.05647462  -0.944
## inccat^4                           -0.02243750  0.05406938  -0.415
## jobsat.L                            0.04922568  0.05909628   0.833
## jobsat.Q                           -0.04252528  0.04981372  -0.854
## jobsat.C                            0.08091540  0.04807390   1.683
## jobsat^4                            0.03884907  0.04605652   0.844
## addresscat.L                                NA          NA      NA
## addresscat.Q                                NA          NA      NA
## addresscat.C                                NA          NA      NA
## addresscat^4                                NA          NA      NA
## carcatvalue.L                      -0.53712998  0.17149614  -3.132
## carcatvalue.Q                       0.31565525  0.13349389   2.365
## carcatvalue.C                               NA          NA      NA
## cardtenure.L                       -0.88437681  0.57045824  -1.550
## cardtenure.Q                        0.51049755  0.39203359   1.302
## cardtenure.C                       -0.49459117  0.33412208  -1.480
## cardtenure^4                        0.07501509  0.30740360   0.244
## cardtenure^5                       -0.26215427  0.28410180  -0.923
## cardtenure^6                        0.06944549  0.26173225   0.265
## cardtenure^7                        0.04181811  0.23974281   0.174
## cardtenure^8                       -0.14618627  0.21798456  -0.671
## cardtenure^9                       -0.36476140  0.19849890  -1.838
## cardtenure^10                      -0.12300332  0.18330987  -0.671
## cardtenure^11                      -0.27249801  0.16800766  -1.622
## cardtenure^12                      -0.17606061  0.15752215  -1.118
## cardtenure^13                       0.00043397  0.15083370   0.003
## cardtenure^14                      -0.09196210  0.14712750  -0.625
## cardtenure^15                      -0.02678449  0.14490211  -0.185
## cardtenure^16                       0.10540772  0.14657754   0.719
## cardtenure^17                       0.13447534  0.14343997   0.938
## cardtenure^18                       0.10143946  0.14243716   0.712
## cardtenure^19                      -0.12386763  0.14500349  -0.854
## cardtenure^20                       0.19364633  0.14551932   1.331
## cardtenure^21                       0.05750672  0.14579059   0.394
## cardtenure^22                      -0.25131236  0.14775027  -1.701
## cardtenure^23                       0.06041752  0.14892165   0.406
## cardtenure^24                       0.07105478  0.14700246   0.483
## cardtenure^25                       0.37338447  0.14539277   2.568
## cardtenure^26                       0.01194626  0.14595676   0.082
## cardtenure^27                       0.16246080  0.14713587   1.104
## cardtenure^28                       0.10821231  0.14733221   0.734
## cardtenure^29                       0.01158058  0.14724177   0.079
## cardtenure^30                       0.05039167  0.14827082   0.340
```

```
## cardtenure^31                          -0.26281379  0.15177497  -1.732
## cardtenure^32                          -0.06373685  0.15250729  -0.418
## cardtenure^33                           0.01491576  0.15079061   0.099
## cardtenure^34                           0.05087551  0.13688226   0.372
## cardtenure^35                           0.28035264  0.13525980   2.073
## cardtenure^36                           0.11916023  0.15344377   0.777
## cardtenure^37                           0.14300040  0.13737893   1.041
## cardtenure^38                           0.01905715  0.14064833   0.135
## cardtenure^39                          -0.01258357  0.14297808  -0.088
## cardtenure^40                           0.09044948  0.15244941   0.593
## card2tenure.L                           0.36313777  0.49197915   0.738
## card2tenure.Q                          -0.35786356  0.36283555  -0.986
## card2tenure.C                           0.44137303  0.32405987   1.362
## card2tenure^4                          -0.17082085  0.29882640  -0.572
## card2tenure^5                          -0.05631070  0.27373097  -0.206
## card2tenure^6                          -0.08388046  0.24346458  -0.345
## card2tenure^7                          -0.11409103  0.21409588  -0.533
## card2tenure^8                           0.11351180  0.18819630   0.603
## card2tenure^9                           0.22964563  0.16528114   1.389
## card2tenure^10                          0.17975280  0.14863535   1.209
## card2tenure^11                         -0.18414073  0.13597557  -1.354
## card2tenure^12                          0.11560752  0.12921691   0.895
## card2tenure^13                          0.08981021  0.12492896   0.719
## card2tenure^14                          0.03103545  0.12415835   0.250
## card2tenure^15                          0.02689017  0.12493885   0.215
## card2tenure^16                          0.00003673  0.12429860   0.000
## card2tenure^17                         -0.17692937  0.12443987  -1.422
## card2tenure^18                         -0.08281831  0.12493006  -0.663
## card2tenure^19                          0.11048333  0.12560228   0.880
## card2tenure^20                          0.14959990  0.12383303   1.208
## card2tenure^21                          0.17452196  0.12705007   1.374
## card2tenure^22                         -0.01824734  0.12455091  -0.147
## card2tenure^23                          0.02123354  0.12550792   0.169
## card2tenure^24                          0.10521503  0.12528129   0.840
## card2tenure^25                          0.21502681  0.12636988   1.702
## card2tenure^26                          0.19010850  0.12399104   1.533
## card2tenure^27                          0.13128064  0.12393168   1.059
## card2tenure^28                         -0.07475545  0.12745550  -0.587
## card2tenure^29                          0.06541995  0.13200516   0.496
## card2tenure^30                         -0.13986422  0.12494525  -1.119
##                                                   Pr(>|t|)
## (Intercept)                            < 0.0000000000000002 ***
## ed                                                0.959230
## lninc                                         0.0000000777 ***
## lncreddebt                                        0.350536
## othdebt                                           0.552405
## carditems                              < 0.0000000000000002 ***
## cardspent                              < 0.0000000000000002 ***
## card2spent                             < 0.0000000000000002 ***
## tollten                                           0.593396
## wireten                                           0.339414
## total_spent                                             NA
## `region_zone 5`                                   0.858323
## `region_zone 3`                                   0.376740
```

```
## `region_zone 4`                                         0.644813
## `region_zone 2`                                         0.945312
## gender_Male                                             0.706260
## `agecat_>65`                                            0.095979 .
## `agecat_25-34`                                          0.934908
## `agecat_50-64`                                          0.238855
## `agecat_35-49`                                          0.161110
## employ_16                                               0.488364
## employ_1                                                0.493609
## employ_22                                               0.420860
## employ_10                                               0.129666
## employ_11                                               0.238104
## employ_15                                               0.391446
## employ_19                                               0.925259
## employ_8                                                0.037647 *
## employ_4                                                0.139971
## employ_12                                               0.534292
## employ_3                                                0.373608
## employ_27                                               0.743569
## employ_31                                               0.812877
## employ_24                                               0.135899
## employ_29                                               0.922655
## employ_2                                                0.124377
## employ_18                                               0.156681
## employ_5                                                0.025838 *
## employ_28                                               0.325908
## employ_26                                               0.263867
## employ_23                                               0.806139
## employ_13                                               0.629442
## employ_17                                               0.534748
## employ_30                                               0.352293
## employ_14                                               0.099149 .
## employ_6                                                0.498908
## employ_36                                               0.804348
## employ_7                                                0.857975
## employ_9                                                0.050405 .
## employ_21                                               0.638181
## employ_20                                               0.126958
## employ_34                                               0.546308
## employ_35                                               0.391455
## employ_33                                               0.455306
## employ_25                                               0.451386
## employ_39                                               0.768829
## employ_45                                               0.111763
## employ_32                                               0.771563
## employ_40                                               0.649968
## employ_37                                               0.500543
## employ_48                                               0.923344
## employ_38                                               0.789263
## employ_46                                               0.918683
## employ_43                                               0.877537
## employ_47                                               0.902760
## employ_41                                               0.108084
## employ_42                                               0.316469
```

```
## employ_49                                                  0.973125
## employ_44                                                  0.277472
## employ_52                                                  0.558424
## employ_51                                                  0.375629
## `empcat_More than 15`                                            NA
## `empcat_6 to 10`                                                 NA
## `empcat_11 to 15`                                                NA
## `empcat_2 to 5`                                                  NA
## retire_Yes                                                 0.241337
## `spousedcat_High school degree`                            0.807101
## `spousedcat_College degree`                                0.300320
## `spousedcat_Some college`                                  0.311796
## `spousedcat_Did not complete high school`                  0.413384
## `spousedcat_Post-undergraduate degree`                     0.674012
## homeown_Own                                                0.211320
## `hometype_Condominium/Townhouse`                           0.612910
## `hometype_Single-family`                                   0.230218
## `hometype_Mobile Home`                                     0.464948
## address_2                                                  0.569615
## address_30                                                 0.527806
## address_3                                                  0.470638
## address_31                                                 0.351382
## address_21                                                 0.149057
## address_20                                                 0.086779 .
## address_19                                                 0.440729
## address_14                                                 0.693287
## address_5                                                  0.279872
## address_9                                                  0.768412
## address_32                                                 0.601685
## address_29                                                 0.208034
## address_8                                                  0.397744
## address_12                                                 0.250764
## address_10                                                 0.489794
## address_38                                                 0.638294
## address_4                                                  0.841342
## address_57                                                 0.587929
## address_16                                                 0.876471
## address_27                                                 0.240186
## address_6                                                  0.578214
## address_35                                                 0.834428
## address_25                                                 0.304840
## address_36                                                 0.169271
## address_22                                                 0.974743
## address_24                                                 0.657739
## address_13                                                 0.439645
## address_11                                                 0.970977
## address_1                                                  0.632247
## address_23                                                 0.784499
## address_26                                                 0.653388
## address_39                                                 0.188131
## address_42                                                 0.372971
## address_17                                                 0.970474
## address_34                                                 0.903950
## address_18                                                 0.900969
```

```
## address_28                                          0.879961
## address_40                                          0.266014
## address_45                                          0.368330
## address_15                                          0.653895
## address_46                                          0.334554
## address_50                                          0.384032
## address_33                                          0.236899
## address_37                                          0.164427
## address_7                                           0.216415
## address_43                                          0.969851
## address_41                                          0.866344
## address_44                                          0.297207
## address_47                                          0.826891
## address_51                                          0.479244
## address_49                                          0.445646
## address_55                                          0.522295
## address_48                                          0.894338
## address_52                                          0.149092
## address_53                                          0.447113
## address_54                                          0.954679
## carown_Lease                                        0.259661
## `carown_N/A`                                        0.000893 ***
## cartype_Import                                      0.642809
## `cartype_N/A`                                             NA
## commutebike_Yes                                     0.143871
## reason_Convenience                                  0.009000 **
## reason_Prices                                       0.014265 *
## reason_Service                                      0.348042
## reason_Other                                        0.165358
## polview_Moderate                                    0.346582
## `polview_Slightly conservative`                     0.560043
## `polview_Slightly liberal`                          0.147168
## polview_Liberal                                     0.833795
## `polview_Extremely liberal`                         0.414809
## `polview_Extremely conservative`                    0.001747 **
## polcontrib_Yes                                      0.924120
## vote_No                                             0.599332
## card_Visa                                           0.521260
## card_Discover                                       0.751264
## card_Other                                          0.771072
## `card_American Express`               < 0.0000000000000002 ***
## cardtype_Other                                      0.017440 *
## cardtype_Gold                                       0.096001 .
## cardtype_Platinum                                   0.544335
## cardbenefit_Other                                   0.141095
## `cardbenefit_Cash back`                             0.249949
## `cardbenefit_Airline miles`                         0.338553
## cardfee_Yes                                         0.598320
## card2_Discover                                      0.237813
## card2_Mastercard                                    0.037072 *
## `card2_American Express`                        0.0000611986 ***
## card2_Visa                                          0.148387
## card2type_None                                      0.146331
## card2type_Gold                                      0.891029
```

```
## card2type_Other                                    0.686364
## `card2benefit_Airline miles`                       0.362225
## card2benefit_Other                                 0.560618
## `card2benefit_Cash back`                           0.708963
## card2fee_Yes                                       0.051382 .
## active_Yes                                         0.617070
## `bfast_Energy bar`                                 0.917812
## bfast_Oatmeal                                      0.309823
## churn_No                                           0.131685
## tollfree_No                                        0.373615
## equip_No                                           0.328922
## callcard_No                                        0.101895
## wireless_Yes                                       0.168291
## multline_No                                        0.940333
## voice_No                                           0.243888
## pager_No                                           0.094776 .
## internet_Other                                     0.498735
## internet_DSL                                       0.461114
## `internet_Cable modem`                             0.483095
## `internet_Dia-up`                                  0.993743
## callid_Yes                                         0.589798
## callwait_No                                        0.121695
## forward_No                                         0.460818
## confer_No                                          0.498765
## ebill_Yes                                          0.723423
## owntv_No                                           0.133281
## ownvcr_No                                          0.451224
## owndvd_No                                          0.861886
## owncd_Yes                                          0.243614
## ownpda_Yes                                         0.126448
## ownpc_Yes                                          0.619171
## ownipod_No                                         0.943395
## owngame_No                                         0.266852
## ownfax_Yes                                         0.663626
## news_Yes                                           0.511306
## response_03_Yes                                    0.053645 .
## edcat.L                                            0.479026
## edcat.Q                                            0.298062
## edcat.C                                            0.741871
## edcat^4                                            0.375397
## inccat.L                                           0.004287 **
## inccat.Q                                       0.0000007749 ***
## inccat.C                                           0.345434
## inccat^4                                           0.678187
## jobsat.L                                           0.404921
## jobsat.Q                                           0.393343
## jobsat.C                                           0.092444 .
## jobsat^4                                           0.399007
## addresscat.L                                             NA
## addresscat.Q                                             NA
## addresscat.C                                             NA
## addresscat^4                                             NA
## carcatvalue.L                                      0.001752 **
## carcatvalue.Q                                      0.018110 *
```

```
## carcatvalue.C                                         NA
## cardtenure.L                                    0.121170
## cardtenure.Q                                    0.192949
## cardtenure.C                                    0.138899
## cardtenure^4                                    0.807225
## cardtenure^5                                    0.356208
## cardtenure^6                                    0.790772
## cardtenure^7                                    0.861539
## cardtenure^8                                    0.502507
## cardtenure^9                                    0.066214 .
## cardtenure^10                                   0.502260
## cardtenure^11                                   0.104915
## cardtenure^12                                   0.263784
## cardtenure^13                                   0.997705
## cardtenure^14                                   0.531982
## cardtenure^15                                   0.853362
## cardtenure^16                                   0.472116
## cardtenure^17                                   0.348571
## cardtenure^18                                   0.476411
## cardtenure^19                                   0.393036
## cardtenure^20                                   0.183374
## cardtenure^21                                   0.693277
## cardtenure^22                                   0.089054 .
## cardtenure^23                                   0.684990
## cardtenure^24                                   0.628875
## cardtenure^25                                   0.010270 *
## cardtenure^26                                   0.934773
## cardtenure^27                                   0.269609
## cardtenure^28                                   0.462711
## cardtenure^29                                   0.937316
## cardtenure^30                                   0.733982
## cardtenure^31                                   0.083441 .
## cardtenure^32                                   0.676029
## cardtenure^33                                   0.921210
## cardtenure^34                                   0.710160
## cardtenure^35                                   0.038280 *
## cardtenure^36                                   0.437468
## cardtenure^37                                   0.297991
## cardtenure^38                                   0.892229
## cardtenure^39                                   0.929874
## cardtenure^40                                   0.553017
## card2tenure.L                                   0.460498
## card2tenure.Q                                   0.324062
## card2tenure.C                                   0.173290
## card2tenure^4                                   0.567607
## card2tenure^5                                   0.837026
## card2tenure^6                                   0.730471
## card2tenure^7                                   0.594142
## card2tenure^8                                   0.546447
## card2tenure^9                                   0.164800
## card2tenure^10                                  0.226616
## card2tenure^11                                  0.175762
## card2tenure^12                                  0.371026
## card2tenure^13                                  0.472261
```

```
## card2tenure^14                                        0.802629
## card2tenure^15                                        0.829604
## card2tenure^16                                        0.999764
## card2tenure^17                                        0.155180
## card2tenure^18                                        0.507431
## card2tenure^19                                        0.379127
## card2tenure^20                                        0.227106
## card2tenure^21                                        0.169647
## card2tenure^22                                        0.883532
## card2tenure^23                                        0.865665
## card2tenure^24                                        0.401066
## card2tenure^25                                        0.088933 .
## card2tenure^26                                        0.125314
## card2tenure^27                                        0.289544
## card2tenure^28                                        0.557566
## card2tenure^29                                        0.620220
## card2tenure^30                                        0.263052
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.202 on 3211 degrees of freedom
## Multiple R-squared:  0.9703, Adjusted R-squared:  0.9677
## F-statistic: 364.8 on 288 and 3211 DF,  p-value: < 0.00000000000000022
```

This model still has variables that are not significant

```
fit2 <- lm(total_spent_sqrt~ed+lncreddebt+carditems+cardspent+card2spent+`agecat_>65`+`agecat_50-64`+
             `agecat_35-49`+employ_1+employ_10+employ_11+employ_15+employ_8+employ_12+employ_27+
             employ_2+employ_18+employ_15+employ_28+employ_26+employ_7+employ_9+employ_20+employ_34+empl
             employ_38+`spousedcat_College degree`+`carown_N/A`+`card_American Express`+cardtype_Other+
             cardtype_Platinum+card2_Discover+card2_Mastercard+`card2_American Express`+card2_Visa+
             churn_No+voice_No+pager_No,data = Dev)
summary(fit2)
```

```
##
## Call:
## lm(formula = total_spent_sqrt ~ ed + lncreddebt + carditems +
##     cardspent + card2spent + `agecat_>65` + `agecat_50-64` +
##     `agecat_35-49` + employ_1 + employ_10 + employ_11 + employ_15 +
##     employ_8 + employ_12 + employ_27 + employ_2 + employ_18 +
##     employ_15 + employ_28 + employ_26 + employ_7 + employ_9 +
##     employ_20 + employ_34 + employ_35 + employ_38 + `spousedcat_College degree` +
##     `carown_N/A` + `card_American Express` + cardtype_Other +
##     cardtype_Platinum + card2_Discover + card2_Mastercard + `card2_American Express` +
##     card2_Visa + churn_No + voice_No + pager_No, data = Dev)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -7.1335 -0.5344  0.2814  0.8302  2.4875
##
## Coefficients:
##                               Estimate Std. Error t value
## (Intercept)                  9.8515618  0.1687412  58.383
## ed                          -0.0104987  0.0074234  -1.414
## lncreddebt                   0.1066813  0.0179933   5.929
```

```
## carditems                     0.2097505  0.0079854  26.267
## cardspent                     0.0176011  0.0001563 112.583
## card2spent                    0.0218682  0.0002092 104.510
## `agecat_>65`                 -0.0451464  0.0656936  -0.687
## `agecat_50-64`                0.1357498  0.0638009   2.128
## `agecat_35-49`                0.2074566  0.0601877   3.447
## employ_1                     -0.0365809  0.0812227  -0.450
## employ_10                    -0.1206429  0.1205264  -1.001
## employ_11                    -0.1190177  0.1124955  -1.058
## employ_15                    -0.0564655  0.1369030  -0.412
## employ_8                     -0.1418240  0.1125437  -1.260
## employ_12                    -0.0169334  0.1352393  -0.125
## employ_27                    -0.1063087  0.2361558  -0.450
## employ_2                     -0.1365370  0.0869772  -1.570
## employ_18                    -0.2428051  0.1711075  -1.419
## employ_28                    -0.3359295  0.2932945  -1.145
## employ_26                    -0.2610638  0.2153703  -1.212
## employ_7                      0.0830489  0.1097625   0.757
## employ_9                     -0.1704133  0.1147773  -1.485
## employ_20                    -0.2649660  0.1640813  -1.615
## employ_34                    -0.1492537  0.2559158  -0.583
## employ_35                    -0.2796039  0.3745426  -0.747
## employ_38                     0.1238703  0.3330160   0.372
## `spousedcat_College degree`   0.0597293  0.0793961   0.752
## `carown_N/A`                 -0.1350345  0.0700827  -1.927
## `card_American Express`       0.4898189  0.0620085   7.899
## cardtype_Other                0.0854191  0.0514157   1.661
## cardtype_Platinum            -0.0106240  0.0508898  -0.209
## card2_Discover               -0.0446165  0.0926348  -0.482
## card2_Mastercard             -0.1649586  0.0883788  -1.866
## `card2_American Express`      0.4229675  0.0966373   4.377
## card2_Visa                   -0.1175534  0.0892676  -1.317
## churn_No                     -0.0871622  0.0517297  -1.685
## voice_No                     -0.1190916  0.0575243  -2.070
## pager_No                      0.0370644  0.0615685   0.602
##                                         Pr(>|t|)
## (Intercept)                  < 0.0000000000000002 ***
## ed                                      0.157372
## lncreddebt                    0.00000000334775247 ***
## carditems                    < 0.0000000000000002 ***
## cardspent                    < 0.0000000000000002 ***
## card2spent                   < 0.0000000000000002 ***
## `agecat_>65`                            0.491986
## `agecat_50-64`                          0.033432 *
## `agecat_35-49`                          0.000574 ***
## employ_1                                0.652466
## employ_10                               0.316913
## employ_11                               0.290139
## employ_15                               0.680036
## employ_8                                0.207694
## employ_12                               0.900364
## employ_27                               0.652621
## employ_2                                0.116553
## employ_18                               0.155983
```

```
## employ_28                                0.252137
## employ_26                                0.225533
## employ_7                                 0.449327
## employ_9                                 0.137706
## employ_20                                0.106435
## employ_34                                0.559787
## employ_35                                0.455404
## employ_38                                0.709942
## `spousedcat_College degree`              0.451925
## `carown_N/A`                             0.054088 .
## `card_American Express`      0.00000000000000374 ***
## cardtype_Other                           0.096735 .
## cardtype_Platinum                        0.834644
## card2_Discover                           0.630093
## card2_Mastercard                         0.062057 .
## `card2_American Express`     0.00001239599924314 ***
## card2_Visa                               0.187971
## churn_No                                 0.092087 .
## voice_No                                 0.038500 *
## pager_No                                 0.547212
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.229 on 3462 degrees of freedom
## Multiple R-squared:  0.9666, Adjusted R-squared:  0.9662
## F-statistic:  2707 on 37 and 3462 DF,  p-value: < 0.00000000000000022
```

To get the best model, a stepwise regression is done

Stepwise regression is used to get better valid predictors

```r
step(fit2,direction = "both")
```

```
## Start:  AIC=1480.2
## total_spent_sqrt ~ ed + lncreddebt + carditems + cardspent +
##     card2spent + `agecat_>65` + `agecat_50-64` + `agecat_35-49` +
##     employ_1 + employ_10 + employ_11 + employ_15 + employ_8 +
##     employ_12 + employ_27 + employ_2 + employ_18 + employ_15 +
##     employ_28 + employ_26 + employ_7 + employ_9 + employ_20 +
##     employ_34 + employ_35 + employ_38 + `spousedcat_College degree` +
##     `carown_N/A` + `card_American Express` + cardtype_Other +
##     cardtype_Platinum + card2_Discover + card2_Mastercard + `card2_American Express` +
##     card2_Visa + churn_No + voice_No + pager_No
##
##                               Df Sum of Sq    RSS    AIC
## - employ_12                    1       0.0  5227.7 1478.2
## - cardtype_Platinum            1       0.1  5227.7 1478.2
## - employ_38                    1       0.2  5227.9 1478.3
## - employ_15                    1       0.3  5227.9 1478.4
## - employ_27                    1       0.3  5228.0 1478.4
## - employ_1                     1       0.3  5228.0 1478.4
## - card2_Discover               1       0.4  5228.0 1478.4
## - employ_34                    1       0.5  5228.2 1478.5
## - pager_No                     1       0.5  5228.2 1478.6
```

```
## - `agecat_>65`                 1       0.7  5228.4 1478.7
## - employ_35                    1       0.8  5228.5 1478.8
## - `spousedcat_College degree`  1       0.9  5228.5 1478.8
## - employ_7                     1       0.9  5228.5 1478.8
## - employ_10                    1       1.5  5229.2 1479.2
## - employ_11                    1       1.7  5229.3 1479.3
## - employ_28                    1       2.0  5229.6 1479.5
## - employ_26                    1       2.2  5229.9 1479.7
## - employ_8                     1       2.4  5230.1 1479.8
## - card2_Visa                   1       2.6  5230.3 1480.0
## <none>                                      5227.7 1480.2
## - ed                           1       3.0  5230.7 1480.2
## - employ_18                    1       3.0  5230.7 1480.2
## - employ_9                     1       3.3  5231.0 1480.4
## - employ_2                     1       3.7  5231.4 1480.7
## - employ_20                    1       3.9  5231.6 1480.8
## - cardtype_Other               1       4.2  5231.8 1481.0
## - churn_No                     1       4.3  5231.9 1481.1
## - card2_Mastercard             1       5.3  5232.9 1481.7
## - `carown_N/A`                 1       5.6  5233.3 1482.0
## - voice_No                     1       6.5  5234.1 1482.5
## - `agecat_50-64`               1       6.8  5234.5 1482.8
## - `agecat_35-49`               1      17.9  5245.6 1490.2
## - `card2_American Express`     1      28.9  5256.6 1497.5
## - lncreddebt                   1      53.1  5280.7 1513.6
## - `card_American Express`      1      94.2  5321.9 1540.7
## - carditems                    1    1041.8  6269.5 2114.3
## - card2spent                   1   16492.7 21720.4 6463.2
## - cardspent                    1   19139.3 24367.0 6865.6
##
## Step:  AIC=1478.22
## total_spent_sqrt ~ ed + lncreddebt + carditems + cardspent +
##     card2spent + `agecat_>65` + `agecat_50-64` + `agecat_35-49` +
##     employ_1 + employ_10 + employ_11 + employ_15 + employ_8 +
##     employ_27 + employ_2 + employ_18 + employ_28 + employ_26 +
##     employ_7 + employ_9 + employ_20 + employ_34 + employ_35 +
##     employ_38 + `spousedcat_College degree` + `carown_N/A` +
##     `card_American Express` + cardtype_Other + cardtype_Platinum +
##     card2_Discover + card2_Mastercard + `card2_American Express` +
##     card2_Visa + churn_No + voice_No + pager_No
##
##                               Df Sum of Sq     RSS    AIC
## - cardtype_Platinum            1       0.1  5227.7 1476.3
## - employ_38                    1       0.2  5227.9 1476.4
## - employ_15                    1       0.2  5227.9 1476.4
## - employ_1                     1       0.3  5228.0 1476.4
## - employ_27                    1       0.3  5228.0 1476.4
## - card2_Discover               1       0.4  5228.0 1476.5
## - employ_34                    1       0.5  5228.2 1476.6
## - pager_No                     1       0.5  5228.2 1476.6
## - `agecat_>65`                 1       0.7  5228.4 1476.7
## - employ_35                    1       0.8  5228.5 1476.8
## - `spousedcat_College degree`  1       0.8  5228.5 1476.8
## - employ_7                     1       0.9  5228.6 1476.8
```

```
## - employ_10                   1       1.5  5229.2 1477.2
## - employ_11                   1       1.7  5229.4 1477.3
## - employ_28                   1       2.0  5229.7 1477.5
## - employ_26                   1       2.2  5229.9 1477.7
## - employ_8                    1       2.4  5230.1 1477.8
## - card2_Visa                  1       2.6  5230.3 1478.0
## <none>                                     5227.7 1478.2
## - ed                          1       3.0  5230.7 1478.2
## - employ_18                   1       3.0  5230.7 1478.2
## - employ_9                    1       3.3  5231.0 1478.4
## - employ_2                    1       3.7  5231.4 1478.7
## - employ_20                   1       3.9  5231.6 1478.8
## - cardtype_Other              1       4.2  5231.8 1479.0
## - churn_No                    1       4.3  5232.0 1479.1
## - card2_Mastercard            1       5.3  5233.0 1479.7
## - `carown_N/A`                1       5.6  5233.3 1480.0
## + employ_12                   1       0.0  5227.7 1480.2
## - voice_No                    1       6.5  5234.1 1480.5
## - `agecat_50-64`              1       6.8  5234.5 1480.8
## - `agecat_35-49`              1      18.0  5245.7 1488.2
## - `card2_American Express`    1      28.9  5256.6 1495.5
## - lncreddebt                  1      53.1  5280.7 1511.6
## - `card_American Express`     1      94.3  5322.0 1538.8
## - carditems                   1    1042.2  6269.9 2112.5
## - card2spent                  1   16496.6 21724.3 6461.8
## - cardspent                   1   19140.5 24368.2 6863.8
##
## Step:  AIC=1476.26
## total_spent_sqrt ~ ed + lncreddebt + carditems + cardspent +
##     card2spent + `agecat_>65` + `agecat_50-64` + `agecat_35-49` +
##     employ_1 + employ_10 + employ_11 + employ_15 + employ_8 +
##     employ_27 + employ_2 + employ_18 + employ_28 + employ_26 +
##     employ_7 + employ_9 + employ_20 + employ_34 + employ_35 +
##     employ_38 + `spousedcat_College degree` + `carown_N/A` +
##     `card_American Express` + cardtype_Other + card2_Discover +
##     card2_Mastercard + `card2_American Express` + card2_Visa +
##     churn_No + voice_No + pager_No
##
##                             Df Sum of Sq     RSS    AIC
## - employ_38                  1       0.2  5228.0 1474.4
## - employ_15                  1       0.2  5228.0 1474.4
## - employ_27                  1       0.3  5228.0 1474.5
## - employ_1                   1       0.3  5228.1 1474.5
## - card2_Discover             1       0.4  5228.1 1474.5
## - employ_34                  1       0.5  5228.3 1474.6
## - pager_No                   1       0.5  5228.3 1474.6
## - `agecat_>65`               1       0.7  5228.5 1474.7
## - employ_35                  1       0.8  5228.6 1474.8
## - `spousedcat_College degree` 1      0.8  5228.6 1474.8
## - employ_7                   1       0.9  5228.6 1474.8
## - employ_10                  1       1.5  5229.3 1475.3
## - employ_11                  1       1.7  5229.4 1475.4
## - employ_28                  1       2.0  5229.7 1475.6
## - employ_26                  1       2.2  5230.0 1475.7
```

```
## - employ_8                    1       2.4  5230.1 1475.8
## - card2_Visa                  1       2.6  5230.4 1476.0
## <none>                                    5227.7 1476.3
## - employ_18                   1       3.0  5230.8 1476.3
## - ed                          1       3.0  5230.8 1476.3
## - employ_9                    1       3.3  5231.1 1476.5
## - employ_2                    1       3.7  5231.4 1476.7
## - employ_20                   1       3.9  5231.7 1476.9
## - churn_No                    1       4.3  5232.0 1477.1
## - cardtype_Other              1       5.1  5232.8 1477.6
## - card2_Mastercard            1       5.3  5233.0 1477.8
## - `carown_N/A`                1       5.6  5233.3 1478.0
## + cardtype_Platinum           1       0.1  5227.7 1478.2
## + employ_12                   1       0.0  5227.7 1478.2
## - voice_No                    1       6.5  5234.2 1478.6
## - `agecat_50-64`              1       6.8  5234.6 1478.8
## - `agecat_35-49`              1      18.0  5245.7 1486.3
## - `card2_American Express`    1      28.9  5256.6 1493.6
## - lncreddebt                  1      53.0  5280.8 1509.6
## - `card_American Express`     1      94.3  5322.0 1536.8
## - carditems                   1    1043.7  6271.4 2111.3
## - card2spent                  1   16527.6 21755.3 6464.8
## - cardspent                   1   19163.2 24391.0 6865.1
##
## Step:  AIC=1474.4
## total_spent_sqrt ~ ed + lncreddebt + carditems + cardspent +
##     card2spent + `agecat_>65` + `agecat_50-64` + `agecat_35-49` +
##     employ_1 + employ_10 + employ_11 + employ_15 + employ_8 +
##     employ_27 + employ_2 + employ_18 + employ_28 + employ_26 +
##     employ_7 + employ_9 + employ_20 + employ_34 + employ_35 +
##     `spousedcat_College degree` + `carown_N/A` + `card_American Express` +
##     cardtype_Other + card2_Discover + card2_Mastercard + `card2_American Express` +
##     card2_Visa + churn_No + voice_No + pager_No
##
##                              Df Sum of Sq     RSS    AIC
## - employ_15                   1       0.3  5228.2 1472.6
## - employ_27                   1       0.3  5228.3 1472.6
## - employ_1                    1       0.3  5228.3 1472.6
## - card2_Discover              1       0.3  5228.3 1472.6
## - employ_34                   1       0.5  5228.5 1472.8
## - pager_No                    1       0.6  5228.5 1472.8
## - `agecat_>65`                1       0.7  5228.6 1472.8
## - `spousedcat_College degree` 1       0.8  5228.8 1473.0
## - employ_35                   1       0.9  5228.8 1473.0
## - employ_7                    1       0.9  5228.8 1473.0
## - employ_10                   1       1.5  5229.5 1473.4
## - employ_11                   1       1.7  5229.7 1473.5
## - employ_28                   1       2.0  5229.9 1473.7
## - employ_26                   1       2.2  5230.2 1473.9
## - employ_8                    1       2.4  5230.4 1474.0
## - card2_Visa                  1       2.6  5230.6 1474.1
## <none>                                    5228.0 1474.4
## - ed                          1       3.0  5231.0 1474.4
## - employ_18                   1       3.1  5231.0 1474.4
```

```
## - employ_9                    1       3.3  5231.3 1474.6
## - employ_2                    1       3.7  5231.7 1474.9
## - employ_20                   1       3.9  5231.9 1475.0
## - churn_No                    1       4.3  5232.2 1475.3
## - cardtype_Other              1       5.1  5233.0 1475.8
## - card2_Mastercard            1       5.2  5233.2 1475.9
## - `carown_N/A`                1       5.6  5233.5 1476.1
## + employ_38                   1       0.2  5227.7 1476.3
## + cardtype_Platinum           1       0.1  5227.9 1476.4
## + employ_12                   1       0.0  5227.9 1476.4
## - voice_No                    1       6.5  5234.5 1476.8
## - `agecat_50-64`              1       6.8  5234.8 1477.0
## - `agecat_35-49`              1      18.0  5246.0 1484.4
## - `card2_American Express`    1      29.0  5257.0 1491.8
## - lncreddebt                  1      53.2  5281.1 1507.8
## - `card_American Express`     1      94.3  5322.3 1535.0
## - carditems                   1    1043.5  6271.5 2109.4
## - card2spent                  1   16527.4 21755.4 6462.8
## - cardspent                   1   19167.8 24395.8 6863.8
##
## Step:  AIC=1472.57
## total_spent_sqrt ~ ed + lncreddebt + carditems + cardspent +
##     card2spent + `agecat_>65` + `agecat_50-64` + `agecat_35-49` +
##     employ_1 + employ_10 + employ_11 + employ_8 + employ_27 +
##     employ_2 + employ_18 + employ_28 + employ_26 + employ_7 +
##     employ_9 + employ_20 + employ_34 + employ_35 + `spousedcat_College degree` +
##     `carown_N/A` + `card_American Express` + cardtype_Other +
##     card2_Discover + card2_Mastercard + `card2_American Express` +
##     card2_Visa + churn_No + voice_No + pager_No
##
##                              Df Sum of Sq     RSS    AIC
## - employ_1                    1       0.3  5228.5 1470.8
## - employ_27                   1       0.3  5228.5 1470.8
## - card2_Discover              1       0.4  5228.6 1470.8
## - employ_34                   1       0.5  5228.7 1470.9
## - pager_No                    1       0.6  5228.8 1470.9
## - `agecat_>65`                1       0.7  5228.9 1471.1
## - employ_35                   1       0.8  5229.0 1471.1
## - `spousedcat_College degree` 1       0.9  5229.1 1471.1
## - employ_7                    1       0.9  5229.1 1471.2
## - employ_10                   1       1.5  5229.7 1471.6
## - employ_11                   1       1.6  5229.8 1471.7
## - employ_28                   1       2.0  5230.2 1471.9
## - employ_26                   1       2.2  5230.4 1472.0
## - employ_8                    1       2.3  5230.5 1472.1
## - card2_Visa                  1       2.6  5230.8 1472.3
## - employ_18                   1       3.0  5231.2 1472.6
## <none>                                     5228.2 1472.6
## - ed                          1       3.0  5231.2 1472.6
## - employ_9                    1       3.2  5231.4 1472.7
## - employ_2                    1       3.6  5231.8 1473.0
## - employ_20                   1       3.9  5232.1 1473.2
## - churn_No                    1       4.3  5232.5 1473.4
## - cardtype_Other              1       5.1  5233.3 1474.0
```

```
## - card2_Mastercard            1       5.2  5233.5 1474.1
## - `carown_N/A`                1       5.6  5233.8 1474.3
## + employ_15                   1       0.3  5228.0 1474.4
## + employ_38                   1       0.2  5228.0 1474.4
## + cardtype_Platinum           1       0.1  5228.2 1474.5
## + employ_12                   1       0.0  5228.2 1474.6
## - voice_No                    1       6.6  5234.8 1475.0
## - `agecat_50-64`              1       6.6  5234.8 1475.0
## - `agecat_35-49`              1      17.8  5246.0 1482.4
## - `card2_American Express`    1      29.0  5257.2 1489.9
## - lncreddebt                  1      53.0  5281.2 1505.9
## - `card_American Express`     1      94.6  5322.8 1533.3
## - carditems                   1    1044.9  6273.1 2108.3
## - card2spent                  1   16536.7 21764.9 6462.4
## - cardspent                   1   19174.4 24402.6 6862.7
##
## Step:  AIC=1470.76
## total_spent_sqrt ~ ed + lncreddebt + carditems + cardspent +
##     card2spent + `agecat_>65` + `agecat_50-64` + `agecat_35-49` +
##     employ_10 + employ_11 + employ_8 + employ_27 + employ_2 +
##     employ_18 + employ_28 + employ_26 + employ_7 + employ_9 +
##     employ_20 + employ_34 + employ_35 + `spousedcat_College degree` +
##     `carown_N/A` + `card_American Express` + cardtype_Other +
##     card2_Discover + card2_Mastercard + `card2_American Express` +
##     card2_Visa + churn_No + voice_No + pager_No
##
##                               Df Sum of Sq     RSS    AIC
## - employ_27                    1       0.3  5228.8 1469.0
## - card2_Discover               1       0.4  5228.9 1469.0
## - employ_34                    1       0.5  5229.0 1469.1
## - pager_No                     1       0.6  5229.1 1469.1
## - `agecat_>65`                 1       0.6  5229.1 1469.1
## - employ_35                    1       0.8  5229.3 1469.3
## - `spousedcat_College degree`  1       0.9  5229.3 1469.3
## - employ_7                     1       1.0  5229.5 1469.4
## - employ_10                    1       1.4  5229.9 1469.7
## - employ_11                    1       1.6  5230.1 1469.8
## - employ_28                    1       1.9  5230.4 1470.1
## - employ_26                    1       2.2  5230.7 1470.2
## - employ_8                     1       2.2  5230.7 1470.2
## - card2_Visa                   1       2.6  5231.1 1470.5
## - employ_18                    1       2.9  5231.4 1470.7
## <none>                                      5228.5 1470.8
## - ed                           1       3.0  5231.5 1470.8
## - employ_9                     1       3.1  5231.6 1470.9
## - employ_2                     1       3.4  5231.9 1471.0
## - employ_20                    1       3.8  5232.3 1471.3
## - churn_No                     1       4.2  5232.7 1471.6
## - cardtype_Other               1       5.1  5233.6 1472.2
## - card2_Mastercard             1       5.2  5233.7 1472.2
## - `carown_N/A`                 1       5.6  5234.1 1472.5
## + employ_1                     1       0.3  5228.2 1472.6
## + employ_15                    1       0.2  5228.3 1472.6
## + employ_38                    1       0.2  5228.3 1472.6
```

```
## + cardtype_Platinum            1       0.1  5228.4 1472.7
## + employ_12                    1       0.0  5228.5 1472.8
## - voice_No                     1       6.6  5235.1 1473.2
## - `agecat_50-64`               1       7.5  5236.0 1473.8
## - `agecat_35-49`               1      18.6  5247.1 1481.2
## - `card2_American Express`     1      29.1  5257.6 1488.2
## - lncreddebt                   1      53.2  5281.7 1504.2
## - `card_American Express`      1      94.6  5323.1 1531.5
## - carditems                    1    1045.2  6273.7 2106.6
## - card2spent                   1   16545.7 21774.2 6461.9
## - cardspent                    1   19177.0 24405.5 6861.2
##
## Step:  AIC=1468.96
## total_spent_sqrt ~ ed + lncreddebt + carditems + cardspent +
##     card2spent + `agecat_>65` + `agecat_50-64` + `agecat_35-49` +
##     employ_10 + employ_11 + employ_8 + employ_2 + employ_18 +
##     employ_28 + employ_26 + employ_7 + employ_9 + employ_20 +
##     employ_34 + employ_35 + `spousedcat_College degree` + `carown_N/A` +
##     `card_American Express` + cardtype_Other + card2_Discover +
##     card2_Mastercard + `card2_American Express` + card2_Visa +
##     churn_No + voice_No + pager_No
##
##                               Df Sum of Sq     RSS    AIC
## - card2_Discover               1       0.4  5229.1 1467.2
## - employ_34                    1       0.5  5229.3 1467.3
## - pager_No                     1       0.6  5229.3 1467.3
## - `agecat_>65`                 1       0.7  5229.4 1467.4
## - employ_35                    1       0.8  5229.6 1467.5
## - `spousedcat_College degree`  1       0.8  5229.6 1467.5
## - employ_7                     1       1.0  5229.8 1467.6
## - employ_10                    1       1.4  5230.2 1467.9
## - employ_11                    1       1.5  5230.3 1468.0
## - employ_28                    1       1.9  5230.7 1468.2
## - employ_26                    1       2.1  5230.9 1468.4
## - employ_8                     1       2.2  5231.0 1468.4
## - card2_Visa                   1       2.5  5231.3 1468.7
## - employ_18                    1       2.9  5231.7 1468.9
## <none>                                      5228.8 1469.0
## - ed                           1       3.0  5231.8 1469.0
## - employ_9                     1       3.1  5231.9 1469.0
## - employ_2                     1       3.4  5232.2 1469.2
## - employ_20                    1       3.8  5232.5 1469.5
## - churn_No                     1       4.2  5233.0 1469.8
## - cardtype_Other               1       5.1  5233.9 1470.4
## - card2_Mastercard             1       5.2  5234.0 1470.4
## - `carown_N/A`                 1       5.6  5234.3 1470.7
## + employ_27                    1       0.3  5228.5 1470.8
## + employ_1                     1       0.3  5228.5 1470.8
## + employ_38                    1       0.2  5228.6 1470.8
## + employ_15                    1       0.2  5228.6 1470.8
## + cardtype_Platinum            1       0.1  5228.7 1470.9
## + employ_12                    1       0.0  5228.8 1471.0
## - voice_No                     1       6.7  5235.4 1471.4
## - `agecat_50-64`               1       7.4  5236.1 1471.9
```

```
## - `agecat_35-49`              1      18.6  5247.3 1479.4
## - `card2_American Express`     1      29.1  5257.9 1486.4
## - lncreddebt                   1      53.1  5281.9 1502.3
## - `card_American Express`      1      94.3  5323.1 1529.5
## - carditems                    1    1045.2  6274.0 2104.8
## - card2spent                   1   16545.6 21774.4 6459.9
## - cardspent                    1   19178.8 24407.6 6859.5
##
## Step:  AIC=1467.19
## total_spent_sqrt ~ ed + lncreddebt + carditems + cardspent +
##     card2spent + `agecat_>65` + `agecat_50-64` + `agecat_35-49` +
##     employ_10 + employ_11 + employ_8 + employ_2 + employ_18 +
##     employ_28 + employ_26 + employ_7 + employ_9 + employ_20 +
##     employ_34 + employ_35 + `spousedcat_College degree` + `carown_N/A` +
##     `card_American Express` + cardtype_Other + card2_Mastercard +
##     `card2_American Express` + card2_Visa + churn_No + voice_No +
##     pager_No
##
##                              Df Sum of Sq     RSS    AIC
## - employ_34                   1       0.5  5229.6 1465.5
## - pager_No                    1       0.6  5229.7 1465.6
## - `agecat_>65`                1       0.7  5229.8 1465.6
## - employ_35                   1       0.8  5229.9 1465.7
## - `spousedcat_College degree` 1       0.8  5230.0 1465.8
## - employ_7                    1       1.0  5230.1 1465.9
## - employ_10                   1       1.4  5230.5 1466.1
## - employ_11                   1       1.5  5230.6 1466.2
## - employ_28                   1       1.9  5231.1 1466.5
## - employ_8                    1       2.2  5231.3 1466.6
## - employ_26                   1       2.2  5231.3 1466.7
## - employ_18                   1       2.9  5232.1 1467.2
## <none>                                     5229.1 1467.2
## - employ_9                    1       3.2  5232.3 1467.3
## - card2_Visa                  1       3.3  5232.4 1467.4
## - employ_2                    1       3.3  5232.5 1467.4
## - ed                          1       3.3  5232.5 1467.4
## - employ_20                   1       3.7  5232.9 1467.7
## - churn_No                    1       4.3  5233.4 1468.1
## - cardtype_Other              1       5.0  5234.2 1468.6
## - `carown_N/A`                1       5.6  5234.7 1468.9
## + card2_Discover              1       0.4  5228.8 1469.0
## + employ_1                    1       0.3  5228.9 1469.0
## + employ_27                   1       0.3  5228.9 1469.0
## + employ_15                   1       0.2  5228.9 1469.0
## + employ_38                   1       0.2  5228.9 1469.0
## + cardtype_Platinum           1       0.1  5229.1 1469.2
## + employ_12                   1       0.0  5229.1 1469.2
## - voice_No                    1       6.6  5235.8 1469.6
## - `agecat_50-64`              1       7.3  5236.4 1470.1
## - card2_Mastercard            1       8.4  5237.6 1470.8
## - `agecat_35-49`              1      18.5  5247.6 1477.5
## - lncreddebt                  1      52.9  5282.0 1500.4
## - `card2_American Express`    1      66.5  5295.6 1509.4
## - `card_American Express`     1      95.5  5324.7 1528.5
```

```
## - carditems                    1    1045.1  6274.3 2102.9
## - card2spent                   1   16559.1 21788.3 6460.1
## - cardspent                    1   19179.0 24408.1 6857.5
##
## Step:  AIC=1465.52
## total_spent_sqrt ~ ed + lncreddebt + carditems + cardspent +
##     card2spent + `agecat_>65` + `agecat_50-64` + `agecat_35-49` +
##     employ_10 + employ_11 + employ_8 + employ_2 + employ_18 +
##     employ_28 + employ_26 + employ_7 + employ_9 + employ_20 +
##     employ_35 + `spousedcat_College degree` + `carown_N/A` +
##     `card_American Express` + cardtype_Other + card2_Mastercard +
##     `card2_American Express` + card2_Visa + churn_No + voice_No +
##     pager_No
##
##                               Df Sum of Sq     RSS    AIC
## - pager_No                     1       0.6  5230.2 1463.9
## - employ_35                    1       0.8  5230.4 1464.0
## - `agecat_>65`                 1       0.8  5230.4 1464.1
## - `spousedcat_College degree`  1       0.9  5230.5 1464.1
## - employ_7                     1       1.0  5230.7 1464.2
## - employ_10                    1       1.4  5231.0 1464.4
## - employ_11                    1       1.5  5231.1 1464.5
## - employ_28                    1       1.9  5231.5 1464.8
## - employ_26                    1       2.1  5231.8 1464.9
## - employ_8                     1       2.1  5231.8 1464.9
## - employ_18                    1       2.9  5232.5 1465.4
## <none>                                      5229.6 1465.5
## - employ_9                     1       3.1  5232.7 1465.6
## - ed                           1       3.2  5232.9 1465.7
## - card2_Visa                   1       3.2  5232.9 1465.7
## - employ_2                     1       3.3  5232.9 1465.7
## - employ_20                    1       3.7  5233.3 1466.0
## - churn_No                     1       4.3  5234.0 1466.4
## - cardtype_Other               1       5.1  5234.7 1466.9
## + employ_34                    1       0.5  5229.1 1467.2
## + card2_Discover               1       0.4  5229.3 1467.3
## - `carown_N/A`                 1       5.7  5235.3 1467.3
## + employ_1                     1       0.3  5229.4 1467.3
## + employ_27                    1       0.3  5229.4 1467.4
## + employ_38                    1       0.2  5229.4 1467.4
## + employ_15                    1       0.2  5229.4 1467.4
## + cardtype_Platinum            1       0.1  5229.6 1467.5
## + employ_12                    1       0.0  5229.6 1467.5
## - voice_No                     1       6.6  5236.2 1467.9
## - `agecat_50-64`               1       7.3  5236.9 1468.4
## - card2_Mastercard             1       8.4  5238.0 1469.1
## - `agecat_35-49`               1      18.5  5248.1 1475.9
## - lncreddebt                   1      52.5  5282.2 1498.5
## - `card2_American Express`     1      66.6  5296.2 1507.8
## - `card_American Express`      1      95.4  5325.1 1526.8
## - carditems                    1    1045.4  6275.0 2101.3
## - card2spent                   1   16573.2 21802.8 6460.5
## - cardspent                    1   19179.0 24408.7 6855.6
##
```

```
## Step:  AIC=1463.9
## total_spent_sqrt ~ ed + lncreddebt + carditems + cardspent +
##     card2spent + `agecat_>65` + `agecat_50-64` + `agecat_35-49` +
##     employ_10 + employ_11 + employ_8 + employ_2 + employ_18 +
##     employ_28 + employ_26 + employ_7 + employ_9 + employ_20 +
##     employ_35 + `spousedcat_College degree` + `carown_N/A` +
##     `card_American Express` + cardtype_Other + card2_Mastercard +
##     `card2_American Express` + card2_Visa + churn_No + voice_No
##
##                              Df Sum of Sq     RSS    AIC
## - `agecat_>65`                1       0.8  5231.0 1462.4
## - employ_35                   1       0.8  5231.0 1462.4
## - `spousedcat_College degree`  1       0.9  5231.1 1462.5
## - employ_7                    1       1.1  5231.3 1462.6
## - employ_10                   1       1.4  5231.6 1462.8
## - employ_11                   1       1.4  5231.6 1462.9
## - employ_28                   1       1.9  5232.1 1463.2
## - employ_8                    1       2.1  5232.3 1463.3
## - employ_26                   1       2.1  5232.3 1463.3
## - employ_18                   1       2.9  5233.1 1463.8
## <none>                                     5230.2 1463.9
## - employ_9                    1       3.1  5233.3 1464.0
## - card2_Visa                  1       3.2  5233.4 1464.0
## - employ_2                    1       3.3  5233.5 1464.1
## - employ_20                   1       3.7  5233.9 1464.4
## - ed                          1       3.7  5233.9 1464.4
## - churn_No                    1       4.1  5234.3 1464.6
## - cardtype_Other              1       5.1  5235.3 1465.3
## + pager_No                    1       0.6  5229.6 1465.5
## + employ_34                   1       0.5  5229.7 1465.6
## - `carown_N/A`                1       5.6  5235.8 1465.6
## + card2_Discover              1       0.4  5229.8 1465.7
## + employ_1                    1       0.3  5229.9 1465.7
## + employ_38                   1       0.3  5229.9 1465.7
## + employ_27                   1       0.3  5229.9 1465.7
## + employ_15                   1       0.2  5230.0 1465.8
## + cardtype_Platinum           1       0.1  5230.1 1465.9
## + employ_12                   1       0.0  5230.2 1465.9
## - voice_No                    1       6.5  5236.7 1466.3
## - `agecat_50-64`              1       7.1  5237.3 1466.7
## - card2_Mastercard            1       8.3  5238.5 1467.5
## - `agecat_35-49`              1      18.2  5248.4 1474.1
## - lncreddebt                  1      52.5  5282.7 1496.9
## - `card2_American Express`    1      66.6  5296.8 1506.2
## - `card_American Express`     1      95.5  5325.7 1525.2
## - carditems                   1    1047.2  6277.4 2100.7
## - card2spent                  1   16591.2 21821.4 6461.5
## - cardspent                   1   19202.4 24432.6 6857.0
##
## Step:  AIC=1462.43
## total_spent_sqrt ~ ed + lncreddebt + carditems + cardspent +
##     card2spent + `agecat_50-64` + `agecat_35-49` + employ_10 +
##     employ_11 + employ_8 + employ_2 + employ_18 + employ_28 +
##     employ_26 + employ_7 + employ_9 + employ_20 + employ_35 +
```

```
##     `spousedcat_College degree` + `carown_N/A` + `card_American Express` +
##     cardtype_Other + card2_Mastercard + `card2_American Express` +
##     card2_Visa + churn_No + voice_No
##
##                               Df Sum of Sq     RSS    AIC
## - `spousedcat_College degree`  1       0.9  5231.8 1461.0
## - employ_35                   1       1.0  5231.9 1461.1
## - employ_7                    1       1.0  5232.0 1461.1
## - employ_10                   1       1.5  5232.5 1461.4
## - employ_11                   1       1.7  5232.7 1461.5
## - employ_28                   1       2.2  5233.1 1461.9
## - employ_8                    1       2.2  5233.1 1461.9
## - employ_26                   1       2.5  5233.4 1462.1
## <none>                                     5231.0 1462.4
## - employ_2                    1       3.0  5234.0 1462.5
## - card2_Visa                  1       3.2  5234.2 1462.6
## - employ_18                   1       3.2  5234.2 1462.6
## - employ_9                    1       3.3  5234.3 1462.6
## - ed                          1       3.5  5234.5 1462.8
## - employ_20                   1       4.0  5235.0 1463.1
## - churn_No                    1       5.1  5236.0 1463.8
## - cardtype_Other              1       5.1  5236.1 1463.8
## + `agecat_>65`                1       0.8  5230.2 1463.9
## + employ_34                   1       0.6  5230.3 1464.0
## + pager_No                    1       0.5  5230.4 1464.1
## - `carown_N/A`                1       5.6  5236.6 1464.2
## + card2_Discover              1       0.4  5230.6 1464.2
## + employ_27                   1       0.4  5230.6 1464.2
## + employ_15                   1       0.3  5230.7 1464.2
## + employ_38                   1       0.2  5230.8 1464.3
## + employ_1                    1       0.1  5230.9 1464.4
## + cardtype_Platinum           1       0.1  5230.9 1464.4
## + employ_12                   1       0.0  5231.0 1464.4
## - voice_No                    1       6.7  5237.7 1464.9
## - card2_Mastercard            1       8.3  5239.3 1466.0
## - `agecat_50-64`              1      12.3  5243.3 1468.7
## - `agecat_35-49`              1      27.0  5257.9 1478.4
## - lncreddebt                  1      52.0  5282.9 1495.0
## - `card2_American Express`    1      66.6  5297.6 1504.7
## - `card_American Express`     1      95.6  5326.6 1523.8
## - carditems                   1    1050.9  6281.9 2101.2
## - card2spent                  1   16610.5 21841.5 6462.7
## - cardspent                   1   19209.1 24440.1 6856.1
##
## Step:  AIC=1461.01
## total_spent_sqrt ~ ed + lncreddebt + carditems + cardspent +
##     card2spent + `agecat_50-64` + `agecat_35-49` + employ_10 +
##     employ_11 + employ_8 + employ_2 + employ_18 + employ_28 +
##     employ_26 + employ_7 + employ_9 + employ_20 + employ_35 +
##     `carown_N/A` + `card_American Express` + cardtype_Other +
##     card2_Mastercard + `card2_American Express` + card2_Visa +
##     churn_No + voice_No
##
##                               Df Sum of Sq     RSS    AIC
```

```
## - employ_35                    1       0.9  5232.8 1459.6
## - employ_7                     1       1.0  5232.9 1459.7
## - employ_10                    1       1.4  5233.2 1459.9
## - employ_11                    1       1.6  5233.5 1460.1
## - employ_28                    1       2.1  5234.0 1460.4
## - employ_8                     1       2.2  5234.1 1460.5
## - employ_26                    1       2.5  5234.3 1460.6
## - ed                           1       3.0  5234.8 1461.0
## <none>                                      5231.8 1461.0
## - employ_2                     1       3.0  5234.9 1461.0
## - employ_18                    1       3.2  5235.0 1461.1
## - card2_Visa                   1       3.2  5235.1 1461.2
## - employ_9                     1       3.3  5235.1 1461.2
## - employ_20                    1       4.1  5235.9 1461.7
## - churn_No                     1       5.0  5236.9 1462.4
## + `spousedcat_College degree`  1       0.9  5231.0 1462.4
## - cardtype_Other               1       5.1  5237.0 1462.4
## + `agecat_>65`                 1       0.8  5231.1 1462.5
## + employ_34                    1       0.7  5231.2 1462.6
## + pager_No                     1       0.6  5231.3 1462.6
## - `carown_N/A`                 1       5.6  5237.4 1462.7
## + card2_Discover               1       0.4  5231.5 1462.8
## + employ_27                    1       0.3  5231.5 1462.8
## + employ_15                    1       0.3  5231.5 1462.8
## + employ_38                    1       0.2  5231.7 1462.9
## + employ_1                     1       0.1  5231.7 1462.9
## + cardtype_Platinum            1       0.0  5231.8 1463.0
## + employ_12                    1       0.0  5231.8 1463.0
## - voice_No                     1       6.8  5238.7 1463.6
## - card2_Mastercard             1       8.3  5240.2 1464.6
## - `agecat_50-64`               1      12.5  5244.3 1467.3
## - `agecat_35-49`               1      27.0  5258.9 1477.0
## - lncreddebt                   1      51.5  5283.4 1493.3
## - `card2_American Express`     1      67.0  5298.9 1503.6
## - `card_American Express`      1      95.3  5327.1 1522.2
## - carditems                    1    1050.4  6282.3 2099.4
## - card2spent                   1   16611.8 21843.6 6461.0
## - cardspent                    1   19217.7 24449.5 6855.5
##
## Step:  AIC=1459.63
## total_spent_sqrt ~ ed + lncreddebt + carditems + cardspent +
##     card2spent + `agecat_50-64` + `agecat_35-49` + employ_10 +
##     employ_11 + employ_8 + employ_2 + employ_18 + employ_28 +
##     employ_26 + employ_7 + employ_9 + employ_20 + `carown_N/A` +
##     `card_American Express` + cardtype_Other + card2_Mastercard +
##     `card2_American Express` + card2_Visa + churn_No + voice_No
##
##                               Df Sum of Sq     RSS    AIC
## - employ_7                     1       1.1  5233.8 1458.3
## - employ_10                    1       1.4  5234.1 1458.5
## - employ_11                    1       1.6  5234.4 1458.7
## - employ_28                    1       2.1  5234.9 1459.0
## - employ_8                     1       2.2  5234.9 1459.1
## - employ_26                    1       2.4  5235.2 1459.3
```

```
## - ed                          1       2.9  5235.6 1459.5
## - employ_2                    1       3.0  5235.8 1459.6
## <none>                                     5232.8 1459.6
## - employ_18                   1       3.1  5235.9 1459.7
## - card2_Visa                  1       3.2  5236.0 1459.8
## - employ_9                    1       3.2  5236.0 1459.8
## - employ_20                   1       4.0  5236.8 1460.3
## + employ_35                   1       0.9  5231.8 1461.0
## + `agecat_>65`                1       0.9  5231.9 1461.0
## - churn_No                    1       5.1  5237.9 1461.0
## + `spousedcat_College degree`  1       0.8  5231.9 1461.1
## - cardtype_Other              1       5.2  5238.0 1461.1
## + employ_34                   1       0.6  5232.1 1461.2
## + pager_No                    1       0.6  5232.2 1461.2
## - `carown_N/A`                1       5.5  5238.3 1461.3
## + card2_Discover              1       0.4  5232.4 1461.4
## + employ_27                   1       0.3  5232.5 1461.4
## + employ_15                   1       0.3  5232.5 1461.4
## + employ_38                   1       0.2  5232.6 1461.5
## + employ_1                    1       0.1  5232.7 1461.6
## + cardtype_Platinum           1       0.0  5232.7 1461.6
## + employ_12                   1       0.0  5232.8 1461.6
## - voice_No                    1       6.8  5239.6 1462.2
## - card2_Mastercard            1       8.2  5240.9 1463.1
## - `agecat_50-64`              1      12.7  5245.5 1466.1
## - `agecat_35-49`              1      27.4  5260.2 1475.9
## - lncreddebt                  1      51.2  5284.0 1491.7
## - `card2_American Express`    1      67.2  5299.9 1502.3
## - `card_American Express`     1      95.6  5328.4 1521.0
## - carditems                   1    1051.0  6283.7 2098.2
## - card2spent                  1   16612.1 21844.8 6459.2
## - cardspent                   1   19219.2 24452.0 6853.8
##
## Step:  AIC=1458.34
## total_spent_sqrt ~ ed + lncreddebt + carditems + cardspent +
##     card2spent + `agecat_50-64` + `agecat_35-49` + employ_10 +
##     employ_11 + employ_8 + employ_2 + employ_18 + employ_28 +
##     employ_26 + employ_9 + employ_20 + `carown_N/A` + `card_American Express` +
##     cardtype_Other + card2_Mastercard + `card2_American Express` +
##     card2_Visa + churn_No + voice_No
##
##                              Df Sum of Sq     RSS    AIC
## - employ_10                   1       1.5  5235.3 1457.3
## - employ_11                   1       1.8  5235.6 1457.5
## - employ_28                   1       2.2  5236.0 1457.8
## - employ_8                    1       2.4  5236.2 1457.9
## - employ_26                   1       2.5  5236.3 1458.0
## - ed                          1       2.9  5236.7 1458.3
## <none>                                     5233.8 1458.3
## - card2_Visa                  1       3.2  5237.1 1458.5
## - employ_2                    1       3.2  5237.1 1458.5
## - employ_18                   1       3.3  5237.1 1458.5
## - employ_9                    1       3.5  5237.3 1458.7
## - employ_20                   1       4.2  5238.0 1459.1
```

```
## + employ_7                    1       1.1  5232.8 1459.6
## + employ_35                   1       1.0  5232.9 1459.7
## + `spousedcat_College degree`  1       0.9  5232.9 1459.7
## - churn_No                    1       5.1  5239.0 1459.8
## + `agecat_>65`                1       0.8  5233.0 1459.8
## - cardtype_Other              1       5.3  5239.1 1459.9
## + employ_34                   1       0.7  5233.2 1459.9
## + pager_No                    1       0.6  5233.2 1459.9
## - `carown_N/A`                1       5.4  5239.2 1459.9
## + employ_27                   1       0.4  5233.5 1460.1
## + employ_15                   1       0.3  5233.5 1460.1
## + card2_Discover              1       0.3  5233.5 1460.1
## + employ_38                   1       0.2  5233.7 1460.2
## + employ_1                    1       0.1  5233.7 1460.2
## + cardtype_Platinum           1       0.0  5233.8 1460.3
## + employ_12                   1       0.0  5233.8 1460.3
## - voice_No                    1       6.8  5240.6 1460.9
## - card2_Mastercard            1       8.2  5242.1 1461.8
## - `agecat_50-64`              1      13.1  5246.9 1465.1
## - `agecat_35-49`              1      28.3  5262.1 1475.2
## - lncreddebt                  1      51.1  5285.0 1490.4
## - `card2_American Express`    1      67.3  5301.2 1501.1
## - `card_American Express`     1      96.3  5330.1 1520.1
## - carditems                   1    1049.9  6283.7 2096.2
## - card2spent                  1   16612.4 21846.3 6457.4
## - cardspent                   1   19224.4 24458.2 6852.7
##
## Step:  AIC=1457.35
## total_spent_sqrt ~ ed + lncreddebt + carditems + cardspent +
##     card2spent + `agecat_50-64` + `agecat_35-49` + employ_11 +
##     employ_8 + employ_2 + employ_18 + employ_28 + employ_26 +
##     employ_9 + employ_20 + `carown_N/A` + `card_American Express` +
##     cardtype_Other + card2_Mastercard + `card2_American Express` +
##     card2_Visa + churn_No + voice_No
##
##                              Df Sum of Sq     RSS    AIC
## - employ_11                   1       1.6  5237.0 1456.4
## - employ_28                   1       2.1  5237.5 1456.8
## - employ_8                    1       2.2  5237.5 1456.8
## - employ_26                   1       2.4  5237.8 1457.0
## - ed                          1       2.8  5238.1 1457.2
## <none>                                     5235.3 1457.3
## - employ_2                    1       3.0  5238.4 1457.4
## - card2_Visa                  1       3.1  5238.5 1457.4
## - employ_18                   1       3.1  5238.5 1457.4
## - employ_9                    1       3.2  5238.6 1457.5
## - employ_20                   1       4.0  5239.4 1458.0
## + employ_10                   1       1.5  5233.8 1458.3
## + employ_7                    1       1.2  5234.1 1458.5
## + employ_35                   1       0.9  5234.4 1458.7
## + `agecat_>65`                1       0.9  5234.4 1458.7
## + `spousedcat_College degree`  1       0.8  5234.5 1458.8
## - churn_No                    1       5.3  5240.6 1458.9
## - cardtype_Other              1       5.3  5240.6 1458.9
```

```
## + pager_No                    1       0.7  5234.7 1458.9
## + employ_34                   1       0.6  5234.7 1458.9
## - `carown_N/A`                1       5.6  5241.0 1459.1
## + card2_Discover              1       0.3  5235.0 1459.1
## + employ_27                   1       0.3  5235.0 1459.1
## + employ_15                   1       0.3  5235.1 1459.2
## + employ_38                   1       0.2  5235.2 1459.2
## + employ_1                    1       0.1  5235.3 1459.3
## + cardtype_Platinum           1       0.1  5235.3 1459.3
## + employ_12                   1       0.0  5235.3 1459.3
## - voice_No                    1       6.7  5242.0 1459.8
## - card2_Mastercard            1       8.2  5243.5 1460.8
## - `agecat_50-64`              1      12.5  5247.8 1463.7
## - `agecat_35-49`              1      27.3  5262.6 1473.5
## - lncreddebt                  1      51.5  5286.8 1489.6
## - `card2_American Express`    1      67.7  5303.1 1500.3
## - `card_American Express`     1      96.3  5331.7 1519.2
## - carditems                   1    1056.5  6291.8 2098.7
## - card2spent                  1   16617.2 21852.5 6456.4
## - cardspent                   1   19225.5 24460.9 6851.1
##
## Step:  AIC=1456.42
## total_spent_sqrt ~ ed + lncreddebt + carditems + cardspent +
##     card2spent + `agecat_50-64` + `agecat_35-49` + employ_8 +
##     employ_2 + employ_18 + employ_28 + employ_26 + employ_9 +
##     employ_20 + `carown_N/A` + `card_American Express` + cardtype_Other +
##     card2_Mastercard + `card2_American Express` + card2_Visa +
##     churn_No + voice_No
##
##                              Df Sum of Sq     RSS    AIC
## - employ_8                    1       2.0  5238.9 1455.7
## - employ_28                   1       2.0  5239.0 1455.8
## - employ_26                   1       2.3  5239.3 1456.0
## - employ_2                    1       2.8  5239.7 1456.3
## - ed                          1       2.8  5239.7 1456.3
## - employ_9                    1       3.0  5239.9 1456.4
## <none>                                     5237.0 1456.4
## - employ_18                   1       3.0  5239.9 1456.4
## - card2_Visa                  1       3.1  5240.1 1456.5
## - employ_20                   1       3.8  5240.8 1457.0
## + employ_11                   1       1.6  5235.3 1457.3
## + employ_7                    1       1.4  5235.6 1457.5
## + employ_10                   1       1.3  5235.6 1457.5
## + `agecat_>65`                1       1.1  5235.8 1457.7
## + employ_35                   1       0.9  5236.0 1457.8
## - cardtype_Other              1       5.1  5242.1 1457.8
## + `spousedcat_College degree` 1       0.8  5236.2 1457.9
## + pager_No                    1       0.6  5236.3 1458.0
## + employ_34                   1       0.6  5236.3 1458.0
## - churn_No                    1       5.6  5242.6 1458.2
## + card2_Discover              1       0.3  5236.6 1458.2
## - `carown_N/A`                1       5.7  5242.6 1458.2
## + employ_27                   1       0.3  5236.7 1458.2
## + employ_15                   1       0.2  5236.7 1458.3
```

```
## + employ_38                   1       0.2  5236.8 1458.3
## + employ_1                    1       0.1  5236.9 1458.4
## + cardtype_Platinum           1       0.1  5236.9 1458.4
## + employ_12                   1       0.0  5237.0 1458.4
## - voice_No                    1       6.8  5243.7 1458.9
## - card2_Mastercard            1       8.2  5245.1 1459.9
## - `agecat_50-64`              1      12.0  5248.9 1462.4
## - `agecat_35-49`              1      26.4  5263.3 1472.0
## - lncreddebt                  1      51.1  5288.0 1488.4
## - `card2_American Express`    1      67.9  5304.8 1499.5
## - `card_American Express`     1      95.9  5332.9 1517.9
## - carditems                   1    1054.9  6291.8 2096.7
## - card2spent                  1   16615.6 21852.6 6454.4
## - cardspent                   1   19272.8 24509.8 6856.1
##
## Step:  AIC=1455.74
## total_spent_sqrt ~ ed + lncreddebt + carditems + cardspent +
##     card2spent + `agecat_50-64` + `agecat_35-49` + employ_2 +
##     employ_18 + employ_28 + employ_26 + employ_9 + employ_20 +
##     `carown_N/A` + `card_American Express` + cardtype_Other +
##     card2_Mastercard + `card2_American Express` + card2_Visa +
##     churn_No + voice_No
##
##                              Df Sum of Sq     RSS    AIC
## - employ_28                   1       2.0  5240.9 1455.1
## - employ_26                   1       2.2  5241.2 1455.2
## - employ_2                    1       2.5  5241.4 1455.4
## - ed                          1       2.7  5241.6 1455.5
## - employ_9                    1       2.8  5241.7 1455.6
## - employ_18                   1       2.9  5241.8 1455.6
## <none>                                     5238.9 1455.7
## - card2_Visa                  1       3.0  5242.0 1455.8
## - employ_20                   1       3.7  5242.6 1456.2
## + employ_8                    1       2.0  5237.0 1456.4
## + employ_7                    1       1.5  5237.4 1456.7
## + employ_11                   1       1.4  5237.5 1456.8
## + `agecat_>65`                1       1.2  5237.7 1456.9
## + employ_10                   1       1.2  5237.7 1456.9
## - cardtype_Other              1       5.1  5244.0 1457.1
## + employ_35                   1       0.9  5238.1 1457.2
## + `spousedcat_College degree` 1       0.8  5238.1 1457.2
## + pager_No                    1       0.6  5238.3 1457.3
## + employ_34                   1       0.6  5238.4 1457.4
## - churn_No                    1       5.5  5244.4 1457.4
## - `carown_N/A`                1       5.6  5244.5 1457.5
## + card2_Discover              1       0.3  5238.6 1457.5
## + employ_27                   1       0.3  5238.7 1457.6
## + employ_38                   1       0.2  5238.7 1457.6
## + employ_15                   1       0.2  5238.8 1457.6
## + cardtype_Platinum           1       0.0  5238.9 1457.7
## + employ_1                    1       0.0  5238.9 1457.7
## + employ_12                   1       0.0  5238.9 1457.7
## - voice_No                    1       6.8  5245.8 1458.3
## - card2_Mastercard            1       8.1  5247.0 1459.1
```

```
## - `agecat_50-64`              1      11.6  5250.5 1461.5
## - `agecat_35-49`              1      25.4  5264.4 1470.7
## - lncreddebt                  1      50.6  5289.5 1487.4
## - `card2_American Express`    1      67.7  5306.6 1498.7
## - `card_American Express`     1      96.0  5334.9 1517.3
## - carditems                   1    1054.6  6293.5 2095.6
## - card2spent                  1   16614.1 21853.0 6452.5
## - cardspent                   1   19293.6 24532.5 6857.3
##
## Step:  AIC=1455.06
## total_spent_sqrt ~ ed + lncreddebt + carditems + cardspent +
##     card2spent + `agecat_50-64` + `agecat_35-49` + employ_2 +
##     employ_18 + employ_26 + employ_9 + employ_20 + `carown_N/A` +
##     `card_American Express` + cardtype_Other + card2_Mastercard +
##     `card2_American Express` + card2_Visa + churn_No + voice_No
##
##                             Df Sum of Sq     RSS    AIC
## - employ_26                  1       2.2  5243.1 1454.5
## - employ_2                   1       2.4  5243.4 1454.7
## - ed                         1       2.5  5243.4 1454.7
## - employ_9                   1       2.7  5243.6 1454.9
## - employ_18                  1       2.8  5243.7 1454.9
## - card2_Visa                 1       2.9  5243.9 1455.0
## <none>                                    5240.9 1455.1
## - employ_20                  1       3.6  5244.5 1455.5
## + employ_28                  1       2.0  5238.9 1455.7
## + employ_8                   1       1.9  5239.0 1455.8
## + employ_7                   1       1.6  5239.3 1456.0
## + `agecat_>65`               1       1.5  5239.5 1456.1
## + employ_11                  1       1.4  5239.5 1456.2
## + employ_10                  1       1.1  5239.8 1456.3
## + employ_35                  1       0.8  5240.1 1456.5
## - cardtype_Other             1       5.1  5246.1 1456.5
## + `spousedcat_College degree` 1      0.8  5240.1 1456.5
## + pager_No                   1       0.6  5240.3 1456.6
## + employ_34                  1       0.5  5240.4 1456.7
## - `carown_N/A`               1       5.5  5246.4 1456.8
## - churn_No                   1       5.6  5246.5 1456.8
## + card2_Discover             1       0.3  5240.6 1456.8
## + employ_27                  1       0.3  5240.7 1456.9
## + employ_38                  1       0.2  5240.7 1456.9
## + employ_15                  1       0.1  5240.8 1457.0
## + cardtype_Platinum          1       0.0  5240.9 1457.0
## + employ_1                   1       0.0  5240.9 1457.1
## + employ_12                  1       0.0  5240.9 1457.1
## - voice_No                   1       6.7  5247.6 1457.5
## - card2_Mastercard           1       7.9  5248.9 1458.4
## - `agecat_50-64`             1      11.8  5252.7 1460.9
## - `agecat_35-49`             1      26.1  5267.0 1470.4
## - lncreddebt                 1      50.1  5291.0 1486.3
## - `card2_American Express`   1      68.4  5309.3 1498.4
## - `card_American Express`    1      96.6  5337.5 1517.0
## - carditems                  1    1054.2  6295.1 2094.5
## - card2spent                 1   16612.1 21853.0 6450.5
```

```
## - cardspent                    1   19295.5 24536.4 6855.9
##
## Step:  AIC=1454.52
## total_spent_sqrt ~ ed + lncreddebt + carditems + cardspent +
##     card2spent + `agecat_50-64` + `agecat_35-49` + employ_2 +
##     employ_18 + employ_9 + employ_20 + `carown_N/A` + `card_American Express` +
##     cardtype_Other + card2_Mastercard + `card2_American Express` +
##     card2_Visa + churn_No + voice_No
##
##                               Df Sum of Sq     RSS    AIC
## - employ_2                     1       2.3  5245.4 1454.1
## - ed                           1       2.3  5245.4 1454.1
## - employ_9                     1       2.6  5245.7 1454.3
## - employ_18                    1       2.7  5245.8 1454.3
## - card2_Visa                   1       2.8  5245.9 1454.4
## <none>                                      5243.1 1454.5
## - employ_20                    1       3.5  5246.6 1454.9
## + employ_26                    1       2.2  5240.9 1455.1
## + employ_28                    1       1.9  5241.2 1455.2
## + employ_8                     1       1.8  5241.3 1455.3
## + `agecat_>65`                 1       1.8  5241.3 1455.3
## + employ_7                     1       1.6  5241.4 1455.4
## + employ_11                    1       1.3  5241.8 1455.7
## + employ_10                    1       1.1  5242.0 1455.8
## + `spousedcat_College degree`  1       0.8  5242.3 1456.0
## + employ_35                    1       0.8  5242.3 1456.0
## - cardtype_Other               1       5.3  5248.4 1456.1
## + pager_No                     1       0.7  5242.4 1456.1
## + employ_34                    1       0.5  5242.6 1456.2
## - `carown_N/A`                 1       5.6  5248.7 1456.2
## - churn_No                     1       5.6  5248.7 1456.2
## + card2_Discover               1       0.4  5242.7 1456.3
## + employ_27                    1       0.2  5242.9 1456.4
## + employ_38                    1       0.2  5242.9 1456.4
## + employ_15                    1       0.1  5243.0 1456.4
## + cardtype_Platinum            1       0.0  5243.1 1456.5
## + employ_12                    1       0.0  5243.1 1456.5
## + employ_1                     1       0.0  5243.1 1456.5
## - voice_No                     1       6.7  5249.8 1457.0
## - card2_Mastercard             1       8.0  5251.1 1457.9
## - `agecat_50-64`               1      11.6  5254.7 1460.2
## - `agecat_35-49`               1      26.9  5270.0 1470.5
## - lncreddebt                   1      49.4  5292.5 1485.3
## - `card2_American Express`     1      68.2  5311.3 1497.8
## - `card_American Express`      1      97.8  5340.9 1517.2
## - carditems                    1    1052.3  6295.4 2092.7
## - card2spent                   1   16625.3 21868.4 6451.0
## - cardspent                    1   19298.5 24541.6 6854.6
##
## Step:  AIC=1454.09
## total_spent_sqrt ~ ed + lncreddebt + carditems + cardspent +
##     card2spent + `agecat_50-64` + `agecat_35-49` + employ_18 +
##     employ_9 + employ_20 + `carown_N/A` + `card_American Express` +
##     cardtype_Other + card2_Mastercard + `card2_American Express` +
```

```
##     card2_Visa + churn_No + voice_No
##
##                             Df Sum of Sq     RSS    AIC
## - employ_9                   1       2.4  5247.8 1453.7
## - ed                         1       2.5  5248.0 1453.8
## - employ_18                  1       2.6  5248.0 1453.8
## - card2_Visa                 1       2.9  5248.3 1454.0
## <none>                                    5245.4 1454.1
## - employ_20                  1       3.3  5248.8 1454.3
## + employ_2                   1       2.3  5243.1 1454.5
## + employ_26                  1       2.1  5243.4 1454.7
## + employ_7                   1       1.9  5243.6 1454.8
## + employ_28                  1       1.9  5243.6 1454.8
## + employ_8                   1       1.6  5243.8 1455.0
## + `agecat_>65`               1       1.3  5244.1 1455.2
## + employ_11                  1       1.1  5244.3 1455.3
## + employ_10                  1       0.9  5244.5 1455.5
## + `spousedcat_College degree`  1     0.9  5244.6 1455.5
## + employ_35                  1       0.8  5244.7 1455.6
## + pager_No                   1       0.7  5244.8 1455.6
## - cardtype_Other             1       5.3  5250.8 1455.7
## - churn_No                   1       5.4  5250.8 1455.7
## - `carown_N/A`               1       5.5  5250.9 1455.8
## + employ_34                  1       0.5  5245.0 1455.8
## + card2_Discover             1       0.3  5245.1 1455.9
## + employ_38                  1       0.2  5245.2 1455.9
## + employ_27                  1       0.2  5245.2 1456.0
## + employ_15                  1       0.1  5245.4 1456.0
## + cardtype_Platinum          1       0.0  5245.4 1456.1
## + employ_12                  1       0.0  5245.4 1456.1
## + employ_1                   1       0.0  5245.4 1456.1
## - voice_No                   1       6.8  5252.3 1456.7
## - card2_Mastercard           1       8.0  5253.4 1457.4
## - `agecat_50-64`             1      12.1  5257.5 1460.1
## - `agecat_35-49`             1      26.5  5272.0 1469.7
## - lncreddebt                 1      50.0  5295.4 1485.3
## - `card2_American Express`   1      68.7  5314.2 1497.6
## - `card_American Express`    1      98.1  5343.6 1517.0
## - carditems                  1    1051.6  6297.1 2091.6
## - card2spent                 1   16623.0 21868.5 6449.0
## - cardspent                  1   19300.1 24545.6 6853.2
##
## Step:  AIC=1453.67
## total_spent_sqrt ~ ed + lncreddebt + carditems + cardspent +
##     card2spent + `agecat_50-64` + `agecat_35-49` + employ_18 +
##     employ_20 + `carown_N/A` + `card_American Express` + cardtype_Other +
##     card2_Mastercard + `card2_American Express` + card2_Visa +
##     churn_No + voice_No
##
##                             Df Sum of Sq     RSS    AIC
## - employ_18                  1       2.4  5250.3 1453.3
## - ed                         1       2.5  5250.3 1453.3
## - card2_Visa                 1       3.0  5250.8 1453.7
## <none>                                    5247.8 1453.7
```

```
## - employ_20                    1       3.2  5251.0 1453.8
## + employ_9                     1       2.4  5245.4 1454.1
## + employ_2                     1       2.1  5245.7 1454.3
## + employ_7                     1       2.1  5245.7 1454.3
## + employ_26                    1       2.0  5245.8 1454.3
## + employ_28                    1       1.8  5246.0 1454.5
## + `agecat_>65`                 1       1.5  5246.3 1454.7
## + employ_8                     1       1.4  5246.4 1454.7
## + employ_11                    1       1.0  5246.9 1455.0
## + `spousedcat_College degree`  1       0.9  5247.0 1455.1
## + employ_10                    1       0.8  5247.0 1455.1
## + employ_35                    1       0.8  5247.1 1455.2
## + pager_No                     1       0.7  5247.2 1455.2
## - cardtype_Other               1       5.3  5253.2 1455.2
## - churn_No                     1       5.5  5253.3 1455.3
## - `carown_N/A`                 1       5.5  5253.4 1455.4
## + employ_34                    1       0.4  5247.4 1455.4
## + card2_Discover               1       0.4  5247.4 1455.4
## + employ_38                    1       0.3  5247.6 1455.5
## + employ_27                    1       0.2  5247.6 1455.6
## + employ_15                    1       0.1  5247.8 1455.6
## + employ_12                    1       0.0  5247.8 1455.6
## + cardtype_Platinum            1       0.0  5247.8 1455.6
## + employ_1                     1       0.0  5247.8 1455.7
## - voice_No                     1       6.8  5254.6 1456.2
## - card2_Mastercard             1       8.1  5255.9 1457.0
## - `agecat_50-64`               1      11.4  5259.2 1459.2
## - `agecat_35-49`               1      25.2  5273.0 1468.4
## - lncreddebt                   1      50.2  5298.0 1485.0
## - `card2_American Express`     1      68.0  5315.9 1496.8
## - `card_American Express`      1      98.6  5346.4 1516.8
## - carditems                    1    1052.2  6300.0 2091.3
## - card2spent                   1   16653.4 21901.3 6452.2
## - cardspent                    1   19298.7 24546.5 6851.3
##
## Step:  AIC=1453.3
## total_spent_sqrt ~ ed + lncreddebt + carditems + cardspent +
##     card2spent + `agecat_50-64` + `agecat_35-49` + employ_20 +
##     `carown_N/A` + `card_American Express` + cardtype_Other +
##     card2_Mastercard + `card2_American Express` + card2_Visa +
##     churn_No + voice_No
##
##                               Df Sum of Sq     RSS    AIC
## - ed                           1       2.4  5252.6 1452.9
## <none>                                      5250.3 1453.3
## - card2_Visa                   1       3.0  5253.3 1453.3
## - employ_20                    1       3.1  5253.3 1453.3
## + employ_18                    1       2.4  5247.8 1453.7
## + employ_9                     1       2.2  5248.0 1453.8
## + employ_7                     1       2.2  5248.1 1453.8
## + employ_2                     1       2.0  5248.3 1454.0
## + employ_26                    1       1.9  5248.3 1454.0
## + `agecat_>65`                 1       1.8  5248.4 1454.1
## + employ_28                    1       1.8  5248.5 1454.1
```

```
## + employ_8                    1       1.3  5248.9 1454.4
## + employ_11                   1       0.9  5249.4 1454.7
## + `spousedcat_College degree`  1       0.8  5249.4 1454.7
## + employ_35                   1       0.7  5249.5 1454.8
## + employ_10                   1       0.7  5249.5 1454.8
## + pager_No                    1       0.7  5249.6 1454.8
## - `carown_N/A`                1       5.4  5255.7 1454.9
## - cardtype_Other              1       5.4  5255.7 1454.9
## + card2_Discover              1       0.4  5249.8 1455.0
## + employ_34                   1       0.4  5249.8 1455.0
## - churn_No                    1       5.7  5256.0 1455.1
## + employ_38                   1       0.3  5250.0 1455.1
## + employ_27                   1       0.2  5250.1 1455.2
## + employ_12                   1       0.1  5250.2 1455.3
## + cardtype_Platinum           1       0.0  5250.2 1455.3
## + employ_1                    1       0.0  5250.2 1455.3
## + employ_15                   1       0.0  5250.2 1455.3
## - voice_No                    1       6.8  5257.1 1455.8
## - card2_Mastercard            1       7.8  5258.1 1456.5
## - `agecat_50-64`              1      10.9  5261.2 1458.6
## - `agecat_35-49`              1      25.2  5275.5 1468.1
## - lncreddebt                  1      50.0  5300.2 1484.4
## - `card2_American Express`    1      68.2  5318.5 1496.5
## - `card_American Express`     1      98.2  5348.5 1516.2
## - carditems                   1    1053.0  6303.2 2091.0
## - card2spent                  1   16660.2 21910.4 6451.7
## - cardspent                   1   19296.3 24546.6 6849.3
##
## Step:  AIC=1452.87
## total_spent_sqrt ~ lncreddebt + carditems + cardspent + card2spent +
##     `agecat_50-64` + `agecat_35-49` + employ_20 + `carown_N/A` +
##     `card_American Express` + cardtype_Other + card2_Mastercard +
##     `card2_American Express` + card2_Visa + churn_No + voice_No
##
##                              Df Sum of Sq     RSS    AIC
## <none>                                     5252.6 1452.9
## - employ_20                   1       3.0  5255.6 1452.9
## - card2_Visa                  1       3.1  5255.7 1452.9
## + ed                          1       2.4  5250.3 1453.3
## + employ_18                   1       2.3  5250.3 1453.3
## + employ_7                    1       2.2  5250.4 1453.4
## + employ_9                    1       2.2  5250.4 1453.4
## + employ_2                    1       2.2  5250.5 1453.4
## + employ_26                   1       1.8  5250.8 1453.6
## + employ_28                   1       1.6  5251.0 1453.8
## + `agecat_>65`                1       1.6  5251.0 1453.8
## - churn_No                    1       4.6  5257.2 1453.9
## + employ_8                    1       1.2  5251.4 1454.1
## + pager_No                    1       1.1  5251.5 1454.2
## + employ_11                   1       0.9  5251.7 1454.3
## - voice_No                    1       5.1  5257.7 1454.3
## + card2_Discover              1       0.7  5251.9 1454.4
## + employ_35                   1       0.6  5252.0 1454.4
## + employ_10                   1       0.6  5252.0 1454.4
```

```
## - `carown_N/A`                 1       5.4  5258.0 1454.5
## + `spousedcat_College degree`  1       0.3  5252.3 1454.6
## + employ_34                    1       0.3  5252.3 1454.6
## - cardtype_Other               1       5.7  5258.3 1454.7
## + employ_38                    1       0.3  5252.3 1454.7
## + employ_27                    1       0.2  5252.5 1454.8
## + employ_12                    1       0.1  5252.5 1454.8
## + cardtype_Platinum            1       0.1  5252.6 1454.8
## + employ_15                    1       0.0  5252.6 1454.9
## + employ_1                     1       0.0  5252.6 1454.9
## - card2_Mastercard             1       8.0  5260.6 1456.2
## - `agecat_50-64`               1      11.1  5263.7 1458.3
## - `agecat_35-49`               1      24.7  5277.3 1467.3
## - lncreddebt                   1      48.4  5301.0 1483.0
## - `card2_American Express`     1      66.2  5318.9 1494.7
## - `card_American Express`      1      96.0  5348.7 1514.3
## - carditems                    1    1053.0  6305.7 2090.4
## - card2spent                   1   16685.7 21938.3 6454.2
## - cardspent                    1   19297.1 24549.7 6847.8
```

```
##
## Call:
## lm(formula = total_spent_sqrt ~ lncreddebt + carditems + cardspent +
##     card2spent + `agecat_50-64` + `agecat_35-49` + employ_20 +
##     `carown_N/A` + `card_American Express` + cardtype_Other +
##     card2_Mastercard + `card2_American Express` + card2_Visa +
##     churn_No + voice_No, data = Dev)
##
## Coefficients:
##             (Intercept)               lncreddebt
##                 9.61480                  0.10063
##               carditems                cardspent
##                 0.20976                  0.01760
##              card2spent           `agecat_50-64`
##                 0.02188                  0.14611
##          `agecat_35-49`                employ_20
##                 0.20964                 -0.22938
##            `carown_N/A`   `card_American Express`
##                -0.13242                  0.48074
##          cardtype_Other         card2_Mastercard
##                 0.09432                 -0.12687
## `card2_American Express`               card2_Visa
##                 0.45245                 -0.08044
##                churn_No                 voice_No
##                -0.08618                 -0.08570
```

```r
# Model given after stepwise regression
fit3 <- lm(formula = total_spent_sqrt ~ ed + lncreddebt + carditems +
              cardspent + card2spent + `agecat_50-64` + `agecat_35-49` +
              employ_27 + employ_2 + employ_18 + employ_20 + employ_35 +
              employ_38 + `spousedcat_College degree` + `card_American Express` +
              cardtype_Other + card2_Discover + card2_Mastercard + `card2_American Express` +
              card2_Visa + churn_No + voice_No, data = Dev)

summary(fit3)
```

```
##
## Call:
## lm(formula = total_spent_sqrt ~ ed + lncreddebt + carditems +
##     cardspent + card2spent + `agecat_50-64` + `agecat_35-49` +
##     employ_27 + employ_2 + employ_18 + employ_20 + employ_35 +
##     employ_38 + `spousedcat_College degree` + `card_American Express` +
##     cardtype_Other + card2_Discover + card2_Mastercard + `card2_American Express` +
##     card2_Visa + churn_No + voice_No, data = Dev)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -7.1074 -0.5386  0.2799  0.8329  2.3786
##
## Coefficients:
##                               Estimate Std. Error t value
## (Intercept)                  9.8121837  0.1608975  60.984
## ed                          -0.0096017  0.0073155  -1.313
## lncreddebt                   0.1036246  0.0178974   5.790
## carditems                    0.2091833  0.0079377  26.353
## cardspent                    0.0176037  0.0001557 113.046
## card2spent                   0.0218613  0.0002082 104.979
## `agecat_50-64`               0.1436550  0.0540743   2.657
## `agecat_35-49`               0.2122901  0.0520915   4.075
## employ_27                   -0.0918786  0.2342787  -0.392
## employ_2                    -0.1005082  0.0847328  -1.186
## employ_18                   -0.2148387  0.1692938  -1.269
## employ_20                   -0.2303926  0.1623446  -1.419
## employ_35                   -0.2633954  0.3725225  -0.707
## employ_38                    0.1232113  0.3302258   0.373
## `spousedcat_College degree`  0.0599422  0.0791880   0.757
## `card_American Express`      0.4936784  0.0618364   7.984
## cardtype_Other               0.0891504  0.0485091   1.838
## card2_Discover              -0.0542003  0.0924477  -0.586
## card2_Mastercard            -0.1728031  0.0882132  -1.959
## `card2_American Express`     0.4172470  0.0964420   4.326
## card2_Visa                  -0.1214151  0.0890886  -1.363
## churn_No                    -0.1000348  0.0501904  -1.993
## voice_No                    -0.0966060  0.0486506  -1.986
##                                         Pr(>|t|)
## (Intercept)                 < 0.0000000000000002 ***
## ed                                       0.18943
## lncreddebt                   0.00000000766740541 ***
## carditems                   < 0.0000000000000002 ***
## cardspent                   < 0.0000000000000002 ***
## card2spent                  < 0.0000000000000002 ***
## `agecat_50-64`                           0.00793 **
## `agecat_35-49`               0.00004697653193402 ***
## employ_27                                0.69495
## employ_2                                 0.23563
## employ_18                                0.20452
## employ_20                                0.15594
## employ_35                                0.47958
## employ_38                                0.70909
## `spousedcat_College degree`              0.44912
```

```
## `card_American Express`      0.00000000000000191 ***
## cardtype_Other                           0.06618 .
## card2_Discover                           0.55772
## card2_Mastercard                         0.05020 .
## `card2_American Express`     0.00001558252915094 ***
## card2_Visa                               0.17302
## churn_No                                 0.04633 *
## voice_No                                 0.04714 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.229 on 3477 degrees of freedom
## Multiple R-squared:  0.9665, Adjusted R-squared:  0.9662
## F-statistic:  4554 on 22 and 3477 DF,  p-value: < 0.00000000000000022
```

This model is further perfected

```r
fit4 <- lm(formula = total_spent_sqrt ~ ed + lncreddebt + carditems +
             cardspent + card2spent + `agecat_50-64` + `agecat_35-49` +
             employ_27 + employ_20 + employ_35 +
             employ_38 + `spousedcat_College degree` + `card_American Express` +
             cardtype_Other + card2_Mastercard + `card2_American Express` +
             card2_Visa + churn_No + voice_No, data = Dev)

summary(fit4)
```

```
##
## Call:
## lm(formula = total_spent_sqrt ~ ed + lncreddebt + carditems +
##     cardspent + card2spent + `agecat_50-64` + `agecat_35-49` +
##     employ_27 + employ_20 + employ_35 + employ_38 + `spousedcat_College degree` +
##     `card_American Express` + cardtype_Other + card2_Mastercard +
##     `card2_American Express` + card2_Visa + churn_No + voice_No,
##     data = Dev)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -7.0997 -0.5428  0.2811  0.8330  2.3815
##
## Coefficients:
##                              Estimate Std. Error t value
## (Intercept)                 9.7730615  0.1513974  64.552
## ed                         -0.0102060  0.0072601  -1.406
## lncreddebt                  0.1035349  0.0178796   5.791
## carditems                   0.2091921  0.0079376  26.355
## cardspent                   0.0176020  0.0001557 113.041
## card2spent                  0.0218646  0.0002082 105.034
## `agecat_50-64`              0.1432748  0.0539414   2.656
## `agecat_35-49`              0.2107877  0.0520674   4.048
## employ_27                  -0.0795043  0.2341763  -0.340
## employ_20                  -0.2195111  0.1622356  -1.353
## employ_35                  -0.2518119  0.3724794  -0.676
## employ_38                   0.1313887  0.3301675   0.398
## `spousedcat_College degree`  0.0601314  0.0791876   0.759
## `card_American Express`      0.4868521  0.0607292   8.017
```

```
## cardtype_Other              0.0896234  0.0484950   1.848
## card2_Mastercard           -0.1302023  0.0551274  -2.362
## `card2_American Express`    0.4591346  0.0687865   6.675
## card2_Visa                 -0.0814766  0.0560996  -1.452
## churn_No                   -0.1013564  0.0501444  -2.021
## voice_No                   -0.0981436  0.0486394  -2.018
##                                        Pr(>|t|)
## (Intercept)                < 0.0000000000000002 ***
## ed                                      0.15988
## lncreddebt                  0.00000000763240270 ***
## carditems                  < 0.0000000000000002 ***
## cardspent                  < 0.0000000000000002 ***
## card2spent                 < 0.0000000000000002 ***
## `agecat_50-64`                          0.00794 **
## `agecat_35-49`              0.00005269935624220 ***
## employ_27                               0.73425
## employ_20                               0.17613
## employ_35                               0.49906
## employ_38                               0.69069
## `spousedcat_College degree`             0.44769
## `card_American Express`     0.00000000000000147 ***
## cardtype_Other                          0.06467 .
## card2_Mastercard                        0.01824 *
## `card2_American Express`    0.00000000002870199 ***
## card2_Visa                              0.14649
## churn_No                                0.04333 *
## voice_No                                0.04369 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.229 on 3480 degrees of freedom
## Multiple R-squared:  0.9664, Adjusted R-squared:  0.9662
## F-statistic:  5272 on 19 and 3480 DF,  p-value: < 0.00000000000000022
```

The VIF is being studied for this model

```
library(car)
```

```
## Loading required package: carData
```

```
##
## Attaching package: 'car'
```

```
## The following object is masked from 'package:DescTools':
##
##     Recode
```

```
## The following object is masked from 'package:dplyr':
##
##     recode
```

```
(VIF <- vif(fit4))
```

```
##                          ed                  lncreddebt
##                    1.269992                    1.142734
##                   carditems                   cardspent
```

```
##                    1.632949                   2.750462
##                   card2spent              `agecat_50-64`
##                    1.798863                   1.211168
##              `agecat_35-49`                  employ_27
##                    1.165784                   1.008946
##                   employ_20                  employ_35
##                    1.011285                   1.007726
##                   employ_38 `spousedcat_College degree`
##                    1.006858                   1.062990
##     `card_American Express`              cardtype_Other
##                    1.356257                   1.003342
##            card2_Mastercard    `card2_American Express`
##                    1.425677                   1.470036
##                  card2_Visa                   churn_No
##                    1.405798                   1.096339
##                    voice_No
##                    1.148020
```

```
# VIF of all variables into consideration are below 5
(tolerance <- 1/VIF)
```

```
##                          ed                 lncreddebt
##                   0.7874064                  0.8750942
##                   carditems                  cardspent
##                   0.6123891                  0.3635753
##                  card2spent              `agecat_50-64`
##                   0.5559068                  0.8256494
##              `agecat_35-49`                  employ_27
##                   0.8577921                  0.9911336
##                   employ_20                  employ_35
##                   0.9888412                  0.9923334
##                   employ_38 `spousedcat_College degree`
##                   0.9931886                  0.9407423
##     `card_American Express`              cardtype_Other
##                   0.7373231                  0.9966689
##            card2_Mastercard    `card2_American Express`
##                   0.7014211                  0.6802556
##                  card2_Visa                   churn_No
##                   0.7113399                  0.9121268
##                    voice_No
##                   0.8710650
```

```
# The tolerance level for all the levels are above the required 0.2
```

The model is put to test for the accuracy on the model and the testing data

```
dev1 <- data.frame(cbind(Dev, pred_sq = predict(fit4, newdata=Dev)))
dev1$pred <- (dev1$pred_sq)^2
val1 <- data.frame(cbind(Test, pred_sq = predict(fit4, newdata=Test)))
val1$pred <- (val1$pred_sq)^2
```

Now to verify all the measures to see model accuracy

MAPE

```
DescTools::MAPE(x=dev1$pred,dev1$total_spent)
```

```
## [1] 0.1105021
```

```
DescTools::MAPE(x=val1$pred,val1$total_spent)
```

```
## [1] 0.1090822
```

The MAPE value for both are same and are well under the required levels

RMSE

```
DescTools::RMSE(x=dev1$pred,dev1$total_spent)
```

```
## [1] 66.253
```

```
DescTools::RMSE(x=val1$pred,val1$total_spent)
```

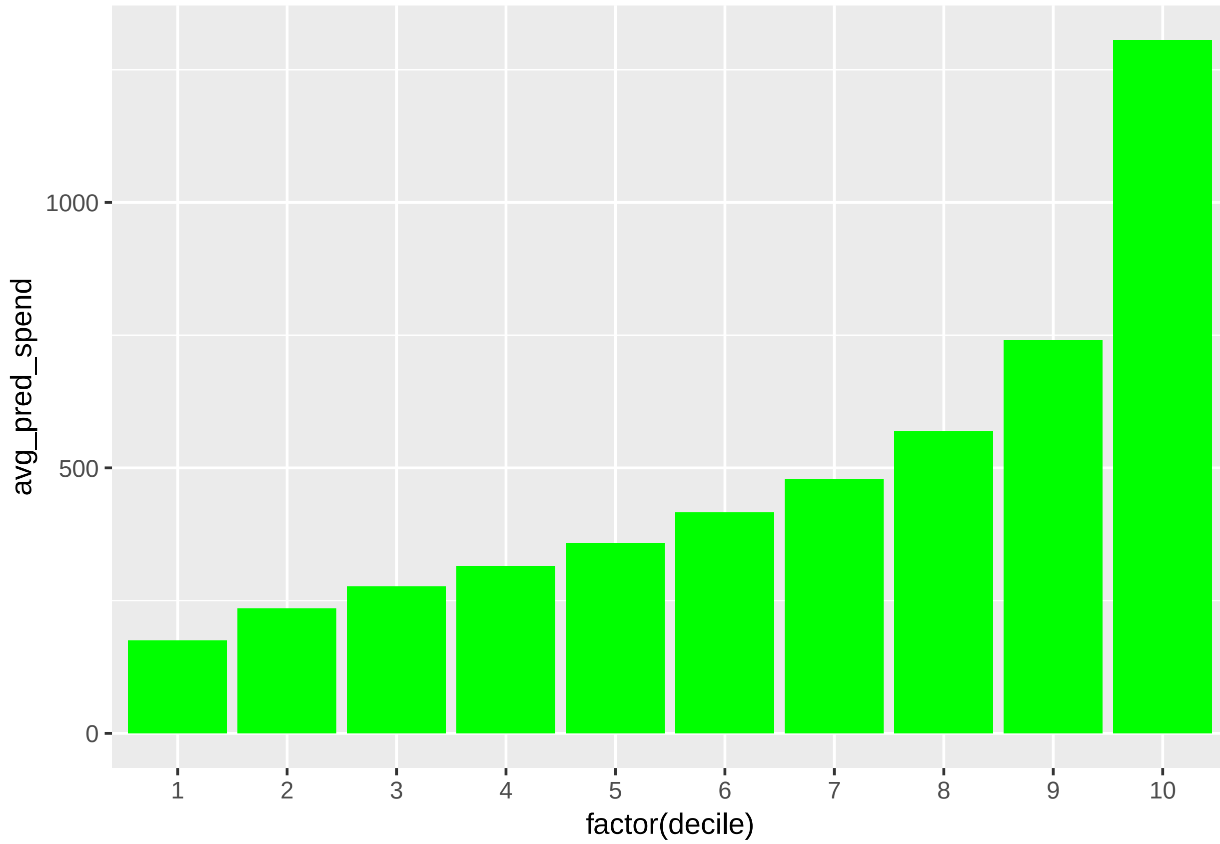
```
## [1] 62.54341
```

Given the range of the total spent, RMSE is low and hence the model is a good fit

Decile analysis

creating the deciles

```
decLocations <- quantile(val1$pred, probs = seq(0.1,0.9,by=0.1))
val1$decile <- findInterval(val1$pred,c(-Inf,decLocations, Inf))

val_decile <- val1 %>% group_by(decile) %>%
  dplyr::summarise(count = n(),
                   avg_pred_spend = mean(pred),
                   avg_spend = mean(total_spent)
  ) %>%
  dplyr::arrange(desc(decile))

require(ggplot2)
Decile_Plot <- ggplot(val_decile)+aes(x = factor(decile),y = avg_pred_spend) +
  geom_bar(stat = "identity",fill = "green")
plot(Decile_Plot)
```



The decile plot reveals that there is a staircase effect created by the model. Thus, making it a reliable model to use.

Of the factors that influence the balanec in a positive way, the top 5 factors are having an American Express card as the primary and secondary card, being in the age category of 35-49 or 50-64 and number of items on primary card. The factors that have a negative impact on the balance value is being with an employer for 35 or 20 years. The negative predictors can be made better.