

# Capstone Project - Factors that Affect the Housing Ownership

## Instructions

This document is to show how the data processes and transforms, also

how to apply the method and then analysis the data for the project

## Factors that Affect the Housing Ownership

### load required R libraries

```
{ r echo=TRUE} install.packages("knitr") library(knitr) install.packages("rmarkdown") library(rmarkdown) install.packages("sas7bdat") library(sas7bdat) install.packages("dplyr") library(dplyr)
```

### LOAD data from the HADS website

Check for attribute, Correlation using pearson for numeric variable or spearman for character variable

and remove the attributes with more than lots of missing data

check attrible of the dataset and check for missing variable

since control number is primary variable, there is no missing within

```
temp <- tempfile()
download.file("https://www.huduser.gov/portal/datasets/hads/hads2013n_ASCII.zip",temp)
hads2013n <- read.csv(unz(temp, "thads2013n.txt"), header = FALSE, skip = 1)
hads2013n.names <- readLines(unz(temp, "thads2013n.txt"), n = 1)
names(hads2013n) <- unlist(strsplit(hads2013n.names, ","))
str(hads2013n)
```

```
## 'data.frame': 64535 obs. of 99 variables:
## $ CONTROL : Factor w/ 64535 levels "'100003130103'",...: 1 2 3 4 5 6 7 8 9 10 ...
## $ AGE1 : int 82 50 53 67 26 56 50 26 60 26 ...
## $ METRO3 : Factor w/ 5 levels "'1'", "'2'", "'3'",...: 3 5 5 5 1 2 1 4 5 4 ...
## $ REGION : Factor w/ 4 levels "'1'", "'2'", "'3'",...: 1 3 3 3 3 3 3 4 4 2 ...
## $ LMED : int 73738 55846 55846 55846 60991 62066 60991 52322 50296 63221 ...
## $ FMR : int 956 1100 1100 949 737 657 988 773 1125 552 ...
## $ L30 : int 15738 17165 13750 13750 14801 13170 16646 13489 13115 13338 ...
## $ L50 : int 26213 28604 22897 22897 24628 21924 27713 22471 21859 22199 ...
## $ L80 : int 40322 45744 36614 36614 39421 35073 44340 35929 34939 35501 ...
## $ IPOV : int 11067 24218 15470 13964 15492 12005 18050 15992 15452 12005 ...
## $ BEDRMS : int 2 4 4 3 2 1 3 2 3 1 ...
## $ BUILT : int 2006 1980 1985 1985 1980 1985 1985 1980 1985 1985 ...
## $ STATUS : Factor w/ 2 levels "'1'", "'3'",...: 1 1 1 1 1 1 1 1 1 1 ...
## $ TYPE : int 1 1 1 1 1 1 1 1 1 1 ...
## $ VALUE : int 40000 130000 150000 200000 -6 -6 260000 -6 170000 -6 ...
## $ VACANCY : int -6 -6 -6 -6 -6 -6 -6 -6 -6 ...
## $ TENURE : Factor w/ 4 levels "'-6'", "'1'", "'2'",...: 2 2 2 2 3 3 2 3 2 3 ...
## $ NUNITS : int 1 1 1 1 100 32 1 8 1 24 ...
## $ ROOMS : int 6 6 7 6 4 3 6 5 7 3 ...
## $ WEIGHT : num 3117 2151 2214 2365 2315 ...
## $ PER : int 1 4 2 2 2 1 3 2 2 1 ...
## $ ZINC2 : int 18021 122961 27974 32220 96874 14987 69962 32000 118987 47987 ...
## $ 7ADFO : Factor w/ 4 levels "'-6'", "'1'", "'2'",...: 2 2 2 2 2 2 2 2 2 2 ...
```

```

## $ ZSMHC : int 533 487 1405 279 759 695 1165 976 1156 1100 ...
## $ STRUCTURETYPE : int 1 1 1 1 5 4 1 3 1 4 ...
## $ OWNRENT : Factor w/ 2 levels "1","2": 1 1 1 1 2 2 1 2 1 2 ...
## $ UTILITY : num 169 245 159 179 146 ...
## $ OTHERCOST : num 213.8 58.3 37.5 70.7 12.5 ...
## $ COST06 : num 649 1168 1193 1579 759 ...
## $ COST12 : num 803 1670 1773 2351 759 ...
## $ COST08 : num 697 1325 1375 1820 759 ...
## $ COSTMED : num 615 1059 1068 1412 759 ...
## $ TOTSAL : int 0 123000 28000 0 96900 15000 70001 20000 107000 48000 ...
## $ ASSISTED : int -9 -9 -9 -9 0 1 -9 0 -9 0 ...
## $ GLMED : int 73738 55846 55846 55846 60991 62066 60991 52322 50296 63221 ...
## $ GL30 : int 15738 17165 13750 13750 14801 13170 16646 13489 13115 13338 ...
## $ GL50 : int 26213 28604 22897 22897 24628 21924 27713 22471 21859 22199 ...
## $ GL80 : int 40322 45744 36614 36614 39421 35073 44340 35929 34939 35501 ...
## $ APLMED : num 51617 55846 44677 44677 48793 ...
## $ ABL30 : num 20235 19911 19938 17875 16651 ...
## $ ABL50 : num 33702 33181 33201 29766 27707 ...
## $ ABL80 : num 51843 53063 53090 47598 44349 ...
## $ ABLMED : num 66364 64781 64781 58080 54892 ...
## $ BURDEN : num 0.3549 0.0475 0.6027 0.1039 0.094 ...
## $ INCRELAMIPCT : num 34.9 220.2 62.6 72.1 198.5 ...
## $ INCRELAMICAT : int 2 7 4 4 7 2 7 4 7 6 ...
## $ INCRELPOVPCT : num 163 508 181 231 625 ...
## $ INCRELPOVCAT : int 3 4 3 4 4 2 4 4 4 4 ...
## $ INCRELFMRPCT : num 47.1 279.5 63.6 84.9 328.6 ...
## $ INCRELFMRCAT : int 1 3 2 2 3 2 3 3 3 3 ...
## $ COST06RELAMIPCT : num 39.1 72.1 73.7 108.7 55.3 ...
## $ COST06RELAMICAT : int 2 4 4 6 3 3 7 4 6 5 ...
## $ COST06RELPOVPCT : num 234 193 309 452 196 ...
## $ COST06RELPOVCAT : int 4 3 4 4 3 4 4 4 4 4 ...
## $ COST06RELFMRPCT : num 67.8 106.1 108.5 166.4 103 ...
## $ COST06RELFMRCAT : int 2 3 3 3 3 3 3 3 3 3 ...
## $ COST08RELAMIPCT : num 42 81.8 84.9 125.4 55.3 ...
## $ COST08RELAMICAT : int 2 4 5 7 3 3 7 4 6 5 ...
## $ COST08RELPOVPCT : num 252 219 355 521 196 ...
## $ COST08RELPOVCAT : int 4 4 4 4 3 4 4 4 4 4 ...
## $ COST08RELFMRPCT : num 72.9 120.4 125 191.8 103 ...
## $ COST08RELFMRCAT : int 2 3 3 3 3 3 3 3 3 3 ...
## $ COST12RELAMIPCT : num 48.4 103.1 109.5 161.9 55.3 ...
## $ COST12RELAMICAT : int 2 6 6 7 3 3 7 4 7 5 ...
## $ COST12RELPOVPCT : num 290 276 458 673 196 ...
## $ COST12RELPOVCAT : int 4 4 4 4 3 4 4 4 4 4 ...
## $ COST12RELFMRPCT : num 84 152 161 248 103 ...
## $ COST12RELFMRCAT : int 2 3 3 3 3 3 3 3 3 3 ...
## $ COSTMedRELAMIPCT : num 37.1 65.4 65.9 97.2 55.3 ...
## $ COSTMedRELAMICAT : int 2 4 4 5 3 3 6 4 5 5 ...
## $ COSTMedRELPOVPCT : num 222 175 276 404 196 ...
## $ COSTMedRELPOVCAT : int 4 3 4 4 3 4 4 4 4 4 ...
## $ COSTMedRELFMRPCT : num 64.3 96.3 97.1 148.8 103 ...
## $ COSTMedRELFMRCAT : int 2 2 2 3 3 3 3 3 3 3 ...
## $ FMTZADEQ : Factor w/ 4 levels "1","2","3","4": 2 2 2 2 2 2 2 2 2 2 ...
## $ FMTMETRO3 : Factor w/ 2 levels "1","2": 1 1 1 1 2 1 2 1 1 1 ...
## $ FMTBULT : Factor w/ 7 levels "1","2","3","4","5","6","7": 6 4 4 4 4 4 4 4 4 4 ...
## $ FMTSTRUCTURETYPE : Factor w/ 7 levels "1","2","3","4","5","6","7": 2 2 2 2 6 5 2 4 2 5 ...
## $ FMTBEDRMS : Factor w/ 5 levels "1","2","3","4","5": 3 5 5 4 3 2 4 3 4 2 ...

## $ FMTOWNRENT : Factor w/ 2 levels "1 Owner","2 Renter": 1 1 1 1 2 2 1 2 1 2 ...
## $ FMTCOST06RELPOVCAT : Factor w/ 5 levels "1","2","3","4","5": 5 4 5 5 4 5 5 5 5 5 ...
## $ FMTCOST08RELPOVCAT : Factor w/ 5 levels "1","2","3","4","5": 5 5 5 5 4 5 5 5 5 5 ...
## $ FMTCOST12RELPOVCAT : Factor w/ 5 levels "1","2","3","4","5": 5 5 5 5 4 5 5 5 5 5 ...
## $ FMTCOSTMEDRELPOVCAT : Factor w/ 5 levels "1","2","3","4","5": 5 4 5 5 4 5 5 5 5 5 ...
## $ FMTINCRELPOVCAT : Factor w/ 5 levels "1","2","3","4","5": 4 5 4 5 3 5 5 5 5 5 ...
## $ FMTCOST06RELFMRCAT : Factor w/ 3 levels "1","2","3": 2 3 3 3 3 3 3 3 3 3 ...
## $ FMTCOST08RELFMRCAT : Factor w/ 3 levels "1","2","3": 2 3 3 3 3 3 3 3 3 3 ...
## $ FMTCOST12RELFMRCAT : Factor w/ 3 levels "1","2","3": 2 3 3 3 3 3 3 3 3 3 ...
## $ FMTCOSTMEDRELFMRCAT : Factor w/ 3 levels "1","2","3": 2 2 2 3 3 3 3 3 3 3 ...
## $ FMTINCRELFMRCAT : Factor w/ 4 levels "1","2","3","4": 2 4 3 3 4 3 4 4 4 4 ...
## $ FMTCOST06RELAMICAT : Factor w/ 7 levels "1","2","3","4","5","6","7": 2 4 4 6 3 3 7 4 6 5 ...
## $ FMTCOST08RELAMICAT : Factor w/ 7 levels "1","2","3","4","5","6","7": 2 4 5 7 3 3 7 4 6 5 ...
## $ FMTCOST12RELAMICAT : Factor w/ 7 levels "1","2","3","4","5","6","7": 2 6 6 7 3 3 7 4 7 5 ...
## $ FMTCOSTMEDRELAMICAT : Factor w/ 7 levels "1","2","3","4","5","6","7": 2 4 4 5 3 3 6 4 5 5 ...
## $ FMTINCRELAMICAT : Factor w/ 8 levels "1","2","3","4","5","6","7","8": 3 8 5 5 8 3 8 5 8 7 ...
## $ FMTASSISTED : Factor w/ 3 levels "1","2","3": 1 1 1 1 2 3 1 2 1 2 ...
## $ FMTBURDEN : Factor w/ 5 levels "1","2","3","4","5": 3 2 4 2 2 4 2 3 2 2 ...
## $ FMTREGION : Factor w/ 2 levels "1","2": 1 1 1 1 1 1 1 2 2 1 ...
## $ FMTSTATUS : Factor w/ 1 level "1": 1 1 1 1 1 1 1 1 1 1 ...

```

apply change for missing and error records and Duplicated Column

notice by reading data description variable like BEDRMS and FMTBEDRMS are same but in different format

and by comparing the unformatted and formatted the column, it shows formatted columns are more reasonable

to be used.

```
## check FMTOWNRENT and TENURE : 2 option, either Owner or Renter
str(hads2013n$FMTOWNRENT)
```

```
## Factor w/ 2 levels "1 Owner","2 Renter": 1 1 1 1 2 2 1 2 1 2 ...
```

```
str(hads2013n$TENURE)
```

```
## Factor w/ 4 levels "1","2","3","4": 2 2 2 3 3 3 2 3 2 3 ...
```

```
table(hads2013n$FMTOWNRENT)

##
##  '1 Owner' '2 Renter'
##      37146      27389

table(hads2013n$TENURE)

##
##  '-6'  '1'  '2'  '3'
##  4438 35852 23358  887

## Hence TENURE contains some error '-6'/missing rows. we will use FMTOWNRENT and Drop TENURE
hads2013n_c <- subset(hads2013n,select = -c(TENURE))

## Warning: closing unused connection 5 (C:\Users\THOMAS~1\AppData\Local\Temp
## \RtmpQDUVUZ\file1b7027c973ec:thads2013n.txt)

## check BEDRMS and FMTBEDRMS
str(hads2013n$BEDRMS)

##  int [1:64535] 2 4 4 3 2 1 3 2 3 1 ...

str(hads2013n$FMTBEDRMS)

##  Factor w/ 5 levels "'0 Studio'", "'1 1BR'",...: 3 5 5 4 3 2 4 3 4 2 ...

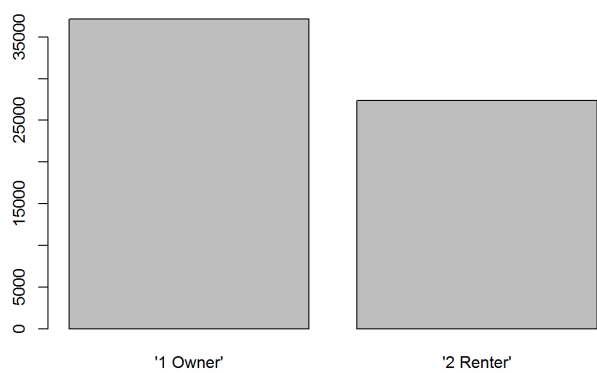
table(hads2013n$BEDRMS)

##
##      0      1      2      3      4      5      6      7
##  622  9821 16401 24850 10189  2209   392    51

table(hads2013n$FMTBEDRMS)

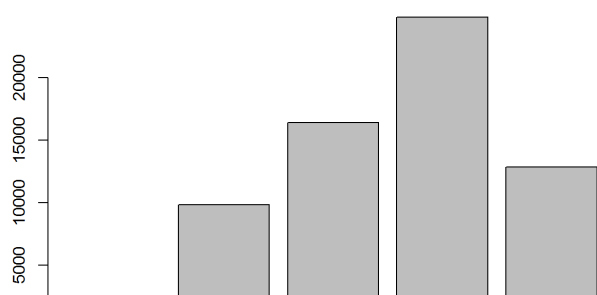
##
##  '0 Studio'  '1 1BR'  '2 2BR'  '3 3BR'  '4 4BR+'
##      622      9821      16401      24850      12841

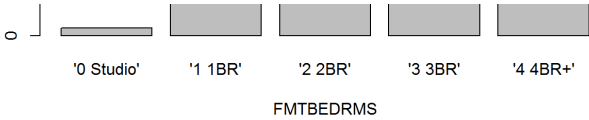
plot(hads2013n_c$FMTOWNRENT)
```



```
## FMTBEDRMS is under sorted by group, drop BEDRMS
hads2013n_c <- subset(hads2013n_c,select = -c(BEDRMS))

plot(hads2013n_c$FMTBEDRMS,xlab="FMTBEDRMS")
```





```
## check ASSISTED & FMTASSISTED: 2 options not assisted or assisted.
str(hads2013n$ASSISTED)

## int [1:64535] -9 -9 -9 -9 0 1 -9 0 -9 0 ...

str(hads2013n$FMTASSISTED)

## Factor w/ 3 levels "'0 Not Assisted'",...: 1 1 1 1 2 3 1 2 1 2 ...

table(hads2013n$ASSISTED)

##
##      -9      0      1
## 40290 17627  6618

table(hads2013n$FMTASSISTED)

##
##      '.' '0 Not Assisted' '1 Assisted'
##    40290      17627      6618

## since test try to remove ALL missing rows from ASSISTED and notice all Owner is removed
## we can say all assist is for renter only. we can remove both column ASSISTED & FMTASSISTED
hads2013n_c2 <- hads2013n[hads2013n$ASSISTED > -1,]
table(hads2013n_c2$FMTASSISTED)

##
##      '.' '0 Not Assisted' '1 Assisted'
##    40290      17627      6618

table(hads2013n_c2$FMTASSISTED)

##
##      '.' '0 Not Assisted' '1 Assisted'
##         0      17627      6618

table(hads2013n_c2$FMTOWNRENT)

##
## '1 Owner' '2 Renter'
##         0      24245

rm(hads2013n_c2)
hads2013n_c <- subset(hads2013n_c,select = -c(ASSISTED, FMTASSISTED))

## REGION & FMTREGION :The four Census regions. Drop FMTREGION
str(hads2013n$REGION)

## Factor w/ 4 levels "'1'", "'2'", "'3'",...: 1 3 3 3 3 3 3 4 4 2 ...

str(hads2013n$FMTREGION)

## Factor w/ 2 levels "'-5'", "'West'": 1 1 1 1 1 1 1 2 2 1 ...

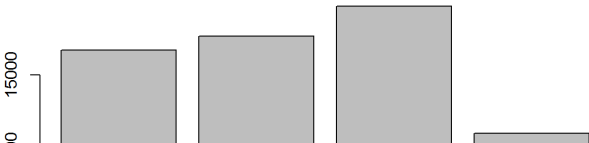
table(hads2013n$REGION)

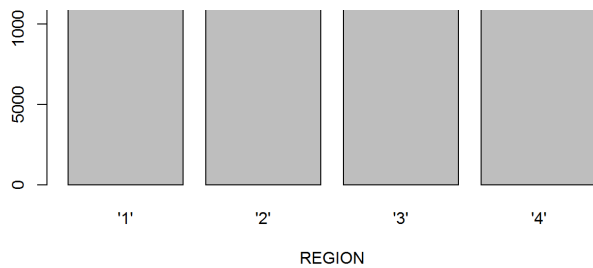
##
## '1' '2' '3' '4'
## 16519 17400 19260 11356

table(hads2013n$FMTREGION)

##
## '-5' 'West'
## 53179 11356

hads2013n_c <- subset(hads2013n_c,select = -c(FMTREGION))
plot(hads2013n_c$REGION,xlab="REGION")
```





```
## METRO3 & FMTMETRO3 :indicate whether a unit is in a central city, suburb, or outside a
## metropolitan area. Drop FMTMETRO3
str(hads2013n$METRO3)
```

```
## Factor w/ 5 levels "1","2","3",...: 3 5 5 5 1 2 1 4 5 4 ...
```

```
str(hads2013n$FMTMETRO3)
```

```
## Factor w/ 2 levels "-5","Central City": 1 1 1 1 2 1 2 1 1 1 ...
```

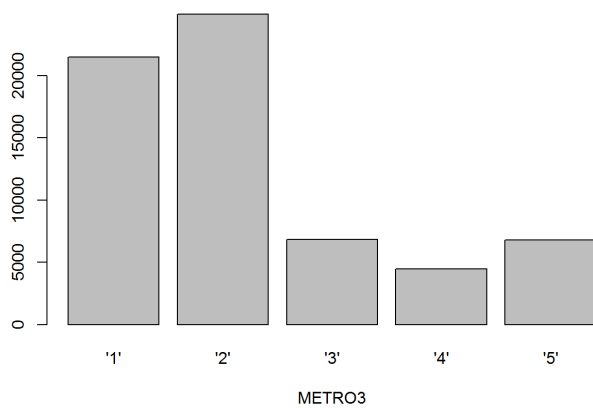
```
table(hads2013n$METRO3)
```

```
##
##      '1'   '2'   '3'   '4'   '5'
## 21493 24936  6851  4462  6793
```

```
table(hads2013n$FMTMETRO3)
```

```
##
##      '-5' 'Central City'
##    43042         21493
```

```
hads2013n_c <- subset(hads2013n_c,select = -c(FMTMETRO3))
plot(hads2013n_c$METRO3,xlab="METRO3")
```



```
##For the Variable below are Mortgage payments at 6, 8, and 12 percent interest rates
##COST12 Housing cost at 12 percent interest
##COST06 Housing cost at 6 percent interest
##COST08 Housing cost at 8 percent interest
##COSTMED Housing cost at Median interest
```

```
##COST06RELFMRCAT, FMTCOST06RELFMRCAT, COST06RELFMRPCT
##Cost06 Relative to FMR (Category), Cost06 Relative to FMR (Percent)
str(hads2013n$COST06RELFMRCAT)
```

```
## int [1:64535] 2 3 3 3 3 3 3 3 3 ...

str(hads2013n$FMTCOST06RELFMRCAT)

## Factor w/ 3 levels "1 LTE 50% FMR",...: 2 3 3 3 3 3 3 3 3 ...

str(hads2013n$COST06RELFMRPCT)

## num [1:64535] 67.8 106.1 108.5 166.4 103 ...

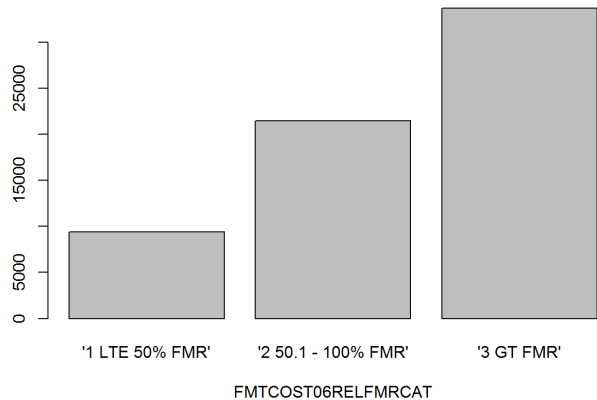
table(hads2013n$COST06RELFMRCAT)

##
##      1      2      3
## 9405 21440 33690

table(hads2013n$FMTCOST06RELFMRCAT)

##
##      '1 LTE 50% FMR' '2 50.1 - 100% FMR'      '3 GT FMR'
##           9405           21440           33690

hads2013n_c <- subset(hads2013n_c,select = -c(COST06RELFMRCAT))
plot(hads2013n_c$FMTCOST06RELFMRCAT,xlab="FMTCOST06RELFMRCAT")
```



```
##COST06RELAMICAT, FMTCOST06RELAMICAT, COST06RELAMIPCT
##Cost06 Relative to Median Income (Category), Cost06 Relative to Median Income (Category),
##Cost06 Relative to Median Income (Percent)
## Drop COST06RELAMICAT
str(hads2013n$COST06RELAMICAT)

## int [1:64535] 2 4 4 6 3 3 7 4 6 5 ...

str(hads2013n$FMTCOST06RELAMICAT)

## Factor w/ 7 levels "1 LTE 30% AMI",...: 2 4 4 6 3 3 7 4 6 5 ...

str(hads2013n$COST06RELAMIPCT)

## num [1:64535] 39.1 72.1 73.7 108.7 55.3 ...

table(hads2013n$COST06RELAMICAT)

##
##      1      2      3      4      5      6      7
## 8789 11479  7612 11939  7668  5157 11891

table(hads2013n$FMTCOST06RELAMICAT)

##
##      '1 LTE 30% AMI' '2 30 - 50% AMI' '3 50 - 60% AMI'
##           8789           11479           7612
##      '4 60 - 80% AMI' '5 80 - 100% AMI' '6 100 - 120% AMI'
##           11939           7668           5157
##      '7 120% AMI'
##
```

```
##      / 4438 Pct1 Y
##      11891
```

```
hads2013n_c <- subset(hads2013n_c,select = -c(COST06RELAMICAT))

##COST06RELPOVCAT      Cost06 Relative to Poverty Income (Category)
##FMTCOST06RELPOVCAT    Cost06 Relative to Poverty Income (Category)
##COST06RELPOVPCT      Cost06 Relative to Poverty Income (Percent)
str(hads2013n$COST06RELPOVCAT)
```

```
## int [1:64535] 4 3 4 4 3 4 4 4 4 4 ...
```

```
str(hads2013n$FMTCOST06RELPOVCAT)
```

```
## Factor w/ 5 levels "'','1 LTE Poverty'",...: 5 4 5 5 4 5 5 5 5 ...
```

```
str(hads2013n$COST06RELPOVPCT)
```

```
## num [1:64535] 234 193 309 452 196 ...
```

```
table(hads2013n$COST06RELPOVCAT)
```

```
##
##      -9      1      2      3      4
## 4438  6802  6114  7254 39927
```

```
table(hads2013n$FMTCOST06RELPOVCAT)
```

```
##
##      '1 LTE Poverty' '2 100-150% Poverty'
##      4438          6802          6114
## '3 150-200% Poverty' '4 200%+ Poverty'
##      7254          39927
```

```
##4438 missing at FMTCOST06RELPOVCAT, 6.88% of the rows to the entire dataset
## check the impact if we Remove the missing row
hads2013n_c2 <- hads2013n[hads2013n$COST06RELPOVCAT > 0,]
table(hads2013n_c2$COST06RELPOVCAT)
```

```
##
##      1      2      3      4
## 6802  6114  7254 39927
```

```
table(hads2013n_c2$FMTOWNRENT)
```

```
##
## '1 Owner' '2 Renter'
##      35852      24245
```

```
table(hads2013n$FMTOWNRENT)
```

```
##
## '1 Owner' '2 Renter'
##      37146      27389
```

```
## comparing the result owner drop from 37146 to 35852 and renter drop from 27389 to 24245
## so we dropp the missing rows and remove the testing dataset and the COST06RELPOVCAT column
rm(hads2013n_c2)
hads2013n_c <- hads2013n_c[hads2013n$COST06RELPOVCAT > 0,]
hads2013n_c <- subset(hads2013n_c,select = -c(COST06RELPOVCAT))
```

```
##SAME to all the following
#COST08RELFMRCAT      Cost08 Relative to FMR (Category)
#FMTCOST08RELFMRCAT  Cost08 Relative to FMR (Category)
#COST08RELFMRPCT      Cost08 Relative to FMR (Percent)
str(hads2013n_c$COST08RELFMRCAT)
```

```
## int [1:60097] 2 3 3 3 3 3 3 3 3 ...
```

```
str(hads2013n_c$FMTCOST08RELFMRCAT)
```

```
## Factor w/ 3 levels "'1 LTE 50% FMR'",...: 2 3 3 3 3 3 3 3 3 ...
```

```
str(hads2013n_c$COST08RELFMRPCT)
```

```
## num [1:60097] 72.9 120.4 125 191.8 103 ...
```

```
table(hads2013n_c$COST08RELFMRCAT)
```

```
##
##      1      2      3
## 8362 17590 34145
```

```
table(hads2013n_c$FMTCOST08RELFMRCAT)
```

```
##
##      '1 LTE 50% FMR' '2 50.1 - 100% FMR'      '3 GT FMR'
##      8362           17590                   34145
```

```
hads2013n_c <- subset(hads2013n_c,select = -c(COST08RELFMRCAT))

##COST08RELAMICAT Cost08 Relative to Median Income (Category)
##FMTCOST08RELAMICAT Cost08 Relative to Median Income (Category)
##COST08RELAMIPCT Cost08 Relative to Median Income (Percent)
str(hads2013n_c$COST08RELAMICAT)
```

```
## int [1:60097] 2 4 5 7 3 3 7 4 6 5 ...
```

```
str(hads2013n_c$FMTCOST08RELAMICAT)
```

```
## Factor w/ 7 levels "'1 LTE 30% AMI'",...: 2 4 5 7 3 3 7 4 6 5 ...
```

```
str(hads2013n_c$COST08RELAMIPCT)
```

```
## num [1:60097] 42 81.8 84.9 125.4 55.3 ...
```

```
table(hads2013n_c$COST08RELAMICAT)
```

```
##
##      1      2      3      4      5      6      7
## 7867  9251  6451 10014  7083  5181 14250
```

```
table(hads2013n_c$FMTCOST08RELAMICAT)
```

```
##
##      '1 LTE 30% AMI' '2 30 - 50% AMI' '3 50 - 60% AMI'
##      7867          9251          6451
##      '4 60 - 80% AMI' '5 80 - 100% AMI' '6 100 - 120% AMI'
##      10014          7083          5181
##      '7 120% AMI +'
##      14250
```

```
hads2013n_c <- subset(hads2013n_c,select = -c(COST08RELAMICAT))

##COST08RELPOVCAT Cost08 Relative to Poverty Income (Category)
##FMTCOST08RELPOVCAT Cost08 Relative to Poverty Income (Category)
##COST08RELPOVPCT Cost08 Relative to Poverty Income (Percent)
str(hads2013n_c$COST08RELPOVCAT)
```

```
## int [1:60097] 4 4 4 4 3 4 4 4 4 4 ...
```

```
str(hads2013n_c$FMTCOST08RELPOVCAT)
```

```
## Factor w/ 5 levels "'','1 LTE Poverty'",...: 5 5 5 5 4 5 5 5 5 ...
```

```
str(hads2013n_c$COST08RELPOVPCT)
```

```
## num [1:60097] 252 219 355 521 196 ...
```

```
table(hads2013n_c$COST08RELPOVCAT)
```

```
##
##      1      2      3      4
## 6547  5557  6657 41336
```

```
table(hads2013n_c$FMTCOST08RELPOVCAT)
```

```
##
##      '. '      '1 LTE Poverty' '2 100-150% Poverty'
##      0          6547          5557
##      '3 150-200% Poverty' '4 200%+ Poverty'
##      6657          41336
```

```
hads2013n_c <- subset(hads2013n_c,select = -c(COST08RELPOVCAT))

##COST12RELFMRCAT Cost12 Relative to FMR (Category)
##FMTCOST12RELFMRCAT Cost12 Relative to FMR (Category)
##COST12RELFMRPCT Cost12 Relative to FMR (Percent)
str(hads2013n_c$COST12RELFMRCAT)
```

```
## int [1:60097] 2 3 3 3 3 3 3 3 3 ...
```

```
str(hads2013n_c$FMTCOST12RELFMRCAT)
```

```
## Factor w/ 3 levels "'1 LTE 50% FMR'",...: 2 3 3 3 3 3 3 3 3 ...
```

```
str(hads2013n_c$COST12RELFMRPCT)
```

```
## num [1:60097] 84 152 161 248 103 ...
```



```
table(hads2013n_c$COST12RELFMRCAT)
```

```
##
##      1      2      3
## 7823 14731 37543
```

```
table(hads2013n_c$FMTTCOST12RELFMRCAT)
```

```
##
##      '1 LTE 50% FMR' '2 50.1 - 100% FMR'      '3 GT FMR'
##           7823           14731           37543
```

```
hads2013n_c <- subset(hads2013n_c,select = -c(COST12RELFMRCAT))

##COST12RELAMICAT   Cost12 Relative to Median Income (Category)
##FMTTCOST12RELAMICAT   Cost12 Relative to Median Income (Category)
##COST12RELAMIPCT   Cost12 Relative to Median Income (Percent)
str(hads2013n_c$COST12RELAMICAT)
```

```
## int [1:60097] 2 6 6 7 3 3 7 4 7 5 ...
```

```
str(hads2013n_c$FMTTCOST12RELAMICAT)
```

```
## Factor w/ 7 levels "1 LTE 30% AMI",...: 2 6 6 7 3 3 7 4 7 5 ...
```

```
str(hads2013n_c$COST12RELAMIPCT)
```

```
## num [1:60097] 48.4 103.1 109.5 161.9 55.3 ...
```

```
table(hads2013n_c$COST12RELAMICAT)
```

```
##
##      1      2      3      4      5      6      7
## 7417 7799 5429 8578 5810 5064 20000
```

```
table(hads2013n_c$FMTTCOST12RELAMICAT)
```

```
##
##      '1 LTE 30% AMI' '2 30 - 50% AMI' '3 50 - 60% AMI'
##           7417           7799           5429
##      '4 60 - 80% AMI' '5 80 - 100% AMI' '6 100 - 120% AMI'
##           8578           5810           5064
##      '7 120% AMI +'
##           20000
```

```
hads2013n_c <- subset(hads2013n_c,select = -c(COST12RELAMICAT))

##COST12RELPOVCAT   Cost12 Relative to Poverty Income (Category)
##FMTTCOST12RELPOVCAT   Cost12 Relative to Poverty Income (Category)
##COST12RELPOVPCT   Cost12 Relative to Poverty Income (Percent)
str(hads2013n_c$COST12RELPOVCAT)
```

```
## int [1:60097] 4 4 4 4 3 4 4 4 4 4 ...
```

```
str(hads2013n_c$FMTTCOST12RELPOVCAT)
```

```
## Factor w/ 5 levels "1 LTE Poverty",...: 5 5 5 5 4 5 5 5 5 ...
```

```
str(hads2013n_c$COST12RELPOVPCT)
```

```
## num [1:60097] 290 276 458 673 196 ...
```

```
table(hads2013n_c$COST12RELPOVCAT)
```

```
##
##      1      2      3      4
## 6254 4906 5710 43227
```

```
table(hads2013n_c$FMTTCOST12RELPOVCAT)
```

```
##
##      '1 LTE Poverty' '2 100-150% Poverty'
##           0           6254           4906
##      '3 150-200% Poverty' '4 200%+ Poverty'
##           5710           43227
```

```
hads2013n_c <- subset(hads2013n_c,select = -c(COST12RELPOVCAT))

##COSTMedRELFMRCAT   CostMed Relative to FMR (Category)
##FMTTCOSTMedRELFMRCAT   CostMed Relative to FMR (Category)
##COSTMedRELFMRPCT   CostMed Relative to FMR (Percent)
str(hads2013n_c$COSTMedRELFMRCAT)
```

```
## int [1:60097] 2 2 2 3 3 3 3 3 3 ...
```

```
str(hads2013n_c$FMTCOSTMEDRELFMRCAT)
```

```
## Factor w/ 3 levels "'1 LTE 50% FMR'",...: 2 2 2 3 3 3 3 3 3 ...
```

```
str(hads2013n_c$COSTMedRELFMRPCT)
```

```
## num [1:60097] 64.3 96.3 97.1 148.8 103 ...
```

```
table(hads2013n_c$COSTMedRELFMRCAT)
```

```
##
##      1      2      3
## 9234 21770 29093
```

```
table(hads2013n_c$FMTCOSTMEDRELFMRCAT)
```

```
##
##      '1 LTE 50% FMR' '2 50.1 - 100% FMR'      '3 GT FMR'
##      9234          21770          29093
```

```
hads2013n_c <- subset(hads2013n_c,select = -c(COSTMedRELFMRCAT))
```

```
##COSTMedRELAMICAT CostMed Relative to Median Income (Category)
##FMTCOSTMEDRELAMICAT CostMed Relative to Median Income (Category)
##COSTMedRELAMIPCT CostMed Relative to Median Income (Percent)
str(hads2013n_c$COSTMedRELAMICAT)
```

```
## int [1:60097] 2 4 4 5 3 3 6 4 5 5 ...
```

```
str(hads2013n_c$FMTCOSTMEDRELAMICAT)
```

```
## Factor w/ 7 levels "'1 LTE 30% AMI'",...: 2 4 4 5 3 3 6 4 5 5 ...
```

```
str(hads2013n_c$COSTMedRELAMIPCT)
```

```
## num [1:60097] 37.1 65.4 65.9 97.2 55.3 ...
```

```
table(hads2013n_c$COSTMedRELAMICAT)
```

```
##
##      1      2      3      4      5      6      7
## 8536 11621 7622 11503 7149 4515 9151
```

```
table(hads2013n_c$FMTCOSTMEDRELAMICAT)
```

```
##
##      '1 LTE 30% AMI' '2 30 - 50% AMI' '3 50 - 60% AMI'
##      8536          11621          7622
##      '4 60 - 80% AMI' '5 80 - 100% AMI' '6 100 - 120% AMI'
##      11503          7149          4515
##      '7 120% AMI +'
##      9151
```

```
hads2013n_c <- subset(hads2013n_c,select = -c(COSTMedRELAMICAT))
```

```
##COSTMedRELPOVCAT CostMed Relative to Poverty Income (Category)
##FMTCOSTMEDRELPOVCAT CostMed Relative to Poverty Income (Category)
##COSTMedRELPOVPCT CostMed Relative to Poverty Income (Percent)
str(hads2013n_c$COSTMedRELPOVCAT)
```

```
## int [1:60097] 4 3 4 4 3 4 4 4 4 ...
```

```
str(hads2013n_c$FMTCOSTMEDRELPOVCAT)
```

```
## Factor w/ 5 levels "'','"','1 LTE Poverty'",...: 5 4 5 5 4 5 5 5 5 ...
```

```
str(hads2013n_c$COSTMedRELPOVPCT)
```

```
## num [1:60097] 222 175 276 404 196 ...
```

```
table(hads2013n_c$COSTMedRELPOVCAT)
```

```
##
##      1      2      3      4
## 7042 6634 7917 38504
```

```
table(hads2013n_c$FMTCOSTMEDRELPOVCAT)
```

```
##
##      '.'      '1 LTE Poverty' '2 100-150% Poverty'
##      0      7042      6634
## '3 150-200% Poverty' '4 200%+ Poverty'
##      7917      38504
```

```
hads2013n_c <- subset(hads2013n_c,select = ~(COSTMedRELPOVCAT))
```

```
##VALUE Current market value of unit
str(hads2013n_c$VALUE)
```

```
## int [1:60097] 40000 130000 150000 200000 -6 -6 260000 -6 170000 -6 ...
```

```
table(hads2013n_c$VALUE,useNA = 'always')
```

```
##
##      -6      1    10000    20000    30000    40000    50000    60000    70000
## 24245    434    553    515    609    712    948    853    1041
## 80000    90000    100000    110000    120000    130000    140000    150000    160000
## 1476    1274    1721    659    1141    1341    931    1858    768
## 170000    180000    190000    200000    210000    220000    230000    240000    250000
## 760    1346    579    1777    326    468    767    430    1394
## 260000    270000    280000    290000    300000    310000    320000    330000    340000
## 320    249    618    240    1413    119    272    408    129
## 350000    360000    370000    380000    390000    400000    410000    420000    430000
## 984    126    112    340    127    918    57    80    192
## 440000    450000    460000    470000    480000    490000    500000    510000    520000
## 71    534    63    32    146    39    626    8    30
## 530000    540000    550000    560000    570000    580000    590000    600000    610000
## 77    24    248    27    15    64    15    410    5
## 620000    630000    640000    650000    660000    670000    680000    690000    700000
## 23    41    14    223    8    11    38    15    221
## 710000    720000    730000    740000    750000    760000    770000    780000    790000
## 3    15    17    5    174    6    2    16    10
## 800000    810000    820000    830000    840000    850000    860000    870000    880000
## 204    3    4    16    3    106    1    4    16
## 890000    900000    910000    920000    930000    940000    950000    960000    970000
## 2    125    1    3    1    1    32    1    1
## 980000    990000    1000000    1010000    1020000    1040000    1050000    1080000    1100000
## 6    3    199    1    4    1    7    1    36
## 1130000    1200000    1240000    1250000    1300000    1350000    1400000    2520000    <NA>
## 1    70    1    13    23    4    1    296    0
```

```
##Noticed that there are 24245 and 434 are error with value (-6 or 1),
24679/nrow(hads2013n_c)
```

```
## [1] 0.4106528
```

```
##41% of this column in the dataset is missing.
##Test to drop the affect of dropping the error rows.
hads2013n_c2 <- hads2013n[hads2013n$VALUE > 2,]
table(hads2013n_c2$VALUE,useNA = 'always')
```

```
##
## 10000    20000    30000    40000    50000    60000    70000    80000    90000
## 598    552    645    763    979    908    1089    1524    1317
## 100000    110000    120000    130000    140000    150000    160000    170000    180000
## 1774    696    1178    1381    967    1918    818    798    1381
## 190000    200000    210000    220000    230000    240000    250000    260000    270000
## 603    1817    333    487    785    441    1421    334    257
## 280000    290000    300000    310000    320000    330000    340000    350000    360000
## 636    258    1444    123    280    415    131    991    134
## 370000    380000    390000    400000    410000    420000    430000    440000    450000
## 119    345    132    925    60    81    203    75    541
## 460000    470000    480000    490000    500000    510000    520000    530000    540000
## 66    41    149    41    653    9    33    77    25
## 550000    560000    570000    580000    590000    600000    610000    620000    630000
## 254    32    18    65    16    418    6    24    45
## 640000    650000    660000    670000    680000    690000    700000    710000    720000
## 14    227    9    11    39    15    225    3    15
## 730000    740000    750000    760000    770000    780000    790000    800000    810000
## 18    5    180    6    2    17    11    212    3
## 820000    830000    840000    850000    860000    870000    880000    890000    900000
## 5    16    5    108    1    4    17    6    126
```

```
##      910000  920000  930000  940000  950000  960000  970000  980000  990000
##      1      3      1      1      35      1      1      7      3
## 1000000 1010000 1020000 1040000 1050000 1080000 1100000 1130000 1200000
##      204      1      4      1      7      1      36      1      71
## 1230000 1240000 1250000 1300000 1350000 1400000 2520000 <NA>
##      1      1      13      23      4      1      324      0
```

```
table(hads2013n_c2$FMTOWNRENT,useNA = 'always')
```

```
##
## '1 Owner' '2 Renter'      <NA>
##      36675      0      0
```

```
table(hads2013n_c$FMTOWNRENT,useNA = 'always')
```

```
##
## '1 Owner' '2 Renter'      <NA>
##      35852      24245      0
```

```
## by comparing two tables, all the Renter got removed from the removing the error rows.
## so we drop the value column
rm(hads2013n_c2)
hads2013n_c <- subset(hads2013n_c,select = -c(VALUE))
```

```
## INCRELFMRCAT HH Income Relative to FMR (Category)
## FMTINCRELFMRCAT HH Income Relative to FMR (Category)
## INCRELFMRPCT HH Income Relative to FMR (Percent)
table(hads2013n_c$INCRELFMRCAT,useNA = 'always')
```

```
##
##      1      2      3 <NA>
## 15960 14007 30130      0
```

```
table(hads2013n_c$FMTINCRELFMRCAT,useNA = 'always')
```

```
##
##      '1 LTE 50% FMR' '2 50.1 - 100% FMR'
##      0      15960      14007
## '3 GT FMR'      <NA>
##      30130      0
```

```
hads2013n_c <- subset(hads2013n_c,select = -c(INCRELFMRCAT))
```

```
##INCRELPOVCAT HH Income Relative to Poverty Income (Category)
##FMTINCRELPOVCAT HH Income Relative to Poverty Income (Category)
##INCRELPOVPCT HH Income Relative to Poverty Income (Percent)
table(hads2013n_c$INCRELPOVCAT,useNA = 'always')
```

```
##
##      1      2      3      4 <NA>
## 11810 6082 5702 36503      0
```

```
table(hads2013n_c$FMTINCRELPOVCAT,useNA = 'always')
```

```
##
##      '1 LTE Poverty' '2 100-150% Poverty'
##      0      11810      6082
## '3 150-200% Poverty' '4 200%+ Poverty'
##      5702      36503      0
```

```
hads2013n_c <- subset(hads2013n_c,select = -c(INCRELPOVCAT))
```

```
##STATUS Interview status
table(hads2013n_c$STATUS,useNA = 'always')
```

```
##
## '1' '3' <NA>
## 60097      0      0
```

```
hads2013n_c <- subset(hads2013n_c,select = -c(STATUS))
```

```
## FMTSTATUS Occupancy Status
table(hads2013n_c$FMTSTATUS,useNA = 'always')
```

```
##
## '-5' <NA>
## 60097      0
```

```
hads2013n_c <- subset(hads2013n_c,select = -c(FMTSTATUS))
```

```
##STRUCTURETYPE Recoded structure type
##TYPE Structure type
##FMTSTRUCTURETYPE Structure Type
str(hads2013n_c$STRUCTURETYPE)
```

```
cat(hads2013n_c$STRUCTURETYPE)
```

```
## int [1:60097] 1 1 1 1 5 4 1 3 1 4 ...
```

```
str(hads2013n_c$TYPE)
```

```
## int [1:60097] 1 1 1 1 1 1 1 1 1 ...
```

```
str(hads2013n_c$FMTSTRUCTURETYPE)
```

```
## Factor w/ 7 levels "'','1 Single Family'",...: 2 2 2 2 6 5 2 4 2 5 ...
```

```
table(hads2013n_c$STRUCTURETYPE,useNA = 'always')
```

```
##
##      -9      1      2      3      4      5      6 <NA>
##      1 39376  5583  6389  2391  4075  2282      0
```

```
table(hads2013n_c$TYPE,useNA = 'always')
```

```
##
##      1      2      3      4      5      6      7      9 <NA>
## 57655 1928  354      4      7     21    10   118      0
```

```
table(hads2013n_c$FMTSTRUCTURETYPE,useNA = 'always')
```

```
##
##           '1 Single Family'   '2 2-4 units'   '3 5-19 units'
##           1           39376           5583           6389
## '4 20-49 units'   '5 50+ units'   '6 Mobile Home'           <NA>
##           2391           4075           2282           0
```

```
hads2013n_c <- hads2013n_c[hads2013n_c$STRUCTURETYPE > 0,]
hads2013n_c <- subset(hads2013n_c,select = -c(STRUCTURETYPE))
```

```
##VACANCY Vacancy status
## all the of the rows are -6 drop column
str(hads2013n_c$VACANCY)
```

```
## int [1:60096] -6 -6 -6 -6 -6 -6 -6 -6 -6 ...
```

```
table(hads2013n_c$VACANCY,useNA = 'always')
```

```
##
##      -6 <NA>
## 60096      0
```

```
hads2013n_c <- subset(hads2013n_c,select = -c(VACANCY))
```

```
##BUILT Year unit was built
##FMTBUILT YEAR UNIT WAS BUILT
## notice that the group of '-5' in FMTBUILT is the the year before, we name it group 'before 1940'
str(hads2013n_c$BUILT)
```

```
## int [1:60096] 2006 1980 1985 1985 1980 1985 1985 1980 1985 1985 ...
```

```
str(hads2013n_c$FMTBUILT)
```

```
## Factor w/ 7 levels "'-5','1940-1959'",...: 6 4 4 4 4 4 4 4 4 ...
```

```
table(hads2013n_c$BUILT,useNA = 'always')
```

```
##
## 1919 1920 1930 1940 1950 1960 1970 1975 1980 1985 1990 1995 2000 2001 2002
## 3904 2658 2643 3636 6777 7264 5192 5847 3750 3861 2366 4776  872  726  649
## 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 <NA>
##  683  764  798  773  661  474  335  245  189  221  32      0
```

```
table(hads2013n_c$FMTBUILT,useNA = 'always')
```

```
##
##      '-5'   '1940-1959'   '1960-1979'   '1980-1989'   '1990-1999'
##      9205      10413      18303           7611           7142
## '2000-2009' 'After 2010'           <NA>
##      6735           687           0
```

```
levels(hads2013n_c$FMTBUILT)[match("' -5'",levels(hads2013n_c$FMTBUILT))] <- "'Before 1940'"
table(hads2013n_c$FMTBUILT,useNA = 'always')
```

```
##
## 'Before 1940'   '1940-1959'   '1960-1979'   '1980-1989'   '1990-1999'
##      9205      10413      18303           7611           7142
## '2000-2009' 'After 2010'           <NA>
```

```
##           6735           687           0
```

```
hads2013n_c <- subset(hads2013n_c,select = -c(BUILT))
```

```
##ZADEQ ADEQUACY OF UNIT
##FMTZADEQ ADEQUACY OF UNIT
str(hads2013n_c$ZADEQ)
```

```
## Factor w/ 4 levels "-6","-1","-2",...: 2 2 2 2 2 2 3 2 2 ...
```

```
str(hads2013n_c$FMTZADEQ)
```

```
## Factor w/ 4 levels "-5","1 Adequate",...: 2 2 2 2 2 2 3 2 2 ...
```

```
table(hads2013n_c$ZADEQ,useNA = 'always')
```

```
##
##  '-6'  '1'  '2'  '3' <NA>
##    0 56787  2148 1161    0
```

```
table(hads2013n_c$FMTZADEQ,useNA = 'always')
```

```
##
##           '-5'           '1 Adequate' '2 Moderately Inadequ'
##           0           56787           2148
## '3 Severely Indadequa'
##           1161           0
```

```
hads2013n_c <- subset(hads2013n_c,select = -c(ZADEQ))
```

```
##OWNRENT      Tenure (adjusted)
##FMTOWNRENT    Owner/Renter Status (adjusted)
str(hads2013n_c$OWNRENT)
```

```
## Factor w/ 2 levels "'1","2": 1 1 1 1 2 2 1 2 1 2 ...
```

```
str(hads2013n_c$FMTOWNRENT)
```

```
## Factor w/ 2 levels "'1 Owner","2 Renter": 1 1 1 1 2 2 1 2 1 2 ...
```

```
table(hads2013n_c$OWNRENT,useNA = 'always')
```

```
##
##  '1'  '2' <NA>
## 35852 24244    0
```

```
table(hads2013n_c$FMTOWNRENT,useNA = 'always')
```

```
##
##  '1 Owner' '2 Renter'      <NA>
##    35852    24244          0
```

```
hads2013n_c <- subset(hads2013n_c,select = -c(OWNRENT))
```

```
## Drop control since it's identical Variable
hads2013n_c <- subset(hads2013n_c,select = -c(CONTROL))
```

## test for the correlation between variables

## peason for numeric variable and spearman for category variable

```
##check the variable is either factor or numeric
nums <- sapply(hads2013n_c, is.numeric)
cate<-sapply(hads2013n_c, is.factor)
str(hads2013n_c[,cate])
```

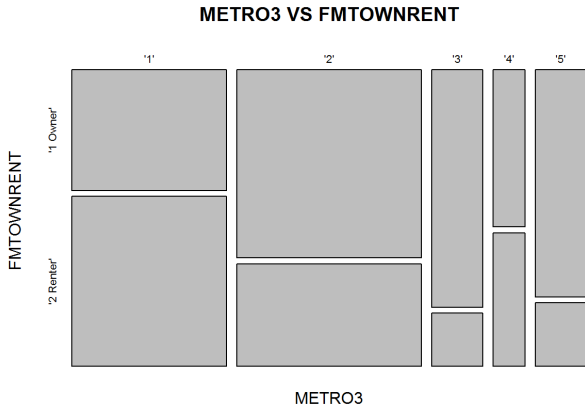
```
## 'data.frame':   60096 obs. of  23 variables:
## $ METRO3      : Factor w/ 5 levels "'1","2","3",...: 3 5 5 5 1 2 1 4 5 4 ...
## $ REGION      : Factor w/ 4 levels "'1","2","3",...: 1 3 3 3 3 3 3 4 4 2 ...
## $ FMTZADEQ     : Factor w/ 4 levels "'-5","1 Adequate",...: 2 2 2 2 2 2 2 3 2 2 ...
## $ FMTBUILT     : Factor w/ 7 levels "'Before 1940",...: 6 4 4 4 4 4 4 4 4 4 ...
## $ FMTSTRUCTURETYPE : Factor w/ 7 levels "'1 Single Family",...: 2 2 2 2 6 5 2 4 2 5 ...
## $ FMTBEDRMS    : Factor w/ 5 levels "'0 Studio","1 1BR",...: 3 5 5 4 3 2 4 3 4 2 ...
## $ FMTOWNRENT   : Factor w/ 2 levels "'1 Owner","2 Renter": 1 1 1 1 2 2 1 2 1 2 ...
## $ FMTCOST06RELPOVCAT : Factor w/ 5 levels "'1 LTE Poverty",...: 5 4 5 5 4 5 5 5 5 5 ...
## $ FMTCOST08RELPOVCAT : Factor w/ 5 levels "'1 LTE Poverty",...: 5 5 5 5 4 5 5 5 5 5 ...
## $ FMTCOST12RELPOVCAT : Factor w/ 5 levels "'1 LTE Poverty",...: 5 5 5 5 4 5 5 5 5 5 ...
## $ FMTCOSTMEDRELPOVCAT : Factor w/ 5 levels "'1 LTE Poverty",...: 5 4 5 5 4 5 5 5 5 5 ...
## $ FMTTNCREIDPOVCAT : Factor w/ 5 levels "'1 LTE Poverty",...: 4 4 4 5 5 3 5 5 5 5
```

```
## $ FMTINCRELFPMRCAT : Factor w/ 3 levels "'1 LTE 50% FMR'",...: 2 3 3 3 3 3 3 3 3 ...
## $ FMTCOST08RELFMRCAT : Factor w/ 3 levels "'1 LTE 50% FMR'",...: 2 3 3 3 3 3 3 3 3 ...
## $ FMTCOST12RELFMRCAT : Factor w/ 3 levels "'1 LTE 50% FMR'",...: 2 3 3 3 3 3 3 3 3 ...
## $ FMTCOSTMEDRELFMRCAT : Factor w/ 3 levels "'1 LTE 50% FMR'",...: 2 2 2 3 3 3 3 3 3 ...
## $ FMTINCRELFMRCAT : Factor w/ 4 levels "'1 LTE 50% FMR'",...: 2 4 3 3 4 3 4 4 4 ...
## $ FMTCOST06RELAMICAT : Factor w/ 7 levels "'1 LTE 30% AMI'",...: 2 4 4 6 3 3 7 4 6 5 ...
## $ FMTCOST08RELAMICAT : Factor w/ 7 levels "'1 LTE 30% AMI'",...: 2 4 5 7 3 3 7 4 6 5 ...
## $ FMTCOST12RELAMICAT : Factor w/ 7 levels "'1 LTE 30% AMI'",...: 2 6 6 7 3 3 7 4 7 5 ...
## $ FMTCOSTMEDRELAMICAT : Factor w/ 7 levels "'1 LTE 30% AMI'",...: 2 4 4 5 3 3 6 4 5 5 ...
## $ FMTINCRELAMICAT : Factor w/ 8 levels "'1 LTE 30% AMI'",...: 3 8 5 5 8 3 8 5 8 7 ...
## $ FMTBURDEN : Factor w/ 5 levels "'1 Less than 30%'",...: 3 2 4 2 2 4 2 3 2 2 ...
```

```
str(hads2013n_c[,nums])
```

```
## 'data.frame': 60096 obs. of 47 variables:
## $ AGE1 : int 82 50 53 67 26 56 50 26 60 26 ...
## $ LMED : int 73738 55846 55846 55846 60991 62066 60991 52322 50296 63221 ...
## $ FMR : int 956 1100 1100 949 737 657 988 773 1125 552 ...
## $ L30 : int 15738 17165 13750 13750 14801 13170 16646 13489 13115 13338 ...
## $ L50 : int 26213 28604 22897 22897 24628 21924 27713 22471 21859 22199 ...
## $ L80 : int 40322 45744 36614 36614 39421 35073 44340 35929 34939 35501 ...
## $ IPOV : int 11067 24218 15470 13964 15492 12005 18050 15992 15452 12005 ...
## $ TYPE : int 1 1 1 1 1 1 1 1 1 1 ...
## $ NUNITS : int 1 1 1 1 100 32 1 8 1 24 ...
## $ ROOMS : int 6 6 7 6 4 3 6 5 7 3 ...
## $ WEIGHT : num 3117 2151 2214 2365 2315 ...
## $ PER : int 1 4 2 2 2 1 3 2 2 1 ...
## $ ZINC2 : int 18021 122961 27974 32220 96874 14987 69962 32000 118987 47987 ...
## $ ZSMHC : int 533 487 1405 279 759 695 1165 976 1156 1100 ...
## $ UTILITY : num 169 245 159 179 146 ...
## $ OTHERCOST : num 213.8 58.3 37.5 70.7 12.5 ...
## $ COST06 : num 649 1168 1193 1579 759 ...
## $ COST12 : num 803 1670 1773 2351 759 ...
## $ COST08 : num 697 1325 1375 1820 759 ...
## $ COSTMED : num 615 1059 1068 1412 759 ...
## $ TOTSAL : int 0 123000 28000 0 96900 15000 70001 20000 107000 48000 ...
## $ GLMED : int 73738 55846 55846 55846 60991 62066 60991 52322 50296 63221 ...
## $ GL30 : int 15738 17165 13750 13750 14801 13170 16646 13489 13115 13338 ...
## $ GL50 : int 26213 28604 22897 22897 24628 21924 27713 22471 21859 22199 ...
## $ GL80 : int 40322 45744 36614 36614 39421 35073 44340 35929 34939 35501 ...
## $ APLMED : num 51617 55846 44677 44677 48793 ...
## $ ABL30 : num 20235 19911 19938 17875 16651 ...
## $ ABL50 : num 33702 33181 33201 29766 27707 ...
## $ ABL80 : num 51843 53063 53090 47598 44349 ...
## $ ABLMED : num 66364 64781 64781 50800 54892 ...
## $ BURDEN : num 0.3549 0.0475 0.6027 0.1039 0.094 ...
## $ INCRELAMIPCT : num 34.9 220.2 62.6 72.1 198.5 ...
## $ INCRELAMICAT : int 2 7 4 4 7 2 7 4 7 6 ...
## $ INCRELPVPCT : num 163 508 181 231 625 ...
## $ INCRELFMRPCT : num 47.1 279.5 63.6 84.9 328.6 ...
## $ COST06RELAMIPCT : num 39.1 72.1 73.7 108.7 55.3 ...
## $ COST06RELPVPCT : num 234 193 309 452 196 ...
## $ COST06RELFMRPCT : num 67.8 106.1 108.5 166.4 103 ...
## $ COST08RELAMIPCT : num 42 81.8 84.9 125.4 55.3 ...
## $ COST08RELPVPCT : num 252 219 355 521 196 ...
## $ COST08RELFMRPCT : num 72.9 120.4 125 191.8 103 ...
## $ COST12RELAMIPCT : num 48.4 103.1 109.5 161.9 55.3 ...
## $ COST12RELPVPCT : num 290 276 458 673 196 ...
## $ COST12RELFMRPCT : num 84 152 161 248 103 ...
## $ COSTMedRELAMIPCT : num 37.1 65.4 65.9 97.2 55.3 ...
## $ COSTMedRELPVPCT : num 222 175 276 404 196 ...
## $ COSTMedRELFMRPCT : num 64.3 96.3 97.1 148.8 103 ...
```

```
###
##for Category vs Category Variables, simpson method is use to test the correlation
plot(table(hads2013n_c$METRO3,hads2013n_c$FMTOWNRENT),main="METRO3 VS FMTOWNRENT",xlab="METRO3",ylab="FMTOWNRENT")
```



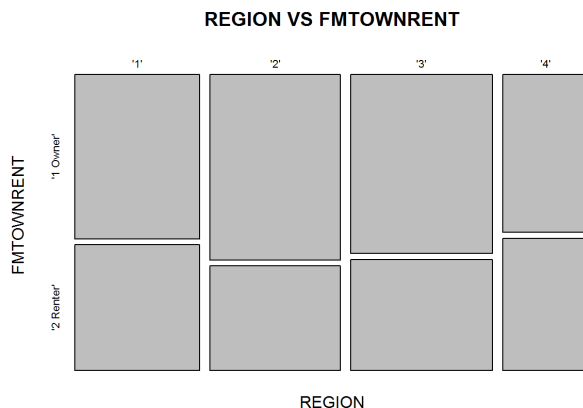
```
summary(glm(FMTOWNRENT ~ METRO3,family=binomial(link = "logit"),data=hads2013n_c))
```

```
##
## Call:
```

```
## glm(formula = FMTOWNRENT ~ METRO3, family = binomial(link = "logit"),
## data = hads2013n_c)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.3257  -0.9334  -0.7021   1.0360   1.8427
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  0.34203    0.01447   23.64 <2e-16 ***
## METRO3'2'    -0.94727    0.01990  -47.61 <2e-16 ***
## METRO3'3'    -1.83758    0.03517  -52.24 <2e-16 ***
## METRO3'4'    -0.50426    0.03464  -14.56 <2e-16 ***
## METRO3'5'    -1.61674    0.03354  -48.20 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 81054  on 60095  degrees of freedom
## Residual deviance: 75688  on 60091  degrees of freedom
## AIC: 75698
##
## Number of Fisher Scoring iterations: 4
```

*#This difference appears large enough to suggest that a relationship does  
#exist between METRO3 and FMTOWNRENT.*

```
plot(table(hads2013n_c$REGION,hads2013n_c$FMTOWNRENT),main="REGION VS FMTOWNRENT",xlab="REGION",ylab="FMTOWNRENT")
```



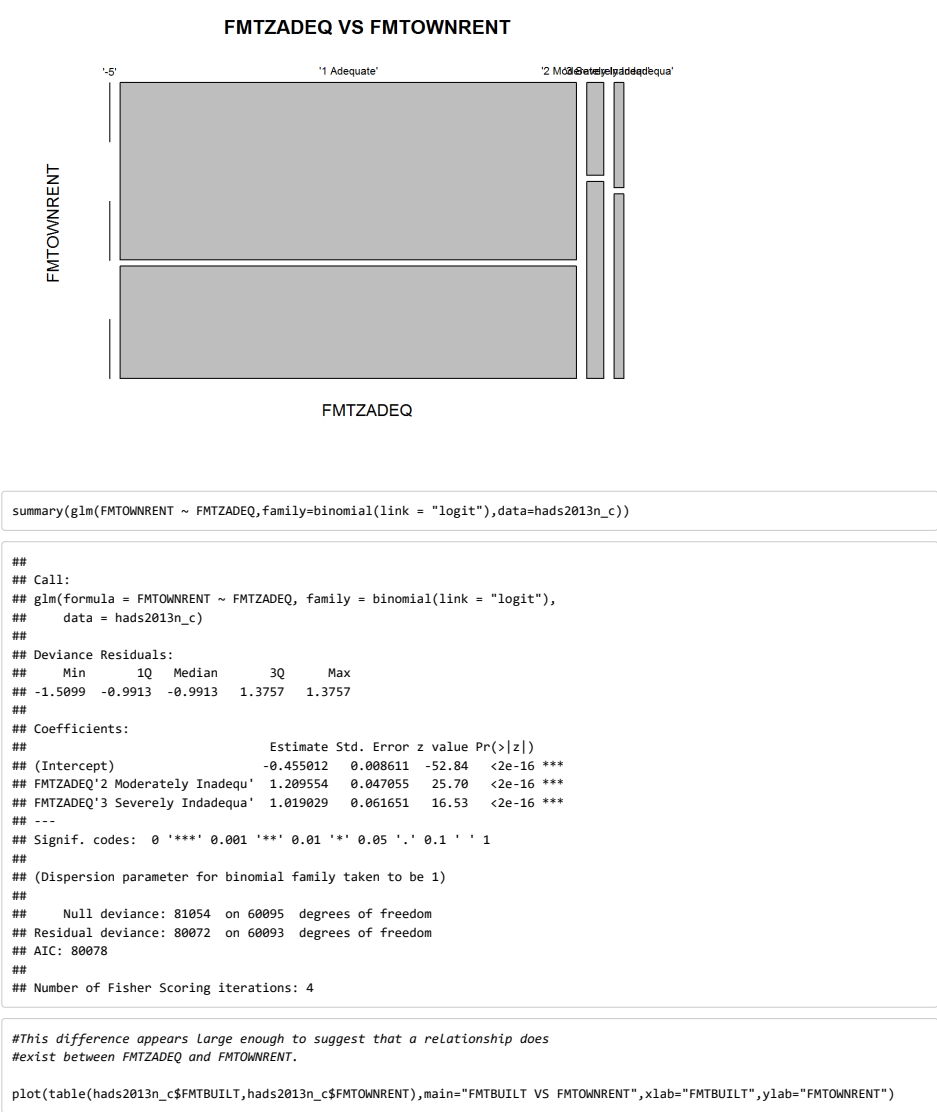
```
summary(glm(FMTOWNRENT ~ REGION,family=binomial(link = "logit"),data=hads2013n_c))
```

```
##
## Call:
## glm(formula = FMTOWNRENT ~ REGION, family = binomial(link = "logit"),
## data = hads2013n_c)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.1039  -0.9834  -0.9465   1.2921   1.4275
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -0.26564    0.01619 -16.413 < 2e-16 ***
## REGION'2'    -0.30530    0.02301 -13.271 < 2e-16 ***
## REGION'3'    -0.20943    0.02240  -9.348 < 2e-16 ***
## REGION'4'     0.09017    0.02527   3.569 0.000359 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 81054  on 60095  degrees of freedom
## Residual deviance: 80720  on 60092  degrees of freedom
## AIC: 80728
##
## Number of Fisher Scoring iterations: 4
```

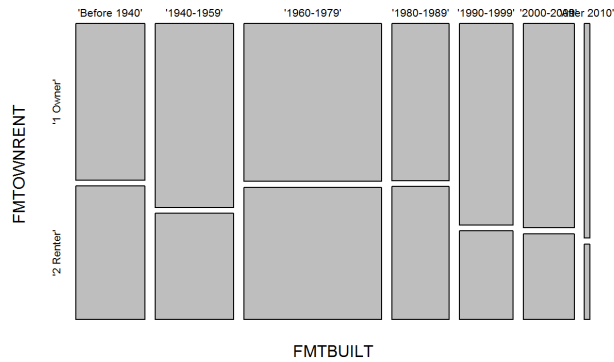
*#This difference appears slightly enough to suggest that a relationship does  
#exist between METRO3 and FMTOWNRENT.*

```
plot(table(hads2013n_c$FMTZADEQ,hads2013n_c$FMTOWNRENT),main="FMTZADEQ VS FMTOWNRENT",xlab="FMTZADEQ",ylab="FMTOWNRENT")
```





FMTBUILT VS FMTOWNRENT

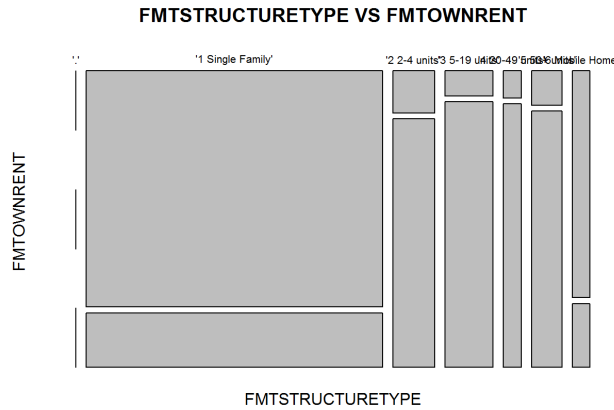


```
summary(glm(FMTOWNRENT ~ FMTBUILT,family=binomial(link = "logit"),data=hads2013n_c))
```

```
##
## Call:
## glm(formula = FMTOWNRENT ~ FMTBUILT, family = binomial(link = "logit"),
##      data = hads2013n_c)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.1110  -1.1032  -0.8544   1.2535   1.6401
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    -0.158287    0.020911  -7.570 3.75e-14 ***
## FMTBUILT'1940-1959' -0.388711    0.029170 -13.326 < 2e-16 ***
## FMTBUILT'1960-1979' -0.018867    0.025642  -0.736  0.462
## FMTBUILT'1980-1989' -0.007907    0.031088  -0.254  0.799
## FMTBUILT'1990-1999' -0.661557    0.033119 -19.975 < 2e-16 ***
## FMTBUILT'2000-2009' -0.709243    0.033913 -20.913 < 2e-16 ***
## FMTBUILT'After 2010' -0.884808    0.089400  -9.897 < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 81054  on 60095  degrees of freedom
## Residual deviance: 79879  on 60089  degrees of freedom
## AIC: 79893
##
## Number of Fisher Scoring iterations: 4
```

```
#This difference appears large enough to suggest that a relationship does
#exist between FMTBUILT and FMTOWNRENT.

plot(table(hads2013n_c$FMTSTRUCTURETYPE,hads2013n_c$FMTOWNRENT),main="FMTSTRUCTURETYPE VS FMTOWNRENT",xlab="FMTSTRUCTURETYPE",ylab="FMTOWNRENT")
```

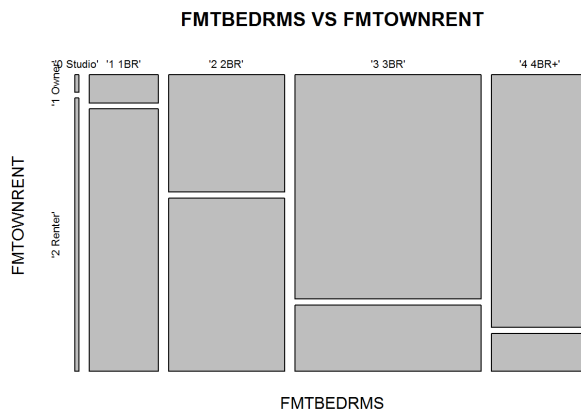


```
summary(glm(FMTOWNRENT ~ FMTSTRUCTURETYPE,family=binomial(link = "logit"),data=hads2013n_c))
```

```
##
## Call:
## glm(formula = FMTOWNRENT ~ FMTSTRUCTURETYPE, family = binomial(link = "logit"),
##      data = hads2013n_c)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.2139  -0.6439  -0.6439   0.5018   1.8305
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    -1.46800     0.01292 -113.639 < 2e-16 ***
## FMTSTRUCTURETYPE'2 2-4 units'    3.24026     0.04012   80.768 < 2e-16 ***
## FMTSTRUCTURETYPE'3 5-19 units'    3.82841     0.04639   82.531 < 2e-16 ***
## FMTSTRUCTURETYPE'4 20-49 units'    3.73746     0.07135   52.381 < 2e-16 ***
## FMTSTRUCTURETYPE'5 50+ units'    3.47680     0.05020   69.262 < 2e-16 ***
## FMTSTRUCTURETYPE'6 Mobile Home'    0.19968     0.05220    3.826  0.00013 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 81054  on 60095  degrees of freedom
## Residual deviance: 53207  on 60090  degrees of freedom
## AIC: 53219
##
## Number of Fisher Scoring iterations: 4
```

*#This difference appears large enough to suggest that a relationship does  
#exist between FMTSTRUCTURETYPE and FMTOWNRENT.*

```
plot(table(hads2013n_c$FMTBEDRMS,hads2013n_c$FMTOWNRENT),main="FMTBEDRMS VS FMTOWNRENT",xlab="FMTBEDRMS",ylab="FMTOWNRENT")
```



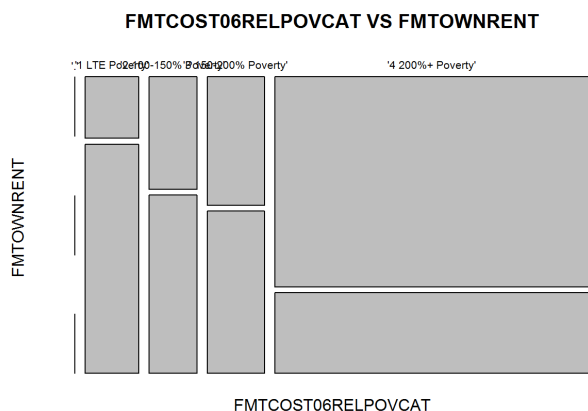
```
summary(glm(FMTOWNRENT ~ FMTBEDRMS,family=binomial(link = "logit"),data=hads2013n_c))
```

```
##
## Call:
## glm(formula = FMTOWNRENT ~ FMTBEDRMS, family = binomial(link = "logit"),
##      data = hads2013n_c)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.3763  -0.7200  -0.5299   1.0171   2.0163
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)     2.7621     0.1882  14.673 < 2e-16 ***
## FMTBEDRMS'1 1BR'  -0.5305     0.1917  -2.768  0.00565 **
## FMTBEDRMS'2 2BR'  -2.3725     0.1890 -12.553 < 2e-16 ***
## FMTBEDRMS'3 3BR'  -3.9798     0.1889 -21.070 < 2e-16 ***
## FMTBEDRMS'4 4BR+' -4.6545     0.1901 -24.481 < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 81054  on 60095  degrees of freedom
```

```
## Residual deviance: 60792 on 60091 degrees of freedom
## AIC: 60802
##
## Number of Fisher Scoring iterations: 5
```

*#This difference appears large enough to suggest that a relationship does exist between FMTBEDRMS and FMTOWNRENT.*

```
plot(table(hads2013n_c$FMTCOST06RELPOVCAT,hads2013n_c$FMTOWNRENT),main="FMTCOST06RELPOVCAT VS FMTOWNRENT",xlab="FMTCOST06RELPOVCAT",ylab="FMTOWNRENT")
```

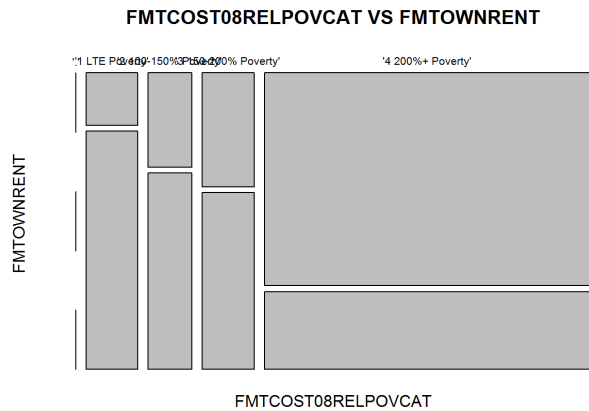


```
summary(glm(FMTOWNRENT ~ FMTCOST06RELPOVCAT,family=binomial(link = "logit"),data=hads2013n_c))
```

```
##
## Call:
## glm(formula = FMTOWNRENT ~ FMTCOST06RELPOVCAT, family = binomial(link = "logit"),
## data = hads2013n_c)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.7637  -0.8062  -0.8062   1.0801   1.6014
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)      1.31822    0.02971   44.37 <2e-16 ***
## FMTCOST06RELPOVCAT'2 100-150% Poverty' -0.85475    0.03966  -21.55 <2e-16 ***
## FMTCOST06RELPOVCAT'3 150-200% Poverty' -1.08502    0.03797  -28.58 <2e-16 ***
## FMTCOST06RELPOVCAT'4 200%+ Poverty'    -2.27543    0.03174  -71.68 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 81054 on 60095 degrees of freedom
## Residual deviance: 72280 on 60092 degrees of freedom
## AIC: 72288
##
## Number of Fisher Scoring iterations: 4
```

*#This difference appears large enough to suggest that a relationship does exist between FMTCOST06RELPOVCAT and FMTOWNRENT.*

```
plot(table(hads2013n_c$FMTCOST08RELPOVCAT,hads2013n_c$FMTOWNRENT),main='FMTCOST08RELPOVCAT VS FMTOWNRENT',xlab='FMTCOST08RELPOVCAT',ylab='FMTOWNRENT')
```

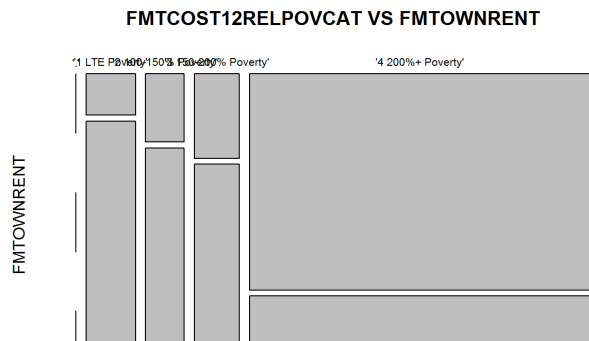


```
summary(glm(FMTOWNRENT ~ FMTCOST08RELPOVCAT,family=binomial(link = "logit"),data=hads2013n_c))
```

```
##
## Call:
## glm(formula = FMTOWNRENT ~ FMTCOST08RELPOVCAT, family = binomial(link = "logit"),
##      data = hads2013n_c)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.8507  -0.7899  -0.7899   0.9974   1.6229
##
## Coefficients:
##              Estimate Std. Error z value
## (Intercept)      1.51372    0.03214   47.09
## FMTCOST08RELPOVCAT'2 100-150% Poverty'  -0.78117    0.04306  -18.14
## FMTCOST08RELPOVCAT'3 150-200% Poverty'  -1.07447    0.04079  -26.34
## FMTCOST08RELPOVCAT'4 200%+ Poverty'     -2.51861    0.03401  -74.06
##              Pr(>|z|)
## (Intercept)      <2e-16 ***
## FMTCOST08RELPOVCAT'2 100-150% Poverty'  <2e-16 ***
## FMTCOST08RELPOVCAT'3 150-200% Poverty'  <2e-16 ***
## FMTCOST08RELPOVCAT'4 200%+ Poverty'     <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 81054  on 60095  degrees of freedom
## Residual deviance: 70152  on 60092  degrees of freedom
## AIC: 70160
##
## Number of Fisher Scoring iterations: 4
```

*#This difference appears large enough to suggest that a relationship does exist between FMTCOST08RELPOVCAT and FMTOWNRENT.*

```
plot(table(hads2013n_c$FMTCOST12RELPOVCAT,hads2013n_c$FMTOWNRENT),main='FMTCOST12RELPOVCAT VS FMTOWNRENT',xlab='FMTCOST12RELPOVCAT',ylab='FMTOWNRENT')
```





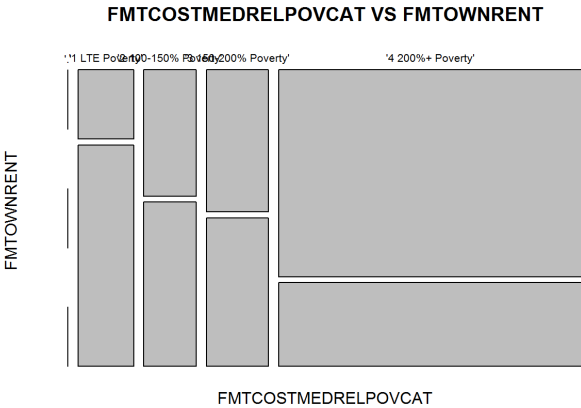
FMTCOST12RELPOVCAT

```
summary(glm(FMTOWNRENT ~ FMTCOST12RELPOVCAT,family=binomial(link = "logit"),data=hads2013n_c))

##
## Call:
## glm(formula = FMTOWNRENT ~ FMTCOST12RELPOVCAT, family = binomial(link = "logit"),
## data = hads2013n_c)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.9759  -0.7695  -0.7695   0.8295   1.6502
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)      1.79887    0.03623   49.65 <2e-16 ***
## FMTCOST12RELPOVCAT'2 100-150% Poverty'  -0.61868    0.04946  -12.51 <2e-16 ***
## FMTCOST12RELPOVCAT'3 150-200% Poverty'  -0.90867    0.04649  -19.55 <2e-16 ***
## FMTCOST12RELPOVCAT'4 200%+ Poverty'    -2.86438    0.03787  -75.64 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 81054  on 60095  degrees of freedom
## Residual deviance: 66548  on 60092  degrees of freedom
## AIC: 66556
##
## Number of Fisher Scoring iterations: 4

#This difference appears large enough to suggest that a relationship does
#exist between FMTCOST12RELPOVCAT and FMTOWNRENT.

plot(table(hads2013n_c$FMTCOSTMEDRELPOVCAT,hads2013n_c$FMTOWNRENT),main='FMTCOSTMEDRELPOVCAT VS FMTOWNRENT',xlab='FMTCOSTMED
RELPOVCAT',ylab='FMTOWNRENT')
```



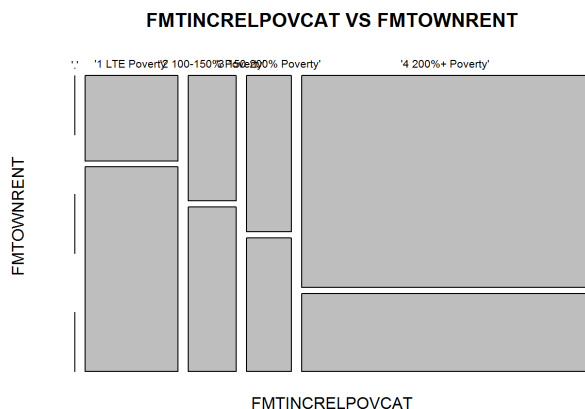
```
summary(glm(FMTOWNRENT ~ FMTCOSTMEDRELPOVCAT,family=binomial(link = "logit"),data=hads2013n_c))

##
## Call:
## glm(formula = FMTOWNRENT ~ FMTCOSTMEDRELPOVCAT, family = binomial(link = "logit"),
## data = hads2013n_c)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.6944  -0.8237  -0.8237   1.1583   1.5785
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)      1.16367    0.02798   41.59 <2e-16 ***
## FMTCOSTMEDRELPOVCAT'2 100-150% Poverty'  -0.89925    0.03737  -24.06 <2e-16 ***
## FMTCOSTMEDRELPOVCAT'3 150-200% Poverty'  -0.89925    0.03737  -24.06 <2e-16 ***
## FMTCOSTMEDRELPOVCAT'4 200%+ Poverty'    -2.86438    0.03787  -75.64 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 81054  on 60095  degrees of freedom
## Residual deviance: 66548  on 60092  degrees of freedom
## AIC: 66556
##
## Number of Fisher Scoring iterations: 4
```

```
## FMTCOSTMEDRELPOVCAT'3 150-200% Poverty' -1.11845 0.03590 -31.16
## FMTCOSTMEDRELPOVCAT'4 200%+ Poverty' -2.07029 0.03016 -68.64
## Pr(>|z|)
## (Intercept) <2e-16 ***
## FMTCOSTMEDRELPOVCAT'2 100-150% Poverty' <2e-16 ***
## FMTCOSTMEDRELPOVCAT'3 150-200% Poverty' <2e-16 ***
## FMTCOSTMEDRELPOVCAT'4 200%+ Poverty' <2e-16 ***
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 81054 on 60095 degrees of freedom
## Residual deviance: 73991 on 60092 degrees of freedom
## AIC: 73999
##
## Number of Fisher Scoring iterations: 4
```

*#This difference appears large enough to suggest that a relationship does  
#exist between FMTCOSTMEDRELPOVCAT and FMTOWNRENT.*

```
plot(table(hads2013n_c$FMTINCRELPOVCAT,hads2013n_c$FMTOWNRENT),main='FMTINCRELPOVCAT VS FMTOWNRENT',xlab='FMTINCRELPOVCAT',y  
lab='FMTOWNRENT')
```

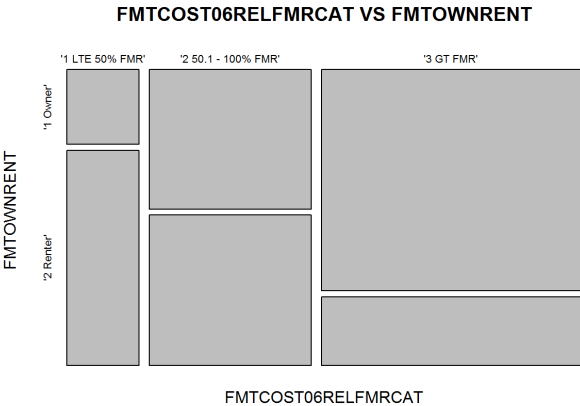


```
summary(glm(FMTOWNRENT ~ FMTINCRELPOVCAT,family=binomial(link = "logit"),data=hads2013n_c))
```

```
##
## Call:
## glm(formula = FMTOWNRENT ~ FMTINCRELPOVCAT, family = binomial(link = "logit"),
## data = hads2013n_c)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.5638  -0.7921  -0.7921   1.0638   1.6200
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)      0.87405    0.02019   43.29 <2e-16
## FMTINCRELPOVCAT'2 100-150% Poverty' -0.60074    0.03283  -18.30 <2e-16
## FMTINCRELPOVCAT'3 150-200% Poverty' -1.03010    0.03337  -30.87 <2e-16
## FMTINCRELPOVCAT'4 200%+ Poverty' -1.87251    0.02338  -80.07 <2e-16
##
## (Intercept) ***
## FMTINCRELPOVCAT'2 100-150% Poverty' ***
## FMTINCRELPOVCAT'3 150-200% Poverty' ***
## FMTINCRELPOVCAT'4 200%+ Poverty' ***
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 81054 on 60095 degrees of freedom
## Residual deviance: 73029 on 60092 degrees of freedom
## AIC: 73037
##
## Number of Fisher Scoring iterations: 4
```

*#This difference appears large enough to suggest that a relationship does  
#exist between FMTINCRELPOVCAT and FMTOWNRENT.*

```
plot(table(hads2013n_c$FMTCOST06RELFMRCAT,hads2013n_c$FMTOWNRENT),main='FMTCOST06RELFMRCAT VS FMTOWNRENT',xlab='FMTCOST06RELFMRCAT',ylab='FMTOWNRENT')
```

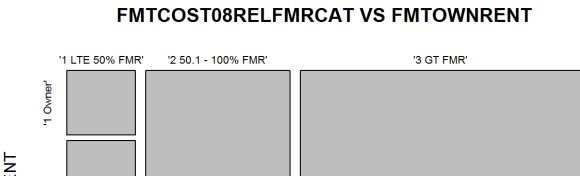


```
summary(glm(FMTOWNRENT ~ FMTCOST06RELFMRCAT,family=binomial(link = "logit"),data=hads2013n_c))
```

```
##
## Call:
## glm(formula = FMTOWNRENT ~ FMTCOST06RELFMRCAT, family = binomial(link = "logit"),
##      data = hads2013n_c)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.6473  -0.7355  -0.7355   1.1457   1.6970
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)      1.05901    0.02442   43.36  <2e-16
## FMTCOST06RELFMRCAT'2 50.1 - 100% FMR' -0.98401    0.02827  -34.81  <2e-16
## FMTCOST06RELFMRCAT'3 GT FMR'      -2.22843    0.02778  -80.22  <2e-16
##
## (Intercept)          ***
## FMTCOST06RELFMRCAT'2 50.1 - 100% FMR' ***
## FMTCOST06RELFMRCAT'3 GT FMR'          ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 81054  on 60095  degrees of freedom
## Residual deviance: 71941  on 60093  degrees of freedom
## AIC: 71947
##
## Number of Fisher Scoring iterations: 4
```

```
#This difference appears large enough to suggest that a relationship does
#exist between FMTCOST06RELFMRCAT and FMTOWNRENT.

plot(table(hads2013n_c$FMTCOST08RELFMRCAT,hads2013n_c$FMTOWNRENT),main='FMTCOST08RELFMRCAT VS FMTOWNRENT',xlab='FMTCOST08REL
FMRCAT',ylab='FMTOWNRENT')
```







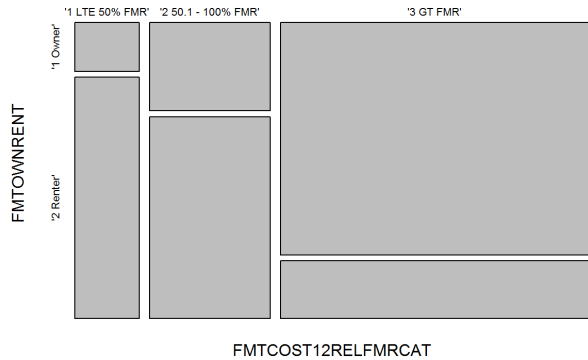
```
summary(glm(FMTOWNRENT ~ FMTCOST08RELFMRCAT,family=binomial(link = "logit"),data=hads2013n_c))
```

```
##
## Call:
## glm(formula = FMTOWNRENT ~ FMTCOST08RELFMRCAT, family = binomial(link = "logit"),
## data = hads2013n_c)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.7367  -0.7031  -0.7031   1.0388   1.7428
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)      1.25777    0.02634   47.75  <2e-16
## FMTCOST08RELFMRCAT'2 50.1 - 100% FMR' -0.92266    0.03046  -30.29  <2e-16
## FMTCOST08RELFMRCAT'3 GT FMR' -2.52919    0.02941  -85.99  <2e-16
##
## (Intercept)          ***
## FMTCOST08RELFMRCAT'2 50.1 - 100% FMR' ***
## FMTCOST08RELFMRCAT'3 GT FMR'      ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 81054  on 60095  degrees of freedom
## Residual deviance: 68635  on 60093  degrees of freedom
## AIC: 68641
##
## Number of Fisher Scoring iterations: 4
```

```
#This difference appears large enough to suggest that a relationship does
#exist between FMTCOST08RELFMRCAT and FMTOWNRENT.

plot(table(hads2013n_c$FMTCOST12RELFMRCAT,hads2013n_c$FMTOWNRENT),main='FMTCOST12RELFMRCAT VS FMTOWNRENT',xlab='FMTCOST12REL
FMRCAT',ylab='FMTOWNRENT')
```

FMTCOST12RELFMRCAT VS FMTOWNRENT



```
summary(glm(FMTOWNRENT ~ FMTCOST12RELFMRCAT,family=binomial(link = "logit"),data=hads2013n_c))
```

```
##
## Call:
## glm(formula = FMTOWNRENT ~ FMTCOST12RELFMRCAT, family = binomial(link = "logit"),
## data = hads2013n_c)
##
```

```

## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.8897  -0.6665  -0.6665   0.8511   1.7964
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)      1.60194    0.03026   52.94  <2e-16
## FMCOST12RELFMRCAT'2 50.1 - 100% FMR' -0.77290    0.03517  -21.98  <2e-16
## FMCOST12RELFMRCAT'3 GT FMR'      -2.99330    0.03291  -90.97  <2e-16
##
## (Intercept)          ***
## FMCOST12RELFMRCAT'2 50.1 - 100% FMR' ***
## FMCOST12RELFMRCAT'3 GT FMR'      ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 81054  on 60095  degrees of freedom
## Residual deviance: 62656  on 60093  degrees of freedom
## AIC: 62662
##
## Number of Fisher Scoring iterations: 4

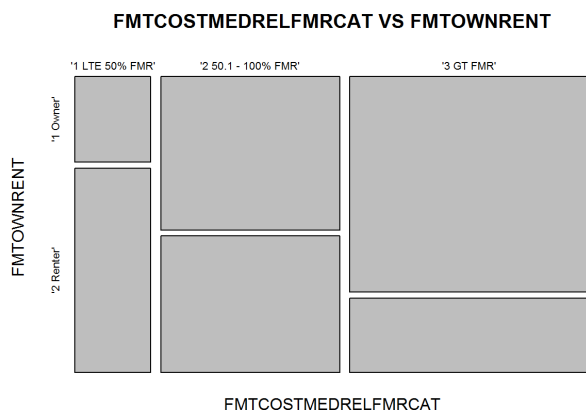
```

*#This difference appears large enough to suggest that a relationship does  
#exist between FMCOST12RELFMRCAT and FMTOWNRENT.*

```

plot(table(hads2013n_c$FMCOSTMEDRELFMRCAT,hads2013n_c$FMTOWNRENT),main='FMCOSTMEDRELFMRCAT VS FMTOWNRENT',xlab='FMCOSTMEDRELFMRCAT',ylab='FMTOWNRENT')

```



```

summary(glm(FMTOWNRENT ~ FMCOSTMEDRELFMRCAT,family=binomial(link = "logit"),data=hads2013n_c))

```

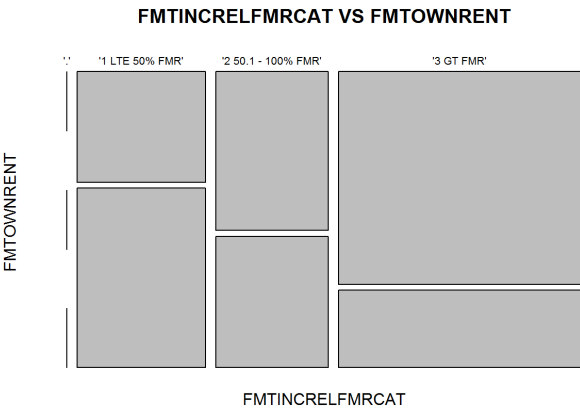
```

##
## Call:
## glm(formula = FMTOWNRENT ~ FMCOSTMEDRELFMRCAT, family = binomial(link = "logit"),
##      data = hads2013n_c)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.5628  -0.7709  -0.7709   1.2270   1.6483
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)      0.87176    0.02282   38.20  <2e-16 ***
## FMCOSTMEDRELFMRCAT'2 50.1 - 100% FMR' -0.98764    0.02656  -37.19  <2e-16 ***
## FMCOSTMEDRELFMRCAT'3 GT FMR'      -1.93313    0.02647  -73.02  <2e-16 ***
##
## (Intercept)          <2e-16 ***
## FMCOSTMEDRELFMRCAT'2 50.1 - 100% FMR' <2e-16 ***
## FMCOSTMEDRELFMRCAT'3 GT FMR'      <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 81054  on 60095  degrees of freedom
## Residual deviance: 74469  on 60093  degrees of freedom
## AIC: 74475
##
## Number of Fisher Scoring iterations: 4

```

```
#This difference appears Large enough to suggest that a relationship does
#exist between FMTCOSTMEDRELFMRCAT and FMTOWNRENT.

plot(table(hads2013n_c$FMTINCRELFMRCAT,hads2013n_c$FMTOWNRENT),main='FMTINCRELFMRCAT VS FMTOWNRENT',xlab='FMTINCRELFMRCAT',y
lab='FMTOWNRENT')
```

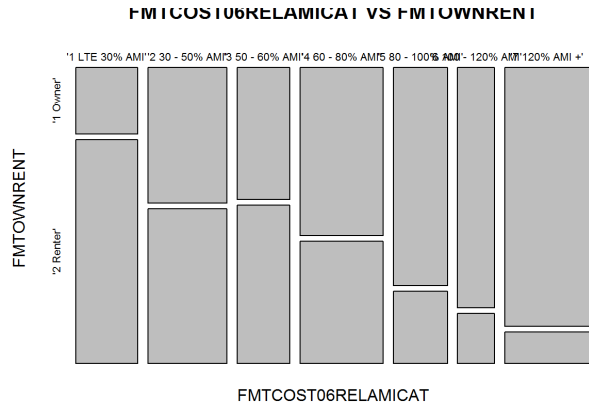


```
summary(glm(FMTOWNRENT ~ FMTINCRELFMRCAT,family=binomial(link = "logit"),data=hads2013n_c))
```

```
##
## Call:
## glm(formula = FMTOWNRENT ~ FMTINCRELFMRCAT, family = binomial(link = "logit"),
## data = hads2013n_c)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.3886  -0.7872  -0.7872   0.9800   1.6264
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)      0.48391    0.01630   29.69  <2e-16
## FMTINCRELFMRCAT'2 50.1 - 100% FMR' -0.67310    0.02353  -28.60  <2e-16
## FMTINCRELFMRCAT'3 GT FMR' -1.49677    0.02087  -71.73  <2e-16
##
## (Intercept)          ***
## FMTINCRELFMRCAT'2 50.1 - 100% FMR' ***
## FMTINCRELFMRCAT'3 GT FMR'      ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 81054  on 60095  degrees of freedom
## Residual deviance: 75440  on 60093  degrees of freedom
## AIC: 75446
##
## Number of Fisher Scoring iterations: 4
```

```
#This difference appears Large enough to suggest that a relationship does
#exist between FMTINCRELFMRCAT and FMTOWNRENT.

plot(table(hads2013n_c$FMTCOST06RELAMICAT,hads2013n_c$FMTOWNRENT),main='FMTCOST06RELAMICAT VS FMTOWNRENT',xlab='FMTCOST06REL
AMICAT',ylab='FMTOWNRENT')
```

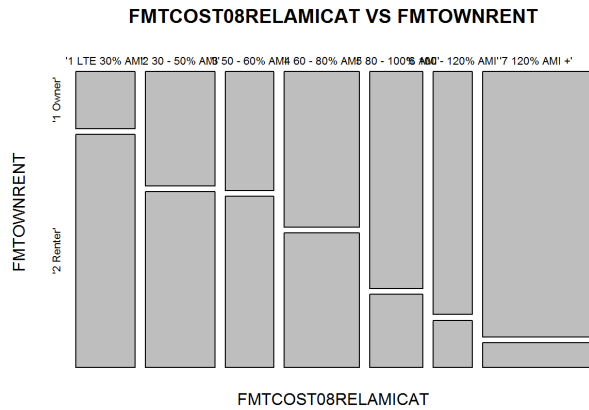


```
summary(glm(FMTOWNNRENT ~ FMTCOST06RELAMICAT,family=binomial(link = "logit"),data=hads2013n_c))
```

```
##
## Call:
## glm(formula = FMTOWNNRENT ~ FMTCOST06RELAMICAT, family = binomial(link = "logit"),
##      data = hads2013n_c)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.7196  -1.0468  -0.4782   1.1002   2.1096
##
## Coefficients:
##              (Intercept)              1.21986      0.02634      46.31      <2e-16
## FMTCOST06RELAMICAT'2 30 - 50% AMI'      -1.08776      0.03280     -33.17      <2e-16
## FMTCOST06RELAMICAT'3 50 - 60% AMI'      -1.03554      0.03567     -29.03      <2e-16
## FMTCOST06RELAMICAT'4 60 - 80% AMI'      -1.53516      0.03264     -47.03      <2e-16
## FMTCOST06RELAMICAT'5 80 - 100% AMI'     -2.32589      0.03791     -61.35      <2e-16
## FMTCOST06RELAMICAT'6 100 - 120% AMI'    -2.78665      0.04613     -60.41      <2e-16
## FMTCOST06RELAMICAT'7 120% AMI +'     -3.33074      0.04015     -82.96      <2e-16
##
## (Intercept)
## FMTCOST06RELAMICAT'2 30 - 50% AMI'
## FMTCOST06RELAMICAT'3 50 - 60% AMI'
## FMTCOST06RELAMICAT'4 60 - 80% AMI'
## FMTCOST06RELAMICAT'5 80 - 100% AMI'
## FMTCOST06RELAMICAT'6 100 - 120% AMI'
## FMTCOST06RELAMICAT'7 120% AMI +'
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 81054  on 60095  degrees of freedom
## Residual deviance: 68266  on 60089  degrees of freedom
## AIC: 68280
##
## Number of Fisher Scoring iterations: 4
```

*#This difference appears large enough to suggest that a relationship does exist between FMTCOST06RELAMICAT and FMTOWNNRENT.*

```
plot(table(hads2013n_c$FMTCOST06RELAMICAT,hads2013n_c$FMTOWNNRENT),main='FMTCOST06RELAMICAT VS FMTOWNNRENT',xlab='FMTCOST06RELAMICAT',ylab='FMTOWNNRENT')
```

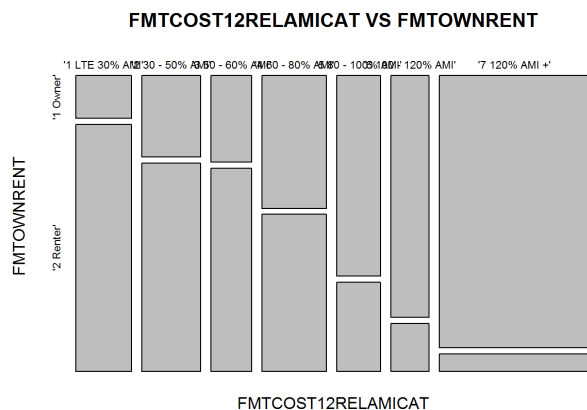


```
summary(glm(FMTOWNRENT ~ FMTCOST08RELAMICAT,family=binomial(link = "logit"),data=hads2013n_c))
```

```
##
## Call:
## glm(formula = FMTOWNRENT ~ FMTCOST08RELAMICAT, family = binomial(link = "logit"),
##      data = hads2013n_c)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.8046  -0.7633  -0.4233   1.0005   2.2168
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)      1.40982    0.02839   49.66 <2e-16
## FMTCOST08RELAMICAT'2 30 - 50% AMI' -0.97843    0.03548  -27.58 <2e-16
## FMTCOST08RELAMICAT'3 50 - 60% AMI' -1.04425    0.03804  -27.45 <2e-16
## FMTCOST08RELAMICAT'4 60 - 80% AMI' -1.55226    0.03475  -44.67 <2e-16
## FMTCOST08RELAMICAT'5 80 - 100% AMI' -2.49399    0.03941  -63.28 <2e-16
## FMTCOST08RELAMICAT'6 100 - 120% AMI' -3.04802    0.04714  -64.65 <2e-16
## FMTCOST08RELAMICAT'7 120% AMI + ' -3.77732    0.04125  -91.57 <2e-16
##
## (Intercept)          ***
## FMTCOST08RELAMICAT'2 30 - 50% AMI' ***
## FMTCOST08RELAMICAT'3 50 - 60% AMI' ***
## FMTCOST08RELAMICAT'4 60 - 80% AMI' ***
## FMTCOST08RELAMICAT'5 80 - 100% AMI' ***
## FMTCOST08RELAMICAT'6 100 - 120% AMI' ***
## FMTCOST08RELAMICAT'7 120% AMI + ' ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 81054  on 60095  degrees of freedom
## Residual deviance: 63701  on 60089  degrees of freedom
## AIC: 63715
##
## Number of Fisher Scoring iterations: 5
```

*#This difference appears Large enough to suggest that a relationship does  
#exist between FMTCOST08RELAMICAT and FMTOWNRENT.*

```
plot(table(hads2013n_c$FMTCOST12RELAMICAT,hads2013n_c$FMTOWNRENT),main='FMTCOST12RELAMICAT VS FMTOWNRENT',xlab='FMTCOST12RELAMICAT',ylab='FMTOWNRENT')
```



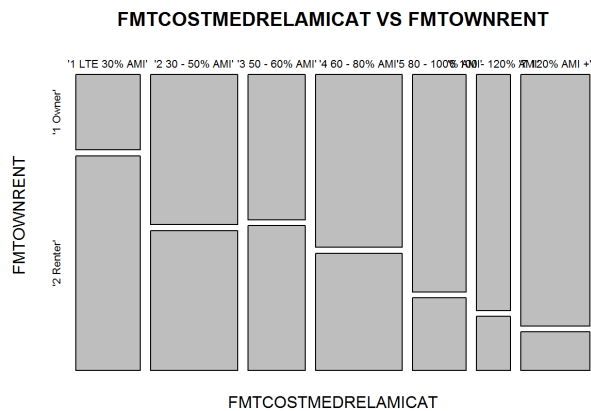
```
summary(glm(FMTOWNRENT ~ FMTCOST12RELAMICAT,family=binomial(link = "logit"),data=hads2013n_c))
```

```
##
## Call:
## glm(formula = FMTOWNRENT ~ FMTCOST12RELAMICAT, family = binomial(link = "logit"),
##      data = hads2013n_c)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.9565  -0.6035  -0.3549   0.8122   2.3648
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)      1.75435    0.03274   53.58 <2e-16
## FMTCOST12RELAMICAT'2 30 - 50% AMI' -0.81452    0.04132  -19.71 <2e-16
## FMTCOST12RELAMICAT'3 50 - 60% AMI' -0.89968    0.04418  -20.36 <2e-16
## FMTCOST12RELAMICAT'4 60 - 80% AMI' -1.58515    0.03927  -40.37 <2e-16
## FMTCOST12RELAMICAT'5 80 - 100% AMI' -2.56342    0.04335  -59.13 <2e-16
## FMTCOST12RELAMICAT'6 100 - 120% AMI' -3.36521    0.04995  -67.37 <2e-16
## FMTCOST12RELAMICAT'7 120% AMI + ' -4.48742    0.04410  -101.77 <2e-16
##
## (Intercept)          ***
```

```
## FMTCOST12RELAMICAT'2 30 - 50% AMI' ***
## FMTCOST12RELAMICAT'3 50 - 60% AMI' ***
## FMTCOST12RELAMICAT'4 60 - 80% AMI' ***
## FMTCOST12RELAMICAT'5 80 - 100% AMI' ***
## FMTCOST12RELAMICAT'6 100 - 120% AMI' ***
## FMTCOST12RELAMICAT'7 120% AMI +' ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 81054  on 60095  degrees of freedom
## Residual deviance: 54846  on 60089  degrees of freedom
## AIC: 54860
##
## Number of Fisher Scoring iterations: 5
```

*#This difference appears Large enough to suggest that a relationship does  
#exist between FMTCOST12RELAMICAT and FMTOWNRENT.*

```
plot(table(hads2013n_c$FMTCOSTMEDRELAMICAT,hads2013n_c$FMTOWNRENT),main='FMTCOSTMEDRELAMICAT VS FMTOWNRENT',xlab='FMTCOSTMEDRELAMICAT',ylab='FMTOWNRENT')
```

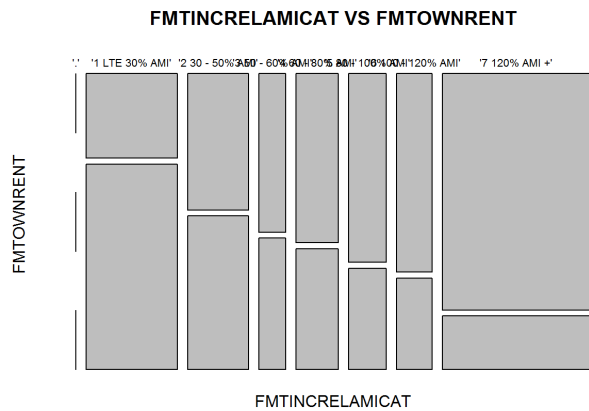


```
summary(glm(FMTOWNRENT ~ FMTCOSTMEDRELAMICAT,family=binomial(link = "logit"),data=hads2013n_c))
```

```
##
## Call:
## glm(formula = FMTOWNRENT ~ FMTCOSTMEDRELAMICAT, family = binomial(link = "logit"),
##      data = hads2013n_c)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.6431  -1.0179  -0.5352   1.1780   2.0071
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)      1.04984    0.02470   42.51  <2e-16
## FMTCOSTMEDRELAMICAT'2 30 - 50% AMI' -1.11957    0.03090  -36.23  <2e-16
## FMTCOSTMEDRELAMICAT'3 50 - 60% AMI' -1.05116    0.03369  -31.20  <2e-16
## FMTCOSTMEDRELAMICAT'4 60 - 80% AMI' -1.43730    0.03116  -46.13  <2e-16
## FMTCOSTMEDRELAMICAT'5 80 - 100% AMI' -2.14641    0.03681  -58.30  <2e-16
## FMTCOSTMEDRELAMICAT'6 100 - 120% AMI' -2.52137    0.04548  -55.44  <2e-16
## FMTCOSTMEDRELAMICAT'7 120% AMI +' -2.92083    0.03944  -74.07  <2e-16
##
## (Intercept)          ***
## FMTCOSTMEDRELAMICAT'2 30 - 50% AMI' ***
## FMTCOSTMEDRELAMICAT'3 50 - 60% AMI' ***
## FMTCOSTMEDRELAMICAT'4 60 - 80% AMI' ***
## FMTCOSTMEDRELAMICAT'5 80 - 100% AMI' ***
## FMTCOSTMEDRELAMICAT'6 100 - 120% AMI' ***
## FMTCOSTMEDRELAMICAT'7 120% AMI +' ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 81054  on 60095  degrees of freedom
## Residual deviance: 71537  on 60089  degrees of freedom
## AIC: 71551
##
## Number of Fisher Scoring iterations: 4
```

*#This difference appears Large enough to suggest that a relationship does  
#exist between FMTCOSTMEDRELAMICAT and FMTOWNRENT.*

```
plot(table(hads2013n_c$FMTINCRELAMICAT,hads2013n_c$FMTOWNRENT),main='FMTINCRELAMICAT VS FMTOWNRENT',xlab='FMTINCRELAMICAT',ylab='FMTOWNRENT')
```

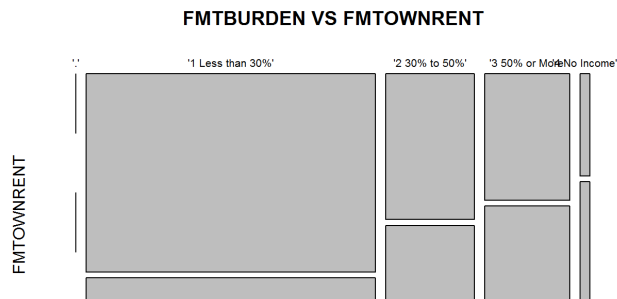


```
summary(glm(FMTOWNRENT ~ FMTINCRELAMICAT,family=binomial(link = "logit"),data=hads2013n_c))
```

```
##
## Call:
## glm(formula = FMTOWNRENT ~ FMTINCRELAMICAT, family = binomial(link = "logit"),
##      data = hads2013n_c)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.5705  -0.9266  -0.6390   1.1274   1.8380
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)      0.88880    0.01978   44.94  <2e-16 ***
## FMTINCRELAMICAT'2 30 - 50% AMI' -0.77005    0.02960  -26.02  <2e-16 ***
## FMTINCRELAMICAT'3 50 - 60% AMI'  -1.07654    0.03867  -27.84  <2e-16 ***
## FMTINCRELAMICAT'4 60 - 80% AMI'  -1.22421    0.03321  -36.86  <2e-16 ***
## FMTINCRELAMICAT'5 80 - 100% AMI' -1.51210    0.03537  -42.75  <2e-16 ***
## FMTINCRELAMICAT'6 100 - 120% AMI' -1.66409    0.03658  -45.49  <2e-16 ***
## FMTINCRELAMICAT'7 120% AMI +'   -2.37384    0.02689  -88.26  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 81054  on 60095  degrees of freedom
## Residual deviance: 71120  on 60089  degrees of freedom
## AIC: 71134
##
## Number of Fisher Scoring iterations: 4
```

```
#This difference appears large enough to suggest that a relationship does
#exist between FMTINCRELAMICAT and FMTOWNRENT.

plot(table(hads2013n_c$FMTBURDEN,hads2013n_c$FMTOWNRENT),main='FMTBURDEN VS FMTOWNRENT',xlab='FMTBURDEN',ylab='FMTOWNRENT')
```





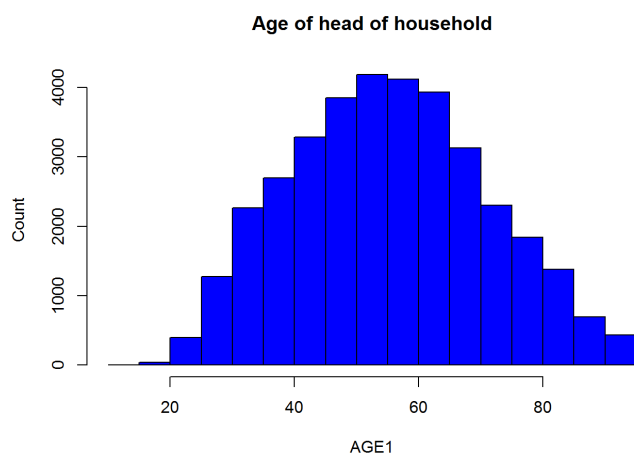
FMTBURDEN

```
summary(glm(FMTOWNRENT ~ FMTBURDEN,family=binomial(link = "logit"),data=hads2013n_c))
```

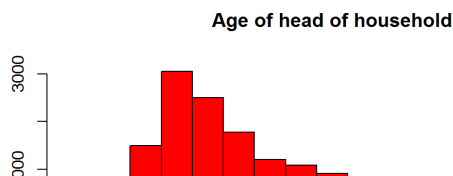
```
##
## Call:
## glm(formula = FMTOWNRENT ~ FMTBURDEN, family = binomial(link = "logit"),
##      data = hads2013n_c)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.4471  -0.8755  -0.8755   1.1799   1.5130
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    -0.76133    0.01119  -68.01  <2e-16 ***
## FMTBURDEN'2 30% to 50%'  0.75549    0.02190   34.50  <2e-16 ***
## FMTBURDEN'3 50% or More'  1.02582    0.02239   45.81  <2e-16 ***
## FMTBURDEN'4 No Income'  1.37601    0.06131   22.44  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 81054  on 60095  degrees of freedom
## Residual deviance: 78049  on 60092  degrees of freedom
## AIC: 78057
##
## Number of Fisher Scoring iterations: 4
```

```
#This difference appears Large enough to suggest that a relationship does
#exist between FMTBURDEN and FMTOWNRENT.
```

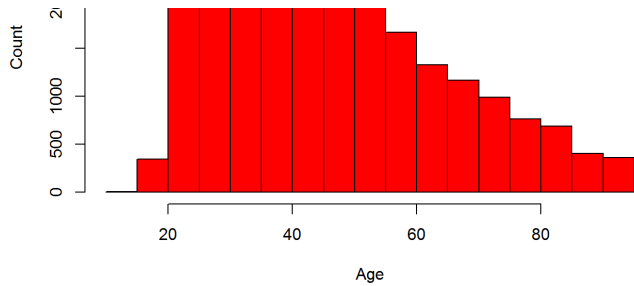
```
##
##
## For Category VS Numeric Variable
hist(hads2013n_c$AGE1[which(hads2013n_c$FMTOWNRENT == "1 Owner")],main= 'Age of head of household', xlab='AGE1', ylab='Count',col = 'blue')
```



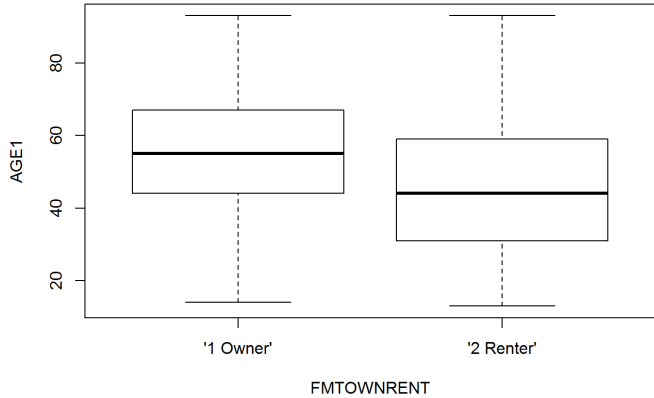
```
hist(hads2013n_c$AGE1[which(hads2013n_c$FMTOWNRENT == "2 Renter")],main= 'Age of head of household', xlab='Age', ylab='Count',col = 'red')
```





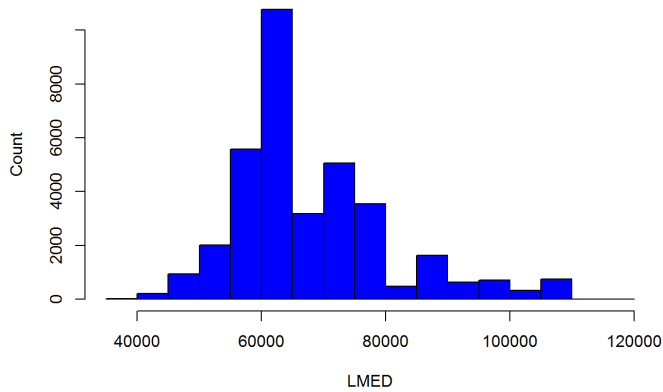


```
plot(hads2013n_c$FMTOWNRENT,hads2013n_c$AGE1,xlab='FMTOWNRENT',ylab='AGE1')
```

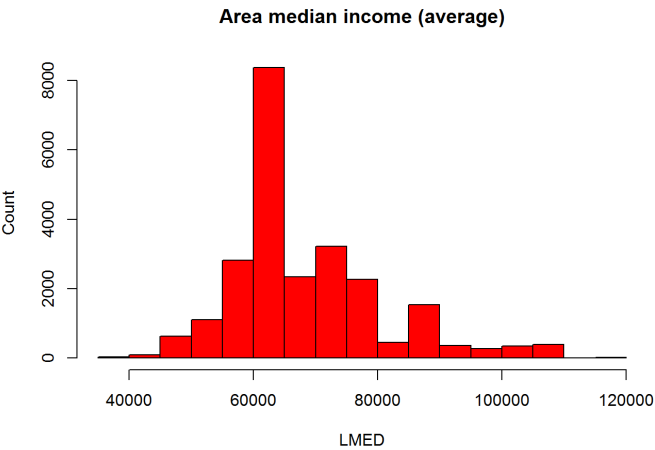


```
hist(hads2013n_c$LMED[which(hads2013n_c$FMTOWNRENT == "1 Owner")],main= 'Area median income (average)', xlab='LMED', ylab='Count',col='blue')
```

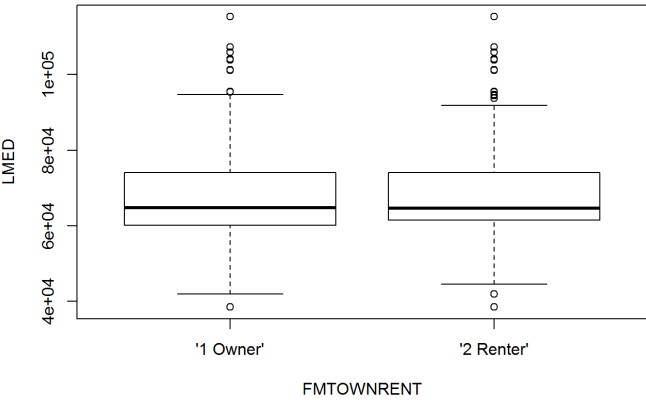
Area median income (average)



```
hist(hads2013n_c$LMED[which(hads2013n_c$FMTOWNRENT == "2 Renter")],main= 'Area median income (average)', xlab='LMED', ylab='Count',col='red')
```

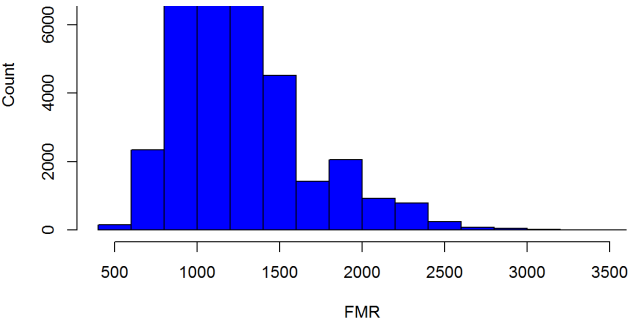


```
plot(hads2013n_c$FMTOWNRENT,hads2013n_c$LMED,xlab='FMTOWNRENT',ylab='LMED')
```



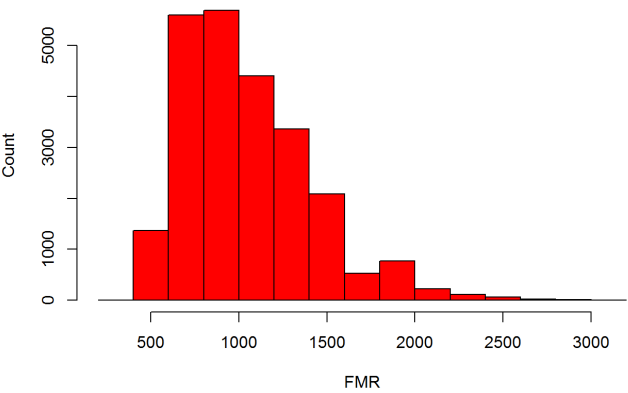
```
hist(hads2013n_c$FMR[which(hads2013n_c$FMTOWNRENT == "1 Owner")],main= 'Fair market rent (average)', xlab='FMR', ylab='Count',col = 'blue')
```



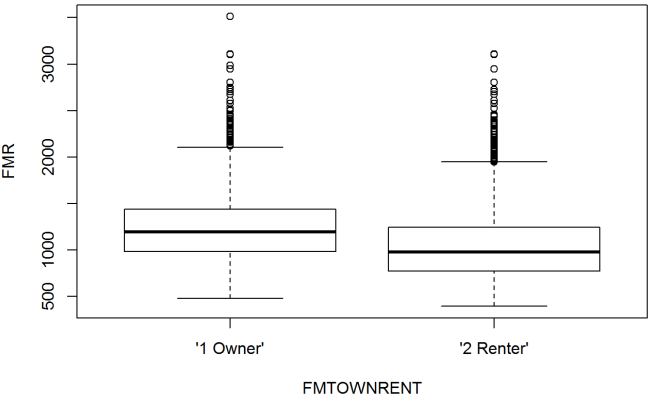


```
hist(hads2013n_c$FMR[which(hads2013n_c$FMTOWNRENT == "'2 Renter'")],main= 'Fair market rent (average)', xlab='FMR', ylab='Count',col = 'red')
```

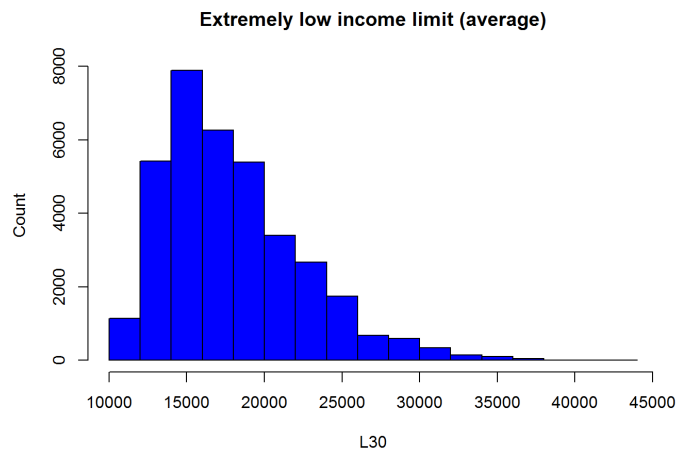
Fair market rent (average)



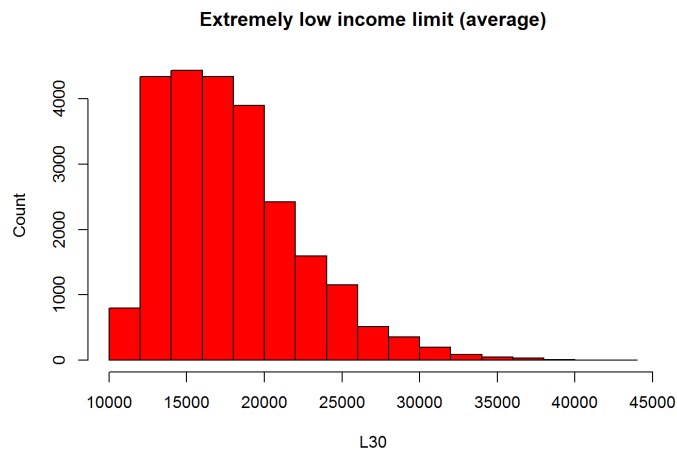
```
plot(hads2013n_c$FMTOWNRENT,hads2013n_c$FMR,xlab='FMTOWNRENT',ylab='FMR')
```



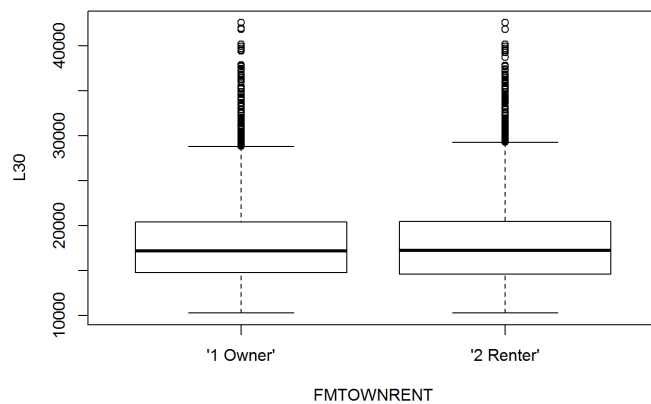
```
hist(hads2013n_c$L30[which(hads2013n_c$FMTOWNRENT == "'1 Owner'")],main= 'Extremely low income limit (average)', xlab='L30', ylab='Count',col = 'blue')
```



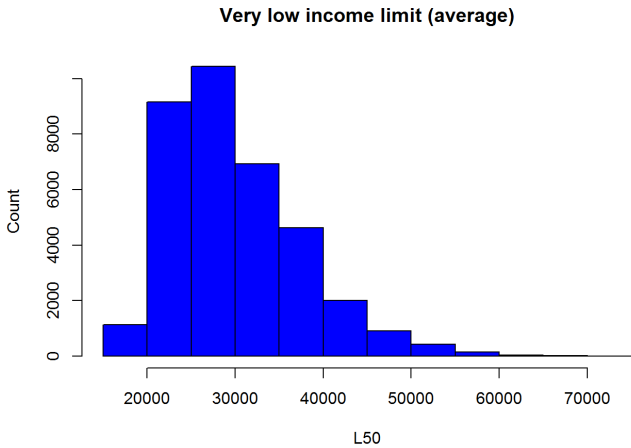
```
hist(hads2013n_c$L30[which(hads2013n_c$FMTOWNRENT == "'2 Renter'")],main= 'Extremely low income limit (average)', xlab='L30',
ylab='Count',col = 'red')
```



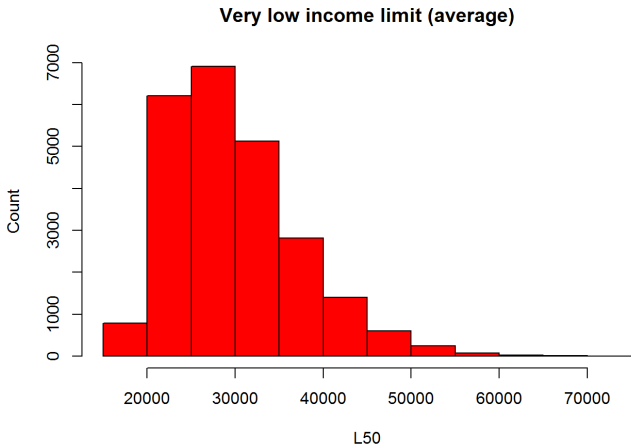
```
plot(hads2013n_c$FMTOWNRENT,hads2013n_c$L30,xlab='FMTOWNRENT',ylab='L30')
```



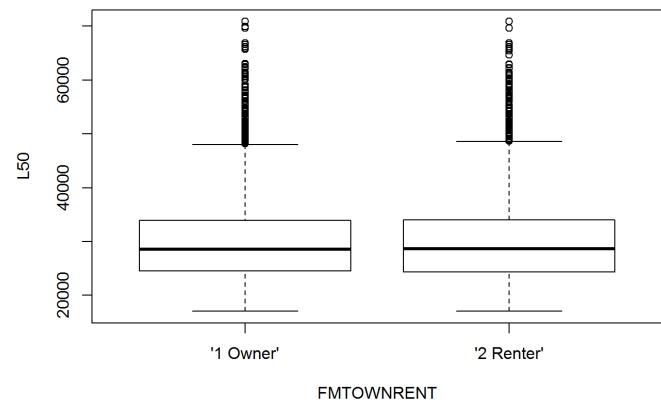
```
hist(hads2013n_c$L50[which(hads2013n_c$FMTOWNRENT == "1 Owner")],main= 'Very low income limit (average)', xlab='L50',
ylab='Count',col = 'blue')
```



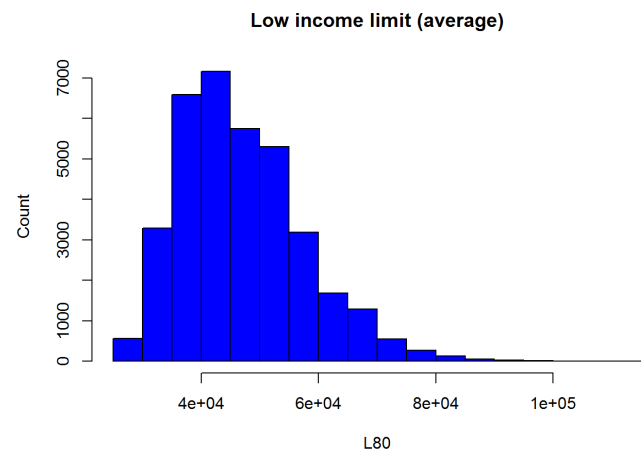
```
hist(hads2013n_c$L50[which(hads2013n_c$FMTOWNRENT == "2 Renter")],main= 'Very low income limit (average)', xlab='L50',
ylab='Count',col = 'red')
```



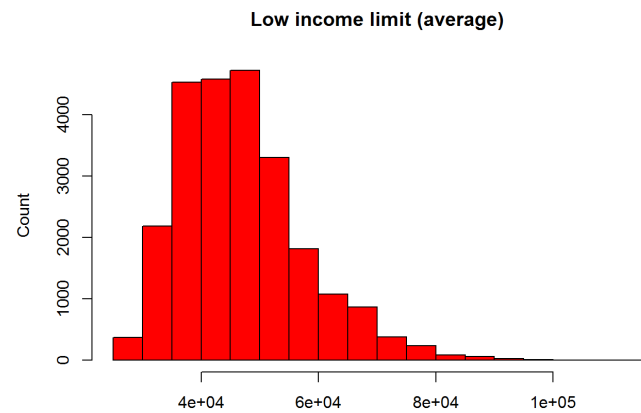
```
plot(hads2013n_c$FMTOWNRENT,hads2013n_c$L50,xlab='FMTOWNRENT',ylab='L50')
```



```
hist(hads2013n_c$L80[which(hads2013n_c$FMTOWNRENT == "1 Owner")],main= 'Low income limit (average)', xlab='L80', ylab='Count',col = 'blue')
```

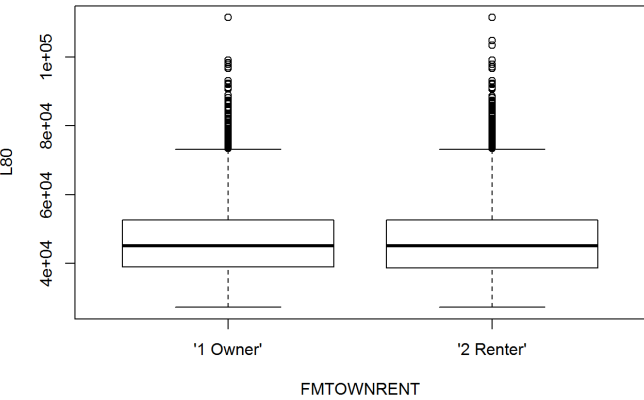


```
hist(hads2013n_c$L80[which(hads2013n_c$FMTOWNRENT == "2 Renter")],main= 'Low income limit (average)', xlab='L80', ylab='Count',col = 'red')
```

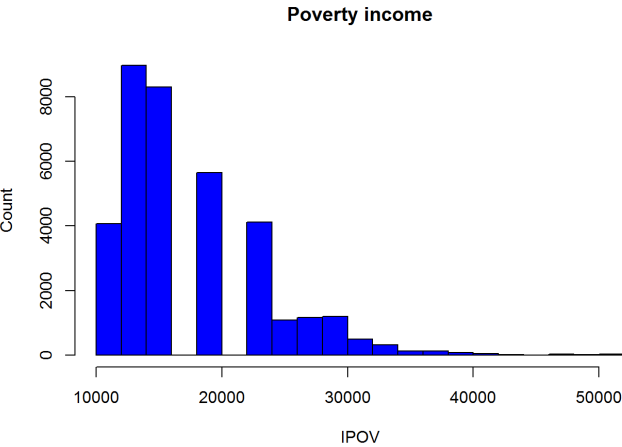


L80

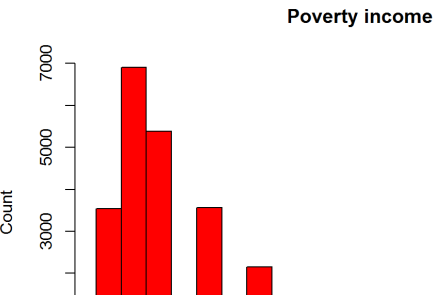
```
plot(hads2013n_c$FMTOWNRENT,hads2013n_c$L80,xlab='FMTOWNRENT',ylab='L80')
```

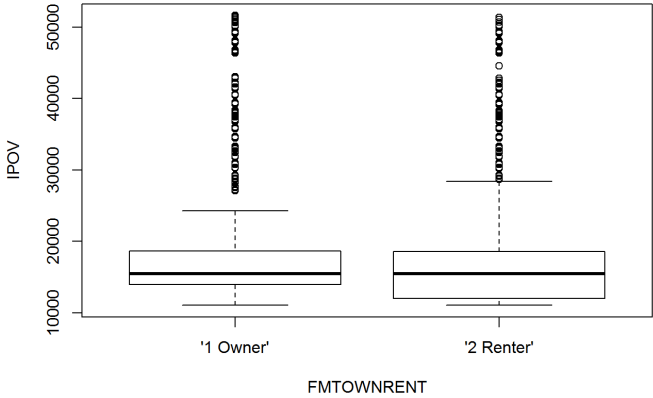
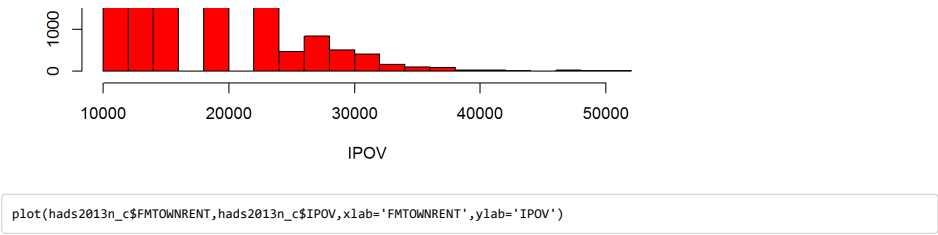


```
hist(hads2013n_c$IPOV[which(hads2013n_c$FMTOWNRENT == "1 Owner")],main= 'Poverty income', xlab='IPOV', ylab='Count',col ='blue')
```

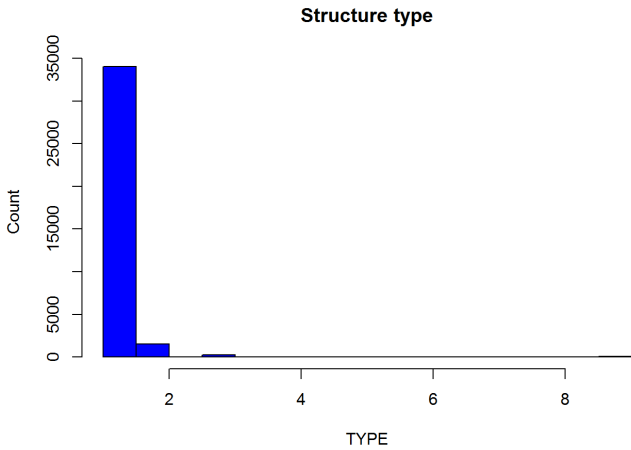


```
hist(hads2013n_c$IPOV[which(hads2013n_c$FMTOWNRENT == "2 Renter")],main= 'Poverty income', xlab='IPOV', ylab='Count',col ='red')
```



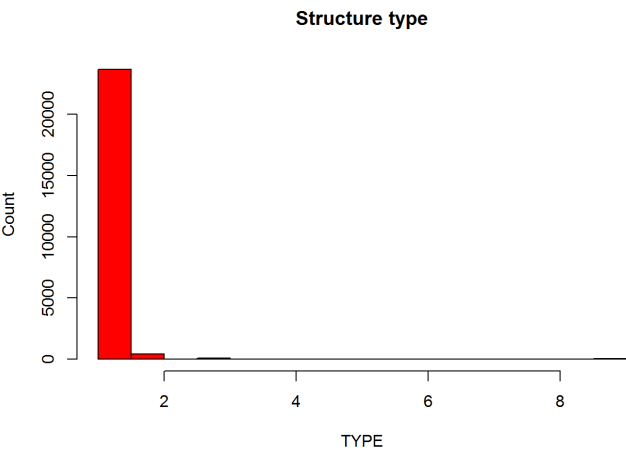


```
hist(hads2013n_c$TYPE[which(hads2013n_c$FMTOWNRENT == "1 Owner")],main= 'Structure type', xlab='TYPE', ylab='Count',col = 'blue')
```

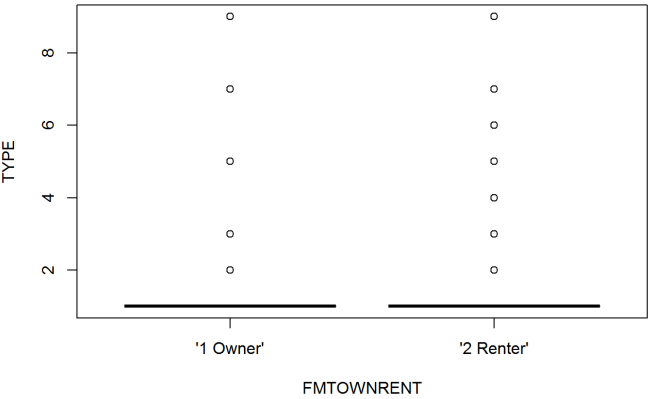


```
hist(hads2013n_c$TYPE[which(hads2013n_c$FMTOWNRENT == "2 Renter")],main= 'Structure type', xlab='TYPE', ylab='Count',col = 'red')
```

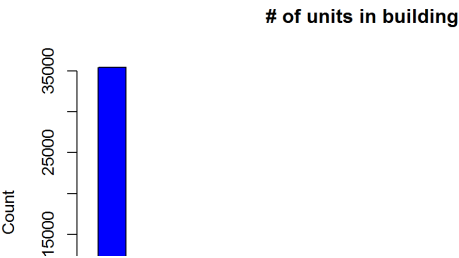


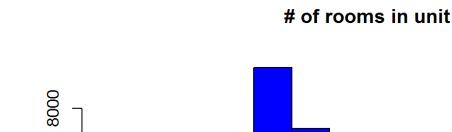
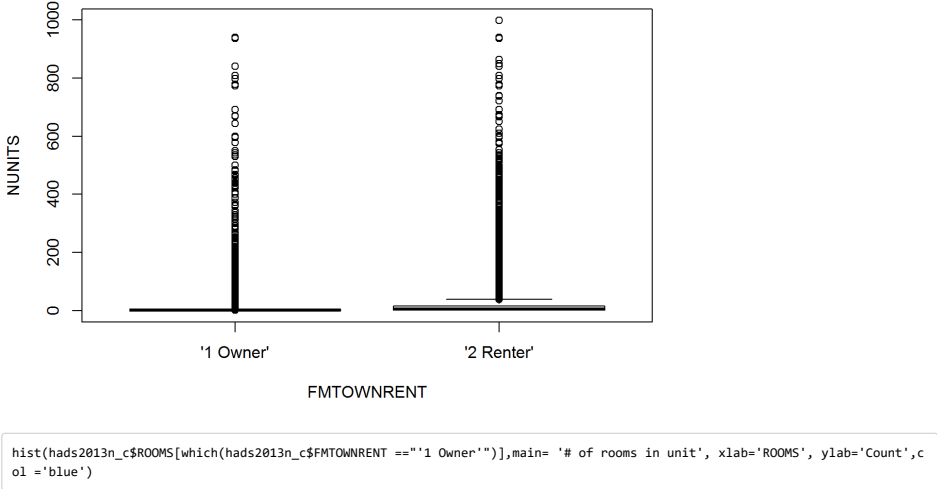
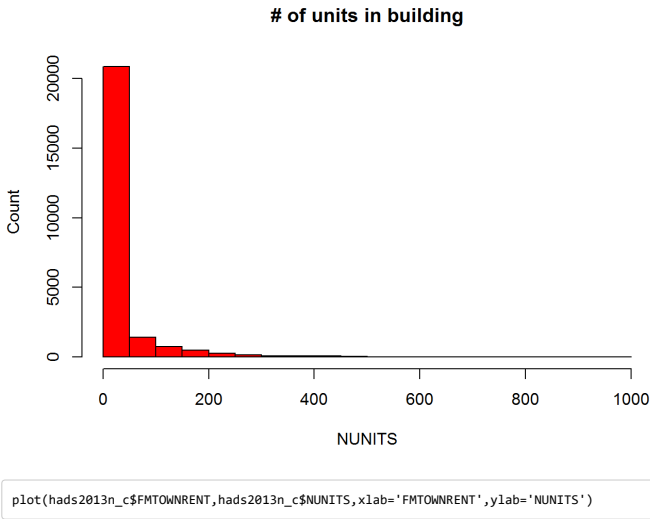
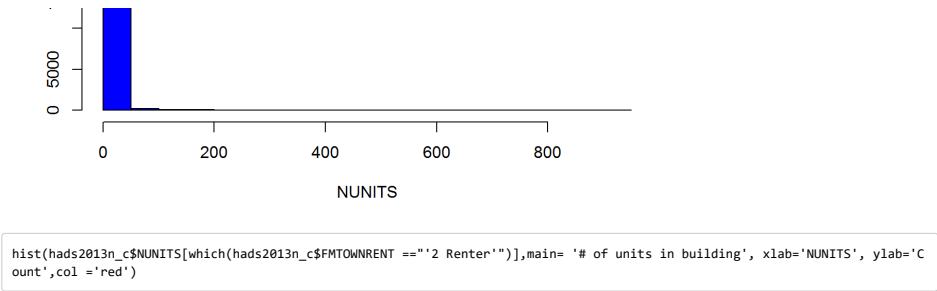


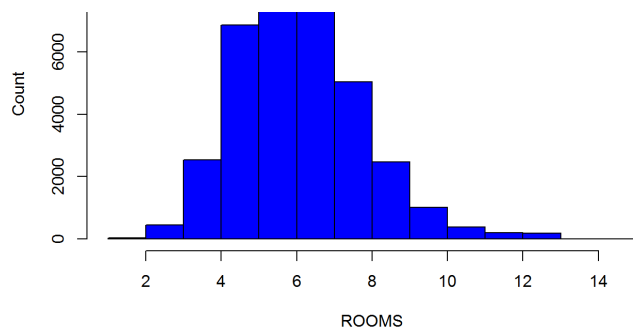
```
plot(hads2013n_c$FMTOWNRENT,hads2013n_c$TYPE,xlab='FMTOWNRENT',ylab='TYPE')
```



```
hist(hads2013n_c$NUNITS[which(hads2013n_c$FMTOWNRENT == "1 Owner")],main= '# of units in building', xlab='NUNITS', ylab='Count',col = 'blue')
```

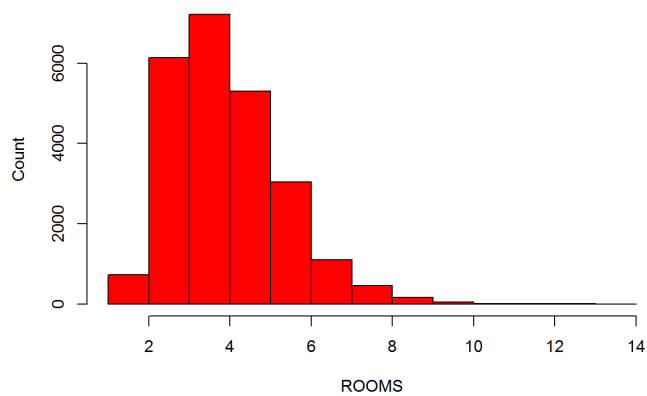




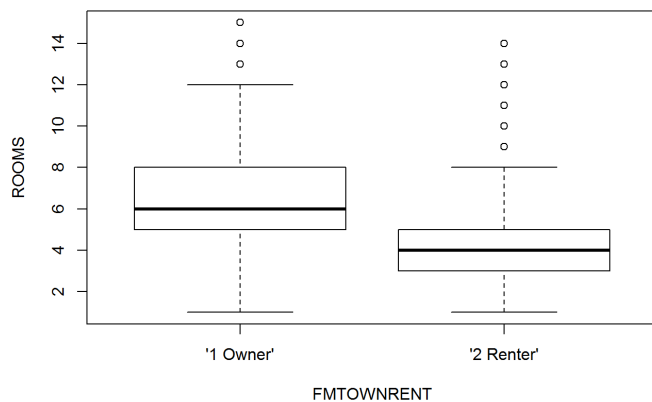


```
hist(hads2013n_c$ROOMS[which(hads2013n_c$FMTOWNRENT == "'2 Renter'")],main= '# of rooms in unit', xlab='ROOMS',
ylab='Count',col='red')
```

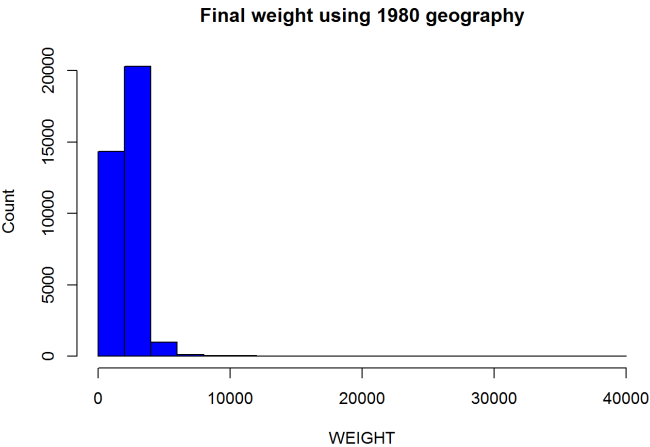
# of rooms in unit



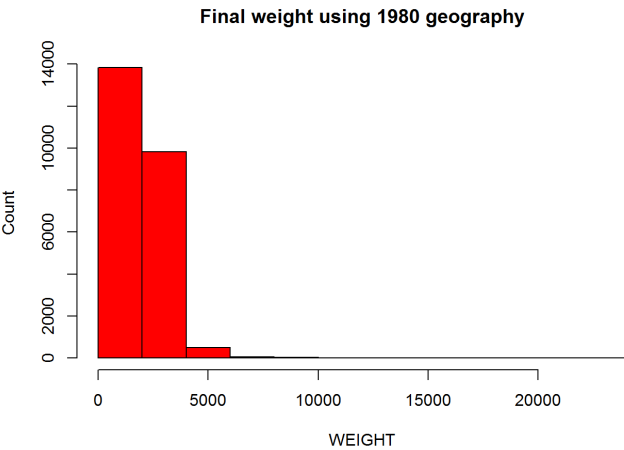
```
plot(hads2013n_c$FMTOWNRENT,hads2013n_c$ROOMS,xlab='FMTOWNRENT',ylab='ROOMS')
```



```
hist(hads2013n_c$WEIGHT[which(hads2013n_c$FMTOWNRENT == "'1 Owner'")],main= 'Final weight using 1980 geography', xlab='WEIGH
T', ylab='Count',col='blue')
```

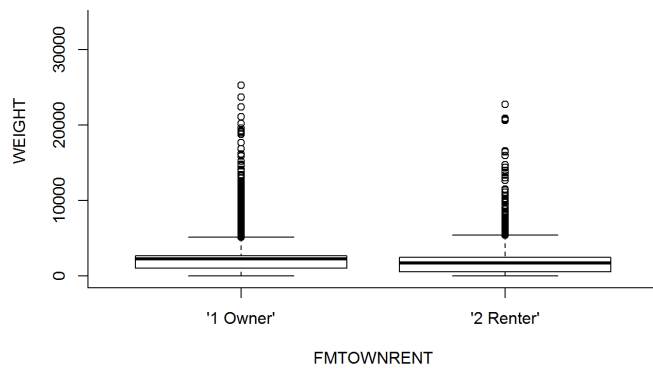


```
hist(hads2013n_c$WEIGHT[which(hads2013n_c$FMTOWNRENT == "'2 Renter'")],main= 'Final weight using 1980 geography', xlab='WEIGHT', ylab='Count',col = 'red')
```



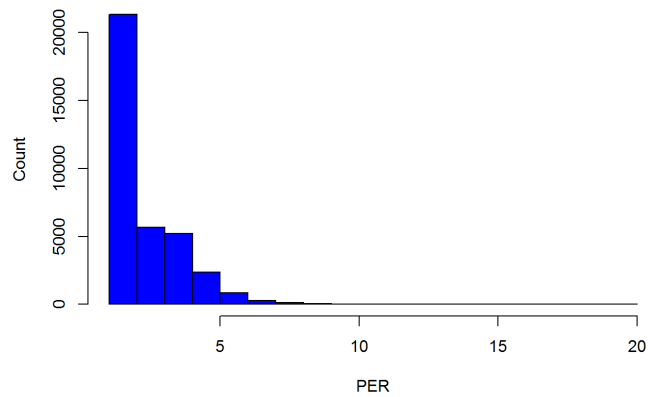
```
plot(hads2013n_c$FMTOWNRENT,hads2013n_c$WEIGHT,xlab='FMTOWNRENT',ylab='WEIGHT')
```





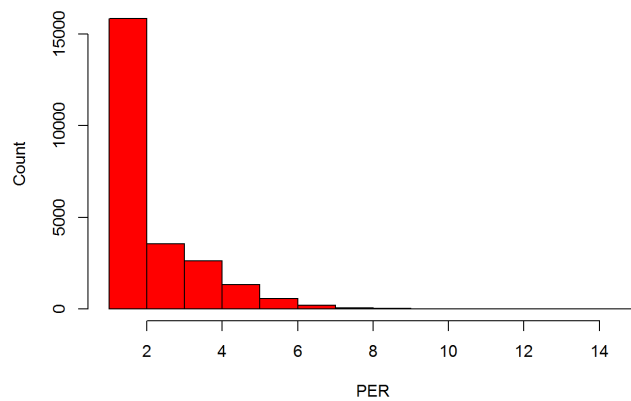
```
hist(hads2013n_c$PER[which(hads2013n_c$FMTOWNRENT == "1 Owner")],main= '# of persons in household', xlab='PER', ylab='Count',col = 'blue')
```

**# of persons in household**

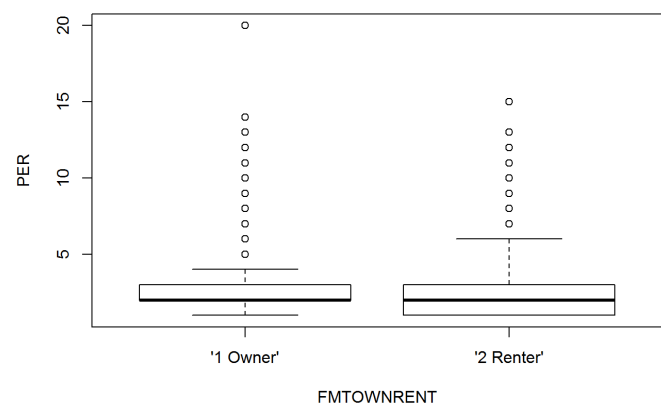


```
hist(hads2013n_c$PER[which(hads2013n_c$FMTOWNRENT == "2 Renter")],main= '# of persons in household', xlab='PER', ylab='Count',col = 'red')
```

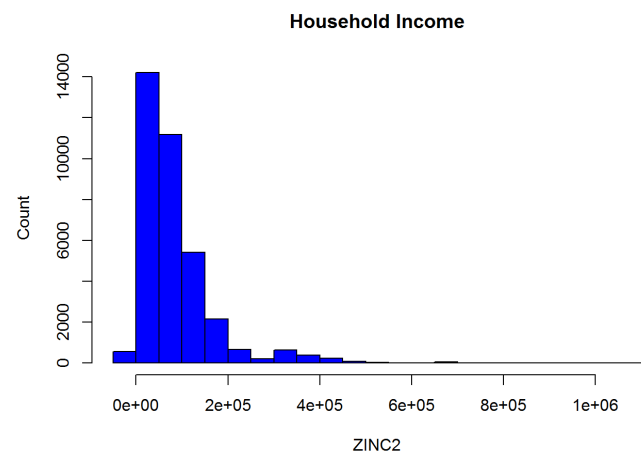
**# of persons in household**



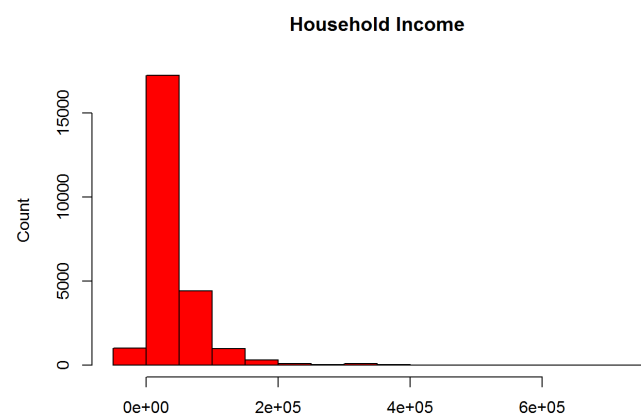
```
plot(hads2013n_c$FMTOWNRENT,hads2013n_c$PER,xlab='FMTOWNRENT',ylab='PER')
```



```
hist(hads2013n_c$ZINC2[which(hads2013n_c$FMTOWNRENT == "'1 Owner'")],main= 'Household Income', xlab='ZINC2', ylab='Count',col
='blue')
```

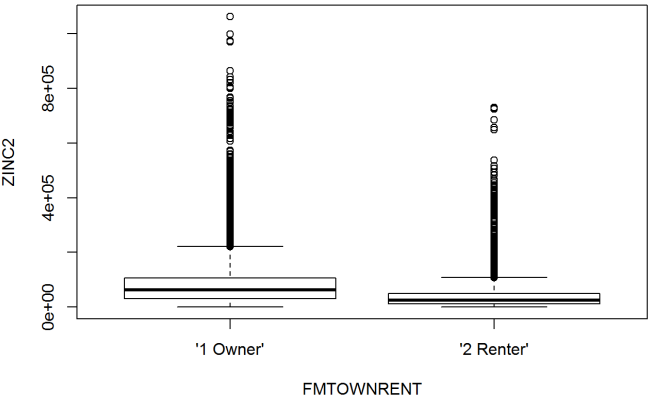


```
hist(hads2013n_c$ZINC2[which(hads2013n_c$FMTOWNRENT == "'2 Renter'")],main= 'Household Income', xlab='ZINC2', ylab='Count',co
l = 'red')
```



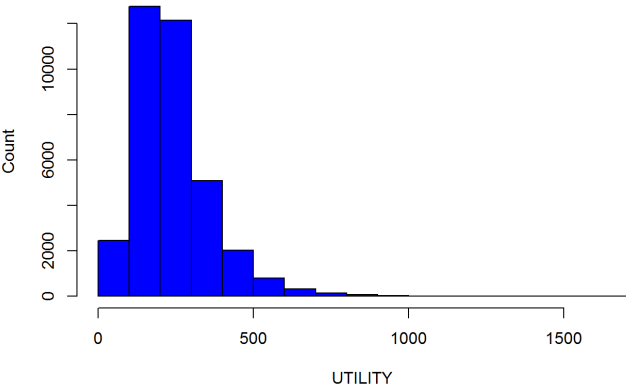
ZINC2

```
plot(hads2013n_c$FMTOWNRENT,hads2013n_c$ZINC2,xlab='FMTOWNRENT',ylab='ZINC2')
```

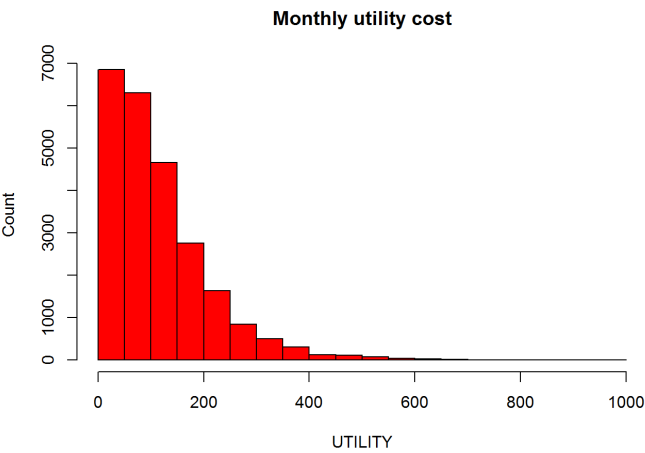


```
hist(hads2013n_c$UTILITY[which(hads2013n_c$FMTOWNRENT == "'1 Owner'")],main= 'Monthly utility cost', xlab='UTILITY', ylab='Count',col = 'blue')
```

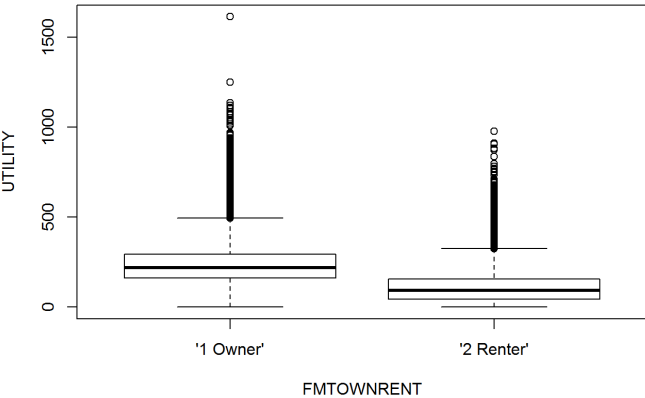
Monthly utility cost



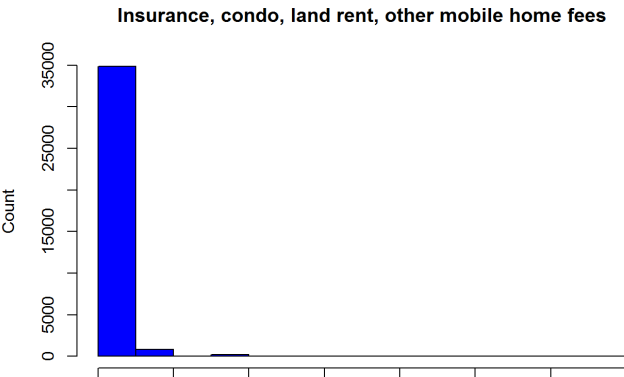
```
hist(hads2013n_c$UTILITY[which(hads2013n_c$FMTOWNRENT == "'2 Renter'")],main= 'Monthly utility cost', xlab='UTILITY', ylab='Count',col = 'red')
```



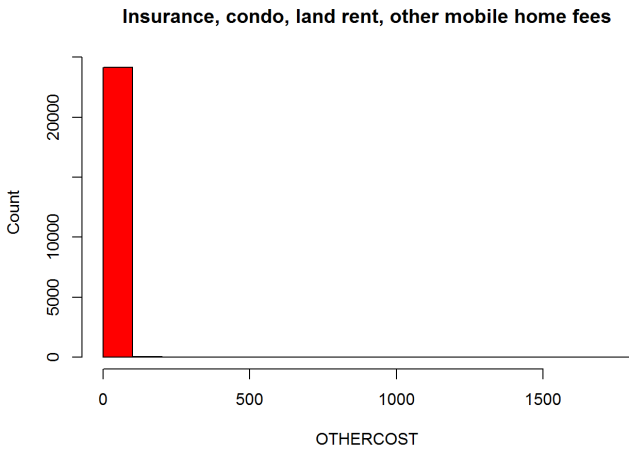
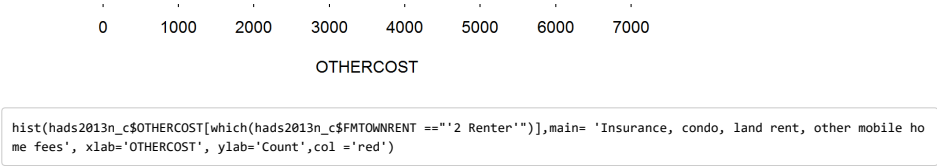
```
plot(hads2013n_c$FMTOWNRENT,hads2013n_c$UTILITY,xlab='FMTOWNRENT',ylab='UTILITY')
```



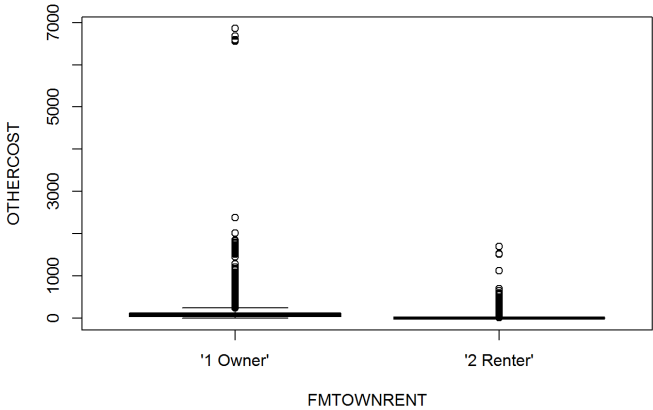
```
hist(hads2013n_c$OTHERCOST[which(hads2013n_c$FMTOWNRENT == "'1 Owner'")],main= 'Insurance, condo, land rent, other mobile home fees', xlab='OTHERCOST', ylab='Count',col ='blue')
```



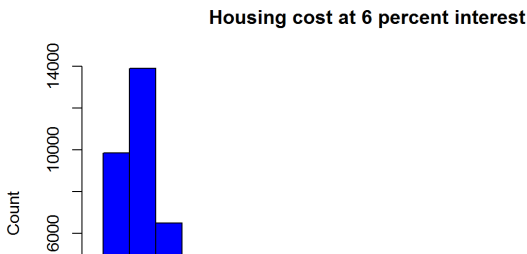


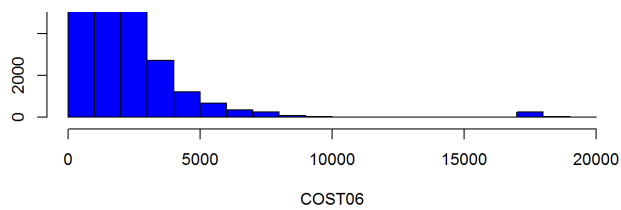


```
plot(hads2013n_c$FMTOWNRENT,hads2013n_c$OTHERCOST,xlab='FMTOWNRENT',ylab='OTHERCOST')
```



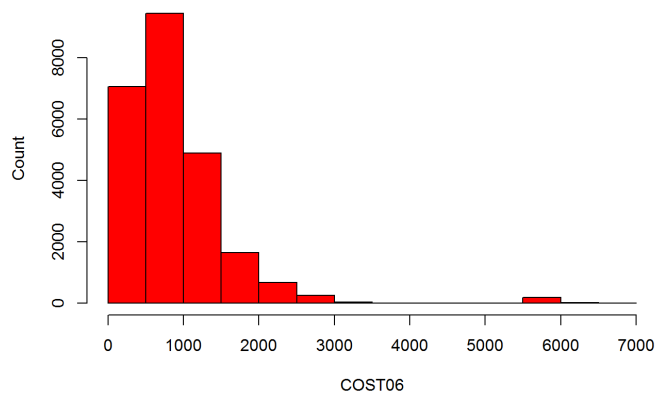
```
hist(hads2013n_c$COST06[which(hads2013n_c$FMTOWNRENT == "'1 Owner'")],main= 'Housing cost at 6 percent interest', xlab='COST06', ylab='Count',col = 'blue')
```



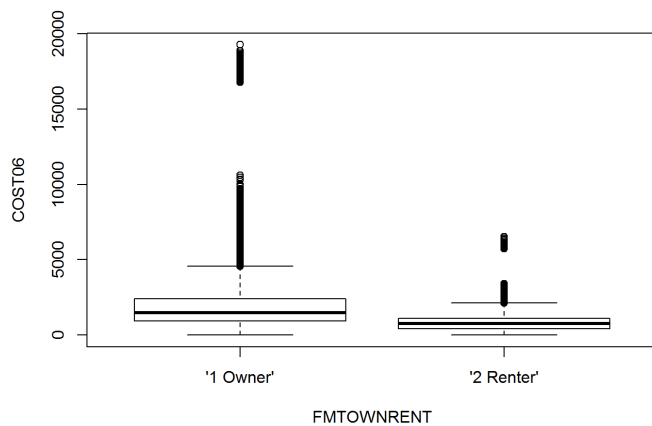


```
hist(hads2013n_c$COST06[which(hads2013n_c$FMTOWNRENT == "'2 Renter'")],main= 'Housing cost at 6 percent interest', xlab='COST06', ylab='Count',col = 'red')
```

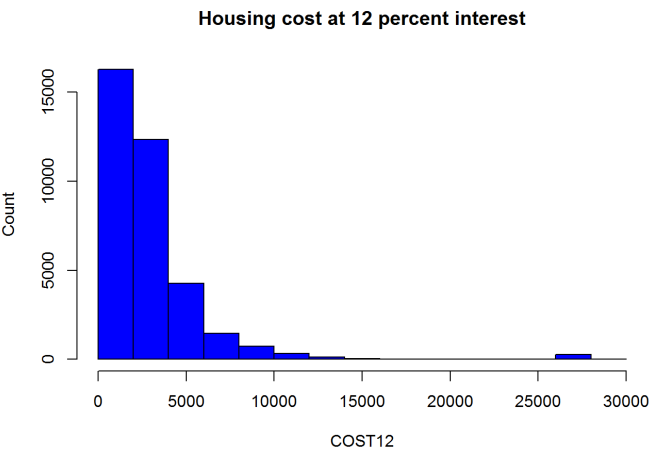
Housing cost at 6 percent interest



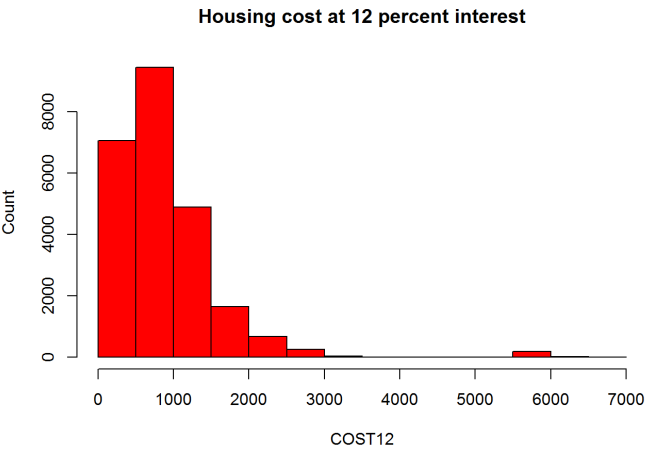
```
plot(hads2013n_c$FMTOWNRENT,hads2013n_c$COST06,xlab='FMTOWNRENT',ylab='COST06')
```



```
hist(hads2013n_c$COST12[which(hads2013n_c$FMTOWNRENT == "'1 Owner'")],main= 'Housing cost at 12 percent interest', xlab='COST12', ylab='Count',col = 'blue')
```

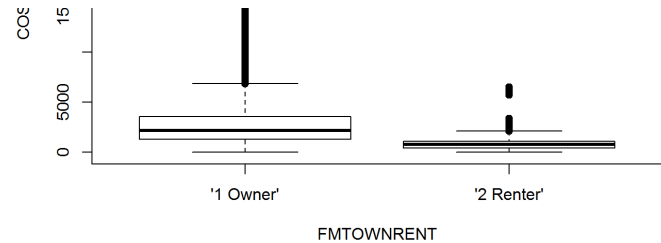


```
hist(hads2013n_c$COST12[which(hads2013n_c$FMTOWNRENT == "2 Renter")],main= 'Housing cost at 12 percent interest', xlab='COST12', ylab='Count',col = 'red')
```



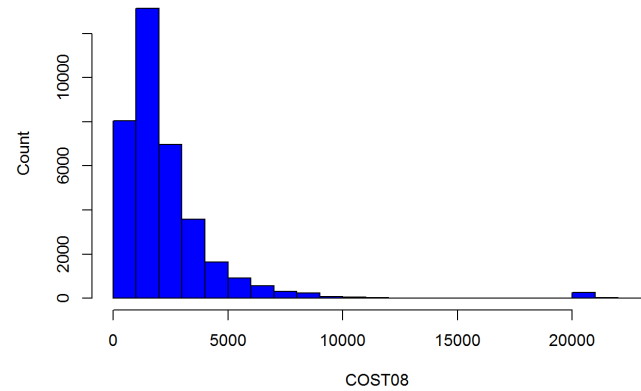
```
plot(hads2013n_c$FMTOWNRENT,hads2013n_c$COST12,xlab='FMTOWNRENT',ylab='COST12')
```





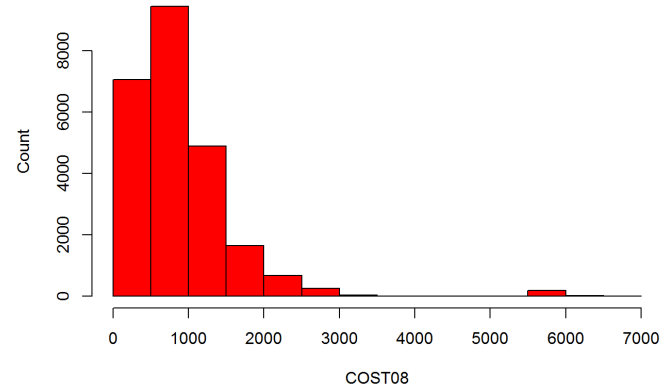
```
hist(hads2013n_c$COST08[which(hads2013n_c$FMTOWNRENT == "1 Owner")],main= 'Housing cost at 8 percent interest', xlab='COST08', ylab='Count',col = 'blue')
```

Housing cost at 8 percent interest



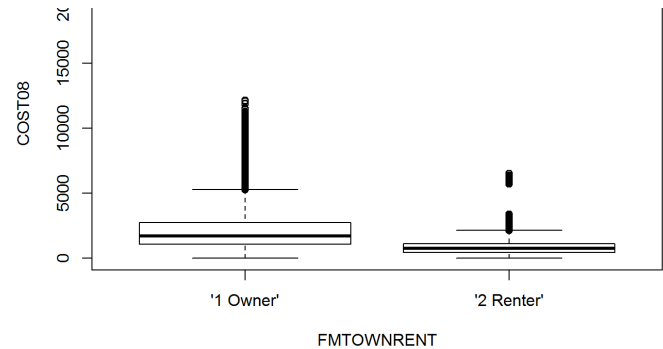
```
hist(hads2013n_c$COST08[which(hads2013n_c$FMTOWNRENT == "2 Renter")],main= 'Housing cost at 8 percent interest', xlab='COST08', ylab='Count',col = 'red')
```

Housing cost at 8 percent interest

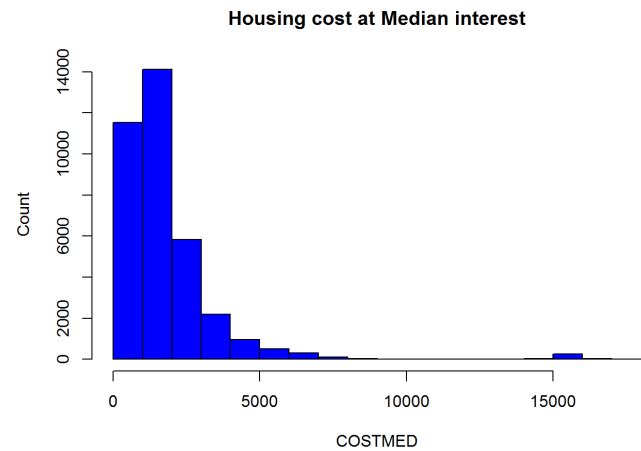


```
plot(hads2013n_c$FMTOWNRENT,hads2013n_c$COST08,xlab='FMTOWNRENT',ylab='COST08')
```

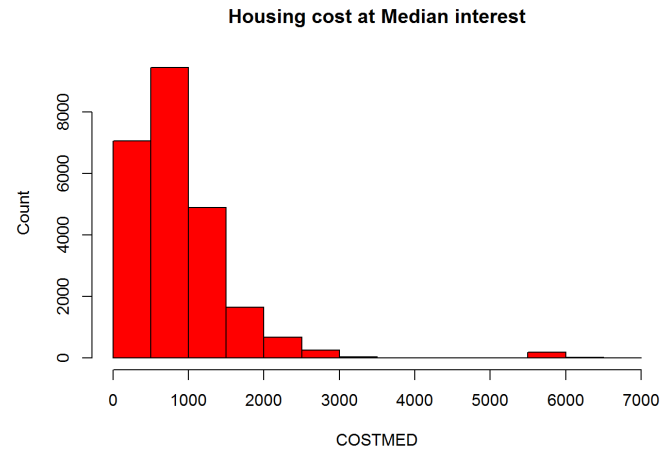




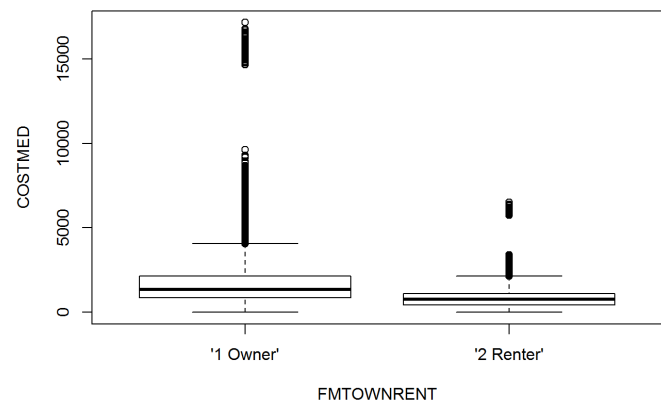
```
hist(hads2013n_c$COSTMED[which(hads2013n_c$FMTOWNRENT == "'1 Owner'")],main= 'Housing cost at Median interest', xlab='COSTMED', ylab='Count',col = 'blue')
```



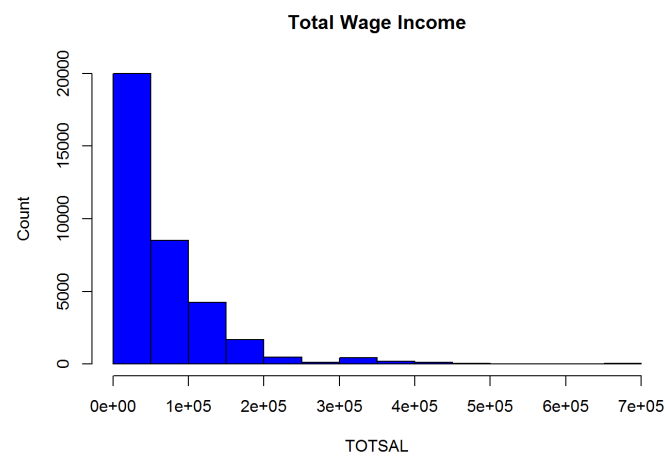
```
hist(hads2013n_c$COSTMED[which(hads2013n_c$FMTOWNRENT == "'2 Renter'")],main= 'Housing cost at Median interest', xlab='COSTMED', ylab='Count',col = 'red')
```



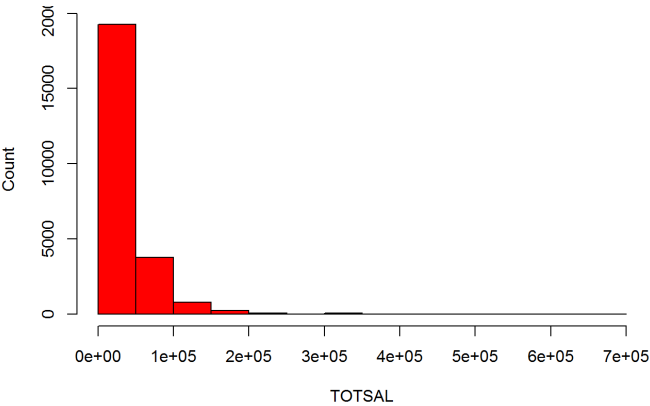
```
plot(hads2013n_c$FMTOWNRENT,hads2013n_c$COSTMED,xlab= 'FMTOWNRENT', ylab= 'COSTMED')
```



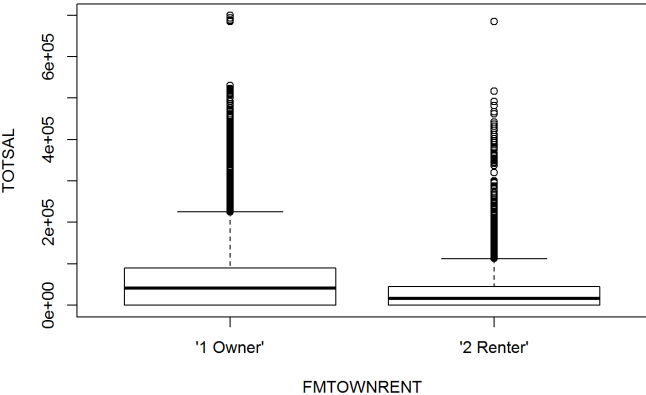
```
hist(hads2013n_c$TOTSAL[which(hads2013n_c$FMTOWNRENT == "1 Owner")],main= 'Total Wage Income', xlab='TOTSAL',
ylab='Count',col ='blue')
```



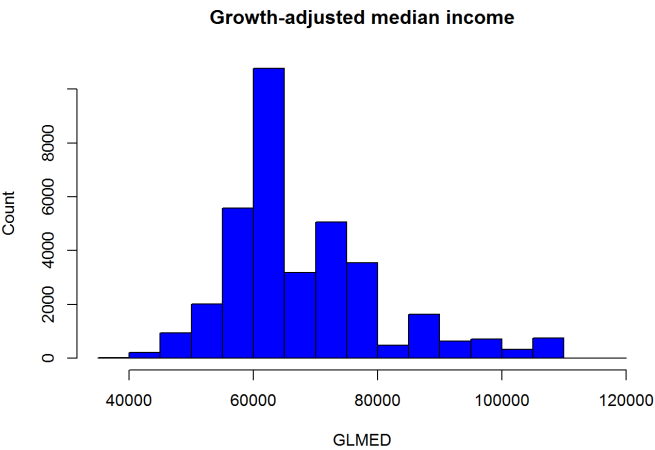
```
hist(hads2013n_c$TOTSAL[which(hads2013n_c$FMTOWNRENT == "2 Renter")],main= 'Total Wage Income', xlab='TOTSAL',
ylab='Count',col ='red')
```



```
plot(hads2013n_c$FMTOWNRENT,hads2013n_c$TOTSAL,xlab='FMTOWNRENT',ylab='TOTSAL')
```

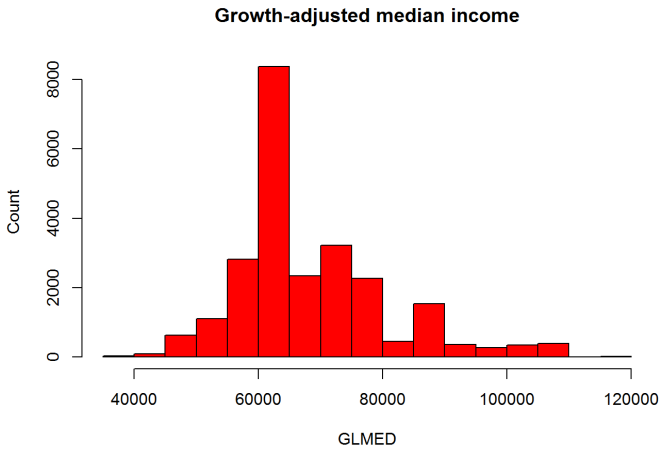


```
hist(hads2013n_c$GLMED[which(hads2013n_c$FMTOWNRENT == "1 Owner")],main= 'Growth-adjusted median income', xlab='GLMED', ylab='Count',col = 'blue')
```

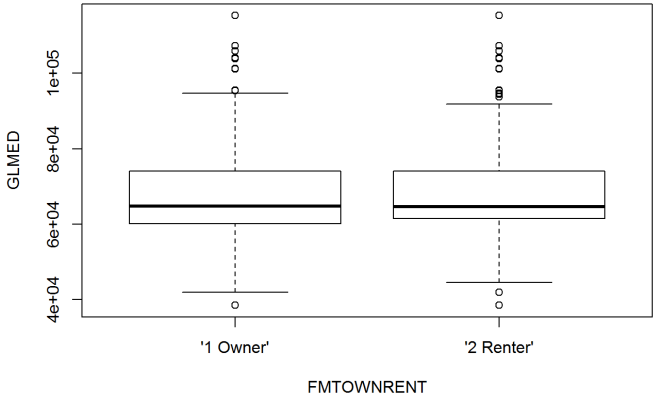


```
plot(hads2013n_c$GLMED[which(hads2013n_c$FMTOWNRENT == "1 Owner")],main= 'Growth-adjusted median income', xlab='GLMED', ylab='Count',col = 'blue')
```

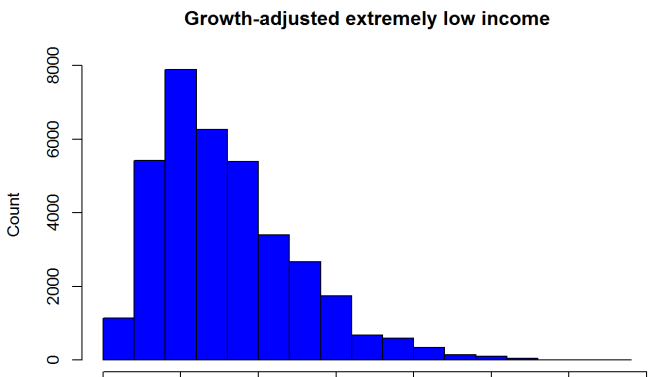
```
hist(hads2013n_c$GLMED[which(hads2013n_c$FMTOWNRENT == "2 Renter")],main= 'Growth-adjusted median income', xlab='GLMED', ylab='Count',col = 'red')
```



```
plot(hads2013n_c$FMTOWNRENT,hads2013n_c$GLMED,xlab='FMTOWNRENT',ylab='GLMED')
```



```
hist(hads2013n_c$GL30[which(hads2013n_c$FMTOWNRENT == "1 Owner")],main= 'Growth-adjusted extremely low income', xlab='GL30', ylab='Count',col = 'blue')
```

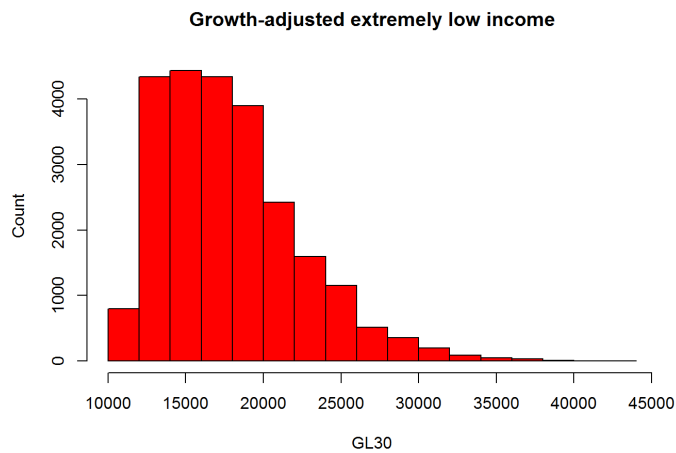




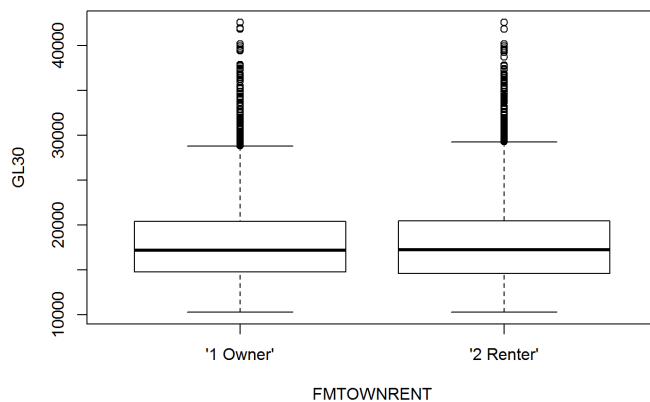
10000 15000 20000 25000 30000 35000 40000 45000

GL30

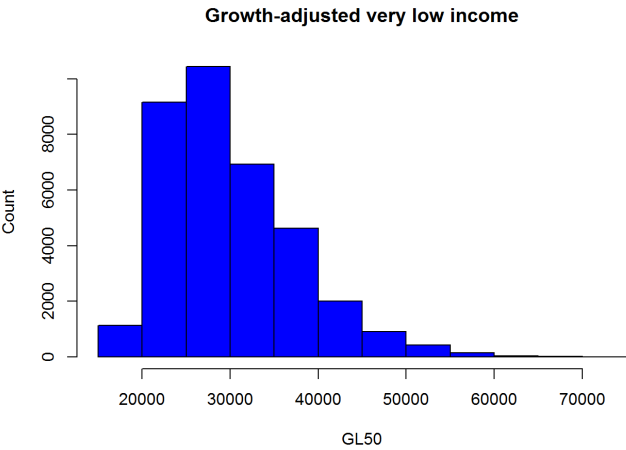
```
hist(hads2013n_c$GL30[which(hads2013n_c$FMTOWNRENT == "'2 Renter'")],main= 'Growth-adjusted extremely low income', xlab='GL30', ylab='Count',col = 'red')
```



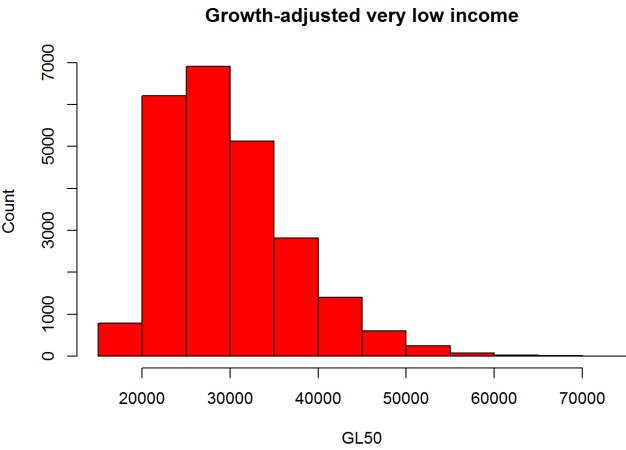
```
plot(hads2013n_c$FMTOWNRENT,hads2013n_c$GL30,xlab='FMTOWNRENT',ylab='GL30')
```



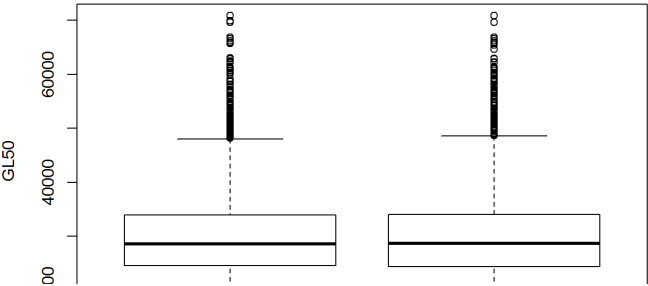
```
hist(hads2013n_c$GL50[which(hads2013n_c$FMTOWNRENT == "'1 Owner'")],main= 'Growth-adjusted very low income', xlab='GL50', ylab='Count',col = 'blue')
```

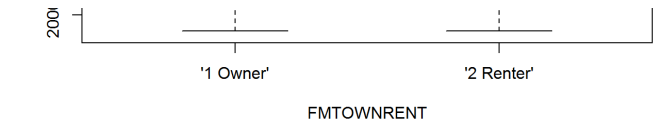


```
hist(hads2013n_c$GL50[which(hads2013n_c$FMTOWNRENT == "'2 Renter'")],main= 'Growth-adjusted very low income', xlab='GL50', ylab='Count',col = 'red')
```

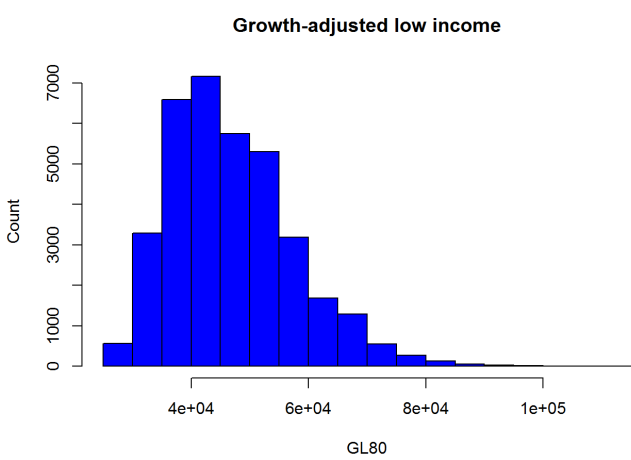


```
plot(hads2013n_c$FMTOWNRENT,hads2013n_c$GL50,xlab= 'FMTOWNRENT',ylab= 'GL50')
```

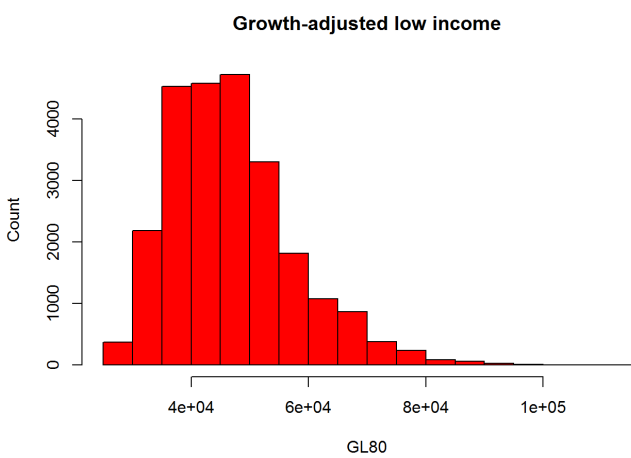




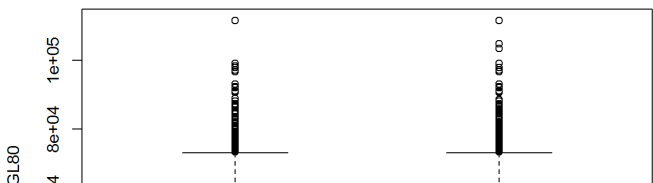
```
hist(hads2013n_c$GL80[which(hads2013n_c$FMTOWNRENT == "1 Owner")],main= 'Growth-adjusted low income', xlab='GL80', ylab='Count',col = 'blue')
```

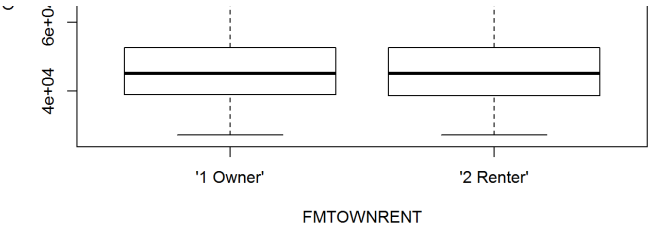


```
hist(hads2013n_c$GL80[which(hads2013n_c$FMTOWNRENT == "2 Renter")],main= 'Growth-adjusted low income', xlab='GL80', ylab='Count',col = 'red')
```



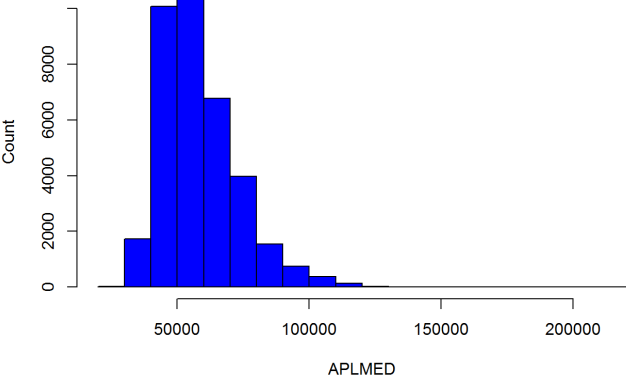
```
plot(hads2013n_c$FMTOWNRENT,hads2013n_c$GL80,xlab='FMTOWNRENT',ylab='GL80')
```





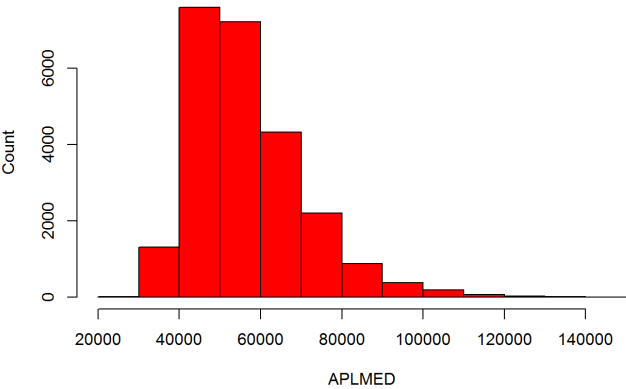
```
hist(hads2013n_c$APLMED[which(hads2013n_c$FMTOWNRENT == "1 Owner")],main= 'Median Income Adjusted for # of Persons',
xlab='APLMED', ylab='Count',col = 'blue')
```

Median Income Adjusted for # of Persons

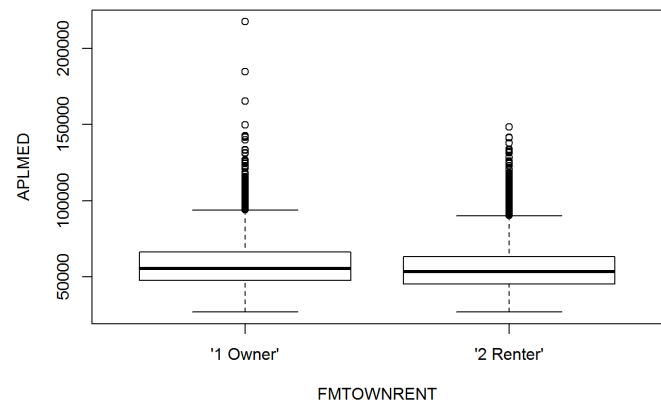


```
hist(hads2013n_c$APLMED[which(hads2013n_c$FMTOWNRENT == "2 Renter")],main= 'Median Income Adjusted for # of Persons',
xlab='APLMED', ylab='Count',col = 'red')
```

Median Income Adjusted for # of Persons

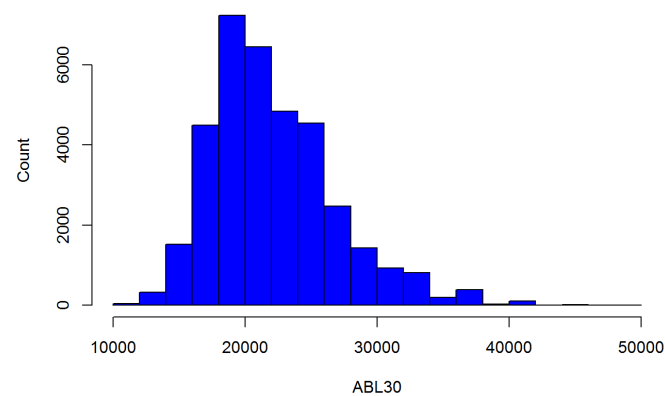


```
plot(hads2013n_c$FMTOWNRENT,hads2013n_c$APLMED,xlab='FMTOWNRENT',ylab='APLMED')
```



```
hist(hads2013n_c$ABL30[which(hads2013n_c$FMTOWNRENT == "1 Owner")],main= 'Extremely Low Income Adjusted for # of Bedrooms',
xlab='ABL30', ylab='Count',col = 'blue')
```

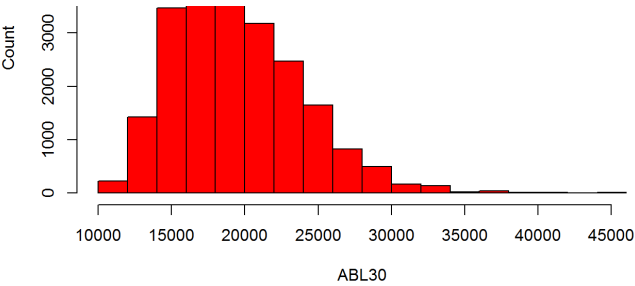
Extremely Low Income Adjusted for # of Bedrooms



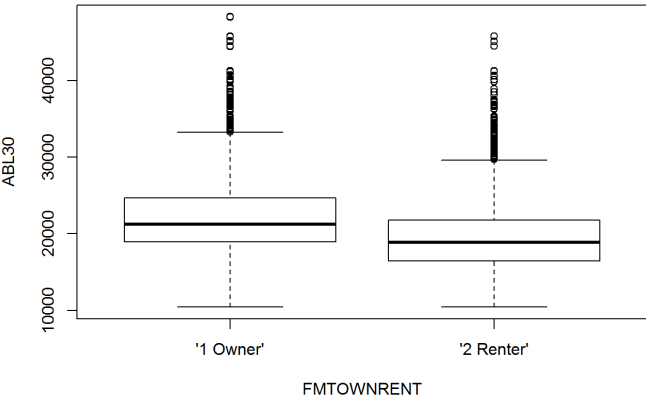
```
hist(hads2013n_c$ABL30[which(hads2013n_c$FMTOWNRENT == "2 Renter")],main= 'Extremely Low Income Adjusted for # of Bedrooms',
xlab='ABL30', ylab='Count',col = 'red')
```

Extremely Low Income Adjusted for # of Bedrooms



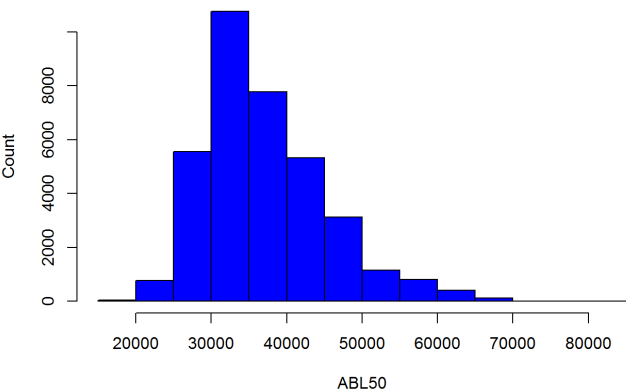


```
plot(hads2013n_c$FMTOWNRENT,hads2013n_c$ABL30,xlab='FMTOWNRENT',ylab='ABL30')
```



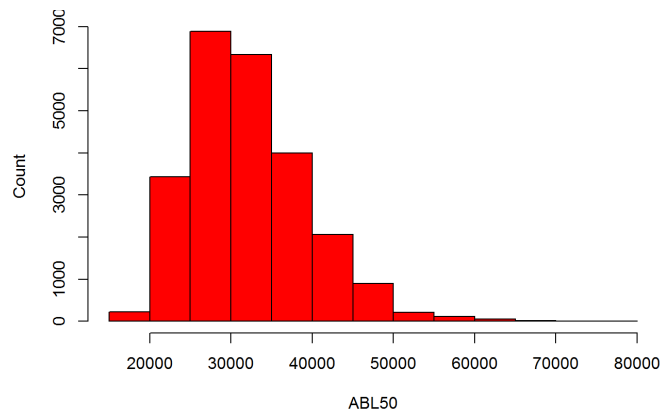
```
hist(hads2013n_c$ABL50[which(hads2013n_c$FMTOWNRENT == "'1 Owner'")],main= 'Very Low Income Adjusted for # of Bedrooms',
xlab='ABL50', ylab='Count',col = 'blue')
```

Very Low Income Adjusted for # of Bedrooms

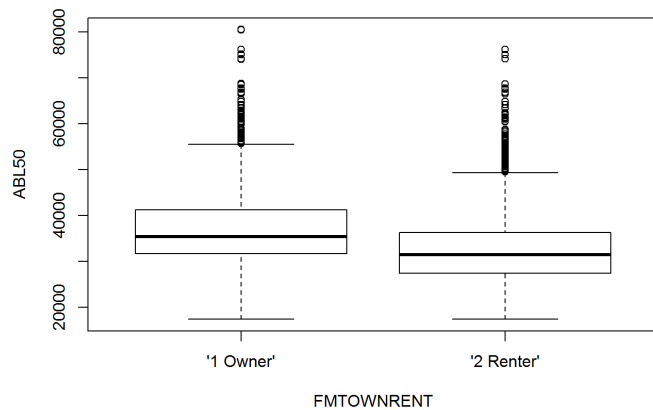


```
hist(hads2013n_c$ABL50[which(hads2013n_c$FMTOWNRENT == "'2 Renter'")],main= 'Very Low Income Adjusted for # of Bedrooms', xla
b='ABL50', ylab='Count',col = 'red')
```

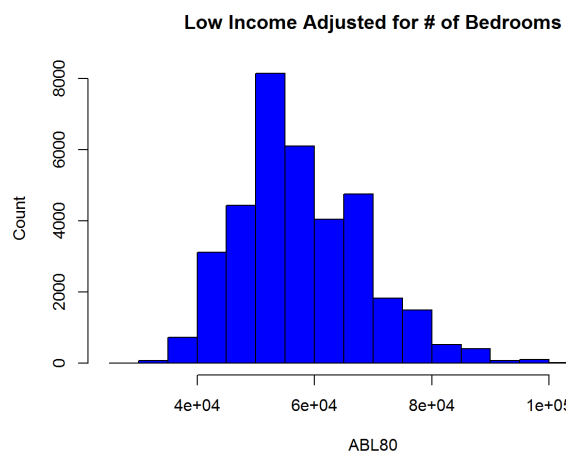
Very Low Income Adjusted for # of Bedrooms



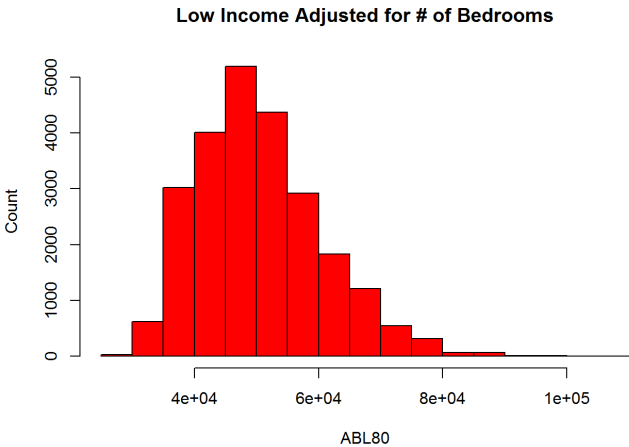
```
plot(hads2013n_c$FMTOWNRENT,hads2013n_c$ABL50,xlab='FMTOWNRENT',ylab='ABL50')
```



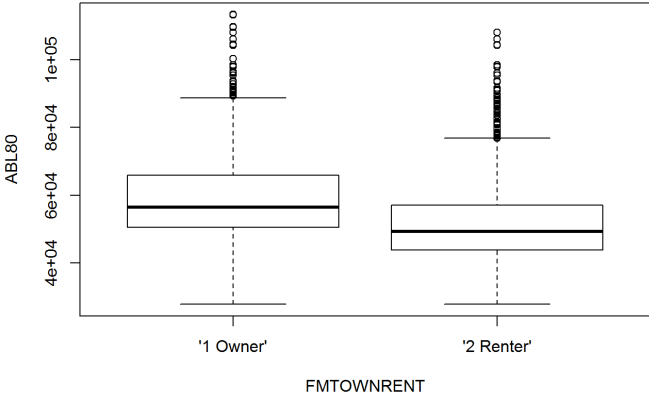
```
hist(hads2013n_c$ABL80[which(hads2013n_c$FMTOWNRENT == "'1 Owner'")],main= 'Low Income Adjusted for # of Bedrooms', xlab='ABL80', ylab='Count',col = 'blue')
```



```
hist(hads2013n_c$ABL80[which(hads2013n_c$FMTOWNRENT == "'2 Renter'")],main= 'Low Income Adjusted for # of Bedrooms', xlab='ABL80', ylab='Count',col = 'red')
```



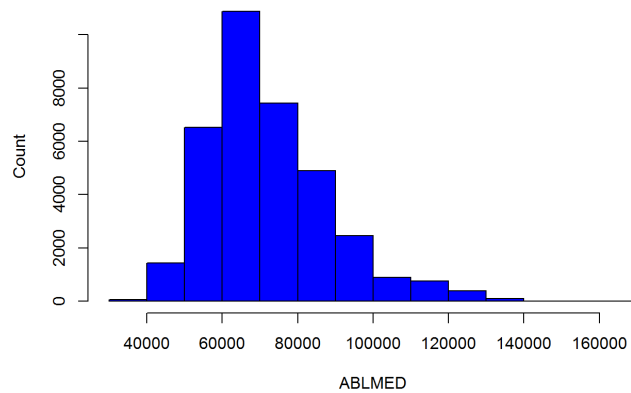
```
plot(hads2013n_c$FMTOWNRENT,hads2013n_c$ABL80,xlab='FMTOWNRENT',ylab='ABL80')
```



```
hist(hads2013n_c$ABLMED[which(hads2013n_c$FMTOWNRENT == "'1 Owner'")],main= 'Median Income Adjusted for # of Bedrooms', xlab='ABLMED', ylab='Count',col = 'blue')
```

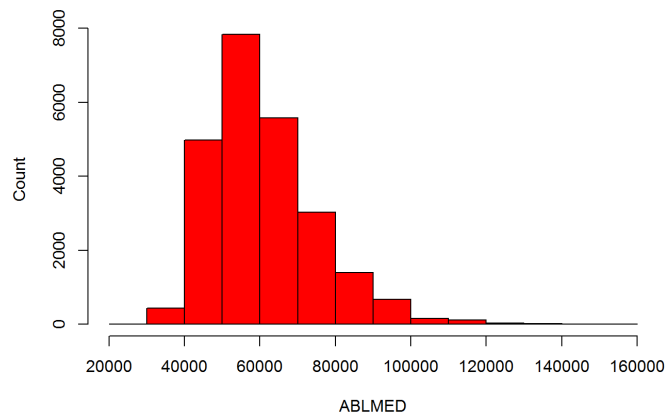


Median Income Adjusted for # of Bedrooms

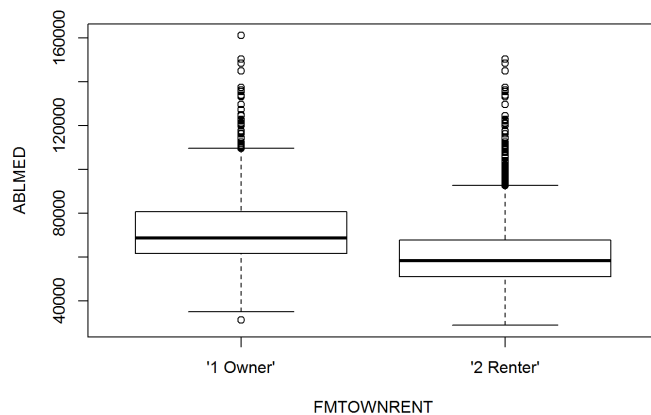


```
hist(hads2013n_c$ABLMED[which(hads2013n_c$FMTOWNRENT == "'2 Renter'")],main= 'Median Income Adjusted for # of Bedrooms',
xlab='ABLMED', ylab='Count',col = 'red')
```

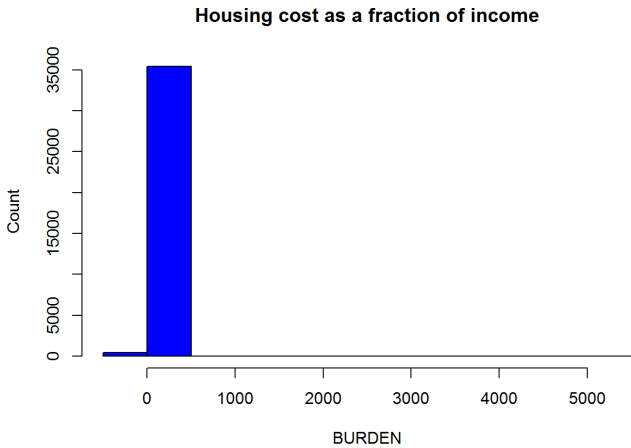
Median Income Adjusted for # of Bedrooms



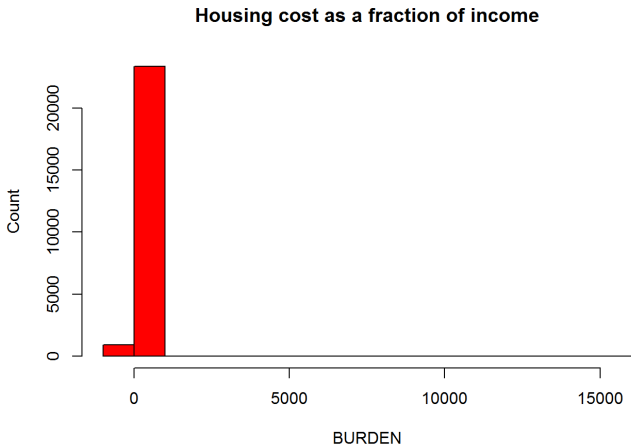
```
plot(hads2013n_c$FMTOWNRENT,hads2013n_c$ABLMED,xlab='FMTOWNRENT',ylab='ABLMED')
```



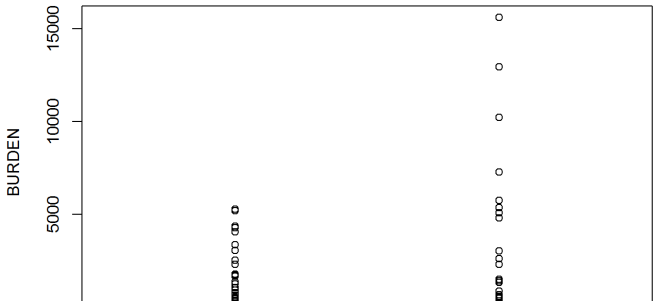
```
hist(hads2013n_c$BURDEN[which(hads2013n_c$FMTOWNRENT == "1 Owner")],main= 'Housing cost as a fraction of income', xlab='BURDEN', ylab='Count',col = 'blue')
```



```
hist(hads2013n_c$BURDEN[which(hads2013n_c$FMTOWNRENT == "2 Renter")],main= 'Housing cost as a fraction of income', xlab='BURDEN', ylab='Count',col = 'red')
```

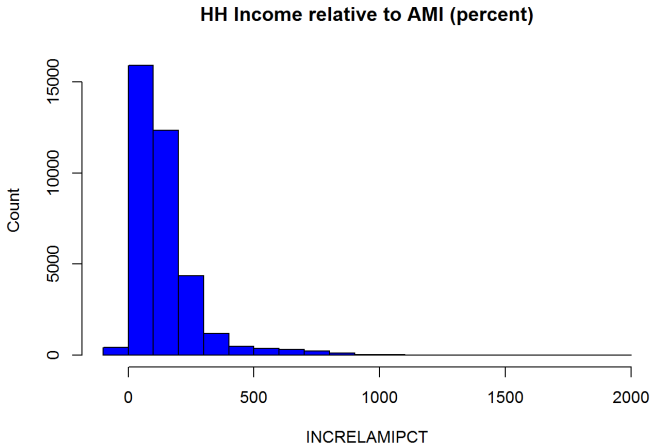


```
plot(hads2013n_c$FMTOWNRENT,hads2013n_c$BURDEN,xlab= 'FMTOWNRENT',ylab='BURDEN')
```

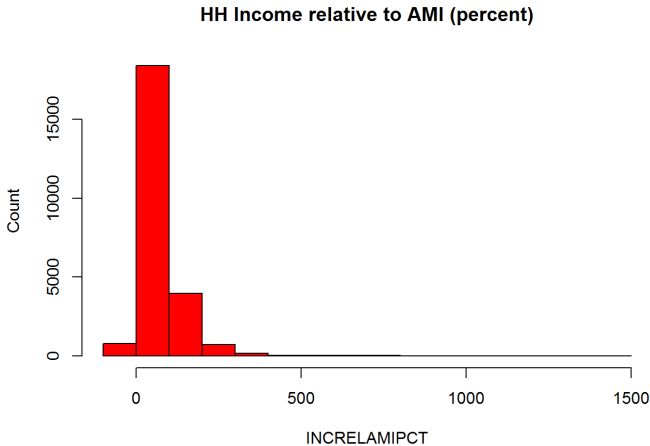




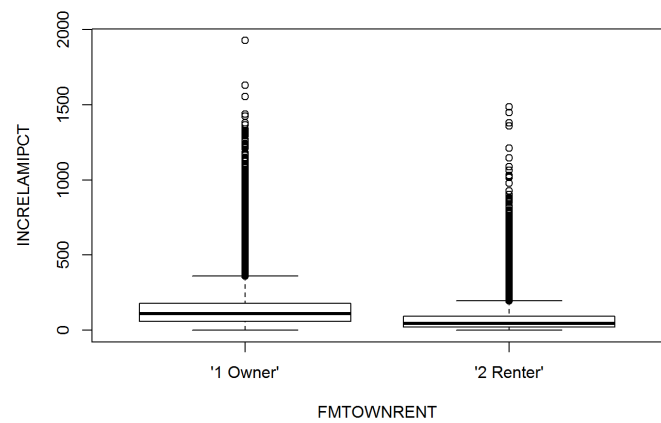
```
hist(hads2013n_c$INCRELAMIPCT[which(hads2013n_c$FMTOWNRENT == "1 Owner")],main= 'HH Income relative to AMI (percent)',
xlab='INCRELAMIPCT', ylab='Count',col = 'blue')
```



```
hist(hads2013n_c$INCRELAMIPCT[which(hads2013n_c$FMTOWNRENT == "2 Renter")],main= 'HH Income relative to AMI (percent)', xla
b='INCRELAMIPCT', ylab='Count',col = 'red')
```

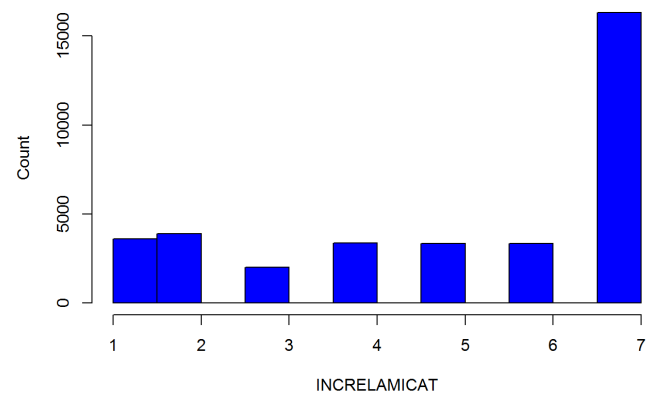


```
plot(hads2013n_c$FMTOWNRENT,hads2013n_c$INCRELAMIPCT,xlab='FMTOWNRENT',ylab='INCRELAMIPCT')
```



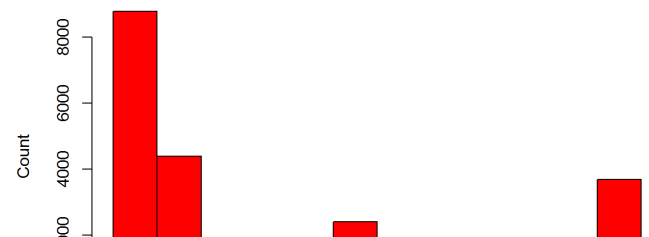
```
hist(hads2013n_c$INCRELAMIPCT[which(hads2013n_c$FMTOWNRENT == "1 Owner")],main= 'HH Income relative to AMI (category)', xlab='INCRELAMIPCT', ylab='Count',col = 'blue')
```

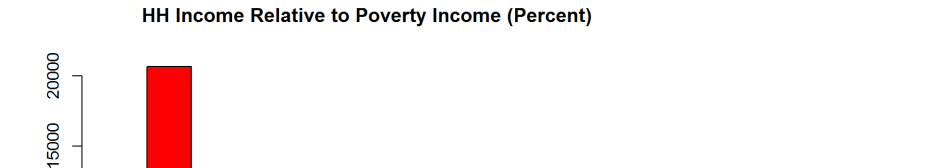
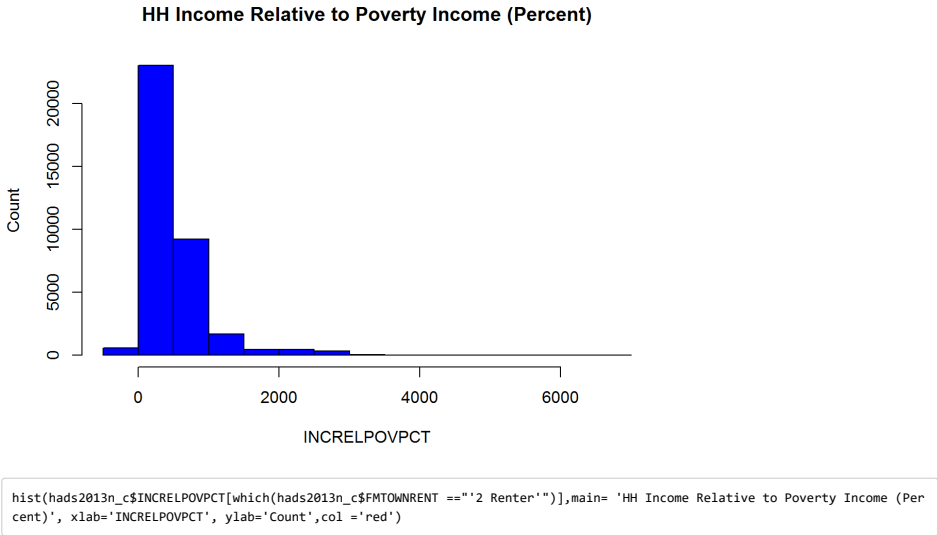
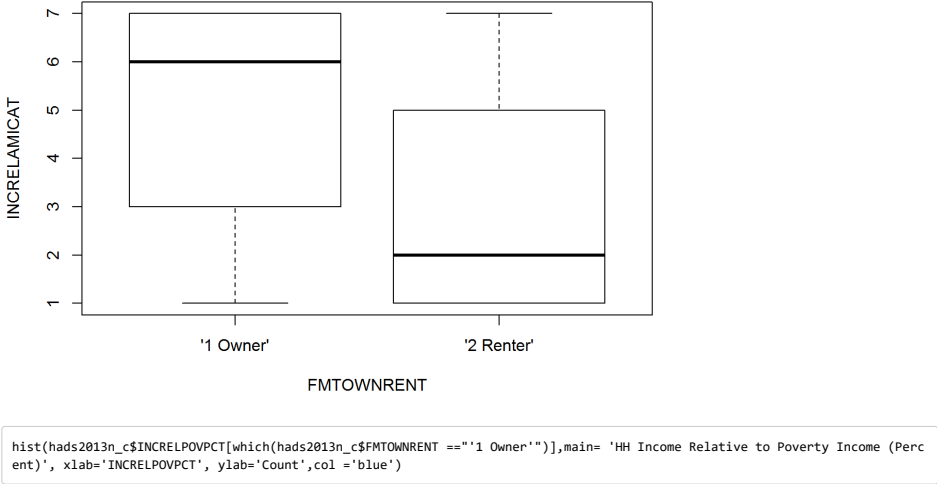
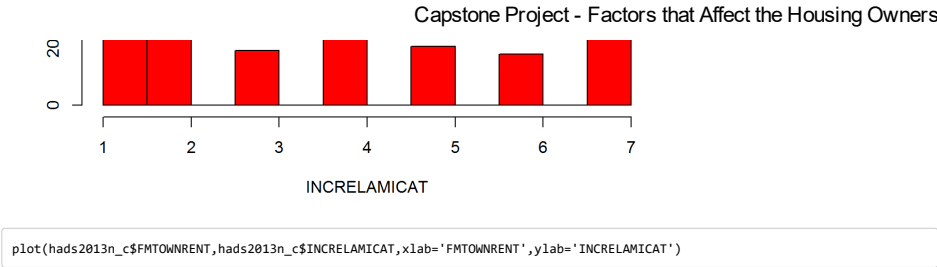
HH Income relative to AMI (category)

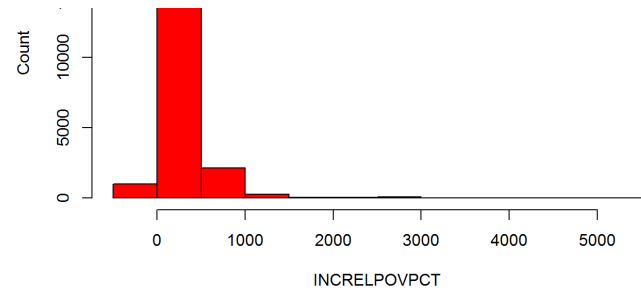


```
hist(hads2013n_c$INCRELAMIPCT[which(hads2013n_c$FMTOWNRENT == "2 Renter")],main= 'HH Income relative to AMI (category)', xlab='INCRELAMIPCT', ylab='Count',col = 'red')
```

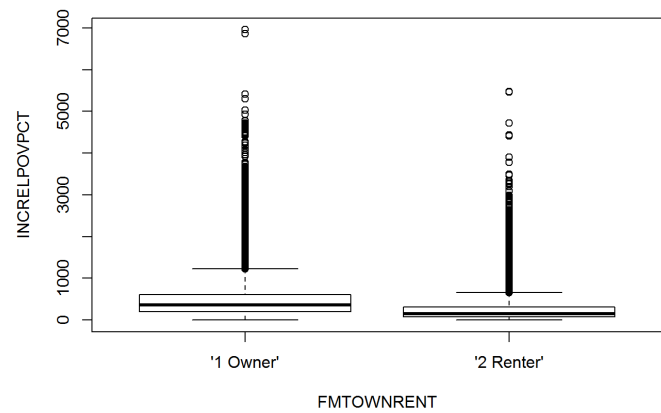
HH Income relative to AMI (category)





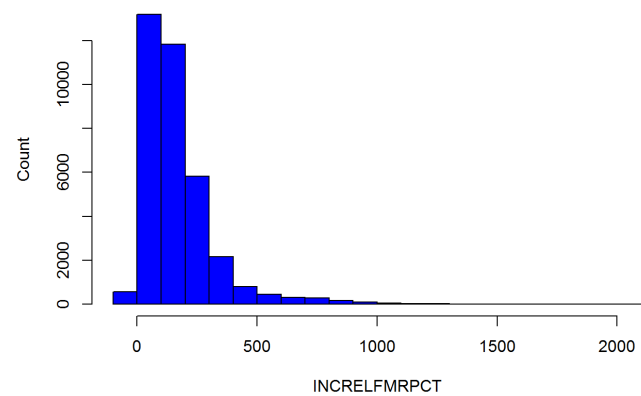


```
plot(hads2013n_c$FMTOWNRENT,hads2013n_c$INCRELPOVPCT,xlab='FMTOWNRENT',ylab='INCRELPOVPCT')
```

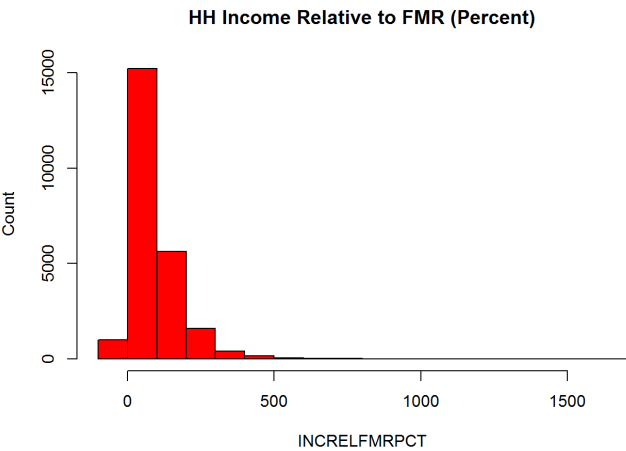


```
hist(hads2013n_c$INCRELFMRPCT[which(hads2013n_c$FMTOWNRENT == "'1 Owner'")],main= 'HH Income Relative to FMR (Percent)',
xlab='INCRELFMRPCT', ylab='Count',col = 'blue')
```

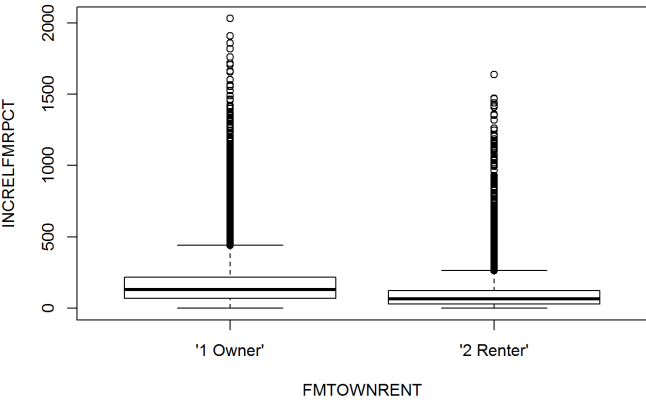
HH Income Relative to FMR (Percent)



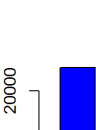
```
hist(hads2013n_c$INCRELFMRPCT[which(hads2013n_c$FMTOWNRENT == "'2 Renter'")],main= 'HH Income Relative to FMR (Percent)', xla
b='INCRELFMRPCT', ylab='Count',col = 'red')
```

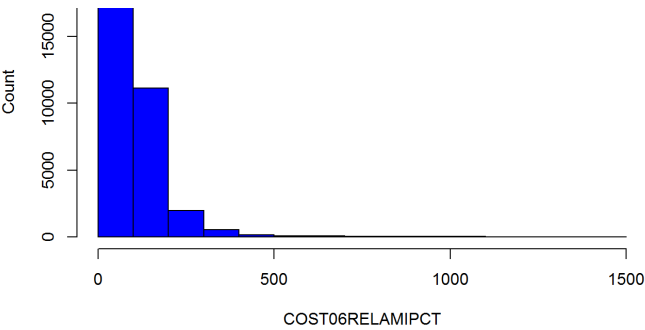


```
plot(hads2013n_c$FMTOWNRENT,hads2013n_c$INCRELFMRPCT,xlab=' FMTOWNRENT',ylab=' INCRELFMRPCT')
```



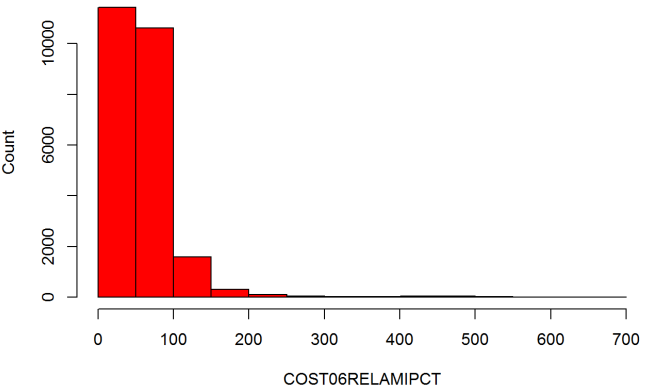
```
hist(hads2013n_c$COST06RELAMIPCT[which(hads2013n_c$FMTOWNRENT == "1 Owner")],main= 'Cost06 Relative to Median Income (Percent)', xlab='COST06RELAMIPCT', ylab='Count',col = 'blue')
```



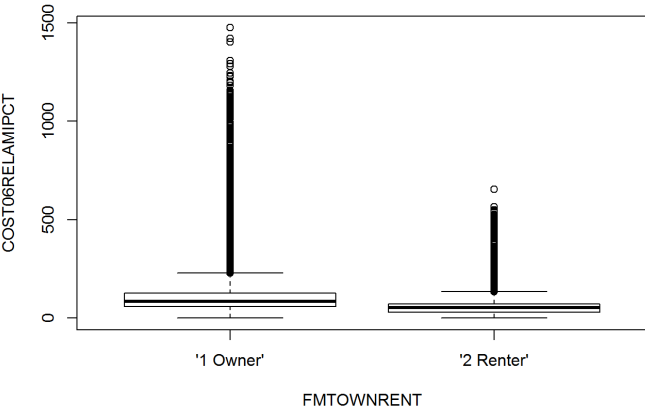


```
hist(hads2013n_c$COST06RELAMIPCT[which(hads2013n_c$FMTOWNRENT == "2 Renter")],main= 'Cost06 Relative to Median Income (Percent)', xlab='COST06RELAMIPCT', ylab='Count',col ='red')
```

Cost06 Relative to Median Income (Percent)

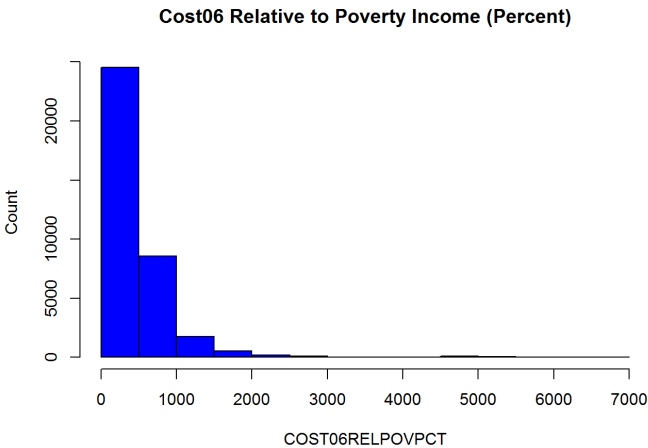


```
plot(hads2013n_c$FMTOWNRENT,hads2013n_c$COST06RELAMIPCT,xlab='FMTOWNRENT',ylab='COST06RELAMIPCT')
```

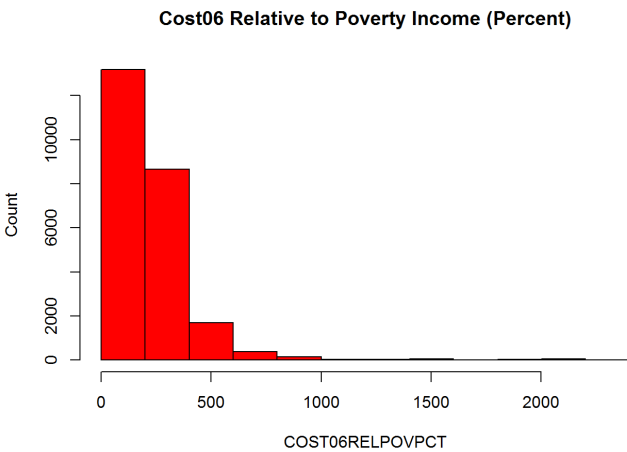


```
hist(hads2013n_c$COST06RELPOVPCT[which(hads2013n_c$FMTOWNRENT == "1 Owner")],main= 'Cost06 Relative to Poverty Income (Percent)', xlab='COST06RELPOVPCT', ylab='Count',col ='blue')
```

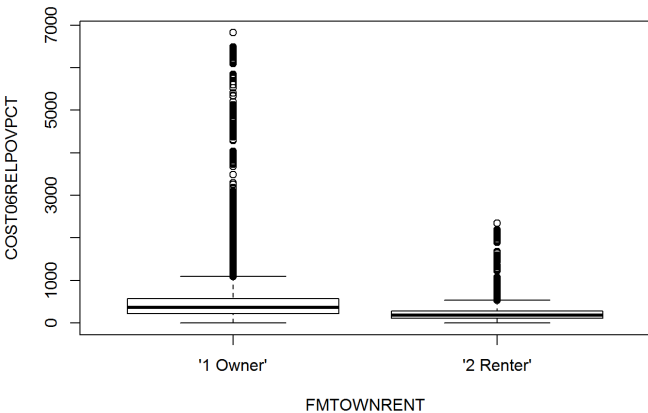




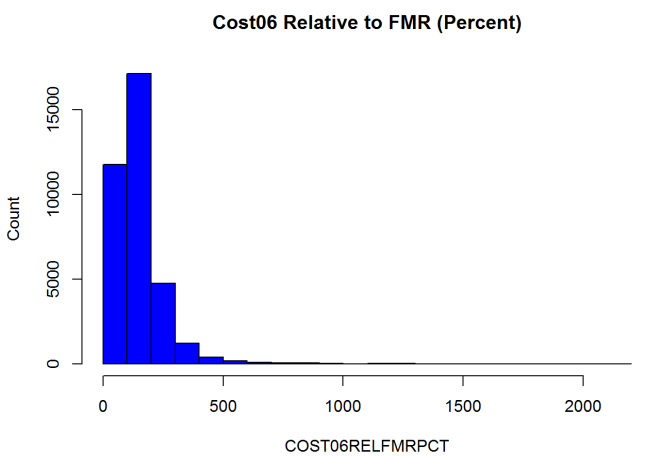
```
hist(hads2013n_c$COST06RELPOVPCT[which(hads2013n_c$FMTOWNRENT == "2 Renter")],main= 'Cost06 Relative to Poverty Income (Per cent)', xlab='COST06RELPOVPCT', ylab='Count',col = 'red')
```



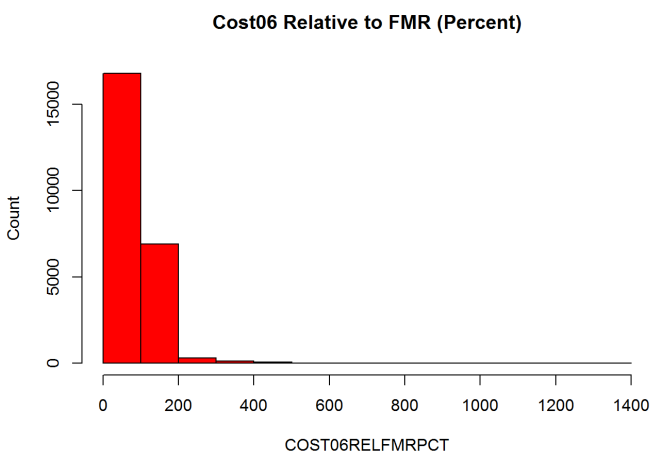
```
plot(hads2013n_c$FMTOWNRENT,hads2013n_c$COST06RELPOVPCT,xlab='FMTOWNRENT',ylab='COST06RELPOVPCT')
```



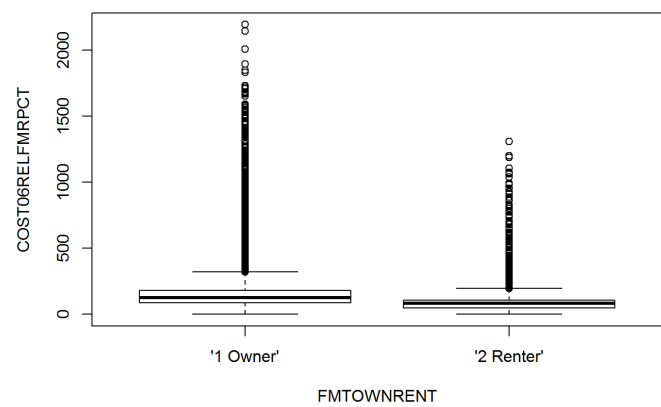
```
hist(hads2013n_c$COST06RELFMRPCT[which(hads2013n_c$FMTOWNRENT == "1 Owner")],main= 'Cost06 Relative to FMR (Percent)',
xlab='COST06RELFMRPCT', ylab='Count',col = 'blue')
```



```
hist(hads2013n_c$COST06RELFMRPCT[which(hads2013n_c$FMTOWNRENT == "2 Renter")],main= 'Cost06 Relative to FMR (Percent)', xla
b='COST06RELFMRPCT', ylab='Count',col = 'red')
```

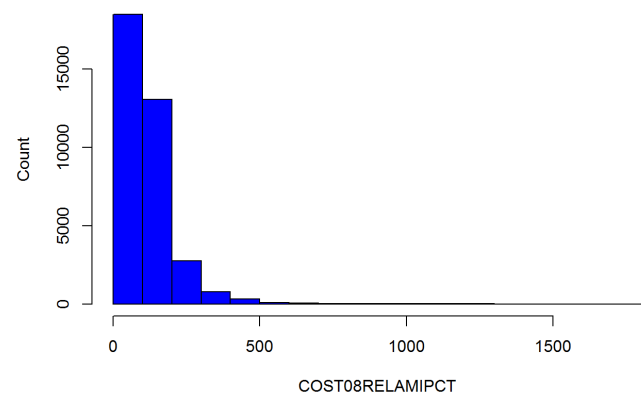


```
p1ot(hads2013n_c$FMTOWNRENT,hads2013n_c$COST06RELFMRPCT,xlab='FMTOWNRENT',ylab='COST06RELFMRPCT')
```



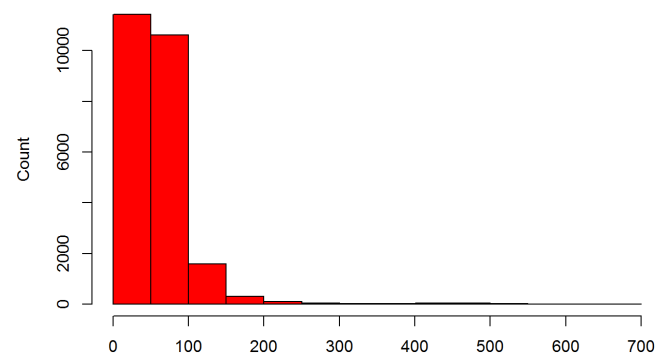
```
hist(hads2013n_c$COST08RELAMIPCT[which(hads2013n_c$FMTOWNRENT == "1 Owner")],main= 'Cost08 Relative to Median Income (Percent)', xlab='COST08RELAMIPCT', ylab='Count',col = 'blue')
```

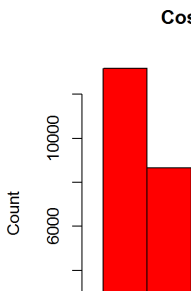
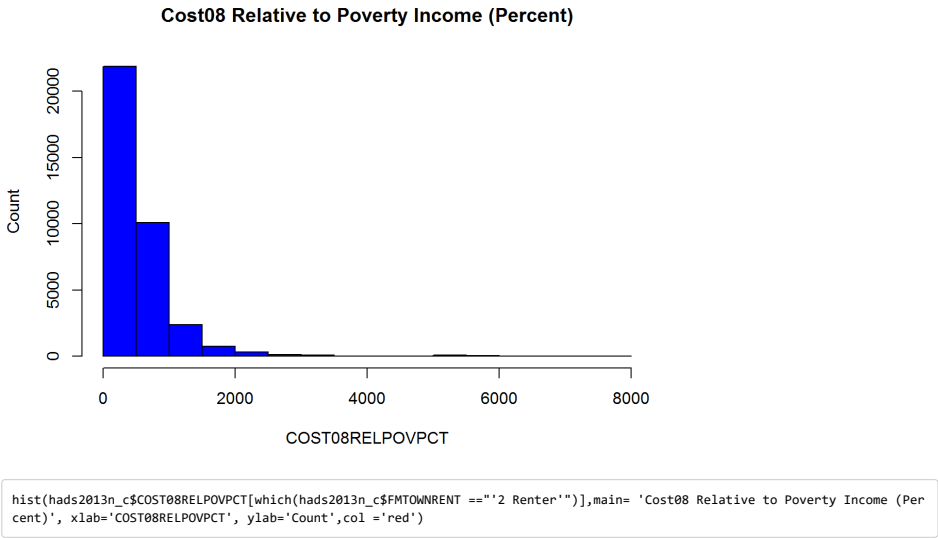
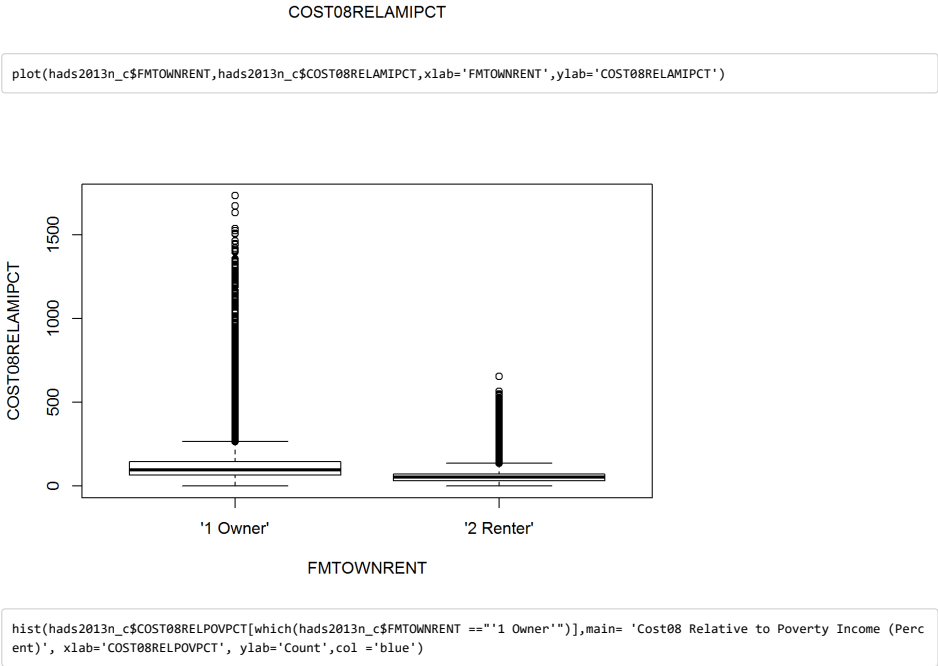
Cost08 Relative to Median Income (Percent)

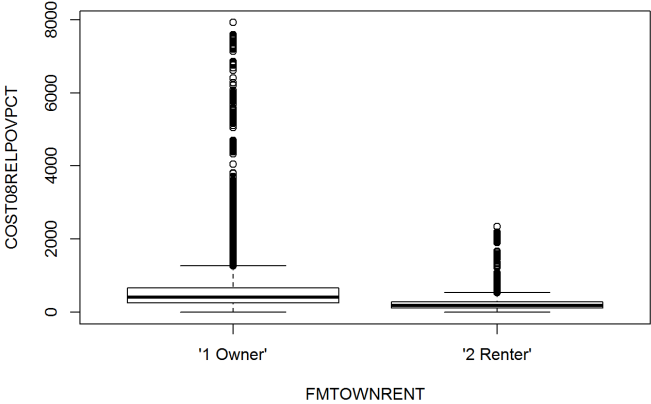
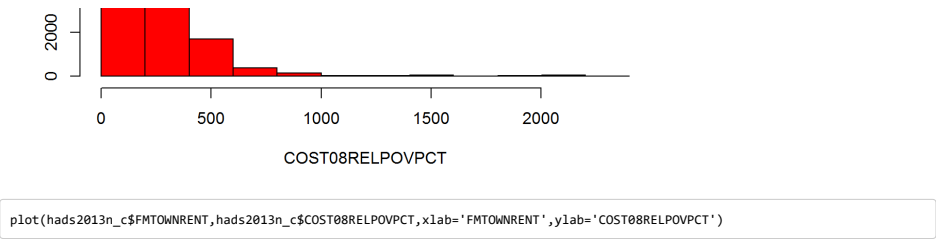


```
hist(hads2013n_c$COST08RELAMIPCT[which(hads2013n_c$FMTOWNRENT == "2 Renter")],main= 'Cost08 Relative to Median Income (Percent)', xlab='COST08RELAMIPCT', ylab='Count',col = 'red')
```

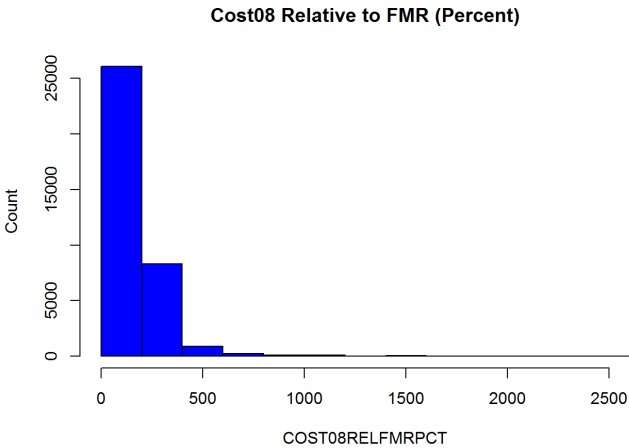
Cost08 Relative to Median Income (Percent)



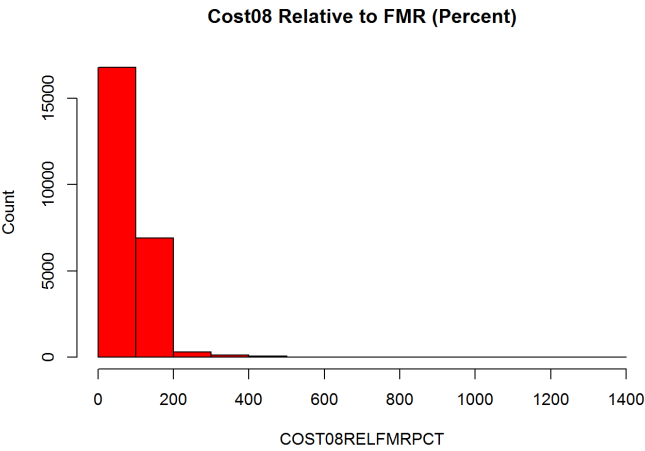




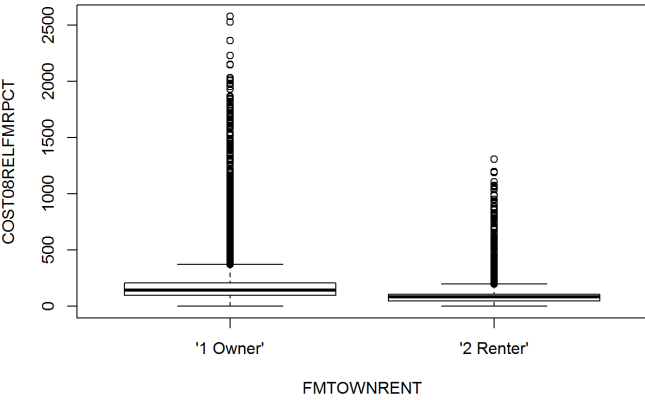
```
hist(hads2013n_c$COST08RELFMRPCT[which(hads2013n_c$FMTOWNRENT == "1 Owner")],main= 'Cost08 Relative to FMR (Percent)', xlab='COST08RELFMRPCT', ylab='Count',col = 'blue')
```



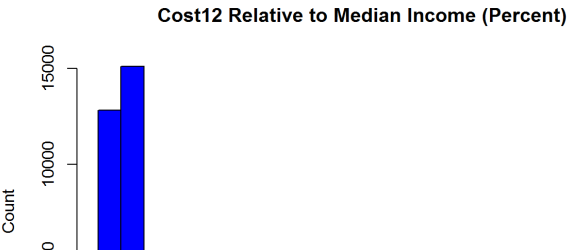
```
hist(hads2013n_c$COST08RELFMRPCT[which(hads2013n_c$FMTOWNRENT == "2 Renter")],main= 'Cost08 Relative to FMR (Percent)', xlab='COST08RELFMRPCT', ylab='Count',col = 'red')
```

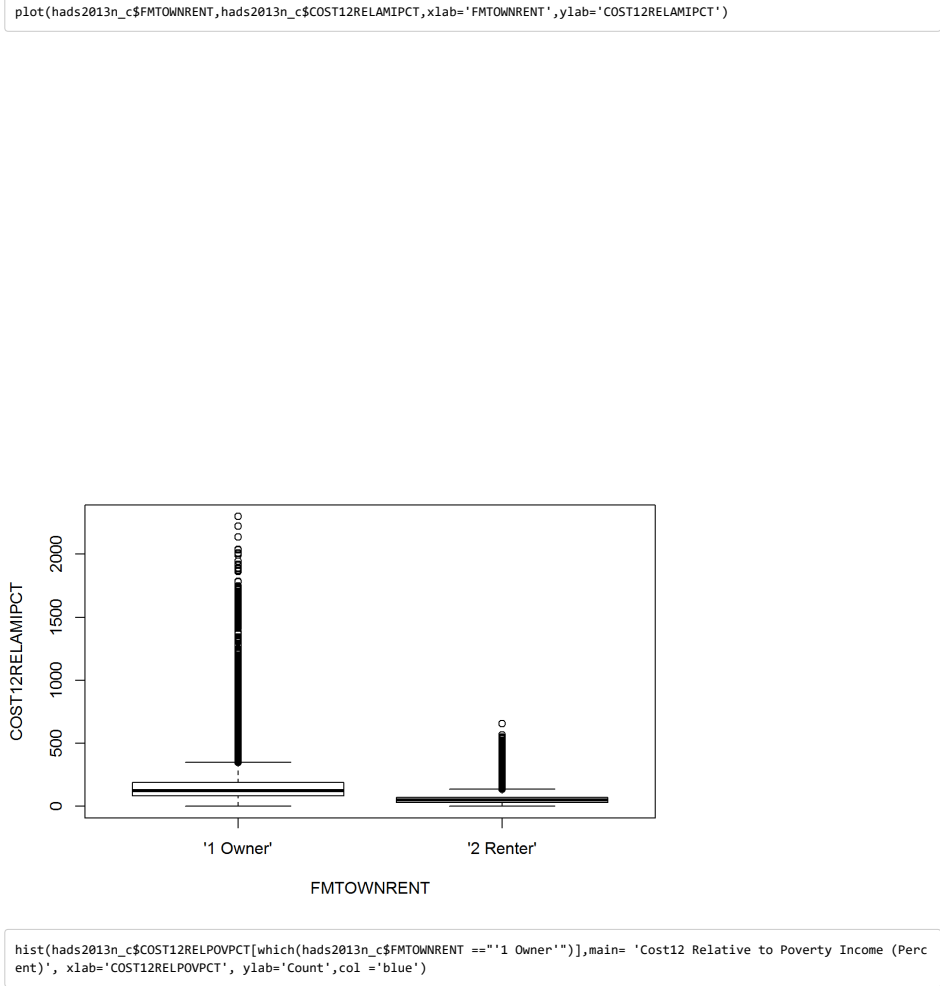
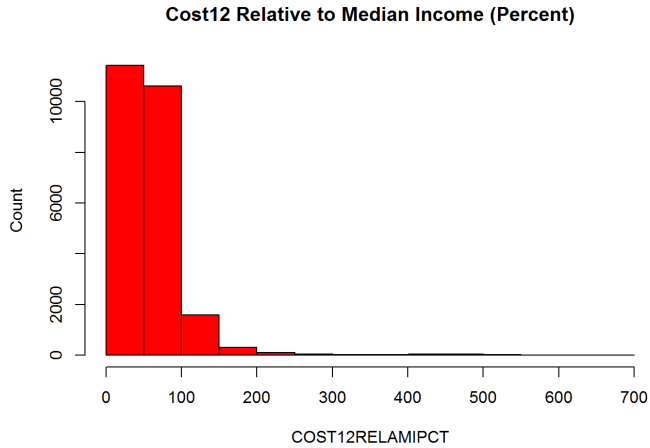
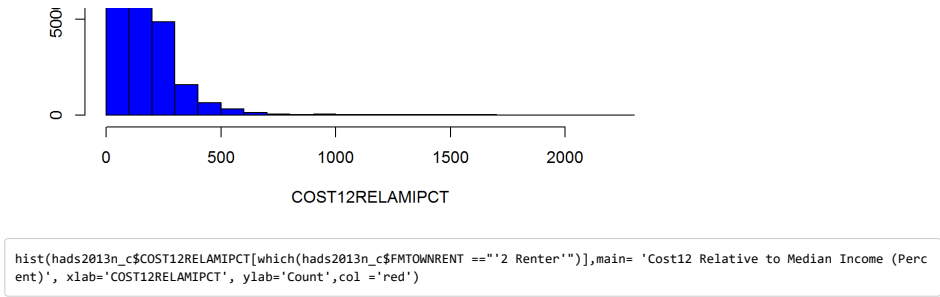


```
plot(hads2013n_c$FMTOWNRENT,hads2013n_c$COST08RELFMRPCT,xlab='FMTOWNRENT',ylab='COST08RELFMRPCT')
```



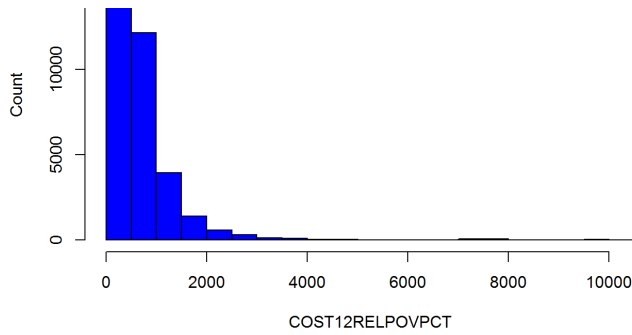
```
hist(hads2013n_c$COST12RELAMIPCT[which(hads2013n_c$FMTOWNRENT == "1 Owner")],main= 'Cost12 Relative to Median Income (Percent)', xlab='COST12RELAMIPCT', ylab='Count',col = 'blue')
```



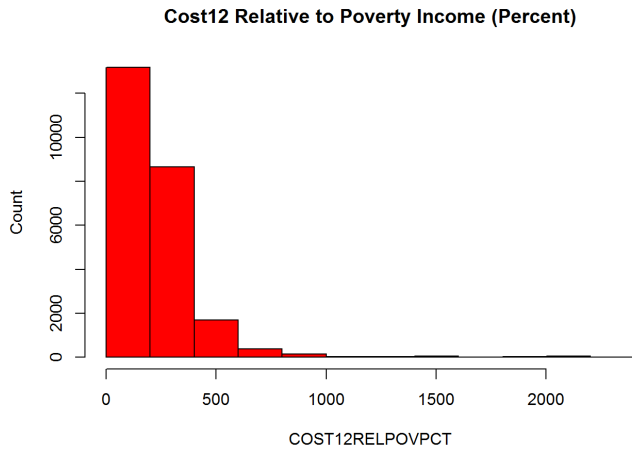


Cost12 Relative to Poverty Income (Percent)

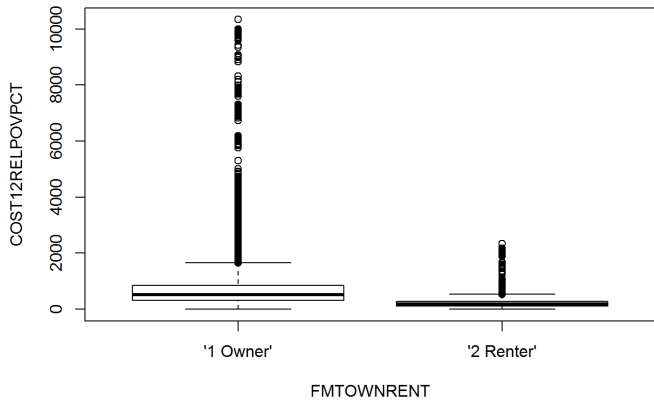




```
hist(hads2013n_c$COST12RELPOVPCT[which(hads2013n_c$FMTOWNRENT == "'2 Renter'")],main= 'Cost12 Relative to Poverty Income (Per cent)', xlab='COST12RELPOVPCT', ylab='Count',col = 'red')
```

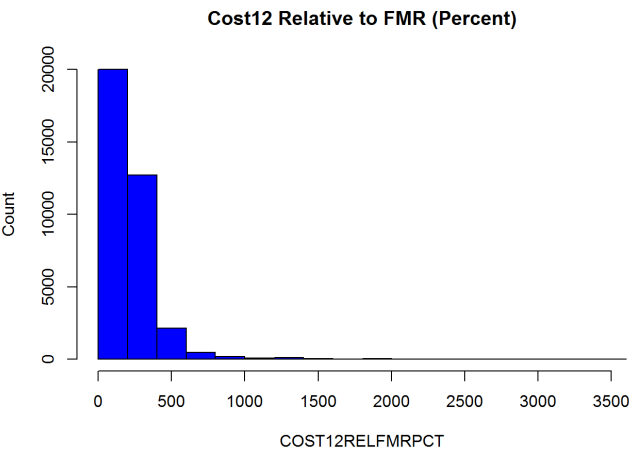


```
plot(hads2013n_c$FMTOWNRENT,hads2013n_c$COST12RELPOVPCT,xlab='FMTOWNRENT',ylab='COST12RELPOVPCT')
```

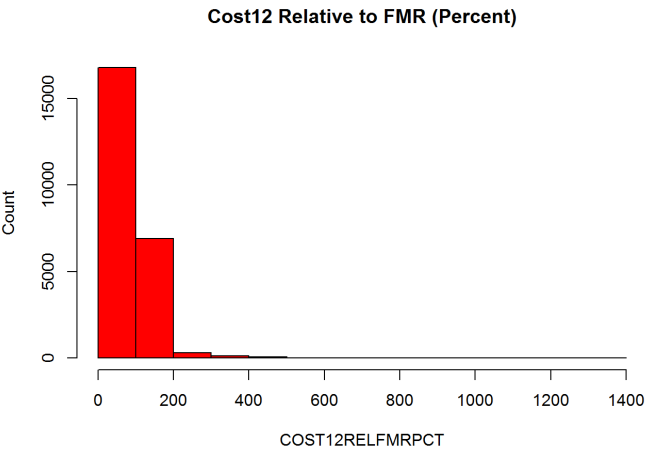


```
hist(hads2013n_c$COST12RELFMRPCT[which(hads2013n_c$FMTOWNRENT == "'1 Owner'")],main= 'Cost12 Relative to FMR (Percent)', xlab='COST12RELFMRPCT', ylab='Count',col = 'blue')
```

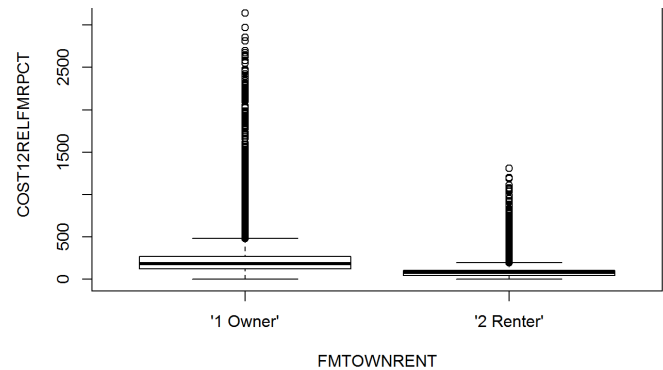




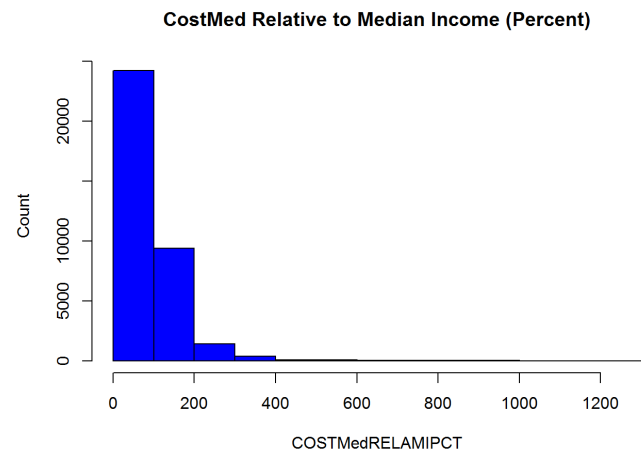
```
hist(hads2013n_c$COST12RELFMRPCT[which(hads2013n_c$FMTOWNRENT == "2 Renter")],main= 'Cost12 Relative to FMR (Percent)', xlab='COST12RELFMRPCT', ylab='Count',col = 'red')
```



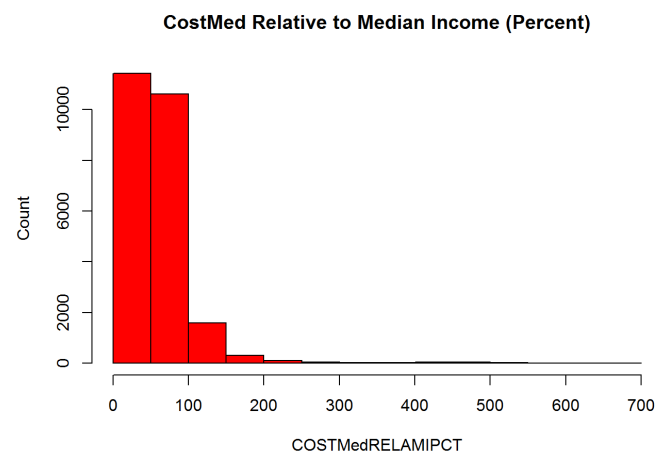
```
plot(hads2013n_c$FMTOWNRENT,hads2013n_c$COST12RELFMRPCT,xlab='FMTOWNRENT',ylab='COST12RELFMRPCT')
```



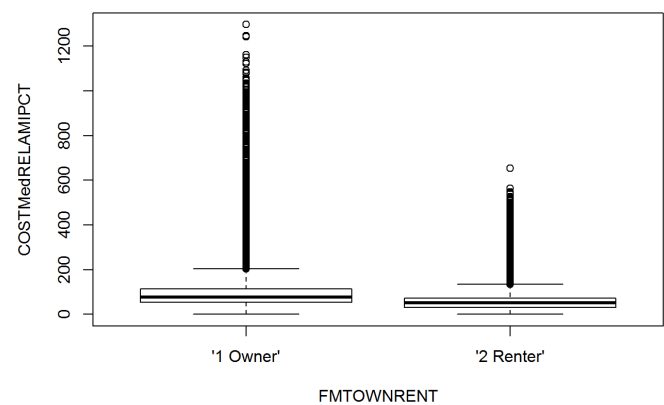
```
hist(hads2013n_c$COSTMedRELAMIPCT[which(hads2013n_c$FMTOWNRENT == "1 Owner")],main= 'CostMed Relative to Median Income (Percent)', xlab='COSTMedRELAMIPCT', ylab='Count',col = 'blue')
```



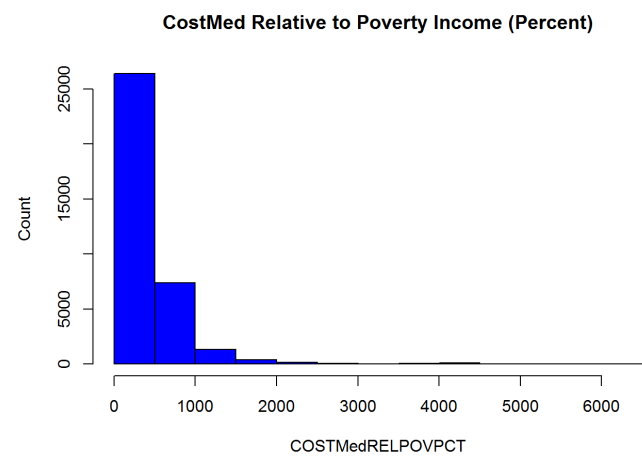
```
hist(hads2013n_c$COSTMedRELAMIPCT[which(hads2013n_c$FMTOWNRENT == "2 Renter")],main= 'CostMed Relative to Median Income (Percent)', xlab='COSTMedRELAMIPCT', ylab='Count',col = 'red')
```



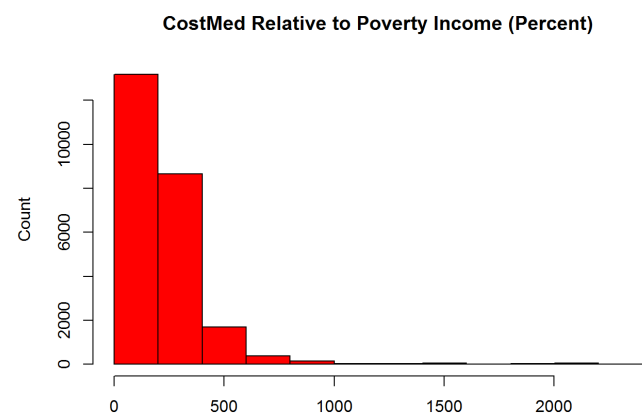
```
plot(hads2013n_c$FMTOWNRENT,hads2013n_c$COSTMedRELAMIPCT,xlab='FMTOWNRENT',ylab='COSTMedRELAMIPCT')
```



```
hist(hads2013n_c$COSTMedRELPOVPCT[which(hads2013n_c$FMTOWNRENT == "'1 Owner'")],main= 'CostMed Relative to Poverty Income (Pe  
rcent)', xlab='COSTMedRELPOVPCT', ylab='Count',col = 'blue')
```

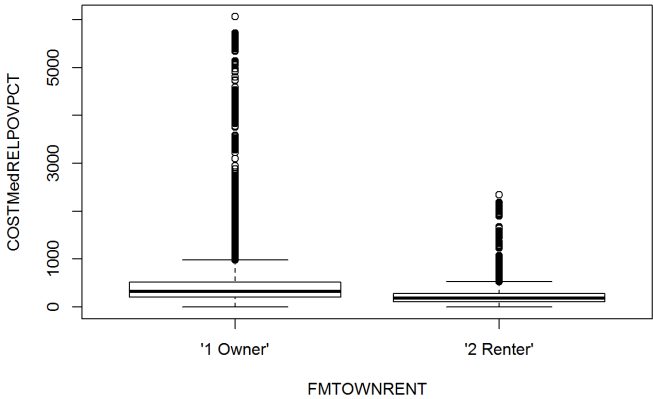


```
hist(hads2013n_c$COSTMedRELPOVPCT[which(hads2013n_c$FMTOWNRENT == "'2 Renter'")],main= 'CostMed Relative to Poverty Income (P  
ercent)', xlab='COSTMedRELPOVPCT', ylab='Count',col = 'red')
```



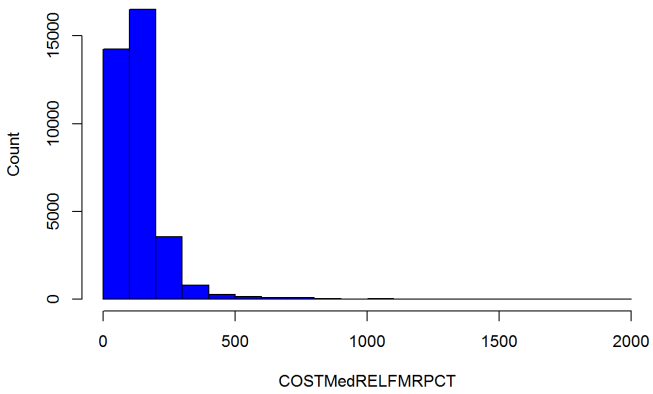
COSTMedRELPOVPCT

```
p1ot(hads2013n_c$FMTOWNRENT,hads2013n_c$COSTMedRELPOVPCT,xlab='FMTOWNRENT',ylab='COSTMedRELPOVPCT')
```

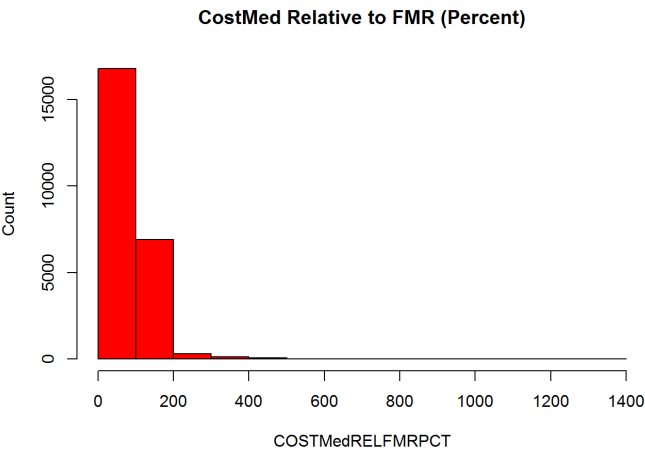


```
hist(hads2013n_c$COSTMedRELMPCT[which(hads2013n_c$FMTOWNRENT == "1 Owner")],main= 'CostMed Relative to FMR (Percent)', xlab='COSTMedRELMPCT', ylab='Count',col = 'blue')
```

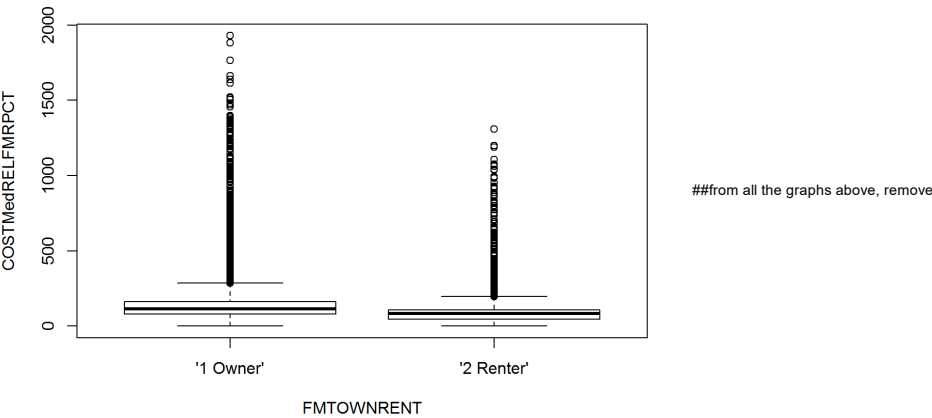
CostMed Relative to FMR (Percent)



```
hist(hads2013n_c$COSTMedRELMPCT[which(hads2013n_c$FMTOWNRENT == "2 Renter")],main= 'CostMed Relative to FMR (Percent)', xlab='COSTMedRELMPCT', ylab='Count',col = 'red')
```



```
plot(hads2013n_c$FMTOWNRENT,hads2013n_c$COSTMedRELFMRPCT,xlab='FMTOWNRENT',ylab='COSTMedRELFMRPCT')
```

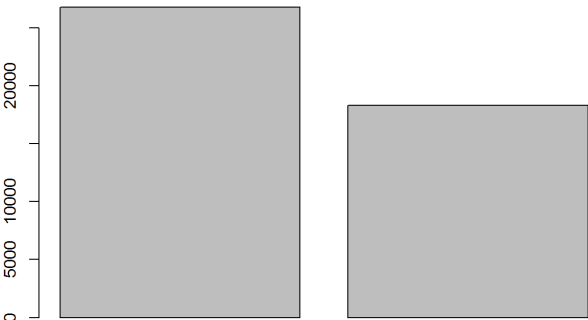


low significant attributes, ## atributes that contains excessive outlier

```
hads2013n_c <- subset(hads2013n_c,select = -c(REGION,FMTZADEQ,LMED, L30,L50,L80,IPOV,TYPE,NUNITS,ZINC2,OTHERCOST,COST06,COST12,COST08,COSTMED,GL30,GL50,GL80,APLMED,ABL50,BURDEN,INCRELAMIPCT,INCRELPOVPCT,INCRELFMRPCT,COST06RELAMIPCT,COST06RELPVPCT,COST06RELFMRPCT,COST08RELAMIPCT,COST08RELPVPCT,COST08RELFMRPCT,COST12RELAMIPCT,COST12RELPVPCT,COST12RELFMRPCT,COSTMedRELAMIPCT,COSTMedRELPVPCT,COSTMedRELFMRPCT))
```

Separate the dataset into 75% of training and 25%of testing

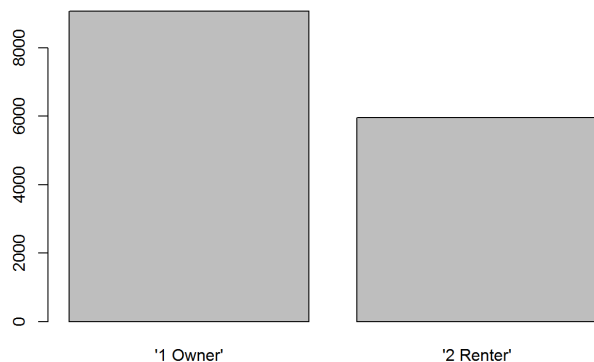
```
set.seed(888)
sample <- sample.int(n = nrow(hads2013n_c), size = floor(.75*nrow(hads2013n_c)), replace = F)
train <- hads2013n_c[sample, ]
test <- hads2013n_c[-sample, ]
plot(train$FMTOWNRENT)
```



'1 Owner'

'2 Renter'

```
plot(test$FMTOWNRENT)
```



Start to do the logistic regression on the training dataset

run the glm

```
fullmod = glm(FMTOWNRENT ~ ., data=train, family=binomial(link = "logit"))
```

```
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
```

```
summary(fullmod)
```

```
##
## Call:
## glm(formula = FMTOWNRENT ~ ., family = binomial(link = "logit"),
##      data = train)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -4.1506   0.0000   0.0000   0.0464   3.7001
##
## Coefficients: (1 not defined because of singularities)
##              Estimate Std. Error z value
## (Intercept)    9.604e+00  6.669e-01  14.401
## AGE1          -3.861e-02  2.358e-03 -16.375
## METRO3'2'      -2.924e-03  9.481e-02  -0.031
## METRO3'3'       1.174e-01  1.407e-01   0.834
## METRO3'4'      -9.494e-02  1.698e-01  -0.559
## METRO3'5'       1.363e-01  1.439e-01   0.947
## FMR            1.182e-03  2.667e-04   4.433
## ROOMS          -3.560e-01  4.106e-02  -8.670
## WEIGHT         -9.996e-05  3.105e-05  -3.219
## PER            2.163e-01  3.339e-02   6.476
## ZSMHC          3.184e-04  3.690e-05   8.629
## UTILITY        -5.940e-03  3.710e-04 -16.009
## TOTSAL         -2.418e-06  6.501e-07  -3.720
## GLMED          -6.353e-05  1.351e-05  -4.702
## ABL30          -2.435e-05  5.774e-05  -0.422
## ABL80          -3.543e-05  1.884e-05  -1.880
## ABLMED         5.240e-05  1.573e-05   3.332
## INCRELAMICAT   -5.771e-02  4.999e-02  -1.154
## FMTBUILT'1940-1959' -2.656e-01  1.213e-01  -2.189
## FMTBUILT'1960-1979'  4.990e-02  1.147e-01   0.435
## FMTBUILT'1980-1989'  6.038e-02  1.375e-01   0.439
## FMTBUILT'1990-1999'  1.680e-01  1.392e-01   1.207
## FMTBUILT'2000-2009'  3.297e-01  1.398e-01   2.359
## FMTBUILT'After 2010' -1.101e-01  3.044e-01  -0.362
## FMTSTRUCTURETYPE'2 2-4 units'  1.576e+00  1.515e-01  10.403
## FMTSTRUCTURETYPE'3 5-19 units'  1.834e+00  1.646e-01  11.148
## FMTSTRUCTURETYPE'4 20-49 units'  1.845e+00  2.016e-01   9.149
## FMTSTRUCTURETYPE'5 50+ units'  1.628e+00  1.530e-01  10.643
## FMTSTRUCTURETYPE'6 Mobile Home' -3.500e+00  1.356e-01 -25.812
## FMTBEDRMS'1 1BR'    -8.758e-03  4.140e-01  -0.021
## FMTBEDRMS'2 2BR'    -9.425e-01  4.487e-01  -2.101
## FMTBEDRMS'3 3BR'    -1.572e+00  5.156e-01  -3.050
## FMTBEDRMS'4 4BR+'   -1.600e+00  6.161e-01  -2.596
## FMTCOST06RELPVOCAT'2 100-150% Poverty'  1.323e+00  5.939e+03   0.000
```

```

## FMTCOST06RELPOVCAT'3 150-200% Poverty' 8.906e-01 6.625e+03 0.000
## FMTCOST06RELPOVCAT'4 200%+ Poverty' 2.810e+00 7.002e+03 0.000
## FMTCOST08RELPOVCAT'2 100-150% Poverty' 1.634e-01 5.750e+03 0.000
## FMTCOST08RELPOVCAT'3 150-200% Poverty' 1.224e+00 6.442e+03 0.000
## FMTCOST08RELPOVCAT'4 200%+ Poverty' 1.215e+00 6.764e+03 0.000
## FMTCOST12RELPOVCAT'2 100-150% Poverty' -2.400e+01 3.919e+03 -0.006
## FMTCOST12RELPOVCAT'3 150-200% Poverty' -4.768e+01 4.430e+03 -0.011
## FMTCOST12RELPOVCAT'4 200%+ Poverty' -7.114e+01 4.621e+03 -0.015
## FMTCOSTMEDRELPOVCAT'2 100-150% Poverty' 2.327e+01 4.192e+03 0.006
## FMTCOSTMEDRELPOVCAT'3 150-200% Poverty' 4.717e+01 4.691e+03 0.010
## FMTCOSTMEDRELPOVCAT'4 200%+ Poverty' 6.822e+01 4.963e+03 0.014
## FMTINCRELPOVCAT'2 100-150% Poverty' -1.296e-01 1.708e-01 -0.759
## FMTINCRELPOVCAT'3 150-200% Poverty' -4.443e-02 2.150e-01 -0.207
## FMTINCRELPOVCAT'4 200%+ Poverty' -1.057e-01 2.428e-01 -0.435
## FMTCOST06RELFMRCAT'2 50.1 - 100% FMR' 3.317e+00 4.060e+03 0.001
## FMTCOST06RELFMRCAT'3 GT FMR' 4.089e+00 4.321e+03 0.001
## FMTCOST08RELFMRCAT'2 50.1 - 100% FMR' -1.961e+00 3.961e+03 0.000
## FMTCOST08RELFMRCAT'3 GT FMR' -2.062e+00 4.190e+03 0.000
## FMTCOST12RELFMRCAT'2 50.1 - 100% FMR' -2.268e+01 2.515e+03 -0.009
## FMTCOST12RELFMRCAT'3 GT FMR' -3.991e+01 2.672e+03 -0.015
## FMTCOSTMEDRELFMRCAT'2 50.1 - 100% FMR' 2.137e+01 2.669e+03 0.008
## FMTCOSTMEDRELFMRCAT'3 GT FMR' 3.846e+01 2.873e+03 0.013
## FMTINCRELFMRCAT'2 50.1 - 100% FMR' -1.008e-01 1.654e-01 -0.609
## FMTINCRELFMRCAT'3 GT FMR' -4.723e-01 2.323e-01 -2.033
## FMTCOST06RELAMICAT'2 30 - 50% AMI' 1.464e+00 4.564e+03 0.000
## FMTCOST06RELAMICAT'3 50 - 60% AMI' 5.671e+00 4.985e+03 0.001
## FMTCOST06RELAMICAT'4 60 - 80% AMI' 7.861e+00 5.271e+03 0.001
## FMTCOST06RELAMICAT'5 80 - 100% AMI' 1.224e+01 5.505e+03 0.002
## FMTCOST06RELAMICAT'6 100 - 120% AMI' 4.619e+01 5.984e+03 0.008
## FMTCOST06RELAMICAT'7 120% AMI + ' 4.523e+01 6.404e+03 0.007
## FMTCOST08RELAMICAT'2 30 - 50% AMI' 8.348e-03 4.365e+03 0.000
## FMTCOST08RELAMICAT'3 50 - 60% AMI' 2.287e+00 4.827e+03 0.000
## FMTCOST08RELAMICAT'4 60 - 80% AMI' 3.719e-01 5.073e+03 0.000
## FMTCOST08RELAMICAT'5 80 - 100% AMI' 1.797e+00 5.279e+03 0.000
## FMTCOST08RELAMICAT'6 100 - 120% AMI' 2.910e+00 5.464e+03 0.001

## FMTCOST08RELAMICAT'7 120% AMI + ' -2.078e+00 5.755e+03 0.000
## FMTCOST12RELAMICAT'2 30 - 50% AMI' -2.301e+01 2.800e+03 -0.008
## FMTCOST12RELAMICAT'3 50 - 60% AMI' -4.715e+01 3.156e+03 -0.015
## FMTCOST12RELAMICAT'4 60 - 80% AMI' -6.984e+01 3.310e+03 -0.021
## FMTCOST12RELAMICAT'5 80 - 100% AMI' -9.563e+01 3.436e+03 -0.028
## FMTCOST12RELAMICAT'6 100 - 120% AMI' -1.206e+02 3.528e+03 -0.034
## FMTCOST12RELAMICAT'7 120% AMI + ' -1.590e+02 3.796e+03 -0.042
## FMTCOSTMEDRELAMICAT'2 30 - 50% AMI' 2.201e+01 3.101e+03 0.007
## FMTCOSTMEDRELAMICAT'3 50 - 60% AMI' 4.221e+01 3.393e+03 0.012
## FMTCOSTMEDRELAMICAT'4 60 - 80% AMI' 6.396e+01 3.606e+03 0.018
## FMTCOSTMEDRELAMICAT'5 80 - 100% AMI' 8.458e+01 3.774e+03 0.022
## FMTCOSTMEDRELAMICAT'6 100 - 120% AMI' 8.932e+01 4.213e+03 0.021
## FMTCOSTMEDRELAMICAT'7 120% AMI + ' 1.101e+02 4.573e+03 0.024
## FMTINCRELAMICAT'2 30 - 50% AMI' -2.807e-01 1.600e-01 -1.755
## FMTINCRELAMICAT'3 50 - 60% AMI' -6.720e-02 2.089e-01 -0.322
## FMTINCRELAMICAT'4 60 - 80% AMI' -1.558e-01 1.686e-01 -0.924
## FMTINCRELAMICAT'5 80 - 100% AMI' 1.337e-01 1.713e-01 0.780
## FMTINCRELAMICAT'6 100 - 120% AMI' -5.471e-02 1.462e-01 -0.374
## FMTINCRELAMICAT'7 120% AMI + ' NA NA NA
## FMTBURDEN'2 30% to 50%' -1.344e-01 1.082e-01 -1.242
## FMTBURDEN'3 50% or More' 5.949e-02 1.329e-01 0.448
## FMTBURDEN'4 No Income' -6.130e-01 2.367e-01 -2.590

## Pr(>|z|)
## (Intercept) < 2e-16 ***
## AGE1 < 2e-16 ***
## METRO3'2' 0.975398
## METRO3'3' 0.404093
## METRO3'4' 0.576160
## METRO3'5' 0.343403
## FMR 9.31e-06 ***
## ROOMS < 2e-16 ***
## WEIGHT 0.001287 ***
## PER 9.40e-11 ***
## ZSMHC < 2e-16 ***
## UTILITY < 2e-16 ***
## TOTSAL 0.000200 ***
## GLMED 2.58e-06 ***
## ABL30 0.673249
## ABL80 0.060059 .
## ABLMED 0.000862 ***
## INCRELAMICAT 0.248323
## FMTBUILT'1940-1959' 0.028565 *
## FMTBUILT'1960-1979' 0.663515
## FMTBUILT'1980-1989' 0.660606
## FMTBUILT'1990-1999' 0.227270
## FMTBUILT'2000-2009' 0.018348 *
## FMTBUILT'After 2010' 0.717527
## FMTSTRUCTURETYPE'2 2-4 units' < 2e-16 ***
## FMTSTRUCTURETYPE'3 5-19 units' < 2e-16 ***
## FMTSTRUCTURETYPE'4 20-49 units' < 2e-16 ***
## FMTSTRUCTURETYPE'5 50+ units' < 2e-16 ***
## FMTSTRUCTURETYPE'6 Mobile Home' < 2e-16 ***
## FMTBEDRMS'1 1BR' 0.983120
## FMTBEDRMS'2 2BR' 0.035660 *
## FMTBEDRMS'3 3BR' 0.002291 **
## FMTBEDRMS'4 4BR+' 0.009419 **
## FMTCOST06RELPOVCAT'2 100-150% Poverty' 0.999822
## FMTCOST06RELPOVCAT'3 150-200% Poverty' 0.999893
## FMTCOST06RELPOVCAT'4 200%+ Poverty' 0.999680
## FMTCOST08RELPOVCAT'2 100-150% Poverty' 0.999977
## FMTCOST08RELPOVCAT'3 150-200% Poverty' 0.999848
## FMTCOST08RELPOVCAT'4 200%+ Poverty' 0.999857
## FMTCOST12RELPOVCAT'2 100-150% Poverty' 0.995114
## FMTCOST12RELPOVCAT'3 150-200% Poverty' 0.991412
## FMTCOST12RELPOVCAT'4 200%+ Poverty' 0.987718
## FMTCOSTMEDRELPOVCAT'2 100-150% Poverty' 0.995571
## FMTCOSTMEDRELPOVCAT'3 150-200% Poverty' 0.991978
## FMTCOSTMEDRELPOVCAT'4 200%+ Poverty' 0.989034
## FMTINCRELPOVCAT'2 100-150% Poverty' 0.447855
## FMTINCRELPOVCAT'3 150-200% Poverty' 0.836328

```

```
## FMTINCRELPOVCAT'4 200%+ Poverty' 0.663220
## FMTCOST06RELFMRCAT'2 50.1 - 100% FMR' 0.999348
## FMTCOST06RELFMRCAT'3 GT FMR' 0.999245
## FMTCOST08RELFMRCAT'2 50.1 - 100% FMR' 0.999605
## FMTCOST08RELFMRCAT'3 GT FMR' 0.999607
## FMTCOST12RELFMRCAT'2 50.1 - 100% FMR' 0.992804
## FMTCOST12RELFMRCAT'3 GT FMR' 0.988082
## FMTCOSTMEDRELFMRCAT'2 50.1 - 100% FMR' 0.993612
## FMTCOSTMEDRELFMRCAT'3 GT FMR' 0.989322
## FMTINCRELFMRCAT'2 50.1 - 100% FMR' 0.542235
## FMTINCRELFMRCAT'3 GT FMR' 0.042010 *

## FMTCOST06RELAMICAT'2 30 - 50% AMI' 0.999744
## FMTCOST06RELAMICAT'3 50 - 60% AMI' 0.999092
## FMTCOST06RELAMICAT'4 60 - 80% AMI' 0.998810
## FMTCOST06RELAMICAT'5 80 - 100% AMI' 0.998225
## FMTCOST06RELAMICAT'6 100 - 120% AMI' 0.993842
## FMTCOST06RELAMICAT'7 120% AMI + ' 0.994365
## FMTCOST08RELAMICAT'2 30 - 50% AMI' 0.999998
## FMTCOST08RELAMICAT'3 50 - 60% AMI' 0.999622
## FMTCOST08RELAMICAT'4 60 - 80% AMI' 0.999942
## FMTCOST08RELAMICAT'5 80 - 100% AMI' 0.999728
## FMTCOST08RELAMICAT'6 100 - 120% AMI' 0.999575
## FMTCOST08RELAMICAT'7 120% AMI + ' 0.999712
## FMTCOST12RELAMICAT'2 30 - 50% AMI' 0.993443
## FMTCOST12RELAMICAT'3 50 - 60% AMI' 0.988078
## FMTCOST12RELAMICAT'4 60 - 80% AMI' 0.983170
## FMTCOST12RELAMICAT'5 80 - 100% AMI' 0.977797
## FMTCOST12RELAMICAT'6 100 - 120% AMI' 0.972736
## FMTCOST12RELAMICAT'7 120% AMI + ' 0.966590
## FMTCOSTMEDRELAMICAT'2 30 - 50% AMI' 0.994336
## FMTCOSTMEDRELAMICAT'3 50 - 60% AMI' 0.990074
## FMTCOSTMEDRELAMICAT'4 60 - 80% AMI' 0.985848
## FMTCOSTMEDRELAMICAT'5 80 - 100% AMI' 0.982118
## FMTCOSTMEDRELAMICAT'6 100 - 120% AMI' 0.983086
## FMTCOSTMEDRELAMICAT'7 120% AMI + ' 0.980786
## FMTINCRELAMICAT'2 30 - 50% AMI' 0.079332
## FMTINCRELAMICAT'3 50 - 60% AMI' 0.747708
## FMTINCRELAMICAT'4 60 - 80% AMI' 0.355447
## FMTINCRELAMICAT'5 80 - 100% AMI' 0.435234
## FMTINCRELAMICAT'6 100 - 120% AMI' 0.708215
## FMTINCRELAMICAT'7 120% AMI + ' NA
## FMTBURDEN'2 30% to 50%' 0.214195
## FMTBURDEN'3 50% or More' 0.654380
## FMTBURDEN'4 No Income' 0.009611 **
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 60874.2 on 45071 degrees of freedom
## Residual deviance: 5917.5 on 44982 degrees of freedom
## AIC: 6097.5
##
## Number of Fisher Scoring iterations: 22
```

```
#AIC 6097.5
```

```
## check the summary and notice lots of variance are significant
## and create reduce model by remove all non-significant variables
redmod = glm(FMTOWNRENT ~ AGE1 + FMR + ROOMS + WEIGHT + PER + ZSMHC +
  UTILITY + TOTSAL + GLMED + ABL80 + ABLMED + FMTINCRELAMICAT + FMTBURDEN, family=binomial(link = "logit"), data=train)
summary(redmod)
```

```
##
## Call:
## glm(formula = FMTOWNRENT ~ AGE1 + FMR + ROOMS + WEIGHT + PER +
## ZSMHC + UTILITY + TOTSAL + GLMED + ABL80 + ABLMED + FMTINCRELAMICAT +
## FMTBURDEN, family = binomial(link = "logit"), data = train)
##
## Deviance Residuals:
## Min 1Q Median 3Q Max
## -3.3290 -0.5621 -0.1848 0.4922 4.5650
##
## Coefficients:
## Estimate Std. Error z value Pr(>|z|)
```



```
## (Intercept)          6.621e+00  1.730e-01  38.275 < 2e-16
## AGE1                -4.962e-02  9.477e-04 -52.358 < 2e-16
## FMR                  7.564e-04  8.910e-05   8.489 < 2e-16
## ROOMS               -4.311e-01  1.752e-02 -24.604 < 2e-16
## WEIGHT              -2.853e-04  1.291e-05 -22.090 < 2e-16
## PER                 2.488e-01  1.156e-02  21.525 < 2e-16
## ZSMHC               -4.896e-05  2.419e-05  -2.024  0.043
## UTILITY             -7.143e-03  1.717e-04 -41.602 < 2e-16
## TOTSAL              -1.984e-06  4.150e-07  -4.782  1.74e-06
## GLMED               5.729e-05  3.144e-06  18.220 < 2e-16
## ABL80               -3.141e-06  4.706e-06  -0.667  0.504
## ABLMED              -7.067e-05  3.916e-06 -18.047 < 2e-16
## FMTINCRELAMICAT'2 30 - 50% AMI' -5.212e-01  5.157e-02 -10.107 < 2e-16
## FMTINCRELAMICAT'3 50 - 60% AMI' -6.985e-01  6.714e-02 -10.404 < 2e-16
## FMTINCRELAMICAT'4 60 - 80% AMI' -8.216e-01  6.277e-02 -13.088 < 2e-16
## FMTINCRELAMICAT'5 80 - 100% AMI' -9.562e-01  6.870e-02 -13.917 < 2e-16
## FMTINCRELAMICAT'6 100 - 120% AMI' -1.146e+00  7.178e-02 -15.962 < 2e-16
## FMTINCRELAMICAT'7 120% AMI + ' -1.467e+00  7.107e-02 -20.636 < 2e-16
## FMTBURDEN'2 30% to 50%'         3.066e-01  4.276e-02  7.170  7.49e-13
## FMTBURDEN'3 50% or More'        2.243e-01  5.704e-02  3.932  8.43e-05
## FMTBURDEN'4 No Income'        -5.139e-01  1.056e-01  -4.866  1.14e-06
##
## (Intercept)          ***
## AGE1                 ***
## FMR                  ***
## ROOMS                ***
## WEIGHT               ***
## PER                  ***
## ZSMHC                *
## UTILITY              ***
## TOTSAL               ***
## GLMED                ***
## ABL80                ***
## ABLMED               ***
## FMTINCRELAMICAT'2 30 - 50% AMI' ***
## FMTINCRELAMICAT'3 50 - 60% AMI' ***
## FMTINCRELAMICAT'4 60 - 80% AMI' ***
## FMTINCRELAMICAT'5 80 - 100% AMI' ***
## FMTINCRELAMICAT'6 100 - 120% AMI' ***
## FMTINCRELAMICAT'7 120% AMI + ' ***
## FMTBURDEN'2 30% to 50%'         ***
## FMTBURDEN'3 50% or More'        ***
## FMTBURDEN'4 No Income'        ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 60874 on 45071 degrees of freedom
## Residual deviance: 33444 on 45051 degrees of freedom
## AIC: 33486
##
## Number of Fisher Scoring iterations: 6
```

```
#AIC 33486
```

```
## go back and use backstep on fullmod
## #stepwise selection method is applying in next step.
backwards = step(fullmod) #backwards stepwise selection
```

```
## Start: AIC=6097.53
## FMTOWNRENT ~ AGE1 + METRO3 + FMR + ROOMS + WEIGHT + PER + ZSMHC +
## UTILITY + TOTSAL + GLMED + ABL30 + ABL80 + ABLMED + INCRELAMICAT +
## FMTBUILT + FMTSTRUCTURETYPE + FMTBEDRMS + FMTOST06RELPOVCAT +
## FMTOST08RELPOVCAT + FMTOST12RELPOVCAT + FMTOSTMEDRELPOVCAT +
## FMTINCRELPPOVCAT + FMTOST06RELFMRCAT + FMTOST08RELFMRCAT +
## FMTOST12RELFMRCAT + FMTOSTMEDRELFMRCAT + FMTINCRELFMRCAT +
## FMTOST06RELAMICAT + FMTOST08RELAMICAT + FMTOST12RELAMICAT +
## FMTOSTMEDRELAMICAT + FMTINCRELAMICAT + FMTBURDEN
```

```
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
```

[illegible]

```
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
```

	Df	Deviance	AIC
## - FMTCOST08RELAMICAT	6	5917.5	6085.5
## - FMTCOST06RELAMICAT	6	5917.5	6085.5
## - FMTCOST08RELPOVCAT	3	5917.5	6091.5
## - FMTCOST06RELPOVCAT	3	5917.5	6091.5
## - METRO3	4	5920.1	6092.1
## - FMTINCRELPOVCAT	3	5918.3	6092.3
## - FMTINCRELAMICAT	6	5924.3	6092.3
## - FMTCOST06RELFMRCAT	2	5917.5	6093.5
## - FMTCOST08RELFMRCAT	2	5917.5	6093.5
## - ABL30	1	5917.7	6095.7
## <none>		5917.5	6097.5
## - ABL80	1	5921.1	6099.1
## - FMTINCRELFMRCAT	2	5923.5	6099.5
## - FMTBURDEN	3	5928.1	6102.1
## - FMTBUILT	6	5937.5	6105.5
## - WEIGHT	1	5928.1	6106.1
## - ABLMED	1	5928.6	6106.6
## - TOTSAL	1	5931.8	6109.8
## - FMR	1	5937.0	6115.0
## - GLMED	1	5939.5	6117.5
## - PER	1	5959.8	6137.8
## - FMTBEDRMS	4	5968.7	6140.7
## - FMTCOSTMEDRELFMRCAT	2	5971.0	6147.0
## - ZSMHC	1	5990.2	6168.2
## - ROOMS	1	5995.3	6173.3
## - FMTCOST12RELFMRCAT	2	6022.0	6198.0
## - FMTCOSTMEDRELPOVCAT	3	6160.3	6334.3
## - AGE1	1	6202.7	6380.7
## - UTILITY	1	6205.5	6383.5
## - FMTCOSTMEDRELAMICAT	6	6238.3	6406.3
## - FMTCOST12RELPOVCAT	3	6648.9	6822.9
## - FMTSTRUCTURETYPE	5	7194.3	7364.3
## - FMTCOST12RELAMICAT	6	8517.2	8685.2

```
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
```

```
##
## Step: AIC=6085.53
## FMTOWNRENT ~ AGE1 + METRO3 + FMR + ROOMS + WEIGHT + PER + ZSMHC +
## UTILITY + TOTSAL + GLMED + ABL30 + ABL80 + ABLMED + FMTBUILT +
## FMTSTRUCTURETYPE + FMTBEDRMS + FMTCOST06RELPOVCAT + FMTCOST08RELPOVCAT +
## FMTCOST12RELPOVCAT + FMTCOSTMEDRELPOVCAT + FMTINCRELPOVCAT +
## FMTCOST06RELFMRCAT + FMTCOST08RELFMRCAT + FMTCOST12RELFMRCAT +
## FMTCOSTMEDRELFMRCAT + FMTINCRELFMRCAT + FMTCOST06RELAMICAT +
## FMTCOST12RELAMICAT + FMTCOSTMEDRELAMICAT + FMTINCRELAMICAT +
## FMTBURDEN
```

```
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
```



```
##
## Step: AIC=6073.53
## FMTOWNRENT ~ AGE1 + METRO3 + FMR + ROOMS + WEIGHT + PER + ZSMHC +
## UTILITY + TOTSL + GLMED + ABL30 + ABL80 + ABLMED + FMTBUILT +
## FMTSTRUCTURETYPE + FMTBEDRMS + FMTCOST06RELPOVCAT + FMTCOST08RELPOVCAT +
## FMTCOST12RELPOVCAT + FMTCOSTMEDRELPOVCAT + FMTINCRELPOVCAT +
## FMTCOST06RELFMRCAT + FMTCOST08RELFMRCAT + FMTCOST12RELFMRCAT +
## FMTCOSTMEDRELFMRCAT + FMTINCRELFMRCAT + FMTCOST12RELAMICAT +
## FMTCOSTMEDRELAMICAT + FMTINCRELAMICAT + FMTBURDEN
```

93/106

```
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
```

	Df	Deviance	AIC
## - FMTCOST08RELPOVCAT	3	5917.5	6067.5
## - FMTCOST06RELPOVCAT	3	5917.5	6067.5
## - METRO3	4	5920.1	6068.1
## - FMTINCRELPOVCAT	3	5918.3	6068.3
## - FMTINCRELAMICAT	6	5924.3	6068.3
## - FMTCOST06RELFMRCAT	2	5917.5	6069.5
## - FMTCOST08RELFMRCAT	2	5917.5	6069.5
## - ABL30	1	5917.7	6071.7
## <none>		5917.5	6073.5
## - ABL80	1	5921.1	6075.1
## - FMTINCRELFMRCAT	2	5923.5	6075.5
## - FMTBURDEN	3	5928.1	6078.1
## - FMTBUILT	6	5937.5	6081.5
## - WEIGHT	1	5928.1	6082.1
## - ABLMED	1	5928.6	6082.6
## - TOTSAL	1	5931.8	6085.8
## - FMR	1	5937.0	6091.0
## - GLMED	1	5939.5	6093.5
## - PER	1	5959.8	6113.8
## - FMTBEDRMS	4	5968.7	6116.7
## - FMTCOSTMEDRELFMRCAT	2	5971.0	6123.0
## - ZSMHC	1	5990.2	6144.2
## - ROOMS	1	5995.3	6149.3
## - FMTCOST12RELFMRCAT	2	6022.0	6174.0
## - FMTCOSTMEDRELPOVCAT	3	6160.3	6310.3
## - AGE1	1	6202.7	6356.7
## - UTILITY	1	6205.5	6359.5
## - FMTCOST12RELPOVCAT	3	6648.9	6798.9
## - FMTSTRUCTURETYPE	5	7194.3	7340.3
## - FMTCOSTMEDRELAMICAT	6	10830.8	10974.8
## - FMTCOST12RELAMICAT	6	15920.6	16064.6

```
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
```

```
##
## Step: AIC=6067.53
## FMTOWNRENT ~ AGE1 + METRO3 + FMR + ROOMS + WEIGHT + PER + ZSMHC +
## UTILITY + TOTSAL + GLMED + ABL30 + ABL80 + ABLMED + FMTBUILT +
## FMTSTRUCTURETYPE + FMTBEDRMS + FMTCOST06RELPOVCAT + FMTCOST12RELPOVCAT +
## FMTCOSTMEDRELPOVCAT + FMTINCRELPOVCAT + FMTCOST06RELFMRCAT +
## FMTCOST08RELFMRCAT + FMTCOST12RELFMRCAT + FMTCOSTMEDRELFMRCAT +
## FMTINCRELFMRCAT + FMTCOST12RELAMICAT + FMTCOSTMEDRELAMICAT +
## FMTINCRELAMICAT + FMTBURDEN
```



```
##  
## Step: AIC=6061.53  
## FMTOWNRENT + AGE1 + METRO3 + FMR + ROOMS + WEIGHT + PER + ZSMHC +  
## UTILITY + TOTSL + GLMED + ABL30 + ABL80 + ABUMED + FMTBUILT +  
## FMTSTRUCTURETYPE + FMTBEDRMS + FMTCOST12RELPOVCAT + FMTCOSTMEDRELPOVCAT +  
## FMTINCRELPPOVCAT + FMTCOST06RELFMRCAT + FMTCOST198RELFMRCAT +  
## FMTCOST12RELFMRCAT + FMTCOSTMEDRELFMRCAT + FMTINCRELFMRCAT +  
## FMTCOST12RELAMICAT + FMTCOSTMEDRELAMICAT + FMTINCRELAMICAT +  
## FMTBURDEN
```

file:///A:/Users/Thomason PC/Desktop/Ryerson/CKME136/Capstone Project R nextp.html



```
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
```

```
##           Df Deviance   AIC
## - METRO3      4  5920.1 6056.1
## - FMTINCRELPVOCAT 3  5918.3 6056.3
## - FMTINCRELAMICAT 6  5924.3 6056.3
## - FMTCOST06RELFMRCAT 2  5917.5 6057.5
## - FMTCOST08RELFMRCAT 2  5917.5 6057.5
## - ABL30        1  5917.7 6059.7
## <none>          5917.5 6061.5
## - ABL80        1  5921.1 6063.1
## - FMTINCRELFMRCAT 2  5923.5 6063.5
## - FMTBURDEN    3  5928.1 6066.1
## - FMTBUILT     6  5937.5 6069.5
## - WEIGHT      1  5928.1 6070.1
## - ABLMED      1  5928.6 6070.6
## - TOTSAL      1  5931.8 6073.8
## - FMR         1  5937.0 6079.0
## - GLMED       1  5939.5 6081.5
## - PER         1  5959.8 6101.8
## - FMTBEDRMS   4  5968.7 6104.7
## - FMTCOSTMEDRELFMRCAT 2  5971.0 6111.0
## - ZSMHC       1  5990.2 6132.2
## - ROOMS       1  5995.3 6137.3
## - FMTCOST12RELFMRCAT 2  6022.0 6162.0
## - AGE1        1  6202.7 6344.7
## - UTILITY     1  6205.5 6347.5
## - FMTCOST12RELPOVCAT 3  7167.6 7305.6
## - FMTSTRUCTURETYPE 5  7194.3 7328.3
## - FMTCOSTMEDRELPOVCAT 3  7338.9 7476.9
## - FMTCOSTMEDRELAMICAT 6 10830.8 10962.8
## - FMTCOST12RELAMICAT 6 15920.6 16052.6
```

```
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
```

```
##
## Step: AIC=6056.11
## FMTOWNRENT ~ AGE1 + FMR + ROOMS + WEIGHT + PER + ZSMHC + UTILITY +
##   TOTSAL + GLMED + ABL30 + ABL80 + ABLMED + FMTBUILT + FMTSTRUCTURETYPE +
##   FMTBEDRMS + FMTCOST12RELPOVCAT + FMTCOSTMEDRELPOVCAT + FMTINCRELPVOCAT +
##   FMTCOST06RELFMRCAT + FMTCOST08RELFMRCAT + FMTCOST12RELFMRCAT +
##   FMTCOSTMEDRELFMRCAT + FMTINCRELFMRCAT + FMTCOST12RELAMICAT +
##   FMTCOSTMEDRELAMICAT + FMTINCRELAMICAT + FMTBURDEN
```

[illegible]

##	Df	Deviance	AIC
## - FMTINCRELPVOCAT	3	5920.9	6050.9
## - FMTINCRELAMICAT	6	5926.9	6050.9
## - FMTCOST06RELFMRCAT	2	5920.1	6052.1
## - FMTCOST08RELFMRCAT	2	5920.1	6052.1
## - ABL30	1	5920.2	6054.2
## <none>		5920.1	6056.1
## - FMTINCRELFMRCAT	2	5926.2	6058.2
## - ABL80	1	5924.4	6058.4
## - FMTBURDEN	3	5930.6	6060.6

```

## - WEIGHT      1  5929.4  6063.4
## - FMTBUILT     6  5940.9  6064.9
## - ABLMED       1  5930.9  6064.9
## - TOTSAL       1  5934.4  6068.4
## - FMR          1  5938.8  6072.8
## - GLMED        1  5942.5  6076.5
## - PER          1  5962.6  6096.6
## - FMTBEDRMS    4  5970.7  6098.7
## - FMTCOSTMEDRELFMRCAT 2  5973.6  6105.6
## - ZSMHC        1  5992.8  6126.8
## - ROOMS        1  5997.6  6131.6
## - FMTCOST12RELFMRCAT 2  6024.4  6156.4
## - AGE1         1  6206.1  6340.1
## - UTILITY      1  6209.9  6343.9
## - FMTCOST12RELPOVCAT 3  7173.9  7303.9
## - FMTSTRUCTURETYPE 5  7284.7  7410.7
## - FMTCOSTMEDRELPOVCAT 3  7343.3  7473.3
## - FMTCOSTMEDRELAMICAT 6 10848.4 10972.4
## - FMTCOST12RELAMICAT 6 15973.0 16097.0

```

```
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
```

```

##
## Step: AIC=6050.87
## FMTOWNRENT ~ AGE1 + FMR + ROOMS + WEIGHT + PER + ZSMHC + UTILITY +
##   TOTSAL + GLMED + ABL30 + ABL80 + ABLMED + FMTBUILT + FMTSTRUCTURETYPE +
##   FMTBEDRMS + FMTCOST12RELPOVCAT + FMTCOSTMEDRELPOVCAT + FMTCOST06RELFMRCAT +
##   FMTCOST08RELFMRCAT + FMTCOST12RELFMRCAT + FMTCOSTMEDRELFMRCAT +
##   FMTINCRELFMRCAT + FMTCOST12RELAMICAT + FMTCOSTMEDRELAMICAT +
##   FMTINCRELAMICAT + FMTBURDEN

```

```

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

```

```
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
```

	Df	Deviance	AIC
## - FMTCOST06RELFMRCAT	2	5920.9	6046.9
## - FMTCOST08RELFMRCAT	2	5920.9	6046.9
## - FMTINCRELAMICAT	6	5930.2	6048.2
## - ABL30	1	5921.0	6049.0
## <none>		5920.9	6050.9
## - ABL80	1	5925.2	6053.2
## - FMTINCRELFMRCAT	2	5928.1	6054.1
## - FMTBURDEN	3	5931.2	6055.2
## - WEIGHT	1	5930.2	6058.2
## - FMTBUILT	6	5941.5	6059.5
## - ABLMED	1	5931.8	6059.8
## - TOTSAL	1	5935.1	6063.1
## - FMR	1	5939.5	6067.5
## - GLMED	1	5943.4	6071.4
## - FMTBEDRMS	4	5971.4	6093.4
## - PER	1	5968.0	6096.0
## - FMTCOSTMEDRELFMRCAT	2	5974.5	6100.5
## - ZSMHC	1	5993.9	6121.9
## - ROOMS	1	5998.3	6126.3
## - FMTCOST12RELFMRCAT	2	6025.3	6151.3
## - UTILITY	1	6211.0	6339.0
## - AGE1	1	6212.3	6340.3
## - FMTCOST12RELPOVCAT	3	7182.5	7306.5
## - FMTSTRUCTURETYPE	5	7287.0	7407.0
## - FMTCOSTMEDRELPOVCAT	3	7347.0	7471.0
## - FMTCOSTMEDRELAMICAT	6	10855.5	10973.5
## - FMTCOST12RELAMICAT	6	15986.7	16104.7

```
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
```

```
##
## Step: AIC=6046.87
## FMTOWNRENT ~ AGE1 + FMR + ROOMS + WEIGHT + PER + ZSMHC + UTILITY +
## TOTSAL + GLMED + ABL30 + ABL80 + ABLMED + FMTBUILT + FMTSTRUCTURETYPE +
## FMTBEDRMS + FMTCOST12RELPOVCAT + FMTCOSTMEDRELPOVCAT + FMTCOST08RELFMRCAT +
## FMTCOST12RELFMRCAT + FMTCOSTMEDRELFMRCAT + FMTINCRELFMRCAT +
## FMTCOST12RELAMICAT + FMTCOSTMEDRELAMICAT + FMTINCRELAMICAT +
## FMTBURDEN
```

[illegible]

##		Df	Deviance	AIC
##	- FMTCOST08RELFMRCAT	2	5920.9	6042.9
##	- FMTINCRELFMRCAT	6	5930.2	6044.2
##	- ABL30	1	5921.0	6045.0
##	<none>		5920.9	6046.9
##	- ABL80	1	5925.2	6049.2
##	- FMTINCRELFMRCAT	2	5928.1	6050.1
##	- FMTBURDEN	3	5931.2	6051.2
##	- WEIGHT	1	5930.2	6054.2
##	- FMTBUILT	6	5941.5	6055.5
##	- ABLMED	1	5931.8	6055.8
##	- TOTSAL	1	5935.1	6059.1
##	- FMR	1	5939.5	6063.5
##	- GLMED	1	5943.4	6067.4
##	- FMTBEDRMS	4	5971.4	6089.4
##	- PER	1	5968.0	6092.0
##	- ZSMHC	1	5993.9	6117.9
##	- ROOMS	1	5998.3	6122.3
##	- FMTCOST12RELFMRCAT	2	6025.3	6147.3
##	- FMTCOSTMEDRELFMRCAT	2	6070.1	6192.1
##	- UTILITY	1	6211.0	6335.0
##	- AGE1	1	6212.3	6336.3
##	- FMTCOST12RELPOVCAT	3	7182.5	7302.5
##	- FMTSTRUCTURETYPE	5	7287.0	7403.0
##	- FMTCOSTMEDRELPOVCAT	3	7347.0	7467.0
##	- FMTCOSTMEDRELAMICAT	6	10855.5	10969.5
##	- FMTCOST12RELAMICAT	6	15986.7	16108.7

```
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
```



```
##
## Step: AIC=6040.2
## FMTOWNRRENT ~ AGE1 + FMR + ROOMS + WEIGHT + PER + ZSMHC + UTILITY +
## TOTAL + GLMED + ABL30 + ABL80 + ABLMED + FMTBUILT + FMTSTRUCTURETYPE +
## FMTBEDRMS + FMTCOST12RELPOVCAT + FMTCOSTMEDRELPOVCAT + FMTCOST12RELFMRCAT +
## FMTCOSTMEDRELFMRCAT + FMTINCRELFMRCAT + FMTCOST12RELAMICAT +
## FMTCOSTMEDRELAMICAT + FMTBURDEN
```

##	Df	Deviance	AIC
## - ABL30	1	5930.4	6038.4
## <none>		5930.2	6040.2

##		Df	Deviance	AIC
##	<none>		5930.4	6038.4
##	- FMTBURDEN	3	5940.3	6042.3
##	- WEIGHT	1	5940.1	6046.1
##	- FMTBUILT	6	5951.8	6047.8
##	- ABLMED	1	5942.9	6048.9
##	- ABL80	1	5943.0	6049.0
##	- TOTSAL	1	5945.5	6051.5
##	- FMR	1	5950.8	6056.8
##	- GLMED	1	5952.8	6058.8
##	- FMTINCRFLMRCAT	2	5962.5	6066.5
##	- FMTBEDRMS	4	5983.1	6083.1
##	- PER	1	5983.2	6089.2
##	- ZSMHC	1	5999.9	6105.9
##	- ROOMS	1	6009.7	6115.7
##	- FMT COST212RELFLMRCAT	2	6144.1	6248.1
##	- FMT COST2MEDRELFLMRCAT	2	6180.4	6284.4
##	- UTILITY	1	6223.4	6329.4
##	- AGE1	1	6223.6	6329.6
##	- FMT COST212RELPOVCAT	3	7202.1	7304.1



```
## - FMTSTRUCTURETYPE      5      7302.0  7400.0
## - FMTCOSTMEDRELPOVCAT    3      7360.2  7462.2
## - FMTCOSTMEDRELAMICAT    6     10863.2 10959.2
## - FMTCOSTI2RELAMICAT     6     16012.8 16108.8
```

```
summary(backwards)
```

```
##
## Call:
## glm(formula = FMTOWNRENT ~ AGE1 + FMR + ROOMS + WEIGHT + PER +
##       ZSMHC + UTILITY + TOTSAL + GLMED + ABL80 + ABLMED + FMTBUILT +
##       FMTSTRUCTURETYPE + FMTBEDRMS + FMTCOSTI2RELPOVCAT + FMTCOSTMEDRELPOVCAT +
##       FMTCOSTI2RELFMRCAT + FMTCOSTMEDRELFMRCAT + FMTINCRELFMRCAT +
##       FMTCOSTI2RELAMICAT + FMTCOSTMEDRELAMICAT + FMTBURDEN, family = binomial(link = "logit"),
##       data = train)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -4.147    0.000    0.000    0.047    3.713
##
## Coefficients:
##              Estimate Std. Error z value
## (Intercept)      9.449e+00  6.285e-01  15.034
## AGE1             -3.845e-02  2.317e-03 -16.595
## FMR               1.075e-03  2.377e-04  4.522
## ROOMS            -3.576e-01  4.086e-02  -8.752
## WEIGHT           -9.366e-05  3.031e-05  -3.090
## PER              2.232e-01  3.086e-02  7.233
## ZSMHC             3.066e-04  3.633e-05  8.439
## UTILITY          -5.969e-03  3.696e-04 -16.150
## TOTSAL           -2.397e-06  6.270e-07  -3.823
## GLMED            -6.339e-05  1.334e-05  -4.752
## ABL80            -4.174e-05  1.175e-05  -3.554
## ABLMED            5.068e-05  1.426e-05  3.553
## FMTBUILT'1940-1959' -2.781e-01  1.201e-01  -2.316
## FMTBUILT'1960-1979'  5.018e-02  1.125e-01  0.446
## FMTBUILT'1980-1989'  6.123e-02  1.346e-01  0.455
## FMTBUILT'1990-1999'  1.753e-01  1.366e-01  1.284
## FMTBUILT'2000-2009'  3.289e-01  1.378e-01  2.386
## FMTBUILT'After 2010' -1.150e-01  3.027e-01  -0.380
## FMTSTRUCTURETYPE'2 2-4 units' 1.552e+00  1.490e-01  10.417
## FMTSTRUCTURETYPE'3 5-19 units' 1.808e+00  1.616e-01  11.191
## FMTSTRUCTURETYPE'4 20-49 units' 1.802e+00  1.977e-01  9.117
## FMTSTRUCTURETYPE'5 50+ units'  1.597e+00  1.478e-01  10.803
## FMTSTRUCTURETYPE'6 Mobile Home' -3.467e+00  1.324e-01 -26.193
## FMTBEDRMS'1 1BR'      8.353e-03  4.115e-01  0.020
## FMTBEDRMS'2 2BR'     -9.277e-01  4.454e-01  -2.083
## FMTBEDRMS'3 3BR'     -1.543e+00  5.093e-01  -3.030
## FMTBEDRMS'4 4BR'     -1.559e+00  6.068e-01  -2.569
## FMTCOSTI2RELPOVCAT'2 100-150% Poverty' -2.368e+01  2.369e+03  -0.010
## FMTCOSTI2RELPOVCAT'3 150-200% Poverty' -4.688e+01  2.643e+03  -0.018
## FMTCOSTI2RELPOVCAT'4 200%+ Poverty'    -6.996e+01  2.772e+03  -0.025
## FMTCOSTMEDRELPOVCAT'2 100-150% Poverty' 2.442e+01  2.369e+03  0.010
## FMTCOSTMEDRELPOVCAT'3 150-200% Poverty' 4.847e+01  2.643e+03  0.018
## FMTCOSTMEDRELPOVCAT'4 200%+ Poverty'    7.103e+01  2.772e+03  0.026
## FMTCOSTI2RELFMRCAT'2 50.1 - 100% FMR' -2.350e+01  1.598e+03  -0.015
## FMTCOSTI2RELFMRCAT'3 GT FMR'          -4.061e+01  1.690e+03  -0.024
## FMTCOSTMEDRELFMRCAT'2 50.1 - 100% FMR' 2.356e+01  1.598e+03  0.015
## FMTCOSTMEDRELFMRCAT'3 GT FMR'          4.119e+01  1.690e+03  0.024
## FMTINCRELFMRCAT'2 50.1 - 100% FMR'    -2.948e-01  1.074e-01  -2.746
## FMTINCRELFMRCAT'3 GT FMR'            -6.916e-01  1.236e-01  -5.594
## FMTCOSTI2RELAMICAT'2 30 - 50% AMI'     -2.275e+01  1.786e+03  -0.013
## FMTCOSTI2RELAMICAT'3 50 - 60% AMI'     -4.591e+01  1.975e+03  -0.023
## FMTCOSTI2RELAMICAT'4 60 - 80% AMI'     -6.954e+01  2.082e+03  -0.033
## FMTCOSTI2RELAMICAT'5 80 - 100% AMI'    -9.451e+01  2.160e+03  -0.044
## FMTCOSTI2RELAMICAT'6 100 - 120% AMI'   -1.189e+02  2.235e+03  -0.053
## FMTCOSTI2RELAMICAT'7 120% AMI + '     -1.614e+02  2.428e+03  -0.066
## FMTCOSTMEDRELAMICAT'2 30 - 50% AMI'    2.317e+01  1.786e+03  0.013
## FMTCOSTMEDRELAMICAT'3 50 - 60% AMI'    4.887e+01  1.975e+03  0.025
## FMTCOSTMEDRELAMICAT'4 60 - 80% AMI'    7.185e+01  2.082e+03  0.035
## FMTCOSTMEDRELAMICAT'5 80 - 100% AMI'   9.745e+01  2.160e+03  0.045
## FMTCOSTMEDRELAMICAT'6 100 - 120% AMI'  1.364e+02  2.342e+03  0.058
## FMTCOSTMEDRELAMICAT'7 120% AMI + '    1.556e+02  2.428e+03  0.064
## FMTBURDEN'2 30% to 50%'               -1.085e-01  1.067e-01  -1.017
## FMTBURDEN'3 50% or More'              1.534e-01  1.255e-01  1.222
## FMTBURDEN'4 No Income'                -4.489e-01  2.275e-01  -1.973
##
##              Pr(>|z|)
## (Intercept)      < 2e-16 ***
## AGE1             < 2e-16 ***
## FMR              6.14e-06 ***
## ROOMS            < 2e-16 ***
## WEIGHT           0.002004 **
## PER              4.74e-13 ***
## ZSMHC            < 2e-16 ***
## UTILITY          < 2e-16 ***
## TOTSAL           0.000132 ***
## GLMED            2.01e-06 ***
##
## ABL80            0.000380 ***
## ABLMED           0.000380 ***
## FMTBUILT'1940-1959' 0.020558 *
## FMTBUILT'1960-1979' 0.655596
## FMTBUILT'1980-1989' 0.649068
## FMTBUILT'1990-1999' 0.199204
## FMTBUILT'2000-2009' 0.017026 *
## FMTBUILT'After 2010' 0.703965
## FMTSTRUCTURETYPE'2 2-4 units' < 2e-16 ***
## FMTSTRUCTURETYPE'3 5-19 units' < 2e-16 ***
## FMTSTRUCTURETYPE'4 20-49 units' < 2e-16 ***
## FMTSTRUCTURETYPE'5 50+ units' < 2e-16 ***
## FMTSTRUCTURETYPE'6 Mobile Home' < 2e-16 ***
## FMTBEDRMS'1 1BR'   0.983803
## FMTBEDRMS'2 2BR'   0.037296 *
## FMTBEDRMS'3 3BR'   0.002442 **
## FMTBEDRMS'4 4BR'   0.010212 *
```

```
## FMTCOST12RELPOVCAT'2 100-150% Poverty' 0.992025
## FMTCOST12RELPOVCAT'3 150-200% Poverty' 0.985845
## FMTCOST12RELPOVCAT'4 200%+ Poverty' 0.979861
## FMTCOSTMEDRELPOVCAT'2 100-150% Poverty' 0.991775
## FMTCOSTMEDRELPOVCAT'3 150-200% Poverty' 0.985366
## FMTCOSTMEDRELPOVCAT'4 200%+ Poverty' 0.979553
## FMTCOST12RELFMRCAT'2 50.1 - 100% FMR' 0.988269
## FMTCOST12RELFMRCAT'3 GT FMR' 0.980827
## FMTCOSTMEDRELFMRCAT'2 50.1 - 100% FMR' 0.988239
## FMTCOSTMEDRELFMRCAT'3 GT FMR' 0.980553
## FMTINCRELFMRCAT'2 50.1 - 100% FMR' 0.006040 **
## FMTINCRELFMRCAT'3 GT FMR' 2.22e-08 ***
## FMTCOST12RELAMICAT'2 30 - 50% AMI' 0.989838
## FMTCOST12RELAMICAT'3 50 - 60% AMI' 0.981452
## FMTCOST12RELAMICAT'4 60 - 80% AMI' 0.973357
## FMTCOST12RELAMICAT'5 80 - 100% AMI' 0.965102
## FMTCOST12RELAMICAT'6 100 - 120% AMI' 0.957587
## FMTCOST12RELAMICAT'7 120% AMI + ' 0.947012
## FMTCOSTMEDRELAMICAT'2 30 - 50% AMI' 0.989647
## FMTCOSTMEDRELAMICAT'3 50 - 60% AMI' 0.980256
## FMTCOSTMEDRELAMICAT'4 60 - 80% AMI' 0.972474
## FMTCOSTMEDRELAMICAT'5 80 - 100% AMI' 0.964016
## FMTCOSTMEDRELAMICAT'6 100 - 120% AMI' 0.953564
## FMTCOSTMEDRELAMICAT'7 120% AMI + ' 0.948897
## FMTBURDEN'2 30% to 50%' 0.308950
## FMTBURDEN'3 50% or More' 0.221566
## FMTBURDEN'4 No Income' 0.048496 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 60874.2 on 45071 degrees of freedom
## Residual deviance: 5930.4 on 45018 degrees of freedom
## AIC: 6038.4
##
## Number of Fisher Scoring iterations: 22
```

## After the Logistic is created, compare it with the test data

```
predTst <- predict(backwards, test, type="response")
thresh <- 0.5 # threshold for categorizing predicted probabilities
predFac <- cut(predTst, breaks=c(-Inf, thresh, Inf), labels=c("'1 Owner'", "'2 Renter'"))
cTab <- table(test$FMTOWNRENT, predFac, dnn=c("actual", "predicted"))
addmargins(cTab)
```

```
##          predicted
## actual    '1 Owner' '2 Renter' Sum
## '1 Owner'      8838      233  9071
## '2 Renter'      154      5799 5953
## Sum           8992      6032 15024
```

```
addmargins(prop.table(cTab))
```

```
##          predicted
## actual    '1 Owner' '2 Renter' Sum
## '1 Owner' 0.58825879 0.01550852 0.60376731
## '2 Renter' 0.01025027 0.38598243 0.39623269
## Sum      0.59850905 0.40149095 1.00000000
```

```
## Predictive rate
spt <- prop.table(cTab)
sp = spt[1,1]+spt[2,2]
sp*100
```

```
## [1] 97.42412
```