

## ##Syntax

```
> BRFSS2013_Data <- read.csv("C:/Users/James/Desktop/BRFSS2013_Data.csv")
> View(BRFSS2013_Data)
> names(BRFSS2013_Data)
 [1] "X"          "X_STATE"    "NUMADULT"   "MENTHLTH"   "ADDEPEV2"
 [6] "MARITAL"    "CHILDREN"   "EDUCA"      "EMPLOY1"    "INCOME2"
[11] "RENTHOM1"   "SEX"        "QLMENTL2"   "X_STSTR"    "X_STRWT"
[16] "X_RAWRAKE"  "X_WT2RAKE"  "X_IMPRACE"  "X_LLCPWT2"  "X_LLCPWT"
[21] "X_PRACE1"   "X_HISPANC"  "X_RACE"     "X_RACEG21"  "X_RACEGR3"
[26] "X_RACE_G1"  "X_AGE5YR"   "X_AGE_G"    "X_CHLDCNT"  "X_EDUCAG"
[31] "X_INCOMG"   "X_AGE80"    "X_IMPEDUC"  "X_IMPMRTL"  "X_IMPHOME"
```

## ##CLEANING UP DATA OBSERVATIONS

### ## CLEANING OUTCOME (MENTHLTH):

```
> BRFSS2013_Data$MENTHLTH[BRFSS2013_Data$MENTHLTH==88] <- 0
> BRFSS2013_Data$MENTHLTH[BRFSS2013_Data$MENTHLTH==77] <- NA
> BRFSS2013_Data$MENTHLTH[BRFSS2013_Data$MENTHLTH==99] <- NA
> summary(BRFSS2013_Data$MENTHLTH)
   Min. 1st Qu.  Median    Mean 3rd Qu.  Max.    NA's
0.000  0.000  0.000  3.371  2.000 30.000   8635
```

### ## CLEANING EXPOSURE (RENTHOM1):

```
> BRFSS2013_Data$RENTHOM1[BRFSS2013_Data$RENTHOM1==7] <- NA
> BRFSS2013_Data$RENTHOM1[BRFSS2013_Data$RENTHOM1==9] <- NA
```

### ## CLEANING HOUSEHOLD SIZE:

```
> BRFSS2013_Data$CHILDREN[BRFSS2013_Data$CHILDREN==88] <- 0
> summary(BRFSS2013_Data$CHILDREN)
   Min. 1st Qu.  Median    Mean 3rd Qu.  Max.    NA's
0.0000 0.0000 0.0000 0.8823 1.0000 99.0000   437
> BRFSS2013_Data$CHILDREN[BRFSS2013_Data$CHILDREN==99] <- NA
> summary(BRFSS2013_Data$CHILDREN)
   Min. 1st Qu.  Median    Mean 3rd Qu.  Max.    NA's
0.0000 0.0000 0.0000 0.5163 1.0000 24.0000   2263
> summary(BRFSS2013_Data$NUMADULT)
   Min. 1st Qu.  Median    Mean 3rd Qu.  Max.    NA's
1.00  1.00  2.00  1.81  2.00  45.00 131709
> BRFSS2013_Data["HOUSEHOLD SIZE"] <- NA
> BRFSS2013_Data$HOUSEHOLD SIZE <-
BRFSS2013_Data$NUMADULT+BRFSS2013_Data$CHILDREN
```

### ## CLEANING HOUSEHOLD INCOME

```
> BRFSS2013_Data$INCOME2[BRFSS2013_Data$INCOME2==77] <- 99
> BRFSS2013_Data$INCOME2[is.na(BRFSS2013_Data$INCOME2)] <- 99
```

### ## CLEANING EDUCATION

```
> BRFSS2013_Data$EDUCA[BRFSS2013_Data$EDUCA==3] <- NA
```

```
> BRFSS2013_Data$EDUCA[BRFSS2013_Data$EDUCA==9] <- NA
```

### ## CLEANING UP EMPLOYMENT

```
> BRFSS2013_Data$EMPLOY1[BRFSS2013_Data$EMPLOY1==6] <- NA
```

```
> BRFSS2013_Data$EMPLOY1[BRFSS2013_Data$EMPLOY1==9] <- NA
```

### ## CLEANING UP MARITAL STATUS

```
> BRFSS2013_Data$MARITAL[BRFSS2013_Data$MARITAL==6] <- NA
```

```
> BRFSS2013_Data$MARITAL[BRFSS2013_Data$MARITAL==9] <- NA
```

### ##SINGLING OUT STATES WE NEED AND AGES

```
> ELIGIBLE <-
```

```
BRFSS2013_Data[which(BRFSS2013_Data$X_STATE==1|BRFSS2013_Data$X_STATE==5|BRFSS2013_Data$X_STATE==6|BRFSS2013_Data$X_STATE==15|BRFSS2013_Data$X_STATE==17|BRFSS2013_Data$X_STATE==20|BRFSS2013_Data$X_STATE==22|BRFSS2013_Data$X_STATE==31|BRFSS2013_Data$X_STATE==35|BRFSS2013_Data$X_STATE==40|BRFSS2013_Data$X_STATE==45|BRFSS2013_Data$X_STATE==55),]
```

```
> ELIGIBLE1 <-subset(ELIGIBLE, X_AGE80 >=35)
```

### ##CREATE NEW DATASET.

```
>write.csv(ELIGIBLE1, "C:/Users/James/Desktop/ELIGIBLE1.csv")
```

### ##IMPORT ELIGIBLE1 DATASET INTO R

```
> ELIGIBLE1 <- read.csv("C:/Users/James/Desktop/ELIGIBLE1.csv")
```

```
> View(ELIGIBLE1)
```

### ##SELECTING DATA:

```
> cleanobservations <- ELIGIBLE1[c(2,3,5,7,9:13,26,29,33,37)]
```

```
> names(cleanobservations)
```

```
[1] "X" "X_STATE" "MENTHLTH"  
[4] "MARITAL" "EDUCA" "EMPLOY1"  
[7] "INCOME2" "RENTHOM1" "SEX"  
[10] "X_RACEGR3" "X_AGE_G" "X_AGE80"  
[13] "HOUSEHOLD SIZE"
```

```
##EXCLUDING MISSING CASES FOR EDUCA, MARITAL, HOUSEHOLD SIZE, EMPLOY1
```

```
final <-  
cleanobservations[complete.cases(cleanobservations[c("EDUCA", "MARITAL", "HOUSE  
HOLD SIZE", "EMPLOY1"))],]
```

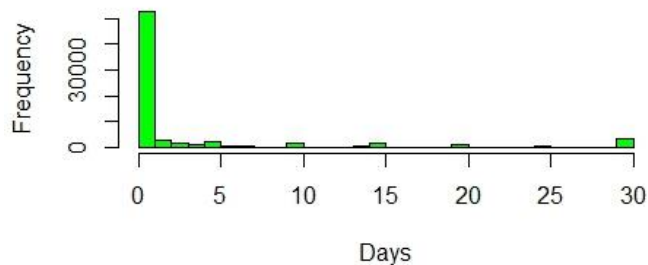
```
##CREATING A NEW AND SMALLER DATASET.
```

```
> write.csv(final, "C:/Users/James/Desktop/NEWDATA.csv")
```

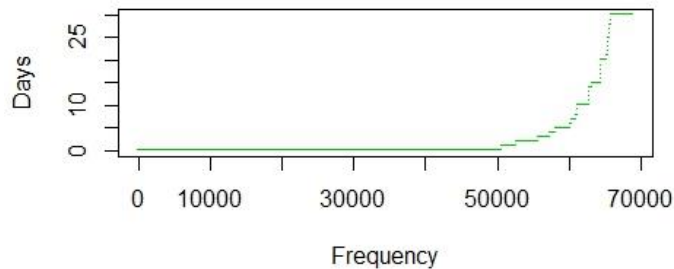
```
##IMPORT FINALDATA DATASET INTO R.
```

```
## BEFORE MAKING VARIABLES INTO CATEGORIES, CREATE HISTOGRAM TO SHOW HOW WE  
CANNOT USE LINEAR REGRESSION. THERE ARE TOO MANY ZEROS.
```

```
> hist(NEWDATA$MENTHLTH, col="green", main="", xlab="Days",  
ylab="Frequency", breaks=30)
```



```
> plot(sort(NEWDATA$MENTHLTH), xlab="Days", ylab="Frequency", col="green3",  
pch=".")
```



##Finding the frequencies for the outcome.

```
> MENTHLTH2 <-cut(NEWDATA$MENTHLTH, br=c(-1,0,13,30))
> summary(MENTHLTH2)
```

(-1,0]	(0,13]	(13,30]	NA's
50666	12164	6099	1041

## RECODING for REGRESSION MODELS:

##CATEGORICAL OUTCOME:

```
> MENTHLTH5 <-cut(NEWDATA$MENTHLTH, br=c(-1,13,30))
> is.factor(MENTHLTH5)
[1] TRUE
> table(MENTHLTH5)
MENTHLTH5
(-1,13] (13,30]
  62830    6099
> summary(MENTHLTH5)
(-1,13] (13,30] NA's
  62830    6099  1041
> levels(MENTHLTH5)<- c("0-13 unhealthy days", "14-30 unhealthy days")
> levels(MENTHLTH5)
[1] "0-13 unhealthy days" "14-30 unhealthy days"
> MENTHLTH1=factor(MENTHLTH5, c("14-30 unhealthy days","0-13 unhealthy
days"))
> levels(MENTHLTH1)
[1] "14-30 unhealthy days" "0-13 unhealthy days"
> summary(MENTHLTH1)
14-30 unhealthy days  0-13 unhealthy days      NA's
          6099              62830          1041
```

## CATEGORICAL EXPOSURE:

```
> RENTHOM12 <- factor(NEWDATA$RENTHOM1 , labels=c("1", "2",
"3"), levels=1:3)
> levels(RENTHOM12)
[1] "1" "2" "3"
> RENTHOM1Summary <- RENTHOM12
> levels(RENTHOM1Summary)<- c("own home", "non-homeownership", "non-
homeownership")
> levels(RENTHOM1Summary)
[1] "own home"      "non-homeownership"
> RENTHOM1Summary1= factor(RENTHOM1Summary, c("non-homeownership", "own
home"))
```

```
> summary(RENTHOM1Summary1)
non-homeownership      own home      NA's
      10942             58350         678
```

## ##Covariates

### ##CATEGORICAL SEX

```
> SEX2 <- factor(NEWDATA$SEX , labels=c("1", "2"), levels=1:2)
> levels(SEX2) <- c("Male","Female")
> levels(SEX2)
[1] "Male" "Female"
> summary(SEX2)
  Male Female
25877 44093
> SEXNEW <-SEX2
> summary(SEXNEW)
  Male Female
25877 44093
```

### ##CATEGORICAL HOUSEHOLD SIZE

```
> is.factor(NEWDATA$HOUSEHOLD SIZE)
[1] FALSE
> HOUSEHOLD SIZE23 <-cut(NEWDATA$HOUSEHOLD SIZE, br=c(0,2,4,6,16))
> summary(HOUSEHOLD SIZE23)
(0,2] (2,4] (4,6] (6,16]
54088 12080 3296 506
> levels(HOUSEHOLD SIZE23) <- c("1 or 2", "3 or 4", "5 or 6", "7 and more")
> summary(HOUSEHOLD SIZE23)
  1 or 2  3 or 4  5 or 6  7 and more
54088 12080 3296 506
> levels(HOUSEHOLD SIZE23) <- c("1 or 2", "3 or 4", "5 or 6", "7 and more")
> HOUSEHOLD SIZENEW <- HOUSEHOLD SIZE23
> summary(HOUSEHOLD SIZENEW)
  1 or 2  3 or 4  5 or 6  7 and more
54088 12080 3296 506
> levels(HOUSEHOLD SIZENEW)
[1] "1 or 2" "3 or 4" "5 or 6" "7 and more"
```

### ##CATEGORICAL AGE

### ##CLEANING UP AGE GROUPS (Deleting categories(or levels) within a variable)

```
> is.factor(NEWDATA$X_AGE_G)
[1] FALSE
> AGEGR <-cut(NEWDATA$X_AGE_G, br=c(0,1,2,3,4,5,6))
> summary(AGEGR)
(0,1] (1,2] (2,3] (3,4] (4,5] (5,6]
  0      0 6792 11583 18892 32703
> AGEGROUPNEW <- AGEGR[AGEGR != "1"]
```

```

> AGEGROUPNEW <- factor(AGEGROUPNEW)
> levels(AGEGROUPNEW)
[1] "(2,3]" "(3,4]" "(4,5]" "(5,6]"
> summary(AGEGROUPNEW)
(2,3] (3,4] (4,5] (5,6]
 6792 11583 18892 32703
> levels(AGEGROUPNEW) <- c("35-44", "45-54", "55-64", "65-80")
> summary(AGEGROUPNEW)
35-44 45-54 55-64 65-80
 6792 11583 18892 32703

```

#### ##CATEGORICAL HOUSEHOLD INCOME

```

> INCOME21 <-cut(NEWDATA$INCOME2, br=c(0,1,2,3,4,5,6,7,8,99))
> is.factor(INCOME21)
[1] TRUE
> summary(INCOME21)
(0,1] (1,2] (2,3] (3,4] (4,5] (5,6] (6,7] (7,8] (8,99]
 2754  3661  4742  5993  7523  9699  9925 16625  9048
> levels(INCOME21) <- c("<25000", "<25000", "<25000", "<25000", "25000 to <
50000", "25000 to < 50000", ">=50000", ">=50000", "unknown")
> summary(INCOME21)
<25000 25000 to < 50000 >=50000 unknown
 17150 17222 26550 9048
> INCOMENEW <-INCOME21
> summary(INCOMENEW)
<25000 25000 to < 50000 >=50000 unknown
 17150 17222 26550 9048
> levels(INCOMENEW)
[1] "<25000" "25000 to < 50000" ">=50000" "unknown"

```

#### ##CATEGORICAL RACE

```

> is.factor(NEWDATA$X_RACEGR3)
[1] FALSE
> RACEGR <-cut(NEWDATA$X_RACEGR3, br=c(0,1,2,3,4,5,9))
> levels(RACEGR) <- c("Non-Hispanic white", "Non-Hispanic black", "Others",
"Others", "Hispanic", "Others")
> summary(RACEGR)
Non-Hispanic white Non-Hispanic black Others Hispanic
 55182 5808 5636 3341
NA's
 3

```

#### ##CATEGORICAL EDUCATION

```

> is.factor(NEWDATA$EDUCA)
[1] FALSE
> EDUCA1 <- cut(NEWDATA$EDUCA, br=c(0,1,2,4,5,6))
> is.factor(EDUCA1)
[1] TRUE
> summary(EDUCA1)
(0,1] (1,2] (2,4] (4,5] (5,6]
  89   2201  22560  20126  24994
> levels(EDUCA1)
[1] "(0,1]" "(1,2]" "(2,4]" "(4,5]" "(5,6]"
> levels(EDUCA1) <- c("Less than High School", "Less than High School", "High
School Graduate to < 4 years of college", "High School Graduate to < 4 years
of college", ">= 4 years of college")
> summary(EDUCA1)
                Less than High School
                2290
High School Graduate to < 4 years of college
                42686
                >= 4 years of college
                24994

> EDUCANEW <- EDUCA1
> summary(EDUCANEW)
                Less than High School
                2290
High School Graduate to < 4 years of college
                42686
                >= 4 years of college
                24994

> levels(EDUCANEW)
[1] "Less than High School"
[2] "High School Graduate to < 4 years of college"
[3] ">= 4 years of college"

```

## ## CATEGORICAL EMPLOYMENT

```

> EMPLOY13 <- factor(NEWDATA$EMPLOY1, labels=c("1", "2", "3", "4", "5", "6",
"7", "8", "9"), levels=1:9)
> summary(EMPLOY13)
  1    2    3    4    5    6    7    8    9
23525 6114 1413 1026 4583    0 27963 5346    0
> EMPLOY1NEW <- EMPLOY13[EMPLOY13 != "6"]
> EMPLOY1NEW <- factor(EMPLOY1NEW)
> levels(EMPLOY1NEW)
[1] "1" "2" "3" "4" "5" "7" "8"
> levels(EMPLOY1NEW)

```

```
[1] "1" "2" "3" "4" "5" "7" "8"
> summary(EMPLOY1NEW)
  1    2    3    4    5    7    8
23525 6114 1413 1026 4583 27963 5346
> levels(EMPLOY1NEW) <- c("Employed", "Employed", "Unemployed", "Unemployed",
"Homemaker", "Retired", "Unable to work")
> summary(EMPLOY1NEW)
      Employed      Unemployed      Homemaker      Retired
      29639           2439           4583           27963
Unable to work
      5346
```

## ## CATEGORICAL MARITAL STATUS

```
> MARITAL0 <- factor(NEWDATA$MARITAL, labels=c("1", "2", "3", "4", "5", "6",
"9"), levels=1:7)
> summary(MARITAL0)
  1    2    3    4    5    6    9
40312 10053 13135 1071 5399  0    0
> MARITALNEW <- MARITAL0[MARITAL0 != "6"]
> MARITALNEW <- factor(MARITALNEW)
> levels(MARITALNEW)
[1] "1" "2" "3" "4" "5"
> levels(MARITALNEW) <- c("Married", "Divorced/widowed/Separated",
"Divorced/widowed/Separated", "Divorced/widowed/Separated", "Never married")
> summary(MARITALNEW)
      Married Divorced/widowed/Separated
      40312           24259
Never married
      5399
```

## ##LINEAR REGRESSION RESULTS

### ##UNADJUSTED LINEAR REGRESSION

```
> LINEARunadjusted <- lm(NEWDATA$MENTHLTH ~ RENTHOM1Summary1, data=NEWDATA)
> summary(LINEARunadjusted)
```

```
Call:
lm(formula = NEWDATA$MENTHLTH ~ RENTHOM1Summary1, data = NEWDATA)
```

```
Residuals:
    Min       1Q   Median       3Q      Max
-4.843 -2.502 -2.502 -1.502  27.498
```

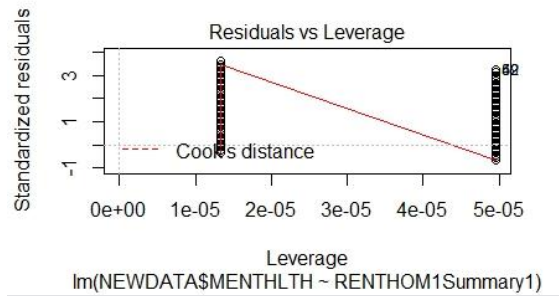
```
Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    4.84295    0.06900   70.19  <2e-16 ***
RENTHOM1Summary1own home -2.34130    0.07515  -31.16  <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Residual standard error: 7.143 on 68273 degrees of freedom
(1695 observations deleted due to missingness)
Multiple R-squared:  0.01402, Adjusted R-squared:  0.014
F-statistic: 970.6 on 1 and 68273 DF, p-value: < 2.2e-16
```

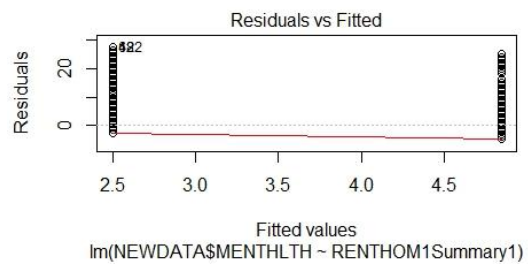


## ## Unadjusted Linear Regression

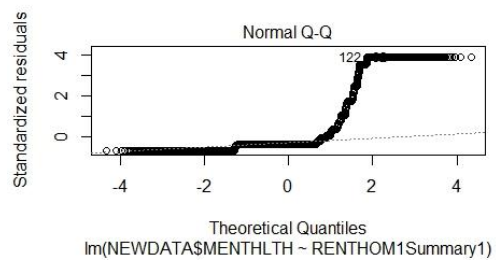
## ## Residuals vs Leverage



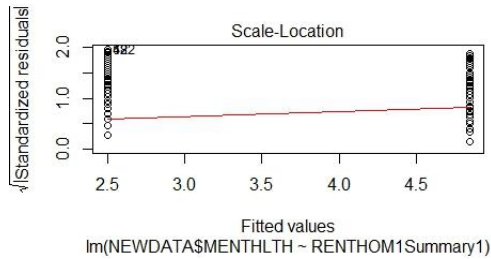
## ## Residuals vs Fitted



## ## QQ-Plot



## ## Scale-Location



## ## ADJUSTED LINEAR REGRESSION

```
> LINEAR <- lm(NEWDATA$MENTHLTH ~ RENTHOM1Summary1 + SEXNEW +  
HOUSEHOLDSIZENEW + AGEGROUPNEW + INCOMENEW + RACEGR + EDUCANEW + EMPLOY1NEW +  
MARITALNEW, data=NEWDATA)  
> summary(LINEAR)
```

Call:

```
lm(formula = NEWDATA$MENTHLTH ~ RENTHOM1Summary1 + SEXNEW + HOUSEHOLDSIZENEW  
+  
AGEGROUPNEW + INCOMENEW + RACEGR + EDUCANEW + EMPLOY1NEW +  
MARITALNEW, data = NEWDATA)
```

Residuals:

Min	1Q	Median	3Q	Max
-12.4947	-2.6391	-1.7973	-0.6391	30.2418

Coefficients:

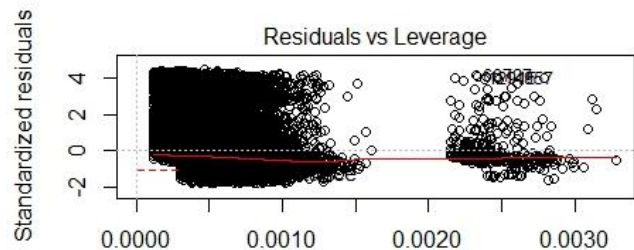
	Estimate	Std. Error
(Intercept)	3.98022	0.20987
RENTHOM1Summary1own home	-0.52694	0.07810
SEXNEWFemale	0.70342	0.05633
HOUSEHOLDSIZENEW3 or 4	0.20485	0.08089
HOUSEHOLDSIZENEW5 or 6	-0.01004	0.13875
HOUSEHOLDSIZENEW7 and more	-0.14821	0.31302
AGEGROUPNEW45-54	0.01464	0.10946
AGEGROUPNEW55-64	-0.61036	0.11170
AGEGROUPNEW65-80	-1.68083	0.12212
INCOMENEW25000 to < 50000	-0.84188	0.07967
INCOMENEW>=50000	-1.24836	0.08514
INCOMENEWUnknown	-1.08262	0.09483
RACEGRNon-Hispanic black	-0.49936	0.09831
RACEGROthers	0.15761	0.09780
RACEGRHispanic	0.04569	0.12914
EDUCANEWHigh School Graduate to < 4 years of college	-0.29761	0.15822
EDUCANEW>= 4 years of college	-0.47143	0.16505
EMPLOY1NEWUnemployed	2.79866	0.14826
EMPLOY1NEWHomemaker	0.20677	0.11558
EMPLOY1NEWRetired	0.46089	0.07616
EMPLOY1NEWUnable to work	7.09242	0.11211
MARITALNEWDivorced/widowed/Separated	0.49909	0.06482
MARITALNEWNever married	0.06960	0.10749
	t value	Pr(> t )
(Intercept)	18.965	< 2e-16 ***
RENTHOM1Summary1own home	-6.747	1.52e-11 ***
SEXNEWFemale	12.487	< 2e-16 ***
HOUSEHOLDSIZENEW3 or 4	2.533	0.01132 *

HOUSEHOLDSIZENEW5 or 6	-0.072	0.94231	
HOUSEHOLDSIZENEW7 and more	-0.473	0.63587	
AGEGROUPNEW45-54	0.134	0.89357	
AGEGROUPNEW55-64	-5.464	4.66e-08	***
AGEGROUPNEW65-80	-13.764	< 2e-16	***
INCOMENEW25000 to < 50000	-10.568	< 2e-16	***
INCOMENEW>=50000	-14.662	< 2e-16	***
INCOMENEWUnknown	-11.416	< 2e-16	***
RACEGRNon-Hispanic black	-5.080	3.79e-07	***
RACEGROthers	1.611	0.10708	
RACEGRHispanic	0.354	0.72350	
EDUCANEWHigh School Graduate to < 4 years of college	-1.881	0.05998	.
EDUCANEW>= 4 years of college	-2.856	0.00429	**
EMPLOY1NEWUnemployed	18.876	< 2e-16	***
EMPLOY1NEWHomemaker	1.789	0.07362	.
EMPLOY1NEWRetired	6.052	1.44e-09	***
EMPLOY1NEWUnable to work	63.263	< 2e-16	***
MARITALNEWDivorced/widowed/separated	7.700	1.38e-14	***
MARITALNEWNever married	0.648	0.51728	

---  
 Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

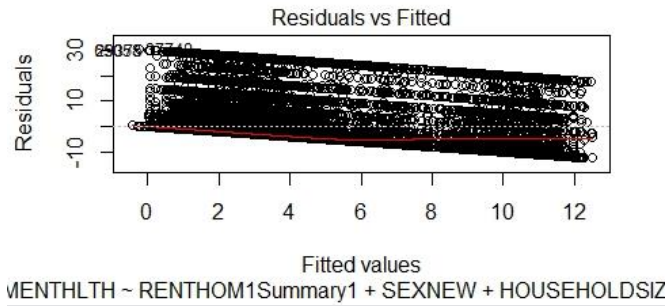
Residual standard error: 6.803 on 68250 degrees of freedom  
 (1697 observations deleted due to missingness)  
 Multiple R-squared: 0.1058, Adjusted R-squared: 0.1056  
 F-statistic: 367.2 on 22 and 68250 DF, p-value: < 2.2e-16

### ##Residuals vs. Leverage

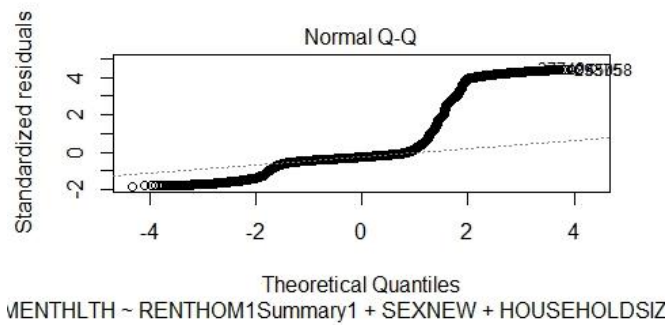


Leverage  
 MENTHLTH ~ RENTHOM1Summary1 + SEXNEW + HOUSEHOLDSIZ

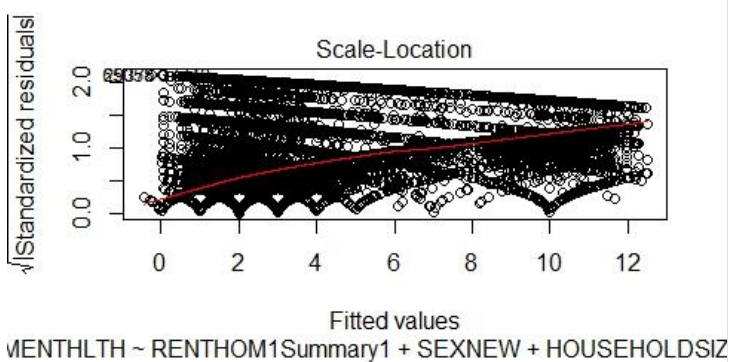
### ##Residuals vs. Fitted



## ## Normal Q-Q Plot



## ## Scale-Location



## ## LOGISTIC REGRESSION RESULTS

```
> logistic <- glm(MENTHLTH1 ~ RENTHOM1Summary1 + SEXNEW + HOUSEHOLDSIZENEW +
  AGEGROUPNEW + INCOMENEW + RACEGR + EDUCANEW + EMPLOY1NEW + MARITALNEW, data =
  NEWDATA, family = "binomial")
> summary(logistic)
```

```
Call:
glm(formula = MENTHLTH1 ~ RENTHOM1Summary1 + SEXNEW + HOUSEHOLDSIZENEW +
  AGEGROUPNEW + INCOMENEW + RACEGR + EDUCANEW + EMPLOY1NEW +
  MARITALNEW, family = "binomial", data = NEWDATA)
```

```
Deviance Residuals:
    Min       1Q   Median       3Q      Max
-2.8209   0.2878   0.3429   0.4038   1.3493
```

```
Coefficients:
```

(Intercept)	Estimate
RENTHOM1Summary1own home	2.02322
SEXNEWFemale	0.15001
HOUSEHOLDSIZENEW3 or 4	-0.30950
HOUSEHOLDSIZENEW5 or 6	-0.09150
HOUSEHOLDSIZENEW7 and more	0.03047
AGEGROUPNEW45-54	0.03426
AGEGROUPNEW55-64	0.02780
AGEGROUPNEW65-80	0.29392
INCOMENEW25000 to < 50000	0.76668
INCOMENEW>=50000	0.29811
INCOMENEWUnknown	0.58566
RACEGRNon-Hispanic black	0.37532
RACEGROthers	0.22042
RACEGRHispanic	-0.07359
EDUCANEWHigh School Graduate to < 4 years of college	-0.08808
EDUCANEW>= 4 years of college	0.14235
EMPLOY1NEWUnemployed	0.30530
EMPLOY1NEWHomemaker	-0.95791
EMPLOY1NEWRetired	-0.14086
EMPLOY1NEWUnable to work	-0.21061
MARITALNEWDivorced/widowed/Separated	-1.76203
MARITALNEWNever married	-0.25570
	Std. Error
(Intercept)	0.09895
RENTHOM1Summary1own home	0.03631
SEXNEWFemale	0.03179
HOUSEHOLDSIZENEW3 or 4	0.04174
HOUSEHOLDSIZENEW5 or 6	0.07284
HOUSEHOLDSIZENEW7 and more	0.15264
AGEGROUPNEW45-54	0.05426
AGEGROUPNEW55-64	0.05600
AGEGROUPNEW65-80	0.06334
INCOMENEW25000 to < 50000	0.04040
INCOMENEW>=50000	0.04614
INCOMENEWUnknown	0.04886
RACEGRNon-Hispanic black	0.04829
RACEGROthers	0.05069
RACEGRHispanic	0.06177
EDUCANEWHigh School Graduate to < 4 years of college	0.06883
EDUCANEW>= 4 years of college	0.07479
EMPLOY1NEWUnemployed	0.06037
EMPLOY1NEWHomemaker	0.06463
EMPLOY1NEWRetired	0.04571
EMPLOY1NEWUnable to work	0.04410
MARITALNEWDivorced/widowed/Separated	0.03509

MARITALNEWNever married	0.05421	
	z value	
(Intercept)	20.446	
RENTHOM1Summary1own home	4.131	
SEXNEWFemale	-9.736	
HOUSEHOLDSIZENEW3 or 4	-2.192	
HOUSEHOLDSIZENEW5 or 6	0.418	
HOUSEHOLDSIZENEW7 and more	0.224	
AGEGROUPNEW45-54	0.512	
AGEGROUPNEW55-64	5.248	
AGEGROUPNEW65-80	12.105	
INCOMENEW25000 to < 50000	7.380	
INCOMENEW>=50000	12.694	
INCOMENEWUnknown	7.682	
RACEGRNon-Hispanic black	4.565	
RACEGROthers	-1.452	
RACEGRHispanic	-1.426	
EDUCANEWHigh School Graduate to < 4 years of college	2.068	
EDUCANEW>= 4 years of college	4.082	
EMPLOY1NEWUnemployed	-15.867	
EMPLOY1NEWHomemaker	-2.180	
EMPLOY1NEWRetired	-4.607	
EMPLOY1NEWUnable to work	-39.955	
MARITALNEWDivorced/widowed/Separated	-7.288	
MARITALNEWNever married	-1.355	
	Pr(> z )	
(Intercept)	< 2e-16	***
RENTHOM1Summary1own home	3.60e-05	***
SEXNEWFemale	< 2e-16	***
HOUSEHOLDSIZENEW3 or 4	0.0284	*
HOUSEHOLDSIZENEW5 or 6	0.6757	
HOUSEHOLDSIZENEW7 and more	0.8224	
AGEGROUPNEW45-54	0.6084	
AGEGROUPNEW55-64	1.54e-07	***
AGEGROUPNEW65-80	< 2e-16	***
INCOMENEW25000 to < 50000	1.59e-13	***
INCOMENEW>=50000	< 2e-16	***
INCOMENEWUnknown	1.57e-14	***
RACEGRNon-Hispanic black	5.00e-06	***
RACEGROthers	0.1466	
RACEGRHispanic	0.1539	
EDUCANEWHigh School Graduate to < 4 years of college	0.0386	*
EDUCANEW>= 4 years of college	4.46e-05	***
EMPLOY1NEWUnemployed	< 2e-16	***
EMPLOY1NEWHomemaker	0.0293	*
EMPLOY1NEWRetired	4.08e-06	***
EMPLOY1NEWUnable to work	< 2e-16	***
MARITALNEWDivorced/widowed/Separated	3.16e-13	***
MARITALNEWNever married	0.1753	

---  
 signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 40843 on 68272 degrees of freedom  
 Residual deviance: 36478 on 68250 degrees of freedom  
 (1697 observations deleted due to missingness)  
 AIC: 36524

Number of Fisher Scoring iterations: 5

## Verifying over-dispersion

```

> qchisq(0.95,68250)
[1] 68863.86
> with(logistic, null deviance - deviance)
Error: unexpected symbol in "with(logistic, null deviance"
> with(logistic, null.deviance - deviance)
[1] 4365.181
> with(logistic, df.null - df.residual)
[1] 17
> with(logistic, pchisq(null.deviance - deviance, df.null - df.residual,
lower.tail = FALSE))
[1] 0
> logLik(logistic)
'log Lik.' -18238.98 (df=23)

```

```
## Find odds ratios, 95% CIs, and p-values. Then, make it into a table.
```

```

> cbind(OR=exp(coef(logistic)), exp(confint(logistic)),
p=summary(logistic)$coef[, 4])
Waiting for profiling to be done...

```

	OR
(Intercept)	7.5626206
RENTHOM1Summary1own home	1.1618511
SEXNEWFemale	0.7338163
HOUSEHOLDSIZENEW3 or 4	0.9125578
HOUSEHOLDSIZENEW5 or 6	1.0309361
HOUSEHOLDSIZENEW7 and more	1.0348500
AGEGROUPNEW45-54	1.0281898
AGEGROUPNEW55-64	1.3416767
AGEGROUPNEW65-80	2.1526125
INCOMENEW25000 to < 50000	1.3473129
INCOMENEW>=50000	1.7961810
INCOMENEWUnknown	1.4554512
RACEGRNon-Hispanic black	1.2466004
RACEGROthers	0.9290542
RACEGRHispanic	0.9156860
EDUCANEWHigh School Graduate to < 4 years of college	1.1529795
EDUCANEW>= 4 years of college	1.3570288
EMPLOY1NEWUnemployed	0.3836957
EMPLOY1NEWHomemaker	0.8686134
EMPLOY1NEWRetired	0.8100910
EMPLOY1NEWUnable to work	0.1716955
MARITALNEWDivorced/widowed/Separated	0.7743718
MARITALNEWNever married	0.9291536
	2.5 %
(Intercept)	6.2345150
RENTHOM1Summary1own home	1.0818093
SEXNEWFemale	0.6893760
HOUSEHOLDSIZENEW3 or 4	0.8410894
HOUSEHOLDSIZENEW5 or 6	0.8950361
HOUSEHOLDSIZENEW7 and more	0.7737515
AGEGROUPNEW45-54	0.9241766
AGEGROUPNEW55-64	1.2018292
AGEGROUPNEW65-80	1.9009454
INCOMENEW25000 to < 50000	1.2448940
INCOMENEW>=50000	1.6410103
INCOMENEWUnknown	1.3232089
RACEGRNon-Hispanic black	1.1347510

RACEGROthers	0.8418971
RACEGRHispanic	0.8121636
EDUCANEWHigh School Graduate to < 4 years of college	1.0062789
EDUCANEW>= 4 years of college	1.1708477
EMPLOY1NEWUnemployed	0.3411362
EMPLOY1NEWHomemaker	0.7662269
EMPLOY1NEWRetired	0.7406277
EMPLOY1NEWUnable to work	0.1574674
MARITALNEWDivorced/widowed/Separated	0.7229128
MARITALNEWNever married	0.8359695
	97.5 %
(Intercept)	9.1892303
RENTHOM1Summary1own home	1.2472962
SEXNEWFemale	0.7808706
HOUSEHOLDSIZENEW3 or 4	0.9906377
HOUSEHOLDSIZENEW5 or 6	1.1909158
HOUSEHOLDSIZENEW7 and more	1.4089047
AGEGROUPNEW45-54	1.1432554
AGEGROUPNEW55-64	1.4969068
AGEGROUPNEW65-80	2.4367329
INCOMENEW25000 to < 50000	1.4585053
INCOMENEW>=50000	1.9663508
INCOMENEWUnknown	1.6025599
RACEGRNon-Hispanic black	1.3712476
RACEGROthers	1.0269851
RACEGRHispanic	1.0347040
EDUCANEWHigh School Graduate to < 4 years of college	1.3180244
EDUCANEW>= 4 years of college	1.5698073
EMPLOY1NEWUnemployed	0.4322406
EMPLOY1NEWHomemaker	0.9872087
EMPLOY1NEWRetired	0.8859837
EMPLOY1NEWUnable to work	0.1871853
MARITALNEWDivorced/widowed/Separated	0.8295129
MARITALNEWNever married	1.0339434

	p
(Intercept)	6.506957e-93
RENTHOM1Summary1own home	3.604993e-05
SEXNEWFemale	2.115975e-22
HOUSEHOLDSIZENEW3 or 4	2.838053e-02
HOUSEHOLDSIZENEW5 or 6	6.757459e-01
HOUSEHOLDSIZENEW7 and more	8.224212e-01
AGEGROUPNEW45-54	6.084197e-01
AGEGROUPNEW55-64	1.535243e-07
AGEGROUPNEW65-80	9.988067e-34
INCOMENEW25000 to < 50000	1.585066e-13
INCOMENEW>=50000	6.371018e-37
INCOMENEWUnknown	1.565470e-14
RACEGRNon-Hispanic black	5.001362e-06
RACEGROthers	1.465504e-01
RACEGRHispanic	1.538543e-01
EDUCANEWHigh School Graduate to < 4 years of college	3.862902e-02
EDUCANEW>= 4 years of college	4.461760e-05
EMPLOY1NEWUnemployed	1.075737e-56
EMPLOY1NEWHomemaker	2.928803e-02
EMPLOY1NEWRetired	4.078797e-06
EMPLOY1NEWUnable to work	0.000000e+00
MARITALNEWDivorced/widowed/Separated	3.155503e-13
MARITALNEWNever married	1.752946e-01

```

> logistictable <- cbind(OR=exp(coef(logistic)), exp(confint(logistic)),
p=summary(logistic)$coef[, 4])
waiting for profiling to be done...
> write.csv(logistictable, "C:/Users/James/Desktop/logittable.csv")

```



```
## Verifying fit of logistic regression (with predictors vs. without
predictors)
> gchisq(0.95,68250)
[1] 68863.86
> with(logistic, null deviance - deviance)
Error: unexpected symbol in "with(logistic, null deviance"
> with(logistic, null.deviance - deviance)
[1] 4180.224
> with(logistic, df.null - df.residual)
[1] 17
> with(logistic, pchisq(null.deviance - deviance, df.null - df.residual,
lower.tail = FALSE))
[1] 0
```

## ## Log-Likelihood

```
> logLik(logistic)
'log Lik.' -18238.98 (df=23)
```

## ## POISSON REGRESSION:

```
> POISSON <- glm(NEWDATA$MENTHLTH ~ RENTHOM1Summary1 + SEXNEW +
HOUSEHOLDSIZENEW + AGEGROUPNEW + INCOMENEW + RACEGR + EDUCANEW + EMPLOY1NEW +
MARITALNEW, family=poisson(link="log"))
> summary(POISSON)
```

```
Call:
glm(formula = NEWDATA$MENTHLTH ~ RENTHOM1Summary1 + SEXNEW +
HOUSEHOLDSIZENEW + AGEGROUPNEW + INCOMENEW + RACEGR + EDUCANEW +
EMPLOY1NEW + MARITALNEW, family = poisson(link = "log"))
```

```
Deviance Residuals:
    Min       1Q   Median       3Q      Max
-6.433  -2.191  -1.893  -1.369   12.330
```

## Coefficients:

	Estimate
(Intercept)	1.225527
RENTHOM1Summary1own home	-0.109092
SEXNEWFemale	0.281583
HOUSEHOLDSIZENEW3 or 4	0.094658
HOUSEHOLDSIZENEW5 or 6	0.042891
HOUSEHOLDSIZENEW7 and more	-0.001799
AGEGROUPNEW45-54	-0.005941
AGEGROUPNEW55-64	-0.170818
AGEGROUPNEW65-80	-0.558701
INCOMENEW25000 to < 50000	-0.222877
INCOMENEW>=50000	-0.424066
INCOMENEWUnknown	-0.300508
RACEGRNon-Hispanic black	-0.137655
RACEGROthers	0.051453
RACEGRHispanic	0.018781
EDUCANEWHigh School Graduate to < 4 years of college	-0.106373
EDUCANEW>= 4 years of college	-0.193155
EMPLOY1NEWUnemployed	0.709654
EMPLOY1NEWHomemaker	0.075313
EMPLOY1NEWRetired	0.123635
EMPLOY1NEWUnable to work	1.252717

MARITALNEWDivorced/widowed/Separated	0.181248
MARITALNEWNever married	0.053332
	Std. Error
(Intercept)	0.015769
RENTHOM1Summary1own home	0.005775
SEXNEWFemale	0.005119
HOUSEHOLDSIZENEW3 or 4	0.006569
HOUSEHOLDSIZENEW5 or 6	0.011451
HOUSEHOLDSIZENEW7 and more	0.024516
AGEGROUPNEW45-54	0.008573
AGEGROUPNEW55-64	0.008808
AGEGROUPNEW65-80	0.010060
INCOMENEW25000 to < 50000	0.006607
INCOMENEW>=50000	0.007486
INCOMENEWUnknown	0.007992
RACEGRNon-Hispanic black	0.007553
RACEGROthers	0.008053
RACEGRHispanic	0.009982
EDUCANEWHigh School Graduate to < 4 years of college	0.010848
EDUCANEW>= 4 years of college	0.011823
EMPLOY1NEWUnemployed	0.009838
EMPLOY1NEWHomemaker	0.010660
EMPLOY1NEWRetired	0.007427
EMPLOY1NEWUnable to work	0.006944
MARITALNEWDivorced/widowed/Separated	0.005652
MARITALNEWNever married	0.008621
	z value
(Intercept)	77.715
RENTHOM1Summary1own home	-18.891
SEXNEWFemale	55.002
HOUSEHOLDSIZENEW3 or 4	14.409
HOUSEHOLDSIZENEW5 or 6	3.746
HOUSEHOLDSIZENEW7 and more	-0.073
AGEGROUPNEW45-54	-0.693
AGEGROUPNEW55-64	-19.395
AGEGROUPNEW65-80	-55.535
INCOMENEW25000 to < 50000	-33.736
INCOMENEW>=50000	-56.646
INCOMENEWUnknown	-37.603
RACEGRNon-Hispanic black	-18.225
RACEGROthers	6.389
RACEGRHispanic	1.882
EDUCANEWHigh School Graduate to < 4 years of college	-9.806
EDUCANEW>= 4 years of college	-16.338
EMPLOY1NEWUnemployed	72.133
EMPLOY1NEWHomemaker	7.065
EMPLOY1NEWRetired	16.647
EMPLOY1NEWUnable to work	180.398
MARITALNEWDivorced/widowed/Separated	32.069
MARITALNEWNever married	6.186
	Pr(> z )
(Intercept)	< 2e-16 ***
RENTHOM1Summary1own home	< 2e-16 ***
SEXNEWFemale	< 2e-16 ***
HOUSEHOLDSIZENEW3 or 4	< 2e-16 ***
HOUSEHOLDSIZENEW5 or 6	0.00018 ***
HOUSEHOLDSIZENEW7 and more	0.94149
AGEGROUPNEW45-54	0.48832
AGEGROUPNEW55-64	< 2e-16 ***
AGEGROUPNEW65-80	< 2e-16 ***
INCOMENEW25000 to < 50000	< 2e-16 ***
INCOMENEW>=50000	< 2e-16 ***
INCOMENEWUnknown	< 2e-16 ***
RACEGRNon-Hispanic black	< 2e-16 ***

RACEGROthers	1.67e-10	***
RACEGRHispanic	0.05990	.
EDUCANEWHigh School Graduate to < 4 years of college	< 2e-16	***
EDUCANEW>= 4 years of college	< 2e-16	***
EMPLOY1NEWUnemployed	< 2e-16	***
EMPLOY1NEWHomemaker	1.61e-12	***
EMPLOY1NEWRetired	< 2e-16	***
EMPLOY1NEWUnable to work	< 2e-16	***
MARITALNEWDivorced/widowed/separated	< 2e-16	***
MARITALNEWNever married	6.16e-10	***

---  
 Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for poisson family taken to be 1)

Null deviance: 692950 on 68272 degrees of freedom  
 Residual deviance: 600314 on 68250 degrees of freedom  
 (1697 observations deleted due to missingness)  
 AIC: 667567

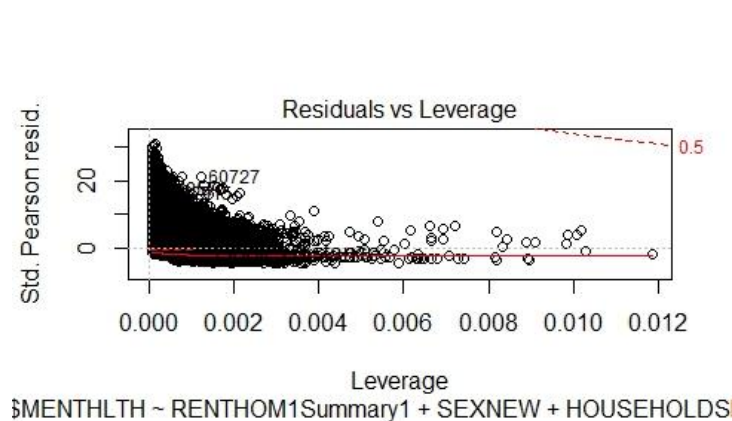
Number of Fisher Scoring iterations: 7

## Log-likelihood

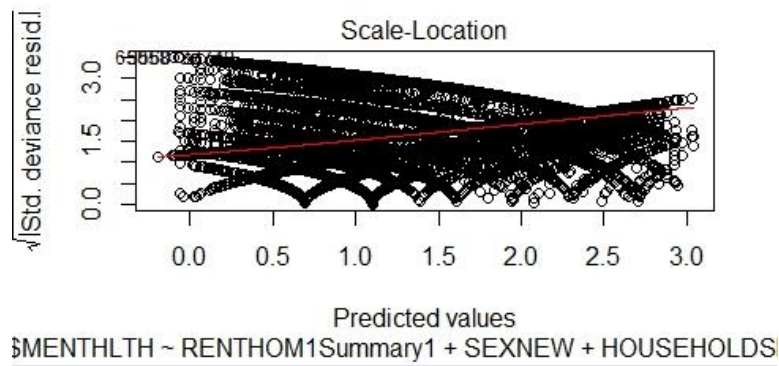
```
> logLik(POISSON)
'log Lik.' -333760.5 (df=23)
```

## RESIDUALS VS LEVERAGE: POISSON

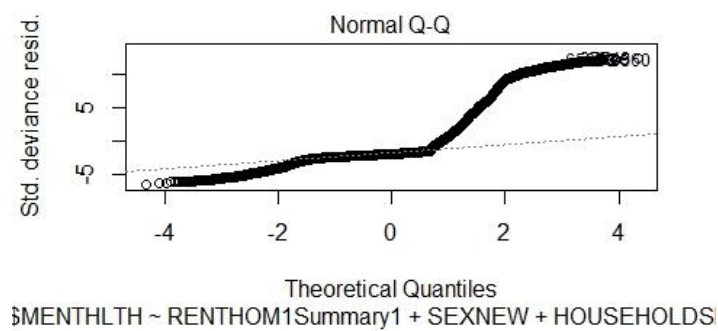
```
> plot(POISSON)
```



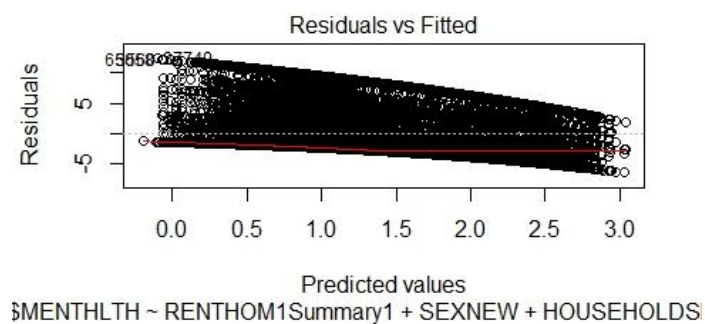
## SCALE LOCATION: POISSON



### ## QQ Plot: Poisson Regression



### ## RESIDUALS VS. FITTED: Poisson Regression



## ## QUASI POISSON

```
> QPOISSON <- glm(NEWDATA$MENTHLTH ~ RENTHOM1Summary1 + SEXNEW +  
HOUSEHOLDSIZENew + AGEGRouPNEW + INCOMENew + RACEGR + EDUCANew + EMPLOY1NEW +  
MARITALNEW, family=quasipoisson(link="log"))  
> summary(QPOISSON)
```

Call:

```
glm(formula = NEWDATA$MENTHLTH ~ RENTHOM1Summary1 + SEXNEW +  
HOUSEHOLDSIZENew + AGEGRouPNEW + INCOMENew + RACEGR + EDUCANew +  
EMPLOY1NEW + MARITALNEW, family = quasipoisson(link = "log"))
```

Deviance Residuals:

Min	1Q	Median	3Q	Max
-6.433	-2.191	-1.893	-1.369	12.330

Coefficients:

	Estimate
(Intercept)	1.225527
RENTHOM1Summary1own home	-0.109092
SEXNEWFemale	0.281583
HOUSEHOLDSIZENew3 or 4	0.094658
HOUSEHOLDSIZENew5 or 6	0.042891
HOUSEHOLDSIZENew7 and more	-0.001799
AGEGRouPNEW45-54	-0.005941
AGEGRouPNEW55-64	-0.170818
AGEGRouPNEW65-80	-0.558701
INCOMENew25000 to < 50000	-0.222877
INCOMENew>=50000	-0.424066
INCOMENewUnknown	-0.300508
RACEGRNon-Hispanic black	-0.137655
RACEGROthers	0.051453
RACEGRHispanic	0.018781
EDUCANewHigh School Graduate to < 4 years of college	-0.106373
EDUCANew>= 4 years of college	-0.193155
EMPLOY1NEWUnemployed	0.709654
EMPLOY1NEWHomemaker	0.075313
EMPLOY1NEWRetired	0.123635
EMPLOY1NEWUnable to work	1.252717
MARITALNEWDivorced/widowed/Separated	0.181248
MARITALNEWNever married	0.053332
	Std. Error
(Intercept)	0.064534
RENTHOM1Summary1own home	0.023632
SEXNEWFemale	0.020951
HOUSEHOLDSIZENew3 or 4	0.026884
HOUSEHOLDSIZENew5 or 6	0.046861
HOUSEHOLDSIZENew7 and more	0.100328
AGEGRouPNEW45-54	0.035084
AGEGRouPNEW55-64	0.036044
AGEGRouPNEW65-80	0.041171
INCOMENew25000 to < 50000	0.027036
INCOMENew>=50000	0.030637
INCOMENewUnknown	0.032704
RACEGRNon-Hispanic black	0.030910
RACEGROthers	0.032956
RACEGRHispanic	0.040849
EDUCANewHigh School Graduate to < 4 years of college	0.044394

```

EDUCANEW>= 4 years of college      0.048383
EMPLOY1NEWUnemployed              0.040261
EMPLOY1NEWHomemaker              0.043626
EMPLOY1NEWRetired                 0.030394
EMPLOY1NEWUnable to work          0.028418
MARITALNEWDivorced/widowed/Separated 0.023129
MARITALNEWNever married           0.035280
                                     t value
(Intercept)                      18.990
RENTHOM1Summary1own home         -4.616
SEXNEWFemale                     13.440
HOUSEHOLDSIZENEW3 or 4           3.521
HOUSEHOLDSIZENEW5 or 6           0.915
HOUSEHOLDSIZENEW7 and more       -0.018
AGEGROUPNEW45-54                 -0.169
AGEGROUPNEW55-64                 -4.739
AGEGROUPNEW65-80                -13.570
INCOMENEW25000 to < 50000        -8.244
INCOMENEW>=50000                 -13.842
INCOMENEWUnknown                 -9.189
RACEGRNon-Hispanic black         -4.453
RACEGROthers                     1.561
RACEGRHispanic                   0.460
EDUCANEWHigh School Graduate to < 4 years of college -2.396
EDUCANEW>= 4 years of college    -3.992
EMPLOY1NEWUnemployed             17.626
EMPLOY1NEWHomemaker              1.726
EMPLOY1NEWRetired                4.068
EMPLOY1NEWUnable to work         44.082
MARITALNEWDivorced/widowed/Separated 7.836
MARITALNEWNever married          1.512
                                     Pr(>|t|)
(Intercept)                      < 2e-16 ***
RENTHOM1Summary1own home         3.91e-06 ***
SEXNEWFemale                     < 2e-16 ***
HOUSEHOLDSIZENEW3 or 4           0.00043 ***
HOUSEHOLDSIZENEW5 or 6           0.36004
HOUSEHOLDSIZENEW7 and more       0.98569
AGEGROUPNEW45-54                 0.86553
AGEGROUPNEW55-64                 2.15e-06 ***
AGEGROUPNEW65-80                 < 2e-16 ***
INCOMENEW25000 to < 50000        < 2e-16 ***
INCOMENEW>=50000                 < 2e-16 ***
INCOMENEWUnknown                 < 2e-16 ***
RACEGRNon-Hispanic black         8.47e-06 ***
RACEGROthers                     0.11846
RACEGRHispanic                   0.64569
EDUCANEWHigh School Graduate to < 4 years of college 0.01657 *
EDUCANEW>= 4 years of college    6.55e-05 ***
EMPLOY1NEWUnemployed             < 2e-16 ***
EMPLOY1NEWHomemaker              0.08429 .
EMPLOY1NEWRetired                4.75e-05 ***
EMPLOY1NEWUnable to work         < 2e-16 ***
MARITALNEWDivorced/widowed/Separated 4.71e-15 ***
MARITALNEWNever married          0.13062
---
signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for quasipoisson family taken to be 16.74745)

Null deviance: 692950  on 68272  degrees of freedom
Residual deviance: 600314  on 68250  degrees of freedom
(1697 observations deleted due to missingness)
AIC: NA

```

Number of Fisher Scoring iterations: 7

## ## Negative binomial regression

```
> library(MASS)
> NEGBIN <- glm.nb(NEWDATA$MENTHLTH ~ RENTHOM1Summary1 + SEXNEW +
HOUSEHOLDSIZENEW + AGEGROUPNEW + INCOMENEW + RACEGR + EDUCANEW + EMPLOY1NEW +
MARITALNEW)
> summary(NEGBIN)
```

```
Call:
glm.nb(formula = NEWDATA$MENTHLTH ~ RENTHOM1Summary1 + SEXNEW +
HOUSEHOLDSIZENEW + AGEGROUPNEW + INCOMENEW + RACEGR + EDUCANEW +
EMPLOY1NEW + MARITALNEW, init.theta = 0.1008499705, link = log)
```

```
Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.0501  -0.8016  -0.7620  -0.2568   2.3387
```

### Coefficients:

	Estimate
(Intercept)	1.22937
RENTHOM1Summary1own home	-0.13509
SEXNEWFemale	0.31974
HOUSEHOLDSIZENEW3 or 4	0.14724
HOUSEHOLDSIZENEW5 or 6	0.10050
HOUSEHOLDSIZENEW7 and more	0.06929
AGEGROUPNEW45-54	-0.01664
AGEGROUPNEW55-64	-0.14374
AGEGROUPNEW65-80	-0.48265
INCOMENEW25000 to < 50000	-0.25275
INCOMENEW>=50000	-0.43374
INCOMENEWUnknown	-0.35167
RACEGRNon-Hispanic black	-0.03777
RACEGROthers	0.09323
RACEGRHispanic	0.10627
EDUCANEWHigh School Graduate to < 4 years of college	-0.17815
EDUCANEW>= 4 years of college	-0.28065
EMPLOY1NEWUnemployed	0.70483
EMPLOY1NEWHomemaker	0.04898
EMPLOY1NEWRetired	0.10268
EMPLOY1NEWUnable to work	1.26914
MARITALNEWDivorced/widowed/Separated	0.21986
MARITALNEWNever married	0.10793
	Std. Error
(Intercept)	0.09882
RENTHOM1Summary1own home	0.03677
SEXNEWFemale	0.02674
HOUSEHOLDSIZENEW3 or 4	0.03823
HOUSEHOLDSIZENEW5 or 6	0.06554
HOUSEHOLDSIZENEW7 and more	0.14759
AGEGROUPNEW45-54	0.05163

AGEGROUPNEW55-64	0.05273
AGEGROUPNEW65-80	0.05775
INCOMENEW25000 to < 50000	0.03759
INCOMENEW>=50000	0.04026
INCOMENEWUnknown	0.04483
RACEGRNon-Hispanic black	0.04638
RACEGROthers	0.04624
RACEGRHispanic	0.06086
EDUCANEWHigh School Graduate to < 4 years of college	0.07433
EDUCANEW>= 4 years of college	0.07762
EMPLOY1NEWUnemployed	0.06941
EMPLOY1NEWHomemaker	0.05469
EMPLOY1NEWRetired	0.03620
EMPLOY1NEWUnable to work	0.05240
MARITALNEWDivorced/widowed/Separated	0.03069
MARITALNEWNever married	0.05079
	z value
(Intercept)	12.440
RENTHOM1Summary1own home	-3.674
SEXNEWFemale	11.956
HOUSEHOLDSIZENEW3 or 4	3.851
HOUSEHOLDSIZENEW5 or 6	1.533
HOUSEHOLDSIZENEW7 and more	0.469
AGEGROUPNEW45-54	-0.322
AGEGROUPNEW55-64	-2.726
AGEGROUPNEW65-80	-8.358
INCOMENEW25000 to < 50000	-6.724
INCOMENEW>=50000	-10.773
INCOMENEWUnknown	-7.845
RACEGRNon-Hispanic black	-0.814
RACEGROthers	2.016
RACEGRHispanic	1.746
EDUCANEWHigh School Graduate to < 4 years of college	-2.397
EDUCANEW>= 4 years of college	-3.615
EMPLOY1NEWUnemployed	10.155
EMPLOY1NEWHomemaker	0.896
EMPLOY1NEWRetired	2.836
EMPLOY1NEWUnable to work	24.220
MARITALNEWDivorced/widowed/Separated	7.163
MARITALNEWNever married	2.125
	Pr(> z )
(Intercept)	< 2e-16 ***
RENTHOM1Summary1own home	0.000239 ***
SEXNEWFemale	< 2e-16 ***
HOUSEHOLDSIZENEW3 or 4	0.000117 ***
HOUSEHOLDSIZENEW5 or 6	0.125204
HOUSEHOLDSIZENEW7 and more	0.638723
AGEGROUPNEW45-54	0.747301
AGEGROUPNEW55-64	0.006409 **
AGEGROUPNEW65-80	< 2e-16 ***
INCOMENEW25000 to < 50000	1.77e-11 ***
INCOMENEW>=50000	< 2e-16 ***
INCOMENEWUnknown	4.34e-15 ***
RACEGRNon-Hispanic black	0.415484
RACEGROthers	0.043789 *
RACEGRHispanic	0.080773 .
EDUCANEWHigh School Graduate to < 4 years of college	0.016545 *
EDUCANEW>= 4 years of college	0.000300 ***
EMPLOY1NEWUnemployed	< 2e-16 ***
EMPLOY1NEWHomemaker	0.370489
EMPLOY1NEWRetired	0.004567 **
EMPLOY1NEWUnable to work	< 2e-16 ***
MARITALNEWDivorced/widowed/Separated	7.91e-13 ***
MARITALNEWNever married	0.033573 *



```

---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for Negative Binomial(0.1008) family taken to be 1)

Null deviance: 41357  on 68272  degrees of freedom
Residual deviance: 38712  on 68250  degrees of freedom
(1697 observations deleted due to missingness)
AIC: 198784

Number of Fisher Scoring iterations: 1

      Theta:  0.100850
    Std. Err.: 0.000930

2 x log-likelihood:  -198735.522000

```

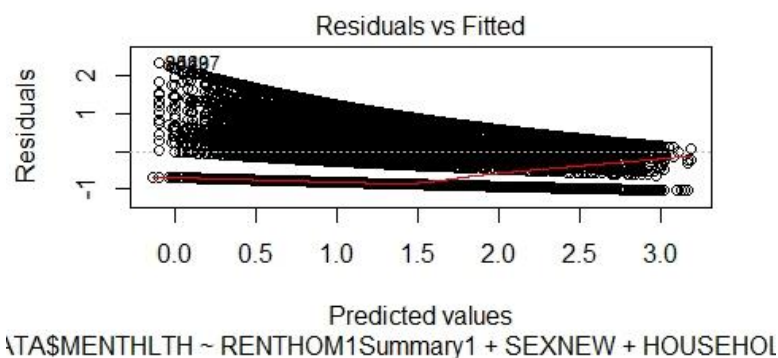
```
## Log-likelihood
```

```

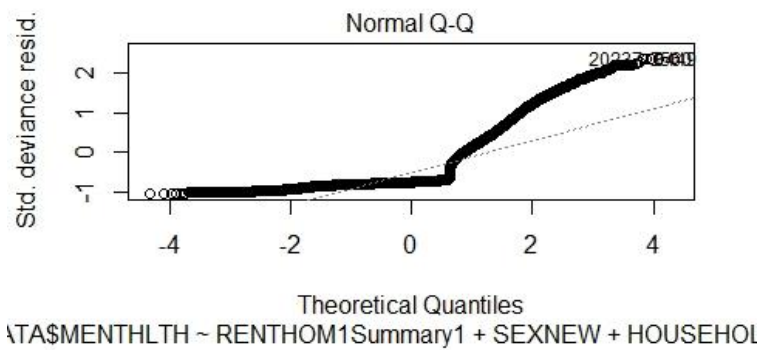
> logLik(NEGBIN)
'log Lik.' -99367.76 (df=24)

```

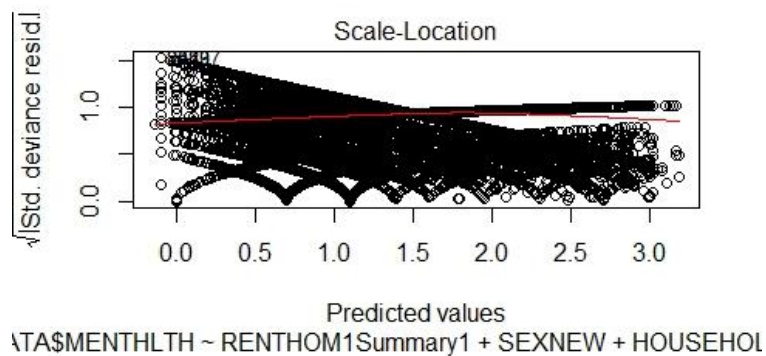
```
## RESIDUALS VS FITTED: Negative binomial
```



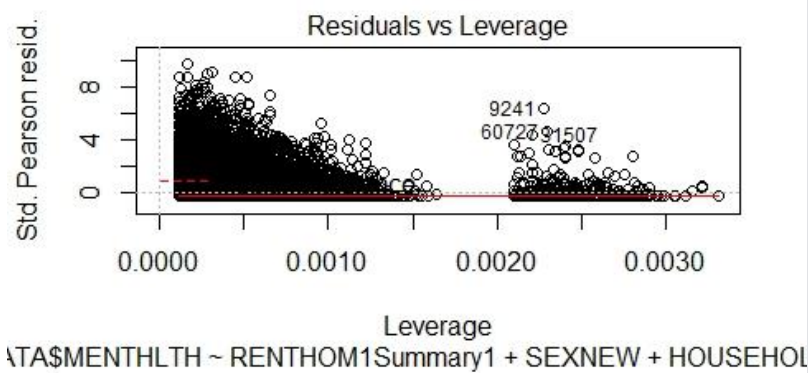
```
## QQ PLOT: Negative binomial
```



## SCALE- LOCATION: Negative binomial



## RESIDUALS VS Leverage: Negative binomial



## Install packages for zero-inflated models

```
> install.packages("ggplot2")
```

```
> install.packages("pscl")
> install.packages("boot")
> library(pscl)
```

## ## Zero-inflated Poisson Model

```
> ZIPoisson <- zeroinfl(NEWDATA$MENTHLTH ~ RENTHOM1Summary1 + SEXNEW +
HOUSEHOLDSIZENEW + AGEGROUPNEW + INCOMENEW + RACEGR + EDUCANEW + EMPLOY1NEW +
MARITALNEW, dist="poisson")
> summary(ZIPoisson)
```

Call:

```
zeroinfl(formula = NEWDATA$MENTHLTH ~ RENTHOM1Summary1 + SEXNEW +
HOUSEHOLDSIZENEW + AGEGROUPNEW + INCOMENEW + RACEGR + EDUCANEW +
EMPLOY1NEW + MARITALNEW, dist = "poisson")
```

Pearson residuals:

Min	1Q	Median	3Q	Max
-1.6478	-0.5512	-0.4546	-0.3311	12.7064

Count model coefficients (poisson with log link):

	Estimate
(Intercept)	2.486550
RENTHOM1Summary1own home	-0.045803
SEXNEWFemale	-0.029553
HOUSEHOLDSIZENEW3 or 4	-0.030893
HOUSEHOLDSIZENEW5 or 6	-0.089621
HOUSEHOLDSIZENEW7 and more	-0.094529
AGEGROUPNEW45-54	-0.003706
AGEGROUPNEW55-64	-0.039591
AGEGROUPNEW65-80	-0.083705
INCOMENEW25000 to < 50000	-0.112349
INCOMENEW>=50000	-0.217196
INCOMENEWUnknown	-0.015222
RACEGRNon-Hispanic black	-0.045536
RACEGROthers	0.076977
RACEGRHispanic	0.068825
EDUCANEWHigh School Graduate to < 4 years of college	-0.082557
EDUCANEW>= 4 years of college	-0.220436
EMPLOY1NEWUnemployed	0.297438
EMPLOY1NEWHomemaker	0.076263
EMPLOY1NEWRetired	0.128586
EMPLOY1NEWUnable to work	0.543249
MARITALNEWDivorced/widowed/Separated	0.095432
MARITALNEWNever married	-0.001573
	Std. Error
(Intercept)	0.015614
RENTHOM1Summary1own home	0.005733
SEXNEWFemale	0.005088
HOUSEHOLDSIZENEW3 or 4	0.006505
HOUSEHOLDSIZENEW5 or 6	0.011422
HOUSEHOLDSIZENEW7 and more	0.024378
AGEGROUPNEW45-54	0.008575
AGEGROUPNEW55-64	0.008766
AGEGROUPNEW65-80	0.010000
INCOMENEW25000 to < 50000	0.006558
INCOMENEW>=50000	0.007381
INCOMENEWUnknown	0.007925
RACEGRNon-Hispanic black	0.007519
RACEGROthers	0.008053
RACEGRHispanic	0.009890
EDUCANEWHigh School Graduate to < 4 years of college	0.010831
EDUCANEW>= 4 years of college	0.011824
EMPLOY1NEWUnemployed	0.009841
EMPLOY1NEWHomemaker	0.010653

EMPLOY1NEWRetired	0.007451
EMPLOY1NEWUnable to work	0.006836
MARITALNEWDivorced/widowed/Separated	0.005544
MARITALNEWNever married	0.008609
	z value
(Intercept)	159.248
RENTHOM1Summary1own home	-7.989
SEXNEWFemale	-5.809
HOUSEHOLDSIZENEW3 or 4	-4.749
HOUSEHOLDSIZENEW5 or 6	-7.846
HOUSEHOLDSIZENEW7 and more	-3.878
AGEGROUPNEW45-54	-0.432
AGEGROUPNEW55-64	-4.516
AGEGROUPNEW65-80	-8.370
INCOMENEW25000 to < 50000	-17.133
INCOMENEW>=50000	-29.427
INCOMENEWUnknown	-1.921
RACEGRNon-Hispanic black	-6.056
RACEGROthers	9.559
RACEGRHispanic	6.959
EDUCANEWHigh School Graduate to < 4 years of college	-7.622
EDUCANEW>= 4 years of college	-18.643
EMPLOY1NEWUnemployed	30.225
EMPLOY1NEWHomemaker	7.159
EMPLOY1NEWRetired	17.257
EMPLOY1NEWUnable to work	79.472
MARITALNEWDivorced/widowed/Separated	17.212
MARITALNEWNever married	-0.183
	Pr(> z )
(Intercept)	< 2e-16 ***
RENTHOM1Summary1own home	1.36e-15 ***
SEXNEWFemale	6.29e-09 ***
HOUSEHOLDSIZENEW3 or 4	2.05e-06 ***
HOUSEHOLDSIZENEW5 or 6	4.29e-15 ***
HOUSEHOLDSIZENEW7 and more	0.000105 ***
AGEGROUPNEW45-54	0.665607
AGEGROUPNEW55-64	6.29e-06 ***
AGEGROUPNEW65-80	< 2e-16 ***
INCOMENEW25000 to < 50000	< 2e-16 ***
INCOMENEW>=50000	< 2e-16 ***
INCOMENEWUnknown	0.054760 .
RACEGRNon-Hispanic black	1.40e-09 ***
RACEGROthers	< 2e-16 ***
RACEGRHispanic	3.42e-12 ***
EDUCANEWHigh School Graduate to < 4 years of college	2.50e-14 ***
EDUCANEW>= 4 years of college	< 2e-16 ***
EMPLOY1NEWUnemployed	< 2e-16 ***
EMPLOY1NEWHomemaker	8.14e-13 ***
EMPLOY1NEWRetired	< 2e-16 ***
EMPLOY1NEWUnable to work	< 2e-16 ***
MARITALNEWDivorced/widowed/Separated	< 2e-16 ***
MARITALNEWNever married	0.854992

Zero-inflation model coefficients (binomial with logit link):

	Estimate
(Intercept)	0.934557
RENTHOM1Summary1own home	0.127629
SEXNEWFemale	-0.491124
HOUSEHOLDSIZENEW3 or 4	-0.206977
HOUSEHOLDSIZENEW5 or 6	-0.211747
HOUSEHOLDSIZENEW7 and more	-0.130026
AGEGROUPNEW45-54	0.013478
AGEGROUPNEW55-64	0.228143
AGEGROUPNEW65-80	0.711614

INCOMENEW25000 to < 50000	0.179567
INCOMENEW>=50000	0.311106
INCOMENEWUnknown	0.458008
RACEGRNon-Hispanic black	0.127878
RACEGROthers	0.051851
RACEGRHispanic	0.063804
EDUCANEWHigh School Graduate to < 4 years of college	-0.008404
EDUCANEW>= 4 years of college	-0.091265
EMPLOY1NEWUnemployed	-0.648381
EMPLOY1NEWHomemaker	0.005111
EMPLOY1NEWRetired	-0.034353
EMPLOY1NEWUnable to work	-1.277703
MARITALNEWDivorced/widowed/Separated	-0.146758
MARITALNEWNever married	-0.116406
	Std. Error
(Intercept)	0.070600
RENTHOM1Summary1own home	0.025771
SEXNEWFemale	0.020083
HOUSEHOLDSIZENEW3 or 4	0.026839
HOUSEHOLDSIZENEW5 or 6	0.044872
HOUSEHOLDSIZENEW7 and more	0.100999
AGEGROUPNEW45-54	0.034843
AGEGROUPNEW55-64	0.036173
AGEGROUPNEW65-80	0.040757
INCOMENEW25000 to < 50000	0.027030
INCOMENEW>=50000	0.029222
INCOMENEWUnknown	0.033668
RACEGRNon-Hispanic black	0.033095
RACEGROthers	0.033828
RACEGRHispanic	0.043330
EDUCANEWHigh School Graduate to < 4 years of college	0.053772
EDUCANEW>= 4 years of college	0.056288
EMPLOY1NEWUnemployed	0.045502
EMPLOY1NEWHomemaker	0.038993
EMPLOY1NEWRetired	0.027362
EMPLOY1NEWUnable to work	0.035172
MARITALNEWDivorced/widowed/Separated	0.022686
MARITALNEWNever married	0.036014
	z value
(Intercept)	13.237
RENTHOM1Summary1own home	4.953
SEXNEWFemale	-24.454
HOUSEHOLDSIZENEW3 or 4	-7.712
HOUSEHOLDSIZENEW5 or 6	-4.719
HOUSEHOLDSIZENEW7 and more	-1.287
AGEGROUPNEW45-54	0.387
AGEGROUPNEW55-64	6.307
AGEGROUPNEW65-80	17.460
INCOMENEW25000 to < 50000	6.643
INCOMENEW>=50000	10.646
INCOMENEWUnknown	13.604
RACEGRNon-Hispanic black	3.864
RACEGROthers	1.533
RACEGRHispanic	1.473
EDUCANEWHigh School Graduate to < 4 years of college	-0.156
EDUCANEW>= 4 years of college	-1.621
EMPLOY1NEWUnemployed	-14.250
EMPLOY1NEWHomemaker	0.131
EMPLOY1NEWRetired	-1.255
EMPLOY1NEWUnable to work	-36.327
MARITALNEWDivorced/widowed/Separated	-6.469
MARITALNEWNever married	-3.232
	Pr(> z )
(Intercept)	< 2e-16 ***

RENTHOM1Summary1own home	7.33e-07	***
SEXNEWFemale	< 2e-16	***
HOUSEHOLDSIZENEW3 or 4	1.24e-14	***
HOUSEHOLDSIZENEW5 or 6	2.37e-06	***
HOUSEHOLDSIZENEW7 and more	0.197957	
AGEGROUPNEW45-54	0.698890	
AGEGROUPNEW55-64	2.84e-10	***
AGEGROUPNEW65-80	< 2e-16	***
INCOMENEW25000 to < 50000	3.07e-11	***
INCOMENEW>=50000	< 2e-16	***
INCOMENEWUnknown	< 2e-16	***
RACEGRNon-Hispanic black	0.000112	***
RACEGROthers	0.125322	
RACEGRHispanic	0.140878	
EDUCANEWHigh School Graduate to < 4 years of college	0.875799	
EDUCANEW>= 4 years of college	0.104930	
EMPLOY1NEWUnemployed	< 2e-16	***
EMPLOY1NEWHomemaker	0.895720	
EMPLOY1NEWRetired	0.209298	
EMPLOY1NEWUnable to work	< 2e-16	***
MARITALNEWDivorced/widowed/separated	9.87e-11	***
MARITALNEWNever married	0.001228	**

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Number of iterations in BFGS optimization: 54

Log-likelihood: -1.475e+05 on 46 Df

```
> logLik(ZIPoisson)
'log Lik.' -147471.5 (df=46)
```

**##Zero-inflated Negative Binomial**

```
> ZIPnegbin <- zeroinfl(NEWDATA$MENTHLTH ~ RENTHOM1Summary1 + SEXNEW +
HOUSEHOLDSIZENEW + AGEGROUPNEW + INCOMENEW + RACEGR + EDUCANEW + EMPLOY1NEW +
MARITALNEW, dist="negbin")
```

```
> summary(ZIPnegbin)
```

Call:

```
zeroinfl(formula = NEWDATA$MENTHLTH ~ RENTHOM1Summary1 + SEXNEW +
HOUSEHOLDSIZENEW + AGEGROUPNEW + INCOMENEW + RACEGR + EDUCANEW +
EMPLOY1NEW + MARITALNEW, dist = "negbin")
```

Pearson residuals:

Min	1Q	Median	3Q	Max
-0.7161	-0.3809	-0.3171	-0.2321	9.0174

Count model coefficients (negbin with log link):

	Estimate
(Intercept)	2.329364
RENTHOM1Summary1own home	-0.053358
SEXNEWFemale	-0.027457
HOUSEHOLDSIZENEW3 or 4	-0.032751
HOUSEHOLDSIZENEW5 or 6	-0.102358
HOUSEHOLDSIZENEW7 and more	-0.088838
AGEGROUPNEW45-54	-0.005365
AGEGROUPNEW55-64	-0.020946
AGEGROUPNEW65-80	-0.054782
INCOMENEW25000 to < 50000	-0.125954
INCOMENEW>=50000	-0.245994
INCOMENEWUnknown	-0.019726
RACEGRNon-Hispanic black	-0.005849
RACEGROthers	0.100953
RACEGRHispanic	0.082244
EDUCANEWHigh School Graduate to < 4 years of college	-0.116373
EDUCANEW>= 4 years of college	-0.294891
EMPLOY1NEWUnemployed	0.331209
EMPLOY1NEWHomemaker	0.076288
EMPLOY1NEWRetired	0.126415
EMPLOY1NEWUnable to work	0.607805
MARITALNEWDivorced/widowed/Separated	0.125993
MARITALNEWNever married	0.012431
Log(theta)	-0.323101
	Std. Error
(Intercept)	0.066010
RENTHOM1Summary1own home	0.024020
SEXNEWFemale	0.020095
HOUSEHOLDSIZENEW3 or 4	0.025069
HOUSEHOLDSIZENEW5 or 6	0.042083
HOUSEHOLDSIZENEW7 and more	0.094499
AGEGROUPNEW45-54	0.032691
AGEGROUPNEW55-64	0.033516
AGEGROUPNEW65-80	0.038635
INCOMENEW25000 to < 50000	0.026179
INCOMENEW>=50000	0.027841
INCOMENEWUnknown	0.033329
RACEGRNon-Hispanic black	0.030893
RACEGROthers	0.033041
RACEGRHispanic	0.040962
EDUCANEWHigh School Graduate to < 4 years of college	0.050141
EDUCANEW>= 4 years of college	0.052783
EMPLOY1NEWUnemployed	0.040971
EMPLOY1NEWHomemaker	0.039489
EMPLOY1NEWRetired	0.028138
EMPLOY1NEWUnable to work	0.028724
MARITALNEWDivorced/widowed/Separated	0.021705
MARITALNEWNever married	0.034329
Log(theta)	0.020033
	z value
(Intercept)	35.288
RENTHOM1Summary1own home	-2.221
SEXNEWFemale	-1.366
HOUSEHOLDSIZENEW3 or 4	-1.306
HOUSEHOLDSIZENEW5 or 6	-2.432
HOUSEHOLDSIZENEW7 and more	-0.940
AGEGROUPNEW45-54	-0.164



AGEGROUPNEW55-64	-0.625	
AGEGROUPNEW65-80	-1.418	
INCOMENEW25000 to < 50000	-4.811	
INCOMENEW>=50000	-8.836	
INCOMENEWUnknown	-0.592	
RACEGRNon-Hispanic black	-0.189	
RACEGROthers	3.055	
RACEGRHispanic	2.008	
EDUCANEWHigh School Graduate to < 4 years of college	-2.321	
EDUCANEW>= 4 years of college	-5.587	
EMPLOY1NEWUnemployed	8.084	
EMPLOY1NEWHomemaker	1.932	
EMPLOY1NEWRetired	4.493	
EMPLOY1NEWUnable to work	21.160	
MARITALNEWDivorced/widowed/Separated	5.805	
MARITALNEWNever married	0.362	
Log(theta)	-16.129	
(Intercept)	< 2e-16	***
RENTHOM1Summary1own home	0.02632	*
SEXNEWFemale	0.17183	
HOUSEHOLDSIZENEW3 or 4	0.19140	
HOUSEHOLDSIZENEW5 or 6	0.01500	*
HOUSEHOLDSIZENEW7 and more	0.34717	
AGEGROUPNEW45-54	0.86963	
AGEGROUPNEW55-64	0.53201	
AGEGROUPNEW65-80	0.15621	
INCOMENEW25000 to < 50000	1.50e-06	***
INCOMENEW>=50000	< 2e-16	***
INCOMENEWUnknown	0.55396	
RACEGRNon-Hispanic black	0.84982	
RACEGROthers	0.00225	**
RACEGRHispanic	0.04466	*
EDUCANEWHigh School Graduate to < 4 years of college	0.02029	*
EDUCANEW>= 4 years of college	2.31e-08	***
EMPLOY1NEWUnemployed	6.27e-16	***
EMPLOY1NEWHomemaker	0.05337	.
EMPLOY1NEWRetired	7.04e-06	***
EMPLOY1NEWUnable to work	< 2e-16	***
MARITALNEWDivorced/widowed/Separated	6.44e-09	***
MARITALNEWNever married	0.71727	
Log(theta)	< 2e-16	***

#### Zero-inflation model coefficients (binomial with logit link):

	Estimate
(Intercept)	0.70678
RENTHOM1Summary1own home	0.13373
SEXNEWFemale	-0.53574
HOUSEHOLDSIZENEW3 or 4	-0.23647
HOUSEHOLDSIZENEW5 or 6	-0.26043
HOUSEHOLDSIZENEW7 and more	-0.16261
AGEGROUPNEW45-54	0.01743
AGEGROUPNEW55-64	0.25053
AGEGROUPNEW65-80	0.76048
INCOMENEW25000 to < 50000	0.17404
INCOMENEW>=50000	0.29028
INCOMENEWUnknown	0.49280
RACEGRNon-Hispanic black	0.13000
RACEGROthers	0.07391
RACEGRHispanic	0.07838
EDUCANEWHigh School Graduate to < 4 years of college	-0.03299
EDUCANEW>= 4 years of college	-0.16059
EMPLOY1NEWUnemployed	-0.65695
EMPLOY1NEWHomemaker	0.02023



EMPLOY1NEWRetired	-0.01820
EMPLOY1NEWUnable to work	-1.31637
MARITALNEWDivorced/widowed/Separated	-0.13176
MARITALNEWNever married	-0.12606
	Std. Error
(Intercept)	0.07859
RENTHOM1Summary1own home	0.02887
SEXNEWFemale	0.02214
HOUSEHOLDSIZENEW3 or 4	0.03030
HOUSEHOLDSIZENEW5 or 6	0.05175
HOUSEHOLDSIZENEW7 and more	0.11573
AGEGROUPNEW45-54	0.04033
AGEGROUPNEW55-64	0.04136
AGEGROUPNEW65-80	0.04601
INCOMENEW25000 to < 50000	0.02983
INCOMENEW>=50000	0.03226
INCOMENEWUnknown	0.03657
RACEGRNon-Hispanic black	0.03710
RACEGROthers	0.03740
RACEGRHispanic	0.04829
EDUCANEWHigh School Graduate to < 4 years of college	0.05903
EDUCANEW>= 4 years of college	0.06191
EMPLOY1NEWUnemployed	0.05261
EMPLOY1NEWHomemaker	0.04344
EMPLOY1NEWRetired	0.02990
EMPLOY1NEWUnable to work	0.04135
MARITALNEWDivorced/widowed/Separated	0.02488
MARITALNEWNever married	0.04057
	z value
(Intercept)	8.994
RENTHOM1Summary1own home	4.632
SEXNEWFemale	-24.202
HOUSEHOLDSIZENEW3 or 4	-7.804
HOUSEHOLDSIZENEW5 or 6	-5.032
HOUSEHOLDSIZENEW7 and more	-1.405
AGEGROUPNEW45-54	0.432
AGEGROUPNEW55-64	6.057
AGEGROUPNEW65-80	16.528
INCOMENEW25000 to < 50000	5.834
INCOMENEW>=50000	8.997
INCOMENEWUnknown	13.474
RACEGRNon-Hispanic black	3.504
RACEGROthers	1.976
RACEGRHispanic	1.623
EDUCANEWHigh School Graduate to < 4 years of college	-0.559
EDUCANEW>= 4 years of college	-2.594
EMPLOY1NEWUnemployed	-12.488
EMPLOY1NEWHomemaker	0.466
EMPLOY1NEWRetired	-0.609
EMPLOY1NEWUnable to work	-31.837
MARITALNEWDivorced/widowed/Separated	-5.297
MARITALNEWNever married	-3.107
	Pr(> z )
(Intercept)	< 2e-16 ***
RENTHOM1Summary1own home	3.62e-06 ***
SEXNEWFemale	< 2e-16 ***
HOUSEHOLDSIZENEW3 or 4	5.99e-15 ***
HOUSEHOLDSIZENEW5 or 6	4.84e-07 ***
HOUSEHOLDSIZENEW7 and more	0.159994
AGEGROUPNEW45-54	0.665603
AGEGROUPNEW55-64	1.38e-09 ***
AGEGROUPNEW65-80	< 2e-16 ***
INCOMENEW25000 to < 50000	5.41e-09 ***
INCOMENEW>=50000	< 2e-16 ***

INCOMENewUnknown	< 2e-16	***
RACEGRNon-Hispanic black	0.000459	***
RACEGROthers	0.048108	*
RACEGRHispanic	0.104551	
EDUCANEWHigh School Graduate to < 4 years of college	0.576240	
EDUCANEW<= 4 years of college	0.009492	**
EMPLOY1NEWUnemployed	< 2e-16	***
EMPLOY1NEWHomemaker	0.641482	
EMPLOY1NEWRetired	0.542611	
EMPLOY1NEWUnable to work	< 2e-16	***
MARITALNEWDivorced/widowed/separated	1.18e-07	***
MARITALNEWNever married	0.001890	**

---  
 Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Theta = 0.7239  
 Number of iterations in BFGS optimization: 57  
 Log-likelihood: -9.625e+04 on 47 Df

## ## Log-likelihood

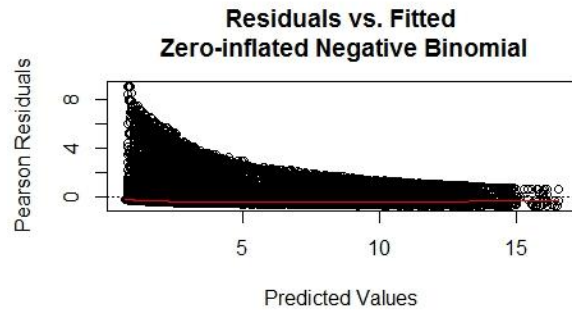
```
> logLik(ZIPnegbin)
'log Lik.' -96255 (df=47)
```

## Vuong Test to see if zero-inflated negative binomial is better than negative binomial

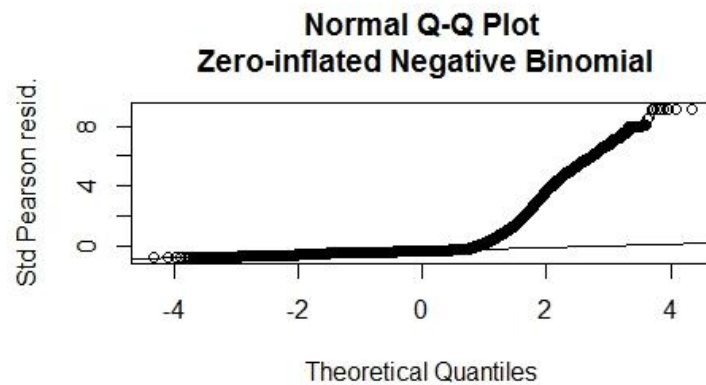
```
> vuong(ZIPnegbin, NEGBIN)
Vuong Non-Nested Hypothesis Test-Statistic: -86.54948
(test-statistic is asymptotically distributed N(0,1) under the
null that the models are indistinguishable)
in this case:
model2 > model1, with p-value < 2.22e-16
```

## ## Plotting zero inflated negative binomial diagnostic plots

```
> plotpanel1 <- function(fit, ...){
+   plot(x=predict(fit), y=residuals(fit, type="pearson"),
+        xlab="Predicted Values", ylab="Pearson Residuals", ...)
+   abline(h=0, lty=3)
+   lines(lowess(x=predict(fit), y=resid(fit, type="pearson")),
+         col="red")
+ }
> plotpanel1(ZIPnegbin, main="Residuals vs. Fitted\n Zero-inflated Negative
Binomial")
```



```
> plotpanel2 <- function(fit, ...){
+   resids <- scale(residuals(fit, type="pearson"))
+   qqnorm(resids, ylab="Std Pearson resid.", ...)
+   qqline(resids)
+ }
> plotpanel2(ZIPnegbin, main="Normal Q-Q Plot\n Zero-inflated Negative
Binomial")
```



```
##Binding the results together.
```

Error in summary(LINEAR\$RENTHOM1Summary1)\$coef :

\$ operator is invalid for atomic vectors

```
> lineartable <- cbind(coef(LINEAR), summary(LINEAR)$coef[,2],  
p=summary(LINEAR)$coef[, 4], summary(LINEAR)$r.squared)
```

(Intercept)	2.02321777	0.09895371
RENTHOM1Summary1own home	0.15001449	0.03631047
SEXNEWFemale	-0.30949662	0.03178869
HOUSEHOLDSIZENEW3 or 4	-0.09150387	0.04174473
HOUSEHOLDSIZENEW5 or 6	0.03046726	0.07284022
HOUSEHOLDSIZENEW7 and more	0.03425651	0.15263647
AGEGROUPNEW45-54	0.02779975	0.05426136
AGEGROUPNEW55-64	0.29392008	0.05600310
AGEGROUPNEW65-80	0.76668220	0.06333820
INCOMENEW25000 to < 50000	0.29811218	0.04039560
INCOMENEW>=50000	0.58566274	0.04613639
INCOMENEWUnknown	0.37531598	0.04885615
RACEGRNon-Hispanic black	0.22042019	0.04828767
RACEGROthers	-0.07358821	0.05068671
RACEGRHispanic	-0.08808176	0.06176631
EDUCANEWHigh School Graduate to < 4 years of college	0.14234950	0.06883049
EDUCANEW>= 4 years of college	0.30529759	0.07478811
EMPLOY1NEWUnemployed	-0.95790547	0.06037165
EMPLOY1NEWHomemaker	-0.14085710	0.06462557
EMPLOY1NEWRetired	-0.21060872	0.04571170
EMPLOY1NEWUnable to work	-1.76203243	0.04410016
MARITALNEWDivorced/widowed/Separated	-0.25570322	0.03508749
MARITALNEWNever married	-0.07348124	0.05421425

		p
(Intercept)	6.506957e-93	-18238.98
RENTHOM1Summary1own home	3.604993e-05	-18238.98
SEXNEWFemale	2.115975e-22	-18238.98
HOUSEHOLDSIZENEW3 or 4	2.838053e-02	-18238.98
HOUSEHOLDSIZENEW5 or 6	6.757459e-01	-18238.98
HOUSEHOLDSIZENEW7 and more	8.224212e-01	-18238.98
AGEGROUPNEW45-54	6.084197e-01	-18238.98
AGEGROUPNEW55-64	1.535243e-07	-18238.98
AGEGROUPNEW65-80	9.988067e-34	-18238.98
INCOMENEW25000 to < 50000	1.585066e-13	-18238.98
INCOMENEW>=50000	6.371018e-37	-18238.98
INCOMENEWUnknown	1.565470e-14	-18238.98
RACEGRNon-Hispanic black	5.001362e-06	-18238.98
RACEGROthers	1.465504e-01	-18238.98
RACEGRHispanic	1.538543e-01	-18238.98
EDUCANEWHigh School Graduate to < 4 years of college	3.862902e-02	-18238.98
EDUCANEW>= 4 years of college	4.461760e-05	-18238.98
EMPLOY1NEWUnemployed	1.075737e-56	-18238.98
EMPLOY1NEWHomemaker	2.928803e-02	-18238.98
EMPLOY1NEWRetired	4.078797e-06	-18238.98
EMPLOY1NEWUnable to work	0.000000e+00	-18238.98
MARITALNEWDivorced/widowed/Separated	3.155503e-13	-18238.98
MARITALNEWNever married	1.752946e-01	-18238.98

```
> logistictable <- cbind(coef(logistic), summary(logistic)$coef[,2],  
p=summary(logistic)$coef[, 4], summary(logistic)$AIC, logLik(logistic))  
> cbind(coef(logistic), summary(logistic)$coef[,2],  
p=summary(logistic)$coef[, 4], AIC(logistic), logLik(logistic))
```

(Intercept)	2.02321777	0.09895371
RENTHOM1Summary1own home	0.15001449	0.03631047
SEXNEWFemale	-0.30949662	0.03178869
HOUSEHOLDSIZENEW3 or 4	-0.09150387	0.04174473
HOUSEHOLDSIZENEW5 or 6	0.03046726	0.07284022
HOUSEHOLDSIZENEW7 and more	0.03425651	0.15263647
AGEGROUPNEW45-54	0.02779975	0.05426136

AGEGROUPNEW55-64	0.29392008	0.05600310
AGEGROUPNEW65-80	0.76668220	0.06333820
INCOMENEW25000 to < 50000	0.29811218	0.04039560
INCOMENEW>=50000	0.58566274	0.04613639
INCOMENEWUnknown	0.37531598	0.04885615
RACEGRNon-Hispanic black	0.22042019	0.04828767
RACEGROthers	-0.07358821	0.05068671
RACEGRHispanic	-0.08808176	0.06176631
EDUCANEWHigh School Graduate to < 4 years of college	0.14234950	0.06883049
EDUCANEW>= 4 years of college	0.30529759	0.07478811
EMPLOY1NEWUnemployed	-0.95790547	0.06037165
EMPLOY1NEWHomemaker	-0.14085710	0.06462557
EMPLOY1NEWRetired	-0.21060872	0.04571170
EMPLOY1NEWUnable to work	-1.76203243	0.04410016
MARITALNEWDivorced/widowed/Separated	-0.25570322	0.03508749
MARITALNEWNever married	-0.07348124	0.05421425

		p
(Intercept)	6.506957e-93	36523.97
RENTHOM1Summary1own home	3.604993e-05	36523.97
SEXNEWFemale	2.115975e-22	36523.97
HOUSEHOLDSIZENEW3 or 4	2.838053e-02	36523.97
HOUSEHOLDSIZENEW5 or 6	6.757459e-01	36523.97
HOUSEHOLDSIZENEW7 and more	8.224212e-01	36523.97
AGEGROUPNEW45-54	6.084197e-01	36523.97
AGEGROUPNEW55-64	1.535243e-07	36523.97
AGEGROUPNEW65-80	9.988067e-34	36523.97
INCOMENEW25000 to < 50000	1.585066e-13	36523.97
INCOMENEW>=50000	6.371018e-37	36523.97
INCOMENEWUnknown	1.565470e-14	36523.97
RACEGRNon-Hispanic black	5.001362e-06	36523.97
RACEGROthers	1.465504e-01	36523.97
RACEGRHispanic	1.538543e-01	36523.97
EDUCANEWHigh School Graduate to < 4 years of college	3.862902e-02	36523.97
EDUCANEW>= 4 years of college	4.461760e-05	36523.97
EMPLOY1NEWUnemployed	1.075737e-56	36523.97
EMPLOY1NEWHomemaker	2.928803e-02	36523.97
EMPLOY1NEWRetired	4.078797e-06	36523.97
EMPLOY1NEWUnable to work	0.000000e+00	36523.97
MARITALNEWDivorced/widowed/Separated	3.155503e-13	36523.97
MARITALNEWNever married	1.752946e-01	36523.97

(Intercept)	-18238.98
RENTHOM1Summary1own home	-18238.98
SEXNEWFemale	-18238.98
HOUSEHOLDSIZENEW3 or 4	-18238.98
HOUSEHOLDSIZENEW5 or 6	-18238.98
HOUSEHOLDSIZENEW7 and more	-18238.98
AGEGROUPNEW45-54	-18238.98
AGEGROUPNEW55-64	-18238.98
AGEGROUPNEW65-80	-18238.98
INCOMENEW25000 to < 50000	-18238.98
INCOMENEW>=50000	-18238.98
INCOMENEWUnknown	-18238.98
RACEGRNon-Hispanic black	-18238.98
RACEGROthers	-18238.98
RACEGRHispanic	-18238.98
EDUCANEWHigh School Graduate to < 4 years of college	-18238.98
EDUCANEW>= 4 years of college	-18238.98
EMPLOY1NEWUnemployed	-18238.98
EMPLOY1NEWHomemaker	-18238.98
EMPLOY1NEWRetired	-18238.98
EMPLOY1NEWUnable to work	-18238.98
MARITALNEWDivorced/widowed/Separated	-18238.98
MARITALNEWNever married	-18238.98

```
> logistictable <- cbind(coef(logistic), summary(logistic)$coef[,2],
p=summary(logistic)$coef[, 4], AIC(logistic), logLik(logistic))
> POISSON <- glm(NEWDATA$MENTHLTH ~ RENTHOM1Summary1 + SEXNEW +
HOUSEHOLDSIZENEW + AGEGRUPONEW + INCOMENEW + RACEGR + EDUCANEW + EMPLOY1NEW +
MARITALNEW, family=poisson(link="log"))
> cbind(coef(POISSON), summary(POISSON)$coef[,2], p=summary(POISSON)$coef[,
4], AIC(POISSON), logLik(POISSON))
```

(Intercept)	1.225527463	0.015769462
RENTHOM1Summary1own home	-0.109091612	0.005774683
SEXNEWFemale	0.281583342	0.005119495
HOUSEHOLDSIZENEW3 or 4	0.094658473	0.006569301
HOUSEHOLDSIZENEW5 or 6	0.042891452	0.011450847
HOUSEHOLDSIZENEW7 and more	-0.001799483	0.024515952
AGEGRUPONEW45-54	-0.005941004	0.008573111
AGEGRUPONEW55-64	-0.170817857	0.008807515
AGEGRUPONEW65-80	-0.558700622	0.010060390
INCOMENEW25000 to < 50000	-0.222877479	0.006606519
INCOMENEW>=50000	-0.424065532	0.007486259
INCOMENEWUnknown	-0.300507648	0.007991511
RACEGRNon-Hispanic black	-0.137654766	0.007553175
RACEGROthers	0.051453339	0.008053006
RACEGRHispanic	0.018780920	0.009981855
EDUCANEWHigh School Graduate to < 4 years of college	-0.106372646	0.010848042
EDUCANEW>= 4 years of college	-0.193155483	0.011822761
EMPLOY1NEWUnemployed	0.709654455	0.009838078
EMPLOY1NEWHomemaker	0.075313197	0.010660440
EMPLOY1NEWRetired	0.123635270	0.007426993
EMPLOY1NEWUnable to work	1.252716853	0.006944177
MARITALNEWDivorced/widowed/Separated	0.181247702	0.005651818
MARITALNEWNever married	0.053331615	0.008620871

		p
(Intercept)	0.000000e+00	667566.9
RENTHOM1Summary1own home	1.343405e-79	667566.9
SEXNEWFemale	0.000000e+00	667566.9
HOUSEHOLDSIZENEW3 or 4	4.528292e-47	667566.9
HOUSEHOLDSIZENEW5 or 6	1.798905e-04	667566.9
HOUSEHOLDSIZENEW7 and more	9.414874e-01	667566.9
AGEGRUPONEW45-54	4.883214e-01	667566.9
AGEGRUPONEW55-64	8.578735e-84	667566.9
AGEGRUPONEW65-80	0.000000e+00	667566.9
INCOMENEW25000 to < 50000	1.715716e-249	667566.9
INCOMENEW>=50000	0.000000e+00	667566.9
INCOMENEWUnknown	0.000000e+00	667566.9
RACEGRNon-Hispanic black	3.283376e-74	667566.9
RACEGROthers	1.666109e-10	667566.9
RACEGRHispanic	5.990313e-02	667566.9
EDUCANEWHigh School Graduate to < 4 years of college	1.064068e-22	667566.9
EDUCANEW>= 4 years of college	5.331444e-60	667566.9
EMPLOY1NEWUnemployed	0.000000e+00	667566.9
EMPLOY1NEWHomemaker	1.609214e-12	667566.9
EMPLOY1NEWRetired	3.195044e-62	667566.9
EMPLOY1NEWUnable to work	0.000000e+00	667566.9
MARITALNEWDivorced/widowed/Separated	1.196301e-225	667566.9
MARITALNEWNever married	6.157832e-10	667566.9

(Intercept)	-333760.5
RENTHOM1Summary1own home	-333760.5
SEXNEWFemale	-333760.5
HOUSEHOLDSIZENEW3 or 4	-333760.5
HOUSEHOLDSIZENEW5 or 6	-333760.5
HOUSEHOLDSIZENEW7 and more	-333760.5
AGEGRUPONEW45-54	-333760.5
AGEGRUPONEW55-64	-333760.5



```

AGEGROUPNEW65-80 -333760.5
INCOMENEW25000 to < 50000 -333760.5
INCOMENEW>=50000 -333760.5
INCOMENEWUnknown -333760.5
RACEGRNon-Hispanic black -333760.5
RACEGROthers -333760.5
RACEGRHispanic -333760.5
EDUCANEWHigh School Graduate to < 4 years of college -333760.5
EDUCANEW>= 4 years of college -333760.5
EMPLOY1NEWUnemployed -333760.5
EMPLOY1NEWHomemaker -333760.5
EMPLOY1NEWRetired -333760.5
EMPLOY1NEWUnable to work -333760.5
MARITALNEWDivorced/widowed/Separated -333760.5
MARITALNEWNever married -333760.5
> poissontable <- cbind(coef(POISSON), summary(POISSON)$coef[,2],
p=summary(POISSON)$coef[, 4], AIC(POISSON), logLik(POISSON))
> cbind(coef(NEGBIN), summary(NEGBIN)$coef[,2], p=summary(NEGBIN)$coef[, 4],
AIC(NEGBIN), logLik(NEGBIN))

```

```

(Intercept) 1.22936878 0.09882441
RENTHOM1Summary1own home -0.13509315 0.03677329
SEXNEWFemale 0.31974145 0.02674277
HOUSEHOLDSIZENEW3 or 4 0.14724130 0.03823034
HOUSEHOLDSIZENEW5 or 6 0.10049835 0.06554410
HOUSEHOLDSIZENEW7 and more 0.06928941 0.14758608
AGEGROUPNEW45-54 -0.01663632 0.05163356
AGEGROUPNEW55-64 -0.14373582 0.05272559
AGEGROUPNEW65-80 -0.48265407 0.05774899
INCOMENEW25000 to < 50000 -0.25275001 0.03759045
INCOMENEW>=50000 -0.43374092 0.04026341
INCOMENEWUnknown -0.35167015 0.04482861
RACEGRNon-Hispanic black -0.03776562 0.04637914
RACEGROthers 0.09322933 0.04624220
RACEGRHispanic 0.10627032 0.06085736
EDUCANEWHigh School Graduate to < 4 years of college -0.17814510 0.07433029
EDUCANEW>= 4 years of college -0.28064717 0.07762352
EMPLOY1NEWUnemployed 0.70482846 0.06940971
EMPLOY1NEWHomemaker 0.04897583 0.05468753
EMPLOY1NEWRetired 0.10267790 0.03620433
EMPLOY1NEWUnable to work 1.26914091 0.05240077
MARITALNEWDivorced/widowed/Separated 0.21985633 0.03069482
MARITALNEWNever married 0.10792763 0.05078571

```

```

p
(Intercept) 1.586423e-35 198783.5
RENTHOM1Summary1own home 2.390862e-04 198783.5
SEXNEWFemale 6.027035e-33 198783.5
HOUSEHOLDSIZENEW3 or 4 1.174322e-04 198783.5
HOUSEHOLDSIZENEW5 or 6 1.252035e-01 198783.5
HOUSEHOLDSIZENEW7 and more 6.387232e-01 198783.5
AGEGROUPNEW45-54 7.473014e-01 198783.5
AGEGROUPNEW55-64 6.408537e-03 198783.5
AGEGROUPNEW65-80 6.390278e-17 198783.5
INCOMENEW25000 to < 50000 1.770668e-11 198783.5
INCOMENEW>=50000 4.638064e-27 198783.5
INCOMENEWUnknown 4.337483e-15 198783.5
RACEGRNon-Hispanic black 4.154844e-01 198783.5
RACEGROthers 4.378852e-02 198783.5
RACEGRHispanic 8.077282e-02 198783.5
EDUCANEWHigh School Graduate to < 4 years of college 1.654486e-02 198783.5
EDUCANEW>= 4 years of college 2.997784e-04 198783.5
EMPLOY1NEWUnemployed 3.160755e-24 198783.5
EMPLOY1NEWHomemaker 3.704892e-01 198783.5
EMPLOY1NEWRetired 4.567286e-03 198783.5

```

EMPLOY1NEWUnable to work	1.373372e-129	198783.5
MARITALNEWDivorced/widowed/Separated	7.913138e-13	198783.5
MARITALNEWNever married	3.357347e-02	198783.5

(Intercept)	-99367.76
RENTHOM1Summary1own home	-99367.76
SEXNEWFemale	-99367.76
HOUSEHOLDSIZENEW3 or 4	-99367.76
HOUSEHOLDSIZENEW5 or 6	-99367.76
HOUSEHOLDSIZENEW7 and more	-99367.76
AGEGROUPNEW45-54	-99367.76
AGEGROUPNEW55-64	-99367.76
AGEGROUPNEW65-80	-99367.76
INCOMENEW25000 to < 50000	-99367.76
INCOMENEW>=50000	-99367.76
INCOMENEWUnknown	-99367.76
RACEGRNon-Hispanic black	-99367.76
RACEGROthers	-99367.76
RACEGRHispanic	-99367.76
EDUCANEWHigh School Graduate to < 4 years of college	-99367.76
EDUCANEW>= 4 years of college	-99367.76
EMPLOY1NEWUnemployed	-99367.76
EMPLOY1NEWHomemaker	-99367.76
EMPLOY1NEWRetired	-99367.76
EMPLOY1NEWUnable to work	-99367.76
MARITALNEWDivorced/widowed/Separated	-99367.76
MARITALNEWNever married	-99367.76

```

> negbintable <- cbind(coef(NEGBIN), summary(NEGBIN)$coef[,2],
p=summary(NEGBIN)$coef[, 4], AIC(NEGBIN), logLik(NEGBIN))
> write.csv(negbintable, "C:/Users/James/Desktop/NEGBINTABLE.csv")
> total <- rbind(lineartable, logistictable, poissontable, negbintable)

```

## Make into CSV file.

```
> write.csv(total, "C:/Users/James/Desktop/HOMEOWNTABLE.csv")
```