Group_2_Analysis

Group 2

2022/3/8

```
library(tidyverse)
library(skimr)
library(GGally)
library(ggfortify)
library(dplyr)
library(nnet)
library(caret)
library(olsrr)
library(MASS)
```

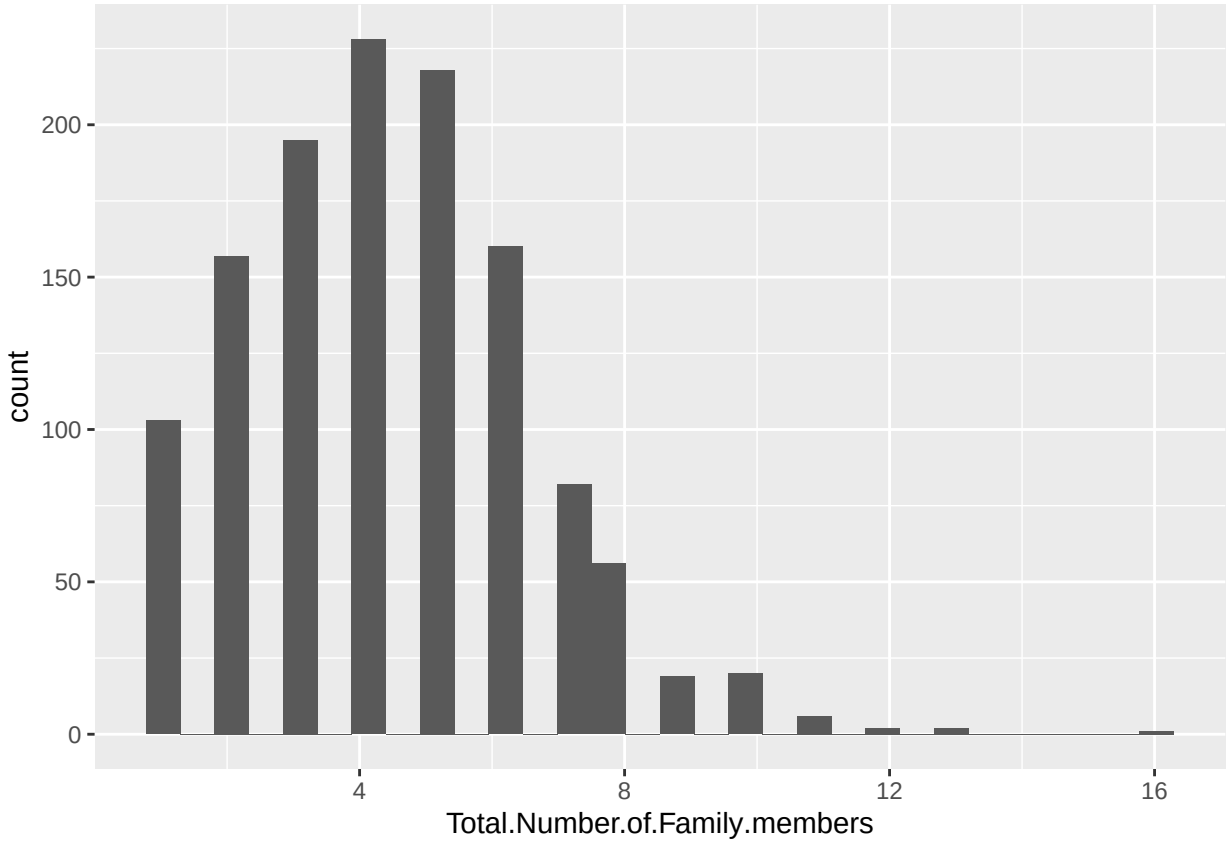
```
# import data
fies = read.csv("dataset2.csv")
```

```
# data summary
fies %>%
  summary()
```

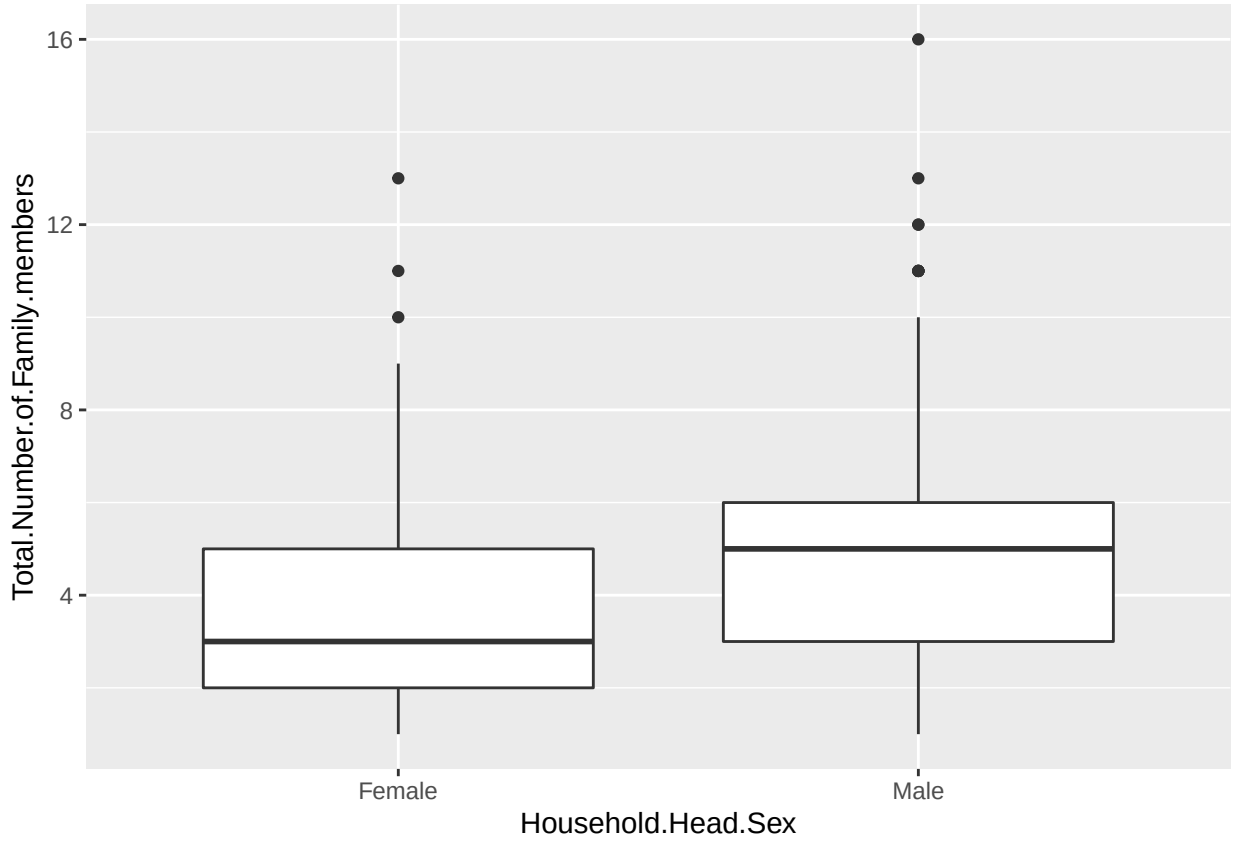
```
 Total.Household.Income    Region          Total.Food.Expenditure
 Min.   :  18784        Length:1249        Min.   : 10488
 1st Qu.:  92101        Class :character   1st Qu.: 43751
 Median : 140483        Mode  :character   Median : 62590
 Mean   : 216685                           Mean   : 70760
 3rd Qu.: 230402                           3rd Qu.: 86708
 Max.   :2891788                           Max.   :413844
 Household.Head.Sex Household.Head.Age Type.of.Household
 Length:1249        Min.   :15.00      Length:1249
 Class :character   1st Qu.:41.00      Class :character
 Mode  :character   Median :51.00      Mode  :character
                    Mean   :51.37
                    3rd Qu.:61.00
                    Max.   :87.00
 Total.Number.of.Family.members House.Floor.Area   House.Age
 Min.   : 1.000                 Min.   :  5.00   Min.   :  0.00
 1st Qu.: 3.000                 1st Qu.: 20.00   1st Qu.:  8.00
 Median : 4.000                 Median : 36.00   Median : 14.00
 Mean   : 4.395                 Mean   : 48.95   Mean   : 16.49
 3rd Qu.: 6.000                 3rd Qu.: 60.00   3rd Qu.: 22.00
 Max.   :16.000                 Max.   :750.00   Max.   :105.00
 Number.of.bedrooms  Electricity
 Min.   :0.000      Min.   :0.0000
 1st Qu.:1.000      1st Qu.:1.0000
```

```
 Median :2.000      Median :1.0000
 Mean   :1.785      Mean   :0.8559
 3rd Qu.:2.000      3rd Qu.:1.0000
 Max.   :7.000      Max.   :1.0000
```

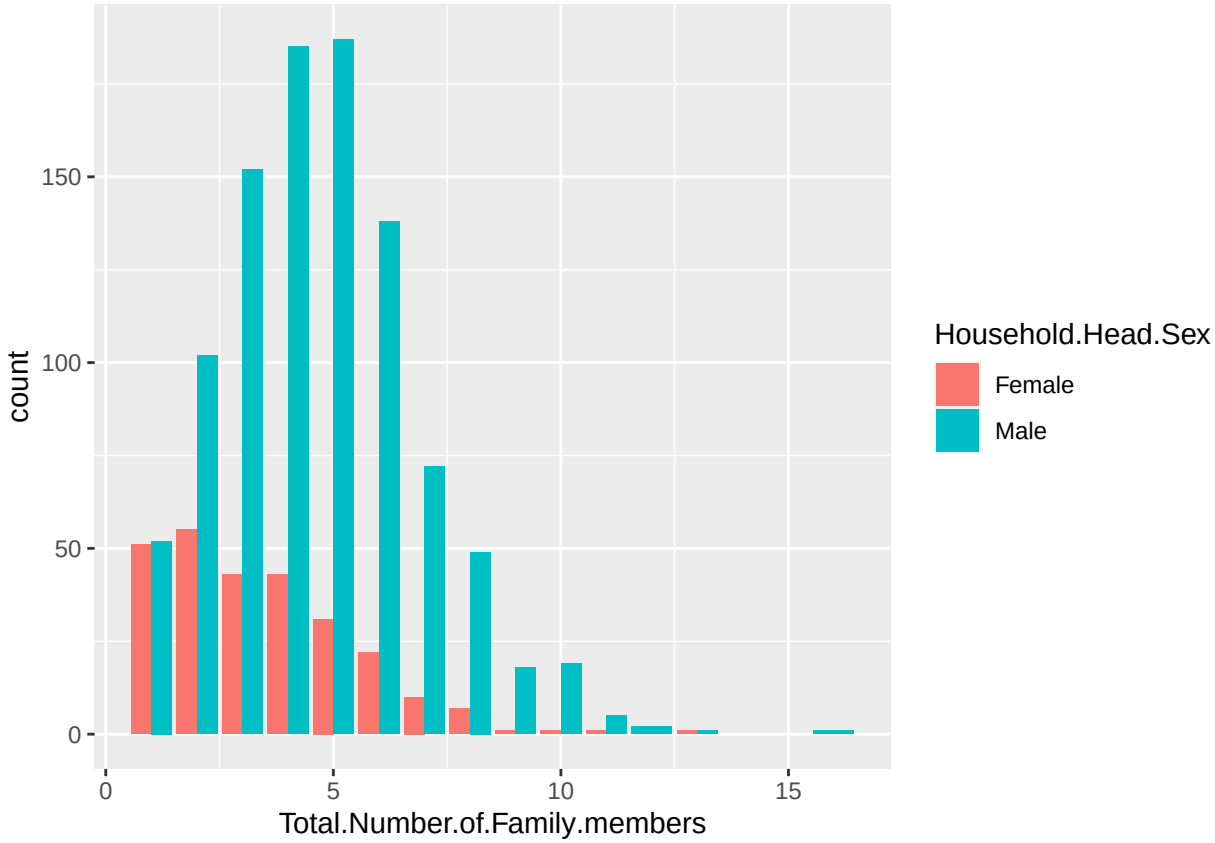
```
#plot the distribution of "Total.Number.of.Family.members"
ggplot(fies, aes(x=Total.Number.of.Family.members)) +
  geom_histogram()
```



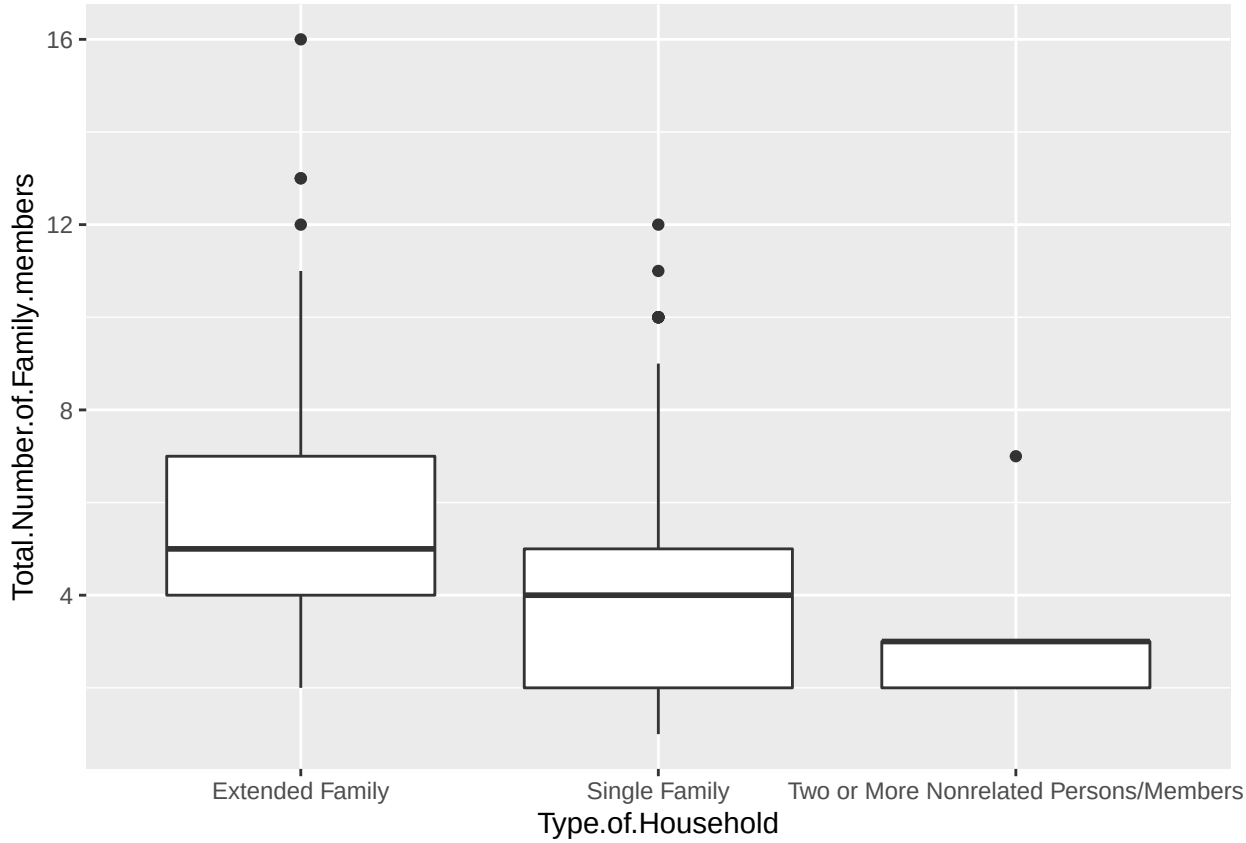

```
#boxplot of "household.Head.Sex" vs "Total.Number.of.Family.members"
ggplot(fies, aes(x=Household.Head.Sex, y=Total.Number.of.Family.members)) +
  geom_boxplot()
```



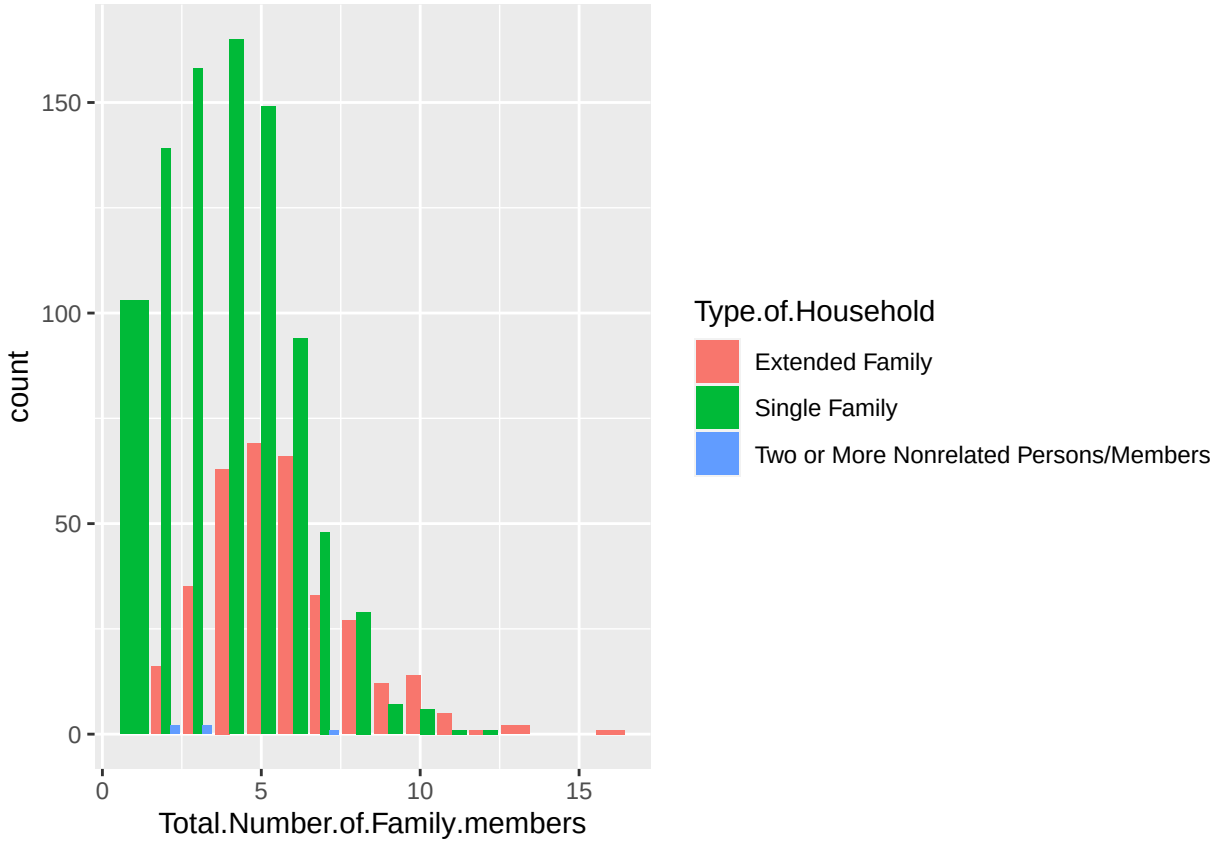

```
#barplot of distribution of
#"Total.Number.of.Family.members" by "Household.Head.Sex"
ggplot(fies, aes(group=Household.Head.Sex,
                 x=Total.Number.of.Family.members,
                 fill=Household.Head.Sex)) +
  geom_bar(position="dodge", stat="count")
```



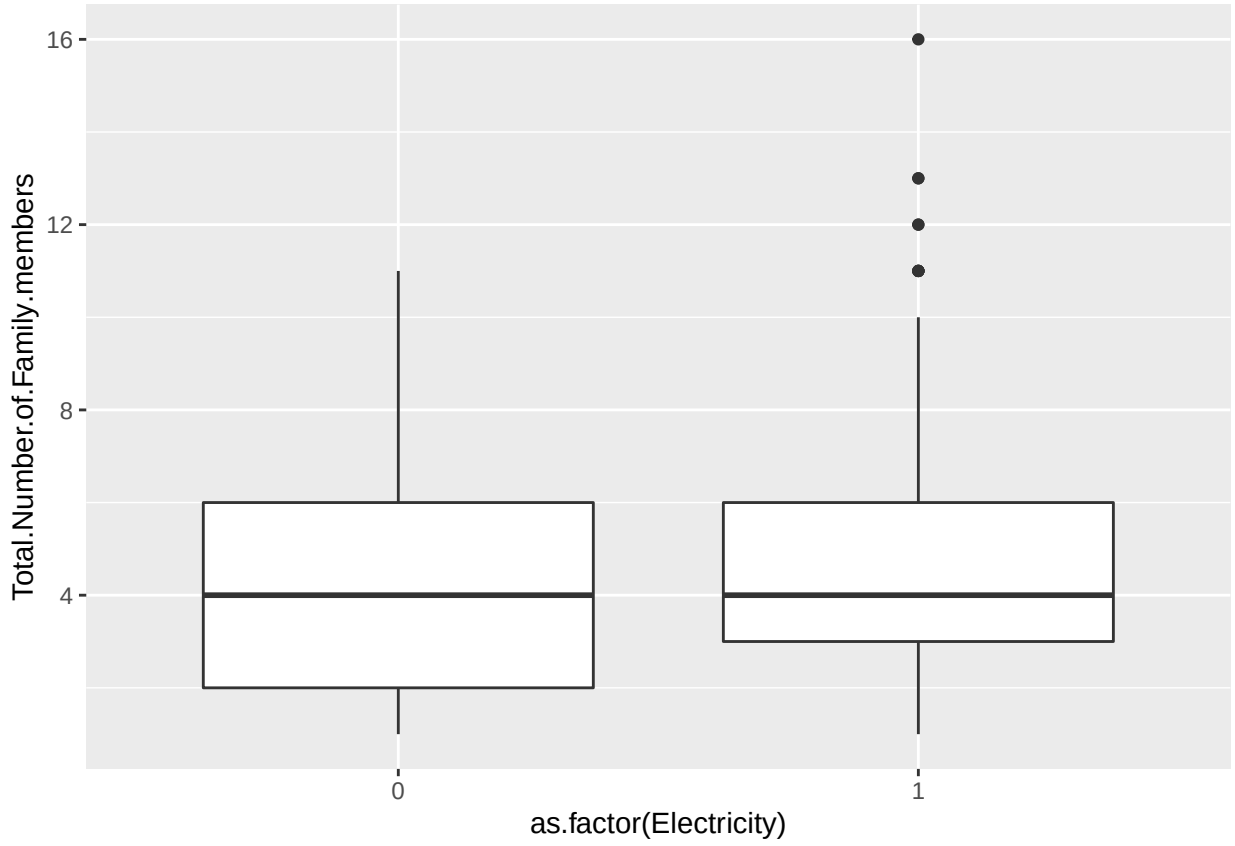

```
#boxplot of "Type.of.Household" vs "Total.Number.of.Family.members"
ggplot(fies, aes(x=Type.of.Household, y=Total.Number.of.Family.members)) +
  geom_boxplot()
```




```
#barplot of distribution of
#"Total.Number.of.Family.members" by "Type.of.Household"
ggplot(fies, aes(group=Type.of.Household,
                 x=Total.Number.of.Family.members,
                 fill=Type.of.Household)) +
  geom_bar(position="dodge", stat="count")
```

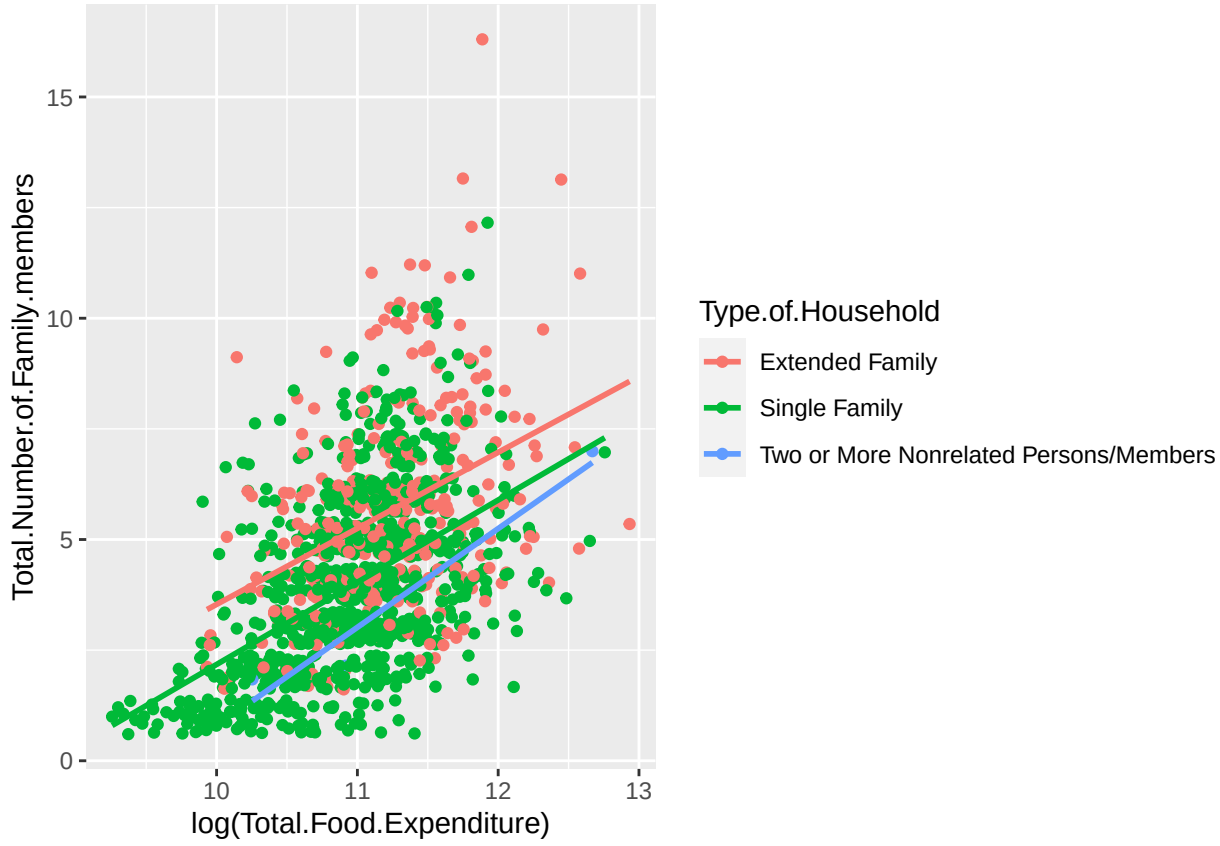



```
#boxplot of "Electricity" vs "Total.Number.of.Family.members"
ggplot(fies, aes(x=as.factor(Electricity), y=Total.Number.of.Family.members)) +
  geom_boxplot()
```

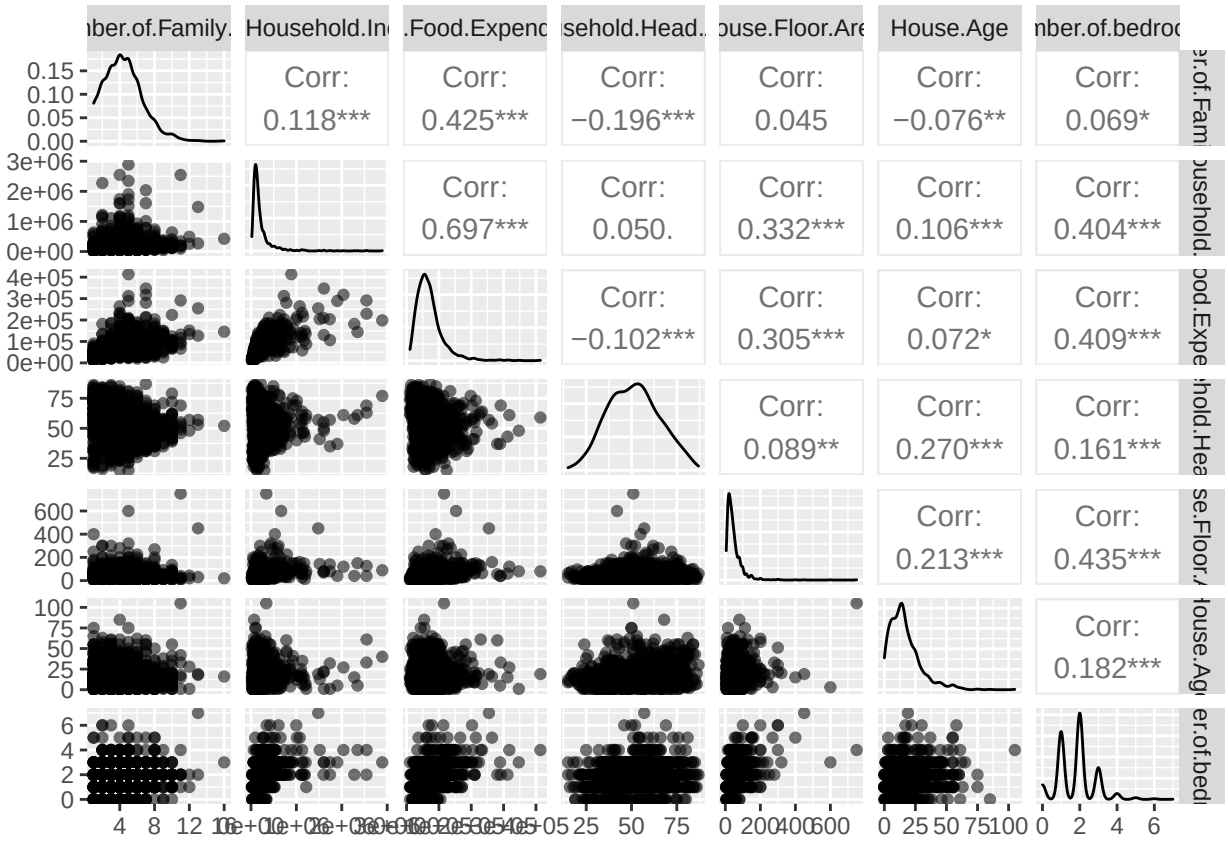



```
ggplot(fies, aes(y=Total.Number.of.Family.members,
                 x=log(Total.Household.Income),
                color = Household.Head.Sex)) +
  geom_jitter()+
  geom_smooth(method = "lm", se = FALSE)
```


```
ggplot(fies, aes(y=Total.Number.of.Family.members,
                 x=log(Total.Food.Expenditure),
                color = Type.of.Household)) +
  geom_jitter()+
  geom_smooth(method = "lm", se = FALSE)
```




```
# continuous variables in the dataset
con_var = c("Total.Household.Income", "Total.Food.Expenditure",
            "Household.Head.Age","House.Floor.Area", "House.Age",
            "Number.of.bedrooms")

# discrete variables in the dataset
dis_var = c("Household.Head.Sex", "Type.of.Household",
            "Electricity", "Total.Number.of.Family.members")

ggpairs(fies[, c("Total.Number.of.Family.members", con_var)], aes(alpha = 0.4))
```



Poisson Regression

Full Model

```
model_pr = glm(Total.Number.of.Family.members ~
                 log(Total.Food.Expenditure) +log(Total.Household.Income) +
                 House.Age +Number.of.bedrooms +Household.Head.Sex +
                 Type.of.Household + Electricity,
                   family = poisson, data=fies)
```

```
summary(model_pr)
```

```
Call:
glm(formula = Total.Number.of.Family.members ~ log(Total.Food.Expenditure) +
    log(Total.Household.Income) + Household.Head.Age + House.Floor.Area +
    House.Age + Number.of.bedrooms + Household.Head.Sex + Type.of.Household +
    Electricity, family = poisson, data = fies)

Deviance Residuals:
     Min        1Q    Median        3Q       Max
-2.37083  -0.58304  -0.08994   0.42434   2.68580
```

```
Coefficients:
                                                    Estimate Std. Error
(Intercept)                                        -3.057e+00  3.467e-01
log(Total.Food.Expenditure)                         7.146e-01  4.622e-02
log(Total.Household.Income)                        -2.487e-01  3.380e-02
Household.Head.Age                                 -3.518e-03  1.113e-03
House.Floor.Area                                   -7.159e-05  3.042e-04
House.Age                                          -2.272e-03  1.174e-03
Number.of.bedrooms                                 -3.188e-02  1.703e-02
Household.Head.SexMale                              1.915e-01  3.757e-02
Type.of.HouseholdSingle Family                     -3.242e-01  3.078e-02
Type.of.HouseholdTwo or More Nonrelated Persons/Members -4.543e-01  2.440e-01
Electricity                                        -7.449e-02  4.138e-02
                                                   z value Pr(>|z|)
(Intercept)                                         -8.816  < 2e-16 ***
log(Total.Food.Expenditure)                         15.460  < 2e-16 ***
log(Total.Household.Income)                         -7.359 1.86e-13 ***
Household.Head.Age                                  -3.161  0.00157 **
House.Floor.Area                                    -0.235  0.81395
House.Age                                           -1.934  0.05309 .
Number.of.bedrooms                                  -1.871  0.06128 .
Household.Head.SexMale                               5.097 3.44e-07 ***
Type.of.HouseholdSingle Family                     -10.535  < 2e-16 ***
Type.of.HouseholdTwo or More Nonrelated Persons/Members  -1.862  0.06266 .
Electricity                                         -1.800  0.07185 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for poisson family taken to be 1)

    Null deviance: 1373.63  on 1248  degrees of freedom
Residual deviance:  755.45  on 1238  degrees of freedom
AIC: 4806.3

Number of Fisher Scoring iterations: 4
```

```
model_pr_opt = step(model_pr)
```

```
Start:  AIC=4806.3
Total.Number.of.Family.members ~ log(Total.Food.Expenditure) +
    log(Total.Household.Income) + Household.Head.Age + House.Floor.Area +
    House.Age + Number.of.bedrooms + Household.Head.Sex + Type.of.Household +
    Electricity

                             Df Deviance    AIC
- House.Floor.Area            1   755.50 4804.4
<none>                            755.45 4806.3
- Electricity                 1   758.65 4807.5
- Number.of.bedrooms          1   758.96 4807.8
- House.Age                   1   759.24 4808.1
- Household.Head.Age          1   765.43 4814.3
- Household.Head.Sex          1   782.39 4831.2
- log(Total.Household.Income) 1   811.29 4860.2
- Type.of.Household           2   865.08 4911.9
```

```
- log(Total.Food.Expenditure)  1  1001.16 5050.0

Step:  AIC=4804.36
Total.Number.of.Family.members ~ log(Total.Food.Expenditure) +
    log(Total.Household.Income) + Household.Head.Age + House.Age +
    Number.of.bedrooms + Household.Head.Sex + Type.of.Household +
    Electricity

                            Df Deviance    AIC
<none>                           755.50 4804.4
- Electricity                1   758.70 4805.6
- House.Age                  1   759.52 4806.4
- Number.of.bedrooms         1   759.61 4806.5
- Household.Head.Age         1   765.47 4812.3
- Household.Head.Sex         1   782.46 4829.3
- log(Total.Household.Income) 1   813.53 4860.4
- Type.of.Household           2   865.08 4909.9
- log(Total.Food.Expenditure) 1  1001.80 5048.7
```

```
summary(model_pr_opt)
```

```
Call:
glm(formula = Total.Number.of.Family.members ~ log(Total.Food.Expenditure) +
    log(Total.Household.Income) + Household.Head.Age + House.Age +
    Number.of.bedrooms + Household.Head.Sex + Type.of.Household +
    Electricity, family = poisson, data = fies)

Deviance Residuals:
     Min        1Q    Median        3Q       Max
-2.36516  -0.58649  -0.09467   0.42940   2.69982

Coefficients:
                                                    Estimate Std. Error
(Intercept)                                        -3.047728   0.344602
log(Total.Food.Expenditure)                         0.715107   0.046178
log(Total.Household.Income)                        -0.250049   0.033332
Household.Head.Age                                 -0.003516   0.001113
House.Age                                          -0.002311   0.001161
Number.of.bedrooms                                 -0.033017   0.016331
Household.Head.SexMale                              0.191592   0.037574
Type.of.HouseholdSingle Family                     -0.324078   0.030772
Type.of.HouseholdTwo or More Nonrelated Persons/Members -0.455303   0.244003
Electricity                                        -0.074456   0.041379
                                                   z value Pr(>|z|)
(Intercept)                                         -8.844  < 2e-16 ***
log(Total.Food.Expenditure)                         15.486  < 2e-16 ***
log(Total.Household.Income)                         -7.502 6.29e-14 ***
Household.Head.Age                                  -3.159  0.00158 **
House.Age                                           -1.990  0.04658 *
Number.of.bedrooms                                  -2.022  0.04321 *
Household.Head.SexMale                               5.099 3.41e-07 ***
Type.of.HouseholdSingle Family                     -10.532  < 2e-16 ***
Type.of.HouseholdTwo or More Nonrelated Persons/Members  -1.866  0.06205 .
Electricity                                         -1.799  0.07196 .
```

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

(Dispersion parameter for poisson family taken to be 1)

    Null deviance: 1373.6  on 1248  degrees of freedom
Residual deviance:  755.5  on 1239  degrees of freedom
AIC: 4804.4

Number of Fisher Scoring iterations: 4
```

Improve a little. Next, try some more complex model.

Model_1

```
model_1 = glm(Total.Number.of.Family.members ~
                log(Total.Food.Expenditure) * Type.of.Household +
                 log(Total.Household.Income) *   Household.Head.Sex
            + Household.Head.Age +
                House.Age +Number.of.bedrooms + Electricity,
                 family = poisson, data=fies)
```

```
summary(model_1)
```

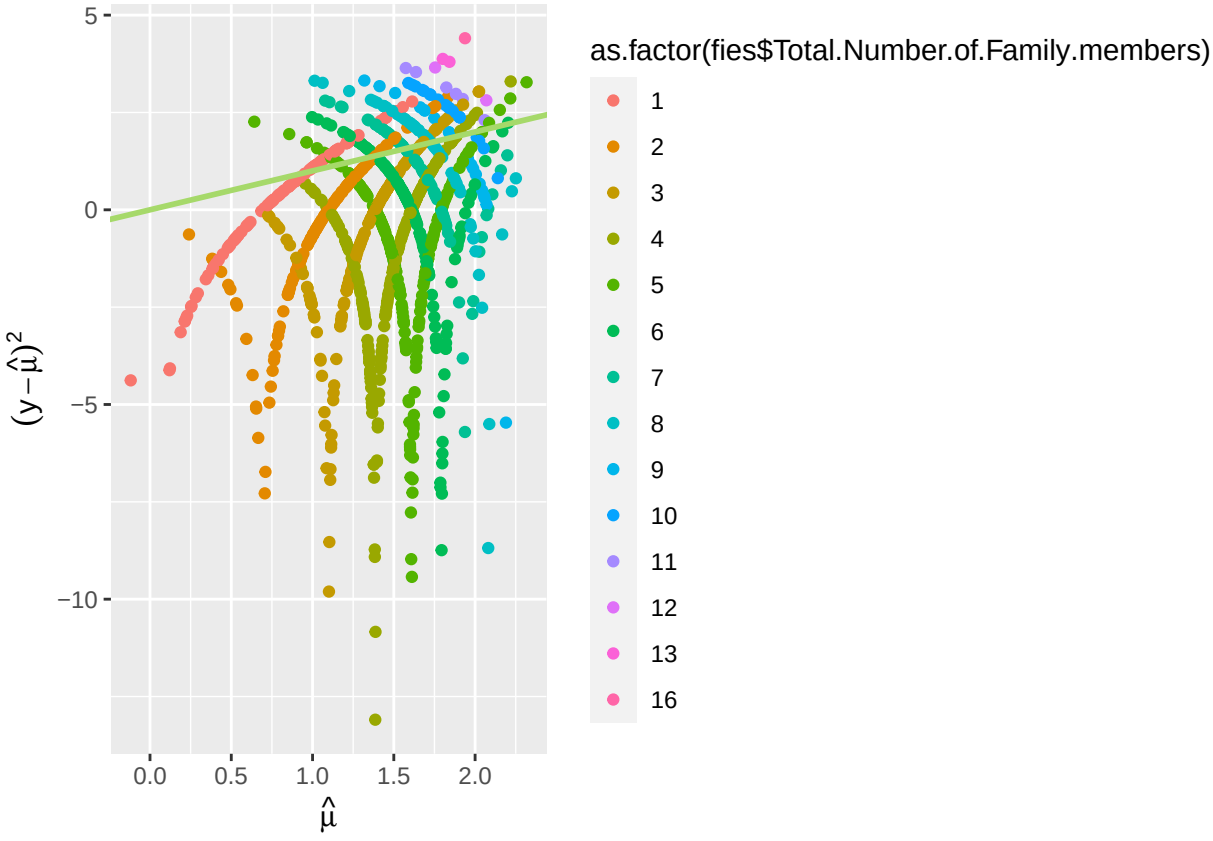
```
Call:
glm(formula = Total.Number.of.Family.members ~ log(Total.Food.Expenditure) *
    Type.of.Household + log(Total.Household.Income) * Household.Head.Sex +
    Household.Head.Age + House.Age + Number.of.bedrooms + Electricity,
    family = poisson, data = fies)

Deviance Residuals:
     Min        1Q    Median        3Q       Max
-2.19336  -0.57457  -0.09103   0.42497   2.93197

Coefficients:
                                                                                Estimate
(Intercept)                                                                    -2.692197
log(Total.Food.Expenditure)                                                     0.608418
Type.of.HouseholdSingle Family                                                 -2.044798
Type.of.HouseholdTwo or More Nonrelated Persons/Members                        -4.468784
log(Total.Household.Income)                                                    -0.181050
Household.Head.SexMale                                                          1.202931
Household.Head.Age                                                             -0.003240
House.Age                                                                      -0.002176
Number.of.bedrooms                                                             -0.030502
Electricity                                                                    -0.078352
log(Total.Food.Expenditure):Type.of.HouseholdSingle Family                      0.153405
log(Total.Food.Expenditure):Type.of.HouseholdTwo or More Nonrelated Persons/Members  0.351841
log(Total.Household.Income):Household.Head.SexMale                             -0.083912
                                                                            Std. Error
(Intercept)                                                                   0.675098
```

```
log(Total.Food.Expenditure)                                                    0.060748
Type.of.HouseholdSingle Family                                                 0.639886
Type.of.HouseholdTwo or More Nonrelated Persons/Members                        3.224211
log(Total.Household.Income)                                                    0.049172
Household.Head.SexMale                                                         0.555312
Household.Head.Age                                                             0.001124
House.Age                                                                      0.001163
Number.of.bedrooms                                                             0.016369
Electricity                                                                    0.041544
log(Total.Food.Expenditure):Type.of.HouseholdSingle Family                     0.056999
log(Total.Food.Expenditure):Type.of.HouseholdTwo or More Nonrelated Persons/Members  0.277561
log(Total.Household.Income):Household.Head.SexMale                             0.045941
                                                                             z value
(Intercept)                                                                   -3.988
log(Total.Food.Expenditure)                                                   10.015
Type.of.HouseholdSingle Family                                                -3.196
Type.of.HouseholdTwo or More Nonrelated Persons/Members                       -1.386
log(Total.Household.Income)                                                   -3.682
Household.Head.SexMale                                                         2.166
Household.Head.Age                                                            -2.882
House.Age                                                                     -1.871
Number.of.bedrooms                                                            -1.863
Electricity                                                                   -1.886
log(Total.Food.Expenditure):Type.of.HouseholdSingle Family                     2.691
log(Total.Food.Expenditure):Type.of.HouseholdTwo or More Nonrelated Persons/Members  1.268
log(Total.Household.Income):Household.Head.SexMale                            -1.827
                                                                            Pr(>|z|)
(Intercept)                                                                 6.67e-05
log(Total.Food.Expenditure)                                                  < 2e-16
Type.of.HouseholdSingle Family                                              0.001396
Type.of.HouseholdTwo or More Nonrelated Persons/Members                     0.165744
log(Total.Household.Income)                                                 0.000231
Household.Head.SexMale                                                      0.030294
Household.Head.Age                                                          0.003947
House.Age                                                                   0.061397
Number.of.bedrooms                                                          0.062401
Electricity                                                                 0.059293
log(Total.Food.Expenditure):Type.of.HouseholdSingle Family                  0.007116
log(Total.Food.Expenditure):Type.of.HouseholdTwo or More Nonrelated Persons/Members 0.204935
log(Total.Household.Income):Household.Head.SexMale                          0.067773

(Intercept)                                                                 ***
log(Total.Food.Expenditure)                                                 ***
Type.of.HouseholdSingle Family                                              **
Type.of.HouseholdTwo or More Nonrelated Persons/Members
log(Total.Household.Income)                                                 ***
Household.Head.SexMale                                                      *
Household.Head.Age                                                          **
House.Age                                                                   .
Number.of.bedrooms                                                          .
Electricity                                                                 .
log(Total.Food.Expenditure):Type.of.HouseholdSingle Family                  **
log(Total.Food.Expenditure):Type.of.HouseholdTwo or More Nonrelated Persons/Members
log(Total.Household.Income):Household.Head.SexMale                          .
```

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

(Dispersion parameter for poisson family taken to be 1)

    Null deviance: 1373.63  on 1248  degrees of freedom
Residual deviance:  745.13  on 1236  degrees of freedom
AIC: 4800

Number of Fisher Scoring iterations: 4
```

Underdispersion

```
ggplot(model_1, aes(x=log(fitted(model_1)),
              y=log((fies$Total.Number.of.Family.members-fitted(model_1))^2),
              color = as.factor(fies$Total.Number.of.Family.members)))+
geom_jitter() +
geom_abline(slope=1, intercept=0, col="#a6d96a", size=1) +
ylab(expression((y-hat(mu))^2)) + xlab(expression(hat(mu)))
```



Underdispersion. Try quasipoisson.

```
X2 <- sum(resid(model_1, type = "pearson")^2)
dp <- X2 / model_1$df.res
dp
```

```
[1] 0.6153645
```

```
drop1(model_1, test = "F")
```

```
Single term deletions

Model:
Total.Number.of.Family.members ~ log(Total.Food.Expenditure) *
    Type.of.Household + log(Total.Household.Income) * Household.Head.Sex +
    Household.Head.Age + House.Age + Number.of.bedrooms + Electricity
                                                Df Deviance    AIC F value
<none>                                               745.13 4800.0
Household.Head.Age                               1   753.43 4806.3 13.7731
House.Age                                        1   748.68 4801.5  5.8871
Number.of.bedrooms                               1   748.61 4801.5  5.7800
Electricity                                      1   748.64 4801.5  5.8225
log(Total.Food.Expenditure):Type.of.Household    2   753.22 4804.1  6.7078
log(Total.Household.Income):Household.Head.Sex   1   748.45 4801.3  5.5098
                                                  Pr(>F)
<none>
Household.Head.Age                              0.0002154 ***
House.Age                                       0.0153943 *
Number.of.bedrooms                              0.0163562 *
Electricity                                     0.0159673 *
log(Total.Food.Expenditure):Type.of.Household   0.0012663 **
log(Total.Household.Income):Household.Head.Sex 0.0190669 *
---
Signif. codes:  0 ’***’ 0.001 ’**’ 0.01 ’*’ 0.05 ’.’ 0.1 ’ ’ 1
```

All coefficients become significant.

Model_2 (quasipoisson)

```
model_2 = glm(Total.Number.of.Family.members ~
                log(Total.Food.Expenditure) * Type.of.Household +
                 log(Total.Household.Income) *   Household.Head.Sex
            + Household.Head.Age +
                House.Age +Number.of.bedrooms + Electricity,
                 family = quasipoisson(link = "log"), data=fies)
```

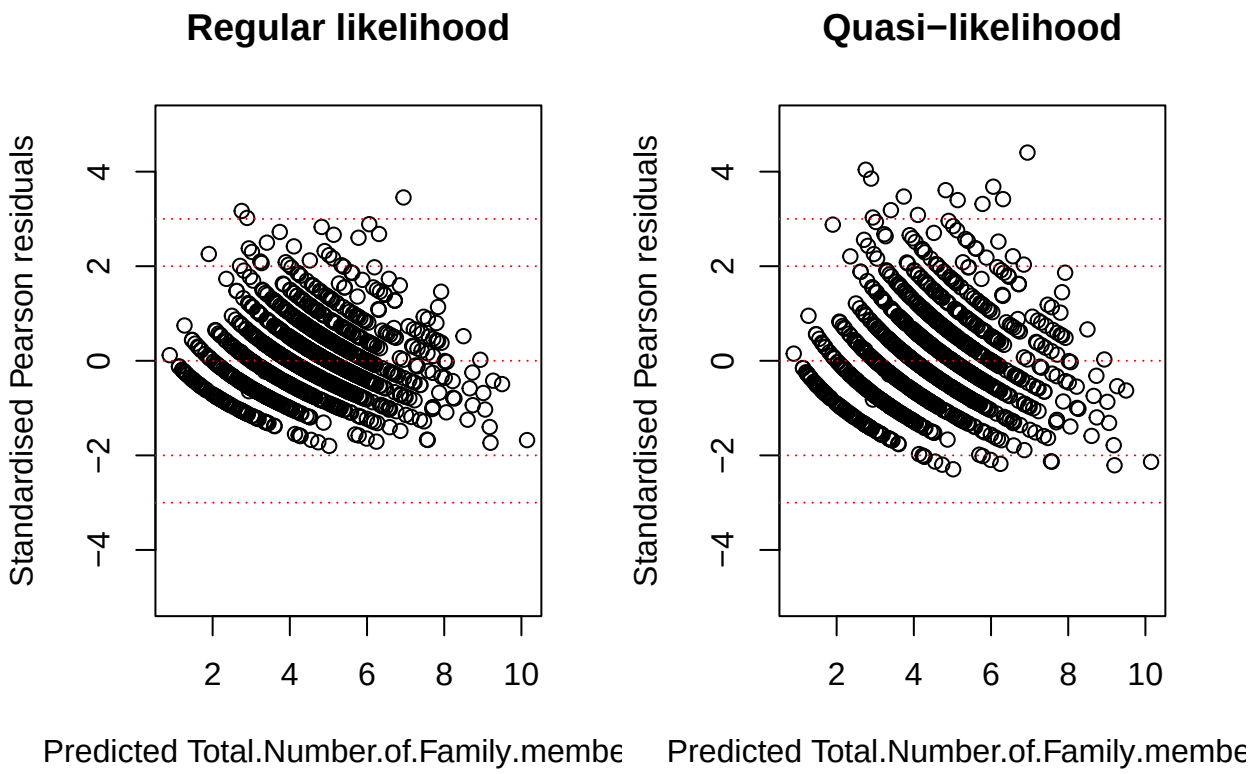
```
summary(model_2)
```

```
Call:
glm(formula = Total.Number.of.Family.members ~ log(Total.Food.Expenditure) *
    Type.of.Household + log(Total.Household.Income) * Household.Head.Sex +
    Household.Head.Age + House.Age + Number.of.bedrooms + Electricity,
    family = quasipoisson(link = "log"), data = fies)
```

```
Deviance Residuals:
     Min        1Q    Median        3Q       Max
-2.19336  -0.57457  -0.09103   0.42497   2.93197

Coefficients:
                                                                                 Estimate
(Intercept)                                                                     -2.6921970
log(Total.Food.Expenditure)                                                      0.6084182
Type.of.HouseholdSingle Family                                                  -2.0447979
Type.of.HouseholdTwo or More Nonrelated Persons/Members                         -4.4687836
log(Total.Household.Income)                                                     -0.1810497
Household.Head.SexMale                                                           1.2029308
Household.Head.Age                                                              -0.0032397
House.Age                                                                       -0.0021759
Number.of.bedrooms                                                              -0.0305018
Electricity                                                                     -0.0783524
log(Total.Food.Expenditure):Type.of.HouseholdSingle Family                       0.1534054
log(Total.Food.Expenditure):Type.of.HouseholdTwo or More Nonrelated Persons/Members  0.3518405
log(Total.Household.Income):Household.Head.SexMale                              -0.0839125
                                                                                Std. Error
(Intercept)                                                                      0.5295830
log(Total.Food.Expenditure)                                                      0.0476543
Type.of.HouseholdSingle Family                                                   0.5019609
Type.of.HouseholdTwo or More Nonrelated Persons/Members                          2.5292430
log(Total.Household.Income)                                                      0.0385733
Household.Head.SexMale                                                           0.4356164
Household.Head.Age                                                               0.0008817
House.Age                                                                        0.0009125
Number.of.bedrooms                                                               0.0128404
Electricity                                                                      0.0325892
log(Total.Food.Expenditure):Type.of.HouseholdSingle Family                       0.0447134
log(Total.Food.Expenditure):Type.of.HouseholdTwo or More Nonrelated Persons/Members  0.2177335
log(Total.Household.Income):Household.Head.SexMale                               0.0360389
                                                                                t value
(Intercept)                                                                      -5.084
log(Total.Food.Expenditure)                                                      12.767
Type.of.HouseholdSingle Family                                                   -4.074
Type.of.HouseholdTwo or More Nonrelated Persons/Members                          -1.767
log(Total.Household.Income)                                                      -4.694
Household.Head.SexMale                                                            2.761
Household.Head.Age                                                               -3.674
House.Age                                                                        -2.385
Number.of.bedrooms                                                               -2.375
Electricity                                                                      -2.404
log(Total.Food.Expenditure):Type.of.HouseholdSingle Family                        3.431
log(Total.Food.Expenditure):Type.of.HouseholdTwo or More Nonrelated Persons/Members   1.616
log(Total.Household.Income):Household.Head.SexMale                               -2.328
                                                                                Pr(>|t|)
(Intercept)                                                                     4.27e-07
log(Total.Food.Expenditure)                                                      < 2e-16
Type.of.HouseholdSingle Family                                                  4.92e-05
Type.of.HouseholdTwo or More Nonrelated Persons/Members                         0.077501
log(Total.Household.Income)                                                     2.98e-06
Household.Head.SexMale                                                          0.005840
```

```
Household.Head.Age                                                                    0.000249
House.Age                                                                             0.017247
Number.of.bedrooms                                                                    0.017679
Electricity                                                                           0.016352
log(Total.Food.Expenditure):Type.of.HouseholdSingle Family                            0.000622
log(Total.Food.Expenditure):Type.of.HouseholdTwo or More Nonrelated Persons/Members 0.106366
log(Total.Household.Income):Household.Head.SexMale                                    0.020052

(Intercept)                                                                           ***
log(Total.Food.Expenditure)                                                           ***
Type.of.HouseholdSingle Family                                                        ***
Type.of.HouseholdTwo or More Nonrelated Persons/Members                               .
log(Total.Household.Income)                                                           ***
Household.Head.SexMale                                                                **
Household.Head.Age                                                                    ***
House.Age                                                                             *
Number.of.bedrooms                                                                    *
Electricity                                                                           *
log(Total.Food.Expenditure):Type.of.HouseholdSingle Family                            ***
log(Total.Food.Expenditure):Type.of.HouseholdTwo or More Nonrelated Persons/Members
log(Total.Household.Income):Household.Head.SexMale                                    *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for quasipoisson family taken to be 0.6153671)

    Null deviance: 1373.63  on 1248  degrees of freedom
Residual deviance:  745.13  on 1236  degrees of freedom
AIC: NA

Number of Fisher Scoring iterations: 4
```

```r
# Residual plots vs. predicted
pred <- predict(model_1, type = "response")
stand.resid <- rstandard(model = model_1, type = "pearson") # Standardised Pearson residuals
par(mfrow=c(1,2))
plot(x = pred, y = stand.resid, xlab = "Predicted Total.Number.of.Family.members", ylab = "Standardised
main = "Regular likelihood", ylim = c(-5,5))
abline(h = c(-3, -2, 0, 2, 3), lty = "dotted", col = "red")

pred <- predict(model_2, type = "response")
stand.resid <- rstandard(model = model_2, type = "pearson") # Standardised Pearson residuals
plot(x = pred, y = stand.resid, xlab = "Predicted Total.Number.of.Family.members", ylab = "Standardised
main = "Quasi-likelihood", ylim = c(-5,5))
abline(h = c(-3, -2, 0, 2, 3), lty = "dotted", col = "red")
```



???

Negative binomial (not used)

```
model_3 = glm.nb(Total.Number.of.Family.members ~
                  log(Total.Food.Expenditure) * Type.of.Household +
                   log(Total.Household.Income) *   Household.Head.Sex
               + Household.Head.Age +
                   House.Age +Number.of.bedrooms + Electricity,
                    data=fies)
```

```
summary(model_3)
```

```
Call:
glm.nb(formula = Total.Number.of.Family.members ~ log(Total.Food.Expenditure) *
    Type.of.Household + log(Total.Household.Income) * Household.Head.Sex +
    Household.Head.Age + House.Age + Number.of.bedrooms + Electricity,
    data = fies, init.theta = 112480.0161, link = log)

Deviance Residuals:
     Min        1Q    Median        3Q       Max
-2.19333  -0.57456  -0.09103   0.42496   2.93184

Coefficients:
```

	Estimate
(Intercept)	-2.692209
log(Total.Food.Expenditure)	0.608420
Type.of.HouseholdSingle Family	-2.044807
Type.of.HouseholdTwo or More Nonrelated Persons/Members	-4.468755
log(Total.Household.Income)	-0.181050
Household.Head.SexMale	1.202934
Household.Head.Age	-0.003240
House.Age	-0.002176
Number.of.bedrooms	-0.030502
Electricity	-0.078353
log(Total.Food.Expenditure):Type.of.HouseholdSingle Family	0.153406
log(Total.Food.Expenditure):Type.of.HouseholdTwo or More Nonrelated Persons/Members	0.351838
log(Total.Household.Income):Household.Head.SexMale	-0.083913
	Std. Error
(Intercept)	0.675114
log(Total.Food.Expenditure)	0.060750
Type.of.HouseholdSingle Family	0.639902
Type.of.HouseholdTwo or More Nonrelated Persons/Members	3.224277
log(Total.Household.Income)	0.049173
Household.Head.SexMale	0.555324
Household.Head.Age	0.001124
House.Age	0.001163
Number.of.bedrooms	0.016369
Electricity	0.041545
log(Total.Food.Expenditure):Type.of.HouseholdSingle Family	0.057001
log(Total.Food.Expenditure):Type.of.HouseholdTwo or More Nonrelated Persons/Members	0.277567
log(Total.Household.Income):Household.Head.SexMale	0.045942
	z value
(Intercept)	-3.988
log(Total.Food.Expenditure)	10.015
Type.of.HouseholdSingle Family	-3.195
Type.of.HouseholdTwo or More Nonrelated Persons/Members	-1.386
log(Total.Household.Income)	-3.682
Household.Head.SexMale	2.166
Household.Head.Age	-2.882
House.Age	-1.871
Number.of.bedrooms	-1.863
Electricity	-1.886
log(Total.Food.Expenditure):Type.of.HouseholdSingle Family	2.691
log(Total.Food.Expenditure):Type.of.HouseholdTwo or More Nonrelated Persons/Members	1.268
log(Total.Household.Income):Household.Head.SexMale	-1.826
	Pr(>|z|)
(Intercept)	6.67e-05
log(Total.Food.Expenditure)	< 2e-16
Type.of.HouseholdSingle Family	0.001396
Type.of.HouseholdTwo or More Nonrelated Persons/Members	0.165756
log(Total.Household.Income)	0.000232
Household.Head.SexMale	0.030297
Household.Head.Age	0.003947
House.Age	0.061402
Number.of.bedrooms	0.062405
Electricity	0.059298
log(Total.Food.Expenditure):Type.of.HouseholdSingle Family	0.007118

```
log(Total.Food.Expenditure):Type.of.HouseholdTwo or More Nonrelated Persons/Members 0.204948
log(Total.Household.Income):Household.Head.SexMale                                  0.067778

(Intercept)                                                                        ***
log(Total.Food.Expenditure)                                                        ***
Type.of.HouseholdSingle Family                                                     **
Type.of.HouseholdTwo or More Nonrelated Persons/Members
log(Total.Household.Income)                                                        ***
Household.Head.SexMale                                                             *
Household.Head.Age                                                                 **
House.Age                                                                          .
Number.of.bedrooms                                                                 .
Electricity                                                                        .
log(Total.Food.Expenditure):Type.of.HouseholdSingle Family                         **
log(Total.Food.Expenditure):Type.of.HouseholdTwo or More Nonrelated Persons/Members
log(Total.Household.Income):Household.Head.SexMale                                 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for Negative Binomial(112480) family taken to be 1)

    Null deviance: 1373.6  on 1248  degrees of freedom
Residual deviance:  745.1  on 1236  degrees of freedom
AIC: 4802

Number of Fisher Scoring iterations: 1

              Theta:  112480
          Std. Err.:  457538
Warning while fitting theta: iteration limit reached

 2 x log-likelihood:  -4774.004
```

```
res.sq <- residuals(model_1, type = "response")^2
set1 <- data.frame(res.sq, mu.hat = model_1$fitted.values)
fit.lin <- lm(formula = res.sq ~ mu.hat, data = set1)
fit.quad <- lm(formula = res.sq ~ mu.hat + I(mu.hat^2), data = set1)

summary(fit.quad)
```

```
Call:
lm(formula = res.sq ~ mu.hat + I(mu.hat^2), data = set1)

Residuals:
   Min     1Q Median     3Q    Max
-7.914 -2.276 -1.188  0.376 76.829

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.71206    0.99574   0.715   0.4747
mu.hat       0.06518    0.43168   0.151   0.8800
I(mu.hat^2)  0.08297    0.04469   1.857   0.0636 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Residual standard error: 5.085 on 1246 degrees of freedom
Multiple R-squared:  0.05956,   Adjusted R-squared:  0.05805
F-statistic: 39.46 on 2 and 1246 DF,  p-value: < 2.2e-16
```

```
plot(set1$mu.hat, y = set1$res.sq, xlab = "Predicted count",
ylab = "Squared Residual")
curve(expr = predict(fit.lin, newdata = data.frame(mu.hat = x), type = "response"),
col = "blue", add = TRUE, lty = "solid")
curve(expr = predict(fit.quad, newdata = data.frame(mu.hat = x), type = "response"),
col = "red", add = TRUE, lty = "dashed")
legend("topleft", legend = c("Linear", "Quadratic"), col = c( "blue", "red"),
lty = c("solid", "dashed"), bty = "n")
```