Data Analysis for Motor insurance Data
"GLM model using R software"

Ahmed Sabeel
sabeel2010@gmail.com

June 20, 2020

```
## Warning:  package 'xtable' was built under R version 3.6.3
## Warning:  package 'car' was built under R version 3.6.3
## Loading required package:  carData
## Warning:  package 'carData' was built under R version 3.6.1
##
## Please cite as:
##  Hlavac, Marek (2018).  stargazer:  Well-Formatted Regression and Summary
Statistics Tables.
##  R package version 5.2.2.  https://CRAN.R-project.org/package=stargazer
## Warning:  package 'ggplot2' was built under R version 3.6.3
## Installing package into 'C:/Users/USER/Documents/R/win-library/3.6'
## (as 'lib' is unspecified)
## Error in contrib.url(repos, "source"):  trying to use CRAN without setting
a mirror
## Warning:  package 'writexl' was built under R version 3.6.3
```

Descritive Statistics

Read the data from the motor insurance data into R file from Excel and SPSS data files

```
## re-encoding from CP1252

##                          25%
## Mean                  1.4009
## Standard Deviation    0.7424
## Minimum               1.0000
## First Quantile        1.0000
## Median                1.0000
## Third Quantile        2.0000
```

```
## 90% Quantile       2.0000
## 95% Quantile       3.0000
## 99% Quantile       4.0000
## Maximum            8.0000
##                         25%
## Mean                10682.13
## Standard Deviation  24491.88
## Minimum                50.00
## First Quantile       1215.00
## Median               4690.25
## Third Quantile      10717.00
## 90% Quantile        22672.40
## 95% Quantile        36283.20
## 99% Quantile       105912.68
## Maximum            450000.00
```

Frequency model of vehicle data

Performing GLM Calcs Using Poisson, negative Binomail and Quasi Poison Distributions

```
##
## Call:
## glm(formula = q11 ~ q1 + q2 + q3 + q4 + q5 + q6 + q7 + q8 + q9 +
##     q10 + offset(log(q13)), family = poisson(link = log), data = df)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -0.6845  -0.3752  -0.3120   0.4044   3.4725
##
## Coefficients: (1 not defined because of singularities)
##                              Estimate Std. Error z value Pr(>|z|)
## (Intercept)                 0.2806970  0.1008713   2.783  0.00539 **
## q1Female                    0.0030637  0.0340926   0.090  0.92840
## q225-30                     0.0901379  0.0728279   1.238  0.21583
## q231-60                     0.0285726  0.0708009   0.404  0.68653
## q260 and above              0.0617605  0.0784817   0.787  0.43132
## q3Engineer and Programmer -0.0398702  0.0361986  -1.101  0.27071
## q3Medical Professional    -0.0751102  0.0530076  -1.417  0.15649
## q3Business man/woman       0.0019969  0.0641862   0.031  0.97518
## q3Student                  0.0448961  0.0749669   0.599  0.54925
## q4Khartoum North           0.0057222  0.0308848   0.185  0.85301
```

```
## q4Om durman               -0.0191246  0.0305997  -0.625  0.53198
## q53-5                      0.0263977  0.0425686   0.620  0.53518
## q56-10                     0.0090218  0.0453691   0.199  0.84238
## q510 and above            -0.0051855  0.0440196  -0.118  0.90623
## q6Sudan                   -0.0213439  0.0497578  -0.429  0.66795
## q6Japan                   -0.0366583  0.0956355  -0.383  0.70149
## q6Germany                 -0.0086957  0.1788220  -0.049  0.96122
## q6Czech                    0.1294084  0.1713107   0.755  0.45001
## q7Toyota                   0.1170225  0.1213524   0.964  0.33489
## q7Giad                            NA         NA      NA       NA
## q7Kia                      0.0137680  0.0548314   0.251  0.80174
## q7Skoda                   -0.2045909  0.2249116  -0.910  0.36301
## q7Mitsubitishi             0.0305572  0.1655473   0.185  0.85356
## q7Merceds                  0.0863028  0.1349550   0.639  0.52250
## q8Tuson                    0.0187434  0.1112942   0.168  0.86626
## q8Visto                    0.0147937  0.0864250   0.171  0.86409
## q8Click                    0.0089661  0.0532089   0.169  0.86618
## q8Fabia                    0.0905589  0.1398095   0.648  0.51716
## q8Lancer                  -0.0154165  0.1653561  -0.093  0.92572
## q8Corolla                 -0.0281341  0.0890704  -0.316  0.75211
## q8Merceds                 -0.0670339  0.1404492  -0.477  0.63316
## q8Hilux                   -0.0112728  0.1321850  -0.085  0.93204
## q8Land Cruiser             0.0633796  0.1341903   0.472  0.63670
## q96-10                    -0.0438863  0.0320845  -1.368  0.17136
## q911-20                   -0.0840520  0.0367245  -2.289  0.02210 *
## q921 and above            -0.0872137  0.0996980  -0.875  0.38169
## q10Medium                  0.0571968  0.0681620   0.839  0.40140
## q10Large                  -0.0009503  0.1290537  -0.007  0.99412
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for poisson family taken to be 1)
##
##     Null deviance: 1415.6  on 4522  degrees of freedom
## Residual deviance: 1389.0  on 4486  degrees of freedom
##   (206 observations deleted due to missingness)
## AIC: 11496
##
## Number of Fisher Scoring iterations: 4
## Analysis of Deviance Table
##
## Model: poisson, link: log
```

```
##
## Response: q11
##
## Terms added sequentially (first to last)
##
##
##      Df Deviance Resid. Df Resid. Dev Pr(>Chi)
## NULL                  4522     1415.6
## q1    1   0.0101      4521     1415.6   0.9201
## q2    3   4.0502      4518     1411.5   0.2561
## q3    4   3.8108      4514     1407.7   0.4322
## q4    2   0.9116      4512     1406.8   0.6339
## q5    3   0.7351      4509     1406.1   0.8649
## q6    4   3.7842      4505     1402.3   0.4360
## q7    5   3.5158      4500     1398.8   0.6210
## q8    9   3.7105      4491     1395.1   0.9294
## q9    3   5.0339      4488     1390.0   0.1693
## q10   2   1.0016      4486     1389.0   0.6061

## Warning in theta.ml(Y, mu, sum(w), w, limit = control$maxit, trace = control$trace
> :  iteration limit reached
## Warning in theta.ml(Y, mu, sum(w), w, limit = control$maxit, trace = control$trace
> :  iteration limit reached

##
## Call:
## glm.nb(formula = q11 ~ q1 + q2 + q3 + q4 + q5 + q6 + q7 + q8 +
##     q9 + q10, data = df, init.theta = 64015.20452, link = log)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -0.6845  -0.3752  -0.3120   0.4044   3.4725
##
## Coefficients: (1 not defined because of singularities)
##                              Estimate Std. Error z value Pr(>|z|)
## (Intercept)                 0.2806970  0.1008731   2.783  0.00539 **
## q1Female                    0.0030638  0.0340932   0.090  0.92840
## q225-30                     0.0901379  0.0728295   1.238  0.21584
## q231-60                     0.0285726  0.0708023   0.404  0.68654
## q260 and above              0.0617604  0.0784833   0.787  0.43133
## q3Engineer and Programmer -0.0398702  0.0361992  -1.101  0.27072
## q3Medical Professional    -0.0751103  0.0530084  -1.417  0.15650
## q3Business man/woman       0.0019969  0.0641875   0.031  0.97518
```

```
## q3Student                   0.0448958  0.0749690   0.599  0.54927
## q4Khartoum North            0.0057223  0.0308854   0.185  0.85301
## q4Om durman                -0.0191246  0.0306003  -0.625  0.53198
## q53-5                       0.0263978  0.0425695   0.620  0.53518
## q56-10                      0.0090218  0.0453700   0.199  0.84238
## q510 and above             -0.0051855  0.0440204  -0.118  0.90623
## q6Sudan                    -0.0213439  0.0497586  -0.429  0.66796
## q6Japan                    -0.0366582  0.0956380  -0.383  0.70150
## q6Germany                  -0.0086955  0.1788268  -0.049  0.96122
## q6Czech                     0.1294083  0.1713447   0.755  0.45010
## q7Toyota                    0.1170226  0.1213578   0.964  0.33491
## q7Giad                             NA         NA      NA       NA
## q7Kia                       0.0137679  0.0548323   0.251  0.80174
## q7Skoda                    -0.2045903  0.2249390  -0.910  0.36307
## q7Mitsubitishi              0.0305573  0.1655521   0.185  0.85356
## q7Merceds                   0.0863028  0.1349600   0.639  0.52252
## q8Tuson                     0.0187435  0.1112962   0.168  0.86626
## q8Visto                     0.0147937  0.0864265   0.171  0.86409
## q8Click                     0.0089661  0.0532097   0.169  0.86619
## q8Fabia                     0.0905585  0.1398119   0.648  0.51717
## q8Lancer                   -0.0154166  0.1653593  -0.093  0.92572
## q8Corolla                  -0.0281342  0.0890731  -0.316  0.75211
## q8Merceds                  -0.0670340  0.1404514  -0.477  0.63317
## q8Hilux                    -0.0112729  0.1321877  -0.085  0.93204
## q8Land Cruiser              0.0633795  0.1341932   0.472  0.63671
## q96-10                     -0.0438863  0.0320851  -1.368  0.17137
## q911-20                    -0.0840521  0.0367252  -2.289  0.02210 *
## q921 and above             -0.0872138  0.0997003  -0.875  0.38170
## q10Medium                   0.0571968  0.0681631   0.839  0.40140
## q10Large                   -0.0009504  0.1290560  -0.007  0.99412
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for Negative Binomial(64015.2) family taken to be 1)
##
##     Null deviance: 1415.6  on 4522  degrees of freedom
## Residual deviance: 1389.0  on 4486  degrees of freedom
##   (206 observations deleted due to missingness)
## AIC: 11498
##
## Number of Fisher Scoring iterations: 1
##
```

```
##
##               Theta:  64015
##           Std. Err.:  141902
## Warning while fitting theta: iteration limit reached
##
##  2 x log-likelihood:  -11421.69

## Warning in anova.negbin(model1, test = "Chisq"):  tests made without re-estimating
'theta'

## Analysis of Deviance Table
##
## Model: Negative Binomial(64015.2), link: log
##
## Response: q11
##
## Terms added sequentially (first to last)
##
##
##      Df Deviance Resid. Df Resid. Dev Pr(>Chi)
## NULL                  4522     1415.5
## q1    1   0.0101      4521     1415.5   0.9201
## q2    3   4.0501      4518     1411.5   0.2561
## q3    4   3.8107      4514     1407.7   0.4322
## q4    2   0.9116      4512     1406.8   0.6340
## q5    3   0.7350      4509     1406.0   0.8649
## q6    4   3.7842      4505     1402.2   0.4360
## q7    5   3.5157      4500     1398.7   0.6210
## q8    9   3.7105      4491     1395.0   0.9294
## q9    3   5.0338      4488     1390.0   0.1693
## q10   2   1.0015      4486     1389.0   0.6061
##             summary.model1..coef.summary.model1..coef...4.....0.1..4.
## (Intercept)                                                 0.005391278
## q911-20                                                     0.022098019
##
## Call:
## glm(formula = q11 ~ q1 + q2 + q3 + q4 + q5 + q6 + q7 + q8 + q9 +
##     q10, family = "quasipoisson", data = df)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -0.6845  -0.3752  -0.3120   0.4044   3.4725
##
```

```
## Coefficients: (1 not defined because of singularities)
##                              Estimate Std. Error t value Pr(>|t|)
## (Intercept)                 0.2806970  0.0629498   4.459 8.43e-06 ***
## q1Female                    0.0030637  0.0212758   0.144 0.885508
## q225-30                     0.0901379  0.0454490   1.983 0.047398 *
## q231-60                     0.0285726  0.0441840   0.647 0.517877
## q260 and above              0.0617605  0.0489773   1.261 0.207374
## q3Engineer and Programmer -0.0398702  0.0225901  -1.765 0.077642 .
## q3Medical Professional    -0.0751102  0.0330800  -2.271 0.023221 *
## q3Business man/woman        0.0019969  0.0400561   0.050 0.960243
## q3Student                   0.0448961  0.0467839   0.960 0.337283
## q4Khartoum North            0.0057222  0.0192740   0.297 0.766565
## q4Om durman                -0.0191246  0.0190961  -1.001 0.316643
## q53-5                       0.0263977  0.0265654   0.994 0.320428
## q56-10                      0.0090218  0.0283131   0.319 0.750012
## q510 and above             -0.0051855  0.0274709  -0.189 0.850286
## q6Sudan                    -0.0213439  0.0310519  -0.687 0.491889
## q6Japan                    -0.0366583  0.0596823  -0.614 0.539098
## q6Germany                  -0.0086957  0.1115957  -0.078 0.937894
## q6Czech                     0.1294084  0.1069082   1.210 0.226165
## q7Toyota                    0.1170225  0.0757313   1.545 0.122360
## q7Giad                             NA         NA      NA       NA
## q7Kia                       0.0137680  0.0342181   0.402 0.687439
## q7Skoda                    -0.2045909  0.1403584  -1.458 0.145012
## q7Mitsubitishi              0.0305572  0.1033115   0.296 0.767414
## q7Merceds                   0.0863028  0.0842201   1.025 0.305546
## q8Tuson                     0.0187434  0.0694543   0.270 0.787275
## q8Visto                     0.0147937  0.0539344   0.274 0.783875
## q8Click                     0.0089661  0.0332055   0.270 0.787159
## q8Fabia                     0.0905589  0.0872496   1.038 0.299359
## q8Lancer                   -0.0154165  0.1031922  -0.149 0.881248
## q8Corolla                  -0.0281341  0.0555853  -0.506 0.612782
## q8Merceds                  -0.0670339  0.0876488  -0.765 0.444430
## q8Hilux                    -0.0112728  0.0824914  -0.137 0.891310
## q8Land Cruiser              0.0633796  0.0837429   0.757 0.449188
## q96-10                     -0.0438863  0.0200227  -2.192 0.028443 *
## q911-20                    -0.0840520  0.0229183  -3.667 0.000248 ***
## q921 and above             -0.0872137  0.0622176  -1.402 0.161058
## q10Medium                   0.0571968  0.0425372   1.345 0.178813
## q10Large                   -0.0009503  0.0805373  -0.012 0.990586
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
## (Dispersion parameter for quasipoisson family taken to be 0.3894515)
##
##     Null deviance: 1415.6  on 4522  degrees of freedom
## Residual deviance: 1389.0  on 4486  degrees of freedom
##   (206 observations deleted due to missingness)
## AIC: NA
##
## Number of Fisher Scoring iterations: 4
## Analysis of Deviance Table
##
## Model: quasipoisson, link: log
##
## Response: q11
##
## Terms added sequentially (first to last)
##
##
##      Df Deviance Resid. Df Resid. Dev Pr(>Chi)
## NULL                  4522     1415.6
## q1    1   0.0101      4521     1415.6  0.87233
## q2    3   4.0502      4518     1411.5  0.01546 *
## q3    4   3.8108      4514     1407.7  0.04421 *
## q4    2   0.9116      4512     1406.8  0.31026
## q5    3   0.7351      4509     1406.1  0.59610
## q6    4   3.7842      4505     1402.3  0.04548 *
## q7    5   3.5158      4500     1398.8  0.10797
## q8    9   3.7105      4491     1395.1  0.39006
## q9    3   5.0339      4488     1390.0  0.00480 **
## q10   2   1.0016      4486     1389.0  0.27641
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Combine results of the three Distributions

```
##
## ================================================================================
##                                          Dependent variable:
##                              ---------------------------------------------------
##                                                    q11
##                               Poisson         negative        glm: quasipoisson
```

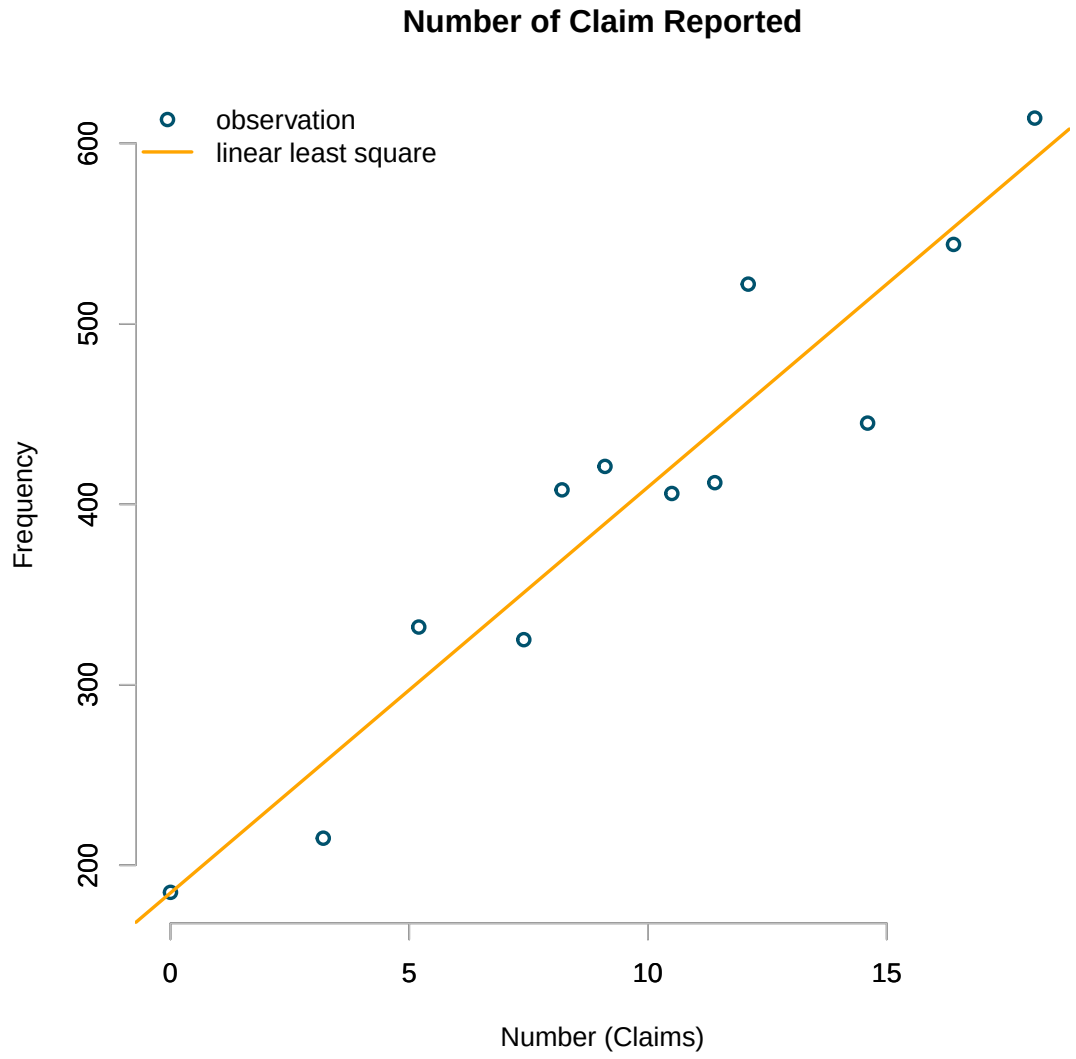
```
##                                          binomial          link = log
##                              (1)             (2)                (3)
## ------------------------------------------------------------------------
## q1Female                    0.003           0.003              0.003
##                            (0.03)          (0.03)             (0.02)
## q225-30                     0.09            0.09               0.09*
##                            (0.07)          (0.07)             (0.05)
## q231-60                     0.03            0.03               0.03
##                            (0.07)          (0.07)             (0.04)
## q260 and above              0.06            0.06               0.06
##                            (0.08)          (0.08)             (0.05)
## q3Engineer and Programmer  -0.04           -0.04              -0.04
##                            (0.04)          (0.04)             (0.02)
## q3Medical Professional     -0.08           -0.08              -0.08*
##                            (0.05)          (0.05)             (0.03)
## q3Business man/woman        0.002           0.002              0.002
##                            (0.06)          (0.06)             (0.04)
## q3Student                   0.04            0.04               0.04
##                            (0.07)          (0.07)             (0.05)
## q4Khartoum North            0.01            0.01               0.01
##                            (0.03)          (0.03)             (0.02)
## q4Om durman                -0.02           -0.02              -0.02
##                            (0.03)          (0.03)             (0.02)
## q53-5                       0.03            0.03               0.03
##                            (0.04)          (0.04)             (0.03)
## q56-10                      0.01            0.01               0.01
##                            (0.05)          (0.05)             (0.03)
## q510 and above             -0.01           -0.01              -0.01
##                            (0.04)          (0.04)             (0.03)
## q6Sudan                    -0.02           -0.02              -0.02
##                            (0.05)          (0.05)             (0.03)
## q6Japan                    -0.04           -0.04              -0.04
##                            (0.10)          (0.10)             (0.06)
## q6Germany                  -0.01           -0.01              -0.01
##                            (0.18)          (0.18)             (0.11)
## q6Czech                     0.13            0.13               0.13
##                            (0.17)          (0.17)             (0.11)
## q7Toyota                    0.12            0.12               0.12
##                            (0.12)          (0.12)             (0.08)
## q7Giad
##
## q7Kia                       0.01            0.01               0.01
```

```
##                              (0.05)           (0.05)              (0.03)
## q7Skoda                       -0.20            -0.20               -0.20
##                              (0.22)           (0.22)              (0.14)
## q7Mitsubitishi                0.03             0.03                0.03
##                              (0.17)           (0.17)              (0.10)
## q7Merceds                     0.09             0.09                0.09
##                              (0.13)           (0.13)              (0.08)
## q8Tuson                       0.02             0.02                0.02
##                              (0.11)           (0.11)              (0.07)
## q8Visto                       0.01             0.01                0.01
##                              (0.09)           (0.09)              (0.05)
## q8Click                       0.01             0.01                0.01
##                              (0.05)           (0.05)              (0.03)
## q8Fabia                       0.09             0.09                0.09
##                              (0.14)           (0.14)              (0.09)
## q8Lancer                      -0.02            -0.02               -0.02
##                              (0.17)           (0.17)              (0.10)
## q8Corolla                     -0.03            -0.03               -0.03
##                              (0.09)           (0.09)              (0.06)
## q8Merceds                     -0.07            -0.07               -0.07
##                              (0.14)           (0.14)              (0.09)
## q8Hilux                       -0.01            -0.01               -0.01
##                              (0.13)           (0.13)              (0.08)
## q8Land Cruiser                0.06             0.06                0.06
##                              (0.13)           (0.13)              (0.08)
## q96-10                        -0.04            -0.04               -0.04*
##                              (0.03)           (0.03)              (0.02)
## q911-20                      -0.08*           -0.08*             -0.08***
##                              (0.04)           (0.04)              (0.02)
## q921 and above                -0.09            -0.09               -0.09
##                              (0.10)           (0.10)              (0.06)
## q10Medium                     0.06             0.06                0.06
##                              (0.07)           (0.07)              (0.04)
## q10Large                     -0.001           -0.001              -0.001
##                              (0.13)           (0.13)              (0.08)
## Constant                     0.28**           0.28**             0.28***
##                              (0.10)           (0.10)              (0.06)
## ------------------------------------------------------------------------
## Observations                  4,523            4,523               4,523
## Log Likelihood             -5,710.82        -5,711.85
## theta                                 64,015.20 (141,902.20)
## Akaike Inf. Crit.          11,495.63        11,497.70
```

```
## ================================================================================
## Note:                                          *p<0.05; **p<0.01; ***p<0.001
```

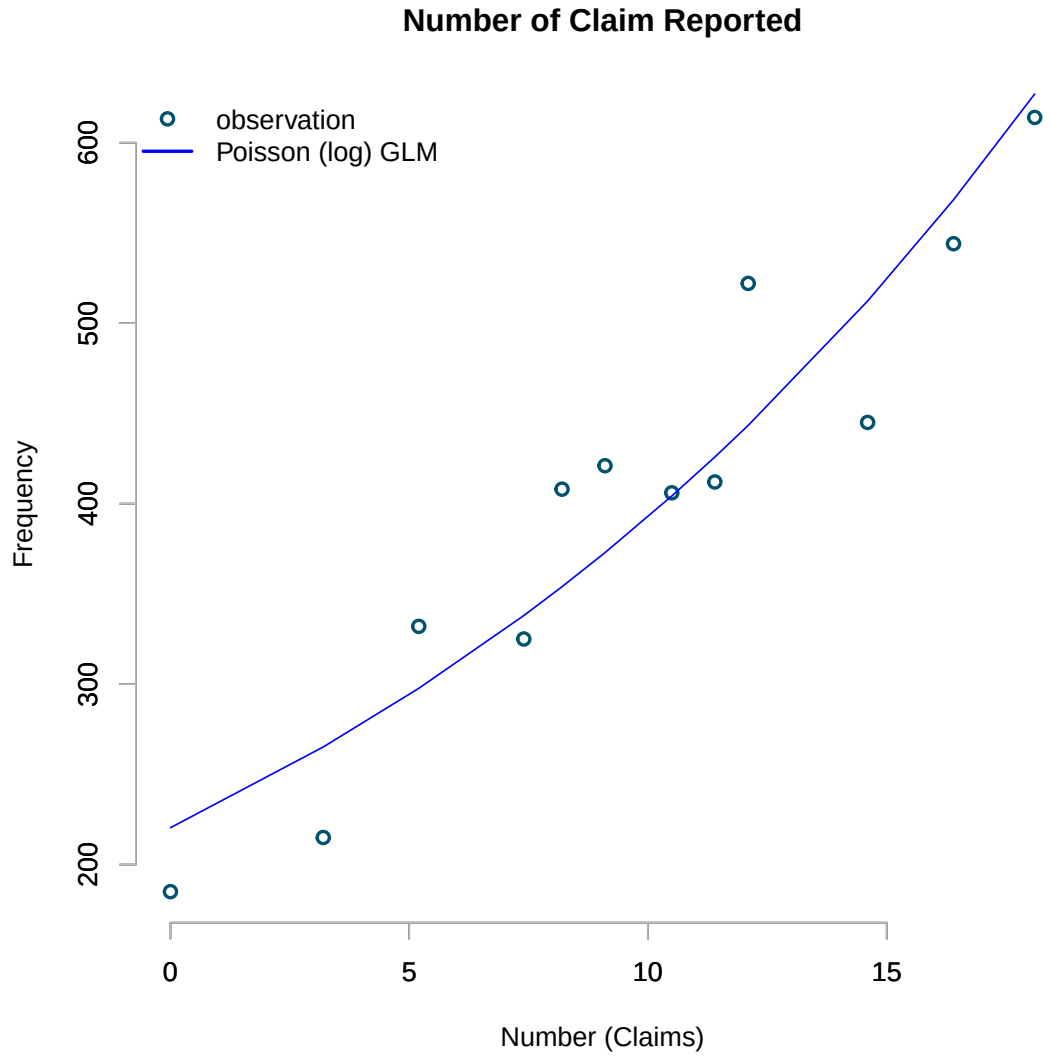
Plotting and comparing frequency models

```
##
## Call:
## lm(formula = frequency ~ claim, data = reported)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -67.970 -26.810  -4.563  30.720  65.243
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  184.682     25.039   7.376 2.38e-05 ***
## claim         22.486      2.287   9.832 1.86e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 40.48 on 10 degrees of freedom
## Multiple R-squared:  0.9063,Adjusted R-squared:  0.8969
## F-statistic: 96.67 on 1 and 10 DF,  p-value: 1.856e-06
```



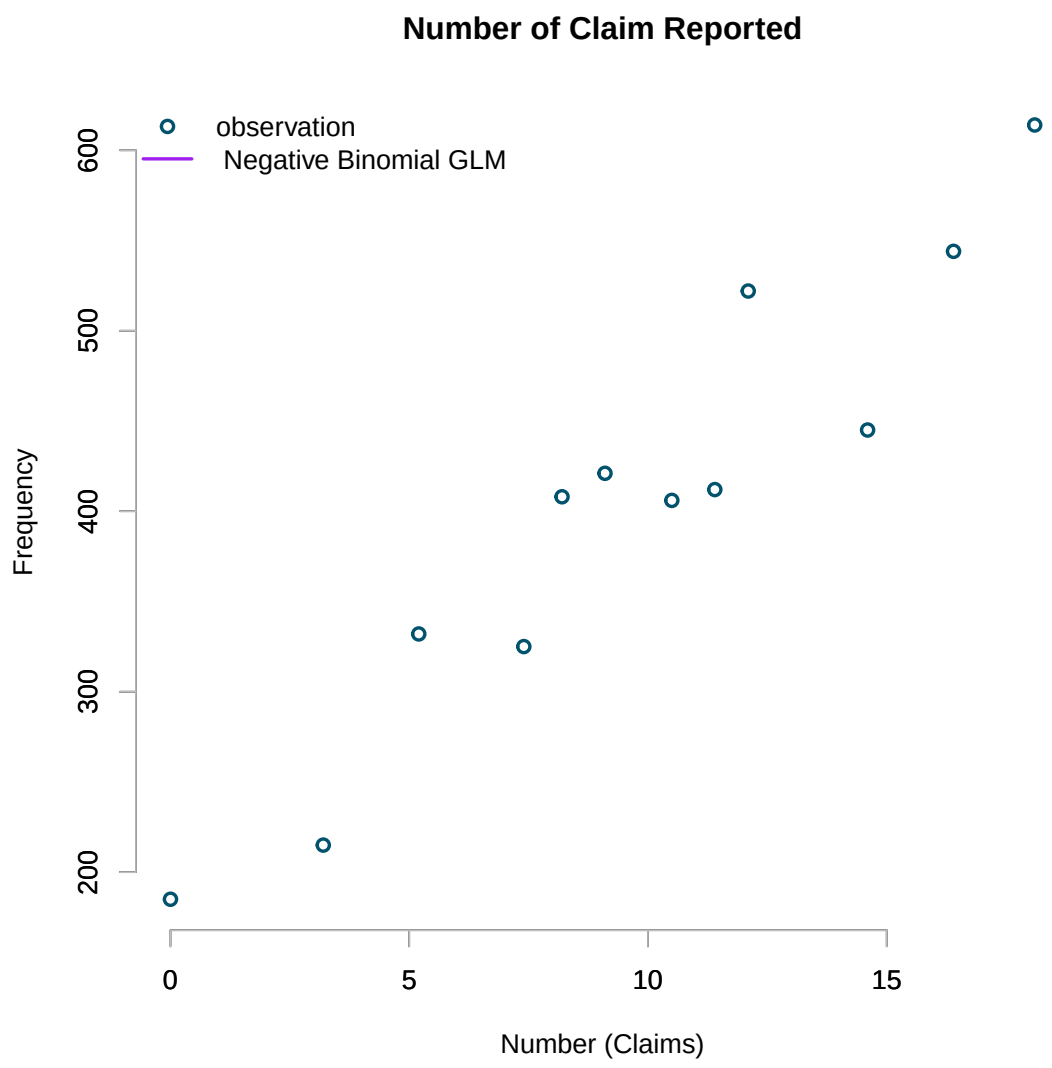

```
##
## Call:
## glm(formula = frequency ~ claim, family = poisson(link = "log"),
##     data = reported)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -3.1870  -1.3855  -0.5965   2.0769   3.6313
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) 5.395581   0.035643  151.38   <2e-16 ***
## claim       0.057754   0.002927   19.73   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## (Dispersion parameter for poisson family taken to be 1)
##
##     Null deviance: 460.137  on 11  degrees of freedom
## Residual deviance:  58.557  on 10  degrees of freedom
## AIC: 155.95
##
## Number of Fisher Scoring iterations: 4
```

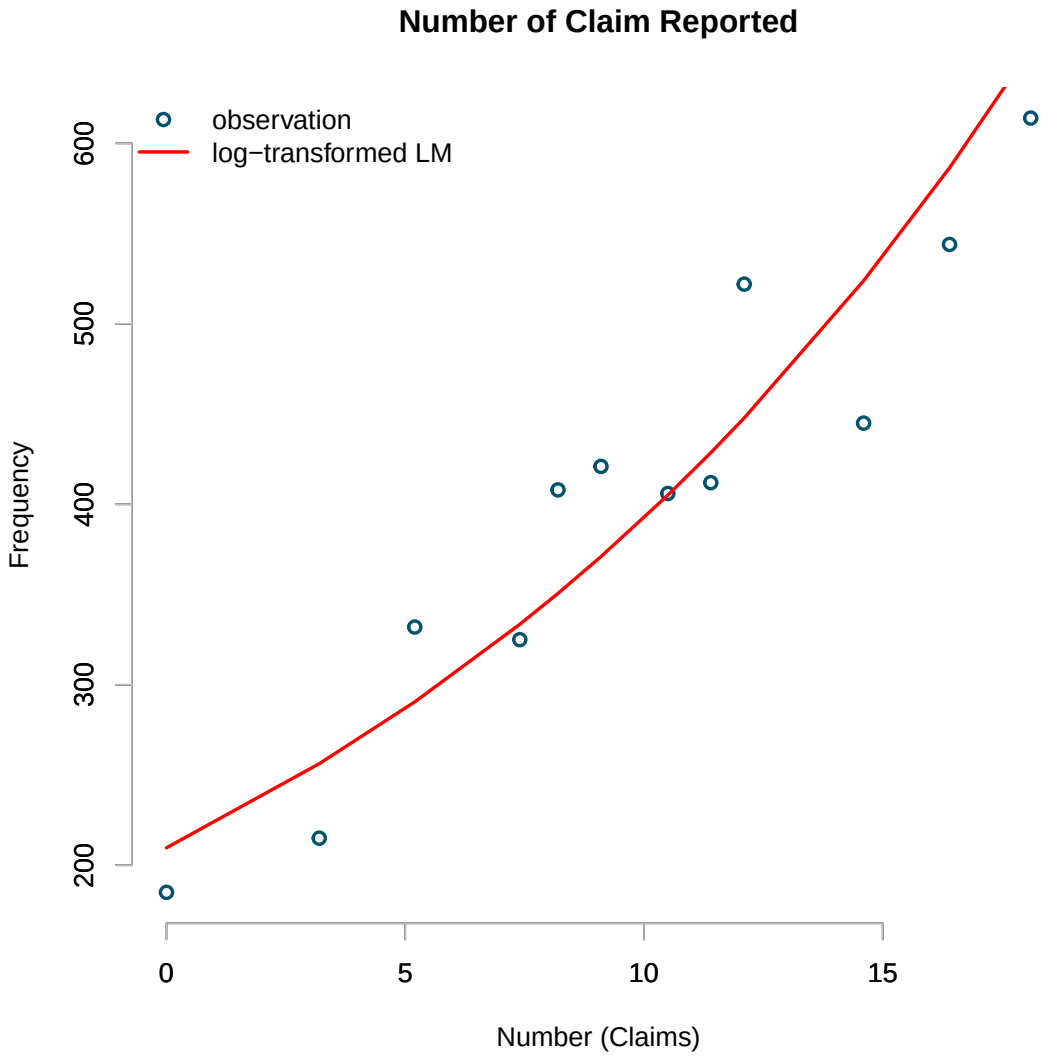


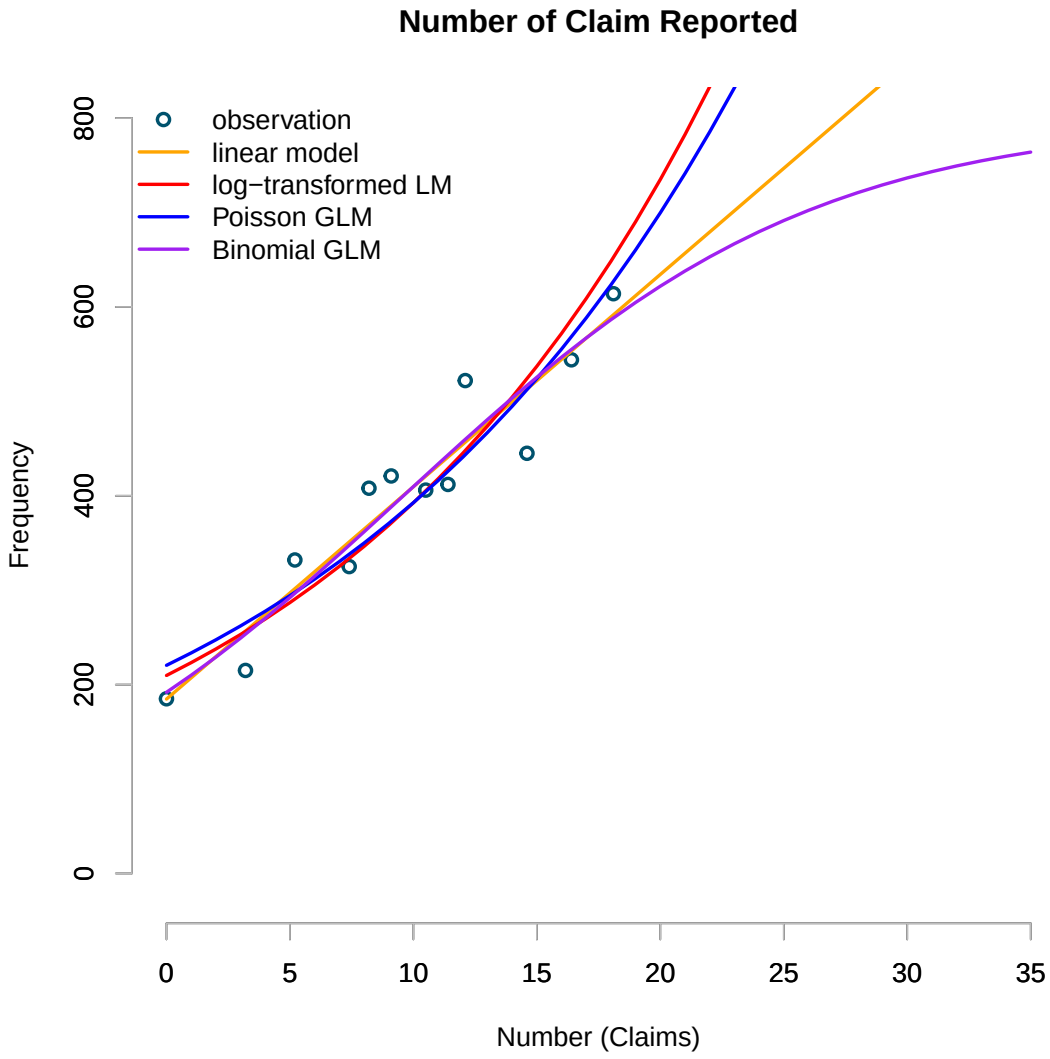

```
##
## Call:
## glm(formula = cbind(frequency, opportunity) ~ claim, family = binomial(link = "logit"),
##     data = reported)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -5.2276  -1.7614  -0.7076   2.3871   4.4963
```

```
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) -1.153631   0.047975  -24.05   <2e-16 ***
## claim        0.120225   0.004419   27.20   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 909.399  on 11  degrees of freedom
## Residual deviance:  91.051  on 10  degrees of freedom
## AIC: 179.46
##
## Number of Fisher Scoring iterations: 3
```



```
##
## Call:
## glm(formula = log(frequency) ~ claim, family = gaussian(link = "identity"),
##     data = reported)
##
## Deviance Residuals:
##      Min        1Q    Median        3Q       Max
## -0.16742  -0.07946  -0.02447   0.13633   0.16145
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 5.337369   0.079241  67.356 1.27e-14 ***
## claim       0.062715   0.007238   8.665 5.81e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 0.01640839)
##
##     Null deviance: 1.39610  on 11  degrees of freedom
## Residual deviance: 0.16408  on 10  degrees of freedom
## AIC: -11.453
##
## Number of Fisher Scoring iterations: 2
```


Number of Claim Reported
observation
log−transformed LM
Frequency
200
300
400
500
600
0
5
10
15
Number (Claims)


Number of Claim Reported
observation
linear model
log−transformed LM
Poisson GLM
Binomial GLM
Frequency
0
200
400
600
800
0
5
10
15
20
25
30
35
Number (Claims)

For Claim Severity Estimations

Perform serverity distribution for GLM using Gamma and log Normal Distributions

```
##
## Call:
## glm(formula = log(q12) ~ q1 + q2 + q3 + q4 + q5 + q6 + q7 + q8 +
##     q9 + q10, family = gaussian(link = "identity"), data = df)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -4.7029  -1.1945   0.0898   0.9341   4.4748
##
## Coefficients: (1 not defined because of singularities)
##                             Estimate Std. Error t value Pr(>|t|)
## (Intercept)                 8.439667   0.158067  53.393  < 2e-16 ***
## q1Female                    0.126314   0.053930   2.342 0.019215 *
## q225-30                     0.306095   0.115393   2.653 0.008015 **
## q231-60                     0.114365   0.110986   1.030 0.302855
## q260 and above              0.190465   0.123419   1.543 0.122845
## q3Engineer and Programmer -0.049495   0.056693  -0.873 0.382690
## q3Medical Professional    -0.189596   0.081677  -2.321 0.020316 *
## q3Business man/woman       0.081573   0.102815   0.793 0.427590
## q3Student                  0.059496   0.122708   0.485 0.627800
## q4Khartoum North          -0.116261   0.049127  -2.367 0.017997 *
## q4Om durman                0.002302   0.048188   0.048 0.961898
## q53-5                     -0.040415   0.067565  -0.598 0.549758
## q56-10                    -0.096075   0.071659  -1.341 0.180076
## q510 and above            -0.041682   0.069362  -0.601 0.547913
## q6Sudan                   -0.194637   0.077847  -2.500 0.012446 *
## q6Japan                    0.503190   0.148810   3.381 0.000727 ***
## q6Germany                  0.461292   0.290929   1.586 0.112904
## q6Czech                    0.078562   0.293251   0.268 0.788789
## q7Toyota                  -0.045487   0.193878  -0.235 0.814518
## q7Giad                           NA         NA      NA       NA
## q7Kia                      0.146211   0.086478   1.691 0.090955 .
## q7Skoda                    0.676600   0.374188   1.808 0.070645 .
## q7Mitsubitishi            -0.070780   0.260826  -0.271 0.786121
## q7Merceds                  0.217569   0.221605   0.982 0.326258
## q8Tuson                    0.356724   0.177524   2.009 0.044551 *
## q8Visto                   -0.308656   0.135133  -2.284 0.022412 *
## q8Click                    0.033088   0.083646   0.396 0.692441
## q8Fabia                   -0.308853   0.225777  -1.368 0.171395
```

```
## q8Lancer                   -0.130904   0.260034  -0.503 0.614699
## q8Corolla                  -0.118648   0.144797  -0.819 0.412595
## q8Merceds                   0.643989   0.219044   2.940 0.003299 **
## q8Hilux                    -0.057221   0.211901  -0.270 0.787144
## q8Land Cruiser              0.220413   0.216313   1.019 0.308281
## q96-10                     -0.439952   0.051091  -8.611  < 2e-16 ***
## q911-20                    -0.627682   0.057882 -10.844  < 2e-16 ***
## q921 and above             -0.977646   0.157173  -6.220 5.42e-10 ***
## q10Medium                   0.058611   0.106497   0.550 0.582105
## q10Large                   -0.326918   0.204788  -1.596 0.110477
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 1.784297)
##
##     Null deviance: 8571.0  on 4522  degrees of freedom
## Residual deviance: 8004.4  on 4486  degrees of freedom
##   (206 observations deleted due to missingness)
## AIC: 15493
##
## Number of Fisher Scoring iterations: 2
## Analysis of Deviance Table
##
## Model: gaussian, link: identity
##
## Response: log(q12)
##
## Terms added sequentially (first to last)
##
##
##      Df Deviance Resid. Df Resid. Dev  Pr(>Chi)
## NULL                  4522     8571.0
## q1    1    5.082      4521     8565.9  0.091484 .
## q2    3   25.927      4518     8540.0  0.002265 **
## q3    4   13.976      4514     8526.0  0.097896 .
## q4    2   13.075      4512     8512.9  0.025630 *
## q5    3    4.787      4509     8508.1  0.443139
## q6    4   88.258      4505     8419.9 4.673e-10 ***
## q7    5   11.333      4500     8408.6  0.273481
## q8    9  143.773      4491     8264.8 1.241e-13 ***
## q9    3  251.032      4488     8013.8 < 2.2e-16 ***
## q10   2    9.397      4486     8004.4  0.071842 .
```

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

## Warning in log(ifelse(y == 0, 1, y/mu)):  NaNs produced
## Warning:  step size truncated due to divergence
## Warning in log(ifelse(y == 0, 1, y/mu)):  NaNs produced
## Warning in log(ifelse(y == 0, 1, y/mu)):  NaNs produced
## Warning in log(ifelse(y == 0, 1, y/mu)):  NaNs produced
## Warning in log(ifelse(y == 0, 1, y/mu)):  NaNs produced
## Warning in log(ifelse(y == 0, 1, y/mu)):  NaNs produced
## Warning in log(ifelse(y == 0, 1, y/mu)):  NaNs produced
## Warning in log(ifelse(y == 0, 1, y/mu)):  NaNs produced
## Warning in log(ifelse(y == 0, 1, y/mu)):  NaNs produced
## Warning:  glm.fit:  algorithm did not converge

##
## Call:
## glm(formula = q12 ~ q1 + q2 + q3 + q4 + q5 + q6 + q7 + q8 + q9 +
##     q10, family = Gamma(link = "identity"), data = df, start = c(1,
##     1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
##     1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1))
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -3.0352  -1.5338  -0.6911   0.0697   7.1450
##
## Coefficients: (1 not defined because of singularities)
##                             Estimate Std. Error t value Pr(>|t|)
## (Intercept)                11090.13    1872.19   5.924 3.38e-09 ***
## q1Female                     510.81     697.69   0.732   0.4641
## q225-30                     2967.65    1383.59   2.145   0.0320 *
## q231-60                     1337.37    1212.71   1.103   0.2702
## q260 and above               674.68    1379.73   0.489   0.6249
## q3Engineer and Programmer   -544.84     701.97  -0.776   0.4377
## q3Medical Professional     -1183.63    1040.15  -1.138   0.2552
## q3Business man/woman         772.19    1677.51   0.460   0.6453
## q3Student                   -687.75    1599.32  -0.430   0.6672
## q4Khartoum North            -844.34     601.76  -1.403   0.1606
## q4Om durman                  -31.93     624.71  -0.051   0.9592
## q53-5                       -949.94     885.68  -1.073   0.2835
## q56-10                      -933.54     945.15  -0.988   0.3233
## q510 and above              -554.76     929.24  -0.597   0.5505
## q6Sudan                    -1683.95     797.35  -2.112   0.0347 *
```

```
## q6Japan                       7524.72    2758.31   2.728   0.0064 **
## q6Germany                    11335.57    6695.18   1.693   0.0905 .
## q6Czech                        579.69    3253.09   0.178   0.8586
## q7Toyota                     -2862.86    3041.44  -0.941   0.3466
## q7Giad                             NA         NA      NA       NA
## q7Kia                          386.40    1078.14   0.358   0.7201
## q7Skoda                      12191.86    7500.22   1.626   0.1041
## q7Mitsubitishi                -787.27    4983.77  -0.158   0.8745
## q7Merceds                     -152.73    2397.39  -0.064   0.9492
## q8Tuson                       3264.70    3392.63   0.962   0.3360
## q8Visto                       -306.46    1449.23  -0.211   0.8325
## q8Click                       1004.25     991.89   1.012   0.3114
## q8Fabia                      -9513.54    6494.56  -1.465   0.1430
## q8Lancer                     -5207.12    4659.29  -1.118   0.2638
## q8Corolla                    -1760.07    1942.13  -0.906   0.3648
## q8Merceds                    10874.64    6983.16   1.557   0.1195
## q8Hilux                       -824.61    3483.59  -0.237   0.8129
## q8Land Cruiser                3462.48    3932.01   0.881   0.3786
## q96-10                       -5005.27     843.97  -5.931 3.24e-09 ***
## q911-20                      -6449.39     854.66  -7.546 5.40e-14 ***
## q921 and above              -10009.61    1476.37  -6.780 1.36e-11 ***
## q10Medium                     1160.14    1215.91   0.954   0.3401
## q10Large                       142.10    3477.87   0.041   0.9674
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for Gamma family taken to be 4.017963)
##
##     Null deviance: 8698.7  on 4522  degrees of freedom
## Residual deviance: 7981.7  on 4486  degrees of freedom
##   (206 observations deleted due to missingness)
## AIC: 91920
##
## Number of Fisher Scoring iterations: 25
```

Combine results of the three Distributions

```
##
## ================================================
##                              Dependent variable:
##                          ---------------------------
```

```
##                               log(q12)          q12
##                                normal       glm: Gamma
##                                                link = identity
##                                 (1)             (2)
## ------------------------------------------------------------
## q1Female                       0.13*          510.81
##                               (0.05)         (697.69)
## q225-30                        0.31**        2,967.65*
##                               (0.12)        (1,383.59)
## q231-60                        0.11          1,337.37
##                               (0.11)        (1,212.71)
## q260 and above                 0.19           674.68
##                               (0.12)        (1,379.73)
## q3Engineer and Programmer     -0.05           -544.84
##                               (0.06)         (701.97)
## q3Medical Professional        -0.19*         -1,183.63
##                               (0.08)        (1,040.15)
## q3Business man/woman           0.08           772.19
##                               (0.10)        (1,677.51)
## q3Student                      0.06           -687.75
##                               (0.12)        (1,599.32)
## q4Khartoum North              -0.12*          -844.34
##                               (0.05)         (601.76)
## q4Om durman                   0.002           -31.93
##                               (0.05)         (624.71)
## q53-5                         -0.04           -949.94
##                               (0.07)         (885.68)
## q56-10                        -0.10           -933.54
##                               (0.07)         (945.15)
## q510 and above                -0.04           -554.76
##                               (0.07)         (929.24)
## q6Sudan                       -0.19*        -1,683.95*
##                               (0.08)         (797.35)
## q6Japan                       0.50***       7,524.72**
##                               (0.15)        (2,758.31)
## q6Germany                      0.46          11,335.57
##                               (0.29)        (6,695.18)
## q6Czech                        0.08           579.69
##                               (0.29)        (3,253.09)
## q7Toyota                      -0.05           -2,862.86
##                               (0.19)        (3,041.44)
## q7Giad
```

```
##
## q7Kia                          0.15          386.40
##                               (0.09)       (1,078.14)
## q7Skoda                        0.68         12,191.86
##                               (0.37)       (7,500.22)
## q7Mitsubitishi                -0.07          -787.27
##                               (0.26)       (4,983.77)
## q7Merceds                      0.22          -152.73
##                               (0.22)       (2,397.39)
## q8Tuson                       0.36*          3,264.70
##                               (0.18)       (3,392.63)
## q8Visto                       -0.31*          -306.46
##                               (0.14)       (1,449.23)
## q8Click                        0.03          1,004.25
##                               (0.08)        (991.89)
## q8Fabia                       -0.31          -9,513.54
##                               (0.23)       (6,494.56)
## q8Lancer                      -0.13          -5,207.12
##                               (0.26)       (4,659.29)
## q8Corolla                     -0.12          -1,760.07
##                               (0.14)       (1,942.13)
## q8Merceds                     0.64**         10,874.64
##                               (0.22)       (6,983.16)
## q8Hilux                       -0.06          -824.61
##                               (0.21)       (3,483.59)
## q8Land Cruiser                 0.22          3,462.48
##                               (0.22)       (3,932.01)
## q96-10                      -0.44***      -5,005.27***
##                               (0.05)        (843.97)
## q911-20                     -0.63***      -6,449.39***
##                               (0.06)        (854.66)
## q921 and above              -0.98***      -10,009.61***
##                               (0.16)       (1,476.37)
## q10Medium                      0.06          1,160.14
##                               (0.11)       (1,215.91)
## q10Large                      -0.33           142.10
##                               (0.20)       (3,477.87)
## Constant                     8.44***       11,090.13***
##                               (0.16)       (1,872.19)
## ------------------------------------------------------
## Observations                   4,523           4,523
## Log Likelihood               -7,709.75      -45,923.19
```

```
## Akaike Inf. Crit.           15,493.49       91,920.38
## ================================================================
## Note:                       *p<0.05; **p<0.01; ***p<0.001
```