Stats*3510 - Assignment 3
Graham Eckel – 0679576

Q2
a)
p.model=glm(ncrash~lexpos+DE+adol, data=dfDriver, family=poisson)
summary(p.model)

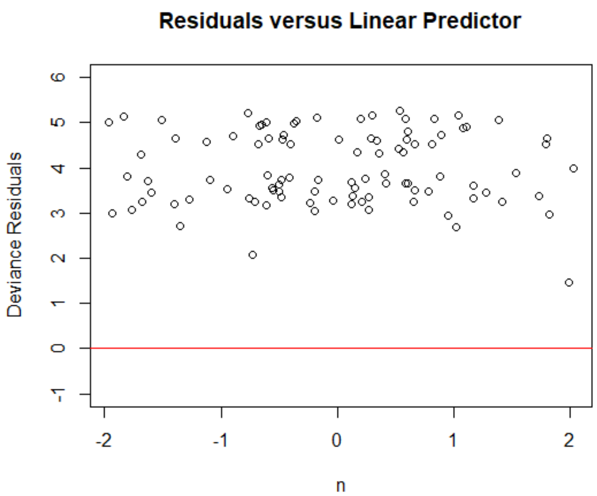
```
Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept) 1.637724   0.026730   61.27   <2e-16 ***
DE          0.001442   0.023984    0.06    0.952
adol        1.371114   0.027797   49.33   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for poisson family taken to be 1)

    Null deviance: 2979.021  on 96  degrees of freedom
Residual deviance:   95.191  on 94  degrees of freedom
AIC: 665.38

Number of Fisher Scoring iterations: 4
```

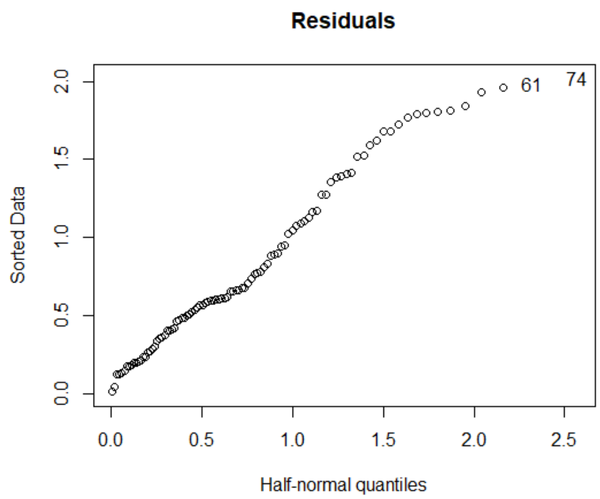
b)
We adjust the number of crashes/near crashes by the number of miles driven because crashes are dependent on the number of miles driven.

c)
The estimated intercept of the log mean number of crashes is 1.6377 per 1000 miles when the driver is not an adolescent and has not taken drivers education.
The estimated intercept of the mean number of crashes is e^1.6377 = 5.1433 per 1000 miles when the driver is not an adolescent and has not taken drivers education

d)
The estimated coefficient for drivers education of the log mean number of crashes is 0.001442 per 1000 miles for adolescents. That is adolescents who take drivers education add 0.001442 to their log mean number of crashes. Adolescents multiply their mean number of crashes per 1000 miles by 1.001443 if they take drivers education.

e)





There is not a mean of 0, but there doesn't appear to be many outliers or pattern. In the Residuals plot, we can see that they are not normally distributed.

f)
yhat = e^(1.637724+1.371114) = 20.26384

The estimated number of expected crashes for an adolescent who did not take drivers education is 20.26384 per 1000 miles.

g)
The estimated number of expected crashes for an adolescent who did not take drivers education per 7000 miles is 20.26384*7 = 141.8469.