cincinnati models

```
library(DT)
library(data.table)
```

```
## Warning: package 'data.table' was built under R version 3.5.2
```

```
library(magrittr)
library(sf)
```

```
## Warning: package 'sf' was built under R version 3.5.2
```

```
## Linking to GEOS 3.7.2, GDAL 2.4.2, PROJ 5.2.0
```

```
library(pscl)
```

```
## Classes and Methods for R developed in the
## Political Science Computational Laboratory
## Department of Political Science
## Stanford University
## Simon Jackman
## hurdle and zeroinfl functions by Achim Zeileis
```

```
library(countreg)
```

```
## Loading required package: MASS
```

```
##
## Attaching package: 'countreg'
```

```
## The following objects are masked from 'package:pscl':
##
##     hurdle, hurdle.control, hurdletest, zeroinfl, zeroinfl.control
```

```
library(GISTools)
```

```
## Loading required package: maptools
```

```
## Warning: package 'maptools' was built under R version 3.5.2
```

```
## Loading required package: sp
```

```
## Checking rgeos availability: TRUE
```

```
## Loading required package: RColorBrewer
```

```
## Loading required package: rgeos
```

```
## Warning: package 'rgeos' was built under R version 3.5.2
```

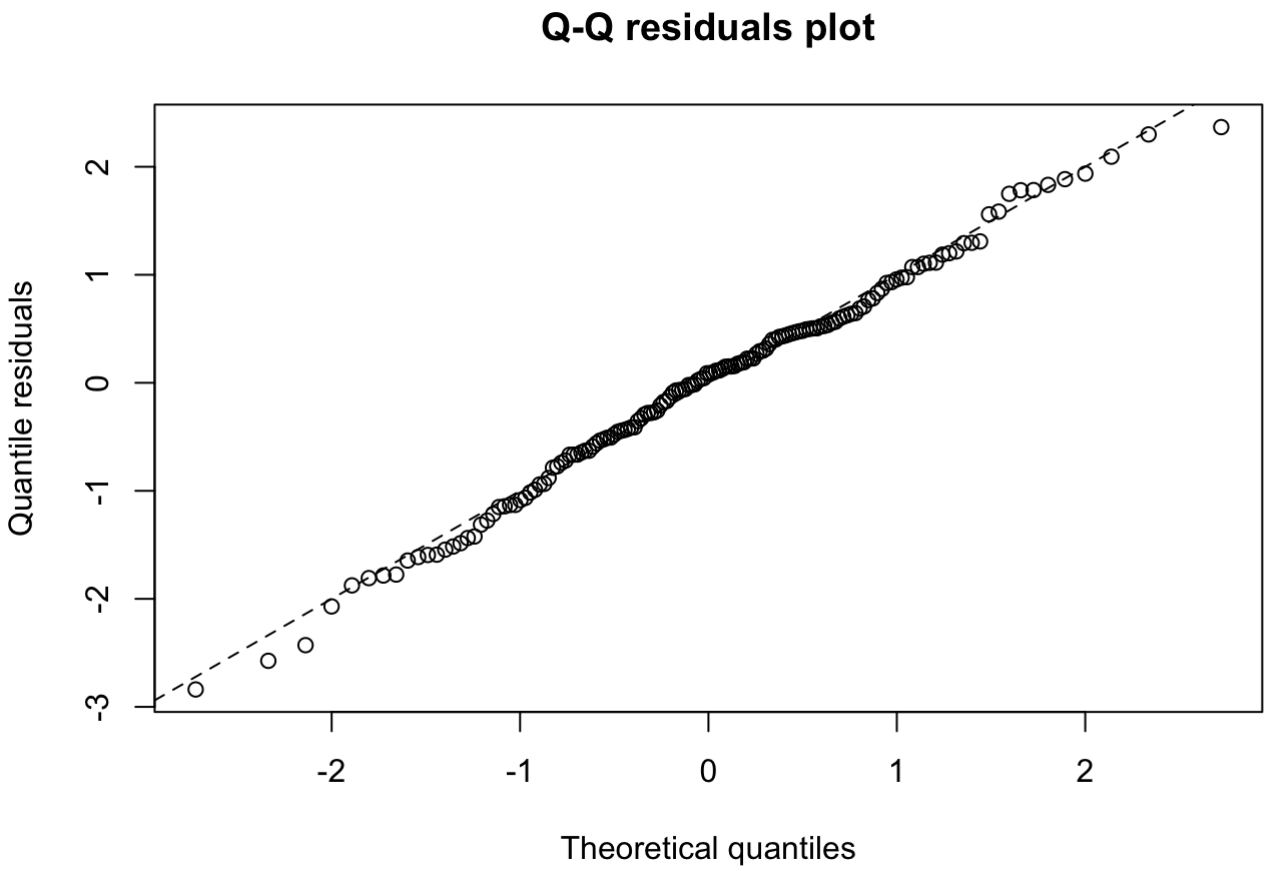
```
## rgeos version: 0.5-2, (SVN revision 621)
##  GEOS runtime version: 3.7.2-CAPI-1.11.2
##  Linking to sp version: 1.3-1
##  Polygon checking: TRUE
```

```
library(gtfsr)
library(sp)
library(stringi)
```

```
chicago_final<-fread("/Users/11kolop/Desktop/cincinnati_final.csv")[,-c(1:2)]
dat.hom.chicago<-chicago_final[chicago_final$ofns_desc=="criminal homicide",]
dat.hom.chicago$transp<-as.numeric(as.character(dat.hom.chicago$transp))
pca <- princomp(na.omit(dat.hom.chicago)[,c(4:15)], cor = TRUE)

mod.zero.inflated.poisson<-zeroinfl(n ~ foreign_share2010 + share_black2010 + share_hisp
2010 + singleparent_share2010+mail_return_rate2010 + scale(transp/sqmi), data = na.omit
(dat.hom.chicago),dist="poisson")

mod.poisson<-glm(n ~ foreign_share2010 + share_black2010 + share_hisp2010 + singleparent
_share2010+mail_return_rate2010 + scale(transp/sqmi), data = na.omit(dat.hom.chicago),fa
mily="poisson")

mod.pca.zero.inflated.poisson<-zeroinfl(na.omit(dat.hom.chicago)$n~scale(na.omit(dat.ho
m.chicago)$transp/na.omit(dat.hom.chicago)$sqmi)+pca$scores[,1] + pca$scores[,2]+pca$sco
res[,3] + pca$scores[,4],dist="poisson")

mod.pca.poisson<-glm(na.omit(dat.hom.chicago)$n~scale(na.omit(dat.hom.chicago)$transp/n
a.omit(dat.hom.chicago)$sqmi)+pca$scores[,1] + pca$scores[,2]+pca$scores[,3] + pca$score
s[,4],family="poisson")

mod.zero.inflated.nb<-zeroinfl(n ~ foreign_share2010 + share_black2010 + share_hisp2010
 + singleparent_share2010+mail_return_rate2010 + scale(transp/sqmi), data = na.omit(dat.
hom.chicago),dist="negbin")

mod.pca.zero.inflated.nb<-zeroinfl(na.omit(dat.hom.chicago)$n~scale(na.omit(dat.hom.chic
ago)$transp/na.omit(dat.hom.chicago)$sqmi)+pca$scores[,1] + pca$scores[,2]+pca$scores[,3
] + pca$scores[,4],dist="negbin")

BIC(mod.poisson,mod.zero.inflated.poisson,mod.pca.poisson,mod.pca.zero.inflated.poisson,
mod.zero.inflated.nb,mod.pca.zero.inflated.nb)
```

```
##                                  df       BIC
## mod.poisson                       7 1149.9749
## mod.zero.inflated.poisson        14  802.7173
## mod.pca.poisson                   6 1067.0552
## mod.pca.zero.inflated.poisson    12  756.5324
## mod.zero.inflated.nb             15  613.3802
## mod.pca.zero.inflated.nb         13  590.4468
```

```
qqrplot(mod.pca.zero.inflated.nb)
```

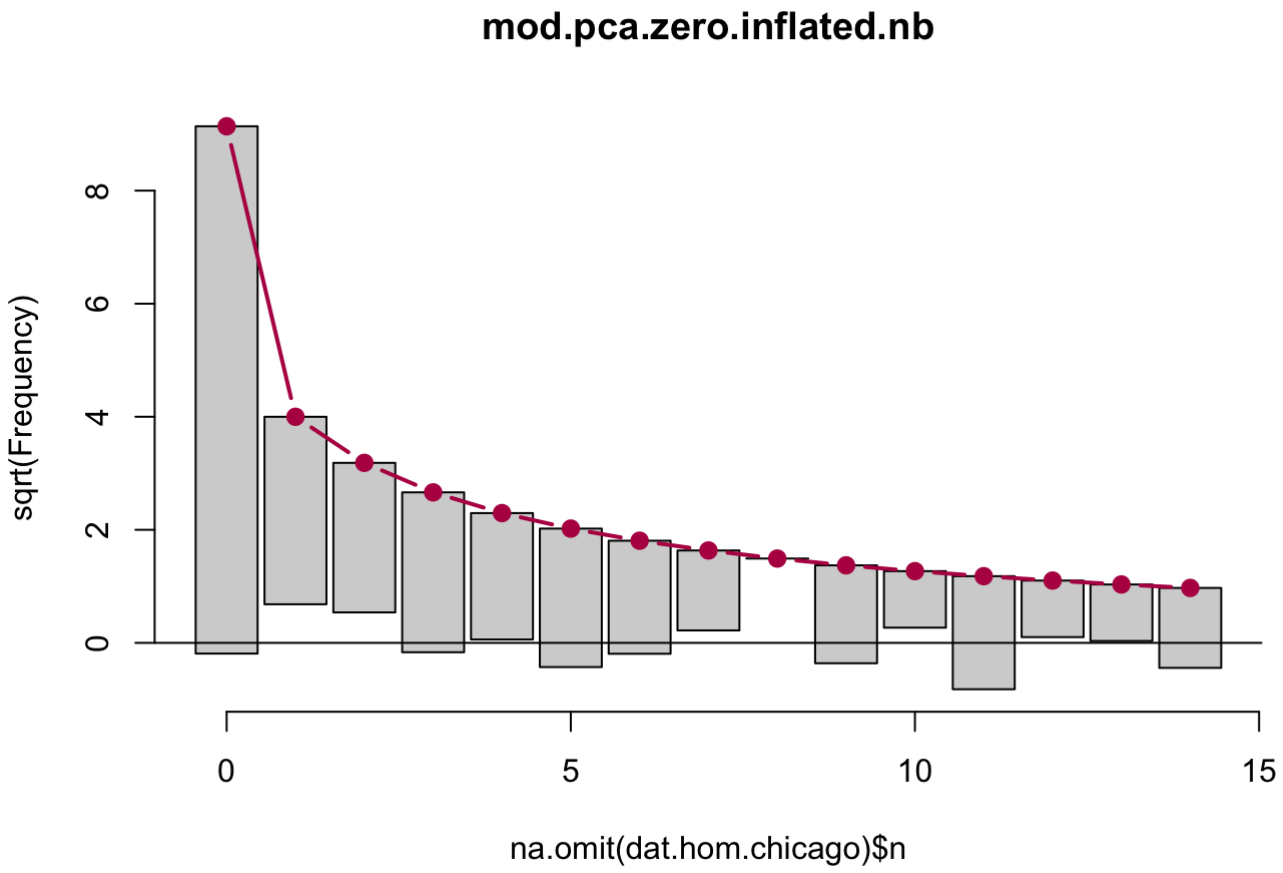



```
summary(mod.pca.zero.inflated.nb)
```

```
##
## Call:
## zeroinfl(formula = na.omit(dat.hom.chicago)$n ~ scale(na.omit(dat.hom.chicago)$trans
p/na.omit(dat.hom.chicago)$sqmi) +
##     pca$scores[, 1] + pca$scores[, 2] + pca$scores[, 3] + pca$scores[,
##     4], dist = "negbin")
##
## Pearson residuals:
##     Min      1Q  Median      3Q     Max
## -0.8916 -0.4972 -0.2242  0.1042  3.6077
##
## Count model coefficients (negbin with log link):
##                                                                      Estimate
## (Intercept)                                                           1.10020
## scale(na.omit(dat.hom.chicago)$transp/na.omit(dat.hom.chicago)$sqmi) -0.01247
## pca$scores[, 1]                                                      -0.30492
## pca$scores[, 2]                                                      -0.21185
## pca$scores[, 3]                                                      -0.81269
## pca$scores[, 4]                                                       0.01948
## Log(theta)                                                           -0.17684
##                                                                      Std. Error
## (Intercept)                                                             0.18511
## scale(na.omit(dat.hom.chicago)$transp/na.omit(dat.hom.chicago)$sqmi)    0.24481
## pca$scores[, 1]                                                         0.07972
## pca$scores[, 2]                                                         0.12769
## pca$scores[, 3]                                                         0.24418
## pca$scores[, 4]                                                         0.18122
## Log(theta)                                                              0.23801
##                                                                      z value
## (Intercept)                                                            5.944
## scale(na.omit(dat.hom.chicago)$transp/na.omit(dat.hom.chicago)$sqmi)  -0.051
## pca$scores[, 1]                                                       -3.825
## pca$scores[, 2]                                                       -1.659
## pca$scores[, 3]                                                       -3.328
## pca$scores[, 4]                                                        0.107
## Log(theta)                                                            -0.743
##                                                                      Pr(>|z|)
## (Intercept)                                                          2.79e-09
## scale(na.omit(dat.hom.chicago)$transp/na.omit(dat.hom.chicago)$sqmi) 0.959370
## pca$scores[, 1]                                                      0.000131
## pca$scores[, 2]                                                      0.097101
## pca$scores[, 3]                                                      0.000874
## pca$scores[, 4]                                                      0.914398
## Log(theta)                                                           0.457506
##
## (Intercept)                                                          ***
## scale(na.omit(dat.hom.chicago)$transp/na.omit(dat.hom.chicago)$sqmi)
## pca$scores[, 1]                                                      ***
## pca$scores[, 2]                                                      .
## pca$scores[, 3]                                                      ***
## pca$scores[, 4]
## Log(theta)
##
```

```
## Zero-inflation model coefficients (binomial with logit link):
##                                                                    Estimate
## (Intercept)                                                       -2.90925
## scale(na.omit(dat.hom.chicago)$transp/na.omit(dat.hom.chicago)$sqmi) -5.17843
## pca$scores[, 1]                                                    0.39672
## pca$scores[, 2]                                                    0.45400
## pca$scores[, 3]                                                    0.08843
## pca$scores[, 4]                                                   -0.48183
##                                                                   Std. Error
## (Intercept)                                                          1.31177
## scale(na.omit(dat.hom.chicago)$transp/na.omit(dat.hom.chicago)$sqmi)    1.91389
## pca$scores[, 1]                                                      0.29379
## pca$scores[, 2]                                                      0.38098
## pca$scores[, 3]                                                      0.50255
## pca$scores[, 4]                                                      0.47923
##                                                                   z value
## (Intercept)                                                        -2.218
## scale(na.omit(dat.hom.chicago)$transp/na.omit(dat.hom.chicago)$sqmi)  -2.706
## pca$scores[, 1]                                                     1.350
## pca$scores[, 2]                                                     1.192
## pca$scores[, 3]                                                     0.176
## pca$scores[, 4]                                                    -1.005
##                                                                   Pr(>|z|)
## (Intercept)                                                        0.02657
## scale(na.omit(dat.hom.chicago)$transp/na.omit(dat.hom.chicago)$sqmi)  0.00682
## pca$scores[, 1]                                                    0.17690
## pca$scores[, 2]                                                    0.23339
## pca$scores[, 3]                                                    0.86033
## pca$scores[, 4]                                                    0.31469
##
## (Intercept)                                                       *
## scale(na.omit(dat.hom.chicago)$transp/na.omit(dat.hom.chicago)$sqmi) **
## pca$scores[, 1]
## pca$scores[, 2]
## pca$scores[, 3]
## pca$scores[, 4]
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Theta = 0.8379
## Number of iterations in BFGS optimization: 27
## Log-likelihood: -262.5 on 13 Df
```

```
rootogram(mod.pca.zero.inflated.nb)
```


mod.pca.zero.inflated.nb
sqrt(Frequency)
0
2
4
6
8
0
5
10
15
na.omit(dat.hom.chicago)$n