Analiza kriminala i socio-ekonomskih faktora

Bruno Ćorić, Filip Škrlec, Jelena Matečić, Iva Zekić

11/6/2020

Učitavanje podataka

Imamo dva skupa podataka kriminala i socio-ekonomskih faktora za grad Chicago.

```
crimeDataset <- read.csv("crime_datasets/Crimes_-_One_year_prior_to_present.csv"
                         ,stringsAsFactors = F,na.strings = "")

povertyDataset <- read.csv("crime_datasets/Chicago_poverty_and_crime.csv"
                           ,stringsAsFactors = F, na.strings="")

head(crimeDataset)
```

```
##      CASE.    DATE..OF.OCCURRENCE                BLOCK IUCR PRIMARY.DESCRIPTION
## 1 JD388829 10/04/2020 08:31:00 PM 086XX S CARPENTER ST 0560             ASSAULT
## 2 JD346990 08/26/2020 01:33:00 PM  011XX N DEARBORN ST 0890               THEFT
## 3 JD403530 10/18/2020 03:50:00 PM     049XX W ADAMS ST 0460             BATTERY
## 4 JD141525 02/05/2020 02:54:00 PM   030XX N HALSTED ST 0860               THEFT
## 5 JD366829 08/26/2020 02:19:00 AM 021XX W CULLERTON ST 0890               THEFT
## 6 JD205528 04/09/2020 02:00:00 PM   029XX S ARCHER AVE 1320     CRIMINAL DAMAGE
##   SECONDARY.DESCRIPTION LOCATION.DESCRIPTION ARREST DOMESTIC BEAT WARD FBI.CD
## 1                SIMPLE            APARTMENT      N        N  613   21    08A
## 2         FROM BUILDING            APARTMENT      N        N 1824    2     06
## 3                SIMPLE               STREET      N        N 1533   28    08B
## 4          RETAIL THEFT           DRUG STORE      N        N 1933   44     06
## 5         FROM BUILDING            APARTMENT      N        N 1234   25     06
## 6            TO VEHICLE               STREET      N        N  913   11     14
##   X.COORDINATE Y.COORDINATE LATITUDE LONGITUDE                     LOCATION
## 1      1170827      1847522 41.73707 -87.64972 (41.737074199, -87.64972468)
## 2           NA           NA       NA        NA                         <NA>
## 3           NA           NA       NA        NA                         <NA>
## 4           NA           NA       NA        NA                         <NA>
## 5           NA           NA       NA        NA                         <NA>
## 6      1168260      1885596 41.84161 -87.65803 (41.841609341, -87.65803375)
```

```
head(povertyDataset)
```

```
##   Community.Area Community.Area.Name Assault..Homicide. Firearm.related
## 1              1         Rogers Park                7.7             5.2
## 2              2         West Ridge                 5.8             3.7
## 3              3             Uptown                 5.4             4.6
## 4              4      Lincoln Square                5.0             6.1
## 5              5        North Center                1.0             1.0
## 6              6           Lake View                1.4             1.8
##   Below.Poverty.Level Crowded.Housing Dependency No.High.School.Diploma
```

```
## 1                22.7             7.9       28.8                     18.1
## 2                15.1             7.0       38.3                     19.6
## 3                22.7             4.6       22.2                     13.6
## 4                 9.5             3.1       25.6                     12.5
## 5                 7.1             0.2       25.5                      5.4
## 6                10.5             1.2       16.5                      2.9
##   Per.Capita.Income Unemployment
## 1             23714          7.5
## 2             21375          7.9
## 3             32355          7.7
## 4             35503          6.8
## 5             51615          4.5
## 6             58227          4.7
```

Faktorizirat ćemo podatke koje bi bilo logično faktorizirati kao što su podaci u stupcu Arrest,Domestic.

```
crimeDataset$ARREST <- as.factor(crimeDataset$ARREST)
crimeDataset$DOMESTIC <- as.factor(crimeDataset$DOMESTIC)
```

Provjeravamo fale li nam neki podaci u najbitnim kategorijama u oba dataseta.

```
s <-c(1,2,3,4,5,6,8,9)
sum(is.na(crimeDataset[s]))
```

```
## [1] 0
```

```
sum(is.na(povertyDataset))
```

```
## [1] 0
```

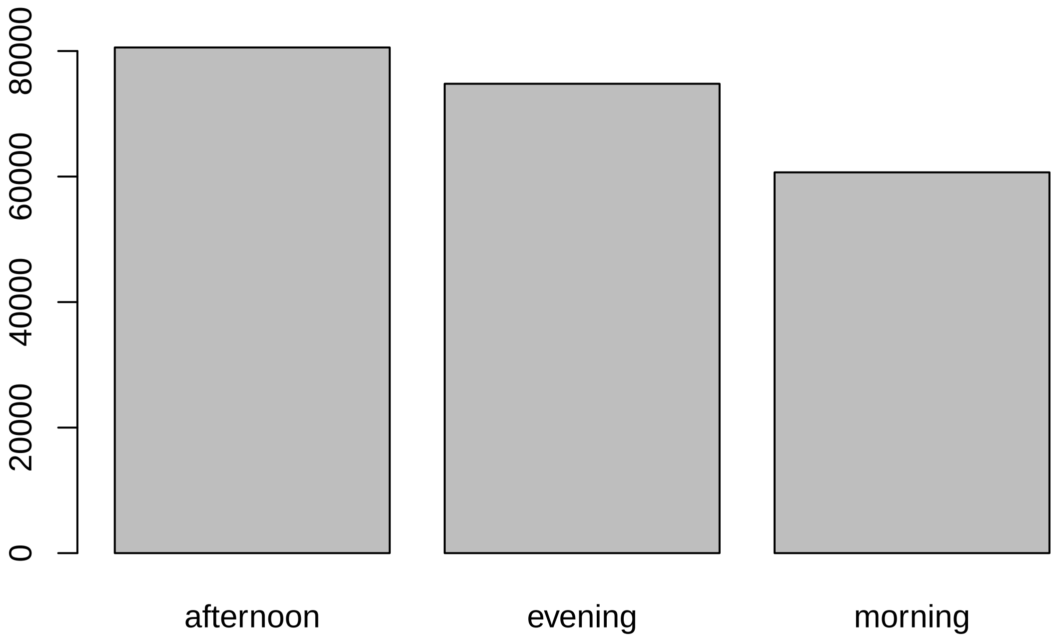
Razlika učestalosti zločina ovisno o tome koje je doba dana

Podijelit ćemo dan na 3 dijela. Od 5 do 13 će biti prvi dio dana. Od 13 do 21 drugi dio dana, a od 20 do 5 treći dio dana.

```
timeOfTheDay <- mdy_hms(crimeDataset$DATE..OF.OCCURRENCE) %>% hour
timeOfTheDay <- sapply(timeOfTheDay, function(x) {
  if(x >= 5 & x < 13) {
    "morning"
  } else if(x >= 13 & x < 20) {
    "afternoon"
  } else {
    "evening"
  }
},simplify="vector")
timeOfTheDay <- as.factor(timeOfTheDay)
crimeDataset$TIME.OF.DAY <- timeOfTheDay
timeOfTheDayCount <- crimeDataset %>% group_by(TIME.OF.DAY) %>% tally
head(crimeDataset[c("DATE..OF.OCCURRENCE","TIME.OF.DAY")])
```

```
##      DATE..OF.OCCURRENCE TIME.OF.DAY
## 1 10/04/2020 08:31:00 PM     evening
## 2 08/26/2020 01:33:00 PM   afternoon
## 3 10/18/2020 03:50:00 PM   afternoon
## 4 02/05/2020 02:54:00 PM   afternoon
## 5 08/26/2020 02:19:00 AM     evening
## 6 04/09/2020 02:00:00 PM   afternoon
```

```r
barplot(table(crimeDataset$TIME.OF.DAY))
```



Napravit ćemo goodnes of fit test nad brojem kriminala koji se dogodio ujutro, popodne i navečer. Nulta hipoteza testa je da je očekivana proporcija 1/3 za broj kriminala u određenom dijelu dana, tj. da se ne razlikuje broj kriminala s obzirom na vrijeme.

```r
chisq.test(timeOfTheDayCount$n)
```

```
##
##  Chi-squared test for given probabilities
##
## data:  timeOfTheDayCount$n
## X-squared = 2909.8, df = 2, p-value < 2.2e-16
```

Odbacujemo nultu hipotezu i zaključujemo da da su proporcije različite.

Napravit ćemo test o homogenosti u kojem želimo viditi postoji li razlika u količini zločina s obzirom na doba dana.

Napravit ćemo test homogenosti u kojem ćemo provjeriti je li broj zločina opasnih po život jednak za sva 3 doba dana. Zločine koje smo uzeli da su opasni po život nalaze se u varijabli `dangCrimes`.

dangerousCrimes	Freq
dangerous	103354
less dangerous	112678

	dangerous	less dangerous
afternoon	36047	44524
evening	40592	34196
morning	26715	33958

```
##
##  Pearson's Chi-squared test
```

```
##
## data:  dangerous
## X-squared = 1904.6, df = 2, p-value < 2.2e-16
```

Zaključujemo da se razlikuje količina opasnih i neopasnih zločina ovisno o tome koje je doba dana.

Je li učestalost krađa veća od učestalosti kriminala vezanih za narkotike?

U varijablu 'description' odvajamo objekt PRIMARY.DESCRIPTION iz našeg skupa podataka crimeDataset. Zatim tu varijablu pretvara u skup podataka ali na način da se sve iste vrijednosti prebroje i upišu u stupaca Freq. U varijable krada i narkotici odvojimo podatke vezane za te zlocine. U varijable broj__krada i broj__narkotici zatim stavljamo sumu retka freq iz prethodnih varijabli.

```r
description <- crimeDataset$PRIMARY.DESCRIPTION

description <- as.data.frame(table(description))

krada <- description[c(18, 28, 31),]
narkotici <- description[c(19, 23),]

broj_krada <- sum(krada$Freq)
broj_narkotici <- sum(narkotici$Freq)

cat("Učestalost krađa: ")
```

```
## Učestalost krađa:
```

```r
broj_krada
```

```
## [1] 63083
```

```r
cat("\n")
```

```r
cat("Učestalost kriminala vezanih za narkotike: ")
```

```
## Učestalost kriminala vezanih za narkotike:
```

```r
broj_narkotici
```

```
## [1] 8268
```

```r
cat("\n")
```

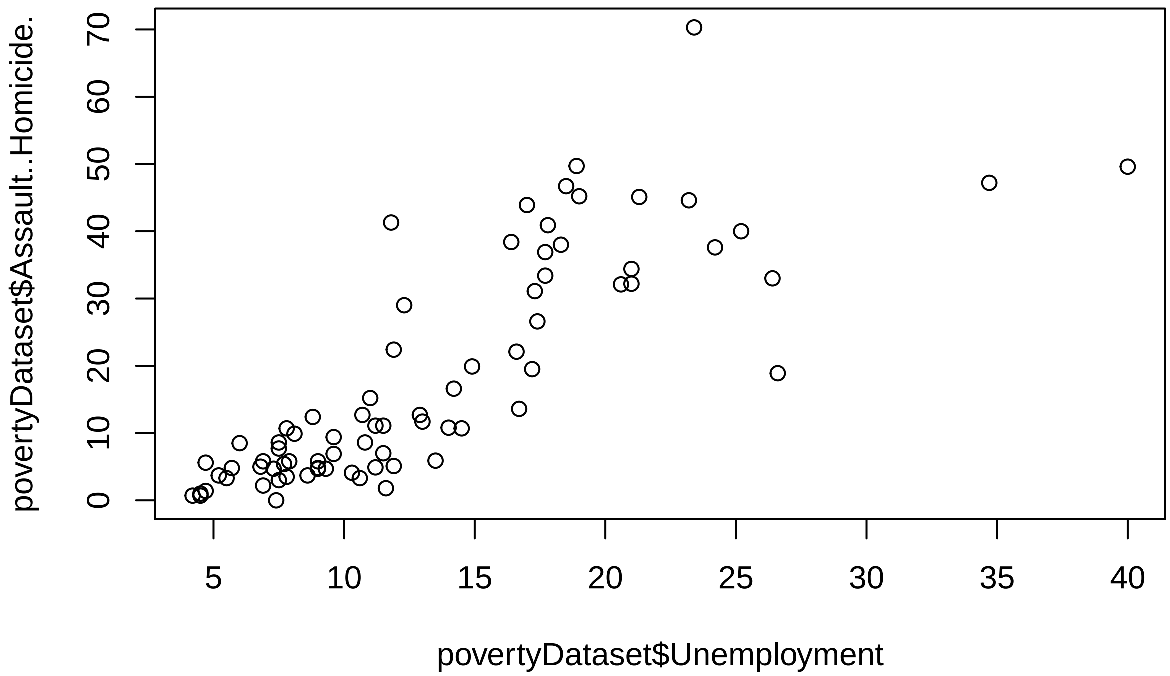
Veza između socio-ekonomskih faktora i pojedine kategorije kriminala

Ispitivati ćemo različite varijable koje bi mogle utjecati na "Assault Homicide" i "Firearm related" kategroije kriminala. Varijable koje ćemo promatrati su (postotci predstavljaju postotak broja stanovništva za određeni kvart): - postotak stanovništva koji su siromašni - postotak stanovništva koji žive u prenatrpanoj kući - postotak ljudi mlađih od 16 ili starijih od 64 koji su financijski ovisni o nekome - postotak ljudi bez diplome srednje škole - dohodak po stanovniku - postotak ljudi koji nisu zaposleni

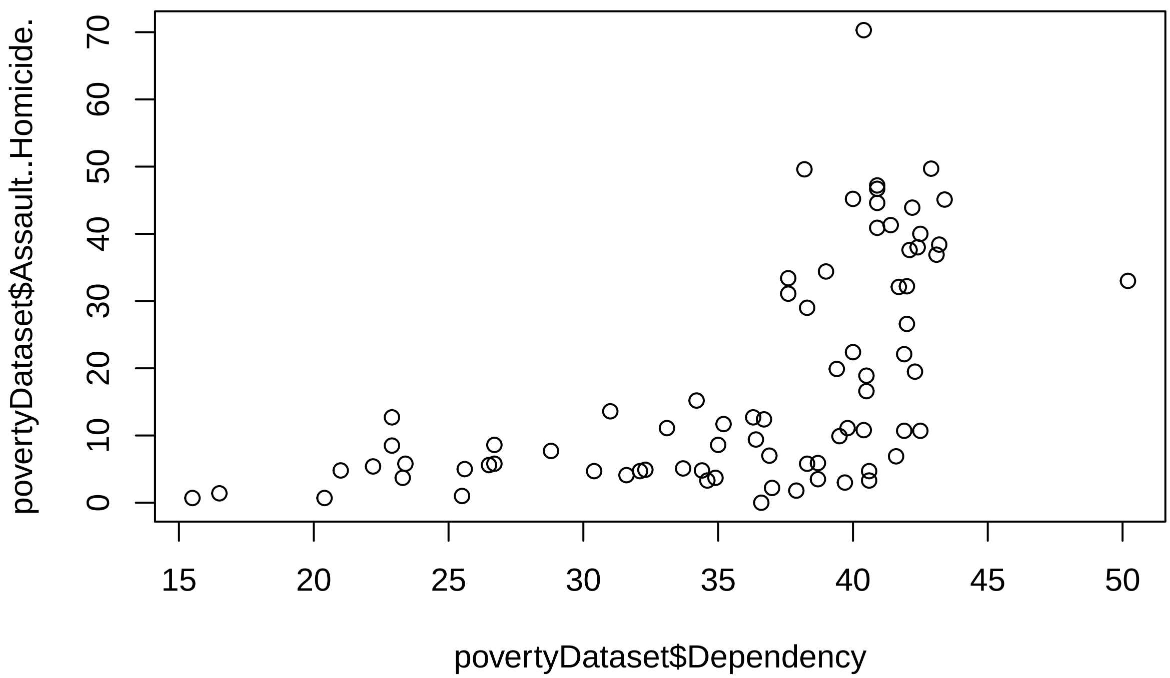
Nacratat ćemo nekoliko grafova kako bi dobili uvid u to kakav odnos imaju varijable.

Vidimo linearan efekt kod nezaposlenosti i siromaštva. Dependency izgleda kao eksponencijalna funkcija dok Per Capita Income kao logaritamska.

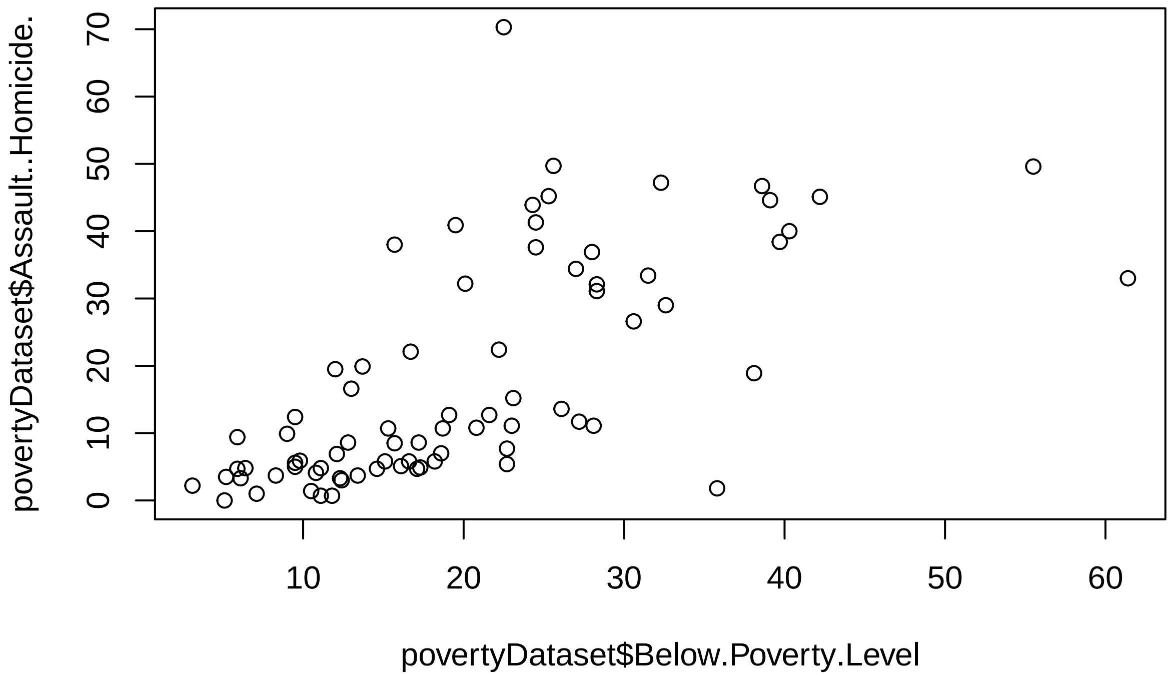
```r
plot(povertyDataset$Unemployment, povertyDataset$Assault..Homicide.)
```



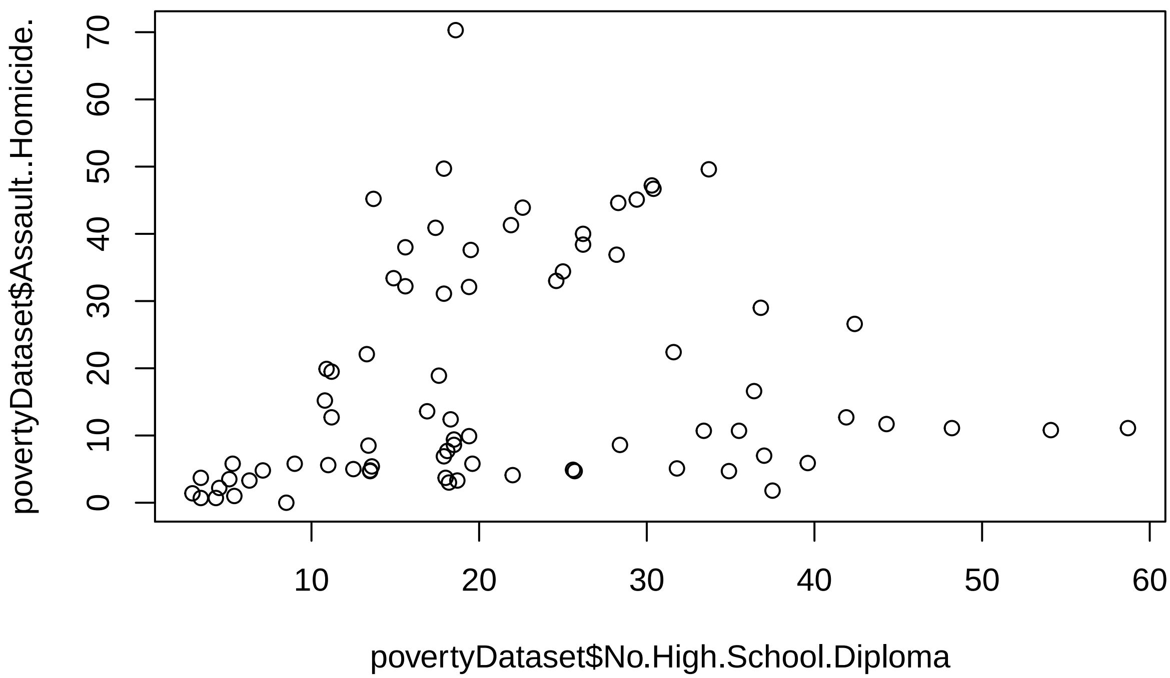

```
plot(povertyDataset$Dependency, povertyDataset$Assault..Homicide.)
```



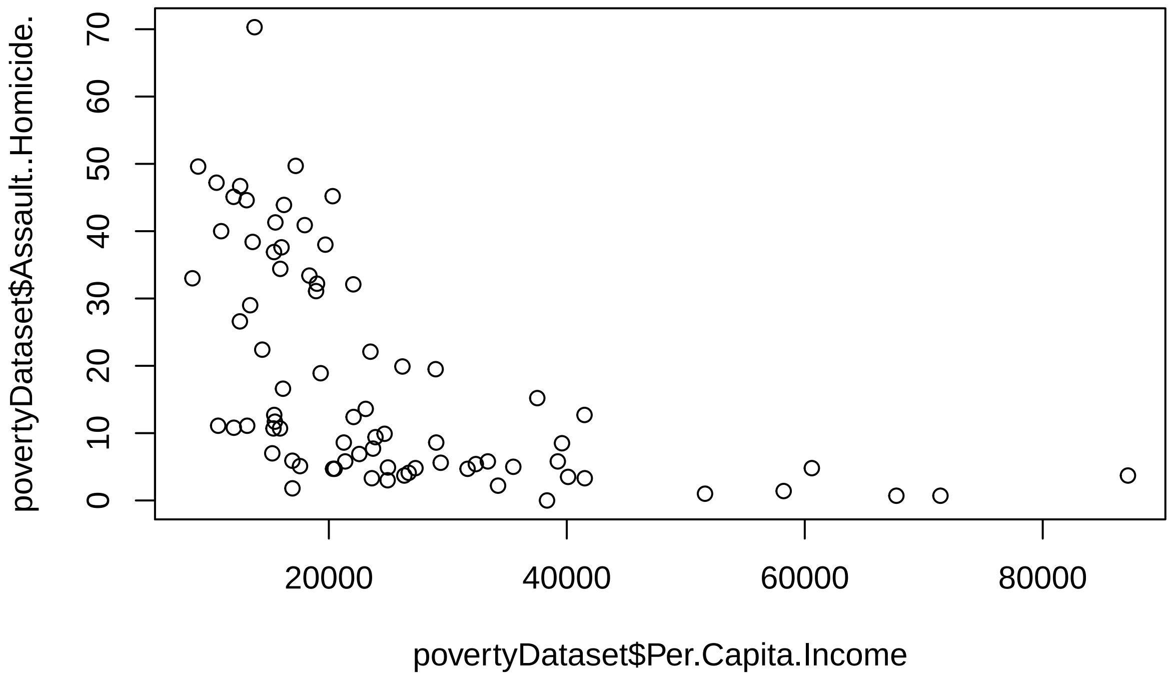

```
plot(povertyDataset$Below.Poverty.Level, povertyDataset$Assault..Homicide.)
```



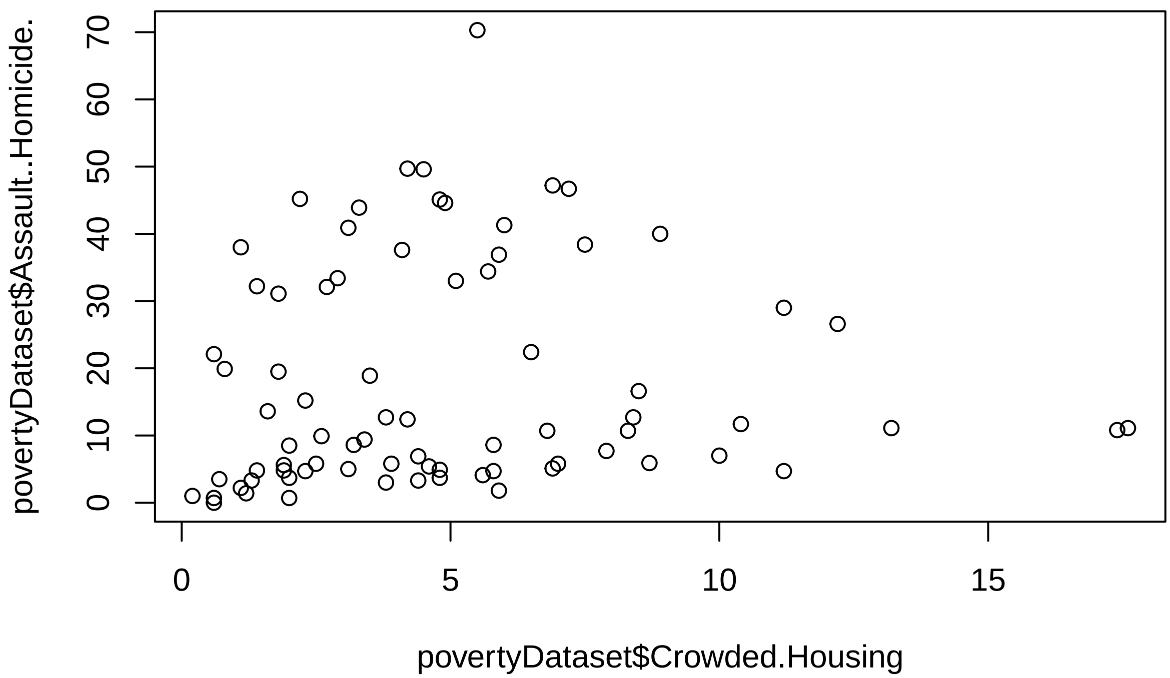

```
plot(povertyDataset$No.High.School.Diploma, povertyDataset$Assault..Homicide.)
```




```
plot(povertyDataset$Per.Capita.Income, povertyDataset$Assault..Homicide.)
```




```
plot(povertyDataset$Crowded.Housing, povertyDataset$Assault..Homicide.)
```



Neke varijable su jako korelirane. Što je i bilo za očekivati.

```
cor(povertyDataset$Firearm.related, povertyDataset$Assault..Homicide.)
```

```
## [1] 0.9671702
```

```
cor(povertyDataset$No.High.School.Diploma, povertyDataset$Crowded.Housing)
```

```
## [1] 0.905274
```

```r
cor(povertyDataset$Below.Poverty.Level, povertyDataset$Unemployment)
```

```
## [1] 0.763817
```

```r
cor(povertyDataset$Dependency, povertyDataset$Per.Capita.Income)
```

```
## [1] -0.7565786
```

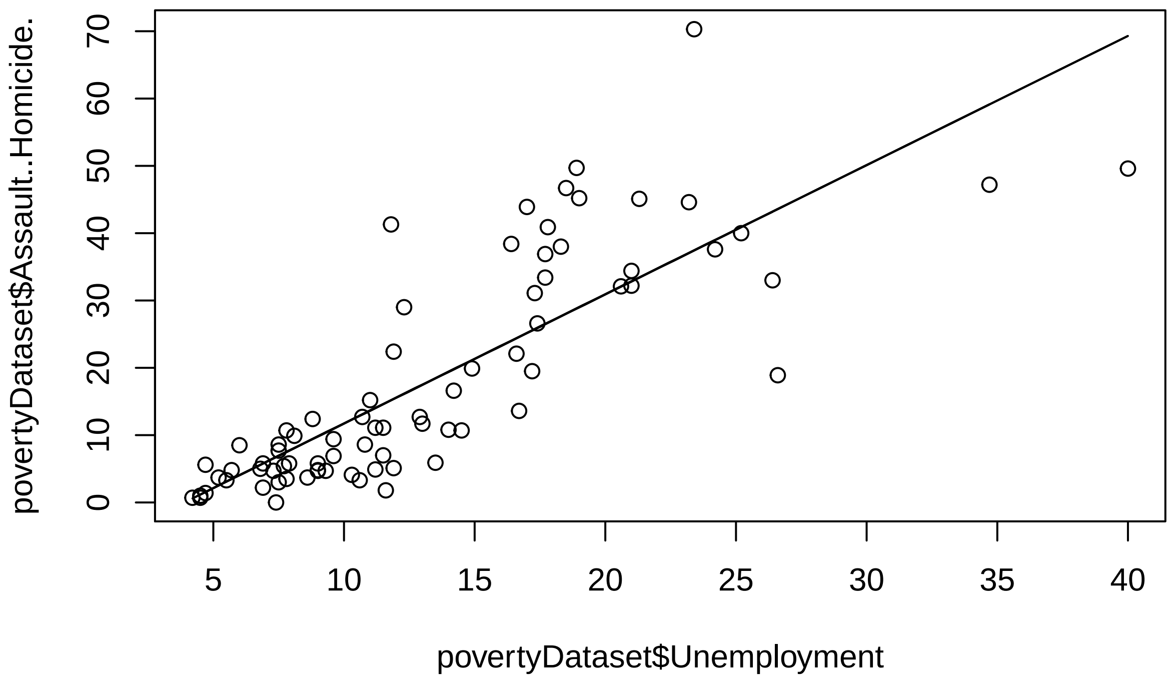
Jednostavne regresije

Izvodjiti ćemo neke zanimljivije jednostavnije modele.

Pošto su Assault Homicide i Firearm related jako korelirani, modeli za njih su jako slični.

Prvo procjenjujemo ubojstva pomoću varijable koja prikazuje nezaposlenost. Dobivamo mjeru kvalitete prilagodbe R^2 = 0.664 što je jako dobro za predviđanje sa samo jednom varijablom, a i očito je iz grafa.

```r
fit.AssaultUnemployment <- lm(Assault..Homicide.~Unemployment,data=povertyDataset)
plot(povertyDataset$Unemployment, povertyDataset$Assault..Homicide.)
lines(povertyDataset$Unemployment, fit.AssaultUnemployment$fitted.values)
```




```r
summary(fit.AssaultUnemployment)
```

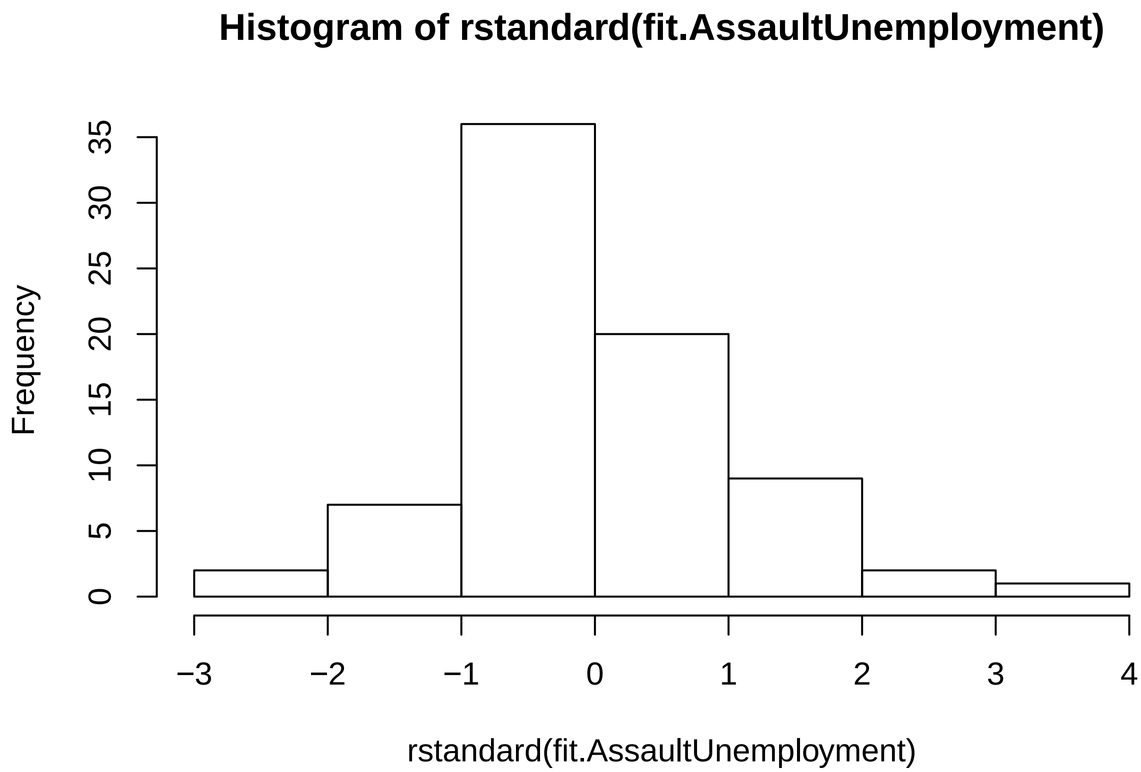
```
##
## Call:
## lm(formula = Assault..Homicide. ~ Unemployment, data = povertyDataset)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -24.684  -5.110  -0.898   2.974  32.856
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   -7.4617     2.3689   -3.15  0.00235 **
## Unemployment   1.9190     0.1576   12.17  < 2e-16 ***
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 9.664 on 75 degrees of freedom
## Multiple R-squared:  0.664,	Adjusted R-squared:  0.6595
## F-statistic: 148.2 on 1 and 75 DF,  p-value: < 2.2e-16
```

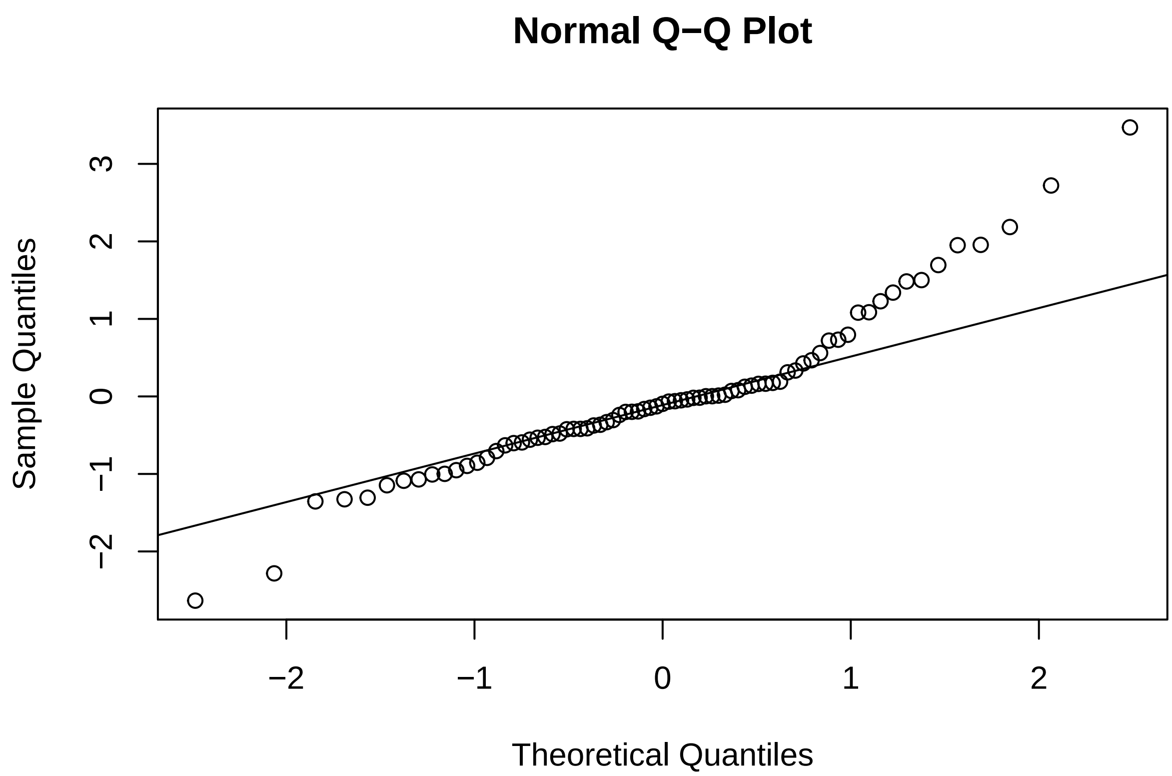
```r
ks.test(rstandard(fit.AssaultUnemployment), 'pnorm')
```

```
##
##  One-sample Kolmogorov-Smirnov test
##
## data:  rstandard(fit.AssaultUnemployment)
## D = 0.16491, p-value = 0.0268
## alternative hypothesis: two-sided
```

```r
hist(rstandard(fit.AssaultUnemployment))
```

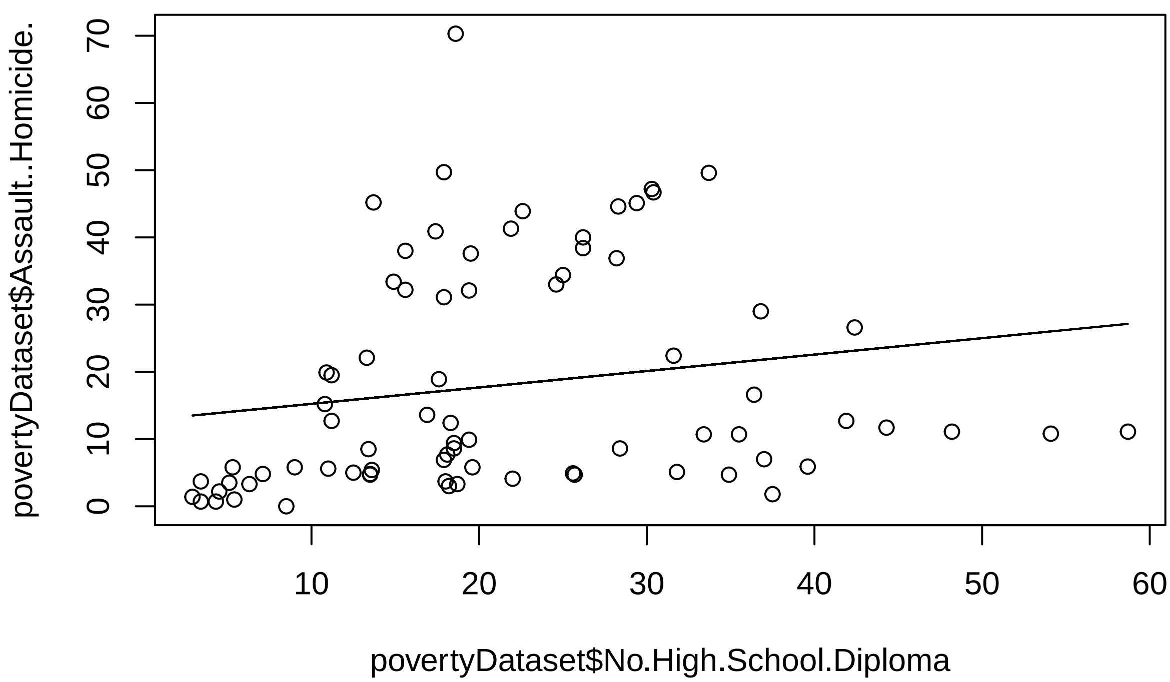



```r
qqnorm(rstandard(fit.AssaultUnemployment))
qqline(rstandard(fit.AssaultUnemployment))
```



Sada koristimo varijablu No High School Diploma. Uočavamo da reziduali nisu ni približno distribuirani po normalnoj distribuciji te je pretpostavka linearne regresije narušena što možda upućuje da nam treba neki složeniji model.

```
fit.AssaultDiploma <- lm(Assault..Homicide.
                         ~No.High.School.Diploma,data=povertyDataset)
plot(povertyDataset$No.High.School.Diploma, povertyDataset$Assault..Homicide.)
lines(povertyDataset$No.High.School.Diploma, fit.AssaultDiploma$fitted.values)
```



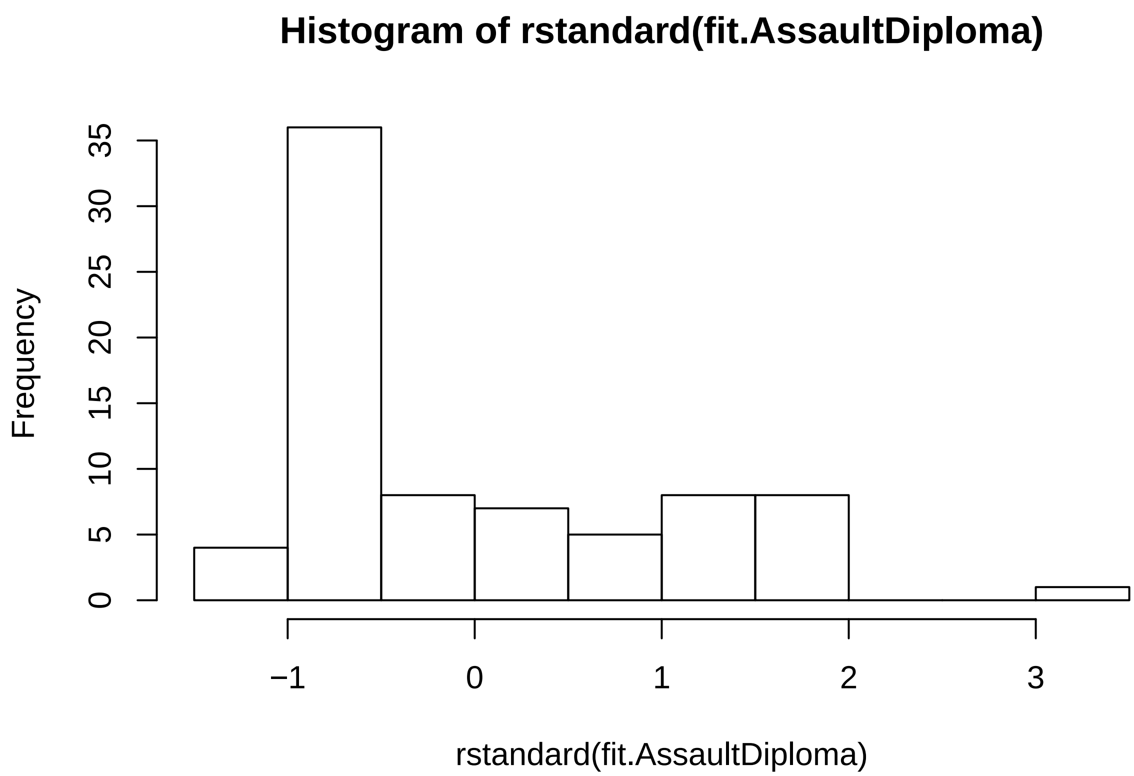
```r
summary(fit.AssaultDiploma)
```

```
##
## Call:
## lm(formula = Assault..Homicide. ~ No.High.School.Diploma, data = povertyDataset)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -20.154 -11.916  -8.712  14.568  52.963
##
## Coefficients:
##                        Estimate Std. Error t value Pr(>|t|)
## (Intercept)             12.7925     3.7804   3.384  0.00114 **
## No.High.School.Diploma   0.2443     0.1522   1.605  0.11262
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 16.39 on 75 degrees of freedom
## Multiple R-squared:  0.03322,    Adjusted R-squared:  0.02033
## F-statistic: 2.577 on 1 and 75 DF,  p-value: 0.1126
```

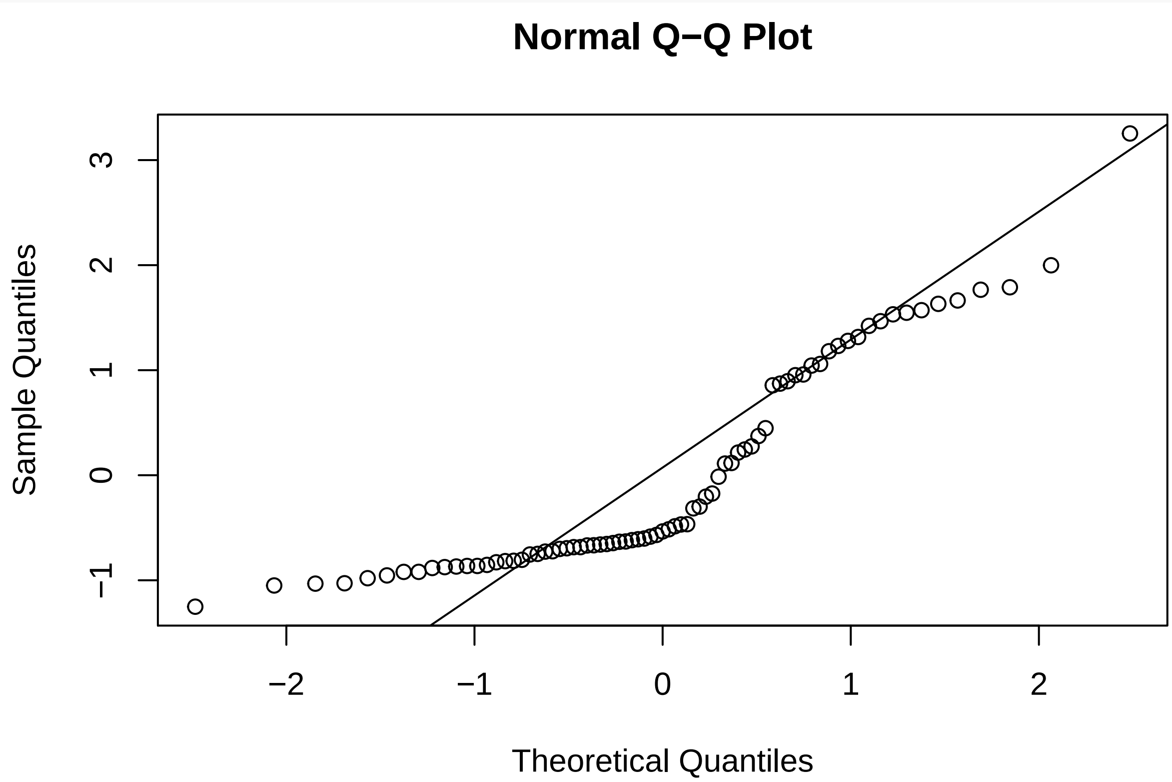
```r
ks.test(rstandard(fit.AssaultDiploma), 'pnorm')
```

```
##
##  One-sample Kolmogorov-Smirnov test
##
## data:  rstandard(fit.AssaultDiploma)
## D = 0.23783, p-value = 0.0002572
## alternative hypothesis: two-sided
```

```r
hist(rstandard(fit.AssaultDiploma))
```



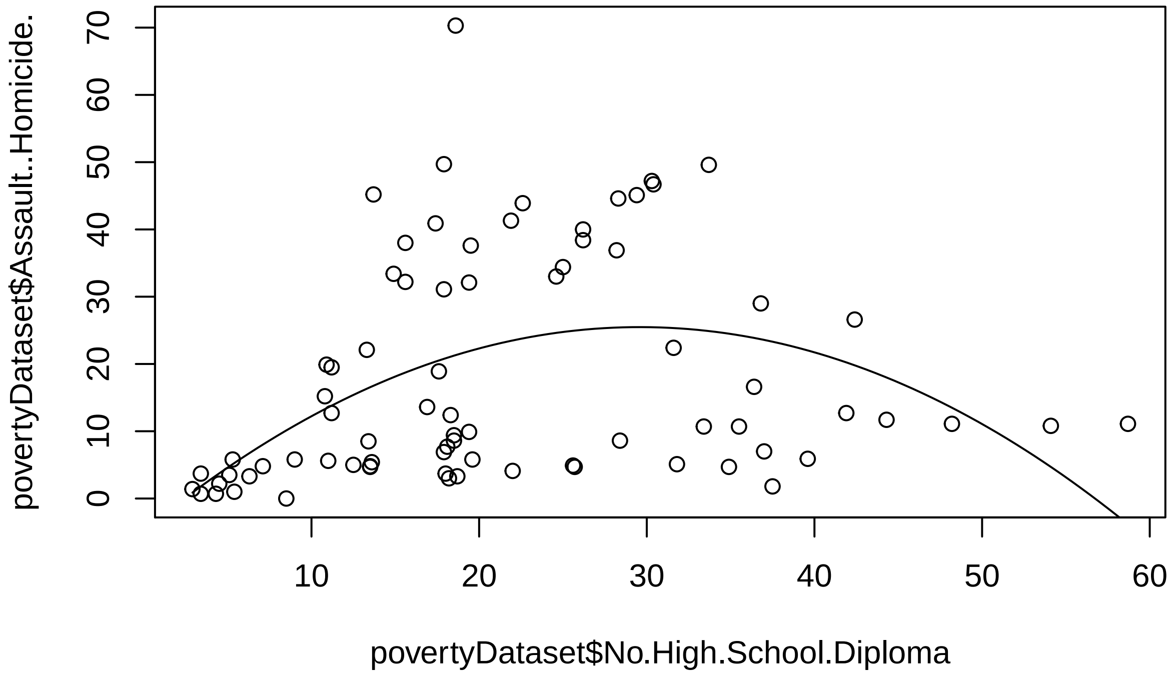

```r
qqnorm(rstandard(fit.AssaultDiploma))
qqline(rstandard(fit.AssaultDiploma))
```



Ako koristimo polinomijalnu regresiju dobivamo puno bolje rezultate

```r
fit.AssaultDiplomaSq <- lm(Assault..Homicide.~No.High.School.Diploma+
                             I(No.High.School.Diploma^2),data=povertyDataset)

plot(povertyDataset$No.High.School.Diploma, povertyDataset$Assault..Homicide.)
curve(predict(fit.AssaultDiplomaSq,
              newdata=data.frame(No.High.School.Diploma=x)),add=T)
```




```r
summary(fit.AssaultDiplomaSq)
```

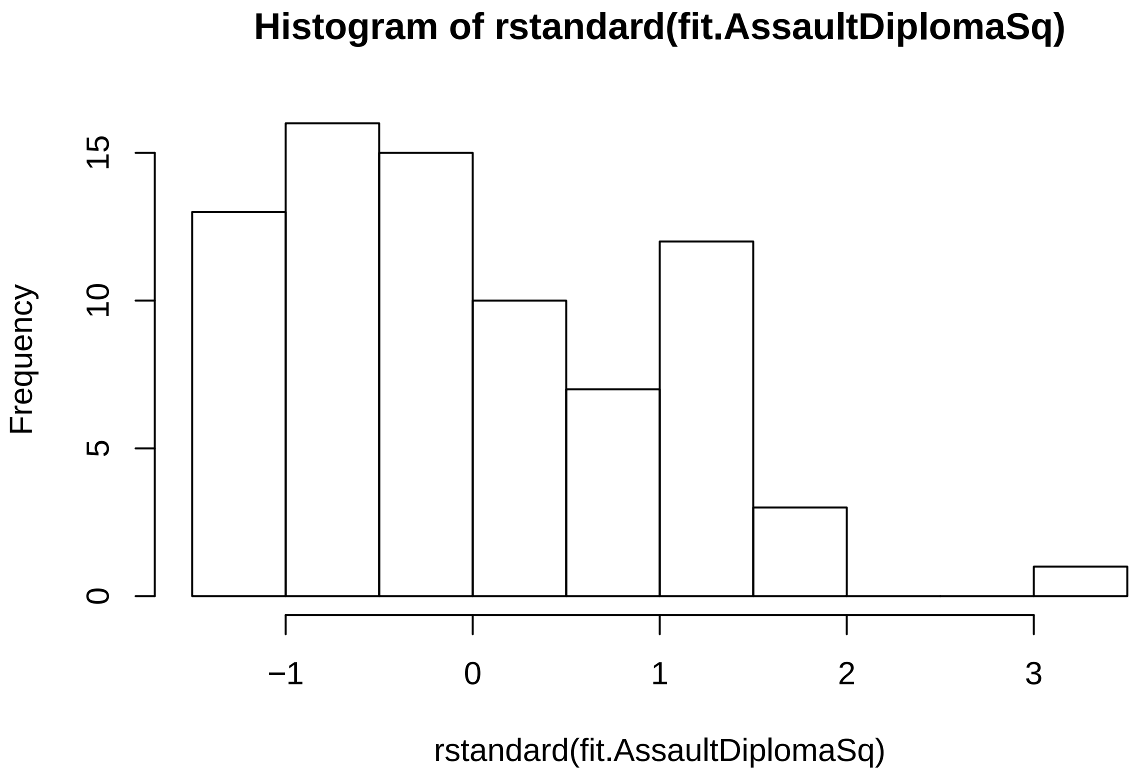
```
##
## Call:
## lm(formula = Assault..Homicide. ~ No.High.School.Diploma + I(No.High.School.Diploma^2),
##     data = povertyDataset)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -21.512 -11.842  -2.399  10.336  48.988
##
## Coefficients:
##                              Estimate Std. Error t value Pr(>|t|)
## (Intercept)                 -4.761998   5.634127  -0.845 0.400720
## No.High.School.Diploma       2.044636   0.476875   4.288 5.38e-05 ***
## I(No.High.School.Diploma^2) -0.034560   0.008755  -3.947 0.000178 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 15 on 74 degrees of freedom
## Multiple R-squared:  0.2014, Adjusted R-squared:  0.1798
## F-statistic:  9.33 on 2 and 74 DF,  p-value: 0.0002436
```

```r
ks.test(rstandard(fit.AssaultDiplomaSq), 'pnorm')
```

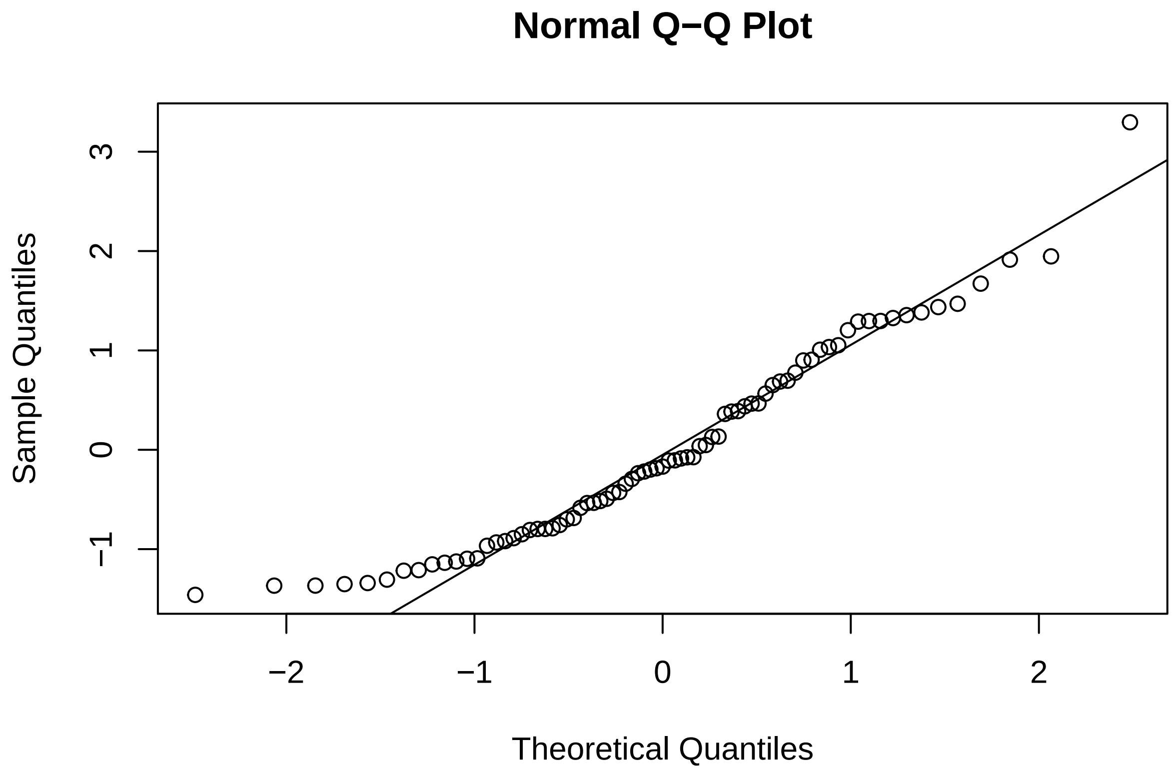
```
##
```

```
##  One-sample Kolmogorov-Smirnov test
##
## data:  rstandard(fit.AssaultDiplomaSq)
## D = 0.10108, p-value = 0.3853
## alternative hypothesis: two-sided
```

```r
hist(rstandard(fit.AssaultDiplomaSq))
```

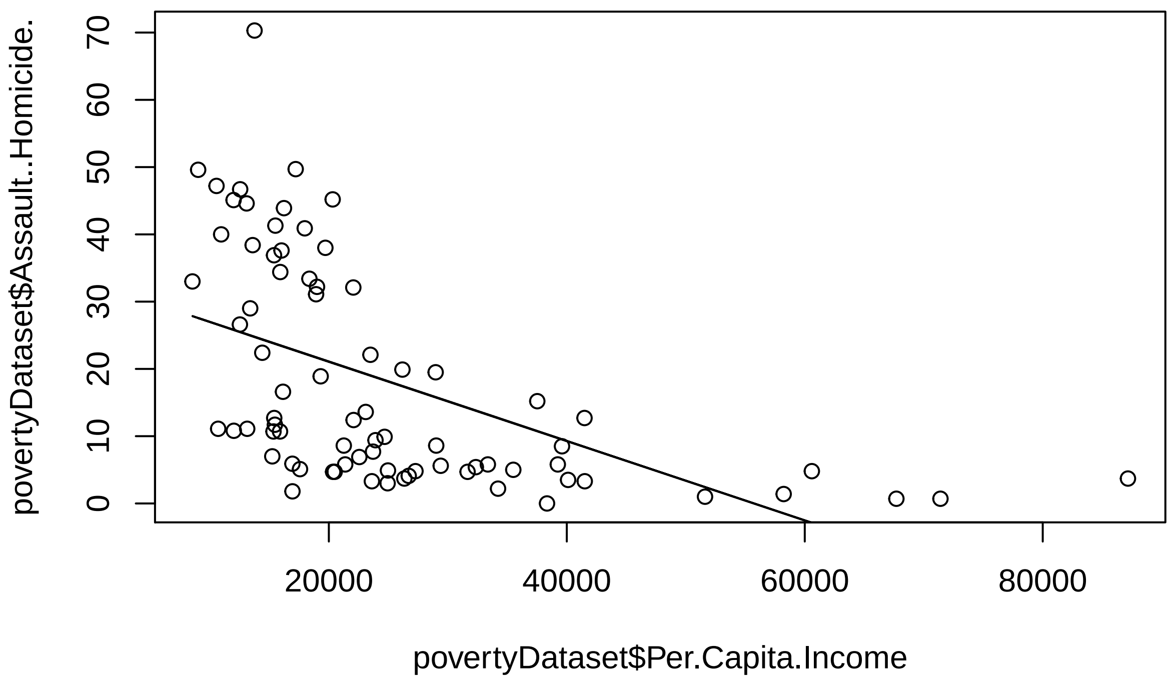



```r
qqnorm(rstandard(fit.AssaultDiplomaSq))
qqline(rstandard(fit.AssaultDiplomaSq))
```



Korištenjem dohodka po glavi dobivamo ne toliko dobar model, ali iz grafa možemo uočiti koliko ima manje zločina u prosječno bogatijim kvartovima.

```
fit.AssaultIncome <- lm(Assault..Homicide.~Per.Capita.Income,data=povertyDataset)
plot(povertyDataset$Per.Capita.Income, povertyDataset$Assault..Homicide.)
lines(povertyDataset$Per.Capita.Income, fit.AssaultIncome$fitted.values)
```



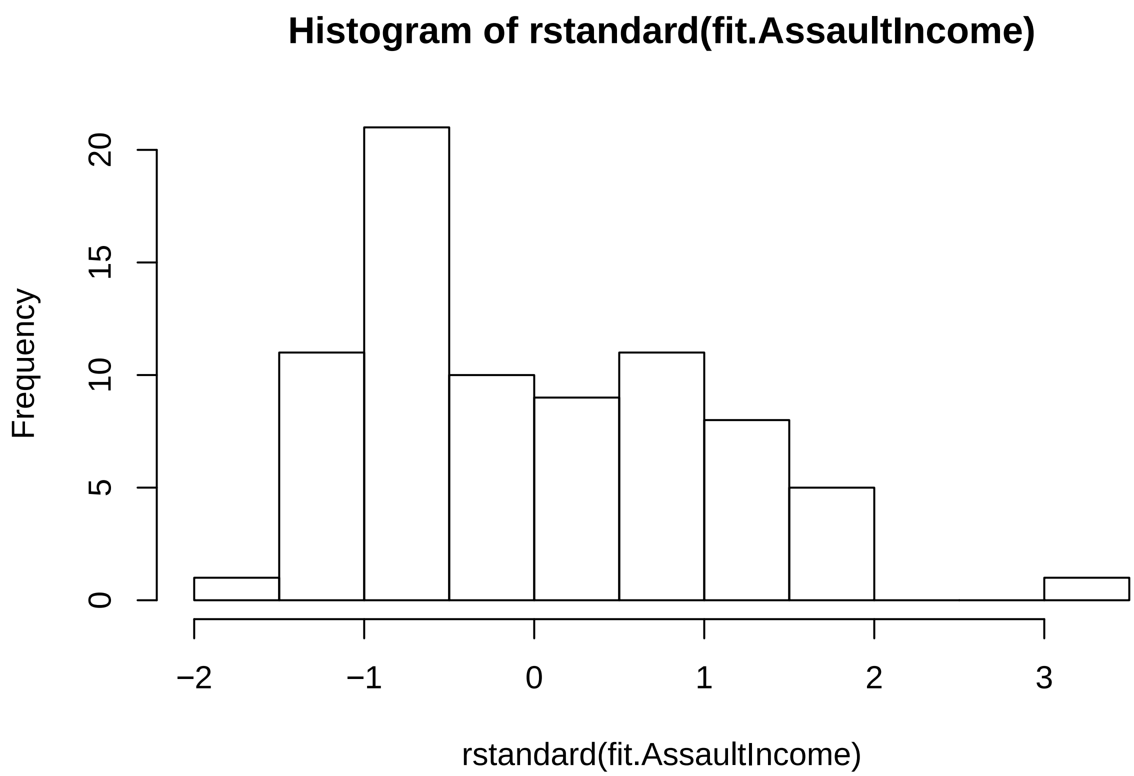
```r
summary(fit.AssaultIncome)
```

```
##
## Call:
## lm(formula = Assault..Homicide. ~ Per.Capita.Income, data = povertyDataset)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -21.087 -11.986  -3.928  10.910  45.534
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)       32.8833950  3.1573233  10.415 3.19e-16 ***
## Per.Capita.Income -0.0005901  0.0001082  -5.452 6.11e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 14.11 on 75 degrees of freedom
## Multiple R-squared:  0.2838, Adjusted R-squared:  0.2743
## F-statistic: 29.72 on 1 and 75 DF,  p-value: 6.114e-07
```

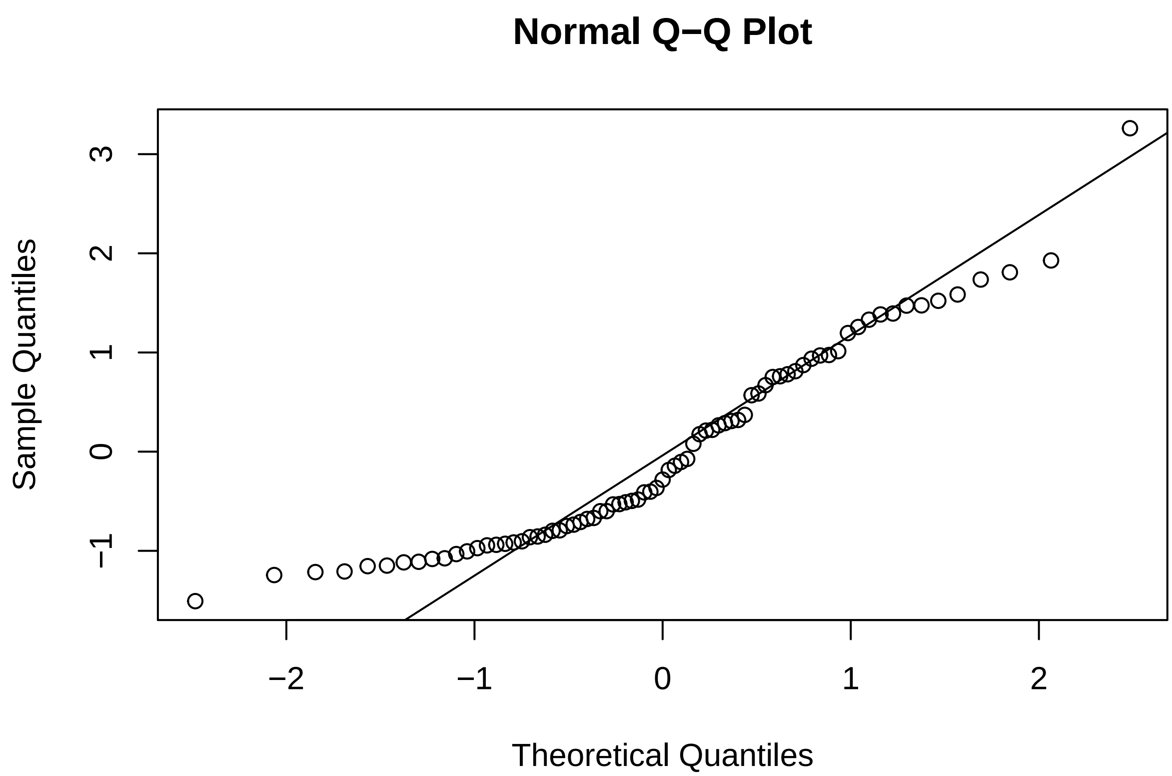
```r
ks.test(rstandard(fit.AssaultIncome), 'pnorm')
```

```
##
##  One-sample Kolmogorov-Smirnov test
##
## data:  rstandard(fit.AssaultIncome)
## D = 0.14008, p-value = 0.0883
## alternative hypothesis: two-sided
```

```r
hist(rstandard(fit.AssaultIncome))
```

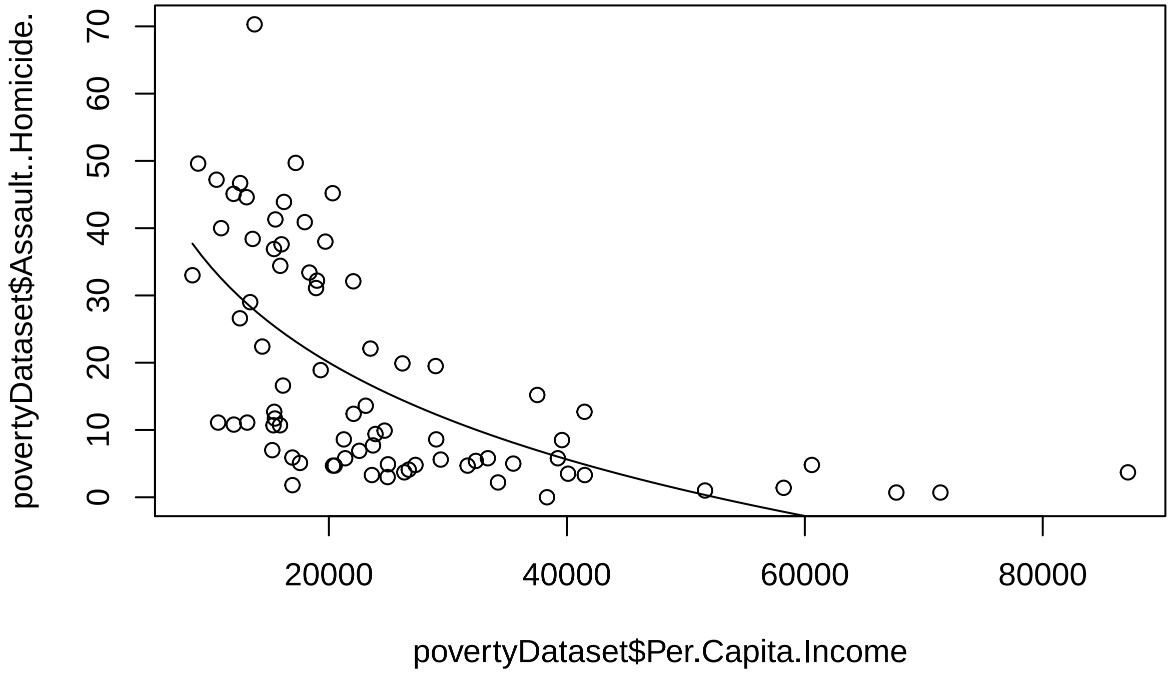



```
qqnorm(rstandard(fit.AssaultIncome))
qqline(rstandard(fit.AssaultIncome))
```



Primjenom transformacije logaritmom nad ulaznim podacima Per Capita Income dobivamo puno bolji rezultat.

```r
fit.AssaultIncome <- lm(Assault..Homicide.~log(Per.Capita.Income),data=povertyDataset)
plot(povertyDataset$Per.Capita.Income, povertyDataset$Assault..Homicide.)
curve(predict(fit.AssaultIncome,
              newdata=data.frame(Per.Capita.Income=x)),add=T)
```




```r
summary(fit.AssaultIncome)
```

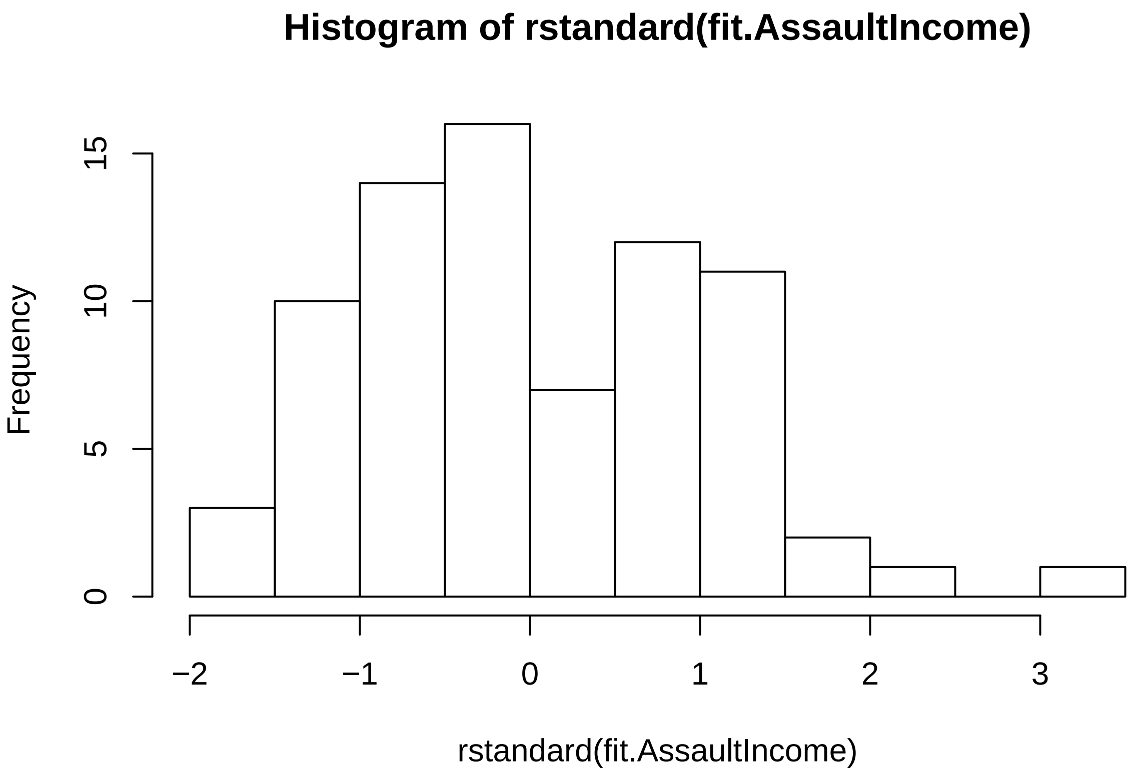
```
##
## Call:
## lm(formula = Assault..Homicide. ~ log(Per.Capita.Income), data = povertyDataset)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -21.918 -10.160  -3.111   9.930  42.504
##
## Coefficients:
##                        Estimate Std. Error t value Pr(>|t|)
## (Intercept)             225.645     30.203   7.471 1.20e-10 ***
## log(Per.Capita.Income)  -20.762      3.017  -6.881 1.55e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 13.05 on 75 degrees of freedom
## Multiple R-squared:  0.387,  Adjusted R-squared:  0.3788
## F-statistic: 47.35 on 1 and 75 DF,  p-value: 1.55e-09
```

```r
ks.test(rstandard(fit.AssaultIncome), 'pnorm')
```

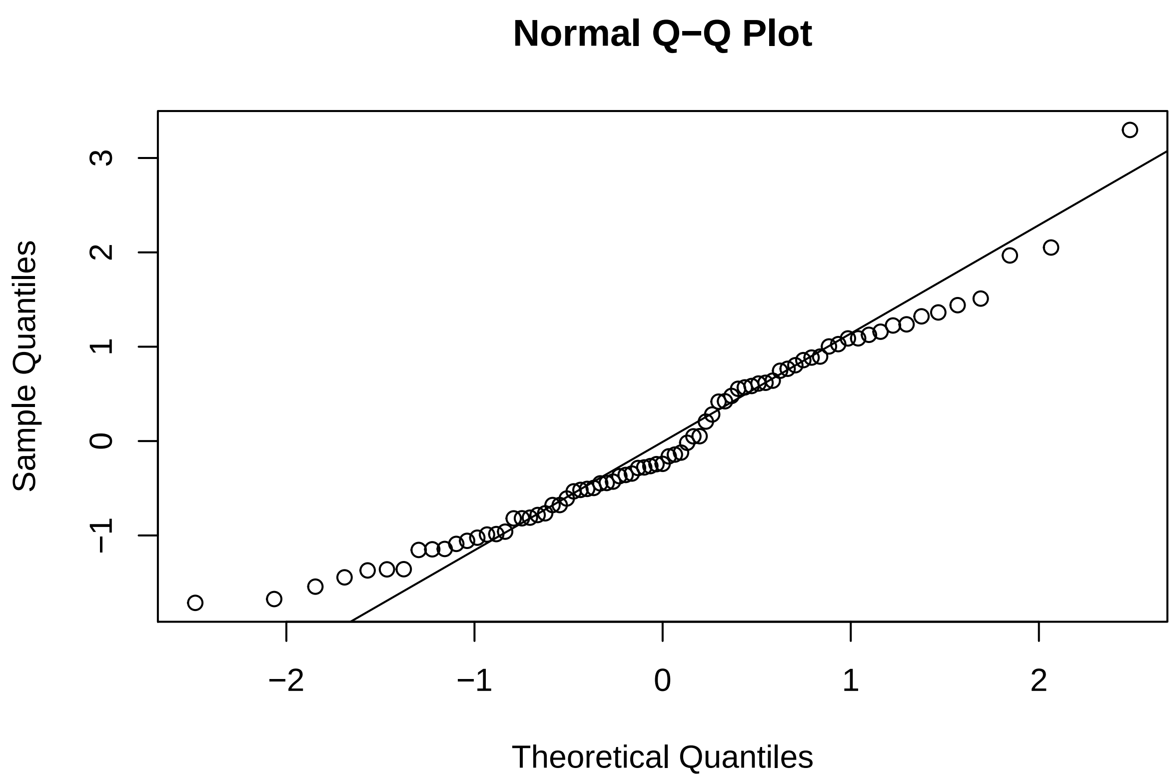
```
##
##  One-sample Kolmogorov-Smirnov test
##
## data:  rstandard(fit.AssaultIncome)
## D = 0.10187, p-value = 0.3759
```

```
## alternative hypothesis: two-sided
```

```r
hist(rstandard(fit.AssaultIncome))
```




```r
qqnorm(rstandard(fit.AssaultIncome))
qqline(rstandard(fit.AssaultIncome))
```



Višestruka regresija

Prije procjene modela višestruke regresije trebamo provjeriti jesu li varijable međusobno zavisne. Ako nemaju vrlo visoku korelaciju možemo ih koristiti zajedno u modeliranju. Već smo pokazali neke varijable koje imaju veliku korelaciju.

Kao što smo mogli očekivati nezaposlenost i siromaštvo objašnjavaju iste efekte u podacima te nećemo dobiti puno bolji model nego samo sa korištenjem siromaštva. Testirali smo i s drugim varijablama u kombinaciji sa nezaposlenošću i zaključili smo da

```
fit1 <- lm(povertyDataset$Assault..Homicide.~povertyDataset$Unemployment +
             povertyDataset$Below.Poverty.Level)
summary(fit1)
```

```
##
## Call:
## lm(formula = povertyDataset$Assault..Homicide. ~ povertyDataset$Unemployment +
##     povertyDataset$Below.Poverty.Level)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -24.871  -4.487  -0.780   2.902  34.466
##
## Coefficients:
##                                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)                         -8.0310     2.4306  -3.304  0.00147 **
## povertyDataset$Unemployment          1.7258     0.2441   7.069  7.3e-10 ***
## povertyDataset$Below.Poverty.Level   0.1548     0.1493   1.036  0.30334
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 9.659 on 74 degrees of freedom
## Multiple R-squared:  0.6688, Adjusted R-squared:  0.6598
## F-statistic:  74.7 on 2 and 74 DF,  p-value: < 2.2e-16
```

Nakon izrade više modela sa varijablom nezaposlenost zaključili smo da nezaposlenost sama po sebi već jako dobro objašnjava podatke te kombiniranjem još s No High School DIploma, Dependency i Per Capita income dobivamo najbolje rezultate.

Kada smo uključili i logaritmom transformirani Per Capita Income i samo Per Capita Income dobili bolji model nego samo s logaritmom transformiranim. Nismo sigurni zašto je to tako.

```
fit2 <- lm(povertyDataset$Assault..Homicide.~povertyDataset$Unemployment +
             povertyDataset$No.High.School.Diploma
           + log(povertyDataset$Per.Capita.Income)
           + povertyDataset$Per.Capita.Income + exp(povertyDataset$Dependency))
summary(fit2)
```

```
##
## Call:
## lm(formula = povertyDataset$Assault..Homicide. ~ povertyDataset$Unemployment +
##     povertyDataset$No.High.School.Diploma + log(povertyDataset$Per.Capita.Income) +
##     povertyDataset$Per.Capita.Income + exp(povertyDataset$Dependency))
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -16.3658  -4.5086  -0.1477   3.5721  21.5639
```

```
##
## Coefficients:
##                                        Estimate Std. Error t value Pr(>|t|)
## (Intercept)                           7.173e+02  1.078e+02   6.656 4.99e-09
## povertyDataset$Unemployment           5.956e-01  2.399e-01   2.483   0.0154
## povertyDataset$No.High.School.Diploma -1.152e+00  1.615e-01  -7.135 6.64e-10
## log(povertyDataset$Per.Capita.Income) -7.097e+01  1.077e+01  -6.589 6.59e-09
## povertyDataset$Per.Capita.Income       1.106e-03  2.312e-04   4.783 9.07e-06
## exp(povertyDataset$Dependency)        -6.109e-21  1.368e-21  -4.467 2.92e-05
##
## (Intercept)                           ***
## povertyDataset$Unemployment           *
## povertyDataset$No.High.School.Diploma ***
## log(povertyDataset$Per.Capita.Income) ***
## povertyDataset$Per.Capita.Income      ***
## exp(povertyDataset$Dependency)        ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 7.423 on 71 degrees of freedom
## Multiple R-squared:  0.8123, Adjusted R-squared:  0.7991
## F-statistic: 61.45 on 5 and 71 DF,  p-value: < 2.2e-16
```