Lab3_YZ_EDA

Yulia Zamriy

March 18, 2018

```
#install.packages("kableExtra")
#install.packages("viridisLite")
#install.packages("viridis")
#install.packages("Hmisc")
library(knitr)
library(kableExtra)
library(Hmisc)
```

```
## Loading required package: lattice
```

```
## Loading required package: survival
```

```
## Loading required package: Formula
```

```
## Loading required package: ggplot2
```

```
##
## Attaching package: 'Hmisc'
```

```
## The following objects are masked from 'package:base':
##
##     format.pval, units
```

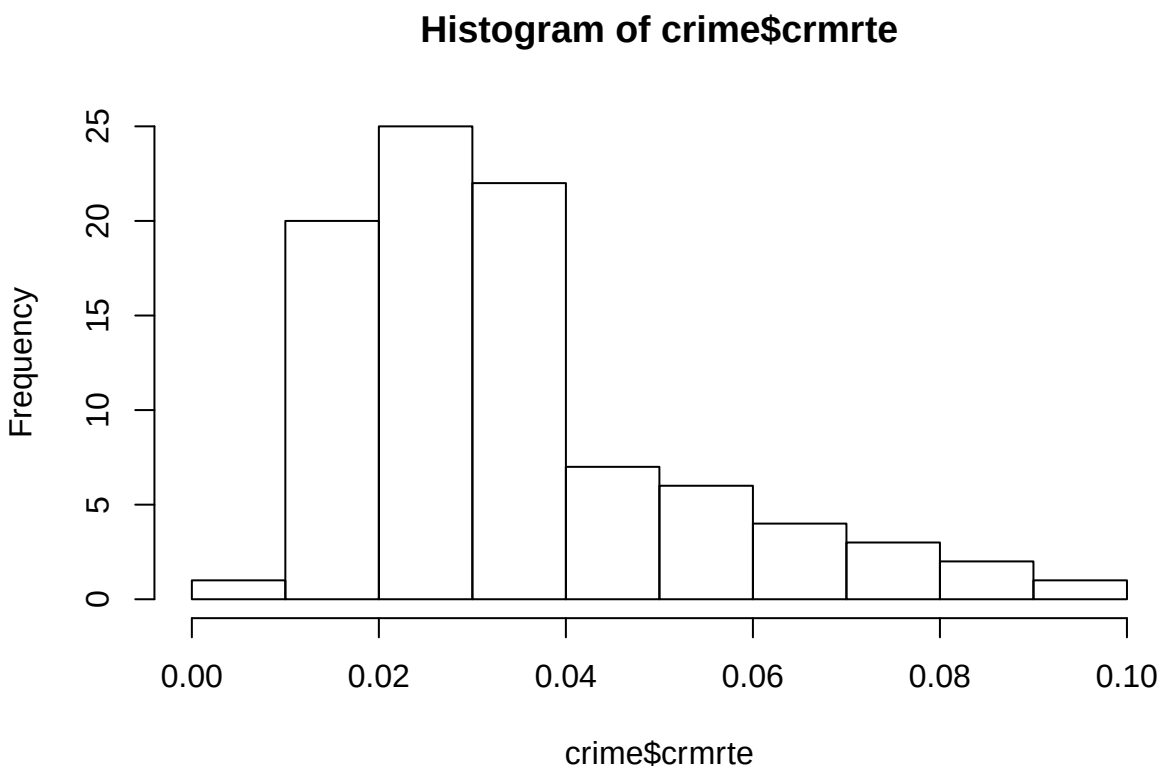
```
library(reshape2)
library(ggplot2)
```

```
#setwd("/home/yulia/Documents/MIDS/W203/Lab_3/")
crime <- read.csv("crime_v2.csv", stringsAsFactors = FALSE)
crime <- na.omit(crime)
```

```
summary(crime$crmrte)
```

```
##     Min.  1st Qu.   Median     Mean  3rd Qu.     Max.
## 0.005533 0.020927 0.029986 0.033400 0.039642 0.098966
```

```
hist(crime$crmrte)
```




```
crime$prbconv <- as.numeric(crime$prbconv)
```

```
summary(crime$prbarr)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.09277 0.20568 0.27095 0.29492 0.34438 1.09091
```

```
summary(crime$prbconv)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.06838 0.34541 0.45283 0.55128 0.58886 2.12121
```

```
summary(crime$prbpris)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.1500  0.3648  0.4234  0.4108  0.4568  0.6000
```

```
nrow(crime[crime$prbarr >= 1,])
```

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

```
nrow(crime[crime$prbconv >= 1,])
```

```
## [1] 10
```

```
crime$exclude <- 0
crime[crime$prbarr > 1,]$exclude <- 1
crime[crime$prbconv > 1,]$exclude <- 1
table(crime$exclude)
```

```
##
##  0  1
## 81 10
```

```
summary(crime$avgsen)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   5.380   7.340   9.100   9.647  11.420  20.700
```

```
summary(crime$polpc)
```

```
##      Min.   1st Qu.    Median      Mean   3rd Qu.      Max.
## 0.0007459 0.0012308 0.0014853 0.0017022 0.0018768 0.0090543
```

```
summary(crime$density)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.00002 0.54741 0.96226 1.42884 1.56824 8.82765
```

```
summary(crime$taxpc)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   25.69   30.66   34.87   38.06   40.95  119.76
```

```
mean(crime$west)
```

```
## [1] 0.2527473
```

```
mean(crime$central)
```

```
## [1] 0.3736264
```

```
mean(crime$urban)
```

```
## [1] 0.08791209
```

```
summary(crime$pctmin80)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   1.284   9.845  24.312  25.495  38.142  64.348
```

```
summary(crime$wcon)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   193.6   250.8   281.4   285.4   314.8   436.8
```

```
summary(crime$wtuc)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   187.6   374.6   406.5   411.7   443.4   613.2
```

```
summary(crime$wtrd)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   154.2   190.9   203.0   211.6   225.1   354.7
```

```
summary(crime$wfir)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   170.9   286.5   317.3   322.1   345.4   509.5
```

```
summary(crime$wser)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
```

```
##   133.0   229.7   253.2   275.6   280.5  2177.1
```

```
summary(crime$wmfg)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   157.4   288.9   320.2   335.6   359.6   646.9
```

```
summary(crime$wfed)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   326.1   400.2   449.8   442.9   478.0   598.0
```

```
summary(crime$wsta)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   258.3   329.3   357.7   357.5   382.6   499.6
```

```
summary(crime$wloc)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   239.2   297.3   308.1   312.7   329.2   388.1
```

```
summary(crime$mix)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.01961 0.08073 0.10186 0.12884 0.15175 0.46512
```

```
summary(crime$pctymle)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.06216 0.07443 0.07771 0.08396 0.08350 0.24871
```

```
crime[crime$wser > 2000,]$exclude <- 1
crime_sub <- subset(crime, exclude == 0)
crime_sub$exclude <- NULL
```

```
# Prepare a .RData for easier sharing and usage.
ind_variables <- c(
  'prbarr', 'prbconv', 'prbpris', 'avgsen',
  'polpc', 'density', 'taxpc', 'west', 'central', 'urban', 'pctmin80', 'wcon',
  'wtuc', 'wtrd', 'wfir', 'wser', 'wmfg', 'wfed', 'wsta', 'wloc', 'mix',
  'pctymle'
)
var_labels <- c(
  'probability of arrest', 'probability of conviction',
  'probability of prison sentence', 'avg. sentence, days',
  'police per capita', 'people per sq. mile', 'tax revenue per capita',
  '=1 if in western N.C.', '=1 if in central N.C.', '=1 if in SMSA',
  'perc. minority, 1980', 'weekly wage, construction',
  'wkly wge, trns, util, commun', 'wkly wge, whlesle, retail trade',
  'wkly wge, fin, ins, real est', 'wkly wge, service industry',
  'wkly wge, manufacturing', 'wkly wge, fed employees',
  'wkly wge, state employees', 'wkly wge, local gov emps',
  'offense mix: face-to-face/other', 'percent young male'
)
impact <- c("Negative" , "Negative", "Negative", "Negative",
            "Negative", "Positive", "Negative",
            "Unclear", "Unclear", "Unclear", "Unclear",
            "Negative","Negative","Negative",
```

```
            "Negative", "Negative", "Negative", "Negative",
            "Negative", "Negative", "Unclear","Positive")
control <- c("Yes", "Yes", "Yes", "Yes",
             "Yes", "No", "Yes",
             "No", "No", "No","No",
             "Yes", "Yes", "Yes",
             "Yes", "Yes", "Yes", "Yes",
             "Yes", "Yes", "No", "No")
desc <- data.frame(ind_variables, var_labels, impact, control)
colnames(desc) <- c("Explanatory Variables",
                    "Explanation",
                    "Expected Impact on Crime Rate",
                    "Can Gov Impact on This?")
# col_labels <-  c(ind_variables = "Explanatory Variables",
#                  var_labels = "Explanation")
# desc <- upData(desc, labels = col_labels)
```

```
kable(desc, booktabs = TRUE) %>%
  kable_styling(latex_options = c("scale_down"),
                full_width = FALSE) %>%
  row_spec(0, bold = TRUE) %>%
  column_spec(1, width = "8em") %>%
  column_spec(3, width = "10em") %>%
  column_spec(4, width = "9em")
```

Explanatory Variables	Explanation	Expected Impact on Crime Rate	Can Gov Impact on This?
prbarr	probability of arrest	Negative	Yes
prbconv	probability of conviction	Negative	Yes
prbpris	probability of prison sentence	Negative	Yes
avgsen	avg. sentence, days	Negative	Yes
polpc	police per capita	Negative	Yes
density	people per sq. mile	Positive	No
taxpc	tax revenue per capita	Negative	Yes
west	=1 if in western N.C.	Unclear	No
central	=1 if in central N.C.	Unclear	No
urban	=1 if in SMSA	Unclear	No
pctmin80	perc. minority, 1980	Unclear	No
wcon	weekly wage, construction	Negative	Yes
wtuc	wkly wge, trns, util, commun	Negative	Yes
wtrd	wkly wge, whlesle, retail trade	Negative	Yes
wfir	wkly wge, fin, ins, real est	Negative	Yes
wser	wkly wge, service industry	Negative	Yes
wmfg	wkly wge, manufacturing	Negative	Yes
wfed	wkly wge, fed employees	Negative	Yes
wsta	wkly wge, state employees	Negative	Yes
wloc	wkly wge, local gov emps	Negative	Yes
mix	offense mix: face-to-face/other	Unclear	No
pctymle	percent young male	Positive	No

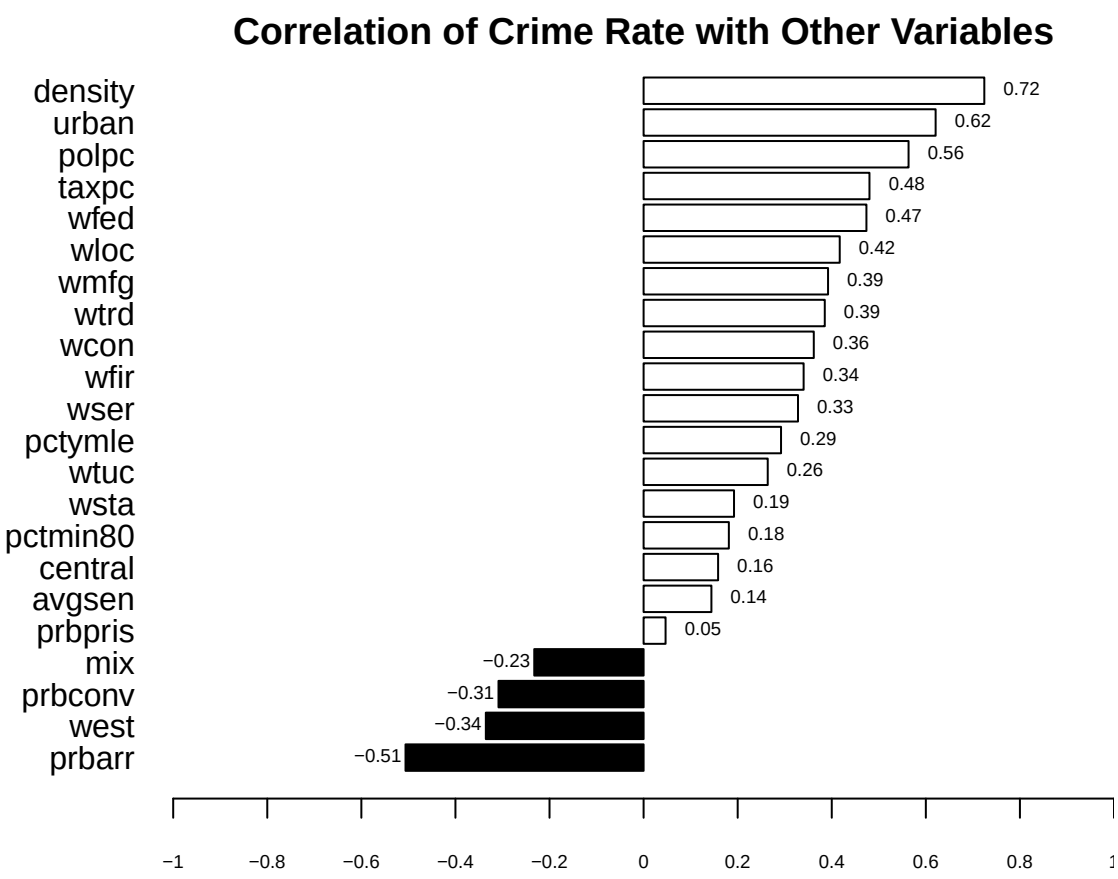
```r
crime_cor <- cor(crime_sub)[3,-c(1,2,3)]
```

```
## Warning in cor(crime_sub): the standard deviation is zero
```

```r
crime_cor <- crime_cor[order(crime_cor)]
negative <- ifelse(crime_cor < 0, 1,0)
```

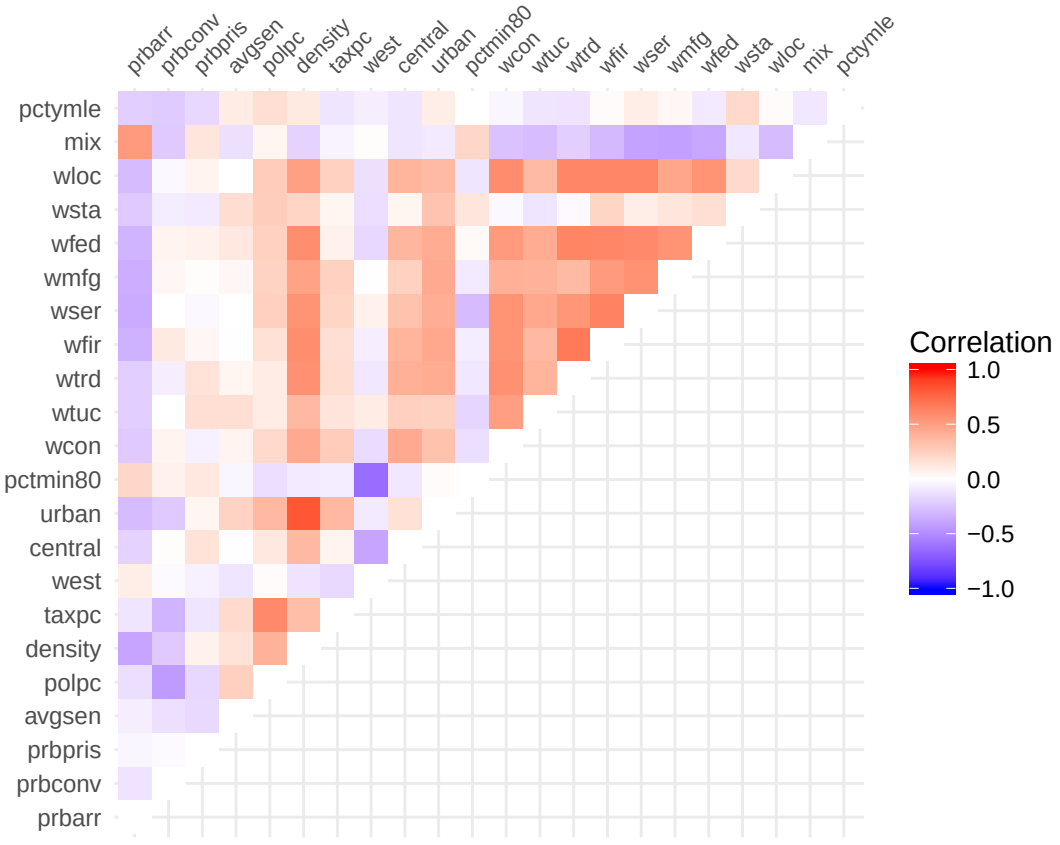
```r
crime_cor_lab <- ifelse(crime_cor < 0, crime_cor-0.15, crime_cor)

par(mar = c(2,8,1,0))
b <- barplot(crime_cor,
        col = negative,
        horiz = TRUE,
        las = 1,
        xaxt = "n",
        xlim = c(-1,1),
        main = "Correlation of Crime Rate with Other Variables")
text(x = crime_cor_lab,
     y = b,
     label = round(crime_cor,2),
     pos = 4,
     cex = 0.6)
axis(1,
     at = seq(-1,1, by = 0.2),
     labels = seq(-1,1, by = 0.2),
     cex.axis = 0.6)
```



```
cor_mat <- round(cor(crime_sub[-c(1:3)]),2)
get_upper_tri <- function(cor_mat){
    cor_mat[lower.tri(cor_mat)]<- NA
    return(cor_mat)
}
cor_mat_upper <- get_upper_tri(cor_mat)
cor_mat_upper2 <- melt(cor_mat_upper, na.rm = TRUE)
cor_mat_upper2[cor_mat_upper2$value == 1,]$value <- 0
```

```
ggplot(data = cor_mat_upper2, aes(Var1, Var2, fill = value)) +
  geom_tile() +
  scale_fill_gradient2(low = "blue", high = "red", mid = "white",
                       midpoint = 0, limit = c(-1,1), space = "Lab",
                       name = "Correlation") +
  theme_minimal() +
  scale_x_discrete(position = "top") +
  theme(axis.text.x = element_text(angle = 45, vjust = 1, size = 8, hjust = 0),
        axis.title.x=element_blank(),
        axis.title.y=element_blank()) +
  coord_fixed()
```




```
ind_vars_all <- c("prbarr", "prbconv", "prbpris", "avgsen", "polpc", "density", "taxpc",
                  "west", "central", "urban", "pctmin80", "wcon", "wtuc", "wtrd", "wfir",
                  "wser", "wmfg", "wfed", "wsta", "wloc", "mix", "pctymle")
ind_vars1 <- c("polpc", "taxpc","wfed","pctymle","avgsen")
```

```r
crmrte_formula1 <- as.formula(paste("crmrte ~", paste(ind_vars1, collapse = "+"), sep = ""))
crmrte_lm1 <- lm(crmrte_formula1, data = crime_sub)
summary(crmrte_lm1)
```

```
##
## Call:
## lm(formula = crmrte_formula1, data = crime_sub)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.022297 -0.007113 -0.001875  0.005518  0.041679
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept) -7.550e-02  1.305e-02  -5.788 1.56e-07 ***
## polpc        6.450e+00  3.543e+00   1.820 0.072698 .
## taxpc        5.527e-04  1.368e-04   4.040 0.000128 ***
## wfed         1.364e-04  2.394e-05   5.698 2.25e-07 ***
## pctymle      2.735e-01  6.222e-02   4.396 3.58e-05 ***
## avgsen      -4.702e-04  6.093e-04  -0.772 0.442709
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.01243 on 75 degrees of freedom
## Multiple R-squared:  0.5921, Adjusted R-squared:  0.5649
## F-statistic: 21.77 on 5 and 75 DF,  p-value: 2.116e-13
```

```r
crmrte_formula_all <- as.formula(paste("crmrte ~", paste(ind_vars_all, collapse = "+"), sep = ""))
crmrte_lm0 <- lm(crmrte ~ 1,
                 data = crime_sub)
crmrte_lm_all <- lm(crmrte_formula_all,
                 data = crime_sub)
crmrte_lm_step <- step(crmrte_lm0, scope=list(lower=crmrte_lm0, upper=crmrte_lm_all), direction="both")
```

```
## Start:  AIC=-642.35
## crmrte ~ 1
##
##            Df Sum of Sq      RSS     AIC
## + density   1 0.0149124 0.013512 -700.59
## + urban     1 0.0109548 0.017470 -679.78
## + polpc     1 0.0090180 0.019407 -671.26
## + prbarr    1 0.0072819 0.021143 -664.32
## + taxpc     1 0.0065539 0.021871 -661.58
## + wfed      1 0.0063783 0.022046 -660.93
## + wloc      1 0.0049416 0.023483 -655.82
## + wmfg      1 0.0043708 0.024054 -653.87
## + wtrd      1 0.0042168 0.024208 -653.36
## + wcon      1 0.0037203 0.024704 -651.71
## + wfir      1 0.0032896 0.025135 -650.31
## + west      1 0.0031966 0.025228 -650.01
## + wser      1 0.0030620 0.025363 -649.58
## + prbconv   1 0.0027045 0.025720 -648.45
## + pctymle   1 0.0024263 0.025998 -647.58
## + wtuc      1 0.0019795 0.026445 -646.20
```

```
## + mix       1 0.0015357 0.026889 -644.85
## + wsta      1 0.0010504 0.027374 -643.40
## + pctmin80  1 0.0009339 0.027491 -643.06
## + central   1 0.0007132 0.027711 -642.41
## <none>                  0.028425 -642.35
## + avgsen    1 0.0005906 0.027834 -642.05
## + prbpris   1 0.0000620 0.028363 -640.53
##
## Step:  AIC=-700.59
## crmrte ~ density
##
##            Df Sum of Sq      RSS     AIC
## + polpc     1 0.0025162 0.010996 -715.28
## + taxpc     1 0.0017529 0.011759 -709.84
## + west      1 0.0017479 0.011764 -709.81
## + pctmin80  1 0.0016935 0.011819 -709.43
## + prbarr    1 0.0016749 0.011837 -709.31
## + pctymle   1 0.0012898 0.012222 -706.71
## + prbconv   1 0.0005883 0.012924 -702.19
## + central   1 0.0003696 0.013143 -700.83
## <none>                  0.013512 -700.59
## + wfir      1 0.0002716 0.013241 -700.23
## + mix       1 0.0002641 0.013248 -700.19
## + wser      1 0.0002014 0.013311 -699.80
## + wloc      1 0.0001459 0.013366 -699.47
## + wfed      1 0.0001150 0.013397 -699.28
## + wmfg      1 0.0000737 0.013438 -699.03
## + wcon      1 0.0000624 0.013450 -698.96
## + urban     1 0.0000565 0.013456 -698.93
## + wtrd      1 0.0000415 0.013471 -698.84
## + avgsen    1 0.0000402 0.013472 -698.83
## + wsta      1 0.0000369 0.013475 -698.81
## + prbpris   1 0.0000014 0.013511 -698.60
## + wtuc      1 0.0000003 0.013512 -698.59
## - density   1 0.0149124 0.028425 -642.35
##
## Step:  AIC=-715.28
## crmrte ~ density + polpc
##
##            Df Sum of Sq       RSS     AIC
## + pctmin80  1 0.0022066 0.0087894 -731.42
## + west      1 0.0020796 0.0089164 -730.26
## + prbarr    1 0.0017331 0.0092629 -727.17
## + pctymle   1 0.0008602 0.0101358 -719.88
## + mix       1 0.0005616 0.0104344 -717.52
## + taxpc     1 0.0003196 0.0106764 -715.67
## + central   1 0.0003188 0.0106772 -715.66
## <none>                  0.0109960 -715.28
## + wser      1 0.0002604 0.0107356 -715.22
## + wfir      1 0.0001295 0.0108665 -714.24
## + wfed      1 0.0001124 0.0108836 -714.11
## + prbpris   1 0.0000999 0.0108961 -714.02
## + wloc      1 0.0000587 0.0109373 -713.71
## + wcon      1 0.0000451 0.0109509 -713.61
```

```
## + wmfg      1 0.0000360 0.0109599 -713.54
## + prbconv   1 0.0000308 0.0109652 -713.51
## + avgsen    1 0.0000169 0.0109791 -713.40
## + urban     1 0.0000134 0.0109826 -713.38
## + wsta      1 0.0000134 0.0109826 -713.38
## + wtrd      1 0.0000071 0.0109889 -713.33
## + wtuc      1 0.0000042 0.0109917 -713.31
## - polpc     1 0.0025162 0.0135122 -700.59
## - density   1 0.0084106 0.0194066 -671.26
##
## Step:  AIC=-731.42
## crmrte ~ density + polpc + pctmin80
##
##            Df Sum of Sq       RSS     AIC
## + prbarr    1 0.0026529 0.0061365 -758.52
## + mix       1 0.0012220 0.0075674 -741.55
## + pctymle   1 0.0008047 0.0079847 -737.20
## + west      1 0.0003953 0.0083941 -733.15
## + taxpc     1 0.0003105 0.0084789 -732.33
## <none>                  0.0087894 -731.42
## + central   1 0.0002093 0.0085801 -731.37
## + wcon      1 0.0001509 0.0086385 -730.82
## + wsta      1 0.0001454 0.0086440 -730.77
## + wloc      1 0.0001198 0.0086696 -730.53
## + wtuc      1 0.0001032 0.0086862 -730.38
## + wfir      1 0.0000774 0.0087120 -730.14
## + wmfg      1 0.0000705 0.0087188 -730.07
## + wtrd      1 0.0000479 0.0087414 -729.86
## + prbconv   1 0.0000397 0.0087497 -729.79
## + wfed      1 0.0000376 0.0087518 -729.77
## + prbpris   1 0.0000254 0.0087640 -729.66
## + urban     1 0.0000161 0.0087733 -729.57
## + avgsen    1 0.0000136 0.0087758 -729.55
## + wser      1 0.0000080 0.0087814 -729.49
## - pctmin80  1 0.0022066 0.0109960 -715.28
## - polpc     1 0.0030293 0.0118186 -709.43
## - density   1 0.0087003 0.0174897 -677.69
##
## Step:  AIC=-758.52
## crmrte ~ density + polpc + pctmin80 + prbarr
##
##            Df Sum of Sq       RSS     AIC
## + wsta      1 0.0005495 0.0055871 -764.12
## + prbconv   1 0.0004111 0.0057254 -762.14
## + taxpc     1 0.0003642 0.0057723 -761.48
## + pctymle   1 0.0003575 0.0057790 -761.39
## + central   1 0.0002684 0.0058681 -760.15
## + wfir      1 0.0002325 0.0059040 -759.65
## <none>                  0.0061365 -758.52
## + mix       1 0.0001459 0.0059906 -758.47
## + wser      1 0.0001077 0.0060288 -757.96
## + wcon      1 0.0000977 0.0060389 -757.82
## + wtrd      1 0.0000848 0.0060517 -757.65
## + west      1 0.0000771 0.0060594 -757.55
```

```
## + wtuc      1 0.0000612 0.0060753 -757.34
## + wloc      1 0.0000259 0.0061106 -756.87
## + avgsen    1 0.0000219 0.0061146 -756.81
## + prbpris   1 0.0000131 0.0061235 -756.70
## + urban     1 0.0000044 0.0061321 -756.58
## + wmfg      1 0.0000027 0.0061338 -756.56
## + wfed      1 0.0000025 0.0061340 -756.56
## - prbarr    1 0.0026529 0.0087894 -731.42
## - pctmin80  1 0.0031264 0.0092629 -727.17
## - polpc     1 0.0032299 0.0093664 -726.27
## - density   1 0.0046085 0.0107450 -715.15
##
## Step:  AIC=-764.12
## crmrte ~ density + polpc + pctmin80 + prbarr + wsta
##
##            Df Sum of Sq       RSS     AIC
## + pctymle   1 0.0004877 0.0050993 -769.52
## + prbconv   1 0.0004208 0.0051663 -768.46
## + central   1 0.0002845 0.0053025 -766.36
## + taxpc     1 0.0002424 0.0053447 -765.71
## + mix       1 0.0001704 0.0054166 -764.63
## + wfir      1 0.0001480 0.0054390 -764.30
## <none>                  0.0055871 -764.12
## + wser      1 0.0001209 0.0054662 -763.89
## + west      1 0.0000701 0.0055169 -763.15
## + wloc      1 0.0000526 0.0055345 -762.89
## + wcon      1 0.0000412 0.0055459 -762.72
## + wtrd      1 0.0000383 0.0055488 -762.68
## + wtuc      1 0.0000115 0.0055755 -762.29
## + urban     1 0.0000097 0.0055774 -762.26
## + avgsen    1 0.0000033 0.0055837 -762.17
## + wmfg      1 0.0000032 0.0055838 -762.17
## + wfed      1 0.0000020 0.0055850 -762.15
## + prbpris   1 0.0000017 0.0055853 -762.15
## - wsta      1 0.0005495 0.0061365 -758.52
## - prbarr    1 0.0030569 0.0086440 -730.77
## - pctmin80  1 0.0035567 0.0091438 -726.22
## - polpc     1 0.0036777 0.0092647 -725.16
## - density   1 0.0047594 0.0103465 -716.21
##
## Step:  AIC=-769.52
## crmrte ~ density + polpc + pctmin80 + prbarr + wsta + pctymle
##
##            Df Sum of Sq       RSS     AIC
## + taxpc     1 0.0004804 0.0046189 -775.54
## + prbconv   1 0.0002573 0.0048420 -771.71
## + central   1 0.0001756 0.0049237 -770.36
## + mix       1 0.0001619 0.0049374 -770.13
## + wser      1 0.0001249 0.0049744 -769.53
## <none>                  0.0050993 -769.52
## + wtrd      1 0.0001076 0.0049918 -769.25
## + wfir      1 0.0001047 0.0049946 -769.20
## + wloc      1 0.0000827 0.0050167 -768.84
## + wcon      1 0.0000744 0.0050249 -768.71
```

```
## + west      1 0.0000599 0.0050394 -768.48
## + wtuc      1 0.0000470 0.0050524 -768.27
## + prbpris   1 0.0000239 0.0050754 -767.90
## + urban     1 0.0000156 0.0050837 -767.77
## + wfed      1 0.0000143 0.0050850 -767.75
## + avgsen    1 0.0000077 0.0050916 -767.64
## + wmfg      1 0.0000003 0.0050990 -767.53
## - pctymle   1 0.0004877 0.0055871 -764.12
## - wsta      1 0.0006797 0.0057790 -761.39
## - prbarr    1 0.0025997 0.0076990 -738.15
## - polpc     1 0.0033241 0.0084234 -730.87
## - pctmin80  1 0.0034778 0.0085772 -729.40
## - density   1 0.0048519 0.0099512 -717.37
##
## Step:  AIC=-775.54
## crmrte ~ density + polpc + pctmin80 + prbarr + wsta + pctymle +
##     taxpc
##
##            Df Sum of Sq       RSS     AIC
## + prbconv   1 0.0001729 0.0044460 -776.63
## + wser      1 0.0001433 0.0044757 -776.09
## + wfir      1 0.0001170 0.0045019 -775.61
## + central   1 0.0001137 0.0045053 -775.55
## <none>                  0.0046189 -775.54
## + wfed      1 0.0001080 0.0045110 -775.45
## + mix       1 0.0001038 0.0045151 -775.38
## + wtrd      1 0.0000931 0.0045258 -775.18
## + wloc      1 0.0000619 0.0045570 -774.63
## + wtuc      1 0.0000521 0.0045668 -774.45
## + prbpris   1 0.0000491 0.0045698 -774.40
## + wcon      1 0.0000389 0.0045800 -774.22
## + avgsen    1 0.0000200 0.0045989 -773.89
## + wmfg      1 0.0000050 0.0046139 -773.62
## + west      1 0.0000031 0.0046158 -773.59
## + urban     1 0.0000000 0.0046189 -773.54
## - taxpc     1 0.0004804 0.0050993 -769.52
## - wsta      1 0.0005294 0.0051483 -768.75
## - pctymle   1 0.0007257 0.0053447 -765.71
## - polpc     1 0.0011968 0.0058157 -758.87
## - prbarr    1 0.0024863 0.0071053 -742.65
## - pctmin80  1 0.0033610 0.0079800 -733.25
## - density   1 0.0043224 0.0089414 -724.03
##
## Step:  AIC=-776.63
## crmrte ~ density + polpc + pctmin80 + prbarr + wsta + pctymle +
##     taxpc + prbconv
##
##            Df Sum of Sq       RSS     AIC
## + mix       1 0.0001685 0.0042776 -777.75
## + wfed      1 0.0001633 0.0042827 -777.66
## <none>                  0.0044460 -776.63
## + central   1 0.0001013 0.0043447 -776.49
## + wser      1 0.0000962 0.0043498 -776.40
## + wtrd      1 0.0000906 0.0043554 -776.29
```

```
## + wloc      1 0.0000895 0.0043566 -776.27
## + wcon      1 0.0000866 0.0043594 -776.22
## + wtuc      1 0.0000597 0.0043863 -775.72
## + wfir      1 0.0000563 0.0043897 -775.66
## - prbconv   1 0.0001729 0.0046189 -775.54
## + prbpris   1 0.0000237 0.0044224 -775.06
## + avgsen    1 0.0000213 0.0044248 -775.01
## + west      1 0.0000030 0.0044430 -774.68
## + urban     1 0.0000001 0.0044460 -774.63
## + wmfg      1 0.0000000 0.0044460 -774.63
## - taxpc     1 0.0003960 0.0048420 -771.71
## - pctymle   1 0.0005216 0.0049676 -769.64
## - wsta      1 0.0005290 0.0049750 -769.52
## - polpc     1 0.0009407 0.0053867 -763.08
## - prbarr    1 0.0026539 0.0071000 -740.71
## - pctmin80  1 0.0034644 0.0079104 -731.96
## - density   1 0.0039593 0.0084053 -727.04
##
## Step:  AIC=-777.75
## crmrte ~ density + polpc + pctmin80 + prbarr + wsta + pctymle +
##     taxpc + prbconv + mix
##
##            Df Sum of Sq       RSS     AIC
## + wser      1 0.0001917 0.0040858 -779.47
## + central   1 0.0001073 0.0041703 -777.81
## <none>                  0.0042776 -777.75
## + wfir      1 0.0000884 0.0041891 -777.45
## + wfed      1 0.0000687 0.0042089 -777.07
## + wtrd      1 0.0000634 0.0042142 -776.96
## + wcon      1 0.0000551 0.0042224 -776.80
## + prbpris   1 0.0000517 0.0042259 -776.74
## + wloc      1 0.0000504 0.0042272 -776.71
## - mix       1 0.0001685 0.0044460 -776.63
## + avgsen    1 0.0000434 0.0042342 -776.58
## + wtuc      1 0.0000261 0.0042514 -776.25
## + wmfg      1 0.0000152 0.0042623 -776.04
## + west      1 0.0000014 0.0042762 -775.78
## + urban     1 0.0000007 0.0042769 -775.77
## - prbconv   1 0.0002375 0.0045151 -775.38
## - taxpc     1 0.0003162 0.0045938 -773.98
## - pctymle   1 0.0004636 0.0047411 -771.42
## - wsta      1 0.0005608 0.0048384 -769.78
## - polpc     1 0.0010517 0.0053292 -761.95
## - prbarr    1 0.0017791 0.0060567 -751.58
## - pctmin80  1 0.0036278 0.0079053 -730.01
## - density   1 0.0038399 0.0081175 -727.86
##
## Step:  AIC=-779.47
## crmrte ~ density + polpc + pctmin80 + prbarr + wsta + pctymle +
##     taxpc + prbconv + mix + wser
##
##            Df Sum of Sq       RSS     AIC
## + wfed      1 0.0002446 0.0038412 -782.47
## + wloc      1 0.0002093 0.0038766 -781.73
```

```
## + wtrd      1 0.0001747 0.0039111 -781.01
## + wcon      1 0.0001733 0.0039125 -780.98
## <none>                  0.0040858 -779.47
## + avgsen    1 0.0000816 0.0040042 -779.10
## + wtuc      1 0.0000765 0.0040094 -779.00
## + central   1 0.0000736 0.0040122 -778.94
## + prbpris   1 0.0000556 0.0040303 -778.58
## - prbconv   1 0.0001832 0.0042690 -777.92
## + wfir      1 0.0000154 0.0040704 -777.77
## - wser      1 0.0001917 0.0042776 -777.75
## + urban     1 0.0000011 0.0040847 -777.49
## + west      1 0.0000005 0.0040853 -777.48
## + wmfg      1 0.0000003 0.0040855 -777.48
## - mix       1 0.0002640 0.0043498 -776.40
## - taxpc     1 0.0003266 0.0044124 -775.24
## - pctymle   1 0.0004852 0.0045711 -772.38
## - wsta      1 0.0005858 0.0046716 -770.62
## - polpc     1 0.0011196 0.0052054 -761.85
## - prbarr    1 0.0017431 0.0058289 -752.69
## - pctmin80  1 0.0030857 0.0071715 -735.90
## - density   1 0.0037412 0.0078271 -728.81
##
## Step:  AIC=-782.47
## crmrte ~ density + polpc + pctmin80 + prbarr + wsta + pctymle +
##     taxpc + prbconv + mix + wser + wfed
##
##            Df  Sum of Sq       RSS     AIC
## + wloc      1 0.00013675 0.0037045 -783.41
## + central   1 0.00011475 0.0037265 -782.93
## + avgsen    1 0.00011288 0.0037284 -782.89
## <none>                   0.0038412 -782.47
## + wcon      1 0.00008649 0.0037547 -782.31
## - mix       1 0.00010872 0.0039500 -782.21
## + wfir      1 0.00006923 0.0037720 -781.94
## + wtrd      1 0.00006162 0.0037796 -781.78
## + prbpris   1 0.00004477 0.0037965 -781.42
## + wtuc      1 0.00004237 0.0037989 -781.37
## + wmfg      1 0.00000688 0.0038344 -780.61
## + urban     1 0.00000493 0.0038363 -780.57
## + west      1 0.00000461 0.0038366 -780.57
## - prbconv   1 0.00020362 0.0040449 -780.29
## - wfed      1 0.00024460 0.0040858 -779.47
## - wser      1 0.00036763 0.0042089 -777.07
## - taxpc     1 0.00051158 0.0043528 -774.34
## - wsta      1 0.00057915 0.0044204 -773.09
## - pctymle   1 0.00068422 0.0045255 -771.19
## - polpc     1 0.00072116 0.0045624 -770.53
## - prbarr    1 0.00175426 0.0055955 -754.00
## - pctmin80  1 0.00224434 0.0060856 -747.20
## - density   1 0.00242465 0.0062659 -744.83
##
## Step:  AIC=-783.41
## crmrte ~ density + polpc + pctmin80 + prbarr + wsta + pctymle +
##     taxpc + prbconv + mix + wser + wfed + wloc
```

```
##
##             Df  Sum of Sq       RSS     AIC
## + central    1 0.00017964 0.0035248 -785.43
## + wfir       1 0.00014102 0.0035635 -784.55
## <none>                    0.0037045 -783.41
## + avgsen     1 0.00008785 0.0036166 -783.35
## - mix        1 0.00010513 0.0038096 -783.14
## - wloc       1 0.00013675 0.0038412 -782.47
## + wcon       1 0.00003243 0.0036721 -782.12
## + wtuc       1 0.00003157 0.0036729 -782.10
## + prbpris    1 0.00002876 0.0036757 -782.04
## - wfed       1 0.00017208 0.0038766 -781.73
## + west       1 0.00001439 0.0036901 -781.72
## + wtrd       1 0.00001357 0.0036909 -781.70
## + urban      1 0.00001258 0.0036919 -781.68
## + wmfg       1 0.00000991 0.0036946 -781.62
## - prbconv    1 0.00020938 0.0039139 -780.95
## - taxpc      1 0.00044978 0.0041543 -776.12
## - wser       1 0.00048360 0.0041881 -775.47
## - wsta       1 0.00064506 0.0043495 -772.40
## - pctymle    1 0.00068652 0.0043910 -771.63
## - polpc      1 0.00072524 0.0044297 -770.92
## - prbarr     1 0.00175711 0.0054616 -753.96
## - pctmin80   1 0.00224701 0.0059515 -747.00
## - density    1 0.00232887 0.0060333 -745.90
##
## Step:  AIC=-785.43
## crmrte ~ density + polpc + pctmin80 + prbarr + wsta + pctymle +
##     taxpc + prbconv + mix + wser + wfed + wloc + central
##
##             Df  Sum of Sq       RSS     AIC
## + wfir       1 0.00011178 0.0034131 -786.04
## + avgsen     1 0.00009496 0.0034299 -785.64
## <none>                    0.0035248 -785.43
## - mix        1 0.00008880 0.0036136 -785.42
## + wcon       1 0.00008385 0.0034410 -785.38
## + prbpris    1 0.00004286 0.0034820 -784.42
## + wtuc       1 0.00003498 0.0034899 -784.24
## + west       1 0.00003474 0.0034901 -784.23
## + wtrd       1 0.00002683 0.0034980 -784.05
## + wmfg       1 0.00001160 0.0035132 -783.70
## + urban      1 0.00000012 0.0035247 -783.43
## - central    1 0.00017964 0.0037045 -783.41
## - wloc       1 0.00020164 0.0037265 -782.93
## - prbconv    1 0.00020376 0.0037286 -782.88
## - wfed       1 0.00020418 0.0037290 -782.87
## - taxpc      1 0.00039804 0.0039229 -778.77
## - wser       1 0.00048695 0.0040118 -776.95
## - pctymle    1 0.00058311 0.0041080 -775.03
## - wsta       1 0.00066960 0.0041944 -773.34
## - polpc      1 0.00074103 0.0042659 -771.98
## - prbarr     1 0.00181156 0.0053364 -753.84
## - pctmin80   1 0.00216312 0.0056880 -748.67
## - density    1 0.00247660 0.0060014 -744.33
```

```
##
## Step:  AIC=-786.04
## crmrte ~ density + polpc + pctmin80 + prbarr + wsta + pctymle +
##     taxpc + prbconv + mix + wser + wfed + wloc + central + wfir
##
##            Df  Sum of Sq       RSS     AIC
## + avgsen    1 0.00011424 0.0032988 -786.80
## + wcon      1 0.00010948 0.0033036 -786.68
## - mix       1 0.00007612 0.0034892 -786.26
## + wtrd      1 0.00008600 0.0033271 -786.11
## <none>                   0.0034131 -786.04
## - wfir      1 0.00011178 0.0035248 -785.43
## + prbpris   1 0.00004255 0.0033705 -785.06
## + wtuc      1 0.00004159 0.0033715 -785.04
## - prbconv   1 0.00013696 0.0035500 -784.86
## - central   1 0.00015040 0.0035635 -784.55
## + west      1 0.00001594 0.0033971 -784.42
## + wmfg      1 0.00000532 0.0034077 -784.17
## + urban     1 0.00000003 0.0034130 -784.04
## - wfed      1 0.00026217 0.0036752 -782.05
## - wloc      1 0.00026736 0.0036804 -781.93
## - wser      1 0.00037024 0.0037833 -779.70
## - taxpc     1 0.00044536 0.0038584 -778.11
## - wsta      1 0.00056625 0.0039793 -775.61
## - pctymle   1 0.00061730 0.0040304 -774.58
## - polpc     1 0.00063456 0.0040476 -774.23
## - prbarr    1 0.00178180 0.0051949 -754.02
## - pctmin80  1 0.00211698 0.0055300 -748.95
## - density   1 0.00258540 0.0059985 -742.37
##
## Step:  AIC=-786.8
## crmrte ~ density + polpc + pctmin80 + prbarr + wsta + pctymle +
##     taxpc + prbconv + mix + wser + wfed + wloc + central + wfir +
##     avgsen
##
##            Df  Sum of Sq       RSS     AIC
## + wcon      1 0.00012802 0.0031708 -788.01
## + wtrd      1 0.00010877 0.0031901 -787.52
## <none>                   0.0032988 -786.80
## + wtuc      1 0.00007849 0.0032203 -786.75
## - mix       1 0.00010350 0.0034023 -786.30
## - avgsen    1 0.00011424 0.0034131 -786.04
## + west      1 0.00003680 0.0032620 -785.71
## - wfir      1 0.00013106 0.0034299 -785.64
## + prbpris   1 0.00003217 0.0032667 -785.59
## - prbconv   1 0.00013745 0.0034363 -785.49
## - central   1 0.00015530 0.0034541 -785.07
## + wmfg      1 0.00000862 0.0032902 -785.01
## + urban     1 0.00000406 0.0032948 -784.90
## - wloc      1 0.00023823 0.0035370 -783.15
## - wfed      1 0.00030601 0.0036048 -781.61
## - wser      1 0.00041757 0.0037164 -779.15
## - wsta      1 0.00048634 0.0037852 -777.66
## - taxpc     1 0.00049065 0.0037895 -777.57
```

```
## - pctymle   1 0.00065820 0.0039570 -774.06
## - polpc     1 0.00067429 0.0039731 -773.73
## - prbarr    1 0.00170045 0.0049993 -755.13
## - pctmin80  1 0.00202229 0.0053211 -750.07
## - density   1 0.00265432 0.0059531 -740.98
##
## Step:  AIC=-788.01
## crmrte ~ density + polpc + pctmin80 + prbarr + wsta + pctymle +
##     taxpc + prbconv + mix + wser + wfed + wloc + central + wfir +
##     avgsen + wcon
##
##            Df  Sum of Sq       RSS     AIC
## + wtrd      1 0.00008550 0.0030853 -788.22
## <none>                   0.0031708 -788.01
## + prbpris   1 0.00005727 0.0031135 -787.48
## - mix       1 0.00011075 0.0032816 -787.22
## + wtuc      1 0.00003982 0.0031310 -787.03
## - wcon      1 0.00012802 0.0032988 -786.80
## + west      1 0.00002897 0.0031418 -786.75
## - avgsen    1 0.00013278 0.0033036 -786.68
## - wloc      1 0.00013953 0.0033103 -786.52
## + wmfg      1 0.00000781 0.0031630 -786.21
## + urban     1 0.00000201 0.0031688 -786.06
## - wfir      1 0.00016267 0.0033335 -785.95
## - prbconv   1 0.00016954 0.0033403 -785.79
## - central   1 0.00021696 0.0033878 -784.64
## - wfed      1 0.00023706 0.0034079 -784.17
## - wsta      1 0.00036909 0.0035399 -781.09
## - taxpc     1 0.00038155 0.0035524 -780.80
## - wser      1 0.00046152 0.0036323 -779.00
## - pctymle   1 0.00060623 0.0037770 -775.83
## - polpc     1 0.00070240 0.0038732 -773.80
## - prbarr    1 0.00173281 0.0049036 -754.69
## - pctmin80  1 0.00207374 0.0052445 -749.25
## - density   1 0.00265946 0.0058303 -740.67
##
## Step:  AIC=-788.22
## crmrte ~ density + polpc + pctmin80 + prbarr + wsta + pctymle +
##     taxpc + prbconv + mix + wser + wfed + wloc + central + wfir +
##     avgsen + wcon + wtrd
##
##            Df  Sum of Sq       RSS     AIC
## <none>                   0.0030853 -788.22
## - wloc      1 0.00007981 0.0031651 -788.15
## - wtrd      1 0.00008550 0.0031708 -788.01
## - wcon      1 0.00010474 0.0031901 -787.52
## + prbpris   1 0.00004828 0.0030370 -787.50
## + wtuc      1 0.00004816 0.0030371 -787.49
## - prbconv   1 0.00011554 0.0032008 -787.24
## + west      1 0.00003212 0.0030532 -787.07
## - mix       1 0.00012948 0.0032148 -786.89
## - avgsen    1 0.00015250 0.0032378 -786.31
## + wmfg      1 0.00000165 0.0030837 -786.26
## + urban     1 0.00000128 0.0030840 -786.25
```

```
## - wfed      1 0.00016038 0.0032457 -786.11
## - wfir      1 0.00022806 0.0033134 -784.44
## - central   1 0.00023160 0.0033169 -784.36
## - wsta      1 0.00028053 0.0033658 -783.17
## - taxpc     1 0.00035592 0.0034412 -781.38
## - wser      1 0.00044293 0.0035282 -779.35
## - pctymle   1 0.00064360 0.0037289 -774.87
## - polpc     1 0.00078342 0.0038687 -771.89
## - prbarr    1 0.00174189 0.0048272 -753.96
## - pctmin80  1 0.00212547 0.0052108 -747.77
## - density   1 0.00239617 0.0054815 -743.67
```

```
summary(crmrte_lm_step)
```

```
##
## Call:
## lm(formula = crmrte ~ density + polpc + pctmin80 + prbarr + wsta +
##     pctymle + taxpc + prbconv + mix + wser + wfed + wloc + central +
##     wfir + avgsen + wcon + wtrd, data = crime_sub)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.013045 -0.004003 -0.001198  0.003880  0.018825
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  1.446e-02  1.542e-02   0.938 0.352050
## density      5.555e-03  7.941e-04   6.995 2.04e-09 ***
## polpc        9.407e+00  2.352e+00   4.000 0.000169 ***
## pctmin80     3.620e-04  5.495e-05   6.588 1.04e-08 ***
## prbarr      -5.667e-02  9.502e-03  -5.964 1.22e-07 ***
## wsta        -5.249e-05  2.193e-05  -2.393 0.019686 *
## pctymle      1.456e-01  4.015e-02   3.625 0.000579 ***
## taxpc        2.413e-04  8.949e-05   2.696 0.008993 **
## prbconv     -9.038e-03  5.884e-03  -1.536 0.129549
## mix         -2.121e-02  1.304e-02  -1.626 0.108938
## wser        -8.551e-05  2.843e-05  -3.007 0.003783 **
## wfed         4.192e-05  2.316e-05   1.810 0.075116 .
## wloc         5.446e-05  4.266e-05   1.277 0.206446
## central     -4.111e-03  1.890e-03  -2.175 0.033420 *
## wfir        -5.645e-05  2.616e-05  -2.158 0.034750 *
## avgsen      -6.384e-04  3.618e-04  -1.765 0.082474 .
## wcon         3.500e-05  2.393e-05   1.462 0.148584
## wtrd         5.199e-05  3.935e-05   1.321 0.191185
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.006998 on 63 degrees of freedom
## Multiple R-squared:  0.8915, Adjusted R-squared:  0.8622
## F-statistic: 30.44 on 17 and 63 DF,  p-value: < 2.2e-16
```