

Appendix B: Data Analysis

Casey DeLano and Jimmy DeLano

December 4, 2019

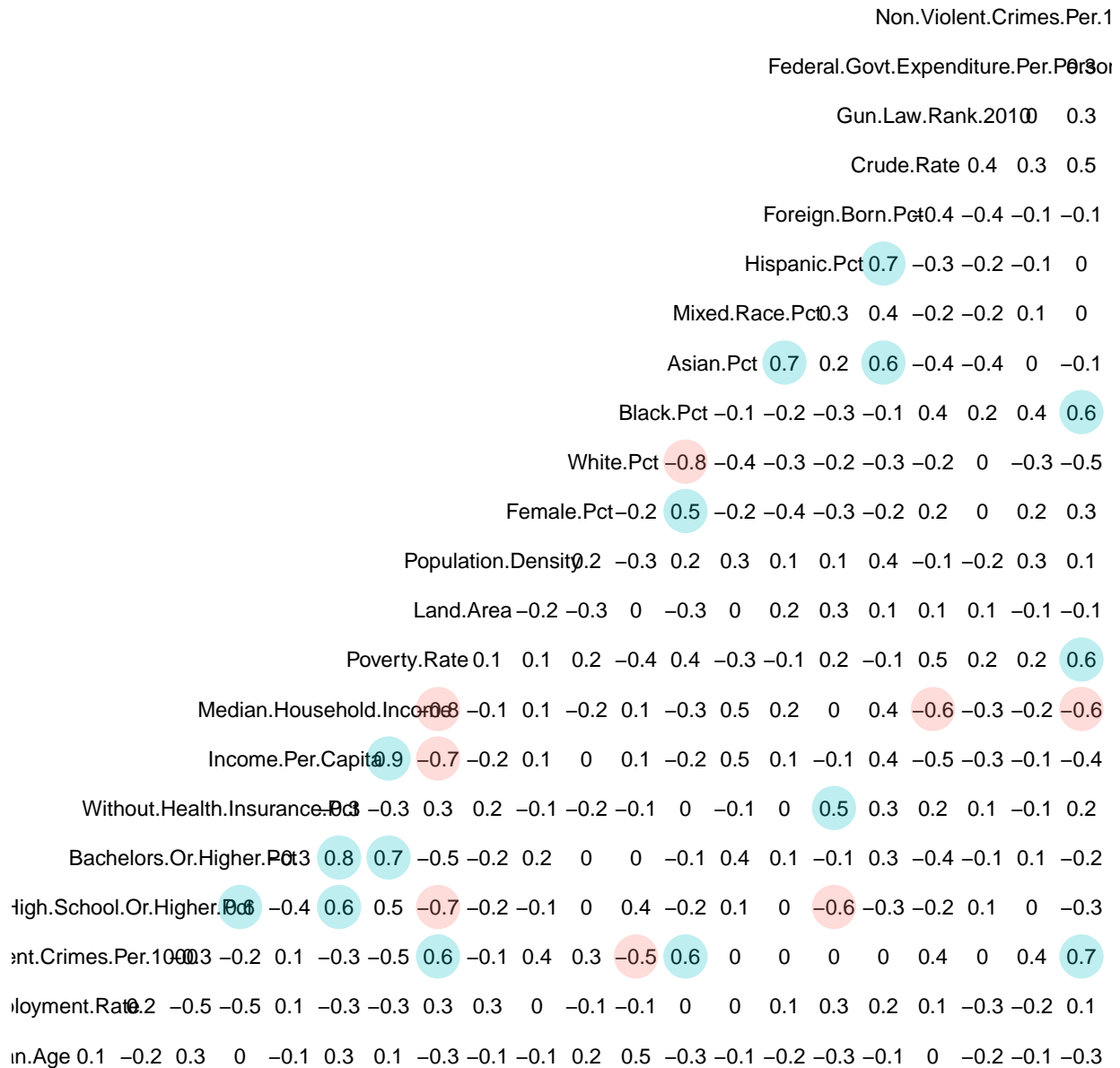
Step 0. Read in Data

```
#read in data
final <- read.csv("C:/Users/Casey DeLano/Desktop/STAT 346/final.csv")
attach(final)
```

Step 1. Data Exploration

We have too many variables to realistically create a scatter plot matrix that we can interpret. Instead, let's create a matrix of each variable's correlation with all of the other variables. We'll highlight correlations that are above 0.5 or below -0.5 so that we can easily locate potential issues.

```
subset = subset(final, select=-c(County, Population, Voter.Group.2008))
ggcorr(subset, geom = "blank", label = TRUE, hjust = 0.75) +
  geom_point(size = 10, aes(color = coefficient > 0, alpha = abs(coefficient) > 0.5)) +
  scale_alpha_manual(values = c("TRUE" = 0.25, "FALSE" = 0)) +
  guides(color = FALSE, alpha = FALSE)
```



This is a great way to visualize the colinearity. Let's use the variance inflation factor analysis more robustly determine which, if any variables, should be omitted.

```
vif0 = lm(Crude.Rate~., data=final) #regress on all variables
vif = update(vif0, .~-County-Population, data=final)
vif(vif)
```

##		GVIF	Df	GVIF^(1/(2*Df))
##	Median.Age	4.037992	1	2.009476
##	Unemployment.Rate	2.380761	1	1.542971
##	Violent.Crimes.Per.1000	3.376292	1	1.837469
##	High.School.Or.Higher.Pct	7.981172	1	2.825097
##	Bachelors.Or.Higher.Pct	10.451276	1	3.232843
##	Without.Health.Insurance.Pct	2.560025	1	1.600008
##	Income.Per.Capita	15.884940	1	3.985592
##	Median.Household.Income	13.986549	1	3.739859
##	Poverty.Rate	9.068891	1	3.011460
##	Land.Area	1.691588	1	1.300611
##	Population.Density	2.108346	1	1.452014
##	Female.Pct	2.362695	1	1.537106
##	White.Pct	20.422208	1	4.519094
##	Black.Pct	17.616782	1	4.197235
##	Asian.Pct	6.005307	1	2.450573
##	Mixed.Race.Pct	4.123775	1	2.030708
##	Hispanic.Pct	7.579778	1	2.753140
##	Foreign.Born.Pct	8.910203	1	2.984996
##	Gun.Law.Rank.2010	1.979908	1	1.407092
##	Federal.Govt.Expenditure.Per.Person	1.541478	1	1.241563
##	Non.Violent.Crimes.Per.1000	3.541806	1	1.881969
##	Voter.Group.2008	2.580552	2	1.267442

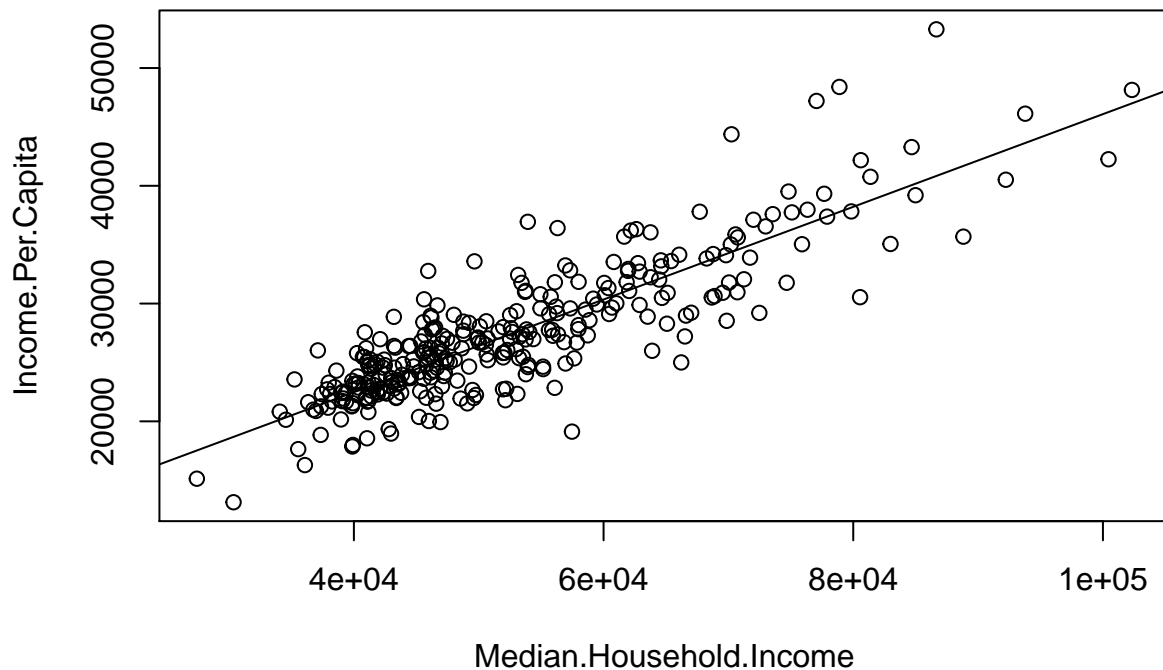
Notice that five variables have VIF values over 10: *Bachelors.Or.Higher.Pct*, *Income.Per.Capita*, *Median.Household.Income*, *White.Pct*, and *Black.Pct*.

Notice that *Income.Per.Capita* and *Median.Household.Income* are very correlated. Let's take a look at a plot of the two.

```
cor(Income.Per.Capita,Median.Household.Income)
```

```
## [1] 0.8602874
```

```
plot(Income.Per.Capita~Median.Household.Income)
abline(lm(Income.Per.Capita~Median.Household.Income))
```



The relationship between these variables is super linear. Recall from the correlation matrix that *Median.Household.Income* better predicts *Crude.Rate*. So let's omit *Income.Per.Capita* from future analysis.

```
cor(Median.Household.Income, Crude.Rate)
```

```
## [1] -0.5840673
```

```
cor(Income.Per.Capita, Crude.Rate)
```

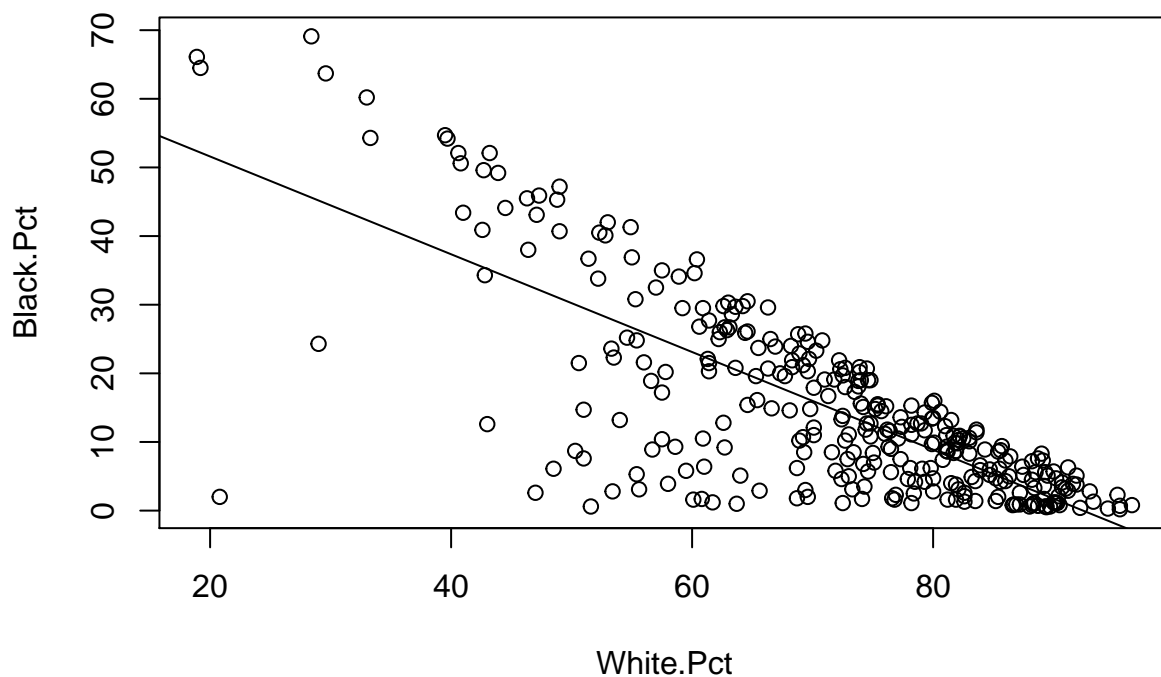
```
## [1] -0.4945271
```

Also from the correlation matrix, we notice that *White.Pct* and *Black.Pct* are also highly correlated. Let's repeat the above process for these variables.

```
cor(Black.Pct, White.Pct)
```

```
## [1] -0.7671012
```

```
plot(Black.Pct~White.Pct)
abline(lm(Black.Pct~White.Pct))
```



```
cor(Black.Pct, Crude.Rate)
```

```
## [1] 0.4322242
```

```
cor(White.Pct, Crude.Rate)
```

```
## [1] -0.2029136
```

Since *Black.Pct* is more highly predictive of *Crude.Rate* than *White.Pct*, let's drop *White.Pct* from future analysis.

Now let's re-run our VIF analysis without these two variables:

```
vif2.0 = lm(Crude.Rate~., data=final) #regress on all variables
vif2 = update(vif2.0, ~.-County-Population-White.Pct-Income.Per.Capita, data=final)
vif(vif2)
```

##		GVIF	Df	GVIF^(1/(2*Df))
##	Median.Age	2.461801	1	1.569013
##	Unemployment.Rate	2.378788	1	1.542332
##	Violent.Crimes.Per.1000	3.353532	1	1.831265
##	High.School.Or.Higher.Pct	6.765412	1	2.601040
##	Bachelors.Or.Higher.Pct	5.773020	1	2.402711
##	Without.Health.Insurance.Pct	2.516775	1	1.586435

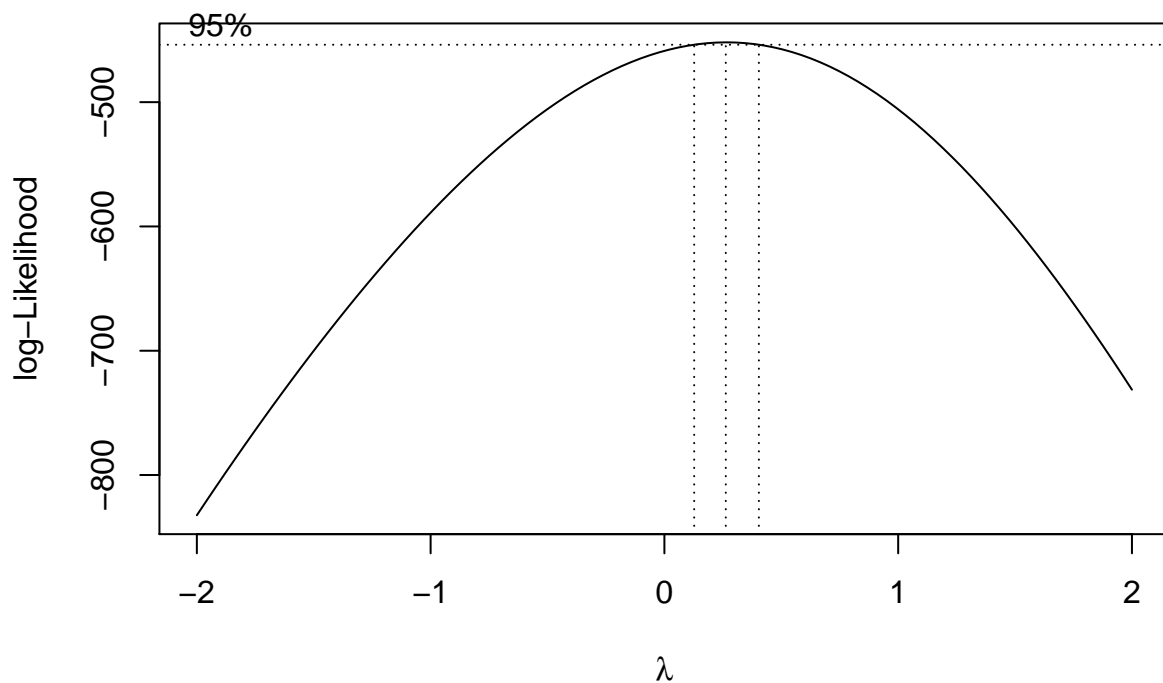
## Median.Household.Income	9.658589	1	3.107827
## Poverty.Rate	8.997157	1	2.999526
## Land.Area	1.470137	1	1.212492
## Population.Density	2.108047	1	1.451912
## Female.Pct	2.301163	1	1.516958
## Black.Pct	3.616364	1	1.901674
## Asian.Pct	4.501203	1	2.121604
## Mixed.Race.Pct	3.065863	1	1.750960
## Hispanic.Pct	7.469613	1	2.733059
## Foreign.Born.Pct	8.767332	1	2.960968
## Gun.Law.Rank.2010	1.902122	1	1.379175
## Federal.Govt.Expenditure.Per.Person	1.529962	1	1.236916
## Non.Violent.Crimes.Per.1000	3.449458	1	1.857272
## Voter.Group.2008	2.472353	2	1.253942

All of the VIF values are under 10, which is the “rule of thumb” cutoff; this solved the issue of *Bachelors.Or.Higher.Pct* because it was highly correlated with our income variables. So now we can move on to regressions.

Step 2. Creating the Best Interpretive Model

Let’s work on creating the best model that is still simple enough to interpret. We’ll start by creating a basic linear model and using boxcox to see if we need to transform our response variable.

```
full0 = lm(Crude.Rate~., data=final) #regress on all variables
full = update(full0, .~-County-Population-White.Pct-Income.Per.Capita, data=final) #remove county (name)
boxcox(full)
```

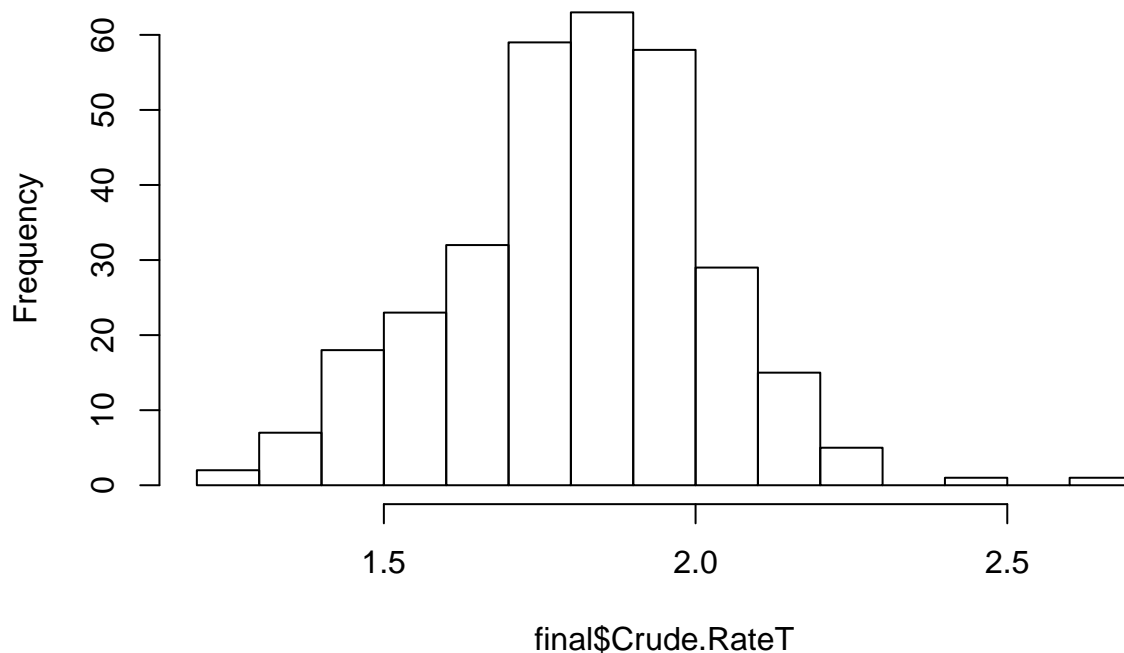


```
#locator(1)
```

The boxcox command suggests raising our response variable, *Crude.Rate* to the power 0.26. Let's raise it to the power $\frac{1}{4}$ since this is still within the recommended interval and the interpretation is a bit easier. We'll just add the letter "T" to the end of our variable *Crude.Rate* to indicate that it's been transformed. This transformation makes *Crude.Rate* much more normally distributed (compared to its right skewedness before).

```
final$Crude.RateT = final$Crude.Rate^(1/4)
hist(final$Crude.RateT)
```

Histogram of final\$Crude.RateT



Now we can create a basic linear model regressed on all our (remaining) variables. We'll then use minimum BIC stepwise to find which predictor variables are most relevant.

```
fullT0 = lm(Crude.RateT~.,data=final)
fullT = update(fullT0, ~.-Crude.Rate-County-Population-White.Pct-Income.Per.Capita, data=final)

stepfull = step(fullT,trace=0, k=log(nrow(final)))
summary(stepfull)
```

```
##
## Call:
## lm(formula = Crude.RateT ~ Median.Age + Unemployment.Rate + Violent.Crimes.Per.1000 +
##      High.School.Or.Higher.Pct + Without.Health.Insurance.Pct +
##      Female.Pct + Black.Pct + Foreign.Born.Pct + Gun.Law.Rank.2010 +
##      Federal.Govt.Expenditure.Per.Person, data = final)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.27446 -0.07309  0.00091  0.06700  0.39695
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   3.6264396   0.5247698   6.911 2.87e-11
## Median.Age     0.0105042   0.0019726   5.325 1.97e-07
## Unemployment.Rate 0.0125719  0.0035357   3.556 0.000437
## Violent.Crimes.Per.1000 0.0156875  0.0029845   5.256 2.78e-07
```



```
## High.School.Or.Higher.Pct      -0.0061079  0.0017563  -3.478  0.000580
## Without.Health.Insurance.Pct    0.0106760  0.0015056   7.091  9.49e-12
## Female.Pct                     -0.0411993  0.0098548  -4.181  3.81e-05
## Black.Pct                      0.0036128  0.0006931   5.212  3.47e-07
## Foreign.Born.Pct               -0.0131767  0.0010226 -12.886  < 2e-16
## Gun.Law.Rank.2010              0.0039595  0.0006131   6.458  4.24e-10
## Federal.Govt.Expenditure.Per.Person 0.0052091  0.0010919   4.770  2.87e-06
##
## (Intercept)                    ***
## Median.Age                     ***
## Unemployment.Rate              ***
## Violent.Crimes.Per.1000        ***
## High.School.Or.Higher.Pct      ***
## Without.Health.Insurance.Pct   ***
## Female.Pct                     ***
## Black.Pct                      ***
## Foreign.Born.Pct               ***
## Gun.Law.Rank.2010              ***
## Federal.Govt.Expenditure.Per.Person ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1132 on 302 degrees of freedom
## Multiple R-squared:  0.7117, Adjusted R-squared:  0.7022
## F-statistic: 74.56 on 10 and 302 DF,  p-value: < 2.2e-16
```

Step 3. Creating the Best Predictive Model

Now let's move on to creating a model that is more complex but has better predictability. We'll do this in JMP.

```
write.table(final, "C:/Users/Casey DeLano/Desktop/STAT 346/final_jmp.txt", sep="\t", row.names=F)
```

We use *Crude.Rate* as our *Y* variable and all of the predictor variables except for *County*, *Population*, *White.Pct*, and *Income.Per.Capita*. We ran a model with all interactions and squared terms using forward stepwise minimum BIC. The following is the model that resulted:

```
knitr::include_graphics("summary1.png")
```

Fit Group

Response Crude.Rate

Effect Summary

Source	LogWorth	PValue
Poverty.Rate*Black.Pct	11.310	0.00000
Violent.Crimes.Per.1000	8.522	0.00000
Without.Health.Insurance.Pct	6.965	0.00000
Federal.Govt.Expenditure.Per.Person	6.299	0.00000
Violent.Crimes.Per.1000*Foreign.Born.Pct	5.960	0.00000
High.School.Or.Higher.Pct*High.School.Or.Higher.Pct	5.655	0.00000
Gun.Law.Rank.2010	5.511	0.00000
Population.Density*Population.Density	5.279	0.00001
Foreign.Born.Pct	4.914	0.00001 ^
Median.Age	4.696	0.00002
Gun.Law.Rank.2010*Federal.Govt.Expenditure.Per.Person	4.458	0.00003
Black.Pct*Non.Violent.Crimes.Per.1000	3.579	0.00026
Median.Household.Income*Foreign.Born.Pct	3.486	0.00033
Without.Health.Insurance.Pct*Non.Violent.Crimes.Per.1000	3.134	0.00073
High.School.Or.Higher.Pct*Federal.Govt.Expenditure.Per.Person	2.850	0.00141
Unemployment.Rate*Black.Pct	2.599	0.00252
Land.Area*Black.Pct	2.450	0.00355
Voter.Group.2008{Democrat&Swing-Republican}	2.205	0.00624
Bachelors.Or.Higher.Pct	2.137	0.00729
Black.Pct	1.722	0.01895 ^
Without.Health.Insurance.Pct*Without.Health.Insurance.Pct	1.699	0.02002
Black.Pct*Gun.Law.Rank.2010	1.251	0.05606
Bachelors.Or.Higher.Pct*Bachelors.Or.Higher.Pct	1.125	0.07500
Unemployment.Rate	1.053	0.08846 ^
Female.Pct	0.949	0.11240
Land.Area	0.939	0.11516 ^
Median.Household.Income	0.602	0.24996 ^
Poverty.Rate	0.582	0.26169 ^
Non.Violent.Crimes.Per.1000	0.251	0.56166 ^
High.School.Or.Higher.Pct	0.204	0.62576 ^
Population.Density	0.129	0.74278 ^

[Remove](#) [Add](#) [Edit](#) ☐ FDR ('^' denotes effects with containing effects above them)

Summary of Fit

RSquare	0.797668
RSquare Adj	0.775347
Root Mean Square Error	2.541447
Mean of Response	11.72093

```
knitr::include_graphics("parameter_estimates1.png")
```

Fit Group

Response Crude.Rate

Analysis of Variance

Source	DF	Sum of Squares	Prob > F
C. Total	312	8970.2478	<.0001*

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob> t
Intercept	27.066416	14.71881	1.84	0.0670
Median.Age	0.2362448	0.054467	4.34	<.0001*
Unemployment.Rate	-0.193072	0.112939	-1.71	0.0885
Violent.Crimes.Per.1000	0.6225759	0.101599	6.13	<.0001*
High.School.Or.Higher.Pct	-0.035858	0.073443	-0.49	0.6258
Bachelors.Or.Higher.Pct	-0.113727	0.042073	-2.70	0.0073*
Without.Health.Insurance.Pct	0.2402403	0.044055	5.45	<.0001*
Median.Household.Income	-4.774e-5	4.141e-5	-1.15	0.2500
Poverty.Rate	-0.119545	0.106293	-1.12	0.2617
Land.Area	-0.000306	0.000194	-1.58	0.1152
Population.Density	5.4373e-5	0.000166	0.33	0.7428
Female.Pct	-0.398557	0.250272	-1.59	0.1124
Black.Pct	0.0555712	0.023545	2.36	0.0189*
Foreign.Born.Pct	-0.204317	0.045874	-4.45	<.0001*
Gun.Law.Rank.2010	0.0777511	0.01633	4.76	<.0001*
Federal.Govt.Expenditure.Per.Person	0.1367284	0.026574	5.15	<.0001*
Non.Violent.Crimes.Per.1000	-0.006214	0.010694	-0.58	0.5617
Voter.Group.2008{Democrat&Swing-Republican}	-0.653145	0.237027	-2.76	0.0062*
(High.School.Or.Higher.Pct-85.9383)*(High.School.Or.Higher.Pct-85.9383)	0.0216841	0.004486	4.83	<.0001*
(Bachelors.Or.Higher.Pct-28.6131)*(Bachelors.Or.Higher.Pct-28.6131)	0.0031691	0.001773	1.79	0.0750
(Without.Health.Insurance.Pct-17.1304)*(Without.Health.Insurance.Pct-17.1304)	-0.011943	0.005105	-2.34	0.0200*
(Population.Density-1213.25)*(Population.Density-1213.25)	-2.952e-8	6.356e-9	-4.64	<.0001*
(Unemployment.Rate-9.74249)*(Black.Pct-14.962)	-0.020884	0.00685	-3.05	0.0025*
(Poverty.Rate-14.1003)*(Black.Pct-14.962)	0.0219558	0.003041	7.22	<.0001*
(Land.Area-1252.27)*(Black.Pct-14.962)	-5.024e-5	0.000017	-2.94	0.0035*
(Violent.Crimes.Per.1000-4.97379)*(Foreign.Born.Pct-10.4837)	0.0527217	0.01058	4.98	<.0001*
(Median.Household.Income-52334.8)*(Foreign.Born.Pct-10.4837)	7.5736e-6	2.082e-6	3.64	0.0003*
(Black.Pct-14.962)*(Gun.Law.Rank.2010-19.0415)	-0.002015	0.00105	-1.92	0.0561
(High.School.Or.Higher.Pct-85.9383)*(Federal.Govt.Expenditure.Per.Person-10.5352)	-0.02076	0.006439	-3.22	0.0014*
(Gun.Law.Rank.2010-19.0415)*(Federal.Govt.Expenditure.Per.Person-10.5352)	0.0107775	0.002562	4.21	<.0001*
(Without.Health.Insurance.Pct-17.1304)*(Non.Violent.Crimes.Per.1000-70.7369)	-0.004759	0.001394	-3.41	0.0007*
(Black.Pct-14.962)*(Non.Violent.Crimes.Per.1000-70.7369)	-0.001852	0.000501	-3.70	0.0003*

Effect Tests

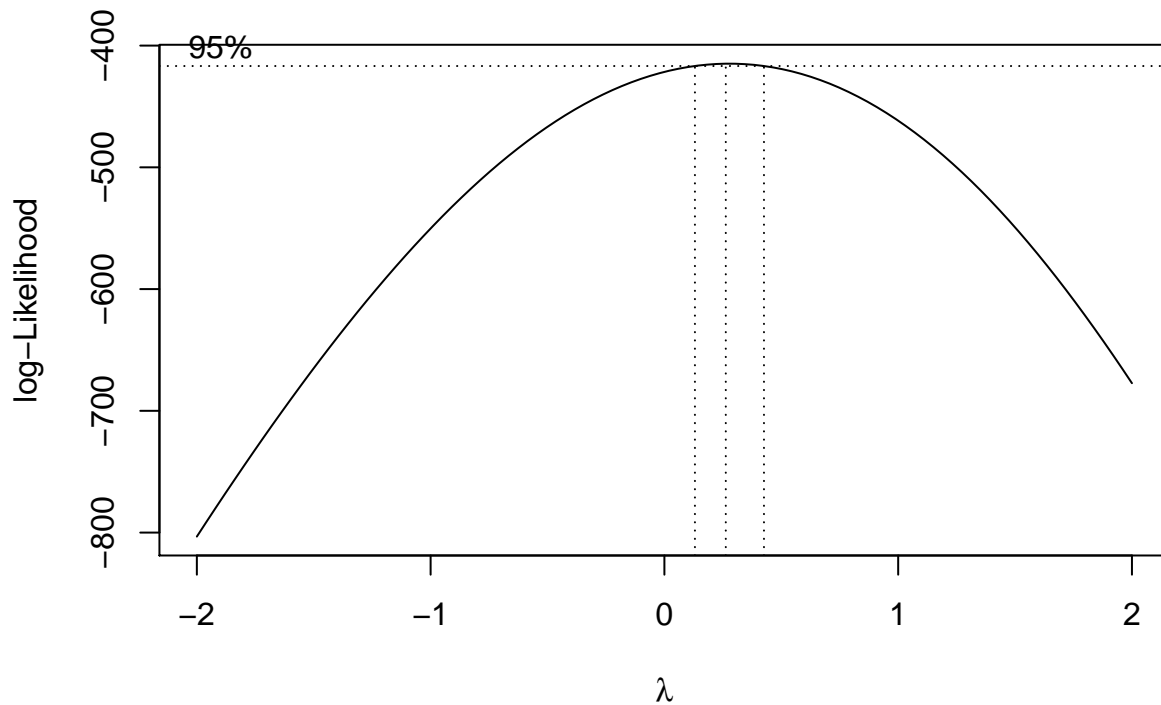
Effect Details

We entered this model into R.

```
jmp.model <- lm(Crude.Rate~Poverty.Rate*Black.Pct+Violent.Crimes.Per.1000+Without.Health.Insurance.Pct+)
```

Similarly to what we did with our simpler model, let's see if we should transform *Crude.Rate* in this model:

```
boxcox(jmp.model)
```



```
#locator(1)
```

Very similarly to before, we get the recommended transformation is raising *Crude.Rate* to the power 0.27. Let's stick to $\frac{1}{4}$ for the same reasons. We re-ran the forward minimum BIC stepwise model in JMP in the exact same way as before except for using *Crude.Rate^T* as our *Y*. The following is the model that resulted:

```
knitr::include_graphics("crudeRateT_summary.png")
```

Fit Group

Response Crude.RateT

Effect Summary

Source	LogWorth	
Foreign.Born.Pct	9.728	
Without.Health.Insurance.Pct	8.273	
Violent.Crimes.Per.1000*Foreign.Born.Pct	7.206	
Black.Pct	7.108	
Population.Density*Foreign.Born.Pct	6.136	
Foreign.Born.Pct*Foreign.Born.Pct	5.813	
Black.Pct*Gun.Law.Rank.2010	5.739	
Gun.Law.Rank.2010	5.723	
Unemployment.Rate*High.School.Or.Higher.Pct	5.415	
Federal.Govt.Expenditure.Per.Person	5.241	
Median.Age	4.842	
Without.Health.Insurance.Pct*Hispanic.Pct	4.662	
Gun.Law.Rank.2010*Federal.Govt.Expenditure.Per.Person	3.604	
Poverty.Rate*Voter.Group.2008{Democrat&Swing-Republican}	3.418	
Without.Health.Insurance.Pct*Non.Violent.Crimes.Per.1000	3.386	
Female.Pct	3.028	
Violent.Crimes.Per.1000	2.964	
High.School.Or.Higher.Pct*Median.Household.Income	2.911	
Female.Pct*Foreign.Born.Pct	2.741	
Median.Household.Income*Voter.Group.2008{Democrat&Swing-Republican}	2.558	
Median.Household.Income	2.183	
High.School.Or.Higher.Pct*High.School.Or.Higher.Pct	2.047	
Mixed.Race.Pct*Hispanic.Pct	1.900	
Mixed.Race.Pct	1.842	
Mixed.Race.Pct*Federal.Govt.Expenditure.Per.Person	1.753	
Population.Density	1.158	
Poverty.Rate	1.001	
Unemployment.Rate	0.691	
High.School.Or.Higher.Pct	0.321	
Voter.Group.2008{Democrat&Swing-Republican}	0.258	
Hispanic.Pct	0.178	
Non.Violent.Crimes.Per.1000	0.169	
High.School.Or.Higher.Pct*Without.Health.Insurance.Pct	0.129	

[Remove](#) [Add](#) [Edit](#) ☐ FDR ('^' denotes effects with containing effects above them)

Summary of Fit

RSquare	0.829726
RSquare Adj	0.809586
Root Mean Square Error	0.0095

```
knitr::include_graphics("crudeRateT_parameter_estimates.png")
```

Report: Fit Model - JMP Pro

Fit Group

Response Crude.RateT

Analysis of Variance

Source	DF	Squares	Prob > F
--------	----	---------	----------

Parameter Estimates

Term	Estimate	Std Error	t Ratio
Intercept	3.0247629	0.491426	6.16
Median.Age	0.0091251	0.002066	4.42
Unemployment.Rate	0.0045143	0.003545	1.27
Violent.Crimes.Per.1000	0.0114594	0.003471	3.30
High.School.Or.Higher.Pct	-0.001691	0.002376	-0.71
Without.Health.Insurance.Pct	0.0102551	0.001702	6.03
Median.Household.Income	-4.129e-6	1.508e-6	-2.74
Poverty.Rate	-0.005459	0.003305	-1.65
Population.Density	1.2044e-5	6.608e-6	1.84
Female.Pct	-0.029117	0.008707	-3.34
Black.Pct	0.0041667	0.000755	5.52
Mixed.Race.Pct	0.0143871	0.005841	2.46
Hispanic.Pct	0.0005141	0.00118	0.43
Foreign.Born.Pct	-0.014276	0.002158	-6.62
Gun.Law.Rank.2010	0.0027114	0.000557	4.87
Federal.Govt.Expenditure.Per.Person	0.004435	0.000959	4.63
Non.Violent.Crimes.Per.1000	0.000159	0.000381	0.42
Voter.Group.2008{Democrat&Swing-Republican}	-0.005858	0.009856	-0.59
(Unemployment.Rate-9.74249)*(High.School.Or.Higher.Pct-85.9383)	0.0025716	0.000546	4.71
(High.School.Or.Higher.Pct-85.9383)*(High.School.Or.Higher.Pct-85.9383)	0.0006543	0.000249	2.63
(High.School.Or.Higher.Pct-85.9383)*(Without.Health.Insurance.Pct-17.1304)	-0.000121	0.000369	-0.33
(High.School.Or.Higher.Pct-85.9383)*(Median.Household.Income-52334.8)	4.8892e-7	1.497e-7	3.27
(Without.Health.Insurance.Pct-17.1304)*(Hispanic.Pct-14.307)	-0.000524	0.000121	-4.33
(Mixed.Race.Pct-2.95048)*(Hispanic.Pct-14.307)	0.0011917	0.000475	2.51
(Violent.Crimes.Per.1000-4.97379)*(Foreign.Born.Pct-10.4837)	0.0017771	0.000319	5.57
(Population.Density-1213.25)*(Foreign.Born.Pct-10.4837)	-1.382e-6	2.727e-7	-5.07
(Female.Pct-51.0709)*(Foreign.Born.Pct-10.4837)	-0.003154	0.001002	-3.15
(Foreign.Born.Pct-10.4837)*(Foreign.Born.Pct-10.4837)	0.0003533	0.000072	4.91
(Black.Pct-14.962)*(Gun.Law.Rank.2010-19.0415)	-0.000165	3.382e-5	-4.88
(Mixed.Race.Pct-2.95048)*(Federal.Govt.Expenditure.Per.Person-10.5352)	-0.001547	0.000648	-2.39
(Gun.Law.Rank.2010-19.0415)*(Federal.Govt.Expenditure.Per.Person-10.5352)	0.0003304	0.000089	3.71
(Without.Health.Insurance.Pct-17.1304)*(Non.Violent.Crimes.Per.1000-70.7369)	-0.000192	5.361e-5	-3.58
(Median.Household.Income-52334.8)*(Voter.Group.2008{Democrat&Swing-Republican}-0.65495)	4.2725e-6	1.415e-6	3.02
(Poverty.Rate-14.1003)*(Voter.Group.2008{Democrat&Swing-Republican}-0.65495)	0.0134856	0.00375	3.60

Effect Tests

Effect Details

Again, we wrote this model into R.

```
jmp.modelT = lm(Crude.RateT~Foreign.Born.Pct+Without.Health.Insurance.Pct+Violent.Crimes.Per.1000*Forei
```

Let's quickly find the R^2 values for our two models. (Note that the R^2 for the JMP model is different when we run it in R than when JMP created it.)

```
summary(stepfull)$r.squared
```

```
## [1] 0.7117095
```

```
summary(jmp.modelT)$r.squared
```

```
## [1] 0.8164592
```

Step 4. Cross-Validation

Now let's run a cross-validation for both of the models. We'll create a test set that is 25% of our total data and a training set that is the remaining 75%. We'll loop through this entire process 10 times so that we can take an average of the resulting SSE, R^2 and MSE for the training and test sets.

```
#create empty vectors so we can add values in for loop
results_simple = NULL
results_complex = NULL

for (i in 1:10){
  #create training and test sets
  smp_size_i <- floor(0.75 * nrow(final))
  set.seed(i)
  train_ind_i <- sample(seq_len(nrow(final)), size = smp_size_i)
  train_i <- final[train_ind_i, ]
  test_i <- final[-train_ind_i, ]

  #create models with training sets
  lm_simple = lm(Crude.RateT~Median.Age+Unemployment.Rate+Violent.Crimes.Per.1000+High.School.Or.Higher
  lm_complex = lm(Crude.RateT~Foreign.Born.Pct+Without.Health.Insurance.Pct+Violent.Crimes.Per.1000*For

  #predict the model on the test set
  preds.simple = predict(lm_simple, test_i)
  preds.complex = predict(lm_complex, test_i)

  #calculate SSE, R^2, and MSEs for simple
  results_simple$sse[i] = sum((test_i$Crude.RateT-preds.simple)^2)
  results_simple$r2[i] = cor(test_i$Crude.RateT, preds.simple)^2
  results_simple$train.mse[i] = mean(lm_simple$residuals^2)
  results_simple$test.mse[i] = results_simple$sse[i]/(nrow(test_i))

  #calculate SSE, R^2, and MSEs for complex
  results_complex$sse[i] = sum((test_i$Crude.RateT-preds.complex)^2)
  results_complex$r2[i] = cor(test_i$Crude.RateT, preds.complex)^2
  results_complex$train.mse[i] = mean(lm_complex$residuals^2)
  results_complex$test.mse[i] = results_complex$sse[i]/(nrow(test_i))
```

```

}

#calculate means of sse, R^2, and mse and put into a table
results_simple = data.frame(results_simple)
results_complex = data.frame(results_complex)
means_simple = apply(results_simple, 2, mean)
means_complex = apply(results_complex, 2, mean)
vector = cbind(means_simple, means_complex)
kable(vector)

```

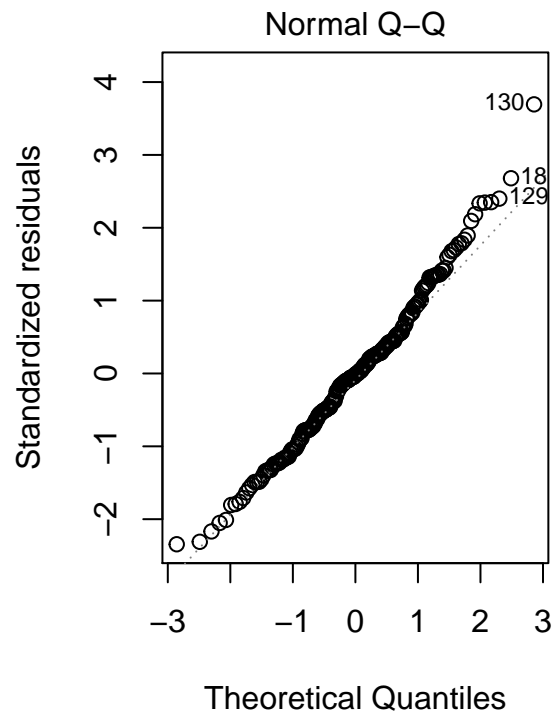
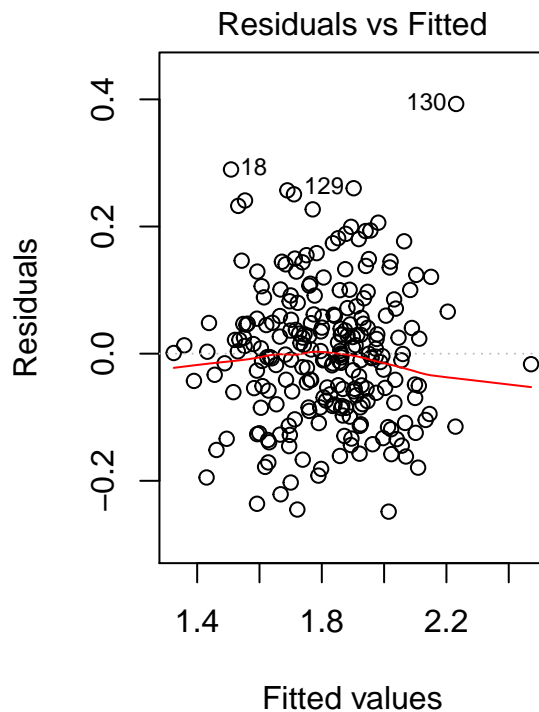
	means_simple	means_complex
sse	1.1172060	0.8391551
r2	0.7028361	0.7775684
train.mse	0.0120176	0.0075293
test.mse	0.0141418	0.0106222

Step 5. Diagnostics

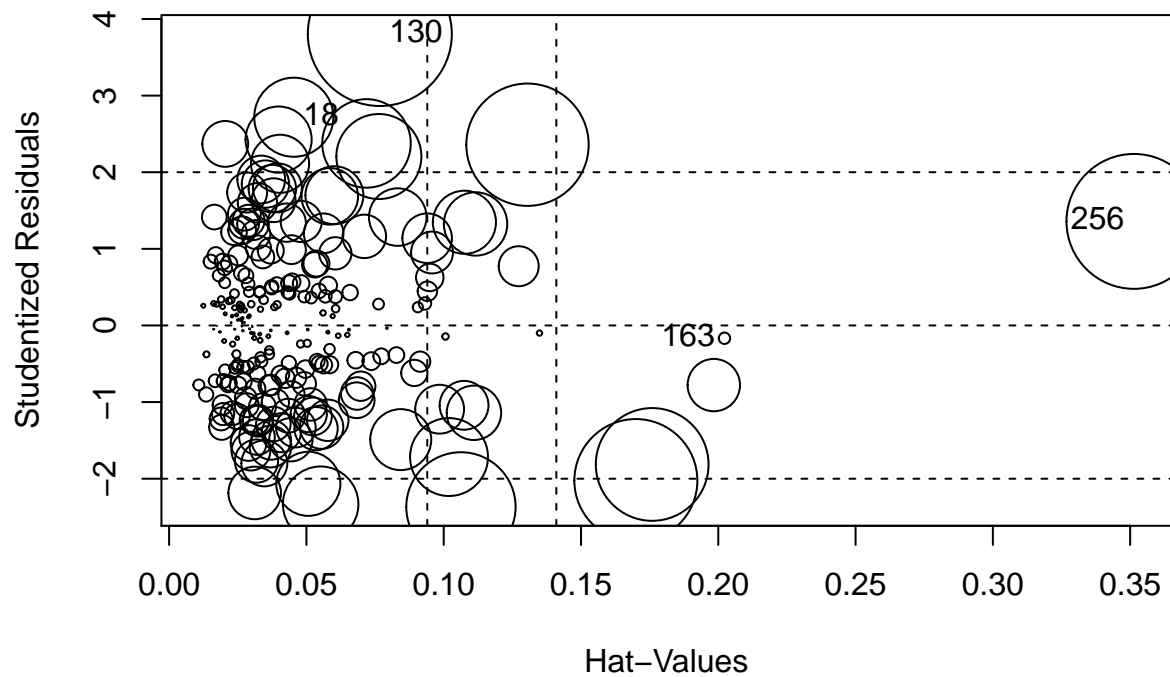
```

#simple model
par(mfrow=c(1,2))
plot(lm_simple,1) #residual plot
plot(lm_simple,2) #normal plot

```

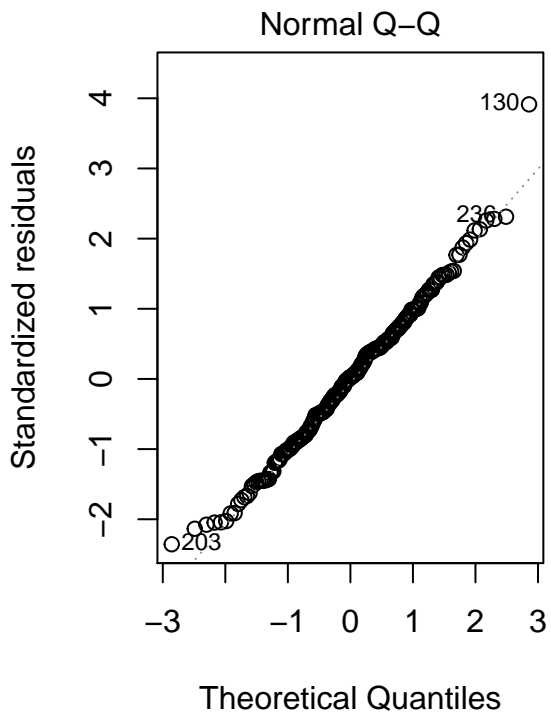
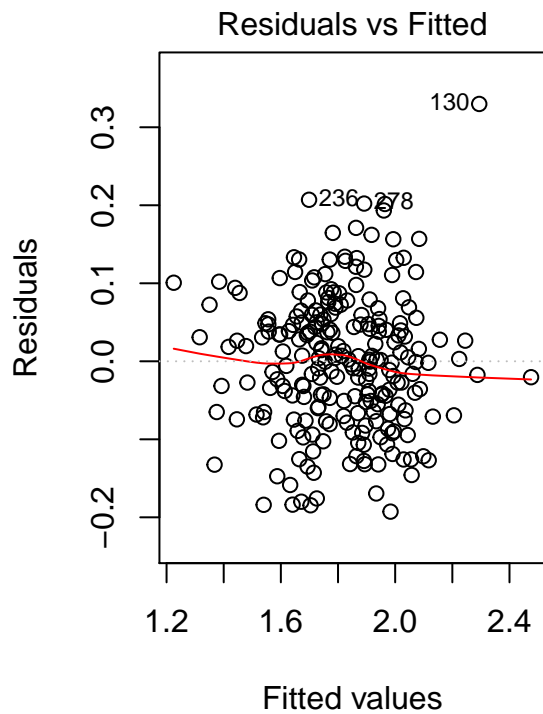



```
par(mfrow=c(1,1))
influencePlot(lm_simple) #influence plot
```

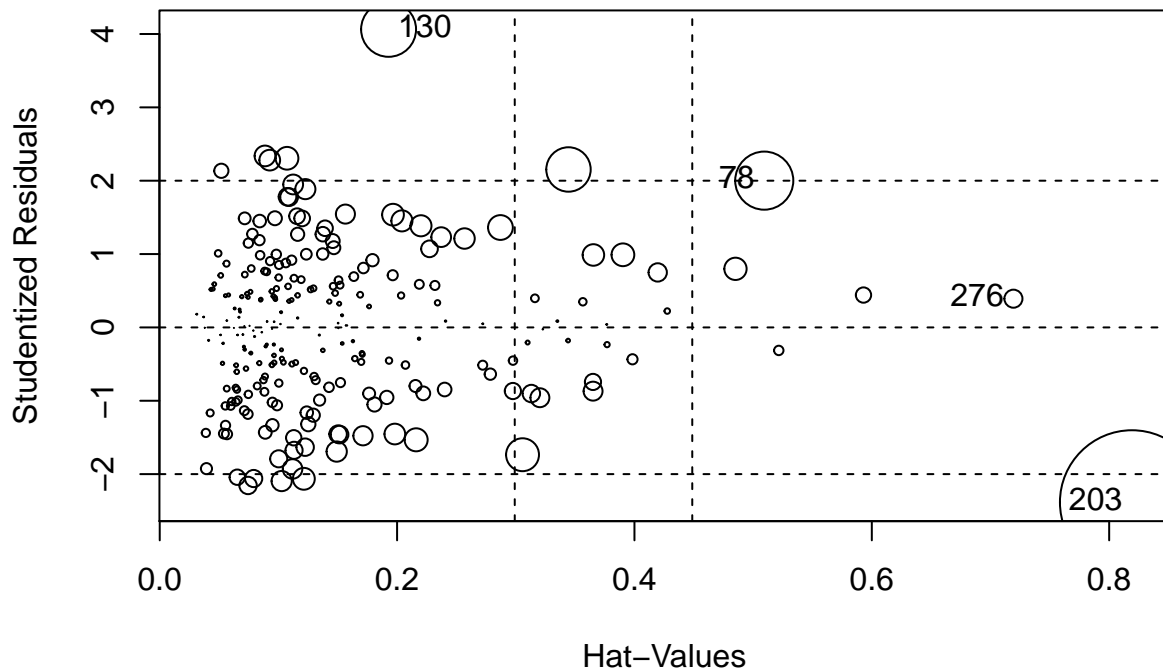


##	StudRes	Hat	CookD
## 130	3.8041332	0.07672751	0.1031017788
## 256	1.3583222	0.35137986	0.0905225425
## 163	-0.1670623	0.20225033	0.0006460771
## 18	2.7179304	0.04539853	0.0310484163

```
#complex model
par(mfrow=c(1,2))
plot(lm_complex,1) #residual plot
plot(lm_complex,2) #normal plot
```



```
par(mfrow=c(1,1))
influencePlot(lm_complex) #influence plot
```



```
##      StudRes      Hat      CookD
## 130  4.0635763 0.1928288 0.10455752
## 78   2.0010003 0.5093743 0.11700545
## 276  0.3916652 0.7193314 0.01128103
## 203 -2.3834178 0.8192001 0.71850087
```

```
#rerun simple model without influential points
final2_s = final[-c(130,256,163,18),]
stepfull2 = lm(formula(stepfull), data=final2_s)
summary(stepfull2)
```

```
##
## Call:
## lm(formula = formula(stepfull), data = final2_s)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.280175 -0.071474  0.001299  0.069791  0.251706
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    3.5209186   0.5132257   6.860 3.98e-11
## Median.Age      0.0099063   0.0019235   5.150 4.74e-07
## Unemployment.Rate 0.0124439   0.0034408   3.617 0.000351
## Violent.Crimes.Per.1000 0.0159934   0.0031090   5.144 4.88e-07
```

```
## High.School.Or.Higher.Pct      -0.0057674  0.0017189  -3.355  0.000896
## Without.Health.Insurance.Pct    0.0107030  0.0014742   7.260  3.38e-12
## Female.Pct                     -0.0390043  0.0096662  -4.035  6.94e-05
## Black.Pct                      0.0033618  0.0006926   4.854  1.96e-06
## Foreign.Born.Pct               -0.0133374  0.0010039 -13.286  < 2e-16
## Gun.Law.Rank.2010              0.0039199  0.0005992   6.542  2.64e-10
## Federal.Govt.Expenditure.Per.Person 0.0041430  0.0012305   3.367  0.000860
##
## (Intercept)                    ***
## Median.Age                     ***
## Unemployment.Rate              ***
## Violent.Crimes.Per.1000        ***
## High.School.Or.Higher.Pct      ***
## Without.Health.Insurance.Pct    ***
## Female.Pct                     ***
## Black.Pct                      ***
## Foreign.Born.Pct               ***
## Gun.Law.Rank.2010              ***
## Federal.Govt.Expenditure.Per.Person ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1099 on 298 degrees of freedom
## Multiple R-squared:  0.7036, Adjusted R-squared:  0.6937
## F-statistic: 70.74 on 10 and 298 DF,  p-value: < 2.2e-16
```

```
summary(stepfull)
```

```
##
## Call:
## lm(formula = Crude.RateT ~ Median.Age + Unemployment.Rate + Violent.Crimes.Per.1000 +
##      High.School.Or.Higher.Pct + Without.Health.Insurance.Pct +
##      Female.Pct + Black.Pct + Foreign.Born.Pct + Gun.Law.Rank.2010 +
##      Federal.Govt.Expenditure.Per.Person, data = final)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.27446 -0.07309  0.00091  0.06700  0.39695
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   3.6264396  0.5247698   6.911 2.87e-11
## Median.Age     0.0105042  0.0019726   5.325 1.97e-07
## Unemployment.Rate 0.0125719  0.0035357   3.556 0.000437
## Violent.Crimes.Per.1000 0.0156875  0.0029845   5.256 2.78e-07
## High.School.Or.Higher.Pct -0.0061079  0.0017563  -3.478 0.000580
## Without.Health.Insurance.Pct 0.0106760  0.0015056   7.091 9.49e-12
## Female.Pct    -0.0411993  0.0098548  -4.181 3.81e-05
## Black.Pct      0.0036128  0.0006931   5.212 3.47e-07
## Foreign.Born.Pct -0.0131767  0.0010226 -12.886 < 2e-16
## Gun.Law.Rank.2010 0.0039595  0.0006131   6.458 4.24e-10
## Federal.Govt.Expenditure.Per.Person 0.0052091  0.0010919   4.770 2.87e-06
##
## (Intercept)                    ***
```

```
## Median.Age ***
## Unemployment.Rate ***
## Violent.Crimes.Per.1000 ***
## High.School.Or.Higher.Pct ***
## Without.Health.Insurance.Pct ***
## Female.Pct ***
## Black.Pct ***
## Foreign.Born.Pct ***
## Gun.Law.Rank.2010 ***
## Federal.Govt.Expenditure.Per.Person ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1132 on 302 degrees of freedom
## Multiple R-squared:  0.7117, Adjusted R-squared:  0.7022
## F-statistic: 74.56 on 10 and 302 DF,  p-value: < 2.2e-16
```

```
assert(sign(summary(stepfull2)$coef) == sign(summary(stepfull)$coef)) #assert the signs of the coefficients
#rerun simple model without influential points
final2_c = final[-c(130,78,276,203),]
jmp.modelT2 = lm(formula(jmp.modelT), data=final2_c)
summary(jmp.modelT2)
```

```
##
## Call:
## lm(formula = formula(jmp.modelT), data = final2_c)
##
## Residuals:
```

	Min	1Q	Median	3Q	Max
	-0.288146	-0.058276	0.000422	0.061035	0.270513

```
##
## Coefficients:
```

	Estimate
(Intercept)	4.535e+00
Foreign.Born.Pct	1.338e-01
Without.Health.Insurance.Pct	4.864e-02
Violent.Crimes.Per.1000	-9.533e-03
Black.Pct	7.164e-03
Population.Density	9.727e-06
Gun.Law.Rank.2010	1.898e-03
Unemployment.Rate	-1.399e-01
High.School.Or.Higher.Pct	-4.827e-02
Federal.Govt.Expenditure.Per.Person	2.880e-03
Median.Age	7.685e-03
Hispanic.Pct	1.833e-03
Poverty.Rate	8.600e-03
Voter.Group.2008Republican	1.255e+00
Voter.Group.2008Swing	3.825e-01
Non.Violent.Crimes.Per.1000	3.813e-03
Female.Pct	7.545e-03
Median.Household.Income	-5.236e-05
Mixed.Race.Pct	1.031e-02
Foreign.Born.Pct:Violent.Crimes.Per.1000	1.958e-03

## Foreign.Born.Pct:Population.Density	-5.672e-07
## Black.Pct:Gun.Law.Rank.2010	-1.732e-04
## Unemployment.Rate:High.School.Or.Higher.Pct	1.689e-03
## Without.Health.Insurance.Pct:Hispanic.Pct	-3.701e-04
## Gun.Law.Rank.2010:Federal.Govt.Expenditure.Per.Person	3.094e-04
## Poverty.Rate:Voter.Group.2008Republican	-3.843e-02
## Poverty.Rate:Voter.Group.2008Swing	-1.065e-02
## Without.Health.Insurance.Pct:Non.Violent.Crimes.Per.1000	-2.075e-04
## High.School.Or.Higher.Pct:Median.Household.Income	6.133e-07
## Foreign.Born.Pct:Female.Pct	-2.994e-03
## Voter.Group.2008Republican:Median.Household.Income	-1.285e-05
## Voter.Group.2008Swing:Median.Household.Income	-3.945e-06
## Hispanic.Pct:Mixed.Race.Pct	1.397e-03
## Federal.Govt.Expenditure.Per.Person:Mixed.Race.Pct	-1.515e-03
## Without.Health.Insurance.Pct:High.School.Or.Higher.Pct	-2.286e-04
##	Std. Error
## (Intercept)	1.287e+00
## Foreign.Born.Pct	5.415e-02
## Without.Health.Insurance.Pct	3.373e-02
## Violent.Crimes.Per.1000	4.608e-03
## Black.Pct	1.152e-03
## Population.Density	9.481e-06
## Gun.Law.Rank.2010	1.024e-03
## Unemployment.Rate	4.064e-02
## High.School.Or.Higher.Pct	1.234e-02
## Federal.Govt.Expenditure.Per.Person	2.841e-03
## Median.Age	2.064e-03
## Hispanic.Pct	3.379e-03
## Poverty.Rate	5.272e-03
## Voter.Group.2008Republican	2.717e-01
## Voter.Group.2008Swing	1.806e-01
## Non.Violent.Crimes.Per.1000	1.052e-03
## Female.Pct	1.380e-02
## Median.Household.Income	1.264e-05
## Mixed.Race.Pct	1.376e-02
## Foreign.Born.Pct:Violent.Crimes.Per.1000	3.652e-04
## Foreign.Born.Pct:Population.Density	3.078e-07
## Black.Pct:Gun.Law.Rank.2010	3.506e-05
## Unemployment.Rate:High.School.Or.Higher.Pct	4.726e-04
## Without.Health.Insurance.Pct:Hispanic.Pct	1.223e-04
## Gun.Law.Rank.2010:Federal.Govt.Expenditure.Per.Person	9.013e-05
## Poverty.Rate:Voter.Group.2008Republican	8.558e-03
## Poverty.Rate:Voter.Group.2008Swing	5.662e-03
## Without.Health.Insurance.Pct:Non.Violent.Crimes.Per.1000	5.471e-05
## High.School.Or.Higher.Pct:Median.Household.Income	1.341e-07
## Foreign.Born.Pct:Female.Pct	1.059e-03
## Voter.Group.2008Republican:Median.Household.Income	3.178e-06
## Voter.Group.2008Swing:Median.Household.Income	1.975e-06
## Hispanic.Pct:Mixed.Race.Pct	5.371e-04
## Federal.Govt.Expenditure.Per.Person:Mixed.Race.Pct	6.559e-04
## Without.Health.Insurance.Pct:High.School.Or.Higher.Pct	3.678e-04
##	t value Pr(> t)
## (Intercept)	3.523 0.000500
## Foreign.Born.Pct	2.471 0.014080

## Without.Health.Insurance.Pct	1.442	0.150424
## Violent.Crimes.Per.1000	-2.069	0.039524
## Black.Pct	6.220	1.85e-09
## Population.Density	1.026	0.305843
## Gun.Law.Rank.2010	1.853	0.064954
## Unemployment.Rate	-3.443	0.000665
## High.School.Or.Higher.Pct	-3.911	0.000116
## Federal.Govt.Expenditure.Per.Person	1.014	0.311633
## Median.Age	3.724	0.000238
## Hispanic.Pct	0.543	0.587813
## Poverty.Rate	1.631	0.103942
## Voter.Group.2008Republican	4.619	5.94e-06
## Voter.Group.2008Swing	2.118	0.035100
## Non.Violent.Crimes.Per.1000	3.626	0.000343
## Female.Pct	0.547	0.585003
## Median.Household.Income	-4.143	4.57e-05
## Mixed.Race.Pct	0.749	0.454467
## Foreign.Born.Pct:Violent.Crimes.Per.1000	5.361	1.76e-07
## Foreign.Born.Pct:Population.Density	-1.843	0.066399
## Black.Pct:Gun.Law.Rank.2010	-4.940	1.36e-06
## Unemployment.Rate:High.School.Or.Higher.Pct	3.575	0.000414
## Without.Health.Insurance.Pct:Hispanic.Pct	-3.026	0.002710
## Gun.Law.Rank.2010:Federal.Govt.Expenditure.Per.Person	3.433	0.000690
## Poverty.Rate:Voter.Group.2008Republican	-4.491	1.05e-05
## Poverty.Rate:Voter.Group.2008Swing	-1.882	0.060955
## Without.Health.Insurance.Pct:Non.Violent.Crimes.Per.1000	-3.792	0.000184
## High.School.Or.Higher.Pct:Median.Household.Income	4.575	7.22e-06
## Foreign.Born.Pct:Female.Pct	-2.828	0.005030
## Voter.Group.2008Republican:Median.Household.Income	-4.044	6.83e-05
## Voter.Group.2008Swing:Median.Household.Income	-1.997	0.046801
## Hispanic.Pct:Mixed.Race.Pct	2.601	0.009788
## Federal.Govt.Expenditure.Per.Person:Mixed.Race.Pct	-2.310	0.021648
## Without.Health.Insurance.Pct:High.School.Or.Higher.Pct	-0.622	0.534755
##		
## (Intercept)	***	
## Foreign.Born.Pct	*	
## Without.Health.Insurance.Pct		
## Violent.Crimes.Per.1000	*	
## Black.Pct	***	
## Population.Density		
## Gun.Law.Rank.2010	.	
## Unemployment.Rate	***	
## High.School.Or.Higher.Pct	***	
## Federal.Govt.Expenditure.Per.Person		
## Median.Age	***	
## Hispanic.Pct		
## Poverty.Rate		
## Voter.Group.2008Republican	***	
## Voter.Group.2008Swing	*	
## Non.Violent.Crimes.Per.1000	***	
## Female.Pct		
## Median.Household.Income	***	
## Mixed.Race.Pct		
## Foreign.Born.Pct:Violent.Crimes.Per.1000	***	

```
## Foreign.Born.Pct:Population.Density .
## Black.Pct:Gun.Law.Rank.2010 ***
## Unemployment.Rate:High.School.Or.Higher.Pct ***
## Without.Health.Insurance.Pct:Hispanic.Pct **
## Gun.Law.Rank.2010:Federal.Govt.Expenditure.Per.Person ***
## Poverty.Rate:Voter.Group.2008Republican ***
## Poverty.Rate:Voter.Group.2008Swing .
## Without.Health.Insurance.Pct:Non.Violent.Crimes.Per.1000 ***
## High.School.Or.Higher.Pct:Median.Household.Income ***
## Foreign.Born.Pct:Female.Pct **
## Voter.Group.2008Republican:Median.Household.Income ***
## Voter.Group.2008Swing:Median.Household.Income *
## Hispanic.Pct:Mixed.Race.Pct **
## Federal.Govt.Expenditure.Per.Person:Mixed.Race.Pct *
## Without.Health.Insurance.Pct:High.School.Or.Higher.Pct
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.09077 on 274 degrees of freedom
## Multiple R-squared:  0.8212, Adjusted R-squared:  0.799
## F-statistic: 37.01 on 34 and 274 DF,  p-value: < 2.2e-16
```

```
summary(jmp.modelT)
```

```
##
## Call:
## lm(formula = Crude.RateT ~ Foreign.Born.Pct + Without.Health.Insurance.Pct +
##   Violent.Crimes.Per.1000 * Foreign.Born.Pct + Black.Pct +
##   Population.Density * Foreign.Born.Pct + Foreign.Born.Pct *
##   Foreign.Born.Pct + Black.Pct * Gun.Law.Rank.2010 + Gun.Law.Rank.2010 +
##   Unemployment.Rate * High.School.Or.Higher.Pct + Federal.Govt.Expenditure.Per.Person +
##   Median.Age + Without.Health.Insurance.Pct * Hispanic.Pct +
##   Gun.Law.Rank.2010 * Federal.Govt.Expenditure.Per.Person +
##   Poverty.Rate * Voter.Group.2008 + Without.Health.Insurance.Pct *
##   Non.Violent.Crimes.Per.1000 + Female.Pct + Violent.Crimes.Per.1000 +
##   High.School.Or.Higher.Pct * Median.Household.Income + Female.Pct *
##   Foreign.Born.Pct + Median.Household.Income * Voter.Group.2008 +
##   Median.Household.Income + High.School.Or.Higher.Pct * High.School.Or.Higher.Pct +
##   Mixed.Race.Pct * Hispanic.Pct + Mixed.Race.Pct + Mixed.Race.Pct *
##   Federal.Govt.Expenditure.Per.Person + Population.Density +
##   Poverty.Rate + Unemployment.Rate + High.School.Or.Higher.Pct +
##   Voter.Group.2008 + Hispanic.Pct + Non.Violent.Crimes.Per.1000 +
##   High.School.Or.Higher.Pct * Without.Health.Insurance.Pct,
##   data = final)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.291728 -0.058135 -0.000541  0.056915  0.315059
##
## Coefficients:
##                                     Estimate
## (Intercept)                      4.656e+00
## Foreign.Born.Pct                  1.755e-01
## Without.Health.Insurance.Pct      4.456e-02
```


## Violent.Crimes.Per.1000	-1.004e-02
## Black.Pct	7.714e-03
## Population.Density	1.242e-05
## Gun.Law.Rank.2010	1.601e-03
## Unemployment.Rate	-1.492e-01
## High.School.Or.Higher.Pct	-5.054e-02
## Federal.Govt.Expenditure.Per.Person	2.771e-03
## Median.Age	8.124e-03
## Hispanic.Pct	1.579e-03
## Poverty.Rate	9.618e-03
## Voter.Group.2008Republican	1.310e+00
## Voter.Group.2008Swing	4.112e-01
## Non.Violent.Crimes.Per.1000	3.012e-03
## Female.Pct	1.013e-02
## Median.Household.Income	-5.570e-05
## Mixed.Race.Pct	1.419e-02
## Foreign.Born.Pct:Violent.Crimes.Per.1000	2.007e-03
## Foreign.Born.Pct:Population.Density	-9.683e-07
## Black.Pct:Gun.Law.Rank.2010	-1.760e-04
## Unemployment.Rate:High.School.Or.Higher.Pct	1.782e-03
## Without.Health.Insurance.Pct:Hispanic.Pct	-3.518e-04
## Gun.Law.Rank.2010:Federal.Govt.Expenditure.Per.Person	3.491e-04
## Poverty.Rate:Voter.Group.2008Republican	-4.073e-02
## Poverty.Rate:Voter.Group.2008Swing	-1.187e-02
## Without.Health.Insurance.Pct:Non.Violent.Crimes.Per.1000	-1.716e-04
## High.School.Or.Higher.Pct:Median.Household.Income	6.428e-07
## Foreign.Born.Pct:Female.Pct	-3.755e-03
## Voter.Group.2008Republican:Median.Household.Income	-1.314e-05
## Voter.Group.2008Swing:Median.Household.Income	-4.091e-06
## Hispanic.Pct:Mixed.Race.Pct	1.186e-03
## Federal.Govt.Expenditure.Per.Person:Mixed.Race.Pct	-1.694e-03
## Without.Health.Insurance.Pct:High.School.Or.Higher.Pct	-2.177e-04
##	Std. Error
## (Intercept)	1.282e+00
## Foreign.Born.Pct	5.222e-02
## Without.Health.Insurance.Pct	3.339e-02
## Violent.Crimes.Per.1000	4.623e-03
## Black.Pct	1.186e-03
## Population.Density	9.452e-06
## Gun.Law.Rank.2010	1.060e-03
## Unemployment.Rate	4.113e-02
## High.School.Or.Higher.Pct	1.254e-02
## Federal.Govt.Expenditure.Per.Person	2.931e-03
## Median.Age	2.128e-03
## Hispanic.Pct	3.313e-03
## Poverty.Rate	5.062e-03
## Voter.Group.2008Republican	2.726e-01
## Voter.Group.2008Swing	1.741e-01
## Non.Violent.Crimes.Per.1000	1.073e-03
## Female.Pct	1.395e-02
## Median.Household.Income	1.303e-05
## Mixed.Race.Pct	1.377e-02
## Foreign.Born.Pct:Violent.Crimes.Per.1000	3.318e-04
## Foreign.Born.Pct:Population.Density	2.792e-07

```

## Black.Pct:Gun.Law.Rank.2010 3.599e-05
## Unemployment.Rate:High.School.Or.Higher.Pct 4.796e-04
## Without.Health.Insurance.Pct:Hispanic.Pct 1.220e-04
## Gun.Law.Rank.2010:Federal.Govt.Expenditure.Per.Person 9.293e-05
## Poverty.Rate:Voter.Group.2008Republican 8.481e-03
## Poverty.Rate:Voter.Group.2008Swing 5.310e-03
## Without.Health.Insurance.Pct:Non.Violent.Crimes.Per.1000 5.591e-05
## High.School.Or.Higher.Pct:Median.Household.Income 1.387e-07
## Foreign.Born.Pct:Female.Pct 1.029e-03
## Voter.Group.2008Republican:Median.Household.Income 3.227e-06
## Voter.Group.2008Swing:Median.Household.Income 1.944e-06
## Hispanic.Pct:Mixed.Race.Pct 4.825e-04
## Federal.Govt.Expenditure.Per.Person:Mixed.Race.Pct 6.729e-04
## Without.Health.Insurance.Pct:High.School.Or.Higher.Pct 3.632e-04
## t value Pr(>|t|)
## (Intercept) 3.633 0.000334
## Foreign.Born.Pct 3.361 0.000886
## Without.Health.Insurance.Pct 1.334 0.183194
## Violent.Crimes.Per.1000 -2.173 0.030658
## Black.Pct 6.505 3.61e-10
## Population.Density 1.314 0.189848
## Gun.Law.Rank.2010 1.510 0.132067
## Unemployment.Rate -3.629 0.000339
## High.School.Or.Higher.Pct -4.032 7.16e-05
## Federal.Govt.Expenditure.Per.Person 0.946 0.345209
## Median.Age 3.818 0.000166
## Hispanic.Pct 0.476 0.634124
## Poverty.Rate 1.900 0.058486
## Voter.Group.2008Republican 4.805 2.54e-06
## Voter.Group.2008Swing 2.362 0.018865
## Non.Violent.Crimes.Per.1000 2.807 0.005350
## Female.Pct 0.727 0.468020
## Median.Household.Income -4.274 2.64e-05
## Mixed.Race.Pct 1.030 0.303954
## Foreign.Born.Pct:Violent.Crimes.Per.1000 6.049 4.69e-09
## Foreign.Born.Pct:Population.Density -3.469 0.000606
## Black.Pct:Gun.Law.Rank.2010 -4.890 1.71e-06
## Unemployment.Rate:High.School.Or.Higher.Pct 3.715 0.000246
## Without.Health.Insurance.Pct:Hispanic.Pct -2.883 0.004248
## Gun.Law.Rank.2010:Federal.Govt.Expenditure.Per.Person 3.756 0.000210
## Poverty.Rate:Voter.Group.2008Republican -4.803 2.56e-06
## Poverty.Rate:Voter.Group.2008Swing -2.236 0.026129
## Without.Health.Insurance.Pct:Non.Violent.Crimes.Per.1000 -3.068 0.002366
## High.School.Or.Higher.Pct:Median.Household.Income 4.634 5.52e-06
## Foreign.Born.Pct:Female.Pct -3.651 0.000312
## Voter.Group.2008Republican:Median.Household.Income -4.073 6.07e-05
## Voter.Group.2008Swing:Median.Household.Income -2.105 0.036197
## Hispanic.Pct:Mixed.Race.Pct 2.458 0.014596
## Federal.Govt.Expenditure.Per.Person:Mixed.Race.Pct -2.518 0.012372
## Without.Health.Insurance.Pct:High.School.Or.Higher.Pct -0.599 0.549384
##
## (Intercept) ***
## Foreign.Born.Pct ***
## Without.Health.Insurance.Pct

```

```

## Violent.Crimes.Per.1000 *
## Black.Pct ***
## Population.Density
## Gun.Law.Rank.2010
## Unemployment.Rate ***
## High.School.Or.Higher.Pct ***
## Federal.Govt.Expenditure.Per.Person
## Median.Age ***
## Hispanic.Pct
## Poverty.Rate .
## Voter.Group.2008Republican ***
## Voter.Group.2008Swing *
## Non.Violent.Crimes.Per.1000 **
## Female.Pct
## Median.Household.Income ***
## Mixed.Race.Pct
## Foreign.Born.Pct:Violent.Crimes.Per.1000 ***
## Foreign.Born.Pct:Population.Density ***
## Black.Pct:Gun.Law.Rank.2010 ***
## Unemployment.Rate:High.School.Or.Higher.Pct ***
## Without.Health.Insurance.Pct:Hispanic.Pct **
## Gun.Law.Rank.2010:Federal.Govt.Expenditure.Per.Person ***
## Poverty.Rate:Voter.Group.2008Republican ***
## Poverty.Rate:Voter.Group.2008Swing *
## Without.Health.Insurance.Pct:Non.Violent.Crimes.Per.1000 **
## High.School.Or.Higher.Pct:Median.Household.Income ***
## Foreign.Born.Pct:Female.Pct ***
## Voter.Group.2008Republican:Median.Household.Income ***
## Voter.Group.2008Swing:Median.Household.Income *
## Hispanic.Pct:Mixed.Race.Pct *
## Federal.Govt.Expenditure.Per.Person:Mixed.Race.Pct *
## Without.Health.Insurance.Pct:High.School.Or.Higher.Pct
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.09413 on 278 degrees of freedom
## Multiple R-squared:  0.8165, Adjusted R-squared:  0.794
## F-statistic: 36.37 on 34 and 278 DF,  p-value: < 2.2e-16

```

```

assert(sign(summary(jmp.modelT2)$coef) == sign(summary(jmp.modelT)$coef)) #assert the signs of the coef.

```