Highway Exmaple #1: Stepwise Selection

For this data from Weisberg, the response variable (Y) is accident rate (per million vehicle miles) and there are 13 potential predictors (X's).

We first look at "traditional" (hypothesis testing) stepwise selection. While this approach is very common in other software programs, there is no "built-in" function in R, so we use add1() and drop1(). drop1() is not really necessary, because we can get the same information from summary().

We also look at AIC stepwise selection using the step() function.

Traditional and AIC stepwise selection can select different models. Backwards, forwards or stepwise selection can select different models.

```
highway <- read.csv("~/Dropbox/STAT512/Lectures/MultReg3/MR3_Highway.csv")
str(highway)
```

```
## 'data.frame':    39 obs. of  15 variables:
##  $ nr  : int  1 2 3 4 5 6 7 8 9 10 ...
##  $ rate: num  4.58 2.86 3.02 2.29 1.61 6.87 3.85 6.12 3.29 5.88 ...
##  $ len : num  4.99 16.11 9.75 10.65 20.01 ...
##  $ adt : int  69 73 49 61 28 30 46 25 43 23 ...
##  $ trks: int  8 8 10 13 12 6 8 9 12 7 ...
##  $ slim: int  55 60 60 65 70 55 55 55 50 50 ...
##  $ lwid: int  12 12 12 12 12 12 12 12 12 12 ...
##  $ shld: int  10 10 10 10 10 10 8 10 4 5 ...
##  $ itg : num  1.2 1.43 1.54 0.94 0.65 0.34 0.47 0.38 0.95 0.12 ...
##  $ sigs: num  0 0 0 0 0 1.84 0.7 0.38 1.39 1.21 ...
##  $ acpt: num  4.6 4.4 4.7 3.8 2.2 24.8 11 18.5 7.5 8.2 ...
##  $ lane: int  8 4 4 6 4 4 4 4 4 4 ...
##  $ fai : int  1 1 1 1 1 0 0 0 0 0 ...
##  $ pa  : int  0 0 0 0 0 1 1 1 1 1 ...
##  $ ma  : int  0 0 0 0 0 0 0 0 0 0 ...
```

```
highway <- highway[,-1]
cor(highway)
```

```
##              rate         len         adt         trks         slim
## rate  1.000000000 -0.46528958 -0.02856981 -0.512522209 -0.68098362
## len  -0.465289581  1.00000000 -0.27156858  0.495943140  0.18624323
## adt  -0.028569805 -0.27156858  1.00000000 -0.096682243  0.24415659
## trks -0.512522209  0.49594314 -0.09668224  1.000000000  0.29618435
## slim -0.680983625  0.18624323  0.24415659  0.296184352  1.00000000
## lwid -0.005619311 -0.31065010  0.12787838 -0.155271368  0.09869275
## shld -0.386907190 -0.10492613  0.45730677  0.006134713  0.68900862
## itg  -0.024840883 -0.24756221  0.90370136 -0.067231363  0.24128154
## sigs  0.564479507 -0.32179111  0.14547472 -0.450259060 -0.41021992
## acpt  0.752025471 -0.23870585 -0.22397976 -0.360266223 -0.68152051
## lane -0.032978589 -0.20250417  0.82392951 -0.153323636  0.26452003
## fai  -0.207610359 -0.02971796  0.75947658  0.142998680  0.46494781
## pa   -0.161539823 -0.15153022  0.02878198 -0.051501441  0.04442617
## ma    0.337848301  0.12959484 -0.46482808 -0.101414053 -0.42395015
##              lwid         shld         itg         sigs         acpt
## rate -0.005619311 -0.386907190 -0.02484088  0.56447951  0.75202547
## len  -0.310650098 -0.104926126 -0.24756221 -0.32179111 -0.23870585
```

```
## adt   0.127878383  0.457306775  0.90370136  0.14547472 -0.22397976
## trks -0.155271368  0.006134713 -0.06723136 -0.45025906 -0.36026622
## slim  0.098692754  0.689008617  0.24128154 -0.41021992 -0.68152051
## lwid  1.000000000 -0.042896157  0.10288256  0.04201636 -0.04201339
## shld -0.042896157  1.000000000  0.37502186 -0.13405785 -0.42495127
## itg   0.102882555  0.375021856  1.00000000  0.06950741 -0.20015776
## sigs  0.042016363 -0.134057847  0.06950741  1.00000000  0.49869347
## acpt -0.042013391 -0.424951272 -0.20015776  0.49869347  1.00000000
## lane  0.095722771  0.481771423  0.69791301  0.24968579 -0.20877852
## fai   0.043701886  0.400236275  0.80840799 -0.24565788 -0.34266141
## pa    0.225073088  0.366881280 -0.13041305  0.29622041 -0.22766102
## ma   -0.282038037 -0.623064920 -0.35559279 -0.07017857  0.51348782
##             lane         fai          pa          ma
## rate -0.03297859 -0.20761036 -0.16153982  0.33784830
## len  -0.20250417 -0.02971796 -0.15153022  0.12959484
## adt   0.82392951  0.75947658  0.02878198 -0.46482808
## trks -0.15332364  0.14299868 -0.05150144 -0.10141405
## slim  0.26452003  0.46494781  0.04442617 -0.42395015
## lwid  0.09572277  0.04370189  0.22507309 -0.28203804
## shld  0.48177142  0.40023627  0.36688128 -0.62306492
## itg   0.69791301  0.80840799 -0.13041305 -0.35559279
## sigs  0.24968579 -0.24565788  0.29622041 -0.07017857
## acpt -0.20877852 -0.34266141 -0.22766102  0.51348782
## lane  1.00000000  0.59151376  0.17431556 -0.51295178
## fai   0.59151376  1.00000000 -0.37377250 -0.27116307
## pa    0.17431556 -0.37377250  1.00000000 -0.68920244
## ma   -0.51295178 -0.27116307 -0.68920244  1.00000000
```

```
#Full Model
FullModel <- lm(rate ~ ., data = highway)
#Equivalent to:
#FullModel <- lm(rate ~ len + adt + trks + slim + lwid +
#+ shld + itg + sigs + acpt + lane + fai + pa + ma,data=highway)
summary(FullModel)
```

```
##
## Call:
## lm(formula = rate ~ ., data = highway)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.99564 -0.62039 -0.05676  0.61741  2.54998
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 13.658212   6.872719   1.987   0.0579 .
## len         -0.064751   0.033369  -1.940   0.0637 .
## adt         -0.004038   0.033945  -0.119   0.9063
## trks        -0.100150   0.114726  -0.873   0.3910
## slim        -0.123778   0.081683  -1.515   0.1422
## lwid        -0.133813   0.597917  -0.224   0.8247
## shld         0.014113   0.162174   0.087   0.9313
## itg         -0.475478   1.282742  -0.371   0.7140
## sigs         0.713644   0.525213   1.359   0.1864
## acpt         0.066588   0.042569   1.564   0.1303
```

```
## lane         0.026675   0.283834   0.094   0.9259
## fai          0.543592   1.728270   0.315   0.7557
## pa          -1.009777   1.105612  -0.913   0.3698
## ma          -0.548025   0.975623  -0.562   0.5793
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.198 on 25 degrees of freedom
## Multiple R-squared:  0.7605, Adjusted R-squared:  0.636
## F-statistic: 6.107 on 13 and 25 DF,  p-value: 5.733e-05
```

```
#Null Model
NullModel <- lm(rate ~ 1, data = highway)
```

"Traditional" Forward Selection "By Hand"

add1() shows the partial F-test p-value for each variable. Add the most significant variable using update() and do add1() again until all remaining variables are non-significant at a pre-specified cutoff (0.05 here).

```
Model1 <- NullModel
add1(Model1, scope = FullModel, test = "F")
```

```
## Single term additions
##
## Model:
## rate ~ 1
##        Df Sum of Sq     RSS    AIC F value    Pr(>F)
## <none>              149.886 54.506
## len     1    32.449 117.437 46.991 10.2237 0.0028385 **
## adt     1     0.122 149.764 56.474  0.0302 0.8629277
## trks    1    39.372 110.514 44.622 13.1817 0.0008505 ***
## slim    1    69.508  80.378 32.204 31.9962 1.833e-06 ***
## lwid    1     0.005 149.881 56.505  0.0012 0.9729162
## shld    1    22.438 127.449 50.182  6.5139 0.0149650 *
## itg     1     0.092 149.794 56.482  0.0228 0.8806802
## sigs    1    47.759 102.127 41.543 17.3029 0.0001817 ***
## acpt    1    84.767  65.119 23.994 48.1636 3.408e-08 ***
## lane    1     0.163 149.723 56.464  0.0403 0.8420245
## fai     1     6.460 143.426 54.788  1.6666 0.2047213
## pa      1     3.911 145.975 55.475  0.9914 0.3258704
## ma      1    17.108 132.778 51.780  4.7674 0.0354212 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Model1 <- update(Model1, ~ . + acpt)
add1(Model1, scope = FullModel, test = "F")
```

```
## Single term additions
##
## Model:
## rate ~ acpt
##        Df Sum of Sq    RSS    AIC F value   Pr(>F)
## <none>              65.119 23.994
## len     1   12.9806 52.139 17.323  8.9627 0.004957 **
```

```
## adt     1    3.0871 62.032 24.099  1.7916 0.189124
## trks    1   10.0532 55.066 19.454  6.5724 0.014678 *
## slim    1    7.9430 57.176 20.920  5.0012 0.031618 *
## lwid    1    0.1013 65.018 25.933  0.0561 0.814118
## shld    1    0.8293 64.290 25.494  0.4644 0.499945
## itg     1    2.4664 62.653 24.488  1.4172 0.241655
## sigs    1    7.1603 57.959 21.451  4.4475 0.041975 *
## lane    1    2.4108 62.708 24.522  1.3840 0.247143
## fai     1    0.4259 64.693 25.738  0.2370 0.629322
## pa      1    0.0148 65.104 25.985  0.0082 0.928485
## ma      1    0.4750 64.644 25.708  0.2645 0.610161
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Model1 <- update(Model1, ~ . + len)
add1(Model1, scope = FullModel, test = "F")
```

```
## Single term additions
##
## Model:
## rate ~ acpt + len
##        Df Sum of Sq    RSS    AIC F value Pr(>F)
## <none>              52.139 17.323
## adt     1    0.3062 51.832 19.094  0.2067 0.6521
## trks    1    2.9834 49.155 17.025  2.1242 0.1539
## slim    1    7.2920 44.847 13.448  5.6910 0.0226 *
## lwid    1    0.8546 51.284 18.679  0.5833 0.4502
## shld    1    3.2651 48.873 16.801  2.3383 0.1352
## itg     1    0.2262 51.912 19.154  0.1525 0.6985
## sigs    1    3.4703 48.668 16.637  2.4957 0.1232
## lane    1    0.3814 51.757 19.037  0.2579 0.6148
## fai     1    0.0457 52.093 19.289  0.0307 0.8619
## pa      1    0.4613 51.677 18.977  0.3124 0.5798
## ma      1    0.1771 51.962 19.191  0.1193 0.7319
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Model1 <- update(Model1, ~ . + slim)
add1(Model1, scope = FullModel, test = "F")
```

```
## Single term additions
##
## Model:
## rate ~ acpt + len + slim
##        Df Sum of Sq    RSS    AIC F value Pr(>F)
## <none>              44.847 13.448
## adt     1   0.93309 43.913 14.628  0.7224 0.4013
## trks    1   2.40652 42.440 13.297  1.9279 0.1740
## lwid    1   0.38786 44.459 15.109  0.2966 0.5896
## shld    1   0.01982 44.827 15.431  0.0150 0.9031
## itg     1   0.87112 43.975 14.683  0.6735 0.4175
## sigs    1   2.51322 42.333 13.198  2.0185 0.1645
## lane    1   1.30168 43.545 14.299  1.0164 0.3205
## fai     1   1.47651 43.370 14.142  1.1575 0.2896
## pa      1   1.21471 43.632 14.377  0.9466 0.3375
```

stop here

```
## ma      1   0.00346 44.843 15.445  0.0026 0.9594
```

```
summary(Model1)
```

```
##
## Call:
## lm(formula = rate ~ acpt + len + slim, data = highway)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.0170 -0.8166  0.1136  0.7886  2.3766
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  9.32557    2.61664   3.564 0.001079 **
## acpt         0.10144    0.02726   3.721 0.000693 ***
## len         -0.07712    0.02486  -3.102 0.003787 **
## slim        -0.10240    0.04292  -2.386 0.022597 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.132 on 35 degrees of freedom
## Multiple R-squared:  0.7008, Adjusted R-squared:  0.6751
## F-statistic: 27.33 on 3 and 35 DF,  p-value: 2.756e-09
```

"Traditional" Backward Selection "By Hand"

drop1() shows the partial F-test p-value for each variable. Remove the least significant variable using update() and do drop1() again until all remaining variables is significant at a pre-specified cutoff (0.05 here). As a note, drop1() is not really necessary, because we can get the same information from summary().

```
Model2 <- FullModel
drop1(Model2, test = "F")
```

```
## Single term deletions
##
## Model:
## rate ~ len + adt + trks + slim + lwid + shld + itg + sigs + acpt +
##     lane + fai + pa + ma
##        Df Sum of Sq    RSS    AIC F value  Pr(>F)
## <none>              35.894 24.763
## len     1    5.4061 41.300 28.235  3.7654 0.06368 .
## adt     1    0.0203 35.914 22.785  0.0142 0.90625
## trks    1    1.0941 36.988 23.934  0.7620 0.39100
## slim    1    3.2969 39.191 26.190  2.2963 0.14223
## lwid    1    0.0719 35.966 22.841  0.0501 0.82473
## shld    1    0.0109 35.905 22.775  0.0076 0.93134
## itg     1    0.1973 36.091 22.977  0.1374 0.71400
## sigs    1    2.6508 38.544 25.542  1.8463 0.18635
## acpt    1    3.5130 39.407 26.405  2.4468 0.13034
## lane    1    0.0127 35.906 22.777  0.0088 0.92587
## fai     1    0.1420 36.036 22.917  0.0989 0.75573
## pa      1    1.1976 37.091 24.043  0.8342 0.36980
## ma      1    0.4530 36.347 23.252  0.3155 0.57931
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Model2 <- update(Model2, ~ . -shld)
drop1(Model2, test = "F")
```

```
## Single term deletions
##
## Model:
## rate ~ len + adt + trks + slim + lwid + itg + sigs + acpt + lane +
##     fai + pa + ma
##        Df Sum of Sq    RSS    AIC F value  Pr(>F)
## <none>              35.905 22.775
## len     1    5.6461 41.551 26.471  4.0886 0.05358 .
## adt     1    0.0160 35.921 20.792  0.0116 0.91512
## trks    1    1.3528 37.257 22.217  0.9796 0.33142
## slim    1    7.6597 43.564 28.316  5.5467 0.02634 *
## lwid    1    0.1755 36.080 20.965  0.1271 0.72438
## itg     1    0.1989 36.103 20.990  0.1441 0.70737
## sigs    1    3.0823 38.987 23.987  2.2320 0.14722
## acpt    1    5.1901 41.095 26.040  3.7584 0.06347 .
## lane    1    0.0123 35.917 20.788  0.0089 0.92551
## fai     1    0.1578 36.062 20.946  0.1143 0.73802
## pa      1    1.4020 37.307 22.269  1.0152 0.32293
## ma      1    0.4848 36.389 21.298  0.3511 0.55862
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Model2 <- update(Model2, ~ . -lane)
drop1(Model2, test = "F")
```

```
## Single term deletions
##
## Model:
## rate ~ len + adt + trks + slim + lwid + itg + sigs + acpt + fai +
##     pa + ma
##        Df Sum of Sq    RSS    AIC F value  Pr(>F)
## <none>              35.917 20.788
## len     1    5.6613 41.578 24.497  4.2558 0.04885 *
## adt     1    0.0068 35.924 18.796  0.0051 0.94335
## trks    1    1.4436 37.360 20.325  1.0852 0.30677
## slim    1    7.6961 43.613 26.360  5.7855 0.02328 *
## lwid    1    0.1856 36.102 18.989  0.1395 0.71167
## itg     1    0.2200 36.137 19.026  0.1654 0.68746
## sigs    1    3.3097 39.227 22.226  2.4880 0.12636
## acpt    1    5.2062 41.123 24.067  3.9137 0.05818 .
## fai     1    0.1741 36.091 18.977  0.1308 0.72038
## pa      1    1.3897 37.307 20.269  1.0447 0.31581
## ma      1    0.4973 36.414 19.324  0.3738 0.54603
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Model2 <- update(Model2, ~ . -adt)
drop1(Model2, test = "F")
```

```
## Single term deletions
```

```
##
## Model:
## rate ~ len + trks + slim + lwid + itg + sigs + acpt + fai + pa +
##     ma
##        Df Sum of Sq    RSS    AIC F value  Pr(>F)
## <none>              35.924 18.796
## len     1    5.6571 41.581 22.499  4.4093 0.04488 *
## trks    1    1.4372 37.361 18.325  1.1202 0.29893
## slim    1    7.7285 43.652 24.395  6.0238 0.02059 *
## lwid    1    0.1832 36.107 16.994  0.1428 0.70834
## itg     1    0.4396 36.363 17.270  0.3426 0.56300
## sigs    1    3.3734 39.297 20.296  2.6294 0.11611
## acpt    1    5.2294 41.153 22.096  4.0759 0.05317 .
## fai     1    0.1755 36.099 16.986  0.1368 0.71427
## pa      1    1.5503 37.474 18.443  1.2084 0.28102
## ma      1    0.5000 36.424 17.335  0.3897 0.53750
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Model2 <- update(Model2, ~ . -fai)
drop1(Model2, test = "F")
```

```
## Single term deletions
##
## Model:
## rate ~ len + trks + slim + lwid + itg + sigs + acpt + pa + ma
##        Df Sum of Sq    RSS    AIC F value  Pr(>F)
## <none>              36.099 16.986
## len     1    5.6122 41.711 20.621  4.5085 0.04238 *
## trks    1    1.3763 37.476 16.445  1.1056 0.30171
## slim    1    7.6380 43.737 22.471  6.1359 0.01932 *
## lwid    1    0.1860 36.285 15.186  0.1495 0.70188
## itg     1    0.2754 36.375 15.282  0.2213 0.64160
## sigs    1    3.2337 39.333 18.331  2.5978 0.11785
## acpt    1    5.1980 41.297 20.232  4.1758 0.05018 .
## pa      1    3.5259 39.625 18.620  2.8325 0.10311
## ma      1    0.9744 37.074 16.024  0.7828 0.38356
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Model2 <- update(Model2, ~ . -lwid)
drop1(Model2, test = "F")
```

```
## Single term deletions
##
## Model:
## rate ~ len + trks + slim + itg + sigs + acpt + pa + ma
##        Df Sum of Sq    RSS    AIC F value  Pr(>F)
## <none>              36.285 15.186
## len     1    5.4363 41.722 18.631  4.4947 0.04238 *
## trks    1    1.2995 37.585 14.558  1.0744 0.30824
## slim    1    7.9013 44.187 20.869  6.5326 0.01590 *
## itg     1    0.2611 36.546 13.466  0.2159 0.64557
## sigs    1    3.4668 39.752 16.745  2.8663 0.10081
## acpt    1    5.0490 41.334 18.267  4.1744 0.04991 *
```

```
## pa      1    3.5885 39.874 16.864  2.9669 0.09528 .
## ma      1    0.8738 37.159 14.114  0.7225 0.40207
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Model2 <- update(Model2, ~ . -itg)
drop1(Model2, test = "F")
```

```
## Single term deletions
##
## Model:
## rate ~ len + trks + slim + sigs + acpt + pa + ma
##        Df Sum of Sq    RSS    AIC F value  Pr(>F)
## <none>              36.546 13.466
## len     1    5.1775 41.724 16.633  4.3918 0.04437 *
## trks    1    1.2638 37.810 12.792  1.0720 0.30851
## slim    1    7.7925 44.339 19.004  6.6099 0.01516 *
## sigs    1    3.2074 39.754 14.746  2.7206 0.10916
## acpt    1    5.5731 42.119 17.001  4.7273 0.03744 *
## pa      1    4.0057 40.552 15.522  3.3978 0.07486 .
## ma      1    0.6127 37.159 12.114  0.5198 0.47635
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Model2 <- update(Model2, ~ . -ma)
drop1(Model2, test = "F")
```

```
## Single term deletions
##
## Model:
## rate ~ len + trks + slim + sigs + acpt + pa
##        Df Sum of Sq    RSS    AIC F value  Pr(>F)
## <none>              37.159 12.114
## len     1    6.4378 43.597 16.346  5.5440 0.02484 *
## trks    1    1.0562 38.215 11.207  0.9095 0.34739
## slim    1    7.1968 44.356 17.019  6.1976 0.01818 *
## sigs    1    3.9001 41.059 14.007  3.3586 0.07618 .
## acpt    1    4.9675 42.127 15.008  4.2778 0.04677 *
## pa      1    3.7692 40.928 13.882  3.2459 0.08103 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Model2 <- update(Model2, ~ . -trks)
drop1(Model2, test = "F")
```

```
## Single term deletions
##
## Model:
## rate ~ len + slim + sigs + acpt + pa
##        Df Sum of Sq    RSS    AIC F value  Pr(>F)
## <none>              38.215 11.207
## len     1   10.5678 48.783 18.729  9.1256 0.00484 **
## slim    1    7.4907 45.706 16.188  6.4684 0.01585 *
## sigs    1    5.4166 43.632 14.377  4.6774 0.03791 *
## acpt    1    5.2241 43.439 14.204  4.5112 0.04126 *
## pa      1    4.1181 42.333 13.198  3.5561 0.06816 .
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Model2 <- update(Model2, ~ . -pa)
drop1(Model2, test = "F")
```

```
## Single term deletions
##
## Model:
## rate ~ len + slim + sigs + acpt
##        Df Sum of Sq    RSS    AIC F value   Pr(>F)
## <none>              42.333 13.198
## len     1    9.1881 51.521 18.859  7.3794 0.010299 *
## slim    1    6.3349 48.668 16.637  5.0879 0.030641 *
## sigs    1    2.5132 44.847 13.448  2.0185 0.164500
## acpt    1   12.5355 54.869 21.314 10.0679 0.003195 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Model2 <- update(Model2, ~ . -sigs)
drop1(Model2, test = "F")
```

```
## Single term deletions
##
## Model:
## rate ~ len + slim + acpt
##        Df Sum of Sq    RSS    AIC F value    Pr(>F)
## <none>              44.847 13.448
## len     1    12.330 57.176 20.920  9.6225 0.0037874 **
## slim    1     7.292 52.139 17.323  5.6910 0.0225972 *
## acpt    1    17.744 62.591 24.449 13.8483 0.0006933 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

stop here

```
summary(Model2)
```

```
##
## Call:
## lm(formula = rate ~ len + slim + acpt, data = highway)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.0170 -0.8166  0.1136  0.7886  2.3766
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  9.32557    2.61664   3.564 0.001079 **
## len         -0.07712    0.02486  -3.102 0.003787 **
## slim        -0.10240    0.04292  -2.386 0.022597 *
## acpt         0.10144    0.02726   3.721 0.000693 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.132 on 35 degrees of freedom
## Multiple R-squared:  0.7008, Adjusted R-squared:  0.6751
## F-statistic: 27.33 on 3 and 35 DF,  p-value: 2.756e-09
```

Forward, Backward and Stepwise Selection based on AIC

```
#trace=0 suppresses the step by step output
Model3 <- step(FullModel, direction = "backward", trace = 1)
```

```
## Start:  AIC=24.76
## rate ~ len + adt + trks + slim + lwid + shld + itg + sigs + acpt +
##     lane + fai + pa + ma
##
##        Df Sum of Sq    RSS    AIC
## - shld  1    0.0109 35.905 22.775
## - lane  1    0.0127 35.906 22.777
## - adt   1    0.0203 35.914 22.785
## - lwid  1    0.0719 35.966 22.841
## - fai   1    0.1420 36.036 22.917
## - itg   1    0.1973 36.091 22.977
## - ma    1    0.4530 36.347 23.252
## - trks  1    1.0941 36.988 23.934
## - pa    1    1.1976 37.091 24.043
## <none>               35.894 24.763
## - sigs  1    2.6508 38.544 25.542
## - slim  1    3.2969 39.191 26.190
## - acpt  1    3.5130 39.407 26.405
## - len   1    5.4061 41.300 28.235
##
## Step:  AIC=22.77
## rate ~ len + adt + trks + slim + lwid + itg + sigs + acpt + lane +
##     fai + pa + ma
##
##        Df Sum of Sq    RSS    AIC
## - lane  1    0.0123 35.917 20.788
## - adt   1    0.0160 35.921 20.792
## - fai   1    0.1578 36.062 20.946
## - lwid  1    0.1755 36.080 20.965
## - itg   1    0.1989 36.103 20.990
## - ma    1    0.4848 36.389 21.298
## - trks  1    1.3528 37.257 22.217
## - pa    1    1.4020 37.307 22.269
## <none>               35.905 22.775
## - sigs  1    3.0823 38.987 23.987
## - acpt  1    5.1901 41.095 26.040
## - len   1    5.6461 41.551 26.471
## - slim  1    7.6597 43.564 28.316
##
## Step:  AIC=20.79
## rate ~ len + adt + trks + slim + lwid + itg + sigs + acpt + fai +
##     pa + ma
##
##        Df Sum of Sq    RSS    AIC
## - adt   1    0.0068 35.924 18.796
## - fai   1    0.1741 36.091 18.977
## - lwid  1    0.1856 36.102 18.989
## - itg   1    0.2200 36.137 19.026
```

```
## - ma    1    0.4973 36.414 19.324
## - pa    1    1.3897 37.307 20.269
## - trks  1    1.4436 37.360 20.325
## <none>              35.917 20.788
## - sigs  1    3.3097 39.227 22.226
## - acpt  1    5.2062 41.123 24.067
## - len   1    5.6613 41.578 24.497
## - slim  1    7.6961 43.613 26.360
##
## Step:  AIC=18.8
## rate ~ len + trks + slim + lwid + itg + sigs + acpt + fai + pa +
##     ma
##
##        Df Sum of Sq    RSS    AIC
## - fai   1    0.1755 36.099 16.986
## - lwid  1    0.1832 36.107 16.994
## - itg   1    0.4396 36.363 17.270
## - ma    1    0.5000 36.424 17.335
## - trks  1    1.4372 37.361 18.325
## - pa    1    1.5503 37.474 18.443
## <none>              35.924 18.796
## - sigs  1    3.3734 39.297 20.296
## - acpt  1    5.2294 41.153 22.096
## - len   1    5.6571 41.581 22.499
## - slim  1    7.7285 43.652 24.395
##
## Step:  AIC=16.99
## rate ~ len + trks + slim + lwid + itg + sigs + acpt + pa + ma
##
##        Df Sum of Sq    RSS    AIC
## - lwid  1    0.1860 36.285 15.186
## - itg   1    0.2754 36.375 15.282
## - ma    1    0.9744 37.074 16.024
## - trks  1    1.3763 37.476 16.445
## <none>              36.099 16.986
## - sigs  1    3.2337 39.333 18.331
## - pa    1    3.5259 39.625 18.620
## - acpt  1    5.1980 41.297 20.232
## - len   1    5.6122 41.711 20.621
## - slim  1    7.6380 43.737 22.471
##
## Step:  AIC=15.19
## rate ~ len + trks + slim + itg + sigs + acpt + pa + ma
##
##        Df Sum of Sq    RSS    AIC
## - itg   1    0.2611 36.546 13.466
## - ma    1    0.8738 37.159 14.114
## - trks  1    1.2995 37.585 14.558
## <none>              36.285 15.186
## - sigs  1    3.4668 39.752 16.745
## - pa    1    3.5885 39.874 16.864
## - acpt  1    5.0490 41.334 18.267
## - len   1    5.4363 41.722 18.631
## - slim  1    7.9013 44.187 20.869
```

```
##
## Step:  AIC=13.47
## rate ~ len + trks + slim + sigs + acpt + pa + ma
##
##        Df Sum of Sq    RSS    AIC
## - ma    1    0.6127 37.159 12.114
## - trks  1    1.2638 37.810 12.792
## <none>              36.546 13.466
## - sigs  1    3.2074 39.754 14.746
## - pa    1    4.0057 40.552 15.522
## - len   1    5.1775 41.724 16.633
## - acpt  1    5.5731 42.119 17.001
## - slim  1    7.7925 44.339 19.004
##
## Step:  AIC=12.11
## rate ~ len + trks + slim + sigs + acpt + pa
##
##        Df Sum of Sq    RSS    AIC
## - trks  1    1.0562 38.215 11.207
## <none>              37.159 12.114
## - pa    1    3.7692 40.928 13.882
## - sigs  1    3.9001 41.059 14.007
## - acpt  1    4.9675 42.127 15.008
## - len   1    6.4378 43.597 16.346
## - slim  1    7.1968 44.356 17.019
##
## Step:  AIC=11.21
## rate ~ len + slim + sigs + acpt + pa
##
##        Df Sum of Sq    RSS    AIC
## <none>              38.215 11.207
## - pa    1    4.1181 42.333 13.198
## - acpt  1    5.2241 43.439 14.204
## - sigs  1    5.4166 43.632 14.377
## - slim  1    7.4907 45.706 16.188
## - len   1   10.5678 48.783 18.729
```

```
summary(Model3)
```

```
##
## Call:
## lm(formula = rate ~ len + slim + sigs + acpt + pa, data = highway)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.93807 -0.80142 -0.00392  0.80743  2.44918
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  9.94408    2.58209   3.851 0.000512 ***
## len         -0.07405    0.02451  -3.021 0.004840 **
## slim        -0.10510    0.04132  -2.543 0.015851 *
## sigs         0.79736    0.36868   2.163 0.037907 *
## acpt         0.06428    0.03026   2.124 0.041256 *
## pa          -0.77443    0.41067  -1.886 0.068156 .
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.076 on 33 degrees of freedom
## Multiple R-squared:  0.745,  Adjusted R-squared:  0.7064
## F-statistic: 19.29 on 5 and 33 DF,  p-value: 5.996e-09
```

```
Model4 <- step(NullModel, direction = "forward",
               scope = list( lower = NullModel, upper = FullModel),
               trace = 0)
summary(Model4)
```

```
##
## Call:
## lm(formula = rate ~ acpt + len + slim + sigs + pa, data = highway)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.93807 -0.80142 -0.00392  0.80743  2.44918
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  9.94408    2.58209   3.851 0.000512 ***
## acpt         0.06428    0.03026   2.124 0.041256 *
## len         -0.07405    0.02451  -3.021 0.004840 **
## slim        -0.10510    0.04132  -2.543 0.015851 *
## sigs         0.79736    0.36868   2.163 0.037907 *
## pa          -0.77443    0.41067  -1.886 0.068156 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.076 on 33 degrees of freedom
## Multiple R-squared:  0.745,  Adjusted R-squared:  0.7064
## F-statistic: 19.29 on 5 and 33 DF,  p-value: 5.996e-09
```

same

```
Model5 <- step(NullModel, direction = "both",
               scope = list( lower = NullModel, upper = FullModel),
               trace = 0)
summary(Model5)
```

```
##
## Call:
## lm(formula = rate ~ acpt + len + slim + sigs + pa, data = highway)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.93807 -0.80142 -0.00392  0.80743  2.44918
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  9.94408    2.58209   3.851 0.000512 ***
## acpt         0.06428    0.03026   2.124 0.041256 *
## len         -0.07405    0.02451  -3.021 0.004840 **
## slim        -0.10510    0.04132  -2.543 0.015851 *
## sigs         0.79736    0.36868   2.163 0.037907 *
```

same as other AIC

```
## pa          -0.77443    0.41067  -1.886 0.068156 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.076 on 33 degrees of freedom
## Multiple R-squared:  0.745,  Adjusted R-squared:  0.7064
## F-statistic: 19.29 on 5 and 33 DF,  p-value: 5.996e-09
```