Stepwise Regression Based on AIC Improvement

Rather than building our regression model step by step manually, we can use the `step` function in R - it takes a starting model, and then uses forwards or backwards procedures (or a combination of both) to produce the best model.

We need to install the `MASS` library.

Let's apply the procedure to `model0` and `model1` as our limits - we can specify the stepwise procedure with the parameter "direction":

```
> library (MASS)

> steplimitsboth <- step(model0, scope = list (upper = model1), direction =
"both")
```

- Let’s focus on the combined method that adds predictors which improve model fit, and removes ones that don’t - based on minimising AIC:

```
> summary(steplimitsboth)

Call:
lm(formula = House_price ~ Crime + Population, data = data)

Residuals:
     Min       1Q   Median       3Q      Max 
-27192.2  -6161.4   -555.2   6203.4  24061.0 

Coefficients:
              Estimate Std. Error t value Pr(>|t|)    
(Intercept)  1.736e+05  1.243e+04  13.973  < 2e-16 ***
Crime       -3.343e+02  1.147e+02  -2.915  0.00388 ** 
Population   6.662e-01  2.442e-01   2.729  0.00682 ** 
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 9366 on 247 degrees of freedom
Multiple R-squared:  0.06084,	Adjusted R-squared:  0.05323 
F-statistic:     8 on 2 and 247 DF,  p-value: 0.0004301

> AIC(steplimitsboth)
[1] 5286.855
```

We can see the procedure has settled on the model with Crime and Population. AIC value is 5286.855. In this case the stepwise model is the same as what we arrived at manually.