Here are the outputs and explanation of each commands:

|  |
| --- |
| > ols\_step\_all\_possible(lmod)  Index N Predictors R-Square Adj. R-Square  4 1 1 abdom 0.6621178 0.6607663  3 2 1 chest 0.4940475 0.4920237  5 3 1 hip 0.3915004 0.3890664  1 4 1 weight 0.3759604 0.3734643  6 5 1 thigh 0.3150402 0.3123003  2 6 1 neck 0.2415614 0.2385276  9 7 2 weight abdom 0.7187265 0.7164672  13 8 2 neck abdom 0.6966835 0.6942472  19 9 2 abdom hip 0.6931164 0.6906515  16 10 2 chest abdom 0.6732290 0.6706043  20 11 2 abdom thigh 0.6715991 0.6689614  12 12 2 neck chest 0.5034718 0.4994836  17 13 2 chest hip 0.4998938 0.4958769  18 14 2 chest thigh 0.4990351 0.4950113  8 15 2 weight chest 0.4952243 0.4911698  10 16 2 weight hip 0.3967085 0.3918627  14 17 2 neck hip 0.3936755 0.3888054  21 18 2 hip thigh 0.3915012 0.3866137  11 19 2 weight thigh 0.3793032 0.3743177  7 20 2 weight neck 0.3769908 0.3719867  15 21 2 neck thigh 0.3348112 0.3294684  30 22 3 weight abdom thigh 0.7236399 0.7202969  23 23 3 weight neck abdom 0.7235932 0.7202496  26 24 3 weight chest abdom 0.7187280 0.7153255  29 25 3 weight abdom hip 0.7187273 0.7153248  35 26 3 neck abdom hip 0.7151147 0.7116686  38 27 3 chest abdom hip 0.6995469 0.6959124  36 28 3 neck abdom thigh 0.6989308 0.6952888  32 29 3 neck chest abdom 0.6984257 0.6947777  41 30 3 abdom hip thigh 0.6950646 0.6913759  39 31 3 chest abdom thigh 0.6807157 0.6768533  27 32 3 weight chest hip 0.5247033 0.5189538  34 33 3 neck chest thigh 0.5141260 0.5082485  33 34 3 neck chest hip 0.5141009 0.5082231  28 35 3 weight chest thigh 0.5131574 0.5072682  22 36 3 weight neck chest 0.5036208 0.4976162  40 37 3 chest hip thigh 0.5002576 0.4942123  31 38 3 weight hip thigh 0.3968386 0.3895423  24 39 3 weight neck hip 0.3967218 0.3894240  37 40 3 neck hip thigh 0.3936991 0.3863648  25 41 3 weight neck thigh 0.3800204 0.3725206  46 42 4 weight neck abdom thigh 0.7276750 0.7232649  51 43 4 weight abdom hip thigh 0.7252947 0.7208460  45 44 4 weight neck abdom hip 0.7240225 0.7195532  49 45 4 weight chest abdom thigh 0.7239150 0.7194440  42 46 4 weight neck chest abdom 0.7236976 0.7192230  55 47 4 neck abdom hip thigh 0.7198238 0.7152865  48 48 4 weight chest abdom hip 0.7187299 0.7141749  52 49 4 neck chest abdom hip 0.7161606 0.7115640  56 50 4 chest abdom hip thigh 0.7014961 0.6966620  53 51 4 neck chest abdom thigh 0.7006379 0.6957900  50 52 4 weight chest hip thigh 0.5284036 0.5207664  43 53 4 weight neck chest hip 0.5274308 0.5197778  44 54 4 weight neck chest thigh 0.5201578 0.5123871  54 55 4 neck chest hip thigh 0.5156677 0.5078242  47 56 4 weight neck hip thigh 0.3968542 0.3870867  60 57 5 weight neck abdom hip thigh 0.7311966 0.7257331  58 58 5 weight neck chest abdom thigh 0.7283531 0.7228318  61 59 5 weight chest abdom hip thigh 0.7253709 0.7197890  57 60 5 weight neck chest abdom hip 0.7240481 0.7184394  62 61 5 neck chest abdom hip thigh 0.7206337 0.7149556  59 62 5 weight neck chest hip thigh 0.5315287 0.5220069  63 63 6 weight neck chest abdom hip thigh 0.7314640 0.7248877  Mallow's Cp  4 60.268352  3 213.608032  5 307.167386  1 321.345361  6 376.926242  2 443.964947  9 10.621203  13 30.732228  19 33.986659  16 52.131001  20 53.618018  12 207.009796  17 210.274172  18 211.057588  8 214.534449  10 304.415792  14 307.182919  21 309.166640  11 320.295541  7 322.405277  15 360.887997  30 8.138374  23 8.180981  26 12.619776  29 12.620394  35 15.916358  38 30.119738  36 30.681843  32 31.142662  41 34.209166  39 47.300527  27 189.639085  34 199.289379  33 199.312263  28 200.173050  22 208.873835  40 211.942290  31 306.297052  24 306.403650  37 309.161422  25 321.641245  46 6.456915  51 8.628625  45 9.789331  49 9.887432  42 10.085775  55 13.620048  48 14.618089  52 16.962154  56 30.341381  53 31.124354  50 188.263140  43 189.150708  44 195.786192  54 199.882797  47 308.282818  60 5.243999  58 7.838290  61 10.559131  57 11.765933  62 14.881093  59 187.411972  63 7.000000 |
|  |
| |  | | --- | | > | |

***The above command shows all step possible regression on the 6 predictors. Basically, if there are k potential independent variables then there are 2k possible subset that need to be tested. In this case there are 6 independent variables (predictor), and so, there will be 64 tests that need to be done. We already know that larger the R-square, better the model. Cp is the measure of variance so; the smaller values are desirable. Looking at the data, we can say that model number 60 has the highest R-squared and lowest Cp and hence that will be the model that should be chosen by above analysis.***

|  |
| --- |
| > ols\_step\_forward\_p(lmod,details=TRUE)  Forward Selection Method  ---------------------------  Candidate Terms:  1. weight  2. neck  3. chest  4. abdom  5. hip  6. thigh  We are selecting variables based on p value...  Forward Selection: Step 1  + abdom  Model Summary  --------------------------------------------------------------  R 0.814 RMSE 4.514  R-Squared 0.662 Coef. Var 23.837  Adj. R-Squared 0.661 MSE 20.380  Pred R-Squared 0.652 MAE 3.631  --------------------------------------------------------------  RMSE: Root Mean Square Error  MSE: Mean Square Error  MAE: Mean Absolute Error  ANOVA  -----------------------------------------------------------------------  Sum of  Squares DF Mean Square F Sig.  -----------------------------------------------------------------------  Regression 9984.086 1 9984.086 489.903 0.0000  Residual 5094.931 250 20.380  Total 15079.017 251  -----------------------------------------------------------------------  Parameter Estimates  --------------------------------------------------------------------------------------------  model Beta Std. Error Std. Beta t Sig lower upper  --------------------------------------------------------------------------------------------  (Intercept) -35.197 2.462 -14.294 0.000 -40.046 -30.347  abdom 0.585 0.026 0.814 22.134 0.000 0.533 0.637  --------------------------------------------------------------------------------------------  Forward Selection: Step 2  + weight  Model Summary  --------------------------------------------------------------  R 0.848 RMSE 4.127  R-Squared 0.719 Coef. Var 21.792  Adj. R-Squared 0.716 MSE 17.033  Pred R-Squared 0.709 MAE 3.344  --------------------------------------------------------------  RMSE: Root Mean Square Error  MSE: Mean Square Error  MAE: Mean Absolute Error  ANOVA  -----------------------------------------------------------------------  Sum of  Squares DF Mean Square F Sig.  -----------------------------------------------------------------------  Regression 10837.688 2 5418.844 318.13 0.0000  Residual 4241.328 249 17.033  Total 15079.017 251  -----------------------------------------------------------------------  Parameter Estimates  --------------------------------------------------------------------------------------------  model Beta Std. Error Std. Beta t Sig lower upper  --------------------------------------------------------------------------------------------  (Intercept) -41.348 2.413 -17.136 0.000 -46.101 -36.596  abdom 0.915 0.053 1.273 17.419 0.000 0.812 1.019  weight -0.136 0.019 -0.517 -7.079 0.000 -0.174 -0.098  --------------------------------------------------------------------------------------------  Forward Selection: Step 3  + thigh  Model Summary  --------------------------------------------------------------  R 0.851 RMSE 4.099  R-Squared 0.724 Coef. Var 21.645  Adj. R-Squared 0.720 MSE 16.803  Pred R-Squared 0.712 MAE 3.327  --------------------------------------------------------------  RMSE: Root Mean Square Error  MSE: Mean Square Error  MAE: Mean Absolute Error  ANOVA  -----------------------------------------------------------------------  Sum of  Squares DF Mean Square F Sig.  -----------------------------------------------------------------------  Regression 10911.778 3 3637.259 216.46 0.0000  Residual 4167.238 248 16.803  Total 15079.017 251  -----------------------------------------------------------------------  Parameter Estimates  --------------------------------------------------------------------------------------------  model Beta Std. Error Std. Beta t Sig lower upper  --------------------------------------------------------------------------------------------  (Intercept) -48.039 3.987 -12.049 0.000 -55.891 -40.186  abdom 0.917 0.052 1.276 17.578 0.000 0.815 1.020  weight -0.170 0.025 -0.643 -6.834 0.000 -0.219 -0.121  thigh 0.209 0.100 0.142 2.100 0.037 0.013 0.405  --------------------------------------------------------------------------------------------  Forward Selection: Step 4  + neck  Model Summary  --------------------------------------------------------------  R 0.853 RMSE 4.077  R-Squared 0.728 Coef. Var 21.530  Adj. R-Squared 0.723 MSE 16.625  Pred R-Squared 0.715 MAE 3.289  --------------------------------------------------------------  RMSE: Root Mean Square Error  MSE: Mean Square Error  MAE: Mean Absolute Error  ANOVA  -----------------------------------------------------------------------  Sum of  Squares DF Mean Square F Sig.  -----------------------------------------------------------------------  Regression 10972.624 4 2743.156 165.001 0.0000  Residual 4106.393 247 16.625  Total 15079.017 251  -----------------------------------------------------------------------  Parameter Estimates  -------------------------------------------------------------------------------------------  model Beta Std. Error Std. Beta t Sig lower upper  -------------------------------------------------------------------------------------------  (Intercept) -38.283 6.460 -5.926 0.000 -51.006 -25.559  abdom 0.924 0.052 1.285 17.757 0.000 0.821 1.026  weight -0.144 0.028 -0.545 -5.106 0.000 -0.199 -0.088  thigh 0.191 0.099 0.130 1.924 0.055 -0.005 0.387  neck -0.366 0.191 -0.115 -1.913 0.057 -0.743 0.011  -------------------------------------------------------------------------------------------  Forward Selection: Step 5  + hip  Model Summary  --------------------------------------------------------------  R 0.855 RMSE 4.059  R-Squared 0.731 Coef. Var 21.433  Adj. R-Squared 0.726 MSE 16.477  Pred R-Squared 0.716 MAE 3.274  --------------------------------------------------------------  RMSE: Root Mean Square Error  MSE: Mean Square Error  MAE: Mean Absolute Error  ANOVA  -----------------------------------------------------------------------  Sum of  Squares DF Mean Square F Sig.  -----------------------------------------------------------------------  Regression 11025.726 5 2205.145 133.833 0.0000  Residual 4053.291 246 16.477  Total 15079.017 251  -----------------------------------------------------------------------  Parameter Estimates  ------------------------------------------------------------------------------------------  model Beta Std. Error Std. Beta t Sig lower upper  ------------------------------------------------------------------------------------------  (Intercept) -26.393 9.232 -2.859 0.005 -44.576 -8.210  abdom 0.955 0.055 1.328 17.492 0.000 0.847 1.062  weight -0.109 0.034 -0.415 -3.226 0.001 -0.176 -0.043  thigh 0.290 0.113 0.197 2.562 0.011 0.067 0.514  neck -0.459 0.197 -0.144 -2.324 0.021 -0.848 -0.070  hip -0.233 0.130 -0.215 -1.795 0.074 -0.488 0.023  ------------------------------------------------------------------------------------------  No more variables to be added.  Variables Entered:  + abdom  + weight  + thigh  + neck  + hip  Final Model Output  ------------------  Model Summary  --------------------------------------------------------------  R 0.855 RMSE 4.059  R-Squared 0.731 Coef. Var 21.433  Adj. R-Squared 0.726 MSE 16.477  Pred R-Squared 0.716 MAE 3.274  --------------------------------------------------------------  RMSE: Root Mean Square Error  MSE: Mean Square Error  MAE: Mean Absolute Error  ANOVA  -----------------------------------------------------------------------  Sum of  Squares DF Mean Square F Sig.  -----------------------------------------------------------------------  Regression 11025.726 5 2205.145 133.833 0.0000  Residual 4053.291 246 16.477  Total 15079.017 251  -----------------------------------------------------------------------  Parameter Estimates  ------------------------------------------------------------------------------------------  model Beta Std. Error Std. Beta t Sig lower upper  ------------------------------------------------------------------------------------------  (Intercept) -26.393 9.232 -2.859 0.005 -44.576 -8.210  abdom 0.955 0.055 1.328 17.492 0.000 0.847 1.062  weight -0.109 0.034 -0.415 -3.226 0.001 -0.176 -0.043  thigh 0.290 0.113 0.197 2.562 0.011 0.067 0.514  neck -0.459 0.197 -0.144 -2.324 0.021 -0.848 -0.070  hip -0.233 0.130 -0.215 -1.795 0.074 -0.488 0.023  ------------------------------------------------------------------------------------------  Selection Summary  --------------------------------------------------------------------------  Variable Adj.  Step Entered R-Square R-Square C(p) AIC RMSE  --------------------------------------------------------------------------  1 abdom 0.6621 0.6608 60.2684 1478.8012 4.5144  2 weight 0.7187 0.7165 10.6212 1434.5921 4.1272  3 thigh 0.7236 0.7203 8.1384 1432.1511 4.0992  4 neck 0.7277 0.7233 6.4569 1430.4445 4.0774  5 hip 0.7312 0.7257 5.2440 1429.1646 4.0592  -------------------------------------------------------------------------- |
|  |
| |  | | --- | | > | |

***The way forward selection works is that, as the analysis goes, it will keep adding variables to the model and have analysis (ANOVA table, parameter estimation etc.) for it. We can analyze it at each step and pick the best model. The basic criteria of choosing the best model is to look for the one that has smallest MSE, Mallow’s Cp or AIC, and largest R-squared value. Looking at the above output, the best model is the one containing the variables (predictors) weight, neck, abdom, hip, thigh, as it fits the criteria. In the backward step regression, it uses backward elimination technique, where it eliminates variables one by one. The same (kind of) output will be generated in the backward selection model. The same concept applies to stepwise regression where all variables are entered in a stepwise manner until no variables are left to enter. The same criteria apply to all of the above-mentioned regression methods. Note that, these three, forward, backward, and stepwise regression methods, the predictors are added and/or deleted based on their p values.***

***Looking at all outcomes, it looks like the model containing variables weight, neck, abdom, hip, thigh is the best one to choose.***