Puffin

Hartaig Singh

September 8, 2017

Import Data Set

puffin <- read.table("puffin.txt",header = T)

Multiple Linear Regression

model1 <- lm(nesting~.,puffin)  
summary(model1)

##   
## Call:  
## lm(formula = nesting ~ ., data = puffin)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -4.0166 -2.1088 0.2293 1.2505 6.9881   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 10.117840 3.185028 3.177 0.00323 \*\*   
## grass -0.007408 0.019459 -0.381 0.70586   
## soil 0.209211 0.077238 2.709 0.01062 \*   
## angle 0.082389 0.077796 1.059 0.29727   
## distance -0.366571 0.057473 -6.378 3.18e-07 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 2.647 on 33 degrees of freedom  
## Multiple R-squared: 0.8792, Adjusted R-squared: 0.8645   
## F-statistic: 60.03 on 4 and 33 DF, p-value: 1.113e-14

**Interpretation**

*Grass:* For every percent increase in grass covered, the number of nests for every 9m^2 decreases by .007 on average holding all other variables constant

*Soil:* For every cm increase in mean soil depth, the number of nests for every 9m^2 increases by .21 on average holding all other variables constant

*Angle:* For every degree increase in angle of slope, the number of nests for every 9m^2 increases by .082 on average holding all other variables constant

*Distance:* For every meter increase in distance from edge, the number of nests for every 9m^2 decreases by .367 on average holding all other variables constant

*R^2:* 87.92% of the variation in number of nests is explained by the predictors grass, soil, angle, and distance

Residual Plot

resid(model1)

## 1 2 3 4 5 6   
## -4.01660727 -3.03636552 0.53950437 -2.86622617 0.62314898 -0.02854612   
## 7 8 9 10 11 12   
## -1.34563187 0.01382399 0.25197081 3.91370685 -3.57731769 3.22078352   
## 13 14 15 16 17 18   
## 0.32385289 0.57899805 2.80687713 1.28851327 -2.35822494 0.99019380   
## 19 20 21 22 23 24   
## 0.20654609 -2.11925911 6.98806138 1.10202778 1.13664065 3.74255804   
## 25 26 27 28 29 30   
## -2.12431115 -3.04595358 -1.89783084 -1.62440818 0.69624086 2.02317028   
## 31 32 33 34 35 36   
## -2.07725967 3.43878484 2.15540692 -0.95093663 -3.49539157 -2.44933205   
## 37 38   
## -1.14033024 2.11312207

Studentized Residuals

rstandard(model1)

## 1 2 3 4 5   
## -1.636072313 -1.251923929 0.222927886 -1.172371252 0.300402446   
## 6 7 8 9 10   
## -0.011526946 -0.536500636 0.005765554 0.106151443 1.631354785   
## 11 12 13 14 15   
## -1.449986156 1.322476646 0.133129956 0.232151787 1.100440096   
## 16 17 18 19 20   
## 0.505409776 -0.923741187 0.396433812 0.081869427 -0.847347881   
## 21 22 23 24 25   
## 2.865571943 0.437621398 0.468245590 1.554733566 -0.838199676   
## 26 27 28 29 30   
## -1.190998896 -0.756636802 -0.705362747 0.285020864 0.826023136   
## 31 32 33 34 35   
## -0.844247789 1.361894548 0.839921377 -0.371677871 -1.355376236   
## 36 37 38   
## -0.993616273 -0.462568589 0.843205298

Deleted studentized residuals

rstudent(model1)

## 1 2 3 4 5   
## -1.680697239 -1.263171095 0.219689691 -1.179291412 0.296221180   
## 6 7 8 9 10   
## -0.011350975 -0.530628500 0.005677528 0.104548566 1.675425259   
## 11 12 13 14 15   
## -1.475628017 1.338229672 0.131132537 0.228794182 1.104085601   
## 16 17 18 19 20   
## 0.499630614 -0.921631019 0.391313944 0.080627627 -0.843638731   
## 21 22 23 24 25   
## 3.255826129 0.432195692 0.462635834 1.590348548 -0.834331332   
## 26 27 28 29 30   
## -1.198863733 -0.751632778 -0.699889325 0.281015249 0.821953168   
## 31 32 33 34 35   
## -0.840483945 1.380456121 0.836082485 -0.366771560 -1.373458489   
## 36 37 38   
## -0.993418746 -0.456990014 0.839423235

Difference between studentized residuals and deleted studentized residual

rstandard(model1)-rstudent(model1)

## 1 2 3 4 5   
## 4.462493e-02 1.124717e-02 3.238194e-03 6.920160e-03 4.181267e-03   
## 6 7 8 9 10   
## -1.759714e-04 -5.872136e-03 8.802603e-05 1.602877e-03 -4.407047e-02   
## 11 12 13 14 15   
## 2.564186e-02 -1.575303e-02 1.997419e-03 3.357605e-03 -3.645506e-03   
## 16 17 18 19 20   
## 5.779162e-03 -2.110167e-03 5.119868e-03 1.241800e-03 -3.709150e-03   
## 21 22 23 24 25   
## -3.902542e-01 5.425706e-03 5.609756e-03 -3.561498e-02 -3.868343e-03   
## 26 27 28 29 30   
## 7.864837e-03 -5.004024e-03 -5.473422e-03 4.005615e-03 4.069968e-03   
## 31 32 33 34 35   
## -3.763844e-03 -1.856157e-02 3.838892e-03 -4.906311e-03 1.808225e-02   
## 36 37 38   
## -1.975264e-04 -5.578576e-03 3.782064e-03

which.max(abs(rstandard(model1)-rstudent(model1)))

## 21   
## 21

Obersvation 21 has the largest difference between studentized residuals and deleted studentized residuals. This means that observation 21 may be influential since its presence changes the predicted value the most in the data set.

DFFITS

dffits(model1)

## 1 2 3 4 5   
## -0.678300663 -0.552780867 0.097435594 -0.489968542 0.234896855   
## 6 7 8 9 10   
## -0.004289905 -0.179263773 0.002657959 0.051636903 0.781867881   
## 11 12 13 14 15   
## -0.574344052 0.570412153 0.056313177 0.081464094 0.306956151   
## 16 17 18 19 20   
## 0.139847685 -0.253106389 0.137483528 0.025646545 -0.292827694   
## 21 22 23 24 25   
## 1.375854767 0.140213745 0.201370817 0.728004949 -0.251971894   
## 26 27 28 29 30   
## -0.320748668 -0.253846288 -0.396873842 0.117415978 0.337237614   
## 31 32 33 34 35   
## -0.333798309 0.435058228 0.212021258 -0.097540966 -0.318702094   
## 36 37 38   
## -0.389120675 -0.178923189 0.285902387

which(dffits(model1)>3\*sqrt((4+1)/(38-4-1)))

## 21   
## 21

Observation 21 has DFFIT greater than the specified cutoff

DFBETAS

dfbetas(model1)

## (Intercept) grass soil angle distance  
## 1 -0.010925482 0.1657259173 -0.091745493 -0.177394590 0.1149261143  
## 2 0.301103830 -0.0746277001 -0.221443826 -0.221348831 -0.0399122887  
## 3 0.091022898 -0.0340417160 -0.048013600 -0.029941300 -0.0348998440  
## 4 -0.413960930 0.3618832829 0.067417064 0.192890680 0.2349991370  
## 5 0.062306852 -0.1325543835 0.186061660 -0.196694663 -0.1940552648  
## 6 0.001596876 -0.0008565932 -0.003145047 0.001679342 0.0012659952  
## 7 0.064542428 -0.0524464997 -0.109395852 0.062458570 0.0336678762  
## 8 0.001290045 -0.0009869870 -0.001661313 0.000546077 0.0009241504  
## 9 0.001100887 -0.0348927413 0.008909895 0.001369672 0.0088105292  
## 10 -0.061599732 -0.1560258099 -0.261407119 0.317474423 0.5080008342  
## 11 -0.283328571 0.2795162726 -0.057883115 0.097901897 0.3087969400  
## 12 -0.288919917 0.0657237805 0.152223025 0.291360019 0.0949873701  
## 13 -0.008825616 0.0442967583 -0.030311119 0.030663796 0.0239922992  
## 14 0.072072527 -0.0156331005 -0.043685272 -0.024445058 -0.0227119482  
## 15 0.084757525 0.1735674157 -0.139160326 -0.012556628 0.0224810057  
## 16 0.014545685 0.0539324274 0.020461073 -0.062544589 -0.0320810597  
## 17 -0.124526293 -0.0431322566 0.140143028 0.027779214 -0.0362109910  
## 18 0.009377186 0.0668774755 -0.078853927 0.027484367 0.0705287148  
## 19 -0.003385649 -0.0033213385 -0.006478999 0.009794380 0.0158972416  
## 20 0.012082185 -0.0272425335 0.030830711 -0.127393467 -0.0068267251  
## 21 -0.656543610 0.0151413266 0.667264610 0.404487049 -0.0544354516  
## 22 -0.009833117 0.0977540330 -0.065490023 0.063261318 0.0415336868  
## 23 0.002139639 0.0325461229 0.125491230 -0.125328884 -0.1297901611  
## 24 0.623607572 -0.5518445561 0.004519881 -0.444003783 -0.4311817263  
## 25 -0.057004416 -0.1333891845 0.046814635 0.072463694 0.0265745676  
## 26 -0.110389884 0.2023849290 -0.031428444 0.080995936 0.0228197282  
## 27 0.136371945 -0.0982489519 -0.102619219 0.010081328 -0.0541098831  
## 28 0.111943400 0.2182298103 -0.258152580 0.062580902 0.0152773770  
## 29 0.027830648 -0.0185743584 -0.046027157 0.055475546 0.0127796352  
## 30 0.005679502 -0.0418971025 -0.091888509 0.196927349 0.0731050796  
## 31 -0.172658071 -0.1249062291 0.084694715 0.130186559 0.1290708067  
## 32 0.216249630 0.0991338887 -0.005251762 -0.248571664 -0.2240820235  
## 33 -0.015135843 0.0851755637 0.071609940 -0.076297600 -0.0600099191  
## 34 -0.022771648 -0.0498954642 0.042722445 0.004367610 -0.0164047225  
## 35 0.130705094 -0.1355216148 0.005461837 -0.089475458 -0.1586625132  
## 36 0.227656831 -0.2784597624 0.106181756 -0.240308804 -0.2990719828  
## 37 0.109566940 -0.0435921220 0.017271319 -0.116456034 -0.1429383460  
## 38 -0.109748015 0.0323454566 0.006929675 0.065095953 0.1600855781

Removing observation 21 would decrease the coefficients for grass, soil, angle, and increase for distance

To generalize, if an observation has a positive DFBETA, removing the observation would decrease the beta value. The opposite holds for a negative DFBETA.

Measures of Influence

influence.measures(model1)

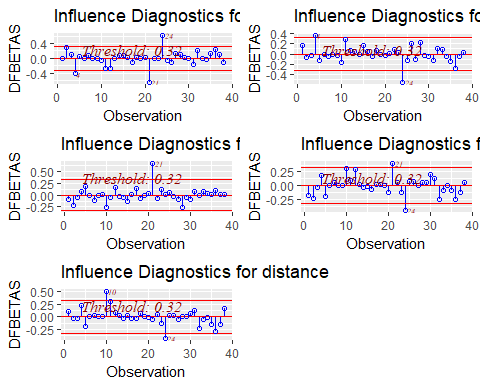
## Influence measures of  
## lm(formula = nesting ~ ., data = puffin) :  
##   
## dfb.1\_ dfb.grss dfb.soil dfb.angl dfb.dstn dffit cov.r cook.d  
## 1 -0.01093 0.165726 -0.09175 -0.177395 0.114926 -0.67830 0.889 8.72e-02  
## 2 0.30110 -0.074628 -0.22144 -0.221349 -0.039912 -0.55278 1.090 6.00e-02  
## 3 0.09102 -0.034042 -0.04801 -0.029941 -0.034900 0.09744 1.385 1.96e-03  
## 4 -0.41396 0.361883 0.06742 0.192891 0.234999 -0.48997 1.106 4.75e-02  
## 5 0.06231 -0.132554 0.18606 -0.196695 -0.194055 0.23490 1.874 1.13e-02  
## 6 0.00160 -0.000857 -0.00315 0.001679 0.001266 -0.00429 1.333 3.80e-06  
## 7 0.06454 -0.052446 -0.10940 0.062459 0.033668 -0.17926 1.244 6.57e-03  
## 8 0.00129 -0.000987 -0.00166 0.000546 0.000924 0.00266 1.422 1.46e-06  
## 9 0.00110 -0.034893 0.00891 0.001370 0.008811 0.05164 1.448 5.50e-04  
## 10 -0.06160 -0.156026 -0.26141 0.317474 0.508001 0.78187 0.933 1.16e-01  
## 11 -0.28333 0.279516 -0.05788 0.097902 0.308797 -0.57434 0.966 6.37e-02  
## 12 -0.28892 0.065724 0.15222 0.291360 0.094987 0.57041 1.050 6.36e-02  
## 13 -0.00883 0.044297 -0.03031 0.030664 0.023992 0.05631 1.378 6.54e-04  
## 14 0.07207 -0.015633 -0.04369 -0.024445 -0.022712 0.08146 1.303 1.37e-03  
## 15 0.08476 0.173567 -0.13916 -0.012557 0.022481 0.30696 1.042 1.87e-02  
## 16 0.01455 0.053932 0.02046 -0.062545 -0.032081 0.13985 1.210 4.00e-03  
## 17 -0.12453 -0.043132 0.14014 0.027779 -0.036211 -0.25311 1.100 1.29e-02  
## 18 0.00938 0.066877 -0.07885 0.027484 0.070529 0.13748 1.279 3.88e-03  
## 19 -0.00339 -0.003321 -0.00648 0.009794 0.015897 0.02565 1.283 1.36e-04  
## 20 0.01208 -0.027243 0.03083 -0.127393 -0.006827 -0.29283 1.171 1.73e-02  
## 21 -0.65654 0.015141 0.66726 0.404487 -0.054435 1.37585 0.329 2.93e-01  
## 22 -0.00983 0.097754 -0.06549 0.063261 0.041534 0.14021 1.252 4.03e-03  
## 23 0.00214 0.032546 0.12549 -0.125329 -0.129790 0.20137 1.342 8.31e-03  
## 24 0.62361 -0.551845 0.00452 -0.444004 -0.431182 0.72800 0.964 1.01e-01  
## 25 -0.05700 -0.133389 0.04681 0.072464 0.026575 -0.25197 1.143 1.28e-02  
## 26 -0.11039 0.202385 -0.03143 0.080996 0.022820 -0.32075 1.003 2.03e-02  
## 27 0.13637 -0.098249 -0.10262 0.010081 -0.054110 -0.25385 1.190 1.31e-02  
## 28 0.11194 0.218230 -0.25815 0.062581 0.015277 -0.39687 1.429 3.20e-02  
## 29 0.02783 -0.018574 -0.04603 0.055476 0.012780 0.11742 1.353 2.84e-03  
## 30 0.00568 -0.041897 -0.09189 0.196927 0.073105 0.33724 1.227 2.30e-02  
## 31 -0.17266 -0.124906 0.08469 0.130187 0.129071 -0.33380 1.211 2.25e-02  
## 32 0.21625 0.099134 -0.00525 -0.248572 -0.224082 0.43506 0.960 3.68e-02  
## 33 -0.01514 0.085176 0.07161 -0.076298 -0.060010 0.21202 1.114 9.07e-03  
## 34 -0.02277 -0.049895 0.04272 0.004368 -0.016405 -0.09754 1.223 1.95e-03  
## 35 0.13071 -0.135522 0.00546 -0.089475 -0.158663 -0.31870 0.923 1.98e-02  
## 36 0.22766 -0.278460 0.10618 -0.240309 -0.299072 -0.38912 1.156 3.03e-02  
## 37 0.10957 -0.043592 0.01727 -0.116456 -0.142938 -0.17892 1.302 6.56e-03  
## 38 -0.10975 0.032345 0.00693 0.065096 0.160086 0.28590 1.167 1.65e-02  
## hat inf  
## 1 0.1401   
## 2 0.1607   
## 3 0.1644   
## 4 0.1472   
## 5 0.3861 \*  
## 6 0.1250   
## 7 0.1024   
## 8 0.1798   
## 9 0.1961   
## 10 0.1788   
## 11 0.1316   
## 12 0.1537   
## 13 0.1557   
## 14 0.1125   
## 15 0.0717   
## 16 0.0727   
## 17 0.0701   
## 18 0.1099   
## 19 0.0919   
## 20 0.1075   
## 21 0.1515 \*  
## 22 0.0952   
## 23 0.1593   
## 24 0.1732   
## 25 0.0836   
## 26 0.0668   
## 27 0.1024   
## 28 0.2433   
## 29 0.1486   
## 30 0.1441   
## 31 0.1362   
## 32 0.0903   
## 33 0.0604   
## 34 0.0661   
## 35 0.0511   
## 36 0.1330   
## 37 0.1329   
## 38 0.1039

Observations 5 and 21 are flagged as being influential

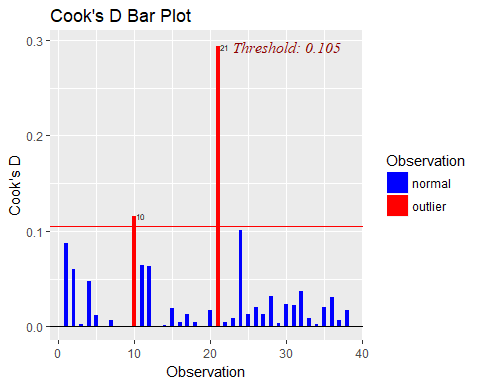
Obersvation 5 is flagged because it has a hat value > 3\*p/n and observation 21 is flagged because of its DFFIT value > 3(sqrt((4+1)/(38-4-1)))

Outlier Analysis

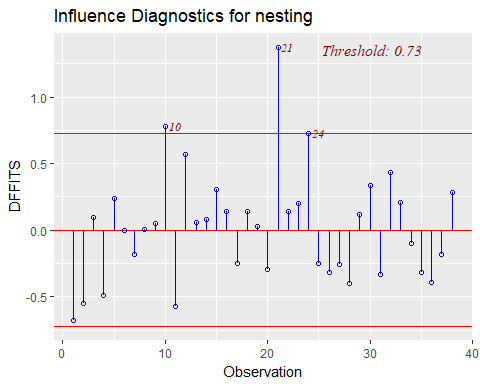
ols\_dfbetas\_panel(model1)



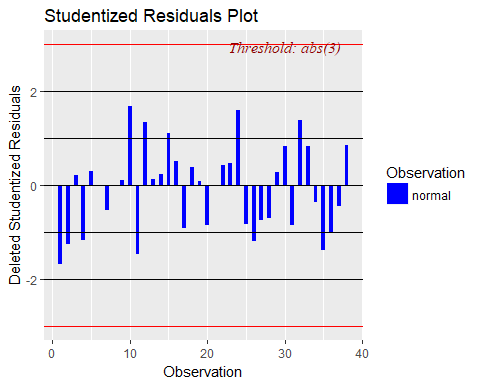
ols\_cooksd\_barplot(model1)



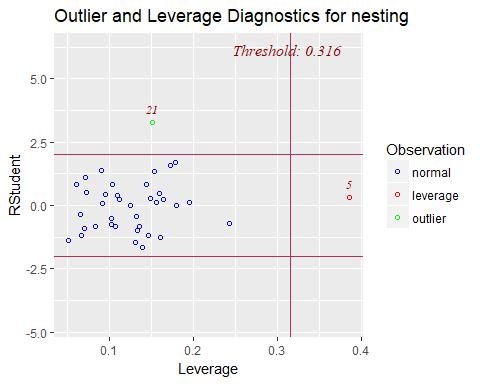
ols\_dffits\_plot(model1)



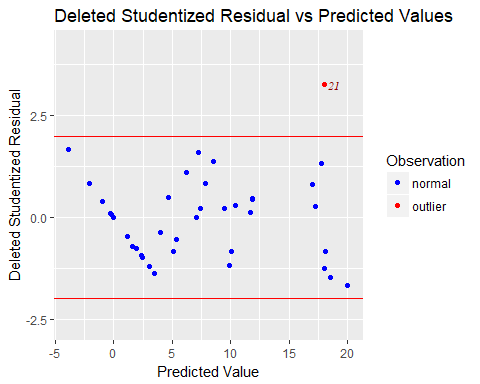
ols\_srsd\_plot(model1)



ols\_rsdlev\_plot(model1)



ols\_dsrvsp\_plot(model1)



I would consider observations 5, 21, 10, 24 as influential because observation 24 has fairly high DFBETAS for grass, angle, and distance and borderline high DFFITS

Observation 10 has high DFBETA for distance, high cooks distance, and high DFFITS

Observation 5 also has high leverage