Exercicio\_1.R

a56843646

Sat Dec 16 09:09:10 2017

#Questao 3  
  
hbat400<-read.csv(file="./hbat\_clima.csv")  
fit.min.hbat400<- glm(evasao~1,family = binomial(link = "logit"),data=hbat400)  
  
stp.hbat400<-step(fit.min.hbat400,  
 direction="both",  
 scope=(~  
 JS1 + JS2 + JS3 + JS4 + JS5 +   
 OC1 + OC2 + OC3 + OC4 +   
 EP1 + EP2 + EP3 + EP4 +   
 AC1 + AC2 + AC3 + AC4 +   
 SI1 + SI2 + SI3 + SI4  
 ))

## Start: AIC=494.03  
## evasao ~ 1  
##   
## Df Deviance AIC  
## + SI4 1 442.19 446.19  
## + SI1 1 447.25 451.25  
## + SI2 1 453.08 457.08  
## + SI3 1 458.27 462.27  
## + EP2 1 479.82 483.82  
## + OC2 1 480.33 484.33  
## + EP4 1 481.71 485.71  
## + EP3 1 482.56 486.56  
## + OC3 1 483.27 487.27  
## + OC4 1 483.36 487.36  
## + EP1 1 483.85 487.85  
## + AC4 1 485.42 489.42  
## + AC3 1 485.48 489.48  
## + AC1 1 486.83 490.83  
## <none> 492.03 494.03  
## + OC1 1 490.60 494.60  
## + JS1 1 490.93 494.93  
## + AC2 1 491.07 495.07  
## + JS3 1 491.38 495.38  
## + JS2 1 492.01 496.01  
## + JS4 1 492.02 496.02  
## + JS5 1 492.02 496.02  
##   
## Step: AIC=446.19  
## evasao ~ SI4  
##   
## Df Deviance AIC  
## + SI1 1 433.20 439.20  
## + SI2 1 438.03 444.03  
## + SI3 1 439.15 445.15  
## + JS5 1 439.22 445.22  
## <none> 442.19 446.19  
## + JS2 1 440.54 446.54  
## + JS4 1 440.76 446.76  
## + AC1 1 441.05 447.05  
## + OC3 1 441.11 447.11  
## + AC3 1 441.25 447.25  
## + AC4 1 441.30 447.30  
## + EP4 1 441.92 447.92  
## + JS3 1 441.96 447.96  
## + EP3 1 441.97 447.97  
## + EP2 1 442.04 448.04  
## + EP1 1 442.06 448.06  
## + AC2 1 442.07 448.07  
## + OC2 1 442.11 448.11  
## + JS1 1 442.18 448.18  
## + OC4 1 442.19 448.19  
## + OC1 1 442.19 448.19  
## - SI4 1 492.03 494.03  
##   
## Step: AIC=439.2  
## evasao ~ SI4 + SI1  
##   
## Df Deviance AIC  
## + JS5 1 429.86 437.86  
## + JS2 1 431.18 439.18  
## <none> 433.20 439.20  
## + SI3 1 431.67 439.67  
## + JS4 1 431.85 439.85  
## + AC4 1 432.51 440.51  
## + SI2 1 432.55 440.55  
## + AC3 1 432.61 440.61  
## + AC1 1 432.64 440.64  
## + OC3 1 432.71 440.71  
## + JS3 1 432.91 440.91  
## + AC2 1 432.96 440.96  
## + OC4 1 433.05 441.05  
## + OC1 1 433.11 441.11  
## + OC2 1 433.13 441.13  
## + EP3 1 433.16 441.16  
## + EP1 1 433.19 441.19  
## + EP2 1 433.20 441.20  
## + EP4 1 433.20 441.20  
## + JS1 1 433.20 441.20  
## - SI1 1 442.19 446.19  
## - SI4 1 447.25 451.25  
##   
## Step: AIC=437.86  
## evasao ~ SI4 + SI1 + JS5  
##   
## Df Deviance AIC  
## <none> 429.86 437.86  
## + SI3 1 428.39 438.39  
## + JS1 1 428.40 438.40  
## + AC4 1 428.94 438.94  
## + AC1 1 429.03 439.03  
## + OC3 1 429.07 439.07  
## + SI2 1 429.13 439.13  
## - JS5 1 433.20 439.20  
## + AC3 1 429.24 439.24  
## + JS2 1 429.60 439.60  
## + AC2 1 429.67 439.67  
## + EP3 1 429.69 439.69  
## + JS4 1 429.76 439.76  
## + JS3 1 429.78 439.78  
## + OC1 1 429.78 439.78  
## + OC4 1 429.79 439.79  
## + EP1 1 429.79 439.79  
## + OC2 1 429.83 439.83  
## + EP2 1 429.84 439.84  
## + EP4 1 429.85 439.85  
## - SI1 1 439.22 445.22  
## - SI4 1 445.59 451.59

summary(stp.hbat400)

##   
## Call:  
## glm(formula = evasao ~ SI4 + SI1 + JS5, family = binomial(link = "logit"),   
## data = hbat400)  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -1.6170 -0.8363 -0.5655 1.0423 2.2941   
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -5.516995 0.878738 -6.278 3.42e-10 \*\*\*  
## SI4 0.681819 0.178001 3.830 0.000128 \*\*\*  
## SI1 0.640242 0.218880 2.925 0.003444 \*\*   
## JS5 -0.010668 0.005873 -1.816 0.069295 .   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for binomial family taken to be 1)  
##   
## Null deviance: 492.03 on 399 degrees of freedom  
## Residual deviance: 429.86 on 396 degrees of freedom  
## AIC: 437.86  
##   
## Number of Fisher Scoring iterations: 5

#Questao 4  
summary(glm(evasao~SI1,family = binomial(link = "logit"),data=hbat400))

##   
## Call:  
## glm(formula = evasao ~ SI1, family = binomial(link = "logit"),   
## data = hbat400)  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -1.1206 -0.7234 -0.7234 1.2353 2.1801   
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -5.4946 0.8370 -6.564 5.23e-11 \*\*\*  
## SI1 1.0719 0.1848 5.800 6.62e-09 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for binomial family taken to be 1)  
##   
## Null deviance: 492.03 on 399 degrees of freedom  
## Residual deviance: 447.25 on 398 degrees of freedom  
## AIC: 451.25  
##   
## Number of Fisher Scoring iterations: 4

summary(glm(evasao~SI2,family = binomial(link = "logit"),data=hbat400))

##   
## Call:  
## glm(formula = evasao ~ SI2, family = binomial(link = "logit"),   
## data = hbat400)  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -1.0944 -0.7388 -0.7388 1.2628 2.1133   
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -5.0014 0.7920 -6.315 2.70e-10 \*\*\*  
## SI2 0.9606 0.1752 5.484 4.15e-08 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for binomial family taken to be 1)  
##   
## Null deviance: 492.03 on 399 degrees of freedom  
## Residual deviance: 453.08 on 398 degrees of freedom  
## AIC: 457.08  
##   
## Number of Fisher Scoring iterations: 4

summary(glm(evasao~SI3,family = binomial(link = "logit"),data=hbat400))

##   
## Call:  
## glm(formula = evasao ~ SI3, family = binomial(link = "logit"),   
## data = hbat400)  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -1.2182 -0.7211 -0.7211 1.1372 2.2807   
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -3.1785 0.4526 -7.022 2.18e-12 \*\*\*  
## SI3 0.6548 0.1182 5.540 3.03e-08 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for binomial family taken to be 1)  
##   
## Null deviance: 492.03 on 399 degrees of freedom  
## Residual deviance: 458.27 on 398 degrees of freedom  
## AIC: 462.27  
##   
## Number of Fisher Scoring iterations: 3

summary(glm(evasao~SI4,family = binomial(link = "logit"),data=hbat400))

##   
## Call:  
## glm(formula = evasao ~ SI4, family = binomial(link = "logit"),   
## data = hbat400)  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -1.3627 -0.9766 -0.6607 1.0028 2.2009   
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -4.1656 0.5607 -7.429 1.10e-13 \*\*\*  
## SI4 0.9183 0.1454 6.316 2.68e-10 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for binomial family taken to be 1)  
##   
## Null deviance: 492.03 on 399 degrees of freedom  
## Residual deviance: 442.19 on 398 degrees of freedom  
## AIC: 446.19  
##   
## Number of Fisher Scoring iterations: 4

cor(hbat400[,c(16,18,21)])

## SI1 SI2 SI3  
## SI1 1.0000000 0.7327186 0.5800353  
## SI2 0.7327186 1.0000000 0.6249696  
## SI3 0.5800353 0.6249696 1.0000000