|  |
| --- |
| > # Loading the data  > library(readxl)  > healthdata <- read\_excel("1555054100\_hospitalcosts.xlsx")  > View(healthdata)  > head(healthdata)  # A tibble: 6 x 6  AGE FEMALE LOS RACE TOTCHG APRDRG  *<dbl>* *<dbl>* *<dbl>* *<dbl>* *<dbl>* *<dbl>*  1 17 1 2 1 2660 560  2 17 0 2 1 1689 753  3 17 1 7 1 20060 930  4 17 1 1 1 736 758  5 17 1 1 1 1194 754  6 17 0 0 1 3305 347  > #EDA  > str(healthdata)  Classes ‘tbl\_df’, ‘tbl’ and 'data.frame': 500 obs. of 6 variables:  $ AGE : num 17 17 17 17 17 17 17 16 16 17 ...  $ FEMALE: num 1 0 1 1 1 0 1 1 1 1 ...  $ LOS : num 2 2 7 1 1 0 4 2 1 2 ...  $ RACE : num 1 1 1 1 1 1 1 1 1 1 ...  $ TOTCHG: num 2660 1689 20060 736 1194 ...  $ APRDRG: num 560 753 930 758 754 347 754 754 753 758 ...  > summary(healthdata)  AGE FEMALE LOS RACE TOTCHG APRDRG  Min. : 0.000 Min. :0.000 Min. : 0.000 Min. :1.000 Min. : 532 Min. : 21.0  1st Qu.: 0.000 1st Qu.:0.000 1st Qu.: 2.000 1st Qu.:1.000 1st Qu.: 1216 1st Qu.:640.0  Median : 0.000 Median :1.000 Median : 2.000 Median :1.000 Median : 1536 Median :640.0  Mean : 5.086 Mean :0.512 Mean : 2.828 Mean :1.078 Mean : 2774 Mean :616.4  3rd Qu.:13.000 3rd Qu.:1.000 3rd Qu.: 3.000 3rd Qu.:1.000 3rd Qu.: 2530 3rd Qu.:751.0  Max. :17.000 Max. :1.000 Max. :41.000 Max. :6.000 Max. :48388 Max. :952.0  NA's :1  > # Missing value imputation  > sum(is.na(healthdata$RACE)==TRUE) # only one missing value  [1] 1  > table(healthdata$RACE) # Maximunm people belongs to the RACE ,1  1 2 3 4 5 6  484 6 1 3 3 2  > healthdata$RACE[is.na(healthdata$RACE)==TRUE] <- 1  > summary(healthdata)  AGE FEMALE LOS RACE TOTCHG APRDRG  Min. : 0.000 Min. :0.000 Min. : 0.000 Min. :1.000 Min. : 532 Min. : 21.0  1st Qu.: 0.000 1st Qu.:0.000 1st Qu.: 2.000 1st Qu.:1.000 1st Qu.: 1216 1st Qu.:640.0  Median : 0.000 Median :1.000 Median : 2.000 Median :1.000 Median : 1536 Median :640.0  Mean : 5.086 Mean :0.512 Mean : 2.828 Mean :1.078 Mean : 2774 Mean :616.4  3rd Qu.:13.000 3rd Qu.:1.000 3rd Qu.: 3.000 3rd Qu.:1.000 3rd Qu.: 2530 3rd Qu.:751.0  Max. :17.000 Max. :1.000 Max. :41.000 Max. :6.000 Max. :48388 Max. :952.0  > #Univariate analysis (boxplot,density plot)  > boxplot(healthdata$AGE) # No outliers  > bx <-boxplot(healthdata$LOS) #Outliers are present  > quantile(healthdata$LOS, seq(0,1,0.02)) # 4% is 1 and 96% is 7  0% 2% 4% 6% 8% 10% 12% 14% 16% 18% 20% 22% 24% 26% 28% 30% 32%  0.00 0.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 2.00 2.00 2.00 2.00 2.00 2.00 2.00  34% 36% 38% 40% 42% 44% 46% 48% 50% 52% 54% 56% 58% 60% 62% 64% 66%  2.00 2.00 2.00 2.00 2.00 2.00 2.00 2.00 2.00 2.00 2.00 2.00 2.00 2.00 2.00 3.00 3.00  68% 70% 72% 74% 76% 78% 80% 82% 84% 86% 88% 90% 92% 94% 96% 98% 100%  3.00 3.00 3.00 3.00 3.00 3.00 3.00 3.00 4.00 4.00 4.00 4.00 5.00 6.00 7.00 10.04 41.00  > bx$stats  [,1]  [1,] 1  [2,] 2  [3,] 2  [4,] 3  [5,] 4  > healthdata$LOS <- ifelse(healthdata$LOS >7, 7,healthdata$LOS)  > healthdata$LOS <- ifelse(healthdata$LOS <1, 1,healthdata$LOS)  > bx1<- boxplot(healthdata$TOTCHG)  > bx1$stats  [,1]  [1,] 532.0  [2,] 1215.5  [3,] 1536.5  [4,] 2530.5  [5,] 4412.0  > quantile(healthdata$TOTCHG, seq(0,1,0.02)) # 96% is 10588  0% 2% 4% 6% 8% 10% 12% 14% 16% 18% 20%  532.00 615.98 628.00 742.58 850.12 933.60 1084.76 1096.00 1110.20 1134.82 1156.80  22% 24% 26% 28% 30% 32% 34% 36% 38% 40% 42%  1173.78 1197.80 1233.96 1263.72 1274.40 1281.00 1290.66 1309.64 1344.96 1381.60 1405.74  44% 46% 48% 50% 52% 54% 56% 58% 60% 62% 64%  1419.12 1445.08 1500.04 1536.50 1604.92 1671.30 1721.64 1768.42 1805.80 1878.04 1937.68  66% 68% 70% 72% 74% 76% 78% 80% 82% 84% 86%  2024.70 2093.20 2182.80 2242.76 2417.12 2557.08 2783.16 2932.80 3124.36 3591.36 3916.54  88% 90% 92% 94% 96% 98% 100%  4443.80 5886.30 7439.16 8660.88 10587.56 14175.38 48388.00  > healthdata$TOTCHG <- ifelse(healthdata$TOTCHG >=10587,10587,healthdata$TOTCHG)  > bx2<- boxplot(healthdata$APRDRG)  > bx2$stats  [,1]  [1,] 560  [2,] 640  [3,] 640  [4,] 751  [5,] 911  > quantile(healthdata$APRDRG, seq(0,1,0.02)) # 4% is 96 and 96% is 758  0% 2% 4% 6% 8% 10% 12% 14% 16% 18% 20% 22% 24% 26%  21.00 53.00 96.80 139.00 247.08 341.30 422.00 612.32 633.00 639.00 640.00 640.00 640.00 640.00  28% 30% 32% 34% 36% 38% 40% 42% 44% 46% 48% 50% 52% 54%  640.00 640.00 640.00 640.00 640.00 640.00 640.00 640.00 640.00 640.00 640.00 640.00 640.00 640.00  56% 58% 60% 62% 64% 66% 68% 70% 72% 74% 76% 78% 80% 82%  640.00 640.00 640.00 640.00 640.00 640.00 640.00 640.00 712.80 751.00 753.00 753.00 753.00 753.00  84% 86% 88% 90% 92% 94% 96% 98% 100%  754.00 754.00 754.00 754.00 755.00 758.00 758.00 776.70 952.00  > healthdata$APRDRG <- ifelse(healthdata$APRDRG >=758,758,healthdata$APRDRG)  > healthdata$APRDRG <- ifelse(healthdata$APRDRG <=96,96,healthdata$APRDRG)  > plot(density(healthdata$AGE)) # right skewed, mean greater than median  > plot(density(healthdata$LOS))  > plot(density(healthdata$TOTCHG)) #right skewed  > plot(density(healthdata$APRDRG)) #left skewed  > #Bivariate analysis  > scatter.smooth(healthdata$AGE, healthdata$TOTCHG)  > # Converting to factors  > healthdata\_copy <- healthdata  > # Question-1 :-age category of people who frequent the hospital and has the maximum expenditure  > healthdata$AGE <-as.factor(healthdata$AGE)  > summary(healthdata$AGE)  0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17  307 10 1 3 2 2 2 3 2 2 4 8 15 18 25 29 29 38  > table(healthdata$AGE)  0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17  307 10 1 3 2 2 2 3 2 2 4 8 15 18 25 29 29 38  > # 307 patients, from age group 0-1, which is the maximum  > hist(healthdata$AGE)  > hist(healthdata\_copy$AGE)  > library("dplyr")  Attaching package: ‘dplyr’  The following objects are masked from ‘package:stats’:  filter, lag  The following objects are masked from ‘package:base’:  intersect, setdiff, setequal, union  > healthdata %>% group\_by(AGE) %>% summarise(total\_expenditure=sum(TOTCHG)) %>% arrange(desc(total\_expenditure))  # A tibble: 18 x 2  AGE total\_expenditure  *<fct>* *<dbl>*  1 0 632711  2 17 123916  3 15 93133  4 16 69149  5 14 64474  6 12 46775  7 1 37744  8 13 31135  9 3 26894  10 9 21147  11 5 18507  12 6 17928  13 10 17532  14 4 15992  15 11 14250  16 7 10087  17 2 7298  18 8 4741  > # Question-2 diagnosis-related group that has maximum hospitalization and expenditure  > healthdata %>% group\_by(APRDRG) %>% summarise(total\_stay=sum(LOS),total\_exp=sum(TOTCHG)) %>%  + arrange(desc(total\_stay,total\_exp))  # A tibble: 46 x 3  APRDRG total\_stay total\_exp  *<dbl>* *<dbl>* *<dbl>*  1 640 651 437978  2 753 99 78444  3 758 96 100969  4 754 87 59150  5 96 49 148273  6 751 32 21666  7 626 25 23227  8 614 21 25006  9 636 20 23224  10 639 18 12612  # ... with 36 more rows  > The diagnosis group, 640 has the maximum hospitalization and expenditure  > Question-3 needs to analyze if the race of the patient is related to the hospitalization costs  > healthdata %>% group\_by(RACE) %>% summarise(race\_cost=sum(TOTCHG))  # A tibble: 6 x 2  RACE race\_cost  *<dbl>* *<dbl>*  1 1 1209347  2 2 25213  3 3 3041  4 4 7034  5 5 6080  6 6 2698  > hist(healthdata$RACE)  > race\_cost <-cor(healthdata$RACE,healthdata$TOTCHG)  > race\_cost  [1] -0.01198213  > Question-4 analyze the severity of the hosp costs by age and gender for the proper allocation of resources.  > healthdata %>% group\_by(FEMALE,AGE) %>% summarise(tot\_cost=sum(TOTCHG)) %>% arrange(desc(tot\_cost))  # A tibble: 31 x 3  # Groups: FEMALE [2]  FEMALE AGE tot\_cost  *<dbl>* *<fct>* *<dbl>*  1 0 0 351968  2 1 0 280743  3 1 17 76008  4 0 15 54154  5 0 17 47908  6 1 14 41679  7 1 16 41370  8 1 15 38979  9 0 1 34622  10 1 12 31222  # ... with 21 more rows  > hosp\_female\_age=healthdata %>% group\_by(FEMALE,AGE) %>% summarise(tot\_cost=sum(TOTCHG)) %>% arrange(desc(tot\_cost))  > # age group, 0-1 & Male has the most hospital cost  > plot(hosp\_female\_age)  > #Question-5 if the length of stay can be predicted from age, gender, and race.  > # Ho - LOS is independent of age, gender and race  > # H1 - LOS is dependent on age, gender and race  > dev.off()  null device  1  > par(mfrow=c(1,3))  > scatter.smooth(healthdata$AGE, healthdata$LOS)  > scatter.smooth(healthdata$FEMALE, healthdata$LOS)  There were 20 warnings (use warnings() to see them)  > scatter.smooth(healthdata$RACE,healthdata$LOS)  There were 35 warnings (use warnings() to see them)  > mod1=lm(healthdata$LOS ~ AGE +FEMALE +RACE, data=healthdata)  > mod1  Call:  lm(formula = healthdata$LOS ~ AGE + FEMALE + RACE, data = healthdata)  Coefficients:  (Intercept) AGE1 AGE2 AGE3 AGE4 AGE5 AGE6 AGE7  2.51986 -0.79265 -0.55777 0.71743 -0.64497 -0.14497 -0.05777 -1.55777  AGE8 AGE9 AGE10 AGE11 AGE12 AGE13 AGE14 AGE15  -1.07672 -0.55777 0.39863 -1.10611 -0.26241 -0.47751 0.13118 0.15425  AGE16 AGE17 FEMALE RACE  -0.77420 -0.01461 0.17440 0.03791  > summary(mod1)  Call:  lm(formula = healthdata$LOS ~ AGE + FEMALE + RACE, data = healthdata)  Residuals:  Min 1Q Median 3Q Max  -2.2752 -0.7322 -0.5578 0.4422 4.7453  Coefficients:  Estimate Std. Error t value Pr(>|t|)  (Intercept) 2.51986 0.17283 14.580 < 2e-16 \*\*\*  AGE1 -0.79265 0.46419 -1.708 0.08835 .  AGE2 -0.55777 1.44401 -0.386 0.69947  AGE3 0.71743 0.83583 0.858 0.39113  AGE4 -0.64497 1.02184 -0.631 0.52822  AGE5 -0.14497 1.02184 -0.142 0.88724  AGE6 -0.05777 1.02367 -0.056 0.95502  AGE7 -1.55777 0.83794 -1.859 0.06363 .  AGE8 -1.07672 1.02483 -1.051 0.29395  AGE9 -0.55777 1.02367 -0.545 0.58609  AGE10 0.39863 0.72541 0.550 0.58290  AGE11 -1.10611 0.51654 -2.141 0.03275 \*  AGE12 -0.26241 0.38153 -0.688 0.49193  AGE13 -0.47751 0.35250 -1.355 0.17617  AGE14 0.13118 0.30447 0.431 0.66677  AGE15 0.15425 0.28128 0.548 0.58369  AGE16 -0.77420 0.28477 -2.719 0.00679 \*\*  AGE17 -0.01461 0.24949 -0.059 0.95332  FEMALE 0.17440 0.13672 1.276 0.20270  RACE 0.03791 0.12648 0.300 0.76451  ---  Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1  Residual standard error: 1.44 on 480 degrees of freedom  Multiple R-squared: 0.0497, Adjusted R-squared: 0.01208  F-statistic: 1.321 on 19 and 480 DF, p-value: 0.1639  > step(mod1)  Start: AIC=384.45  healthdata$LOS ~ AGE + FEMALE + RACE  Df Sum of Sq RSS AIC  - AGE 17 47.946 1043.72 373.97  - RACE 1 0.186 995.96 382.55  - FEMALE 1 3.376 999.15 384.15  <none> 995.77 384.45  Step: AIC=373.97  healthdata$LOS ~ FEMALE + RACE  Df Sum of Sq RSS AIC  - RACE 1 0.0021 1043.7 371.97  - FEMALE 1 4.1302 1047.8 373.94  <none> 1043.7 373.97  Step: AIC=371.97  healthdata$LOS ~ FEMALE  Df Sum of Sq RSS AIC  - FEMALE 1 4.1291 1047.8 371.94  <none> 1043.7 371.97  Step: AIC=371.94  healthdata$LOS ~ 1  Call:  lm(formula = healthdata$LOS ~ 1, data = healthdata)  Coefficients:  (Intercept)  2.548  > cor(healthdata\_copy)  AGE FEMALE LOS RACE TOTCHG APRDRG  AGE 1.00000000 0.23397895 -0.046677757 0.016063637 0.13275329 0.14329907  FEMALE 0.23397895 1.00000000 0.062773500 -0.038693127 -0.11327798 0.25528993  LOS -0.04667776 0.06277350 1.000000000 -0.001000405 0.60272759 0.01947788  RACE 0.01606364 -0.03869313 -0.001000405 1.000000000 -0.01198213 -0.04377122  TOTCHG 0.13275329 -0.11327798 0.602727592 -0.011982132 1.00000000 -0.51545941  APRDRG 0.14329907 0.25528993 0.019477878 -0.043771220 -0.51545941 1.00000000  > Question-6 agency wants to find the variable that mainly affects hospital costs.  > healthmod1 <- lm(TOTCHG ~ ., healthdata)  > healthmod1  Call:  lm(formula = TOTCHG ~ ., data = healthdata)  Coefficients:  (Intercept) AGE1 AGE2 AGE3 AGE4 AGE5 AGE6 AGE7  3975.279 -296.830 2036.385 5313.249 4191.538 3644.949 3240.168 -143.662  AGE8 AGE9 AGE10 AGE11 AGE12 AGE13 AGE14 AGE15  208.569 6121.939 211.047 633.543 1559.560 1187.110 1002.441 1275.108  AGE16 AGE17 FEMALE LOS RACE APRDRG  1653.871 1408.295 -371.441 1065.313 -168.073 -7.044  > summary(healthmod1)  Call:  lm(formula = TOTCHG ~ ., data = healthdata)  Residuals:  Min 1Q Median 3Q Max  -4382.2 -524.1 -59.2 244.9 5762.1  Coefficients:  Estimate Std. Error t value Pr(>|t|)  (Intercept) 3975.2795 304.8195 13.041 < 2e-16 \*\*\*  AGE1 -296.8301 417.4013 -0.711 0.477345  AGE2 2036.3854 1219.0202 1.671 0.095473 .  AGE3 5313.2487 697.5642 7.617 1.40e-13 \*\*\*  AGE4 4191.5378 862.2601 4.861 1.59e-06 \*\*\*  AGE5 3644.9494 876.1814 4.160 3.77e-05 \*\*\*  AGE6 3240.1682 874.6424 3.705 0.000236 \*\*\*  AGE7 -143.6615 719.0772 -0.200 0.841733  AGE8 208.5685 856.0638 0.244 0.807617  AGE9 6121.9389 866.5842 7.064 5.73e-12 \*\*\*  AGE10 211.0469 610.5346 0.346 0.729739  AGE11 633.5426 432.1533 1.466 0.143302  AGE12 1559.5595 317.8393 4.907 1.27e-06 \*\*\*  AGE13 1187.1102 297.7045 3.988 7.72e-05 \*\*\*  AGE14 1002.4412 254.6525 3.937 9.50e-05 \*\*\*  AGE15 1275.1078 234.4485 5.439 8.58e-08 \*\*\*  AGE16 1653.8713 239.5223 6.905 1.61e-11 \*\*\*  AGE17 1408.2946 207.7340 6.779 3.57e-11 \*\*\*  FEMALE -371.4412 115.3926 -3.219 0.001374 \*\*  LOS 1065.3133 38.0198 28.020 < 2e-16 \*\*\*  RACE -168.0729 105.6133 -1.591 0.112181  APRDRG -7.0439 0.3965 -17.763 < 2e-16 \*\*\*  ---  Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1  Residual standard error: 1199 on 478 degrees of freedom  Multiple R-squared: 0.7722, Adjusted R-squared: 0.7622  F-statistic: 77.16 on 21 and 478 DF, p-value: < 2.2e-16  > step(healthmod1)  Start: AIC=7110.68  TOTCHG ~ AGE + FEMALE + LOS + RACE + APRDRG  Df Sum of Sq RSS AIC  <none> 687078094 7110.7  - RACE 1 3640292 690718386 7111.3  - FEMALE 1 14893685 701971779 7119.4  - AGE 17 390085100 1077163194 7301.5  - APRDRG 1 453557650 1140635744 7362.1  - LOS 1 1128529069 1815607163 7594.5  Call:  lm(formula = TOTCHG ~ AGE + FEMALE + LOS + RACE + APRDRG, data = healthdata)  Coefficients:  (Intercept) AGE1 AGE2 AGE3 AGE4 AGE5 AGE6 AGE7  3975.279 -296.830 2036.385 5313.249 4191.538 3644.949 3240.168 -143.662  AGE8 AGE9 AGE10 AGE11 AGE12 AGE13 AGE14 AGE15  208.569 6121.939 211.047 633.543 1559.560 1187.110 1002.441 1275.108  AGE16 AGE17 FEMALE LOS RACE APRDRG  1653.871 1408.295 -371.441 1065.313 -168.073 -7.044  > healthmod2 <- lm(TOTCHG ~ AGE+FEMALE+LOS+APRDRG,healthdata)  > healthmod2  Call:  lm(formula = TOTCHG ~ AGE + FEMALE + LOS + APRDRG, data = healthdata)  Coefficients:  (Intercept) AGE1 AGE2 AGE3 AGE4 AGE5 AGE6 AGE7  3762.261 -263.307 2077.905 5332.730 4220.280 3682.899 3281.520 -107.614  AGE8 AGE9 AGE10 AGE11 AGE12 AGE13 AGE14 AGE15  148.380 6157.799 236.862 627.557 1569.376 1162.296 988.119 1278.616  AGE16 AGE17 FEMALE LOS APRDRG  1619.799 1418.112 -363.653 1064.664 -6.995  > summary(healthmod2)  Call:  lm(formula = TOTCHG ~ AGE + FEMALE + LOS + APRDRG, data = healthdata)  Residuals:  Min 1Q Median 3Q Max  -4358.9 -513.3 -61.5 248.0 5762.6  Coefficients:  Estimate Std. Error t value Pr(>|t|)  (Intercept) 3762.262 274.295 13.716 < 2e-16 \*\*\*  AGE1 -263.307 417.536 -0.631 0.528588  AGE2 2077.905 1220.689 1.702 0.089359 .  AGE3 5332.730 698.572 7.634 1.24e-13 \*\*\*  AGE4 4220.280 863.449 4.888 1.39e-06 \*\*\*  AGE5 3682.899 877.257 4.198 3.21e-05 \*\*\*  AGE6 3281.520 875.654 3.748 0.000200 \*\*\*  AGE7 -107.614 719.869 -0.149 0.881229  AGE8 148.380 856.595 0.173 0.862551  AGE9 6157.799 867.676 7.097 4.62e-12 \*\*\*  AGE10 236.862 611.295 0.387 0.698576  AGE11 627.557 432.828 1.450 0.147741  AGE12 1569.376 318.287 4.931 1.13e-06 \*\*\*  AGE13 1162.296 297.771 3.903 0.000108 \*\*\*  AGE14 988.119 254.900 3.876 0.000121 \*\*\*  AGE15 1278.616 234.813 5.445 8.28e-08 \*\*\*  AGE16 1619.799 238.945 6.779 3.57e-11 \*\*\*  AGE17 1418.112 207.974 6.819 2.78e-11 \*\*\*  FEMALE -363.653 115.473 -3.149 0.001739 \*\*  LOS 1064.664 38.078 27.960 < 2e-16 \*\*\*  APRDRG -6.995 0.396 -17.665 < 2e-16 \*\*\*  ---  Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1  Residual standard error: 1201 on 479 degrees of freedom  Multiple R-squared: 0.771, Adjusted R-squared: 0.7614  F-statistic: 80.63 on 20 and 479 DF, p-value: < 2.2e-16  > library(car)  Loading required package: carData  Attaching package: ‘car’  The following object is masked from ‘package:dplyr’:  recode  > vif(healthmod2) # No mulitcollinearity , variance of regrssion coeff is not inflated due the correlation between  GVIF Df GVIF^(1/(2\*Df))  AGE 1.728256 17 1.016222  FEMALE 1.155197 1 1.074801  LOS 1.053635 1 1.026467  APRDRG 1.571916 1 1.253761  > par(mfrow=c(2,2))  > plot(healthmod2)  > durbinWatsonTest(healthmod2)  lag Autocorrelation D-W Statistic p-value  1 0.03815566 1.920708 0.346  Alternative hypothesis: rho != 0  > hist(residuals(healthmod2))  > plot(healthdata$TOTCHG, residuals(healthmod2))  > dev.off()  null device  1  > plot(healthdata$TOTCHG, residuals(healthmod2))  > library(predictmeans)  Loading required package: lme4  Loading required package: Matrix  Registered S3 methods overwritten by 'lme4':  method from  cooks.distance.influence.merMod car  influence.merMod car  dfbeta.influence.merMod car  dfbetas.influence.merMod car  Loading required package: nlme  Attaching package: ‘nlme’  The following object is masked from ‘package:lme4’:  lmList  The following object is masked from ‘package:dplyr’:  collapse  Loading required package: parallel  > cooksd=CookD(healthmod2)  Error in LOObeta - beta0 : non-numeric argument to binary operator  > # Non constant variance test  > ncvTest(healthmod2) # pvalue is less, so Ho is rejected and variance is not constant  Non-constant Variance Score Test  Variance formula: ~ fitted.values  Chisquare = 287.7652, Df = 1, p = < 2.22e-16 |
|  |
| |  | | --- | | > | |