MVA Project

2022-05-14

##读取文档  
Heart1 = read.table("C:/Users/zhuan/Desktop/清华大学/2022春 大三下/多元统计分析/庄成 2019080344 MVA大作业/heart\_2020\_cleaned.csv", header = T, sep = ",")  
Heart = read.table("C:/Users/zhuan/Desktop/清华大学/2022春 大三下/多元统计分析/庄成 2019080344 MVA大作业/heart\_2020\_cleaned\_processed.csv", header = T, sep = ",")  
Heartstd = scale(Heart[,])

head(Heartstd)

## HeartDisease BMI Smoking AlcoholDrinking Stroke  
## [1,] -0.3059536 -1.84474727 1.1934717 -0.2703193 -0.19804  
## [2,] -0.3059536 -1.25633616 -0.8378891 -0.2703193 5.04947  
## [3,] -0.3059536 -0.27460211 1.1934717 -0.2703193 -0.19804  
## [4,] -0.3059536 -0.64747225 -0.8378891 -0.2703193 -0.19804  
## [5,] -0.3059536 -0.72613683 -0.8378891 -0.2703193 -0.19804  
## [6,] 3.2684598 0.08568169 1.1934717 -0.2703193 -0.19804  
## PhysicalHealth MentalHealth DiffWalking Sex AgeCategory Race  
## [1,] -0.04675098 3.2810637 -0.4015777 -0.9517094 0.1361843 -0.4748376  
## [2,] -0.42406912 -0.4900378 -0.4015777 -0.9517094 1.5388035 -0.4748376  
## [3,] 2.09138515 3.2810637 -0.4015777 1.0507376 0.6972320 -0.4748376  
## [4,] -0.42406912 -0.4900378 -0.4015777 -0.9517094 1.2582796 -0.4748376  
## [5,] 3.09756685 -0.4900378 2.4901704 -0.9517094 -0.7053873 -0.4748376  
## [6,] 0.33056716 -0.4900378 2.4901704 -0.9517094 1.2582796 0.2965224  
## Diabetic PhysicalActivity GenHealth SleepTime Asthma KidneyDisease  
## [1,] 2.25151 0.5382552 0.3883066 -1.46035124 2.5415108 -0.1955541  
## [2,] -0.41046 0.5382552 0.3883066 -0.06760042 -0.3934655 -0.1955541  
## [3,] 2.25151 0.5382552 -1.5293897 0.62877498 2.5415108 -0.1955541  
## [4,] -0.41046 -1.8578489 -0.5705416 -0.76397583 -0.3934655 -0.1955541  
## [5,] -0.41046 0.5382552 0.3883066 0.62877498 -0.3934655 -0.1955541  
## [6,] -0.41046 -1.8578489 -1.5293897 3.41427661 -0.3934655 -0.1955541  
## SkinCancer  
## [1,] 3.1184143  
## [2,] -0.3206748  
## [3,] -0.3206748  
## [4,] 3.1184143  
## [5,] -0.3206748  
## [6,] -0.3206748

library(psych)  
p = ncol(Heart); describe(Heart[,c(1:(p-2))])

## vars n mean sd median trimmed mad min max range skew  
## HeartDisease 1 319795 0.09 0.28 0.00 0.00 0.00 0.00 1 1.00 2.96  
## BMI 2 319795 0.30 0.07 0.29 0.29 0.06 0.13 1 0.87 1.33  
## Smoking 3 319795 0.41 0.49 0.00 0.39 0.00 0.00 1 1.00 0.36  
## AlcoholDrinking 4 319795 0.07 0.25 0.00 0.00 0.00 0.00 1 1.00 3.43  
## Stroke 5 319795 0.04 0.19 0.00 0.00 0.00 0.00 1 1.00 4.85  
## PhysicalHealth 6 319795 0.11 0.27 0.00 0.03 0.00 0.00 1 1.00 2.60  
## MentalHealth 7 319795 0.13 0.27 0.00 0.06 0.00 0.00 1 1.00 2.33  
## DiffWalking 8 319795 0.14 0.35 0.00 0.05 0.00 0.00 1 1.00 2.09  
## Sex 9 319795 0.48 0.50 0.00 0.47 0.00 0.00 1 1.00 0.10  
## AgeCategory 10 319795 0.58 0.27 0.62 0.59 0.34 0.08 1 0.92 -0.26  
## Race 11 319795 0.27 0.22 0.17 0.21 0.00 0.17 1 0.83 2.10  
## Diabetic 12 319795 0.08 0.19 0.00 0.03 0.00 0.00 1 1.00 2.36  
## PhysicalActivity 13 319795 0.78 0.42 1.00 0.84 0.00 0.00 1 1.00 -1.32  
## GenHealth 14 319795 0.72 0.21 0.80 0.73 0.30 0.20 1 0.80 -0.45  
## SleepTime 15 319795 0.30 0.06 0.29 0.30 0.06 0.04 1 0.96 0.68  
## Asthma 16 319795 0.13 0.34 0.00 0.04 0.00 0.00 1 1.00 2.15  
## kurtosis se  
## HeartDisease 6.78 0  
## BMI 3.89 0  
## Smoking -1.87 0  
## AlcoholDrinking 9.76 0  
## Stroke 21.54 0  
## PhysicalHealth 5.53 0  
## MentalHealth 4.40 0  
## DiffWalking 2.36 0  
## Sex -1.99 0  
## AgeCategory -1.02 0  
## Race 3.28 0  
## Diabetic 4.98 0  
## PhysicalActivity -0.26 0  
## GenHealth -0.33 0  
## SleepTime 7.85 0  
## Asthma 2.61 0

##数据集可视化  
library(GGally)

## 载入需要的程辑包：ggplot2

##   
## 载入程辑包：'ggplot2'

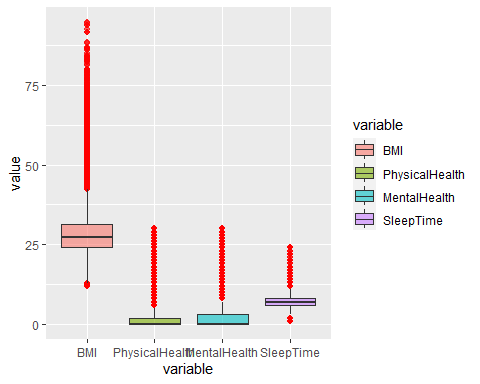
## The following objects are masked from 'package:psych':  
##   
## %+%, alpha

## Registered S3 method overwritten by 'GGally':  
## method from   
## +.gg ggplot2

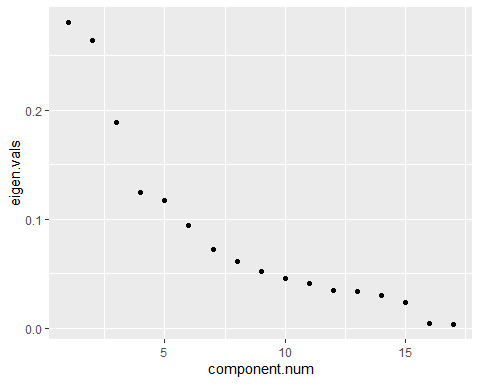
library(corrgram)  
library(reshape2)  
#cov(Heart[,-1])  
#corrgram(Heart, order=TRUE, main="correlation between Heart Disease Factors", lower.panel=panel.pts, upper.panel = panel.cor)  
melt.data=melt(Heart1)

## Using HeartDisease, Smoking, AlcoholDrinking, Stroke, DiffWalking, Sex, AgeCategory, Race, Diabetic, PhysicalActivity, GenHealth, Asthma, KidneyDisease, SkinCancer as id variables

ggplot(data = melt.data, aes(x=variable, y=value, fill=variable))+geom\_boxplot(alpha = 0.6, outlier.color = "red", outlier.shape = 20, outlier.size = 3)



##未标准化数据  
library(ggplot2)  
Heartpca = prcomp(Heart[,2:18])  
evals <- data.frame(Heartpca$sdev^2)  
names(evals) <-'eigen.vals'  
evals$component.num <- as.integer(seq(nrow(evals)))  
ggplot(evals,aes(x=component.num,y=eigen.vals)) + geom\_point()



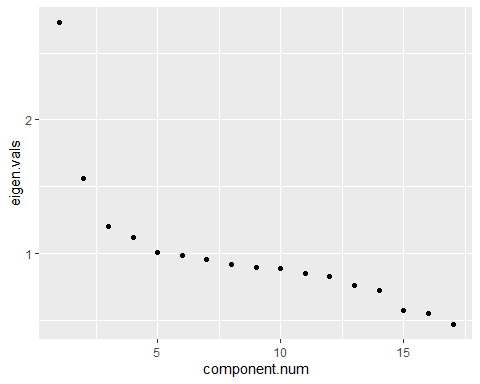
Gamma <- Heartpca$rot  
round(Gamma,3)

## PC1 PC2 PC3 PC4 PC5 PC6 PC7 PC8 PC9  
## BMI 0.019 0.007 0.026 -0.016 -0.007 0.009 -0.012 -0.006 -0.004  
## Smoking 0.761 -0.214 -0.591 0.069 -0.085 -0.012 -0.050 -0.069 -0.072  
## AlcoholDrinking 0.035 -0.024 -0.087 0.011 -0.089 0.044 0.312 0.913 0.156  
## Stroke 0.064 0.023 0.045 -0.049 0.081 0.009 -0.027 -0.015 0.101  
## PhysicalHealth 0.216 0.111 0.188 -0.235 0.101 0.263 0.196 -0.131 0.297  
## MentalHealth 0.109 0.095 0.035 -0.215 -0.169 0.339 0.578 -0.220 0.191  
## DiffWalking 0.333 0.192 0.313 -0.297 0.368 0.319 -0.276 0.261 -0.450  
## Sex 0.075 -0.904 0.404 -0.096 0.001 -0.002 0.011 0.006 0.035  
## AgeCategory 0.149 0.069 0.043 0.082 0.497 -0.299 -0.225 0.042 0.525  
## Race -0.010 0.010 0.037 -0.017 -0.150 0.126 -0.039 -0.065 -0.380  
## Diabetic 0.071 0.038 0.065 -0.039 0.078 0.000 -0.078 -0.057 0.133  
## PhysicalActivity -0.406 -0.230 -0.558 -0.515 0.378 0.234 -0.052 -0.008 0.020  
## GenHealth -0.188 -0.076 -0.134 0.137 -0.069 -0.106 -0.055 0.090 -0.184  
## SleepTime -0.005 0.001 -0.001 0.014 0.016 -0.017 -0.014 0.008 0.005  
## Asthma 0.083 0.110 0.055 -0.706 -0.402 -0.541 -0.135 0.046 0.032  
## KidneyDisease 0.054 0.026 0.051 -0.046 0.078 -0.003 -0.009 -0.038 0.110  
## SkinCancer 0.055 0.002 0.020 -0.020 0.461 -0.498 0.607 -0.084 -0.383  
## PC10 PC11 PC12 PC13 PC14 PC15 PC16 PC17  
## BMI -0.012 -0.014 -0.023 -0.037 -0.058 -0.037 0.976 0.198  
## Smoking 0.001 0.024 -0.010 -0.005 0.004 0.024 0.002 0.004  
## AlcoholDrinking -0.153 -0.042 -0.010 -0.016 -0.020 -0.001 0.007 0.000  
## Stroke -0.103 -0.168 -0.480 0.835 -0.047 0.062 0.013 -0.003  
## PhysicalHealth -0.350 0.605 0.164 0.065 0.070 0.320 0.013 0.008  
## MentalHealth 0.348 -0.474 0.115 -0.008 0.069 0.074 -0.002 0.020  
## DiffWalking 0.277 -0.057 0.007 -0.028 0.015 0.031 -0.021 -0.003  
## Sex 0.028 -0.039 0.025 -0.002 0.010 0.010 -0.006 0.001  
## AgeCategory -0.004 -0.335 0.389 0.035 0.179 0.051 0.025 -0.017  
## Race -0.700 -0.416 0.322 0.075 0.183 0.064 0.002 0.005  
## Diabetic -0.228 -0.269 -0.278 -0.348 -0.698 0.376 -0.056 -0.011  
## PhysicalActivity -0.057 -0.014 0.005 -0.012 -0.016 -0.033 0.012 0.005  
## GenHealth 0.256 0.003 0.068 0.080 0.217 0.856 0.063 -0.003  
## SleepTime 0.000 0.000 0.000 0.005 0.006 0.011 -0.198 0.980  
## Asthma 0.029 -0.012 0.039 0.007 0.023 0.026 -0.011 0.002  
## KidneyDisease -0.169 -0.109 -0.623 -0.401 0.617 0.044 0.003 -0.003  
## SkinCancer -0.073 0.056 -0.042 -0.006 -0.053 -0.002 0.008 -0.002

summary(Heartpca)

## Importance of components:  
## PC1 PC2 PC3 PC4 PC5 PC6 PC7  
## Standard deviation 0.5295 0.5138 0.4338 0.35317 0.34250 0.30722 0.26870  
## Proportion of Variance 0.1906 0.1795 0.1279 0.08481 0.07976 0.06418 0.04909  
## Cumulative Proportion 0.1906 0.3701 0.4980 0.58286 0.66263 0.72680 0.77590  
## PC8 PC9 PC10 PC11 PC12 PC13 PC14  
## Standard deviation 0.24731 0.22861 0.21326 0.20253 0.18699 0.18206 0.17426  
## Proportion of Variance 0.04159 0.03554 0.03093 0.02789 0.02377 0.02254 0.02065  
## Cumulative Proportion 0.81749 0.85302 0.88395 0.91184 0.93561 0.95815 0.97880  
## PC15 PC16 PC17  
## Standard deviation 0.15393 0.06352 0.05875  
## Proportion of Variance 0.01611 0.00274 0.00235  
## Cumulative Proportion 0.99491 0.99765 1.00000

##标准化数据  
Heartstdpca = prcomp(Heartstd[,2:18])  
stdevals <- data.frame(Heartstdpca$sdev^2)  
names(stdevals) <-'eigen.vals'  
stdevals$component.num <- as.integer(seq(nrow(stdevals)))  
library(ggplot2)  
ggplot(stdevals,aes(x=component.num,y=eigen.vals)) + geom\_point()



Gammastd <- Heartstdpca$rot  
round(Gammastd,3)

## PC1 PC2 PC3 PC4 PC5 PC6 PC7 PC8 PC9  
## BMI 0.213 -0.150 0.315 -0.318 0.207 -0.451 0.049 -0.065 -0.048  
## Smoking 0.168 0.062 -0.488 -0.374 -0.096 -0.081 0.061 -0.070 -0.094  
## AlcoholDrinking -0.033 -0.081 -0.538 -0.111 -0.282 -0.292 0.410 0.263 -0.141  
## Stroke 0.196 0.132 -0.026 0.033 -0.105 0.538 0.270 -0.371 -0.499  
## PhysicalHealth 0.416 -0.122 -0.135 0.116 -0.024 0.138 -0.180 -0.038 0.219  
## MentalHealth 0.213 -0.395 -0.290 0.248 0.062 0.072 -0.096 0.072 0.136  
## DiffWalking 0.426 0.037 0.016 0.052 -0.094 0.004 -0.126 -0.068 0.031  
## Sex -0.061 0.038 -0.094 -0.714 0.223 0.279 -0.021 -0.232 0.349  
## AgeCategory 0.215 0.565 -0.003 0.049 -0.017 -0.089 -0.077 0.034 -0.163  
## Race 0.009 -0.344 0.292 -0.147 -0.353 0.286 0.162 0.071 -0.043  
## Diabetic 0.262 0.093 0.330 -0.191 0.062 -0.175 0.235 0.232 -0.284  
## PhysicalActivity -0.298 0.032 -0.084 0.081 0.345 0.176 0.256 0.035 -0.047  
## GenHealth -0.460 0.073 0.016 0.054 0.001 -0.019 0.053 0.018 -0.070  
## SleepTime -0.048 0.298 0.134 0.093 -0.536 -0.151 0.281 -0.322 0.512  
## Asthma 0.138 -0.220 0.002 0.244 0.362 -0.197 0.536 -0.451 0.098  
## KidneyDisease 0.200 0.125 0.108 0.011 0.116 0.316 0.415 0.591 0.354  
## SkinCancer 0.083 0.413 -0.173 0.138 0.340 0.000 -0.013 0.010 0.131  
## PC10 PC11 PC12 PC13 PC14 PC15 PC16 PC17  
## BMI -0.173 -0.283 -0.145 0.369 0.415 -0.033 -0.053 0.182  
## Smoking 0.085 0.520 0.314 0.331 0.165 0.037 0.200 0.044  
## AlcoholDrinking 0.032 -0.366 -0.260 -0.250 -0.004 -0.024 -0.054 0.007  
## Stroke -0.245 -0.210 -0.152 0.222 -0.024 -0.037 0.055 0.031  
## PhysicalHealth -0.038 -0.102 0.066 -0.317 0.103 -0.290 0.303 0.613  
## MentalHealth -0.148 -0.228 0.334 0.346 -0.220 0.252 -0.438 0.048  
## DiffWalking 0.015 -0.020 -0.120 -0.281 0.272 0.726 0.150 -0.257  
## Sex -0.025 -0.159 -0.065 -0.206 -0.218 0.111 -0.198 0.024  
## AgeCategory 0.104 0.136 0.015 -0.149 0.110 -0.026 -0.674 0.267  
## Race 0.664 -0.039 0.107 0.039 0.197 0.023 -0.167 0.111  
## Diabetic -0.031 -0.092 0.487 -0.180 -0.477 0.098 0.185 0.047  
## PhysicalActivity -0.124 -0.157 0.510 -0.261 0.553 0.024 -0.022 -0.052  
## GenHealth -0.020 0.043 -0.096 0.120 -0.087 0.544 0.132 0.652  
## SleepTime -0.157 -0.148 0.258 0.114 0.016 0.000 0.022 -0.021  
## Asthma 0.195 0.306 -0.134 -0.148 -0.139 0.022 -0.095 0.051  
## KidneyDisease -0.181 0.219 -0.228 0.185 0.063 0.024 0.008 0.019  
## SkinCancer 0.559 -0.406 0.001 0.305 -0.063 0.009 0.249 -0.058

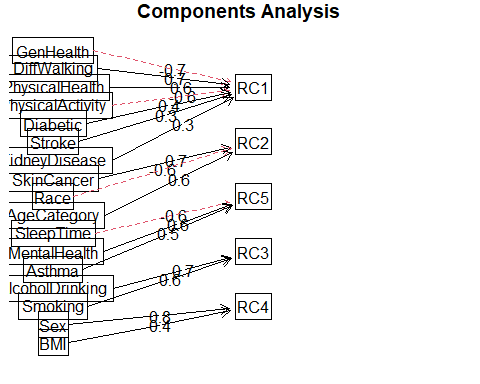
summary(Heartstdpca)

## Importance of components:  
## PC1 PC2 PC3 PC4 PC5 PC6 PC7  
## Standard deviation 1.6522 1.24987 1.09491 1.05732 1.00356 0.99219 0.97856  
## Proportion of Variance 0.1606 0.09189 0.07052 0.06576 0.05924 0.05791 0.05633  
## Cumulative Proportion 0.1606 0.25246 0.32298 0.38874 0.44798 0.50589 0.56222  
## PC8 PC9 PC10 PC11 PC12 PC13 PC14  
## Standard deviation 0.95849 0.94723 0.93995 0.92049 0.90839 0.87111 0.85171  
## Proportion of Variance 0.05404 0.05278 0.05197 0.04984 0.04854 0.04464 0.04267  
## Cumulative Proportion 0.61626 0.66904 0.72101 0.77085 0.81939 0.86403 0.90670  
## PC15 PC16 PC17  
## Standard deviation 0.75452 0.74019 0.68476  
## Proportion of Variance 0.03349 0.03223 0.02758  
## Cumulative Proportion 0.94019 0.97242 1.00000

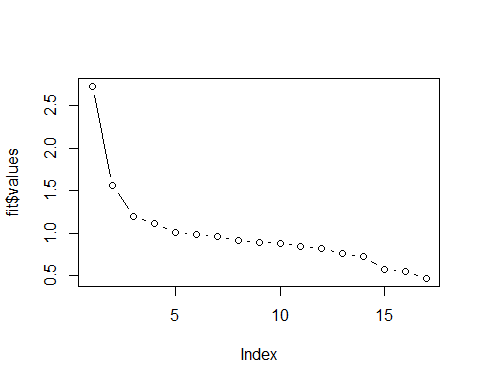
#FA验证  
library(psych)  
fit = principal(Heart[,2:18],nfactors = 5, rotate="varimax")  
#fit = principal(Heart[,2:18],nfactors = 5, rotate="varimax",n.obs = dim(Heart[,2:18]), scores = T, method = 'Bartlett')  
fit

## Principal Components Analysis  
## Call: principal(r = Heart[, 2:18], nfactors = 5, rotate = "varimax")  
## Standardized loadings (pattern matrix) based upon correlation matrix  
## RC1 RC2 RC5 RC3 RC4 h2 u2 com  
## BMI 0.30 -0.19 0.18 -0.34 0.39 0.43 0.57 3.9  
## Smoking 0.28 0.13 -0.01 0.55 0.36 0.53 0.47 2.4  
## AlcoholDrinking -0.02 -0.05 -0.02 0.67 0.02 0.45 0.55 1.0  
## Stroke 0.35 0.12 -0.08 0.03 -0.05 0.15 0.85 1.4  
## PhysicalHealth 0.64 0.01 0.32 0.12 -0.09 0.53 0.47 1.6  
## MentalHealth 0.25 -0.15 0.58 0.27 -0.21 0.54 0.46 2.3  
## DiffWalking 0.70 0.06 0.09 -0.02 -0.05 0.51 0.49 1.1  
## Sex -0.12 0.01 -0.07 0.11 0.78 0.64 0.36 1.1  
## AgeCategory 0.40 0.57 -0.36 -0.09 -0.02 0.63 0.37 2.6  
## Race 0.09 -0.65 -0.07 -0.09 0.01 0.44 0.56 1.1  
## Diabetic 0.44 -0.01 -0.06 -0.36 0.23 0.38 0.62 2.5  
## PhysicalActivity -0.56 0.21 0.12 -0.04 0.02 0.38 0.62 1.4  
## GenHealth -0.73 0.00 -0.24 0.00 -0.09 0.59 0.41 1.2  
## SleepTime 0.09 -0.02 -0.62 0.02 -0.27 0.47 0.53 1.4  
## Asthma 0.10 0.05 0.53 -0.15 -0.11 0.33 0.67 1.4  
## KidneyDisease 0.31 0.17 0.03 -0.18 0.05 0.16 0.84 2.4  
## SkinCancer 0.07 0.67 0.01 -0.04 0.00 0.46 0.54 1.0  
##   
## RC1 RC2 RC5 RC3 RC4  
## SS loadings 2.61 1.37 1.35 1.18 1.11  
## Proportion Var 0.15 0.08 0.08 0.07 0.07  
## Cumulative Var 0.15 0.23 0.31 0.38 0.45  
## Proportion Explained 0.34 0.18 0.18 0.15 0.15  
## Cumulative Proportion 0.34 0.52 0.70 0.85 1.00  
##   
## Mean item complexity = 1.8  
## Test of the hypothesis that 5 components are sufficient.  
##   
## The root mean square of the residuals (RMSR) is 0.09   
## with the empirical chi square 691867.9 with prob < 0   
##   
## Fit based upon off diagonal values = 0.49

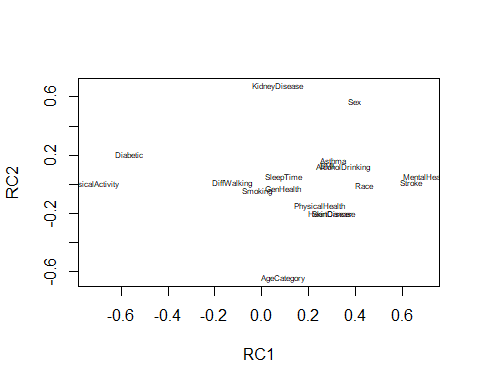
fa.diagram(fit)



plot(fit$values, type = "b")



plot(fit$loadings, type="n")  
text(fit$loadings, labels = names(Heart), cex=.5)



#fit$scores

#LDA  
#library(ICSNP)  
library(MASS)  
library(ggplot2)  
#HotellingsT2(Heart[Heart$HeartDisease=='0', c(2:18)], Heart[Heart$HeartDisease=='1', c(2:18)])  
subdata = Heart[,]  
#LDA  
L = lda(HeartDisease~.,subdata);L

## Call:  
## lda(HeartDisease ~ ., data = subdata)  
##   
## Prior probabilities of groups:  
## 0 1   
## 0.91440454 0.08559546   
##   
## Group means:  
## BMI Smoking AlcoholDrinking Stroke PhysicalHealth MentalHealth  
## 0 0.2975715 0.3962458 0.07056925 0.02626341 0.09854719 0.1276259  
## 1 0.3099799 0.5858693 0.04168341 0.16034048 0.26027472 0.1547255  
## DiffWalking Sex AgeCategory Race Diabetic PhysicalActivity  
## 0 0.1175767 0.4645717 0.5584571 0.2713516 0.06795487 0.7881350  
## 1 0.3663464 0.5895956 0.7872553 0.2469648 0.17476163 0.6389143  
## GenHealth SleepTime Asthma KidneyDisease SkinCancer  
## 0 0.7345248 0.2955590 0.1297406 0.02846571 0.08494231  
## 1 0.5532167 0.2973398 0.1802141 0.12621927 0.18193110  
##   
## Coefficients of linear discriminants:  
## LD1  
## BMI -0.527446780  
## Smoking 0.211696264  
## AlcoholDrinking -0.159238203  
## Stroke 1.893176212  
## PhysicalHealth 0.353823442  
## MentalHealth 0.005689991  
## DiffWalking 0.459650735  
## Sex 0.516513810  
## AgeCategory 1.565608158  
## Race -0.099724277  
## Diabetic 0.987102780  
## PhysicalActivity -0.021148933  
## GenHealth -1.673648829  
## SleepTime 0.149888648  
## Asthma 0.167453900  
## KidneyDisease 1.065390268  
## SkinCancer 0.267328464

yhat = predict(L, subdata)$class  
subdata$HeartDisease.pred = yhat  
tab = table(true=Heart$HeartDisease, pred = yhat);tab

## pred  
## true 0 1  
## 0 284645 7777  
## 1 21395 5978

aper = sum(tab[row(tab)!=col(tab)])/sum(tab);aper

## [1] 0.09122094

#use cross validation  
Lcv = lda(HeartDisease~., subdata, CV = TRUE)  
tabcv = table(pred=Lcv$class, true = Heart$HeartDisease);tabcv

## true  
## pred 0 1  
## 0 284645 21395  
## 1 7777 5978

apercv = sum(tabcv[row(tabcv)!=col(tabcv)])/sum(tabcv);apercv

## [1] 0.09122094

#logistic模型  
log.fit = glm(HeartDisease~., data=Heart, family = binomial(link=logit))  
summary(log.fit)

##   
## Call:  
## glm(formula = HeartDisease ~ ., family = binomial(link = logit),   
## data = Heart)  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -2.1854 -0.4089 -0.2462 -0.1364 3.6191   
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -4.23609 0.06850 -61.841 < 2e-16 \*\*\*  
## BMI 0.99824 0.10638 9.383 < 2e-16 \*\*\*  
## Smoking 0.36949 0.01430 25.844 < 2e-16 \*\*\*  
## AlcoholDrinking -0.23198 0.03352 -6.920 4.51e-12 \*\*\*  
## Stroke 1.04585 0.02264 46.200 < 2e-16 \*\*\*  
## PhysicalHealth 0.07529 0.02476 3.041 0.002360 \*\*   
## MentalHealth 0.13893 0.02648 5.247 1.55e-07 \*\*\*  
## DiffWalking 0.20969 0.01810 11.583 < 2e-16 \*\*\*  
## Sex 0.72669 0.01450 50.116 < 2e-16 \*\*\*  
## AgeCategory 3.57394 0.03937 90.785 < 2e-16 \*\*\*  
## Race -0.13738 0.03592 -3.824 0.000131 \*\*\*  
## Diabetic 0.85999 0.03141 27.380 < 2e-16 \*\*\*  
## PhysicalActivity 0.02357 0.01605 1.468 0.142003   
## GenHealth -2.49053 0.04229 -58.889 < 2e-16 \*\*\*  
## SleepTime -0.61374 0.10377 -5.914 3.33e-09 \*\*\*  
## Asthma 0.27893 0.01922 14.509 < 2e-16 \*\*\*  
## KidneyDisease 0.57336 0.02442 23.477 < 2e-16 \*\*\*  
## SkinCancer 0.12663 0.01941 6.525 6.80e-11 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for binomial family taken to be 1)  
##   
## Null deviance: 186906 on 319794 degrees of freedom  
## Residual deviance: 145470 on 319777 degrees of freedom  
## AIC: 145506  
##   
## Number of Fisher Scoring iterations: 6

anova(log.fit)

## Analysis of Deviance Table  
##   
## Model: binomial, link: logit  
##   
## Response: HeartDisease  
##   
## Terms added sequentially (first to last)  
##   
##   
## Df Deviance Resid. Df Resid. Dev  
## NULL 319794 186906  
## BMI 1 805.0 319793 186101  
## Smoking 1 3579.9 319792 182521  
## AlcoholDrinking 1 607.9 319791 181913  
## Stroke 1 6800.1 319790 175113  
## PhysicalHealth 1 3953.3 319789 171159  
## MentalHealth 1 197.9 319788 170962  
## DiffWalking 1 2883.4 319787 168078  
## Sex 1 1903.5 319786 166175  
## AgeCategory 1 14296.0 319785 151879  
## Race 1 3.6 319784 151875  
## Diabetic 1 1515.3 319783 150360  
## PhysicalActivity 1 35.3 319782 150324  
## GenHealth 1 4018.8 319781 146306  
## SleepTime 1 36.4 319780 146269  
## Asthma 1 221.3 319779 146048  
## KidneyDisease 1 535.7 319778 145512  
## SkinCancer 1 42.0 319777 145470

Note that the echo = FALSE parameter was added to the code chunk to prevent printing of the R code that generated the plot.