Winter 2023
SurvMeth 686
Project #1

Cheng, Chia Wen

Abstract

This paper analyzes a provided dataset regarding personal health conditions and habits. We are interested in estimating the probability of having coronary heart disease, or CHD, with a concise model incorporating the most impactful factors. The stepwise function and literature review are applied to the process of identifying predictors. Six independent variables compose of the best model for prediction: $tobacco$, ldl, $famhist$, $typea$, $obesity$, and age. Likelihood ratio test and the manually computed R-squared specify that our model is not a good fit to our data. Nevertheless, other possible models perform similarly poorly in goodness of fit as ours, and our model is not statistically significantly worse than the model with the lowest AIC score. Prediction of this model performs fairly well with a high overall accuracy of nearly 0.75 computed by confusionMatrix, and a large area under the curve of the ROC of 0.79. Family history is the most critical predictor as it owns the greatest impact on our outcome variable, probability of having CHD, with the greatest magnitude for every one-unit change.

Preliminary Analysis

The dependent variable we are interested in is a dichotomous variable composed of 0 and 1, where 0 indicates an absent of coronary heart disease (hereinafter called 'CHD' or 'the disease'), and 1 signals a present of the disease. According to the nature of CHD as the dependent variable, a logistic model with $lambda = 0$ used to predict the probability of having CHD may be appropriate, as logistic regression is a specific type of generalized linear model (GLM) designed to model data that has a binomial distribution.[i]

There are 420 observations in the provided dataset, which is considered sufficient in sample size in terms of representativeness of population. All observations are complete without missing values. The only non-numeric variable we have so far is the "famhist," marking family history of CHD. However, for the purpose of predicting, I re-encode "Present" as 1 and "Absent" as 0. "Age" is left numeric because continuous variables usually provide more valuable information than categorical variables do. Grouping ages may result in loss of some insightful information. Variables "adiposity" and "typea" are left skewed; "age" is distinct; and all the other predictors except for "famhist" are right skewed with more observations on the left side. The dummy "famhist" has a mean of 0.4214, revealing slightly more observations in this dataset have no family history of CHD.

Predictor Selection

To decide what predictors to be involved, I first regress two GLMs, one with all variables as independent variables, and the other with intercept only. The model with intercept only has a greater AIC than the one with every variable. In addition, the residual deviance of 428.74 is smaller than the null deviance of 540.05. This further proves the preference of including predictors to the model in this case over including nothing.

This paper utilizes two methods in selecting the predictors: looking for a best model statistically, and reviewing literature as an auxiliary. I use the stepwise function to select the best GLM. The procedure adds or removes independent variables one at a time considering the variable's statistical significance—it either adds the most significant variable or removes the least significant variable.[ii] The model with "tobacco," "ldl," "famhist," 'typea," and "age" as predictors is the best one with the lowest AIC of 444.37, in light of R programming. All of the five regressors as well as the intercept are statistically significant at 95% confidence level, implying a concise model without redundant variables. The residual deviance is 432.37 on 414 degrees of freedom.

Since the stepwise regression does not consider all possible models, and the fact that its return should be treated as a "suggestion" instead of a valid conclusion on which independent variables should be selected, and which can potentially be removed, conducting literature review helps select predictors in this analysis to ideally strike the limitation of the stepwise regression. According to the Centers for Disease Control and Prevention (CDC), "overweight, physical inactivity, unhealthy eating, and smoking tobacco are risk factors for coronary artery disease, or CAD." "A family history of heart disease also increases your risk for CAD, especially a family history of having heart disease at an early age (50 or younger)."[iii] I add "obesity" to the model to better reflect the risk of "overweight."

Model Goodness of Fit

With the six predictors, five selected by the stepwise function and one by myself with assistance of prior studies, several other possible models are tried. I tried adding in interaction terms, such as $age * lbl$, $age * obesity$, and $ldl * obesity$ to see if these improve our model. The results turn out that none of these models with interaction terms is able to achieve an even smaller AIC score comparing to the plain model. I also tried transforming some independent variables by squaring age, $famhist$, and $obesity$, respectively. Again, none of them is able to reach an AIC score lower than the plain model. Thus, I move forward with the model composed of $tobacco$, ldl, $famhist$, $typea$, $obesity$, and age.

Now that a model containing $tobacco$, ldl, $famhist$, $typea$, $obesity$, and age is decided, model goodness of fit is to be examined. Due to the limitation of our model prediction method, GLM, normality of residual distribution is not a good way of examination; instead, I implement both likelihood ratio test and manually computed R-squared for GLM, and compare AIC scores.

1. The likelihood ratio test implies that the model without “obesity” fits better with data we have than the one with “obesity,” where a smaller maximum log-likelihood is found in the former model. This comes align with our assumption because the programming selects the former model as the best fitted one. However, a p-value of 0.2644 shows that the latter model is not statistically significant worse than the former one.
2. An appropriate calculation for GLM to obtain an R-squared is by using the formula:

$$1-\frac{residual\ deviance}{null\ deviance}$$

   The ratio of residual deviance over null deviance is 0.7983038, expressing that the residual deviance is not explained in the null.[iv] Our R-squared is then 0.2016962, which is a low satisfaction of the match between our data and our model.
3. The AIC score of 445.12 is slightly higher than the score of 444.37, pointing out that our model with the 6 predictors performs worse than the machine-selected best-fitted model composed of the 5 predictors. Our model also shares a higher AIC score than the one containing “sbp,” “tobacco,” “ldl,” “famhist,” “typea,” and “age,” which is 444.83. Nevertheless, it is lower than the AIC scores of the rest possible models stepwise function exercised.

Therefore, I conclude that our model is not a good fit to our data. Nonetheless, the machine-selected model turns out with similar low goodness of fit as our model of selection. The fact that our model of selection is not statistically significantly worse than the one our machine selects is sufficient for us to move forward with it.

<u>Prediction Accuracy Assessment</u>

Prediction accuracy is assessed in two ways:

1. With the function confusionMatrix in R programming, we get the following results (the threshold of 0.5 is used as I would like to have significantly more false positives and substantially less false negatives for prevention and data tracking[v]):

$$Sensitivity\ (true\ positive\ rate): 0.5347$$

$$Specificity\ (true\ negative\ rate): 0.8514$$

$$Positively\ Predicted\ Value: 0.6525$$

$$Negatively\ Predicted\ Value: 0.7781$$

   The overall accuracy is 0.7429, indicating a fair to good performance of model prediction, with 0.8 taken as the threshold for good prediction performance.
2. The area under curve (AUC) of the ROC curve is 0.7904, indicating well predictions of the logistic model.

To conclude, our model performed well in prediction accuracy.

Coefficient Interpretation

From the output of our logistic regression model, we can interpret the coefficients as the following:

- in the case of our current data, the mean probability of having CHD is 0.00493, holding all other regressors constant; it is statistically significant at 95% confidence level; and
- on average, holding all other regressors constant, every one-unit increase in tobacco per day is estimated to be associated with an increase in the odds-ratio of having CHD by 1.09; or, the log odds of having CHD increases by 0.521 for every one-unit increase in tobacco per day; it is statistically significant at 95% confidence level; and
- on average, holding all other regressors constant, every one-unit increase in ldl is estimated to be associated with an increase in the odds-ratio by 1.2; or, the log odds of having CHD increases by 0.546 for every one-unit increase in ldl; it is statistically significant at 95% confidence level; and
- on average, holding all other regressors constant, a person with family history of CHD is estimated to be associated with a higher odds-ratio by 2.51 comparing to one without family history of the disease; or, the log odds of having CHD is higher by 0.715 for every person with family history of CHD comparing to those without; it is statistically significant at 95% confidence level; and
- on average, holding all other regressors constant, every one-score increase in Type A personality test is estimated to be associated with an increase in the odds-ratio by 1.03; or, the log odds of having CHD increases by 0.507 for every one-unit increase in Type A personality test score; it is statistically significant at 95% confidence level; and
- on average, holding all other regressors constant, every one-unit increase in a person's BMI is estimated to be associated with an increase in the odds-ratio by 0.967; or, the log odds of having CHD increases by 0.492 for every one-unit increase in BMI; it is not statistically significant at 90% confidence level; and
- on average, holding all other regressors constant, every one-year increase in a person's age is estimated to be associated with an increase in the odds-ratio by 1.05; or, the log odds of having CHD increases by 0.513 for every one-year increase in age; it is statistically significant at 95% confidence level.

Critical Predictor

Family history is a critical predictor among all variables because of the greatest magnitude of its change in odds-ratio for every one-unit change in it, although being a dichotomous variable.

[i] Stephen Roecker and Tom D'Avello. (March 2020.) Logistic Regression. Retrieved on February 15, 2023, from http://ncss-tech.github.io/stats_for_soil_survey/chapters/7_generalized_linear_models/7_generalized_linear_models.html#:~:text=Logistic%20regression%20is%20a%20specific,link%20transform%20is%20generally%20used.

[ii] Jim Frost. Guide to Stepwise Regression and Best Subsets Regression. Retrieved on February 15, 2023, from https://statisticsbyjim.com/regression/guide-stepwise-best-subsets-regression/.

[iii] Centers for Disease Control and Prevention. (July 2021.) Coronary Artery Disease (CAD). Retrieved on February 13, 2023, from https://www.cdc.gov/heartdisease/coronary_ad.htm#:~:text=Overweight%2C%20physical%20inactivity%2C%20unhealthy%20eating,age%20(50%20or%20younger).

[iv] How to calculate goodness of fit in glm. Cross Validated. Retrieved on February 16, 2023, from https://stats.stackexchange.com/questions/46345/how-to-calculate-goodness-of-fit-in-glm-r.

[v] Shad Griffin. (July 2020.) Determining a Cut-Off or Threshold When Working With a Binary Dependent (Target) Variable. Meium. Retrieved on February 16, 2023, from https://medium.com/swlh/determining-a-cut-off-or-threshold-when-working-with-a-binary-dependent-target-variable-7c2342cf2a7c.

SurvMeth 686 Project 1-Appendix

Cheng, Chia Wen

2023-02-17

```
setwd("C:/Users/Angela/    (cwcheng@umich.edu)/0. study abroad/academic/9. 2023 Winter/8. SurvMeth 686
chd <- read.csv('chd.csv')

## packages
library(ggplot2)
library(MASS)
library(maxLik)
library(lmtest)
library(pROC)
library(caret)


## examine data
str(chd)
```

```
## 'data.frame':    420 obs. of  10 variables:
##  $ sbp      : int  128 160 162 136 170 130 150 122 132 132 ...
##  $ tobacco  : num  0 3 7 5.8 0.4 4 0.18 4.18 2.8 7.2 ...
##  $ ldl      : num  2.63 9.19 7.67 5.9 4.11 2.4 4.14 9.05 4.79 3.65 ...
##  $ adiposity: num  23.9 26.5 34.3 27.6 42.1 ...
##  $ famhist  : chr  "Absent" "Present" "Present" "Absent" ...
##  $ typea    : int  45 39 33 65 56 60 53 44 50 56 ...
##  $ obesity  : num  21.6 28.2 30.8 25.7 33.1 ...
##  $ alcohol  : num  6.54 14.4 0 14.4 2.06 ...
##  $ age      : int  57 54 62 59 57 40 44 52 48 34 ...
##  $ chd      : int  0 1 0 0 0 0 0 1 0 0 ...
```

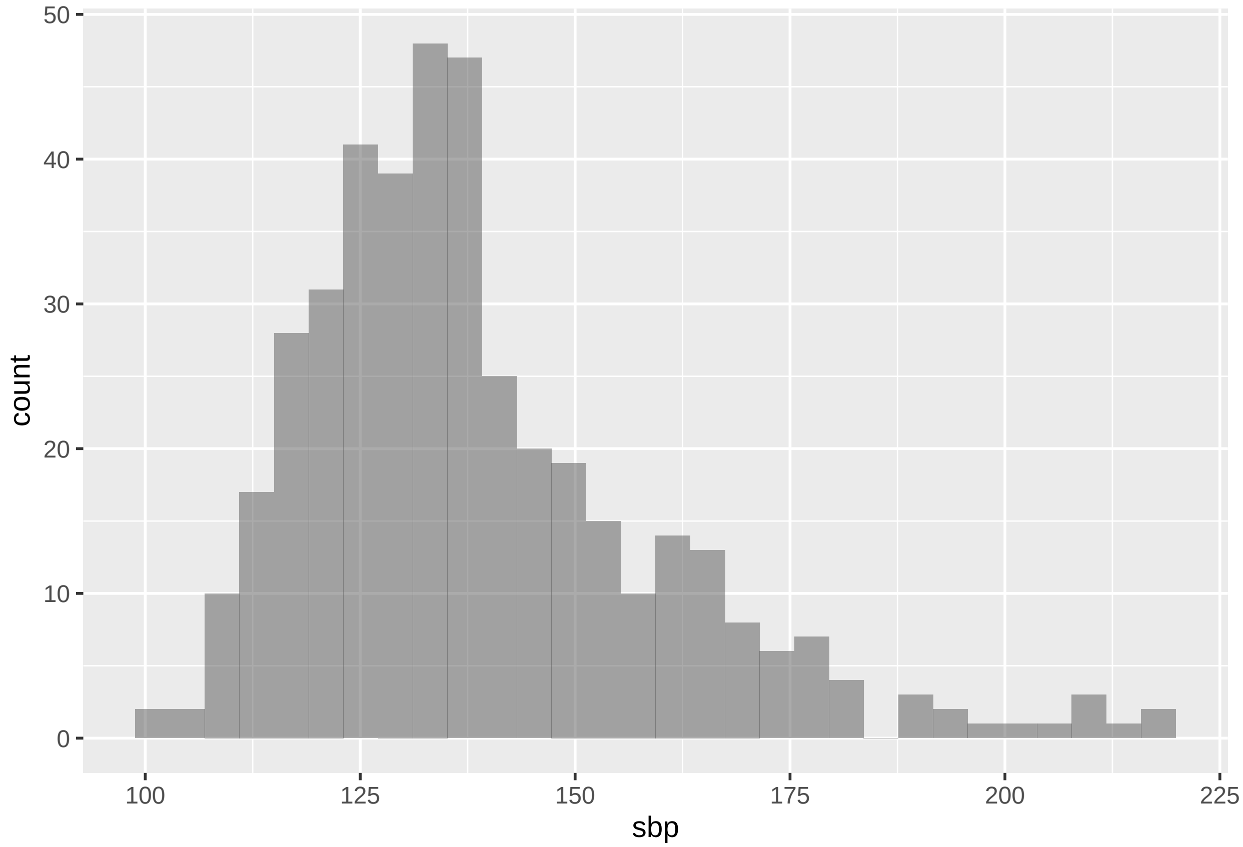
```
## sbp: int; tobbacco: num; ldl: num; adiposity: num; famhist: chr; typea: int; obesity: num;
## alcohol: num; age: int; chd: int

## preliminary analysis
is.na(chd) ## no missing value
```

```
##          sbp tobacco   ldl adiposity famhist typea obesity alcohol   age   chd
##   [1,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##   [2,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##   [3,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##   [4,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##   [5,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##   [6,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##   [7,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##   [8,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##   [9,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [10,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [11,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [12,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [13,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [14,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [15,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [16,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [17,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [18,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [19,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [20,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [21,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [22,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [23,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [24,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [25,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [26,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [27,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [28,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [29,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [30,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [31,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [32,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [33,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [34,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [35,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [36,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [37,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [38,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [39,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [40,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [41,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [42,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [43,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [44,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [45,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [46,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [47,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [48,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [49,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [50,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [51,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [52,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [53,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [54,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [55,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [56,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [57,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [58,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [59,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [60,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [61,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [62,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [63,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [64,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [65,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [66,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [67,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [68,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [69,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [70,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [71,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [72,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [73,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [74,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [75,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [76,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [77,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [78,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [79,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [80,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [81,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [82,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [83,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [84,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [85,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [86,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [87,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [88,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [89,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [90,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [91,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [92,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [93,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [94,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [95,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [96,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [97,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [98,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
##  [99,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [100,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [101,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [102,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [103,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [104,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [105,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [106,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [107,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [108,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [109,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [110,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [111,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [112,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [113,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [114,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [115,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [116,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [117,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [118,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [119,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [120,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [121,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [122,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [123,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [124,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [125,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [126,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [127,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [128,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [129,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [130,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [131,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [132,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [133,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [134,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [135,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [136,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [137,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [138,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [139,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [140,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [141,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [142,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [143,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [144,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [145,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [146,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [147,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [148,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [149,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [150,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [151,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [152,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [153,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [154,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [155,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [156,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [157,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [158,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [159,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [160,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [161,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [162,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [163,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [164,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [165,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [166,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [167,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [168,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [169,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [170,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [171,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [172,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [173,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [174,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [175,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [176,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [177,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [178,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [179,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [180,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [181,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [182,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [183,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [184,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [185,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [186,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [187,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [188,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [189,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [190,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [191,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [192,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [193,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [194,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [195,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [196,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [197,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [198,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [199,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [200,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [201,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [202,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [203,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [204,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [205,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [206,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [207,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [208,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [209,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [210,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [211,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [212,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [213,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [214,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [215,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [216,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [217,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [218,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [219,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [220,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [221,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [222,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [223,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [224,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [225,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [226,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [227,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [228,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [229,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [230,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [231,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [232,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [233,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [234,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [235,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [236,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [237,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [238,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [239,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [240,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [241,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [242,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [243,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [244,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [245,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [246,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [247,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [248,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [249,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [250,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [251,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [252,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [253,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [254,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [255,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [256,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [257,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [258,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [259,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [260,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [261,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [262,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [263,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [264,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [265,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [266,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [267,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [268,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [269,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [270,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [271,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [272,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [273,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [274,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [275,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [276,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [277,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [278,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [279,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [280,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [281,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [282,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [283,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [284,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [285,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [286,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [287,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [288,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [289,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [290,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [291,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [292,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [293,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [294,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [295,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [296,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [297,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [298,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [299,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [300,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [301,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [302,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [303,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [304,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [305,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [306,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [307,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [308,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [309,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [310,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [311,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [312,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [313,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [314,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [315,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [316,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [317,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [318,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [319,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [320,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [321,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [322,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [323,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [324,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [325,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [326,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [327,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [328,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [329,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [330,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [331,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [332,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [333,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [334,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [335,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [336,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [337,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [338,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [339,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [340,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [341,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [342,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [343,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [344,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [345,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [346,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [347,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [348,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [349,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [350,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [351,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [352,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [353,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [354,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [355,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [356,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [357,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [358,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [359,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [360,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [361,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [362,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [363,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [364,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [365,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [366,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [367,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [368,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [369,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [370,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [371,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [372,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [373,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [374,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [375,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [376,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [377,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [378,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [379,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [380,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [381,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [382,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [383,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [384,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [385,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [386,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [387,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [388,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [389,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [390,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [391,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [392,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [393,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [394,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [395,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [396,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [397,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [398,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [399,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [400,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [401,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [402,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [403,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [404,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [405,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [406,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [407,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [408,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [409,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [410,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [411,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [412,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [413,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [414,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [415,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [416,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [417,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [418,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [419,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
## [420,] FALSE   FALSE FALSE     FALSE   FALSE FALSE   FALSE   FALSE FALSE FALSE
```

```
summary(chd$sbp)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max. 
##   101.0   124.0   134.0   138.5   148.0   218.0
```

```
ggplot(chd, aes(x=sbp, color=sbp, fill=sbp)) +
  geom_histogram(position="identity", alpha=0.5) ## sbp is right skewed
```



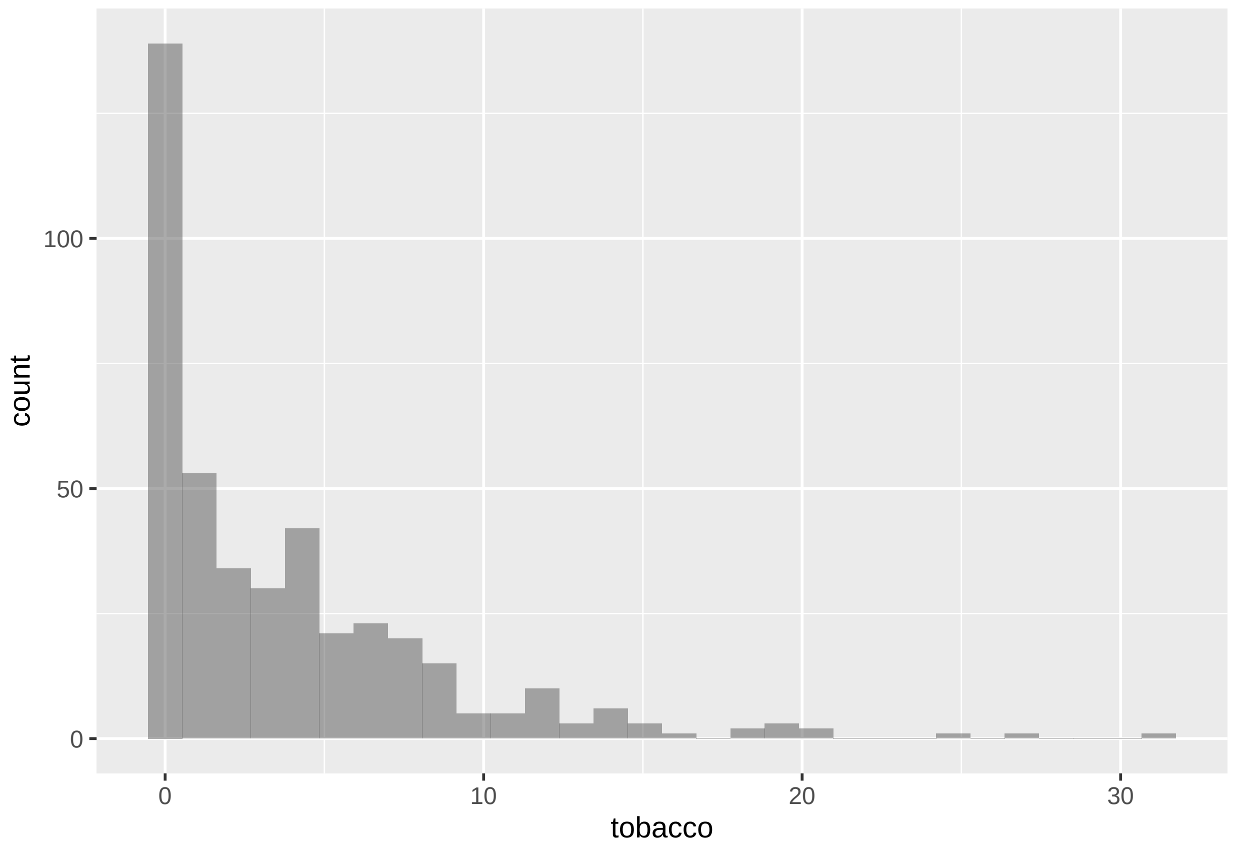
```
summary(chd$chd) ## the mean is 0.3429, , indicating a greater number in chd=0
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0000  0.0000  0.0000  0.3429  1.0000  1.0000
```

```
summary(chd$tobacco)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0000  0.0675  2.0850  3.7348  5.6000 31.2000
```

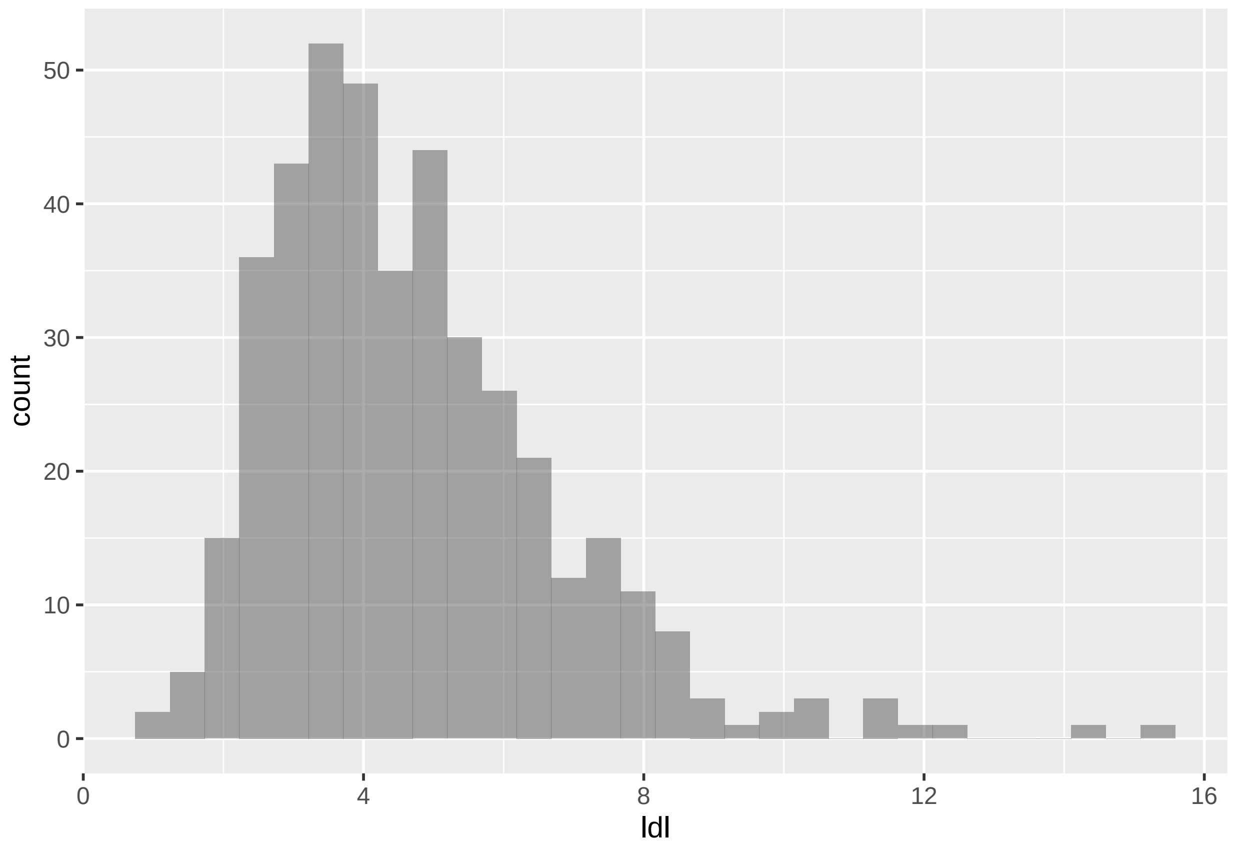
```
ggplot(chd, aes(x=tobacco, color=tobacco, fill=tobacco)) +
  geom_histogram(position="identity", alpha=0.5) ## highly right skewed with a maximum at 31.2 and mean
```



```
summary(chd$ldl)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max. 
##   0.980   3.255   4.325   4.726   5.803  15.330
```

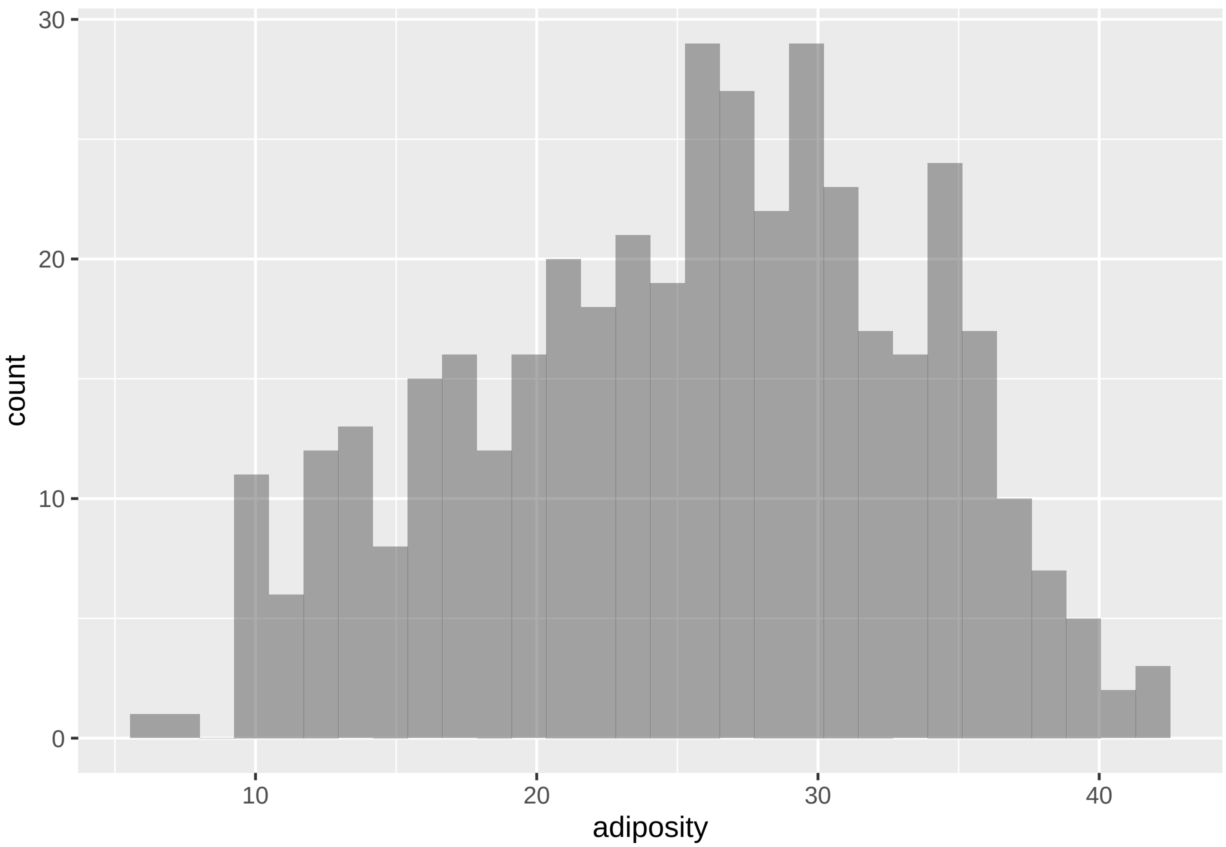
```
ggplot(chd, aes(x=ldl, color=ldl, fill=ldl)) +
  geom_histogram(position="identity", alpha=0.5) ## right skewed with a mean at 4.726 and a maximum at 
```



```
summary(chd$adiposity)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    6.74   19.94   26.25   25.46   31.30   42.49
```

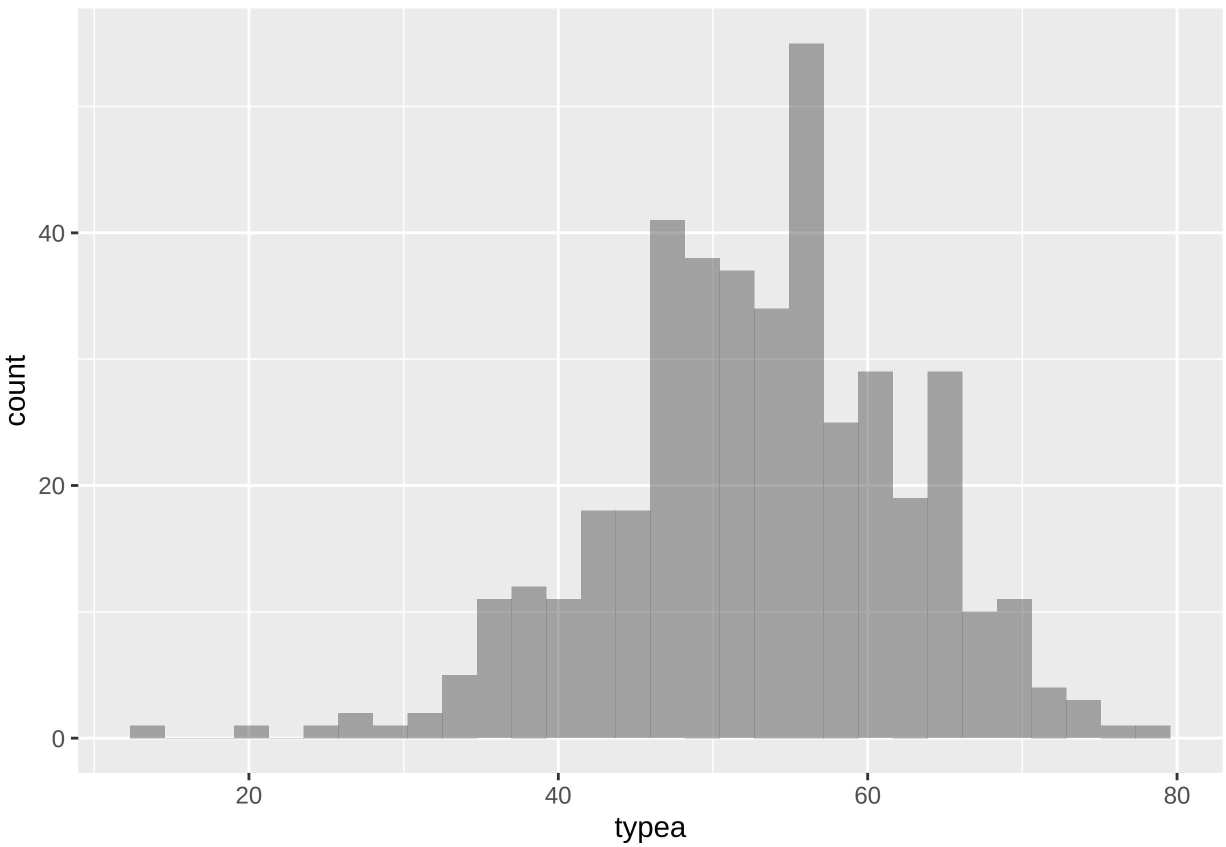
```
ggplot(chd, aes(x=adiposity, color=adiposity, fill=adiposity)) +
  geom_histogram(position="identity", alpha=0.5) ## slightly left skewed with a min of 6.74, a max of 4
```



```
summary(chd$typea)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max. 
##   13.00   47.00   53.00   52.97   60.00   78.00
```

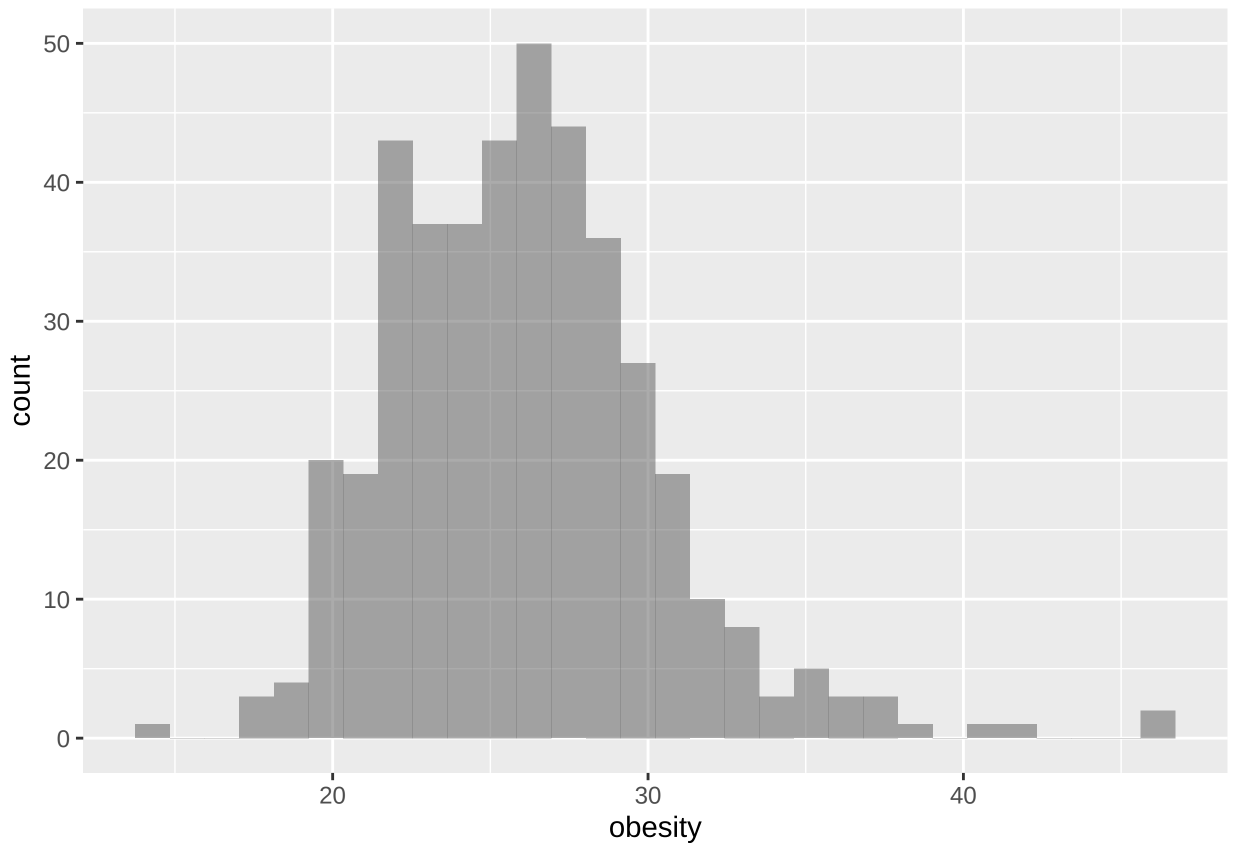
```
ggplot(chd, aes(x=typea, color=typea, fill=typea)) +
  geom_histogram(position="identity", alpha=0.5) ## slightly left skewed with a min of 13, a max of 78,
```



```
summary(chd$obesity)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max. 
##   14.70   22.95   25.89   26.07   28.49   46.58 
```

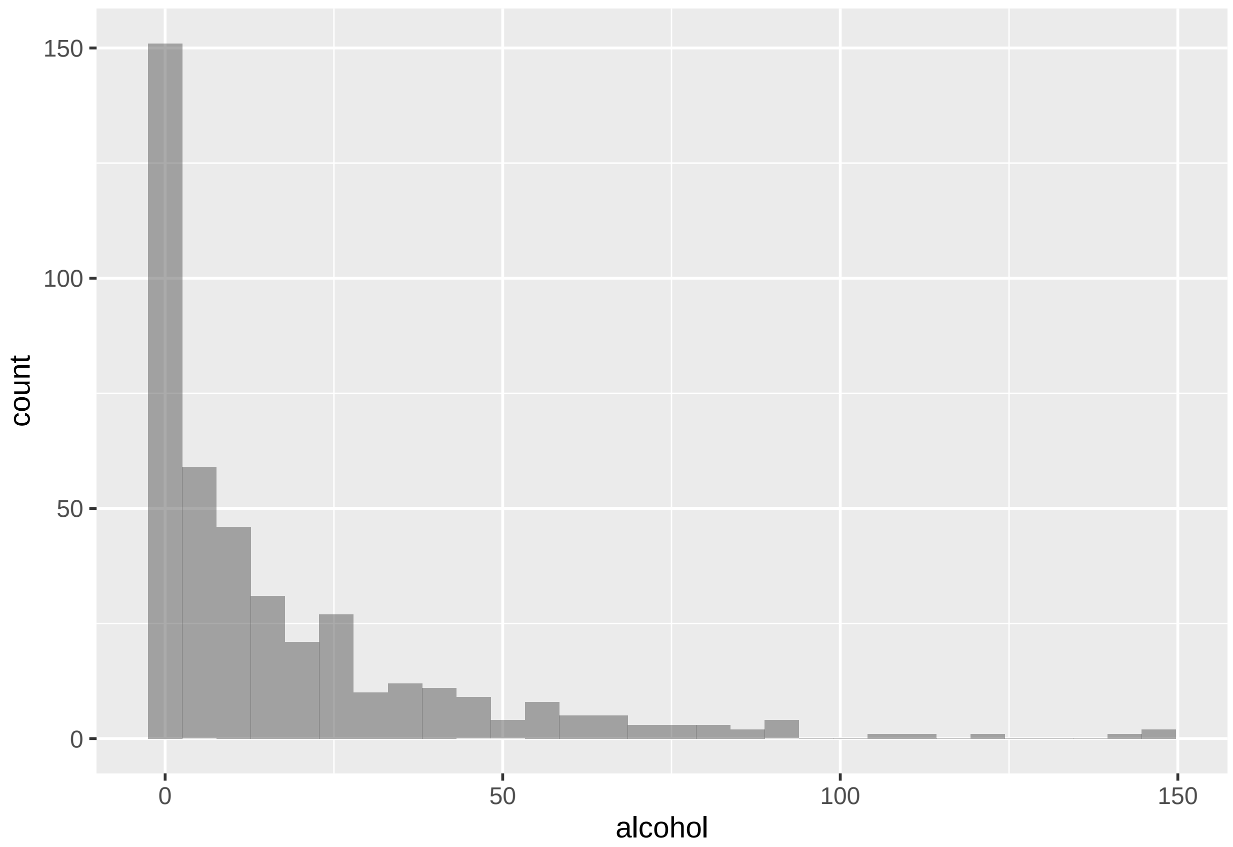
```
ggplot(chd, aes(x=obesity, color=obesity, fill=obesity)) +
  geom_histogram(position="identity", alpha=0.5) ## slightly right skewed with a min of 14.7, a max of
```



```
summary(chd$alcohol)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    0.00    0.51    7.66   17.17   24.04  147.19
```

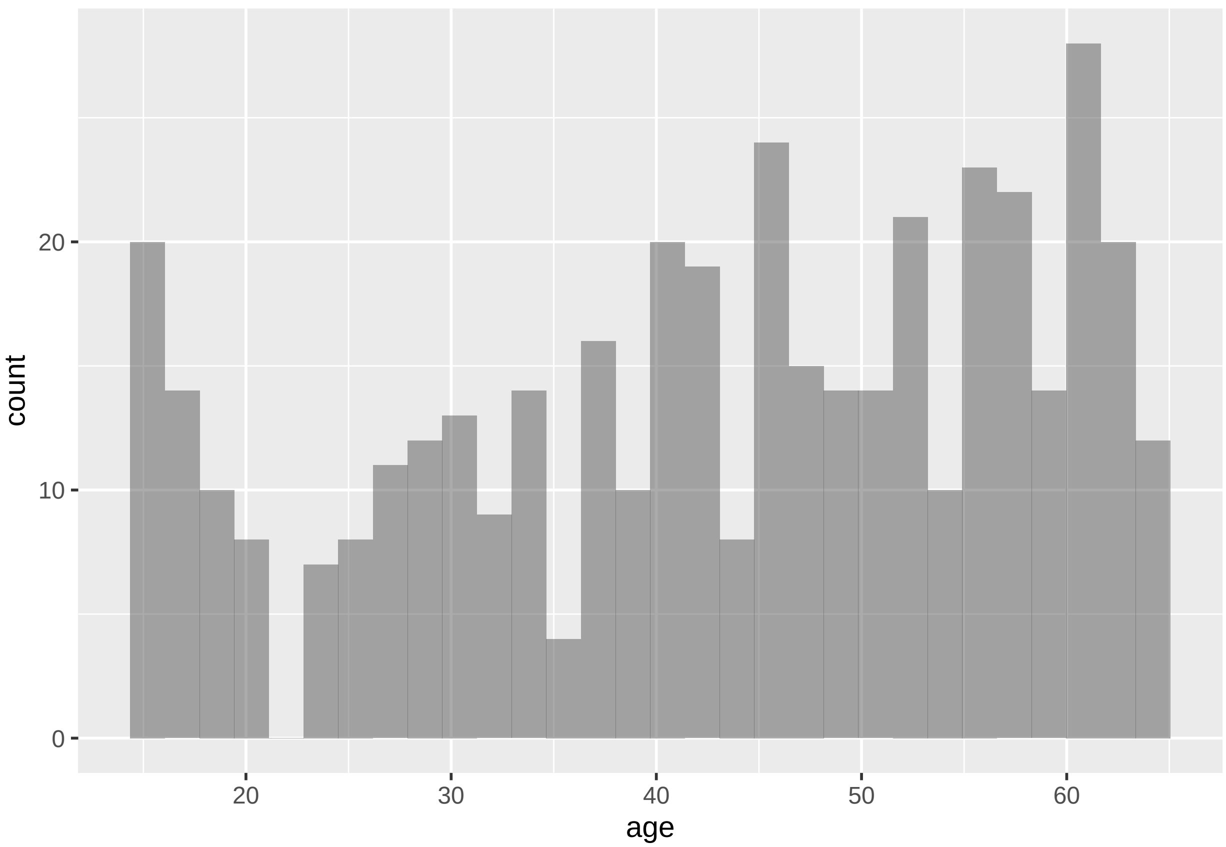
```
ggplot(chd, aes(x=alcohol, color=alcohol, fill=alcohol)) +
  geom_histogram(position="identity", alpha=0.5) ## highly right skewed with a min of 0, a max of 147.1
```



```
summary(chd$age)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   15.00   32.00   45.00   43.07   55.00   64.00
```

```
ggplot(chd, aes(x=age, color=age, fill=age)) +
  geom_histogram(position="identity", alpha=0.5) ## very distinct
```



```
## create a new data frame to encode 'famhist' to dummy
chd_1 <- chd
chd_1$famhist[chd_1$famhist == "Present"] <- 1
chd_1$famhist[chd_1$famhist == "Absent"] <- 0
chd_1$famhist <- as.numeric(chd_1$famhist)
summary(chd_1$famhist) ## mean at 0.4214, indicating more observations in famhist=0 (absent)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0000  0.0000  0.0000  0.4214  1.0000  1.0000
```

```
## predictor selection
fullmod <- glm(chd ~ ., data = chd_1, family = binomial)
summary(fullmod)
```

```
##
## Call:
## glm(formula = chd ~ ., family = binomial, data = chd_1)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -1.7495  -0.8349  -0.4333   0.8938   2.5643
##
## Coefficients:
##                Estimate Std. Error z value Pr(>|z|)
## (Intercept) -5.979369   1.369834  -4.365 1.27e-05 ***
## sbp          0.007746   0.005958   1.300 0.193526
## tobacco      0.081893   0.027356   2.994 0.002758 **
## ldl          0.178967   0.062625   2.858 0.004267 **
## adiposity    0.014893   0.030536   0.488 0.625744
## famhist      0.920423   0.239991   3.835 0.000125 ***
## typea        0.030187   0.012873   2.345 0.019026 *
## obesity     -0.055503   0.045343  -1.224 0.220925
## alcohol      0.001774   0.004672   0.380 0.704129
## age          0.044239   0.012707   3.482 0.000498 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 540.05  on 419  degrees of freedom
## Residual deviance: 428.74  on 410  degrees of freedom
## AIC: 448.74
##
## Number of Fisher Scoring iterations: 5
```

```
reducedmod <- glm(chd ~ 1, data = chd_1, family = binomial)
summary(reducedmod)
```

```
##
## Call:
## glm(formula = chd ~ 1, family = binomial, data = chd_1)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -0.9164  -0.9164  -0.9164   1.4632   1.4632
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -0.6506     0.1028  -6.329 2.47e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 540.05  on 419  degrees of freedom
## Residual deviance: 540.05  on 419  degrees of freedom
## AIC: 542.05
##
## Number of Fisher Scoring iterations: 4
```

```
mod <- glm(chd ~ ., data = chd_1, family = binomial)
stepAIC(mod, trace = 1, direction = "both") ## model chosen was with "tobacco + ldl + famhist + typea +
```

```
## Start:  AIC=448.74
## chd ~ sbp + tobacco + ldl + adiposity + famhist + typea + obesity +
##     alcohol + age
##
##             Df Deviance    AIC
## - alcohol    1   428.88 446.88
## - adiposity  1   428.97 446.97
## - obesity    1   430.28 448.28
## - sbp        1   430.44 448.44
## <none>           428.74 448.74
## - typea      1   434.42 452.42
## - ldl        1   437.34 455.34
## - tobacco    1   438.37 456.37
## - age        1   441.28 459.28
## - famhist    1   443.70 461.70
##
## Step:  AIC=446.88
## chd ~ sbp + tobacco + ldl + adiposity + famhist + typea + obesity +
##     age
##
##             Df Deviance    AIC
## - adiposity  1   429.14 445.14
## - obesity    1   430.46 446.46
## - sbp        1   430.73 446.73
## <none>           428.88 446.88
## + alcohol    1   428.74 448.74
## - typea      1   434.63 450.63
## - ldl        1   437.34 453.34
## - tobacco    1   439.26 455.26
## - age        1   441.29 457.29
## - famhist    1   444.17 460.17
##
## Step:  AIC=445.14
## chd ~ sbp + tobacco + ldl + famhist + typea + obesity + age
##
##             Df Deviance    AIC
## - obesity    1   430.83 444.83
## - sbp        1   431.12 445.12
## <none>           429.14 445.14
## + adiposity  1   428.88 446.88
## + alcohol    1   428.97 446.97
## - typea      1   434.73 448.73
## - ldl        1   438.75 452.75
## - tobacco    1   439.65 453.65
## - famhist    1   444.58 458.58
## - age        1   448.09 462.09
##
## Step:  AIC=444.83
## chd ~ sbp + tobacco + ldl + famhist + typea + age
##
##             Df Deviance    AIC
## - sbp        1   432.37 444.37
## <none>           430.83 444.83
## + obesity    1   429.14 445.14
## + adiposity  1   430.46 446.46
## + alcohol    1   430.67 446.67
## - typea      1   436.03 448.03
## - ldl        1   439.01 451.01
## - tobacco    1   441.47 453.47
## - famhist    1   445.96 457.96
## - age        1   448.99 460.99
##
## Step:  AIC=444.37
## chd ~ tobacco + ldl + famhist + typea + age
##
##             Df Deviance    AIC
## <none>           432.37 444.37
## + sbp        1   430.83 444.83
## + obesity    1   431.12 445.12
## + alcohol    1   432.06 446.06
## + adiposity  1   432.20 446.20
## - typea      1   437.33 447.33
## - ldl        1   440.96 450.96
## - tobacco    1   443.21 453.21
## - famhist    1   447.32 457.32
## - age        1   455.15 465.15
```

```
##
## Call:  glm(formula = chd ~ tobacco + ldl + famhist + typea + age, family = binomial,
##     data = chd_1)
##
## Coefficients:
## (Intercept)      tobacco          ldl      famhist        typea          age
##    -5.98647      0.08444      0.16578      0.91050      0.02794      0.04939
##
## Degrees of Freedom: 419 Total (i.e. Null);  414 Residual
## Null Deviance:       540
## Residual Deviance: 432.4     AIC: 444.4
```

```
summary(stepAIC(mod, trace = 1, direction = "both"))
```

```
## Start:  AIC=448.74
## chd ~ sbp + tobacco + ldl + adiposity + famhist + typea + obesity +
##     alcohol + age
##
##             Df Deviance    AIC
## - alcohol    1   428.88 446.88
## - adiposity  1   428.97 446.97
## - obesity    1   430.28 448.28
## - sbp        1   430.44 448.44
## <none>           428.74 448.74
## - typea      1   434.42 452.42
## - ldl        1   437.34 455.34
## - tobacco    1   438.37 456.37
## - age        1   441.28 459.28
## - famhist    1   443.70 461.70
##
## Step:  AIC=446.88
## chd ~ sbp + tobacco + ldl + adiposity + famhist + typea + obesity +
##     age
##
##             Df Deviance    AIC
## - adiposity  1   429.14 445.14
## - obesity    1   430.46 446.46
## - sbp        1   430.73 446.73
## <none>           428.88 446.88
## + alcohol    1   428.74 448.74
## - typea      1   434.63 450.63
## - ldl        1   437.34 453.34
## - tobacco    1   439.26 455.26
## - age        1   441.29 457.29
## - famhist    1   444.17 460.17
##
## Step:  AIC=445.14
## chd ~ sbp + tobacco + ldl + famhist + typea + obesity + age
##
##             Df Deviance    AIC
## - obesity    1   430.83 444.83
## - sbp        1   431.12 445.12
## <none>           429.14 445.14
## + adiposity  1   428.88 446.88
## + alcohol    1   428.97 446.97
## - typea      1   434.73 448.73
## - ldl        1   438.75 452.75
## - tobacco    1   439.65 453.65
## - famhist    1   444.58 458.58
## - age        1   448.09 462.09
##
## Step:  AIC=444.83
## chd ~ sbp + tobacco + ldl + famhist + typea + age
##
##             Df Deviance    AIC
## - sbp        1   432.37 444.37
## <none>           430.83 444.83
## + obesity    1   429.14 445.14
## + adiposity  1   430.46 446.46
## + alcohol    1   430.67 446.67
## - typea      1   436.03 448.03
## - ldl        1   439.01 451.01
## - tobacco    1   441.47 453.47
## - famhist    1   445.96 457.96
## - age        1   448.99 460.99
##
## Step:  AIC=444.37
## chd ~ tobacco + ldl + famhist + typea + age
##
##             Df Deviance    AIC
## <none>           432.37 444.37
## + sbp        1   430.83 444.83
## + obesity    1   431.12 445.12
## + alcohol    1   432.06 446.06
## + adiposity  1   432.20 446.20
## - typea      1   437.33 447.33
## - ldl        1   440.96 450.96
## - tobacco    1   443.21 453.21
## - famhist    1   447.32 457.32
## - age        1   455.15 465.15
```

```
##
## Call:
## glm(formula = chd ~ tobacco + ldl + famhist + typea + age, family = binomial,
##     data = chd_1)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -1.8638  -0.8128  -0.4311   0.9263   2.6135
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) -5.98647    0.95421  -6.274 3.52e-10 ***
## tobacco      0.08444    0.02665   3.168 0.001534 **
## ldl          0.16578    0.05770   2.873 0.004064 **
## famhist      0.91050    0.23743   3.835 0.000126 ***
## typea        0.02794    0.01272   2.197 0.028023 *
## age          0.04939    0.01079   4.577 4.72e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 540.05  on 419  degrees of freedom
## Residual deviance: 432.37  on 414  degrees of freedom
## AIC: 444.37
##
## Number of Fisher Scoring iterations: 5
```

```
## explore different possible models
mod_int_1 <- glm(chd ~ tobacco + ldl + famhist + typea + obesity + age + ldl*obesity, data = chd_1, fami
summary(mod_int_1)
```

```
##
## Call:
## glm(formula = chd ~ tobacco + ldl + famhist + typea + obesity +
##     age + ldl * obesity, family = binomial, data = chd_1)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -1.7721  -0.8170  -0.4306   0.9039   2.4769
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) -3.91451    2.34982  -1.666 0.095738 .
## tobacco      0.08507    0.02684   3.170 0.001525 **
## ldl         -0.09929    0.42621  -0.233 0.815793
## famhist      0.92712    0.23870   3.884 0.000103 ***
## typea        0.02870    0.01278   2.246 0.024710 *
## obesity     -0.08725    0.08615  -1.013 0.311177
## age          0.05101    0.01080   4.725  2.3e-06 ***
## ldl:obesity  0.01062    0.01583   0.671 0.502085
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 540.05  on 419  degrees of freedom
## Residual deviance: 430.67  on 412  degrees of freedom
## AIC: 446.67
##
## Number of Fisher Scoring iterations: 5
```

```
mod_int_2 <- glm(chd ~ tobacco + ldl + famhist + typea + obesity + age + ldl*age, data = chd_1, family
summary(mod_int_2)
```

```
##
## Call:
## glm(formula = chd ~ tobacco + ldl + famhist + typea + obesity +
##     age + ldl * age, family = binomial, data = chd_1)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -1.8510  -0.8130  -0.4329   0.8940   2.5579
##
## Coefficients:
##               Estimate Std. Error z value Pr(>|z|)
## (Intercept) -5.3878291  1.5217146  -3.541 0.000399 ***
## tobacco      0.0844972  0.0267492   3.159 0.001584 **
## ldl          0.2035399  0.2470537   0.824 0.410014
## famhist      0.9197822  0.2382243   3.861 0.000113 ***
## typea        0.0287266  0.0127710   2.249 0.024490 *
## obesity     -0.0331998  0.0300121  -1.106 0.268634
## age          0.0523645  0.0256088   2.045 0.040876 *
## ldl:age     -0.0004009  0.0050718  -0.079 0.937003
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 540.05  on 419  degrees of freedom
## Residual deviance: 431.11  on 412  degrees of freedom
## AIC: 447.11
##
## Number of Fisher Scoring iterations: 5
```

```
mod_int_3 <- glm(chd ~ tobacco + ldl + famhist + typea + obesity + age + age*obesity, data = chd_1, fami
summary(mod_int_3)
```

```
##
## Call:
## glm(formula = chd ~ tobacco + ldl + famhist + typea + obesity +
##     age + age * obesity, family = binomial, data = chd_1)
##
## Deviance Residuals:
```

```
##     Min       1Q   Median       3Q      Max
## -1.8527  -0.8115  -0.4345   0.8935   2.5228
##
## Coefficients:
##                Estimate Std. Error z value Pr(>|z|)
## (Intercept) -4.8847523  3.2121906  -1.521 0.128337
## tobacco      0.0843979  0.0267588   3.154 0.001610 **
## ldl          0.1846511  0.0608390   3.035 0.002405 **
## famhist      0.9210051  0.2382357   3.866 0.000111 ***
## typea        0.0288457  0.0127512   2.262 0.023685 *
## obesity     -0.0503231  0.1259544  -0.400 0.689499
## age          0.0415311  0.0650417   0.639 0.523129
## obesity:age  0.0003606  0.0025726   0.140 0.888529
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 540.05  on 419  degrees of freedom
## Residual deviance: 431.10  on 412  degrees of freedom
## AIC: 447.1
##
## Number of Fisher Scoring iterations: 5
```

```
mod_sq_1 <- glm(chd ~ tobacco + ldl + famhist + typea + obesity^2 + age, data = chd_1, family = binomial
summary(mod_sq_1)
```

```
##
## Call:
## glm(formula = chd ~ tobacco + ldl + famhist + typea + obesity^2 +
##     age, family = binomial, data = chd_1)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -1.8538  -0.8127  -0.4306   0.8925   2.5488
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) -5.30715    1.12552  -4.715 2.41e-06 ***
## tobacco      0.08449    0.02675   3.159 0.001584 **
## ldl          0.18462    0.06080   3.037 0.002393 **
## famhist      0.92036    0.23814   3.865 0.000111 ***
## typea        0.02880    0.01274   2.259 0.023855 *
## obesity     -0.03318    0.03001  -1.106 0.268851
## age          0.05053    0.01079   4.684 2.81e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 540.05  on 419  degrees of freedom
## Residual deviance: 431.12  on 413  degrees of freedom
## AIC: 445.12
##
```

```
## Number of Fisher Scoring iterations: 5
```

```
mod_sq_2 <- glm(chd ~ tobacco + ldl + famhist^2 + typea + obesity + age, data = chd_1, family = binomial
summary(mod_sq_2)
```

```
##
## Call:
## glm(formula = chd ~ tobacco + ldl + famhist^2 + typea + obesity +
##     age, family = binomial, data = chd_1)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -1.8538  -0.8127  -0.4306   0.8925   2.5488
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) -5.30715    1.12552  -4.715 2.41e-06 ***
## tobacco      0.08449    0.02675   3.159 0.001584 **
## ldl          0.18462    0.06080   3.037 0.002393 **
## famhist      0.92036    0.23814   3.865 0.000111 ***
## typea        0.02880    0.01274   2.259 0.023855 *
## obesity     -0.03318    0.03001  -1.106 0.268851
## age          0.05053    0.01079   4.684 2.81e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 540.05  on 419  degrees of freedom
## Residual deviance: 431.12  on 413  degrees of freedom
## AIC: 445.12
##
## Number of Fisher Scoring iterations: 5
```

```
mod_sq_3 <- glm(chd ~ tobacco + ldl + famhist + typea + obesity + age^2, data = chd_1, family = binomial
summary(mod_sq_3)
```

```
##
## Call:
## glm(formula = chd ~ tobacco + ldl + famhist + typea + obesity +
##     age^2, family = binomial, data = chd_1)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -1.8538  -0.8127  -0.4306   0.8925   2.5488
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) -5.30715    1.12552  -4.715 2.41e-06 ***
## tobacco      0.08449    0.02675   3.159 0.001584 **
## ldl          0.18462    0.06080   3.037 0.002393 **
## famhist      0.92036    0.23814   3.865 0.000111 ***
## typea        0.02880    0.01274   2.259 0.023855 *
## obesity     -0.03318    0.03001  -1.106 0.268851    
## age          0.05053    0.01079   4.684 2.81e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 
## (Dispersion parameter for binomial family taken to be 1)
## 
##     Null deviance: 540.05  on 419  degrees of freedom
## Residual deviance: 431.12  on 413  degrees of freedom
## AIC: 445.12
## 
## Number of Fisher Scoring iterations: 5
```

```
# none of them reaches a lower AIC score than the one without any interaction terms or transformation o

## model fit
mod_fin <- glm(chd ~ tobacco + ldl + famhist + typea + obesity + age, data = chd_1, family = binomial)
summary(mod_fin)
```

```
## 
## Call:
## glm(formula = chd ~ tobacco + ldl + famhist + typea + obesity + 
##     age, family = binomial, data = chd_1)
## 
## Deviance Residuals: 
##     Min       1Q   Median       3Q      Max  
## -1.8538  -0.8127  -0.4306   0.8925   2.5488  
## 
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)    
## (Intercept) -5.30715    1.12552  -4.715 2.41e-06 ***
## tobacco      0.08449    0.02675   3.159 0.001584 ** 
## ldl          0.18462    0.06080   3.037 0.002393 ** 
## famhist      0.92036    0.23814   3.865 0.000111 ***
## typea        0.02880    0.01274   2.259 0.023855 *  
## obesity     -0.03318    0.03001  -1.106 0.268851    
## age          0.05053    0.01079   4.684 2.81e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 
## (Dispersion parameter for binomial family taken to be 1)
## 
##     Null deviance: 540.05  on 419  degrees of freedom
## Residual deviance: 431.12  on 413  degrees of freedom
## AIC: 445.12
## 
## Number of Fisher Scoring iterations: 5
```

```
#Log-likelihood ratio test
lrtest(stepAIC(mod, trace = 1, direction = "both"), mod_fin)
```

```
## Start:  AIC=448.74
## chd ~ sbp + tobacco + ldl + adiposity + famhist + typea + obesity + 
##     alcohol + age
##
##             Df Deviance    AIC
## - alcohol    1   428.88 446.88
## - adiposity  1   428.97 446.97
## - obesity    1   430.28 448.28
## - sbp        1   430.44 448.44
## <none>           428.74 448.74
## - typea      1   434.42 452.42
## - ldl        1   437.34 455.34
## - tobacco    1   438.37 456.37
## - age        1   441.28 459.28
## - famhist    1   443.70 461.70
##
## Step:  AIC=446.88
## chd ~ sbp + tobacco + ldl + adiposity + famhist + typea + obesity +
##     age
##
##             Df Deviance    AIC
## - adiposity  1   429.14 445.14
## - obesity    1   430.46 446.46
## - sbp        1   430.73 446.73
## <none>           428.88 446.88
## + alcohol    1   428.74 448.74
## - typea      1   434.63 450.63
## - ldl        1   437.34 453.34
## - tobacco    1   439.26 455.26
## - age        1   441.29 457.29
## - famhist    1   444.17 460.17
##
## Step:  AIC=445.14
## chd ~ sbp + tobacco + ldl + famhist + typea + obesity + age
##
##             Df Deviance    AIC
## - obesity    1   430.83 444.83
## - sbp        1   431.12 445.12
## <none>           429.14 445.14
## + adiposity  1   428.88 446.88
## + alcohol    1   428.97 446.97
## - typea      1   434.73 448.73
## - ldl        1   438.75 452.75
## - tobacco    1   439.65 453.65
## - famhist    1   444.58 458.58
## - age        1   448.09 462.09
##
## Step:  AIC=444.83
## chd ~ sbp + tobacco + ldl + famhist + typea + age
##
##             Df Deviance    AIC
## - sbp        1   432.37 444.37
## <none>           430.83 444.83
## + obesity    1   429.14 445.14
## + adiposity  1   430.46 446.46
## + alcohol    1   430.67 446.67
## - typea      1   436.03 448.03
## - ldl        1   439.01 451.01
## - tobacco    1   441.47 453.47
## - famhist    1   445.96 457.96
## - age        1   448.99 460.99
##
## Step:  AIC=444.37
## chd ~ tobacco + ldl + famhist + typea + age
##
##             Df Deviance    AIC
## <none>          432.37 444.37
## + sbp        1   430.83 444.83
## + obesity    1   431.12 445.12
## + alcohol    1   432.06 446.06
## + adiposity  1   432.20 446.20
## - typea      1   437.33 447.33
## - ldl        1   440.96 450.96
## - tobacco    1   443.21 453.21
## - famhist    1   447.32 457.32
## - age        1   455.15 465.15
```

```
## Likelihood ratio test
##
## Model 1: chd ~ tobacco + ldl + famhist + typea + age
## Model 2: chd ~ tobacco + ldl + famhist + typea + obesity + age
##   #Df  LogLik Df  Chisq Pr(>Chisq)
## 1   6 -216.18
## 2   7 -215.56  1 1.2455     0.2644
```

```
#Manually computed R-squared
with(summary(mod_fin), 1 - deviance/null.deviance)
```

```
## [1] 0.2016962
```

```
with(summary(mod_fin), deviance/null.deviance)
```

```
## [1] 0.7983038
```

```
#AIC scores
summary(stepAIC(mod, trace = 1, direction = "both"))
```

```
## Start:  AIC=448.74
## chd ~ sbp + tobacco + ldl + adiposity + famhist + typea + obesity +
##     alcohol + age
##
##             Df Deviance    AIC
## - alcohol    1   428.88 446.88
## - adiposity  1   428.97 446.97
## - obesity    1   430.28 448.28
## - sbp        1   430.44 448.44
## <none>          428.74 448.74
## - typea      1   434.42 452.42
## - ldl        1   437.34 455.34
## - tobacco    1   438.37 456.37
## - age        1   441.28 459.28
## - famhist    1   443.70 461.70
##
## Step:  AIC=446.88
## chd ~ sbp + tobacco + ldl + adiposity + famhist + typea + obesity +
##     age
##
##             Df Deviance    AIC
## - adiposity  1   429.14 445.14
## - obesity    1   430.46 446.46
## - sbp        1   430.73 446.73
## <none>           428.88 446.88
## + alcohol    1   428.74 448.74
## - typea      1   434.63 450.63
## - ldl        1   437.34 453.34
## - tobacco    1   439.26 455.26
## - age        1   441.29 457.29
## - famhist    1   444.17 460.17
##
## Step:  AIC=445.14
## chd ~ sbp + tobacco + ldl + famhist + typea + obesity + age
##
##             Df Deviance    AIC
## - obesity    1   430.83 444.83
## - sbp        1   431.12 445.12
## <none>           429.14 445.14
## + adiposity  1   428.88 446.88
## + alcohol    1   428.97 446.97
## - typea      1   434.73 448.73
## - ldl        1   438.75 452.75
## - tobacco    1   439.65 453.65
## - famhist    1   444.58 458.58
## - age        1   448.09 462.09
##
## Step:  AIC=444.83
## chd ~ sbp + tobacco + ldl + famhist + typea + age
##
##             Df Deviance    AIC
## - sbp        1   432.37 444.37
## <none>           430.83 444.83
## + obesity    1   429.14 445.14
## + adiposity  1   430.46 446.46
## + alcohol    1   430.67 446.67
## - typea      1   436.03 448.03
## - ldl        1   439.01 451.01
## - tobacco    1   441.47 453.47
## - famhist    1   445.96 457.96
## - age        1   448.99 460.99
##
## Step:  AIC=444.37
## chd ~ tobacco + ldl + famhist + typea + age
```

```
##
##             Df Deviance    AIC
## <none>           432.37 444.37
## + sbp        1   430.83 444.83
## + obesity    1   431.12 445.12
## + alcohol    1   432.06 446.06
## + adiposity  1   432.20 446.20
## - typea      1   437.33 447.33
## - ldl        1   440.96 450.96
## - tobacco    1   443.21 453.21
## - famhist    1   447.32 457.32
## - age        1   455.15 465.15
```

```
##
## Call:
## glm(formula = chd ~ tobacco + ldl + famhist + typea + age, family = binomial,
##     data = chd_1)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -1.8638  -0.8128  -0.4311   0.9263   2.6135
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) -5.98647    0.95421  -6.274 3.52e-10 ***
## tobacco      0.08444    0.02665   3.168 0.001534 **
## ldl          0.16578    0.05770   2.873 0.004064 **
## famhist      0.91050    0.23743   3.835 0.000126 ***
## typea        0.02794    0.01272   2.197 0.028023 *
## age          0.04939    0.01079   4.577 4.72e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 540.05  on 419  degrees of freedom
## Residual deviance: 432.37  on 414  degrees of freedom
## AIC: 444.37
##
## Number of Fisher Scoring iterations: 5
```

```
summary(mod_fin)
```

```
##
## Call:
## glm(formula = chd ~ tobacco + ldl + famhist + typea + obesity +
##     age, family = binomial, data = chd_1)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -1.8538  -0.8127  -0.4306   0.8925   2.5488
##
## Coefficients:
```

```
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) -5.30715    1.12552  -4.715 2.41e-06 ***
## tobacco      0.08449    0.02675   3.159 0.001584 **
## ldl          0.18462    0.06080   3.037 0.002393 **
## famhist      0.92036    0.23814   3.865 0.000111 ***
## typea        0.02880    0.01274   2.259 0.023855 *
## obesity     -0.03318    0.03001  -1.106 0.268851
## age          0.05053    0.01079   4.684 2.81e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 540.05  on 419  degrees of freedom
## Residual deviance: 431.12  on 413  degrees of freedom
## AIC: 445.12
##
## Number of Fisher Scoring iterations: 5
```

```
## prediction accuracy
#Get the predicted values (probabilities)
pdata <- predict(mod_fin, type = "response")
#Convert probability to 0/1 prediction for each case
pclass.5 = as.factor(as.numeric(pdata>0.5))
#Look at the cross-tabulation
table(pclass.5,chd_1$chd)
```

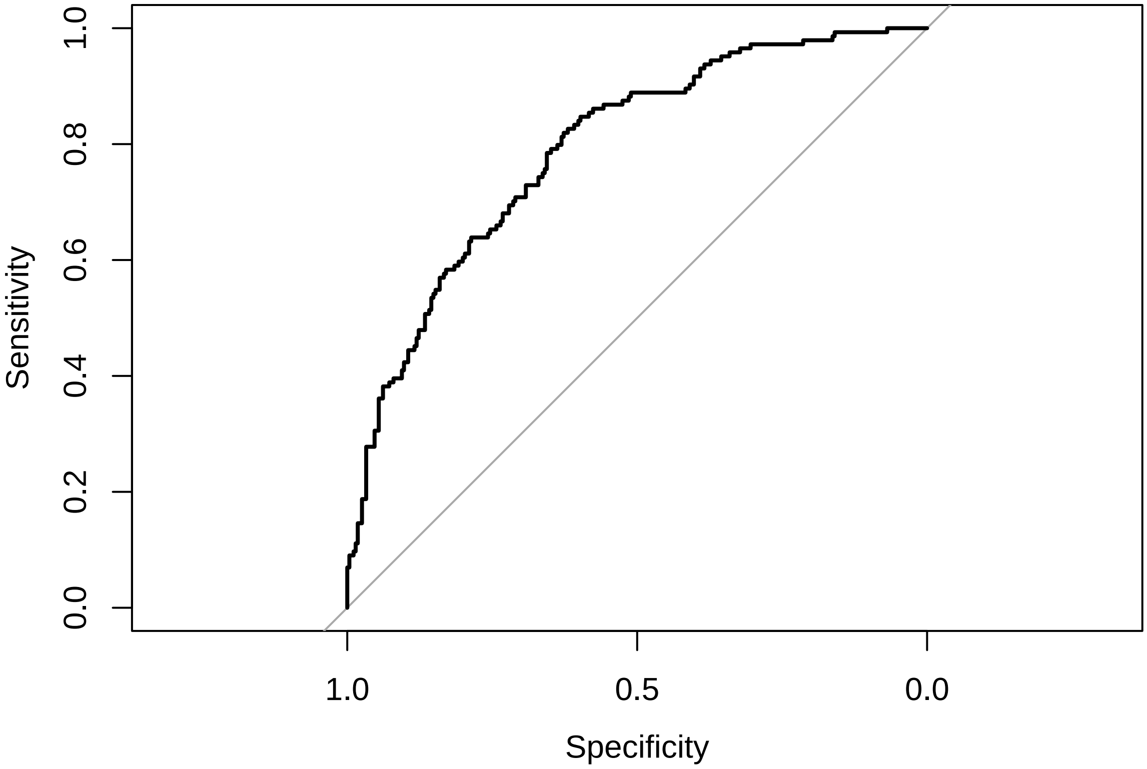
```
##
## pclass.5   0   1
##        0 235  67
##        1  41  77
```

```
#Confusion matrix and related rates/statistics
confusionMatrix(pclass.5,
                reference=as.factor(chd_1$chd),
                positive="1")
```

```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction   0   1
##          0 235  67
##          1  41  77
##
##                Accuracy : 0.7429
##                  95% CI : (0.6982, 0.784)
##     No Information Rate : 0.6571
##     P-Value [Acc > NIR] : 9.587e-05
##
##                   Kappa : 0.4036
##
##  Mcnemar's Test P-Value : 0.01614
##
```

```
##            Sensitivity : 0.5347
##            Specificity : 0.8514
##         Pos Pred Value : 0.6525
##         Neg Pred Value : 0.7781
##             Prevalence : 0.3429
##         Detection Rate : 0.1833
##   Detection Prevalence : 0.2810
##      Balanced Accuracy : 0.6931
##
##       'Positive' Class : 1
##
```

```
#ROC-AUC
chd_1$phat <- predict(mod_fin, chd_1, type = "response")
roc(factor(chd_1$chd), chd_1$phat, plot=TRUE, auc.print=TRUE)
```



```
##
## Call:
## roc.default(response = factor(chd_1$chd), predictor = chd_1$phat,      plot = TRUE, auc.print = TRUE)
##
## Data: chd_1$phat in 276 controls (factor(chd_1$chd) 0) < 144 cases (factor(chd_1$chd) 1).
## Area under the curve: 0.7904
```

```r
## coefficient interpretation
intercept <- 1/(1+exp(-(mod_fin$coefficients[1]))) ## 0.00493
tobacco <- exp(mod_fin$coefficients[2]) ## 1.09
ldl <- exp(mod_fin$coefficients[3]) ## 1.2
famhist <- exp(mod_fin$coefficients[4]) ## 2.51
typea <- exp(mod_fin$coefficients[5]) ## 1.03
obesity <- exp(mod_fin$coefficients[6]) ## 0.967
age <- exp(mod_fin$coefficients[7]) ## 1.05
tobacco_1 <- 1/(1+exp(-(mod_fin$coefficients[2]))) ## 0.521
ldl_1 <- 1/(1+exp(-(mod_fin$coefficients[3]))) ## 0.546
famhist_1 <- 1/(1+exp(-(mod_fin$coefficients[4]))) ## 0.715
typea_1 <- 1/(1+exp(-(mod_fin$coefficients[5]))) ## 0.507
obesity_1 <- 1/(1+exp(-(mod_fin$coefficients[6]))) ## 0.492
age_1 <- 1/(1+exp(-(mod_fin$coefficients[7]))) ## 0.513

## important predictors
```