Laboration report in Machine Learning

Computer lab 1 block 1

732A99

Sigme Cinar
Duc Tran
William Wiik

Division of Statistics and Machine Learning
Department of Computer Science
Linköping University

13 november 2023

Contents

1 Assignment 1. Handwritten digit recognition with K-nearest neighbors.

The data in this task is from the file optdigits.csv. Data consists of 3822 handwritten digits from 0 to 9 and are stored as images of size 8x8.

1.1 1.1

Question: Import the data into R and divide it into training, validation and test sets (50%/25%/25%) by using the partitioning principle specified in the lecture slides.

Answer: The code used is presented as follows:

```
# Read in data
data <- read.csv("optdigits.csv")

# Renaming the response variable and changing it to a factor variable
data <- rename(data, y=X0.26)
data$y <- as.factor(data$y)

# Partitioning training data (50%)
n=dim(data)[1]
set.seed(12345)
id=sample(1:n, floor(n*0.5))
train=data[id,]

# Partitioning validation data (25%)
id1=setdiff(1:n, id)
set.seed(12345)
id2=sample(id1, floor(n*0.25))
valid=data[id2,]

# Partitioning test data (25%)
id3=setdiff(id1,id2)
test=data[id3,]
```

1.2 1.2

Question: Use training data to fit 30-nearest neighbor classifier with function kknn() and kernel="rectangular" from package kknn and estimate

- Confusion matrices for the training and test data (use table())
- Misclassification errors for the training and test data

Answer: The confusion matrix for the model trained on training data with k=30 and evaluated on training data is presented in table 1.

```
# kknn on training data and evaluation on training data
model_kknn_train <- kknn(formula=y~., train=train, test=train, kernel="rectangular", k=30)
conf_mat_train <- table(train$y, model_kknn_train$fitted.values)
acc_train <- sum(diag(conf_mat_train)) / sum(conf_mat_train)
miss_train <- 1-acc_train
```

```r
# kknn on training data and evaluation on test data
model_kknn_test <- kknn(formula=y~., train=train, test=test, kernel="rectangular", k=30)
conf_mat_test <- table(test$y, model_kknn_test$fitted.values)
acc_test <- sum(diag(conf_mat_test)) / sum(conf_mat_test)
miss_test <- 1-acc_test

# Rows are true values, columns are model prediction
kable(conf_mat_train, caption = "Confusion matrix for training data, model
      predictions by columns and true value by rows.")
```

Table 1: Confusion matrix for training data, model predictions by columns and true value by rows.

	0	1	2	3	4	5	6	7	8	9
0	177	0	0	0	1	0	0	0	0	0
1	0	174	9	0	0	0	1	0	1	3
2	0	0	170	0	0	0	0	1	2	0
3	0	0	0	197	0	2	0	1	0	0
4	0	1	0	0	166	0	2	6	2	2
5	0	0	0	0	0	183	1	2	0	11
6	0	0	0	0	0	0	200	0	0	0
7	0	1	0	1	0	1	0	192	0	0
8	0	10	0	1	0	0	2	0	190	2
9	0	3	0	4	2	0	0	2	4	181

```r
miss_train
```

```
## [1] 0.04238619
```

The misclassification error on training data from table 1 is around 4.24%. The two number with the highest number of wrong predictions are 8 and 9, with 15 wrong predictions each. The model also struggles to predict number 1 and 5 where both had 14 wrong predictions. Examining the most common misclassification, the number 5 was predicted as 9 a total of eleven times and the number 8 was predicted as 1 a total of ten times.

The easiest numbers to predict are 0 and 6. For 0 the model correctly predicted 177 out of 178 numbers and did not predict the number 0 for any other number. For the number 6, the model correctly predicted all 200 numbers, but it did however predict six other numbers as 6 incorrectly.

The confusion matrix for the model trained on training data with k=30 and evaluated on test data is presented in table 2.

```r
kable(conf_mat_test, caption = "Confusion matrix for test data, model
      predictions by columns and true value by rows.")
```

Table 2: Confusion matrix for test data, model predictions by columns and true value by rows.

	0	1	2	3	4	5	6	7	8	9
0	97	0	0	0	0	0	1	0	0	0
1	0	91	3	0	0	0	0	0	0	3
2	0	0	93	1	0	0	0	0	1	0

	0	1	2	3	4	5	6	7	8	9
3	0	0	0	95	0	0	0	2	1	0
4	1	0	0	0	89	0	1	5	1	3
5	0	1	0	1	0	79	1	0	0	5
6	0	0	0	0	0	0	94	0	0	0
7	0	2	0	0	0	1	0	91	1	0
8	0	3	0	1	0	0	1	0	86	0
9	0	0	0	4	0	0	0	2	1	94

```
miss_test
```

```
## [1] 0.04916318
```

The misclassification error on test data from table 2 is around 4.92%. The two number with the highest number of wrong predictions are 4 and 5. The number 4 had ten wrong predictions and the number 5 had eight wrong predictions.

The easiest numbers to predict are 0 and 6. For 0 the model correctly predicted 97 out of 98 number and the misclassification was as number 6.
For the number 6, the model correctly predicted all 94 numbers, but it did however predict four times for other numbers as 6 incorrectly.

The overall prediction quality from the model is good, especially for the numbers 0 and 6.

1.3 1.3

Question: Find any 2 cases of digit "8" in the training data which were easiest to classify and 3 cases that were hardest to classify (i.e. having highest and lowest probabilities of the correct class). Reshape features for each of these cases as matrix 8x8 and visualize the corresponding digits (by using e.g. heatmap() function with parameters Colv=NA and Rowv=NA) and comment on whether these cases seem to be hard or easy to recognize visually.

Answer: The code used to find the 2 digits that are hardest to classify were found with the code as follows

```
y <- train$y
fit_y <- model_kknn_train$fitted.values
# probabilities given from number 0 to 9, index 9 = number 8.
prob_8 <- model_kknn_train$prob[, 9]

# Data frame consisting of true value of y, model prediction and the models
# probability that the number is 8.
data_8 <- data.frame(y = y, fit_y = fit_y, prob = prob_8)
data_8$observation_id <- rownames(data_8)

# Only observations with the label 8 is kept.
data_8 <- data_8[data_8$y == "8", ]
head(arrange(data_8, prob), 2)
```

```
##   y fit_y      prob observation_id
## 1 8     6 0.1000000           1624
## 2 8     1 0.1666667           1663
```

From the output, observation 1624 and 1663 were hardest to classify as 8 from the model. The three observations that were easiest to identify as 8 were found with the code as follows

```
tail(arrange(data_8, prob), 3)
```

```
##     y fit_y prob observation_id
## 203 8     8    1           1810
## 204 8     8    1           1811
## 205 8     8    1           1864
```

From the output observation 1810, 1811, and 1864 were three observations that were easiest to identify as 8 with a probability from the model as 100% (in total there were 49 observations that had 100% probability).

A function that reshapes each observation to a 8x8 cases and then visualizing the result in a heatmap was done with the code as follows

```
# Change colour palette to black and white
colfunc <- colorRampPalette(c("white", "black"))

plot_8 <- function(index){
  title <- paste0("Obs: ", index)
  # Reshapes the observations to a 8x8
  plot <- as.matrix(train[index, -65]) # Remove response variable
  plot <- matrix(plot, nrow=8, byrow=TRUE)
  heatmap(plot, col=colfunc(16), Colv=NA, Rowv=NA, main=title, margins=c(2,2))
}
```

The heatmaps for observations 1624 and 1663 are presented in figure 1.

```
plot_8(1624)
plot_8(1663)
```

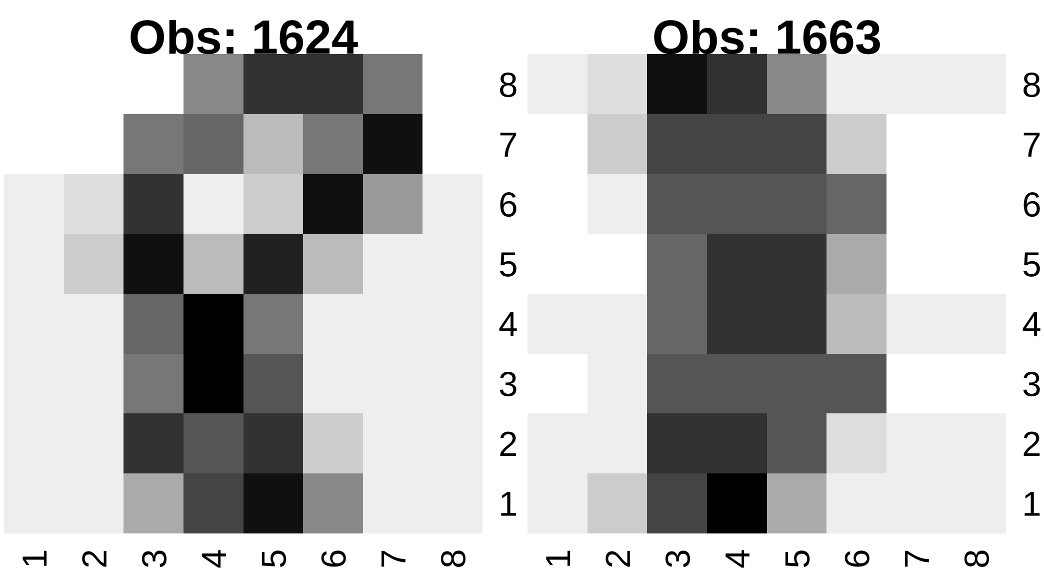


Figure 1: Heatmap for two observations that were hard to classify: 1624 and 1663.

In figure 1, it is hard to visually recognize what number the observations 1624 and 1663 are.

The heatmaps for observations 1810, 1811, and 1864 are presented in figure 2.

```
plot_8(1810)
plot_8(1811)
plot_8(1864)
```

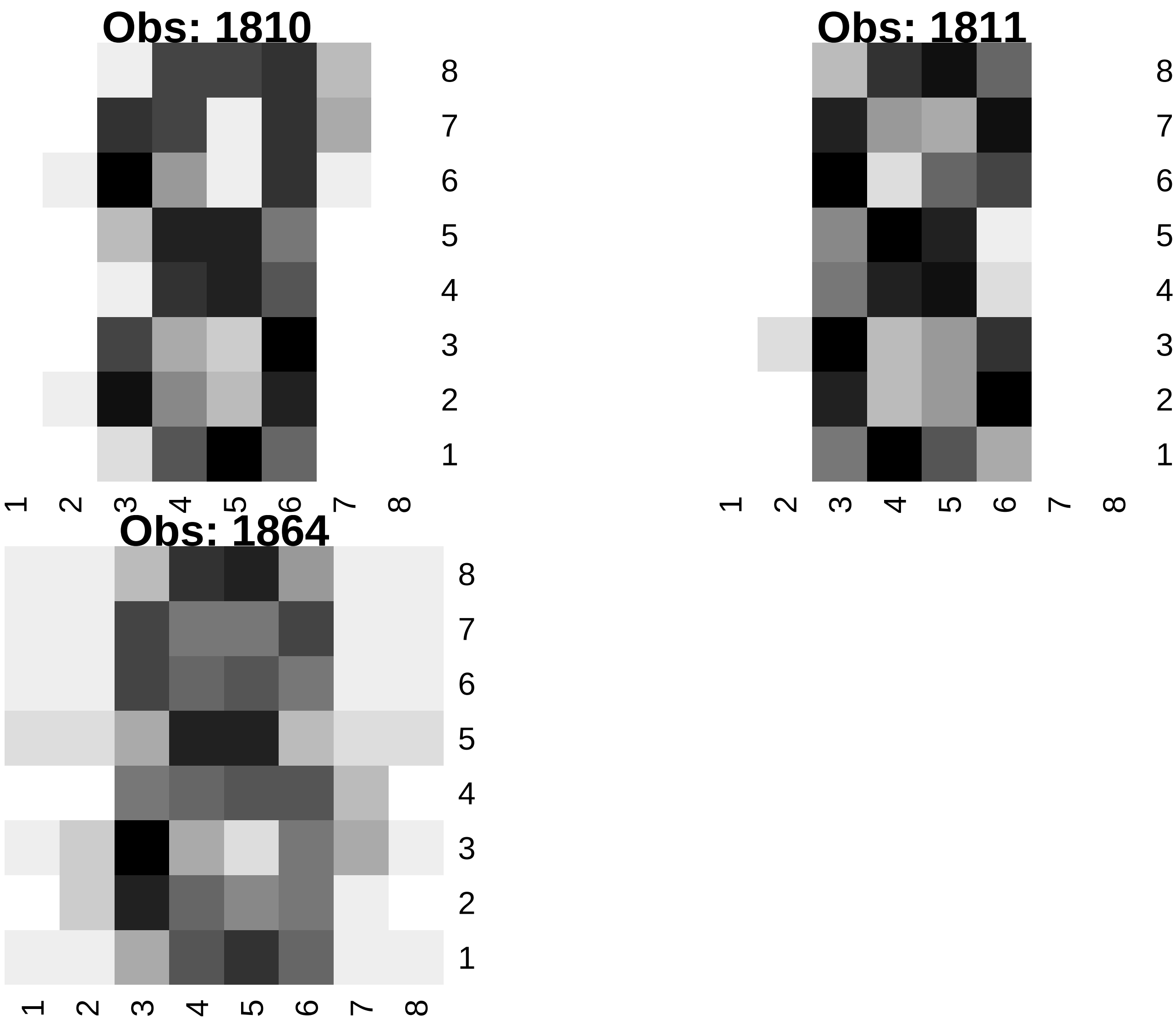


Figure 2: Heatmap for three observations that were easy to classify: 1810, 1811, and 1864.

In figure 2, is is easy to visually recognize that observations 1810, 1811, and 1864 are of the number 8.

1.4 1.4

Question: Fit a K-nearest neighbor classifiers to the training data for different values of $K = 1, 2, ..., 30$ and plot the dependence of the training and validation misclassification errors on the value of K (in the same plot). How does the model complexity change when K increases and how does it affect the training and validation errors? Report the optimal k according to this plot. Finally, estimate the test error for the model having the optimal K, compare it with the training and validation errors and make necessary conclusions about the model quality.

Answer: The code to create K-nearest neighbor clasifiers for different k and the misclassification error on training and validation is as follows

```
fit_kknn <- function(k){
  model_kknn_train <- kknn(formula=y~., train=train, test=train, kernel="rectangular", k=k)
  # Confusion matrix for train data
  conf_mat_train <- table(model_kknn_train$fitted.values, train$y)
  acc_train <- sum(diag(conf_mat_train)) / sum(conf_mat_train)
  # Missclassification for training data
  miss_train <- 1-acc_train

  model_kknn_valid <- kknn(formula=y~., train=train, test=valid, kernel="rectangular", k=k)
  # Confusion matrix for validation data
  conf_mat_valid <- table(model_kknn_valid$fitted.values, valid$y)
  acc_valid <- sum(diag(conf_mat_valid)) / sum(conf_mat_valid)
  # Missclassification for validation data
  miss_valid <- 1-acc_valid

  result <- c(miss_train, miss_valid)
  return(result)
}

# Missclassification for k=1,...,30 for training and validation data
result <- data.frame(train = 0, valid = 0)
for(i in 1:30){
  model <- fit_kknn(i)
  result[i,1] <- model[1]
  result[i,2] <- model[2]
}
result$index <- 1:30
```

The plot showing the dependence of misclassification error for training and validation data is presented in figure 3.

```
ggplot(result, aes(x=index)) +
  geom_line(aes(y=train, colour="train")) +
  geom_point(aes(y=train, colour="train")) +
  geom_line(aes(y=valid, colour="valid")) +
  geom_point(aes(y=valid, colour="valid")) +
  scale_color_manual(name = "Data",
                     values = c("train" = "steelblue", "valid" = "indianred")) +
  scale_x_continuous(breaks = c(seq(from=0, to=30, by=5))) +
```

```r
  scale_y_continuous(limits = c(0, 0.06)) +
  theme_bw() +
  labs(x = "k",
       y = "Missclassification rate")
```

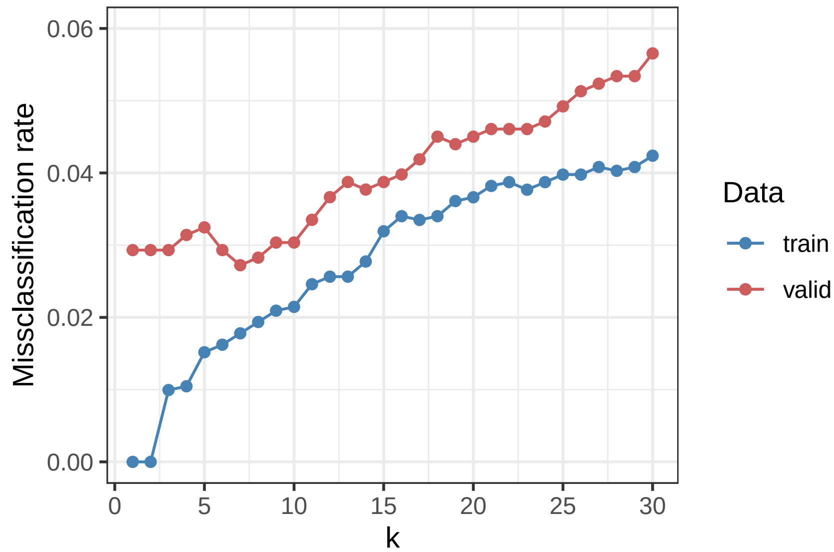


Figure 3: Misclassification errors for training and validation data for k=1,. . . ,30 on model trained on training data.

In figure 3, the model complexity is highest when k = 1 and decreases with larger k. With k = 1, the prediction in the model is by the observation in training data closest to the value we want to predict and when k is equal to the number of observations, the model will always predict the same value (except situations with ties). The training error increases when the model complexity decreases, this is because the model will be less overfitted on the training data with larger k and after some k, the model will be underfitted. Validation error also generally increased when the model complexity decreased, however at some k, the model will be between overfitted and underfitted which will give lowest validation error. This happened at k=7, which is considered to be the optimal k.

```r
which(result$valid == min(result$valid))
```

```
## [1] 7
```

With k=7, the error for the test data is calculated. The errors for training, validation, and test data are compared in table 3.

```r
model_test_7 <- kknn(formula = y~., train = train, test = test, kernel = "rectangular", k=7)

conf_mat_test <- table(model_test_7$fitted.values, test$y)
acc_test <- sum(diag(conf_mat_test)) / sum(conf_mat_test)
miss_test <- 1-acc_test

table_data <- cbind(training=result[7, 1], validtion=result[7, 2], test=miss_test)
kable(table_data, digits=3, caption="Misclassification error k=7 for different data.")
```

Table 3: Misclassification error k=7 for different data.

training	validtion	test
0.018	0.027	0.039

In table 3, the error for training is the lowest, followed by validation, and then test. Since the error for training is decently lower the model is a bit overfitted on training data. The difference between validation and test can be interpreted that for k=7 the error is smallest for the validation data, but it might not be the best since there is a bias when picking k from validation data.

1.5 1.5

Question: Fit K-nearest neighbor classifiers to the training data for different values of $K = 1, 2, ..., 30$, compute the error for the validation data as cross-entropy (when computing log of probabilities add a small constant within log, e.g. 1e-15, to avoid numerical problems) and plot the dependence of the validation error on the value of K. What is the optimal K value here? Assuming that response has multinomial distribution, why might the cross-entropy be a more suitable choice of the error function than the misclassification error for this problem?

Answer: The code used to compute cross-entropy for validation data

```
cross_entropy <- function(k){
  model_kknn_valid <-
    kknn(formula = y~.,
         train = train,
         test = valid,
         kernel = "rectangular",
         k=k)

  y <- as.integer(valid$y)

  prob <- c()
  for(i in 1:length(y)){
    prob[i] <- model_kknn_valid$prob[i, y[i]]
  }

  value <- -sum(log(prob + 1e-15))
  return(value)
}

result <- c()
for(i in 1:30){
  model <- cross_entropy(i)
  result[i] <- model
}
```

The cross-entropy for different k is presented in figure 4.

```
plot_data <- data.frame(index=1:30, result)
ggplot(plot_data, aes(x=index, y=result)) +
  geom_point(color="forestgreen") +
```

```r
  geom_line(color="forestgreen") +
  scale_x_continuous(breaks = c(seq(from=0, to=30, by=5))) +
  theme_bw() +
  labs(x="K",
       y="Cross-entropy")
```

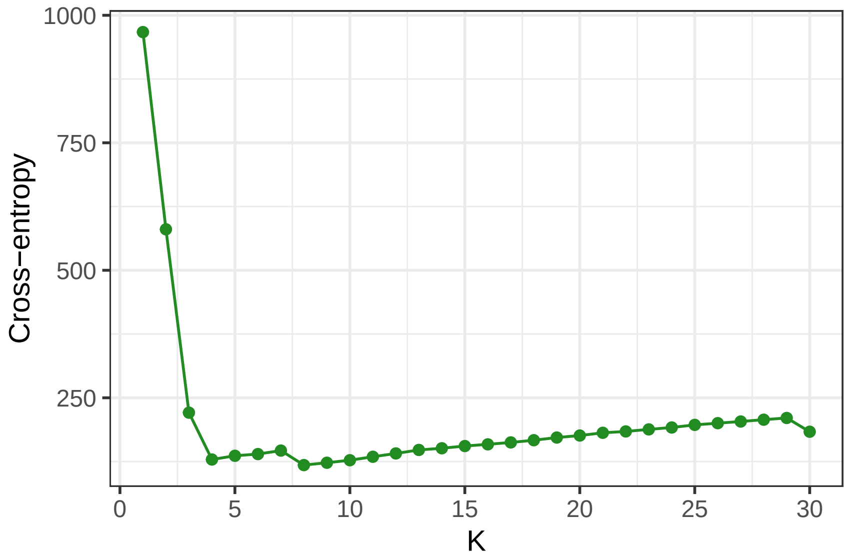


Figure 4: Cross-entropy error for validation data for different values of k for kknn models.

```r
which(min(result) == result)
```

```
## [1] 8
```

From figure 4, the model with k=8 has the lowest value for cross-entropy and is considered to be the best k.

If the response has multinomial distribution, the maximum likelihood estimation is:

$$L(Y_i = C_1, Y_i = C_2, ..., Y_i = C_m|\theta) = \prod_{i=1}^{N} p_\theta(Y_i = C_m) \tag{1}$$

where Y_i is observation i, N is number of observations, and C_m is class m.

The log-likelihood of equation 1 is:

$$logL(Y_i = C_1, Y_i = C_2, ..., Y_i = C_m) = \sum_{i=1}^{N} logp_\theta(Y_i = C_m) \tag{2}$$

```r
df2 <- read.csv("parkinsons.csv")

# Shuffle the data
set.seed(123)
df2 <- df2[sample(nrow(df2)), ]
```

1.5.1 Question 1,2)

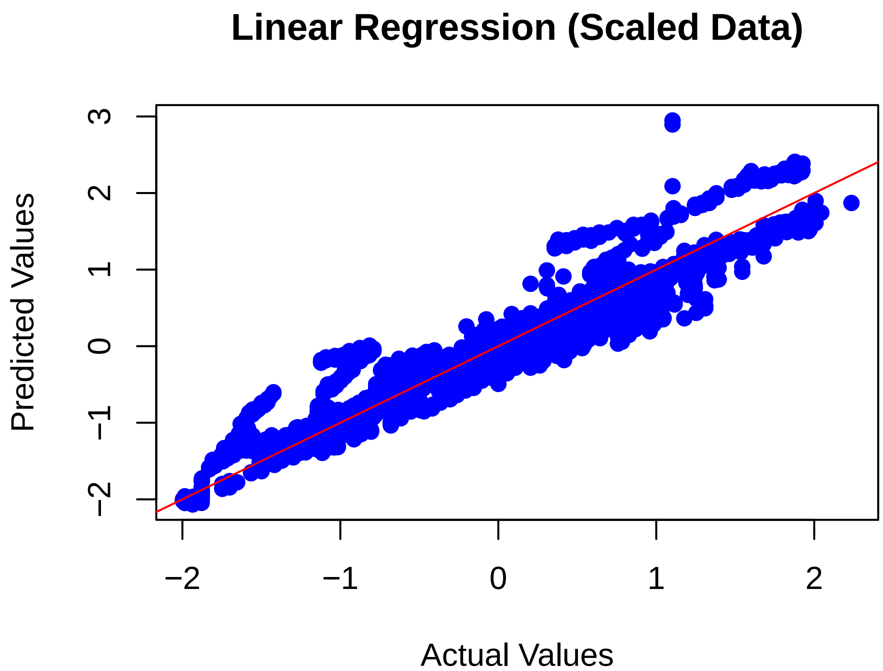
```r
set.seed(123)
# Split train and test
train_indices <- createDataPartition(df2$motor_UPDRS, p = 0.6, list = FALSE)
train_data <- df2[train_indices, ]
test_data <- df2[-train_indices, ]

predictor_cols <- setdiff(names(train_data), "motor_UPDRS")
scaler <- preProcess(train_data)
trainS <- predict(scaler, train_data)
testS <- predict(scaler, test_data)

#train_sd <- apply(train_data, 2, sd)

# Linear regression model
lm_model <- lm(motor_UPDRS ~ ., data = trainS)

# Predictions on the test data
trainS_x <- trainS[, predictor_cols]
testS_x <- testS[, predictor_cols]

predS_train <- predict(lm_model, newdata = trainS_x)
predS_test <- predict(lm_model, newdata = testS_x)

mse_train <- mean((trainS$motor_UPDRS - predS_train)^2)
mse_test <- mean((testS$motor_UPDRS - predS_test)^2)

cat("Mean Squared Error (MSE) on the training data:", mse_train, "\n")
```

```
## Mean Squared Error (MSE) on the training data: 0.09478013
```

```r
cat("Mean Squared Error (MSE) on the test data:", mse_test, "\n")
```

```
## Mean Squared Error (MSE) on the test data: 0.09364679
```

```r
plot(testS$motor_UPDRS, predS_test, main = "Linear Regression (Scaled Data)",
     xlab = "Actual Values", ylab = "Predicted Values", pch = 19, col = "blue")
abline(a = 0, b = 1, col = "red")
```



1.5.2 Question 3)

```
# Define the functions
Loglikelihood <- function(theta, std){
  n <- nrow(trainS_x)
  prediction <- as.matrix(trainS_x) %*% as.matrix(theta)
  actual <- trainS$motor_UPDRS
  res <- actual-prediction
  likelihood <- (-(n/2) * log(2*pi*std^2) - (1/(2*std^2)) * sum(res^2))
  return(likelihood)
}

Ridge <- function(theta, std, lambda){
  likelihood_ridge <- -Loglikelihood(theta, std) + (lambda/2)*sum(theta^2)
  return(likelihood_ridge)
}

#optim() function minimizes
RidgeOpt <- function(lambda){
  # Define a new function to optimize
  my_fnc <- function(parameters){
    theta <- parameters[1:(length(parameters)-1)]
    std <- parameters[length(parameters)]
    return(Ridge(theta, std, lambda))
  }
  initial_values <- c(rep(0, ncol(trainS_x)), 1)
```

```r
  optimal_values <- optim(par = initial_values, fn = my_fnc, method = "BFGS")$par
  optimal_theta <- optimal_values[1:length(predictor_cols)]
  optimal_std <- optimal_values[length(predictor_cols) + 1]
  optimal_lambda <- optimal_values[length(optimal_values)]

  result_list <- list(theta = optimal_theta, std = optimal_std, lambda = optimal_lambda)
  return(result_list)
}
```

Formula for the degree of freedom in ridge regression is as follows:

$$df(\lambda) = trace(X(X^TX + \lambda I)^{-1}X^T)$$

```r
library(psych)
DF <- function(lambda){
  X <- as.matrix(trainS_x)
  dof <- tr(X %*% (solve(t(X) %*% X + lambda*diag(ncol(trainS_x)))) %*% t(X))
  return(dof)
}
```

1.5.3 Question 4)

```r
lambda <- 1
result_ridge_1 <- RidgeOpt(lambda)

optimal_theta_1 <- result_ridge_1$theta
optimal_std_1 <- result_ridge_1$std
optimal_lambda_1 <- result_ridge_1$lambda

ridge_pred_train_1 <- as.matrix(trainS_x) %*% as.matrix(optimal_theta_1)
ridge_pred_test_1 <- as.matrix(testS_x) %*% as.matrix(optimal_theta_1)

mse_ridge_train_1 <- mean((trainS$motor_UPDRS - ridge_pred_train_1)^2)
mse_ridge_test_1 <- mean((testS$motor_UPDRS - ridge_pred_test_1)^2)

cat("Lambda:", lambda, "\n")
```

```
## Lambda: 1
```

```r
cat("Mean Squared Error (MSE) ridge regression on training data:", mse_ridge_train_1, "\n")
```

```
## Mean Squared Error (MSE) ridge regression on training data: 0.09481838
```

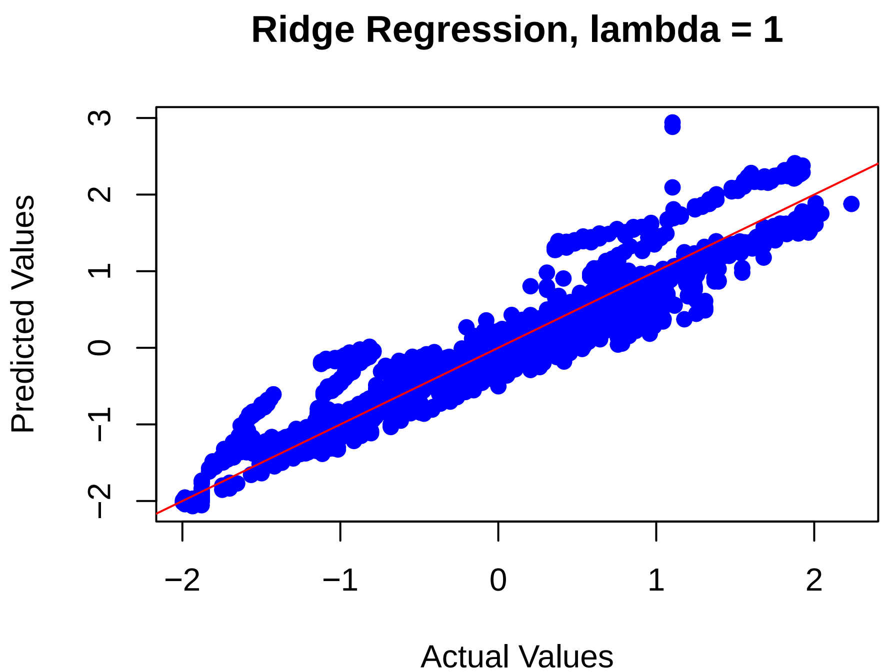
```r
cat("Mean Squared Error (MSE) ridge regression on test data:", mse_ridge_test_1, "\n")
```

```
## Mean Squared Error (MSE) ridge regression on test data: 0.09360224
```

```r
cat("Degree of freedom:", DF(lambda), "\n")
```

```
## Degree of freedom: 18.84537
```

```r
plot(testS$motor_UPDRS, ridge_pred_test_1, main = "Ridge Regression, lambda = 1",
     xlab = "Actual Values", ylab = "Predicted Values", pch = 19, col = "blue")
abline(a = 0, b = 1, col = "red")
```




```r
lambda <- 100
result_ridge_100 <- RidgeOpt(lambda)

optimal_theta_100 <- result_ridge_100$theta
optimal_std_100 <- result_ridge_100$std
optimal_lambda_100 <- result_ridge_100$lambda

ridge_pred_train_100 <- as.matrix(trainS_x) %*% as.matrix(optimal_theta_100)
ridge_pred_test_100 <- as.matrix(testS_x) %*% as.matrix(optimal_theta_100)

mse_ridge_train_100 <- mean((trainS$motor_UPDRS - ridge_pred_train_100)^2)
mse_ridge_test_100 <- mean((testS$motor_UPDRS - ridge_pred_test_100)^2)

cat("Lambda:", lambda, "\n")
```

```
## Lambda: 100
```

```r
cat("Mean Squared Error (MSE) ridge regression on training data:", mse_ridge_train_100, "\n")
```

```
## Mean Squared Error (MSE) ridge regression on training data: 0.09487141
```

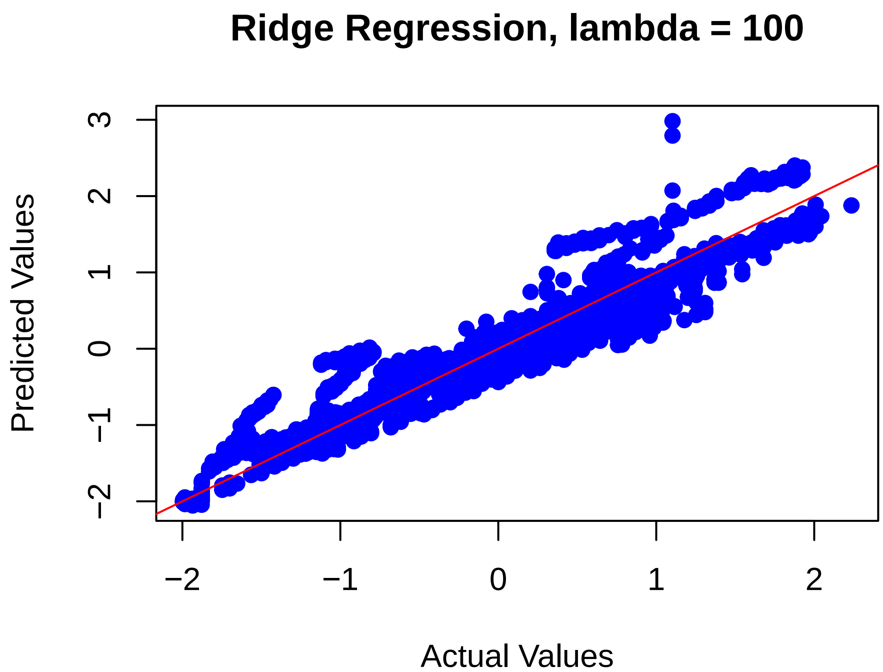
```r
cat("Mean Squared Error (MSE) ridge regression on test data:", mse_ridge_test_100, "\n")
```

```
## Mean Squared Error (MSE) ridge regression on test data: 0.09361211
```

```
cat("Degree of freedom:", DF(lambda), "\n")
```

```
## Degree of freedom: 14.61243
```

```
plot(testS$motor_UPDRS, ridge_pred_test_100, main = "Ridge Regression, lambda = 100",
     xlab = "Actual Values", ylab = "Predicted Values", pch = 19, col = "blue")
abline(a = 0, b = 1, col = "red")
```




```
lambda <- 1000
result_ridge_1000 <- RidgeOpt(lambda)

optimal_theta_1000 <- result_ridge_1000$theta
optimal_std_1000 <- result_ridge_1000$std
optimal_lambda_1000 <- result_ridge_1000$lambda

ridge_pred_train_1000 <- as.matrix(trainS_x) %*% as.matrix(optimal_theta_1000)
ridge_pred_test_1000 <- as.matrix(testS_x) %*% as.matrix(optimal_theta_1000)

mse_ridge_train_1000 <- mean((trainS$motor_UPDRS - ridge_pred_train_1000)^2)
mse_ridge_test_1000 <- mean((testS$motor_UPDRS - ridge_pred_test_1000)^2)

cat("Lambda:", lambda, "\n")
```

```
## Lambda: 1000
```

```
cat("Mean Squared Error (MSE) ridge regression on training data:", mse_ridge_train_1000, "\n")
```

```
## Mean Squared Error (MSE) ridge regression on training data: 0.09621063
```

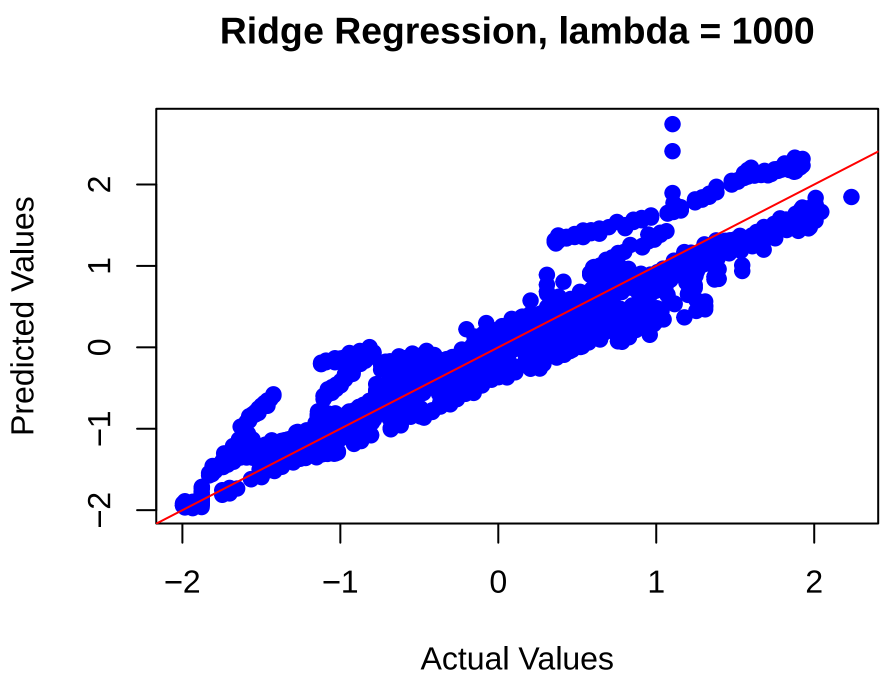
```r
cat("Mean Squared Error (MSE) ridge regression on test data:", mse_ridge_test_1000, "\n")
```

```
## Mean Squared Error (MSE) ridge regression on test data: 0.09423032
```

```r
cat("Degree of freedom:", DF(lambda), "\n")
```

```
## Degree of freedom: 9.222909
```

```r
plot(testS$motor_UPDRS, ridge_pred_test_1000, main = "Ridge Regression, lambda = 1000",
     xlab = "Actual Values", ylab = "Predicted Values", pch = 19, col = "blue")
abline(a = 0, b = 1, col = "red")
```



As lambda increases, the degrees of freedom decreases and the MSE increases

2 Assignment 3. Logistic regression and basis function expansion

2.1 Data

The data contains information about the onset of diabetes within 5 years in Pima Indians given medical details. The variables are:

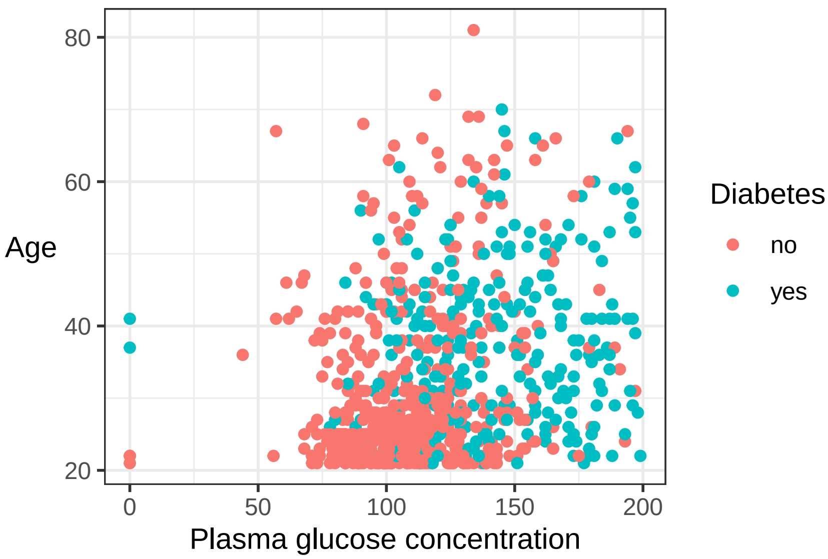
- Number of times pregnant
- Plasma glucose concentration a 2 hours in an oral glucose tolerance test.
- Diastolic blood pressure (mm Hg).
- Triceps skinfold thickness (mm).
- 2-Hour serum insulin (mu U/ml).
- Body mass index (weight in kg/(height in m)^2).
- Diabetes pedigree function.
- Age (years).
- Diabetes (0=no or 1=yes).

```r
diabetes_df <- read.csv("pima-indians-diabetes.csv", header=FALSE)

colnames(diabetes_df) <- c("times_pregnant", "plasma_glucose_conc",
                        "diastolic_blood_pressure", "triceps_skinfold_thickness",
                        "serum_insulin","body_mass_index", "diabetes_pedigree",
                        "age", "diabetes")

diabetes_df$diabetes <- ifelse(diabetes_df$diabetes == 0, "no", "yes")
diabetes_df$diabetes <- as.factor(diabetes_df$diabetes)
```

2.2 3.1

Question: Make a scatterplot showing a Plasma glucose concentration on Age where observations are colored by Diabetes levels.

```r
library(ggplot2)

ggplot(diabetes_df, aes(x = plasma_glucose_conc, y = age, color = diabetes)) +
  geom_point() +
  theme_bw() +
  theme(axis.title.y = element_text(angle = 0,vjust = 0.5)) +
  labs(colour = "Diabetes",
       x = "Plasma glucose concentration",
       y = "Age")
```



Question: Do you think that Diabetes is easy to classify by a standard logistic regression model that uses these two variables as features? Motivate your answer.

We don't think these two variables are good variables to classify Diabetes because there is no clear relationship between age, plasma glucose concentration with Diabetes.

2.3 3.2

Question:

Train a logistic regression model with $y =$ Diabetes as target $x_1 =$ Plasma glucose concentration and $x_2 =$ Age as features and make a prediction for all observations by using $r = 0.5$ as the classification threshold. Report the probabilistic equation of the estimated model and compute also the training misclassification error.

The probabilistic equation:

$$$$

```
model <- glm(diabetes ~ plasma_glucose_conc + age, data = diabetes_df,
             family =  "binomial")

pred <- predict(model, newdata = diabetes_df, type = "response")

# Using 0.5 as the classification threshold
pred <- ifelse(pred > 0.5, "yes", "no")

confusion <- table(diabetes_df$diabetes, pred)
misclass_rate <- (confusion[1,2] + confusion[2,1]) / sum(confusion)


knitr::kable(as.data.frame(round(misclass_rate,2)), col.names = "Misclassification error",
             caption = "Misclassification error")
```

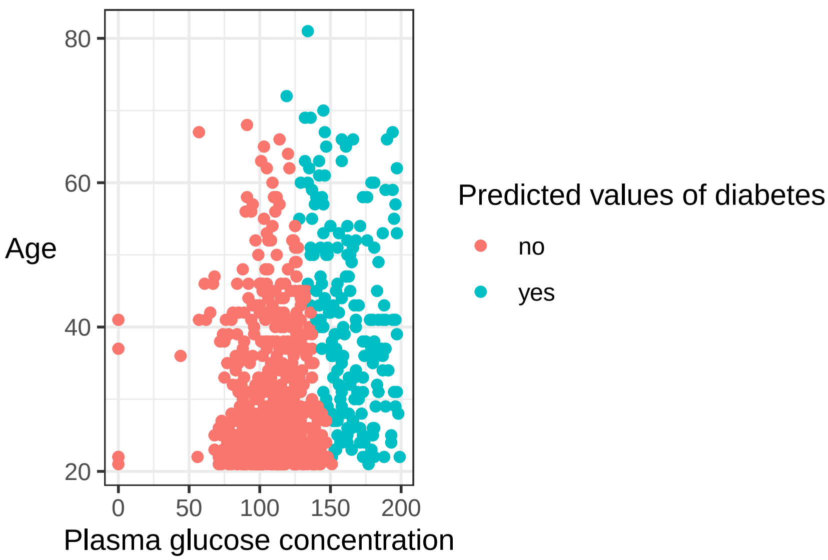
Table 4: Misclassification error

Misclassification error
0.26

Question:

Make a scatter plot of the same kind as in step 1 but showing the predicted values of Diabetes as a color instead.

```
diabetes_df_pred <- diabetes_df
diabetes_df_pred$pred <- pred

ggplot(diabetes_df_pred, aes(x = plasma_glucose_conc, y = age, color = pred)) +
  geom_point() +
  theme_bw() +
  theme(axis.title.y = element_text(angle = 0,vjust = 0.5)) +
  labs(colour = "Predicted values of diabetes",
       x = "Plasma glucose concentration",
       y = "Age")
```



Question:

Comment on the quality of the classification by using these results.

2.4 3.3

Question:

Use the model estimated in step 2 to a) report the equation of the decision boundary between the two classes b) add a curve showing this boundary to the scatter plot in step 2.

The decision boundary equation:

$$\beta_0 + \beta_1 \cdot \text{plasma glucos econc}$$

```
summary(model)
```

```
##
## Call:
## glm(formula = diabetes ~ plasma_glucose_conc + age, family = "binomial",
##     data = diabetes_df)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -2.3367  -0.7775  -0.5087   0.8367   3.1630
##
## Coefficients:
##                      Estimate Std. Error z value Pr(>|z|)
## (Intercept)         -5.912449   0.462620  -12.78  < 2e-16 ***
## plasma_glucose_conc  0.035644   0.003290   10.83  < 2e-16 ***
## age                  0.024778   0.007374    3.36 0.000778 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 993.48  on 767  degrees of freedom
## Residual deviance: 797.36  on 765  degrees of freedom
## AIC: 803.36
##
## Number of Fisher Scoring iterations: 4
```

```
slope <- coef(model)[2]/(-coef(model)[3])
intercept <- coef(model)[1]/(-coef(model)[3])


diabetes_df
```

```
##     times_pregnant plasma_glucose_conc diastolic_blood_pressure
## 1                6                 148                       72
## 2                1                  85                       66
## 3                8                 183                       64
## 4                1                  89                       66
## 5                0                 137                       40
## 6                5                 116                       74
## 7                3                  78                       50
## 8               10                 115                        0
## 9                2                 197                       70
## 10               8                 125                       96
## 11               4                 110                       92
## 12              10                 168                       74
## 13              10                 139                       80
```

```
## 14                1               189                 60
## 15                5               166                 72
## 16                7               100                  0
## 17                0               118                 84
## 18                7               107                 74
## 19                1               103                 30
## 20                1               115                 70
## 21                3               126                 88
## 22                8                99                 84
## 23                7               196                 90
## 24                9               119                 80
## 25               11               143                 94
## 26               10               125                 70
## 27                7               147                 76
## 28                1                97                 66
## 29               13               145                 82
## 30                5               117                 92
## 31                5               109                 75
## 32                3               158                 76
## 33                3                88                 58
## 34                6                92                 92
## 35               10               122                 78
## 36                4               103                 60
## 37               11               138                 76
## 38                9               102                 76
## 39                2                90                 68
## 40                4               111                 72
## 41                3               180                 64
## 42                7               133                 84
## 43                7               106                 92
## 44                9               171                110
## 45                7               159                 64
## 46                0               180                 66
## 47                1               146                 56
## 48                2                71                 70
## 49                7               103                 66
## 50                7               105                  0
## 51                1               103                 80
## 52                1               101                 50
## 53                5                88                 66
## 54                8               176                 90
## 55                7               150                 66
## 56                1                73                 50
## 57                7               187                 68
## 58                0               100                 88
## 59                0               146                 82
## 60                0               105                 64
## 61                2                84                  0
## 62                8               133                 72
```

```
## 63                5               44                   62
## 64                2              141                   58
## 65                7              114                   66
## 66                5               99                   74
## 67                0              109                   88
## 68                2              109                   92
## 69                1               95                   66
## 70                4              146                   85
## 71                2              100                   66
## 72                5              139                   64
## 73               13              126                   90
## 74                4              129                   86
## 75                1               79                   75
## 76                1                0                   48
## 77                7               62                   78
## 78                5               95                   72
## 79                0              131                    0
## 80                2              112                   66
## 81                3              113                   44
## 82                2               74                    0
## 83                7               83                   78
## 84                0              101                   65
## 85                5              137                  108
## 86                2              110                   74
## 87               13              106                   72
## 88                2              100                   68
## 89               15              136                   70
## 90                1              107                   68
## 91                1               80                   55
## 92                4              123                   80
## 93                7               81                   78
## 94                4              134                   72
## 95                2              142                   82
## 96                6              144                   72
## 97                2               92                   62
## 98                1               71                   48
## 99                6               93                   50
## 100               1              122                   90
## 101               1              163                   72
## 102               1              151                   60
## 103               0              125                   96
## 104               1               81                   72
## 105               2               85                   65
## 106               1              126                   56
## 107               1               96                  122
## 108               4              144                   58
## 109               3               83                   58
## 110               0               95                   85
## 111               3              171                   72
```

```
## 112         8      155        62
## 113         1       89        76
## 114         4       76        62
## 115         7      160        54
## 116         4      146        92
## 117         5      124        74
## 118         5       78        48
## 119         4       97        60
## 120         4       99        76
## 121         0      162        76
## 122         6      111        64
## 123         2      107        74
## 124         5      132        80
## 125         0      113        76
## 126         1       88        30
## 127         3      120        70
## 128         1      118        58
## 129         1      117        88
## 130         0      105        84
## 131         4      173        70
## 132         9      122        56
## 133         3      170        64
## 134         8       84        74
## 135         2       96        68
## 136         2      125        60
## 137         0      100        70
## 138         0       93        60
## 139         0      129        80
## 140         5      105        72
## 141         3      128        78
## 142         5      106        82
## 143         2      108        52
## 144        10      108        66
## 145         4      154        62
## 146         0      102        75
## 147         9       57        80
## 148         2      106        64
## 149         5      147        78
## 150         2       90        70
## 151         1      136        74
## 152         4      114        65
## 153         9      156        86
## 154         1      153        82
## 155         8      188        78
## 156         7      152        88
## 157         2       99        52
## 158         1      109        56
## 159         2       88        74
## 160        17      163        72
```

```
## 161                    4                   151                        90
## 162                    7                   102                        74
## 163                    0                   114                        80
## 164                    2                   100                        64
## 165                    0                   131                        88
## 166                    6                   104                        74
## 167                    3                   148                        66
## 168                    4                   120                        68
## 169                    4                   110                        66
## 170                    3                   111                        90
## 171                    6                   102                        82
## 172                    6                   134                        70
## 173                    2                    87                         0
## 174                    1                    79                        60
## 175                    2                    75                        64
## 176                    8                   179                        72
## 177                    6                    85                        78
## 178                    0                   129                       110
## 179                    5                   143                        78
## 180                    5                   130                        82
## 181                    6                    87                        80
## 182                    0                   119                        64
## 183                    1                     0                        74
## 184                    5                    73                        60
## 185                    4                   141                        74
## 186                    7                   194                        68
## 187                    8                   181                        68
## 188                    1                   128                        98
## 189                    8                   109                        76
## 190                    5                   139                        80
## 191                    3                   111                        62
## 192                    9                   123                        70
## 193                    7                   159                        66
## 194                   11                   135                         0
## 195                    8                    85                        55
## 196                    5                   158                        84
## 197                    1                   105                        58
## 198                    3                   107                        62
## 199                    4                   109                        64
## 200                    4                   148                        60
## 201                    0                   113                        80
## 202                    1                   138                        82
## 203                    0                   108                        68
## 204                    2                    99                        70
## 205                    6                   103                        72
## 206                    5                   111                        72
## 207                    8                   196                        76
## 208                    5                   162                       104
## 209                    1                    96                        64
```

```
## 210          7          184          84
## 211          2           81          60
## 212          0          147          85
## 213          7          179          95
## 214          0          140          65
## 215          9          112          82
## 216         12          151          70
## 217          5          109          62
## 218          6          125          68
## 219          5           85          74
## 220          5          112          66
## 221          0          177          60
## 222          2          158          90
## 223          7          119           0
## 224          7          142          60
## 225          1          100          66
## 226          1           87          78
## 227          0          101          76
## 228          3          162          52
## 229          4          197          70
## 230          0          117          80
## 231          4          142          86
## 232          6          134          80
## 233          1           79          80
## 234          4          122          68
## 235          3           74          68
## 236          4          171          72
## 237          7          181          84
## 238          0          179          90
## 239          9          164          84
## 240          0          104          76
## 241          1           91          64
## 242          4           91          70
## 243          3          139          54
## 244          6          119          50
## 245          2          146          76
## 246          9          184          85
## 247         10          122          68
## 248          0          165          90
## 249          9          124          70
## 250          1          111          86
## 251          9          106          52
## 252          2          129          84
## 253          2           90          80
## 254          0           86          68
## 255         12           92          62
## 256          1          113          64
## 257          3          111          56
## 258          2          114          68
```

```
## 259                    1                  193                            50
## 260                   11                  155                            76
## 261                    3                  191                            68
## 262                    3                  141                             0
## 263                    4                   95                            70
## 264                    3                  142                            80
## 265                    4                  123                            62
## 266                    5                   96                            74
## 267                    0                  138                             0
## 268                    2                  128                            64
## 269                    0                  102                            52
## 270                    2                  146                             0
## 271                   10                  101                            86
## 272                    2                  108                            62
## 273                    3                  122                            78
## 274                    1                   71                            78
## 275                   13                  106                            70
## 276                    2                  100                            70
## 277                    7                  106                            60
## 278                    0                  104                            64
## 279                    5                  114                            74
## 280                    2                  108                            62
## 281                    0                  146                            70
## 282                   10                  129                            76
## 283                    7                  133                            88
## 284                    7                  161                            86
## 285                    2                  108                            80
## 286                    7                  136                            74
## 287                    5                  155                            84
## 288                    1                  119                            86
## 289                    4                   96                            56
## 290                    5                  108                            72
## 291                    0                   78                            88
## 292                    0                  107                            62
## 293                    2                  128                            78
## 294                    1                  128                            48
## 295                    0                  161                            50
## 296                    6                  151                            62
## 297                    2                  146                            70
## 298                    0                  126                            84
## 299                   14                  100                            78
## 300                    8                  112                            72
## 301                    0                  167                             0
## 302                    2                  144                            58
## 303                    5                   77                            82
## 304                    5                  115                            98
## 305                    3                  150                            76
## 306                    2                  120                            76
## 307                   10                  161                            68
```

```
## 308          0          137          68
## 309          0          128          68
## 310          2          124          68
## 311          6           80          66
## 312          0          106          70
## 313          2          155          74
## 314          3          113          50
## 315          7          109          80
## 316          2          112          68
## 317          3           99          80
## 318          3          182          74
## 319          3          115          66
## 320          6          194          78
## 321          4          129          60
## 322          3          112          74
## 323          0          124          70
## 324         13          152          90
## 325          2          112          75
## 326          1          157          72
## 327          1          122          64
## 328         10          179          70
## 329          2          102          86
## 330          6          105          70
## 331          8          118          72
## 332          2           87          58
## 333          1          180           0
## 334         12          106          80
## 335          1           95          60
## 336          0          165          76
## 337          0          117           0
## 338          5          115          76
## 339          9          152          78
## 340          7          178          84
## 341          1          130          70
## 342          1           95          74
## 343          1            0          68
## 344          5          122          86
## 345          8           95          72
## 346          8          126          88
## 347          1          139          46
## 348          3          116           0
## 349          3           99          62
## 350          5            0          80
## 351          4           92          80
## 352          4          137          84
## 353          3           61          82
## 354          1           90          62
## 355          3           90          78
## 356          9          165          88
```

```
## 357                 1                  125                           50
## 358                13                  129                            0
## 359                12                   88                           74
## 360                 1                  196                           76
## 361                 5                  189                           64
## 362                 5                  158                           70
## 363                 5                  103                          108
## 364                 4                  146                           78
## 365                 4                  147                           74
## 366                 5                   99                           54
## 367                 6                  124                           72
## 368                 0                  101                           64
## 369                 3                   81                           86
## 370                 1                  133                          102
## 371                 3                  173                           82
## 372                 0                  118                           64
## 373                 0                   84                           64
## 374                 2                  105                           58
## 375                 2                  122                           52
## 376                12                  140                           82
## 377                 0                   98                           82
## 378                 1                   87                           60
## 379                 4                  156                           75
## 380                 0                   93                          100
## 381                 1                  107                           72
## 382                 0                  105                           68
## 383                 1                  109                           60
## 384                 1                   90                           62
## 385                 1                  125                           70
## 386                 1                  119                           54
## 387                 5                  116                           74
## 388                 8                  105                          100
## 389                 5                  144                           82
## 390                 3                  100                           68
## 391                 1                  100                           66
## 392                 5                  166                           76
## 393                 1                  131                           64
## 394                 4                  116                           72
## 395                 4                  158                           78
## 396                 2                  127                           58
## 397                 3                   96                           56
## 398                 0                  131                           66
## 399                 3                   82                           70
## 400                 3                  193                           70
## 401                 4                   95                           64
## 402                 6                  137                           61
## 403                 5                  136                           84
## 404                 9                   72                           78
## 405                 5                  168                           64
```

```
## 406                2               123                 48
## 407                4               115                 72
## 408                0               101                 62
## 409                8               197                 74
## 410                1               172                 68
## 411                6               102                 90
## 412                1               112                 72
## 413                1               143                 84
## 414                1               143                 74
## 415                0               138                 60
## 416                3               173                 84
## 417                1                97                 68
## 418                4               144                 82
## 419                1                83                 68
## 420                3               129                 64
## 421                1               119                 88
## 422                2                94                 68
## 423                0               102                 64
## 424                2               115                 64
## 425                8               151                 78
## 426                4               184                 78
## 427                0                94                  0
## 428                1               181                 64
## 429                0               135                 94
## 430                1                95                 82
## 431                2                99                  0
## 432                3                89                 74
## 433                1                80                 74
## 434                2               139                 75
## 435                1                90                 68
## 436                0               141                  0
## 437               12               140                 85
## 438                5               147                 75
## 439                1                97                 70
## 440                6               107                 88
## 441                0               189                104
## 442                2                83                 66
## 443                4               117                 64
## 444                8               108                 70
## 445                4               117                 62
## 446                0               180                 78
## 447                1               100                 72
## 448                0                95                 80
## 449                0               104                 64
## 450                0               120                 74
## 451                1                82                 64
## 452                2               134                 70
## 453                0                91                 68
## 454                2               119                  0
```

```
## 455           2          100           54
## 456          14          175           62
## 457           1          135           54
## 458           5           86           68
## 459          10          148           84
## 460           9          134           74
## 461           9          120           72
## 462           1           71           62
## 463           8           74           70
## 464           5           88           78
## 465          10          115           98
## 466           0          124           56
## 467           0           74           52
## 468           0           97           64
## 469           8          120            0
## 470           6          154           78
## 471           1          144           82
## 472           0          137           70
## 473           0          119           66
## 474           7          136           90
## 475           4          114           64
## 476           0          137           84
## 477           2          105           80
## 478           7          114           76
## 479           8          126           74
## 480           4          132           86
## 481           3          158           70
## 482           0          123           88
## 483           4           85           58
## 484           0           84           82
## 485           0          145            0
## 486           0          135           68
## 487           1          139           62
## 488           0          173           78
## 489           4           99           72
## 490           8          194           80
## 491           2           83           65
## 492           2           89           90
## 493           4           99           68
## 494           4          125           70
## 495           3           80            0
## 496           6          166           74
## 497           5          110           68
## 498           2           81           72
## 499           7          195           70
## 500           6          154           74
## 501           2          117           90
## 502           3           84           72
## 503           6            0           68
```

```
## 504                 7                  94                        64
## 505                 3                  96                        78
## 506                10                  75                        82
## 507                 0                 180                        90
## 508                 1                 130                        60
## 509                 2                  84                        50
## 510                 8                 120                        78
## 511                12                  84                        72
## 512                 0                 139                        62
## 513                 9                  91                        68
## 514                 2                  91                        62
## 515                 3                  99                        54
## 516                 3                 163                        70
## 517                 9                 145                        88
## 518                 7                 125                        86
## 519                13                  76                        60
## 520                 6                 129                        90
## 521                 2                  68                        70
## 522                 3                 124                        80
## 523                 6                 114                         0
## 524                 9                 130                        70
## 525                 3                 125                        58
## 526                 3                  87                        60
## 527                 1                  97                        64
## 528                 3                 116                        74
## 529                 0                 117                        66
## 530                 0                 111                        65
## 531                 2                 122                        60
## 532                 0                 107                        76
## 533                 1                  86                        66
## 534                 6                  91                         0
## 535                 1                  77                        56
## 536                 4                 132                         0
## 537                 0                 105                        90
## 538                 0                  57                        60
## 539                 0                 127                        80
## 540                 3                 129                        92
## 541                 8                 100                        74
## 542                 3                 128                        72
## 543                10                  90                        85
## 544                 4                  84                        90
## 545                 1                  88                        78
## 546                 8                 186                        90
## 547                 5                 187                        76
## 548                 4                 131                        68
## 549                 1                 164                        82
## 550                 4                 189                       110
## 551                 1                 116                        70
## 552                 3                  84                        68
```

```
## 553                6             114                 88
## 554                1              88                 62
## 555                1              84                 64
## 556                7             124                 70
## 557                1              97                 70
## 558                8             110                 76
## 559               11             103                 68
## 560               11              85                 74
## 561                6             125                 76
## 562                0             198                 66
## 563                1              87                 68
## 564                6              99                 60
## 565                0              91                 80
## 566                2              95                 54
## 567                1              99                 72
## 568                6              92                 62
## 569                4             154                 72
## 570                0             121                 66
## 571                3              78                 70
## 572                2             130                 96
## 573                3             111                 58
## 574                2              98                 60
## 575                1             143                 86
## 576                1             119                 44
## 577                6             108                 44
## 578                2             118                 80
## 579               10             133                 68
## 580                2             197                 70
## 581                0             151                 90
## 582                6             109                 60
## 583               12             121                 78
## 584                8             100                 76
## 585                8             124                 76
## 586                1              93                 56
## 587                8             143                 66
## 588                6             103                 66
## 589                3             176                 86
## 590                0              73                  0
## 591               11             111                 84
## 592                2             112                 78
## 593                3             132                 80
## 594                2              82                 52
## 595                6             123                 72
## 596                0             188                 82
## 597                0              67                 76
## 598                1              89                 24
## 599                1             173                 74
## 600                1             109                 38
## 601                1             108                 88
```

```
## 602       6   96    0
## 603       1  124   74
## 604       7  150   78
## 605       4  183    0
## 606       1  124   60
## 607       1  181   78
## 608       1   92   62
## 609       0  152   82
## 610       1  111   62
## 611       3  106   54
## 612       3  174   58
## 613       7  168   88
## 614       6  105   80
## 615      11  138   74
## 616       3  106   72
## 617       6  117   96
## 618       2   68   62
## 619       9  112   82
## 620       0  119    0
## 621       2  112   86
## 622       2   92   76
## 623       6  183   94
## 624       0   94   70
## 625       2  108   64
## 626       4   90   88
## 627       0  125   68
## 628       0  132   78
## 629       5  128   80
## 630       4   94   65
## 631       7  114   64
## 632       0  102   78
## 633       2  111   60
## 634       1  128   82
## 635      10   92   62
## 636      13  104   72
## 637       5  104   74
## 638       2   94   76
## 639       7   97   76
## 640       1  100   74
## 641       0  102   86
## 642       4  128   70
## 643       6  147   80
## 644       4   90    0
## 645       3  103   72
## 646       2  157   74
## 647       1  167   74
## 648       0  179   50
## 649      11  136   84
## 650       0  107   60
```

```
## 651                1                 91                         54
## 652                1                117                         60
## 653                5                123                         74
## 654                2                120                         54
## 655                1                106                         70
## 656                2                155                         52
## 657                2                101                         58
## 658                1                120                         80
## 659               11                127                        106
## 660                3                 80                         82
## 661               10                162                         84
## 662                1                199                         76
## 663                8                167                        106
## 664                9                145                         80
## 665                6                115                         60
## 666                1                112                         80
## 667                4                145                         82
## 668               10                111                         70
## 669                6                 98                         58
## 670                9                154                         78
## 671                6                165                         68
## 672                1                 99                         58
## 673               10                 68                        106
## 674                3                123                        100
## 675                8                 91                         82
## 676                6                195                         70
## 677                9                156                         86
## 678                0                 93                         60
## 679                3                121                         52
## 680                2                101                         58
## 681                2                 56                         56
## 682                0                162                         76
## 683                0                 95                         64
## 684                4                125                         80
## 685                5                136                         82
## 686                2                129                         74
## 687                3                130                         64
## 688                1                107                         50
## 689                1                140                         74
## 690                1                144                         82
## 691                8                107                         80
## 692               13                158                        114
## 693                2                121                         70
## 694                7                129                         68
## 695                2                 90                         60
## 696                7                142                         90
## 697                3                169                         74
## 698                0                 99                          0
## 699                4                127                         88
```

```
## 700       4       118       70
## 701       2       122       76
## 702       6       125       78
## 703       1       168       88
## 704       2       129        0
## 705       4       110       76
## 706       6        80       80
## 707      10       115        0
## 708       2       127       46
## 709       9       164       78
## 710       2        93       64
## 711       3       158       64
## 712       5       126       78
## 713      10       129       62
## 714       0       134       58
## 715       3       102       74
## 716       7       187       50
## 717       3       173       78
## 718      10        94       72
## 719       1       108       60
## 720       5        97       76
## 721       4        83       86
## 722       1       114       66
## 723       1       149       68
## 724       5       117       86
## 725       1       111       94
## 726       4       112       78
## 727       1       116       78
## 728       0       141       84
## 729       2       175       88
## 730       2        92       52
## 731       3       130       78
## 732       8       120       86
## 733       2       174       88
## 734       2       106       56
## 735       2       105       75
## 736       4        95       60
## 737       0       126       86
## 738       8        65       72
## 739       2        99       60
## 740       1       102       74
## 741      11       120       80
## 742       3       102       44
## 743       1       109       58
## 744       9       140       94
## 745      13       153       88
## 746      12       100       84
## 747       1       147       94
## 748       1        81       74
```

```
## 749              3            187                       70
## 750              6            162                       62
## 751              4            136                       70
## 752              1            121                       78
## 753              3            108                       62
## 754              0            181                       88
## 755              8            154                       78
## 756              1            128                       88
## 757              7            137                       90
## 758              0            123                       72
## 759              1            106                       76
## 760              6            190                       92
## 761              2             88                       58
## 762              9            170                       74
## 763              9             89                       62
## 764             10            101                       76
## 765              2            122                       70
## 766              5            121                       72
## 767              1            126                       60
## 768              1             93                       70
##     triceps_skinfold_thickness serum_insulin body_mass_index diabetes_pedigree
## 1                           35             0            33.6             0.627
## 2                           29             0            26.6             0.351
## 3                            0             0            23.3             0.672
## 4                           23            94            28.1             0.167
## 5                           35           168            43.1             2.288
## 6                            0             0            25.6             0.201
## 7                           32            88            31.0             0.248
## 8                            0             0            35.3             0.134
## 9                           45           543            30.5             0.158
## 10                           0             0             0.0             0.232
## 11                           0             0            37.6             0.191
## 12                           0             0            38.0             0.537
## 13                           0             0            27.1             1.441
## 14                          23           846            30.1             0.398
## 15                          19           175            25.8             0.587
## 16                           0             0            30.0             0.484
## 17                          47           230            45.8             0.551
## 18                           0             0            29.6             0.254
## 19                          38            83            43.3             0.183
## 20                          30            96            34.6             0.529
## 21                          41           235            39.3             0.704
## 22                           0             0            35.4             0.388
## 23                           0             0            39.8             0.451
## 24                          35             0            29.0             0.263
## 25                          33           146            36.6             0.254
## 26                          26           115            31.1             0.205
## 27                           0             0            39.4             0.257
## 28                          15           140            23.2             0.487
```

```
## 29                          19          110            22.2            0.245
## 30                           0            0            34.1            0.337
## 31                          26            0            36.0            0.546
## 32                          36          245            31.6            0.851
## 33                          11           54            24.8            0.267
## 34                           0            0            19.9            0.188
## 35                          31            0            27.6            0.512
## 36                          33          192            24.0            0.966
## 37                           0            0            33.2            0.420
## 38                          37            0            32.9            0.665
## 39                          42            0            38.2            0.503
## 40                          47          207            37.1            1.390
## 41                          25           70            34.0            0.271
## 42                           0            0            40.2            0.696
## 43                          18            0            22.7            0.235
## 44                          24          240            45.4            0.721
## 45                           0            0            27.4            0.294
## 46                          39            0            42.0            1.893
## 47                           0            0            29.7            0.564
## 48                          27            0            28.0            0.586
## 49                          32            0            39.1            0.344
## 50                           0            0             0.0            0.305
## 51                          11           82            19.4            0.491
## 52                          15           36            24.2            0.526
## 53                          21           23            24.4            0.342
## 54                          34          300            33.7            0.467
## 55                          42          342            34.7            0.718
## 56                          10            0            23.0            0.248
## 57                          39          304            37.7            0.254
## 58                          60          110            46.8            0.962
## 59                           0            0            40.5            1.781
## 60                          41          142            41.5            0.173
## 61                           0            0             0.0            0.304
## 62                           0            0            32.9            0.270
## 63                           0            0            25.0            0.587
## 64                          34          128            25.4            0.699
## 65                           0            0            32.8            0.258
## 66                          27            0            29.0            0.203
## 67                          30            0            32.5            0.855
## 68                           0            0            42.7            0.845
## 69                          13           38            19.6            0.334
## 70                          27          100            28.9            0.189
## 71                          20           90            32.9            0.867
## 72                          35          140            28.6            0.411
## 73                           0            0            43.4            0.583
## 74                          20          270            35.1            0.231
## 75                          30            0            32.0            0.396
## 76                          20            0            24.7            0.140
## 77                           0            0            32.6            0.391
```

```
## 78                          33           0          37.7          0.370
## 79                           0           0          43.2          0.270
## 80                          22           0          25.0          0.307
## 81                          13           0          22.4          0.140
## 82                           0           0           0.0          0.102
## 83                          26          71          29.3          0.767
## 84                          28           0          24.6          0.237
## 85                           0           0          48.8          0.227
## 86                          29         125          32.4          0.698
## 87                          54           0          36.6          0.178
## 88                          25          71          38.5          0.324
## 89                          32         110          37.1          0.153
## 90                          19           0          26.5          0.165
## 91                           0           0          19.1          0.258
## 92                          15         176          32.0          0.443
## 93                          40          48          46.7          0.261
## 94                           0           0          23.8          0.277
## 95                          18          64          24.7          0.761
## 96                          27         228          33.9          0.255
## 97                          28           0          31.6          0.130
## 98                          18          76          20.4          0.323
## 99                          30          64          28.7          0.356
## 100                         51         220          49.7          0.325
## 101                          0           0          39.0          1.222
## 102                          0           0          26.1          0.179
## 103                          0           0          22.5          0.262
## 104                         18          40          26.6          0.283
## 105                          0           0          39.6          0.930
## 106                         29         152          28.7          0.801
## 107                          0           0          22.4          0.207
## 108                         28         140          29.5          0.287
## 109                         31          18          34.3          0.336
## 110                         25          36          37.4          0.247
## 111                         33         135          33.3          0.199
## 112                         26         495          34.0          0.543
## 113                         34          37          31.2          0.192
## 114                          0           0          34.0          0.391
## 115                         32         175          30.5          0.588
## 116                          0           0          31.2          0.539
## 117                          0           0          34.0          0.220
## 118                          0           0          33.7          0.654
## 119                         23           0          28.2          0.443
## 120                         15          51          23.2          0.223
## 121                         56         100          53.2          0.759
## 122                         39           0          34.2          0.260
## 123                         30         100          33.6          0.404
## 124                          0           0          26.8          0.186
## 125                          0           0          33.3          0.278
## 126                         42          99          55.0          0.496
```

```
## 127                    30     135    42.9    0.452
## 128                    36      94    33.3    0.261
## 129                    24     145    34.5    0.403
## 130                     0       0    27.9    0.741
## 131                    14     168    29.7    0.361
## 132                     0       0    33.3    1.114
## 133                    37     225    34.5    0.356
## 134                    31       0    38.3    0.457
## 135                    13      49    21.1    0.647
## 136                    20     140    33.8    0.088
## 137                    26      50    30.8    0.597
## 138                    25      92    28.7    0.532
## 139                     0       0    31.2    0.703
## 140                    29     325    36.9    0.159
## 141                     0       0    21.1    0.268
## 142                    30       0    39.5    0.286
## 143                    26      63    32.5    0.318
## 144                     0       0    32.4    0.272
## 145                    31     284    32.8    0.237
## 146                    23       0     0.0    0.572
## 147                    37       0    32.8    0.096
## 148                    35     119    30.5    1.400
## 149                     0       0    33.7    0.218
## 150                    17       0    27.3    0.085
## 151                    50     204    37.4    0.399
## 152                     0       0    21.9    0.432
## 153                    28     155    34.3    1.189
## 154                    42     485    40.6    0.687
## 155                     0       0    47.9    0.137
## 156                    44       0    50.0    0.337
## 157                    15      94    24.6    0.637
## 158                    21     135    25.2    0.833
## 159                    19      53    29.0    0.229
## 160                    41     114    40.9    0.817
## 161                    38       0    29.7    0.294
## 162                    40     105    37.2    0.204
## 163                    34     285    44.2    0.167
## 164                    23       0    29.7    0.368
## 165                     0       0    31.6    0.743
## 166                    18     156    29.9    0.722
## 167                    25       0    32.5    0.256
## 168                     0       0    29.6    0.709
## 169                     0       0    31.9    0.471
## 170                    12      78    28.4    0.495
## 171                     0       0    30.8    0.180
## 172                    23     130    35.4    0.542
## 173                    23       0    28.9    0.773
## 174                    42      48    43.5    0.678
## 175                    24      55    29.7    0.370
```

```
## 176                        42        130        32.7        0.719
## 177                         0          0        31.2        0.382
## 178                        46        130        67.1        0.319
## 179                         0          0        45.0        0.190
## 180                         0          0        39.1        0.956
## 181                         0          0        23.2        0.084
## 182                        18         92        34.9        0.725
## 183                        20         23        27.7        0.299
## 184                         0          0        26.8        0.268
## 185                         0          0        27.6        0.244
## 186                        28          0        35.9        0.745
## 187                        36        495        30.1        0.615
## 188                        41         58        32.0        1.321
## 189                        39        114        27.9        0.640
## 190                        35        160        31.6        0.361
## 191                         0          0        22.6        0.142
## 192                        44         94        33.1        0.374
## 193                         0          0        30.4        0.383
## 194                         0          0        52.3        0.578
## 195                        20          0        24.4        0.136
## 196                        41        210        39.4        0.395
## 197                         0          0        24.3        0.187
## 198                        13         48        22.9        0.678
## 199                        44         99        34.8        0.905
## 200                        27        318        30.9        0.150
## 201                        16          0        31.0        0.874
## 202                         0          0        40.1        0.236
## 203                        20          0        27.3        0.787
## 204                        16         44        20.4        0.235
## 205                        32        190        37.7        0.324
## 206                        28          0        23.9        0.407
## 207                        29        280        37.5        0.605
## 208                         0          0        37.7        0.151
## 209                        27         87        33.2        0.289
## 210                        33          0        35.5        0.355
## 211                        22          0        27.7        0.290
## 212                        54          0        42.8        0.375
## 213                        31          0        34.2        0.164
## 214                        26        130        42.6        0.431
## 215                        32        175        34.2        0.260
## 216                        40        271        41.8        0.742
## 217                        41        129        35.8        0.514
## 218                        30        120        30.0        0.464
## 219                        22          0        29.0        1.224
## 220                         0          0        37.8        0.261
## 221                        29        478        34.6        1.072
## 222                         0          0        31.6        0.805
## 223                         0          0        25.2        0.209
## 224                        33        190        28.8        0.687
```

```
## 225                           15          56           23.6            0.666
## 226                           27          32           34.6            0.101
## 227                            0           0           35.7            0.198
## 228                           38           0           37.2            0.652
## 229                           39         744           36.7            2.329
## 230                           31          53           45.2            0.089
## 231                            0           0           44.0            0.645
## 232                           37         370           46.2            0.238
## 233                           25          37           25.4            0.583
## 234                            0           0           35.0            0.394
## 235                           28          45           29.7            0.293
## 236                            0           0           43.6            0.479
## 237                           21         192           35.9            0.586
## 238                           27           0           44.1            0.686
## 239                           21           0           30.8            0.831
## 240                            0           0           18.4            0.582
## 241                           24           0           29.2            0.192
## 242                           32          88           33.1            0.446
## 243                            0           0           25.6            0.402
## 244                           22         176           27.1            1.318
## 245                           35         194           38.2            0.329
## 246                           15           0           30.0            1.213
## 247                            0           0           31.2            0.258
## 248                           33         680           52.3            0.427
## 249                           33         402           35.4            0.282
## 250                           19           0           30.1            0.143
## 251                            0           0           31.2            0.380
## 252                            0           0           28.0            0.284
## 253                           14          55           24.4            0.249
## 254                           32           0           35.8            0.238
## 255                            7         258           27.6            0.926
## 256                           35           0           33.6            0.543
## 257                           39           0           30.1            0.557
## 258                           22           0           28.7            0.092
## 259                           16         375           25.9            0.655
## 260                           28         150           33.3            1.353
## 261                           15         130           30.9            0.299
## 262                            0           0           30.0            0.761
## 263                           32           0           32.1            0.612
## 264                           15           0           32.4            0.200
## 265                            0           0           32.0            0.226
## 266                           18          67           33.6            0.997
## 267                            0           0           36.3            0.933
## 268                           42           0           40.0            1.101
## 269                            0           0           25.1            0.078
## 270                            0           0           27.5            0.240
## 271                           37           0           45.6            1.136
## 272                           32          56           25.2            0.128
## 273                            0           0           23.0            0.254
```

```
## 274                          50          45          33.2          0.422
## 275                           0           0          34.2          0.251
## 276                          52          57          40.5          0.677
## 277                          24           0          26.5          0.296
## 278                          23         116          27.8          0.454
## 279                           0           0          24.9          0.744
## 280                          10         278          25.3          0.881
## 281                           0           0          37.9          0.334
## 282                          28         122          35.9          0.280
## 283                          15         155          32.4          0.262
## 284                           0           0          30.4          0.165
## 285                           0           0          27.0          0.259
## 286                          26         135          26.0          0.647
## 287                          44         545          38.7          0.619
## 288                          39         220          45.6          0.808
## 289                          17          49          20.8          0.340
## 290                          43          75          36.1          0.263
## 291                          29          40          36.9          0.434
## 292                          30          74          36.6          0.757
## 293                          37         182          43.3          1.224
## 294                          45         194          40.5          0.613
## 295                           0           0          21.9          0.254
## 296                          31         120          35.5          0.692
## 297                          38         360          28.0          0.337
## 298                          29         215          30.7          0.520
## 299                          25         184          36.6          0.412
## 300                           0           0          23.6          0.840
## 301                           0           0          32.3          0.839
## 302                          33         135          31.6          0.422
## 303                          41          42          35.8          0.156
## 304                           0           0          52.9          0.209
## 305                           0           0          21.0          0.207
## 306                          37         105          39.7          0.215
## 307                          23         132          25.5          0.326
## 308                          14         148          24.8          0.143
## 309                          19         180          30.5          1.391
## 310                          28         205          32.9          0.875
## 311                          30           0          26.2          0.313
## 312                          37         148          39.4          0.605
## 313                          17          96          26.6          0.433
## 314                          10          85          29.5          0.626
## 315                          31           0          35.9          1.127
## 316                          22          94          34.1          0.315
## 317                          11          64          19.3          0.284
## 318                           0           0          30.5          0.345
## 319                          39         140          38.1          0.150
## 320                           0           0          23.5          0.129
## 321                          12         231          27.5          0.527
## 322                          30           0          31.6          0.197
```

```
## 323                                       20            0             27.4             0.254
## 324                                       33           29             26.8             0.731
## 325                                       32            0             35.7             0.148
## 326                                       21          168             25.6             0.123
## 327                                       32          156             35.1             0.692
## 328                                        0            0             35.1             0.200
## 329                                       36          120             45.5             0.127
## 330                                       32           68             30.8             0.122
## 331                                       19            0             23.1             1.476
## 332                                       16           52             32.7             0.166
## 333                                        0            0             43.3             0.282
## 334                                        0            0             23.6             0.137
## 335                                       18           58             23.9             0.260
## 336                                       43          255             47.9             0.259
## 337                                        0            0             33.8             0.932
## 338                                        0            0             31.2             0.343
## 339                                       34          171             34.2             0.893
## 340                                        0            0             39.9             0.331
## 341                                       13          105             25.9             0.472
## 342                                       21           73             25.9             0.673
## 343                                       35            0             32.0             0.389
## 344                                        0            0             34.7             0.290
## 345                                        0            0             36.8             0.485
## 346                                       36          108             38.5             0.349
## 347                                       19           83             28.7             0.654
## 348                                        0            0             23.5             0.187
## 349                                       19           74             21.8             0.279
## 350                                       32            0             41.0             0.346
## 351                                        0            0             42.2             0.237
## 352                                        0            0             31.2             0.252
## 353                                       28            0             34.4             0.243
## 354                                       12           43             27.2             0.580
## 355                                        0            0             42.7             0.559
## 356                                        0            0             30.4             0.302
## 357                                       40          167             33.3             0.962
## 358                                       30            0             39.9             0.569
## 359                                       40           54             35.3             0.378
## 360                                       36          249             36.5             0.875
## 361                                       33          325             31.2             0.583
## 362                                        0            0             29.8             0.207
## 363                                       37            0             39.2             0.305
## 364                                        0            0             38.5             0.520
## 365                                       25          293             34.9             0.385
## 366                                       28           83             34.0             0.499
## 367                                        0            0             27.6             0.368
## 368                                       17            0             21.0             0.252
## 369                                       16           66             27.5             0.306
## 370                                       28          140             32.8             0.234
## 371                                       48          465             38.4             2.137
```

```
## 372                                   23          89          0.0          1.731
## 373                                   22          66         35.8          0.545
## 374                                   40          94         34.9          0.225
## 375                                   43         158         36.2          0.816
## 376                                   43         325         39.2          0.528
## 377                                   15          84         25.2          0.299
## 378                                   37          75         37.2          0.509
## 379                                    0           0         48.3          0.238
## 380                                   39          72         43.4          1.021
## 381                                   30          82         30.8          0.821
## 382                                   22           0         20.0          0.236
## 383                                    8         182         25.4          0.947
## 384                                   18          59         25.1          1.268
## 385                                   24         110         24.3          0.221
## 386                                   13          50         22.3          0.205
## 387                                   29           0         32.3          0.660
## 388                                   36           0         43.3          0.239
## 389                                   26         285         32.0          0.452
## 390                                   23          81         31.6          0.949
## 391                                   29         196         32.0          0.444
## 392                                    0           0         45.7          0.340
## 393                                   14         415         23.7          0.389
## 394                                   12          87         22.1          0.463
## 395                                    0           0         32.9          0.803
## 396                                   24         275         27.7          1.600
## 397                                   34         115         24.7          0.944
## 398                                   40           0         34.3          0.196
## 399                                    0           0         21.1          0.389
## 400                                   31           0         34.9          0.241
## 401                                    0           0         32.0          0.161
## 402                                    0           0         24.2          0.151
## 403                                   41          88         35.0          0.286
## 404                                   25           0         31.6          0.280
## 405                                    0           0         32.9          0.135
## 406                                   32         165         42.1          0.520
## 407                                    0           0         28.9          0.376
## 408                                    0           0         21.9          0.336
## 409                                    0           0         25.9          1.191
## 410                                   49         579         42.4          0.702
## 411                                   39           0         35.7          0.674
## 412                                   30         176         34.4          0.528
## 413                                   23         310         42.4          1.076
## 414                                   22          61         26.2          0.256
## 415                                   35         167         34.6          0.534
## 416                                   33         474         35.7          0.258
## 417                                   21           0         27.2          1.095
## 418                                   32           0         38.5          0.554
## 419                                    0           0         18.2          0.624
## 420                                   29         115         26.4          0.219
```

```
## 421                                      41          170          45.3          0.507
## 422                                      18           76          26.0          0.561
## 423                                      46           78          40.6          0.496
## 424                                      22            0          30.8          0.421
## 425                                      32          210          42.9          0.516
## 426                                      39          277          37.0          0.264
## 427                                       0            0           0.0          0.256
## 428                                      30          180          34.1          0.328
## 429                                      46          145          40.6          0.284
## 430                                      25          180          35.0          0.233
## 431                                       0            0          22.2          0.108
## 432                                      16           85          30.4          0.551
## 433                                      11           60          30.0          0.527
## 434                                       0            0          25.6          0.167
## 435                                       8            0          24.5          1.138
## 436                                       0            0          42.4          0.205
## 437                                      33            0          37.4          0.244
## 438                                       0            0          29.9          0.434
## 439                                      15            0          18.2          0.147
## 440                                       0            0          36.8          0.727
## 441                                      25            0          34.3          0.435
## 442                                      23           50          32.2          0.497
## 443                                      27          120          33.2          0.230
## 444                                       0            0          30.5          0.955
## 445                                      12            0          29.7          0.380
## 446                                      63           14          59.4          2.420
## 447                                      12           70          25.3          0.658
## 448                                      45           92          36.5          0.330
## 449                                      37           64          33.6          0.510
## 450                                      18           63          30.5          0.285
## 451                                      13           95          21.2          0.415
## 452                                       0            0          28.9          0.542
## 453                                      32          210          39.9          0.381
## 454                                       0            0          19.6          0.832
## 455                                      28          105          37.8          0.498
## 456                                      30            0          33.6          0.212
## 457                                       0            0          26.7          0.687
## 458                                      28           71          30.2          0.364
## 459                                      48          237          37.6          1.001
## 460                                      33           60          25.9          0.460
## 461                                      22           56          20.8          0.733
## 462                                       0            0          21.8          0.416
## 463                                      40           49          35.3          0.705
## 464                                      30            0          27.6          0.258
## 465                                       0            0          24.0          1.022
## 466                                      13          105          21.8          0.452
## 467                                      10           36          27.8          0.269
## 468                                      36          100          36.8          0.600
## 469                                       0            0          30.0          0.183
```

```
## 470                                41        140          46.1          0.571
## 471                                40          0          41.3          0.607
## 472                                38          0          33.2          0.170
## 473                                27          0          38.8          0.259
## 474                                 0          0          29.9          0.210
## 475                                 0          0          28.9          0.126
## 476                                27          0          27.3          0.231
## 477                                45        191          33.7          0.711
## 478                                17        110          23.8          0.466
## 479                                38         75          25.9          0.162
## 480                                31          0          28.0          0.419
## 481                                30        328          35.5          0.344
## 482                                37          0          35.2          0.197
## 483                                22         49          27.8          0.306
## 484                                31        125          38.2          0.233
## 485                                 0          0          44.2          0.630
## 486                                42        250          42.3          0.365
## 487                                41        480          40.7          0.536
## 488                                32        265          46.5          1.159
## 489                                17          0          25.6          0.294
## 490                                 0          0          26.1          0.551
## 491                                28         66          36.8          0.629
## 492                                30          0          33.5          0.292
## 493                                38          0          32.8          0.145
## 494                                18        122          28.9          1.144
## 495                                 0          0           0.0          0.174
## 496                                 0          0          26.6          0.304
## 497                                 0          0          26.0          0.292
## 498                                15         76          30.1          0.547
## 499                                33        145          25.1          0.163
## 500                                32        193          29.3          0.839
## 501                                19         71          25.2          0.313
## 502                                32          0          37.2          0.267
## 503                                41          0          39.0          0.727
## 504                                25         79          33.3          0.738
## 505                                39          0          37.3          0.238
## 506                                 0          0          33.3          0.263
## 507                                26         90          36.5          0.314
## 508                                23        170          28.6          0.692
## 509                                23         76          30.4          0.968
## 510                                 0          0          25.0          0.409
## 511                                31          0          29.7          0.297
## 512                                17        210          22.1          0.207
## 513                                 0          0          24.2          0.200
## 514                                 0          0          27.3          0.525
## 515                                19         86          25.6          0.154
## 516                                18        105          31.6          0.268
## 517                                34        165          30.3          0.771
## 518                                 0          0          37.6          0.304
```

```
## 519                                 0           0          32.8              0.180
## 520                                 7         326          19.6              0.582
## 521                                32          66          25.0              0.187
## 522                                33         130          33.2              0.305
## 523                                 0           0           0.0              0.189
## 524                                 0           0          34.2              0.652
## 525                                 0           0          31.6              0.151
## 526                                18           0          21.8              0.444
## 527                                19          82          18.2              0.299
## 528                                15         105          26.3              0.107
## 529                                31         188          30.8              0.493
## 530                                 0           0          24.6              0.660
## 531                                18         106          29.8              0.717
## 532                                 0           0          45.3              0.686
## 533                                52          65          41.3              0.917
## 534                                 0           0          29.8              0.501
## 535                                30          56          33.3              1.251
## 536                                 0           0          32.9              0.302
## 537                                 0           0          29.6              0.197
## 538                                 0           0          21.7              0.735
## 539                                37         210          36.3              0.804
## 540                                49         155          36.4              0.968
## 541                                40         215          39.4              0.661
## 542                                25         190          32.4              0.549
## 543                                32           0          34.9              0.825
## 544                                23          56          39.5              0.159
## 545                                29          76          32.0              0.365
## 546                                35         225          34.5              0.423
## 547                                27         207          43.6              1.034
## 548                                21         166          33.1              0.160
## 549                                43          67          32.8              0.341
## 550                                31           0          28.5              0.680
## 551                                28           0          27.4              0.204
## 552                                30         106          31.9              0.591
## 553                                 0           0          27.8              0.247
## 554                                24          44          29.9              0.422
## 555                                23         115          36.9              0.471
## 556                                33         215          25.5              0.161
## 557                                40           0          38.1              0.218
## 558                                 0           0          27.8              0.237
## 559                                40           0          46.2              0.126
## 560                                 0           0          30.1              0.300
## 561                                 0           0          33.8              0.121
## 562                                32         274          41.3              0.502
## 563                                34          77          37.6              0.401
## 564                                19          54          26.9              0.497
## 565                                 0           0          32.4              0.601
## 566                                14          88          26.1              0.748
## 567                                30          18          38.6              0.412
```

```
## 568                                    32          126          32.0            0.085
## 569                                    29          126          31.3            0.338
## 570                                    30          165          34.3            0.203
## 571                                     0            0          32.5            0.270
## 572                                     0            0          22.6            0.268
## 573                                    31           44          29.5            0.430
## 574                                    17          120          34.7            0.198
## 575                                    30          330          30.1            0.892
## 576                                    47           63          35.5            0.280
## 577                                    20          130          24.0            0.813
## 578                                     0            0          42.9            0.693
## 579                                     0            0          27.0            0.245
## 580                                    99            0          34.7            0.575
## 581                                    46            0          42.1            0.371
## 582                                    27            0          25.0            0.206
## 583                                    17            0          26.5            0.259
## 584                                     0            0          38.7            0.190
## 585                                    24          600          28.7            0.687
## 586                                    11            0          22.5            0.417
## 587                                     0            0          34.9            0.129
## 588                                     0            0          24.3            0.249
## 589                                    27          156          33.3            1.154
## 590                                     0            0          21.1            0.342
## 591                                    40            0          46.8            0.925
## 592                                    50          140          39.4            0.175
## 593                                     0            0          34.4            0.402
## 594                                    22          115          28.5            1.699
## 595                                    45          230          33.6            0.733
## 596                                    14          185          32.0            0.682
## 597                                     0            0          45.3            0.194
## 598                                    19           25          27.8            0.559
## 599                                     0            0          36.8            0.088
## 600                                    18          120          23.1            0.407
## 601                                    19            0          27.1            0.400
## 602                                     0            0          23.7            0.190
## 603                                    36            0          27.8            0.100
## 604                                    29          126          35.2            0.692
## 605                                     0            0          28.4            0.212
## 606                                    32            0          35.8            0.514
## 607                                    42          293          40.0            1.258
## 608                                    25           41          19.5            0.482
## 609                                    39          272          41.5            0.270
## 610                                    13          182          24.0            0.138
## 611                                    21          158          30.9            0.292
## 612                                    22          194          32.9            0.593
## 613                                    42          321          38.2            0.787
## 614                                    28            0          32.5            0.878
## 615                                    26          144          36.1            0.557
## 616                                     0            0          25.8            0.207
```

```
## 617                         0          0          28.7         0.157
## 618                        13         15          20.1         0.257
## 619                        24          0          28.2         1.282
## 620                         0          0          32.4         0.141
## 621                        42        160          38.4         0.246
## 622                        20          0          24.2         1.698
## 623                         0          0          40.8         1.461
## 624                        27        115          43.5         0.347
## 625                         0          0          30.8         0.158
## 626                        47         54          37.7         0.362
## 627                         0          0          24.7         0.206
## 628                         0          0          32.4         0.393
## 629                         0          0          34.6         0.144
## 630                        22          0          24.7         0.148
## 631                         0          0          27.4         0.732
## 632                        40         90          34.5         0.238
## 633                         0          0          26.2         0.343
## 634                        17        183          27.5         0.115
## 635                         0          0          25.9         0.167
## 636                         0          0          31.2         0.465
## 637                         0          0          28.8         0.153
## 638                        18         66          31.6         0.649
## 639                        32         91          40.9         0.871
## 640                        12         46          19.5         0.149
## 641                        17        105          29.3         0.695
## 642                         0          0          34.3         0.303
## 643                         0          0          29.5         0.178
## 644                         0          0          28.0         0.610
## 645                        30        152          27.6         0.730
## 646                        35        440          39.4         0.134
## 647                        17        144          23.4         0.447
## 648                        36        159          37.8         0.455
## 649                        35        130          28.3         0.260
## 650                        25          0          26.4         0.133
## 651                        25        100          25.2         0.234
## 652                        23        106          33.8         0.466
## 653                        40         77          34.1         0.269
## 654                         0          0          26.8         0.455
## 655                        28        135          34.2         0.142
## 656                        27        540          38.7         0.240
## 657                        35         90          21.8         0.155
## 658                        48        200          38.9         1.162
## 659                         0          0          39.0         0.190
## 660                        31         70          34.2         1.292
## 661                         0          0          27.7         0.182
## 662                        43          0          42.9         1.394
## 663                        46        231          37.6         0.165
## 664                        46        130          37.9         0.637
## 665                        39          0          33.7         0.245
```

## 666	45	132	34.8	0.217
## 667	18	0	32.5	0.235
## 668	27	0	27.5	0.141
## 669	33	190	34.0	0.430
## 670	30	100	30.9	0.164
## 671	26	168	33.6	0.631
## 672	10	0	25.4	0.551
## 673	23	49	35.5	0.285
## 674	35	240	57.3	0.880
## 675	0	0	35.6	0.587
## 676	0	0	30.9	0.328
## 677	0	0	24.8	0.230
## 678	0	0	35.3	0.263
## 679	0	0	36.0	0.127
## 680	17	265	24.2	0.614
## 681	28	45	24.2	0.332
## 682	36	0	49.6	0.364
## 683	39	105	44.6	0.366
## 684	0	0	32.3	0.536
## 685	0	0	0.0	0.640
## 686	26	205	33.2	0.591
## 687	0	0	23.1	0.314
## 688	19	0	28.3	0.181
## 689	26	180	24.1	0.828
## 690	46	180	46.1	0.335
## 691	0	0	24.6	0.856
## 692	0	0	42.3	0.257
## 693	32	95	39.1	0.886
## 694	49	125	38.5	0.439
## 695	0	0	23.5	0.191
## 696	24	480	30.4	0.128
## 697	19	125	29.9	0.268
## 698	0	0	25.0	0.253
## 699	11	155	34.5	0.598
## 700	0	0	44.5	0.904
## 701	27	200	35.9	0.483
## 702	31	0	27.6	0.565
## 703	29	0	35.0	0.905
## 704	0	0	38.5	0.304
## 705	20	100	28.4	0.118
## 706	36	0	39.8	0.177
## 707	0	0	0.0	0.261
## 708	21	335	34.4	0.176
## 709	0	0	32.8	0.148
## 710	32	160	38.0	0.674
## 711	13	387	31.2	0.295
## 712	27	22	29.6	0.439
## 713	36	0	41.2	0.441
## 714	20	291	26.4	0.352

```
## 715                                  0          0        29.5         0.121
## 716                                 33        392        33.9         0.826
## 717                                 39        185        33.8         0.970
## 718                                 18          0        23.1         0.595
## 719                                 46        178        35.5         0.415
## 720                                 27          0        35.6         0.378
## 721                                 19          0        29.3         0.317
## 722                                 36        200        38.1         0.289
## 723                                 29        127        29.3         0.349
## 724                                 30        105        39.1         0.251
## 725                                  0          0        32.8         0.265
## 726                                 40          0        39.4         0.236
## 727                                 29        180        36.1         0.496
## 728                                 26          0        32.4         0.433
## 729                                  0          0        22.9         0.326
## 730                                  0          0        30.1         0.141
## 731                                 23         79        28.4         0.323
## 732                                  0          0        28.4         0.259
## 733                                 37        120        44.5         0.646
## 734                                 27        165        29.0         0.426
## 735                                  0          0        23.3         0.560
## 736                                 32          0        35.4         0.284
## 737                                 27        120        27.4         0.515
## 738                                 23          0        32.0         0.600
## 739                                 17        160        36.6         0.453
## 740                                  0          0        39.5         0.293
## 741                                 37        150        42.3         0.785
## 742                                 20         94        30.8         0.400
## 743                                 18        116        28.5         0.219
## 744                                  0          0        32.7         0.734
## 745                                 37        140        40.6         1.174
## 746                                 33        105        30.0         0.488
## 747                                 41          0        49.3         0.358
## 748                                 41         57        46.3         1.096
## 749                                 22        200        36.4         0.408
## 750                                  0          0        24.3         0.178
## 751                                  0          0        31.2         1.182
## 752                                 39         74        39.0         0.261
## 753                                 24          0        26.0         0.223
## 754                                 44        510        43.3         0.222
## 755                                 32          0        32.4         0.443
## 756                                 39        110        36.5         1.057
## 757                                 41          0        32.0         0.391
## 758                                  0          0        36.3         0.258
## 759                                  0          0        37.5         0.197
## 760                                  0          0        35.5         0.278
## 761                                 26         16        28.4         0.766
## 762                                 31          0        44.0         0.403
## 763                                  0          0        22.5         0.142
```

```
## 764                          48          180          32.9             0.171
## 765                          27            0          36.8             0.340
## 766                          23          112          26.2             0.245
## 767                           0            0          30.1             0.349
## 768                          31            0          30.4             0.315
##     age diabetes
## 1    50      yes
## 2    31       no
## 3    32      yes
## 4    21       no
## 5    33      yes
## 6    30       no
## 7    26      yes
## 8    29       no
## 9    53      yes
## 10   54      yes
## 11   30       no
## 12   34      yes
## 13   57       no
## 14   59      yes
## 15   51      yes
## 16   32      yes
## 17   31      yes
## 18   31      yes
## 19   33       no
## 20   32      yes
## 21   27       no
## 22   50       no
## 23   41      yes
## 24   29      yes
## 25   51      yes
## 26   41      yes
## 27   43      yes
## 28   22       no
## 29   57       no
## 30   38       no
## 31   60       no
## 32   28      yes
## 33   22       no
## 34   28       no
## 35   45       no
## 36   33       no
## 37   35       no
## 38   46      yes
## 39   27      yes
## 40   56      yes
## 41   26       no
## 42   37       no
## 43   48       no
```

```
## 44   54       yes
## 45   40        no
## 46   25       yes
## 47   29        no
## 48   22        no
## 49   31       yes
## 50   24        no
## 51   22        no
## 52   26        no
## 53   30        no
## 54   58       yes
## 55   42        no
## 56   21        no
## 57   41       yes
## 58   31        no
## 59   44        no
## 60   22        no
## 61   21        no
## 62   39       yes
## 63   36        no
## 64   24        no
## 65   42       yes
## 66   32        no
## 67   38       yes
## 68   54        no
## 69   25        no
## 70   27        no
## 71   28       yes
## 72   26        no
## 73   42       yes
## 74   23        no
## 75   22        no
## 76   22        no
## 77   41        no
## 78   27        no
## 79   26       yes
## 80   24        no
## 81   22        no
## 82   22        no
## 83   36        no
## 84   22        no
## 85   37       yes
## 86   27        no
## 87   45        no
## 88   26        no
## 89   43       yes
## 90   24        no
## 91   21        no
## 92   34        no
```

```
## 93   42          no
## 94   60         yes
## 95   21          no
## 96   40          no
## 97   24          no
## 98   22          no
## 99   23          no
## 100  31         yes
## 101  33         yes
## 102  22          no
## 103  21          no
## 104  24          no
## 105  27          no
## 106  21          no
## 107  27          no
## 108  37          no
## 109  25          no
## 110  24         yes
## 111  24         yes
## 112  46         yes
## 113  23          no
## 114  25          no
## 115  39         yes
## 116  61         yes
## 117  38         yes
## 118  25          no
## 119  22          no
## 120  21          no
## 121  25         yes
## 122  24          no
## 123  23          no
## 124  69          no
## 125  23         yes
## 126  26         yes
## 127  30          no
## 128  23          no
## 129  40         yes
## 130  62         yes
## 131  33         yes
## 132  33         yes
## 133  30         yes
## 134  39          no
## 135  26          no
## 136  31          no
## 137  21          no
## 138  22          no
## 139  29          no
## 140  28          no
## 141  55          no
```

```
## 142  38         no
## 143  22         no
## 144  42        yes
## 145  23         no
## 146  21         no
## 147  41         no
## 148  34         no
## 149  65         no
## 150  22         no
## 151  24         no
## 152  37         no
## 153  42        yes
## 154  23         no
## 155  43        yes
## 156  36        yes
## 157  21         no
## 158  23         no
## 159  22         no
## 160  47        yes
## 161  36         no
## 162  45         no
## 163  27         no
## 164  21         no
## 165  32        yes
## 166  41        yes
## 167  22         no
## 168  34         no
## 169  29         no
## 170  29         no
## 171  36        yes
## 172  29        yes
## 173  25         no
## 174  23         no
## 175  33         no
## 176  36        yes
## 177  42         no
## 178  26        yes
## 179  47         no
## 180  37        yes
## 181  32         no
## 182  23         no
## 183  21         no
## 184  27         no
## 185  40         no
## 186  41        yes
## 187  60        yes
## 188  33        yes
## 189  31        yes
## 190  25        yes
```

```
## 191   21          no
## 192   40          no
## 193   36         yes
## 194   40         yes
## 195   42          no
## 196   29         yes
## 197   21          no
## 198   23         yes
## 199   26         yes
## 200   29         yes
## 201   21          no
## 202   28          no
## 203   32          no
## 204   27          no
## 205   55          no
## 206   27          no
## 207   57         yes
## 208   52         yes
## 209   21          no
## 210   41         yes
## 211   25          no
## 212   24          no
## 213   60          no
## 214   24         yes
## 215   36         yes
## 216   38         yes
## 217   25         yes
## 218   32          no
## 219   32         yes
## 220   41         yes
## 221   21         yes
## 222   66         yes
## 223   37          no
## 224   61          no
## 225   26          no
## 226   22          no
## 227   26          no
## 228   24         yes
## 229   31          no
## 230   24          no
## 231   22         yes
## 232   46         yes
## 233   22          no
## 234   29          no
## 235   23          no
## 236   26         yes
## 237   51         yes
## 238   23         yes
## 239   32         yes
```

```
## 240   27          no
## 241   21          no
## 242   22          no
## 243   22         yes
## 244   33         yes
## 245   29          no
## 246   49         yes
## 247   41          no
## 248   23          no
## 249   34          no
## 250   23          no
## 251   42          no
## 252   27          no
## 253   24          no
## 254   25          no
## 255   44         yes
## 256   21         yes
## 257   30          no
## 258   25          no
## 259   24          no
## 260   51         yes
## 261   34          no
## 262   27         yes
## 263   24          no
## 264   63          no
## 265   35         yes
## 266   43          no
## 267   25         yes
## 268   24          no
## 269   21          no
## 270   28         yes
## 271   38         yes
## 272   21          no
## 273   40          no
## 274   21          no
## 275   52          no
## 276   25          no
## 277   29         yes
## 278   23          no
## 279   57          no
## 280   22          no
## 281   28         yes
## 282   39          no
## 283   37          no
## 284   47         yes
## 285   52         yes
## 286   51          no
## 287   34          no
## 288   29         yes
```

```
## 289  26        no
## 290  33        no
## 291  21        no
## 292  25       yes
## 293  31       yes
## 294  24       yes
## 295  65        no
## 296  28        no
## 297  29       yes
## 298  24        no
## 299  46       yes
## 300  58        no
## 301  30       yes
## 302  25       yes
## 303  35        no
## 304  28       yes
## 305  37        no
## 306  29        no
## 307  47       yes
## 308  21        no
## 309  25       yes
## 310  30       yes
## 311  41        no
## 312  22        no
## 313  27       yes
## 314  25        no
## 315  43       yes
## 316  26        no
## 317  30        no
## 318  29       yes
## 319  28        no
## 320  59       yes
## 321  31        no
## 322  25       yes
## 323  36       yes
## 324  43       yes
## 325  21        no
## 326  24        no
## 327  30       yes
## 328  37        no
## 329  23       yes
## 330  37        no
## 331  46        no
## 332  25        no
## 333  41       yes
## 334  44        no
## 335  22        no
## 336  26        no
## 337  44        no
```

```
## 338   44         yes
## 339   33         yes
## 340   41         yes
## 341   22          no
## 342   36          no
## 343   22          no
## 344   33          no
## 345   57          no
## 346   49          no
## 347   22          no
## 348   23          no
## 349   26          no
## 350   37         yes
## 351   29          no
## 352   30          no
## 353   46          no
## 354   24          no
## 355   21          no
## 356   49         yes
## 357   28         yes
## 358   44         yes
## 359   48          no
## 360   29         yes
## 361   29         yes
## 362   63          no
## 363   65          no
## 364   67         yes
## 365   30          no
## 366   30          no
## 367   29         yes
## 368   21          no
## 369   22          no
## 370   45         yes
## 371   25         yes
## 372   21          no
## 373   21          no
## 374   25          no
## 375   28          no
## 376   58         yes
## 377   22          no
## 378   22          no
## 379   32         yes
## 380   35          no
## 381   24          no
## 382   22          no
## 383   21          no
## 384   25          no
## 385   25          no
## 386   24          no
```

```
## 387   35       yes
## 388   45       yes
## 389   58       yes
## 390   28        no
## 391   42        no
## 392   27       yes
## 393   21        no
## 394   37        no
## 395   31       yes
## 396   25        no
## 397   39        no
## 398   22       yes
## 399   25        no
## 400   25       yes
## 401   31       yes
## 402   55        no
## 403   35       yes
## 404   38        no
## 405   41       yes
## 406   26        no
## 407   46       yes
## 408   25        no
## 409   39       yes
## 410   28       yes
## 411   28        no
## 412   25        no
## 413   22        no
## 414   21        no
## 415   21       yes
## 416   22       yes
## 417   22        no
## 418   37       yes
## 419   27        no
## 420   28       yes
## 421   26        no
## 422   21        no
## 423   21        no
## 424   21        no
## 425   36       yes
## 426   31       yes
## 427   25        no
## 428   38       yes
## 429   26        no
## 430   43       yes
## 431   23        no
## 432   38        no
## 433   22        no
## 434   29        no
## 435   36        no
```

```
## 436  29        yes
## 437  41         no
## 438  28         no
## 439  21         no
## 440  31         no
## 441  41        yes
## 442  22         no
## 443  24         no
## 444  33        yes
## 445  30        yes
## 446  25        yes
## 447  28         no
## 448  26         no
## 449  22        yes
## 450  26         no
## 451  23         no
## 452  23        yes
## 453  25         no
## 454  72         no
## 455  24         no
## 456  38        yes
## 457  62         no
## 458  24         no
## 459  51        yes
## 460  81         no
## 461  48         no
## 462  26         no
## 463  39         no
## 464  37         no
## 465  34         no
## 466  21         no
## 467  22         no
## 468  25         no
## 469  38        yes
## 470  27         no
## 471  28         no
## 472  22         no
## 473  22         no
## 474  50         no
## 475  24         no
## 476  59         no
## 477  29        yes
## 478  31         no
## 479  39         no
## 480  63         no
## 481  35        yes
## 482  29         no
## 483  28         no
## 484  23         no
```

```
## 485   31        yes
## 486   24        yes
## 487   21         no
## 488   58         no
## 489   28         no
## 490   67         no
## 491   24         no
## 492   42         no
## 493   33         no
## 494   45        yes
## 495   22         no
## 496   66         no
## 497   30         no
## 498   25         no
## 499   55        yes
## 500   39         no
## 501   21         no
## 502   28         no
## 503   41        yes
## 504   41         no
## 505   40         no
## 506   38         no
## 507   35        yes
## 508   21         no
## 509   21         no
## 510   64         no
## 511   46        yes
## 512   21         no
## 513   58         no
## 514   22         no
## 515   24         no
## 516   28        yes
## 517   53        yes
## 518   51         no
## 519   41         no
## 520   60         no
## 521   25         no
## 522   26         no
## 523   26         no
## 524   45        yes
## 525   24         no
## 526   21         no
## 527   21         no
## 528   24         no
## 529   22         no
## 530   31         no
## 531   22         no
## 532   24         no
## 533   29         no
```

```
## 534  31        no
## 535  24        no
## 536  23       yes
## 537  46        no
## 538  67        no
## 539  23        no
## 540  32       yes
## 541  43       yes
## 542  27       yes
## 543  56       yes
## 544  25        no
## 545  29        no
## 546  37       yes
## 547  53       yes
## 548  28        no
## 549  50        no
## 550  37        no
## 551  21        no
## 552  25        no
## 553  66        no
## 554  23        no
## 555  28        no
## 556  37        no
## 557  30        no
## 558  58        no
## 559  42        no
## 560  35        no
## 561  54       yes
## 562  28       yes
## 563  24        no
## 564  32        no
## 565  27        no
## 566  22        no
## 567  21        no
## 568  46        no
## 569  37        no
## 570  33       yes
## 571  39        no
## 572  21        no
## 573  22        no
## 574  22        no
## 575  23        no
## 576  25        no
## 577  35        no
## 578  21       yes
## 579  36        no
## 580  62       yes
## 581  21       yes
## 582  27        no
```

```
## 583  62        no
## 584  42        no
## 585  52       yes
## 586  22        no
## 587  41       yes
## 588  29        no
## 589  52       yes
## 590  25        no
## 591  45       yes
## 592  24        no
## 593  44       yes
## 594  25        no
## 595  34        no
## 596  22       yes
## 597  46        no
## 598  21        no
## 599  38       yes
## 600  26        no
## 601  24        no
## 602  28        no
## 603  30        no
## 604  54       yes
## 605  36       yes
## 606  21        no
## 607  22       yes
## 608  25        no
## 609  27        no
## 610  23        no
## 611  24        no
## 612  36       yes
## 613  40       yes
## 614  26        no
## 615  50       yes
## 616  27        no
## 617  30        no
## 618  23        no
## 619  50       yes
## 620  24       yes
## 621  28        no
## 622  28        no
## 623  45        no
## 624  21        no
## 625  21        no
## 626  29        no
## 627  21        no
## 628  21        no
## 629  45        no
## 630  21        no
## 631  34       yes
```

```
## 632  24        no
## 633  23        no
## 634  22        no
## 635  31        no
## 636  38       yes
## 637  48        no
## 638  23        no
## 639  32       yes
## 640  28        no
## 641  27        no
## 642  24        no
## 643  50       yes
## 644  31        no
## 645  27        no
## 646  30        no
## 647  33       yes
## 648  22       yes
## 649  42       yes
## 650  23        no
## 651  23        no
## 652  27        no
## 653  28        no
## 654  27        no
## 655  22        no
## 656  25       yes
## 657  22        no
## 658  41        no
## 659  51        no
## 660  27       yes
## 661  54        no
## 662  22       yes
## 663  43       yes
## 664  40       yes
## 665  40       yes
## 666  24        no
## 667  70       yes
## 668  40       yes
## 669  43        no
## 670  45        no
## 671  49        no
## 672  21        no
## 673  47        no
## 674  22        no
## 675  68        no
## 676  31       yes
## 677  53       yes
## 678  25        no
## 679  25       yes
## 680  23        no
```

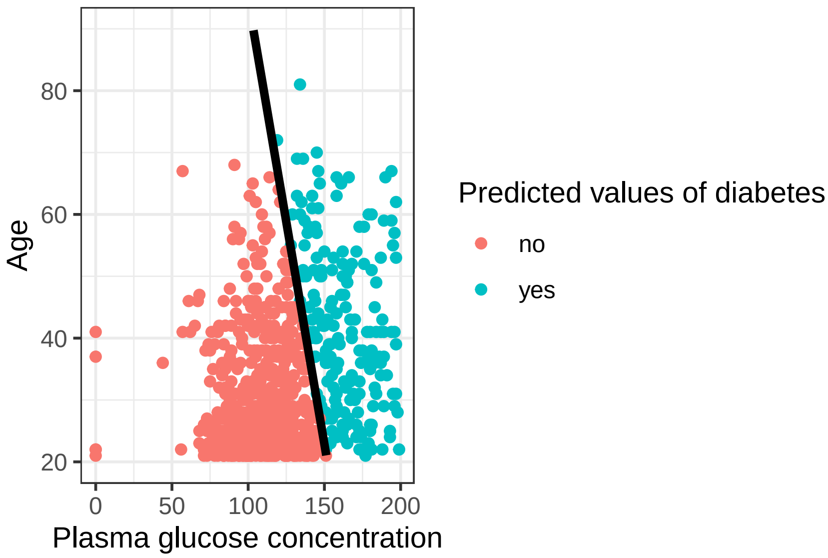
```
## 681  22        no
## 682  26       yes
## 683  22        no
## 684  27       yes
## 685  69        no
## 686  25        no
## 687  22        no
## 688  29        no
## 689  23        no
## 690  46       yes
## 691  34        no
## 692  44       yes
## 693  23        no
## 694  43       yes
## 695  25        no
## 696  43       yes
## 697  31       yes
## 698  22        no
## 699  28        no
## 700  26        no
## 701  26        no
## 702  49       yes
## 703  52       yes
## 704  41        no
## 705  27        no
## 706  28        no
## 707  30       yes
## 708  22        no
## 709  45       yes
## 710  23       yes
## 711  24        no
## 712  40        no
## 713  38       yes
## 714  21        no
## 715  32        no
## 716  34       yes
## 717  31       yes
## 718  56        no
## 719  24        no
## 720  52       yes
## 721  34        no
## 722  21        no
## 723  42       yes
## 724  42        no
## 725  45        no
## 726  38        no
## 727  25        no
## 728  22        no
## 729  22        no
```

```
## 730  22       no
## 731  34      yes
## 732  22      yes
## 733  24      yes
## 734  22       no
## 735  53       no
## 736  28       no
## 737  21       no
## 738  42       no
## 739  21       no
## 740  42      yes
## 741  48      yes
## 742  26       no
## 743  22       no
## 744  45      yes
## 745  39       no
## 746  46       no
## 747  27      yes
## 748  32       no
## 749  36      yes
## 750  50      yes
## 751  22      yes
## 752  28       no
## 753  25       no
## 754  26      yes
## 755  45      yes
## 756  37      yes
## 757  39       no
## 758  52      yes
## 759  26       no
## 760  66      yes
## 761  22       no
## 762  43      yes
## 763  33       no
## 764  63       no
## 765  27       no
## 766  30       no
## 767  47      yes
## 768  23       no
```

```
 #coef(model)[1] + t(as.matrix(coef(model)[2:3])) %*% as.matrix(diabetes_df[,c("plasma_glucose_conc","ag

ggplot(diabetes_df_pred, aes(x = plasma_glucose_conc, y = age, color = pred)) +
  geom_point() +
  theme_bw() +
 stat_function(fun = ({function(x) (-coef(model)[1] - coef(model)[2]*x)/ coef(model)[3] }),
               size=1.5, color = "black") +
  ylim(20,90) +
  labs(colour = "Predicted values of diabetes",
       x = "Plasma glucose concentration",
```

```
       y = "Age")
```

```
## Warning: Removed 76 row(s) containing missing values (geom_path).
```



Question:

Comment whether the decision boundary seems to catch the data distribution well.

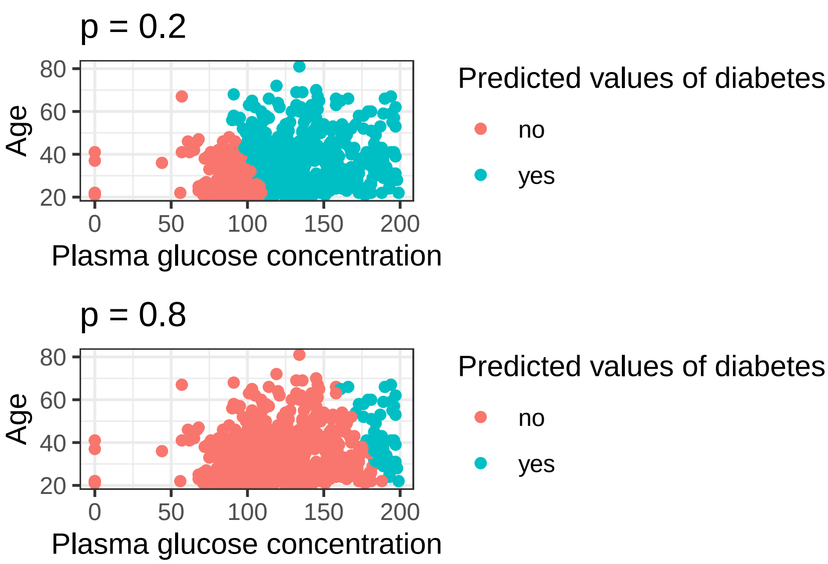
2.5 3.4

Question:

Make same kind of plots as in step 2 but use thresholds $r = 0.2$ and $r = 0.8$. By using these plots

```
library("ggpubr")

# Using 0.2 as the classification threshold
pred <- predict(model, newdata = diabetes_df, type = "response")
pred <- ifelse(pred > 0.2, "yes", "no")
diabetes_df_pred$pred <- pred


p1 <- ggplot(diabetes_df_pred, aes(x = plasma_glucose_conc, y = age, color = pred)) +
      geom_point() +
      theme_bw() +
      labs(colour = "Predicted values of diabetes",
           x = "Plasma glucose concentration",
           y = "Age") +
      ggtitle("p = 0.2")


# Using 0.8 as the classification threshold
pred <- predict(model, newdata = diabetes_df, type = "response")
```

```
pred <- ifelse(pred > 0.8, "yes", "no")
diabetes_df_pred$pred <- pred


p2 <- ggplot(diabetes_df_pred, aes(x = plasma_glucose_conc, y = age, color = pred)) +
      geom_point() +
      theme_bw() +
      labs(colour = "Predicted values of diabetes",
           x = "Plasma glucose concentration",
           y = "Age") +
  ggtitle("p = 0.8")


ggarrange(p1, p2,  ncol = 1, nrow = 2)
```



Question:

Comment on what happens with the prediction when r value changes.

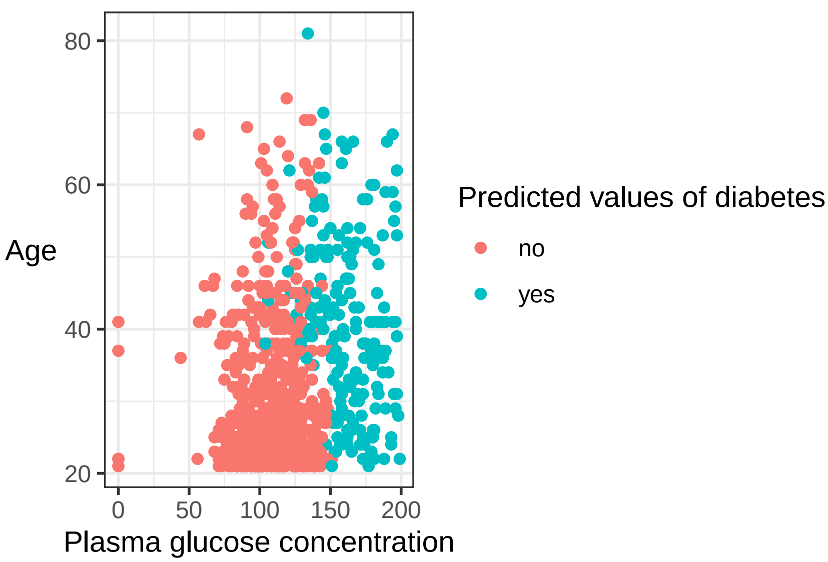
2.6 3.5

Question:

Perform a basis function expansion trick by computing new features $z_1 = x_1^4$, $z_2 = x_1^3 x^2$, $z_3 = x_1^2 x_2^2$, $z_4 = x_1 x_2^3$, $z_5 = x_2^4$, adding them to the data set and then computing a logistic regression model with y as target and $x_1, x_2, z_1, ..., z_5$ as features. Create a scatterplot of the same kind as in step 2 for this model.

```
diabetes_df$z1  <- diabetes_df$times_pregnant^4
diabetes_df$z2  <- diabetes_df$times_pregnant^3 * diabetes_df$plasma_glucose_conc^2
diabetes_df$z3  <- diabetes_df$times_pregnant^2 * diabetes_df$plasma_glucose_conc^2
diabetes_df$z4  <- diabetes_df$times_pregnant * diabetes_df$plasma_glucose_conc^3
diabetes_df$z5  <- diabetes_df$plasma_glucose_conc^4
```

```r
model <- glm(diabetes ~ plasma_glucose_conc + age + z1 + z2 + z3 + z4 + z5, data = diabetes_df,
             family =  "binomial")

pred <- predict(model, newdata = diabetes_df, type = "response")

# Using 0.5 as the classification threshold
pred <- ifelse(pred > 0.5, "yes", "no")

diabetes_df_pred <- diabetes_df
diabetes_df_pred$pred <- pred

ggplot(diabetes_df_pred, aes(x = plasma_glucose_conc, y = age, color = pred)) +
  geom_point() +
  theme_bw() +
  theme(axis.title.y = element_text(angle = 0,vjust = 0.5)) +
  labs(colour = "Predicted values of diabetes",
       x = "Plasma glucose concentration",
       y = "Age")
```



Question:

Compute the training misclassification rate. What can you say about the quality of this model compared to the previous logistic regression model? How have the basis expansion trick affected the shape of the decision boundary and the prediction accuracy?

3 Statement of Contribution

We worked on the assignment individually for the computer labs (to be more efficient when asking questions), Duc on task 1, Sigme on task 2, and William on task 3. We later solved all assignment individually and compared and discussed our solutions before dividing the task of writing the laboration report.

3.1 Question 1

Text written by Duc.

3.2 Question 2

Text written by Sigme.

3.3 Question 3

Text written by William.

4 Appendix

The code used in this laboration report are summarised in the code as follows:

```
library(ggplot2)
library(kknn)
library(dplyr)
library(knitr)
library(caret)
library(psych)
knitr::opts_chunk$set(
  echo = TRUE,
  fig.width = 4.5,
  fig.height = 3)
# Read in data
data <- read.csv("optdigits.csv")

# Renaming the response variable and changing it to a factor variable
data <- rename(data, y=X0.26)
data$y <- as.factor(data$y)

# Partitioning training data (50%)
n=dim(data)[1]
set.seed(12345)
id=sample(1:n, floor(n*0.5))
train=data[id,]

# Partitioning validation data (25%)
id1=setdiff(1:n, id)
set.seed(12345)
id2=sample(id1, floor(n*0.25))
valid=data[id2,]

# Partitioning test data (25%)
id3=setdiff(id1,id2)
test=data[id3,]
# kknn on training data and evaluation on training data
model_kknn_train <- kknn(formula=y~., train=train, test=train, kernel="rectangular", k=30)
```

```r
conf_mat_train <- table(train$y, model_kknn_train$fitted.values)
acc_train <- sum(diag(conf_mat_train)) / sum(conf_mat_train)
miss_train <- 1-acc_train

# kknn on training data and evaluation on test data
model_kknn_test <- kknn(formula=y~., train=train, test=test, kernel="rectangular", k=30)
conf_mat_test <- table(test$y, model_kknn_test$fitted.values)
acc_test <- sum(diag(conf_mat_test)) / sum(conf_mat_test)
miss_test <- 1-acc_test

# Rows are true values, columns are model prediction
kable(conf_mat_train, caption = "Confusion matrix for training data, model
      predictions by columns and true value by rows.")
miss_train
kable(conf_mat_test, caption = "Confusion matrix for test data, model
      predictions by columns and true value by rows.")
miss_test
y <- train$y
fit_y <- model_kknn_train$fitted.values
# probabilities given from number 0 to 9, index 9 = number 8.
prob_8 <- model_kknn_train$prob[, 9]

# Data frame consisting of true value of y, model prediction and the models
# probability that the number is 8.
data_8 <- data.frame(y = y, fit_y = fit_y, prob = prob_8)
data_8$observation_id <- rownames(data_8)

# Only observations with the label 8 is kept.
data_8 <- data_8[data_8$y == "8", ]
head(arrange(data_8, prob), 2)
tail(arrange(data_8, prob), 3)
# Change colour palette to black and white
colfunc <- colorRampPalette(c("white", "black"))

plot_8 <- function(index){
  title <- paste0("Obs: ", index)
  # Reshapes the observations to a 8x8
  plot <- as.matrix(train[index, -65]) # Remove response variable
  plot <- matrix(plot, nrow=8, byrow=TRUE)
  heatmap(plot, col=colfunc(16), Colv=NA, Rowv=NA, main=title, margins=c(2,2))
}
plot_8(1624)
plot_8(1663)
plot_8(1810)
plot_8(1811)
plot_8(1864)
fit_kknn <- function(k){
  model_kknn_train <- kknn(formula=y~., train=train, test=train, kernel="rectangular", k=k)
  # Confusion matrix for train data
```

```
  conf_mat_train <- table(model_kknn_train$fitted.values, train$y)
  acc_train <- sum(diag(conf_mat_train)) / sum(conf_mat_train)
  # Missclassification for training data
  miss_train <- 1-acc_train

  model_kknn_valid <- kknn(formula=y~., train=train, test=valid, kernel="rectangular", k=k)
  # Confusion matrix for validation data
  conf_mat_valid <- table(model_kknn_valid$fitted.values, valid$y)
  acc_valid <- sum(diag(conf_mat_valid)) / sum(conf_mat_valid)
  # Missclassification for validation data
  miss_valid <- 1-acc_valid

  result <- c(miss_train, miss_valid)
  return(result)
}

# Missclassification for k=1,...,30 for training and validation data
result <- data.frame(train = 0, valid = 0)
for(i in 1:30){
  model <- fit_kknn(i)
  result[i,1] <- model[1]
  result[i,2] <- model[2]
}
result$index <- 1:30
ggplot(result, aes(x=index)) +
  geom_line(aes(y=train, colour="train")) +
  geom_point(aes(y=train, colour="train")) +
  geom_line(aes(y=valid, colour="valid")) +
  geom_point(aes(y=valid, colour="valid")) +
  scale_color_manual(name = "Data",
                     values = c("train" = "steelblue", "valid" = "indianred")) +
  scale_x_continuous(breaks = c(seq(from=0, to=30, by=5))) +
  scale_y_continuous(limits = c(0, 0.06)) +
  theme_bw() +
  labs(x = "k",
       y = "Missclassification rate")
which(result$valid == min(result$valid))
model_test_7 <- kknn(formula = y~., train = train, test = test, kernel = "rectangular", k=7)

conf_mat_test <- table(model_test_7$fitted.values, test$y)
acc_test <- sum(diag(conf_mat_test)) / sum(conf_mat_test)
miss_test <- 1-acc_test

table_data <- cbind(training=result[7, 1], validtion=result[7, 2], test=miss_test)
kable(table_data, digits=3, caption="Misclassification error k=7 for different data.")
cross_entropy <- function(k){
  model_kknn_valid <-
    kknn(formula = y~.,
         train = train,
```

```r
         test = valid,
         kernel = "rectangular",
         k=k)

  y <- as.integer(valid$y)

  prob <- c()
  for(i in 1:length(y)){
    prob[i] <- model_kknn_valid$prob[i, y[i]]
  }

  value <- -sum(log(prob + 1e-15))
  return(value)
}

result <- c()
for(i in 1:30){
  model <- cross_entropy(i)
  result[i] <- model
}
plot_data <- data.frame(index=1:30, result)
ggplot(plot_data, aes(x=index, y=result)) +
  geom_point(color="forestgreen") +
  geom_line(color="forestgreen") +
  scale_x_continuous(breaks = c(seq(from=0, to=30, by=5))) +
  theme_bw() +
  labs(x="K",
       y="Cross-entropy")
which(min(result) == result)
df2 <- read.csv("parkinsons.csv")

# Shuffle the data
set.seed(123)
df2 <- df2[sample(nrow(df2)), ]
set.seed(123)
# Split train and test
train_indices <- createDataPartition(df2$motor_UPDRS, p = 0.6, list = FALSE)
train_data <- df2[train_indices, ]
test_data <- df2[-train_indices, ]

predictor_cols <- setdiff(names(train_data), "motor_UPDRS")
scaler <- preProcess(train_data)
trainS <- predict(scaler, train_data)
testS <- predict(scaler, test_data)

#train_sd <- apply(train_data, 2, sd)

# Linear regression model
lm_model <- lm(motor_UPDRS ~ ., data = trainS)
```

```r
# Predictions on the test data
trainS_x <- trainS[, predictor_cols]
testS_x <- testS[, predictor_cols]

predS_train <- predict(lm_model, newdata = trainS_x)
predS_test <- predict(lm_model, newdata = testS_x)

mse_train <- mean((trainS$motor_UPDRS - predS_train)^2)
mse_test <- mean((testS$motor_UPDRS - predS_test)^2)

cat("Mean Squared Error (MSE) on the training data:", mse_train, "\n")
cat("Mean Squared Error (MSE) on the test data:", mse_test, "\n")
plot(testS$motor_UPDRS, predS_test, main = "Linear Regression (Scaled Data)",
     xlab = "Actual Values", ylab = "Predicted Values", pch = 19, col = "blue")
abline(a = 0, b = 1, col = "red")
# Define the functions
Loglikelihood <- function(theta, std){
  n <- nrow(trainS_x)
  prediction <- as.matrix(trainS_x) %*% as.matrix(theta)
  actual <- trainS$motor_UPDRS
  res <- actual-prediction
  likelihood <- (-(n/2) * log(2*pi*std^2) - (1/(2*std^2)) * sum(res^2))
  return(likelihood)
}

Ridge <- function(theta, std, lambda){
  likelihood_ridge <- -Loglikelihood(theta, std) + (lambda/2)*sum(theta^2)
  return(likelihood_ridge)
}

#optim() function minimizes
RidgeOpt <- function(lambda){
  # Define a new function to optimize
  my_fnc <- function(parameters){
    theta <- parameters[1:(length(parameters)-1)]
    std <- parameters[length(parameters)]
    return(Ridge(theta, std, lambda))
  }
  initial_values <- c(rep(0, ncol(trainS_x)), 1)

  optimal_values <- optim(par = initial_values, fn = my_fnc, method = "BFGS")$par
  optimal_theta <- optimal_values[1:length(predictor_cols)]
  optimal_std <- optimal_values[length(predictor_cols) + 1]
  optimal_lambda <- optimal_values[length(optimal_values)]

  result_list <- list(theta = optimal_theta, std = optimal_std, lambda = optimal_lambda)
  return(result_list)
}
```

```r
library(psych)
DF <- function(lambda){
  X <- as.matrix(trainS_x)
  dof <- tr(X %*% (solve(t(X) %*% X + lambda*diag(ncol(trainS_x)))) %*% t(X))
  return(dof)
}
lambda <- 1
result_ridge_1 <- RidgeOpt(lambda)

optimal_theta_1 <- result_ridge_1$theta
optimal_std_1 <- result_ridge_1$std
optimal_lambda_1 <- result_ridge_1$lambda

ridge_pred_train_1 <- as.matrix(trainS_x) %*% as.matrix(optimal_theta_1)
ridge_pred_test_1 <- as.matrix(testS_x) %*% as.matrix(optimal_theta_1)

mse_ridge_train_1 <- mean((trainS$motor_UPDRS - ridge_pred_train_1)^2)
mse_ridge_test_1 <- mean((testS$motor_UPDRS - ridge_pred_test_1)^2)

cat("Lambda:", lambda, "\n")
cat("Mean Squared Error (MSE) ridge regression on training data:", mse_ridge_train_1, "\n")
cat("Mean Squared Error (MSE) ridge regression on test data:", mse_ridge_test_1, "\n")
cat("Degree of freedom:", DF(lambda), "\n")
plot(testS$motor_UPDRS, ridge_pred_test_1, main = "Ridge Regression, lambda = 1",
     xlab = "Actual Values", ylab = "Predicted Values", pch = 19, col = "blue")
abline(a = 0, b = 1, col = "red")
lambda <- 100
result_ridge_100 <- RidgeOpt(lambda)

optimal_theta_100 <- result_ridge_100$theta
optimal_std_100 <- result_ridge_100$std
optimal_lambda_100 <- result_ridge_100$lambda

ridge_pred_train_100 <- as.matrix(trainS_x) %*% as.matrix(optimal_theta_100)
ridge_pred_test_100 <- as.matrix(testS_x) %*% as.matrix(optimal_theta_100)

mse_ridge_train_100 <- mean((trainS$motor_UPDRS - ridge_pred_train_100)^2)
mse_ridge_test_100 <- mean((testS$motor_UPDRS - ridge_pred_test_100)^2)

cat("Lambda:", lambda, "\n")
cat("Mean Squared Error (MSE) ridge regression on training data:", mse_ridge_train_100, "\n")
cat("Mean Squared Error (MSE) ridge regression on test data:", mse_ridge_test_100, "\n")
cat("Degree of freedom:", DF(lambda), "\n")
plot(testS$motor_UPDRS, ridge_pred_test_100, main = "Ridge Regression, lambda = 100",
     xlab = "Actual Values", ylab = "Predicted Values", pch = 19, col = "blue")
abline(a = 0, b = 1, col = "red")
lambda <- 1000
result_ridge_1000 <- RidgeOpt(lambda)
```

```
optimal_theta_1000 <- result_ridge_1000$theta
optimal_std_1000 <- result_ridge_1000$std
optimal_lambda_1000 <- result_ridge_1000$lambda

ridge_pred_train_1000 <- as.matrix(trainS_x) %*% as.matrix(optimal_theta_1000)
ridge_pred_test_1000 <- as.matrix(testS_x) %*% as.matrix(optimal_theta_1000)

mse_ridge_train_1000 <- mean((trainS$motor_UPDRS - ridge_pred_train_1000)^2)
mse_ridge_test_1000 <- mean((testS$motor_UPDRS - ridge_pred_test_1000)^2)

cat("Lambda:", lambda, "\n")
cat("Mean Squared Error (MSE) ridge regression on training data:", mse_ridge_train_1000, "\n")
cat("Mean Squared Error (MSE) ridge regression on test data:", mse_ridge_test_1000, "\n")
cat("Degree of freedom:", DF(lambda), "\n")
plot(testS$motor_UPDRS, ridge_pred_test_1000, main = "Ridge Regression, lambda = 1000",
     xlab = "Actual Values", ylab = "Predicted Values", pch = 19, col = "blue")
abline(a = 0, b = 1, col = "red")

diabetes_df <- read.csv("pima-indians-diabetes.csv", header=FALSE)

colnames(diabetes_df) <- c("times_pregnant", "plasma_glucose_conc",
                        "diastolic_blood_pressure", "triceps_skinfold_thickness",
                        "serum_insulin","body_mass_index", "diabetes_pedigree",
                        "age", "diabetes")

diabetes_df$diabetes <- ifelse(diabetes_df$diabetes == 0, "no", "yes")
diabetes_df$diabetes <- as.factor(diabetes_df$diabetes)



library(ggplot2)

ggplot(diabetes_df, aes(x = plasma_glucose_conc, y = age, color = diabetes)) +
  geom_point() +
  theme_bw() +
  theme(axis.title.y = element_text(angle = 0,vjust = 0.5)) +
  labs(colour = "Diabetes",
       x = "Plasma glucose concentration",
       y = "Age")


model <- glm(diabetes ~ plasma_glucose_conc + age, data = diabetes_df,
             family =  "binomial")

pred <- predict(model, newdata = diabetes_df, type = "response")

# Using 0.5 as the classification threshold
pred <- ifelse(pred > 0.5, "yes", "no")
```

```r
confusion <- table(diabetes_df$diabetes, pred)
misclass_rate <- (confusion[1,2] + confusion[2,1]) / sum(confusion)


knitr::kable(as.data.frame(round(misclass_rate,2)), col.names = "Misclassification error",
             caption = "Misclassification error")


diabetes_df_pred <- diabetes_df
diabetes_df_pred$pred <- pred

ggplot(diabetes_df_pred, aes(x = plasma_glucose_conc, y = age, color = pred)) +
  geom_point() +
  theme_bw() +
  theme(axis.title.y = element_text(angle = 0,vjust = 0.5)) +
  labs(colour = "Predicted values of diabetes",
       x = "Plasma glucose concentration",
       y = "Age")



summary(model)

slope <- coef(model)[2]/(-coef(model)[3])
intercept <- coef(model)[1]/(-coef(model)[3])


diabetes_df


 #coef(model)[1] + t(as.matrix(coef(model)[2:3])) %*% as.matrix(diabetes_df[,c("plasma_glucose_conc","ag

ggplot(diabetes_df_pred, aes(x = plasma_glucose_conc, y = age, color = pred)) +
  geom_point() +
  theme_bw() +
 stat_function(fun = ({function(x) (-coef(model)[1] - coef(model)[2]*x)/ coef(model)[3] }),
               size=1.5, color = "black") +
  ylim(20,90) +
  labs(colour = "Predicted values of diabetes",
       x = "Plasma glucose concentration",
       y = "Age")


library("ggpubr")

# Using 0.2 as the classification threshold
pred <- predict(model, newdata = diabetes_df, type = "response")
pred <- ifelse(pred > 0.2, "yes", "no")
diabetes_df_pred$pred <- pred
```

```r
p1 <- ggplot(diabetes_df_pred, aes(x = plasma_glucose_conc, y = age, color = pred)) +
      geom_point() +
      theme_bw() +
      labs(colour = "Predicted values of diabetes",
           x = "Plasma glucose concentration",
           y = "Age") +
      ggtitle("p = 0.2")


# Using 0.8 as the classification threshold
pred <- predict(model, newdata = diabetes_df, type = "response")
pred <- ifelse(pred > 0.8, "yes", "no")
diabetes_df_pred$pred <- pred


p2 <- ggplot(diabetes_df_pred, aes(x = plasma_glucose_conc, y = age, color = pred)) +
      geom_point() +
      theme_bw() +
      labs(colour = "Predicted values of diabetes",
           x = "Plasma glucose concentration",
           y = "Age") +
  ggtitle("p = 0.8")


ggarrange(p1, p2,  ncol = 1, nrow = 2)


diabetes_df$z1  <- diabetes_df$times_pregnant^4
diabetes_df$z2  <- diabetes_df$times_pregnant^3 * diabetes_df$plasma_glucose_conc^2
diabetes_df$z3  <- diabetes_df$times_pregnant^2 * diabetes_df$plasma_glucose_conc^2
diabetes_df$z4  <- diabetes_df$times_pregnant * diabetes_df$plasma_glucose_conc^3
diabetes_df$z5  <- diabetes_df$plasma_glucose_conc^4


model <- glm(diabetes ~ plasma_glucose_conc + age + z1 + z2 + z3 + z4 + z5, data = diabetes_df,
             family =  "binomial")

pred <- predict(model, newdata = diabetes_df, type = "response")

# Using 0.5 as the classification threshold
pred <- ifelse(pred > 0.5, "yes", "no")

diabetes_df_pred <- diabetes_df
diabetes_df_pred$pred <- pred

ggplot(diabetes_df_pred, aes(x = plasma_glucose_conc, y = age, color = pred)) +
  geom_point() +
```

```r
  theme_bw() +
  theme(axis.title.y = element_text(angle = 0,vjust = 0.5)) +
  labs(colour = "Predicted values of diabetes",
       x = "Plasma glucose concentration",
       y = "Age")
```