

Práctica 2

Jorge Cutipa Musaja

30 de diciembre, 2018

- 1. Descripción del dataset
- 2. Integración y selección de los datos de interés a analizar
- 3. Limpieza de los datos
 - 3.1. ¿Los datos contienen ceros o elementos vacíos? ¿Cómo gestionarías cada uno de estos casos?
 - 3.2. Identificación y tratamiento de valores extremos.
- 4. Análisis de los datos
 - 4.1. Selección de los grupos de datos que se quieren analizar/comparar
 - 4.2. Comprobación de la normalidad y homogeneidad de la varianza
 - 4.3. Aplicación de pruebas estadísticas para comparar los grupos de datos
- 5. Representación de los resultados a partir de tablas y gráficas.
- 6. Resolución del problema.

1. Descripción del dataset

En abril de 1912, durante su viaje inaugural, el Titanic se hundió después de chocar con un iceberg, matando a 1502 de 2224 pasajeros y tripulantes. Una de las razones por las que el naufragio llevó a tal pérdida de vidas humanas fue que no había suficientes botes salvavidas para los pasajeros y la tripulación. Aunque siempre habrá algún elemento de azar a considerar, algunos grupos de personas tenían más probabilidades de sobrevivir al hundimiento que otros.

El objetivo de esta práctica es elaborar un árbol de clasificación, para predecir que grupos de pasajeros y tripulantes tienen - o mejor dicho, tenían - una mayor probabilidad de sobrevivir al hundimiento del Titanic.

Una descripción de los tipos de datos disponibles es lo que visualizaremos a continuación:

```
glimpse(datos)
```

```
## Observations: 891
## Variables: 12
## $ PassengerId <int> 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15,...
## $ Survived <int> 0, 1, 1, 1, 0, 0, 0, 0, 1, 1, 1, 1, 0, 0, 0, 1, 0,...
## $ Pclass <int> 3, 1, 3, 1, 3, 3, 1, 3, 3, 2, 3, 1, 3, 3, 3, 2, 3,...
## $ Name <chr> "Braund, Mr. Owen Harris", "Cumings, Mrs. John Bra...
## $ Sex <chr> "male", "female", "female", "female", "male", "mal...
## $ Age <dbl> 22, 38, 26, 35, 35, NA, 54, 2, 27, 14, 4, 58, 20, ...
## $ SibSp <int> 1, 1, 0, 1, 0, 0, 0, 3, 0, 1, 1, 0, 0, 1, 0, 0, 4,...
## $ Parch <int> 0, 0, 0, 0, 0, 0, 0, 1, 2, 0, 1, 0, 0, 5, 0, 0, 1,...
## $ Ticket <chr> "A/5 21171", "PC 17599", "STON/O2. 3101282", "1138...
## $ Fare <dbl> 7.2500, 71.2833, 7.9250, 53.1000, 8.0500, 8.4583, ...
## $ Cabin <chr> "", "C85", "", "C123", "", "", "E46", "", "", "", ...
## $ Embarked <chr> "S", "C", "S", "S", "S", "Q", "S", "S", "S", "C", ...
```

Como se observa, Survived y Pclass han sido consideradas como variables numéricas, cuando en realidad son de tipo *clase (factor)*. Corregiremos esto.

```
datos$Survived <- factor(datos$Survived)
datos$Pclass <- factor(datos$Pclass)
```

En resumen, en total, se tienen 891 registros y 12 variables, incluyendo a la variable objetivo: Survived.

2. Integración y selección de los datos de interés a analizar

Como realizaremos un árbol de clasificación, no es necesario hacer una preselección de variables relevantes. Lo único que haremos será eliminar variables que no aportan información, como el nombre de los pasajeros, número de ticket, entre otros.

```
datos <- datos[,colnames(datos)!="PassengerId"]
datos <- datos[,colnames(datos)!="Name"]
datos <- datos[,colnames(datos)!="Ticket"]
datos <- datos[,colnames(datos)!="Cabin"]
```

Ahora, en total, se tienen 891 registros y 8 variables, incluyendo a la variable objetivo: Survived.

3. Limpieza de los datos

3.1. ¿Los datos contienen ceros o elementos vacíos? ¿Cómo gestionarías cada uno de estos casos?

Podemos crear una función para identificar las variables con elementos vacíos

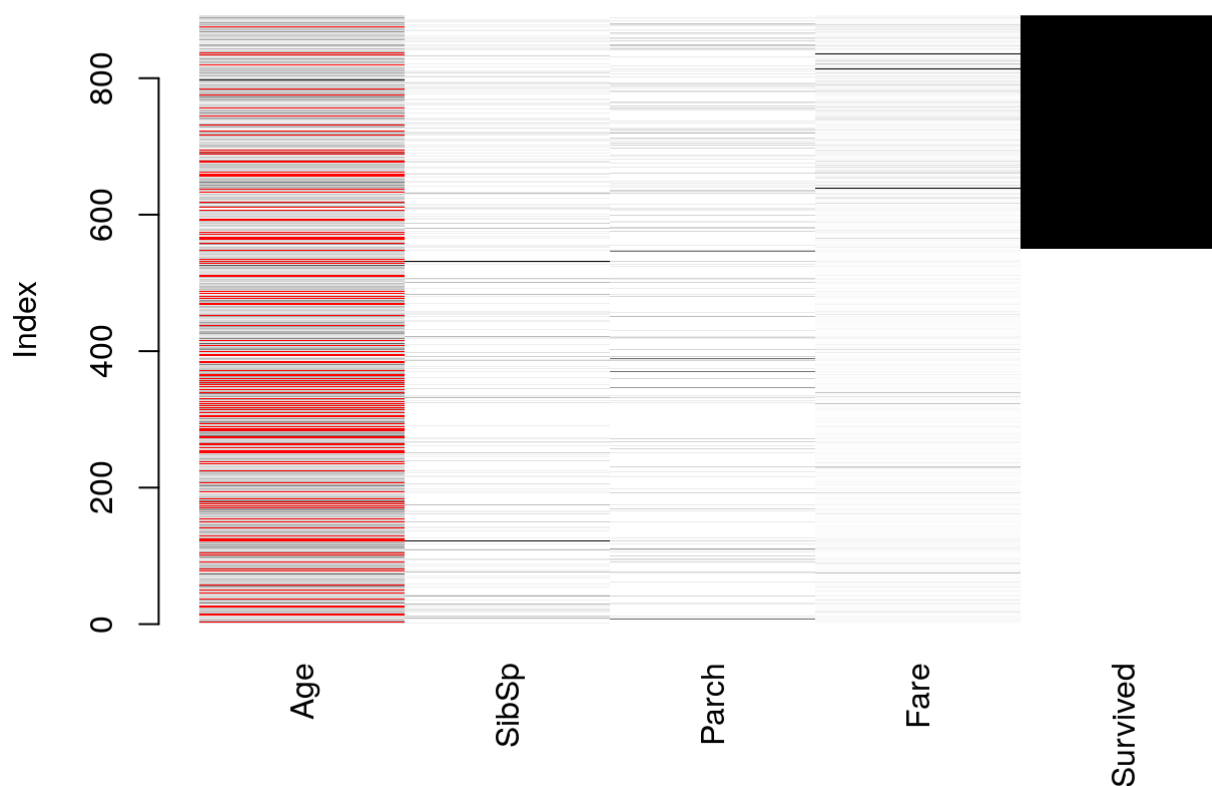
```
datos.na = data.frame(sapply(datos, complete.cases))
sapply(-datos.na+1, sum) #La suma de cada variable nos indicará cuántos valores perdidos tiene cada variable, si es que los tuviese
```

```
## Survived    Pclass      Sex      Age      SibSp      Parch      Fare Embarked
##           0         0         0      177         0         0         0         0
```

Como se observa, el número de valores perdidos para Age es relativamente alto. Una opción sería imputar los valores faltantes en Age con la ayuda de algún algoritmo como k-means; otra opción es eliminar todos los registros con valores perdidos en Age, pues aún eliminando estos registros, tendríamos suficiente muestra para construir nuestro árbol de clasificación.

Optaremos por esta última alternativa, no sin antes verificar que los valores perdidos no siguen algún patrón. Para ello, utilizaremos una matrixplot de variables cuantitativas.

```
datos.num <- datos[sapply(datos, is.numeric)]
datos.num$Survived <- datos$Survived
matrixplot(datos.num, sortby = "Survived")
```



Como podemos observar, ordenando los datos respecto a nuestra variable objetivo, los valores perdidos en Age (representados en color rojo) no siguen ningún patrón. Por tanto, procederemos a eliminar los registros con valores perdidos en Age.

```
datos <- filter(datos, is.na(datos$Age)==F)
```

Ahora, en total, se tienen 714 registros y 8 variables, incluyendo a la variable objetivo: Survived.

Además, debemos verificar que, para el caso de las variables categóricas, todos sus valores están correctamente etiquetados.

```
apply(datos, 2, table)
```

```

## $Survived
##
##    0    1
## 424 290
##
## $Pclass
##
##    1    2    3
## 186 173 355
##
## $Sex
##
## female   male
##    261    453
##
## $Age
##
##  0.42  0.67  0.75  0.83  0.92  1.00  2.00  3.00  4.00  5.00  6.00  7.00
##    1    1    2    2    1    7   10    6   10    4    3    3
##  8.00  9.00 10.00 11.00 12.00 13.00 14.00 14.50 15.00 16.00 17.00 18.00
##    4    8    2    4    1    2    6    1    5   17   13   26
## 19.00 20.00 20.50 21.00 22.00 23.00 23.50 24.00 24.50 25.00 26.00 27.00
##   25   15    1   24   27   15    1   30    1   23   18   18
## 28.00 28.50 29.00 30.00 30.50 31.00 32.00 32.50 33.00 34.00 34.50 35.00
##   25    2   20   25    2   17   18    2   15   15    1   18
## 36.00 36.50 37.00 38.00 39.00 40.00 40.50 41.00 42.00 43.00 44.00 45.00
##   22    1    6   11   14   13    2    6   13    5    9   12
## 45.50 46.00 47.00 48.00 49.00 50.00 51.00 52.00 53.00 54.00 55.00 55.50
##    2    3    9    9    6   10    7    6    1    8    2    1
## 56.00 57.00 58.00 59.00 60.00 61.00 62.00 63.00 64.00 65.00 66.00 70.00
##    4    2    5    2    4    3    4    2    2    3    1    2
## 70.50 71.00 74.00 80.00
##    1    2    1    1
##
## $SibSp
##
##    0    1    2    3    4    5
## 471 183  25  12  18    5
##
## $Parch
##
##    0    1    2    3    4    5    6
## 521 110  68    5    4    5    1
##
## $Fare
##
##  0.0000  4.0125  5.0000  6.2375  6.4375  6.4500  6.4958  6.7500
##    7      1      1      1      1      1      2      2
##  6.9750  7.0458  7.0500  7.0542  7.1250  7.1417  7.2250  7.2292
##    2      1      6      2      4      1      6      8
##  7.2500  7.4958  7.5208  7.5500  7.6500  7.7333  7.7417  7.7500
##   10      3      1      2      4      2      1     14
##  7.7750  7.7958  7.8000  7.8542  7.8750  7.8792  7.8875  7.8958
##   14      6      1     13      1      1      1     23
##  7.9250  8.0292  8.0500  8.1583  8.3000  8.3625  8.4042  8.4333
##   18      1     29      1      1      1      1      1
##  8.5167  8.6542  8.6625  8.6833  8.8500  9.0000  9.2167  9.2250
##    1      1     12      1      1      2      1      2
##  9.3500  9.4750  9.4833  9.5000  9.5875  9.8250  9.8375  9.8417
##    2      1      1      8      2      2      1      1
##  9.8458 10.1708 10.4625 10.5000 10.5167 11.1333 11.2417 11.5000
##    1      1      2     24      1      3      2      4
## 12.0000 12.2750 12.2875 12.3500 12.4750 12.5250 12.6500 12.8750
##    1      1      1      2      4      1      1      1
## 13.0000 13.4167 13.5000 13.7917 13.8583 14.0000 14.1083 14.4000
##   41      1      4      1      1      1      1      2
## 14.4542 14.4583 14.5000 15.0000 15.0458 15.2458 15.5000 15.5500
##    6      1      5      1      1      2      2      1
## 15.7417 15.7500 15.8500 15.9000 16.0000 16.1000 16.7000 17.4000
##    2      1      4      2      1      6      2      1
## 17.8000 18.0000 18.7500 18.7875 19.2583 19.5000 20.2125 20.2500
##    2      3      3      2      4      2      2      2
## 20.5250 20.5750 21.0000 21.0750 22.0250 22.5250 23.0000 24.0000
##    3      2      6      4      1      1      4      2
## 24.1500 25.5875 25.9292 26.0000 26.2500 26.2833 26.2875 26.3875
##    5      1      2     30      6      1      3      1
## 26.5500 27.0000 27.7208 27.7500 27.9000 28.5000 28.7125 29.0000
##   13      2      3      4      6      1      1      2
## 29.1250 29.7000 30.0000 30.0708 30.5000 30.6958 31.0000 31.2750
##    5      2      5      2      4      1      2      7
## 31.3875 32.3208 32.5000 33.0000 33.5000 34.0208 34.3750 34.6542
##    4      1      1      2      1      1      4      1
## 35.5000 36.7500 37.0042 38.5000 39.0000 39.4000 39.6000 39.6875

```

##	3	2	2	1	4	1	1	6
##	40.1250	41.5792	46.9000	47.1000	49.5000	49.5042	50.4958	51.4792
##	1	3	6	1	1	2	1	1
##	51.8625	52.0000	52.5542	53.1000	55.0000	55.4417	55.9000	56.4958
##	1	5	3	5	1	1	2	4
##	56.9292	57.0000	57.9792	59.4000	61.1750	61.3792	61.9792	63.3583
##	2	2	2	1	1	1	1	1
##	65.0000	66.6000	69.3000	71.0000	71.2833	73.5000	75.2500	76.2917
##	2	2	2	2	1	5	1	1
##	76.7292	77.2875	77.9583	78.2667	78.8500	79.2000	79.6500	80.0000
##	3	2	3	2	2	3	3	2
##	81.8583	82.1708	83.1583	83.4750	86.5000	89.1042	90.0000	91.0792
##	1	1	3	2	3	1	4	2
##	93.5000	106.4250	108.9000	110.8833	113.2750	120.0000	133.6500	134.5000
##	2	2	2	3	3	4	1	2
##	135.6333	146.5208	151.5500	153.4625	164.8667	211.3375	211.5000	227.5250
##	3	1	4	3	2	3	1	3
##	247.5208	262.3750	263.0000	512.3292				
##	2	2	4	3				
##								
##	\$Embarked							
##								
##	C	Q	S					
##	2	130	28	554				

Como se observa, para el caso de Embarked, dos de sus valores no están etiquetados. Elimaremos los dos registros correspondientes.

```
datos <- filter(datos, datos$Embarked=="C" | datos$Embarked=="Q" | datos$Embarked=="S")
```

Por tanto, en total, se tienen 712 registros y 8 variables, incluyendo a la variable objetivo: Survived.

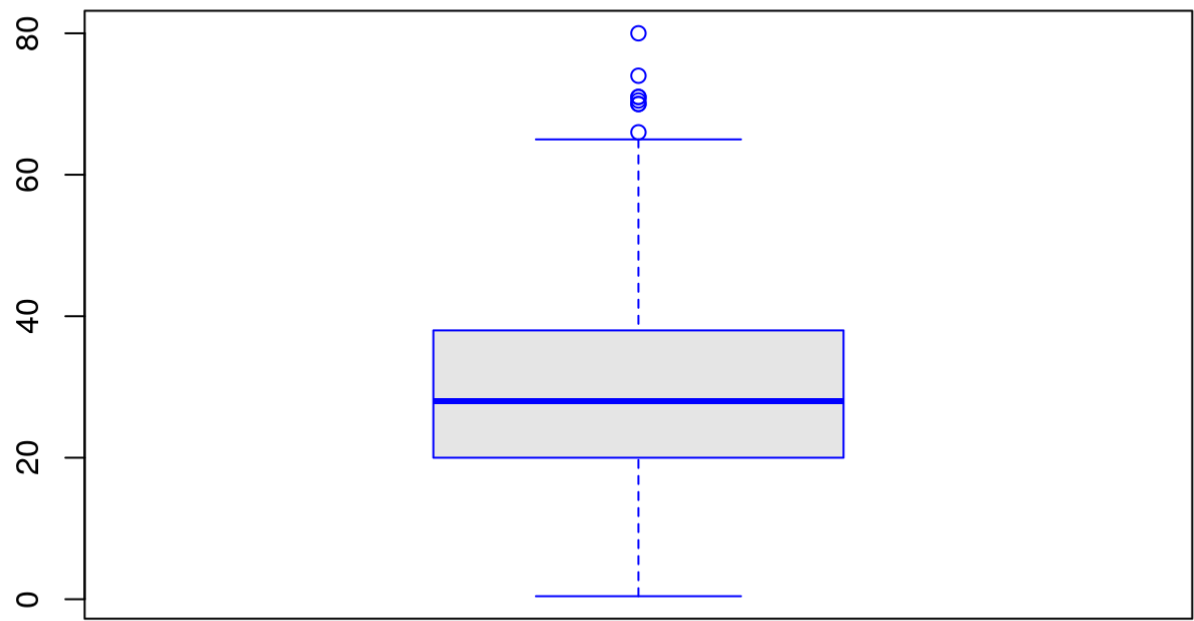
3.2. Identificación y tratamiento de valores extremos.

Para identificar valores extremos realizaremos boxplot de cada variable numérica en el dataset.

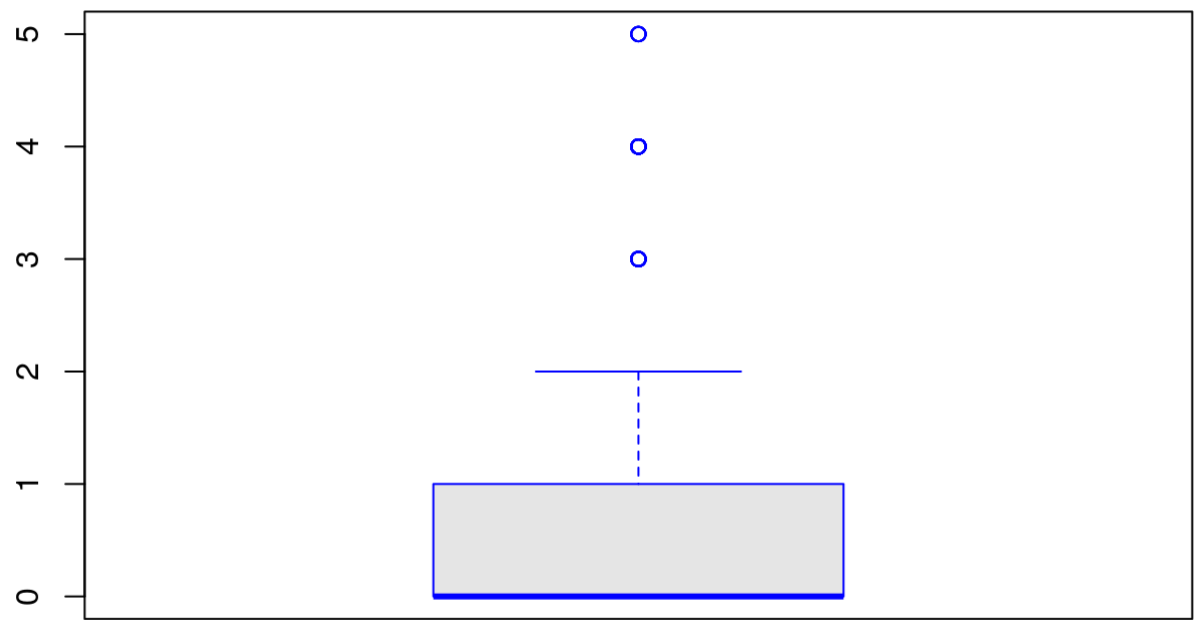
```
datos.num <- datos[sapply(datos, is.numeric)]

for (i in 1:dim(datos.num)[2]){
  boxplot(datos.num[,i], main=names(datos.num)[i], border = "blue", col = "grey90")
}
```

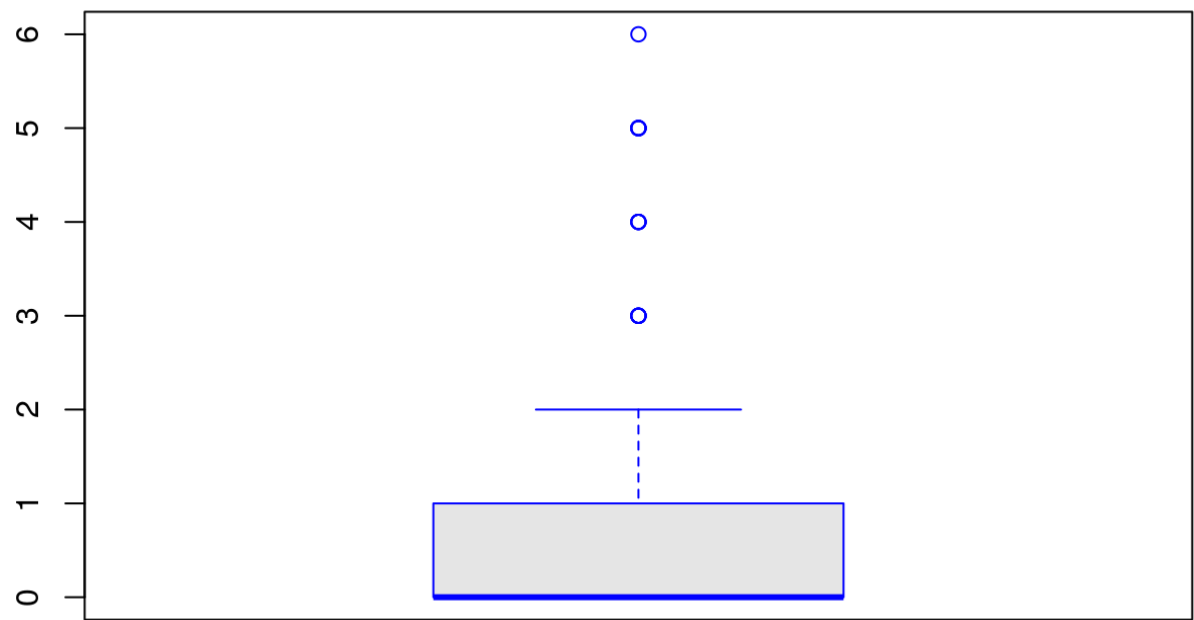
Age

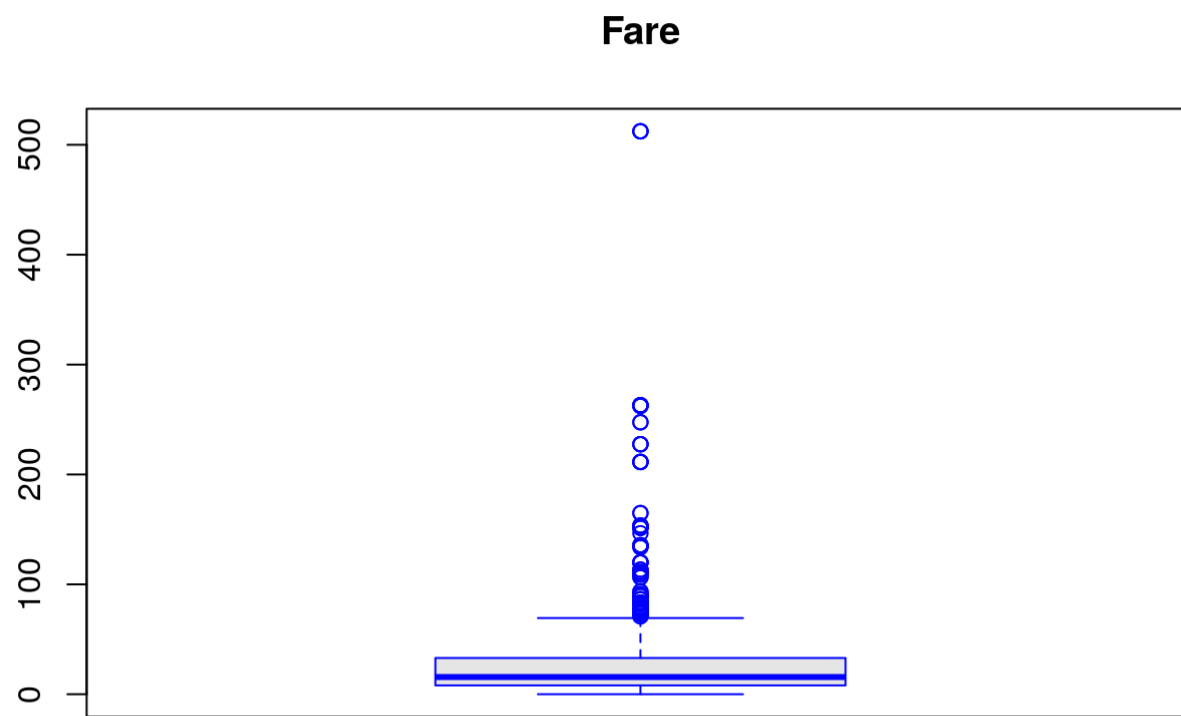


SibSp



Parch





Gráficamente, podemos interpretar que para Age y Fare se observa una gran cantidad de valores extremos; sin embargo, será mejor inspeccionar al detalle cuántos valores extremos tenemos en estas dos variables.

```
#Valores extremos para Age
boxplot.stats(datos$Age)$out
```

```
## [1] 66.0 71.0 70.5 71.0 80.0 70.0 70.0 74.0
```

```
#Valores extremos para Fare
boxplot.stats(datos$Fare)$out
```

```
## [1] 71.2833 263.0000 82.1708 76.7292 83.4750 73.5000 263.0000
## [8] 77.2875 247.5208 73.5000 77.2875 79.2000 146.5208 113.2750
## [15] 76.2917 90.0000 83.4750 90.0000 86.5000 512.3292 79.6500
## [22] 153.4625 135.6333 77.9583 78.8500 91.0792 151.5500 247.5208
## [29] 151.5500 108.9000 83.1583 262.3750 164.8667 134.5000 135.6333
## [36] 153.4625 134.5000 263.0000 75.2500 135.6333 211.5000 227.5250
## [43] 73.5000 120.0000 113.2750 90.0000 120.0000 263.0000 81.8583
## [50] 89.1042 91.0792 90.0000 78.2667 151.5500 86.5000 108.9000
## [57] 93.5000 106.4250 71.0000 106.4250 110.8833 79.6500 110.8833
## [64] 79.6500 79.2000 78.2667 153.4625 77.9583 76.7292 73.5000
## [71] 113.2750 133.6500 73.5000 512.3292 76.7292 211.3375 110.8833
## [78] 227.5250 151.5500 227.5250 211.3375 512.3292 78.8500 262.3750
## [85] 71.0000 86.5000 120.0000 77.9583 211.3375 79.2000 120.0000
## [92] 93.5000 83.1583 164.8667 83.1583
```

Podemos observar que en el caso de Fare es donde se tiene un elevado número de valores extremos. Una solución podría ser recortar nuevamente la muestra, en los registros correspondientes a los valores extremos de Fare; sin embargo, debemos recordar que el problema de los valores extremos se extiende cuando debemos hacer contrastes de hipótesis, correlaciones, regresiones, etc. En nuestro caso, un árbol de clasificación es un método computacional, y si bien es cierto que se rige bajo ciertas normas estadísticas, que no existan valores extremos en las variables no es un requisito para su elaboración. Por tanto, mantendremos el dataset sin ningún cambio.

4. Análisis de los datos

4.1. Selección de los grupos de datos que se quieren analizar/comparar

Seleccionamos los conjuntos de entrenamiento (train) y prueba (test):

```
set.seed(64)
ids <- createDataPartition(datos$Survived,
                           p = 0.7,
                           list = F)

train <- datos[ids,]
test <- datos[-ids,]
```

Por tanto, del dataset original de 712 registros, se han tomado 499 registros para el dataset train y 213 registros para el dataset test.

4.2. Comprobación de la normalidad y homogeneidad de la varianza

No aplica esta comprobación para el caso del árbol de clasificación. Lo importante es tener un dataset balanceado en el target respecto al dataset original, que ya se ha logrado en el apartado anterior. Y se puede corroborar con la siguientes tablas:

```
prop.table(table(datos$Survived))
```

```
##  
##           0           1  
## 0.5955056 0.4044944
```

```
prop.table(table(train$Survived))
```

```
##  
##           0           1  
## 0.5951904 0.4048096
```

4.3. Aplicación de pruebas estadísticas para comparar los grupos de datos

Elaboraremos el árbol de clasificación

```
tree <- C5.0(Survived ~ ., data=train)
```

Mostramos los resultados del entrenamiento

```
summary(tree)
```

```

##
## Call:
## C5.0.formula(formula = Survived ~ ., data = train)
##
##
## C5.0 [Release 2.07 GPL Edition]      Sun Dec 30 10:04:24 2018
## -----
##
## Class specified by attribute `outcome'
##
## Read 499 cases (8 attributes) from undefined.data
##
## Decision tree:
##
## Sex = female:
## :...Pclass in {1,2}: 1 (115/5)
## :   Pclass = 3:
## :     :...Fare > 23: 0 (15/1)
## :       Fare <= 23:
## :         :...Age <= 17: 1 (14/1)
## :           Age > 17:
## :             :...Embarked in {C,Q}: 0 (7/1)
## :               Embarked = S:
## :                 :...SibSp <= 0:
## :                   :...Fare <= 16.1: 1 (20/7)
## :                     :   Fare > 16.1: 0 (2)
## :                       SibSp > 0:
## :                         :...Fare <= 15.0458: 0 (6)
## :                           Fare > 15.0458: 1 (6/2)
## Sex = male:
## :...Pclass in {2,3}:
## :   :...Age > 6: 0 (228/22)
## :     Age <= 6:
## :       :...SibSp <= 2: 1 (9)
## :         SibSp > 2: 0 (4)
## Pclass = 1:
## :...Embarked = Q: 0 (1)
##   Embarked = C:
##     :...SibSp <= 0:
##       :   :...Fare <= 263: 0 (15/5)
##         :     Fare > 263: 1 (2)
##         :   SibSp > 0:
##           :     :...Fare <= 93.5: 1 (6)
##             :       Fare > 93.5: 0 (2)
## Embarked = S:
## :...Age > 53: 0 (9)
##   Age <= 53:
##     :...Parch > 1: 1 (2)
##       Parch <= 1:
##         :...Parch > 0: 0 (4/1)
##           Parch <= 0:
##             :...Age > 47: 1 (5)
##               Age <= 47:
##                 :...Age <= 27: 1 (2)
##                   Age > 27: 0 (25/6)
##
##
## Evaluation on training data (499 cases):
##
##      Decision Tree
##      -----
##      Size      Errors
##
##      22   51(10.2%)   <<
##
##
##      (a)   (b)   <-classified as
##      ----   ----
##      282   15   (a): class 0
##      36   166  (b): class 1
##
##
## Attribute usage:
##
## 100.00% Pclass
## 100.00% Sex
## 68.74% Age
## 22.85% Embarked
## 19.04% Fare
## 14.43% SibSp
## 7.62% Parch
##

```



```
##  
## Time: 0.0 secs
```

Aplicamos el árbol de clasificación al test

```
newdata <- test[,colnames(test)!="Survived"]  
prediccion <- predict(tree, newdata=newdata, type="class")
```

Mostramos los resultados de aplicar el árbol de clasificación al test

```
confusionMatrix(data=prediccion, reference=test$Survived, positive="1")
```

```
## Confusion Matrix and Statistics  
##  
##           Reference  
## Prediction    0    1  
##           0 106  33  
##           1  21  53  
##  
##           Accuracy : 0.7465  
##           95% CI : (0.6825, 0.8034)  
##    No Information Rate : 0.5962  
##    P-Value [Acc > NIR] : 3.047e-06  
##  
##           Kappa : 0.4613  
##  McNemar's Test P-Value : 0.1344  
##  
##           Sensitivity : 0.6163  
##           Specificity : 0.8346  
##    Pos Pred Value : 0.7162  
##    Neg Pred Value : 0.7626  
##    Prevalence : 0.4038  
##    Detection Rate : 0.2488  
##    Detection Prevalence : 0.3474  
##    Balanced Accuracy : 0.7255  
##  
##    'Positive' Class : 1  
##
```

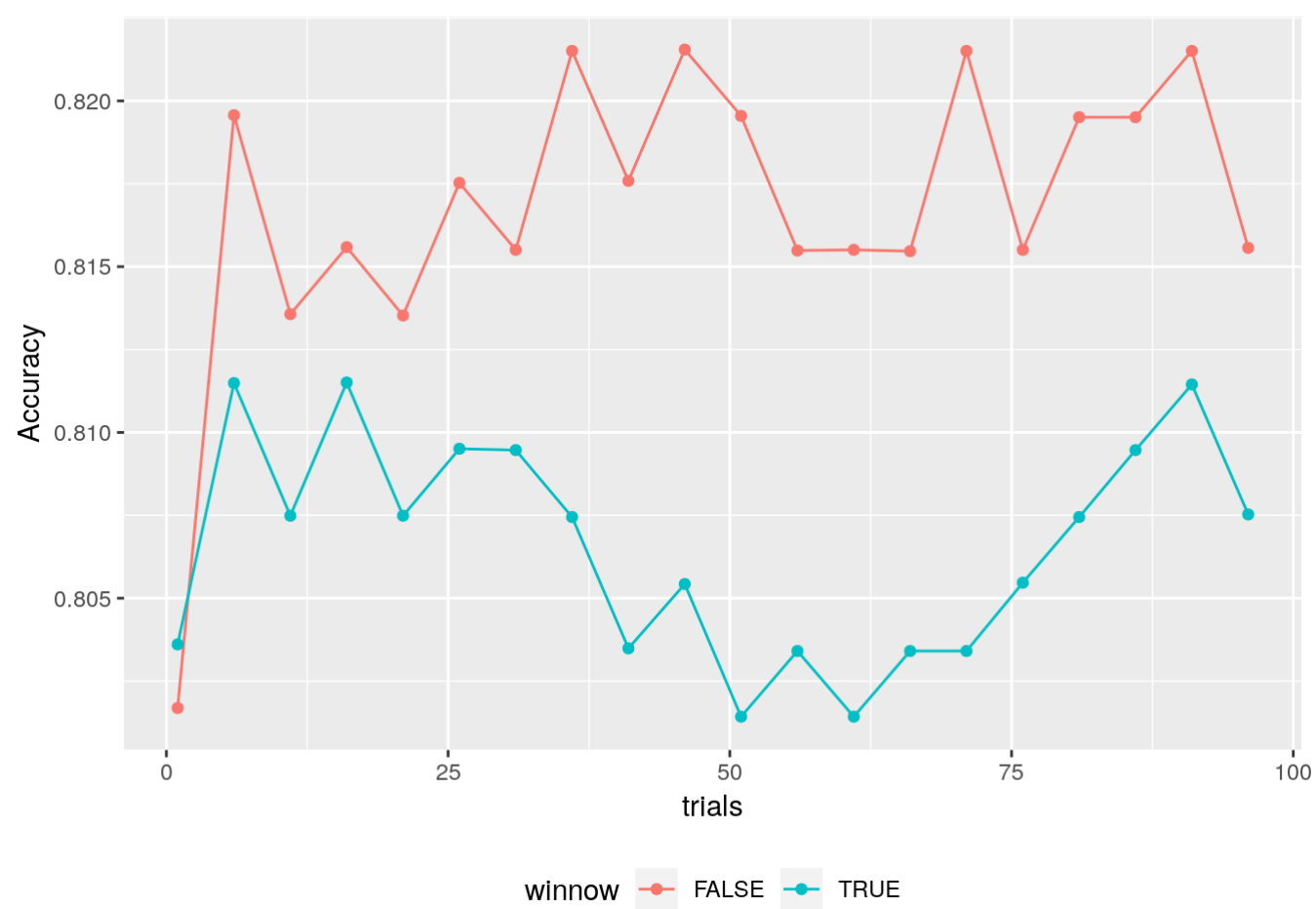
A fin de mejorar la predicción, podemos incorporar un bosque, sin embargo, a fin de elegir el mejor número de árboles para nuestro bosque, incorporaremos un método de validación cruzada para elegir el número de árboles más adecuado.

```
ctrl <- trainControl(method = "cv", number = 5)  
  
grid <- expand.grid(model = "tree",  
                   trials = seq(1,100,5),  
                   winnow = c(FALSE, TRUE))  
  
set.seed(64)  
modelos <- train(Survived ~ ., data = train,  
                 method = "C5.0",  
                 trControl = ctrl,  
                 tuneGrid = grid)  
  
modelos$bestTune
```

```
##   trials model winnow  
## 10     46  tree FALSE
```

Podemos representar gráficamente los resultados

```
ggplot(data = modelos$results, aes(x = trials, y = Accuracy, color = winnow)) +  
  geom_line() +  
  geom_point() +  
  theme_gray() +  
  theme(legend.position = "bottom")
```



Aplicaremos entonces un bosque con 46 árboles.

```
bosque <- C5.0(Survived ~ ., data=train, trials=46)
```

Mostramos los resultados del entrenamiento

```
summary(bosque)
```

```

##
## Call:
## C5.0.formula(formula = Survived ~ ., data = train, trials = 46)
##
##
## C5.0 [Release 2.07 GPL Edition]      Sun Dec 30 10:04:27 2018
## -----
##
## Class specified by attribute `outcome'
##
## Read 499 cases (8 attributes) from undefined.data
##
## ----- Trial 0: -----
##
## Decision tree:
##
## Sex = female:
## :...Pclass in {1,2}: 1 (115/5)
## :   Pclass = 3:
## :     :...Fare > 23: 0 (15/1)
## :       Fare <= 23:
## :         :...Age <= 17: 1 (14/1)
## :           Age > 17:
## :             :...Embarked in {C,Q}: 0 (7/1)
## :               Embarked = S:
## :                 :...SibSp <= 0:
## :                   :...Fare <= 16.1: 1 (20/7)
## :                     :   Fare > 16.1: 0 (2)
## :                       SibSp > 0:
## :                         :...Fare <= 15.0458: 0 (6)
## :                           Fare > 15.0458: 1 (6/2)
## Sex = male:
## :...Pclass in {2,3}:
## :   :...Age > 6: 0 (228/22)
## :     Age <= 6:
## :       :...SibSp <= 2: 1 (9)
## :         SibSp > 2: 0 (4)
## Pclass = 1:
## :...Embarked = Q: 0 (1)
##   Embarked = C:
##     :...SibSp <= 0:
##       :   :...Fare <= 263: 0 (15/5)
##         :     Fare > 263: 1 (2)
##       :   SibSp > 0:
##         :     :...Fare <= 93.5: 1 (6)
##           :       Fare > 93.5: 0 (2)
## Embarked = S:
## :...Age > 53: 0 (9)
##   Age <= 53:
##     :...Parch > 1: 1 (2)
##       Parch <= 1:
##         :...Parch > 0: 0 (4/1)
##           Parch <= 0:
##             :...Age > 47: 1 (5)
##               Age <= 47:
##                 :...Age <= 27: 1 (2)
##                   Age > 27: 0 (25/6)
##
## ----- Trial 1: -----
##
## Decision tree:
##
## Pclass = 1:
## :...Sex = female:
## :   :...Age <= 9: 0 (2.9)
## :     Age > 9: 1 (50.4/2.9)
## :   Sex = male:
## :     :...Age <= 57: 1 (74.3/26.5)
## :       Age > 57: 0 (8.6/0.8)
## Pclass in {2,3}:
## :...Sex = female:
## :   :...Pclass = 2:
## :     :...Age <= 56: 1 (44/5.9)
## :       Age > 56: 0 (2.9)
## :     Pclass = 3:
## :       :...Embarked in {S,Q}: 0 (72.1/22.4)
## :         Embarked = C: 1 (8.4/1.6)
## Sex = male:
## :...SibSp > 1: 0 (14.8/0.8)
##   SibSp <= 1:
##     :...Embarked = Q: 0 (10.7/2.9)
##       Embarked = C:
##         :...Pclass = 2: 0 (5.4)

```

```

##          : Pclass = 3: 1 (26.4/10.9)
##          Embarked = S:
##          :...SibSp <= 0:
##              :...Fare <= 40.125: 0 (139.1/33.2)
##              : Fare > 40.125: 1 (7.4/1.6)
##              SibSp > 0:
##              :...Age <= 32.5: 1 (22.1/8.6)
##              Age > 32.5: 0 (9.3)
##
## ----- Trial 2: -----
##
## Decision tree:
##
## Sex = female: 1 (174.4/59.5)
## Sex = male:
## :...SibSp > 1: 0 (15/2)
##     SibSp <= 1:
##         :...Pclass = 2: 0 (63.5/18.6)
##         Pclass = 3:
##             :...Fare <= 41.5792: 0 (147.2/49.5)
##             : Fare > 41.5792: 1 (6.3/1.4)
##             Pclass = 1:
##                 :...Age > 53: 0 (13.1/3.8)
##                 Age <= 53:
##                     :...Fare <= 25.9292: 0 (5.6)
##                     Fare > 25.9292:
##                         :...Fare <= 27: 1 (12.4/1.4)
##                         Fare > 27: 0 (61.7/25.4)
##
## ----- Trial 3: -----
##
## Decision tree:
##
## Embarked = C:
## :...Sex = female: 1 (26/3.7)
## : Sex = male:
## : :...Pclass = 1: 1 (36.6/12.6)
## : Pclass = 2: 0 (3.9)
## : Pclass = 3:
## : :...Age <= 29: 1 (25.2/9.7)
## : Age > 29: 0 (7.3)
## Embarked in {S,Q}:
## :...Pclass in {1,2}:
##     :...Sex = male:
##     : :...Age <= 13: 1 (7.3)
##     : : Age > 13: 0 (114.3/42.7)
##     : Sex = female:
##     : :...SibSp <= 0: 1 (27.5/2.9)
##     : SibSp > 0:
##     : :...SibSp > 1: 1 (2.8)
##     : SibSp <= 1:
##     : :...Fare <= 135.6333: 1 (24.3/8)
##     : Fare > 135.6333: 0 (7.5/0.6)
## Pclass = 3:
## :...SibSp > 1: 0 (22.5/2.4)
##     SibSp <= 1:
##         :...Age <= 8: 1 (10)
##         Age > 8:
##             :...SibSp > 0: 0 (39.4/12.5)
##             SibSp <= 0:
##                 :...Parch > 1: 0 (7.3/1.2)
##                 Parch <= 1:
##                     :...Embarked = Q: 0 (15.4/7.6)
##                     Embarked = S:
##                         :...Sex = female:
##                         :...Parch > 0: 1 (2.4)
##                         : Parch <= 0:
##                         : :...Fare <= 7.65: 1 (3.6)
##                         : Fare > 7.65: 0 (31.6/8.5)
##                         Sex = male:
##                         :...Age <= 23.5: 0 (26.4/4)
##                         Age > 23.5:
##                             :...Age > 32.5: 0 (18.6/4)
##                             Age <= 32.5:
##                                 :...Age <= 30.5: 0 (21.6/8)
##                                 Age > 30.5: 1 (17.4/1.1)
##
## ----- Trial 4: -----
##
## Decision tree:
##
## Sex = female: 1 (175.4/72.7)
## Sex = male:

```

```

## :...Pclass = 3: 0 (152.7/62.7)
##   Pclass = 2:
##     :...Age <= 9: 1 (4.3)
##     :   Age > 9: 0 (57.3/23.2)
##     Pclass = 1:
##       :...Embarked = Q: 0 (1)
##       Embarked = S:
##         :...Age <= 53: 1 (60.3/22.9)
##         :   Age > 53: 0 (5.3)
##         Embarked = C:
##           :...Parch > 1: 0 (3.5)
##           Parch <= 1:
##             :...Fare <= 29.7: 0 (5.5)
##             Fare > 29.7: 1 (33.8/13.7)
##
## ----- Trial 5: -----
##
## Decision tree:
##
## Pclass = 2:
## :...Age <= 20: 1 (15.5/2.1)
## :   Age > 20:
## :     :...Sex = female: 1 (30.9/14.3)
## :     Sex = male:
## :       :...Age <= 31: 0 (14)
## :       Age > 31: 1 (34.4/15)
## Pclass = 1:
## :...Sex = female:
## :   :...Parch <= 1: 1 (23.5)
## :   :   Parch > 1: 0 (9.8/3.4)
## :   Sex = male:
## :     :...Embarked = Q: 0 (0.9)
## :     Embarked = C:
## :       :...Fare <= 29.7: 0 (5)
## :       :   Fare > 29.7:
## :         :   :...Fare <= 37.0042: 1 (6.6)
## :         :   Fare > 37.0042: 0 (30.5/11.7)
## :         Embarked = S:
## :           :...Fare <= 26: 0 (5.6)
## :           Fare > 26:
## :             :...Fare <= 26.2875: 1 (8.1)
## :             Fare > 26.2875:
## :               :...SibSp <= 0: 0 (30.5/12.2)
## :               SibSp > 0: 1 (23.7/9.9)
## Pclass = 3:
## :...SibSp > 1: 0 (23.2/4.6)
##   SibSp <= 1:
##     :...Age <= 10: 1 (11.9/1.2)
##     Age > 10:
##       :...Parch > 0: 0 (42.2/13.5)
##       Parch <= 0:
##         :...Embarked = C: 0 (28.4/11.6)
##         Embarked = Q: 1 (17.7/7)
##         Embarked = S:
##           :...Sex = female: 0 (48.7/19.6)
##           Sex = male:
##             :...SibSp > 0: 1 (12.1/4.1)
##             SibSp <= 0:
##               :...Age <= 23.5: 0 (22.2/5.1)
##               Age > 23.5: 1 (53.6/23.2)
##
## ----- Trial 6: -----
##
## Decision tree:
##
## Sex = female: 1 (172.4/77.6)
## Sex = male:
## :...Pclass = 2:
## :   :...Age <= 9: 1 (3.4)
## :   :   Age > 9: 0 (61.1/21.3)
## :   Pclass = 1:
## :     :...Age > 53: 0 (12.1/3.6)
## :     :   Age <= 53:
## :       :   :...Fare <= 25.9292: 0 (4.5)
## :       :   Fare > 25.9292:
## :         :   :...Fare <= 27: 1 (14.7/1.6)
## :         :   Fare > 27:
## :           :   :...Age <= 28.5: 1 (25.6/7.9)
## :           :   Age > 28.5: 0 (50.4/18.4)
## :   Pclass = 3:
## :     :...Age > 32.5: 0 (30.4/4.4)
## :     Age <= 32.5:
## :       :...SibSp > 0: 0 (32.6/13)

```

```

##          SibSp <= 0:
##          :...Embarked in {C,Q}: 1 (28.9/11.6)
##          Embarked = S:
##          :...Age > 30.5: 1 (16.1/3.1)
##          Age <= 30.5:
##          :...Fare <= 7.1417: 1 (10.7/1.8)
##          Fare > 7.1417: 0 (36.2/7)
##
## ----- Trial 7: -----
##
## Decision tree:
##
## Pclass = 1:
## :...Sex = female: 1 (29.1/6.3)
## :   Sex = male:
## :     :...SibSp > 0: 1 (37.6/15.7)
## :       SibSp <= 0:
## :         :...Fare > 35.5: 0 (27.2/7.4)
## :           Fare <= 35.5:
## :             :...Fare <= 26: 0 (4.3)
## :               Fare > 26: 1 (37.1/11.8)
## Pclass in {2,3}:
## :...Age <= 6: 1 (19.8/4.2)
##   Age > 6:
##     :...SibSp > 1: 0 (20.8/4.4)
##       SibSp <= 1:
##         :...Embarked = Q: 0 (25.1/8.2)
##           Embarked = C:
##             :...Age > 29: 0 (10.1)
##               Age <= 29:
##                 :...Age <= 19: 0 (9.4/1.5)
##                 Age > 19: 1 (26.7/10)
##           Embarked = S:
##             :...SibSp > 0: 0 (68.1/25.3)
##               SibSp <= 0:
##                 :...Pclass = 2: 0 (56.3/27.6)
##                   Pclass = 3:
##                     :...Sex = male: 0 (78.8/30)
##                       Sex = female:
##                         :...Fare > 16.1: 0 (6.2)
##                           Fare <= 16.1:
##                             :...Fare <= 7.65: 1 (3.1)
##                             Fare > 7.65:
##                               :...Age <= 25: 0 (21.4/4.8)
##                               Age > 25: 1 (17.9/5.5)
##
## ----- Trial 8: -----
##
## Decision tree:
##
## Pclass = 2:
## :...Sex = male:
## :   :...Fare <= 13: 1 (37.5/16.6)
## :     Fare > 13: 0 (22/6.9)
## :   Sex = female:
## :     :...Parch > 1: 1 (6.4)
## :       Parch <= 1:
## :         :...Age <= 43: 1 (26.1/5.1)
## :           Age > 43: 0 (12.6/3.4)
## Pclass = 1:
## :...Sex = female: 1 (26.9/7.1)
## :   Sex = male:
## :     :...Embarked = Q: 0 (1.1)
## :       Embarked = C:
## :         :...Parch > 1: 0 (2.6)
## :           Parch <= 1:
## :             :...Fare <= 29.7: 0 (4.9)
## :               Fare > 29.7: 1 (34.7/12.8)
## :           Embarked = S:
## :             :...Age > 53: 0 (6.7)
## :               Age <= 53:
## :                 :...Age > 47: 1 (9.6/1.7)
## :                   Age <= 47:
## :                     :...Fare > 57.9792: 0 (8.5/1.1)
## :                       Fare <= 57.9792:
## :                         :...Age <= 45: 1 (34.1/15)
## :                           Age > 45: 0 (5.3)
## Pclass = 3:
## :...SibSp > 1: 0 (25.1/5.6)
##   SibSp <= 1:
##     :...SibSp > 0:
##       :...Age <= 14: 1 (6)
##       Age > 14: 0 (52.1/22.9)

```

```

##      SibSp <= 0:
##      :...Age > 32.5: 0 (34.1/6.9)
##      Age <= 32.5:
##      :...Parch > 0: 1 (15/5.5)
##      Parch <= 0:
##      :...Age <= 17: 0 (11.9/5.4)
##      Age > 17:
##      :...Age <= 30.5: 0 (92.5/40.2)
##      Age > 30.5: 1 (23.4/8.5)
##
## ----- Trial 9: -----
##
## Decision tree:
##
## Sex = female:
## :...Pclass in {1,2}: 1 (67.8/21.2)
## :   Pclass = 3:
## :     :...Fare > 20.575: 0 (23.7/4.7)
## :     Fare <= 20.575:
## :       :...SibSp > 2: 1 (5.2)
## :       SibSp <= 2:
## :         :...Age <= 17: 1 (14.4/2.1)
## :         Age > 17: 0 (70.4/31.1)
## Sex = male:
## :...Pclass = 3: 0 (140.6/66.4)
##   Pclass = 2:
##   :...Age <= 9: 1 (4.1)
##   :   Age > 9: 0 (67.2/22.9)
##   Pclass = 1:
##   :...Age > 53: 0 (12.6/3.4)
##   Age <= 53:
##   :...Parch > 0: 1 (15.8/7.5)
##   Parch <= 0:
##   :...Age <= 47: 0 (64.3/22.5)
##   Age > 47: 1 (12.9/2.5)
##
## ----- Trial 10: -----
##
## Decision tree:
##
## Parch > 1:
## :...Fare <= 19.5: 1 (7.8/0.2)
## :   Fare > 19.5: 0 (49.6/15.4)
## Parch <= 1:
## :...Sex = female: 1 (133.4/51.5)
##   Sex = male:
##   :...Pclass = 2:
##   :     :...Age <= 9: 1 (3.7)
##   :     :   Age > 9: 0 (63.5/25.1)
##   :     Pclass = 3:
##   :       :...Age <= 29: 1 (93.4/38.7)
##   :       :   Age > 29: 0 (44.7/19.8)
##   :       Pclass = 1:
##   :         :...Embarked = Q: 0 (0.8)
##   :         Embarked = C:
##   :           :...Parch > 0: 1 (7.9/2.4)
##   :           :   Parch <= 0:
##   :             :   :...Age <= 24.5: 0 (5.1)
##   :             :   Age > 24.5: 1 (27.9/11.6)
##   :         Embarked = S:
##   :           :...Fare > 57.9792: 0 (11.5/1.3)
##   :           Fare <= 57.9792:
##   :             :...Fare <= 26: 0 (4.2)
##   :             Fare > 26:
##   :               :...Parch > 0: 1 (2.3)
##   :               Parch <= 0:
##   :                 :...Fare <= 26.2875: 1 (6)
##   :                 Fare > 26.2875:
##   :                   :...Age <= 27: 1 (3.3)
##   :                   Age > 27: 0 (33.9/13.2)
##
## ----- Trial 11: -----
##
## Decision tree:
##
## Sex = male:
## :...Embarked in {C,Q}: 0 (90.4/39.3)
## :   Embarked = S:
## :     :...SibSp <= 0: 0 (168.1/76)
## :     SibSp > 0: 1 (63.9/30)
## Sex = female:
## :...Embarked = C: 1 (17.8/5.1)
##   Embarked = Q: 0 (11.8/5.2)

```

```

##      Embarked = S:
##      :...Age <= 16: 0 (19.4/4.6)
##      Age > 16:
##      :...Fare <= 7.65: 1 (4.6)
##      Fare > 7.65:
##      :...Parch > 0: 1 (48.3/19.4)
##      Parch <= 0:
##      :...Fare <= 11.2417: 0 (41.6/15.9)
##      Fare > 11.2417: 1 (33.2/9.3)
##
## ----- Trial 12: -----
##
## Decision tree:
##
## Sex = female:
## :...SibSp <= 0: 1 (93.7/37.3)
## :   SibSp > 0:
## :     :...Embarked = C: 1 (7)
## :     Embarked = Q: 0 (1.1)
## :     Embarked = S:
## :       :...Age <= 30: 0 (43.8/11.7)
## :       Age > 30: 1 (28.7/9.6)
## Sex = male:
## :...Embarked in {C,Q}: 0 (90/42.8)
## Embarked = S:
## :...Fare > 61.3792: 0 (15.3/3.4)
## Fare <= 61.3792:
## :...SibSp <= 0:
## :   :...Fare <= 13: 0 (119.6/56.5)
## :   Fare > 13: 1 (42.7/19.2)
## :   SibSp > 0:
## :     :...Age <= 10: 1 (11.1/3.9)
## :     Age > 10:
## :       :...Fare <= 9.5875: 1 (11.3/3)
## :       Fare > 9.5875: 0 (34.6/12.4)
##
## ----- Trial 13: -----
##
## Decision tree:
##
## SibSp > 1:
## :...Fare <= 110.8833: 0 (28.1/7.5)
## :   Fare > 110.8833: 1 (3.8)
## SibSp <= 1:
## :...Embarked = Q: 0 (23/10.1)
## Embarked = C:
## :...Pclass = 2: 0 (4.1/1.7)
## :   Pclass = 1:
## :     :...Fare <= 29.7: 0 (4/0.2)
## :     Fare > 29.7: 1 (43.9/15.3)
## :     Pclass = 3:
## :       :...Age <= 29: 1 (35.3/14.8)
## :       Age > 29: 0 (5.3)
## Embarked = S:
## :...Age <= 8: 1 (17.6/3)
## Age > 8:
## :...SibSp > 0:
## :   :...Age > 38: 0 (28.4/8.7)
## :   Age <= 38:
## :     :...Parch > 4: 1 (4.1)
## :     Parch <= 4:
## :       :...Parch <= 0: 1 (42.9/18.7)
## :       Parch > 0: 0 (25.3/8.2)
## SibSp <= 0:
## :...Sex = female:
## :   :...Pclass in {1,2}: 1 (20.8/4.1)
## :   Pclass = 3:
## :     :...Fare <= 16.1: 1 (42.2/19.4)
## :     Fare > 16.1: 0 (5.9)
## Sex = male:
## :...Pclass = 2:
## :   :...Fare <= 13: 1 (41.1/18.7)
## :   Fare > 13: 0 (6.2)
## Pclass = 1:
## :...Fare > 35.5: 0 (6.8)
## :   Fare <= 35.5:
## :     :...Age <= 54: 1 (27.5/10.5)
## :     Age > 54: 0 (3.8)
## Pclass = 3:
## :...Age <= 30.5:
## :   :...Fare <= 7.1417: 1 (12.6/2.8)
## :   Fare > 7.1417: 0 (32.2/6.6)
## Age > 30.5:

```



```

##                :...Fare <= 7.7958: 0 (3.1)
##                Fare > 7.7958: 1 (31.2/8.6)
##
## ----- Trial 14: -----
##
## Decision tree:
##
## Sex = female:
## :...Pclass in {1,2}: 1 (64.7/20.9)
## :   Pclass = 3: 0 (107.5/48.2)
## Sex = male:
## :...Pclass in {2,3}: 0 (221.1/93.8)
##   Pclass = 1:
##     :...Embarked in {S,Q}: 0 (63.5/27.6)
##     Embarked = C:
##       :...Fare > 263: 1 (3.4)
##       Fare <= 263:
##         :...Fare <= 93.5: 1 (30.9/13.3)
##         Fare > 93.5: 0 (8)
##
## ----- Trial 15: -----
##
## Decision tree:
##
## Sex = female: 1 (173.3/76.8)
## Sex = male:
## :...Pclass = 1:
##   :...SibSp <= 0: 0 (67.5/29.2)
##   :   SibSp > 0: 1 (39.2/17.5)
##   Pclass = 2:
##     :...Age <= 9: 1 (3.8)
##     :   Age > 9: 0 (66.5/25.1)
##     Pclass = 3:
##       :...SibSp > 1: 0 (7.7)
##       SibSp <= 1:
##         :...Age <= 8: 1 (7.6)
##         Age > 8:
##           :...Age <= 18: 0 (9.2)
##           Age > 18:
##             :...Embarked in {C,Q}: 1 (36/15.3)
##             Embarked = S:
##               :...Fare > 8.05: 0 (19.9/4.5)
##               Fare <= 8.05:
##                 :...Fare <= 7.8958: 0 (36.3/16.1)
##                 Fare > 7.8958: 1 (31.9/8.1)
##
## ----- Trial 16: -----
##
## Decision tree:
##
## Pclass = 2:
## :...Parch <= 1: 0 (98.6/46.4)
## :   Parch > 1: 1 (7.6/1)
## Pclass = 1:
## :...Sex = female: 1 (21.1/7.7)
## :   Sex = male:
## :     :...Embarked = Q: 0 (0.7)
## :     Embarked = C:
## :       :...Fare > 263: 1 (3.6)
## :       :   Fare <= 263:
## :         :   :...Fare <= 93.5: 1 (30.7/13.9)
## :         :   Fare > 93.5: 0 (7.6)
## :         Embarked = S:
## :           :...Age > 53: 0 (4.9)
## :           Age <= 53:
## :             :...Parch > 0: 1 (8.6/3.6)
## :             Parch <= 0:
## :               :...Age <= 47: 0 (42.8/18.5)
## :               Age > 47: 1 (8.7)
## Pclass = 3:
## :...SibSp > 1: 0 (22.8/6.6)
##   SibSp <= 1:
##     :...Age <= 10: 1 (12.4/1.6)
##     Age > 10:
##       :...Sex = female: 0 (93.1/40.4)
##       Sex = male:
##         :...Age <= 18: 0 (8.4)
##         Age > 18:
##           :...Age > 32.5: 0 (24.3/5.4)
##           Age <= 32.5:
##             :...Embarked in {C,Q}: 1 (34.9/16.1)
##             Embarked = S:
##               :...Age > 30.5: 1 (18.8/3.7)

```

```

##                               Age <= 30.5:
##                               :...Age > 25: 0 (8.8)
##                               Age <= 25:
##                               :...SibSp > 0: 1 (8.1/1.1)
##                               SibSp <= 0:
##                               :...Fare <= 7.1417: 1 (12/2.5)
##                               Fare > 7.1417: 0 (20.5/6.5)
##
## ----- Trial 17: -----
##
## Decision tree:
##
## Sex = male:
## :...Embarked in {C,Q}: 0 (88.8/35.9)
## :   Embarked = S:
## :     :...Pclass = 1:
## :       :...Age <= 53: 1 (59.2/26)
## :       :   Age > 53: 0 (4.4)
## :       Pclass = 2:
## :       :...Age <= 9: 1 (4.7)
## :       :   Age > 9: 0 (59.8/27.8)
## :       Pclass = 3:
## :       :...SibSp > 1: 0 (5.7)
## :       :   SibSp <= 1:
## :       :     :...Age <= 11: 1 (4.9)
## :       :     Age > 11:
## :       :       :...Fare <= 40.125: 0 (82.4/36.1)
## :       :       Fare > 40.125: 1 (5.3/0.9)
## Sex = female:
## :...Pclass = 1: 1 (20.2/8.3)
##   Pclass = 2:
##   :...Age <= 56: 1 (43.8/9.5)
##   :   Age > 56: 0 (4.3)
##   Pclass = 3:
##   :...Fare > 23: 0 (17.4/3.7)
##   :   Fare <= 23:
##   :     :...SibSp > 2: 1 (5.1)
##   :     :   SibSp <= 2:
##   :     :     :...Age <= 17: 1 (18.3/2.7)
##   :     :     Age > 17:
##   :     :       :...Embarked in {C,Q}: 0 (11.6/1.5)
##   :     :       Embarked = S:
##   :     :         :...SibSp <= 0: 1 (44.1/19.2)
##   :     :         SibSp > 0: 0 (19/7.9)
##
## ----- Trial 18: -----
##
## Decision tree:
##
## Pclass = 2: 1 (109.6/44.8)
## Pclass = 1:
## :...Embarked = Q: 0 (0.6)
## :   Embarked = C:
## :     :...Sex = female: 1 (3)
## :     :   Sex = male:
## :     :     :...Fare <= 29.7: 0 (4.1)
## :     :     :   Fare > 29.7: 1 (40.5/16.5)
## :     Embarked = S:
## :       :...SibSp > 0:
## :       :     :...SibSp > 1: 1 (2.4)
## :       :     :   SibSp <= 1:
## :       :     :     :...Fare <= 61.3792: 1 (15.7/6.5)
## :       :     :     :   Fare > 61.3792: 0 (23.7/4.8)
## :       :     SibSp <= 0:
## :       :       :...Sex = female: 1 (2.6)
## :       :       :   Sex = male:
## :       :       :     :...Fare > 35.5: 0 (7.9)
## :       :       :     :   Fare <= 35.5:
## :       :       :       :...Fare <= 26: 0 (5.3)
## :       :       :       :   Fare > 26: 1 (28.8/10.8)
## Pclass = 3:
## :...Embarked = Q: 0 (23.8/11.2)
##   Embarked = C:
##   :...Fare <= 15.0458: 0 (32.5/12.4)
##   :   Fare > 15.0458: 1 (11.6/1.1)
##   Embarked = S:
##   :...SibSp > 1: 0 (16.9/4.6)
##   :   SibSp <= 1:
##   :     :...SibSp > 0:
##   :     :       :...Age <= 38: 1 (39.4/14.5)
##   :     :       :   Age > 38: 0 (4.1)
##   :     SibSp <= 0:
##   :       :...Parch > 3: 0 (2.8)

```

```

##          Parch <= 3:
##          :...Fare <= 7.65: 1 (20.8/7.3)
##          Fare > 7.65:
##          :...Age <= 25: 0 (44.7/16.2)
##          Age > 25: 1 (58.2/25.8)
##
## ----- Trial 19: -----
##
## Decision tree:
##
## Sex = female: 1 (170.3/77.4)
## Sex = male:
## :...Embarked = Q: 0 (12.6/5)
##   Embarked = C:
##   :...Pclass = 2: 0 (3)
##   :   Pclass in {1,3}:
##   :   :...SibSp > 0: 1 (19.9/9.5)
##   :   :   SibSp <= 0:
##   :   :   :...Age <= 35: 1 (38.2/14.8)
##   :   :   :   Age > 35: 0 (16.2/3.7)
##   Embarked = S:
##   :...Age <= 11: 1 (13.2/2.8)
##   :   Age > 11:
##   :   :...Pclass = 3: 0 (95.4/40.2)
##   :   :   Pclass = 1:
##   :   :   :...Fare <= 26: 0 (4.7)
##   :   :   :   Fare > 26:
##   :   :   :   :...Fare <= 26.2875: 1 (5.6)
##   :   :   :   :   Fare > 26.2875: 0 (53.8/18.8)
##   :   :   Pclass = 2:
##   :   :   :...Fare > 13: 0 (17.4/5)
##   :   :   :   Fare <= 13:
##   :   :   :   :...Age <= 33: 0 (22.3/5.4)
##   :   :   :   :   Age > 33: 1 (26.4/10.1)
##
## ----- Trial 20: -----
##
## Decision tree:
##
## Parch > 1:
## :...Fare <= 19.5: 1 (6.7/0.7)
## :   Fare > 19.5: 0 (49.2/13.5)
## Parch <= 1:
## :...Sex = female:
## :   :...Pclass = 1: 1 (7.6)
## :   :   Pclass in {2,3}:
## :   :   :...Age <= 17: 1 (17.8/3.1)
## :   :   :   Age > 17:
## :   :   :   :...Embarked in {C,Q}: 0 (11.1/1.5)
## :   :   :   :   Embarked = S:
## :   :   :   :   :...Fare > 11.2417: 1 (43.5/16.6)
## :   :   :   :   :   Fare <= 11.2417:
## :   :   :   :   :   :...Age <= 19: 1 (3.7)
## :   :   :   :   :   :   Age > 19: 0 (45.9/13.5)
## :   Sex = male:
## :   :...Parch > 0:
## :   :   :...SibSp > 3: 0 (2.5)
## :   :   :   SibSp <= 3:
## :   :   :   :...Age <= 10: 1 (8)
## :   :   :   :   Age > 10: 0 (24.8/9.7)
## :   :   Parch <= 0:
## :   :   :...Pclass in {2,3}: 0 (190.5/84.5)
## :   :   :   Pclass = 1:
## :   :   :   :...Age <= 28.5: 1 (18.5/6.1)
## :   :   :   :   Age > 28.5:
## :   :   :   :   :...Fare <= 26: 0 (4.3)
## :   :   :   :   :   Fare > 26:
## :   :   :   :   :   :...Fare <= 26.2875: 1 (5)
## :   :   :   :   :   :   Fare > 26.2875:
## :   :   :   :   :   :   :...Age <= 47: 0 (38.7/9.5)
## :   :   :   :   :   :   :   Age > 47: 1 (21.3/8.5)
##
## ----- Trial 21: -----
##
## Decision tree:
##
## Sex = female:
## :...Pclass = 1: 1 (20.9/7.7)
## :   Pclass = 3:
## :   :...Embarked in {C,Q}: 1 (20.1/9.2)
## :   :   Embarked = S:
## :   :   :...Fare <= 7.65: 1 (4.8)
## :   :   :   Fare > 7.65: 0 (84.7/34.5)

```

```

## : Pclass = 2:
## : ...Parch > 1: 1 (8)
## :   Parch <= 1:
## :     ...Embarked in {C,Q}: 1 (3.5)
## :       Embarked = S:
## :         ...SibSp <= 0: 1 (14.4/3.8)
## :           SibSp > 0: 0 (19.3/7.1)
## Sex = male:
## ...Pclass = 2: 0 (69/31.1)
##   Pclass = 3:
##     ...SibSp > 0: 0 (32.2/13.6)
##       SibSp <= 0:
##         ...Age <= 32.5: 1 (91/35.6)
##           Age > 32.5: 0 (21.8/6.3)
##   Pclass = 1:
##     ...Age > 61: 0 (6.3)
##       Age <= 61:
##         ...Embarked = Q: 0 (0.4)
##           Embarked = S:
##             ...Parch > 1: 1 (3.2)
##               Parch <= 1:
##                 ...Fare <= 57.9792: 1 (48.2/23.5)
##                   Fare > 57.9792: 0 (9.1/1.2)
##             Embarked = C:
##               ...Fare > 263: 1 (4.7)
##                 Fare <= 263:
##                   ...Fare <= 93.5: 1 (29.1/12.5)
##                     Fare > 93.5: 0 (8.2)
##
## ----- Trial 22: -----
##
## Decision tree:
##
## Sex = female: 1 (172.7/78.7)
## Sex = male:
## ...SibSp <= 0:
##   ...Pclass in {2,3}: 0 (170.7/79.3)
##     Pclass = 1:
##       ...Fare > 263: 1 (4.2)
##         Fare <= 263:
##           ...Fare > 76.7292: 0 (8.9)
##             Fare <= 76.7292:
##               ...Age <= 27: 1 (6.5)
##                 Age > 27: 0 (51.2/16.6)
##   SibSp > 0:
##     ...SibSp > 2: 0 (4.4)
##       SibSp <= 2:
##         ...Age <= 13: 1 (9.9)
##           Age > 13:
##             ...Age <= 19: 0 (6.1)
##               Age > 19:
##                 ...Age <= 20.5: 1 (7.2)
##                   Age > 20.5: 0 (57.1/25.6)
##
## ----- Trial 23: -----
##
## Decision tree:
##
## Pclass = 2: 1 (108.2/47.1)
## Pclass = 1:
##   ...SibSp <= 0: 0 (73.5/33.5)
##     SibSp > 0:
##       ...SibSp > 1: 1 (4.2/0.3)
##         SibSp <= 1:
##           ...Fare <= 135.6333: 1 (38.2/15.9)
##             Fare > 135.6333: 0 (10.7/1.7)
##   Pclass = 3:
##     ...SibSp > 1: 0 (20.2/5.4)
##       SibSp <= 1:
##         ...Sex = female: 0 (99.3/42.9)
##           Sex = male:
##             ...Parch > 0: 1 (15/5.9)
##               Parch <= 0:
##                 ...Fare > 40.125: 1 (6.3/1.3)
##                   Fare <= 40.125:
##                     ...Fare <= 8.05: 1 (105.6/47.5)
##                       Fare > 8.05: 0 (17.9/2.8)
##
## ----- Trial 24: -----
##
## Decision tree:
##
## Embarked = Q: 0 (24/9.8)

```

```

## Embarked = S:
## :...SibSp > 1: 0 (23.9/9.2)
## :   SibSp <= 1:
## :     :...SibSp <= 0:
## :       :...Sex = female: 1 (69.5/30.4)
## :       :   Sex = male: 0 (175.9/77.8)
## :       SibSp > 0:
## :         :...Age <= 38: 1 (86.3/40.7)
## :         Age > 38: 0 (26.9/9.7)
## Embarked = C:
## :...Sex = female: 1 (17.3/4.6)
##   Sex = male:
##   :...Fare > 263: 1 (4)
##   Fare <= 263:
##   :...Fare > 93.5: 0 (6.8)
##   Fare <= 93.5:
##   :...Pclass = 2: 0 (2.6)
##   Pclass in {1,3}:
##   :...SibSp > 0: 1 (14.1/4.2)
##   SibSp <= 0:
##   :...Age <= 32.5: 1 (34.1/13.7)
##   Age > 32.5: 0 (13.6/2.2)
##
## ----- Trial 25: -----
##
## Decision tree:
##
## Pclass = 3: 0 (264.1/118.5)
## Pclass = 2:
## :...Age <= 20: 1 (16.4/2.6)
## :   Age > 20:
## :     :...Sex = female: 1 (36.4/16.1)
## :     Sex = male:
## :       :...Age <= 31: 0 (13.2)
## :       Age > 31: 1 (42.2/19.9)
## Pclass = 1:
## :...Sex = female: 1 (23.2/8.4)
##   Sex = male:
##   :...Parch > 0: 1 (17/7.1)
##   Parch <= 0:
##   :...Age > 57: 0 (6)
##   Age <= 57:
##   :...Age <= 24.5: 0 (5.2)
##   Age > 24.5:
##   :...Age <= 27: 1 (7.6)
##   Age > 27:
##   :...Age <= 47: 0 (49.4/19.7)
##   Age > 47: 1 (18.2/4.3)
##
## ----- Trial 26: -----
##
## Decision tree:
##
## Sex = female: 1 (176/84.2)
## Sex = male:
## :...Pclass = 2: 0 (72.5/29.1)
##   Pclass = 3:
##   :...SibSp <= 0: 1 (118.6/57.2)
##   :   SibSp > 0: 0 (30.5/13.4)
##   Pclass = 1:
##   :...Embarked = Q: 0 (0.4)
##   Embarked = C:
##   :...Fare <= 263: 0 (33.8/14.7)
##   :   Fare > 263: 1 (3.8)
##   Embarked = S:
##   :...Age > 53: 0 (6.3)
##   Age <= 53:
##   :...SibSp > 0: 1 (23.1/10.9)
##   SibSp <= 0:
##   :...Fare <= 35.5: 1 (28.9/9.3)
##   Fare > 35.5: 0 (5.2)
##
## ----- Trial 27: -----
##
## Decision tree:
##
## Embarked = Q: 0 (24.1/11)
## Embarked = S:
## :...Pclass = 2:
## :   :...Parch <= 1: 0 (96.6/43.4)
## :   :   Parch > 1: 1 (6.2/1.7)
## :   Pclass = 3:
## :     :...SibSp > 1: 0 (16.6/4.7)

```

```

## : : SibSp <= 1:
## : : ...Age <= 8: 1 (9.3)
## : : Age > 8: 0 (176.6/78.7)
## : Pclass = 1:
## : ...SibSp > 0:
## : : ...SibSp <= 1: 0 (36/14.4)
## : : SibSp > 1: 1 (3.3)
## : SibSp <= 0:
## : : ...Sex = female: 1 (3.5)
## : Sex = male:
## : : ...Fare > 35.5: 0 (5.8)
## : Fare <= 35.5:
## : : ...Age <= 54: 1 (28.5/10)
## : Age > 54: 0 (3.9)
## Embarked = C:
## : ...Sex = female: 1 (16.5/4.8)
## Sex = male:
## : ...Pclass = 2: 0 (2.5)
## Pclass in {1,3}:
## : ...Fare > 263: 1 (3.6)
## Fare <= 263:
## : ...Fare > 93.5: 0 (6.3)
## Fare <= 93.5:
## : ...SibSp <= 0: 0 (45.8/21.6)
## SibSp > 0: 1 (13.7/4.4)
##
## ----- Trial 28: -----
##
## Decision tree:
##
## Sex = female: 1 (180.5/80.6)
## Sex = male:
## : ...Pclass = 2:
## : ...Age <= 9: 1 (4.1)
## : Age > 9: 0 (65.6/27.6)
## Pclass = 3:
## : ...Fare <= 40.125: 0 (143.4/69.2)
## : Fare > 40.125: 1 (7.4/1.7)
## Pclass = 1:
## : ...Embarked = C: 1 (35.8/16.5)
## Embarked = Q: 0 (0.3)
## Embarked = S:
## : ...Age > 53: 0 (5.4)
## Age <= 53:
## : ...Fare <= 25.9292: 0 (4.3)
## Fare > 25.9292:
## : ...Fare <= 26.2875: 1 (5.5)
## Fare > 26.2875:
## : ...Parch > 1: 1 (2.5)
## Parch <= 1:
## : ...Parch > 0: 0 (5.4/2)
## Parch <= 0:
## : ...Age > 47: 1 (7.1)
## Age <= 47:
## : ...Age <= 28.5: 1 (6.5/1.1)
## Age > 28.5: 0 (25/6.9)
##
## ----- Trial 29: -----
##
## Decision tree:
##
## Parch > 1:
## : ...Fare <= 13.5: 1 (4)
## : Fare > 13.5: 0 (55.9/18.8)
## Parch <= 1:
## : ...Parch > 0:
## : ...Age <= 6: 1 (15.3/0.7)
## : Age > 6: 0 (53.9/24.3)
## Parch <= 0:
## : ...Sex = female:
## : ...Fare > 11.2417: 1 (36.9/9.9)
## : Fare <= 11.2417:
## : : ...Embarked in {C,Q}: 1 (11.1/4.4)
## : Embarked = S:
## : : ...Fare <= 7.65: 1 (4.6)
## : Fare > 7.65: 0 (47.7/16.1)
## Sex = male:
## : ...Embarked = C: 0 (56.4/26.2)
## Embarked = Q: 1 (11.1/5.4)
## Embarked = S:
## : ...SibSp > 0: 1 (39.3/17.3)
## SibSp <= 0:
## : ...Age > 30.5:

```

```

##          :...Age <= 32.5: 1 (21.3/3.7)
##          :   Age > 32.5: 0 (74.6/36.2)
##          Age <= 30.5:
##          :...Fare <= 7.1417: 1 (13.2/3.3)
##          :   Fare > 7.1417:
##          :...Age <= 19: 1 (21.2/7.3)
##          :   Age > 19: 0 (32.5/3.8)
##
## ----- Trial 30: -----
##
## Decision tree:
##
## Embarked = C: 1 (88.7/37.9)
## Embarked = Q: 0 (25.6/10.3)
## Embarked = S:
## :...Pclass = 2:
##   :...Sex = female: 1 (42.3/15.2)
##   :   Sex = male:
##   :     :...Age <= 9: 1 (3.4)
##   :     :   Age > 9: 0 (62/28.2)
##   Pclass = 1:
##   :...Sex = female: 1 (21.5/7.7)
##   :   Sex = male:
##   :     :...Age > 53: 0 (4.4)
##   :     :   Age <= 53:
##   :     :     :...Parch > 1: 1 (2.8)
##   :     :     :   Parch <= 1:
##   :     :     :     :...Fare <= 35.5: 1 (28/9.9)
##   :     :     :     :   Fare > 35.5: 0 (24.7/9.6)
##   Pclass = 3:
##   :...SibSp > 1: 0 (18/4.3)
##   :   SibSp <= 1:
##   :     :...Age <= 8: 1 (9)
##   :     :   Age > 8:
##   :     :     :...Age <= 17: 0 (9.3)
##   :     :     :   Age > 17:
##   :     :     :     :...SibSp > 0: 0 (39.1/17.7)
##   :     :     :     :   SibSp <= 0:
##   :     :     :     :     :...Parch <= 1: 1 (113.6/52.3)
##   :     :     :     :     :   Parch > 1: 0 (6.6/1.2)
##
## ----- Trial 31: -----
##
## Decision tree:
##
## Pclass = 1:
## :...Age <= 57: 1 (112.8/50.7)
## :   Age > 57: 0 (8.7/2.5)
## Pclass = 2:
## :...Sex = female: 1 (43/15.7)
## :   Sex = male:
## :     :...Fare <= 13: 1 (43.8/19.7)
## :     :   Fare > 13: 0 (22.8/8.1)
## Pclass = 3:
## :...Embarked in {S,Q}: 0 (225.4/99.6)
##   Embarked = C:
##   :...Age <= 29: 1 (35.6/14.6)
##   :   Age > 29: 0 (6.9)
##
## ----- Trial 32: -----
##
## Decision tree:
##
## Sex = female: 1 (175.3/79.8)
## Sex = male:
## :...Age <= 11: 1 (16.9/3.9)
## :   Age > 11:
## :     :...Pclass in {1,2}: 0 (162.7/70.7)
## :     :   Pclass = 3:
## :     :     :...SibSp > 0: 0 (24.2/10.8)
## :     :     :   SibSp <= 0:
## :     :     :     :...Embarked = C: 1 (21.6/9.9)
## :     :     :     :   Embarked = Q: 0 (10.8/5.3)
## :     :     :     :   Embarked = S:
## :     :     :     :     :...Age <= 30.5: 0 (50.4/14.5)
## :     :     :     :     :   Age > 30.5:
## :     :     :     :     :     :...Age <= 32.5: 1 (17.7/2.7)
## :     :     :     :     :     :   Age > 32.5: 0 (19.4/6.5)
##
## ----- Trial 33: -----
##
## Decision tree:
##

```

```

## Pclass = 3: 0 (267.2/120)
## Pclass = 2:
## :...Age <= 15: 1 (6.1)
## :   Age > 15:
## :     :...SibSp > 0: 0 (37.9/16.3)
## :       SibSp <= 0:
## :         :...Sex = female: 1 (14.4/4.7)
## :           Sex = male: 0 (49.8/23.4)
## Pclass = 1:
## :...Sex = female: 1 (23.4/8.8)
##   Sex = male:
##     :...SibSp > 0: 1 (36/15.6)
##       SibSp <= 0:
##         :...Fare > 35.5: 0 (22.7/7)
##           Fare <= 35.5:
##             :...Fare <= 26: 0 (4.5)
##               Fare > 26: 1 (37.1/13.3)
##
## ----- Trial 34: -----
##
## Decision tree:
##
## Sex = female:
## :...Pclass in {1,2}: 1 (66.8/25.8)
## :   Pclass = 3:
## :     :...Fare > 23: 0 (18.3/4)
## :       Fare <= 23:
## :         :...SibSp > 2: 1 (5.1)
## :           SibSp <= 2:
## :             :...Age <= 17: 1 (17.1/2.4)
## :               Age > 17:
## :                 :...Embarked in {C,Q}: 0 (12.6/1.3)
## :                   Embarked = S:
## :                     :...SibSp <= 0: 1 (45/19.2)
## :                       SibSp > 0: 0 (20.2/7.6)
## Sex = male:
## :...Parch > 0:
## :   :...SibSp > 2: 0 (4.6)
## :     :   SibSp <= 2:
## :       :   :...Age <= 13: 1 (12.9)
## :         :     Age > 13: 0 (27.7/11.4)
## :       Parch <= 0:
## :         :...Age > 32.5: 0 (108.5/45)
## :           Age <= 32.5:
## :             :...Age <= 19: 0 (22.4/6)
## :               Age > 19:
## :                 :...Embarked in {C,Q}: 1 (45/19.3)
## :                   Embarked = S:
## :                     :...Age > 30.5: 1 (29.3/8.4)
## :                       Age <= 30.5:
## :                         :...Fare <= 7.925: 1 (33.2/14.6)
## :                           Fare > 7.925: 0 (28.4/5.1)
##
## ----- Trial 35: -----
##
## Decision tree:
##
## Fare <= 15.55:
## :...Age > 32.5: 0 (44/4.2)
## :   Age <= 32.5:
## :     :...Parch > 0: 1 (23.3/9.2)
## :       Parch <= 0:
## :         :...Sex = female:
## :           :...Embarked in {C,Q}: 1 (12.5/4.3)
## :             :   Embarked = S:
## :               :   :...Fare <= 7.65: 1 (4.2)
## :                 :     Fare > 7.65:
## :                   :       :...Age <= 25: 0 (26.5/5.4)
## :                     :         Age > 25: 1 (13.4/3.5)
## :                   Sex = male:
## :                     :...Fare > 8.05: 0 (22.9)
## :                       Fare <= 8.05:
## :                         :...SibSp > 0: 1 (10.6/2.2)
## :                           SibSp <= 0:
## :                             :...Age <= 28.5: 0 (59.2/14)
## :                               Age > 28.5: 1 (30.8/13.2)
## Fare > 15.55:
## :...Parch > 1: 0 (56.6/21.6)
## :   Parch <= 1:
## :     :...Sex = female: 1 (39.7/4)
## :       Sex = male:
## :         :...Parch > 0: 1 (26.9/8.7)
## :           Parch <= 0:

```



```

##      :...Age <= 24.5: 0 (7.8)
##      Age > 24.5:
##      :...Pclass = 2: 0 (14/5.1)
##      Pclass = 3: 1 (10.8/3.9)
##      Pclass = 1:
##      :...Age <= 27: 1 (8.9)
##      Age > 27:
##      :...Age > 57: 0 (4.6)
##      Age <= 57:
##      :...Age > 47: 1 (22.1/4.1)
##      Age <= 47:
##      :...Fare <= 27: 1 (13.7/2.7)
##      Fare > 27: 0 (38.4/12.3)
##
## ----- Trial 36: -----
##
## Decision tree:
##
## Sex = female:
## :...Pclass in {1,3}:
## :   :...Pclass = 1: 1 (25.8/9.3)
## :   :   Pclass = 3: 0 (117.4/54.8)
## :   Pclass = 2:
## :   :...Age <= 56: 1 (33.5)
## :   :   Age > 56: 0 (4.9)
## Sex = male:
## :...Pclass = 2: 0 (33.6/2.4)
##   Pclass in {1,3}:
##   :...Pclass = 1:
##   :   :...Embarked = Q: 0 (0.6)
##   :   :   Embarked = C:
##   :   :   :...Fare <= 29.7: 0 (4)
##   :   :   :   Fare > 29.7: 1 (37.6/14.1)
##   :   :   Embarked = S:
##   :   :   :...Age > 53: 0 (6.5)
##   :   :   :   Age <= 53:
##   :   :   :   :...Parch > 1: 1 (3.2)
##   :   :   :   :   Parch <= 1:
##   :   :   :   :   :...Fare <= 57.9792: 1 (50.1/19.2)
##   :   :   :   :   :   Fare > 57.9792: 0 (9.6/1.6)
##   :   Pclass = 3:
##   :   :...SibSp > 1: 0 (9.7)
##   :   :   SibSp <= 1:
##   :   :   :...Age > 32.5: 0 (15.4)
##   :   :   :   Age <= 32.5:
##   :   :   :   :...Age <= 13: 1 (8/1.2)
##   :   :   :   :   Age > 13:
##   :   :   :   :   :...Age <= 19: 0 (13.3)
##   :   :   :   :   :   Age > 19:
##   :   :   :   :   :   :...Embarked in {C,Q}: 0 (40.2/19)
##   :   :   :   :   :   :   Embarked = S:
##   :   :   :   :   :   :   :...Age > 30.5: 1 (18.1/6.6)
##   :   :   :   :   :   :   :   Age <= 30.5:
##   :   :   :   :   :   :   :   :...SibSp > 0: 1 (11.6/4.6)
##   :   :   :   :   :   :   :   :   SibSp <= 0:
##   :   :   :   :   :   :   :   :   :...Fare <= 7.1417: 1 (12.7/2.7)
##   :   :   :   :   :   :   :   :   :   Fare > 7.1417: 0 (34.2)
##
## ----- Trial 37: -----
##
## Decision tree:
##
## Sex = female:
## :...Pclass in {1,3}:
## :   :...Fare <= 135.6333: 1 (138.5/52.9)
## :   :   Fare > 135.6333: 0 (16/5.3)
## :   Pclass = 2:
## :   :...Age <= 56: 1 (28.4)
## :   :   Age > 56: 0 (4.1)
## Sex = male:
## :...Age <= 7: 1 (15.9/3.6)
##   Age > 7:
##   :...Pclass = 2: 0 (26.4)
##     Pclass in {1,3}:
##     :...Age <= 19: 0 (19.9/1.4)
##     :   Age > 19:
##     :   :...Embarked = Q: 0 (11.7/5.2)
##     :   :   Embarked = C:
##     :   :   :...Fare > 263: 1 (4.7)
##     :   :   :   Fare <= 263:
##     :   :   :   :...Age <= 23: 1 (16/1.9)
##     :   :   :   :   Age > 23: 0 (52.9/19)
##     :   :   Embarked = S:

```

```

##          :...Pclass = 3: 0 (84.6/24.1)
##          Pclass = 1:
##          :...Fare <= 26: 0 (6.8)
##          Fare > 26:
##          :...Fare <= 26.2875: 1 (5.1)
##          Fare > 26.2875: 0 (59.1/23.3)
##
## ----- Trial 38: -----
##
## Decision tree:
##
## Sex = male:
## :...Embarked = Q: 0 (8.4)
## :   Embarked in {S,C}:
## :   :...Pclass = 2: 0 (26.5/4.4)
## :   Pclass = 1:
## :   :...SibSp > 0: 1 (44.8/16.5)
## :   :   SibSp <= 0:
## :   :   :...Fare <= 26: 0 (5.8)
## :   :   Fare > 26:
## :   :   :...Fare <= 35.5: 1 (41.7/14.7)
## :   :   Fare > 35.5: 0 (26.9/7.6)
## :   Pclass = 3:
## :   :...SibSp > 1: 0 (8.6)
## :   SibSp <= 1:
## :   :...Age > 32.5: 0 (10.4)
## :   Age <= 32.5:
## :   :...Embarked = C: 1 (34/15.7)
## :   Embarked = S:
## :   :...Age > 30.5: 1 (20.2/7.4)
## :   Age <= 30.5:
## :   :...SibSp > 0: 1 (18.3/8)
## :   SibSp <= 0:
## :   :...Fare <= 7.1417: 1 (14.1/4.4)
## :   Fare > 7.1417: 0 (29.1/0.9)
## Sex = female:
## :...Pclass in {1,2}:
## :   :...Age <= 2: 0 (4.7/0.2)
## :   Age > 2: 1 (53.3/8.1)
## Pclass = 3:
## :...Age <= 5: 1 (10.4/1.2)
## Age > 5:
## :...Age <= 11: 0 (7.6)
## Age > 11:
## :...Embarked in {C,Q}: 0 (27.3/11)
## Embarked = S:
## :...Fare > 32.5: 0 (5.1)
## Fare <= 32.5:
## :...Fare > 31.275: 1 (5.3)
## Fare <= 31.275:
## :...SibSp > 2: 1 (4.4)
## SibSp <= 2:
## :...Parch > 3: 0 (4.7)
## Parch <= 3:
## :...Fare <= 7.65: 1 (4.7)
## Fare > 7.65:
## :...SibSp > 0: 0 (25/8.7)
## SibSp <= 0:
## :...Age > 52: 1 (3)
## Age <= 52:
## :...Age > 34: 0 (5.1)
## Age <= 34:
## :...Parch <= 0: 0 (34.4/14.2)
## Parch > 0: 1 (5)
##
## ----- Trial 39: -----
##
## Decision tree:
##
## Sex = female:
## :...Pclass in {1,2}: 1 (42.8/3.8)
## :   Pclass = 3:
## :   :...Fare > 23: 0 (20.5/4.4)
## :   Fare <= 23:
## :   :...Age <= 17: 1 (24.5/2.4)
## :   Age > 17:
## :   :...SibSp > 0: 0 (28.6/13.2)
## :   SibSp <= 0:
## :   :...Fare > 19.5: 0 (3.6)
## :   Fare <= 19.5:
## :   :...Parch > 1: 1 (5.1)
## :   Parch <= 1:
## :   :...Embarked = S: 1 (45.9/17)

```

```

## :                               Embarked in {C,Q}: 0 (10.9)
## Sex = male:
## :...Embarked = Q: 0 (7)
##   Embarked in {S,C}:
##     :...Age <= 7: 1 (18.5/4.5)
##       Age > 7:
##         :...Pclass = 2: 0 (18.5)
##           Pclass in {1,3}:
##             :...Age <= 19: 0 (24.2/1.7)
##               Age > 19:
##                 :...Embarked = S:
##                   :...Age > 53: 0 (11.1)
##                     :   Age <= 53:
##                       :     :...Age > 47: 1 (12.3/2.8)
##                         :       Age <= 47:
##                           :         :...Fare > 61.3792: 0 (6.5)
##                             :           Fare <= 61.3792:
##                               :             :...Fare <= 52: 0 (112.1/32.8)
##                                 :               Fare > 52: 1 (15.6/3.9)
##                               Embarked = C:
##                                 :...Fare > 263: 1 (5.7)
##                                   Fare <= 263:
##                                     :...Fare > 93.5: 0 (6.3)
##                                       Fare <= 93.5:
##                                         :...Parch > 0: 1 (9.5/2.2)
##                                           Parch <= 0:
##                                             :...Fare <= 29.7: 0 (36.5/11.9)
##                                               Fare > 29.7: 1 (21.4/5.9)
##
## ----- Trial 40: -----
##
## Decision tree:
##
## Fare > 108.9: 1 (24.8/3.3)
## Fare <= 108.9:
## :...Sex = female:
##   :...Pclass in {1,2}: 1 (21.7)
##     :   Pclass = 3:
##       :     :...Embarked = C: 1 (12.5/4)
##         :       Embarked = Q: 0 (14.5/7.1)
##           :       Embarked = S:
##             :         :...Fare <= 10.5167: 0 (61.3/22.3)
##               :           Fare > 10.5167:
##                 :             :...Parch <= 1: 1 (31.5/5.7)
##                   :               Parch > 1: 0 (24.5/9.1)
##             Sex = male:
##               :...SibSp > 1: 0 (18.2/1.8)
##                 SibSp <= 1:
##                   :...Fare <= 7.1417: 0 (14.9)
##                     Fare > 7.1417:
##                       :...Age > 49: 0 (34.3/6.4)
##                         Age <= 49:
##                           :...Pclass = 2: 0 (17.3/4.7)
##                             Pclass = 1:
##                               :...Age > 47: 1 (6.6)
##                                 :   Age <= 47:
##                                   :     :...Age > 45: 0 (9.7)
##                                     :       Age <= 45:
##                                       :         :...Fare <= 26.2875: 1 (8.8)
##                                         :           Fare > 26.2875:
##                                           :             :...Parch > 0: 1 (4.8/0.7)
##                                             :               Parch <= 0:
##                                                 :                 :...Age <= 28.5: 1 (22.2/7.3)
##                                                   :                   Age > 28.5: 0 (38.2/10.1)
##                             Pclass = 3:
##                               :...Embarked = Q: 0 (3.3)
##                                 Embarked = C:
##                                   :...Age > 29: 0 (5.2)
##                                     :   Age <= 29:
##                                       :     :...SibSp <= 0: 1 (28.4/8.9)
##                                         :       SibSp > 0: 0 (9.1/2.3)
##                               Embarked = S:
##                                 :...Age <= 11: 1 (3.6)
##                                   Age > 11:
##                                     :...Parch > 0: 0 (6.7)
##                                       Parch <= 0:
##                                         :...SibSp > 0: 1 (16.6/7)
##                                           SibSp <= 0:
##                                             :...Age <= 30.5: 0 (21.5)
##                                               Age > 30.5: 1 (23.9/8.5)
##
## ----- Trial 41: -----
##

```

[illegible]

```

##                                     Age > 30.5: 1 (25/6.6)
##
## ----- Trial 43: -----
##
## Decision tree:
##
## Sex = female:
## :...Pclass in {1,2}: 1 (18.6)
## :   Pclass = 3:
## :     :...Fare > 23: 0 (13.1)
## :       Fare <= 23:
## :         :...Age <= 17: 1 (20.5/5.1)
## :           Age > 17: 0 (109.3/46.7)
## Sex = male:
## :...Pclass = 2:
## :   :...Age <= 9: 1 (7.4)
## :     Age > 9: 0 (44.1)
## :   Pclass in {1,3}:
## :     :...Embarked = Q: 0 (3.7)
## :       Embarked = S:
## :         :...Age > 53: 0 (10.7)
## :           Age <= 53:
## :             :...SibSp > 2: 0 (7.3)
## :               SibSp <= 2:
## :                 :...Age <= 13: 1 (8.6)
## :                   Age > 13: 0 (142.9/46.9)
## :       Embarked = C:
## :         :...Age <= 19: 0 (9.2/0.4)
## :           Age > 19:
## :             :...Fare > 263: 1 (4.3)
## :               Fare <= 263:
## :                 :...Fare > 93.5: 0 (5.5)
## :                   Fare <= 93.5:
## :                     :...SibSp > 0: 1 (18/4.2)
## :                       SibSp <= 0:
## :                         :...Fare > 77.9583: 0 (5.2)
## :                           Fare <= 77.9583:
## :                             :...Age <= 29: 1 (29.7/7)
## :                               Age > 29: 0 (22.7/6.1)
##
## ----- Trial 44: -----
##
## Decision tree:
##
## Sex = male:
## :...Pclass = 1:
## :   :...Embarked = Q: 0 (0.6)
## :     Embarked = S:
## :       :...Age > 53: 0 (8.9)
## :         Age <= 53:
## :           :...Age > 47: 1 (20.7/2.9)
## :             Age <= 47:
## :               :...Age <= 28.5: 1 (16.1/2.1)
## :                 Age > 28.5:
## :                   :...Fare <= 27: 1 (23.3/7.5)
## :                     Fare > 27: 0 (24.9/4)
## :       Embarked = C:
## :         :...Fare <= 29.7: 0 (5)
## :           Fare > 29.7:
## :             :...Fare <= 37.0042: 1 (10.1)
## :               Fare > 37.0042:
## :                 :...Fare > 263: 1 (3.7)
## :                   Fare <= 263:
## :                     :...Fare <= 93.5: 1 (25/9.3)
## :                       Fare > 93.5: 0 (7.6)
## :   Pclass in {2,3}:
## :     :...Age > 32: 0 (35.5)
## :       Age <= 32:
## :         :...SibSp > 2: 0 (7)
## :           SibSp <= 2:
## :             :...Age <= 8: 1 (13.6)
## :               Age > 8:
## :                 :...Pclass = 2: 0 (19.5)
## :                   Pclass = 3:
## :                     :...SibSp > 0: 0 (21.3/2)
## :                       SibSp <= 0:
## :                         :...Age <= 30.5: 0 (48.5/13)
## :                           Age > 30.5: 1 (25.7/7.4)
## Sex = female:
## :...Pclass in {1,2}: 1 (15.9)
## :   Pclass = 3:
## :     :...Fare > 23: 0 (11.2)
## :       Fare <= 23:

```

```

##      :...Age <= 17: 1 (13.2)
##      Age > 17:
##      :...Embarked = Q: 0 (4.3)
##      Embarked in {S,C}:
##      :...Fare <= 7.65: 1 (9.4)
##      Fare > 7.65:
##      :...Age > 29: 1 (35/7.8)
##      Age <= 29:
##      :...Embarked = C: 1 (4.8)
##      Embarked = S:
##      :...Age > 27: 0 (8.6)
##      Age <= 27:
##      :...Age > 26: 1 (10.7)
##      Age <= 26:
##      :...SibSp > 1: 0 (2.8)
##      SibSp <= 1:
##      :...Age <= 19: 1 (8)
##      Age > 19: 0 (34.2/11.3)
##
## ----- Trial 45: -----
##
## Decision tree:
##
## Sex = female:
## :...Pclass in {1,2}: 1 (12.7)
## :   Pclass = 3:
## :     :...Age > 36.5: 0 (26.9/3.9)
## :     Age <= 36.5:
## :       :...Fare > 23: 0 (7.2)
## :       Fare <= 23:
## :         :...Age <= 17: 1 (10.6)
## :         Age > 17:
## :           :...Embarked = C: 1 (3.8)
## :           Embarked = Q: 0 (3.5)
## :           Embarked = S:
## :             :...Age > 29: 1 (17.9)
## :             Age <= 29:
## :               :...SibSp <= 0: 1 (45.5/14.4)
## :               SibSp > 0: 0 (13.1)
##
## Sex = male:
## :...Pclass = 2:
## :   :...Age <= 9: 1 (5.1)
## :   Age > 9: 0 (30.2)
## :   Pclass in {1,3}:
## :     :...Embarked = Q: 0 (9.4)
## :     Embarked = S:
## :       :...Fare <= 7.8958: 0 (34.9)
## :       Fare > 7.8958:
## :         :...Pclass = 3:
## :         :   :...Fare <= 51.4792: 0 (42.1/13.1)
## :         :   Fare > 51.4792: 1 (6.9/1.1)
## :         :   Pclass = 1:
## :         :     :...Age > 53: 0 (7.1)
## :         :     Age <= 53:
## :         :       :...Fare <= 25.9292: 0 (4.6)
## :         :       Fare > 25.9292:
## :         :         :...Fare <= 26.2875: 1 (8.1)
## :         :         Fare > 26.2875:
## :         :           :...Fare > 57.9792: 0 (19.2/4.9)
## :         :           Fare <= 57.9792:
## :         :             :...Parch > 0: 1 (7.4)
## :         :             Parch <= 0:
## :         :               :...Age > 47: 1 (11)
## :         :               Age <= 47:
## :         :                 :...Age <= 27: 1 (5.8)
## :         :                 Age > 27: 0 (32.5/8.2)
##
## Embarked = C:
## :...Parch > 0: 1 (21.1/8.4)
## :   Parch <= 0:
## :     :...Age > 57: 0 (6)
## :     Age <= 57:
## :       :...Fare > 77.9583: 0 (19.4/3.9)
## :       Fare <= 77.9583:
## :         :...Fare > 47.1: 1 (5.8)
## :         Fare <= 47.1:
## :           :...SibSp > 0: 0 (6)
## :           SibSp <= 0:
## :             :...Fare > 37.0042: 0 (4.8)
## :             Fare <= 37.0042:
## :               :...Fare > 28.5: 1 (8.1)
## :               Fare <= 28.5:
## :                 :...Pclass = 1: 0 (2.8)
## :                 Pclass = 3:

```

```

##          :...Fare <= 7.55: 0 (18.6/7)
##          Fare > 7.55: 1 (15.3/3.9)
##
##
## Evaluation on training data (499 cases):
##
## Trial      Decision Tree
## -----
##      Size      Errors
##
##      0      22    51(10.2%)
##      1      16   107(21.4%)
##      2       9    97(19.4%)
##      3      23    71(14.2%)
##      4      10    99(19.8%)
##      5      23   171(34.3%)
##      6      14    93(18.6%)
##      7      18   120(24.0%)
##      8      23   114(22.8%)
##      9      12    74(14.8%)
##     10      17   158(31.7%)
##     11      10   129(25.9%)
##     12      12   125(25.1%)
##     13      25   154(30.9%)
##     14       7    92(18.4%)
##     15      12   122(24.4%)
##     16      22   120(24.0%)
##     17      18    76(15.2%)
##     18      22   201(40.3%)
##     19      14   108(21.6%)
##     20      17    84(16.8%)
##     21      20   161(32.3%)
##     22      11    86(17.2%)
##     23      11   268(53.7%)
##     24      13   108(21.6%)
##     25      12   118(23.6%)
##     26      11   194(38.9%)
##     27      18   124(24.8%)
##     28      15    88(17.6%)
##     29      16   160(32.1%)
##     30      16   162(32.5%)
##     31       8   140(28.1%)
##     32       9   104(20.8%)
##     33      10   110(22.0%)
##     34      16   103(20.6%)
##     35      21   106(21.2%)
##     36      21    88(17.6%)
##     37      15   101(20.2%)
##     38      28    86(17.2%)
##     39      22    58(11.6%)
##     40      26    81(16.2%)
##     41      15    76(15.2%)
##     42      26   139(27.9%)
##     43      18    64(12.8%)
##     44      30    52(10.4%)
##     45      33    56(11.2%)
## boost          30( 6.0%)  <<
##
##
##      (a)  (b)  <-classified as
##      ----  ----
##      289    8   (a): class 0
##      22   180   (b): class 1
##
##
## Attribute usage:
##
## 100.00% Pclass
## 100.00% Sex
## 100.00% Age
## 100.00% SibSp
## 100.00% Parch
## 100.00% Fare
## 100.00% Embarked
##
##
## Time: 0.1 secs

```

Aplicamos el bosque de decisión al test

```

newdata <- test[,colnames(test)!="Survived"]
prediccion <- predict(bosque, newdata=newdata, type="class")

```

Mostramos los resultados de aplicar el bosque de decisión al test

```
confusionMatrix(data=prediccion, reference=test$Survived, positive="1")
```

```
## Confusion Matrix and Statistics
##
##              Reference
## Prediction    0    1
##              0 107  32
##              1  20  54
##
##              Accuracy : 0.7559
##              95% CI : (0.6925, 0.812)
##      No Information Rate : 0.5962
##      P-Value [Acc > NIR] : 7.064e-07
##
##              Kappa : 0.4813
##  Mcnemar's Test P-Value : 0.1272
##
##              Sensitivity : 0.6279
##              Specificity : 0.8425
##      Pos Pred Value : 0.7297
##      Neg Pred Value : 0.7698
##      Prevalence : 0.4038
##      Detection Rate : 0.2535
##      Detection Prevalence : 0.3474
##      Balanced Accuracy : 0.7352
##
##      'Positive' Class : 1
##
```

Como se observa, hemos ganado algunos puntos en el poder de predicción.

5. Representación de los resultados a partir de tablas y gráficas.

De <https://www.kaggle.com/c/titanic> (<https://www.kaggle.com/c/titanic>) se descargaron dos dataset. El primero es el que hemos venido utilizando para hacer el train y test. El segundo dataset es el siguiente:

```
glimpse(aplicar)
```

```
## Observations: 418
## Variables: 11
## $ PassengerId <int> 892, 893, 894, 895, 896, 897, 898, 899, 900, 901, ...
## $ Pclass      <int> 3, 3, 2, 3, 3, 3, 3, 2, 3, 3, 3, 1, 1, 2, 1, 2, 2,...
## $ Name        <chr> "Kelly, Mr. James", "Wilkes, Mrs. James (Ellen Nee...
## $ Sex         <chr> "male", "female", "male", "male", "female", "male"...
## $ Age         <dbl> 34.5, 47.0, 62.0, 27.0, 22.0, 14.0, 30.0, 26.0, 18...
## $ SibSp       <int> 0, 1, 0, 0, 1, 0, 0, 1, 0, 2, 0, 0, 1, 1, 1, 1, 0,...
## $ Parch       <int> 0, 0, 0, 0, 1, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0,...
## $ Ticket      <chr> "330911", "363272", "240276", "315154", "3101298",...
## $ Fare        <dbl> 7.8292, 7.0000, 9.6875, 8.6625, 12.2875, 9.2250, 7...
## $ Cabin       <chr> "", "", "", "", "", "", "", "", "", "", "B...
## $ Embarked    <chr> "Q", "S", "Q", "S", "S", "S", "Q", "S", "C", "S", ...
```

Como puede observarse, no contiene a la variable Survived. Es decir, que una vez que utilicemos el bosque modelado no podremos elaborar una matriz de confusión ni tampoco evaluar el Accuracy; sin embargo, la predicción que obtengamos nos servirá para participar en la competencia de Kaggle.

```
aplicar$Pclass <- factor(aplicar$Pclass)
prediccion <- predict(bosque, newdata=aplicar, type="class")
write.csv(prediccion, "prediccion.csv")
```

6. Resolución del problema.

A manera de conclusión, podemos afirmar que hemos contruido un bosque cuyo poder de acierto en la predicción es de alrededor del 76%.