1. **Node number 1**: 891 observations, complexity param=0.4444444

predicted class=0 expected loss=0.3838384 P(node) =1

class counts: 549 342

probabilities: 0.616 0.384

left son=2 (577 obs) right son=3 (314 obs)

Primary splits:

sex splits as RL, improve=124.426300, (0 missing)

fare < 10.48125 to the left, improve= 37.941940, (0 missing)

age < 6.5 to the right, improve= 8.814172, (177 missing)

Surrogate splits:

fare < 77.6229 to the left, agree=0.679, adj=0.089, (0 split)

**Node number 2**: 577 observations, complexity param=0.02339181

predicted class=0 expected loss=0.1889081 P(node) =0.647587

class counts: 468 109

probabilities: 0.811 0.189

left son=4 (553 obs) right son=5 (24 obs)

Primary splits:

age < 6.5 to the right, improve=10.78893, (124 missing)

fare < 26.26875 to the left, improve=10.21672, (0 missing)

**Node number 3**: 314 observations

predicted class=1 expected loss=0.2579618 P(node) =0.352413

class counts: 81 233

probabilities: 0.258 0.742

**Node number 4**: 553 observations

predicted class=0 expected loss=0.1681736 P(node) =0.620651

class counts: 460 93

probabilities: 0.832 0.168

**Node number 5**: 24 observations

predicted class=1 expected loss=0.3333333 P(node) =0.02693603

class counts: 8 16

probabilities: 0.333 0.667

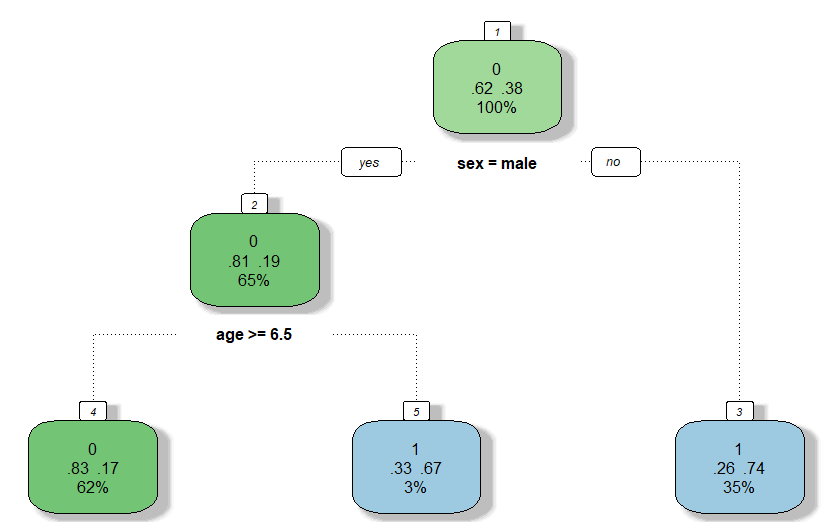
**Node Number 1:** There are 891 observations, the predicted class is 0, class counts is 549 and 342 for probabilities 0.616 and 0.384 respectively. Since predicted class is 0(not survived), so in this case, 549 is not survived, and 342 is survived. For the primary splits, sex(improve=124.4) should be the primary splits since fare have improve of 37.9 and age has improve of 8.8.

**Node Number 2**: There are 577 observations, and the predicted class is 0(not survived), with 468 not survived(81.1%) and 109 survived(18.9%). For primary splits, age is the best selection since it has improve of 10.8 and the other one, which is fare, has improve of 10.2.

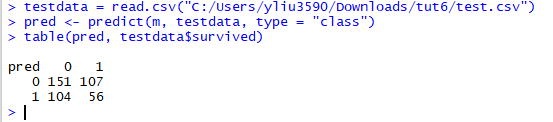
**Node Number 3**: There are 314 observations, and the predicted class is 1(survived), with 233 survived(74.2%) and 81 not survived(25.8%). There is no split in this node.

**Node Number 4**: There are 553 observations, the predicted class is 0 (not survived), there are 460 not survived(83.2%) and 93 survived(16.8%). There is no split in this node.

**Node number 5**: : There are 24 observations, the predicted class is 1 (survived), there are 8 not survived(33.3%) and 16 survived(66.7%). There is no split in this node.

1. 

For each node, the top number is the majority of this node, for 0 as not survived, and 1 as survived. The middle number is the percentage of survived and no survived, according to the top number, you can determine which percentage is survived or not survived (the bigger percentage is the majority of the node). And the bottom number describe the percentage of this node’s observation among the total observation. And below the node is the decision that you could determine which node you should go according to the primary split.

1. 

We use the test file to test our model, and we use the programme above to get the table. From the table the actual class should be the same as the predicted class, but there are 104 and 107 which the actual class is not the same as the predicted class. To calculate the accuracy, we get 0.495, which shows that this is not a good model.