資料輸出

```
#---根據車齡與新增順序---#
> E55 <- read.csv(file = "car4.csv", head=TRUE, sep = ",")
> E55 <- subset(E55 ,select = c(CARTYPE,CARAGE,CARNO,CARCOLOR,ADDR1,COMPA
NY))

> E55$CARAGE <- ifelse(E55$CARAGE == "2005",1,0);
#-----取得總筆數-----#
> n<-nrow(E55)
> set.seed(8000)
#----將數據順序重新排列
> E63<-E55[sample(n),]

#----取出樣本數的 SLR
> SLR <- sample(seq_len(n),size = round(0.7*n))

#----訓練資料與測試資料比例: 70%建模,30%驗證
> trainE55 <- E63[SLR,]
> testE55  <- E63[-SLR,]

#---建立決策樹模型
> dtreeM <- rpart(formula = CARAGE ~ .,data = trainE55,
                method = "class",control = rpart.control(cp = 0.001))

> dtreeM
n=13907

node), split, n, loss, yval, (yprob)
      * denotes terminal node

1) root 13358 2143 0 (0.8395718 0.1604282)
  2) CARNO=0002-T7,0003-KB,0005-K8,0005-WP,0007-F8,0008-G9,0009-LV,0011-E9,001
,... <truncated>
#---
##----
###------
####-------
#####--------(略)
```

```
##################################################################
> S600 <- predict(dtreeM, newdata = testE55, type = "class")
> AMG_GT <- table(testE55$CARAGE, S600, dnn = c("實際", "預測"))
> AMG_GT
     預測
實際    0    1
   0 4819  184
   1  816  141

#---賓果率---#
> AMG_GT [4] / sum(AMG_GT[, 2])
[1] 0.4338462

#---非賓果率---#
> AMG_GT [1] / sum(AMG_GT[, 1])
[1] 0.8551908

#---整體準確率---#
> accuracy <- sum(diag(AMG_GT)) / sum(AMG_GT)
> accuracy
[1] 0.8322148 ≈ 80%




##################################################################

> length(which(E55[,5] == 2005))
[1] 3099

> (length((which(E55[,5] == 2005)))/( sum(table(E55$CARAGE))))
[1] 0.1626516 ≈ 20%




##-各廠牌汽車百分比-##

> length(grep(c("國瑞|豐田|TOYOTA|Toyota|VIOS|CAMRY"),(dt $ CARTYPE))) / length(dt$CARTYPE)
[1] 0.1944934 ≈ 20%
> length(grep(c("中華|三菱"),(dt $ CARTYPE))) / length(dt$CARTYPE)
[1] 0.1478331 ≈ 15%
```

```
> length(grep(c("日產|NISSA"),(dt $ CARTYPE))) / length(dt$CARTYPE)
[1] 0.09357226 ≈ 9%
> length(grep(c("三陽"),(dt $ CARTYPE))) / length(dt$CARTYPE)
[1] 0.03649268 ≈ 4%
> length(grep(c("本田"),(dt $ CARTYPE))) / length(dt$CARTYPE)
[1] 0.02914381 ≈ 3%
> length(grep(c("拖車|曳引車"),(dt $ CARTYPE))) / length(dt$CARTYPE)
[1] 0.01409372 ≈ 1%
> length(grep(c("遊覽"),(dt $ CARTYPE))) / length(dt$CARTYPE)
[1] 0.003976443 ≈ 0.3%
> length(grep(c("賓士|寶馬|BENZ|BMW|benz|bmw|AUDI|audi|保時捷"),(dt $ CARTY
PE))) / length(dt$CARTYPE)
[1] 0.1283536 ≈ 13%


###############################################################################
```