RAZAVET Mael

Exercise 2 – Data Mining

Question 1:

names(iris)

returns

[1] "Sepal.Length" "Sepal.Width" "Petal.Length" "Petal.Width" "Species"

ncol(iris) returns 5. There are 5 different attributes.

nrow(iris) returns 150. There are 150 instances.

Question 2:

summary(iris)

returns

Sepal.Length Sepal.Width Petal.Length Petal.Width Species

Min. :4.300 Min. :2.000 Min. :1.000 Min. :0.100 setosa :50

1st Qu.:5.100 1st Qu.:2.800 1st Qu.:1.600 1st Qu.:0.300 versicolor:50

Median :5.800 Median :3.000 Median :4.350 Median :1.300 virginica :50

Mean :5.843 Mean :3.057 Mean :3.758 Mean :1.199

3rd Qu.:6.400 3rd Qu.:3.300 3rd Qu.:5.100 3rd Qu.:1.800

Max. :7.900 Max. :4.400 Max. :6.900 Max. :2.500

Question 3:

irisSubset <- iris[40:85,]

save(irisSubset, file="irisSubset.RData")

Question 4:

rm(irisSubset)

load("irisSubset.RData")

Question 5:

irisSubset[order(irisSubset$Sepal.Length, decreasing=TRUE),]

Question 6:

irisSubsetSepal <- iris[iris$Sepal.Length < 5.4, ]

Question 7:

max(irisSubsetSepal$Sepal.Length) is 5.3

max(irisSubsetSepal$Sepal.Width) is 4.1

max(irisSubsetSepal$Petal.Length) is 4.5

max(irisSubsetSepal$Petal.Width) is 1.7

max(irisSubsetSepal$Species) is returning an error as this attribute is a categorical variable.

Question 8:

minMaxSpeciesType <- function(specyType, attributeName){

return(list(min(attributeName[iris$Species == specyType]), max(attributeName[iris$Species == specyType])))

}

Example:

minMaxSpeciesType('setosa', iris$Petal.Width)

returns

[[1]]

[1] 0.1

[[2]]

[1] 0.6