HW9.R

xboxv

2020-04-12

```
# Step 1
load(url('https://sites.google.com/a/rdatamining.com/www/data/titanic.raw.rdata?attredirects=1'))
t <-titanic.raw
total <-nrow(t)

#Compute the percentage of people that survived.
surv <-sum(t$Survived =="Yes")
percsurvived <-((surv/total)*100)
percsurvived

## [1] 32.3035

#Compute the percentage of people that were children.
childcount <-sum(t$Age =="Child")
childpct <-((childcount/total)*100)
childpct

## [1] 4.952294

#Compute the percentage of people that were female.
femalecount <-sum(t$Sex =="Female")
femalepct <-((femalecount/total)*100)
femalepct

## [1] 21.35393

#Finally, compute the percentage of people that were in first class.
fccount <-sum(t$Class =="1st")
fcpct <- ((fccount/total)*100)
fcpct

## [1] 14.76602

# Step 2
#What percentage of children survived?
childrenSurvived <- sum(t$Age=="Child"&t$Survived=="Yes")/sum(t$Age=="Child")*100
childrenSurvived

## [1] 52.29358

#What percentage of female survived?
femalesurvived <- sum(t$Sex=="Female"&t$Survived=="Yes")/sum(t$Sex=="Female")
```

```r
*100
femalesurvived

## [1] 73.19149

#What percentage of first-class people survived?
firstSurvived <- sum(t$Class=="1st"&t$Survived=="Yes")/sum(t$Class=="1st")*10
0
firstSurvived

## [1] 62.46154

#What percentage of third-class people survived?
Thirdsurvived <- sum(t$Class=="3rd"&t$Survived=="Yes")/sum(t$Class=="3rd")*10
0
Thirdsurvived

## [1] 25.21246

# Step 3 writing a function
# fucking no clue...

myFunction <- function(class,sex,age,surv)
{
  i <- 0
  for (i in length(t)){
    if(class=='1st' && sex=='Female' && age=='Adult' && surv=='Yes'){
      df <- data.frame(class,sex,age,surv)
      return(df)
    }
  }
}

#Function calling with different arguments to check the functionality of it
myFunction('1st','Female','Adult','Yes')

##   class    sex   age surv
## 1   1st Female Adult  Yes

myFunction('1st','Male','Child','Yes')
myFunction('3rd','Female','Adult','No')


percentFunction <- function()
{
  num <- length(df)
  return((num/2201)*100)
}

percentFunction()
```

```
## [1] 0.04543389
```

```r
#step 4 using aRules

library(DT)
```

```
## Warning: package 'DT' was built under R version 3.6.3
```

```r
library(data.table)
```

```
## Warning: package 'data.table' was built under R version 3.6.3
```

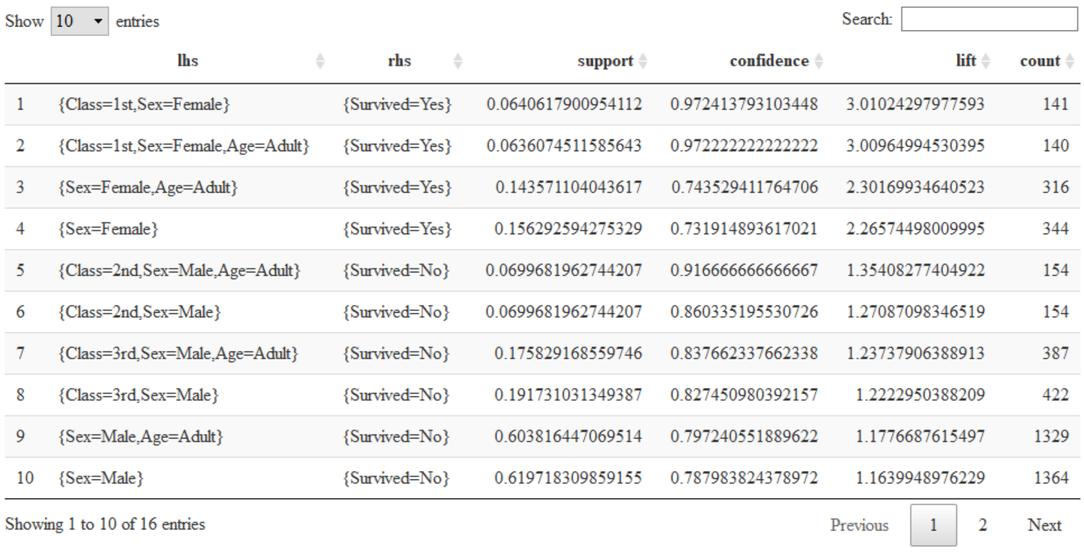
```r
library(arules)
```

```
## Warning: package 'arules' was built under R version 3.6.3

## Loading required package: Matrix

##
## Attaching package: 'arules'

## The following objects are masked from 'package:base':
##
##     abbreviate, write
```

```r
titanic <- as(t, 'transactions')
summary( itemFrequency(titanic) )
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.04952 0.16413 0.32190 0.40000 0.60825 0.95048
```

```r
# train apriori
rules <- apriori(titanic,parameter = list(support = 0.05, confidence = 0.7, minlen = 2, maxlen = 5), appearance = list(rhs = c('Survived=No', 'Survived=Yes'), default = 'lhs'),control = list(verbose = FALSE))

# convert object of class rules to data frame
rules_dt <- data.table(lhs = labels(lhs(rules)), rhs = labels(rhs(rules)), quality(rules) )[order(-lift),]
DT::datatable(rules_dt)
```



Show 10 entries

Search:

	lhs	rhs	support	confidence	lift	count
1	{Class=1st,Sex=Female}	{Survived=Yes}	0.0640617900954112	0.972413793103448	3.01024297977593	141
2	{Class=1st,Sex=Female,Age=Adult}	{Survived=Yes}	0.0636074511585643	0.972222222222222	3.00964994530395	140
3	{Sex=Female,Age=Adult}	{Survived=Yes}	0.143571104043617	0.743529411764706	2.30169934640523	316
4	{Sex=Female}	{Survived=Yes}	0.156292594275329	0.731914893617021	2.26574498009995	344
5	{Class=2nd,Sex=Male,Age=Adult}	{Survived=No}	0.0699681962744207	0.916666666666667	1.35408277404922	154
6	{Class=2nd,Sex=Male}	{Survived=No}	0.0699681962744207	0.860335195530726	1.27087098346519	154
7	{Class=3rd,Sex=Male,Age=Adult}	{Survived=No}	0.175829168559746	0.837662337662338	1.23737906388913	387
8	{Class=3rd,Sex=Male}	{Survived=No}	0.191731031349387	0.827450980392157	1.2222950388209	422
9	{Sex=Male,Age=Adult}	{Survived=No}	0.603816447069514	0.797240551889622	1.1776687615497	1329
10	{Sex=Male}	{Survived=No}	0.619718309859155	0.787983824378972	1.1639948976229	1364

Showing 1 to 10 of 16 entries

Previous 1 2 Next

```
#Three most important and useful rules

#Contains a value for lift which is large enough to make this a good rule, an
d conditions which will be a valid condition to check

# {Class=1st,Survived=Yes}  => {Age=Adult}    0.08950477 0.9704433  1.0210066

#Contains large value for lift, and conditions which will be a valid conditio
n to check

# {Class=3rd,Survived=No}   => {Age=Adult}    0.21626533 0.9015152  0.9484870

#Contains maximum value for lift, and conditions which will be a valid condit
ion to check

# {Sex=Female,Age=Adult}    => {Survived=Yes} 0.14357110 0.7435294  2.3016993
```