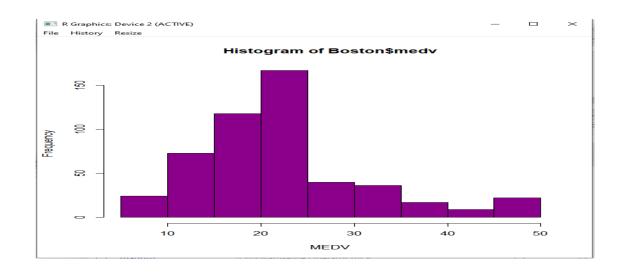
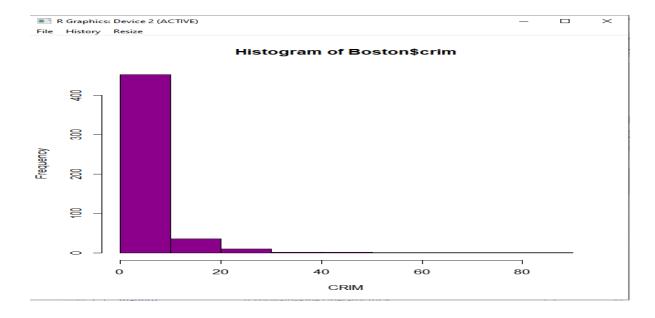
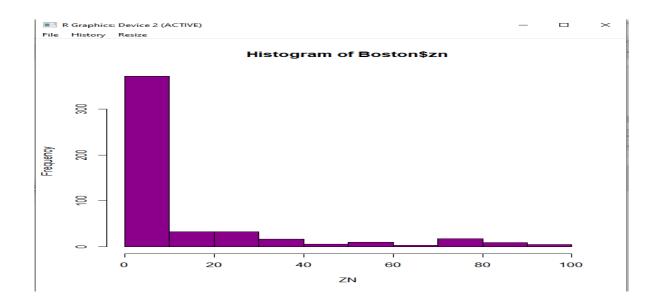
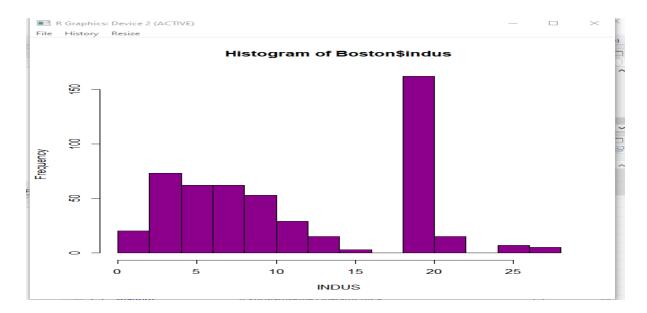
## Question 2.

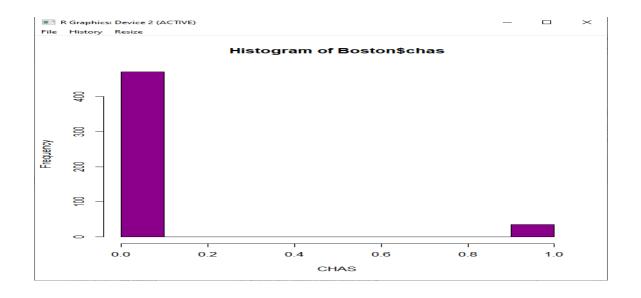
a)

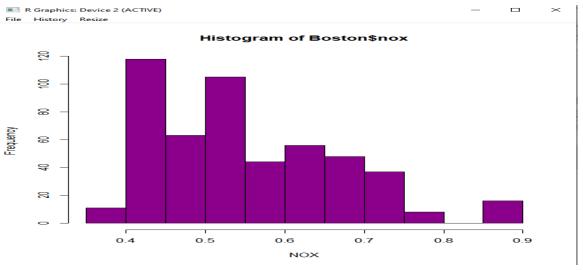


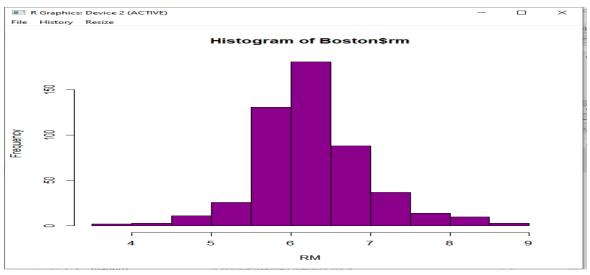


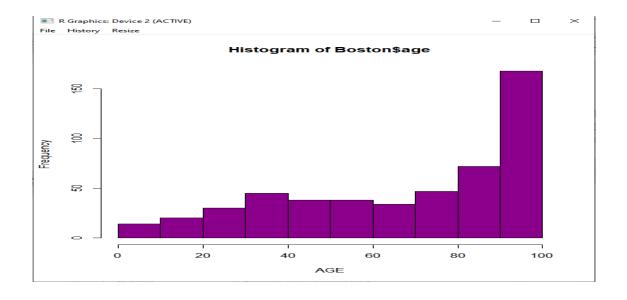


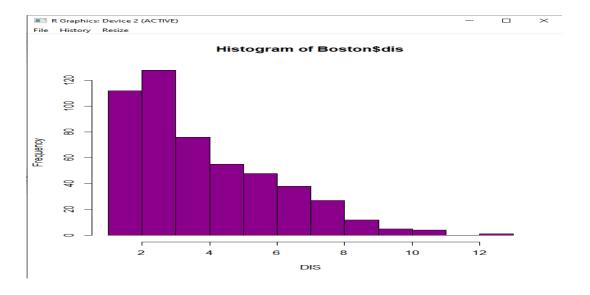


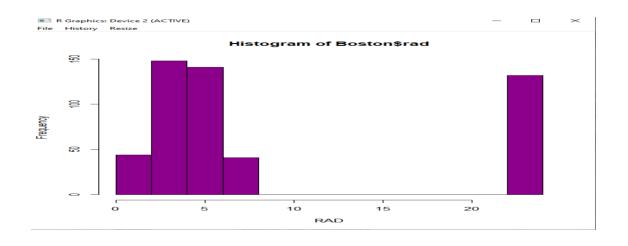


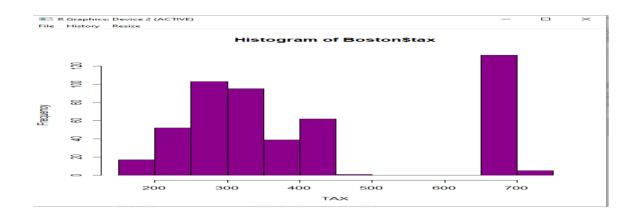


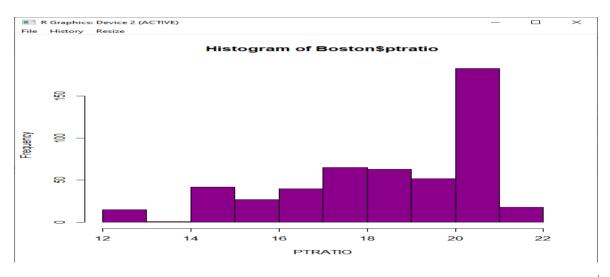


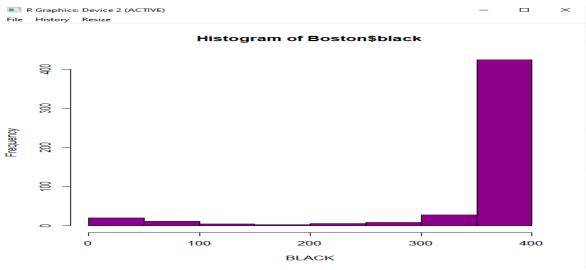


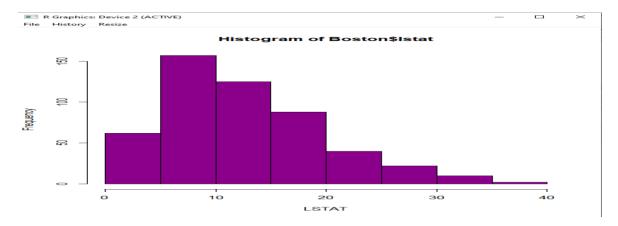












Values for grouping categories were chosen based on mean , quartile, maximum and mininum values of the column data

```
Console ~/ 🙈
> summary(Boston)
     crim
                                               indus
                                                                  chas
Min. : 0.00632
1st Qu.: 0.08204
                                          Min. : 0.46
1st Qu.: 5.19
                                                                    :0.00000
                      Min.
                                  0.00
                                                            Min.
                                                                                 Min.
                                                                                         :0.3850
                              :
                                                            1st Qu.: 0.00000
                      1st Qu.:
                                  0.00
                                                                                 1st Qu.:0.4490
 Median : 0.25651
                      Median :
                                  0.00
                                          Median: 9.69
                                                            Median :0.00000
                                                                                 Median :0.5380
Mean : 3.61352
3rd Qu.: 3.67708
                      Mean
                              : 11.36
                                          Mean
                                                 :11.14
                                                            Mean
                                                                    :0.06917
                                                                                 Mean
                                                                                         :0.5547
                      3rd Qu.: 12.50
                                          3rd Qu.:18.10
                                                            3rd Qu.: 0.00000
                                                                                 3rd Qu.: 0.6240
        :88.97620
                              :100.00
                                                  :27.74
                                                                    :1.00000
                                                                                 мах.
                                                                                         :0.8710
                      Max.
                                          Max.
 Max.
                                                            Max.
       rm
                                            dis
                                                               rad
                                                                                   tax
                                                                                                   ptratio
                        age
                                      Min. : 1.130
1st Qu.: 2.100
Min.
                   Min.
                                                          Min.
                                                                             Min.
                                                                                               Min. :12.60
1st Qu.:17.40
Min. :3.561
1st Qu.:5.886
                   Min. : 2.90
1st Qu.: 45.02
                                                                 : 1.000
                                                                                     :187.0
                                                          1st Ou.: 4.000
                                                                             1st Ou.:279.0
                             77.50
                                      Median : 3.207
 Median :6.208
                   Median :
                                                          Median :
                                                                    5.000
                                                                             Median :330.0
                                                                                                Median :19.05
        :6.285
                   Mean
                             68.57
                                      Mean
                                              : 3.795
                                                          Mean
                                                                  : 9.549
                                                                             Mean
                                                                                    :408.2
                                                                                                Mean
                                                                                                        :18.46
                   3rd Qu.: 94.08
Max. :100.00
                                      3rd Qu.: 5.188
Max. :12.127
 3rd Qu.:6.623
                                                          3rd Qu.:24.000
                                                                             3rd Qu.:666.0
                                                                                                3rd Qu.:20.20
                                                                  :24.000
                                                                                      :711.0
         :8.780
                   Max.
                                      Max.
                                                          мах.
                                                                             мах.
                                                                                                мах.
                                                                                                        :22.00
Max.
     hlack
```

```
Console ~/ 🙈
                                                 Mean : 3.207
                                                                          meuran .
                         meuran .
                                                                                                    euran .
                                                                                                                            euran .15.
                                  : 68.57
                                                                                    : 9.549
                        Mean
                                                 Mean
                                                                          Mean
                                                                                                   Mean
                                                                                                            :408.2
                                                                                                                          Mean
                                                                                                                                    :18.46
                        3rd Qu.: 94.08
                                                 3rd Qu.: 5.188
                                                                          3rd Qu.:24.000
 3rd Qu.:6.623
                                                                                                   3rd Qu.:666.0
                                                                                                                          3rd Qu.:20.20
Max. ..
black
           :8.780
                                  :100.00
                                                            :12.127
                                                                                    :24.000
                                                                                                             :711.0
                               lstat
                                                       medv
                         Min. : 1.73
1st Qu.: 6.95
 Min.
               0.32
                                                 Min.
 1st Qu.:375.38
                                                 1st Qu.:17.02
 Median :391.44
                         Median :11.36
Mean :12.65
                                                 Median :21.20
          :356.67
 Mean
                                                 Mean :22.53
 3rd Qu.:396.23
                          3rd Qu.:16.95
                                                 3rd Qu.:25.00
Max. :396.90 Max. :37.97 Max. :50.00

> Boston[["crim"]] <- ordered(cut(Boston[["crim"]], c(0,0.1,3.61,5,90)), labels = c('verylowcrime','lowcrime','moderatecrime','highcrime'))

> Boston[["zn"]] <- ordered(cut(Boston[["zn"]], c(0,11,12,20,100)), labels = c('nonresidential','smallresidence','moderateresidence','largeresidence'))
```

b) Item frequency Plot of binary incidence matrix of Boston housing data is as follows:

## Parameters considered:

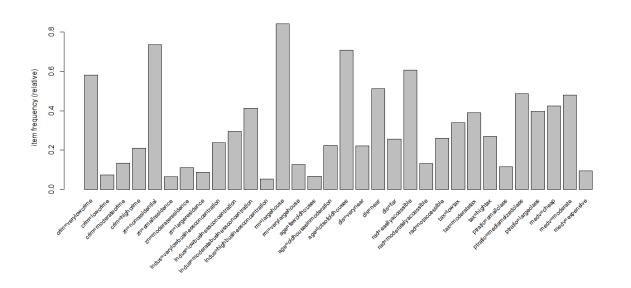
1. Support = 0.05

■ R Graphics: Device 2 (ACTIVE)

— 

X

File History Resize



After Applying apriori algorithm with following parameters:

Support = 0.02, confidence = 0.8

```
c)
  Console ~/ 🙈
                                                                                                         \neg
 > summary(rulesLowCrimeNearCity)
 set of 0 rules
 > summary(rules_LowCrime)
 set of 2648 rules
 rule length distribution (lhs + rhs):sizes
   2 3 4 5 6 7 8 9 10
9 116 444 798 760 395 108 17 1
    Min. 1st Qu. Median
                            Mean 3rd Qu.
                                             мах.
    2.00 5.00
                  5.00
                            5.47
                                    6.00
 summary of quality measures:
                                           lift
     support
                      confidence
                                                          count
                                           :1.377
                                                      Min. : 11.0
1st Qu.: 14.0
   Min. :0.02174
                     Min. :0.8000
                                     Min.
   1st Qu.:0.02767
                    1st Qu.:1.0000
                                     1st Qu.:1.721
   Median :0.03755
                    Median :1.0000
                                     Median :1.721
                                                      Median: 19.0
         :0.05020
                    Mean :0.9776
                                     Mean :1.683
                                                      Mean : 25.4
  Mean
   3rd Qu.:0.05534
                                     3rd Qu.:1.721
                                                      3rd Qu.: 28.0
                    3rd Qu.:1.0000
         :0.47036
                           :1.0000
  Max.
                    Max.
                                     Max.
                                             :1.721
                                                      Max.
                                                            :238.0
 mining info:
       data ntransactions support confidence
  boston_tr
                      506
                             0.02
                                          0.8
```

Student should choose house away from his work as crime rate decreases.

d)

```
> summary(rulesLowPupil_TeacherRatio)
set of 296 rules
rule length distribution (lhs + rhs):sizes
 3 4 5 6 7 8 9 10
 6 29 64 84 69 34 9 1
   Min. 1st Qu. Median
                           Mean 3rd Ou.
                                           Max.
  3.000 5.000 6.000 6.095 7.000 10.000
summary of quality measures:
    support
                     confidence
                                         lift
 Min. :0.02174
                   Min. :0.8000 Min. :6.979
1st Qu.:0.9333 1st Qu.:8.143
Median :1.0000 Median :8.724
                                                    Min. :11.00
 1st Qu.:12.00
                                                    Median :14.00
 Mean :0.03149 Mean :0.9636
3rd Qu.:0.03360 3rd Qu.:1.0000
                                    Mean :8.407
                                                    Mean :15.93
                                    3rd Qu.:8.724
                                                    3rd Qu.:17.00
       :0.05929 Max. :1.0000
                                                          :30.00
 Max.
                                   Max.
                                           :8.724
                                                    Max.
mining info:
      data ntransactions support confidence
 boston_tr
                     506
                            0.02
> inspect(head(sort(rulesLowPupil_TeacherRatio, by ='lift'),n = 6))
                                                                    support confidence
                                                                                           lift count
[1] {crim=lowcrime,
                                         => {ptratio=smallclass} 0.02371542
                                                                                     1 8.724138
    zn=smallresidence}
                                                                                                   12
[2] {crim=lowcrime,
     indus=verylowbusinessconcentration} => {ptratio=smallclass} 0.02371542
                                                                                     1 8.724138
                                                                                                   12
[3] {crim=lowcrime,
                                         => {ptratio=smallclass} 0.02371542
     tax=lowtax}
                                                                                     1 8.724138
                                                                                                   12
[4] {crim=lowcrime,
     zn=smallresidence,
     indus=verylowbusinessconcentration} => {ptratio=smallclass} 0.02371542
[5] {crim=lowcrime,
```

He would get small residence for low crime . for getting school with low teacher-pupil ratio they would get small residence. Chances of getting a good residence is very low.

=> {ptratio=smallclass} 0.02371542

=> {ptratio=smallclass} 0.02371542

1 8.724138

1 8.724138

12

12

zn=smallresidence,

zn=smallresidence.

rad=easilyaccessible}

tax=lowtax}

[6] {crim=lowcrime,

## Question 3

A random data that resembles the original data with different column 'flag' . A flag was 0 for reference data and 1 for original data.

Decision Tree model was built.