Using R or Python, preprocess your dataset to run with Apriori and then run the Apriori algorithm.  Change the support and confidence measures until you get optimal results.  Then answer the following questions on one single-spaced page:

1. How did you adjust support and confidence to achieve optimal results?

Student Alcohol Consumption is consisting with 649 observations and with 33 variables, where the data set was related to student education and grading, which factors are impacting the grade of the student. In fact, to make a closer relation and define aa best relations and minimize the vision of the dataset, we identified the cor and cov of the dataset and identified the best related attributes which helps in identifying the best output helps the students to improve the society to make a student and make government to implement best practices to make student educated. We used Chisq test , and provided the ab line to get the best relations and which what factors relates to best matters relates to result. Now as the dataset, being a good minner, it is always important to reduce the dataset to increase the dataset and observation, to make the dataset most effective, to get effective results and in order to identify which value collides and interprets best value in resolving best results, to make them most effective, and make them identified with apriori algorithm and generating rules based on confidence and support we can identify the best attributes as a group best related and make them results with best results regarding the question asked by people and government. In this Dataset, we had taken min len as 2, and support to 0.3 and confidence s 0.97 because to resolve with minimal results reason was to get as much as less results and make them effective to output and if the results are less collecting more attributes then we consider that your are making on right track for results. We made couple of combination with length ,support and confidence and identified this identifies the best minimal results which would suffice the requirement.   
  
code:

str(student\_por)

data1 = student\_por[-1]

col\_names <- names(data1)

data1[,col\_names] <- lapply(data1[,col\_names] , factor)

rules.all<-apriori(data1)

inspect(rules.)

rules <- apriori(data1,control = list(verbose=F),

parameter = list(minlen=2, supp=0.3, conf=0.97))

quality(rules) <- round(quality(rules), digits=3)

rules.sorted <- sort(rules, by="confidence")

inspect(rules.sorted)

subset.matrix <- is.subset(rules.sorted, rules.sorted)

subset.matrix[lower.tri(subset.matrix, diag = T)] <- NA

redundant <- colSums(subset.matrix, na.rm = T) >= 1

which(redundant)

rules.pruned <- rules.sorted[!redundant]

inspect(rules.pruned)

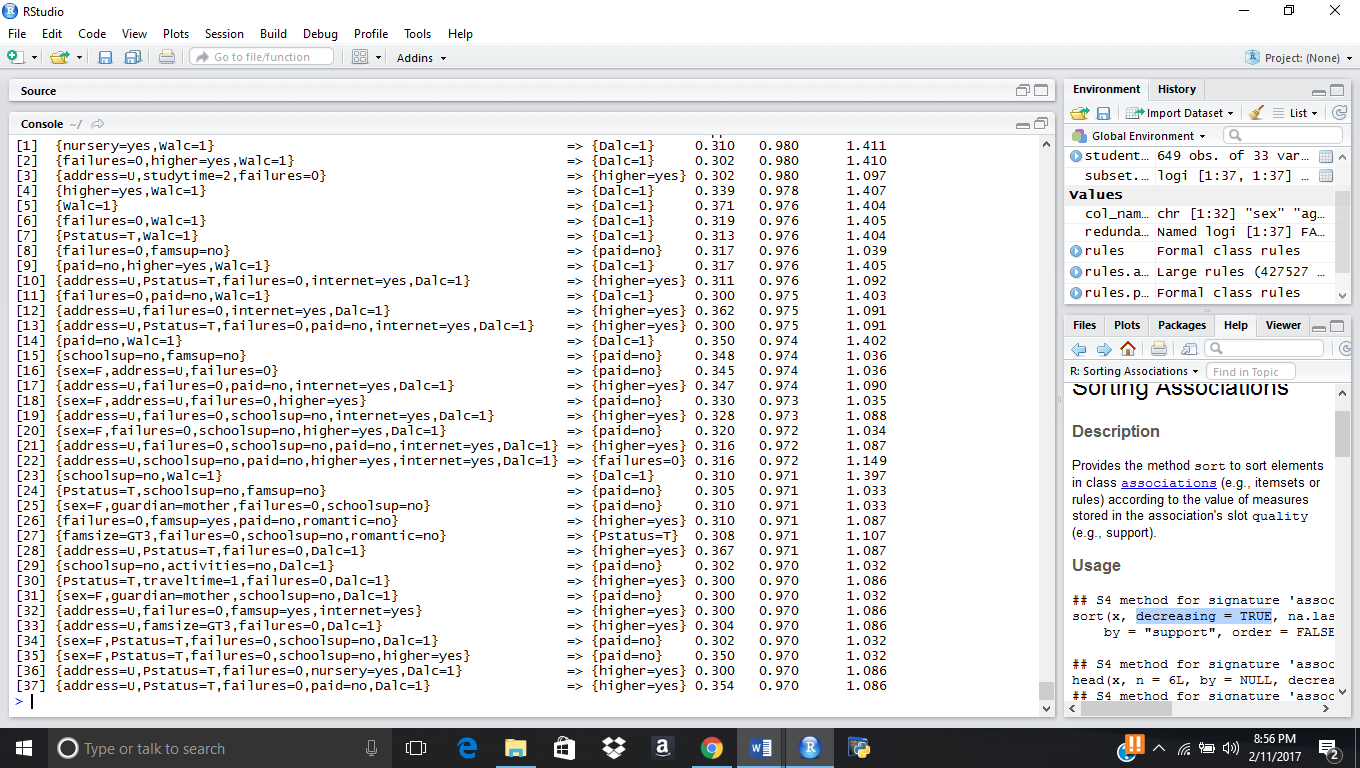
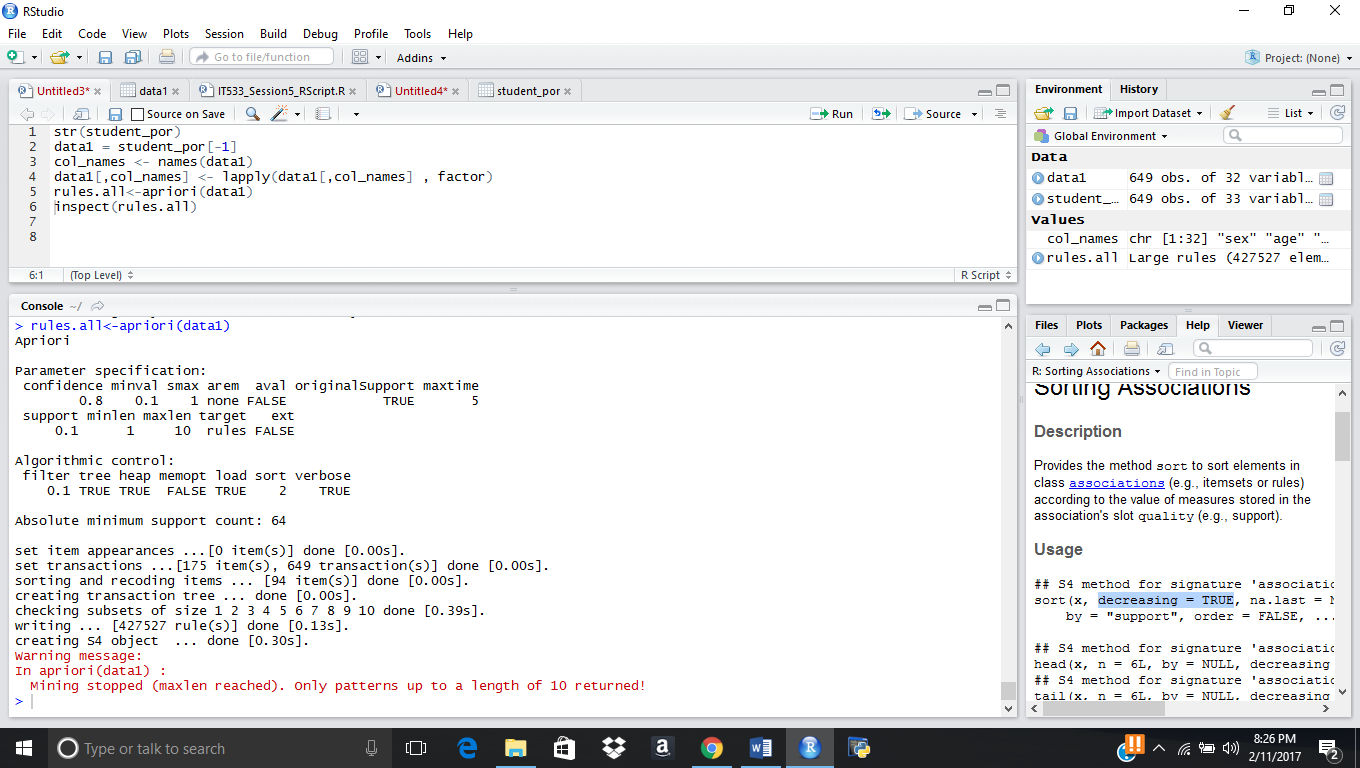
inspect(rules.pruned[1:5])

plot(rules)

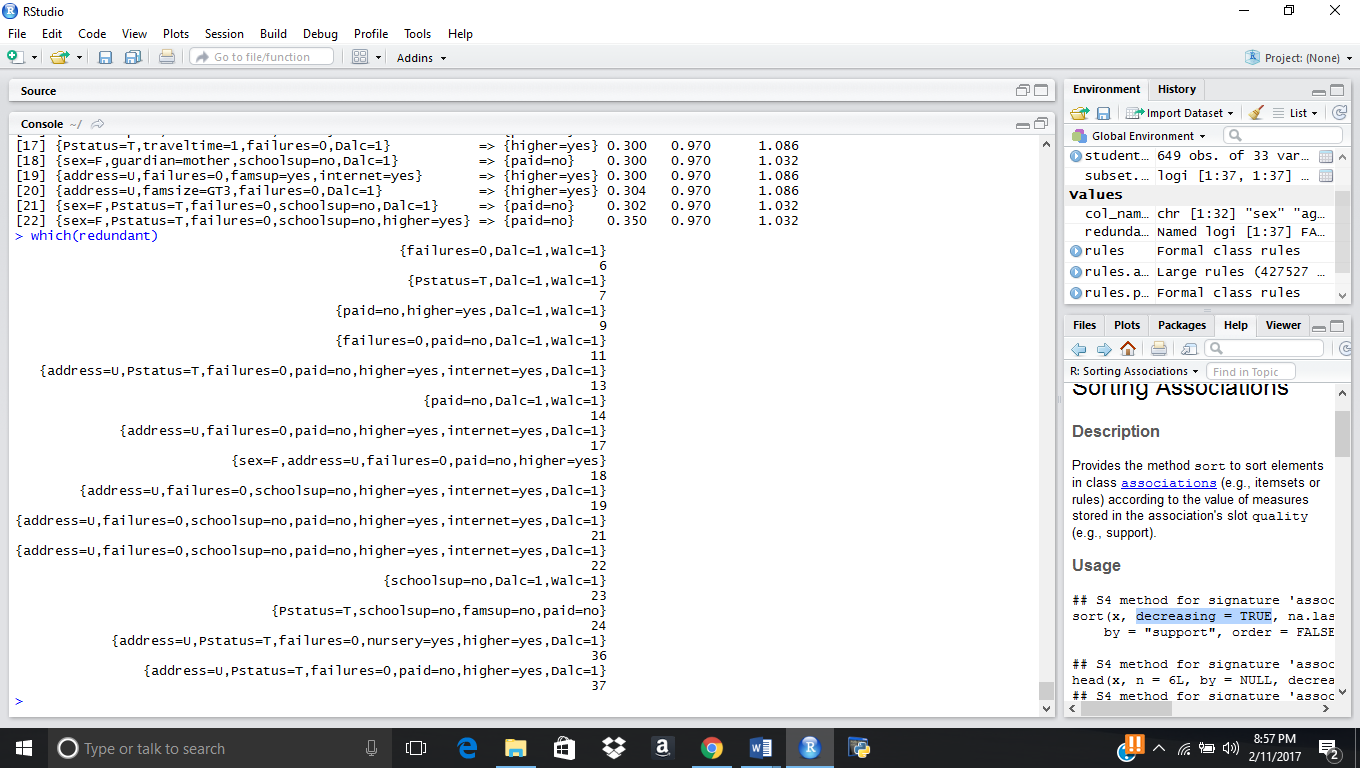
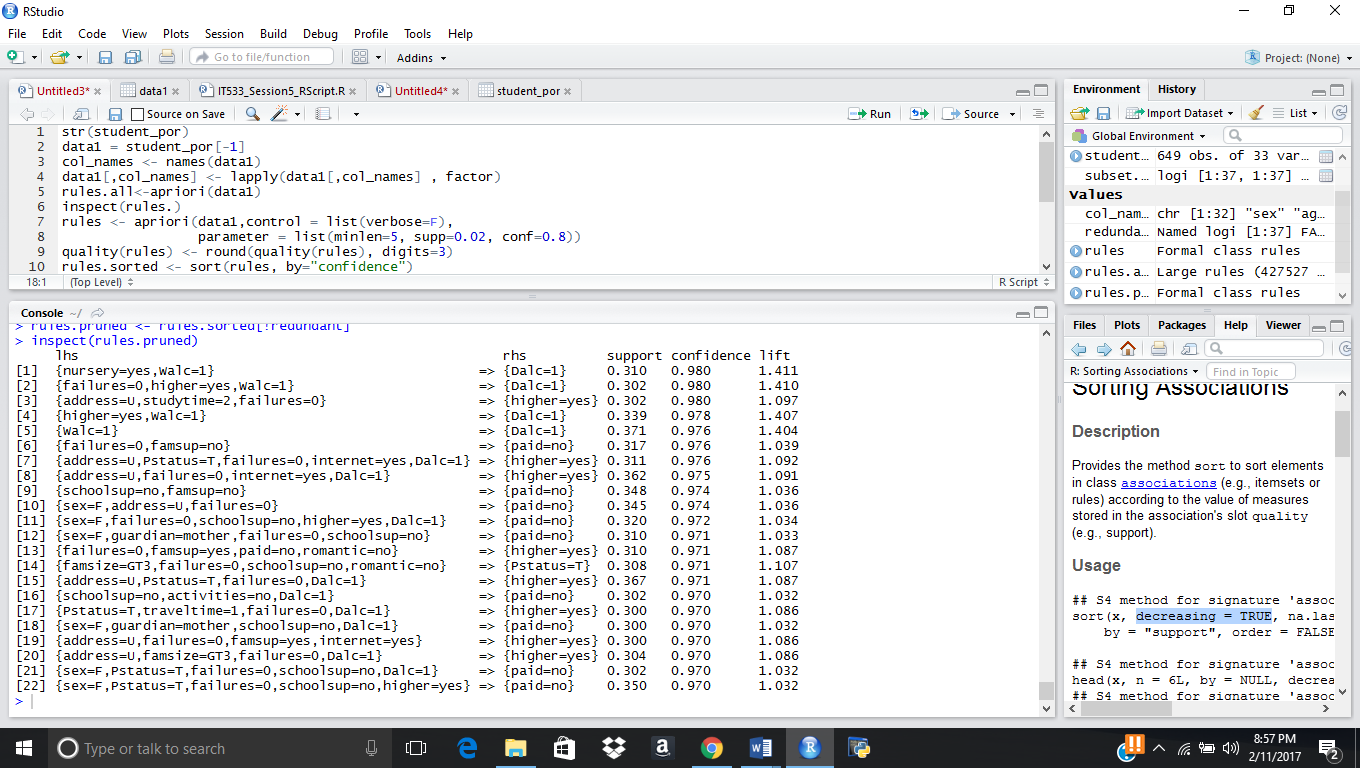
plot(rules, method = "grouped")

plot(rules, method = "graph")

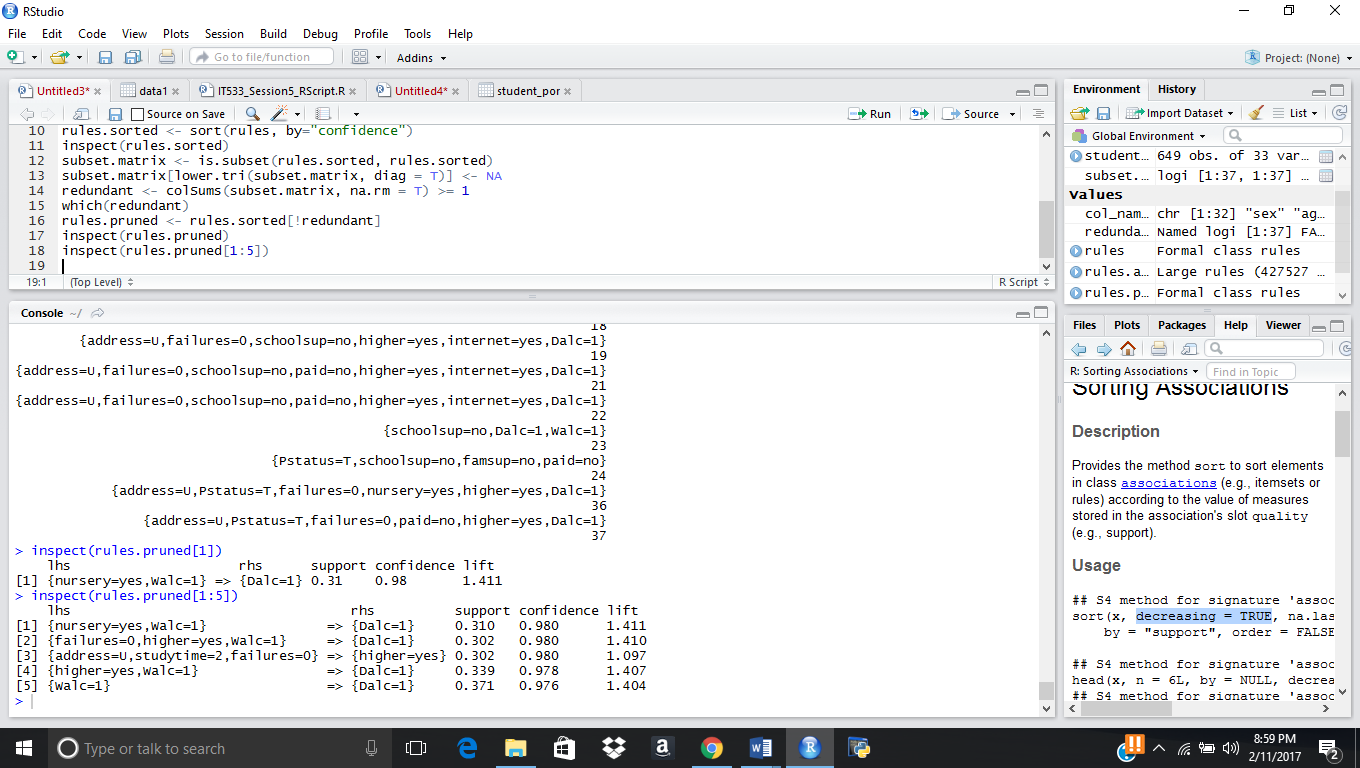
plot(rules, method = "paracoord", control = list(reorder = TRUE))  
  
By taking Default Value like without any constraint in consideration below is the screenshot which resolves with results.

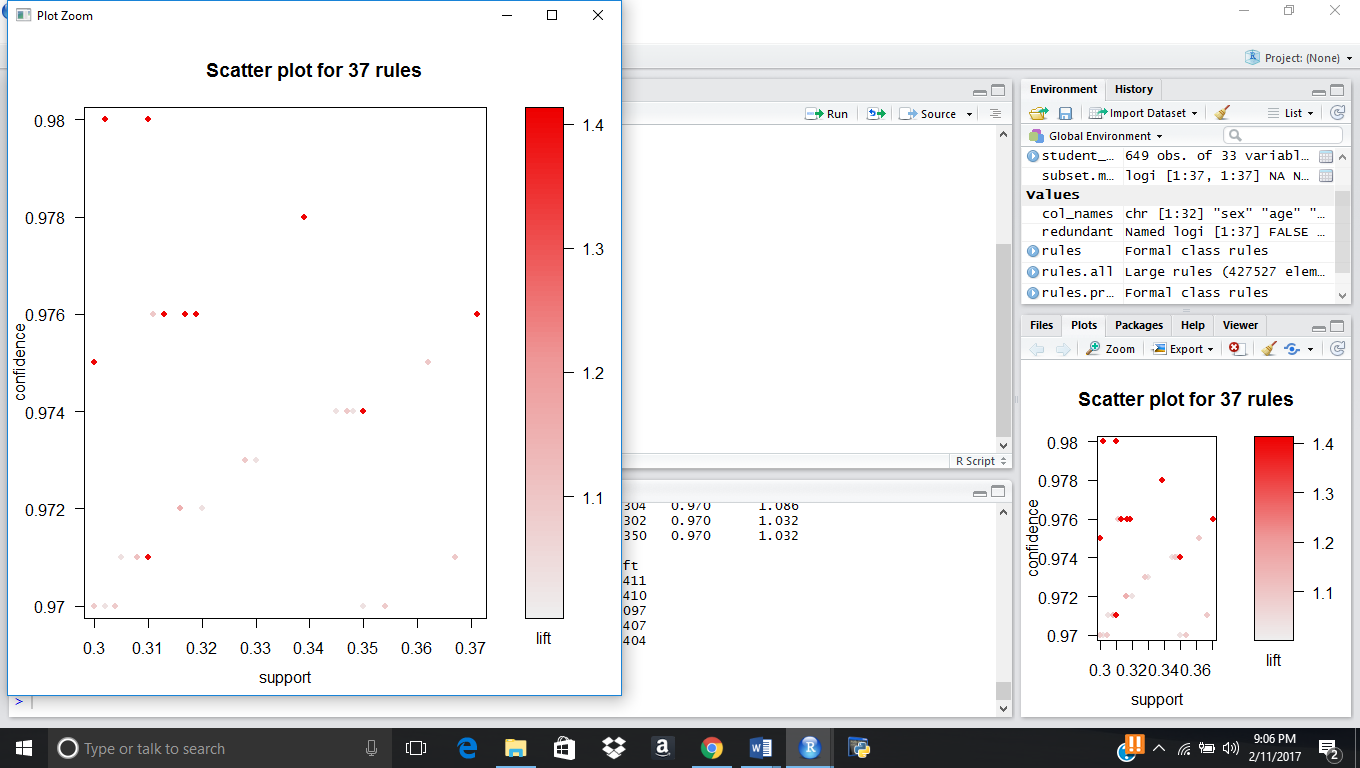
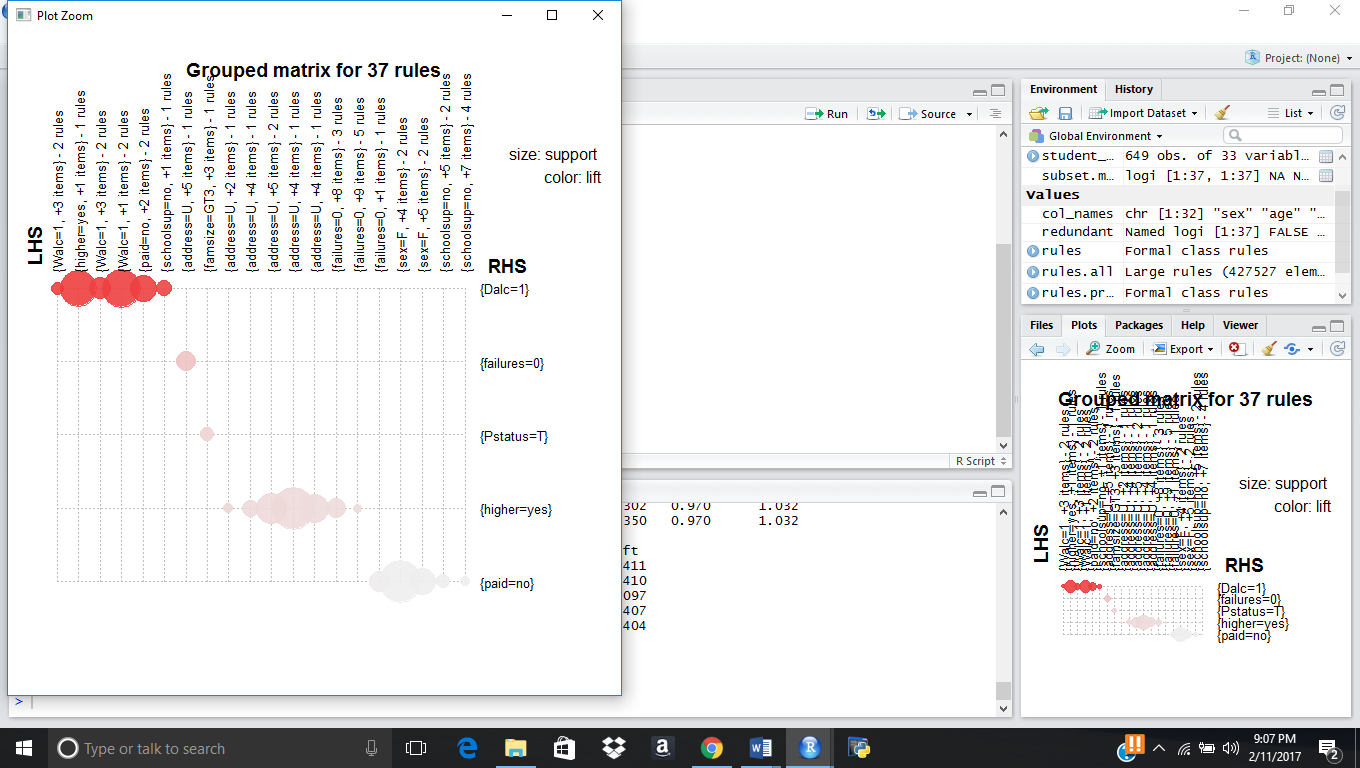


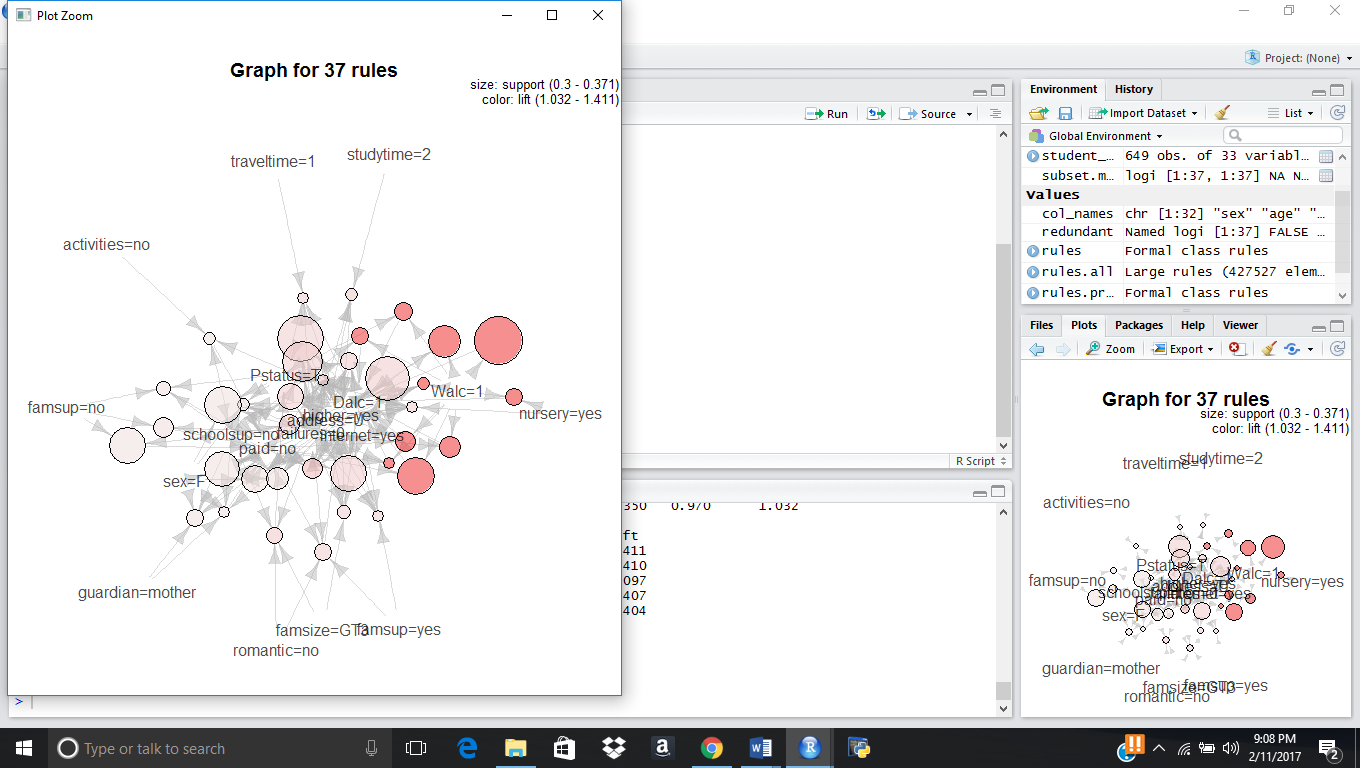
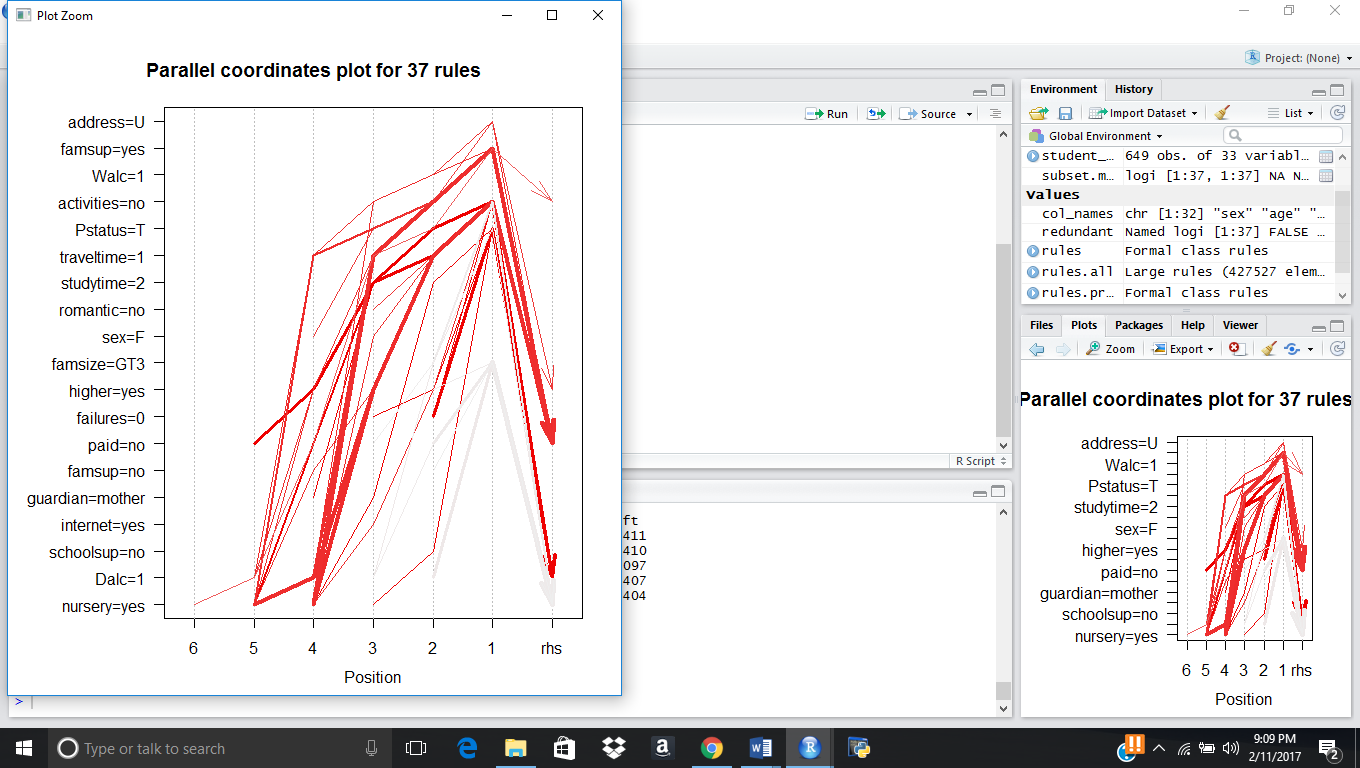
After applying the rules of apriori algorithm it was resolved with this output which was identified with screenshot when condition is added.



After the redundancy it was identified that it was decreased to 22 attributes and where when cor and cov was been identified we where able to interpret the same output what we are able to do with this algorithm and results with length of 2 and redundant value was been removed.



Here the rules are been pruned to get very near to identify the value, and make them plotted with graph.  
  




2.Show your work.  What do the results tell you about the data?  
  
 In order to identify best results for students we had identify with health factor of student how it is impact with education and grade of the student, if the health is 5 and then taking with 5 values of the output , we identified with confidence is 1 which mean student with high health is having to have good grade, and now we can use common sense, which factor defact the health to make grade reduced. Attaching the screenshot and script we visualized.  **R Code:**  
  
rules <- apriori(data1,

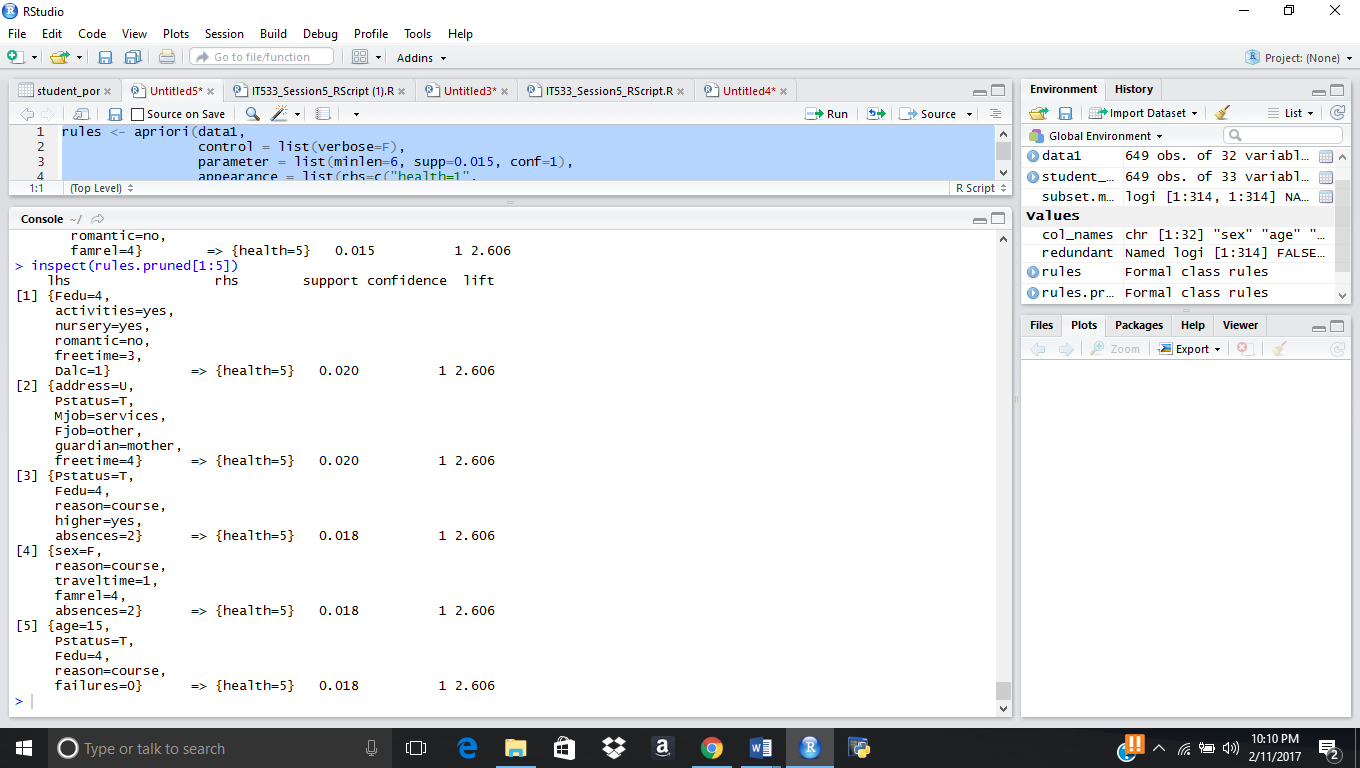
control = list(verbose=F),

parameter = list(minlen=6, supp=0.015, conf=1),

appearance = list(rhs=c("health=1",

"health=2","health=3",

"health=4","health=5"),

 default="lhs"))

Varying with different len and support and con we interpret different results and making the varied changes ot it, it decreases the data set values make it more effective dataset.

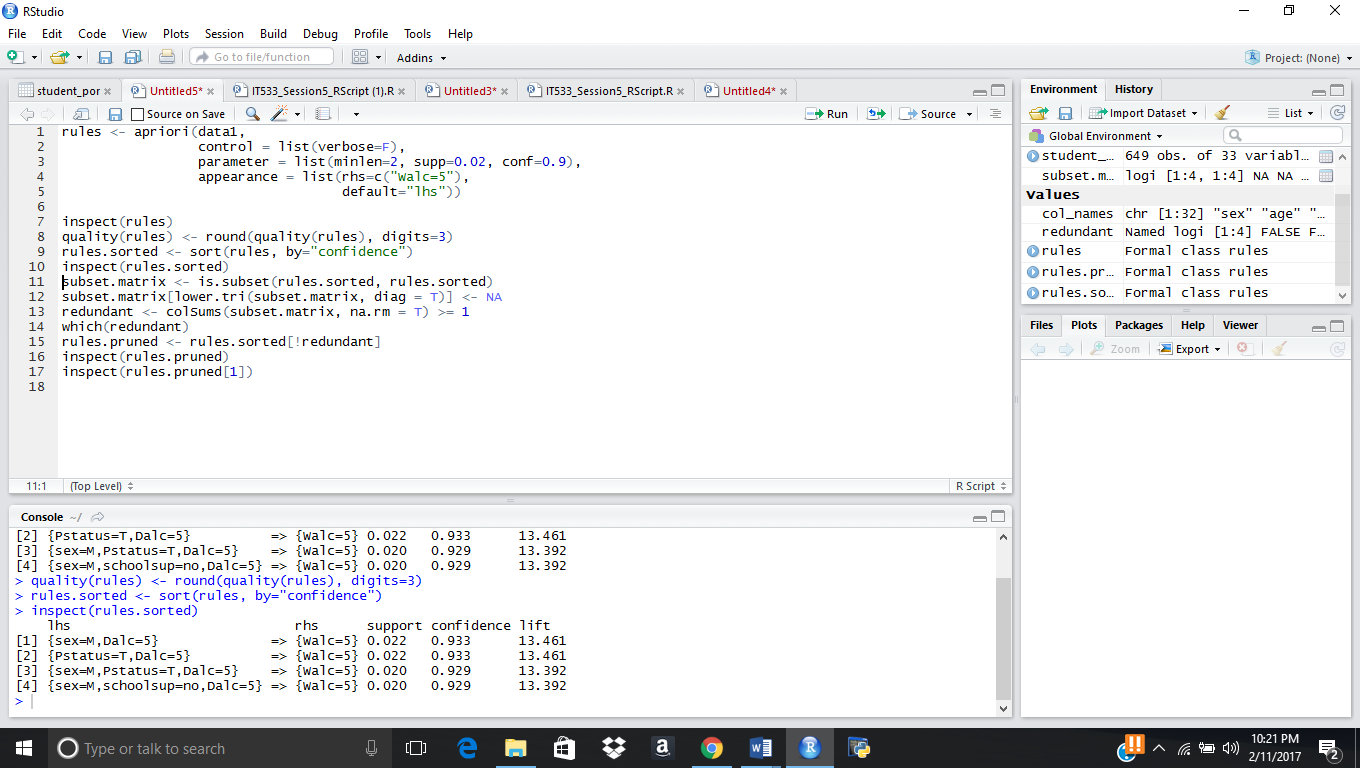
rules <- apriori(data1,

control = list(verbose=F),

parameter = list(minlen=2, supp=0.02, conf=0.9),

appearance = list(rhs=c("Walc=5"),

default="lhs"))



lhs rhs support confidence lift

[1] {sex=M,Dalc=5} => {Walc=5} 0.022 0.933 13.461

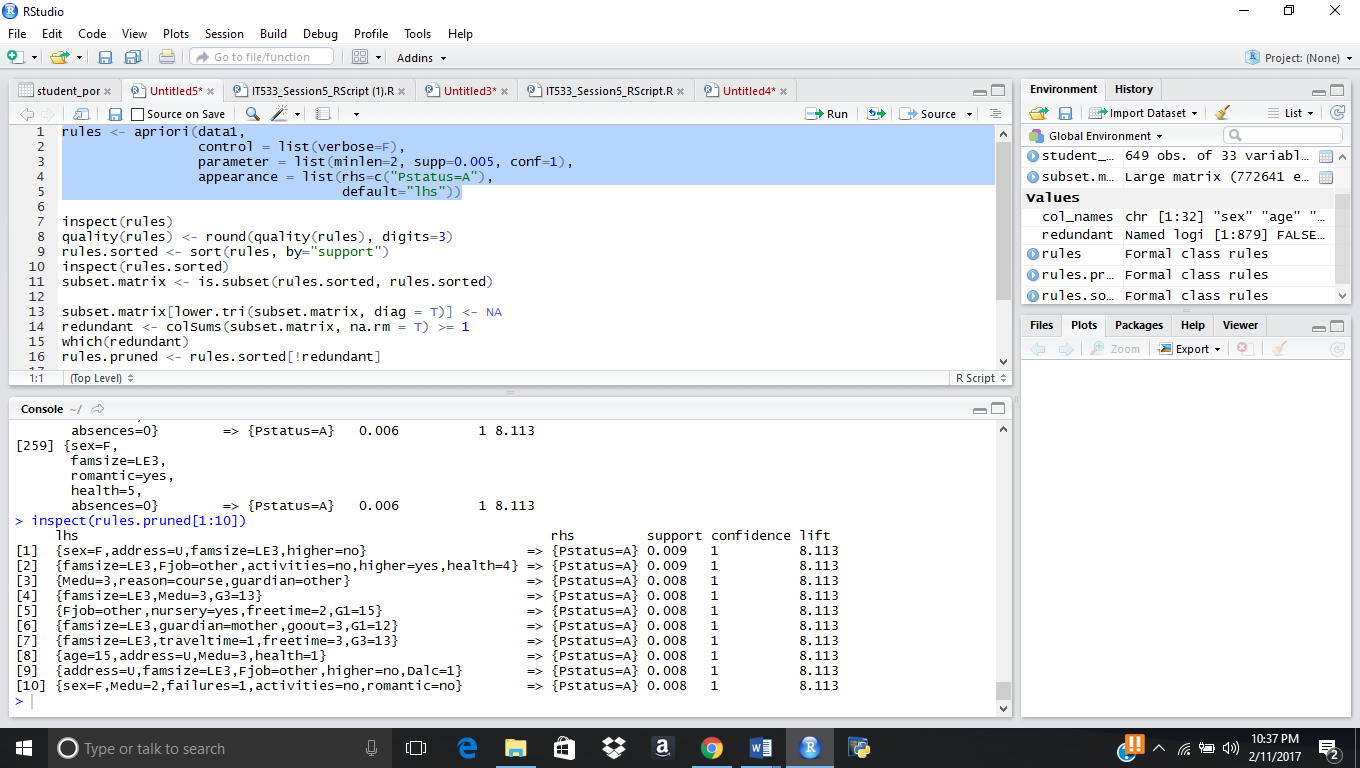
Interpreting with Different values to get closer and identified with lift to get greater impact.

rules <- apriori(data1,

control = list(verbose=F),

parameter = list(minlen=2, supp=0.005, conf=1),

appearance = list(rhs=c("Pstatus=A"),

 default="lhs"))

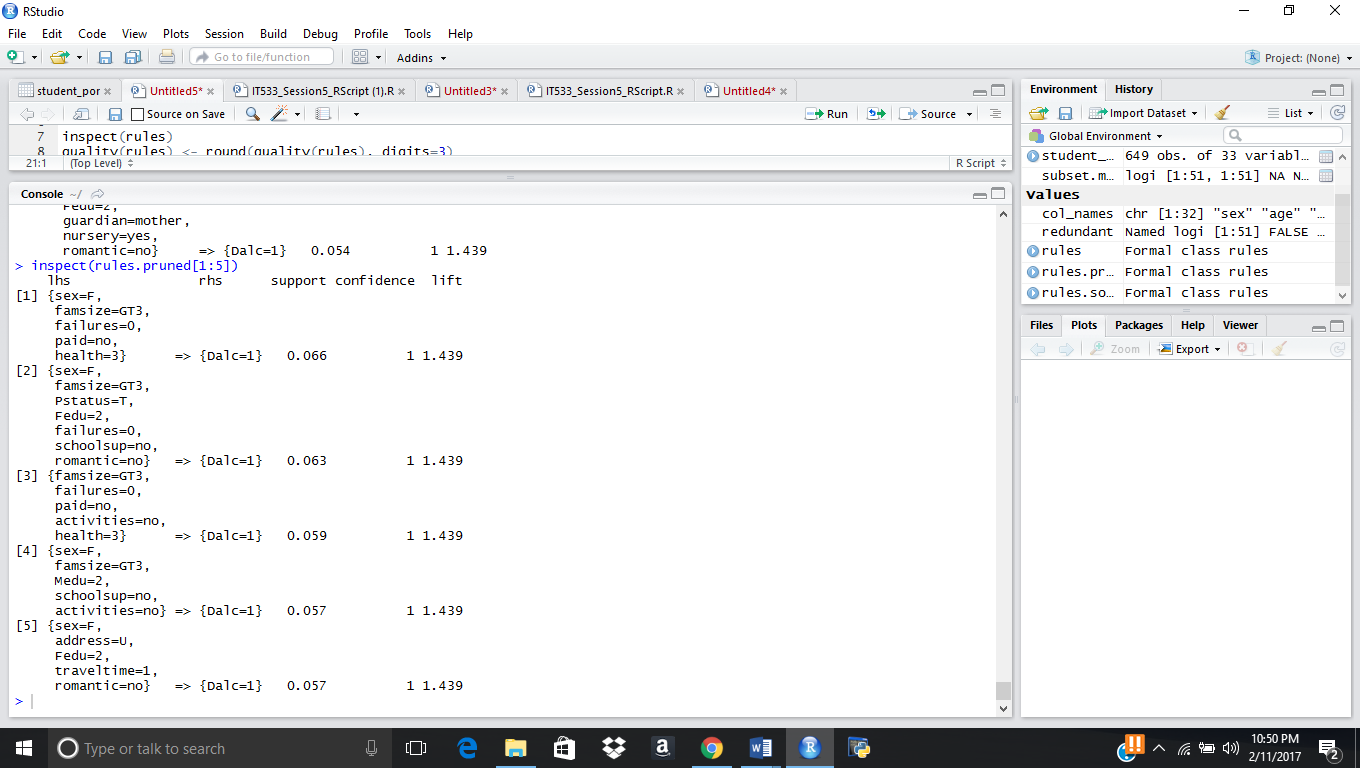
rules <- apriori(data1,

control = list(verbose=F),

parameter = list(minlen=2, supp=0.05, conf=1),

appearance = list(rhs=c("Walc=1","Dalc=1"),

default="lhs"))



Above are the results with pruned and make them with closer values and used for interpreter by changing with different attributes confidence and support with best association rules.