PART B

(PART B: TO BE COMPLETED BY STUDENTS)

***(Students must submit the soft copy as per following segments within two hours of the practical. The soft copy must be uploaded on the Blackboard or emailed to the concerned lab in charge faculties at the end of the practical in case the there is no Black board access available)***

|  |  |
| --- | --- |
| Roll No.: B032 | Name: Naman Garg |
| Class: B Tech CS (B div) | Batch: B2 |
| Date of Experiment: 4/9/2020 | Date of Submission: 4/9/2020 |
| Grade: |  |

**B.1 Association Rule Code written by student:**

***(Paste your Association Rules code completed during the 2 hours of practical in the lab here)***

("D:/naman/5th Semester/DM/Labs/Experiment 9")

install.packages("arules")

library("arules")

data("Groceries")

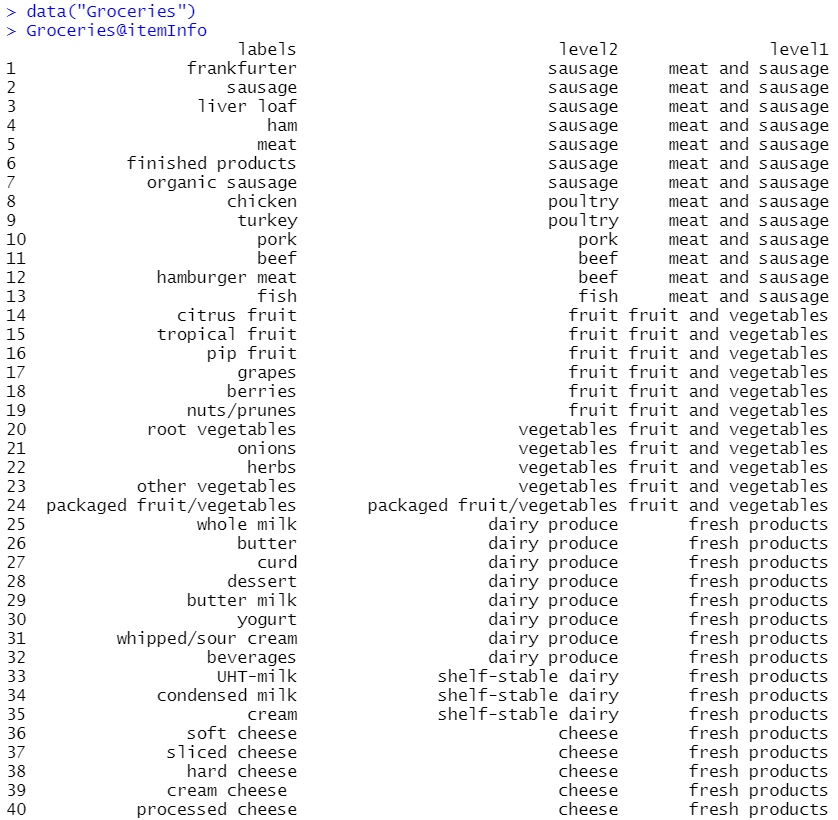
Groceries@itemInfo

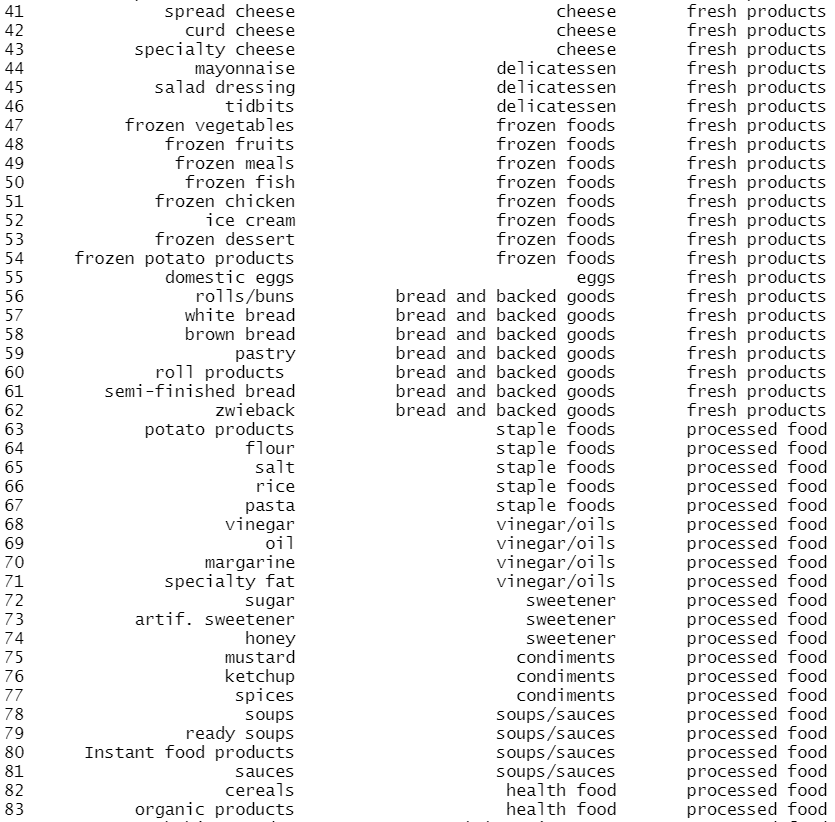
rules<-apriori(Groceries,parameter = list(supp = 0.001, conf = 0.9))

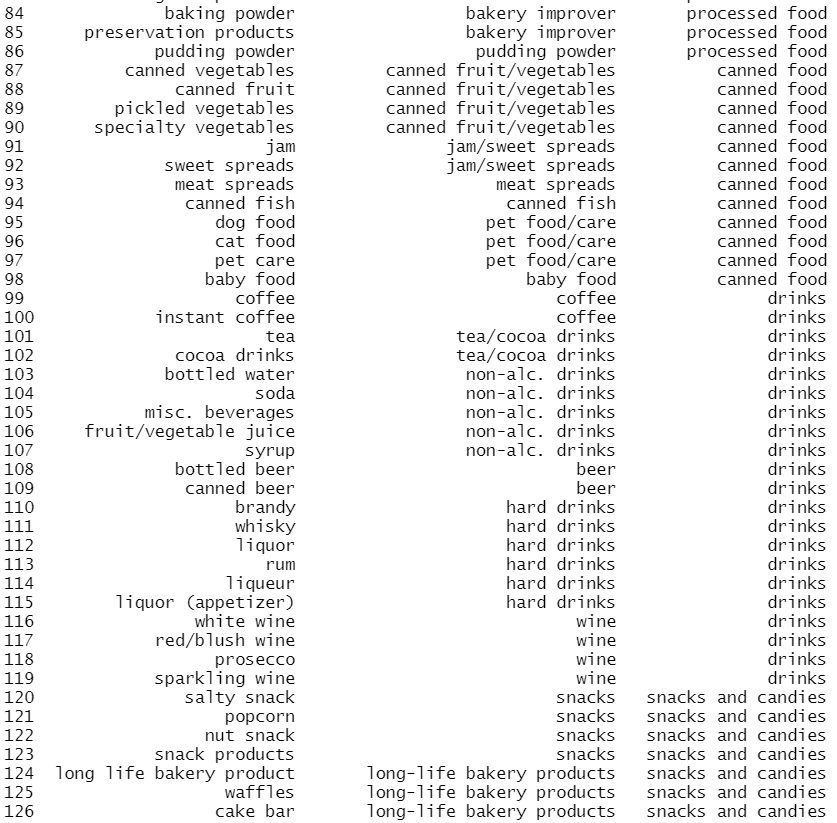
inspect(head(rules))

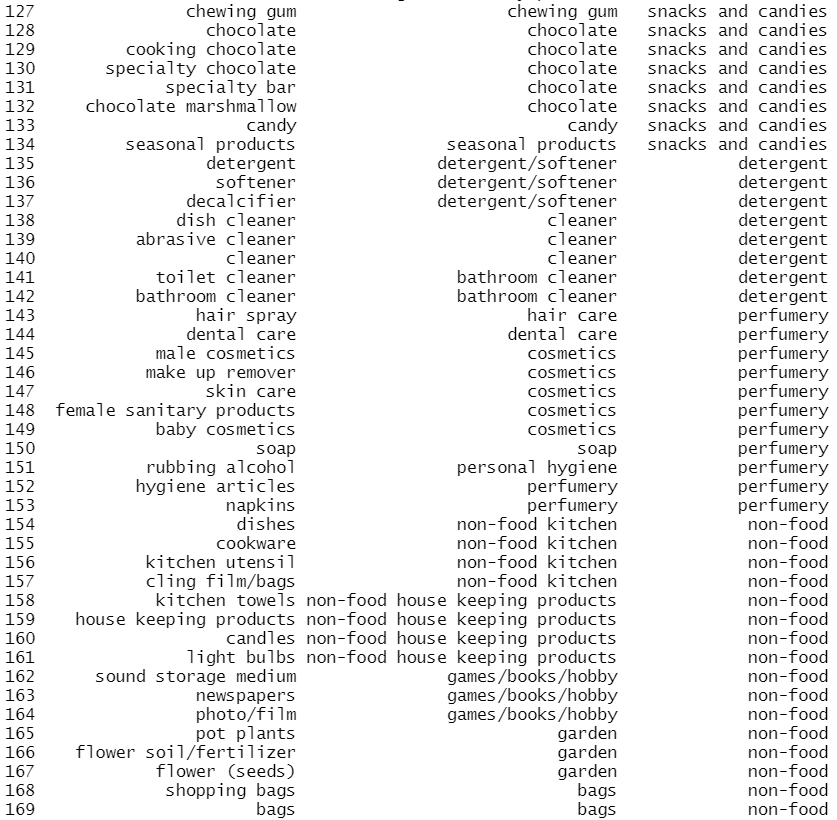
**B.2 Input and Output:**

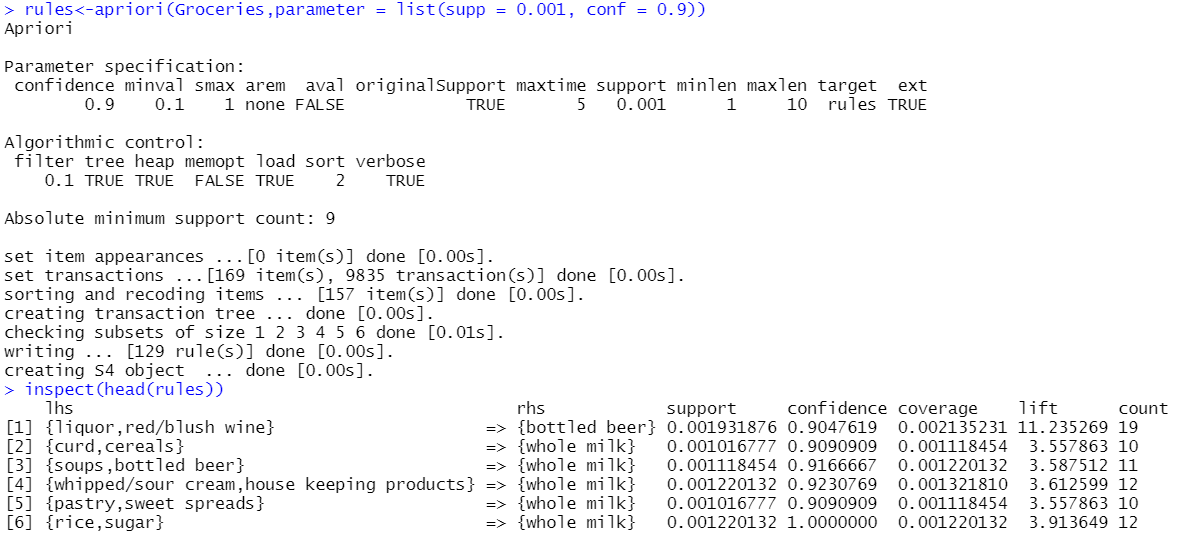
***(Paste your program input and output in following format, If there is error then paste the specific error in the output part. In case of error with due permission of the faculty extension can be given to submit the error free code with output in due course of time. Students will be graded accordingly.)***











**B.3 Observations and learning:**

Apriori Property

All non-empty subset of frequent itemset must be frequent. The key concept of Apriori algorithm is its anti-monotonicity of support measure. Apriori assumes that - “All subsets of a frequent itemset must be frequent (Apriori propertry). If an itemset is infrequent, all its supersets will be infrequent.”

Association rules analysis is a technique to uncover how items are associated to each other. There are three common ways to measure association:

* Measure 1: Support. This says how popular an itemset is, as measured by the proportion of transactions in which an itemset appears.
* Measure 2: Confidence. This says how likely item Y is purchased when item X is purchased, expressed as {X -> Y}. This is measured by the proportion of transactions with item X, in which item Y also appears.
* Measure 3: Lift. This says how likely item Y is purchased when item X is purchased, while controlling for how popular item Y is. A lift value of 1 implies no association between items. A lift value greater than 1 means that item Y is likely to be bought if item X is bought, while a value less than 1 means that item Y is unlikely to be bought if item X is bought.

In the experiment, we used a dataset on grocery transactions from the arules R library. It contains actual transactions at a grocery outlet over 30 days.

**B.4 Conclusion:**

* Understood the concept of Data Mining by implementing some data mining algorithm.
* Understood the Association Rules in Mining.
* Understood how to create Association Rules by Apriori Algorithm.