Super-market Data visualisation with "R"

Abstraction

The data visualisation with association rule going to show that, How to prepare data-set as a shopping list from csv file?. Next step is to overview of most frequent item in list, It gives graphical overview and data Visualisation graphs. Generate rule of data set and create the plot for most Frequent Item using "paracoord" method and also make a plot using "two-key plot" method

Create transaction data set

For data visualisation with association rule, data must be in list view. For that create a data set from csv file and generate basket list for further data visualisation. Basket list will nearly 0.2 million row, but our association visualisation perform on 25,000 transaction row. And 2319 item.

```
 # Read CSV file from local computer and assign to the data variable
> data  <- read.csv( file  ="/home/vikas/Documents/acadamic/system-simulation/data.csv",sep = ";")

# Create Group from basket id for data set
> group_data <- group_by(data,basket_id)

# Make a list of item base on grop set
> basket <- summarise(group_data,count = n(),basket_list = list(article_name))

# Filter data first 25000 basket list analysis
> basket_list <- basket$basket_list[1:25000]

# Convert list in to transaction
> retail_transaction <- as(basket_list,"transactions")
```

Item frequency plot overview

Item frequency plot going to show the most frequent item in basket list with support. Overview of top 8 most frequent item in basket list shows on plot. The most frequent item is "Tweenzers with magnifying glass" has support 0.03 it means there are 3% of basket has this item.

```
> itemFrequencyPlot(retail_transaction,topN = 8)
```

support = σ{item in set} /

σ{total Transaction}

Opration on Data

- 25,000 (Row)

- 2319 item (column)
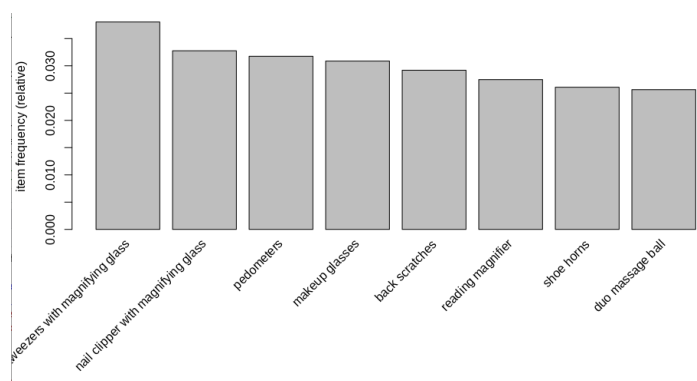
Plot 1: Generate rule & plot for most frequent item using paracoord method

Generate rule with support 0.0002 and confidance 0.66 and minimum item in basket is 2 and maximum item in basket is 3 that generate 54 rule, from that get subset of most frequent item "tweezers with magnifying glass" and filter first 4 rules.
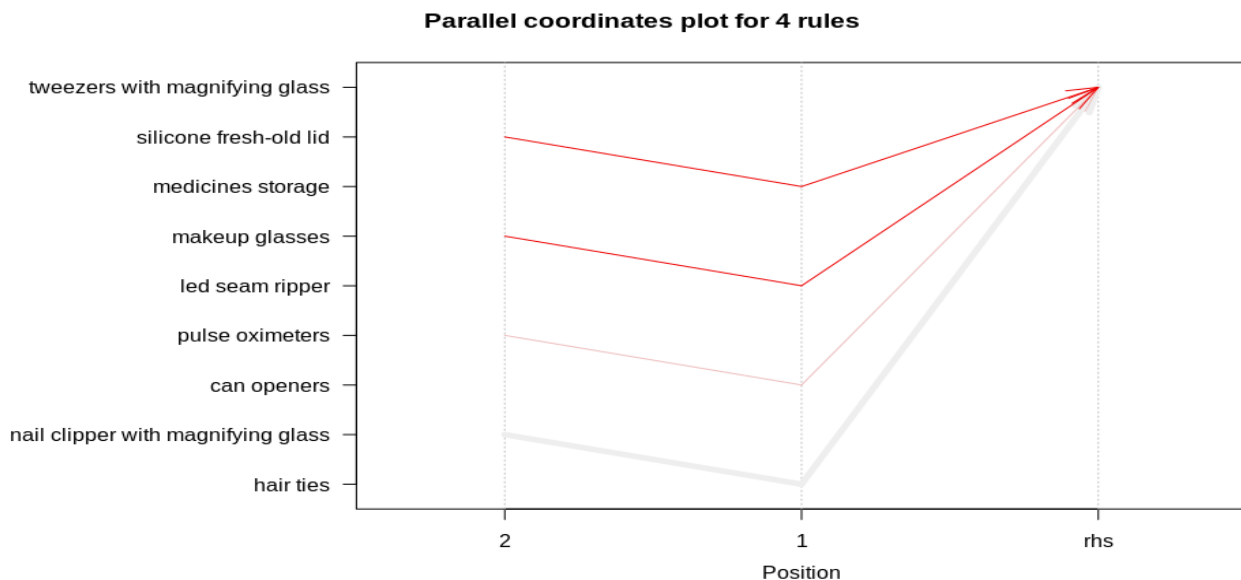
```
> basket_rule <- apriori(retail_transaction, parameter = list(support = 0.0002, conf
                  = 0.66, minlen = 2, maxlen = 3))
```

```
#filter rule apply on tweezer with magnifying glass
> tweezer_rule <- subset(basket_rule, subset = rhs %in% "tweezers with magnifying
                  glass")
> tweezer_rule_head <- head(sort(tweezer_rule ,by = "lift"),4)
> plot(tweezer_rule_head,method="paracoord")
```

Parallel coordinates plot for 4 rules

Plot 2: Generate rule for basket and create plot with two-key plot method

For entire basket overview use the tow-key plot method, It gives more flexibility to show item list on pertiquler region point with support and confidance value

```
> plot(basket_rule,method = "two-key plot",interactive = TRUE)
```

selected region show the inspect of item with support and confidance
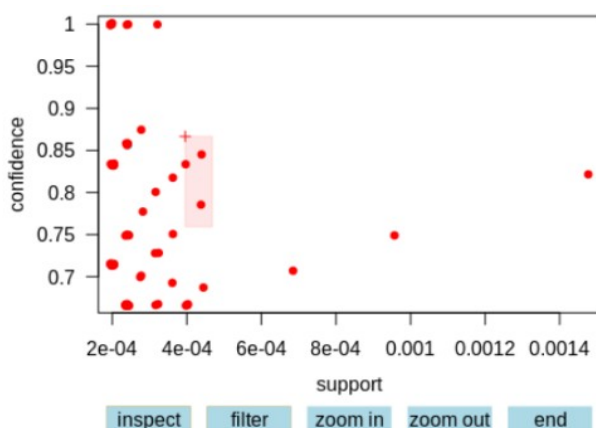
Appendix

code :

```r
install.packages("arules")
install.packages("tidyverse")
install.packages("arulesViz")

library(arulesViz)
library(arules)
library(tidyverse)

# Read CSV file from local computer and assign to the data variablr
data  <- read.csv( file = "/home/vikas/Documents/acadamic/system-
simulation/Exercise2/data.csv",sep = ";")

# Create Group from basket id for data set
group_data <- group_by(data,basket_id)

# Make a list of item base on grop set
basket <- summarise(group_data,count = n(),basket_list = list(article_name))
# Filter data first 25000 basket list analysis
basket_list <- basket$basket_list[1:25000]
basket_list[1:2]

# Convert list in to transaction
retail_transaction <- as(basket_list,"transactions")
# Test plot function and set of rulse
itemFrequencyPlot(retail_transaction,topN = 8)
apriori(retail_transaction)

# Create Rule for basket set
basket_rule <- apriori(retail_transaction, parameter = list(support = 0.0002, conf =
0.66, minlen = 2, maxlen = 3))

# filter and select first 6 rule from top
basket_rule_six <- head(sort(basket_rule ,by = "lift"),6)
inspect(head(sort(basket_rule ,by = "lift"),6))
# Create plot with method two-key plot
plot(basket_rule,method = "two-key plot",interactive = TRUE)
```

```r
#Filter rule apply on tweezer with magnifing glass
tweezer_rule <- subset(basket_rule, subset = rhs %in% "tweezers with magnifying glass")
tweezer_rule_head <- head(sort(tweezer_rule ,by = "lift"),4)
inspect(tweezer_rule_head)
# Create plot with method paracoord
plot(tweezer_rule_head,method="paracoord")


# Create Extra plote for vizulization
plot(basket_rule,methods = "grouped", control = list(10))
plot(basket_rule_six, method="grouped")
```