МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ

ФЕДЕРАЛЬНОЕ государственное БЮДЖЕТНОЕ

образовательное учреждение

высшего образования

«НОВОСИБИРСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Кафедра Вычислительной техники



**ОТЧЁТ**

**по лабораторной работе № 3**

**«Ассоциативный анализ данных»**

**по дисциплине: «***Компьютерные технологии анализа и обработки данных***»**

Выполнил:Проверил:

Студенты гр. «АММ2-21», «АВТФ»  
 *«Арнольд Э.В»* *«*к.т.н., доцент*»*

*«Антонов С.С» «Альсова О.К.»*

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

(подпись) (подпись)

Новосибирск

2021

1. Задание к работе:

1. Проведите ассоциативный анализ данных согласно варианту задания (см. данные в csv-файлах). Приведены данные о продажах бытовой и электронной техники. Всего 1230 записей (покупателей). Данные закодированы 1 – покупка совершена, 0 – нет. Также приведена информация о кредите (был ли взят кредит на покупку).

2. Выявите ассоциативные взаимосвязи в данных.

3. Постройте сетевой график правил ассоциации.

4. Сделайте выводы о закономерностях в данных на основе проведенного анализа.

Ход работы

Загрузим исходные данные и взглянем на загруженную таблицу:

*require("arules")*

*require("arulesViz")*

*Sys.setlocale("LC\_ALL", "Russian")*

*Sys.setlocale("LC\_CTYPE", "en\_RU.UTF-8")*

*data <- read.delim(file = "C:\\Users\\Dmitriy\\YandexDisk\\нгту\\КТАД-2021\\lab3\_7.csv", sep = ",", header = TRUE, row.names = 1)*

*dataN <- data[, -11]*

*isCredit <- data$deb*

*isCredit[isCredit == 1] <- "With deb"*

*isCredit[isCredit == 0] <- "Without deb"*

*View(dataN)*

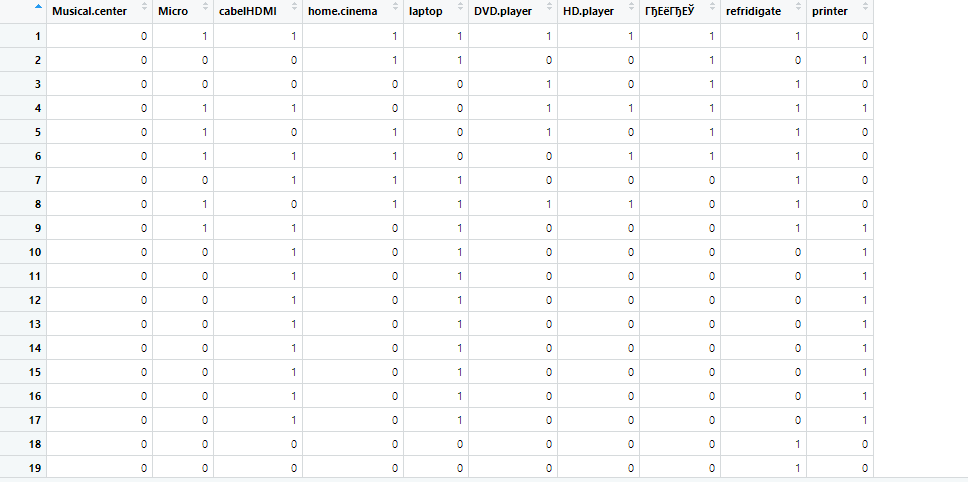


Рис. 1 – Загруженная таблица с данными.

Cделаем преобразование исходных данных в файл транзакций, считаем транзакции из файла в переменную и посмотрим сводную информацию о наборе транзакций:

*itemsList <- sapply(1:nrow(dataN), function(i)*

*paste(c(isCredit[i], colnames(dataN[i, dataN[i, ] == 1])),*

*collapse = ",", sep = "\n"))*

*write(itemsList, file = "C://basket.csv")*

*trans <- read.transactions("C://basket.csv", format = "basket", sep = ",")*

*summary(trans)*

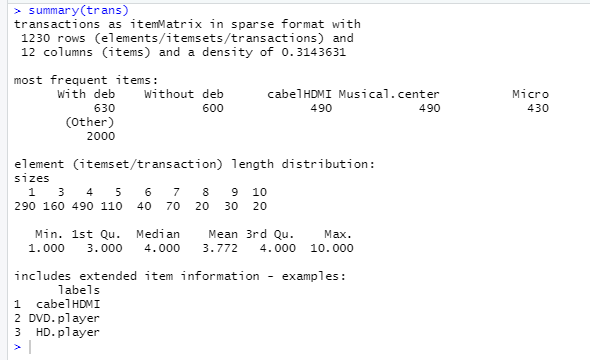
**

Рис. 2 – Сводная информация о наборе транзакций.

Построим частотную диаграмму транзакций:

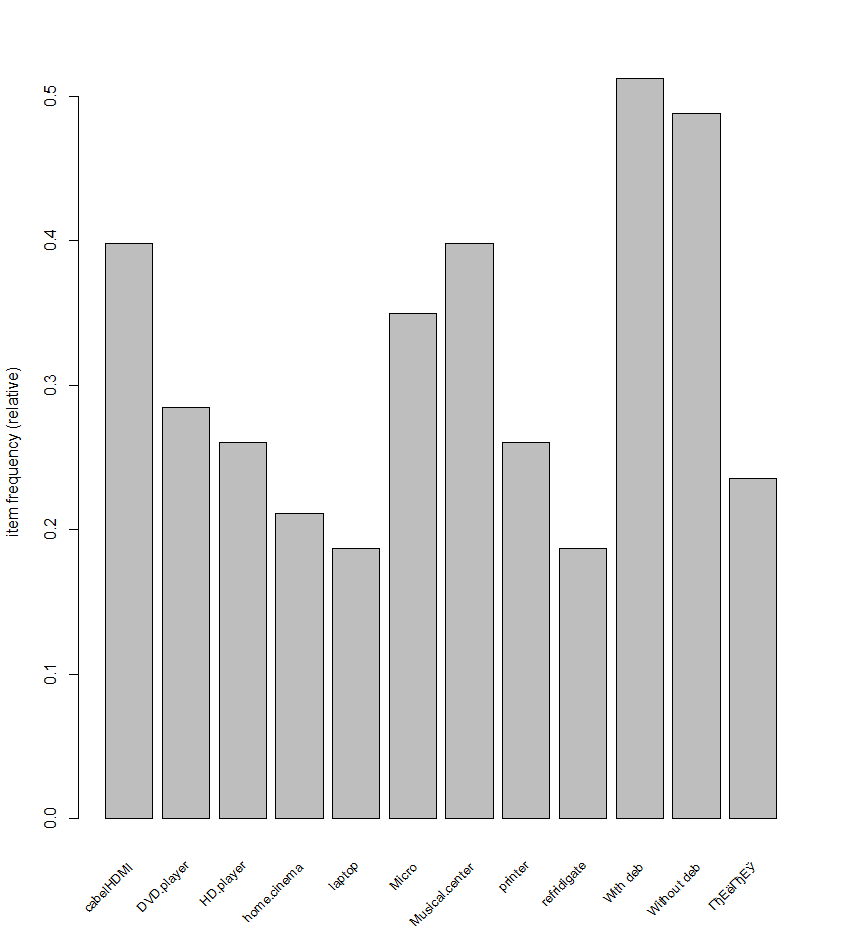
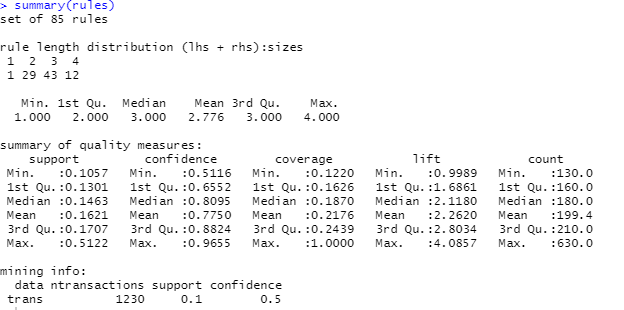


Рис. 3 – Частотная диаграмма транзакций

Составим ассоциативное правило следующей командой с минимальной поддержкой 0.1 и минимальной достоверностью 0.5:

*rules <- apriori(trans, parameter = list(support = 0.1, confidence = 0.5))*

Просмотр сводной информации о полученных ассоциативных правилах осуществляется с помощью команды *summary(rules)*.



Построим сеть ассоциативных правил при помощи команды:

*plot(head(sort(rulesWithCredit, by = "support"), 84), method = "paracoord")*

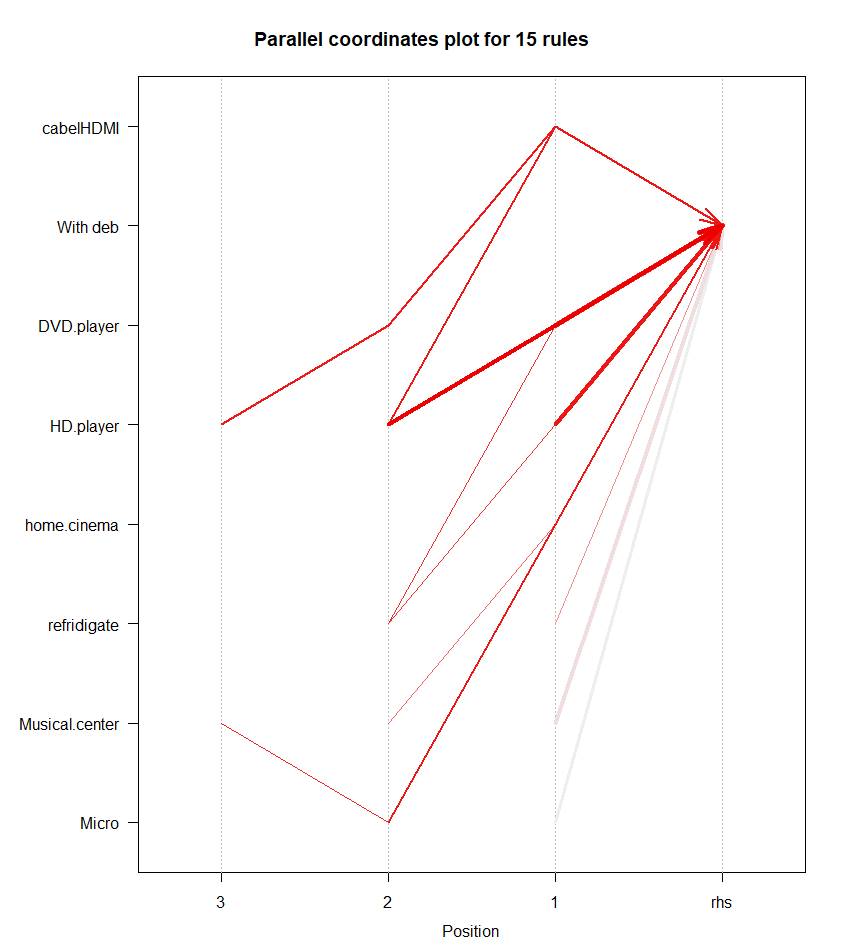
**

Рисунок 4 – Сеть ассоциативных правил для транзакций “С кредитом”

Создадим ассоциативное правило для транзакций с кредитом и без кредита.

*rulesWithCredit <- subset(rules, subset = rhs %in% "With deb")*

*rulesWithoutCredit <- subset(rules, subset = rhs %in% "Without deb")*

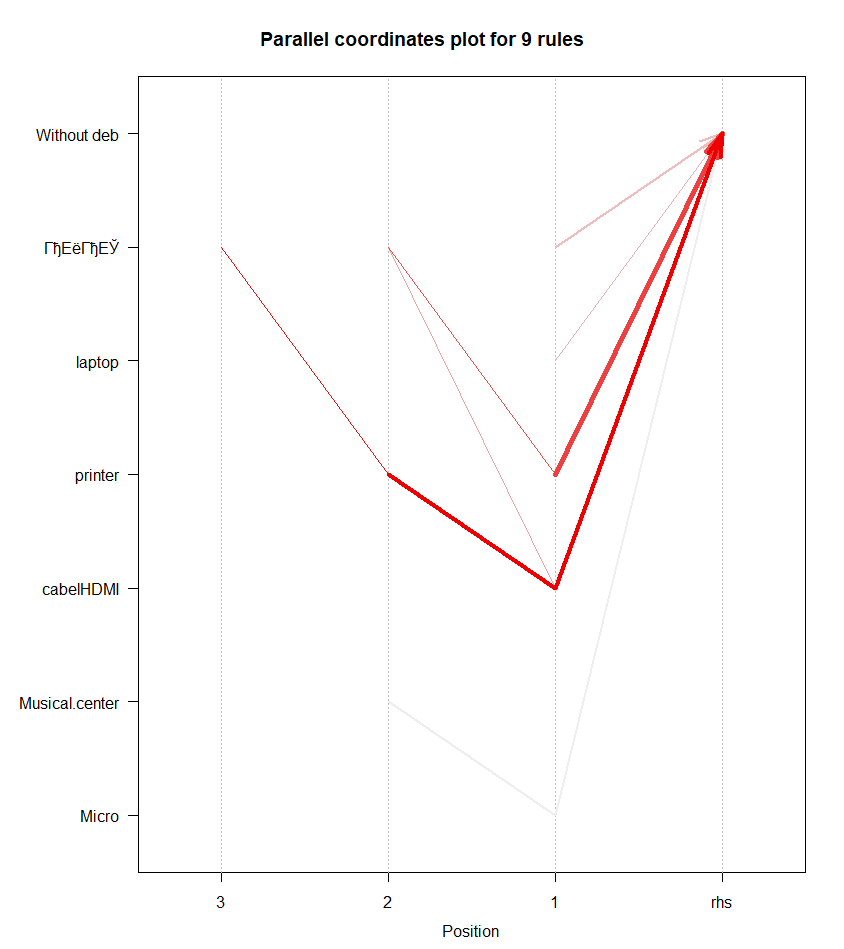


Рисунок 5 – Сеть ассоциативных правил для транзакций “Без кредитом”

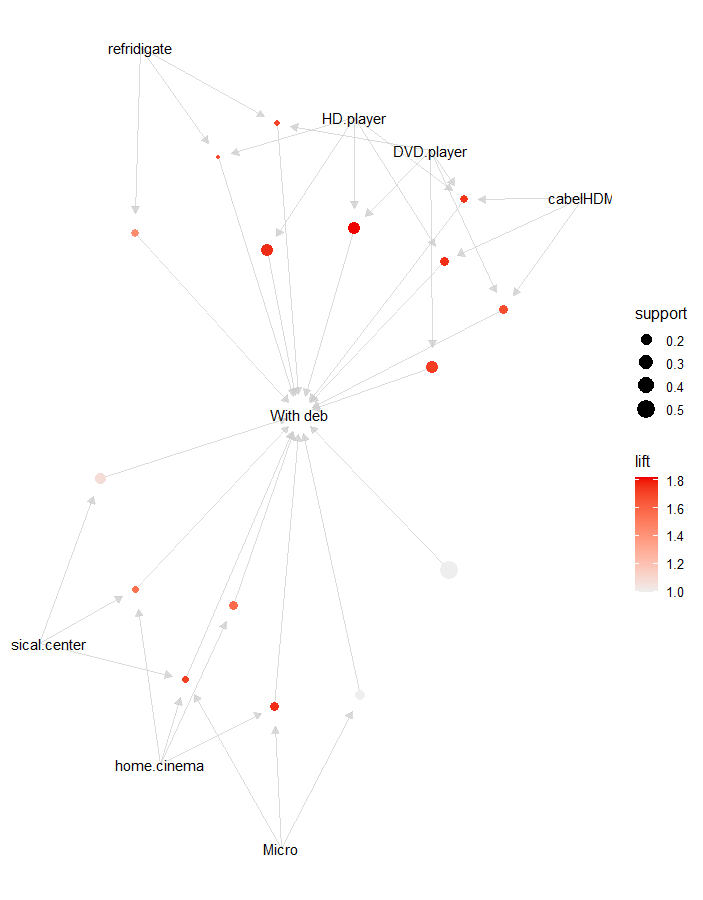


Рисунок 6 – Граф ассоциативных правил “С кредитом”

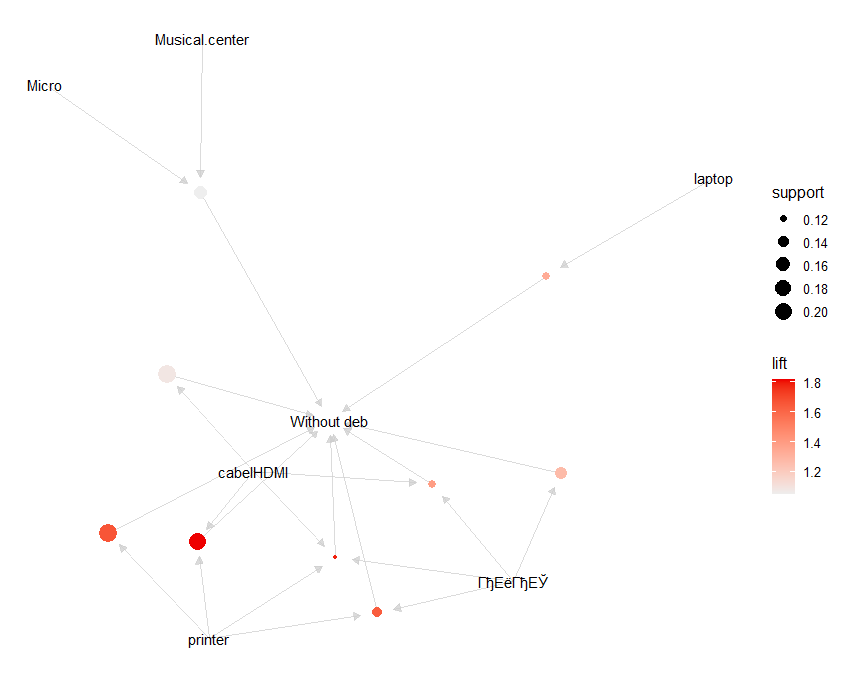


Рисунок 7 – Граф ассоциативных правил “Без кредита”

Как видно из рисунка 7 и рисунка 6 можно сделать вывод об о том что чаще в кредит покупают телевизор , dvd-player и кабель hdmi, а музыкальный центр, микрофон и холдильник предпочитают покупать без кредита.

**Вывод:** в ходе лабораторной работы были получены навыки ассоциативного анализа в программной среде R. Проведен ассоциативный анализ о продаже домашней электроники, выявлены ассоциативные связи между данными, построены сетевой график и граф ассоциативных правил, сделаны выводы о проделанной работе.