МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКО ФЕДЕРАЦИИ

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ

УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ

**«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ НИЖЕГОРОДСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ИМ. Н.И. ЛОБАЧЕВСКОГО»**

Дзержинский филиал

**Кафедра** Прикладная информатика

**ОТЧЕТ**

**по дисциплине «Интеллектуальный анализ данных»**

**к лабораторной работе № 3**

**Тема «R»**

Выполнил:

Студент гр. № 2721Б1ПИ

Погодин Дмитрий Евгеньевич

Проверил:

Ст.преподаватель, к.т.н.

\_\_\_\_\_\_\_\_\_/Нажимов А.В./ \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Дзержинск 2024

Оглавление

[**Лабораторная работа №3** 3](#_Toc135380445)

[**Вывод** 10](#_Toc135380446)

# **Лабораторная работа №3**

**Ассоциативные правила**

**Цель**

Построить график, импортировать библиотеки для построения, применить ассоциативные правила, воспользовавшись пакетом arules и сделать выводы.

**Ход выполнения работы**

1. Добавим библиотеки cluster, ggplot2, factoextra, arules. Для этого в пакетах найдём их по поиску и загрузим.
2. Данные, которые мы используем, – данные транзакций покупки продуктов, в которых имеется 169 уникальных элементов товаров: groceries.csv.

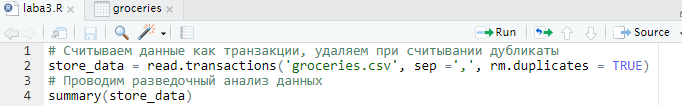


Рис.1. Код для описания переменных набора

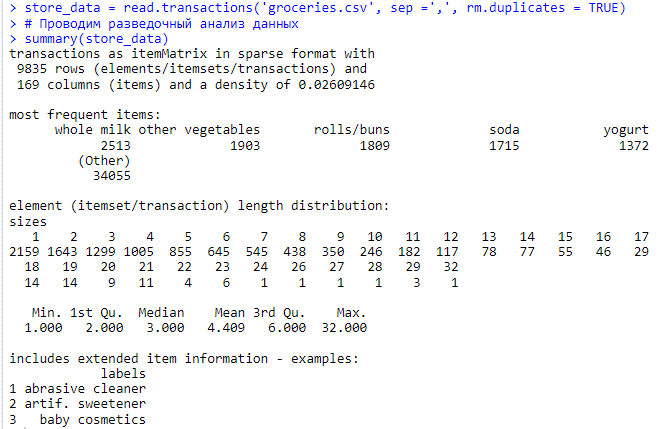


Рис. 2. Описание переменных набора

По полученным данным можно заметить, что наиболее встречающийся продукт это whole milk, который встречается 2513 раз. Также ещё 5 популярных продуктов это vegetables (1903 times), rolls/buns (1809 times), soda (1715 times), yogurt (1372 times).

Самое большое число покупок имеет 1-ая транзакция (2159 раз), именно столько человек купили 1 продукт.

Самое наименьшее количество имеют несколько последних транзакций: 24, 26, 27, 28, 32.

Согласно полученным данным в среднем покупатели приобретают 4 продукта.

Основная тематика продуктов (спецификация), которые покупают покупатели:

Abrasive cleaner (Абразивный очиститель),

Artif sweetener (Искусственный подсластитель),

Baby cosmetics (Детская косметика)

1. Построим на графике наиболее часто встречающиеся в транзакциях продукты:

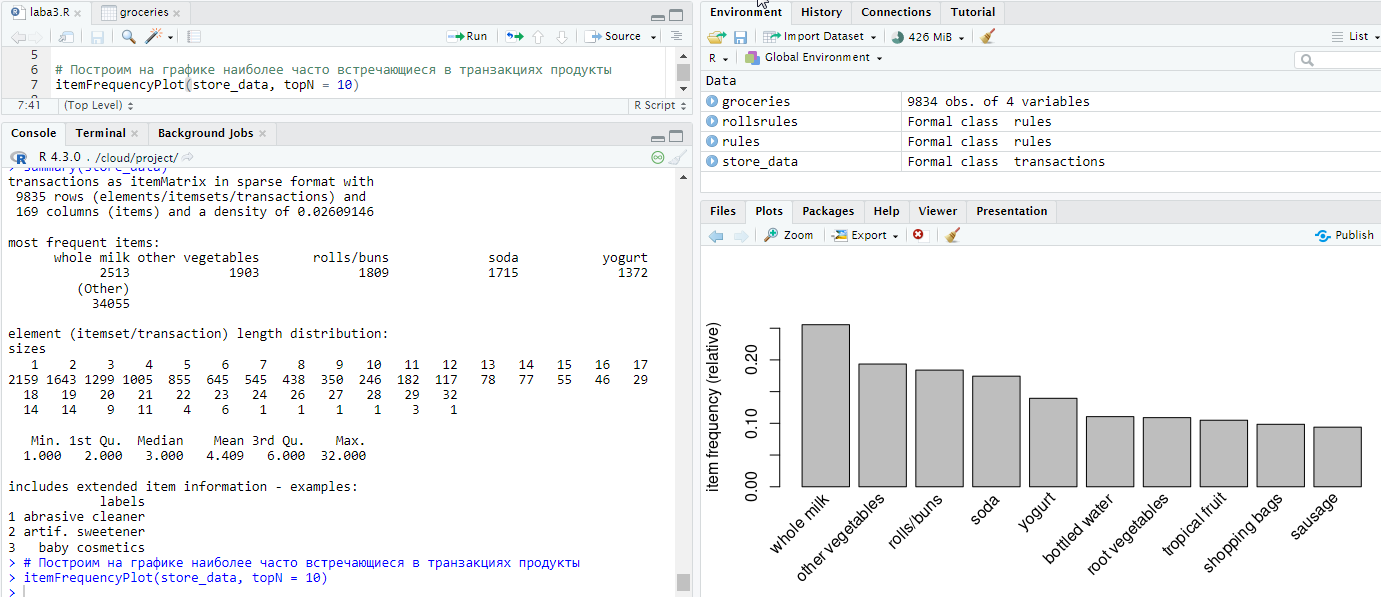


Рис. 3. График + код + консоль для наиболее часто встречающихся продуктов в транзакциях

Благодаря графику мы можем наглядно видеть, самые популярные продукты в транзакциях, а именно, whole milk имеет частоту 0.25, vegetables 0,20, самое минимальное из 10 самых популярных имеет sausage 0,7.

1. Применим метод apriori к транзакциям набора store\_data:

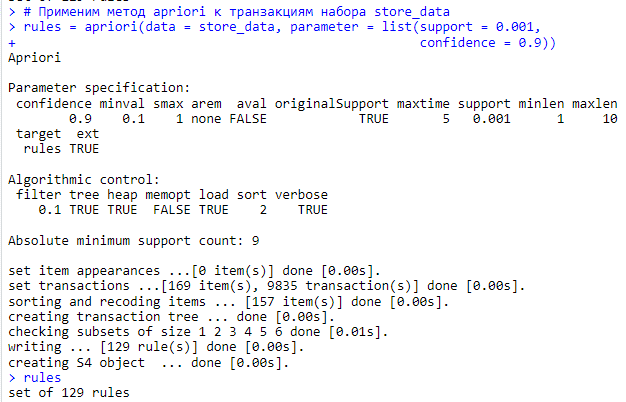


Рис.4. Применение метода apriori к транзакциям набора store\_data

1. Выведем первые 3 правила, упорядоченные по убыванию параметра lift

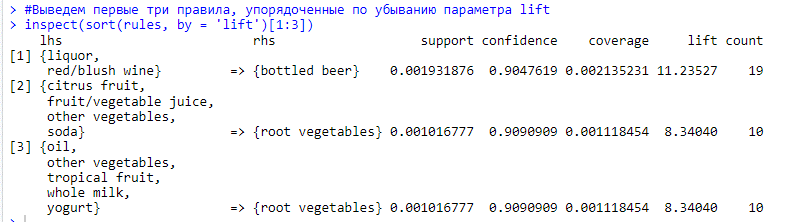


Рис.5. Первые 3 правила

Из полученных 3 правил можно заметить, что:

- если покупают ликёр или красное вино, то скорее всего ещё купят пиво в бутылках, но не всегда, на это нам указывают коэффициенты support, confidence, coverage, lift=11, что показывает степень зависимости левой части правила и правой, count=19, количество продуктов.

- если покупают цитрусовые фрукты, фруктовый или овощной сок, то скорее всего купят корнеплоды, но не всегда.

- если покупают масло, другие овощи, цельное молоко, йогурт, то скорее всего купят корнеплоды, но не всегда.

Из полученных 3 правил заметно, что у 2 и 3 правила одинаковые коэффициенты support, confidence, lift, count.

1. Анализ правил:

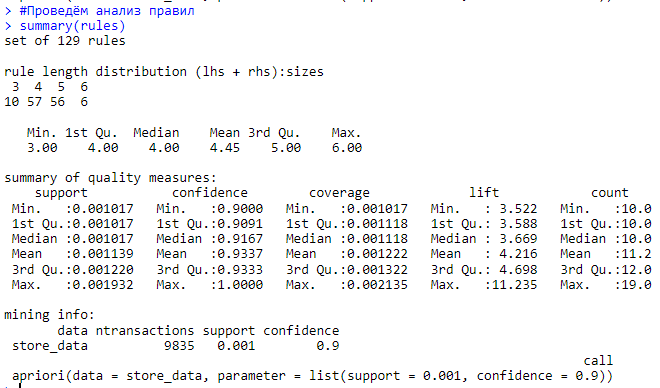


Рис.6. Анализ правил

- Всего 129 правил

- В среднем 4 правила, максимальное 6, минимальное 3

- Всего транзакций 9835

1. Выберем правила, в которые входит продукт rolls/buns

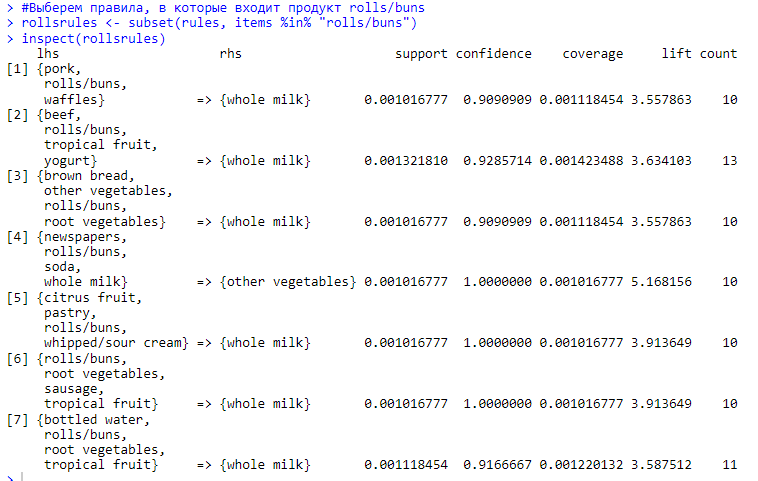


Рис.7. Правила, в которые входит продукт rolls/buns

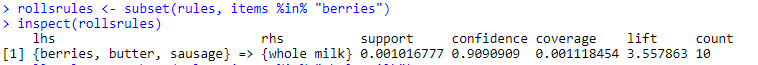


Рис.8. Правила, в которые входит продукт berries

Исходя из этого можно сделать вывод, что berries (ягоды) очень редко покупают с другими продуктами

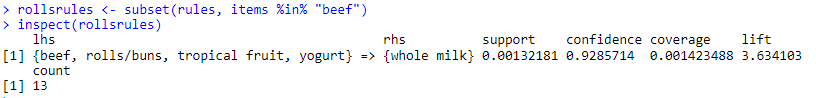


Рис.9. Правила, в которые входит продукт beef (говядина)

Говядину также редко покупают с другими продуктами.

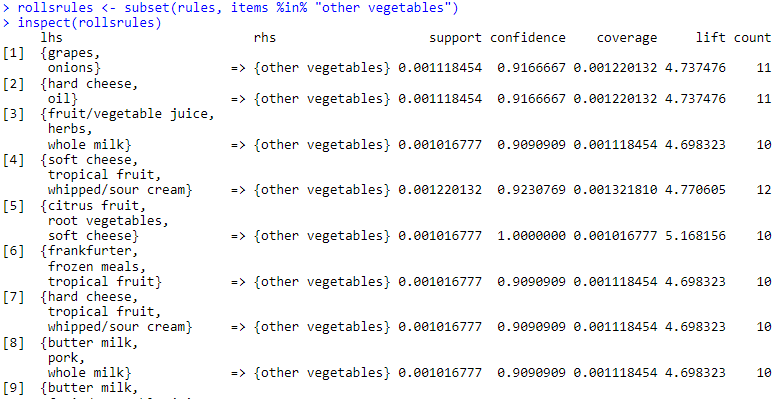


Рис.10. Правила, в которые входит продукт other vegetables

Всего в other vegetables входит 64 правило, из этого следует, что овощи покупают часто с другими продуктами.

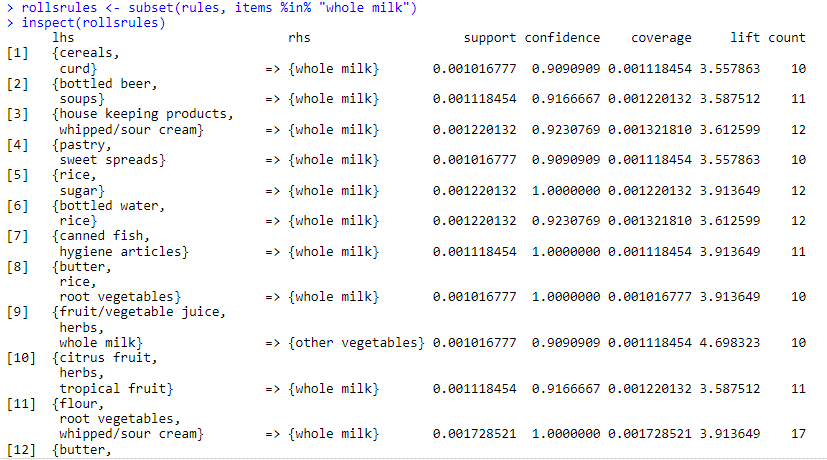


Рис.11. Правила, в которые входит продукт whole milk

Всего в whole milk 102 правила, следовательно с ними берут очень много продуктов, потому что молоко часто используют для приготовления блюд.

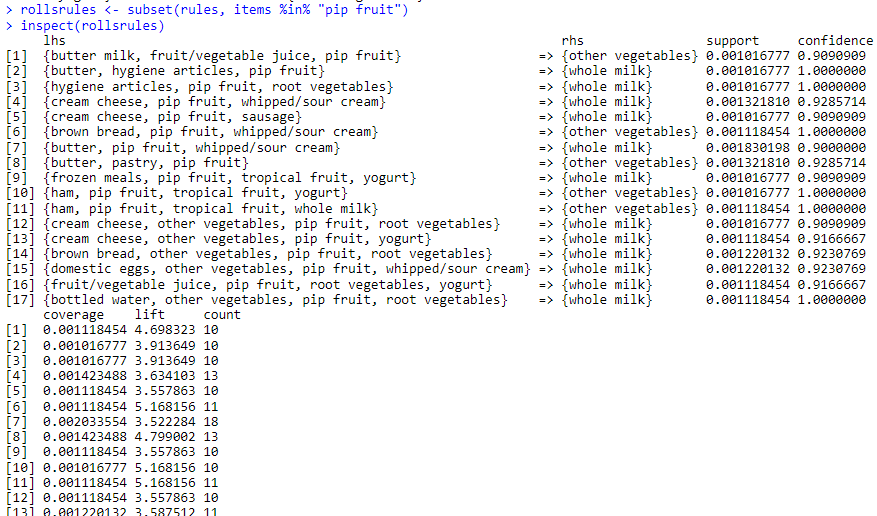


Рис.12. Правила, в которые входит продукт pip fruit

Pip fruit довольно часто берут с другими продуктами, особенно с цельным молоком.

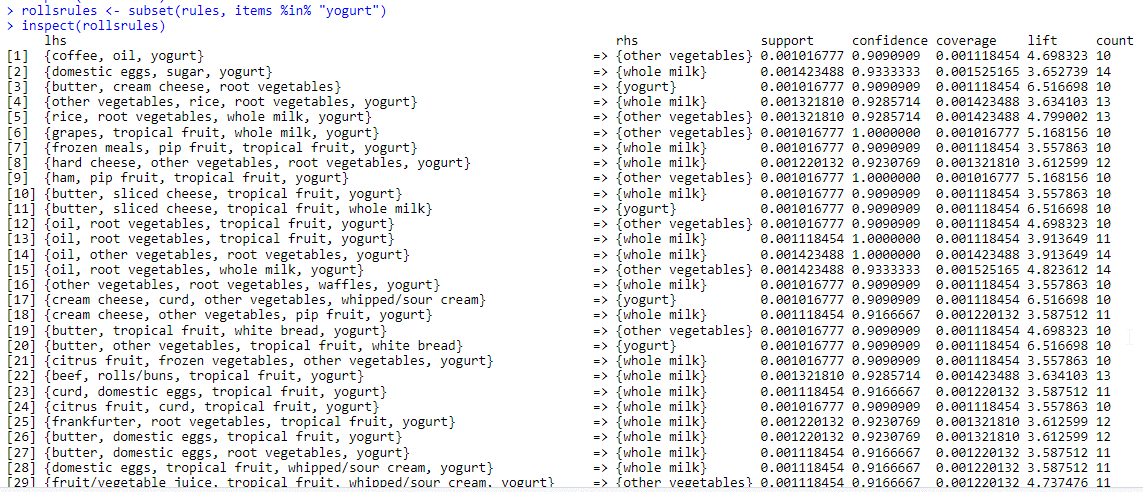


Рис.13. Правила, в которые входит продукт yogurt

Всего 38 правил, йогурт берут чисто с цельным молоком и овощами.

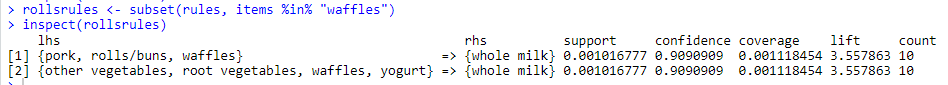


Рис.14. Правила, в которые входит продукт waffles

Вафли берут редко, и в основном с цельным молоком, но не всегда.

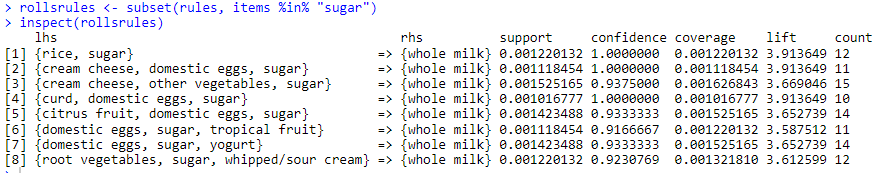


Рис.15. Правила, в которые входит продукт sugar

Сахар покупают нечасто, и зачастую с сахаром берут цельное молоко.

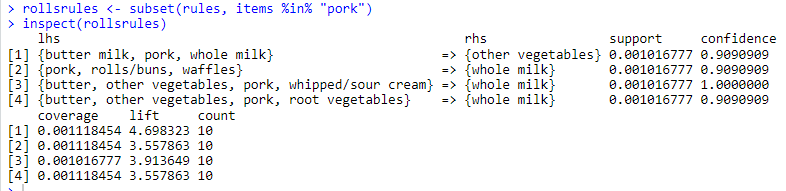


Рис.16. Правила, в которые входит продукт pork

Судя по полученным данным, если берут свинину, то с ней покупают овощи и масло, а также могут взять цельное молоко, но не всегда.

# **Вывод**

В ходе выполнения всех лабораторных работ мною был проведён анализ полученных данных с помощью метода ассоциативных правил.