

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ
Федеральное государственное автономное образовательное учреждение
высшего образования
«Южно-Уральский государственный университет
(национальный исследовательский университет)»
Высшая школа электроники и компьютерных наук
Кафедра системного программирования

ОТЧЕТ
по лабораторной работе №1
Технологии Аналитической обработки информации

Выполнил,
студент группы КЭ-403
Исхаков М.Р.
Проверил:
Гоглачев А.И.

Челябинск, 2025 г.

ФОРМУЛИРОВКА ЗАДАНИЯ

В рамках выполнения лабораторной работы необходимо:

1. Разработайте программу, которая выполняет поиск частых наборов объектов в заданном наборе данных с помощью алгоритма Apriori (или одной из его модификаций). Список результирующих наборов должен содержать как наборы, так и значение поддержки для каждого набора. Параметрами программы являются набор, порог поддержки и способ упорядочивания результирующего списка наборов (по убыванию значения поддержки или лексикографическое).

2. Проведите эксперименты на наборе данных **baskets.csv** (сведения о покупках в супермаркете). В экспериментах варьируйте пороговое значение поддержки (например: 1%, 3%, 5%, 10%, 15%).

3. Выполните визуализацию результатов экспериментов в виде следующих диаграмм:

- сравнение быстродействия на фиксированном наборе данных при изменении порога поддержки;
- количество частых наборов объектов различной длины на фиксированном наборе данных при изменяемом пороге поддержки.

4. Подготовьте отчет о выполнении задания и загрузите отчет в формате PDF в систему. Отчет должен представлять собой связный и структурированный документ со следующими разделами:

- формулировка задания;
- гиперссылка на каталог репозитория с исходными текстами, наборами данных и др. сопутствующими материалами;
- рисунки с результатами визуализации;
- пояснения, раскрывающие смысл полученных результатов

Исходные коды для задания представлены в репозитории:
<https://github.com/SMarkls/analysis>

РИСУНКИ С РЕЗУЛЬТАТАМИ ВИЗУАЛИЗАЦИИ

В ходе выполнения поиска частых наборов объектов в заданном наборе данных с помощью алгоритма априори были получены результаты и отображены на рисунках 1 и 2.

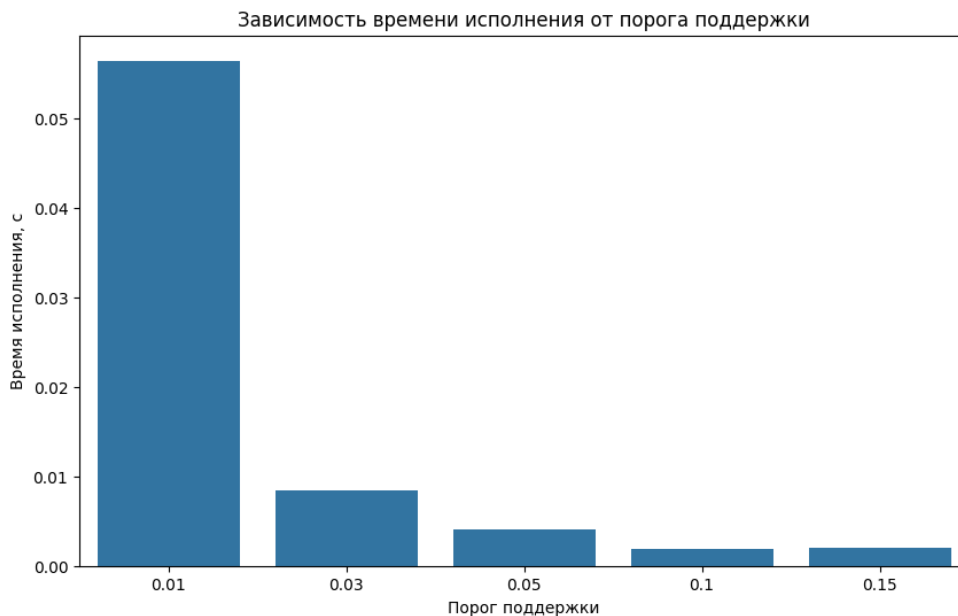


Рисунок 1 - Зависимость времени исполнения от порога поддержки

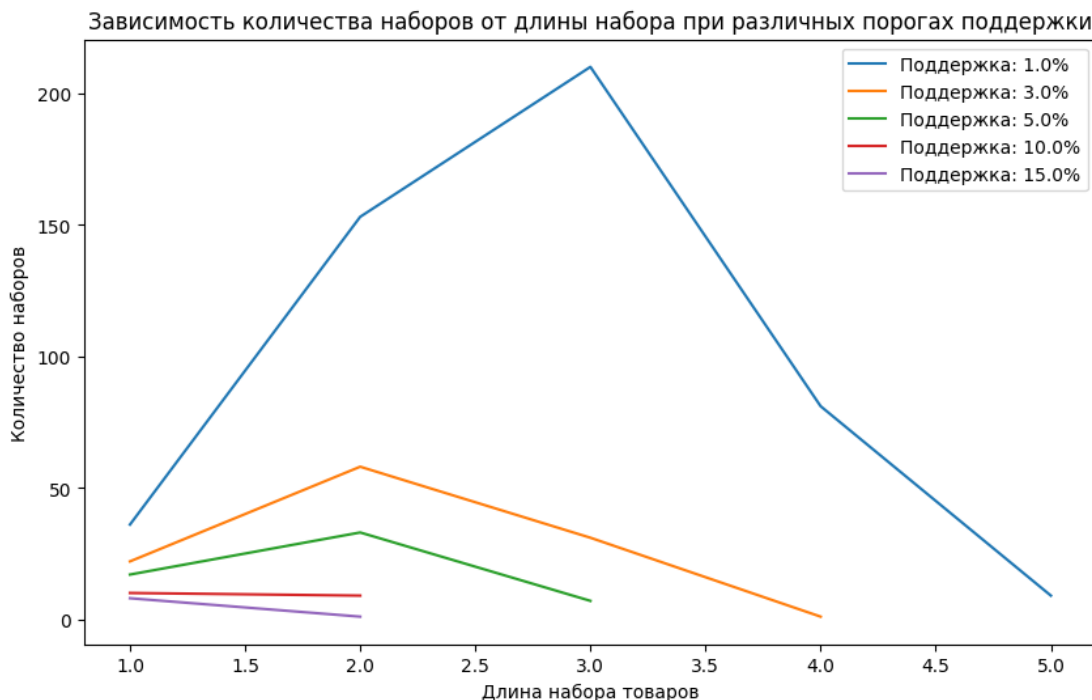


Рисунок 2 - Зависимость количества наборов от длина набора при различных порогах поддержки

ВЫВОДЫ ИЗ ВИЗУАЛИЗАЦИИ

Время выполнения значительно увеличивается при низких порогах поддержки, т.к. количество наборов существенно выше.

При низком пороге поддержки обнаруживается значительно больше частых наборов всех длин. С увеличением порога поддержки количество обнаруживаемых наборов резко снижается.