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**

о выполнении практического задания № 1 по дисциплине

«Технологии аналитической обработки информации»

Выполнил:

студент группы КЭ-402 Риконина В.В.

Проверил: Юртин А. А.

Челябинск – 2025

# ЗАДАНИЕ

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

# ВЫПОЛНЕНИЕ

Исходный код программы представлен в репозитории: https://github.com/veronirik/Analytical-information-processing-technologies

Результаты выполнения представлены в виде графиков.

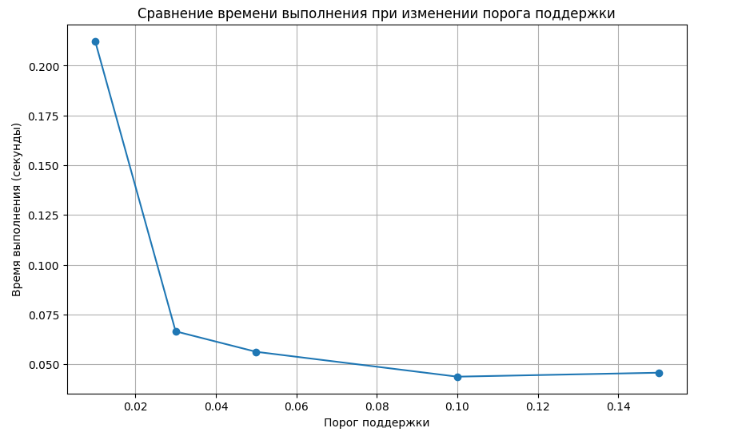


Рисунок 1 – График зависимости времени выполнения от значения min\_support

По данному графику можно заметить, что чем больше значение min\_support, тем меньше времени затрачивается на работу программы. Это происходит из-за того, что при большем min\_support находится меньше частых наборов, вследствие чего алгоритм выполняет меньше операций.

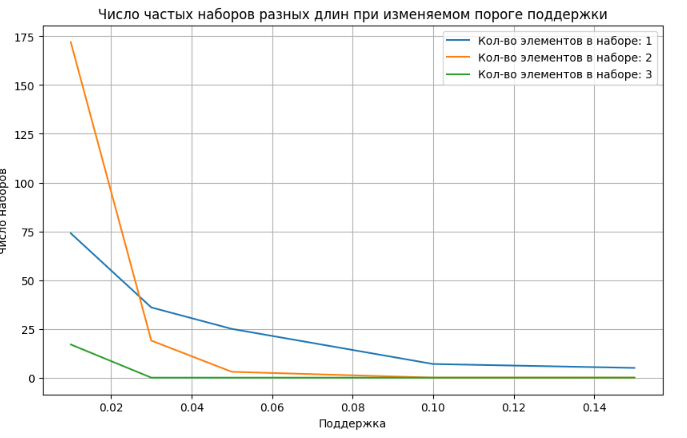


Рисунок 2 – График зависимости количества частотных наборов разной длины от значения min\_support

По данному графику можно заметить, что чем больше min\_support, тем меньше становится различных частотных наборов разной длины. Это связано с тем, что при увеличении min\_support алгоритм apriori начинает искать более связанные наборы продуктов, отсеивая редко встречающиеся.

Это происходит из-за того, что аргумент min\_support является минимальной долей транзакций, в которых встречается набор товаров. Это значит, что увеличение данного аргумента напрямую влияет на количество найденных частых товаров, что видно на полученных графиках.