**Технологии аналитической обработки информации**

**Задание 1.**

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 изменяемом пороге 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.

**Результат.**

На рисунке 1 представлена реализация метода для нахождения частых наборов с сортировкой по значению поддержки.

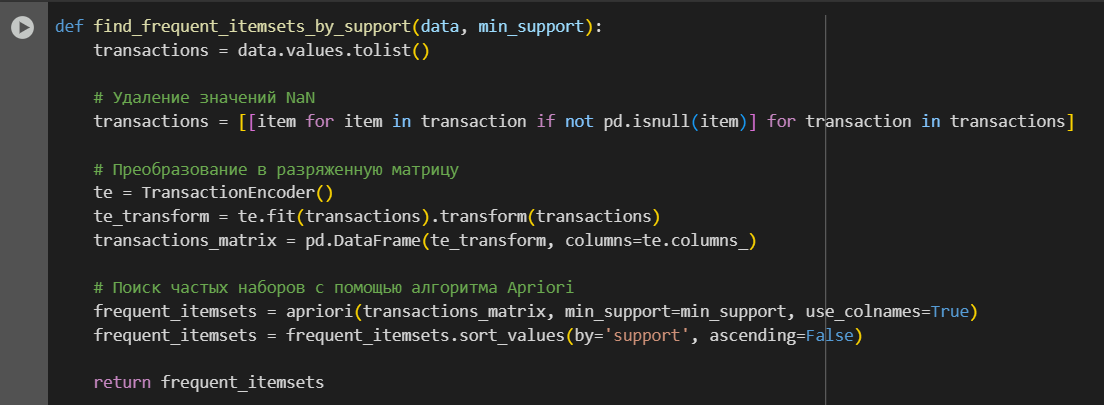
****

Рисунок 1 – Метод find\_frequent\_itemsets\_by\_support

На рисунке 2 представлен вызов метода для нахождения частых наборов с сортировкой по значению поддержки.

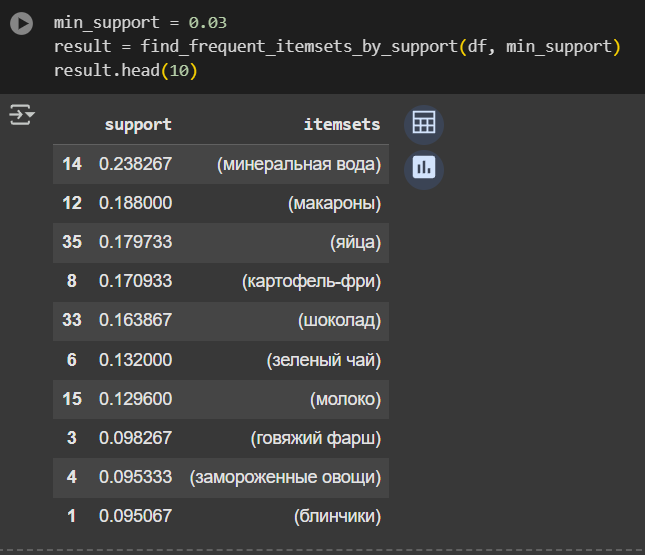
****

Рисунок 2 – Вызов метода

На рисунке 3 представлена реализация метода для нахождения частых наборов с лексикографической сортировкой.

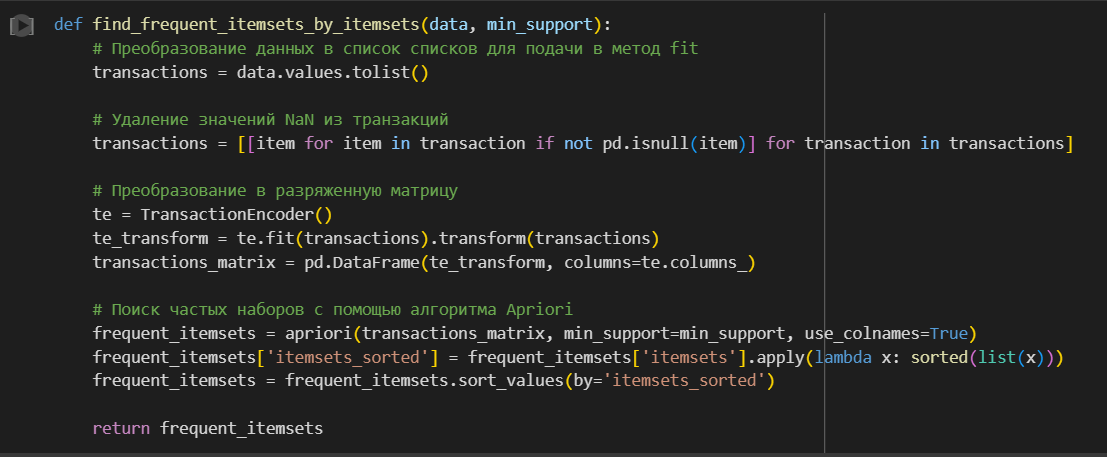


Рисунок 3 – Метод find\_frequent\_itemsets\_by\_itemset

На рисунке 4 представлен вызов метода для нахождения частых наборов с лексикографической сортировкой.

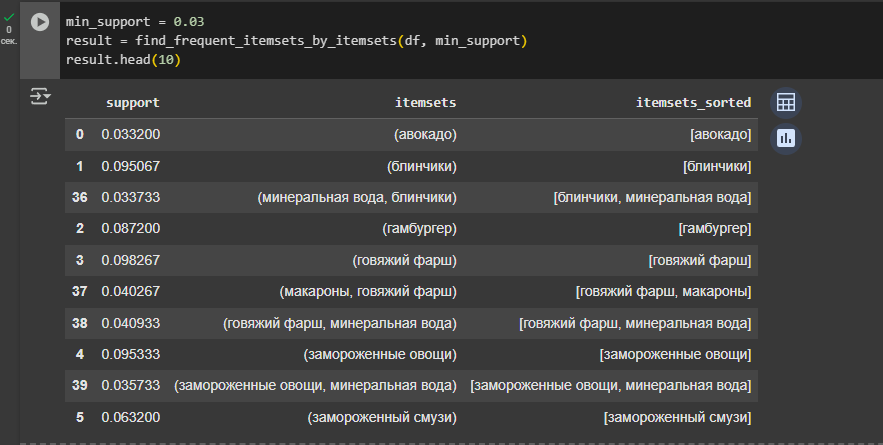


Рисунок 4 – Вызов метода find\_frequent\_itemsets\_by\_itemset

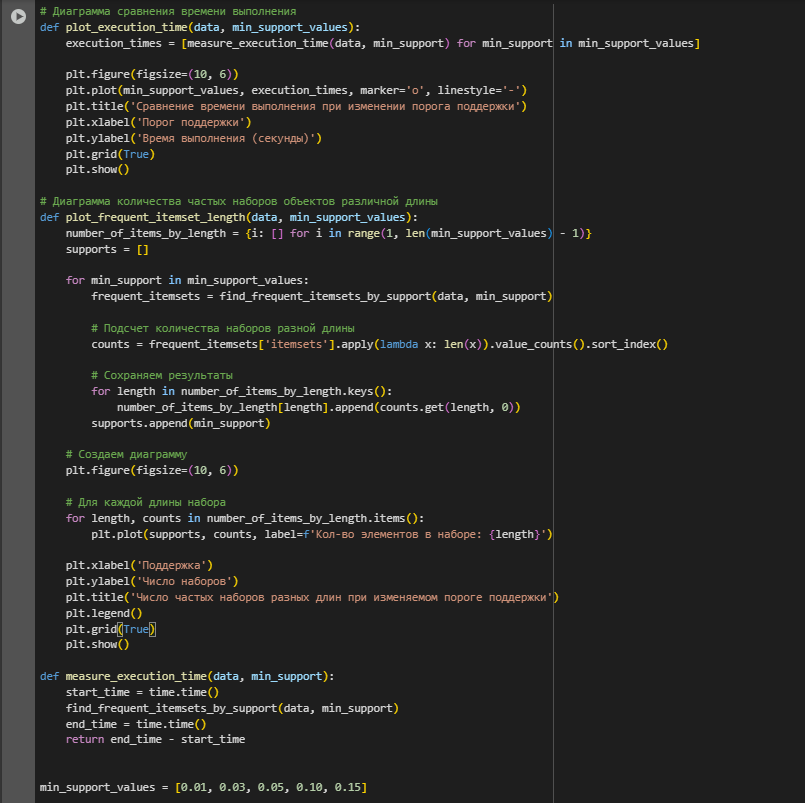


Рисунок 5 – Методы построения графиков



Рисунок 6 – График зависимости поддержки от времени выполнения

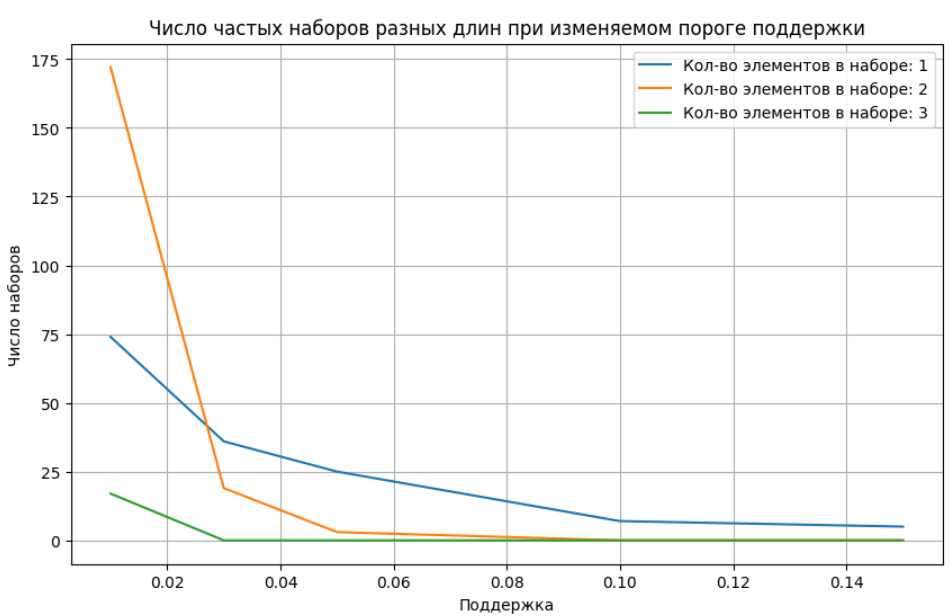


Рисунок 7 – Графики зависимости поддержки от числа наборов с разным количеством элементов