|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  |  | | --- | --- | --- | |  |  |  | | МИНОБРНАУКИ РОССИИ | | | | Федеральное государственное бюджетное образовательное учреждение  высшего образования  **«МИРЭА – Российский технологический университет»**  **РТУ МИРЭА** | | | | |
| Институт информационных технологий (ИИТ) | |
| Кафедра Прикладной Математики (ПМ) | |

|  |  |  |  |
| --- | --- | --- | --- |
| **ОТЧЁТ ПО ПРАКТИЧЕСКОЙ РАБОТЕ №8** | | | |
| **по дисциплине «Технологии и инструментарий анализа больших данных»** | | | |
|  | | | |
| Выполнил студент группы ИКБО-14-20 | | Вежновец Ф. Ю. | |
|  | |  | |
| Принял: асистент | | Горячев А. А. | |
| Практические работы выполнены | «\_\_\_»\_\_\_\_\_\_\_2023 г. | | (подпись студента) | |
| «Зачтено» | «\_\_\_»\_\_\_\_\_\_\_2023 г. | | (подпись руководителя) | |
|  |  | |  | |

Москва 2023

Задание 1

Загрузить данные Market\_Basket\_Optimisation.csv.

Реализация

Результат работы программы и код предоставлена на рисунке 1.

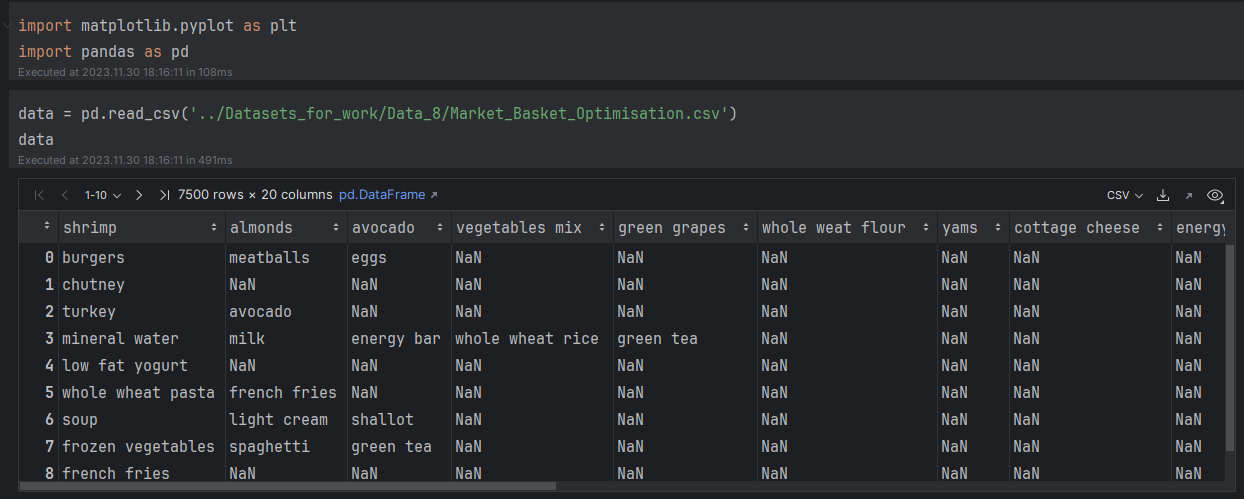


Рисунок 1 – Импорт данных

Задание 2

Визуализировать данные (отразить на гистограммах относительную и фактическую частоту встречаемости для 20 наиболее популярных товаров).

Реализация

Результат работы программы и код предоставлена на рисунке 2 – 4.

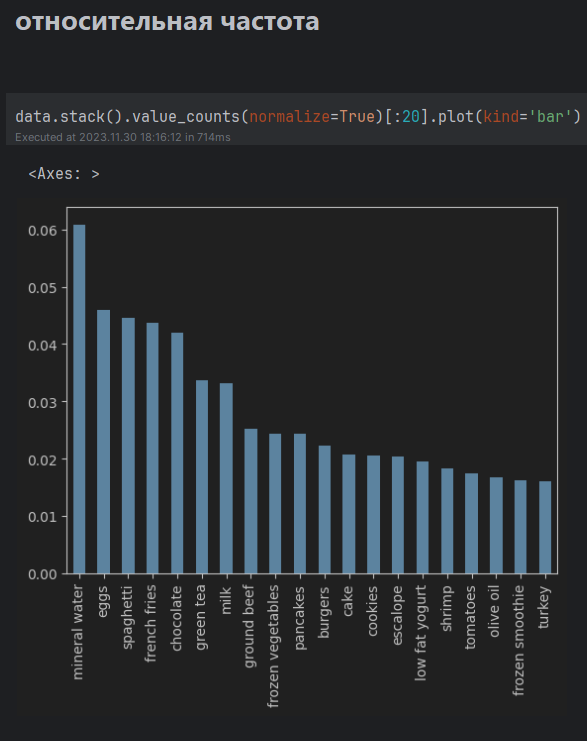


Рисунок 2 – Относительная частота

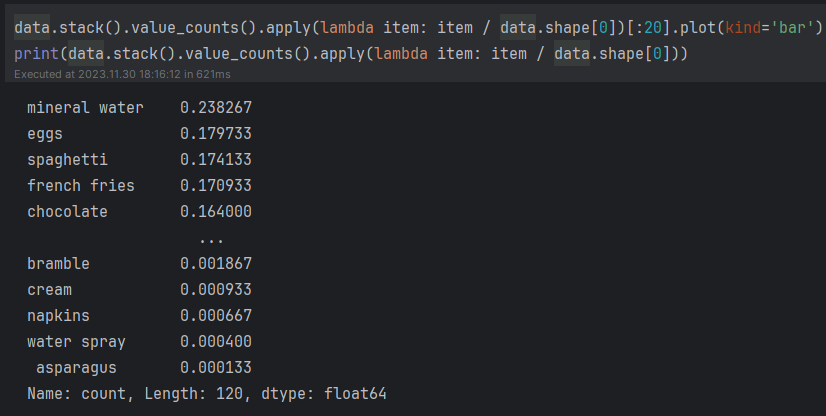


Рисунок 3 – Фактическая частота

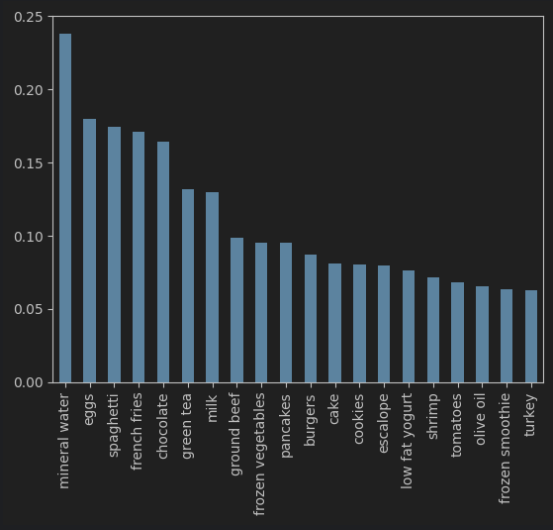


Рисунок 4 – Фактическая частота

Задание 3

Применить алгоритм Apriori, используя 3 разные библиотеки (apriori\_python, apyori, efficient\_apriori). Подобрать гиперпараметры для алгоритмов так, чтобы выводилось порядка 10 наилучших правил.

Реализация

Результат работы программы и код предоставлена на рисунке 5 – 9.

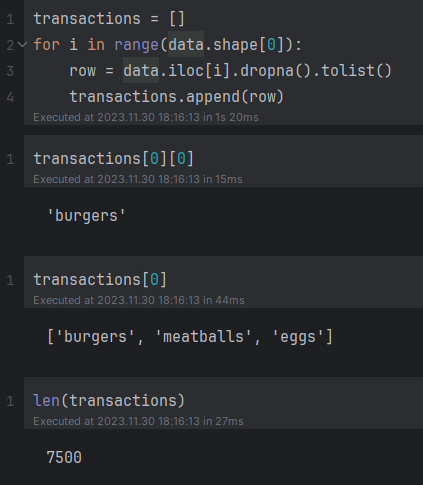


Рисунок 5 – Получаем транзакции

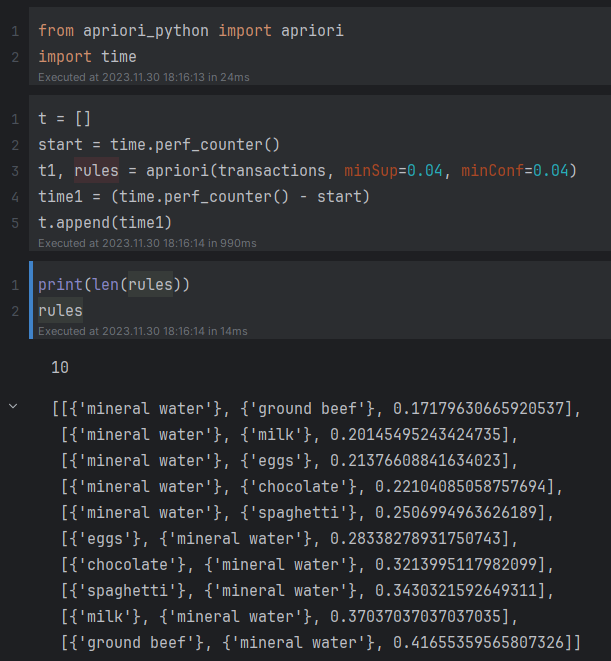


Рисунок 6 – Первый алгоритм априори

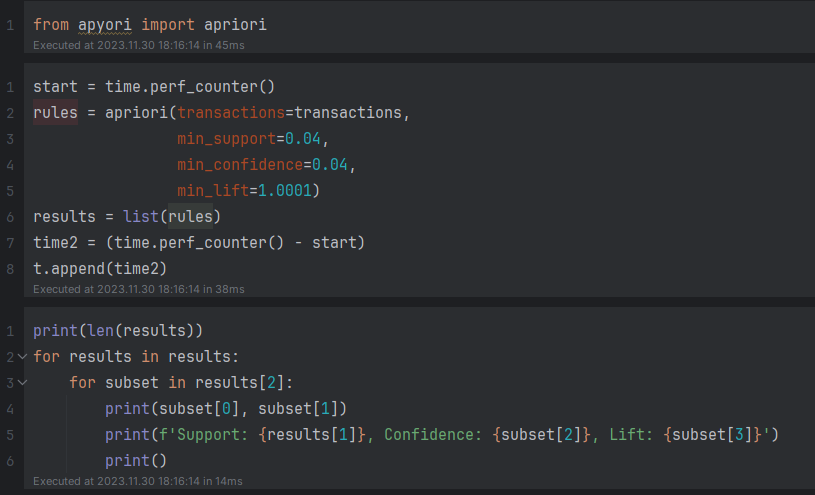


Рисунок 7 – Второй алгоритм априори

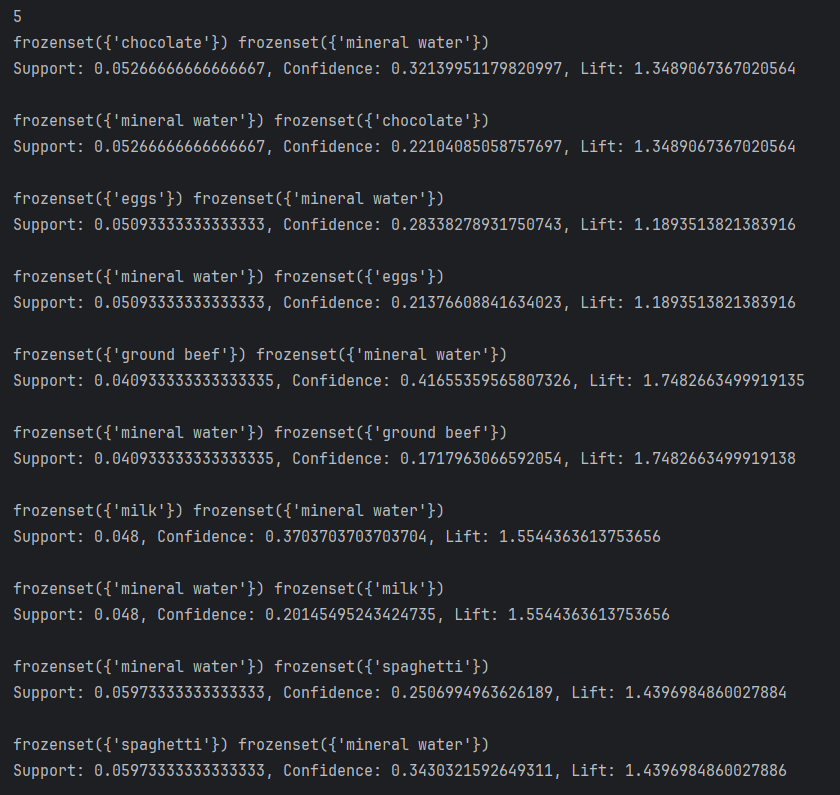


Рисунок 8 – Второй алгоритм априори вывод

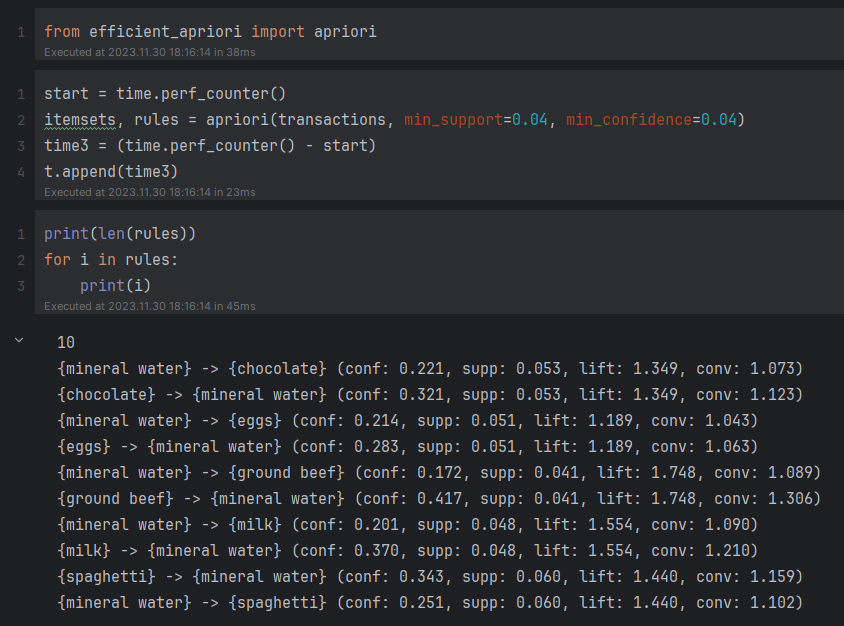


Рисунок 9 – Третий алгоритм априори

Задание 4

Применить алгоритм FP-Growth из библиотеки fpgrowth\_py. Подобрать гиперпараметры для алгоритма так, чтобы выводилось порядка 10 наилучших правил.

Реализация

Результат работы программы и код предоставлена на рисунке 10.

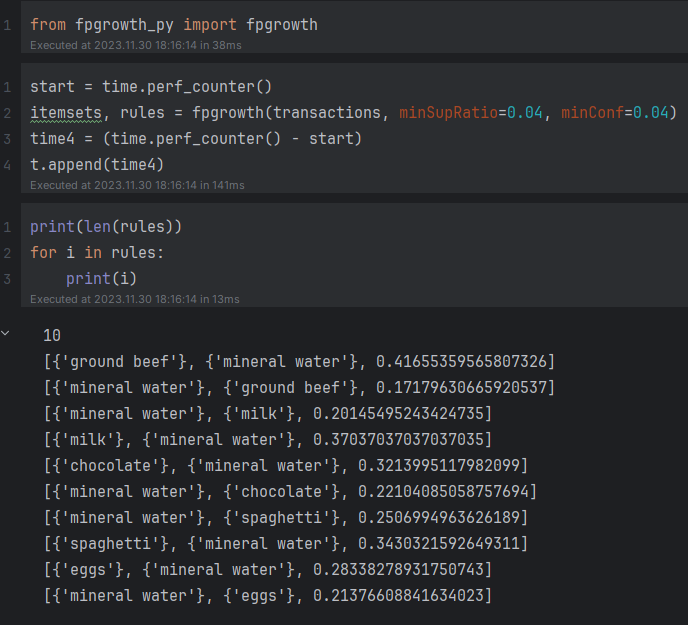


Рисунок 10 – Алгоритм FP-Growth

Задание 5

Сравнить время выполнения всех алгоритмов и построить гистограмму.

Реализация

Результат работы программы и код предоставлена на рисунке 11.



Рисунок 11 – Сравнение

Задание 6

Загрузить данные data.csv.

Реализация

Результат работы программы и код предоставлена на рисунке 12.

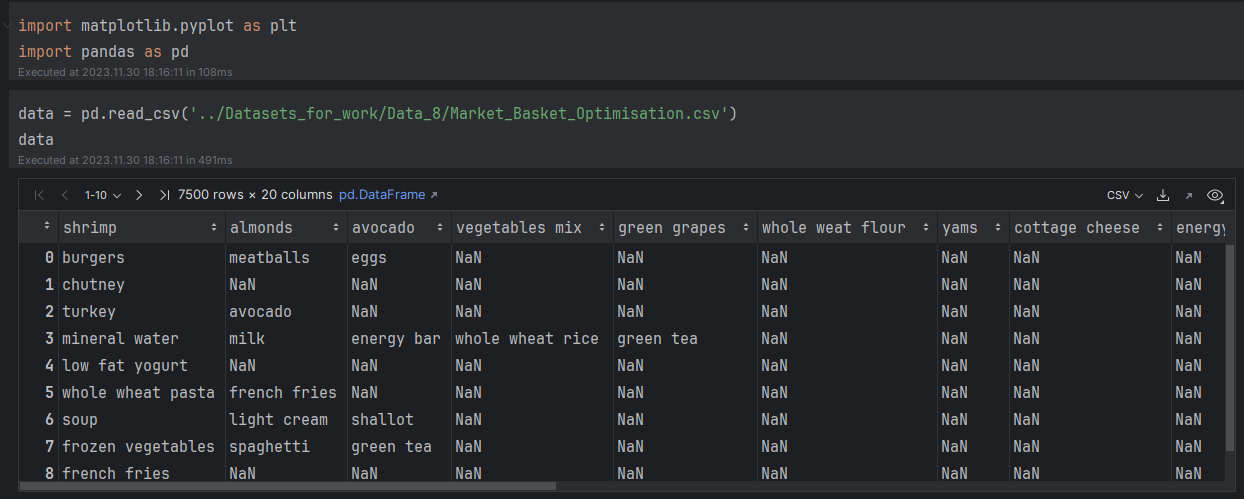


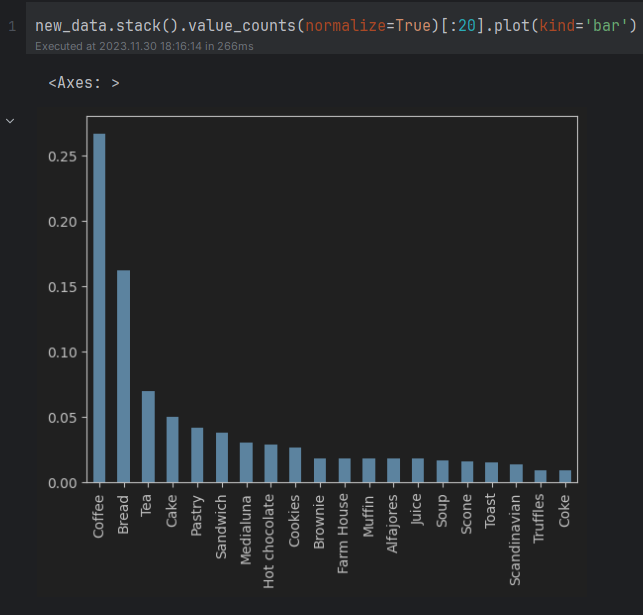
Рисунок 12 – Импорт данных

**Задание 7**

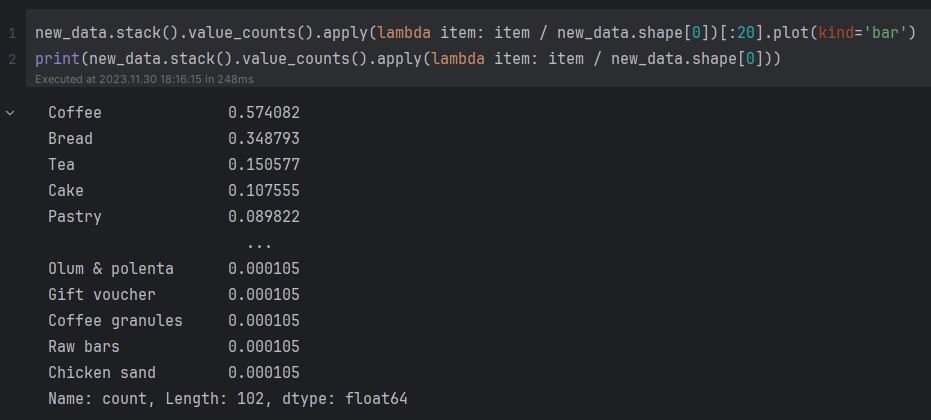
Визуализировать данные (отразить на гистограммах относительную и фактическую частоту встречаемости для 20 наиболее популярных товаров).

**Реализация**

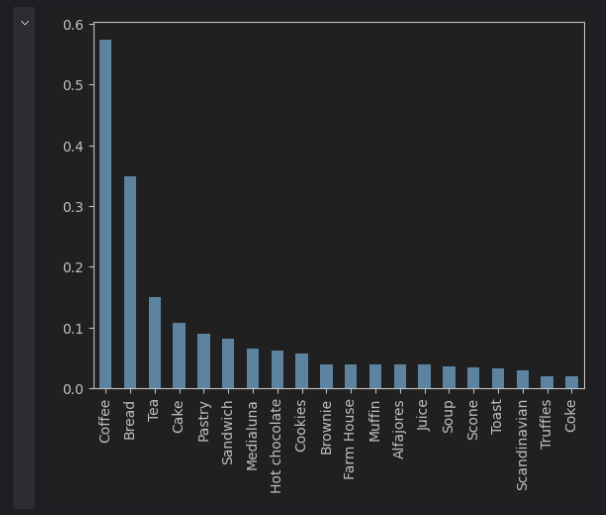
*Результат работы программы и код предоставлена на рисунке 13 –* 15*.*

****

**Рисунок 13 – Относительная частота**

****

**Рисунок 14 – Фактическая частота**

****

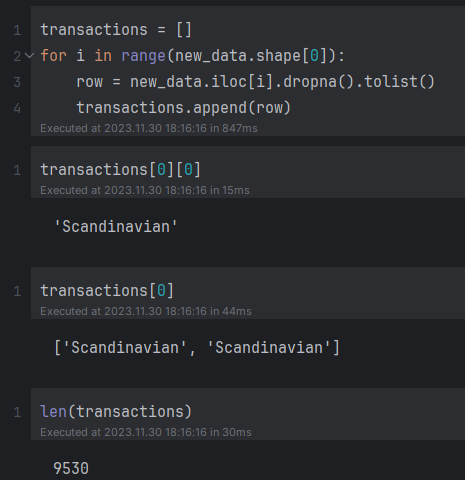
**Рисунок 15 – Фактическая частота**

**Задание 3**

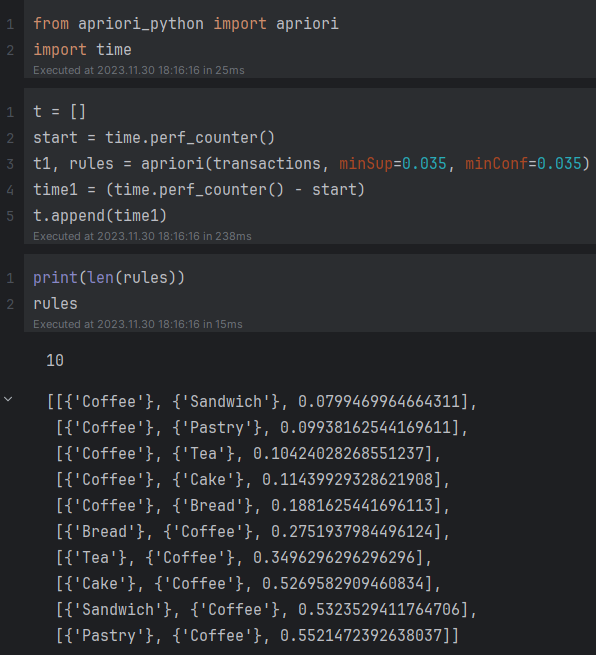
Применить алгоритм Apriori, используя 3 разные библиотеки (apriori\_python, apyori, efficient\_apriori). Подобрать гиперпараметры для алгоритмов так, чтобы выводилось порядка 10 наилучших правил.

**Реализация**

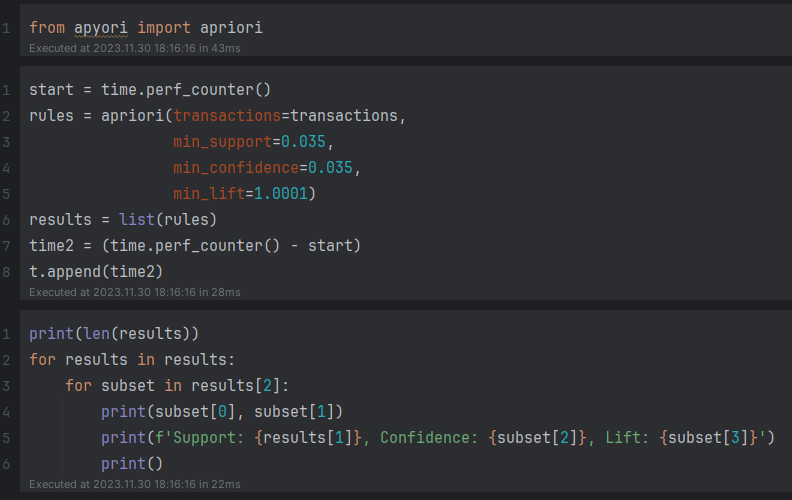
*Результат работы программы и код предоставлена на рисунке* 16 *–* 20*.*

****

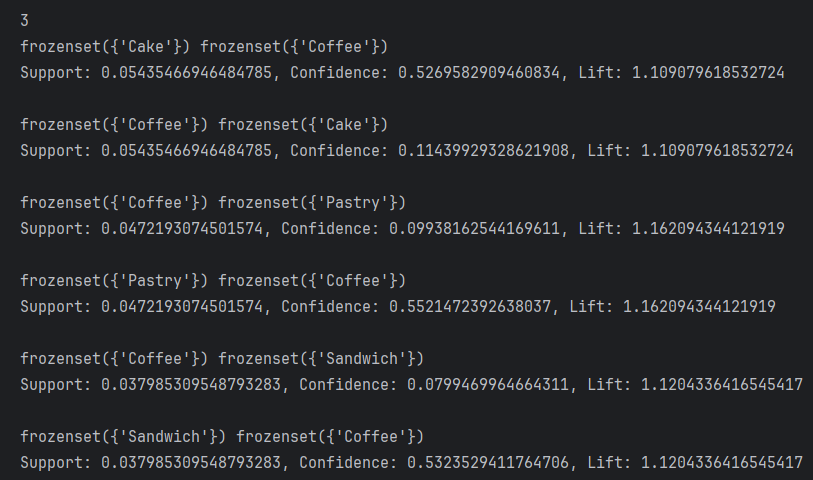
**Рисунок 16 – Получаем транзакции**

****

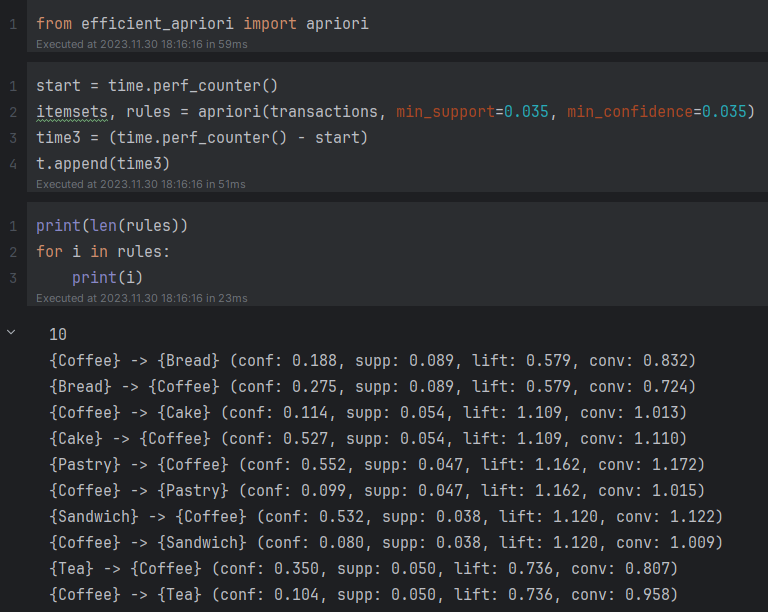
**Рисунок 17 – Первый алгоритм априори**

****

**Рисунок 18 – Второй алгоритм априори**

****

**Рисунок 19 – Второй алгоритм априори вывод**

****

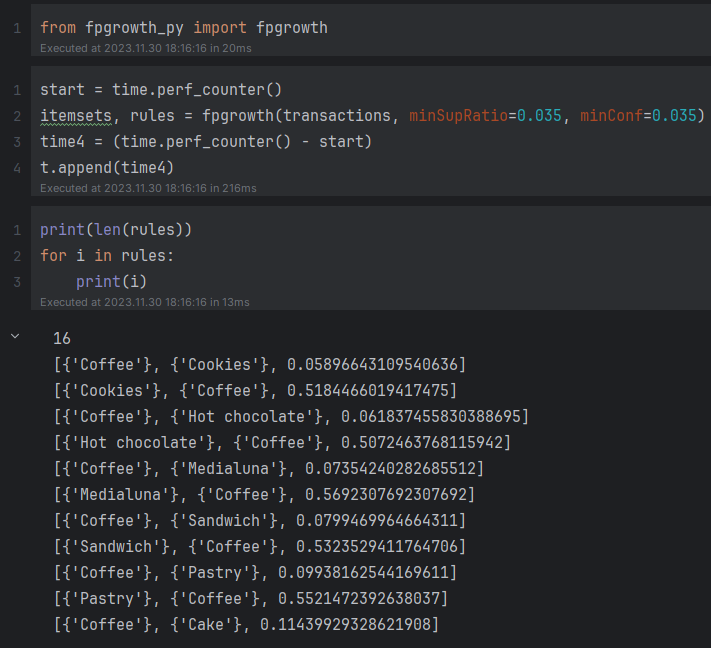
**Рисунок 20 – Третий алгоритм априори**

**Задание 4**

Применить алгоритм FP-Growth из библиотеки fpgrowth\_py. Подобрать гиперпараметры для алгоритма так, чтобы выводилось порядка 10 наилучших правил.

**Реализация**

*Результат работы программы и код предоставлена на рисунке* 21*.*

****

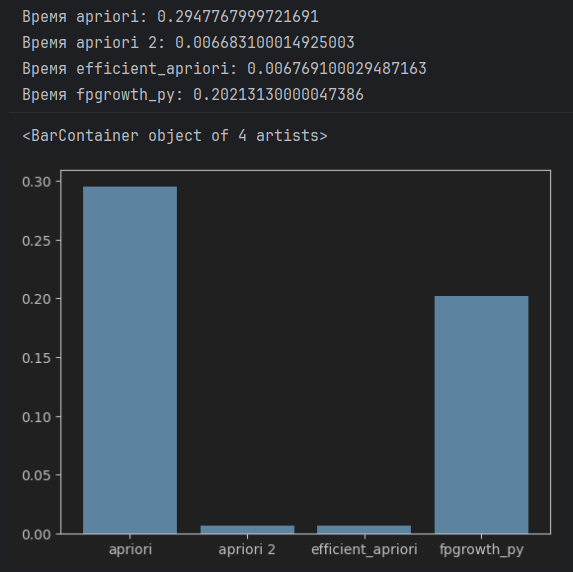
**Рисунок 21 – Алгоритм FP-Growth**

**Задание 5**

Сравнить время выполнения всех алгоритмов и построить гистограмму.

**Реализация**

*Результат работы программы и код предоставлена на рисунке* 22*.*

****

**Рисунок 22 – Сравнение**