|  |  |
| --- | --- |
|  | |
| МИНистерство Науки и высшего образования российской федерации | |
| Федеральное государственное бюджетное образовательное учреждениевысшего образования"МИРЭА - Российский технологический университет"РТУ МИРЭА | |
| Институт информационных технологий (ИТ) | |
| Кафедра прикладной математики (ПМ) | |

|  |  |
| --- | --- |
| **ОТЧЕТ**  **ПО ПРАКТИЧЕСКИМ РАБОТАМ** | |
| **по дисциплине** | |
| «Технологии и инструментарий анализа больших данных» | |
|  | |
| Выполнили студенты группы ИВБО-07-19 | Гридасов Е.А. |
| Принял ассистент кафедры ПМ | Горячев А.А. |

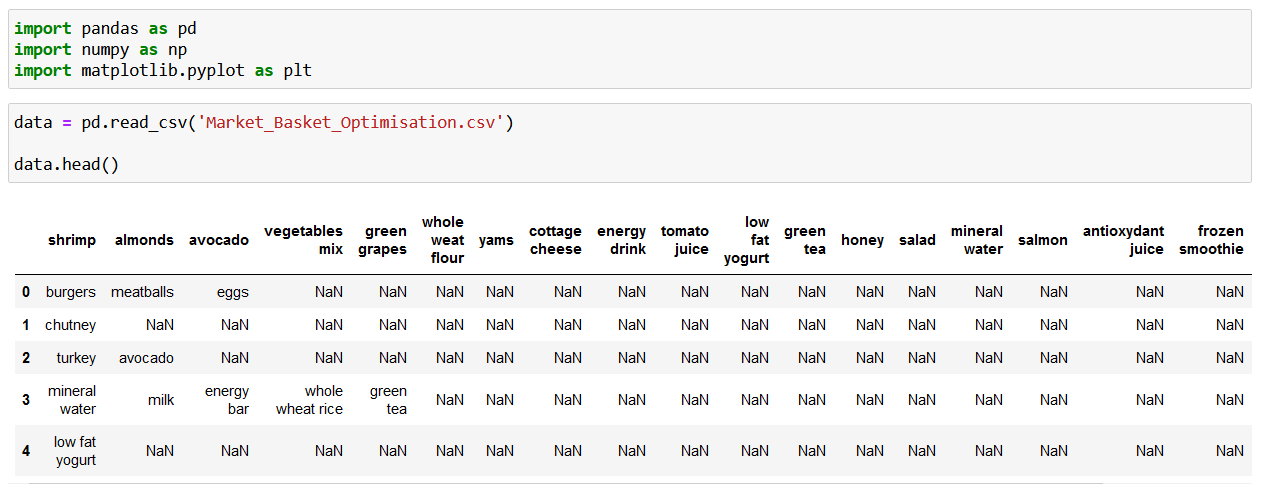
|  |  |  |
| --- | --- | --- |
| Практическая работа выполнена | «\_\_»\_\_\_\_\_\_\_\_\_2022 г. |  |
|  |  |  |
| «Зачтено» | «\_\_»\_\_\_\_\_\_\_\_\_2022 г. |  |

Москва 2022

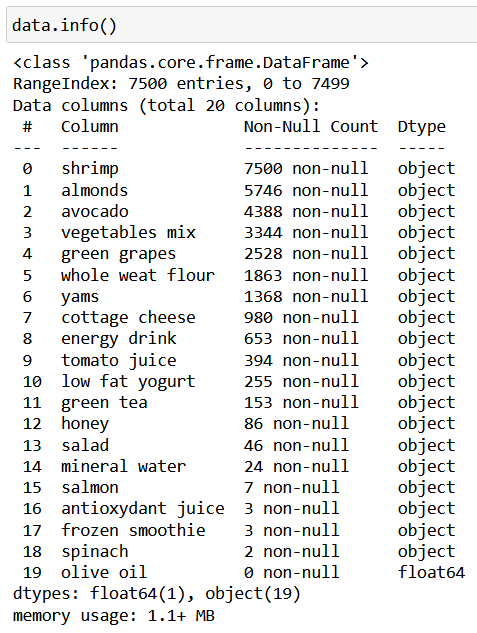
# ВЫПОЛНЕНИЕ ЗАДАЧ

Целью работы было изучение ассоциативных правил с помощью библиотек Python.

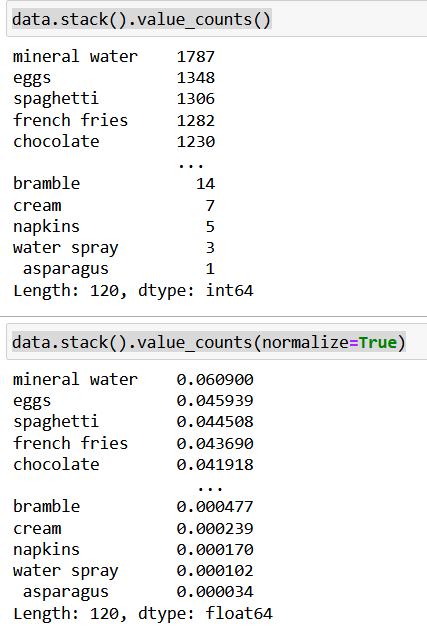
Для выполнения задания был скачан датасет Market\_Basket\_Optimization, для которого были применены необходимые методы.



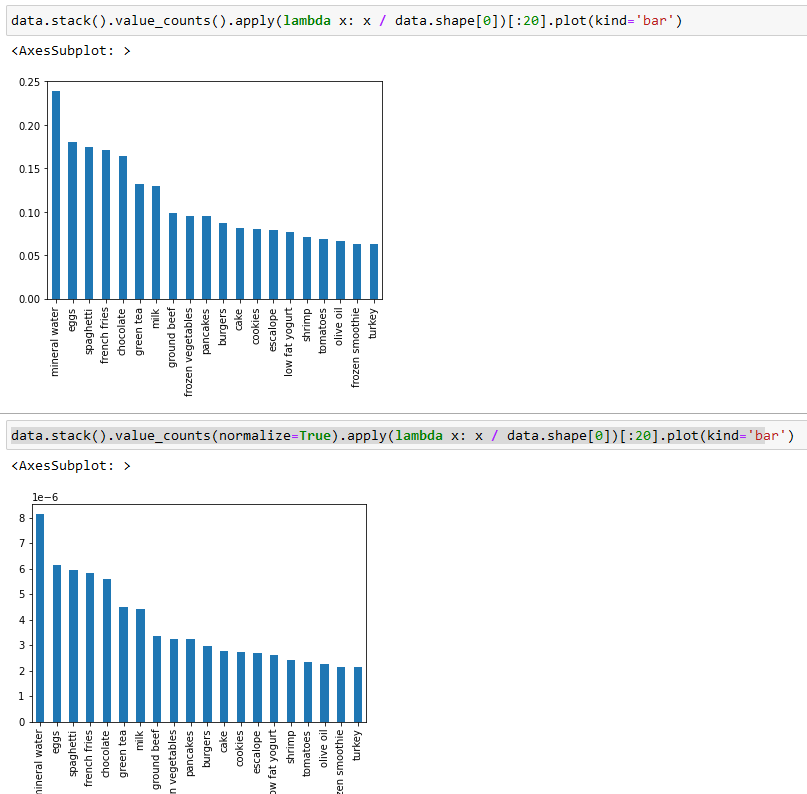
**Рисунок 1 – Датасет Market\_Basket\_Optimization**



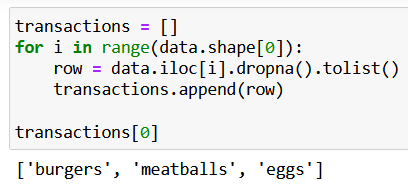
**Рисунок 2 — Изучение датасета**



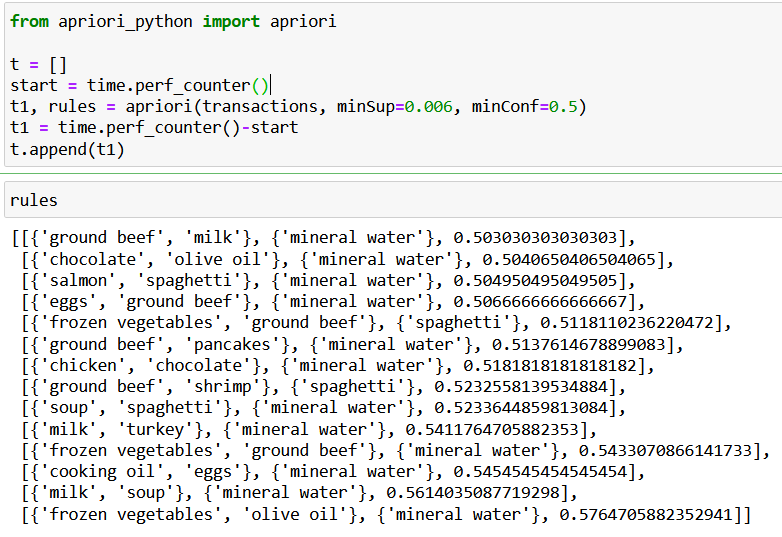
**Рисунок 3 — Изучение частоты встречаемых значений**



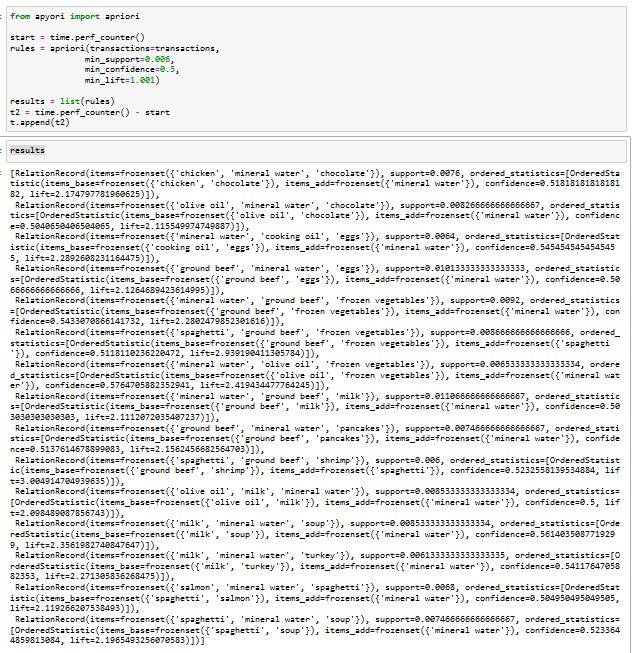
**Рисунок 4 — Визуализация частоты значений**



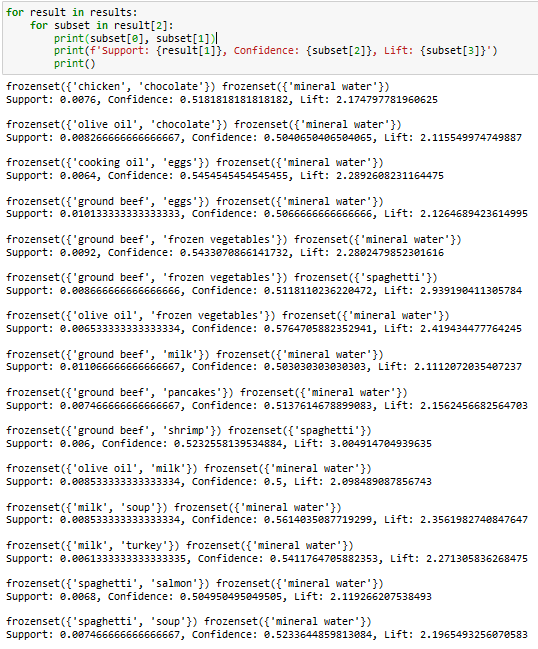
**Рисунок 5 — Создание транзакций**



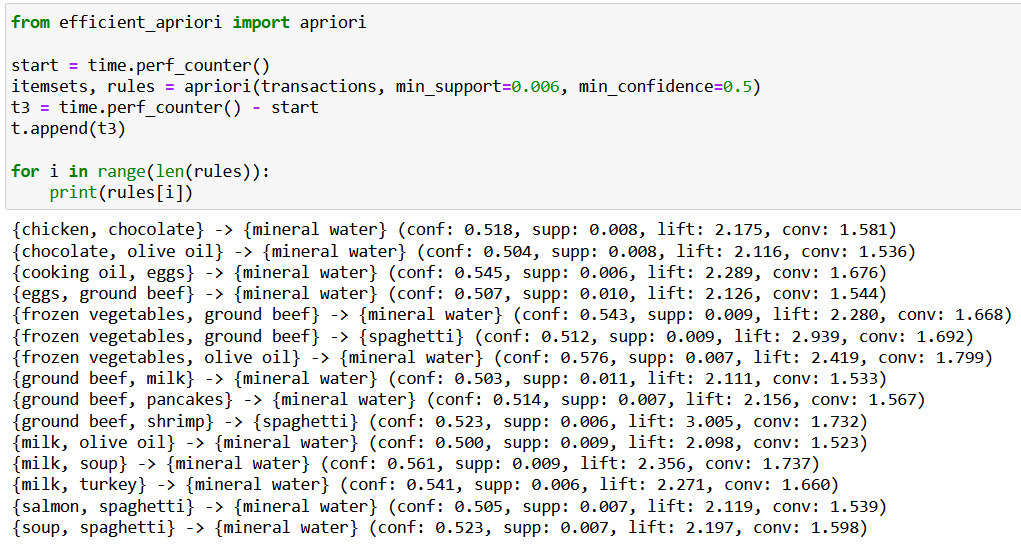
**Рисунок 6 — Применение apriori\_python**



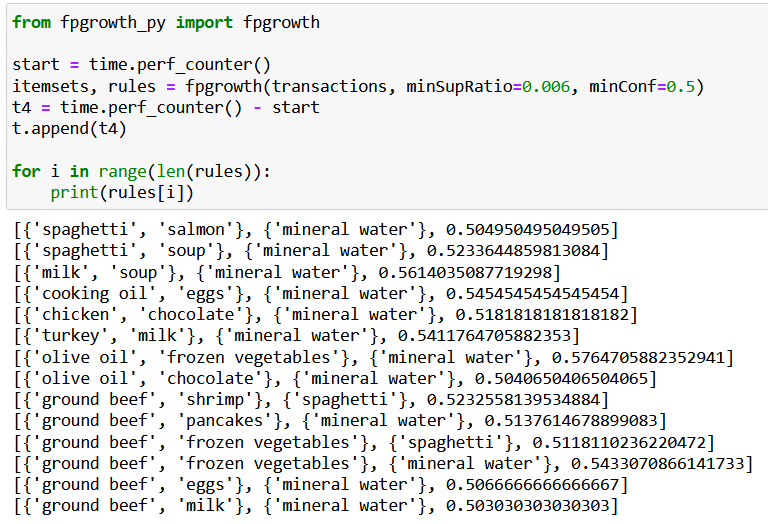
**Рисунок 7 — Применение apyori**



**Рисунок 8 — Форматированный вывод**



**Рисунок 9 — Применение efficient\_apriori**

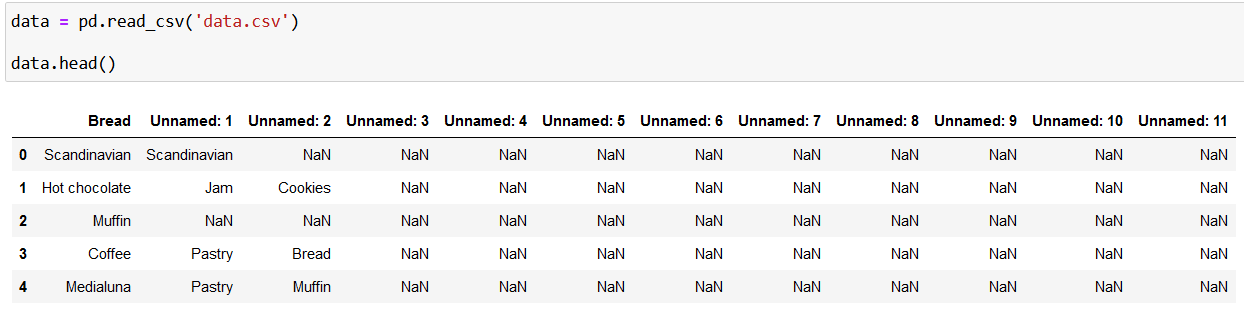


**Рисунок 10 — Применение fpgrowth\_py**

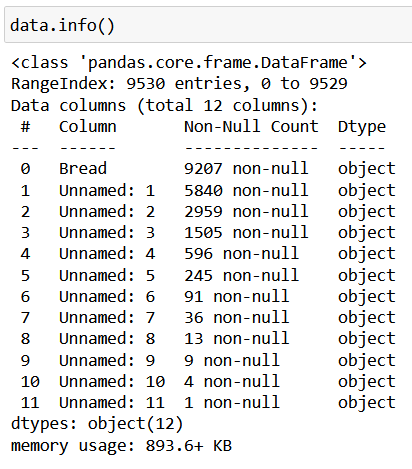


**Рисунок 11 — Сравнение времени исполнения методов**

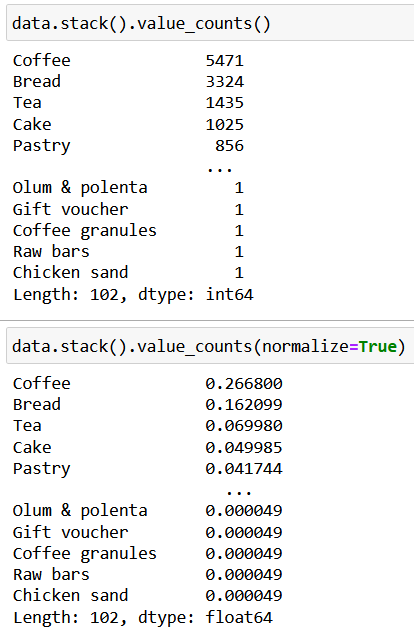
Далее был скачан датасет data, с которым были проведены такие же манипуляции.



**Рисунок 12 – Датасет data**



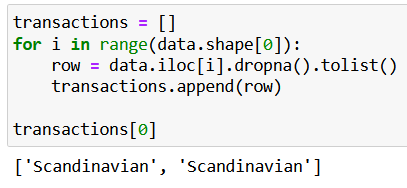
**Рисунок 13 — Изучение датасета**



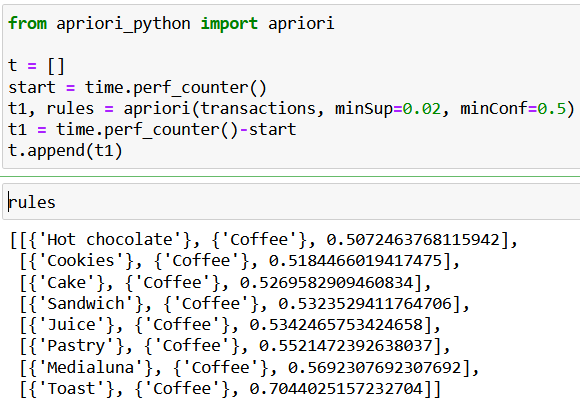
**Рисунок 14 — Изучение частоты встречаемых значений**



**Рисунок 15 — Визуализация частоты значений**



**Рисунок 16 — Создание транзакций**



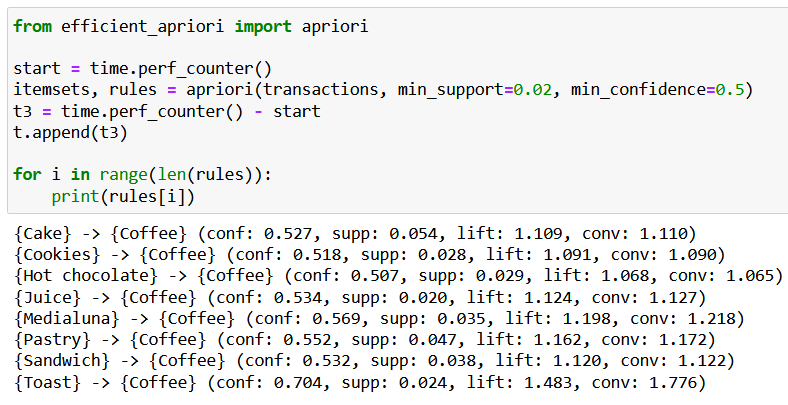
**Рисунок 17 — Применение apriori\_python**



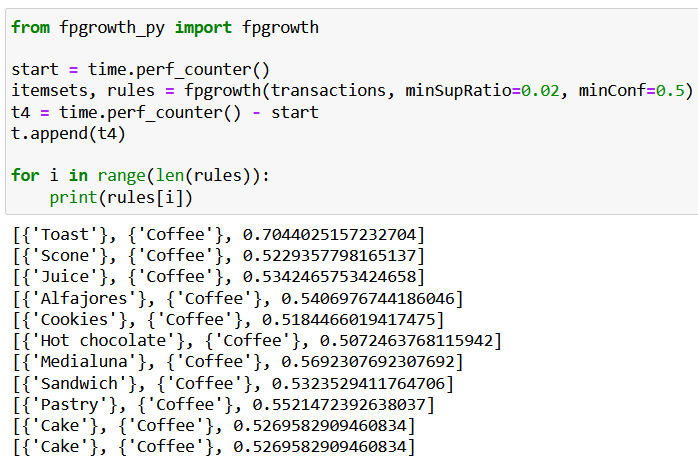
**Рисунок 18 — Применение apyori**



**Рисунок 19 — Форматированный вывод**



**Рисунок 20 — Применение efficient\_apriori**



**Рисунок 21 — Применение fpgrowth\_py**



**Рисунок 11 — Сравнение времени исполнения методов**

По результату были сделаны выводы: скорость исполнения методов на основе ассоциативных правил во многом зависит от количества транзакций и размера датасета.

# ВЫВОДЫ

В ходе выполнения работы были изучены библиотеки matplotlib, plotly, sklearn, scipy. Освоена тема ассоциативных правил.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Python Documentation [Электронный ресурс] - <https://www.python.org/>
2. Jupyter Documentation [Электронный ресурс] - <https://docs.jupyter.org/en/latest/index.html>
3. Pandas Documentation [Электронный ресурс] - <https://pandas.pydata.org/docs/getting_started/index.html>
4. Matplotlib Documentation [Электронный ресурс] - <https://matplotlib.org/>
5. Plotly Documentation [Электронный ресурс] - <https://plotly.com/python/>
6. Numpy Documentation [Электронный ресурс] - <https://numpy.org/doc/>
7. Stats Documentation [Электронный ресурс] - <https://docs.scipy.org/doc/scipy/reference/stats.html>