Задания

1. Применить для тестовых вариантов и вариантов из репозиториев различные алгоритмы поиска ассоциативных правил при одинаковых начальных условиях (при одинаковых пороговых значениях для поддержки и достоверности) и сравнить полученные результаты. Для тестовых вариантов выполнить ручные расчеты (например, с применением MS Excel) и расчеты с применением программных библиотек на языке Python. Для вариантов из репозиториев выполнить расчеты с применением программных библиотек на языке Python.

В качестве алгоритмов поиска ассоциативных правил использовать алгоритмы:

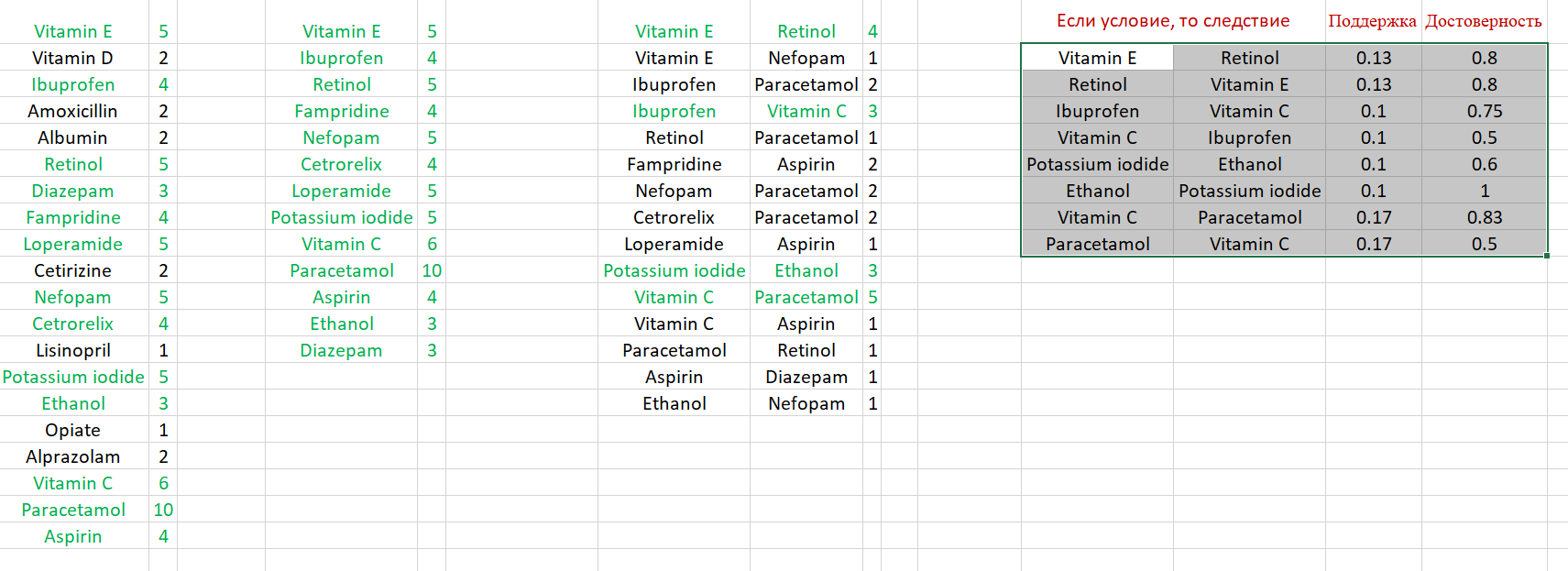
−Apriori (https://pypi.org/project/apriori-python/);

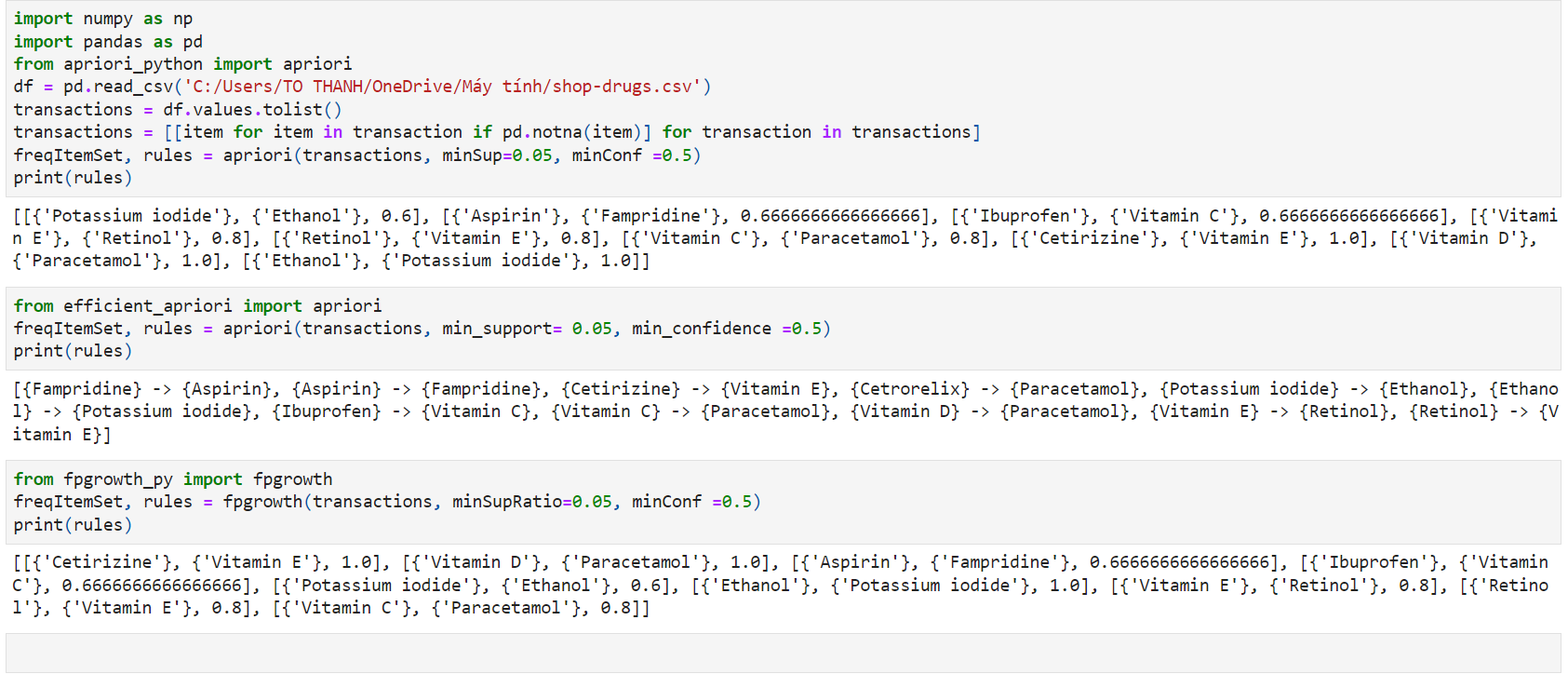
−Efficient Apriori (https://pypi.org/project/efficient-apriori/);

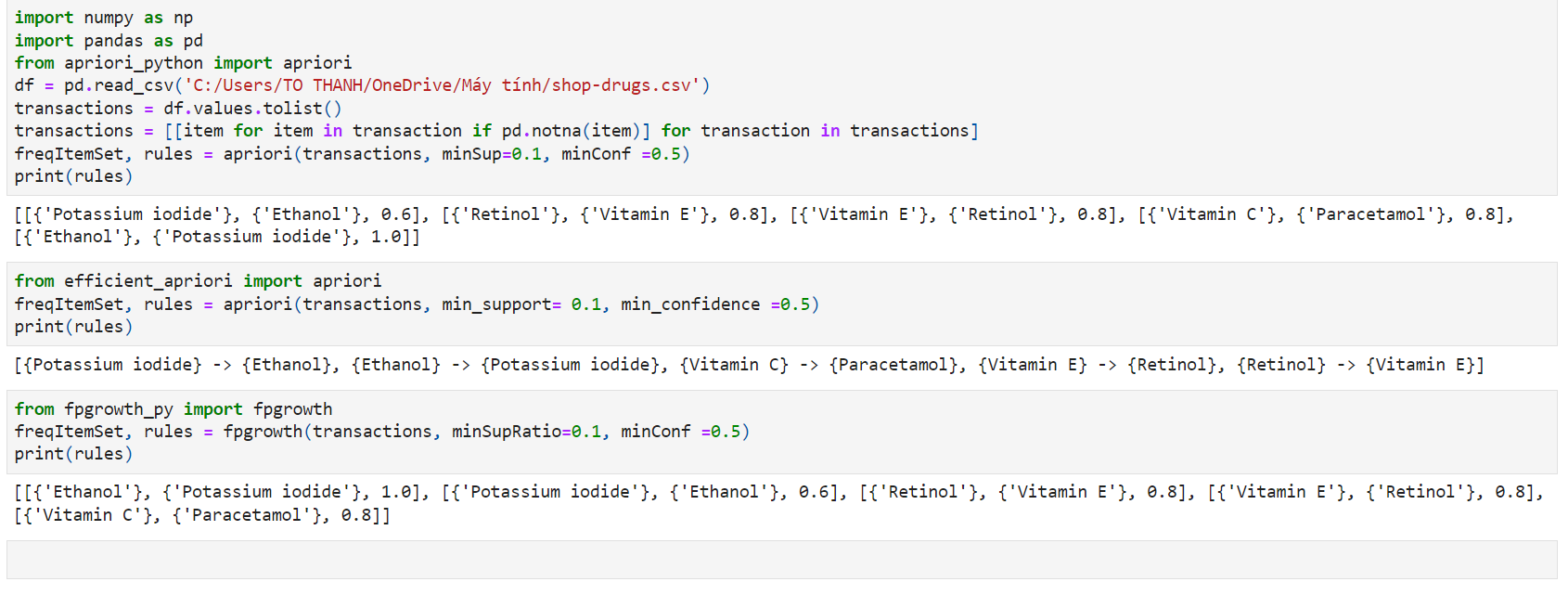
−FPGrowth (<https://pypi.org/project/fpgrowth-py/>).

Для тестовых вариантов:

Вариант 2. Сформировать набор из 30 транзакций на основе чеков покупок в аптеке.



****



После ручного расчета мы видим, что правил ассоциации гораздо больше по сравнению с использованием функций Python.

то есть, {Ibuprofel}->{Vitamin C}, и minConf = 0.5 {Vitamin C}-> {Ibuprofel}

{Paracetamol}->{Vitamin C}. Остальные правила ассоциации верны.

Для вариантов из репозиториев:





При одинаковых пороговых значениях для поддержки и достоверности: Для транзакций из больших хранилищ данных из-за большего количества транзакций, когда мы увеличиваем уровень поддержки с 0,05 до 0,1, правило ассоциации будет равно 0.

При уровне поддержки = 0,05 и достоверности = 0,5, мы видим, что тестовый вариант имеет большее количество ассоциативных правил.

2. Сформировать базы ассоциативных правил с уровнем минимальной достоверности 60% и 80%. Вычислить для ассоциативных правил поддержку, достоверность, значимость

Из 30 транзакций на основе чеков покупок в аптеке базы ассоциативных правил с уровнем минимальной достоверности 60% это:

{Potassium iodide} -> {Ethanol}

{Ethanol} -> {Potassium iodide}

{Vitamin C} -> {Paracetamol}

{Vitamin E} -> {Retinol}

{Retinol} -> {Vitamin E}

{Ibuprofen} -> {Vitamin C}

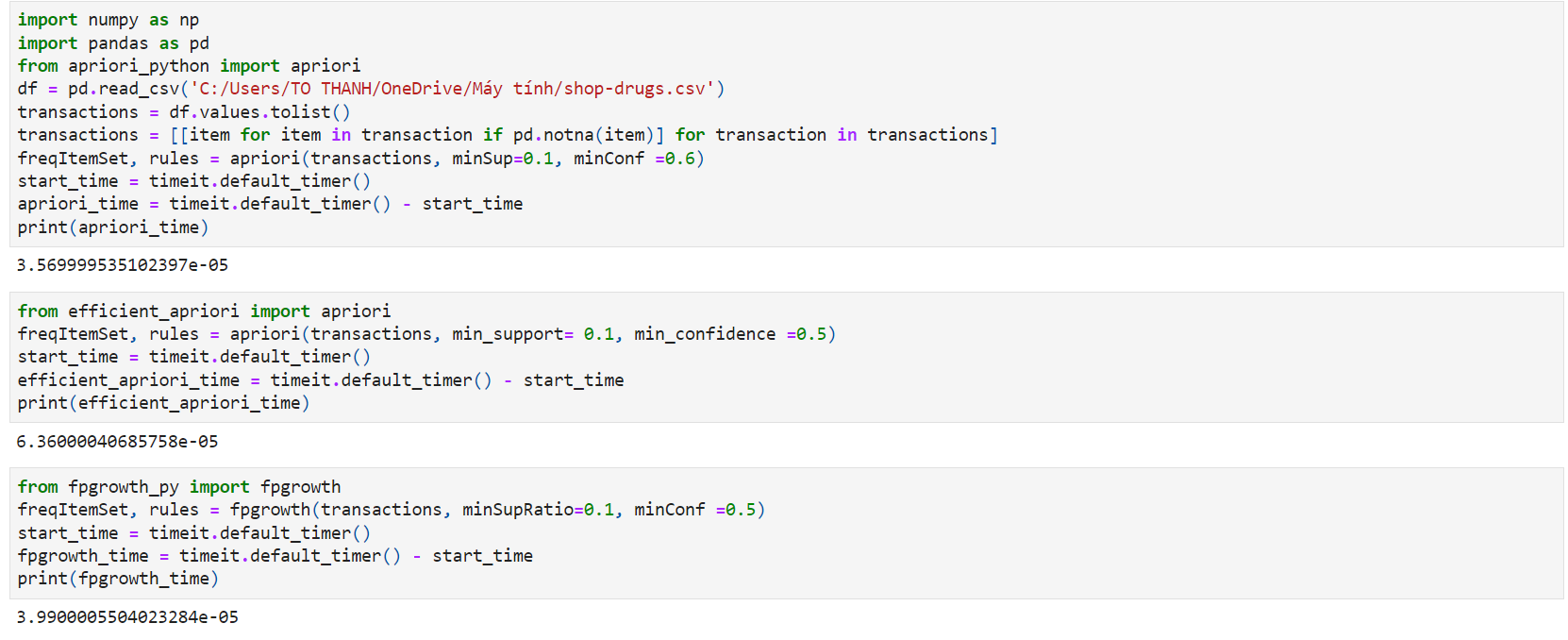
Из 30 транзакций на основе чеков покупок в аптеке базы ассоциативных правил с уровнем минимальной достоверности 80% это:

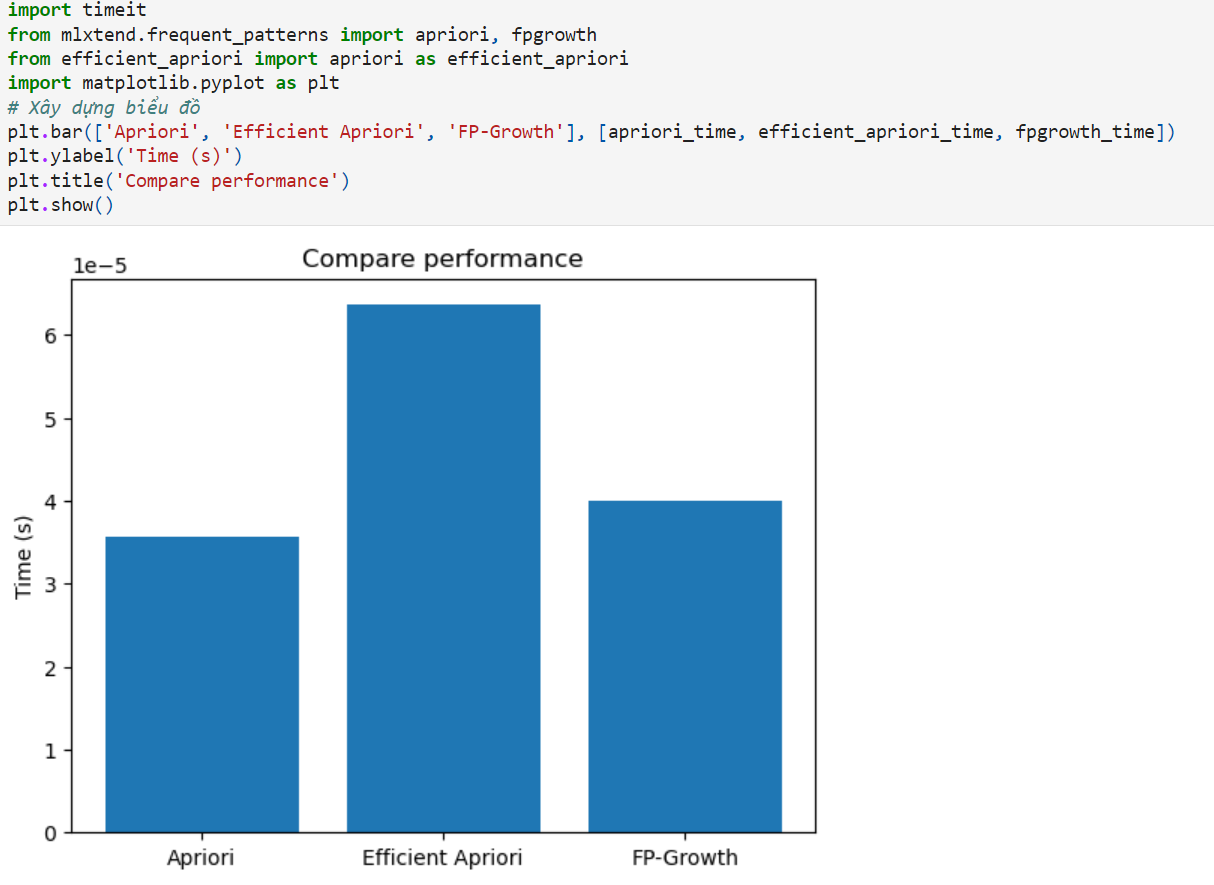
{Vitamin E} -> {Retinol}, {Vitamin C} -> {Paracetamol}

{Retinol} -> {Vitamin E}, {Ethanol} -> {Potassium iodide}

|  |  |  |  |
| --- | --- | --- | --- |
| Если условие, то следствие | Поддержка S | Достоверность С | *L* **(Лифт)** |
| {Potassium iodide} -> {Ethanol} | 0.1 | 0.6 | 6 |
| {Ethanol} -> {Potassium iodide} | 0.1 | 1 | 6 |
| {Vitamin C} -> {Paracetamol} | 0.17 | 0.83 | 2.5 |
| {Vitamin E} -> {Retinol} | 0.13 | 0.8 | 4.8 |
| {Retinol} -> {Vitamin E} | 0.13 | 0.8 | 4.8 |
| {Ibuprofen} -> {Vitamin C} | 0.1 | 0.75 | 3.75 |

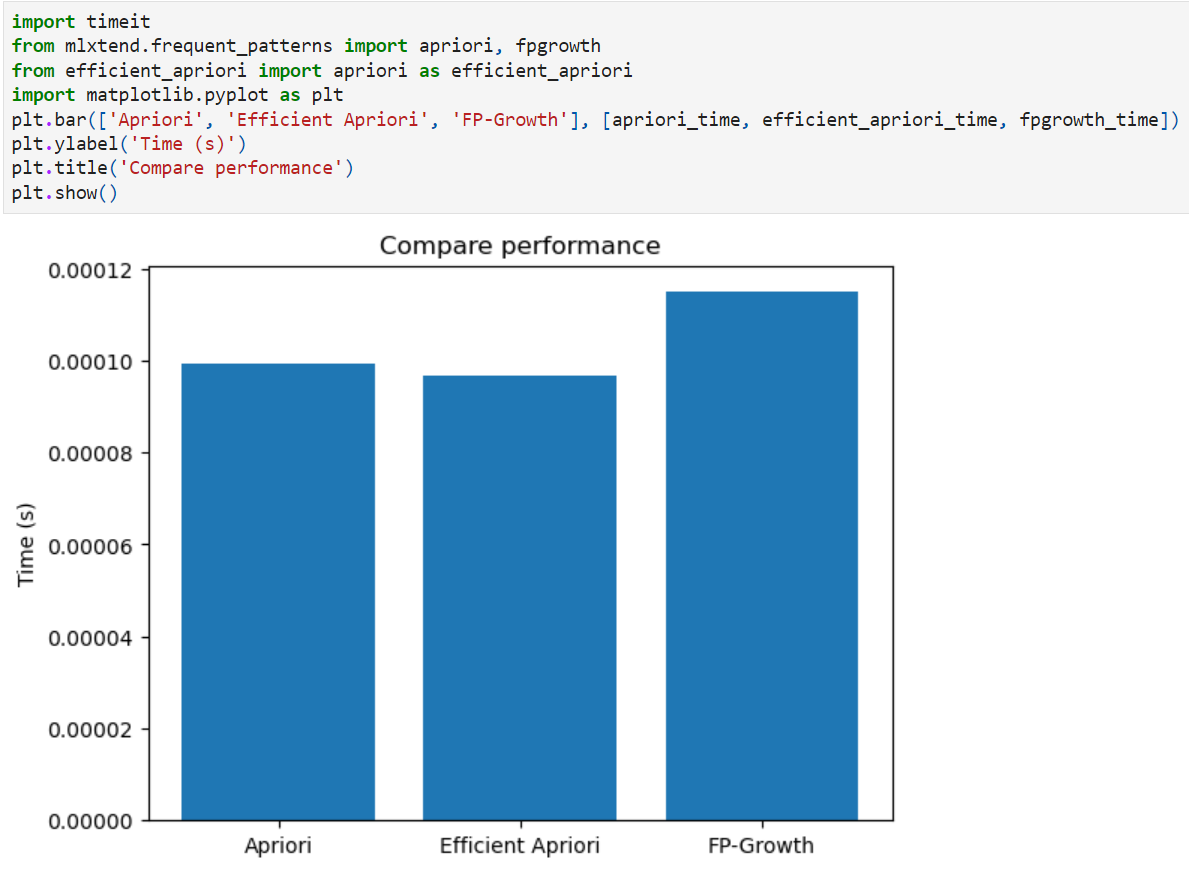
3. Оценить время формирования искомых ассоциативных правил с применением различных алгоритмов и построить диаграммы, позволяющие выполнить сравнительный анализ.





Для тестовых вариантов:





Для вариантов из репозиториев:

4. Выполнить визуализацию ассоциативных правил

