

Аффинитивный анализ чеков покупок

Автор: Шакиров Тимур ИУ5-55Б

Сбор и подготовка данных

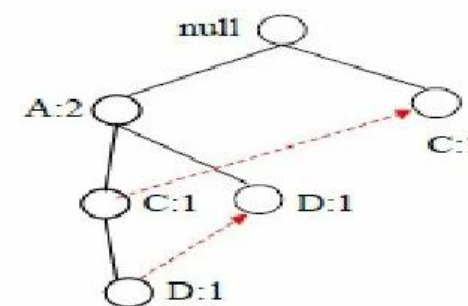
ID Транзакции		Дата	Продукт	Кол-во	Цена, Руб.	Итоговая стоимость, Руб.
0	191	2024-11-21	Колбаса	4.000	250.00	1000.00000
1	191	2024-11-21	Бананы	5.000	90.00	450.00000
2	191	2024-11-21	Огурцы	4.000	100.00	400.00000
3	191	2024-11-21	Сливки	3.000	70.00	210.00000
4	191	2024-11-21	Алкоголь	4.000	300.00	1200.00000

- Сбор данных
- Предобработка данных
- Свёрдживание данных
- Использование Google Colab

Методы и алгоритмы

- Apriori Algorithm
- Association Rule Mining
- FP-growth Algorithm
- Статистические методы

TID	Items
1	{A,B}
2	{B,C,D}
3	{A,C,D,E}
4	{A,D,E}
5	{A,B,C}
6	{A,B,C,D}
7	{A}
8	{A,B,C}
9	{A,B,D}
10	{B,C,E}



Transaction ID	Items Bought
2000	A,B,C
1000	A,C
4000	A,D
5000	B,E,F

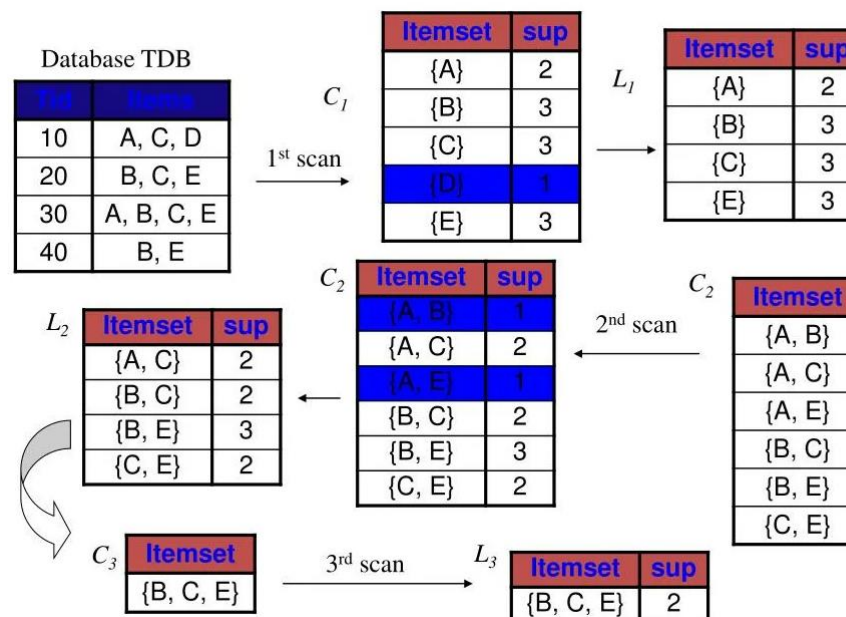
Min. support 50%
Min. confidence 50%

Frequent Itemset	Support
{A}	75%
{B}	50%
{C}	50%
{A,C}	50%

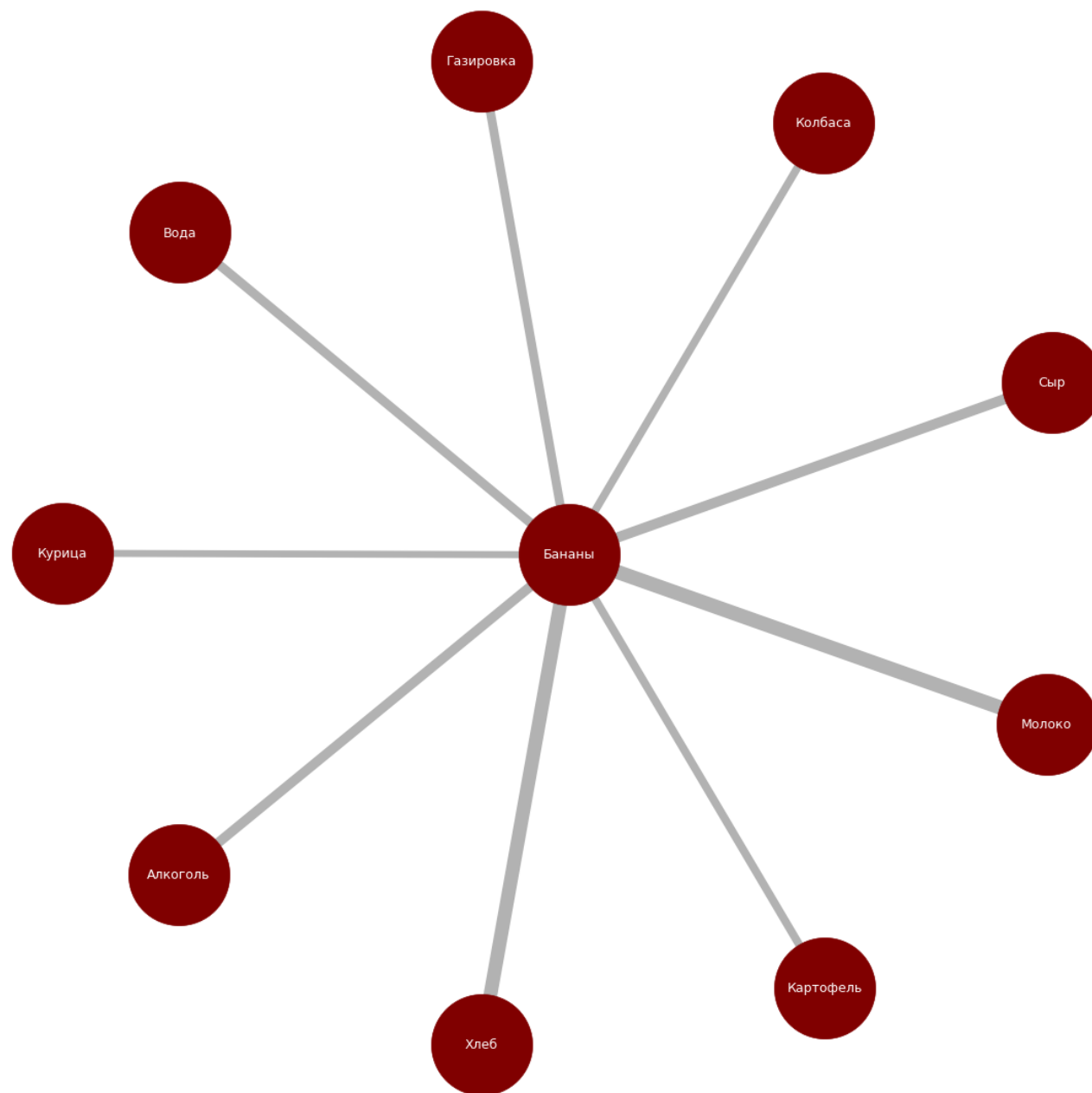
For rule $A \Rightarrow C$:

support = support($\{A \cup C\}$) = 50%

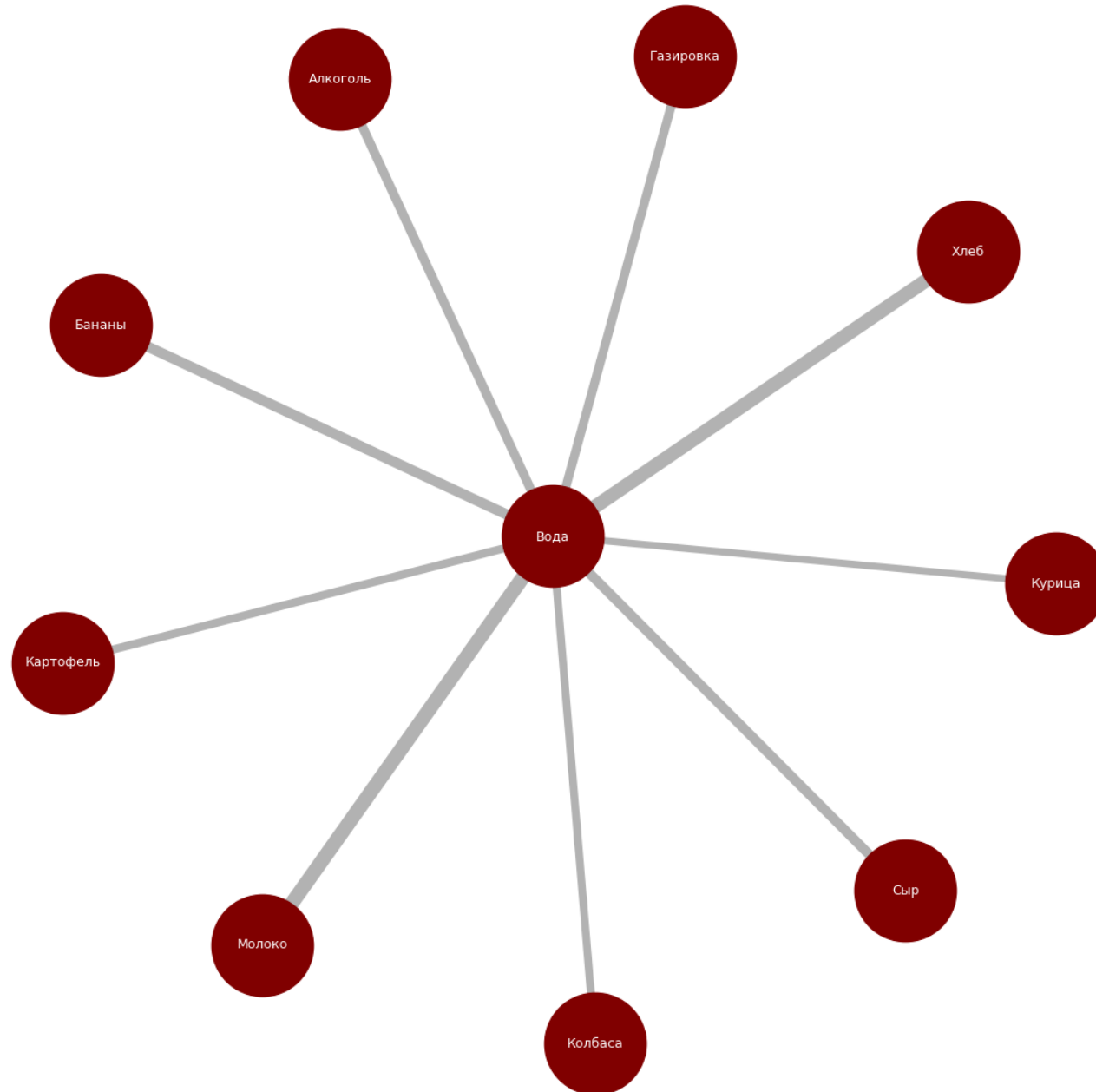
confidence = support($\{A \cup C\}$) / support($\{A\}$) = 66.6%



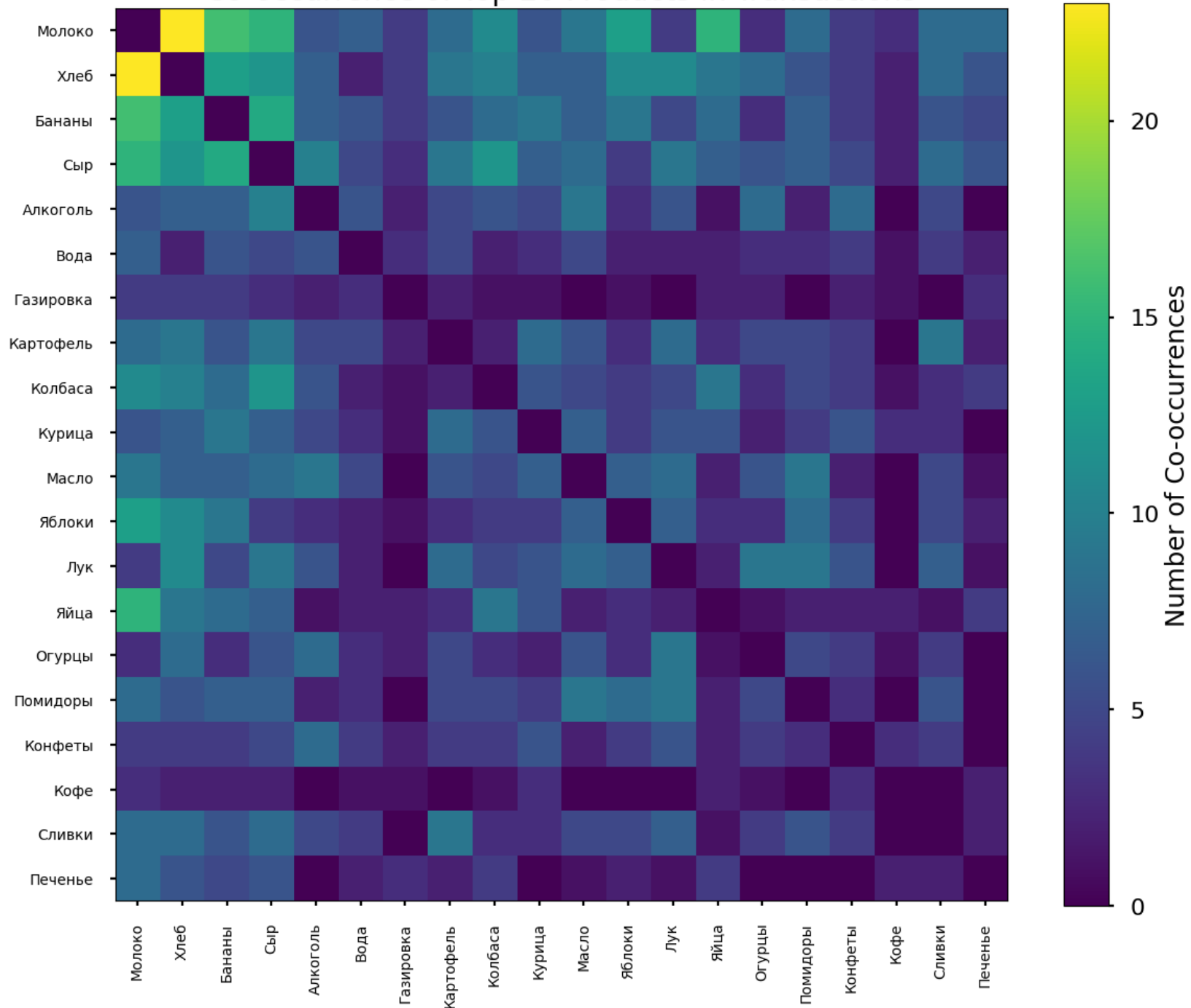
10 самых покупаемых продуктов с Бананы

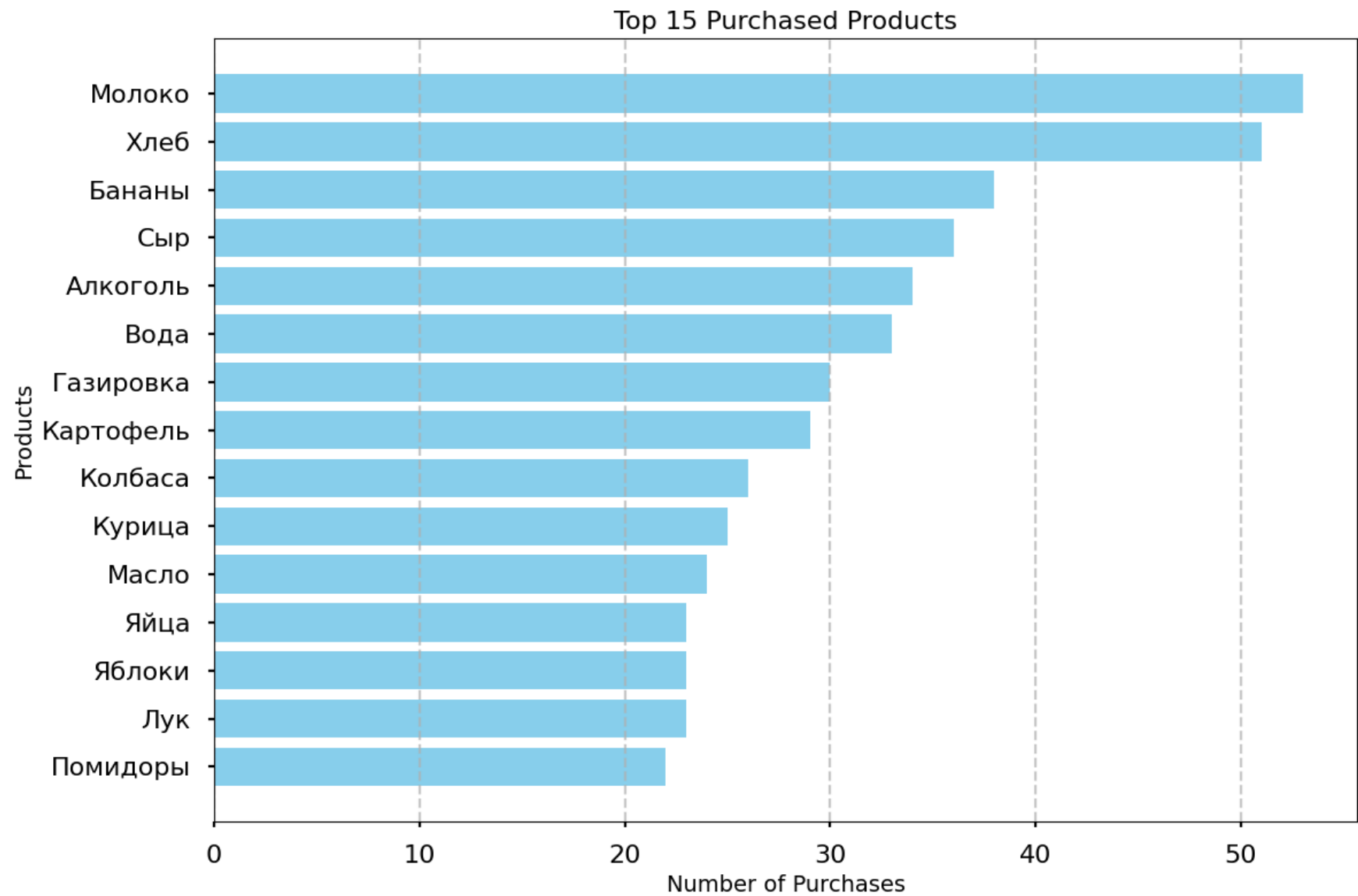


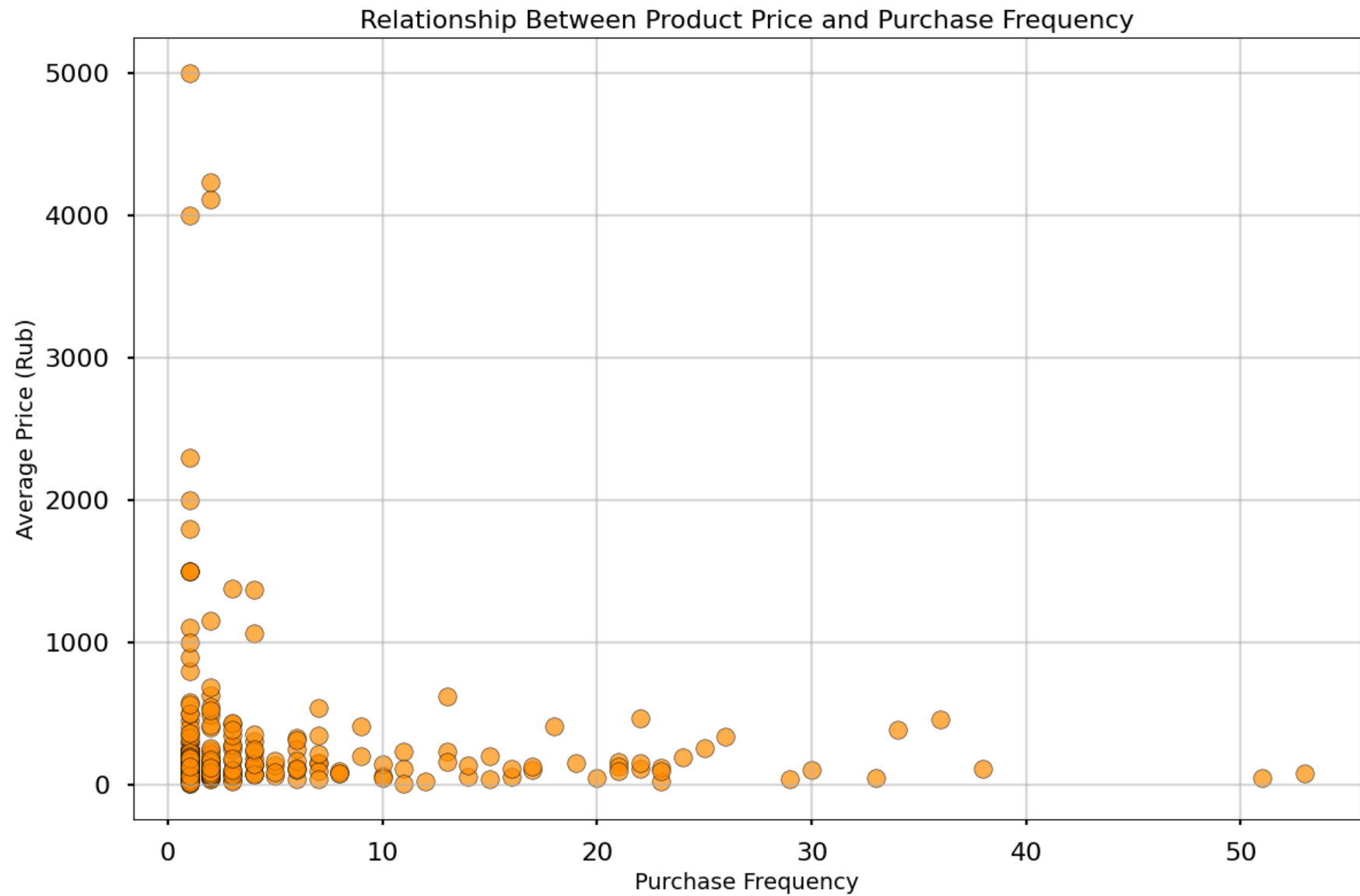
10 самых покупаемых продуктов с Вода



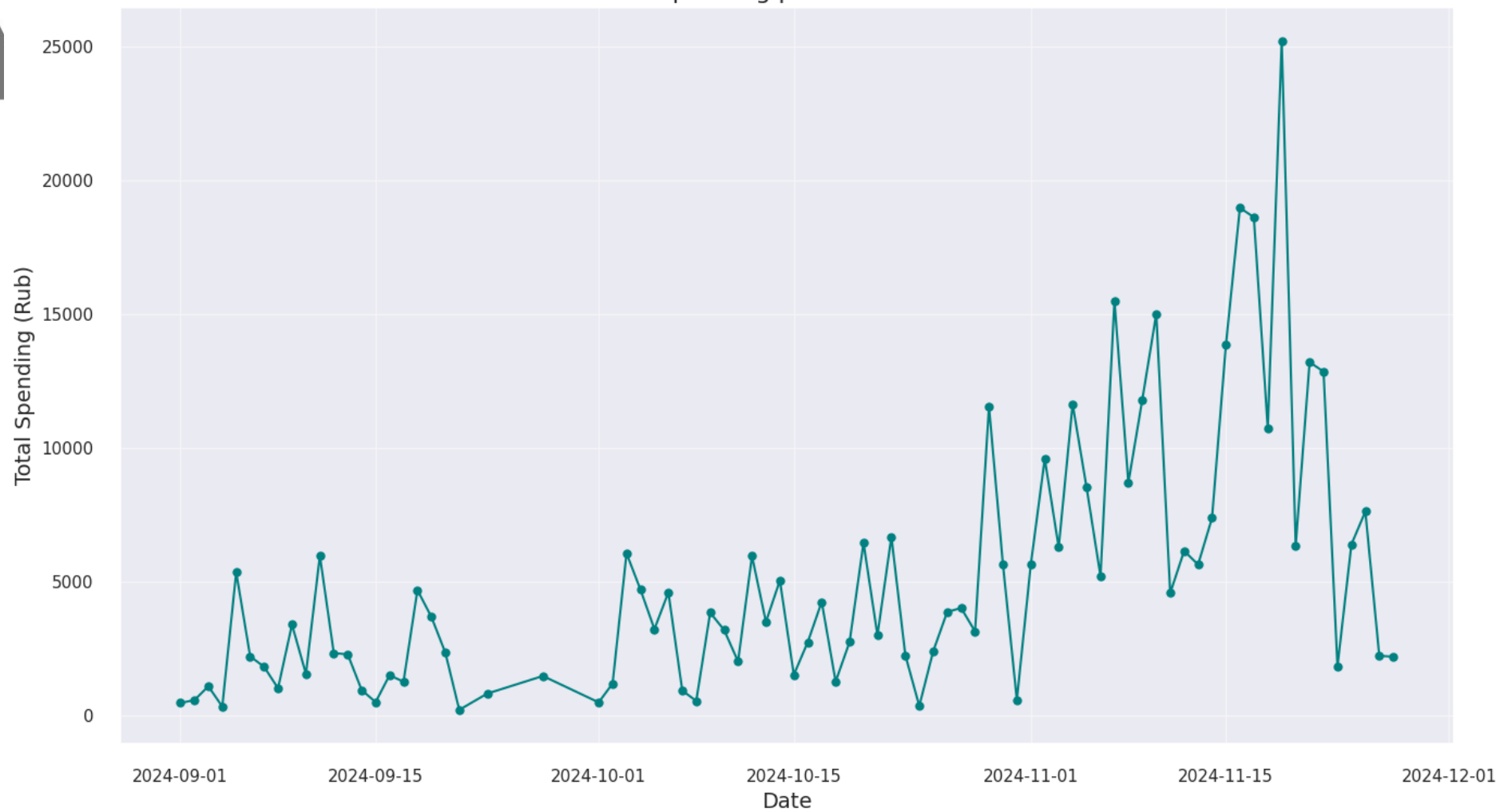
Co-occurrence of Top 20 Products in Transactions



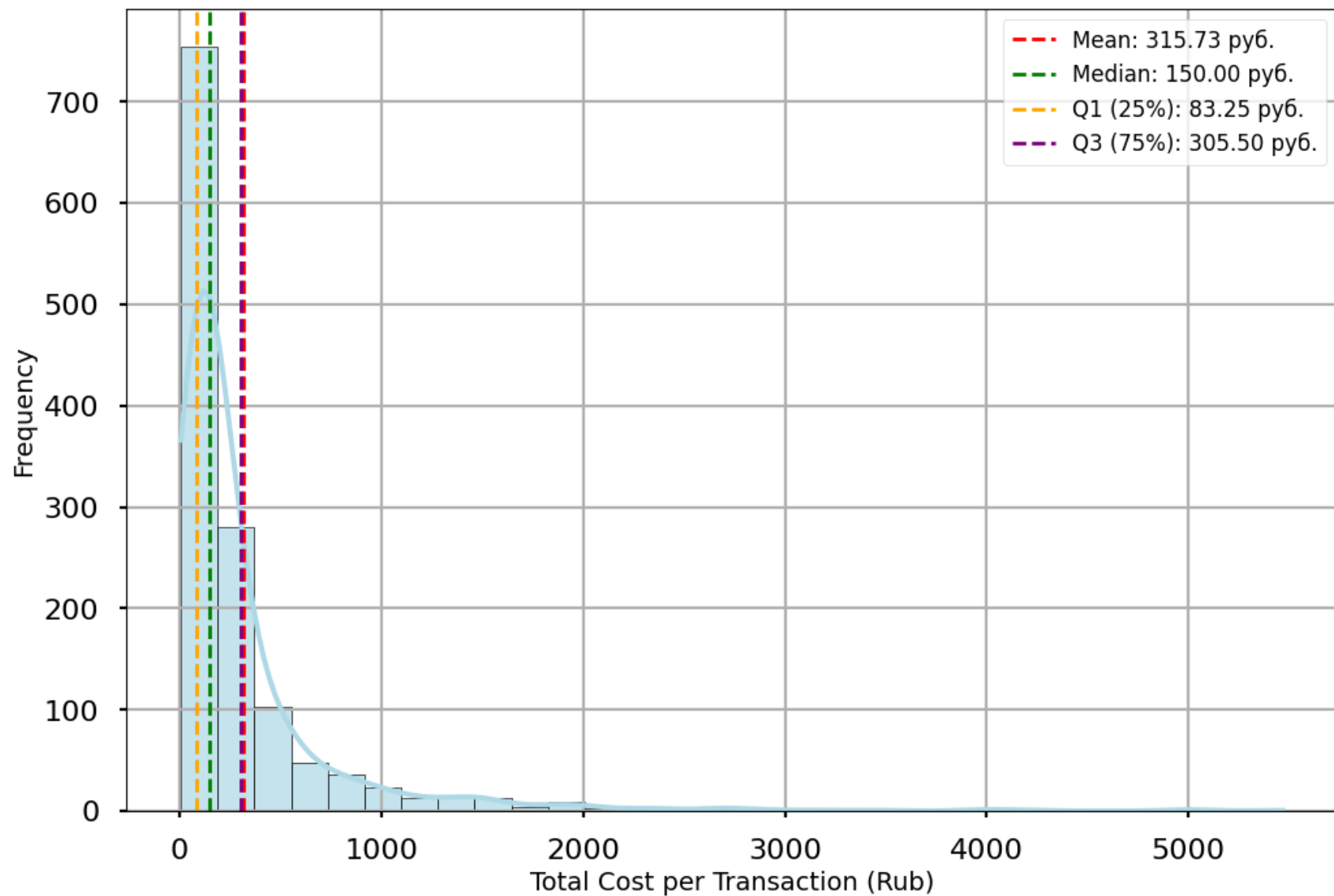




Trends in Total Spending per Transaction Over Time

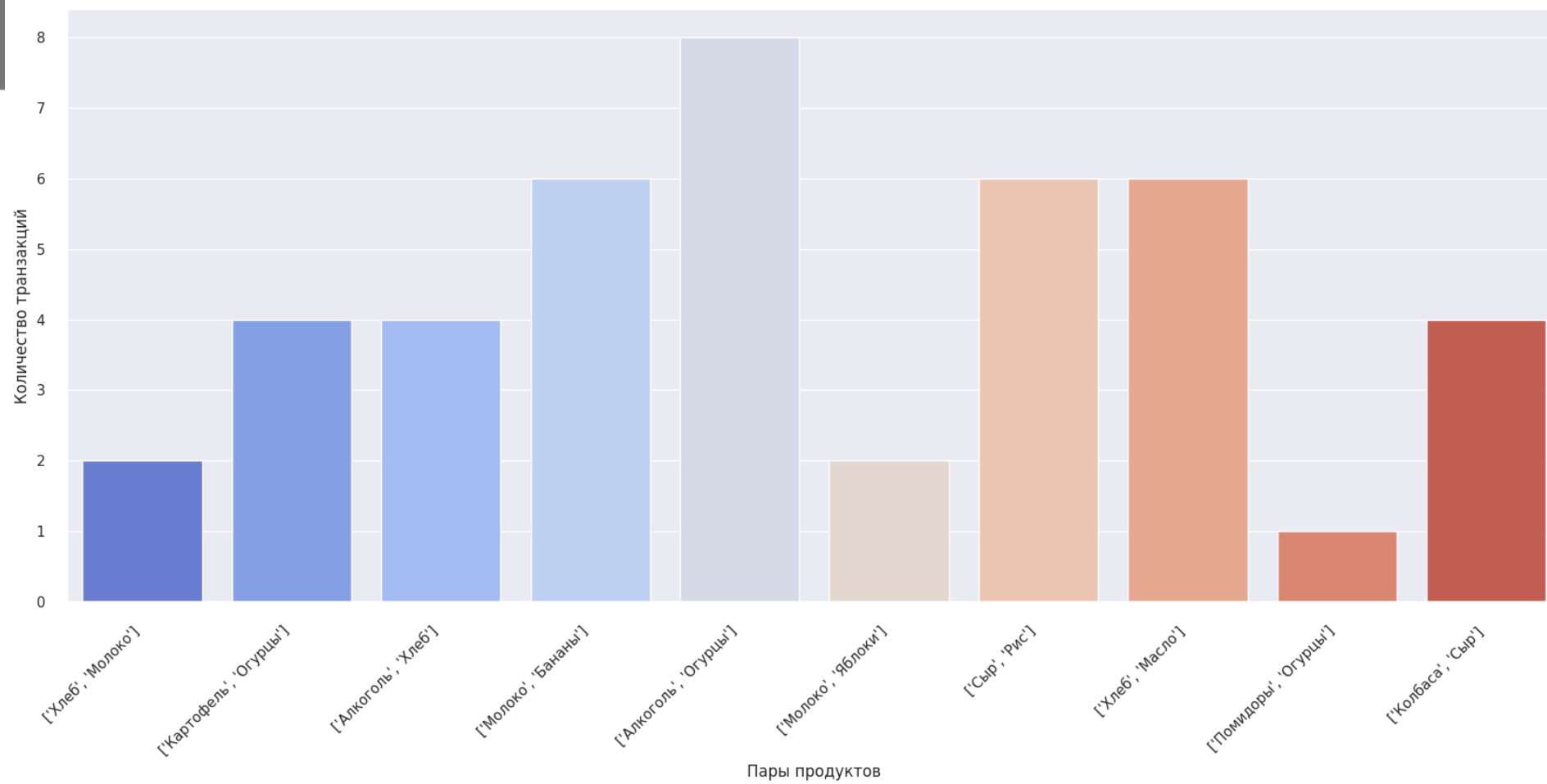


Enhanced Distribution of Total Transaction Costs



	antecedents	consequents	antecedent support	consequent support	support	confidence	lift	representativity	leverage	conviction	zhangs_metric	jaccard	certainty	kulczynski	Количество продуктов (antecedents)	Левая часть (antecedents)	Правая часть (consequents)
1568	(Бананы, Помидоры, Морковь)	(Яблоки)	0.012270	0.070552	0.01227	1.000000	14.173913	1.0	0.011404	inf	0.940994	0.173913	1.000000	0.586957	3	Бананы,Помидоры,Морковь	Яблоки
736	(Картофель, Конфеты)	(Морковь)	0.012270	0.061350	0.01227	1.000000	16.300000	1.0	0.011517	inf	0.950311	0.200000	1.000000	0.600000	2	Картофель,Конфеты	Морковь
1751	(Яйца, Бананы, Молоко, Хлеб)	(Колбаса)	0.012270	0.079755	0.01227	1.000000	12.538462	1.0	0.011291	inf	0.931677	0.153846	1.000000	0.576923	4	Яйца,Бананы,Молоко,Хлеб	Колбаса
1512	(Молоко, Бананы, Помидоры)	(Масло)	0.012270	0.073620	0.01227	1.000000	13.583333	1.0	0.011367	inf	0.937888	0.166667	1.000000	0.583333	3	Молоко,Бананы,Помидоры	Масло
1500	(Лук, Бананы, Яблоки)	(Хлеб)	0.012270	0.156442	0.01227	1.000000	6.392157	1.0	0.010350	inf	0.854037	0.078431	1.000000	0.539216	3	Лук,Бананы,Яблоки	Хлеб
...
601	(Молоко)	(Бананы, Морковь)	0.162577	0.021472	0.01227	0.075472	3.514825	1.0	0.008779	1.058407	0.854396	0.071429	0.055184	0.323450	1	Молоко	Бананы,Морковь
607	(Молоко)	(Бананы, Печенье)	0.162577	0.015337	0.01227	0.075472	4.920755	1.0	0.009776	1.065043	0.951465	0.074074	0.061071	0.437736	1	Молоко	Бананы,Печенье
613	(Молоко)	(Бананы, Помидоры)	0.162577	0.021472	0.01227	0.075472	3.514825	1.0	0.008779	1.058407	0.854396	0.071429	0.055184	0.323450	1	Молоко	Бананы,Помидоры
620	(Молоко)	(Сливки, Бананы)	0.162577	0.018405	0.01227	0.075472	4.100629	1.0	0.009278	1.061725	0.902930	0.072727	0.058137	0.371069	1	Молоко	Сливки,Бананы
1839	(Молоко)	(Яйца, Сыр, Колбаса, Хлеб)	0.162577	0.015337	0.01227	0.075472	4.920755	1.0	0.009776	1.065043	0.951465	0.074074	0.061071	0.437736	1	Молоко	Яйца,Сыр,Колбаса,Хлеб

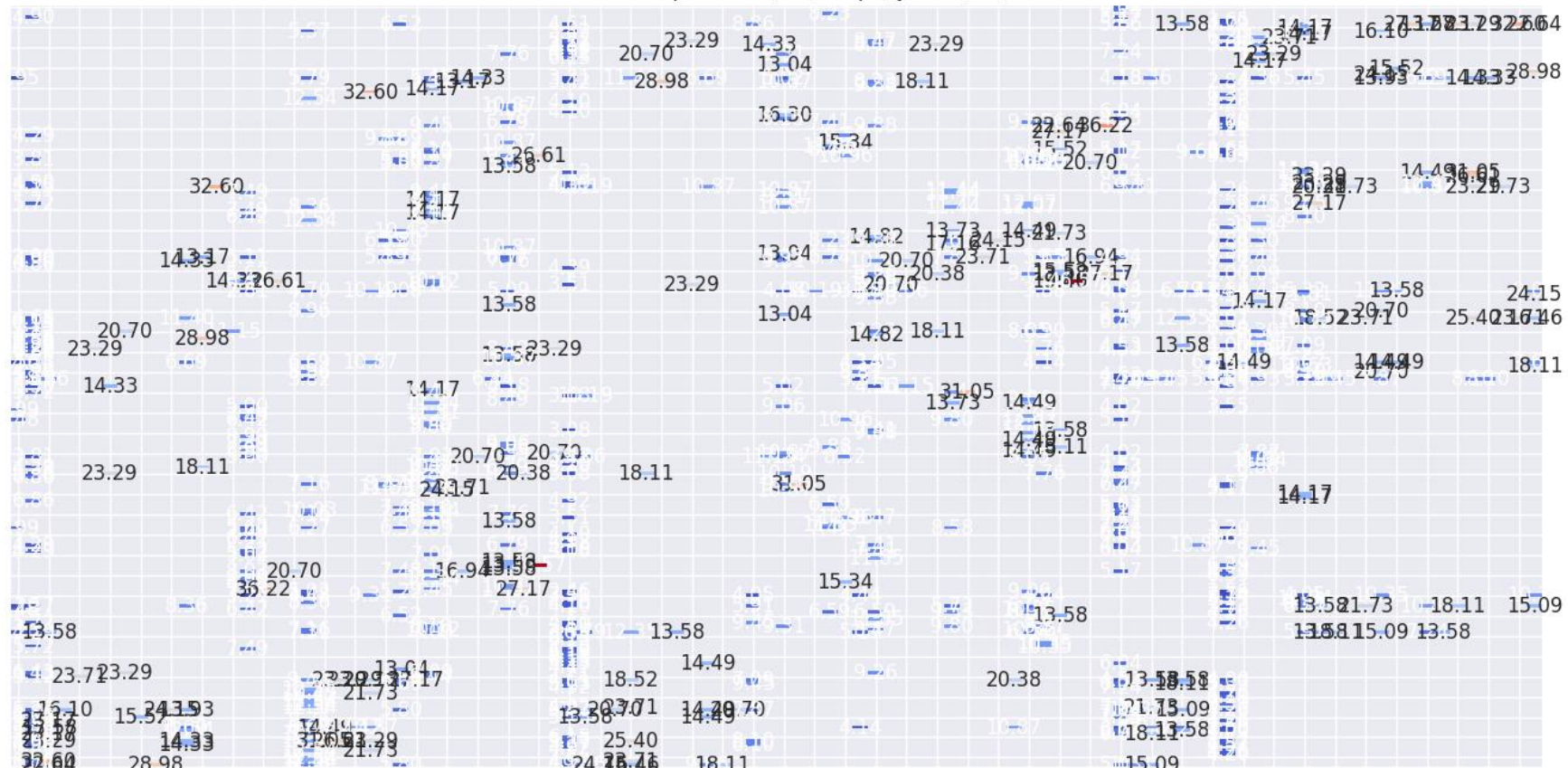
1840 rows x 17 columns



Тепловая карта ассоциаций продуктов (Lift)

Левая часть (antecedents)

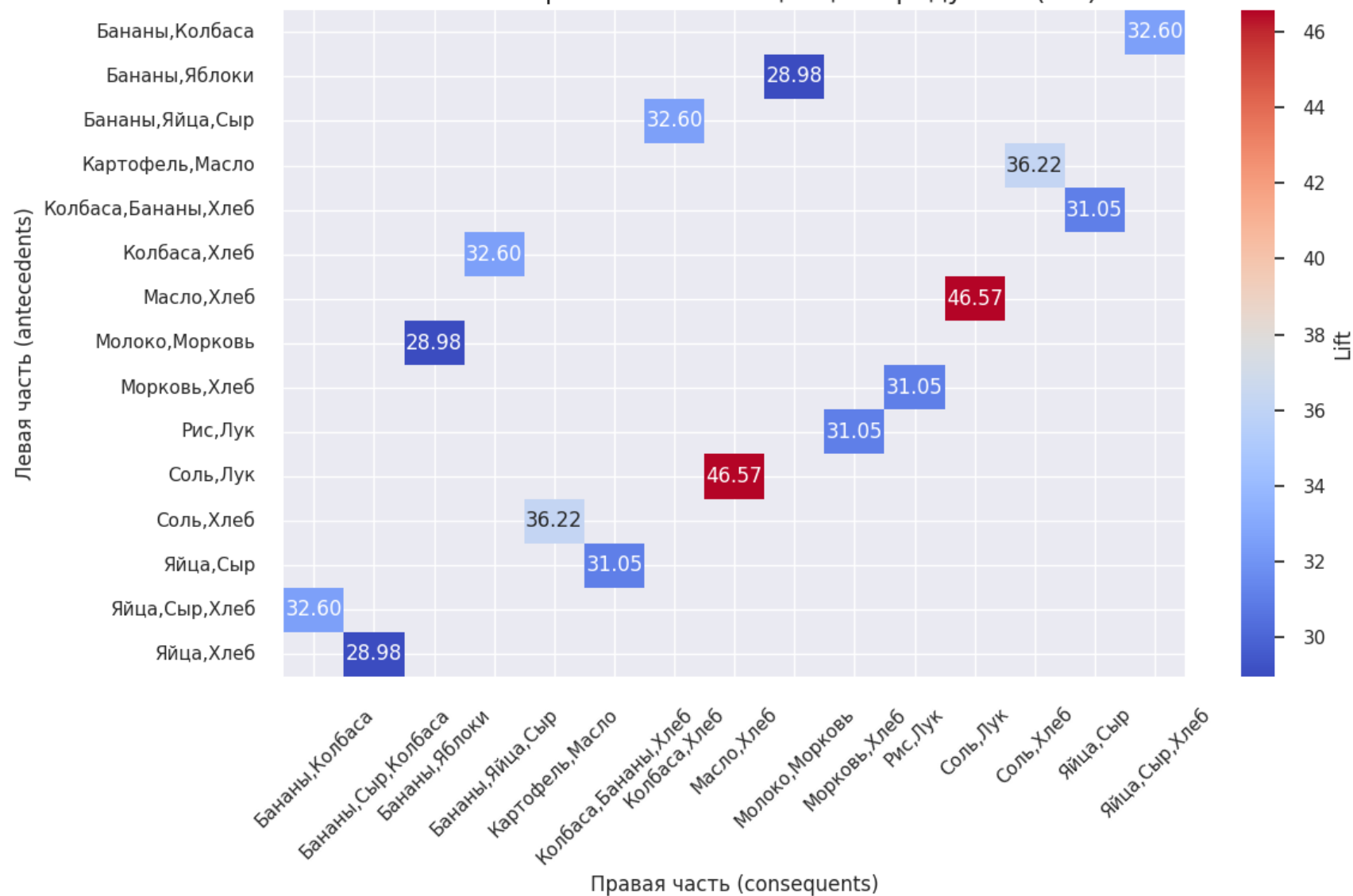
Алкоголь, Колбаса
Бананы, Колбаса, Сыр, Хлеб
Бананы, Морковь, Яблоки
Бананы, Сыр
Бананы, Яблоки, Хлеб
Говядина, Хлеб
Картофель, Масло, Соль
Картофель, Свинина
Колбаса, Бананы, Молоко, Хлеб
Конфеты, Лук
Курица, Морковь
Лук, Колбаса
Лук, Рис, Морковь
Масло, Курица
Молоко, Бананы
Молоко, Бананы, Хлеб
Молоко, Масло
Молоко, Помидоры, Масло
Молоко, Сыр, Хлеб
Морковь, Хлеб
Огурцы, Морковь
Перец, Масло
Помидоры, Лук
Помидоры, Яблоки
Сахар, Бананы
Свинина, Лук
Сливки, Бананы
Сливки, Хлеб
Соль, Масло, Лук
Сыр, Бананы
Сыр, Масло
Сыр, Хлеб, Печенье
Творог, Сыр
Яблоки, Хлеб
Яйца, Бананы, Молоко, Колбаса
Яйца, Колбаса, Хлеб
Яйца, Печенье
Яйца, Сыр, Молоко, Хлеб

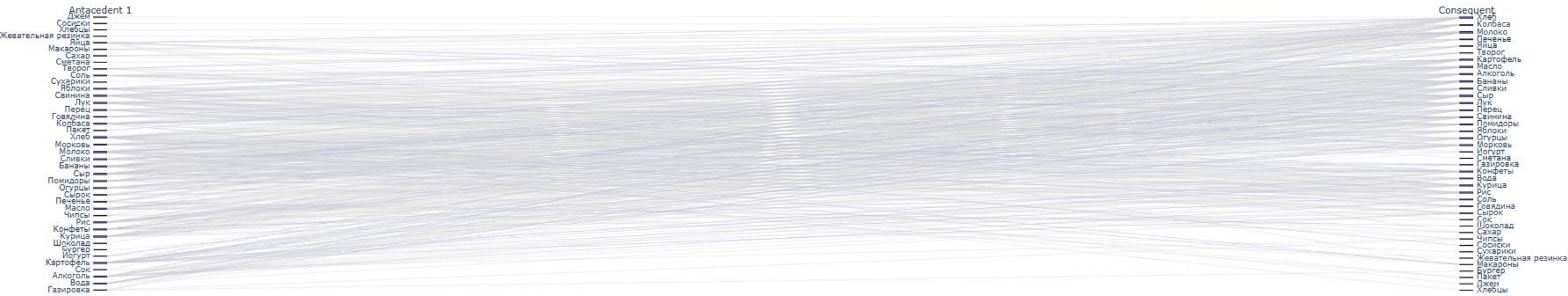


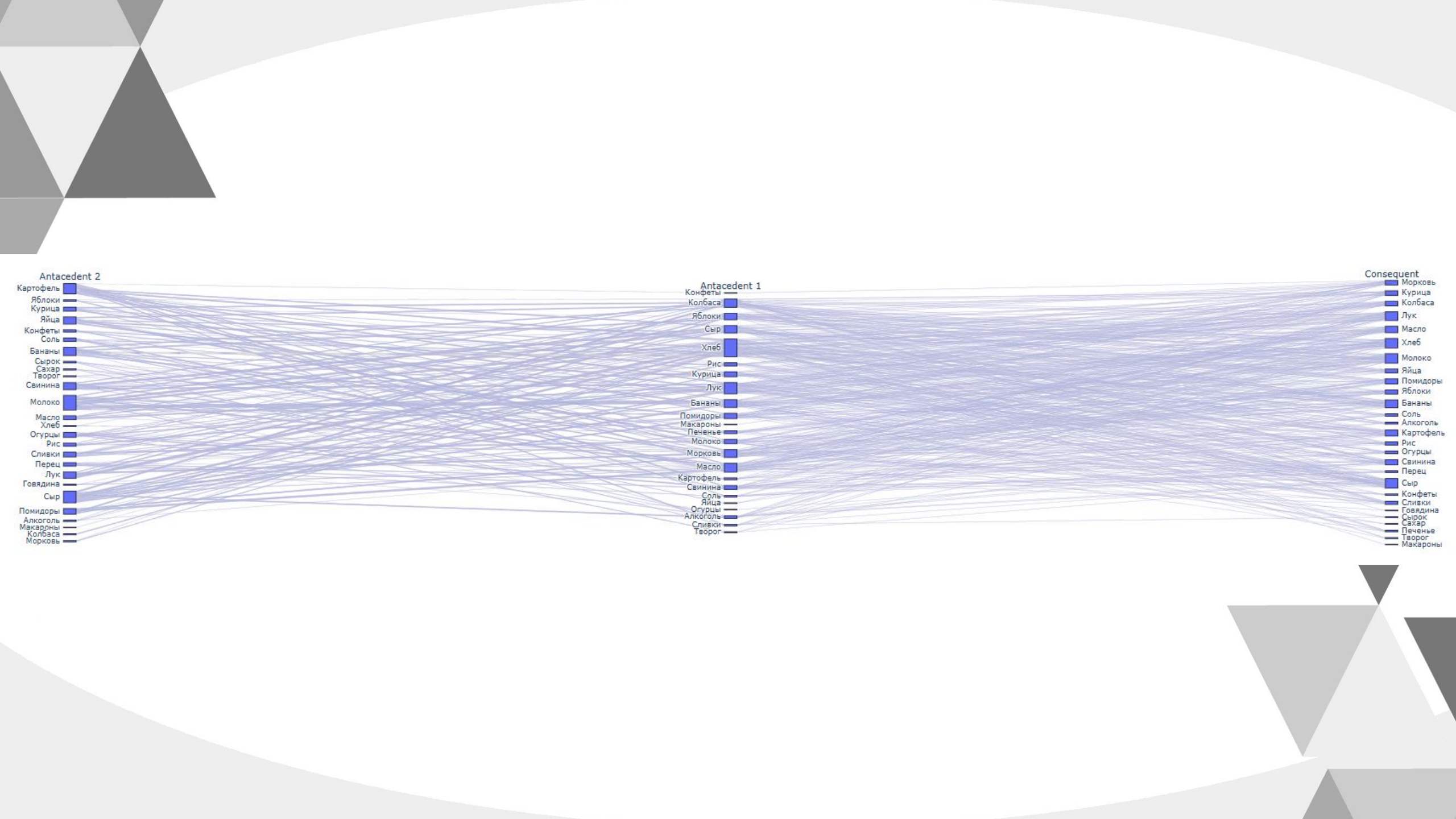
Правая часть (consequents)

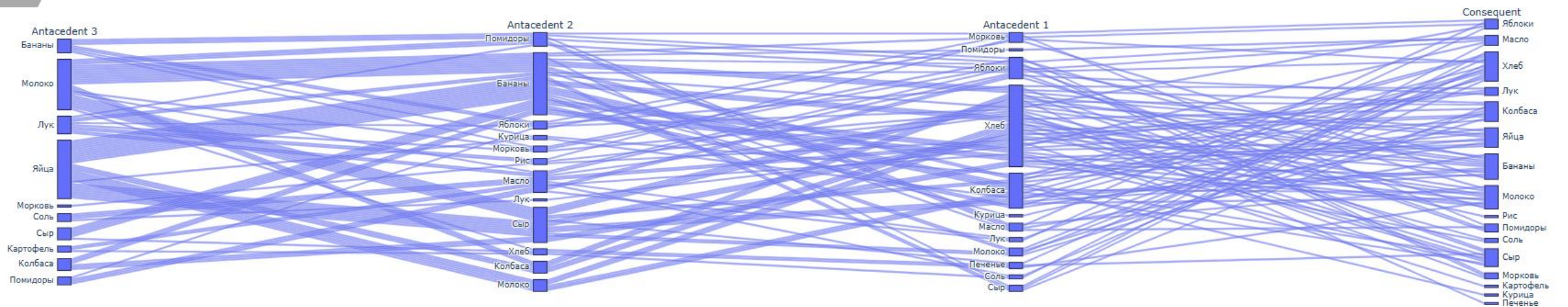
Алкоголь
Бананы, Колбаса
Бананы, Лук
Бананы, Морковь
Бананы, Сыр
Бананы, Сыр, Хлеб
Бананы, Яблоки
Говядина
Картофель, Масло
Картофель, Хлеб
Колбаса, Бананы, Хлеб
Колбаса, Сыр, Хлеб
Конфеты
Курица, Колбаса
Лук, Морковь
Лук, Яблоки
Масло
Масло, Хлеб
Молоко
Молоко, Бананы, Колбаса
Молоко, Колбаса
Молоко, Морковь
Молоко, Помидоры
Молоко, Сыр, Колбаса
Молоко, Хлеб
Морковь
Морковь, Яблоки
Перец
Помидоры
Помидоры, Масло
Помидоры, Яблоки
Рис, Лук
Рис, Хлеб
Свинина
Соль
Соль, Масло
Сыр
Сыр, Колбаса
Сыр, Хлеб
Творог
Хлеб, Печенье
Яблоки, Морковь
Яйца
Яйца, Бананы, Колбаса
Яйца, Бананы, Хлеб
Яйца, Колбаса, Хлеб
Яйца, Молоко
Яйца, Молоко, Хлеб
Яйца, Сыр, Колбаса
Яйца, Сыр, Хлеб

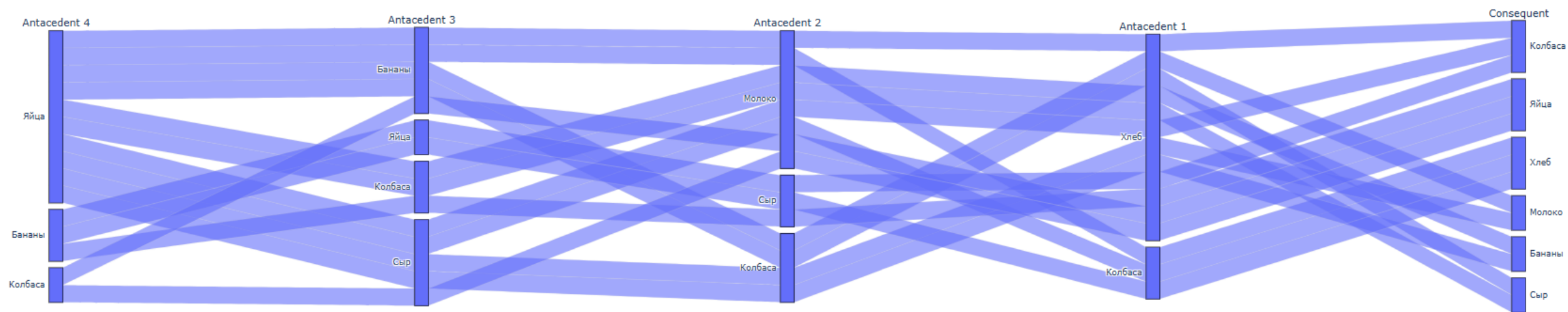
Тепловая карта топ-15 ассоциаций продуктов (Lift)

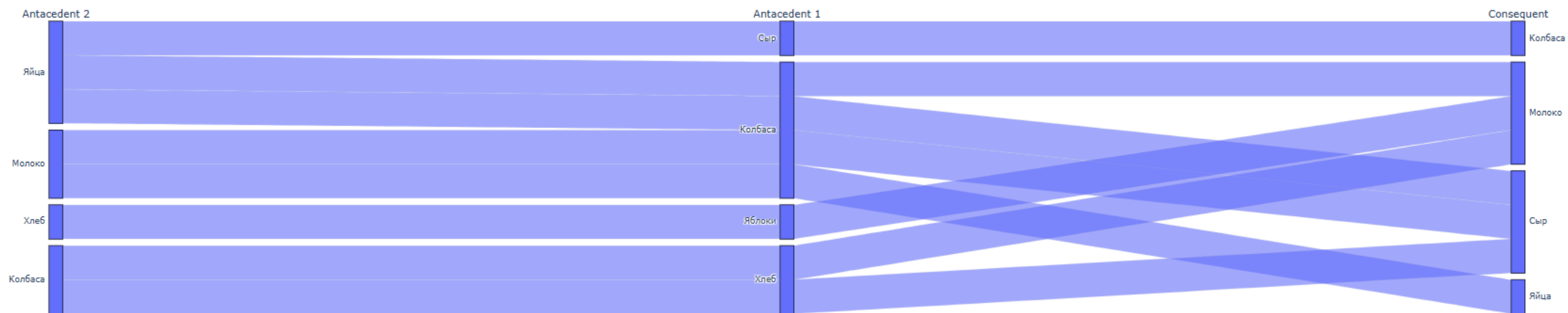




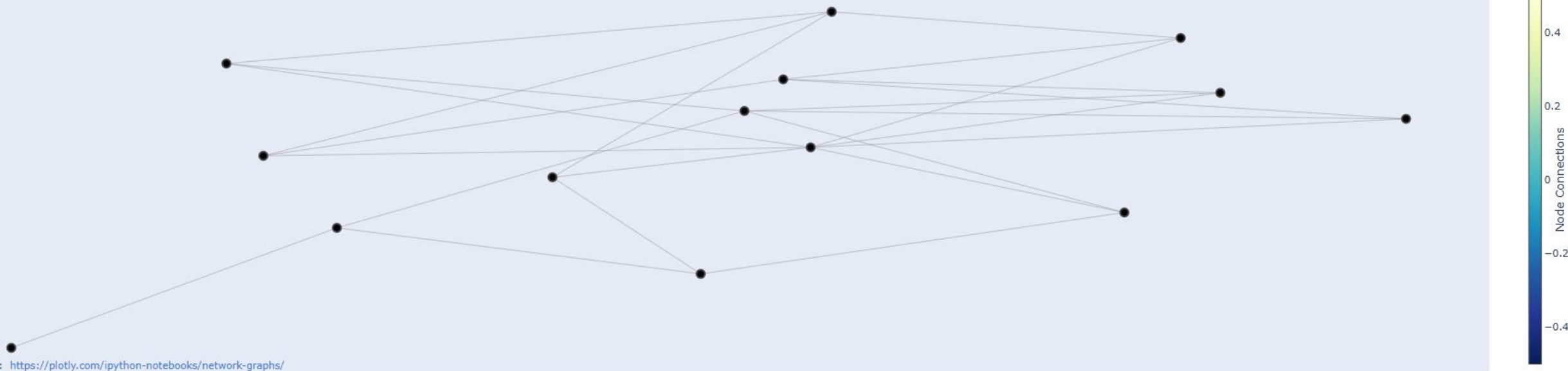








Network graph made with Python



Python code: <https://plotly.com/ipython-notebooks/network-graphs/>

The background features a decorative geometric pattern in the corners, consisting of various shades of gray triangles. The top-left and bottom-right corners have clusters of triangles, while the top-right and bottom-left corners have curved, light gray shapes.

Спасибо за внимание