

```
In [10]: %run тестовое_задание_WiseTech_для_скрипта.py
```

# Определение характерных параметров для анализа торговли финансовыми инструментами

- 1. Общее описание задачи:
  - быстро и грубо оценить влияние каждого параметра на результаты
  - выбрать 5-6 строчек с наиболее характерными параметрами и целесообразными результатами

Цель:  
отобрать строки с наиболее характерными параметрами и целесообразными результатами для более детального дальнейшего анализа

- 2. Импортируем библиотеки
- 3. Создаем функции
- 4. Загружаем данные
- 5. Оцениваем влияние параметров на результаты

None  
Матрица показывает, что присутствует как прямая, так и обратная зависимость результатов от параметров. При прямой зависимости увеличение значения параметра приводит к росту значения результата. При обратной зависимости увеличение значения параметра приводит к снижению значения результата. Отсутствует зависимость результата от параметра7 - причина в том, что этот параметр имеет одно единственное значение в предоставленном датафрейме.

Также мы можем увидеть различную степень зависимости результата от значений параметров. Определим рамки степени зависимости. Примем зависимость до 50% включительно как слабую, а зависимость более 50% как сильную.

## 5.1 Выведем на экран оценку влияния каждого параметра на результаты

	result	параметр1	параметр2	параметр3	параметр4	параметр5	параметр6	параметр7	параметр8
0	profit	обратная сильная	прямая сильная	обратная слабая	обратная слабая	обратная слабая	прямая сильная	отсутствует	обратная слабая

	result	параметр1	параметр2	параметр3	параметр4	параметр5	параметр6	параметр7	параметр8
1	trades	обратная сильная	прямая сильная	обратная слабая	обратная слабая	обратная сильная	прямая слабая	отсутствует	обратная слабая
2	profitabletrades	прямая сильная	обратная сильная	прямая слабая	прямая слабая	прямая слабая	обратная сильная	отсутствует	прямая слабая
3	profitability	прямая слабая	обратная слабая	прямая слабая	обратная слабая	обратная слабая	обратная слабая	отсутствует	обратная слабая
4	matozhidanie	прямая сильная	обратная сильная	прямая слабая	прямая слабая	прямая сильная	обратная слабая	отсутствует	прямая слабая
5	riskfactor	прямая слабая	обратная сильная	прямая слабая	прямая слабая	прямая слабая	обратная слабая	отсутствует	прямая слабая

## 5.2 Выгрузим таблицу в файл Excel. Назовем его $\in \text{fluence}_{on\_rest}$

Выгружен файл [influence\\_on\\_result.csv](#)

Выгружен файл с влиянием каждого параметра на результаты [influence\\_on\\_result.xlsx](#)

## 5.3 Визуализируем зависимость результатов от значени параметров

## 5.4 Исследуем аномальные значения

- колонка *profitability* имеет три аномальных значения, причем все они расположены далеко за пределами нормальных значений
- параметр3 и параметр6 имеют по три аномальных значения

## 6. Записи с наиболее характерными параметрами и целесообразными результатами

- Колонка Profit не покажет нам объективные данные, если мы не владеем суммой сделок в этих операциях. Большой показатель абсолютной прибыли может быть достигнут за счет значительной суммы денежных средств, привлеченных в сделку. Но при этом возможно, что с точки зрения отношения прибыли к сумме вложения эта сделка не будет выглядеть экономически эффективной
- Также нам не принесет новых знаний данные о количестве сделок в колонке *trades* - много их или мало - не дает нам понимания их влияния на общий результат.

А анализ относительных величин позволяет корректно оценивать сделки. Результаты прибыли и количества сделок учтены в этих относительных результатах.

- Поэтому для анализа и выявления записей с наиболее характерными параметрами и целесообразными результатами мы воспользуемся колонками не с абсолютными, а с относительными результатами. Оставим следующие колонки:

- 'ProfitableTrades'
- 'Profitability'
- 'MatOzhidanie'
- 'RiskFactor'

6.1 Отсортируем 6 записей, в которых в каждой из указанных выше колонок результаты показывают одни из своих самых высоких значений при том, что в трех других колонках результаты тоже являются одними из самых высоких.

Важно различать, что одни из самых высоких - это не тоже самое, что максимальные результаты для каждой колонки.

Технически мы произведем последовательный перебор и отсеивание невысоких значений в отобранных четырех колонках, пока не останется нужное нам количество записей - его мы приняли равным 6

6.2 Выведем на экран отсортированные данные

	index	profit	trades	profitabletrades	profitability	matozhidanie	riskfactor	параметр1	параметр2	параметр3	параметр4	параметр5	параметр
0	30	43.44325	14	0.750000	6.016160	3.103095	3.720440	1.4	3.0	-0.8	0.6	2.6	4
1	34	40.61175	12	0.791667	10.716085	3.384315	3.709320	2.2	3.5	-0.4	0.2	2.6	3
2	35	39.77855	11	0.818182	7.508665	3.616230	3.795110	2.2	3.0	-0.6	0.3	2.6	4
3	36	39.67800	13	0.846154	11.134815	3.052150	3.695445	2.2	3.0	-0.4	0.1	2.2	3
4	37	38.37620	12	0.912879	12.216150	3.198015	3.940080	1.8	4.0	-0.2	0.1	1.4	1
5	40	32.51535	10	0.900000	54.910705	3.251535	4.259445	1.8	3.0	-0.2	0.1	1.0	1

6.3 Выгрузим таблицу в файл Excel. Назовем его *characteristic.xlsx*

Выгружен файл **characteristic.csv**

Выгружен файл с наиболее характерными параметрами и целесообразными результатами **characteristic.xlsx**

7. Наблюдения относительно данных

- Прибыль показывает значения в диапазоне до 50 единиц и в диапазоне от 65 до 86 единиц. Диапазон от 50 до 65 единиц заполнен единичными сделками
- Количество сделок имеет разрыв в диапазоне от 44 до 67 сделок
- Profitability имеет два аномальных значения в размере 55 и 197 единиц, но эти позиции имеют минимальные значения прибыли
- Параметр 2 не принимал значения, равные 4.5 и 5.0

- Параметр 3 обычно принимал значения, равные -0.8
- Параметр 6 не принимал значения в диапазоне от 16 до 34
- Параметр 7 принимал единственное значение, значения, равное 28.1 - Оценили зависимость влияния каждого параметра на результаты

## 8. Выводы и заключения

### 8.1 - Оценили зависимость влияния каждого параметра на результаты

- Визуализировали матрицу корреляций
- Визуализировали графики зависимостей результата от параметров
- Сохранили в виде файла Excel -  $\in$  *fluence<sub>on</sub>rest.xlsx*

### 8.2 - Исследовали аномальные значения. Они частично пересекаются с записями в итоговом файле *characteristic*.

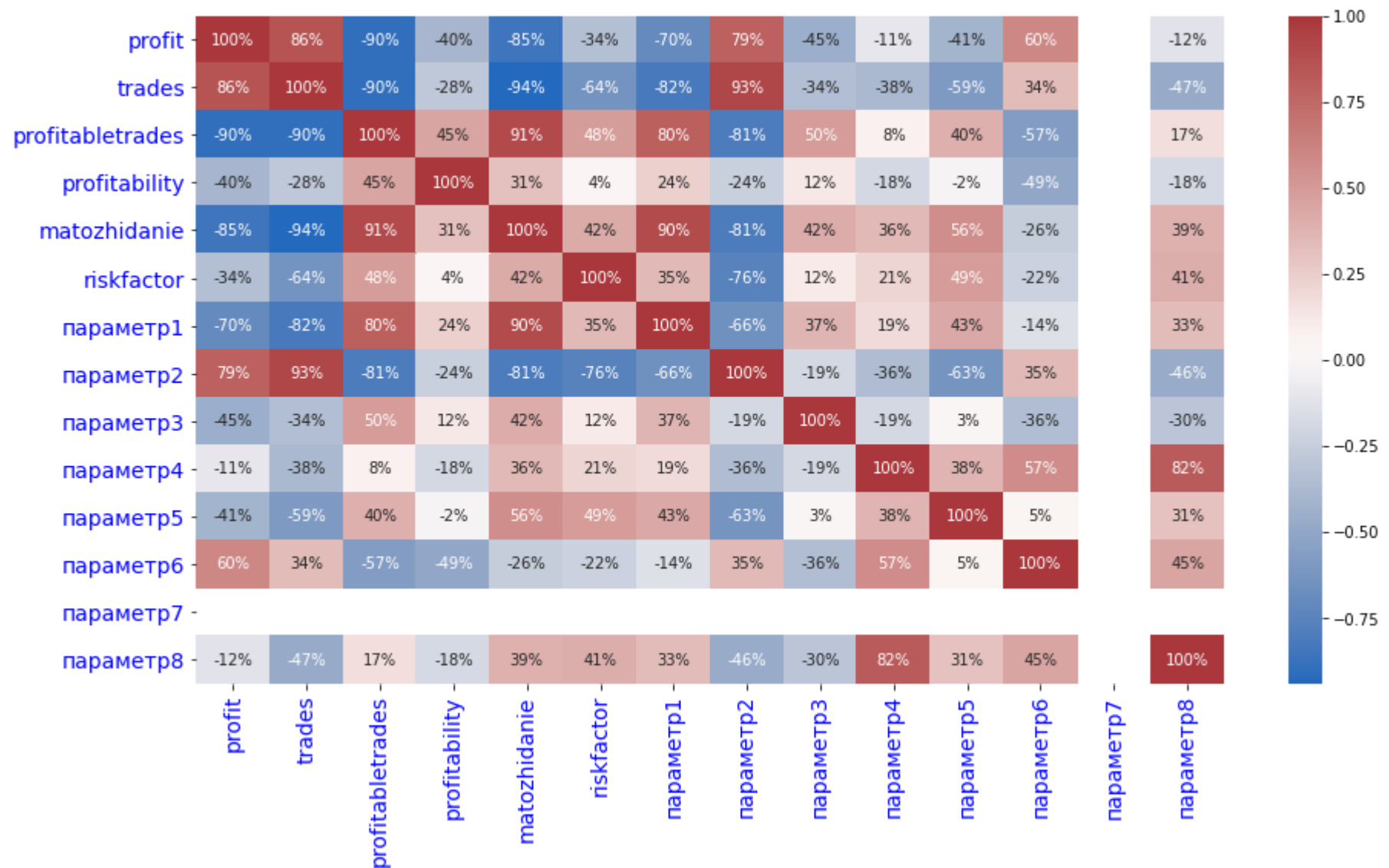
Это дает возможность отдельно обратить внимание:

- на влияние на результат параметра 3
- на результаты в записи с номером 40

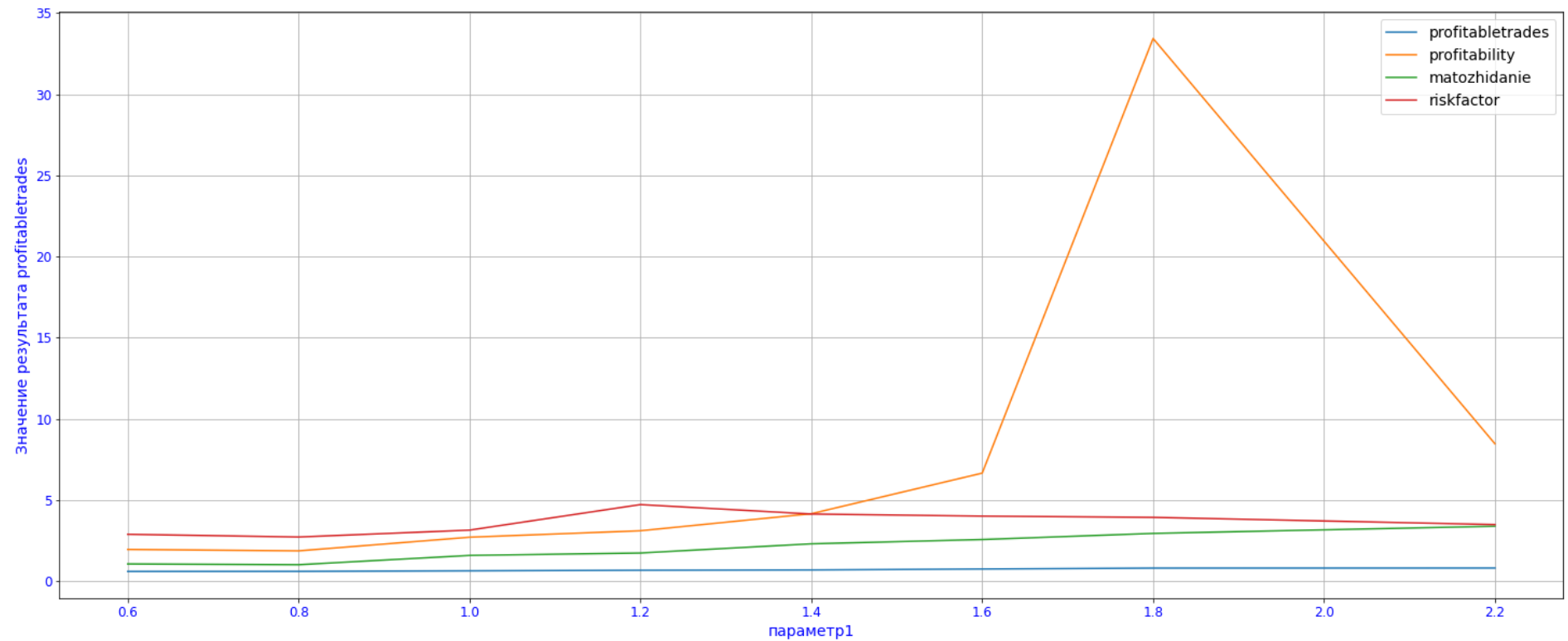
### 8.3 - Для анализа и выявления записей с наиболее характерными параметрами и целесообразными результатами использовали значения в колонках с относительными результатами:

- 'ProfitableTrades'
- 'Profitability'
- 'MatOzhidanie'
- 'RiskFactor'
- Отсортировали 6 записей, в которых в каждой из указанных выше колонок результаты показывают одни из своих самых высоких значений при том, что в трех других колонках результаты тоже являются одними из самых высоких.
- Сохранили в виде файла Excel - *characteristic.xlsx*

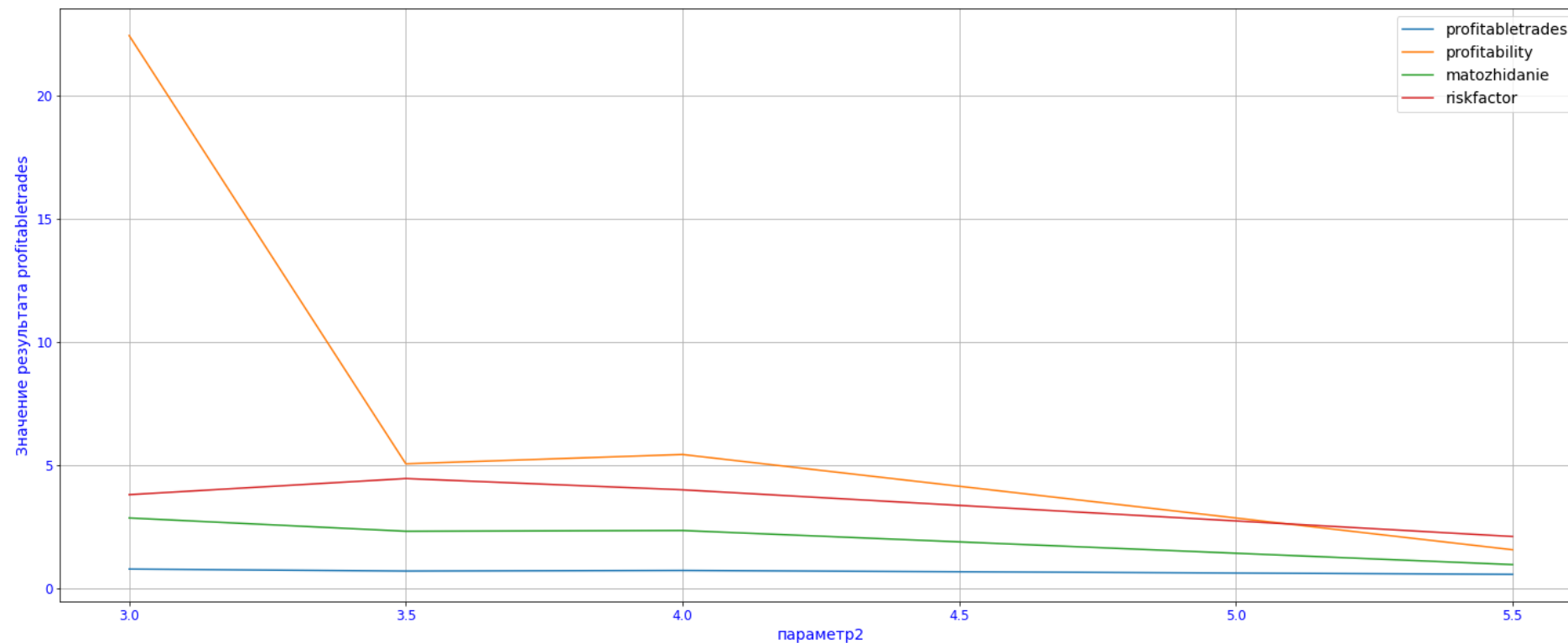
# Матрица корреляции



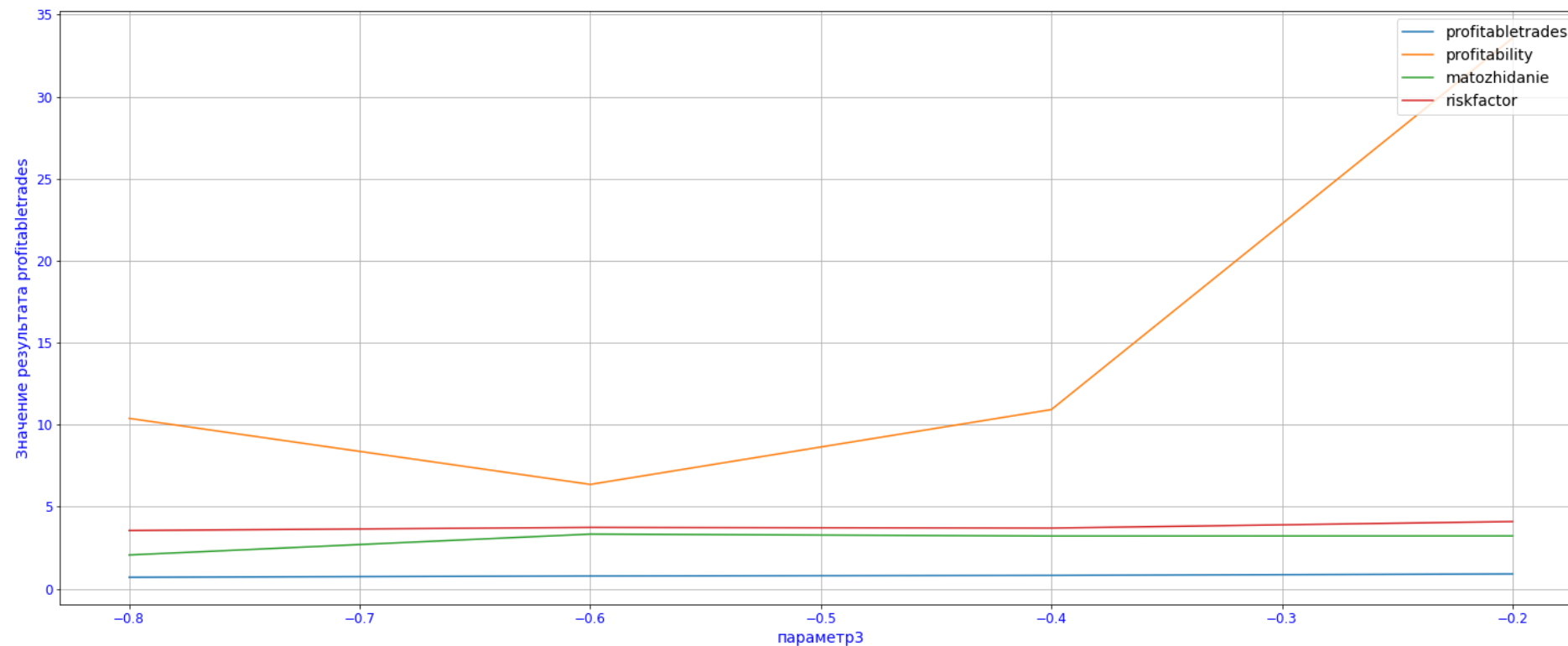
## Значение результата profitabletrades в зависимости от параметр1



## Значение результата profitabletrades в зависимости от параметр2

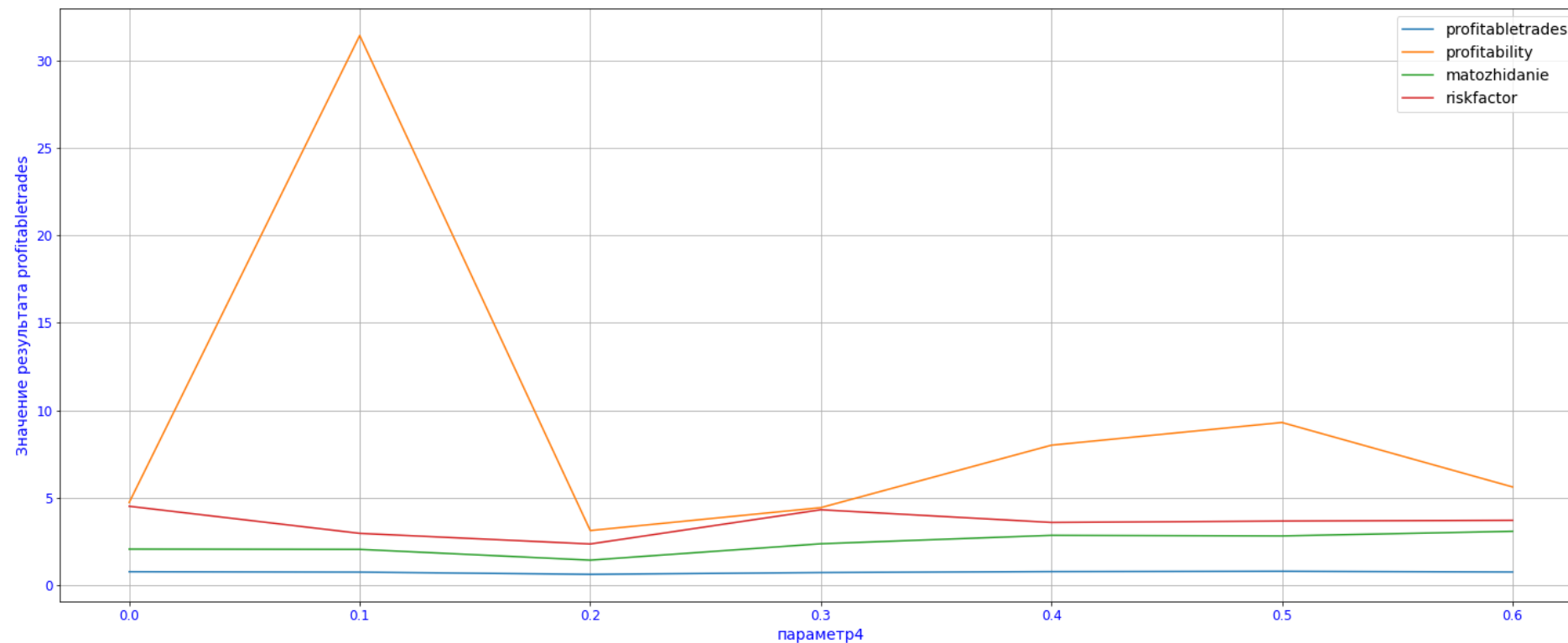


## Значение результата profitabletrades в зависимости от параметр3

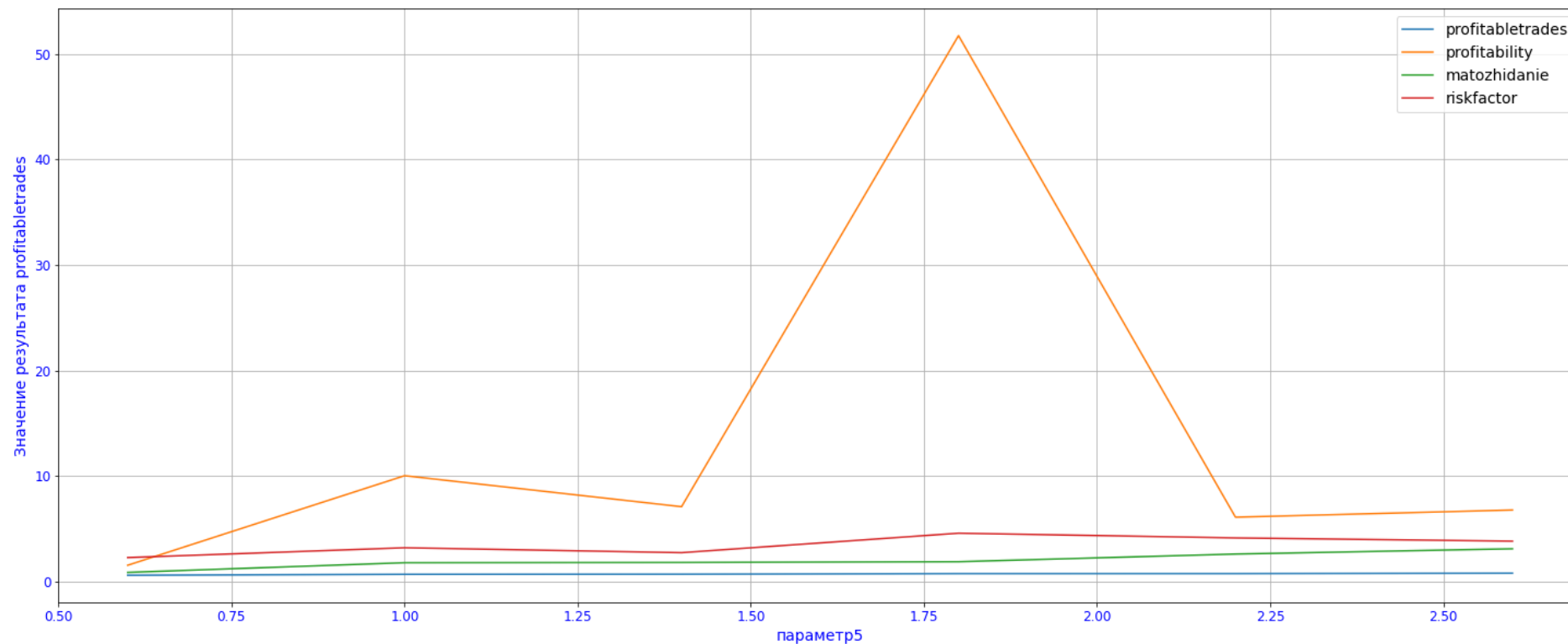




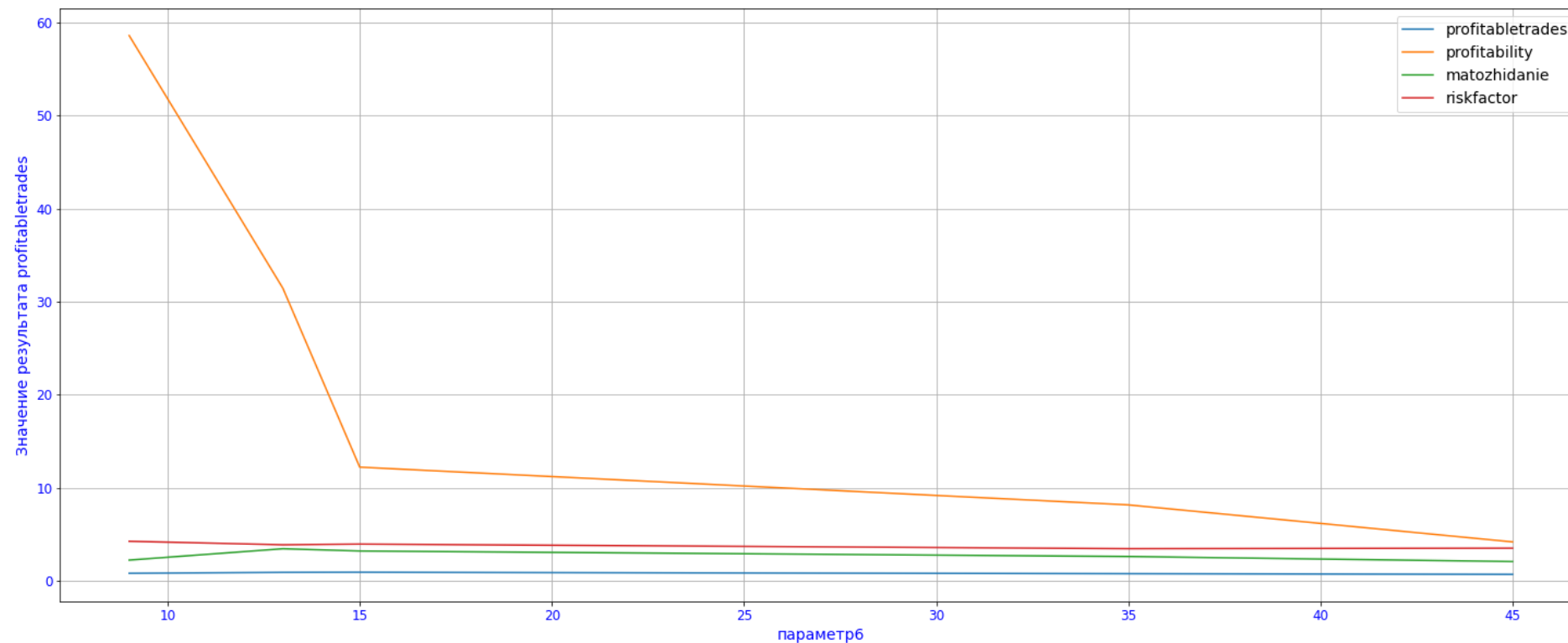
## Значение результата profitabletrades в зависимости от параметр4



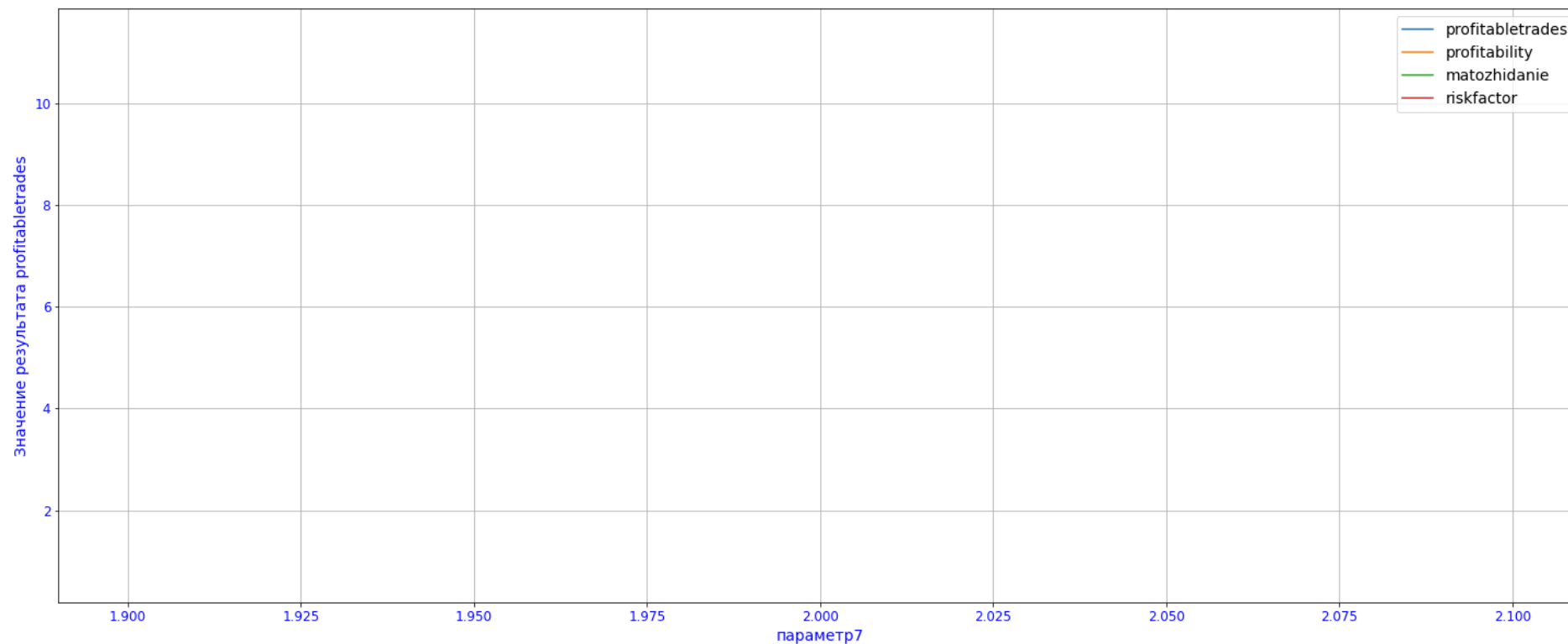
## Значение результата profitabletrades в зависимости от параметр5



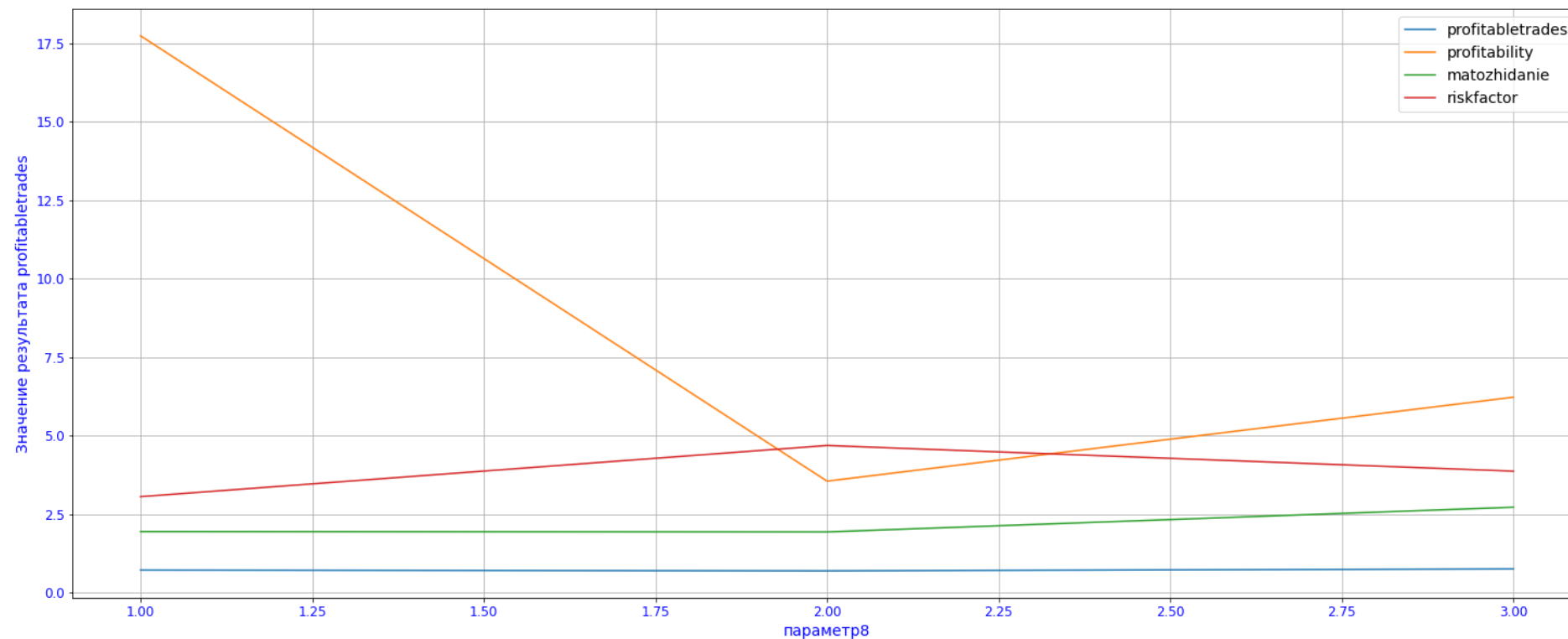
## Значение результата profitabletrades в зависимости от параметр6



## Значение результата profitabletrades в зависимости от параметр7



## Значение результата profitabletrades в зависимости от параметр8



In [ ]: