

### ***С какими инструментами обработки данных вы знакомы?***

Основной инструмент - python.

А именно:

- pandas для анализа и подготовки данных;
- matplotlib, seaborn для построения графиков.

Для работы с БД предпочитаю использовать, либо нативные консольные тулзы, либо встроенный инструмент в PyCharm'e (Database tool).

### ***Есть ли у вас опыт разработки? Какой?***

Разработка backend-а приложений на python ~ 2 года.

### ***Что вы понимаете под понятием Data Mining?***

Это анализ существующих данных для выявления закономерностей и поиска параметров для обучения модели.

Именно тех параметров которые решат проблему бизнеса.

### ***Какие инструменты и методы помогают вам экономить время и избегать ошибок при работе?***

Хорошо настроенная IDE, имеющая автодополнение, различные статанализаторы и другие упрощающие жизнь штуки.

Также крайне желательно ревьюить себя, будь то написанный код или проработанная архитектура/гипотеза.

В процессе ревью, часто выявляются ранее пропущенные моменты, которые необходимо учесть/проработать.

Ещё важно правильно декомпозировать задачу. Атомарные задачи решаются проще и быстрее.

### ***На каких ресурсах вы получаете информацию и новости по используемым технологиям?***

С python'ом всё достаточно просто, комьюнити давно прорешало большинство возможных проблем/сложностей, и это всё легко гуглится, или ищется в официальной документации.

С Data Science сложнее, так как практическая часть начала активно развиваться относительно не так давно, и многие вещи сложно просто взять и нагуглить.

И очень радует, что есть [ODS](#), где можно спросить свой вопрос и получить как минимум ссылку на полезную статью.

Так же полезно читать, то что спрашивают все остальные. Не знаешь, что может из этого пригодится в будущем.

Помимо этого читаю статьи на [Towards Data Science](#) и слежу за новостями в различных телеграм-каналах (например [Machinelearning](#), [Big Data Science](#)).

### ***Расскажите нам о себе и о ваших последних работах.***

Попал в IT-сферу ~5 лет назад.

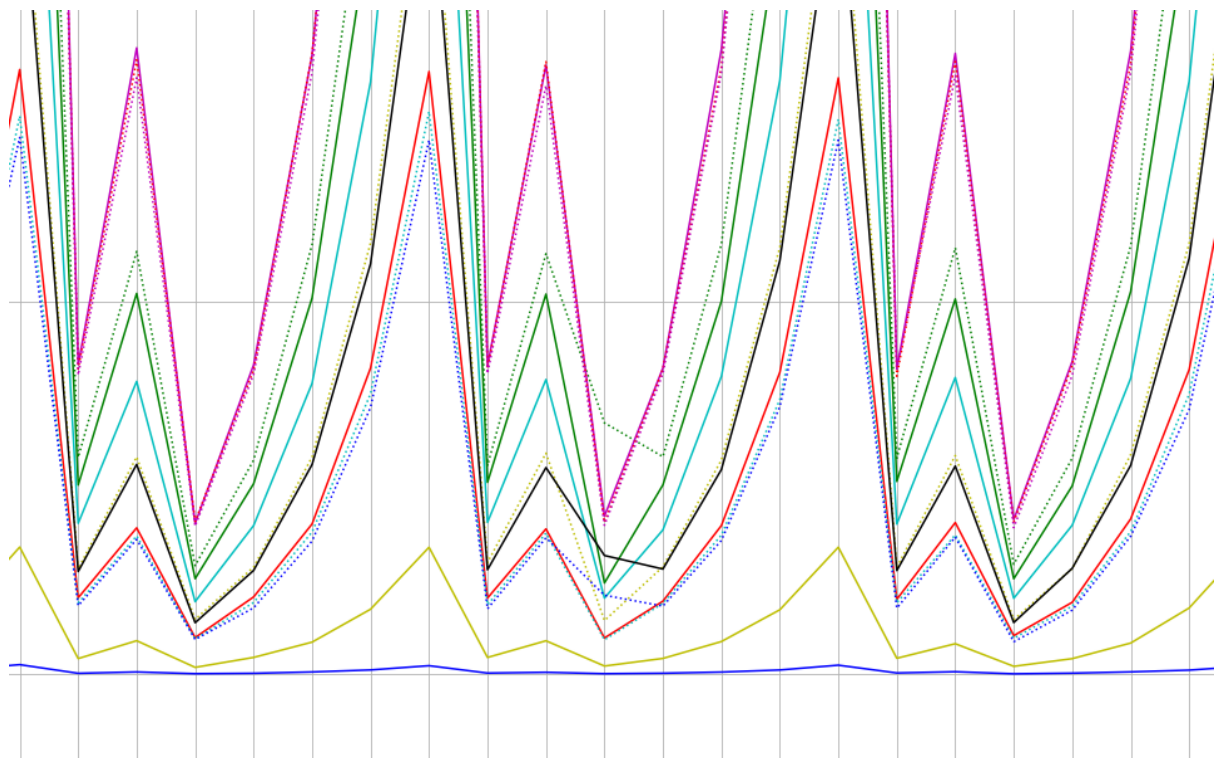
Изначально занимался разработкой автотестов на Python.

Далее перешел в backend-разработку, где неплохо освоил python и работу различными инструментами упрощающими жизнь (Docker, Kubernetes, CI, ...).

Опыта продуктовой разработки в Data Science нет, но для себя чётко решил двигаться в этом направлении, что и делаю )

## Задачи

1. Видимо выпускник ожидал столкнуться с нормальным распределением. Ожидая вероятность встретить женщину как  $1/2$ , при этом уточнив, что среди первых 30 человек будет женщина он ожидал вероятность неблагоприятного исхода как  $1/30$ . Вероятно не было учтено географическое положение бара (военная часть рядом).
2. График получился огромный, поэтому прикладываю отдельным файлом.
3. Проанализировал транзакции и они цикличны в недельный период. Но есть аномалия за 2021-03-10.



Как видно у троих клиентов транзакций больше чем ожидалось.  
Предположительно сбой (атака) был 10 марта в период с 0:00 по 21:00 у клиентов с идентификаторами: 7, 16, 27.

