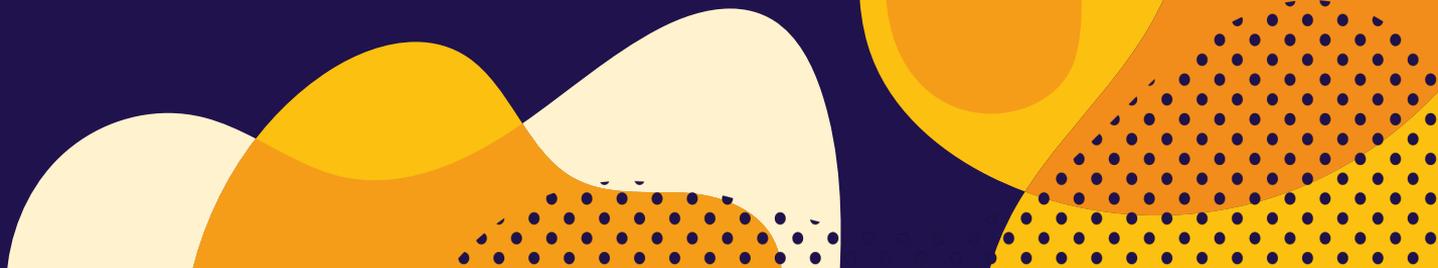


ХАКАТОНОВЫЕ МОНСТРЫ - 16 команда



Команда



Игнатенко
Никита



Абрамов
Артем



Керский
Лев



Яковлев
Тимофей



Удалов
Макар



Никитин
Владимир

Исследовательский вопрос

Какие показатели влияют на продуктивность сотрудника?



Мотивация

Изучить данные представленные в датасете о продуктивности работников, а после дать рекомендации, как можно улучшить данные показатели.



Структура данных

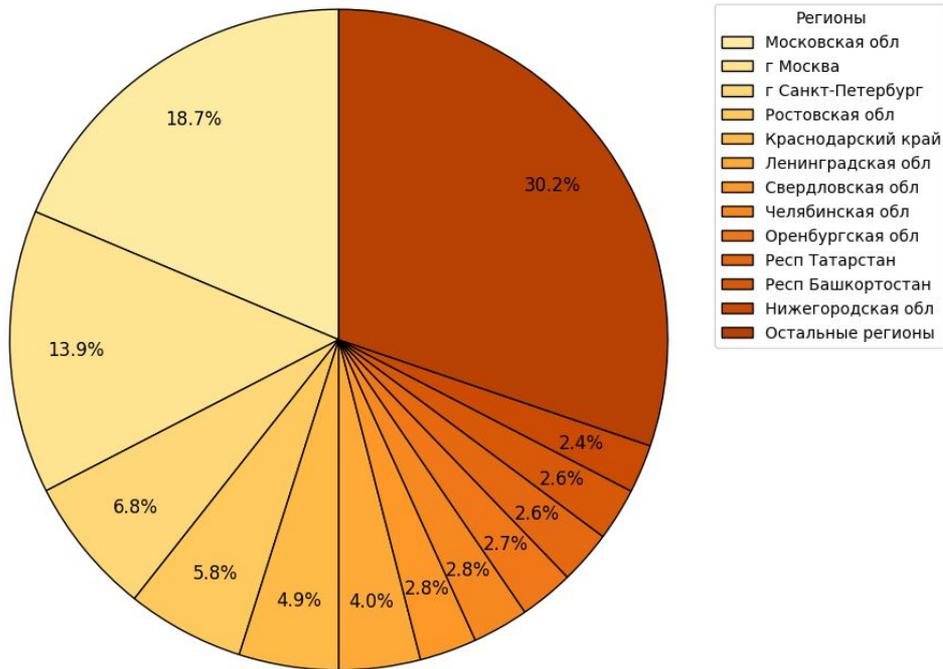
Информация

Данные о работе
сотрудников колл-центра

>170 тыс. строк

87 регионов

С 2024-09-01 до
2025-01-31



Подготовка данных

Формула продуктивности

Продуктивность P можно выразить следующим образом:

$$P = \frac{(C + U)}{(T_C + T_U)}$$

где:

- C — количество обработанных звонков,
- U — количество обработанных пользователей в чатах,
- T_C — общее время, затраченное на все звонки,
- T_U — общее время, затраченное на все чаты.

Подготовка данных

Ввели новую переменную
“продуктивность”

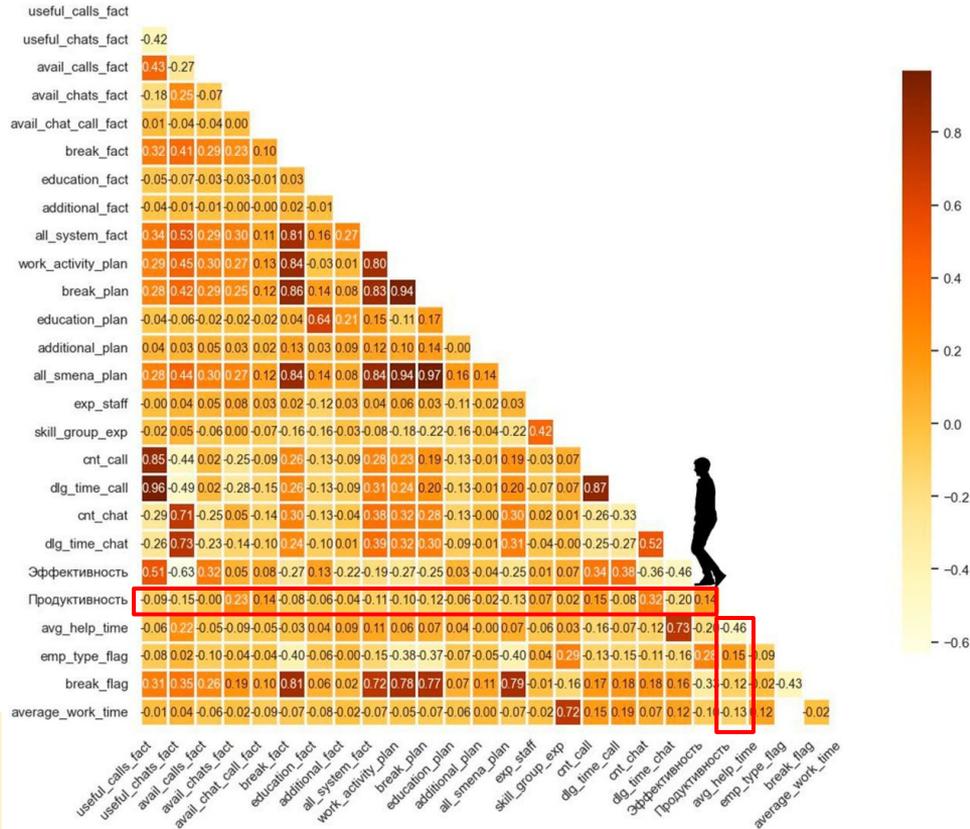
Очистка от выбросов

Перевод текстовых значений
к эквивалентным числовым

Удаление NaN

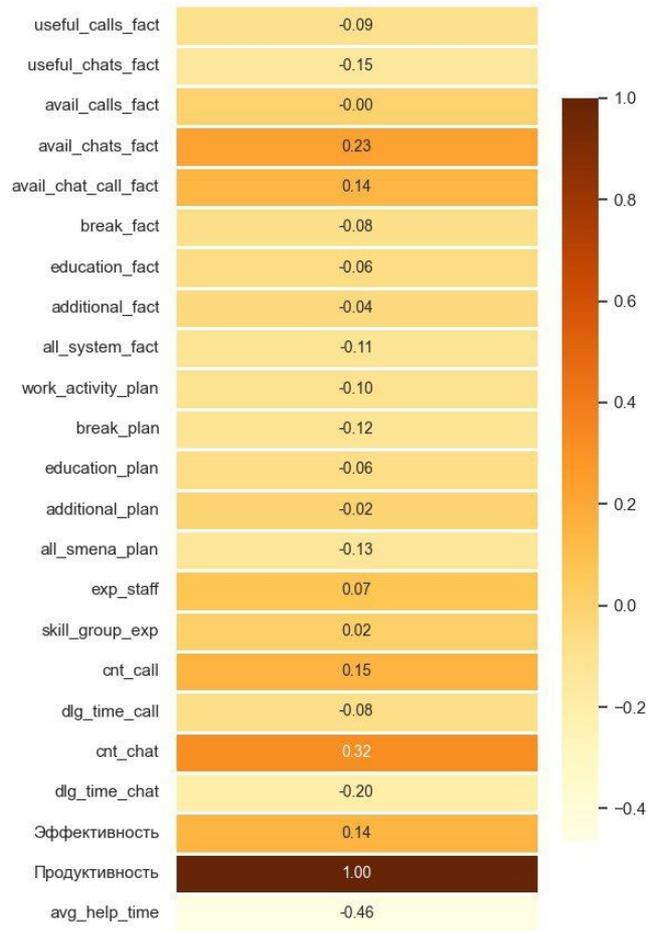
Предварительный анализ

Матрица корреляции числовых параметров



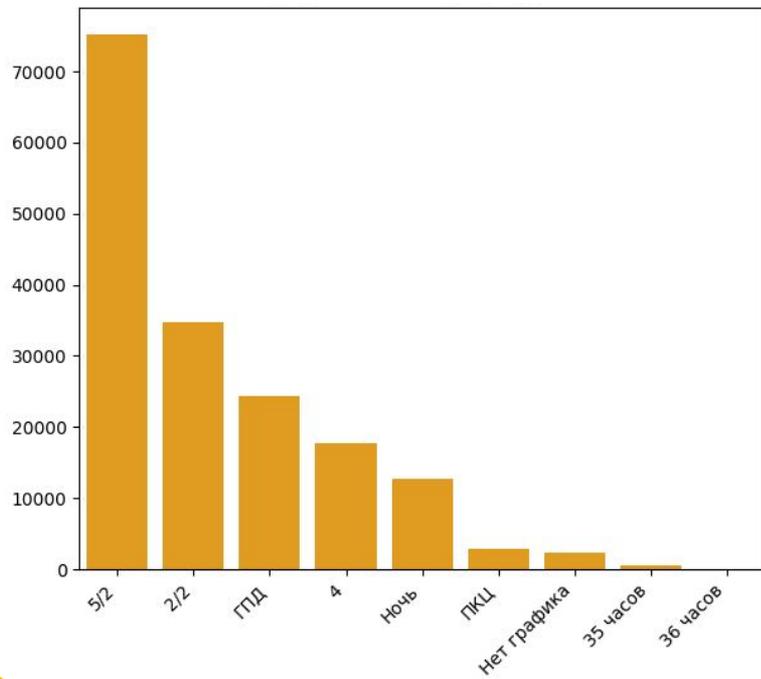
Учтем все переменные из таблицы вместе с переменной “Продуктивность” в матрице корреляции.

Матрица корреляции продуктивности

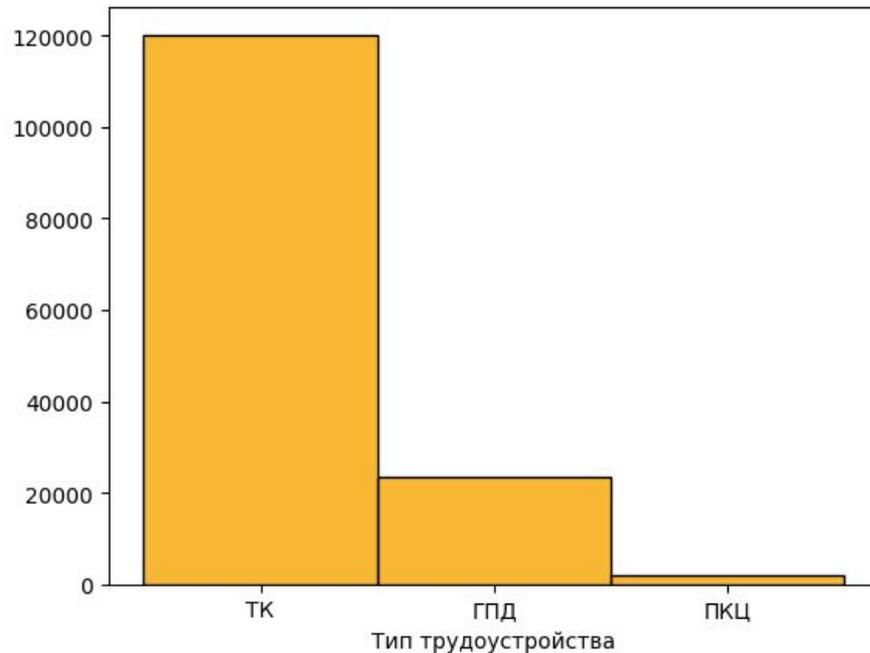


Распределение по:

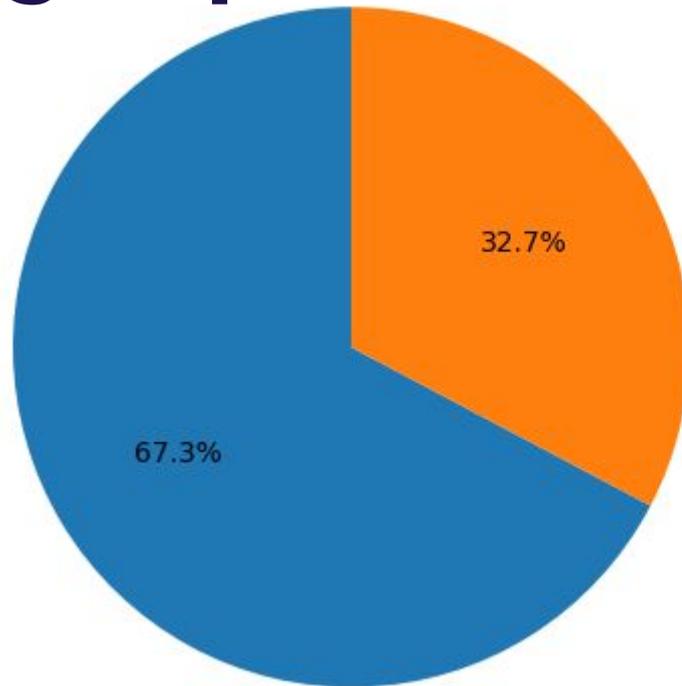
графику работы



типу трудоустройства



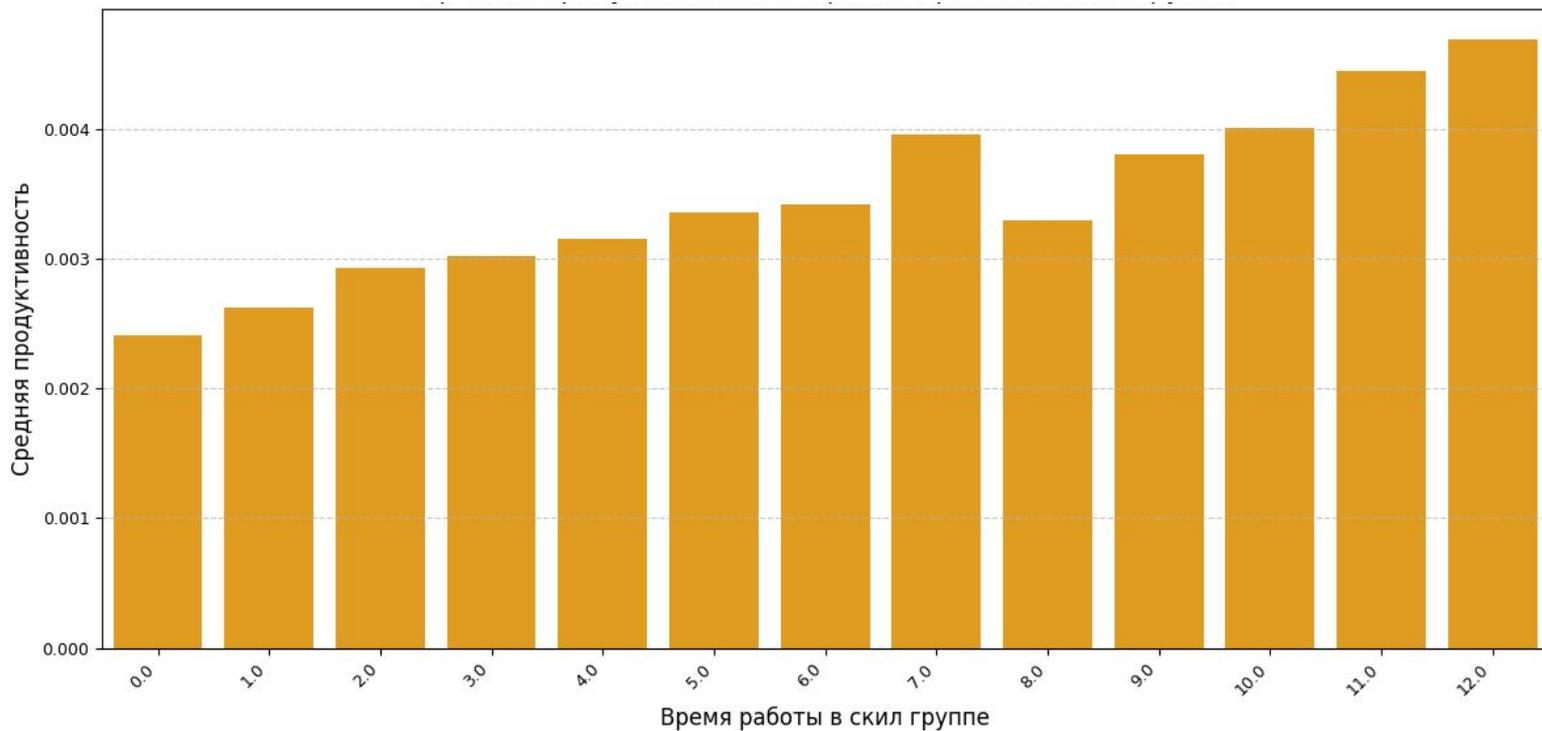
Процент людей, которые хотя бы раз меняли skill group



- Люди, которые хотя бы раз меняли skill group
- Люди, которые не меняли skill group

Людей, которые не меняли skill group примерно в два раза меньше.

Средняя продуктивность от времени работы в скил группе (конкретно для Москвы)

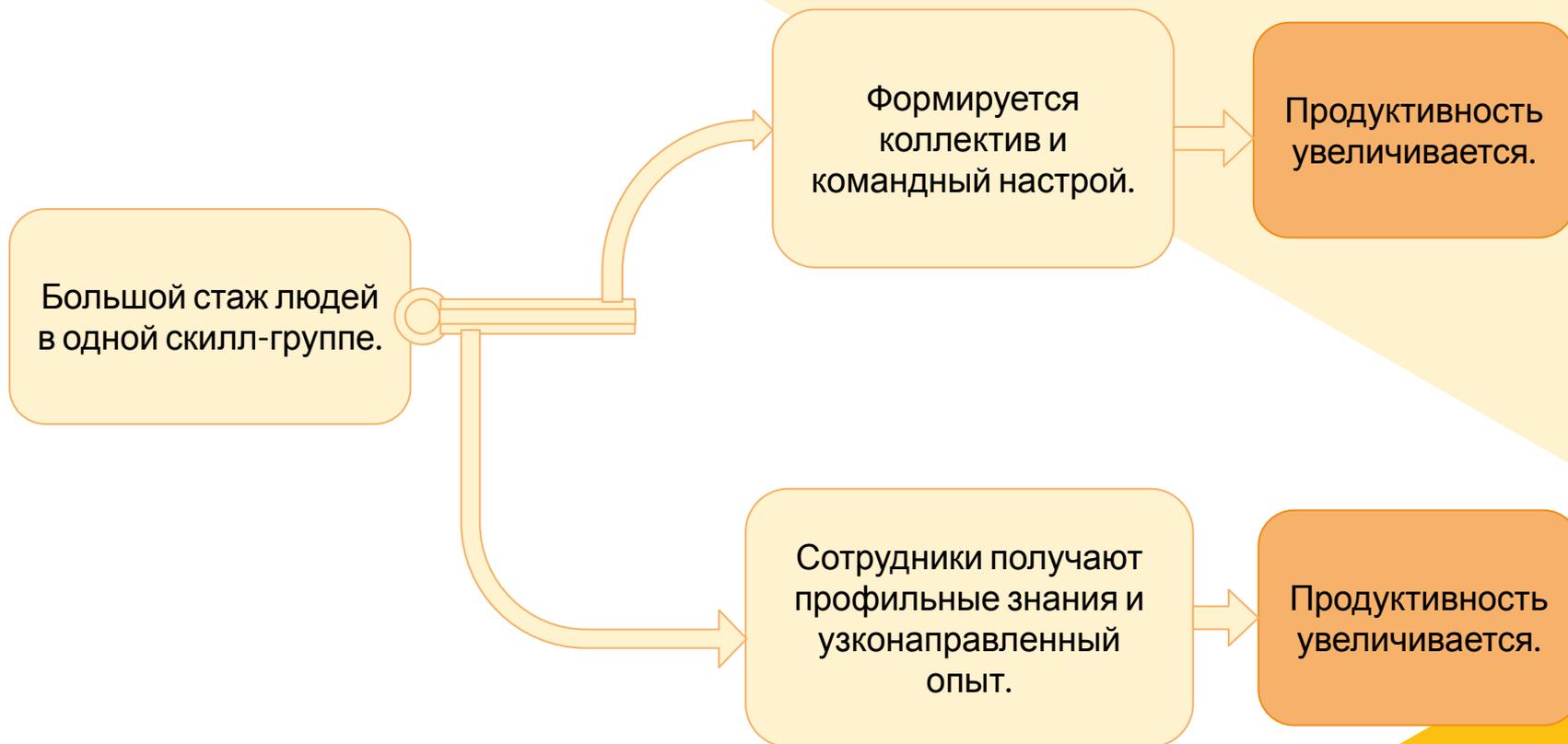


Гипотеза:

**Гипотеза - продуктивность
сотрудника связана с его
стажем в текущей скилл-
группе.**



Механизм гипотезы



Математическая модель 1

ЦЕЛЬ:

Оценить взаимосвязь между **стажем сотрудника в его скилл-группе** и его **продуктивностью**.

МЕТОД:

Применение статистического теста (корреляция Пирсона).

Математическая модель 1

Корреляция Пирсона:

Значения p-value
оказались
приемлемыми только у
некоторых регионов.

К статистически
значимым можно
отнести только
густонаселенные
регионы.

Регион	p-value
Москва г	< 0.01
Самарская обл	< 0.01
Калининградская обл	< 0.01
Ростовская обл	< 0.01
...	...
Новосибирская обл	0.0819
Саратовская обл	0.0914
Чувашская республика	0.1097
Тюменская обл	0.1406

Математическая модель 2

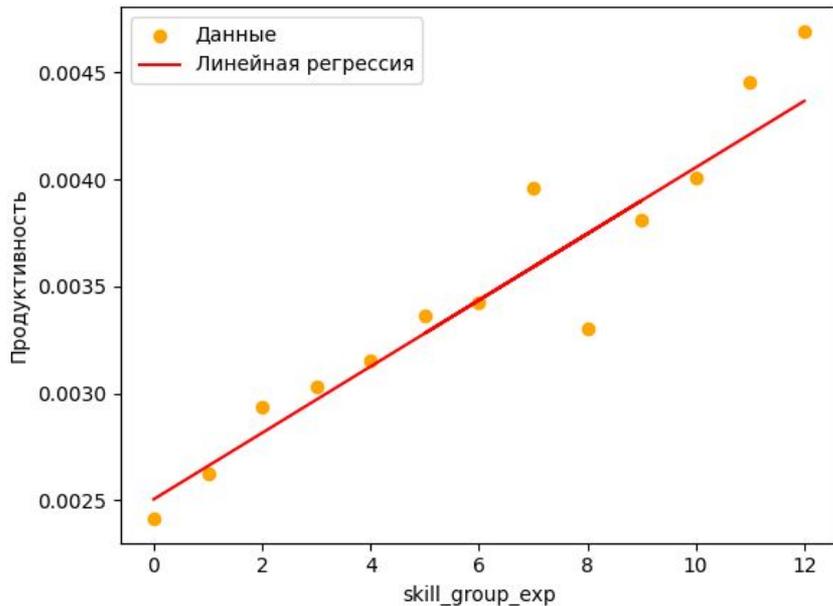
ЦЕЛЬ:

Подтвердить взаимосвязь между **стажем сотрудника в его скилл-группе** и его **продуктивностью в крупнонаселенных городах**.

МЕТОД:

Линейная регрессия.

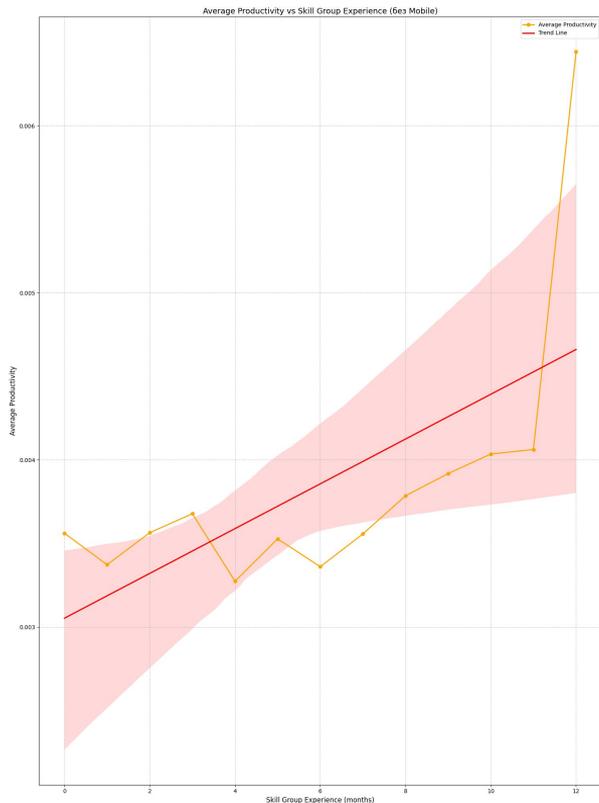
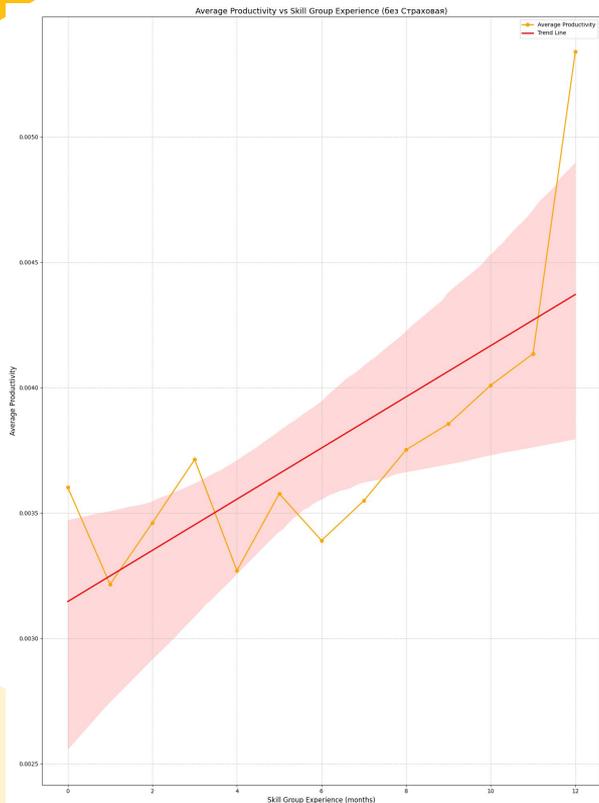
Линейная регрессия/ Результат



Мы взяли крупные города и **медианное значение продуктивности** в этой скилл группе. Вот что мы получили:

- $R^2 = 0.894$
- F-stat=67
- Корреляция : 0.90
- p-value < 0.01

Подтверждение регрессии



На этих графиках вы видите зависимость продуктивности в скилл группе без какого-либо бизнес направления, ведь если в каком-то направлении оно не подтвердится, то гипотеза окажется ложной.

Вывод:

гипотеза подтвердилась частично

Продуктивность сотрудника положительно коррелирует с его стажем в текущей скилл-группе, но это корректно только для густонаселенных регионов.

В густонаселенных городах больше текучка кадров, из-за чего лучше отображается зависимость продуктивности от времени нахождения в команде.



Применение

- Рекомендуем не переводить человека из команды, где он давно работает
- Формирование новых постоянных групп
- Штат team-менеджеров для повышения продуктивности внутри команд
- Уменьшить текучку кадров в больших городах: удерживать хорошего специалиста и развивать его внутри команды

Ограничения и перспективы

- **Употребим только для крупных городов.**
- **Не хватает дополнительных данных (например, возраст, зарплата, причина перехода и прочие параметры).**
- **Заполнение всего датасета без пропусков (из-за них наша гипотеза не подтверждается на все сто).**
- **Анализ прибыли компании и количества клиентов, необходимость найма новых сотрудников, расширение сферы влияния.**

Время вопросов!



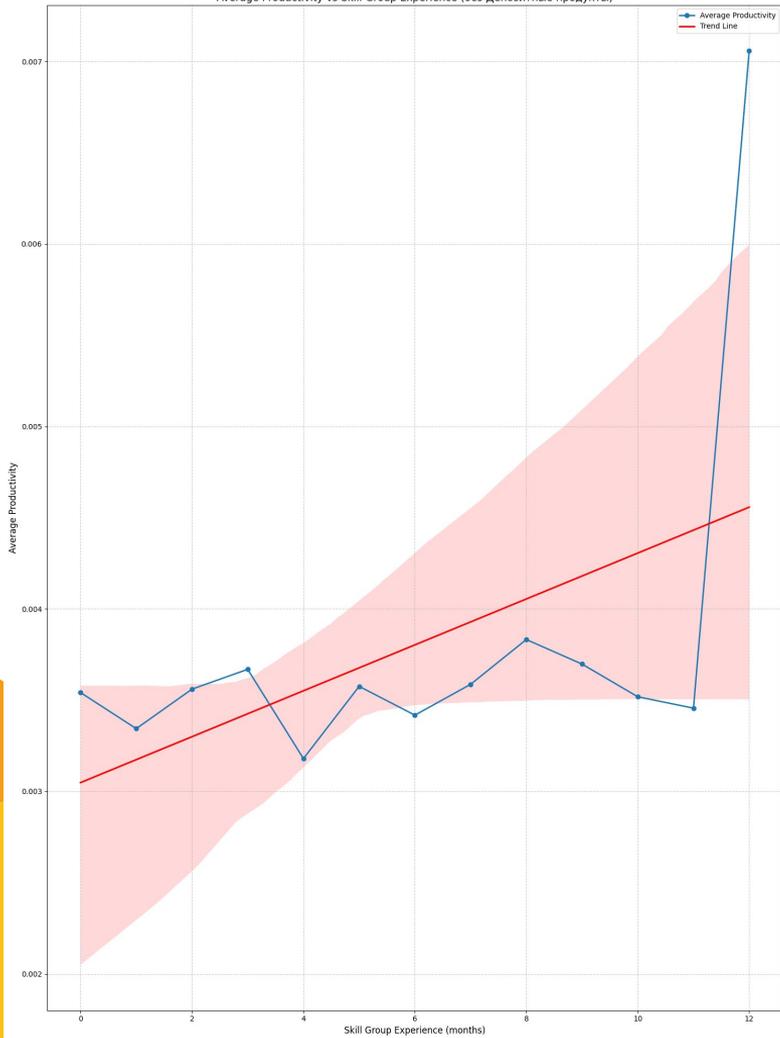
GitHub нашего проекта:



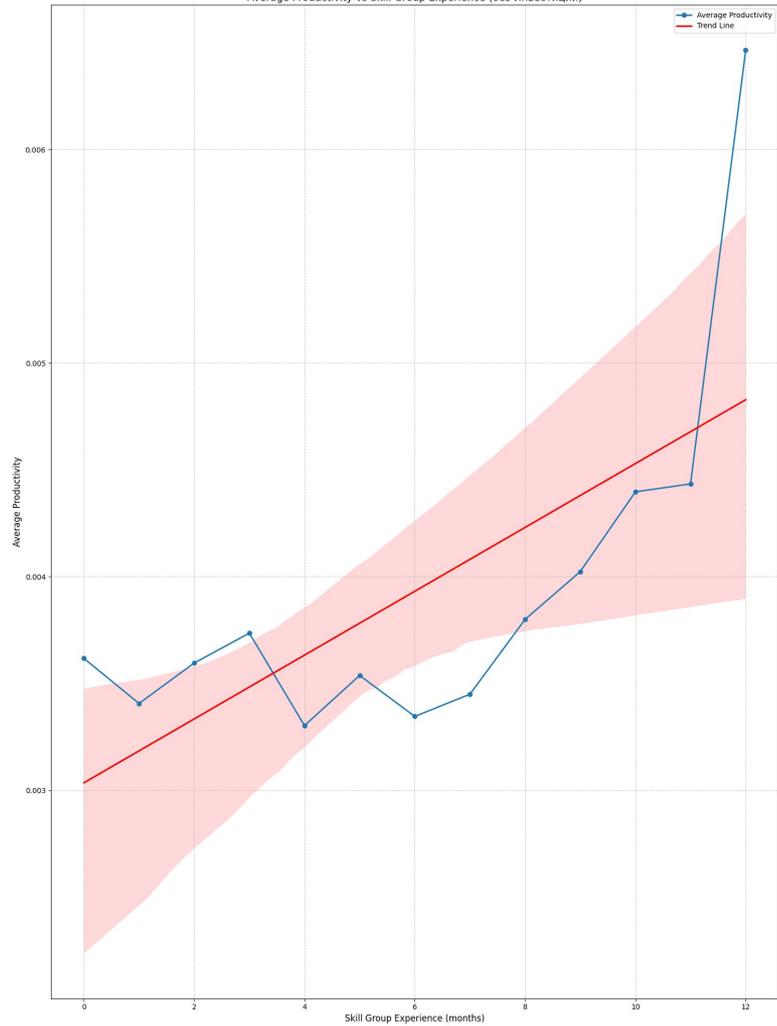
По всем вопросам:



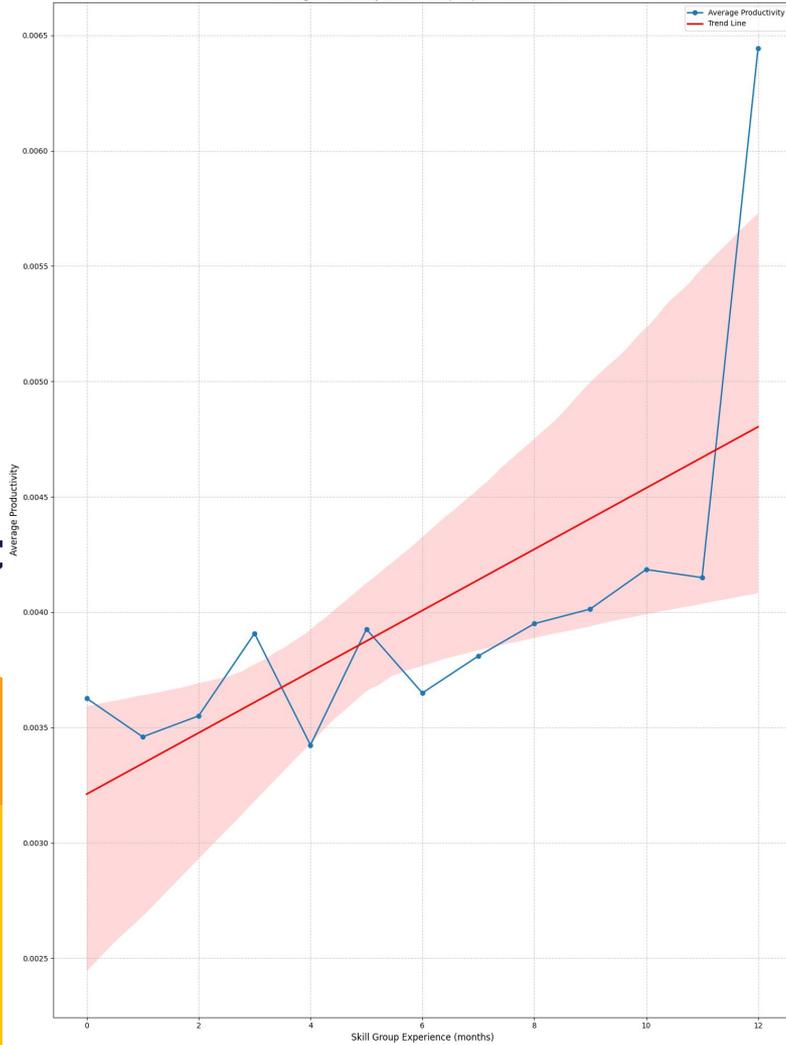
Average Productivity vs Skill Group Experience (без Депозитные продукты)



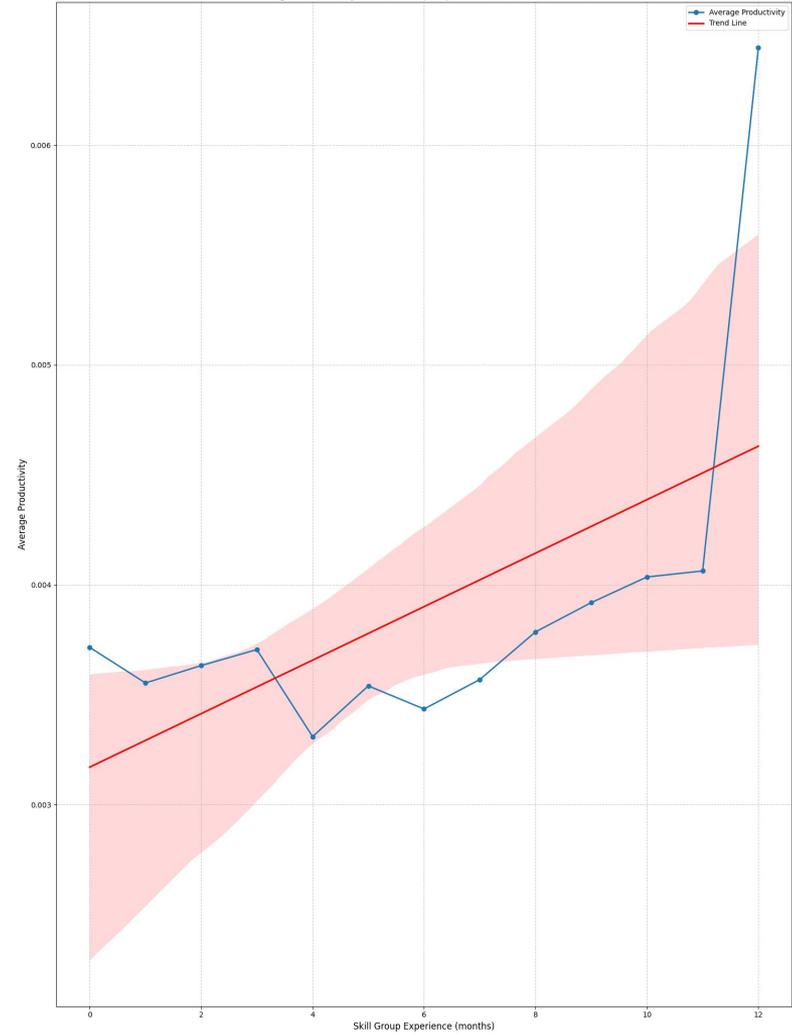
Average Productivity vs Skill Group Experience (без Инвестиции)



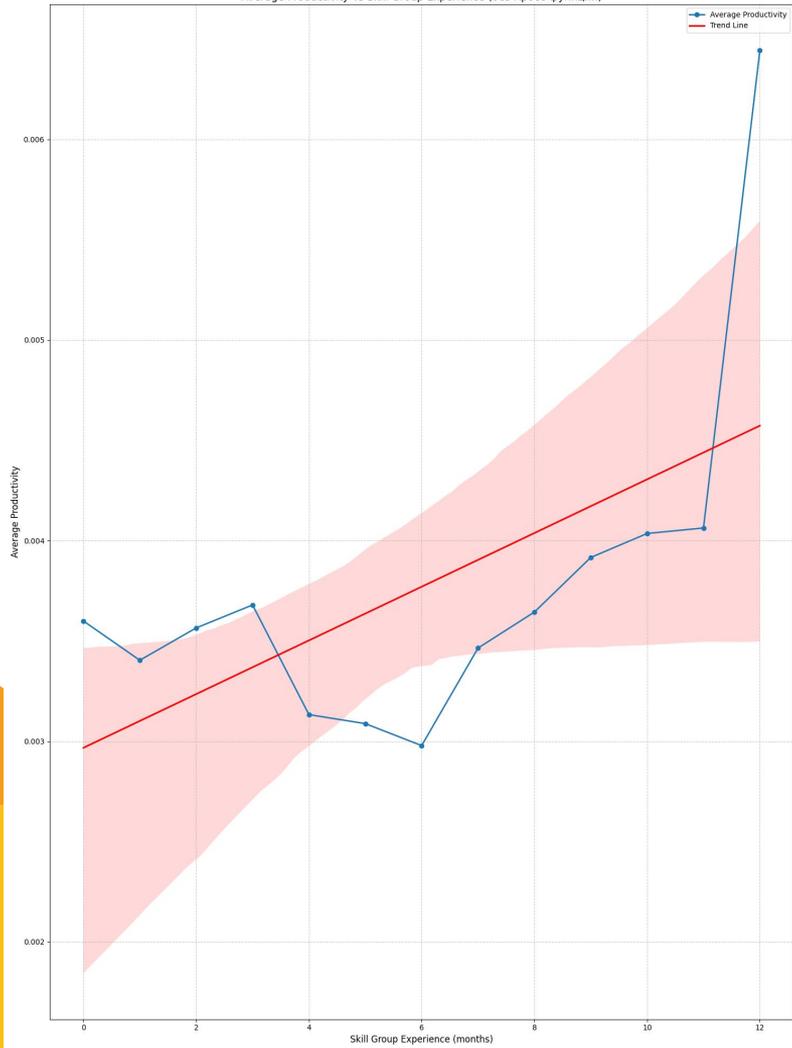
Average Productivity vs Skill Group Experience (6ea SME)



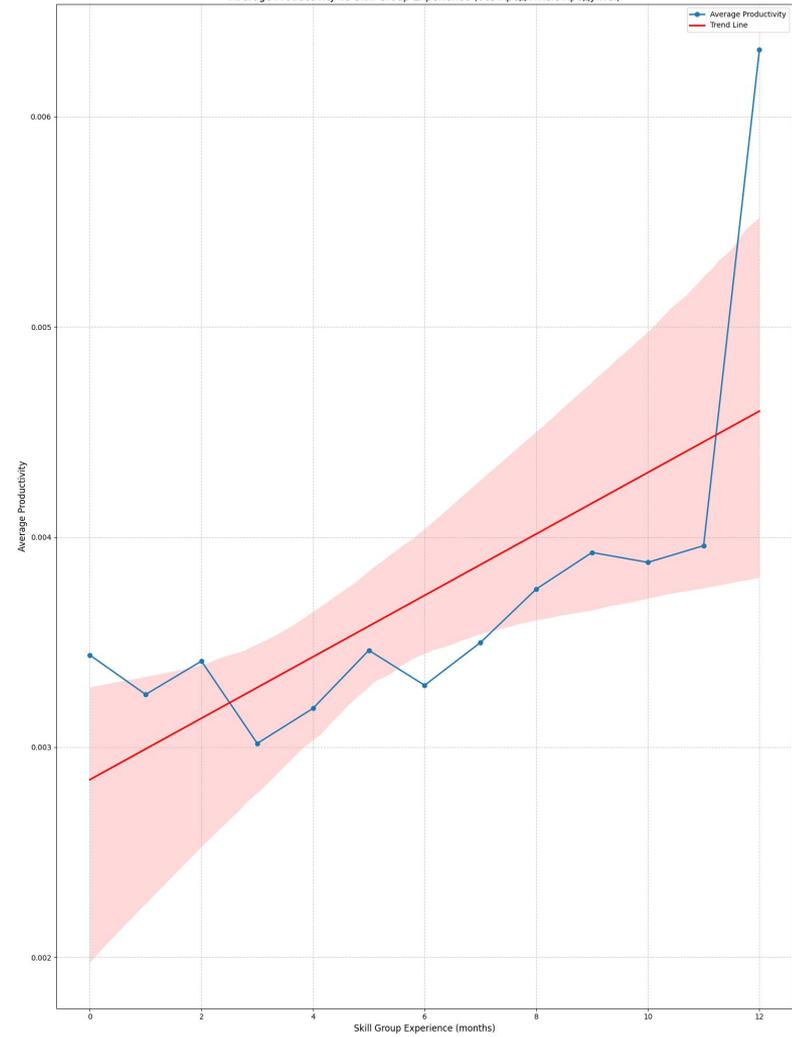
Average Productivity vs Skill Group Experience (6ea Premium Банк)



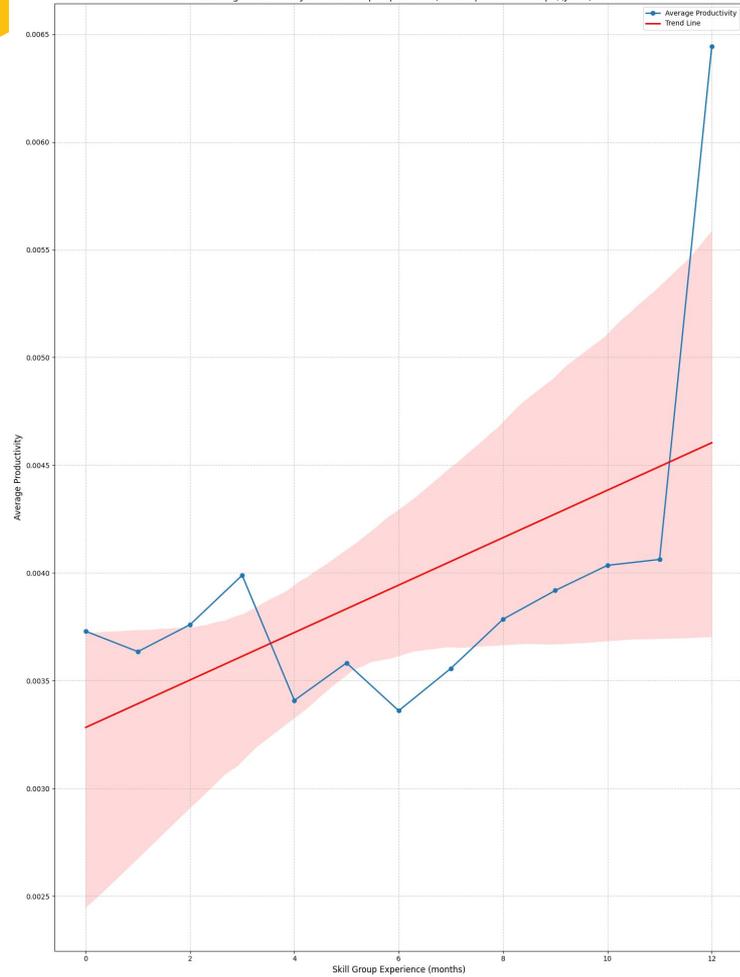
Average Productivity vs Skill Group Experience (без Кросс-функции)

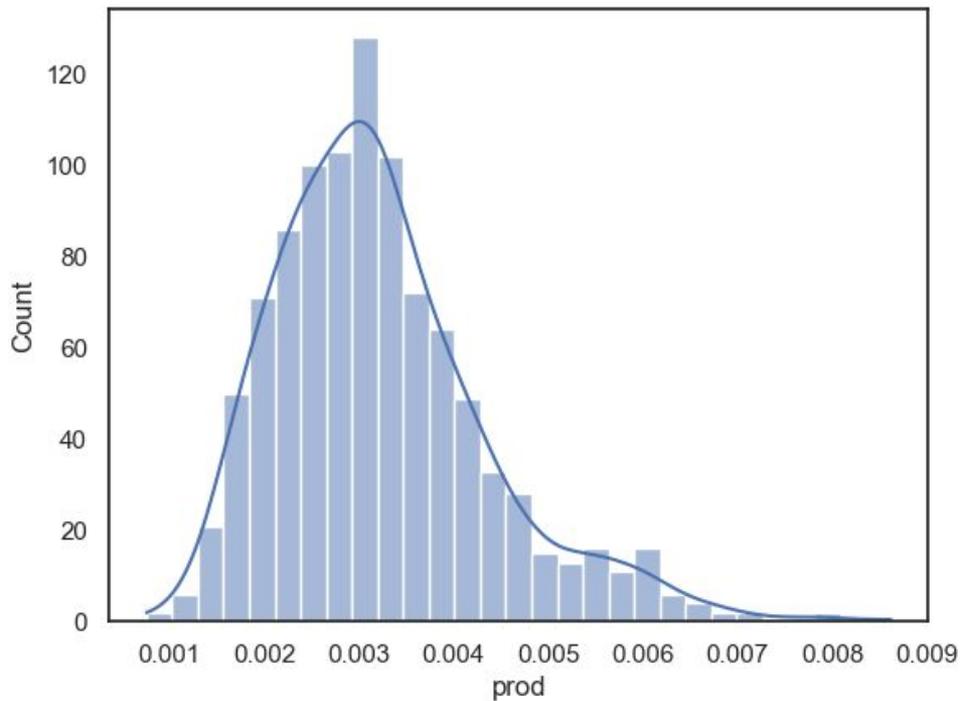


Average Productivity vs Skill Group Experience (без Кредитные продукты)

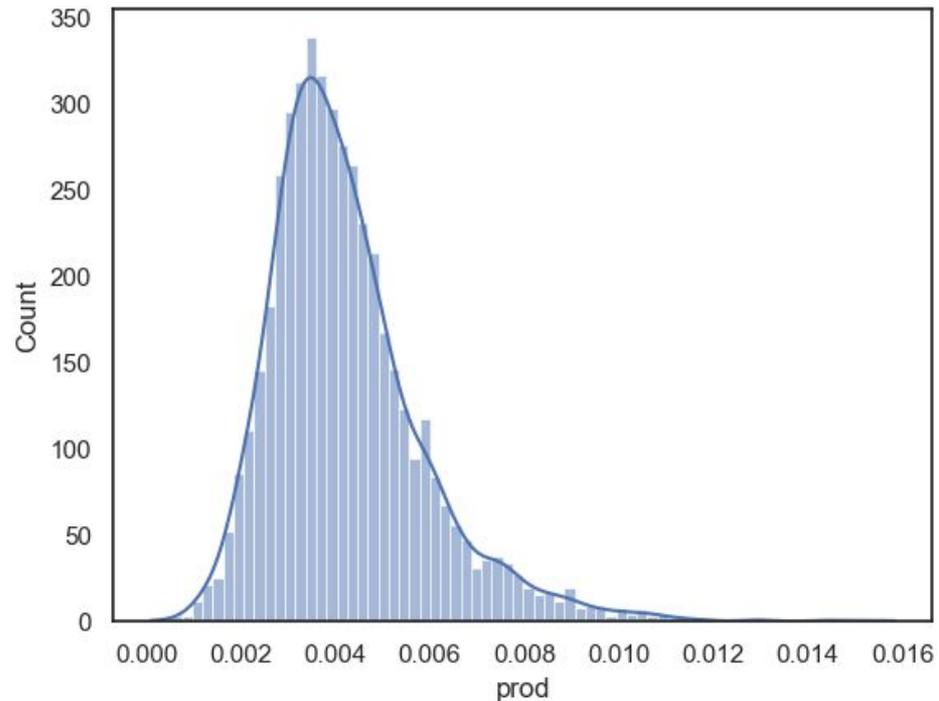


Average Productivity vs Skill Group Experience (без Нефинансовые продукты)

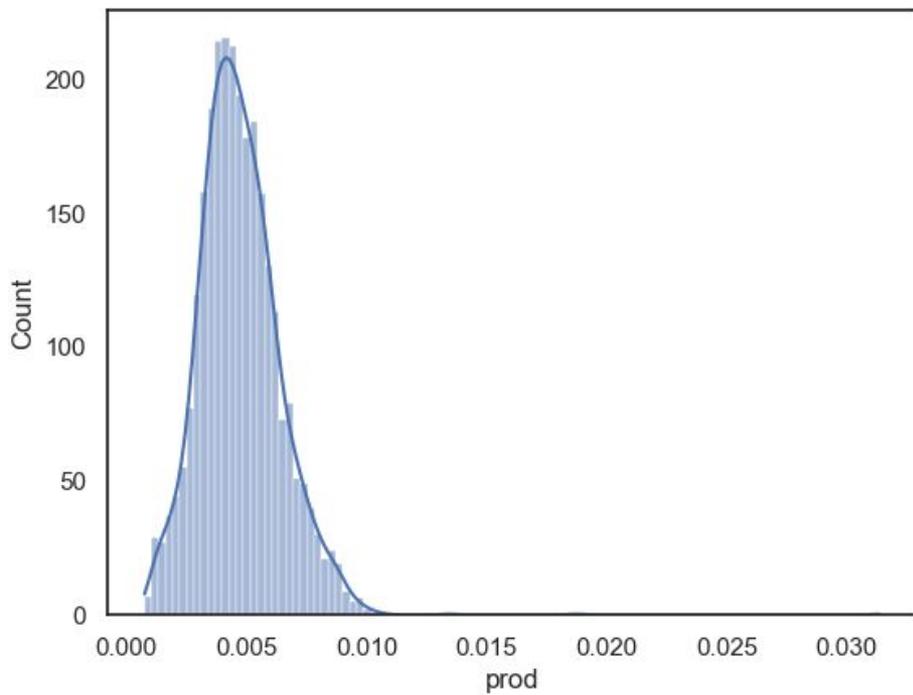




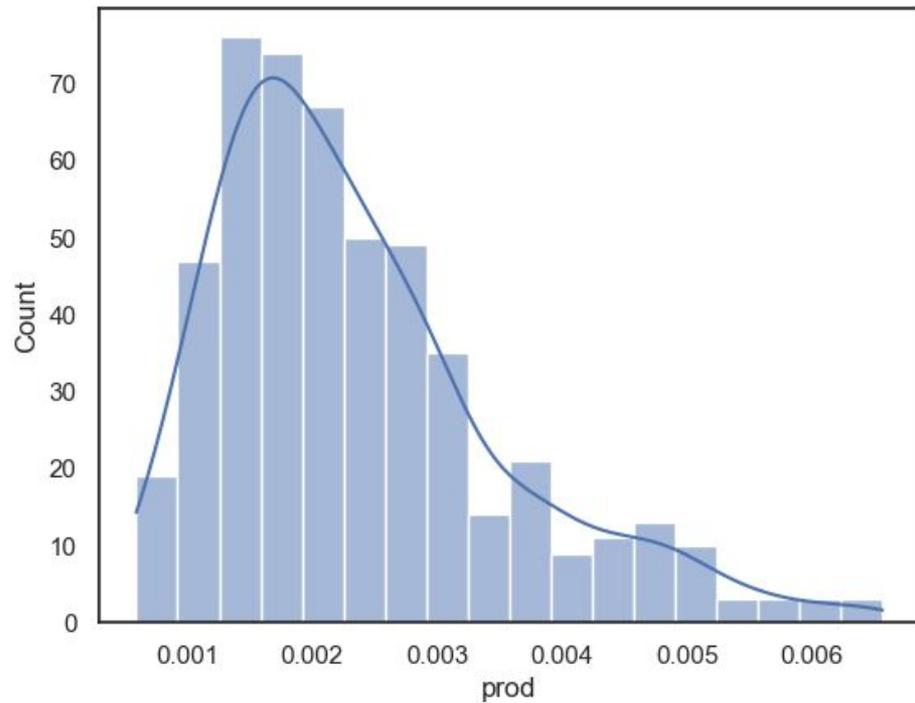
нефинансовые продукты



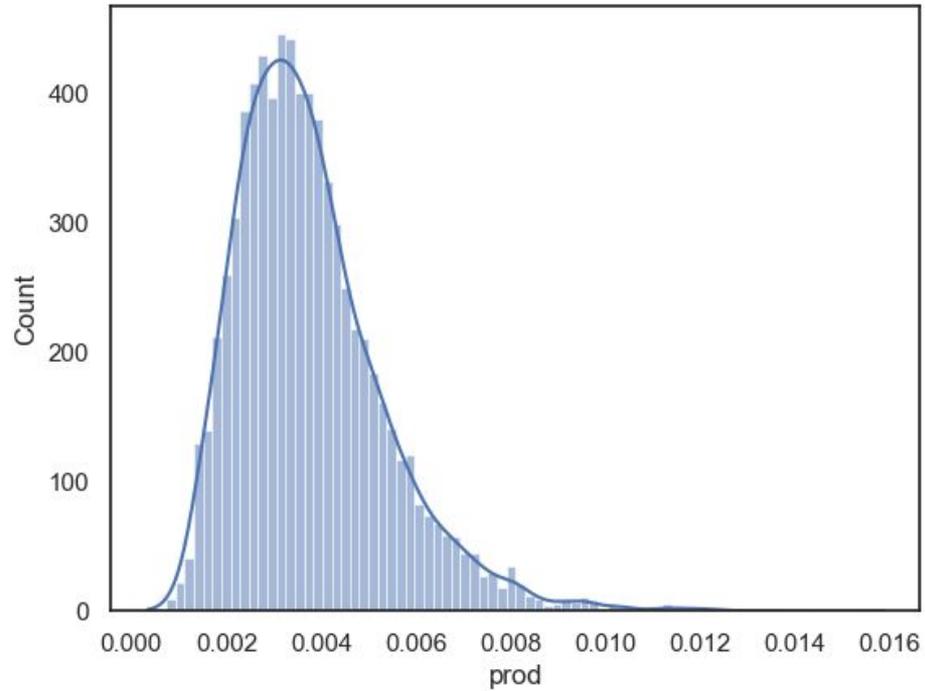
кредитные продукты



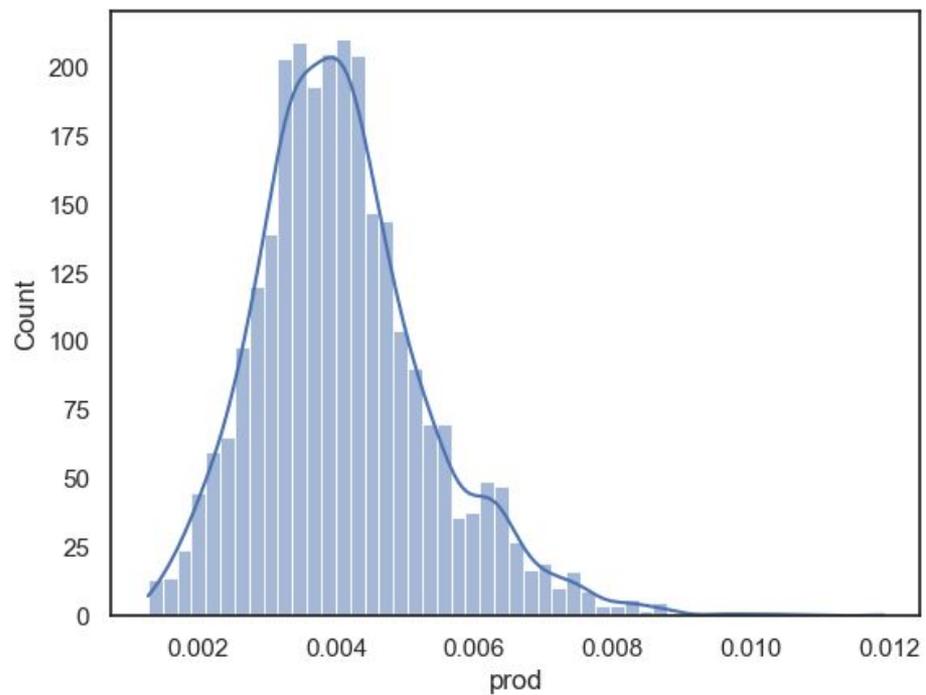
страховая



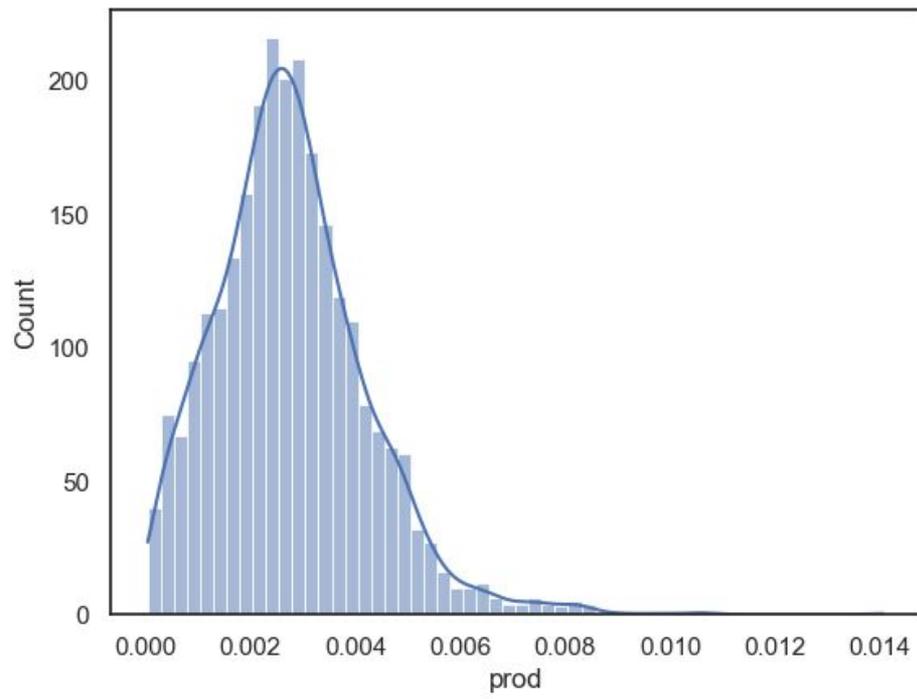
Premium банк



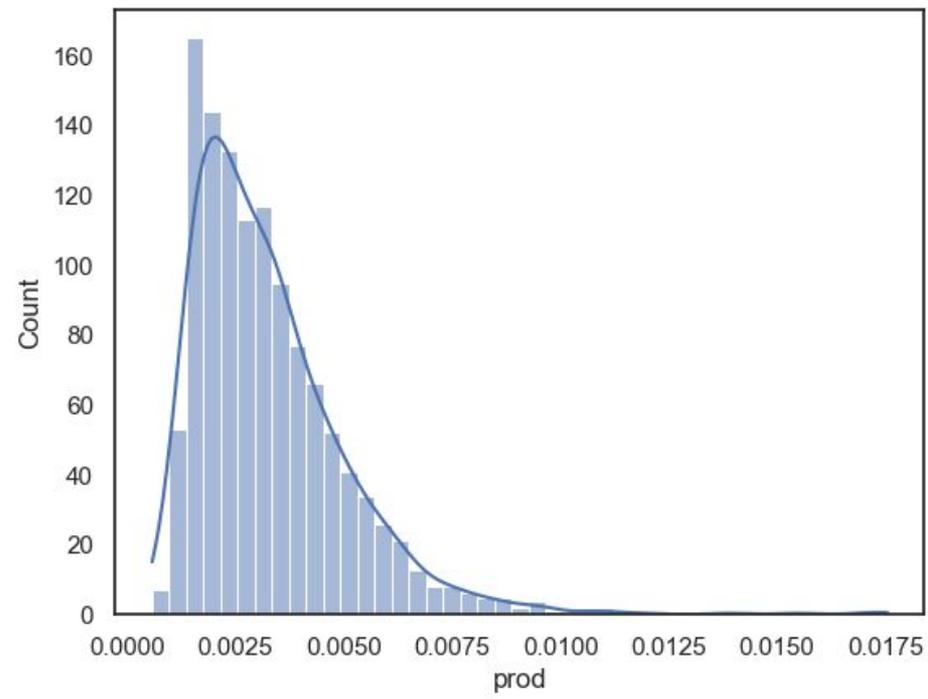
Депозитные продукты



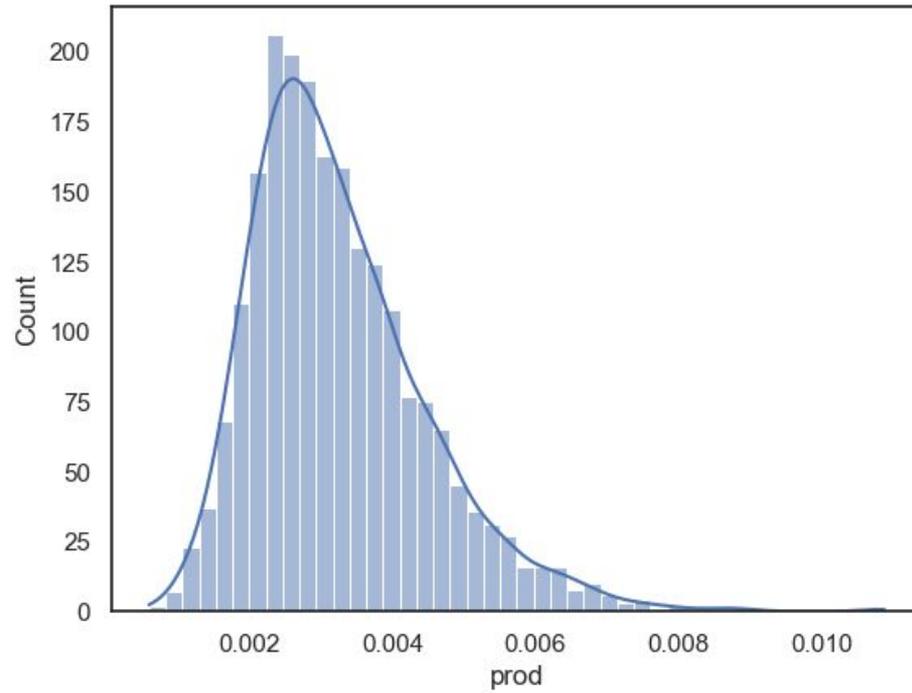
T - mobile



SME



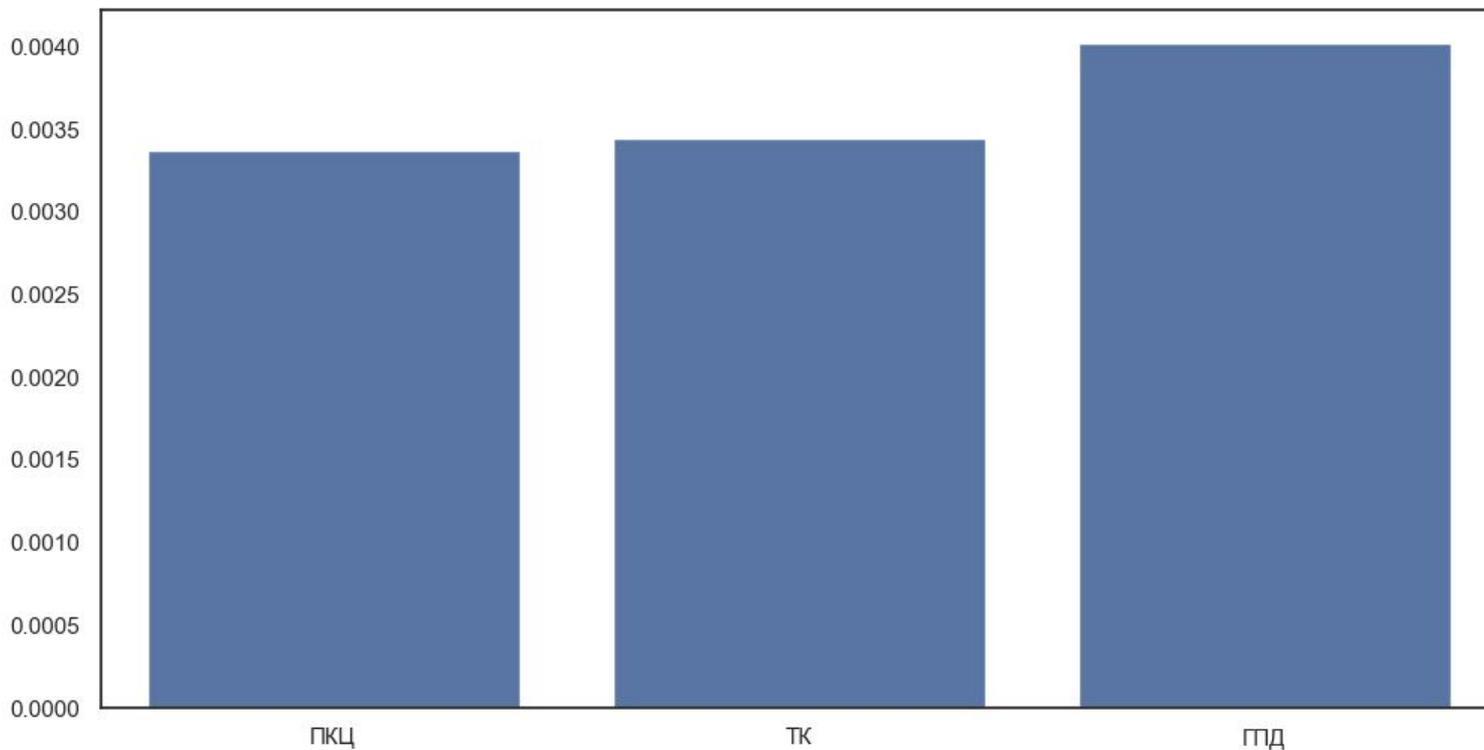
Кросс-функции



Инвестиции

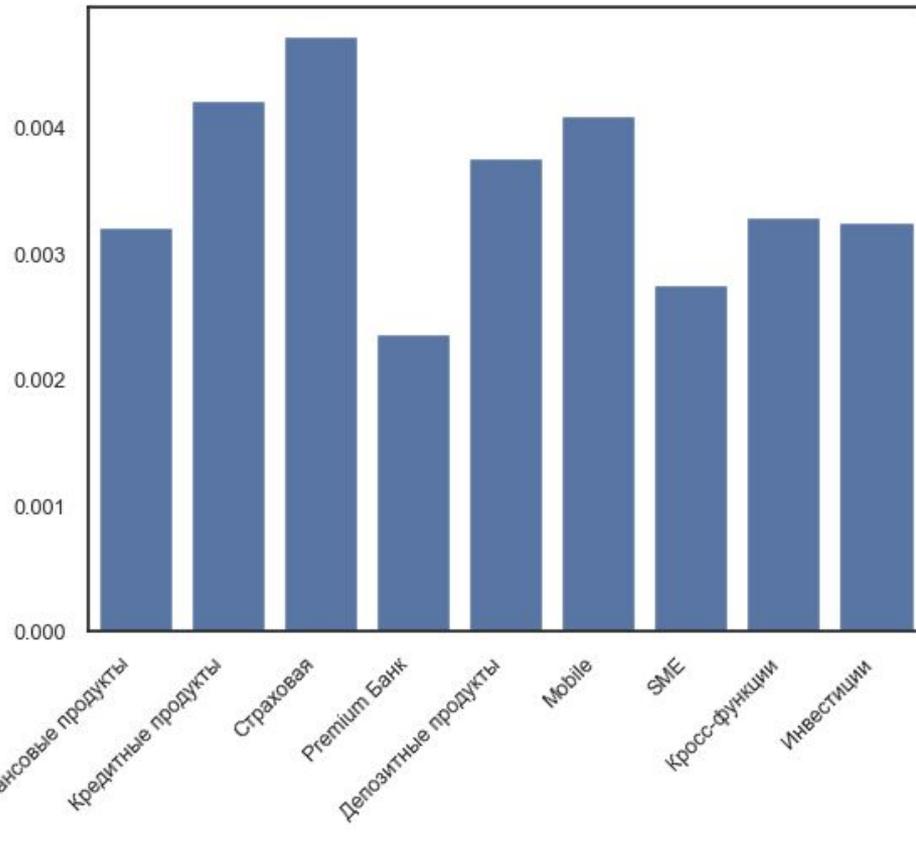
1	Регион	Коэффициент корреля...	P-value
39	Новосибирская обл	0.074442853	0.0819463...
40	Саратовская обл	0.055898511	0.09140561
41	Чувашская республика Чувашия	-0.71564235	0.1097924...
42	Тюменская обл	-0.120063915	0.1406531...
43	Псковская обл	0.941174671	0.2194469...
3	Москва г	0.625416584	6.27986E-21
4	Самарская обл	0.391417574	1.10998E-19
5	Иркутская обл	0.540921098	4.05708E-18
6	г Москва	0.133766024	8.28018E-17

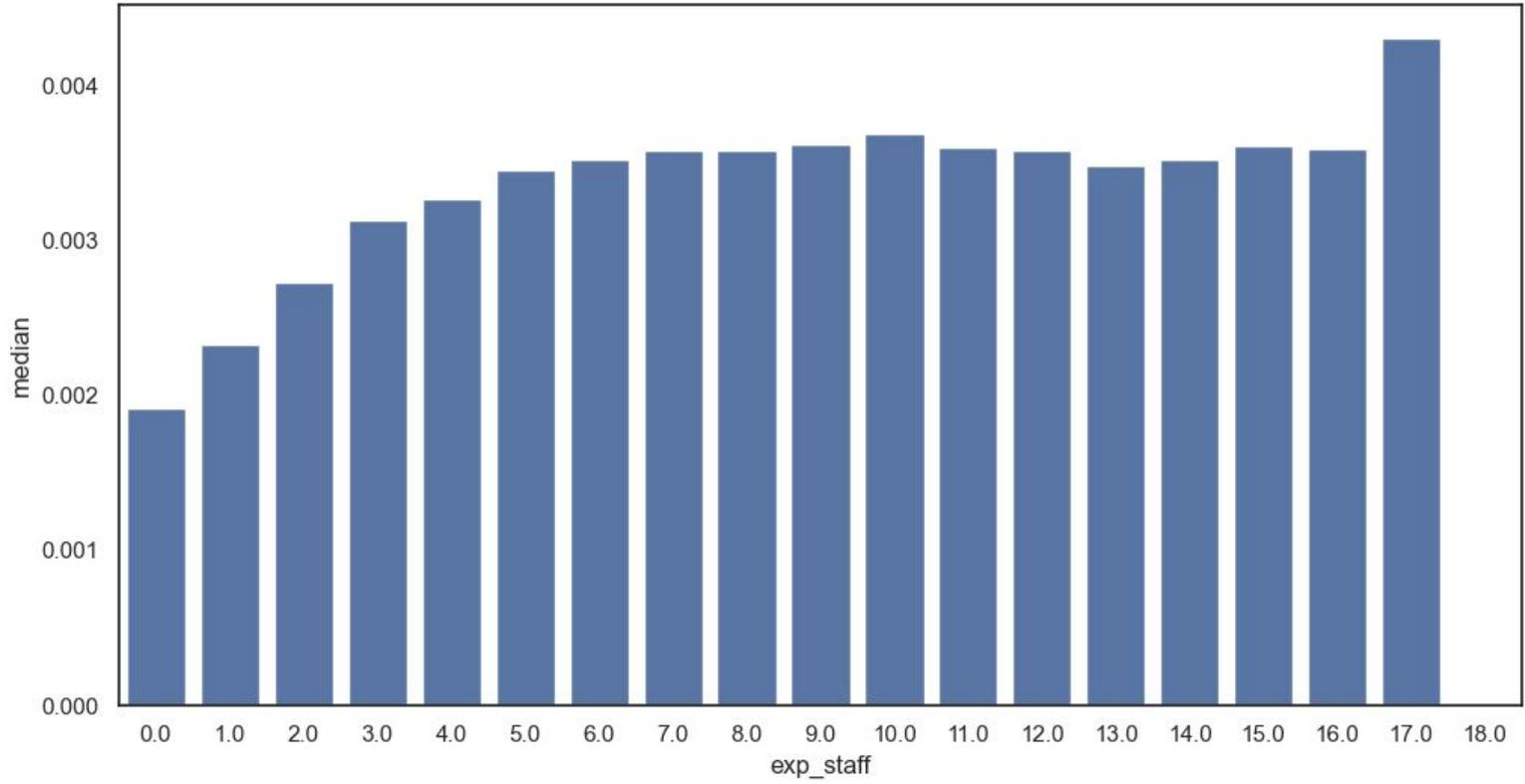
Продуктивность от типа трудоустройства

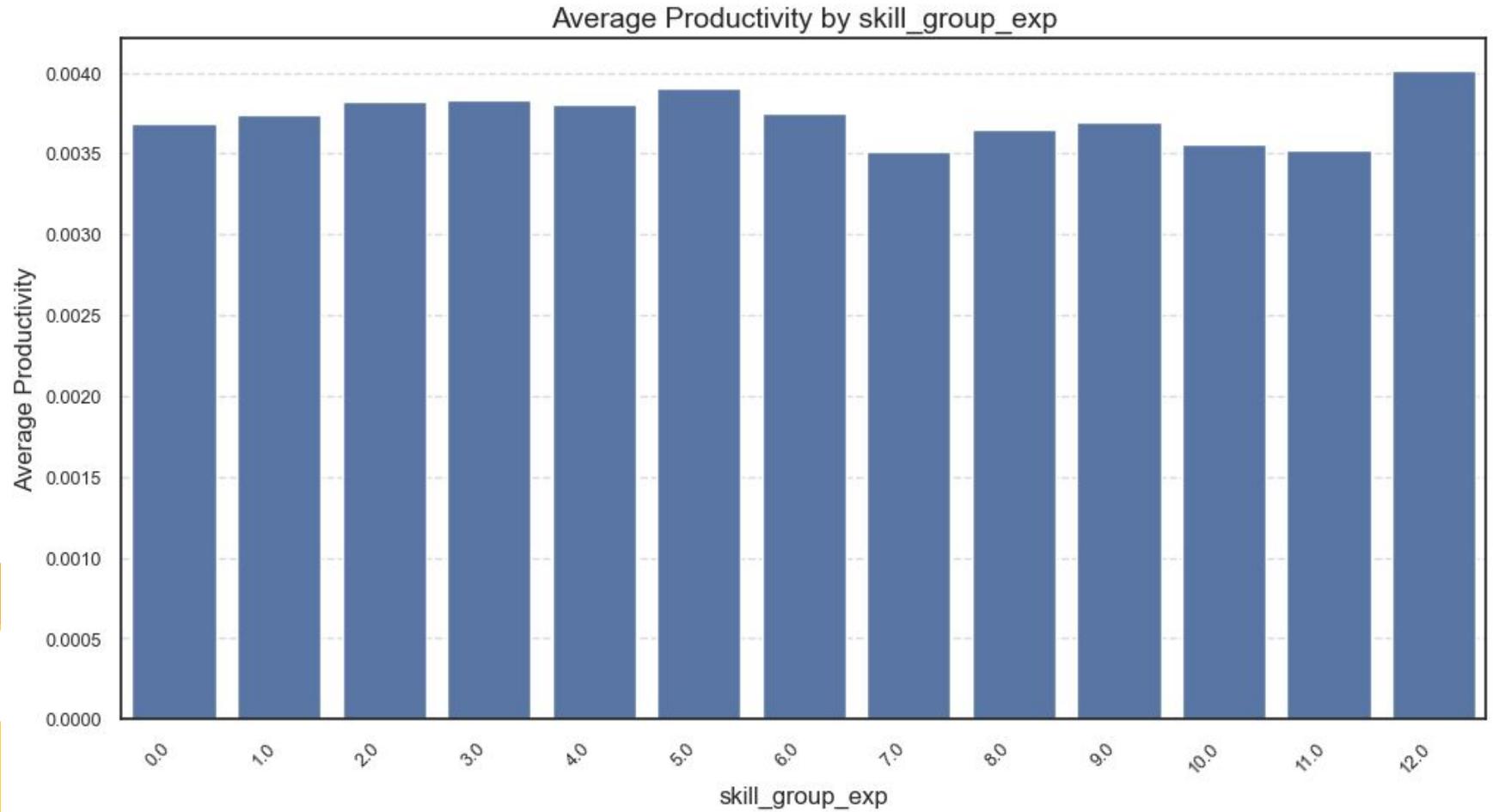


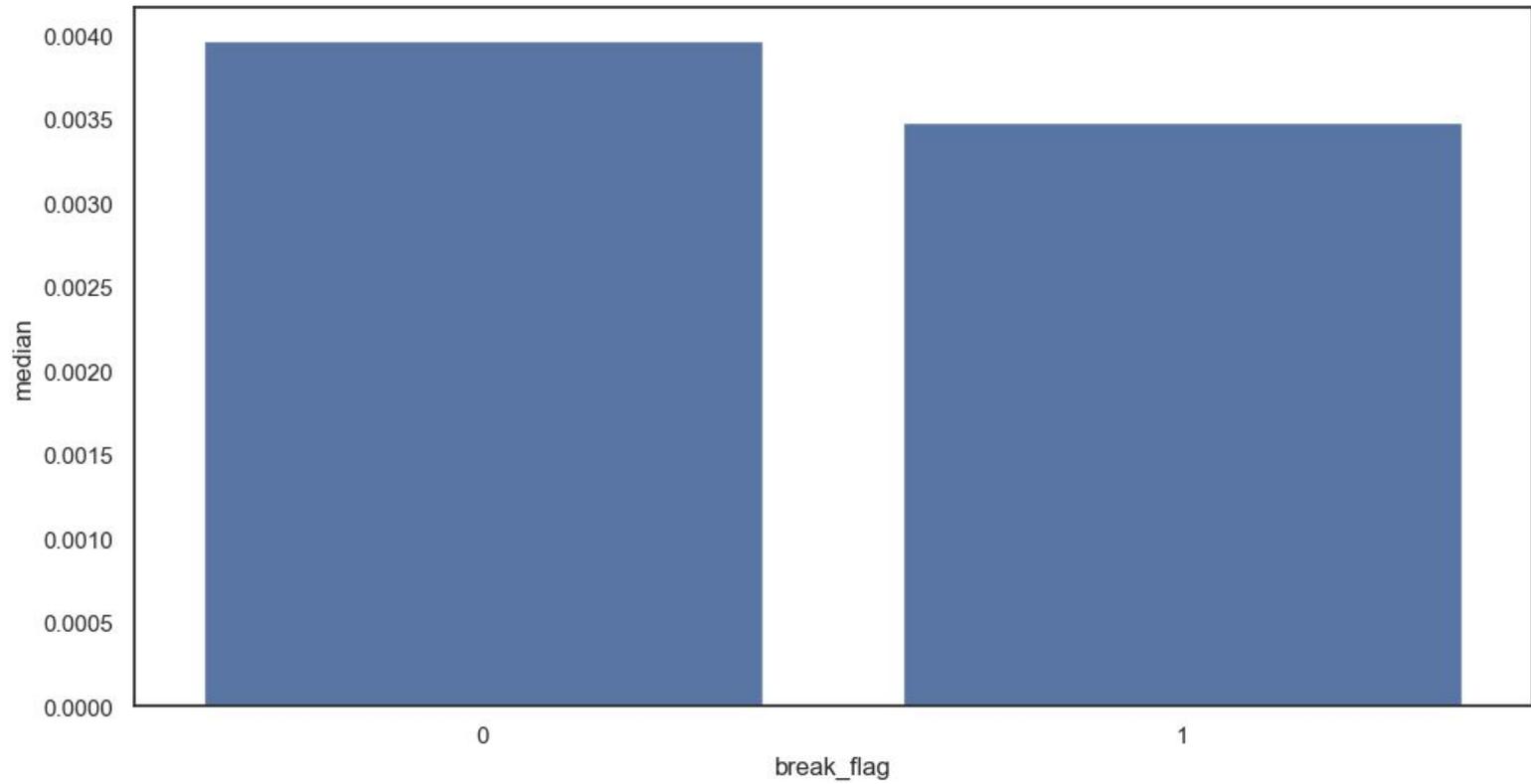
Продуктивность

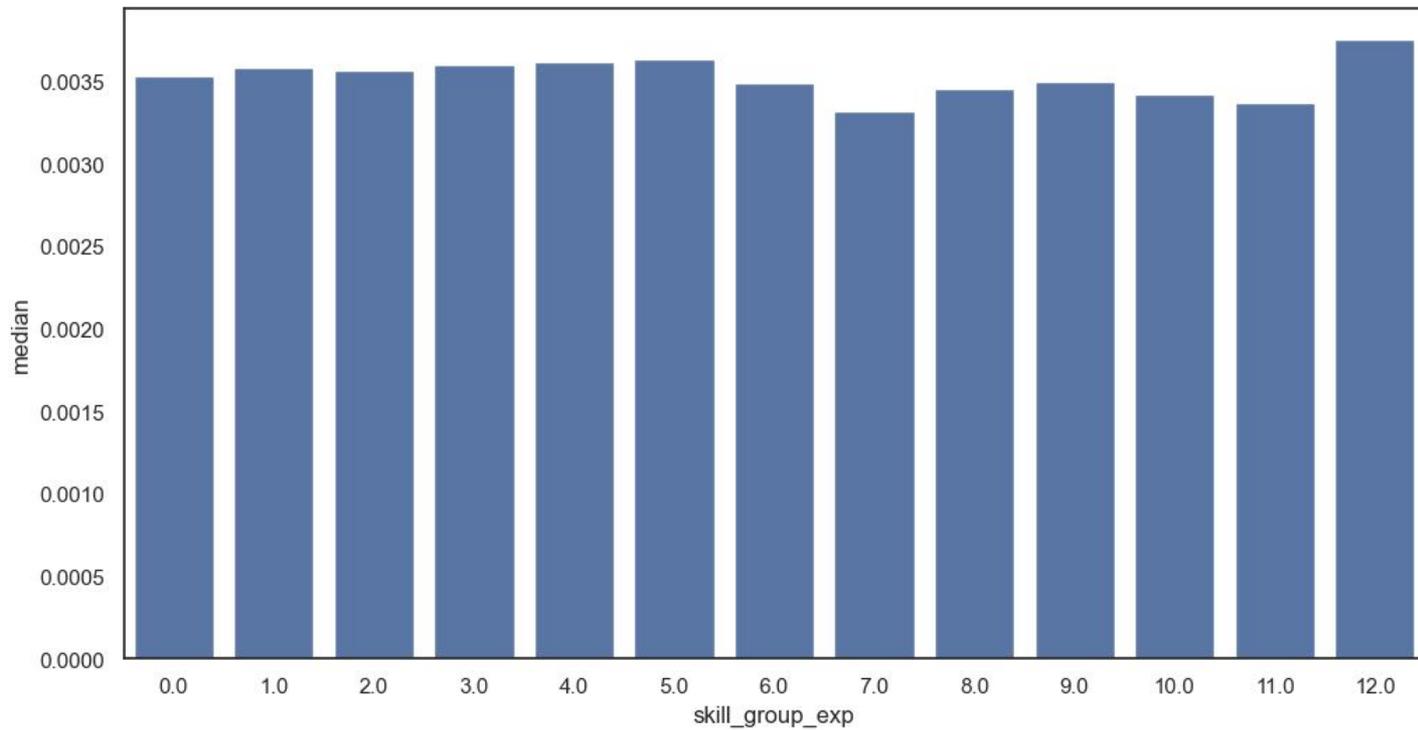
групп





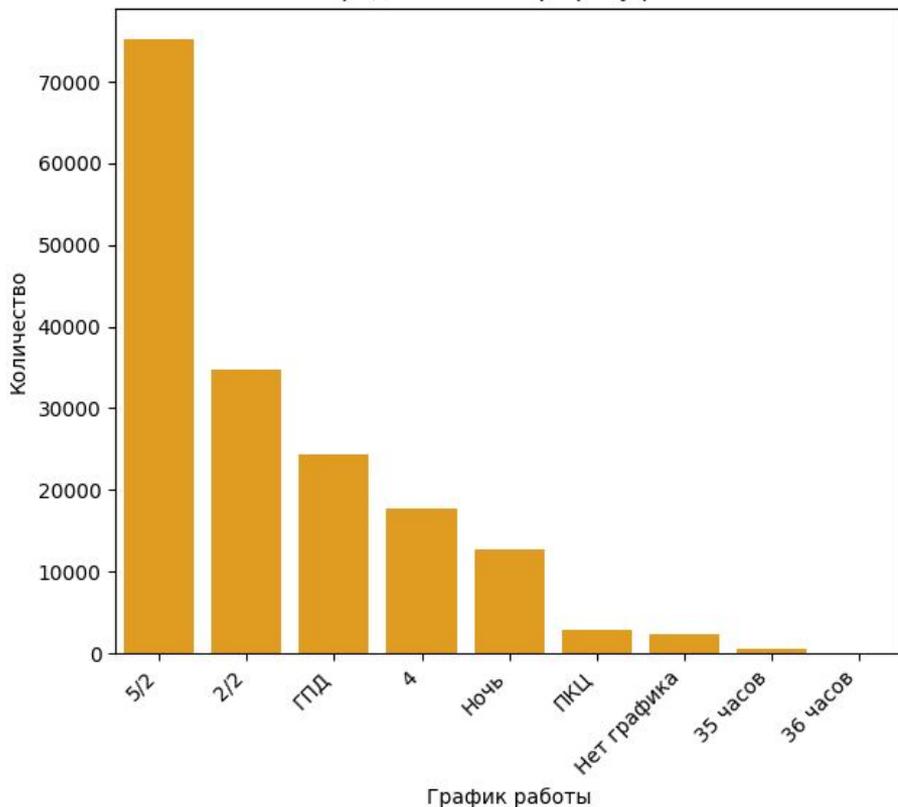






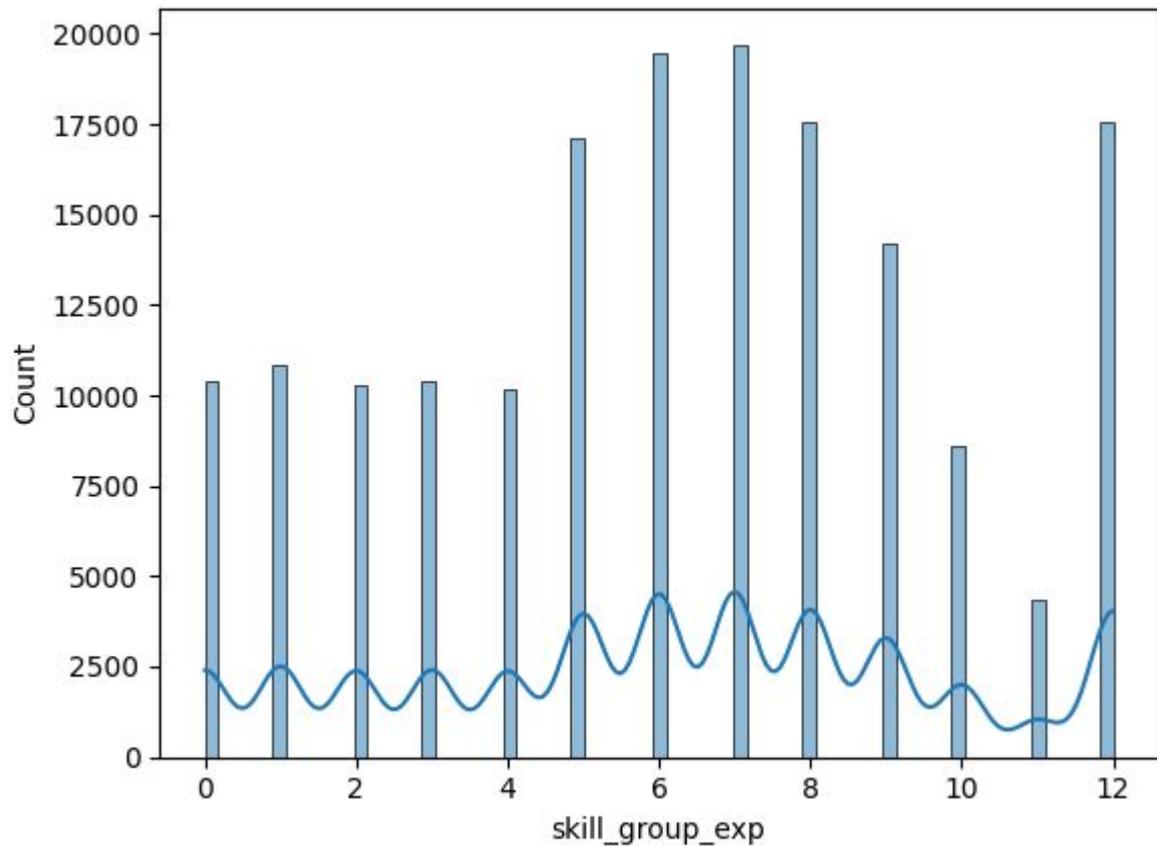
Структура Данных и подготовка

Распределение по графику работы

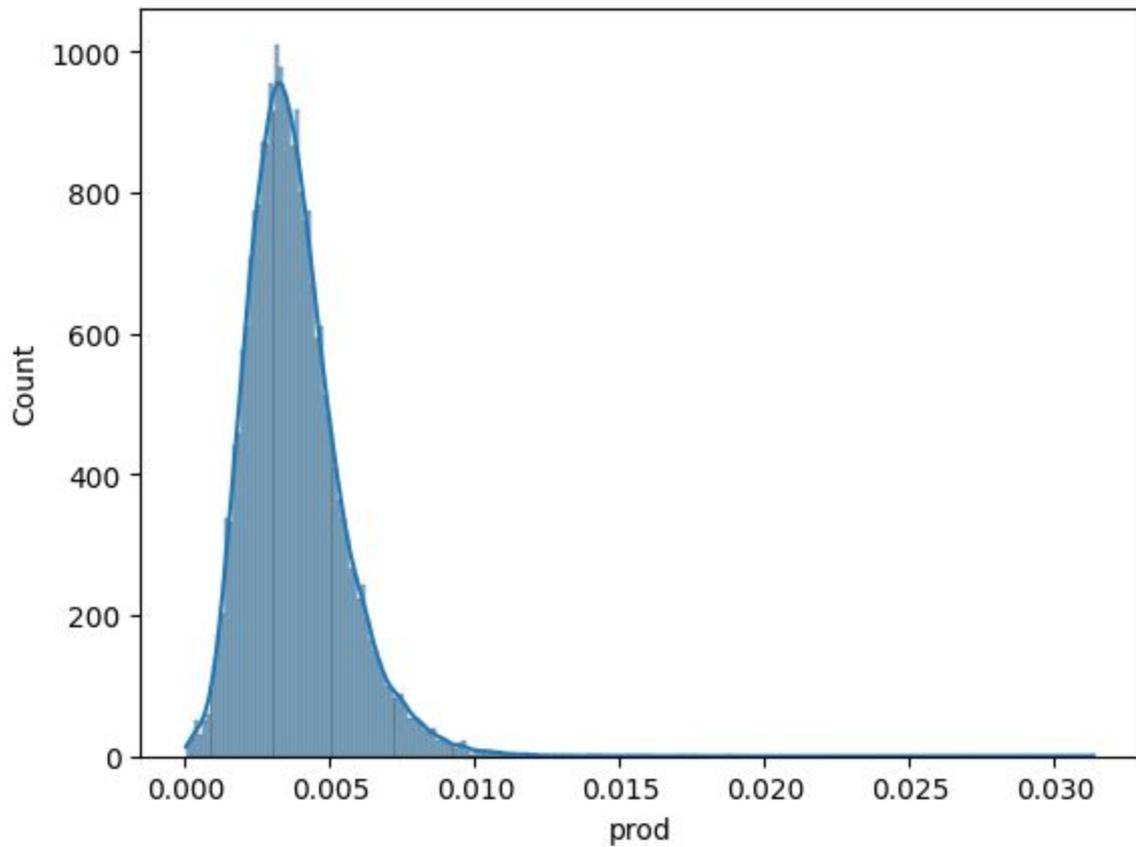


Здесь можно увидеть, по какому графику чаще всего работают люди. Больше 70 тыс. чел. работают по графику 5/2, а самый маленький график по кол-во людей - 36 часов.

Shapiro-Wilk test statistics: 0.9592

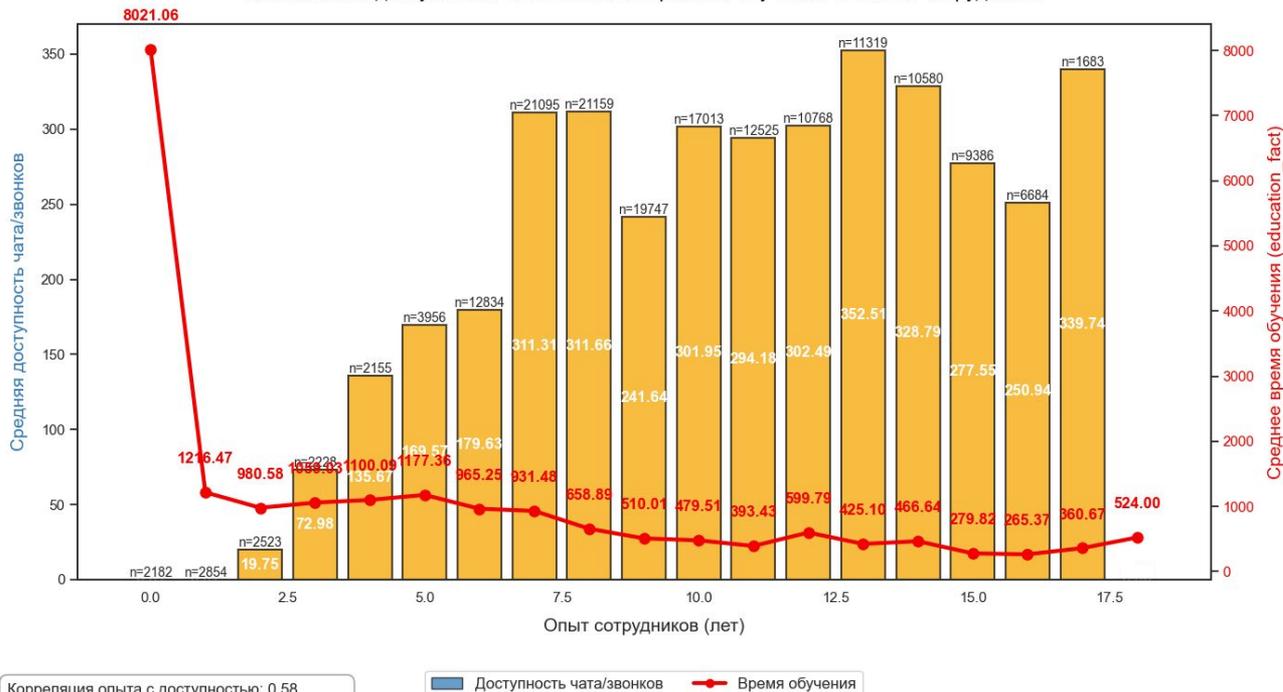


Shapiro-Wilk test statistics: 0.9441



Первичный анализ

Зависимость доступности чата/звонков и времени обучения от опыта сотрудников

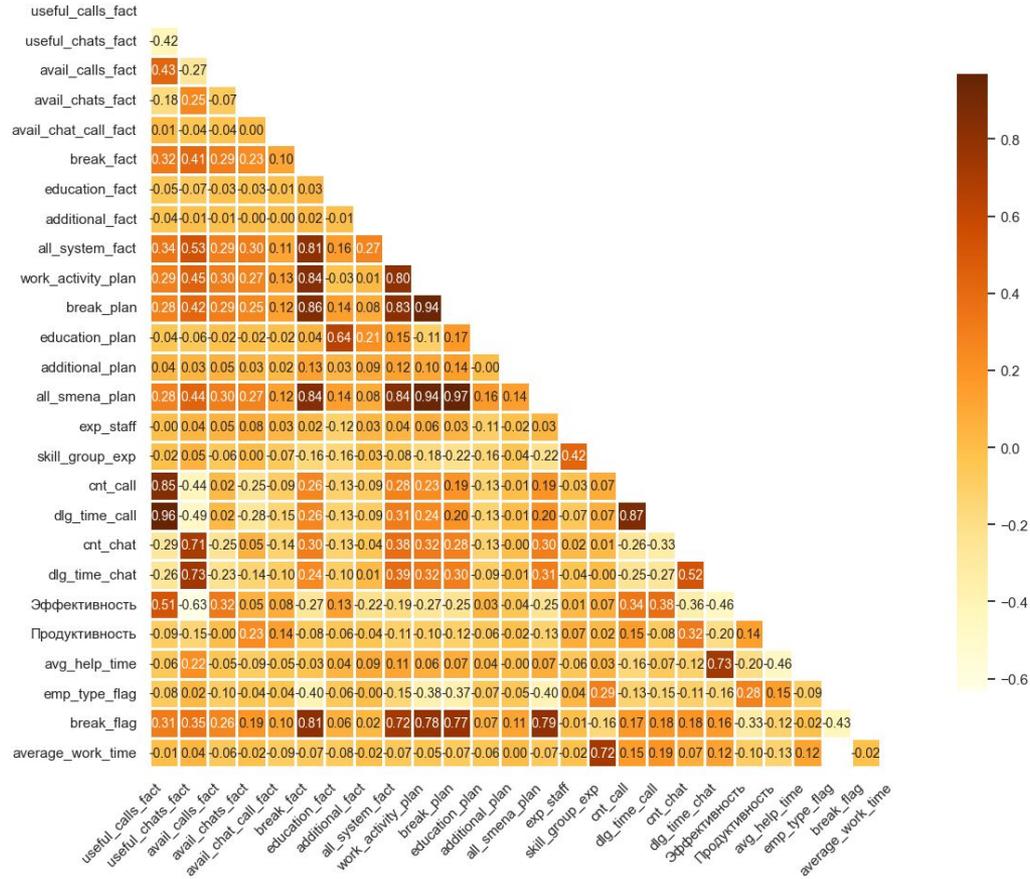


Корреляция опыта с доступностью: 0.58
Корреляция опыта с временем обучения: -0.53

■ Доступность чата/звонков — Время обучения

На данном графике видно, что люди в начале обучения практически не работают на коммуникациях с абонентами, а по мере обучения и роста стажа они понемногу вливаются в рабочий процесс.

Матрица корреляции числовых параметров



Переменные

calendar_dt — дата;

- id_employee — уникальный идентификатор сотрудника;
- useful_calls_fact — время (секунды), когда сотрудник обрабатывал звонки клиентов;
- useful_chats_fact — время (секунды), когда сотрудник отвечал клиентам в чатах.2 / 5

Время ожидания (когда сотрудник доступен для помощи, но обращения не поступают):

- avail_calls_fact — время (секунды), когда сотрудник был доступен для звонков, но никто не звонил;
- avail_chats_fact — время (секунды), когда сотрудник был доступен для чатов, но никто не писал;
- avail_chat_call_fact — время (секунды), когда сотрудник был доступен и для звонков, и для чатов, но коммуникаций не было.

Дополнительные активности в течение дня:

- break_fact — сколько времени (секунд) сотрудник был на перерыве (перерывах);

переменные

- `education_fact` — сколько времени (секунд) сотрудник провел на обучении;
- `additional_fact` — сколько времени сотрудник участвовал в других активностях (например, встречах с руководителем);
- `all_system_fact` — общее время (секунды), которое сотрудник провел в системе.

Тип трудоустройства:

- `emp_type` — форма работы сотрудника:
 - ТК — работает по сменам (обычный трудовой договор);
 - ГПД — работает тогда, когда ему удобно (индивидуальный предприниматель);
 - ПКЦ — работает через партнерский колл-центр (банк заранее заказывает нужное количество сотрудников на определенный день, а партнеры обеспечивают их наличие).

Эти показатели касаются сотрудников с трудовым договором (ТК):

- `work_activity_plan` — запланированное время (секунды) для работы с клиентами;
- `break_plan` — запланированное время (секунды) на перерывы;
- `education_plan` — запланированное время (секунды) на обучение;
- `additional_plan` — запланированное время (секунды) на дополнительные активности;
- `all_shifts_plan` — общее запланированное время (секунды) смены;
- `grafik` — рабочий график сотрудника (только для ТК).

Информация о сотруднике

Ниже — сущности, которые представляют собой иерархию:

- `business_line` — бизнес-направление, в котором работает сотрудник. Верхнеуровневая сущность. Включает в себя несколько направлений;
- `partnarietie` — конкретное подразделение в бизнес-направлении. Уровень ниже. Включает в себя несколько групп прогнозирования;
- `grid_group` — группа прогнозирования (внутренний термин колл-центра).

Включает в себя несколько `skill_group`;

- `skill_group` (скилл группа — какие навыки есть у сотрудника и какие вопросы он может решать (только для сотрудников, кто устроен по трудовому договору, ТК);
- `planning_group_nm` — группа планирования (аналог скилл-группы) для сотрудников ГПД и ПКЦ.3 / 5

Опыт и стаж:

- `exp_staff` — сколько месяцев сотрудник работает в компании;
- `skill_group_exp` — сколько месяцев сотрудник работает в своей текущей скилл-группе или группе планирования (в зависимости от типа трудоустройства);
- `hire_dt` — дата найма сотрудника.

Количество обращений и затраченное время:

Плотность распределения продуктивности и опыта

