

Тебе срочно нужен кредит!

Команда
Garlic Toasts

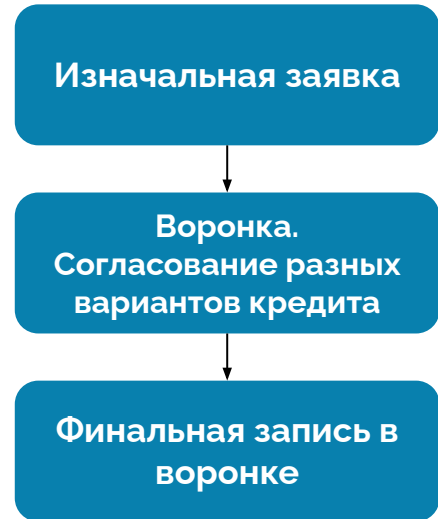


Структура данных

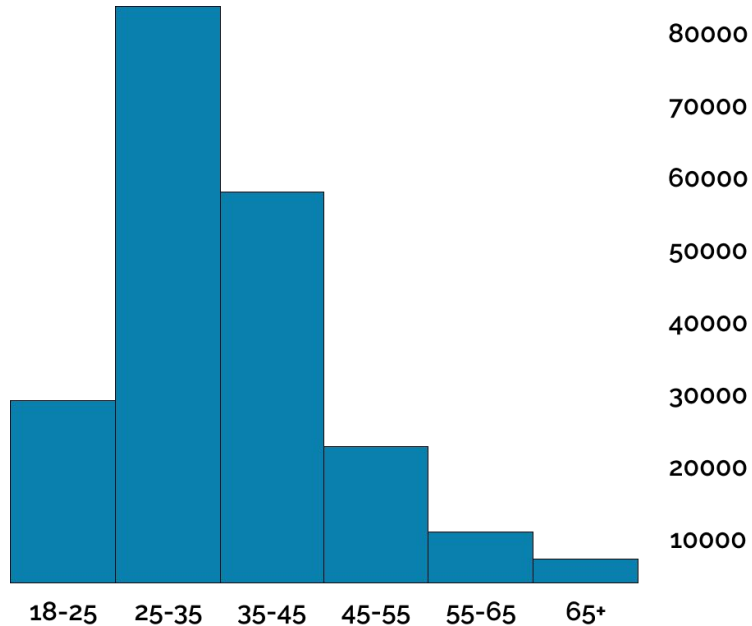
1. Заявки клиентов банка Тинькофф на кредитование
2. Данные за период **март 2021 - ноябрь 2022**
3. **29** столбцов
4. **226 тыс.** согласований, **211 тыс.** заявок, **200 тыс.** уникальных клиентов
5. Информация о поле, возрасте и финансовом положении заявителей на кредитные продукты

Согласования → Заявки

1. Переход от базы данных согласований к БД заявок
2. За результат заявки считается последний этап согласования с клиентом (финальная запись из воронки)
3. **226 тыс.** согласований → **211 тыс.** заявок
4. Конверсия **15%** (дошли до оформления кредита)



Возраст заявителей



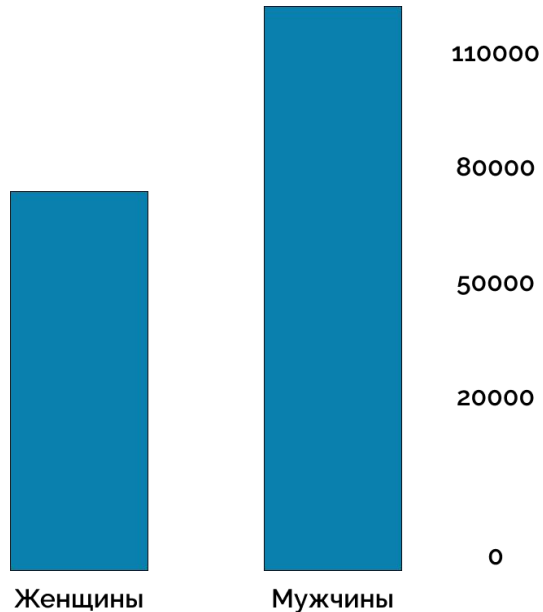
Когортное распределение возрастов заявителей

Очистка данных (8 случаев):

- Значения < 18
- Пустые поля

Средний возраст: 36

Пол клиентов



Гендерное распределение клиентов

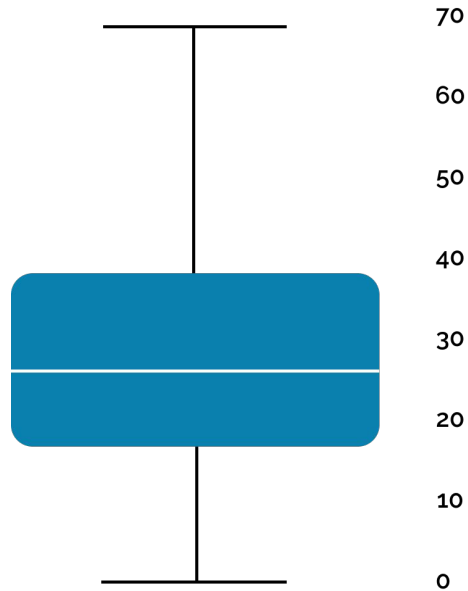
Очистка данных (46 случаев):

- Пустые поля

Мужчин: 60% (114 тыс.)

Женщин: 40% (76 тыс.)

Зарплата клиентов



Распределение дохода клиентов (y.e.)

Очистка данных (6449 случаев):

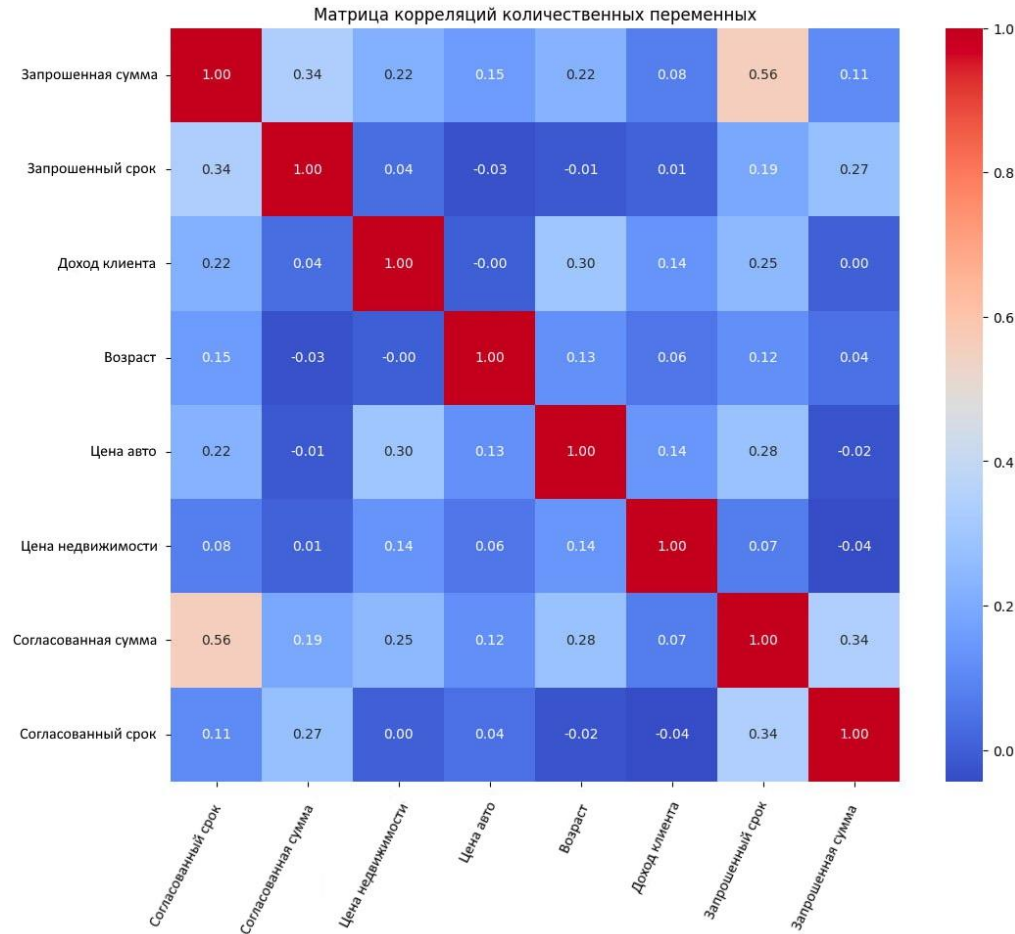
- Пустые поля
- $Q \pm 1.5 \text{ IQR}$

Медианная зарплата: 25 y.e.

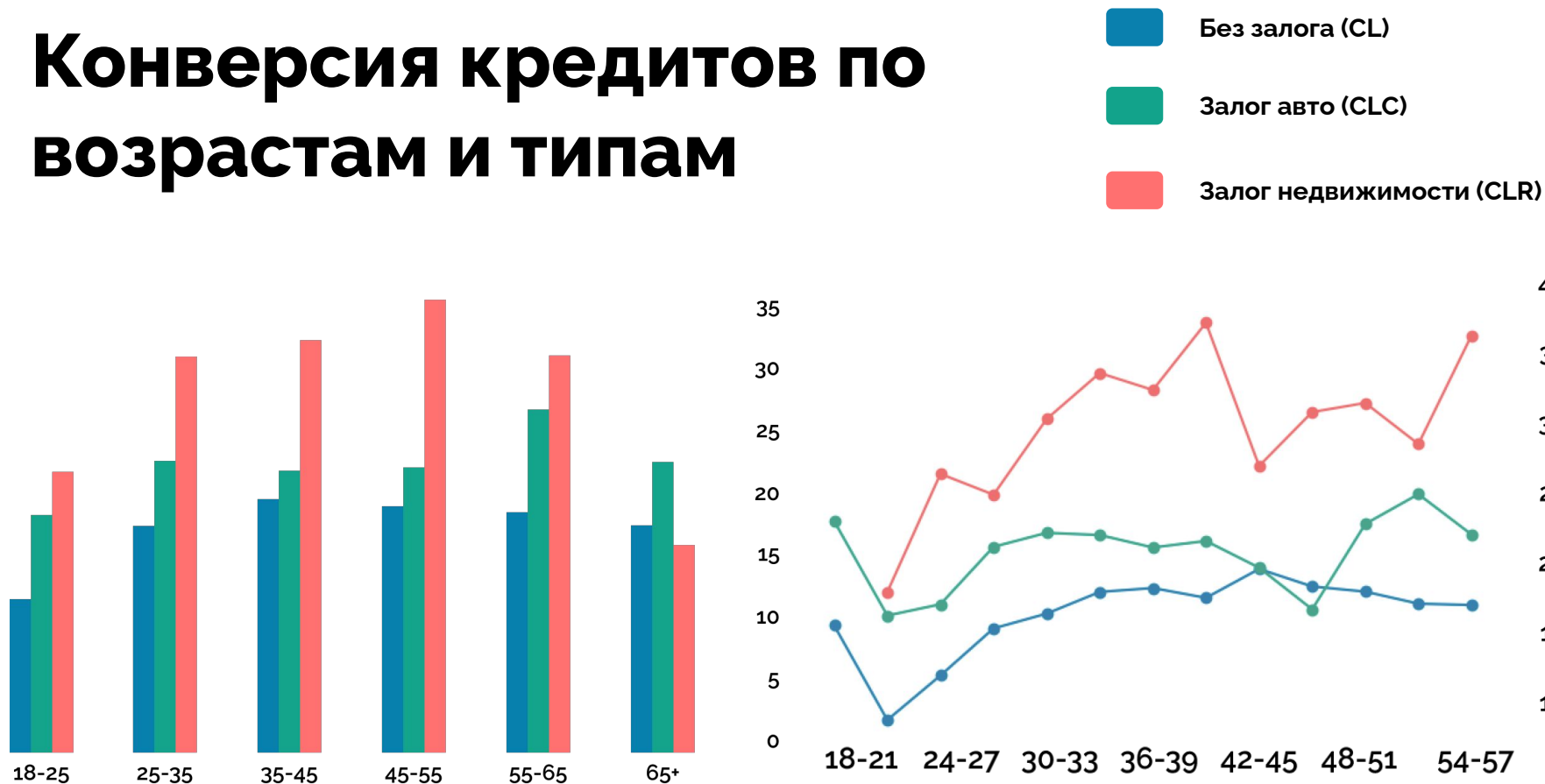
Итого: ~206 тыс. заявок после очистки

Корреляция

Значимые корреляции не обнаружены,
требуется более глубокий анализ



Конверсия кредитов по возрастам и типам



Исследовательский вопрос

Какие характеристики клиента определяют, какой вид кредитования для него наиболее актуален?



Гипотеза

Чем старше человек, тем выше вероятность, что он возьмёт кредит с залогом недвижимости

Только среди трудоспособного населения (18 - 55 лет)

Механизм



Нулевая гипотеза

Среднее значение вероятности
взятия кредита **равно** в каждой
возрастной когорте

Альтернативная гипотеза

Среднее значение вероятности
взятия кредита **не равно** в
каждой возрастной когорте

Тест Краскела-Уоллиса (не требует норм.
распределения)

p-value: ~0.4

Нулевая гипотеза **подтверждается**

Разделение по полу

Мужчины

p-value: ~0.6

Нулевая гипотеза
подтверждается

Женщины

p-value: ~0.4

Нулевая гипотеза
подтверждается

Разделение по наличию авто

Есть авто

p-value: ~0.8

Нулевая гипотеза
подтверждается

Нет авто

p-value: ~0.4

Нулевая гипотеза
подтверждается

Разделение по наличию кредитной карты

Есть кредитка

p-value: ~0.7

Нулевая гипотеза
подтверждается

Нет кредитки

p-value: ~0.7

Нулевая гипотеза
подтверждается

Интерпретация

Вывод

Гипотеза **опровергается**.

Взятие кредита с залогом недвижимости не зависит от возраста

Устойчивость

Проведена проверка устойчивости в подгруппах по полу, наличию авто и кредитной карты

Ограничения

Данные за определенный период

Только трудоспособное население

Ограниченная выборка

Ограничения и перспективы

Ограничения

1. Период **март 2021 - ноябрь 2022**
2. Клиенты Тинькофф банка
3. Рассматривается только трудоспособное население. Мало данных по старшему поколению.

Перспективы

Провести исследование среди клиентов старше трудоспособного возраста, чтобы выявить факторы взятия кредита в этой возрастной группе

Policy implication

1. Не таргетируем кредит с залогом недвижимости по какой-то возрастной группе
2. Банку следует развивать инструменты распознавания наличия и стоимости недвижимости у заемщика для таргетинга, и не опираться на возраст.



[Материалы исследования](#)