



Selling Pandas



1. Иванов Герман
2. Хлопотных Андрей
3. Михайлов Андрей
4. Шершнеv Михаил
5. Лоцинин Николай
6. Кром Илья

Алё! Ну как там с деньгами?



Структура данных

Информация

Датасет о перекрестных продажах
кредитов, клиентам Тинькофф

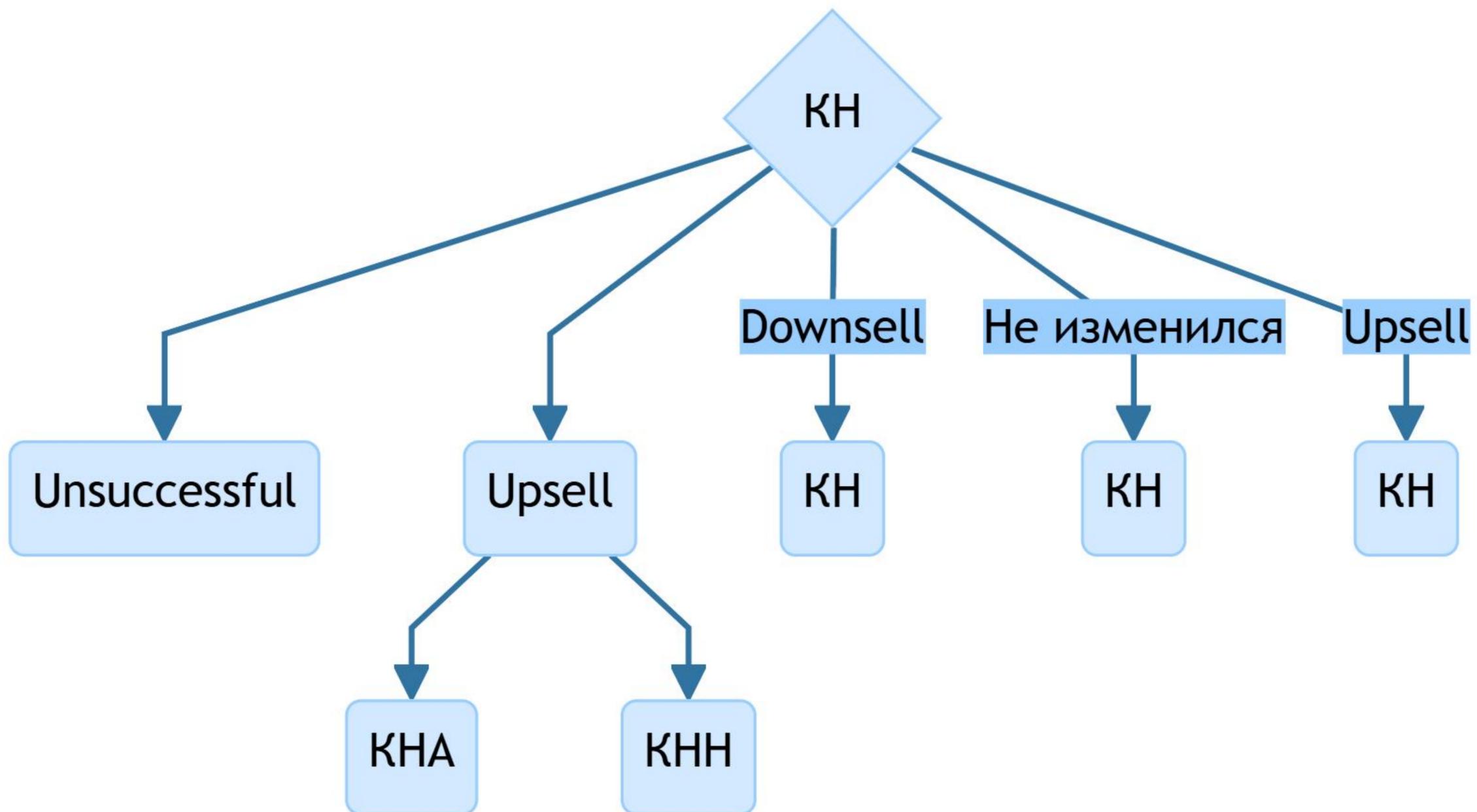
226 тыс. записей, 211 тыс. заявок

Данные с марта 2021 года по ноябрь
2022





Предварительный анализ





Очистка от выбросов

Очистка

По IQR

226 тыс. строк



183 тыс. строк





Матрица корреляций

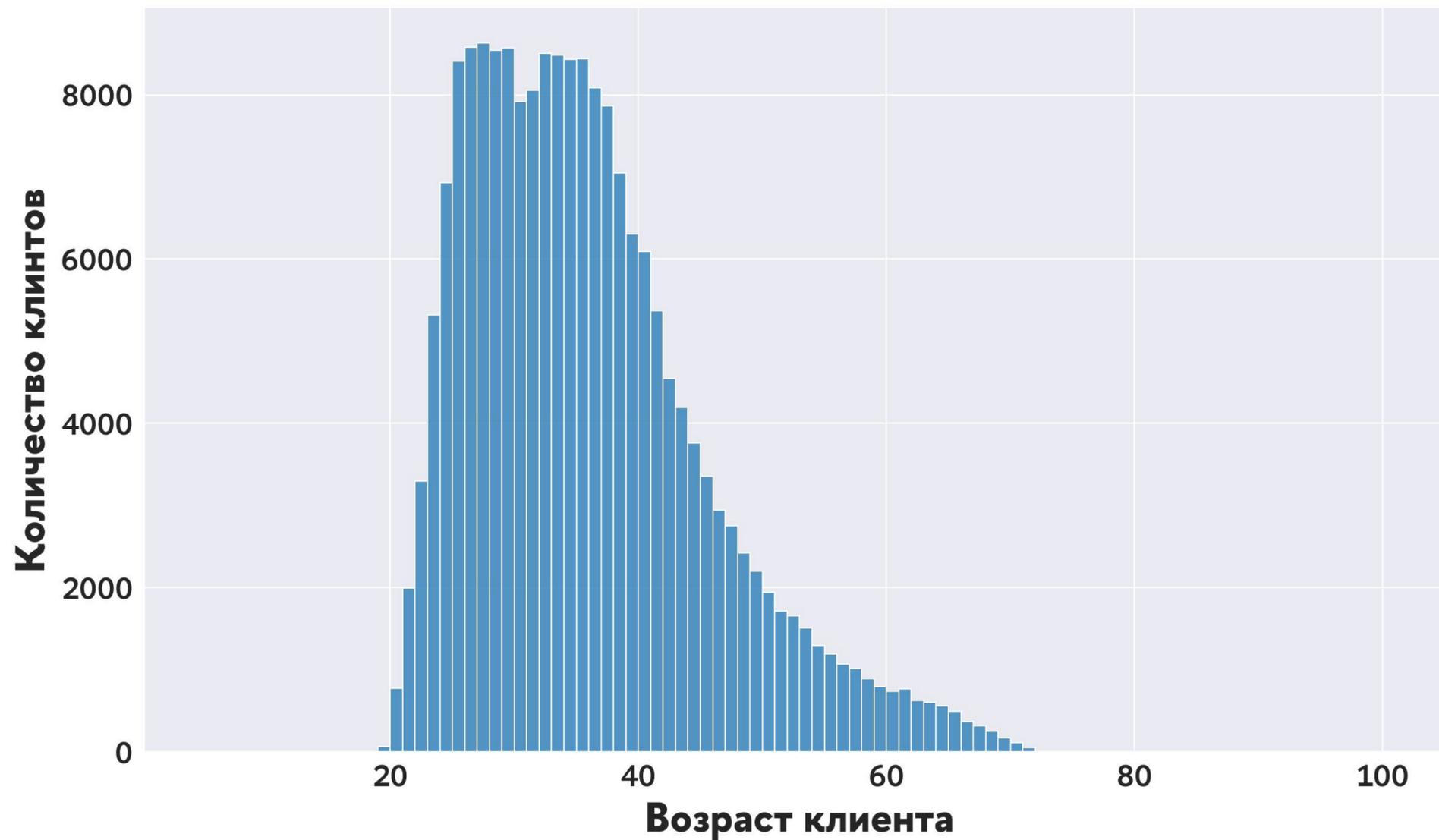
Запрошенная сумма	1	0.096	0.11	0.14	0.2	0.051	0.57	0.084	-0.082
Запрошенный срок	0.096	1	0.009	-0.011	-0.007	0.015	0.16	0.29	-0.062
Доход в мес.	0.11	0.009	1	-0.003	0.18	0.082	0.11	-0.004	-0.022
Возраст	0.14	-0.011	-0.003	1	0.12	0.017	0.12	0.032	-0.08
Цена машины	0.2	-0.007	0.18	0.12	1	0.097	0.3	-0.04	-0.03
Цена недвижимости	0.051	0.015	0.082	0.017	0.097	1	0.028	-0.048	-0.011
Согласованная сумма	0.57	0.16	0.11	0.12	0.3	0.028	1	0.28	-0.24
Согласованный срок	0.084	0.29	-0.004	0.032	-0.04	-0.048	0.28	1	-0.28
Согласованная ставка	-0.082	-0.062	-0.022	-0.08	-0.03	-0.011	-0.24	-0.28	1
	Запрошенная сумма	Запрошенный срок	Доход в мес.	Возраст	Цена машины	Цена недвижимости	Согласованная сумма	Согласованный срок	Согласованная ставка





Предварительный анализ

Распределение возрастов



Summary statistics for the age distribution:

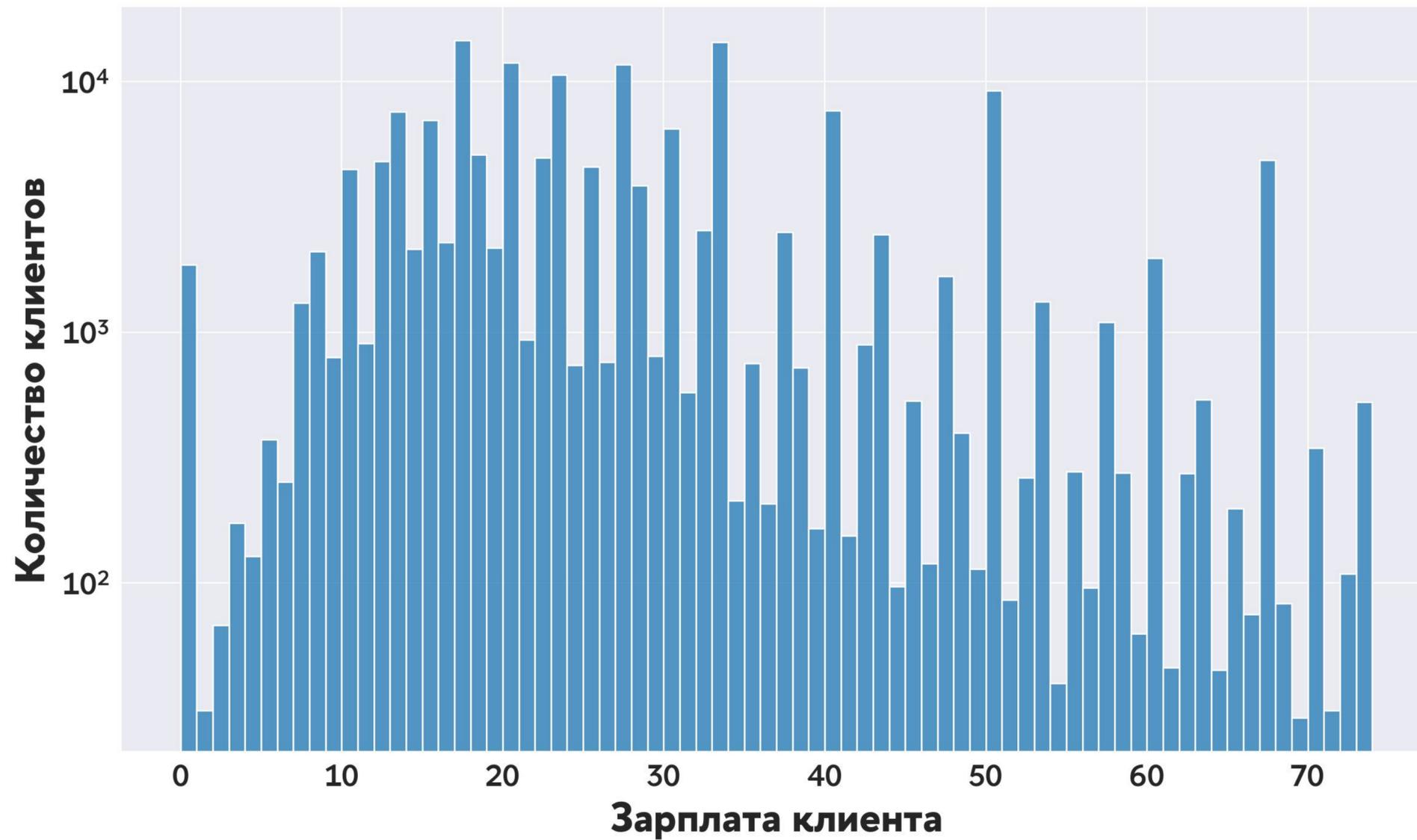
- Среднее: 35.5
- Медиана: 34.0
- Стандартное отклонение: 9.6





Предварительный анализ

Распределение зарплат



Summary statistics for the salary distribution:

- Среднее: 27.6
- Медиана: 24.0
- Стандартное отклонение: 14.7





Обработка данных

Добавление переменных

is_upsell

Фильтрация

Оставляем только тех, кто:

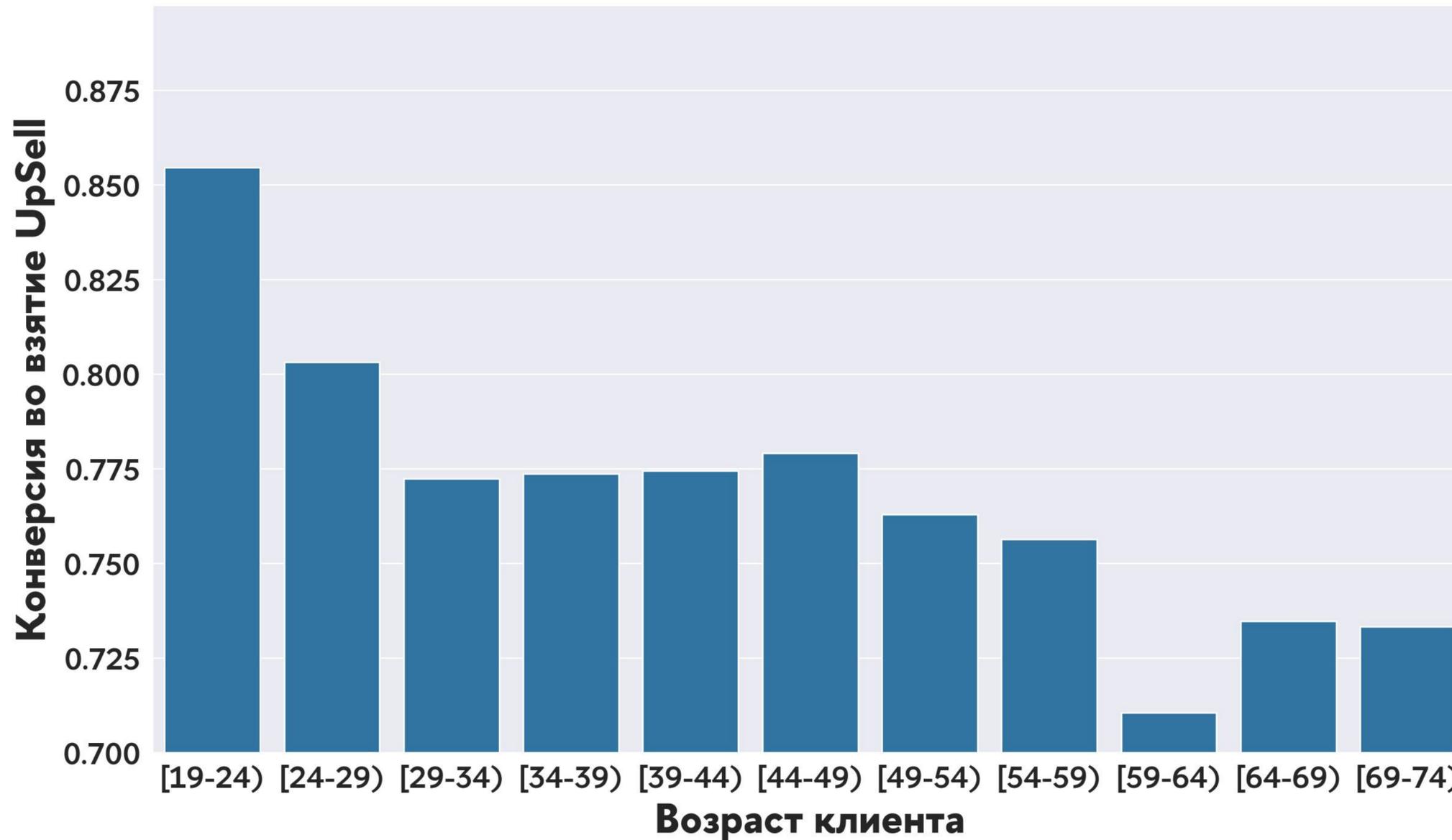
- запросили кредит наличными
- возраст от 18 лет
- был согласован продукт КНА, КНН или КН





Предварительный анализ

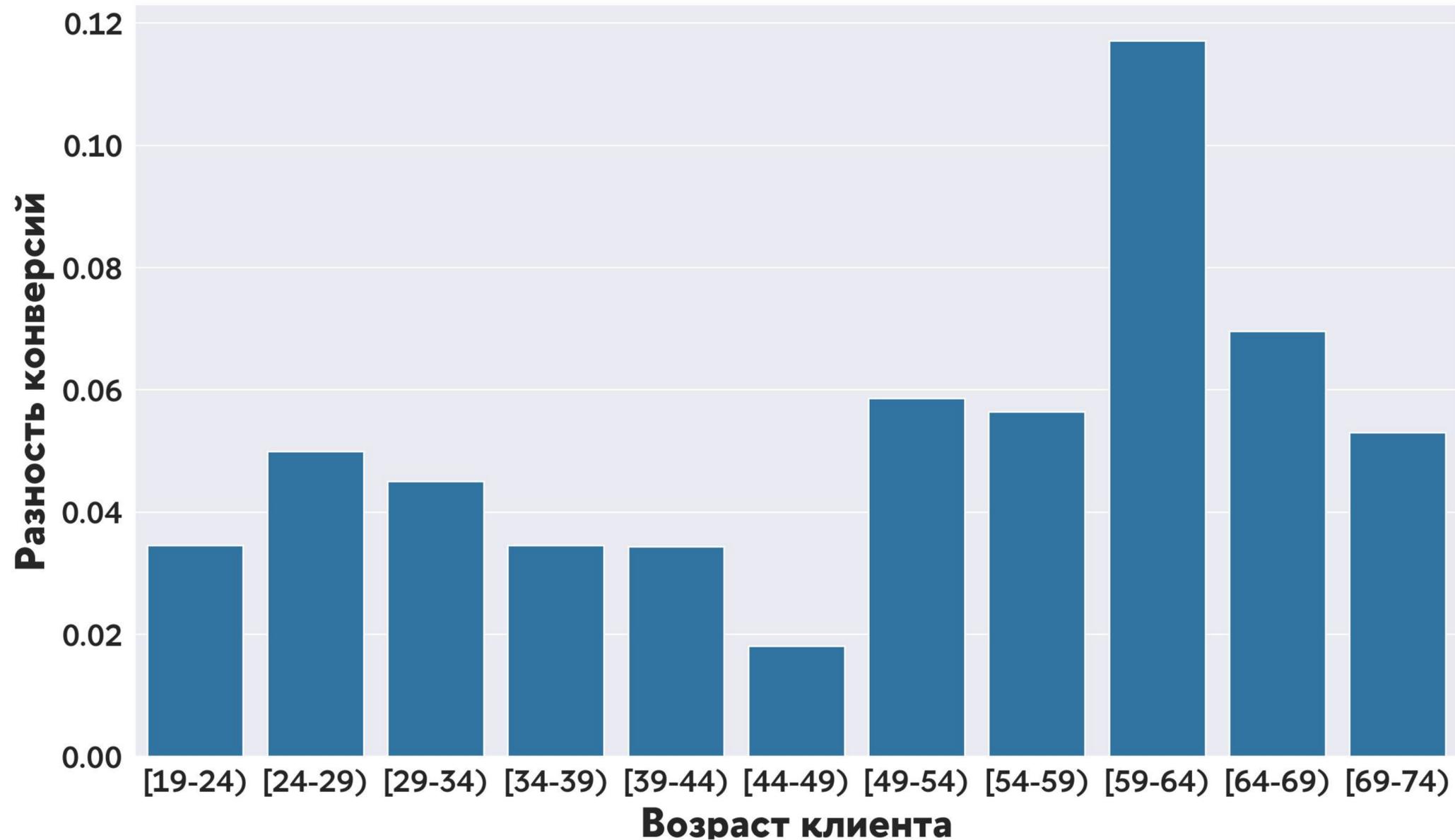
Распределение конверсии во взятие UpSell по возрастам





Предварительный анализ

Распределение разности конверсий по возрастам



Мы посчитали конверсию в принятие кредита для тех, кому предлагался Upsell (y_1) и для тех, кому не предлагался Upsell (y_2).

По оси y находится разность $y_2 - y_1$





Исследовательский вопрос и гипотеза

Исследовательский вопрос

Какие люди чаще **соглашаются** на Upsell предложения, а какие **отказываются**?

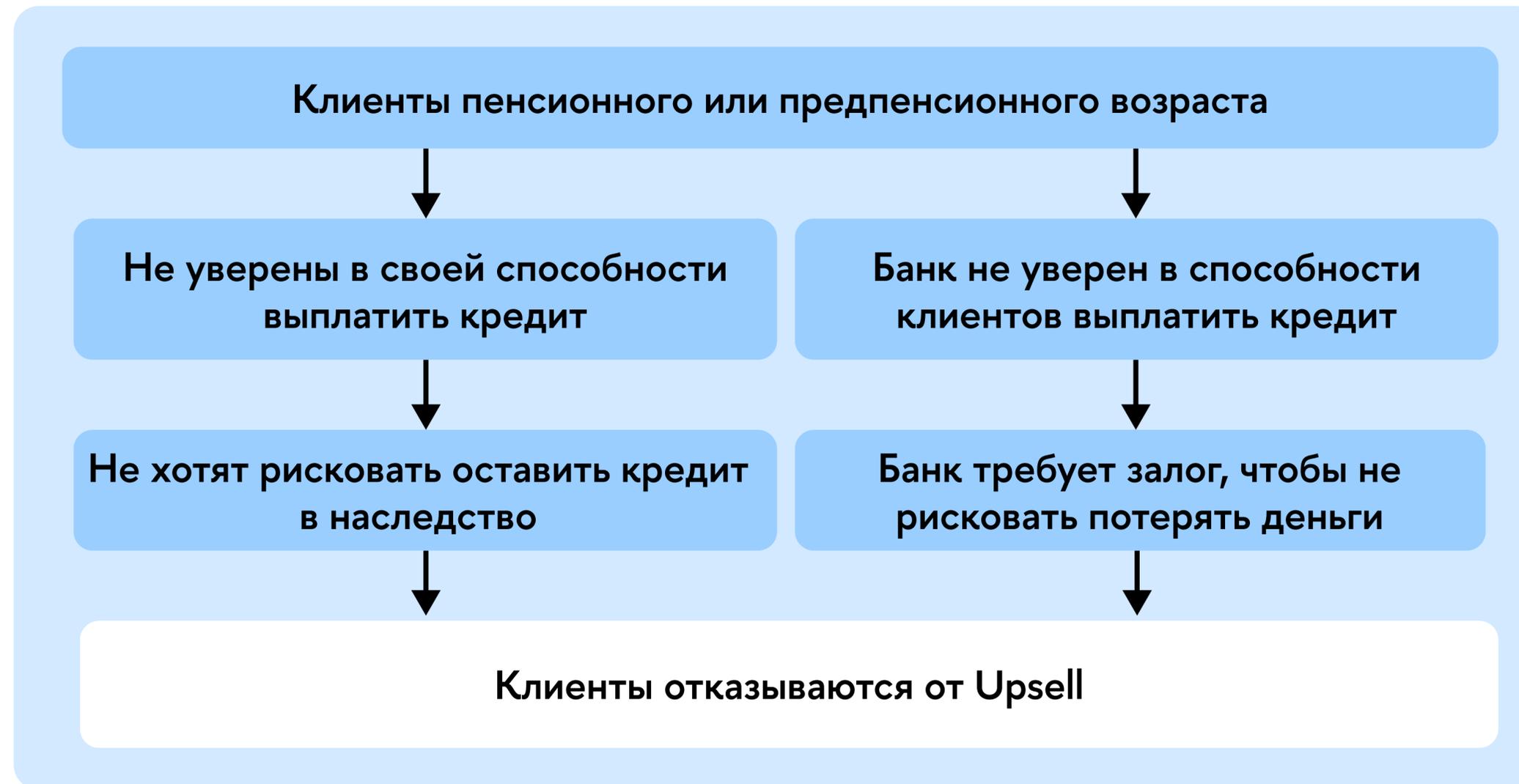
Гипотеза

Люди предпенсионного и пенсионного возрастов **менее** склонны к **принятию** Upsell предложения.





Механизм





Математическая модель

Gr1 < 50 лет
Gr2 ≥ 50 лет

Кредит **был** активирован Кредит **не был** активирован

Upsell был предложен			→ <input checked="" type="checkbox"/> Значимый: Gr1 > Gr2 p-value: 0,004
Upsell не был предложен			





Математическая модель

Gr1 < 50 лет
Gr2 ≥ 50 лет

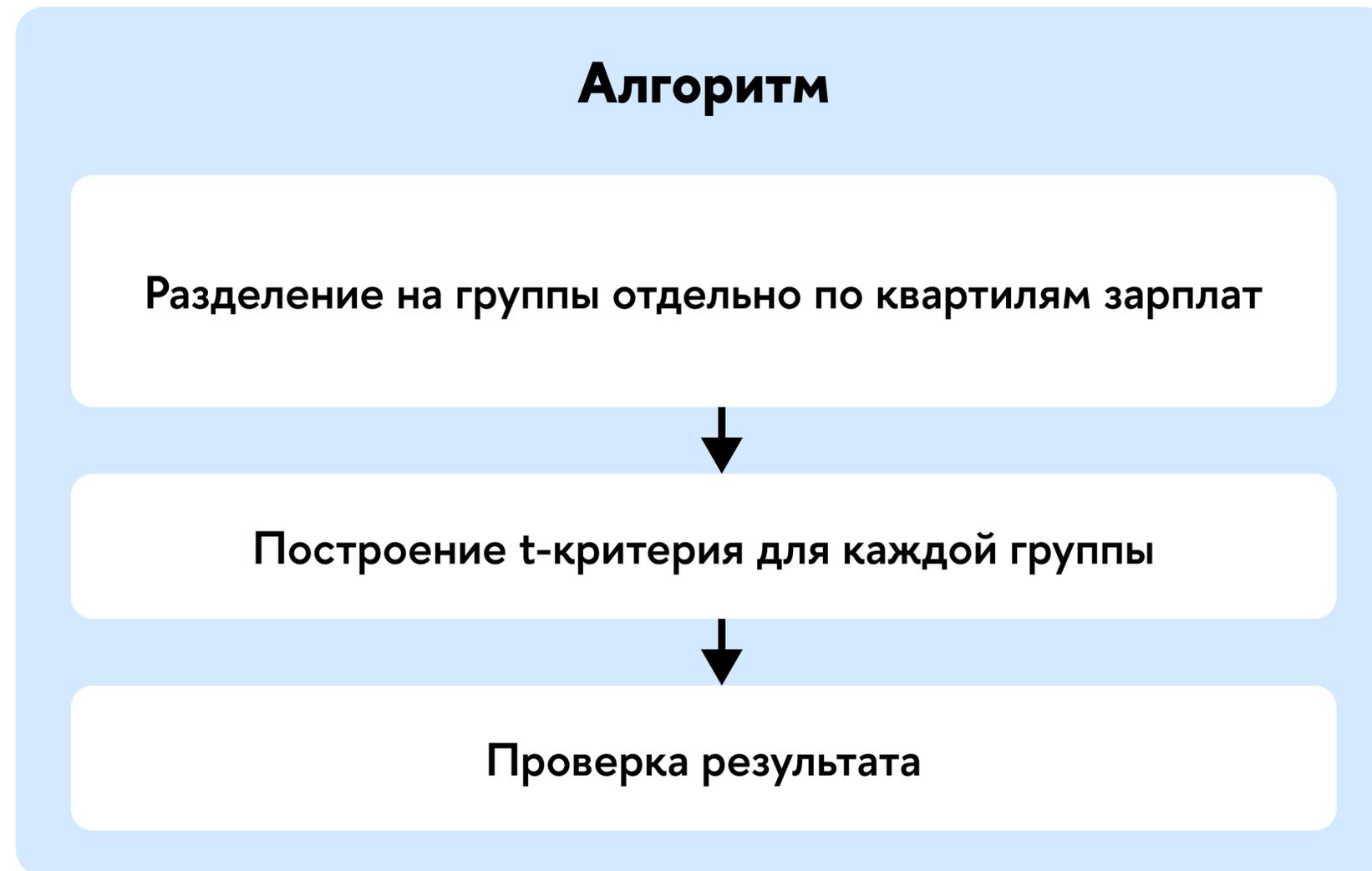
	Кредит был активирован	Кредит не был активирован
Upsell был предложен		
Upsell не был предложен		

→ **X** Не значимый: Gr1 ≈ Gr2
p-value: 0,63





Проверка модели на устойчивость





Разделение на группы по зарплатным квартилям

Результат

Группа	Объем	p-value	Значимость:
Первый квартиль	11551	<0.01	✓
Второй квартиль	7395	<0.01	✓
Третий квартиль	9595	<0.01	✓
Четвёртый квартиль	9301	<0.01	✓





Интерпретация



Гипотеза подтвердилась

Возрастные люди менее склонны к тому,
чтобы брать кредит

Произведена проверка устойчивости



Разделение по квантилям

зарплаты



Перспективы и ограничения

Ограничения

Применимо для
клиентов Тинькофф
Банка

Ограниченные временные
рамки в наборе данных

Перспективы

Учитывать причину отказа для более качественного
предсказания

Учитывать, ушел ли клиент в дефолт или в другой банк

Добавить новые кредитные продукты

Исследовать другие признаки и их комбинации





Policy implication

Меньше предлагать Upsell более возрастному контингенту, т.к. они менее склонны принимать его



Экономия затрат банка на колл-центр, а также экономия времени его сотрудников

Предлагать финансовые курсы для людей старшего возраста, чтобы они не боялись брать Upsell





Спасибо за внимание!

Остались ли вопросы?