



# КЛИЕНТСКАЯ АНАЛИТИКА В РИТЕЙЛЕ

Константинов Александр

Учебный проект



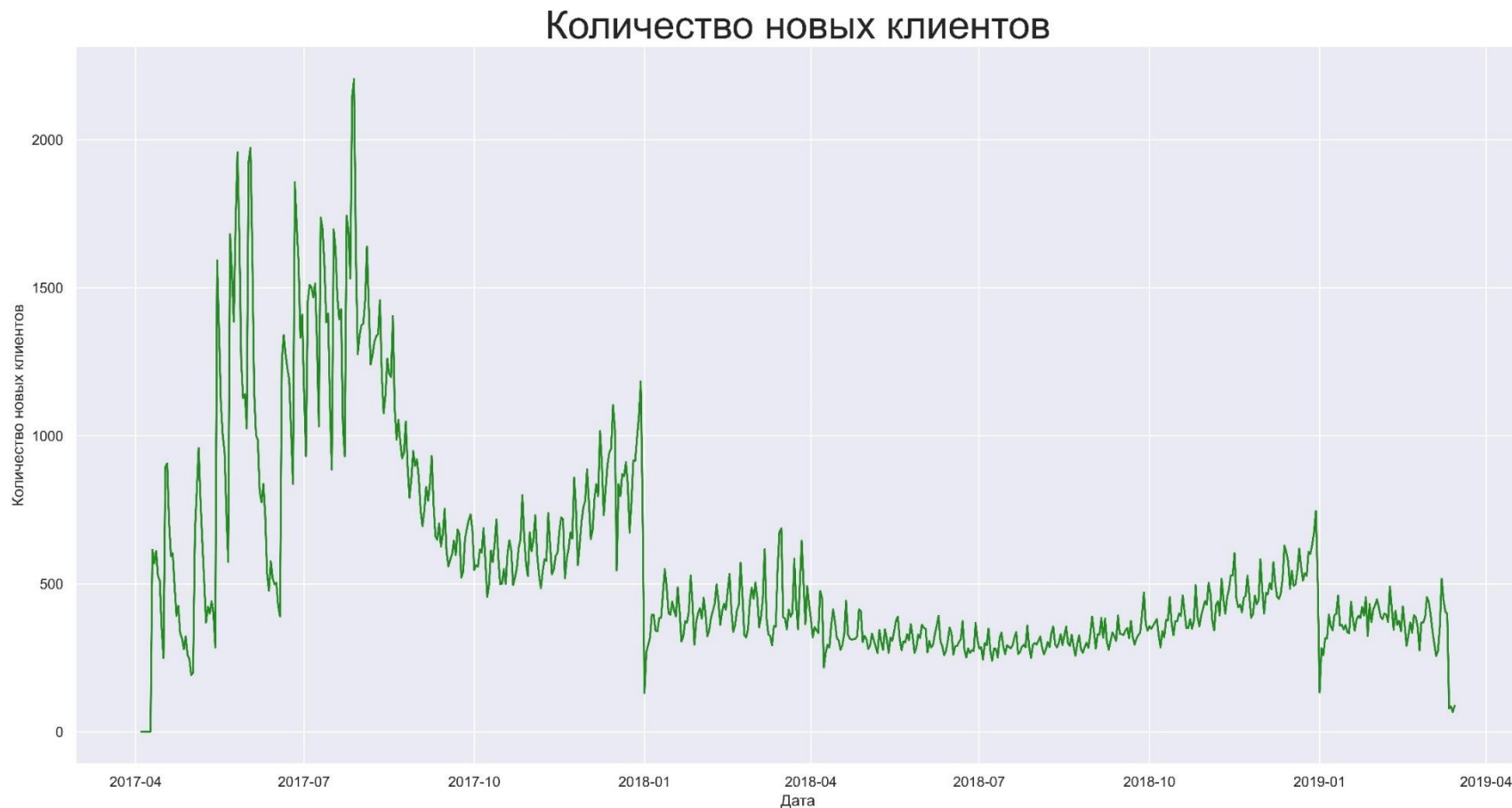
## Общая информация о датасете

- Количество клиентов – **400 162**
- Данные о клиентах, получивших карту лояльности с **04.04.2017 по 15.03.2019**
- Гендерный состав: мужчины – **66 807**, женщины – **147 649**, не указан пол – **185 706**
- Очевидно недостоверные данные о возрасте клиента – **1 145**. Например, значения -937, 119, 115 и т. д.
- Информация о транзакциях: количество транзакций – **8 045 201**, количество строк с покупками – **45 786 568**, период – с **21.11.2018 по 18.03.2019**
- Принимая в расчет, что данные могут быть неполными, но выражающими общие тенденции генеральной совокупности, стоит понимать, что цены актуальны для 2018-2019 гг., поэтому месячные расходы клиентов в районе 3000-5000 рублей реалистичны в силу уровня цен, также некоторые транзакции могут не учитываться, так как клиенты могли не пользоваться картой лояльности при некоторых покупках.

## Содержание

- Разведочный анализ данных
- Сегментация клиентов
- Метрики по новым клиентам

## Разведочный анализ данных (1/5): количество новых клиентов падает



Количество новых клиентов переживает спад с каждым годом.

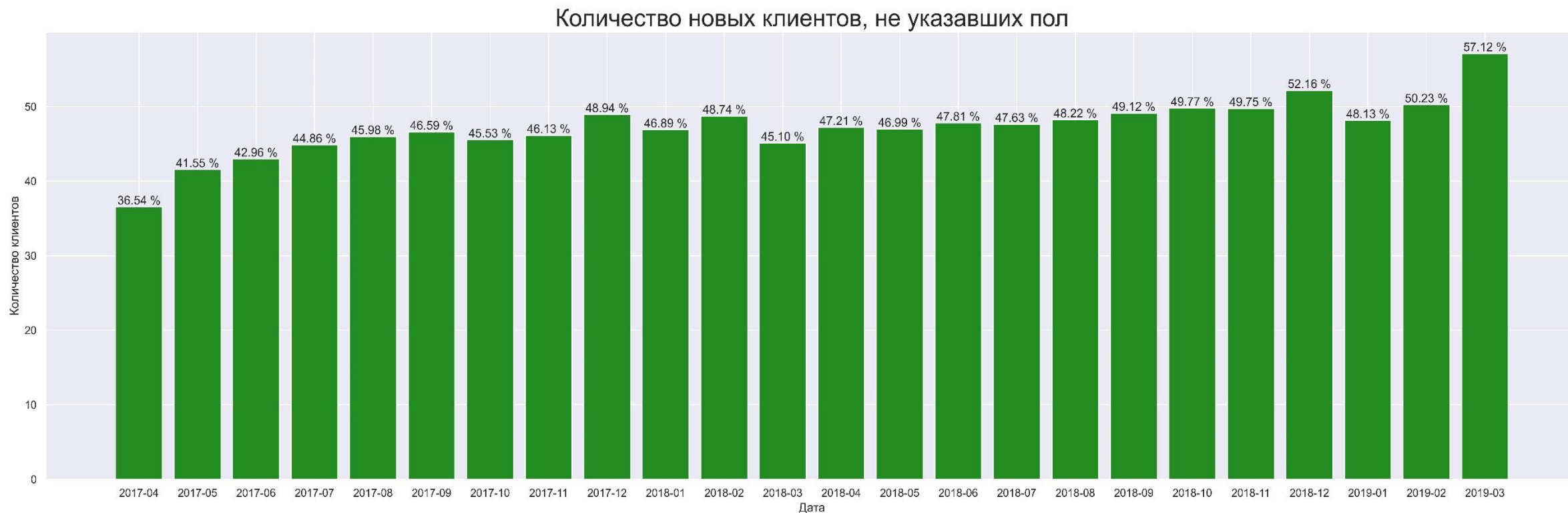
Можно предполагать, что тенденция с падением количества новых клиентов продолжится.

За **2017 год** (с апреля) количество новых клиентов равняется **236 762 чел.**

За **2018 год** количество оформивших карту лояльности составляет **136 999 чел.**

За **2019 год** (с января по март) оформило **26 401 чел.**, что свидетельствует о серьезном спаде.

## Разведочный анализ данных (2/5): новые клиенты чаще скрывают личную информацию



**Доля новых клиентов, которые решили не указывать гендер при регистрации, растет с каждым месяцем.**

Проблема может казаться незначимой, но может затруднять аналитику.

Возможно, что между значимыми для бизнеса метриками и тем, указывает гендер или нет, имеется некоторая значимая связь.

Но ее можно обнаружить только при более полной информации о клиенте.

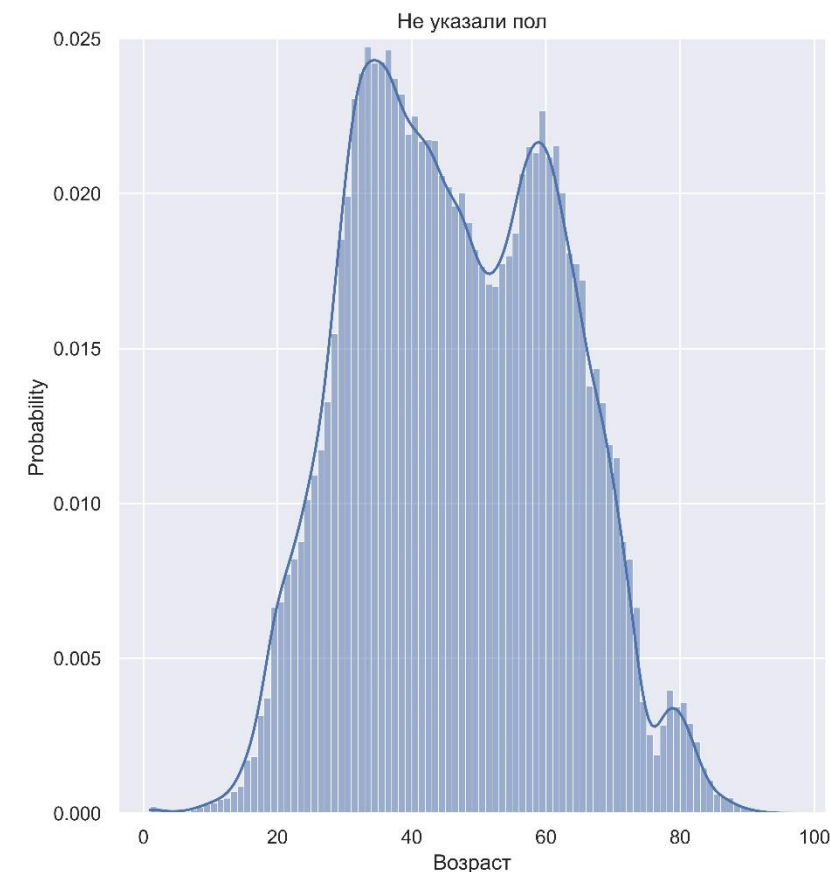
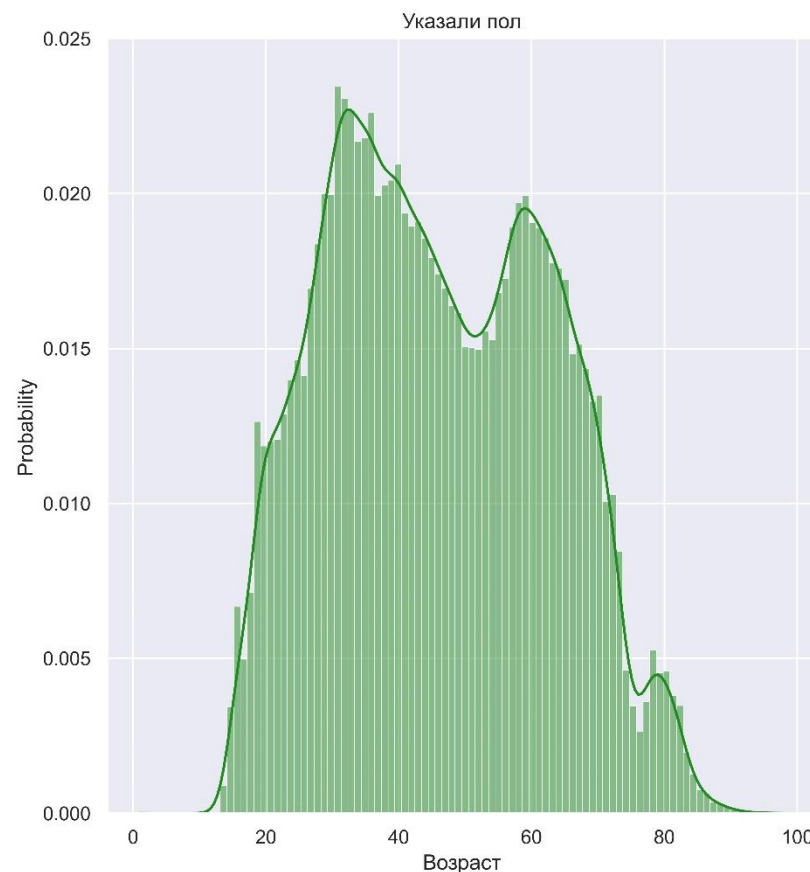
## Разведочный анализ данных (3/5): разница между распределениями возрастов отсутствует

**Разница в возрастном распределении** между теми, кто указал свой гендер, и проигнорировавшими эту возможность **незначительна.**

Однако, можно заметить, что пик в районе 30-40 лет более острый среди тех, кто не указал свой пол при оформлении.

Возможно, наблюдаемые пропуски в данных связаны не с желанием клиентов, а с работой сервиса.

Возрастное распределение

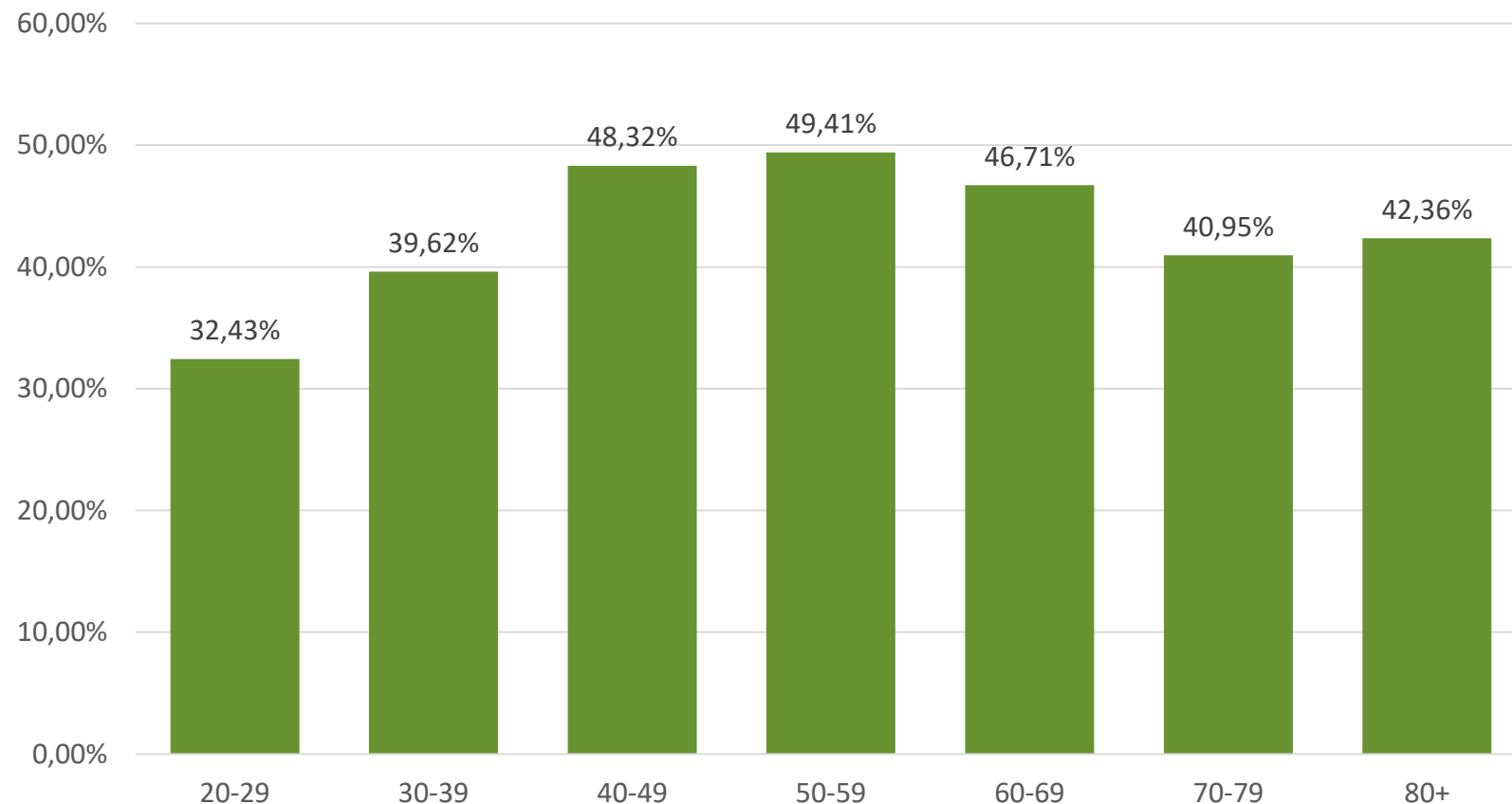


## Разведочный анализ данных (4/5): молодые клиенты чаще указывают свой пол при оформлении

Можем видеть, что **клиенты среднего возраста**, действительно, чаще не указывают свой гендер.

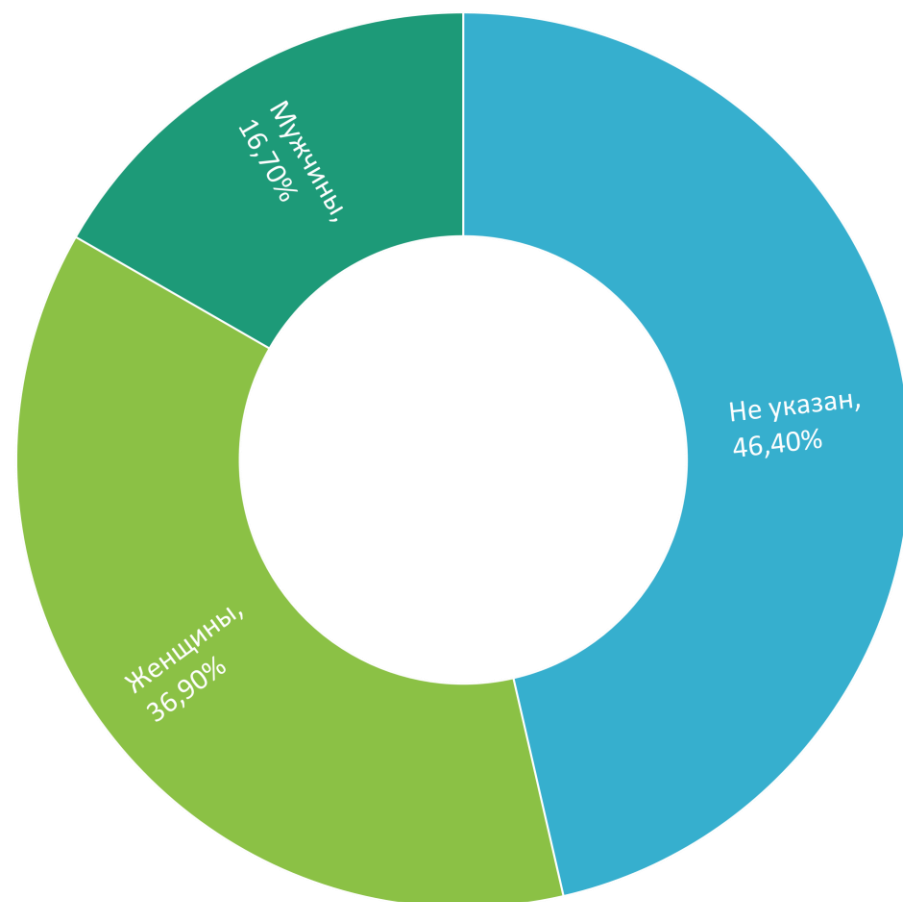
**Разница в 17 процентов** между клиентами до 30 и клиентами между 50 и 59 годами выглядит значимой.

Доля клиентов с неизвестным полом



# Разведочный анализ данных (5/5): доля женщин среди клиентов подавляющая, но имеет тенденцию к снижению

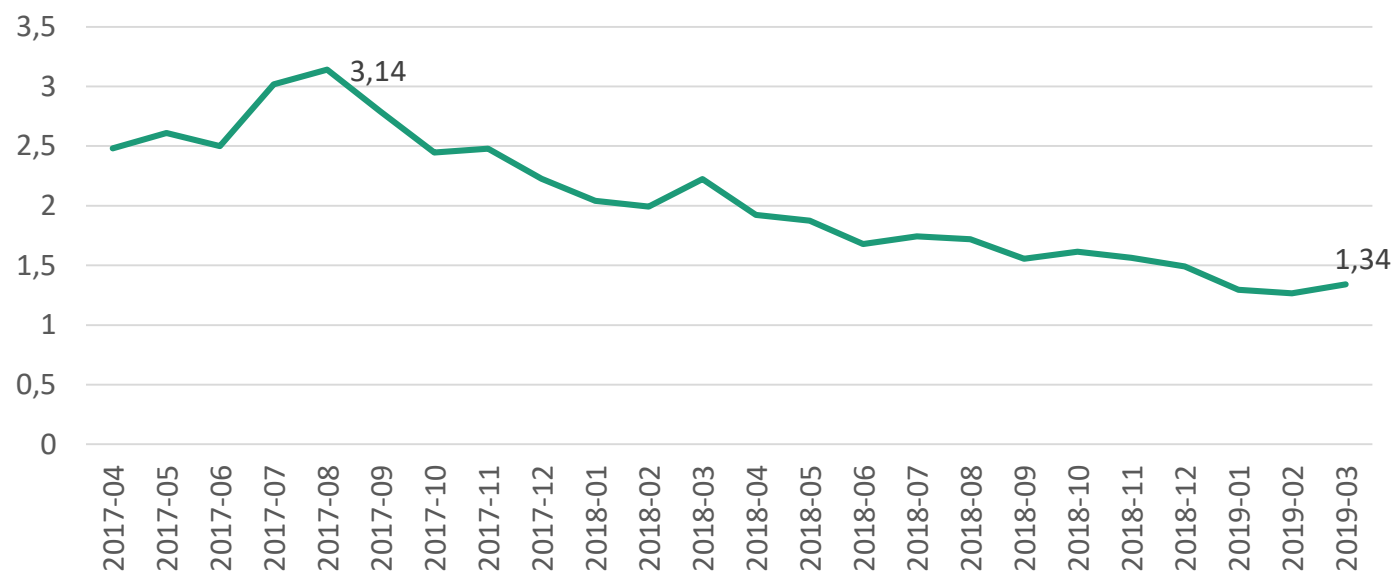
Доля клиентов по гендеру



Среди клиентов с известным полом **женщин в 2,25 раза больше.**

Доля женщин среди новых клиентов больше в любой временной период, но **разрыв заметно снижается**. В последнем месяце их на **34 процента** больше, по сравнению с августом 2017 г., когда их было больше на **214 процентов**.

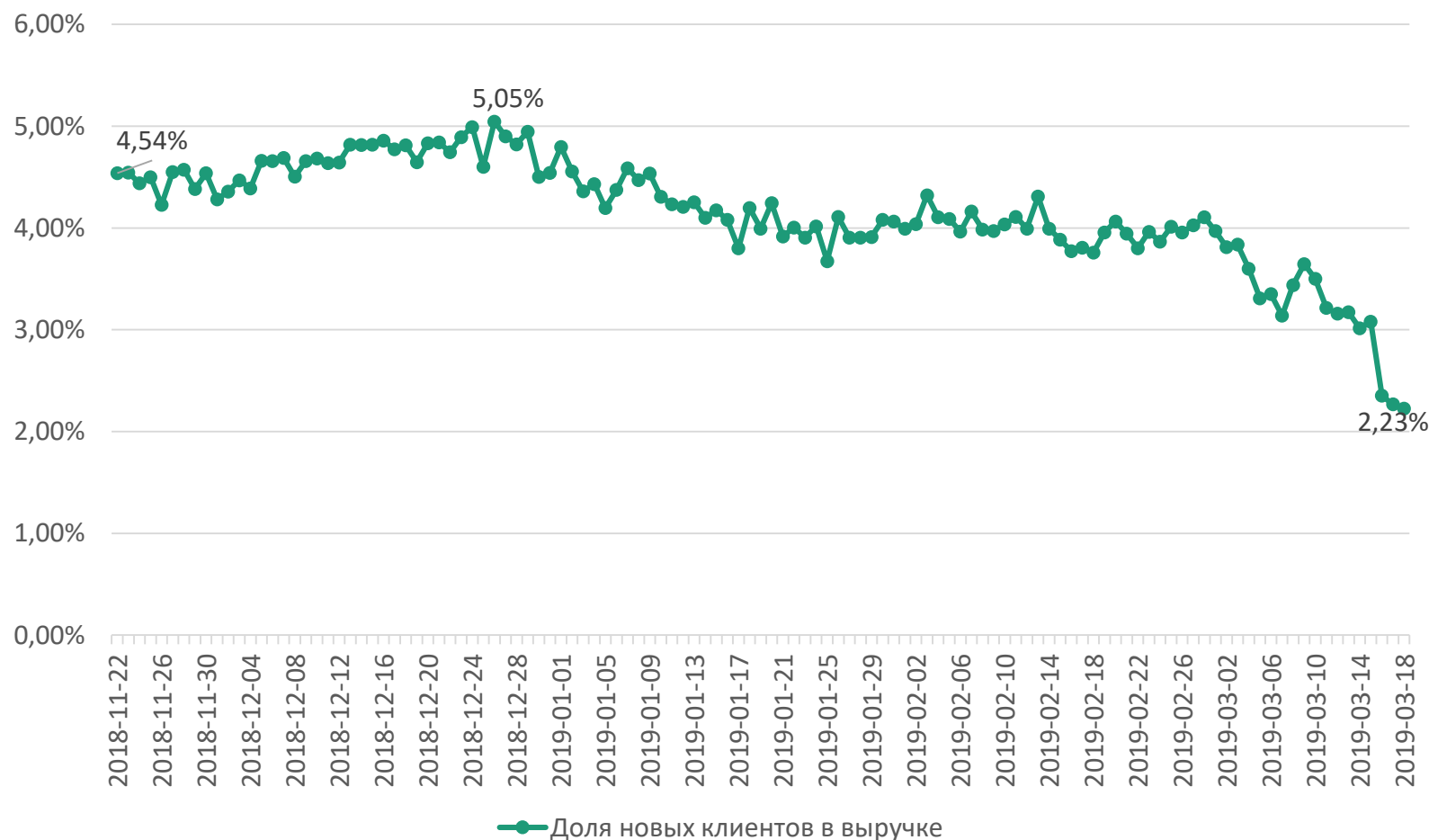
Во сколько раз среди новых клиентов больше женщин, чем мужчин





## Сегментация клиентов (1/4): доля новых клиентов в выручке снижается

Доля новых клиентов в выручке



**Доля новых клиентов** (пользуются картой лояльности первый месяц) **в выручке снижается.**

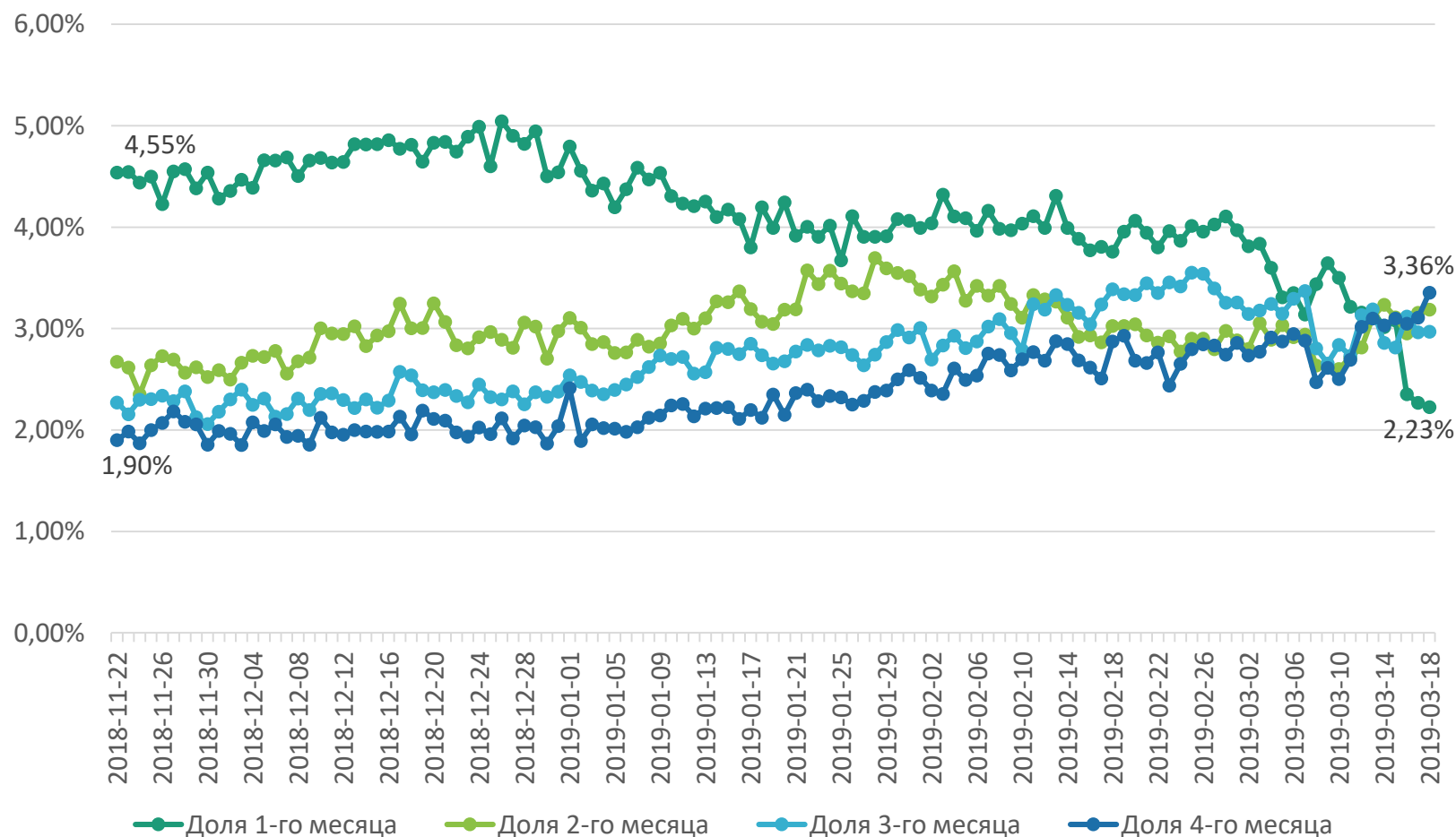
Максимальное значение зафиксировано 26 декабря 2018 г. на уровне **5,05 процентов** от выручки.

Минимальное значение зафиксировано 18 марта 2019 г. на уровне **2,23 процента.**

В марте доля новых клиентов в выручке составила **3,25 процента.**

## Сегментация клиентов (2/4): доля клиентов 2-го, 3-го и 4-го месяцев в выручке растет, обгоняя клиентов 1-ой недели

Доля новых клиентов в выручке



Динамика среди клиентов 2-го, 3-го и 4-го месяцев пользования картой лояльности демонстрирует альтернативную картину (наблюдается рост).

Среди этих клиентов, которых также можно назвать новыми, доля в выручке составляет примерно **3 процента** для каждой из групп.

Таким образом, мы наблюдаем процесс, который можно назвать **равномеризацией долей в выручке** для новых клиентов.

## Сегментация клиентов (3/4): средний чек новых клиентов демонстрирует рост в марте

Средний чек в первый месяц



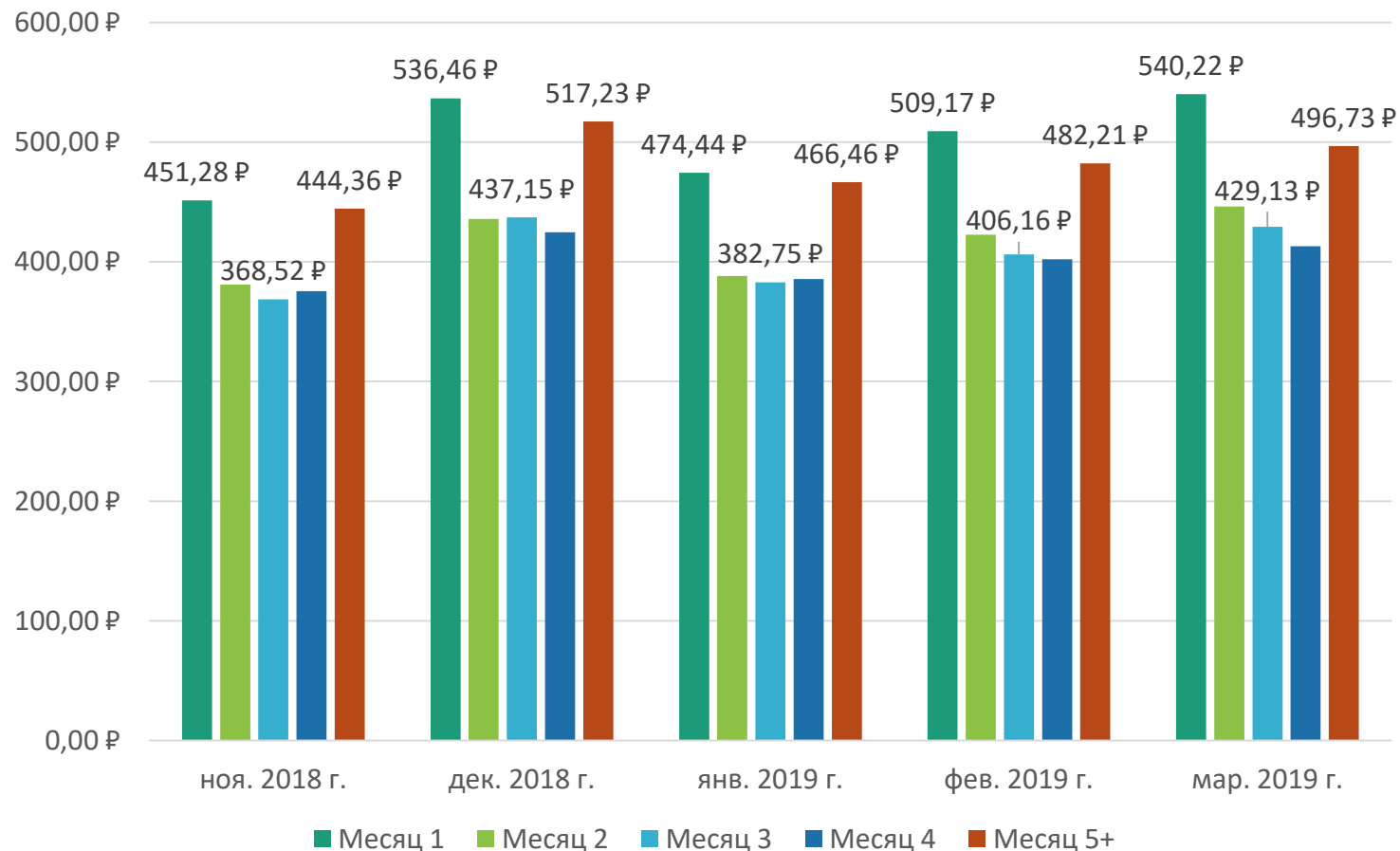
Динамика среди клиентов 2-го, 3-го и 4-го месяцев пользования картой лояльности демонстрирует альтернативную картину (**наблюдается рост**).

Среди этих клиентов, которых также можно назвать новыми, доля в выручке составляет примерно **3 процента** для каждой из групп.

Таким образом, мы наблюдаем процесс, который можно назвать **равномеризацией долей в выручке** для новых клиентов.

## Сегментация клиентов (4/4): средний чек новых клиентов демонстрирует рост в марте

Сравнение средних чеков по месяцам



На диаграмме можно увидеть, что **наибольший средний чек среди клиентов первого месяца использования карты лояльности.**

**Последующие месяцы** демонстрируют спад, который не зависит от временного

**Средний чек** лояльных клиентов (свыше 4 месяцев) **выше**, чем новых клиентов.

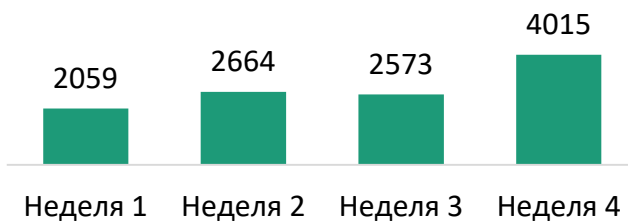
Однако средний чек клиентов первого месяца оказывается выше в каждый месяц.

# Метрики по новым клиентам: отчет за январь 2019 г.

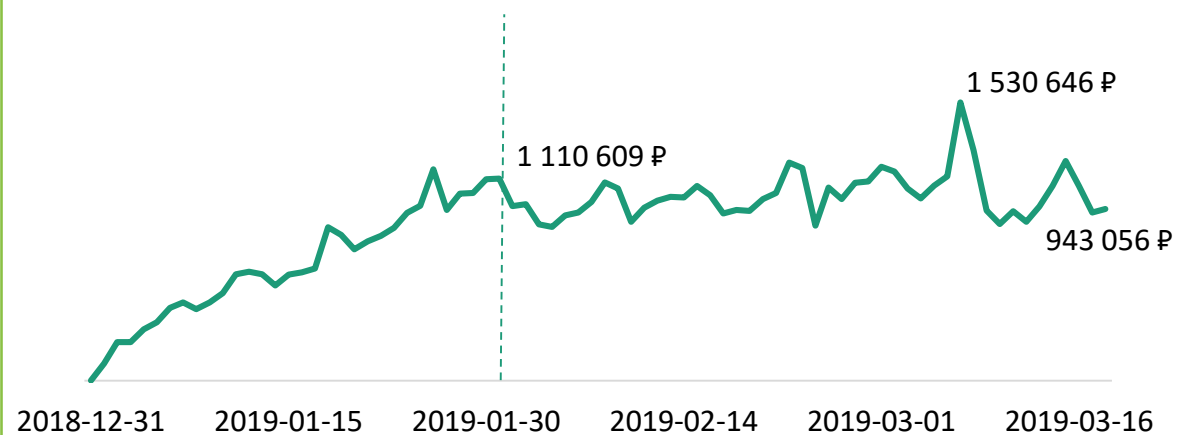
Количество новых клиентов

11 311 чел.

Количество новых клиентов по неделям



Выручка от клиентов, получивших карту в янв. 2019 г.



Средний чек клиента в первый месяц

503,50 руб.

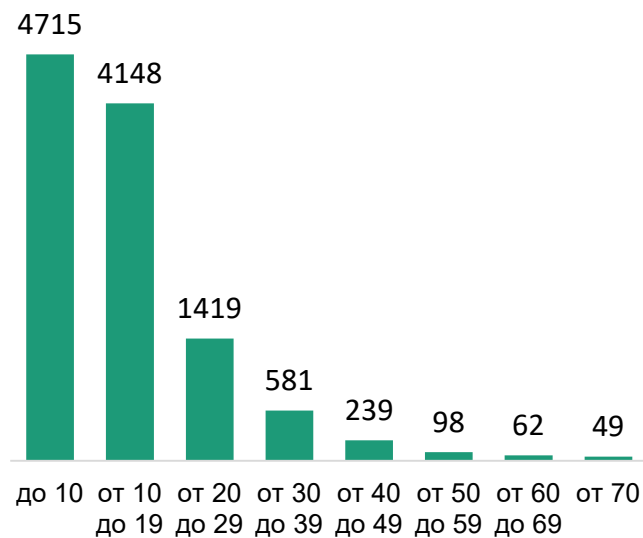
Среднее количество покупок в первый месяц

14,2

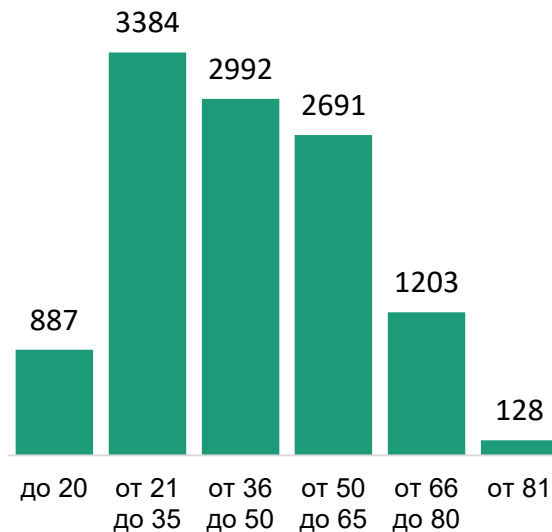
Retention rate для клиентов, в первый месяц

97,2 %

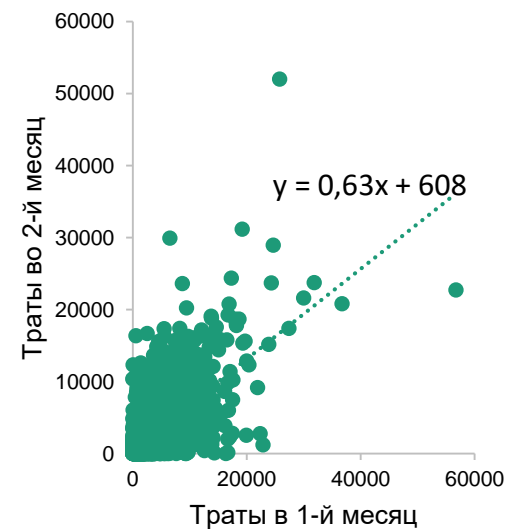
Группировка клиентов по количеству совершенных покупок



Распределение по возрасту



Соотношение трат в первые два месяца



Пол клиентов

