

Анализ данных e-commerce платформы

РЕТ-ПРОЕКТ ПО АНАЛИЗУ ПОВЕДЕНИЯ КЛИЕНТОВ,
ВЫРУЧКИ, ДОСТАВКИ И ПОЛЬЗОВАТЕЛЬСКИХ ОЦЕНОК

ИНСТРУМЕНТЫ: PYTHON (PANDAS), SQL (SQLITE), EXCEL,
POWER BI

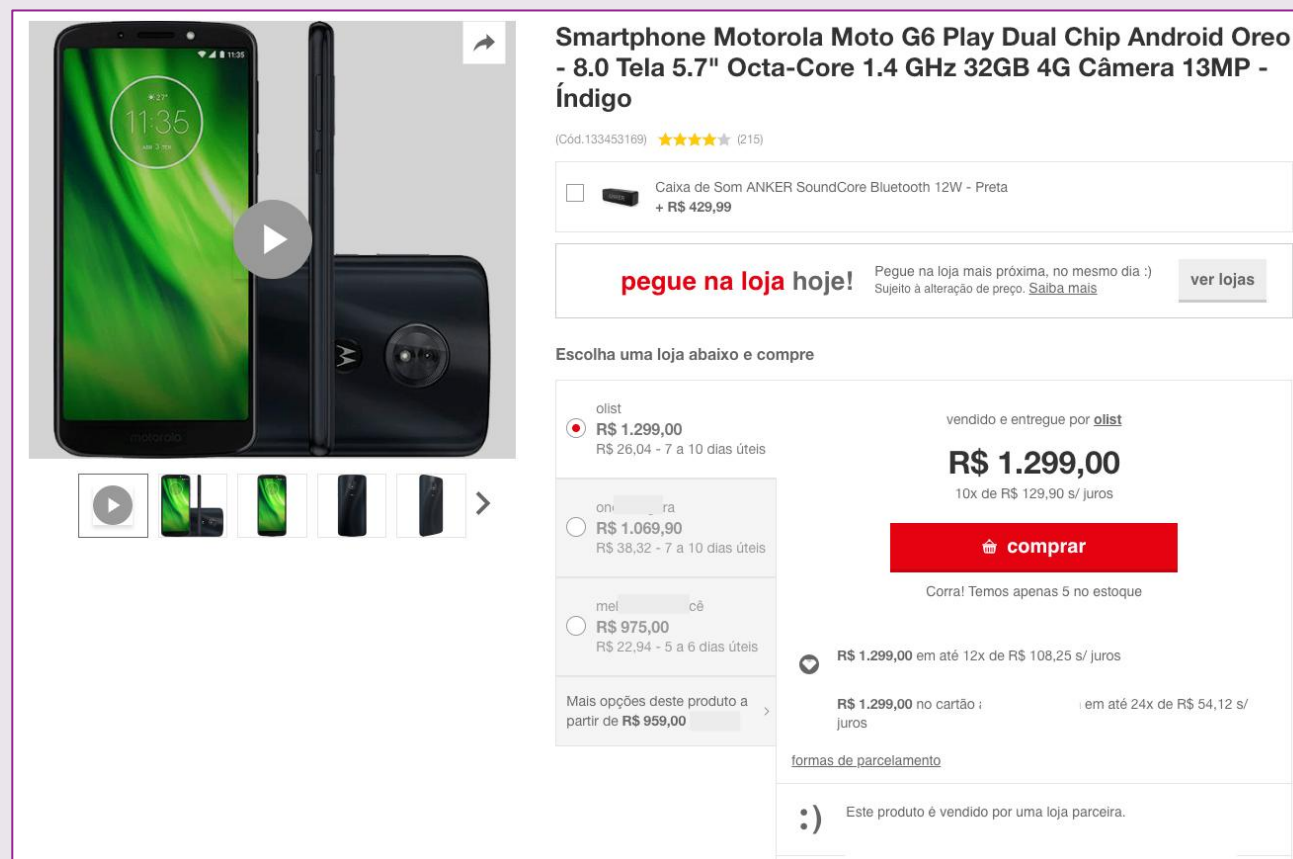
БЕЛОУСОВА АННА, JUNIOR DATA ANALYST

Описание продукта

Онлайн-платформа электронной коммерции, где клиенты оформляют заказы с доставкой товаров от разных продавцов, на заказ можно оставить отзыв с оценкой.

Бизнес заинтересован в:

- росте повторных покупок
- увеличении выручки
- улучшении клиентского опыта и рейтингов



Smartphone Motorola Moto G6 Play Dual Chip Android Oreo - 8.0 Tela 5.7" Octa-Core 1.4 GHz 32GB 4G Câmera 13MP - Índigo

(Cód.133453169) ★★★★★ (215)

☐ Caixa de Som ANKER SoundCore Bluetooth 12W - Preta + R\$ 429,99

pegue na loja hoje! Pegue na loja mais próxima, no mesmo dia :) Sujeito à alteração de preço. [Saiba mais](#) [ver lojas](#)

Escolha uma loja abaixo e compre

loja	preço	entrega
olist	R\$ 1.299,00 R\$ 26,04 - 7 a 10 dias úteis	vendido e entregue por olist
onipara	R\$ 1.069,90 R\$ 38,32 - 7 a 10 dias úteis	
melcê	R\$ 975,00 R\$ 22,94 - 5 a 6 dias úteis	

Mais opções deste produto a partir de R\$ 959,00

R\$ 1.299,00
10x de R\$ 129,90 s/ juros

[comprar](#)

Corra! Temos apenas 5 no estoque

☒ R\$ 1.299,00 em até 12x de R\$ 108,25 s/ juros

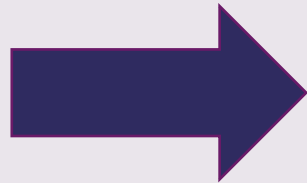
☐ R\$ 1.299,00 no cartão : em até 24x de R\$ 54,12 s/ juros

[formas de parcelamento](#)

:) Este produto é vendido por uma loja parceira.

Вопросы

1. Как ведут себя клиенты после первой покупки?
2. Как распределяется выручка между разными группами клиентов?
3. Как формируется выручка и какие категории товаров вносят наибольший вклад?
4. Есть ли сезонность в продажах?
5. Насколько эффективно работает доставка и как часто происходят задержки?
6. Какие факторы влияют на рейтинг заказов?

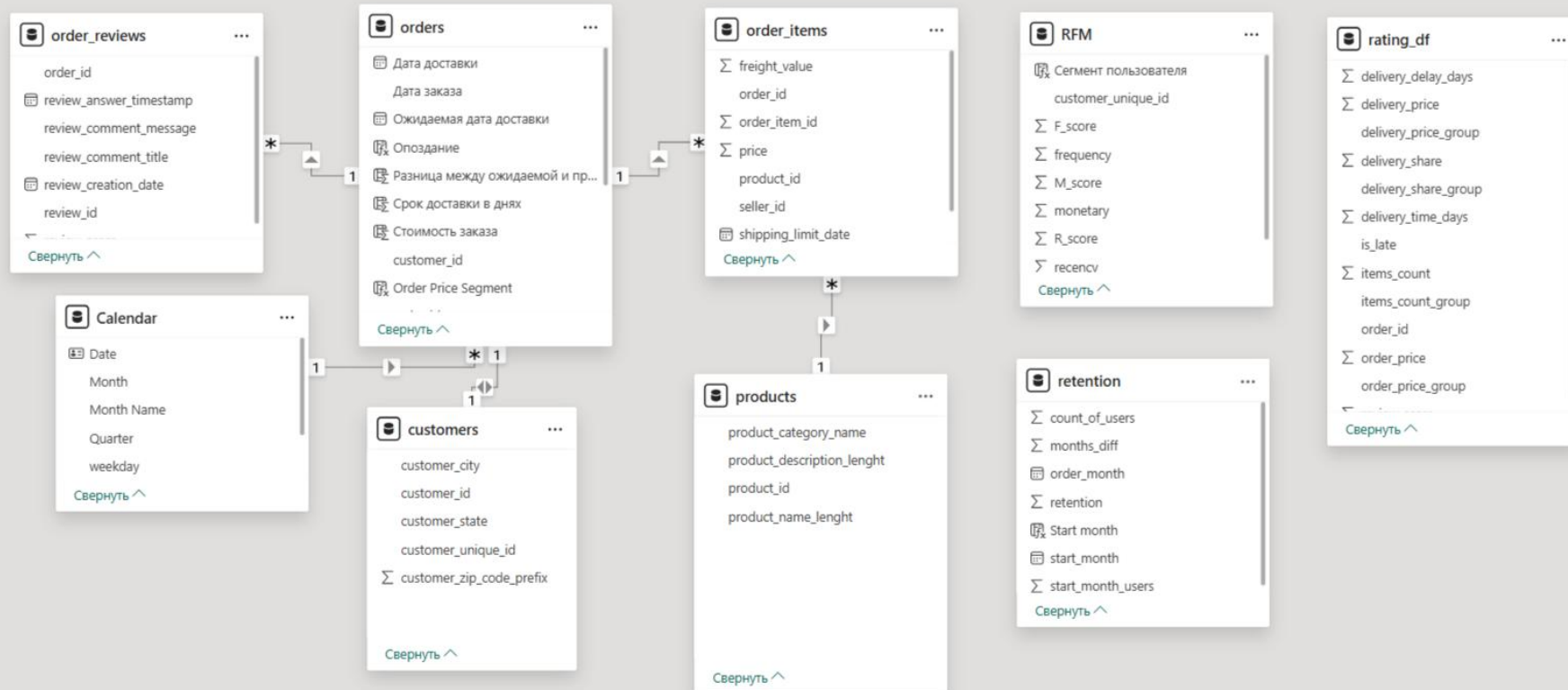


Задачи

1. Рассчитать месячный retention клиентов
2. Провести RFM-сегментацию клиентов
3. Проанализировать выручку и её структуру
4. Выявить сезонность продаж
5. Проанализировать задержки доставки
6. Найти факторы, влияющие на рейтинг заказов

Модель данных

- Один клиент может иметь несколько заказов
- Один заказ может содержать несколько товаров
- Для каждого заказа может быть оставлен отзыв



orders - информация о заказах и датах доставки
order_reviews - пользовательские оценки заказов
order_items - товары и стоимость заказов
customers - данные о клиентах
products - данные о категориях товаров

Retention

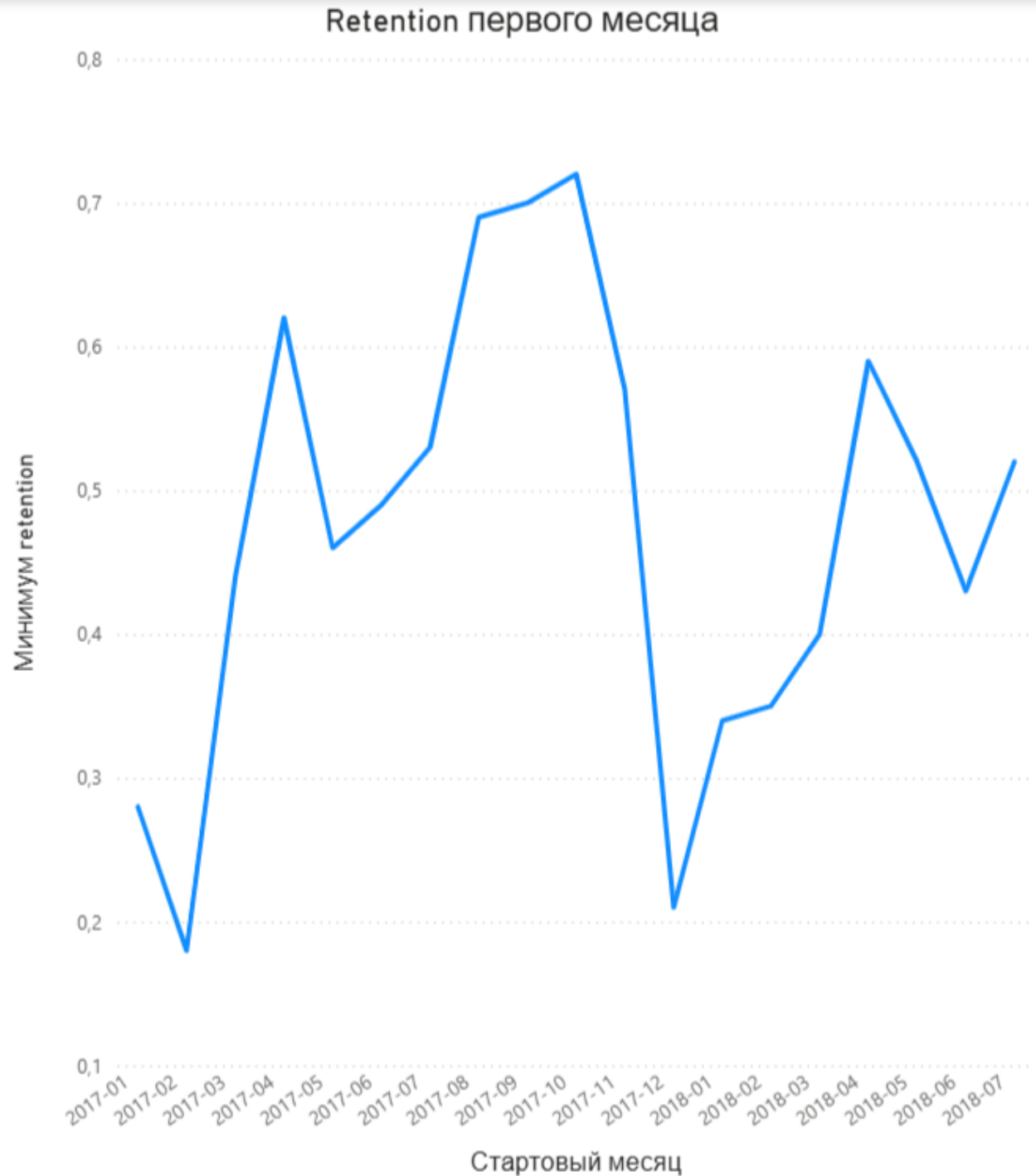
Возвращаемость клиентов после первой покупки крайне низкая: retention первого месяца стабильно ниже 1% (0,18–0,72%). Отсутствие резких провалов во времени указывает на системную проблему удержания, а не на влияние разовых факторов.



- *Продукт успешно привлекает новых клиентов, но не формирует привычку повторных покупок*
- *Потенциал роста — увеличение LTV за счёт повторных заказов, а не привлечение трафика*

←

Месячный retention											
Start month	0	1	2	3	4	5	6	7	8	9	10
2016-09	100,00										
2016-10	100,00						0,38			0,38	
2016-12	100,00	100,00									
2017-01	100,00	0,28	0,28	0,14	0,42	0,14	0,42	0,14	0,14		0,4
2017-02	100,00	0,18	0,31	0,12	0,43	0,12	0,25	0,18	0,12	0,18	0,1
2017-03	100,00	0,44	0,36	0,40	0,36	0,16	0,16	0,32	0,32	0,08	0,3
2017-04	100,00	0,62	0,22	0,18	0,27	0,27	0,35	0,31	0,31	0,18	0,2
2017-05	100,00	0,46	0,46	0,29	0,29	0,32	0,41	0,14	0,26	0,26	0,2
2017-06	100,00	0,49	0,40	0,43	0,30	0,40	0,36	0,23	0,13	0,20	0,3
2017-07	100,00	0,53	0,35	0,24	0,29	0,21	0,32	0,11	0,19	0,27	0,2
2017-08	100,00	0,69	0,35	0,27	0,35	0,52	0,30	0,27	0,15	0,15	0,2
2017-09	100,00	0,70	0,55	0,27	0,45	0,22	0,22	0,25	0,27	0,17	0,2
2017-10	100,00	0,72	0,25	0,09	0,23	0,21	0,21	0,37	0,28	0,18	0,2
2017-11	100,00	0,57	0,37	0,17	0,17	0,18	0,11	0,18	0,13	0,06	
2017-12	100,00	0,21	0,28	0,34	0,26	0,21	0,17	0,02	0,19		
2018-01	100,00	0,34	0,37	0,29	0,29	0,16	0,18	0,23			
2018-02	100,00	0,35	0,40	0,30	0,25	0,22	0,21				
2018-03	100,00	0,40	0,30	0,30	0,12	0,12					
2018-04	100,00	0,59	0,30	0,24	0,14						
2018-05	100,00	0,52	0,26	0,18							
2018-06	100,00	0,43	0,27								
2018-07	100,00	0,52									
2018-08	100,00										



RFM-сегментация

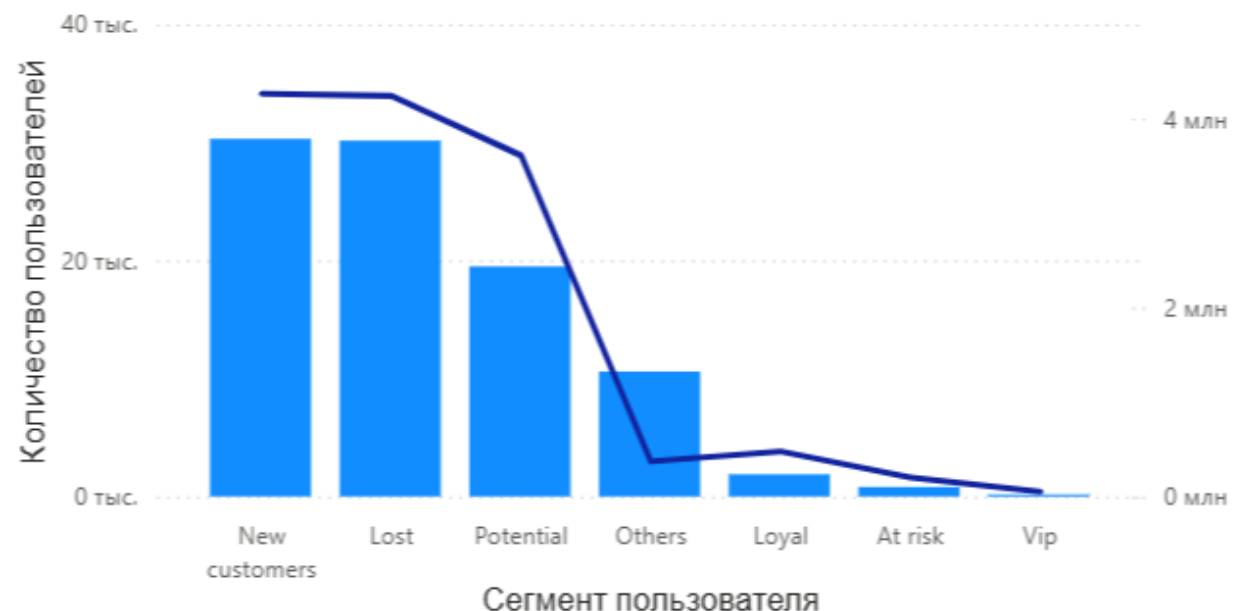
Самый крупный сегмент клиентской базы — новые пользователи, за которыми следует сегмент потерянных клиентов. При этом сегмент *Potential* является третьим по размеру и имеет более высокий средний чек, чем новые и потерянные пользователи.



- *Основной риск бизнеса — отток после первой покупки*
- *Сегмент *Potential* — ключевая точка роста: клиенты покупали недавно и на значимую сумму, но не стали повторными*
- *VIP и лояльные клиенты требуют персонализированных стратегий удержания*

Количество пользователей в сегментах

● Количество пользователей ● Выручка



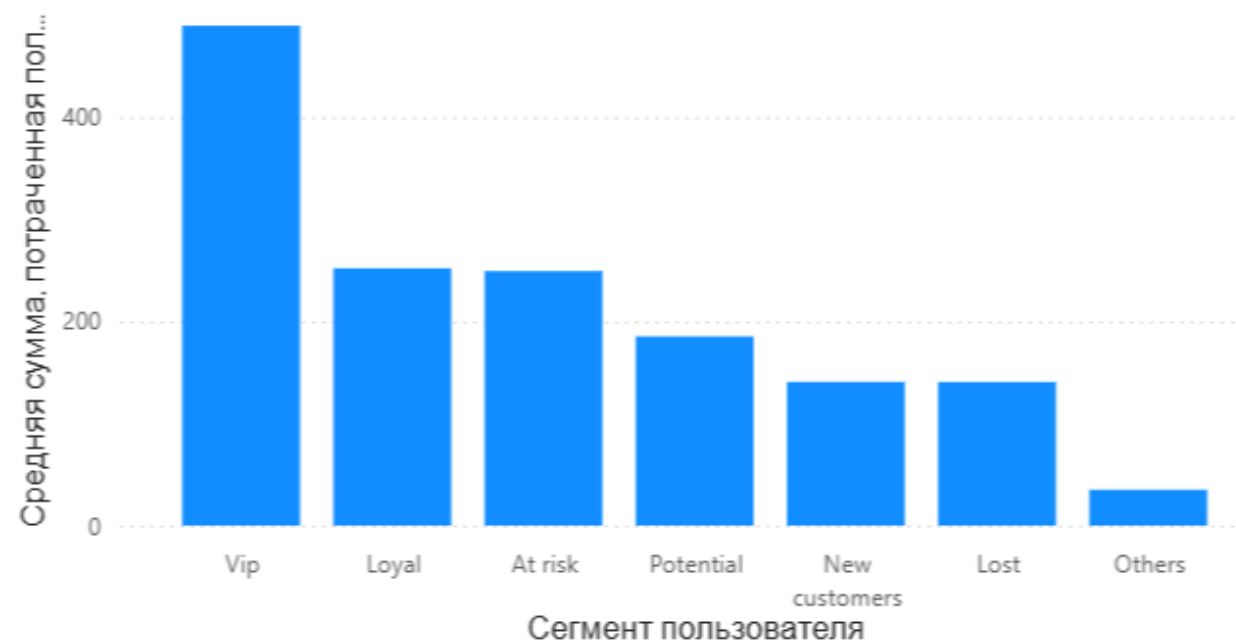
Количество пользователей в сегментах

F_score	1	2	3	Всего
1	30157	30090	30310	90557
2	749	903	921	2573
3	63	61	104	228
Всего	30969	31054	31335	93358

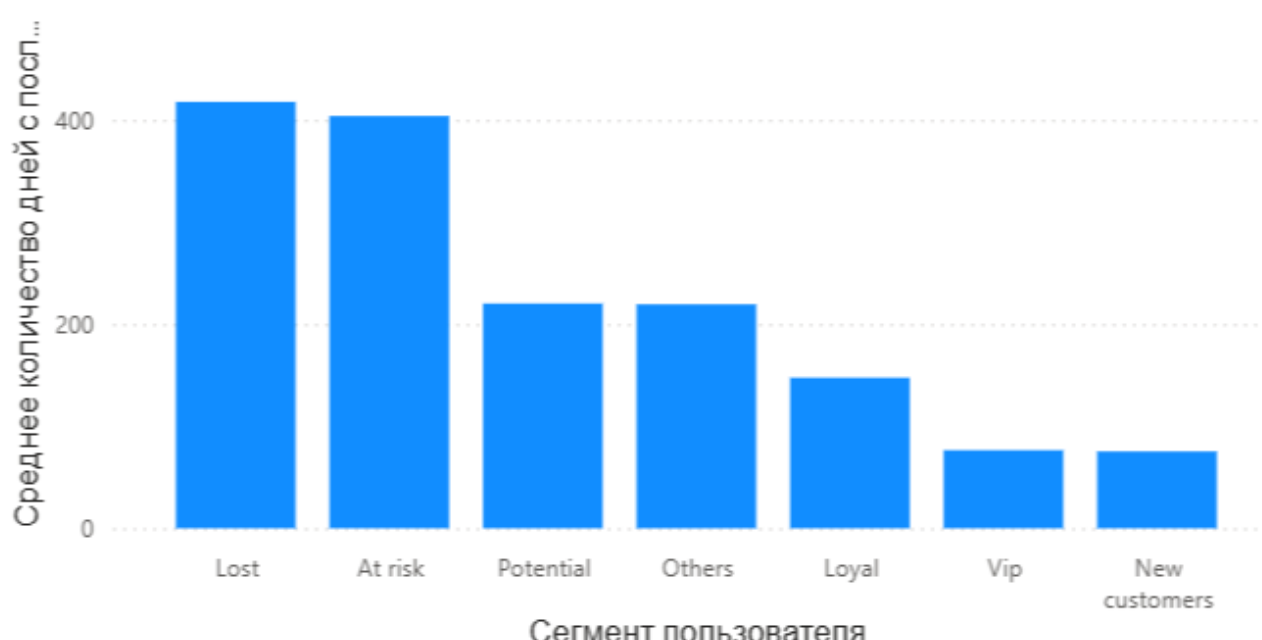
Описание сегментов:

1. New customers - 1 заказ, сделанный недавно
2. Lost - 1 заказ, сделанный давно
3. Potential - Клиенты с одной покупкой средней давности, но с высоким чеком
4. Loyal - Лояльные активные клиенты, регулярно совершают покупки, но не всегда самые свежие или частые.
5. At risk - Ранее активные клиенты, которые давно не покупали.
6. VIP - Самые ценные клиенты: покупали совсем недавно и делают это часто.

Средняя сумма, потраченная пользователем по сегментам



Среднее количество дней с последнего заказа по сегментам

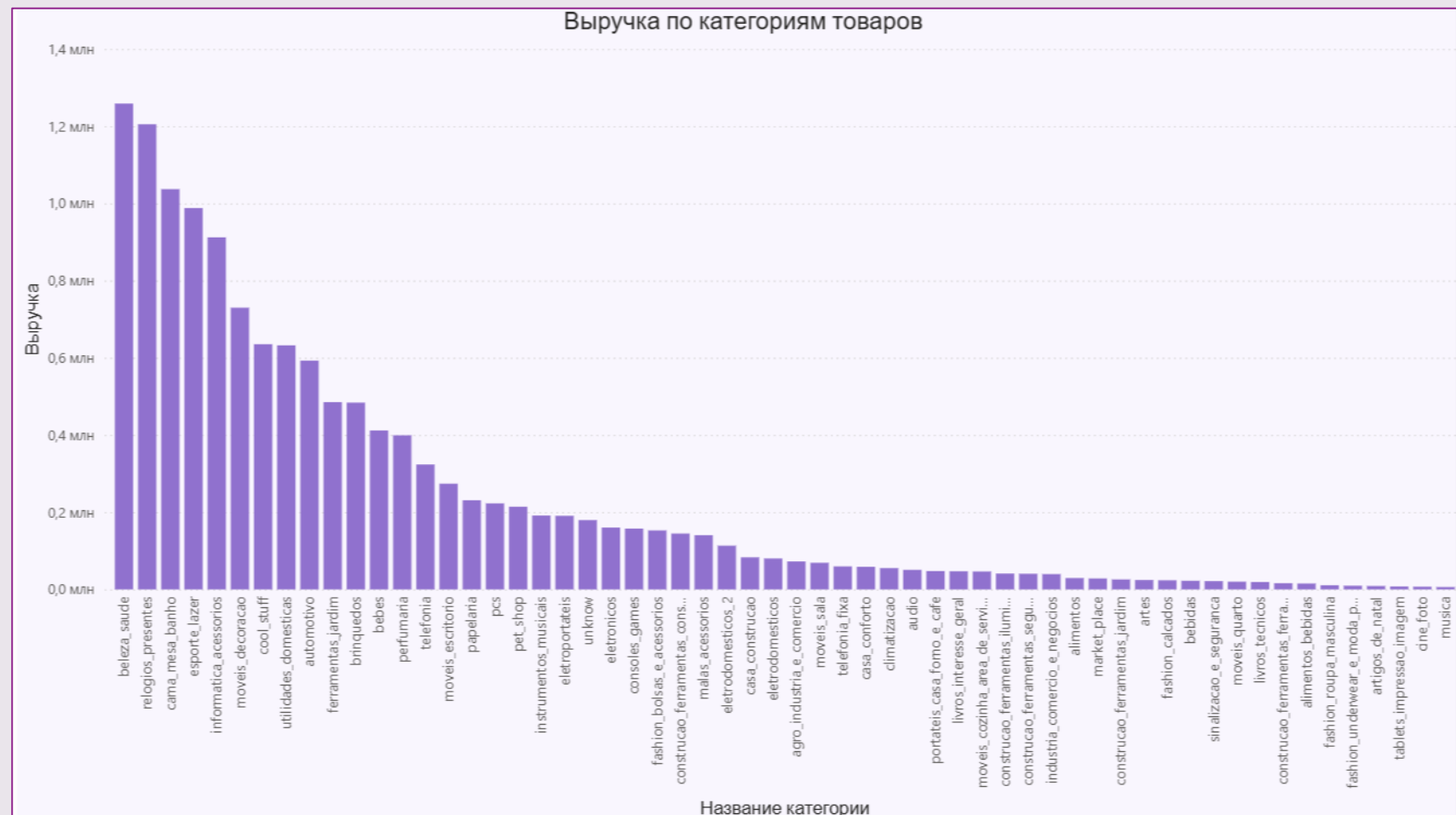


Выручка

Основную выручку формируют 5 категорий:

1. Красота и здоровье
2. Часы и подарки
3. Кровать, стол, ванна
4. Спорт и досуг
5. Компьютерные аксессуары.

- **Данные категории — ключевые драйверы выручки, в них целесообразно тестировать акции, персональные предложения и улучшения клиентского опыта**



Сезонность

В первые месяцы продукт находился, скорее всего, на стадии выхода на рынок, с января 2017 года наблюдается устойчивый рост. Пик продаж приходится на ноябрь 2017 года, после чего показатели выходят на плато.



- *Рост выручки обеспечивается увеличением количества заказов, а не ростом среднего чека*
- *Средний чек стабилен (124–149 R\$),*
- *Выраженный ноябрьский пик может быть использован для сезонных маркетинговых кампаний*

13,22 млн

Выручка в R\$

93,36 тыс.

Количество пользователей

96 тыс.

Количество заказов

137,04

Средний чек в R\$

141,62

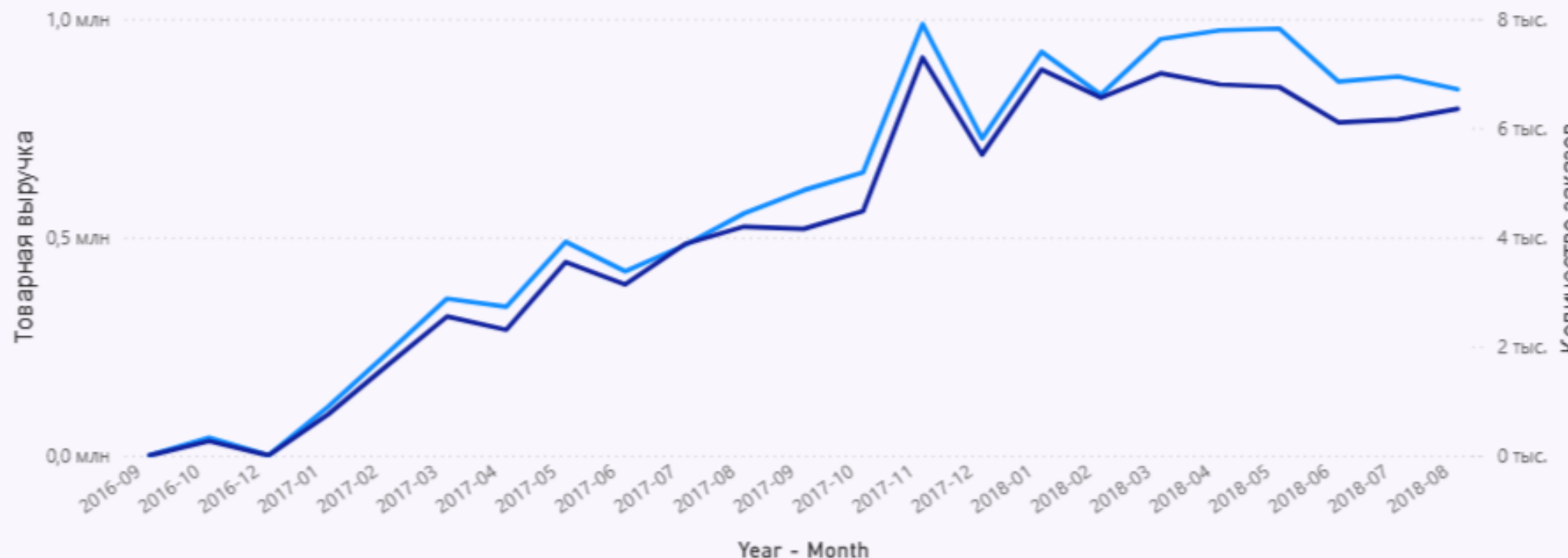
ARPU

141,62

ARPPU

Выручка и количество заказов по месяцам

Товарная выручка ● Количество заказов



Средний чек по месяцам



Доставка

Доля заказов с опозданием нестабильна и подвержена сильным колебаниям. В отдельных штатах (AL, MA, SE) уровень опозданий существенно выше среднего.



- *Нестабильность доставки снижает предсказуемость клиентского опыта.*
- *Возможные причины: некорректная оценка сроков доставки или локальные логистические проблемы.*
- *Фокус на проблемных регионах позволит снизить долю опозданий без масштабных изменений всей логистики*

10,62

Средняя задержка доставки
в днях

12,50

Средний срок доставки в
днях

99 ТЫС.

Количество заказов

6535

количество опоздавших зака...

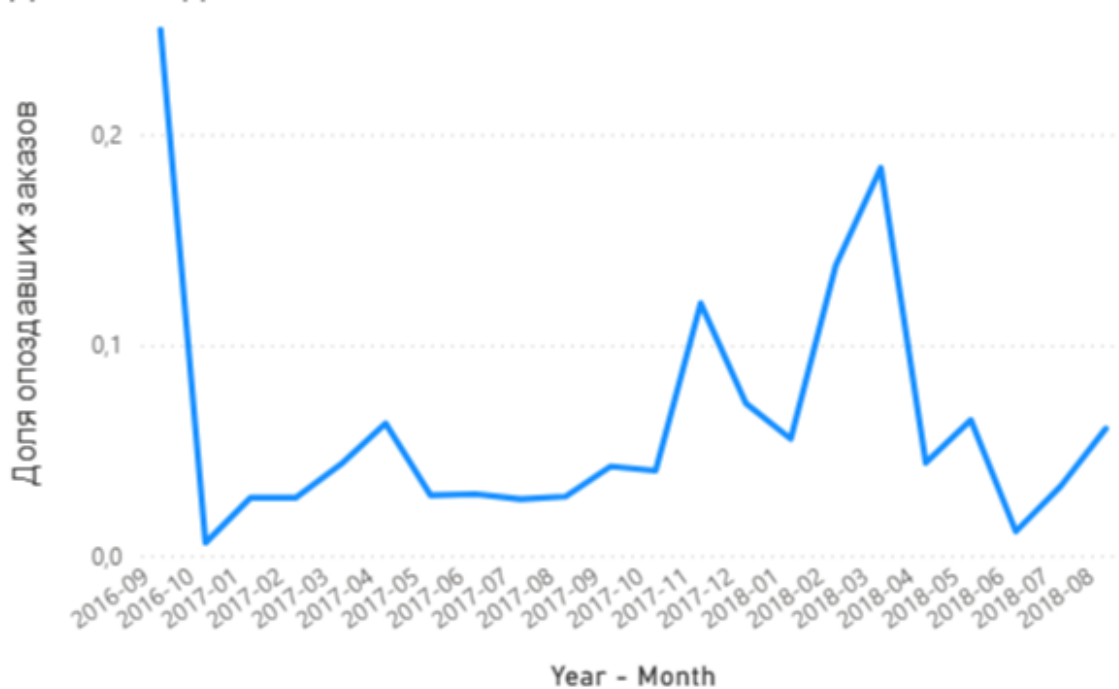
0,07

Доля опоздавших заказов

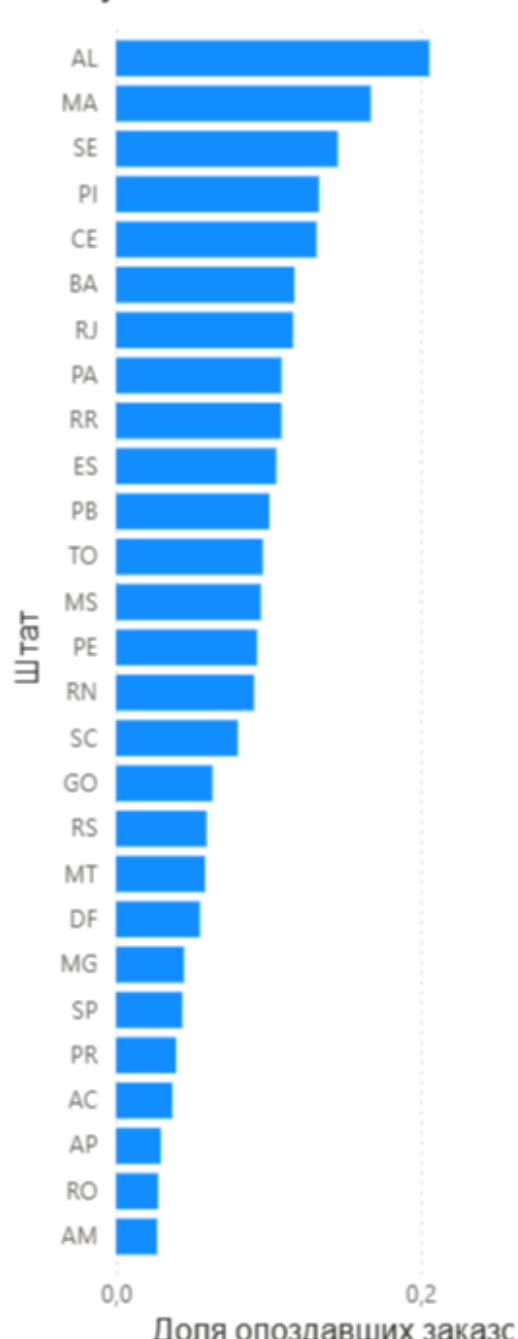
Средний срок доставки по Year - Month



Доля опоздавших заказов по Year - Month



Доля опоздавших заказов по Штату



Средний срок доставки по штату



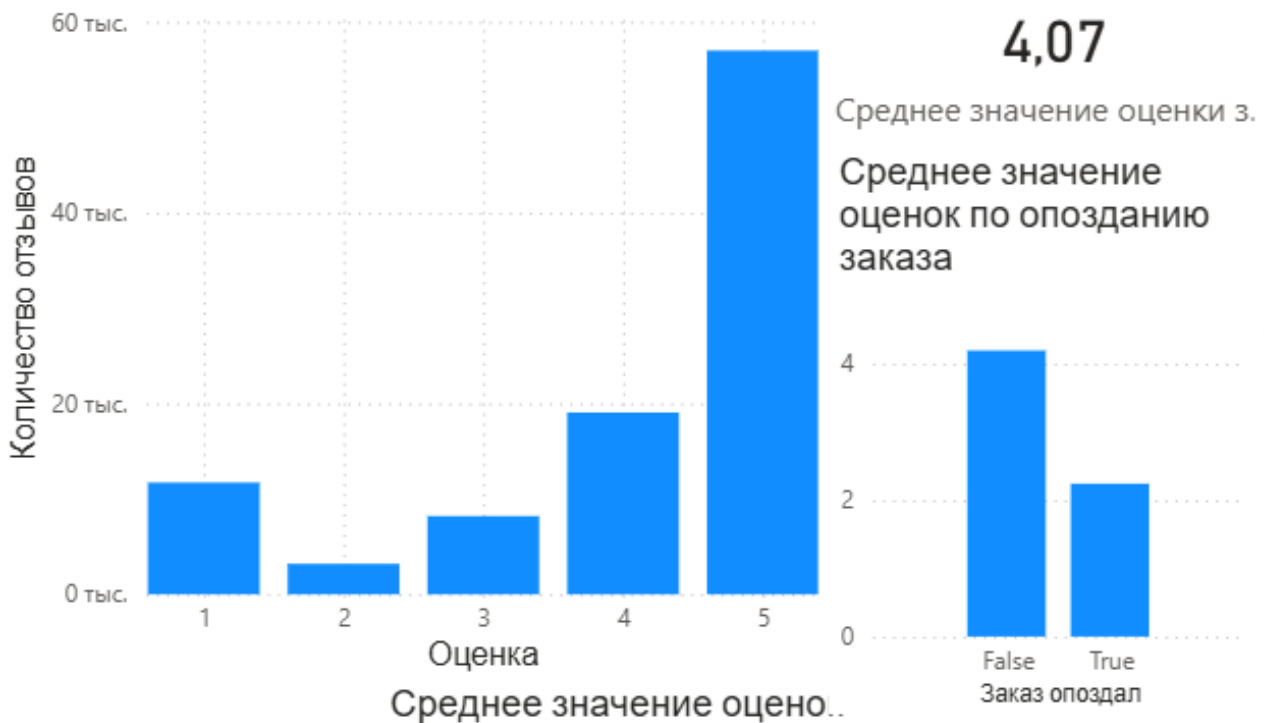
Рейтинг

Более дорогие заказы в среднем получают более низкие оценки, что указывает на рост ожиданий клиентов с увеличением суммы покупки. Заказы с большим количеством товаров и высокой стоимостью доставки также оцениваются хуже. Самый сильный негативный фактор — опоздание доставки.

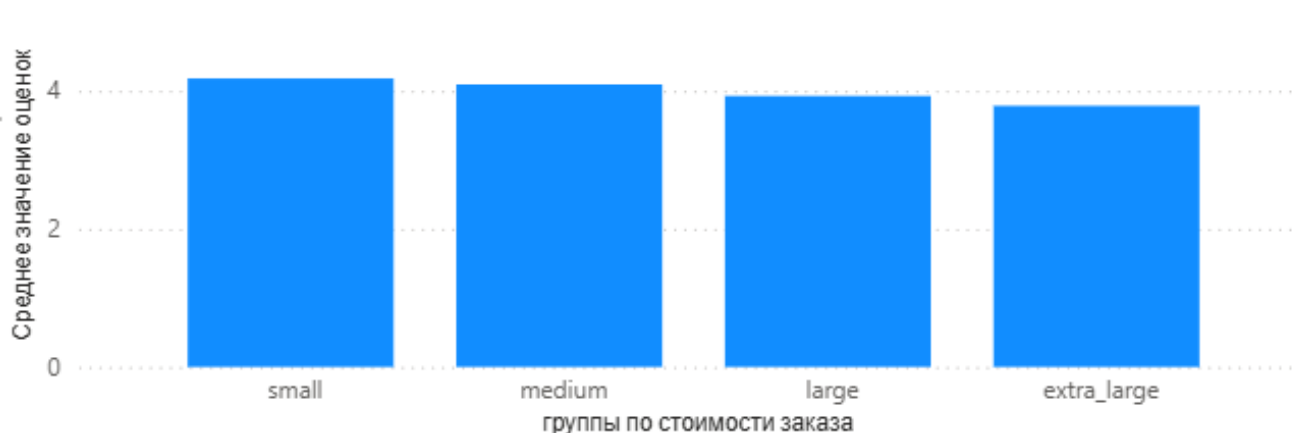


- *Клиенты готовы платить больше, но ожидают соответствующего уровня сервиса*
- *Скорость и надёжность доставки напрямую влияют на пользовательскую удовлетворённость*
- *Снижение доли опоздавших заказов — самый эффективный способ повышения рейтинга без изменения продукта*

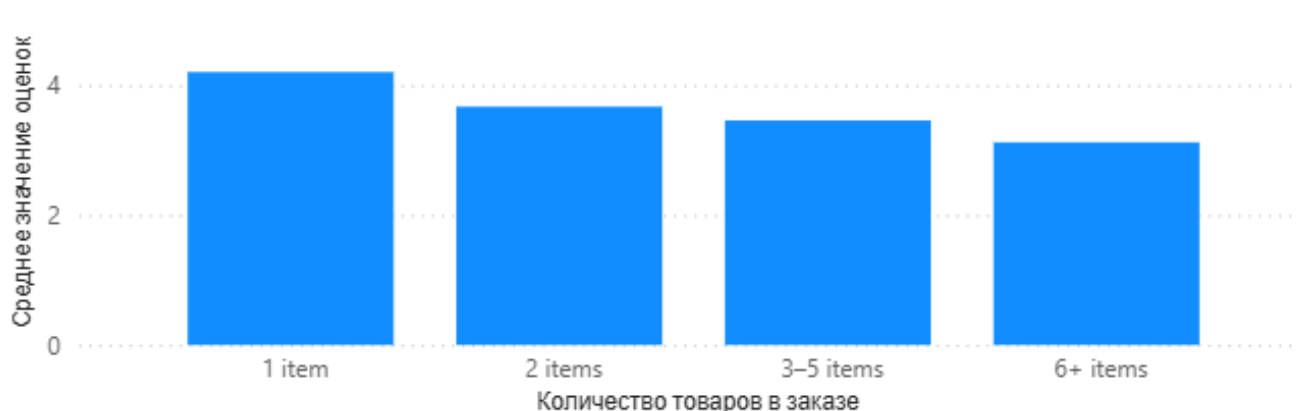
Количественное распределение оценок



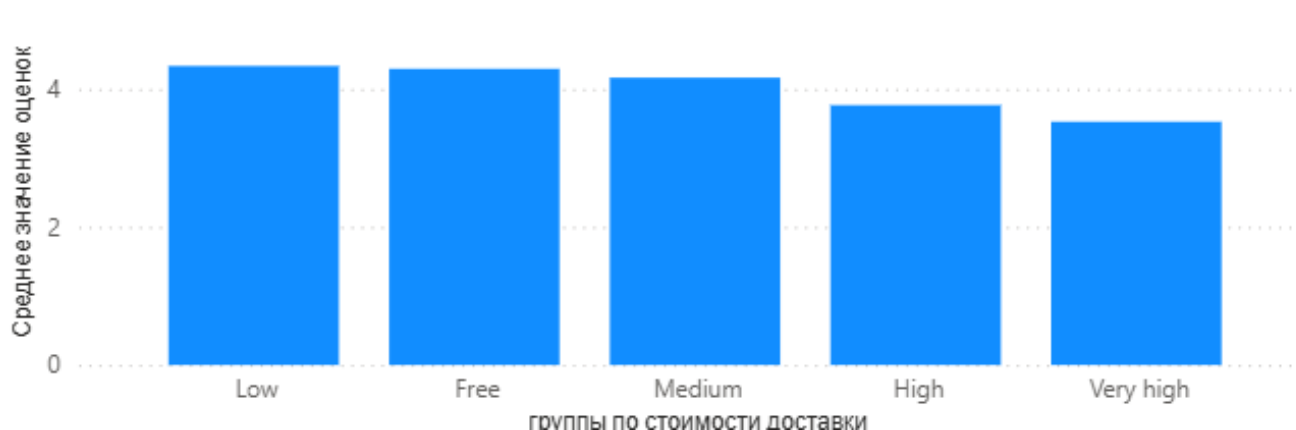
Среднее значение оценок по группам стоимости заказа



Среднее значение оценок по Количеству товаров в заказе



Среднее значение оценок по доли стоимости доставки



Рекомендации

1.
Сфокусировать
ся на
удержании
новых
клиентов:

- Персональн
ые акции
- Напоминания
- Коммуникац
ия после 1ой
покупки

2.
Снизить долю
опоздавших
заказов, начав с
проблемных
регионов (AL,
MA, SE)

3.
Использовать
сегмент
Potential как
приоритетную
группу для
повторных
продаж

4.
Усилить
контроль
доставки для
дорогих
заказов, где
ожидания
клиентов выше

5.
Развивать
ключевые
категории,
формирующие
основную
выручку, через
промо и
улучшение
сервиса

Контакты

Анна Белоусова

Junior Data Analyst

tg: @belousovaanya

email: belousova-2001@yandex.ru

GitHub: [belousovaanya2-hue/ecommerce-analysis](https://github.com/belousovaanya2-hue/ecommerce-analysis): E-commerce data analysis: RFM, retention, revenue, delivery and ratings

Dataset: [Brazilian E-Commerce Public Dataset by Olist](#)