1. Представь, что ты главный аналитик в компании и собираешь дашборд с ключевыми метриками (от 4 до 6) для СЕО, на основании которых он будет судить о "здоровье" бизнеса, эффективности текущих операций и т.д**. каждый день**.

 Исходя из специфики индустрии и бизнес-модели СберМаркета, какие метрики ты бы отразил? Коротко можно объяснить почему.

Ответ:

**Количество новых пользователей (первая регистрация):** так как рынок доставки находится на стадии активного роста, не смотря на лидирующие позиции СберМаркета важно и далее привлекать больше новых клиентов.

**Доля потерянных клиентов** (тех, кто не сделал повторный или следующий заказ в ожидаемые сроки): этот показатель отражает удовлетворенность качеством услуги

**Доля доставок с задержкой от общего числа доставок:** ключевое преимущество сервиса – быстрая доставка, важно отслеживать, когда происходит сбой и как часто.

**Доля замен и отмен от общего количества проданных продуктов:**  так как ненайденные товары – это проблема, отслеживаем динамику этого показателя.

**Посещение сайта (трафик)** – чтобы выявить провалы в рекламе, выдаче, работе сайта.

Конверсия посещения в сайта в дальнейшие действия клиента на сайте.

**Выполнение месячного плана (накопительным итогом)** сравнение доли выполнения с прогнозной долей на сегодняшний день.

2. Отчет по исследованию среза транзакций

Выводы

1 Период

Реестр содержит заказы с 2018-08-26 по 2019-12-18. В начале и конце периода

неполные месяцы, но там почти полные недели. 2018-08-26 – это воскресенье и данные

есть прямо с полуночи, а вот конец периода 2019-12-18 – это среда и не полный день.

Обрезаю три последних дня, чтобы не искажать статистику по неделям, а по месяцам

анализирую с учетом неполных месяцев по краям.

2 Аномалии

Три случая отрицательных значений в стоимости доставки, не существенно для

анализа данных, но стоит обратить внимание в целом с точки зрения возможных

убытков из-за ошибок.

Есть и очень дорогие доставки, но их менее 1% случаев - не похоже на ошибки

(заказы большие по сумме и весу).

Выявлены заказы в нулевым весом или очень маленьким весом (менее 50 граммов) -

возможно, это редкая ошибка в заполнении веса заказов (доли процента от общего

количества заказов) и 26 заказов с нулевым количеством продуктов (но есть

стоимость, вес и даже платная доставка).

Обнаружены 9 заказов с нулевой стоимостью, при этом у заказов есть вес (от 5 г до

18 кг) и товары в составе (от 3 до 18 товаров). 7 таких заказов отправлены ретейлеру

№1 (из них 6 - в город №1, один - в город №2), ещё два заказа отправлены в город

№13 ритейлерам №8 и 15.

Аналогичный анализ надо провести для заказов менее 100 и, возможно, 500 рублей.

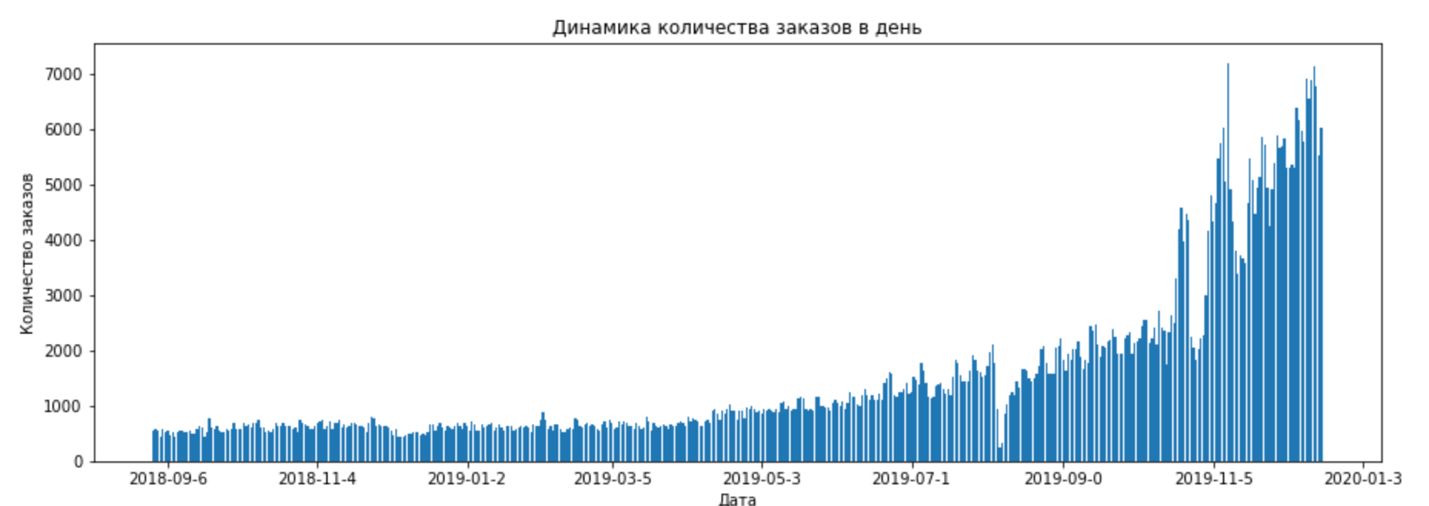
Доля этих заказов очень небольшая, но интересно из изучить с точки зрения рисков

мошенничества.

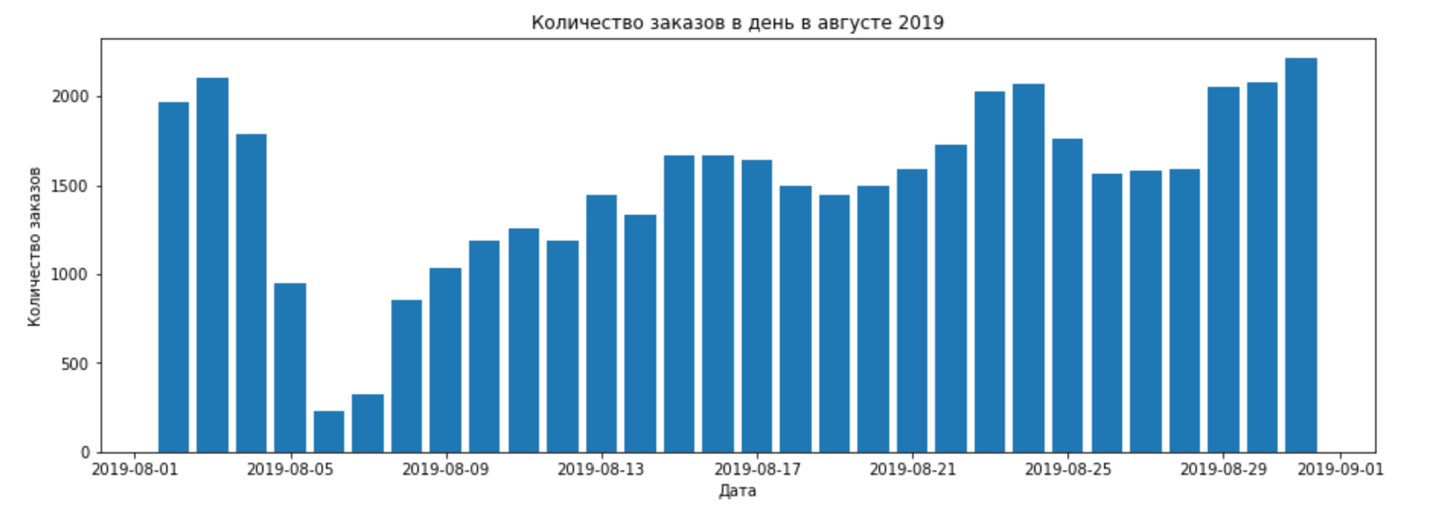
3 Описание тренда

Рост количества пользователей и заказов начался в середине 2019 года, а затем в

октябре произошел резкий скачок.



6-7 августа 2019 есть нетипичный для всего периода провал в количестве заказов.



Возможно, что 6 августа (или с конца дня 5 августа) были перебои в работе сайта или

мобильного приложения. В течение 7-8 августа объем заказов восстановился.

4 Дни недели



Есть выраженная корреляция между днем недели (пн-вс) и такими показателями как средняя сумма заказа, средний вес заказа и среднее количество позиций в заказе, и заметная корреляция с количеством заказов.

Не знаю, как эта информация может быть полезна бизнесу. Возможно, полезно более подробно изучить эту корреляцию для планирования нагрузки на курьеров.

5 Когорты

Хотя у нас период обрезан сначала, но всё-таки он длинный и интересно было

посмотреть на когорты клиентов. Я выделила когорты по месяцу первого заказа и

проанализировала количество размещенных заказов каждой когортой. Ожидаемо,

что в пиковый период пришло больше всего именно новых пользователей и они

разместили много первых заказов.

Интересно, что количество заказов на пользователя в месяц мало менялось с конца

2018 года. Но некоторое повышение произошло как раз в районе середины 2019 года

и максимальной величины этот показатель достиг именно в месяцы пикового роста -

ноябрь 2020 года. То есть тут есть и повышение повторных заказов за короткий

период. Можно углубитсья и покопаться, кто именно размещал больше повторных

заказов.

Также ожидаемо, что среднее количество заказов на пользователя коллелирует с

возрастом когорты, то есть клиенты магазина лояльны и продолжают снова и снова

размещать заказы. Тут немного странно, что у клиентов августа-2018 этот показатель

аномально высокий (почти в два раза выше сентября), но у нас августа обрезанный,

так что выводы не сделать.

Далее можно сравнить когорты за одинаковые периоды. Взять, например,

трехмесячные периоды, и сравнить когорты по количеству заказов на человека за

первые три месяца. В связи с ограниченным временем в рамках данного отчета я

дальше углубляться не стала.

С 2018-09 количество новых пользователей в месяц снижается постепенно вплоть до

2019-04, после чего начинается резкий всплеск. Нельзя исключать влияние на отчет

того, что у нас обрезанный период и "новые клиенты" могут быть не очень новые в

начале отчета, но потом уже явно виден приток новых клиентов с апреля 2019,

поэтому те маркетинговые усилия, которые проводились в это время явно были

правильные и привлекли много новых клиентов.

6 Отмены заказов

В таблице с отменами все id продуктов разные, хотя не может быть, чтобы каждый

продукт отменили только по одному разу и больше никогда с продуктом такого не

случалось. Возможно, тут какой-то динамический id продукта, который присваивается

при помещении в корзину или при подтверждении заказа. Было бы полезно иметь

статистику по постоянным id товаров, выявлять конкретные товары, по которым есть

проблемы с доступностью и работать с поставщиками для улучшения доступности

ассортимента, или выявлять "рискованных" поставщиков и работать с

альтернативными.

Соединив общую таблицу заказов с реестром отмен, удалось составить рейтинг

магазинов по доле отменных продуктов от общего числа заказанных продуктов за

весь период и в динамике по месяцам, то есть можно посмотреть отдельно по

каждому магазину, как меняется доля отменных продуктов от общего числа

продуктов и при росте доли адресно работать с магазинами по улучшению этого

показателя.

Например, самая высокая доля отмененных продуктов в целом за период в магазинах

№128, 238, 258, 257, 57. А растет эта доля в магазинах № 8, 57, 92. Соответственно,

магазин №57 нуждается в проверке в первую очередь.

Аналогичный рейтинг можно составить и по заменам.

7 Отрицательные выводы

Отрицательный результат - это тоже иногда интересный результат.

Нельзя сказать, что есть какая-то тенденция в среднем количестве продуктов в

заказе в целом за период и в среднем количестве продуктов в заказе. Если копать

дальше, я бы посмотрела эти показатель по когортам, по городам, по магазинам и по

ретейлерам, чтобы выявить признаки, коррелирующие с повторными заказами.

Повторные заказы - это хорошо.

8 Промо-коды

Доля заказов с промокодами от общего числа заказов в целом за период: 13,5%.

Доля скидки от общей суммы заказов со скидками: 8,3%

Хотя ежемесячная сумма промокодов растет в течение всего периода, но средняя

сумма промокодов на количество заказов уменьшилась, а самая минимальная она в

пиковый месяц ноябрь. Получается, что промокоды на этот пик не повлияли.

1. Что ещё интересно посмотреть, но я не успела:

* Посмотреть корреляцию между прокодами и другими показателями (количество заказов, количество продуктов в заказе и др).
* периодичность заказов (время между первым и следующим заказами) по каждому пользователю;
* заказы с большими суммами: есть ли общие признаки и тенденции;
* уточнить, item\_total указана за вычетом промокода или нет. Если нет, то оценить долю очень маленьких заказов после вычета промокода из суммы
* показатели продаж в разрезе магазинов, ретейлеров, городов

Я не сделала аналитику по городам (из-за дефицита времени), но полезно посмотреть на динамику различных показателей по городам. Ожидаемо, города №1 и 2 будут идти с большим отрывом в абсолютных показателях, но можно посмотреть относительные приросты.

* Какие города/ретейлеры дали наибольший прирост в пиковое время, узнать, за счет

чего, и попробовать повторить из опыт на других локациях. Также интересно

отслеживать динамику различных показателей в разрезе городов и ретейлеров,

чтобы вовремя заметить негативные тенденции и принять меры.

* среднее время доставки в разрезе магазинов (а также городов и ретейлеров),
* корреляцию среднего времени с другими показателями продаж.