



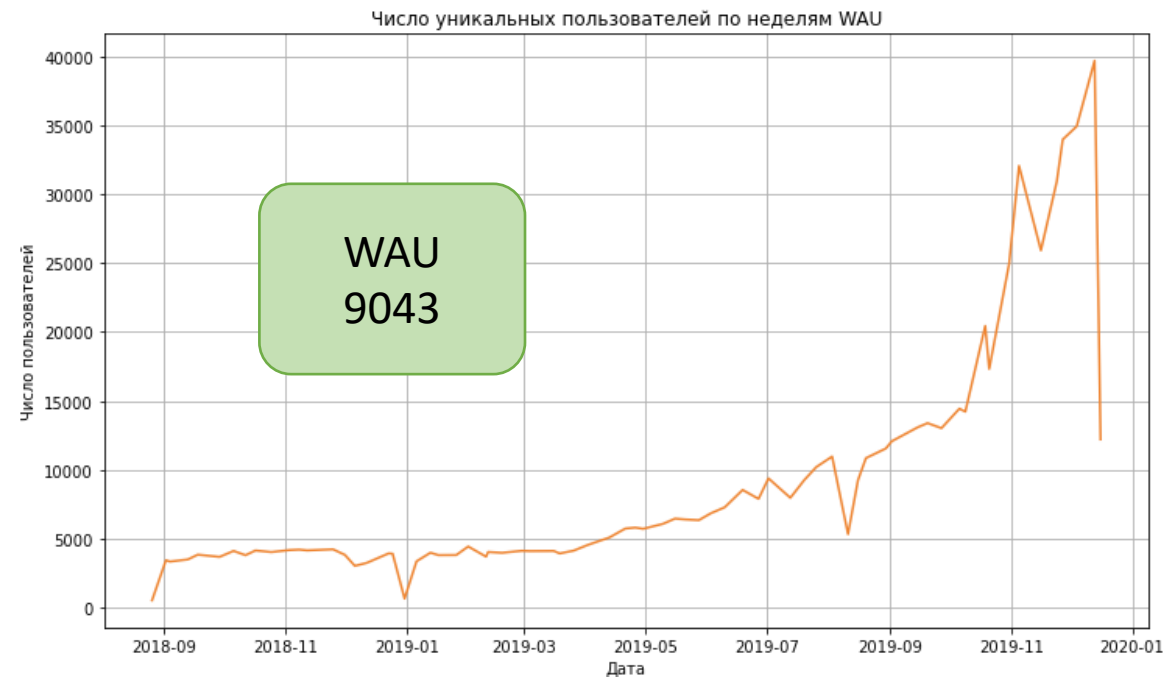
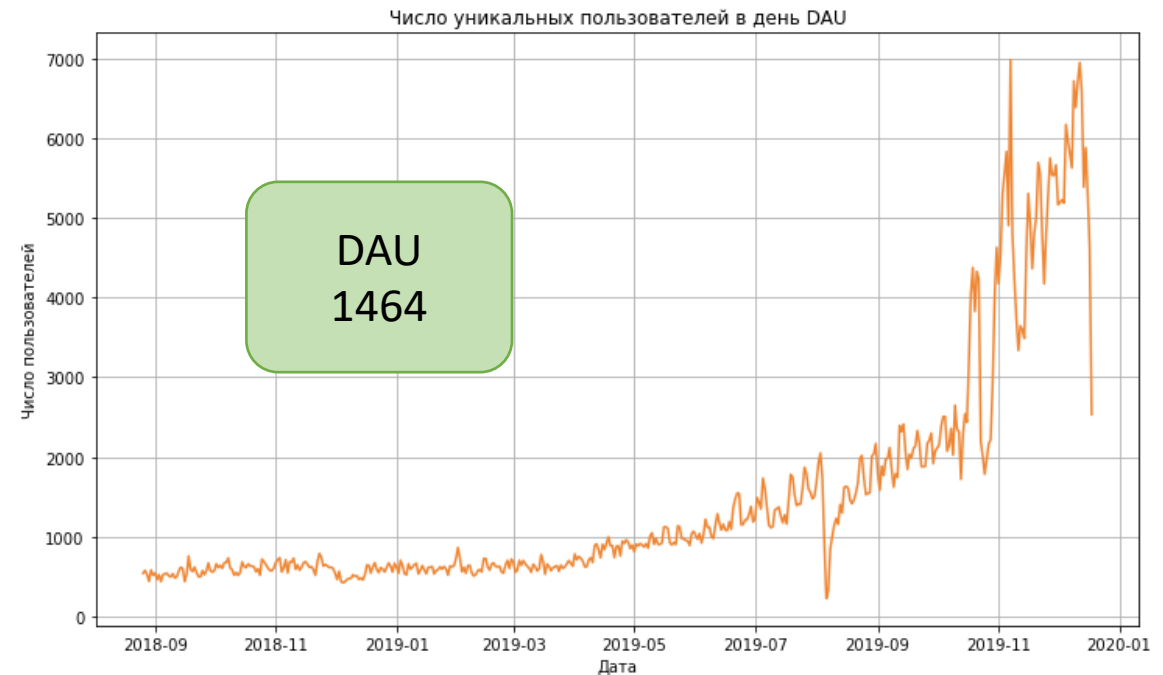
## Тестовое задание 2: Инсайты данных по логам базы

Захаров Дмитрий

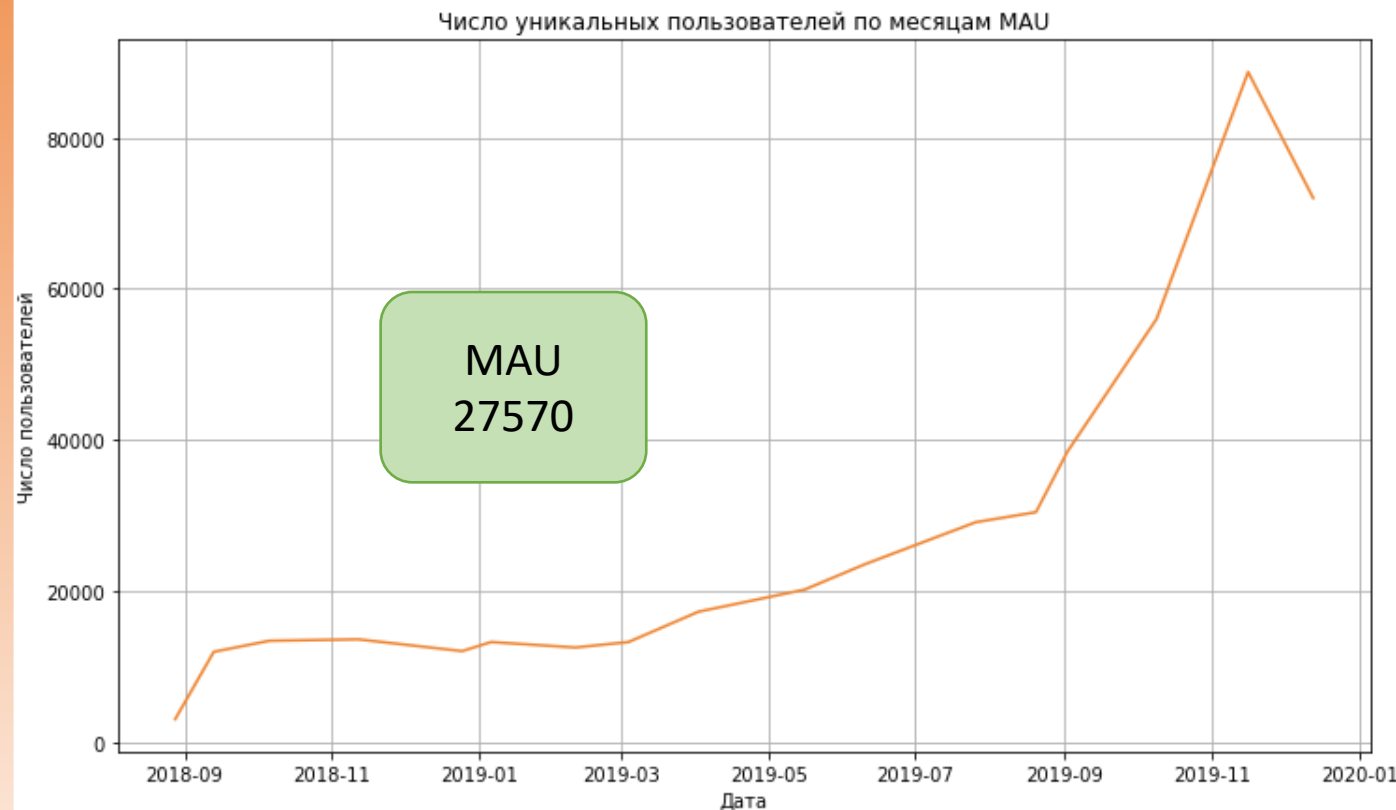
# Уникальные пользователи

По неделям наблюдаем выбивающиеся снижения в январе(праздники) и августе, ноябре 2019 года.

Августовское и ноябрьское снижение сложно объяснить, это мог быть сбой в течение недели.



# Уникальные пользователи



Число уникальных пользователей растет от месяца к месяцу, значение первого и последнего месяца досчитаны не полностью, учитываем это. По неделям наблюдаем выбивающиеся снижения в январе(праздники) и августе, ноябре 2019 года.

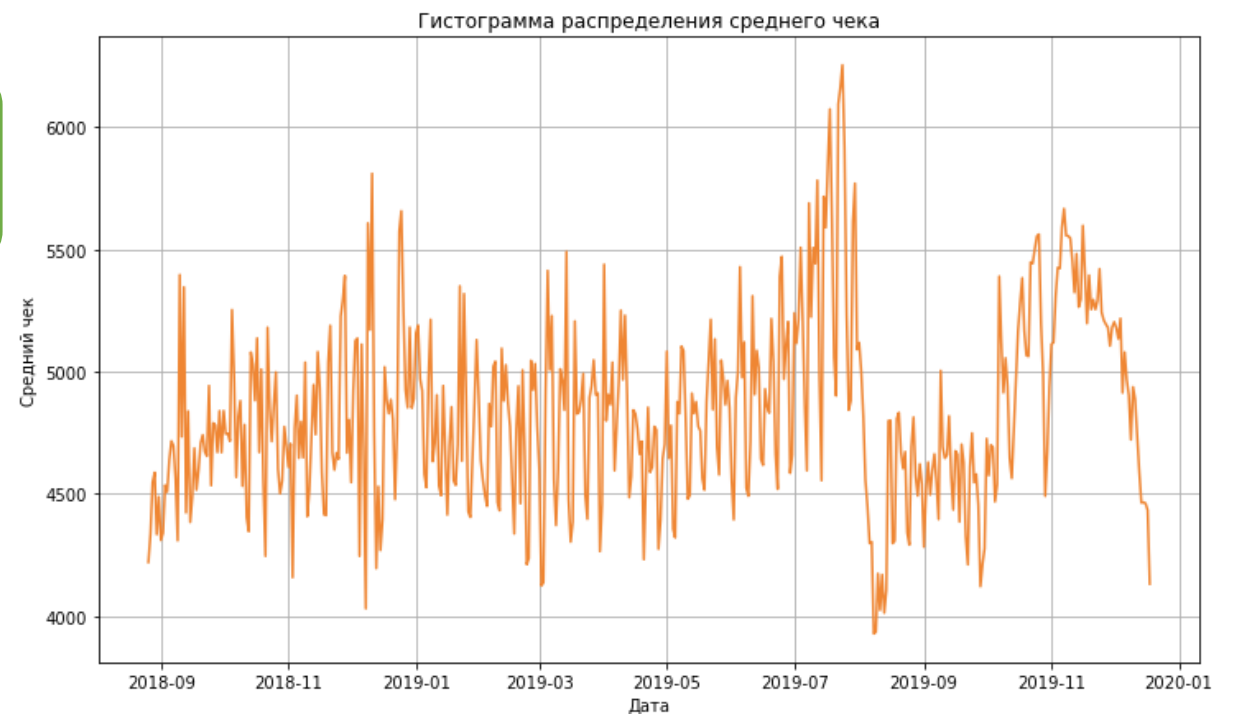
# Средний чек

Средний  
чек  
4936.18

По неделям наблюдаем выбивающиеся снижения в январе(праздники) и августе, ноябре 2019 года.

Трудно сказать, является ли такая особенность сезонной, так как не имеем данных за прошлый год. Из того, что значения не отросли до прежних максимумов, а зафиксировались в новом интервале, можем сказать, что результат - следствие изменений в продукте.

Предположили рост стоимости доставки, но это не подтвердилось, средняя стоимость доставки в этот период снижалась.



## Топ ритейлеров

Выделили топ ритейлеров, самый крупный обслуживает 224214 клиентов со средним чеком 5145 рублей. Самый мелкий 1023 клиента со средним чеком 3599 рублей.

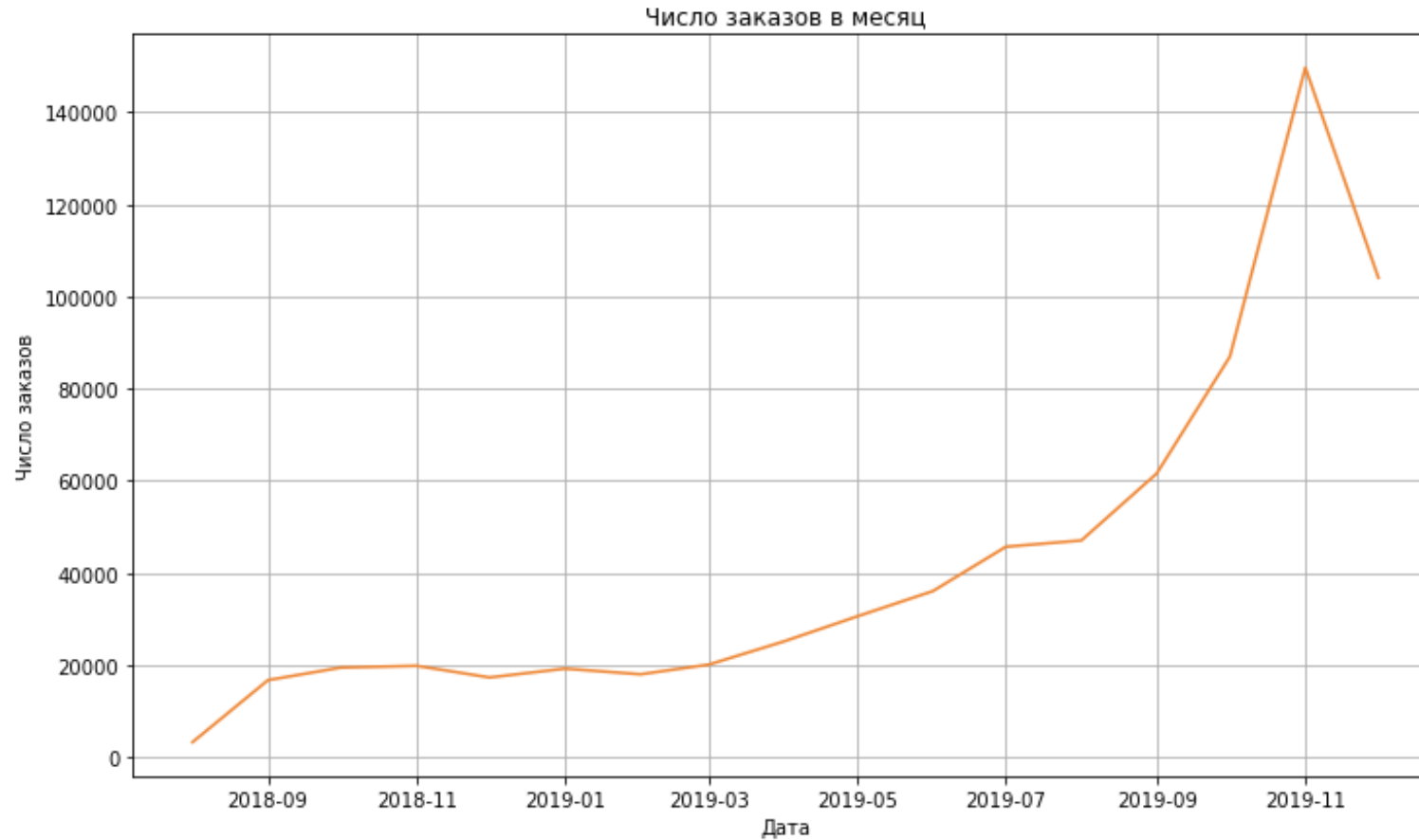
| retailer_id | count_orders | unique_users | total_sum  | total_mean  |
|-------------|--------------|--------------|------------|-------------|
| 1           | 619200       | 224214       | 3185986562 | 5145.327135 |
| 8           | 53612        | 29930        | 179163981  | 3341.863407 |
| 15          | 45348        | 29258        | 182855341  | 4032.269141 |
| 16          | 2273         | 1023         | 8181123    | 3599.262209 |

## Топ городов

Выделили топ городов, в городе 1 имеем 176015 клиентов и самый высокий средний чек 5500 рублей.

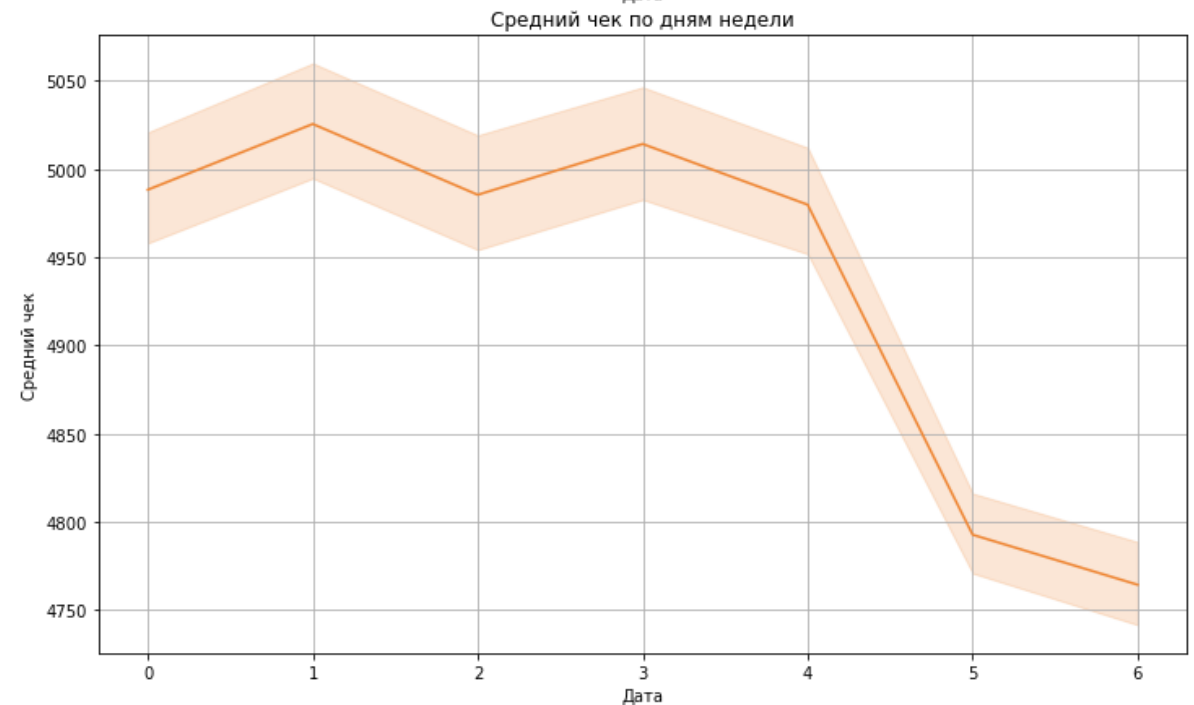
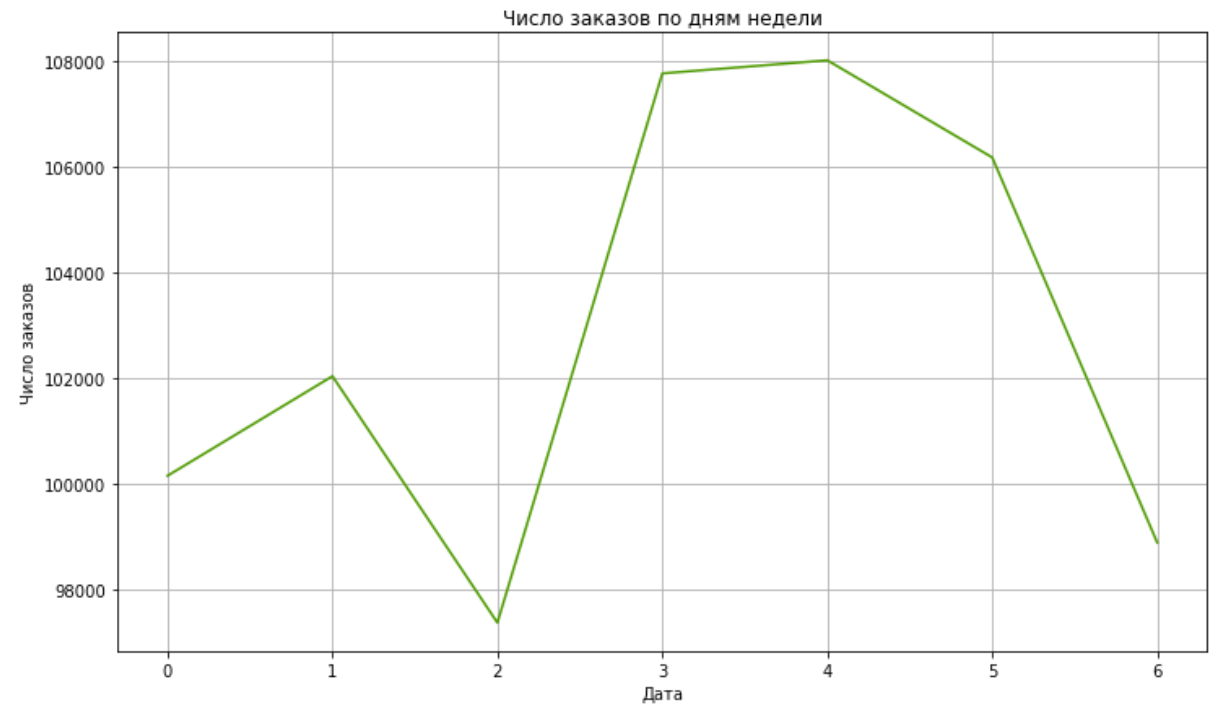
| city | count_orders | unique_users | total_sum  | total_mean  |
|------|--------------|--------------|------------|-------------|
| 1    | 478861       | 176015       | 2633935595 | 5500.417856 |
| 7    | 76360        | 21819        | 281712395  | 3689.266566 |
| 10   | 46789        | 15816        | 181328178  | 3875.444613 |
| 2    | 39784        | 13948        | 155608511  | 3911.333978 |
| 5    | 29386        | 11671        | 115071114  | 3915.848159 |
| 8    | 28677        | 11600        | 111014967  | 3871.219688 |
| 13   | 20576        | 7551         | 77516247   | 3767.313715 |

# Рост числа заказов



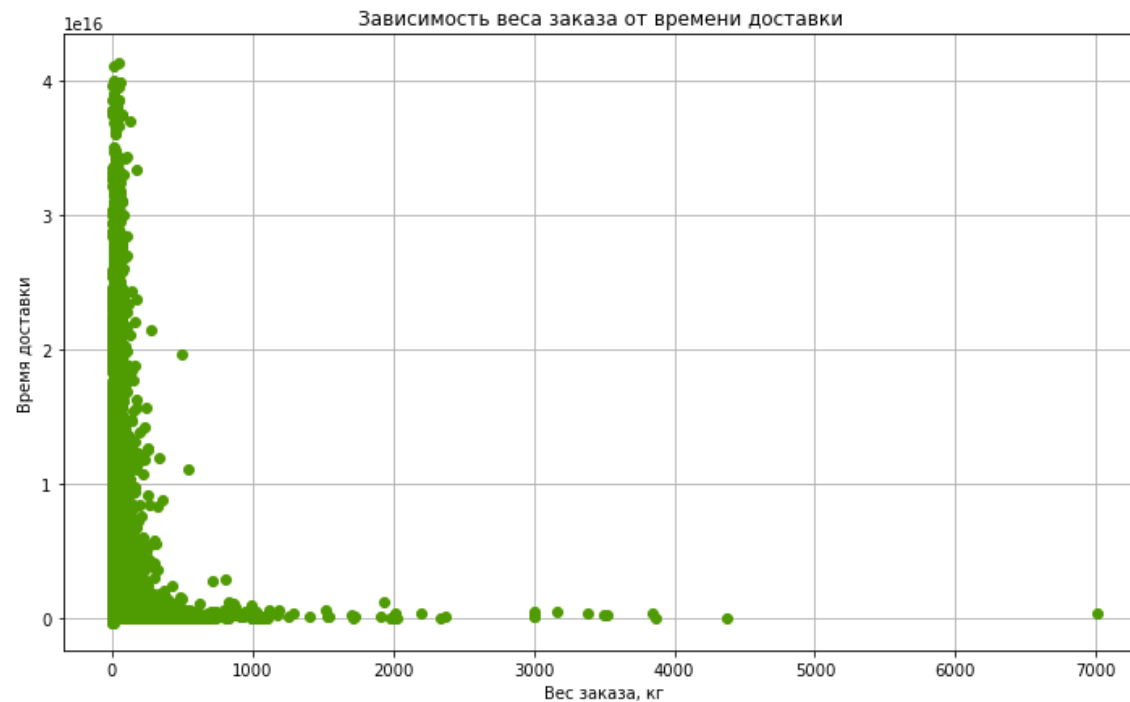
- Число заказов в сентябре 2018: 16708
- Число заказов в ноябре 2019: 149706
- Рост составляет 896.01%

- Средний чек снижается в период выходных(с 5000 до 4800 рублей).
- Максимальное число заказов получаем в четверг и пятницу(108000), минимальное в среду(около 98000) и воскресенье(98000).

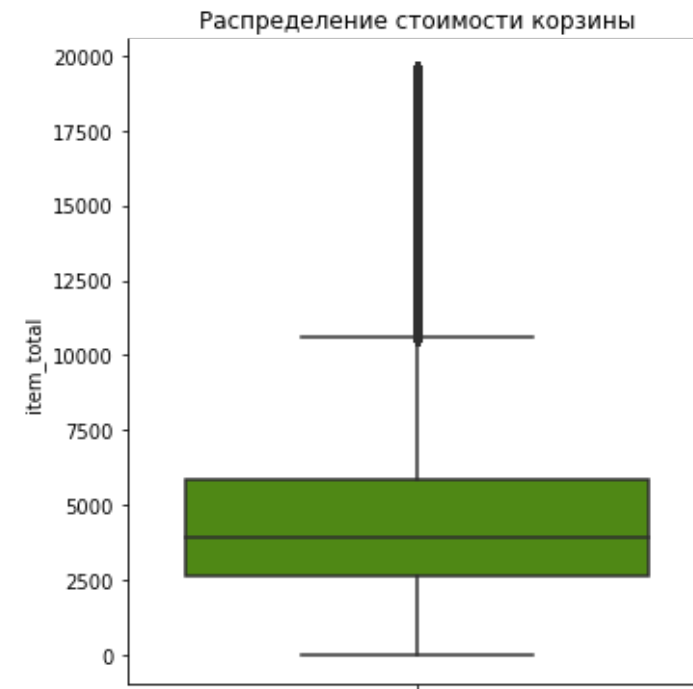
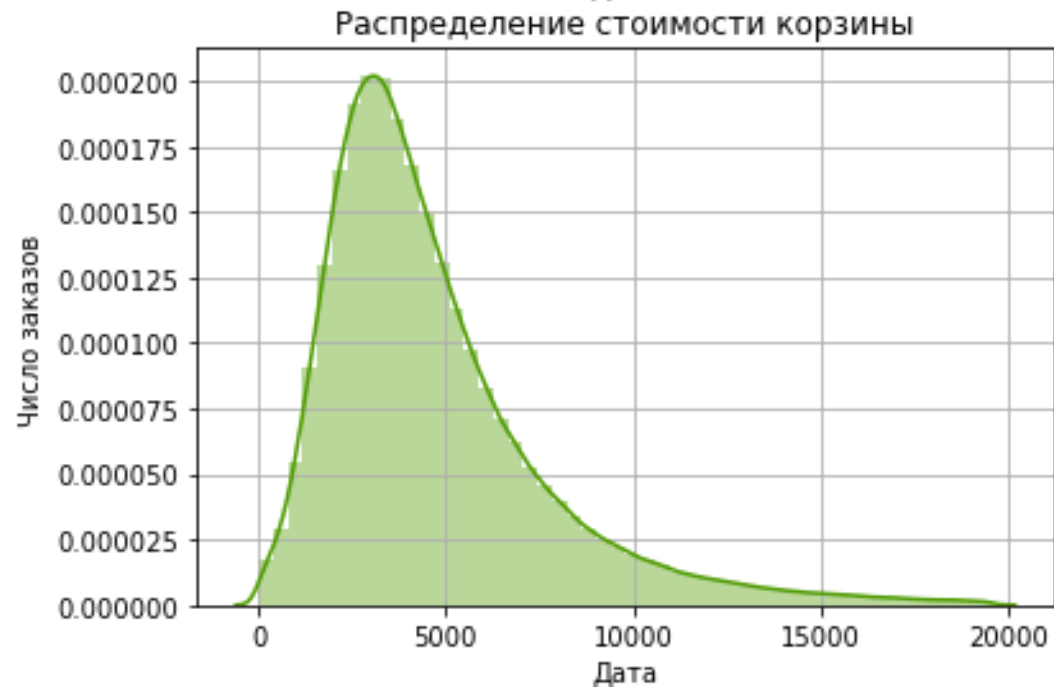
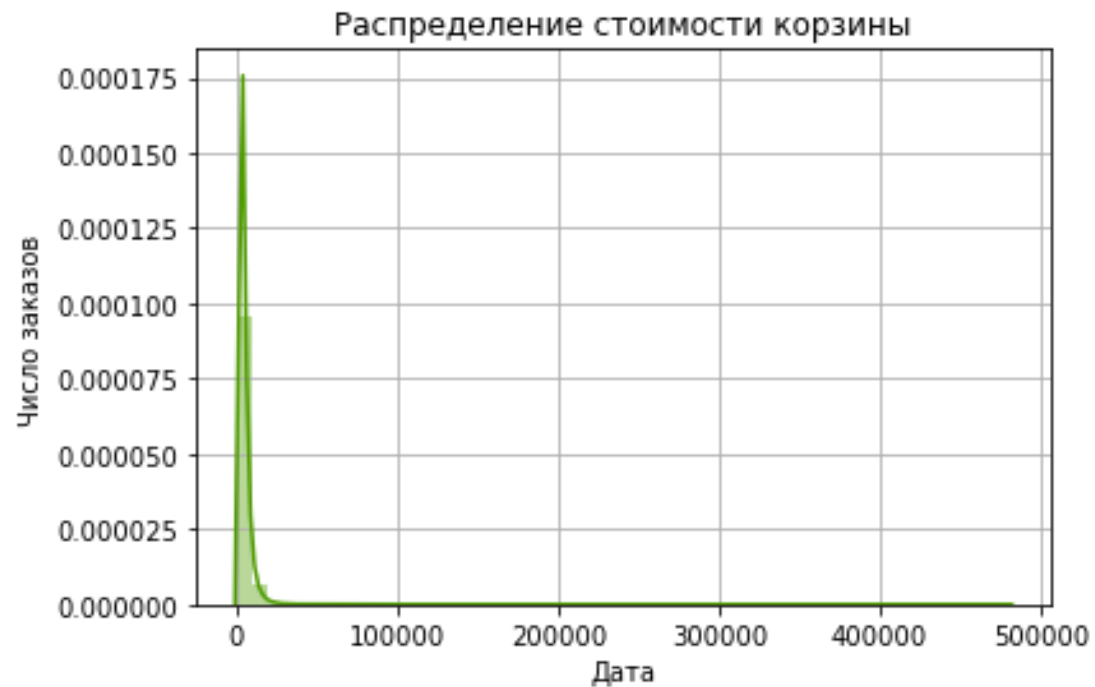


# Доставка

- В среднем от создания заказа до доставки проходит 4 дня 08:20 минут. Медианно доставка происходит за 1 день 01:18 час.
- В среднем от заявленного интервала курьеры опаздывают на 43:15 минуты. Медианно доставка опаздывает на 50:52 минут.



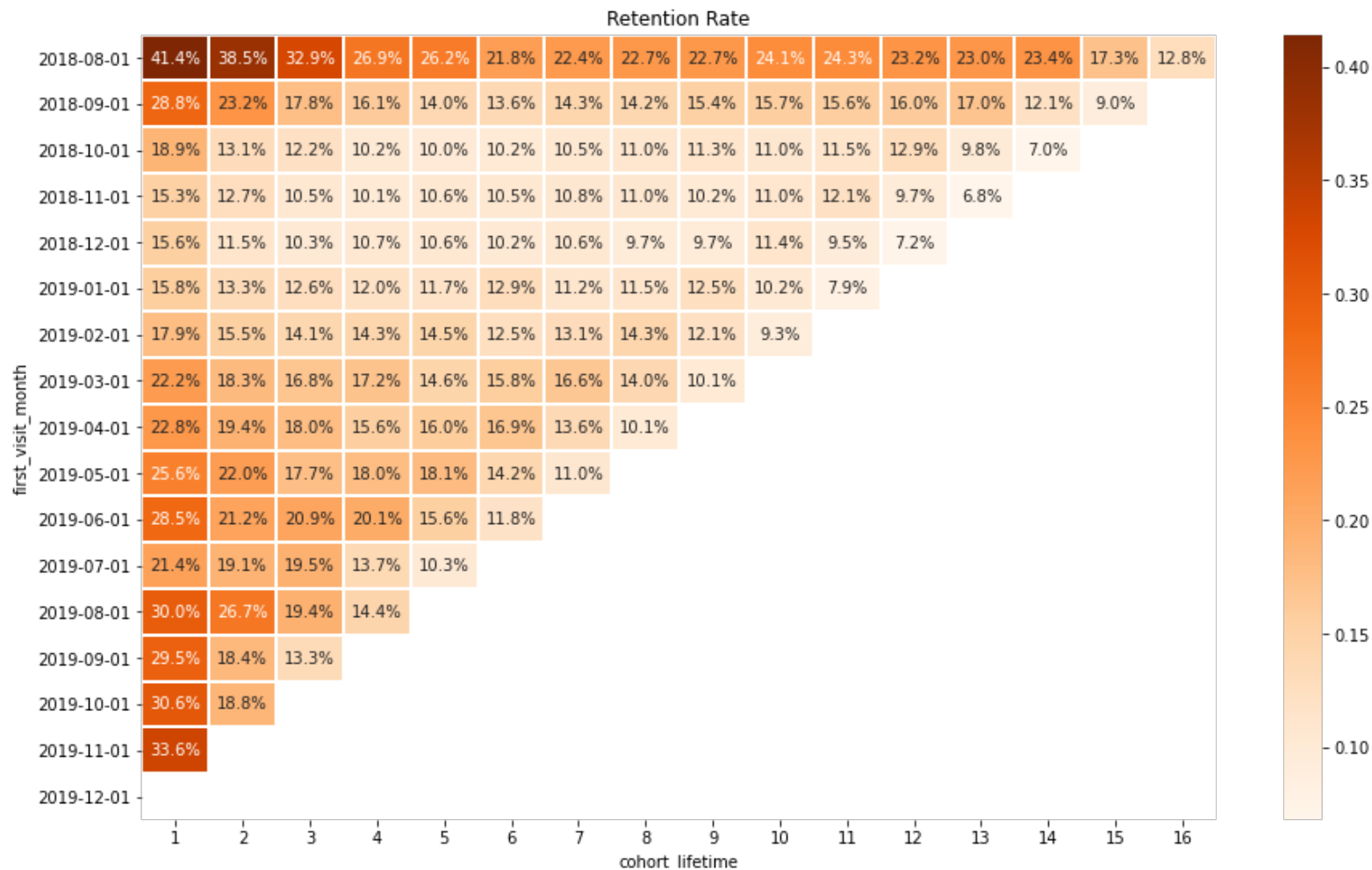




Имеем распределение стоимости корзины с длинным хвостом.

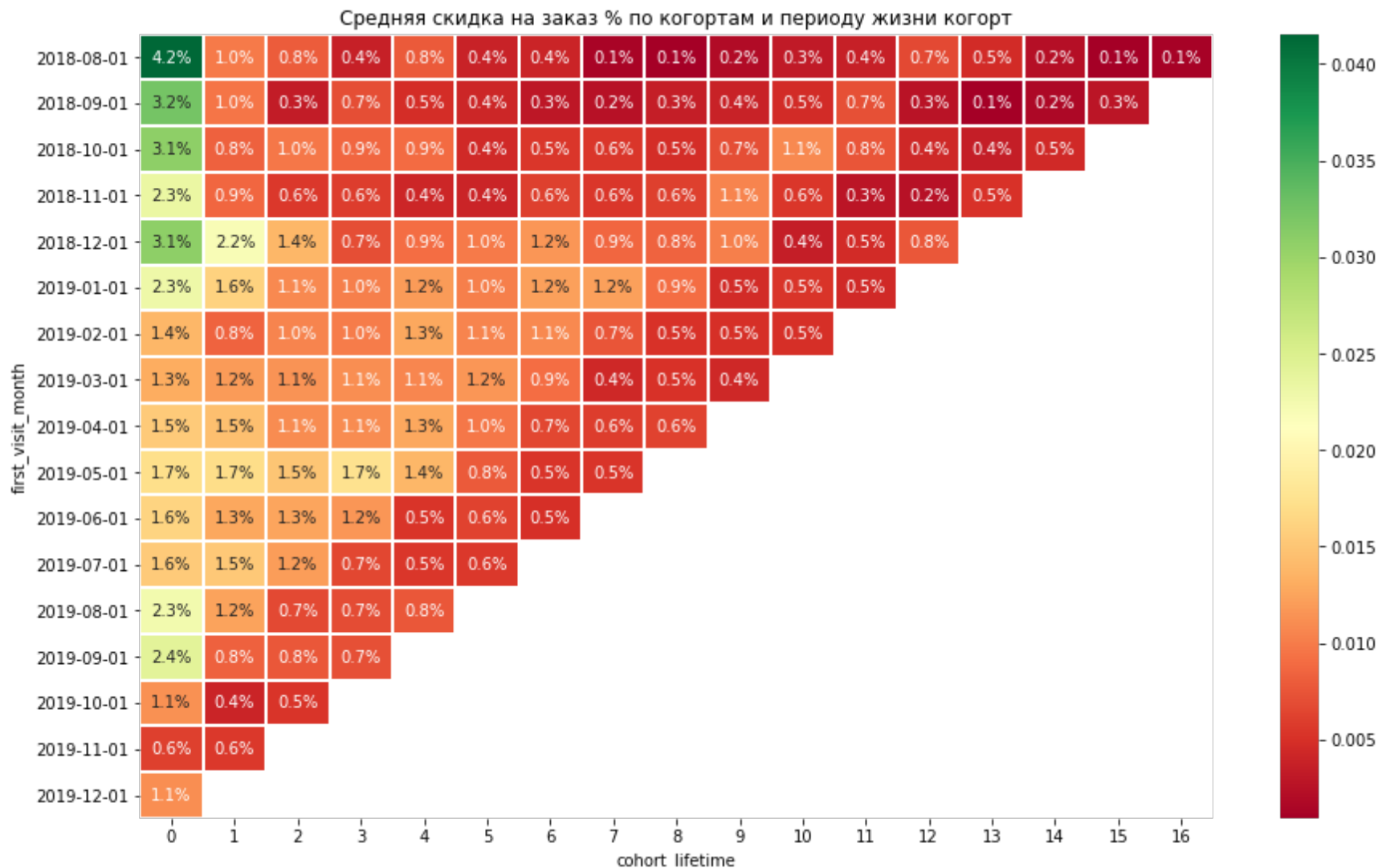
Суммарная стоимость корзин составила 87.32% от первоначальной при фильтрации 95 персентиля веса заказа(63957.8 г).

# Retention Rate



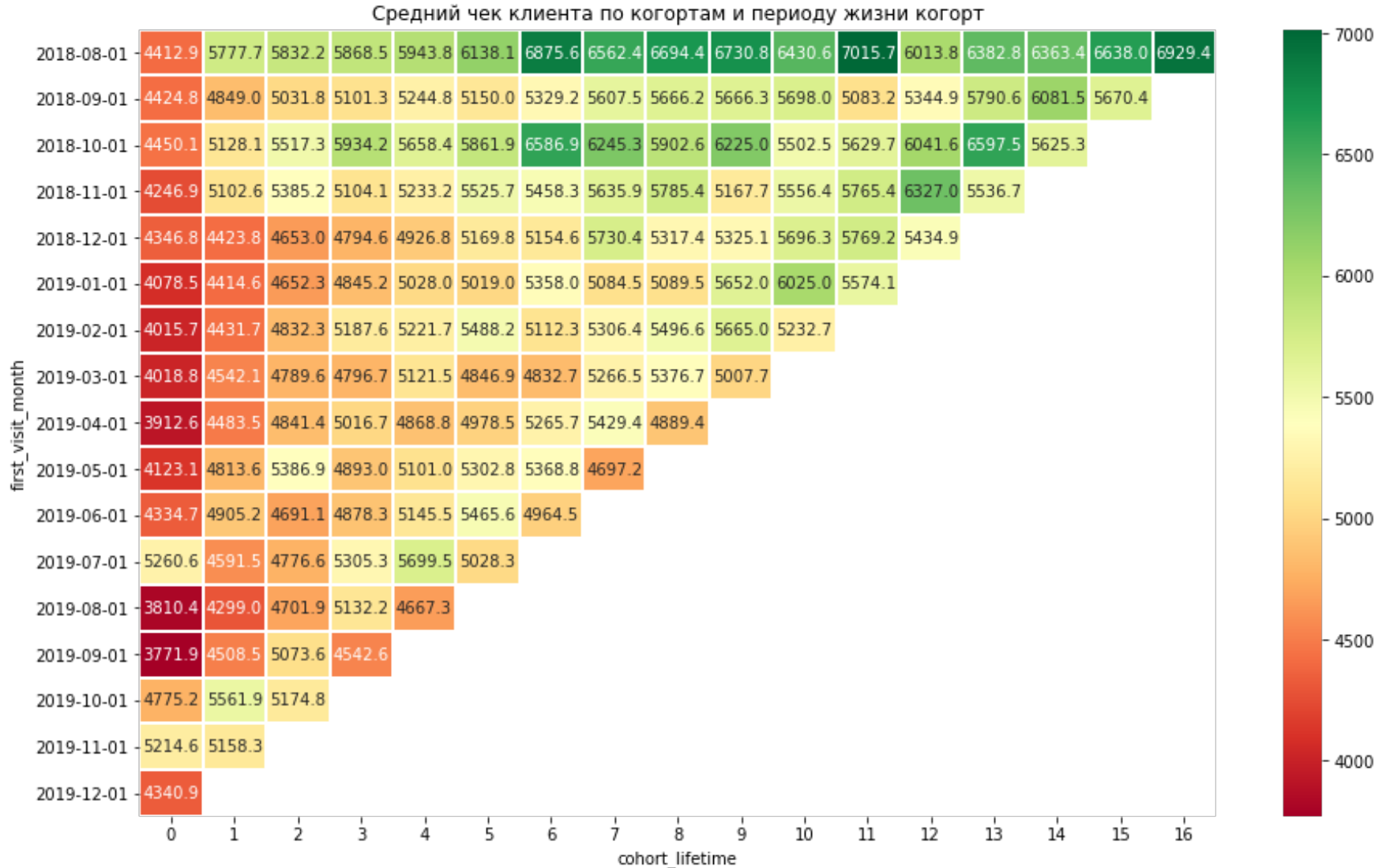
Выделим августовскую когорту, в ней самые высокие RR на протяжении всей жизни, 41.4% клиентов вернулись во второй месяц. Пользователи этой когорты лояльны к продукту, возможно, мы предоставляем им особые условия. Эти значения существенно отличаются от сентябрьской-январской когорты, в дальнейшем, видимо, была изменена стратегия и уже в ноябрьской когорте 2019 года мы получили RR 33.6%

# Скидка на заказ



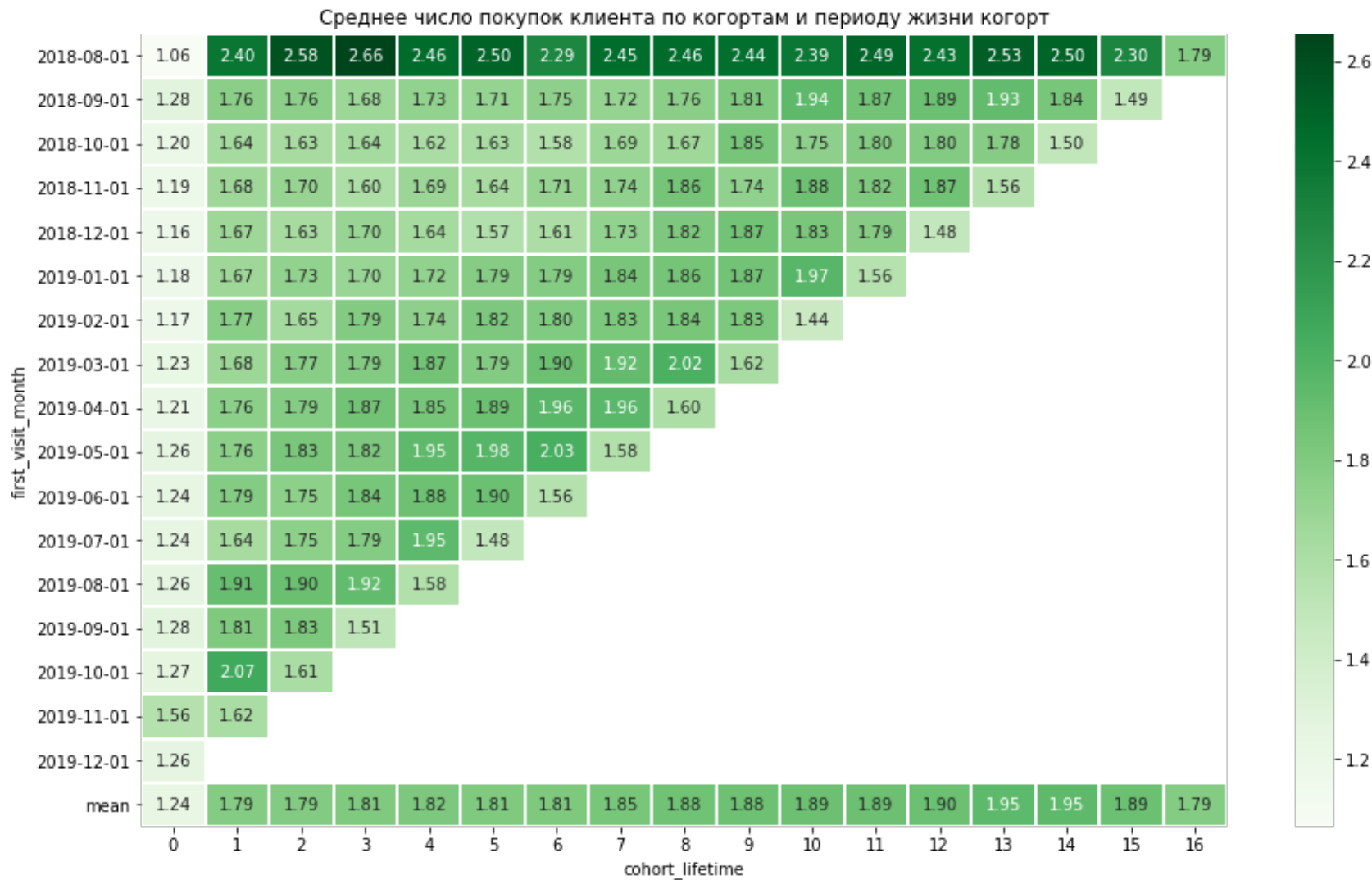
Зафиксировали повышенный процент скидки у августовской, сентябрьской и октябрьской когорт 2018 года. Возможно, повышенный процент скидки и экономия пользователей стали одним из факторов их лояльности. Со второго месяца мы не предоставляем существенных скидок, однако эффект, судя по удержанию сохраняется. Интересно, что ноябрьской когорте мы в среднем предложили низкую скидку, однако имеем высокий коэффициент удержания. Вероятно, есть еще инструменты лояльности аудитории.

# Средний чек по когортам



Наблюдаем рост среднего чека во всех когортах на периоде жизни когорт, средний чек самой лояльной нашей когорты выбивается из остальных, имеем высокий средний чек в августовской корте 208 года. Также видим сниженный средний чек у августовской и сентябрьской когорте 2019 года и повышенный у июньской и ноябрьской 2019 года.

# Число покупок на клиента



Во всех когортах наблюдаем существенное повышение среднего числа покупок во второй месяц существования и медленного повышения на всем периоде жизни когорт, лояльная когорта 2018 года показывает высокое число покупок, в среднем больше 2 каждый месяц.

# Выводы

- Необходимо разобраться в причинах провала среднего чека и числа пользователей в августе и ноябре.
- Стоит посмотреть на сколько нагружены курьеры выполнением тяжелых заказов, можно ввести стандарт веса на курьера, также можно продлевать доставку для таких оптовых клиентов, либо просить их делать заказ заранее. Выполнение таких заказов не должно сказываться на доставке базовым пользователям. Существует вариант редактирования аудитории сервиса путем введения базового веса доставки и числа курьеров на заказ, так станет меньше оптовых заказчиков, но наша доставка может ускориться, за счет этого можно увеличить удовлетворенность пользователей и обрести новых.

# Выводы

- Выделили когорту 2018-08, в ней высокий коэффициент удержания пользователя, высокий средний чек и высокое число покупок на клиента, им была предоставлена повышенная скидка в первый месяц, это могло стать причиной такой лояльности, стоит также посмотреть применение на них других инструментов повышения лояльности аудитории и попробовать сделать наших новых клиентов лояльнее сервису.
- При отмене или замене товара в заказе мы изменяем 7.96% от количества товаров, значение небольшое, но это может стать негативным пользовательским опытом. Основная идея в предоставлении информации о том, что товары заканчиваются, для этого необходимо договориться с магазинами об обновлении базы. Хорошо показывать клиенту предупреждение, также можно предложить ему самому выбрать продукт замены или вариант отмены товара. Есть вариант добавить бонусы в качестве извинений за то, что мы не смогли забрать товар, можем измерять удовлетворенность пользователей и на основании этого принимать решения. Также хорошо предупредить пользователя о поступлении товара, так мы сможем получить дополнительный заказ пользователя.



# Особенности данных

- У нас есть 17660 дублей в таблице `addresses`, дублирующие записи добавляются не подряд и имеют одинаковые `id`. Возможно, наша проблема в базе или при редактировании адреса, но не изменении его, мы все равно добавляем запись в базу с прежним идентификатором. Стоит проверить алгоритм добавления нового адреса.
- В данных обнаружены пропуски, 46 значений отсутствует в колонках координат доставки (`lat,lon`), возможно, это объясняется последующим удалением адреса доставки пользователем. 4 пропуска в колонках интервала доставки (`starts_at,ends_at`) могут объясняться сбоем, либо удалением из базы таких интервалов.
- В `total_cost` имеем подозрительные значения с отрицательной стоимостью доставки, при этом промокоды не применялись, это могла быть ошибка при добавлении в базу данных, знак был не учтен.