1. Проверьте возможные ошибки в данных.

* Указаны даты покупок после текущей
* Есть заказы с нулевой или отрицательной стоимостью (44 шт.)

1. Проанализируйте, какой период данных выгружен.

Диапазон дат от 2031-12-20 до 2001-01-20

1. Посчитайте количество строк, количество заказов и количество уникальных пользователей, кто совершал заказы.

Количество строк: 2002804 строк

Количество заказов: 2002804 заказов

Количество уникальных пользователей: 1 015 119

1. По годам и месяцам посчитайте средний чек, среднее количество заказов на пользователя. Сделайте вывод, как изменялись эти показатели год от года.

* По среднему чеку год от года наблюдается определенная сезонность с увеличением среднего чека в период Августа – Ноября, и небольшой просадкой в марте - мае
* По удельному количеству заказов картина год от года существенно не меняется. Распределение внутри года показывает что в зонах с большим количеством заказов количество заказов в пересчете на человека – уменьшается.

1. Найдите количество пользователей, кот покупал в одном году и перестал покупать в следующем.

Таких пользователей 1 015 119 (это все пользователи!)

Код SQL:

**select** **count**( **distinct** user\_id) **from** orders

**except**

**select** **distinct** user\_id

**from** orders o1

**where** **exists**(**select** 1 **from** orders o2

**where** o2.user\_id = o1.user\_id

**and** **extract**(**year** **from** o2.o\_date) + 1 = **extract**(**year** **from** o1.o\_date)

)

1. Найти ID самого активного по кол-ву покупок пользователя.

Самый активный пользователь: User\_id 765861 3183 заказов

Код SQL

**select** user\_id, **count**(id\_o) **from** orders **group** **by** user\_id **order** **by** **COUNT**(id\_o) **DESC**

1. Найдите коэффициенты сезонности по месяцам.