

ДЗ №1. Анализ проводился по файлу: orders_20190822

2. Проанализировать, какой период данных выгружен:

- 2 года: с 01.01.2016 по 31.12.2017

3. Посчитать кол-во строк, кол-во заказов и кол-во уникальных пользователей, кот совершили заказы.

- Количество строк: 2 002 804

- Количество заказов: 2 002 804

- Количество уникальных пользователей: 1 015 118. Не стал учитывать пользователя с 0 id, посчитал это ошибкой в данных

4. По годам посчитать средний чек, среднее кол-во заказов на пользователя, сделать вывод , как изменились это показатели Год от года

- Средний чек в 2016 году – 2100,21

- Средний чек в 2017 году – 2398,62

В расчетах игнорировал заказы с суммой меньше 50 (поскольку исследуем интернет-магазин, а заказы с суммой более 50, например 100/200/300 приводят к значительному росту выборки, которую мы не можем игнорировать), принимая во внимание, что данных заказов 3,3 тыс. на более, чем 2 млн записей. Из расчетов были исключены выбросы по 2016 году, 2 значения : 819 096,6 и 616 994. Из расчетов были исключены выбросы по 2017 году, 3 значения: 540 482,6 , 521 336,2 и 392 205,1

- Среднее количество заказов на пользователя в 2016: 1,93
- Среднее количество заказов на пользователя в 2017: 1,74

При расчете среднего количества заказов на пользователя аналогично расчету среднего чека, были исключены: заказы с выбросами по цене, заказы с суммой менее 50.

Учитывая, что среднее количество заказов на пользователей упало (при это количество пользователей и количество заказов возросло), при одновременном росте среднего чека, могут быть сделаны следующие выводы:

- пользователи стали покупать больше одних и тех же товаров в рамках одного заказа;
- пользователи сменили предпочтения в пользу более дорогих аналогов;
- стоимость товаров возросла

Однако, учитывая, что 81% пользователей, которые покупали в 2016, перестали покупать в 2017 году(см пункт 5), при этом рост пользователей в 2017 составил 47%, возможно, речь идет о переориентировании на другой сегмент рынка с более высокой потребительской способностью.

5. Найти кол-во пользователей, кот покупали в одном году и перестали покупать в следующем.

360 225 покупали в 2016 и перестали покупать в 2017

6. id самого активного пользователя

Наиболее активный пользователь по стоимости заказа: id 1959313 – сумма заказа: 616 994, однако, заказ всего 1, что может указывать на фейковые данные (ошибки в базе, брошенную корзину прочее).

В базе также имеется еще порядка 5 пользователей с большими суммами заказов, при этом каждый из них совершил всего 1 заказ.

Поэтому наиболее активным пользователем признан: id 4243798 – сумма заказов: 196 140, количество заказов - 2.