

Задание:

- К вам приходит руководитель продукта “Интернет-магазин”
- И спрашивает про удержание клиентов и повторные покупки
- У него есть опасение, что из-за проблем с доставкой товаров в магазине у нас стали больше оттекать пользователи
- Он просит вас сравнить, на сколько сейчас все хуже, чем в прошлом году?

Данные лежат в папке занятия на сервере:

- есть транзакции (покупки) в интернет магазине, в том числе id клиента, id покупки, timestamp покупки и тп
- orders.csv, customers.csv

Ваши задачи:

1. Проверить гипотезу руководителя продукта
2. Дать руководителю продукта инструмент для самостоятельного принятия решений по качеству удержания пользовательской базы

Решение:

Первое, что приходит в голову – проверить, как изменялось количество заказов в разрезе лет.

	123 year	123 sum_orders
1	2 014	969
2	2 015	1 038
3	2 016	1 315
4	2 017	1 687

На первый взгляд, количество заказов растет с каждым годом.

Поскольку нас интересует этот (2017) и прошлый (2016) года, то возьмем список уникальных пользователей с 2014 по 2015 год и проверим, сколько из них совершали покупки в 2016 году, а также число вновь прибывших пользователей. Аналогичное сделаем для 2017 года.

	Число пользователей, повторно сделавших покупку	Число новых пользователей
2015	807	231
2016	1211	104
2017	1668	19

Хм, с одной стороны, в годовом разрезе оттока пользователей не наблюдается. Однако, число новых пользователей заметно снижается. Может сарафанное радио разнесло весть о нашей ужасной доставке? Сделаем более углубленный анализ. Помимо годового разреза, рассмотрим, какой тип доставки выбирали пользователи, делавшие покупки повторно, а также новые пользователи.

		Число пользователей, повторно сделавших покупку	В процентном соотношении ко всем повторным пользователям	Число новых пользователей	В процентах к новым пользователям
2015	First	112	0,138	31	0,134
	Second	157	0,195	49	0,212
	Standard	495	0,613	141	0,610
	Same Day	43	0,053	10	0,043
2016	First	197	0,163	16	0,154
	Second	222	0,183	22	0,212
	Standard	720	0,595	62	0,596
	Same Day	72	0,059	2	0,019
2017	First	280	0,168	4	0,210
	Second	321	0,192	3	0,157
	Standard	978	0,586	12	0,632
	Same Day	89	0,053	0	0

Доставкой типа «First» пользуются все больше людей, полагаю, что с ней все в порядке. Доставкой типа «Second» с переменным успехом пользуются наши постоянные клиенты, а вот новые пользователи отдают предпочтение другому типу доставки.

Доставкой типа «Standard» пользуется большинство и постоянных, и новых пользователей.

Доставкой типа «Same Day» пользуются наименьшее количество пользователей, а из новых практически никто. Возможно, это связано с тем, что доставка в тот же день осуществляется только в город, в котором расположен магазин (если магазин расположен только в одном городе), а возможно в ней есть проблема. Выясним это!

Итак, доставку «Same Day» выбирали люди из 108 различных городов. Следовательно, если есть проблема с доставкой, то это проблема с доставкой «Same Day».

Рассмотрим проблему не со стороны доставки, а со стороны категории пользователей. Составим для них такую же таблицу:

		Число пользователей, повторно сделавших покупку	В процентном соотношении ко всем повторным пользователям	Число новых пользователей	В процентах к новым пользователям
2015	Consumer	428	0,530	107	0,463
	Corporate	237	0,294	75	0,325
	Home Office	142	0,176	49	0,212
2016	Consumer	593	0,490	58	0,558
	Corporate	390	0,322	32	0,308
	Home Office	228	0,188	14	0,135
2017	Consumer	863	0,517	13	0,684
	Corporate	488	0,293	5	0,263
	Home Office	317	0,190	1	0,053

Приток новых пользователей значительно снижается во всех категориях.

Численно количество постоянных клиентов увеличивается, то есть те, кто приходят, все-таки остаются :) Но из процентного соотношения видно, что количество Consumer растет сильнее.