

Анализ программы лояльности Строительного ритейла



Выпускной проект. DA_59

2023г.

© Кирилл Непомнящих

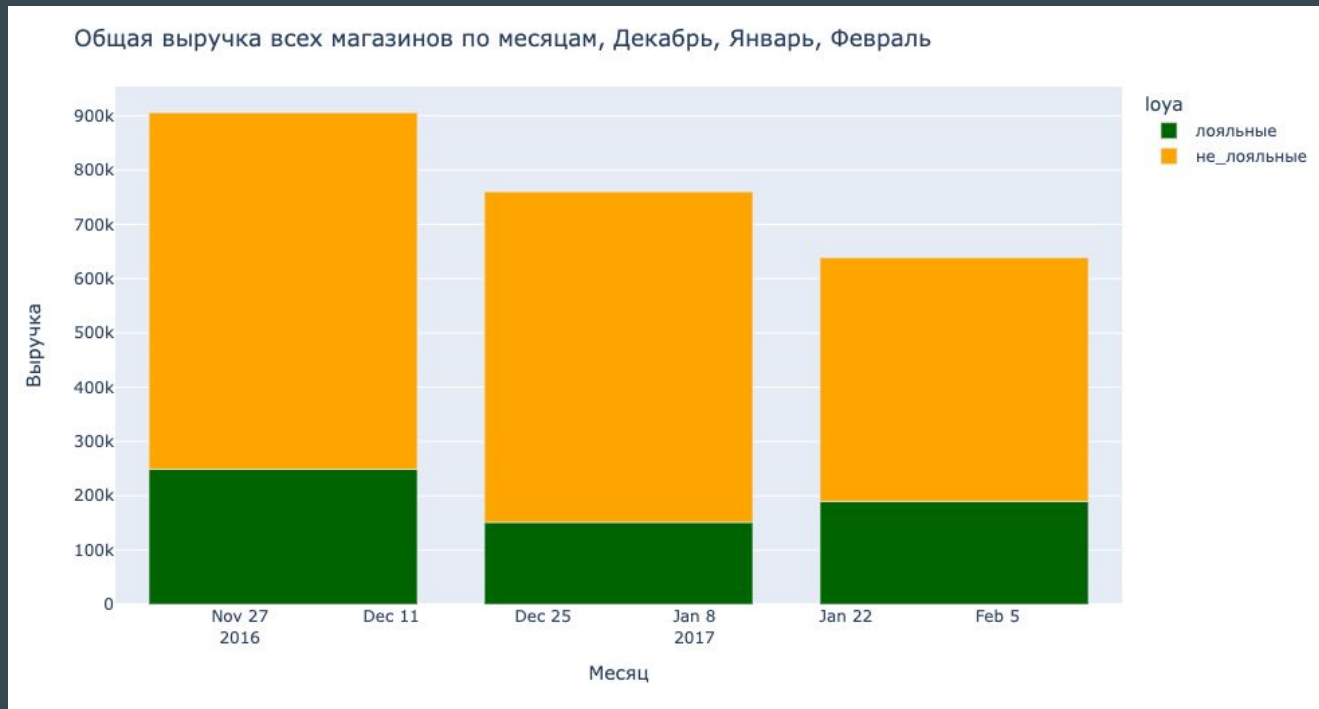
Цель и задача проекта

- Провести анализ программы лояльности;
- Проверить статистические гипотезы.
 - 1. Сумма чека участников программы лояльности отличается от среднего чека остальных покупателей
 - 2. Кол-во чеков на участника программы лояльности отличается от кол-ва чеков остальных покупателей
 - 3. Среднее кол-во позиций товара в чеке участников программы лояльности отличается от среднего кол-ва позиций товаров в чеке остальных покупателей

Вывод:

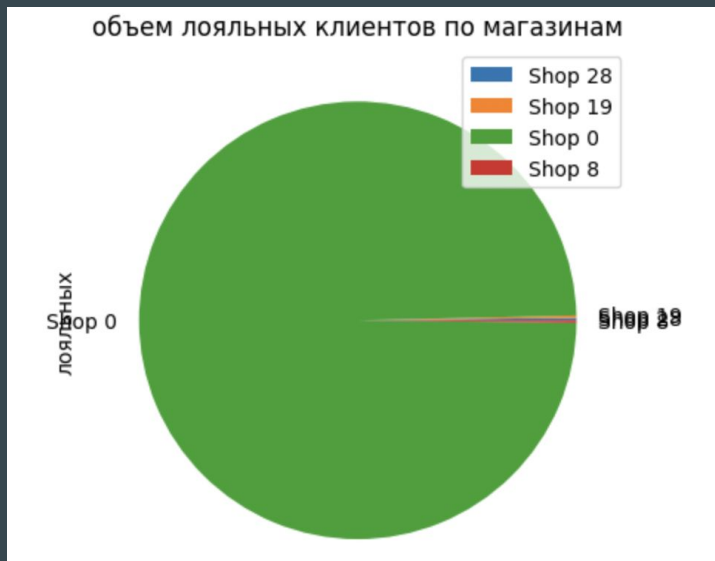
- Лояльные и нелояльные покупатели в среднем приносят выручку одинаковое кол-во на покупателя.
- При этом, Лояльные делают больше чеков на покупателя группы
- Кол-во позиций товаров в чеке также не имеет различий между участниками программы и остальных

Итоги исследования:



- Данные по магазинам представлены за 3 месяца (Декабрь, Январь, Февраль).
Магазин скорее всего находится в Великобритании

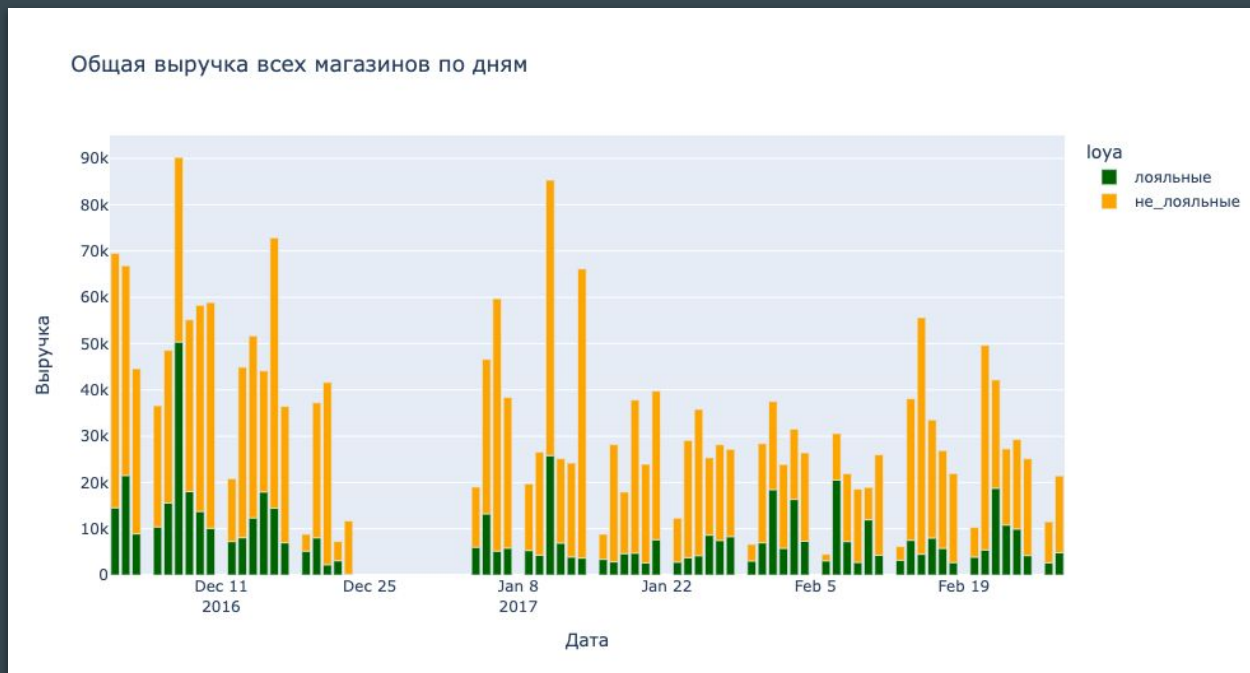
Итоги исследования:



shop_id	всего	лояльных	доля_лояльных	доля_не_лояльных
Shop 28	1	1	100.000000	0.000000
Shop 19	3	1	33.333333	66.666667
Shop 0	1873	558	29.791778	70.208222
Shop 8	11	1	9.090909	90.909091
Shop 23	1	0	0.000000	100.000000

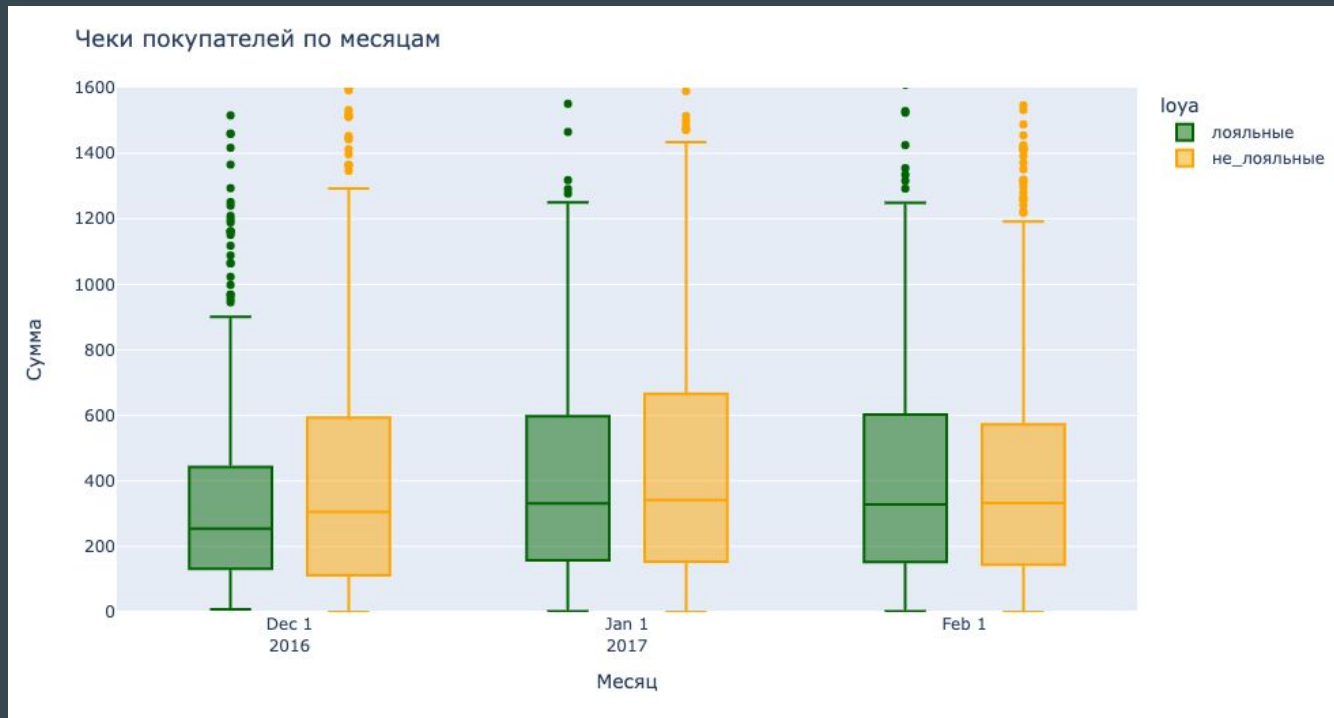
- Анализ программы лояльности показал что программа действует в 4х магазинах. Из них 3 магазина имеют малое кол-во покупок.
- Выделяется Shop 0, его мы и взяли для анализа программы

Итоги исследования:



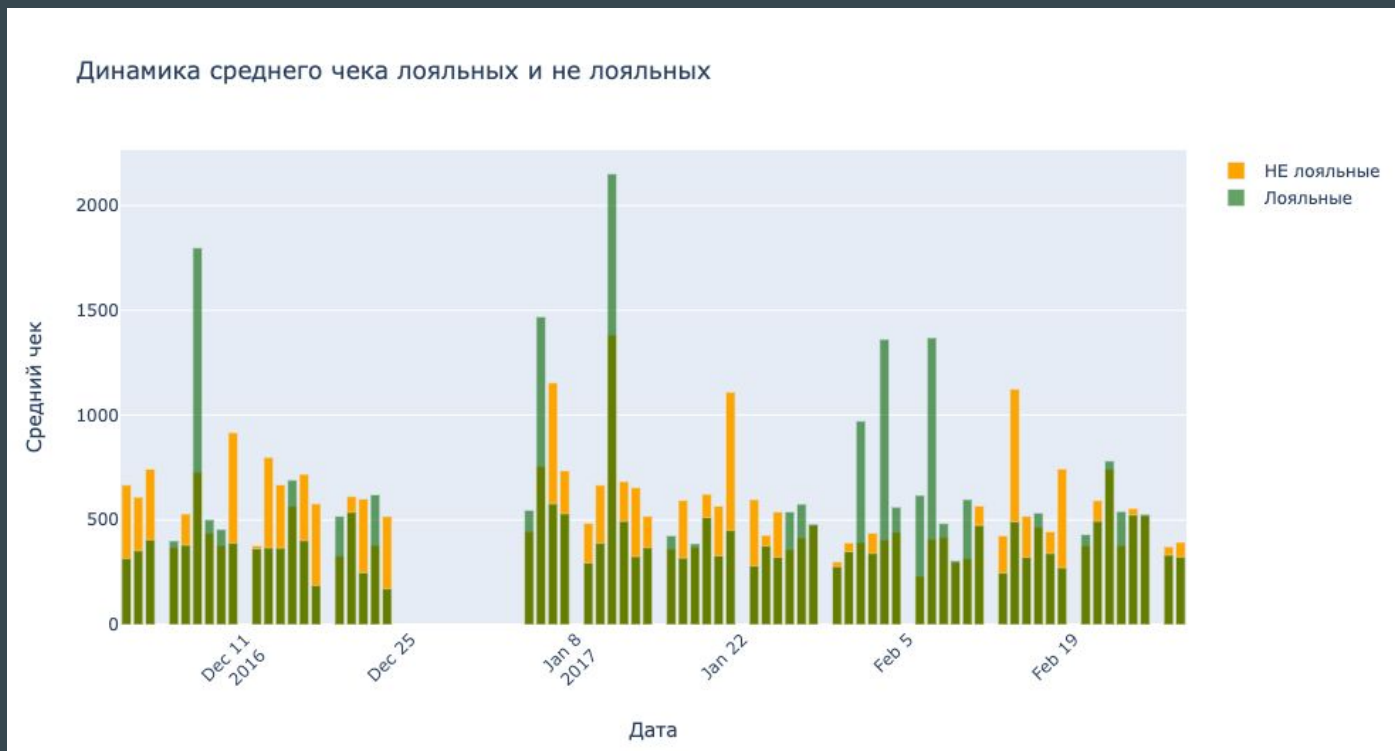
Исходя из исследования, а также временной период за который предоставлены данные, делаем вывод что пользователи программы лояльности делают выручки меньше чем остальные. Для подтверждения проверили гипотезы и выяснили:

Итоги исследования:



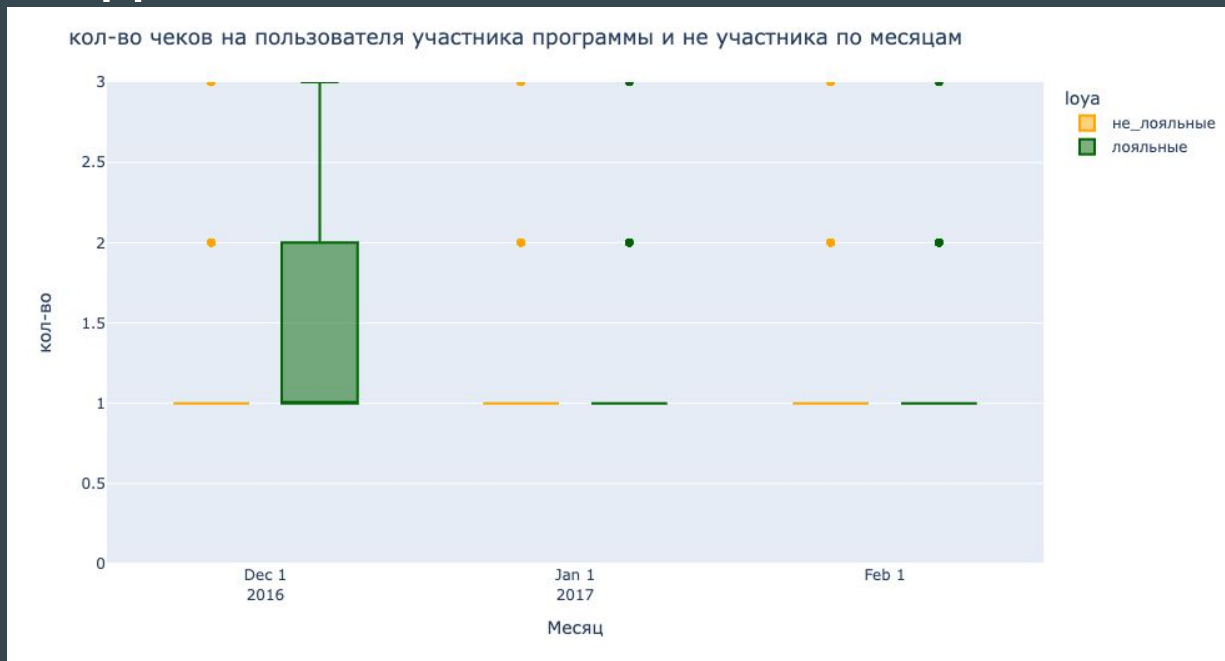
- Лояльные покупатели приносят выручки по медиане чеков меньше чем НЕ лояльный

Итоги исследования:



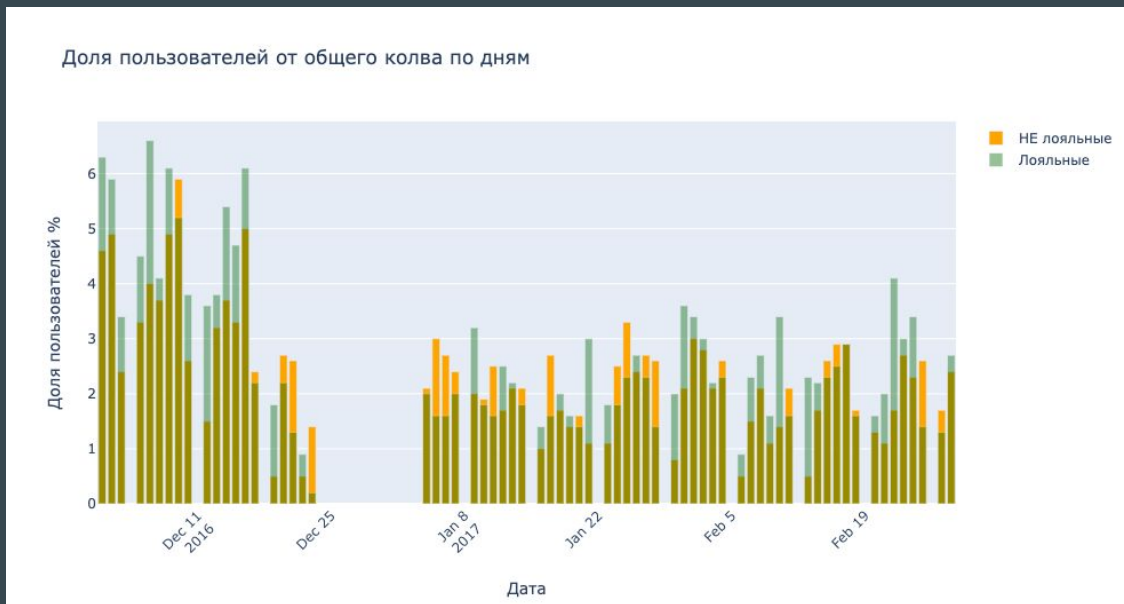
- Динамика по средним показателям имеет резкие движения у лояльных. Дни акций, скидок персональных. В остальные дни, лидирует группа НЕ_Лояльные

Итоги исследования:



- медианное кол-во чеков у лояльного и не лояльного покупателя = 1 чеку.
 - Но в декабре Лояльные сделали больше чеков
- Есть статистические различия по количеству чеков на покупателя среди групп Лояльные и Не лояльные

Итоги исследования:



- В преддверии нового года, ввиду специальных предложений для лояльных покупателей, Лояльные покупатели активнее НЕ лояльных (Относительно своей группы)
- В январе ситуация меняется на противоположную, и в феврале активность лояльных опять берет верх

Гипотезы и проверка гипотез

Гипотеза 1: Чеки участников программы лояльности

- H_0 : Не отличается от среднего чека остальных покупателей
- H_1 : Отличается от среднего чека остальных покупателей.

Проверка:

- Критический уровень статистической значимости $\alpha = 0.05$
- p -значение: 0.07017049025923001
- Не получилось отвергнуть нулевую гипотезу, статистически значимых различий нет

Гипотезы и проверка гипотез

Гипотеза 2: Кол-во чеков на участника программы лояльности

- H_0 : Не отличается от кол-ва чеков остальных покупателей
- H_1 : Отличается от чеков покупателей не участвующих в программе лояльности

Проверка:

- Критический уровень статистической значимости $\alpha = 0.05$
- p -значение: $4.94766718518552e-12$
- Отвергаем нулевую гипотезу, есть статистически значимые различия
- среднее кол-во всех чеков на покупателя ЛОЯЛЬНЫЕ больше чем среднее кол-во всех чеков на покупателя НЕ_ЛОЯЛЬНЫЕ на 15.8 %

Гипотезы и проверка гипотез

Гипотеза 3: Среднее кол-во позиций товара в чеке участников программы лояльности

- H_0 : Не отличается от среднего кол-ва позиций товаров в чеке остальных
- H_1 : Больше чем позиций товаров в чеке остальных покупателей

Проверка:

- Критический уровень статистической значимости $\alpha = 0.05$
- p -значение: 0.9465697291410925
- Не получилось отвергнуть нулевую гипотезу, статистически значимых различий нет

Рекомендации

- В данных присутствуют покупатели, которые возвращали товар, купленный ранее. В результате учитывая возврат товара - по некоторым клиентам выручка отрицательная. В результате данные рассчитаны без учета возврата товара.
- Отсутствует информация как менялась цена прайса, в результате в расчете принята медиана цена товара. Что искажает результат. Цена товара для лояльных и нелояльных клиентов принята одной и той же. У меня есть предположение, что для лояльных клиентов применялись скидки. И Если учесть дисконт для покупателей участвующих в программе, то показатели такие как выручка для Лояльных может оказаться еще меньше.
- За 3 месяца нельзя сделать выводы о том, была ли эффективна или неэффективна программа. Не понятно когда была первая покупка. Если участники программы лояльности в эти 3 месяца впервые стали клиентами магазина, то данные показатели могут быть вполне успешными.

Источник данных

- Выгрузка операций торговой сети
 - Файл данных retail_dataset.csv
 - Файл product_codes.csv