

# Подготовка и анализ A/B теста

Проект спринта №7  
Студент: Гольцов Игорь

# Содержание

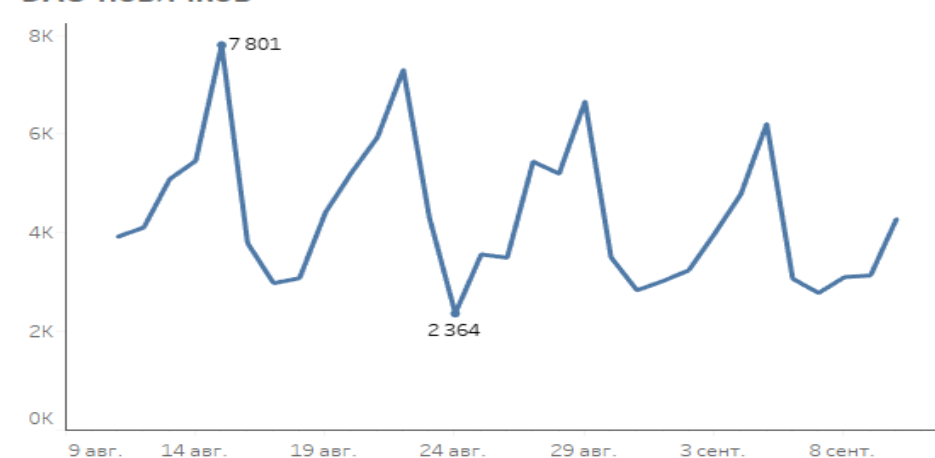
- Подготовка теста (слайды 3 - 6)
- Техническое задание (слайд 7)
- Проведение теста (слайд 8)
- Анализ результатов (слайды 9 – 13)

Ссылка на дашборд: [Подготовка и проведение АВ теста](#)

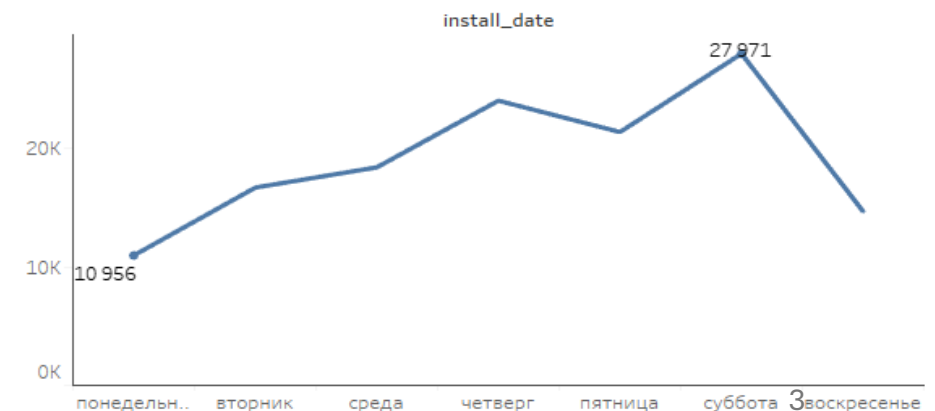
# Подготовка теста

- Среднее количество привлеченных пользователей в день: 4324
- На графике динамики наблюдается сезонность: в течении недели от понедельника число привлеченных пользователей растет и достигает пика в суббота

DAU новичков

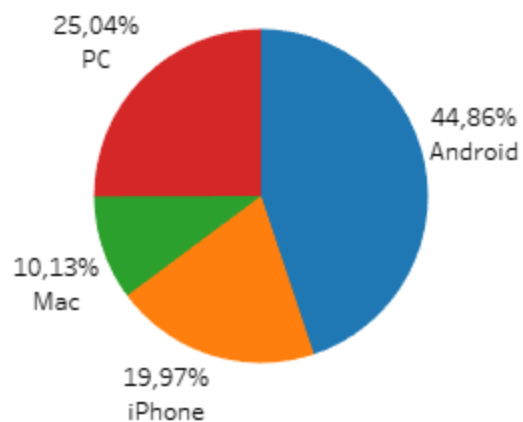


DAU новичков по дням недели

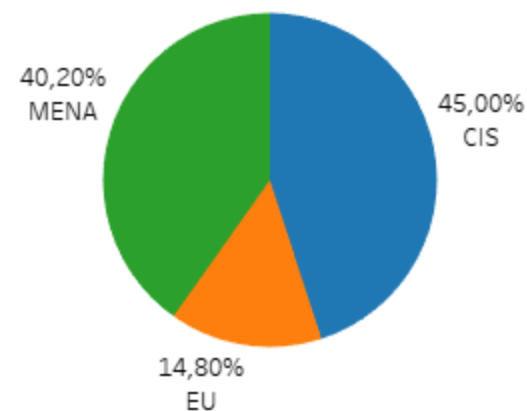


# Подготовка теста

Новички по устройствам



Новички по регионам



- По типу устройства, с большим отрывом, лидируют пользователи Android (45%).
- По регионам почти одинаковую долю имеют регионы CIS (45%) и MENA (40%), EU отстает с долей 15%

# Подготовка теста

Показатель конверсии в покупку ноутбука по историческим данным составляет **0,5%**

# Подготовка теста

Распределение выручки имеет форму Гамма-распределения, с удлинением вправо хвостом, что говорит о большом кол-ве значений (выбросов), отличающихся от среднего.

Параметры распределения:

- Среднее значение: 5421
- Медианна: 3649
- СКО: 4697

Сильное отличие среднего значения и медианного, так же говорит о большом разбросе.

Гистограмма выручки

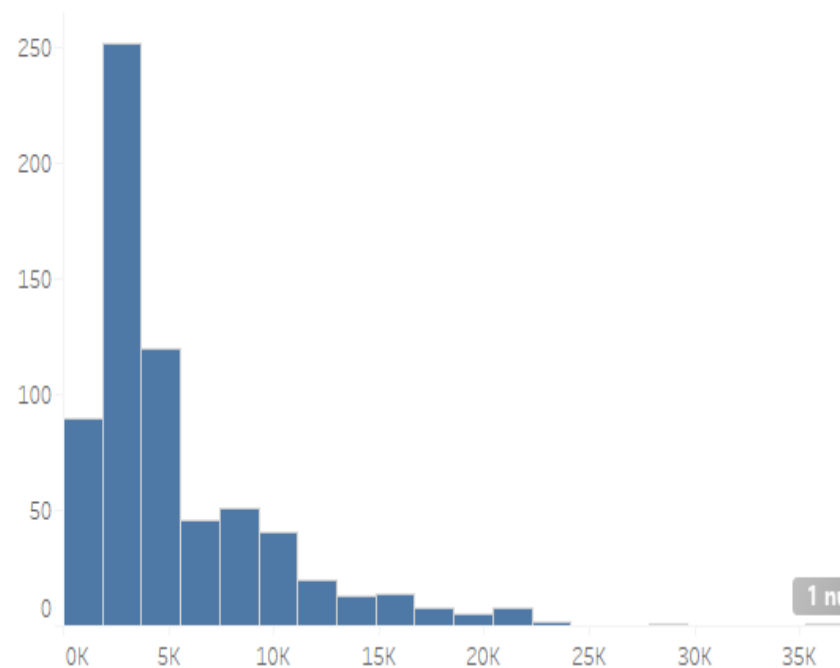
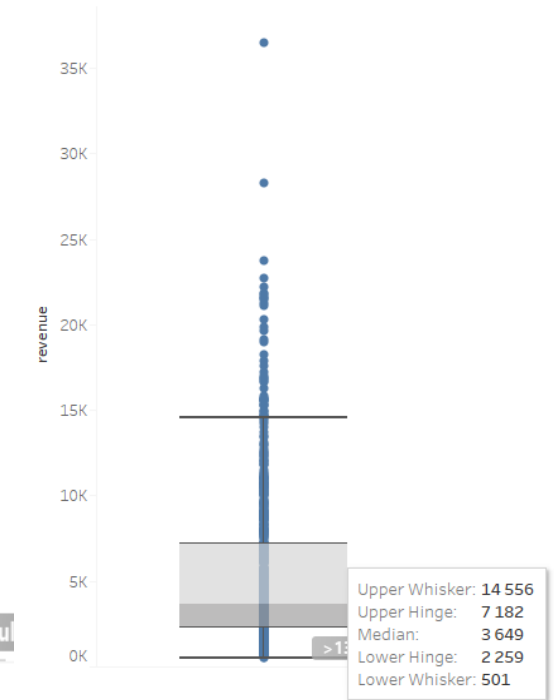


Диаграмма размаха выручки



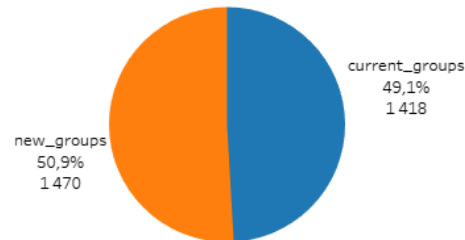
# Техническое задание АВ теста

- Название теста: "gaming\_laptops\_test"
- Назначение теста: проверка гипотезы влияния добавления отдельной категории с игровыми ноутбуками на целевые метрики
- Тестируемые метрики: конверсия в покупку ноутбука (ожидается рост на 100%), средний чек в категории «Компьютеры» (ожидается, что не поменяется)
- Число тестовых групп, их названия, доля пользователей для каждой группы: 2 группы (контрольная - "current\_groups" и тестовая "new\_groups") с одинаковым числом пользователей и долями
- Длительность теста и нужное число пользователей:
  - Для статистической значимости необходимо минимум 4716 пользователей в каждой группе.
  - Длительность теста 1 неделя - при необходимом кол-ве 4716 на каждую группу (всего 9432) и ежедневном кол-ве новых пользователей 4673, достаточно будет 1 недели, что бы набрать необходимое кол-во и при этом учесть сезонность всех дней недели.
- Дата запуска теста: 14.10.2020 (запуск теста планируется после проведения праздника "Chinese Moon Festival", для исключения влияния праздника на поведение пользователей, при этом остается достаточно времени для подготовки.
- Дата остановки теста: 20.10.2020

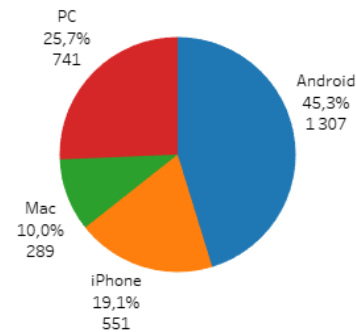
# Проведение теста

- На момент среза приняло участие 2888 пользователей
- Согласно Z тесту пропорций, нулевая гипотеза о том, что группы распределены равномерно не отвергается (  $p\text{-value} = 0.33338$ , что выше уровня значимости)
- Существенных отличий в распределении участников по регионам и устройствам от исторических данных не выявлено
- Необходимое количество участников еще не было набрано, следовательно тест продолжается

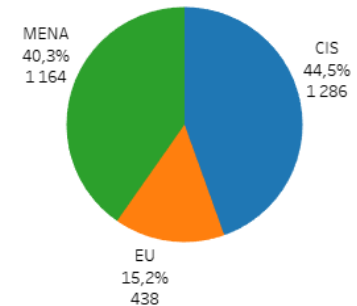
Баланс групп



Участники по устройствам



Участники по регионам

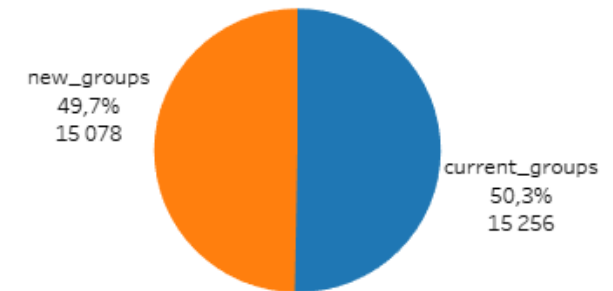




# Анализ результатов теста

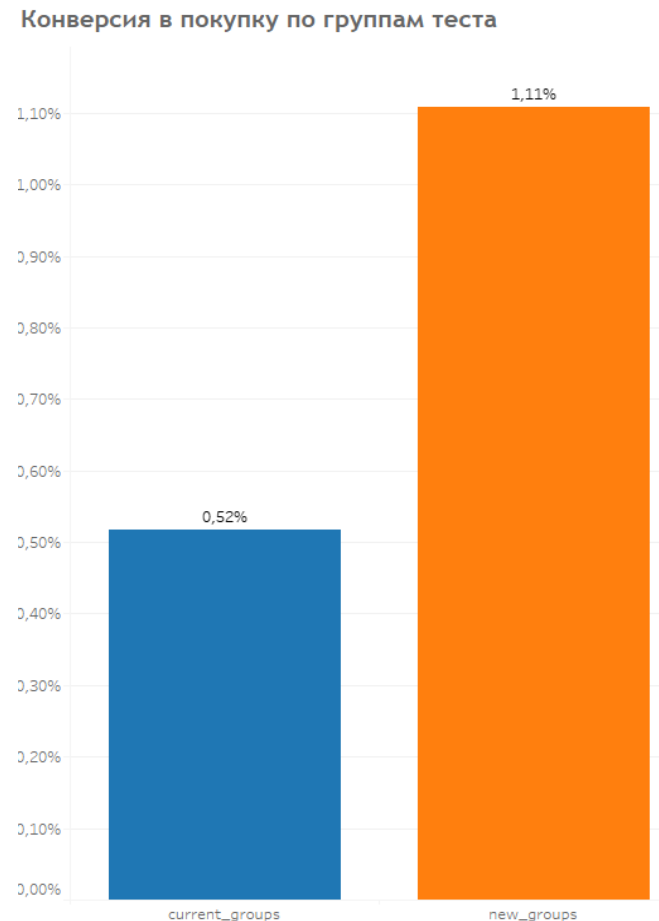
- Общее количество пользователей, принявших участие в тесте, составило 30334. Это достаточно для статически значимого сравнения.
- Распределение равномерное (согласно z тесту пропорций)

Размеры групп



# Анализ результатов теста

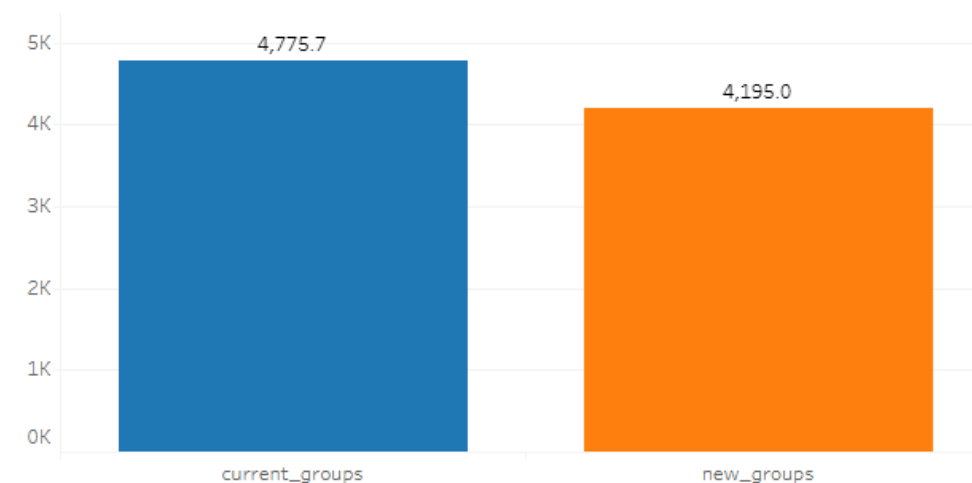
- Конверсия в покупку в тестовой группе, более чем в два раза, превышает конверсию в контрольной группе. (1,11% vs 0,52%)
- Согласно Z тесту пропорций p-value составило 0.00000, что может говорить, что результаты статистически значимы
- По части конверсии цель теста выполнена, т.к. превышение более чем в два раза



# Анализ результатов теста

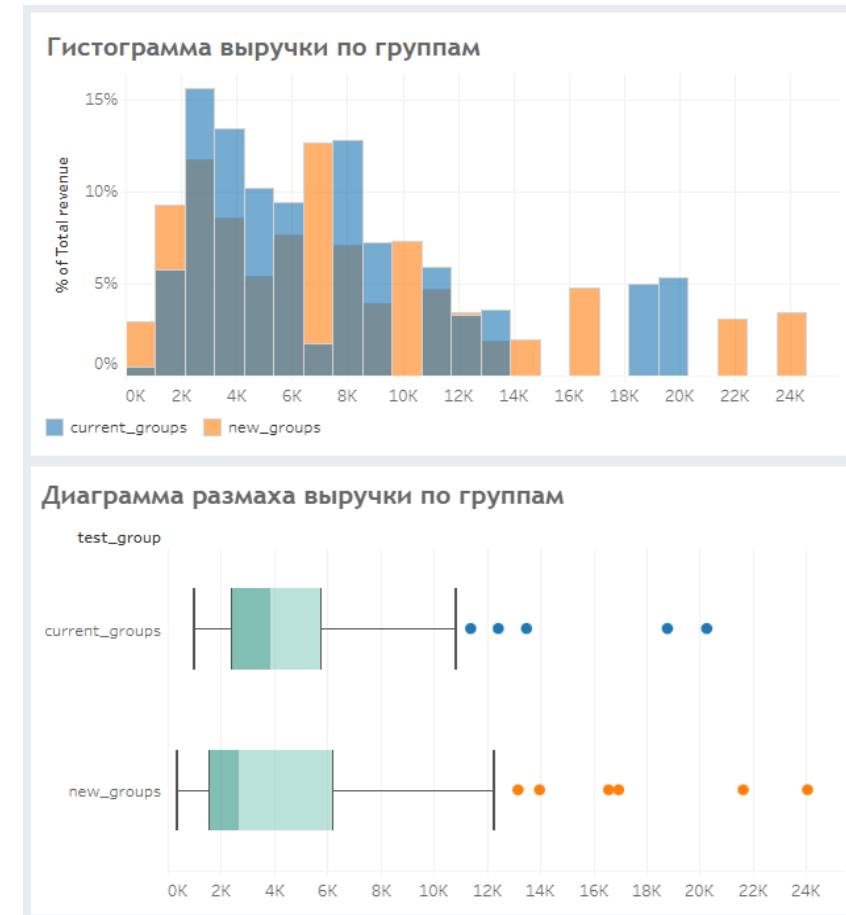
- Средний чек в тестовой группе оказался ниже (на 12%), чем в контрольной группе.
- Согласно t-тесту, p-value равняется 0.26938, т.е. средние чеки в обеих группах не отличаются
- Согласно тесту Манна-Уитни, p-value равняется 0.01016, что говорит о том, что данные получены из разных генеральных совокупностей
- Данные по тестам отличаются из-за существенного влияния выбросов
- По части среднего чека, ожидаемая цель не выполнена, т.к. наблюдается падение в показателе

Средняя выручка по группам теста



# Анализ результатов теста

- На диаграмме размаха наблюдаются выбросы в обеих группах.
- Медианное значения в обеих группах сильно отличаются. В тестовой группе медианное значение 2682 (против 3864 в контрольной группе), но за счет выбросов различие в средних значениях не такое явное
- Выбросы нельзя игнорировать, т.к. они не случайны, а показывают покупателей, совершивших, самые дорогие покупки.



# Бизнес-выводы и рекомендации

- Введение отдельной категории игровых ноутбуков оказало положительное влияние на конверсию, чуть более чем в два раза. Это может говорить о том, что для целевой аудитории, навигация стала понятнее и они чаще находили подходящий вариант для совершения покупки.
- При этом можно выделить отрицательное влияние на средний чек, такие изменения нельзя игнорировать и цель эксперимента в полном объеме не была выполнена.
- Причины снижения среднего чека, могут заключаться в том, что более широкая аудитория была привлечена в категорию (в том числе более импульсные и недорогие покупки), стало покрытым больше сегментов. Возможно ранее пользователи искали игровые ноутбуки целенаправленно, пользователей было меньше, но покупки они делали дороже.
- Еще одной причиной может быть отличие среднего чека внутри категории, если игровые ноутбуки в целом дешевле, то увеличение их проданного кол-ва, отрицательно влияет на средний чек по категории компьютерной техники в целом.
- С точки зрения общей рекомендации, т.к. охват платящей аудитории оказался больше и совокупная выручка оказалась выше, я бы охарактеризовал результат как в целом положительный. Можно сфокусироваться на выделении преимуществ более дорогих позиций внутри категории, для переориентирования покупателей и увеличения среднего чека. Так же стоило бы дополнительно оценить маржу и итоговую прибыль по группам, для оценки общей прибыльности эксперимента и так же сопоставить результат с долгосрочной стратегией компании (устраивает ли увеличение покупок в низкоценовом сегменте)