

ПРОВЕДЕНИЕ АНАЛИЗА ДЛЯ "ВСЁ.ТЕХНИКА"

Подлеснов Андрей

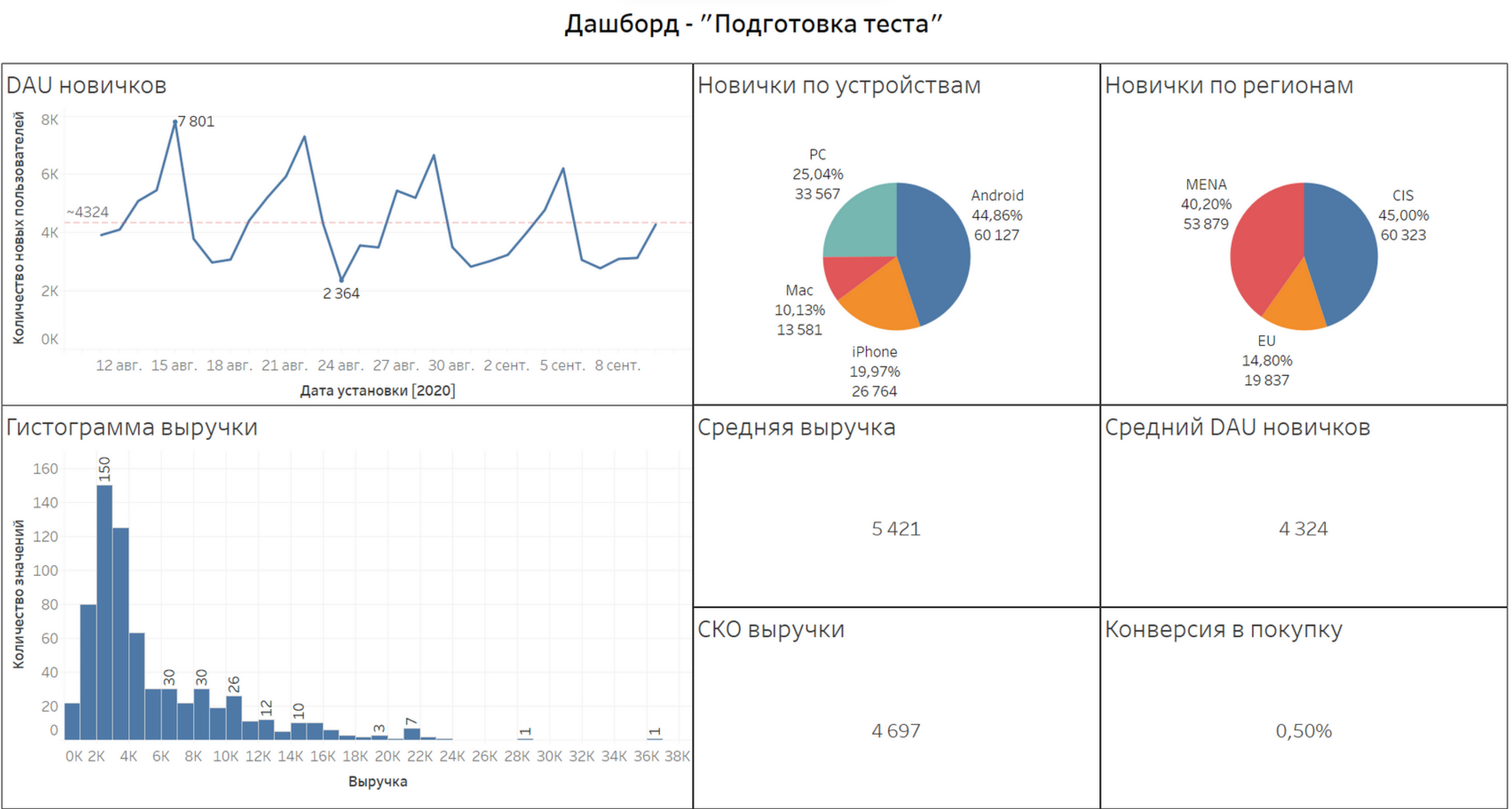
Цели анализа

Основная цель анализа заключается в проведении полного цикла A/B-тестирования для принятия решения о внедрении в маркетплейс "Все.техника" отдельной категории "Игровые ноутбуки", потенциально способной увеличить количество продаж компьютерной техники.

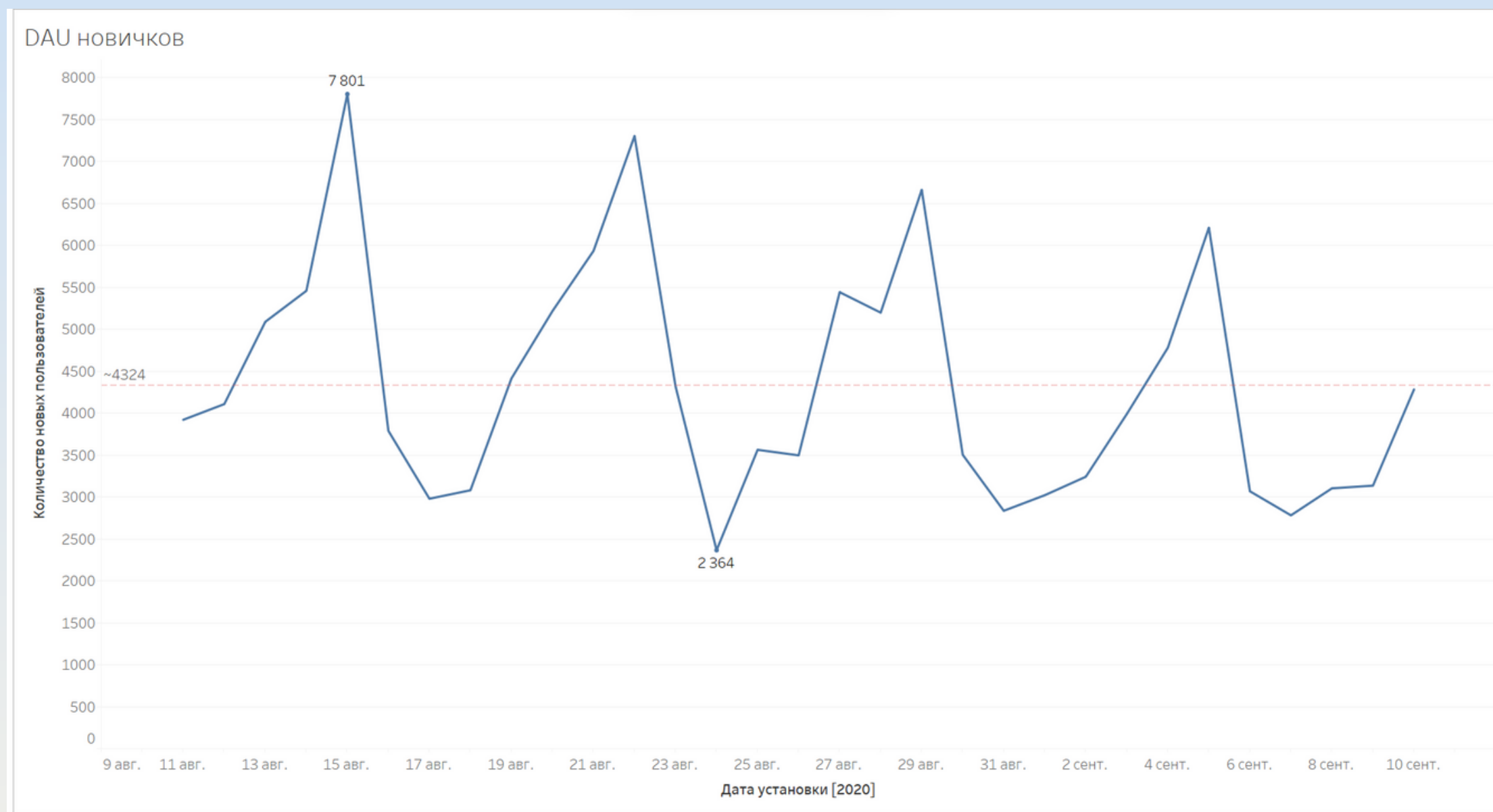
Предполагаемые результаты тестирования:

- Увеличение метрики "Конверсия в покупку" по категории "Компьютерная техника" на 100%
- Значение метрики "Средний чек" по категории "Ноутбуки" останется неизменным

Дашборд - "Подготовка теста"



DAU новичков

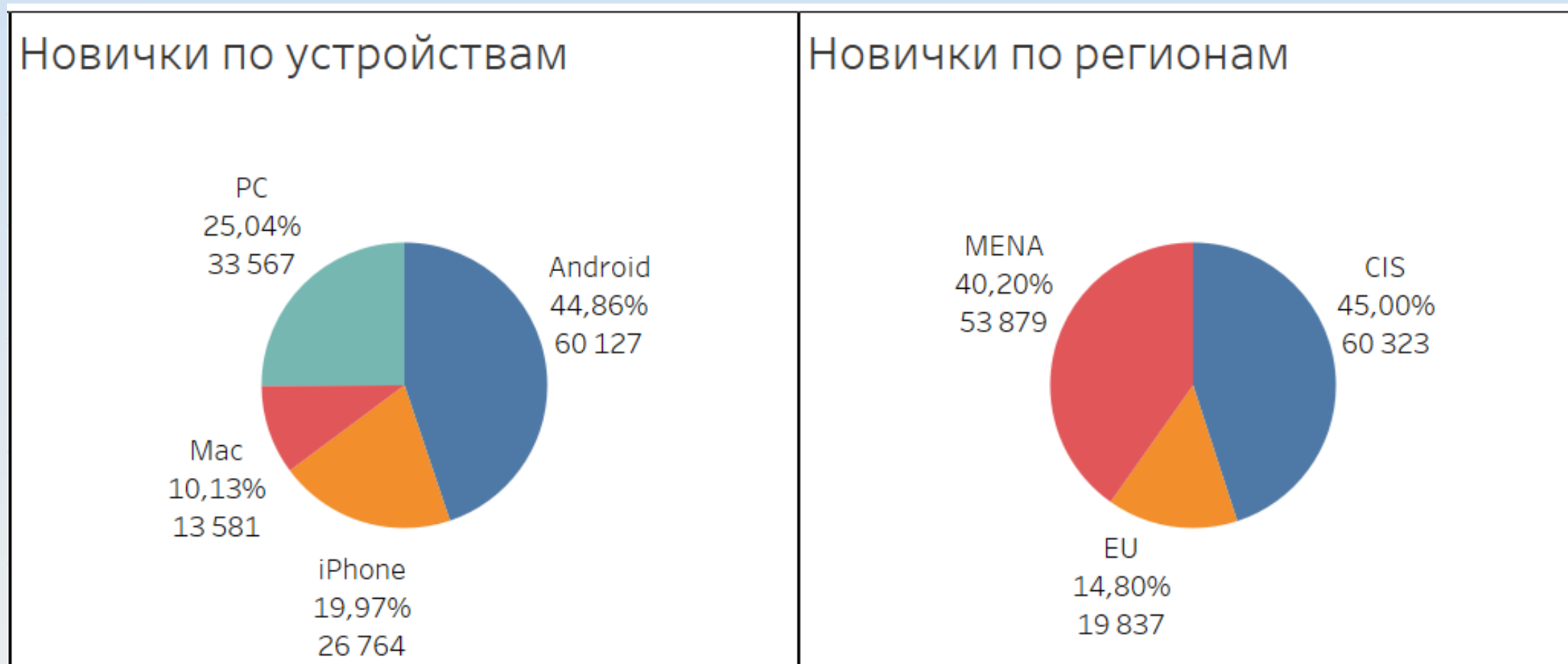


На представленном графике изображено количество новых пользователей, зарегистрировавшихся в период с 9 августа по 10 сентября.

Не трудно заметить, что объем новых пользователей является цикличным (сезонным) в зависимости от дня недели. Например, по субботам - 15.08; 22.08; 29.08, наблюдается рост новых пользователей, а по понедельникам - 17.08; 24.08; 31.08, происходит резкий спад пользователей. Среднее количество новых пользователей в день составляет около 4324 человек.

Таким образом, при проведении тестирования требуется следить за недельной динамикой, чтобы получить более точные результаты.

Распределение пользователей

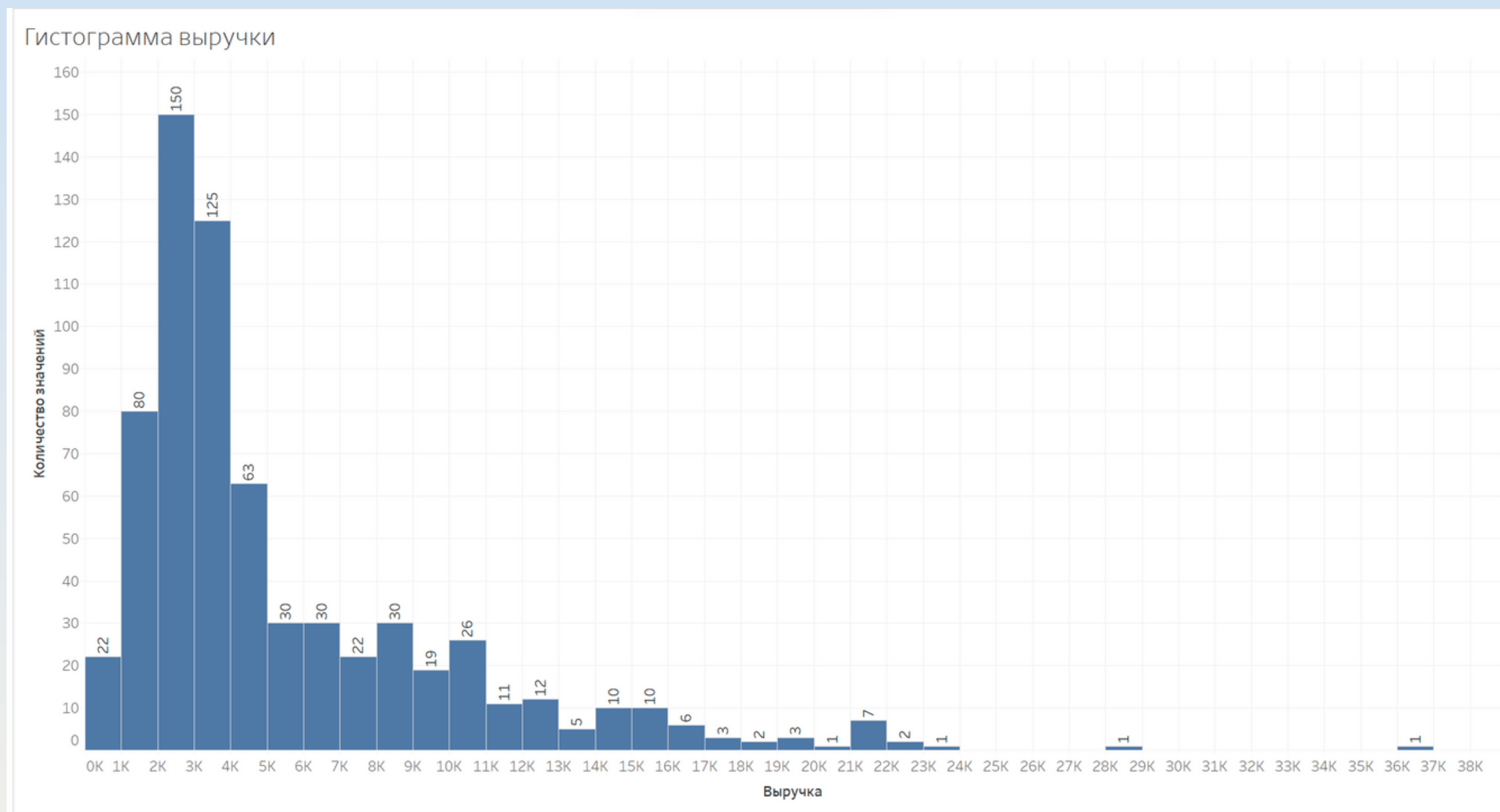


На круговых диаграммах представлены распределения пользователей по устройствам и по регионам.

Самым популярным устройством среди новых пользователей является "Android" - составляет ~45% от общего числа пользователей. Общая доля пользователей, использующих мобильное приложение составляет ~65%.

Относительно регионов видно, что основная масса новых пользователей (~85%) заказывают продукции из "CIS" и "MENA", 45% и 40.2% соответственно.

Распределение выручки



На гистограмме изображено распределение выручки с шагом в 1000 рублей.

На рисунке четко видно, что основная часть новых пользователей делают покупки на сумму от 1000 до 3000 рублей. Покупки на сумму до 1000 рублей и более 3000 рублей среди новых пользователей резко снижаются. Присутствуют так же единичные крупные покупки, например, на 36 тыс. рублей.

Можно предположить, что новые пользователи либо хотят получше ознакомиться с маркетплейсом, перед крупной покупкой, либо не могут найти нужный им товар, например, "игровой ноутбук".

Показатели выручки

Средняя выручка 5 421	Средний DAU новичков 4 324
СКО выручки 4 697	Конверсия в покупку 0,50%

Слева представлены основные показатели выручки среди новых пользователей.

Средняя выручка составляет 5421 рубль, в то время как среднеквадратичное отклонение (СКО) равно 4697 рублям. Высокое значение СКО говорит о том, что сумма выручки с покупок среди новых пользователей имеет сильный разброс значений, это так же было заметно на гистограмме в предыдущем слайде.

Конверсия в покупку составляет ~0.005, что равно 0.5%. Показание метрики говорит о том, что новые пользователи неохотно покупают технику.

"Подготовка теста" - краткие выводы:

О пользователях:

- В среднем привлекается 4324 пользователей в день
- Наибольшее количество новых пользователей привлекается по субботам, наименьшее - по понедельникам.
- Новые пользователи в большей степени предпочитают пользоваться Android (45%), в меньшей Mac (10%).
- Большая часть пользователей относится к регионам CIS и MENA - 45% и 40% соответственно.

О выручке:

- Конверсия в покупку по категории "Компьютерная техника" составляет 0.5%.
- В основном пользователи делают заказ на сумму от 1000 до 3000 рублей.
- Средняя выручка равна 5421 рублю.
- Значение СКО равняется 4697 рублям.
- Имеется большой разброс по сумме покупок у новых пользователей.
- В основном новые пользователи совершают покупки на сумму от 1000 до 3000 рублей.
- Присутствуют аномальные значения - единичные случаи высокой выручки.

Параметры теста:

Название теста: gaming_laptops_test

Назначение теста: Принятие решения о внедрении в маркетплейс "Все.техника" отдельной категории "Игровые ноутбуки"

Метрики и планируемые изменения:

- Конверсия в покупку в категории "Компьютерная техника" - увеличение метрики на 100%
- Средний чек по категории "Ноутбуки" - изменение метрики на 0% (без изменений)

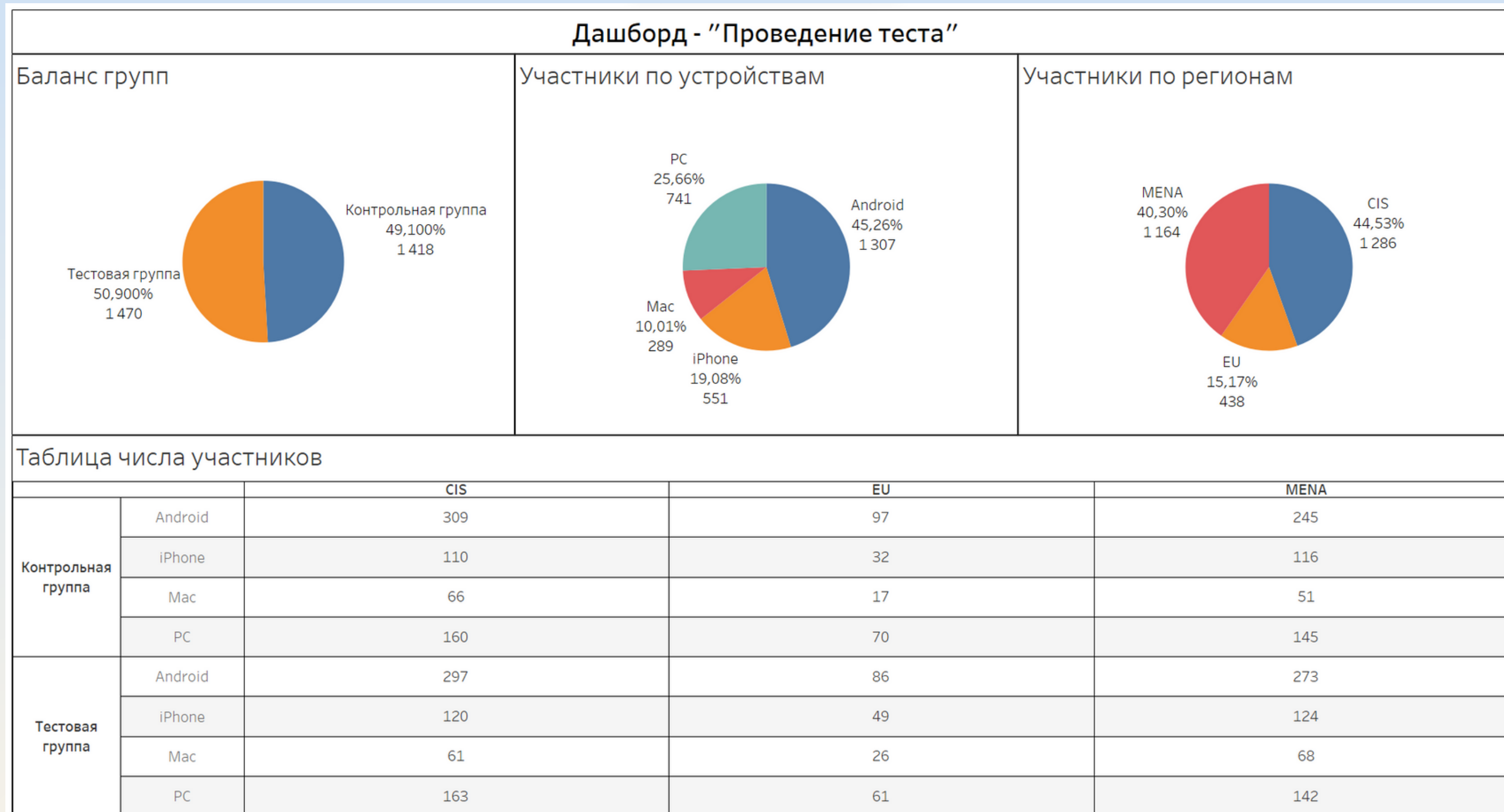
Тестовые группы:

- new_groups - группа пользователей, у которых добавили новую категорию. 50%
- current_groups - группа пользователей, у которых не произошло изменений. 50%
- Требуемое число пользователей - 9432 человек.

Время теста:

- Длительность теста - 7 дней
- Дата запуска теста - 14.10.2020
- Дата остановки теста - 21.10.2020

Дашборд - "Проведение теста"

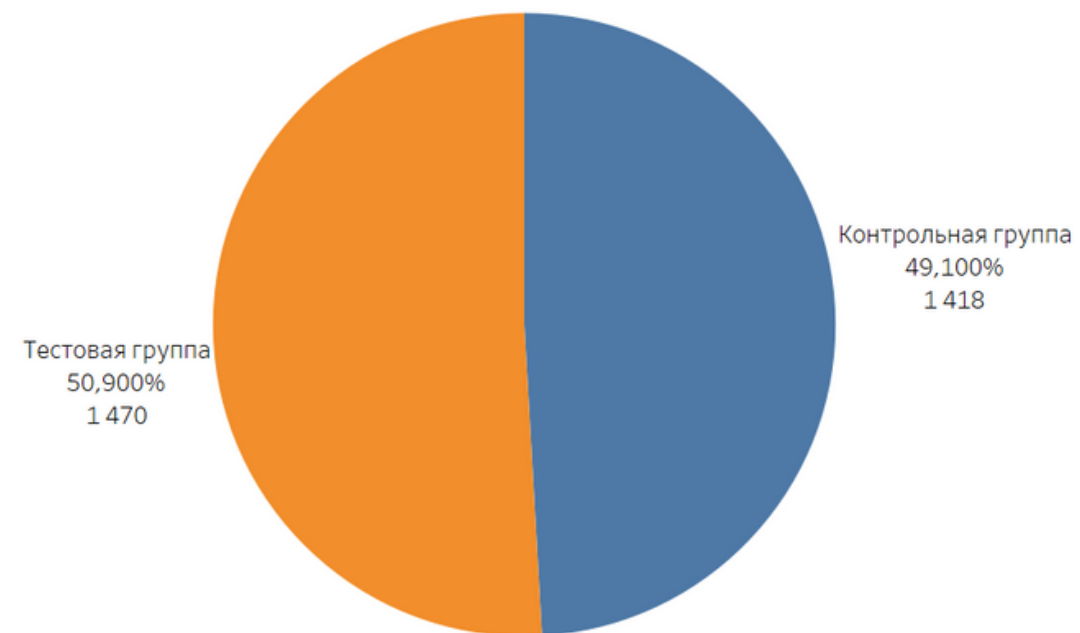


Ссылка на полную версию дашборда:

https://public.tableau.com/views/Project_4_/sheet13?:language=en-US&publish=yes&:display_count=n&:origin=viz_share_link

Баланс групп

Баланс групп



Слева представлена диаграмма, описывающая распределение пользователей по тестовым группам за 14.10.2020.

На рисунке видно, что тестовая группа имеет 50.9% пользователей, контрольная - 49.1%. Разница между долями составляет 1.8%, что не является критичным значением, однако стоит провести Z-тест, чтобы удостовериться в этом.

Z-тест: Проведение теста

p_0 (hypothesized population proportion)

0,5

p (observed sample proportion)

0,509

n (sample size)

2888

CALCULATE

z-statistic: 0.96732

p-value (one-tailed): 0.16669

p-value (two-tailed): 0.33338

95% C.I. = [0.4908, 0.5272]

Одновыборочный Z-тест используется для определения предполагаемой доли с фактической в одной из тестовых групп.

В качестве предполагаемой доли пользователей (p_0) тестовой группы возьмем 0.5, что соответствует предполагаемым 50%.

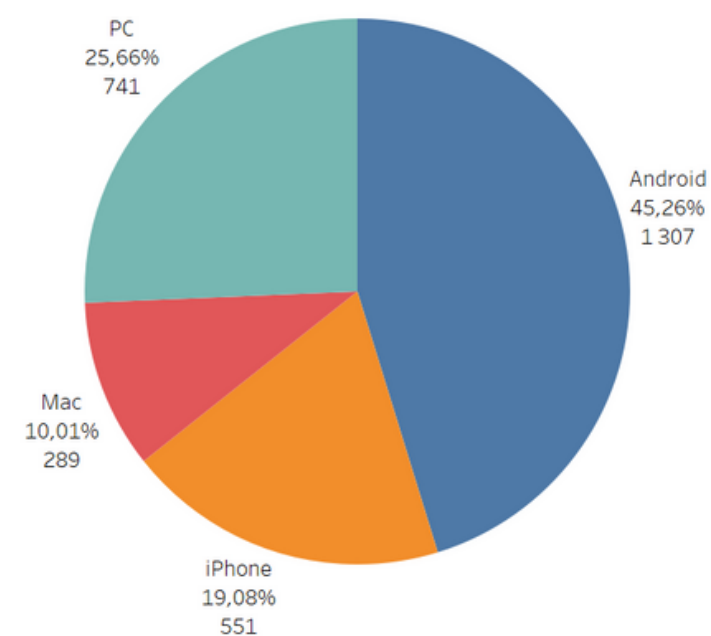
В качестве фактической доли (p) тестовой группы возьмем значение из диаграммы "Баланс групп" равное 0.509 (50.9%).

Размер выборки равен количеству пользователей, которые участвуют в тестировании (1470+1418).

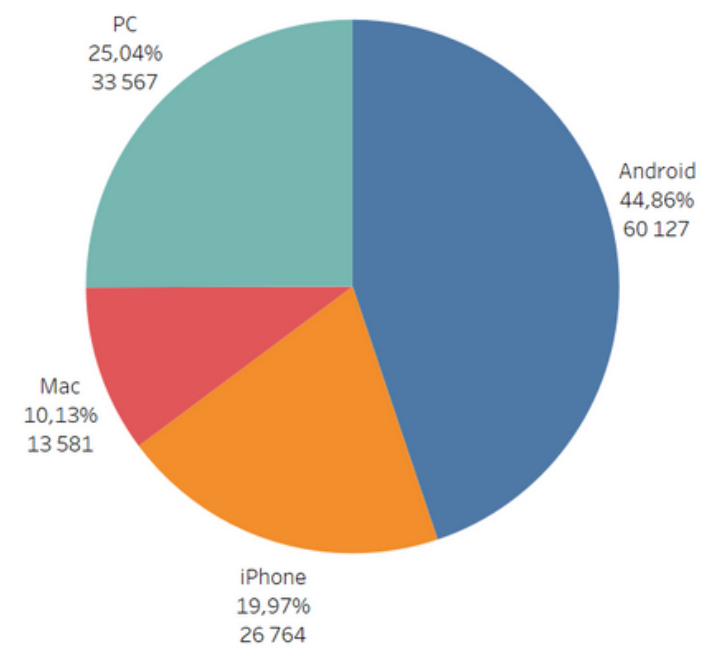
В данном случае, мы смотрим на значение двухстороннего z-теста, равное 0.33338. По причине того, что полученное значение меньше указанной доли, можно сделать вывод, что группы распределены равномерно.

Распределение по устройствам

Участники по устройствам



Новички по устройствам



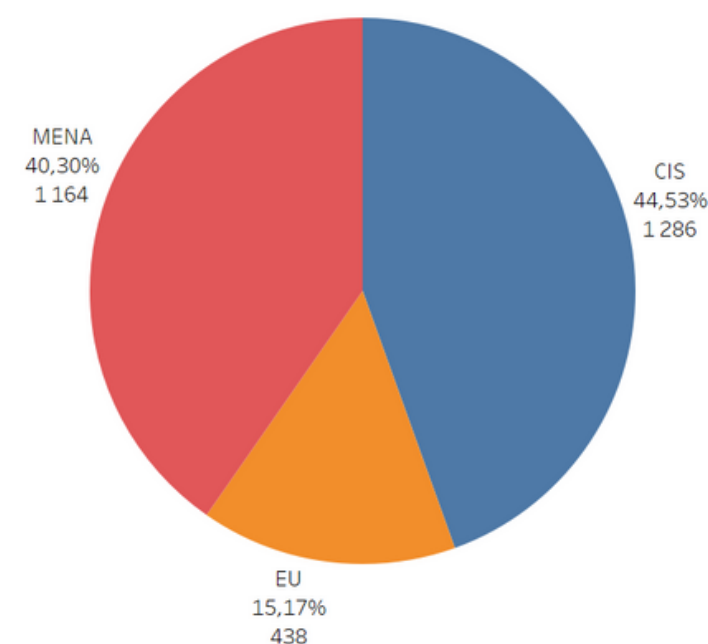
Слева представлена диаграмма, описывающая распределение участников теста по устройствам за 14.10.2020, справа - распределение пользователей по устройствам, но на основе исторических данных.

Все доли графиков примерно равны, отличие между первым днем теста и историческими данными составляет менее 1%.

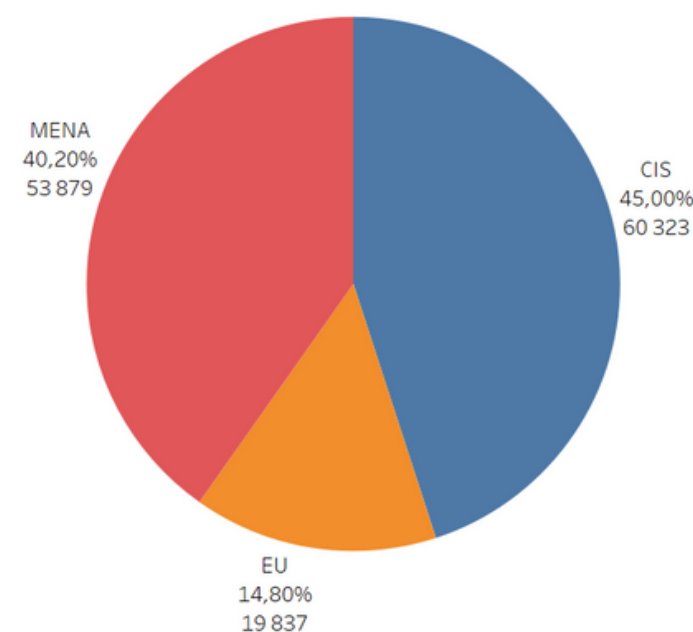
Таким образом, в первый день теста доли пользователей относительно устройства, практически, совпадают с долями, основанными на исторических данных.

Распределение по регионам

Участники по регионам



Новички по регионам



При сравнении долей по регионам, наблюдается аналогичная ситуация.

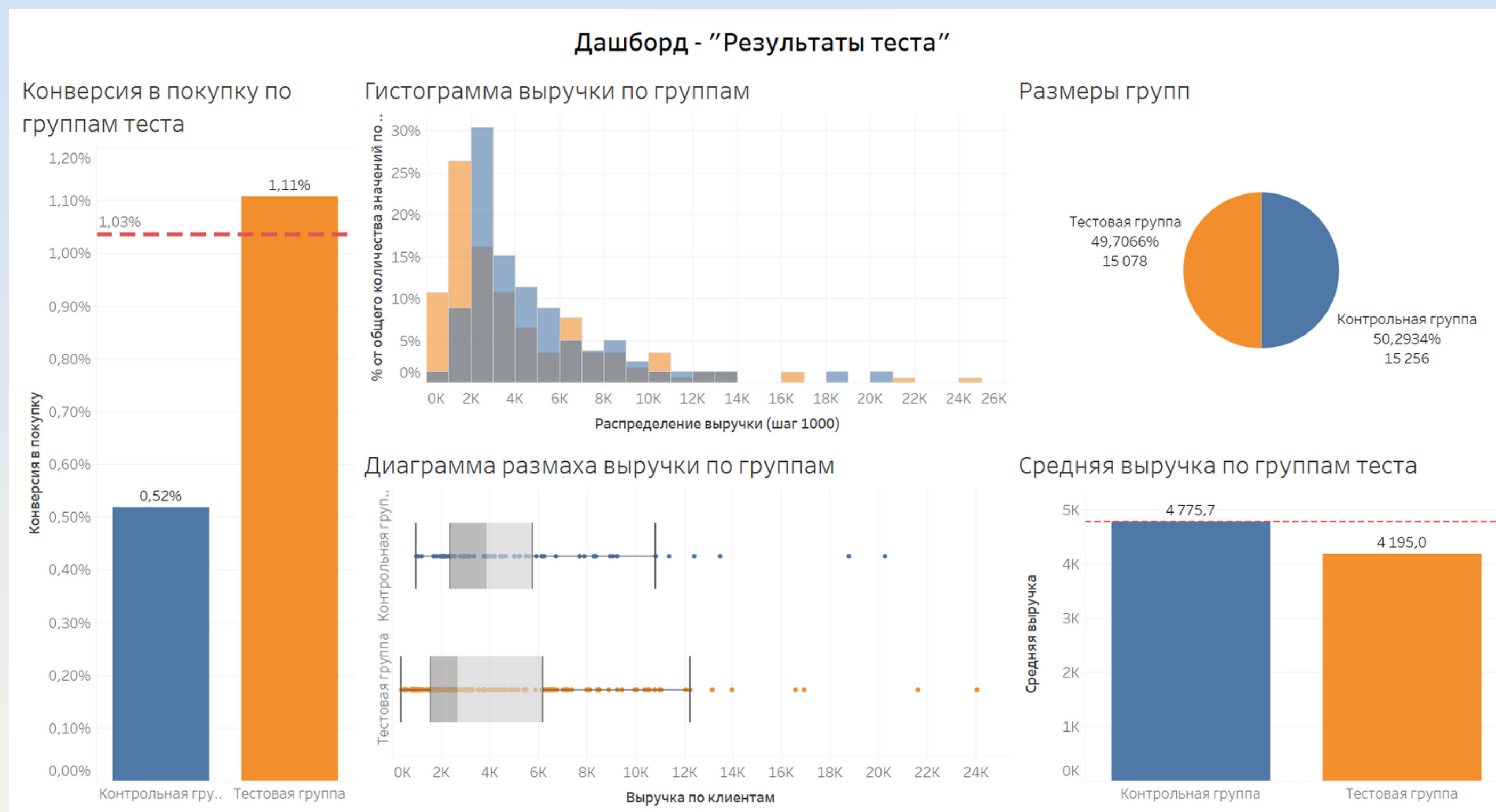
Все доли графиков примерно равны, отличие между первым днем теста и историческими данными составляет менее 1%.

Таким образом, в первый день теста доли пользователей относительно регионов, практически, совпадают с долями, основанными на исторических данных.

"Проведение теста" - краткие выводы:

- Тестовая и контрольная группы распределены равномерно с учетом проведенного Z-теста.
- Распределение пользователей по устройствам при проведении тестирования, практически, совпадает с распределением, основанном на исторических данных.
- Распределение пользователей по регионам при проведении тестирования, практически, совпадает с распределением, основанном на исторических данных.
- Сбор данных проводится в нормальном режиме и не было набрано минимально необходимое количество пользователей - 4716 человек. Остановка теста не требуется.

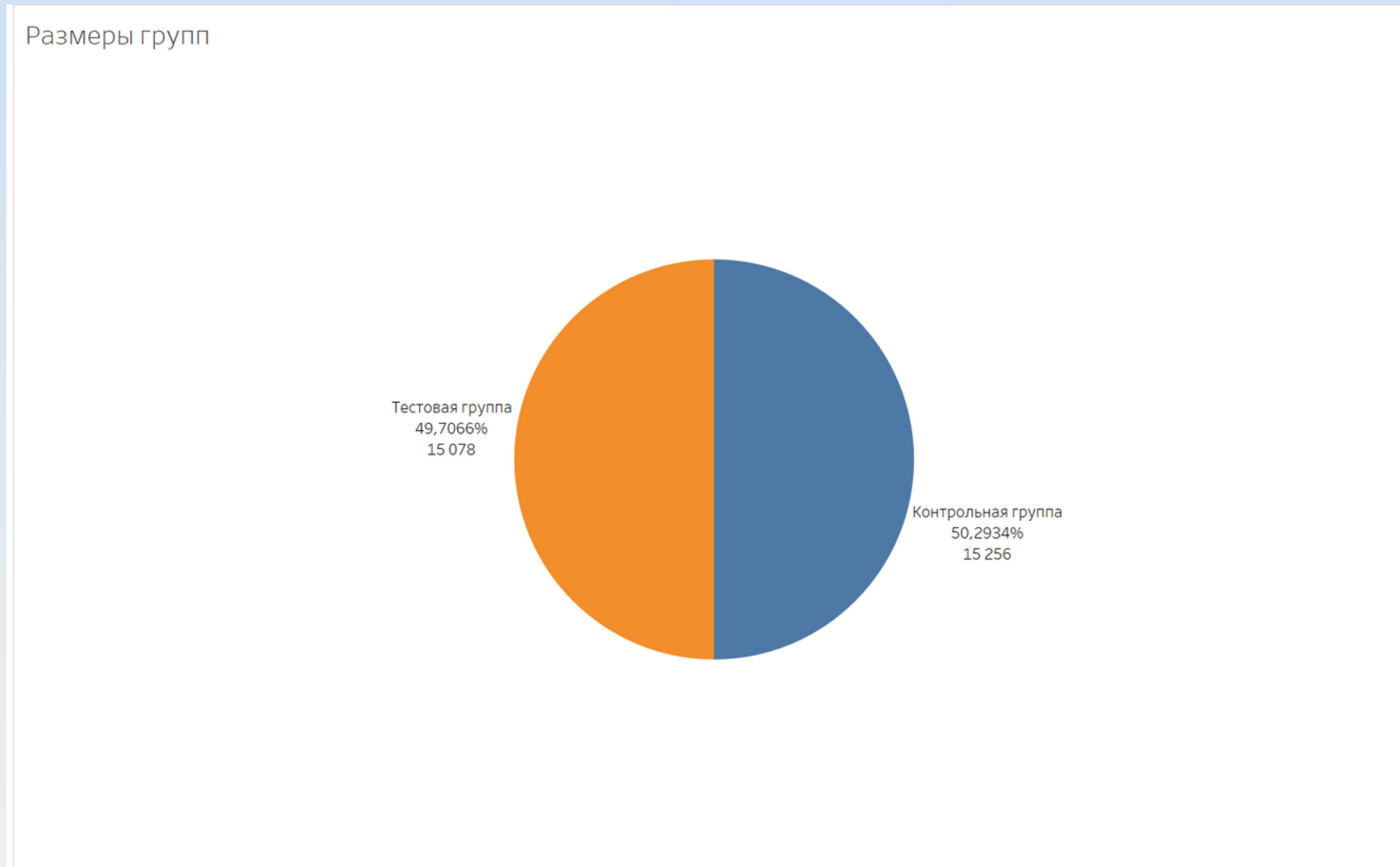
Дашборд - "Результаты теста"



Ссылка на полную версию дашборда:

https://public.tableau.com/views/Project_4_/sheet21?:language=en-US&publish=yes&:display_count=n&:origin=viz_share_link

Размеры групп



На диаграмме изображено распределение пользователей по группам, участвовавших в тестировании. У тестовой группы была добавлена категория "Игровые ноутбуки", у контрольной не было изменений.

На первый взгляд, группы распределены равномерно, т.к. отличия составляют около 0.5%, однако стоит это проверить с помощью Z-теста.

Z-тест: Результаты теста

p_0 (hypothesized population proportion)

0,5

p (observed sample proportion)

0,497066

n (sample size)

30334

CALCULATE

z-statistic: -1.02201

p-value (one-tailed): 0.15339

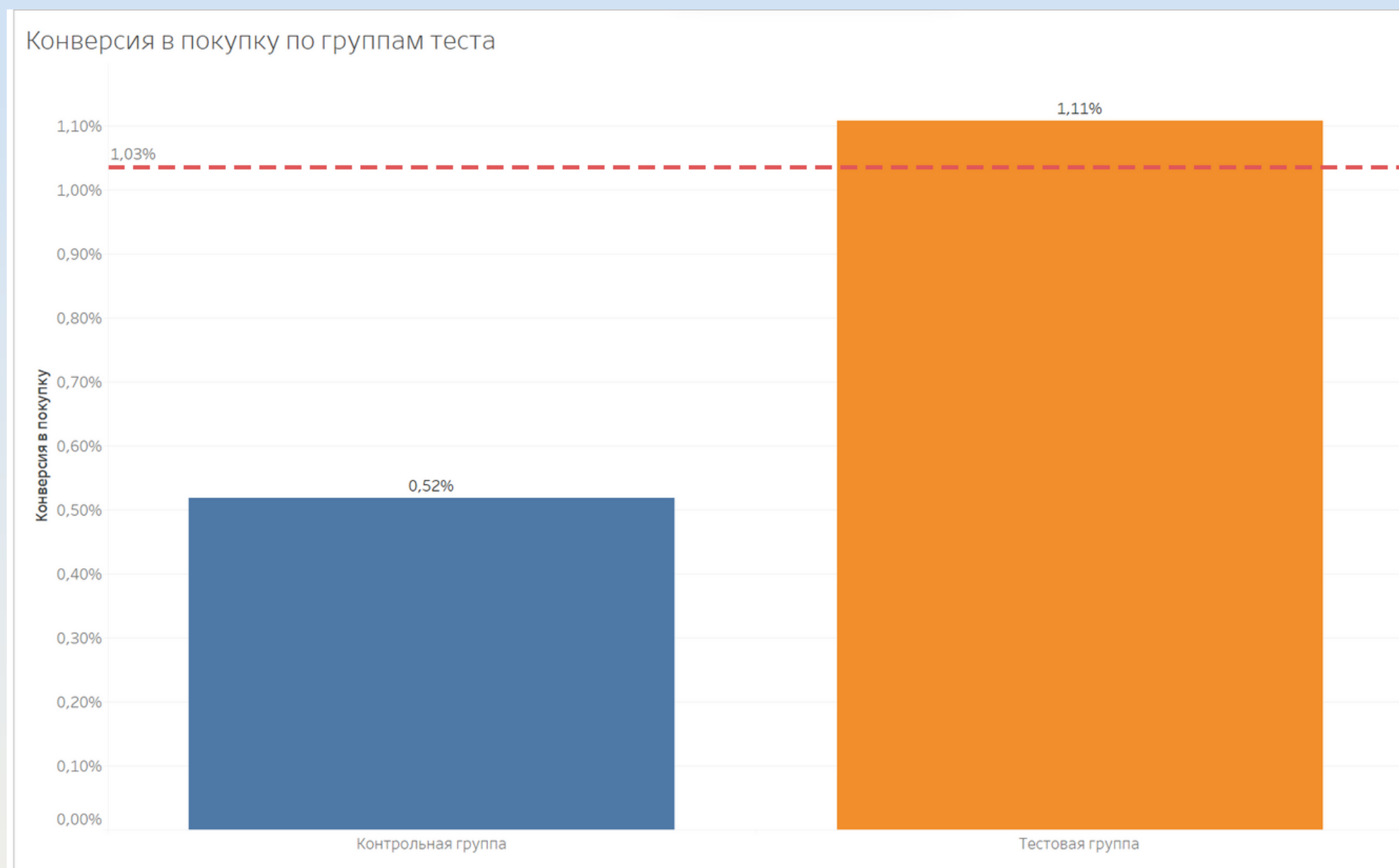
p-value (two-tailed): 0.30678

95% C.I. = [0.4914, 0.5027]

Проводим аналогичный Z-тест, как при проведении тестирования.

Двусторонний тест (p-value two-tailed) равен 0.30678, что меньше указанного значения "p", значит делаем вывод, что фактическая доля тестовой группы равна планируемой (50%).

Конверсия в покупку

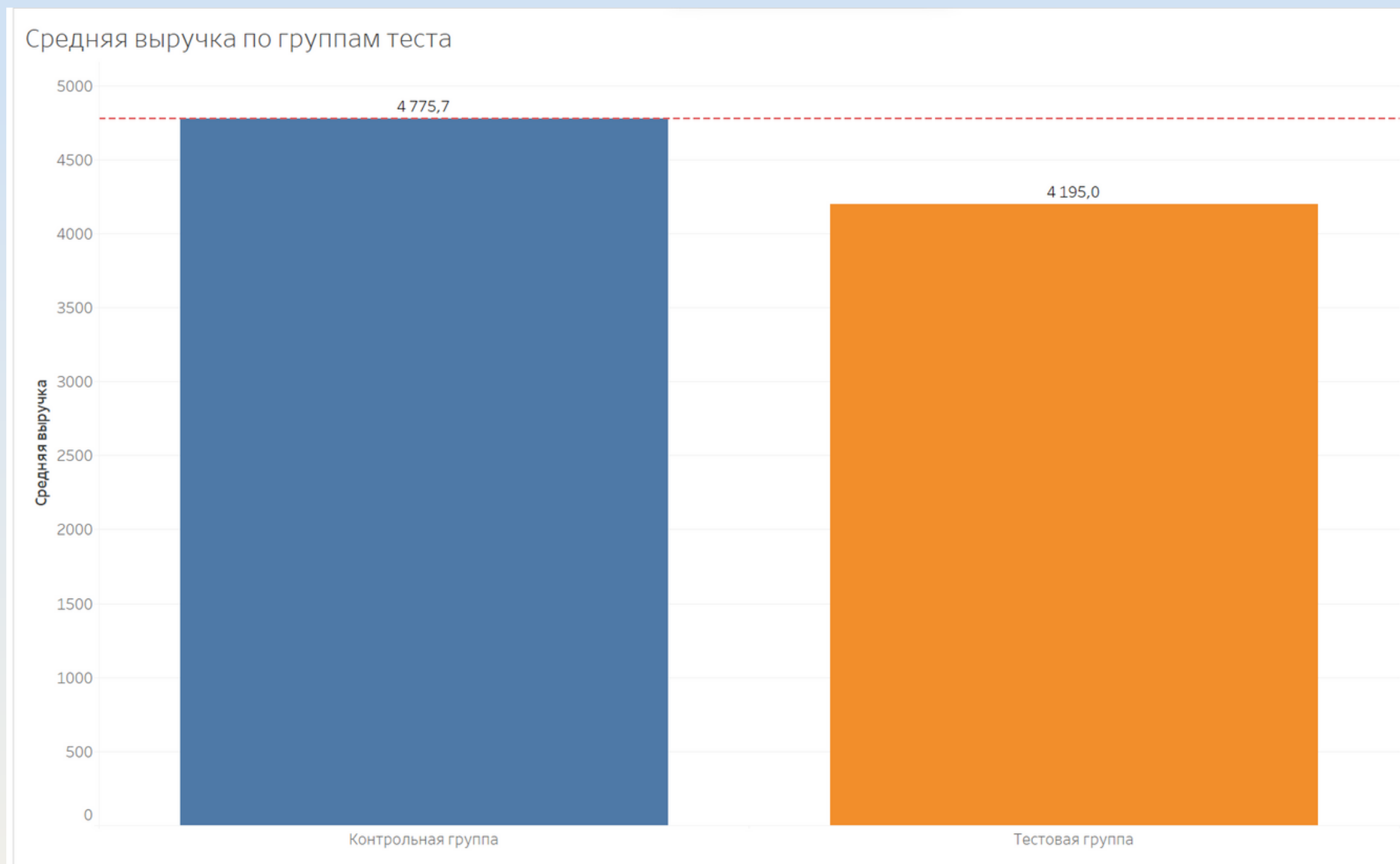


На графике представлена метрика "Конверсия в покупку" по тестовым группам.

По результатам проведенного тестирования конверсия в покупку в тестовой группе выросла на ~114%, в то время как прогнозируемый рост равнялся 100%.

Вывод: Фактический результат оказался лучше прогнозируемого.

Средняя выручка

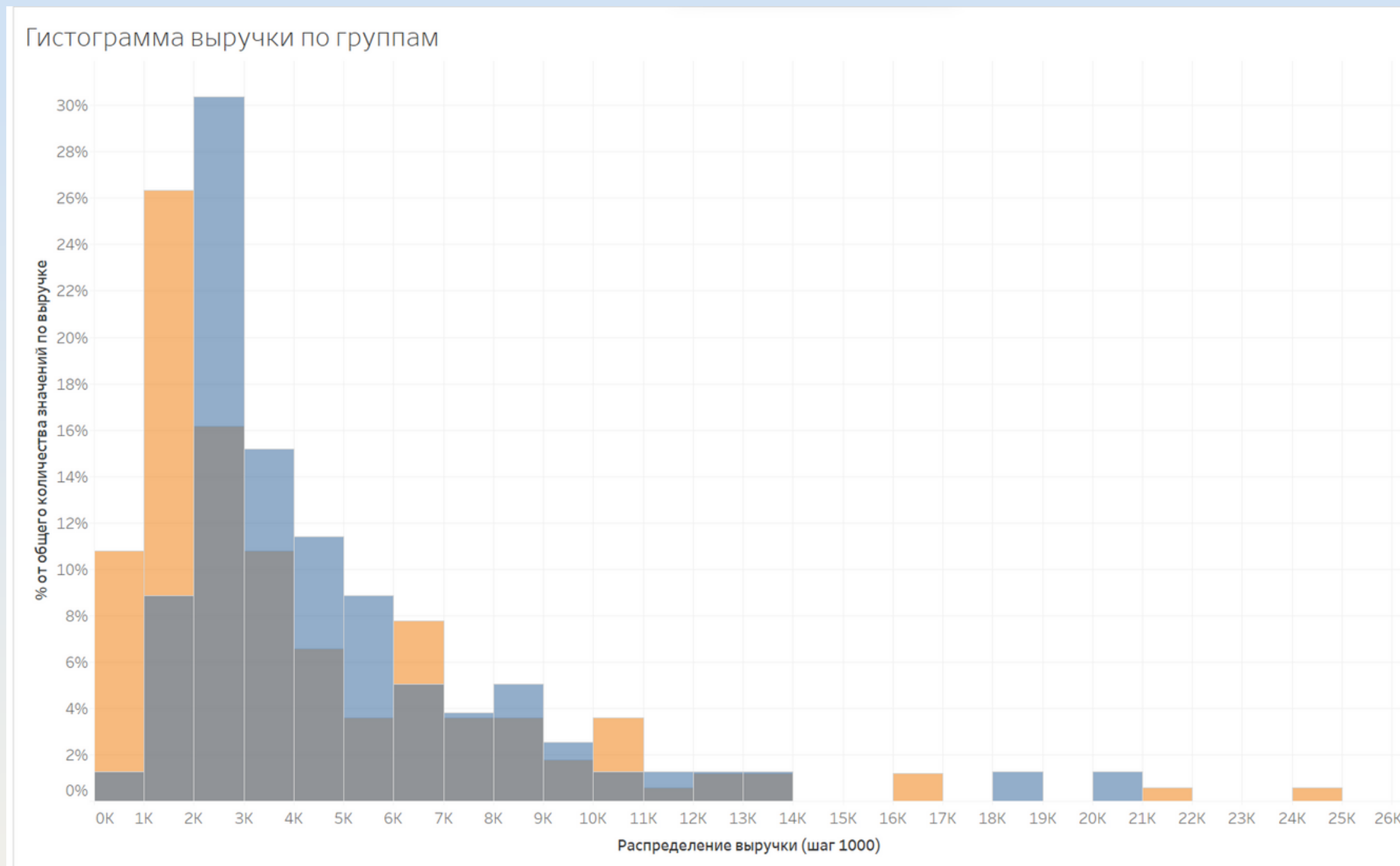


На графике представлена метрика "Средняя выручка" по тестовым группам.

По результатам проведенного тестирования средняя выручка в тестовой группе снизилась на ~12%, в то время как изменение не прогнозировалось.

Вывод: Нововведение негативно сказалось на средней выручке.

Гистограмма выручки



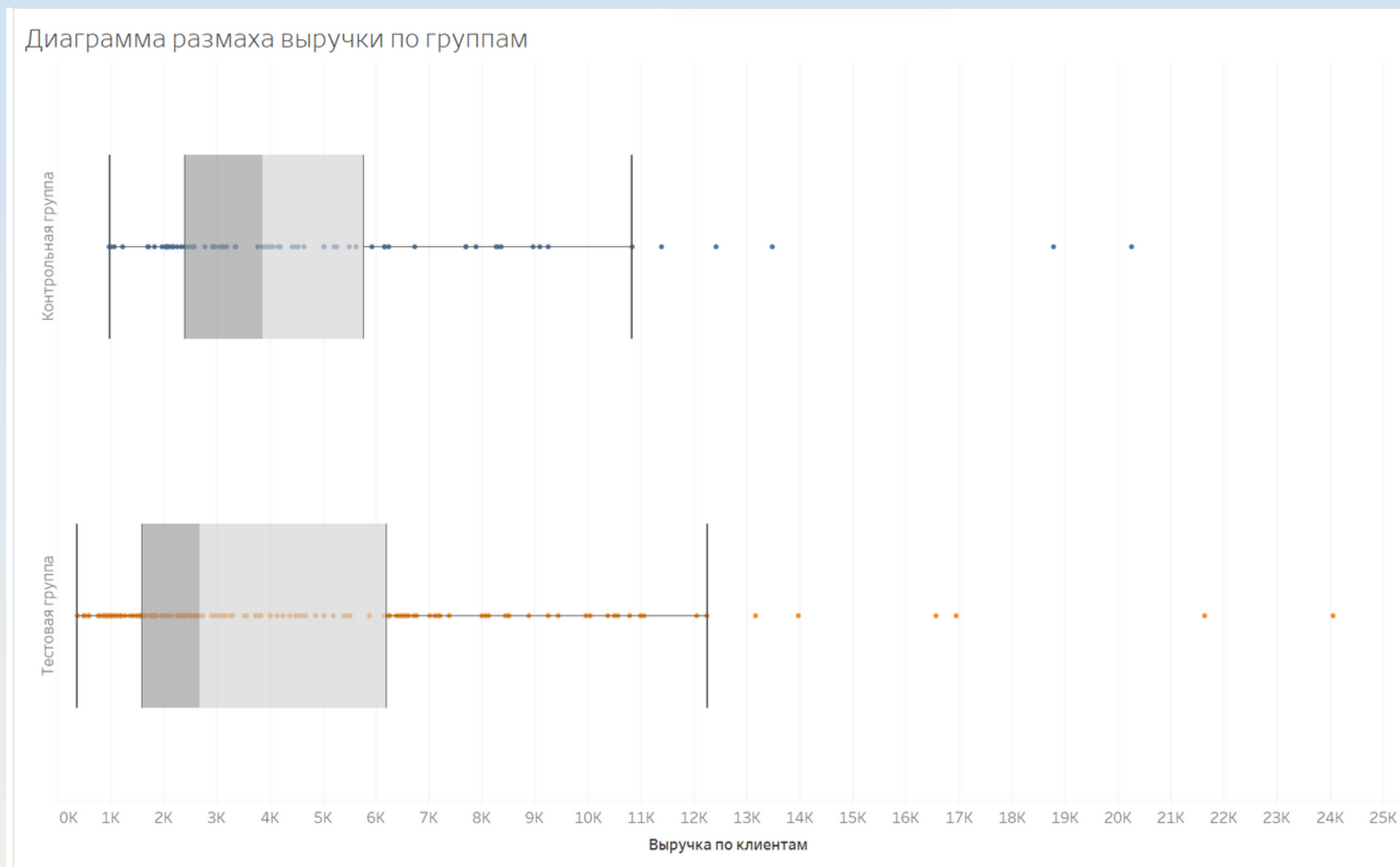
На графике представлена гистограмма по распределению выручки, а также доля каждого промежутка выручки от общей числа покупок.

На графике видно, что в тестовой группе пользователи нацелены больше на покупки до 2000 рублей, в то время как в контрольной группе основная доля покупок приходится на сумму от 2000 до 4000 рублей.

Также заметны выбросы, которые следует оценить при помощи диаграммы размаха выручки.

Вывод: с введением новой категории, пользователи стали чаще совершать недорогие покупки.

Диаграмма размаха выручки



На графике представлена диаграмма размаха, используемая при анализе аномальных значений.

Первое, что стоит заметить, так это расширение значений выручки, пользователи стали покупать как более дешевые товары, так более дорогие.

Если обратить внимание на смещенную влево медиану, то можно сделать вывод, что пользователи тестовой группы больше склоняются к покупке недорогих товаров.

Также можно заметить, что в тестовой группе стали совершать более дорогие покупки и чуть чаще, чем в контрольной.

Вывод: пользователи тестовой группы стали более склонны к покупке недорогих товаров.

Статистические тесты: Z-тест для пропорций

p_1 (sample 1 proportion)
0,52

n_1 (sample 1 size)
15256

p_2 (sample 2 proportion)
1,11

n_2 (sample 2 size)
15078

CALCULATE

z-statistic: -131.84207

p-value (one-tailed): 0.00000

p-value (two-tailed): 0.00000

Для оценки конверсии в покупку используем двухвыборочный Z-тест для пропорций. Он позволит оценить, изменилась ли конверсия в покупку при внедрении новой категории.

Значение p-value (двухсторонний) равно нулю, что меньше уровня значимости, равного 0.05, значит имеются различия в группах, конверсия изменилась. Помимо этого, конверсия в покупку выросла в тестовой группе и превысила плановое значение, что является положительным эффектом для продукта.

Статистические тесты: t-тест

The t -value is -1.10702. The p -value is .269377. The result is *not* significant at $p < .05$.

Want to know how to report this t -test result in your work? (Opens in a new tab so you calculation.)

[How to report a \$t\$ -test result \(APA\)](#)

Calculate T and P Values

Reset

Т. к. средняя выручка не измеряется в долях, то понадобится использование t -теста. Он применяется в случаях, когда нужно определить различия между количественными значениями метрик.

В данном случае p -value = 0.269377, что меньше уровня значимости равного 0.05.

Таким образом, по проведенному t -тесту средняя выручка по тестовым группам одинаковая, но стоит провести тест Манна-Уитни, чтобы удостовериться в этом. Причиной дополнительного теста является большое количество выбросов, а так же сильное снижение метрики (почти на треть), в то время как количество пользователей в 2 выборках примерно одинаковое.

Статистические тесты: тест Манна-Уитни

Remember to select significance level and whether your hypothesis is one or two-tailed.

The z -score is -2.57143. The p -value is .01016. The result is significant at $p < .05$.

Тест Манна-Уитни позволит определить, взяты ли выборки из одной генеральной совокупности, что укажет на целесообразность сравнения метрик средней выручки.

По результатам теста p -value = 0.01016, что меньше уровня значимости 0.05. Это говорит о том, что выборки взяты из разных генеральных совокупностей и сравнение значений средней выручки является относительно бессмысленным.

РЕКОМЕНДАЦИИ

В ходе проведения тестирования было выявлено, что конверсия в покупку выросла в тестовой группе и превысила прогнозируемые значения. Однако, она могла вырасти по причине более частых покупок среди недорогих товаров, возможно комплектующих к игровым ноутбукам. Дополнительно гипотеза подкрепляется фактом того, что в тестовой группе медианное значение выручки снизилось. Проведение статистических тестов показало, что полученные значения по метрике "Средняя выручка" указывают на различия в параметрах тестовых групп, иными словами вполне возможно, что проходило сравнение средних значений у пользователей из разных регионов или пользователей, использующие разные устройства.

Таким образом, идея внедрения отдельной категории является перспективной, т.к. был достигнут результат по одной из метрик, однако требуется провести дополнительное тестирование, чтобы более точно определить целевую аудиторию. Возможно, что нововведение может положительно сказаться на какой-то определенной категории пользователей, а на других - отрицательно.

Вывод: Сейчас внедрять категорию не стоит, требуются дополнительные исследования.

СПАСИБО ЗА ВНИМАНИЕ!