

Исследование для компании «Мегафон»

Содержание исследования

Оценка статистических характеристик
количественных показателей, представленных
в данных по двум группам абонентов

Постановка задачи исследования

Все абоненты будут разделены на две группы:

- пользователей, которых всё устраивает (оценка 9-10) - **группа А**
- пользователей, которые недовольны качеством предоставляемых услуг (оценка 1-8) - **группа В**

Постановка задачи исследования

Если показатели у двух групп отличаются и пользователям из группы А действительно, по техническим характеристикам, предоставляются услуги лучшего качества, то компании, скорее всего потребуется повысить качество предоставляемых услуг, чтобы для каждого абонента услуги предоставлялись на равных условиях.



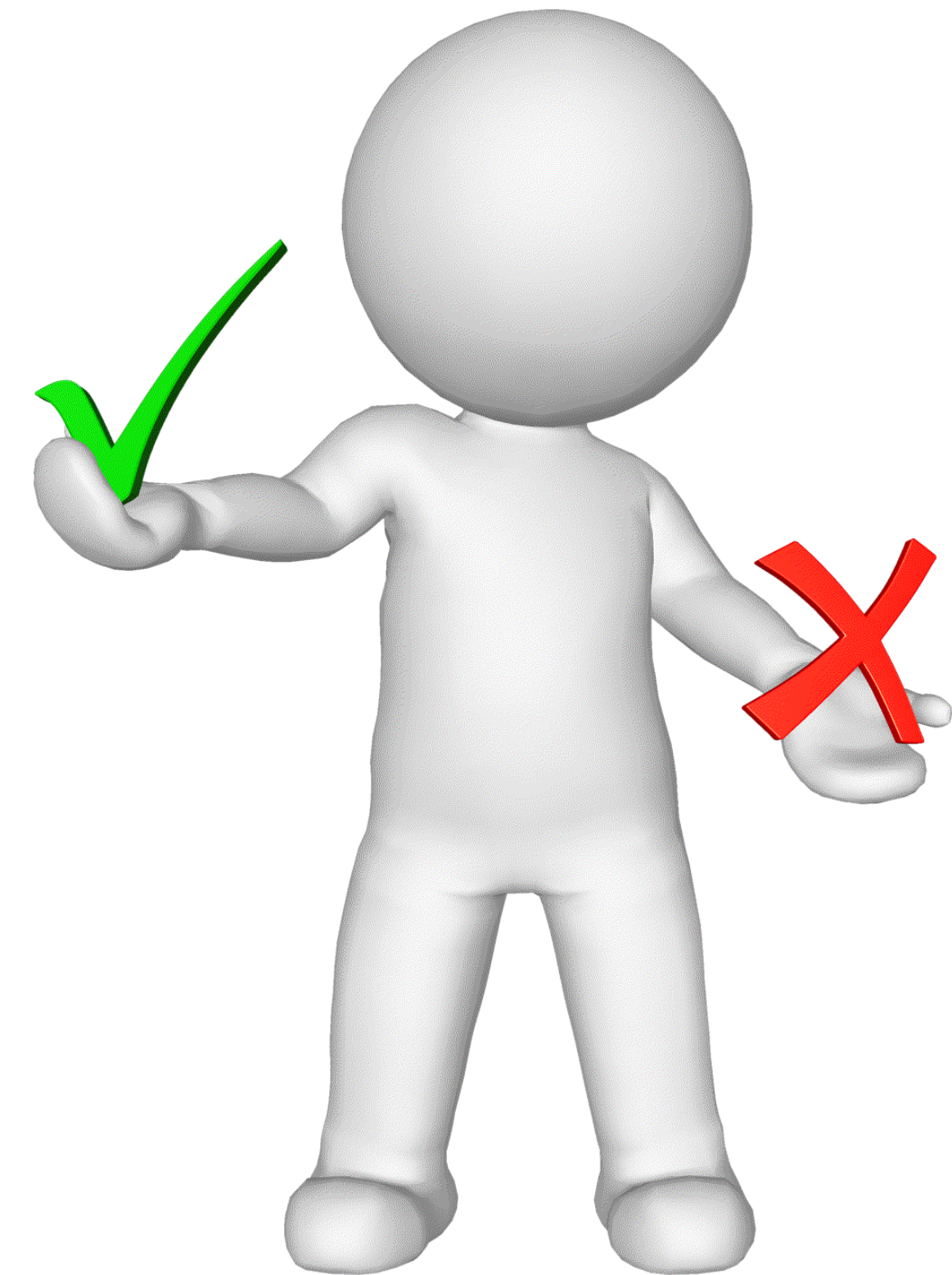
Постановка задачи исследования

Если же статистически значимой разницы между группами не будет наблюдаться, то можно будет сделать вывод, что дело не в технических проблемах и пользователям не нравится что-то еще.



Постановка задачи исследования

Необходимо оценить характеристики всех метрик на предмет наличия статистически значимой разницы между группой А и группой В, для определения вектора дальнейшего развития компании.



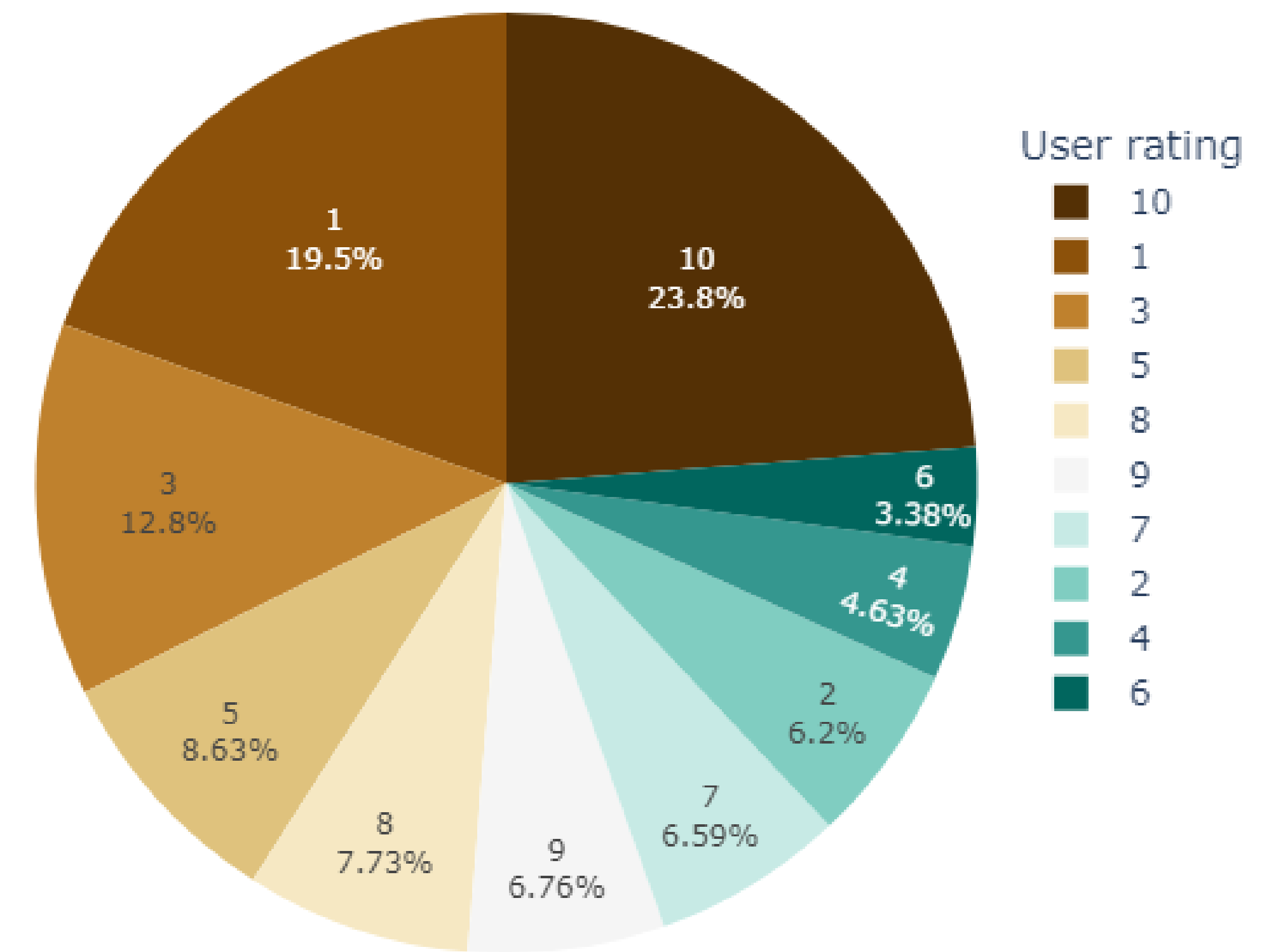
Результаты разведочного анализа

Если разделить категории ответов на группы в процентном соотношении среди пользователей, им будут соответствовать значения:

- группа А - 30.56%
- группа В - 69.44%

То есть, фактически, **70% пользователей участвующих в опросе не довольный** качеством предоставляемых услуг.

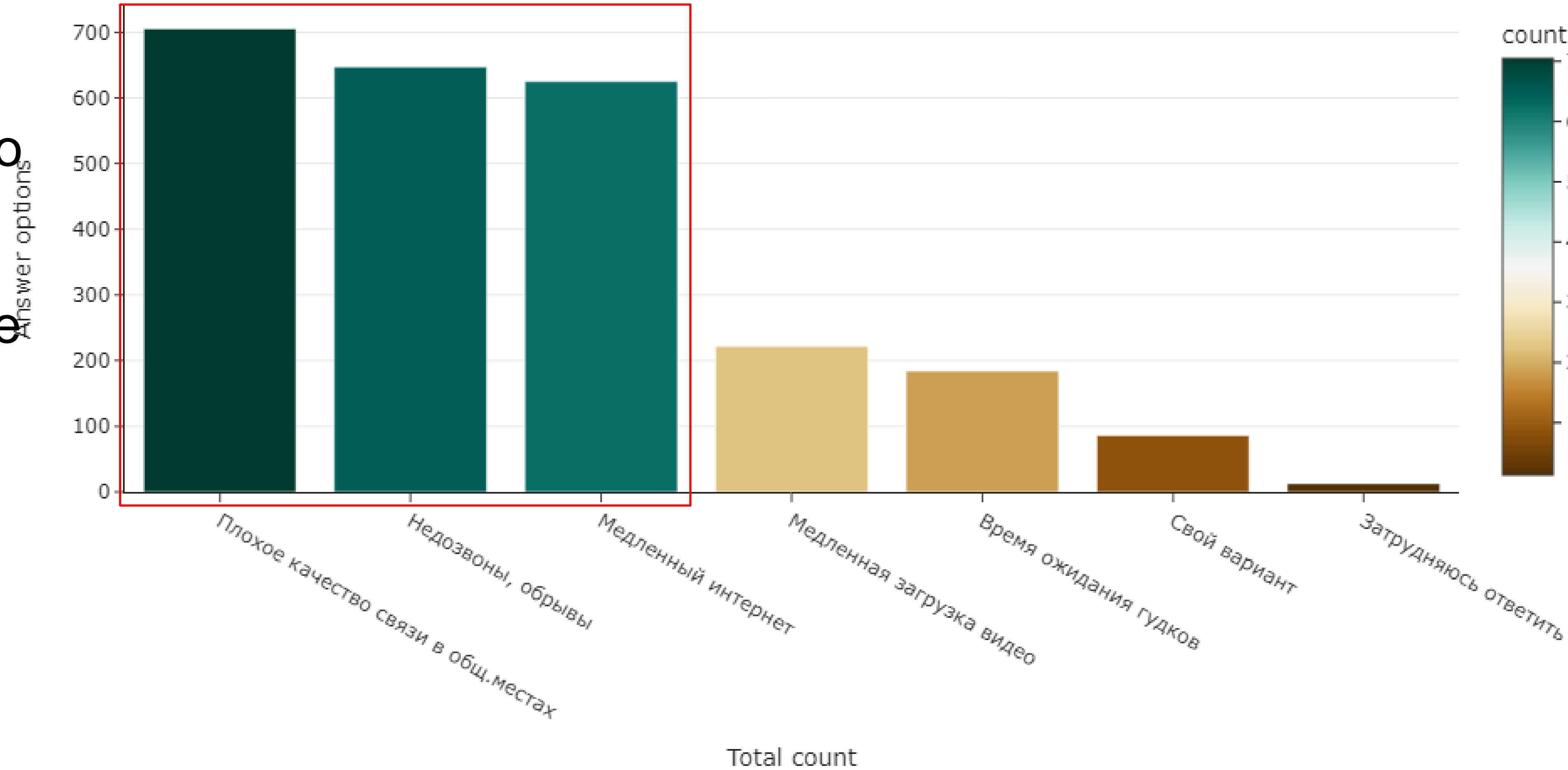
Distribution of answers to the first question



Результаты разведочного анализа

Distribution of user responses to the second question

Больше всего опрошенным абонентам не нравятся



Результаты разведочного анализа

Показатели метрики **Total**

Traffic(MB) по группам

почти одинаковы,

отличается только

медианное значение.

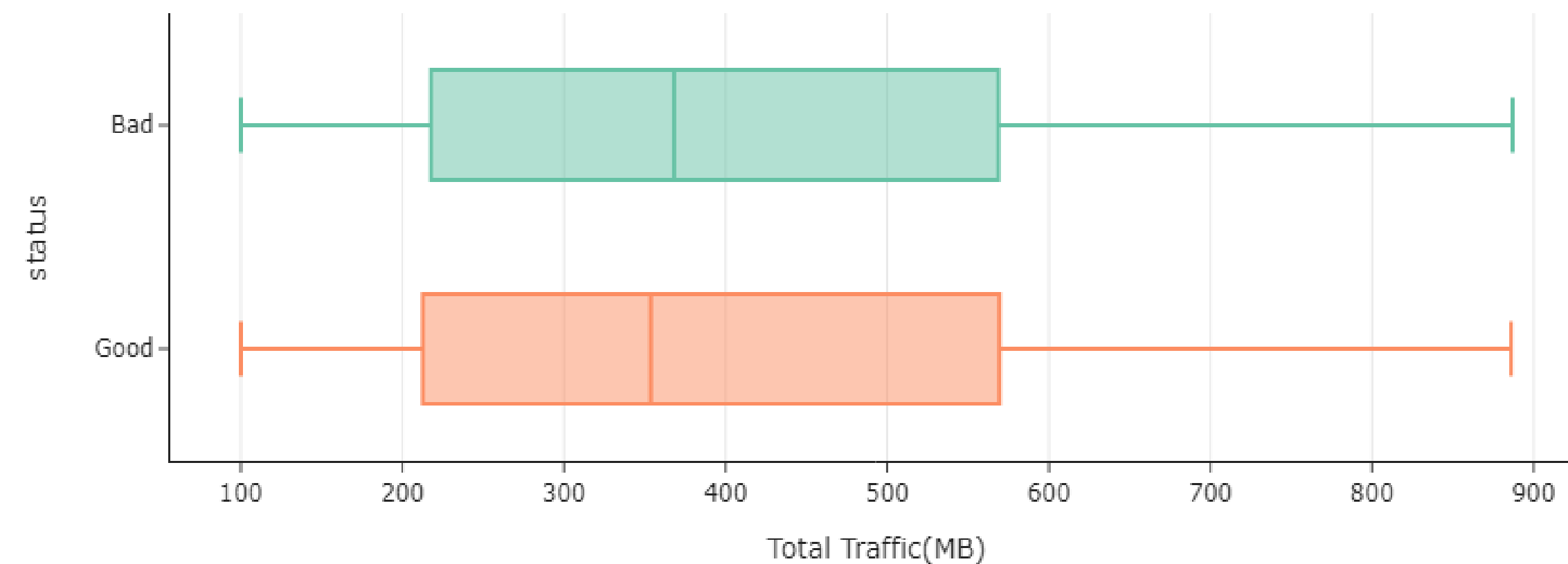
Объем трафика

передачи данных у

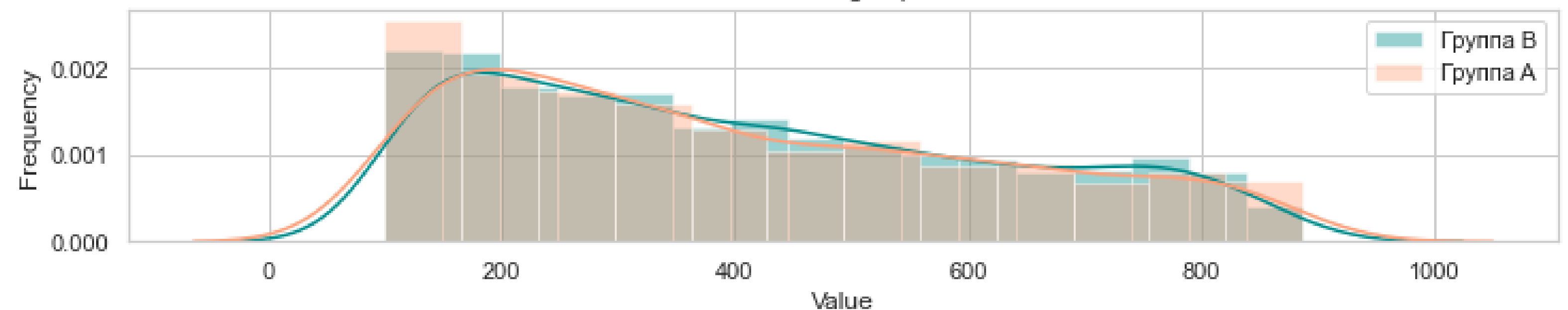
обоих групп не

отличается.

Statistical characteristics by parameter Total Traffic(MB) groups



Hist of distribution Total Traffic(MB) for two groups



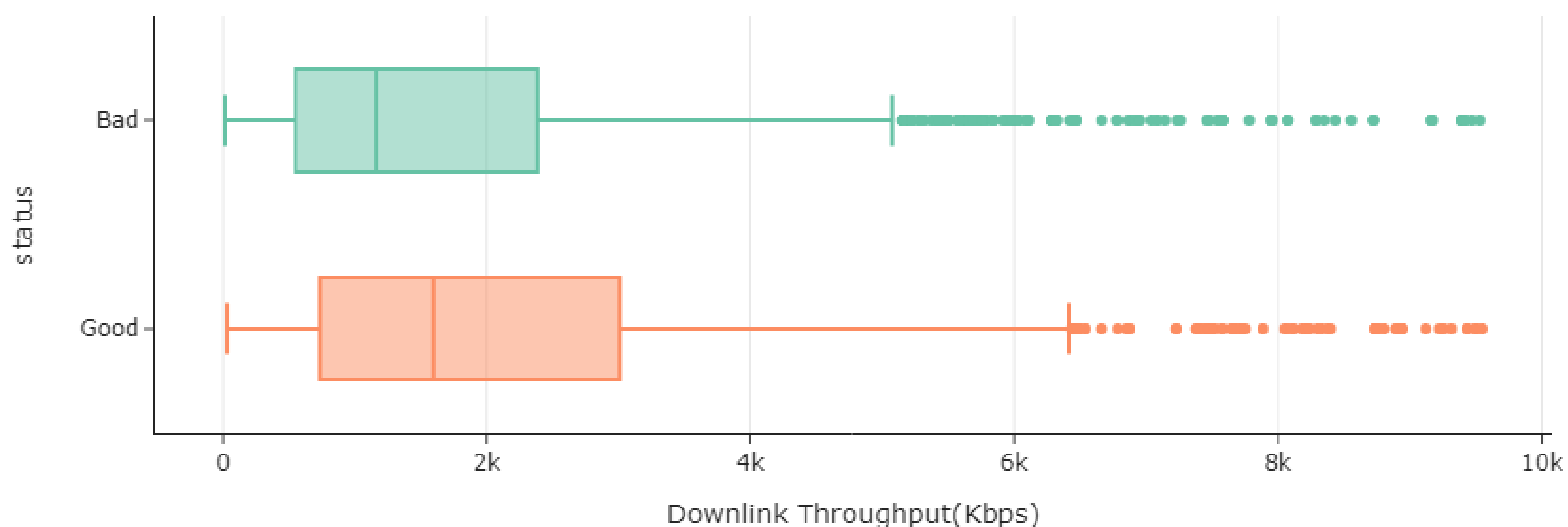
Результаты разведочного анализа

Downlink

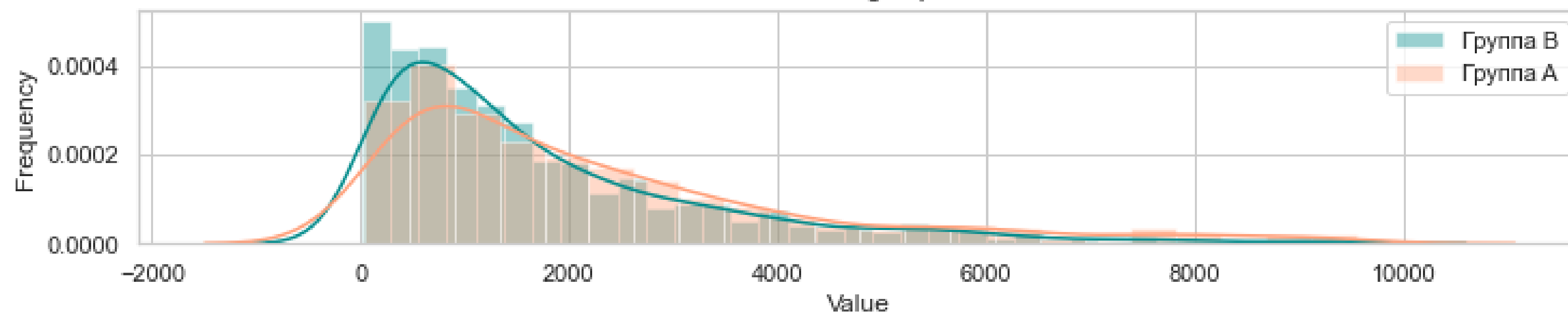
Throughput(Kbps) имеет большой разброс от верхнего квартиля до максимума.

Средняя скорость "к абоненту" выше у группы с хорошими оценками.

Statistical characteristics by parameter Downlink Throughput(Kbps) groups



Hist of distribution Downlink Throughput(Kbps) for two groups

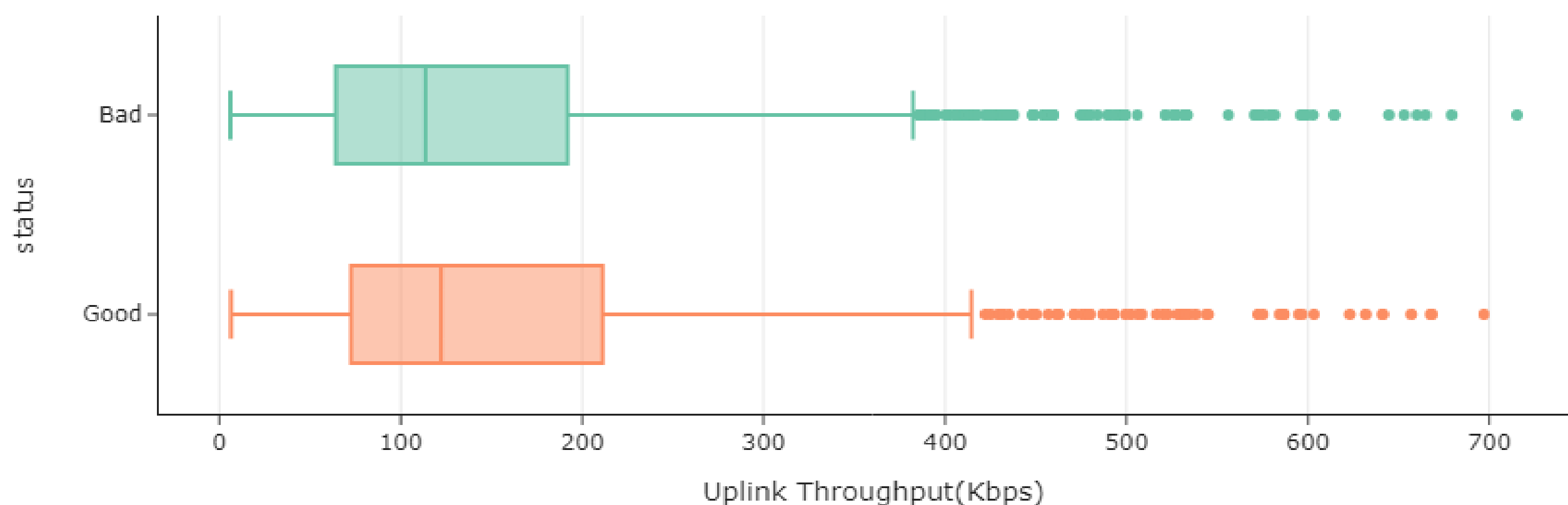


Результаты разведочного анализа

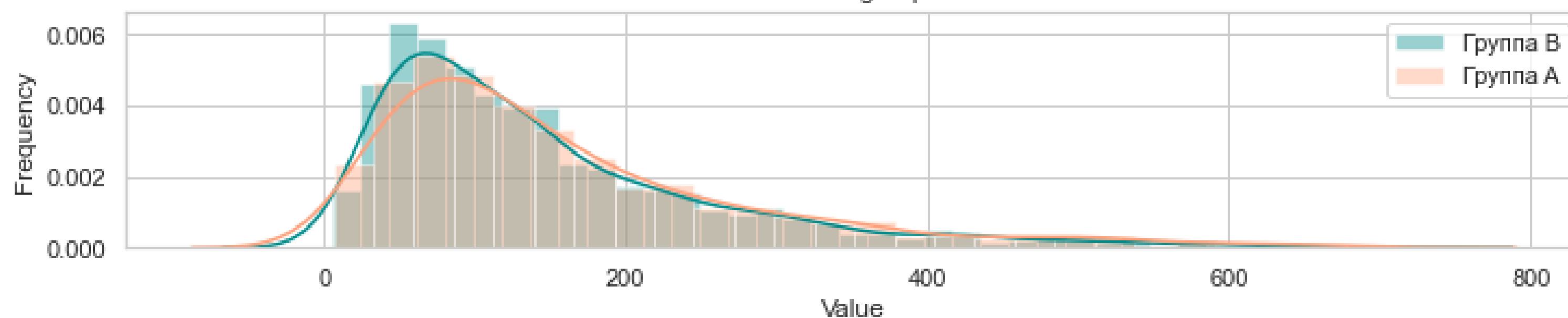
Uplink Throughput(Kbps)

МКР и максимум выше у группы с хорошими оценками, значит, можно сделать вывод, что средняя скорость «от абонента» выше у группы с хорошими оценками.

Statistical characteristics by parameter Uplink Throughput(Kbps) groups



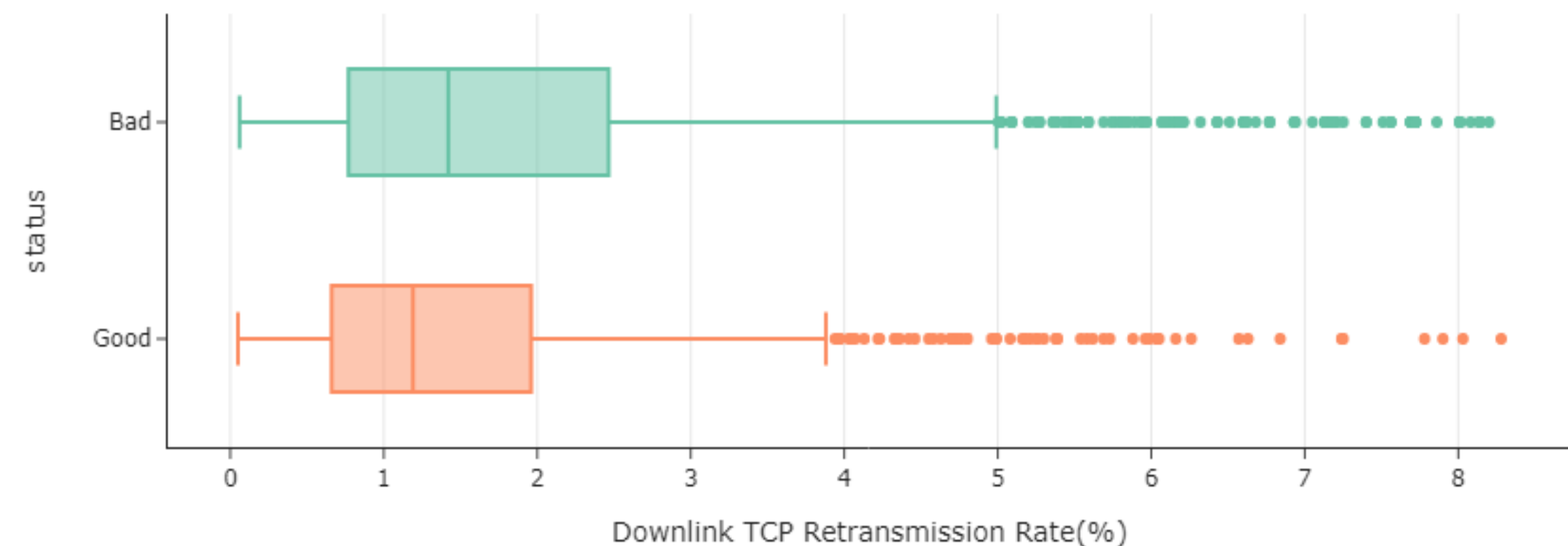
Hist of distribution Uplink Throughput(Kbps) for two groups



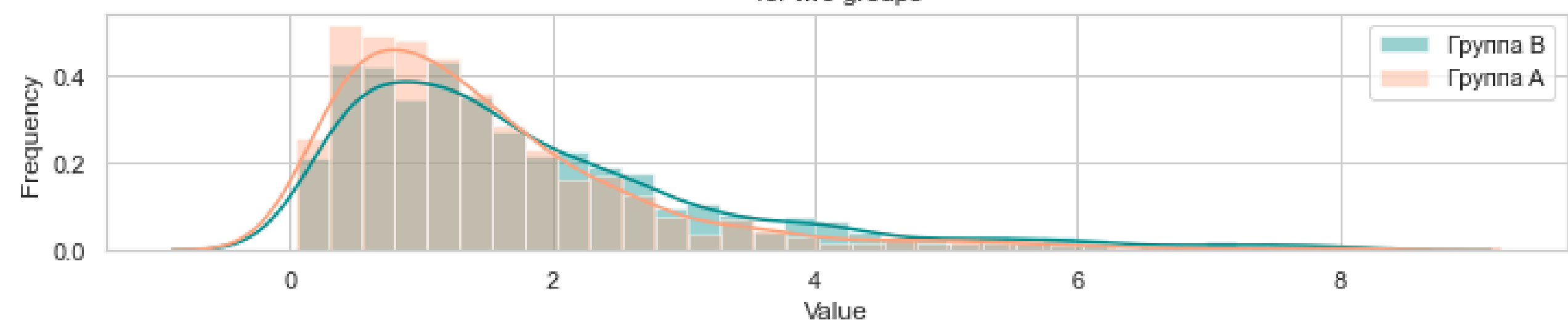
Результаты разведочного анализа

По этим графикам по статистическим характеристикам уже лидирует группа В, но это закономерно, т.к. чем выше параметр Downlink TCP Retransmission Rate(%) тем хуже, пакет переотправляется, снижается полезная скорость.

Statistical characteristics by parameter Downlink TCP Retransmission Rate(%) groups



Hist of distribution Downlink TCP Retransmission Rate(%) for two groups



Результаты разведочного анализа

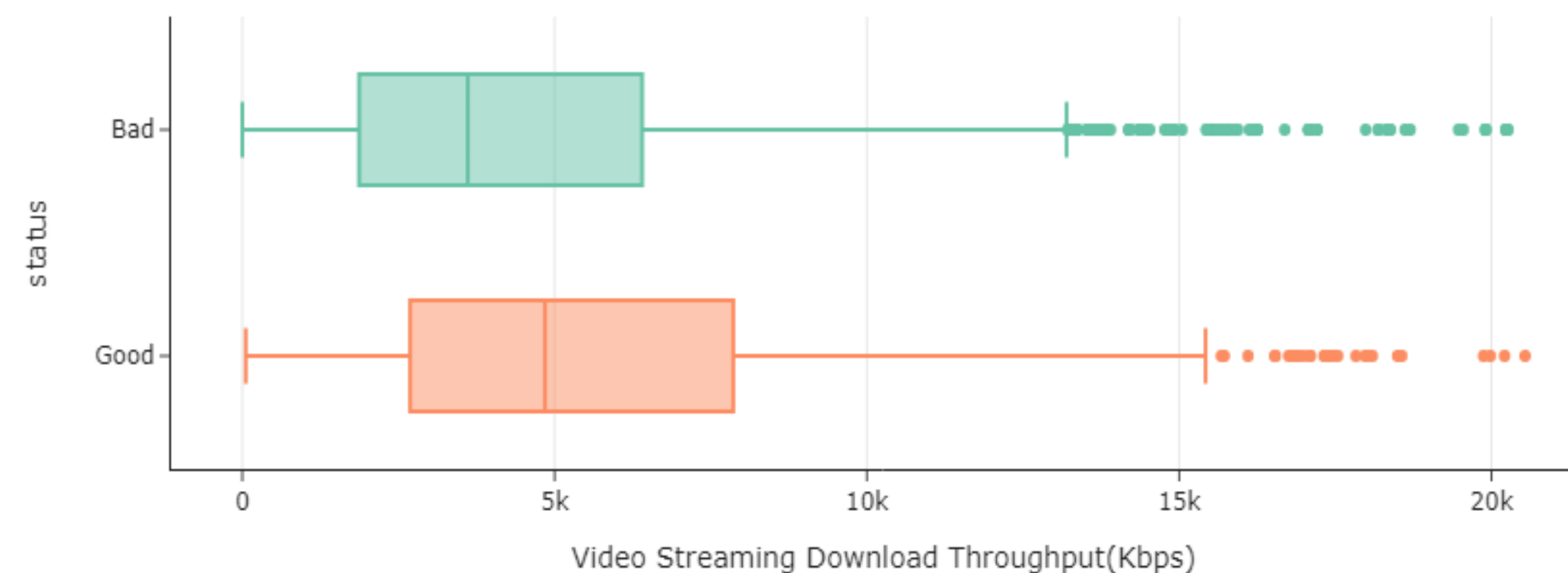
Video Streaming Download Throughput(Kbps)

Характеристики у группы с хорошими оценками для этой метрики выше, чем у группы В.

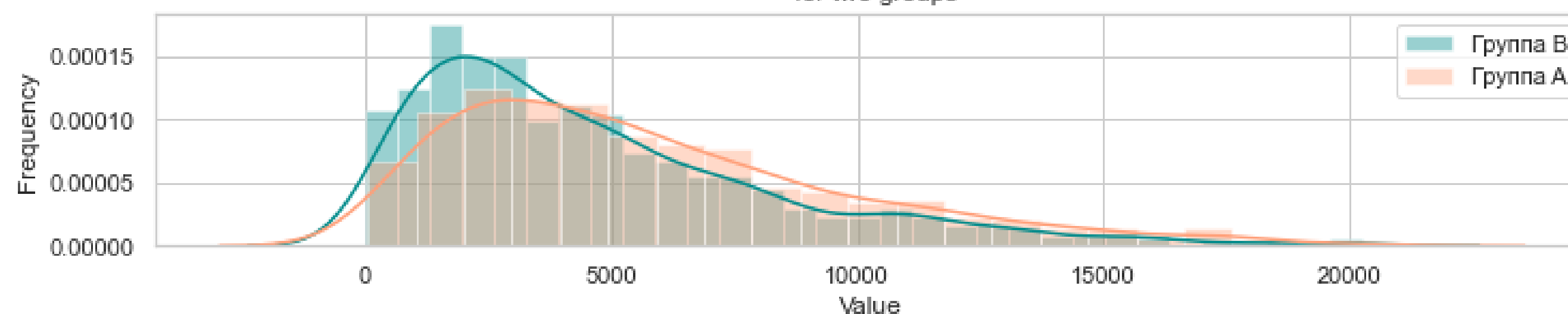
Значит, у абонентов из группы

А скорость загрузки видео выше, чем у группы В.

Statistical characteristics by parameter Video Streaming Download Throughput(Kbps) groups



Hist of distribution Video Streaming Download Throughput(Kbps) for two groups



Результаты разведочного анализа

Статистические

характеристики метрики **Video**

Streaming xKB Start

Delay(ms) выше по группе В и

ниже у группы А, что можно

интерпретировать следующим

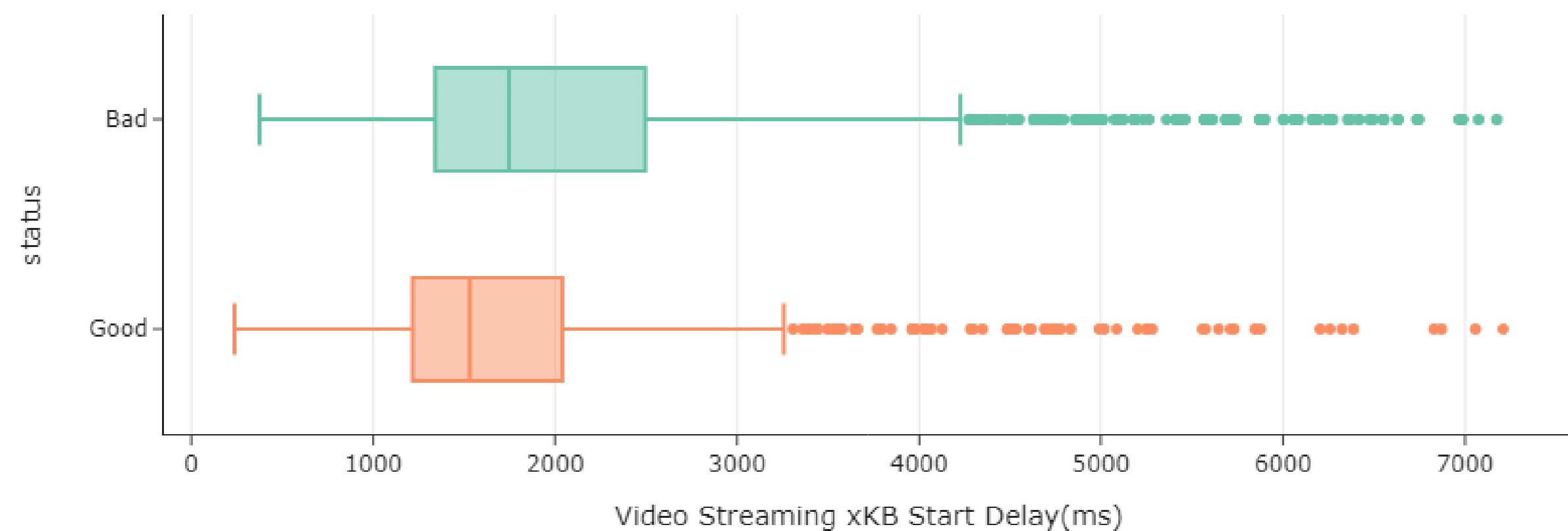
образом: у абонентов группы

В задержка старта

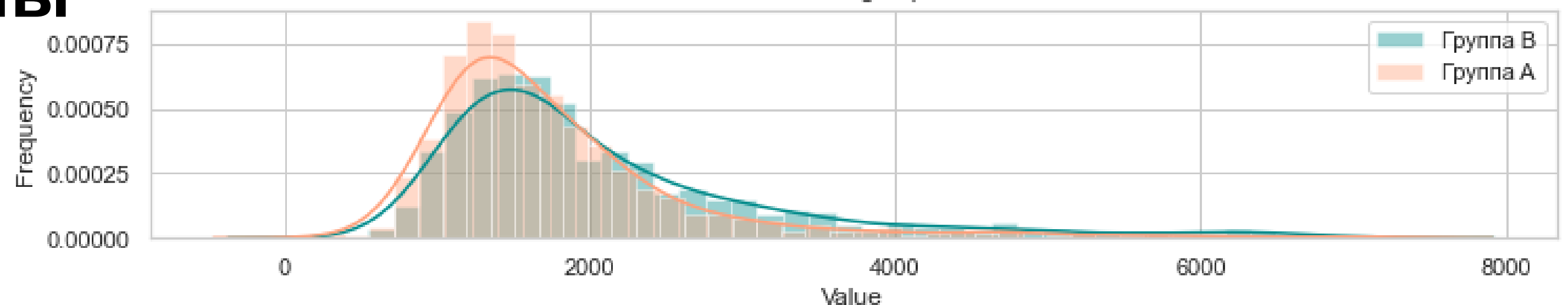
воспроизведения видео

выше, чем группы А.

Statistical characteristics by parameter Video Streaming xKB Start Delay(ms) groups



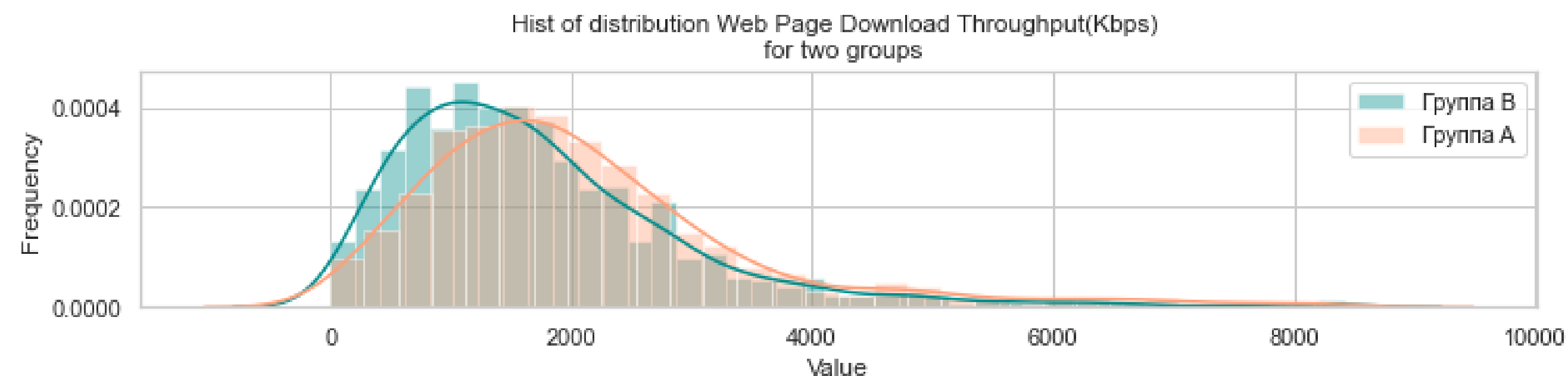
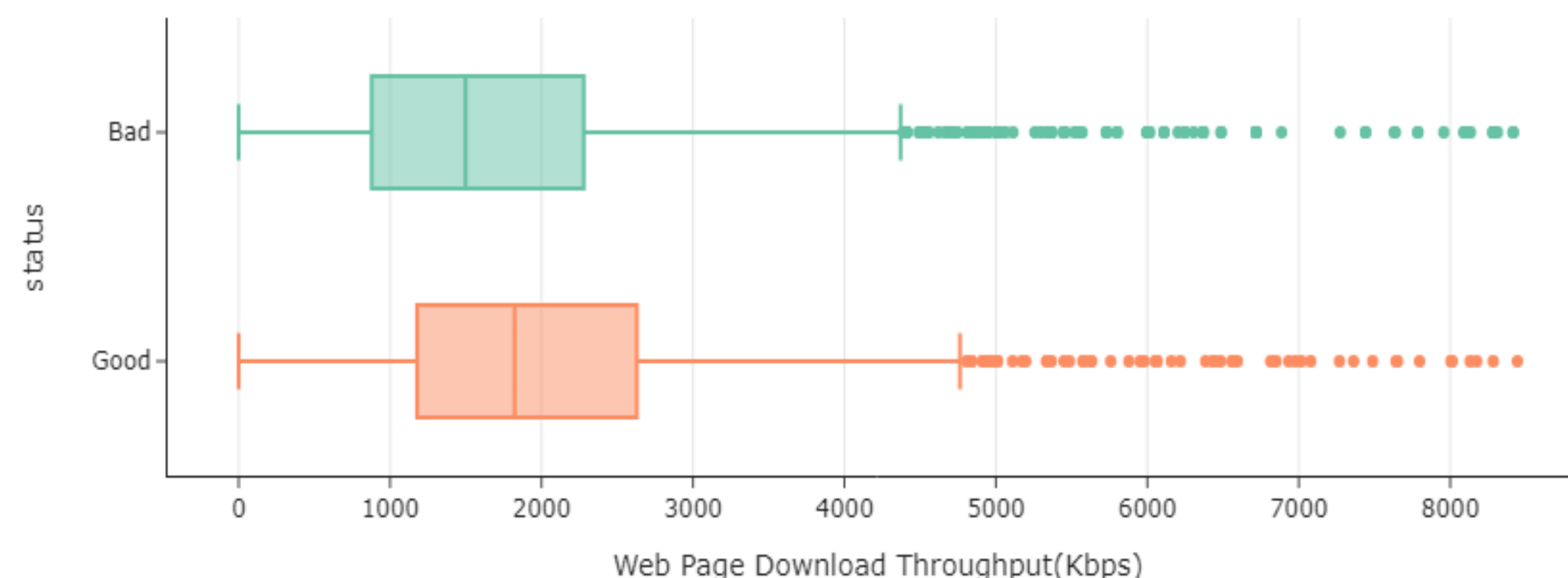
Hist of distribution Video Streaming xKB Start Delay(ms) for two groups



Результаты разведочного анализа

Характеристики метрики **Web Page Download Throughput(Kbps)** выше у группы с хорошими оценками, а значит **скорость загрузки web-страниц через браузер у группы с хорошими оценками выше.**

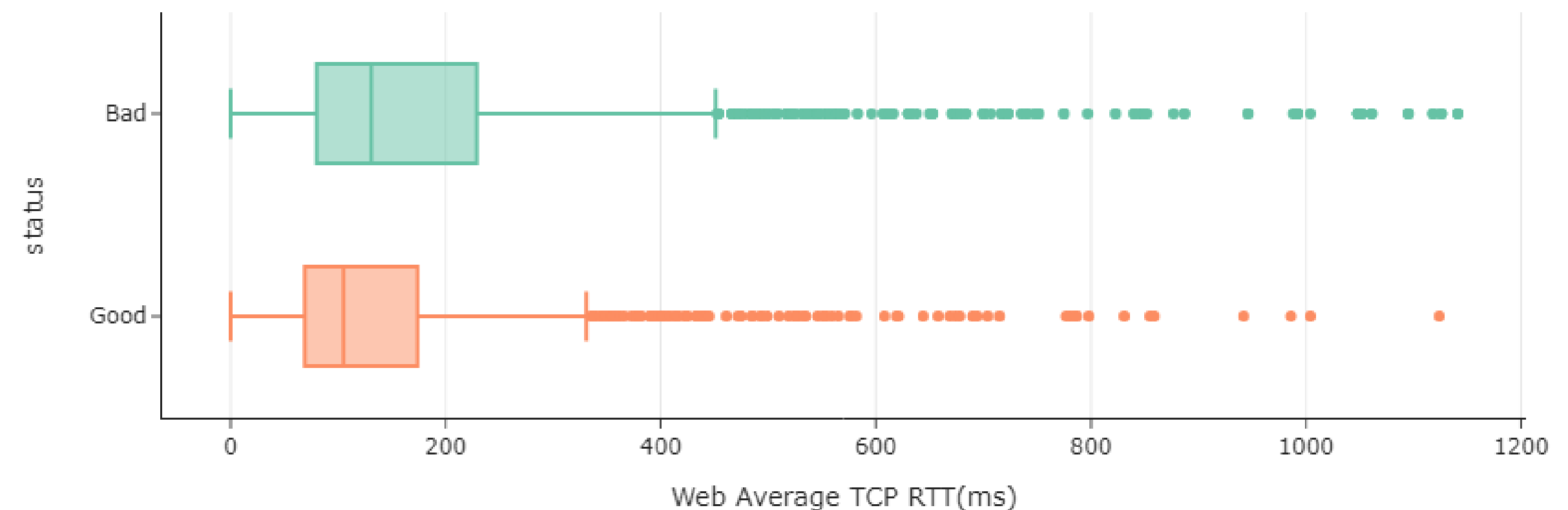
Statistical characteristics by parameter Web Page Download Throughput(Kbps) groups



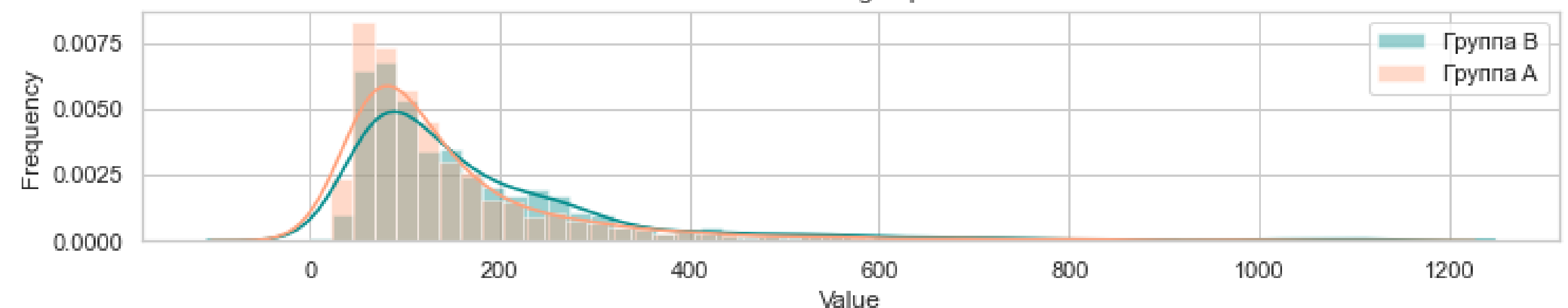
Результаты разведочного анализа

Судя по графикам
статистических характеристик
метрики **Web Average TCP
RTT(ms)** у группы В пинг при
просмотре web-страниц
выше, чем у группы с
хорошими оценками.

Statistical characteristics by parameter Web Average TCP RTT(ms) groups

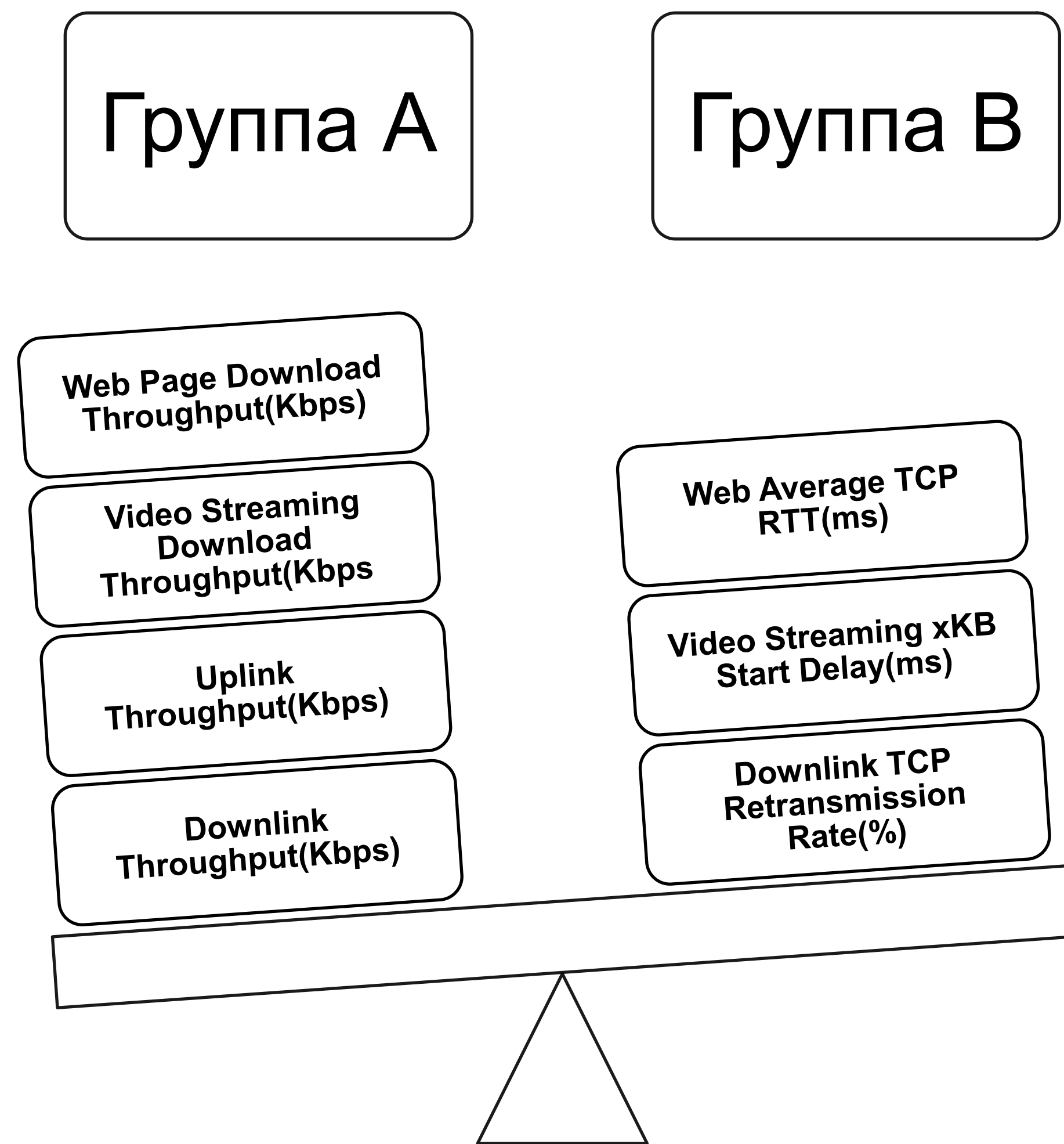


Hist of distribution Web Average TCP RTT(ms)
for two groups



Результаты разведочного анализа

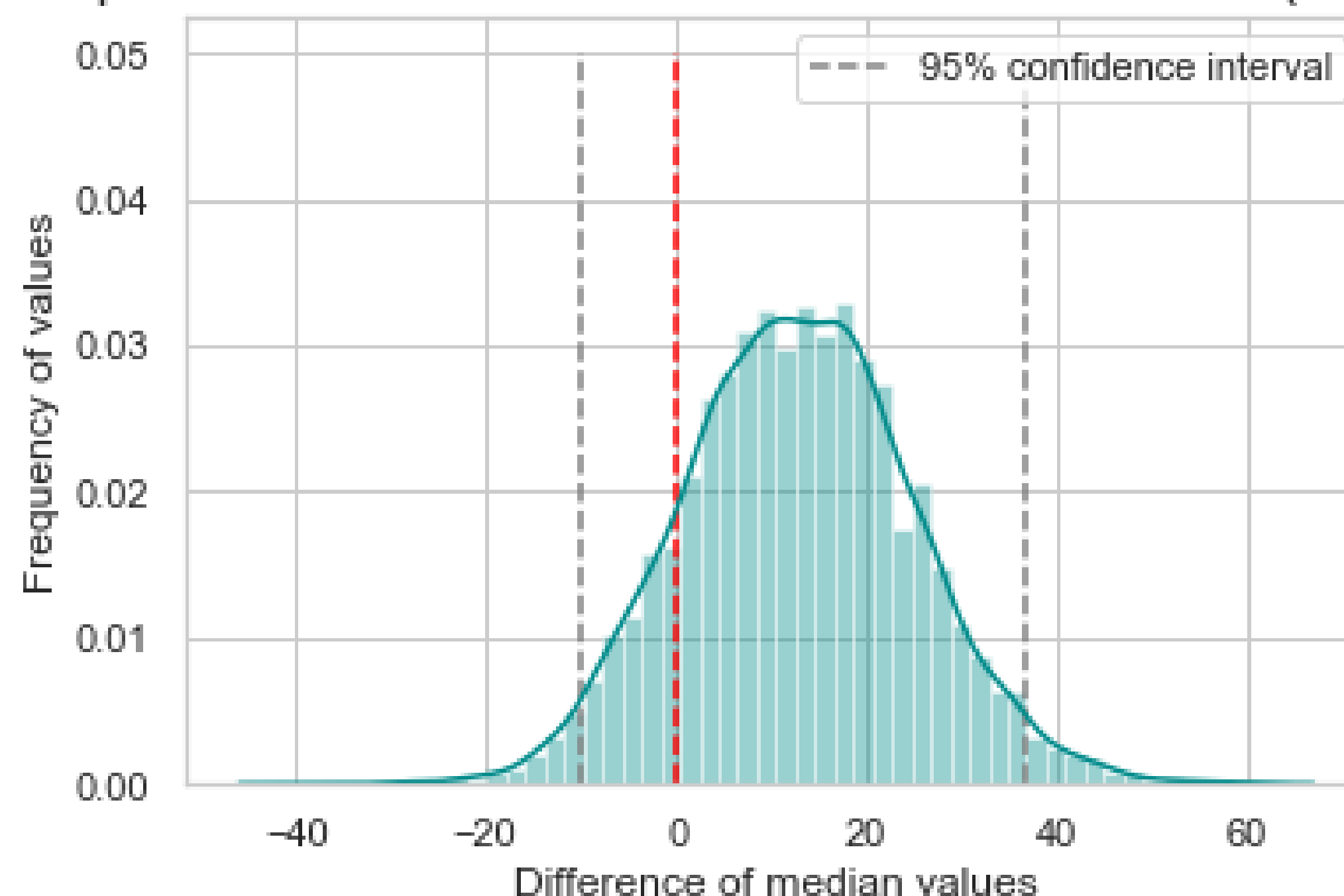
- Объем трафика у групп не отличается
- Средняя скорость «от абонента» и «к абоненту», скорость загрузки видео, а также скорость загрузки web-страниц выше у группы А.
- Частота переотправок пакетов «к абоненту», задержка старта воспроизведения видео и пинг при загрузке web-страниц выше у группы В, что не является хорошим знаком, т.к. чем выше значения этих метрик – тем хуже.



Результаты статистического теста

В ходе теста становится понятно, что не объем полного трафика влияет на удовлетворенность клиентов, более того, в ходе разведочного анализа было выдвинуто предположение, что объем полного трафика для двух групп не отличается, теперь мы это подтвердили на практике.

Bootstrap-distribution of median differences Total Traffic(MB) by groups

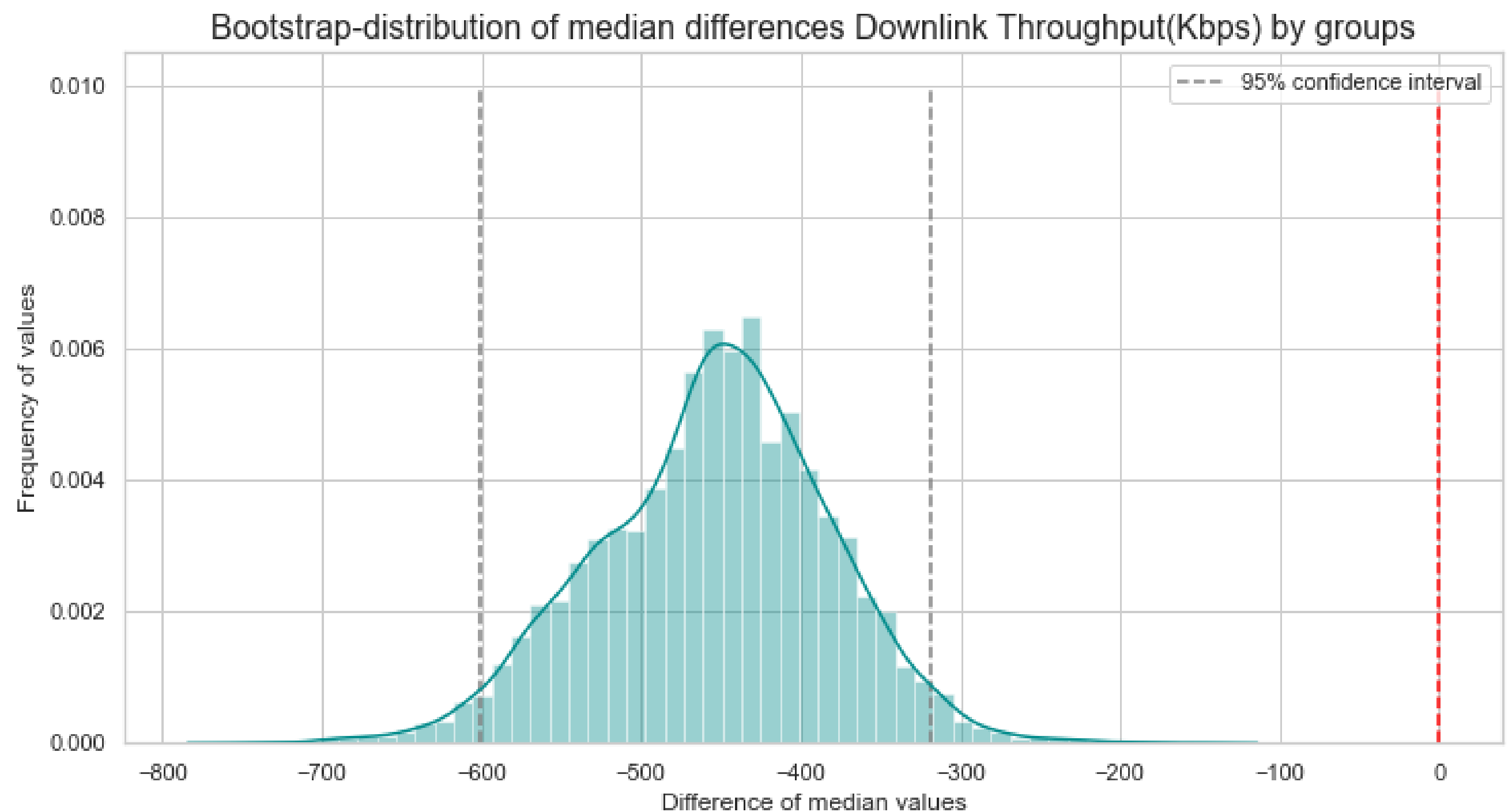


• **$H_0: A=B$**

• **$H_A: A \neq B$**

Результаты статистического теста

Имеется статистически значимая разница в медианных значениях средней скорости "к абоненту" между группами. В ходе тестирования установлено, что средняя скорость "к абоненту" у выше у группы А, что подтверждает предположение, выдвинутое в ходе разведочного анализа.

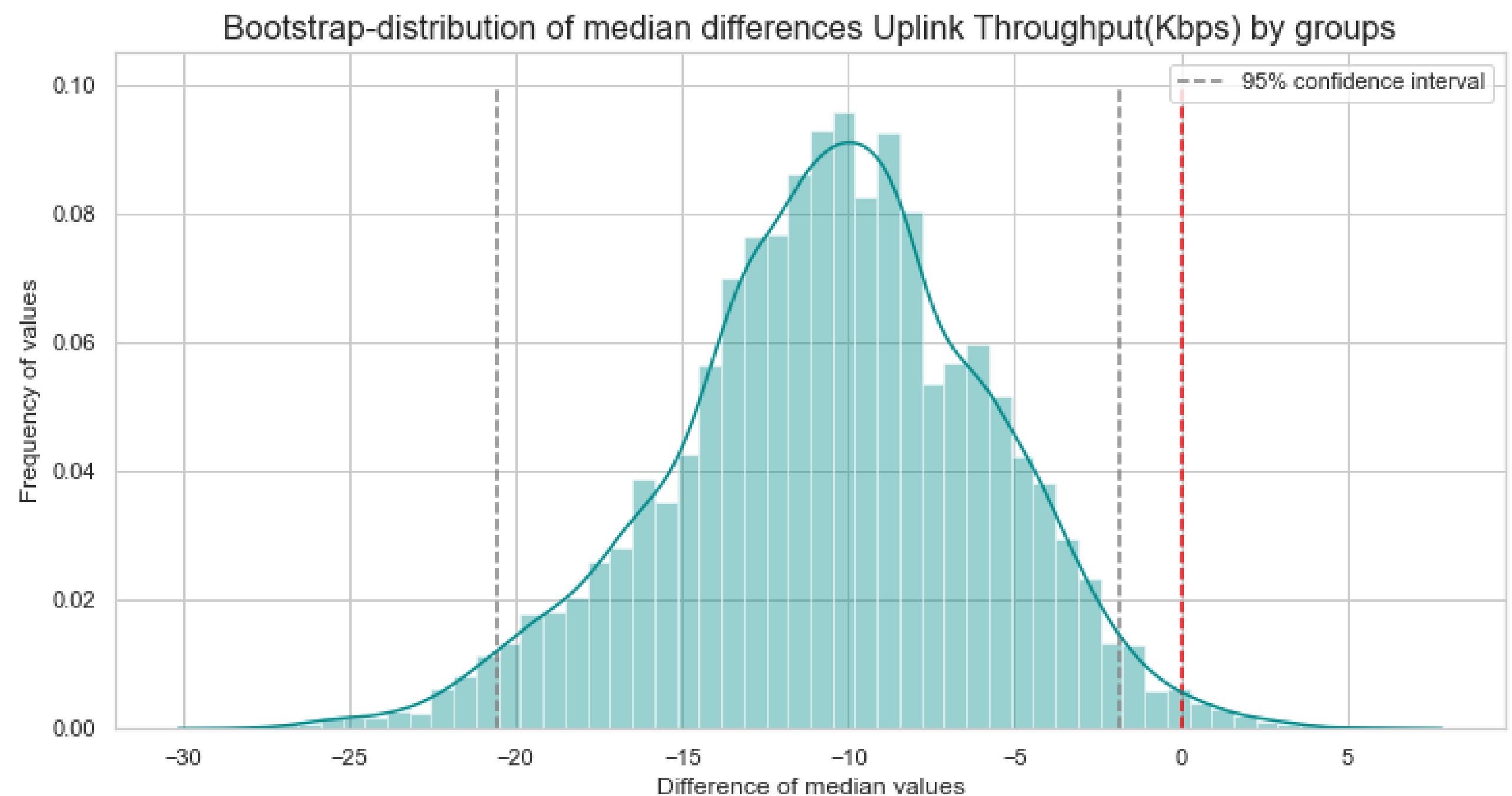


• $H_0: A=B$

• $H_A: A>B$

Результаты статистического теста

Имеется статистически значимая разница в медианных значениях средней скорости «от абонента» между группами. В ходе тестирования установлено, что средняя скорость «от абонента» у выше у группы А, что подтверждает предположение, выдвинутое в ходе разведочного анализа.

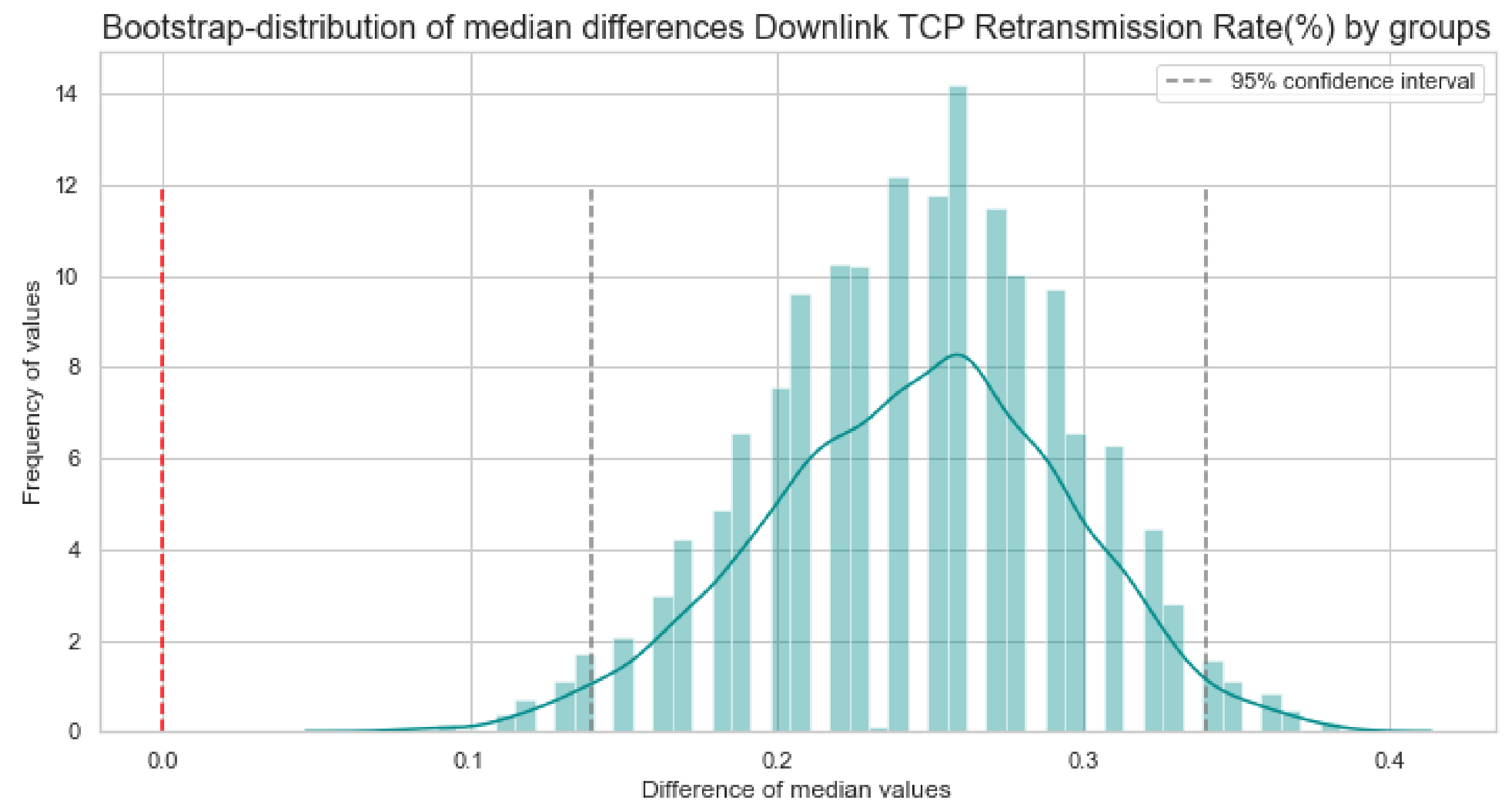


• $H_0: A=B$

• $H_A: A>B$

Результаты статистического теста

Наблюдается статистически значимая разница в значениях частоты переотправок пакетов «к абоненту», поэтому можно сделать вывод, что нулевую гипотезу мы можем отвергнуть в пользу альтернативной. У группы В частота переотправок пакетов «к абоненту» выше, чем у группы А.

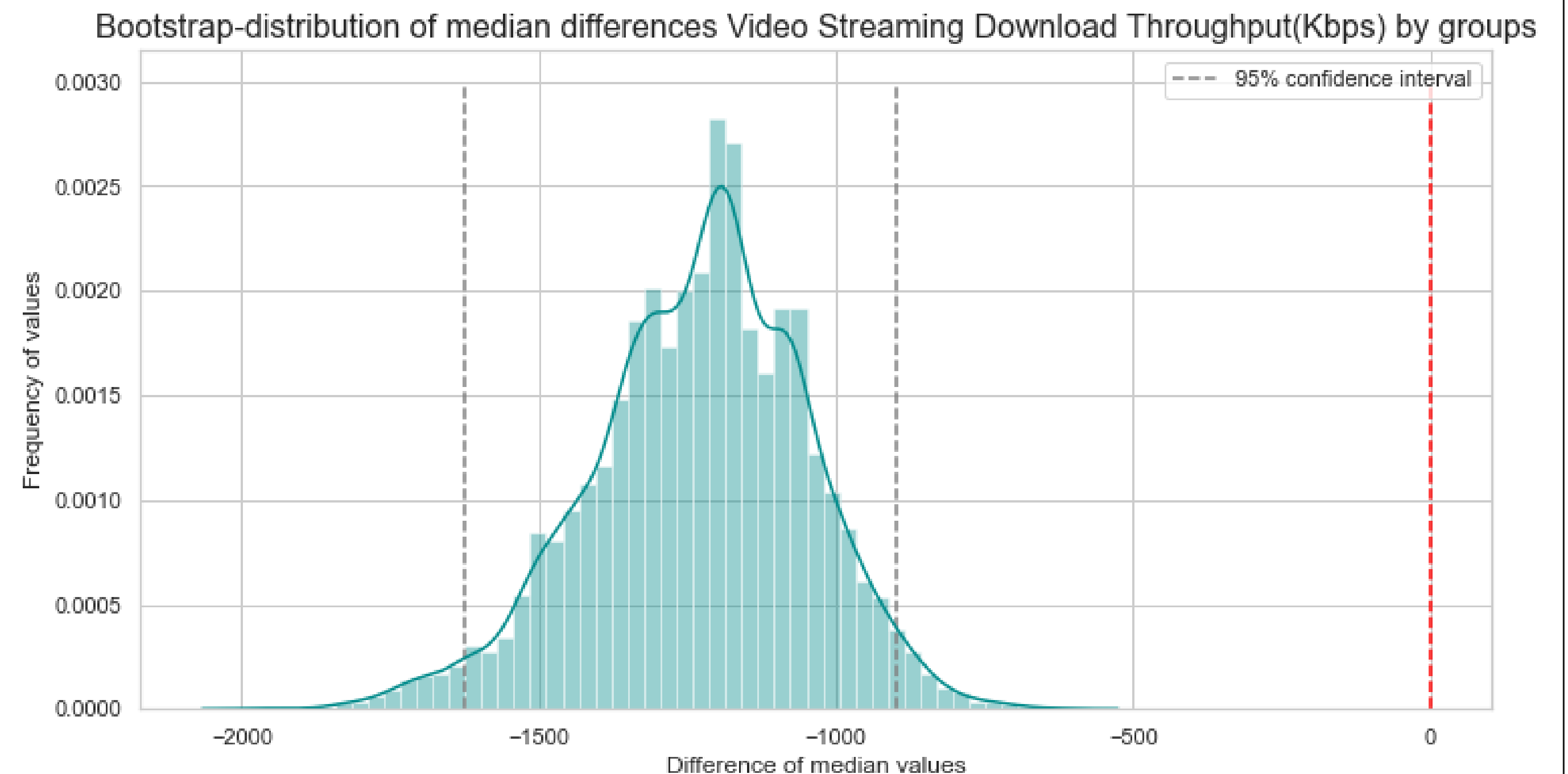


• $H_0: A=B$

• $H_A: A < B$

Результаты статистического теста

Статистически значимая
разница наблюдается, т.к. 0 не
входит в доверительный
интервал, поэтому можно
сделать вывод, что нулевую
гипотезу мы можем отвергнуть
в пользу альтернативной.
Скорость загрузки потокового
видео выше у группы A.

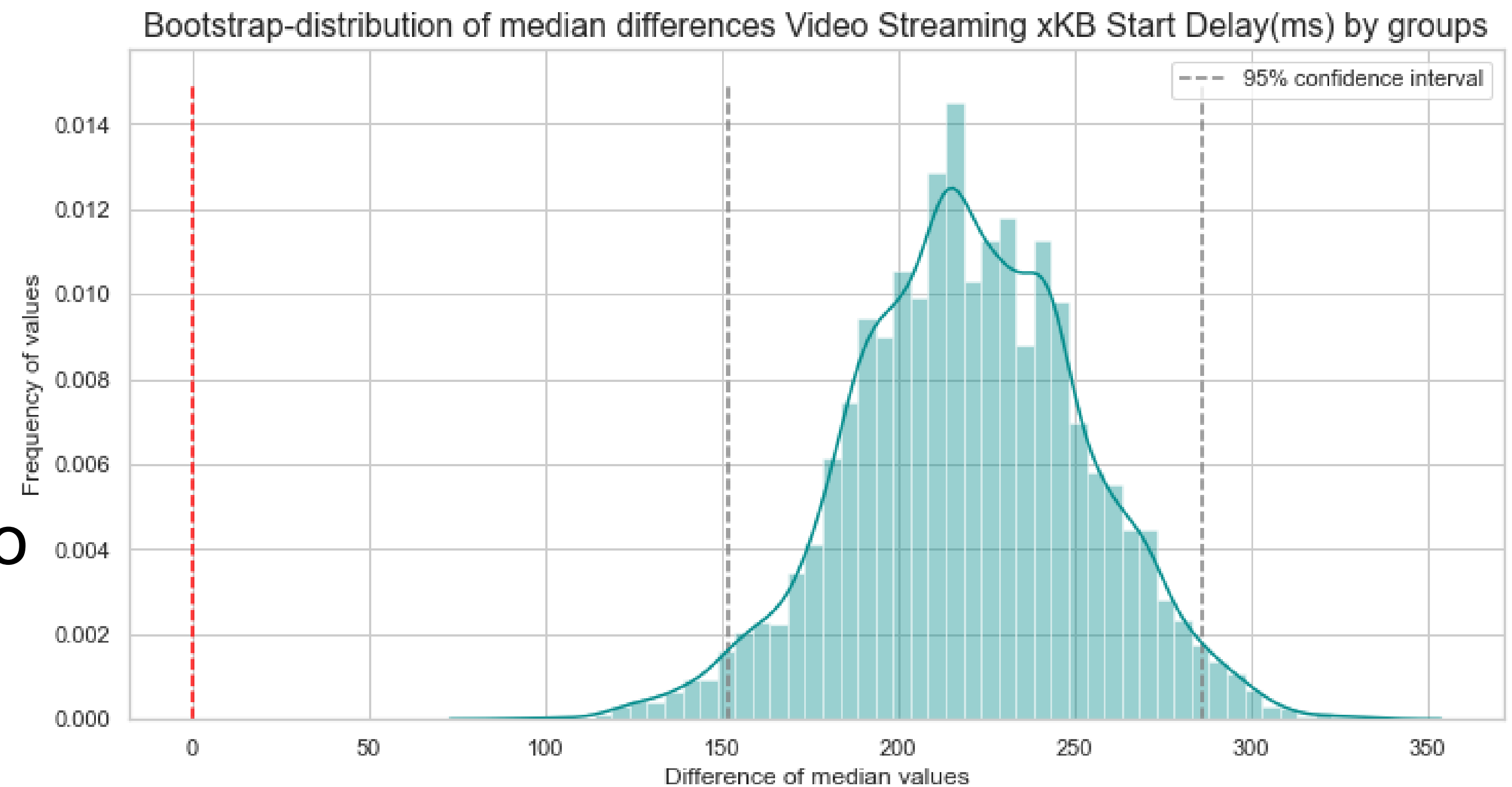


• $H_0: A=B$

• $H_A: A>B$

Результаты статистического теста

Наблюдается статистически значимая разница медианной задержки старта воспроизведения видео. У абонентов группы В задержка старта воспроизведения видео выше, чем у второй группы, что полностью совпадает с выводом разведочного анализа.

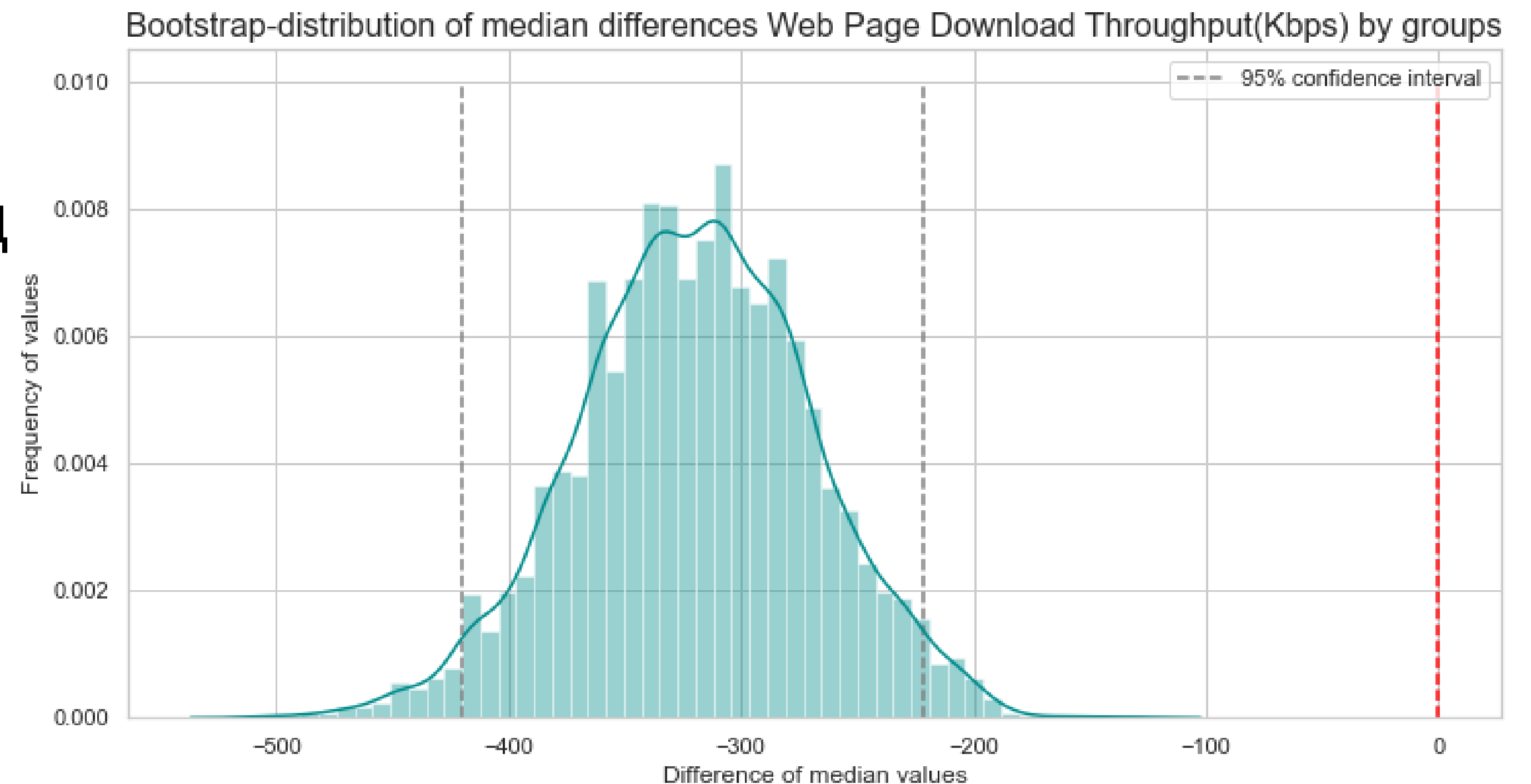


• $H_0: A=B$

• $H_A: A < B$

Результаты статистического теста

Имеется статистически значимая разница медианной скорости загрузки web-страниц через браузер, 0 не входит в доверительный интервал, значит, нулевую гипотезу мы отвергаем в пользу альтернативной. Скорость загрузки web-страниц через браузер у группы А выше, чем у группы В.

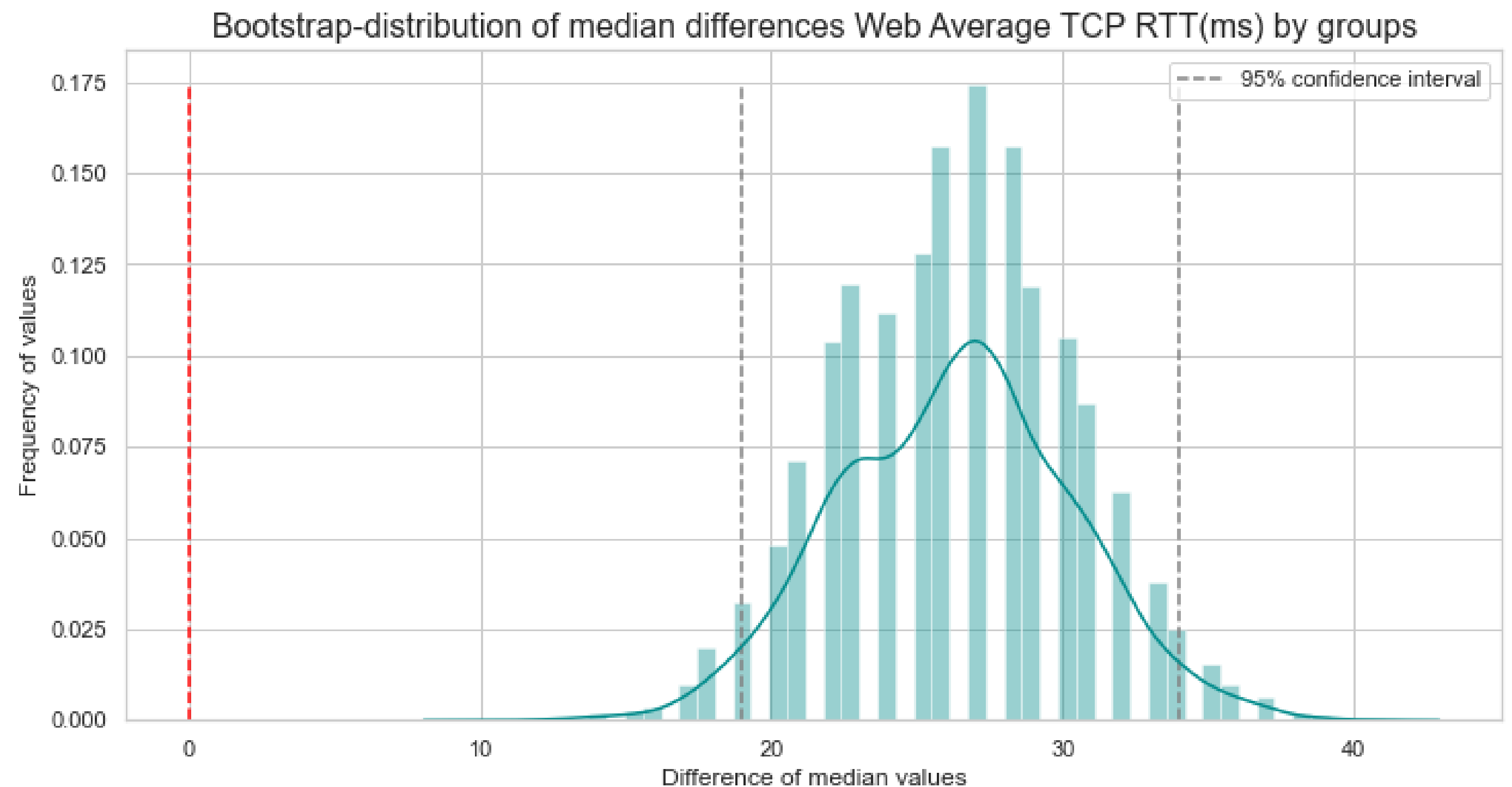


• $H_0: A=B$

• $H_A: A>B$

Результаты статистического теста

Имеется статистически значимая разница медианны пинга при просмотре web-страниц, т.к. 0 не входит в доверительный интервал, что значит, что нулевую гипотезу мы отвергаем в пользу альтернативной. У группы В пинг при просмотре web-страниц выше, чем у группы А.



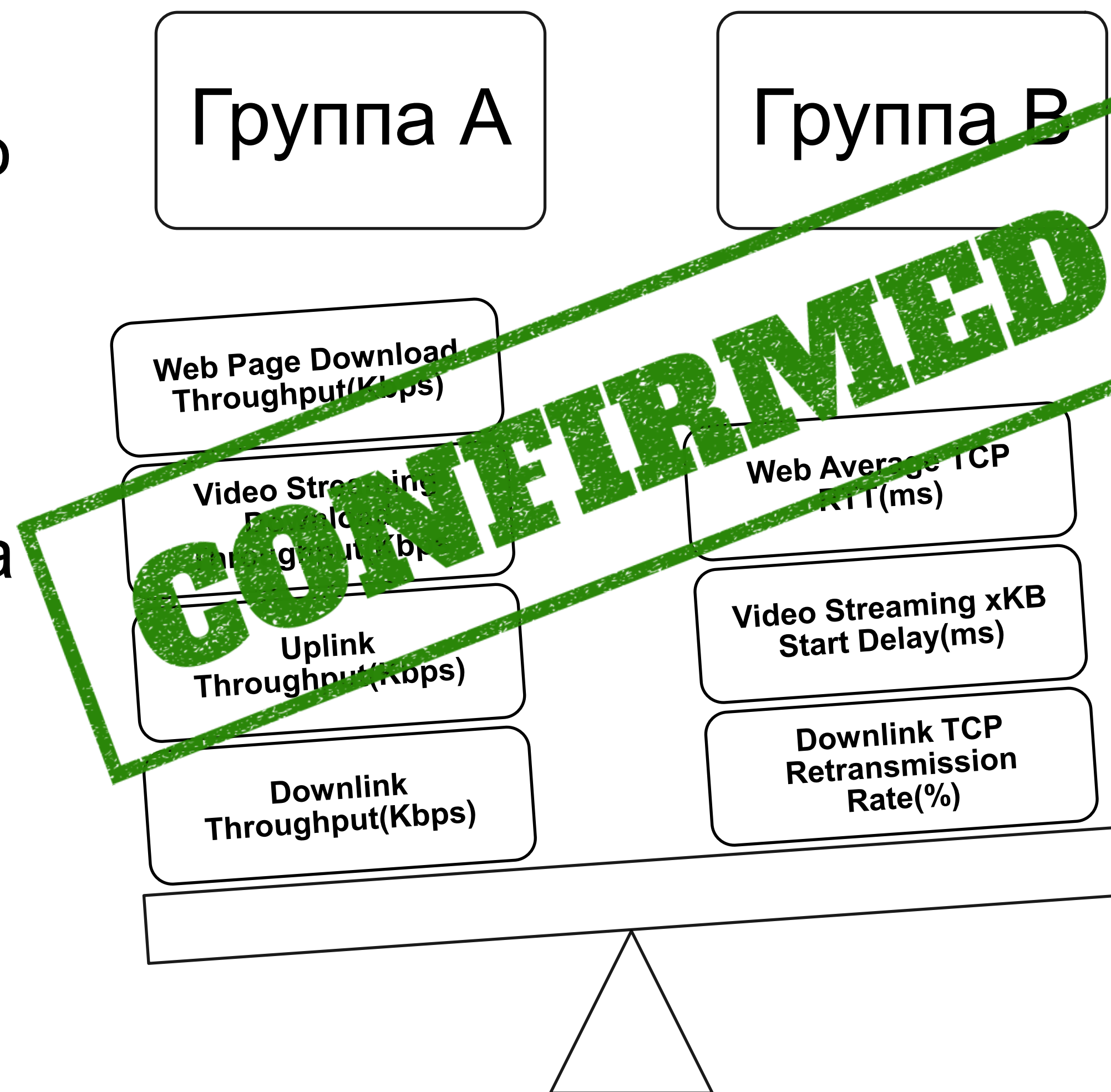
• $H_0: A=B$

• $H_A: A < B$

Вывод по предмету анализа

Выводы, сделанные в ходе разведочного анализа – подтвердились на практике.

С вероятностью 95% можно утверждать, что все метрики, за исключением объема полного трафика характеризующие качество предоставляемых услуг имеют статистически значимое влияние на удовлетворенность клиентов.



Вывод по проведенной работе

Из всего вышеописанного можно заключить, что все задачи из первого раздела исследования с постановкой задач - выполнены.

А одно из двух предположений - подтверждено, а именно, то что, если показатели метрик у групп отличаются и пользователям из группы А действительно, по техническим характеристикам, предоставляются услуги лучшего качества, нежели второй группе, руководству компании придется предпринять меры для исправления этой ситуации.

Вывод по проведенной работе

На основании сделанных выводов, руководству компании потребуется принять меры для исправления данной ситуации, потому что между группами пользователей действительно имеется статистически значимая разница в качестве предоставляемых услуг.

Это позволит добиться хорошего уровня удовлетворенности пользователей, завоевать доверие клиентов на рынке услуг сотовой связи, чтобы они могли рекомендовать продукты компании своим знакомым, покупать дополнительные услуги, что в целом, позволит бизнесу занять лидирующее место на рынке, и как следствие повысить итоговую выручку.

Спасибо за внимание!