



# Структура данных



Данные в период с  
01.06.2024 по 21.06.2024.

**21** день

**3** недели

**7 000 000** строк

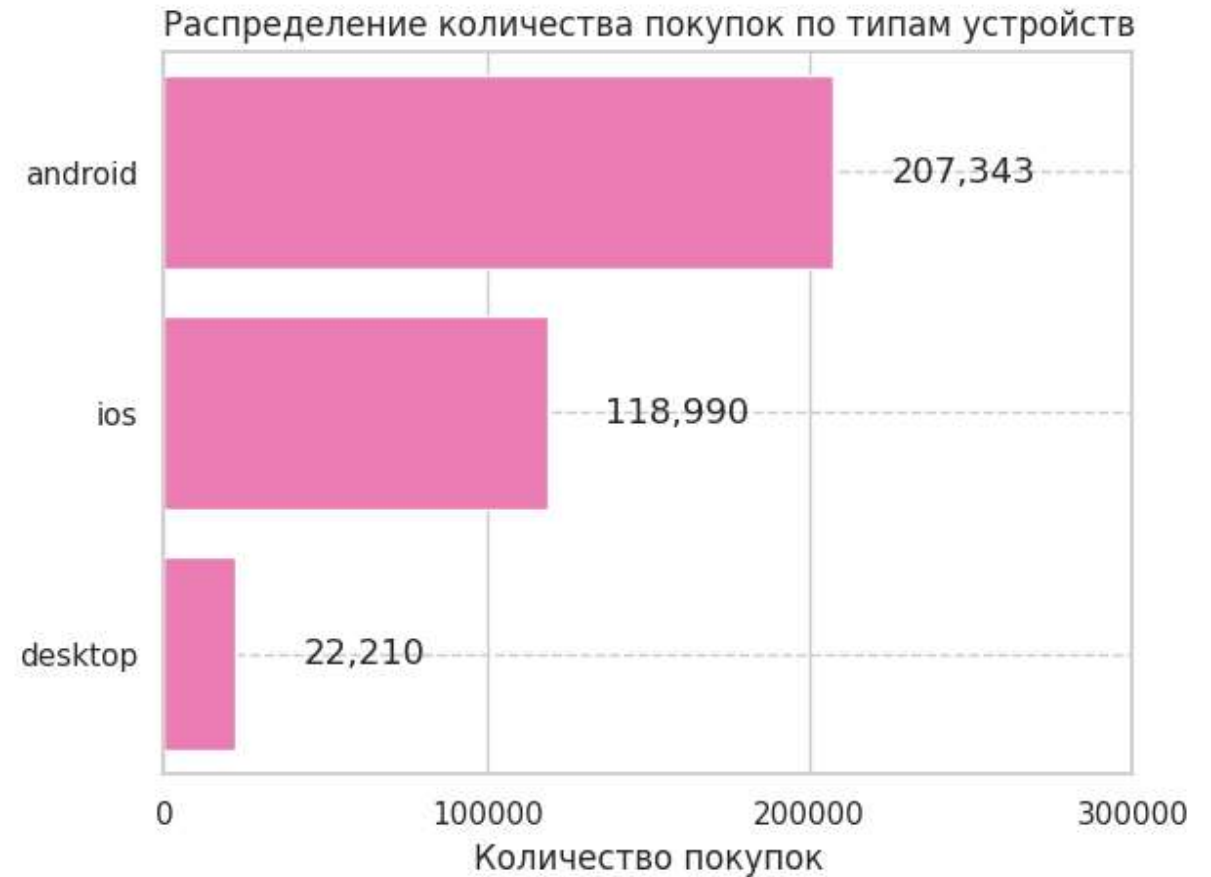
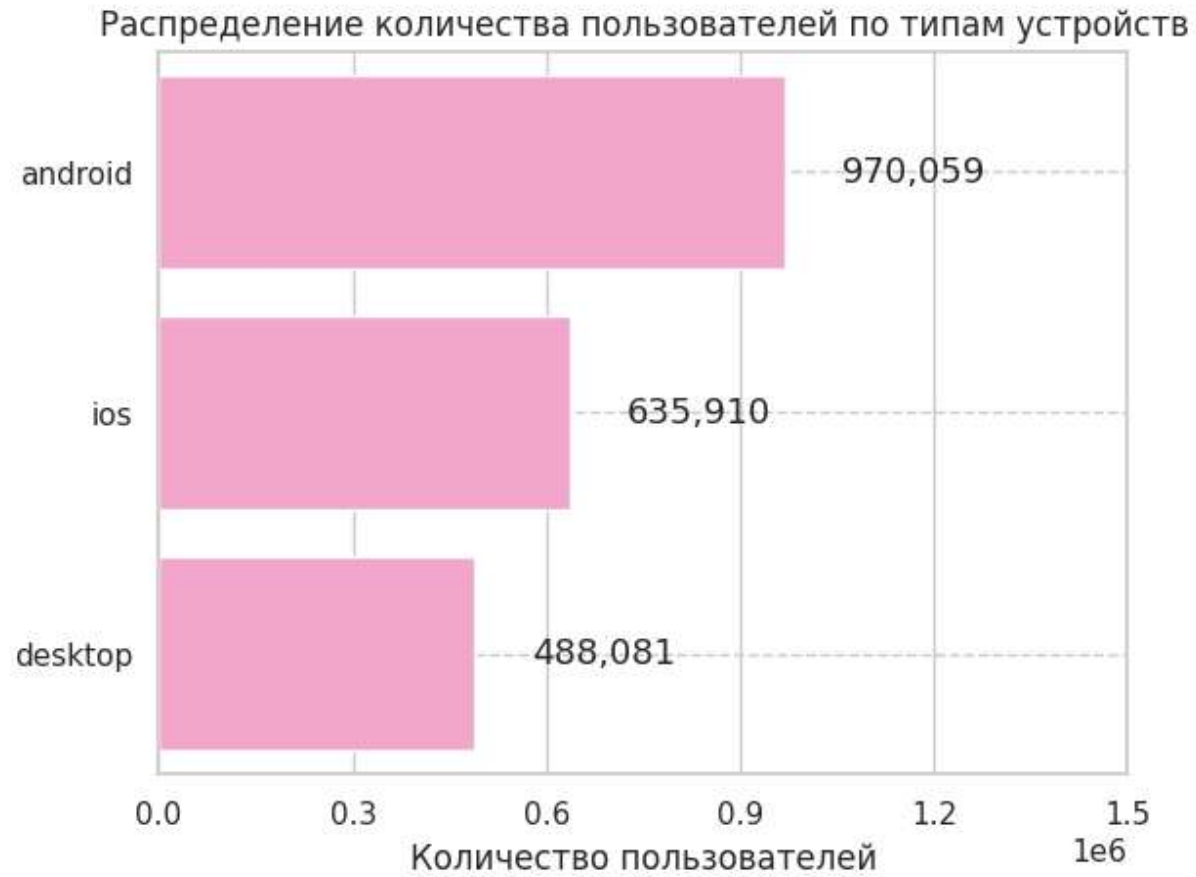
**2 108 035** пользователей

**90 375** товаров

# Предварительный анализ



# Предварительный анализ



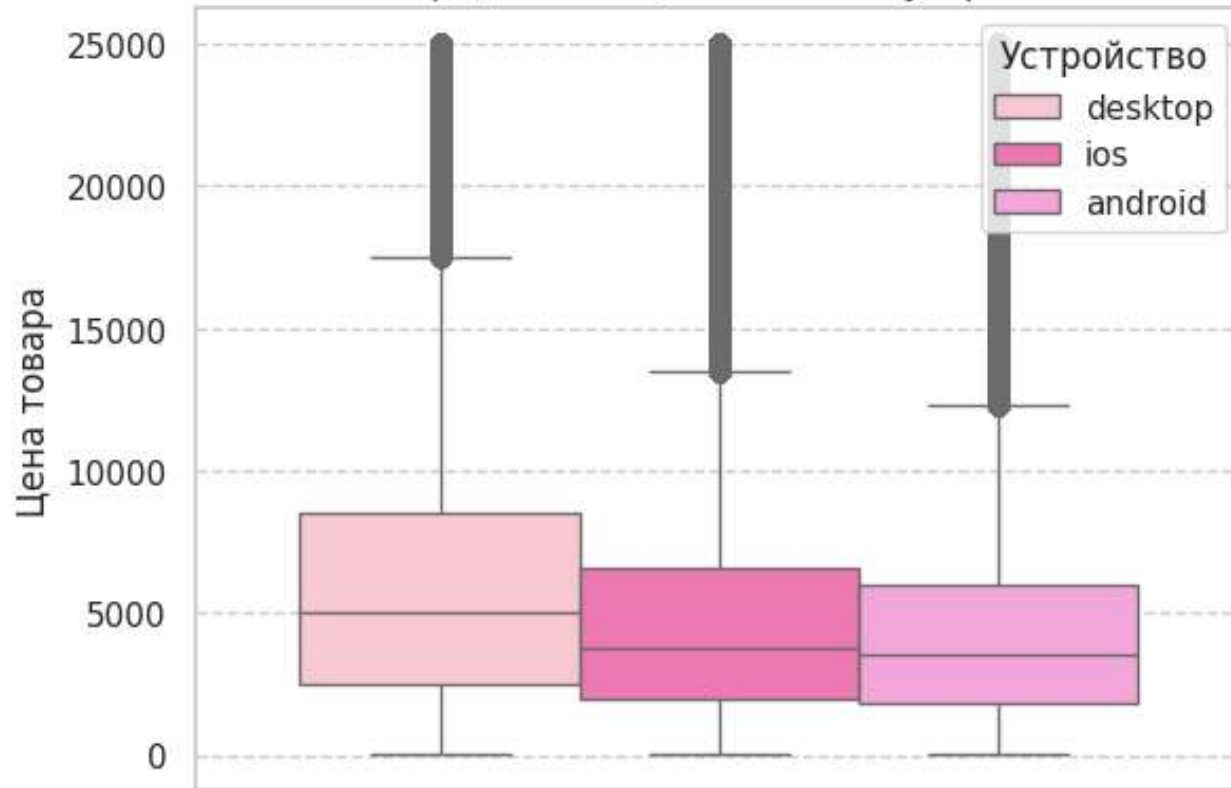
# Предварительный анализ



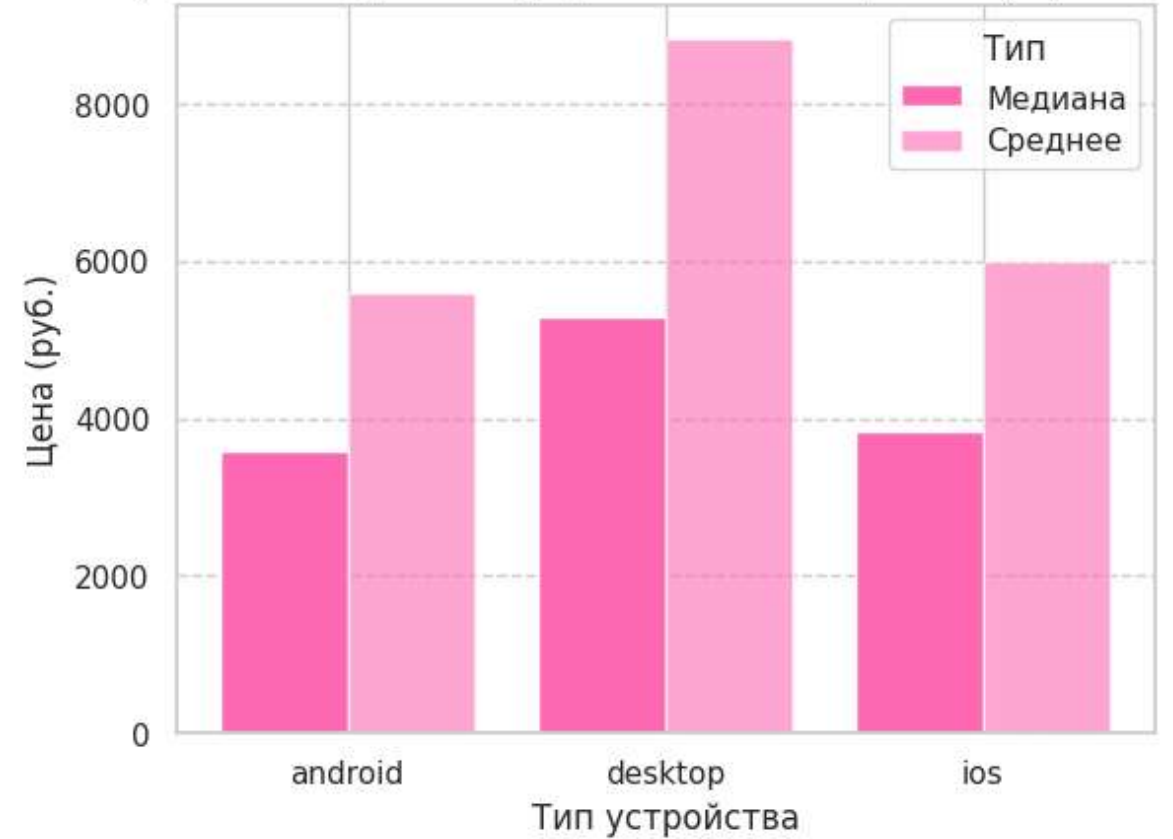
Конверсия - отношение пользователей, которые совершили нужное нам целевое действие, к числу пользователей на предыдущем этапе

# Предварительный анализ

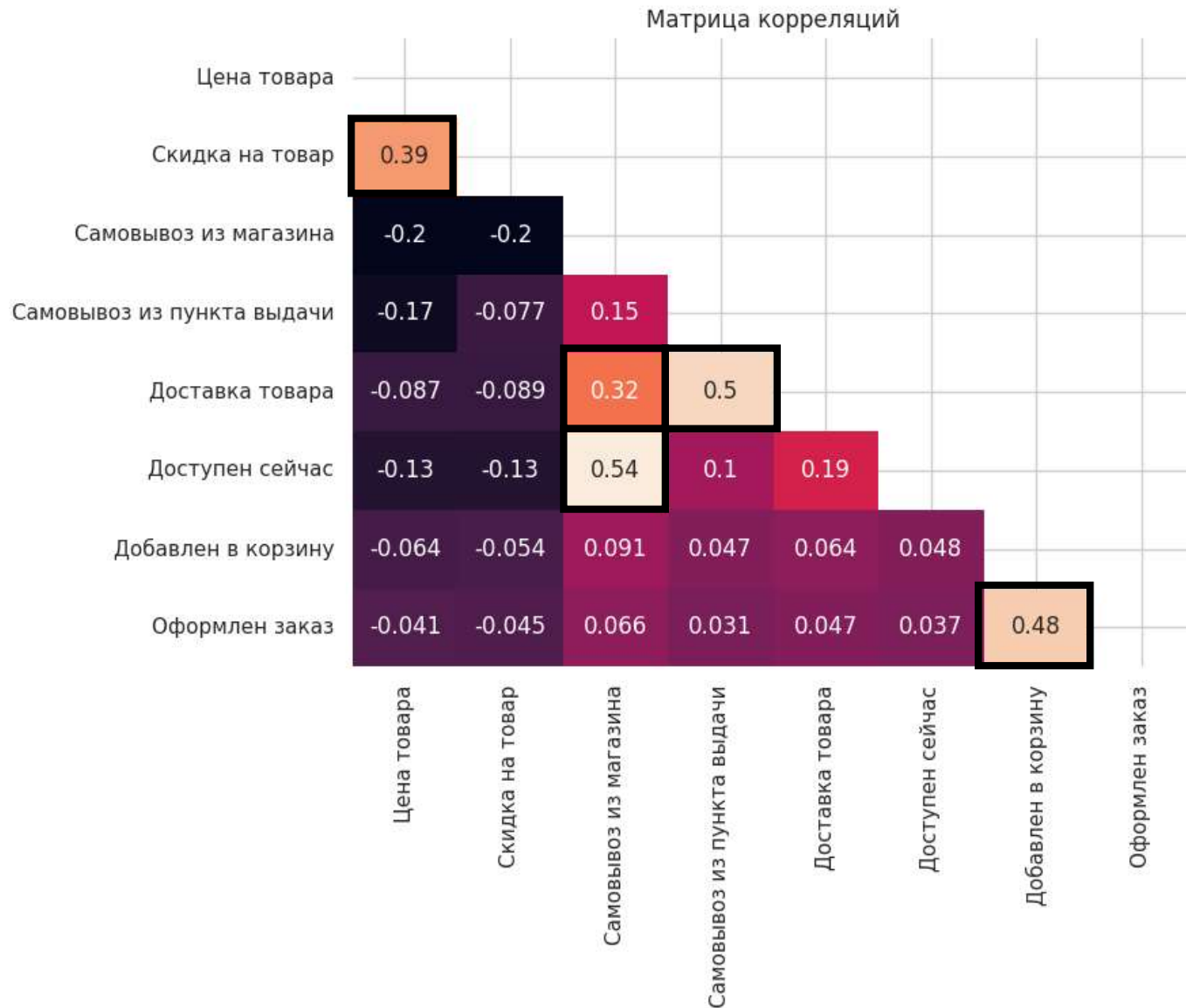
Распределение цен по типам устройств



Сравнение медианы и среднего значения цены по устройствам



# Корреляционный анализ



- Корреляции очевидны и (или) незначительны
- Корреляции между бинарными переменными легко объяснимы

# Подготовка базы данных



**34**  
дубликата\*



**11**  
пустых значений

Последний день заполнен не полностью

Возможно **влияние**  
на исследование



остается

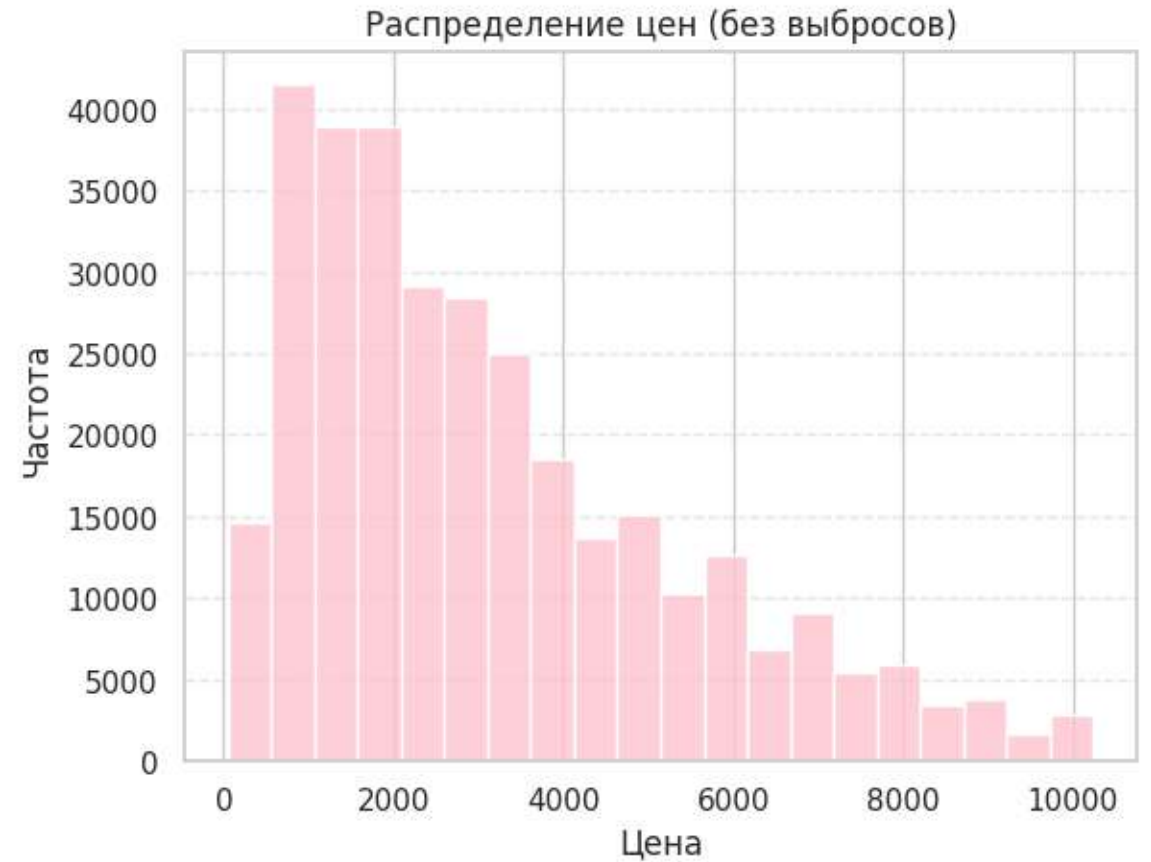
**6 937 646** строк

\*Доп. 6 пользователей покупали товары с других устройств



# Подготовка базы данных

Удаляем выбросы с помощью метода IQR\*



\*Межквартильный размах определяется как разница между нижним (Q1) и верхним (Q3) квартилями. Выбросы определяются как значения, находящиеся за пределами  $1,5 \text{ IQR}$  от Q1 и Q3.

# Исследовательский вопрос



Как средняя цена покупки зависит от платформы пользователя?

## Гипотеза



Пользователи сайта обычно оформляют заказы на товары с более высокой средней ценой по сравнению с пользователями приложения.

# Механизм гипотезы

Более **зрелая**<sup>1</sup> и финансово  
**обеспеченная**<sup>2</sup> аудитория

+

Более **широкий** экран  
монитора пользователя.

+

Более тщательный выбор



**Пользователь совершает  
более дорогую покупку**

\*Прим. 1:  
Из исследования Digital 2021: Global Overview

\*Прим. 2:  
Из исследования "finexpertiza"

# Математическая модель

t-критерий Стьюдента

$$dis = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$$

$$t_{stat} = \frac{x_1 - x_2}{\sqrt{dis}}$$

$x_i$  — выборочное среднее значение для  $i$  группы

$\sigma_i$  — выборочное стандартное отклонение  $i$  группы

$n_i$  — размер выборки  $i$  группы

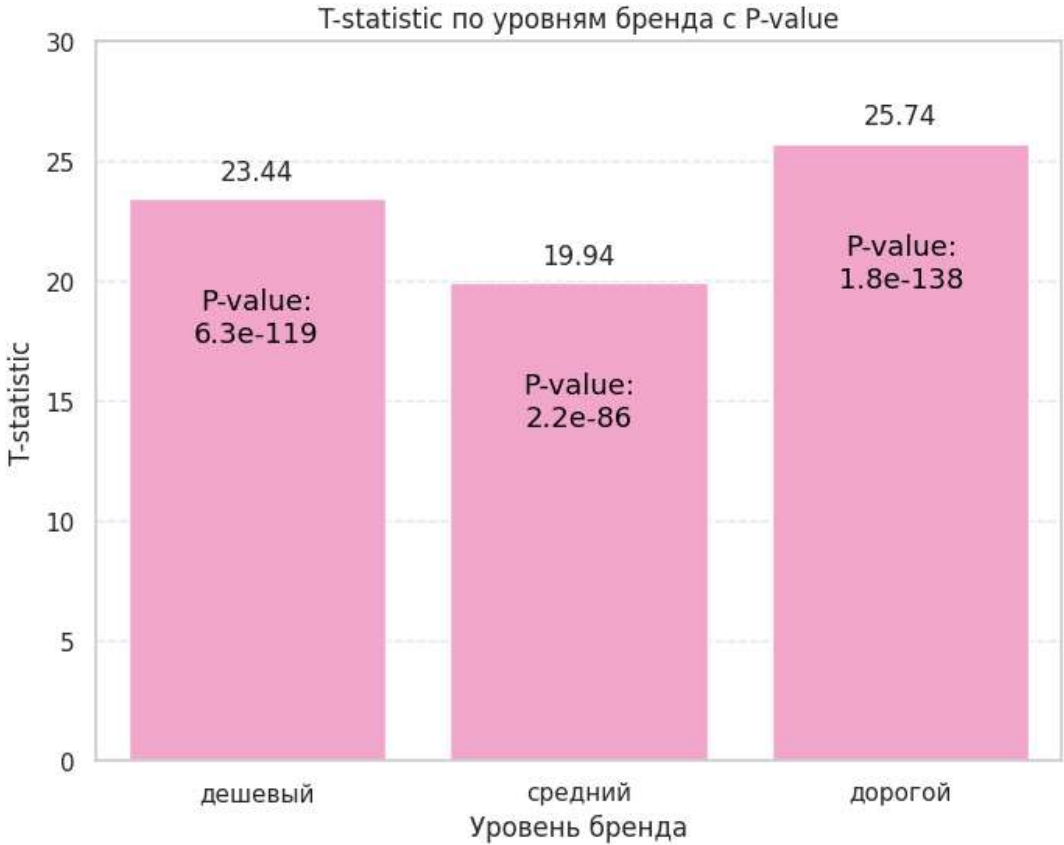
**dis** — разница дисперсий двух групп

# Математическая модель

Результат для всей выборки

| Показатель                 | Значение |
|----------------------------|----------|
| Абсолютная разница средних | 808      |
| T-statistic                | 45       |
| P-value                    | ≈0       |

Результат для разных уровней бренда



P-value < 0.05



Результат является статистически значимым

# Интерпретация



Гипотеза **подтвердилась**, пользователи сайта покупают товары с **большей средней ценой** в отличии от пользователей мобильных устройств

# Ограничения

Особенности этого  
месяца

Недостаточно данных  
о пользователях

Нельзя считать, что наши выводы будут работать  
на других временах года

# Перспективы



Большой **охват** по  
времени



Данные за другое  
**время** года



Покупки из  
**магазина**



Информация  
о **клиентах**



# Практическое применение



Провести A/B тест на более дорогие рекомендации

Расширить функционал сайта

- Персонализированные рекомендации
- 3D обзор товаров

Специальные предложения для платформы

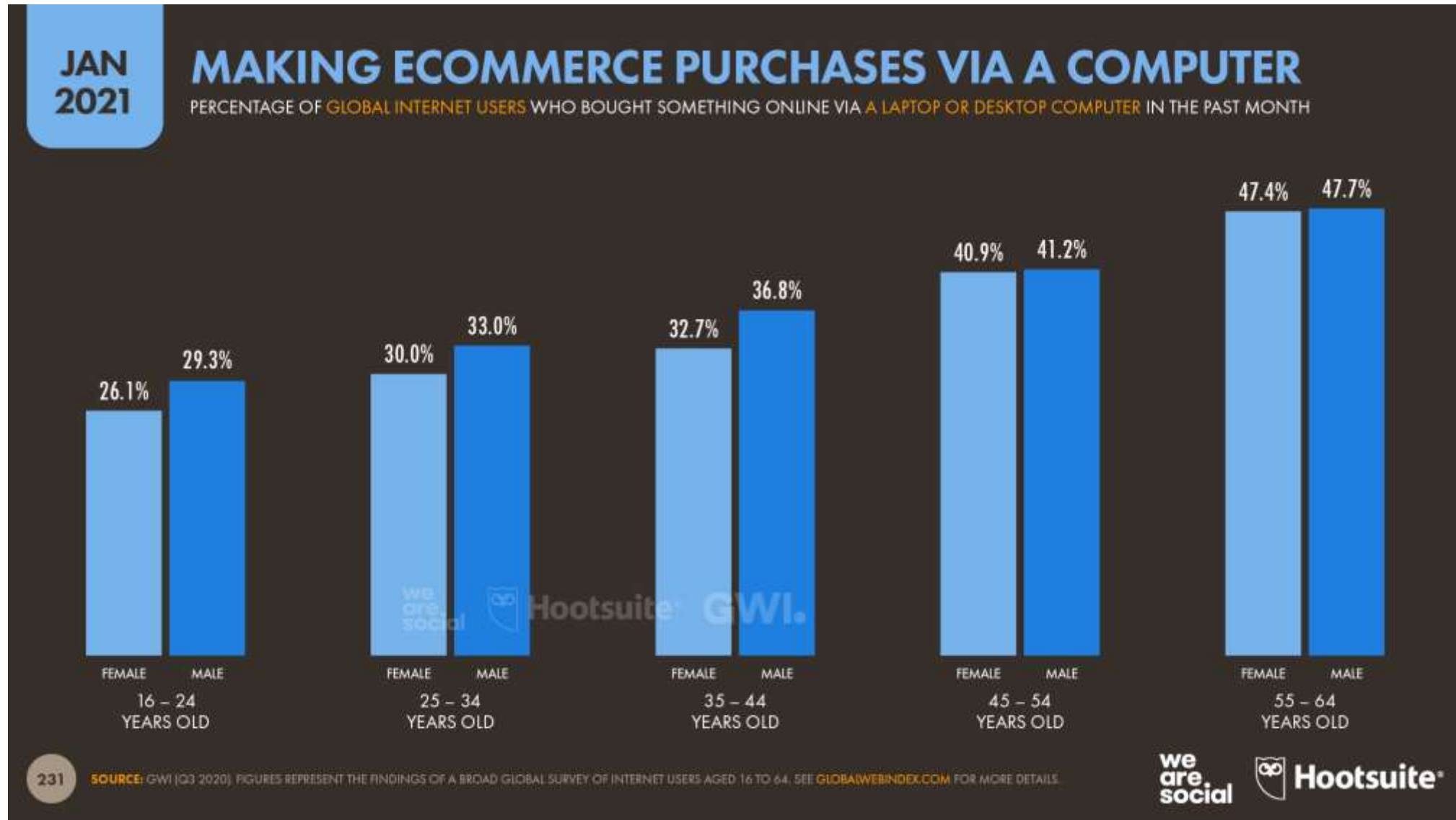
- Купоны на скидки на дорогие товары
- Программа Trade-in

Примечание?  
для экспертов вход по пропуску

# Приложение 1

<https://datareportal.com/reports/digital-2021-global-overview-report>

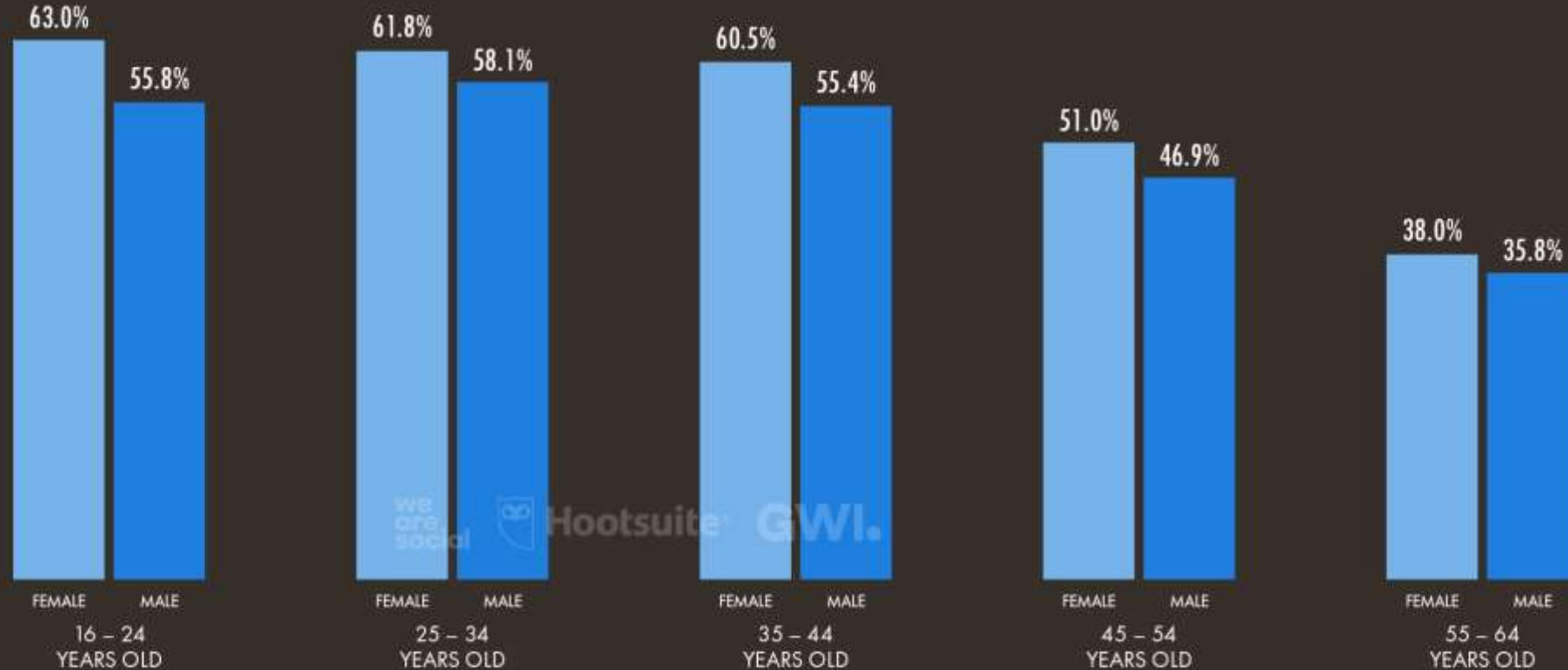
Исследование всемирной сети.



JAN  
2021

## MOBILE ECOMMERCE ADOPTION BY AGE AND GENDER

PERCENTAGE OF GLOBAL INTERNET USERS WHO BOUGHT SOMETHING ONLINE VIA A MOBILE DEVICE IN THE PAST MONTH



# Приложение 2

<https://finexpertiza.ru/press-service/researches/2022/vozh-zarpl-maksimum/>

Исследование зарплат граждан России

Меньше всего получают несовершеннолетние работники в возрасте до 18 лет - 25,1 тыс. руб. после вычета НДФЛ, а также молодые люди 18-19 лет - 33,4 тыс. руб. В студенческом возрасте 20-24 лет заработки увеличиваются на 26,4%, до 42,2 тыс. руб., к 25-29 годам - еще на 30%, до 54,9 тыс. руб., в 30-34 года - на 4,7%, до 57,5 тыс. руб., достигая своего пика. В 35-39 лет зарплаты держатся примерно на этом же уровне, опускаясь лишь на 1,7% - до 56,5 тыс. руб., в 40-44 года - уже на 4,2%, до 54,2 тыс. руб., в 45-49 лет - на 5,2%, до 51,3 тыс. руб.

# Приложение 3

| Переменная  | Объяснение   |
|-------------|--|
| device_type | тип устройства (mobileapp_android — мобильное приложение (Android), mobileapp_ios — мобильное приложение (IOS), desktop — десктопная версия сайта) |