



Подписываться или нет?

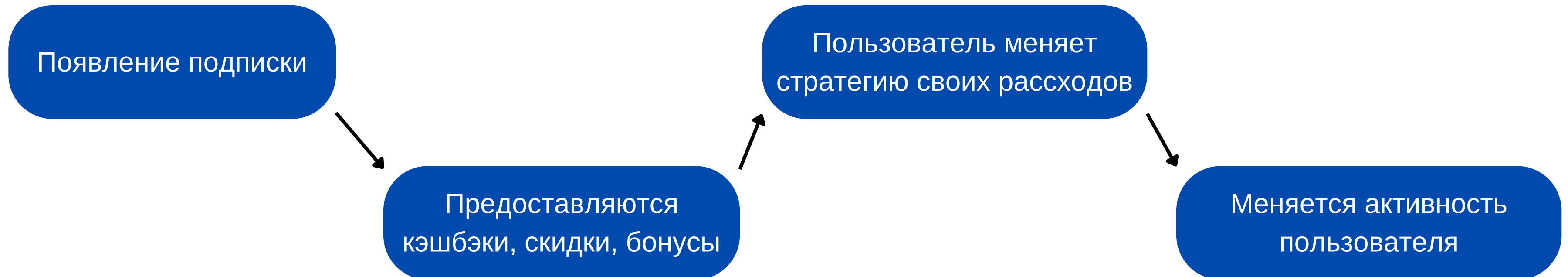
Исследовательский проект

Карташов Игорь
Смирнов Александр
Иванов Герман
Тукан Фёдор

Теоретический блок

Исследовательский вопрос: зависит ли активность и поведение пользователя от изменения статуса подписки?

Гипотеза: изменение статуса подписки стимулирует пользователя к изменениям в его поведении, активности покупок.



Теоретический блок

Исследуемые показатели

| Объясняемая переменная | В чём измеряется | Обозначение |
|-----------------------------------|-------------------|-------------|
| Стоимость покупки, кол-во покупок | Числовые значения | price |

| Объясняющая переменная | В чём измеряется | Обозначение |
|---|-----------------------|--------------------------------|
| Подписки | Категориальные данные | bundle |
| Категории товаров и услуг: путешествия, кино, спорт, топливо... | Бинарные данные | travel, cinema, sport, fuel... |

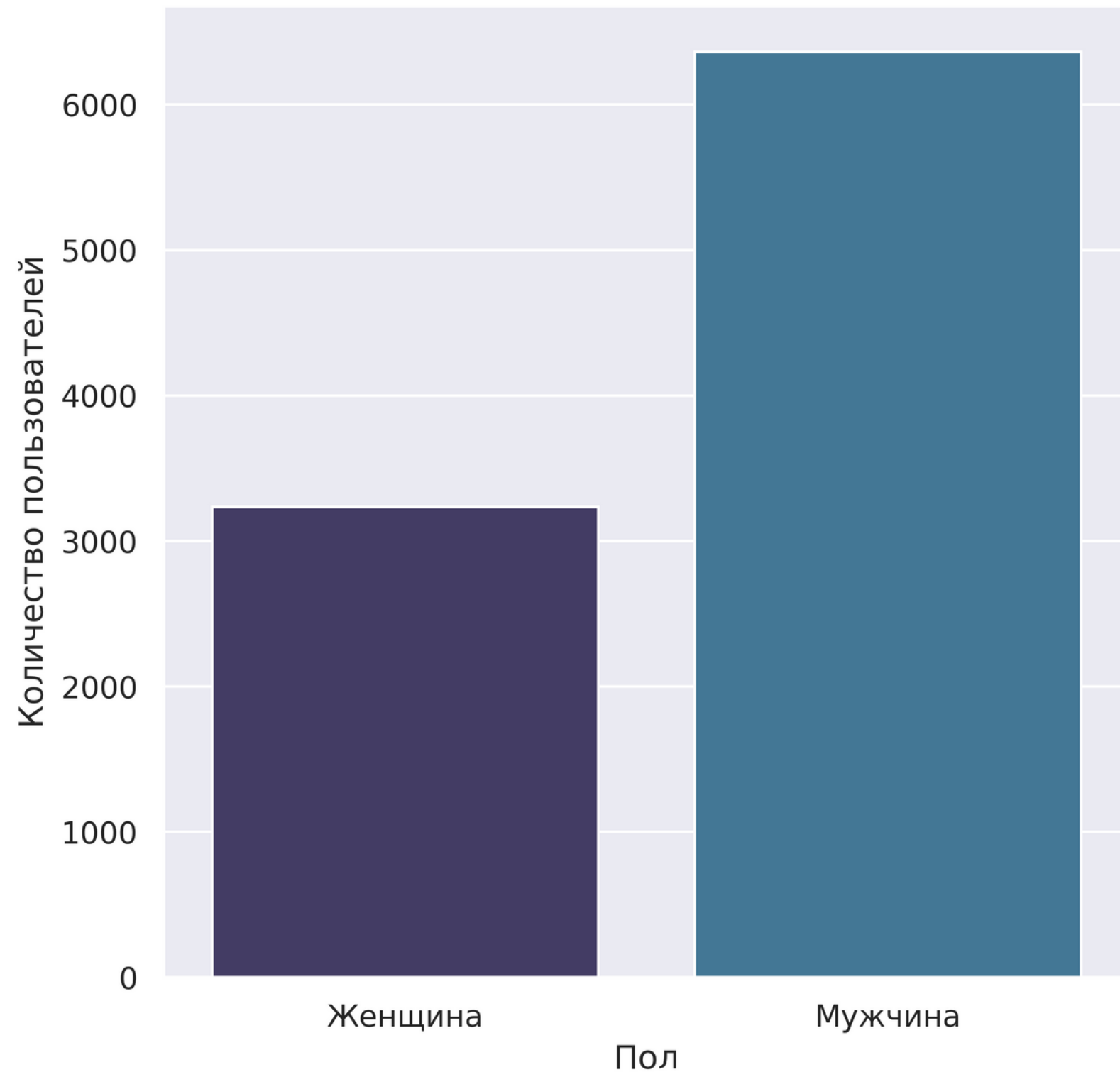
Предварительный анализ

Предобработка данных

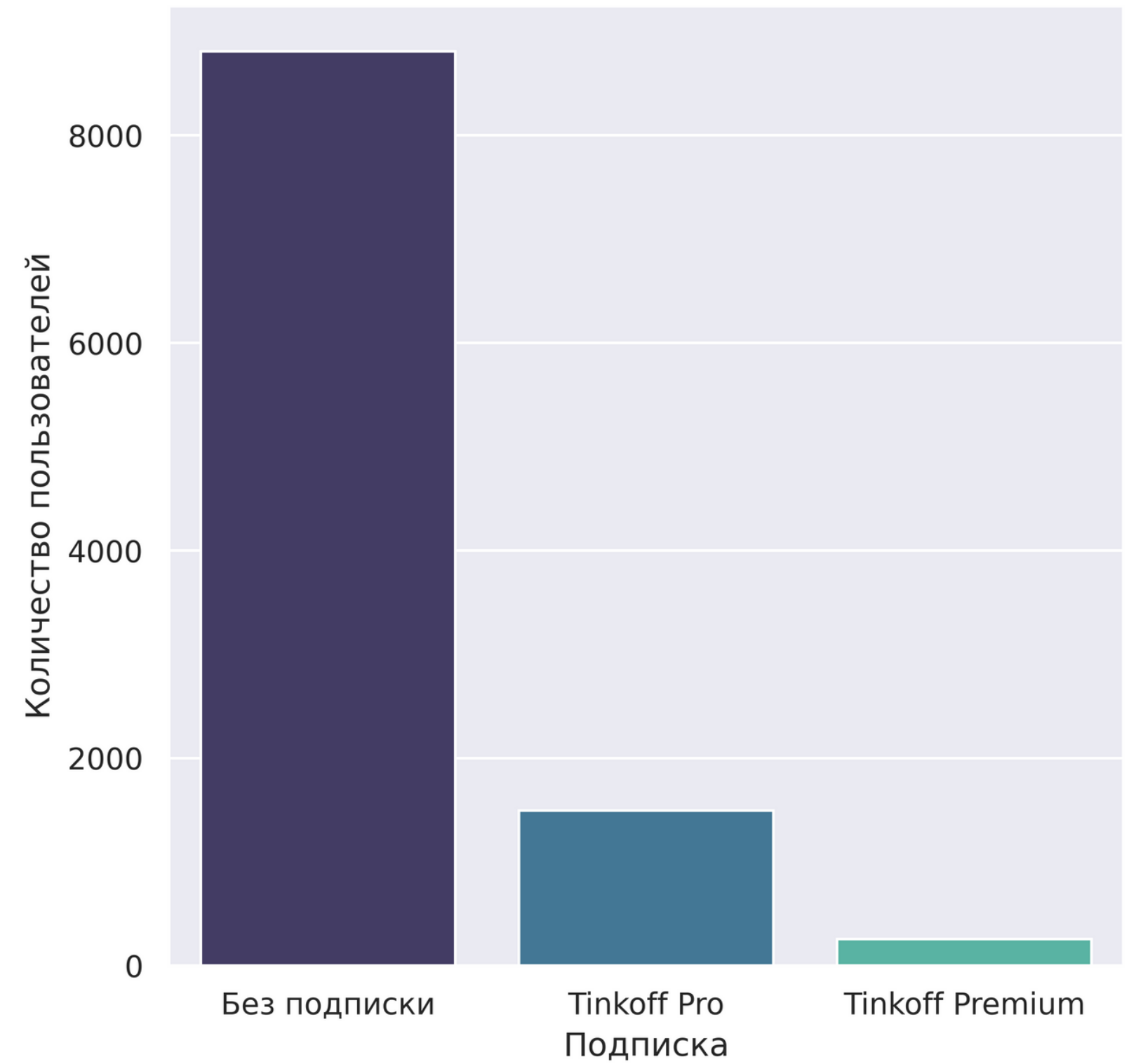
- 1) Исключаем строки с пропусками, несуществующими услугами и бронированием.
- 2) Исключаем строки пользователя с id:82385:
 - 1084 покупки (0 бронирований)
 - потратил более 7 млн на travel
 - Premium подписка



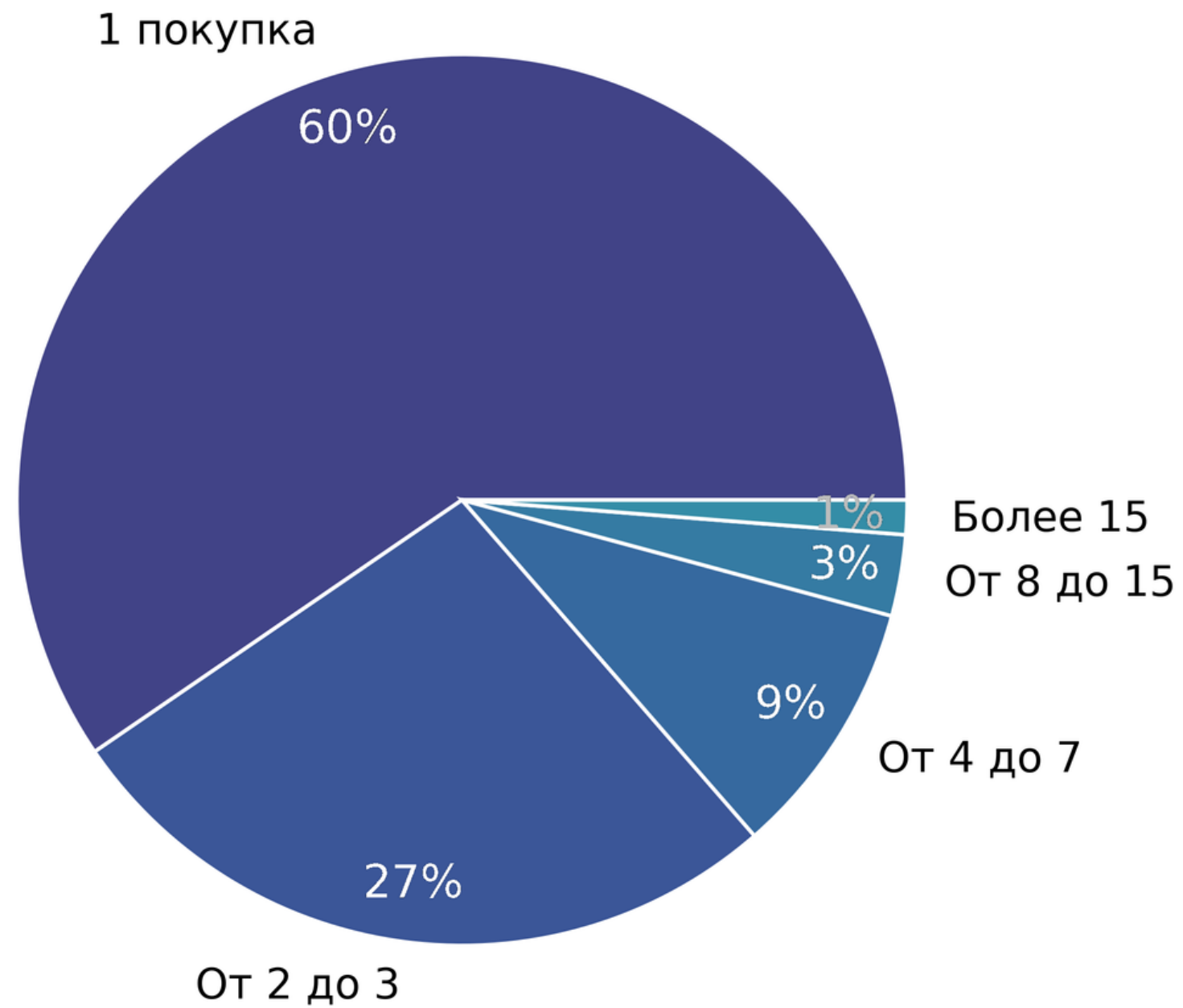
Распределение по полу



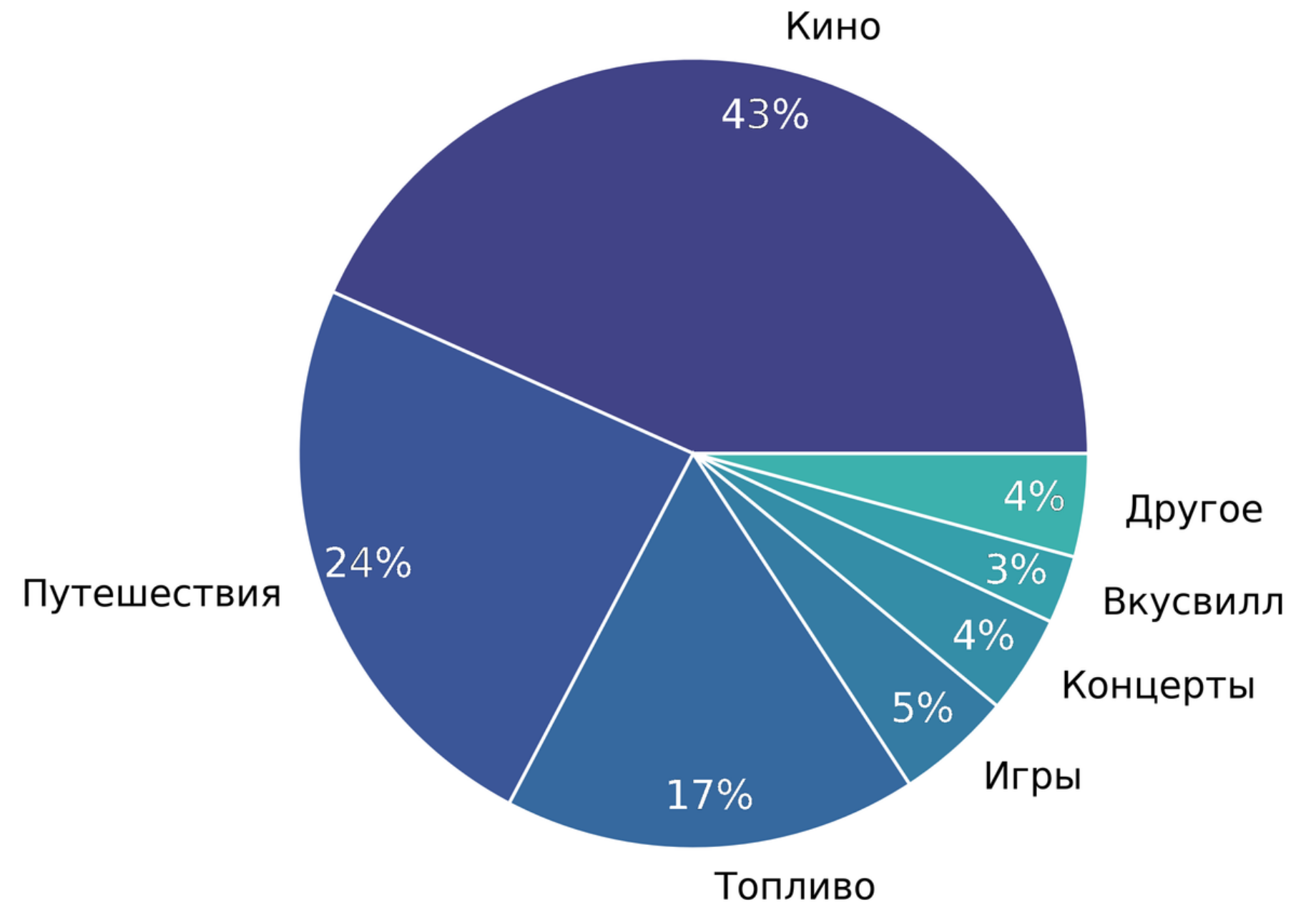
Популярность подписки



Распределение по количеству покупок



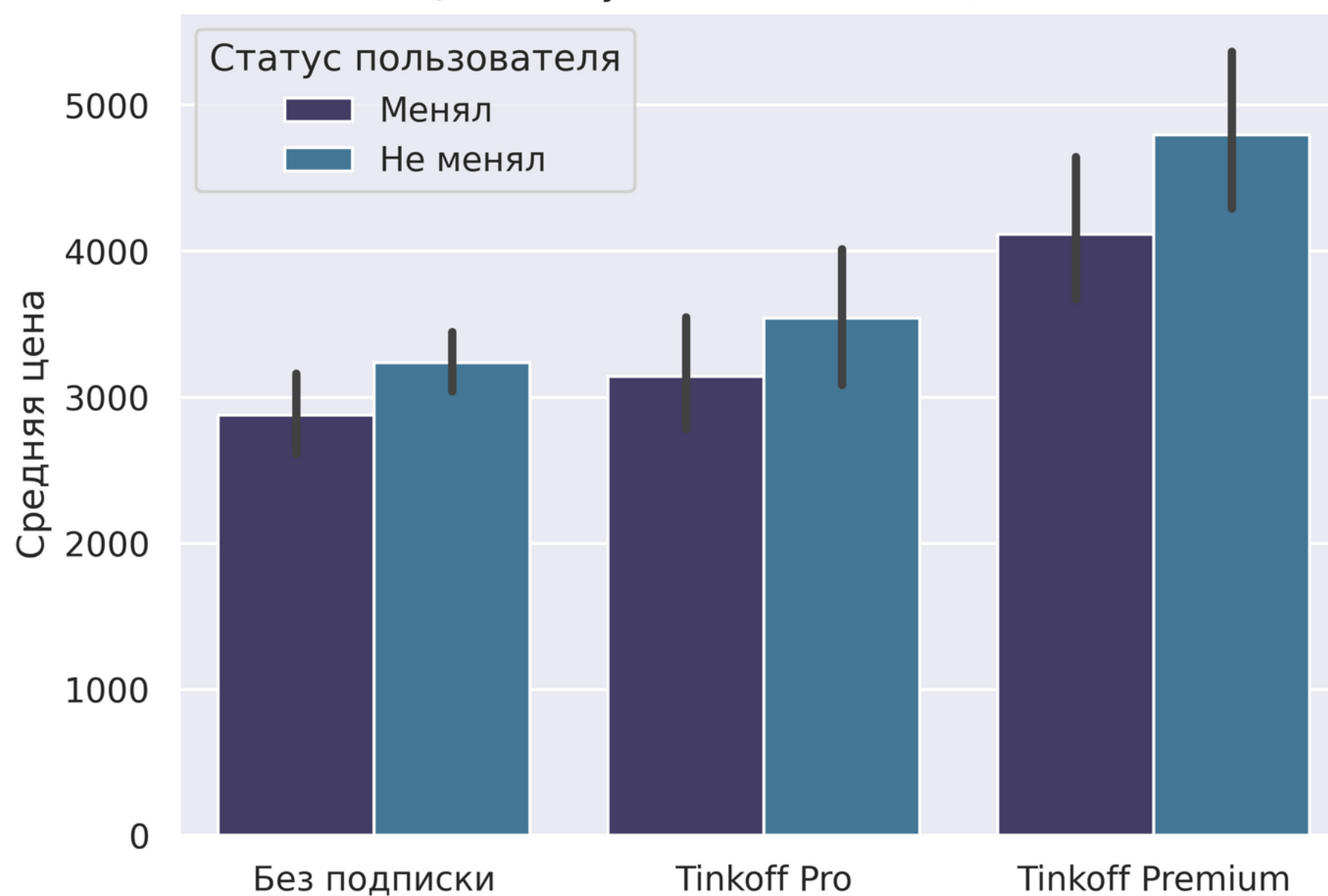
Распределение по категориям



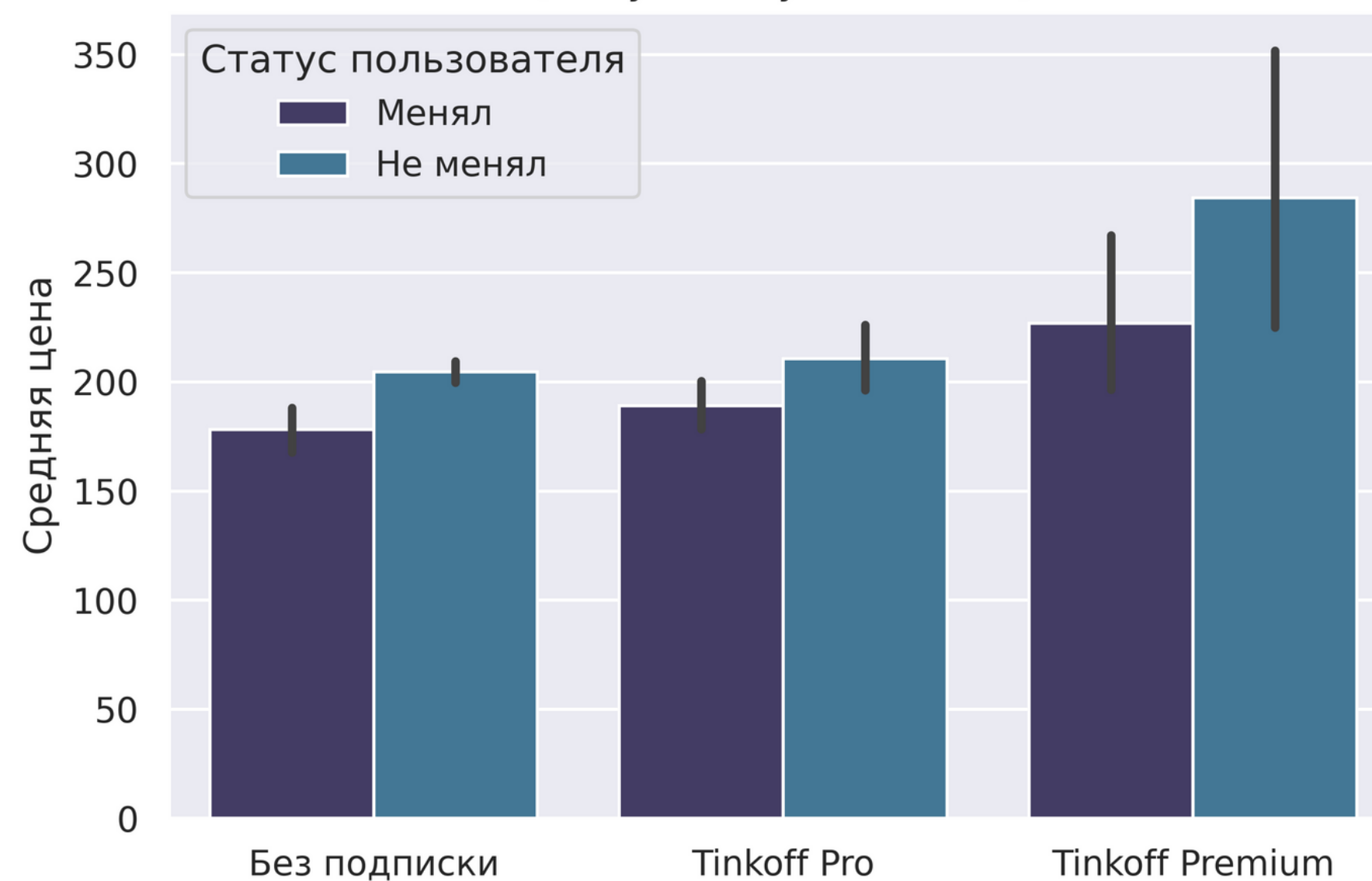
Математическое подтверждение

Динамика средней стоимости покупок

Цена за путешествия по подпискам



Цена по подпискам (без учета путешествий)



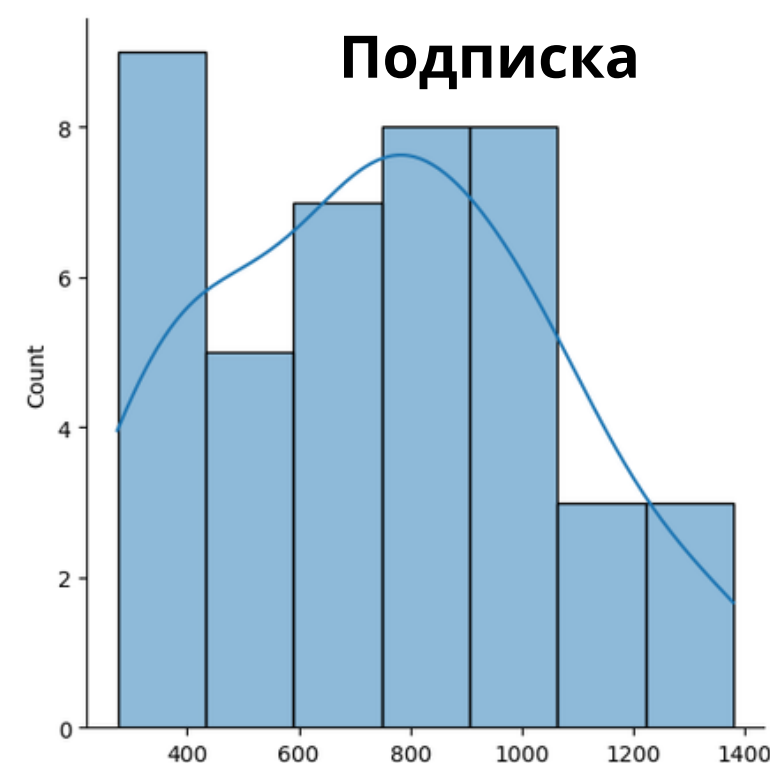
Математическое подтверждение

test Шапиро-Уилко

=> в выборках отсутствует нормальное распределение

U-test Манна-Уитни

Рассматриваем на примере динамического поведения пользователей



H_0 : Средний чек пользователей подписавшихся и отписавшихся одинаков

H_1 : Средний чек подписавшихся и отписавшихся различается

| Поведение | Размер выборки | Средний чек |
|---------------|----------------|-------------|
| Подписавшиеся | 217 | 278.67 |
| Отписавшиеся | 299 | 255.72 |

$p\text{-value}=0.65 > 0.05 \longrightarrow H_0$ подтверждается

Математическое подтверждение

test Шапиро-Уилко

=> в выборках отсутствует нормальное распределение

U-test Манна-Уитни

Рассматриваем на примере динамического поведения пользователей



H_0 : Средний чек пользователей с категорией путешествия одинаков

H_1 : Средний чек различается

| Поведение | Размер выборки | Средний чек |
|---------------|----------------|-------------|
| Подписавшиеся | 213 | 2597.52 |
| Отписавшиеся | 512 | 520.65 |

$p\text{-value}=0.22 > 0.05 \longrightarrow H_0$ подтверждается

Без изменений статуса подписки (статические пользователи)

| | Кино | Путешествия | Топливо | Игры | Концерты | Вкусвилл |
|-----------------|--------|-------------|---------|-------|----------|----------|
| Без подписки | 47.05% | 17.56% | 17.10% | 5.77% | 4.60% | 3.19% |
| Tinkoff Pro | 46.69% | 20.04% | 13.54% | 4.93% | 6.26% | 4.09% |
| Tinkoff Premium | 17.36% | 70.50% | 3.77% | 1.26% | 3.35% | 1.67% |

Подписавшиеся пользователи

| | Кино | Путешествия | Топливо | Игры | Концерты | Вкусвилл |
|-----------------|--------|-------------|---------|-------|----------|----------|
| Без подписки | 42.90% | 24.45% | 27.19% | 0.41% | 1.37% | 1.23% |
| Tinkoff Pro | 42.77% | 27.75% | 21.10% | 2.02% | 1.88% | 3.18% |
| Tinkoff Premium | 28.08% | 50.00% | 18.49% | 0.68% | 0.00% | 0.00% |

Отписавшиеся пользователи

| | Кино | Путешествия | Топливо | Игры | Концерты | Вкусвилл |
|-----------------|--------|-------------|---------|-------|----------|----------|
| Без подписки | 38.94% | 23.88% | 28.75% | 3.35% | 1.90% | 1.29% |
| Tinkoff Pro | 51.83% | 18.94% | 21.24% | 1.76% | 2.98% | 0.54% |
| Tinkoff Premium | 18.75% | 65.00% | 13.75% | 0.42% | 0.42% | 0.00% |

Вывод

1. Наличие подписки положительно влияет на средний чек пользователя
2. Изменение статуса подписки стимулирует клиентов пользоваться определенными видами товаров и услуг.
3. Средний чек пользователей подписавшихся и отписавшихся одинаков.

Ограничения

- “Топорное” деление динамической выборки на подписчиков и отписчиков из-за дефицита пользователей, менявших статус подписки, и ограниченности рассматриваемого таймлайна (всего год).
- Невозможно точно определить причины изменений в поведении пользователей из-за отсутствия более подробных данных о них (сколько денег на счету, условия по карте и т.д.).

Перспективы

1. Расширить выборку за счет использования других возрастов, провести аналогичный анализ для нее.
2. Добавить информацию за прошлые года за счет Tinkoff Лайфстайл (расширить таймлайн).
3. Добавить больше данных о пользователях (уточнить место, откуда производится покупка).

Применение

Исследуя поведение пользователей, банк может предлагать им конкретные виды товаров и услуг, дополнительный кэшбек на них и т.д.

Например:

- Tinkoff Premium популярна для путешествий, поэтому нужно предлагать более выгодные условия для этой категории.
- Статические пользователи без подписки и с Tinkoff Pro достаточно активно пользуются всеми видами услуг, для них стоит распределить кэшбэк на все категории.

Команда



Иванов Герман



Карташов Игорь



Смирнов Александр



Тукан Фёдор