

Ненужные вещи

Датасет содержит данные о событиях, совершенных в мобильном приложении "Ненужные вещи". В нем пользователи продают свои ненужные вещи, размещая их на доске объявлений. В датасете содержатся данные пользователей, впервые совершивших действия в приложении после 7 октября 2019 года.

Цель исследования- Ответить на вопросы заказчика:

- 1. Какие пользователи склонны часто возвращаться в мобильное приложение?**
- 2. Какие пользователи часто делают целевое событие (contacts_show)**
- 3. Как различается время между распространенными событиями пользователей, (например, различаются ли пользователи по группам на основе времени, которое проходит с момента первого использования мобильного приложения и до использования функции поиска)?**

Шипилов Роман Андреевич

Общий вывод

1. Какие пользователи склонны часто возвращаться в мобильное приложение? - Первая группа
2. Какие пользователи часто делают целевое событие (contacts_show) - Вторая группа
3. Как различается время между распространенными событиями пользователей, (например, различаются ли пользователи по группам на основе времени, которое проходит с момента первого использования мобильного приложения и до использования функции поиска)?

Пользователям из 2 группы относительно других групп меньше всего/не показывали рекомендованные объявления. Пользователи из 2 группы предпочитают листать фото карточек. Пользователи из группы 0 часто пользуются поиском, видимо не находят в поиске того, чего искали. Пользователи из группы 1 чаще проводят время в приложении, им часто показывают рекомендованные объявления.

Скорее всего дело в рекомендациях системы. Советую проверить на сколько актуальные объявления показывают пользователям.

Структура данных

```
1 print('total users:', mobile_dataset['user_id'].nunique())
```

```
total users: 4293
```

Общее количество пользователей в датафрейме - 4293

```
1 print('date start:', mobile_dataset['event_time'].min())
2 print('date end:', mobile_dataset['event_time'].max())
```

```
date start: 2019-10-07 00:00:00
```

```
date end: 2019-11-03 23:58:13
```

Данные в датафрейме начинаются с 2019-10-07 00:00:00 и заканчиваются 2019-11-03 23:58:13

```
1 df = mobile_dataset.merge(mobile_sources, on='user_id')
2
3 df.groupby('source')['user_id'].nunique()
```

```
source
```

```
google    1129
```

```
other     1230
```

```
yandex    1934
```

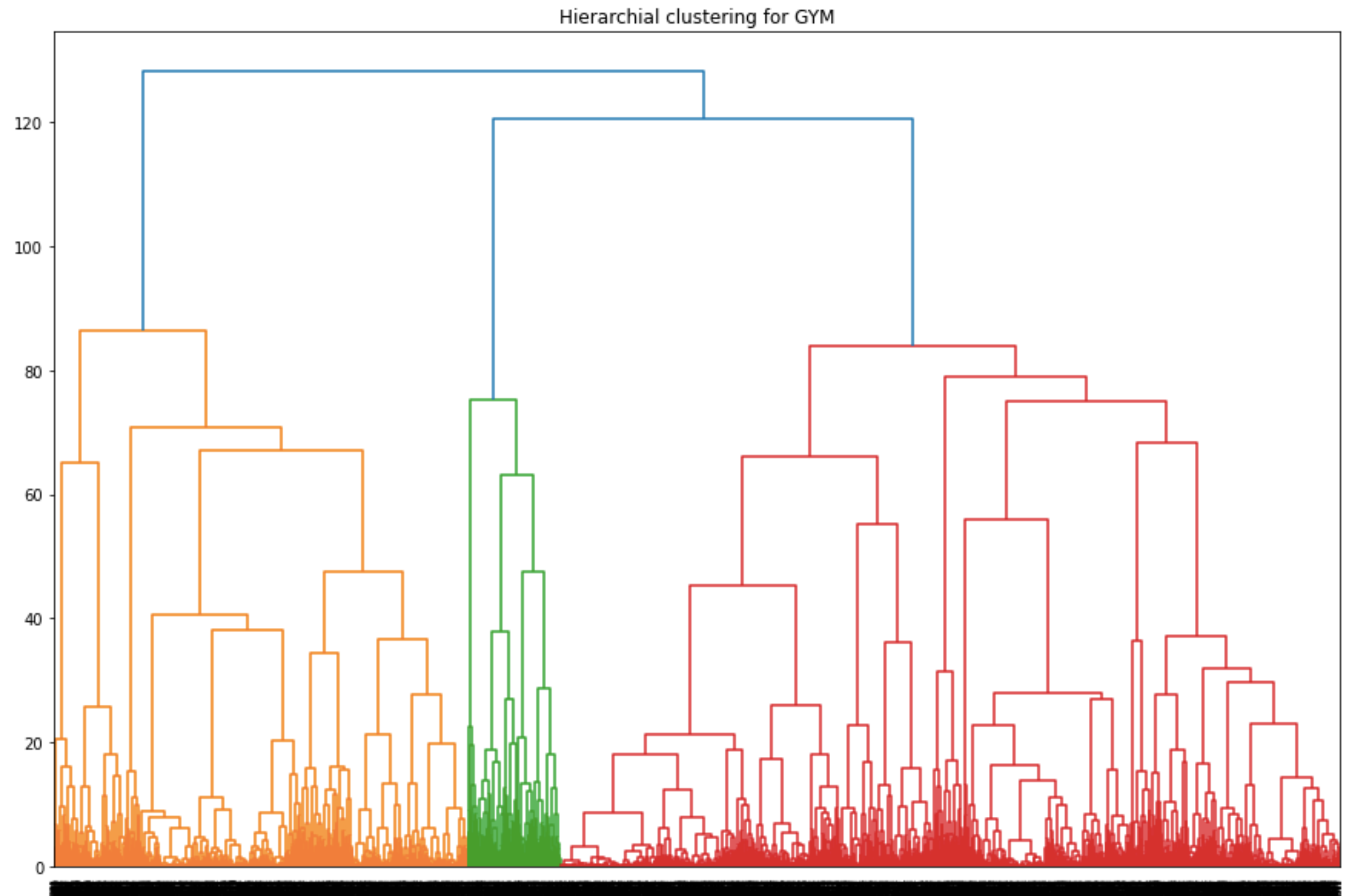
```
Name: user_id, dtype: int64
```

Так пользователи распределены по источникам

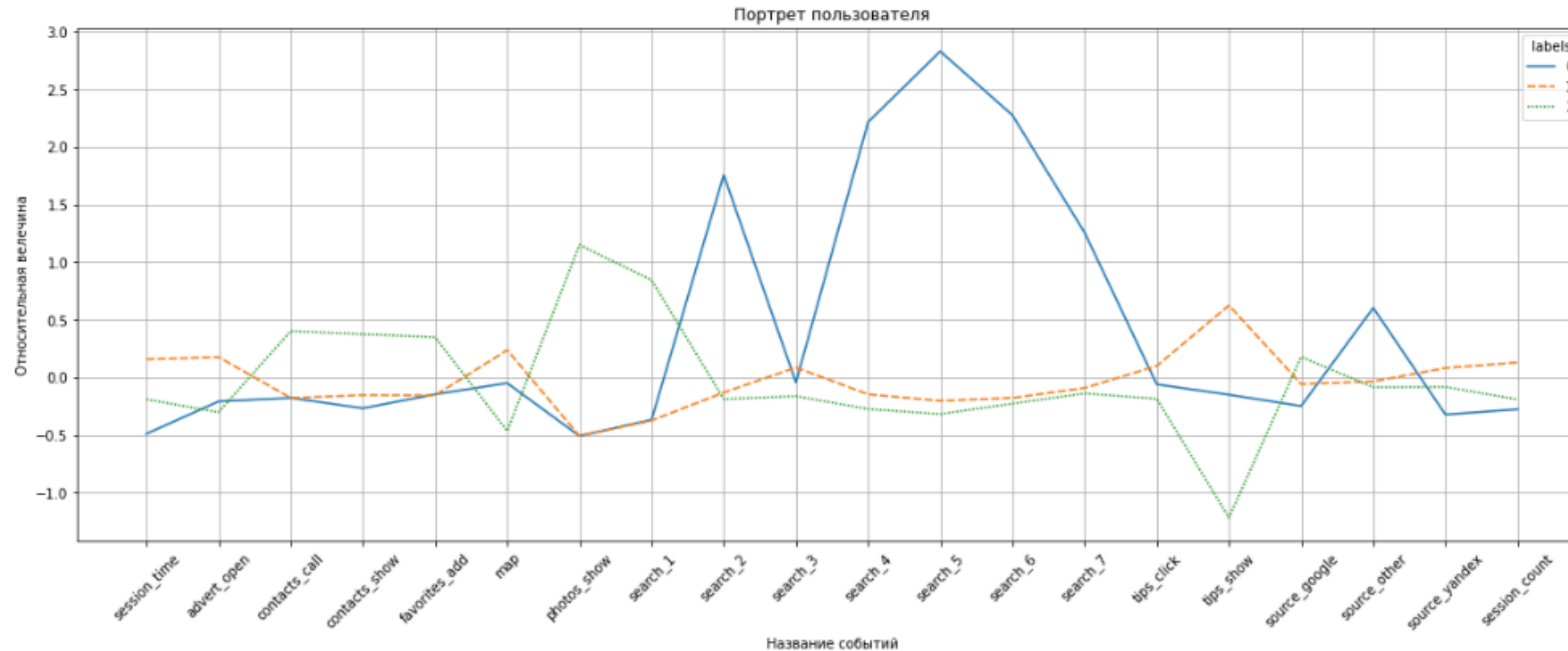
Кластеризация

Разделение пользователей на группы на основе их действий

- На основании действий пользователей, выделились 3 основные группы



Портрет пользователя

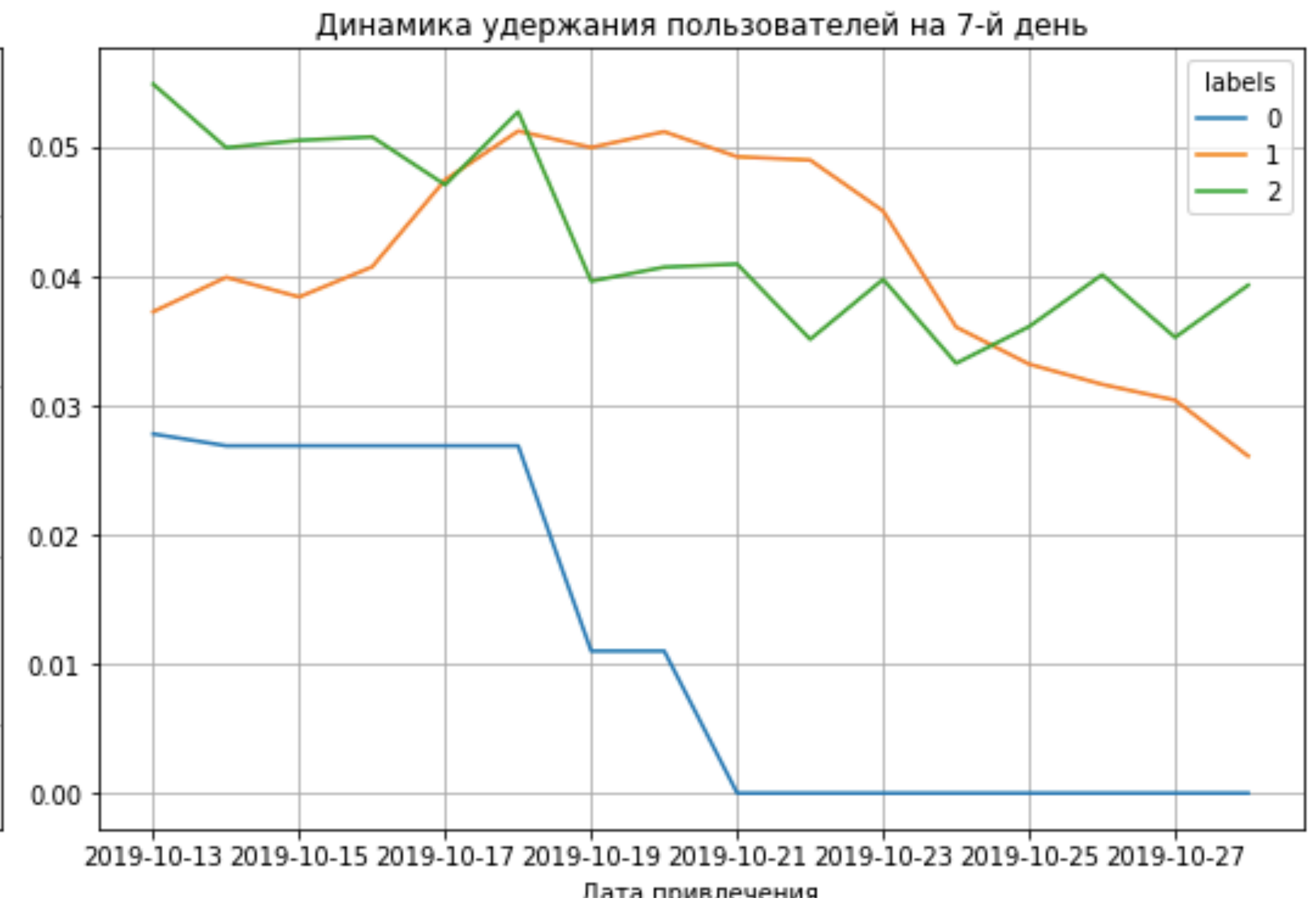
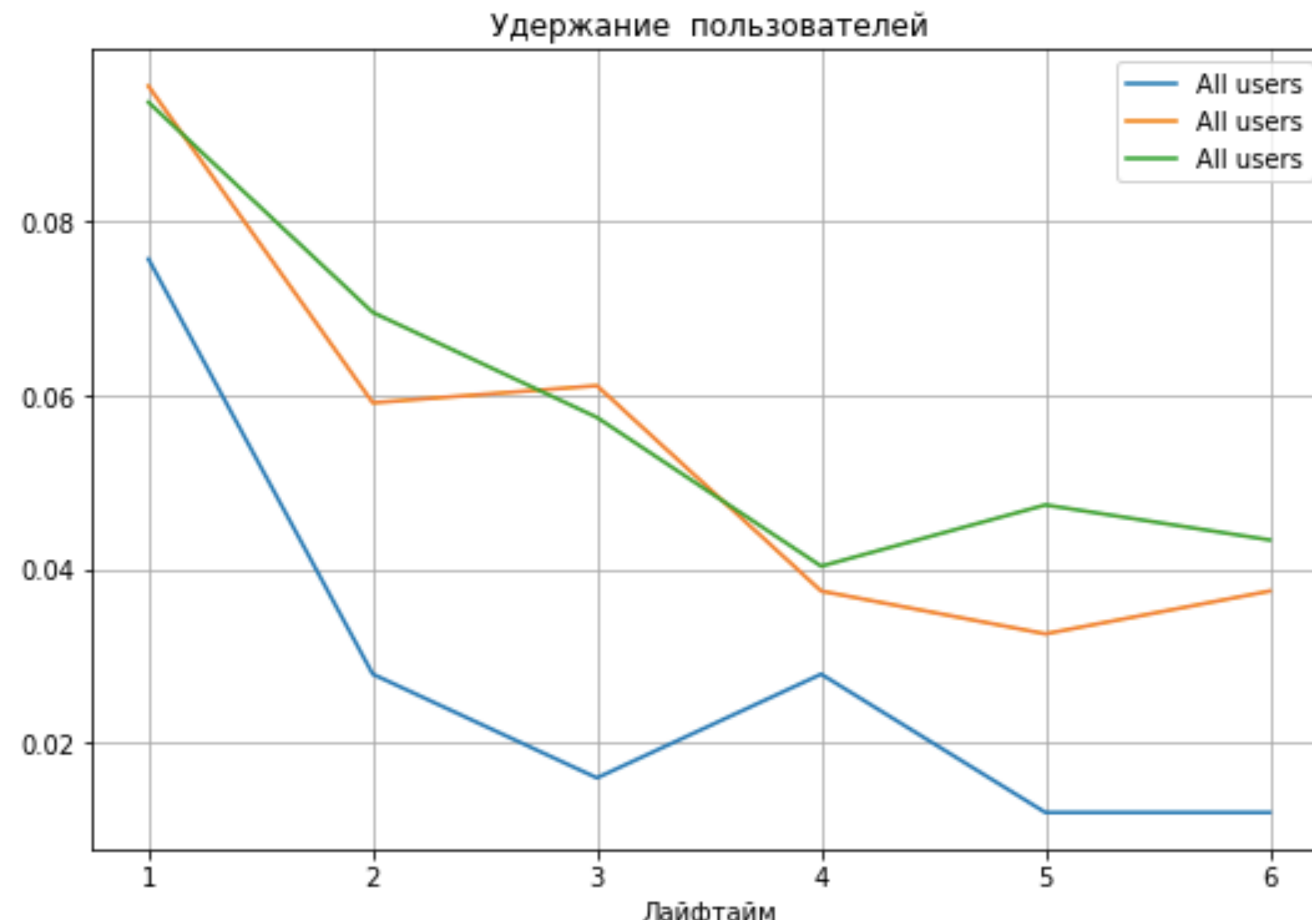


- Пользователям из 2 группы относительно других групп меньше всего/не показывали рекомендованные объявления.
- Пользователи из 2 группы предпочитают листать фото карточек.
- Пользователи из группы 0 часто пользуются поиском, видимо не находят в поиске того, чего искали.
- Пользователи из группы 1 чаще проводят время в приложении, им часто показывают рекомендованные объявления.

Удержание

Какие пользователи склонны часто возвращаться в мобильное приложение?

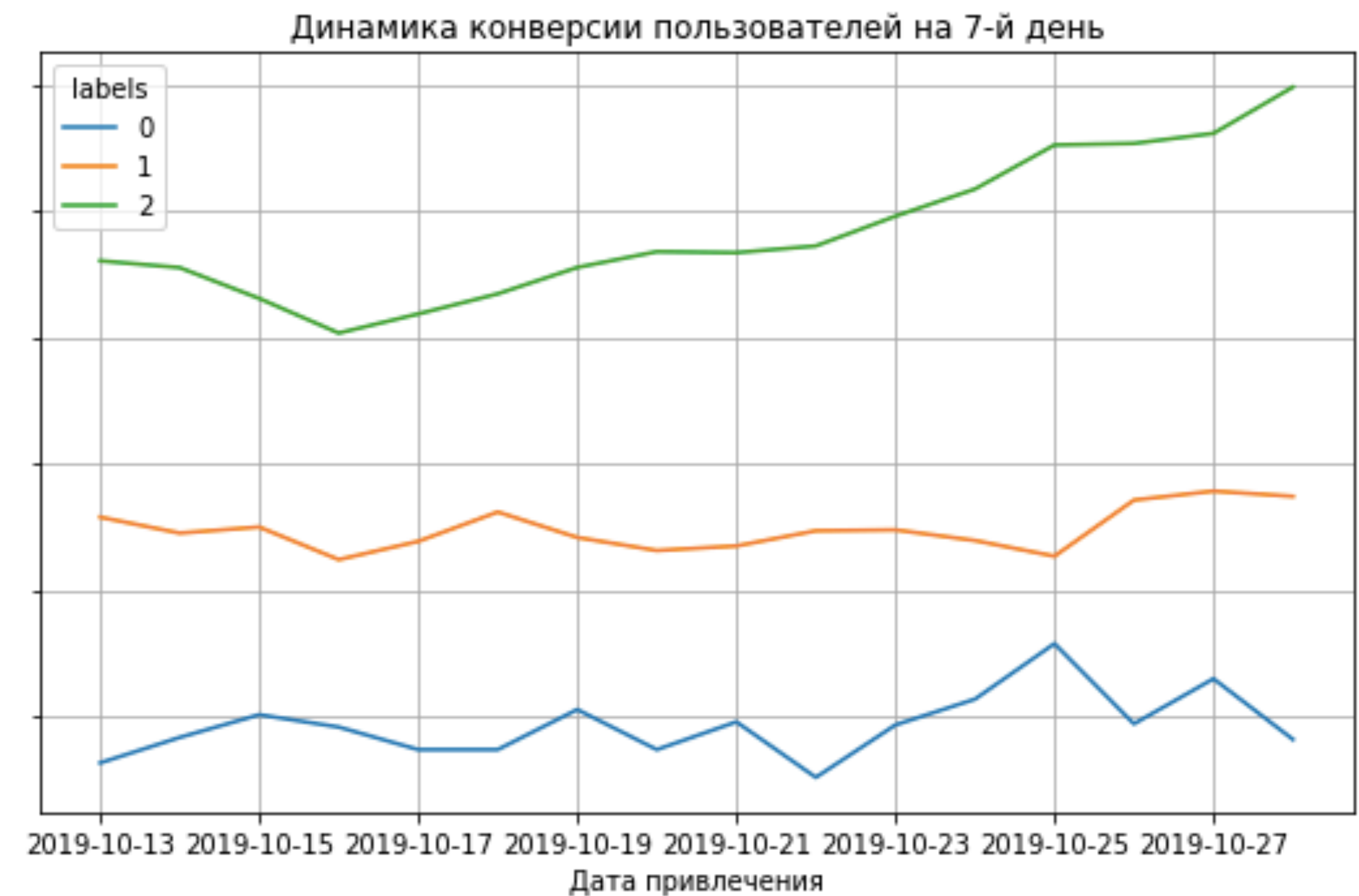
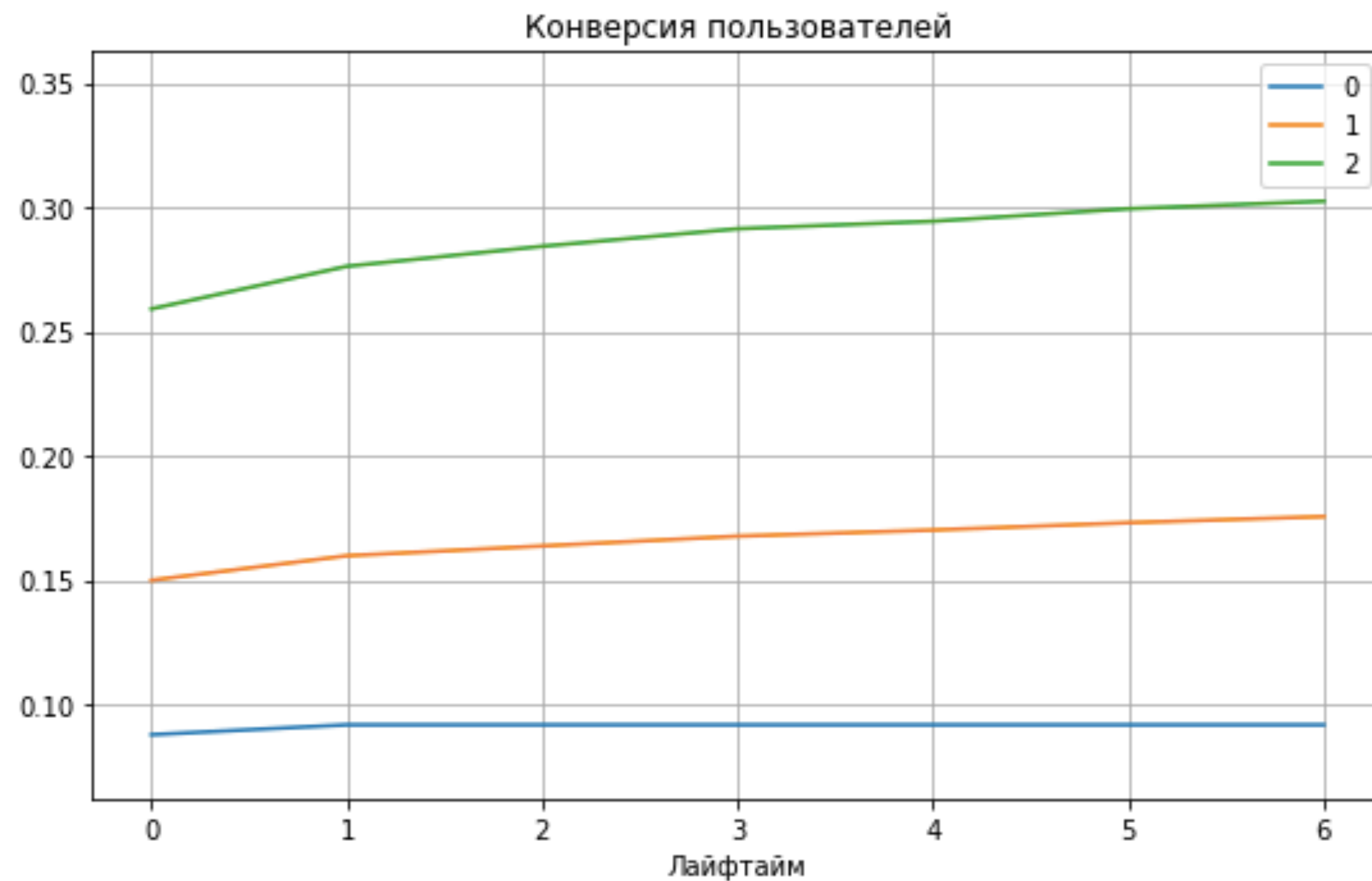
- Пользователи из 2 группы склонны возвращаться.



Конверсия

Какие пользователи часто делают целевое событие (contacts_show)

- Пользователи из 2 группы чаще делают целевое событие.



Время

Как различается время между распространенными событиями пользователей, (например, различаются ли пользователи по группам на основе времени, которое проходит с момента первого использования мобильного приложения и до использования функции поиска)?

- У группы 0 в среднем 25 секунд
- У группы 1 в среднем 311 секунд.
- У группы 2 в среднем 216 секунд.

**Пользователи из группы 0 быстрее всех начинают пользоваться поиском.
Пользователи из группы 1 дольше всех доходят до поиска.**

КОНЕЦ

Спасибо за внимание