**Цели домашнего задания:**

**1.** Научиться строить матрицы переходов и графов.

**2.** Научиться выделять кластеры на основании вовлечённости в разные функции или экраны приложения.

**3.** Научиться строить bar-charts вовлечённости для каждого кластера.

**Задание**

Вам представлены данные clickstream некоторого приложения (mobile-app-data.csv), которые показывают, как пользователи перемещаются между основными функциями приложения, включая onboarding (onb), редактор материалов (tool) и т .д.

**Что нужно сделать:**

**1.** Повторите обработку данных в Python, как показано в практической части видеоурока.

**2**. Постройте матрицы переходов и графы.

**3.** Выделите кластеры на основании вовлечённости в разные функции или экраны приложения.

**4.** Постройте bar-charts вовлечённости для каждого кластера.

**5.** Разберите полученные результаты. Какие выводы можно сделать? Вовлечение в какие функции приложения как влияет на конверсии в различных кластерах?

**Рекомендации по выполнению:**

**1.** Результатом работы должен быть ноутбук на Python с проведённым анализом — построенные матрицы переходов, графы, карты кластеров, профили вовлечения кластеров пользователей в топ-10 функций.

**2**. Ниже приведено описание групп и некоторых отдельных событий из датасета. Помните, что события можно группировать, удалять малозначимые, выделять значимые в новые группы.

**3.** Уделите особое внимание интерпретации результатов — что вы видите на полученных графиках, о чём это говорит. Опишите вашу интерпретацию прямо в ноутбуке возле результатов.

**Описание полей датасета:**

* **event\_timestamp** — временна́я метка в формате int, указывающая, когда пользователь совершил действие в приложении. Указывает не на конкретную дату и время, а на порядок событий относительно друг друга.
* **event\_name** — действия, которые пользователи совершают в приложении.
* **user\_pseudo\_id** — id пользователя.

**Описание событий**

Если во второй колонке указано конкретное событие, то в третьей описано именно оно. Если ячейка пустая, характеристика дана группе событий.

|  |  |  |
| --- | --- | --- |
| **Группа событий** | **Событие** | **Описание** |
| basket |  | Корзина |
| catalog |  | Действия в каталоге |
| chout | chout\_link | Пользователь совершает покупку через ссылку |
| edit |  | Пользователь редактирует фотографии |
| first | first\_open | Пользователь впервые запускает приложение после установки |
| login | login | Пользователь входит в аккаунт |
| logout | logout | Пользователь выходит из своего аккаунта |
| onb |  | Онбординг (приветственные экраны) |
| os |  | Системные логи |
| phs |  | Работа с фотографиями |
| preview |  | Действия в режиме предпросмотра |
| reg |  | Регистрация |
| remove | remove | Пользователь удаляет все данные приложения, включая настройки и данные для входа. Только Android |
| screen | screen\_view | Выполняется переход на другой экран |
| sys |  | Системные логи |
| tool |  | Действия с редактором материалов |
| upload |  | События, связанные с загрузкой |
| user | user\_engagement | Периодически регистрируется, если ресурс (приложение или веб-страница) находится в активном режиме |

**Критерии оценки**

**«Зачет»** — сделанный Python пайплайн обработки данных по аналогии с классной работой, есть анализ полученных результатов.  
**«На доработку»** — есть обработка, но нет анализа и интерпретации полученных результатов.