#### ТЕХНИЧЕСКАЯ ДОКУМЕНТАЦИЯ К ПРОЕКТУ:

«Анализ пользовательского взаимодействия с карточками статей в Яндекс.Дзен»

ОГЛАВЛЕНИЕ

1. Вводные данные *стр. 2*
2. Техническое задание *стр. 3*
3. Макет дашборда *стр. 4*

#### ССЫЛКА НА ДАШБОРД (Tableau):

[Анализ пользовательского взаимодействия с карточками статей Яндекс.Дзен](https://public.tableau.com/views/vbout_YaPP10_dash/__1?:language=en-US&:display_count=n&:origin=viz_share_link)

#### ВВОДНЫЕ ДАННЫЕ

**Анализ пользовательского взаимодействия с карточками статей в**[**Яндекс.Дзен**](https://zen.yandex.ru/)

Каждую карточку определяют её **тема** и **источник** (у него тоже есть тема). Примеры тем: «Красота и здоровье», «Россия», «Путешествия».

Пользователей системы характеризует **возрастная категория**. Скажем, «26-30» или «45+».

Есть **три способа взаимодействия** пользователей с системой:

* Карточка отображена для пользователя (show);
* Пользователь кликнул на карточку (click);
* Пользователь просмотрел статью карточки (view).

**Регулярные вопросы:**

* Сколько взаимодействий пользователей с карточками происходит в системе с разбивкой по темам карточек?
* Как много карточек генерируют источники с разными темами?
* Как соотносятся темы карточек и темы источников?

**Дашборд** должен основываться на **пайплайне**, который:

* берет данные из таблицы, в которой хранятся сырые данные
* трансформирует данные
* укладывает их в агрегирующую таблицу

**Пайплайн** будет разработан вас **дата-инженерами**.

#### ТЕХНИЧЕСКОЕ ЗАДАНИЕ

|  |  |
| --- | --- |
| **Бизнес-задача:** | анализ взаимодействия пользователей с карточками Яндекс.Дзен |
| **Частота использования дашборда:** | 1 раз в неделю |
| **Основные пользователи:** | менеджеры по анализу контента |
| **Состав данных для дашборда:** | * История событий по темам карточек (два графика - абсолютные числа и процентное соотношение) * Разбивка событий по темам источников * Таблица соответствия тем источников темам карточек |
| **Параметры группировки данных:** | * Дата и время * Тема карточки * Тема источника * Возрастная группа |
| **Характер данных:** | * История событий по темам карточек — абсолютные величины с разбивкой по минутам * Разбивка событий по темам источников — относительные величины (% событий) * Соответствия тем источников темам карточек - абсолютные величины |
| **Важность:** | все графики имеют равную важность |
| **Источники данных для дашборда:** | cырые данные о событиях взаимодействия пользователей с карточками (таблица log\_raw) |
| **База данных, в которой будут храниться агрегированные данные:** | дополнительные агрегированные таблицы в БД zen |
| **Частота обновления данных:** | один раз в сутки, в полночь по UTC |

#### МАКЕТ ДАШБОРДА

