

Дашборд для анализа взаимодействия пользователей с карточками Яндекс Дзен

Подготовил Дерябин Глеб

Содержание

- Цель и задачи
- Техническое задание
- Исходные данные
- Дашборд
- Графики
- Ответы на вопросы
- Выводы

Цель и задачи

Цель: создать дашборд, при помощи которого отдел маркетинга сможет анализировать взаимодействие пользователей с карточками

Задачи:

1. Совместно разработать ТЗ и макет дашборда
2. Подключиться к базе zen и забрать агрегированную таблицу dash_visits, собранную дата-инженерами
3. Разработать дашборд в Tableau согласно макету

Техническое задание

Бизнес-задача: анализ взаимодействия пользователей с карточками Яндекс.Дзен;

Насколько часто предполагается пользоваться дашбордом: не реже, чем раз в неделю;

Кто будет основным пользователем дашборда: 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;

По каким параметрам данные должны группироваться:

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;

Источники данных для дашборда: дата-инженеры обещали подготовить для вас агрегирующую таблицу dash_visits. Вот её структура:

Таблица хранится в специально подготовленной для вас базе данных zen;

Частота обновления данных: один раз в сутки, в полночь по UTC;

Макет дашборда

Название дашборда		Фильтр по дате и времени	
Описание дашборда		Фильтр по темам карточек	Фильтр по возрастным категориям
График истории взаимодействия «События по темам карточек» (абсолютные значения, stacked area chart)	График истории взаимодействия «% событий по темам карточек» (% от общего, stacked area chart)	График разбивки событий по темам источников «События по темам источников» (относительные значения, pie chart)	
Таблица соответствия тем карточек темам источников «Темы источников - темы карточек». В колонках должны быть указаны темы источников, в строках – темы карточек. На пересечении должно быть указано абсолютное количество событий. Ячейки таблицы должны иметь окраску в зависимости от числа взаимодействий (highlight table).			

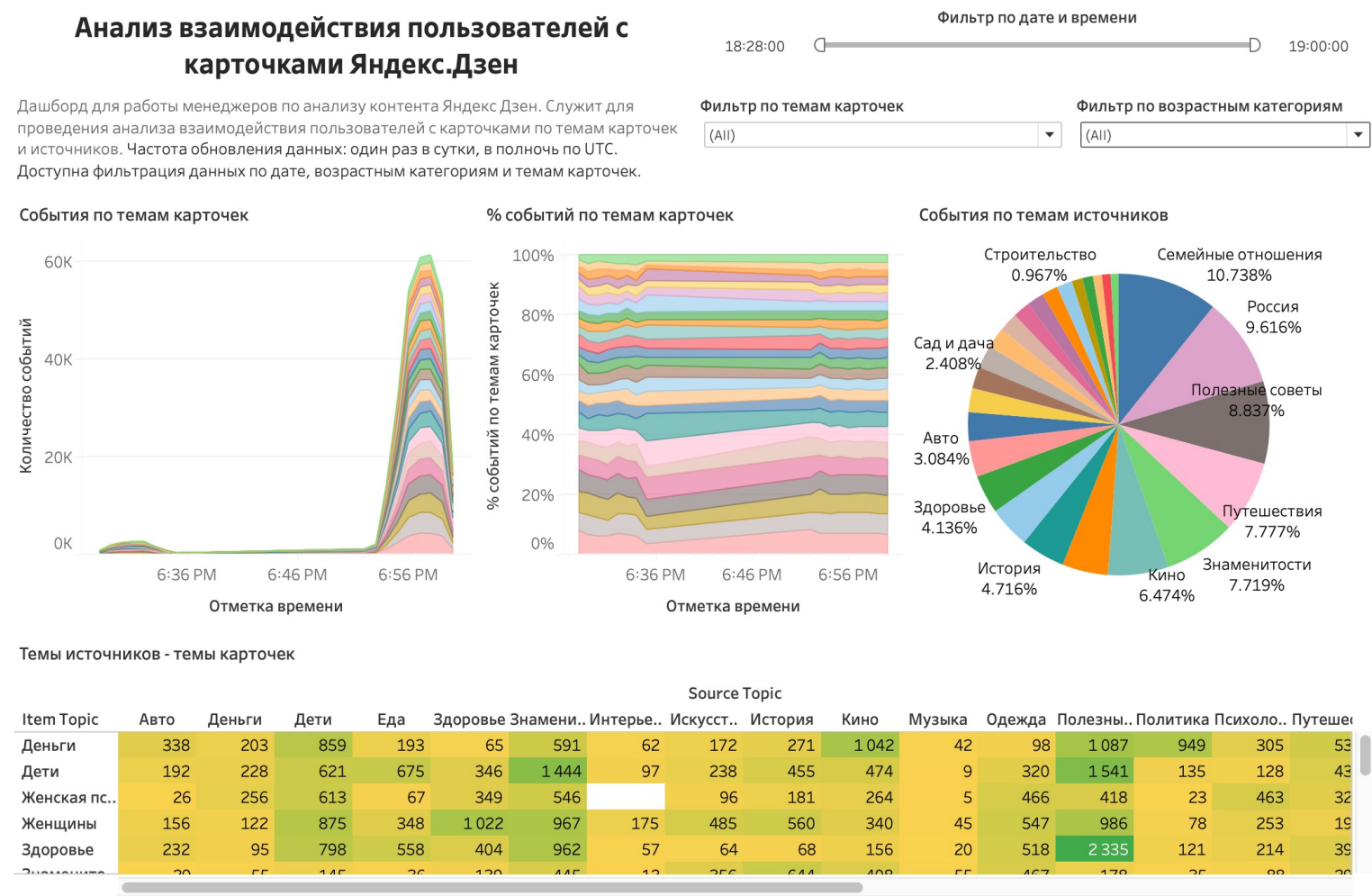
Исходные данные

Используя данные о взаимодействиях пользователей Яндекс Дзен с карточками, дата-инженеры создали агрегированную таблицу. Для пробного использования в ней собраны данные за 24 сентября 2019 года с 18:30 до 19:00

record_id	item_topic	source_topic	age_segment	dt	visits
1040597	Деньги	Авто	18-25	2019-09-24 18:32:00	3
1040598	Деньги	Авто	18-25	2019-09-24 18:35:00	1
1040599	Деньги	Авто	18-25	2019-09-24 18:54:00	4
1040600	Деньги	Авто	18-25	2019-09-24 18:55:00	17
1040601	Деньги	Авто	18-25	2019-09-24 18:56:00	27

Дашборд

Дашборд расположен [по ссылке](#)



События по темам карточек

События по темам карточек

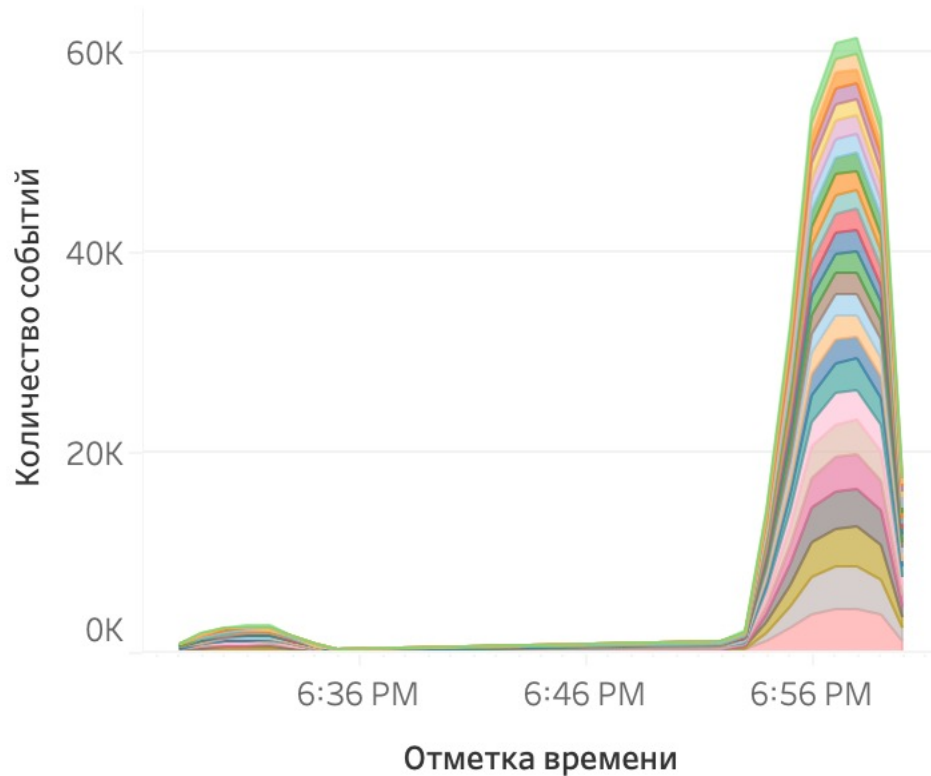


График истории взаимодействия «События по темам карточек» в виде зон с накоплением по абсолютным значениям (кол-во взаимодействий)

Вывод: пик взаимодействий с карточками 24 сентября пришелся на период 18:50-19:00. Суммарное кол-во событий достигло 60 тыс.

% событий по темам карточек

% событий по темам карточек

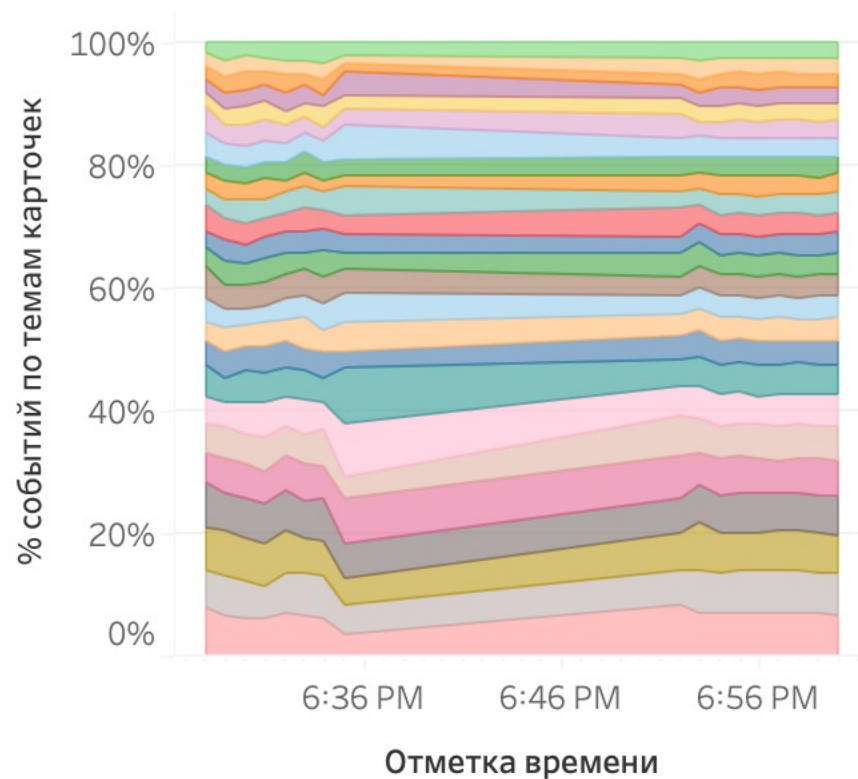


График истории взаимодействия «% событий по темам карточек» в виде зон с накоплением по относительным значениям (кол-во взаимодействий)

Вывод: в 18:36 замечен спад по темам карточек наука, отношения, факты, общество и другие, зато есть рост по темам история, скандалы и психология. После картина меняется на обратную.

События по темам источников

События по темам источников



График разбивки событий по темам источников в виде круговой диаграммы по относительным значениям

Вывод: исходя из долей можно определить как самые популярные темы источников: семейные отношения, Россия, советы, путешествия и знаменитости, так и самые непопулярные: финансы, музыка, строительство, технологии, деньги. Возможно, это слишком »сложные темы» для вечернего чтения.

Темы источников - темы карточек

Таблица соответствия тем карточек темам источников. В пересечениях столбцов и строк указаны абсолютные значения количества взаимодействий. Ячейки таблицы имеют окраску в зависимости от значений.

Темы источников - темы карточек

Item Topic	Source Topic															
	Авто	Деньги	Дети	Еда	Здоровье	Знамени..	Интерье..	Искусст..	История	Кино	Музыка	Одежда	Полезны..	Политика	Психоло..	Путешес
Деньги	338	203	859	193	65	591	62	172	271	1 042	42	98	1 087	949	305	53
Дети	192	228	621	675	346	1 444	97	238	455	474	9	320	1 541	135	128	43
Женская пс..	26	256	613	67	349	546		96	181	264	5	466	418	23	463	32
Женщины	156	122	875	348	1 022	967	175	485	560	340	45	547	986	78	253	19
Здоровье	232	95	798	558	404	962	57	64	68	156	20	518	2 335	121	214	39
Знамениты	20	55	145	26	120	445	12	256	644	408	55	467	178	25	88	28

Ответы на вопросы

- Сколько взаимодействий пользователей с карточками происходит в системе с разбивкой по темам карточек?

Ответить на вопрос можно при помощи графика «События по темам карточек». Например, в 18:58 с карточками по теме наука взаимодействовали 4372 раза.

- Как много карточек генерируют источники с разными темами?

Больше всего взаимодействий генерируют карточки с источником «Семейные отношения» - 10.7% от общего кол-ва.

- Как соотносятся темы карточек и темы источников?

Темы карточек связаны с темами источников по контексту. Например, карточки по теме Деньги генерируют источники по темам Россия, полезные советы, кино, политика, дети

Выводы

В результате работы над проектом создан интерактивный дашборд, при помощи которого можно исследовать взаимодействие пользователей с карточками Яндекс Дзен в ретроспективе, в разрезах по возрасту пользователей, темам карточек и источникам.

Например, при общем взгляде на имеющиеся данные из дашборда видно:

- Кол-во взаимодействий с карточками с 18:53 до 19:00 было максимальным и составило 60 тыс.
- Больше всего взаимодействовали с карточками по теме «Наука» (от 3 до 8% от общего числа)
- Карточки от источников с темой «Семейные отношения» получают больше всего взаимодействий (10.7%)
- Карточки «общество» от источников с темой Россия получили 3471 взаимодействие