

Анализ взаимодействия пользователей с карточками статей Яндекс.Дзен

Воробьёва К.

-
- ▶ Задачи исследования
 - ▶ Источник данных
 - ▶ Взаимодействие пользователей с карточками статей по темам
 - ▶ Разбивка событий по темам источников
 - ▶ Соотношение тем карточек и тем источников
 - ▶ Дашборд «Анализ взаимодействия пользователей с карточкам...»
 - ▶ Выводы
 - ▶ Контакты

Задачи:

1) Необходимо определить:

- ▶ 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?

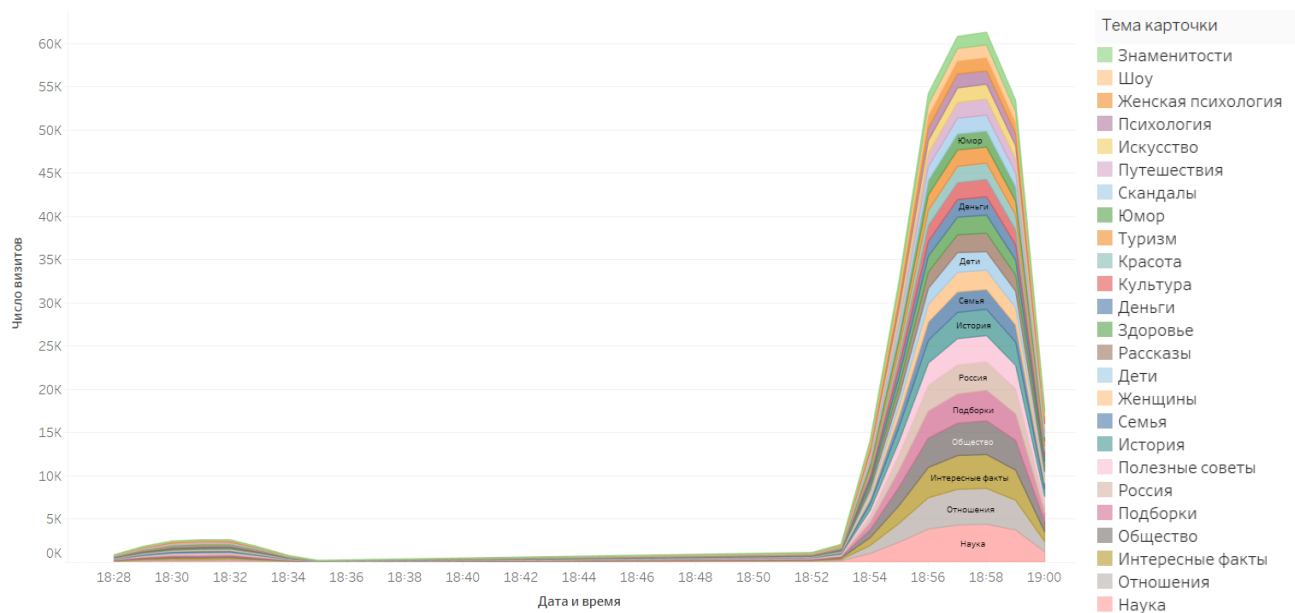
2) Построить дашборд с ежедневным обновлением информации

Источник данных

Данные для анализа и построения дашборда основываются на пайплайне, разработанном дата-инженером, который берет сырые данные из таблиц взаимодействия пользователей с карточками статей Яндекс.Дзен, трансформирует данные и укладывает в агрегирующую таблицу.

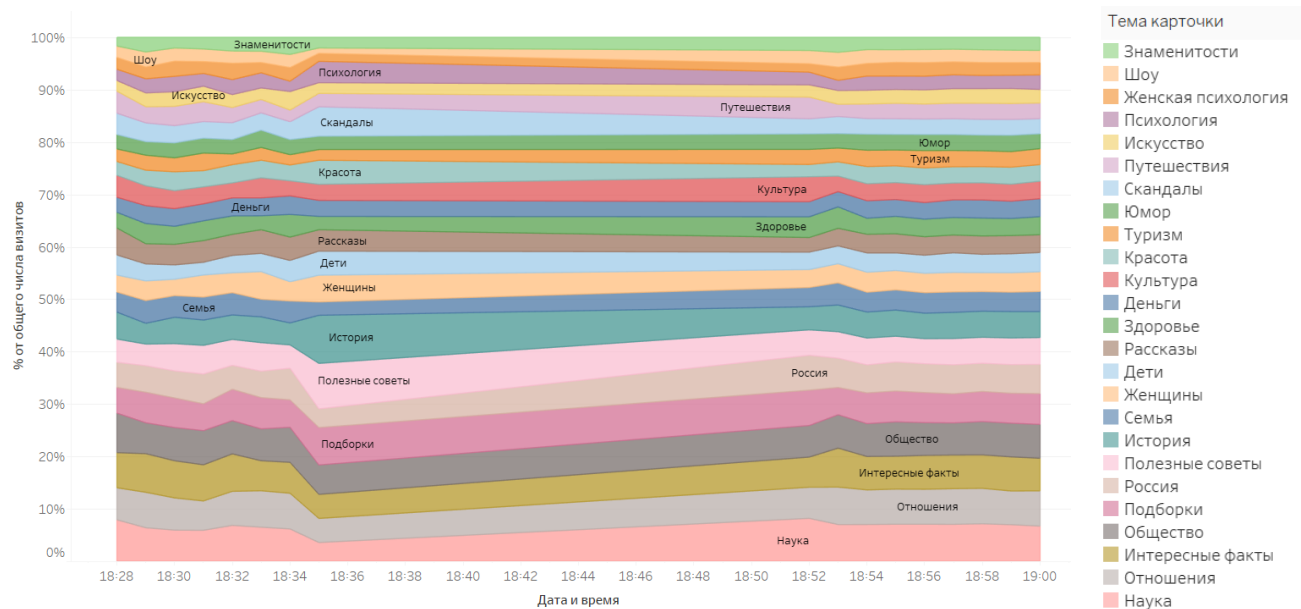
Частота обновления данных: один раз в сутки, в полночь по UTC.

Взаимодействие пользователей с карточками статей по темам



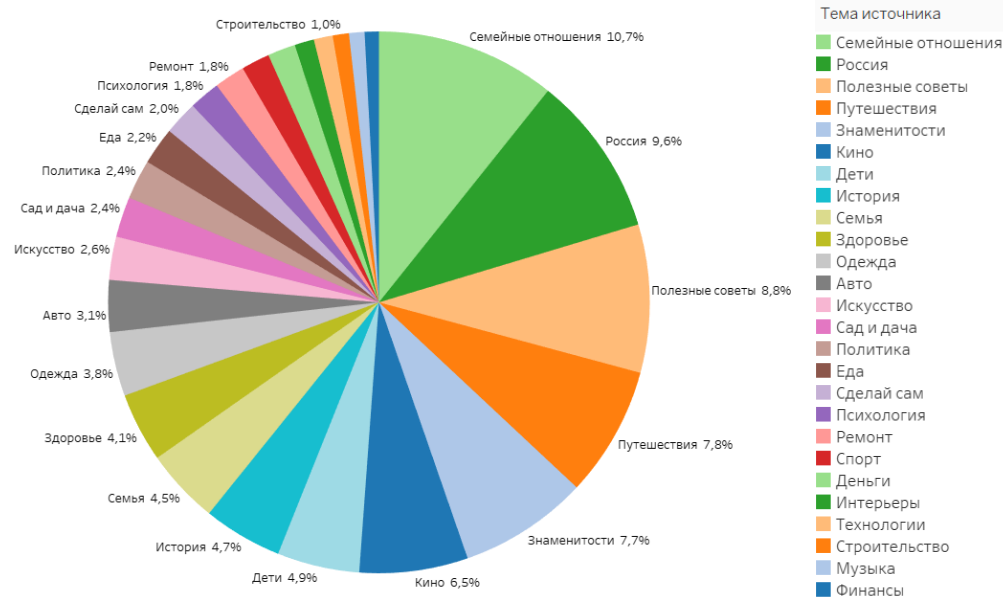
В пике (61247 событий) максимальное число событий у карточек с темами «Наука» - 4372, «Отношения» - 4145 и «Интересные факты» - 3910.

Взаимодействие пользователей с карточками статей по темам



Из 25 тем карточек, 50% событий приходится на 9 тем: «Наука», «Отношения», «Интересные факты», «Общество», «Подборки», «Россия», «Полезные советы», «История» и «Семья».

Разбивка событий по темам источников



Наибольшее число событий генерируют источники с темами «Семейные отношения» - 10.7%, «Россия» - 9.6%, «Полезные советы» - 8.8%. Меньше всего генерируют источники с темами «Финансы» - 0.8%, «Музыка» - 0.9% и «Строительство» - 1%

Соотношение тем карточек и тем источников

Тема карточки	Тема источника								
	Путешествия	Семейные отношения	Знаменитости	История	Семья	Деньги	Россия	Полезные советы	Кино
Наука	1 823	1 711	1 554	1 102	809	197	1 782	1 865	3 279
Отношения	1 269	1 030	2 040	679	194	281	2 278	2 716	672
Интересные факты	1 448	1 323	720	1 273	611	145	2 567	2 023	1 033
Общество	891	2 727	829	1 116	1 416	350	3 471	1 313	578
Подборки	1 491	1 386	1 353	597	424	217	877	2 795	1 403
Россия	1 079	1 962	1 579	1 104	608	229	2 847	730	919
Полезные советы	2 088	1 781	271	437	811	916	403	1 413	143
История	2 643	1 344	921	636	696	114	1 491	859	1 091
Семья	340	1 489	689	266	1 131	458	519	1 293	486
Женщины	192	2 270	967	560	988	122	397	986	340
Дети	431	1 184	1 444	455	558	228	547	1 541	474
Рассказы	4 587	875	615	568	539	454	421	358	357

Наибольшее число событий сгенерировали карточки с темой «Рассказы» и источником «Путешествия» - 4587. Высокие показатели генерации событий у карточек с темой «Общество» и источником «Россия» - 3471 и карточек с темой «Наука» и источником «Кино» - 3279.

Соотношение тем карточек и тем источников

Тема карточки	Тема источника											
	Сад и дача	Ремонт	Сделай сам	Финансы	Авто	Строительство	Политика	Интерьеры	Музыка	Одежда	Еда	Технологии
Туризм	523	403	688	83	161	74	307	279	77	248	382	38
Юмор	321	201	204	41	224	55	17	17	67	384	194	11
Скандалы	24	13	56	13	14		374	93	132	281	48	273
Путешествия	499	451	127	77	343	85	60	26	33	382	357	44
Искусство	43	172	128	4	145	53	121	164	223	209	7	58
Психология	86	34	45	76	19	88	224	46	19	183	62	23
Женская психолог..	49	13	19	5	26		23		5	466	67	39
Шоу		7		354	4	8	522	30	98	393	1	
Знаменитости	21	3	6	59	29	50	35	12	55	467	36	158

Для большинства тем карточек и источников существуют сгенерированные события, но есть темы, не имеющие «точек соприкосновения»:

- Карточки с темой «Шоу» не имели среди источников тем «Сад и дача», «Сделай сам» и «Технологии»
- Карточки с темой «Женская психология» не имели источников с темой «Строительство» и «Интерьеры»
- Карточки с темой «Скандалы» не имели среди источников тему «Строительство»

Дашборд

«Анализ взаимодействия пользователей с карточками Яндекс.Дзен»

Анализ взаимодействия пользователей с карточками Яндекс.Дзен

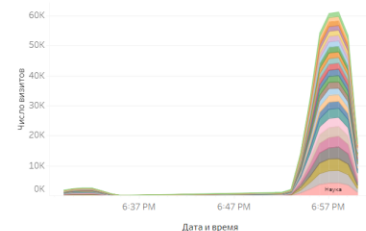
На графиках представлена история взаимодействия пользователей по темам карточек (в абсолютных и относительных величинах), разбивка событий по темам источников в абсолютных величинах и таблица соответствия тем карточек источникам с указанием числа взаимодействий в абсолютных величинах.
Фильтры данных - дата и время событий, тема карточек и возрастной сегмент пользователей.
Данные о взаимодействии пользователей с карточками статей обновляются один раз в сутки.

Дата и время
6:28 11 PM
7:00:00 PM

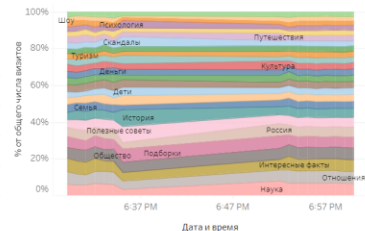
Тема карточек
(All)

Возрастной сегмент
(All)

События по темам карточек



% событий по темам карточек



События по темам источников

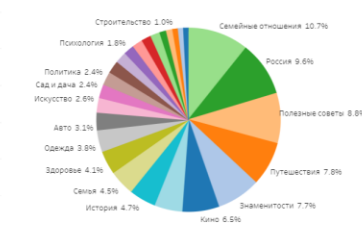


Таблица соответствия тем карточек темам источников

Тема карточки	ID	Тема источника											
		Путешествия	Семейные отношения	Знания	История	Семья	Деньги	Россия	Полезные советы	Кино	Искусство	Дети	Зд
Наука		1,818	1,709	1,549	1,097	804	196	1,779	1,856	3,268	510	927	
Отношения		1,266	1,008	2,037	678	194	280	2,270	2,707	671	373	241	
Интересные факты		1,444	1,320	717	1,272	606	145	2,560	2,018	1,030	694	326	
Общество		885	2,719	827	1,113	1,415	390	3,456	1,308	575	218	1,005	
Подборки		1,488	1,382	1,350	595	424	217	875	2,787	1,398	257	319	

- На графиках представлена история взаимодействия пользователей по темам карточек (в абсолютных и относительных величинах), разбивка событий по темам источников в относительных величинах и таблица соответствия тем карточек источникам с указанием числа взаимодействий в абсолютных величинах.
- Фильтры данных - дата и время событий, тема карточек и возрастной сегмент пользователей.
- Данные о взаимодействии пользователей с карточками статей обновляются один раз в сутки.

[Ссылка на дашборд](#)

Выводы:

- ▶ Из 25 тем карточек, 50% событий приходится на 9 тем: «Наука», «Отношения», «Интересные факты», «Общество», «Подборки», «Россия», «Полезные советы», «История» и «Семья». При этом в пике генерации наибольшее число событий у карточек с темами «Наука», «Отношения», «Интересные факты».
- ▶ Наибольшее число событий генерируют источники с темами «Семейные отношения» - 10.7%, «Россия» - 9.6%, «Полезные советы» - 8.8%. Меньше всего генерируют источники с темами «Финансы» - 0.8%, «Музыка» - 0.9% и «Строительство» - 1%
- ▶ Что касается соотношения тем карточек и источников, Наибольшее число событий сгенерировали карточки с темой «Рассказы» и источником «Путешествия» - 4587. Высокие показатели генерации событий у карточек с темой «Общество» и источником «Россия» - 3471 и карточек с темой «Наука» и источником «Кино» - 3279. Для большинства тем карточек и источников существуют сгенерированные события, но есть темы, не имеющие «точек соприкосновения»: карточки с темой «Шоу» не имели среди источников тем «Сад и дача», «Сделай сам» и «Технологии», карточки с темой «Женская психология» не имели источников с темой «Строительство» и «Интерьеры», карточки с темой «Скандалы» не имели среди источников тему «Строительство»
- ▶ Построен дашборд с ежедневным обновлением информации

Спасибо за внимание!

vorobeva@ya-team.ru