

Анализ взаимодействия пользователей с карточками статей Яндекс.Дзен

Исследование для менеджеров по анализу контента Яндекс.Дзен

Описание проекта

Автоматизация анализа пользовательского взаимодействия с карточками статей Яндекс. Дзен и построении дашборда для менеджеров по анализу контента.

Техническое задание:

Бизнес-задача: анализ взаимодействия пользователей с карточками Дзена;
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;

Структура данных источника:

Каждую карточку статьи определяет тема `item_topic` и источник (у него тоже есть тема `source_topic`);
Пользователей системы характеризует возрастная категория `age_segment`: 18-25, 26-30, 31-35, 36-40, 41-45, 45+;
Взаимодействие с карточками статей характеризуется показателем `visits` – количество событий.

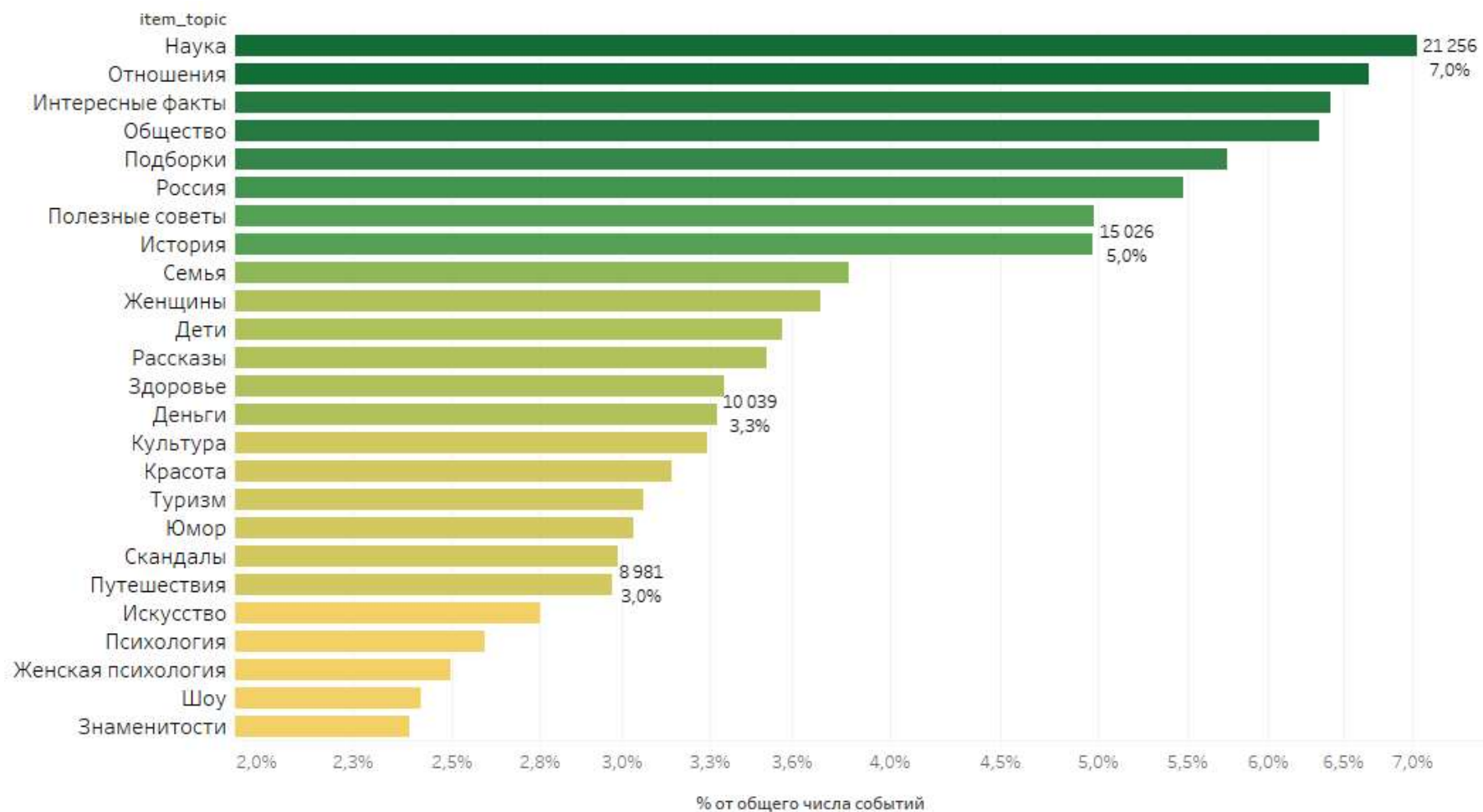
Источники данных для дашборда: дашборд основывается на пайплайне, который берет данные из таблицы, в которой хранятся сырые данные, трансформирует данные и укладывает их в агрегирующую таблицу `dash_visits`. Пайплайн разработан командой дата-инженеров. Частота обновления данных: один раз в сутки, в полночь по UTC;

Выполнение проекта

Для выполнения проекта были сделаны следующие шаги:

1. Создание коннекта к базе данных;
 2. Выполнение sql-запроса на выгрузку данных из таблицы `dash_visits`, подготовленной дата-инженерами в файл, пригодный для создания дашборда;
 3. Формирован дашборда в Tableau public;
 4. Публикация дашборда на сайте Tableau public
 5. Подготовка презентация с детализацией анализа пользовательского взаимодействия с карточками статей Яндекс. Дзен.
-

Число событий по темам карточек (абсолютные и относительные величины)*:



Большей популярностью пользуются карточки статей по таким темам как: Наука, Отношения, Интересные факты, Общество, Подборки и Россия (от 5,5 % до 7 % событий приходится на данные темы). Самые не популярные статьи на темы: Знаменитости, Шоу, Женская психология, Психология и Искусство (менее 3 % событий).

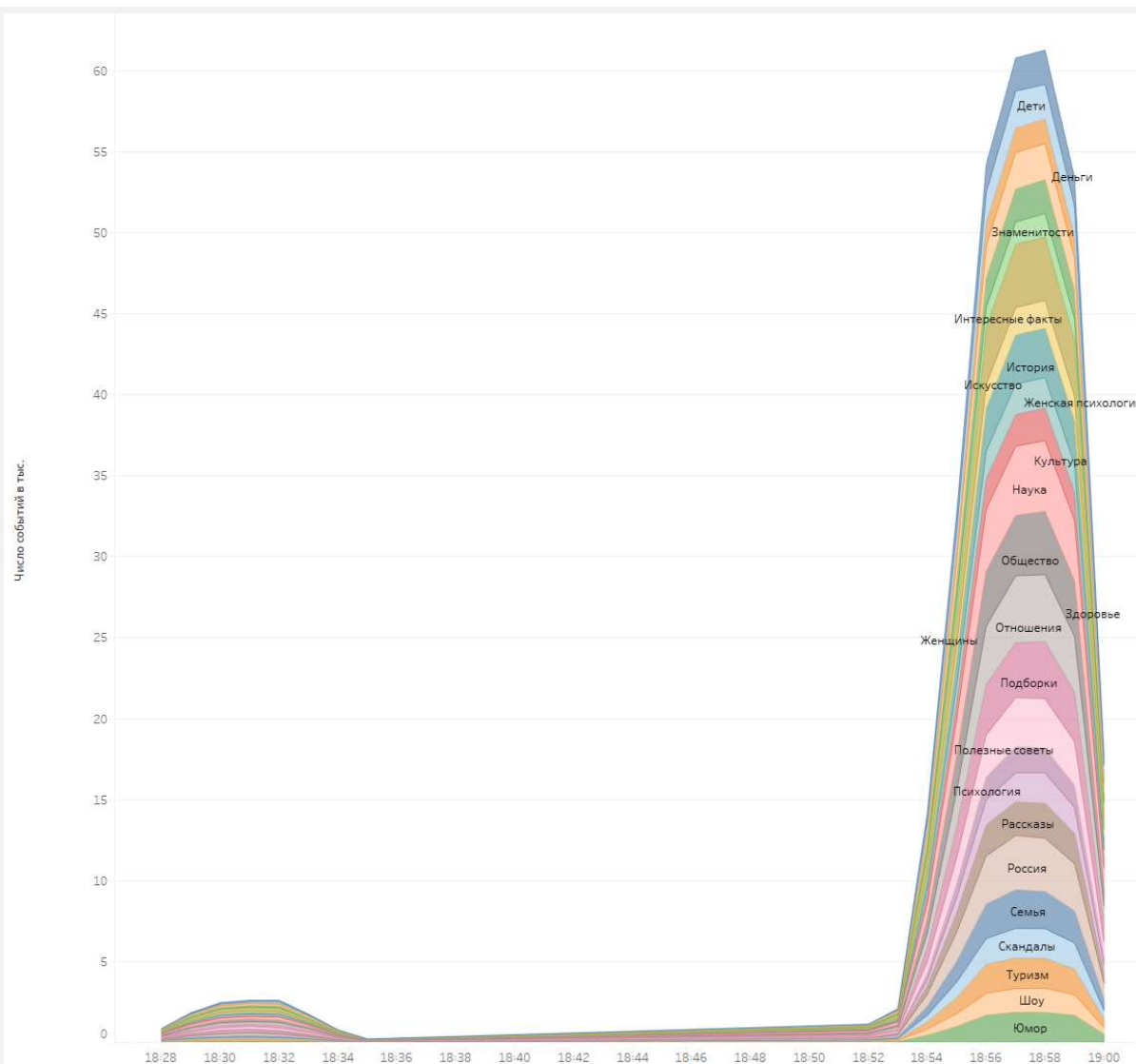
* На основе выгрузки за период с 18:28 до 19:00, 24.09.2019 г.

История событий по темам карточек статей (абсолютные величины)*

Наибольшее число событий происходит в период с 18:53 до 19:00, пик событий приходится на 18:58.

В 18:58 лидеры по числу событий карточки с темами:

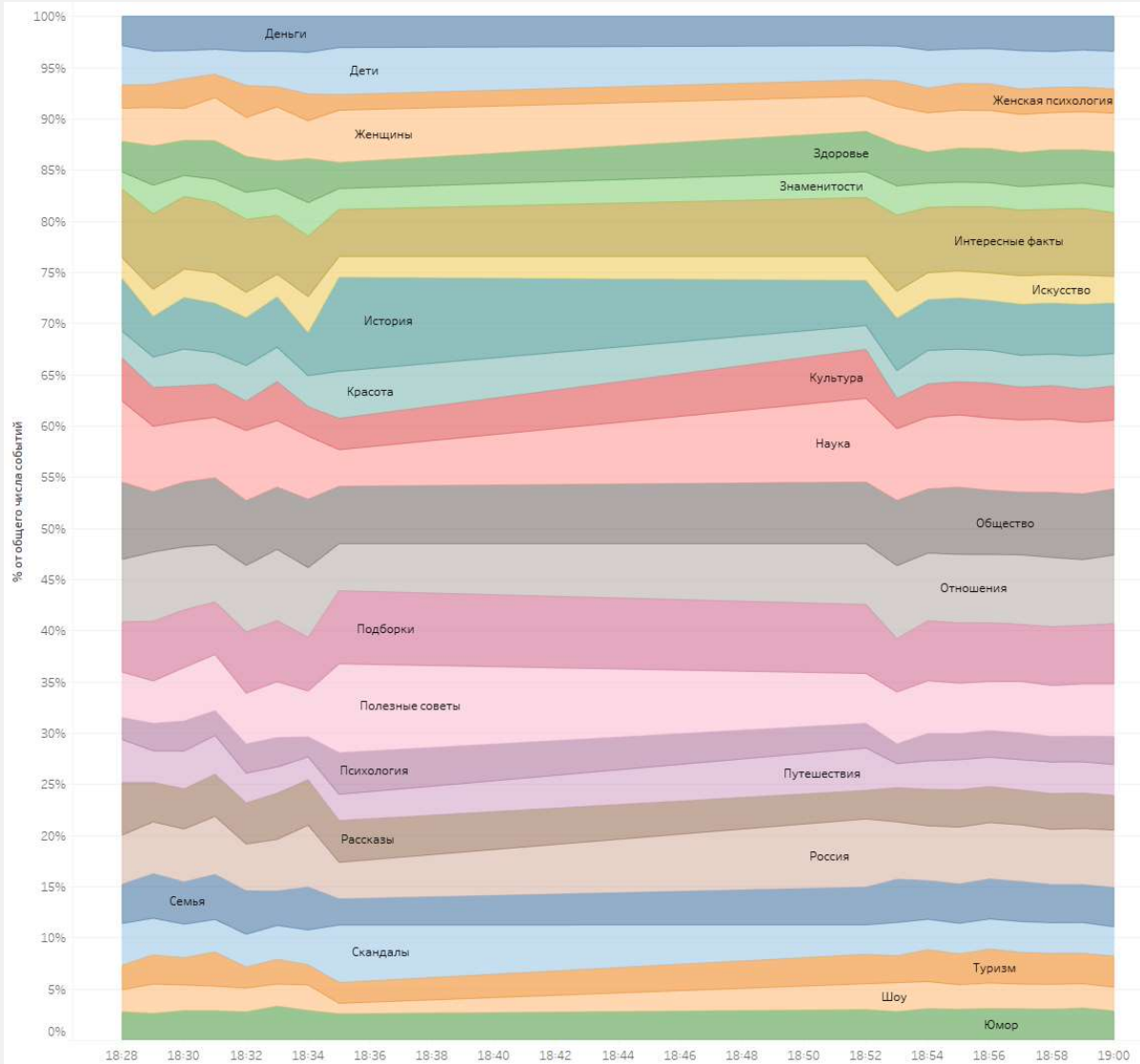
- Наука (4,372),
- Отношения (4,145),
- Интересные факты (3,910),
- Общество (3,897),
- Подборки (3,520),
- Россия (3,296).



* На основе выгрузки за период с 18:28 до 19:00, 24.09.2019 г.

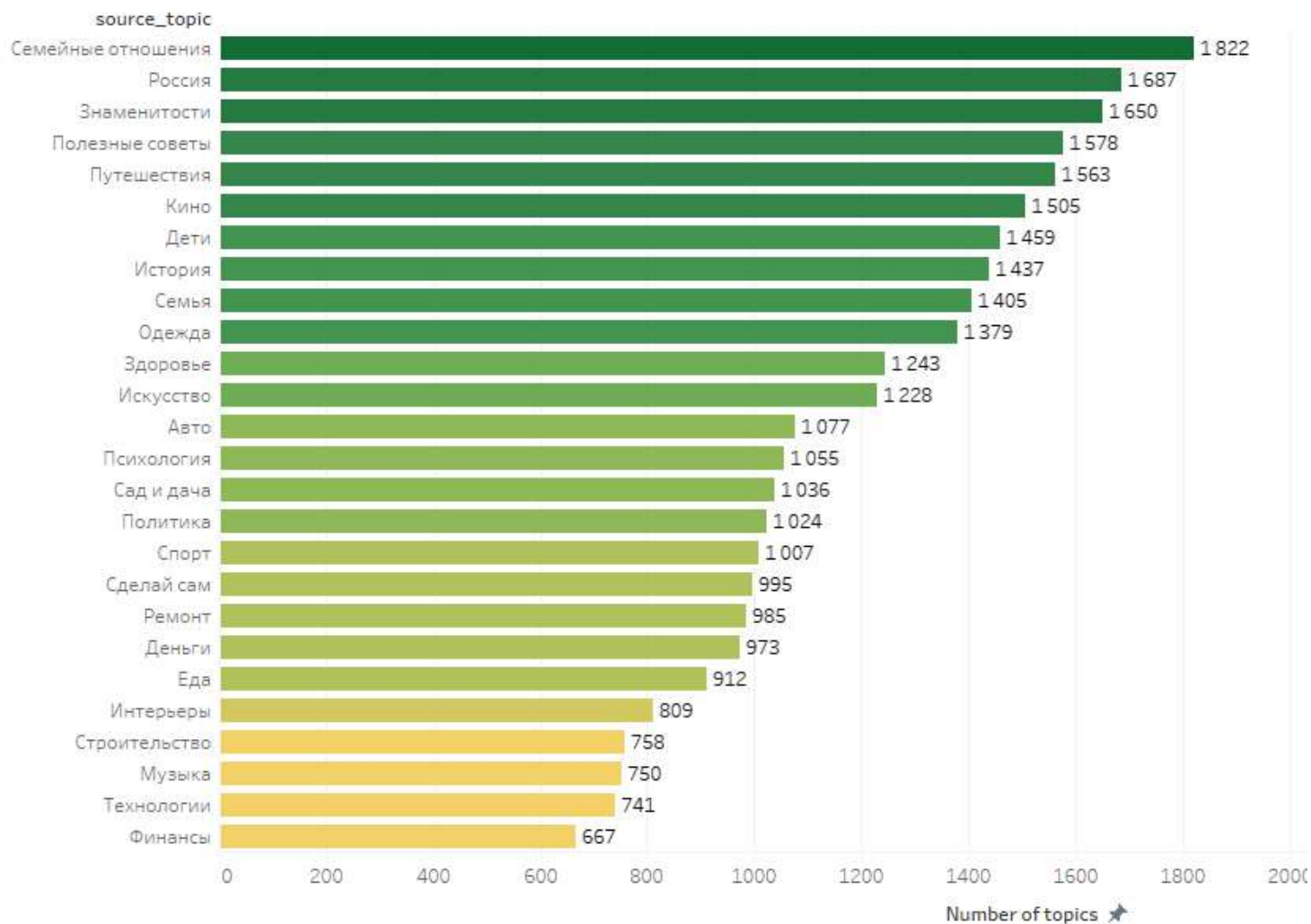
История событий по темам карточек статей (относительные величины)*

В период с 18:28 по 18:36⁵ видны колебания событий по темам, к концу часа(19:00) относительные величины событий по темам карточек стабилизируются.



* На основе выгрузки за период с 18:28 до 19:00, 24.09.2019 г.

Число карточек по темам источников (абсолютные величины)*:



Больше всего публикаций идет от источников с темами: Семейные отношения, Россия, Знаменитости, Полезные советы, Путешествия (до 1,5 тыс. статей). Что наполовину пересекается с популярными темами карточек статей на Яндекс. Дзен. Реже всего присутствуют статьи от источников с темами: Финансы, технологии, Музыка, Строительство (до 800 статей).

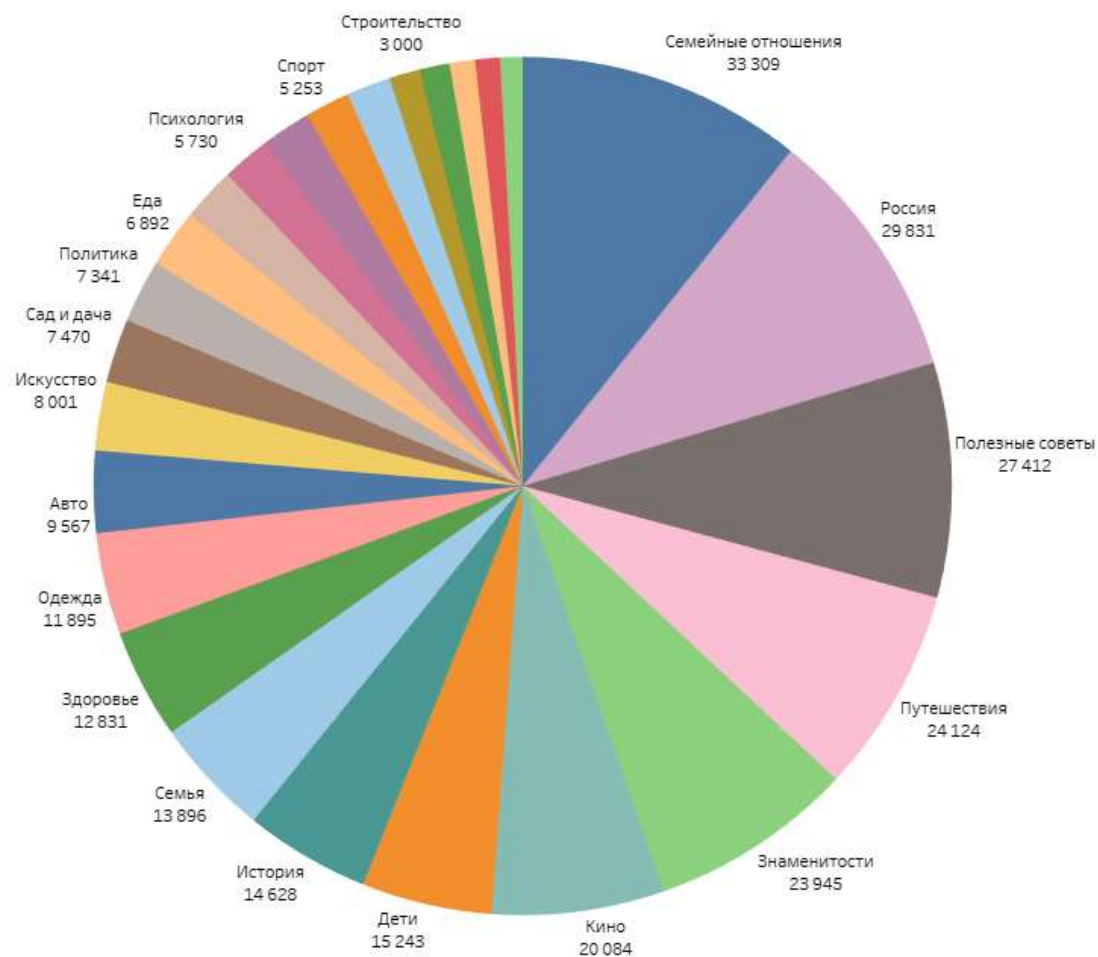
* На основе выгрузки за период с 18:28 до 19:00, 24.09.2019 г.

События по темам источников (абсолютные величины)*

ТОП-5 тем источников,
с карточками статей,
которых у пользователей
происходит больше всего
взаимодействий:

1. Семейные отношения;
2. Россия
3. Полезные советы
4. Путешествия
5. Знаменитости

*Ранее отмечалось, что
этим источниками
создано больше всего
карточек статей*



* На основе выгрузки за период с 18:28 до 19:00, 24.09.2019 г.

Темы карточек статей от ТОП-5 тем источников:

item_topic	source_topic				
	Знамени..	Полезные советы	Путешес..	Россия	Семейны е отнош..
Наука	76	78	78	82	72
Отношения	77	73	73	85	73
Интересные факты	65	78	73	83	81
Подборки	71	85	78	63	79
Общество	64	77	69	84	82
Россия	76	60	70	77	88
История	69	62	82	70	75
Полезные советы	54	76	76	57	76
Здоровье	67	84	52	62	65
Рассказы	61	53	94	58	62
Дети	67	78	53	61	68
Юмор	51	73	63	64	75
Красота	76	68	63	53	65
Деньги	59	72	58	67	68
Женщины	70	69	43	58	82
Культура	79	44	53	76	69
Семья	64	76	51	55	73
Путешествия	59	63	68	59	70
Искусство	70	56	62	62	69
Туризм	53	59	58	69	77
Скандалы	81	22	57	83	68
Психология	61	66	44	66	62
Женская психология	58	49	50	59	81
Знаменитости	58	36	55	83	62
Шоу	64	21	40	51	80

#	Тема карточек	Число статей по ТОП-5 источникам
1	Наука	386
2	Отношения	381
3	Интересные факты	379
4	Подборки	376
5	Общество	376
6	Россия	371

ТОП-5 тем источников больше всего генерируют карточки статей с такими же темами, исключением стали карточки статей с темой «Интересные факты», источников с точно такой же темой нет.

* ТОП-5 тем источников определены на основе выгрузки событий за период с 18:28 до 19:00, 24.09.2019 г.

Таблица соответствия тем карточек к темам источников:

	source_topic																										
	Авто	Деньги	Дети	Еда	Здоровье	Знаменитости	Интерьеры	Искусство	История	Кино	Музыка	Одежда	Полезные советы	Политика	Психология	Путешествия	Ремонт	Россия	Сад и дача	Сделай сам	Семейные отношения	Семья	Спорт	Строительство	Технологии	Финансы	
item_topic	1																										
Наука	1606	197	930	91	629	1554	186	510	1102	3279	403	393	1865	244	291	1823	415	1782	227	358	1711	809	421	99	749	62	
Отношения	478	281	242	573	1173	2040	862	374	679	672	139	1428	2716	830	197	1269	510	2278	944	517	1030	194	312	513	119	296	
Интересные факты	1254	145	327	391	2090	720	54	697	1273	1033	325	837	2023	340	449	1448	175	2567	825	341	1323	611	186	234	219	55	
Общество	379	350	1007	160	1181	829	64	218	1116	578	79	355	1313	1209	919	891	181	3471	76	162	2727	1416	338	97	118	406	
Подборки	551	217	319	871	396	1353	122	257	597	1403	190	1612	2795	43	396	1491	479	877	522	364	1386	424	280	361	440	26	
Россия	1885	229	493	23	282	1579	134	355	1104	919	324	276	730	232	38	1079	345	2847	85	114	1962	608	703	179	316	125	
Полезные советы	749	916	707	223	1346	271	427	83	437	143	127	891	1413	51	438	2088	485	403	253	679	1781	811	136	77	294	206	
История	522	114	1047	60	771	921	30	695	636	1091	164	310	859	695	227	2643	168	1491	209	149	1344	696	190	9	286	62	
Семья	50	458	361	1236	275	689	321	288	266	486	41	544	1293	78	176	340	447	519	548	434	1489	1131	4	351	34	38	
Женщины	156	122	875	348	1022	967	175	485	560	340	45	547	986	78	253	192	67	397	125	223	2270	988	210	56	6	6	
Дети	192	228	621	675	346	1444	97	238	455	474	9	320	1541	135	128	431	174	547	633	354	1184	558	21	103	28	53	
Рассказы	78	454	286	28	248	615	47	291	568	357	36	30	358	63	233	4587	123	421	201	92	875	539	222	74	1	82	
Здоровье	232	95	798	558	404	962	57	64	68	156	20	518	2335	121	214	393	145	712	542	548	814	247	197	106	74	19	
Деньги	338	203	859	193	65	591	62	172	271	1042	42	98	1087	949	305	536	362	1132	101	137	696	345	134	68	88	415	
Культура	103	46	578	55	41	1508	124	682	636	1543	178	295	231	308	81	430	154	1474	100	76	892	482	47	86	46	9	
Красота	29	48	703	253	736	1313	185	232	177	472	38	428	1227	282	171	781	172	269	513	466	748	274	28	174	39	56	
Туризм	161	74	406	382	311	269	279	173	618	252	77	248	721	307	133	483	403	867	523	688	1580	283	79	74	38	83	
Юмор	224	79	404	194	124	334	17	205	715	502	67	384	1150	17	140	636	201	621	321	204	1887	584	281	55	11	41	
Скандалы	14	35	844	48	43	1992	93	91	410	707	132	281	54	374	18	284	13	1560	24	56	952	386	597		273	13	
Путешествия	343	363	509	357	187	509	26	555	1018	289	33	382	888	60	126	641	451	588	499	127	709	306	88	85	44	77	
Искусство	145	98	639	7	70	1083	164	512	596	1254	223	209	331	121	98	664	172	632	43	128	754	404	54	53	58	4	
Психология	19	86	1233	62	583	582	46	306	286	219	19	183	862	224	141	261	34	936	86	45	855	656	125	88	23	76	
Женская психология	26	256	613	67	349	546		96	181	264	5	466	418	23	463	321	13	660	49	19	2073	759	26		39	5	
Шоу	4	8	297	1	20	829	30	66	215	2201	98	393	38	522	7	113	7	505			1503	151	141	8		354	
Знаменитости	29	55	145	36	139	445	12	356	644	408	55	467	178	35	88	299	3	2275	21	6	764	234	433	50	158	59	

Общие итоги проекта

По результатам анализа данных за исследуемый период можно сделать выводы:

1. Разным темам источников соответствует огромное разнообразие тем карточек статей Яндекс. Дзен;
2. Пользовательский интерес к карточкам статей не постоянен и сильно меняется в течение времени;
3. Больше популярностью пользуются карточки статей по таким темам как: **Наука, Отношения, Интересные факты, Общество, Подборки и Россия** (от 5,5 % до 7 % событий приходится на данные темы).
4. Самые не популярные статьи на темы: Знаменитости, Шоу, Женская психология, Психология и Искусство (менее 3 % событий).
5. Были определены ТОП-5 тем источников, имеющих больше всего публикаций карточек статей: **Семейные отношения, Россия, Полезные советы, Путешествия и Знаменитости** (до 1,5 тыс. статей). Что наполовину пересекается с популярными темами карточек статей на Яндекс. Дзен.
6. ТОП-5 тем источников больше всего генерируют карточки статей с такими же темами, исключением стали карточки статей с темой «Интересные факты», источников с точно такой же темой нет.
7. Реже всего присутствуют статьи от источников с темами: Финансы, технологии, Музыка, Строительство (до 800 статей).

В ходе выполнения проекта был подготовлен дашборд для менеджеров по анализу контента Яндекс. Дзен:

Дашборд, показывающий взаимодействие пользователей с карточками статей Яндекс. Дзен опубликован на сайте Tableau public и доступен по ссылке :

https://public.tableau.com/app/profile/suriana.safina/viz/Ya_Dzen_dash2/Dashboard1

При работе с дашбордом могут быть использованы следующие фильтры:

Фильтр по времени события с 18:28 до 19:00 24.09.2019 г.;

Фильтр по возрастным категориям пользователей от 18 до 45+;

Фильтр по темам карточек статей;



СПАСИБО

Проект подготовлен Сафиной С.Ф.
suriyana.safina@tedo.ru
+ 7 903 387-97-95