АНАЛИЗ ПОЛЬЗОВАТЕЛЬСКОГО ВЗАИМОДЕЙСТВИЯ С КАРТОЧКАМИ СТАТЕЙ В ЯНДЕКС.ДЗЕНЕ

Михаил Козелков

ОГЛАВЛЕНИЕ:

Описание проекта	. 3
Дашборд	4
Ответ на вопрос «Сколько взаимодействий пользователей с карточка происходит в системе с разбивкой по темам карточек?»	
Ответ на вопрос «Как много карточек генерируют источники с разны темами?»	
Ответ на вопрос «Как соотносятся темы карточек и темы источников?»	7
Заключение и вывод	8

Описание проекта:

Я работаю аналитиком в Яндекс. Дзене. Почти всё моё время занимает анализ пользовательского взаимодействия с карточками статей.

Каждую карточку определяют её тема и источник (у него тоже есть тема).

Пользователей системы характеризует возрастная категория.

Каждую неделю начинающие менеджеры Денис и Валерия задают мне одни и те же вопросы:

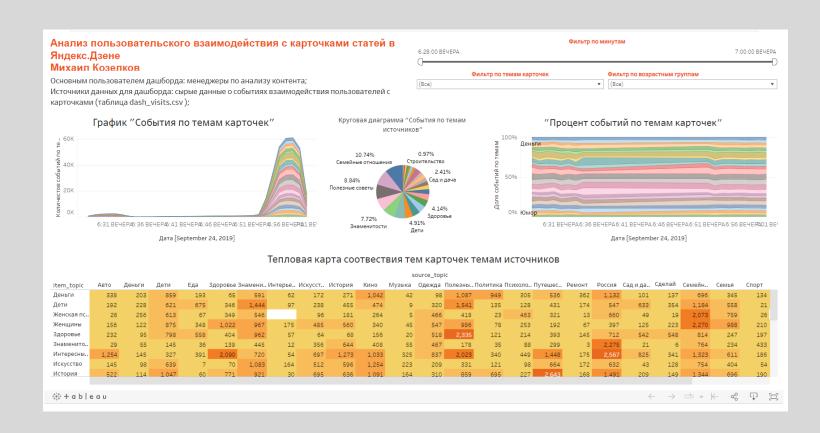
- Сколько взаимодействий пользователей с карточками происходит в системе с разбивкой по темам карточек?
- Как много карточек генерируют источники с разными темами?
- Как соотносятся темы карточек и темы источников?

На шестую неделю работы я решаю, что процесс пора автоматизировать. Для Дениса и Валерии нужно сделать дашборд.

В данной презентации я постараюсь ответить на вопросы менеджеров, используя дашборд:

Данные для Дашборда получены по средствам доступа к базе данных «data-analyst-zen-project-db», предоставленного мне Дата-инженерами Яндекс.Дзен.

С помощью инструментов языка Python и SQL данные были преобразованы в агрегирующую таблицу «dash_visits.csv» с последующим использованием в системе интерактивной аналитики Tableau



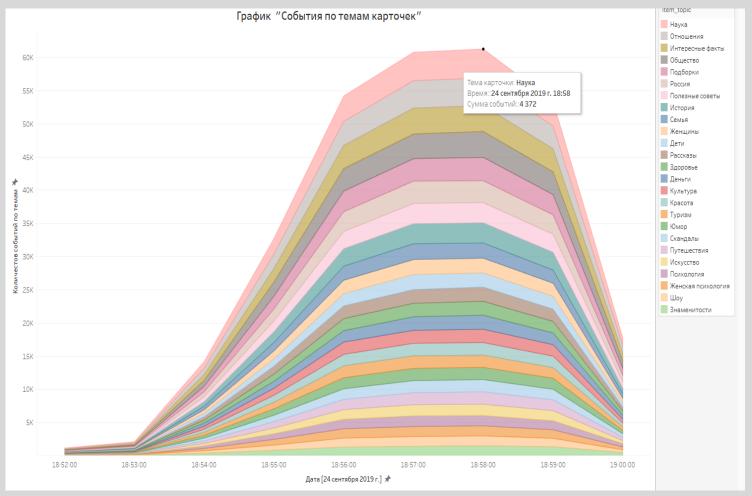
1.Сколько взаимодействий пользователей с карточками происходит в системе с разбивкой по темам карточек?

Пик событий приходится на время **18-58**.

Всего около 62 тыс.

Топ 5 тем

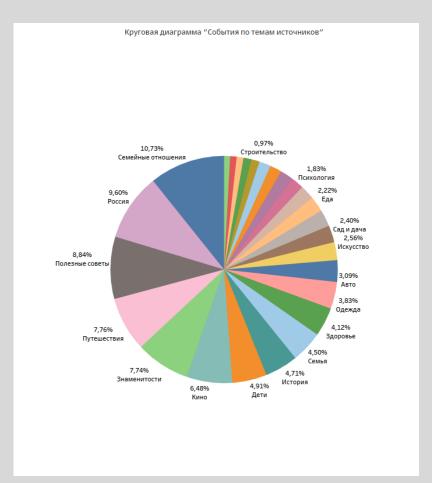
- Наука,
- Отношения,
- Интересные факты,
- Общество,
- Подборки,
 примерно по 4 тыс.
 событий каждая.

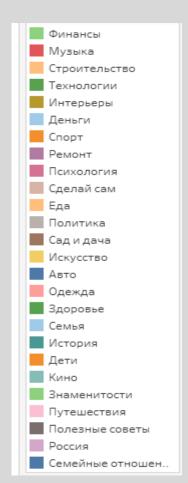


2. Как много карточек генерируют источники с разными темами?

Топ 5 источников:

- Семейные отношения
- Россия
- Полезные советы
- Путешествия
- Знаменитости





Как соотносятся темы карточек и темы источников?

На этот вопрос нам ответит тепловая карта. Можно видеть, что больше всего переходов из источника темы Путешествия в карточки темы Рассказы

	·																				
	source_topic																				
	Полезные												Сади	Сделай	Семейны						
item_topic							Кино	Музыка	Одежда		Политика П				Россия	дача	сам	е отнош	Семья		Строите.
Деньги	184	62	558	60	165	256	1010	40	93	1 052	926	287	515	347	1087	97	134	660	326	127	64
Дети	652	339	1 414	91	226	431	457	9	305	1 483	126	123	418	167	532	604	338	1129	537	18	102
Женская пс	65	337	524		94	169	252	5	443	400	22	442	308	13	638	49	19	1977	727	25	
Женщины	326	970	930	168	461	527	330	42	514	949	77	242	186	65	380	119	213	2 174	959	196	54
Здоровье	531	382	927	53	62	66	144	18	498	2 226	113	201	378	141	687	519	531	780	229	188	99
Знаменито	35	135	432	12	342	624	393	52	442	171	33	85	285	2	2 171	21	6	728	224	412	4!
Интересны	373	2 009	686	54	659	1224	989	309	805	1945	323	416	1 400	166	2 452	774	324	1 261	575	177	229
Искусство	7	64	1038	159	498	569	1 202	219	197	313	116	93	636	165	598	41	122	722	390	52	49
История	57	740	876	26	661	604	1052	160	293	834	673	211	2 571	157	1426	201	141	1 283	664	185	9
Красота	245	694	1 256	181	217	170	457	36	411	1 172	275	166	740	164	256	488	445	711	267	27	171
Культура	54	40	1 444	119	642	601	1 484	166	287	223	299	78	415	146	1410	95	74	851	460	46	83
Наука	89	590	1 497	177	490	1057	3 142	397	379	1790	229	277	1746	401	1718	217	346	1658	782	399	9!
Общество	155	1 129	800	64	211	1075	560	77	337	1 244	1167	875	844	178	3 326	70	158	2 6 1 5	1364	326	96
Отношения	548	1 119	1971	832	349	646	649	133	1377	2611	808	189	1221	489	2 157	907	497	994	182	302	493
Подборки	835	384	1304	110	248	573	1350	185	1535	2 691	41	379	1 429	458	849	500	351	1324	405	262	348
Полезные с	213	1 281	255	411	83	419	135	124	848	1355	51	423	2 013	457	383	235	654	1701	791	131	74
Психология	58	560	559	44	294	274	210	19	176	818	215	137	253	32	887	84	44	816	631	116	8:
Путешеств	338	172	491	25	521	986	275	29	363	852	60	122	607	434	563	480	121	672	296	83	83
Рассказы	28	235	586	44	283	548	344	35	29	344	59	221	4 294	119	399	192	88	848	520	214	73
Россия	23	272	1523	129	334	1062	869	302	265	696	223	38	1040	220	2716	00	111	1886	593	674	170
Семья	183	259	658	302	268	245	461	38	525	1 225	76	163		Тема карточки: Расска			417	1 432	1105	4	337
Скандалы	48	41	1921	82	88	396	679	126	266	51	358	16		Тема источника:: Путец			52	907	370	578	
Туризм	375	301	252	268	167	594	247	76	236	699	299	128	Суг	Сумма событий:: 4 294			661	1515	272	77	7:
Шоу	1	20	800	29	63	209	2 097	94	375	36	497	5	102	7	490			1 447	147	131	
Юмор	184	119	317	17	197	682	488	65	372	1 107	17	134	609	190	591	313	201	1815	563	268	58
	<																				>

Тепловая карта соотвествия тем карточек темам источников

Заключение и вывод:

Исследование показало, что пик событий приходится примерно на 19-00, когда как я полагаю пользователи заканчивают работу и активно читают новости и статьи. На данную дату было зафиксировано 62 тыс. событий, где чаще всего это были темы: Наука, Отношения, Интересные факты, Общество и Подборки.

Больше всего карточек приходит из источников с такими темами как, Семейные отношения, Россия, Полезные советы, Путешествия и Знаменитости. Их доля в сумме составляет почти 44 % от общего числа.

Тепловая карта наглядно показывает, как соотносятся темы карточек и темы источников.

Считаю, что Дашборд отвечает на частые вопросы менеджеров, но так же он получился довольно информативным. В нем есть такие фильтры как: Дата и время, Возраст пользователей и Темы карточек. Четыре графика могут дать исчерпывающую информацию и помочь увидеть нужные показатели.

Ссылка на дашборд: https://public.tableau.com/views/AnalysisYa dzen/Dashboard1?:language=en-US&publish=yes&:display count=n&:origin=viz share link

Контакты: Михаил Козелков

тел. 8-999-777-77-77

Telegram: @tapbich