

Анализ пользовательского взаимодействия с карточками статей Яндекс.Дзен

за 24.09.2019 18:28-19:01

Хабарова Зарина,
email: kurbanova.zarina1995@yandex.ru

2023 год

Оглавление

Анализ пользовательского взаимодействия с карточками статей Яндекс.Дзен

66 14-000000 00.00-00.00

bioRxiv preprint doi: <https://doi.org/10.1101/2020.06.01.320103>; this version posted June 1, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Вступление

Описание проекта: Для информатизации учебного психологического взаимодействия с карточками статей Юлиана Дюва вводится система диалогов. Каждую карточку обрабатывает от темы и вступает (иногда даже есть тема). Примеры это «Книжки и диалоги», «Рецензии», «Путешествия». Психологической системы характеризует возрастная категория. Статьи, «26-36»

Задачи по дисциплине: освоение приемов, обеспечивающих проведение анализа сложившейся ситуации с учетом ее стабильности. Анализ Дви при планировании построения дислокации и отнесения их к установленным нормам.

Цель исследования: разработать алгоритм, по которому менеджеры образовательных учреждений оперативно проводят анализ информационного взаимодействия с корпоративной сетью Яндекса. Для чего необходимо составить программу разработки нового алгоритма на основании анализа данных в той же системе.

Источники данных: даны-индексы погодеопикли ирегитрууюую таблицу *day_vinfo*, на основании этих данных был проведён анализ.

Источник данных

Давление основывается на тиббалайне, который будет брать давление по тиббалайне, в котором находится сердце дитя, трансформировать давление и указывать на в атрофическом тиббалайне. Пайбалайне разработана для-инженером.

Даты-индексы подготовлены агрегирующей таблицей `dash_vists`. Вот её структура:

nscol_id — номер колонки.

Иван, Игорь — тоже изобретатели.

Алекс_Ферс — наш источник,

apo_segment — компактный сегмент.

di — carta e spina,

TABLE 1. *Continued*

Journal of Management Inquiry 23(1) 3-16

Общие выводы по дашборду

- [illegible]

Краткое техническое задание (ТЗ) для создания дашборда:

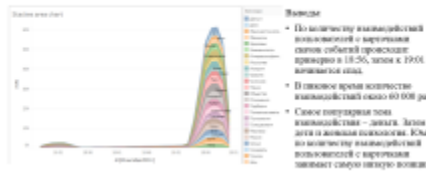
- [illegible]

Макет дашборда

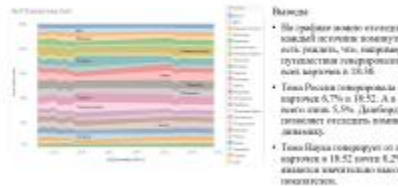
Наименование дисциплины		Виды работ и программы	
Описание дисциплины		Виды работ на курсе	Виды программ на курсе
График лекций и семинаров (с указанием тем лекций) (лекционные материалы, slides and chart)	График работы магистрантов (с указанием по темам, разделам) (work plans, slides and chart)	График проведения семиналов и практических занятий (семинарские задания, практические работы) (chart)	

Таблица, содержащая все материалы, необходимые для выполнения магистерской диссертации, размещены в электронном виде на сайте факультета и публикуются в электронном виде на сайте факультета. На территории факультета также имеется библиотека, содержащая материалы. Работа над диссертацией должна быть завершена в соответствии со сроком, указанным в таблице (publish date).

Количество взаимодействий пользователей с карточками в системе с разбивкой по темам карточек в абсолютных значениях



Как много карточек генерируют источники с разными темами?



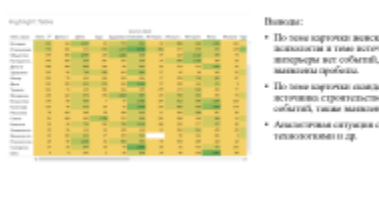
Как соотносятся темы карточек и темы источников?



Как соотносятся темы карточек и темы источников?



Как соотносятся темы карточек и темы источников?



Рекомендации

На основании предоставленной информации для идентификации по адресному телефону даны *links* для просмотра дашборда.

Данный адрес вебсайта предоставляет пользователям возможности по доступу к системе Инициатив. Данная ссылка содержит для просмотра анализа пользовательского взаимодействия с данными из картотеки событий. Как показывает анализ, структура данных для просмотра анализа по каждому из событий имеет следующий вид:

Настоящее число представляется следующим образом: неระบุ, чем-то в процессе.

Построенный дашборд соответствует следующему шаблону: анализ взаимодействия пользователей с картотекой Инициатив. Данный шаблон позволяет пользователям оперативно анализировать данные.

Ссылка на дашборд:

https://public.tableausoftware.com/views/analysis/analysis/analysis/links/links_1874541221490/Default.aspx?&_t=154541221490

Вступление

Описание проекта: Для автоматизации анализа пользовательского взаимодействия с карточками статей Яндекс.Дзен необходимо создать дашборд. 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+».

Задачи исследования: автоматизировать еженедельный процесс анализа пользовательского взаимодействия с карточками статей Яндекс.Дзен при помощи построения дашборда и ответить на поставленные вопросы.

Цель исследования: разработать дашборд, по которому менеджеры самостоятельно смогут более оперативно проводить анализ пользовательского взаимодействия с карточками сетей Яндекс.Дзен. Что значительно сократит трудозатраты действующего аналитика на еженедельный анализ одних и тех же метрик.

Источник данных: дата-инженеры подготовили агрегирующую таблицу `dash_visits`, на основании этих данных был проведен анализ.

Источник данных

Дашборд основывается на пайплайне, который будет брать данные из таблицы, в которых хранятся сырые данные, трансформировать данные и укладывать их в агрегирующую таблицу. Пайплайн разработан дата-инженерами.

Дата-инженеры подготовили агрегирующую таблицу `dash_visits`. Вот её структура:

`record_id` — первичный ключ,

`item_topic` — тема карточки,

`source_topic` — тема источника,

`age_segment` — возрастной сегмент,

`dt` — дата и время,

`visits` — количество событий.

Таблица хранится в специально подготовленной базе данных `zen`.

Общие выводы по дашборду

- В пиковое время количество взаимодействий около 60 000 раз.
- Тема Россия генерировала от всех карточек 6,7% в 18:52. А в 18:57 всего лишь 5,5%. Дашборд позволяет отследить поминутно динамику. Тема Наука генерирует от всех карточек в 18:52 почти 8,2%, что является значительно высоким показателем. Тема Семейные отношения является самой популярной (количество визитов 33309), что составляет 10,74% от общего количества. Количество карточек – 1822. Затем по популярности идет тема Россия (количество визитов 29831), что составляет 9,62% от общего количества. Количество карточек – 1687.
- Карточка Россия в топе по количеству карточек, но если посмотреть в таблицу соотношения, можно увидеть, что топ по перекрёстному, например, не по теме Россия.
- По теме карточки женская психология и теме источника интерьеры нет событий, то есть выявлены пробелы. По теме карточки скандалы и теме источника строительство нет событий, также выявлены пробелы. Аналогичная ситуация с шоу и технологиями и др.

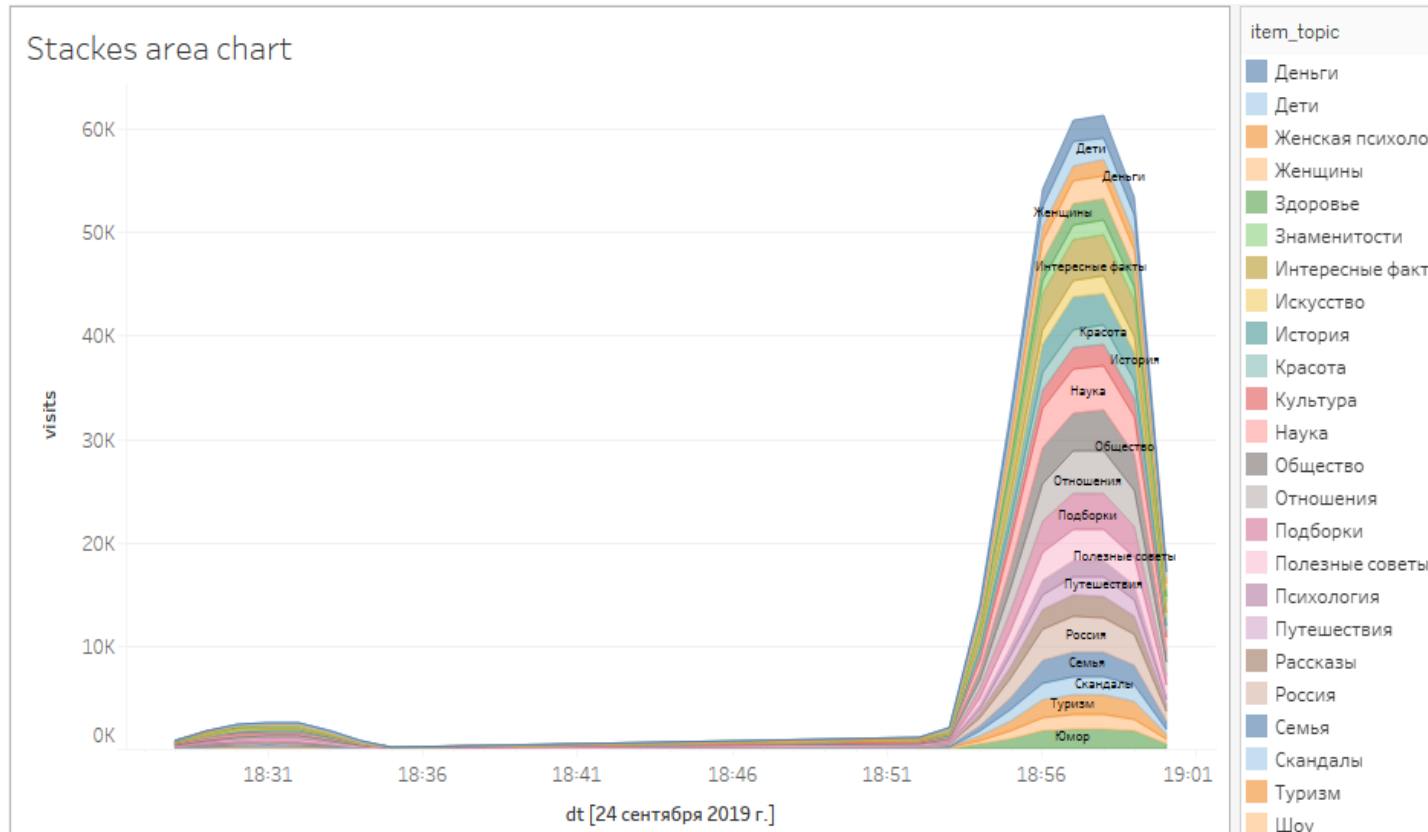
Краткое техническое задание (ТЗ) для создания дашборда:

- 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;
 - Важность: все графики имеют равную важность; 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).			

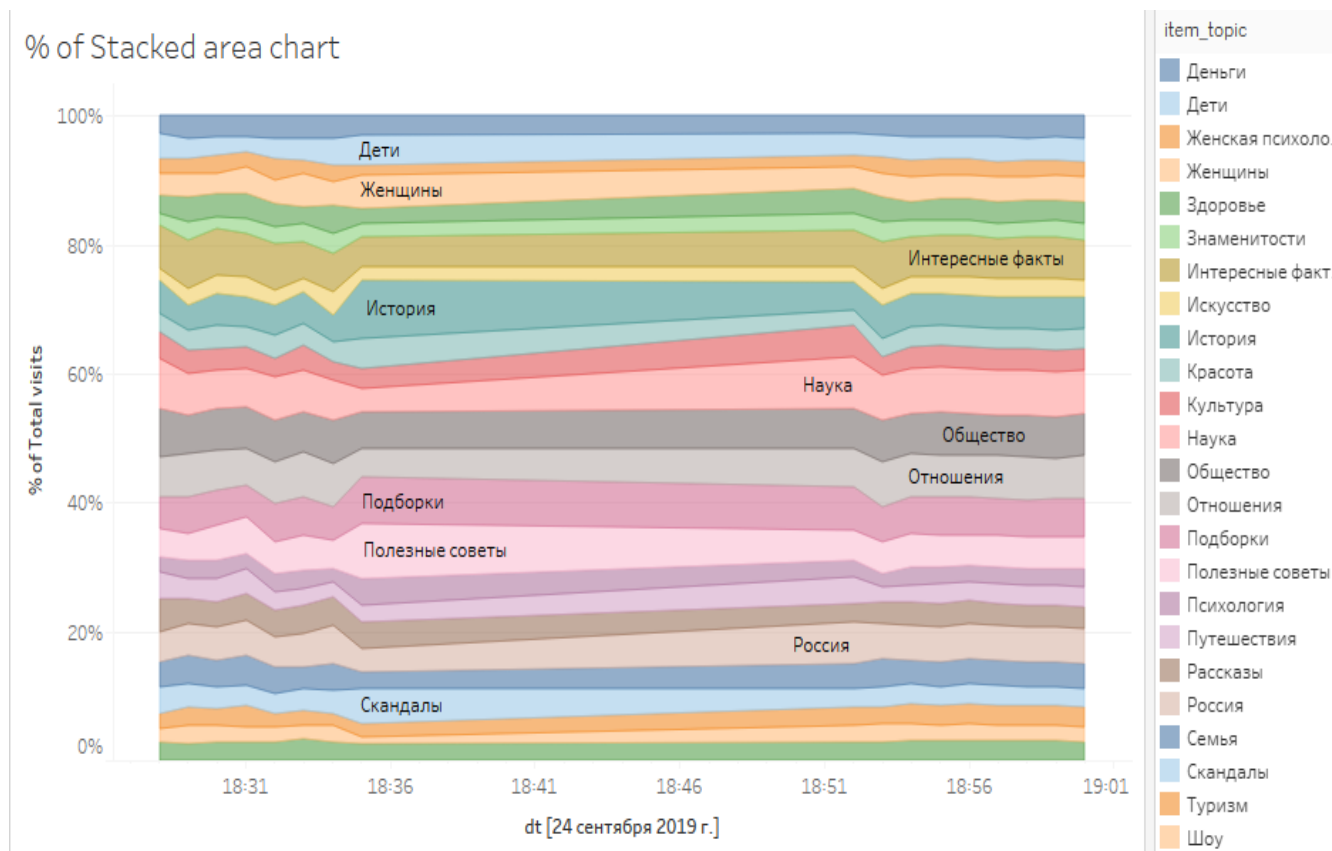
Количество взаимодействий пользователей с карточками в системе с разбивкой по темам карточек в абсолютных значениях



Выводы:

- По количеству взаимодействий пользователей с карточками скачок событий происходит примерно в 18:56, затем к 19:01 начинается спад.
- В пиковое время количество взаимодействий около 60 000 раз.
- Самая популярная тема взаимодействия — деньги. Затем дети и женская психология. Юмор по количеству взаимодействий пользователей с карточками занимает самую низкую позицию.

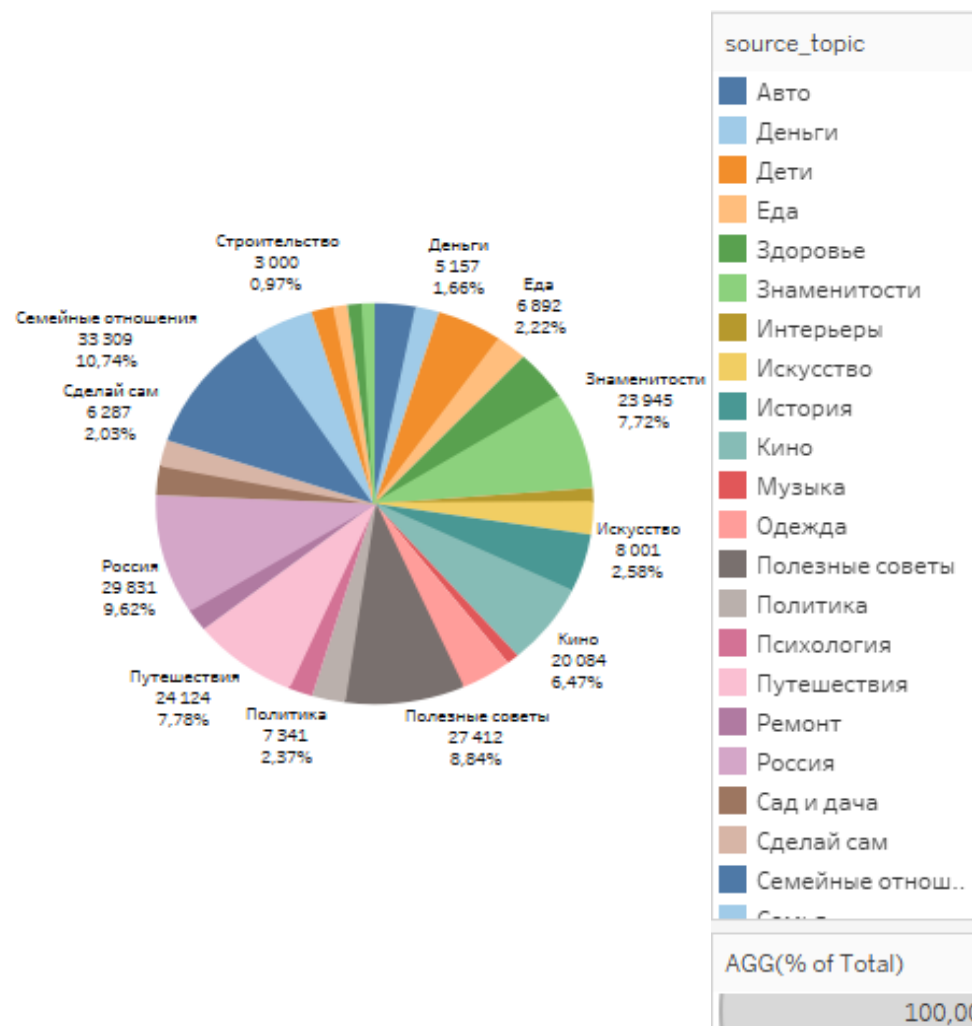
Как много карточек генерируют источники с разными темами?



Выводы:

- На графике можно отследить каждый источник поминутно, то есть увидеть, что, например, путешествия генерировали 3,7% от всех карточек в 18:30.
- Тема Россия генерировала от всех карточек 6,7% в 18:52. А в 18:57 всего лишь 5,5%. Дашборд позволяет отследить поминутно динамику.
- Тема Наука генерирует от всех карточек в 18:52 почти 8,2%, что является значительно высоким показателем.

Как соотносятся темы карточек и темы источников?



Выводы:

- Тема Семейные отношения является самой популярной (количество визитов 33309), что составляет 10,74% от общего количества. Количество карточек – 1822.
- Затем по популярности идет тема Россия (количество визитов 29831), что составляет 9,62% от общего количества. Количество карточек – 1687.
- Количество карточек по теме Финансы наименьшее (примерно 650).
- Посмотрим сопоставление темы карточки с темой источников по таблице соответствия.

Как соотносятся темы карточек и темы источников?

Highlight Table

item_topic	source_topic											
	Психологические	Путешествия	Ремонт	Россия	Сад и дача	Сделай сам	Семейные	Семья	Спорт	Строительство	Технологии	Финансы
Россия	38	1079	345	2847	85	114	1962	608	703	179	316	
Наука	291	1823	415	1782	227	358	1711	809	421	99	749	
Интересные факты	449	1448	175	2567	825	341	1323	611	186	234	219	
Полезные советы	438	2088	485	403	253	679	1781	811	136	77	294	
Подборки	396	1491	479	877	522	364	1386	424	280	361	440	
История	227	2643	168	1491	209	149	1344	696	190	9	286	
Отношения	197	1269	510	2278	944	517	1030	194	312	513	119	
Общество	919	891	181	3471	76	162	2727	1416	338	97	118	
Путешествия	126	641	451	588	499	127	709	306	88	85	44	
Деньги	305	536	362	1132	101	137	696	345	134	68	88	
Здоровье	214	393	145	712	542	548	814	247	197	106	74	
Юмор	140	636	201	621	321	204	1887	584	281	55	11	
Дети	128	431	174	547	633	354	1184	558	21	103	28	
Туризм	133	483	403	867	523	688	1580	283	79	74	38	
Женщины	253	192	67	397	125	223	2270	988	210	56	6	
Искусство	98	664	172	632	43	128	754	404	54	53	58	
Культура	81	430	154	1474	100	76	892	482	47	86	46	
Рассказы	233	4587	123	421	201	92	875	539	222	74	1	

Выводы:

- На слайде представлен фрагмент таблицы по соотношению тем карточек с темами источников из дашборда.
- Карточка Россия в топе по количеству карточек, но если посмотреть в таблицу соотношения, можно увидеть, что топ по перекрёстному, например, не по теме Россия.

Как соотносятся темы карточек и темы источников?

Highlight Table

item_topic	source_topic											
	Авто	Деньги	Дети	Еда	Здоровье	Знамени..	Интерье..	Искусст..	История	Кино	Музыка	Оде
История	522	114	1047	60	771	921	30	695	636	1091	164	
Отношения	478	281	242	573	1173	2040	862	374	679	672	139	
Общество	379	350	1007	160	1181	829	64	218	1116	578	79	
Путешеств..	343	363	509	357	187	509	26	555	1018	289	33	
Деньги	338	203	859	193	65	591	62	172	271	1042	42	
Здоровье	232	95	798	558	404	962	57	64	68	156	20	
Юмор	224	79	404	194	124	334	17	205	715	502	67	
Дети	192	228	621	675	346	1444	97	238	455	474	9	
Туризм	161	74	406	382	311	269	279	173	618	252	77	
Женщины	156	122	875	348	1022	967	175	485	560	340	45	
Искусство	145	98	639	7	70	1083	164	512	596	1254	223	
Культура	103	46	578	55	41	1508	124	682	636	1543	178	
Рассказы	78	454	286	28	248	615	47	291	568	357	36	
Семья	50	458	361	1236	275	689	321	288	266	486	41	
Красота	29	48	703	253	736	1313	185	232	177	472	38	
Знаменито..	29	55	145	36	139	445	12	356	644	408	55	
Женская пс..	26	256	613	67	349	546		96	181	264	5	
Психология	19	86	1233	62	583	582	46	306	286	219	19	
Скандалы	14	35	844	48	43	1992	93	91	410	707	132	
Шоу	4	8	297	1	20	829	30	66	215	2201	98	

Выводы:

- По теме карточки женская психология и теме источника интерьеры нет событий, то есть выявлены пробелы.
- По теме карточки скандалы и теме источника строительство нет событий, также выявлены пробелы.
- Аналогичная ситуация с шоу и технологиями и др.

Рекомендации

На основании предоставленной информации дата-инженерами по агрегирующей таблице dash_visits был построен дашборд.

Данный дашборд необходимо использовать менеджерам по анализу контента Яндекс.Дзен самостоятельно для проведения анализа пользовательского взаимодействия с карточками сетей. Что значительно сократит трудозатраты действующего аналитика на еженедельный анализ одних и тех же метрик.

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

Построенный дашборд соответствует бизнес задаче: анализ взаимодействия пользователей с карточками Яндекс.Дзен и полностью позволит менеджерам оперативно анализировать данные.

Ссылка на дашборд:

https://public.tableau.com/app/profile/zarinakhabarova/viz/dash_visits_16745122141690/Dashboard?publish=yes