

Анализ пользовательского взаимодействия с карточками статей в Яндекс.Дзен

(по данным о событиях взаимодействия пользователей с карточками за 24 сентября 2019 года с 18:28 по 19:00)

Содержание

• Описание проекта Яндекс.Дзен	3
• Цели исследования	4
• Исходные данные	5
• Макет спроектированного дашборда	6
• Описание и функционал дашборда	7
• Соотношение тем карточек и тем источников	8
• Как много карточек генерируют источники с разными темами?	9
• Графики истории взаимодействия	10
• Выводы	11
• Контакты и ссылки	14

Описание проекта Яндекс.Дзен

Ваша лента

Что вам интересно?

Выбирайте, что нравится, и Дзен подстроится под вас

Выберите хотя бы одну тему



Природа
и наука



Музыка



Спорт



Еда и
напитки



Техно-
логии



Отно-
ни!



Путе-
шества



Кино



Домашние
животные



История
и МХК



Дом
и сад



На Яндекс.Дзен есть статьи (карточки). Из одной статьи можно перейти в другую. При переходе карточка, с которой перешли, становится источником.

Под событиями понимаются все виды взаимодействия с карточками:

- карточка отображена для пользователя (show);
- пользователь кликнул на карточку (click);
- пользователь просмотрел статью карточки (view).

Цели исследования

Целью данного исследования является построение дашборда, автоматизирующего процесс решения задачи ответов на следующие вопросы, касающиеся анализа пользовательского взаимодействия с карточками статей Яндекс.Дзен:

- 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?

Исходные данные

```
7]: display(dash_visits.head())  
print(dash_visits.info())
```

	record_id	item_topic	source_topic	age_segment	dt	visits
0	1040597	Деньги	Авто	18-25	2019-09-24 18:32:00	3
1	1040598	Деньги	Авто	18-25	2019-09-24 18:35:00	1
2	1040599	Деньги	Авто	18-25	2019-09-24 18:54:00	4
3	1040600	Деньги	Авто	18-25	2019-09-24 18:55:00	17
4	1040601	Деньги	Авто	18-25	2019-09-24 18:56:00	27

```
<class 'pandas.core.frame.DataFrame'>  
RangeIndex: 30745 entries, 0 to 30744  
Data columns (total 6 columns):  
#   Column          Non-Null Count  Dtype  
---  -  
0   record_id       30745 non-null  int64  
1   item_topic      30745 non-null  object  
2   source_topic    30745 non-null  object  
3   age_segment     30745 non-null  object  
4   dt              30745 non-null  datetime64[ns]  
5   visits          30745 non-null  int64
```

- Дата проведения исследования: 07/05/2022.
- Пайплайн с исходными данными для анализа разработан и предоставлен дата-инженерами Яндекс.
- Пайплайн берет данные из таблицы, в которых хранятся сырые данные о событиях взаимодействия пользователей с карточками, трансформирует их и формирует агрегирующую таблицу.
- Анализируемый временной период: 24/09/2019 с 18:28 по 19:00.
- Всего 30745 записей.

Макет спроектированного дашборда

Анализ взаимодействия пользователей с карточками Яндекс.Дзен

Графики описывают анализ взаимодействия пользователей с карточками статей. Каждую карточку определяют её тема и источник (у которого тоже есть тема). Под событиями понимаются все виды взаимодействия с карточками.

Дата/время 18:28:00 19:00:00

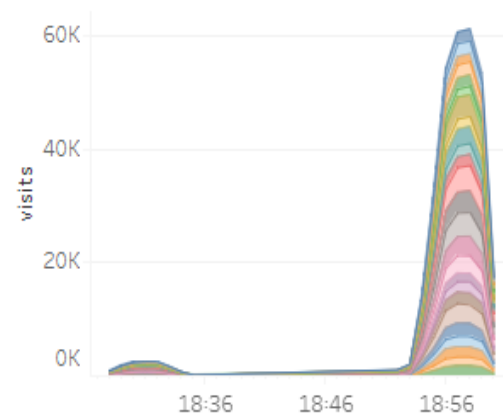
Выбор темы карточек

(All)

Возрастная категория

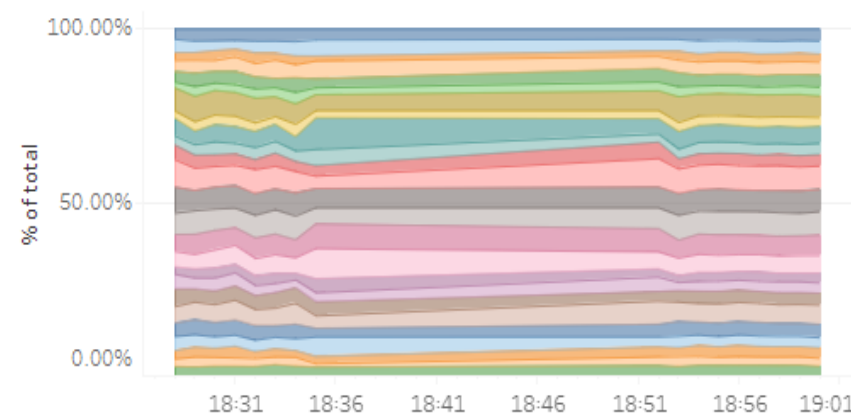
(All)

События по темам карточек



Дата / время

% событий по темам карточек



Дата / время

События по темам источников



Темы источников - темы карточек

item_topic	Путе..	Россия	Семейн..	Знамени..	История	Семья	Деньги	Полезны..	Кино	Искусст..	Дети	Здоровье	Психоло..	Спорт	Сад и да..	Ремонт	Сделай	Финансы	Авто	Строите..	По.
Рассказы	4,587	421	875	615	568	539	454	358	357	291	286	248	233	222	201	123	92	82	78	74	
История	2,643	1,491	1,344	921	636	696	114	859	1,091	695	1,047	771	227	190	209	168	149	62	522	9	
П	2,000	400	1,500	200	400	200	200	1,400	1,400	200	700	1,300	400	400	200	400	200	200	700	77	

Описание и функционал дашборда

- Состав и характер данных для дашборда:
 - История событий по темам карточек (два графика с накоплением - абсолютные величины с разбивкой по минутам и процентное соотношение);
 - Разбивка событий по темам источников - относительные величины (% событий);
 - Таблица соответствия тем источников темам карточек - абсолютные величины;
- Возможность группировки данных по параметрам:
 - Дата и время;
 - Тема карточки;
 - Тема источника;
 - Возрастная группа;
- Частота обновления данных: один раз в сутки, в полночь по UTC.

Соотношение тем карточек и тем источников

item_topic	Путешествия	Россия	Семейные отношения
Рассказы	4 587	421	875
История	2 643	1 491	1 344
Полезные советы	2 088	403	1 781
Наука	1 823	1 782	1 711
Подборки	1 491	877	1 386
Интересные факты	1 448	2 567	1 323
Отношения	1 269	2 278	1 030
Россия	1 079	2 847	1 962
Общество	891	3 471	2 727
Красота	781	269	748
Искусство	664	632	754
Путешествия	641	588	709
Юмор	636	621	1 887
Деньги	536	1 132	696
Туризм	483	867	1 580
Дети	431	547	1 184
Культура	430	1 474	892
Здоровье	393	712	814

Наиболее частые переходы (> 2000) по убыванию:

- Рассказы -> Путешествия (4587)
- Общество -> Россия (3471)
- Наука -> Кино (3279)
- Россия -> Россия (2847)
- Подборки -> Полезные советы (2795)
- Общество -> Семейные отношения (2727)
- Отношения -> Полезные советы (2716)
- История -> Путешествия (2643)
- Интересные факты -> Россия (2567)
- Отношения -> Россия (2278)
- Знаменитости -> Россия (2275)
- Женщины -> Семейные отношения (2270)
- Шоу -> Кино (2201)
- Интересные факты -> Здоровье (2090)
- Полезные советы -> Путешествия (2088)
- Женская психология -> Семейные отношения (2073)
- Отношения -> Знаменитости (2040)
- Интересные факты -> Полезные советы (2023)

Как много карточек генерируют источники с разными темами?

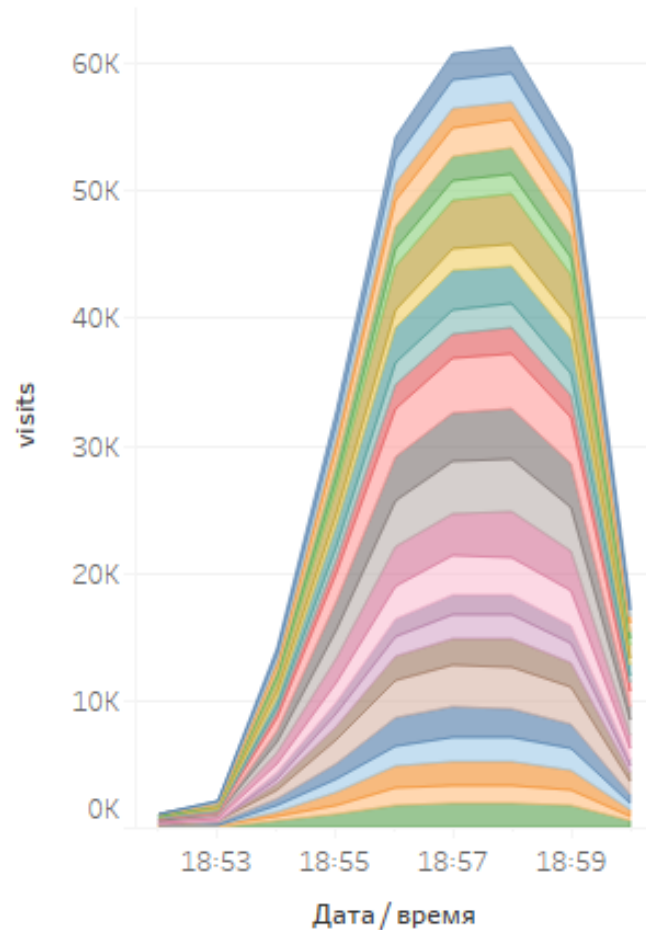


Самые частые источники
(по убыванию популярности):

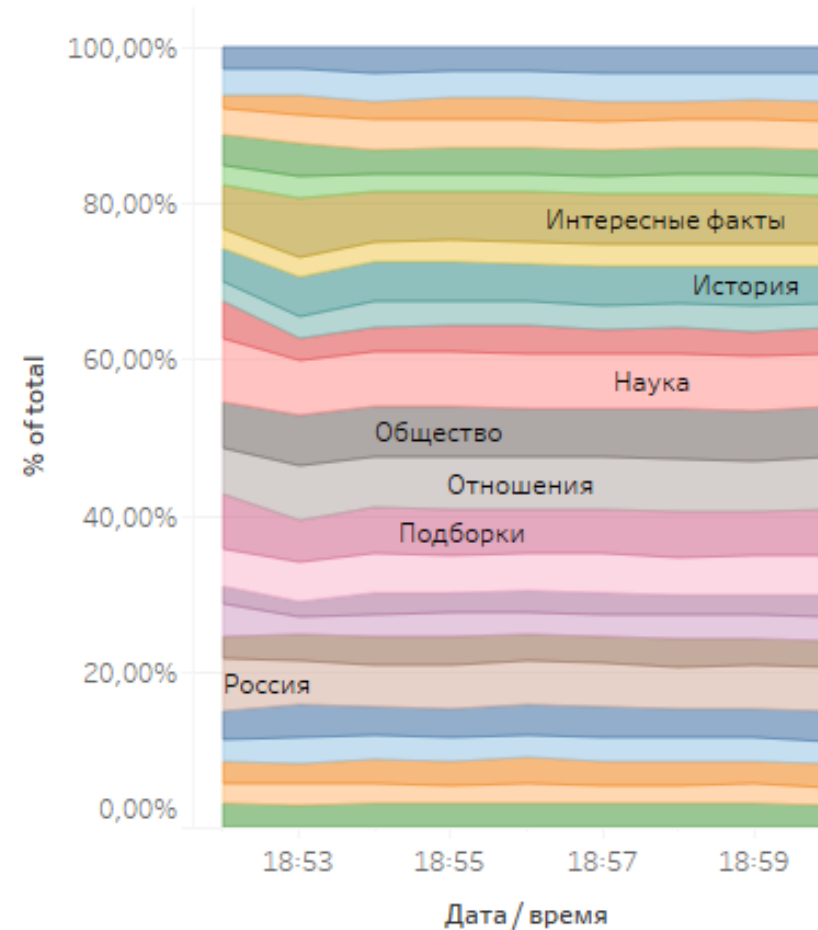
- Семейные отношения (10,7%)
- Россия (9,6%)
- Полезные советы (8,8%)
- Путешествия (7,8%)
- Знаменитости (7,7%)
- Кино (6,5%)
- История (4,7%)
- Дети (4,9%)
- Семья (4,5%)
- Здоровье (4,1%)

Графики истории взаимодействия

События по темам карточек



% событий по темам карточек



Источники, которые генерируют большее количество визитов:

- Интересные факты
- История
- Наука
- Общество
- Отношения
- Подборки
- Россия

Выводы

Топ-10 наиболее частых соотношений тем источников и карточек
для исследуемого промежутка времени 24/09/2019 с 18:28 по 19:00:

- Рассказы -> Путешествия (4587)
- Общество -> Россия (3471)
- Наука -> Кино (3279)
- Россия -> Россия (2847)
- Подборки -> Полезные советы (2795)
- Общество -> Семейные отношения (2727)
- Отношения -> Полезные советы (2716)
- История -> Путешествия (2643)
- Интересные факты -> Россия (2567)
- Отношения -> Россия (2278)

Выводы

Топ-10 наиболее частых тем источников

для исследуемого промежутка времени 24/09/2019 с 18:28 по 19:00:

- Семейные отношения (10,7%)
- Россия (9,6%)
- Полезные советы (8,8%)
- Путешествия (7,8%)
- Знаменитости (7,7%)
- Кино (6,5%)
- История (4,7%)
- Дети (4,9%)
- Семья (4,5%)
- Здоровье (4,1%)

Выводы

Источники, сгенерировавшие большее количество визитов
24/09/2019 с 18:28 по 19:00:

- Интересные факты
- История
- Наука
- Общество
- Отношения
- Подборки
- Россия

Контакты и ссылки

Дашборд можно просмотреть по ссылке:

[KotsNS_dash](#) (created on: 10/05/2022 3:37:23 PM)

Дата проведения исследования: 07/05/2022

Презентацию подготовила
студентка потока da_37
Котс Наталия

email: kotsnat@gmail.com
+7-915-433-33-39