

АВТОМАТИЗАЦИЯ. ДАШБОРД ДЛЯ ЯНДЕКС.ДЗЕН

Выполнила: Мирошниченко Юлия

Тел. 8(910)4290993

Дата: 11.02.2023

ОПИСАНИЕ ПРОЕКТА

Вы работаете аналитиком в Яндекс.Дзене. Почти всё ваше время занимает анализ пользовательского взаимодействия с карточками статей.

Есть три способа взаимодействия пользователей с системой:

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

Каждую неделю начинающие менеджеры задают вам 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 вы решаете, что процесс пора автоматизировать. Для Дениса и Валерии нужно сделать дашборд.

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

ОПИСАНИЕ ДАННЫХ

Источники данных для дашборда: хранятся в агрегирующей таблице dash_visits.

Структура таблицы dash_visits :

- **record_id** — первичный ключ,
- **item_topic** — тема карточки,
- **source_topic** — тема источник,
- **age_segment** — возрастной сегмент,
- **dt** — дата и время,
- **visits** — количество событий.

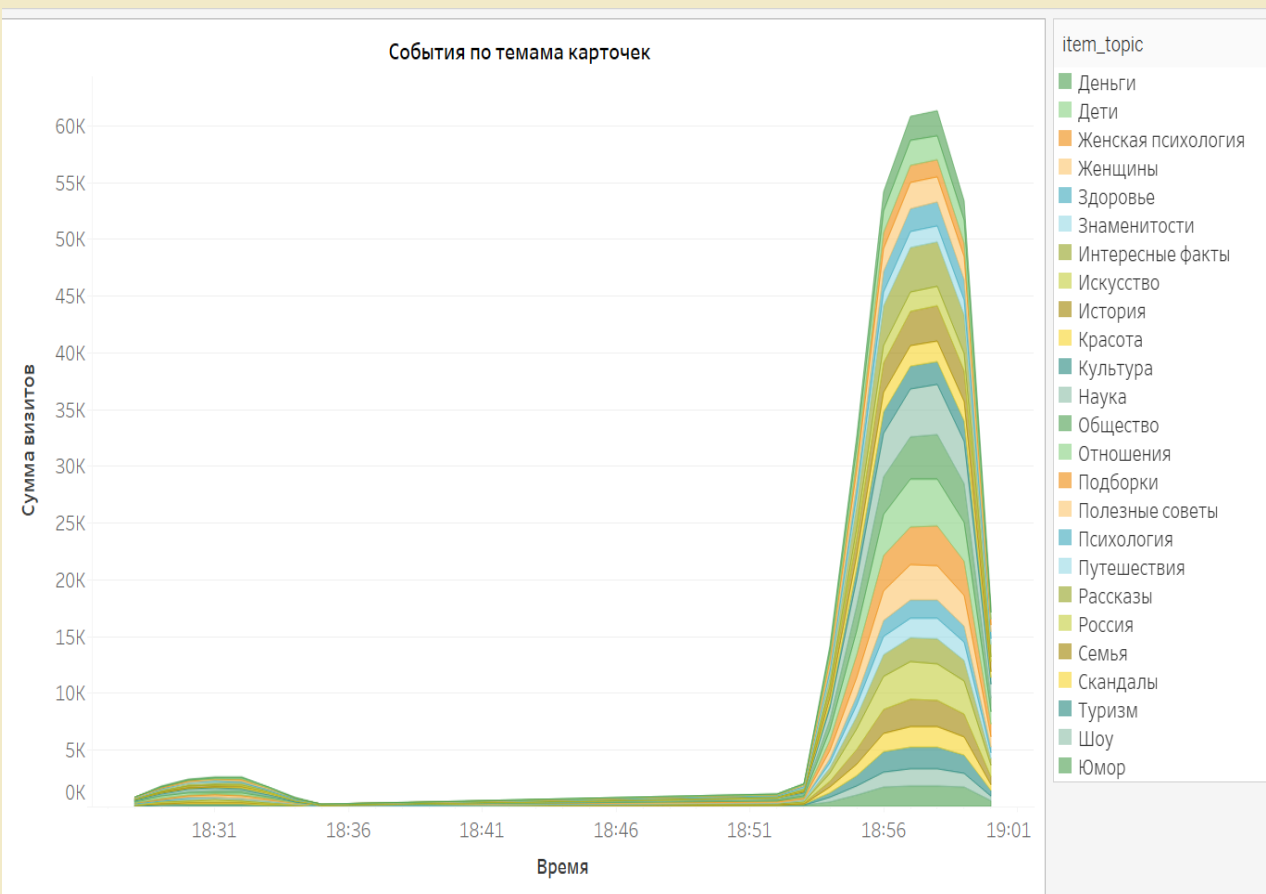
Частота обновления данных: один раз в сутки, в полночь по UTC;

Исследуемый период данных: с 24-09-2019 18:28:00 по 24-09-2019 19:00:00

СТРУКТУРА ДАШБОРДА

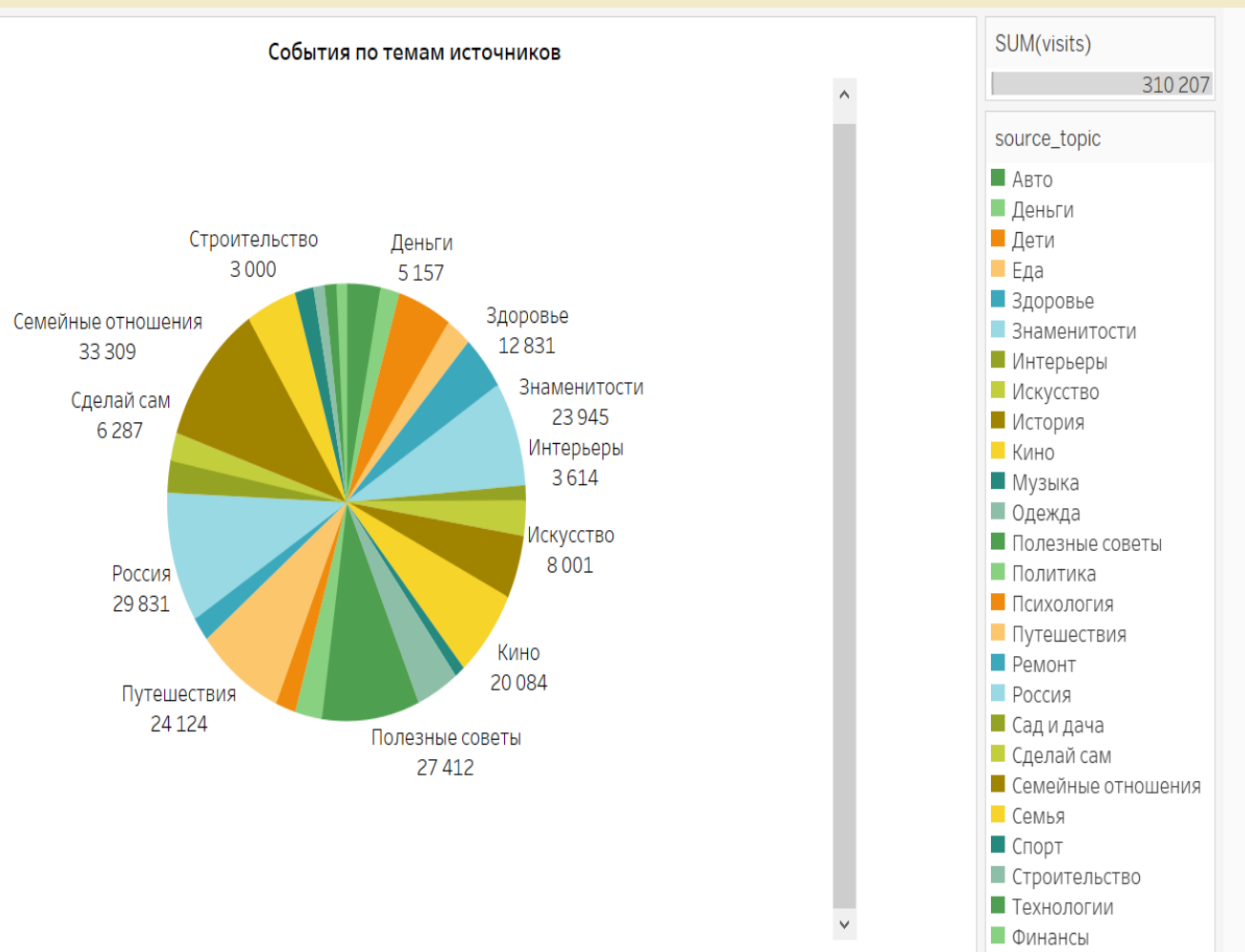
Название дашборда		Фильтр по дате и времени	
Описание дашборда		Фильтр по темам карточек	Фильтр по возрастным категориям
График истории взаимодействия «События по темам карточек» (абсолютные значения, stacked area chart)	График истории взаимодействия «% событий по темам карточек» (% от общего, stacked 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).			

СКОЛЬКО ВЗАИМОДЕЙСТВИЙ ПОЛЬЗОВАТЕЛЕЙ С КАРТОЧКАМИ ПРОИСХОДИТ В СИСТЕМЕ С РАЗБИВКОЙ ПО ТЕМАМ КАРТОЧЕК?



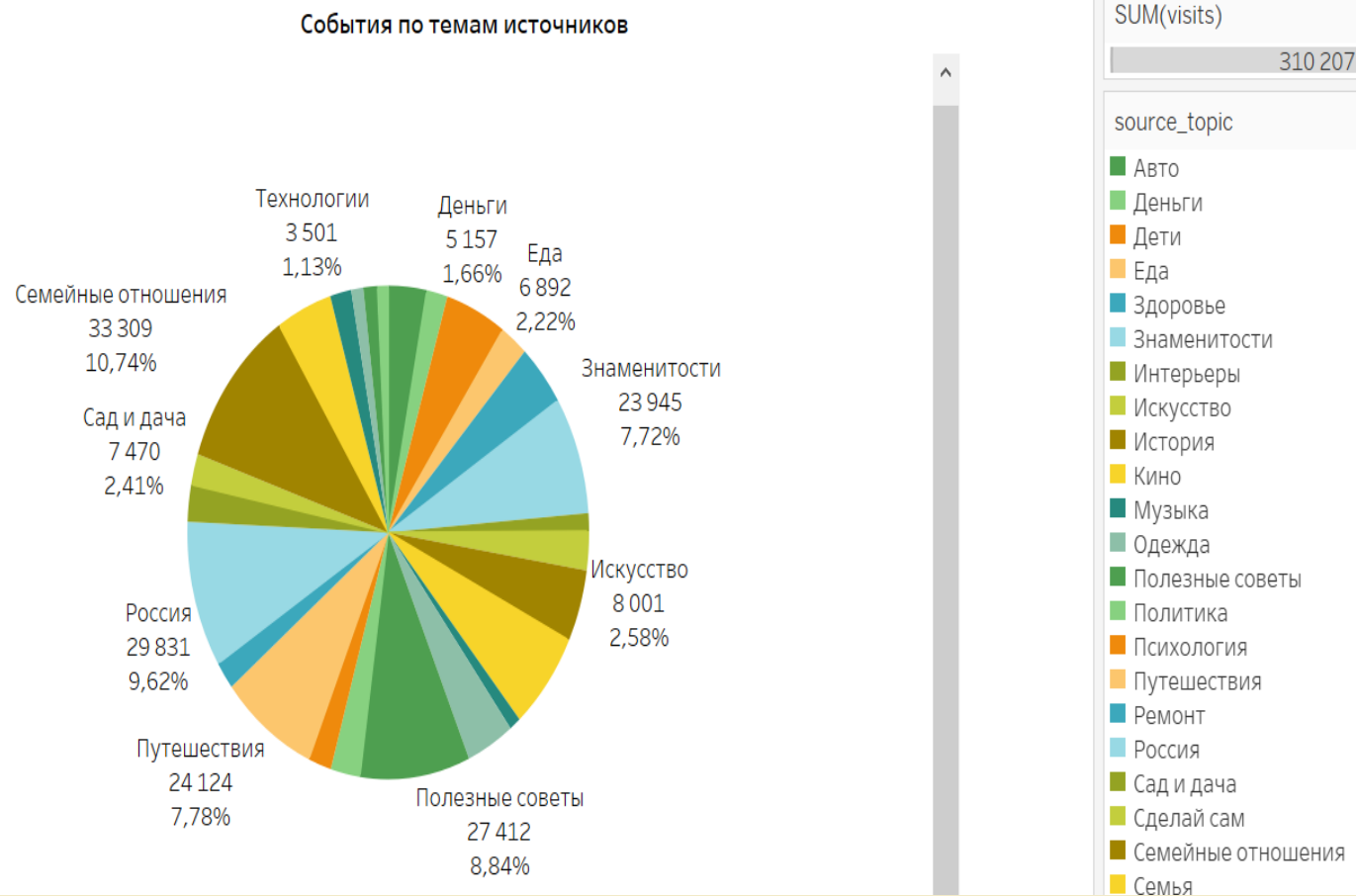
- На пике находятся свыше 60 тысяч событий. Наибольшее число событий приходится на тему карточки – Наука.(4372 события в период времени 18:58:00).
- Также в топ входят следующие темы: Отношения(4145 в период времени события 18:58:00), Интересные (3910 события в период времени 18:58:00), Общество(3897 событий в период времени 18:58:00).
- Меньше всего событий на следующие темы: Знаменитости(1463 события в период времени 18:58:00), Шоу (1459 событий в период времени 18:58:00), Женская психология (1518 событий в период времени 18:58:00).
- Другие темы делят от 1.5-3.5 тысяч событий.

КАК СООТНОСЯТСЯ ТЕМЫ КАРТОЧЕК И ТЕМЫ ИСТОЧНИКОВ?



- Больше всего карточек на тему источника: Россия, Семейные отношения, Знаменитости, Полезные советы, Кино, Путешествия.
- Значительно меньше карточек приходится на следующие темы источника: Строительство, Деньги, Сделай сам, Здоровье.

КАК МНОГО КАРТОЧЕК ГЕНЕРИРУЮТ ИСТОЧНИКИ С РАЗНЫМИ ТЕМАМИ?



- Значительную часть диаграммы занимают: Россия, Семейные отношения, Путешествия, Полезные советы, Знаменитости.
- Остальные темы примерно до 5% диаграммы занимают.

ВЫВОД

В ходе проекта нами был построен дашборд, отвечающие на следующие вопросы бизнеса:

- 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?

Исследуемый период был взят с 24-09-2019 18:28:00 по 24-09-2019 19:00:00.

В период с 18:53: 00 наблюдается резкий рост событий и продолжается данный рост до 18:58: 00.

В период 18:58: 00 зафиксированы максимальные количества событий.

В результате исследования были выявлены самые популярные темы источников: Семейные отношения(10.74%), Россия (9.62%), Полезные советы(8.84%), Путешествия(7.78%), Знаменитости(7.72%), Кино(6.47%). Остальные темы источников занимают до 5%.

ССЫЛКА НА ДАШБОРД:

https://public.tableau.com/app/profile/.83294926/viz/Dashbord_16761295599110/sheet4?publish=yes