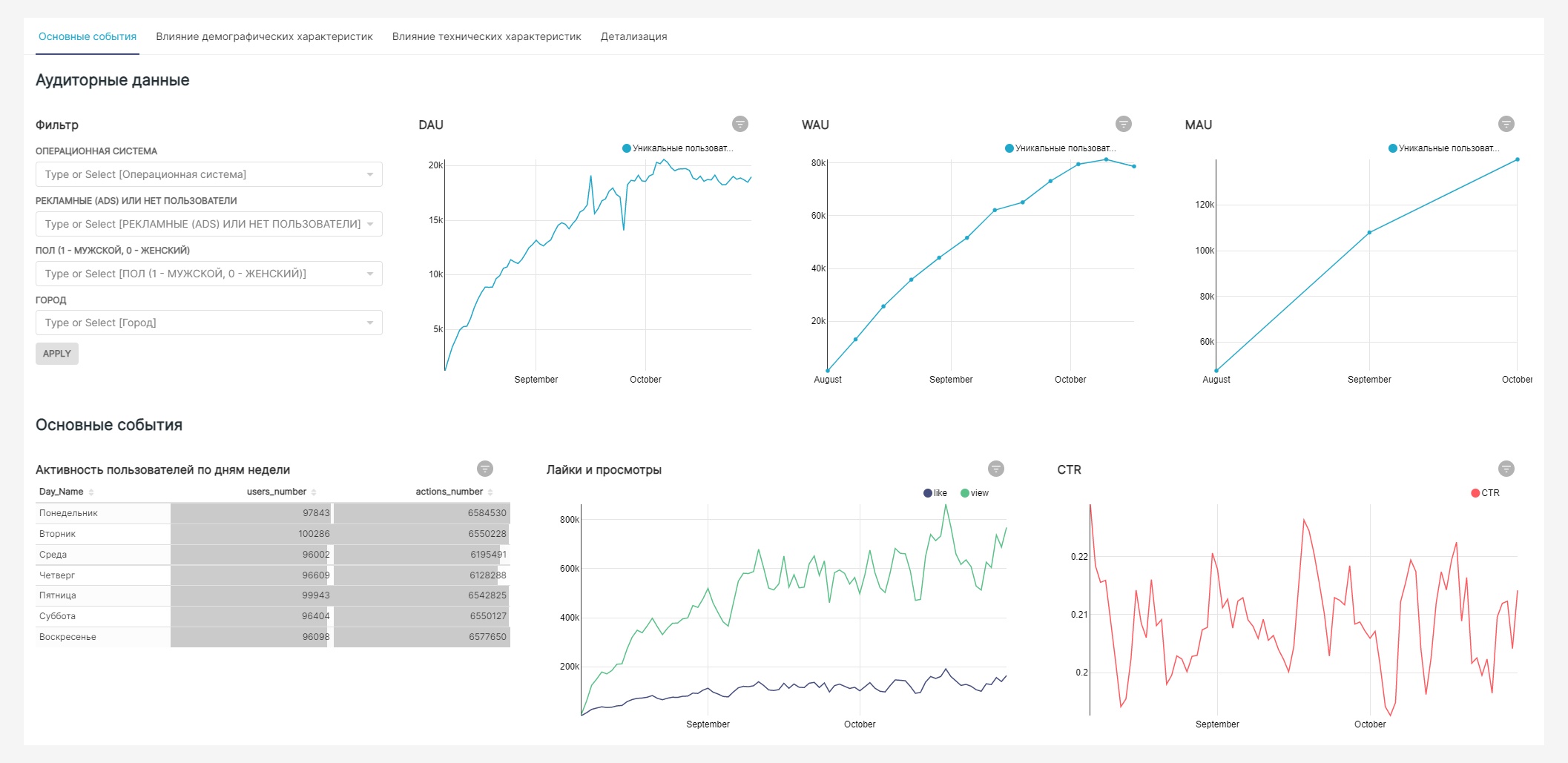
|  |  |
| --- | --- |
| Задание 1.1 | SQL  (Предварительное формирование входного набора данных в Redash) |
| Разработайте дизайн базовых дашбордов по ретроспективным данным, которые покрывают собой ключевые аудиторные метрики и ключевые события нашего продукта. | Группировка данных для вкладки «Влияние демографических характеристик»  SELECT toStartOfDay(toDateTime(time)) AS timestamp,  count(DISTINCT user\_id) AS user\_unique,  count(user\_id) AS count\_action,  country\_gr,  age\_gr,  gender\_label,  source,  os,  action  FROM  (SELECT country,  gender,  age,  action,  os,  source,  time,  user\_id,  multiIf(age < 20, '0-20', age >= 20 and age < 30,'21-30', age >= 30 and age < 40, '31-40', '40+')  AS age\_gr,  multiIf(country LIKE 'Russia', 'Russia', country LIKE 'Kazakhstan','Kazakhstan',  country LIKE 'Belarus', 'Belarus', country LIKE 'Ukrain' , 'Ukrain', 'Other')  AS country\_gr,  multiIf(gender =1, 'муж', 'жен') AS gender\_label  FROM simulator\_20220920.feed\_actions  )  GROUP BY timestamp,  country\_gr,  age\_gr,  gender\_label,  source,  os,  action |
| Группировка данных для вкладки «Основные события», для графика «Активность пользователей по дням недели»  SELECT toDayOfWeek(time) as Day,  count(user\_id)  FROM  simulator\_20220920.feed\_actions  GROUP BY  toDayOfWeek(time) |

Вкладка «Основные события»



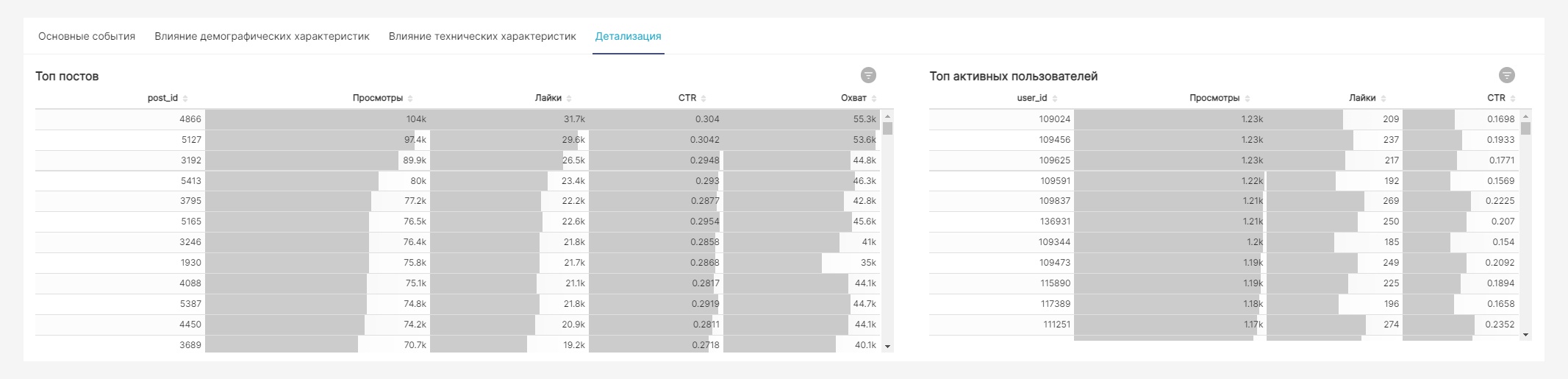
Вкладка «Влияние демографических характеристик»



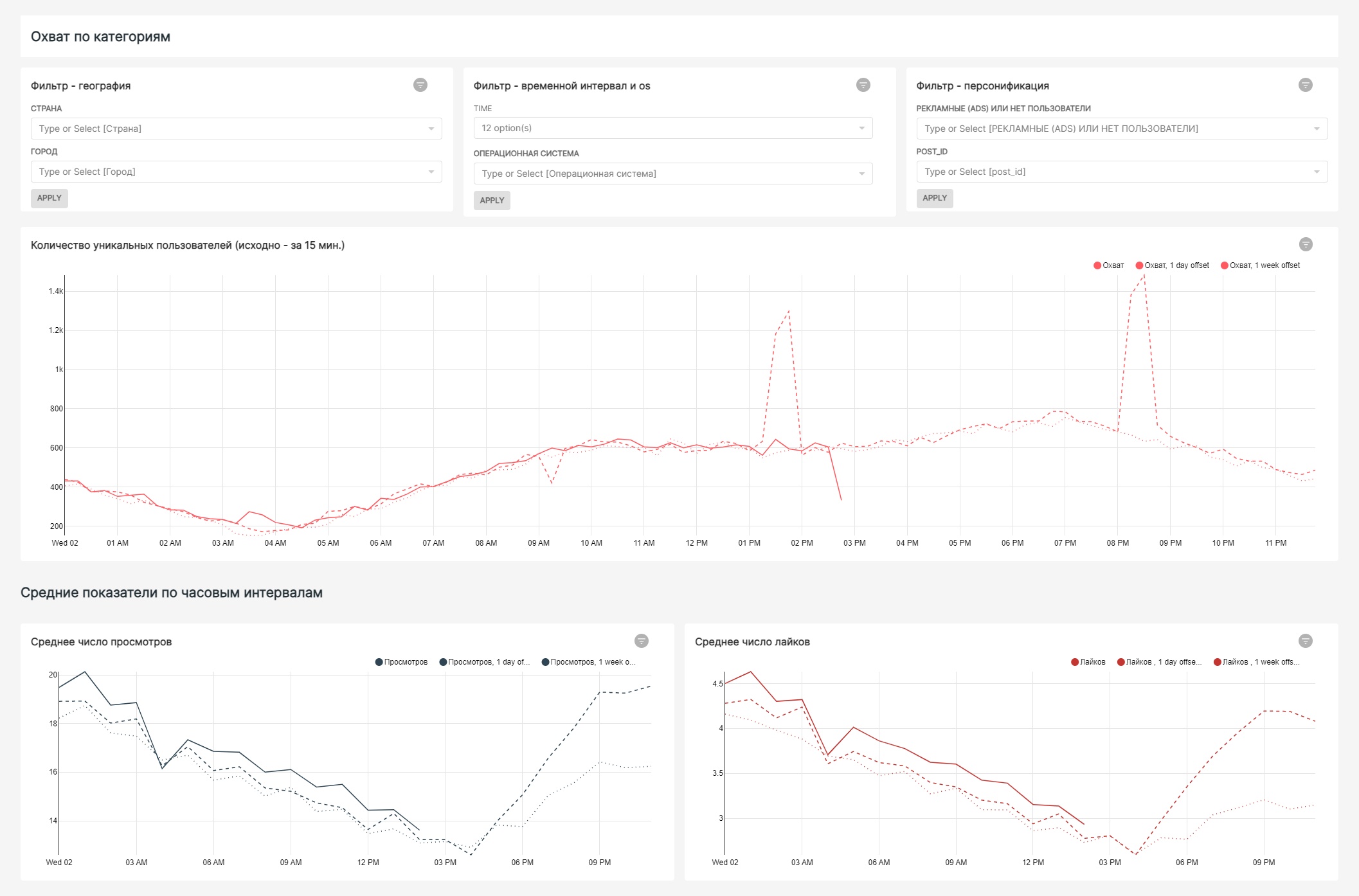
Вкладка «Влияние технических характеристик»



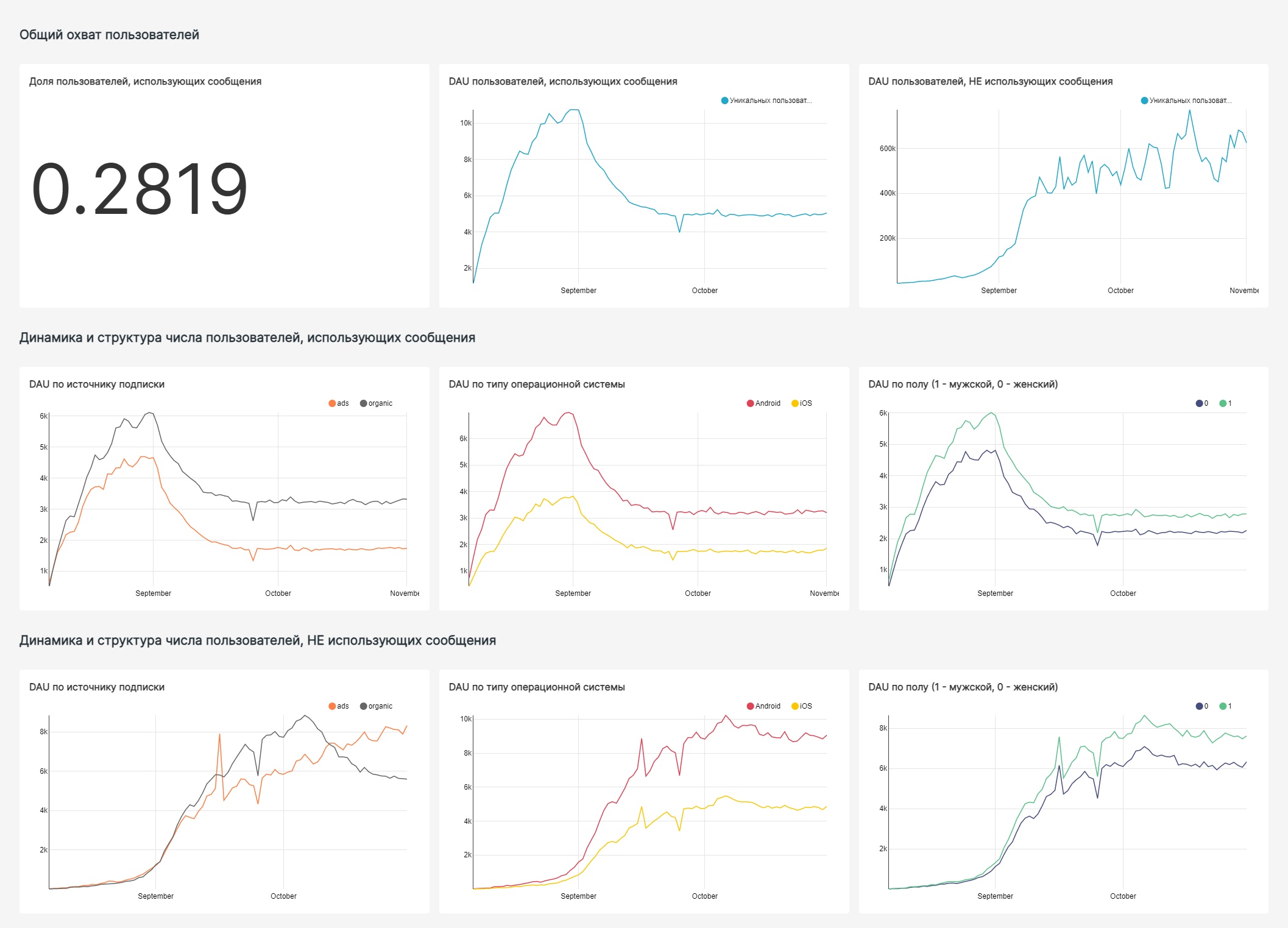
Вкладка «Детализация»



|  |  |
| --- | --- |
| Задание 1.2 | SQL  (Предварительное формирование входного набора данных в Redash) |
| Разработайте дизайн базового дашборда по оперативным данным за сегодняшний день. | SELECT \*  FROM simulator\_20220920.feed\_actions |



|  |  |
| --- | --- |
| Задание 2 | SQL  (Предварительное формирование входного набора данных в Redash) |
| Второй задачей будет составить дашборд, который описывает взаимодействие двух сервисов — ленты и сообщений.  Постарайтесь при помощи дашборда ответить на следующие вопросы:  1. Какая у нашего приложения активная аудитория по дням, т.е. пользователи, которые пользуются и лентой новостей, и сервисом сообщений.  2. Сколько пользователей использует только ленту новостей и не пользуются сообщениями.  Какие еще вопросы вы бы хотели исследовать о связи двух продуктов? Добавляйте графики и описания таких вопросов к себе на дашборд. | Формирование датасета для фактоида «Доля пользователей»  SELECT user\_id,  MAX(message) AS mess  FROM  (SELECT user\_id,  'yes' as message  FROM  (SELECT user\_id  FROM simulator\_20220920.feed\_actions  ) t1  JOIN  (SELECT user\_id  FROM simulator\_20220920.message\_actions  ) t2  USING user\_id  GROUP BY user\_id  UNION ALL  SELECT user\_id,  'no' as message  FROM  (SELECT user\_id  FROM simulator\_20220920.feed\_actions  WHERE user\_id NOT IN  (SELECT user\_id  FROM simulator\_20220920.message\_actions  )  )  )  GROUP BY user\_id |
| Формирование датасета для графиков по пользователям, использующим сообщения  SELECT \*  FROM  (SELECT user\_id,  time, source,  os,  gender  FROM  simulator\_20220920.feed\_actions  ) t1  JOIN  (SELECT  user\_id  FROM  simulator\_20220920.message\_actions  ) t2  USING user\_id |
| Формирование датасета для графиков по пользователям, не использующим сообщения  SELECT  user\_id,  time,  source,  os,  gender  FROM  simulator\_20220920.feed\_actions  WHERE user\_id NOT IN  (SELECT  user\_id  FROM  simulator\_20220920.message\_actions  ) |



Сформированные выводы (были получены на момент 7октября, без внесенных еще организаторами курсов резких колебаний в показателях, рисунки сохранены позднее):

1) общая доля людей, использующих отправку сообщений, составляет пятую часть от всего числа пользователей. При этом динамика показывает, что в начале функционирования приложения практически все пользователи использовали сообщения, но уже к сентябрю количество пользователей, использующих отправку сообщений, упало практически в два раза, оставаясь далее на достаточно стабильном уровне.

2) Примерно в этот же момент резко возрастает количество пользователей, не использующих отправку сообщений (приближаясь в итоге как раз к отношению 1:5). Возможно, в начале сентября произошли какие-то изменения в условиях использования приложения, или каким-то образом изменился контент, что привело к потере интереса к обмену сообщениями. Возможно, конечно, это связано и просто с падением ажиотажного интереса наиболее активной части пользователей к новому продукту. И приходу им на смену менее активной аудитории. Но до появления какой-либо дополнительной информации это все остается на уровне предположений и гипотез.

3) По итогам анализа уже имеющейся информации: не было обнаружено влияния пола пользователя и типа операционной системы и типа подписки (рекламная или нет) на то, использует пользователь сообщения или нет. Также не обнаружено и влияния этих факторов на динамику охвата пользователей разных типов.