

# Аналитика коммуникаций в чатах Яндекс.Практикума

Людмила Богданова  
13 апреля 2023  
<https://t.me/Luwowmafia>

# Оглавление

- Цель
- Исходные данные
- Выводы
- Сообщения и реакции по каналам
- Когорты. Активность по направлениям
- Когорты. Общая активность самых активных
- Гипотеза 1. "Зависимость количества реакций от дня недели и времени суток"
  - Распределение сообщений и реакций по датам
  - Распределение сообщений и реакций по дням недели
  - Распределение сообщений и реакций по времени
- Гипотеза 2. "Зависимость количества реакций от автора сообщения из команды Практикума"
  - Анализ участников команды
  - Успешный автор постов
  - Неуспешный автор постов
  - Есть ли зависимость количества полученных реакций от количества написанных сообщений?
- Гипотеза 3. "Зависимость количества реакций от длины сообщения"
- Гипотеза 4. "Больше реакций получают сообщения участников команды Яндекс.Практикума, чем студентов"
- Рекомендации

# Цель

Проанализировать учебные чаты Яндекс. Практикума, выяснить, когда активность студентов наибольшая и когда лучше публиковать анонсы, чтобы получить больше откликов и реакций

# Исходные данные

Датасет, предоставленный заказчиков, в котором содержится следующая информация: идентификатор сообщения, тип поста, идентификатор пользователя, дата поста, дата ответа, факт вхождения в неизвестную группу(возможно, команда Практикума), дата треда, метка действий пользователя, канал, дата файла, прикрепленные файлы (количество), реакции, длина текста сообщения, количество слов в сообщении

# Выводы

- Рассматриваемый временной период данных характеризуется активностью пользователей с 28.11 по 17.12 (совпадает с начало нового спринта и большим учебным проектом), в это время в большинстве каналов и когорт наблюдается рост числа сообщений и реакций на сообщения
- Больше всего постов и реакций в понедельник и среду, меньше всего - в субботу
- Больше всего реакций получают сообщения, опубликованные в 7.00 и 13.00
- Распределение количества сообщений близко к нормальному, активность сообщений наблюдается с 07.00 до 18.00. Утверждение о том, что в вечернее время студенты больше пишут сообщения и оставляют больше реакций на предоставленных данных не подтверждается.
- Самые популярные каналы: projects, exerciser, teamwork, info. В каналах projects и exerciser доля реакции / сообщения составляет менее 21-29%
- Зависимости количества реакций от количества слов в сообщении не наблюдается. Но высокое количество реакций находится в пределах до 250 слов в сообщении.

# Выводы

- В направлении Data Analysis больше всего сообщений(48,56% от общего количества) и реакций (53,54% от общего количества), а направление Data Lens оставляет практически равное количество сообщений и реакций.
- Во время спринта поведение пользователей в целом схоже с общей картиной: всплески реакций на сообщения, написанные в 07.00, 10.00, 13.00, 17.00, 19.00, количество сообщений распределено нормально с 07.00 до 19.00 - высокая активность по сообщениям.
- Среди 24 возможных участников команды Практикума были отобраны 5 предполагаемых участников , у которых количество реакций превышает количество сообщений и 3 "проблемных" участника команды с низкой долей реакций. Поведение "успешных" в целом совпало с общей картиной откликов, но среди "неуспешных" сообщения, написанные в правильные дни и часы не получали должного количества реакций.

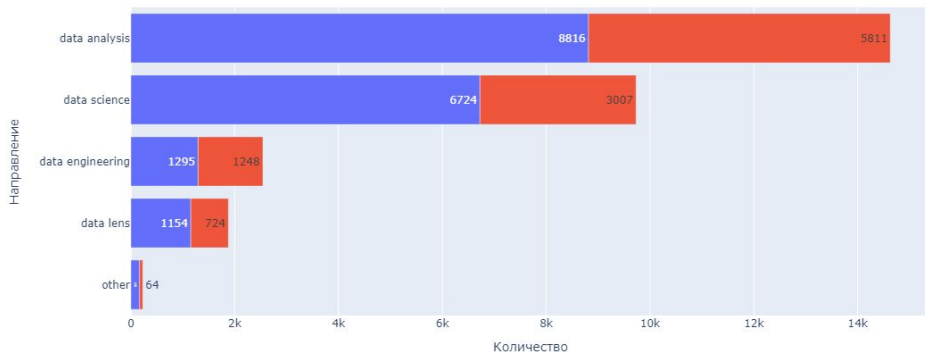
# Сообщения и реакции по каналам

Самые популярные каналы: projects (31% сообщений, 10% реакций), exerciser(22% сообщений, 10% реакций), teamwork (18% сообщений, 19% реакций), info (15% сообщений, 44 % реакций).

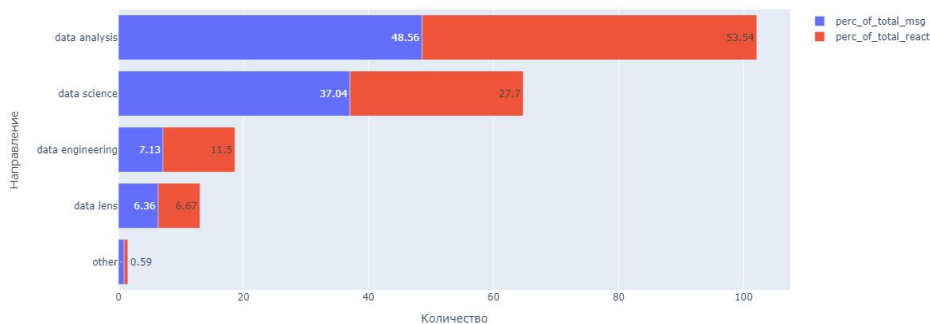
В каналах projects и exerciser доля реакций составляет менее 21-29% (понятно, что в этих каналах чаще спрашивают и отвечают сообщениями), а в каналах информационного характера info, digital\_professions, tutorial, masterskaya, library, data\_edteam\_info - студенты получают информацию и оставляют реакцию о просмотре (чаще всего информация в комментариях не нуждается).

# Когорты. Активность по направлениям

Абсолютное распределение количества сообщений и реакций по направлениям обучения



Относительное распределение количества сообщений и реакций по направлениям обучения, %



В направлении Data Analysis больше всего сообщений (48,56% от общего количества) и реакций (53,54% от общего количества), а направление Data Lens оставляет практически равное количество сообщений и реакций

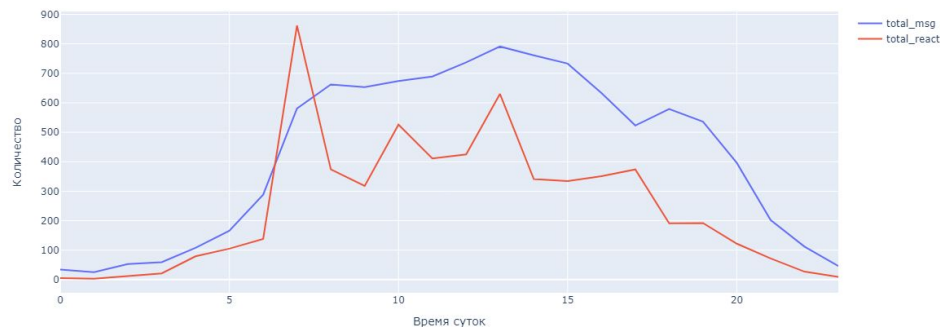


# Когорты. Общая активность самых активных

График зависимости количества сообщений и количества реакций от дня недели



График зависимости количества сообщений и количества реакций от времени суток



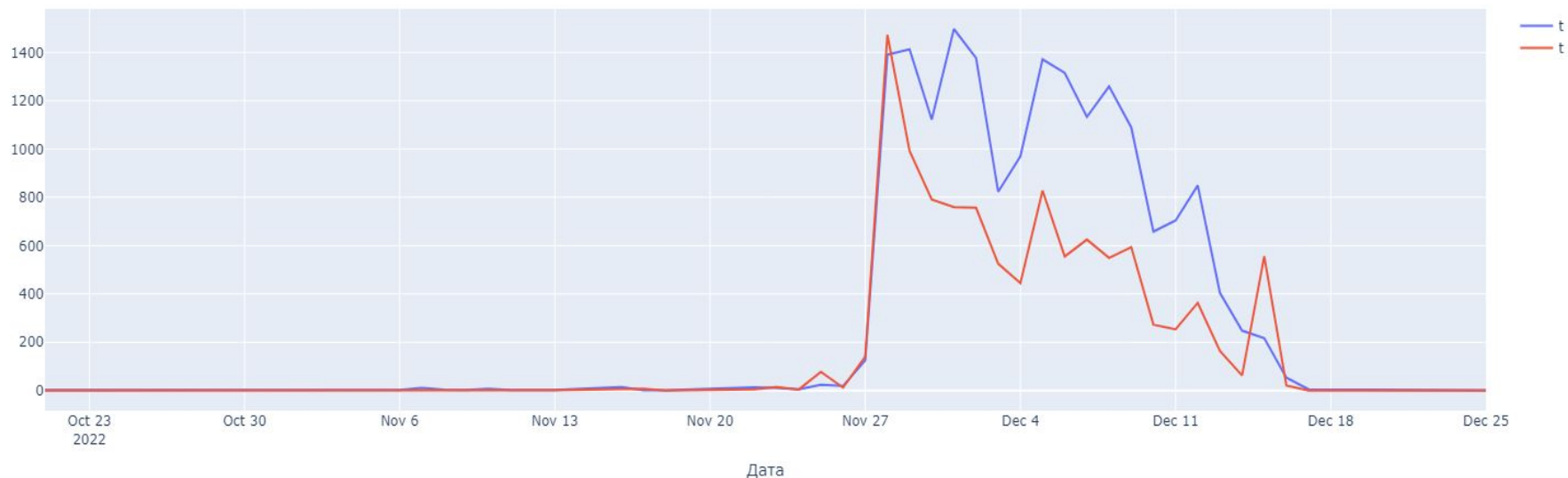
Топ-12 когорт активизируются в среду и низкоактивны в выходные. 07.00 и 13.00 - пики активности, с сохраняющейся активностью до 22.00

## Гипотеза 1. "Зависимость количества реакций от дня недели и времени суток"

Подтверждается: больше всего реакций наблюдается в понедельник и среду, пики реакций на сообщения 07.00, 10.00, 13.00, 17.00

# Распределение сообщений и реакций по датам

График зависимости количества сообщений и количества реакций от даты



Основная активность наблюдается с 28.11 по 17.12, что соответствует вероятному времени спринта (почти ровно три недели с понедельника по субботу)

# Распределение сообщений и реакций по дням недели

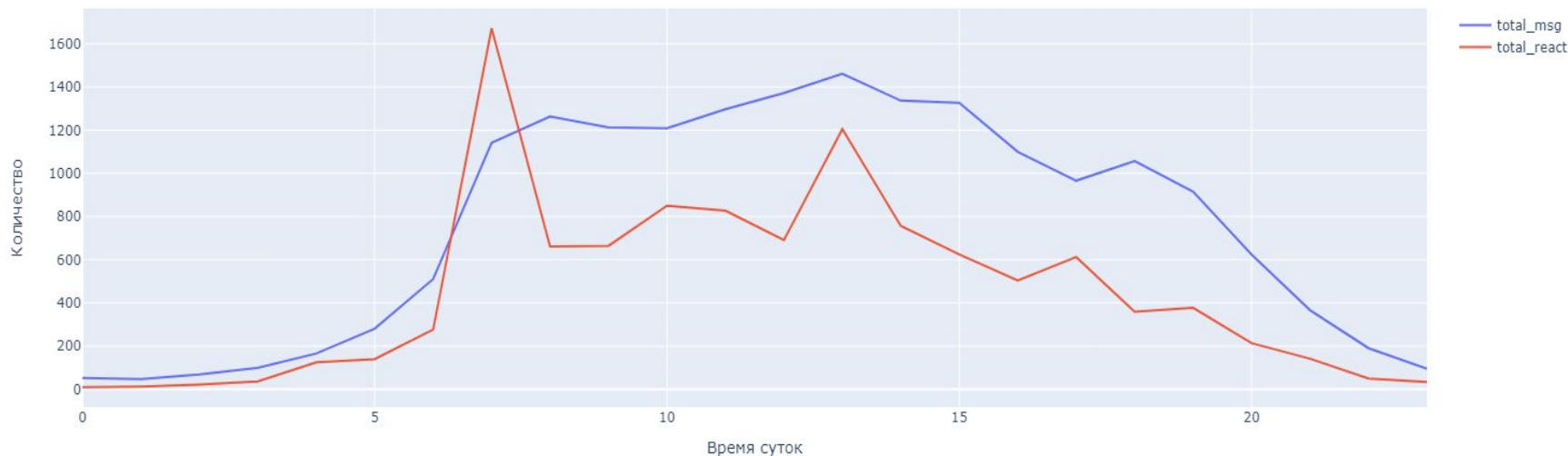
График зависимости количества сообщений и количества реакций от дня недели



Больше всего постов и реакций в понедельник и среду, меньше всего - в субботу

# Распределение сообщений и реакций по времени

График зависимости количества сообщений и количества реакций от времени суток



Пиковые моменты реакций на сообщения, написанные в 7.00 и 13.00. Распределение количества сообщения близко к нормальному, активность сообщений наблюдается с 07.00 до 18.00

## Гипотеза 2. "Зависимость количества реакций от автора сообщения из команды Практикума"

Вероятность такой зависимости высока (есть сообщения от команды в "правильные" часы и дни, но количество реакций все равно небольшое)

# Анализ участников команды

- Для определения возможных участников команды Практикума была построена сводная таблица по активностям в различных когортах (обычно студенты находятся в пределах 1-2 когорт, редко - трёх). Был составлен список из 24 таких кандидатов.
- Среди 24 возможных участников команды Практикума были отобраны 5 предполагаемых участников, у которых количество реакций на сообщения превышает количество сообщений и 3 "проблемных" участника команды с низкой долей реакций. Поведение "успешных" в целом совпало с общей картиной откликов, но среди "неуспешных" сообщения, написанные в правильные дни и часы не получали должного количества реакций.

Далее пример успешного юзера и не успешного

# Успешный автор постов

User U02PLR8QYT0 (отношение количество реакций / количество сообщений = 4.6)

Количество сообщений 143, реакций 662

User U02PLR8QYT0 больше всего пишет сообщений в понедельник, но большинство реакций получают сообщения, написанные в среду (возможно, именно в этот день он публикует анонсы).

Всплески реакции на сообщения, написанные в 07.00, 09.00, 13.00, 15.00, 17.00 и ничего после 18.00

График зависимости количества сообщений и количества реакций от дня недели

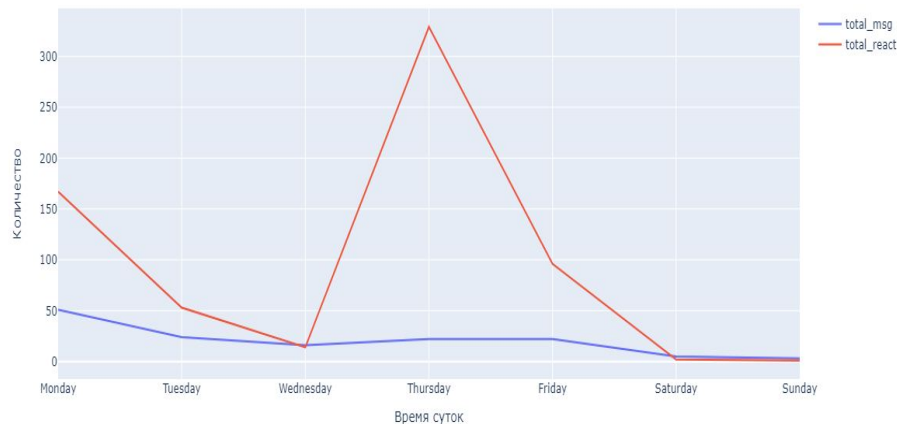
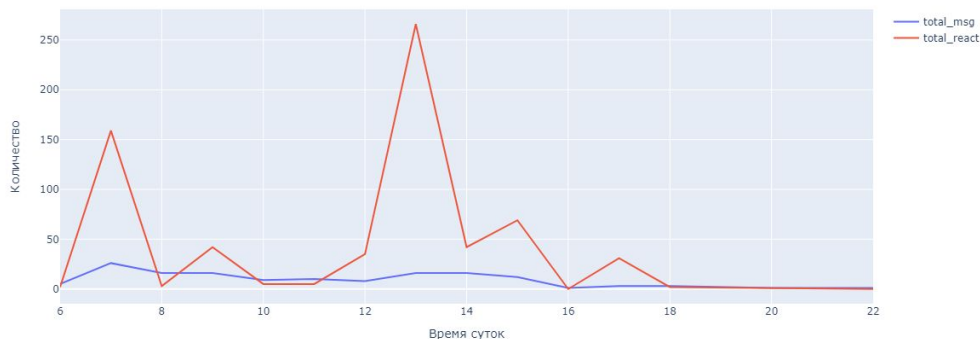


График зависимости количества сообщений и количества реакций от времени суток





# Неуспешный автор постов

User U031P6LDU56 (отношение количество реакций / количество сообщений = 0,41)

Количество сообщений 53, реакций 22

User U031P6LDU56 в основном пишет сообщения в ПН и ЧТ, большинство реакций получают сообщения, написанные в четверг.

Всплески реакции на сообщения, написанные в 08.00, 10.00, 13.00, 17.00.

График зависимости количества сообщений и количества реакций от дня недели

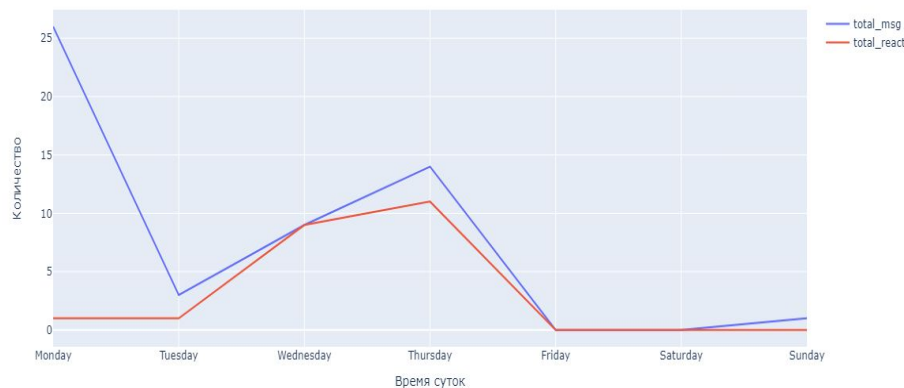
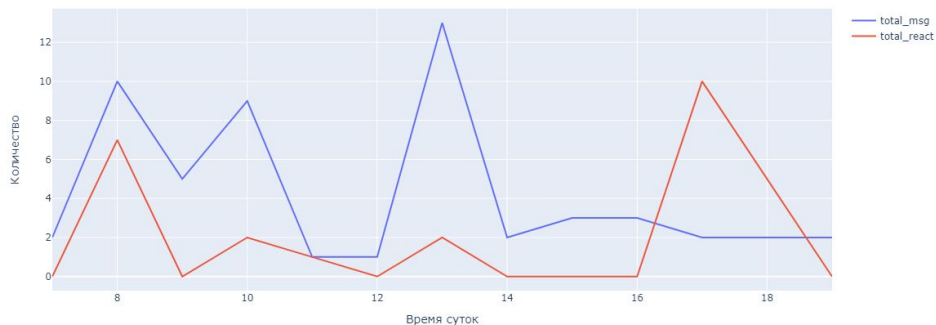
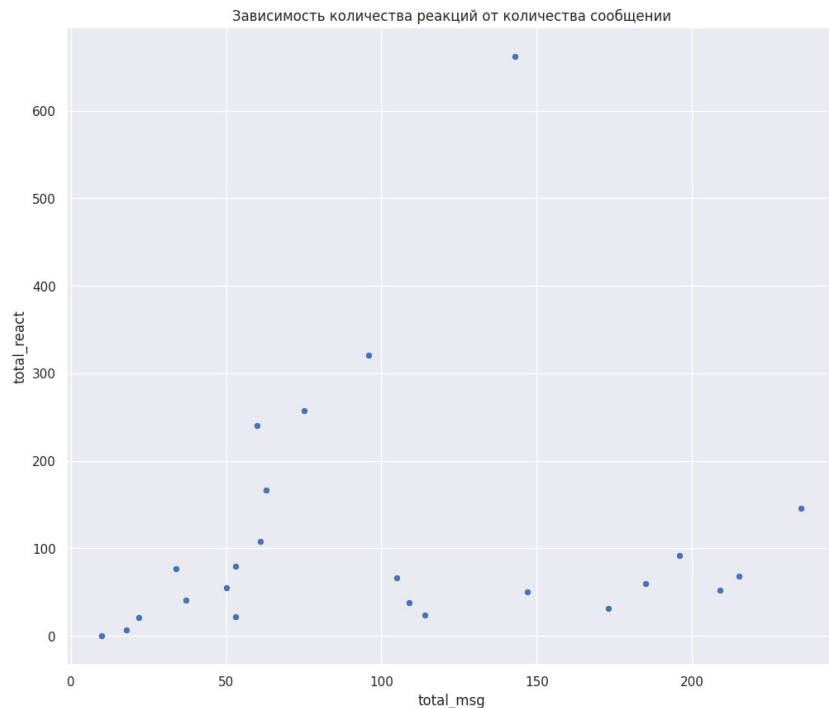


График зависимости количества сообщений и количества реакций от времени суток

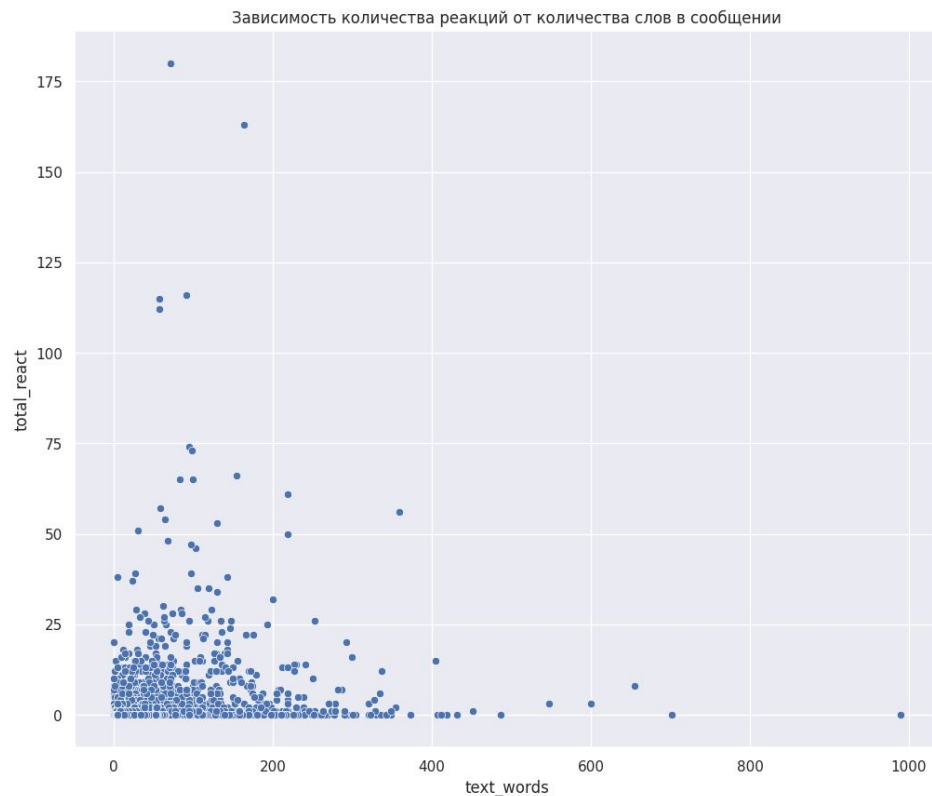


# Есть ли зависимость количества полученных реакций от количества написанных сообщений?



Линейной зависимости нет.  
Участник команды Яндекс.  
Практикума, получивший  
больше всего реакций на  
сообщения, написал среднее  
количество сообщений

### Гипотеза 3. "Зависимость количества реакций от длины сообщения"



Какой-то определенной зависимости количества реакций от количества слов в сообщении не наблюдается. Но высокое количество реакций находится в пределах до 250 слов в сообщении

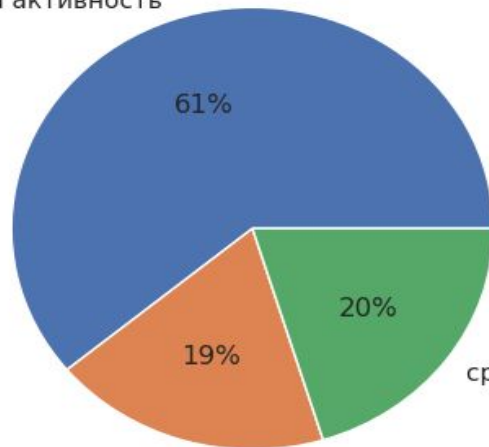
## Гипотеза 4. "Больше реакций получают сообщения участников команды Яндекс.Практикума, чем студентов"

Проанализировав реакции на сообщения пользователей - студентов и команды, видно, что 75% сообщений команды имеют как минимум 1 реакцию, а у сообщения студентов - нет.

# Активность команды и студентов

Активность команды Практикума

высокая активность

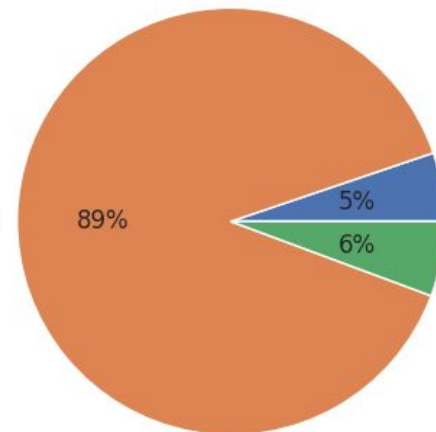


средняя активность

низкая активность

Активность активности студентов

низкая активность



высокая активность

средняя активность

# Рекомендации

- Команде не стоит публиковать анонсы в субботу и воскресенье.
- Несмотря на наличие определенных дней и часов, необходим дополнительный анализ содержания сообщений и личности публикующего из команды Практикума. Возможно, кого-то студенты знают лучше, чем куратора, публикующего анонсы раз в неделю.
- В анонсы-сообщения стоит добавить просьбу “Поставить реакцию” и указать какую, например, “лапку” (если нужно чтобы студенты что-то сделали, нужно попросить их это сделать).
- В целом, реакции можно заменить “глазиком” - количеством просмотров сообщения. Тогда команда будет видеть, сколько человек просмотрело это сообщений, даже без дополнительных действий.