

Представленный датасет содержит информацию о том, как обрабатываются коммуникации клиентов, обращающихся в поддержку Т-Банка в чатах и с помощью звонков. Колл-центр — это команда специально обученных специалистов, которые помогают клиентам решать возникающие вопросы. Ежедневно в банк обращаются сотни тысяч клиентов, поэтому важно, чтобы их вопросы обрабатывались быстро и качественно. Для Т-Банка, как онлайн-банка, , высокий уровень обслуживания — ключевое конкурентное преимущество.

## Как выглядит путь клиента?

Если у клиента возникает вопрос или проблема, сначала он может попробовать решить их самостоятельно: воспользоваться приложением, помощниками, FAQ-страницами и другими инструментами. Многие клиенты на этом этапе находят ответы сами — за счет удобного приложения и понятной навигации.

Но не все вопросы можно решить без помощи. Наш датасет как раз описывает путь тех клиентов, кто не смог справиться самостоятельно и обратился в поддержку в чате или с помощью звонка. Тех, кто нашел решение сам, мы здесь не рассматриваем.

Когда клиент понимает, что ему нужна помощь, у него есть два пути: чат или звонок.

Дальнейшие этапы будут во многом схожи.

1. При обращении клиент сначала попадает на бота (в чате — текстового, по телефону — голосового). Бот старается понять проблему и предложить решение.
2. Если бот не справляется, он переводит клиента на сотрудника поддержки. При этом бот собирает контекст: определяет продукт, по которому возник вопрос, и формулирует суть обращения. Благодаря этому при передаче запроса сотруднику система выбирает наиболее подходящего специалиста с нужными навыками.

После завершения общения клиент получает пуш-уведомление с просьбой оценить обслуживание по шкале CSAT (от 1 до 5) и указать, решена ли его проблема. Эти данные позволяют понять, насколько клиенты удовлетворены качеством работы поддержки.

## Что содержится в датасете?

В датасете представлены коммуникации за 10 рабочих дней — с 2 по 14 июня 2025 года — с клиентами, обратившимися в поддержку Т-Банка, с помощью которых можно оценить, насколько качественно мы обслуживаем клиентов и что может влиять на их негативный опыт.

## Информация по обращению:

appeal\_id — ID обращения

start\_dttm — дата/время начала обращения

appeal\_type\_nm — тип обращения (чат/звонок)

participants\_desc — участники обращения (сотрудник, бот + сотрудник)

appeal\_time\_sec — время от начала обращения до его завершения. С учетом маршрутизации, звонка на линию помощи, удержания клиента, автозакрытия чата по тайм-ауту. То есть это все то время, когда решали вопрос клиента.

employee\_dialog\_time\_sec — время общения с сотрудником

bot\_dialog\_time\_sec — время общения с ботом

repeated\_24h\_flg — флаг, что клиент повторно обратился в поддержку в течение 24 часов после предыдущего обращения

repeated\_48h\_flg — флаг, что клиент повторно обратился в поддержку в течение 48 часов после предыдущего обращения

repeated\_5d\_flg — флаг, что клиент повторно обратился в поддержку в течение 5 дней после предыдущего обращения

domain\_nm — домен / группа продуктов. Переменная для разграничения продуктов

csat\_score — оценка CSAT-опроса

appeal\_cost\_rub\_amt — стоимость обращения. Рассчитывается в рублях

transfer\_cnt — количество трансферов в рамках обращения. Сколько раз клиента переключали между сотрудниками или доменами

task\_cnt — количество созданных тасок. Сколько офлайн-задач было создано в рамках коммуникации. Задачи создают, когда не могут решить проблему в онлайн и нужно участие более опытного сотрудника

problem\_id — ID проблемы. Агрегированная сущность. Через ML-модель склеиваем коммуникацию в одну проблему клиента

problem\_status\_code — статус проблемы (решена ли проблема в ходе коммуникации, (RESOLVED/UNRESOLVED))

repeated\_appeal\_flg — флаг повторной коммуникации по той же проблеме

help\_line\_calls\_cnt — количество звонков на линию помощи. Сколько раз позвонили на линию помощи сотрудники, когда сами не смогли решить вопрос клиента

## Информация по клиенту:

id\_client — уникальный ID-клиента

gender\_cd — пол клиента

client\_segment\_nm — сегментация клиентов (высокодоходные/доходные/малодоходные)

age\_segment\_nm — сегмент по возрасту (дети/подростки/молодежь/взрослые/сеньоры/не определен)

m\_segment\_flg — предприниматели из M-сегмента

s\_segment\_flg — предприниматели из S-сегмента

auto\_segment\_nm — сегмент «Автолюбители». Смотреть таблицу с описанием переменных

travel\_segment\_nm — сегмент «Путешественник». Смотреть таблицу с описанием переменных

client\_type\_nm — тип клиента по сегменту «Премиум»

subscription\_nm — final\_grade\_code

## Исследовательский вопрос

Что влияет на мнение клиента при обращении в поддержку Т-Банка и формирует его положительный или отрицательный опыт?

## Альтернативный вопрос

Есть ли связь между стоимостью обработки обращений и тем, остается клиент доволен или недоволен поддержкой?

## Метрики для исследования клиентского опыта

После общения с поддержкой клиент получает CSAT-опрос и ставит оценку от 1 до 5. В датасете есть переменная:

- csat\_score — оценка, которую поставил клиент за пройденный опрос.

Из этого поля можно посчитать переменную csat\_amount — количество раз, когда каждая оценка была поставлена. Например, если три клиента поставили 5, то для оценки 5 csat\_amount = 3.

Исходя из csat\_score и рассчитанного csat\_amount, можно построить следующие метрики.

### 1. Средняя оценка CSAT

Показывает общий уровень удовлетворенности клиентов.

Берем каждую оценку, умножаем на количество ее повторений, складываем и делим на общее количество оценок.

### 2. Доля пятерок

Показывает долю клиентов, которые остались полностью довольны обслуживанием.

### 3. Доля «единиц»

Показывает долю клиентов, которые остались недовольны.

## Совет для анализа:

- Средний балл дает усредненную картину клиентского опыта.
- Доли пятерок и единиц позволяют явно увидеть, сколько клиентов полностью довольны и сколько максимально недовольны.

### Пример расчета метрик

Предположим, у нас есть 10 клиентов и их оценки в `csat_score`.

**Шаг 1.** Считаем `csat_amount` — сколько раз встречается каждая оценка.

**Шаг 2.** Считаем метрики.

Средний CSAT:

Доля пятерок:

$$\text{CSAT (5)} = 5 / 10 = 50\%$$

Доля единиц:

$$\text{CSAT (1)} = 2 / 10 = 20\%$$

Из примера видно, что средний балл дает усредненную картину клиентского опыта, а доли пятерок и единиц позволяют увидеть, сколько клиентов полностью довольны и сколько недовольны. Эти метрики можно использовать для анализа качества обслуживания и исследования факторов, влияющих на положительный или отрицательный клиентский опыт.