

DAN 



ВЫСШАЯ ШКОЛА
ЭКОНОМИКИ



ТИНЬКОФФ

ИТМО

АКТИВНОСТЬ ПОЛЬЗОВАТЕЛЕЙ НА ПЛАТФОРМЕ «ТИНЬКОФФ ГОРОД»

В конце декабря 2022 года Тинькофф представил платформу «Город», цель которой — дать пользователям приложения Тинькофф возможность легко и удобно найти и заказать нужные им продукты и услуги, а также узнать о событиях и мероприятиях вокруг них. Этой платформе отводится ключевая роль в формировании экосистемы Тинькофф, объединяющей на базе суперприложения банка функции маркетплейса, туроператора, страхового агента, гида, афиши, СМИ, социальной сети и мессенджера. Контент Города отличается в зависимости от истории покупок и активности пользователя: автомобилисту Город покажет напоминание о заправке, человеку, регулярно посещающему барбершоп, — предложение записаться на стрижку. Интеграция с Тинькофф Журналом позволяет наполнить ленту полезным контентом на близкие пользователям темы.



Помимо расширения спектра товаров и услуг, доступ к которым получает пользователь через приложение Тинькофф, развитие платформы «Город» имеет целью и расширение клиентской базы, в том числе ее возрастных границ. Поэтому особый интерес для банка представляет анализ потребительского поведения подростков — пользователей 14–18 лет, особенно учитывая более активное использование ими мобильных приложений. Анализируя особенности потребительского поведения подростков, мы получаем представление о потребительском поведении ядерной целевой аудитории, 25–45 лет, в будущем.

Предшественником Тинькофф Города был Тинькофф Лайфстайл, включавший сервисы, которые позже стали частью Города. В датасете, который вы получили, представлены данные о том, какие покупки совершали подростки 14–18 лет в сервисах Тинькофф Лайфстайла с 1 января 2022 по 1 января 2023 года, а также некоторые характеристики данных пользователей. Одной записью представлен заказ. Всего в базе 24 721 наблюдение (= заказ), которые сделаны 10 079 уникальными клиентами.

Описание переменных представлено ниже.

user_id	Уникальный идентификатор клиента
sex	Пол (М = мужской, F = женский)
age	Возраст клиента
region	Актуальный регион, где клиент находится в данный момент
bundle	Наличие подписки (No bundle = нет подписки, Tinkoff Pro, Tinkoff Premium — соответствующие подписки)
num_sessions	Количество сессий* у пользователя за 2022 год
avg_session	Средняя длина сессии* (в секундах)
fuel, cinema, vkusvill, travel, restaurants, concerts, games, books, theater, sports, beauty, flowers, technology	Бинарные переменные, описывающие категорию заказа (Топливо, Кино, Вкусвилл, Путешествия, Рестораны, Концерты, Игры, Книги, Театр, Спорт, Красота, Цветы, Техника). Например, если заказ в категории «Топливо», переменная fuel будет равна 1, все остальные переменные — 0
order_date	Дата создания заказа
price	Стоимость заказа (price = 0 означает, что осуществлено бронирование)

Используя полученные данные, вы должны ответить на вопрос:

«**Что влияет на активность пользователей в экосистеме Тинькофф?**»

1. Что мы можем использовать в качестве меры активности? (Помните, что банк стремится создать универсальную платформу, которая будет использоваться потребителями для покупок всех возможных товаров и услуг. Соответственно, нас может интересовать как объем покупок, измеренный в количестве заказов и общей сумме заказов, так и их разнообразие с точки зрения категорий). Выберите одну или несколько мер активности пользователей и обоснуйте свой выбор.
2. Какие доступные в базе данных характеристики пользователей могут быть связаны с активностью пользователей? Каких зависимостей вы ожидаете?
3. Чтобы исследовать связь между характеристиками пользователей и их активностью в Тинькофф Городе, вам необходимо будет перейти от базы, где единица наблюдения — заказ, к базе данных, где единицей наблюдения будет пользователь. Агрегируйте данные на уровне пользователей.

* Сессия — это время от открытия приложения до его закрытия, включает в себя все действия, совершенные в мобильном банке.

4. Обратите внимание, что распределение пользователей по числу заказов устроено следующим образом: большинство пользователей сделали всего один заказ, но при этом есть потребители, сделавшие более 10 заказов. На какие группы можно разделить потребителей с точки зрения числа заказов? Как эту классификацию можно использовать в анализе?
5. Учитывая все вышеперечисленное, проведите анализ данных и выявите ключевые факторы, влияющие на активность пользователей на платформе «Тинькофф Город». Обратите внимание на использование пользователями услуги бронирования. Влияет ли использование этой услуги на активность пользователей?
6. Подведите итоги проведенного исследования. Что именно вы хотели выяснить? Какие результаты были получены? Как эти результаты могут быть использованы на практике? Какие ограничения в рамках своего исследования вы видите? Если бы у вас было больше времени и не было ограничений, связанных с доступностью данных, как вы могли бы дополнить и улучшить свое исследование?

Примечание. Пункты выше примерные и отражают программу максимум. Успешная работа не обязательно требует выполнения всех из них — вы можете сфокусироваться на отдельных аспектах. Более того, работа, где глубоко исследуют какой-то один важный аспект, представляет большую ценность, чем работа, где исследуют все, но поверхностно.

Желаем вам плодотворной работы на хакатоне!