

Вебинар по снятию запроса

Проект «Телеком»

Поиск клиентов с неоптимальными
тарифами // когорта 37DA

Алексей Каргин

code reviewer Eng / team lead Ru / middle reviewer Ru / Яндекс Практикум
29 мая 2022 г.



Наши договоренности

- можно на «ты»
- камера включена
- понятное имя и фамилия
- одновременно работает 1 микрофон
- уважительное отношение друг к другу
- глупых вопросов не бывает

Проработка запроса

Проработка запроса

Зачем?

Не всегда задача сформулирована чётко и исчерпывающе. Чтобы решить её самостоятельно, нужно собрать больше информации. Привыкните к тому, что это на вашей ответственности.

Важно!

Вопросы задаем с позиции джуна, который получил от тимлида задачу, и хочет уточнить неясные моменты.

Проработка запроса

1. Уточните, кто конечный **заказчик** результата.
2. Выясните, **зачем** нужен **результат** вашей работы (как в дальнейшем будут использованы данные).
3. Уточните все **параметры**, которые хочет знать заказчик задачи.
Нет ли подводных камней?
4. Проговорите и зафиксируйте **образ идеального** конечного **результата** (отчет, презентация, дашборд).
5. Набросайте **прототип решения** (декомпозиция проекта) — и сверьтесь с заказчиком, подойдет ли оно?
6. Предоставьте заказчику **конечный результат**.

Поиск клиентов с неоптимальными тарифами

Вопросы из Slack (общие)

- Какие параметры хочет видеть заказчик?
- Что стоит отразить в презентации?
- И что стоит сравнивать в гипотезах?
- Как определить активных пользователей: учитывать тех, кто совершал звонки в ноябре с датой регистрации ранее или всех звонивших в ноябре?
- Нужно ли классифицировать клиентов по длительности оказания услуг с момента регистрации в сервисе?
- Надо ли учитывать динамику количества совершаемых исходящих вызовов или их длительность при определении нового тарифа?

Поиск клиентов с неоптимальными тарифами

Цель проекта: Провести мониторинг клиентов, которые экономически неэффективно используют услуги на своих тарифах, и могут снизить свои расходы, если перейдут на другие тарифы, с более выгодными для них условиями.

А зачем? Есть вероятность, что клиенты, которые платят много уйдут к конкурентам. Переведя их на оптимальные тарифы, можно будет снизить вероятность оттока. Таким образом, можно увеличить их LTV, пожертвовав сиюминутной прибылью.

Поиск клиентов с неоптимальными тарифами

- Задачи проекта:** установить сколько активных клиентов используют экономически невыгодные тарифы и как изменится выручка компании, при переводе всех клиентов на оптимальные тарифы. Поэтому есть еще несколько задач, которые мы должны решить:
- изучить динамику работы клиентов Телекома и выделить активных клиентов, по которым будем проводить все расчеты;
 - изучить текущую выручку компании (в разрезе месяца), а также динамику за весь период, включая изучение среднего чека в разрезе тарифа;
 - определить клиентов, которые неэффективно используют услуги Телекома (платят больше, чем могли бы платить на другом тарифе);
 - изучить потери в выручке, если все клиенты перейдут на выгодные тарифы;
 - предложить пути решения проблемы потери выручки;

Дополнительно, такое исследование позволит лучше понять как клиенты потребляют услуги и возможно выдвинуть гипотезы нового ценообразования. Планируется проводить такие исследования раз в квартал.

Поиск клиентов с неоптимальными тарифами

Результат проекта: полученные результаты необходимо представлять в виде:

- тетрадки **JupyterNotebook** для тимлида со всеми расчетами, выводами и пояснениями результатов
- **отчета** для менеджеров коммерческого департамента, по данным которого будет проведена работа с клиентами; отчет должен включать список клиентов, которые неэффективно для себя пользуются услугами Телекома и которым целесообразно предложить смену тарифного плана, а также дополнительную информацию, которая будет полезна в работе;
- **презентации для коммерческого директора**, рассчитанную на небольшое сообщение – пять-семь минут и содержащую ответы на главные вопросы, которые интересуют заказчика;
- **дашборд**, иллюстрирующий потребление услуг клиентами Телекома за весь период времени в разрезе тарифного плана (с возможностью фильтрации по типу и направлению звонка).

Поиск клиентов с неоптимальными тарифами

Вопросы из Slack (биллинг)

- Как вести расчет оплаты, можно за полный месяц или нужно учитывать дату регистрации в сервисе и исходить от этой даты?
- При расчете стоимости оплаты за месяц нам ведь нужны строки где вызовы не пропущены? И нужно ли использовать столбец длительность звонка с учётом времени ожидания?
- Как использовать информацию о количестве звонков и длительность звонка с учётом времени ожидания?
- Как происходит тарификация исходящих звонков: цена по тарифу за длительность звонка без учёта времени ожидания?

Поиск клиентов с неоптимальными тарифами

Расходы клиента состоят из:

- **ежемесячной платы** +
- платы за перерасход по внутренним исходящим звонкам, **с учетом времени ожидания** +
- платы за внешние исходящие звонки, **с учетом времени ожидания** +
- платы 100 руб за оператора

В Телекоме действуют следующие тарифные планы:

План	Ежемесячная плата	Входящие	Исходящие звонки внутри сети — бесплатны до 2000 минут в месяц на всех тарифах. Цена сверх лимита:	Исходящие звонки, кроме внутренних:
А	5 000 ₽	0	0.10 ₽/мин.	0.4 ₽/мин.
В	2 000 ₽	0	0.15 ₽/мин.	0.5 ₽/мин.
С	1 000 ₽	0	0.30 ₽/мин.	0.7 ₽/мин.

Поиск клиентов с неоптимальными тарифами

Счета выставляются на прошедший календарный месяц, поэтому списание можно считать от 1-го числа каждого месяца. Остаток трафика сгорает на начало месяца. Например, клиент начал платить 9-го числа, это значит что за период с 9 по конец месяца у него работают действующие аккумуляторы трафика (например, 2000 бесплатных внутренних минут).

Тарификация происходит на основе кумулятивных показателей за период расчета (**календарный месяц**). Продолжительность звонков за месяц округляется в большую сторону.

Поиск клиентов с неоптимальными тарифами

Вопросы из Slack (данные)

- В каких единицах времени данные о длительности звонков? В столбце `call_duration` - длительность одного успешного звонка, а в столбце `calls_count` - количество попыток дозвонится?
- Чем заменить пропуски в `operator_id` датасета `telecom_dataset` (пропусков 16%)?
- Если длительность звонков измеряется в секундах, то действует ли правило округления (если звонок длится несколько секунд, то платишь как за целую минуту)?
- Стоит ли проводить какую либо работу с аномальными значениями в столбцах количества звонков и длительности?
- Являются ли аномальным пропущенные исходящие вызовы, в которых указана длительность звонка без учёта времени ожидания?

Поиск клиентов с неоптимальными тарифами

Датасеты характеризуют использование услуг «Нупозвони» — провайдера виртуальной телефонии. Его клиенты — колл-центры (юр. лица), которые:

- распределяют входящие вызовы по своим операторам,
- обеспечивают исходящие вызовы от операторов.

	user_id	date	direction	internal	operator_id	is_missed_call	calls_count	call_duration	total_call_duration
0	166377	2019-08-04 00:00:00+03:00	in	False	NaN	True	2	0	4
1	166377	2019-08-05 00:00:00+03:00	out	True	880022.0	True	3	0	5
2	166377	2019-08-05 00:00:00+03:00	out	True	880020.0	True	1	0	1
3	166377	2019-08-05 00:00:00+03:00	out	True	880020.0	False	1	10	18
4	166377	2019-08-05 00:00:00+03:00	out	False	880022.0	True	3	0	25

Поиск клиентов с неоптимальными тарифами

user_id — идентификатор клиента,

date — дата статистики;

direction — направление вызовов: **out** — исходящий вызов, **in** — входящий;

internal — маркер внутренних и внешних вызовов; вызовы между сотрудниками одного колл-центра называют внутренними, остальные — внешними.

operator_id — идентификатор оператора;

is_missed_call — маркер пропущенных вызовов;

calls_count — количество звонков;

call_duration — длительность **всех** звонков без учёта времени ожидания;

total_call_duration — длительность **всех** звонков с учётом времени ожидания.

Примерное содержание проекта

1. Загрузка данных и подготовка к исследованию;
2. Исследовательский анализ данных:
 - Общая характеристика данных, включая анализ количества клиентов и их динамики использования услуг Телекома и параметры выбора активных клиентов;
3. Решение основных вопросов заказчика;
4. Проверка статистических гипотез (две H_0 гипотезы);
5. Вывод и рекомендации для заказчика;
6. Материалы для коллег из коммерческого департамента;
7. Ссылка на дашборд.

Финальный спринт

Я Практикум

Финальный спринт: структура

1: Декомпозиция основного проекта

- детализация всех стадий исследования данных (загрузка, подготовка, исследовательский анализ, статистический анализ с формулировкой гипотез, заключительная часть)

2: Основной проект (10 дней) дедлайн **09 июня**:

- проект
- презентация
- дашборд

3: A/B-тест

4: SQL проект

5: Итоговый созвон:

- обсуждение результатов проекта
- демонстрация интересных решений (опционально)

Важная информация

Поскольку наша задача смоделировать реальную работу, то присылаем на проверку **весь проект, включая дашборд и презентацию.**

Если есть вопросы — **обсуждаем их в канале Слака** нашей группе — чем более активным будет обсуждение, тем проще будет завершить проект.

Общее обсуждение задач и решений — одна из идей финального спринта.

Лучше **не брать** решение из интернета
(я их все знаю + не смогу дать персональную обратную связь).

Удачи!