

# А/В тест для хакатона

## Метрики

1. Основная метрика (целевая): изменение средней эффективности в течении месяца после коммуникации с момента использования ML модели (для предложения «хорошего» курса)
2. Прокси метрика: процент прохождения курсов с момента запуска модели
3. Контр. метрика: использование продукта клиентом (util\_flg). Считаем, что данный показатель не должен упасть.

## Гипотеза

Будем считать двухсторонний А/В тест, чтобы одновременно зафиксировать стат. значимое отрицательное влияние ML модели.

1. Нулевая гипотеза: модель не влияет на основную метрику
2. Альтернатива: изменения есть!

## Мощность/стат. значимость

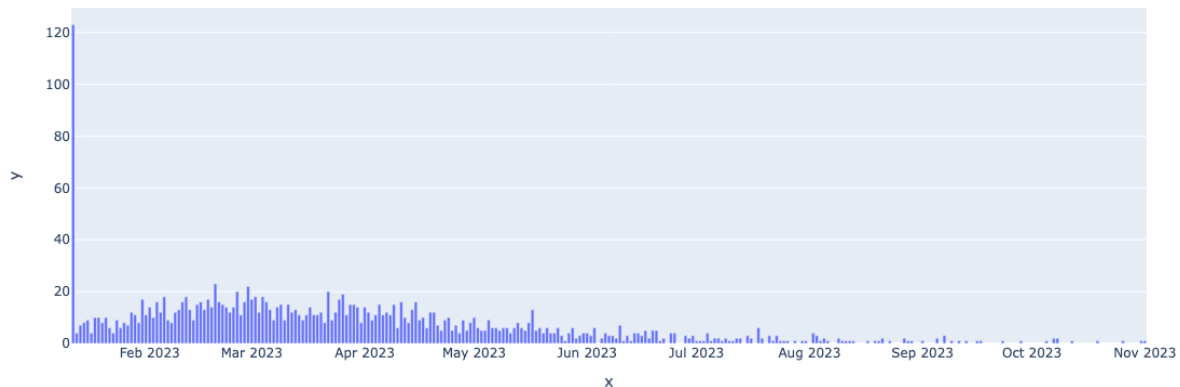
Просто зафиксирую здесь, что следуя историческим результатам, выбираем мощность 80% и статистическую значимость 95%.

## Выбор групп

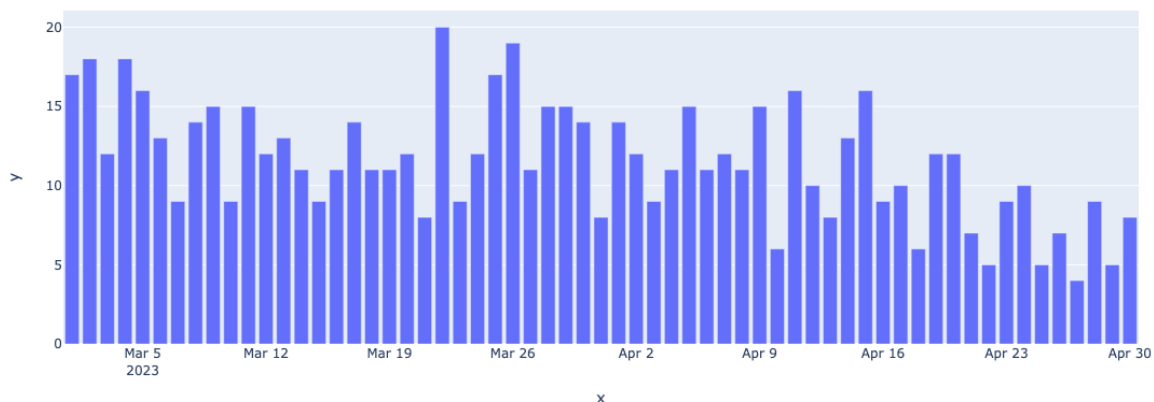
Предлагается разделить всех сотрудников на две группы (случайным образом). То есть в тесте/контроле ~ 1000 человек. При этом каждый человек попадает в тест после того, как модель будет развернута, в день первого просмотра какого-либо курса (начало его прохождения).

## Оценка MDE/длительность теста

Для оценки MDE/длительности теста будем считать, что тестовая и контрольная группы имеют примерно равную дисперсию и размер. Оценим среднее число сотрудников, которые попадают в «тест». Для этого посмотрим на распределение даты начала прохождения своих первых курсов всеми сотрудниками.



Видно, что сотрудники «активно» вовлекаются в прохождение курсов в начале года, поэтому для определения среднего числа попадания в «тест» возьмем Март 2023 и Апрель 2023.



Таким образом, в «тест» в среднем попадают по 11.7 новых сотрудников в день, то есть по 6 сотрудников/день попадают в тестовую и контрольную группы.

Среднее значение основной метрики:  $60.0 \pm 3.4$

Исходя из приведенных выше данных, можно оценить MDE (минимальный прирост по метрике) для 30, 45 и 60 дней

$MDE(30 \text{ дней}) : 1.7\%$

$MDE(45 \text{ дней}) : 1.4\%$

$MDE(60 \text{ дней}) : 1.2\%$