

Построение аналитики для корпоративных клиентов обучающей платформы

SUMMARY

Для любого учебного заведения важно понимать, что происходит по следующим направлениям:

1. Какое количество учащихся
2. Каков прогресс обучения, может быть так, что студент зарегистрировался на курс, но его прогресс слишком медленный и он не успевает до дедлайна. В таком случае ему можно будет направить напоминание или предложить помощь, если возникли технические проблемы или проблемы с усвоением материала.
3. Интересно посмотреть, какие задачи вызывают у студентов наибольшие сложности - на какие задачи тратится наибольшее количество попыток или были попытки решения, но правильный вариант ответа так и не был дан. В таком случае можно пересмотреть подачу учебного материала или предложить в курсе обучения дополнительные более легкие задачи, которые подведут к решению более сложных.
4. Статистика заходов студентов на платформу: плавно или скачкообразно идет процесс обучения. Может в некоторые периоды заходов вообще не было из-за проблем с оборудованием
5. Задачи, к решению которых студенты вообще не приступали: проверить технические моменты. Может эти задачи не видны на платформе или объяснение задачи не понятно написано.
6. Студенты, которые не приступали к учебе (не решили ни одной задачи, ни одного теста, не сделали ни одной попытки): связаться со студентами и уточнить, нет ли технических сложностей при входе на платформу
7. Количество активированных аккаунтов на платформе. Если студент зарегистрировался, но не активировал аккаунт - проверить, что могло послужить причиной

АНАЛИЗ

1. Общее количество студентов на платформе

```
SELECT
    count(id)
FROM
    users
WHERE
    company_id = 1
```

Всего на данный момент на платформе зарегистрировано 133 студента для Компании 1

2. Общий прогресс обучения

Для расчета общего прогресса обучения сначала рассчитаем количество задач, представленное на платформе для Компании 1: 124

Соответственно 133 студента должны каждый решить по 124 задачи.

133 * 124 = 16 492 задачи должно быть решено. Это наша цель - 100%

Прогресс обучения это отношение количества успешно решенных задач к общему количеству (16 492). Здесь мы берем во внимание только одно правильное решение задачи для одного студента (т.к. Студент мог правильно решить одну и ту же задачу несколько раз)

```
WITH
  problems_count AS (
    SELECT
      company_id,
      count(*) AS all_amount
    FROM
      problem_to_company
    WHERE
      company_id = 1
    GROUP BY
      company_id
  ),
  students AS (
    SELECT
      company_id,
      count(id) AS total_users
    FROM
      users
    WHERE
      company_id = 1
    GROUP BY
      company_id
  ),
  agg AS (
    SELECT
      u.email,
      pc.all_amount,
      count(*) AS tries,
      count(
        CASE
          WHEN cs.is_false = 0 THEN 1
        END
      ) AS solutions,
      count(
        DISTINCT CASE
          WHEN cs.is_false = 0 THEN cs.problem_id
        END
      ) AS solved_problems,
      round(
        count(
          DISTINCT CASE
            WHEN cs.is_false = 0 THEN cs.problem_id
          END
        ) * 100.0 / pc.all_amount
      ) AS progress
```

```

FROM
    codesubmit cs
    JOIN users u ON u.id = cs.user_id
    JOIN company c ON c.id = u.company_id
    JOIN problems_count pc ON pc.company_id = u.company_id
WHERE
    c.id = 1
GROUP BY
    1,
    pc.all_amount
)
SELECT
    sum(solved_problems) * 100.00 / (124 * 133)
FROM
    Agg

```

Итоговый прогресс обучения 18,18%

Здесь интересно узнать, каковы дедлайны обучения, если они имеются и посмотреть успевают ли студенты завершить программу к дедлайну при таком темпе ее прохождения.

3. Статистика заходов на платформу

```

SELECT
    date_trunc('week', ue.entry_at),
    count(DISTINCT ue.user_id)
FROM
    userentry ue
    JOIN users u ON u.id = ue.user_id
WHERE
    u.company_id = 1 [[and ue.entry_at::date between {{start_date}} and
    {{end_date}}]]
GROUP BY
    1

```

В ноябре и апреле виден всплеск уникальных заходов на платформу и увеличение количества новых пользователей. Это может быть связано с приемом большого количества новых сотрудников или с необходимостью закрыть сессию к полугодовому дедлайну.

4. Статистика прохождения тестов

Соотношение правильных и неправильных ответов, которые дают студенты.

```

SELECT
    count(t4.is_correct) AS total_answers,
    count(
        CASE
            WHEN t4.is_correct = TRUE THEN 1
        END
    ) AS correct_answers,
    count(

```

```

        CASE
            WHEN t4.is_correct = FALSE THEN 1
        END
    ) AS wrong_answers
FROM
    testresult t
    JOIN test t2 ON t.test_id = t2.id
    JOIN users u ON u.id = t.user_id
    LEFT JOIN testanswer t4 ON t.answer_id = t4.id
WHERE
    u.company_id = 1

```

Соотношение правильных и неправильных ответов на тестовые вопросы примерно 1 : 1. Здесь нужно будет проверить дополнительно формулировки и типы вопросов. Если тип вопроса с бинарным ответом - возможно у учащихся низкий уровень подготовки и они просто угадывают ответы. Либо вопросы сформулированы так, что даже знающие ученики их недопонимают и ошибаются с выбором.

5. Задачи с наибольшим количеством попыток и правильными решениями

```

WITH
    total AS (
        SELECT
            problem_id,
            1 AS is_false
        FROM
            coderun cr
        UNION ALL
        SELECT
            problem_id,
            is_false
        FROM
            codesubmit cs
    )
SELECT
    total.problem_id,
    p.name,
    count(*) AS cnt,
    count(
        CASE
            WHEN is_false = 0 THEN 1
        END
    ) AS solved
FROM
    total
    JOIN problem_to_company ptc ON ptc.problem_id = total.problem_id
    JOIN problem p ON p.id = total.problem_id
WHERE
    ptc.company_id = 1
GROUP BY

```

```
total.problem_id,  
p.name  
ORDER BY  
solved
```

Интересно посмотреть на задачи, для которых было затрачено большое количество попыток по отношению к количеству правильных решений. В чем возникают сложности у студентов? Понятны ли формулировки вопросов в задачах?

6. Активированные аккаунты

Студенты могли зарегистрироваться, но так и не активировали аккаунт. Возможно не пришло письмо с ключом активации или возникли другие технические сложности, которые необходимо устранить.

```
SELECT  
date_trunc('month', date_joined) AS MONTH,  
count(*) AS new_students,  
count(  
DISTINCT CASE  
WHEN is_active = 1 THEN id  
END  
) AS active_students  
FROM  
users  
WHERE  
company_id = 1  
GROUP BY  
1
```

Практически все аккаунты активированы. В отношении тех, которые еще не активированы, уточнить, нет ли технических проблем с активацией.

7. TOP 10 задач с попытками, но без правильных решений

На платформе могут быть задачи, которые студенты пытались решить, но у них не получилось. Проверить формулировки данных задач и подачу учебного материала по теме.

```
WITH  
total AS (  
SELECT  
problem_id,  
1 AS is_false  
FROM  
coderun cr  
UNION ALL  
SELECT  
problem_id,  
is_false  
FROM  
codesubmit cs
```

```

    )
SELECT
    total.problem_id,
    p.name,
    count(*) AS cnt,
    count(
        CASE
            WHEN is_false = 0 THEN 1
        END
    ) AS solved
FROM
    total
    JOIN problem_to_company ptc ON ptc.problem_id = total.problem_id
    JOIN problem p ON p.id = total.problem_id
WHERE
    ptc.company_id = 1
GROUP BY
    total.problem_id,
    p.name
ORDER BY
    solved
LIMIT 10

```

8. Прогресс обучения по студентам

Таблица с прогрессом обучения по каждому студенту.

```

WITH
    problems_count AS (
        SELECT
            company_id,
            count(*) AS all_amount
        FROM
            problem_to_company
        WHERE
            company_id IN (1)
        GROUP BY
            company_id
    )
SELECT
    u.email,
    u.date_joined,
    count(*) AS tries,
    count(
        DISTINCT CASE
            WHEN cs.is_false = 0 THEN cs.problem_id
        END
    ) AS solved_problems,
    pc.all_amount,
    round(

```

```

        count(
            DISTINCT CASE
                WHEN cs.is_false = 0 THEN cs.problem_id
            END
        ) * 100.0 / pc.all_amount
    ) AS progress
FROM
    codesubmit cs
    RIGHT JOIN users u ON u.id = cs.user_id
    JOIN company c ON c.id = u.company_id
    JOIN problems_count pc ON pc.company_id = u.company_id
WHERE
    c.id = 1
GROUP BY
    1,
    2,
    c.name,
    pc.all_amount
ORDER BY
    progress DESC

```

10. Студенты, не приступившие к учебе

Список студентов, которые не решили ни одной задачи, не предприняли ни одной попытки решения и не решили ни одного теста. Быть может у них технические сложности со входом на платформу, которые необходимо устранить.

```

SELECT
    u.username,
    u.email,
    u.date_joined::date
FROM
    users u
    LEFT JOIN codesubmit c ON u.id = c.user_id
    LEFT JOIN coderun cr ON u.id = cr.user_id
    LEFT JOIN teststart ts ON u.id = ts.user_id
WHERE
    c.id IS NULL
    AND cr.id IS NULL
    AND ts.id IS NULL
    AND u.company_id = 1

```

11. Задачи, к которым не приступали

Проверить список задач, видны ли они на платформе.

```

WITH
    agg AS (
        SELECT
            ptc.problem_id AS problem_id,

```

```

        p.name,
        count(cr.id) AS try,
        count(cs.id) AS solved
FROM
    problem_to_company ptc
    JOIN problem p ON p.id = ptc.problem_id
    FULL JOIN coderun cr ON cr.problem_id = ptc.problem_id
    FULL JOIN codesubmit cs ON cs.problem_id = ptc.problem_id
WHERE
    ptc.company_id = 1
GROUP BY
    1,
    2
ORDER BY
    try
)
SELECT
    problem_id
FROM
    Agg
WHERE
    try = 0
    AND solved = 0

```


ВЫВОДЫ

Итоговые результаты:

1. Общее количество студентов от компании — 133.
2. Прогресс обучения — в среднем составляет только 18,18%, что может говорить о низкой вовлеченности или сложностях при прохождении программы.
3. Активность — наблюдаются всплески посещаемости в ноябре и апреле, вероятно, из-за приема новых сотрудников или подготовки к промежуточным дедлайнам.
4. Результаты тестов — примерно 50/50 правильных и неправильных ответов, что может свидетельствовать о низком уровне подготовки, сложных или нечетко сформулированных вопросах, попытках угадывать.
5. Трудности с задачами — выявлены задачи с множеством попыток, но без правильных решений, а также задачи, к которым студенты вовсе не приступали.
6. Студенты, не приступившие к учебе — есть группа студентов, которые зарегистрировались, но не начали обучение, и причины этого стоит выяснить (возможно, технические сложности).

Основные предположения подтвердились:

- Низкий прогресс — действительно у студентов проблемы с темпом прохождения.
- Есть задачи, вызывающие особые трудности, подтверждено числом попыток и отсутствием правильных решений.
- Статистика заходов показывает неравномерную активность, что может свидетельствовать о временных проблемах или несвоевременной коммуникации.
- Имеются студенты, не приступившие к учебе.
- Технические проблемы — не подтвердились в массовом порядке (почти все аккаунты активированы), но могут присутствовать точечно.

Какие бизнес-решения принять и почему:

1. Автоматизация уведомлений и поддержки:

- Напоминания для студентов с низким прогрессом.
- Автооповещения о дедлайнах и помощь при застое.
- Сделать автоматические напоминания студентам, которые зарегистрировались, но не начали обучение.
- Предложить “спящим” студентам помощь или мини-опрос: сталкиваются ли они с техническими или содержательными трудностями?

2. Анализ и улучшение контента:

- Пересмотреть формулировки и структуру сложных задач.
- Добавить подсказки или промежуточные (подводящие) задачи для сложных задач.
- Провести ревизию тестовых вопросов: убрать неоднозначные или "ловушки".

3. Повышение вовлеченности:

- Внедрить геймификацию (бейджи, уровни, рейтинг).
- Мотивационные письма или сессии «вопрос-ответ» с преподавателями.
- Если обучение привязано к дедлайнам — оценить, хватает ли текущего темпа для завершения программы вовремя.
- В случае нехватки времени — предложить корректировку сроков или вводить этапы контроля прогресса.
- Ввести мониторинг студентов с нулевым или критически низким прогрессом, чтобы реагировать точно (наставники, менторы, поддержка).

4. Контроль доступа и интерфейса:

- Провести техаудит задач, к которым нет ни одной попытки — возможно, они просто не отображаются у студентов.
- Тестирование платформы с точки зрения UX — не слишком ли сложно найти или понять задание.

5. Разделение аудитории по уровням:

- Возможно, стоит предложить разные версии курса (начальный/продвинутый) в зависимости от стартового уровня.