

Контекст

Вы работаете с покерной платформой, на которой существует несколько типов разметок игроков:

1. Автоматические разметки:
 - Запускаются ежедневно через Airflow DAG.
 - Игроки получают запись в таблице, если попадают в разметку, и не получают, если не попадают.
2. Ручные разметки:
 - Формируются операторами вручную: при попадании добавляется новая строка в таблицу, если игрок больше не должен там быть - статус меняется на амнистирован
 - Данные хранятся в файлах, обработка которых также выполняется через Airflow DAG.

Примерный объем записей 100 тысяч по разметке в день.

Вопросы по тестовому заданию приветствуются.

База данных ClickHouse

Проблема

Разметки создавались разными командами в разное время и имеют различные форматы.

Задача

Разработать архитектуру сводной таблицы, объединяющей разметки в единую структуру.

Цели

1. Одним простым запросом понимать состояние игрока на конкретный момент времени - в каких разметках у него метка 1
2. Хранить историю игрока: когда попал в разметку, когда перестал попадать, когда попал снова

Примеры форматов данных

Авторазметки

DO ▾ | J2%

	A	B
1	DetectedAt	PlayerID
2	01.02.2025	1
3	01.02.2025	2
4	02.02.2025	1
5	02.02.2025	3
6	03.02.2025	1
7	03.02.2025	2
8		
9		
10		

Игрок 1 был в разметке все 3 дня 1, 2 и 3 февраля
Игрок 2 был 1 и 3 февраля
Игрок 3 только 2 февраля

Ручные

E	F	G
Date	PlayerID	Status
01.02.2025	1	Detected
01.02.2025	2	Detected
02.02.2025	2	Amnisted
02.02.2025	3	Detected
03.02.2025	2	Detected

У ручных разметок другая структура: один раз попавший в разметку игрок там и остается, пока не получит статус *Amnisted*.
Аналог предыдущей таблицы на скрине
Игрок 1 также в разметке все 3 дня с момента детекта 1 февраля
Игрок 2 попал в разметку 1 февраля, был удален из нее 2, снова попал 3
Игрок 3 в разметке с 3 февраля

Требуемый результат

1. Спроектируйте структуру сводной таблицы для объединения разметок.
2. Опишите логику заполнения и обновления таблицы.

Формат ответа

любой