

ЗАНЯТИЕ 3

Аналитические функции. ETL.



Александр Джумурат, онлайн кинотеатр «ivi.ru»



adzhumurat@ivi.ru



adzhumurat

Джумурат
Александр

Разработчик рекомендательной системы,
онлайн-кинотеатр ivi.ru

Цель занятия:

Аналитические(оконные) функции. Extract, Transform, Load.

В КОНЦЕ ЗАНЯТИЯ СМОЖЕТЕ:

- Строить сложные метрики с помощью оконных функций;
- Строить ETL-процессы средствами Postgres

О ЧЁМ ПОГОВОРИМ
И ЧТО СДЕЛАЕМ

1. Оконные (аналитические) функции:
основные принципы
2. Обогащение выборки
дополнительными данными
3. ETL-процедуры

Часть 1

Оконные функции (Window functions)

Оконные функции - способ считать сложные метрики внутри подгрупп

Эти функции не меняют число записей выборки, но добавляют дополнительные метрики.

ROW_NUMBER - проставляет номера строк для каждого элемента внутри окна, сортировка производится по полю **ORDER BY**

SUM - сумма значений поля внутри окна

Если в спецификацию окна добавить **ORDER BY** то простая сумма превращается в сумму с накоплением

COUNT() - счётчик элементов. Что случится, если добавить **ORDER BY** в спецификацию окна?

AVG() - среднее значение. Полезен, например, для нормировки

MIN() - минимальное значение внутри окна

MAX() - максимальное значение внутри окна

Обе функции часто используются вместе -
например, для нормировки.

RANK() - ранг для каждой строки, ранги таких строк с разрывами (gap)

DENSE_RANK() - ранги строиться без “разрывов”

PERCENT_RANK() - конвертирует абсолютные значения в доли

FIRST_VALUE() - первый элемент окна

LAST_VALUE() - последний элемент окна

NTH_VALUE() - элемент окна под номером n

LAG() - вычисляет элемент, который находится на фиксированном расстоянии внутри окна от текущего элемента

Пример

Оконные функции

Часть 2

ETL -Extract, Transform,
Load.

ETL - Extract, Transform, Load, общее название для потоков данных в системах где часть узлов данные производит, а часть читает

В Postgres импорт и экспорт данных отвечает команда `\copy`

Пример использования команды `\copy`

Самостоятельное изучение
Крутая статья на Хабре
про window functions.

ПОЛЕЗНЫЕ МАТЕРИАЛЫ

1. Оконные функции (осторожно: почти час)
2. Документация Postgres по window functions

ДОМАШНЕЕ ЗАДАНИЕ

[Ссылка на github](#)



НЕТОЛОГИЯ
групп

Спасибо за внимание!

Джумурат Александр



adzhumurat@ivi.ru

В