

SQL

Занятие 2

Агрегатные функции. Подзапросы.


Бояр Владислав

Агрегатные функции

Что такое агрегатная функция?

Агрегатная функция вычисляет единственное значение в результате обработки множества строк. Наиболее часто применяется в сочетании с группировкой данных для вычисления значений по каждой группе.

Наиболее популярные агрегатные функции:

- **count**(количество), **sum**(сумма), **max**(максимум), **min**(минимум)
 - **avg**(среднее), **mode** (мода)
 - **string_agg**, **array_agg**, **json_agg** – объединение данных
 - **stddev** (стандартное отклонение), **variance** (дисперсия), **percentile_cont** (перцентиль)
- 




Можно ли применять GROUP BY без агрегатной функции?



Можно ли применять агрегатную функцию без группировки?

Агрегатная функция и группировка

- Агрегатная функция без группировки применяется ко всему набору данных;
 - Группировка сокращает количество строк выборки до количества уникальных значений группируемых полей;
 - `count(*)` – вычисляет общее количество строк выборки
 - `count(column_name)` – вычисляет количество значений поля, за исключением NULL
- 

Место агрегатных функции в запросе

SELECT можно здесь — вычисление и вывод агрегационных значений

FROM ...

WHERE ...

GROUP BY ...

HAVING здесь — фильтрация данных после группировки

ORDER BY и здесь — сортировка на основе агрегационных значений

Подзапросы

Что такое подзапрос?

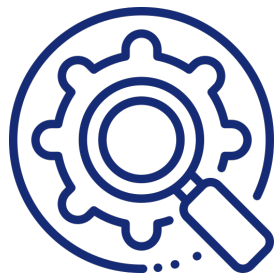
Подзапросы (subqueries) - это запросы, которые вложены в другие запросы.

Результатом подзапроса может быть одно значение или целая таблица.

Подзапросы используются для выполнения фильтрации, сравнения и вычисления значений.



Практика



Практика

1. Вывести уникальные направления перелётов (аэропорт отправления – аэропорт прибытия) двумя способами;
2. Найти аэропорт с наибольшей географией отправления;
3. Посчитать для таблицы с билетами (ticket_flights):
 1. Количество билетов каждого класса;
 2. Сумму, максимум, минимум, среднее, медиану, моду, стандартное отклонение и дисперсию стоимости билетов каждого класса;
4. Найти минимальную, среднюю и максимальную продолжительность полета для каждой модели самолёта. Формат вывода: «часы:минуты»;
5. Найти аэропорты (departure_airport), где количество рейсов превышает среднее количество рейсов по всем аэропортам;
6. Вывести номер бронирования, количество и список пассажиров в бронировании в виде массива. Отсортировать по количеству пассажиров в обратном порядке