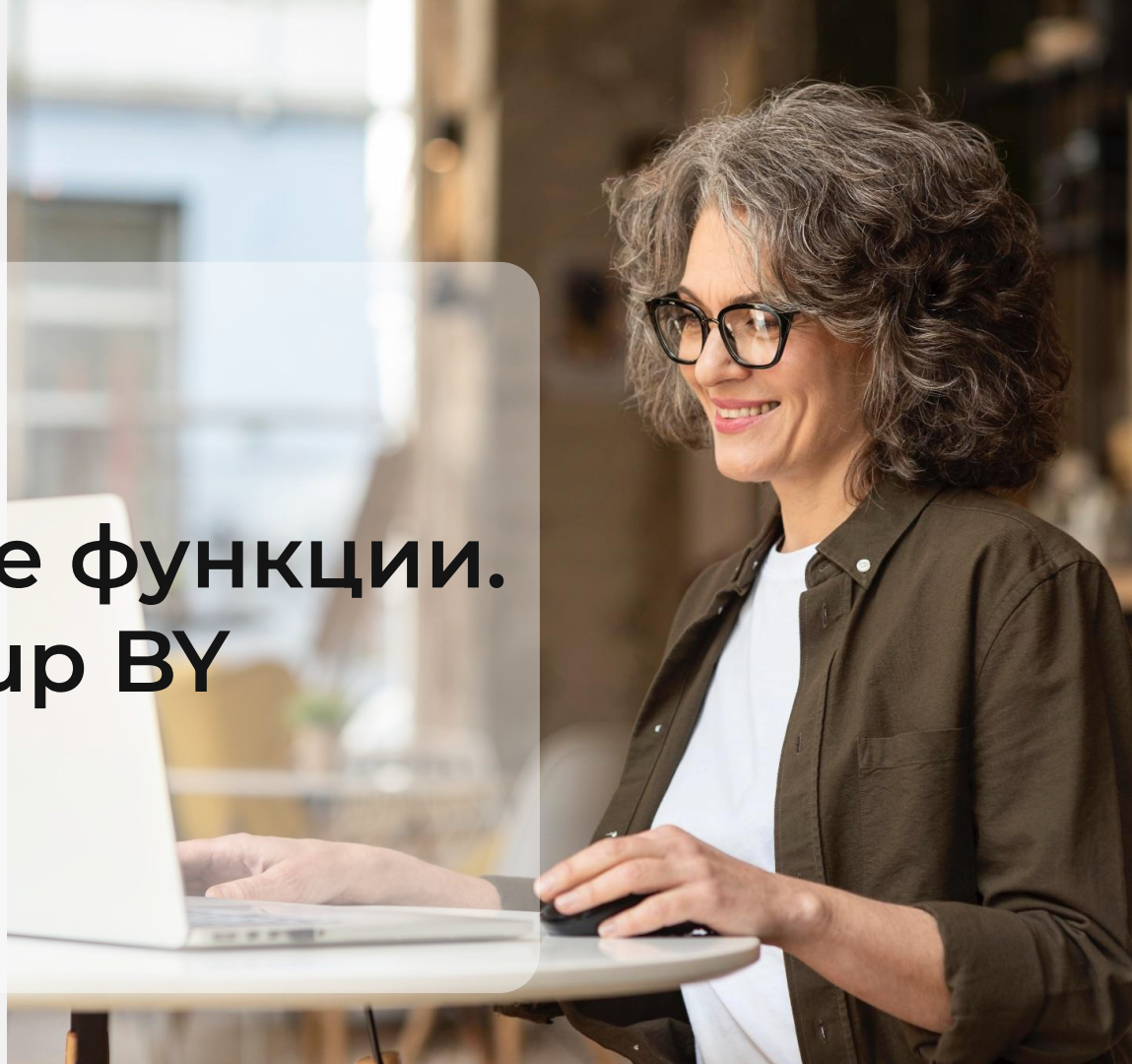


Database

Агрегирующие функции. Оператор Group BY



Преподаватель

Портрет

Имя Фамилия

Текущая должность

Количество лет опыта

Какой у Вас опыт - ключевые кейсы

Самые яркие проекты

Дополнительная информация по вашему усмотрению

Корпоративный e-mail

Социальные сети (по желанию)

Важно

-  Камера должна быть включена на протяжении всего занятия
-  В течение занятия вопросы задавать в чате или когда преподаватель спрашивает, есть ли у Вас вопросы
-  Вести себя уважительно и этично по отношению к остальным участникам занятия
-  Организационные вопросы по обучению решаются с кураторами, а не на тематических занятиях
-  Во время занятия будут интерактивные задания, будьте готовы включить камеру или демонстрацию экрана по просьбе преподавателя

Повторение



Связь между Python, SQL и BI-инструментами



Процесс загрузки данных, создания витрины и подключения к BI-инструменту



Представления (Views) в SQL

План занятия

- Агрегирующие функции Оператор Group BY
- Основные агрегирующие функции
- Оператор GROUP BY



ОСНОВНОЙ БЛОК





**Агрегирующие
функции. Оператор
Group BY**



Агрегация

Это процесс объединения данных, чтобы получить общее представление о наборе данных.

Необходимость агрегации



Агрегация помогает обобщить большие объемы данных и извлечь из них полезную информацию.



Агрегирующие функции

Это функции, которые позволяют производить вычисления над набором строк, возвращая одно итоговое значение.

Операторы сравнения

COUNT()

Подсчитывает количество строк в наборе данных.

Можно использовать COUNT(column_name) для подсчета строк (не NULL).

SUM()

Вычисляет сумму значений в столбце.

Работает только с числовыми типами данных.

AVG()

Вычисляет среднее значение по столбцу.

Игнорирует NULL значения при вычислении среднего.

MIN()

Определяет минимальное значение в столбце.

Работает с числовыми, строковыми и датами.

MAX()

Определяет максимальное значение в столбце.

Работает с числовыми, строковыми и датами.



GROUP_CONCAT

Это специфическая агрегатная функция в MySQL, которая объединяет значения из нескольких строк в одну строку с разделителем.

GROUP_CONCAT



Агрегатные функции игнорируют NULL значения



Можно использовать DISTINCT внутри агрегатных функций, чтобы учитывать только уникальные значения

Подсчитывает количество уникальных клиентов в таблице orders.



```
SELECT COUNT(DISTINCT customer_id) AS unique_customers  
FROM orders;
```



ВОПРОСЫ





ЗАДАНИЕ



Задание

Найдите общее количество товаров quantity в таблице order_details.



Решение

```
SELECT SUM(quantity) FROM order_details
```



Задание

Посчитайте количество уникальных `order_id` в таблице `order_details`.



Решение

```
SELECT COUNT(DISTINCT order_id) FROM
order_details
```



Задание

Перечислите через запятую имена всех сотрудников из таблицы employees.



Решение

```
SELECT GROUP_CONCAT(first_name) FROM
employees
```



Задание

Выведите среднее, минимум и максимум столбца `unit_price` таблицы `order_details`.



Решение

```
SELECT AVG(unit_price), MIN(unit_price),  
MAX(unit_price)  
  
FROM order_details
```





ВОПРОСЫ





Оператор GROUP BY



GROUP BY

Это оператор SQL, который используется для группировки строк, имеющих одинаковые значения в одном или нескольких столбцах.

Использование GROUP BY



SQL Server группирует строки по указанным столбцам.



Все строки с одинаковыми значениями в этих столбцах объединяются в одну группу.



Агрегирующая функция применяется к каждой группе отдельно.

Оператор GROUP BY



Определение

- column_name — столбец, по которому производится группировка.
- AGGREGATE_FUNCTION(column_name) — агрегирующая функция, применяемая к каждой группе.

Код

```
SELECT column_name,  
AGGREGATE_FUNCTION(column_name)  
FROM table_name  
GROUP BY column_name;
```



ВОПРОСЫ





ЗАДАНИЕ



Задание

1. Из таблицы employees посчитать количество сотрудников в каждом городе city.
2. Отсортировать результаты по убыванию.
3. Посчитать общее количество продуктов из таблицы order_details для каждого заказа.
4. Отсортировать по убыванию общего количества продуктов. Для краткости записи в GROUP BY можно не указывать конкретное имя колонки, а указать ее порядковый номер в SELECT.
5. Сделать то же самое в ORDER BY.

Задание

6. Посчитать сколько сотрудников работает в каждой компании из таблицы customers. Учитывать только тех сотрудников, у которых job_id равен 'Purchasing Manager'.
7. Если столбец, который вы хотите использовать для группировки содержит только уникальные неповторяющиеся значения, то в группировке нет смысла - любая агрегатная функция даст один и тот же результат. Попробуйте сгруппировать любую таблицу по первичному ключу и применить агрегатные функции к столбцам.
8. Посчитать количество сотрудников в разрезе компании и занимаемой должности из таблицы employees.



ВОПРОСЫ



Заключение

