

## Задание 9 "Основы СУБД (3)"

### 1. Способ распределения (DISTRIBUTED BY () / RANDOMLY)

```
DISTRIBUTED BY (year, month)
```

Обоснование:

- Данные анализируются по времени → частые фильтры по `year`, `month`
- Хэширование по `year`, `month` обеспечит:
  - Равномерное распределение между сегментами
  - Эффективность JOIN'ов с другими таблицами по дате
  - Минимизацию перемещения данных (data redistribution)

Если `year` и `month` дают мало уникальных комбинаций — можно добавить `channel_name`.

```
DISTRIBUTED BY (year, month, channel_name)
```

### 2. Разбиение на разделы (PARTITION BY RANGE / LIST / HASH)

```
PARTITION BY RANGE (year)
(
    START (2020) END (2031) EVERY (1),
    DEFAULT PARTITION other
);
```

Обоснование:

- Данные естественным образом группируются по году
- Запросы часто фильтруют по году или диапазону лет
- При удалении старых данных — можно просто `DROP PARTITION`
- Ускоряет выполнение запросов (partition pruning)