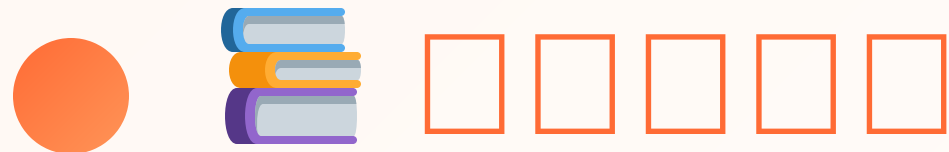


6 -

GROUP BY





1. GROUP BY

1. GROUP BY

1. GROUP BY

1. GROUP BY

1. GROUP BY

2. □□□□

```
SELECT  
    □□□□□□□□ ,  
    □□□□  
FROM □□□□  
GROUP BY □□□□□□□□ ;
```

3. □□□□□□

```
SELECT□□□□□GROUP BY□□□□□□□□□□□□□□□  
GROUP BY□□□□□COUNT□SUM□AVG□□□□□□□□□□  
□□□□□□□□□□□□□□
```

4. □□□□□□□

□□□□□

customer_id	quantity
C001	5
C001	10
C002	2
C002	1

GROUP BY customer_id □□□□

customer_id	SUM(quantity)
C001	15
C002	3



1 ID

```
SELECT
  customer_id,
  COUNT(*) AS
FROM 'data/sales.csv'
GROUP BY customer_id;
```

□ □ □ □ □ □ □

```
SELECT
    customer_id,
    COUNT(*) AS □□□□
FROM 'data/sales.csv'
GROUP BY customer_id
ORDER BY □□□□ DESC;
```


□□2□□□ID□□□□□□□□□□

```
SELECT
    product_id,
    SUM(quantity) AS □□□□□
FROM 'data/sales.csv'
GROUP BY product_id;
```

□□□□□□□□□□□□□□

□□3□□□□□□□□TOP3□□□

```
SELECT
    product_id AS □□ID,
    SUM(quantity) AS □□□□□
FROM 'data/sales.csv'
GROUP BY product_id
ORDER BY □□□□□ DESC
LIMIT 3;
```

□□□□□□□

GROUP BY□□□□□□□□

ORDER BY□□□□□□□□□

LIMIT□□□3□□□□□



1. □□□□□□□□

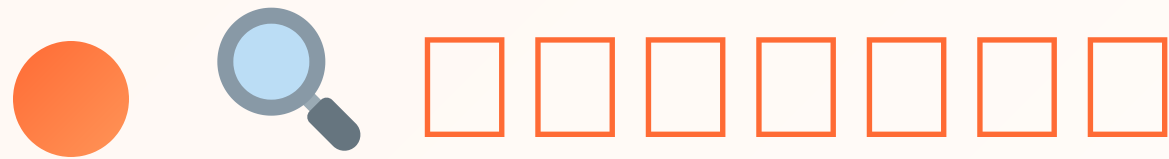
```
SELECT
  customer_id AS □□ID,
  COUNT(*) AS □□□□,
  SUM(quantity) AS □□□□,
  AVG(quantity) AS □□□□□
FROM 'data/sales.csv'
GROUP BY customer_id;
```

2. GROUP BY

```
-- 5이상인 고객별 판매량 집계
SELECT
    customer_id,
    COUNT(*) AS 고객수,
    SUM(quantity) AS 총판매량
FROM 'data/sales.csv'
WHERE quantity >= 5
GROUP BY customer_id;
```

3. □□□□□□□□

```
-- □□□□□□□□  
SELECT  
    order_date AS □□□,  
    COUNT(*) AS □□□□,  
    SUM(quantity) AS □□□□  
FROM 'data/sales.csv'  
GROUP BY order_date  
ORDER BY order_date;
```



1

```
-- 3
SELECT
  customer_id AS ID,
  COUNT(*) AS ,
  SUM(quantity) AS
FROM 'data/sales.csv'
GROUP BY customer_id
HAVING COUNT(*) >= 3
ORDER BY  DESC;
```

□□□2□□□□□□□□□□

```
-- □□□□□□□□□□  
SELECT  
    category AS □□□□,  
    COUNT(*) AS □□□,  
    AVG(price) AS □□□□,  
    MAX(price) AS □□□□  
FROM 'data/products.csv'  
GROUP BY category;
```

● ★ GROUP BY□□□□□

□□□□□□□□□□□□

1. □□□□□□□□

```
SELECT customer_id, quantity
FROM 'data/sales.csv'
ORDER BY customer_id;
```


2. GROUP BY□□□

```
SELECT
    customer_id,
    SUM(quantity) AS □□
FROM 'data/sales.csv'
GROUP BY customer_id;
```

→ □□□□□□□□□□□□□□□□□□

□ □ □ □ □ □ □ □ □ □ □ □



□ □ □ □ □ □ □ □

```
```sql -- product_id□GROUP BY□□□□ SELECT customer_id, product_id, -- □□□□  
□□□□□ COUNT(*) FROM 'data/sales.csv' GROUP BY customer_id; ```
```



□□□□□

```
```sql SELECT customer_id, COUNT(*) AS □□□□ FROM 'data/sales.csv' GROUP  
BY customer_id; ```
```



GROUP BY

1.

```
-- 
SELECT
    product_id,
    SUM(quantity) AS 
FROM 'data/sales.csv'
GROUP BY product_id
ORDER BY  DESC;
```

2. □□□□

```
-- □□□□□□□□□□□□  
SELECT  
    SUBSTR(order_date, 1, 7) AS □□,  
    COUNT(*) AS □□  
FROM 'data/sales.csv'  
GROUP BY SUBSTR(order_date, 1, 7);
```

3. □□□□□□□□□□

```
-- □□×□□□□□□□□  
SELECT  
    customer_id,  
    product_id,  
    SUM(quantity) AS □□□  
FROM 'data/sales.csv'  
GROUP BY customer_id, product_id;
```



GROUP BY



HAVING



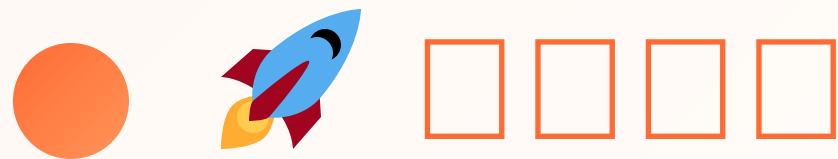
WHEREHAVING





```
-- 表
SELECT
    表名,
    COUNT(*),
    SUM(列名),
    AVG(列名)
FROM 表名
GROUP BY 表名;

-- 表名
SELECT
    表名,
    表名 AS 表名
FROM 表名
GROUP BY 表名
ORDER BY 表名 DESC
LIMIT 10;
```

7 INNER JOIN
ID



1 GROUP BY

```
-- 1. ID
-- 
-- 2. 
-- 
-- 3. 
-- 
```

□□2□□□□□□□□

GROUP BY□□□□□□□□□□□□□□

```
-- □□□□□□□□□□  
SELECT  
    customer_id,  
    ____ AS □□□□,  
    ____ AS □□□□,  
    ____ AS □□□□□,  
    ____ AS □□□□□,  
    ____ AS □□□□□  
FROM 'data/sales.csv'  
GROUP BY ____;
```

□□3□HAVING□□□□

□□□□□□□□□□□□

```
-- 1. 2□□□□□□□□□□  
SELECT product_id, COUNT(*) as □□□□  
FROM 'data/sales.csv'  
GROUP BY product_id  
HAVING ____ >= ____;
```

```
-- 2. □□□□□15□□□□□□  
-- □□□□□□□□□□□□□□□□
```

```
-- 3. □□□□□3□□□□□□□□□□  
-- □□□□□□□□□□□□□□□□
```

4

```
-- 1. ABC
-- 20%
--
```

□□4□□□□

```
-- 2. □□□□□□□□  
-- □□□□□□□□□□□□□□1□□□□□□  
SELECT  
    CASE  
        WHEN SUM(quantity) * 10000 < 50000 THEN '□□□'  
        WHEN SUM(quantity) * 10000 < 200000 THEN '□□□'  
        ELSE '□□□'  
    END AS □□□,  
    COUNT(*) AS □□□  
FROM 'data/sales.csv'  
GROUP BY □□□;
```



```
-- □□□□GROUP BY□  
-- □□×□□□□□□□□□□□□□□□□  
-- □□□□□□□□□□□□□□□□□□□□
```



```
- - [][][][]
- - 1. [][][][]
- -   - [][][][]
- -   - [][][][]
- -   - [][][][]

- - 2. [][][][]
- -   - [][][][]
- -   - [][][][]
- -   - [][][][]DISTINCT[]
```




□□□□□□□□□□□□□□□□

```
-- □□□1□GROUP BY□□□□□□□□  
SELECT customer_id, product_id, COUNT(*)  
FROM 'data/sales.csv'  
GROUP BY customer_id;
```

```
-- □□□2□□□□□□□□□GROUP BY  
SELECT customer_id  
FROM 'data/sales.csv'  
GROUP BY customer_id;
```

```
-- □□□3□HAVING□WHERE□□□□  
SELECT customer_id, COUNT(*) as cnt  
FROM 'data/sales.csv'  
HAVING customer_id = 'C001'  
GROUP BY customer_id;
```



GROUP BY

```
-- 1. 고객별 평균 구매량 구하기
-- 고객별 평균 구매량 구하기
WITH avg_purchase AS (
    SELECT AVG(quantity) as avg_qty FROM 'data/sales.csv'
)
SELECT customer_id, AVG(quantity) as avg_qty
FROM 'data/sales.csv'
GROUP BY customer_id
HAVING AVG(quantity) > (SELECT avg_qty FROM avg_purchase);
```



```
-- 2. 
-- 
-- 
```



```
-- 
-- 1. TOP3
-- 2. TOP3
-- 3. 
-- 4. 
-- 
```



```
-- □□×□□□□□□□□□□
-- □□□□□□□□□□□□□□
SELECT
  customer_id,
  SUM(CASE WHEN order_date = '2024-01-15' THEN quantity ELSE 0 END) AS "1□15□",
  SUM(CASE WHEN order_date = '2024-01-16' THEN quantity ELSE 0 END) AS "1□16□",
  SUM(CASE WHEN order_date = '2024-01-17' THEN quantity ELSE 0 END) AS "1□17□"
  -- □□□□□□□□□□
FROM 'data/sales.csv'
GROUP BY customer_id;
```

● ? FAQ

Q: WHERE和HAVING的区别

A: WHERE = GROUP BY前使用 HAVING = GROUP BY后使用

Q: GROUP BY和SELECT的区别

A: GROUP BY用于分组 SELECT用于查询

Q: 左连接和右连接的区别

A: LEFT JOIN左表为基准 RIGHT JOIN右表为基准