

kokchun giang

**grouping data** to  
arrange identical data  
into group to do  
meaningful  
aggregations

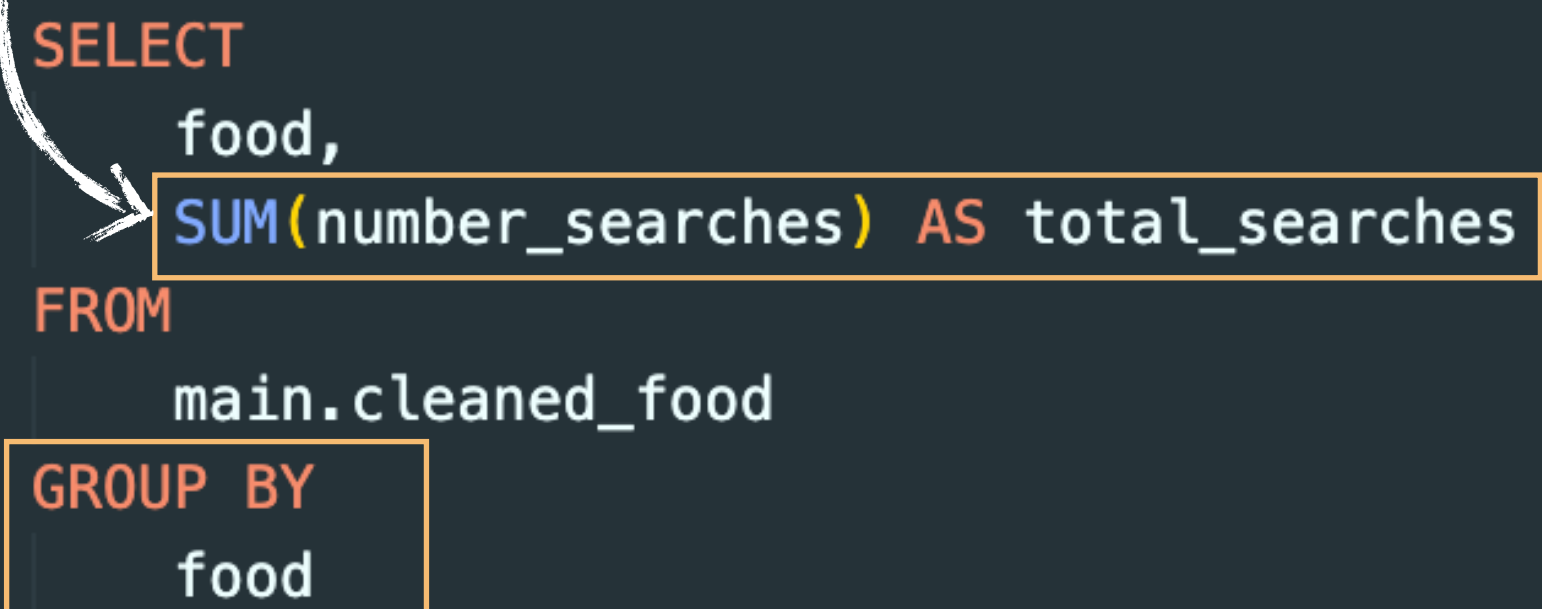




# group by columns

combine with  
aggregate function

```
SELECT  
  food,  
  SUM(number_searches) AS total_searches  
FROM  
  main.cleaned_food  
GROUP BY  
  food
```



groups by  
food columns

the **HAVING** clause to filter group

```
SELECT
    YEAR,
    SUM(number_searches) AS total_searches
FROM
    main.cleaned_food
GROUP BY
    YEAR
HAVING
    total_searches > 300000;
```

condition to filter on after  
aggregation has been  
performed

```
SELECT
    YEAR,
    food,
    SUM(number_searches) AS total_searches
FROM
    main.cleaned_food
GROUP BY
    food,
    YEAR
HAVING
    food = 'pizza';
```

groups by two  
columns, food  
together with year

# duckdb **order of executions**

1. FROM
2. WHERE
3. GROUP BY
4. HAVING
5. SELECT
6. DISTINCT
7. ORDER BY
8. LIMIT/OFFSET

```
4 SELECT  
    YEAR,  
    food,  
    SUM(number_searches) AS total_searches  
1 FROM  
    main.cleaned_food  
2 GROUP BY  
    food,  
    YEAR,  
3 HAVING  
    food = 'pizza';
```