

# Database Report on

## Generating SQL Queries by Large Language Models

### Due on July 3rd, 2024

This report is to explore SQL code generation capabilities of large language models, from natural language query description.

#### Question.

Consider relations of the following restaurant database schema:

MenuItem (mid, mname, mtype, price),    Guest(gid, tableid),  
Order(oid, gid, mid, quantity, date)

(query descriptions)

Q1. Find mid of menuitems such that the number of guests who ordered the menuitem only once is maximum. Report such mid and the maximum number of guests.

Q2. Find guests who always order an item that is most expensive within the table.

Q3. Find the golden guest pair such that their total amount in two visits is maximum than any other guest pairs. Guest pairs are any two guests coming on the same day and occupy one table by the two.

(1) For each of the above query descriptions Q1, Q2, Q3, generate SQL queries by using ChatGPT. Show the generated queries.

(2) For the queries generated from Q1, Q2, Q3.

(2-1) Argue whether each generated query is correctly expressing its query description. Construct a table instance for explaining your answer. If the generated query is wrong, point out its errors, and show a correct SQL query that returns the expected answer for the query description.

(2-2) Revise the natural language prompt so that the correct SQL query or its equivalent query, for the query description is generated.

**Prompt example:** ChatGPT <https://chatgpt.com/>

You are an experienced SQL programmer. Consider relations of the following restaurant database schema:

MenuItem (mid, mname, mtype, price), Guest(gid, tableid),  
Order(oid, gid, mid, quantity, date)

Now write an SQL query Report gid of guests who did not order “Ramen” and “Udon” on the same date.

You are an experienced SQL programmer. Consider tables of the following restaurant database schema:  
Menuitem (mid, mname, mtype, price), Guest(gid, tableid),  
Order(oid, gid, mid, quantity, date)  
Now write an SQL query Report gid of guests who did not order “Ramen” and “Udon” on the same date.



To report the `gid` of guests who did not order both "Ramen" and "Udon" on the same date, you can use the following SQL query:

```
sql 📄 コードをコピーする  
  
SELECT DISTINCT g.gid  
FROM Guest g  
WHERE g.gid NOT IN (  
    SELECT o.gid  
    FROM Order o  
    JOIN MenuItem m1 ON o.mid = m1.mid  
    JOIN MenuItem m2 ON o.mid = m2.mid  
    WHERE m1.mname = 'Ramen' AND m2.mname = 'Udon' AND o.date = o.date  
);
```

Here is an explanation of the query:

