How to Prompt LLMs for Text-to-SQL: A Study in Zero-shot, Single-domain, and Cross-domain Settings

Shuaichen Chang* and **Eric Fosler-Lussier**
The Ohio State University

Table Representation Learning Workshop @ NeurIPS 2023

* Now at AWS AI Lab

Text-to-SQL

LLMs for Text-to-SQL with In-context Learning

Accuracy in the Spider dataset
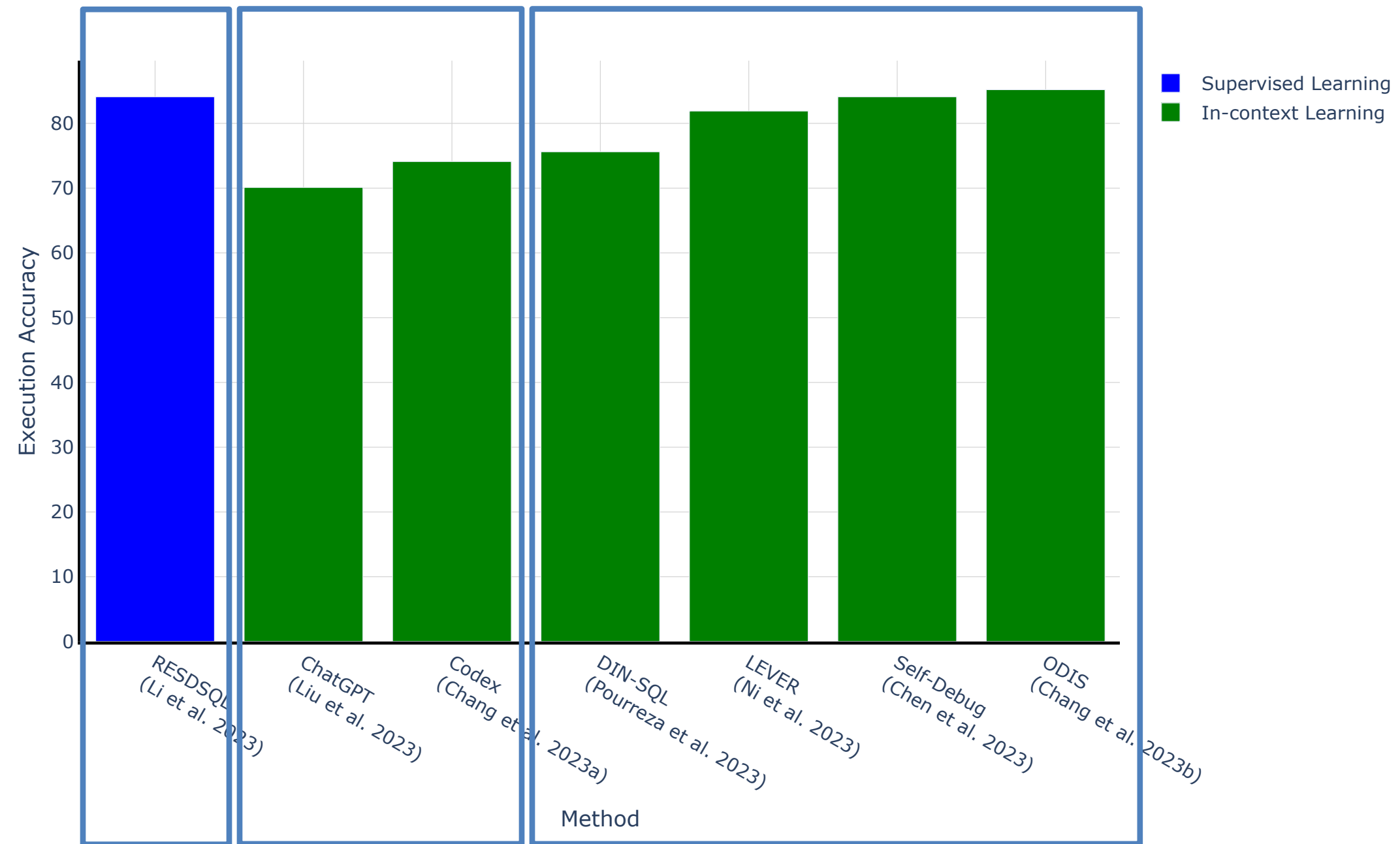
Trained on 7K examples   Zero-shot Prompt   Few-shot Prompt (< 32 demonstrations)
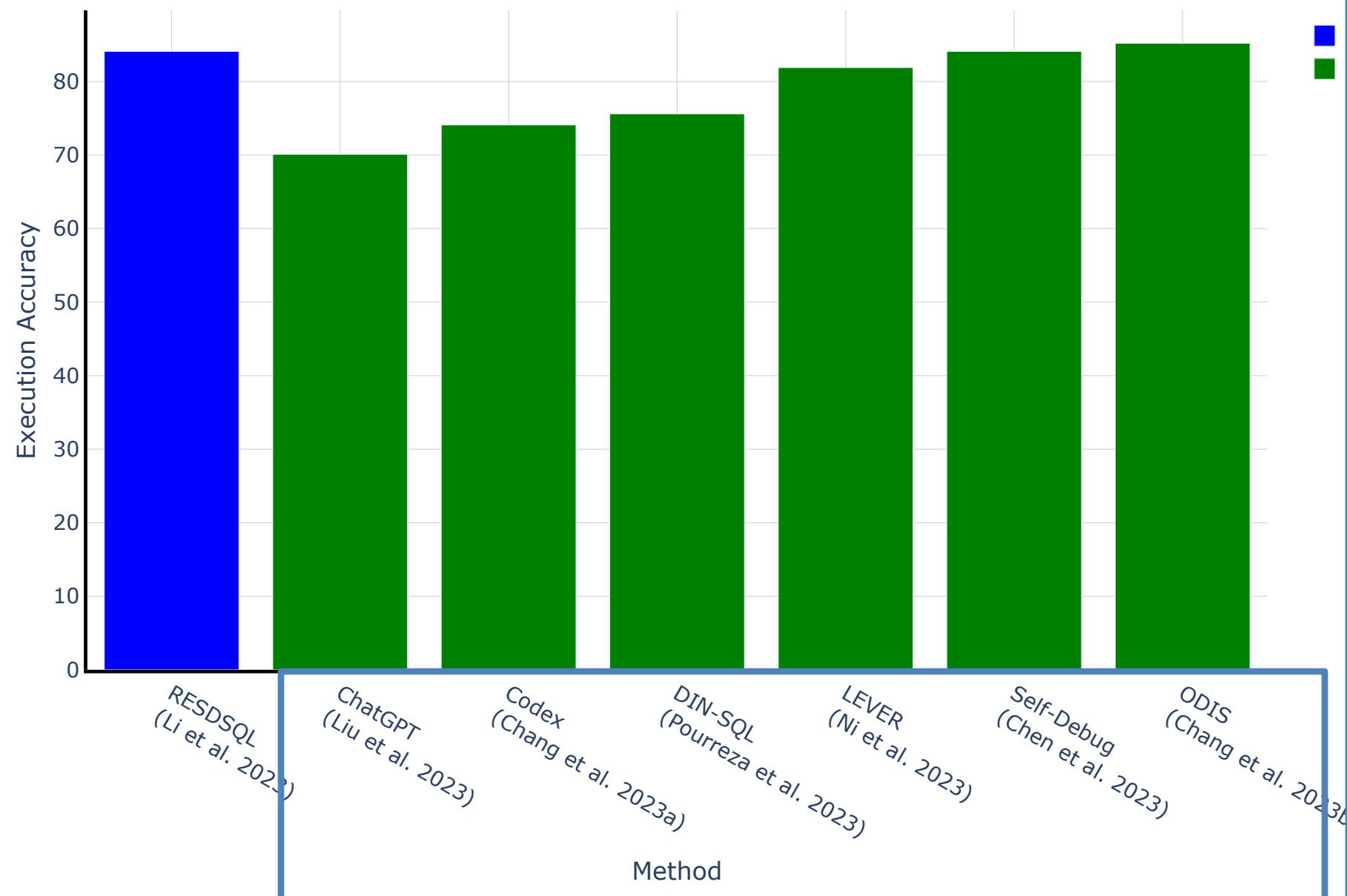
LLMs for Text-to-SQL with In-context Learning

- These studies enhance the LLMs' performance with different approaches: demonstration retrieval or intermediate reasoning (their main proposals)

- However, they also employ different strategies for constructing the prompt text of databases and demonstrations
  - It's hard to compare two work on their main contributions, or prompt constructions
  - Future work still has to explore the effective prompt constructions for Text-to-SQL (prompt engineering)

LLMs for Text-to-SQL with In-context Learning

- Our goals:
  - Study how to represent structured database in unstructured prompt text for LLMs in three common scenarios:
    - Zero-shot text-to-SQL
    - Few-show single-domain text-to-SQL
    - Few-show cross-domain text-to-SQL
  - Study how to construct demonstrations for cross-domain text-to-SQL

Zero-shot Text-to-SQL Prompt

Zero-shot Prompt

Database
```
CREATE TABLE Highschooler (
ID int primary key,
name text,
grade int
);
/*
3 example rows:
SELECT * FROM Highschooler LIMIT 3;
ID      name      grade
1510     Jordan      9
1689     Gabriel     9
1381     Tiffany     9
*/
```

Database Schema

Database Content

Task Instruction
```
-- Using valid SQLite, answer the following questions for the tables
 provided above.
```

Test Question
```
Question: How many high schoolers are there?
SELECT
```

Zero-shot Text-to-SQL Results

- Table relationship and database content are crucial, however, it requires careful representation in the prompt.

		Codex		ChatGPT	
Database Prompt Construction		**# Tokens**	**EX**	**# Tokens**	**EX**
Table Schema	Table(Columns)	147	71.9	115	70.5
	Columns=[]	167	71.8	135	69.1
+Relationship	Columns=[]+ForeignKey	223	73.1	174	71.2
	CreateTable	356	73.1	254	71.7
+Relationship+Content	CreateTable+InsertRow 3	1013	71.9	872	71.8
	CreateTable+SelectRow 3	770	74.1	674	72.1
	CreateTable+SelectCol 3	831	**75.7**	712	**73.6**

Single-domain Text-to-SQL Prompt Example

Few-shot Single-domain Prompt

Database

```
CREATE TABLE Highschooler (
ID int primary key,
name text,
grade int
);
/*
3 example rows:
SELECT * FROM Highschooler LIMIT 3;
ID     name       grade
1510     Jordan      9
1689     Gabriel     9
1381     Tiffany     9
*/
```

Task Instruction

```
-- Using valid SQLite, answer the following questions for the tables
 provided above.
```

Demonstration

```
Question: What is Kyle's id?
SELECT ID FROM Highschooler WHERE name = "Kyle";
```

Test Question

```
Question: How many high schoolers are there?
SELECT
```

In-domain demonstrations are annotated NLQ and SQL pairs from the same database as the test question.

Single-domain Text-to-SQL Results

- The performance of LLMs continues to enhance as the number of in-domain examples increases.
- Table relationship becomes less important when having in-domain examples, however, database content is still important.
- LLMs are robust to database content representation when having in-domain examples.
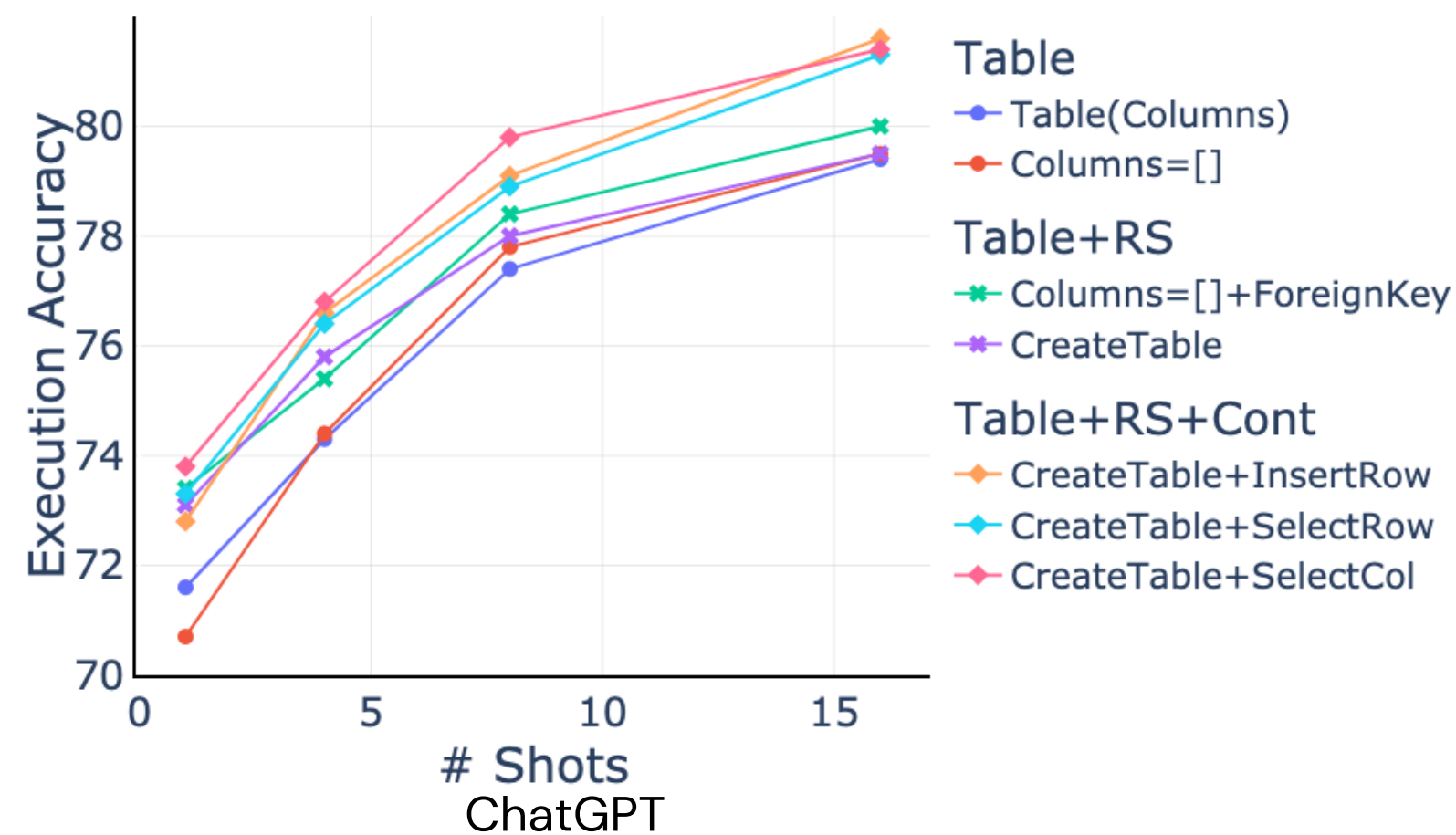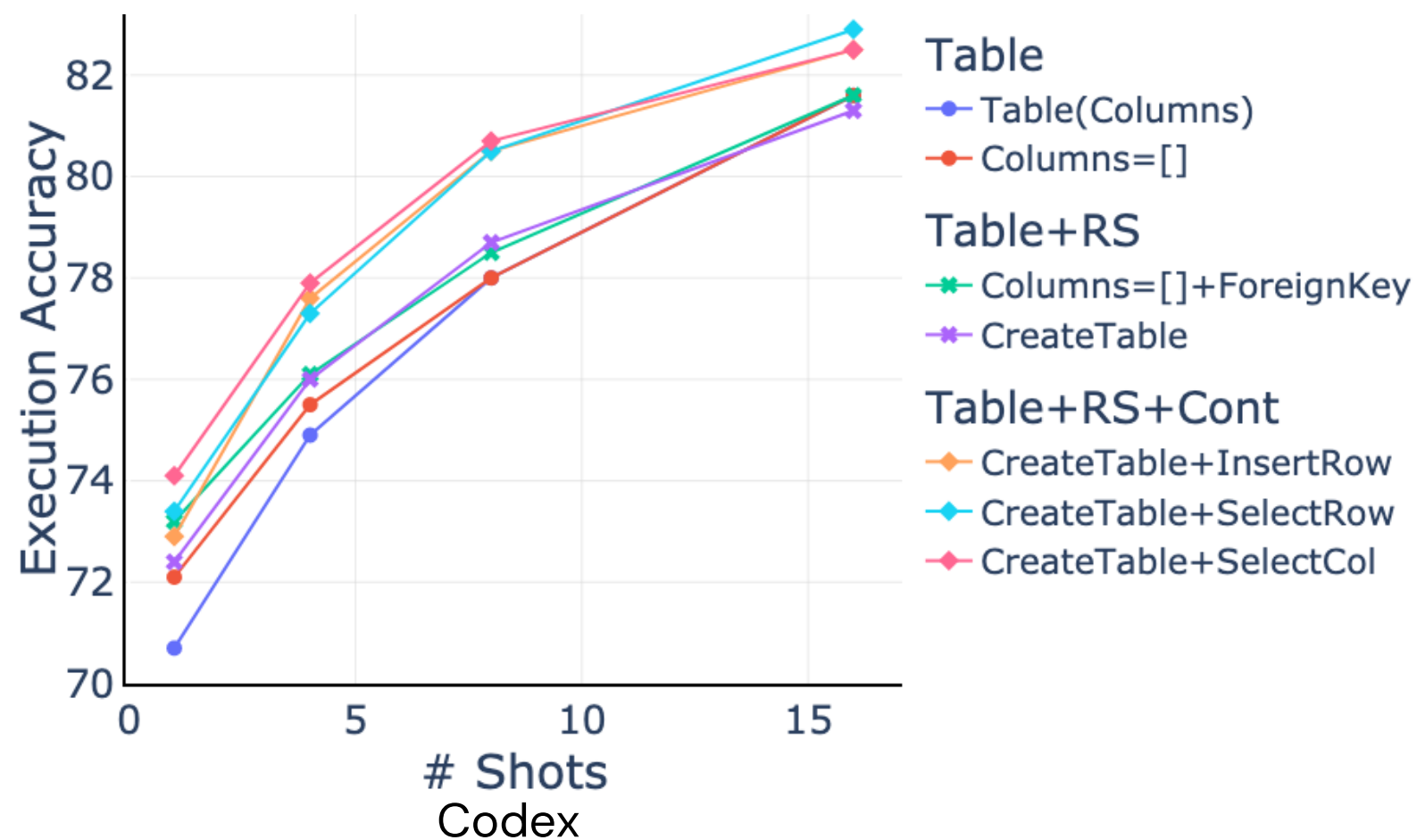
* Demonstration examples were randomly selected and the experiments were repeated three times.

Cross-domain Text-to-SQL Prompt Example

Few-shot Cross-domain Prompt

Demonstration
```
create table track (
track_id int,
name text,
location text,
seating real,
year_opened real,
primary key (track_id)
);
/*
3 example rows:
select * from track limit 3;
track_id    name        location        seating     year_opened
1       Auto Club Speedway      Fontana, CA     92000.0     1997.0
2       Chicagoland Speedway        Joliet, IL      75000.0     2001.0
3       Darlington Raceway      Darlington, SC      63000.0     1950.0
*/

-- Using valid SQLite, answer the following questions for the tables
   provided above.
Question: Show the name and location for all tracks.
select name, location from the track;
```

Database
```
CREATE TABLE Highschooler (
ID int primary key,
name text,
grade int
);
/*
3 example rows:
SELECT * FROM Highschooler LIMIT 3;
ID      name        grade
1510        Jordan      9
1689        Gabriel     9
1381        Tiffany     9
*/
```

Task Instruction
```
-- Using valid SQLite, answer the following questions for the tables
   provided above.
```

Test Question
```
Question: How many high schoolers are there?
SELECT
```

Out-of-domain demonstrations are the NLQ and
SQL pairs along with their respective databases,
which are different from the test database.

Cross-domain Text-to-SQL Results

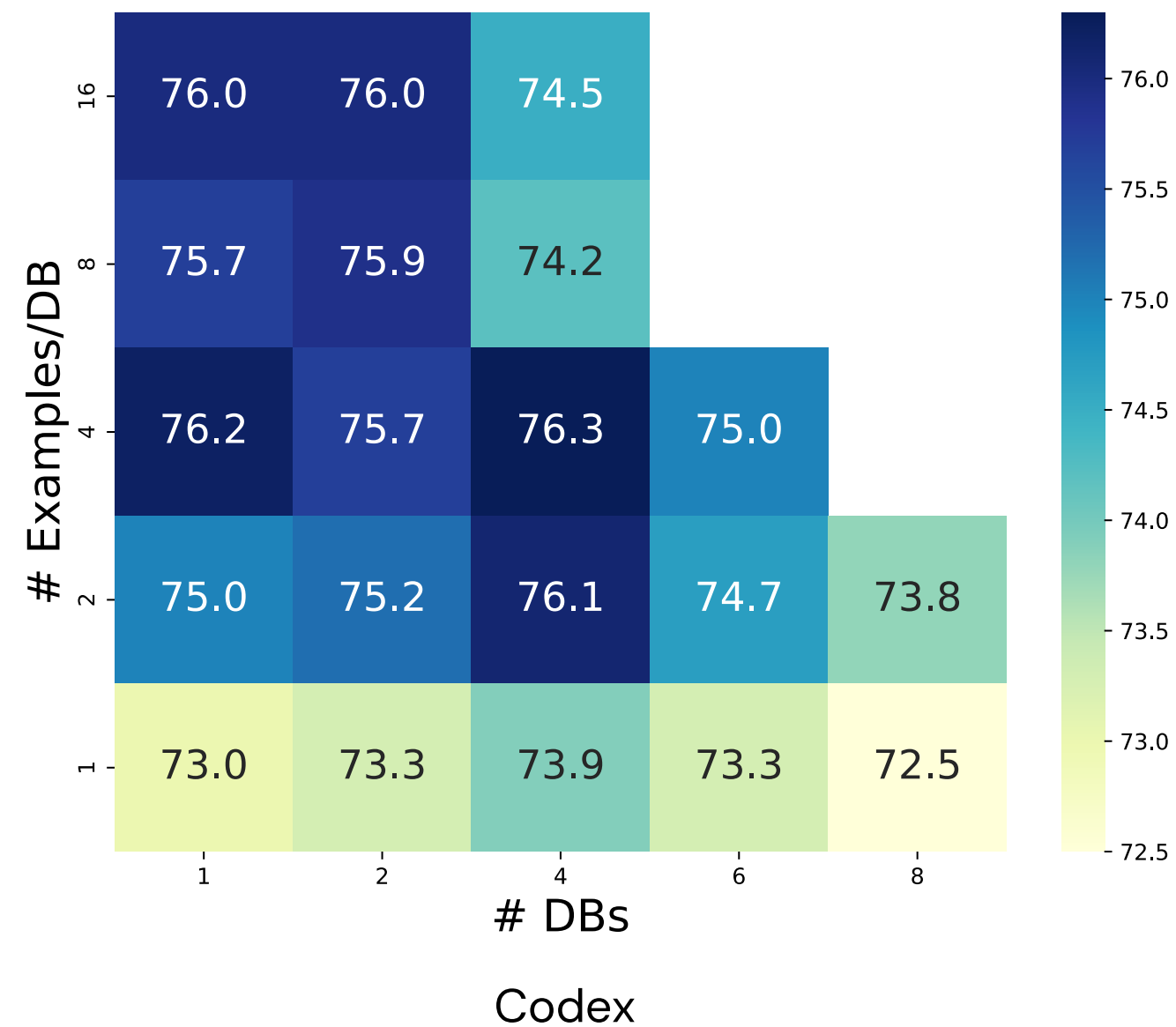- Can out-of-domain demonstrations enhance the performance of LLMs?

Cross-domain Text-to-SQL Results

- We consider a general scenario where the demonstrations contains M databases, each with K examples.
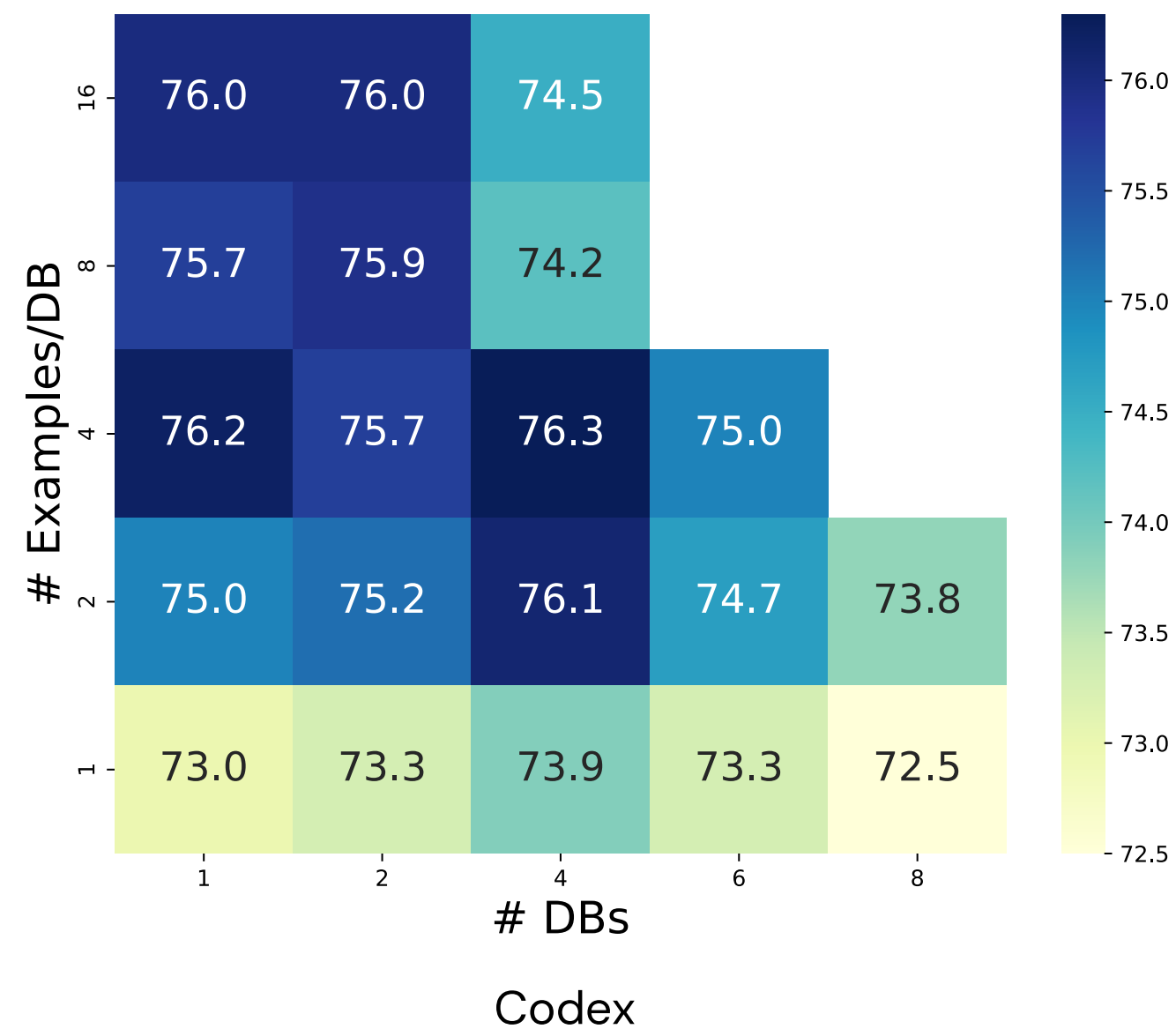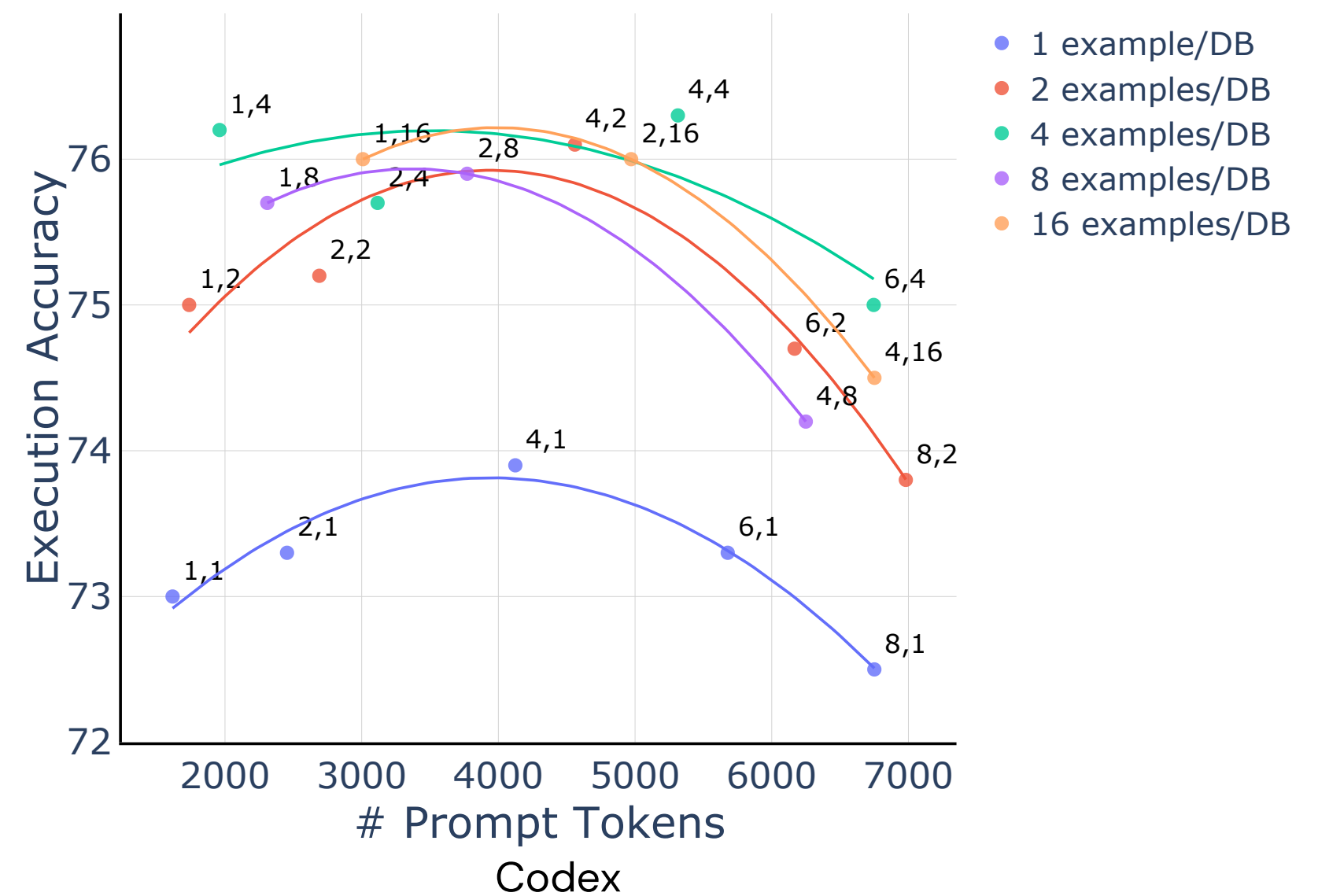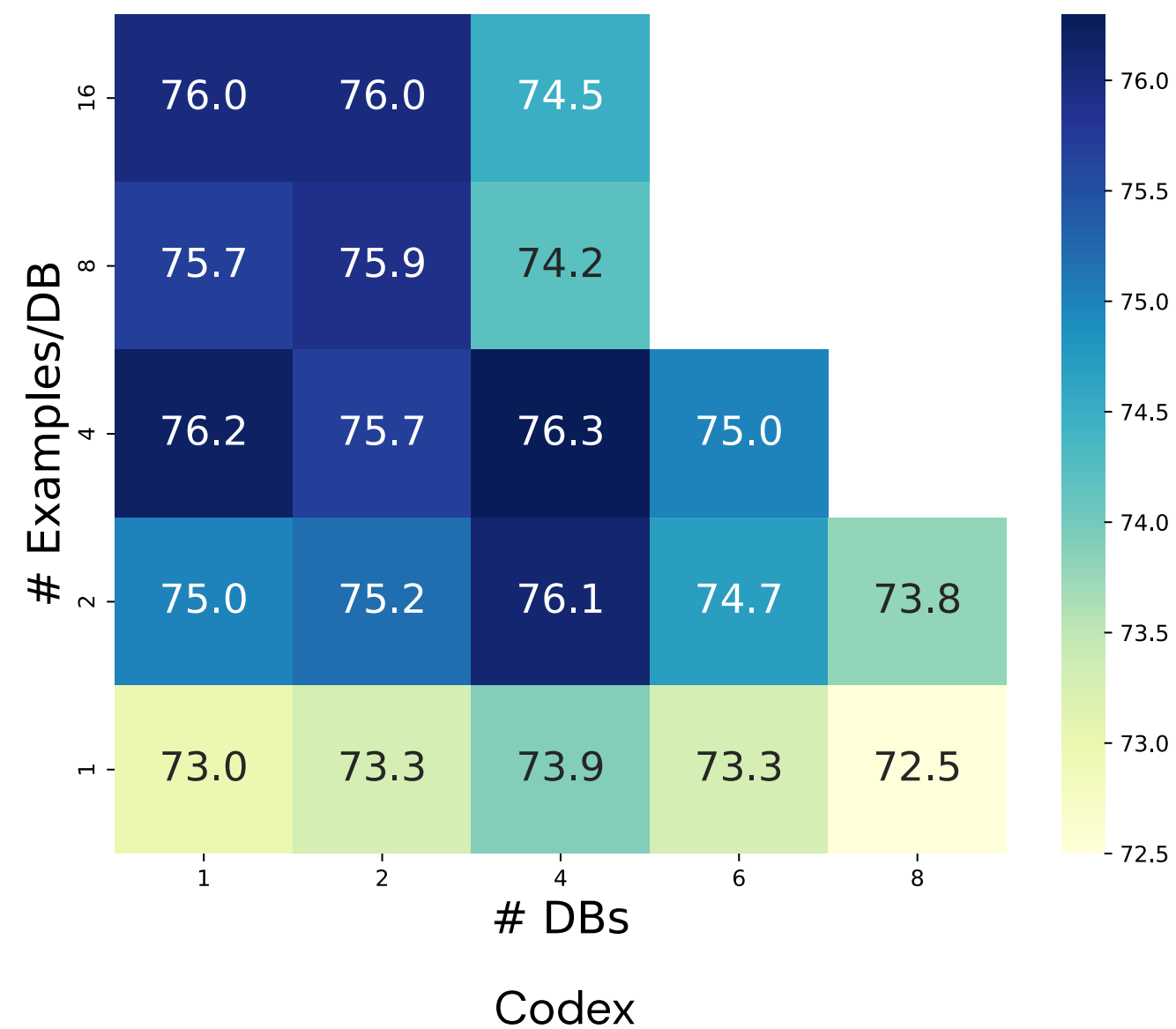
Cross-domain Text-to-SQL Results

- Why does the performance increases and then decreases when more demonstration databases are provided?

Codex

Cross-domain Text-to-SQL Results

- Why does the performance increases and then decreases when more demonstration databases are provided?
  - It is related to the length of the prompt text.

Codex

Codex

Cross-domain Text-to-SQL Results

- Are table relationship and content still important with out-of-domain demonstrations?
  - Table relationship and content are database-specific knowledge.
  - LLMs cannot learn them from out-of-domain demonstrations.

Database Prompt Construction		0-shot	1-shot	2-shot	4-shot	8-shot	16-shot
Table Schema	Table(Columns)	71.9	72.0	73.0	73.2	72.8	73.9
	Columns=[]	71.8	71.9	73.6	74.2	73.7	74.4
+Relationship	Columns=[]+ForeignKey	73.1	73.3	74.5	74.9	74.9	75.2
	CreateTable	73.1	72.1	73.4	73.7	74.1	75.1
+Relationship+Content	CreateTable+InsertRow 3	71.9	72.2	74.1	74.9	74.9	74.8
	CreateTable+SelectRow 3	74.1	73.0	75.0	76.2	75.7	76.0
	CreateTable+SelectCol 3	**75.7**	**74.4**	**75.5**	**76.5**	**76.8**	**76.5**

Takeaways

- Table relationship and database content play a crucial role. However, it requires careful designs in zero-shot and cross-domain settings.
- In-domain demonstrations mitigate LLMs' sensitivity to database constructions, but cannot replace database content knowledge.
- The prompt length has a significant impact on LLMs' performance.

Q&A