



SQL PROJECT (DODGERS SPORTS FACILITY CENTER)

Prepared for: **Portfolio**

Prepared by: **OLUWASEGUN MOHAMMED**, *Data Analyst*

September 23, 2022

EXECUTIVE SUMMARY

Introduction

I'm currently focused on knowing SQL to fetch, manipulate and create data. In this new learning phase, I currently need to put out my new skills about SQL functions I self-learned. This is my first SQL project and I'm passionate about creating more. In this project, I will share the steps I took in analyzing the data, and lastly, I will be transparent with my report.

The Data

This project uses SQL (Structured Query Language) to perform an exploratory data analysis on the "Dodgers Sports Facility Center". This is fictitious (imaginary) client data for a certain company in the US region in Las Vegas. The file is broken down into three tables; "bookings, facilities, and members"

Solution

I made use of PostgreSQL as I'm more familiar with this ORDBMS (Oriental-Relational Database Management System). Also, I made use of the SQL functions to query for results.



EXPLORATORY DATA ANALYSIS

Questions for Analysis

1. Are there duplicated data in the bookings table?
 2. What are the sports facilities available in the Sports center?
 3. What is the cost of facilities to both members and guests?
 4. What are the names of facilities free for members to use?
 5. What are the facilities that charge a fee to members and that fee is less than 1/50th of the monthly cost of maintenance?
 6. Who are the members who joined after the start of September 2012?
 7. What is the signup date of the first member?
 8. What is the signup date of the last member?
 9. How many facilities have a guest cost of 10 or more?
 10. How many slots are booked per facility in the month of September 2012?
 11. Which facility has more than 1,000 slots booked?
 12. What are the start times for the bookings of tennis courts for the date 21st of September, 2012?
 13. What are the start times bookings for the member "David Farrell"?
-

RESULTS

✓ Are there any duplicated data in the bookings table

The screenshot shows a database query interface. At the top, there is a toolbar with various icons for file operations, editing, and execution. Below the toolbar, there are tabs for 'Query' and 'Query History'. The 'Query' tab is active, displaying a SQL query:

```
1 SELECT *
2 FROM cd.bookings
3 GROUP BY bookings.bookid
4 HAVING COUNT(bookings.bookid) > 1
```

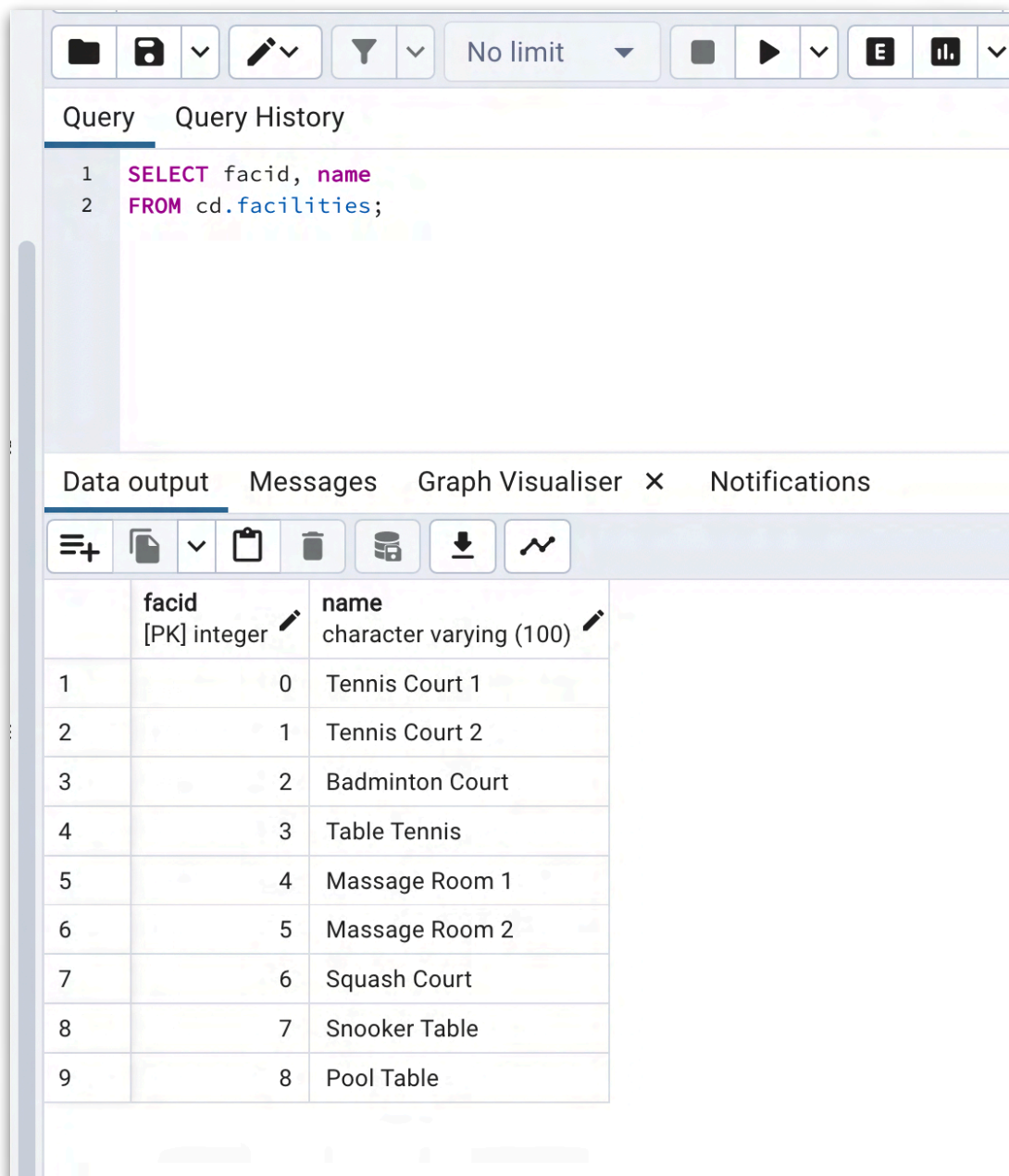
Below the query editor, there are tabs for 'Data output', 'Messages', 'Graph Visualiser', and 'Notifications'. The 'Data output' tab is active, showing a table with the following columns:

bookid	facid	memid	starttime	slots
--------	-------	-------	-----------	-------

The table is currently empty, indicating that no duplicate data was found in the bookings table.

The output of the query returns to be empty, hence, this signifies there were no duplicate data in the bookings table.

✓ What are the sports facilities available in the Sports center?



The screenshot shows a database query tool interface. At the top, there is a toolbar with icons for file operations, editing, filtering, and execution. Below the toolbar, the 'Query' tab is active, displaying a SQL query:

```
1 SELECT facid, name
2 FROM cd.facilities;
```

Below the query editor, the 'Data output' tab is active, showing the results of the query in a table format. The table has two columns: 'facid' (integer, primary key) and 'name' (character varying (100)). There are 9 rows of data, each representing a sports facility.

	facid [PK] integer	name character varying (100)
1	0	Tennis Court 1
2	1	Tennis Court 2
3	2	Badminton Court
4	3	Table Tennis
5	4	Massage Room 1
6	5	Massage Room 2
7	6	Squash Court
8	7	Snooker Table
9	8	Pool Table

There are 9 sports facilities available in the sports center with each allocated to an ID.

☒ What is the cost of facilities to both members and guests?

No limit

Query
Query History

```

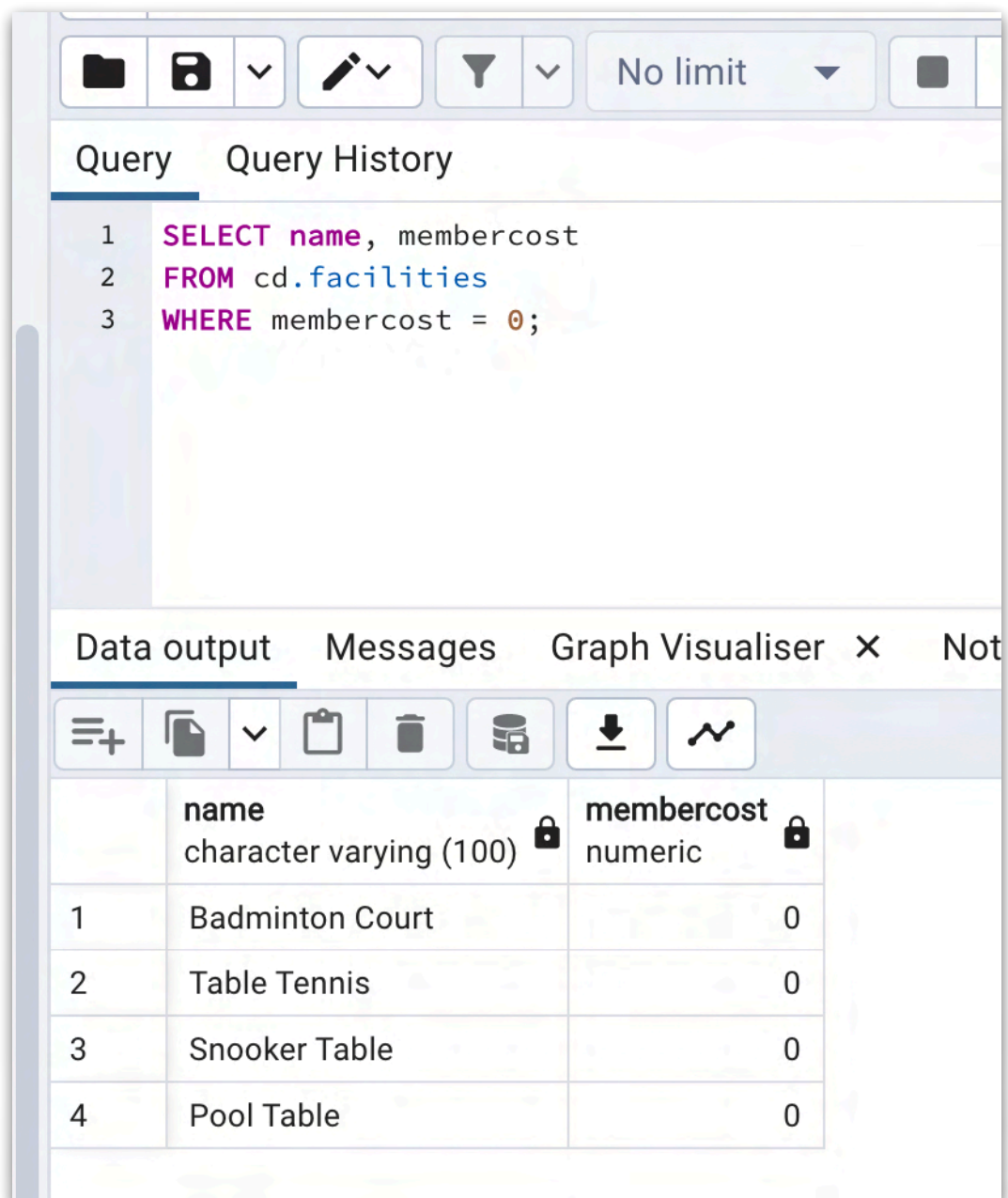
1 SELECT facid, name, membercost, guestcost
2 FROM cd.facilities;
```

Data output
Messages
Graph Visualiser X
Notifications

	facid [PK] integer	name character varying (100)	membercost numeric	guestcost numeric
1	0	Tennis Court 1	5	25
2	1	Tennis Court 2	5	25
3	2	Badminton Court	0	15.5
4	3	Table Tennis	0	5
5	4	Massage Room 1	35	80
6	5	Massage Room 2	35	80
7	6	Squash Court	3.5	17.5
8	7	Snooker Table	0	5
9	8	Pool Table	0	5

It is of notice that guests' cost is more expensive than members' cost.

✓ What are the names of facilities free for members to use?



The screenshot shows a database query interface. At the top, there is a toolbar with icons for file operations, a dropdown menu, a filter icon, and a 'No limit' button. Below the toolbar, there are two tabs: 'Query' and 'Query History'. The 'Query' tab is active, displaying a SQL query:

```
1 SELECT name, membercost
2 FROM cd.facilities
3 WHERE membercost = 0;
```

Below the query editor, there are four tabs: 'Data output', 'Messages', 'Graph Visualiser', and 'Notepad'. The 'Data output' tab is active, showing a table with the results of the query. The table has two columns: 'name' (character varying (100)) and 'membercost' (numeric). The results are as follows:

	name character varying (100)	membercost numeric
1	Badminton Court	0
2	Table Tennis	0
3	Snooker Table	0
4	Pool Table	0

Badminton Court, Table Tennis, Snooker Table, and Pool Table are the 4 facilities free for members of the center to use.

- ☒ What are the facilities that charge a fee to members and that fee is less than 1/50th of the monthly cost of maintenance?

No limit

Query

Query History

```

1 SELECT facid, name, membercost, monthlymaintenance
2 FROM cd.facilities
3 WHERE membercost > 0 AND
4 (membercost < monthlymaintenance/50.0);
5

```

Data output

Messages

Graph Visualiser

×

Notifications

	facid [PK] integer	name character varying (100)	membercost numeric	monthlymaintenance numeric
1	4	Message Room 1	35	3000
2	5	Message Room 2	35	3000

"Message Room 1" and "Message Room 2" are the only facilities charging a fee of less than 1/50th of the monthly cost of maintenance to members.

✓ Who are the members who joined after the start of September 2012?

Query				
<pre> 1 SELECT memid, surname, firstname, joindate 2 FROM cd.members 3 WHERE joindate >= '2012-09-01'; 4 </pre>				
Data output				
	memid [PK] integer	surname character varying (200)	firstname character varying (200)	joindate timestamp without time zone
1	24	Sarwin	Ramnaresh	2012-09-01 08:44:42
2	26	Jones	Douglas	2012-09-02 18:43:05
3	27	Rumney	Henrietta	2012-09-05 08:42:35
4	28	Farrell	David	2012-09-15 08:22:05
5	29	Worthington-Smyth	Henry	2012-09-17 12:27:15
6	30	Purview	Millicent	2012-09-18 19:04:01
7	33	Tupperware	Hyacinth	2012-09-18 19:32:05
8	35	Hunt	John	2012-09-19 11:32:45
9	36	Crumpet	Erica	2012-09-22 08:36:38
10	37	Smith	Darren	2012-09-26 18:08:45

This shows the list of members joining the sport center after the start of September.

✓ What is the signup date of the first member?

The screenshot shows a SQL query editor interface. At the top, there is a toolbar with icons for file operations, query execution, and settings. Below the toolbar, the 'Query' tab is active, displaying the following SQL query:

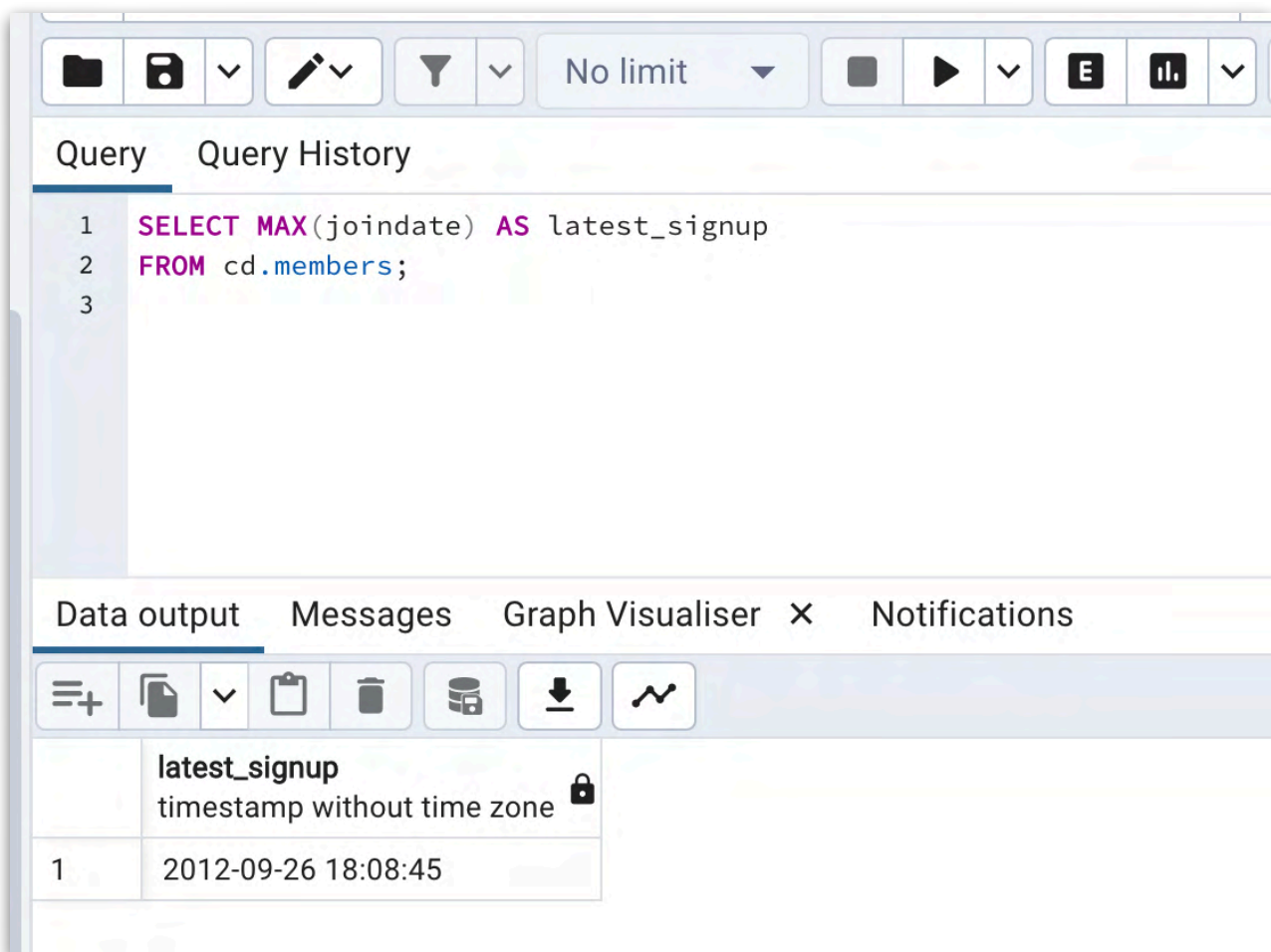
```
1 SELECT MIN(joindate) AS first_signup
2 FROM cd.members;
3
4
```

Below the query editor, the 'Data output' tab is active, showing the results of the query. The results are displayed in a table with one column, 'first_signup', and one row containing the value '2012-07-01 00:00:00'.

	first_signup timestamp without time zone
1	2012-07-01 00:00:00

The query returns the date the first signup was made at the center.

✓ What is the signup date of the last member?



The screenshot shows a SQL query editor interface. At the top, there is a toolbar with icons for file operations, query execution, and settings. Below the toolbar, the 'Query' tab is active, displaying the following SQL query:

```
1 SELECT MAX(joindate) AS latest_signup
2 FROM cd.members;
3
```

Below the query editor, the 'Data output' tab is active, showing the results of the query. The results are displayed in a table with two columns: the column name 'latest_signup' and its data type 'timestamp without time zone'. The first row of data shows the date and time '2012-09-26 18:08:45'.

	latest_signup timestamp without time zone
1	2012-09-26 18:08:45

The query returns the date the last signup was made at the center.

☒ How many facilities have a guest cost of 10 or more?

The screenshot shows a database query interface. At the top, there is a toolbar with icons for file operations, a dropdown menu, a filter icon, a dropdown menu, a text input field with 'No limit', and a play button. Below the toolbar, there are two tabs: 'Query' and 'Query History'. The 'Query' tab is active, displaying a SQL query:

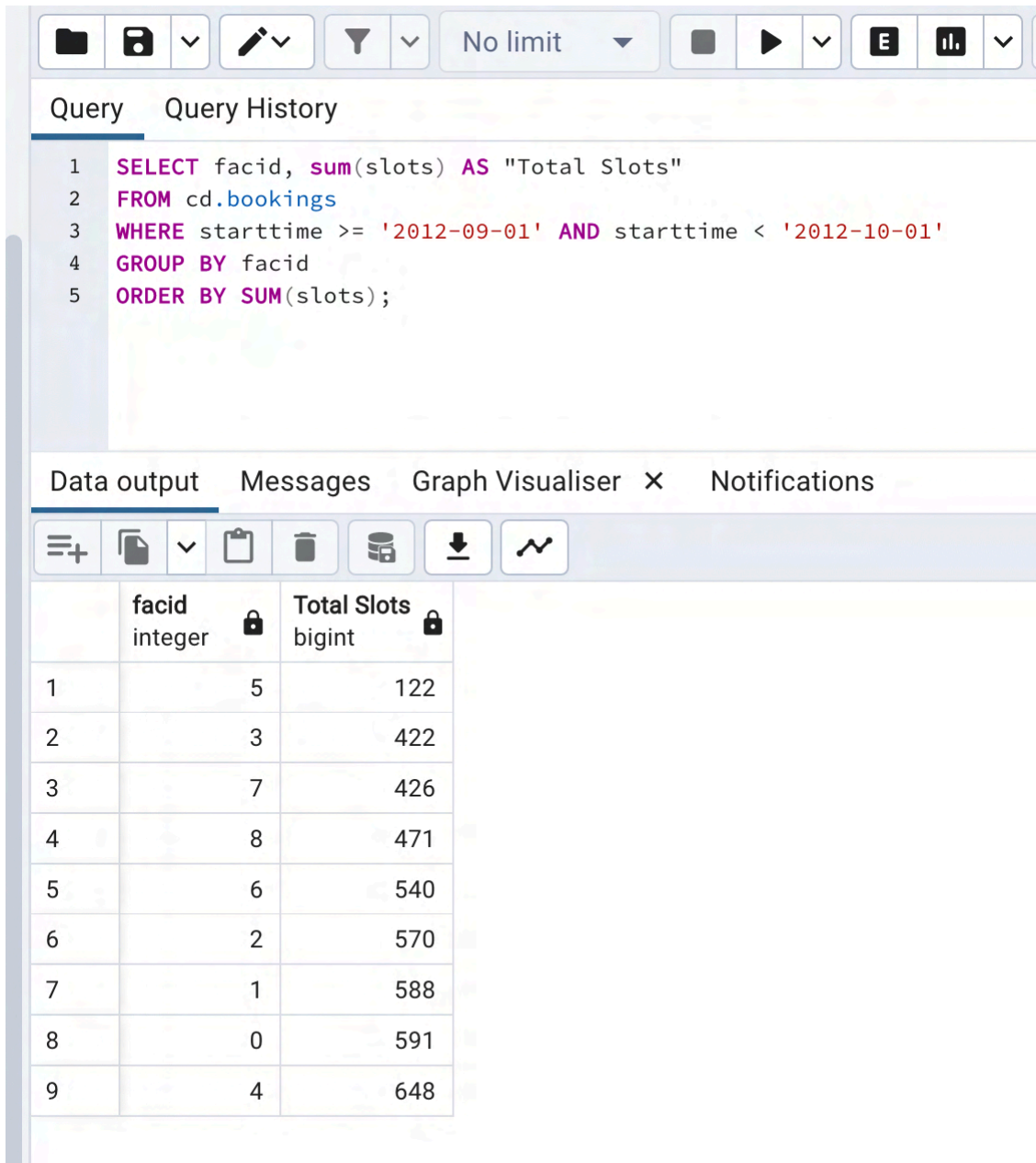
```
1 SELECT COUNT(*)
2 FROM cd.facilities
3 WHERE guestcost >= 10;
4
```

Below the query editor, there are four tabs: 'Data output', 'Messages', 'Graph Visualiser', and 'Notification'. The 'Data output' tab is active, showing a table with the following data:

	count bigint	
1	6	

There are 6 facilities that cost more than \$10 to the guests'.

✓ How many slots are booked per facility in the month of September 2012?



The screenshot shows a database query interface. At the top, there is a toolbar with icons for file operations, filters, and execution. Below the toolbar, the 'Query' tab is active, displaying a SQL query. The query is as follows:

```

1 SELECT facid, sum(slots) AS "Total Slots"
2 FROM cd.bookings
3 WHERE starttime >= '2012-09-01' AND starttime < '2012-10-01'
4 GROUP BY facid
5 ORDER BY SUM(slots);

```

Below the query editor, the 'Data output' tab is active, showing the results of the query in a table. The table has two columns: 'facid' (integer) and 'Total Slots' (bigint). The results are as follows:

	facid integer	Total Slots bigint
1	5	122
2	3	422
3	7	426
4	8	471
5	6	540
6	2	570
7	1	588
8	0	591
9	4	648

This shows the number of slots booked per facility in the month of September.

☒ Which facility has more than 1,000 slots booked?

▼

▼

No limit▼

▼

Query

Query History

```

1  SELECT facid, sum(slots) AS total_slots
2  FROM cd.bookings
3  GROUP BY facid
4  HAVING SUM(slots) > 1000
5  ORDER BY facid;

```

Data output

Messages

Graph Visualiser

×

Notificatio

≡+

▼

↓

	facid integer		total_slots bigint
1		0	1320
2		1	1278
3		2	1209
4		4	1404
5		6	1104

- ☒ What are the start times for the bookings of tennis courts for the date 21st of September, 2012?

No limit

Query

Query History

```

1 SELECT cd.bookings.starttime AS start, cd.facilities.name AS name
2 FROM cd.facilities
3 INNER JOIN cd.bookings
4 ON cd.facilities.facid = cd.bookings.facid
5 WHERE cd.facilities.facid IN (0,1)
6 AND cd.bookings.starttime >= '2012-09-21'
7 AND cd.bookings.starttime < '2012-09-22'
8 ORDER BY cd.bookings.starttime;
9

```

Data output

Messages

Graph Visualiser

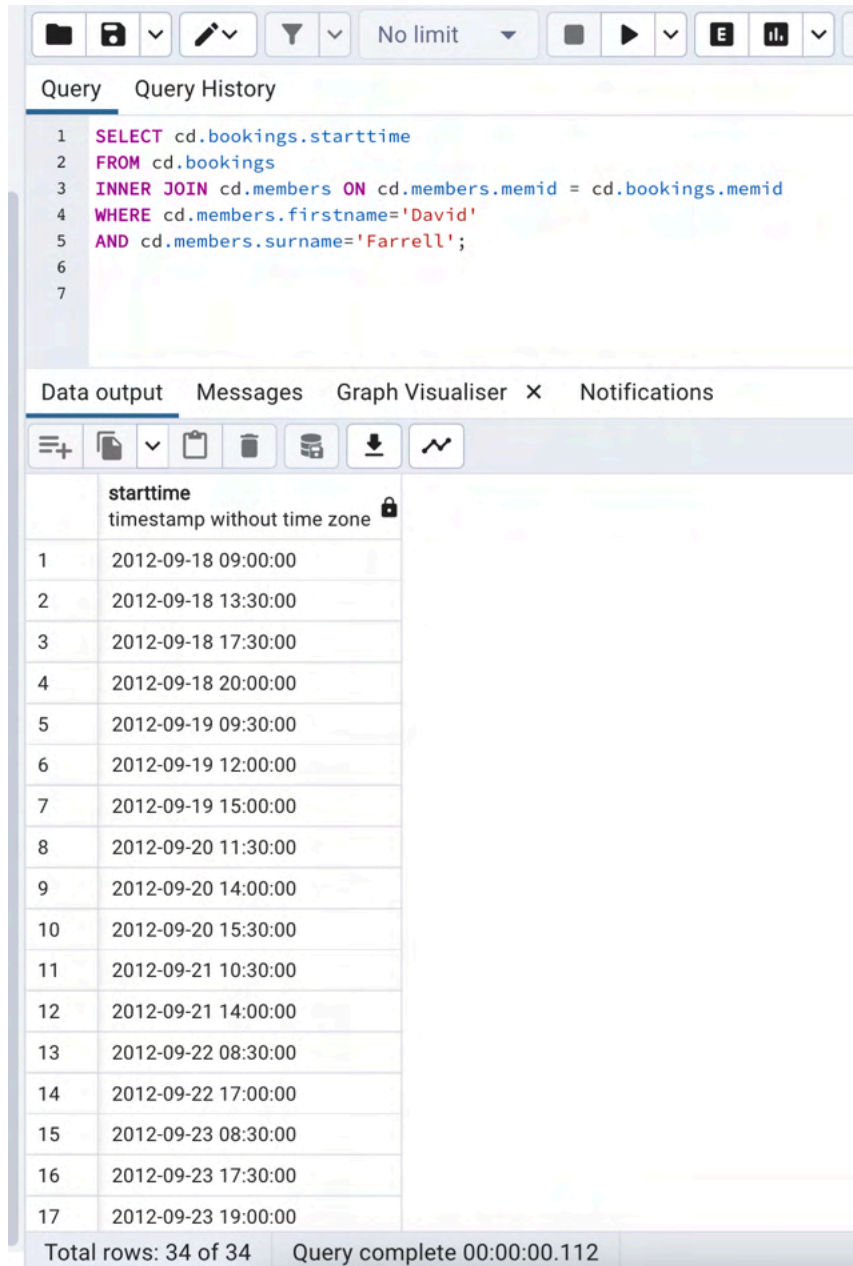
×

Notifications

	start timestamp without time zone	name character varying (100)
1	2012-09-21 08:00:00	Tennis Court 1
2	2012-09-21 08:00:00	Tennis Court 2
3	2012-09-21 09:30:00	Tennis Court 1
4	2012-09-21 10:00:00	Tennis Court 2
5	2012-09-21 11:30:00	Tennis Court 2
6	2012-09-21 12:00:00	Tennis Court 1
7	2012-09-21 13:30:00	Tennis Court 1
8	2012-09-21 14:00:00	Tennis Court 2
9	2012-09-21 15:30:00	Tennis Court 1
10	2012-09-21 16:00:00	Tennis Court 2
11	2012-09-21 17:00:00	Tennis Court 1
12	2012-09-21 18:00:00	Tennis Court 2

Results of Start times of Tennis facilities.

✓ What are the start times bookings for the member “David Farrell”?



The screenshot shows a database query interface. At the top, there is a toolbar with icons for file operations, query execution, and settings. Below the toolbar, the 'Query' tab is active, displaying a SQL query:

```

1 SELECT cd.bookings.starttime
2 FROM cd.bookings
3 INNER JOIN cd.members ON cd.members.memid = cd.bookings.memid
4 WHERE cd.members.firstname='David'
5 AND cd.members.surname='Farrell';
6
7

```

Below the query editor, the 'Data output' tab is active, showing the results of the query. The results are displayed in a table with two columns: an index and the 'starttime' (timestamp without time zone). The table contains 17 rows of data, representing bookings for David Farrell.

	starttime timestamp without time zone
1	2012-09-18 09:00:00
2	2012-09-18 13:30:00
3	2012-09-18 17:30:00
4	2012-09-18 20:00:00
5	2012-09-19 09:30:00
6	2012-09-19 12:00:00
7	2012-09-19 15:00:00
8	2012-09-20 11:30:00
9	2012-09-20 14:00:00
10	2012-09-20 15:30:00
11	2012-09-21 10:30:00
12	2012-09-21 14:00:00
13	2012-09-22 08:30:00
14	2012-09-22 17:00:00
15	2012-09-23 08:30:00
16	2012-09-23 17:30:00
17	2012-09-23 19:00:00

At the bottom of the interface, a status bar indicates 'Total rows: 34 of 34' and 'Query complete 00:00:00.112'.

Start times bookings of David Farrell on all days present at the center.

CONCLUSION

In finding solutions to my analysis, I made use of the SQL functions and the aggregate functions, including the MIN, MAX, COUNT, AVG, HAVING, etc.

The questions answered by this analysis are insightful and can answer the business questions of the organization if needs arrive for driving out Information.

This is my first SQL project, and I'm looking forward to getting more of it done. Thanks for going through it.



THANK YOU !!!
