By Nitish Adhikari

Email id :nitishbuzzpro@gmail.com (mailto:nitishbuzzpro@gmail.com), +91-9650740295

Linkedin : https://www.linkedin.com/in/nitish-adhikari-6b2350248 (https://www.linkedin.com/in/nitish-adhikari-6b2350248)

Project : Query a hotel/club database on PostgreSQL server using Jupyter Notebook IDE

A Data Scientist/ML/AI developer needs SQL in order to handle structured data. This structured data is stored in relational databases. Therefore, in order to query these databases, a data scientist must have a sound knowledge of SQL.

In [ ]:

```
!pip install sqlalchemy
!pip install psycopg2
!pip install ipython-sql
```

In [193]:

```
import sqlalchemy
```

In [194]:

```
#create a postgresql engine
engine = sqlalchemy.create_engine('postgresql://postgres:******@localhost:5432/hotel')
```

In [195]:

```
#Load the sql extension
%load_ext sql
```

```
The sql extension is already loaded. To reload it, use:
  %reload_ext sql
```

In [196]:

```
#Set up the connection
%sql $engine.url
```

retrieve all the information from the cd.facilities table

In [197]:

```
%sql select * from cd.facilities
```

```
   postgresql://postgres:***@localhost:5432/dvdrental
 * postgresql://postgres:***@localhost:5432/hotel
9 rows affected.
```

Out[197]:

facid	name	membercost	guestcost	initialoutlay	monthlymaintenance
0	Tennis Court 1	5	25	10000	200
1	Tennis Court 2	5	25	8000	200
2	Badminton Court	0	15.5	4000	50
3	Table Tennis	0	5	320	10
4	Massage Room 1	35	80	4000	3000
5	Massage Room 2	35	80	4000	3000
6	Squash Court	3.5	17.5	5000	80
7	Snooker Table	0	5	450	15
8	Pool Table	0	5	400	15

print out a list of all of the facilities and their cost to members.

In [198]:

```
%sql select name, membercost from cd.facilities
```

```
   postgresql://postgres:***@localhost:5432/dvdrental
 * postgresql://postgres:***@localhost:5432/hotel
9 rows affected.
```

Out[198]:

name	membercost
Tennis Court 1	5
Tennis Court 2	5
Badminton Court	0
Table Tennis	0
Massage Room 1	35
Massage Room 2	35
Squash Court	3.5
Snooker Table	0
Pool Table	0

produce a list of facilities that charge a fee to members

In [199]:

```
%sql select * from cd.facilities \
where membercost != 0
```

```
   postgresql://postgres:***@localhost:5432/dvdrental
 * postgresql://postgres:***@localhost:5432/hotel
5 rows affected.
```

Out[199]:

facid	name	membercost	guestcost	initialoutlay	monthlymaintenance
0	Tennis Court 1	5	25	10000	200
1	Tennis Court 2	5	25	8000	200
4	Massage Room 1	35	80	4000	3000
5	Massage Room 2	35	80	4000	3000
6	Squash Court	3.5	17.5	5000	80

produce a list of facilities that charge a fee to members, and that fee is less than 1/50th of the monthly maintenance cost

In [200]:

```
%sql select * from cd.facilities \
where membercost != 0 and membercost < monthlymaintenance/50
```

```
   postgresql://postgres:***@localhost:5432/dvdrental
 * postgresql://postgres:***@localhost:5432/hotel
2 rows affected.
```

Out[200]:

facid	name	membercost	guestcost	initialoutlay	monthlymaintenance
4	Massage Room 1	35	80	4000	3000
5	Massage Room 2	35	80	4000	3000

produce a list of all facilities with the word 'Tennis' in their name

In [201]:

```
%sql select * from cd.facilities \
where name like '%Tennis%'
```

```
   postgresql://postgres:***@localhost:5432/dvdrental
 * postgresql://postgres:***@localhost:5432/hotel
3 rows affected.
```

Out[201]:

facid	name	membercost	guestcost	initialoutlay	monthlymaintenance
0	Tennis Court 1	5	25	10000	200
1	Tennis Court 2	5	25	8000	200
3	Table Tennis	0	5	320	10

retrieve the details of facilities with ID 1 and 5

In [202]:

```
%sql select * from cd.facilities \
where facid in (1,5)
```

```
   postgresql://postgres:***@localhost:5432/dvdrental
 * postgresql://postgres:***@localhost:5432/hotel
2 rows affected.
```

Out[202]:

facid	name	membercost	guestcost	initialoutlay	monthlymaintenance
1	Tennis Court 2	5	25	8000	200
5	Massage Room 2	35	80	4000	3000

produce a list of members who joined after the start of September 2012. Return the memid, surname, firstname, and joindate of the members

In [203]:

```
%sql select memid, surname, firstname, joindate from cd.members \
WHERE joindate >= '2012-09-01'
```

```
   postgresql://postgres:***@localhost:5432/dvdrental
 * postgresql://postgres:***@localhost:5432/hotel
10 rows affected.
```

Out[203]:

memid	surname	firstname	joindate
24	Sarwin	Ramnaresh	2012-09-01 08:44:42
26	Jones	Douglas	2012-09-02 18:43:05
27	Rumney	Henrietta	2012-09-05 08:42:35
28	Farrell	David	2012-09-15 08:22:05
29	Worthington-Smyth	Henry	2012-09-17 12:27:15
30	Purview	Millicent	2012-09-18 19:04:01
33	Tupperware	Hyacinth	2012-09-18 19:32:05
35	Hunt	John	2012-09-19 11:32:45
36	Crumpet	Erica	2012-09-22 08:36:38
37	Smith	Darren	2012-09-26 18:08:45

produce an ordered list of the first 10 surnames in the members table. The list would not contain duplicates

In [204]:

```
%sql select distinct(surname) from cd.members \
order by surname \
limit 10
```

```
   postgresql://postgres:***@localhost:5432/dvdrental
 * postgresql://postgres:***@localhost:5432/hotel
10 rows affected.
```

Out[204]:

surname
Bader
Baker
Boothe
Butters
Coplin
Crumpet
Dare
Farrell
Genting
GUEST

retrieve signup date of the last member

In [205]:

```
%sql select max(joindate) as latest_signup from cd.members
```

```
   postgresql://postgres:***@localhost:5432/dvdrental
 * postgresql://postgres:***@localhost:5432/hotel
1 rows affected.
```

Out[205]:

latest_signup
2012-09-26 18:08:45

Produce a count of the number of facilities that have a cost to guests of 10 or more

In [206]:

```
%sql select count(*) from cd.facilities \
where guestcost >= 10
```

```
   postgresql://postgres:***@localhost:5432/dvdrental
 * postgresql://postgres:***@localhost:5432/hotel
1 rows affected.
```

Out[206]:

count
6

Produce a list of the total number of slots booked per facility in the month of September 2012. Produce an output table consisting of facility id and slots, sorted by the number of slots

In [207]:

```
%sql SELECT facid, sum(slots) AS total_slots FROM cd.bookings \
WHERE starttime >= '2012-09-01' AND starttime < '2012-10-01' \
GROUP BY facid \
ORDER BY SUM(slots)
```

```
   postgresql://postgres:***@localhost:5432/dvdrental
 * postgresql://postgres:***@localhost:5432/hotel
9 rows affected.
```

Out[207]:

facid	total_slots
5	122
3	422
7	426
8	471
6	540
2	570
1	588
0	591
4	648

Produce a list of facilities with more than 1000 slots booked. Produce an output table consisting of facility id and total slots, sorted by facility id.

In [208]:

```
%sql SELECT facid, sum(slots) AS total_slots FROM cd.bookings \
group by facid \
order by  facid
```

```
   postgresql://postgres:***@localhost:5432/dvdrental
 * postgresql://postgres:***@localhost:5432/hotel
9 rows affected.
```

Out[208]:

facid	total_slots
0	1320
1	1278
2	1209
3	830
4	1404
5	228
6	1104
7	908
8	910

Produce a list of the start times for bookings for tennis courts, for the date '2012-09-21'.Return a list of start time and facility name pairings, ordered by the time.

In [209]:

```
%sql SELECT starttime, name FROM cd.bookings \
inner join cd.facilities on cd.bookings.facid = cd.facilities.facid \
WHERE cd.facilities.facid IN (0,1) \
AND cd.bookings.starttime >= '2012-09-21' \
AND cd.bookings.starttime < '2012-09-22' \
order by starttime
```

```
   postgresql://postgres:***@localhost:5432/dvdrental
 * postgresql://postgres:***@localhost:5432/hotel
12 rows affected.
```

Out[209]:

starttime	name
2012-09-21 08:00:00	Tennis Court 1
2012-09-21 08:00:00	Tennis Court 2
2012-09-21 09:30:00	Tennis Court 1
2012-09-21 10:00:00	Tennis Court 2
2012-09-21 11:30:00	Tennis Court 2
2012-09-21 12:00:00	Tennis Court 1
2012-09-21 13:30:00	Tennis Court 1
2012-09-21 14:00:00	Tennis Court 2
2012-09-21 15:30:00	Tennis Court 1
2012-09-21 16:00:00	Tennis Court 2
2012-09-21 17:00:00	Tennis Court 1
2012-09-21 18:00:00	Tennis Court 2

complete!!