

Database Creation

Monday, February 6, 2017



Agenda

- Announcements
- Reading Quiz
- Postgres RDS
- DDL
- Copy Command
- Transforms

Announcements

- Lab 1: grades and comments will be released this evening
- Next class: Lab 2 setup session
- Lab 2 setup guide: <https://github.com/wolfier/CS327E/wiki/Setting-up-Lab-Two>
- Next week: Lab 2 sessions

Q1: The SELECT statement ...

- a) Can have multiple source tables
- b) Retrieves data from the database
- c) Returns a relation
- d) All of the above

Q2: Which query retrieves all columns from the table below named Work?

- a) `SELECT * FROM Work`
- b) `SELECT all FROM Work`
- c) `FROM Work SELECT all`
- d) `FROM Work SELECT *`

work_num	author_num	title
1	1	Jane Eyre
2	1	Villette
3	2	Hound of the Baskervilles
4	2	Lost World, The
5	2	Complete Sherlock Holmes
7	3	Prince and the Pauper
8	3	Tom Sawyer

Q3: What does the WHERE clause do?

- a) Deletes records
- b) Indicates source tables or columns
- c) Joins tables
- d) Filters rows
- e) Groups similar columns

Q4: Which of the following is incorrect?

- a) UPDATE changes the data in the records
- b) INSERT inserts new data into tables in the database
- c) DELETE removes records from the tables in the database
- d) TRUNCATE removes the data but not the data structure
- e) DROP removes the data structure but not the data

Q5: What does NULL mean?

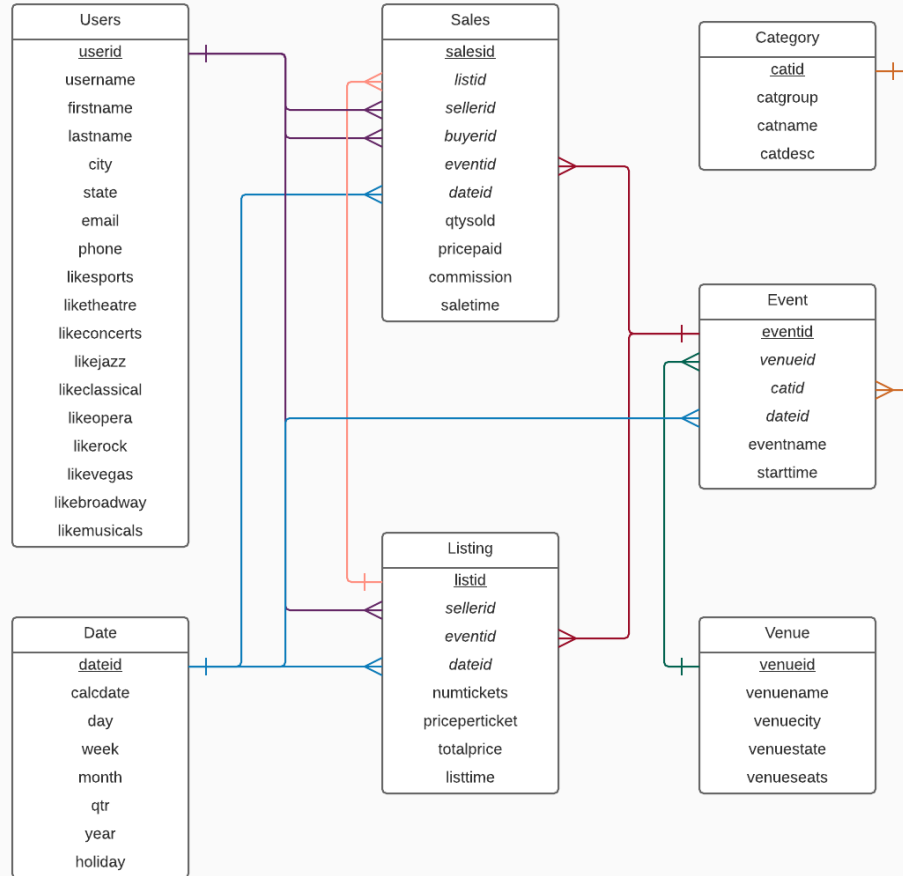
- a) True
- b) Absence of any value
- c) False

Postgres RDS Demo

See Lab 2 setup guide for a step-by-step procedure on how to create the Postgres RDS instance and connect to it from psql:

<https://github.com/wolfier/CS327E/wiki/Setting-up-Lab-Two>

TICKIT ERD



From Physical Diagram to DDL

Users		
PK	userid	int
	username	char(8)
	firstname	varchar(30)
	lastname	varchar(30)
	city	varchar(30)
	state	char(2)
	email	varchar(100)
	phone	char(14)
	likesports	boolean
	liketheatre	boolean
	likeconcerts	boolean
	likejazz	boolean
	likeclassical	boolean
	likeopera	boolean
	likerock	boolean
	likevegas	boolean
	likebroadway	boolean
	likemusicals	boolean



```
1 create table Users(  
2     userid int primary key,  
3     username char(8),  
4     firstname varchar(30),  
5     lastname varchar(30),  
6     city varchar(30),  
7     state char(2),  
8     email varchar(100),  
9     phone char(14),  
10    likesports boolean,  
11    liketheatre boolean,  
12    likeconcerts boolean,  
13    likejazz boolean,  
14    likeclassical boolean,  
15    likeopera boolean,  
16    likerock boolean,  
17    likevegas boolean,  
18    likebroadway boolean,  
19    likemusicals boolean);
```

From Physical Diagram to DDL

Date		
PK	dateid	smallint
	caldate	date
	day	char(3)
	week	smallint
	month	char(5)
	qtr	char(5)
	year	smallint
	holiday	boolean



```
35 create table Date(  
36     dateid smallint primary key,  
37     caldate date not null,  
38     day char(3) not null,  
39     week smallint not null,  
40     month char(5) not null,  
41     qtr char(5) not null,  
42     year smallint not null,  
43     holiday boolean);
```

	A	B	C	D	E	F	G	H
1	1827	1/1/2014	WE	1	JAN	1	2014	TRUE
2	1828	1/2/2014	TH	1	JAN	1	2014	FALSE
3	1829	1/3/2014	FR	1	JAN	1	2014	FALSE
4	1830	1/4/2014	SA	2	JAN	1	2014	FALSE
5	1831	1/5/2014	SU	2	JAN	1	2014	FALSE
6	1832	1/6/2014	MO	2	JAN	1	2014	FALSE
7	1833	1/7/2014	TU	2	JAN	1	2014	FALSE
8	1834	1/8/2014	WE	2	JAN	1	2014	FALSE
9	1835	1/9/2014	TH	2	JAN	1	2014	FALSE

source csv file

dateid	caldate	day	week	month	qtr	year	holiday
1827	2014-01-01	WE	1	JAN	1	2014	t
1828	2014-01-02	TH	1	JAN	1	2014	f
1829	2014-01-03	FR	1	JAN	1	2014	f
1830	2014-01-04	SA	2	JAN	1	2014	f
1831	2014-01-05	SU	2	JAN	1	2014	f
1832	2014-01-06	MO	2	JAN	1	2014	f
1833	2014-01-07	TU	2	JAN	1	2014	f
1834	2014-01-08	WE	2	JAN	1	2014	f
1835	2014-01-09	TH	2	JAN	1	2014	f

target database table

From Physical Diagram to DDL

Event		
PK	eventid	int
FK	venueid	smallint
FK	catid	smallint
FK	dateid	smallint
	eventname	varchar(200)
	starttime	timestamp



```
49 create table Event(  
50     eventid int primary key,  
51     venueid smallint not null references Venue(venueid),  
52     catid smallint not null references Category(catid),  
53     dateid smallint not null references Date(dateid),  
54     eventname varchar(200) not null,  
55     starttime timestamp);
```

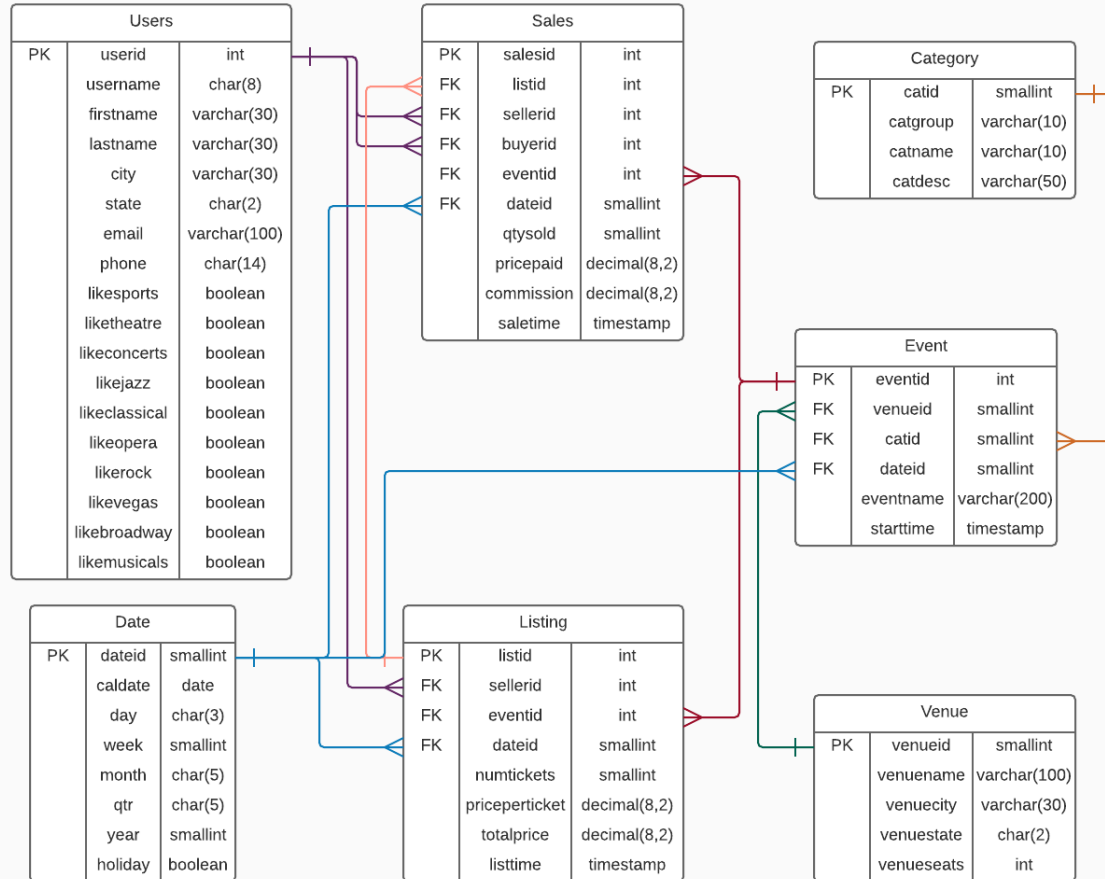
	A	B	C	D	E	F
1	1	305	8	1851	Gotterdammerung	1/25/2008 14:30
2	2	306	8	2114	Boris Godunov	10/15/2008 20:00
3	3	302	8	1935	Salome	4/19/2008 14:30
4	4	309	8	2090	La Cenerentola (Cinderella)	9/21/2008 14:30
5	5	302	8	1982	Il Trovatore	6/5/2008 19:00
6	6	308	8	2109	L Elisir d Amore	10/10/2008 19:30
7	7	309	8	1891	Doctor Atomic	3/6/2008 14:00
8	8	302	8	1832	The Magic Flute	1/6/2008 20:00
9	9	308	8	2087	The Fly	9/18/2008 19:30

source csv file

eventid	venueid	catid	dateid	eventname	starttime
1	305	8	1851	Gotterdammerung	2014-01-25 14:30:00
2	306	8	2114	Boris Godunov	2014-10-15 20:00:00
3	302	8	1935	Salome	2014-04-19 14:30:00
4	309	8	2090	La Cenerentola (Cinderella)	2014-09-21 14:30:00
5	302	8	1982	Il Trovatore	2014-06-05 19:00:00
6	308	8	2109	L Elisir d Amore	2014-10-10 19:30:00
7	309	8	1891	Doctor Atomic	2014-03-06 14:00:00
8	302	8	1832	The Magic Flute	2014-01-06 20:00:00
9	308	8	2087	The Fly	2014-09-18 19:30:00

target database table

Practice Problem 1: Write the DDL for the Sales table based on the below diagram



Practice Problem 1: Which tables must be created before the Sales table?

- a) None
- b) Users, Date, Event, Listing
- c) Users, Date, Category, Venue, Event, Listing
- d) Any of the above (i.e. table creation order doesn't matter)

Demo

See snippets repo on Github for DDL, data load, and transforms:

<https://github.com/cs327e-spring2017/snippets>

Summary

- Use \copy command to load data into tables:
<https://www.postgresql.org/docs/9.6/static/sql-copy.html>
- Use ALTER TABLE command to rename columns, drop columns, drop constraints, create constraints:
<https://www.postgresql.org/docs/9.6/static/sql-altertable.html>
- To delete a parent record, remember to first delete all its child records:
 1. DELETE FROM Events WHERE catid < 6;
 2. DELETE FROM Category WHERE catid < 6;