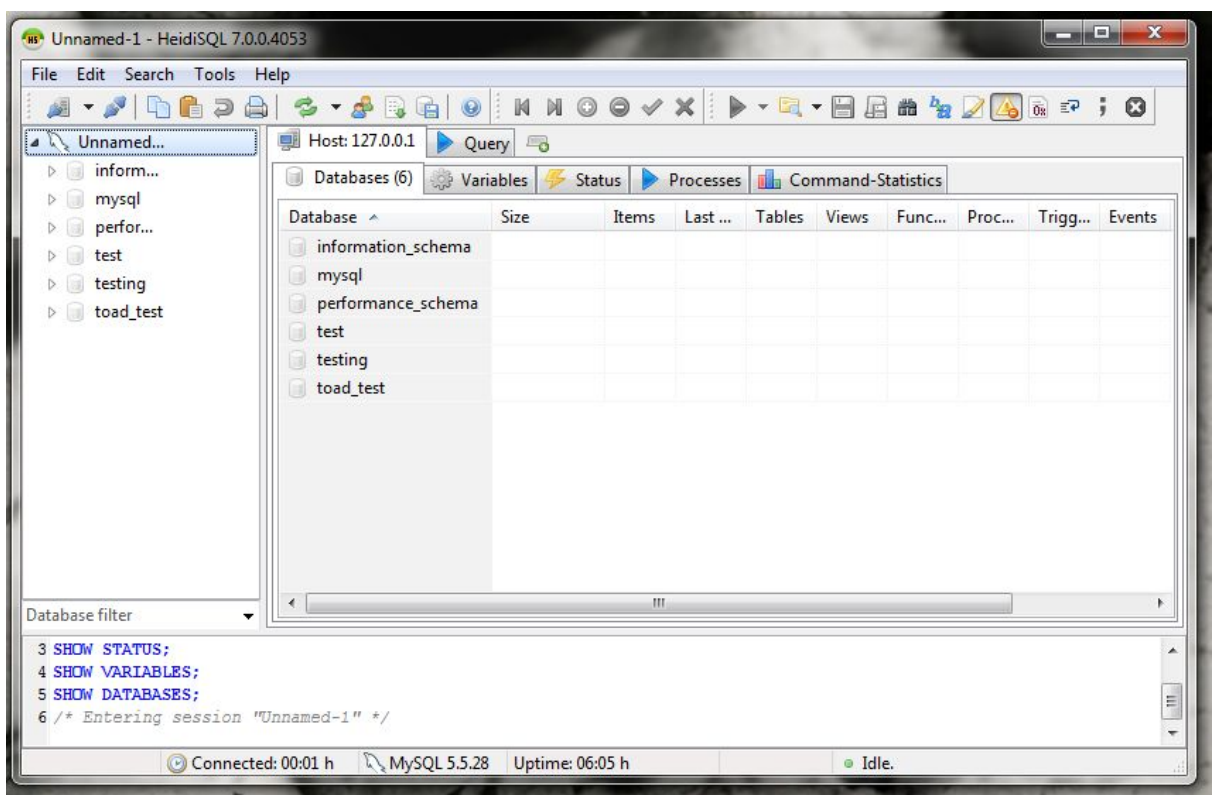


Database HW1
Mahya Jamshidian
Part 1:

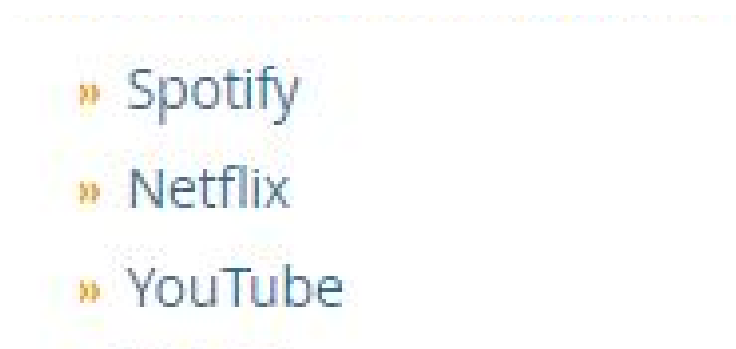
1. MySQL is one of the most popular opensource RDBMSs.



- a. 23 May 1995 - (OS?)
- b. Yes.



- c. Some of the noteworthy customers:



- d. The most preferred language that developers use alongside MySQL:
 - i. PHP - the usage of MySQL is not limited to some specific framework. MySQL is compatible with famous PHP frameworks like Laravel and Drupal.
 - ii. JAVA

2.

- a. One merit of the databases is fast data-access and efficient information retrieval. For this to happen, many complicated data structures are used to implement the efficiency needed as the backbone. Since many users are not trained for such complexity of programming, these structures are hidden beneath layers of abstraction.
- b. In the physical layer of abstraction, the real process of storing the data and lower layers data structures are described. The logical layer of abstraction describes the types of data and their relation. This high-level description provides the designer with simpler usage of the system. Although the implementation of the physical layer may be complicated, the user at the logical level does not need to deal with this complexity. This independence is often known as physical data independence.
- c. The first scheme, which is the most prevailing style of database design, is called the Relational Model. In this approach, each table contains unique and stable columns as well as unique records. Another model is known as the semistructured data model deals with attributes of the same type with different fields. The extensible markup language is a common used example of this style.

3. Four of the drawbacks of using file-processing systems for data storage and retrieval instead of databases are stated below:

- i. **Data redundancy and inconsistency:** File creation at any arbitrary time and probable duplication of data lead to higher storage and access cost. In addition, it may lead to data inconsistency; that is, various copies of the same data may no longer agree
 - ii. **Difficulty in accessing data:** Conventional file-processing environments do not allow needed data to be retrieved in a convenient and efficient manner. More responsive data-retrieval systems are required for general use
 - iii. **Data isolation:** Because data are scattered in various files, and files may be in different formats, writing new application programs to retrieve the appropriate data is difficult.
 - iv. **Security problems:** Not every user of the database system should be able to access all the data. But, since application programs are added to the file-processing system in an ad hoc manner, enforcing such security constraints is difficult.
- b. One can assume the below disadvantages to an RDBMS.
 - i. Expensive software design in comparison to file-system
 - ii. Expensive hardware support as a result of complex software

- iii. Skilled staff so increased expense in training more professional IT people.

- 4. A **super key** is a set of one or more attributes that, taken collectively, allow us to identify uniquely a tuple in the relation. This way, any tuple containing values describing a unique record can be a super key.

But, we are interested in a set of attributes that no proper subset can act as a super key. Such a minimal subset is the **candidate key**.

We shall use the term **primary key** as the candidate key chosen by database designer for uniquely describing records.

- 5. There are three possible explanations for the necessity of NULL value in a database.

- a. **Some attributes are not available in a record:** Sometimes using a default value for a blank space in a record leads to misinterpretation of the data, and by using a NULL value, we announce some missing knowledge regarding an attribute.
- b. **Sometimes in a record, some data is not applicable and therefore, not necessary, and it may seem meaningless.** For example, in a university database, a field corresponding to grade is not applicable to a student who has audited the course.
- c. **By exploiting NULL value, administrators can grant themselves more time for completing the records and fields.** Some databases require more time to be completed or some values are optional.

- 6. The <> ALL statement is valid if the left-hand side value is not equal to any of the right-hand side sub-query. Similarly, the NOT IN statement returns TRUE if the left-hand side value does not exist in the right-hand side set resulted from the sub-query. However, if the subquery returns no record, the NOT IN statement return NULL value whereas ALL statement always returns a TRUE value.

- 7.
 - a. While using nested queries, if one internal query returns a NULL value, the comparison will result in RUN TIME ERROR. In this query, if the budget field satisfying the internal WHERE is a NULL field, the query will halt with an error of RUN-TIME.

SELECT *name* **FROM** *instructor*

WHERE *salary* * 1.5 > (SELECT *budget* **FROM** *department*

WHERE *department.dept_name* = *instructor.dept_name*)

- b. This query returns all the students that have taken all the courses that the "Biology" department has to offer.

```

SELECT S.ID, S.name
FROM student AS S
WHERE NOT EXISTS ((SELECT course id
                    FROM course
                    WHERE dept_name = 'Biology')
EXCEPT
(SELECT T.course id
 FROM takes AS T
 WHERE S.ID = T.ID));

```

The same query can be rewritten with ALL statement as below since \neq ALL is identical to NOT IN:

```

SELECT S.ID, S.Name FROM Student AS S
WHERE NOT EXISTS
    ((SELECT course_id FROM Course
      WHERE (dept_name = "Biology"
            AND course_id <> ALL
                                (SELECT T.course_id FROM Takes AS T
                                WHERE S.ID = T.ID)))

```

8.

- a. In a database, tables on which other tables depend are first to be filled. In the DVD rental database, the below tables must be (or better be) completed for other tables to make sense.
 - i. **Country**, since the city table depends on it for the country_id field. **The city** table can also be further completed for it contains constant records.
 - ii. **Language**, since the film table depends on it for the language_id field.
 - iii. **Actor**, since the film_actor table depends on it for the actor_id field.
 - iv. **Category**, since the film_category table depends on it for category_id.
 - v. **Staff** and **Store** tables can be completed by the institution in the beginning.
- b. In the DVD rental database, these tables can grow in size daily:
 - i. **The film**, **Film_category**, and **film_actor**, since new movies can be described by a new category and can contain different actors and actresses. Note that the actor table may grow but not on a daily basis.
 - ii. **Rental**, by any new transaction, one record is added to this database.
 - iii. **Payment**, by any new transaction, one record is added to this database.
 - iv. **Customer**, any day some new customers may rent movies from different stores.
 - v. **Address**, since new customers and new staff may have distinct addresses.

9.

```

postgres=# SELECT pg_postmaster_start_time(), version(), pg_conf_load_time(), current_database();
 pg_postmaster_start_time | version | pg_conf_load_time | current_database
-----+-----+-----+-----
 2019-10-19 22:24:00.431664+03:30 | PostgreSQL 11.5, compiled by Visual C++ build 1914, 64-bit | 2019-10-19 22:24:00.399691+03:30 | postgres
(1 row)

```

10. S

a.

	Data Output	Explain	Messages	Notifications
	first_name character varying (45)	last_name character varying (45)	city character varying (50)	
1	Dorothy	Taylor	Esfahan	
2	Kathleen	Adams	Arak	
3	Roberta	Harper	Shahr-e Kord	
4	Dianne	Shelton	Tabriz	
5	Harry	Arce	Najafabad	
6	Howard	Fortner	Kermanshah	
7	Oscar	Aquino	Sirjan	
8	Tommy	Collazo	Qomsheh	

b. S

	first_name character varying (45)	last_name character varying (45)	
1	Minnie	Kilmer	
2	Renee	Ball	
3	Thora	Temple	
4	Burt	Dukakis	
5	Liza	Bergman	
6	Sidney	Crowe	
7	Angelina	Astaire	
8	Ed	Mansfield	
9	Ray	Johansson	
10	Laura	Brody	
11	Michelle	Mcconaughey	
12	Frances	Day-Lewis	
13	Michael	Bolger	
14	Burt	Temple	
15	Morgan	Mcdormand	
16	Kevin	Bloom	
17	Rip	Crawford	
18	Rita	Reynolds	
19	Morgan	Williams	
20	Albert	Johansson	
21	Jude	Cruise	
22	Jayne	Neeson	
23	Gregory	Gooding	
24	Burt	Posev	

c. D

	first_name character varying (45)	last_name character varying (45)	rental_date timestamp without time zone	return_date timestamp without time zone
1	Dianne	Shelton	2005-07-08 00:34:25	2005-07-08 22:55:25
2	Kathleen	Adams	2005-08-02 03:48:17	2005-08-02 22:15:17
3	Harry	Arce	2005-08-19 00:22:24	2005-08-19 22:53:24

d. S

Data Output

	first_name character varying (45)	last_name character varying (45)
1	Minnie	Kilmer
2	Renee	Ball
3	Thora	Temple
4	Burt	Dukakis
5	Liza	Bergman
6	Sidney	Crowe
7	Angelina	Astaire
8	Ed	Mansfield
9	Ray	Johansson
10	Laura	Brody
11	Michelle	McConaughey
12	Frances	Day-Lewis
13	Burt	Temple

e. F

Data Output

	name character varying (25)
1	Documentary
2	New
3	Children
4	Music
5	Travel
6	Foreign
7	Drama
8	Horror
9	Action
10	Comedy
11	Family
12	Games

f. D

Data Output									
	film_id integer	title character varying (255)	description text	release_year integer	language_id smallint	rental_duration smallint	rental_rate numeric (4,2)	length smallint	replacement_cost numeric (5,2)
1	35	Arachnophobia Rollercoaster	A Action-Pack...	2006	1	4	2.99	147	

g. F

Data Output				
	first_name character varying (45)	last_name character varying (45)	c1 bigint	c2 bigint
1	Marcia	Dean	27	15
2	Clara	Shaw	24	18
3	Rhonda	Kennedy	24	15
4	Marsha	Douglas	23	14
5	Lena	Jensen	23	9
6	Florence	Woods	23	7
7	Eleanor	Hunt	22	24
8	Wesley	Bull	22	18
9	Sue	Peters	22	18
10	Marion	Snyder	22	17
11	Elizabeth	Brown	22	16
12	Curtis	Irby	22	16
13	Diane	Collins	22	13

h.

>	address
>	category
▼	category_rating
▼	Columns (4)
	category_id
	category_name
	avg_rate
	max_len
>	Constraints