

SQL

For QA Engineers

PART 1

Part 1-

Database Basics

(types, components and terminology)

Why Data is Important?

Hardware-Software-**Data**

What is Database?

Database (dB) is an organized collection of data

Advantages:

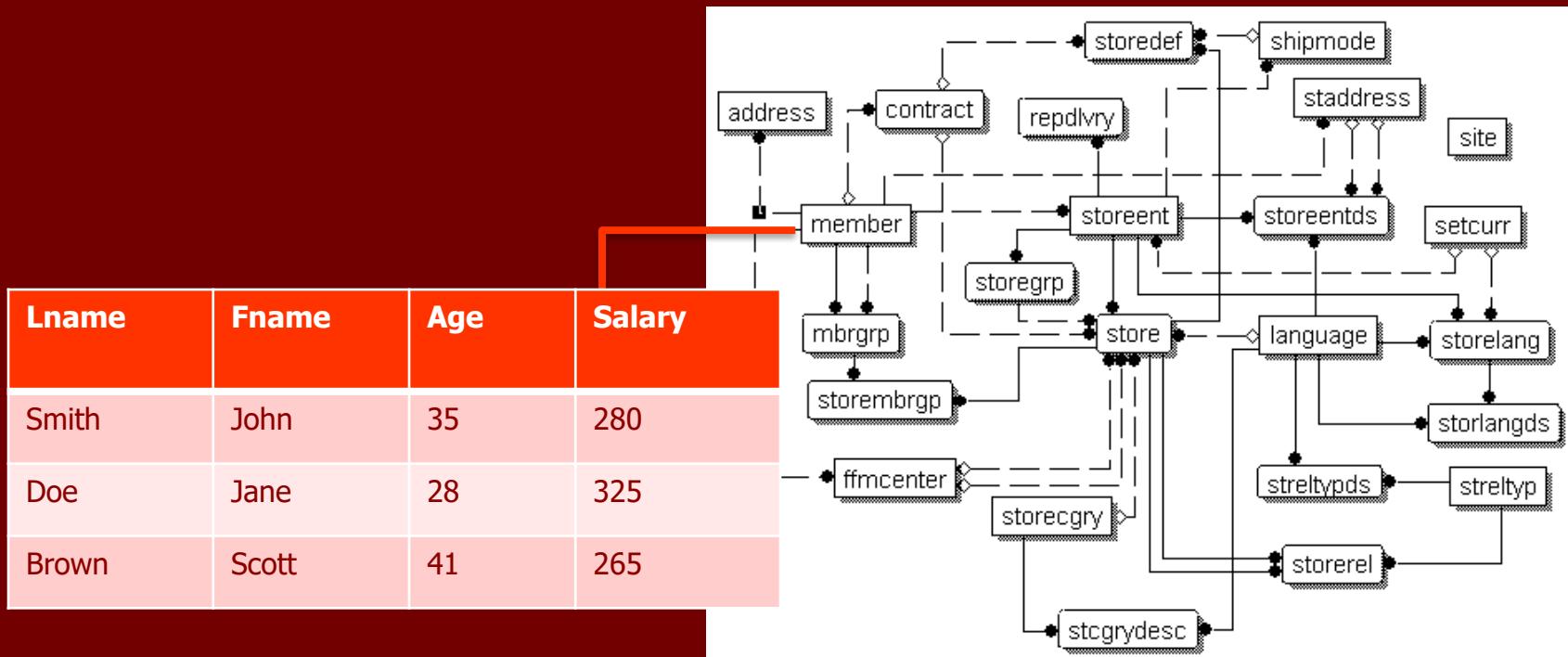
- Control of data accuracy;
- Data consistency;
- Data manipulation (sort, retrieve, update...);
- Data Security;

Disadvantages:

- Complexity, size, cost, maintenance;

Relational Databases

Relational databases use **tables** to store information. It allows you to easily find specific information, sort based on any field, and generate reports that contain only certain fields from each record.

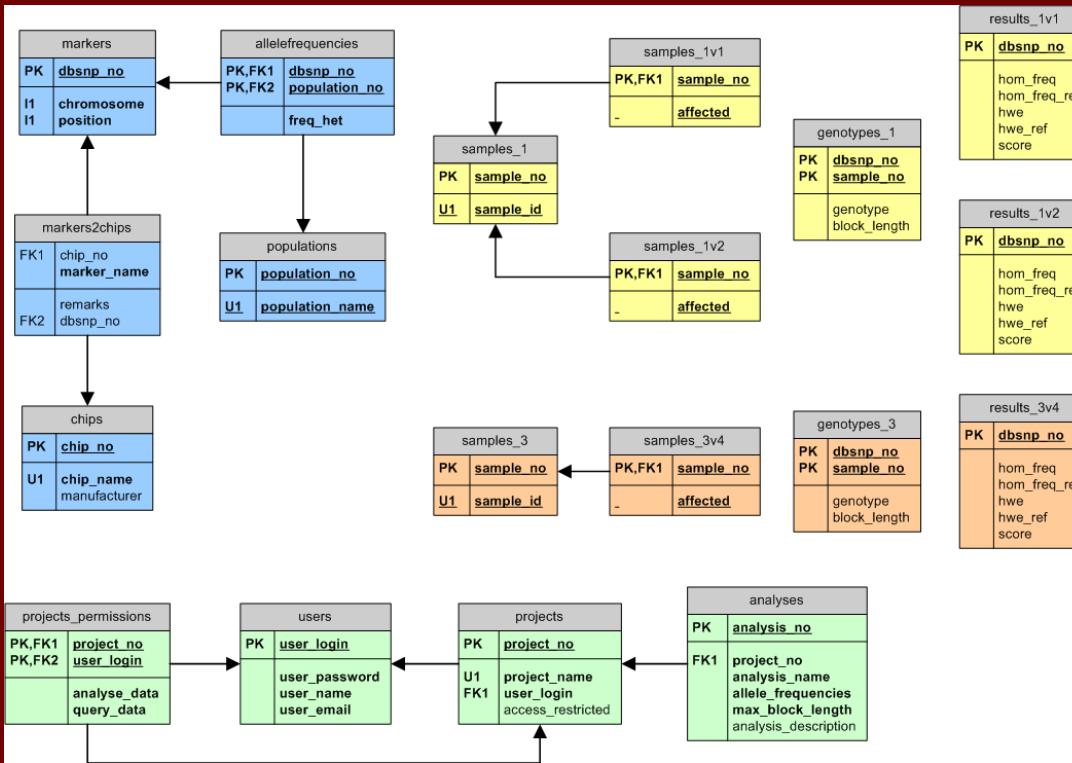


Different Parts of Database

- Schema
- Table
- Field
- Record
- Query
- Report

■ Schema:

- Logical container for database objects (tables, views, triggers) that user creates.



■Table:

Table is a set of data elements (values) that is organized using columns (fields) and rows (records).

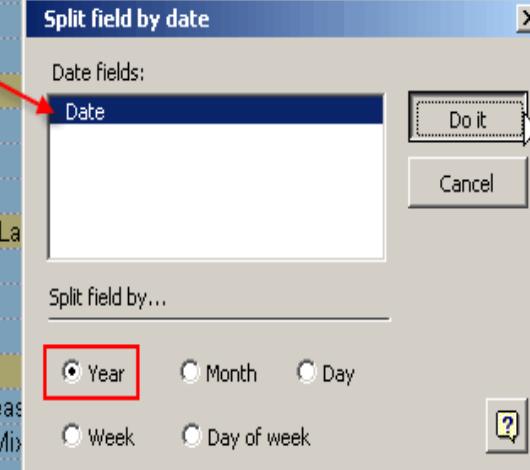
EmployeeAddressTable					
SSN	FirstName	LastName	Address	City	State
512687458	Joe	Smith	83 First Street	Howard	Ohio
758420012	Mary	Scott	842 Vine Ave.	Losantiville	Ohio
102254896	Sam	Jones	33 Elm St.	Paris	New York
876512563	Sarah	Ackerman	440 U.S. 110	Upton	Michigan

Table has a specified number of columns, but can have any number of rows

■ Field:

Field is a
database storage
simplest unit

	A	B	C	D	E	F	G
1							
2	Date	ProductName	Category	UnitPrice			
3	1/2/2003	Chang	Beverages	\$19.00			
4	1/24/2003	Guarana Fantástica					
5	2/3/2003	Sasquatch Ale					
6	2/4/2003	Steeleye Stout					
7	2/7/2003	Côte de Blaye					
8	2/8/2003	Chartreuse verte					
9	2/12/2005	Ipoh Coffee					
10	3/7/2005	Laughing Lumberjack Latte					
11	3/10/2005	Outback Lager					
12	3/15/2005	Rhönbräu Klosterbier					
13	3/16/2005	Lakkalikööri					
14	1/3/2005	Aniseed Syrup					
15	1/4/2004	Chef Anton's Cajun Seafood Gumbo					
16	1/5/2004	Chef Anton's Gumbo Mix					
17	1/6/2004	Grandma's Boysenberry Spread					
18	1/8/2004	Northwoods Cranberry Sauce	Condiments	\$20.00			
19	1/15/2004	Genen Shouyu	Condiments	\$15.50			
20	2/13/2004	Gula Malacca	Condiments	\$19.45			
21	3/1/2004	Sirop d'érape	Condiments	\$28.50			
22	3/3/2004	Vegie-spread	Condiments	\$43.90			
23	3/5/2004	Louisiana Fiery Hot Pepper Sauce	Condiments	\$21.05			
24	3/6/2004	Louisiana Hot Spiced Okra	Condiments	\$17.00			
25	3/17/2004	Original Frankfurter grüne Soße	Condiments	\$13.00			
26	1/16/1999	Pavlova	Confections	\$17.45			
27	1/19/1999	Teatime Chocolate Biscuits	Confections	\$9.20			
28	1/20/1999	Sir Rodney's Marmalade	Confections	\$81.00			
29	1/21/1999	Sir Rodney's Scones	Confections	\$10.00			
30		NuNuCa Nuß-Nougat-Creme	Confections	\$14.00			
31	1/26/1999	Gumbär Gummibärchen	Confections	\$31.23			
32	1/27/1999	Schoggi Schokolade	Confections	\$43.90			
33	2/16/1999	Zaanse koeken	Confections	\$9.50			
34	2/17/1999	Chocolade	Confections	\$12.75			
35	2/18/1999	Maxilaku	Confections	\$20.00			



■ Record:

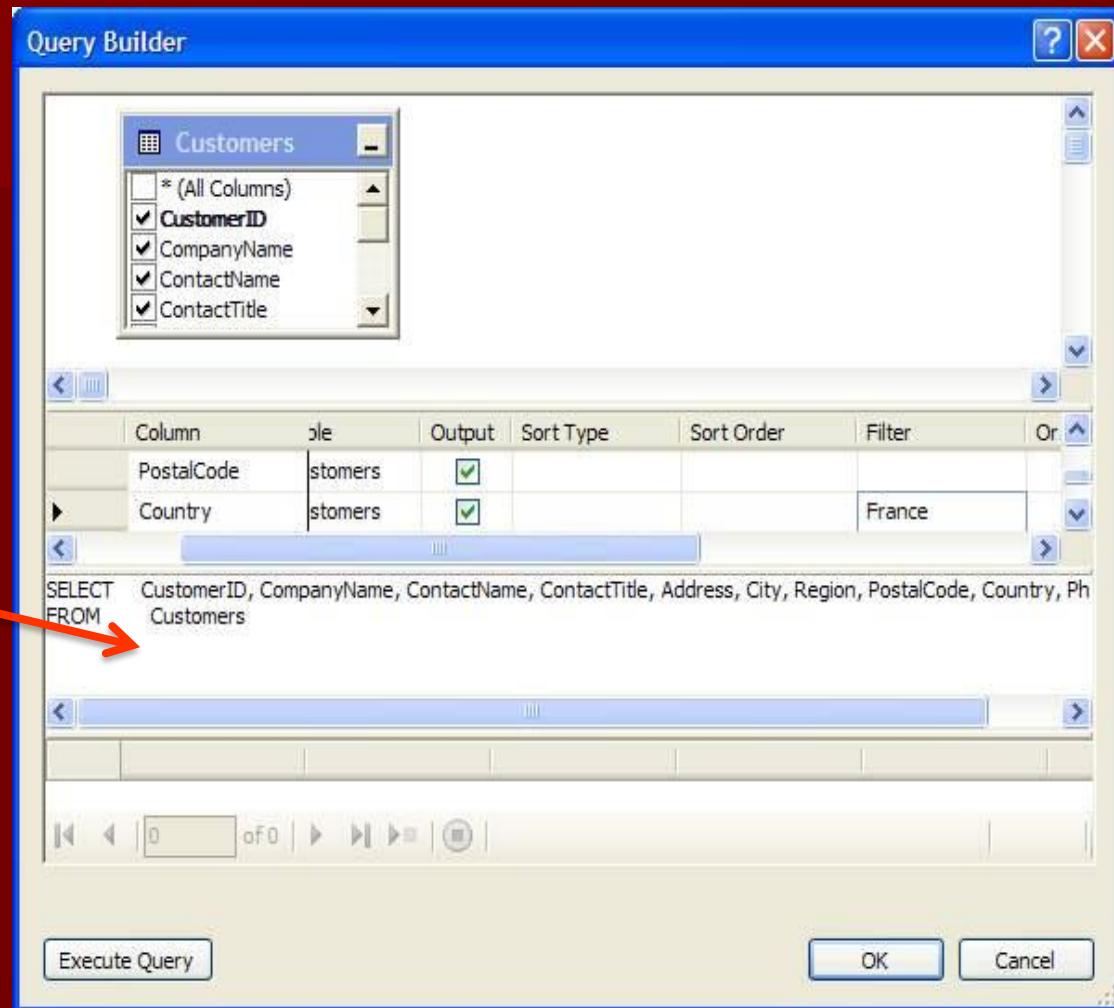
Record is a row and it represents a single structured data set in a table

First Name	Last Name	Address	City	Age	
Mickey	Mouse	123 Fantasy Way	Anaheim	73	
Bat	Man	321 Cavern Ave	Gotham	54	
Wonder	Woman	987 Truth Way	Paradise	39	
Donald	Duck	555 Quack Street	Mallard	65	
Bugs	Bunny	567 Carrot Street	Rascal	58	
Wiley	Coyote	999 Acme Way	Canyon	61	
Cat	Woman	234 Purrfect Street	Hairball	32	
Tweety	Bird	543	Itotitaw	28	

Records

■Query:

Query is your
request to the
database to
retrieve
information



■ Report:

- If the query is a question...
...then the report is its answer
- Reports can be tailored to the needs of the data-user, making the information they extract much more useful

DBMS

- What controls data in a database (DB)?
- How do we get the correct data back?

DBMS (Database Management System) is a software that controls the organization, storage, retrieval, security and integrity of data in a database.

DBMS accepts requests from the application and instructs the operating system to transfer the appropriate data.

DBMS vendors are: Oracle, IBM, Microsoft, and Sybase

Popular Databases

- Oracle
- MS SQL Server
- PostgreSQL (popular open source db)
- MySQL (popular open source db)
- DB2
- ASE

Where a tester can use SQL?

- Getting data for testing;
- Saving data, generated during testing activity;
- Data verifications in databases:
 - to find data;
 - to ensure data integrity;
 - to manipulate test data for specific tests;
- Testing databases;
- Grey Box testing.

Part 2 – Data Storage (Tables, Keys, etc)

Database Tables Example

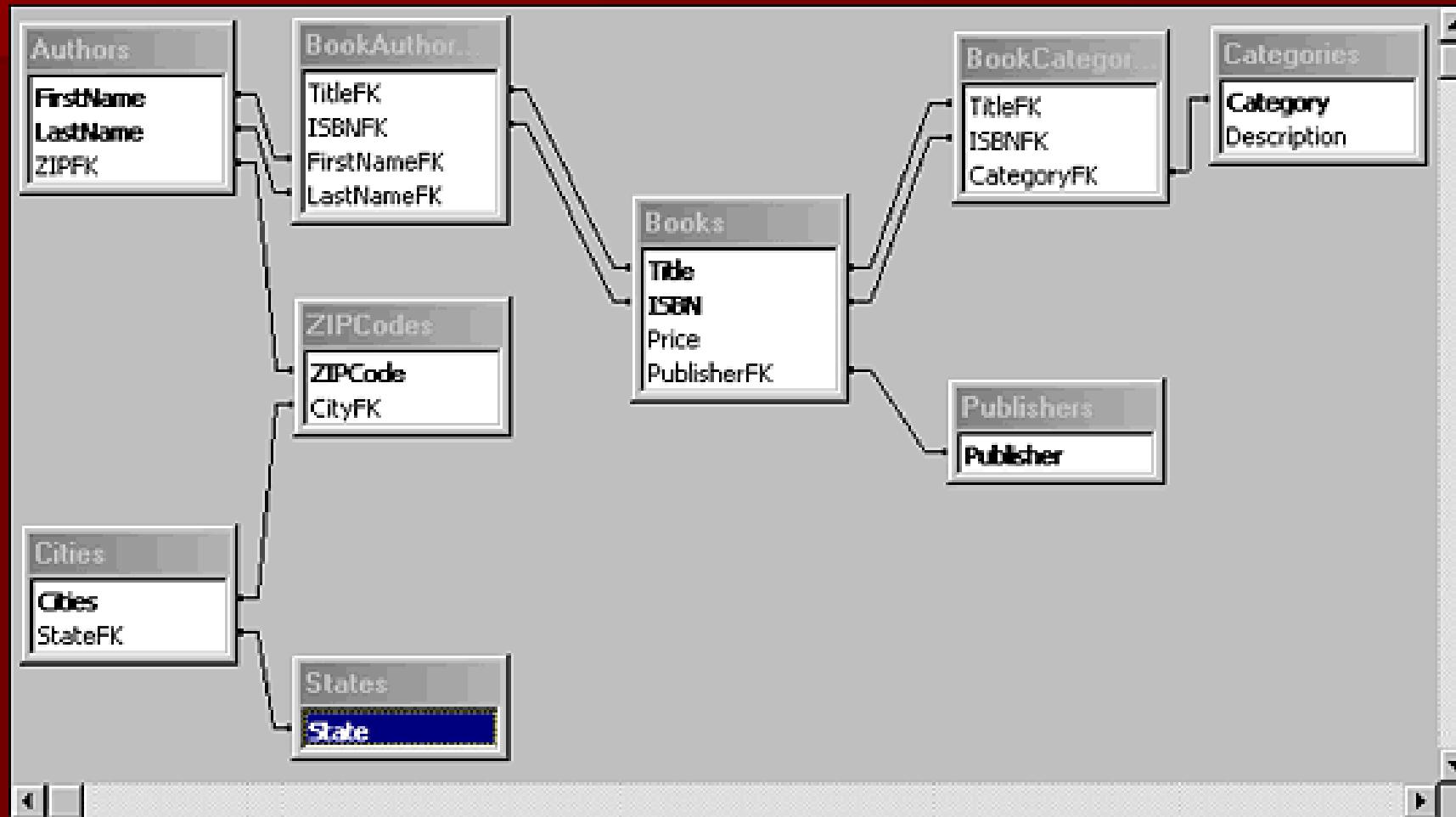


Table “Users”

First Name	Last Name	Sex
Anna	Razgul	F
Boris	Gurich	M
Sri	Lakshmi	M
Girish	Mandari	M

Q: What will happen if I will accidentally add a record again?

A: Duplicate Records!

Table with Duplicate Records

First Name	Last Name	Sex
Anna	Razgul	F
Anna	Razgul	F
Boris	Gurich	M
Sri	Lakshmi	M
Girish	Mandari	M

Q: Is this a problem and how to solve it?

A: Make a record UNIQUE!

Make a Record Unique!

ID	First Name	Last Name	Sex
1	Anna	Razgul	F
2	Boris	Gurich	M
3	Sri	Lakshmi	M
4	Girish	Mandari	M

Q: What is a newly added column "ID" in this table?

A: PRIMARY KEY (PK)

Table “Users”

Each record in the table is now *uniquely* identified

In database terminology we added
PRIMARY KEY (PK)

ID	First Name	Last Name	Sex
1	Anna	Razgul	F
2	Boris	Gurich	M
3	Sri	Lakshmi	M
4	Girish	Mandari	M

"ID" is a primary key (PK) of "Users" table

Primary Key

PRIMARY KEY (PK) is a *unique identifier* of every record in a table

Examples of Primary Keys:

- Social Security Number (SSN)
- Drivers License ID

Can “Last_Name” be a good example of PK?

Add a Column

Now we want to add home phone numbers to each employee.

ID	First Name	Last Name	Sex	Phone
1	Anna	Razgul	F	(111) 111-1111
2	Boris	Gurich	M	(222) 222-2222
3	Sri	Lakshmi	M	(333) 333-3333
4	Anna	Razgul	F	(444) 444-4444

“USERS” table becomes more complicated...

QUESTION:

How can I add Anna's cell phone number?

ID	First Name	Last Name	Sex	Phone
1	Anna	Razgul	F	(111) 111-1111
2	Boris	Gurich	M	(222) 222-2222
3	Sri	Lakshmi	M	(333) 333-3333
4	Anna	Razgul	F	(444) 444-4444

Is there any potential problem I can face?

One value per Record rule

Even if we manage to add data to the existing field, the database will still regard multiple numbers as a single value (atomic value).

ID	First Name	Last Name	Sex	Phone
1	Anna	Razgul	F	(111) 111-1111 (432) 555-4444
2	Boris	Gurich	M	(222) 222-2222
3	Sri	Lakshmi	M	(333) 333-3333
4	Anna	Razgul	F	(444) 444-4444

We cannot have multiple phone numbers listed for a given person (Record) in the same Field (Column)

QUESTION:

Is it a good idea to add another phone column to the table Users? Why?

ID	First Name	Last Name	Sex	PhoneHome	PhoneWork
1	Anna	Razgul	F	(111) 111-1111	(432) 555-4444
2	Boris	Gurich	M	(222) 222-2222	
3	Sri	Lakshmi	M	(333) 333-3333	
4	Anna	Razgul	F	(444) 444-4444	

Most likely 'No'. Why?

Our objective is to:

- Simplify tables (divide larger tables into smaller tables)
- Isolate data so that data modifications can be made in just one table and then propagated through the rest of the database via the defined relationships.

Database Normalization

NORMALIZATION is the process of dividing large tables into smaller tables and defining relationships between them.

SOLUTION:

Is it a good idea to add another phone column to the table “Users”?

First Name	Last Name	Sex	PhoneWork	PhoneCell
Anna	Razgul	F	(111) 111-1111	(432) 555-4444
Boris	Gurich	M	(222) 222-2222	

Solution is **normalization**.

Create two related tables, each serving specific purpose and balance the data and relations across the tables.

Table1 – “Users”

ID	First Name	Last Name	Sex
1	Anna	Razgul	F
2	Boris	Gurich	M
3	Sri	Lakshmi	M
4	Anna	Razgul	F

Table2 – “Phone”

numberID	UserId	PhoneNumber	PhoneType
1001	1	(111) 111 1111	Work
1002	1	(432) 555 4444	Cell
1003	2	(222) 222 2222	Work
1004	3	(333) 333 3333	Work

“PhoneId” is a *primary key* of table “Phone”

Table Relationship

ID	First Name	Last Name	Sex		
1	Anna	Razgul	F		
2	Boris	Gurich	M		
3	Sri	Lakshmi	M		
4	Anna	PhoneId	UserId	PhoneNumber	PhoneType
		1001	1	(111) 111 1111	Work
		1002	1	(432) 555 4444	Cell
		1003	2	(222) 222 2222	Work
		1004	3	(333) 333 3333	Work

We created a link between tables

How to Find Data in Tables?

ID	First Name	Last Name	Sex
1	Anna	Razgul	F
2	Doris	Gurich	M
3	Sri	Lakshmi	M
4	Anna	PhoneId	UserId
		1001	1
		1002	1
		1003	2
		1004	3
		PhoneNumber	PhoneType
		(111) 111 1111	Work
		(432) 555 4444	Cell
		(222) 222 2222	Work
		(333) 333 3333	Work

Primary key “ID” from “Users” table must be present in “Phone” as “UserId”.

What is "UserId" and what if it's not present?

Foreign Key

ID	First Name	Last Name	Sex
1	Anna	Razgul	F
2	Doris	Gurich	M
3	Sri	Lakshmi	M
Phone Id	User Id	Phone Number	Phone Type
1001	1	(111) 111 1111	Work
1002	1	(432) 555 4444	Cell
1003	2	(222) 222 2222	Work
1004	3	(333) 333 3333	Work

“UserId” is a FOREIGN KEY (FK) in “Phone” table

Foreign Key

ID	First Name	Last Name	Sex
1	Anna	Razgul	F
2	Boris	Gurich	M
EntryNumber	OwnerId	PhoneNumber	PhoneType
1001	1	(111) 111 1111	Mobile
1002	2	(222) 222 2222	Home
1003	3	(333) 333 3333	home
1004	1	(432) 555 4444	work

“OwnerID” is a FOREIGN KEY (FK)
in table “Phoneinfo”

Foreign Key

What is a FOREIGN KEY (FK)?

- A foreign key (FK) is a column (or combination of columns) that is used to establish a relationship between the tables;
- Foreign key is usually not unique (one-to-many relation) and shall always point to a primary key.

PK vs. FK

	Primary Key (PK)	Foreign Key (FK)
Purpose	<i>Unique identifier</i> of every record in a table (To make every record in the table unique)	Used to <i>establish a relationship between the tables</i>
Definition	Must be unique within a table; Can't be null;	Usually not unique (one-to-many relation); Always point to a primary key;

Part 3 – Practical SQL

SQL
stands
for
Structured **Q**uery **L**anguage

What is SQL?

SQL is a database computer language designed for the retrieval and management of data in relational databases, database schema creation and modification, and database object access control management.

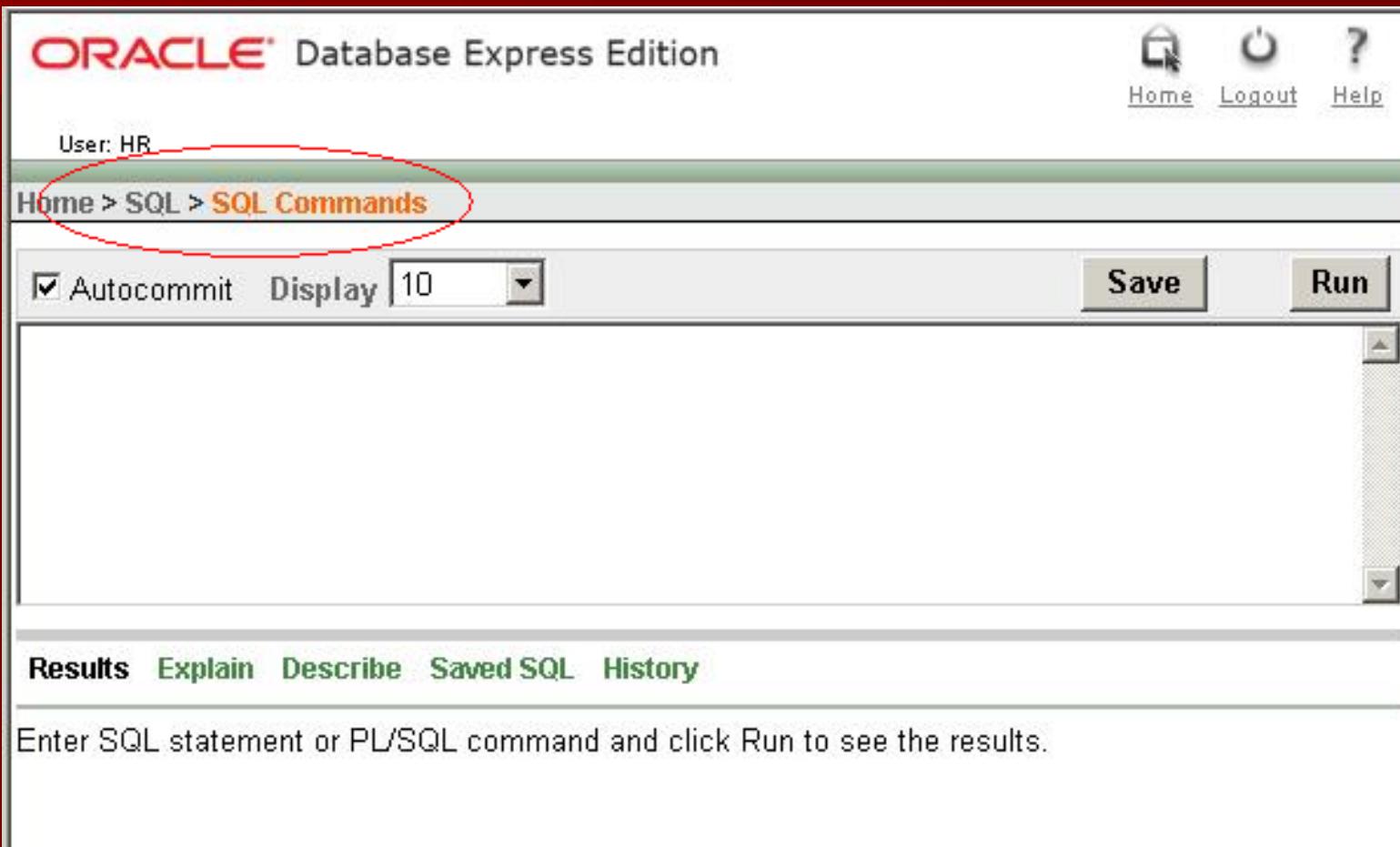
ORACLE

To connect to ORACLE database:

- dB Administrator:
 - Grand user an access (id/psw);
 - Decides which roles and privileges user has;
- User:
 - Log on (connect to a database);
 - Use SQL to communicate with ORACLE database

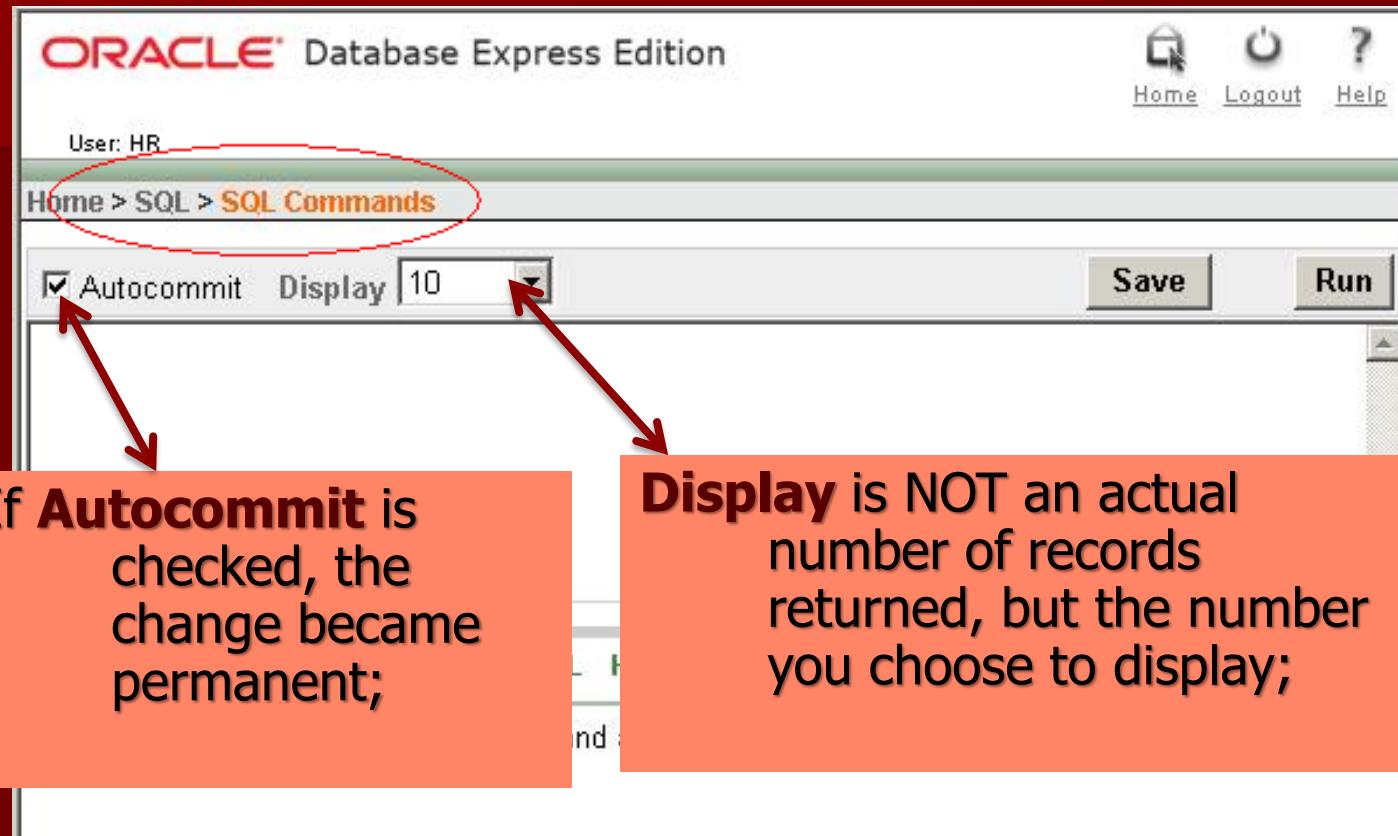
Sooo... We have to learn SQL!

SQL Command Page



The screenshot shows the Oracle Database Express Edition SQL Command Page. At the top, it displays "ORACLE Database Express Edition". Below that, the user is identified as "User: HR". The navigation bar shows "Home > SQL > SQL Commands", with "SQL Commands" highlighted in orange and circled in red. On the left, there are checkboxes for "Autocommit" (which is checked) and "Display" (set to 10), along with "Save" and "Run" buttons. The main area is a large text input field for SQL commands. At the bottom, there is a menu bar with "Results", "Explain", "Describe", "Saved SQL", and "History". A message at the bottom of the page reads: "Enter SQL statement or PL/SQL command and click Run to see the results."

SQL Command Page



Constraint

Constraint is used to define data integrity (restrict the values in a database). Constraint violation... What is it?

Constraint Types (Oracle):

1. NOT NULL constraint - dB value can't be null
2. Unique constraint - multiple rows can't have the same value in the same column
3. Primary Key constraint - combines a NOT NULL and Unique constraints in a single declaration
4. Foreign Key constraint - values in one table match values in another table
5. Check constraint - value in the dB must comply with a specific condition
6. REF constraint - further describe the relationship between REF column and the object it references

Constraint clauses can appear in CREATE/ALTER TABLE, CREATE/ALTER VIEW

DESC

To verify the created table or find column names and datatypes:

DESC[RIBE] <TABLE NAME>

For example:

DESC EMPLOYEES

REVIEW!



- **Database (dB)** is a storage space for content / information (data);
It is an organized collection of data;
- **Relational databases** use **tables** to store information and specify the relations between those tables;
- **Schema** is a logical container for database objects (tables, views, triggers) that user creates;
- **Table** is a container for data elements and relations. It's using columns (**fields**) and rows (**records**);
- **Field** is a database storage simplest unit (table column);
- **Record** is a row and it represents a single structured data set in table;
- **Query** is your request to the database to retrieve *information*;
- **Report** is the returned information to the specified query;
- **DBMS** (Database Management System) is a software that controls the organization, storage, retrieval, security and integrity of data in dB;
- **Popular Databases:** ORACLE, Sybase, MySQL, PostgreSQL, etc.

- PRIMARY KEY (PK) is a *unique identifier* of every record in a table;
- FOREIGN KEY (FK) is a column (or combination of columns) that is used to establish a relationship between the tables; Foreign key is usually not unique (one-to-many relation) and shall always point to a primary key.
- NORMALIZATION is the dB process of organizing the fields and tables of a relationship database to minimize redundancy and dependency. Divide large tables into smaller tables and defining relationships between them.
- SQL stands for Structured Query Language. It is a database computer language, designed to retrieve and manage data, create and modify dB schema, etc.

Interview Questions

1. What is a Database (dB)?
2. What is a Relational database? Advantages?
3. What is a Schema? What are the schema commands?
4. What is a Table? Can we have an empty table?
5. What is Field? What 'part' of the table?
6. What is Record? What 'part' of the table?
7. What is Query?
8. What is Report? How much information can you get in one report?
9. What is DBMS? Name a few popular databases.
10. What is a PRIMARY KEY (PK) / FOREIGN KEY (FK)?