SQL INTERVIEW QUESTIONS

What is the difference between a "where" clause and a "having" clause?

"Where" is a kind of restiriction statement. You use where clause to restrict all the data from DB.Where clause is using before result retrieving. But Having clause is using after retrieving the data.Having clause is a kind of filtering command.

What is the basic form of a SQL statement to read data out of a table?

The basic form to read data out of table is 'SELECT * FROM table_name; ' An answer: 'SELECT * FROM table_name WHERE xyz= 'whatever';' cannot be called basic form because of WHERE clause.

What structure can you implement for the database to speed up table reads?

Follow the rules of DB tuning we have to:

- Properly use indexes (different types of indexes)
- properly locate different DB objects across different tablespaces, files and so on
- create a special space (tablespace) to locate some of the data with special datatype (for example CLOB, LOB and …)

What are the tradeoffs with having indexes?

- Faster selects, slower updates.
- Extra storage space to store indexes. Updates are slower because in addition to updating the table you have to update the index.

What is a "join"?

'Join' used to connect two or more tables logically with or without common field.

What is "normalization"? "Denormalization"? Why do you sometimes want to denormalize?

Normalizing data means eliminating redundant information from a table and organizing the data so that future changes to the table are easier. Denormalization means allowing redundancy in a table. The main benefit of denormalization is improved performance with simplified data retrieval and manipulation. This is done by reduction in the number of joins needed for data processing.

What is a "constraint"?

A constraint allows you to apply simple referential integrity checks to a table. There are four primary types of constraints that are currently supported by SQL Server: PRIMARY/UNIQUE - enforces uniqueness of a particular table column.

DEFAULT - specifies a default value for a column in case an insert operation does not provide one. FOREIGN KEY - validates that every value in a column exists in a column of another table. CHECK - checks that every value stored in a column is in some specified list. Each type of constraint performs a specific type of action. Default is not a constraint. NOT NULL is one more constraint which does not allow values in the specific column to be null. And also it is the only constraint which is not a table level constraint.

What types of index data structures can you have?

An index helps to faster search values in tables. The three most commonly used index-types are: - B-Tree: builds a tree of possible values with a list of row IDs that have the leaf value. Needs a lot of space and is the default index type for most databases. - Bitmap: string of bits for each possible value of the column. Each bit string has one bit for each row. Needs only few space and is very fast.(however, domain of value cannot be large, e.g. SEX(m,f); degree(BS,MS,PHD) - Hash: A hashing algorithm is used to assign a set of characters to represent a text string such as a composite of keys or partial keys, and compresses the underlying data. Takes longer to build and is supported by relatively few databases.

What is a "primary key"?

A PRIMARY INDEX or PRIMARY KEY is something which comes mainly from database theory. From its behavior is almost the same as an UNIQUE INDEX, i.e. there may only be one of each value in this column. If you call such an INDEX PRIMARY instead of UNIQUE, you say something about your table design, which I am not able to explain in few words. Primary Key is a type of a constraint enforcing uniqueness and data integrity for each row of a table. All columns participating in a primary key constraint must possess the NOT NULL property.

What is a "functional dependency"? How does it relate to database table design?

Functional dependency relates to how one object depends upon the other in the database. for example, procedure/function sp2 may be called by procedure sp1. Then we say that sp1 has functional dependency on sp2.

What is a "trigger"?

Triggers are stored procedures created in order to enforce integrity rules in a database. A trigger is executed every time a data-modification operation occurs (i.e., insert, update or delete). Triggers are executed automatically on occurance of one of the data-modification operations. A trigger is a database object directly associated

with a particular table. It fires whenever a specific statement/type of statement is issued against that table. The types of statements are insert,update,delete and query statements. Basically, trigger is a set of SQL statements A trigger is a solution to the restrictions of a constraint. For instance: 1.A database column cannot carry PSEUDO columns as criteria where a trigger can. 2. A database constraint cannot refer old and new values for a row where a trigger can.

Why can a "group by" or "order by" clause be expensive to process?

Processing of "group by" or "order by" clause often requires creation of temporary tables to process the results of the query, which is depending of the resultset can be very expensive.

What is "index covering" of a query?

Index covering means that "Data can be found only using indexes, without touching the tables"

What is a SQL view?

An output of a query can be stored as a view. View acts like small table which meets our criterion. View is a precomplied SQL query which is used to select data from one or more tables. A view is like a table but it doesn't physically take any space. View is a good way to present data in a particular format if you use that query quite often. View can also be used to restrict users from accessing the tables directly.

SQL:

SQL is an English like language consisting of commands to store, retrieve, maintain & regulate access to your database.

SQL*Plus:

SQL*Plus is an application that recognizes & executes SQL commands & specialized SQL*Plus commands that can customize reports, provide help & edit facility & maintain system variables.

NVL:

Null value function converts a null value to a non-null value for the purpose of evaluating an expression. Numeric Functions accept numeric I/P & return numeric values. They are MOD, SQRT, ROUND, TRUNC & POWER.

Date Functions:

Date Functions are ADD_MONTHS, LAST_DAY, NEXT_DAY, MONTHS_BETWEEN & SYSDATE.

Character Functions:

Character Functions are INITCAP, UPPER, LOWER, SUBSTR & LENGTH. Additional functions are GREATEST & LEAST. Group Functions returns results based upon groups of rows rather than one result per row, use group functions. They are AVG, COUNT, MAX, MIN & SUM.

TTITLE & BTITLE:

TTITLE & BTITLE are commands to control report headings & footers.

COLUMN:

COLUMN command define column headings & format data values.

BREAK:

BREAK command clarify reports by suppressing repeated values, skipping lines & allowing for controlled break points.

COMPUTE:

Command control computations on subsets created by the BREAK command.

SET:

SET command changes the system variables affecting the report environment.

SPOOL:

SPOOL command creates a print file of the report.

JOIN:

JOIN is the form of SELECT command that combines info from two or more tables. Types of Joins are Simple (Equijoin & Non-Equijoin), Outer & Self join. Equijoin returns rows from two or more tables joined together based upon an equality condition in the WHERE clause. Non-Equijoin returns rows from two or more tables based upon a relationship other than the equality condition in the WHERE clause. Outer Join combines two or more tables returning those rows from one table that have no direct match in the other table. Self Join joins a table to itself as though it were two separate tables.

Union:

Union is the product of two or more tables.

Intersect:

Intersect is the product of two tables listing only the matching rows.

Minus:

Minus is the product of two tables listing only the non-matching rows.

Correlated Subquery:

Correlated Subquery is a subquery that is evaluated once for each row processed by the parent statement. Parent statement can be Select, Update or Delete. Use CRSQ to answer multipart questions whose answer depends on the value in each row processed by parent statement.

Multiple columns:

Multiple columns can be returned from a Nested Subquery.

Sequences:

Sequences are used for generating sequence numbers without any overhead of locking. Drawback is that after generating a sequence number if the transaction is rolled back, then that sequence number is lost.

Synonyms:

Synonyms is the alias name for table, views, sequences & procedures and are created for reasons of Security and Convenience. Two levels are Public - created by DBA & accessible to all the users. Private - Accessible to creator only. Advantages are referencing without specifying the owner and Flexibility to customize a more meaningful naming convention.

Indexes:

Indexes are optional structures associated with tables used to speed query execution and/or guarantee uniqueness. Create an index if there are frequent retrieval of fewer than 10-15% of the rows in a large table and columns are referenced frequently in the WHERE clause. Implied tradeoff is query speed vs. update speed. Oracle automatically update indexes. Concatenated index max. is 16 columns.

Data types:

Max. columns in a table is 255. Max. Char size is 255, Long is 64K & Number is 38 digits. Cannot Query on a long column. Char, Varchar2 Max. size is 2000 & default is 1 byte. Number(p,s) p is precision range 1 to 38, s is scale -84 to 127. Long Character data of variable length upto 2GB. Date Range from Jan 4712 BC to Dec 4712 AD. Raw Stores Binary data (Graphics Image & Digitized Sound). Max. is 255 bytes. Mslabel Binary format of an OS label. Used primarily with Trusted Oracle.
Order of SQL statement execution Where clause, Group By clause, Having clause, Order By clause & Select.

Transaction:

Transaction is defined as all changes made to the database between successive commits.

Commit:

Commit is an event that attempts to make data in the database identical to the data in the form. It involves writing or posting data to the database and committing data to the database. Forms check the validity of the data in fields and records during a commit. Validity check are uniqueness, consistency and db restrictions.

Posting:

Posting is an event that writes Inserts, Updates & Deletes in the forms to the database but not committing these transactions to the database.

Rollback:

Rollback causes work in the current transaction to be undone.

Savepoint:

Savepoint is a point within a particular transaction to which you may rollback without rolling back the entire transaction.

Set Transaction:

Set Transaction is to establish properties for the current transaction.

Locking:

Locking are mechanisms intended to prevent destructive interaction between users accessing data. Locks are used to achieve.

Consistency:

Assures users that the data they are changing or viewing is not changed until they are thro' with it.

Integrity:

Assures database data and structures reflects all changes made to them in the correct sequence. Locks ensure data integrity and maximum concurrent access to data. Commit statement releases all locks. Types of locks are given below. Data Locks protects data i.e. Table or Row lock. Dictionary Locks protects the structure of database object i.e. ensures table's structure does not change for the duration of the transaction. Internal Locks & Latches protects the internal database structures. They are automatic. Exclusive Lock allows queries on locked table but no other activity is allowed. Share Lock allows concurrent queries but prohibits updates to the locked tables. Row Share allows concurrent access to the locked table but prohibits for a exclusive table lock. Row Exclusive same as Row Share but prohibits locking in shared mode. Shared Row Exclusive locks the whole table and allows users

to look at rows in the table but prohibit others from locking the table in share or updating them. Share Update are synonymous with Row Share.

Deadlock:

Deadlock is a unique situation in a multi user system that causes two or more users to wait indefinitely for a locked resource. First user needs a resource locked by the second user and the second user needs a resource locked by the first user. To avoid dead locks, avoid using exclusive table lock and if using, use it in the same sequence and use Commit frequently to release locks.

Mutating Table:

Mutating Table is a table that is currently being modified by an Insert, Update or Delete statement. Constraining Table is a table that a triggering statement might need to read either directly for a SQL statement or indirectly for a declarative Referential Integrity constraints. Pseudo Columns behaves like a column in a table but are not actually stored in the table. E.g. Currval, Nextval, Rowid, Rownum, Level etc.

SQL*Loader:

SQL*Loader is a product for moving data in external files into tables in an Oracle database. To load data from external files into an Oracle database, two types of input must be provided to SQL*Loader : the data itself and the control file. The control file describes the data to be loaded. It describes the Names and format of the data files, Specifications for loading data and the Data to be loaded (optional). Invoking the loader sqlload username/password controlfilename.

Explain MySQL architecture.

The front layer takes care of network connections and security authentications, the middle layer does the SQL query parsing, and then the query is handled off to the storage engine. A storage engine could be either a default one supplied with MySQL (MyISAM) or a commercial one supplied by a third-party vendor (ScaleDB, InnoDB, etc.)

Explain MySQL locks.

Table-level locks allow the user to lock the entire table, page-level locks allow locking of certain portions of the tables (those portions are referred to as tables), row-level locks are the most granular and allow locking of specific rows.

Explain multi-version concurrency control in MySQL.

Each row has two additional columns associated with it - creation time and deletion time, but instead of storing timestamps, MySQL stores version numbers.

What are MySQL transactions?

A set of instructions/queries that should be executed or rolled back as a single atomic unit.

What's ACID?

Automicity - transactions are atomic and should be treated as one in case of rollback. Consistency - the database should be in consistent state between multiple states in transaction. Isolation - no other queries can access the data modified by a running transaction. Durability - system crashes should not lose the data.

Which storage engines support transactions in MySQL?

Berkeley DB and InnoDB.

How do you convert to a different table type?

ALTER TABLE customers TYPE = InnoDB

How do you index just the first four bytes of the column?

ALTER TABLE customers ADD INDEX (business_name(4))

What's the difference between PRIMARY KEY and UNIQUE in MyISAM?

PRIMARY KEY cannot be null, so essentially PRIMARY KEY is equivalent to UNIQUE NOT NULL.

How do you prevent MySQL from caching a query?

SELECT SQL_NO_CACHE …

What's the difference between query_cache_type 1 and 2?

The second one is on-demand and can be retrieved via SELECT SQL_CACHE … If you're worried about the SQL portability to other servers, you can use SELECT /* SQL_CACHE */ id FROM … - MySQL will interpret the code inside comments, while other servers will ignore it.

What is DDL, DML and DCL?

If you look at the large variety of SQL commands, they can be divided into three large subgroups. Data Definition Language deals with database schemas and descriptions of how the data should reside in the database, therefore language statements like CREATE TABLE or ALTER TABLE belong to DDL. DML deals with data manipulation, and therefore includes most common SQL statements such SELECT, INSERT, etc. Data Control Language includes commands such as GRANT, and mostly concerns with rights, permissions and other controls of the database system.

How do you get the number of rows affected by query?

SELECT COUNT (user_id) FROM users would only return the number of user_id's.

If the value in the column is repeatable, how do you find out the unique values?

Use DISTINCT in the query, such as SELECT DISTINCT user_firstname FROM users; You can also ask for a number of distinct values by saying SELECT COUNT (DISTINCT user_firstname) FROM users;

How do you return the a hundred books starting from 25th?

SELECT book_title FROM books LIMIT 25, 100. The first number in LIMIT is the offset, the second is the number.

You wrote a search engine that should retrieve 10 results at a time, but at the same time you'd like to know how many rows there're total. How do you display that to the user?

SELECT SQL_CALC_FOUND_ROWS page_title FROM web_pages LIMIT 1,10; SELECT FOUND_ROWS(); The second query (not that COUNT() is never used) will tell you how many results there're total, so you can display a phrase "Found 13,450,600 results, displaying 1-10". Note that FOUND_ROWS does not pay attention to the LIMITs you specified and always returns the total number of rows affected by query.

How would you write a query to select all teams that won either 2, 4, 6 or 8 games?

SELECT team_name FROM teams WHERE team_won IN (2, 4, 6, 8)

How would you select all the users, whose phone number is null?

SELECT user_name FROM users WHERE ISNULL(user_phonenumber);

What does this query mean: SELECT user_name, user_isp FROM users LEFT JOIN isps USING (user_id)

It's equivalent to saying SELECT user_name, user_isp FROM users LEFT JOIN isps WHERE users.user_id=isps.user_id

How do you find out which auto increment was assigned on the last insert?

SELECT LAST_INSERT_ID() will return the last value assigned by the auto_increment function. Note that you don't have to specify the table name.

What does –i-am-a-dummy flag to do when starting MySQL?

Makes the MySQL engine refuse UPDATE and DELETE commands where the WHERE clause is not present.

On executing the DELETE statement I keep getting the error about foreign key constraint failing. What do I do?

What it means is that so of the data that you're trying to delete is still alive in another table. Like if you have a table for universities and a table for students, which contains the ID of the university they go to, running a delete on a university table will fail if the students table still contains people enrolled at that university. Proper way to do it would be to delete the offending data first, and then delete the university in question. Quick way would involve running SET foreign_key_checks=0 before the DELETE command, and setting the parameter back to 1 after the DELETE is done. If your foreign key was formulated with ON DELETE CASCADE, the data in dependent tables will be removed automatically.

When would you use ORDER BY in DELETE statement?

When you're not deleting by row ID. Such as in DELETE FROM techinterviews_com_questions ORDER BY timestamp LIMIT 1. This will delete the most recently posted question in the table techinterviews_com_questions.

How can you see all indexes defined for a table?

SHOW INDEX FROM techinterviews_questions;

How would you change a column from VARCHAR(10) to VARCHAR(50)?

ALTER TABLE techinterviews_questions CHANGE techinterviews_content techinterviews_CONTENT VARCHAR(50).

How would you delete a column?

ALTER TABLE techinterviews_answers DROP answer_user_id.

How would you change a table to InnoDB?

ALTER TABLE techinterviews_questions ENGINE innodb;

When you create a table, and then run SHOW CREATE TABLE on it, you occasionally get different results than what you typed in. What does MySQL modify in your newly created tables?

VARCHARs with length less than 4 become CHARs.

CHARs with length more than 3 become VARCHARs.

NOT NULL gets added to the columns declared as PRIMARY KEYs.

Default values such as NULL are specified for each column

How do I find out all databases starting with 'tech' to which I have access to?

SHOW DATABASES LIKE 'tech%';

How do you concatenate strings in MySQL?

CONCAT (string1, string2, string3)

How do you get a portion of a string?

SELECT SUBSTR(title, 1, 10) from techinterviews_questions;

What's the difference between CHAR_LENGTH and LENGTH?

The first is, naturally, the character count. The second is byte count. For the Latin characters the numbers are the same, but they're not the same for Unicode and other encodings.

How do you convert a string to UTF-8?

SELECT (techinterviews_question USING utf8);

What do % and _ mean inside LIKE statement?

% corresponds to 0 or more characters, _ is exactly one character.

What does + mean in REGEXP?

At least one character.

How do you get the month from a timestamp?

SELECT MONTH(techinterviews_timestamp) from techinterviews_questions;

How do you offload the time/date handling to MySQL?

SELECT DATE_FORMAT(techinterviews_timestamp, '%Y-%m-%d') from techinterviews_questions; A similar TIME_FORMAT function deals with time.

How do you add three minutes to a date?

ADDDATE(techinterviews_publication_date, INTERVAL 3 MINUTE)

What's the difference between Unix timestamps and MySQL timestamps?

Internally Unix timestamps are stored as 32-bit integers, while MySQL timestamps are stored in a similar manner, but represented in readable YYYY-MM-DD HH:MM:SS format.

How do you convert between Unix timestamps and MySQL timestamps?

UNIX_TIMESTAMP converts from MySQL timestamp to Unix timestamp, FROM_UNIXTIME converts from Unix timestamp to MySQL timestamp.

What are ENUMs used for in MySQL?

You can limit the possible values that go into the table. CREATE TABLE months (month ENUM 'January', 'February', 'March',…); INSERT months VALUES ('April');

How are ENUMs and SETs represented internally?

As unique integers representing the powers of two, due to storage optimizations.

How do you start and stop MySQL on Windows?

net start MySQL, net stop MySQL

How do you start MySQL on Linux?

/etc/init.d/mysql start

Explain the difference between mysql and mysqli interfaces in PHP?

mysqli is the object-oriented version of mysql library functions.

What's the default port for MySQL Server?

3306

What does tee command do in MySQL?

tee followed by a filename turns on MySQL logging to a specified file. It can be stopped by command notee.

Can you save your connection settings to a conf file?

Yes, and name it ~/.my.conf. You might want to change the permissions on the file to 600, so that it's not readable by others.

How do you change a password for an existing user via mysqladmin?

mysqladmin -u root -p password "newpassword"

Use mysqldump to create a copy of the database?

mysqldump -h mysqlhost -u username -p mydatabasename > dbdump.sql

Have you ever used MySQL Administrator and MySQL Query Browser?

Describe the tasks you accomplished with these tools.

What are some good ideas regarding user security in MySQL?

There is no user without a password. There is no user without a user name. There is no user whose Host column contains % (which here indicates that the user can log in from anywhere in the network or the Internet). There are as few users as possible (in the ideal case only root) who have unrestricted access.

Explain the difference between MyISAM Static and MyISAM Dynamic.

In MyISAM static all the fields have fixed width. The Dynamic MyISAM table would include fields such as TEXT, BLOB, etc. to accommodate the data types with various lengths. MyISAM Static would be easier to restore in case of corruption, since even though you might lose some data, you know exactly where to look for the beginning of the next record.

What does myisamchk do?

It compressed the MyISAM tables, which reduces their disk usage.

Explain advantages of InnoDB over MyISAM?

Row-level locking, transactions, foreign key constraints and crash recovery.

Explain advantages of MyISAM over InnoDB?

Much more conservative approach to disk space management - each MyISAM table is stored in a separate file, which could be compressed then with myisamchk if

needed. With InnoDB the tables are stored in tablespace, and not much further optimization is possible. All data except for TEXT and BLOB can occupy 8,000 bytes at most. No full text indexing is available for InnoDB. TRhe COUNT(*)s execute slower than in MyISAM due to tablespace complexity.

What are HEAP tables in MySQL?

HEAP tables are in-memory. They are usually used for high-speed temporary storage. No TEXT or BLOB fields are allowed within HEAP tables. You can only use the comparison operators = and <=>. HEAP tables do not support AUTO_INCREMENT. Indexes must be NOT NULL.

How do you control the max size of a HEAP table?

MySQL config variable max_heap_table_size.

What are CSV tables?

Those are the special tables, data for which is saved into comma-separated values files. They cannot be indexed.

Explain federated tables.

Introduced in MySQL 5.0, federated tables allow access to the tables located on other databases on other servers.

What is SERIAL data type in MySQL?

BIGINT NOT NULL PRIMARY KEY AUTO_INCREMENT

What happens when the column is set to AUTO INCREMENT and you reach the maximum value for that table?

It stops incrementing. It does not overflow to 0 to prevent data losses, but further inserts are going to produce an error, since the key has been used already.

Explain the difference between BOOL, TINYINT and BIT.

Prior to MySQL 5.0.3: those are all synonyms. After MySQL 5.0.3: BIT data type can store 8 bytes of data and should be used for binary data.

Explain the difference between FLOAT, DOUBLE and REAL.

FLOATs store floating point numbers with 8 place accuracy and take up 4 bytes. DOUBLEs store floating point numbers with 16 place accuracy and take up 8 bytes. REAL is a synonym of FLOAT for now.

If you specify the data type as DECIMAL (5,2), what's the range of values that can go in this table?

999.99 to -99.99. Note that with the negative number the minus sign is considered one of the digits.

What happens if a table has one column defined as TIMESTAMP?

That field gets the current timestamp whenever the row gets altered.

But what if you really want to store the timestamp data, such as the publication date of the article?

Create two columns of type TIMESTAMP and use the second one for your real data.

Explain data type TIMESTAMP DEFAULT CURRENT_TIMESTAMP ON UPDATE CURRENT_TIMESTAMP

The column exhibits the same behavior as a single timestamp column in a table with no other timestamp columns.

What does TIMESTAMP ON UPDATE CURRENT_TIMESTAMP data type do?

On initialization places a zero in that column, on future updates puts the current value of the timestamp in.

Explain TIMESTAMP DEFAULT '2006:09:02 17:38:44' ON UPDATE CURRENT_TIMESTAMP.

A default value is used on initialization, a current timestamp is inserted on update of the row.

If I created a column with data type VARCHAR(3), what would I expect to see in MySQL table?

CHAR(3), since MySQL automatically adjusted the data type.