Database Systems Overview Homework

1. **What database models do you know?**

* Relational (table)
* Hierarchical (tree)
* Network / graph
* Object-oriented

1. **Which are the main functions performed by a Relational Database Management System (RDBMS)?**

* Creating/altering/deleting tables and relationships between them (database schema)
* Adding, changing, deleting, searching and retrieving of data stored in the tables
* Support for the SQL language
* Transaction management (optional)

1. **Define what is "table" in database terms.**

* A table is a set of data elements (values) that is organized using a model of vertical columns (which are identified by their name and have type of the data stored in them) and horizontal rows, the cell being the unit where a row and column intersect. A table is another term for relation but tables allow duplicate rows and relations don’t.

1. **Explain the difference between a primary and a foreign key.**

* **Primary key** is a column in the table that uniquely identifies its rows (usually it’s a number but not obligatory)
* **Foreign key** is an identifier of a record located in another table (usually its primary key but not obligatory)

1. **Explain the different kinds of relationships between tables in relational databases.**

* **Relationship “one-to-many”:** a single record in the first table has many corresponding records in the second table
* **Relationship “many-to-many”:** records in the first table have many corresponding records in the second one and vice versa
* **Relationship “one-to-one”:** a single record in a table corresponds to a single record in the other table
* **Self-relationship:** a column that is considered a foreign key points to another column that is considered a primary key according to some relationship between the data in these columns

1. **When is a certain database schema normalized? What are the advantages of normalized databases?**

* A certain database schema is normalized when the repeating data in the tables of the database is removed. The advantages of normalized databases are: fast searching, sorting and creating indexes, greater overall database organization, reduction of redundant data, data consistency within the database, a much more flexible database design, a better handle on database security.

1. **What are database integrity constraints and when are they used?**

* Database integrity constraints ensure data integrity in the tables. They enforce data rules which cannot be violated. They are used we have to limit the type of data that can go into a table - this ensures the accuracy and reliability of the data in the database.

1. **Point out the pros and cons of using indexes in a database.**

* **Pros:** fast searching of values in a certain column or group of columns
* **Cons:** slow adding and deleting of records in tables

1. **What's the main purpose of the SQL language?**

* The main purpose of the SQL language is managing data held in a relational database management system. This includes creating, altering, deleting tables and other objects in the database, searching, retrieving, inserting, modifying and deleting rows in tables. It consist of DDL (data definition language – creates, changes and deletes tables) and DML (data manipulation language – executes various queries over the data in the database).

1. **What are transactions used for? Give an example.**

* Transactions are used for ensuring that a sequence of database operations will be executed as a single unit – they will all be executed successfully or none of them will be executed. An example for transaction usage is the scenario when two people using the same account want to buy a stock from an internet store at the same time – one of them will succeed but the second one will be pleased to try again later because the database will lock the amount of money in the account.

1. **What is a NoSQL database?**

* A NoSQL database provides a mechanism for storage and retrieval of data that uses looser consistency models rather than traditional relational databases. NoSQL databases are often highly optimized key–value stores intended for simple retrieval and appending operations, with the goal being significant performance benefits in terms of latency and throughput.

1. **Explain the classical non-relational data models.**

* **Document model**
  + Set of documents, e.g. JSON strings
* **Key-value model**
  + Set of key-value pairs
* **Hierarchical key-value**
  + Hierarchy of key-value pairs
* **Wide-column model**
  + Key-value model with schema
* **Object model**
  + Set of OOP-style objects

1. **Give few examples of NoSQL databases and their pros and cons.**

* **MongoDB**
  + **Pros:** Schema-less; Ease of scale-out; Not expensive; Level of consistency can be chosen depending on the value of the data
  + **Cons:** Data size is typically higher; Less flexibity with querying; Does not support transactions
* **Memcached**
  + **Pros:** Caching; Rate limiting; Real-time analytics
  + **Cons:** Serialization; Replication; Not fault tolerant
* **Cassandra**
  + **Pros:** Native compression reduces complexity and growth is easier to manage with new nodes; Better choice for massive scale (both in storage growth and in throughput of inserts); Does multi-datacenter installations well
  + **Cons:** Difficult to setup and manage in smaller environments; Inappropriate for companies constrained by resources