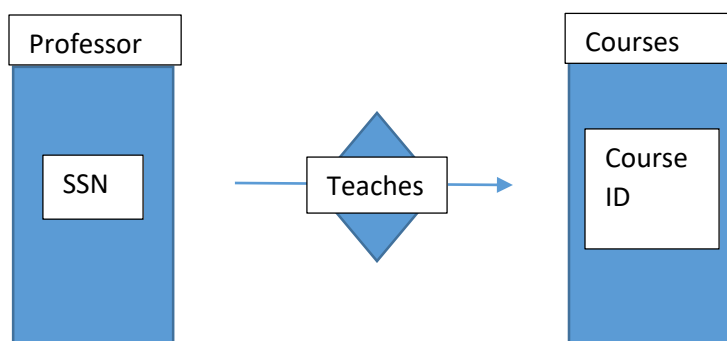


1. List four significant differences between a file-processing system and a DBMS. Why would you choose a database system instead of simply storing data in operating system files? When would it make sense not to use a database system?
  - a. File processing systems have data redundancy and inconsistency as well as security issues. Whereas DBMS have efficient data access, crash recovery and data independence. We would choose using a database system when we need security of the data as well as when we want there to not be redundant data. We don't use database systems when there is a small amount of data, or when efficiency is not a major concern
2. Explain the concept of physical data independence, and its importance in database systems.
  - a. Physical data independence is when a schema is changed without another schema being changed as well. This is very important because it can make the program faster, and it would be more secure.
3. Explain the difference between two-tier and three-tier architectures.
  - a. 2 tier is client and server communication whereas 3-teir is an applications server in between the client and the data source.
4. What is a transaction? Define ACID and provide a brief description of each.
  - a. A transaction is a collection of operations that perform a single logical function. Acid is an acronym for atomicity, consistency, isolation, and durability. Atomicity is when all or none of the operations of the transaction are in the database. Consistency is that the database preserves the consistency of the database. Isolation is when multiple transactions execute separately not knowing of other transactions. Durability is when a transaction successfully completes even if there is a system failure.
5. What is a foreign key constraint? Why are such constraints important? What is referential integrity?
  - a. A foreign key constraint is used to prevent destruction of links between tables. These constraints are very important because it keeps the integrity of the tables intact. Referential integrity is when a column of a table exists in another table.
6. What is the difference between a super key, a candidate key, and a primary key?
  - a. Super key is a way to identify a tuple. A candidate key is a super key that is minimal size. Primary key is a super key that is selected.
7. A university database contains information about professors (identified by social security number, or SSN) and courses (identified by courseid). Professors can teach the same course in several semesters, and only the most recent such offering needs to be recorded. Draw a diagram that describes it (assuming no further constraints hold).



8. Consider the instance of the Students relation shown below
- Age cannot be a candidate key because multiple ages are the same value
  - Sid and login are candidate keys.
  - Sid is valid for a primary key I wouldn't choose an email for a primary key because email addresses can be updated so you would need to change primary keys every time that happens.
9. Given two relations R1 and R2, where R1 contains N1 tuples, R2 contains N2 tuples, and  $N2 > N1 > 0$ , give the minimum and maximum possible sizes (in tuples) for the resulting relation produced by each of the following relational algebra expressions. In each case, state any assumptions about the schemas for R1 and R2 needed to make the expression meaningful:
- $N1 + N2$ , without common columns
  - Same columns only between N1 and N2
  - Only the different columns between N1 and N2
  - All combinations between N1 and N2 (Cartesian product)
10. Consider the following bank database. What are the appropriate primary and foreign keys?
- Primary keys – loan number, customer name, account number, branch name
  - Foreign keys – loan number, account number, branch name.
11. Give an expression in the relational algebra to express each of the following queries
- $\pi_{employee\ name}(\sigma_{Wells\ Fargo}(works))$
  - $\pi_{person\ name}(\sigma_{Wells\ Fargo}(works \bowtie employee)) \cup \pi_{cities}(\sigma_{Wells\ Fargo}(works \bowtie employee))$
  - $\pi(\sigma_{Wells\ Fargo} \wedge salary < 50,000 (works \bowtie employee) (works))$
  - $\pi_{employee\ name}(\sigma_{City}(Works \bowtie Employee))$
  - $\pi_{company\ name}(\sigma_{City}(Company) \cap \sigma_{Wells\ Fargo}(Company))$
12. Consider the following relations, calculate the operations:

- a. Union

7274	Robinson	37
7432	O'Malley	39
9824	Darkes	38
9297	O'Malley	56

- b. Intersection

7432	O'Malley	39
9824	Darkes	38

- c. Set difference

7274	Robinson	37
9297	O'Malley	56

13.

Smith	A	Venus
Black	A	Venus
Black	B	Mars

All four joins, inner, left, right and full are all the same table.

14.

7274	Venus
7432	Venus
9824	Mars
7274	Mars
7432	Earth
9824	Venus

Venus
Mars

7274
9824

Division