**Jennifer Faer**

**IS 5313 Structured Data and Querying**

**Dr. Bill Rosener**

**Fall 2024**

https://billrosener.com/teaching/is4293/assignments/assignment2/assignment2-description.html

Chapter 2: Review Questions

1. Why is the relational model important?

It is the single most important standard in database processing and is used for the design and implementation of almost every commercial database worldwide. This model provides a robust framework for organizing and managing data. (Page 85).

1. Define the term **entity** and give an example of an entity (other than the one from this chapter).

An entity is anything that we want to keep track of in a database. Examples: people (employees, customers), places (addresses, cities, counties, monuments, vacation destinations), and things (products, dates, prices, inventory counts, email addresses, social security numbers). (glossary and page 86.)

1. List the characteristics a table must have to be considered a relation. Define the term **domain**, and explain the significance of the **domain integrity constraint** to a relation.

A relation is a table with the following characteristics:

1. Each row contains data about an entity
2. Each column contains data about an attribute of the entity
3. All entries in a column are of the same data type
4. Each column has a unique name
5. Each cell must hold a single value
6. Column order is not important
7. Row order is not important
8. No two rows are identical

Domain is the type of data allowed in a column, and the domain integrity constraint limits the values of the data. For example, if the column’s domain is defined as an integer, then the database will reject any non-integer entries. (glossary and page 86)

1. Give an example of a relation (other than one from this chapter).

The inventory of an RV repair company:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **PartID** | **PartName** | **PartMfgNumber** | **SellingPrice** | **InventoryCount** |
| 1 | 310 Toilet Bone w/ Upgraded Seat | Dometic 9108923942 | $300.00  | 3 |
| 2 | Safe-t-Alert Propane sensor 30-442-pbr | Safe-T-Alert 30-442-pbr | $80.39  | 1 |
| 3 | 24 in GALVANIZED steel plate, per ft | GALVMET01 | $5.00  | 0 |
| 4 | Digital Thermostat | Airxcel 9430A3382 | $90.00  | 1 |
| 5 | 300 Series Vacuum Breaker w/ Sprayer | Dometic 385319054 | $81.71  | 1 |

1. Give an example of a table that is not a relation (other than one from this chapter).

An anonymized version of our extremely frustrating client database:

|  |  |  |  |
| --- | --- | --- | --- |
| **CustomerID** | **Name** | **Address** | **Email** |
| 1 | Bob Smith | Klamath River RV Park, Klamath CA | bobsmith@example.com, bobsspouse@example.com |
| 2 | Jane Smith | 900 Klamath Beach Rd, Klamath, CA 95548 | bobsspouse@example.com |
| 3 | Tycho | Village Camper Inn | tycho@tychothecat.com |
| 4 | Kepler Faer | Lighthouse Cove RV Park, 900 Sunset Cir, Crescent City | kepler@tychothecat |

1. Under what circumstances can an attribute of a relation be of variable length?

When the data itself does not have a fixed size. Examples: strings for names or email addresses; lists for things like tags or bookmarks; and numerical data such as inventory counts or decimals where the number of decimal places can vary.

1. Explain the use of the terms **file**, **record**, and **field**.

File is the same as a table (or relation), record is the same as a row (or tuple), and field is the same as column (or attribute). (page 86)

1. Explain the use of the terms **relation**, **tuple**, and **attribute**.

See 2.7 above.

1. Under what circumstances can a relation have duplicate rows?

Very large DBMS’s may have duplicate rows because the system might not automatically check for duplicates due to the increased processing times involved. (page 88). Some other scenarios where duplicate rows may happen include older poorly designed databases where the record does not have a primary key, lack of constraints, or issues with importing data from multiple siloed sources into one new all-in-one database (such as consolidating customer data from scheduling software, bookkeeping software, and job tracking software into something like Salesforce).

1. Define the term **unique key** and give an example.

A unique key is a non-repeated key that is used to identify a single row (glossary). An example of a unique key is the primary key given to records in a database, or an individual’s social security number.

1. Define the term **nonunique key** and give an example.

A nonunique key is a key that could identify more than one row (glossary). A nonunique key could be a veterinarian’s CustomerID which is associated with each of the customer’s pets. If they have more than one pet being treated by this vet, the CustomerID may be used multiple times in the table of unique pets. Another example would be a home address, which may be used by everyone living at this home.

1. Give an example of a relation with a unique composite key.

A flight schedule relation could have a unique composite key made of the flight number, departure date, and departure time. Some airlines use the same flight number to refer to the same flight path (example: Los Angeles to Phoenix), and each individual flight can be differentiated by the combination of departure dates and times.

1. Define the terms **candidate key** and **primary key**. Explain the difference between a primary key and a candidate key. Explain the significance of the entity integrity constraint to a primary key.

A candidate key is a potential primary key. It is a unique attribute that could be used to identify a record. Ultimately, one candidate key is selected to become the primary key. (page 90)

The entity integrity constraint requires that the primary key is unique and not null. (glossary) This means that the candidate key must always have a value and is never repeated in any record in the relation.

1. Describe four uses of a primary key.
2. Uniquely identifies records
3. Ensures data integrity (per the entity integrity constraint)
4. Represents the table (or record) in relation to other tables
5. Improves efficiency in processing (only need to search for a single primary key, rather than an entire record)
6. What is a **surrogate key**, and under what circumstances would you use one?

A surrogate key is a short, unchanging, numeric key that has been added to a table to act as a primary key. It has no meaning to users, is generally hidden on reports and forms, and is usually automatically assigned by the DBMS every time a record is added (page 94). A surrogate key is useful when the candidate keys may contain sensitive information (such as a social security number), when candidate keys are too long or complex to be a suitable primary key, or when futureproofing because there may be some concern about the future uniqueness or usefulness of candidate key.

1. How do surrogate keys obtain their values?

Surrogate keys generally obtain their values automatically when the new record is created, such as with AutoNumber in MS Access or the AUTO\_INCREMENT command in SQL. (page 95)

1. Why are the values of surrogate keys normally hidden from users on forms, queries, and reports?

The values of surrogate keys are normally hidden from users because that entity has no meaning to users. (page 94)

1. Explain the term **foreign key** and give an example.

A foreign key in one relation is the primary key of a record in another relation, which is used to denote a relationship between the two relations. (page 96) An example would be the A veterinarian’s table of pets, where one column is the CustomerID to indicate which human client owns which pets.

1. Explain how primary keys and foreign keys are denoted in this book.

Primary keys are denoted with underlining, and foreign keys are denoted with italics. (page 97)

1. Define the term **referential integrity constraint** and give an example of one. How does the referential integrity constraint contribute to database integrity?

The referential integrity constraint says that the foreign key in one table must be the primary key of another table, ensuring reliable links between tables. (glossary)

1. Explain three possible interpretations of a null value.
2. The value is unknown
3. The value is not valid
4. The value is blank
5. Give an example of a null value (other than one from this chapter), and explain each of the three possible interpretations for that value.
6. Define the terms **functional dependency** and **determinant,** using an example not from this book.

In list of customers, a phone number may be null because it is missing an area code, contains a letter, or the information was not provided by the customer.

1. In the following equation, name the functional dependency and identify the determinant(s):

**Area = Length × Width**

Area is functionally dependent on length and width, making length and width the determinants.

(Length, Width) -> Area

2.25 Explain the meaning of the following expression:

**A 🡪 (B, C)**

Given this expression, tell if it is also true that:

**A 🡪 B**

and

**A 🡪 C**

B and C are functionally dependent on A (making A the determinant). If the value of A is known, then it is possible to know the values of both B and C, making the second statements (A -> B and A -> C) true.