What happens on the internet in one minute ?

1) Nearly 640,000 GB of IP address is transferred across the globe.
2) More than 204 million emails are sent.
3) At least 6 million pages are viewed on Facebook around the world.
4) More than 1.3 million video clips are watched on YouTube.
5) 2+ million search queries are made on Google.
6) 47,000 apps are downloaded.
7) 6 new Wikipedia articles are published.

The world is getting flooded with data every minute.

Points to ponder upon…

1) Where do they store this data ?
2) How do they manage this data ?
3) How do they organize this data ?
4) How do they process it , retrieve it , analyze it ?

The answer to these questions lies in understanding the DBMS.

DBMS

Data Base Management Systems

By

Anand Joshi

Why should I study DBMS ?

Motivation:

1) To answer the above questions.

2) Great Job opportunities.

3) Emerging markets-
   - Data mining
   - Data warehousing
   - Database Management
   - Data Analyst
   - Big Data

What is a Database ?

• A **Database** is a collection of logically related data.

• **Database Management** is the task of maintaining databases so that the information is easily available.

What is a DBMS ?

• A DBMS is a software used to store and manage data. The main objectives of any DBMS are to –

- Provide an easy way to store, update, and retrieve data from a database.

- Manage information about users who interact with the DBMS, and the tasks that these users can perform on the data.

DBMS then…

File Processing Systems:

- Data is stored in a text file.
- Examples : A CSV file, Spreadsheet, XML.

Drawbacks:

- Data redundancy and Data inconsistency.
- Difficulty in accessing the data.
- Data isolation – Scattering of data in different files.
- Data integrity might be lost.
- Type Checks.
- Backup , recovery features not present.

DBMS now…

- The newer DBMS solves all the above mentioned problems.
- They are ACID compliant.
  - **Atomic**
  - **Consistent**
  - **Isolated**
  - **Durable**

• **Transaction :** A series of data manipulation operations that must be committed into the database as one whole operation.

• They follow the ACID rules.

• **Atomic :** All operations should execute or no operation.

• **Consistent :** After the operations execute, database is in the consistent state as it was before the execution.

• **Isolated :** Appropriate locks are placed on shared data to isolate the operation on the data.

• **Durability :** Data must be persisted on the hard disk, so that if there is a crash the data is durable.

Data Abstraction

- Abstraction means to hide something.
- Data abstraction hides the complexity of data from the user.
- Three levels of abstraction in a DBMS :
    - Physical (Internal) level.
    - Conceptual (logical) level.
    - View (External) level.

1) Physical Level :
- Lowest level
- How data stored
- Data Structure

Job of the application programmer

2) Logical Level :
- Next higher level
- What data is stored
- Describes in terms of small structures (e.g. tables)
- Used by DBA

Job of the administrator

3) View Level :
- Highest level
- Describes only part of DB
- Low Complexity

Job of the front end developer

Data Models

- A Data model is a description of the data in a database.
- It also provides relationship between data and any constraints.
- Data models can be divided into two categories

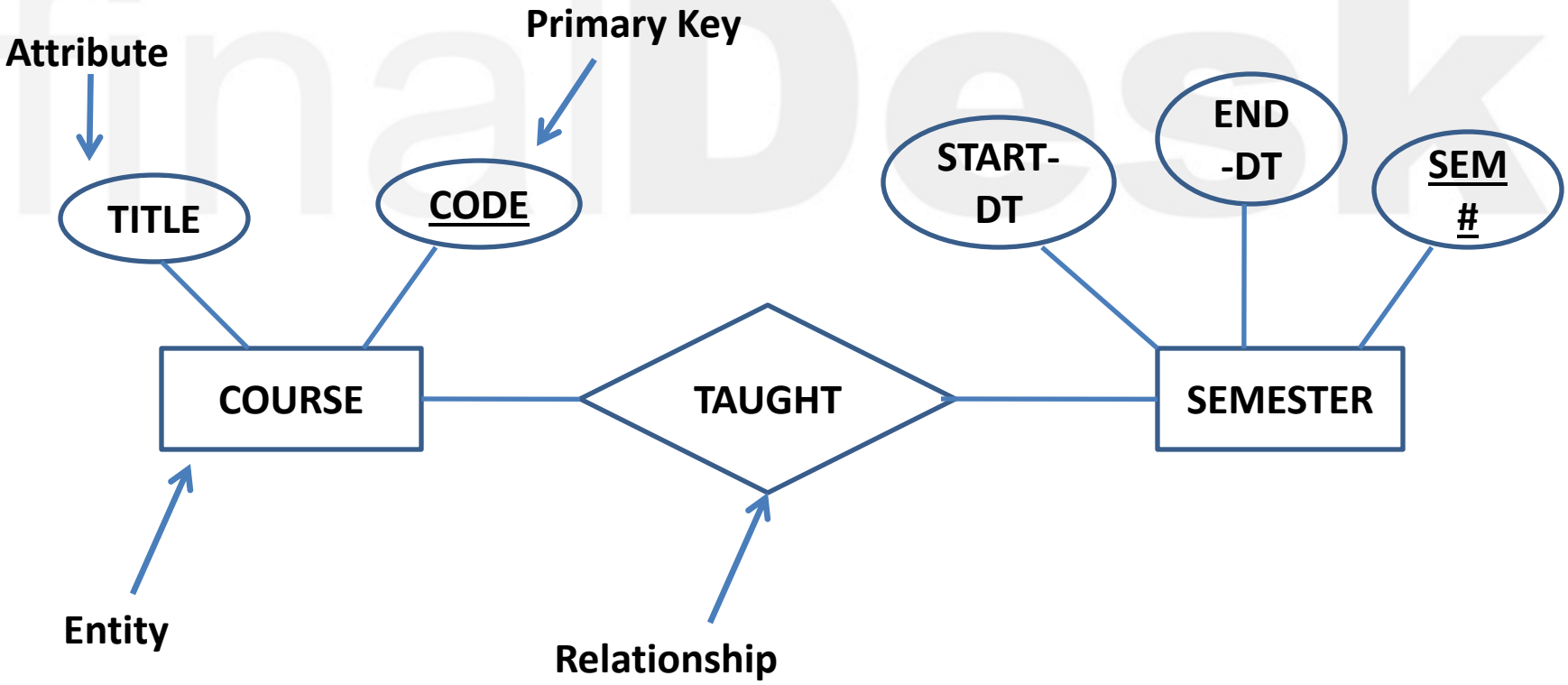
1) **Object based logical model:** Focuses on describing the data, the relationship among the data, and any constraints defined.
   E.g. – ER model.

2) **Record based logical model:** Focuses on specifying the logical structuring of the database.
   E.g. – Relational model.

The ER Model

• The Entity-Relationship (ER) model views the real world as a collection of objects or entities.

• Consider the following ER diagram.

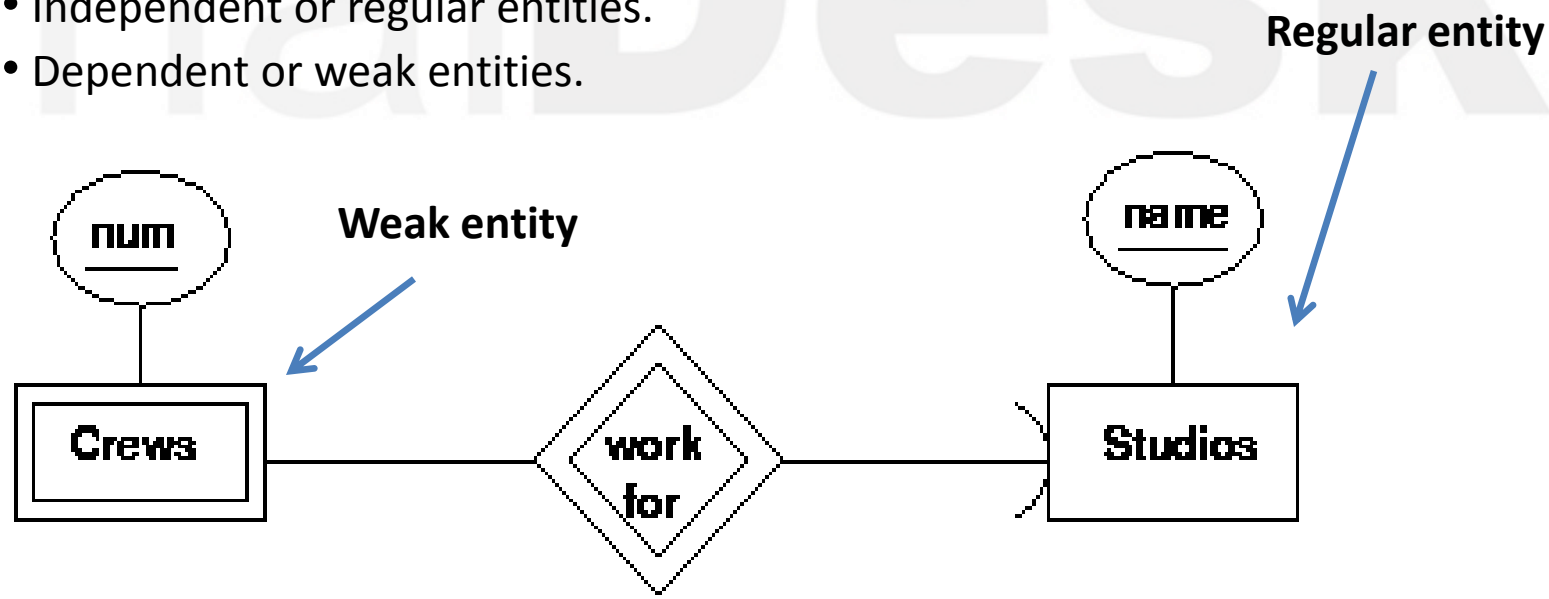


Entity:

• An entity is any object, place, person, or activity about which data is recorded.

• An **entity instance** is an individual occurrence of an entity type. For example, **DBMS** is an instance of the **COURSE** entity type.

• There are two types of entities
  • Independent or regular entities.
  • Dependent or weak entities.

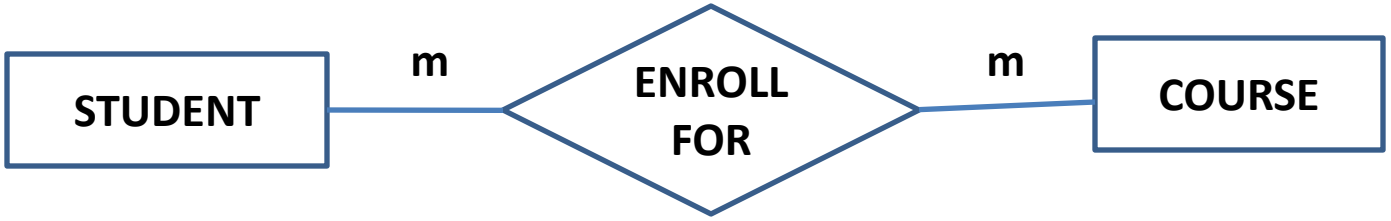


Relationship:

- An association among entities
- Types of relationships.
- **One-to-One relationship:**

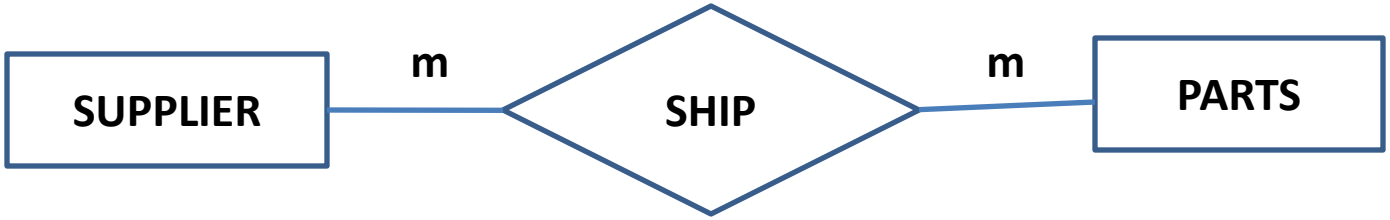
- **One-to-Many (Many-to-One) relationship:**

- **Many-to-Many relationship:**



Q) A supplier supplies certain parts. A particular part is not necessarily supplied by only one supplier. No supplier supplies only a single part. What type of relationship is this ? Draw a diagram to depict the relationship.

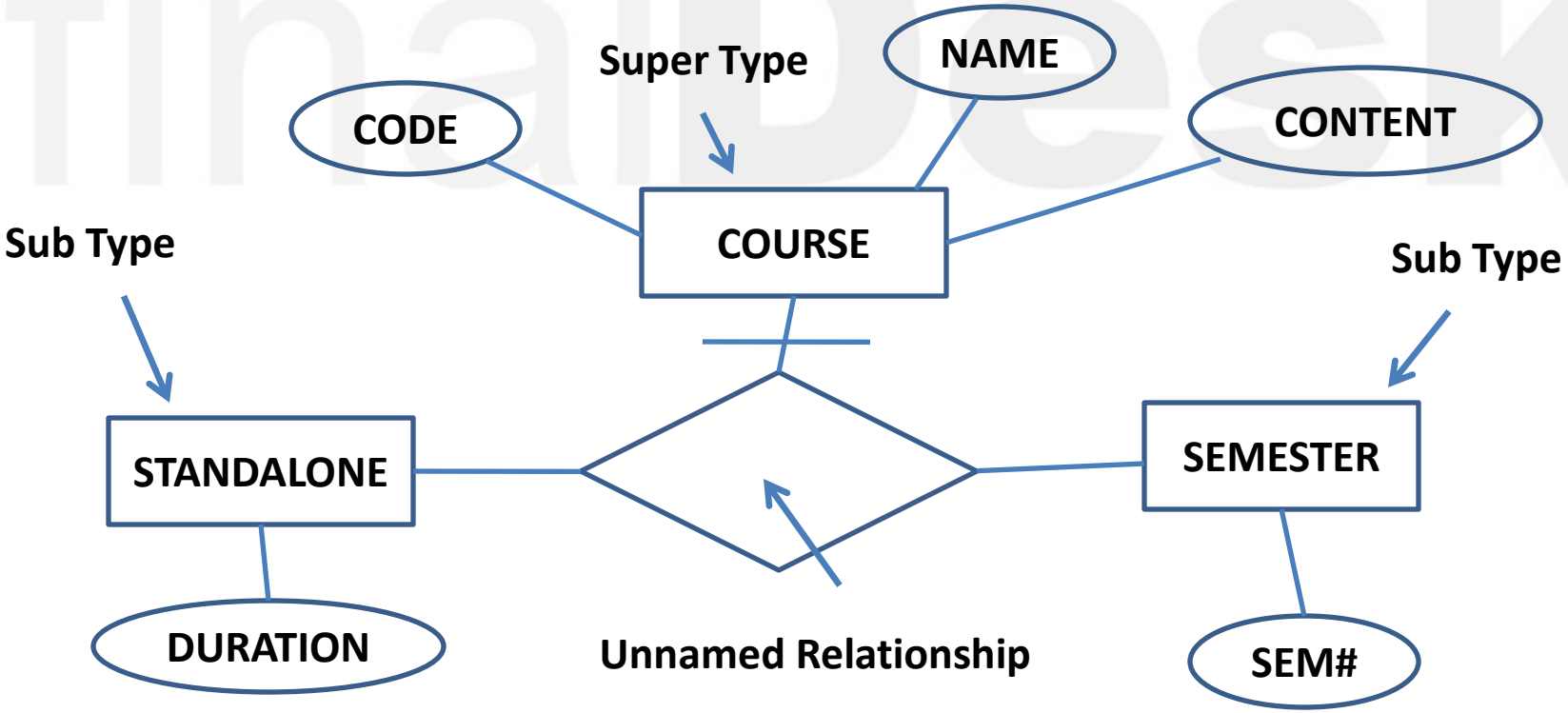
Solution:



Attributes:

- An attribute is a property of a given entity.
- The key properties are underlined.
- A key property or a primary key uniquely identifies an instance of an entity.

Subtypes and Supertypes:

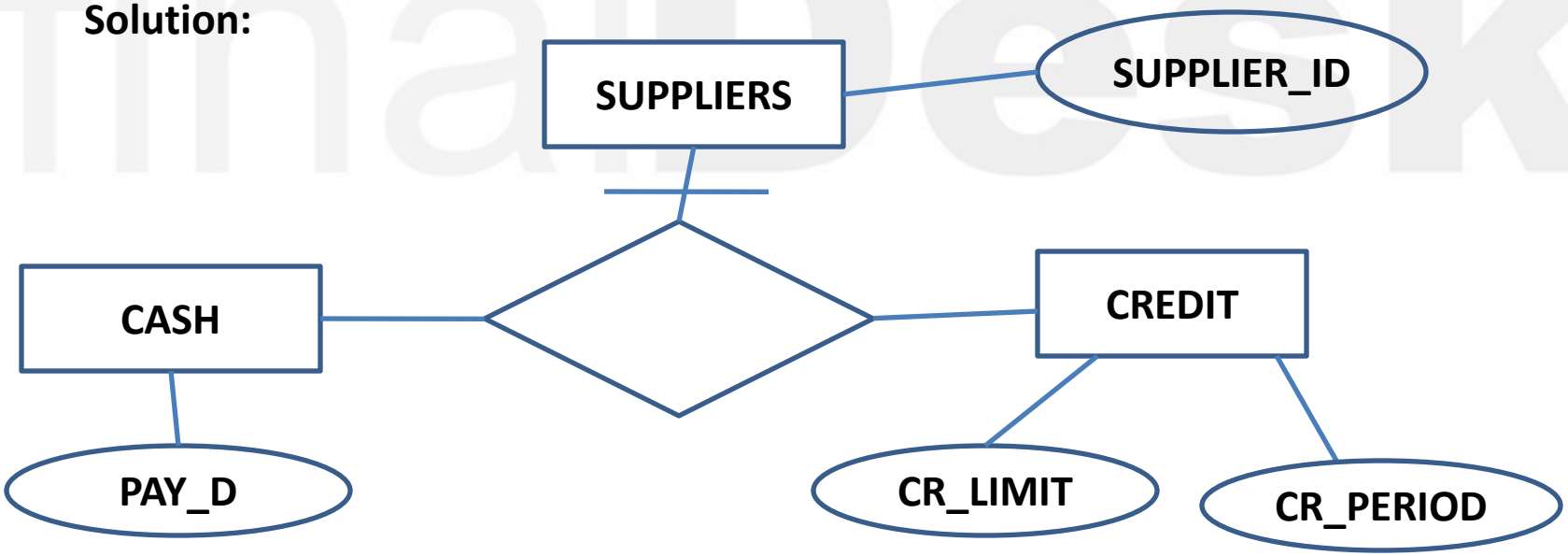


Steps to create an ER Diagram

1) Identify the entities.
2) Find the relationships.
3) Identify the key attributes.
4) Identify the other important attributes.
5) Draw a complete ER diagram.
6) Review of ER diagram.

Q) There are two types of suppliers. One type of supplier allows credit, while the other type insists on payment in cash before delivery. The manufacturer wishes to maintain separate information on these two types of suppliers. For the credit supplier, “credit period” and “credit limit” have to be recorded. For the supplier, “date of payment” has to be stored. Draw a complete ER diagram to represent this information.

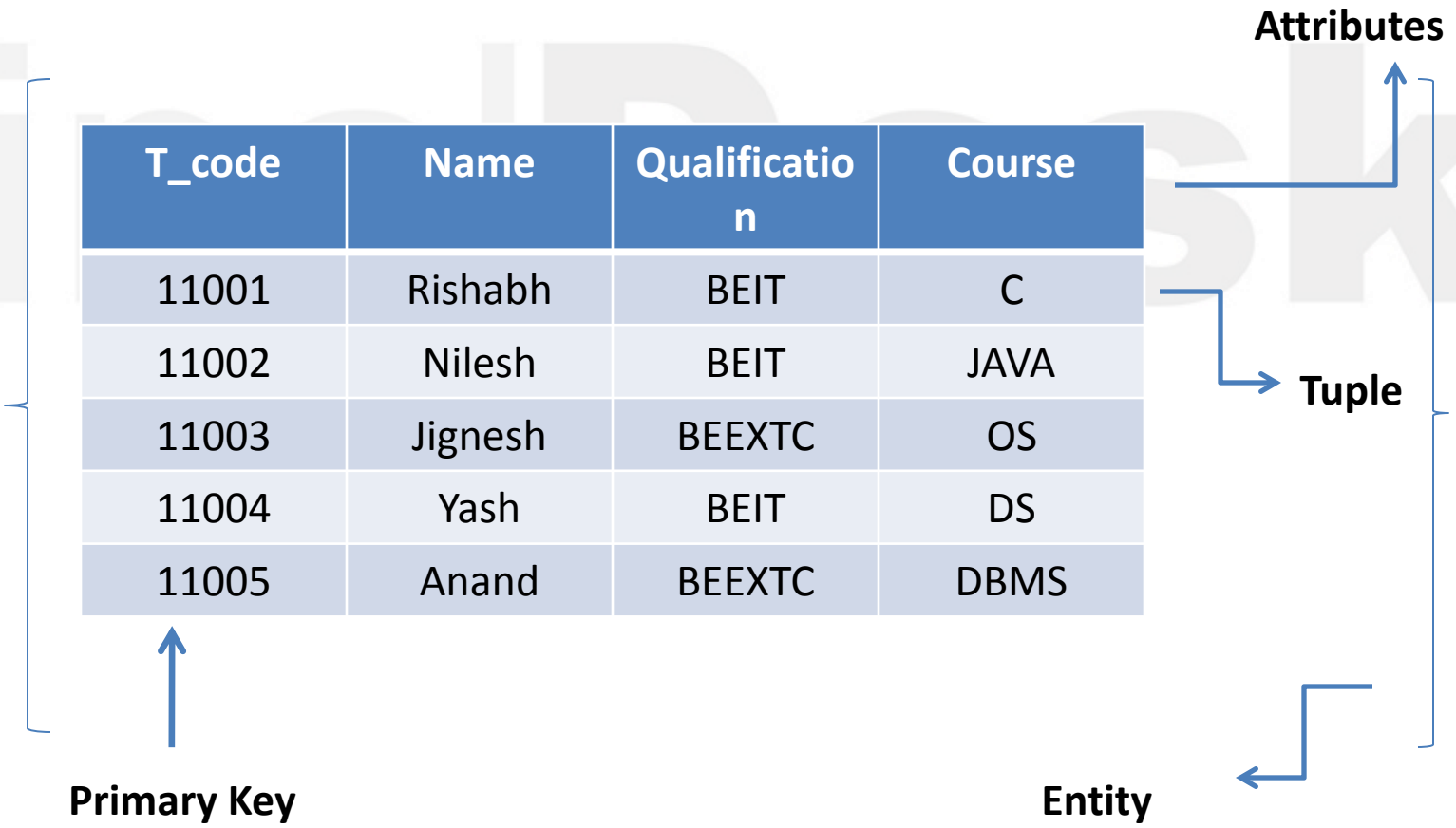
Solution:



Relational Model

• A Relational Model represents all data in the database as simple tables in the row-column format.

• The following table describes the details of trainers.



Important points

- The row (or record) in the table is called a **tuple**.
- The column (or field) in the table is called an **attribute**.
- Number of tuples is called the **cardinality** of the table.
- Number of attributes is called the **degree** of the table.
- A table where every row is different from all other rows is called a **relation**. The term relational database comes from this term.

Representing Relationships in a RDBMS

• Relationships in an RDBMS are represented by common data values stored in two or more tables.

• A column in one table whose value matches the primary key in some other table is called a **foreign key**.

Table 1

Roll no	Name
1	Banti
2	Bablu

Table 2

Course ID	Course name	Roll no
CS01	JAVA	2
CP23	DBMS	1

Foreign Key

Summary

- Data models can be classified as a) Object based logical model b) Record based logical model.
- Entities are represented as rectangles.
- Relationships are represented as diamonds.
- Attributes are represented as ellipses.
- Relationships, whether many-to-many, one-to-one or one-to-many are represented symbolically.
- Weak entities are represented by double-lined boxes.
- In the relational model, data is represented in tables (relations) of rows (tuples) and columns (attributes).
- The number of tuples is called the cardinality of the table, and the number of attributes is called the degree of the table.
- An attribute (or a set of attributes) that is unique in every tuple is called the primary key.
- The Foreign key is a column in one table that matches the primary key of another table.

Contact Info

- trainers@finaldesk.com
- rishabh@finaldesk.com
- nilesh@finaldesk.com
- jignesh@finaldesk.com
- yash@finaldesk.com
- anand@finaldesk.com