Se dice que uno o más atributos de una entidad es un identificador o clave primaria si el valor de dichos atributos determina de forma unívoca cada uno de los elementos de dicha entidad, y no existe ningún subconjunto de él que permita identificar a la entidad de manera única.

En una base de datos se almacenan además de las entidades, las relaciones existentes entre ellas. Estas relaciones se almacenan con punteros que inserta automáticamente el software que la maneja y esto es "transparente" al usuario.

**Niveles de abstracción de la información: vistas y esquemas**

* Nivel de vista: Permite describir diferentes vistas o subesquemas, cada una de las cuales se corresponde con la parte de la base de datos que interesa a un determinado grupo de usuarios. Además, limita el acceso solo a la información de la vista.
* Nivel conceptual: Describe el esquema de la base de datos. En éste se especifica qué información se guarda en la base de datos, incluyendo todos los datos almacenados en ella y las relaciones entre ellos. Este nivel se utiliza en la administración de la base de datos.
* Nivel físico: Describe cómo se almacenan los datos, con las estructuras de datos necesarias para ello.

**Modelos de datos**

Modelo de datos: grupo de herramientas conceptuales que permite describir los datos, sus relaciones, su semántica y sus limitaciones. Ayuda a describir la estructura de una base de datos. Clasificación de los modelos de datos:

* Modelos lógicos basados en objetos: Describen los datos a nivel conceptual y a nivel de vista, permiten una estructuración flexible y especificar limitaciones de los datos. (Modelo entidad-relación)
* Modelos lógicos basados en registros: Describen los datos a nivel conceptual y a nivel de vista, permitiendo especificar la estructura lógica pero no las limitaciones de los datos. (Modelo jerárquico, Modelo en red y Modelo relacional)
* Modelos físicos de los datos: Describen los datos en el nivel de implementación de los sistemas de base de datos.

**Modelo entidad-relación**

* Entidad: objeto que tiene existencia propia, que puede distinguirse de otros y del cuál se quiere almacenar información de ciertas características. Ejemplo: Pepe con DNI 24324450 y que vive en Granada.
* Conjunto de entidades: grupo de entidades del mismo tipo que representa la estructura genérica de una entidad de interés. Las entidades pueden pertenecer a más de un conjunto de entidades. Ejemplo: Cliente.
* Relación: asociación entre varias entidades. Ejemplo: El cliente con DNI 24324450 compra un coche con matrícula 6670 BBC el 7/11/2005.
* Conjunto de relaciones: grupo de relaciones del mismo tipo que representa la estructura genérica de las relaciones entre conjuntos de entidades. Ejemplo: Compra.
* Grado de una relación: número de tipos de entidad que intervienen en un tipo de relación. Ejemplo: En el caso de la relación en la que un cliente compra un coche, el grado de la relación es 2.
* Atributo: unidad básica de información sobre un tipo de entidad o un tipo de relación. Ejemplos: DNI, Nombre, Dirección, Fecha de compra.
* Cardinalidad: expresa el número de entidades con las que puede asociarse o corresponderse una determinada entidad mediante una relación:
  + Uno a Uno (1:1)
  + Uno a muchos (1:N)
  + Muchos a uno (N:1)
  + Muchos a muchos (N:M)

**Sistema de gestión de la base de datos - SGBD ("Data Base Management System" - DBMS):** Conjunto de software destinado a la creación, control y manipulación de la información de una base de datos. Un SGBD debe permitir la realización de las siguientes tareas:

1. Definición del esquema de la base de datos y de los distintos subesquemas.
2. Acceso a los datos desde algún lenguaje de alto nivel.
3. Interrogación (o recuperación de la información) directa en modo conversacional.
4. Organización física de la base de datos y recuperación tras fallos del sistema.

Las tres primeras tareas se realizan mediante dos lenguajes específicos:

* Lenguaje de descripción de datos (DDL, "Data Description Language» ). Se usa para la descripción del esquema y de los subesquemas.
* Lenguaje de manipulación de datos (DML, "Data Manipulation Language»). Se utiliza para el acceso a la base de datos desde lenguajes de alto nivel o en modo conversacional.

El sistema de gestión de la base de datos actúa como intermediario entre los programas de aplicación y el sistema operativo, lo que permite que los programas sean independientes de la estructura física de los datos