



**Universidad Nacional de Ingeniería
Facultad de Ciencias**

Bases de Datos

Introducción a las Base de Datos

Víctor Melchor Espinoza

Dato e Información

Si bien es cierto que son términos interconectados y forman parte de un mismo contexto, poseen significados totalmente distintos.

Básicamente, un dato es un conjunto de símbolos (alfanumérico o imagen), que no está agregado a ningún conocimiento específico, lo cual lo convierte en inutilizable para quien no sabe en que contexto está contenido y lo que representa.

Ejemplo

937 430

Dato e Información

Una información es agregar un determinado conocimiento a un dato.

Una información puede ser interpretada, mientras que un dato solamente puede ser visualizado.

Ejemplo

Población universitaria peruana en el año 2010

937 430

Dato e Información

Jerárquicamente, el dato sirve como materia prima para el alcance de una información, pues es en base a esta que puede ser hecha cualquier interpretación sobre su valor, haciéndose objetivo el conocimiento.

Dato  **Información**  **Conocimiento**

Ejemplo

Pob. Univ. Per. en el año 1996: 390 596

Población universitaria peruana en el año 2010

937 430

se incrementó en 2,4 respecto al censo anterior(1996)

Algunos conceptos

- Una base de datos en su definición más sencilla, es una colección de archivos relacionados.
- De manera más técnica, una base de datos es un conjunto de varios tipos de datos organizados e interrelacionados.
- A manera de teoría de conjuntos, es un conjunto de datos de diferentes ámbitos, organizados sistemáticamente, es decir, siguen ciertas reglas.

Base de Datos

Definición

Una base de datos es un conjunto de datos que describe las actividades de una o varias organizaciones relacionadas.

La base de datos de una universidad puede contener información sobre:

- **Entidades:** como alumnos, profesores y asignaturas.
- **Relaciones** entre esas entidades como la matrícula de los alumnos en las diversas asignaturas, los profesores que imparten cada asignatura

Base de Datos - Concepto

Base de datos: Es una colección de datos interrelacionados, que representan información sobre un dominio específico.

Son conjuntos de registros dispuestos en estructura regular que permite la reorganización de los mismos y la producción de información. Una base de datos normalmente agrupa registros utilizables para un mismo fin.

Una base de datos es usualmente mantenida y accesada a través de un software conocido como **Sistema Gestor de Base de Datos** (SGBD). Normalmente un SGBD adopta un modelo de datos, de forma pura, reducida o extendida. Muchas veces el término base de datos es usado como sinónimo de SGBD.

Ejemplo

Oracle, Paradox, Access, DBase, Interbase , PostgreSQL ,MySQL.

Definición de Manejador de Base de Datos

Un **manejador de base de datos** es conocido con las siglas **SGBD** – Sistema Gestor de Base de Datos o en inglés **DBMS** – Database Manager System.

DBMS es un software que actúa como interfaz entre los datos almacenados en forma binaria en una base de datos y el usuario que desea manejar tales datos.

Sistema Gestor de Base de Datos

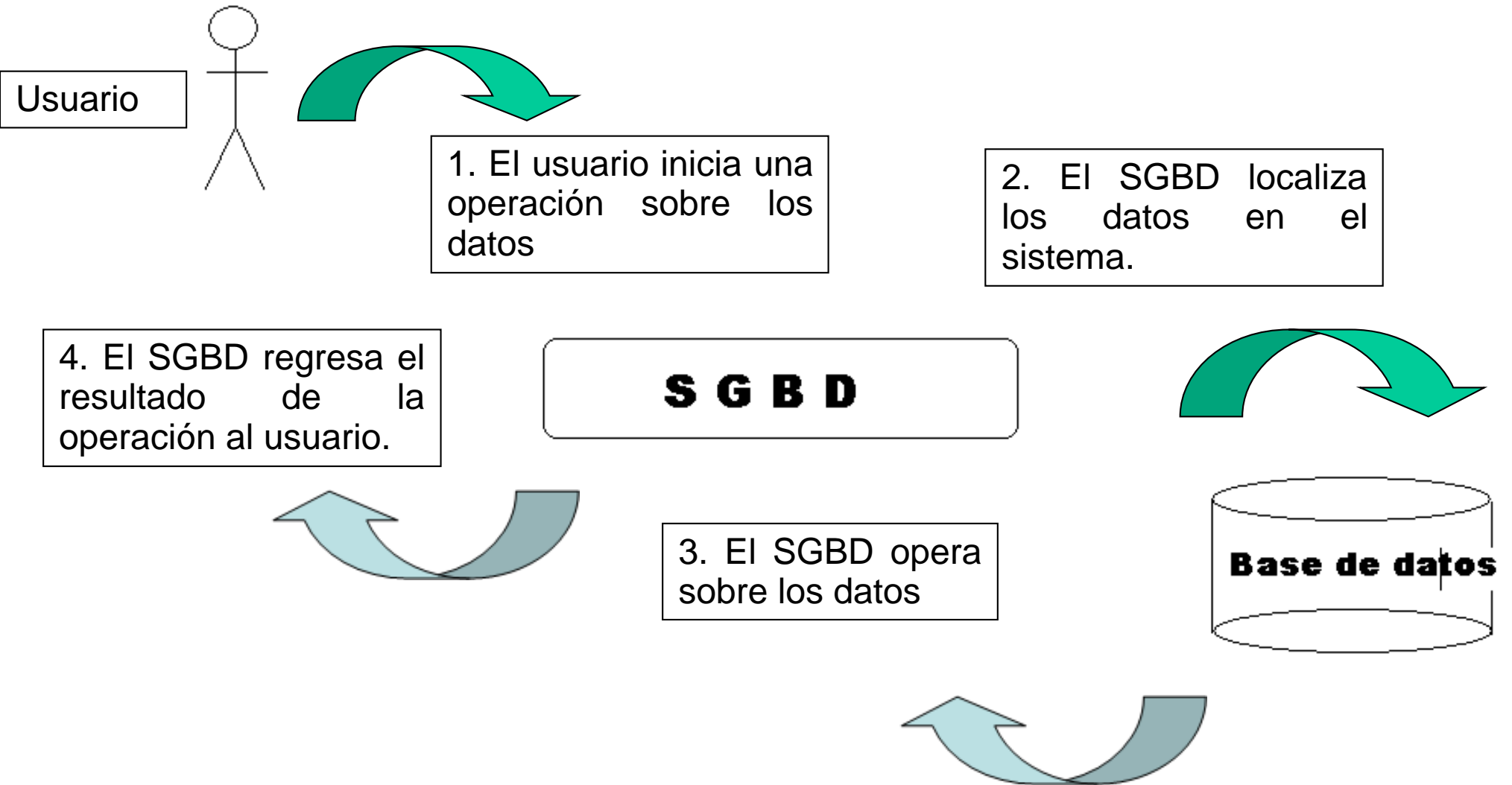
SGBD

Un sistema gestor de base de datos (DBMS) consiste de un conjunto de datos interrelacionados y en un conjunto de programas para tener acceso a esos datos. Los datos describen una empresa concreta.

Ejemplo



Esquema de Manejador de bases de datos.



Manejador bases de datos

Un **sistema manejador de base de datos relacionales** es conocido con las siglas **SGBDR** es el tipo de SGBD dominante hoy en día.

Aplicación de las bases de datos

Las bases de datos son ampliamente usadas.

¿Podría mencionar algunos sectores donde se usan las bases de datos?

Aplicación de las bases de datos

Las siguientes son algunas de sus aplicaciones más representativas:

Banca

Para información de los clientes, cuentas y préstamos así como para transacciones bancarias.

Líneas aéreas

Para reservas e información de planificación. Las líneas aéreas fueron de las primeras en usar las bases de datos de forma distribuida geográficamente (los terminales situados en todo el mundo accedían al sistema de bases de datos centralizado a través de las líneas telefónicas y otras redes de datos).

Aplicación de las bases de datos

Universidades

Para información de los estudiantes, matrículas de las asignaturas y cursos.

Transacciones de tarjetas de crédito

Para compras con tarjeta de crédito y generación mensual de extractos.

Telecomunicaciones

Para guardar un registro de las llamadas realizadas, generación mensual de facturas, manteniendo el saldo de las tarjetas telefónicas de prepago y para almacenar información sobre las redes de comunicaciones.

Aplicación de las bases de datos

Finanzas

Para almacenar información sobre grandes empresas, ventas y compras de documentos formales financieros en la bolsa como acciones y bonos.

Ventas

Para información de clientes, productos y compras.

Producción

Para la gestión de la cadena de producción y para el seguimiento de la producción de elementos en las fábricas, inventarios de elementos en almacenes y pedidos de elementos.

Aplicación de las bases de datos

Recursos humanos

Para información sobre los empleados, salarios, impuestos y beneficios y para la generación de las nóminas.

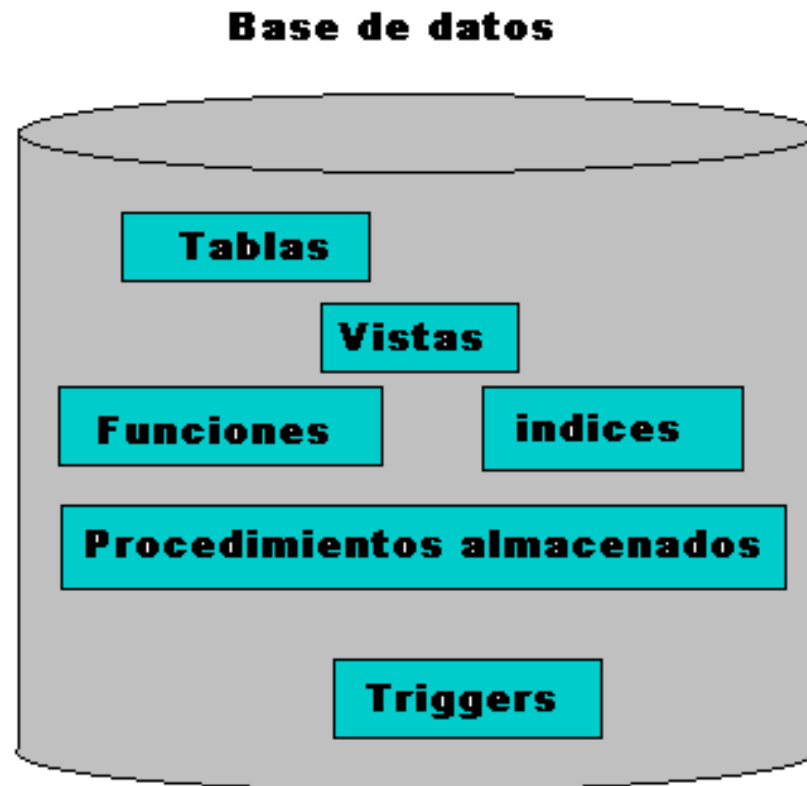
NOTA.

Como esta lista ilustra, las bases de datos forman una parte esencial de casi todas las empresas actuales.

Objetos de una base de datos

Las bases de datos están compuestas básicamente por 6 objetos:

Tablas, Vistas, Funciones, Índices, Procedimientos almacenados y Triggers.



Objetivos de un SGBD

Aislar a los usuarios de los detalles más internos de la base de datos (abstracción de datos).

Proveer independencia de datos a las aplicaciones (la estructura física de almacenamiento es la estrategia de acceso).

Evolución de los SGBD

El primer SGBD diseñado por Charles Bachman en General Electric en 1960 se denominó **Almacén Integrado de Datos – IDS**. Aportó el modelo de datos en red que se normalizó en la conferencia sobre lenguajes de Sistemas de Datos(CODASYL).

A fines de 1960, IBM desarrolló el SGBD **IMS, Sistema de Gestión de Información**, que constituyó la base para el modelo jerárquico de datos.

El sistema SABRE para la realización de reservas en las líneas aéreas fue desarrollado conjuntamente por American Airlines e IBM.

Evolución de los SGBD

En 1970, Edgar Codd, del laboratorio de investigación de San José de IBM propuso el **modelo relacional de datos** que fue un hito en el desarrollo de los sistemas de bases de datos.

Codd ganó el premio Turing en 1981 por su trabajo pionero.

Evolución de los SGBD

En 1980 el lenguaje de Consultas **SQL** para las bases de datos relacionales desarrollado como parte del proyecto System R de IBM se constituye en el lenguaje de consultas estándar.

A fines de 1980, SQL fue normalizado y la norma actual, SQL 1999 ha sido adoptada por ANSI e ISO.

La forma más utilizada de programación concurrente es la ejecución concurrente de programas de bases de datos (transacciones).

Los usuarios escriben los programas como si se fueran a ejecutar de manera independiente y el SGBD los ejecuta de forma concurrente.

Evolución de los SGBD

En 1980 y 1990 se llevó a cabo una gran cantidad de investigación en lenguajes de consulta más potentes y en modelos de datos más ricos.

Un fenómeno interesante es la aparición de varios paquetes de ERP y MRP.

Entre los paquetes más usados hay sistemas de Oracle, People Soft, SAP y Siebel.

Estos identifican un conjunto de tareas frecuentes en las diversas organizaciones y ofrecen una capa general de aplicación para desempeñarlas, la cual puede ser personalizada.

Evolución de los SGBD

Los SGBD han entrado en la era de Internet. Las consultas se generan mediante formularios accesibles por Web y se le da formato a las respuestas mediante un lenguaje de marcas como HTML.

Entre las tendencias futuras tenemos:

- Bibliotecas Digitales.
- Bases de datos multimedia.
- Elaboración del mapa del genoma humano.
- Proyecto de Sistema de Observación de la Tierra de la NASA.

Ventajas de un SGBD

- Rapidez en la manipulación y en el acceso a la información;
- Reducción del esfuerzo humano (desarrollo y utilización);
- Disponibilidad de la información en el tiempo necesario;
- Control integrado de informaciones distribuídas físicamente;
- Reducción de redundancia y de inconsistencia de informaciones;
- Compartición de datos;
- Aplicación automática de restricciones de seguridad;
- Reducción de problemas de integridad.

Desventajas de un SGBD

Sin dispositivos de control adecuados, la seguridad puede quedar comprometida; por ejemplo, en el caso de acceso no autorizado a datos.

La integridad de las informaciones puede quedar comprometida si no hubiera mecanismos de control; por ejemplo en el caso de manipulación concurrente de datos.

La operación del sistema de base de datos y el desarrollo de aplicaciones necesitan ser hechos con mucha precisión para evitar que informaciones no correspondan a la realidad.

La administración del sistema de base de datos puede volverse muy compleja en ambientes distribuidos, con gran volumen de información manipulada por una gran cantidad de usuarios.

SQL. Lenguaje de base de datos

SQL. Es el lenguaje estándar para el manejo de base de datos. ***SQL - Structured Query Language.***

SQL se divide en dos Sub-lenguajes

a) DDL (Data Definition Language).

b) DML (Data Manipulation Language).

Lenguajes de Bases de Datos

Hay dos lenguajes de comandos que nos permiten interactuar con un SGBD:

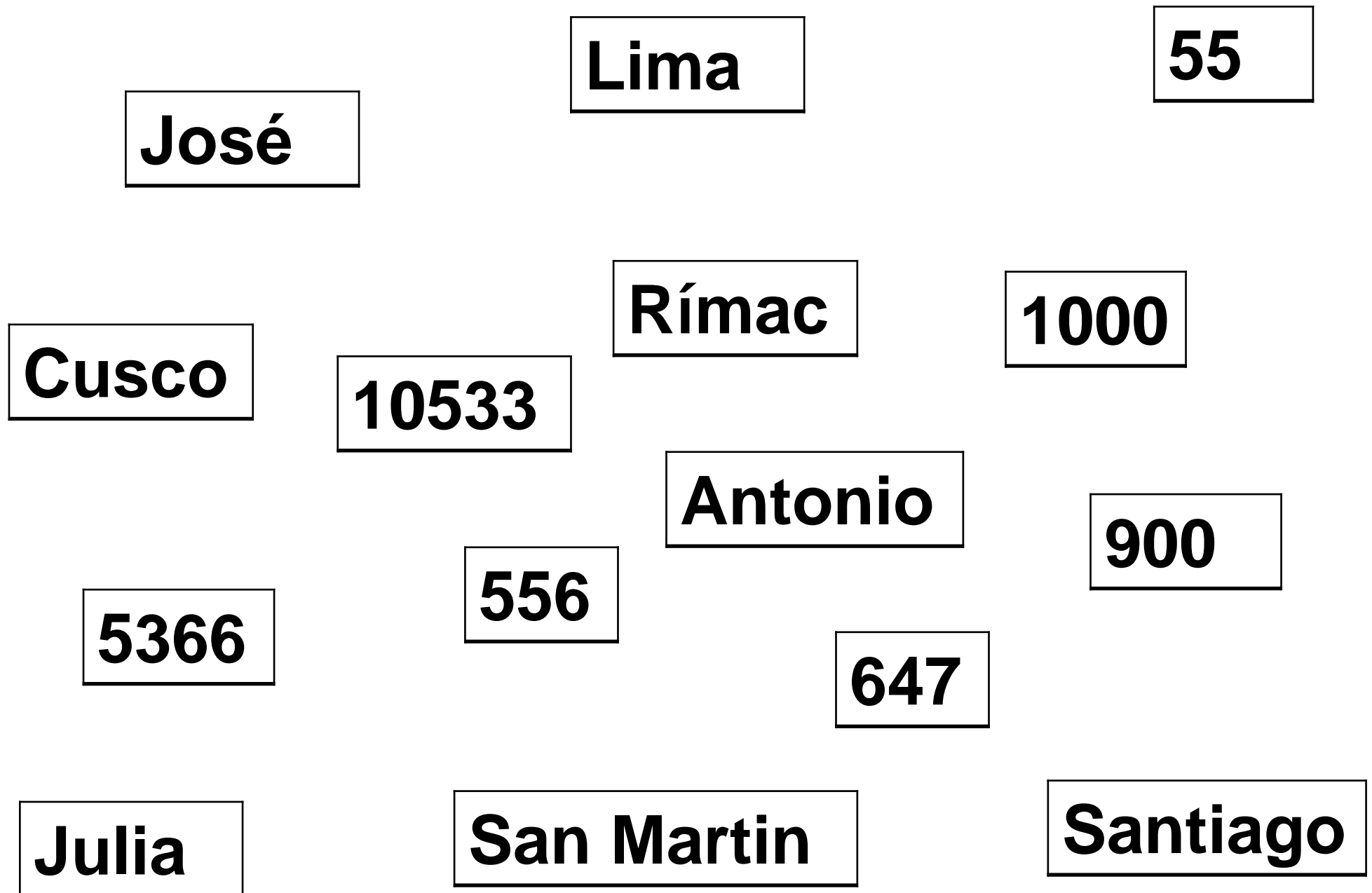
DDL

Se usa para definir los datos. Es un lenguaje para la definición del esquema de la base de datos y otras propiedades de los datos. Agrupa a las operaciones **CREATE DATABASE**, **CREATE TABLE**, **CREATE INDEX**, etc.

DML

Se usa para preguntar por los datos. Es un lenguaje que permite a los usuarios tener acceso a los datos o manipularlos. agrupa a las operaciones **SELECT**, **INSERT**, **UPDATE** y **DELETE**.

Modelo de Datos



Modelo de Datos

Nombre	Distrito	Ciudad	Cuenta	Saldo
José	San Martin	Lima	900	55
Julia	Rímac	Lima	556	1000
Julia	Rímac	Lima	647	5366
Antonio	Santiago	Cusco	647	5366
Antonio	Santiago	Cusco	801	10533

Modelo Jerárquico

Usado en los primeros sistemas de gestión de base de datos mainframe.

José	San Martin	Lima
-------------	-------------------	-------------

Antonio	Santiago	Cusco
----------------	-----------------	--------------

900	55
------------	-----------

647	5366
------------	-------------

801	10533
------------	--------------

Julia	Rímac	Lima
--------------	--------------	-------------

556	1000
------------	-------------

647	5366
------------	-------------

Modelo Jerárquico

Características

La información estaba almacenada en ficheros por orden secuencial.

Este modelo establece mediante punteros, una relación entre los distintos registros de los ficheros.



Problema

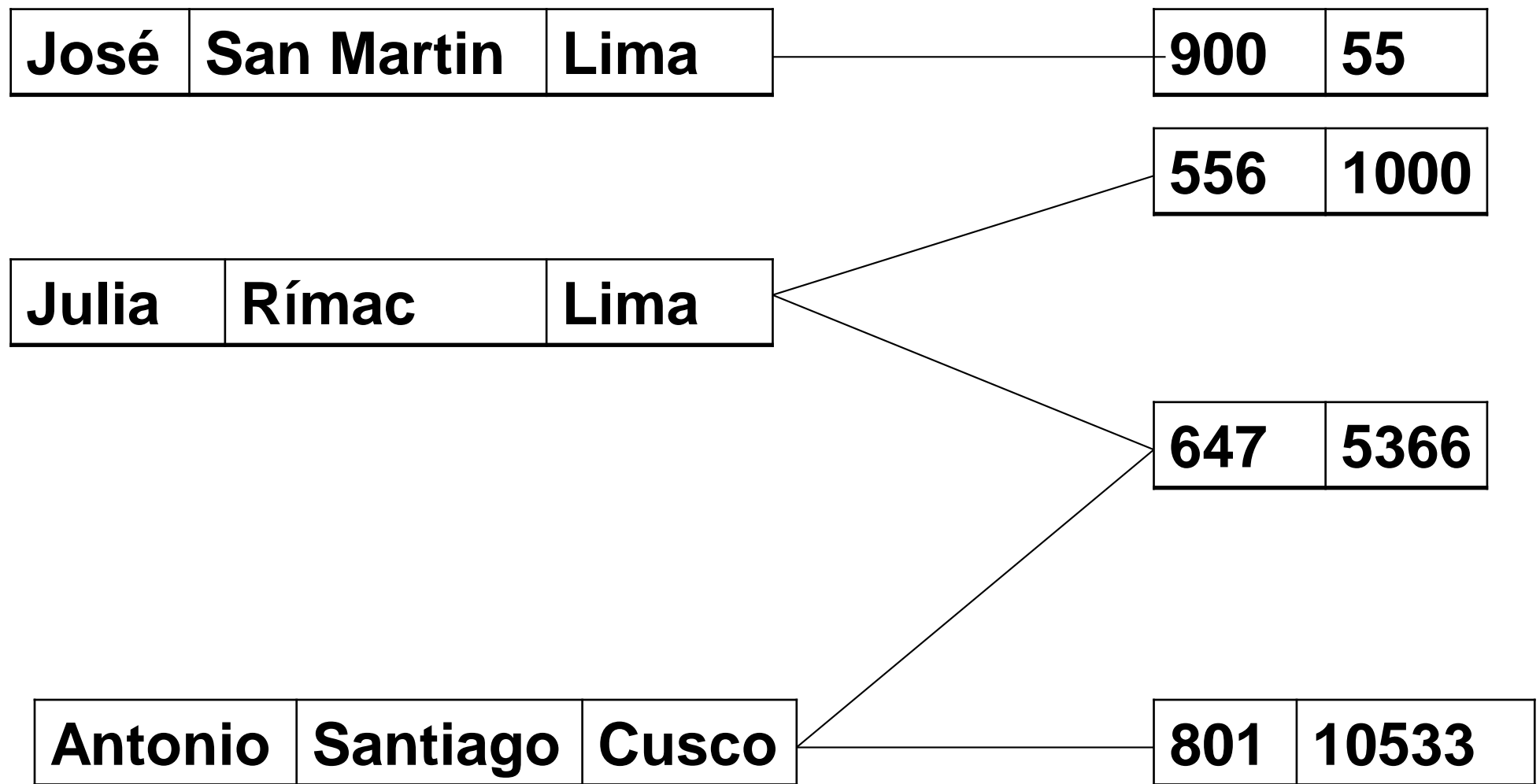
Pesadez de las reorganizaciones de los ficheros. Por ejemplo: una nueva factura.



Modelo de Red

Semejante al modelo jerárquico, con la diferencia de que cada registro padre puede ser asociado a más de un registro hijo.

Los datos se representan por colecciones de registros y sus relaciones.



Modelo de Red

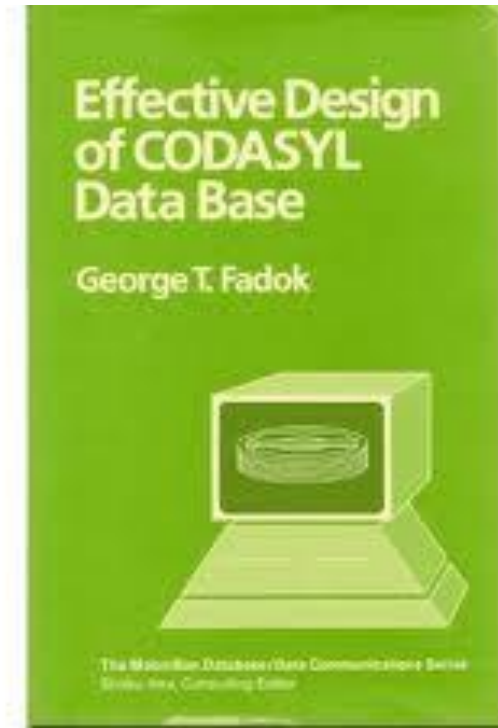
Características

Los registros están enlazados entre sí.

Se evitan reorganizaciones de ficheros.

Problema

Es mas lento que recorrer los ficheros secuencialmente



Modelo Relacional

Concepto creado por Edgar Frank Codd en 1970 basado en la lógica de los predicados y en la teoría de los conjuntos.

Codigo	Nombre	Distrito	Ciudad
015	José	San Martin	Lima
021	Julia	Rímac	Lima
037	Antonio	Santiago	Cusco

Cuenta	Saldo
900	55
556	1000
647	5366
647	5366
801	10533

Cliente	Cuenta
015	900
021	556
021	647
037	647
037	801

¿Por qué usar una Base de Datos Relacional?

Fueron desarrollados para proveer un fácil acceso a los datos, permitiendo que los usuarios utilicen una gran variedad de enfoques en el tratamiento de las informaciones.

Mientras que en una base de datos jerárquica los usuarios necesitan definir las cuestiones de negocios de manera específica, iniciando por la raíz del mismo, en las Bases de Datos Relacionales los usuarios pueden hacer preguntas relacionadas a los negocios a través de varios puntos.

El lenguaje estándar de las Bases de Datos Relacionales es el **Structured Query Language**, o simplemente **SQL**, como es más conocido.