

INTRODUCCIÓ A LES BASES DE DADES



Objectius



- Diferencia entre dades i informació
- Definir que és una base de dades
- Identifica els diferents elements, objectes i estructures d'emmagatzematge físic disponibles en un SGBD corporatiu • Identifica l'estructura d'un diccionari de dades.

Introducció

- La feina d'informàtic és **gestionar dades**. Quan aquestes dades son molt nombroses necessitem fer programes informàtics que ens ajudin a automatitzar el treball sobre aquestes dades.
- En aquest curs veurem la gestió de dades des d'un **punt de vist acadèmic**. Les necessitats de cada projecte podran fer que alguns de les coses que veiem s'hagin d'adaptar.
- **El BigData y els analistes de dades** es una de les **professions de futur**. La base de com ser-ho, ho treballarem en aquest curs.

Què és una dada?

- Una **dada** és la representació ,a través d' una seqüència de símbols, nombres o lletres, d'una variable que pot ser **quantitativa o qualitativa** i indica un valor que se li assigna a les coses.
- Les dades descriuen **fets empírics**. Per examinar-los han de ser organitzats o tabulats, ja que una dada per si mateixa no pot demostrar massa sinó que s' ha d' avaluar el conjunt per examinar els resultats.
- Per exemple:

40 **Edat**

IMPORTANT



Exemples de dades

- **Exemples de dades qualitatives:**

Nom: Francesc Barragan

Ocupació: Docent

Domicili: Avinguda Barcelona

Gènere: Masculí

- **Exemples de dades quantitatives:**

Edat: 47

Alçada: 1.84

Nombre de fills: 1

Anys en el càrrec: 2



Diferència entre dada i informació

- El concepte de dada està estretament lligada al d'informació, però hi ha una diferència fonamental entre ambdós termes. Mentre **les dades refereixen a esdeveniments o fets registrats**, la **informació** està constituïda per aquelles dades en brut que són processades de manera que **generin contingut que pugui ser interpretat pels usuaris**.
- **Les dades no tenen sentit per si mateixes**, però en ser processades i contextualitzades es converteixen **en informació** certa i disponible per conèixer un fenomen, prendre decisions o executar accions.



Pregunta: Què ens interessa tenir dades o informació?

Entitats, atributs i valors

Elements que caracteritzen les informacions:

1. **Les entitats** són els objectes del món real que conceptualitzem. Són identificables, és a dir, distingibles els uns dels altres. I ens interessen algunes (com a mínim una) de les seves propietats.
2. **Els atributs** són les propietats de les entitats que ens interessen.
3. **Els valors** són els continguts concrets dels atributs, les determinacions concretes que assoleixen. Potser que no tingui un contingut concret i és llavors quan parlem de valors **NUL**.

Exercici

- **Exemples de entitat, atributs i valors**

Considerarem que **una pel·lícula concreta és una entitat**, perquè és un objecte del món real, que hem conceptualitzat dins d'una categoria (la dels films cinematogràfics), i que al mateix temps és distingible d'altres entitats de la mateixa categoria (és a dir, d'altres films).

D'aquesta pel·lícula ens interessaran alguns aspectes, que anomenarem **atributs**, com per exemple, **el títol, el director i l'any de producció**.

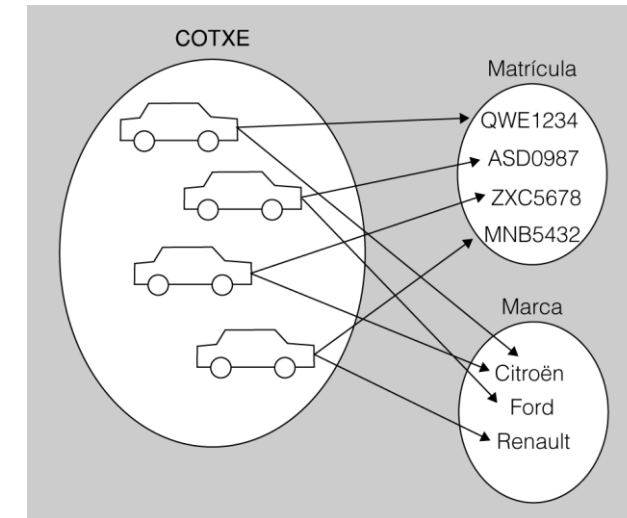
Finalment, aquests atributs adoptaran uns **valors concrets** com ara, i respectivament, Minions El Origen de Gru, Kyle Balda,.. i 2022.



**Podeu
posar més
exemples?**

Representació de les dades: tabular

- Ja sabem que les dades són **informacions representades informàticament**. Per tant, també podríem anomenar *món de les dades* el món de les representacions.
- La representació informàtica més freqüent en l'àmbit de les BD és la **representació tabular**, la qual s'implementa habitualment en fitxers que s'estructuren en **registres i camps**.



Cotxes	
Matrícula	Marca
QWE1234	Citroën
ASD0987	Ford
ZXC5678	Citroën
MNB5432	Renault

Què és una base de dades?

- Conjunt estructurat de dades que representa entitats (fets o ocurrències del món real) i les seves interrelacions (comparteixen algun vincle entre elles).
- No necessàriament el suport ha de tenir suport informàtic (**manuals** o **automatitzats**).

IMPORTANT

- Exemples:

Guia telefònica
Biblioteca



Concepte i definició de base de dades

- Actualment, la majoria de base de dades estan en suport digital i són utilitzades per gran part d'aplicacions informàtiques.

Examples:

Aplicacions que utilitzen BD:

WhatsApp

Facebook

Amazon

Cercador de Google

Empreses de tot tipus:

Instituts

Universitats

Gestors de continguts: Wordpress, Joomla, Drupal, Prestashop



Característiques d'un BD

Les principals característiques d'una base de dades son:

Persistència



Les dades han de durar en el temps
Només s'han de guardar les dades importants

Relacions



Entitats: conjunt de dades d'un tema
Relacions: connexions que existeixen entre les entitats

Compartir



Múltiples usos: diferents maneres d'entrar, mantenir i visualitzar la informació
Múltiples usuaris: moltes persones la puguin utilitzar al mateix temps

Què és un SGBD?

- Per poder treballar amb la informació continguda en una base de dades (crear la seva estructura, afegir, modificar, eliminar dades...consultar dades) hi ha aplicacions especialitzades en aquesta tasca anomenats Sistemes **Gestors de Base de Dades (SGBD)** o **Database Management System (DBMS)** en anglès.
- Més formalment: És una col·lecció de dades interrelacionades i un conjunt de programes per gestionar-les.



Objectius dels SGBD

- ❖ Proporcionar una forma d'emmagatzemar i recuperar la informació d'una Base de Dades de manera pràctica i eficient.
- ❖ Oferir als usuaris una visió abstracte de les dades. El sistema oculta certs detalls en la forma en que s'emmagatzemen i es mantenen les dades.
- ❖ Facilitar als desenvolupadors la creació i manipulació d'informació.

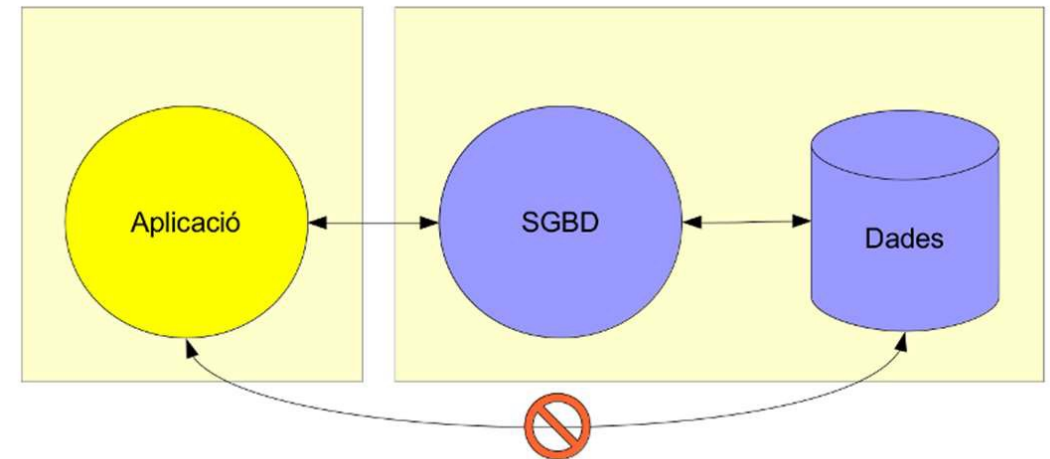


Característiques del SGBD

- Pot ser una aplicació o unes llibreries. En cas de ser aplicació caldrà disposar d'un sistema de comunicació de processos per tal de poder-nos connectar (per exemple model client-servidor).
- Existeixen diferents sistemes de comunicació:
 - per pipes
 - per xarxa
 - per memòria compartida.
- En cas de ser llibreries caldrà enllaçar-les al nostre programa, ja sigui de manera estàtica en temps de compilació del programa o bé de manera dinàmica durant l'execució del mateix.

Característiques del SGBD

- Les aplicacions no accedeixen directament a les dades, li demanen al SGBD (sistema gestor de base de dades) les operacions que volen realitzar.
- Exemple d'operacions poden ser modificació de dades o consulta de dades existents. El SGBD és qui realitza aquestes operacions mantenint la integritat de les dades i aplicant la seguretat que estigui definida al sistema.



Avantatges d'utilitzar un SGBD

- 1 Realitzar consultes no predefinides i complexes**
Proporcionen mecanismes per manipular dades (alta, baixa, modificació i consulta) i definir-ne de noves.
- 2 La independència física de les dades**
Un canvi sobre l'esquema físic no afecta ni als esquemes lògics ni als externs.
- 3 Minimitzar la redundància de dades**
Definim que volem guardar a nivell lògic i el SGBD ens ajuda a definir com s'acaba guardant a nivell físic. Tria la forma d'enregistrar les dades com el manteniment de les relacions, entre aquestes, sovint van orientades a minimitzar la redundància (característica no desitjable).
- 4 Vetllar per la integritat de les dades**
Podem definir regles semàntiques sobre determinades dades, el SGBD exigeix en tot moment el compliment d'aquestes per mantenir la integritat de les dades.
- 5 La independència lògica de les dades**
Un canvi sobre l'esquema lògic no afecta ni als esquemes físics ni als externs.
- 6 Accés concurrent**
Implementa diferents mecanismes per facilitar l'accés concurrent a les dades (múltiples usuaris treballant sobre la mateixa BD).
- 7 Seguretat**
Proporciona mecanismes per la definició d'usuaris i permisos sobre la BD. L'encarregat de gestionar aquests permisos és un usuari especial anomenat Administrador de la BD.
Encriptació

Quin SGBD hem d'escollir?

- El SGBD a escollir anirà en funció dels **requeriments del projecte** i del **suport** que vulguem tenir.
- Principalment tindrem:
 - **Bases de dades propietàries** (pagament per llicència Oracle, MSSQL)
 - **Base de dades lliures** (Postgresql, MariaDB)
- Qualsevol del dos models es podran presentar en modalitat Cloud o On Premise, i algunes com AWS o Snowflake només poden ser contractades en Cloud.



Quadrant de Gartner

- Un bon indicador de com està el mercat de les noves tecnologies és veure el famós **quadrant màgic** de la consultora americana **Gartner**, destinada a la consultoria i investigació en el mercat de les noves tecnologies.
- Agrupa el competidors per:
 - **Líders (leaders)**: Són qui tenen la millor puntuació resultant de combinar l'habilitat d'execució i l'abast de visió.
 - **Aspirants (challengers)**: Caracteritzats per oferir bones funcionalitats i un nombre considerable d'instal·lacions del producte, però sense la visió dels líders.
 - **Visionaris (visionaries)**: Tenen habilitat per anticipar-se a les necessitats del mercat, però en canvi no responen amb una plataforma sòlida per respondre a les necessitats a nivell global.
 - **Nínxols específics (niche players)**: Enfocats a determinades àrees de les tecnologies d'emmagatzematge de BD, però sense



<http://www.gb-advisors.com/es/cuadrante-de-gartner/>

Exemple Universitat (I)

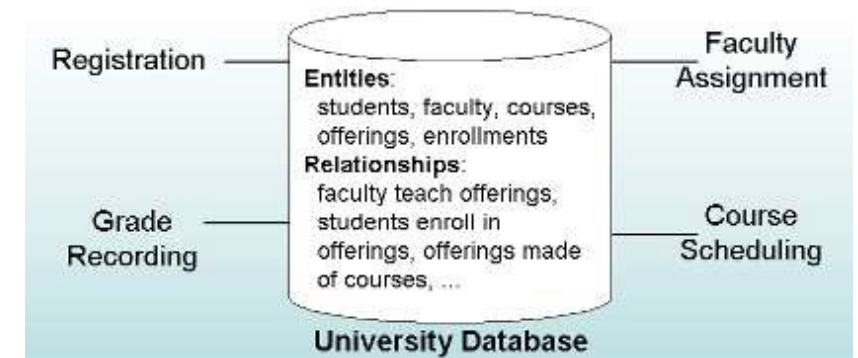


- Dins d'una Universitat volem dur el manteniment d'informació diversa (procés de matrícula, informes de les diferents qualificacions dels alumnes, programació/horari de les matèries i/o cursos, assignació de professors i alumnes als diferents cursos, etc...). Cada submón conté o comparteix diferents entitats (estudiants, facultats, cursos, professors, semestres) i relacions (els estudiants realitzen assignatures, els estudiants estudien en una facultat, la universitat té diferents facultats, cada facultat ofereix una sèrie d'assignatures,...).
- Per definir aquesta base de dades, hem d'especificar l'estructura de cada element que en volem guardar informació, indicant els diferents tipus d'elements d'informació que volem emmagatzemar en cada element.
- Per exemple: en el cas d'un estudiant volem guardar: El nom, el DNI, el N° d'estudiant, les carreres i les assignatures que està matriculat. La manipulació de la base de dades consisteix en l'actualització de les dades i en les consultes que en podem fer.

Exemple Universitat (II)



- Consultes:
 - Obtenir un llistat de tots els alumnes que estan matriculats en una assignatura concreta
 - Quins alumnes van obtenir més d'un 7 en l'assignatura de Base de dades l'any passat.
 - Quants professors diferents de Base de dades hi han hagut al llarg de tots els anys.
- Actualització:
 - Modificar l'adreça d'un alumne
 - Modificar el professor que imparteix una assignatura concreta.
 - Esborrar una assignatura perquè no hi ha cap alumne que s'hi hagi matriculat.
 - Assignar un alumne en una assignatura.



WEBGRAFIA

- Batini, C.; Ceri, S.; Navathe, S.B. (1992). Conceptual Database Design: An Entity-Relationship Approach. Reading, Massachusetts: Addison Wesley.
- Teorey, T.J. (1999). Database Modeling & Design. The Fundamental Principles (3a ed.). San Francisco: Morgan Kaufmann Publishers, Inc.