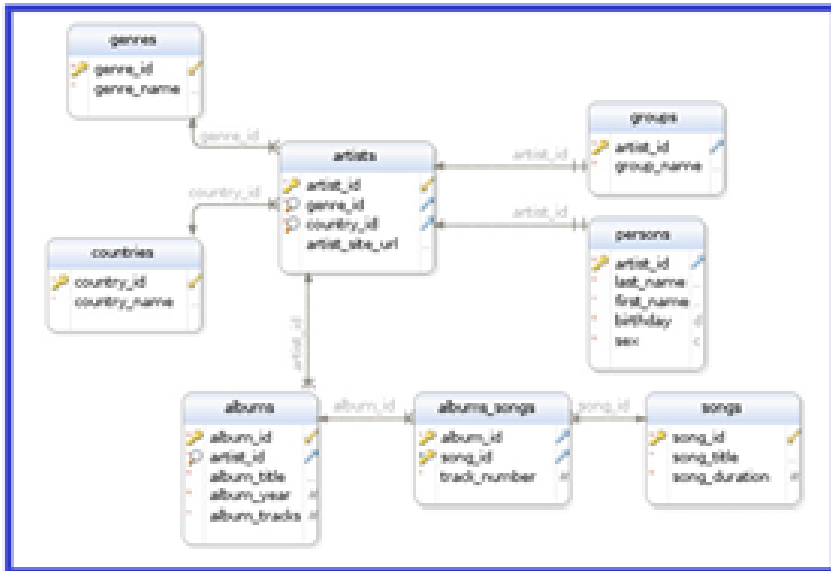
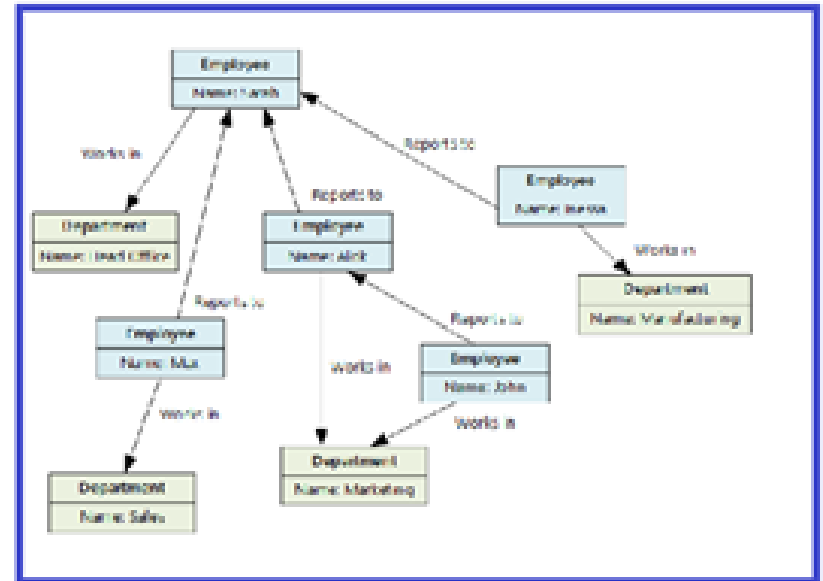


**Системи  
управління  
базами  
даних**



**SQL**

**VS**



**No SQL**

# Найпопулярніші бази даних

1. Oracle
2. MySQL
3. **Microsoft SQL Server**
4. PostgreSQL
5. **MongoDB**
6. Firebase
7. Splunk
8. Redis
9. Elasticsearch
10. SQLite
11. DB2
12. MariaDB
13. Apache Hive
14. SAP HANA

# Класифікація інформаційних систем

*Інформаційна система є сукупністю технічного, програмного та організаційного забезпечення, а також персоналу, призначена для того, щоб своєчасно забезпечувати людей необхідною інформацією .*



Основні компоненти інформаційної системи

## Класифікація інформаційних систем

Галузь застосування	Цільова функція	У міру <u>розподіленості</u>	За охопленням завдань (масштабності)
Виробництво Освіта Охорона здоров'я Наука Військова справа Соціальна сфера Торгівля інші	Керуючі Інформаційно-довідкові Підтримки прийняття рішень	Локальні ( <u>desktop</u> ) Розподілені ( <u>distributed</u> ) Файл-серверні Клієнт-серверні	Персональна інформаційна система Групова інформаційна система Корпоративна інформаційна система

**Керуюча інформаційна система** – цифрова система контролю чи управління деяким реальним об'єктом. Даними системами вирішуються завдання, не пов'язані з необхідністю ухвалення рішення в реальному часі (розрахунок, моделювання, офісні завдання). Дані ІВ вирішують завдання, які мають чітко виражену специфіку.

**Інформаційно-довідкова система** — система, призначена для зберігання, пошуку та обробки інформації, та відповідні організаційні ресурси (людські, технічні, фінансові тощо), які забезпечують та поширюють інформацію. Дані системи призначені для своєчасного забезпечення конкретних інформаційних потреб у межах певної предметної галузі, результатом функціонування інформаційних систем є інформаційна продукція — документи, інформаційні масиви, бази даних та інформаційні послуги.



**Система підтримки прийняття рішень ( Decision Support System , DSS) — комп'ютерна автоматизована система, метою якої є допомога людям, які приймають рішення у складних умовах для повного та об'єктивного аналізу предметної діяльності. Це означає, що вона видає інформацію (у друкованій формі, або на екрані монітора, або звуком), ґрунтуючись на вхідних даних, що допомагає людям швидко та точно оцінити ситуацію та ухвалити рішення. Дані системи виникли внаслідок злиття управлінських інформаційних систем та систем управління базами даних.**

У міру розподілу розрізняють:

***Локальні інформаційні системи ( desktop )*** у яких всі компоненти (БД, СУБД, клієнтські програми) працюють однією комп'ютері.

***Розподілені інформаційні системи ( distributed ),*** в яких компоненти розподілені по кількох комп'ютерах.

***У файл - серверних ІС*** база даних знаходиться на файловому сервері, а СУБД та клієнтські програми знаходяться на робочих станціях.

***У клієнт-серверних ІС*** база даних та СУБД знаходяться на сервері, а на робочих станціях знаходяться клієнтські програми.



Класифікація інформаційних систем щодо охоплення завдань (масштабності) можна визначити так.

***Персональна інформаційна система*** варта вирішення деякого кола завдань однієї людини. Прикладом такої системи є система обробки контенту.

***Групова інформаційна система*** орієнтована колективне використання інформації членами робочої групи чи підрозділи.

***Корпоративна інформаційна система*** в ідеалі охоплює всі інформаційні процеси цілого підприємства, досягаючи їх повної узгодженості, повноти та прозорості. Такі системи у видавничій галузі називають комплексними редакційно-видавничими системами .

**Банк даних** є різновидом ІВ, в якій реалізовані функції централізованого зберігання та накопичення оброблюваної інформації, організованої в одну або декілька баз даних.

**База даних (БД)** являє собою сукупність спеціальним чином організованих даних, що зберігаються в пам'яті обчислювальної системи, що відображають стан об'єктів та їх взаємозв'язків у предметній галузі, що розглядається.

**Словник даних (СД)** являє собою підсистему БНД, призначену для централізованого зберігання інформації про структури даних, взаємозв'язки файлів БД один з одним, типи даних і формати їх подання, належність даних користувачам, коди захисту та розмежування доступу

**Адміністратор бази даних (АБД)** це особа або група осіб, які відповідають за вироблення вимог до БД, її проектування, створення, ефективне використання та супровід.

**Обчислювальна система (ВС)** являє собою сукупність взаємозалежних та узгоджено діючих ЕОМ або процесорів та інших пристроїв, що забезпечують автоматизацію процесів прийому, обробки та видачі інформації споживачам.

**Обслуговуючий персонал** виконує функції підтримки технічних та програмних засобів у працездатному стані .

# Моделі даних

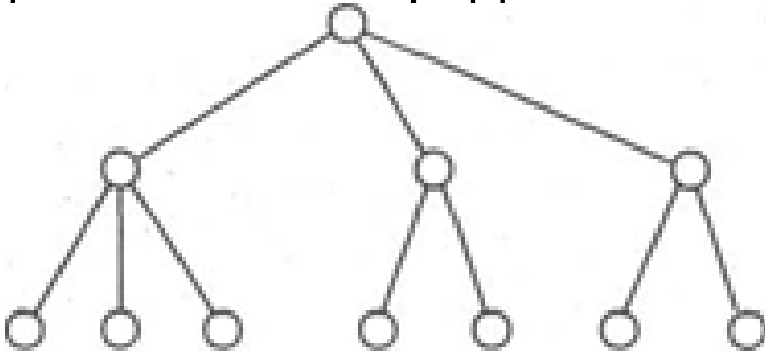
Дані, що зберігаються в базі, мають певну логічну структуру — іншими словами, описуються деякою моделлю подання даних (моделлю даних), що підтримується СУБД.

До класичних належать такі моделі даних :

- Ієрархічна ,
- Мережева ,
- Реляційна,
- Постреляційна ,
- Багатомірна ,
- об'єктно -орієнтована,
- Концептуальні.

# Ієрархічна модель

В ієрархічній моделі зв'язку між даними можна описати за допомогою впорядкованого графа (або дерева).

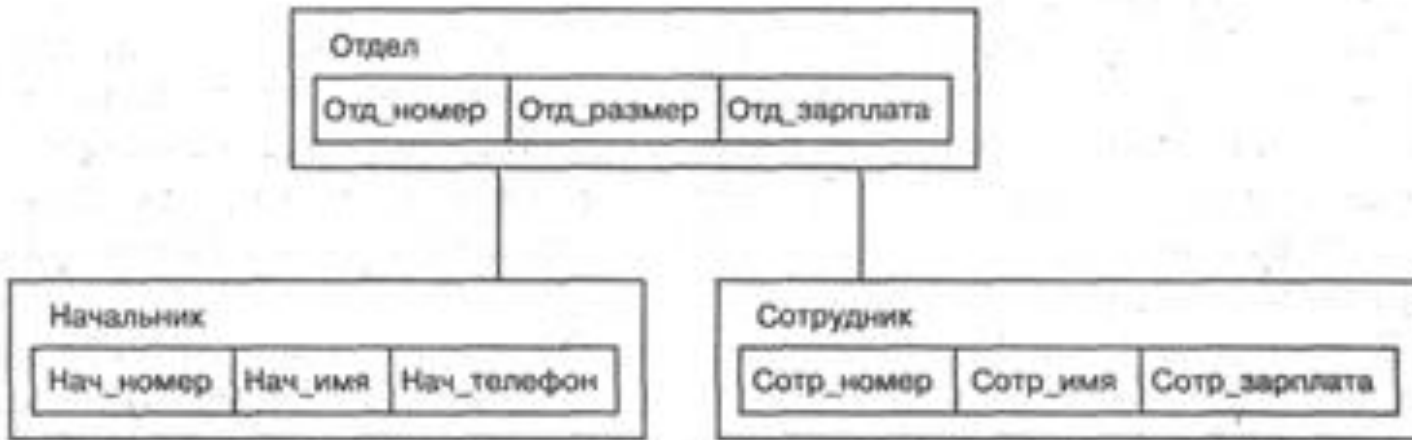


The screenshot shows the Windows Registry Editor. The left pane displays a tree view of the registry structure:

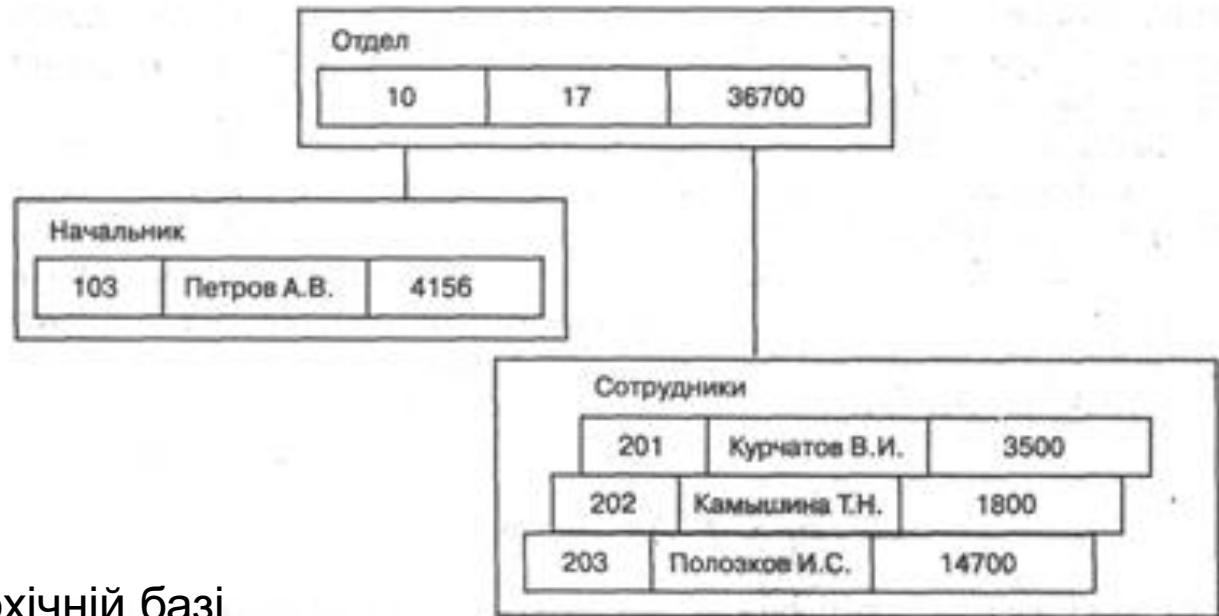
- HKEY\_CLASSES\_ROOT
  - OpenWithList
  - shell
    - Media Info
    - Command
  - removeproperties

The right pane shows a list of registry values:

Имя	Тип	Значение
ab (По умолчанию)	REG_SZ	"C:\Program



Приклад типу «дерево»

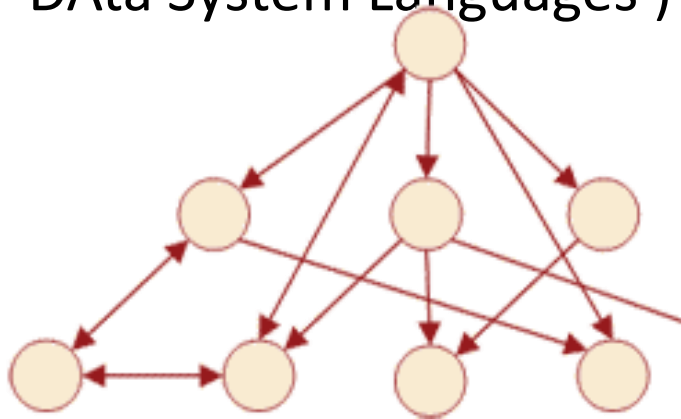


Дані в ієрархічній базі

# Мережева модель

Мережева модель даних є загальною структурою проти ієрархічної.

Основні принципи мережевої моделі даних були розроблені в середині 60-х років, еталонний варіант мережевої моделі даних описаний у звітах робочої групи з мов баз даних ( COncference on DAta System Languages ) CODASYL (1971 н )



Мережева модель даних

Приклад схеми мережевої БД



# Реляційна модель

Реляційна модель даних ґрунтується на понятті ставлення (relation ).

Едгар Франк "Тед" Кодд - (23 серпня 1923 -18 квітня 2003) - британський учений, роботи якого заклали основи теорії реляційних баз даних. У 1970 році видав роботу «A Relational Model of Data for Large Shared Data Banks », яка вважається першою роботою з реляційної моделі даних.

Атрибут  
(стовпець)

ID	Name	Phone	Birthday
1	Andrew	222-31-31	1976
2	Bohdan	264-05-06	1990
3	Ivan	237-01-02	1989

Відношення  
(таблиця)

Кортеж  
(стрічка)

# Постріляційна модель

Класична реляційна модель передбачає неподільність даних, які у полях записів таблиць. Це означає, що у таблиці подається у першій нормальній формі (залежність між атрибутами відносин).

Існує низка випадків, коли це обмеження заважає ефективній реалізації додатків.

У ряді випадків можливі значення атрибуту мають внутрішню структуру. Наприклад, дата народження, номер учбової групи студента. Правила нормалізації реляційної моделі даних потребують декомпозиції такого складного (складеного) атрибуту на кілька простих (атомарних). Це не завжди зручно та наочно.

### Накладні

Номери накладних	Номери покупців
0373	8723
8374	8232
7364	8723

### Накладні-товари

Номери накладних	Назва товару	Кількість
0373	Сир	
0373	Риба	2
8374	Лимонад	1
8374	Сік	6
8374	Печиво	2
7364	Йогурт	1

а)

### Накладні

Номери накладних	Номери покупців	Назва товару	Кількість
0373	8723	Сир	3
		Риба	2
8374	8232	Лимонад	1
		Сік	6
		Печиво	2
7364	8723	Йогурт	1

Структура даних реляційної (а) та постріляційної (б) моделей

б)

Перевагою постріляційної моделі є можливість представлення сукупності пов'язаних реляційних таблиць однією постріляційною таблицею.

Недоліком постріляційної моделі є складність вирішення проблеми забезпечення цілісності та несуперечності даних, що зберігаються.

# Багатовимірна модель

Джерелом для багатовимірних баз даних стали не сучасні технології баз даних, а алгебра багатовимірних матриць, яка використовувалася для ручного аналізу даних з кінця XIX століття.

Моделі	Місяці	Обсяги
Opel	червень	12
Opel	липень	24
Opel	серпень	5
Audi	червень	2
Audi	липень	18
Ford	липень	19

а

Моделі	червень	липень	серпень
Opel	12	24	5
Audi	2	18	–
Ford	–	19	–

б

Реляційне (а) та багатовимірне (б) подання даних

Розглянемо основні поняття багатовимірних моделей даних.

Багатовимірне подання даних поділяється на дві групи:

1. факти – чисельні параметри
2. виміри – текстові параметри надання максимально можливого контексту для фактів.

*Вимір* - упорядкований набір значень, що приймаються конкретним параметром, і відповідної однієї з граней гіперкуба.

*Факти* (заходи) – це дані, які кількісно описують процес, вони безперервні, тобто можуть приймати нескінченну кількість значень.

Сховища даних, як правило, містять такі три типи фактів.

1. Події ( event ), по крайнього заходу, лише на рівні найбільшої гранулярності, зазвичай, моделюють події реального світу, у своїй кожен факт представляє певний екземпляр досліджуваного явища. Прикладами можуть бути продажі, клацання мишею на Web -сторінці чи рух товарів складі .
2. Миттєві знімки ( snapshot ) моделюють стан об'єкта в даний момент часу, такі як рівні наявності товарів у магазині або на складі та кількість користувачів Web -сайту. Один і той самий екземпляр явища реального світу, наприклад, конкретна банка бобів, може виникати у кількох фактах .
3. Сукупні миттєві знімки ( cumulative snapshot ) містять інформацію про діяльність організації за певний відрізок часу. Наприклад, сукупний обсяг продажів за попередній період, включаючи поточний місяць, легко порівняти з показниками за відповідні місяці минулого року.

# Гіперкуб



A 3D hypercube visualization of a data cube. The top face shows the months: березень (March), лютий (February), and січень (January). The front face shows the countries: США (USA), Канада (Canada), and Мексика (Mexico). The side face shows the goods categories: Напої (Beverages), Продукти харчування (Food products), and Інші товари (Other goods). The data values are as follows:

	США	Канада	Мексика
Напої	10 000	2 000	1 000
Продукти харчування	5 000	500	250
Інші товари	5 000	500	250

Приклад гіперкуба

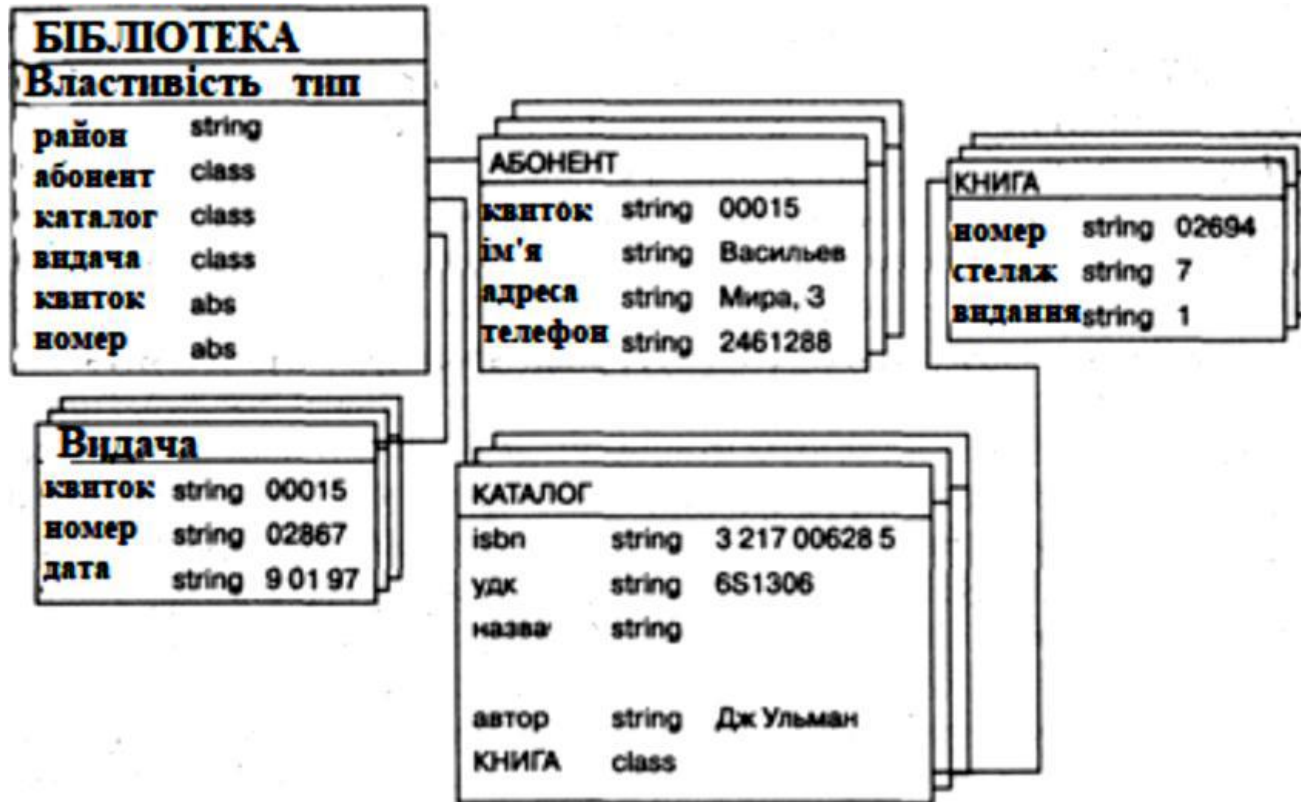
Двовимірний зріз куба для одного заходу

	США	Канада	Мексика
Січень	20 000	4 000	2 000
Лютий	30 000	6 000	3 000
Березень	50 000	10 000	5 000



# Об'єктно-орієнтована модель

У об'єктно-орієнтованій моделі при поданні даних є можливість ідентифікувати окремі записи бази. Між записами бази даних та функціями їх обробки встановлюються взаємозв'язки за допомогою механізмів, подібних до відповідних засобів, а об'єктно-орієнтованих мов програмування.



# Типи даних

Спочатку СУБД застосовувалися переважно на вирішення фінансово-економічних завдань. При цьому незалежно від моделі подання, у базах даних використовувалися такі основні типи даних :

- символльні типи;
- числові типи;
- типи дата/час;
- двійкові типи;
- користувальницькі типи.

**Символьні типи.** Символьні типи використовуються для представлення рядків символів, так і окремих символів.

Тип даних	Призначення
CHAR	Строковий тип. Довжина рядка зумовлена. Використовується переважно для коротких назв.
VARCHAR	Символьний тип довільної довжини. Використовується для розміщення абзаців
TEXT	Використовується для розміщення рядків до 6500 символів (наприклад, стаття).
MEDIUMTEXT	Використовується для розміщення рядків до 16,7 мільйонів символів (наприклад, книга).

**Числові типи.** Числові типи призначені для позначення цілих, речових та грошових типів.

Тип даних	Діапазон значень	Точність - число знаків після коми
INT	від -2 147 483 648 до +2 147 483 647	0
DOUBLE	від -1.797 e-308 до +1.797 e + 308	до 15
TINYINT	від 1 до 255	0
DECIMAL	Числа , що складаються з N цифр с M цифрами в дробовій частині. За замовчуванням N = 30, M = 6	M
MONEY	Для зберігання грошових величин. Припустимо значення NULL. Числа з 20 цифр, з 4 цифрами після коми	4

**Типи дата/час.** Типи дата/час призначені для зберігання часу, дат та дат разом із часом

Тип даних	Призначення
DATE	Тип для представлення дати у вигляді сукупності року, місяця і числа. Значення року може змінюватися в діапазоні від 0001 до 9999 року
TIME	Тип для представлення часу у вигляді сукупності години, хвилин, секунд і часток секунд. Частки секунд зберігаються з точністю до 6 знаків.
TIMESTAMP	Тип для представлення моменту часу конкретної дати. Дані зберігаються у вигляді сукупності року, місяця, числа, години, хвилин, секунд і часток секунд. Частки секунд зберігаються з точністю до 6 знаків.

# Формати подання даних типу дата / час, що визначаються за замовчуванням

Тип даних	Формат, який використовується за умовчанням
DATE	'YYYY-MM-DD'
TIME	'HH: NN: <u>ss.SSS</u> '
TIMESTAMP	'YYYY-MM-DD HH: NN: <u>ss.SSS</u> '
DATETIME	'YYYY-MM-DD HH: NN: <u>ss.SSS</u> '
SMALLDATETIME	'YYYY-MM-DD HH: NN: <u>ss.SSS</u> '

YYYY - чотири цифри , що позначають рік:

MM - дві цифри, що позначають місяць:

DD - дві цифри, що позначають день:

HH - дві цифри, що позначають години:

NN - дві цифри, що позначають хвилини:

ss - дві цифри, що позначають секунди:

SSS - три цифри, що позначають частки секунд.

**Двійкові типи.** Двійкові типи призначені для подання двійкових даних, включаючи зображення та іншу інформацію, що не обробляється власними засобами СУБД

Тип даних	Призначення
BIT	Тип для представлення значень 0 і 1.
BINARY	Теж, що і CHAR, за винятком операцій порівняння.
LONG BINARY	Тип для представлення двійкових даних довільної довжини. Довжина довільна. Обмежена максимальним розміром файлів бази даних (2 <u>Гбайта</u> )



# Системи управління базами даних

*Система управління базами даних (СУБД) - це комплекс мовних та програмних засобів, призначений для створення, ведення та спільного використання БД багатьма користувачами.*

Для роботи з інформацією СУБД, що зберігається в базі даних, надає програмам і користувачам наступні два типи мов:

- мова опису даних — непроцедурна високорівнева мова декларативного типу, призначена для опису логічної структури даних;
- мова маніпулювання даними — сукупність конструкцій, що забезпечують виконання основних операцій роботи з даними: введення, модифікацію та вибірку даних за запитамі.

Найбільшого поширення набули дві стандартизовані мови: QBE ( Query By Example ) — мова запитів за зразком (має властивості мови маніпулювання даними)

SQL ( Structured Query Language ) - структурована мова запитів.

Поєднує властивості властивостей обох типів - опису і маніпулювання даними.

Для створення та управління персональними БД та додатків, що працюють з ними, використовуються такі СУБД, як Access та Visual FoxPro фірми Microsoft , Paradox фірми Borland .

Корпоративна БД створюється, підтримується та функціонує під керуванням сервера БД, наприклад, Microsoft SQL Server або Oracle Server .

# Класифікація СУБД

```
graph TD; A[Класифікація СУБД] --> B[Вид програми]; A --> C[Характер використання]; A --> D[Моделі даних];
```

## Вид програми

повнофункціональні СУБД;  
сервери БД;  
клієнти БД;  
засоби розробки програм роботи з БД.

## Характер використання

локальні;  
розраховані на багато користувачів

## Моделі даних

ієрархічні;  
мережеві;  
реляційні;  
об'єктно орієнтовані;  
інші.

**Повнофункціональні СУБД** є традиційними системами, які спочатку з'явилися для великих машин, потім для мінімашин і для персональних ЕОМ.

До повнофункціональних СУБД належать такі пакети, як Clarion . Database Deweidper , DataEase , DataFlex , dBase IV, Microsoft Access , Microsoft FoxPro та Paradox R:BASE.

**Сервери БД** призначені для організації центрів обробки даних у мережах ЕОМ. Сервери БД реалізують функції управління базами даних, які запитують інші (клієнтські) програми зазвичай за допомогою операторів SQL.

Прикладами серверів БД є такі програми: NetWare SQL ( Novell ), MS SQL Server ( Microsoft ), InterBase ( Borland ), SQLBase Server ( Gupta ), Intelligent Database ( Ingress ).

У ролі **клієнтських програм** для серверів БД у випадку можуть використовуватися різні програми: електронні таблиці, текстові процесори, програми електронної пошти тощо.

**Засоби розробки програм роботи з БД** можуть використовуватися для створення різновидів таких програм:

- клієнтських програм;
- серверів БД та їх окремих компонентів;
- користувацьких додатків.

До засобів розробки додатків користувача відносяться системи: Delphi і Power Builder ( Borland ), Visual Basic ( Microsoft ), SILVERRUN ( Computer Advisers Inc. ), S- Designor (SDP і Powersoft ) та ERwin ( LogicWorks ).

**За характером використання** СУБД ділять на локальні та розраховані на багато користувачів.

*Локальні* СУБД зазвичай забезпечують можливість створення персональних БД та недорогих додатків, що працюють з ними. До персональних СУБД, наприклад, належать Visual FoxPro , Paradox , Clipper , dBase , Access та ін.

Організація функціонування локальної ІС однією комп'ютері серед певної операційної системи (ОС) можлива з допомогою наступних варіантів використання програмних средств:

- "повної" СУБД;
- додатки та «усіченої» (ядра) СУБД;
- незалежної програми.

**СУБД що розраховані на багато користувачів** включають сервер БД і клієнтську частину і, як правило, можуть працювати в неоднорідному обчислювальному середовищі (з різними типами ЕОМ і операційними системами).



Файл-серверні СУБД



Клієнт-серверні СУБД



***По моделі даних*** СУБД, що використовуються, поділяють на ієрархічні, мережеві, реляційні, об'єктно-орієнтовані та інші типи. Деякі СУБД можуть одночасно підтримувати кілька моделей даних.

# Етапи проектування баз даних

Основні етапи проектування баз даних такі:

**Концептуальне проектування (інфологічна модель)** – побудова формалізованої моделі предметної галузі .

**Логічне проектування (датологічна модель)** – відображення інфологічної моделі на моделі даних, що використовуються в конкретній СУБД, наприклад, на реляційній моделі даних .

**Фізичне проектування** - реалізація датологічної моделі засобами конкретної СУБД, і навіть вибір рішень, що з фізичним середовищем зберігання даних.