

# Introduction au Data Mining



# Sommaire

- Introduction aux BDD
- Un Problème et une Solution
- Qu'est –ce que le Data Mining?
- Objectif du Data Mining
- Qu'est-ce qui n'est pas du Data Mining?
- Convergence de trois technologies clés
- Fonctions du Data mining
- Types du Data Mining
- Problèmes



# Qu'est-ce qu'une Base de données


Une base de données est une collection organisée de données.



# Examples

# Collaborateurs

 Search Friends

 Find Friends


 All Friends

 Browse

 Phonebook

 Recently Added

## Lists

 Facebook


 Drew

 San Francisco, CA

 family

 best friends

 co-workers

 yoga buddies

 limited profile

 Create

Create New List

Edit List

Delete List



**Alexandre Roche**  
Facebook co-workers



**Anikka Fragodt**  
Facebook co-workers



**Ankur Pansari**  
Facebook co-workers



**Aubrey Obata**  
Facebook co-workers



**Cameron Marlow**  
Facebook co-workers



**Leah Pearlman**  
Facebook co-workers



**Meredith Chin**  
Facebook co-workers

# Examples

# Informations des Patients

**Edit a Patient - Samantha Jones [894567]**

Ledger Schedule Lookup View Tx Card Dx/Tx Comment Documents Imaging Print Lette...

**Patient** | Patient (more) | Status | Relationships | Billing Parties | Insurance | Fees | Appointments | Referrals | History | Custom

First/Last Name:

Sort Name:

Greeting:  Gender:

Street:

City:

State:  Zip:

Birthdate:  14 years

Phone #s:

Type	Number	Note
Home	(818) 693-0369	

SSN / Gov't ID #:

E-mail:

Mail Method:


Alerts:

Date	Alert
N/A	Requires Penicillin before ...
N/A	Allergic to Latex

Patient ID:

Balance Summary - All Billing Parties:

	Patient	Insurance	Total
Current w/o Credits	\$200.00	\$0.00	\$200.00
Credits	\$0.00	\$0.00	\$0.00
Current w/ Credits	\$200.00	\$0.00	\$200.00
Over 30	\$0.00	\$0.00	\$0.00
Over 60	\$0.00	\$0.00	\$0.00
Over 90	\$200.00	\$0.00	\$200.00
Total Due	\$400.00	\$0.00	\$400.00
Total Past Due	\$200.00	\$0.00	\$200.00
Unbilled Balance	\$400.00	\$0.00	\$400.00



# Exemples Système de Réserveation de vol

TASAR - PRODUCTION

File Edit Booking Inventory Queues Accounting Tools DCS Maintenance Window Help



**Availability**

Origin:  ...    
 Destination:  ...

Departure Date:    
 Duration of Stay:  Days   
 Return Date:

Number of Passengers: Adults:  Children:  Infants:

Booking Type:  Roundtrip  One-Way  Other   
 Show:  Available Only  All Flights   
 Non Revenue  Roundtrip Fares

No	Flight	Dep.	Arr.	Date	Time
1	* NK604	FLL	LGA	Wed, 19Nov2003	07:10
	C 263.00 A 198.00 Y 223.00 B 198.00 H 173.00 Q 148.00 M				
2	* NK1758	FLL	LGA	Wed, 19Nov2003	09:00
	C 263.00 A 198.00 Y 223.00 B 198.00 H 173.00 Q 148.00				
3	* NK770	FLL	LGA	Wed, 19Nov2003	11:00
	C 263.00 A 198.00 Y 223.00 B 198.00 H 173.00 Q 148.00 M				
4	* NK196	FLL	LGA	Wed, 19Nov2003	15:20
	C 263.00 A 198.00 Y 223.00 B 198.00 H 173.00 Q 148.00 M				
5	* NK1582	FLL	LGA	Wed, 19Nov2003	17:50
	C 263.00 A 198.00 Y 223.00 B 198.00 H 173.00 Q 148.00 M				

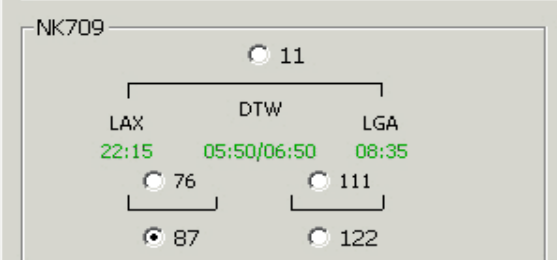
No	Flight	Dep.	Arr.	Date	Time
1	* NK705	LGA	FLL	Wed, 19Nov2003	06:50
	C 263.00 A 198.00 Y 223.00 B 198.00 H 173.00 Q 148.00 M				
2	* NK1581	LGA	FLL	Wed, 19Nov2003	07:00
	C 263.00 A 198.00 Y 223.00 B 198.00 H 173.00 Q 148.00 M				
3	* NK1123	LGA	FLL	Wed, 19Nov2003	09:30
	C 263.00 A 198.00 Y 223.00 B 198.00 H 173.00 Q 148.00 M				
4	* NK197	LGA	FLL	Wed, 19Nov2003	12:30
	C 263.00 A 198.00 Y 223.00 B 198.00 H 173.00 Q 148.00 M				
5	* NK751	LGA	FLL	Wed, 19Nov2003	15:15

Departure Date:    
 Duration of Stay:  Days   
 Return Date:

Non Revenue  Roundtrip Fares

**Passenger Manifest**

Filter: Flight number:  Flight date:



Passenger Name	Locator	Cat.
▶ AMBROSE/DANIEL	HLCJ2T	S
BAGHDASARIAN/NERSIS	HLKCJZ	S
BARARSANI/NIMA	HKGFXV	S
BECK/DAVID	HKKFQP	S
BERBERIAN/ARA	HKZNH D	S
BLACK/JESSERAY	HKZFTV	S
BLACK/THOMASCOLBY	HKZFTV	S
CAMERO/LOUIS	HKWXXJ	S
CHEN/LI	HKLMTM	S
CZEWSKI/RACHEL A	HLJNBP	S

Rows: 87

**AboutTASAR**

## TASAR

Version 1.0.234

Copyright (c) 2003 RRT Ltd

The World's First Airline Reservation System based on a three-tiered network design concept

**This product is licensed to:**  
Spirit Airlines

Visit our web site at: [www.RRTechnologies.net](http://www.RRTechnologies.net)

Login information

User: PETER City: FLL  
 Company: NK Branch: 0001

Groups

Administrators

# Donnée vs. information

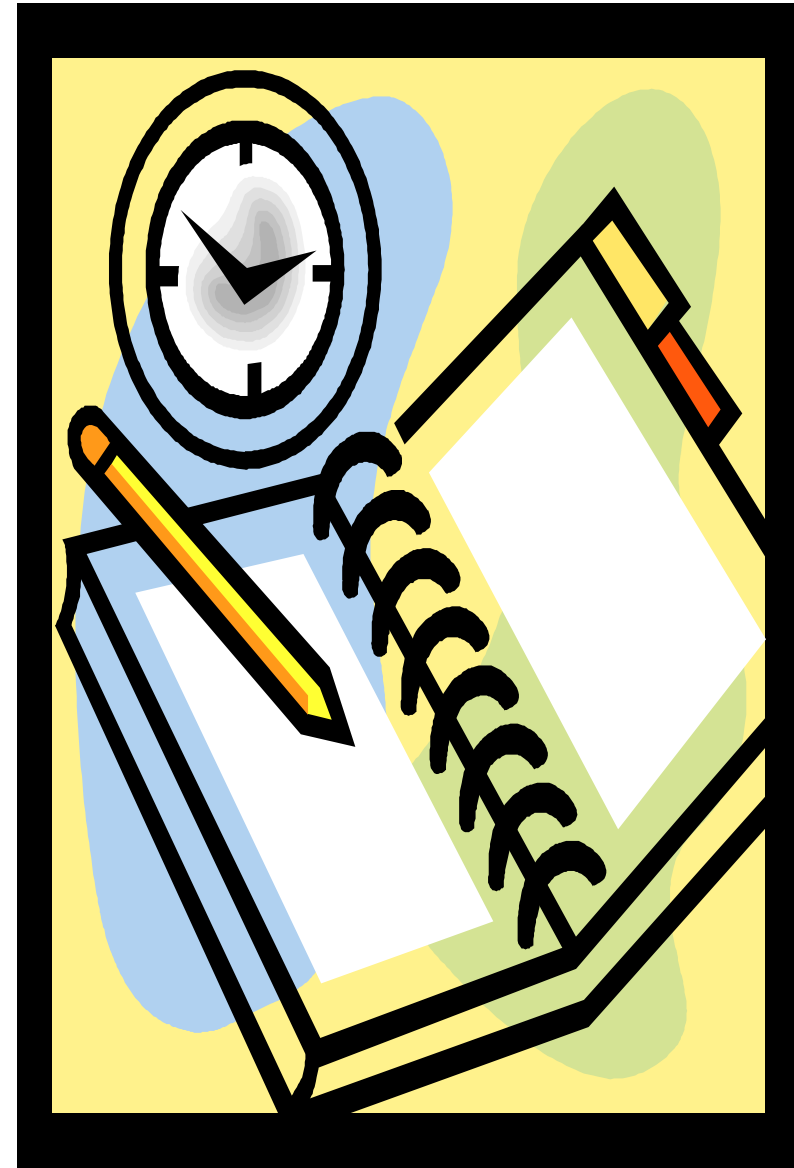
- Qu'est-ce qu'une donnée?
  - Une donnée est une information non traitée.



- Qu'est-ce qu'une **information**?
  - Une Information est une donnée organisée et communiquée de telle façon qu'elle a un sens.
  - La **donnée** est transformée en **information**, et l'information est transformée en **connaissance**.
  - Connaissance, information évaluée et organisée de telle sorte qu'elle peut être utilisée délibérément.

# Pourquoi on a besoin d'une BDD?

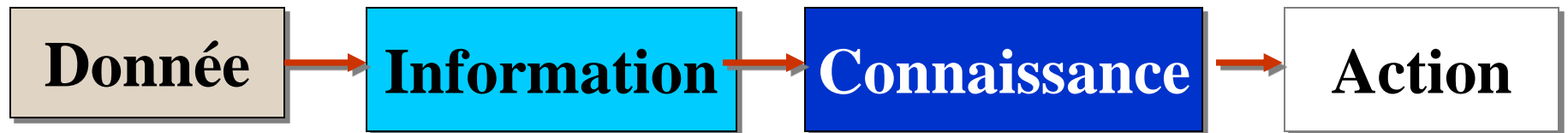
- Garder les enregistrements de nos:
  - Clients
  - Staff
  - Volontaires
- Garder les enregistrements de nos activités et nos interventions
- Garder les ventes
- Développer les rapports
- Améliorer la recherche





# La mission d'un SGBD

Est de transformer



# BDD et SGBD

- **Base de données (BDD):** Collection partagée de données logiquement liées (avec une description de ces données), conçue pour atteindre les besoins en information d'une organisation.
- **Systeme de Gestion de Bases de Données (SGBD):** Un système informatique permettant aux utilisateurs de définir, créer et maintenir la base de données tout en assurant le contrôle d'accès à cette base de données.

# Qui et Comment ?

- C'est le SGBD qui fait ce travail.
- En utilisant des logiciels (programmes) informatiques: Access, FileMaker, Lotus Notes, Oracle or SQL Server, .....
- Ces logiciels contiennent des outils pour ajouter, modifier ou supprimer des données de la BDD, poser des questions (requêtes) à propos des données stockées dans la BDD et produire des rapports et des résumés de certaines données de la BDD.