ECE/CS230
Computer Systems Security

Charalambos (Harrys) Konstantinou

https://sites.google.com/view/ececs230kaust/

Databases

Overview

1. Introduction to Databases

2. Database Security Requirements

3. Attacks to Databases & Best Practices to Secure Databases

4. Data Mining & Big Data

Introduction to Databases

- **Database:** a collection of data and a set of rules to organize the data according to relationships.
  - The user interacts with the data through the set of rules.
  - The format of the file is not concern of to the user.

- **Database administrator:** a person who defines the rules and controls the access to the data.

- **Database manager or Database management system (DBMS)**: software or program the user uses to interact with the database.

Introduction to Databases

Components of a Database:

- **Records** – contain **fields** or **elements**
- **Schema** – logical structure of a database (blueprint of organization)
  - **Subschema** - a sub part of a schemas that describes a different view of the database (schemas may have different subschemas)

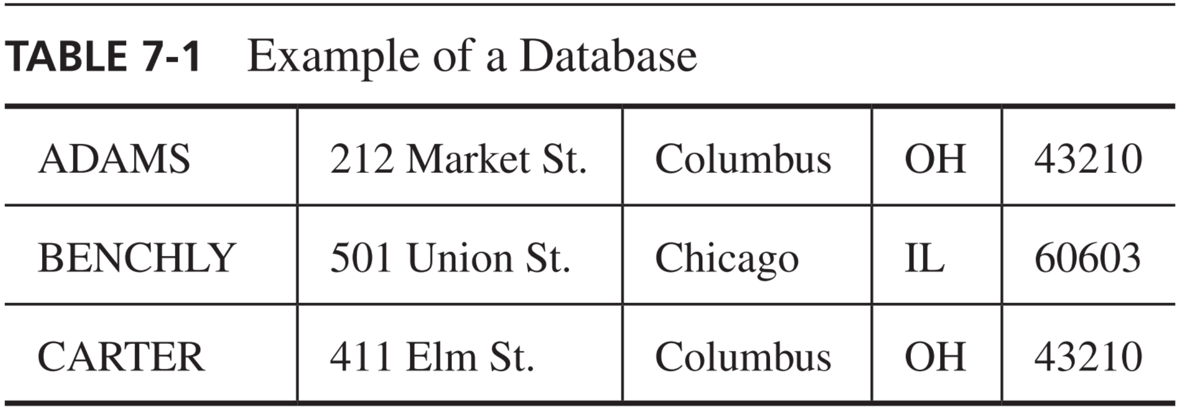


TABLE 7-1 Example of a Database

ADAMS	212 Market St.	Columbus	OH	43210
BENCHLY	501 Union St.	Chicago	IL	60603
CARTER	411 Elm St.	Columbus	OH	43210

Security in Computing 5th Edition by Charles Pfleeger, Shari Pfleeger, Jonathan Margulies

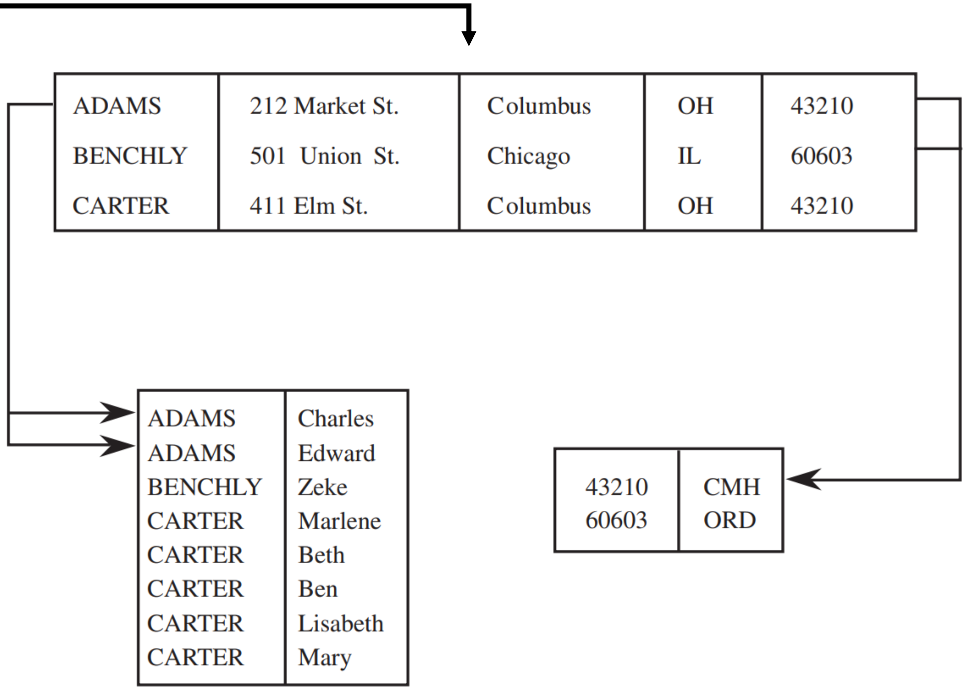
Introduction to Databases

Components of a Database:

- **Subschema**

TABLE 7-2 Schema of Database from Figure 7-1

Name	First	Address	City	State	Zip	Airport
ADAMS	Charles	212 Market St.	Columbus	OH	43210	CMH
ADAMS	Edward	212 Market St.	Columbus	OH	43210	CMH
BENCHLY	Zeke	501 Union St.	Chicago	IL	60603	ORD
CARTER	Marlene	411 Elm St.	Columbus	OH	43210	CMH
CARTER	Beth	411 Elm St.	Columbus	OH	43210	CMH
CARTER	Ben	411 Elm St.	Columbus	OH	43210	CMH
CARTER	Lisabeth	411 Elm St.	Columbus	OH	43210	CMH
CARTER	Mary	411 Elm St.	Columbus	OH	43210	CMH



ADAMS	212 Market St.	Columbus	OH	43210
BENCHLY	501 Union St.	Chicago	IL	60603
CARTER	411 Elm St.	Columbus	OH	43210

ADAMS	Charles
ADAMS	Edward
BENCHLY	Zeke
CARTER	Marlene
CARTER	Beth
CARTER	Ben
CARTER	Lisabeth
CARTER	Mary

43210	CMH
60603	ORD

Security in Computing 5th Edition by Charles Pfleeger, Shari Pfleeger, Jonathan Margulies

Introduction to Databases

Components of a Database:

- **Attribute:** the name of each column in a database
- **Relation**: a set of columns (or attributes)

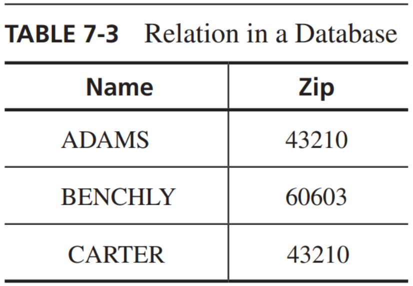


TABLE 7-3 Relation in a Database

Name	Zip
ADAMS	43210
BENCHLY	60603
CARTER	43210

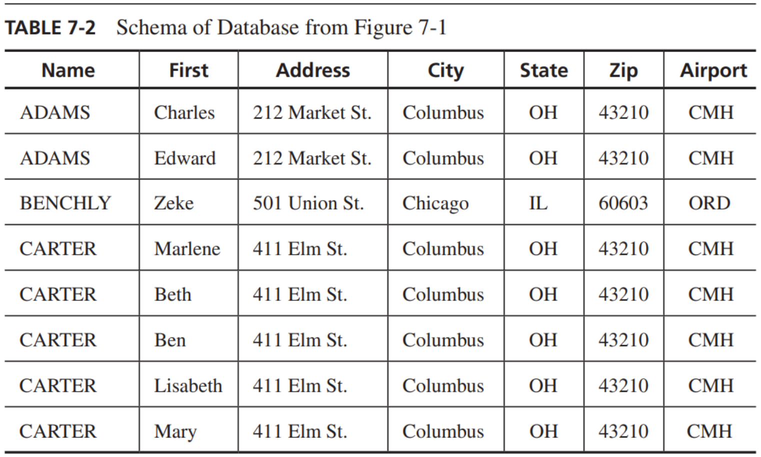


TABLE 7-2 Schema of Database from Figure 7-1

Name	First	Address	City	State	Zip	Airport
ADAMS	Charles	212 Market St.	Columbus	OH	43210	CMH
ADAMS	Edward	212 Market St.	Columbus	OH	43210	CMH
BENCHLY	Zeke	501 Union St.	Chicago	IL	60603	ORD
CARTER	Marlene	411 Elm St.	Columbus	OH	43210	CMH
CARTER	Beth	411 Elm St.	Columbus	OH	43210	CMH
CARTER	Ben	411 Elm St.	Columbus	OH	43210	CMH
CARTER	Lisabeth	411 Elm St.	Columbus	OH	43210	CMH
CARTER	Mary	411 Elm St.	Columbus	OH	43210	CMH

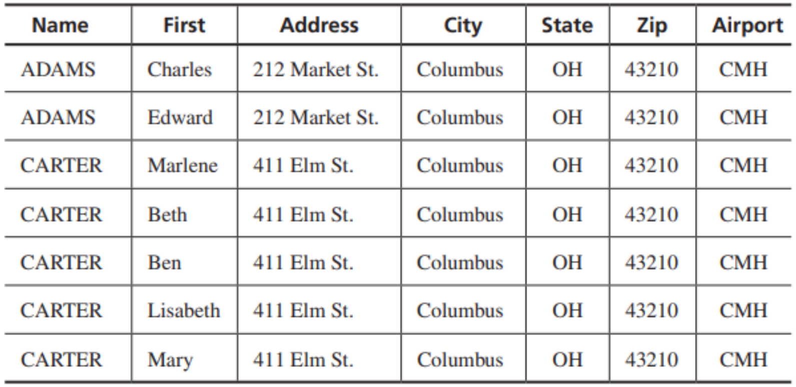
Security in Computing 5th Edition by Charles Pfleeger, Shari Pfleeger, Jonathan Margulies

Introduction to Databases

Components of a Database:

- **Queries**: A command used by an user accessing a DBMS
  - The result of a query is a **Subschema**.
  - Most current Databases (Microsoft SQL, MySQL, SQLite, etc.) are based on the **Structured Query Language (SQL)**
  - *and (^), or (v), other comparisons such as (<,>, etc)*

SELECT (ZIP='43210') ^ (NAME='ADAMS')



Name	First	Address	City	State	Zip	Airport
ADAMS	Charles	212 Market St.	Columbus	OH	43210	CMH
ADAMS	Edward	212 Market St.	Columbus	OH	43210	CMH
CARTER	Marlene	411 Elm St.	Columbus	OH	43210	CMH
CARTER	Beth	411 Elm St.	Columbus	OH	43210	CMH
CARTER	Ben	411 Elm St.	Columbus	OH	43210	CMH
CARTER	Lisabeth	411 Elm St.	Columbus	OH	43210	CMH
CARTER	Mary	411 Elm St.	Columbus	OH	43210	CMH

Security in Computing 5th Edition by Charles Pfleeger, Shari Pfleeger, Jonathan Margulies

Introduction to Databases

SHOW FIRST WHERE (ZIP='43210') ^ (NAME='ADAMS')

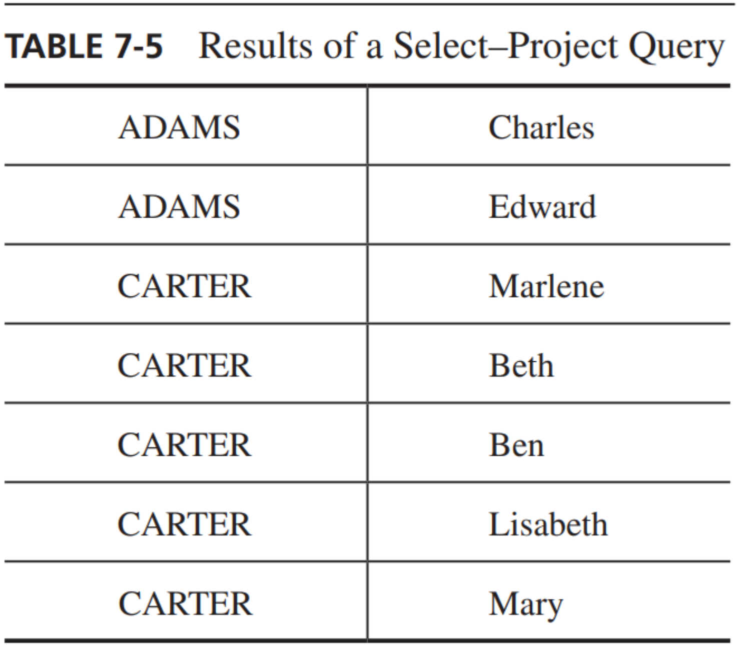


TABLE 7-5 Results of a Select–Project Query

ADAMS	Charles
ADAMS	Edward
CARTER	Marlene
CARTER	Beth
CARTER	Ben
CARTER	Lisabeth
CARTER	Mary

Security in Computing 5th Edition by Charles Pfleeger, Shari Pfleeger, Jonathan Margulies

Learn by Using: SQLite & SQLite DB Browser

- **SQLite** is a a cross-platform relational database management system (RDBMS)

- Software: **DB Browser for SQLite**

See appended video demo.

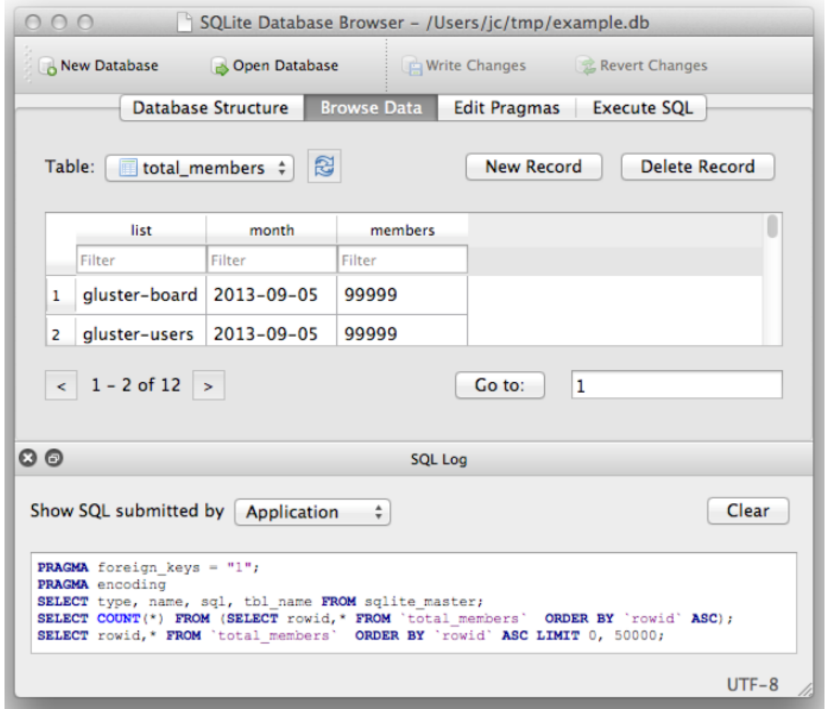


Table: total_members

	list	month	members
	Filter	Filter	Filter
1	gluster-board	2013-09-05	99999
2	gluster-users	2013-09-05	99999

< 1 - 2 of 12 >

Go to: 1

Show SQL submitted by: Application

```
PRAGMA foreign_keys = "1";
PRAGMA encoding
SELECT type, name, sql, tbl_name FROM sqlite_master;
SELECT COUNT(*) FROM (SELECT rowid,* FROM `total_members`  ORDER BY `rowid` ASC);
SELECT rowid,* FROM `total_members`  ORDER BY `rowid` ASC LIMIT 0, 50000;
```

UTF-8

https://sqlitebrowser.org/

Introduction to Databases

Advantages of Using Databases:

1. **Shared access**: users can use one common database and not multiple.
2. **Controlled access**: only authorized users can be allowed to view or modify
3. **Minimal redundancy**: individual users do not have to maintain their own sets of data.
4. **Data consistency**: a value change will be reflected for all users at the same time.
5. **Data integrity**: data is protected against accidental or malicious changes.

Problem using Excel as a database for COVID-19 cases

https://www.bbc.com/news/technology-54423988

Security in Computing 5th Edition by Charles Pfleeger, Shari Pfleeger, Jonathan Margulies

Database Security Requirements

- **Physical database integrity.** Data must be immune from physical problems, such as power failures, and you can reconstruct the database if destroyed.
- **Logical database integrity**. The structure of the database is preserved. With logical integrity of a database, a modification to the value of one field does not affect other.
- **Element integrity**. The data contained in each element are accurate.
- **Auditability**. You can track who or what has accessed (or modified) the database.
- **Access control**. Users are allowed to access only authorized data.
- **User authentication**. Every user is identified, both for the audit trail and for permissions.
- **Availability.** Users can access the database in general and all the data for which they are authorized.

Security in Computing 5th Edition by Charles Pfleeger, Shari Pfleeger, Jonathan Margulies

Database Security Requirements

Physical, Logical & Element Integrity

- **Physical:** Protect against database corruption – outside forces such as fire or power failures.
- **Logical:** Protect against database corruption – illegal programs/software

The best way to protect a database on integrity is to perform ***periodic backups***.

- **Element**: correctness or accuracy of elements inside the database.

The best way to protect a database on element integrity is to:

1. Do **field checks**
2. Perform **access control**
3. Have a **change log**

Security in Computing 5th Edition by Charles Pfleeger, Shari Pfleeger, Jonathan Margulies

Database Security Requirements

Auditability (Integrity)

- Generate audit record of all access made to the database (reads and writes)
- Granularity may become a problem in auditing some systems.
- Sophisticated DBMS may have better granularity.

Access Control (Confidentiality)

- Databases can be logically separated by user access privileges
- The database administrator is in charge of specifying access controls
- **Inference**: users able to access data values from others. In other words, infer or derive sensitive data from non-sensitive data.

Security in Computing 5th Edition by Charles Pfleeger, Shari Pfleeger, Jonathan Margulies

Database Security Requirements

User Authentication (Confidentiality)

- DBMS must perform user, password, and time-of-day checks.
- DBMS must perform their own authentication and not rely on other programs or the OS.

Availability

- DBMS have traits from programs and systems.
- DBMS need to be available when users request data from them.
- Users may want to access the same record at the same time, so mechanisms must be put in place to handle them.

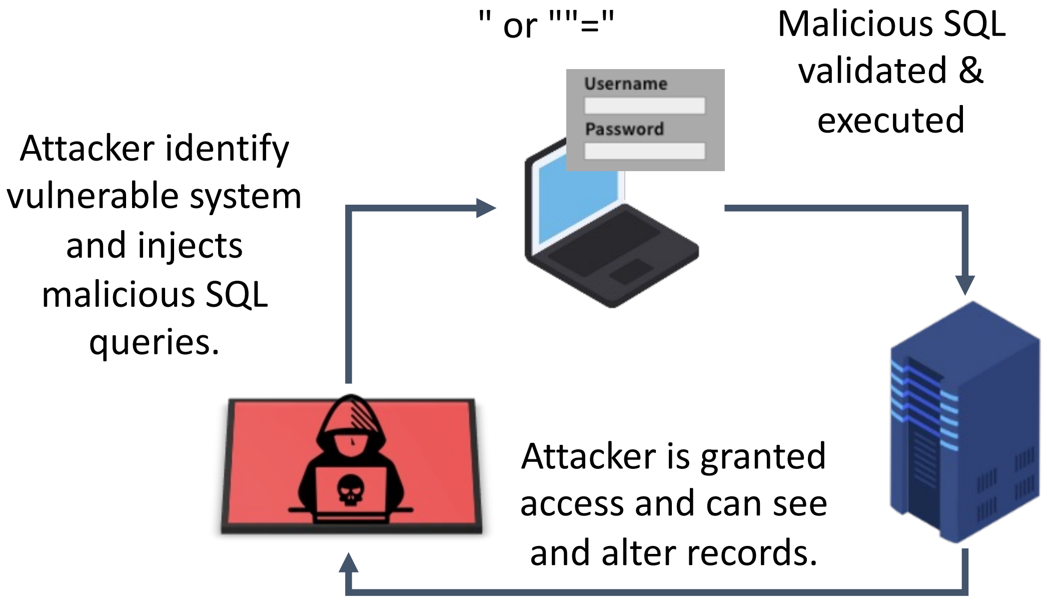
Security in Computing 5th Edition by Charles Pfleeger, Shari Pfleeger, Jonathan Margulies

Attacks to Databases

SQL Injection: common attack where adversary injects SQL queries (malicious code) into fields that connect to a database.

Basic practices to secure fields:

1. Parameterize your Queries instead of directly embedding user input in them.
2. Escape the characters that have a special meaning in SQL.
3. Pattern-check your parameters.
4. Restrict access to sensitive tables with database permission



" or ""="

Malicious SQL validated & executed

Attacker identify vulnerable system and injects malicious SQL queries.

Attacker is granted access and can see and alter records.

Attacks to Databases

Weak Authentication: weak authentication of database credentials or vulnerabilities to brute force attacks.

Best practices for authentication:

- Implement brute force controls (account lockout after many invalid attempts or Use password blacklisting)
- Require or encourage users to regularly change passwords
- Implement multi-factor authentication
- Don’t store user passwords in the clear (use strong hashing algorithms)
- Protect the application database credentials (make sure they are unguessable)

Attacks to Databases

Privilege Abuse & Excessive Privileges: Incorrectly (or by mistake) assigning higher than required privileges to users may put in danger the DBMS.

Users should only be given the minimum access required, no more no less.

Best practices for avoiding Privilege Abuse:

- The rate of user access to data should be limited.
- The DBMS must not have exposed interfaces that allow arbitrary queries and bulk export of data.
- Any arbitrary query, access, or use of the database should be logged, regularly audited, and limited to as few people as possible.
- Role-based controls
- Procedures to update permissions when someone moves or changes roles.

Attacks to Databases

Inadequate logging and weak auditing: logging and auditing are essential to know what changes have been done and by whom in a database.

- It helps recovering original data from maliciously altered data if an attack occurs.
- Also can help investigating how the attack was done and what was changed.

Best practices for Logging and Auditing:

- Clarify the information to collect at the application and database query layer.
- Secure your logged data.
- Implement procedures for auditing the data collected.
- Consider implementing network-based audit.

Attacks to Databases

Denial of service (DoS): DoS attacks can compromise the accessibility and availability of DBMS.

- Resource consumption-based attacks: repeatedly sending complex search queries to exhaust server resources

Cloud-based DoS protection services can be used to defend against this types of attacks.

Inadequate Backup: Not having adequate backup makes the system vulnerable to integrity attacks and states where the data can be unrecoverable.

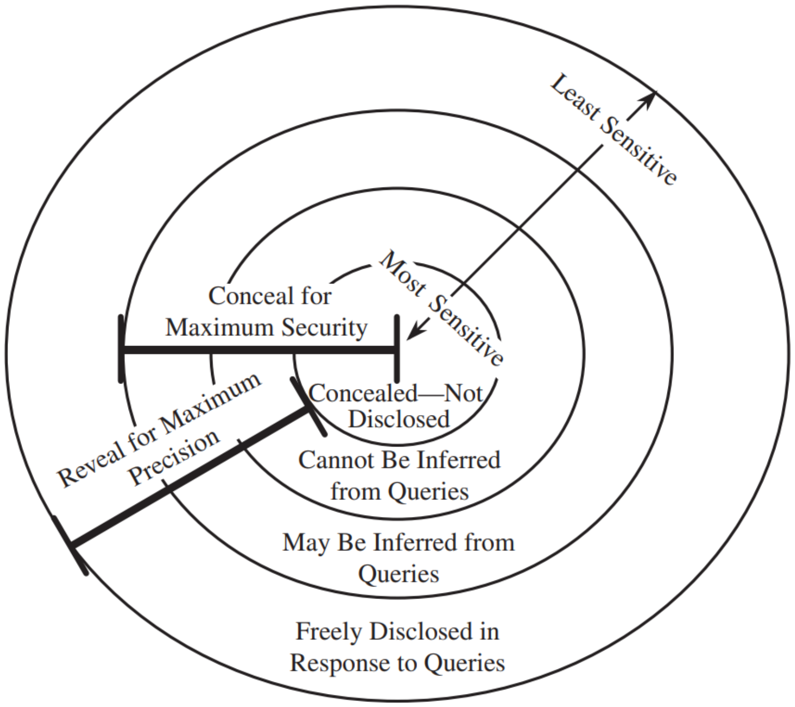
The best practices to perform adequate backup and secure these backups are:

- All backups should be encrypted to protect confidentiality and integrity of the data.
- Have backups offline (not connected to internet)
- Cloud services is **not** the same as **backup**. Make sure to have multiple physical backups that you can rely on.

Security Vs. Precision

Security: Forbids any queries that access sensitive data. (Secure Queries)

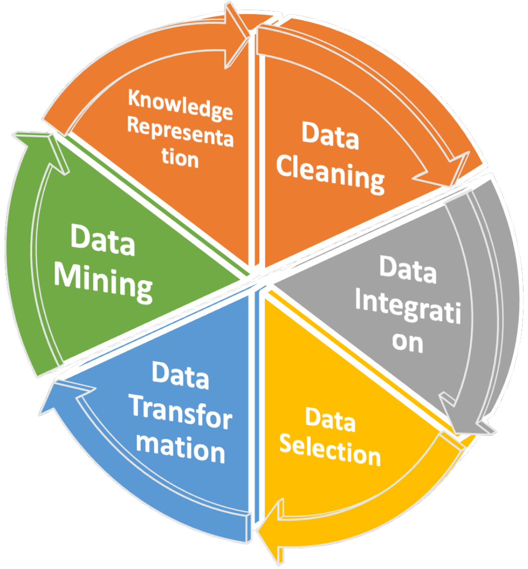
Precision: Aggregated result should reveal as much non-sensitive data as possible. (Precise Queries)


Least Sensitive
Most Sensitive
Conceal for Maximum Security
Reveal for Maximum Precision
Concealed—Not Disclosed
Cannot Be Inferred from Queries
May Be Inferred from Queries
Freely Disclosed in Response to Queries

Security in Computing 5th Edition by Charles Pfleeger, Shari Pfleeger, Jonathan Margulies

Data Mining & Big Data

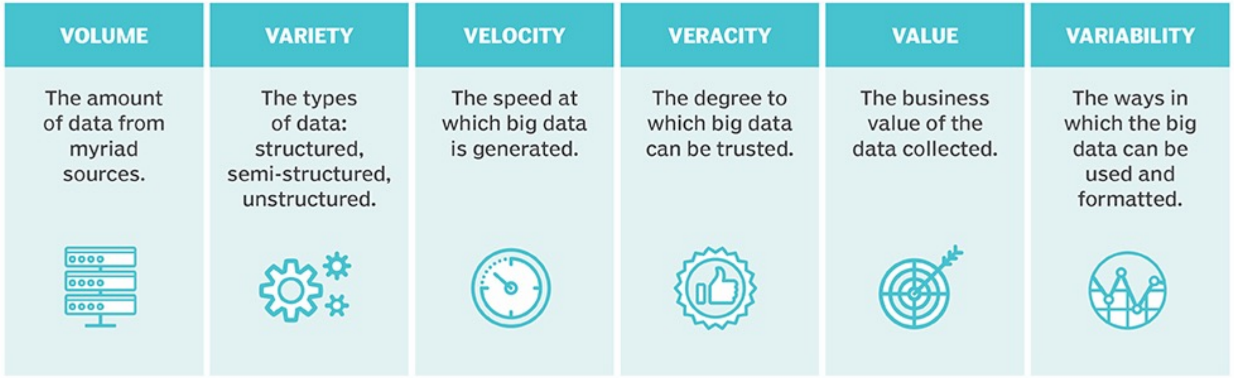
Data Mining: ways of using statistics, machine learning, mathematical models, pattern recognition, and other techniques to discover patterns and relations on large datasets.



Data Mining & Big Data

Big Data: analysis of massive amounts of data, often collected from different resources or databases.

The use of massive amounts of data from varied sources is often referred to as big data.



VOLUME	VARIETY	VELOCITY	VERACITY	VALUE	VARIABILITY
The amount of data from myriad sources.	The types of data: structured, semi-structured, unstructured.	The speed at which big data is generated.	The degree to which big data can be trusted.	The business value of the data collected.	The ways in which the big data can be used and formatted.

https://searchdatamanagement.techtarget.com/definition/big-data

Data Mining & Big Data

Problems in Data Mining & Big Data

1. **Privacy & Sensitivity**
2. **Data correctness & Integrity**
3. **Availability of Data**

Security in Computing 5th Edition by Charles Pfleeger, Shari Pfleeger, Jonathan Margulies