

## Kinds of Data Meant for Mining

### 1. Flat Files:

- **Description:** Simple data files in text or binary format. Data is structured but without inherent relationships between records (like in a relational database).
- **Example:** A CSV (Comma Separated Values) file containing customer data: Name, Age, City\nJohn, 30, New York\nJane, 25, London. A handwritten list of names and addresses.
- **Application:** Used in Data Warehousing to store data, used in carrying data to and from server, etc.

### 2. Relational Databases:

- **Description:** Data organized into tables with rows (records/tuples) and columns (attributes/fields). Relationships between tables are defined using keys. Uses SQL (Structured Query Language) for access.
- **Example:** An Orders table (OrderID, CustomerID, OrderDate) linked to a Customers table (CustomerID, Name, Address).
- **Application:** Data Mining, ROLAP model, etc.

### 3. Data Warehouses:

- **Description:** A central repository of integrated data from multiple sources, designed for analysis and reporting rather than transaction processing. Often uses a dimensional model (star schema, snowflake schema).
- **Example:** A data warehouse combining sales data from different regions, product information, and customer demographics, enabling analysis of sales trends over time.
- **Application:** Business decision making, Data mining, etc.

### 4. Object-Relational Databases:

- **Description:** Combines features of relational databases and object-oriented databases. Supports complex data types (objects) and inheritance.
- **Example:** A database storing product information where each product is an object with properties like name, description, and a list of associated images (also objects).

### 5. Transactional Databases:

- **Description:** Databases that support transactions (sequences of operations treated as a single unit). Emphasize data integrity and consistency (ACID properties: Atomicity, Consistency, Isolation, Durability).
- **Example:** A banking database recording account transactions (deposits, withdrawals). Rollback capabilities are crucial.
- **Application:** Banking, Distributed systems, Object databases, etc.

### 6. Multimedia Databases:

- **Description:** Databases storing audio, video, images, and text. Often use object-oriented or object-relational models to handle complex data.
- **Example:** A database of movies with video files, metadata (title, actors, genre), and user reviews.
- Application: Digital libraries, video-on demand, news-on demand, musical database, etc.

## 7. Spatial Databases:

- **Description:** Databases that store geographic or spatial data (coordinates, shapes, topology). Used for mapping, location-based services, and geographic information systems (GIS).
- **Example:** A database storing map data (roads, buildings, points of interest) for a navigation app.
- Application: Maps, Global positioning, etc.

## 8. Time-Series Databases:

- **Description:** Databases optimized for storing and analyzing data that changes over time. Data is typically indexed by time.
- **Example:** A database storing stock prices, sensor readings from an industrial machine, or website traffic logs.
- Application: eXtremeDB, Graphite, InfluxDB, etc.

## 9. World Wide Web (WWW):

- **Description:** A vast, heterogeneous collection of documents and resources (text, images, videos, etc.) accessible via the internet. Data is often unstructured or semi-structured.
- **Example:** Web pages, social media posts, online product catalogs, and scientific publications. Mining the web involves techniques like web scraping, text mining, and link analysis.
- Application: Online shopping, Job search, Research, studying, etc.