# Assignment 1

**1-28**. **Reread the definitions for data and database in this chapter. Database management systems only recently began to include the capability to store and retrieve more than numeric and textual da ta. What special data storage, retrieval, and maintenance capabilities do images, sound, video, and other advanced data types require that are not required or are simpler with numeric and textual data?**

**A-1-28.** data refers to facts concerning objects and events that could be recorded and stored on computer media. A database is an organized collection of logically related data.

The traditional definition of data now needs to be expanded to reflect a new reality: Databases today are used to store objects such as documents, e-mails, tweets, GPS information, maps, photographic images, sound, and video segments in addition to structured data. This type of data is referred to as unstructured data, or as multimedia data. Big data technologies have emerged as they are characterized by their ability to deal with large volumes of data with a variety of data types arriving at the organizational systems with high velocity.

The capabilities of storage, retrieval, and maintenance of Multi-object majorly depend on below factors:

1. **Storage of Large volume data**: It needs substantial storage capabilities.
2. **Retrieving Complex structure**: has an un-conventional structure that requires additional software and efforts to retrieve data or optimize the access.
3. **Maintenance:** Maintenance needs extra efforts and maybe multiple updates because the entire object needs to be restored rather than its parts to keep data consistent.


**1-29. Table 1-1 shows example metadata for a set of data items. Identify three other columns for these data (i.e., three other metadata characteristics for the listed attributes) and complete the entries of the table in Table 1-1 for these three additional columns.**
**A-1-29.**

| Name | Type | Length | Min | Max | Description | Source | Create date | Last modification date | Author |
|------|------|--------|-----|-----|-------------|--------|-------------|------------------------|--------|
| Course | Alphanumeric | 30 | | | Course ID and Name | Academic Unit | When a new entry is added | When someone modify course in DB | User inserted the data |
| Section | Integer | 1 | 1 | 9 | Section Number | Registrar | | | |
| Semester | Alphanumeric | 10 | | | Semester and Year | Registrar | | | |
| Name | Alphanumeric | 30 | | | Student Name | Student IS | | | |

| ID | Integer | 9 | | | Student ID (SSN) | Student IS | | | |
|---|---|---|---|---|---|---|---|---|---|
| Major | Alphanumeric | 4 | | | Student Major | Student IS | | | |
| GPA | Decimal | 3 | 0.0 | 4.0 | Student Grade Point Avg | Academic Unit | | | |

**1-30. In the section "Disadvantages of File Processing Systems," the statement is made that the disadvantages of file processing systems can also be limitations of databases, depending on how an organization manages its databases. First, why do organizations create multiple databases, not just one all-inclusive database supporting all data processing needs? Second, what organizational and personal factors are at work that might lead an organization to have multiple, independently managed databases (and, hence, not completely follow the database approach)?**

**A-1-30.** There could be multiple factors for organization to use multiple databases:

1. **Data management and grouping**: Sometimes the organization creates multiple databases, and each database is containing data serving a particular purpose or functionality. For e.g., in E-Comm system, organization might want to keep order database separated from product database or customer database.
2. **Access control:** To limit the access of any user over the data, organization might want to keep separate database.
3. **Backup:** Organization might want to keep a copy of databases as backup.

Factors leading organizations to have multiple databases:

1. **Multi-party application to control access**: In a large organization, there could be multiple parties using the data for further smaller applications which does not require the entire data. Also, to ensure there is no data leak from any application, smaller applications are given the access over data they require so that we can maintain control over the access of data. This can lead to maintain multiple databases.
2. **Separating production data from QA data**: Majority of the applications have production data completely insulated from the QA or test data to ensure data authenticity. QA data are multiple copies of production data that are used to development and testing purpose. It leads to multiple database version serving particular purpose.
3. **Security of sensitive data**: To maintain highly sensitive data separated from all the other data, organization might create multiple databases.

**1-38. Consider your business school or other academic unit as a business enterprise.**
   a. **Define several major data entity types and draw a preliminary enterprise data model (similar in notation to Figure 1-3a).**

**A-1-38.** Draw.io ERD

## University

| PK | University ID int NOT NULL |
|----|----|
| | university_name char(50) NOT NULL |
| | Campus_location varchar NOT NULL |

## College

| PK | College_id int NOT NULL |
|----|----|
| FK1 | university_id int NOT NULL |

has a

has a

## Department

| PK | Derpartment_id int NOT NULL |
|----|----|
| FK1 | college_id int NOT NULL |
| | |

offers

## Professor

| PK | professor_id int NOT NULL |
|----|----|
| FK1 | department_id int NOT NULL |
| | |

## Courses

| PK | course_id int NOT NULL |
|----|----|
| FK1 | department_id int NOT NULL |

teaches

## Student

| PK | student_id int NOT NULL |
|----|----|
| FK1 | college_id int NOT NULL |
| | |

Enrolled