Week 1

*Limitations of Files*

* Sequential access: inefficient for large files
* Non-concurrent access: files cannot be updated by multiple people simultaneously
* Cannot span disks: all data in a file is present on a single disk
* Security: Open files have all data visible to the user

*Limitations of DBMSs*

* Resources: Use memory and disk space
* Cost: Generally expensive software
* Overheads: Using a DBMS is impractical for small files and single-user data

*Advantages of databases*

* Security: designer has control over what users can access
* Reliability:
* Easy data administration: Centralised repository of data and query languages allow quick updating/deleting/etc. of all data
* Large data capacity: Databases can store a lot of data
* Powerful querying: Users are able to make any sort of query that they could need
* Transaction: Changes made to the database are ACID: atomic, consistent, isolated and durable
* Concurrent access: Allows multiple users to access simultaneously
* Efficient access: Allows users to access the data they need quickly
* Data independence: Databases are multi-layered – if a certain type of information is changed, it does not affect other related information.

*DBMS Environment*

* Users include:
  + End-users: customers, employees – anyone who needs to use the database
  + Application developers: create the programs that use the database
  + Database designers: create the database in the application domain, physical/logical design including data requirements
  + Database administrators: maintain the resources necessary for an optimised database, create security, backup and recovery procedures

*Database Design Process*

* Requirements analysis: Fact-finding data, services (inserts, updates & queries) and business rules
  + Use questionnaires, interviews, observation, documentation, etc.
* Conceptual database design: Model requirements with Enhanced Entity-Relationship Model
* Logical database design: Convert to Relational Schema and remove redundancies with normalisation (1st, 2nd, 3rd and Boyce-Codd)
* Physical database design: The model is implemented using a DBMS, considering storage and access