**Lesson Summary and Revision Notes: Week 12**

Describe the limitations of a file-based approach for storage and retrieval of data.

* Inconsistency in data format from different applications as the data format depends on the software being used
* The same information being kept in several different places (files) and this leads to redundancy.
* Data inconsistency, a situation where various copies of the same data are conflicting, wastes storage space and duplicates effort.

Briefly explain what is a Relational Database

Relational database (RDB): a way of structuring information in tables, rows, and columns.

* A RDB has the ability to establish links - or relationships–between information by joining tables.
* This makes it easy to understand and gain insights about the relationship between various data points.

Describe the benefits of using a database-approach for storage and retrieval of data

* **Data Integrity**: Database approach to storage of files can ensure data integrity. This means data is consistent and accurate in the database. It is essential as there are multiple databases in a DBMS that contains data that is visible to multiple users.
* **Data Security**: Database approach to storage can help to ensure data security. Only users can access the database and their identity must be authenticated using username and password. A DBMS (database management system) provides a better platform for data privacy thus helping companies to offer an improved data security.
* **Minimized Data Inconsistency**: Data inconsistency occurs between files when various versions of the same data appear in different places. Data consistency is ensured in the database; there is no data redundancy. Besides, any database changes are immediately reflected by all users, and there is no data inconsistency.

**Database Terminology**







Primary Key: A primary key is a field in a table which uniquely identifies each row/record in a database table. Primary keys must contain unique values. A primary key column cannot have NULL values.

Secondary Key: Secondary Key is the key that has not been selected to be the primary key. However, it is considered a candidate key for the primary key. Therefore, a candidate key not selected as a primary key is called secondary key.

Foreign Key: A foreign key is a key used to link two tables together. This is sometimes also called as a referencing key. A Foreign Key is a column or a combination of columns whose values match a Primary Key in a different table.

**Entity Relationship**

Relationships can take several forms:

**one-to-one**, 1:1 (eg: a person has only one passport and a passport is given to one person.)



**one-to-many**, 1:m (eg: a customer can place many orders but a order cannot be placed by many customers)



**many-to-one**, m:1 (eg: many students can study in a single college but a student cannot study in many colleges at the same time)



**many-to-many**, m:m (eg: a student can be assigned to many projects and a project can be assigned to many students.)



Define What is Database Normalization

* A Normalization is a database design technique that reduces data redundancy and eliminates undesirable characteristics.
* Normalization rules divides larger tables into smaller tables and links them using relationships.

What is the Purpose of Database Normalization

* The purpose of Normalisation in to eliminate redundant (repetitive) data and ensure data is stored logically.
* Tables that are not normalised will be larger. As more data is stored, it will be harder to update the database when changes are made and more difficult to extract the required data to answer queries.