Ch6 : Data Normalization

→ **Normalization**

The process of decomposing bad relations by breaking up their attributes into smaller relations (well-structured relations)

→ **Well-structured relations**

A relation that contains minimal data redundancy and allows users to insert, delete, and update rows without causing data inconsistencies.

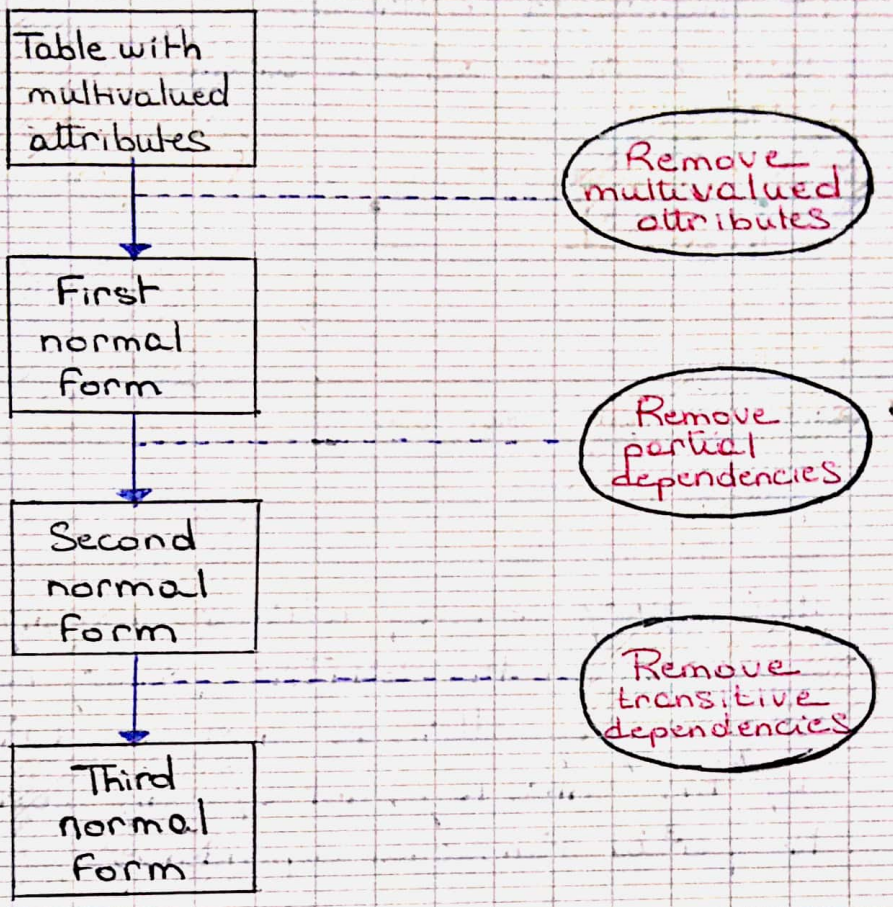
↳ Goal is to avoid anomalies:

→ Insertion Anomaly: adding new rows forces user to create duplicate data.

→ Deletion Anomaly: deleting rows may cause a loss of data that would be needed for future rows

→ Modification Anomaly: Changing data in a row forces changes to other rows because of duplication.

→ Steps in normalization



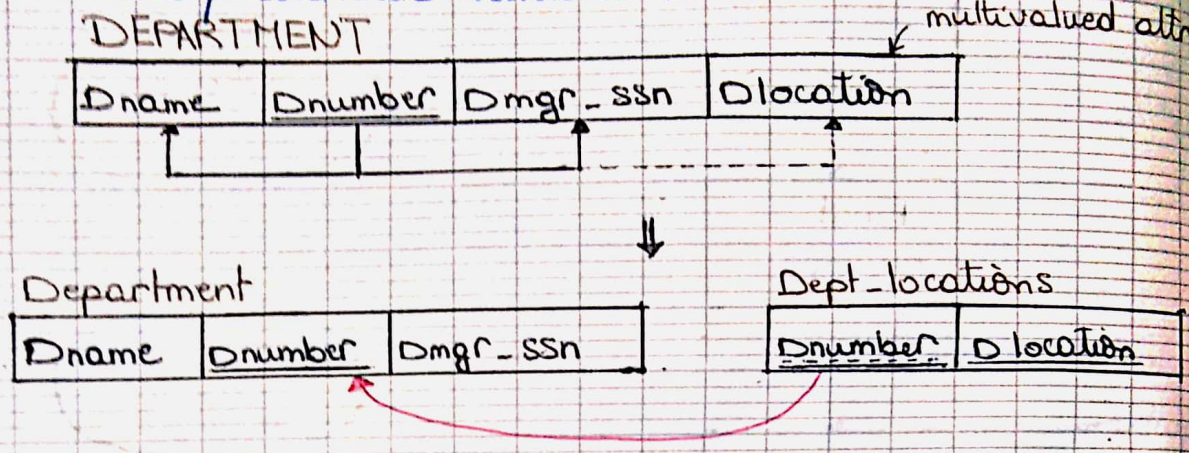
→ Functional Dependencies

The value of one attribute (the determinant) determine the value of another attribute.

هي علاقة تربط بين Attribute و آخر في نفس الجدول.
من خلال normalization سوف نجمع الـ Attributes التي يوجد بينهم Functional Dependencies و نخرجهم في relation منفصلة.

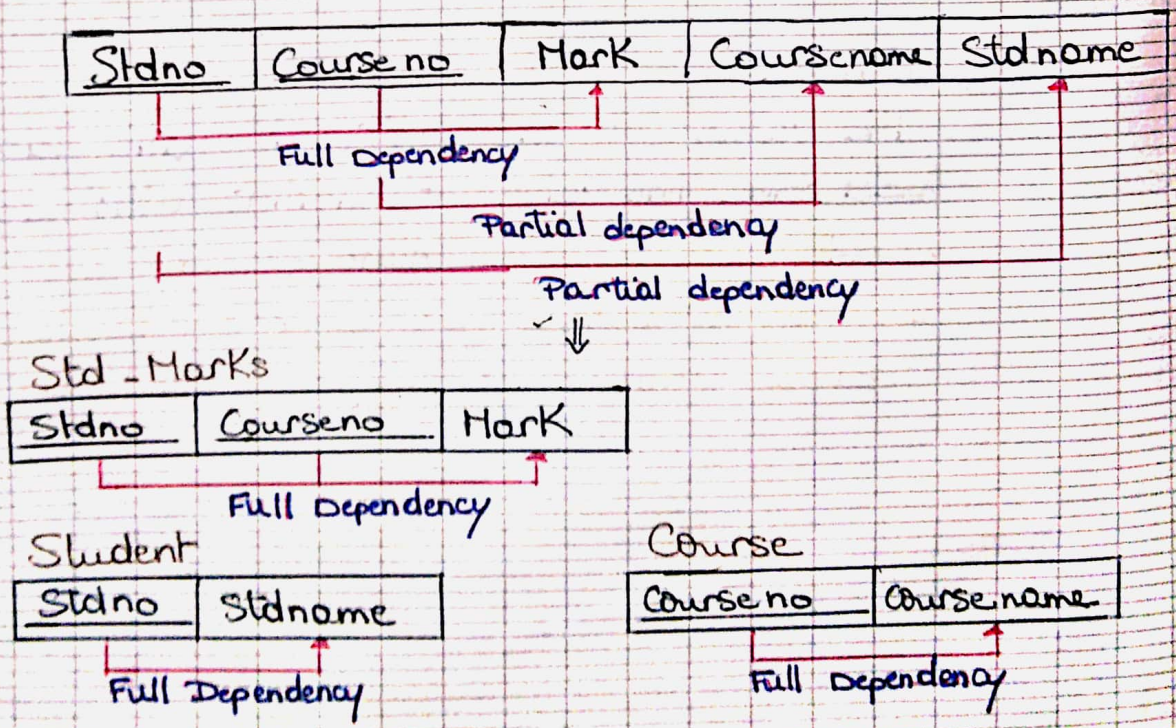
→ First Normal Form

- No multivalued attributes
- Every attribute value is atomic



→ Second Normal Form

- INF + every non_Key attribute is fully functionally dependent on the ENTIRE primary Key
- Every non_Key attribute must be defined by the entire Key, not by only part of the Key
- No partial functional dependencies.



→ Third Normal Form

. 2NF + no transitive dependencies (Functional dependencies on non-primary-Key attribute)

. This is called transitive, because the primary Key is a determinant for another attribute wich in turn is a determinant for a third (non-Key attribute dependant on another non-key attr)

. Solution: Non-Key determinant with transitive dependencies go into a new table; non-Key determinant becomes primary Key in the new relation and stay as foreign Key in the old relation.