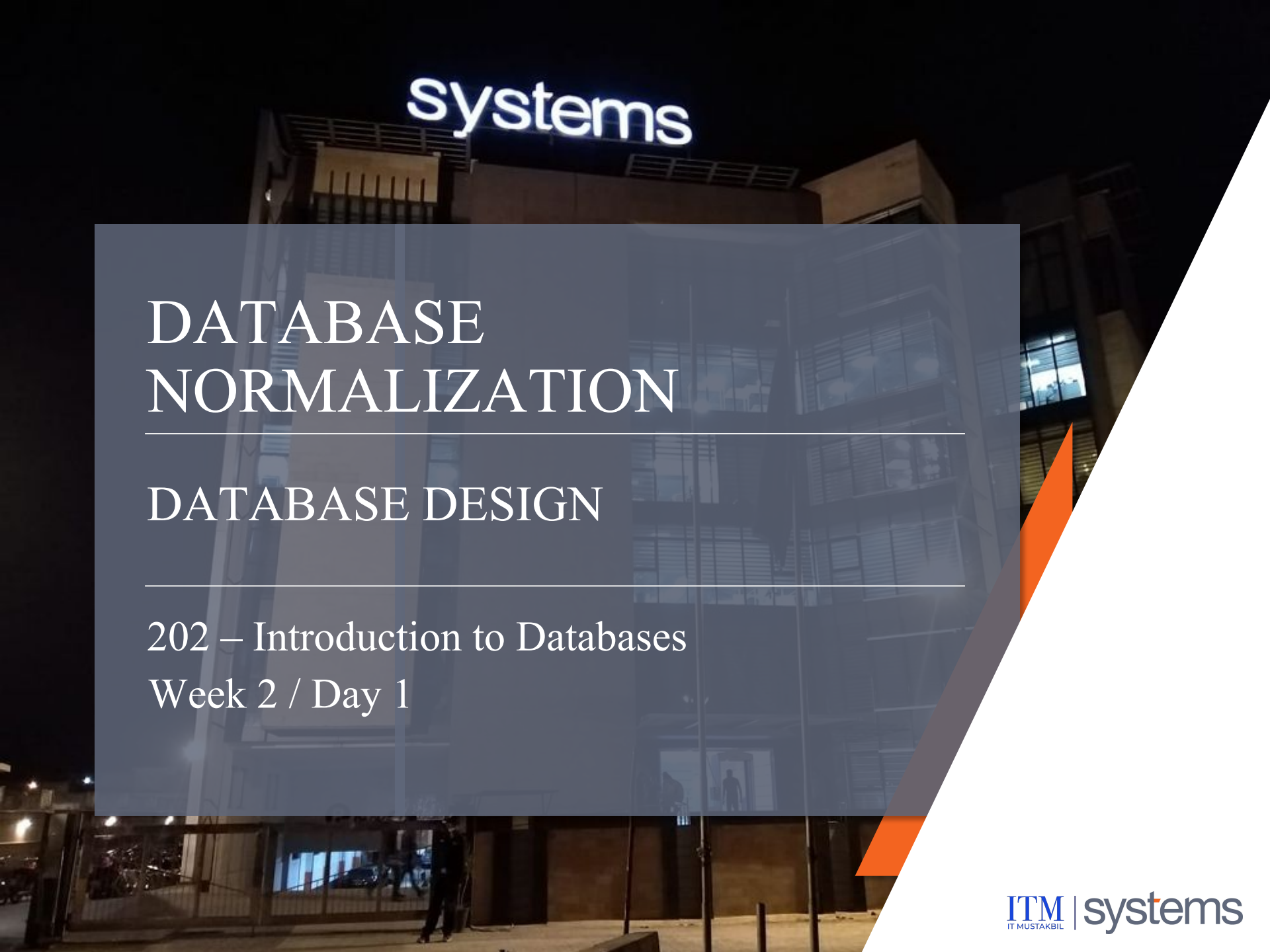


DATABASE NORMALIZATION

DATABASE DESIGN

202 – Introduction to Databases

Week 2 / Day 1

ITM | systems
IT MUSTAKBIL

LEARNING OBJECTIVES & AGENDA

Learning Objectives:

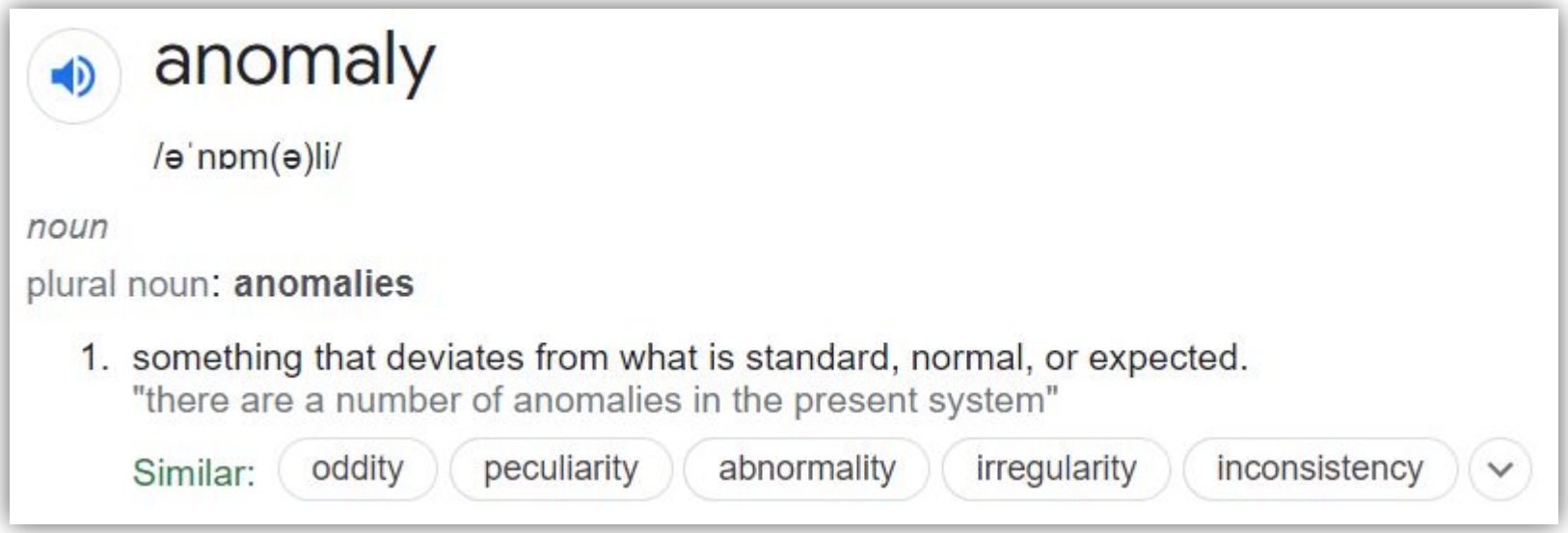
- Understand problems with insert / update / delete anomalies.
- Understand attribute functional dependencies in relations.
- Normalize relations using loss-less decomposition strategies.

Agenda / Sub-Topics:

- Database Anomalies
- Database Normalization
- 1NF
- 2NF
- 3NF

DATABASE ANOMALIES

WHAT ARE ANOMALIES?



anomaly

/əˈnɒm(ə)li/

noun

plural noun: **anomalies**

1. something that deviates from what is standard, normal, or expected.
   "there are a number of anomalies in the present system"

   Similar: oddity | peculiarity | abnormality | irregularity | inconsistency

WHAT ARE ANOMALIES?

- **Database anomalies** are faults in the database that usually emerge as a result of poor planning and storing everything in a flat database.
- Presence of anomalies means database does not accurately reflect real-world needs.
- Anomalies are generally reduced using normalization techniques.

TYPES OF ANOMALIES

- **Insert anomaly** – Inability to add data in the absence of other data.
- **Delete anomaly** – Inability to delete data without deleting other data.
- **Update anomaly** – Data updated in some places while remaining redundant copies not updated, leading to inconsistencies.

INSERT ANOMALY

roll_no	first_name	last_name	mobile	…	club
1111	Ammar	Khan	03212345678	…	Football
1111	Ammar	Khan	03337418529	…	Cricket
2222	Taha	Zafar	03451122334	…	Basket ball
3333	Ibrahim	Jafri	03018956237	…	Cricket
4444	Abdullah	Khawaja	03037485961		

- A school has created a policy that every student must be part of at least 1 extra-curricular club activity from G9 onwards.
- A new student, Abdullah Khawaja, has joined the school in G9. He has not yet selected an extra-curricular club.
- How do we add his record?

DELETE ANOMALY

roll_no	first_name	last_name	mobile	…	club
1111	Ammar	Khan	03212345678	…	Football
1111	Ammar	Khan	03337418529	…	Cricket
2222	Taha	Zafar	03451122334	…	Basket ball
3333	Ibrahim	Jafri	03018956237	…	Cricket
4444	Abdullah	Khawaja	03037485961		

- There is only one student in the basket ball club. The school's administration decides to discontinue this club.
- How do we only delete basket ball from the record?

UPDATE ANOMALY

roll_no	first_name	last_name	mobile	…	club
1111	Ammar	Khan	03212345678	…	Football
1111	Ammar	Khan	03337418529	…	Cricket
2222	Taha	Zafar	03451122334	…	Basket ball
3333	Ibrahim	Jafri	03018956237	…	Cricket
4444	Abdullah	Khawaja	03037485961		

- Ammar's mobile number has changed and asks the school administration to update the number.
- The school administration may only update the first record, without updating the second record, leading to inconsistent data.

DATABASE NORMALIZATION

DATABASE NORMALIZATION IS …

- A database design technique that reduces data redundancy
- Eliminates undesirable characteristics like Insertion, Update and Deletion Anomalies.
- Normalization rules divides larger tables into smaller tables and links them using relationships.
- The purpose of Normalization in SQL is to eliminate redundant (repetitive) data and ensure data is stored logically.

DATABASE NORMALIZATION IS …

- The primary goals of database normalization are:
  1. Eliminating Redundancy: By dividing data into separate tables and linking them through relationships, you avoid storing the same information multiple times. This reduces the storage space required and minimizes the risk of data inconsistencies or anomalies.
  2. Ensuring Data Integrity: Normalization helps maintain data accuracy and integrity by enforcing constraints and rules that prevent incorrect or conflicting data from being entered into the database.

1NF

FIRST NORMAL FORM (1NF) CONDITIONS

- Each column in a table must contain atomic (indivisible) values.
- Each row must have a unique identifier, usually a primary key.

2NF

SECOND NORMAL FORM (2NF) CONDITIONS

- Everything in previous normal forms.
- All non-key attributes must be functionally dependent on the entire primary key.

FUNCTIONAL DEPENDENCIES

- A functional dependency refers to the ability of one attribute being able to completely and uniquely identify another.
- E.g., student_name is functionally dependent on the student_rollno.
- This is because student_name can be completely and uniquely identified using the student_rollno.

FULLY FUNCTIONAL DEPENDENCY

- The functional dependency is a relationship that exists between two attributes. It typically exists between the primary key and non-key attribute within a table.

$$X \rightarrow Y$$

- **Determinant** The left side of FD is known as a determinant.
- **Dependent** The right side of the production is known as a dependent.

FULLY FUNCTIONAL DEPENDENCY

Roll_N0	**Name**	**Marks**
011	Ammar Khan	70
012	Abdullah	35
013	Ali	50
014	Ayesha	80

PARTIAL FUNCTIONAL DEPENDENCY

- Partial Dependency occurs when a non-key attribute is functionally dependent on part of a candidate key.
- Partial dependencies are when one of the primary keys determines another attribute or attributes
- The 2nd Normal Form (2NF) eliminates the Partial Dependency.

PARTIAL FUNCTIONAL DEPENDENCY EXAMPLE

<STUDENTPROJECT>

Student_ID	**Project_No**	**Student_Name**	**Project_Name**
S01	199	Ammar	Geo Location
S02	120	Aamna	Cluster Exploration

PARTIAL FUNCTIONAL DEPENDENCY EXAMPLE

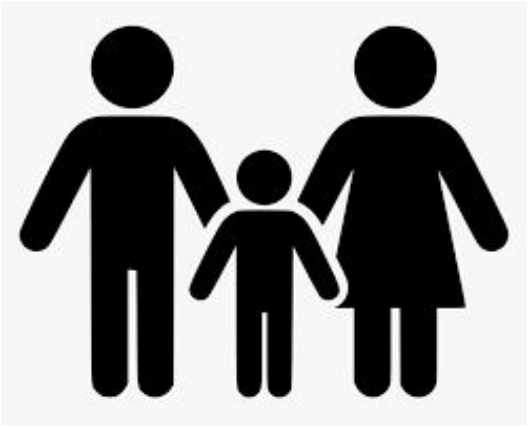
<StudentInfo>

StudentID	**ProjectNo**	**StudentName**
S01	199	Ammar
S02	120	Aamna

<ProjectInfo>

ProjectNo	**ProjectName**
199	Geo Location
120	Cluster Exploration

FULLY FUNCTIONAL DEPENDENCY VS PARTIAL FUNCTIONAL DEPENDENCY



Baby partially depends on his Mother.
Baby Partially depends on his father.

His father died. Sorry!

Baby fully functional depends on mother

3NF

THIRD NORMAL FORM (3NF) CONDITIONS

- Everything in previous normal forms.
- There should be no transitive dependencies.
  - Transitive dependencies occur when a non-key attribute depends on another non-key attribute.

TRANSITIVE FUNCTIONAL DEPENDENCY

- When an indirect relationship causes functional dependency it is called Transitive Dependency.
- A transitive dependency occurs when one non-key attribute is dependent on another non-key attribute.

To achieve 3NF, eliminate the Transitive Dependency.

TRANSITIVE DEPENDENCY EXAMPLE

<MovieListing>

Movie_ID	**Listing_ID**	**Listing_Type**	**DVD_Price ($)**
M08	L09	Crime	180
M03	L05	Drama	250
M05	L09	Crime	180

TRANSITIVE DEPENDENCY EXAMPLE

<Movie>

Movie_ID	DVD_Price ($)
L09	180
L05	250
L09	180

<Listing>

Listing_ID	Listing_Type
L09	Crime
L05	Drama
L09	Crime

NORMALIZATION

NORMALIZATION

- A database design technique that reduces data redundancy
- Eliminates undesirable characteristics like Insertion, Update and Deletion Anomalies.
- Normalization rules divides larger tables into smaller tables and links them using relationships.
- The purpose of Normalization in SQL is to eliminate redundant (repetitive) data and ensure data is stored logically.

1NF (FIRST NORMAL FORM)

- Each table cell should contain a single value.
- Each record needs to be unique.

EMP_ID	EMP_NAME	EMP_PHONE	EMP_CITY
14	Danish	7272826385, 9064738238	Islamabad
20	Irfan	8574783832	Lahore
12	Aamir	7390372389, 8589830302	Multan

1NF (FIRST NORMAL FORM)

EMP_ID	EMP_NAME	EMP_PHONE	EMP_CITY
14	Danish	7272826385	Islamabad
14	Danish	9064738238	Islamabad
20	Ammar	8574783832	Lahore
12	Aamir	7390372389	Multan
12	Aamir	8589830302	Multan

2NF (SECOND NORMAL FORM)

- In the 2NF, relational must be in 1NF.
- In the second normal form, all non-key attributes are fully functional dependent on the primary key (No partial dependency).

Employee No	**Department No**	**Employee Name**	**Department**
1	101	Ali	OBIEE
2	102	Ayesha	COGNOS
3	101	Eisha	OBIEE

2NF (SECOND NORMAL FORM)

<Employee>

Employee_No	Department_No	Employee_Name
1	101	Ali
2	102	Ayesha
3	101	Esha

<Department>

Department_No	Department
101	OBIEE
102	COGNOS

3NF (THIRD NORMAL FORM)

- In the 3NF, relational must be in 2NF.
- There is no transitive functional dependency.

Employee_No	**SalarySlip_No**	**Employee_Name**	**Salary**
1	0001	Aman	50000
2	0002	Irfan	40000
3	0003	Usama	57000

3NF (THIRD NORMAL FORM)

<Employee>

Employee_No	SalarySlip_No	Employee_Name
1	0001	Aman
2	0002	Irfan
3	0003	Usama

<Department>

SalarySlip_No	Salary
0001	50000
0002	40000
0003	57000

ADDITIONAL RESOURCES

- FUNCTIONAL DEPENDENCIES

https://opentextbc.ca/dbdesign01/chapter/chapter-11-functional-dependencies/

- FUNCTIONAL DEPENDENCY VS PARTIAL DEPENDENCY

https://www.geeksforgeeks.org/differentiate-between-partial-dependency-and-fully-functional-dependency/