

ISA

Geographic Information Systems 2026/2027

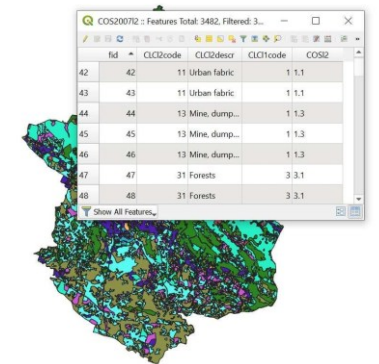
Lesson 04 - Organising geographic data



Topics

- Relational Database principles
- The 3rd normal form
- Structured Query Language (SQL)

GIS are data repositories



Geographic Information Systems - 2026/2027 - Lesson 04 - 3

Relational Databases and GIS

- GIS store **geographic objects** and **non-geographic objects**, their characteristics and relationships among them
- Non-geometric characteristics can be stored in GIS and in **databases** using the same techniques/methods
 - Database techniques/methodologies already exist for about 50 years
 - Using already proven and tested techniques/methods in thousands of Information Systems and improved during decades has many advantages
 - Moreover, some GIS functionalities were developed based on widely adopted database principles
 - **Relational Database**

Geographic Information Systems - 2026/2027 - Lesson 04 - 4

Relational Databases Principles

- Structure and Organization

Fundamental to create efficient & effective databases

Focuses on data arrangement & interconnection

Geographic Information Systems - 2026/2027 - Lesson 04 - 5

Relational Databases Principles

- Structure and Organization
 - Relation (table)
 - represents an entity or a concept in the database
 - data is organized into relations (aka tables) consisting of tuples (rows) and attributes (columns) defined by a **schema**
 - Keys
 - Primary keys: unique identifiers of one relation
 - Foreign keys: reference in one relation to a primary key of another relation
 - Relationships
 - Allows organizing and connecting relations through their attributes
 - Relations can be connected by distinct types of relationships

Geographic Information Systems - 2026/2027 - Lesson 04 - 6

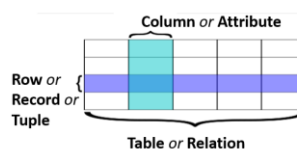
Relation

- Relations are intended to represent objects and/or relations among objects
 - Objects **sharing common characteristics** are represented using **one given relation** – every relation represents objects with a common object type
 - Objects with **different characteristics** are represented using **different relations**
 - In the GIS context, **relation**, **theme** and **layer** can be considered synonymous

Geographic Information Systems - 2026/2027 - Lesson 04 - 7

Relation

- **Relation** is the basic concept of relational databases
- A **relational database** is a set of relations
- A relation is defined by one **schema**
 - serves as a formal description of how the data is organized within the table (consistency and integrity)
 - outlines the logical structure of the table without containing the data itself
- The schema typically includes:
 - Relation name
 - Attribute names (column)
 - Data types for each attribute
 - Constraints (e.g., primary keys, foreign keys, unicity constraints)
 - Relationships with other tables



Geographic Information Systems - 2026/2027 - Lesson 04 - 8

Table and schema - examples

Table

fid	geom	CLC2code	CLC2descr	CLC1code	COSI2
1	Polygon	51	Inland waters	5	5.1
2	Polygon	51	Inland waters	5	5.1
3	Polygon	51	Inland waters	5	5.1
4	Polygon	51	Inland waters	5	5.1
5	Polygon	11	Urban fabric	1	1.1
6	Polygon	11	Urban fabric	1	1.1
7	Polygon	11	Urban fabric	1	1.1
8	Polygon	11	Urban fabric	1	1.1
9	Polygon	11	Urban fabric	1	1.1
10	Polygon	11	Urban fabric	1	1.1
11	Polygon	11	Urban fabric	1	1.1

Schema

Visible	Read Only	Field Name	Alias	Data Type	Allow NULL	Highlight	Number Format	D
☒	☒	fid	fid	Object ID	☐	☐	Numeric	
☒	☐	geom	geom	Geometry	☒	☐		
☒	☐	CLC2code	CLC2code	Long	☒	☐	Numeric	
☒	☐	CLC2descr	CLC2descr	Text	☒	☐		
☒	☐	CLC1code	CLC1code	Long	☒	☐	Numeric	
☒	☐	COSI2	COSI2	Text	☒	☐		

Geographic Information Systems - 2026/2027 - Lesson 04 - 9

Table

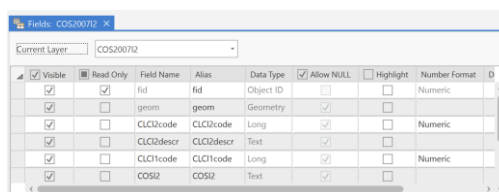
fid	geom	CLC2code	CLC2descr	CLC1code	COSI2
1	Polygon	51	Inland waters	5	5.1
2	Polygon	51	Inland waters	5	5.1
3	Polygon	51	Inland waters	5	5.1
4	Polygon	51	Inland waters	5	5.1
5	Polygon	11	Urban fabric	1	1.1
6	Polygon	11	Urban fabric	1	1.1
7	Polygon	11	Urban fabric	1	1.1
8	Polygon	11	Urban fabric	1	1.1
9	Polygon	11	Urban fabric	1	1.1
10	Polygon	11	Urban fabric	1	1.1
11	Polygon	11	Urban fabric	1	1.1

- One **table** is a set of **attribute** values structured into rows and columns
 - every **object** is described by the values in one **row**
 - every **attribute** (**field**) corresponds to one **column**
- one table **cannot** contain equal rows
- objects may have some attributes without value (**optional attribute**)
 - when an optional attribute does not have value, means its value is **null**
- the **identifiers** (names) of attributes must be unique for each relation schema

Geographic Information Systems - 2026/2027 - Lesson 04 - 10

Attributes

- An **attribute** represents a characteristic of an object
 - values describe one characteristic of a particular object
- Values of one attribute must belong to the same **domain**
- the domain is defined by its **datatype**
 - the domain specifies the set of values that the attribute can take
 - An attribute A_i takes values in a set Domain D_i



Geographic Information Systems - 2026/2027 - Lesson 04 - 11

Attributes - data types

Geopackage	ArcGIS table
• Whole number (integer) • Whole number (integer 64 bits) • Decimal number (real) • Text (string) • JSON (string) – JavaScript Object Notation • Date • Date & Time • Binary object (BLOB) • Boolean	• Short integer (-32,768 to 32,767) • Long integer (-2,147,483,648 to 2,147,483,647) • Float (approximately -3.4E38 to 1.2E38) • Double (approximately -2.2E308 to 1.8E308) • Text (string) • JSON (string) – JavaScript Object Notation • Date (dates, times, or dates and times) • Binary object (BLOB)

Geographic Information Systems - 2026/2027 - Lesson 04 - 12

Relational Databases Principles

- Structure and Organization
 - Relation (table)
 - represents an entity or a concept in the database
 - data is organized into relations (aka tables) consisting of tuples (rows) and attributes (columns) defined by a schema
 - Keys
 - **Primary keys:** unique identifiers of one relation
 - **Foreign keys:** reference in one relation to a primary key of another relation
 - Relationships
 - Allows organizing and connecting relations through their attributes
 - Relations can be connected by distinct types of relationships

Geographic Information Systems - 2026/2027 - Lesson 04 - 13

Primary key of a relation

- **Attribute** with unique identifiers for each tuple (row) in a relation (table)
- Characteristics include:
 - **Uniqueness:** each value in the primary key column must be distinct (no 2 records have the same ID)
 - the value of a primary key uniquely specifies a particular object in a relation
 - **Non-null:** every record has an identifier (cannot contain null values)
 - **Immutability:** Once assigned, primary key values should not be changed to maintain data integrity
 - **Single or composite:** can consist of a single or multiple attributes (composite key)
 - composite primary key is only required when there is not one attribute where each value is unique
 - **Indexing:** automatically indexed, improving query performance
 - **Referential integrity:** serve as the basis for foreign key relationships with other relations

Geographic Information Systems - 2026/2027 - Lesson 04 - 14

Primary key of a relation

- A candidate key of a relation is defined as the **minimal set of attributes** where each value is unique and identifies a single object (or row) of the relation (or table)
 - sometimes natural attributes are good primary keys
 - often, an artificial attribute is added to the objects in order to identify each object uniquely
 - useful to uniquely identify every instance of the relation
 - this kind of keys have no intrinsic meaning
 - for instance, in a table of data about students at a school, they might all be assigned a student identification number
- One table may have several **candidate keys**, but one and only one primary key must be defined

Geographic Information Systems - 2026/2027 - Lesson 04 - 15

Foreign key of a relation

- A foreign key (FK) is an **Attribute** in one relation (table) that references the primary key of another relation
- Creates a connection (**relationship**) between relations
- Characteristics include:
 - **Referential integrity**: values in the foreign key column correspond to existing values in the referenced table's primary key
 - **Relationship enforcement**: establish and maintain links between related relations
 - **Cascading actions**: may allow automatically update or delete related records when the referenced primary key is modified
 - **Null values**: can often accept null values, allowing for optional relationships
 - **Multiple occurrences**: a relation can have multiple foreign keys, referencing different relations
 - **Composite foreign keys**: can also consist of multiple columns

Geographic Information Systems - 2026/2027 - Lesson 04 - 16

Relational Databases Principles

- Structure and Organization
 - Relation (table)
 - represents an entity or a concept in the database
 - data is organized into relations (aka tables) consisting of tuples (rows) and attributes (columns) defined by a schema
 - Keys
 - Primary keys: unique identifiers of one relation
 - Foreign keys: reference in one relation to a primary key of another relation
 - Relationship
 - Allows organizing and connecting relations through their attributes
 - Relations can be connected by distinct types of relationships

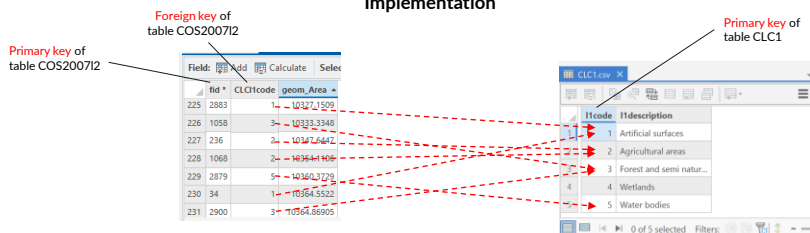
Geographic Information Systems - 2026/2027 - Lesson 04 - 17

Relationships



- 1:1 (one to one)
 - Each parcel has only one owner
 - Each class code of the Corine Land Cover classification has one and only one description and vice versa (applied to all classification levels)

Implementation

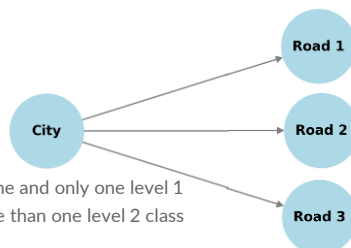


Geographic Information Systems - 2026/2027 - Lesson 04 - 19

Relationships

- 1:M (one to many)

- Each city may have many roads
- Each level 2 class of the CLC classification belongs to one and only one level 1 class but every level 1 class can be subdivided into more than one level 2 class



Implementation

Classes		
Class Name	Credit Hours	Class ID
Introduction to Jazz	4	101
Introduction to Guitar	5	102
Advanced Guitar	4	103
Jazz Band	4	104

Students		
Student ID	Student Name	Date of Birth
1001	Wayne Gentry	3/5/2005
1002	Cathy Baker	10/2/2005
1003	Zayan Dean	12/3/2004
1004	Isa Midead	4/18/2005
1005	Just Wilks	6/4/2005
1006	Sally Driscoll	10/3/2004



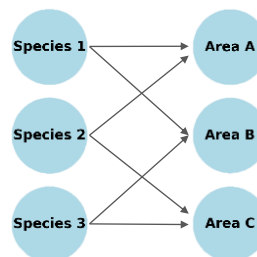
Enrollments			
Class ID	Student ID	Class Grade	Enrollment ID
101	1001	B+	100001
101	1003	A-	100002
101	1004	B	100003
101	1005	A+	100004
102	1002	C+	100005
102	1004	A+	100006
102	1005	C	100007
103	1002	B	100008
103	1002	B-	100009
104	1002	A	100010
104	1005	A-	100011
104	1006	B+	100012

Geographic Information Systems - 2026/2027 - Lesson 04 - 21

Relationships

- M:N (many to many)

- One river can cross more than 1 municipality and 1 municipality can be crossed by more than one river



Implementation

River

fid	CODRIOS	DESIGNACAO	TIPO
6054	626	Ribeira da Apostoia	Principal
883	701	Rio Mondego	Principal
1744	719	Rio Vouga	Principal
2163	719.01	Vala Corujeira	Principal
2041	719.03	Rio Boco	Principal

CODRIOS	dico
719	0101
719	0102
9719	0102
719	0105
9719	0105

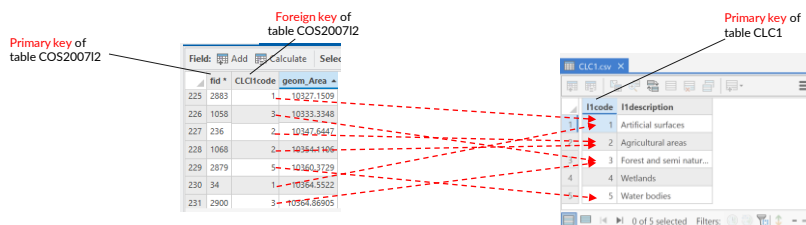
Municipality

fid	Concelho	dico
1292	Águeda	0101
1303	Albergaria-A-Velha	0102
1556	Anadia	0103
210	Arouca	0104
1282	Aveiro	0105

Geographic Information Systems - 2026/2027 - Lesson 04 - 23

Primary keys, Foreign Keys & Relationships

- On relational DBMS, **relationships** among **entities** are represented in **relations** (tables) using **attributes** containing common values, i.e., **primary keys** and **foreign keys**



- foreign key is an attribute (or a set of attributes) in a relation matching the primary key of another relation

Remark: geopackage file format always include a primary key named fid – this attribute cannot be deleted and users cannot change its values (it is protected)

Geographic Information Systems – 2026/2027 - Lesson 04 - 24

Text notation to represent the schema of relations

`<TABLE>(<ATTRIB01>, <ATTRIB02>, ... , <ATTRIB_n>)`

Summary:

- The name outside the parentheses is the name of the relation (layer in the GIS context)
- The attribute names are inside parentheses
- The primary key is underlined attribute
- The foreign key is the dash underlined attribute(s)
- A composite primary key is represented by more than one underlined attributes

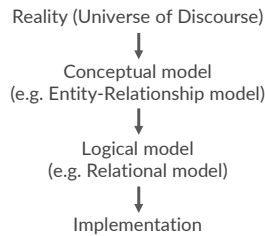
Examples:

- CLC1 (l1code , l1description)
- COS2007I2 (fid , CLC1code , CLC2descr , CLC1code , COSI2)
- River (fid , codrios , designacao , tipo)
- Municipality (fid , concelho , dico)
- RiversInMunic (codrios , dico)

Geographic Information Systems – 2026/2027 - Lesson 04 - 26

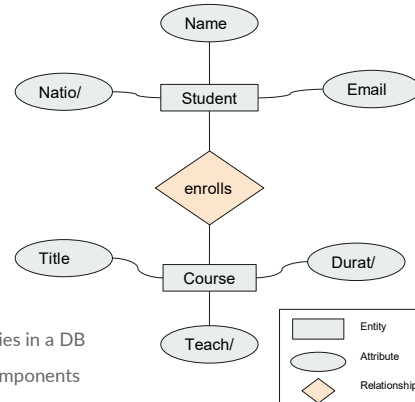
Diagrammatic representation of Databases

Process of database development



- visual representation of the relationships between entities in a DB
- help understanding the structure and identifying key components (primary keys, foreign keys, relationship cardinality)

Entity-Relationship Diagram (ERD)



Geographic Information Systems - 2026/2027 - Lesson 04 - 27

Relational Databases Principles

- Structure and Organization
- Data Integrity and Consistency

Fundamental to ensure data remains accurate, reliable & consistent throughout its lifecycle

Represented by the ACID properties: Atomicity, Consistency, Isolation & Durability

Geographic Information Systems - 2026/2027 - Lesson 04 - 28

Relational Databases Principles

- Structure and Organization
- Data Integrity and Consistency
 - Atomicity: ensures operations are treated as a single, indivisible unit of work
 - A logical group of operations (transaction) must either be completed entirely or not at all
 - Prevents partial updates and avoids the DB to be in a inconsistent state

Geographic Information Systems - 2026/2027 - Lesson 04 - 29

Relational Databases Principles

- Structure and Organization
- Data Integrity and Consistency
 - Atomicity: ensures operations are treated as a single, indivisible unit of work
 - Consistency: maintains logical consistency of data across relations & relationships
 - A set of logical operations must take the database from one valid state to another
 - Ensures all predefined rules and constraints are maintained

Geographic Information Systems - 2026/2027 - Lesson 04 - 30

Relational Databases Principles

- Structure and Organization
- Data Integrity and Consistency
 - Atomicity: ensures operations are treated as a single, indivisible unit of work
 - Consistency: maintains logical consistency of data across relations & relationships
 - Isolation: guarantees concurrent execution leads to equal results as sequential execution
 - ensures concurrent transactions do not interfere with each other
 - maintains data consistency in multi-user environments

Geographic Information Systems - 2026/2027 - Lesson 04 - 31

Relational Databases Principles

- Structure and Organization
- Data Integrity and Consistency
 - Atomicity: ensures operations are treated as a single, indivisible unit of work
 - Consistency: maintains logical consistency of data across relations & relationships
 - Isolation: guarantees concurrent execution leads to equal results as sequential execution
 - Durability: prevents data losses even in the event of power loss, crashes, or errors
 - Ensures committed transactions are permanently recorded
 - Uses logs and recovery techniques to maintain data integrity after system failures

Geographic Information Systems - 2026/2027 - Lesson 04 - 32

Relational Databases Principles

- Structure and Organization
- Data Integrity and Consistency
- Design Principles

Fundamental guidelines that ensure efficient, consistent, and reliable database systems

Geographic Information Systems - 2026/2027 - Lesson 04 - 33

Relational Databases Principles

- Structure and Organization
- Data Integrity and Consistency
- Design Principles
 - Normalization:
 - systematic approach to organizing data
 - reduces redundancy, increases consistency, improves integrity & performance
 - divide data into subject-based tables, respective attributes & establish relationships between tables
 - Progress through various normal forms (1NF, 2NF, 3NF)

Geographic Information Systems - 2026/2027 - Lesson 04 - 34

Normalization

Database normalization is the process of **organizing the attributes and tables** of a relational database to **minimize redundancy & avoid dependencies** among attributes

- Objective:
isolate data so that additions and deletions of rows, and modifications of an attribute value can be made in just one table and then propagated through the rest of the database via the implemented relationships
- make it easier for data updates and keeping the database consistent with reality (trustable)
- Relational database normalization uses the so-called “**normal forms**”

Online Resource: <https://tinyurl.com/ISA-norm>
<https://youtu.be/J-drts33N8g>

Geographic Information Systems - 2026/2027 - Lesson 04 - 36

Normal Forms

- There are 7 normal forms
- Commonly, for relational databases normalization only goes up to the 3rd normal form (3NF)
- Discover & study dependencies within tables to avoid redundant information & anomalies

Geographic Information Systems - 2026/2027 - Lesson 04 - 37

Normal Forms

- There are 7 normal forms
- Commonly, for relational databases normalization only goes up to the 3rd normal form (3NF)
- Discover & study dependencies within tables to avoid redundant information & anomalies

AgriData

FarmID	Crops	Farmer	Equip	EquipMaker	EquipYear
F01	Corn, Soybeans	John Smith	Tractor, Plough	John Deere	2020, 2018
F02	Wheat	Jane Doe	Harvester	Case IH	2019
F01	Corn, wheat	John Smith	Sprayer	John Deere	2021
F03	Soybeans, rice	Oliver Duke	Tractor, Sprayer	Case IH, John Deere	2021, 2022



Unnormalized

- Attribute values are tuples with multiple values

Geographic Information Systems - 2026/2027 - Lesson 04 - 38

First Normal Form

- No duplicate rows (**unique** rows)
- No repeating columns
- Each cell only contains a single value
- Values in a cell must be non-divisible
- Attribute values must be in the same domain

atomic

AgriData

FarmID	Crops	Farmer	Equip	EquipMaker	EquipYear
F01	Corn, Soybeans	John Smith	Tractor, Plough	John Deere	2020, 2018
F02	Wheat	Jane Doe	Harvester	Case IH	2019
F01	Corn, wheat	John Smith	Sprayer	John Deere	2021
F03	Soybeans, rice	Oliver Duke	Tractor, Sprayer	Case IH, John Deere	2021, 2022

Unnormalized

- Attribute values are tuples with multiple values

Geographic Information Systems - 2026/2027 - Lesson 04 - 39

First Normal Form

AgriData

FarmID	Crop	Prod_tons	Farmer	Equip	EquipMaker	EquipYear
F01	Corn	5	John Smith	Tractor	John Deere	2020
F01	Soybeans	8	John Smith	Plough	John Deere	2018
F02	Wheat	10	Jane Doe	Harvester	Case IH	2019
F01	Corn	10	John Smith	Sprayer	John Deere	2021
F01	Wheat	6	John Smith	Sprayer	John Deere	2021
F03	Soybeans	9	Oliver Duke	Tractor	Case IH	2021
F03	Rice	15	Oliver Duke	Sprayer	John Deere	2022

- No duplicate rows (**unique** rows)
- No repeating columns
- Each cell only contains a single value
- Values in a cell must be non-divisible
- Attribute values must be in the same domain

atomic

Geographic Information Systems - 2026/2027 - Lesson 04 - 40

First Normal Form

AgriData

FarmID	Crop	Prod_tons	Farmer	Equip	EquipMaker	EquipYear
F01	Corn	5	John Smith	Tractor	John Deere	2020
F01	Soybeans	8	John Smith	Plough	John Deere	2018
F02	Wheat	10	Jane Doe	Harvester	Case IH	2019
F01	Corn	10	John Smith	Sprayer	John Deere	2021
F01	Wheat	6	John Smith	Sprayer	John Deere	2021
F03	Soybeans	9	Oliver Duke	Tractor	Case IH	2021
F03	Rice	15	Oliver Duke	Sprayer	John Deere	2022

- No duplicate rows (**unique** rows)
- No repeating columns
- Each cell only contains a single value
- Values in a cell must be non-divisible
- Attribute values must be in the same domain

atomic

1NF ✓

Geographic Information Systems - 2026/2027 - Lesson 04 - 41

First Normal Form

AgriData

FarmID	Crop	Prod_tons	Farmer	Equip	EquipMaker	EquipYear
F01	Corn	5	John Smith	Tractor	John Deere	2020
F01	Soybeans	8	John Smith	Plough	John Deere	2018
F02	Wheat	10	Jane Doe	Harvester	Case IH	2019
F01	Corn	10	John Smith	Sprayer	John Deere	2021
F01	Wheat	6	John Smith	Sprayer	John Deere	2021
F03	Soybeans	9	Oliver Duke	Tractor	Case IH	2021
F03	Rice	15	Oliver Duke	Sprayer	John Deere	2022

- No duplicate rows (**unique** rows)
- No repeating columns
- Each cell only contains a single value
- Values in a cell must be non-divisible
- Attribute values must be in the same domain

1NF ✓

But...

creates redundancies

Geographic Information Systems - 2026/2027 - Lesson 04 - 42

Second Normal Form

Production

FarmID	Crop	Prod_tons	Farmer	Equip	EquipMaker	EquipYear
F01	Corn	5	John Smith	Tractor	John Deere	2020
F01	Soybeans	8	John Smith	Plough	John Deere	2018
F02	Wheat	10	Jane Doe	Harvester	Case IH	2019
F01	Corn	10	John Smith	Sprayer	John Deere	2021
F01	Wheat	6	John Smith	Sprayer	John Deere	2021
F03	Soybeans	9	Oliver Duke	Tractor	Case IH	2021
F03	Rice	15	Oliver Duke	Sprayer	John Deere	2022

1NF

- No duplicate rows (**unique** rows)
- No repeating columns
- Each cell only contains a single value
- Values in a cell must be non-divisible
- Attribute values must be in the same domain

- Table must be in 1NF
- No partial dependencies
(all non-key attributes must be functionally dependent on the candidate primary key; only about 1 type of entity)

Geographic Information Systems - 2026/2027 - Lesson 04 - 43

Second Normal Form

Production

FarmID	Crop	Prod_tons	Farmer	Equip	EquipMaker	EquipYear
F01	Corn	5	John Smith	Tractor	John Deere	2020
F01	Soybeans	8	John Smith	Plough	John Deere	2018
F02	Wheat	10	Jane Doe	Harvester	Case IH	2019
F01	Corn	10	John Smith	Sprayer	John Deere	2021
F01	Wheat	6	John Smith	Sprayer	John Deere	2021
F03	Soybeans	9	Oliver Duke	Tractor	Case IH	2021
F03	Rice	15	Oliver Duke	Sprayer	John Deere	2022

FarmID + Crop form a candidate key

1NF

- No duplicate rows (**unique** rows)
- No repeating columns
- Each cell only contains a single value
- Values in a cell must be non-divisible
- Attribute values must be in the same domain

atomic

- Table must be in 1NF
- No partial dependencies
(all non-key attributes must be functionally dependent on the candidate primary key; only about 1 type of entity)

Geographic Information Systems - 2026/2027 - Lesson 04 - 44

Second Normal Form

Production

FarmID	Crop	Prod_tons	Farmer	Equip	EquipMaker	EquipYear
F01	Corn	5	John Smith	Tractor	John Deere	2020
F01	Soybeans	8	John Smith	Plough	John Deere	2018
F02	Wheat	10	Jane Doe	Harvester	Case IH	2019
F01	Corn	10	John Smith	Sprayer	John Deere	2021
F01	Wheat	6	John Smith	Sprayer	John Deere	2021
F03	Soybeans	9	Oliver Duke	Tractor	Case IH	2021
F03	Rice	15	Oliver Duke	Sprayer	John Deere	2022

FarmID + Crop form a candidate key

1NF

- No duplicate rows (**unique** rows)
- No repeating columns
- Each cell only contains a single value
- Values in a cell must be non-divisible
- Attribute values must be in the same domain

atomic

- Table must be in 1NF
- No partial dependencies
(all non-key attributes must be functionally dependent on the candidate primary key; only about 1 type of entity)

PARTIAL DEPENDENCIES

Geographic Information Systems - 2026/2027 - Lesson 04 - 45

Regions

Second Normal Form

Farm_Crops

FarmCropID	FarmID	CropID	Prod_tons
FC01	F01	C01	5
FC02	F01	C02	8
FC03	F02	C03	10
FC04	F01	C01	10
FC05	F01	C03	6
FC06	F03	C02	9
FC07	F03	C04	15

Crops

CropID	CropName
C01	Corn
C02	Soybeans
C03	Wheat
C04	Rice

Farms

FarmID	Farmer
F01	John Smith
F02	Jane Doe
F03	Oliver Duke

Farm_equip

FarmID	Equip	EquipMaker	Equip Year
F01	Tractor	John Deere	2020
F01	Plough	John Deere	2018
F01	Sprayer	John Deere	2021
F02	Harvester	Case IH	2019
F03	Tractor	Case IH	2021
F03	Sprayer	John Deere	2022

1NF

- No duplicate rows (**unique** rows)
- No repeating columns
- Each cell only contains a single value
- Values in a cell must be non-divisible
- Attribute values must be in the same domain

2NF

- Table must be in 1NF
- No partial dependencies

(all non-key attributes must be functionally dependent on the candidate primary key; only about 1 type of entity)

Geographic Information Systems - 2026/2027 - Lesson 04 - 46

Regions

Second Normal Form

Farm_Crops

FarmCropID	FarmID	CropID	Prod_tons
FC01	F01	C01	5
FC02	F01	C02	8
FC03	F02	C03	10
FC04	F01	C01	10
FC05	F01	C03	6
FC06	F03	C02	9
FC07	F03	C04	15

Crops

CropID	CropName
C01	Corn
C02	Soybeans
C03	Wheat
C04	Rice

Farms

FarmID	Farmer
F01	John Smith
F02	Jane Doe
F03	Oliver Duke

Farm_equip

FarmID	Equip	EquipMaker	Equip Year
F01	Tractor	John Deere	2020
F01	Plough	John Deere	2018
F01	Sprayer	John Deere	2021
F02	Harvester	Case IH	2019
F03	Tractor	Case IH	2021
F03	Sprayer	John Deere	2022

1NF

- No duplicate rows (**unique** rows)
- No repeating columns
- Each cell only contains a single value
- Values in a cell must be non-divisible
- Attribute values must be in the same domain

2NF

- Table must be in 1NF
- No partial dependencies

(all non-key attributes must be functionally dependent on the candidate primary key; only about 1 type of entity)

FULL DEPENDENCIES

Geographic Information Systems - 2026/2027 - Lesson 04 - 47

Regions

Second Normal Form

Farm_Crops

FarmID	FarmID	CropID	Prod_tons
FC01	F01	C01	5
FC02	F01	C02	8
FC03	F02	C03	10
FC04	F01	C01	10
FC05	F01	C03	6
FC06	F03	C02	9
FC07	F03	C04	15

Crops

CropID	CropName
C01	Corn
C02	Soybeans
C03	Wheat
C04	Rice

Farms

FarmID	Farmer
F01	John Smith
F02	Jane Doe
F03	Oliver Duke

Farm_equipments

FarmID	Equip	EquipMaker	Equip Year
F01	Tractor	John Deere	2020
F01	Plough	John Deere	2018
F01	Sprayer	John Deere	2021
F02	Harvester	Case IH	2019
F03	Tractor	Case IH	2021
F03	Sprayer	John Deere	2022

1NF

- No duplicate rows (**unique** rows)
- No repeating columns
- Each cell only contains a single value
- Values in a cell must be non-divisible
- Attribute values must be in the same domain

2NF

- Table must be in 1NF
- No partial dependencies
(all non-key attributes must be functionally dependent on the candidate primary key; only about 1 type of entity)

FULL DEPENDENCIES

2NF ✓

Geographic Information Systems - 2026/2027 - Lesson 04 - 48

Third Normal Form

Farm_Crops

FarmID	FarmID	CropID	Prod_tons
FC01	F01	C01	5
FC02	F01	C02	8
FC03	F02	C03	10
FC04	F01	C01	10
FC05	F01	C03	6
FC06	F03	C02	9
FC07	F03	C04	15

Crops

CropID	CropName
C01	Corn
C02	Soybeans
C03	Wheat
C04	Rice

Farms

FarmID	Farmer
F01	John Smith
F02	Jane Doe
F03	Oliver Duke

Farm_equipments

FarmID	Equip	EquipMaker	Equip Year
F01	Tractor	John Deere	2020
F01	Plough	John Deere	2018
F01	Sprayer	John Deere	2021
F02	Harvester	Case IH	2019
F03	Tractor	Case IH	2021
F03	Sprayer	John Deere	2022

1NF

- No duplicate rows (**unique** rows)
- No repeating columns
- Each cell only contains a single value
- Values in a cell must be non-divisible
- Attribute values must be in the same domain

2NF

- Table must be in 1NF
- No partial dependencies
(all non-key attributes must be functionally dependent on the candidate primary key; only about 1 type of entity)

3NF

- Table must be in 2NF
- No transitive dependencies
(no other non key attribute would need to change if a non-key attribute changes)

Geographic Information Systems - 2026/2027 - Lesson 04 - 49

Third Normal Form

Farm_Crops

FarmCropID	FarmID	CropID	Prod_tons
FC01	F01	C01	5
FC02	F01	C02	8
FC03	F02	C03	10
FC04	F01	C01	10
FC05	F01	C03	6
FC06	F03	C02	9
FC07	F03	C04	15

Crops

CropID	CropName
C01	Corn
C02	Soybeans
C03	Wheat
C04	Rice

*EquipmentMaker depends on Equip,
not directly on primary key*

Farms

FarmID	Farmer
F01	John Smith
F02	Jane Doe
F03	Oliver Duke

FarmID	Equip	EquipMaker	Equip Year
F01	Tractor	John Deere	2020
F01	Plough	John Deere	2018
F01	Sprayer	John Deere	2021
F02	Harvester	Case IH	2019
F03	Tractor	Case IH	2021
F03	Sprayer	John Deere	2022

1NF

- No duplicate rows (**unique** rows)
- No repeating columns
- Each cell only contains a single value } **atomic**
- Values in a cell must be non-divisible
- Attribute values must be in the same domain

2NF

- Table must be in 1NF
- No partial dependencies
(all non-key attributes must be functionally dependent on the candidate primary key; only about 1 type of entity)
- Table must be in 2NF
- No transitive dependencies
(no other non key attribute would need to change if a non-key attribute changes)

Geographic Information Systems - 2026/2027 - Lesson 04 - 50

Third Normal Form

Farm_Crops

FarmCropID	FarmID	CropID	Prod_tons
FC01	F01	C01	5
FC02	F01	C02	8
FC03	F02	C03	10
FC04	F01	C01	10
FC05	F01	C03	6
FC06	F03	C02	9
FC07	F03	C04	15

Crops

CropID	CropName
C01	Corn
C02	Soybeans
C03	Wheat
C04	Rice

Farms

FarmID	Farmer
F01	John Smith
F02	Jane Doe
F03	Oliver Duke

Equip

EquipID	Equip	EquipMaker	Equip Year
E01	Tractor	John Deere	2020
E02	Plough	John Deere	2018
E03	Sprayer	John Deere	2021
E04	Harvester	Case IH	2019
E05	Tractor	Case IH	2021
E06	Sprayer	John Deere	2022

Farm equipments

FarmEquipID	FarmID	EquipID
FE01	F01	E01
FE02	F01	E02
FE03	F02	E03
FE04	F01	E04
FE05	F03	E05
FE06	F03	E06

1NF

- No duplicate rows (**unique** rows)
- No repeating columns
- Each cell only contains a single value } **atomic**
- Values in a cell must be non-divisible
- Attribute values must be in the same domain

2NF

- Table must be in 1NF
- No partial dependencies
(all non-key attributes must be functionally dependent on the candidate primary key; only about 1 type of entity)
- Table must be in 2NF
- No transitive dependencies
(no other non key attribute would need to change if a non-key attribute changes)

3NF

Geographic Information Systems - 2026/2027 - Lesson 04 - 51

23

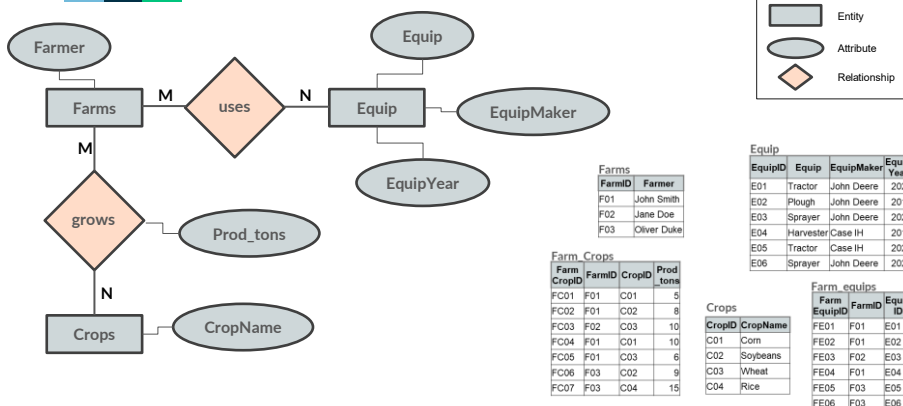
Normal Forms - wrap up

Discover & study dependencies within tables to avoid redundant information and resulting anomalies

- A table is in **First Normal Form (1NF)** if it is **atomic** and values in column are of the same **domain**
- A table is in **Second Normal Form (2NF)** if it is in **1NF** and does not have **partial** dependencies
- A table is in **Third Normal Form (3NF)** if it is in **2NF** and does not have **transitive** dependencies

Geographic Information Systems - 2026/2027 - Lesson 04 - 55

Entity-Relationship Diagram - Conceptual Model



Geographic Information Systems - 2026/2027 - Lesson 04 - 57

Exercise 1

<TABLE> (<ATTRIB01>, <ATTRIB02>, ... , <ATTRIB_n>)

Summary:

- The name outside the parentheses is the name of the relation (layer in the GIS context)
- The attribute names are inside parentheses
- The primary key is underlined attribute
- The foreign key is the dash underlined attribute(s)
- A composite primary key is represented by more than one underlined attributes

Farms
<u>FarmID</u> Farmer
F01 John Smith
F02 Jane Doe
F03 Oliver Duke

Farm Crops
<u>FarmCropID</u> FarmID CropID Prod_tons
FC01 F01 C01 5
FC02 F01 C02 8
FC03 F02 C03 10
FC04 F01 C01 10
FC05 F01 C03 6
FC06 F03 C02 9
FC07 F03 C04 15

Equip
<u>EquipID</u> Equip EquipMaker Equip Year
E01 Tractor John Deere 2020
E02 Plough John Deere 2018
E03 Sprayer John Deere 2021
E04 Harvester Case IH 2019
E05 Tractor Case IH 2021
E06 Sprayer John Deere 2022

Crops
<u>CropID</u> CropName
C01 Corn
C02 Soybeans
C03 Wheat
C04 Rice

Farm equipments
<u>Farm EquipID</u> FarmID Equip ID
FE01 F01 E01
FE02 F01 E02
FE03 F02 E03
FE04 F01 E04
FE05 F03 E05
FE06 F03 E06

1. Please elaborate the diagram of the relational model
2. then elaborate the text notation to represent the diagram of the relational model

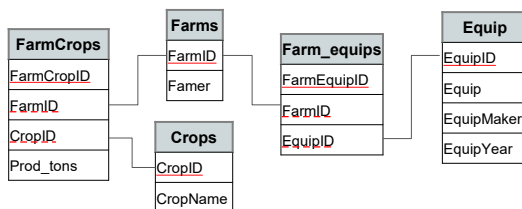
Geographic Information Systems - 2026/2027 - Lesson 04 - 58

Exercise 1

<TABLE> (<ATTRIB01>, <ATTRIB02>, ... , <ATTRIB_n>)

Summary:

- The name outside the parentheses is the name of the relation (layer in the GIS context)
- The attribute names are inside parentheses
- The primary key is underlined attribute
- The foreign key is the dash underlined attribute(s)
- A composite primary key is represented by more than one underlined attributes



Farms
<u>FarmID</u> Farmer
F01 John Smith
F02 Jane Doe
F03 Oliver Duke

Farm Crops
<u>FarmCropID</u> FarmID CropID Prod_tons
FC01 F01 C01 5
FC02 F01 C02 8
FC03 F02 C03 10
FC04 F01 C01 10
FC05 F01 C03 6
FC06 F03 C02 9
FC07 F03 C04 15

Equip
<u>EquipID</u> Equip EquipMaker Equip Year
E01 Tractor John Deere 2020
E02 Plough John Deere 2018
E03 Sprayer John Deere 2021
E04 Harvester Case IH 2019
E05 Tractor Case IH 2021
E06 Sprayer John Deere 2022

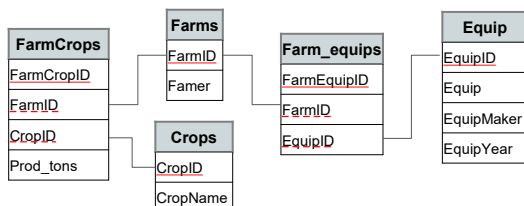
Crops
<u>CropID</u> CropName
C01 Corn
C02 Soybeans
C03 Wheat
C04 Rice

Farm equipments
<u>Farm EquipID</u> FarmID Equip ID
FE01 F01 E01
FE02 F01 E02
FE03 F02 E03
FE04 F01 E04
FE05 F03 E05
FE06 F03 E06

1. Please elaborate the diagram of the relational model
2. then elaborate the text notation to represent the diagram of the relational model

Geographic Information Systems - 2026/2027 - Lesson 04 - 60

Exercise 1



FarmCrops (FarmCropID, FarmID, CropID, Prod_tons)

Farms (FarmID, Farmer)

Farm_equips (FarmEquipID, FarmID, EquipID)

Equip (EquipID, Equip, EquipMaker, EquipYear)

1. Please elaborate the diagram of the relational model
2. Then elaborate the text notation to represent the diagram of the relational model

FarmID	Farmer
F01	John Smith
F02	Jane Doe
F03	Oliver Duke

EquipID	Equip	EquipMaker	EquipYear
E01	Tractor	John Deere	2020
E02	Plough	John Deere	2018
E03	Sprayer	John Deere	2021
E04	Harvester	Case IH	2019
E05	Tractor	Case IH	2021
E06	Sprayer	John Deere	2022

FarmCropID	FarmID	CropID	Prod_tons
FC01	F01	C01	5
FC02	F01	C02	8
FC03	F02	C03	10
FC04	F01	C01	10
FC05	F01	C03	6
FC06	F03	C02	9
FC07	F03	C04	15

CropID	CropName
C01	Com
C02	Soybeans
C03	Wheat
C04	Rice

FarmEquipID	FarmID	EquipID
FE01	F01	E01
FE02	F01	E02
FE03	F02	E03
FE04	F01	E04
FE05	F03	E05
FE06	F03	E06

Geographic Information Systems - 2026/2027 - Lesson 04 - 61

Text notation to represent the schema of relations

<TABLE>(<ATTRIB01>, <ATTRIB02>, ... , <ATTRIB_n>)

Summary:

- The name outside the parentheses is the name of the relation (layer in the GIS context)
- The attribute names are inside parentheses
- The primary key is underlined attribute
- The foreign key is the dash underlined attribute(s)
- A composite primary key is represented by more than one underlined attributes

Geographic Information Systems - 2026/2027 - Lesson 04 - 62

Relational Databases Principles

- Structure and Organization
- Data Integrity and Consistency
- Design Principles
 - Normalization: systematic approach to organizing data
 - Avoid Redundancy: Eliminate duplicate data to maintain data integrity and optimize storage
 - Referential Integrity: relationships between relations must remain consistent
 - foreign keys always referring to valid primary keys
 - Null Values: handle null values adequately (represent missing or unknown information)
 - Logical & Physical Independence: Separate logical data structures from physical storage
 - changes in physical storage should not affect the logical structure

Geographic Information Systems - 2026/2027 - Lesson 04 - 67

Exercise

Suppose that you have to design a database schema concerning remarkable trees in order to be stored as vector data in a GIS.

→ For each tree it is required to record its *common name*, *scientific name*, *family*, *height* and *trunk diameter*.

In the study area there is around one thousand remarkable trees and several species are represented by more than one tree.

What relation schema(s) best fits these requirements?

- A. Tree (scientificName, commonName, family, height, trunkDiameter)
- B. Tree (treeldNumber, commonName, scientificName, family, height, trunkDiameter)
- C. Tree (treeldNumber, scientificName, height, trunkDiameter)
Taxonomy (scientificName, commonName, family)

Geographic Information Systems - 2026/2027 - Lesson 04 - 68

Exercise

Suppose that you have to design a database schema concerning remarkable trees in order to be stored as vector data in a GIS.

→ For each tree it is required to record its *common name*, *scientific name*, *family*, *height* and *trunk diameter*.

In the study area there is around one thousand remarkable trees and several species are represented by more than one tree.

What relation schema(s) best fits these requirements?

- A. Tree (scientificName, commonName , family , height , trunkDiameter)
- B. Tree (treeldNumber, commonName , scientificName , family , height , trunkDiameter)
- C. Tree (treeldNumber, scientificName , height , trunkDiameter)
Taxonomy (scientificName, commonName, family)

Geographic Information Systems - 2026/2027 - Lesson 04 - 69

Credits

The structure and content of the course, including most of the slides and exercises was developed by Professor Graça Abrantes (SIG 2021_2022) and Professor Rui Figueira (SIG 2023_2024)

Geographic Information Systems - 2026/2027 - Lesson 01 - 70