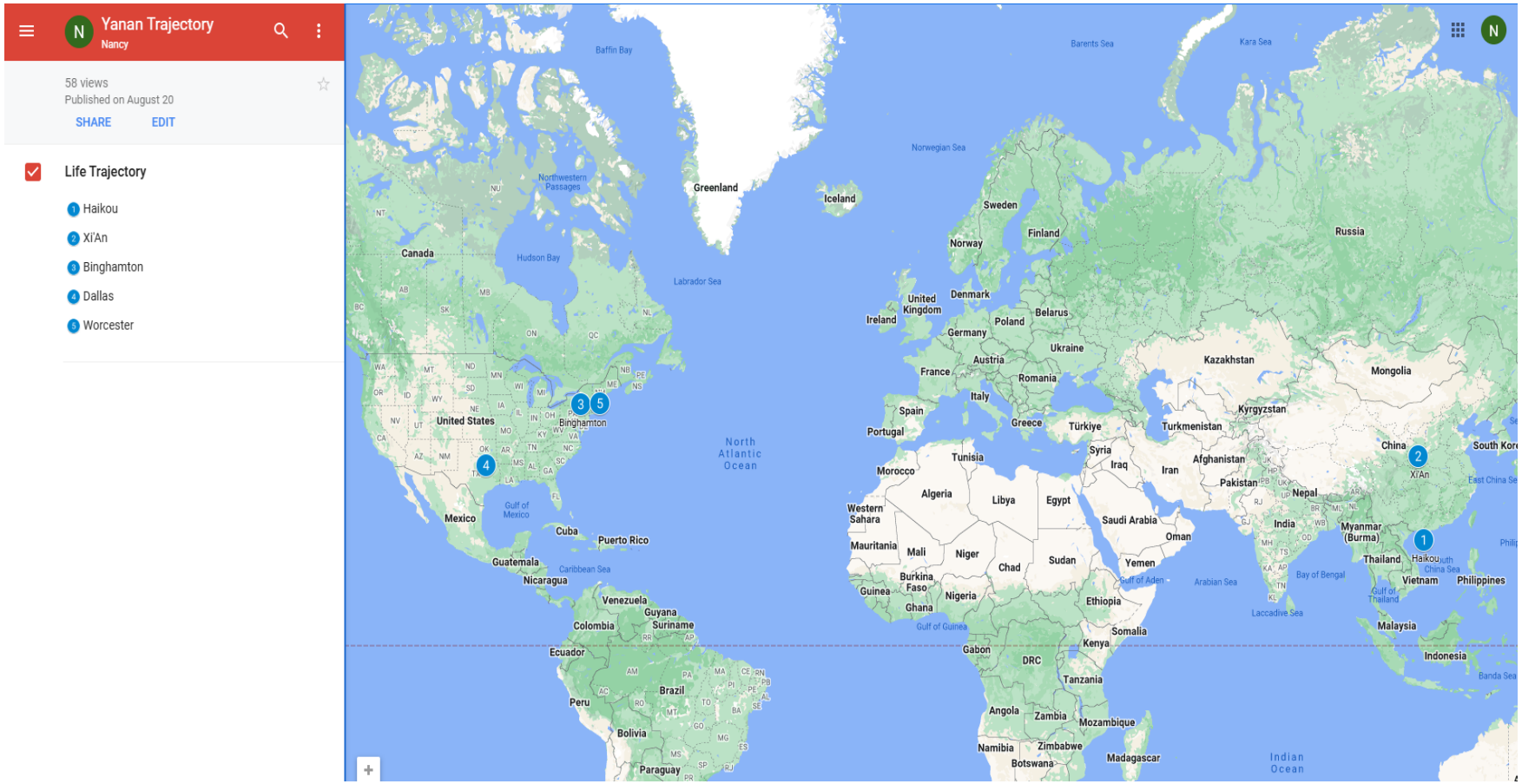
WEEK 01

Instructor: Yanan Wu
TA: Vanchy Li

Spring 2025

YANAN WU – VISITING ASSISTANT PROFESSOR

- Education & Experience



TEACHING

- Intermediate Statistics
- Python Programming in GIS
  - Manipulating spatial data
  - Web mapping
  - Processing Raster
  - Data Analysis
  - Creating Custom Tool
  - Data visualization
- Spatial Database

TA INTRODUCTION

- Office Hours:
- Office Location

HOW ABOUT YOU?

- Your background (e.g., name, major, where you come from)
- What is your funniest thing that happened during your winter break?
- What relevant experience do you have with statistics?
- What are your expectations for this course?

COURSE FORMAT

- Lectures: Instructor & TA
  - Location:
  - Time
- Labs: TA
  - Location
  - Time
- Office Hours
- Office Location

COURSE REQUIREMENTS

- Labs: 11 in total

  For any graded assignment, if the you do not agree with the grade received, the instructor and TA must be notified within one week after the assignment is graded.

- Late policy for lab (excluding midterm and final project)
- One midterm exam & One final project (oral presentation and paper report)

GRADE

WEEK 01

LECTURE SESSION

Instructor: Yanan Wu
TA: Vanchy Li

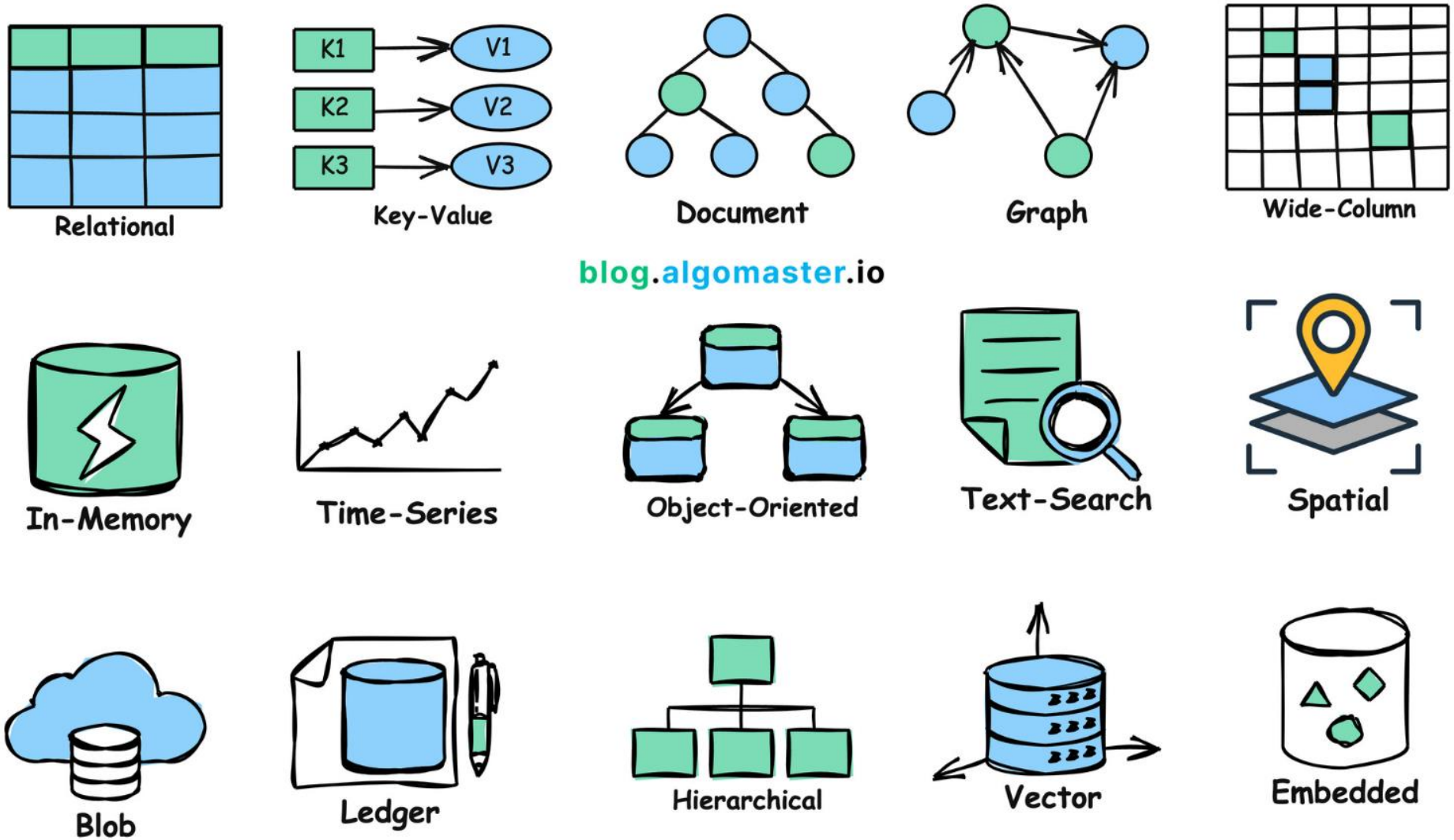
Spring 2025

DATABASE

- A database is an organized collection of structured data, or information, typically stored electronically in a computer that can be easily accessed, managed, and updated

Source

TYPES OF DATABASE



Source

DATABASE MANAGEMENT SYSTEM (DBMS)

- A software system that is designed to manage and organize data in a structured manner.

Key features of DBMS

Store large amounts of data systematically

Implements access control to protect data from unauthorized access

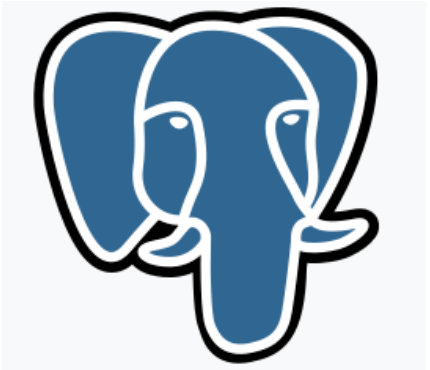
Ensure the consistency of data through constraints

Provide tools to back up the database

OPEN-SOURCE TECHNOLOGIES

- MySQL

- PostgreSQL



- MongoDB

- SQLite

- DynamoDB

- MariaDB

KEY ATTRIBUTES OF A ROBUST DBMS

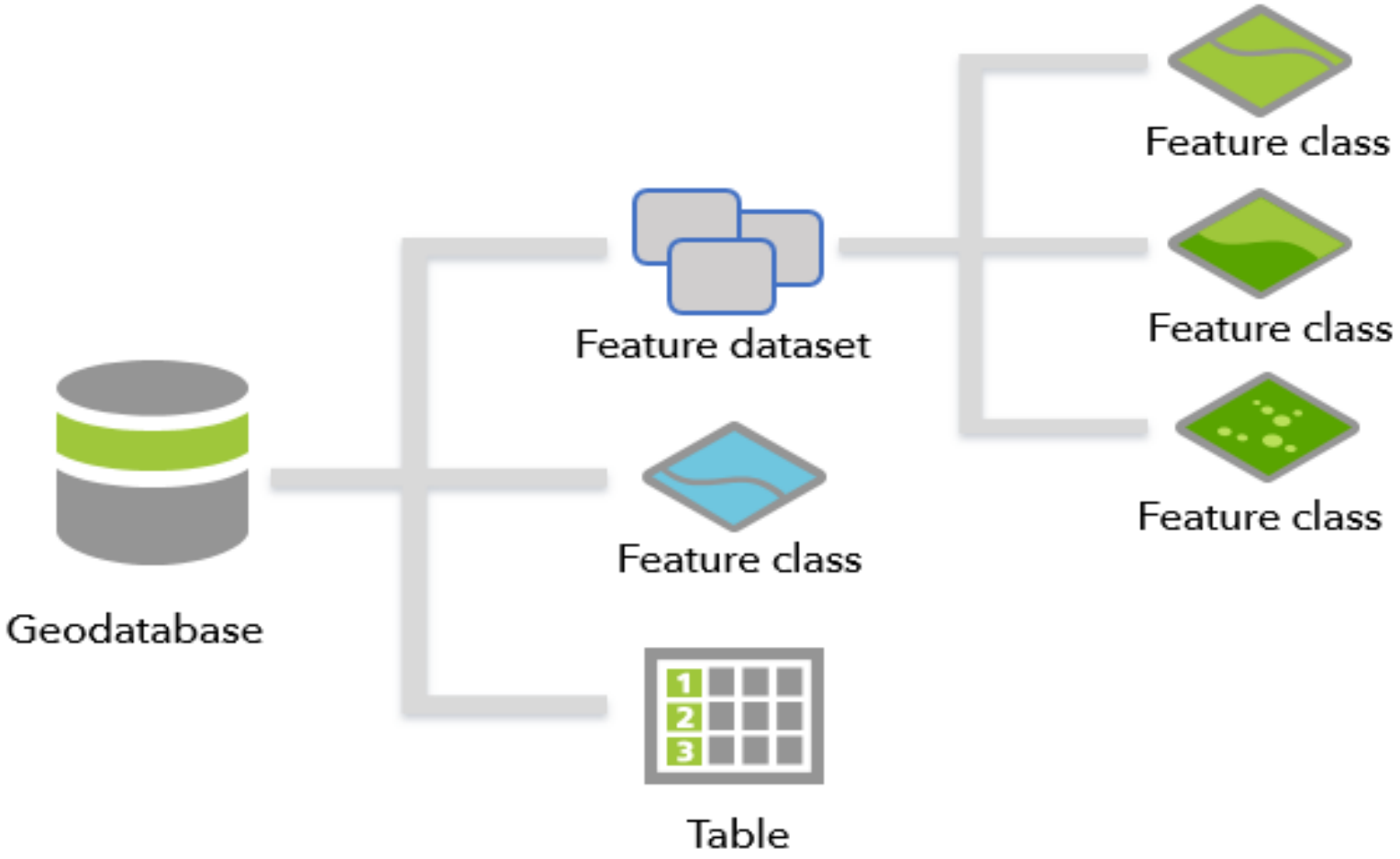
- Reliability
  - The system can function correctly and consistently over time
  - The system must be able to offer continual uninterrupted service
- Consistency
  - Refers to ensuring that the database remains accurate, valid, and follows defined rules or constraints
- Technology-Proof
  - The system is adapted to evolving technologies and its long-term viability

GEODATABASE

- A **Geodatabase** is a specialized type of database designed to store, manage, and analyze spatial and geographic data.

- It serves as a comprehensive framework for organizing spatial data and associated attribute information, allowing users to perform advanced geographic analysis, mapping, and data management tasks.

GEODATABASE IN GIS

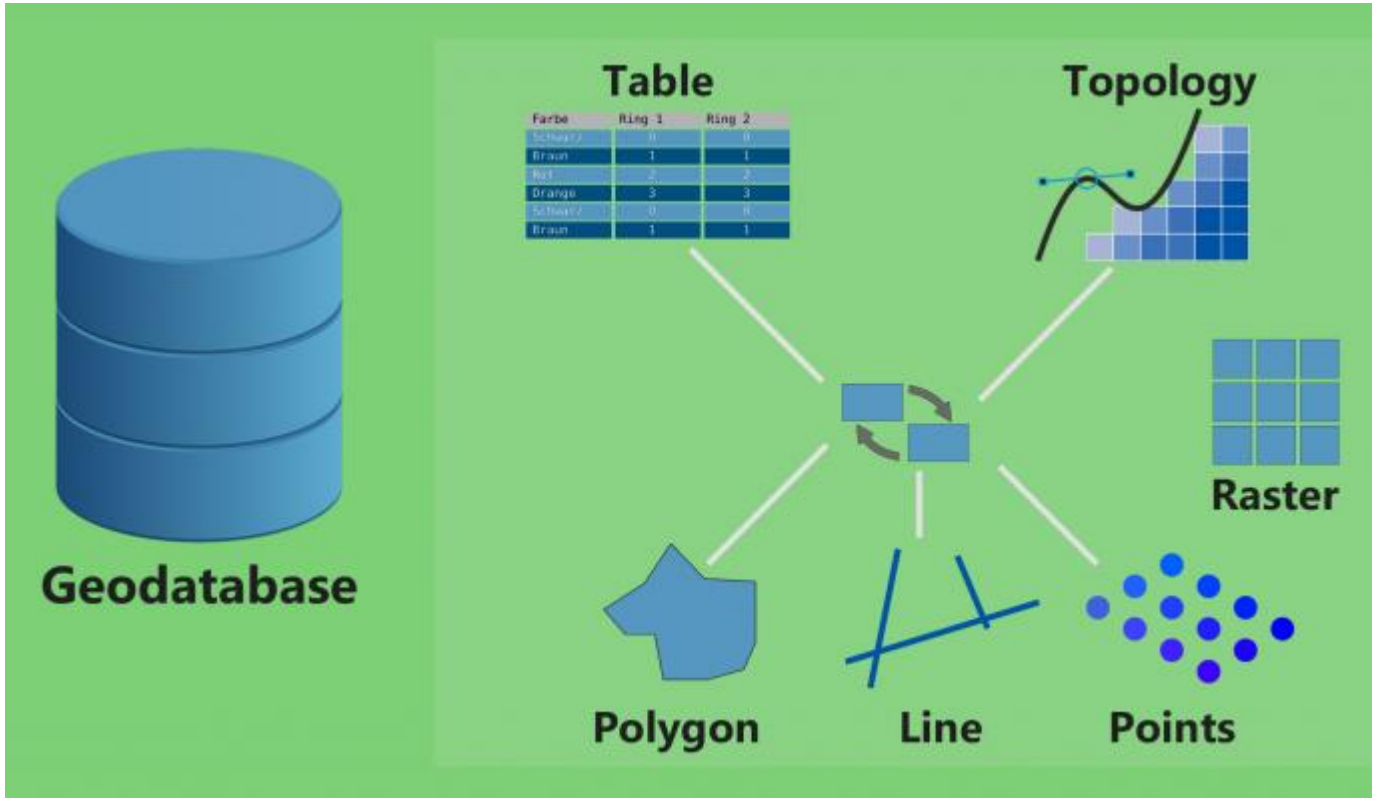


ADVANTAGES OF USING A GEODATABASE

- Centralized Data Management

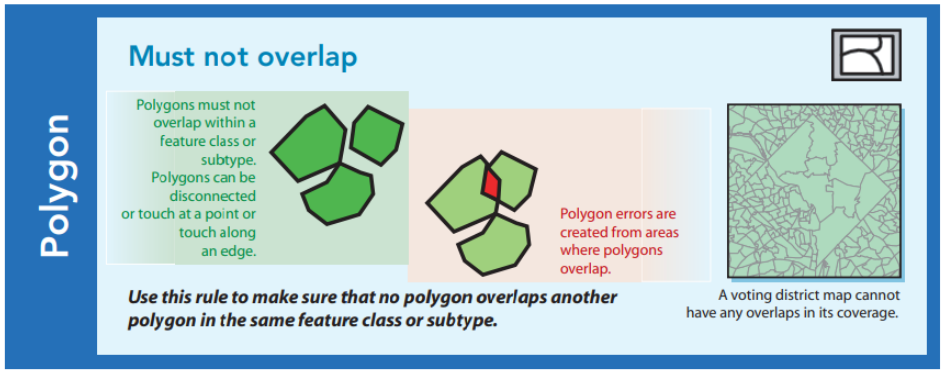
Combines spatial (maps, shapes) and non-spatial (tabular) data in a single database.

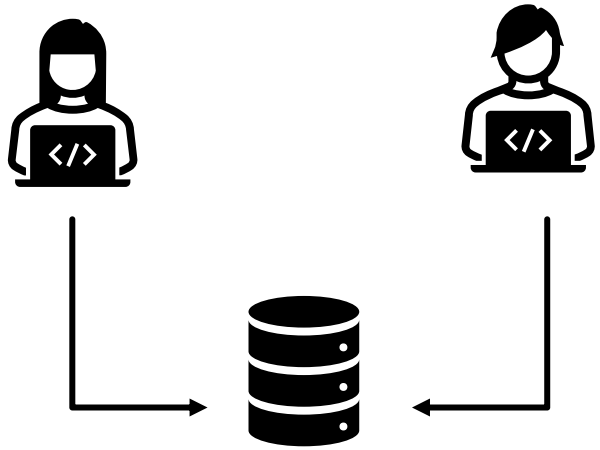
Provides a unified framework for managing diverse data types, including vector, raster, and attribute data.



ADVANTAGES OF USING A GEODATABASE

- Data Integrity and Consistency
  - Supports topology rules to ensure data accuracy (e.g., no overlapping polygons or gaps).
  - Enforces data validation and constraints to maintain reliable relationships between features.
  - Versioning support allows multiple users to work on the same dataset without conflicts.





ADVANTAGES OF USING A GEODATABASE

- Cross-Platform Compatibility
  - Compatible with various GIS software tools, such as ArcGIS, QGIS, and GeoServer.
  - Works with relational database systems like PostgreSQL/PostGIS, Oracle, or SQL Server.

ADVANTAGES OF USING A GEODATABASE

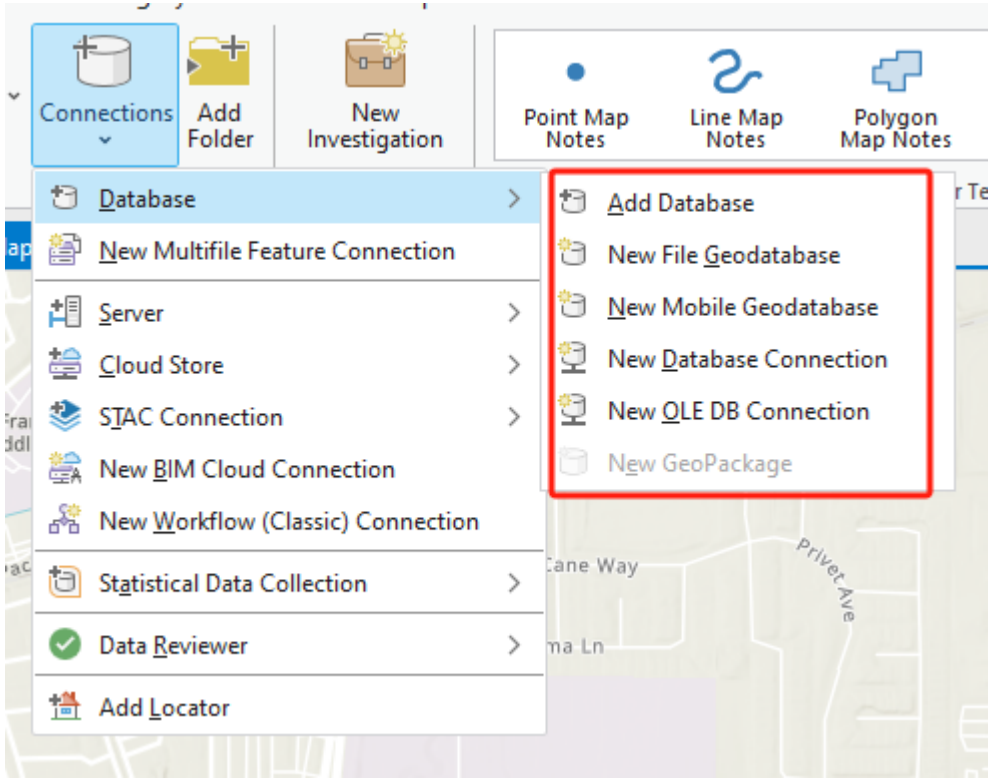
- Scalability
  - Supports various levels of implementation:
    - File Geodatabases: A file geodatabase is stored as multiple files in a folder with a `.gdb` extension. For team collaboration on medium-scale projects.
    - Enterprise Geodatabases: Also known as multiuser geodatabases, enterprise geodatabases are stored in relational databases. They can be virtually unlimited in size and number of users; the limits differ depending on the database management system (DBMS) vendor. For large organizations and multi-user environments.
    - Mobile Geodatabases: A mobile geodatabase is stored in an SQLite database that is entirely contained in a single file and has a `.geodatabase` extension.
- Capable of handling large datasets efficiently in enterprise environments.

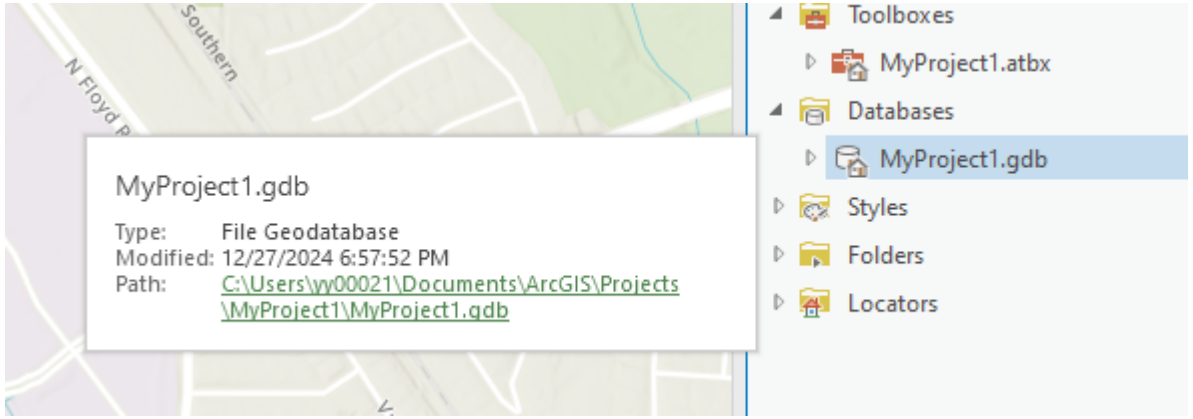
COMPARING TYPES OF GEODATABASES

More detailed comparison of geodatabases

Key characteristics	Enterprise geodatabase	Mobile geodatabase	File geodatabase
Description	A collection of various types of GIS datasets held as tables in a relational database.	A collection of various types of GIS datasets contained within a relational database.	A collection of various types of GIS datasets held in a file system folder.
Number of users	Multiple editors and can support multiple readers.	Single editor and can support multiple readers.	Single editor and can support multiple readers.
Storage format	• Oracle • Microsoft SQL Server • IBM Db2 • PostgreSQL • SAP HANA	All the datasets that belong to one mobile geodatabase are contained in an SQLite database that is stored in a single file.	Each dataset is a separate file on disk. All the datasets that belong to one geodatabase are contained in a single folder.

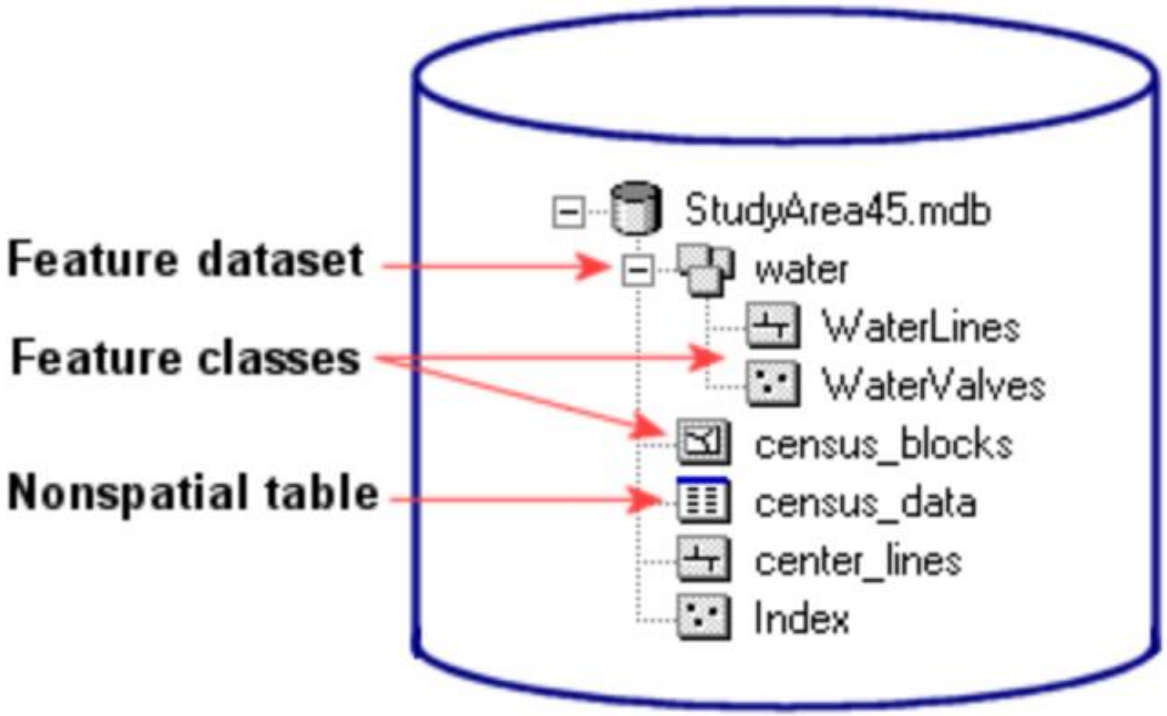
GEODATABASE IN ARCGIS PRO





FEATURE DATASETS AND FEATURE CLASS

- A feature dataset is composed of feature class
- A feature class is a collection of geographic features with the same geometry type, attributes, and spatial reference.



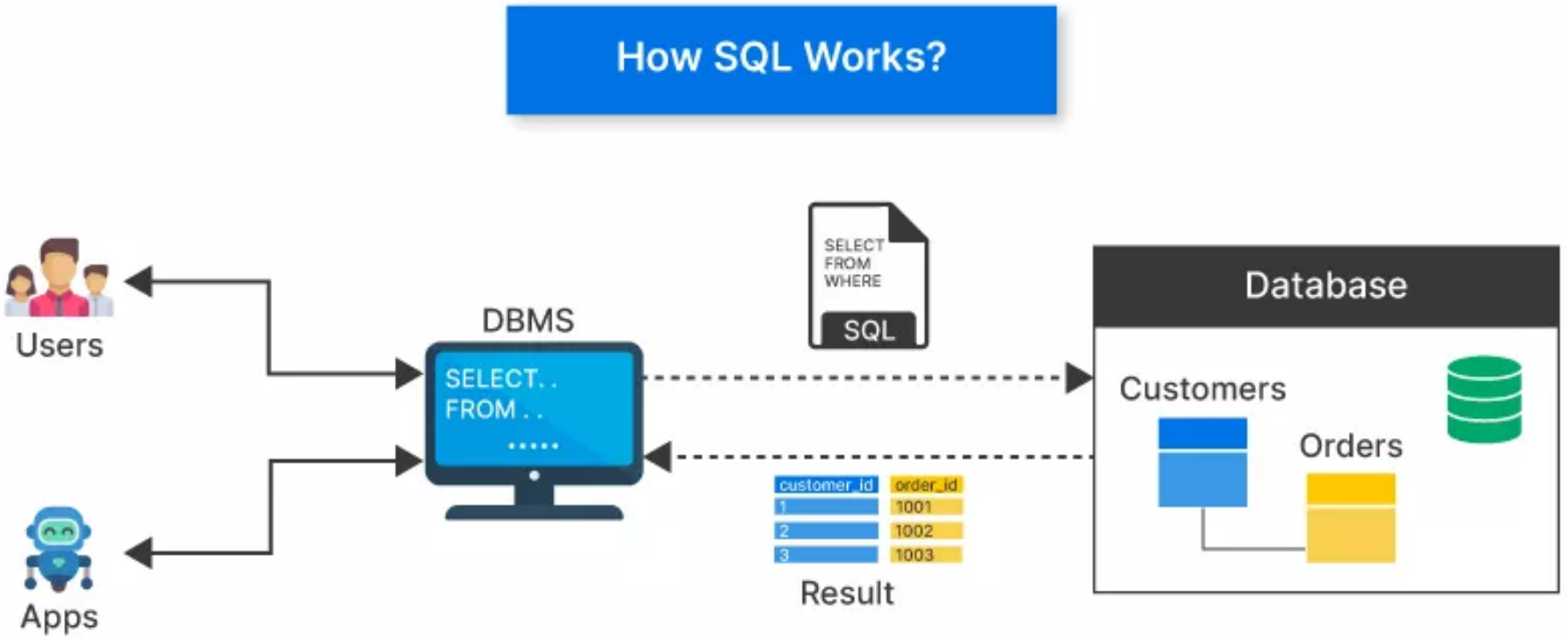
STANDARD QUERY LANGUAGE (SQL)

- SQL is a standard programming language specifically designed for managing and manipulating relational databases.

- It allows users to interact with a DBMS to perform various operations such as querying, updating, and managing database structures.

HOW SQL WORKS

- It serves as a standard interface for interacting with relational database management systems (RDBMS), allowing users to perform a variety of routine tasks and operations on data
  - such as querying, updating, inserting, and deleting records.



Source

KEY POINTS BETWEEN DBMS AND SQL

- SQL is Independent of DBMS: While SQL is a standard, each DBMS may implement it with slight variations or extensions (e.g., PL/SQL in Oracle, T-SQL in SQL Server).
- DBMS Without SQL: Some DBMS types, such as NoSQL databases (e.g., MongoDB), do not rely on SQL but use alternative query languages.
- SQL Requires a DBMS: SQL commands need a DBMS to interpret and execute them. Without a DBMS, SQL commands would have no database to operate on.

SOFTWARES

- PostgreSQL
- What is PostgreSQL?
  - An open-source, powerful, object-relational database system.
  - Known for its reliability, extensibility, and SQL compliance.
- Why PostgreSQL for GIS?
  - Supports structured queries for data management and manipulation.
  - Extensible via plugins, making it adaptable for spatial needs.

SOFTWARE

- What is PostGIS?
  - A spatial database extension for PostgreSQL.
  - Adds support for geographic objects like points, lines, and polygons.

- Key Features of PostGIS:
  - Spatial data types: GEOMETRY and GEOGRAPHY.
  - Spatial functions:
    - Query and manipulate spatial data (ST_Intersects, ST_Buffer, ST_Distance).
    - Perform spatial analyses (e.g., clipping, union, or proximity searches).

WEEK 01

NO LAB

Instructor: Yanan Wu
TA: Vanchy Li

Spring 2025