Week 6 Fri 11/3/16 GIS Data Storage and Structures

Database: A collection of data files that is organized to support data storage, manipulation, and retrieval

Usually built/managed/used by a ‘DBMS’

A ‘spatial’ database is usually at the core of a GIS program

Types of Databases

Flat file systems (one of the simplest forms of storage for a database system)

Relational Databases

Presently most widely used in GIS

Imagine dividing up a big flat file into a series of smaller files or ‘tables’

Each kind of attribute has its own table

Pros

Relatively simple to implement, well understood and ubiquitous in computing today

Can handle many data types

Encourages (requires) careful thinking in the design states

Yet within the scope of a given design you’ve implemented, the relational database does not require prior knowledge about how you will *manipulate* data

Cons

In an era with “big data,” there may be practical limits to the scale at which data can be stored and queried in typical relational databases

Can perform poorly in representing and calculating with heterogeneous and dense relationships whose types may not be well understood ahead of time (compare: graph databases!)

NoSQL databases, including graph databases

Table Terminology

Row = Record = Tuple

Column = Field = Item = Attribute

DBMS: A software package that facilitates data storage, manipulation, and retrieval

Usually has

Data definition language

Data dictionary [“metadata”]

Data entry & update modules

Report generator

Query Language

Week 7 Mon 11/7/16 GIS: Manipulation through Selection

ArcGIS for Desktop and ArcGIS Online

ArcGIS for Desktop == ArcMap OR ArcGIS Pro

These terms and the ecosystems *will change*

“Desktop GIS” used to be more “standalone” but now it’s increasingly a part of a larger “Internet GIS”

Working with Data in a Database: Query & Retrieval

Retrieval: The ability of the DBMS or GIS to get back data that were previously stored, especially to retrieve a subset of data based that fit particular characteristics

Query: An operation that you perform to guide the retrieval – the parameters of the retrieval

One of the key specializations of [a spatial database that underlies] GIS is its ability to query and retrieve based on both **attribute** and **spatial** characteristics

SQL: A formal/mathematical language for manipulating relational databases. Gets things and changes them too

ArcGIS uses SQL

Queries and Practicality:

The results of selections in ArcGIS are “stable” – records remain selected UNTIL you un-select them

Queries are a common way to retrieve a smaller subset of info to work with (consider exporting the selection to new file)

Pros to this?

Database vs Cartographic Approaches

We’ve been making cartographic masks to [de]emphasize some parts of the data

We’ve also been using ‘filters’ to only show some parts of the data

Queries are more like filters. They return a ‘subset’ of the data in a ‘selection’

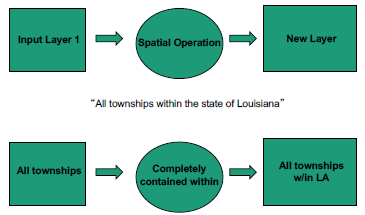
Spatial Queries

Select/retrieve records or objects based upon spatial/geographic characteristics

Spatial queries produce new sets of geographical features which might then be used in building new GIS layers

Many forms of analysis rely on or start with spatial query

Visually ‘modeling’ a spatial query:



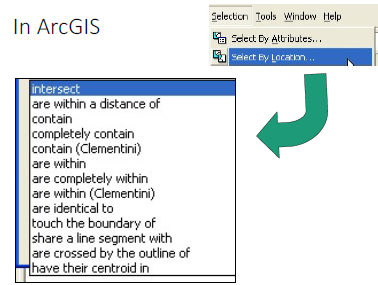
More Examples of Spatial Queries by Location

Adjacency: Retrieves all records that share a boundary

Containment: All features in one layer that are completely within the features of another layer

Proximity: All features within a certain distance of features of another layer [Compare with buffering, later!]

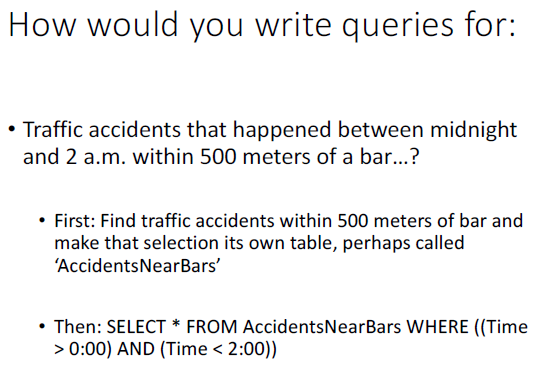
Intersect: All features that intersect the features of another layer



Week 7 Wed 11/9/16 Introduction to Spatial Analysis

Writing Attribute Queries: Use the proper SQL syntax

Writing Spatial Queries: Just write out in words



Relational Algebra: An abstract language of operations that allow you to manipulate tables and their data

These operations may, for example

Combine

Select

Or split tables

And often may be configured by setting parameters

SQL provides you with an implementation of a relational algebra

Selection (restriction): Reduces one table in the record dimension (a selection of a subset of records, for all fields)

AKA filter/where specific rows

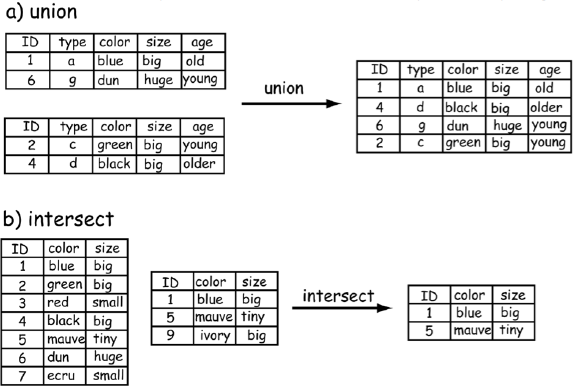
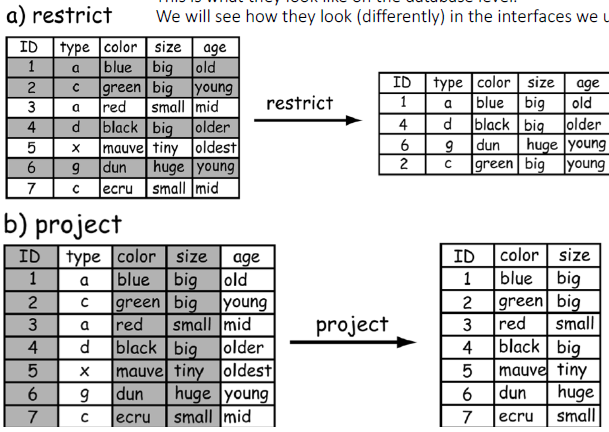
Project: Select specific columns

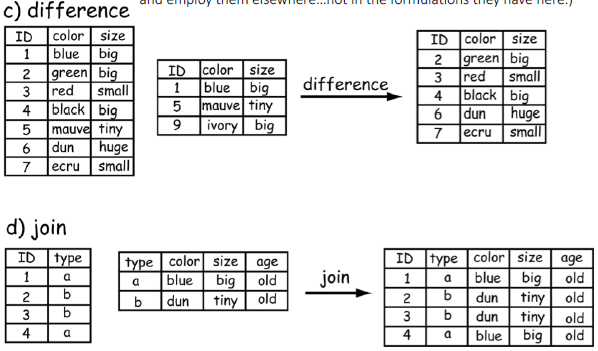
Union: Combine all rows from two different tables into one table

Intersect: Filter/where rows that both tables have into a new table.

Difference: Opposite of intersect. Filter/where rows that neither table has in common

Join: Add columns and rows from one table into another.





Spatial Analysis: Using geographic computational methods to help us discover and understand spatial patterns and relationships within spatial data.

Offers techniques to help ask and answer questions of sptial data

Simple: “How long is the Chattahoochee River?”

More complicated: “If we have a 2 degree temperature increase, how will the spatial distributions of particular species change?”