

# Introduction to Enterprise Databases

IT4GIS

Keith T. Weber, GISP

GIS Director

ISU-GIS Training and Research Center

Pocatello | Idaho Falls | Meridian | Twin Falls

Idaho State UNIVERSITY

---

---

---

---

---

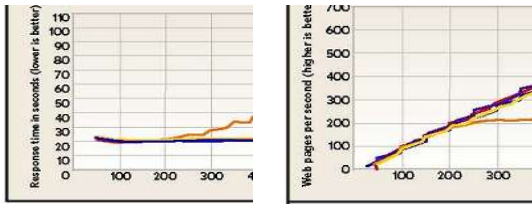
---

---

---

## Concurrent Clients

- GIS for the Enterprise
  - Focus on current/near term concurrent clients



Pocatello | Idaho Falls | Meridian | Twin Falls

Idaho State UNIVERSITY

---

---

---

---

---

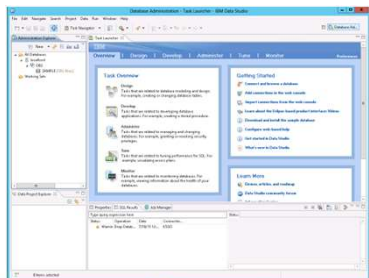
---

---

---

## Database Administration (e.g., IBM DB2)

- GUI based database administration
- Alternatively, command prompt can be used.
- **Do you know what the command prompt is?**



Pocatello | Idaho Falls | Meridian | Twin Falls

Idaho State UNIVERSITY

---

---

---

---

---

---

---

---



## Numeric Data Types



- FOR BIT DATA (boolean)
- BYTE (0-255)
- SMALLINT (-32,768 to 32,767 )
- INTEGER (-2,147,483,648 to 2,147,483,647)
- FLOAT <n>
- DOUBLE PRECISION <n<sub>p</sub>,n<sub>s</sub>>

Pocatello | Idaho Falls | Meridian | Twin Falls

Idaho State  
UNIVERSITY

---

---

---

---

---

---

---

---

## Data Type Parameters Supported in ArcGIS

- FLOAT <n<sub>p</sub>,n<sub>s</sub>>
  - n<sub>precision</sub>(total field length)
  - n<sub>scale</sub> (decimal places)
  - n must be between 1-6 (larger n values need to use DOUBLE)
  - n<sub>p</sub>,n<sub>s</sub> = 5,3 → 26.589 is OK, 256.381 is not
  - Five (5) total characters 2 6 . 5 8 9

Pocatello | Idaho Falls | Meridian | Twin Falls

Idaho State  
UNIVERSITY

---

---

---

---

---

---

---

---

## Parameters (cont'd)

- DOUBLE PRECISION <n<sub>p</sub>,n<sub>s</sub>>
  - n<sub>p</sub> = 7 or more
  - n<sub>s</sub> = 0 or more

Pocatello | Idaho Falls | Meridian | Twin Falls

Idaho State  
UNIVERSITY

---

---

---

---

---

---

---

---

## Character Data Types

- CHARACTER<n>
- VARCHAR<n>



---

---

---

---

---

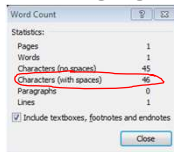
---

---

---

## Parameters (cont'd)

- CHARACTER<n>
  - (AKA, String or Text)
  - Example a field named "URL" with  $n = 46$
  - "<http://giscenter.isu.edu/training/it4gis.htm>"



---

---

---

---

---

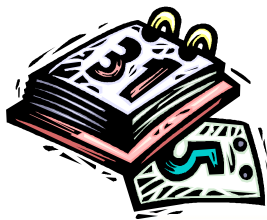
---

---

---

## Special Data Types

- DATE
- TIME
- TIMESTAMP



---

---

---

---

---

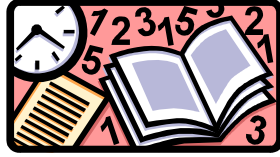
---

---

---

## Special Data Types (cont'd)

- Stored in special *System managed* tables
  - BLOB<n[K|M|G]>
  - CLOB<n[K|M|G]>
  - DBCLOB<n[K|M|G]>
  - GRAPHIC<n>
  - VARGRAPHIC<n>



Pocatello | Idaho Falls | Meridian | Twin Falls

Idaho State  
UNIVERSITY

---

---

---

---

---

---

---

---

## Beware of the 64-bit OBJECTID

- The new Esri 64-bit OBJECTID allows practically unlimited number of records in a geodatabase table
- HOWEVER, it is not supported by older versions of ArcGIS and is not supported in the shapefile. Thus these data:
  - Can only be used in current versions of ArcGIS
  - Do NOT support Open GIS
  - Cannot be exported for use in other software (QGIS, Idrisi Terrset, etc.)

Pocatello | Idaho Falls | Meridian | Twin Falls

Idaho State  
UNIVERSITY

---

---

---

---

---

---

---

---

## Table Data Pages

- All fields with *standard* data types for each record are contained within a single data page.
- There is a maximum of 255 records stored on each page.
- The ART of efficient data modeling is to minimize wasted space on a page while maximizing the proportion of each page written.

Pocatello | Idaho Falls | Meridian | Twin Falls

Idaho State  
UNIVERSITY

---

---

---

---

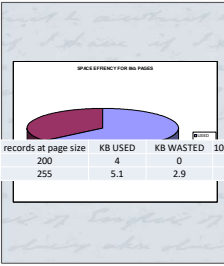
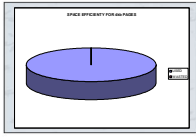
---

---

---

---

## An Example



number of fields	KB per record	page size	KB for 255 records	records at page size	KB USED	KB WASTED	100 GB TABLE SPACE
10	0.02	4	5.1	200	4	0	0 GB WASTED
10	0.02	8	5.1	255	5.1	2.9	29 GB WASTED

Each record contains 10 fields  
 Each record needs 20 bytes (0.02kb)  
 How many records will fit into a 4kb page?

What percentage of a page is written?  
 -based on available space?  
 -based on available records?

one 8kb page

---

---

---

---

---

---

---

---

---

---

---

---

## Storing Vector Coordinates in a ORDBMS

- **DB2 Spatial Extender** (and other spatially enabled databases) lets you integrate geographic data with your existing business data. It includes:
  - Data types such as points, lines, and polygons
  - Functions such as area, endpoint, and intersect
  - An indexing scheme for spatial data
- What about Oracle, MS SQL Server, and PostgreSQL?

---

---

---

---

---

---

---

---

---

---

---

---

## Professional Hints and Tips

- Working toward a security clearance

---

---

---

---

---

---

---

---

---

---

---

---



## Key Concepts

- Understand that while data is stored in tables, these tables span TABLE PAGES
- Understand what PRE-FETCH and CACHE are...and how they differ.
- Understand data types

---

---

---

---

---

---

---

---

## Questions?



---

---

---

---

---

---

---

---

## OBJECT-ORIENTED DESIGN

---

---

---

---

---

---

---

---

## The Early Days...

- Computer programming from the caveman era



---

---

---

---

---

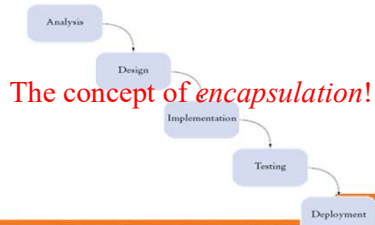
---

---

---

## Why...Object-Oriented

- A brief history of computer programming...
- Early *waterfall* programming



---

---

---

---

---

---

---

---

## What is?

- Fundamentally, we need to know
  - What is a CLASS
  - What is an OBJECT

---

---

---

---

---

---

---

---



## What is a CLASS?

- A *class* is a computer construct representing a concept bound in a cohesive package
  - Some are concrete (i.e., real world)
    - Bank account
    - Rental item
    - Database item
    - Pile
  - Others are abstract
    - Scanner
    - Stream
    - Math

Pocatello | Idaho Falls | Meridian | Twin Falls

Idaho State  
UNIVERSITY

---

---

---

---

---

---

---

---

## Discovering CLASSES

- Simple Rule:
  - Look for *nouns* in descriptions
  - Obviously not all nouns are classes
  - But at least this approach can allow one to create a list of *candidate classes*

Pocatello | Idaho Falls | Meridian | Twin Falls

Idaho State  
UNIVERSITY

---

---

---

---

---

---

---

---

## What is an OBJECT

- An *instance* of a CLASS
- That contains meaningful data
- OBJECTS occupy memory space at runtime
  - If not, they are CLASSES
  - For example: data type vs. double

Pocatello | Idaho Falls | Meridian | Twin Falls

Idaho State  
UNIVERSITY

---

---

---

---

---

---

---

---

### A Little Quiz...

- #1 Class or Object?



**Dog**

*Dog is a generalization of Scooby-Doo*

**Scooby-Doo**

---

---

---

---

---

---

---

---

### A Little Quiz (cont'd)...

- #2 Class or Object?

**The concept of *subclass!***

**Animal**

*Dog is a subclass of the Animal class  
Animal is a generalization of Dog*

**Dog**

**Scooby-Doo**

---

---

---

---

---

---

---

---

### A Little Quiz (cont'd)...

- #3 Class or Object?

**Animal**

**The concept of *polymorphism!***

**Bird**

**Dog**

---

---

---

---

---

---

---

---

Questions so far...



---

---

---

---

---

---

---

---

---

---



### Key Points

- Many classes already exist and are at our disposal when we design a database
- *Inheritance* is an important concept
  - A *subclass* inherits from its *superclass*
  - i.e., a *child* inherits from its *parent*

---

---

---

---

---

---

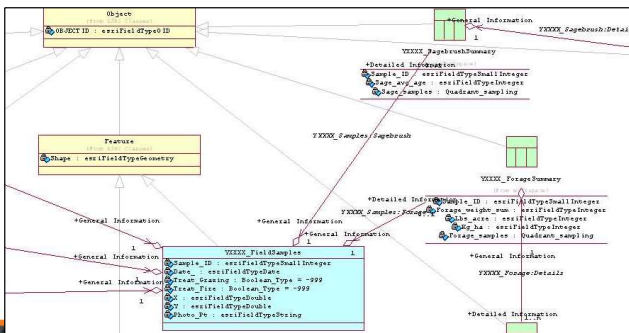
---

---

---

---

### Inheritance in the Geodatabase



---

---

---

---

---

---

---

---

---

---

## Identifying Inheritance

- *Is-a* relationship
- Relationship between a more specialized class (*subclass*) and a generalized class (*superclass*)
- Every...
  - Savings account *is a* bank account
  - DVD rental *is a* rental
  - Dog *is a* mammal
  - Parcel *is a* polygon

Pocatello | Idaho Falls | Meridian | Twin Falls

Idaho State  
UNIVERSITY

---

---

---

---

---

---

---

---

CLASS OR OBJECT

## INSTANTIATION

Pocatello | Idaho Falls | Meridian | Twin Falls

Idaho State  
UNIVERSITY

---

---

---

---

---

---

---

---

## Instantiate into an *Object* (to make into an instance)

- Three features characterize OBJECTS and distinguish OBJECTS from CLASSES:
  - **Identity**: specific attribute (property) settings have been made for the class. This distinguishes it from all other objects.
  - **State**: Describes the data stored in the object **WHERE DID THIS COME FROM?**
  - **Behavior**: describes the method in the object's **interface** through which the object can be used (how do we make the dog bark?)

Pocatello | Idaho Falls | Meridian | Twin Falls

Idaho State  
UNIVERSITY

---

---

---

---

---

---

---

---

## Instantiating the Dog CLASS

- CLASS (DOG)
- Attributes (Properties)
  - NAME = Scooby-Doo
  - HEIGHT = 36
  - WEIGHT = 145
- Inheritance?



Scooby-Doo

Pocatello | Idaho Falls | Meridian | Twin Falls

Idaho State  
UNIVERSITY

---

---

---

---

---

---

---

---

## Key Concepts

- Understand the difference between a CLASS and an OBJECT
- Understand new terms:
  - Encapsulation, polymorphism, superclass, subclass, behavior, attributes, instantiation
- Understand why inheritance is an important part of an object-relational database

Pocatello | Idaho Falls | Meridian | Twin Falls

Idaho State  
UNIVERSITY

---

---

---

---

---

---

---

---

## DATABASE DESIGN CONCEPTS AND PRACTICES

Pocatello | Idaho Falls | Meridian | Twin Falls

Idaho State  
UNIVERSITY

---

---

---

---

---

---

---

---

## Basic Steps in Database Design



- Understand and document the business' needs.
  - Problem statement
  - Business object types
  - Business relationships
  - Business constraints
- Create an ERM
- Data and process inventory
- Integrity
- Populate the database

Pocatello | Idaho Falls | Meridian | Twin Falls

Idaho State  
UNIVERSITY

---

---

---

---

---

---

---

---

## Database Design

- Why spend so much time and effort?
  - Efficiency (speed, storage)
  - Client satisfaction
  - Flexibility
  - Cost savings realized

Pocatello | Idaho Falls | Meridian | Twin Falls

Idaho State  
UNIVERSITY

---

---

---

---

---

---

---

---

## Design Considerations

- Basic steps (described earlier)
- Data types/Data Modeling
- Normalization
- With >1 table, relationships must be examined

Pocatello | Idaho Falls | Meridian | Twin Falls

Idaho State  
UNIVERSITY

---

---

---

---

---

---

---

---

## Relationships

- Determine where relationships exist between tables
- Determine the type of relationship that exists



- One-to-one
- One-to-many
- Many-to-one
- Many-to-Many



---

---

---

---

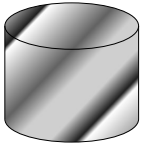
---

---

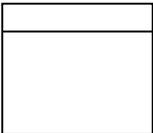
---

---

## Generic Design Symbology



= Database



= Table

---

---

---

---

---

---

---

---

## Generic Table Symbology

Parcels
Parcel_ID
TRS
Value
Zoning

← Table name  
← Divider

← List of all attributes stored in this table as they will appear in the table

---

---

---

---

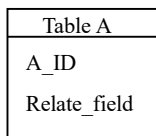
---

---

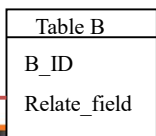
---

---

## Generic Relationship Symbology



- Draw schema of RDB
- Determine relationship fields
- Connect



Pocatello | Idaho Falls | Meridian | Twin Falls

State UNIVERSITY

---

---

---

---

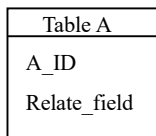
---

---

---

---

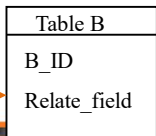
## Symbolizing Relationship Type



- One-to-one  $1..1 \rightarrow 1..1$
- One-to-many  $1..1 \rightarrow 1..M$
- Zero?  $\rightarrow$

1..1

1..M



Pocatello | Idaho Falls | Meridian | Twin Falls

State UNIVERSITY

---

---

---

---

---

---

---

---

## The Relationship Type...

- Also known as
  - Cardinality (ArcGIS terminology)
  - Multiplicity (UML terminology)

Pocatello | Idaho Falls | Meridian | Twin Falls

Idaho State UNIVERSITY

---

---

---

---

---

---

---

---



## Generalized Process



Inception



Elaboration

- Construction
- Transition

---

---

---

---

---

---

---

---

## Professional Hints and Tips

- Getting ready for job interviews
  - Social Media
  - Dress for success

---

---

---

---

---

---

---

---

## Questions...



---

---

---

---

---

---

---

---

### Your Assignment

- Complete the exercise following the README text file in this week's exercise package
- But first, time for another 2-minute Write!

---

---

---

---

---

---

---

---