

# Lecture 1:

# Building a Geodatabase

Sergiy Kostrikov

Geographic Information Systems / Geodatabases for  
Human Geographers — PhD Courses  
Fall 2020

# Reminders

- You should look through the Basic Geodatabase textbook

# Objectives

- Become proficient in the use of GIS tools to conduct spatial analyses.
- Locate and obtain geospatial data resources from federal, state, and local data providers.
- Correctly convert data into a common coordinate system appropriate for a study site and objectives.
- Work with data in formats common to GIS, including shapefiles, rasters, and geodatabases.

# Why are we covering geodata bases?

- It is the native data structure for ArcGIS
- Primary data format used for editing and data management
- Much of the powerful functionality of ArcGIS depends on geodatabase models

## First, Some Design Considerations

# • What kind of geodatabase do you need?

## File geodatabase

- A folder of files that is cross platform
- Datasets can be up to 1 TB in size (with configuration larger

SingleBest default choice



## UserPersonal geodatabase

- Stores data in Microsoft Access
- Limited to 250 — 500 MB for the entire geodatabase
- Only supported on Windows

## Enterprise geodatabase

- Required for multiple simultaneous users

Multiimplemented within Relational Database Management System

User(RDBMS) software

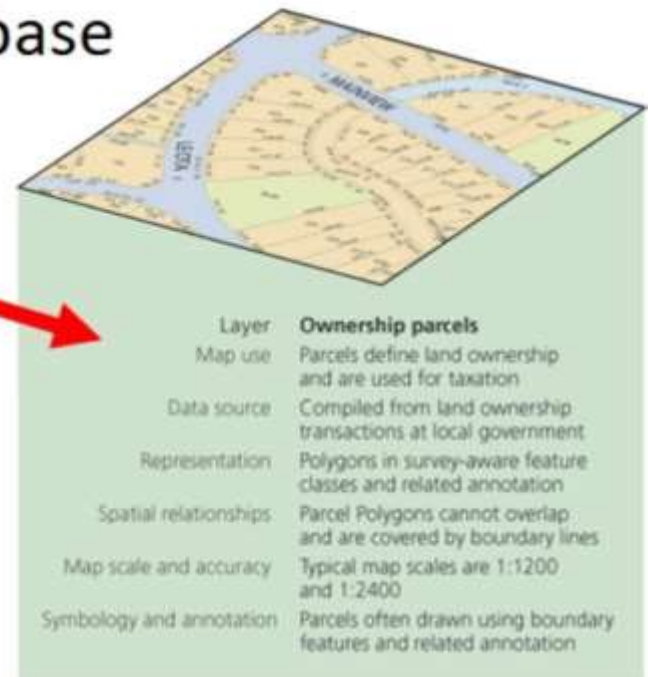
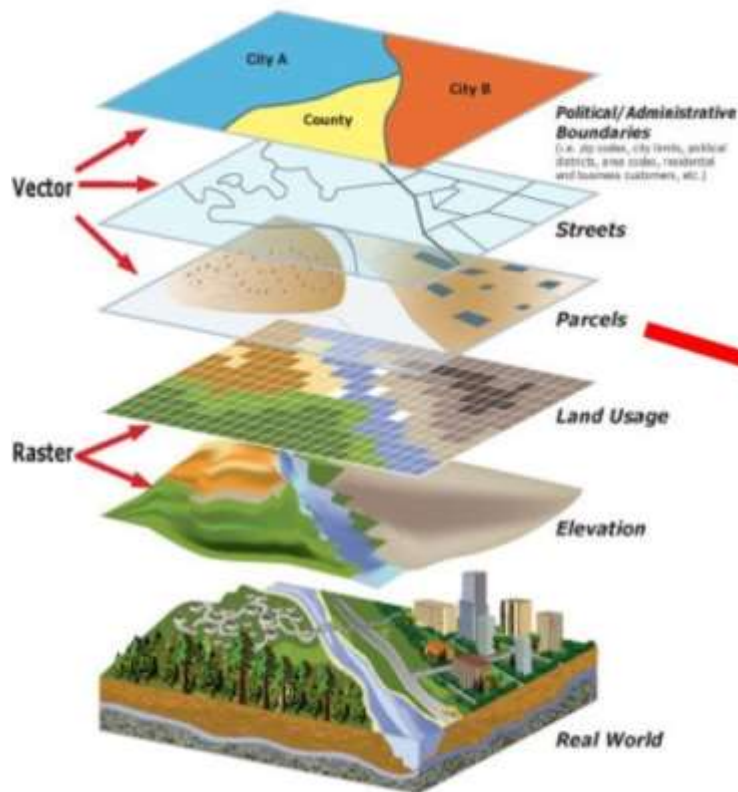
- Required for archive, replication, SQL querying, multi-user editing



What are the thematic layers I need to include?

- Roads, streams, soils, terrain?

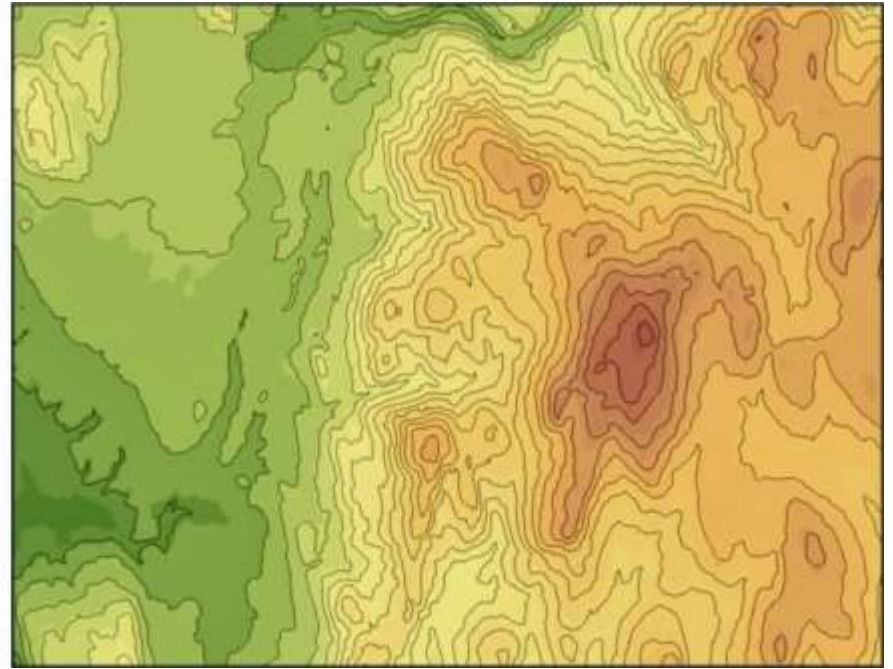
- Develop specification for representing the contents of each thematic layer in the database



# What is the correct geographic representation for each thematic layer?

- Elevation as a raster of elevations or as a vector dataset of elevation contours? Both?
- Can be determined by how you intend to use the data



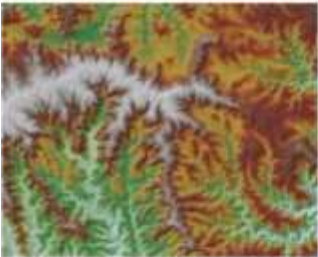


What object types do I need in my database?

## Vector Datasets



## Rasters



## Non-Spatial Tabular Data



# Geodatabase

## Feature Classes



## Raster Datasets



## Object Class (Table)



# What are the attribute fields and their data types for each dataset?

- What columns do I need in the table for each dataset?
- What data type should each column in a table have?
- Numbers (multiple)

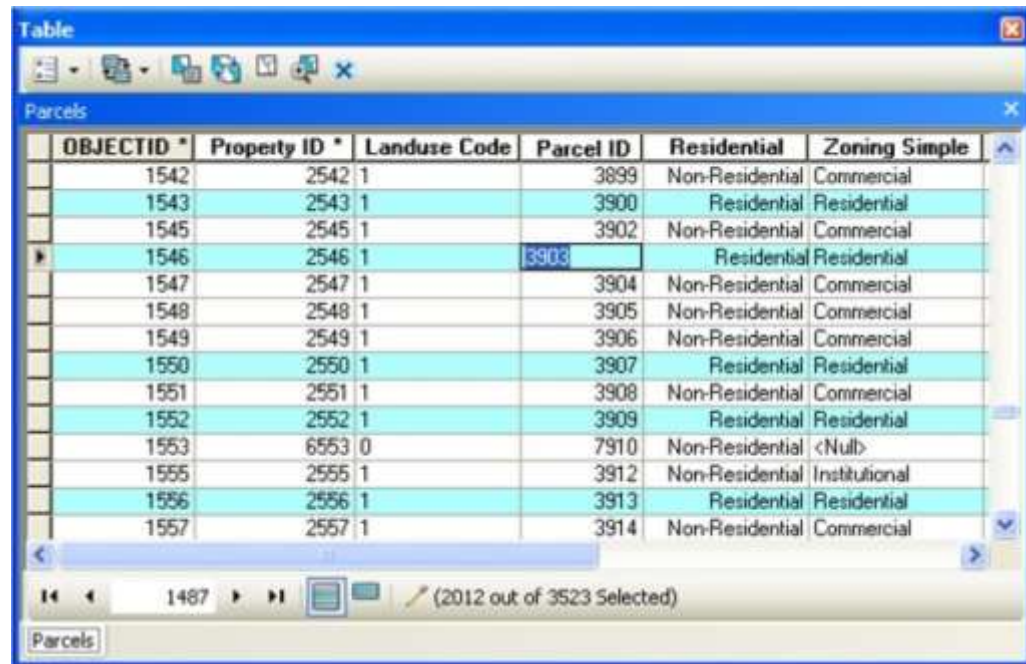
◦ Text ◦ Dates

◦ BLOBs ◦ Object IDs  
(automatic) ◦ Global IDs  
◦ Raster

◦ Geometry

Are there  
important  
relationships I  
need to capture?

- Are there tables containing additional data related to the geospatial location data?



OBJECTID *	Property ID *	Landuse Code	Parcel ID	Residential	Zoning Simple
1542	2542	1	3899	Non-Residential	Commercial
1543	2543	1	3900	Residential	Residential
1545	2545	1	3902	Non-Residential	Commercial
1546	2546	1	3903	Residential	Residential
1547	2547	1	3904	Non-Residential	Commercial
1548	2548	1	3905	Non-Residential	Commercial
1549	2549	1	3906	Non-Residential	Commercial
1550	2550	1	3907	Residential	Residential
1551	2551	1	3908	Non-Residential	Commercial
1552	2552	1	3909	Residential	Residential
1553	6553	0	7910	Non-Residential	<Null>
1555	2555	1	3912	Non-Residential	Institutional
1556	2556	1	3913	Residential	Residential
1557	2557	1	3914	Non-Residential	Commercial

Parent Feature  
Class

Table with related  
information

Relationship Class



ParceffoPermit



Permit

Parcel (origin)(destination)

Parcel 10	Zone	Bl
123		
789		

primary  
key

Permit	Parcel 10	File Oate
.1	456	
2	789	
3	789	
4	123	

Destination  
foreign key

# How should my datasets be organized?

## LocalGovernment.qdb



E) AdministrativeArea  
AdministrativeAreaBoundaryLine

- Are there spatially relatedCountyBoundaryMunicipalBoundary

feature classes that should be  
grouped into a feature  
dataset?

[E StateBoundary

EI Stormwater

Stormwater\_Net

Stormwater\_Net\_Junctions

EI swCasing

SchoolBoundary

- Advanced topics: are you using any of the following?

- Topology
- swDischargePoint
- Network dataset
- swGravityMain
- Terrain dataset (TIN)
- swInlet
- swManhole



swControlValve

swDetention

swFitting



swNetworkStructure

swOpenDrain swPressurePipe

swSystemValve

swVirtualDrainline

swWeirStructure

# Do you need geodatabase elements to facilitate data editing and manage data integrity?

- Coded value domains — require attribute values to be selected from a list
- Subtypes — require features to be one of a set of available subtypes

Coded value domain

TSDatatype

Description

Field type Long integer

Split policy Default value

Merge policy Default value

Code	Description
1	Instantaneous
2	Cumulative
3	Incremental
4	Average
5	Maximum
6	Minimum



Line feature class  
**Boundary**

Subtypes are Water, Right of Way,  
Subdivision



- Relationships allow records in one table only when related records exist in another

Boundary, Parcel Split, Lot Line, Parcel, Private Road
--

## Create a New Geodatabase

- Three methods:
  1. Design and create a new empty geodatabase
  2. Copy and modify an existing geodatabase schema (empty database)

3. Create a copy of the schema and contents from an existing geodatabase

## ArcGIS Data Models

- ESRI has developed a series of geodatabase model templates
- These can be used as-is or as the basis for a more advanced implementation
- To use an existing ESRI data model:
  1. Download the appropriate data model  
<http://support.esri.com/datamodels>
  2. Create an empty file geodatabase

3. Import the schema and set up the appropriate spatial

---

ata eo ata ases co      in -t e-sc ema-o -a- eo ata

reference for its contents. See [http : deskto .arc is.com en  
deskto latest mana easea out-copying-. tm](http://desktop.arcgis.com/en/desktop/latest/management/ease_out-copying-.htm)

4. Load some data
5. Test and refine

## Existing ESRI Data Models

- Address
- Agriculture
- Atmospheric
- Basemap
- Biodiversity
- Defense — Intel
- Energy Utilities
- Environmental Regulated Facilities
- Fire Service

- BroadbandStat
- Building Interior Space
- Carbon Footprint
- Census — Admin Boundaries
- Forest Service
- Forestry
- Geology
- Groundwater
- Health

Learn More:

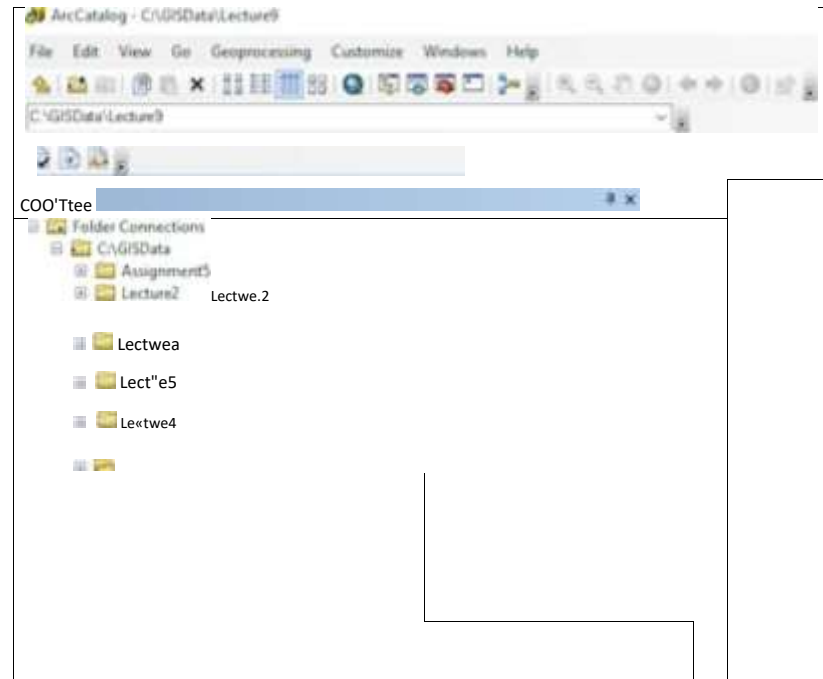
<http://support.esri.com/en/knowledgebase/techarticles/detail/40585>

a

# Creating Geodatabase

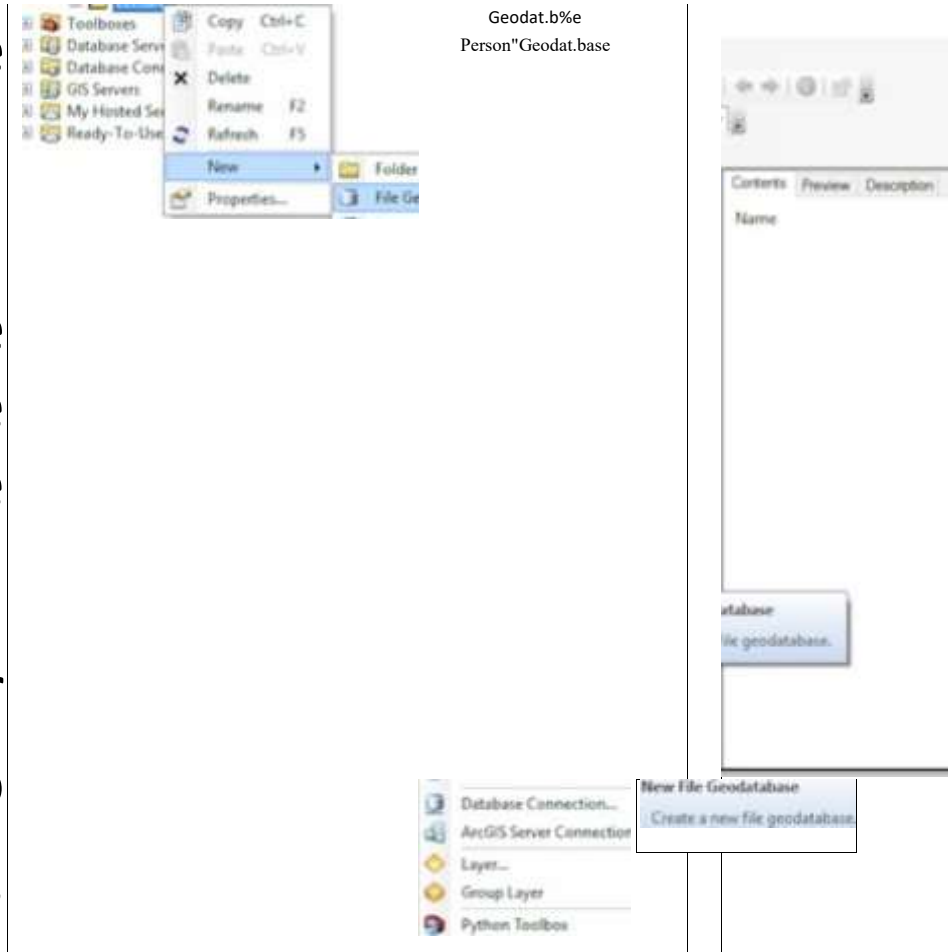
- A useful and more in depth tutorial:  
<http://desktop.arcgis.com/en/desktop/latest/manage-data/geodatabases/a-quick-tour-of-thebuilding-geodatabases-tutorial.htm>

# Create a Create New File Geodatabase



a

1. Open ArcCatalog (or the Catalog Window in ArcMap)
2. Navigate to folder where you want to create the geodatabase in the Catalog tree
3. Right click on the folder where you want to create the geodatabase, point to "New", and click "File Geodatabase"

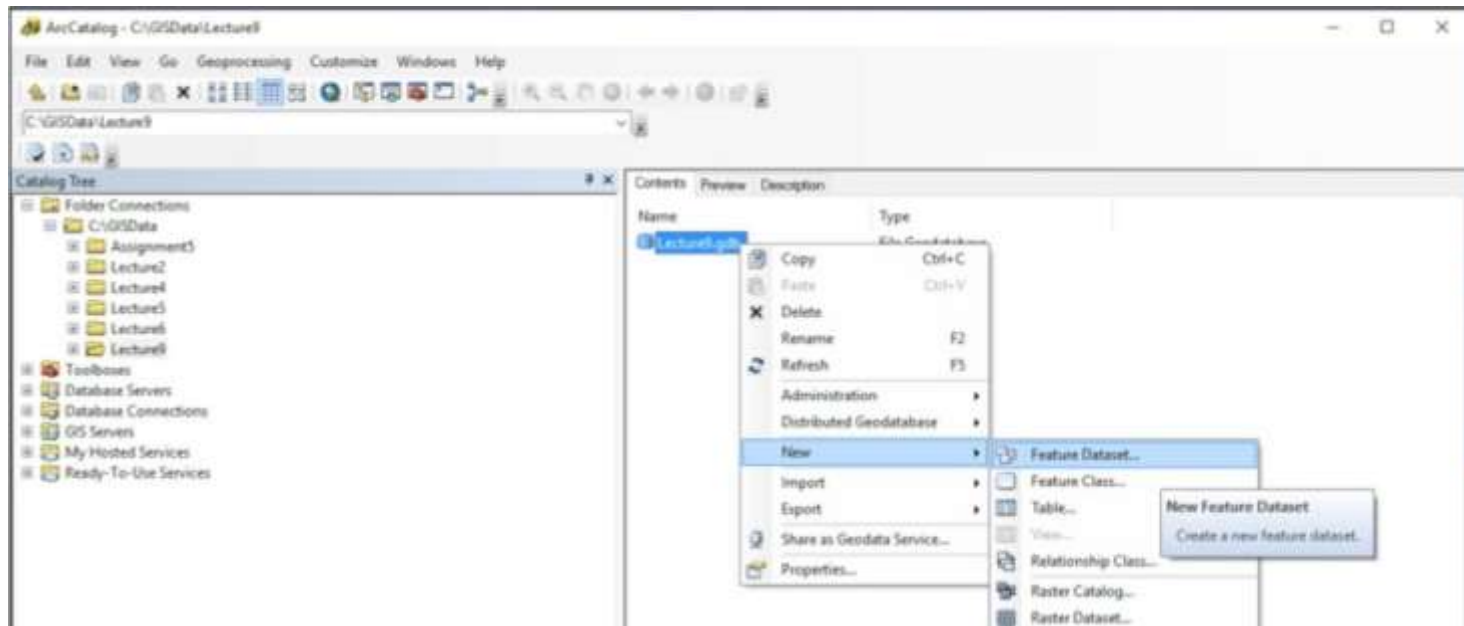


# Create a

NOTE: You can rename a geodatabase by right clicking on the geodatabase and selecting "Rename" from the context menu

## Feature Dataset

1. Right click on the geodatabase
2. Click New --> Feature Dataset





a

3. Give your Feature Dataset a name
4. Choose a coordinate system (all of the feature classes in your feature dataset must use the same coordinate system)

# Create

## Choose Coordinate Systems

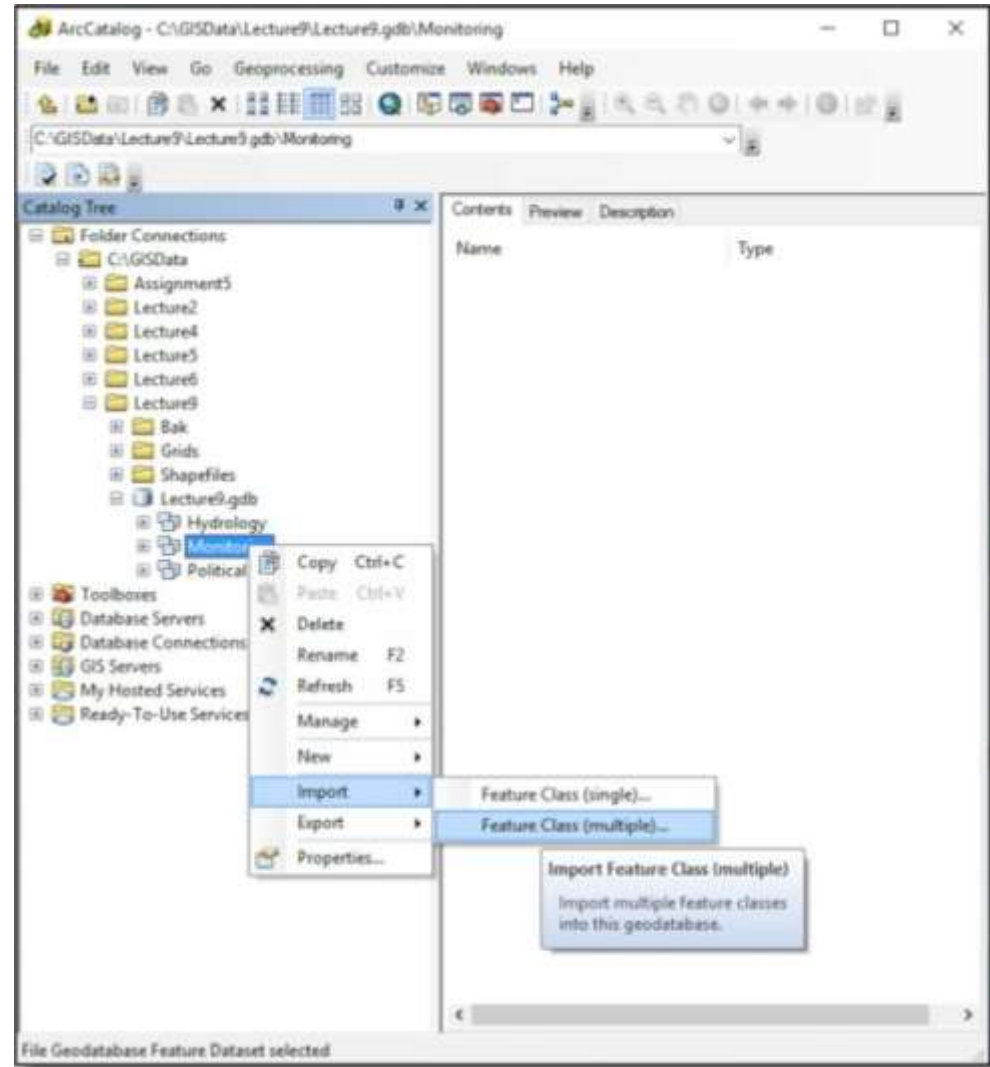
- XY Coordinate System: NAD 1983 UTM Zone 12N
- Z Coordinate System: NAVD 1988



# Feature Classes

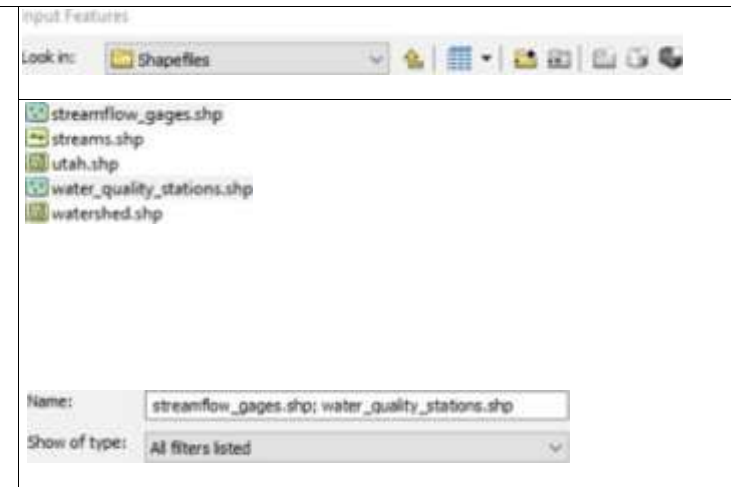
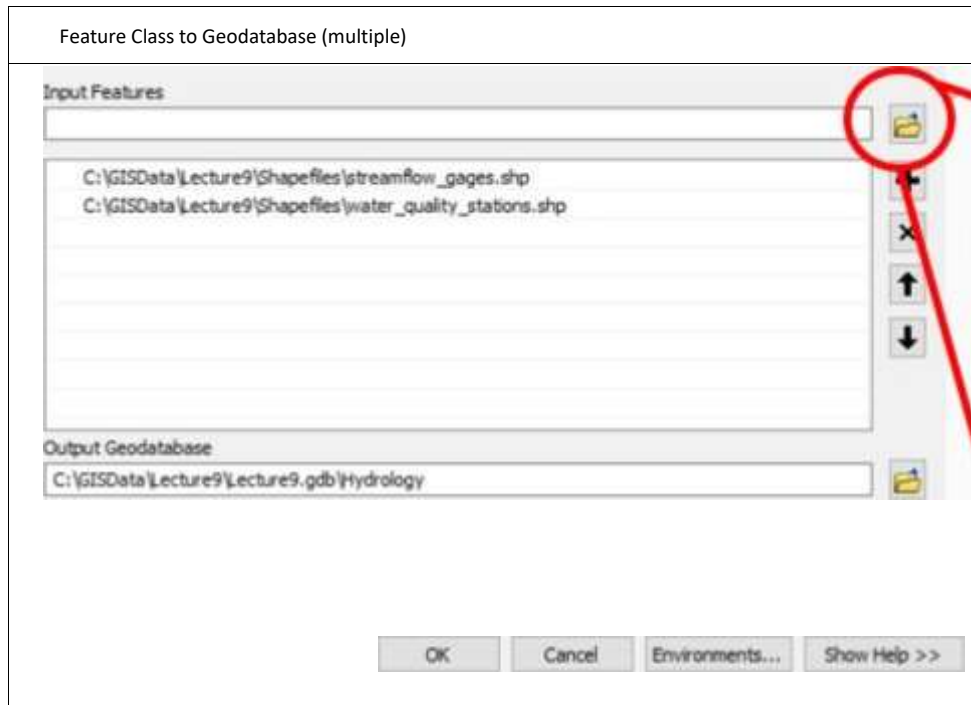
# Create

- Two methods to Create new, empty feature class or Import feature class from shapefiles or layer files
- Right click on Feature Dataset
  - New or Import



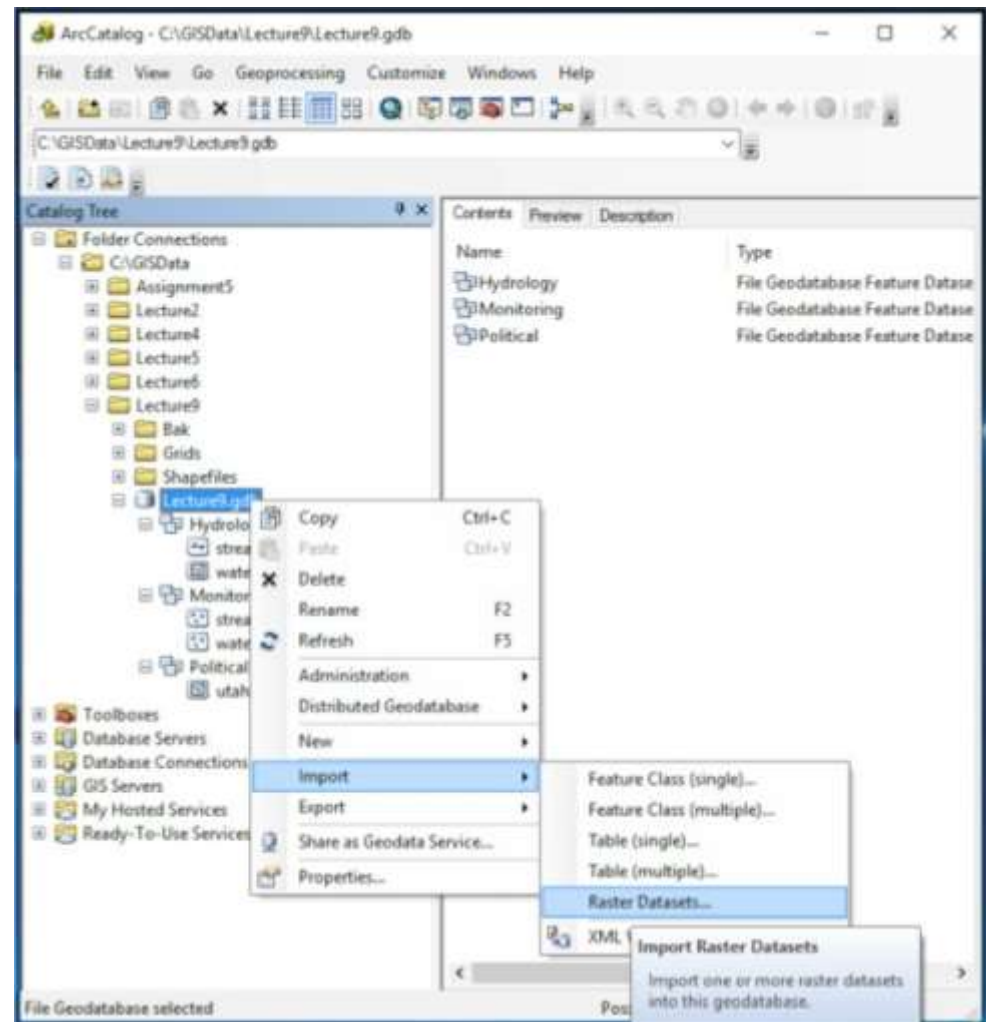
# Import Feature Classes

- Import Feature Class (multiple)
- Choose feature classes to import



# Create a Raster Dataset

- Right click on geodatabase
- Import --> Raster Datasets...

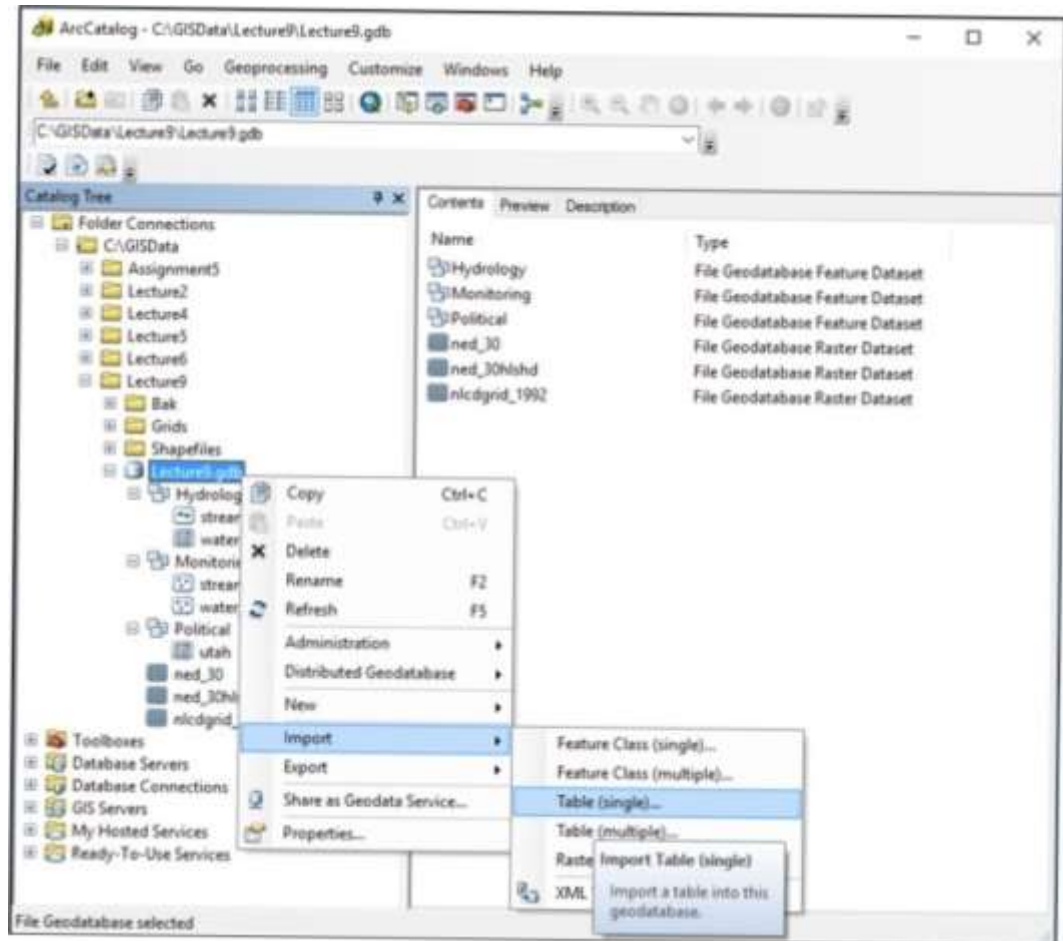


Create

Create a Table



- Right click on geodatabase • New Table: create a new empty table or Import --> Table: import a table from an existing file



# Import Table from File

- Supported formats:
  - Excel (if you have it installed) (.xlsx)
  - Comma separated values (.csv)
  - dBASE (.dbf)
  - A couple of other less common formats

- Must specify:
- Input rows (the input file)
  - Output location (the geodatabase)
  - Output table (name)
  - An expression (optional)
  - Field data type mapping (optional)

## Some Notes about Data Type Mapping

- AgencyCode = Text

Tableto Table

Input Rows  
C:\GISData\Lecture9\Tables\BR\_Discharge.csv

Output Location  
C:\GISData\Lecture9\Lecture9.gdb

Output Table  
discharge

Expression (optional)

Field Map (optional)

- AgencyCode (Text)  
C:\GISData\Lecture9\Tables\BR\_Discharge.csv.AgencyCode (Text)
- SiteCode (Text)  
C:\GISData\Lecture9\Tables\BR\_Discharge.csv.SiteCode (Long)
- Date (Date)  
C:\GISData\Lecture9\Tables\BR\_Discharge.csv.Date (Date)
- DischargeCFS (Double)  
C:\GISData\Lecture9\Tables\BR\_Discharge.csv.DischargeCFS (Long)
- CommentCode (Text)

OK Cancel Environments... Show Help >>

- SiteCode = Text (Length =10)
- Date = Date
- DischargeCFS = Double

(Precision = 8, Scale = 2)

- CommentCode = Text

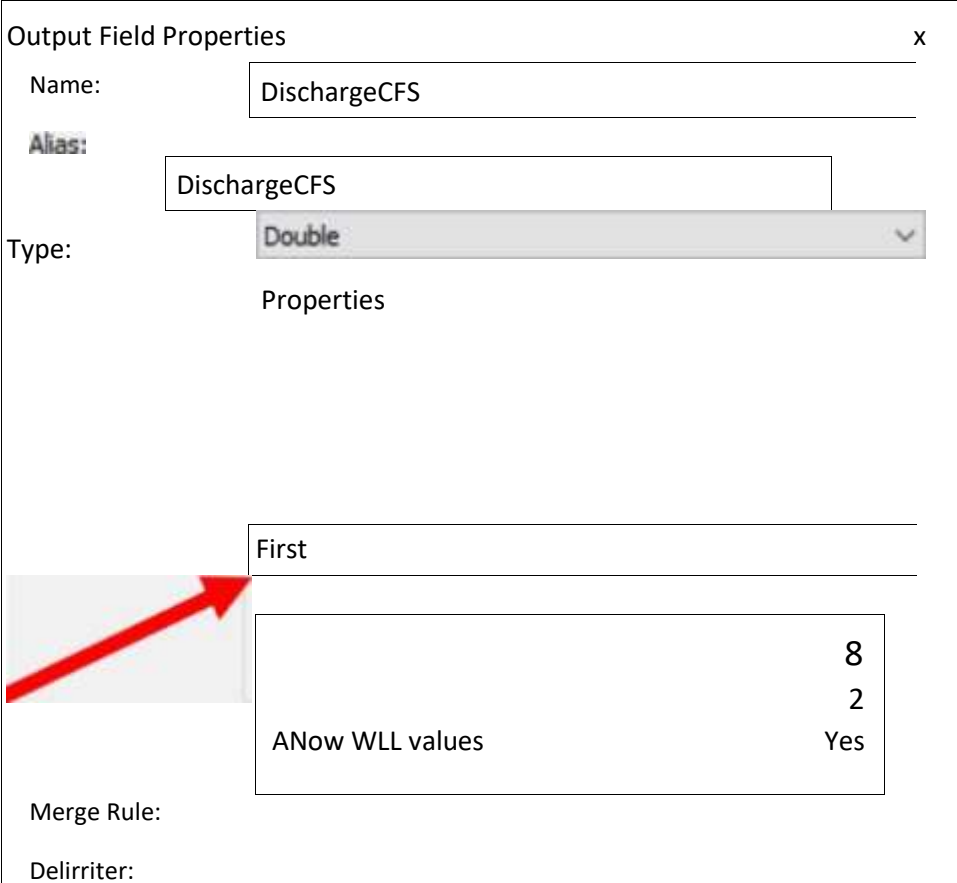
Length: Total number of characters

Precision: The total number of digits stored (total field length)

Scale: The total number of decimal places

Allow Nulls: Can the field be empty?

Right click on the Field name and select "Properties" to modify the data type mapping



Output Field Properties x

Name: DischargeCFS

Alias: DischargeCFS

Type: Double

Properties

First

ANow WLL values	8
	2
	Yes

Merge Rule:

Delirriter:

## Create a Relationship Class

- Right click on geodatabase

• New

-->

Relationship

Class

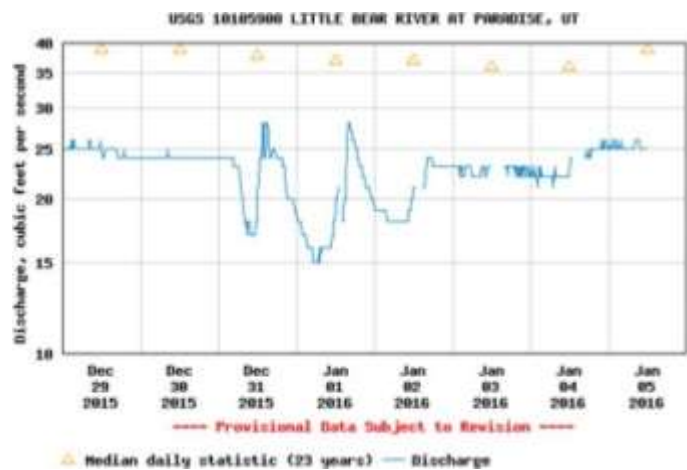


Table					
discharge					
		s.tec	Date	DischargeCFS	
2	USGS	10104600		13 A	
3	USGS	totoeoo		13 A	



Connect streamflow gage locations to discharge values measured at those sites

4	USGS	'0104600		13 A
5	USGS	'0104600		13 LA
	USGS	'0104600		13 A 12 LA
	USGS	'0104600		121A

COcutcf 66305 Selected)

SitesTo

Table

streamflow\_gages

OBJECTID *	Shape *	STATION *	STATION_NA
1	Point	10104600	SOUTH FORK LITTLE BEAR RIVER NEAR AVON, UT
2	Point	10104700	LITTLE BEAR RIVER BELOW DAVENPORT CREEK NEAR AVON, UT
3	Point	10104900	EAST FORK LITTLE BEAR RIVER ABOVE RESERVOIR NEAR AVON, UT
4	Point	10105000	EAST FORK LITTLE BEAR RIVER NEAR AVON, UT
5	Point	10105800	LITTLE BEAR RIVER AT PARADISE, UT
6	Point	10106000	LITTLE BEAR RIVER NEAR PARADISE, UT
7	Point	10107400	LITTLE BEAR RIVER NEAR HYBLIM, UT

(0 out of 9 Selected)

streamflow\_gages discharge

SitesToDischarge

CJ streamflow\_gages (origin)

discharge (destination)

# Origin and Destination

New Relationship Class

- Origin table/Feature

Nartz of tie

dass:

StesToOschæge

class:

"streamflow\_gages"

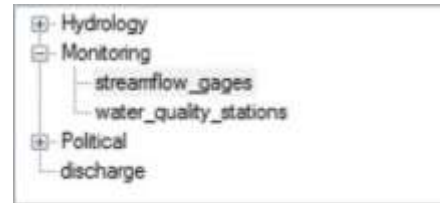
- Destination

table/Feature class:

"discharge"

Select the table/feature classes that will be associated by this relationship class.

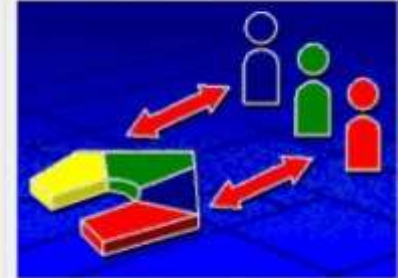
table/feature class:



Destination table/feature class:



A relationship class is a collection of relationships between objects in two tables/feature classes.



Parcels are owned by owners.  
Owners own parcels.

< Back

Next >

Cancel

Example: Can discharge

Relationship Types values exist in  
the

database if there is no

- Simple related gage site where o No referential they were measured? integrity

- Both tables can exist independent of the other

- Composite

- Existence of records in the destination table depend on a related record in the source table

# Labels and Message Propagation

Relationship Class

Select the type of relationship that this relationship class will store:

☐ Simple (peer to peer) relationship

Simple or peer-to-peer relationships are relationships that exist between two or more objects in the database that can exist independent of each other. In this kind of relationship, when the object(s) in the origin table/feature class are deleted, the related object(s) in the destination table/feature class are not deleted.

☒ Composite relationship

Composite relationships are relationships where the lifetime of the object(s) in the destination table/feature class are controlled by the lifetime of their related object in the origin table/feature class. When the object in the origin table/feature class is deleted, the related object(s) in the destination table/feature class are also deleted.

< Back Next > Cancel



- When you select a gage site, do you want related discharge values to be automatically selected?
- When you select discharge values, do you want the related site(s) to be automatically selected?

New Relationship Class

Specify a label for the relationship as it is traversed from the origin table/feature class to the destination table/feature class.

discharge

Specify a label for the relationship as it is traversed from the destination table/feature class to the origin table/feature class.

streamflow\_gages

Which messages be propagated between the object record by this ?

- C) "n to deâhün)
- C) Backward to
- @ecth
- C) messages

< Back

Next >

# Cardinality

1— 1	one to one
is one re ate	record

- : There in each table on both sides of the

I—M	one to man
recor	in eta e
t	on

M — N	man to man
comp ex re ations ip	
t	

relationship

## New Relationship Class

the fry is relationship class (origin - destination).

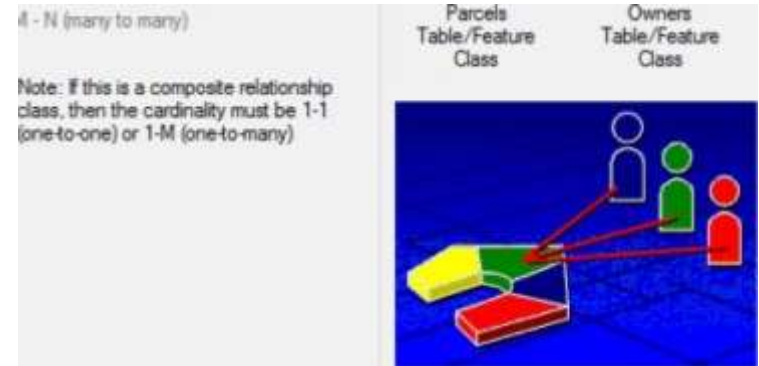
O I • 1 Ole to me)

h a I-M to many) ationship ,each  
c\*ct be n relaed me or.gn to

muhple ble /feature class in

thecan estnat•cn class.

@1-M (one to m.w)



is a

< Back

Next >

Cancel

- : One side of the relationship is related to many records in the table on the other side

- : A table requires a bridge table

# Attributes for the Relationship

- Can be stored in an additional table
- Not required for most relationships
- A new "bridge" table is required for "many-to-many" relationships

New Relationship Class

a new table to be created to include dttdes with then a table wa be created to them.

Do you wish to add &trtb-nes to this relationship class?

☐ Yes. I like to add to this relationship class.
 ☐ No. I do want to add to this class.

# Primary and Foreign Keys

- The attribute columns on which the tables are linked
- Do not have to have the same name
- Must have the same data type

new Relationship

Sea k" tū.'fe\*n d" be t}.

Sect k. n

Select the primary key field in the origin table/feature class:

STATION

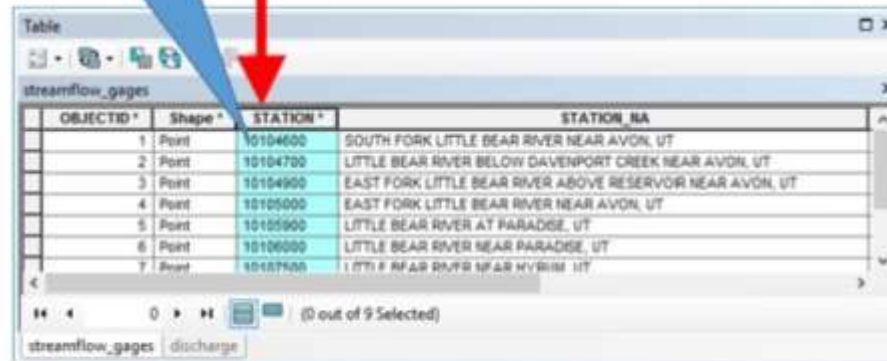
Select the foreign key field in the destination table/feature class that refers to the primary key field in the origin table/feature class:

SteCode

< Back Next > Cancel

Origin  
Primary Key

SitesTo



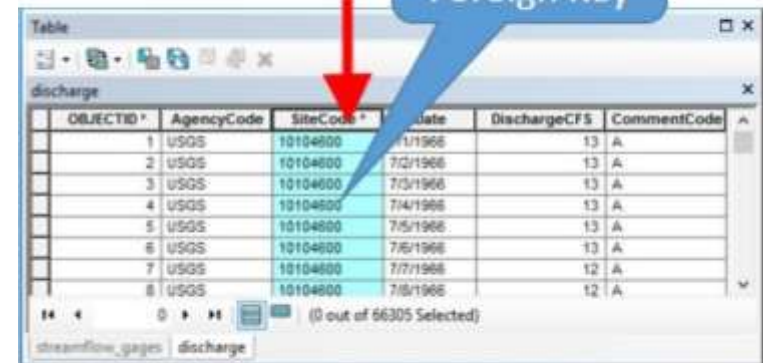
OBJECTID *	Shape *	STATION *	STATION_NA
1	Point	10104600	SOUTH FORK LITTLE BEAR RIVER NEAR AVON, UT
2	Point	10104700	LITTLE BEAR RIVER BELOW DAVENPORT CREEK NEAR AVON, UT
3	Point	10104900	EAST FORK LITTLE BEAR RIVER ABOVE RESERVOIR NEAR AVON, UT
4	Point	10105000	EAST FORK LITTLE BEAR RIVER NEAR AVON, UT
5	Point	10105900	LITTLE BEAR RIVER AT PARADISE, UT
6	Point	10106000	LITTLE BEAR RIVER NEAR PARADISE, UT
7	Point	10107500	LITTLE BEAR RIVER NEAR HYSLOM, UT

streamflow\_gages (origin)

SitesToDischarge

range

Destination  
Foreign Key



OBJECTID *	AgencyCode	SiteCode *	Date	DischargeCF5	CommentCode
1	USGS	10104600	7/1/1966	13	A
2	USGS	10104600	7/2/1966	13	A
3	USGS	10104600	7/3/1966	13	A
4	USGS	10104600	7/4/1966	13	A
5	USGS	10104600	7/5/1966	13	A
6	USGS	10104600	7/6/1966	13	A
7	USGS	10104600	7/7/1966	12	A
8	USGS	10104600	7/8/1966	12	A

discharge (destination)



# Result

EJ 13 Lecture9.gdb

- Three Feature B Hydrology Datasets  
watershed
- Five Feature ClassesE) Monitoring

EI

- Three Raster streamflow\_gages water\_quality\_stations  
Datasets Political
- One Table Utah discharge
- One Relationship Class ned\_30

ned\_30hlshd  
nlcdgrid\_1992  
SitesToDischarge

## Advantages of File Geodatabases vs. Personal Geodatabase and Shapefiles

- Compatible with Windows and Linux platforms
- Optimized performance (outperform in both size limitations and speed)
- Storing both vector and raster datasets together
- Less restrictive editing locks (per table rather than entire database)

- Can use data compression to reduce storage requirements
- Data integrity tools (e.g., coded domains, relationships)
- Simpler to create relationships

## Summary

- Design considerations for geodatabases include
  - Geodatabase type
  - Thematic layers to include
  - Geographic representation of thematic layers
  - Selection of object types, attributes and relationships
- Geodatabases enable organization of thematic layers

(e.g., Feature Datasets)

- ArcGIS provides tools for creating geodatabases and importing data
- A number of geodatabase data models are already available for use
- Geodatabase design and implementation is an iterative process that may need to be refined over time

## Additional Resources

- Geodatabase design steps:

<http://desktop.arcgis.com/en/desktop/latest/manage-data/geodatabases/geodatabase-designsteps.htm>

# Assignment 1— Building a Geodatabase

- Finish building the geodatabase we began today in class o Feature datasets and feature classes
  - Raster datasets
  - Object class (table) and relationship
- Choose a study area of some sort
- Create a new file geodatabase for your study area
- Create 1 or more Feature Datasets
- Create or load at least 2 feature classes

- Download the NED DEM for your study area and load it into your geodatabase