

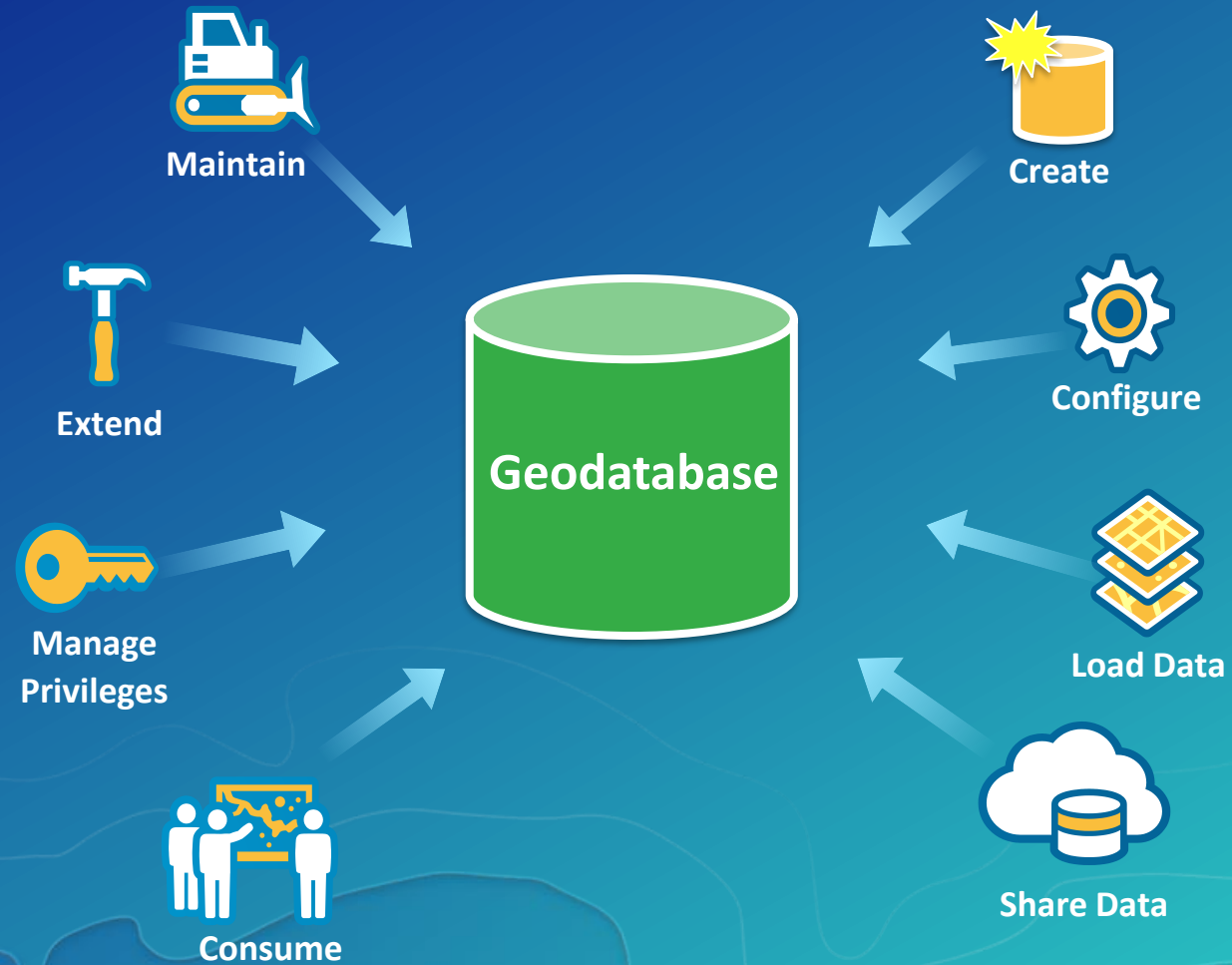
Geodatabase Best Practices

Dave Crawford | Erik Hoel

Geodatabase best practices - outline

- Geodatabase creation
- Data ownership
- Data model
- Data configuration
- Geodatabase behaviors
- Data integrity and validation
- Administration
- Extending
- Performance

Geodatabase Lifecycle



Geodatabase – extending the database

- Database
 - Tables and native geometry types
 - Transaction / security model
- Geodatabase
 - Information model, transaction model, metadata tables, and an API / SDK
- Benefits
 - Improved data quality
 - Editing efficiency
 - Web model
 - API / SDK

Geodatabase – transaction models

- Transaction models
 - Single User - one editor at a time
 - Multi-user - versioning supports long transactions
- Archiving
 - Built on top of versioning
 - Enables historical reference of what has happened to data over time
 - Enables a time slider in ArcMap and Pro

Geodatabase best practices

- **Geodatabase creation**

- Data ownership
- Data model
- Data configuration
- Geodatabase behaviors
- Data integrity and validation
- Administration
- Extending
- Performance

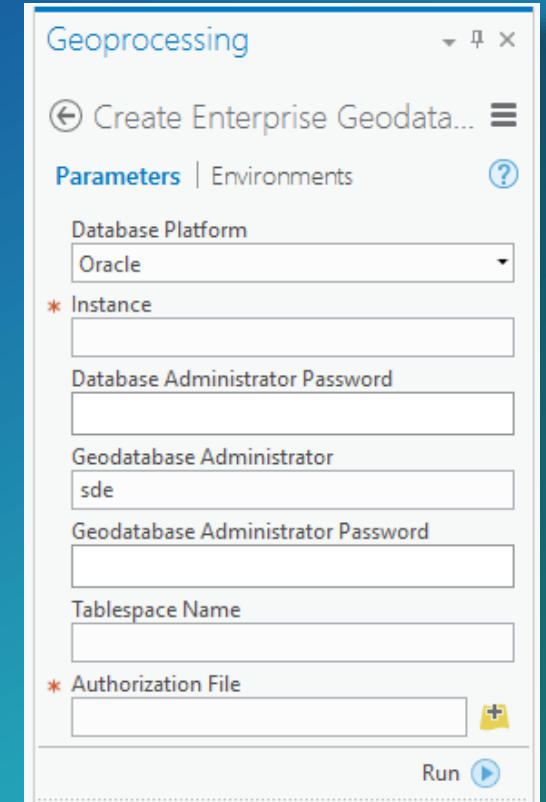
Geodatabase types

- What type to use?
 - Single User
 - File Geodatabase
 - Multi-user
 - Enterprise Geodatabase
 - Database Server Desktop/Workgroup
- Considerations
 - Size
 - Functionality
 - Concurrent viewers/editors
 - Storage platform
 - Plans for publishing/services

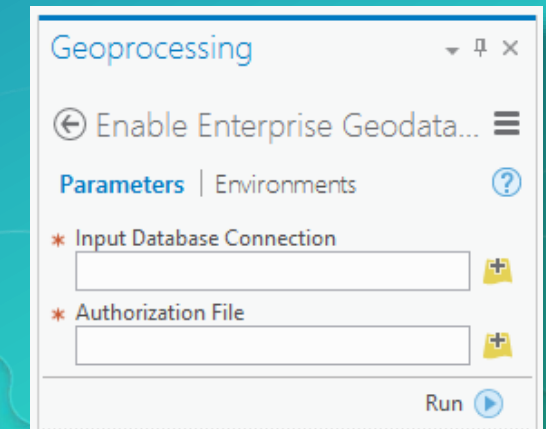
Characteristics	Personal geodatabase	File geodatabase	Desktop geodatabase	Workgroup geodatabase	Enterprise geodatabase
Functionality	Original Desktop format	Improved Desktop format	Distributed data or project-level use	Departmental projects or small organizations	Large capacity and user base
Storage mechanism	Microsoft Access database (.mdb)	File folder; displays .gdb extension in ArcCatalog	Microsoft SQL Server Express	Microsoft SQL Server Express	SQL Server, Oracle, PostgreSQL, DB2, Informix
Storage limit	2 GB per geodatabase; effective limit ~500 MB	1 terabyte (TB) per object, configurable to 256 TB	10 GB per database server	10 GB per database server	Limited by relational database and hardware
User limit	One editor per database	One editor per object	Three concurrent users, one can edit	Ten concurrent users, all can edit	Unlimited
Platform	Windows	Any	Windows	Windows	Any
Licensing: ArcGIS for Desktop	All ArcGIS for Desktop licenses can view, create, edit, and manage.		<ul style="list-style-type: none"> • All ArcGIS for Desktop licenses can view. • Standard or Advanced license required to create, edit, and manage. 		
Licensing: ArcGIS for Server	Not required	Not required	Not required	Workgroup Edition	Enterprise Edition

Geodatabase creation

- Single user geodatabases
 - Create File Geodatabase
- Multi-user geodatabases
 - Create Enterprise Geodatabase
 - All in one (creates database, geodatabase admin account, etc.)
 - Requires access to DBA account
 - Enable Enterprise Geodatabase
 - No access to DBA account
 - Two steps - setup database and then enable the geodatabase



The screenshot shows the 'Create Enterprise Geodatabase' dialog box in the Geoprocessing environment. The 'Parameters' tab is active. The 'Database Platform' is set to 'Oracle'. The 'Instance' field is empty. The 'Database Administrator Password' field is empty. The 'Geodatabase Administrator' is set to 'sde'. The 'Geodatabase Administrator Password' field is empty. The 'Tablespace Name' field is empty. The 'Authorization File' field is empty with a plus icon to its right. A 'Run' button is at the bottom right.



The screenshot shows the 'Enable Enterprise Geodatabase' dialog box in the Geoprocessing environment. The 'Parameters' tab is active. The 'Input Database Connection' field is empty with a plus icon to its right. The 'Authorization File' field is empty with a plus icon to its right. A 'Run' button is at the bottom right.

Geodatabase creation – best practices

- For File Geodatabases
 - Use the geoprocessing tool – supported across the platform
- For Enterprise Geodatabases
 - Use the Create Enterprise Geodatabase geoprocessing tool
 - Oracle – prerequisite of using an existing database
 - Fallback to the Enable Enterprise Geodatabase geoprocessing tool

Geodatabase best practices

- Geodatabase creation
- **Data ownership**
- Data model
- Data configuration
- Geodatabase behaviors
- Data integrity and validation
- Administration
- Extending
- Performance

Data ownership considerations

- POLP – principle of least privilege
 - Database admin / geodatabase admin / data owner / end users

Geodatabase admin responsibilities:

- Ownership and maintenance on system tables
- Compress and DEFAULT version
- Upgrades
- Changes to DBTUNE/log file configurations

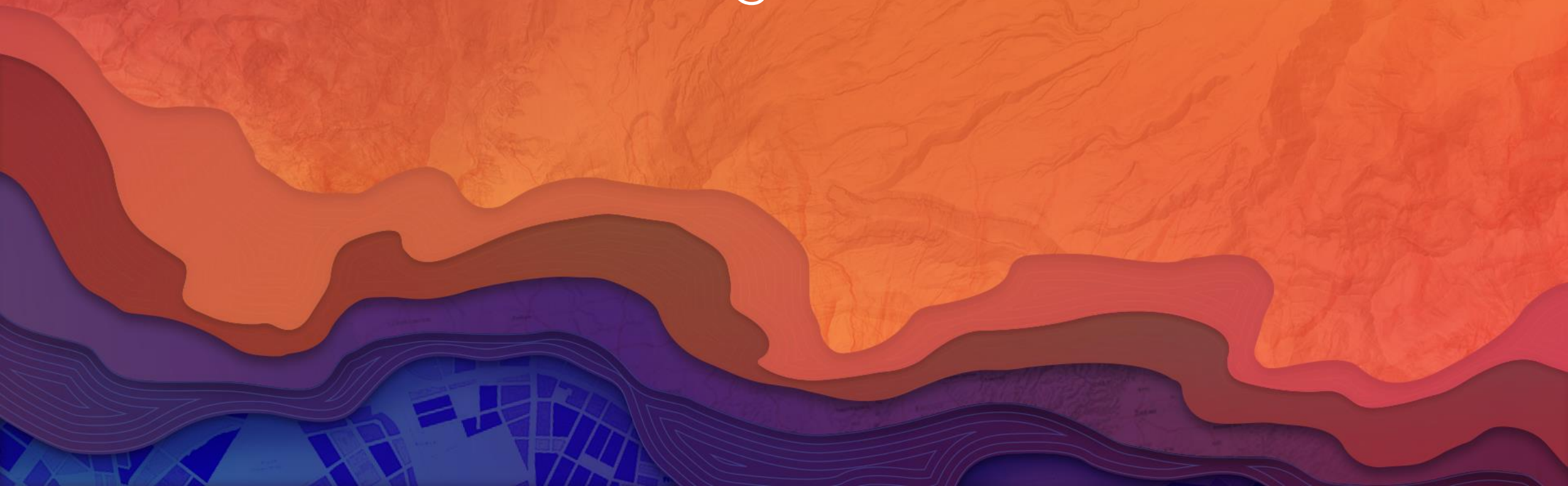
Data owner responsibilities:

- Manage privileges
- Schema changes
- Register data as versioned

Data ownership – best practices

- Grant the fewest permissions to the fewest number of users
- Have a unique Geodatabase Admin user
- Have a non-Geodatabase Admin create data

DEMO – Geodatabase creation and configuration



Geodatabase best practices

- Geodatabase creation
- Data ownership
- **Data model**
- Data configuration
- Geodatabase behaviors
- Data integrity and validation
- Extending
- Performance

Data model

What is a data model?

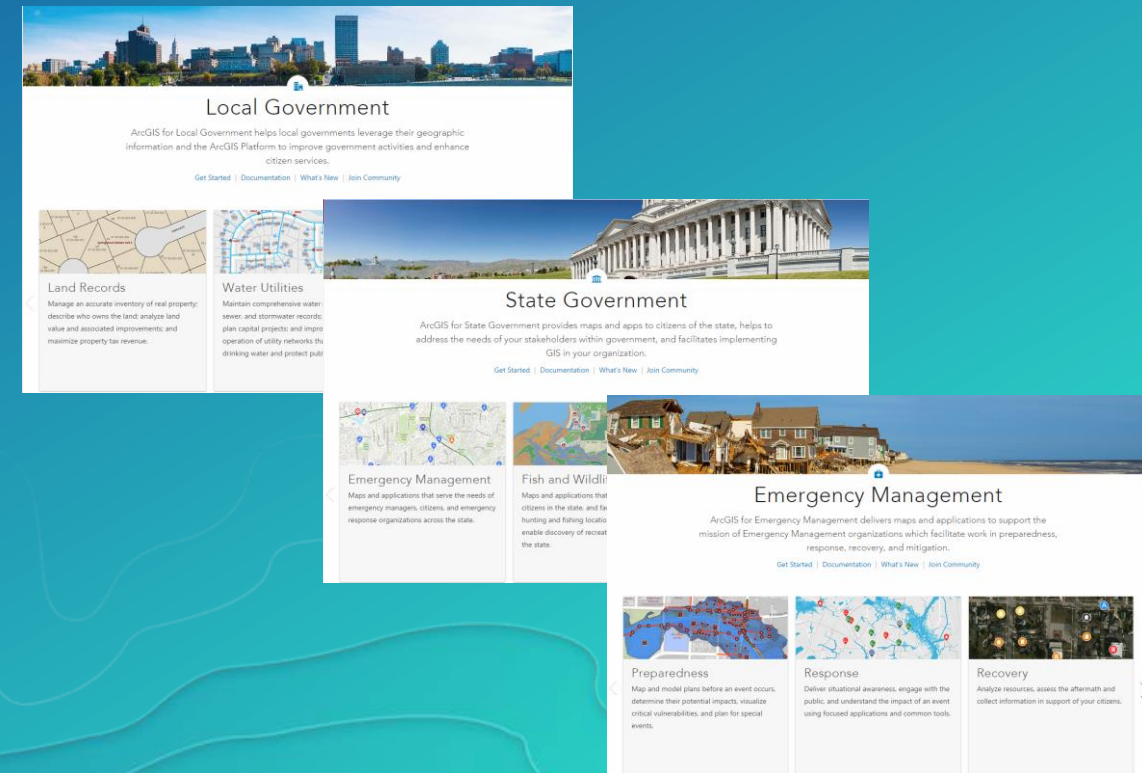
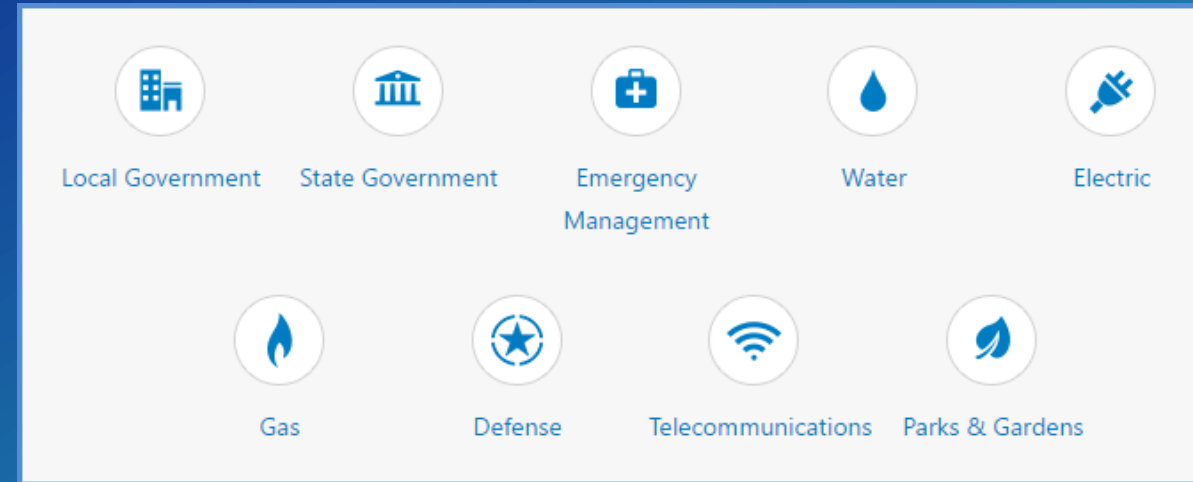
- Collection of things in a Geodatabase
- Feature classes, attributes, subtypes, domains, etc.

Picking a data model

- What data are you representing
- ArcGIS data models
 - <http://solutions.arcgis.com/>
- Other published data models

Custom data model

- Is your data unique?
- Do you just not like an existing data model?



Data model – best practices

Use an ArcGIS Data Model

<http://solutions.arcgis.com/>

Use a data model provided by a business partner

Build a custom data model (if):

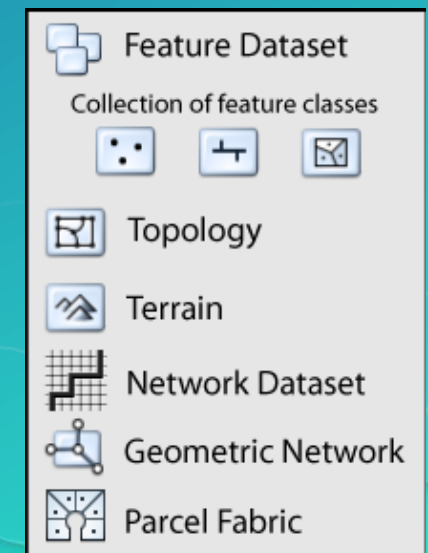
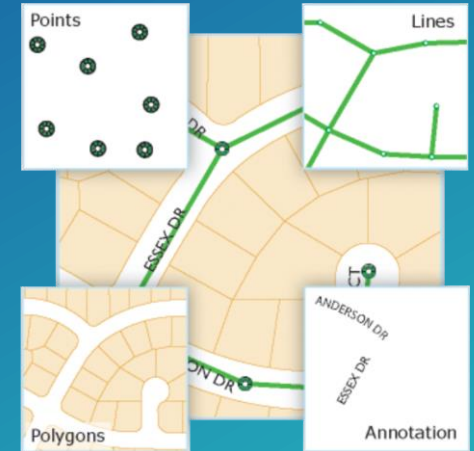
- You have an understanding of your data
- You have an understanding of the geodatabase
- You have an understanding of how data model decisions will impact the entire platform
- You follow the guidelines provided in the following section

Geodatabase best practices

- Geodatabase creation
- Data ownership
- Data model
- **Data configuration**
- Geodatabase behaviors
- Data integrity and validation
- Administration
- Extending
- Performance

Data Configuration – Prototype everything!

- Minimize the number of Feature Classes
 - Group together similar features
 - Find a balance between grouping and null or empty attributes
- Use a Feature Dataset when you want to:
 - use a Controller Dataset
 - enforce common spatial reference / privileges
 - group similar themed classes
- Field data types and field order:
 - Understand your data when choosing field types:
 - Consider field order during data creation
 - Keep text fields short, expand later
 - Will data be entered via web/mobile clients?



Geodatabase Best Practices

- Geodatabase creation
- Data ownership
- Data model
- Data configuration
- **Geodatabase behaviors**
- Data integrity and validation
- Administration
- Extending
- Performance

Subtypes

- Modeling multiple types of features in a single feature class
- Features that shares similar attributes
- Short or Long Integer field - short is almost always appropriate

Default values

- Most fields in a feature class or table support a default value
 - This is a value entered during editing, if the user does not specify a value
 - Raster fields do not support default values
- Default values can be configured at the field level or subtype level
- Once subtypes are created – default values should be configured at subtype level
 - Subtype "Commercial" can have a default value of 2 for a field
 - Subtype "Other" can have a different default value for the same field

Domains

- Define a set of permissible values
- Range domains permit values to be valid in a field when they are within the specified range
- Coded value domains
 - Description (what you see in ArcMap or Pro) and code (the value that is stored)
 - In ArcGIS editing tools, drop downs are provided with valid descriptions
 - Field calculation - code can be used
- Validation will compare actual values against their domain to determine if the entries are correct

Geodatabase behavior – best practices

- Subtypes
 - Use subtypes where possible, instead of adding additional feature classes
 - If features share a large majority of their attributes with other features
 - Group the features into a single class, differentiated by subtype
- Default values
 - Always setup default values to limit the <null> entries in a table
 - Editing using web/mobile? Limits the amount of entry required
- Domains
 - Use domains to simplify editing
 - Use domains to make sure valid values are being entered

Relationship classes

- A persisted relationship between objects in different feature classes or tables
- Visible using the attribute dialog in ArcMap and ArcGIS Pro
- Support rules for data integrity
- No relationship rules (constraints) = no validation
 - Business logic by defining constraints on the classes' objects or features that is accessed during validation
 - Setup after relationship class creation

Relationship classes – best practices

- Use relationship classes for those relationships that will enforce some data integrity
- Create rules to further ensure data integrity
 - These rules are not reactively checked
- Avoid creating relationship classes between two feature classes residing in separate feature datasets

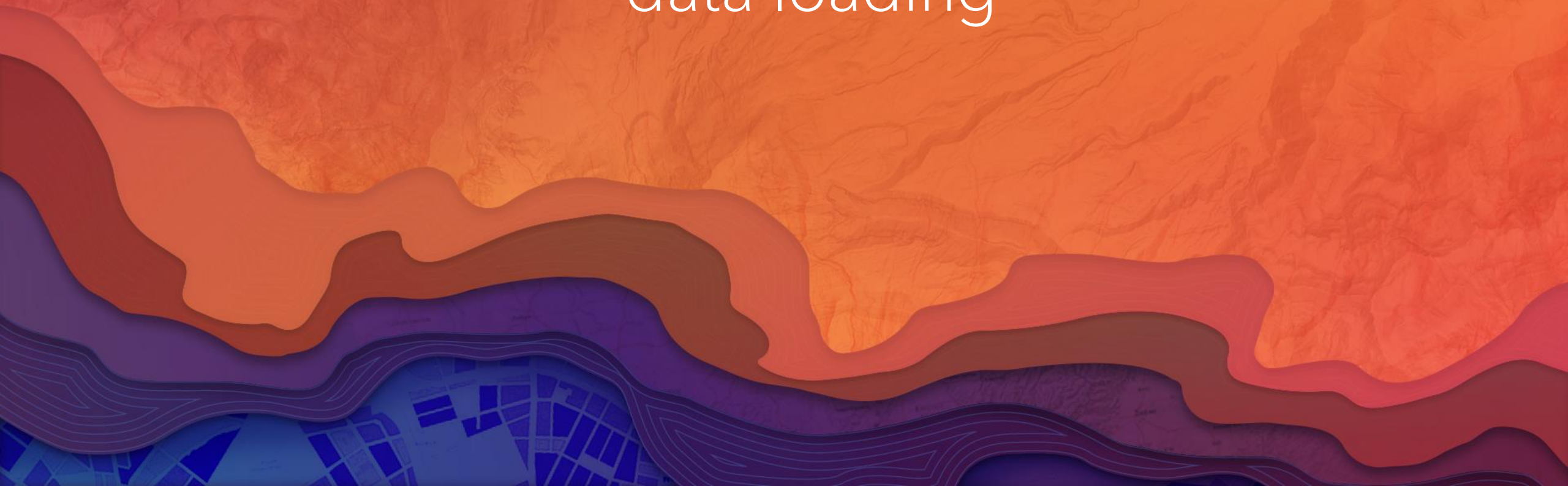
Data loading - tools

- Loading new objects
 - Import/Export (XML Workspace Document, Feature Class to Feature Class)
 - Copy tool or manual copy/paste
- Updating existing objects:
 - Simple data loader vs. object loader
 - Object loader requires an edit session but triggers geodatabase behavior
 - Object loader trigger validation against geodatabase rules during load
 - Object load may be undone as it is inside an edit session
 - Append (requires a destination feature class or table to already exist)

Data loading – best practice

- Load all data before registering as versioned – this limits records in deltas
- If you need to load data into a versioned system, consider unregistering as versioned prior to loading the data

DEMO – Geodatabase behaviors and data loading



Geodatabase best practices

- Geodatabase creation
- Data ownership
- Data model
- Data configuration
- Geodatabase behaviors
- **Data integrity and validation**
- Administration
- Extending
- Performance

Attribute integrity while editing

- The Geodatabase provides functionality to help maintain attribute quality while editing
 - Attribute domains
 - Subtypes
 - Default values for fields
 - Relationship classes

Attribute integrity while editing

- The Geodatabase provides functionality to help maintain attribute quality while editing
 - Attribute domains
 - Subtypes
 - Default values for fields
 - Relationship classes

Domains: Montgomery								
Domain Name	Description	Field Type	Domain Type	Split Policy	Merge Policy		Code	Description
AncillaryRoleDomain		Short	Coded Value Domain	Default	Default		CI	Cast iron
AnnotationStatus	Valid annotation state values.	Short	Coded Value Domain	Duplicate	Default		DI	Ductile iron
BooleanSymbolValue	Valid values are Yes and No.	Short	Coded Value Domain	Duplicate	Default		PVC	PVC
DistDiam	Valid diameters for distribution mains	Float	Coded Value Domain	Default	Default		ACP	Asbestos concrete
EnabledDomain		Short	Coded Value Domain	Default	Default		COP	Copper
FittingType	Valid fitting type codes	Text	Coded Value Domain	Duplicate				
HorizontalAlignment	Valid horizontal symbol alignment values.	Short	Coded Value Domain	Duplicate				
Material	Valid pipe materials	Text	Coded Value Domain	Duplicate				
RotAngle	Valid rotation angles	Long	Range Domain	Duplicate				
TransDiam	Valid diameters for transmission mains	Float	Coded Value Domain	Duplicate				
VerticalAlignment	Valid symbol vertical alignment values.	Short	Coded Value Domain	Duplicate				
ZoningCodes	Simplified zoning codes	Text	Coded Value Domain	Default				

Attributes

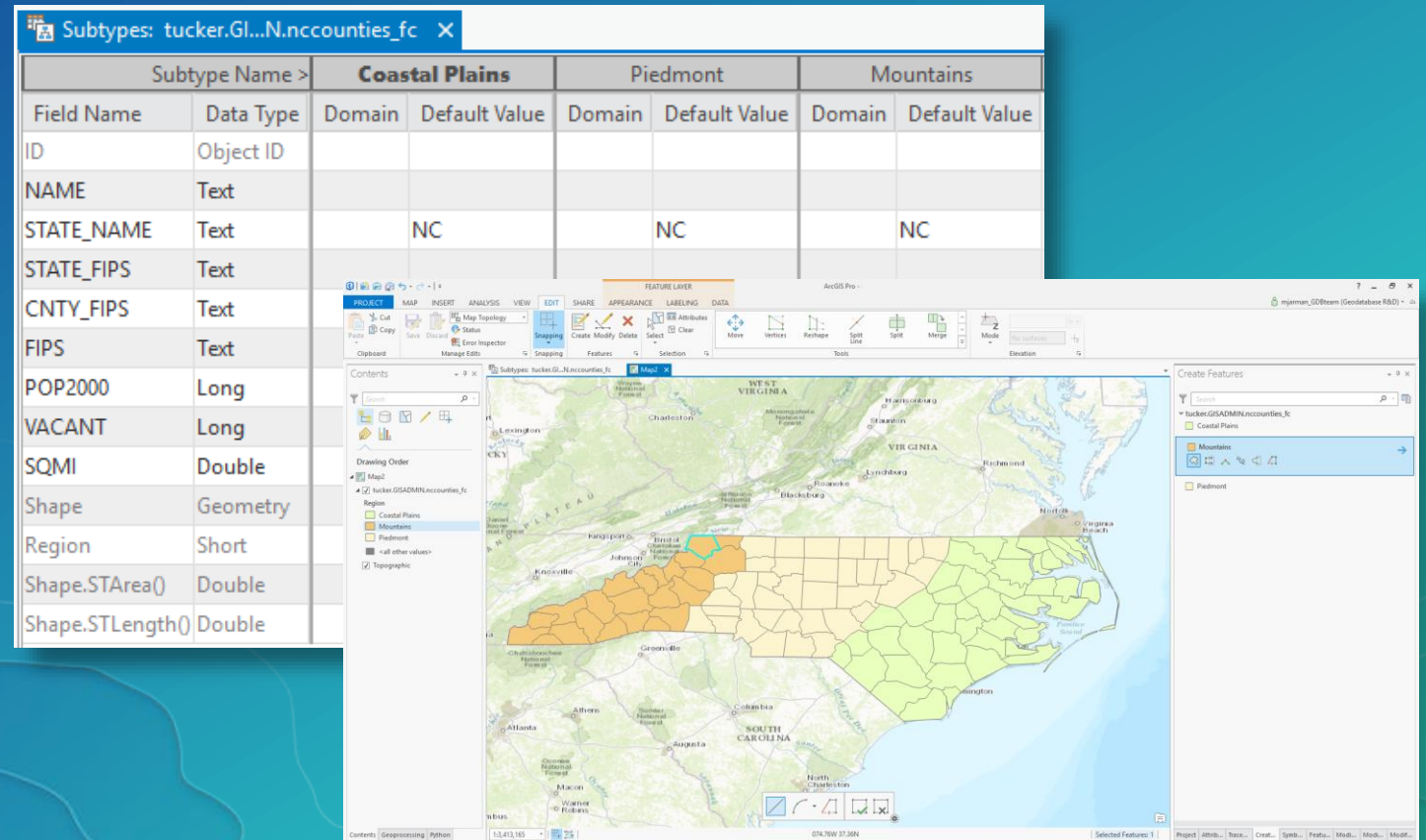
Transmission mains
35

Attributes | Geometry

MUNICIPALITY	<Null>
CONTRACTOR	<Null>
OWNER	<Null>
MATERIAL	Ductile iron
DIAMETER	<Null>
PARALLEL	Asbestos concrete
JOINT_TYPE	Cast iron
RECORDED_LENGTH	Copper
COVER_TYPE	Ductile iron
	PVC
	<Null>

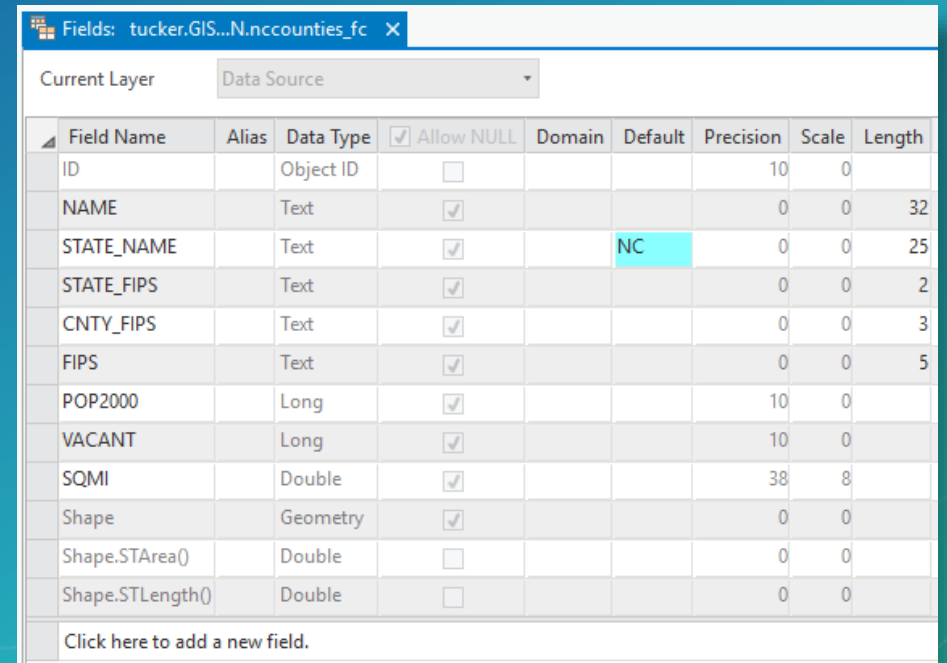
Attribute integrity while editing

- The Geodatabase provides functionality to help maintain attribute quality while editing
 - Attribute domains
 - Subtypes
 - Default values for fields
 - Relationship classes



Attribute integrity while editing

- The Geodatabase provides functionality to help maintain attribute quality while editing
 - Attribute domains
 - Subtypes
 - **Default values for fields**
 - Relationship classes
- **Attribute Assistant**
 - Editor add-in/extension from Solutions team
 - Reduced mouse clicks during editing
 - Accurate attribute info

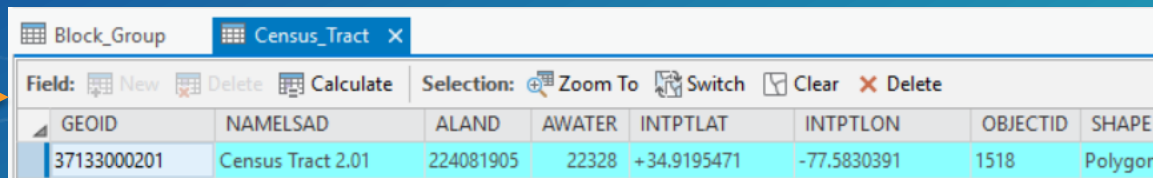


Field Name	Alias	Data Type	<input checked="" type="checkbox"/> Allow NULL	Domain	Default	Precision	Scale	Length
ID		Object ID	<input type="checkbox"/>			10	0	
NAME		Text	<input checked="" type="checkbox"/>			0	0	32
STATE_NAME		Text	<input checked="" type="checkbox"/>		NC	0	0	25
STATE_FIPS		Text	<input checked="" type="checkbox"/>			0	0	2
CNTY_FIPS		Text	<input checked="" type="checkbox"/>			0	0	3
FIPS		Text	<input checked="" type="checkbox"/>			0	0	5
POP2000		Long	<input checked="" type="checkbox"/>			10	0	
VACANT		Long	<input checked="" type="checkbox"/>			10	0	
SQMI		Double	<input checked="" type="checkbox"/>			38	8	
Shape		Geometry	<input checked="" type="checkbox"/>			0	0	
Shape.STArea()		Double	<input type="checkbox"/>			0	0	
Shape.STLength()		Double	<input type="checkbox"/>			0	0	

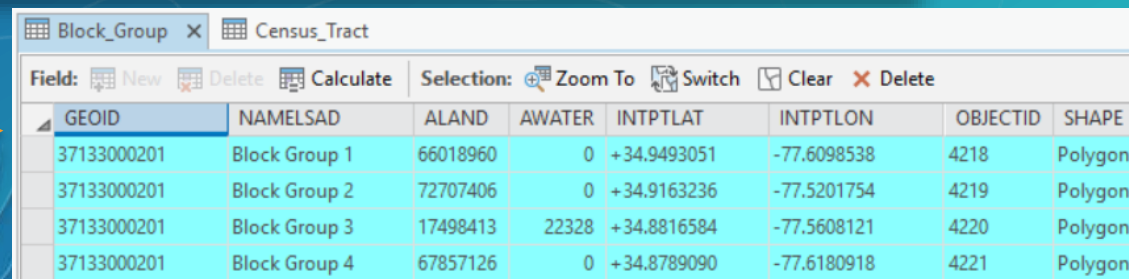
[Click here to add a new field.](#)

Attribute integrity while editing

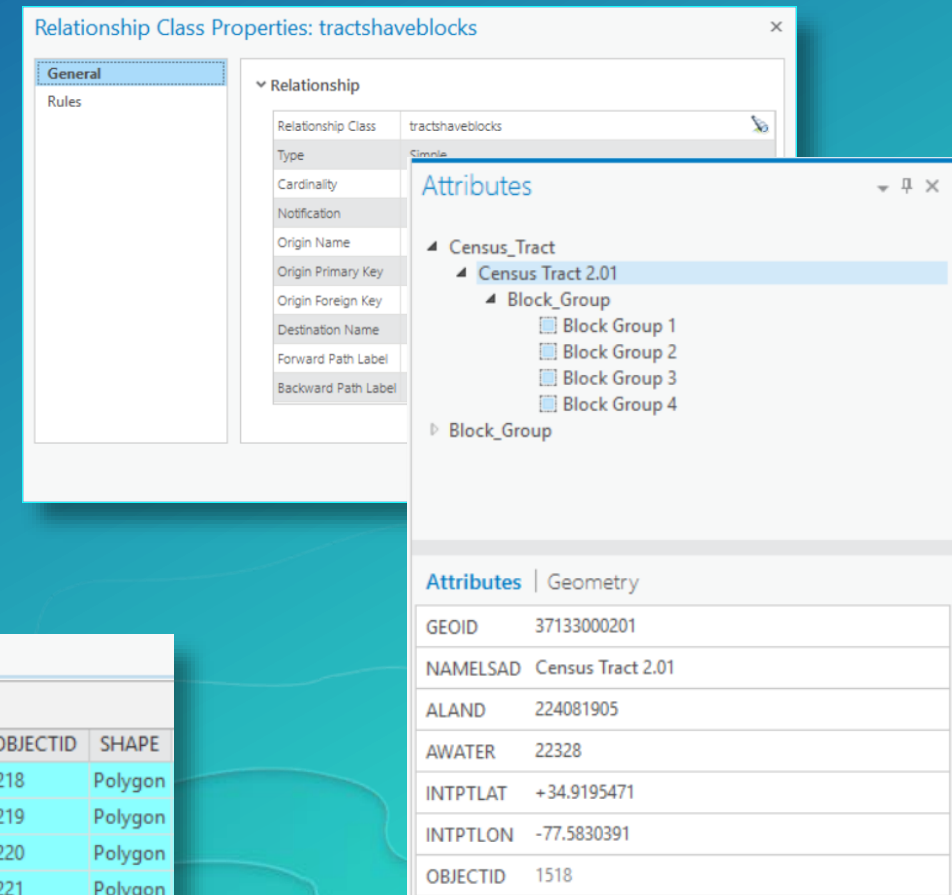
- The Geodatabase provides functionality to help maintain attribute quality while editing
 - Attribute domains
 - Subtypes
 - Default values for fields
 - Relationship classes



GEOID	NAMELSAD	ALAND	AWATER	INTPTLAT	INTPTLON	OBJECTID	SHAPE
37133000201	Census Tract 2.01	224081905	22328	+34.9195471	-77.5830391	1518	Polygon



GEOID	NAMELSAD	ALAND	AWATER	INTPTLAT	INTPTLON	OBJECTID	SHAPE
37133000201	Block Group 1	66018960	0	+34.9493051	-77.6098538	4218	Polygon
37133000201	Block Group 2	72707406	0	+34.9163236	-77.5201754	4219	Polygon
37133000201	Block Group 3	17498413	22328	+34.8816584	-77.5608121	4220	Polygon
37133000201	Block Group 4	67857126	0	+34.8789090	-77.6180918	4221	Polygon



Relationship Class Properties: tractshaveblocks

General

Rules

Relationship

Relationship Class: tractshaveblocks

Type: Simple

Cardinality:

Notification:

Origin Name:

Origin Primary Key:

Origin Foreign Key:

Destination Name:

Forward Path Label:

Backward Path Label:

Attributes

Census_Tract

Census Tract 2.01

Block_Group

Block Group 1

Block Group 2

Block Group 3

Block Group 4

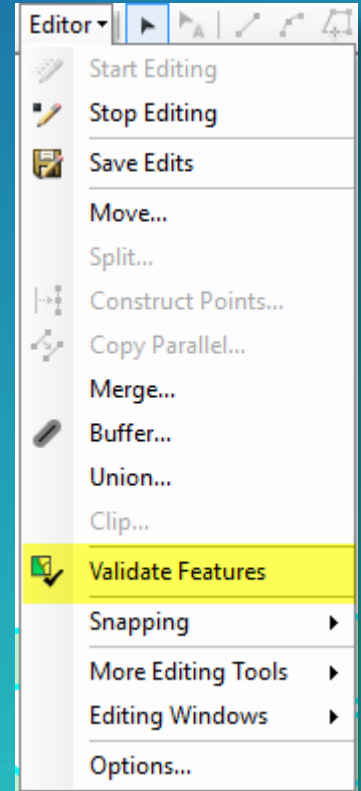
Block_Group

Attributes | Geometry

GEOID	37133000201
NAMELSAD	Census Tract 2.01
ALAND	224081905
AWATER	22328
INTPTLAT	+34.9195471
INTPTLON	-77.5830391
OBJECTID	1518

Data validation

- Verifies an object against any rules that are defined for objects class or subtype
- Even with geodatabase behavior such as attribute domains there is still a need to perform 'Validate Features' to review data that violates rules
- Examples of editing operations that require validation:
 - Editing of range domains
 - Editing attribute domains (coded value/range) via field calculator
 - Bulk loading of data into existing feature classes

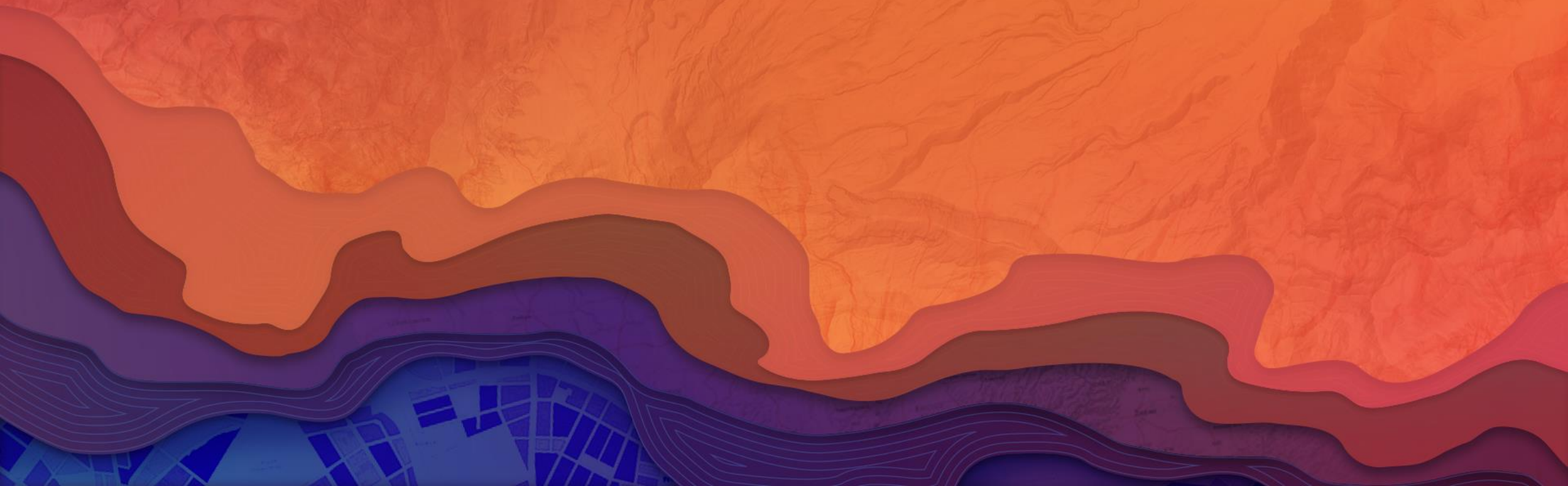


Validate Features - currently only in ArcMap
ArcGIS Pro - use Data Reviewer Extension

Data validation

- Multiple invalid features:
 - Warning/message received will not include the specific reason
 - Tip: make a single sub selection and re-validate
 - Validation process stops at the first rule where a feature is found invalid
 - Always re-validate
- Data Reviewer - ArcMap/Pro extension for a more interactive validation process
- Attribute Assistant - editor add-in/extension from Solutions team

DEMO – Editing and data validation

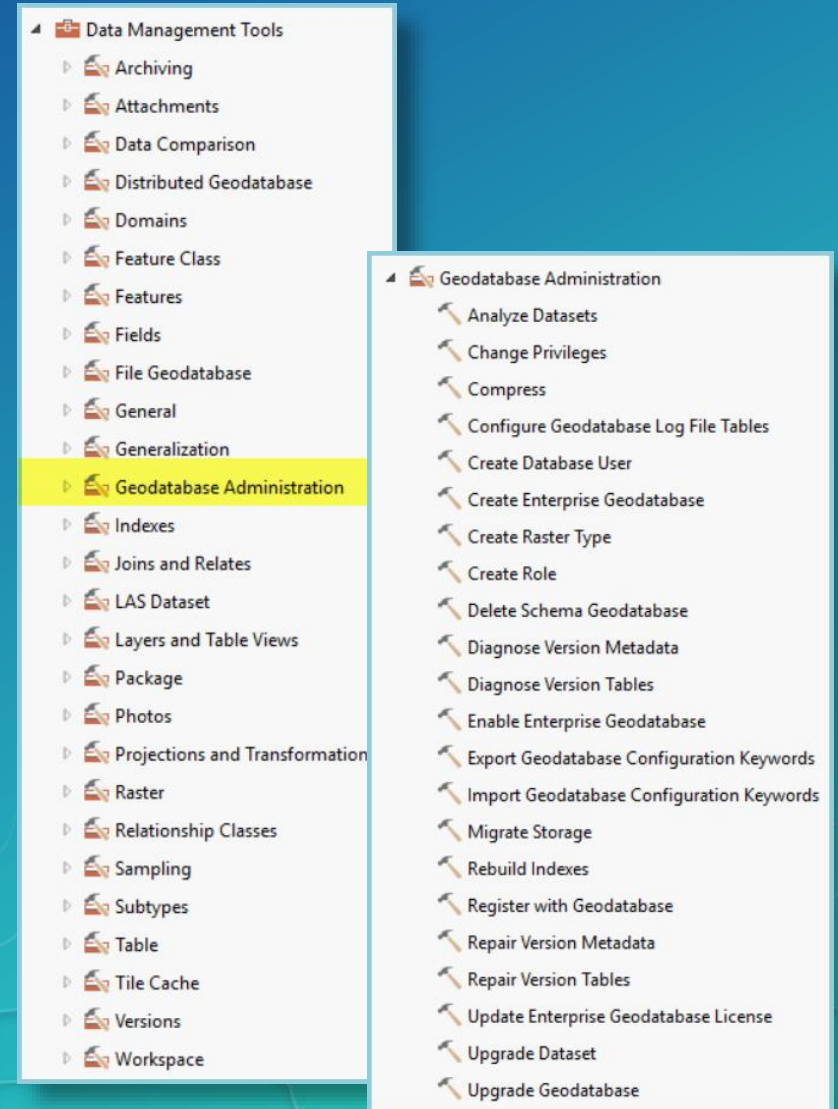


Geodatabase best practices

- Geodatabase creation
- Data ownership
- Data model
- Data configuration
- Geodatabase behaviors
- Data integrity and validation
- **Administration**
 - Extending
 - Performance

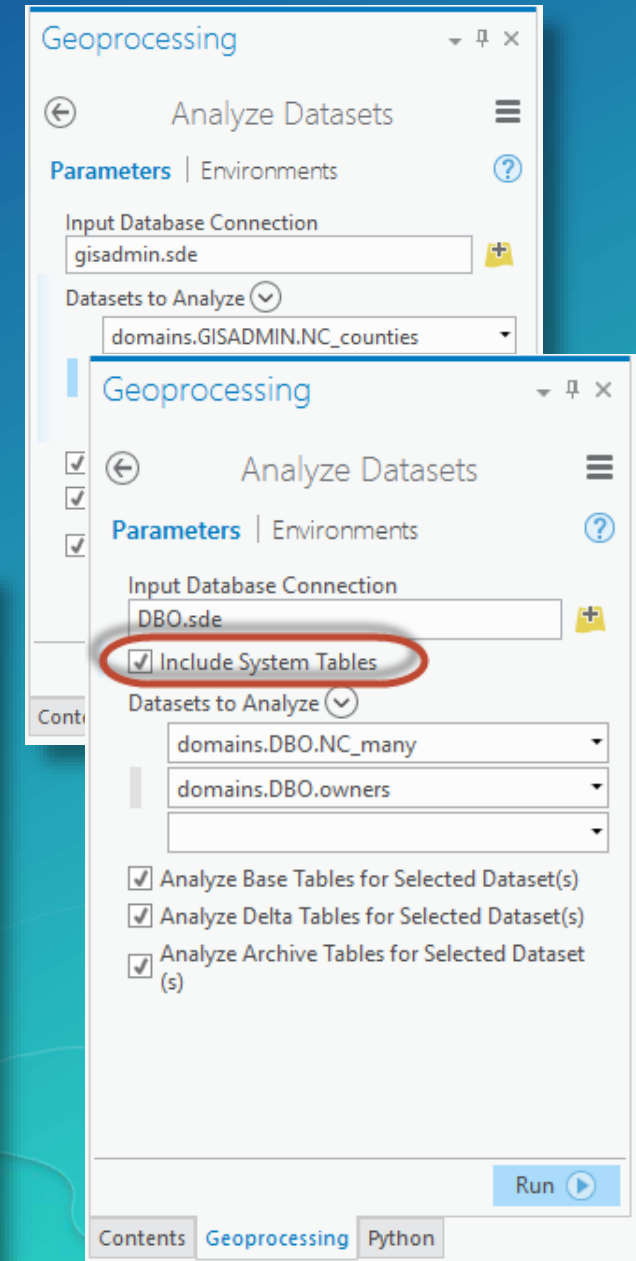
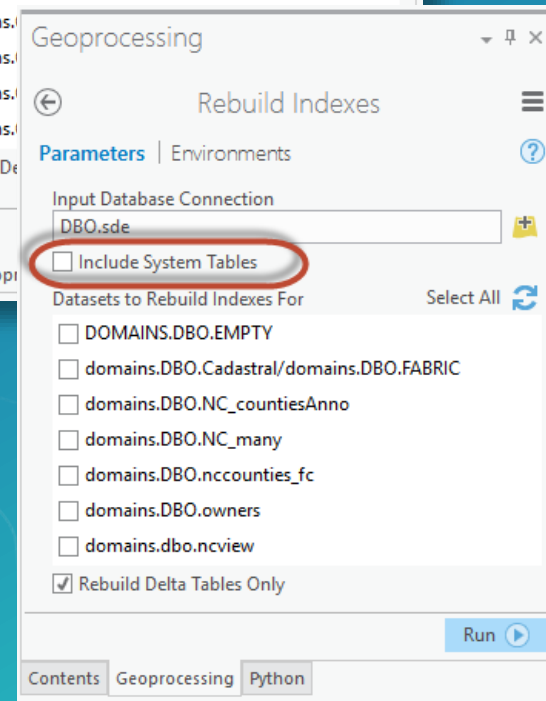
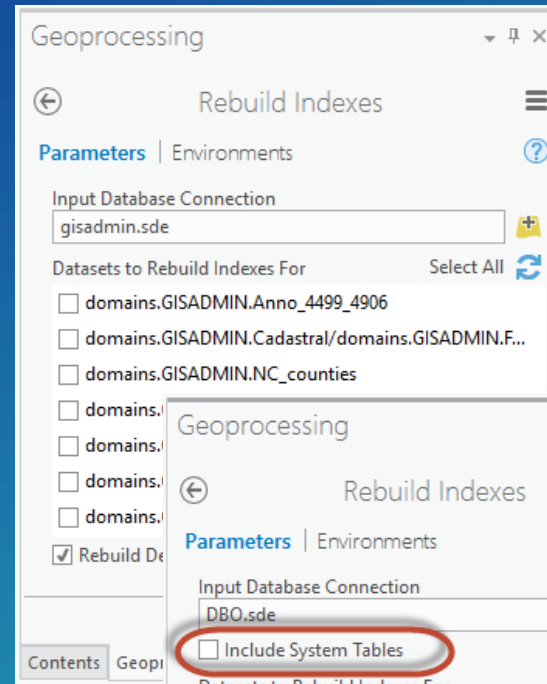
Enterprise geodatabase administration

- Geodatabase Administration toolbox
 - Geodatabase creation/upgrade
- Geodatabase configuration tasks
 - Logfile configuration
 - DBTUNE keyword customization
- Database/Geodatabase admin tasks
 - User/role creation
 - Rebuild indexes
 - Analyze datasets (update statistics)
 - Change privileges
 - Compress (versioned geodatabase)



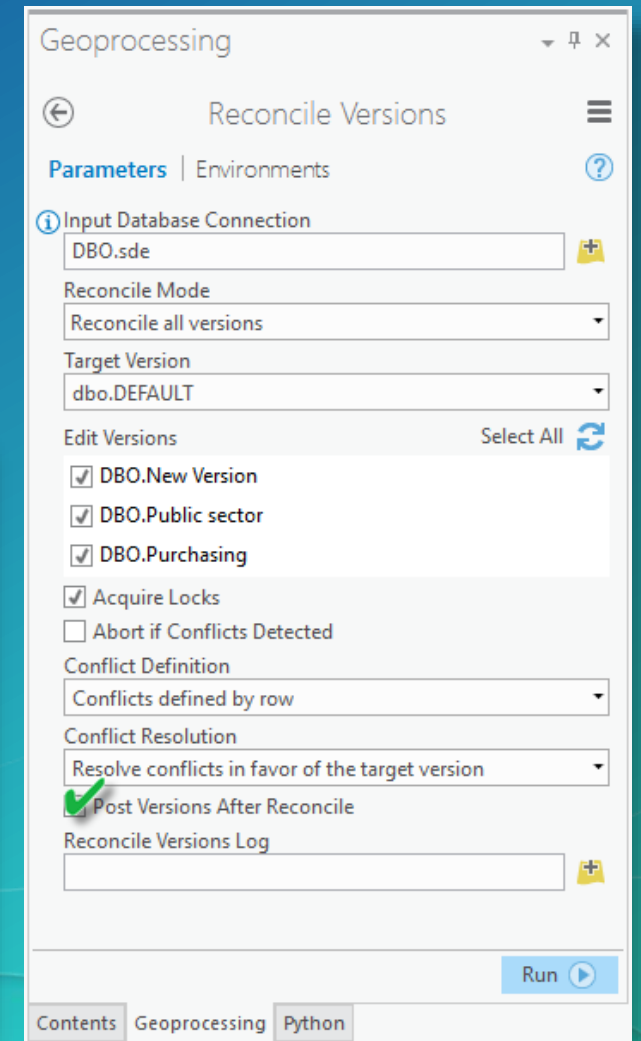
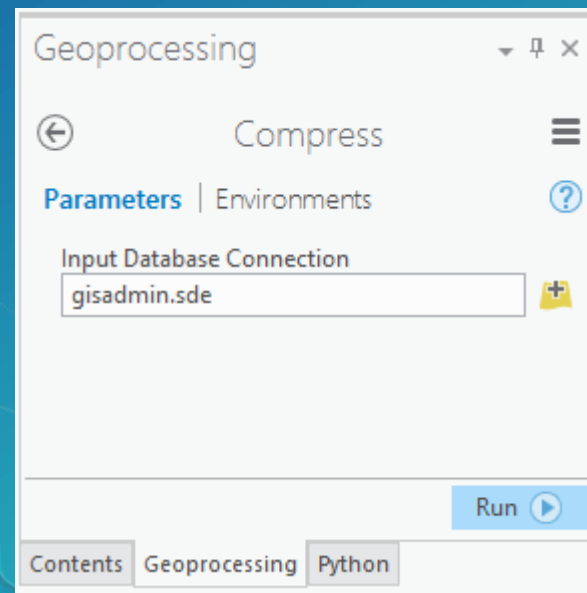
Enterprise geodatabase administration

- Rebuild indexes / analyze datasets
 - Data owner can analyze the tables they own
 - GDB Admin can analyze all tables, including system tables
 - Defaults to rebuild delta tables only



Enterprise geodatabase administration

- In order to maximally compress the geodatabase:
 - Remove locks (disconnect users, stop services consuming gdb datasets)
 - Reconcile, post, and delete versions
 - Synchronize any replicas
- Reconcile Versions tool
 - Batch reconciles when 'Post Versions After Reconcile' option is checked



DEMO – Database maintenance

Performance

- Fat/Wide feature classes vs. narrow feature classes
- Subtypes vs. new feature classes
- Lumping feature classes together
- Relationship class messaging – turn off for simple relationship classes
- Layers in a map
- Prototyping – What you are actually going to do with the data!
- Query Def

Geodatabase best practices - outline

- Geodatabase creation
- Data ownership
- Data model
- Data configuration
- Geodatabase behaviors
- Data integrity and validation
- Administration
- Extending
- Performance

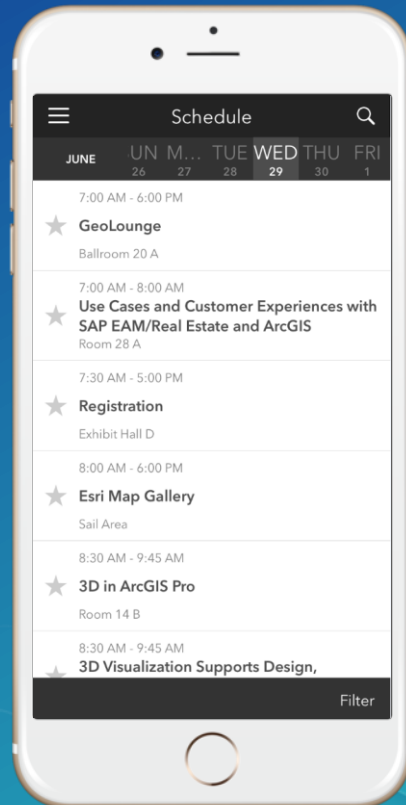
Please take our Survey

Your feedback allows us to help maintain high standards and to help presenters

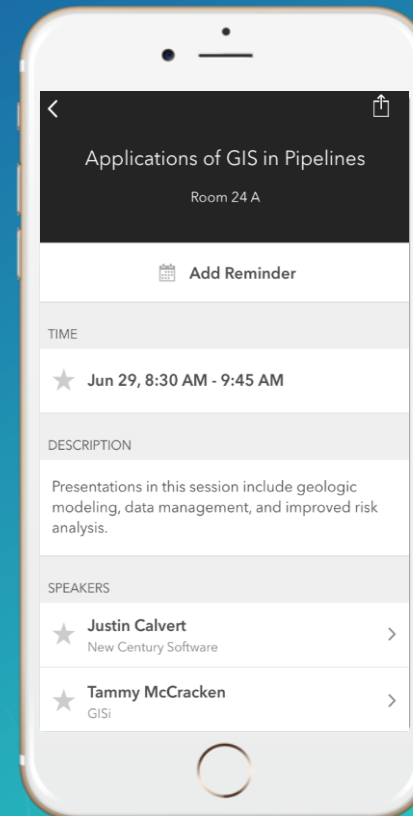
Find your event in the
Esri Events App



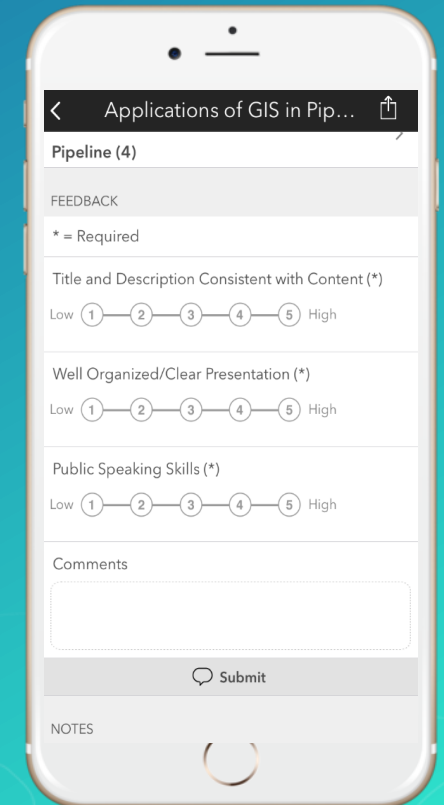
Find the session
you want to review



Scroll down to the
bottom of the session



Answer survey
questions and submit





esri

THE
SCIENCE
OF
WHERE