

Database Tables

get ready to get nerdy!

Terminology

- record = a row in a table
- field = a column in a table
- table = a “spreadsheet” in a database
- database = a collection of tables (a Geodatabase includes other stuff, too)

Tables

- two types of tables in ArcGIS
 - attribute tables
 - connected with explicit spatial data
 - standalone tables
 - not connected with explicit spatial data
 - can have *implicit* spatial data!

Tables

- attribute tables
 - has a set of mandatory fields (feature/object ID - FID/OID, etc.)
- standalone tables
 - has a single mandatory field (object ID - OID)

Tables

- standalone tables
- what do I mean by explicit and implicit spatial data?
 - explicit = connected to feature or spatial object
 - implicit = spatial data that has not been recognized as such
 - GPS coordinates not in a shapefile
 - place names
 - can be turned into explicit data via several processes
 - georeferencing! (later in course)

Tables

- standalone tables
- what do I mean by explicit and implicit spatial data?
 - explicit = connected to feature or spatial object
 - implicit = spatial data that has not been recognized as such
 - GPS coordinates not in a shapefile
 - place names
 - can be turned into explicit data via several processes
 - joining (later in lecture)
 - georeferencing! (later in course)

Records


- represents a single entity
 - attribute table = spatial entity
 - standalone = another entity
- must have a value or <null> for every field
- otherwise, exactly like an Excel spreadsheet row

Fields

- must be well-defined
- big difference from Excel
- means that each field can have only one data type and data length
 - when defining a field, must choose data type and data length

Fields

data types

Type	Definition	Example	Uses
Short	integers -32,000 to 32,000	15, 22, 87	integer, ordinal, nominal data (small numbers)
Long	Really big integers	230942323	integer (large numbers)
Float	smaller decimal numbers	1234023.22	ratio
Double	larger decimal numbers	2.35908E+13	ratio
String / Text	alphanumeric strings	Saint Louis	nominal, ordinal
Date	dates!	3/22/99	dates! (technically nominal)
Blob	binary large objects (images, documents, movies, etc.)		everything else (technically nominal)

Fields

data lengths

Don't worry about it (use default), except for strings

Make string field lengths as big as the biggest string you expect to store!

Lengths are defined by number of characters

Fields

data lengths

STATENAME field should be as long as the longest state name (~25)

Fields

- defining fields
 - traditional goal: choose the data type and data length that is just big enough, but not too big
 - too big = database will get huge
 - not big enough = won't be able to store your data
 - these days
 - most default values are fine
 - only worry about optimization on BIG databases

Fields

- ArcGIS table requirements
 - field names = 13 character max!
 - can make field aliases
 - saved with map project file
 - LAME!

Fields

- calculating fields
 - just like Excel
 - doing an example on the board...

Technology

- file formats
 - ArcGIS supports many types of tables (and databases)
 - DBase (DBF), MySQL, Oracle, Excel 2003 (.xls) (not .xlsx)
 - interface + requirements are always the same
 - abstraction!
 - don't have to worry about the peculiarities of each format

The screenshot shows an attribute table with the following columns: ObjectID, Shape, STATE_NAME, STATE_FIPS, SUB_REGION, STATE_ABBR, POP2000, POP2005, POP00, POP05, POP05_SQMI, and WHITE. Annotations with arrows point to specific fields: 'required' points to ObjectID, STATE_FIPS, and POP00; 'short int' points to STATE_ABBR; 'string (35)' points to STATE_NAME; and 'float' points to POP05_SQMI.

ObjectID	Shape	STATE_NAME	STATE_FIPS	SUB_REGION	STATE_ABBR	POP2000	POP2005	POP00	POP05	POP05_SQMI	WHITE
1	Polygon	Alabama	01	East South Central	AL	4447100	454149	1	88	3162808	
2	Polygon	Alaska	02	Pacific	AK	628932	660008	1	1.1	434534	
3	Polygon	Arizona	04	Mountain	AZ	5130832	588259	45	51.6	3873611	
4	Polygon	California	05	West South Central	CA	2973400	2796234	50.3	52.6	2139598	
5	Polygon	Colorado	08	Mountain	CO	33371648	36529730	214.2	231	2017005	
6	Polygon	Connecticut	09	New England	CT	4301261	4797963	41.3	48.1	3560005	
7	Polygon	Delaware	10	South Atlantic	DE	3405585	3510968	684.6	69.6	2780355	
8	Polygon	District of Columbia	11	South Atlantic	DC	759000	847811	390.5	42.2	584773	
9	Polygon	Florida	12	South Atlantic	FL	15184500	15985000	8075.7	8271	179101	
10	Polygon	Georgia	13	South Atlantic	GA	8011111	8282333	316.6	316.6	2456502	
11	Polygon	Hawaii	15	Pacific	HI	38880	13922	185.3	185.3	1553781	
12	Polygon	Idaho	16	Mountain	ID	1293953	1420234	15.6			
13	Polygon	Illinois	17	East North Central	IL	12419293	12964993	220.4			
14	Polygon	Indiana	18	East North Central	IN	6080485	6277388	168	173.3	1266922	
15	Polygon	Iowa	19	West North Central	IA	2926324	2967823	32	52.7	2748640	
16	Polygon	Kansas	20	West North Central	KS	2688418	2778410	37	33.8	2313944	
17	Polygon	Kentucky	21	East South Central	KY	4041769	4191769	11		944889	
18	Polygon	Louisiana	22	West South Central	LA	4488978	4571111	117.3		358161	
19	Polygon	Maine	23	New England	ME	1274923	1311111	39		238014	
20	Polygon	Maryland	24	South Atlantic	MD	5296486	5634034	782.3	797.7	391308	
21	Polygon	Massachusetts	25	New England	MA	6349927	6474034	80.3	84	474003	
22	Polygon	Michigan	26	East North Central	MI	9938444	10310273	171.1	177.5	7988053	
23	Polygon	Minnesota	27	West North Central	MN	4919479	5257496	58.3	62.3	4400282	
24	Polygon	Mississippi	28	East South Central	MS	2844858	2917751	59.7	61.2	1746099	
25	Polygon	Missouri	29	West North Central	MO	5552111	5852142	80.3	84	474003	
26	Polygon	Montana	30	Mountain	MT	902195	928880	6.1	6.3	917229	
27	Polygon	Nebraska	31	West North Central	NE	1711263	1768255	22.1	22.9	1833261	
28	Polygon	Nevada	32	Mountain	NV	1996257	2448538	18.1	22.1	1501888	
29	Polygon	New Hampshire	33	New England	NH	1237868	1317967	133.4	142.2	1188851	
30	Polygon	New Jersey	34	Middle Atlantic	NJ	8414350	8768091	1114.8	1161.7	8104705	
31	Polygon	New Mexico	35	Mountain	NM	1819048	1916138	15	15.8	1214253	
32	Polygon	New York	36	Middle Atlantic	NY	18976457	19411913	390.3	399.2	1299368	
33	Polygon	North Carolina	37	South Atlantic	NC	8049313	8732955	163	176.9	8004956	
34	Polygon	North Dakota	38	West North Central	ND	643290	657999	9.1	9	993161	
35	Polygon	Ohio	39	East North Central	OH	11353140	11573459	275.2	280.5	9845453	
36	Polygon	Oklahoma	40	West South Central	OK	3450854	3553828	49.4	50.8	2628434	
37	Polygon	Oregon	41	Pacific	OR	3421399	3857289	35.2	37.7	2861623	
38	Polygon	Pennsylvania	42	Middle Atlantic	PA	12281054	12480851	271.1	275.5	10484220	
39	Polygon	Puerto Rico	72	South Atlantic	PR	-99	-99	-99	-99	-99	
40	Polygon	Rhode Island	44	New England	RI	1048319	1074684	684.1	688.3	891191	
41	Polygon	South Carolina	45	South Atlantic	SC	4012012	4258937	129.8	137.6	2695880	
42	Polygon	South Dakota	46	West North Central	SD	754944	793954	9.8	10.2	689404	

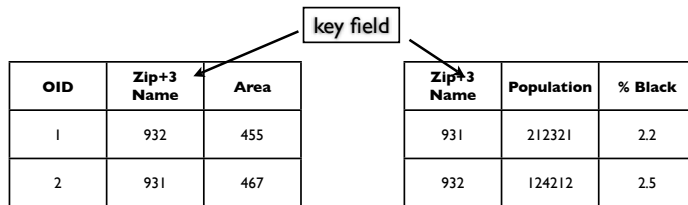
Joining Tables

- situation
 - many times you have a standalone table that has *implicit spatial data*
 - you have feature layer that has the *same spatial data*, but in *explicit form*
 - you want to link them together

Joining Tables

- example
 - census data Zip+3 polygon shapefile
 - attribute table just has Zip+3 name, area, etc.
 - a census data table with lots of demographic info on each Zip+3 polygon
 - must be downloaded separately
 - need to link to the Zip+3 shapefile for spatial display, analysis, operations, etc.
- JOINING IS YOUR ANSWER!

Joining Tables

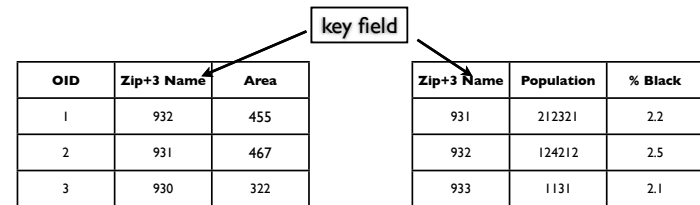


Attribute Table
Destination Table

Standalone Table
Source Table

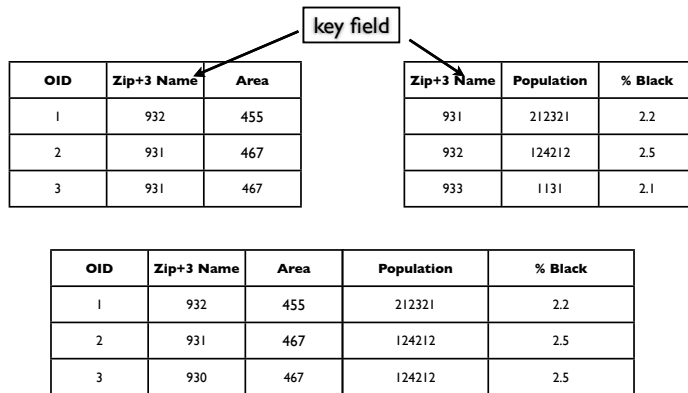
OID	Zip+3 Name	Area	Population	% Black
1	932	455	212321	2.2
2	931	467	124212	2.5

Joining Tables



OID	Zip+3 Name	Area	Population	% Black
1	932	455	212321	2.2
2	931	467	124212	2.5
3	930	322	<null>	<null>

Joining Tables



Joining Tables

- cardinality
 - joining is good for **one-to-one** cardinality and **one-to-many** cardinality
 - why not **many-to-one** or **many-to-many** cardinality?
 - we use relates for these cardinalities

Querying Tables

- Why query?
 - to **select** features (or raster cells) meeting a certain criteria
 - you can then perform operations on the selected features
 - make them into a new layer
 - count them, averages, other stats
 - pattern analysis
 - etc.