Chapter 4. Attribute Data

Objectives:

- Understanding how **tabular data** are stored and used
- Understanding the links between **database management systems** and **tables**
- (Using queries to select records of interest)
- Understanding joins and cardinality concepts
- **Summarizing** tables to get **statistics** on groups
- Learning how to **define fields**
- Editing and calculating fields in tables

A GIS can answer the question: What is where?

- WHAT: Characteristics of attributes or features.
- WHERE: In geographic space.

A GIS links attribute and spatial data

- Attribute Data
 - Flat File
 - Relations

- Map Data
 - Point File
 - Line File
 - Area File
 - Topology
 - Theme

The Two Types of Data in GIS

Spatial data: Describing where things are

AND

Attribute data: Describing what things are

- •Example: A point specified by UTM coordinates
 - •Easting = 50,000 m
 - •Northing = 5,000,000 m
 - •Zone =17
- •This specifies the **location** of a point of the ground
- •The nature of the real-world feature located at this point would be recorded in the attribute data

•Traditionally, geographic data and attributes were recorded on paper too (maps), and these had the same problems as a phone book

Table terminology

Т	itle							Each fie	ld is sp	pecificall	y
								defined before a	and es	stablishe ta can bi	d
									entere	d.	
	III Attribu	utes of Cour	nties							_	미뇌
		NAME	FIPS	AREA	POP1990	POP2000	POP90_SQMI	HOUSEHOLDS	MALES	FEMALES	M 🔺
- 11	Lake of	f the Woods	27077	1784.0634	4076.000000	4651	2	1576	2037	2039	
- 11	Ferry		53019	2280.2319	6295.000000	7199	3	2247	3280	3015	
- 11	Steven	s	53065	2529.9794	30948.000000	40652	12	11241	15454	15494	
- 11	Okano		53047	5306.18	33350.000000	38640	6	12654	16828	16522	
- 11	Pend O	rei	53051	1445.0286	8915.000000	11752	6	3395	4426	4489	
- 11	Bounda			79.2987	8332.000000	10068	7	2857	4252	4080	
- 11	Lincoln	Kecc	ords	46.0908	17481.000000	18859	5	6668	8777	8704	
- 11	Flathea			32.0306	59218.000000	73438	11	22834	29316	29902	
- 11	Glacier		30035	3124.4572	12121.000000	12626	4	3816	5985	6136	
- 11	Toole		30101	1943.2598	5046.000000	4556	3	1922	2486	2560	
- 11	Liberty		30051	1485.9458	2295.000000	2222	2	788	1120	1175	-
	•										
	Record: III I III S Field definitions control *2000 Sele						² *2000 Selecte	d)	Options 👻]	
-	the type of data that can be stored in a field.										

Types of tables

- Attribute table
 - Stores attributes of map features
 - Associated with a spatial data layer
 - Has special fields for spatial information

- Standalone table
 - Stores any tabular data
 - Not associated with spatial data
 - OID instead of FID

III Attributes of US Counties									
	FID	Shape*	NAME	STATE_NAME					
•	0	Polygon	Lake of the Woods	Minnesota					
	1	Polygon	Ferry	Washington					
-	2	Polygon	Stevens	Washington					
	3	Polygon	Okanogan	Washington					
2	4	Polygon	Pend Oreille	Washington					

▦	III Attributes of popestmt									
	OID	FIPS	POP1998	POP1997	POP1996					
E	0	01001	42095	41284	40251					
	1	01003	132828	128820	124257					
	2	01005	26895	26791	26870					
	3	01007	18926	18595	18227					
	4	01009	46266	44930	43548					

Database Management Systems

- Dedicated systems for managing tables of data
- Provide **data management** for agencies, universities, companies, etc.
- Designed for multi-user environments with enhanced security needs
- Focus on data tables with tools for queries, reporting, graphing, etc.

Flat file DBMS

• Flat file

- Stores data as rows of information in files
- **Simple** and robust
- **Inefficient** for search and query



Hierarchical DBMS

- Stores data in multiple tables
- Tables have **defined parent-child relationships**
- **Pre-set hierarchy** of table relationships **designed for specific queries**
- Very efficient for specific queries
- Range of queries **limited** by structure



Relational DBMS

- Stores data in multiple tables
- Table relationships are **defined as needed**
- Very flexible
- Ideal for **open-ended applications** when queries not known beforehand
- Most common type used in GIS applications





How many customers exceed 100 KwH/month? How many service calls has Customer X had?

Which serviceman performed the most calls in December?

How many different customers has each serviceman seen?

Has Serviceman Smith ever visited Customer Jones?

Relation Rules (Codd, 1970)

- Only one value in **each cell** (intersection of row and column)
- All values in a column are about the **same subject**
- Each row is **unique**
- No significance in **column** sequence
- No significance in **row** sequence

Normalization

- This is the process of converting tables to **conform** to Codd's relational rules
- **Split tables** into new tables that can be **joined** at query time
 - The relational join
- Several levels of normalization
 - Forms: 1NF, 2NF, 3NF, etc.
- Normalization creates many **expensive** joins
- **De-normalization** is OK for **performance optimization**

Relational Join

- We use the **relational join** operation because
 - We are using tables that have been transformed by normalization
 - Data created/maintained by different users, but integration needed for queries
 - We want to **combine data** to ask questions that can only be answered by using the data together
- Table joins use **common keys** (column values) filled with the same identifiers
- The table (attribute) join concept has been extended to **geographic** cases

Joining tables

III Att, butes of StateDemog

POP1990

⁹⁹7928

> >

Join tables on

common field

Show: All Selected

STATE FIPS

I I

Destination table

Source table

POP1999 POP90_SQMI HOUSEHOLD

_ _ X

Records (0 out of 51 Select

∣ ≖ ⊁∣

III Attributes of US States								
	FID	Shape*	AREA	STATE_NAME	STATE_FIPS			
	0	Polygon	67290.061	Washington	53			
	1	Polygon	147244.653	Montana	30			
	2	Polygon	32161.925	Maine	23			
	3	Polygon	70812.056	North Dakota	38			
	4	Polygon	77195.055	South Dakota	46			
	5	Polygon	97803.199	Wyoming	56			
	6	Polygon	56088.178	Wisconsin	55 🔪 🔽			
Record: II I O DI Show: All Selected Records (I								

_										
▦	Attributes of US States									
	FID	Shape*	AREA	STATE_NAME	STATE_FIPS	P0P1990	P0P1999	POP90_SQMI		
	0	Polygon	67290.061	Washington	53	4866692	5773907	72		
	1	Polygon	147244.653	Montana	30	799065	884214	5		
	2	Polygon	32161.925	Maine	23	1227928	1248908	38		
	3	Polygon	70812.056	North Dakota	38	638800	637016	9		
	4	Polygon	77195.055	South Dakota	46	696004	739508	9		
	5	Polygon	97803.199	Wyoming	56	453588	492025	5		
	6	Polygon	56088.178	Wisconsin	55	4891769	5251095	87		
	Į 7	Polyaph	83343 643	Idaho	16	1006749	1250247		Υ	
Loipod t										
Record: I I I I I Show: All Selected Records (0 out of 51 Selected.) JOINEO L										
						• • • • • • •	,			

Join facts

- Joins are **temporary relationships** between tables used by a relational DBMS
- Tables **must share a common field** (key)
- Treats the **two tables as a single table**
- Original stored data is **not affected**
- Can be removed when **no longer needed**

Relational databases

Restaurant table

Restaurant	Res- ID	Parcel_no
Jake's Pizza	20	45-98764
Momma's Pie Hut	30	64-56790
Big Burger Barn	40	62-98754

Employee table

	Res-ID	Name	SSN	
	20	Jake Smith	134-56-7689	
*	20 Nancy Gold		229-69-3490	
X	20	Dan Smurt	345-34-8968	
	30	Karen White	776-67-4578	
	40	Judy Lewis	670-45-6890	
	40	Joshua Jones	675-56-4982	

Parcels table

Parcel_no	Address	Value	Owner
45-98764	1104 Maple Ave	67,000	Roger Clark
64-56790	1900 Main St	114,510	Roger Clark
62-98754	9207 Sherry Ave	59,000	Judy Lewis

Store distinct tables Establish relationships between them

One-to-one joins

Destination table

(always imagine on the left)

Attributes of US States								
	FID	Shape*	STATE_NAME	STATE_ABBR				
Þ	0	Polygon	Hawaii	HI				
	1	Polygon	Washington	WA				
	2	Polygon	Montana	МТ				
	3	Polygon	Maine	МЕ				
	4	Polygon	North Dakota	ND				
	5	Polygon	South Dakota	SD				
	6	Polygon	Wyoming	WY				
	7	Polvaon	Wisconsin	WI				

Source table

(always imagine on the right)

quakesum								
STATE	Count_	Sum_DAMAGE	Sι					
CA	218	3705234000						
AK	106	32600000						
MT	62	4220000						
WA	67	3775000						
ID	41	1350000						
HI	63	1100000						
OR	24	760000						

When each record in the destination table matches exactly one record in the source table, we call it a cardinality of one-to-one.

Types of Cardinality

• One-to-one

- States to Governors
- Husbands to wives

• One-to-many

- States to cities
- Districts to schools

(Destination on the left)

- Many-to-one
 - Cities to states
 - Schools to districts
- Many-to-many
 - Students to classes
 - Stores to customers

In evaluating cardinality, always put the destination first.

Rule of Joining

Each record in the destination table must match one and only one record in the source table.

III Attributes of US States								
	FID Shape* STATE_NAME STATE_ABBR							
E	0	Polygon	Hawaii	HI				
	1	Polygon	Washington	WA 🔨				
	2	Polygon	Montana	MT				
	3	Polygon	Maine	ME				
	4	Polygon	North Dakota	ND				
	5	Polygon	South Dakota	SD				
	6	Polygon	Wyoming	WY				
	7	Polvaon	Wisconsin	WI				

quakesu			
STATE	Count_	Sum_DAMAGE	Sum_DE
CA	218	3705234000	
AK	106	32600000	
MT	62	4220000	
WA	67	3775000	
ID	41	1350000	
HI	63	1100000	
OR	24	760000	

Onetoone

Destination table

Source table

				l r	_		
US Counties					田	Attribu	t
Shape*	NAME	STATE_NAM	IS▲			FID	ľ
Polygon	Lake of the Woods	Minnesota	27		H		ľ
Polygon	Ferry	Washington	5 5			1	ŀ
Polygon	Stevens	Washington	55		Н	2	ŀ
Polygon	Okanogan	Washington	53		Н	3	ŀ
Polygon	Pend Oreille	Washington 4	53		Н	4	ŀ
Polygon	Boundary	Idaho	16		Н	5	┝
Polygon	Lincoln	Montana	30		Н	6	┝
Polygon	Flathead	Montana	30 🖵		Н	7	ŀ

	Attribu	tes of US !	States	
	FID	Shape*	STATE_NAME	STATE_ABBR
۲	0	Polygon	Hawaii	HI
	1	Polygon	Washington	WA
	2	Polygon	Montana	MT
	3	Polygon	Maine	ME
	4	Polygon	North Dakota	ND
	5	Polygon	South Dakota	SD
	6	Polygon	Wyoming	WY
	7	Polvaon	Wisconsin	WI



One-to-many

#	Attribu	tes of US :	States	
	FID	Shape*	STATE_NAME	STATE_ABBR
E	0	Polygon	Hawaii	HI
	1	Polygon	Washington	WA
	2	Polygon	Montana	МТ
	3	Polygon	Maine	ME
	4	Polygon	North Dakota	ND
	5	Polygon	South Dakota	SD
	6	Polygon	Wyoming	WY
	7	Polvaon	Wisconsin	WI

Destination table

Source table

Violates the Rule of Joining

Record to join to destination is **ambiguous**

Must use a relate instead

Relates

- Similar to a join except that
 - The tables **remain separate**
 - Items selected in one table may be highlighted in the related table

Related tables

States: Select the New England States

Attributes of states												
OBJECTID Shape *	AREA	STATE_NAME	STATE_FIPS	SUB_REGIO	NSTATE_A	88		- I	• • ••••			
8 Polygon	56088.178	Wisconsin	55	E N Cen	WI		FIND &	. кері	ace			
9 Polygon	83343.643	Idaho	16	Mtn	ID							
10 Polygon	9603.272	Vermont	50	N Eng	VT		Select	BV A	ttribute	95		
11 Polygon	84520.49	Minnesota	27	W N Cen	MN	SQL						
12 Polygon	97073.594	Oregon	41	Pacific	OR		C - I	жШ				
13 Polygon	9259.527	New Hampshire	33	N Eng	NH		Select	HI				
14 Polygon	56257.965	lowa	19	W N Cen	IA		_1					
15 Polygon	8172.561	Massachusetts	25	N Eng	MA		- Clear S	pelect	ION			
16 Polygon	//330.258	Nebraska	31	W N Cen	INE		_	_				
						2	Switch	i Sele	ction			
Record: 14 4	1 > >	Show: All S	elected Recor	rds (6 out of 51	Selected)							
'												
						1	алл н	eld				
naross F	2005 (<mark>, Engla</mark>	nd St	atos		Add H Relate	eld ed Tat	oles		•	
ongress F	Reps (of New	<mark>/ Engla</mark>	nd St	ates		Relate	eld :d Tat	oles		•1	
ongress F	Reps of	<mark>of New</mark>	<mark>/ Engla</mark>	<mark>nd St</mark>	ates		Relate	eld d Tat Sta	oles I tes-T i	o-Disti	► I	: cd
DINGRESS F	Reps (of New	<mark>/ Engla</mark>	nd St	ates	ISTAT	Relate	eld d Tat Sta	oles i tes-T i	o-Disti	► I	: cd
DINGICOUS F Attributes of cd108 OBJECTID Shape 413 Polygon	Reps (of New NAME	/ Engla	nd St	ates	STAT	Relate	eld ed Tat Sta	oles I tes-T i	o-Disti	→ [ricts	: cd
DINGICESS F Attributes of cd108 OBJECTID Shape 413 Polygon 414 Polygon	Carolyn McC	of New NAME Carthy	/ Engla	nd St	ates	STAT		eld ed Tat Sta	oles I tes-T i	o-Disti	► I ricts	: cd
ONGRESS F Attributes of cd108 OBJECTID Shape 413 Polygon 414 Polygon 415 Polygon	Carolyn McC Steve J. Isra Peter T. King	of New NAME Carthy sel	/ Engla	nd St DISTRICTI 3604 3602 3603	ates D STFIPS 36 36 36 36	STAT NY NY NY		eld d Tat Sta	oles I tes-T i	o-Disti	► I	: cd
ONGRESS F Attributes of cd108 OBJECTID * Shape 413 Polygon 414 Polygon 415 Polygon 416 Polygon	Carolyn McC Steve J. Isra Peter T. King James McGo	of New NAME Carthy sel g overn	/ Engla PARTY Democrat Democrat Republican Democrat	nd St DISTRICTI 3604 3602 3603 2503	ates D <u>STFIPS</u> 36 36 36 36 25	STAT NY NY NY MA		eid d Tab	oles I tes-T	o-Disti	→ I	: cd
DINGRESS F Attributes of cd108 OBJECTID • Shape 413 Polygon 414 Polygon 415 Polygon 416 Polygon 417 Polygon	Carolyn McC Steve J. Isra Peter T. King James McGi Rob Simmor	ovem	/ Engla PARTY Democrat Democrat Republican Republican	nd St DISTRICTI 3604 3602 3603 2503 0902	ates D <u>STFIPS</u> 36 36 36 25 09	STAT NY NY MA CT		eid d Tab Sta	oles i tes-T i	o-Disti	▶ I ricts	: cd
Attributes of cd108 OBJECTID	Carolyn McC Steve J. Isra Peter T. King James McGo Rob Simmor Patrick J. Ke	Df New NAME Carthy sel g overn ns ennedy	/ Engla PARTY Democrat Democrat Republican Democrat Republican Democrat	nd St DISTRICTI 3604 3602 3603 2503 0902 4401	ates D STFIPS 36 36 36 25 09 44	STAT NY NY MA CT RI		eid d Tab Sta	oles i tes-T i	o-Disti	▶ I ricts	: cd
Attributes of cd108 OBJECTID • Shape 413 Polygon 414 Polygon 415 Polygon 415 Polygon 416 Polygon 417 Polygon 418 Polygon 419 Polygon	Carolyn McC Carolyn McC Steve J. Isra Peter T. King James McG Rob Simmor Patrick J. Ka James R. La	Df New NAME Carthy ael g overn 1s annedy angevin	/ Engla PARTY Democrat Democrat Republican Democrat Republican Democrat Democrat	nd St DISTRICTI 3604 3602 3603 2503 0902 4401 4402	ates D STFIPS 36 36 36 25 09 44 44	STAT NY NY MA CT RI RI	Add Hi	eld d Tab	oles I tes-T	o-Disti	▶ I	: cd
Attributes of cd108 OBJECTID • Shape 413 Polygon 414 Polygon 415 Polygon 415 Polygon 416 Polygon 417 Polygon 418 Polygon 419 Polygon 420 Polygon	Carolyn McC Carolyn McC Steve J. Isra Peter T. King James McG Rob Simmor Patrick J. Ka James R. La Rosa L. Del	Df New NAME Carthy sel g ovem 1s annedy angevin Lauro	/ Engla PARTY Democrat Democrat Republican Democrat Republican Democrat Democrat Democrat	nd St DISTRICTI 3604 3602 3603 2503 0902 4401 4402 0903	ates D STFIPS 36 36 36 25 09 44 44 09	STAT NY NY MA CT RI RI CT		eld d Tab	oles I tes-T	o-Disti	▶ I	: cd
Attributes of cd108 OBJECTID • Shape 413 Polygon 414 Polygon 415 Polygon 415 Polygon 416 Polygon 417 Polygon 418 Polygon 419 Polygon 420 Polygon	Carolyn McC Carolyn McC Steve J. Isra Peter T. King James McGo Rob Simmor Patrick J. Ke James R. La Rosa L. Del	DI NAME NAME Carthy ael g overn 1s ennedy angevin Lauro	/ Engla PARTY Democrat Democrat Republican Democrat Republican Democrat Democrat Democrat	nd St DISTRICTI 3604 3602 3603 2503 0902 4401 4402 0903	ates D STFIPS 36 36 36 25 09 44 44 09	STAT NY NY MA CT RI RI CT		eld d Tab	oles I tes-T i	o-Disti	▶ I ricts	: cd
Attributes of cd108 OBJECTID • Shape 413 Polygon 414 Polygon 415 Polygon 416 Polygon 416 Polygon 417 Polygon 418 Polygon 419 Polygon 420 Polygon	Carolyn McC Steve J. Isra Peter T. King James McGi Rob Simmor Patrick J. Ke James R. La Rosa L. Del	DI NAME NAME Carthy ael g ovem hs ennedy angevin Lauro	/ Engla PARTY Democrat Democrat Republican Democrat Republican Democrat Democrat Democrat	nd St DISTRICTI 3604 3602 3603 2503 0902 4401 4402 0903 cd (22 cut of 1)	ates 36 36 36 36 36 44 44 44 44 44	STAT NY NY MA CT RI RI CT		eld d Tab	oles I tes-T	o-Disti	▶ I	: cd

Summarizing tables

- Calculate statistics for **groups of features** in a table
- Groups by unique values in the one field
- User chooses statistics to calculate for other fields
- **Produces another table as output** with groups and stats

						_0	×
•	AREA	STATE_NAME	STATE_FIPS	SUB_REGION	STATE_ABBR	POP1990	
	56088.178	Wisconsin	55	E N Cen	WI	4891769	
	83343.643	Idaho	16	Mtn	ID	1006749	
	9603.272	Vermont	50	N Eng	VT	562758	
	84520.49	Minnesota	27	W N Cen	MN	4375099	
	97073.594	Oregon	41	Pacific	OR	2842321	
	9259.527	New Hampshire	33	N Eng	NH	1109252	
	56257.965	lowa	19	W N Cen	IA	2776755	
	8172.561	Massachusetts	25	N Eng	MA	6016425	
	77330.258	Nebraska	31	W N Cen	NE	1578385	

How many people live in each subregion? What is the total area of each subregion?

Summarizing tables

	Attribu	tes of qu	ıakehis					X
Γ	STATE	DEPTH	DEATHS	DAMAGE	MAG	MMI	LOCATION	~
	CA	20	3000	52400000	7.8	11	Near San Francisco, California	-
	CA	0	300	0	0	9	Southern California	
	AK	33	125	31100000	9.23	10	Prince William Sound, Alaska	
	PR	0	116	400000	7.5	9	Northwestern Mona Passage	
	CA	16	115	4000000	6.3	8	Southeast of Long Beach, near Newport Beach, California	
	Н	0	77	0	7.9	10	Near south coast of Hawaii	

Historic major earthquakes



How many earthquakes in each state?

Total deaths and damage in each state?

Average magnitude in each state?

How to summarize

of qua	kehi	is					
STA	TE	DEPTH	DEATHS	DAMAGE	MAG	MMI	LOCATION
мо		0	7	0	7.88	12	New Madrid, Missouri
SN		0	51	0	7.36	12	Northern Sonora, Mexico
AK		0	0	0	8.15	11	Yakutat Bay, Alaska
AK			0	0	8.26	11	Southeast Alaska
AR			7	0	7.68	11	Northeast Arkansas
CA		20	3000	52400000	7.80	11	Near San Francisco, California
CA		16	12	6000000	7.48	11	South of Bakersfield, California
CA		8	65	50500000	6.62	11	North of San Fernando, California
		7	<u>۱</u>				

Right-click State field

Sum Deaths

Sum Damage

Average Mag

Average MMI

Summarize

Summarize creates a new table containing one record for each unique value of the selected field, along with statistics summarizing any of the other fields.

1. Select a <u>fi</u> eld to summarize:
STATE
Choose one or more summary statistics to be included in the output table:
FID ID ID
MAG Minimum Maximum
✓ Average
Variance
3. Specify output table:
C:\MGIS\mgisdata\temp\quakesum.dbf
Summarize on the selected records only
More about Summarize OK Cancel

X

Summarize Output Table

1	Attributes of	quakesu	m					×	
ſ	OID	STATE	Count_	Sum_DAMAGE	Sum_DEATHS	Average_MAG	Average_MMI		
E	6	CA	218	3705234000	3777	5.2575	7.422		
I	0	AK	106	32600000	125	6.5042	3.7358		
L	26	MT	62	4220000	32	2.9737	5.9677		
E	54	WA	67	3775000	15	3.5894	5.8955		
E	13	ID	41	1350000	2	3.721	5.6585		
L	12	HI	63	1100000	79	3.9322	6.4603		
L	40	OR	24	760000	2	3.9646	5.9583		
L	45	SC	24	600000	60	2.0258	6.0833		
L	43	PR	43	400000	116	1.5	5.6512		
L	36	NY	2	200000	0	3.4565	6.3043		
L	17	KY	12	100000	0	4.0525	5.8333	F	
ľ	Record: III 0 DI S Selected Records (0 out of 57 Selected.)								
Count field always generated									
					automatica	allv			

Create map

Could we now create a map of deaths by state?
No, there are no features (yet).

▦	Attributes of	quakesu	m		Sta	Standalone table				
	OID	STATE	Count_	Sum_DAMAGE	Sum_DEATHS	Average_MAG	Average_MMI			
	6	CA	218	3705234000	3777	5.2575	7.422			
	0	AK	106	32600000	125	6.5042	3.7358			
	26	MT	62	4220000	32	2.9737	5.9677			
	54	WA	67	3775000	15	3.5894	5.8955			
	13	ID	41	1350000	2	3.721	5.6585			
	12	HI	63	1100000	79	3.9322	6.4603			
	40	OR	24	760000	2	3.9646	5.9583			
	45	SC	24	600000	60	2.0258	6.0833			
	43	PR	43	400000	116	1.5	5.6512			
	36	NY	23	200000	0	3.4565	6.3043			
	17	KY	12	100000	0	4.0525	5.8333	T		
Re		-· 0	H	Show: All Selecte	ed Records (0 o	ut of 57 Selected.)	– – Option:	s •		

Joining the table

	utes of	quakesu	m			
Summarize	ID	STATE	Count_	Sum_DAMAGE	Sum_DEATHS	A
output toblo	6	CA	218	3705234000	3777	
output table	0	AK	106	32600000	125	
	26	MT	62	4220000	32	
	54	WA	67	3775000	15	
	13	ID	41	1350000	2	
	12	HI	63	1100000	79	
	40	OR	24	760000	2	



						Sta	tes laver	•		
	▦	Attribu	tes of US !							
		FID	Shape*	STATE_NAME	STATE	aunules				
I		0	Polygon	Hawaii	HI		1108229			
		1	Polygon	Washington	WA		4866692			
I		2	Polygon	Montana	MT		799065			
		3	Polygon	Maine	ME		1227928			
		4	Polygon	North Dakota	ND		638800			
		5	Polygon	South Dakota	SD		696004			
		6	Polygon	Wyoming	WY		453588			
		7	Polvaon	Wisconsin	WI		4891769	-1		

US Earthquake Deaths by State



Join summarize output to states layer to create map of deaths

Fields

- Fields have specific types available
- Must be **defined before use**
- Once defined, cannot be changed
- Naming rules
 - No more than 13 characters
 - Use only letters and numbers
 - Must start with a letter
- How is information **actually stored** in fields ...

Maps as Numbers

- GIS requires that both data and maps be **represented as numbers**.
- The GIS places data into the computer's memory in a **physical data structure** (i.e. files and directories).
- Files can be written in **binary** or as **ASCII text**.
- Binary is **faster to read and smaller**, ASCII can be **read by humans and edited** but uses more space.

Binary Notation

•Everything is represented as 0s and 1s in a computer. These two-state forms correspond to yes/no, on/off, open/closed

Binary		Decimal	One to c	One to one correspondence	
0, 1	1 bit	0,1,2,9	Decimal	Binary	
			0	0	
00, 01	2 bits	00, 01,	1	1	
10, 11		97, 99	2	10	
			3	11	
000, 001	3 bits	000, 001,	4	100	
010, 011		002, 003,	5	101	
100, 101		•••	6	?	
110, 111		998, 999			
				wid Tapanhaum EEOS 281 JIMB Eall 2010	
	Binary 0, 1 00, 01 10, 11 000, 001 010, 011 100, 101 110, 111	Binary 1 bit 0, 1 1 bit 00, 01 2 bits 10, 11 3 bits 000, 001 3 bits 010, 011 110 110, 111 110	Binary $0, 1$ Decimal 1 bit $0, 1$ 1 bit $0, 1, 2, \dots 9$ $00, 01$ $10, 11$ 2 bits $00, 01, \dots$ $97, 99$ $000, 001$ $010, 011$ $100, 101$ $110, 111$ 3 bits $000, 001, \dots$ $002, 003, \dots$ $998, 999$	BinaryDecimalOne to 0 0, 11 bit $0, 1, 2, \dots 9$ Decimal000000, 012 bits00, 01,110, 1197, 99233000, 0013 bits000, 001, 4010, 011002, 003, 5100, 10110, 111998, 999	

Binary Notation

Decimal: $72,479 = 70,000 = 7 \times 10^4$ $2,000 = 2 \times 10^3$ $400 = 4 \times 10^2$ $70 = 7 \times 10^1$ $9 = 9 \times 10^0$



Binary:

Note: In binary 1010 + 110 10000

Bits and Bytes



ASCII Encoding

•If computers store everything using 0s and 1s, then how are **characters** represented?

•The **ASCII** (American Standard Code for Information Interchange) code assigns the numbers 0 through 127 to 128 characters, including upper and lower case alphabets plus various special characters, such as white space etc.

•e.g. decimal 85 is assigned to represent upper case U. In binary, 01010101 = 85. Thus the computer represents U using 01010101.

•Files which contain information encoded in ASCII are **easily transferred** and processed by different computers and programs. These are called "ASCII" or "text" files.
ASCII storage

- American Standard Code for Information Interchange (ASCII)
- Stores letters, characters, and symbols as **single 8-bit binary codes**

CAT =
$$\{67, 65, 84\}$$
 decimal = 01000011010000101010100
cat = $\{99, 97, 116\}$ decimal = 011000110110000101110100
148 = $\{49, 52, 56\}$ decimal = 001100010011010000111000

Storing data

- Text data always stored in ASCII format
- Numeric data may be stored in ASCII or binary format
- Binary is generally more efficient

ASCII stores **three letter codes of 1 byte each** = 3 bytes 106 = {49,48,54} decimal = 001100010011000000110110

Binary stores 106 as a single 1-byte binary number 106 = 01101010

Byte storage limits

- A single byte can store a value from 0 to 2⁸-1
- Larger numbers require more bytes
 - 1-byte $2^8-1 = 255$
 - 2-bytes $2^{16}-1 = 65,535$
 - 3-bytes 2^{24} -1 = 16,777,215
 - 4-bytes 2^{32} -1 = 4,294,967,295
- **Signed numbers** require a **bit to store positive or negative**, so storage limits are **smaller**
 - -2 bytes 2^{15} -1 = -32,767 to +32,767
 - 4 bytes 2^{31} -1 = -2,147,483,647 to +2,147,483,647

In base 2: 00000000 = 011111111 = 255 $2^8 = 256$

Integer vs. float storage

Scientific notation 3.2957239 x 10⁴

- Binary stores whole numbers (integers)
- To store **decimal values**, the computer stores a form of scientific notation with a mantissa and an exponent
 - 3.2957239e04 = 32957.239
 - -3.2957239e04 = -32957.239
 - 3.2957239e-04 = 0.00032957239

Float precision

- Large numbers start to lose precision because the number of significant digits in the mantissa is limited.
 - 3.2957239e12 = 3295723900000
- A double-precision floating point allots more storage to the mantissa value
 - 3.295723956249723e12 = 3295723956249.723

Database storage

- Database fields typically are **defined by**:
 - ASCII vs. binary type storage
 - Bytes of storage allocated
 - Integer vs. floating point
- Definition **limits the values** that can be stored
 - Important to match type to storage requirements
 - Try to minimize storage space while making sure all potential values will fit in the field

Text (ASCII) field with 10 bytes"Mississipp"Binary 2-byte signed integer:-32,767 to +32,767Single-precision floating pointx.xxxxxxeyy

About ArcGIS

Chapter 4. Attribute Data

Copyright © 2009 by Maribeth H. Price

David Tenenbaum - EEOS 281 - UMB Fall 2010

Working with tables in ArcGIS

Copyright © 2009 by Maribeth H. Price

Tables in ArcGIS

- Tables contain attribute data
- Many formats, one interface

	FID	Shape*	NAME	STATE_NAME	STATE_FIPS	CNTY_FIF
	0	Polygon	Lake of the Woods	Minnesota	27	077
ſ	1	Polygon	Ferry	Washington	53	019
ĺ	2	Polygon	Stevens	Washington	53	065
ſ	3	Polygon	Okanogan	Washington	53	047
ſ	4	Polygon	Pend Oreille	Washington	53	051
ſ	5	Polygon	Boundary	Idaho	16	021
ľ	6	Polygon	Lincoln	Montana	30	053
l	7	D-1	The share and	1.1	loo	000

Sources of tables

- Dbase files
- INFO files
- ASCII Text files (tab or comma delimited)
- Records from SQL database systems
- Excel worksheets

ArcMap table interface



Adjusting field width

• **Temporary**, does not affect stored file

I Attributes of popestmt								
OID	FIPS	P0P1996	FCP1007	P9P 1996				
0	01001	42095	41284	40251				
1	01003	132828	128820	124257				
2	01005	26895	26791	26870				
3	01007	18926	18595	18227				
4	01009	46266	44930	43548				

Hover over field break to get double arrow, then drag

Field properties tab

Layer Properties ?						? ×		
General Source Selection Display				Fields	Definition Qu	ery Label:	s Joins & Relates	1
4	<u>P</u> rimary Display F	ield:	RISDATA				•	
u	: <u>h</u> oose which field	s will be visible. Clic	k in the alia:	s column to	edit the alias	for any field	i.	
	Name	Alias	Туре	Length	Precision	Scale	Number Format	
	🗹 Shape_1		Polygon					
	🗹 AREA	AREA	Double	8	0	0	Numeric	
	PERIMETER	PERIMETER	Double	8	0	0	Numeric	
	🗹 STANDS2#	STANDS2#	Long	4	0	0	Numeric	
	STANDS2-ID	STANDS2-ID	Long	4	0	0	Numeric	
	🗹 RISDATA	RISDATA	String	10	0	0		
	🗹 DATA	DATA	String	10	0	0	\bigcirc	
	OWNER	OWNER	String	3	0	0		
	COV_TYPE	COVERLITYPE	String	3	0	0		
	SSTAGE96	SSTAG 96	String	4	0	0		<u> </u>
	Select All		_					
		eld allas						
					01	<	Cancel	pply

Shortcut to field properties

Attributes of Counties			Field Propert	ies		<u>? ×</u>
Italite of Counties NAME FIPS ADEA Lake of the Woods 27077 17 Ferry 53019 22 Stevens 53065 25 Okanogan 53047 Ferry Pend Oreille 53051 14 Boundary 16021 12 Lincoln 30053 37 Flathead 30029 52 Glacier 30035 31 Toole 30101 19 Liberty 30051 14 Record: I< 0	ENDEMON ENDEMON	IOLDS MALES 1576 203 2247 328 11241 1545 12654 1682 3395 442 2857 425 6668 877 22834 2931 3816 598 1922 248 788 112 Selected)	Name: <u>N</u> ame: <u>Al</u> ias: <u>Type:</u> Display — <u>Display</u> — <u>Use Fiel</u> <u>Number Fon</u> <u>Data</u> <u>Allow NULI</u> <u>Default Val</u>	POP2000 POP2000 Double Id off d as Primary Displa mat:	ay Field	Numeric
	Propert <u>i</u> es			ОК	Cancel	Apply

Formatting field display

Number Format Category: None Currency Numeric Percentage Custom Rate Fraction Scientific Angle	Rounding Number of decimal places Number of significant digits 3 3 Alignment Left Right 12 Right 12 Show thousands separators Pad with zeros Show plus sign	×
	OK Cancel	

# Attributes of po D X							
FIPS PORMES BORDES							
	06037	9213533	91-				
	17031	5189689	518				
	48201	3206063	31!				
	04013	2784075	265				
	06073	2780592	272				
	06059	2721701	266				
			00000000				
	36047	2267942	228 👻				
-	36047	2267942	22(-				
•	36047	2267942	22€ ▼				
•	36047	2267942 tes ι ρο	22€ ↓				
	36047 Attribu	2267942 tes t ρο POP1998	22€ ↓ ▶ ■ □ × ₽0₽ ▲				
	36047 Attribu FIPS 06037	2267942 tes τ ρο POP1998 9,210,000	22€ ↓ ▶ POP ▲ 91				
4	36047 Attribu FIPS 06037 17031	2267942 tes t po POP1998 9,210,000 5,190,000	22€ ↓ ► POP ▲ 91 518				
	36047 Attribu FIPS 06037 17031 48201	2267942 tes t po 9,210,000 5,190,000 3,210,000	22€ ↓ ► POP ▲ 91 518 315				
	36047 Attribu FIPS 06037 17031 48201 04013	2267942 tes (po 9,210,000 5,190,000 3,210,000 2,780,000	22€ ↓ ► POP ▲ 91 518 318 268				
	36047 Attribu FIPS 06037 17031 48201 04013 06073	2267942 tes (ρο POP1998 9,210,000 5,190,000 3,210,000 2,780,000 2,780,000	22€ ← ► POP ▲ 91 518 318 268 272				

36047

226 🔻

2,270,000

Table appearance

Table Appearance	? ×
Tables	Options 🌆 Find & Replace
Change the appearance of this table.	Select By Attributes
	Select All
Selection color	Switch Selection
The color that selected records are shown in:	Add Field
Highlight color	Related Tables
When only the selected records in a table are being	Create Graph
corresponding features on the map. The records and	Add Table to Layout
features are shown in this color:	C Reload Cache
Table Font: MS Shell Dig	Evport
Table Font Size and Color: 8 💌 📰 🗸	Appearance

Sorting tables

• Has **no effect** on original data

STATE_NAME	STATE_FIPS	POP1990	P0P1999	POP90	SC	МІ нонсеного	MAIES	1
Washington	53	4866692	5773907		1	Sort Ascending	747	
Montana	30	799065	884214		Ē	Sort Descending	769	
Maine	23	1227928	1248908			Summarize	850	
North Dakota	38	638800	637016			Calculate Values	201	
South Dakota	46	696004	739508			Calculate values	498	
Wyoming	56	453588	482025		Σ:	Statistics	007	
Wisconsin	55	4891769	5251093			Freeze/Unfreeze Colun	nn 835	
Idaho	16	1006749	1250247		—			
Vermont	50	562758	593860			Delete Field	492	
Minnesota	27	4375099	4765612	,	_	52 1647853	2145183	
d l	1	1		1		1 1)	Ē

ArcGIS field data types

Geodatabases and shapefiles							
Short	Integers stored as signed 2-byte binary numbers (value range from -32,000 to +32,000)	255 1201					
Long	Integers stored as signed 4-byte binary numbers (value range from -2 billion to +2 billion)	156000					
Float	Floating point values with 8 significant digits in the mantissa	1.2893851e12					
Double	Double-precision floating point values with 16 significant digits in the mantissa	1.11111111111111 1e13					
Text	Alphanumeric strings	'Maple St'					
Date	Date format	07/12/92					
BLOB	Binary large object; any complex binary data including images, documents, etc.						

Field characteristics

- Length
 - The total characters a text field can store
- Precision
 - The total width of digits a numeric field can store
- Scale
 - The number of decimal places

Length = 10 Maple St. Maple Stre

156 1985.128 -1922.5600

0.001 0.00001

Editing and calculating fields

Editing fields

Open Editor toolbar



Type edits in fields

			_
3	Ave_MAG	Ave_MMI	Risk
5	6.5042	3.7358	High
)	3.6143	6.1429	Low
7	5.0113	6.5625	High
)	2.135	5.9	Lo
]	6.35	7	
1	6 3167	85	

Start editing



Save edits, stop editing



Calculating fields

Add Field	
Name: Perc_Hisp	
Type: Float	•
Field Properties	
Precision 6 Scale 2	
	Add Field Name: Perc_Hisp Type: Float Field Properties Field Precision 6 Scale 2

Add a new field if necessary Consider whether you need decimal places!

AVG_SALE87 Perc_Hisp AVG_SALE87 Perc_Hisp 125203 C Sort 87000 C Sort 62980 C Sum	The second secon			Cal	culate
64681 (62007 (Calo	culate Values	or			? ×
Right-click field to calculate	ete Field Perc_Hisp POP1990 POP2000 POP90_SQN RENTER_O SEPARATEI		Type Number String Date	Functions Abs() Atn() Cos() Exp() Fix() Int() Log() Sin() Sqr()	
Enter expressio	Perc_Hisp = [HISPANIC] /	/ [POP2000] *100)	Advanced	× / & + - = Save Load
Copyright © 2009 by Maribeth H. Price	e			David Tenenbaum	

Chapter 5. Queries

Objectives:

- Understanding **queries** and how they are used
- Selected features based on attributes using SQL and Boolean operators
- Selected features based on their **spatial location** with respect to other features
- Applying **selection options**, including the selectable layers and the selection method

What are queries?

- Extract certain records from a map or table
- Records meet certain criteria
 - Aspatial queries
 - All parcels with value greater than \$100,000.
 - Spatial queries
 - All parcels that lie completely within the flood plain

Selecting features of interest



[COV_TYPE] = "TAA"

Selecting aspen stands from a forest vegetation layer.

Using statistics on areas (m²) Minimum: 12,900 Maximum: 750,500 Sum: 10,529,000



Exploring patterns



Are aspen stands randomly scattered or clustered?

Do they occur in **particular portions** of the forest?

What are the distributions of stand densities?

Isolating for more analysis



Are there any mature stands with large trees and open crowns? Where are they?

[TREE_SZ96] = 'L' AND [DENSITY96] = 'A'

Exploring spatial relationships



What fraction of stands are intersected by roads?

What **types of trees** are **adjacent** to aspen stands?

■ Attributes of nextaspenstats						
	COV	TYPE	Count_	Sum_Shape_1_Area		
	TPP		236	41075786.751212		
			33	18307652.679152		
	TAA		85	10529137.894304		
	TBO		5	1296700.771798		
	GRA		5	500328.341753		
	TWS		5	326514.674681		
	TLP		1	35111.344044		
	NFL		1	23186.765917		

Queries involving surfaces

Over what range of elevations do aspen occur?

Do aspen occur above 1500m elevation?





SQL

- Many databases use a **special query language** called **Structured Query Language**
- Can write queries that **work in multiple DBMS** environments
- Queries can be saved and reused
- Nearly always **case-sensitive**

SQL Query Examples

Select By Attributes							
		Query Wizard					
Layer:	Cities	•					
Method :	Create a new selection						
Fields:	Unique sample values						
"CITY_FIPS "CITY_NAM "STATE_FI "STATE_N. "STATE_C "TYPE" "CAPITAL" "ELEVATIC "POP1990"	$\begin{array}{c c} S'' & \bullet & = & \langle \rangle & \text{Like} \\ AE'' & & \rangle & = & \text{And} \\ S'' & & \rangle & \Rightarrow & = & \text{And} \\ AME'' & & \langle \langle \rangle & \bullet & \circ \\ TY'' & & \langle \langle \rangle & \bullet & \circ \\ & & - & \% & () & \text{Not} \\ \end{array}$	10155 10201 10328 11451 11917 12453 13131 14290 17176					
"HOUSEHO	SQL Info	Complete List					
SELECT * FROM cities WHERE:							
Programs may have an interface to help users build SQL expressions							
Clear	Verify Help Load	Save					
	Apply	Close					

Some Valid Queries

SELECT *FROM cities WHERE "POP1990" >= 500000

SELECT *FROM counties WHERE "BEEFCOW_92" < "BEEFCOW_87"

SELECT *FROM parcels WHERE "LU-CODE" = 42 AND "VALUE" > 50000

SELECT *FROM rentals WHERE "RENT" > 700 AND "RENT" < 1500

In most databases, SQL expressions are case-sensitive "Smith" ≠ "SMITH"

Queries as sets



- Let T = [all students in University]
- Let A = [students from New York]
- Let B = [Geography majors]
- Let C = [English majors]

Queries are used to extract subsets (records) of interest from a set (table).

Multiple criteria may be used (such as Geography majors from New York)

Single criteria



- Let T = [all students in University]
- Let A = [students from New York]
- Let B = [Geography majors]
- Let C = [English majors]

Select students from T where [Home_State] = "NY"

Select students from T where [Major] = "Geography"

Double criteria



- Let T = [all students in University]
- Let A = [students from New York]
- Let B = [Geography majors]
- Let C = [English majors]

Select students from T where [Home_State] = "NY" OR [Home_State] = "NJ"

```
Select students from T where
[Home_State] = "NY" AND [Major] = "Geography"
```

AND vs OR?



- Let T = [all students in University]
- Let A = [students from New York]
- Let B = [Geography majors]
- Let C = [English majors]

Select students from T where [Home_State] = "NY" OR [Home_State] = "NJ"

Select students from T where [Home_State] = "NY" AND [Major] = "Geography"

Each condition is **tested separately**. If **AND** is used, then **BOTH must be true**. If **OR** is used, then **either may be true**.
Boolean expressions

AND and OR are known as **Boolean operators**. Boolean operators are used to **evaluate pairs of conditions**.



AND vs OR?



- Let T = [all students in University]
- Let A = [students from New York]
- Let B = [Geography majors]
- Let C = [English majors]



Select students from T where [Home_State] = "NY" AND [Major] = "Geography"

AND vs OR?



- Let T = [all students in University]
- Let A = [students from New York]
- Let B = [Geography majors]
- Let C = [English majors]



Select students from T where [Home_State] = "NY" OR [Major] = "Geography"

What do you get?



- Let T = [all students in University]
- Let A = [students from New York]
- Let B = [Geography majors]
- Let C = [English majors]

•Select students from T where [Major] = "Geography" AND [Major] = English"

• BAND C

•Select students from T where [Major] = "Geography" OR [Major] = "English"

• B OR C

•Select students from T where [State] = "NY" AND [Major] = "English"

• A AND C

•Select students from T where [State = "NY" OR [Major] = "English"

• A OR C

Copyright © 2009 by Maribeth H. Price

Other Boolean operators

Some databases use additional operators besides AND and OR.



What do you get?



A AND B	B AND A
A OR B	B OR A
A XOR B	B XOR A
B XOR C	B XOR C
A NOT B	B NOT A
B NOT C	C NOT B



- Let T = [all students in University]
- Let A = [students from New York]
- Let B = [Geography majors]
- Let C = [English majors]

David Tenenbaum - EEOS 281 - UMB Fall 2010

Commutation of operators

- AND, OR and XOR are **commutative**
 - A AND B == B AND A
 - A OR B == B OR A
 - A XOR B == B XOR A
- NOT is not commutative
 - − A NOT B $\neq \neq$ B NOT A

Order of operations

- Boolean operators have equal order or precedence
- Evaluation occurs from left to right
- Parentheses must be used to change order

What do you get?



- Let T = [all students in University]
- Let A = [students from New York]
- Let B = [Geography majors]
- Let C = [English majors]

A AND B OR C (A AND B) OR C A AND (B OR C) (A OR B) AND C A OR (B AND C)



David Tenenbaum - EEOS 281 - UMB Fall 2010

Searching for partial matches

- Sometimes you need to find **one string within another** rather than an exact match
 - Find all customer names beginning with "Mac" or "Mc"
 - Find all zip codes beginning with 0
- Typically uses a "wildcard" character
 - *Mac* or *Mc*
 - 0*

The Like Operator

- "NAME" LIKE '%(D)%'
 Finds all of the (D) Democrats
- % is (single character) wildcard
- Ignores Don or Danforth
- "NAME" **LIKE** '%New %'
 - Would find <u>New</u> Hampshire and <u>New</u> York, but not <u>New</u>castle or Ken<u>new</u>ick

NAME				
Sonny Callahan (R)				
Terry Everett (R)				
Bob Riley (R)				
Robert B. Aderholt (R)				
Robert E. "Bud" Cramer, Jr. (D)				
Spencer Bachus (R)				
Earl F. Hilliard (D)				
Don Young (R)				
Matt Salmon (R)				
Ed Pastor (D)				
Bob Stump (R)				
John Shadegg (R)				
Jim Kolbe (R)				

Spatial operators

intersect

are within a distance of contain completely contain contain (Clementini) are within are completely within are completely within are within (Clementini) are identical to touch the boundary of share a line segment with are crossed by the outline of have their centroid in

 Spatial queries can employ a number of operators to test the basic conditions of intersection, containment, and proximity.

Basic spatial relationships



• Intersection

- Does the road **cross** the aspen?
- Do two polygons share areas or boundaries?

• Containment

- Is the aspen **inside** a geology unit?
- Is a road **inside** a geology unit?

• Proximity

 How many aspen stands within 200 meters of a road?

- The operators **test relationships** between **two layers at a time**.
 - The target layer is the one containing the features to be selected
 - The source layer is the one containing the features being compared to.

Select the **aspen** stands that are intersected by **roads**.



Select the **roads** that are intersected by **aspen** stands.



Spatial operators



- The use and action of the operators depends on the feature geometry (points, lines polygons)
- Some operators can only be used with certain geometry types
 - Polygons can contain points, but not vice versa

Intersection operators

	intersect
	are within a distance of
	contain
	completely contain
	contain (Clementini)
	are within
	are completely within
	are within (Clementini)
	are identical to
	touch the boundary of
	share a line segment with
	are crossed by the outline of
Г	have their centroid in

• Features **intersect** when any part of one feature touches, crosses, or overlaps another feature.



The lower set includes "special cases" of intersecting features.

Containment operators

intersect	
are within a distance of	
contain	
completely contain	
contain (Clementini)	
are within	
are completely within	
are within (Clementini)	
are identical to	1
touch the boundary of	
share a line segment with	
are crossed by the outline of	
have their centroid in	

- Features that **enclose** all of another feature **contain** it.
- Within is the inverse of contain





Selecting features that contain points



Selecting features that contain lines

Containment operators are affected by geometry type.

Notice that only polygons can contain other polygons.

However, points can contain other points, lines can contain points or lines, and polygons can contain anything.



Selecting features that contain polygons

Types of containment



Within is the inverse. Columbia county is within Oregon. Jefferson county is completely within Oregon.

Contains

- One feature lies inside another and may share a boundary
- Oregon contains Columbia county

Completely contains

- One feature lies inside another without touching the boundary
- Oregon does not completely contain Columbia county, but does completely contain Jefferson county

Clementini operators

- Eliseo Clementini and his coauthors defined a special set of topological relationships concerning containment*.
- Clementini considers the boundary of a polygon to be separate from its inside or outside.
- The Clementini operator is equivalent to the standard operator except when the source feature lies only on the boundary of the target feature.

*Eliseo Clementini, Paolino Di Felice, and Peter van Oosterom, A Small Set of Formal Topological Relationships Suitable for End-User Interaction. Proceedings of the Third International Symposium on Advances in Spatial Databases, pp. 277-295, June 23-25, 1993.

Clementini example



- The Rio Grande River lies on the border of Texas
 - The Contains operator would select the Rio Grande
 - The Clementini Contains operator would NOT select the Rio Grande because the state boundary is not considered part of Texas

Conversely, the Rio Grande is within Texas using the standard operator, but is not within Texas using the Clementini operator.

Proximity operators

- 6		
		intersect
		are within a distance of
		contain
		completely contain
		contain (Clementini)
		are within
		are completely within
		are within (Clementini)
		are identical to
		touch the boundary of
		share a line segment with
		are crossed by the outline of
	Г	have their centroid in
ľ		

Volcanoes within 100 km of an interstate

• This operator tests whether the target features are **within a specified distance** of the source features.





For more information on each operator and how they apply in the case of points, lines, and polygons, consult the Help information for the Select By Location command in ArcGIS.

Copyright © 2009 by Maribeth H. Price

More examples



Select counties that contain state capitals



Select counties that are within 200 miles of Denver

More examples





Select counties that intersect rivers

Select rivers that intersect Texas

More examples





Select cities that are within 20 miles of an interstate highway Select cities that are within counties named Washington

Scale and accuracy issues

• When testing spatial relationships, consider the **possibility** that **features are not exactly located**.



Consider selecting cities that lie on (intersect) rivers.

A single point or line cannot adequately represent location at this scale. Selection becomes a hit or miss affair.

One can use **buffers** to allow a little **room for error**.

Topology issues



Here, Shannon County should contain the Pine Ridge Reservation, but it does not.

- In the real world, certain boundaries coincide.
- This condition **won't hold true for many feature classes** unless they have been **specifically checked and corrected for logical consistency**.
- Keep in mind that your data sets may contain **topological inconsistencies** that may affect your results.

About ArcGIS

Chapter 5. Queries

Copyright © 2009 by Maribeth H. Price

David Tenenbaum - EEOS 281 - UMB Fall 2010

General information about queries

Queries in ArcMap

- Interactive selection
 - Choose features by **pointing to them on the screen**
- Select By Attribute
 - Select features based on **attribute criteria**
- Select By Location
 - Select features based on their **spatial relationships**

Viewing selected features

	III Attributes of States					<u>- 0 ×</u>		
		FID	Shape*	AREA	STATE_NAME	STATE_FIPS	SUB_REGION	STAT 🔺
		14	Polygon	8172.561	Massachusetts	25	N Eng	MA
		15	Polygon	77330.258	Nebraska	31	W N Cen	NE 🔜
		16	Polygon	48561.751	New York	36	Mid Atl	NY
		17	Polygon	45360.118	Pennsylvania	42	Mid Atl	PA
		18	Polygon	4976.566	Connecticut	09	N Eng	CT
		19	Polygon	1044.881	Rhode Island	44	N Eng	RI 💌
	◀							►
Record: II I II Show: All Selected Records (7 out of 51 Selected.)								

States for which POP2000>2 million



Highlighted in table Highlighted in map

Using Selected features





 Once a layer has a query placed upon it, all subsequent operations on that layer use ONLY the selected features.

Volcanoes selected, then buffered

Buffer uses only selected volcanoes

Statistics only include selected volcanoes

Copyright © 2009 by Maribeth H. Price

Selection states





Clear Selection

On toolbar



From table options menu



Copyright © 2009 by Maribeth H. Price

Interactive selection
Interactive Selection



Select Features tool



Hold down shift key to select more than one feature

Draw a rectangle that passes through features to be selected.

Click on feature to select

Selectable Layers

All layers selectable

States selectable



Select by Attributes Select by Location

Copyright © 2009 by Maribeth H. Price

Select By Attributes

Select By Attributes					
		Query Wizard			
Layer:	Cities	•			
Method :	Create a new selection	•			
Fields:		Unique sample values			
"CITY_FIPS "CITY_NAM "STATE_FI "STATE_N "STATE_C "TYPE" "CAPITAL" "ELEVATIC "POP1990"	$\begin{array}{cccccccccccccccccccccccccccccccccccc$	10155 10201 10328 11451 11917 12453 13131 14290 17176			
"HOUSEHO	SQL Info	Complete List			
SELECT * FROM cities WHERE:					
"'POP1990'' > 500000					
Clear	Verify Help Load.	. Save			
	Apply	Close			

Some Valid Queries

SELECT *FROM cities WHERE "POP1990" >= 500000

SELECT *FROM counties WHERE "BEEFCOW_92" < "BEEFCOW_87"

SELECT *FROM parcels WHERE "LU-CODE" = 42 AND "VALUE" > 50000

SELECT *FROM rentals WHERE "RENT" > 700 AND "RENT" < 1500

Note: Shapefile tables use quotes for field names; geodatabase tables use brackets



Copyright © 2009 by Maribeth H. Price

Intersect

Select By Location	? ×
Lets you select features from one or more layers based on whe in relation to the features in another layer.	ere they are located
I want to:	
select features from	
the following layers:	
Urban Areas States	
Counties	
that:	
intersect	
the features in this layer:	
Rivers	
Use selected features (0 features selected)	
Apply a buffer to the features in Rivers	
of: 0.000000 Meters	

Within distance of

Select By Location	? ×			
Lets you select features from one or more layers based on where they are located in relation to the features in another layer.				
I want to:				
select features from				
the following layers:				
Capitals				
☐ Interstates				
that:				
are within a distance of				
the features in this layer:	A state of the sta			
Capitals				
Use selected features (0 features selected)				
Apply a buffer to the features in Capitals				
of: 50 Miles				

Using a selected set

Select By Location Lets you select features from one or more layers based on where th in relation to the features in another layer. I want to:		
	select features from	S MARSHAR
the following layers:		Carlos A TA
	□Interstates ☑Rivers	
	Urban Areas	
ļ	that:	Salandan h
ľ	intersect	L 24 AH 2000 JK
	the features in this layer:	man shares
	States	M LE TYLL
2	 Use selected features (1 features selected) Apply a buffer to the features in States of: 50.000000 Miles 	

Selection methods

Selection methods



Available for all three types of selection

The Boolean Two-Step



facilitates using multiple steps to apply multiple criteria—like using Boolean operators.



- A OR B Create new selection A; Add B to current selection
- A AND B Create new selection A; Select B from current selection
- A NOT B Create new selection A; Remove B from current selection

Creating layers from queries

Copyright © 2009 by Maribeth H. Price

Creating layers

- **Common operation** following a query
- Creates a new layer with only the selected features



Layers based on selections



[cover_type] = 'Aspen'

Still **based on one original file** shared by both layers

Shows only a **selected subset** in the map and in the table

Use as **input for a tool**, e.g. buffer only the aspen stands



Creating layers

SELECT * FROM counties WHERE:

"STATE_NAME" = 'South Dakota'





Create Layer From Selected Features



Next Topic:

Spatial Joins

David Tenenbaum - EEOS 281 - UMB Fall 2010