GIS Rollup and Drilldown: Geospatial BI and Data Aggregation

Brian Hebert

ScribeKey, LLC

May, 2011

Presentation Overview

- Look at geospatial data compilation and viewing using Business Intelligence (BI) and Data Warehousing approaches.
- Show multi-level data aggregation of point and linear data using hierarchical polygon sets.
- Review details of data compilation and presentation workflow.
- GOAL: Help you get more value out of your data.
- This presentation is part of a series looking at the application of BI and Data Warehousing in a geospatial context.
- Related presentations and supporting materials can be found at www.scribekey.com, including Data Profiling, Meta-Layers, Data Dictionary Generation, Data Quality, etc.

The Value of Data Aggregation in Decision Support

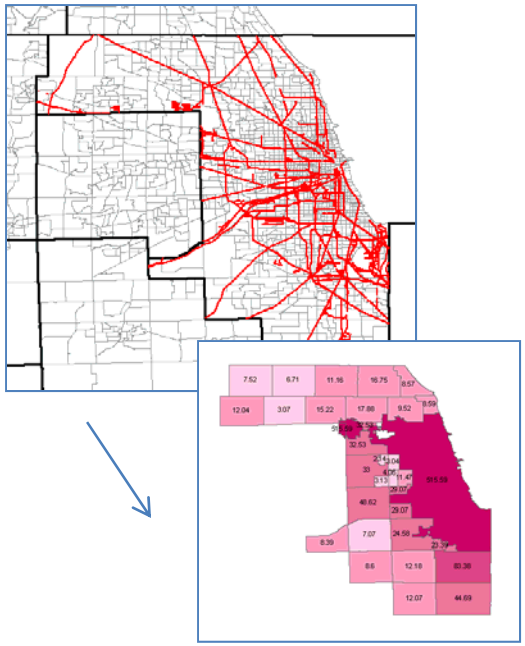
- Data aggregation provides a key component of decision support information systems, AKA, Business Intelligence (BI).
- Provides a smaller, faster, high level summary and simplification of large volumes of data.
- Helps decision makers focus in on what's important.
- Created using standard RDBMS SQL aggregation constructs, SUM, COUNT, and GROUP BY and OLAP technology.

Id	Type	Amt
1	D	64
2	B	24
3	F	95
4	C	98
5	D	54
6	A	26
7	C	29
8	F	84
9	C	56
10	E	59
11	C	31
12	F	77
14	D	52
15	A	8
16	E	27
17	F	82
18	C	38
19	G	57
20	A	13
21	E	68
22	B	87
23	D	43
…	…	…
100,000	F	55

BASE DATA

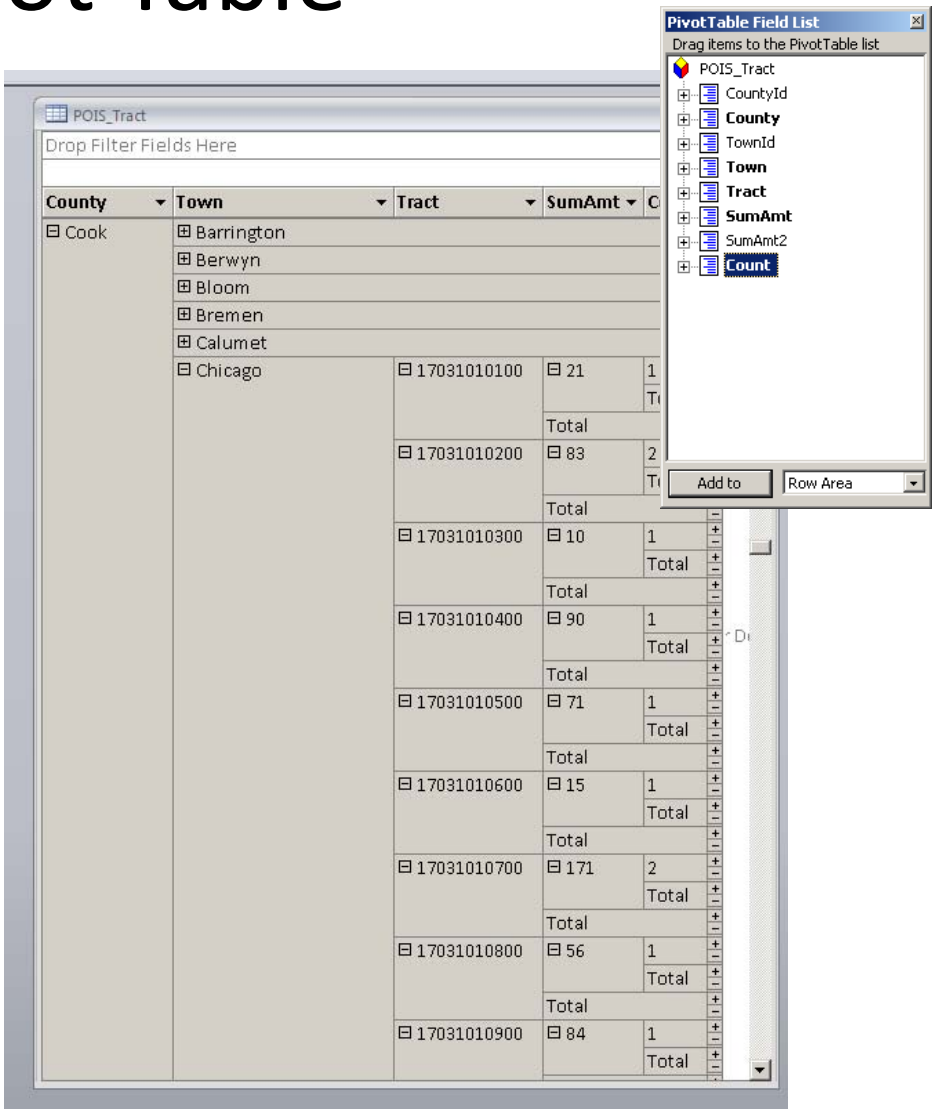
Type	Num	Total
A	15,713	386,623
B	15,631	605,963
C	15,258	613,823
D	15,591	685,496
E	15,739	916,551
F	15,929	1,120,066
G	15,634	1,123,990

AGGREGATE



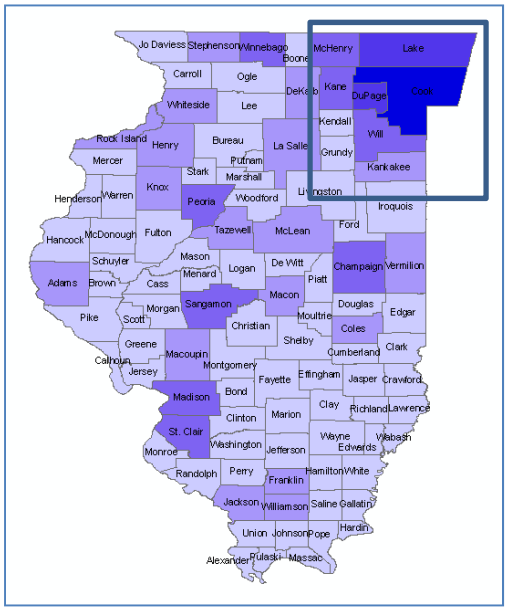
The Pivot Table

- Provides easy-to-use summary and **drilldown/rollup** views of data.
- Very flexible for defining rows, columns, and filters
- Available with MS Excel, Access, and in OLAP environments
- Used extensively in Business Intelligence and Data Warehousing

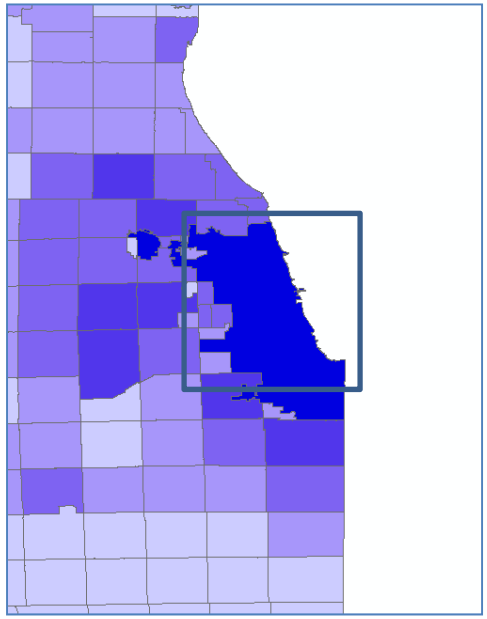


HOW CAN WE MAKE A PIVOT MAP, TO GET THE SAME CAPABILITY WITH GIS?

GOAL: The Pivot Map

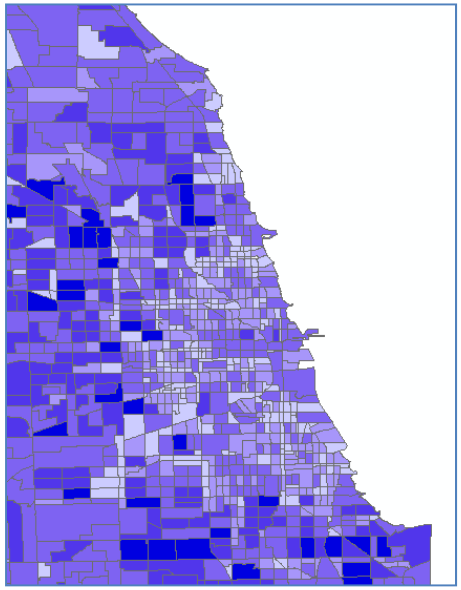


COUNTY



TOWN

Increasingly detailed views



CENSUS TRACT

Applying Pivot Table like view and **Drilldown/Rollup** with hierarchical geography units

www.scribekey.com

Example Case Study: ILLINOIS

- Sample CENSUS data for the State of Illinois.
- Using point and linear data.
- POIs, FOIs, Widgets; think of this as your data.
- Compile linear Railroad feature lengths and display results using shaded polygons.
- Utilities, Infrastructure, Disease Surveillance, Crime, Transportation, Data Holdings, etc.

www.scribekey.com

3 Essential Requirements and Steps

With point data, assume we have **GEOCODED** base info, e.g., sales, crime, disease, critical infrastructure, demographics

1) **SPATIAL JOIN:** Assign points, lines to hierarchical sets of geographic units, as polygons, e.g., country, region, state, marketing areas, legislative disticts, county, city/town, ZipCode, census tract/block-group, block, etc.

2) **AGGREGATION:** Generate rollup views/tables using SQL SUM, COUNT, GROUP BY, as views and tables.

3) **PRESENTATION:** Provide tools for presenting the rollups, e.g., reports, drilldown/pivot tables, drilldown maps, etc.

www.scribekey.com

Designing the Base Table

- For FOIs; the point and linear features of interest, create table with types, numeric quantities, time, and hierarchical geography assignments

Field	Definition	Category
Id	A unique identifier	Id
Type	The feature's type	Type
SubType	The feature's subtype	Type
Amt	A generic quantity	Measure
Amt2	A 2nd generic quantity	Measure
WhenDate	Event date	Time
WhenTime	Event time	Time
Street	Street Name	Address
Street Number	Street Address Number	Address
CountyId	County FIPS	Geography
County	County Name	Geography
TownId	Town FIPS (COSUB)	Geography
Town	Town Name	Geography
Zip	Zip 5	Address
Tract	Census Tract	Geography
BlockGroup	Census Block Group	Geography
X	X (lon)	Location
Y	Y (lat)	Location

IN OLAP TERMINOLOGY, DATA GROUPING CATEGORIES ARE CALLED **DIMENSIONS,** NUMERICAL QUANTITIES ARE CALLED **MEASURES**

www.scribekey.com

Loading the Base Table

- Can use SPATIAL JOIN multiple times, from ArcToolbox, creates new datasets
- For large number of geographies, write custom application with ArcObjects, for each point, cycle through a number of geographies doing a Point-In-Polygon operation
- Economies to be gained with nested, hierarchical IDs, e.g., Census

www.scribekey.com

Base Table Contents

POIs

Id	SubType	Type	Amt	Amt2	Street	StreetNumb	Town	TownId	WhenDate	Zip	WhenTime	x	y	CountyId	County	Tract	BlockGroup
340	6	C	30	369	College Avenue	105	Burton	1700110019	6/14/2006	62347	10:17:00 AM	-91.2090984	39.8864203	17001	Adams	17001010300	170010103003
338	5	D	55	595	Canterbury Court	356	Camp Point	1700110890	8/22/2009	62320	4:42:00 PM	-91.0578296	40.0468339	17001	Adams	17001010200	170010102003
344	2	B	32	241	Sheffield Drive	961	Clayton	1700114780	11/10/2001	62324	7:18:00 AM	-90.9544009	40.0293879	17001	Adams	17001010200	170010102004
346	9	G	90	915	Lilac Lane	711	Columbus	1700115859	1/3/2002	62320	1:04:00 AM	-91.088749	39.955882	17001	Adams	17001010300	170010103002
341	8	F	2	868	Buttonwood Drive	658	Ellington	1700123412	8/1/2008	62305	12:01:00 AM	-91.3049389	39.9827328	17001	Adams	17001010400	170010104002
348	1	B	18	259	Bridle Lane	460	Ellington	1700123412	10/25/2000	62305	6:10:00 PM	-91.354113	39.9540863	17001	Adams	17001000100	170010001001
349	2	B	37	288	Chestnut Street	15	Ellington	1700123412	1/12/2002	62301	8:22:00 AM	-91.3628881	39.9445852	17001	Adams	17001000100	170010001002
351	5	D	48	532	Market Street	763	Ellington	1700123412	2/4/2003	62305	2:07:00 AM	-91.341404	39.941735	17001	Adams	17001000100	170010001005
352	5	D	67	561	Heather Lane	317	Ellington	1700123412	4/21/2005	62305	3:39:00 PM	-91.3251155	39.9622816	17001	Adams	17001010400	170010104001
355	8	F	97	981	Route 29	620	Fall Creek	1700125258	7/19/2008	62360	10:57:00 PM	-91.3359869	39.8089159	17001	Adams	17001010600	170010106004
353	1	B	23	307	Fulton Street	871	Gilmer	1700129288	10/27/2000	62338	8:13:00 PM	-91.1942345	39.9787802	17001	Adams	17001010400	170010104003
381	8	F	79	815	Route 9	600	Honey Creek	1700135957	11/29/2001	62338	10:46:00 PM	-91.240577	40.039376	17001	Adams	17001010100	170010101004
358	2	A	28	354	Old York Road	923	Keene	1700139220	11/28/2001	62349	6:16:00 AM	-91.1996204	40.133538	17001	Adams	17001010100	170010101001
339	1	A	27	341	Linda Lane	909	Liberty	1700143146	11/28/2001	62347	6:15:00 AM	-91.0794361	39.8659839	17001	Adams	17001010300	170010103004
347	5	D	62	661	Route 30	265	McKee	1700145746	3/8/2004	62324	2:36:00 PM	-90.9776437	39.8917092	17001	Adams	17001010300	170010103001
356	4	D	53	579	Franklin Street	815	Melrose	1700148203	2/6/2003	62305	4:31:00 AM	-91.2909702	39.8737832	17001	Adams	17001010600	170010106003
357	1	A	73	178	Lincoln Street	370	Melrose	1700148203	5/22/2000	62305	4:42:00 PM	-91.420627	39.8763449	17001	Adams	17001010600	170010106002
395	7	G	92	928	Cedar Lane	561	Melrose	1700148203	7/17/2007	62305	12:54:00 PM	-91.3411661	39.8959658	17001	Adams	17001010600	170010106001
376	7	F	74	768	6th Street	547	Mendon	1700148320	10/28/2001	62351	8:53:00 PM	-91.2886099	40.0898323	17001	Adams	17001010100	170010101005
378	8	D	49	542	Brook Lane	297	Mendon	1700148320	8/20/2008	62351	3:29:00 PM	-91.3139215	40.1480765	17001	Adams	17001010100	170010101002
396	5	D	12	526	Route 20	756	Northeast	1700153624	9/4/2004	62339	2:06:00 AM	-90.980366	40.1132236	17001	Adams	17001010200	170010102001
337	4	G	99	996	Court Street	161	Payson	1700158278	2/5/2003	62360	3:58:00 AM	-91.2368498	39.8201252	17001	Adams	17001010300	170010103006
350	7	F	92	934	Elmwood Avenue	568	Quincy	1700162380	7/18/2007	62301	9:54:00 PM	-91.3726444	39.9451959	17001	Adams	17001000100	170010001003
380	6	E	24	316	Creek Road	405	Quincy	1700162380	5/13/2005	62301	5:14:00 AM	-91.4124471	39.9400923	17001	Adams	17001000400	170010004003
382	4	C	51	414	Cedar Street	154	Quincy	1700162380	2/16/2003	62301	11:30:00 AM	-91.4137424	39.9183101	17001	Adams	17001000800	170010008002
383	9	D	54	589	Laurel Drive	707	Quincy	1700162380	8/22/2009	62301	4:32:00 PM	-91.3995451	39.9113961	17001	Adams	17001000900	170010009002
384	1	A	10	188	Laurel Street	902	Quincy	1700162380	3/8/2004	62301	6:15:00 AM	-91.3865365	39.9066103	17001	Adams	17001001002	170010010022
385	6	E	81	686	Buckingham Drive	457	Quincy	1700162380	6/25/2006	62301	6:48:00 PM	-91.3592208	39.9049865	17001	Adams	17001001100	170010011003
386	2	F	84	862	High Street	652	Quincy	1700162380	1/1/2002	62305	12:49:00 AM	-91.3488137	39.9304291	17001	Adams	17001001100	170010011001
394	7	F	37	752	Church Street	8	Quincy	1700162380	5/11/2007	62301	8:21:00 AM	-91.3588826	39.920951	17001	Adams	17001001100	170010011002
397	1	E	67	702	Route 6	311	Quincy	1700162380	4/10/2000	62301	7:39:00 AM	-91.368532	39.9375716	17001	Adams	17001000100	170010001004
398	3	B	23	301	2nd Street	29	Quincy	1700162380	10/26/2000	62301	7:51:00 PM	-91.3869104	39.9577092	17001	Adams	17001000201	170010002011
399	7	F	42	476	Washington Street	60	Quincy	1700162380	1/2/2007	62301	9:24:00 AM	-91.3856466	39.9485279	17001	Adams	17001000201	170010002012
400	9	G	97	975	Route 100	777	Quincy	1700162380	7/19/2003	62301	2:57:00 PM	-91.3903968	39.9442374	17001	Adams	17001000202	170010002022
401	5	D	17	574	Meadow Lane	332	Quincy	1700162380	4/6/2005	62301	3:09:00 AM	-91.3823725	39.9441581	17001	Adams	17001000202	170010002021
402	1	F	72	749	Hickory Lane	363	Quincy	1700162380	4/11/2000	62301	8:42:00 AM	-91.3903947	39.9392721	17001	Adams	17001000202	170010002023
403	3	B	28	348	River Road	81	Quincy	1700162380	11/28/2006	62301	9:54:00 PM	-91.3823767	39.9391727	17001	Adams	17001000202	170010002024

Record: 1 of 9803 No Filter Search

WITH BASE TABLE IN PLACE, CAN START TO BUILD AGGREGATE ROLLUPS WITH SQL

Building the Rollups with SQL

- For numeric fields, we use aggregate functions, SUM(), MIN(), MAX(), AVG(), STD().
- Always include COUNT(*).
- Use of GROUP BY to aggregate by Type, Geography, Time, etc.
- There are lots of possible permuations.
- This is what PIVOT TABLE or CUBE does, creates all possible views using the measures and dimensions.

Building the Rollups with SQL (cont.)

```
SELECT
COUNTYID, COUNTY, TYPE,
COUNT(*), SUM(AMT)
[INTO POI_COUNTY]
FROM POI_BASE
GROUP BY
COUNTYID, COUNTY, TYPE
```

Depending on dataset size, use VIEWS or create actual TABLES.

Rollup Hierarchy

POIS_County

CountyId	County	SumAmt	SumAmt2	Count
17001	Adams	2967	33347	60
17003	Alexander	695	7551	13
17005	Bond	601	6900	15
17007	Boone			
17009	Brown			
17011	Bureau			
17013	Calhoun			
17015	Carroll			
17017	Cass			
17019	Champaign			
17021	Christian			
17023	Clark			
17025	Clay			
17027	Clinton			
17029	Coles			
17031	Cook			
17033	Crawford			
17035	Cumberland			
17037	DeKalb			
17039	De Witt			
17041	Douglas			
17043	DuPage			
17045	Edgar			
17047	Edwards			
17049	Effingham			
17051	Fayette			
17053	Ford			
17055	Franklin			
17057	Fulton			
17059	Gallatin			
17061	Greene			
17063	Grundy			
17065	Hamilton			
17067	Hancock			

Record: 1 of 102 No Filt

POIS_Town

CountyId	County	TownId	Town	SumAmt	SumAmt2	Count
17001	Adams	1700110019	Burton	30	369	1
17001	Adams	1700110890	Camp Point	55	595	1
17001	Adams	1700114780	Clayton	32	241	1
17001	Adams					
17001	Adams					
17001	Adams					
17001	Adams					
17001	Adams					
17001	Adams					
17001	Adams					
17001	Adams					
17001	Adams					
17001	Adams					
17001	Adams					
17001	Adams					
17001	Adams					
17001	Adams					
17001	Adams					
17001	Adams					
17001	Adams					
17003	Alexander					
17003	Alexander					
17003	Alexander					
17003	Alexander					
17003	Alexander					
17003	Alexander					
17003	Alexander					
17005	Bond					
17005	Bond					
17005	Bond					
17005	Bond					
17005	Bond					
17005	Bond					

Record: 1 of 1383

POIS_Tract

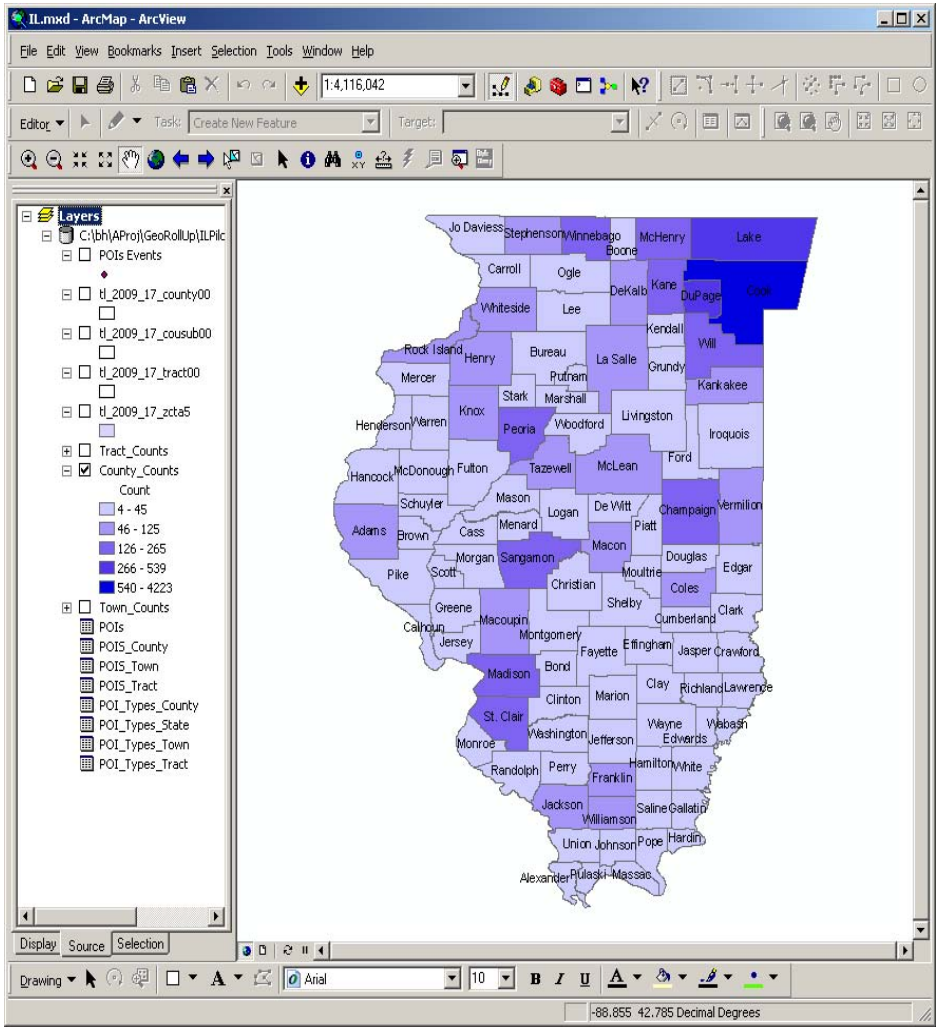
CountyId	County	TownId	Town	Tract	SumAmt	SumAmt2	Count
17001	Adams	1700110019	Burton	17001010300	30	369	1
17001	Adams	1700110890	Camp Point	17001010200	55	595	1
17001	Adams	1700114780	Clayton	17001010200	32	241	1
17001	Adams	1700115859	Columbus	17001010300	90	915	1
17001	Adams	1700123412	Ellington	17001000100	103	1079	3
17001	Adams	1700123412	Ellington	17001010400	69	1429	2
17001	Adams	1700125258	Fall Creek	17001010600	97	981	1
17001	Adams	1700129288	Gilmer	17001010400	23	307	1
17001	Adams	1700135957	Honey Creek	17001010100	79	815	1
17001	Adams	1700139220	Keene	17001010100	28	354	1
17001	Adams	1700143146	Liberty	17001010300	27	341	1
17001	Adams	1700145746	McKee	17001010300	62	661	1
17001	Adams	1700148203	Melrose	17001010600	218	1685	3
17001	Adams	1700148320	Mendon	17001010100	123	1310	2
17001	Adams	1700153624	Northeast	17001010200	12	526	1
17001	Adams	1700158278	Payson	17001010300	99	996	1
17001	Adams	1700162380	Quincy	17001000100	159	1636	2
17001	Adams	1700162380	Quincy	17001000201	65	777	2
17001	Adams	1700162380	Quincy	17001000202	214	2646	4
17001	Adams	1700162380	Quincy	17001000400	91	1209	4
17001	Adams	1700162380	Quincy	17001000500	99	1417	2
17001	Adams	1700162380	Quincy	17001000600	275	2822	5
17001	Adams	1700162380	Quincy	17001000700	96	482	2
17001	Adams	1700162380	Quincy	17001000800	133	1668	3
17001	Adams	1700162380	Quincy	17001000900	110	1191	2
17001	Adams	1700162380	Quincy	17001001001	129	1452	3
17001	Adams	1700162380	Quincy	17001001002	16	661	2
17001	Adams	1700162380	Quincy	17001001100	202	2300	3
17001	Adams	1700163550	Richfield	17001010300	87	887	1
17001	Adams	1700164408	Riverside	17001010500	139	1454	2
17001	Adams	1700177057	Ursa	17001010100	5	141	1
17003	Alexander	1700390612	Cairo	17003957800	145	1698	4

Record: 1 of 3890 No Filter Search

ROLLUPS PROVIDE COUNTY, TOWN, CENSUS TRACT HIERARCHY AND AGGREGATES

ArcMap Presentation Requirements

- Add base polygon layers
- Add rollup tables
- Link through Joins and Relates
- Set symbolization, classification, color shading



www.scribekey.com

Anatomy of Drilldown

- Select high-level parent feature, view attributes
- View related child records, RELATE SELECT
- Drill down to view next level geography
- Turn layers on/off
- Zoom to selection
- Lots of individual operations, mouse clicks, can be consolidated using ArcObjects

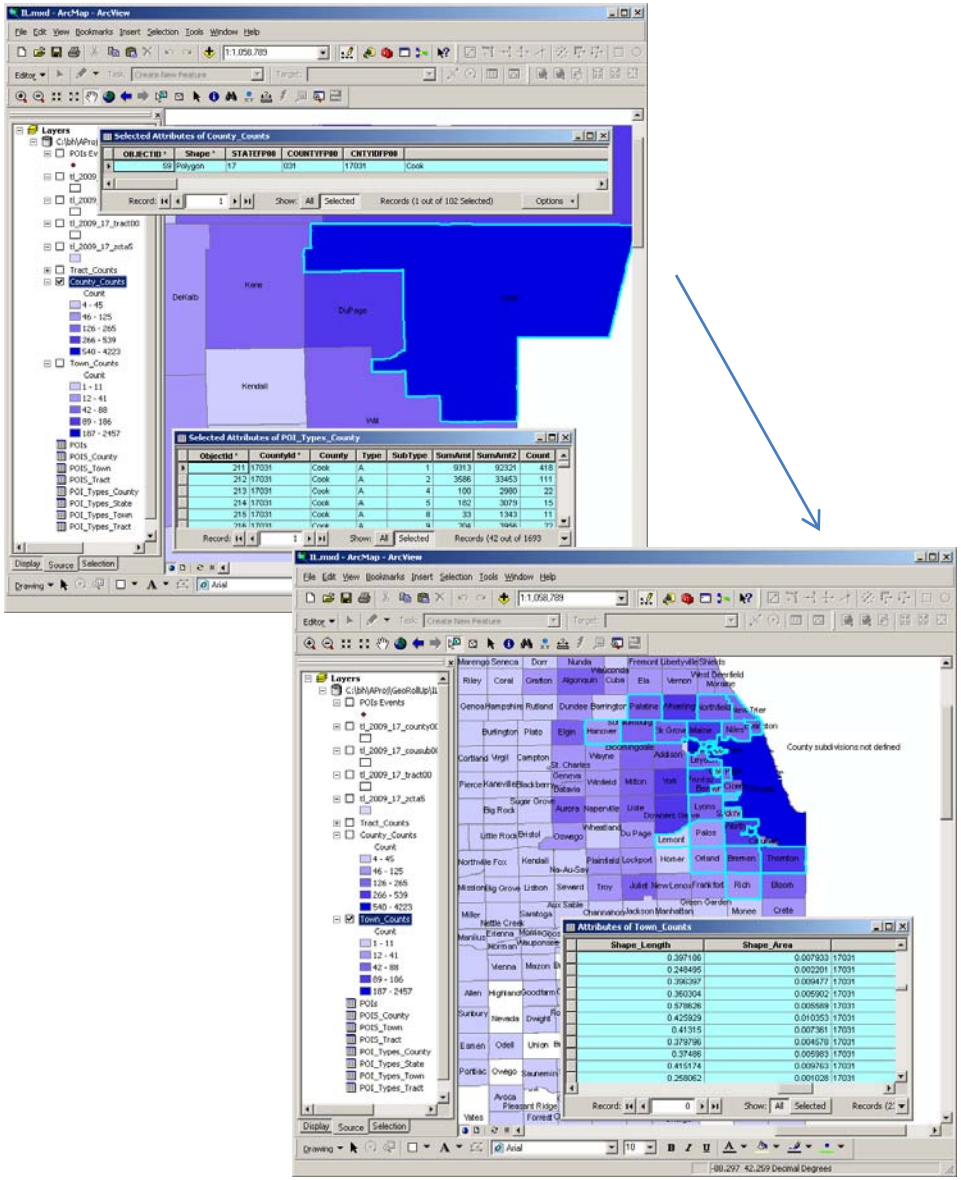
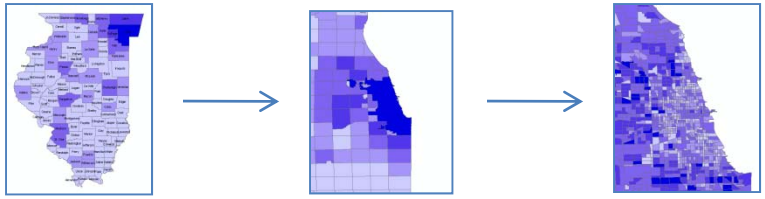


Table Driven Presentation

No.	FeatureAlias	Layer	Key	ParentLayer
1	Counties	tl_2009_17_county00	CNTYIDFP00	NA
2	Towns	tl_2009_17_cousub00	COSBIDFP00	tl_2009_17_county00
3	Tracts	tl_2009_17_tract00	CMTIDFP00	tl_2009_17_cousub00

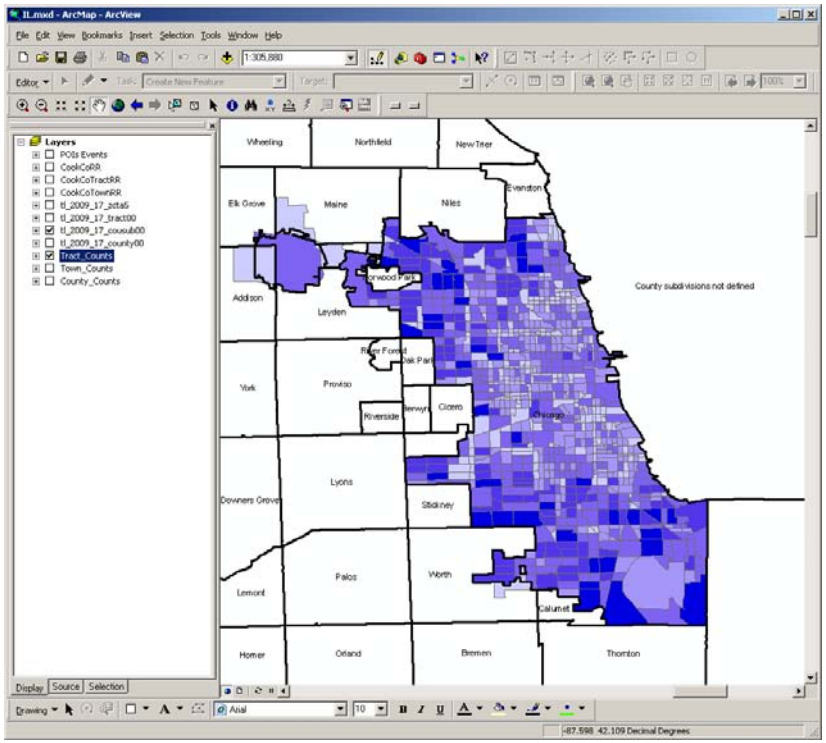
Load from DB



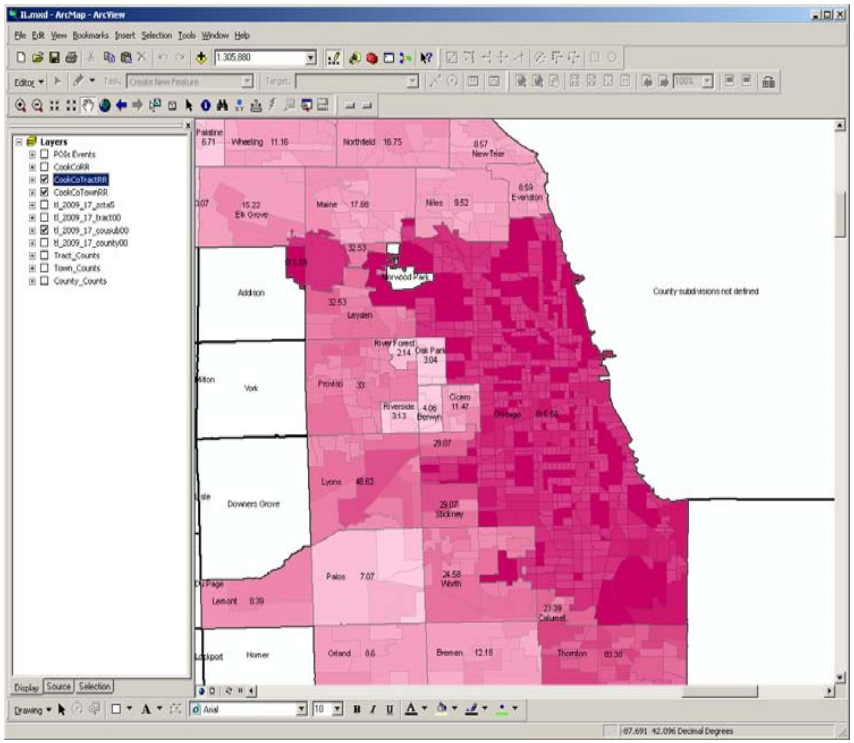
```
C# Code Facsimile: string[,] LayerFamilyTree = new string[3, 4] {
{ "Counties", "tl_2009_17_county00", "CNTYIDFP00",  "NA"},
{ "Towns",    "tl_2009_17_cousub00", "COSBIDFP00",  "tl_2009_17_county00"},
{ "Tracts",    "tl_2009_17_tract00",    "CMTIDFP00",   "tl_2009_17_cousub00"}
};
```

ArcObjects **Drilldown and Rollup** presentation tools can be driven by tabular data which captures the hierarchical relationship between parent and child feature layers. This is a form of metadata typically found in OLAP environments, and is the same information used to capture the relationships with Point-In-Polygon routine.

Additional Presentation Options



Showing only the features in the current family tree with Query Filter



Showing 2 hierarchical levels at the same time with Transparency

Other Considerations

- Using partitioned data
- Use of more advanced formulas in SQL
- Use of linear data lengths
- Use of area data measurements
- Non-hierarchical geography
- Metadata: Meta-Layers

Partitioned Data

- For a variety of reasons, data is often partitioned.
- Example: US Census data, by State, then County
- This data can be used, as-is, without needing to merge it all into one set of large layers.
- Using partitioned data can increase performance considerably.
- Consequences for 3 main GeoRollup activities:

a) SPATIAL JOIN: Perform a 2-pass operation, e.g., find the town first, then the parcel.

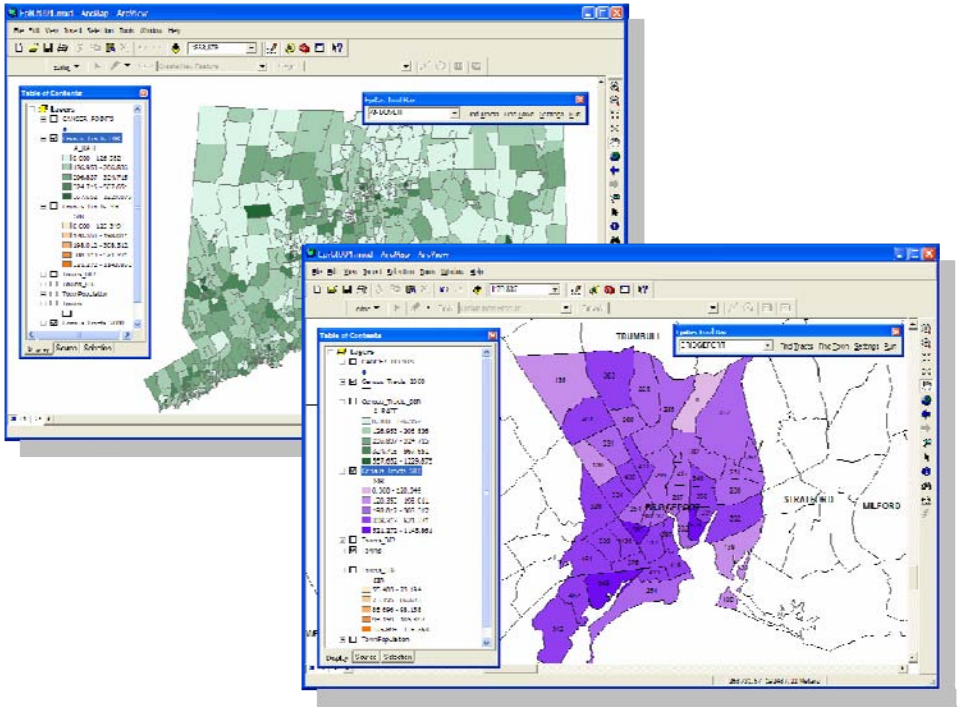
b) AGGREGATION: More granular operation with separate tables, but these can be merged using SQL UNION, etc.

c) PRESENTATION: More granular file and layer management for zoom-to-extents, layers on/off, etc.

www.scribekey.com

More Advanced Formulas

- When you want to see more than simple count(*) or sum()
- Case Study: EPIGIS Disease Surveillance
- Cancer cases by gender, age, cancer type, etc. used with population data.
- Can be expressed in SQL and aggregated by geography.
- Case count aggregation used as numerator over population to present incidence by Town, and drilldown to Census Tract.



```
SELECT Gender, Age, CancerType, Town,
        CensusTract, Count(*) AS Cases,
        Count(*)/Sum(Pop) As Rate
FROM
CancerData
GROUP BY
Gender, Age, CancerType, Town, CensusTract
```

Linear Data: Length Calculation

- Example use case: Need to see length in miles of Railroad related to population.
- May need to break lines at polygon boundaries.
- Lon/Lat data needs to be projected to use Geometry Calculator or can use great circle distance.
- Aggregate data and FEET->MILES conversion using SQL:

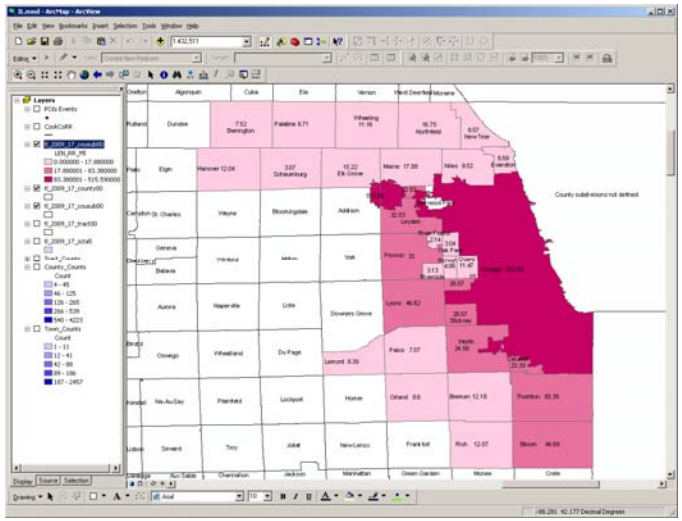
```
SELECT ROUND(SUM(LEN_FT)/5280, 2) as LEN_MI
```

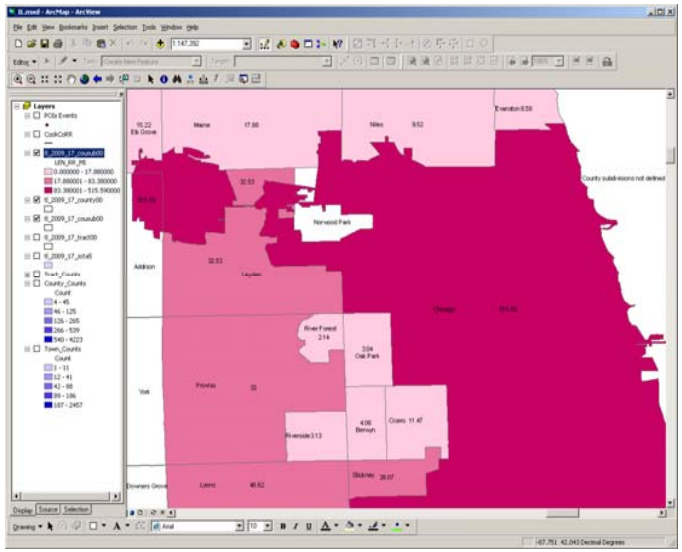
Rail Length In Miles By Town in Cook Country

- Extracted only RAILFLG='Y' from Cook Country Streets
- Projected to ILLINOIS State Plane Feet
- Performed Calculate Geometry Length
- Performed SPATIAL JOIN with Town layer
- Used SQL to perform aggregation:

SELECT
NAME00 AS Town,
Round(Sum(LEN_FT)/5280,2) AS LEN_MI
INTO RRLenByTown
FROM CookCoRR
GROUP BY
NAME00

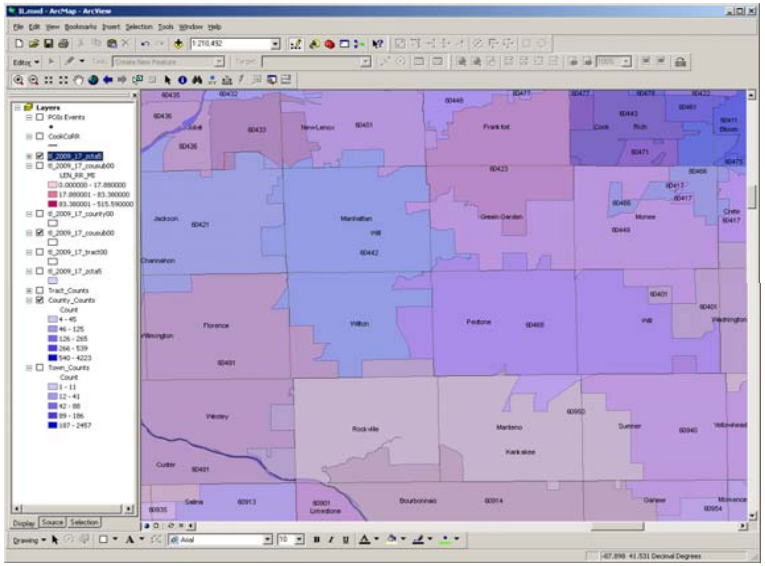
- Added new field to town layer and updated
- Added layer to ArcMap, set Symbology to 3 classes using RR_LEN_MI

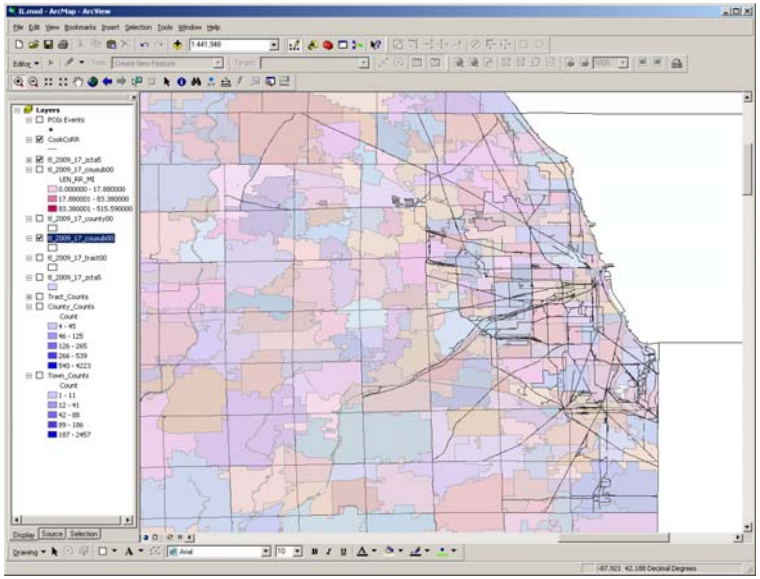




Area Measurements and Non-Hierarchical Sets

- Like points and linear features, areas can be aggregated by size, type, etc.
- Relationship between polygon layers is not always hierarchical, example: ZIP Codes and Towns. Can use intersection and hybrid layer.
- Also, phenomena is not evenly distributed over a polygon area.
- Aggregation transfer approaches can use street network or population polygons as proxy.
- Shared boundary features inclusion issue.





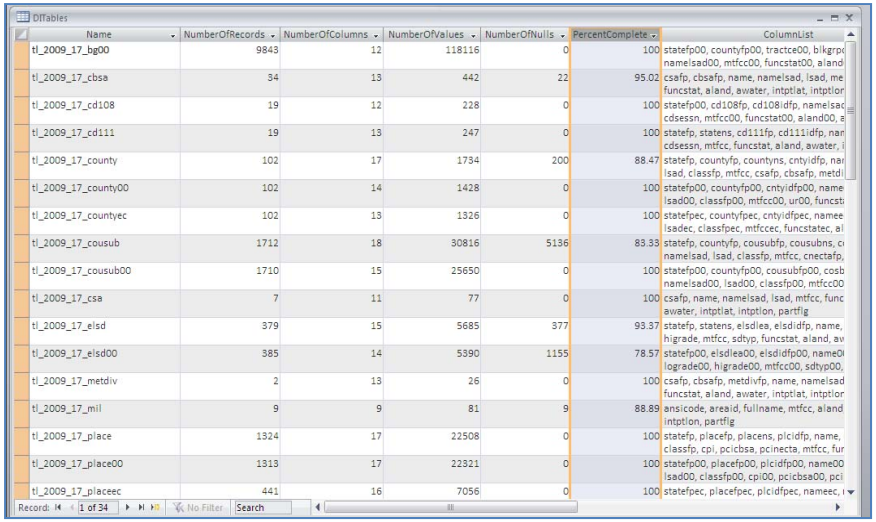
Viewing Metadata: Meta-Layers

- Different kinds of metadata:
  a) Database metadata as name, data type, length of fields
  b) FGDC metadata as info about both layer, e.g., feature geometry type, projection, and individual attributes, domains
  c) Aggregation metadata as summary count(*), sum(), etc. values
- Metadata can be managed and viewed as GIS layers, using polygons to represent coverage areas, and metadata elements loaded into RDBMS tables.

www.scribekey.com

1st Step: Data Profile and FGDC Metadata in DB

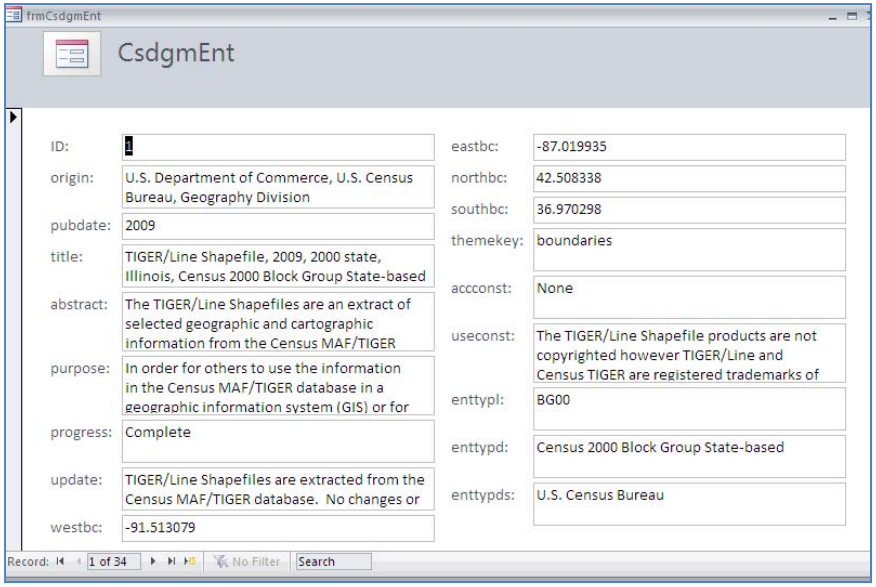
- Data profiling captures structure and contents information from data, e.g., record counts, attribute counts, attribute names, data types, percent complete, etc.
- Import FGDC Entity, Attribute, and Domain values and definitions into tabular database tables.



DITables

Name	NumberOfRecords	NumberOfColumns	NumberOfValues	NumberOfNulls	PercentComplete	ColumnList
tl_2009_17_bg00	9843	12	118116	0	100	statefp00, countyfp00, tractce00, blkgrp0 namelsad00, mtfcc00, funcstat00, aland
tl_2009_17_cbsa	34	13	442	22	95.02	csafp, cbsafp, name, namelsad, lsad, me funcstat, aland, awater, intptlat, intptlor
tl_2009_17_cd108	19	12	228	0	100	statefp00, cd108fp, cd108idfp, namelsa cdsessn, mtfcc00, funcstat00, aland00, a
tl_2009_17_cd111	19	13	247	0	100	statefp, statens, cd111fp, cd111idfp, nar cdsessn, mtfcc, funcstat, aland, awater, i
tl_2009_17_county	102	17	1734	200	88.47	statefp, countyfp, countyns, cntyidfp, nar lsad, classfp, mtfcc, csafp, cbsafp, metdi
tl_2009_17_county00	102	14	1428	0	100	statefp00, countyfp00, cntyidfp00, name lsad00, classfp00, mtfcc00, ur00, funcst
tl_2009_17_countyec	102	13	1326	0	100	statefpec, countyfpec, cntyidfpec, namee lsadec, classfpec, mtfccec, funcstatec, al
tl_2009_17_cousub	1712	18	30816	5136	83.33	statefp, countyfp, cousubfp, cousubns, c namelsad, lsad, classfp, mtfcc, cnectafp,
tl_2009_17_cousub00	1710	15	25650	0	100	statefp00, countyfp00, cousubfp00, cosb namelsad00, lsad00, classfp00, mtfcc00
tl_2009_17_csa	7	11	77	0	100	csafp, name, namelsad, lsad, mtfcc, func awater, intptlat, intptlon, partflg
tl_2009_17_elsd	379	15	5685	377	93.37	statefp, statens, elsdlea, elsdidfp, name, higrade, mtfcc, sdtyp, funcstat, aland, av
tl_2009_17_elsd00	385	14	5390	1155	78.57	statefp00, elsdlea00, elsdidfp00, name0 lograde00, higrade00, mtfcc00, sdtyp00,
tl_2009_17_metdiv	2	13	26	0	100	csafp, cbsafp, metdivfp, name, namelsad funcstat, aland, awater, intptlat, intptlor
tl_2009_17_mil	9	9	81	9	88.89	ansicode, areaid, fullname, mtfcc, aland intptlon, partflg
tl_2009_17_place	1324	17	22508	0	100	statefp, placefp, placens, plcidfp, name, classfp, cpi, pcicbsa, pcinecta, mtfcc, fur
tl_2009_17_place00	1313	17	22321	0	100	statefp00, placefp00, plcidfp00, name00 lsad00, classfp00, cpi00, pcicbsa00, pci
tl_2009_17_placeec	441	16	7056	0	100	statefpec, placefpec, plcidfpec, nameec,

Record: 1 of 34 No Filter Search



frmCsdgmEnt

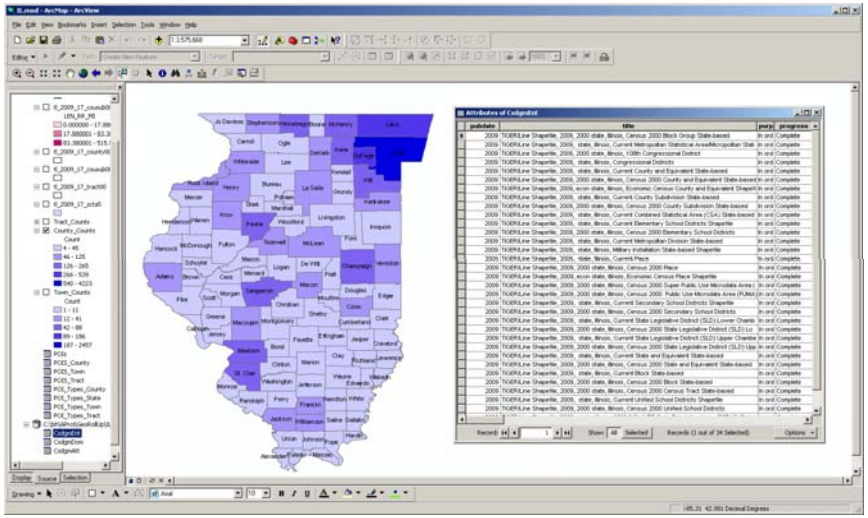
CsdgmEnt

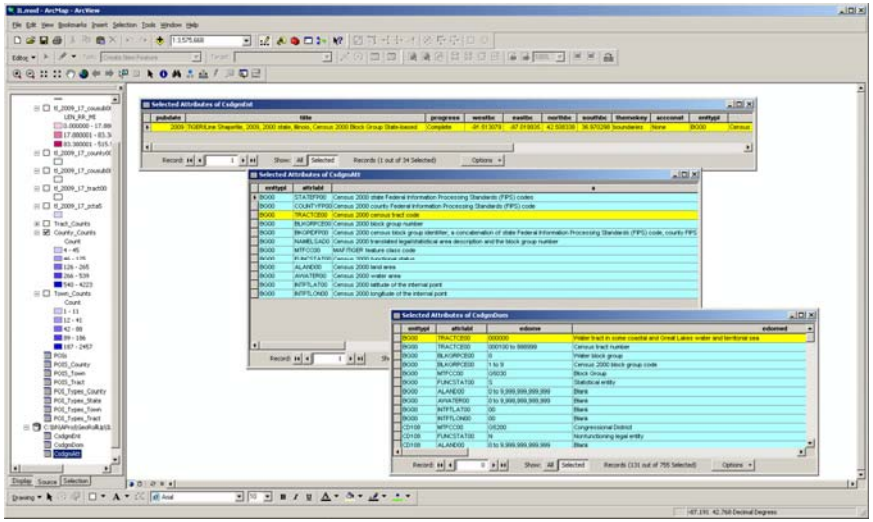
ID:	1
origin:	U.S. Department of Commerce, U.S. Census Bureau, Geography Division
pubdate:	2009
title:	TIGER/Line Shapefile, 2009, 2000 state, Illinois, Census 2000 Block Group State-based
abstract:	The TIGER/Line Shapefiles are an extract of selected geographic and cartographic information from the Census MAF/TIGER
purpose:	In order for others to use the information in the Census MAF/TIGER database in a geographic information system (GIS) or for
progress:	Complete
update:	TIGER/Line Shapefiles are extracted from the Census MAF/TIGER database. No changes or
westbc:	-91.513079
eastbc:	-87.019935
northbc:	42.508338
southbc:	36.970298
themekey:	boundaries
accconst:	None
useconst:	The TIGER/Line Shapefile products are not copyrighted however TIGER/Line and Census TIGER are registered trademarks of
enttypl:	BG00
enttypd:	Census 2000 Block Group State-based
enttypds:	U.S. Census Bureau

Record: 1 of 34 No Filter Search

Add and Join/Relate Metadata Tables in ArcMap: Meta-Layers

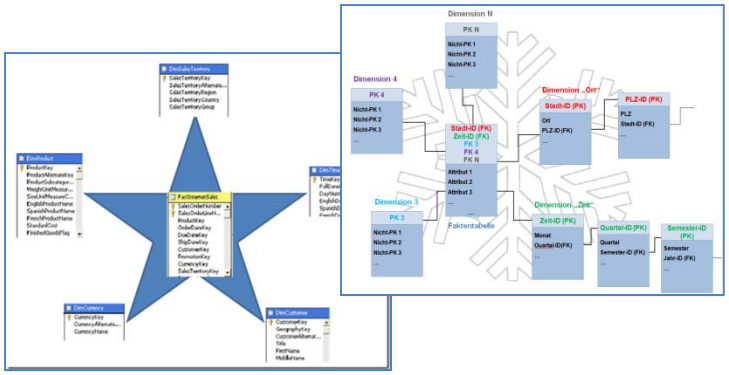
- Integrate data profiling and FGDC metadata in RDBMS
- Data quality, percent complete, etc.
- Add as table to ArcMap
- Metadata available like other tabular data in ArcGIS
- N->1 Attribute->Entity
- HTML Data Dictionary

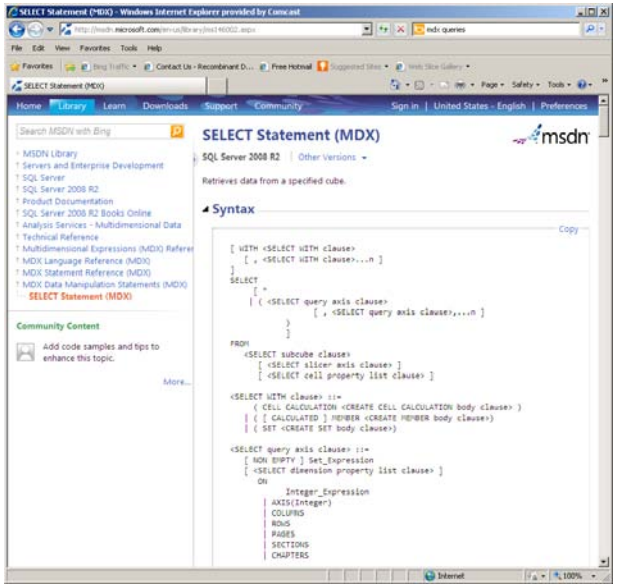


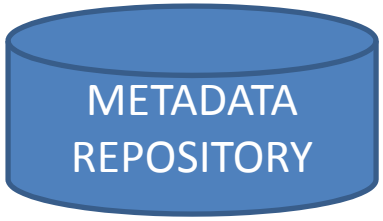


OLAP 101

- Read-only database, not normalized
- Measures and Dimensions
- Star and Snowflake schemas
- MDX, XMLA Cube Query Languages
- Metadata Repository
- There is no middle tier, pure query language behind the interfaces







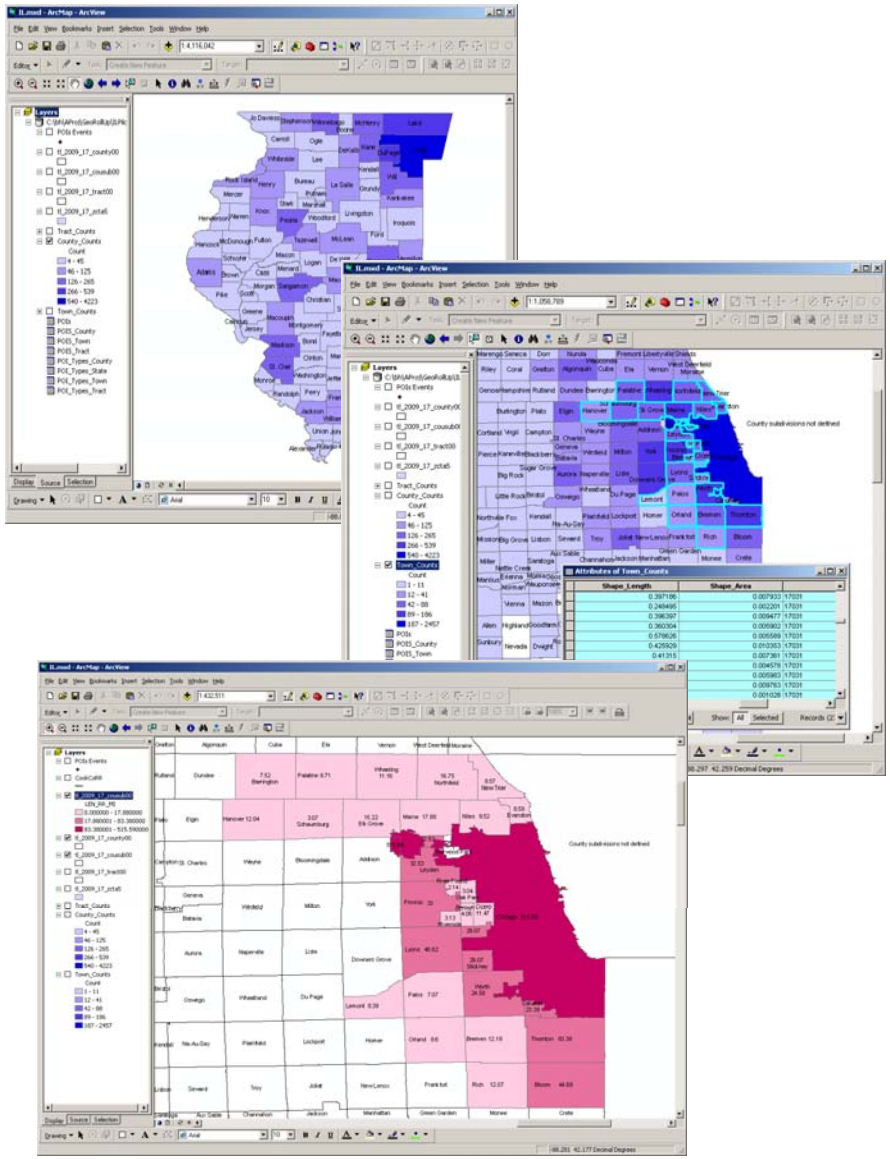
Why is there a GIS/BI Technology Gap?

- GIS has evolved independently of mainstream RDBMS OLTP/OLAP because of the variable length geometry and spatial operations.
- GIS schemas and applications have focused on read-write production databases, not read-only decision support databases.
- GIS standards bodies have focused on XML as opposed to RDBMS schemas for production databases and metadata, again, not decision support databases.
- BI tools and techniques typically don’t handle geometric operations, e.g., spatial join, dissolve, etc.
- There are some fledging solutions, SAS/ESRI Bridge, MS SQL Server RS Map Control, QlikView/Map integration.
- One ultimate goal would be to link and drive a GIS display engine from an OLAP Cube.

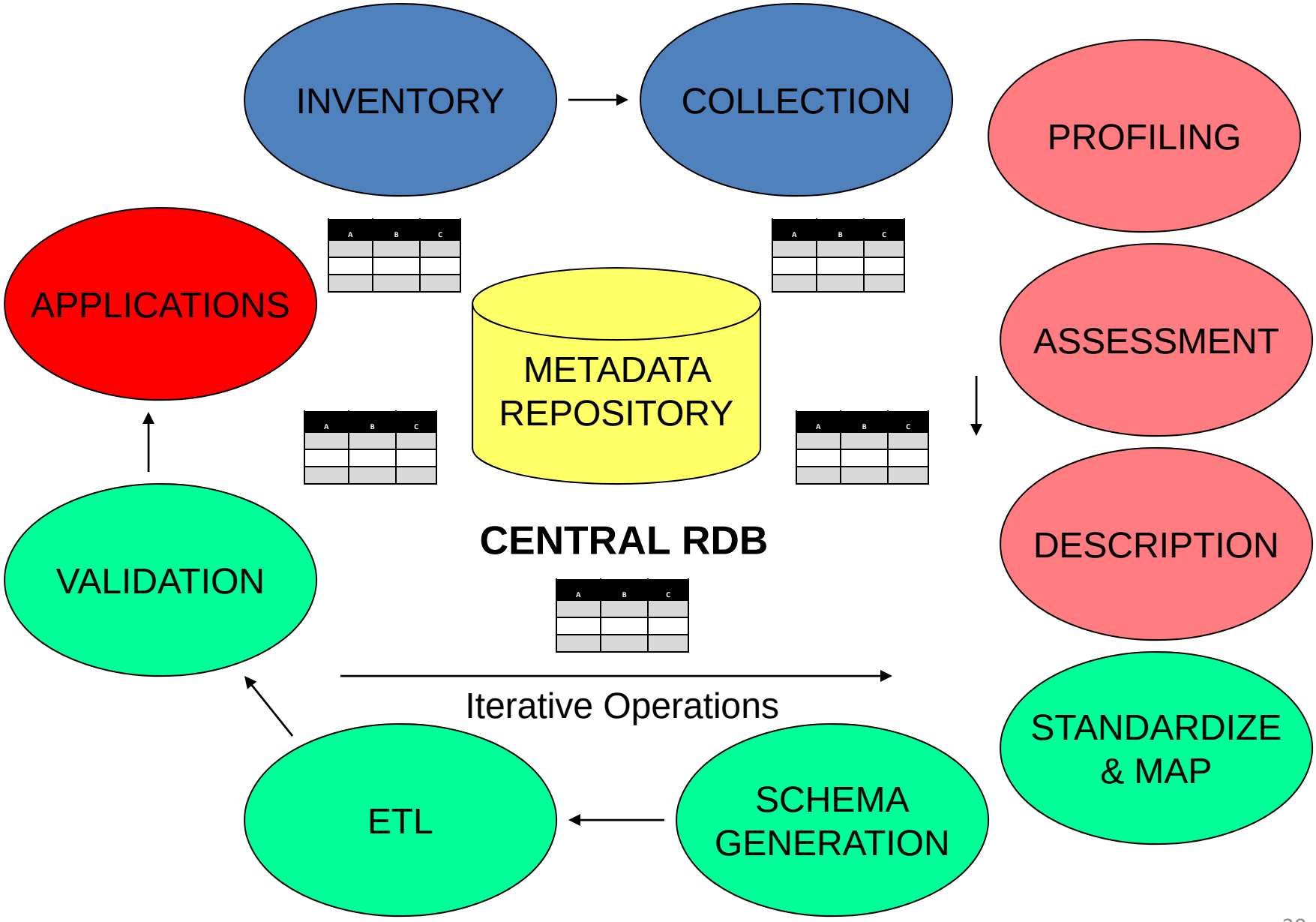
www.scribekey.com

Recap and Take Aways

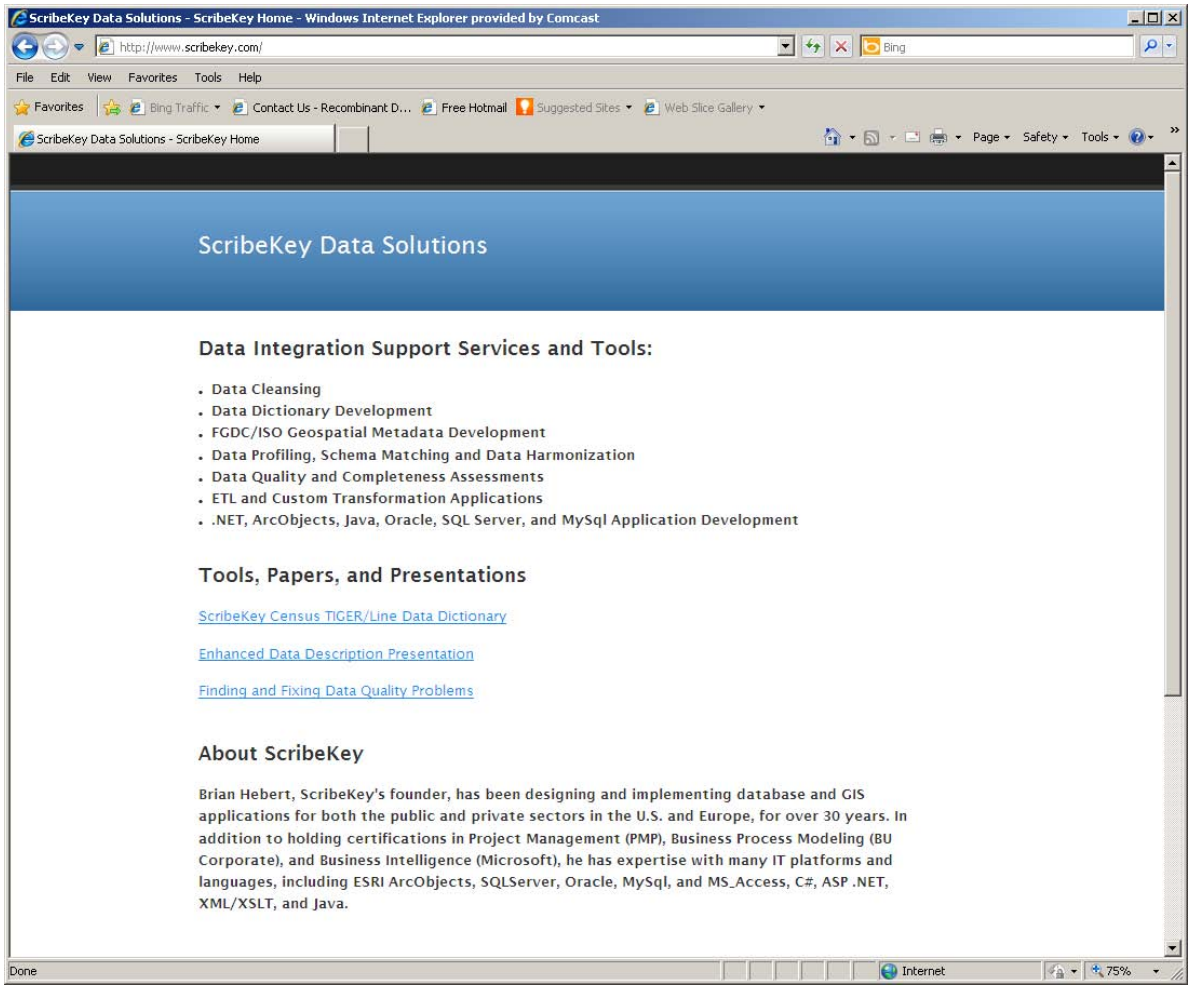
- Applying OLAP data compilation and viewing techniques can provide very rich data views and powerful decision support.
- Try using pivot tables in MS Excel or Access, get familiar with concepts, reproduce similar views in ArcMap.
- Focus on creation of base table, then rollup with SQL COUNT(*), SUM(), GROUP BY.
- Learn about and explore OLAP, BI, Data Warehousing.
- Learn about data profiling and put FGDC metadata into data tables that can be viewed in ArcMap



RDB Metadata Repository and Tools



Thank You
Questions and Answers



www.scribekey.com