Integrating an OLAP Data Cube with ArcMap

Lt. (j.g.) William Winner, NOAA Deb Agarwal, UC Berkeley Catharine van Ingen, Microsoft



Microsoft

Project Origins

- Began at UC Berkeley
 - Proposal to NSF for a watershed data synthesis center
 - Proposal to Microsoft to create a demonstration centered around California water issues (2006)



Microsoft

Original Objectives

- Database requirements
 - Needed large storage capacity
 - Needed ability to merge data from different agencies in different formats
 - Needed to handle different types of data
 - Needed the ability to perform analyses quickly



- An N-dimensional database
 - Traditional DBs can do this too, but...
 - Data must be flattened
 - Done using Aggregations, Group By's and Cross Tabs

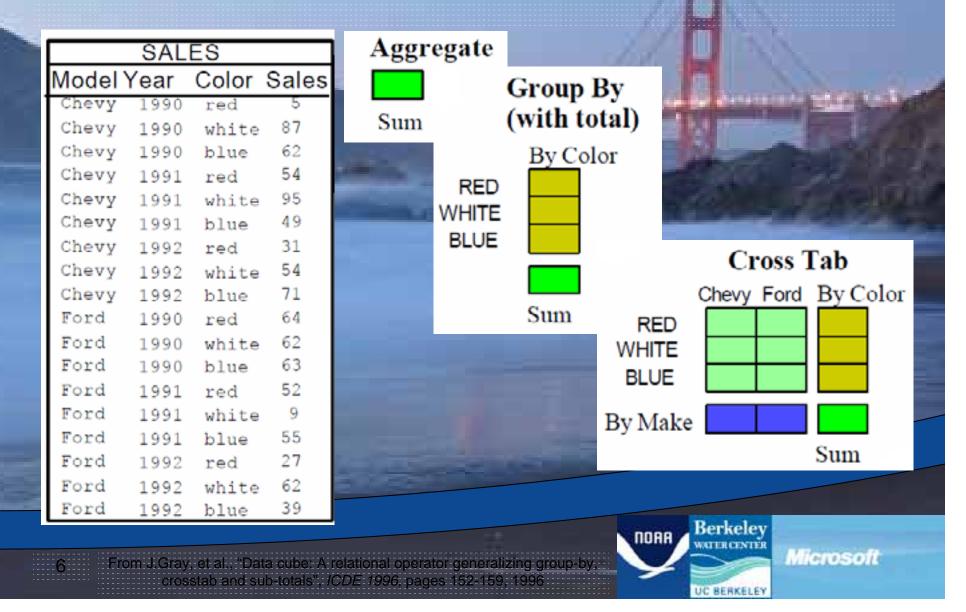


	SAL	ES	
Model `	Year	Color	Sales
Chevy	1990	red	5
Chevy	1990	white	87
Chevy	1990	blue	62
Chevy	1991	red	54
Chevy	1991	white	95
Chevy	1991	blue	49
Chevy	1992	red	31
Chevy	1992	white	54
Chevy	1992	blue	71
Ford	1990	red	64
Ford	1990	white	62
Ford	1990	blue	63
Ford	1991	red	52
Ford	1991	white	9
Ford	1991	blue	55
Ford	1992	red	27
Ford	1992	white	62
Ford	1992	blue	39

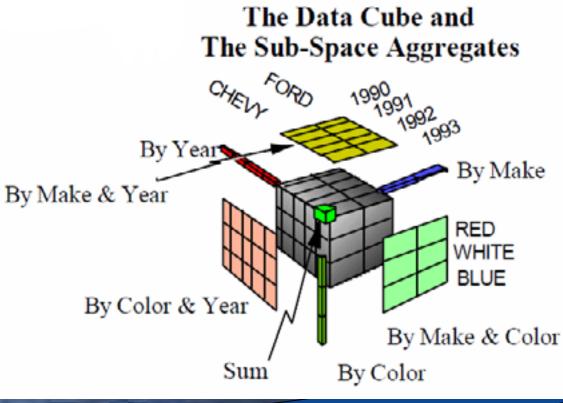


5 From J.Gray, et al., "Data cube: A relational operator generalizing group-by, crosstab and sub-totals", *ICDE 1996*, pages 152-159, 1996

Microsoft



	SAL	ES		
Model	Year	Color	Sales	
Chevy	1990	red	5	
Chevy	1990	white	87	
Chevy	1990	blue	62	
Chevy	1991	red	54	
Chevy	1991	white	95	
Chevy	1991	blue	49	
Chevy	1992	red	31	
Chevy	1992	white	54	By Make &
Chevy	1992	blue	71	
Ford	1990	red	64	
Ford	1990	white	62	
Ford	1990	blue	63	В
Ford	1991	red	52	D.
Ford	1991	white	9	
Ford	1991	blue	55	
Ford	1992	red	27	
Ford	1992	white	62	a second
Ford	1992	blue	39	



From J.Gray, et al., "Data cube: A relational operator generalizing group-by, crosstab and sub-totals", ICDE 1996, pages 152-159, 1996

UC BERKELEY

Microsoft

Berkeley

NUTERCENTER

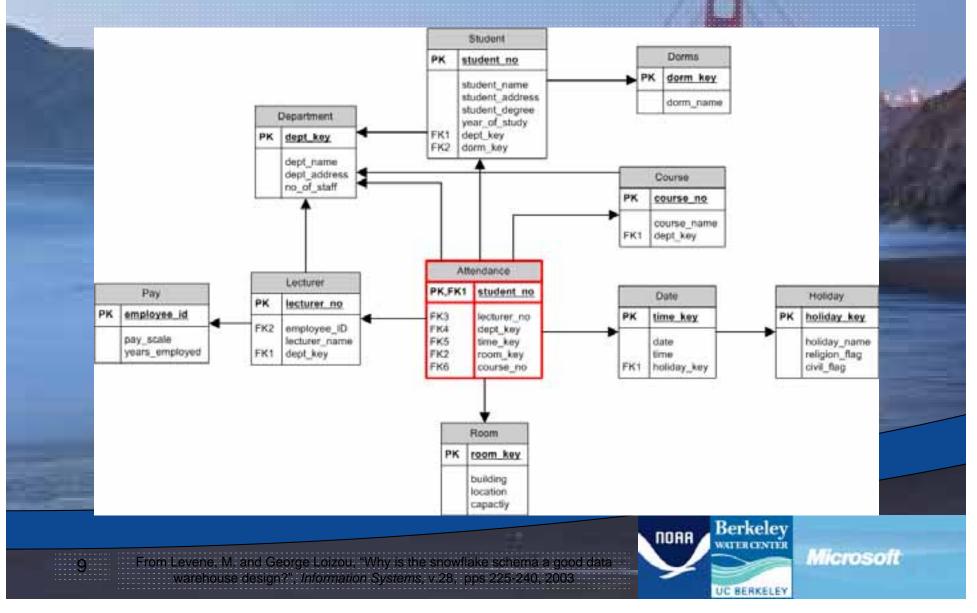
NOAR

How is the Cube stored?

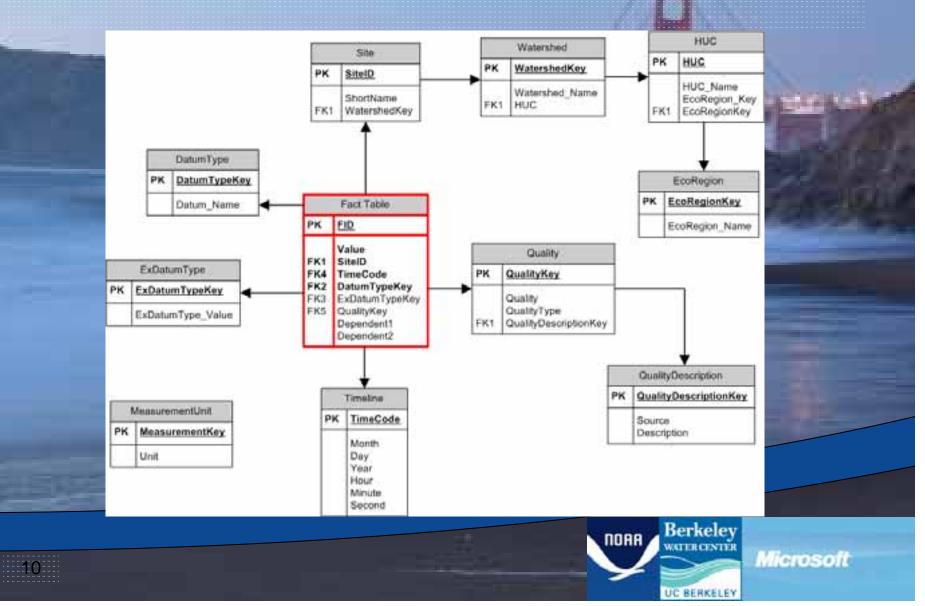
- The Cube is a computed product from a series of relational tables
- Our Cube is based on a Context-Dependent Snowflake Schema



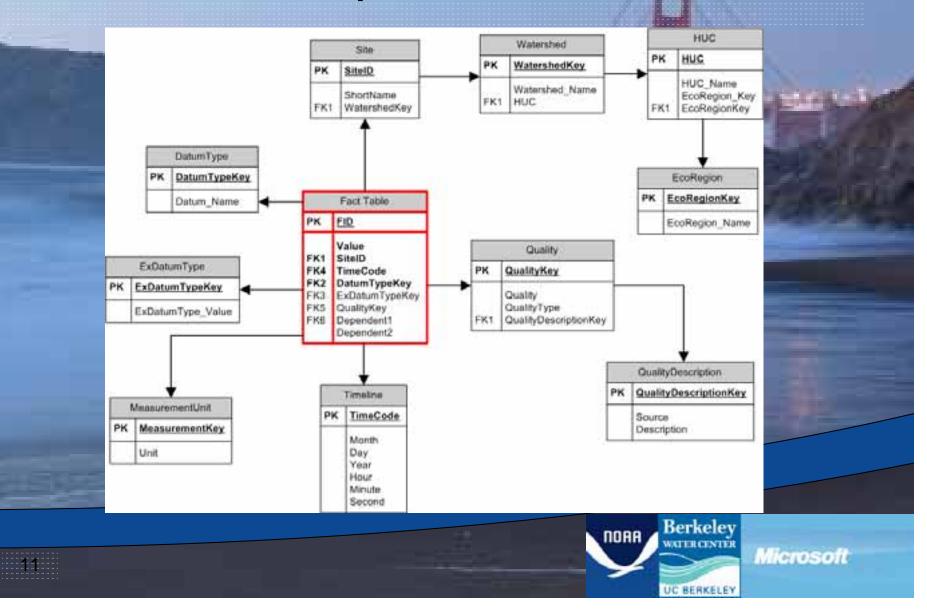
Snowflake Schema



Snowflake Schema



Context-Dependent Snowflake



NMFS Involvement

- Came aboard at the end of 2008
- Saw the potential to use this to help inform Salmonid Recovery Plans
- Saw the potential to include a spatial component with the data



Toolbar Requirements

- Allow for spatial queries
- Flatten data to fit within a shapefile's attribute table
- Provide some way to visually move through the data



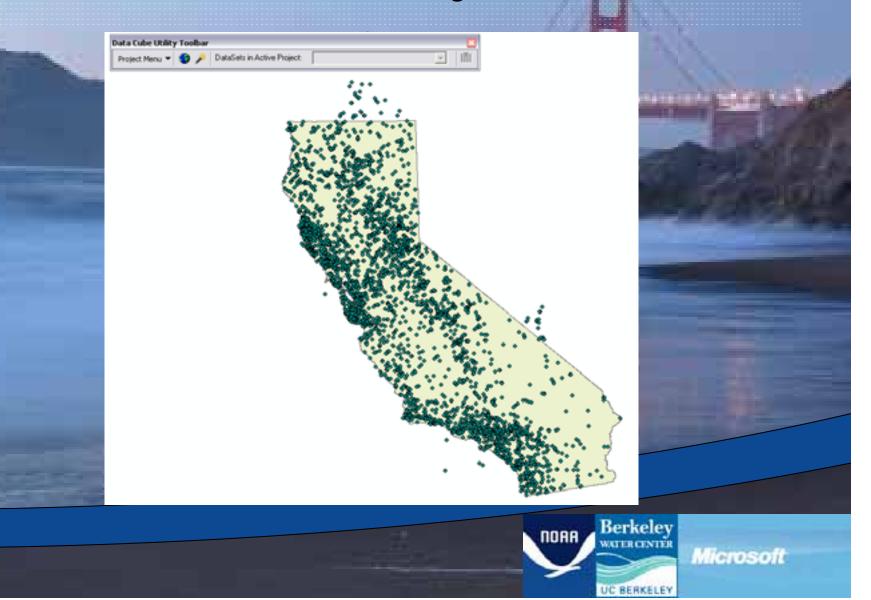
How did we do it?

- The toolbar was written using Visual Studio 2008 along with the ArcGIS API's.
- Also included project management and storage
- Written with just shy of 10,000 lines of designer- and user-generated code



-	L .	be Utility Joli Menu 👻 🜍 🌡		Active Project				
-		ew Project						
		ben						Cardina - Securit
		ive						
		We As						The second s
		ose Project						
	_	oject Manager						and the second second second
		tive Project	,					and in case of the local division in which the local division in which the local division is not the local division in the local div
		ataSets	,	-				and the second
	() Al	out DataCubeUti	ity Toolbar					Summer of the local division of the local di
Constraint Const	2011	17						
		10000	-	-	1000	The state of the s	the state of the s	
		-						

40 4



	Site Selection Wizard			
	Step 1			
	Selection Method © Use Current Selection © Select by Proximity to Another Layer © Select by ID			
	Cancel Back	Next	Proceed to DistaCub	
-				

Step 1 Step 2			
Select Layer to Use: Navarro			22.
Use Buffer	Meters	-	
			•
Cancel	Back Next	Proceed to DataCube	•
	1. N. N.		-
	° %2	a	

Step 1 Step 2 Step 3	zard 			
Select Type(s) of Sil				
Select A				
VOAA S				
SEC Site				
J• 300 516	-			
				8
				- •, E51
Cance	Back.	Next	Proceed to DataCu	be
		N. N.	A	••
		12		,
			· · · · ·	* • • • • • • • • • • • • • • • • • • •

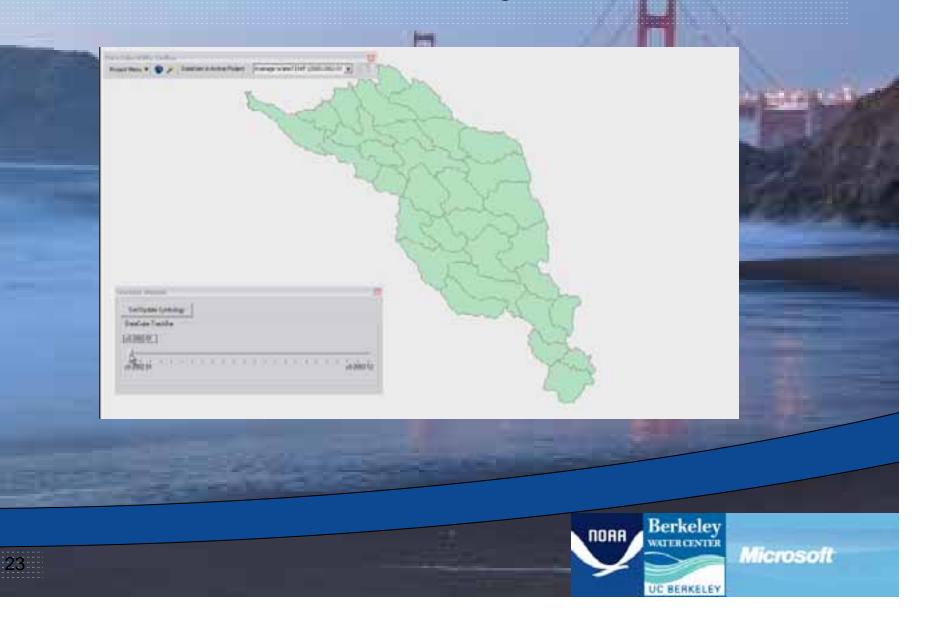
 ✓ 040973 ✓ 046105 ✓ 049051 ✓ 11467800 ✓ 11467850 ✓ 11460000 ✓ FSPS-2099 ✓ FSPS-2104 ✓ FSPS-2107 ✓ FSPS-2110 ✓ FSPS-2111 ✓ MCWAS-11 				
MCWAS-116			5 200 250 2099 2104 2107 2110 2111 5-11	 046105 049051 114678 114678 114660 FSPS-2 FSPS-2 FSPS-2 FSPS-2 FSPS-2 MCWA
Cancel Back Next Proceed to DataCube	Next Proceed to DataCube	Back	Cancel	

The state of the	Data Cube Selection Datasets to Daray		
			and the second s
	<u>e c</u>		State State State
	DatesType EdDatasType v/aesTEMP][Nine]	Heesan Armage	
	Column Selection Tensine 1 Tensine 2 Decade to Month (*) Dispare		
	Film Quelly		
		1	
	Tenine		All of the local division of the local divis
	Tautra File 1		the second se
		728 729 720 700 700 700 700 700 700 ₩	
		03 04 💌	
SIETER	Submit Dumpful	l.	
		ПОАА	Berkeley
			Microsoft

. 110	Shape	Shell	c8-2967-81	18-3842-82	18-2962-83	10.2942-84	18-2842-85	48-2962-86	+9-2942-47	19-2082-88	#8-2982-69	c6-2982-10	18-2982-11	+8-2982-12	#8-2985-81	+8-2983.42
	Port	16-881-2007MA	-30999	-0000	- 2009	-9909	17.548	19.4705	21.0812	20.009	18,7929	15.5162	- 59999	-9998	-2005	-399
. 1	Point.	MPC06-106-21	-9909	-30900	-9999	-99908	/95928	-0000	-8899	-8900	-9980	-09033	-39988	-99988	-9000	-099
. 3	Part	APRO05-T06-20	-9999	-8909	-2009	-9999	-9994	-39999	-8999	-8900	-34999	-2000	-9908	-99999	-0005	-399
- 1	Port.	MPC06-186-00	-3009	-39900	-99999	-9999	12 5839	16.3545	18.6492	17 433	15.2041	12.5995	-9009	-79998	-0093	-399
. 4	Port	MI1CO5-T06-01	-9009	-8999	-8999	-9999	15.0010	18.3099	18.5468	17,7916	16.1152	13.0123	.9999	-9999	-9999	-096
1	Port	MRC06-185-21	-2099	-9999	-99999	-99993	-99993	20999	-99999	-09999	-9099	-0999	-99993	-09999	-9999	-000
1	Point	MRC05-T05-20	-9999	-9999	-9999	-9999	-9999	-0999	-9999	-0999	-9999	-9999	.9999	-99998	-9999	.09
1	Point	MRC05-185-01	-39999	-9999	-99999	-99999	-9999	-9999	.9999	-10909	.9999	-9999	.99999	-99999	-0999	-39
1	Point	MPC05-765-08	-99999	-99999	-9999	-9998	-99993	-9999	-99999	-9999	-99999	-9999	-99999	-99998	-9995	-199
. 1	Port	MPC05-T05-05	-99999	-8995	-9999	-9999	-9999	-9999	-8999	-9900	-9998	-9908	-9998	-99998	-9999	-89
	Part	MRC05-785-03	-9999	-3990	-9999	-9999	-99993	-0099	-8999	-8909	-9999	-2993	-99998	-9900	-0093	-09
. 11	Post	MRC05-105-02	-9999	-9090	-9999	-99995	-9995	-3099	-8999	-6999	-9909	-9999	-9099	-99995	-9995	-99
	Port	MPC05-125-01	-3999	-4099	-9999	-9999	13.8299	15.8442	17.2338	15.6158	14.4054	11.0052	-9999	-99999	-0000	-49
- 13	Port	MRC05-182-28	-9999	38999	-9999	-9999	.0000	-30999	-9999	-8996	-3000	-9999	-9099	-9999	-0095	-00
14	Post	MRC05-182-21	-3009	-9988	-5909	-99999	-9999	.0000	-9999	-3995	-9009	-99019	-99998	-99998	-9999	
11	Port	MRC05-182-28	-99999	-8998	2009	-9909	-99995		-8999	-8900	-30099	-2099	90099	-0094	-0008	
11	Polit	MPC05-183-25	-9999	-30909	-0009	-0008	/9993	-0000	-8800	-8100	-9990	-3903	-30088	-99988	-0000	
17	Plat	AR0005-T82-24	-9999	-8909	-3999	-9999	.0095	-3999	-8999	-8999	-30000	- 2909	-9904	-9999	-4999	-89
10	Fort.	MINCOS-182-23	-33999	-39909	-99999	-9999	-19999	-3000	-3999	-8908	-8999	-9998	-99998	-79991		-00
- 10	Fore	MRC05-T62-22	-9999	-8990	-9999	-9999	-0990	.0099	-3999	.0999	-9699	-9999	.99995	-9699	.0000	-09
31	Fort	MRC06-182-21	-2020	-9999	-99999	-9999	99993	-3999	-8099	-0999	-9999	-9999	-99999	-0094	-0099	-09
	Point	MPC05-T02-09	-9999	-9999	-8999	-9999	10.0105	11.0453	12.0004	12:5742	11.0194	9.9572	.9999	-9999	.0990	.09
	Point	MPIC26-182-08	-9999	-8999	-99999	-99999	10.9682	12,4813	14.1535	15,2504	12.3407	9,8297	-9999	-9999	-0000	-39
	Post	MPC05-103-07	-99999	-9999	-9999	-9998	9.8725	11.7306	13.199	12 8518	12.2219	10.5234	-99993	-99998	-9999	-99
- 24	Ford.	MPC05-TED.OIL	-9999	-89995	-9999	-9999	10.0904	11.0421	13.0078	12.7185	12.2764	10.7279	-9998	-99998	-9999	-89
	Fort	MRC05-782-05	-9999	-3990	-9999	-99999	17.2049	30.0604	21 4009	19.3965	10.4017	15.3053	-99909	-99999	-0093	-09
	Port	MRC05-182-04	-9999	-36360	-9999	-9999	-9995	-9999	-8999	-8999	-9999	-9999	-9099	-9999	-9995	-99
21	Port	MR005-182-03	-39999	-9999	-99999	-9999	-9999	-0000	-3999	-3999	-9999	-0999	-99999	-99999	-0000	-09



Microsoft



Future Goals

- Integrate spatial attributes directly into the cube
- Improve post-query speed
- Allow for other cubes
- Incorporate other data types

