Status of Delivery of California's Official Redistricting Database

Citizen's Redistricting Commission; 3/24/2011

Karin Mac Donald, Statewide Database, UC Berkeley

Robin Johansen, Remcho, Johansen and Purcell

Michael Wagaman, Speaker's Office of Member Services

Doug Yoakam, Senate Minority Caucus

Why are We Talking to You About the Statewide Database?

- (b) The Legislature shall take all steps necessary to ensure that a complete and accurate computerized database is available for redistricting, and that procedures are in place to provide the public ready access to redistricting data and computer software for drawing maps. Upon the commission's formation and until its dissolution, the Legislature shall coordinate these efforts with the commission.

Who Created the Statewide Database?

- A bipartisan agreement after Prop. 140 passed in 1990 led function of creating the redistricting database to be moved out of the Legislature to an independent body

- To create a database that was non-partisan, academic, public, free and always available

- Board of Directors included Republican Bob Naylor, Democrat Jim Wisley and Independent Paul McKaskle

- Not cheap! Legislature invests millions of dollars each decade to create and distribute

What is in the Official Dataset?

- Same data that was produced for 2001 redistricting
  - Census Geography
  - Census Data (PL94-171)
  - Statement of Registration Data for 10 years
  - Statement of Vote for 10 years

- Added for 2011
  - Citizen Voting Age Population (CVAP)
  - Selected Older Elections
  - Documentation
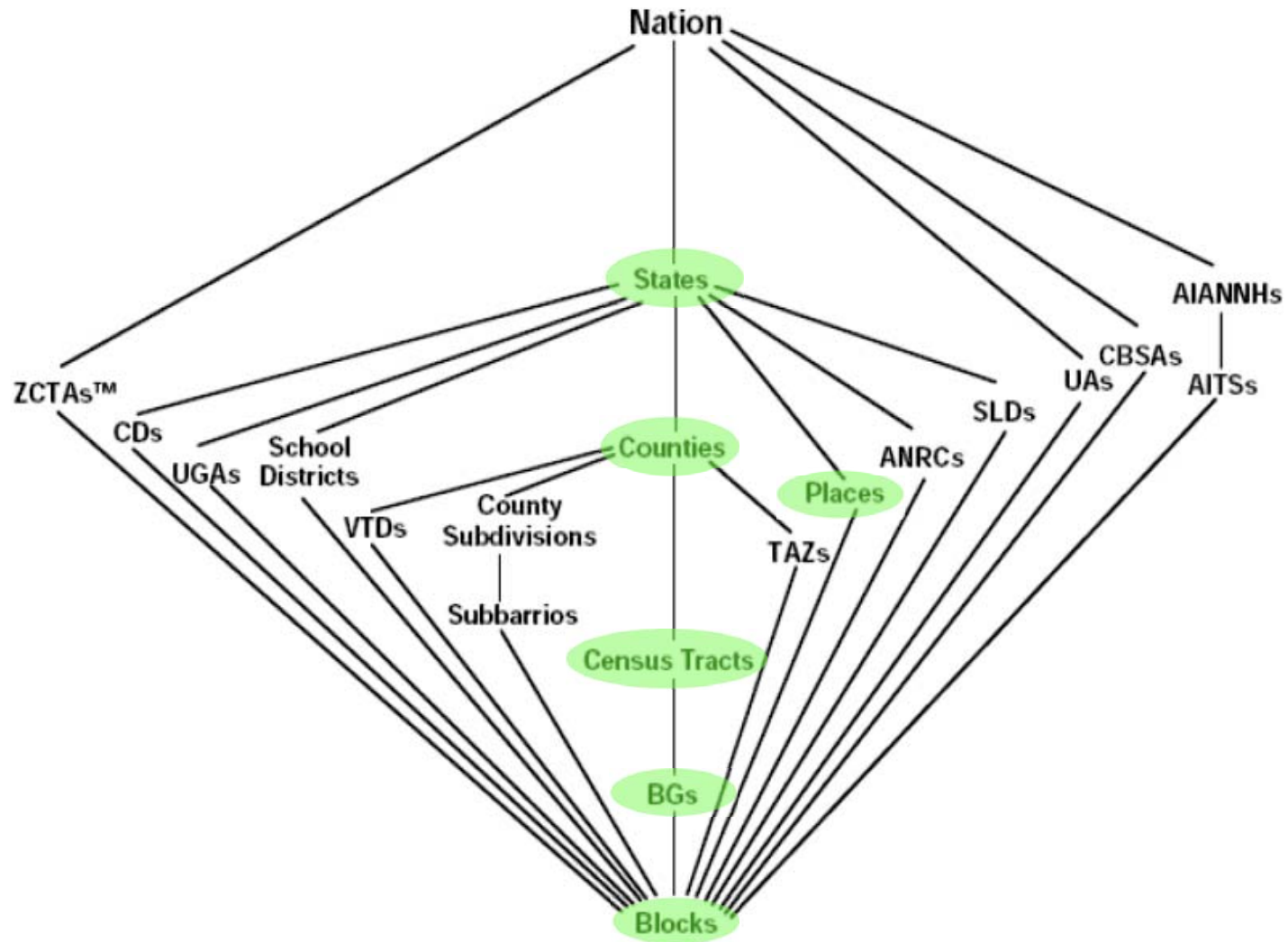
What is in the Official Dataset?

- Does not include
  - American Community Survey
  - Adjusted Census Data (i.e. Prisoners)
  - Racially Polarized Voting Analysis
  - Local Election Results

- Does not preclude Commission's supplemental database
  - For example, planned community of interest database based on public testimony

When Do You Get the Data?

- Anticipated delivery data is April 11th

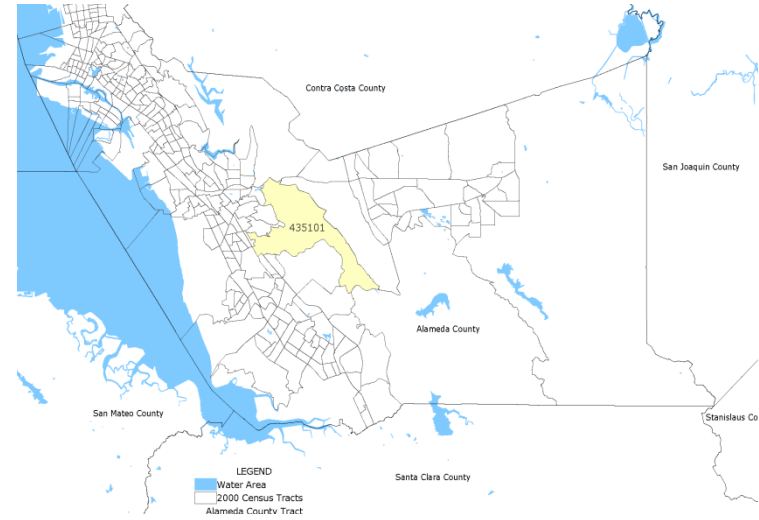- Will be simultaneously released to the public through the Statewide Database's website (http://swdb.berkeley.edu)

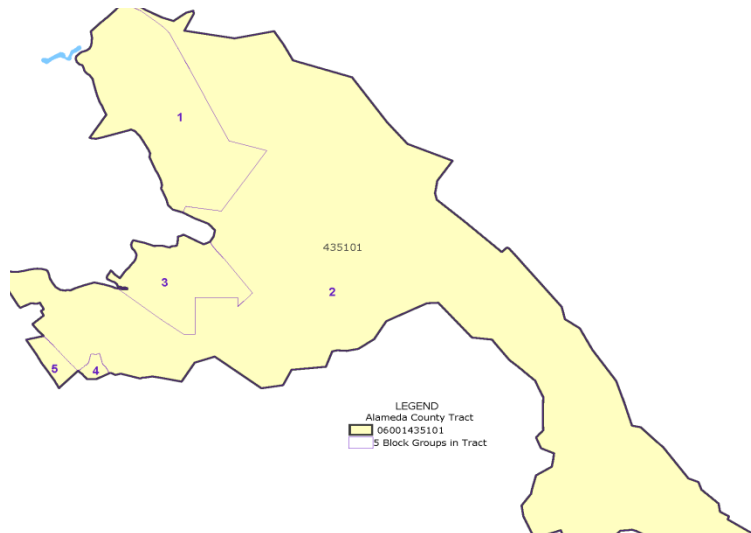# Census Geography

Census Geography

California is divided into counties….

Counties are divided into tracts…

Tracts are divided into block groups…

And block groups are divided into blocks

Census Geography

- **58 Counties**
- **1,523 Census Places**
- **8,057 Census Tracts**
- **23,212 Census Block Groups**
- **710,145 Census Blocks**

Census Data

- Basic Information/ Data that jurisdictions traditionally use for Redistricting

- Census 2010 Redistricting Data (Public Law 94-171, or PL94) contains the count of the U.S. population

- BLOCK-LEVEL dataset

- Includes data on people's race and ethnicity, for both the total and the voting age population

- Information is based on answers to the questions in the Census 2010 Short-Form questionnaire.

- There are 5 detailed tables available in the PL94-171 data product.

Census Data

P1. Race

Universe: Total population

Total:

        Population of one race:

                White alone

                Black or African American alone

                American Indian and Alaska Native alone

                Asian alone

                Native Hawaiian and Other Pacific Islander alone

                Some other race alone

Repeats for the Population of two or more races…..

P2. Hispanic or Latino, and Not Hispanic or Latino by Race

Universe: Total population

Total:
Hispanic or Latino

Not Hispanic or Latino:

        Population of one race:

                White alone

                Black or African American alone

                American Indian and Alaska Native alone

                Asian alone

                Native Hawaiian and Other Pacific Islander alone

                Some other race alone

Repeats for the Population of two or more races…..

Census Data

P3. Race For The Population 18 Years and Over
P1 variables are repeated for the Population 18 Years and Over

P4. Hispanic or Latino, and Not Hispanic or Latino By Race For The Population 18 Years And Over
P2 variables are repeated for the Population 18 Years and Over

H1. Occupancy Status
Universe: Housing units
Total:
        Occupied
        Vacant

Statement of Registration

- Registration data is important to any racially polarized voting (RPV) analysis

Statement of Registration

- Voters are geocoded to individually assign them to their proper census block.

Statement of Registration

	REG	ABS	POLLV	VOTE
California Statewide	codebook by rgprec by rrprec by srprec	codebook by rgprec by rrprec by srprec	codebook by rgprec by rrprec by srprec	codebook by rgprec by rrprec by srprec
County 001 (Alameda)	by rgprec by rrprec by srprec	by rgprec by rrprec by srprec	by rgprec by rrprec by srprec	by rgprec by rrprec by srprec
County 003 (Alpine)	by rgprec by rrprec by srprec	by rgprec by rrprec by srprec	mail-ballot county	by rgprec by rrprec by srprec
County 005 (Amador)	by rgprec by rrprec by srprec	by rgprec by rrprec by srprec	by rgprec by rrprec by srprec	by rgprec by rrprec by srprec

Statement of Registration

Variable	Start	End	Format	Description
TOTREG	16	20	F5.0	Total registration
DEM	21	25	F5.0	Democratic Party Registration
REP	26	30	F5.0	Republican Party Registration
AIP	31	35	F5.0	American Independent Party Registration
PAF	36	40	F5.0	Peace and Freedom Party Registration
MSC	41	45	F5.0	Miscellaneous Registration
LIB	46	50	F5.0	Libertarian Party Registration
NLP	51	55	F5.0	Natural Law Party Registration
GRN	56	60	F5.0	Green Party Registration
REF	61	65	F5.0	Reform Party Registration
DCL	66	70	F5.0	Declined to State Registration
MALE	91	95	F5.0	Male
FEMALE	96	100	F5.0	Female

Statement of Registration

Variable	Start	End	Format	Description
HISPDEM	101	105	F5.0	Latino Democrats
HISPREP	106	110	F5.0	Latino Republicans
HISPDCL	111	115	F5.0	Latino Declined to State
HISPOTH	116	120	F5.0	Latino Other Party

REPEATS for:

Jewish

Korean

Japanese

Chinese

Asian Indian

Vietnamese

Filipino

Statement of Registration

Variable	Start	End	Format	Description
DEMMUNK	291	295	F5.0	Dem Male age unknown
DEMM1824	261	265	F5.0	Dem Male age 18-24
DEMM2534	266	270	F5.0	Dem Male age 25-34
DEMM3544	271	275	F5.0	Dem Male age 35-44
DEMM4554	276	280	F5.0	Dem Male age 45-54
DEMM5564	281	285	F5.0	Dem Male age 55-64
DEMM65PL	286	290	F5.0	Dem Male age 65 Plus
DEMFUNK	326	330	F5.0	Dem female age unknown
DEMF1824	296	300	F5.0	Dem female age 18-24
DEMF2534	301	305	F5.0	Dem female age 25-34
DEMF3544	306	310	F5.0	Dem female age 35-44
DEMF4554	311	315	F5.0	Dem female age 45-54
DEMF5564	316	320	F5.0	Dem female age 55-64
DEMF65PL	321	325	F5.0	Dem female age 65 Plus

REPEATS FOR:
Republican
Declined to State
Other Party

Statement of Registration

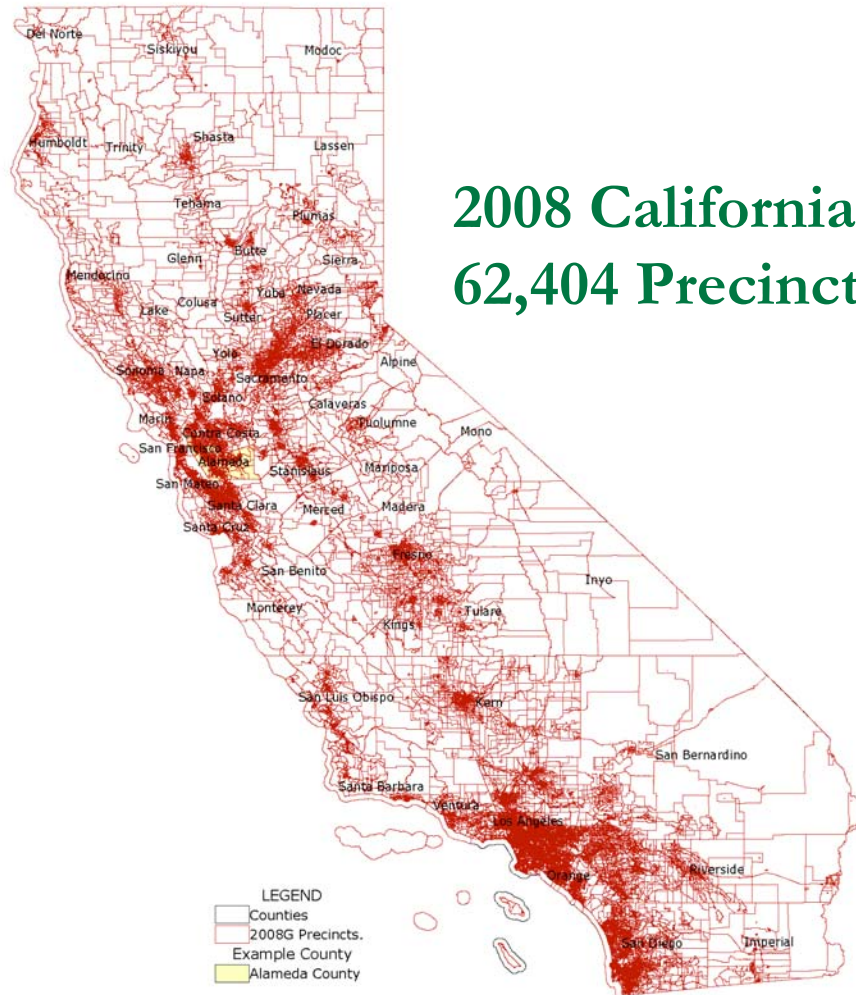| Variable	Start	End	Format	Description
RREG1G	541	545	F5.0	REP registered 1 cycle
RREG2G	546	550	F5.0	REP registered 2 cycles
RREG3G	551	555	F5.0	REP registered 3 cycles
RREG4G	556	560	F5.0	REP registered 4 cycles
RREG5G	561	565	F5.0	REP registered 5 cycles
RREG6G	566	570	F5.0	REP registered 6 cycles
RREG7G	571	575	F5.0	REP registered 7 cycles
RREG8G	576	580	F5.0	REP registered 8 cycles
RREG9G	581	585	F5.0	REP registered 9+ cycles

REPEATS FOR:

Democrat

Declined to State

Other Party

Statement of Vote

- Electoral data is important to any racially polarized voting (RPV) analysis

- Individual level data are not available

- Precincts are the smallest unit for electoral data and many change from election to election

- Precincts do not correspond to census geography

- Need to convert precinct data to census geography

- Involves statistical procedure called ecological inference
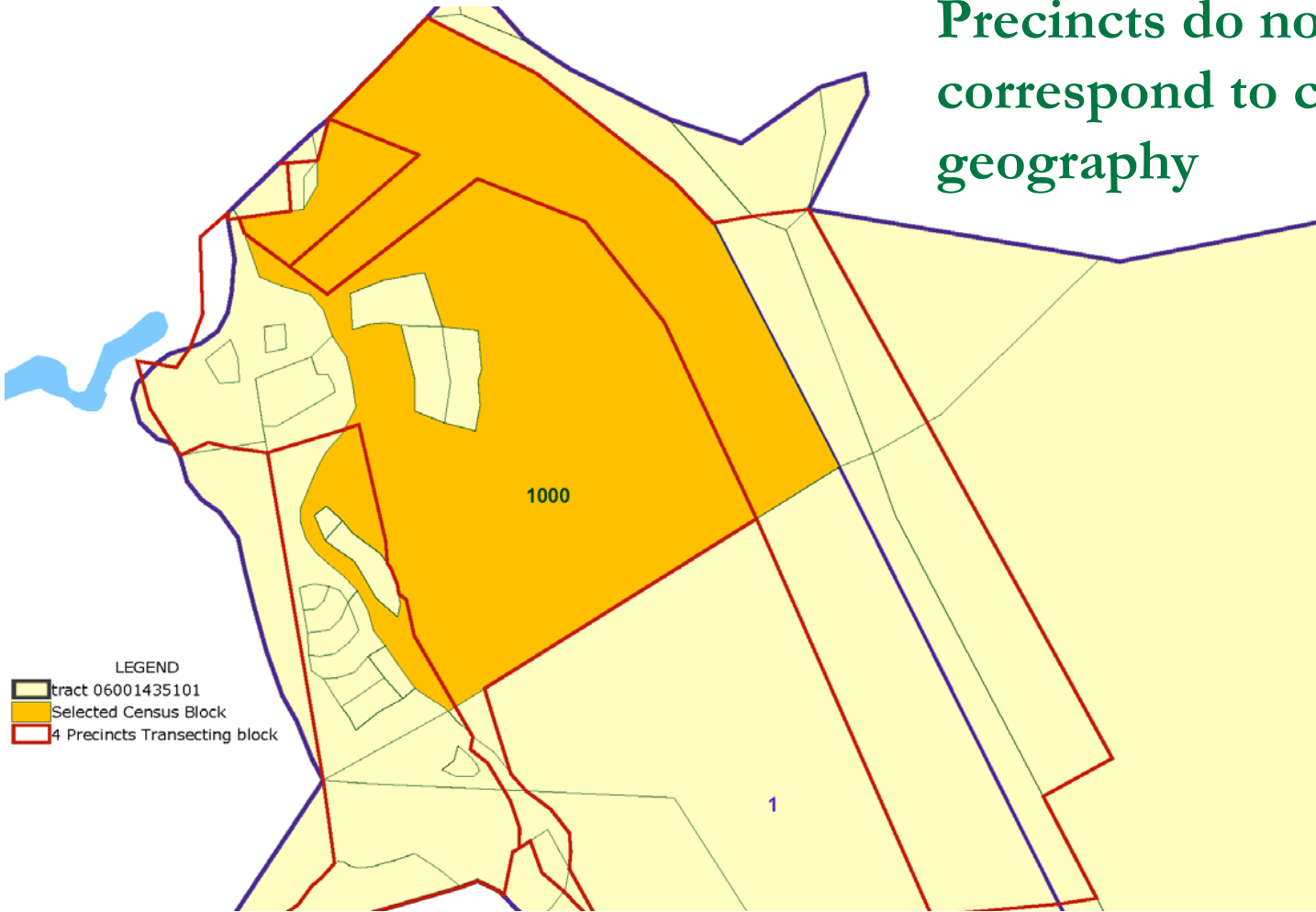
- Process will be documented

Statement of Vote

**2008 California General Election
62,404 Precincts in the State**

Statement of Vote

Precincts do not correspond to census geography

1000

LEGEND
tract 06001435101
Selected Census Block
4 Precincts Transecting block

1

Statement of Vote

- For each election database includes:

  - Turnout
  - Statewide Contests
  - Legislative Contests
  - Propositions

What are the Advantages of the SWDB?

- Everyone uses the same database
  - Saves $$$
  - Transparency builds trust!
  - Other side issue eliminated
  - More time for discussions about lines than data
- Availability to other jurisdictions in CA
  - Saves $$$
  - Cities, counties, special districts – better data!
- Debugging over time by users
  - Error goes down; issues resolved promptly
- Allows public to familiarize themselves with data
  - And can be used between redistrictings
- Avoids building at extreme time pressure
  - Cuts down on data collection issues
  - Leaves time to deal with 'surprises'

Questions?