

Matching Death Master File with Kansas Voter Registration Data

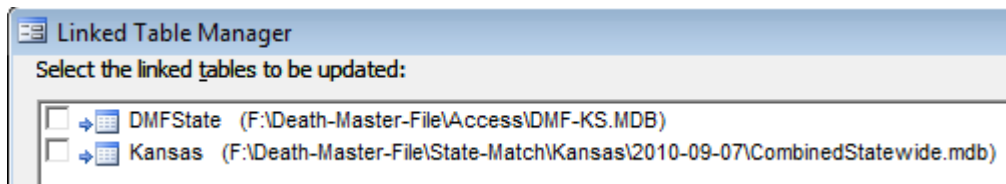
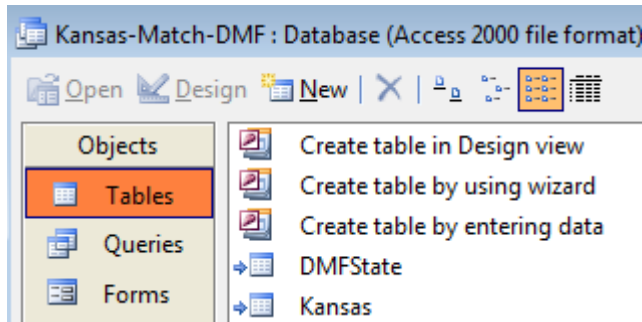
Earl F Glynn
Franklin Center for Government & Public Integrity

Tables.....	1
"Minimal Match"	1
Match Scoring.....	3
Query: 3Make-Score-Matches → Score-Order Table.....	4
Query: 4Make-Score-Zero → Score-Zero Table.....	7

Tables

Access file **Kansas-Match-DMF.mdb**.

In Kansas voter table make sure text_name_last and date_of_birth are indexed "Yes (Duplicates OK)" and that text_registrant_id is the primary key.

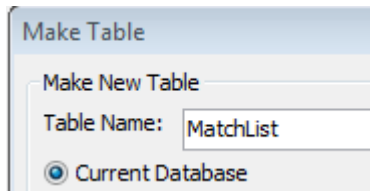
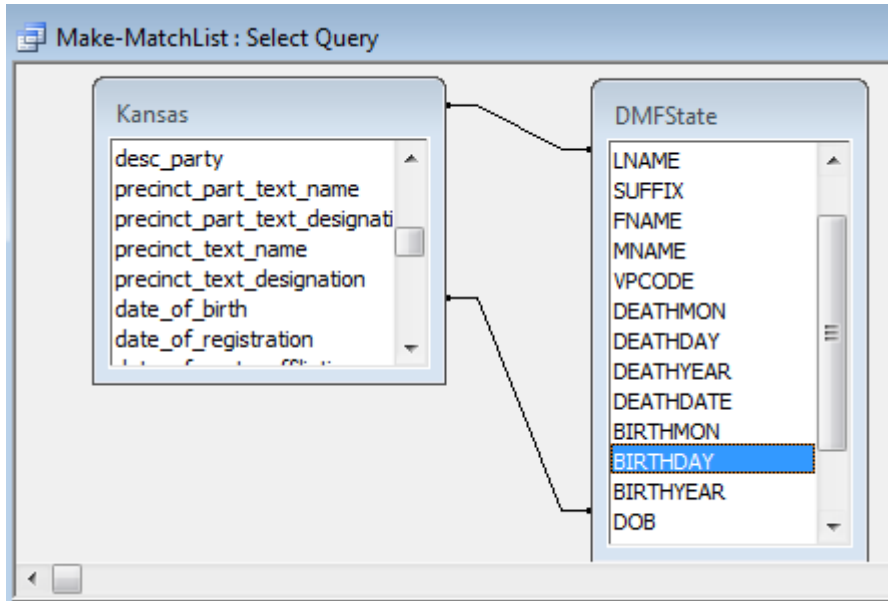


"Minimal Match"

The first filter is to match two files based only on last name and date of birth. This will result in some duplicates and false hits but reduces the size of the file for further consideration..

Join state table with DMFState on last name and date of birth fields in query "**1Make-Matchlist**".

Matching Death Master File with Kansas Voter Registration Data



Add DMFState fields to state table fields:

SSN	VPCODE	date_of_birth	DEATHDATE	STATE	LASTZIP	LUMPZIP
DMFState	DMFState	Kansas	DMFState	DMFState	DMFState	DMFState

...

FNAME	MNAME	LNAME	SUFFIX	text_name_first	text_name_middle	text_name_last	cde_name_suffix
DMFState	DMFState	DMFState	DMFState	Kansas	Kansas	Kansas	Kansas

...

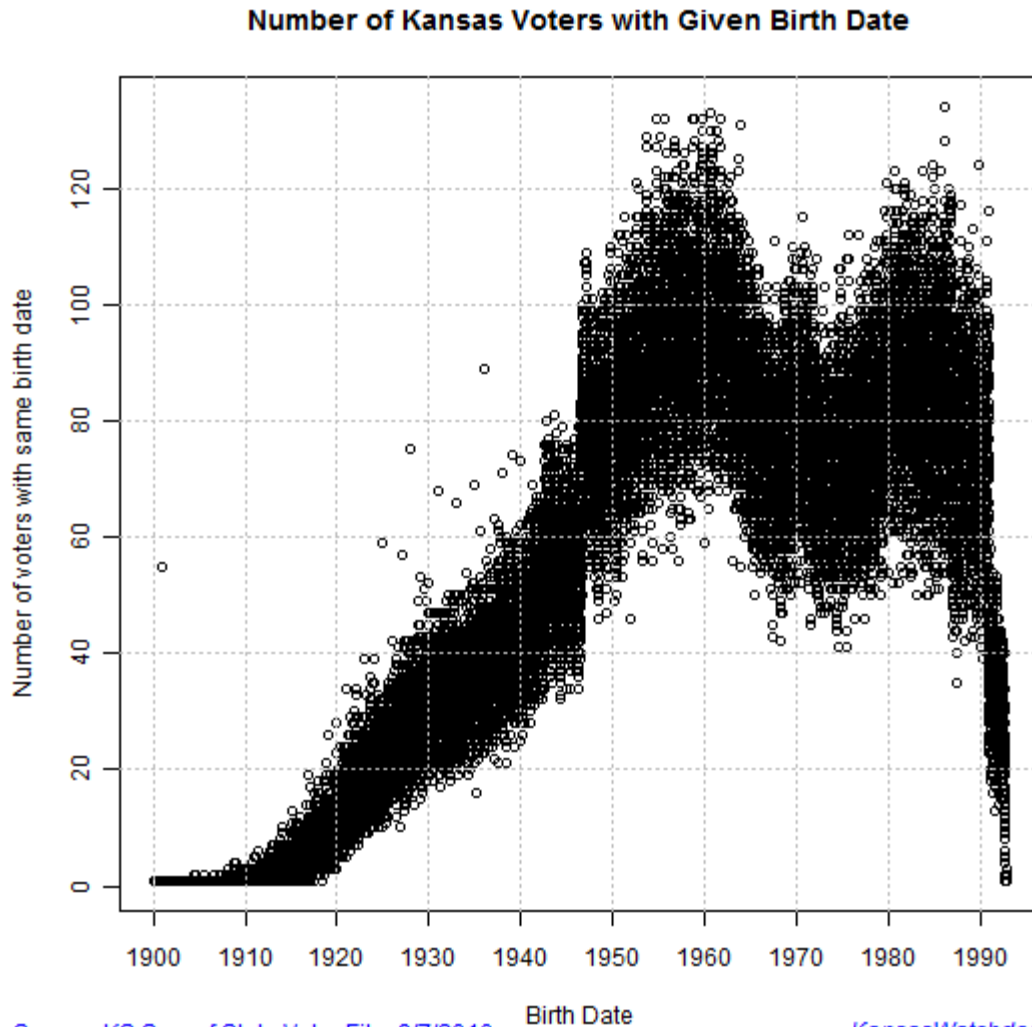
Query took about 20 minutes and creates **MatchList** table.

15,405 records in minimal match

Can result in duplicate SSNs and voter IDs.

A match on the birthdate alone limits the potential pool of voters to consider to no more than about 125. See chart below.

Matching Death Master File with Kansas Voter Registration Data



The 1.7 million Kansas voters share about 30,300 different birthdates for an average of 56.5 voters sharing any given date.

Match Scoring

Query: **2Make-Scoring** → **Scoring** Table

15,405 records

age: Round((([DeathDate]-[date_of_birth])/365.25,1)

FNAMEscore: Iif((Len([text_name_first])>1) And
(Len([FNAME])>1),Iif([text_name_first]=[FNAME],256,0),Iif([text_name_first]=[FNAME],
128,Iif((Left([text_name_first],1)=[FNAME]) Or (Left([FNAME],1)=[text_name_first]),64,0)))

Matching Death Master File with Kansas Voter Registration Data

MixNamescore: Iif(((Len([text_name_first])>1) And ([text_name_first]=[MNAME])) Or ((Len([FNAME])>1) And ([FNAME]=[text_name_middle])),32,0)

Suffixscore: IIF([SUFFIX]=[cde_name_suffix],16,0)

ZIPscore: Iif(([text_res_zip5]=[LASTZIP]) Or ([text_res_zip5]=[LUMPZIP]),8,0)

MNAMEscore: Iif((Len([text_name_middle])>1) And (Len([MNAME])>1),Iif([text_name_middle]=[MNAME],4,0),Iif([text_name_middle]=[MNAME],2,Iif((Left([text_name_middle],1)=[MNAME]) Or (Left([MNAME],1)=[text_name_middle]),1,0))))

TotalScore: [FNAMEscore] +[MixNamescore]+[Suffixscore]+[ZIPscore]+[MNAMEscore]

FNAME match, > 1	FNAME match, = 1	FNAME initials match	FNAME and MNAME mixed	Suffix match	ZIP match	MNAME match, > 1	MNAME match, = 1	MNAME initials match
256	128	64	32	16	8	4	2	1

This binary scoring system allows one to determine what matched and what didn't match. For example, a score of 264 would be 100001000 in binary where the 1 for 256 indicates a first name match, and the 1 for 8 indicates a DMF LASTZIP or LUMLPZIP match with the voter registration ZIP. A larger score isn't necessarily "better". Different scores show different types of matches.

LastVote: Iif(Len([text_election_code_1])=6,CLng(Right([text_election_code_1],4)),Null)

LastYear: Year([DEATHDATE])

VoteOK: Iif((([LastVote]>1900) And ([LastVote] < 2011) And ([LastYear]>1900),Iif([LastVote]<[LastYear],"OK",Iif([LastVote]=[LastYear],"?","Problem"))),"

Query: 3Make-Score-Matches → Score-Order Table

3081 records

LNAME	FNAME	MNAME	SUFFIX	text_name_first	text_name_middle	cde_name_suffix	LASTZIP	LUMPZIP
-------	-------	-------	--------	-----------------	------------------	-----------------	---------	---------

text_res_zip5	text_res_city	county	desc_party	FNAMEscore	MixNamescore	Suffixscore	ZIPscore	MNAMEscore	TotalScore
---------------	---------------	--------	------------	------------	--------------	-------------	----------	------------	------------

TotalScore > 0

LastVote	LastYear	VoteOK	DEATHDATE	age	text_election_code	cde_registrant_stai	SSN	text_registrant_id
----------	----------	--------	-----------	-----	--------------------	---------------------	-----	--------------------

Add index to SSN, text_registrant_id (Duplicates OK)

Multiple dead match voter

20 duplicate SSNs from DMF (Query A1-Make-Duplicate-SSN → DuplicateSSN)

Matching Death Master File with Kansas Voter Registration Data

Resolve manually: Assign Code: Mismatch, Indeterminate, blank

Multiple voters match dead

11 duplicate voter registration numbers (Query A3-Make-Duplicate-ID → DuplicateID) [No overlap with DuplicateSSN]

Resolve manually: Assign Code: Mismatch, Indeterminate, blank

3029 records after resolving duplicate IDs.

Combinations of Matching Scores (Query A5-Make-Score-Order-Counts):

FNAME score	MixName score	Suffix score	ZIP score	MNAME score	Total Score	Raw Count
256	0	0	8	4	268	9
256	0	0	8	2	266	334
256	0	0	8	1	265	147
256	0	0	8	0	264	187
256	0	0	0	4	260	25
256	0	0	0	2	258	502
256	0	0	0	1	257	186
256	0	0	0	0	256	213
128	0	0	8	2	138	8
128	0	0	8	1	137	6
128	0	0	0	2	130	4
128	0	0	0	1	129	6
64	32	0	8	0	104	2
64	0	0	8	4	76	4
64	0	0	8	2	74	48
64	0	0	8	1	73	22
64	0	0	8	0	72	28
64	0	0	0	4	68	2
64	0	0	0	2	66	23
64	0	0	0	1	65	11
64	0	0	0	0	64	64
0	32	0	8	1	41	1
0	32	0	8	0	40	37
0	32	0	0	0	32	31
0	0	16	0	0	16	1
0	0	0	8	4	12	1
0	0	0	8	2	10	36
0	0	0	8	1	9	23
0	0	0	8	0	8	296
0	0	0	0	2	2	532
0	0	0	0	1	1	240
TOTAL						3029

Matching Death Master File with Kansas Voter Registration Data

Each category of scores was reviewed manually looking for mismatches.

TotalScore	Raw Count	Apparent Match	Weak Match	Matches	Mismatch
268	9	9		9	
266	334	334		334	
265	147	147		147	
264	187	175	11	186	1
260	25	25		25	
258	502	502		502	
257	186	186		186	
256	213	171	15	186	27
138	8	8		8	
137	6	6		6	
130	4	4		4	
129	6	6		6	
104	2	2		2	
76	4	4		4	
74	48	48		48	
73	22	22		22	
72	28	27	1	28	
68	2	2		2	
66	23	23		23	
65	11	11		11	
64	64	13		13	51
41	1	1		1	
40	37	36	1	37	
32	31	25		25	6
16	1			0	1
12	1	1		1	
10	36	25	5	30	6
9	23	16	2	18	5
8	296	46	25	71	225
2	532	25	2	27	505
1	240	3	1	4	236
TOTAL	3,029	1,903	63	1,966	1,063

Suggestions for next match:

- Remove periods from middle initials for match with DMF. Would improve MNAMEscore from 1 to 2.
- Assign better score if match is "missing" against some middle name or initial, vs. a mismatch of names or initials.
- Unclear how significant female mismatch on middle name is when first name is not that common since middle names/maiden names can be interchanged for some. E.g, Verify "Mary P Berns" vs "Mary S Berns"
- Consider expanding ZIP match scores. Instead of only one score, have one for exact match, and different proximity scores, perhaps 0 to 25 miles for one score, and 26-100 miles for another. Compute distances based on latitude-longitude values available for each ZIP code.

Matching Death Master File with Kansas Voter Registration Data

But lack of a ZIP match should not exclude from consideration: "Flossie M Glynn" showed an Arkansas ZIP code but she had lived in St. Marys until her death in 2008 -- her SSN indicated a Kansas connection but not the ZIP code.

- If first and second names are both initials only, score higher? All of the score = 138 seem to be good matches.
- Score 104 is a fluke. First and middle names switched, but first initial in voter list matched first initial of other name.
- Scores 65, 66 and 73 are curious lists. Person used initial for first name in voter list, but spelled out name in social security list, or vice versa.
- Score 72 often has first and middle names combined in social security list and separated in voter list, or vice versa.
- Score 10 shows several likely spelling errors in names, or spelling variations, e.g., "Charlotte" vs. "Lottie" and "Gerald" vs. "Jerry."
- Name matches with embedded spaces removed: "Laverne" vs "La Verne"
- Several women in Kansas have voter first name of "Mrs." with middle name of husband's first name.

Verified Dead Registered Voters:

- Former Kansas Governor Wm H Avery listed by Social Security Administration as William H. Avery. Died 11/4/2009 at age 98. Match Score 9 (match on last name, birthdate, middle name and zip code)
- Fr. Heliodore N. Mejak, former pastor Holy Family Catholic Church in Kansas City, KS. Died 12/24/2007. Match Score 266 (match on last name, birthdate, first name, middle initial, zip code)

Query: 4Make-Score-Zero → Score-Zero Table

12,324 records

As above, but TotalScore = 0

Look for data quality problems that may have resulted in zero score.

Matching Death Master File with Kansas Voter Registration Data

Breakdown of Dead Registered Voters

County	Count
Allen	7
Anderson	7
Atchison	42
Barber	16
Barton	11
Bourbon	17
Brown	7
Butler	54
Chase	3
Chautauqua	9
Cherokee	61
Cheyenne	2
Clark	1
Clay	21
Coffey	4
Cowley	55
Crawford	14
Dickinson	6
Doniphan	9
Douglas	31
Edwards	1
Elk	1
Ellis	9
Ellsworth	12
Finney	9
Ford	17
Franklin	30
Geary	17
Gove	2
Graham	11
Grant	7
Greenwood	20
Hamilton	1
Harper	29
Harvey	13
Haskell	8
Hodgeman	3
Jackson	16
Jefferson	58
Jewell	16
Johnson	213

County	Count
Kearny	2
Kingman	23
Kiowa	7
Labette	60
Leavenworth	50
Lincoln	6
Linn	21
Logan	1
Lyon	36
Marion	2
Marshall	6
McPherson	1
Meade	25
Miami	13
Mitchell	9
Montgomery	68
Morton	1
Nemaha	8
Neosho	26
Ness	2
Norton	10
Osage	22
Osborne	1
Ottawa	5
Pawnee	2
Phillips	4
Pratt	11
Reno	96
Republic	2
Rice	4
Riley	12
Rooks	10
Rush	6
Russell	6
Saline	46
Scott	4
Sedgwick	230
Seward	25
Shawnee	55
Sherman	4
Smith	17
Stafford	6
Stanton	1
Sumner	50
Thomas	1
Trego	2
Wallace	1

County	Count
Washington	3
Wilson	6
Woodson	6
Wyandotte	81
TOTAL	1,966

771 are inactive or 39%