An Overview of IPUMS and Linking Historical Data

Hosted by:

CIC and SDC Steering Committee Members

April 23, 2025

IPUMS

What is it? How do we link data over time?

David Van Riper (and Jonathan Schroeder)
vanriper@umn.edu

CIC/SDC Spring Webinar

April 23, 2025

What is IPUMS?

Data provider

- U.S. & international
- Censuses & major surveys
- Integrated across time & space

= Data + websites

+ documentation

+ people

Why use IPUMS?

- Breadth & depth
- Harmonization & enhancements
- Data customization options
- Comprehensive documentation
- User support
- *Free of charge*

Webinar plan

- IPUMS Overview: Projects & Data Types
- Linking geographic units over time

What's in IPUMS?

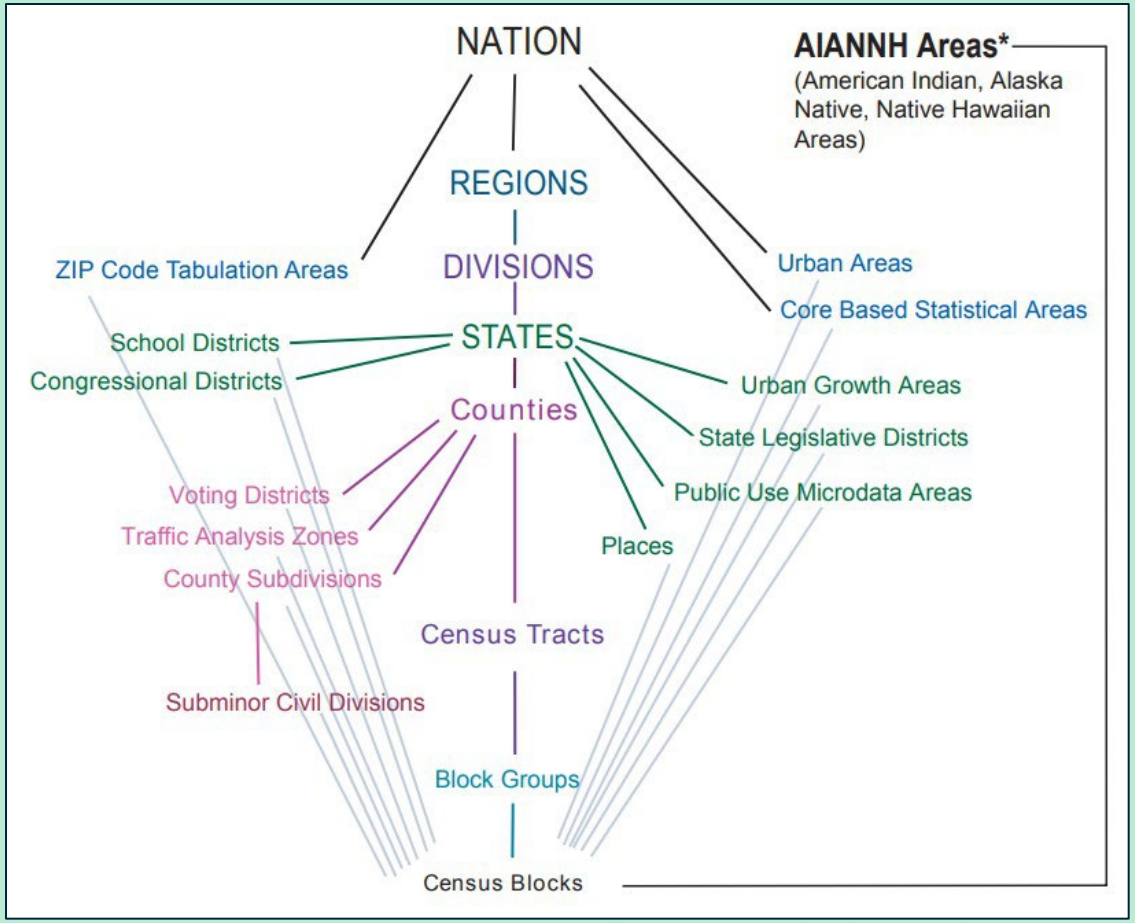
IPUMS USA U.S. Census and American Community Survey microdata from 1850 to the present.	**IPUMS CPS** Current Population Survey microdata including basic monthly surveys and supplements from 1962 to the present.	**IPUMS INTERNATIONAL** World's largest collection of census microdata covering nearly 100 countries, contemporary and historical.
IPUMS GLOBAL HEALTH Health survey data for Africa and Asia, including harmonized data collections for DHS and PMA.	**IPUMS NHGIS** Tabular U.S. Census data and GIS boundary files from 1790 to the present.	**IPUMS IHGIS** Tabular and GIS data from population, housing, and agricultural censuses around the world.
IPUMS TIME USE Historical and contemporary time use data from 1965 to the present.	**IPUMS HEALTH SURVEYS** Historical and contemporary U.S. health survey data from NHIS (1963-present) and MEPS (1996-present).	**IPUMS HIGHER ED** Survey data on the science and engineering workforce in the U.S. from 1993 to the present.

Microdata

Summary data

Observation units of census & survey data

Summary data: *Areas*



https://www.census.gov/geo/reference/hierarchy.html

Microdata: *Respondents*

Households

Persons

+ Flexible classification, modeling & cross-tabulation
- Restricted geographic detail
- Limited sample sizes

Summary data

Table

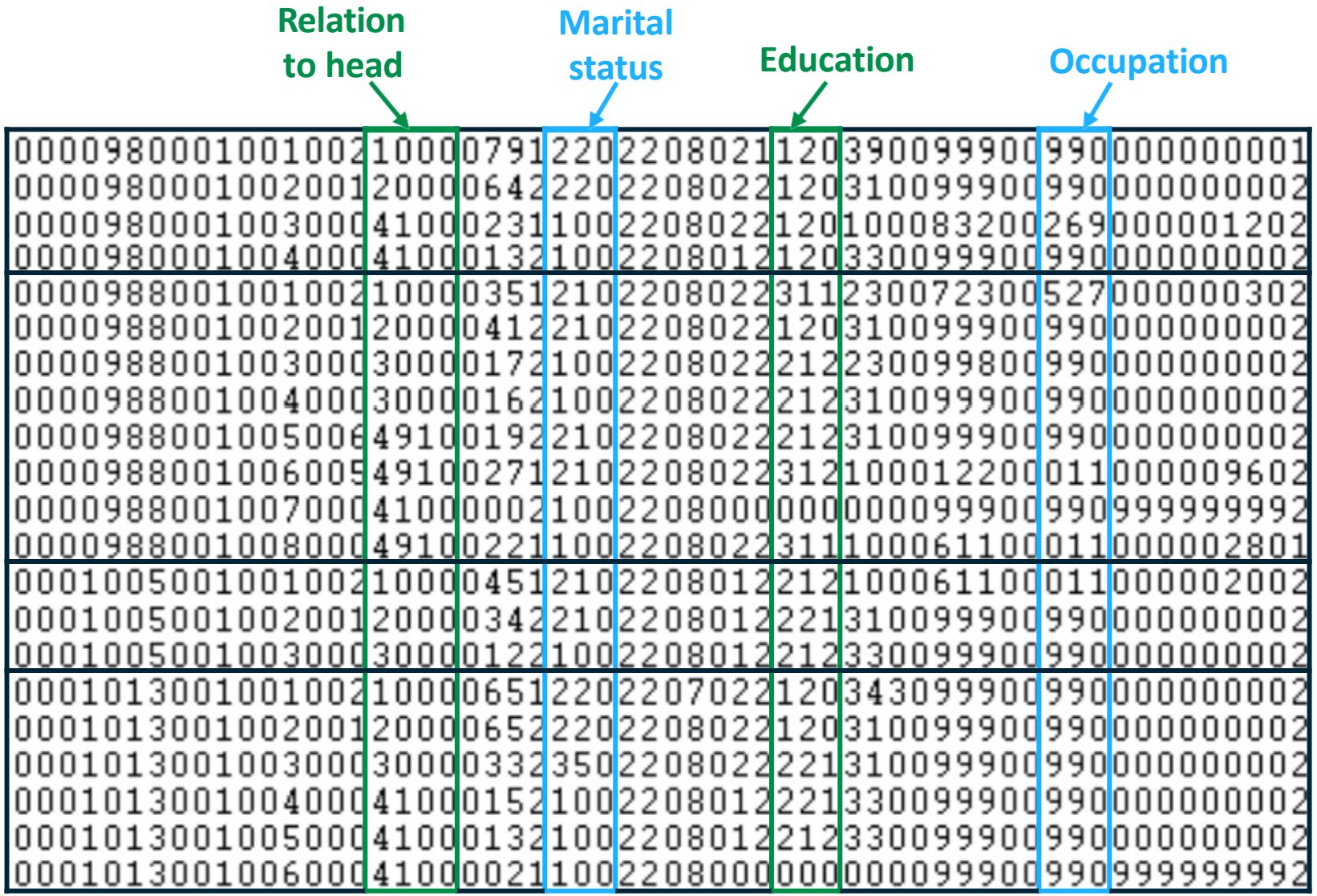
nhgis0214_ds82_1950_tract.csv

GISJOIN	YEAR	STATE	STATEA	COUNTY	COUNTYA	PRETRACTA	TRACTA	POSTTRCTA	AREANAME	BZ8001
G240510000084	1950	Maryland	24	Baltimore Ci	510	<Null>	0008	4	STCTY-24510 TRACT- 0008-4 IN BALTIMORE MD	4124
G240510000085	1950	Maryland	24	Baltimore Ci	510	<Null>	0008	5	STCTY-24510 TRACT- 0008-5 IN BALTIMORE MD	6037
G240510000086	1950	Maryland	24	Baltimore Ci	510	<Null>	0008	6	STCTY-24510 TRACT- 0008-6 IN BALTIMORE MD	6461
G240510000087	1950	Maryland	24	Baltimore Ci	510	<Null>	0008	7	STCTY-24510 TRACT- 0008-7 IN BALTIMORE MD	8430
G240510000091	1950	Maryland	24	Baltimore Ci	510	<Null>	0009	1	STCTY-24510 TRACT- 0009-1 IN BALTIMORE MD	5627
G240510000092	1950	Maryland	24	Baltimore Ci	510	<Null>	0009	2	STCTY-24510 TRACT- 0009-2 IN BALTIMORE MD	2958
G240510000093	1950	Maryland	24	Baltimore Ci	510	<Null>	0009	3	STCTY-24510 TRACT- 0009-3 IN BALTIMORE MD	5268
G240510000094	1950	Maryland	24	Baltimore Ci	510	<Null>	0009	4	STCTY-24510 TRACT- 0009-4 IN BALTIMORE MD	3095
G240510000095	1950	Maryland	24	Baltimore Ci	510	<Null>	0009	5	STCTY-24510 TRACT- 0009-5 IN BALTIMORE MD	2954
G240510000096	1950	Maryland	24	Baltimore Ci	510	<Null>	0009	6	STCTY-24510 TRACT- 0009-6 IN BALTIMORE MD	5205
G240510000097	1950	Maryland	24	Baltimore Ci	510	<Null>	0009	7	STCTY-24510 TRACT- 0009-7 IN BALTIMORE MD	4794
G240510000098	1950	Maryland	24	Baltimore Ci	510	<Null>	0009	8	STCTY-24510 TRACT- 0009-8 IN BALTIMORE MD	7425
G240510000099	1950	Maryland	24	Baltimore Ci	510	<Null>	0009	9	STCTY-24510 TRACT- 0009-9 IN BALTIMORE MD	7768
G240510000101	1950	Maryland	24	Baltimore Ci	510	<Null>	0010	1	STCTY-24510 TRACT- 0010-1 IN BALTIMORE MD	9376
G240510000102	1950	Maryland	24	Baltimore Ci	510	<Null>	0010	2	STCTY-24510 TRACT- 0010-2 IN BALTIMORE MD	7987
G240510000103	1950	Maryland	24	Baltimore Ci	510	<Null>	0010	3	STCTY-24510 TRACT- 0010-3 IN BALTIMORE MD	2441
G240510000111	1950	Maryland	24	Baltimore Ci	510	<Null>	0011	1	STCTY-24510 TRACT- 0011-1 IN BALTIMORE MD	3194
G240510000112	1950	Maryland	24	Baltimore Ci	510	<Null>	0011	2	STCTY-24510 TRACT- 0011-2 IN BALTIMORE MD	6651
G240510000113	1950	Maryland	24	Baltimore Ci	510	<Null>	0011	3	STCTY-24510 TRACT- 0011-3 IN BALTIMORE MD	3396
G240510000114	1950	Maryland	24	Baltimore Ci	510	<Null>	0011	4	STCTY-24510 TRACT- 0011-4 IN BALTIMORE MD	8129
G240510000121	1950	Maryland	24	Baltimore Ci	510	<Null>	0012	1	STCTY-24510 TRACT- 0012-1 IN BALTIMORE MD	3495
G240510000122	1950	Maryland	24	Baltimore Ci	510	<Null>	0012	2	STCTY-24510 TRACT- 0012-2 IN BALTIMORE MD	7746
G240510000123	1950	Maryland	24	Baltimore Ci	510	<Null>	0012	3	STCTY-24510 TRACT- 0012-3 IN BALTIMORE MD	5965
G240510000124	1950	Maryland	24	Baltimore Ci	510	<Null>	0012	4	STCTY-24510 TRACT- 0012-4 IN BALTIMORE MD	5840
G240510000125	1950	Maryland	24	Baltimore Ci	510	<Null>	0012	5	STCTY-24510 TRACT- 0012-5 IN BALTIMORE MD	6813

1 (0 out of *4000 Selected)

nhgis0214_ds82_1950_tract.csv

Microdata


Relation to head
Marital status
Education
Occupation
0000980001001002100007912202208021203900999009900000000001
0000980001002001200006422202208022203100999009900000000002
0000980001003000410002311002208022201000832002690000001202
0000980001004000410001321002208012203300999009900000000002
0000988001001002100003512102208022311230072300527000000302
0000988001002001200004122102208022203100999009900000000002
0000988001003000300001721002208022212300998009900000000002
0000988001004000300001621002208022212310099900990000000002
0000988001005006491001921002208022212310099900990000000002
0000988001006005491002711002208022312100012200011000009602
0000988001007000410000021002208000000000099900990999999992
0000988001008000491002211002208022311100061100011000002801
0001005001001002100004512102208012212100061100011000002002
0001005001002001200003422102208012221310099900990000000002
0001005001003000300001221002208012212330099900990000000002
0001013001001002100006512202207022203430999009900000000002
0001013001002001200006522202208022203100999009900000000002
0001013001003000300003323502208022213100999009900000000002
0001013001004000410001521002208012213300999009900000000002
0001013001005000410001321002208012212330099900990000000002
0001013001006000410000211002208000000000099900990999999992

IPUMS USA

U.S. Census and American Community Survey microdata from 1850 to the present.

IPUMS USA

- U.S. decennial censuses, 1850-2010
- American Community Survey, 2000-2023
- Full count historical census data for 1850-1880, 1900-1950
  - IPUMS MLP links individuals across historical full count censuses

IPUMS CPS

Current Population Survey microdata including basic monthly surveys and supplements from 1962 to the present.

IPUMS CPS

- Current Population Survey
- Labor force survey with topical supplements
- Basic monthly surveys, 1976-present
- Annual Social and Economic Supplement (ASEC), 1962-present
- Short panel: observe individuals up to 8 times over 16-month period

IPUMS TIME USE

U.S. and international time use data
for 1930 to the present.

IPUMS TIME USE

Time Diary Data Structure

START	STOP	ACTIVITY	LOCATION	WITH WHOM
4:00	6:00	Sleep	n/a	n/a
6:00	6:10	Personal Care	n/a	n/a
6:10	6:15	Dressing	Home	Alone
6:15	6:30	Eating Breakfast	Home	Spouse
6:30	6:45	Travel to Train Station	Car	Alone
6:45	6:55	Waiting for the Train	Train Station	Alone
6:55	7:30	Travel to Work	Train	Alone
7:30	10:00	Working	Work	Alone
10:00	11:00	Working	Work	Co-Workers
11:00	11:05	Personal Care	Work	n/a
11:05	12:00	Working	Work	Alone
12:00	12:10	Traveling to Restaurant	Walking	Co-Workers

IPUMS NHGIS

Tabular U.S. Census data
and GIS boundary files from
1790 to the present.

IPUMS NHGIS

- Decennial Census (1790-present) and ACS summary tables and GIS files
- Geographic highlights
  - Counties, 1790-present
  - Census tracts, 1910-present*
  - Census blocks, 1970-present*
- Time series & geographic crosswalks

HARMONIZATION

Microdata Harmonization

- A single, consistent data series from datasets collected in different times and places
- Codes group broadly comparable categories while retaining sample-specific detail
- Denote potential comparability issues

Changes to marital status variable in the National Health Interview Survey

Pre-2004

1 = Married, spouse present

2 = Married, spouse absent

3 = Married, no spouse info

4 = Widowed

5 = Divorced

6 = Separated

7 = Never married

2004-forward

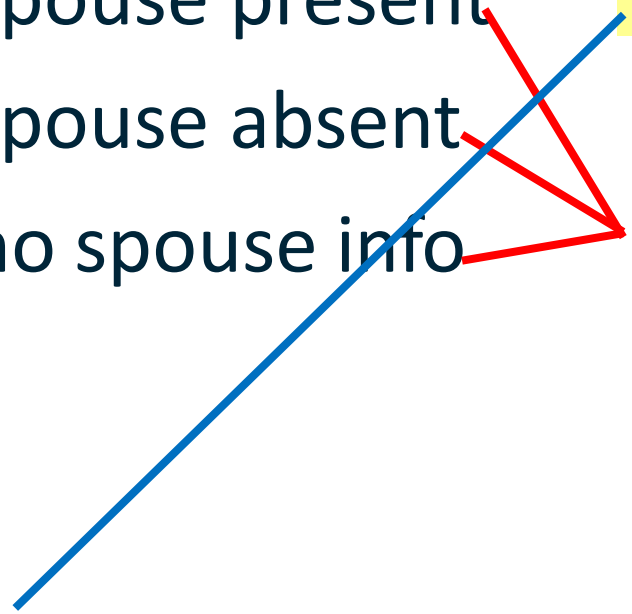
1 = Separated

2 = Divorced

3 = Married

4 = Single/never married

5 = Widowed



Harmonized Marital Status Codes

```
10 = Married
11 = Married, spouse present
12 = Married, spouse absent
13 = Married, no spouse info
20 = Widowed
30 = Divorced
40 = Separated
50 = Never Married
```

The **first digit** captures concepts that are consistent across all samples

The **second digit** contains sample-specific detail

IPUMS NHGIS

KEY NHGIS RESOURCES

Time Series & Crosswalks

TIME SERIES TABLES

WHAT ARE TIME SERIES TABLES?

Nominal integration

nhgis0215_ts_state.csv [Read-Only] - Microsoft Excel non-commercial use

File Home Insert Page Layout Formulas Data Review View Add-Ins

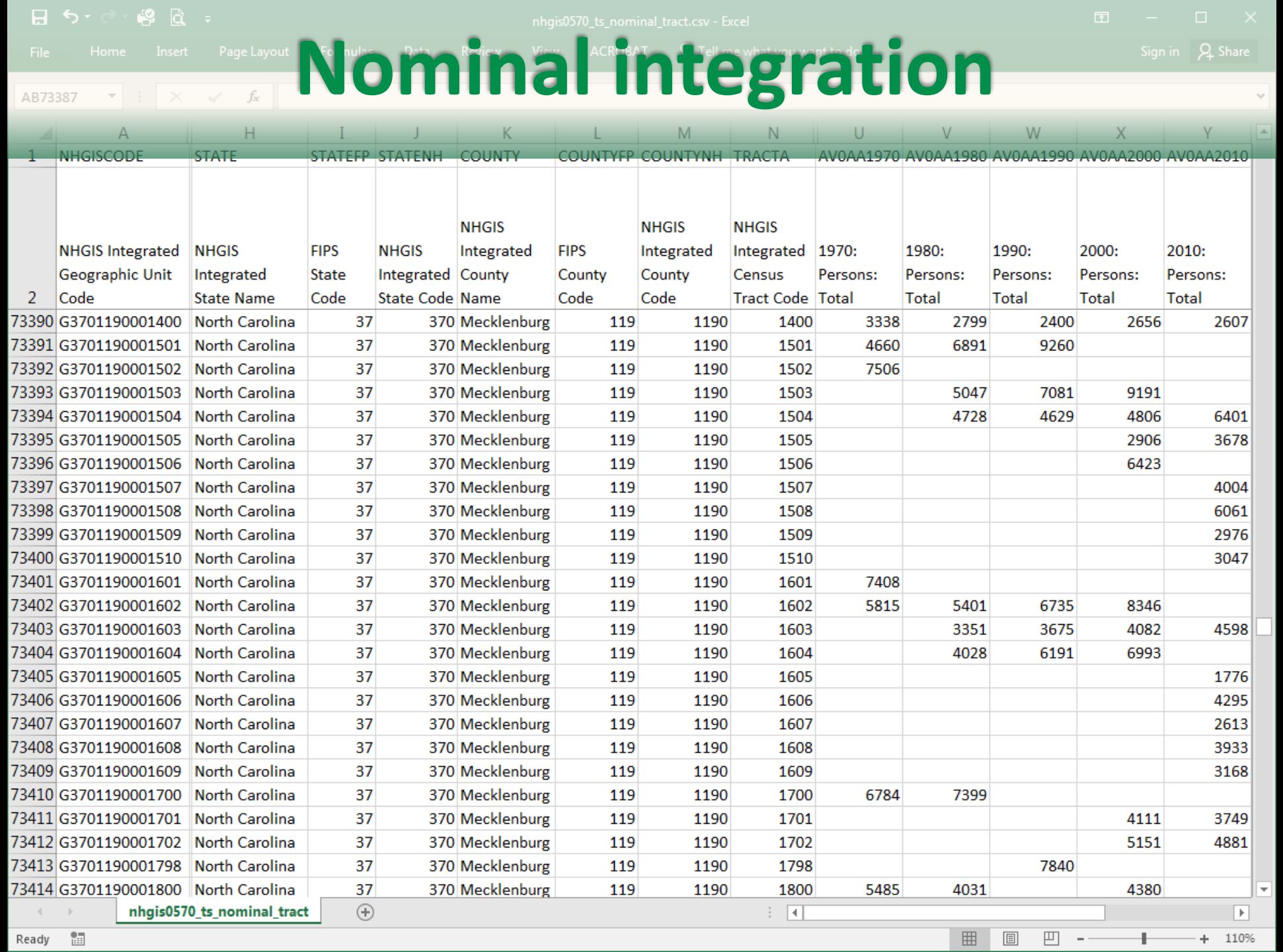
A1

	A	B	C	D	E	F	G	H	I	J
1	GISJOIN	STATE	STATEA	BL1AA1970	BL1AA1980	BL1AA1990	BL1AA2000	BL1AA125	BL1AA125M	BL1AB1970
2	GIS Join Match Code	State Name	State Code	1970: Persons: Male ~ Never married	1980: Persons: Male ~ Never married	1990: Persons: Male ~ Never married	2000: Persons: Male ~ Never married	2008-2012: Persons: Male ~ Never married	Margin of error: 2008-2012: Persons: Male ~ Never married	1970: Persons: Male ~ Now married
3	G010	Alabama	10	324572	375515	406140	444450	584355	4261	792229
4	G020	Alaska	20	38428	53423	67827	78955	105333	1532	69046
5	G040	Arizona	40	170463	286183	409976	581715	876661	5440	410438
6	G050	Arkansas	50	170171	200864	209851	245874	330745	3312	464000
7	G060	California	60	2102965	2938148	4034185	4343790	5778554	18439	4611973
8	G080	Colorado	80	230115	341985	373157	516816	678866	4145	510105
9	G090	Connecticut	90	306085	369471	412921	385888	496941	3768	706908
10	G100	Delaware	100	52624	67052	76997	87154	122851	1650	126479
11	G110	District Of C	110	99908	112013	118273	112516	142162	1558	142605
12	G120	Florida	120	596651	950049	1336469	1702648	2553294	14356	1687636
13	G130	Georgia	130	446917	564696	716125	940487	1299191	7312	1050444
14	G150	Hawaii	150	104625	137448	152188	166715	205985	2060	169920
15	G160	Idaho	160	69395	86794	91380	128613	172169	2427	169553
16	G170	Illinois	170	1115952	1318480	1395167	1495795	1838840	9371	2542632
17	G180	Indiana	180	477592	541264	569459	643222	814989	5228	1222896
18	G190	Iowa	190	271540	301518	281081	318000	375129	2818	665781
19	G200	Kansas	200	223168	241362	244866	281338	345753	3010	542632
20	G210	Kentucky	210	318856	358353	361278	401465	518929	4039	757239
21	G220	Louisiana	220	370838	450684	464285	514633	633702	4170	793188
22	G230	Maine	230	97253	116576	126930	132473	163254	1650	227135
23	G240	Maryland	240	392646	501319	587368	610852	814952	4997	906860
24	G250	Massachuse	250	622162	743825	832633	817954	981462	5728	1243732
25	G260	Michigan	260	872687	1036241	1078995	1159693	1352844	6714	2021198
26	G270	Minnesota	270	403762	484050	504483	591866	716441	4620	849440
27	G280	Mississippi	280	222569	252992	275422	319922	388559	3394	480663

nhgis0215_ts_state

Ready 100%

Nominal integration



nhgis0570_ts_nominal_tract.csv - Excel

	A	H	I	J	K	L	M	N	U	V	W	X	Y
1	NHGISCODE	STATE	STATEFP	STATENH	COUNTY	COUNTYFP	COUNTYNH	TRACTA	AV0AA1970	AV0AA1980	AV0AA1990	AV0AA2000	AV0AA2010
2	NHGIS Integrated Geographic Unit Code	NHGIS Integrated State Name	FIPS State Code	NHGIS Integrated State Code	NHGIS Integrated County Name	FIPS County Code	NHGIS Integrated County Code	NHGIS Integrated Census Tract Code	1970: Persons: Total	1980: Persons: Total	1990: Persons: Total	2000: Persons: Total	2010: Persons: Total
73390	G3701190001400	North Carolina	37	370	Mecklenburg	119	1190	1400	3338	2799	2400	2656	2607
73391	G3701190001501	North Carolina	37	370	Mecklenburg	119	1190	1501	4660	6891	9260		
73392	G3701190001502	North Carolina	37	370	Mecklenburg	119	1190	1502	7506				
73393	G3701190001503	North Carolina	37	370	Mecklenburg	119	1190	1503		5047	7081	9191	
73394	G3701190001504	North Carolina	37	370	Mecklenburg	119	1190	1504		4728	4629	4806	6401
73395	G3701190001505	North Carolina	37	370	Mecklenburg	119	1190	1505				2906	3678
73396	G3701190001506	North Carolina	37	370	Mecklenburg	119	1190	1506				6423	
73397	G3701190001507	North Carolina	37	370	Mecklenburg	119	1190	1507					4004
73398	G3701190001508	North Carolina	37	370	Mecklenburg	119	1190	1508					6061
73399	G3701190001509	North Carolina	37	370	Mecklenburg	119	1190	1509					2976
73400	G3701190001510	North Carolina	37	370	Mecklenburg	119	1190	1510					3047
73401	G3701190001601	North Carolina	37	370	Mecklenburg	119	1190	1601	7408				
73402	G3701190001602	North Carolina	37	370	Mecklenburg	119	1190	1602	5815	5401	6735	8346	
73403	G3701190001603	North Carolina	37	370	Mecklenburg	119	1190	1603		3351	3675	4082	4598
73404	G3701190001604	North Carolina	37	370	Mecklenburg	119	1190	1604		4028	6191	6993	
73405	G3701190001605	North Carolina	37	370	Mecklenburg	119	1190	1605					1776
73406	G3701190001606	North Carolina	37	370	Mecklenburg	119	1190	1606					4295
73407	G3701190001607	North Carolina	37	370	Mecklenburg	119	1190	1607					2613
73408	G3701190001608	North Carolina	37	370	Mecklenburg	119	1190	1608					3933
73409	G3701190001609	North Carolina	37	370	Mecklenburg	119	1190	1609					3168
73410	G3701190001700	North Carolina	37	370	Mecklenburg	119	1190	1700	6784	7399			
73411	G3701190001701	North Carolina	37	370	Mecklenburg	119	1190	1701				4111	3749
73412	G3701190001702	North Carolina	37	370	Mecklenburg	119	1190	1702				5151	4881
73413	G3701190001798	North Carolina	37	370	Mecklenburg	119	1190	1798			7840		
73414	G3701190001800	North Carolina	37	370	Mecklenburg	119	1190	1800	5485	4031		4380	

nhgis0570_ts_nominal_tract

Ready

110%

Geographic standardization

nhgis0570_ts_geog2010_tract.csv - Excel

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	GISJOIN	GEOGYEAR	STATE	STATEA	COUNTY	COUNTYA	TRACTA	CL8AA199	CL8AA199	CL8AA199	CL8AA200	CL8AA200	CL8AA200	CL8AA2010
2	GIS Join Match Code	Geography Year	State Name	State Code	County Name	County Code	Census Tract Code	1990: Persons: Total	Lower bound: 1990: Persons: Total	Upper bound: 1990: Persons: Total	2000: Persons: Total	Lower bound: 2000: Persons: Total	Upper bound: 2000: Persons: Total	2010: Persons: Total
49315	G3701190001400	2010	North Carolina	37	Mecklenburg Co	119	1400	2389.68	2366	2391	2652.31	2640	2653	2607
49316	G3701190001504	2010	North Carolina	37	Mecklenburg Co	119	1504	4629	4629	4629	4806	4806	4806	6401
49317	G3701190001505	2010	North Carolina	37	Mecklenburg Co	119	1505	2901	2901	2901	2906	2906	2906	3678
49318	G3701190001507	2010	North Carolina	37	Mecklenburg Co	119	1507	3419	3419	3419	4791	4791	4791	4004
49319	G3701190001508	2010	North Carolina	37	Mecklenburg Co	119	1508	3662	3662	3662	4400	4400	4400	6061
49320	G3701190001509	2010	North Carolina	37	Mecklenburg Co	119	1509	3213	3213	3213	3229	3229	3229	2976
49321	G3701190001510	2010	North Carolina	37	Mecklenburg Co	119	1510	3146.1	3146	3211	3194.36	3194	3367	3047
49322	G3701190001603	2010	North Carolina	37	Mecklenburg Co	119	1603	3660	3660	3660	4065	4065	4065	4598
49323	G3701190001605	2010	North Carolina	37	Mecklenburg Co	119	1605	1389	1389	1389	1978	1978	1978	1776
49324	G3701190001606	2010	North Carolina	37	Mecklenburg Co	119	1606	3822	3822	3822	4154	4154	4154	4295
49325	G3701190001607	2010	North Carolina	37	Mecklenburg Co	119	1607	1524	1524	1524	2214	2214	2214	2613
49326	G3701190001608	2010	North Carolina	37	Mecklenburg Co	119	1608	3872	3872	3872	4321	4321	4321	3933
49327	G3701190001609	2010	North Carolina	37	Mecklenburg Co	119	1609	2251.16	1687	2334	2445.55	2137	2635	3168
49328	G3701190001701	2010	North Carolina	37	Mecklenburg Co	119	1701	3303	3303	3303	4111	4111	4111	3749
49329	G3701190001702	2010	North Carolina	37	Mecklenburg Co	119	1702	4537	4537	4537	5151	5151	5151	4881
49330	G3701190001801	2010	North Carolina	37	Mecklenburg Co	119	1801	1813.95	1349	1814	1706.96	1702	1707	1249
49331	G3701190001802	2010	North Carolina	37	Mecklenburg Co	119	1802	2775	2775	2775	2673	2673	2673	2892
49332	G3701190001910	2010	North Carolina	37	Mecklenburg Co	119	1910	4277	4277	4277	4614	4614	4614	5117
49333	G3701190001911	2010	North Carolina	37	Mecklenburg Co	119	1911	4613	4613	4613	5041	5041	5041	5441
49334	G3701190001912	2010	North Carolina	37	Mecklenburg Co	119	1912	5096.84	5014	5661	5506.45	5317	5815	6565
49335	G3701190001914	2010	North Carolina	37	Mecklenburg Co	119	1914	3295	3295	3295	3965	3965	3965	3485
49336	G3701190001915	2010	North Carolina	37	Mecklenburg Co	119	1915	4403	4403	4403	5460	5460	5460	6293
49337	G3701190001916	2010	North Carolina	37	Mecklenburg Co	119	1916	3003.8	2465	3050	3005.23	2508	3009	3725
49338	G3701190001917	2010	North Carolina	37	Mecklenburg Co	119	1917	3050.2	3004	3589	3294.77	3291	3792	3567
49339	G3701190001918	2010	North Carolina	37	Mecklenburg Co	119	1918	2842.4	698	3700	3129	3129	3129	3527
49340	G3701190001919	2010	North Carolina	37	Mecklenburg Co	119	1919	3243.6	2386	5388	3837	3837	3837	3379

nhgis0570_ts_geog2010_tract

Ready

Nominally integrated time series

- ~5,700 time series in 273 tables

  Including: marital status, nativity & place of birth, education, employment, means of transportation, travel time, income, poverty

- 8 geographic levels:

  nation, regions, divisions, states, counties, tracts, county subdivisions, places

- Years mainly in 1970-2023 range
  - *Total Population* back to 1790 & *Persons by Sex* back to 1820
  - "Long-form" tables use ACS 2006-10 through 2019-23

Geographically standardized time series

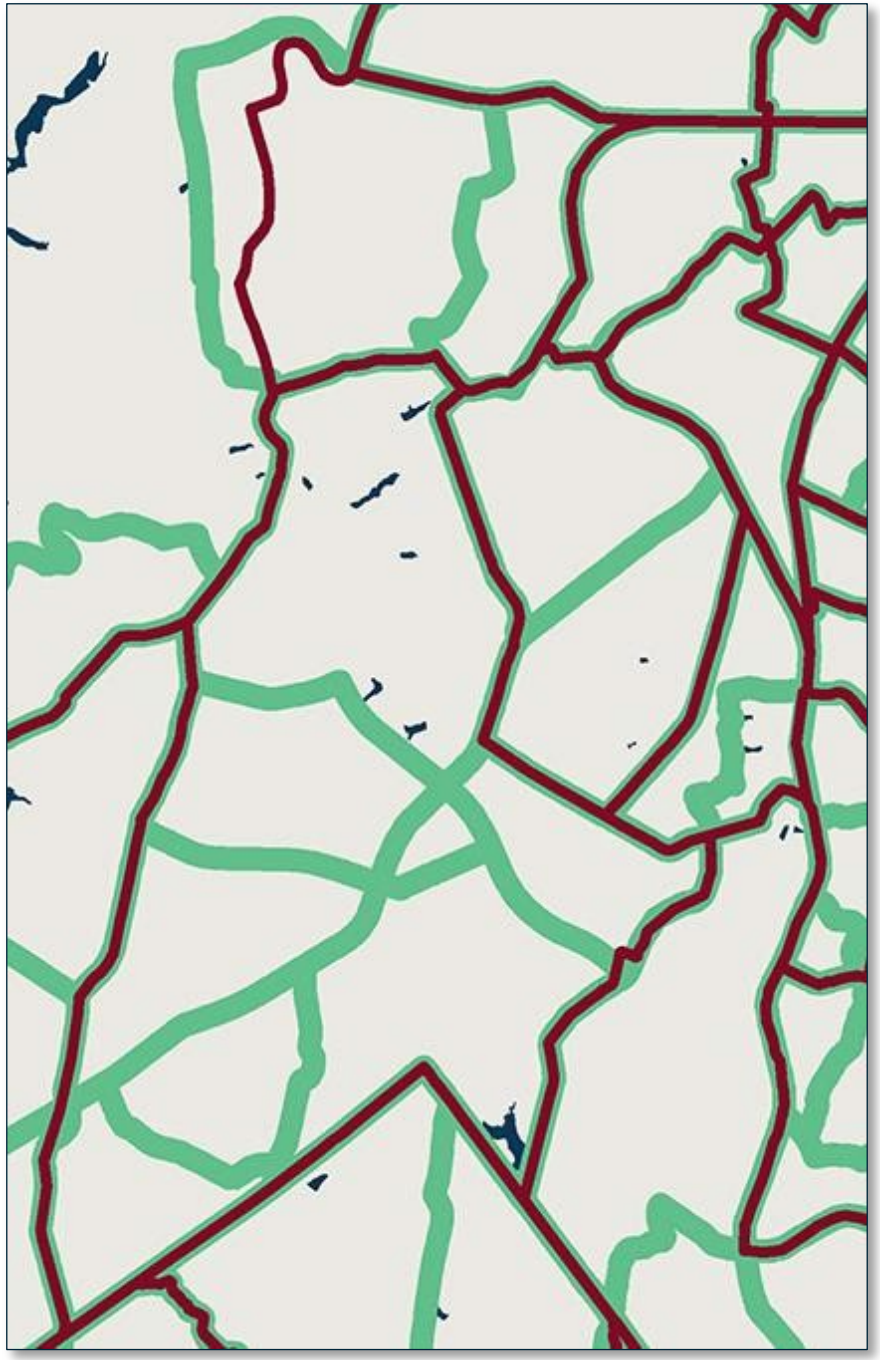
- 1990–2020 data for 2010 census units
- 10 geographic levels: states, counties, tracts, block groups, county subdivisions, places, congressional districts, CBSAs, urban areas, ZCTAs
- ~1,600 time series in 111 tables
- "Short-form" counts only
  - Race, ethnicity, age, sex, household size & relationships, housing occupancy & tenure
  - *Not* income, education, employment, …

I want to analyze tract data across time…

but the tracts don't match…

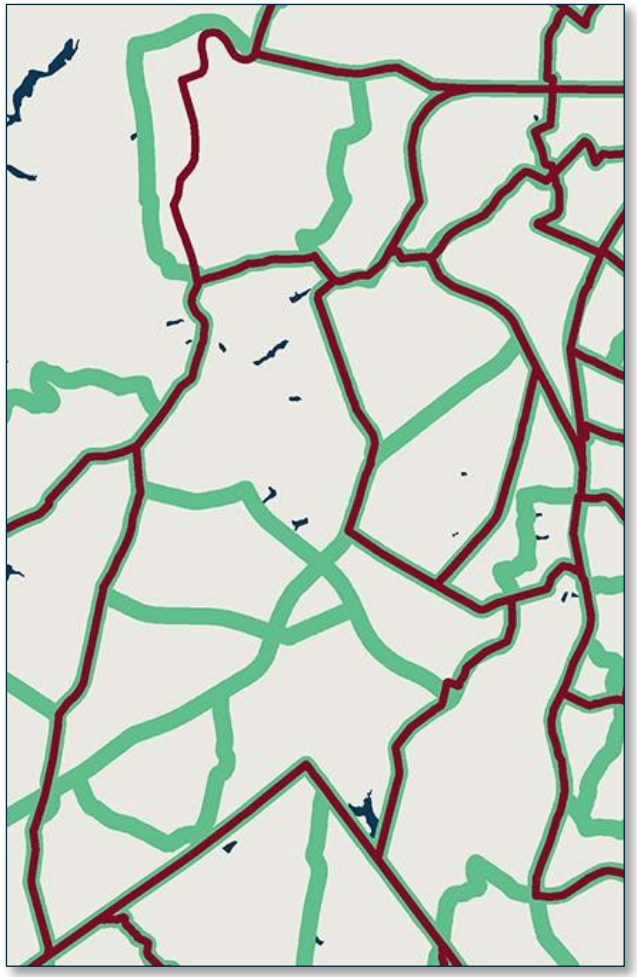
~~*so I need a tract crosswalk*~~

I need data from different years for **one year's tracts**

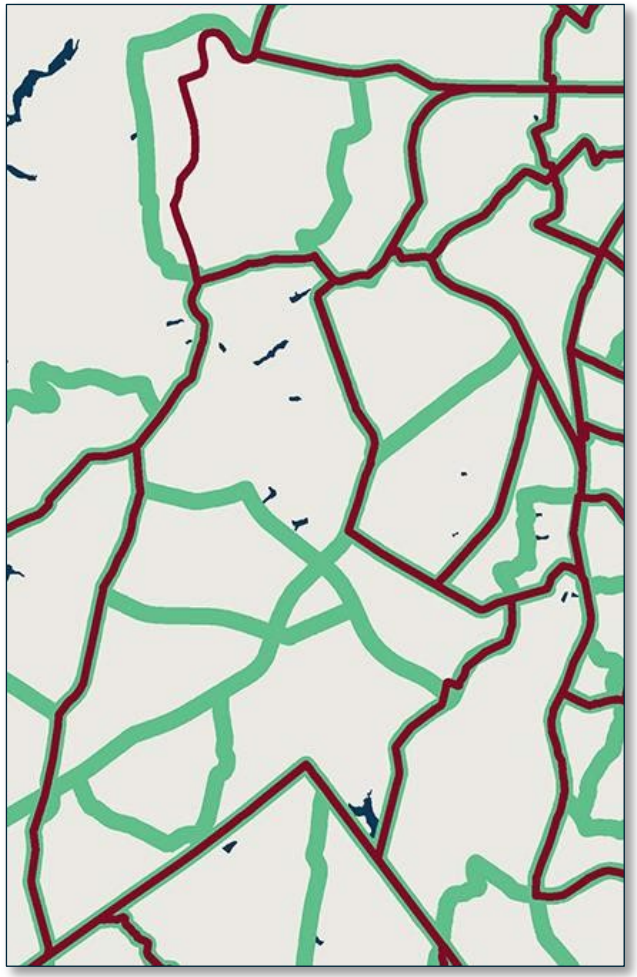


1990 tract

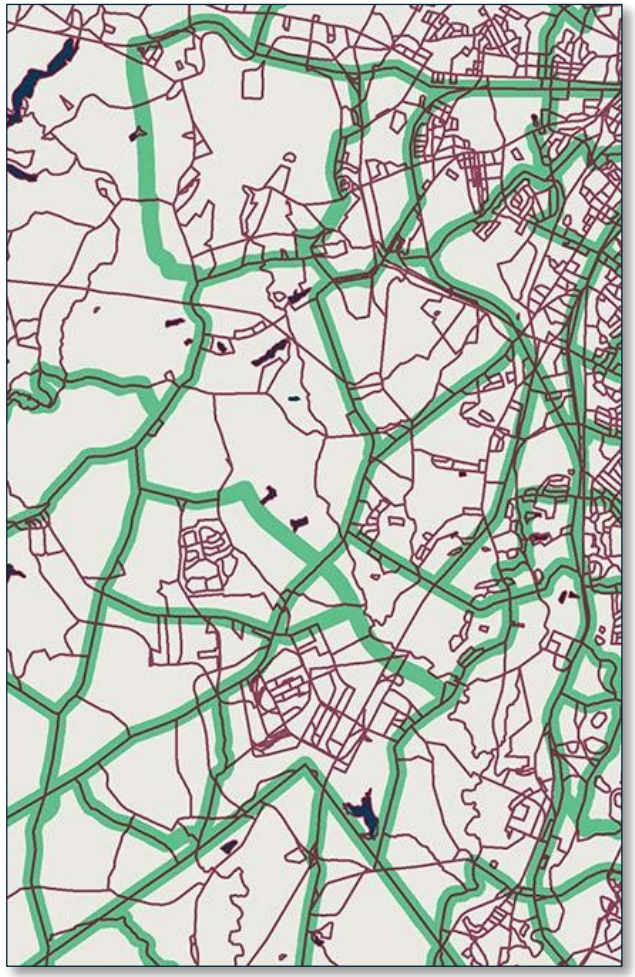
2010 tract



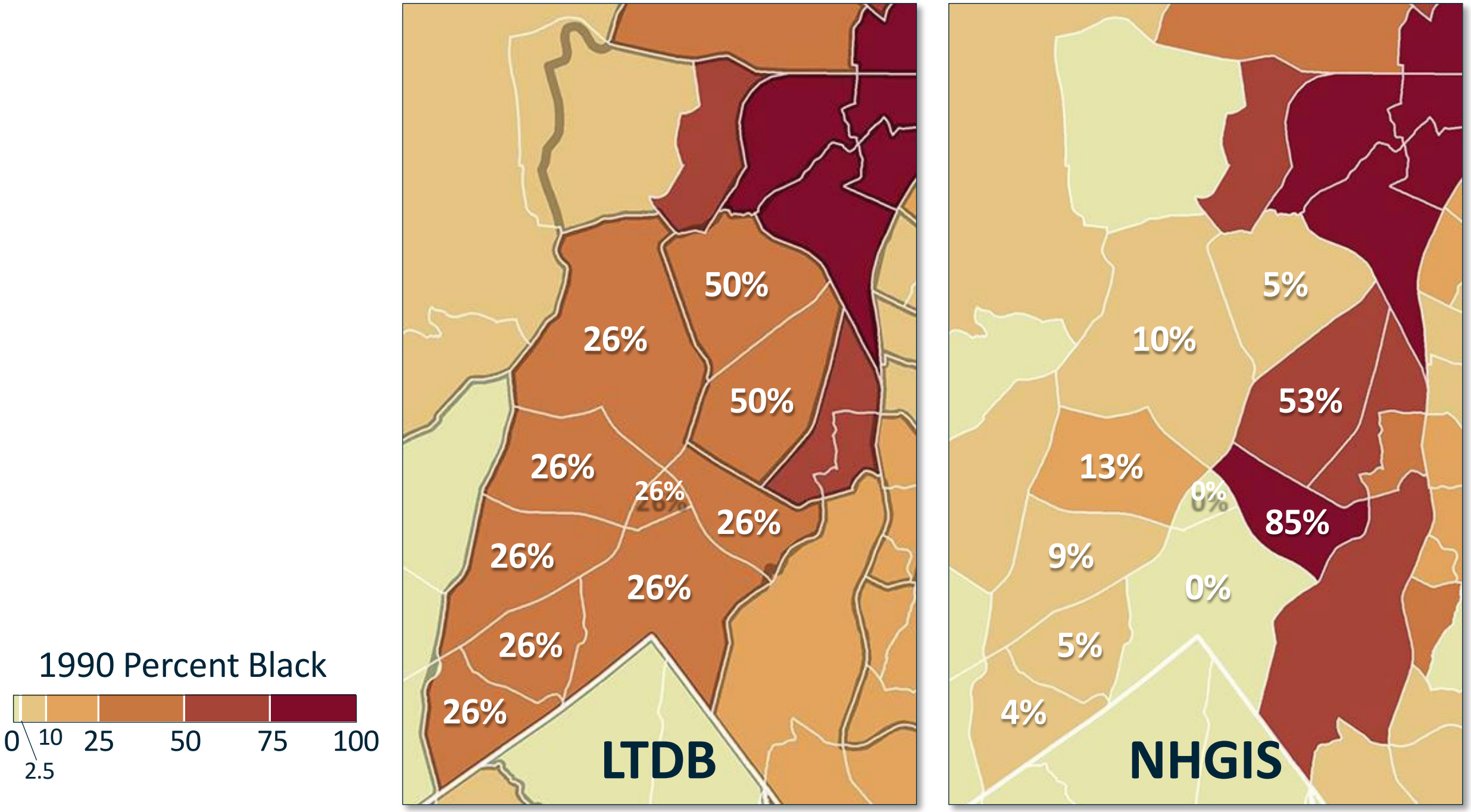
tracts → tracts



tracts → tracts



blocks → tracts


1990 Percent Black
0
2.5
10
25
50
75
100
50%
26%
50%
26%
26%
26%
26%
26%
26%
26%
LTDB
5%
10%
53%
13%
0%
0%
85%
9%
0%
5%
4%
NHGIS

Time Series Tables | IPUM

Secure | https://www.nhgis.org/documentation/time-series#standardization

Documentation at nhgis.org

2. ***Geographically standardized tables*** provide data from multiple times for a single census's geographic units. At this time, NHGIS's standardized time series tables provide 1990, 2000, and 2010 data for *2010 census units*.

To allocate 1990 and 2000 summary data to 2010 census units, NHGIS reaggregates data from the smallest source units for which 1990 and 2000 data are available: census blocks. Where a source block intersects multiple 2010 units, NHGIS applies interpolation to estimate how the source block's characteristics are distributed among the 2010 units.

Detailed methodology:

- 2000 Block Data Standardized to 2010 Geography
- 1990 Block Data Standardized to 2010 Geography

For each standardized statistic, NHGIS also supplies lower and upper bounds based on the spatial relationship between the source units and standard units. For example, if there are three 2000 census blocks that straddle a 2010 census unit's boundary, then it is possible that either all or none of the three blocks' 2000 residents were located in the 2010 unit. The upper bound assumes that all residents and housing units in straddling blocks were located in the 2010 unit, and the lower bound assumes that none were. Bounds for 1990 estimates also take account of additional uncertainty due to accuracy improvements in Census Bureau's geographic data files, which make it impossible to determine exact spatial relationships between 1990 and 2010 units.

NHGIS has not yet implemented standardization for non-count statistics such as medians and quotients. Therefore, currently available standardized tables supply only count statistics.

NHGIS delivers standardized statistics with two decimal digits of precision in order to reduce the size of rounding errors when users sum estimates. Rounding errors may still

Models

- 2000 – 2010
  - Hybrid Binary Dasymetric – Target-density Weighting
    - (2001 NLCD >= 5% impervious surface + 300 ft buffer around residential roads) – water area
    - Allocate 2000 data based on 2010 densities
- 1990 – 2010
  - Cascading density weighting (CDW)

Density weighting across longer periods

- E.g., from 1990 blocks to 2010 blocks…
  - Cascading density weighting
    1. Use 2010 densities to guide *2000* block interpolation
    2. Use estimated 2000 densities to guide 1990 block data

Correspondence info

- 2020 TIGER/Line → 2010 & 2020 units
- 2010 TIGER/Line → 2000 & 2010 units
- 2000 TIGER/Line → 1990 & 2000 units

Accuracy improvements

NHGIS Crosswalks by Geographic Levels and Years

Source Zones	Target Zones	1990 to 2010	2000 to 2010	2010 to 2020	2020 to 2010
Blocks	Blocks, Block Groups, Census Tracts, Counties	X	X	X	X
Block Group Parts	Block Groups, Census Tracts, Counties	X	X		
Block Groups	Block Groups, Census Tracts, Counties			X	X
Census Tracts	Census Tracts, Counties	X	X	X	X

1990 block to 2010 tract example

1990 blk	2010 tract	Weight
270530220108	27053021900	0.1244121774
270530220108	27053022000	0.8755878226
270530220111	27053021900	1.0

1990 block group part to 2010 tract

G27000300148601486050206069999999512016

2010 tract	Wt_pop	Wt_adult	Wt_fam	Wt_hh
27003050218	.3654463488	.180333757	.1972739688	.1869978055
27003050219	.6299307182	819666243	.8027260312	.8130021945
27003050505	0.0	0.0	0.0	0.0

Acknowledgments

Eunice Kennedy Shriver National Institute of Child Health & Human Development
[NIH-R01HD043392,
NIH-R01HD057929,
NIH-R24HD041023]

National Science Foundation
[SES-1324875]

Let's keep in touch!

- User forum
  https://forum.ipums.org/
- Email
  ipums@umn.edu
  nhgis@umn.edu

USING FULL-COUNT CENSUS DATA AND THE MULTIGENERATIONAL LONGITUDINAL PANEL (MLP) TO STUDY URBAN CHANGE

MULTIGENERATIONAL LONGITUDINAL PANEL

- The Multigenerational Longitudinal Panel is a product of IPUMS, in collaboration with *Ancestry.com* and *FamilySearch.org*
- Links individual-level records between censuses through a consistent unique identifier – currently includes data from 1850-1950 (prelim)
- Links can be chained across multiple pairs of census years to yield multi-decade longitudinal data
- High conditional accuracy (96%-99%) matching against records from other sources, although selection into both datasets is problematic
- Full-count census data is also matched to Social Security application and death records

Jonas Helgertz, Joseph Price, Jacob Wellington, Kelly J Thompson, Steven Ruggles & Catherine A. Fitch (2022) A new strategy for linking U.S. historical censuses: A case study for the IPUMS multigenerational longitudinal panel, *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, 55:1, 12-2[illegible]

DOI: 10.1080/01615440.2021.1985027

IPUMS

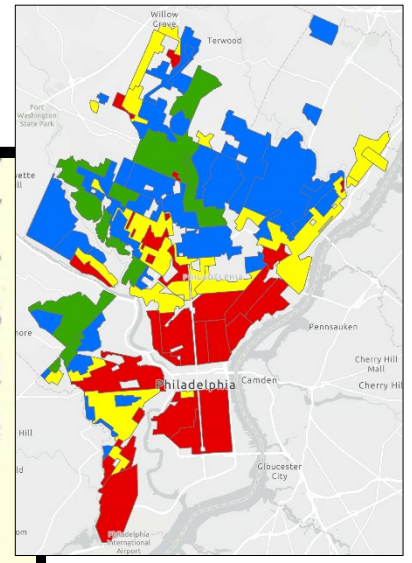
NUMBER OF LINKAGES

IPUMS

10-Year Linkages		
Census Years	**Number of Links**	**Approximate % Start Year**
1850-1860	6,659,383	35%
1860-1870	8,437,331	31%
1870-1880	13,148,961	34%
1900-1910	30,601,820	40%
1910-1920	37,940,159	41%
1920-1930	45,335,479	43%
1930-1940	54,136,642	44%
1940-1950	54,927,229	42%

Number of 20- and 30-year linkages is smaller

OUR PROJECT



below those to which most Americans are accustomed. While the ranking given below may be scientifically wrong from the standpoint of inherent racial characteristics, it registers an opinion or prejudice that is reflected in land values; it is the ranking of races and nationalities with respect to their beneficial effect upon land values. Those having the most favorable come first in the list and those exerting the most detrimental effect appear last.[32]

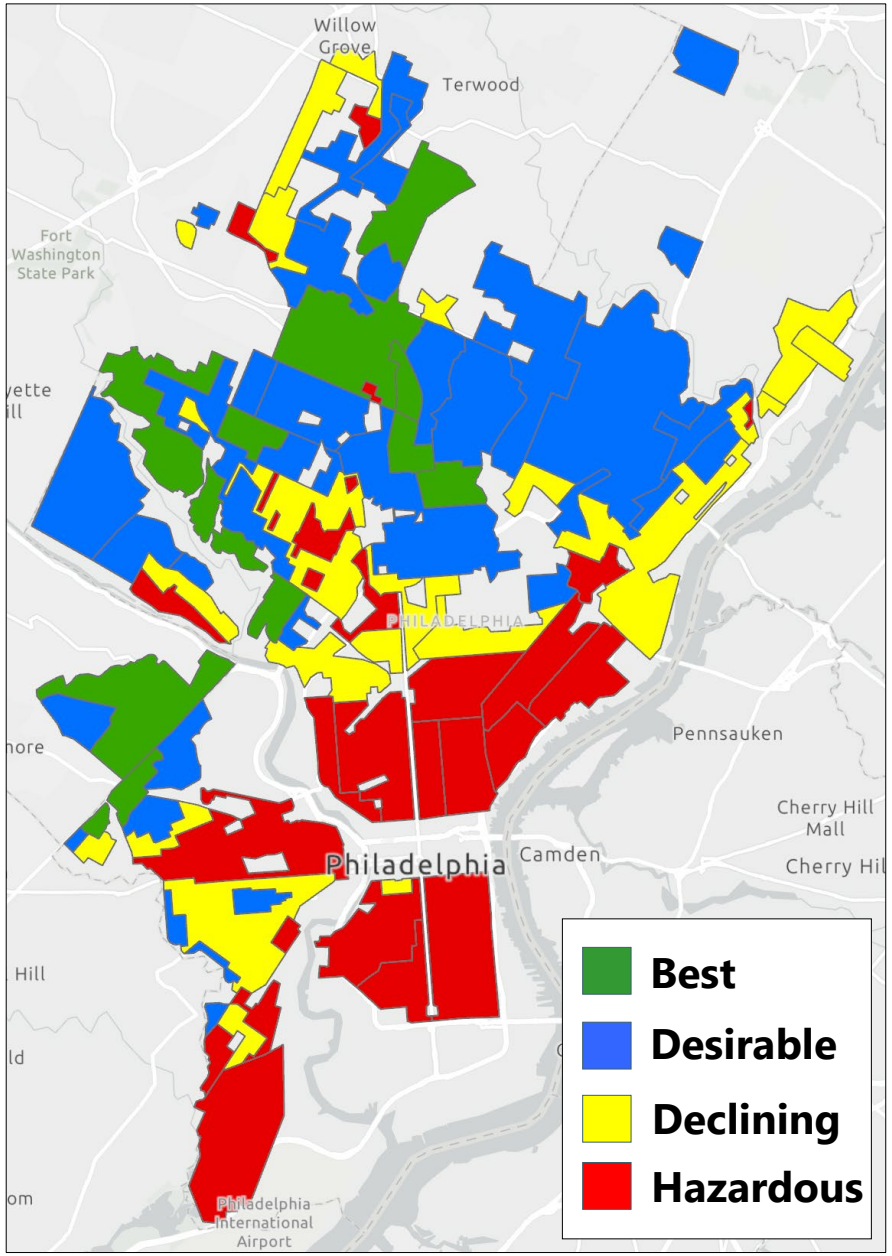
1. English, Germans, Scotch, Irish, Scandinavians
2. North Italians
3. Bohemians or Czechoslovakians
4. Poles
5. Lithuanians
6. Greeks
7. Russian Jews of the lower class
8. South Italians
9. Negroes
10. Mexicans

Hoyt, Homer (1933) *One hundred years of land values in Chicago*. University of Chicago.

The expansion of these racial and national groups has perhaps had a greater effect in promoting a rise in the values of land in the outer areas of the city than it has had upon the older areas, for it has forced or stimulated the old American stock to seek new neighborhoods and has caused them to migrate from their old homes. Even these races themselves have pushed farther from the center of the city, leaving behind an area to be occupied by the least desirable members of their own race. As new immigration from Europe has almost ceased, and as

OUR PROJECT

- Use full-count data to compare HOLC neighborhood population appraisals with counts from the census
- Assess the neighborhood conditions – as measured by HOLC ratings – of immigrants in 1930
- Compare immigrant residential patterns based on country of origin
- Use matched immigrants from the MLP to evaluate the neighborhood transitions of immigrants who moved between 1930 and 1940



HOLC ESTIMATES VERSUS 1940 CENSUS

PRELIMINARY

	Grade A Best	Grade B Desirable	Grade C Declining	Grade D Hazardous
Mean % Black from Census	1.7	1.5	2.4	14.3
Mean % Black from HOLC	0.0	0.2	0.3	18.2
Mean Difference	**1.7**	**1.3**	**2.1**	**-3.9**
Mean % Immigrant from Census	9.9	11.1	13.6	14.3
Mean % Immigrant from HOLC	0.4	2.9	14.1	30.9
Mean Difference	**9.5**	**8.2**	**-0.5**	**-16.6**

- HOLC appraisers slightly underestimated the black population in higher-grade neighborhoods and overestimated it in low-grade neighborhoods
- HOLC appraisers highly underestimated the immigrant population in higher-grade neighborhoods and overestimated it in low-grade neighborhoods

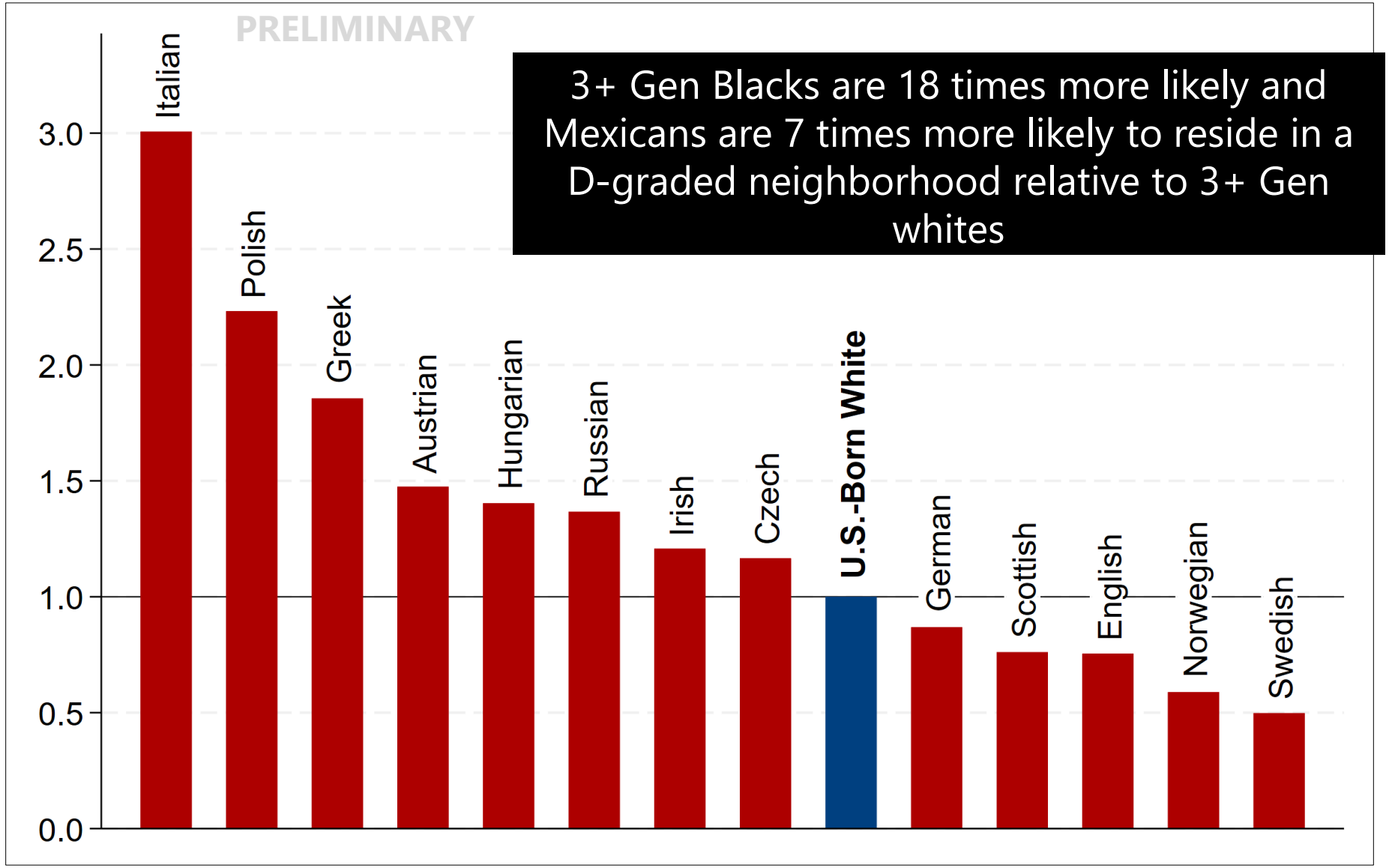
CHARACTERISTICS OF IMMIGRANTS (1940)

PRELIMINARY

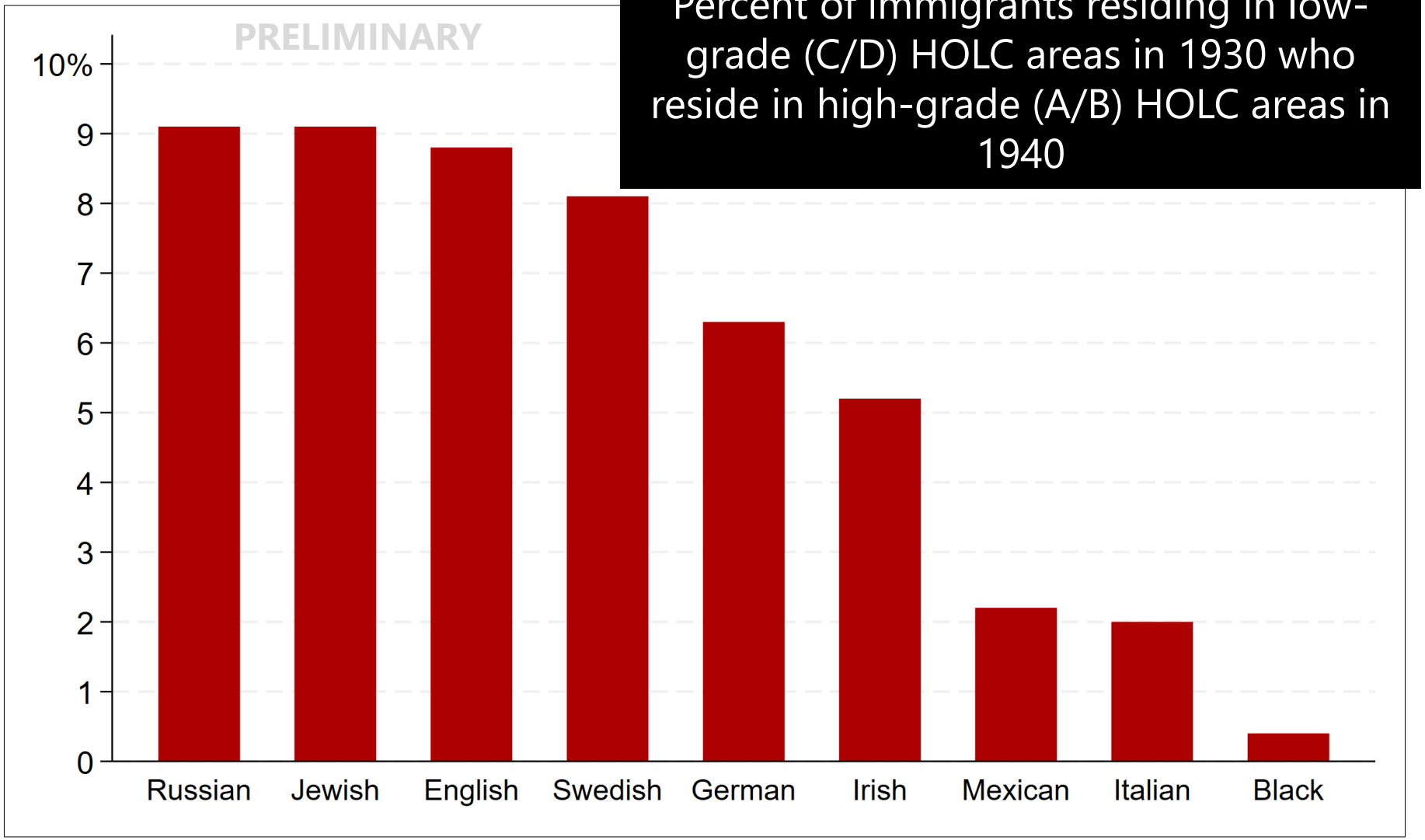
	Male	Child in Home	Max 4th Grade Ed	Own Home	Income
Swedish	56%	54%	8%	62%	840
English	50%	55%	8%	49%	1000
3+ Gen White	50%	63%	8%	45%	520
German	52%	57%	11%	58%	780
Irish	43%	66%	12%	44%	936
Russian	55%	74%	34%	40%	599
Greek	71%	67%	36%	32%	576
3+ Gen Latino	50%	72%	38%	41%	300
Jewish	51%	79%	39%	22%	624
3+ Gen Black	49%	54%	43%	23%	200
Italian	58%	81%	53%	47%	700

Italian, Jewish, and Greek immigrants were notably younger

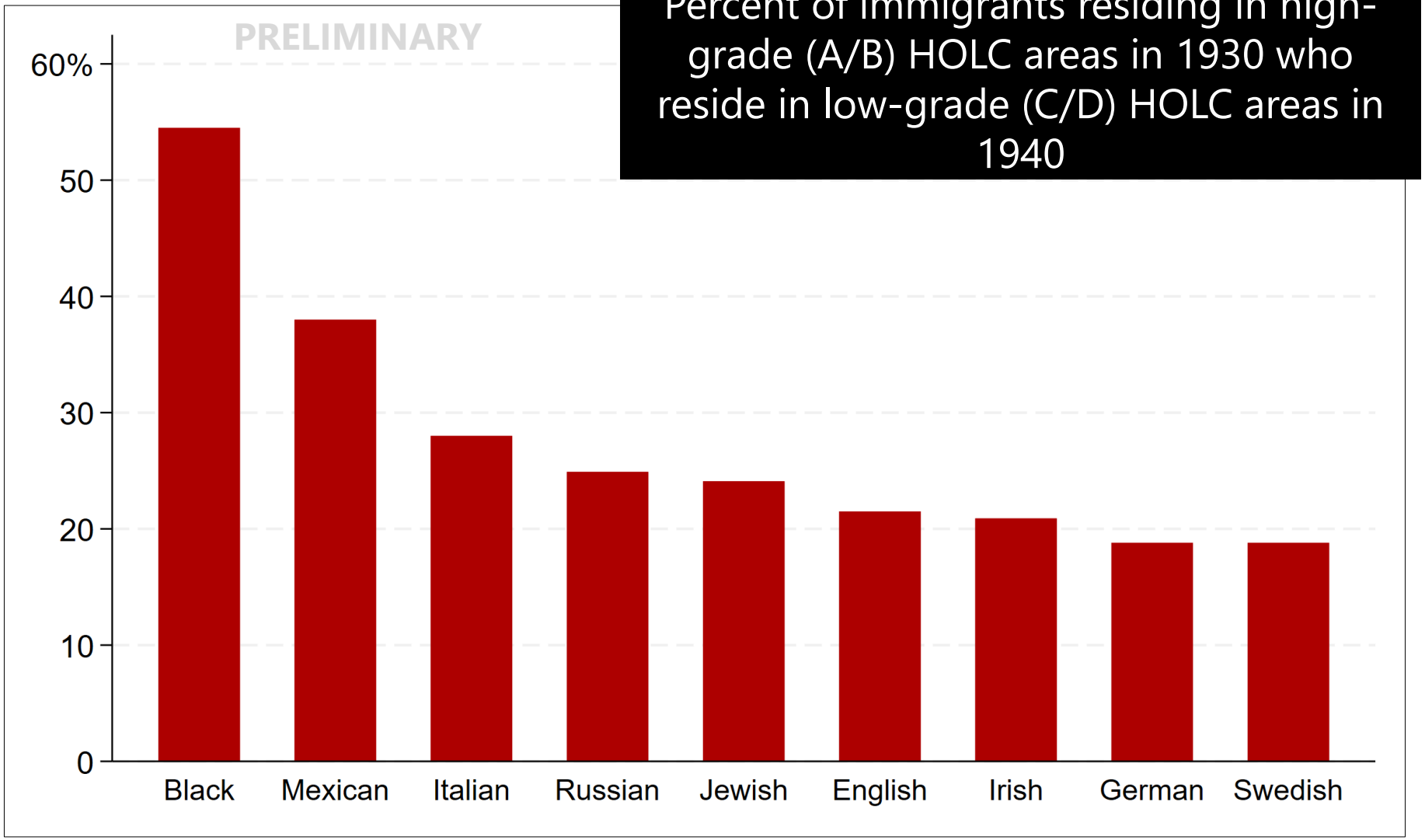
ODDS OF RESIDING IN D-RATED AREA (1940)



IMMIGRANTS TRANSITIONING UP 1930-1940



IMMIGRANTS TRANSITIONING DOWN 1930-1940



NEXT STEPS

- Transition away from the HOLC maps and incorporate other measures of neighborhood quality
- Digitize enumeration district maps for additional cities, including New York
- Use data on birthplace of mother and father to identify second-generation immigrants

USING FULL-COUNT CENSUS DATA AND THE MULTIGENERATIONAL LONGITUDINAL PANEL (MLP) TO STUDY URBAN CHANGE

Matthew H. Ruther
Kentucky State Data Center
University of Louisville
matthew.ruther@louisville.edu

Demographic Trends in Minnesota Homeownership

Jeff Howison, PhD

NOTE: We were unable to feature Jeff Howison's research in this webinar due to time constraints, but we will highlight his work in a future webinar. Stay tuned!
You can access his report here: Minnesota Homeownership

Thank you for joining us!

Mallory Bateman (SDC), mallory.bateman@utah.edu

Doug White (CIC), douglas.white@lsus.edu

David Van Riper, vanriper@umn.edu

Matthew Ruther, matthew.ruther@louisville.edu

Jeff Howison, jeff.howison@state.mn.us