Demographic-GIS Analysis of Spatial Population Distributions in US Urban Areas 1940

Mark Fossett, Sierra Laddusaw, and Walter Kamphoefner with primary assistance from Michael Upchurch. Additional assistance from J. Gilmartin, J. Ren, R. Earl, S. Norris, M. Smith, N. Vargas-Hirschfeld, Cindy Barrahona, and Wan Huang.

Pilot Study Establishes Feasibility & Value of Larger Project

Our pilot study of ten (10) cities in 1940 demonstrates the scientific value and technical feasibility of preparing GIS resources and spatially-referenced population data to support analysis of the spatial distribution of population in historical U.S. urban areas.

The results of this study will be crucial for establishing credibility of a proposal to the National Institutes of Health to extend the project to larger-scale and prepare a comprehensive data base of GIS resources and spatially referenced population summary files covering all major urban areas of the United States in 1940.

The use-case for the larger proposal will be to assess the effects of city and neighborhood context in 1940 on social mobility and later-life mortality. It will involve collaboration with research teams at UC-Berkeley and the University of Minnesota Population Center (UMPC) to prepare restricted-use data sets that link census records from 1940 with records from later censuses 1950-2010 and post-1940 mortality records.

The data products generated by the larger project will enable systematic research on a wide range of topics by demographers, economists, epidemiologists, geographers, historians, regional scientists, sociologists, & urban planners.

The Current State of Urban Research

Recent decades have seen a revolution in quantitative analysis of spatial population distributions in urban areas. The work is possible due to advances in Graphical Information Systems (GIS) software and digital resources for working with spatially referenced data. Unfortunately, the capabilities of contemporary research methods cannot be applied to earlier historical eras because crucial digital resources and spatially referenced data do not exist. This limits both cross sectional research of historical urban patterns and longitudinal studies investigating the origins of contemporary urban patterns.
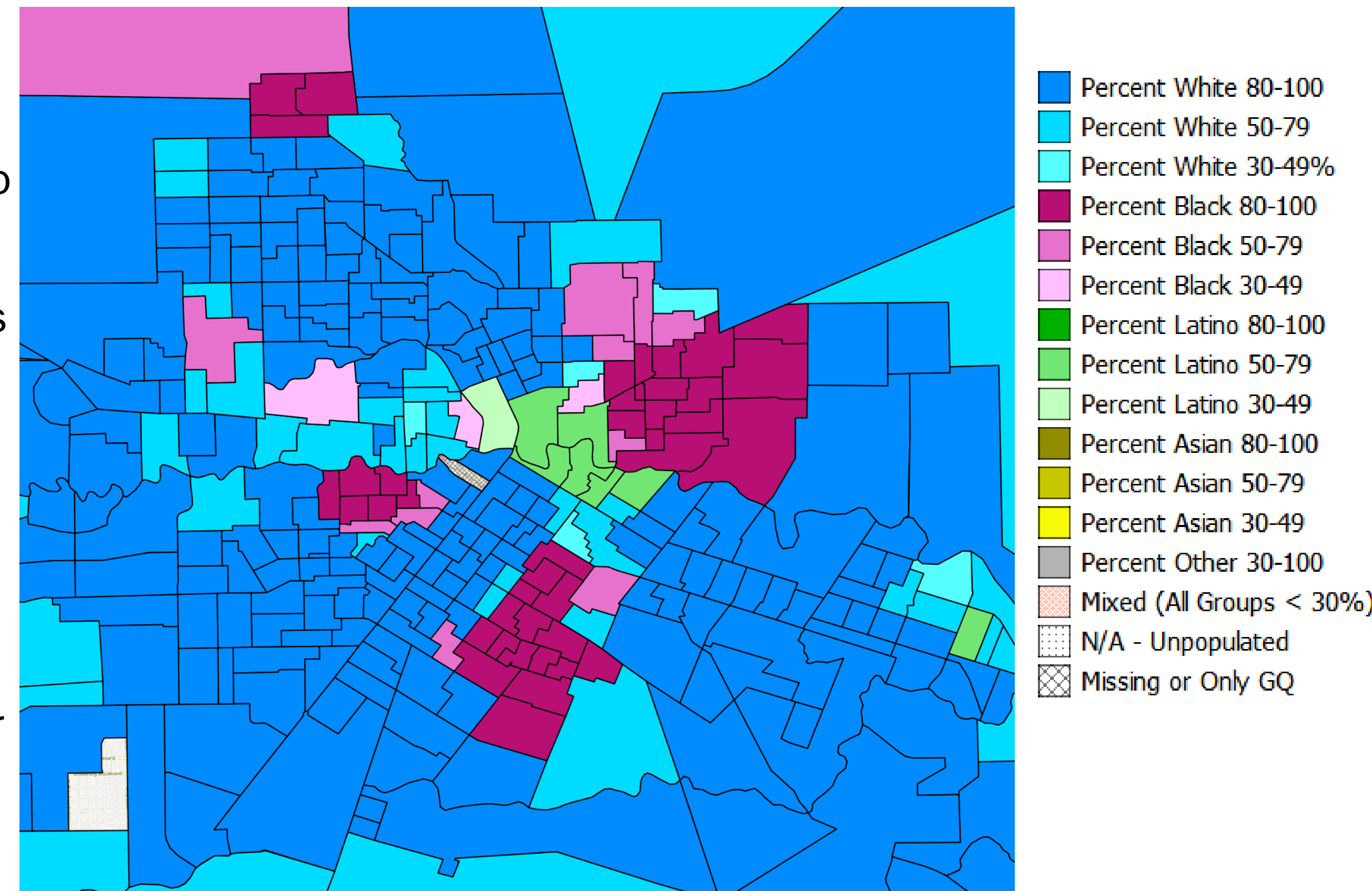
Addressing the Need

The results of our T3 pilot project will allow us to develop a highly credible proposal for a large-scale project to create a comprehensive data base of digital resources for research on historical urban areas. The project will be expensive because the production process is painstaking and labor intensive. Our pilot project demonstrates our research team's ability to prepare high-quality digital resources at acceptable cost to expand opportunities for research on urban spatial patterns.

Houston TX 1940 – Three Groups: Whites, Blacks, & Latinos

Patterns of segregation seen in 1940 persist to the present day.

Black presence is concentrated in the 3rd, 4th, and 5th Wards.

Latino presence in the 2nd Ward expands to the southeast in later decades.

Legend:
- Percent White 80-100
- Percent White 50-79
- Percent White 30-49%
- Percent Black 80-100
- Percent Black 50-79
- Percent Black 30-49
- Percent Latino 80-100
- Percent Latino 50-79
- Percent Latino 30-49
- Percent Asian 80-100
- Percent Asian 50-79
- Percent Asian 30-49
- Percent Other 30-100
- Mixed (All Groups < 30%)
- N/A - Unpopulated
- Missing or Only GQ

Buffalo NY 1940 – Many European Immigrant Groups

Complex spatial patterns are seen for European immigrant populations.

For example, immigrants from Poland and Italy are clearly segregated from each other in separate, spatially distinct enclaves. The level of segregation is higher than previous research could establish.

Blacks are segregated from all groups – overlapping the most with Italian immigrants.

Legend:
- Percent UK & Canada (30%)
- Percent UK & Canada (20-29)
- Percent UK & Cananda (10-19%)
- Percent Yugoslavia (10-19%)
- Percent Austria (10-19%)
- Percent Germany (10-19%)
- Percent Hungary (10-19%)
- Percent Poland (30%)
- Percent Poland (20-29%)
- Percent Poland (10-19%)
- Percent Ireland (30%)
- Percent Ireland (20-29%)
- Percent Ireland (10-19%)
- Percent Italy (30%)
- Percent Italy (20-29%)
- Percent Italy (10-19%)
- Percent Russia (10-19%)
- Mixed (All FB 30%)
- Mixed (All FB 20-29%)
- Mixed (All FB 10-19%)
- Percent Foreign Born < 10%
- Percent Black 50%
- N/A - Unpopulated
- Missing or Only GQ

Methods and Products

The research team created digital boundary files for census enumeration districts for ten (10) cities: Akron OH, Buffalo NY, El Paso TX, Fort Worth TX, Fresno CA, Minneapolis MN, Houston TX, San Antonio TX, Sacramento CA, and Tulsa OK. The boundary files were "hand-crafted" from historical maps and other historical resources. We developed a detailed project workflow manual in consultation with leaders of the University of Minnesota Population Center's National Historical Graphical Information Systems project. NHGIS leaders have agreed to endorse our proposal for the larger project and will host the GIS resources we will produce.

Spatially referenced data are needed to exploit the potential of GIS resources. Our project has processed restricted-use (proprietary) population data under secure computing arrangements to produce spatially referenced population summary files that can be distributed publicly. The potential the data files create for research are illustrated in the two example maps documenting racial-ethnic residential segregation patterns shown in this poster.
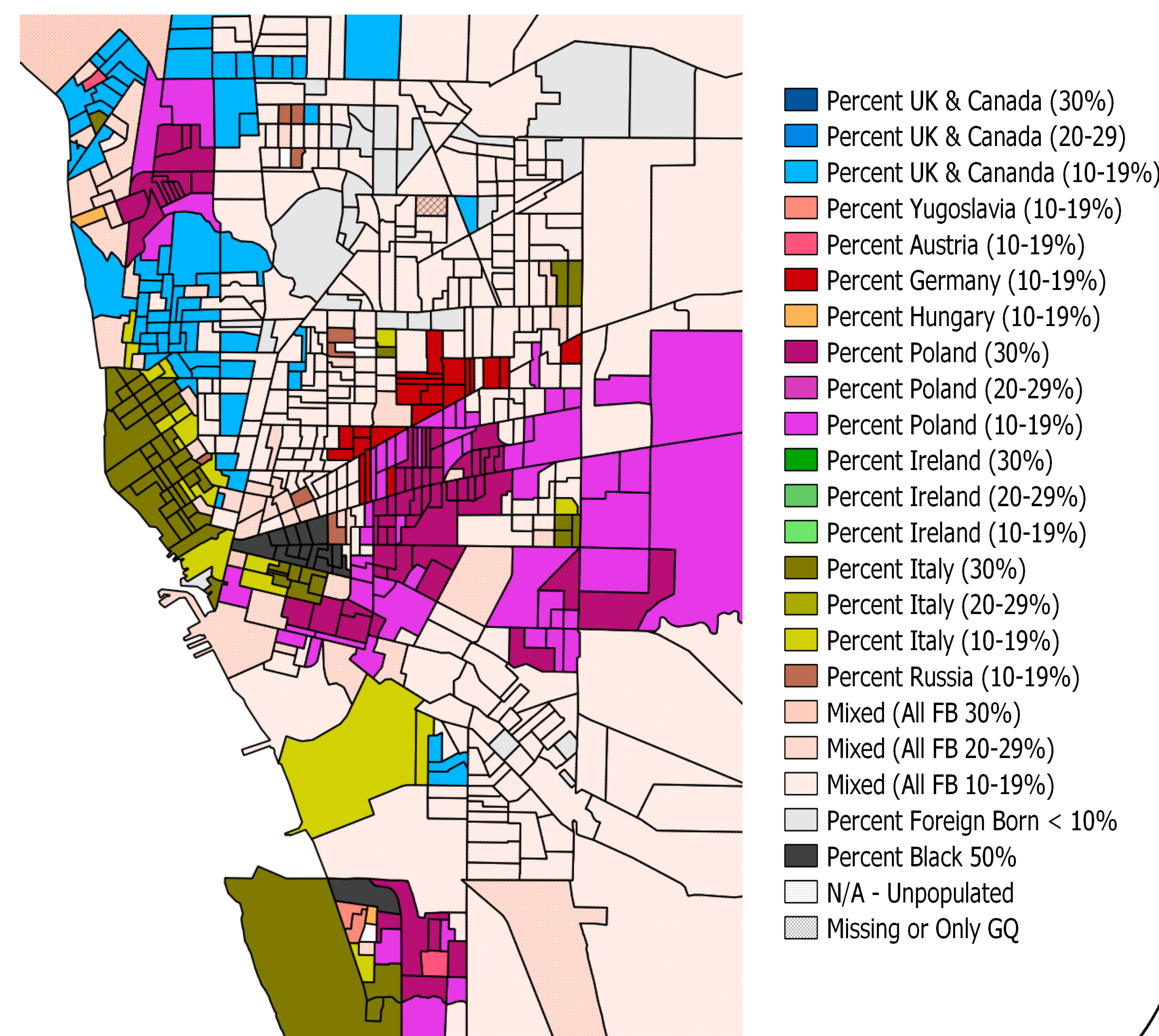
Note that analyses of this detail and quality were not feasible before the pilot project. Our larger proposal will create similar research possibilities for a larger set of cities.

The Next Step – A Larger-Scale Project

The value of digital GIS products and spatially referenced population data sets are obvious and unquestioned. But these products are expensive to produce. Accordingly, a competitive proposal for a larger project must have a compelling use-case to justify the investment of funds.

We will partner with research teams at UC-Berkeley and U. Minnesota to develop a compelling research application. The Berkeley team has developed a unique "linked mortality" file matching persons in the 1940 census to death records post 1940. The UM team is linking records for persons in the 1940 census to their records in later censuses. We will enhance their work by appending contextual data for neighborhoods and communities to the person records in 1940.

The three research teams will use the resulting database to assess the consequences of neighborhood and community conditions and residential segregation for later life mortality and social mobility. The data products will be subject to confidentiality protocols. Accordingly, we will place the files in the Federal Statistical Research Data Center (FSRDC) system and analyze them at the Texas RDC at Texas A&M University.