Matt Hodges
Final Project
Introduction to Geographic Databases
December 14, 2022

Introduction

Certain geographic features can give rise to economic development, chief among them is access to transportation. Cities and towns all along the upper Mississippi river flourished from the mid-nineteenth to the mid-twentieth century, having originally gained an industrial foothold through their proximity to the river as a major form of transportation and then as established industrial centers positioned exceptionally well for the manufacturing boom of the 1940s and 50s. As manufacturing declined in the latter decades of the twentieth century, however, these towns have seen sharp population declines and have struggled to find their new economic identity in the changing industrial landscape.

The aim of my project is to isolate the counties that border the Mississippi river from St. Louis to St. Paul, comparing certain demographic and economic characteristics within this cohort of counties and considering dynamics of intra- and interstate relationships between the counties and their respective states and between the counties and neighboring states. Certain paths of inquiry include:

1. What relationships exist between changes in population and changes in the economic make up of communities? For the communities that saw sharp changes in population, how did their unemployment rate move and the composition of their occupation and industry data change? Which is the driving factor: population or economic composition? Similarly, as the population shifts, were there stark changes in their age cohorts and educational and income statistics?

2. Did the economic data in the counties that border the Mississippi river from St. Louis to St. Paul change more precipitously than the other counties in their states? As a group, how do they compare to all the other counties in the five states they are a part of (Illinois, Iowa, Minnesota, Missouri, and Wisconsin)? Of these states, did some of their counties that border the river perform better than the other counties that border the river? If so, could this indicate better policies by some states versus others?

3. Does proximity to the river correlate to certain demographic and economic changes?

Data

Data were collected from two sources: US Census Bureau, both the decennial census and American Community Survey (ACS) and the Environmental Systems Research Institute (ESRI). Each of these sources provided the following information:

Decennial census: data were collected from decennial years 1970 to 2020 from the data hub Social Explorer, with the following attributes of interest identified: *Total Population*, *Population Density*, *Age*, *Educational Attainment for Population 25 and Over*, *Unemployment Rate*, *Industry*, *Occupation*, and *Household Income*. *Household*

Income was not collected in the 1970 census, so it will not be present in that decade's analysis. Likewise, the latter five features were not in the 2010 or 2020 census data, so the ACS was used for those years for those features (see next). https://www.socialexplorer.com/explore-tables

ACS: because certain features were not present in the 2010 and 2020 decennial census data, the 2006-2010 and 2016-2020 ACS data were used, collecting *Educational Attainment for Population 25 and Over*, *Unemployment Rate*, *Industry*, *Occupation*, and *Household Income* for those two decennial years. https://www.socialexplorer.com/explore-tables

ESRI: for the Mississippi river geographic data, I found shapefile from ESRI, which includes *esriGeometryPolyline* data for all the rivers across the US. The most pertinent attributes include *objectid* and *name*, although some other features like *state*, *miles*, and *shape__len* may prove useful. https://hub.arcgis.com/datasets/esri::usa-rivers-and-streams/explore?location=40.351646%2C-90.677577%2C5.00

Social Explorer: finally, to get as specified country geographic data as possible, TIGER files created by the US Census Bureau were downloaded from a different section of the Social Explorer database for each of the decennial census years, although a 2019 shapefile had to be substituted for 2020 as it appears not to be published yet. https://geodata.socialexplorer.com/collection/519a774f-e50f-43f9-943d-23ff15ac4617

Conveniently, the spatial data of the county and river shapefiles were both projected under SRID:4326 when they were downloaded, with the rivers as MultiLineString shapes and the counties in MultiPolygon form. In each of the county files, the main attributes of interest were the geometric data and the FIPS code, which would be used later to join the geom data to the census demographic and economic data. These tables had a range of row counts from as low as 3137 or as high as 3233, depending on whether territorial counties were included in the data and factoring county changes and additions over time. Likewise, for the river shapefiles, there were only two attributes of interest: the geom and the name of the river. The row count for the rivers table was much higher, with nearly 113,000 values, but that is because this data is broken down by river segments, not the whole length of an entire river—a problem that will be solve by finding distinct values for county names later in the analysis process. That said, when distinct river or stream names are considered in the entire table, there are still 20,543 values counted, coming to nearly six and a half rivers or streams per county on average across the two types of data tables.

For the decennial and ACS census data, only certain attributes were selected upon the initial retrieval of data, ones considered more relevant to the scope of this study. Because the data tracked in the census tends to change from one decennial count to the next, not all the categories were conveniently arranged across the breadth of the tables. For instance, in the 1970 and 1980 census, Family Income was tracked, but that was discontinued in favor of Household Income in 1990, so a decision was made to track Household income only back to 1980, the census in which it was first measured. Similarly, the way educational attainment was recorded

changed, with the 1970 census having 13 categories and the 2000 census having just over half that at seven categories:

SE:T22. Educational Attainment for Population 25 Years and Over		
Total Population 25 Years Old and Over:	12,068	
No school years completed (includes nursery and kindergarten)	426	3.5%
Elementary:	3,872	32.1%
1-4 years	1,001	8.3%
5-6 years	1,198	9.9%
7 years	670	5.6%
8 years	1,003	8.3%
High school:	6,070	50.3%
1-3 years	2,313	19.2%
4 years	3,757	31.1%
College:	1,700	14.1%
1-3 years	933	7.7%
4 years	507	4.2%
5 years or more	260	2.2%

Table 1: 1970 census education attainment categories for Total Population 25 and Over

SE:T40. Educational Attainment For Population 25 Years And Over		
Population 25 years and over:	27,589	
Less Than High School	5,872	21.3%
High School Graduate (includes equivalency)	9,332	33.8%
Some college	7,413	26.9%
Bachelor's degree	3,245	11.8%
Master's degree	1,329	4.8%
Professional school degree	276	1.0%
Doctorate degree	122	0.4%

Table 2:2000 census education attainment categories for Total Population 25 and Over

The resolution to these issues is discussed in the next section of this study, but it should be noted here that each table of census data from the source require substantial standardization to aid in further analysis. Total attributes and row counts of these tables from the source range from 75 to 93 for the former and 3137 to 3233 for the latter, the same row counts as the TIGER files in the spatial data.

Analysis

The first order of business before analysis could occur was to clean and standardize the data. It was already determined that some data did not span the entire timescale (industry classifications and household income levels), so the analysis of those categories would just need to include and note the relevant decades. Attributes such as population, population density, area, and household count did not need any adjustments, except to standardize the names for any minor discrepancies from one census to another. There was another minor issue with both the FIPS attribute and the county and state naming conventions. First, with FIPS, the shapefile data contained a leading zero for FIPS codes that were less than five digits in length, while the census data did not have the leading zero for the same four-digit codes. This would create a problem

later when joining the tables on the FIPS attribute. There were two ways to solve it. First, using Excel, a TEXT formula would change the attribute to text and allow the leading zero to be automatically formatted onto the column:

FIPS	Name	State	Population
1001	Autauga	Alabama	43671
1003	Baldwin	Alabama	140415
1005	Barbour	Alabama	29038
1007	Bibb	Alabama	20826
1009	Blount	Alabama	51024

Table 3: FIPS category in Excel table before added leading zero

FIPS	Name	State	Population
01001	Autauga	Alabama	43671
01003	Baldwin	Alabama	140415
01005	Barbour	Alabama	29038
01007	Bibb	Alabama	20826
01009	Blount	Alabama	51024

Table 4: FIPS category after =TEXT(A2,"00000") applied

This process was easy enough, but because the files were saved in .csv format, they did not always save properly and did not always import into pgAdmin or QGIS correctly. Alternative, in a similar way in QGIS, a new field could be created in the shapefile that deleted the leading zero and made the join possible. Ultimately, the latter method was deployed for ease of use and functionality when joining the tables. To solve the issue with the county and state names, more work was done in Excel. Each county name was formatted like this, *Autauga County*, and the state information looked like this, *Autauga County, Alabama*. A simple SUBSTITUTION formula was used to remove unwanted text, where the unwanted part in the cell was identified and replaced with zero characters. In the end, this left just the county name and the state name for the values of each record.

FIPS	Name	State	Population
1001	Autauga County	Autauga County, Alabama	32259
1003	Baldwin County	Baldwin County, Alabama	78556
1005	Barbour County	Barbour County, Alabama	24756
1007	Bibb County	Bibb County, Alabama	15723
1009	Blount County	Blount County, Alabama	36459

Table 5: Name and State categories before standardization

FIPS	Name	State	Population
1001	Autauga	Alabama	32259
1003	Baldwin	Alabama	78556
1005	Barbour	Alabama	24756
1007	Bibb	Alabama	15723
1009	Blount	Alabama	36459

Table 6: Name and State after =SUBSTITUTE(B2," County","") applied

The final preparations in Excel included calculating labor force participation and unemployment rates and narrowing the data with varying characteristics over the time horizon (i.e., age, industry, income, education, etc.) into consistent category sizes. For this second task, age was put into three categories (under 18, 18 to 54, and 55 plus), education was divided into four categories (less than high school, high school graduate, some college, and four plus years of

college), five industry categories were identified that easily tracked across the time horizon (construction, manufacturing, wholesale, retail, and other), and income levels were put into four categories (under $25,000, $25,000 to $74,999, $75,000 to $149,999, and over $150,000). For the 1970 data, there were 16 attributes, where 1980 and 1990 had 21 (included industry categories), and 2000 through 2020 had the full set of 25 attributes (including the income levels). This became the final version of the census data to be joined to the shapefiles and prepared for analysis. Conceptually, though the tables were most pre-joined in Excel, the data can be thought of distinct in the following breakdown of this ER diagram:

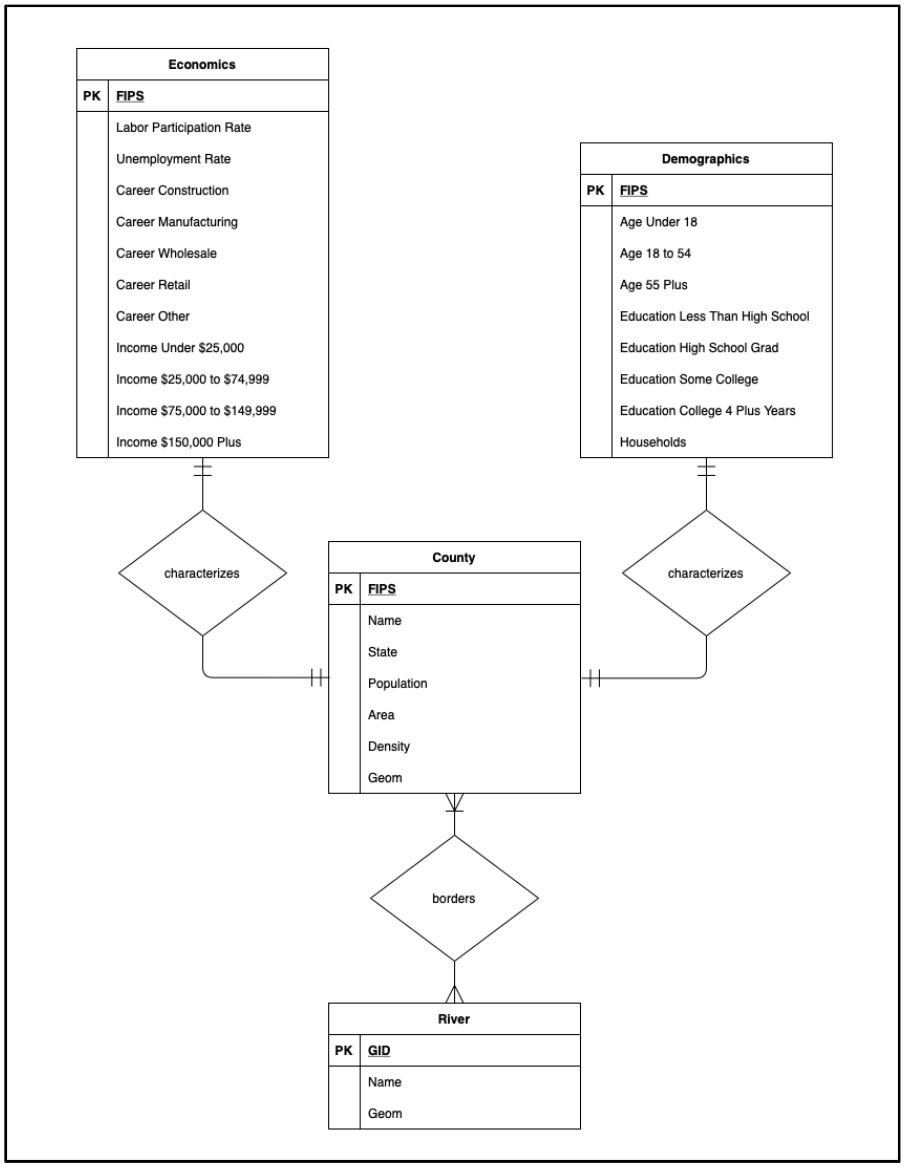


Figure 1: ER diagram with four conceptual tables and listed attributes

These Excel .csv files were now ready to join on the FIPS attribute with the shapefiles. This was a relatively simple process performed in QGIS, where the only attribute kept from the original shapefile was the geom data. Each of these resulting tables formed the basis of analysis and could be joined geographically to the rivers table and to each other to tell the story of the population, economic, and demographic changes over time of the region under review.

Results

Two geographic queries components were required first to identify the counties that border the Mississippi river and second to exclude the counties that were south of St. Louis and north of St. Paul (specifically Minneapolis in the end, as the is the city identified in the data). Here is the query structure that accomplished this task:

```
WHERE ST_Intersects(r.geom,sev.geom)
AND r.name='Mississippi River'
AND ST_AsLatLonText(ST_Centroid(sev.geom)) >=
    ST_AsLatLonText((SELECT the_geom FROM world.cities WHERE name='St. Louis'))
AND ST_AsLatLonText(ST_Centroid(sev.geom)) <=
    ST_AsLatLonText((SELECT the_geom FROM world.cities WHERE name='Minneapolis'))
ORDER BY sev.state
```

Forty-four counties fall within these parameters, twelve from Illinois (the most), ten from Iowa, five from Minnesota (the least), nine from Missouri, and eight from Wisconsin.

Once the geographic challenge was overcome, then it became a matter of identifying and joining the appropriate tables and selecting for the desired statistics. An early success identified the appropriate counties and selected the population for each census:

Iowa Counties	1970	1980	1990	2000	2010	2020
Allamakee	14968	15108	13855	14675	14330	14061
Clayton	20606	21098	19054	18678	18129	17043
Clinton	56749	57122	51040	50149	49116	46460
Des Moines	46982	46203	42614	42351	40325	38910
Dubuque	90609	93745	86403	89143	93653	99266
Jackson	20839	22503	19950	20296	19848	19485
Lee	42996	43106	38687	38052	35862	33555
Louisa	10682	12055	11592	12183	11387	10837
Muscatine	37181	40436	39907	41722	42745	43235
Scott	142687	160022	150979	158668	165224	174669

Table 7: Population change over the decades for each of the Iowa Counties

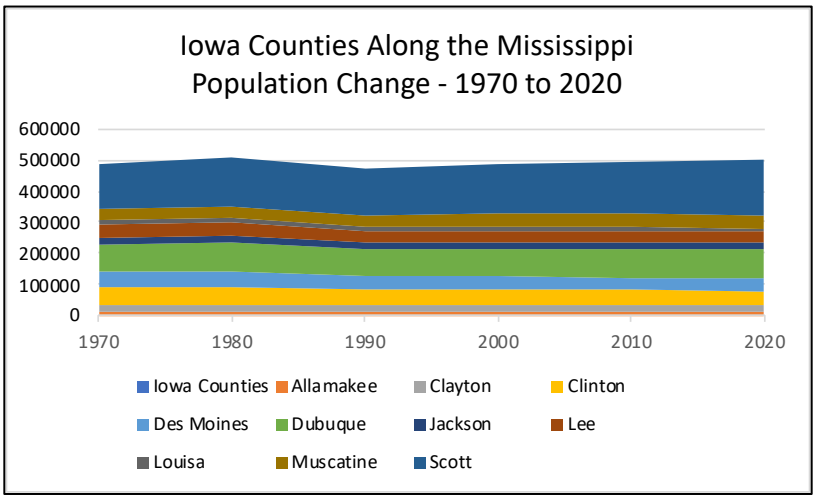
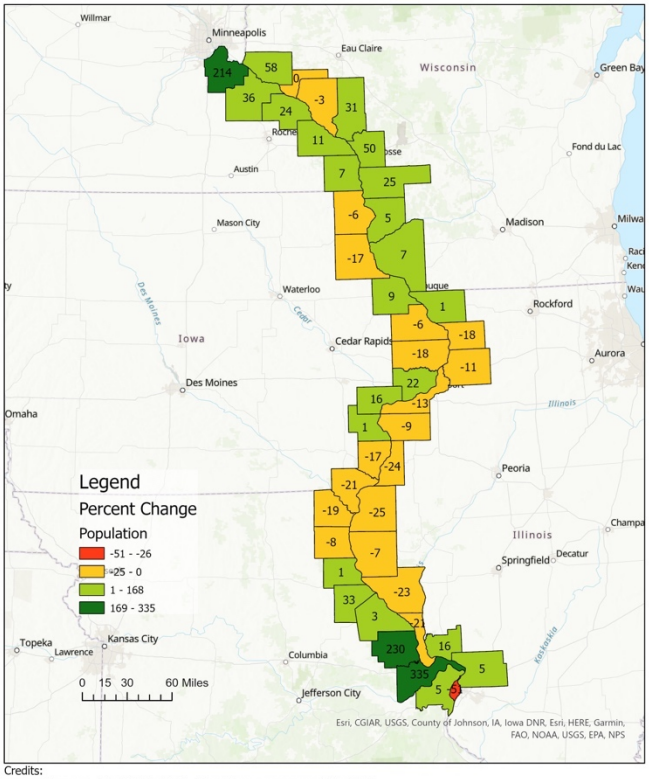


Figure 2: chart representation of the same data in Table 7

Though most of the queries were similar in structure, with each including the county name, the state which the county resides, and the geom information, several challenges popped up for each category of analysis below. Ultimately, for the sake of time, the intended intra- and interstate analysis had to be postponed but the basic structure of the data is there to continue that analysis at a later date.

Population: Of the forty-four counties bordering the Mississippi river from St. Louis to St. Paul, twenty had declining populations and twenty-four saw increases. St. Louis City county (not to be confused with its neighbor St. Louis County) had the steepest decline from 1970 to 2020, seeing just over half of its population (51%) leave. Conversely, two St. Louis suburban counties—Lincoln and St. Charles—saw massive increases of 230% and 335% respectively, the two largest increases of the cohort. Similarly, Dakota county Minnesota saw a substantial increase at 214%, indicating that urban sprawl was likely a present phenomenon over this time horizon. On a state level, there were clear winners and loser from a population perspective. All five of Minnesota's counties saw their populations increase, and six of Wisconsin's eight counties (75% of the counties; although their two declining populations were on the small end) did too. Missouri had six of nine increased (66%), but Iowa and Illinois each had a majority of counties lose population, with the former at six in ten (60%) and the latter nine of twelve (75%). The tri-state area of Illinois, Iowa, and Missouri did not fare well over the period, and west-central Illinois has clearly struggled.

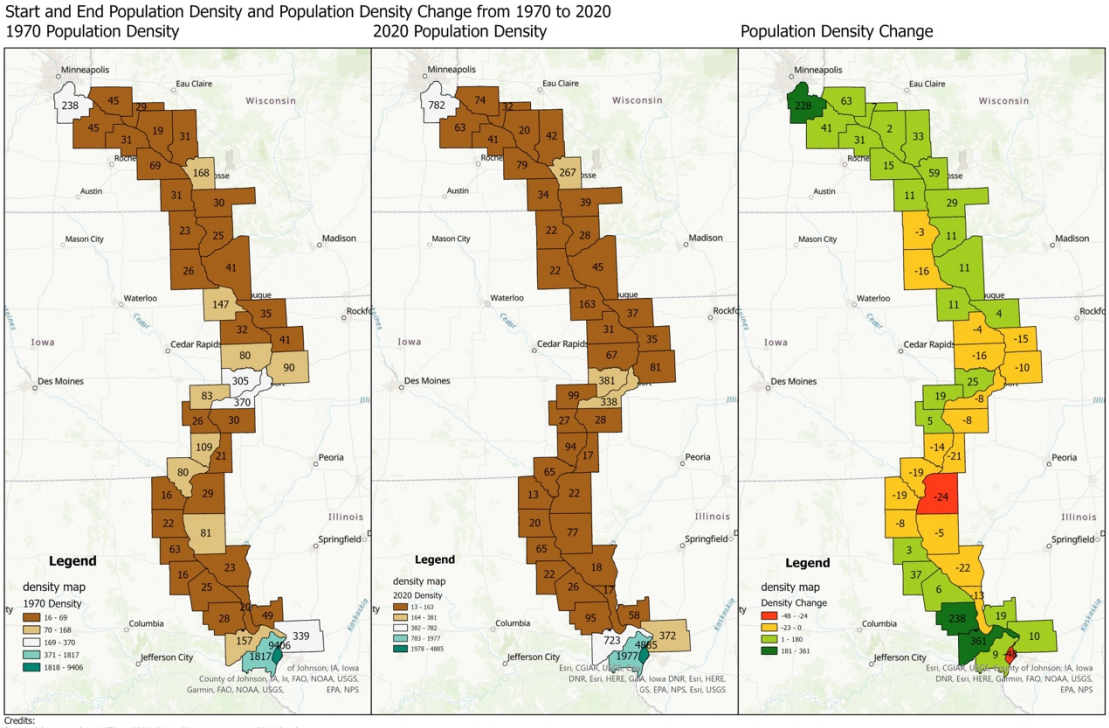
Population Change from 1970 to 2020,
Upper Mississippi River Counties from St. Louis to St. Paul



Credits:
Decennial census data 1970 to 2020, https://www.census.gov/data.html
ACS 2006-2010 & 2016-2020 data, https://www.census.gov/programs-surveys/acs/data/data-tables.html
Map Projection:
NAD83(2011) / Iowa South, EPSG:6464

Density: The population density map tells much the same story as the population numbers, mainly because the area of these counties has been fixed for quite some time. Wisconsin fares better in this analysis for some reason, with no counties showing up with density declines. The tri-state area again demonstrates steep lose, with Hancock county in Illinois, Lee county in Iowa, and Clark county in Missouri seeing 24%, 19%, and 19% declines respectively in their population density.

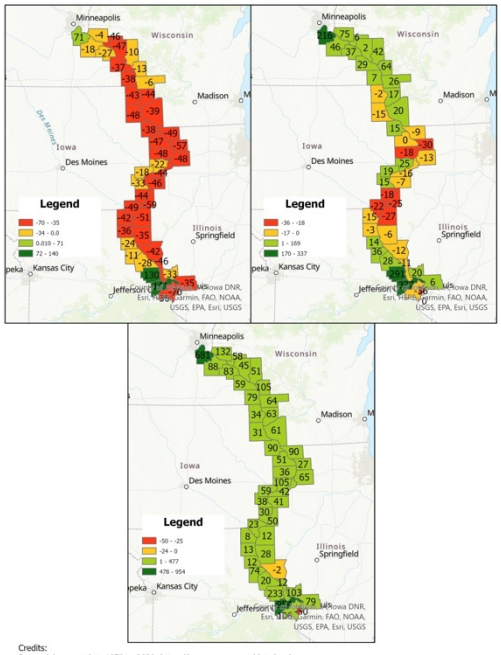
Start and End Population Density and Population Density Change from 1970 to 2020



Credits:
Decennial census data 1970 to 2020, https://www.census.gov/data.html
ACS 2006-2010 & 2016-2020 data, https://www.census.gov/programs-surveys/acs/data/data-tables.html
Map Projection:
NAD83(2011) / Iowa South, EPSG:6464

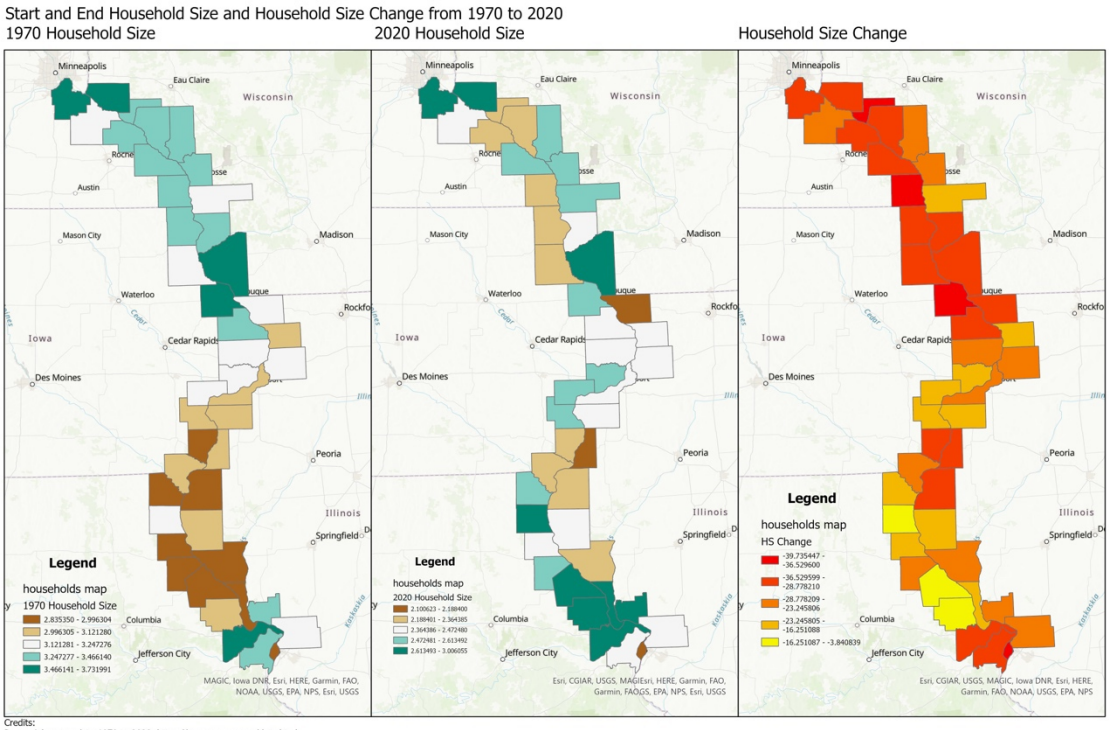
Age: The age analysis tells an interesting tale. For populations under 18, all but three of the counties saw declines in their numbers, those being the counties from the population analysis that had the highest overall increases: Lincoln and St. Charles in Missouri and Dakota in Minnesota. For the middle age group (age 18 to 54) the map looks fairly similar to the population map, in terms of which counties saw increases and which saw declines. Although there was some shuffling between counties, twenty saw declines and twenty-four saw increases. Finally, for ages fifty-five plus, all but two counties saw their elderly populations increase. Of the two that saw a drop, one was at just a 2% decline and the other was St. Louis City county at 50%, so its elderly population at least did not drop as steep as its overall population, even if by only 1%.

Age Group Change from 1970 to 2020,
Under 18 (top left) - 18 to 54 (top right) - 55 plus (bottom)



Credits:
Decennial census data 1970 to 2020, https://www.census.gov/data.html
ACS 2006-2010 & 2016-2020 data, https://www.census.gov/programs-surveys/acs/data/data-tables.html
Map Projection:
NAD83(2011) / Iowa South, EPSG:6464

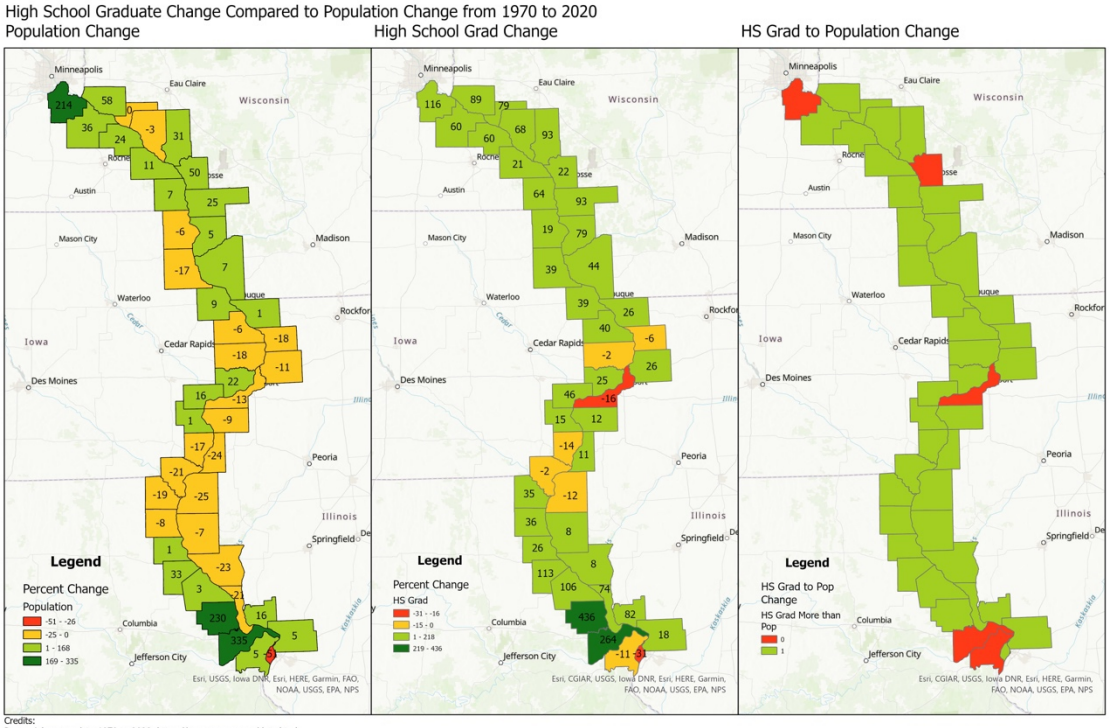
Households: Upon the cursory analysis of this study, there does not seem to be much of a household story to tell. Under declining and aging populations, household sizes tend to decrease, which is true of every county in this cohort, even among the counties that saw their populations increase. In the first decade analyzed, there did seem to be a correlation between household size and cardinal direction, with households increasing in size as populations moved northward, with the St. Louis suburban counties as the exception. In the 2020 numbers, this does not seem to hold true.



Credits:
Decennial census data 1970 to 2020, https://www.census.gov/data.html
ACS 2006-2010 & 2016-2020 data, https://www.census.gov/programs-surveys/acs/data/data-tables.html
Map Projection:
NAD83(2011) / Iowa South, EPSG:6464

Education: For three out of the four categories of education analysis, all the counties moved in unison either downward or upward. For instance, the *Less than High School* category saw between a 31% (Dakota county, Minnesota) and 90% (Calhoun county, Illinois) decreases across the board. This could be both a function of decreasing younger age cohorts and a stronger push to increase high school graduation rates, although more data would need to be analyzed to confirm this. The higher two education categories—*Some College* and *College Four Years Plus*—saw substantial increases, which tracks with the growth of college enrollments over the same time horizon. For the *High School Graduate* category, there were eight counties that saw decreases, all in Illinois, Iowa, or Missouri. Although when compared to the shift in population that number drops to five counties that had worse performing rates than their population, this time with only Iowa not having a county that's high school grad population change rate dipped below its population change rate.

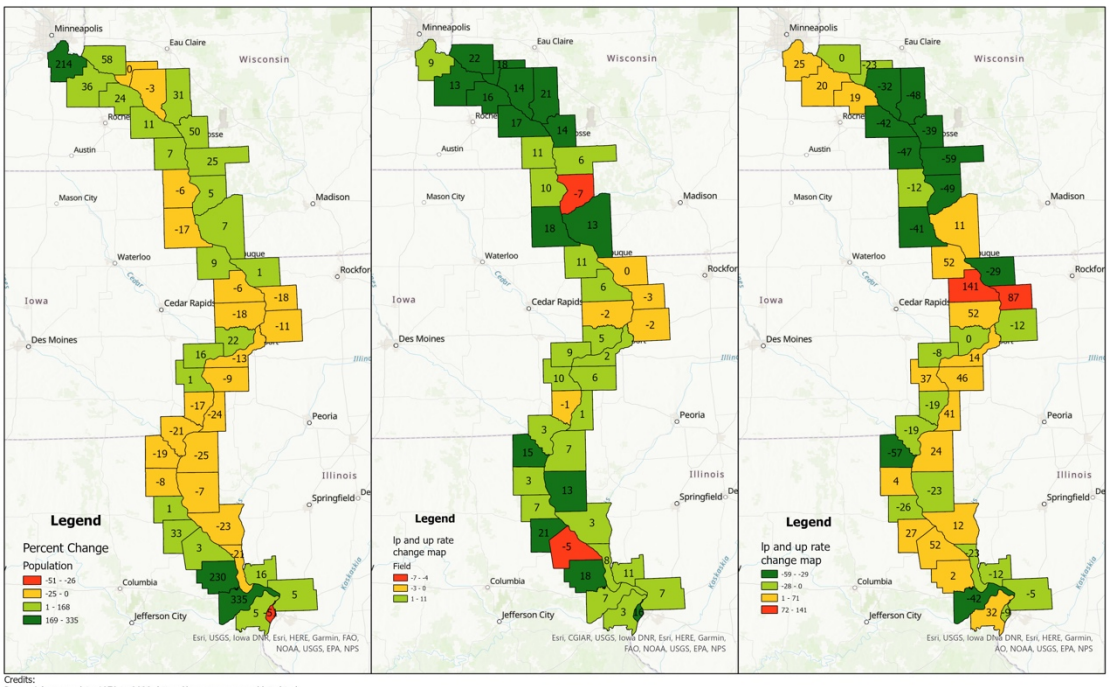
High School Graduate Change Compared to Population Change from 1970 to 2020



Credits:
Decennial census data 1970 to 2020, https://www.census.gov/data.html
ACS 2006-2010 & 2016-2020 data, https://www.census.gov/programs-surveys/acs/data/data-tables.html
Map Projection:
NAD83(2011) / Iowa South, EPSG:6464

Labor Force Participation and Unemployment Rates: Larger trends in the labor force participation rate and the unemployment rate are determined by much more than

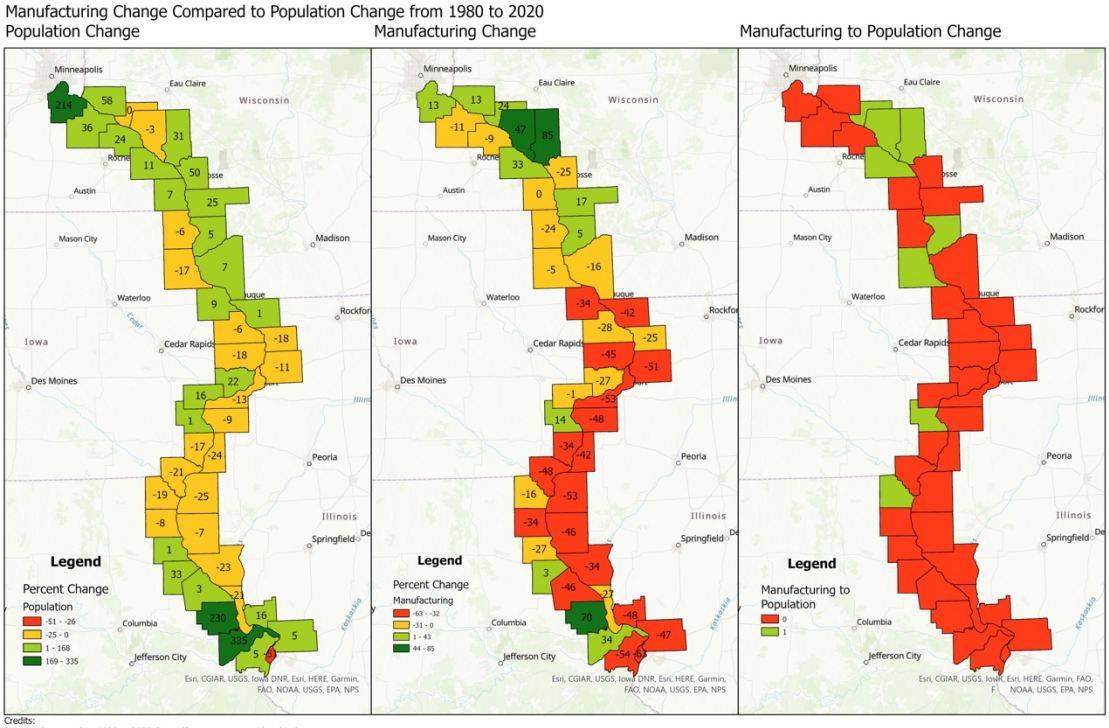
Labor Participation Rate and Unemployment Rate Change from 1970 to 2020



Credits:
Decennial census data 1970 to 2020, https://www.census.gov/data.html
ACS 2006-2010 & 2016-2020 data, https://www.census.gov/programs-surveys/acs/data/data-tables.html
Map Projection:
NAD83(2011) / Iowa South, EPSG:6464

population, so this cursory analysis might not be of much use. For participation, the map does not show any particular rhyme or reason, especially compared to the population change map. Counties in Minnesota, Missouri, and Wisconsin have tended toward higher participation, where Iowa and Illinois have trouble spots. For unemployment, the coloring on the map is reversed, as a decrease in the unemployed population over time is a good thing. Again, this map does not capture the myriad variables that influence employment, but might provide useful analysis if the reader's eyes do not cross too quickly while trying to make heads or tails of it.

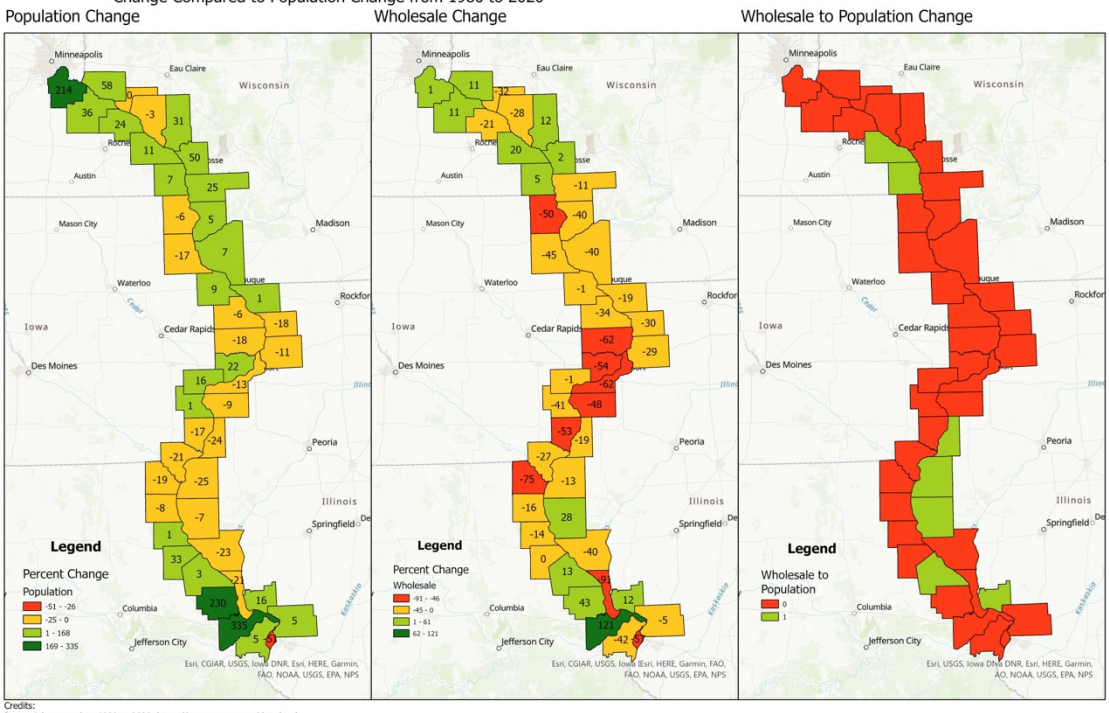
Industry: The industry level analysis provides some of the most interesting finds of this study, with promising results that could furthered by a more robust understanding of the census data. Three of the five categories saw declines over the time horizon in most counties in the cohort: manufacturing, wholesale, and retail. When compared to their population changes, only eight, seven, and eleven counties in those respective categories outpaced their rate of population change. In other words, they either grew by more than their populations, or if they decreased, they did not decline as much as their total populations did. For the construction category, only three counties did not keep pace with their populations, and for the other category, every county saw increases compared to their total population. This is where further analysis would be helpful, could the challenges of tracking census categories over time be resolved.



Credits:
Decennial census data 1980 to 2020, https://www.census.gov/data.html
ACS 2006-2010 & 2016-2020 data, https://www.census.gov/programs-surveys/acs/data/data-tables.html
Map Projection:
NAD83(2011) / Iowa South, EPSG:6464

Wholesale

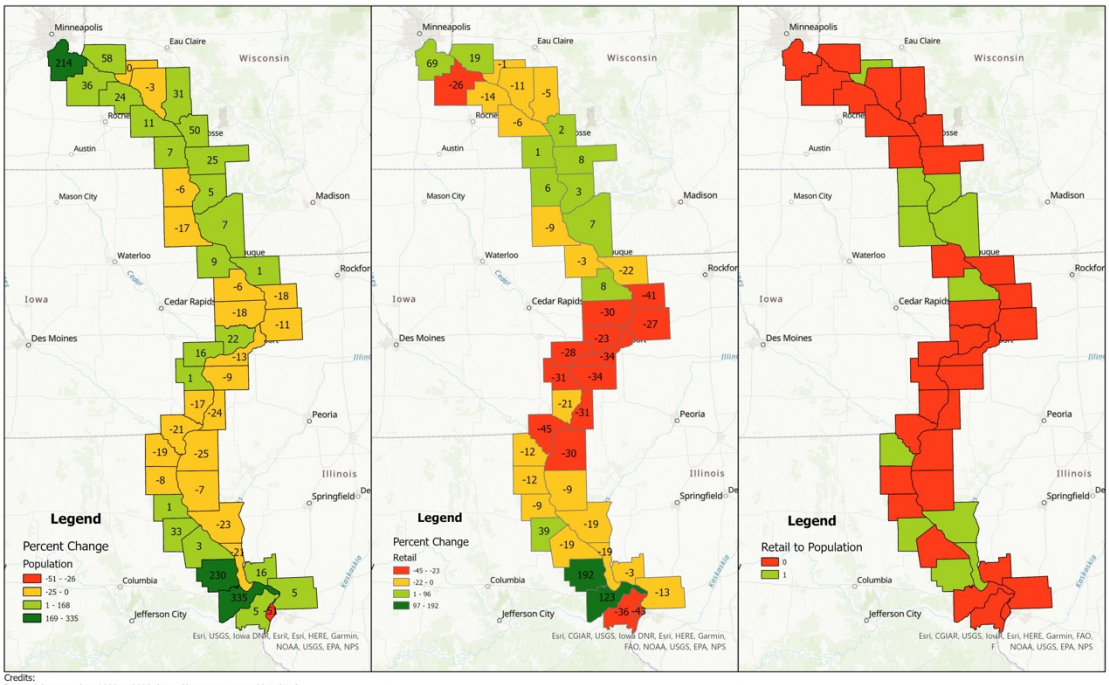
Change Compared to Population Change from 1980 to 2020



Credits:
Decennial census data 1980 to 2020, https://www.census.gov/data.html
ACS 2006-2010 & 2016-2020 data, https://www.census.gov/programs-surveys/acs/data/data-tables.html
Map Projection:
NAD83(2011) / Iowa South, EPSG:6464

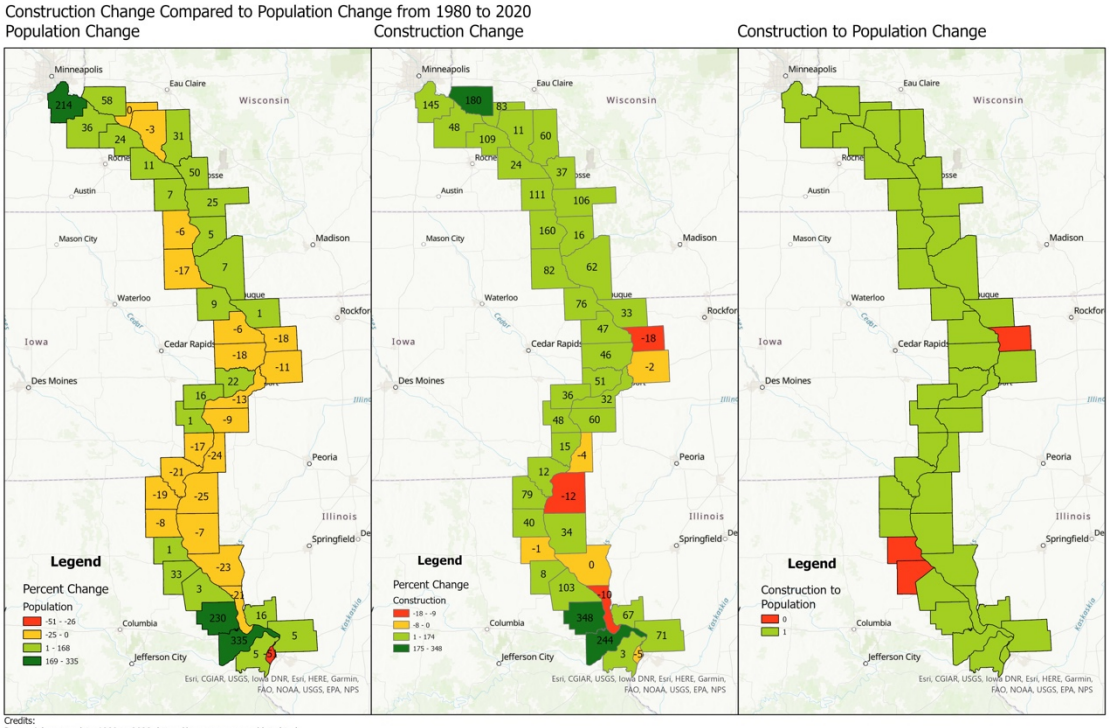
Retail

Change Compared to Population Change from 1980 to 2020



Credits:
Decennial census data 1980 to 2020, https://www.census.gov/data.html
ACS 2006-2010 & 2016-2020 data, https://www.census.gov/programs-surveys/acs/data/data-tables.html
Map Projection:
NAD83(2011) / Iowa South, EPSG:6464

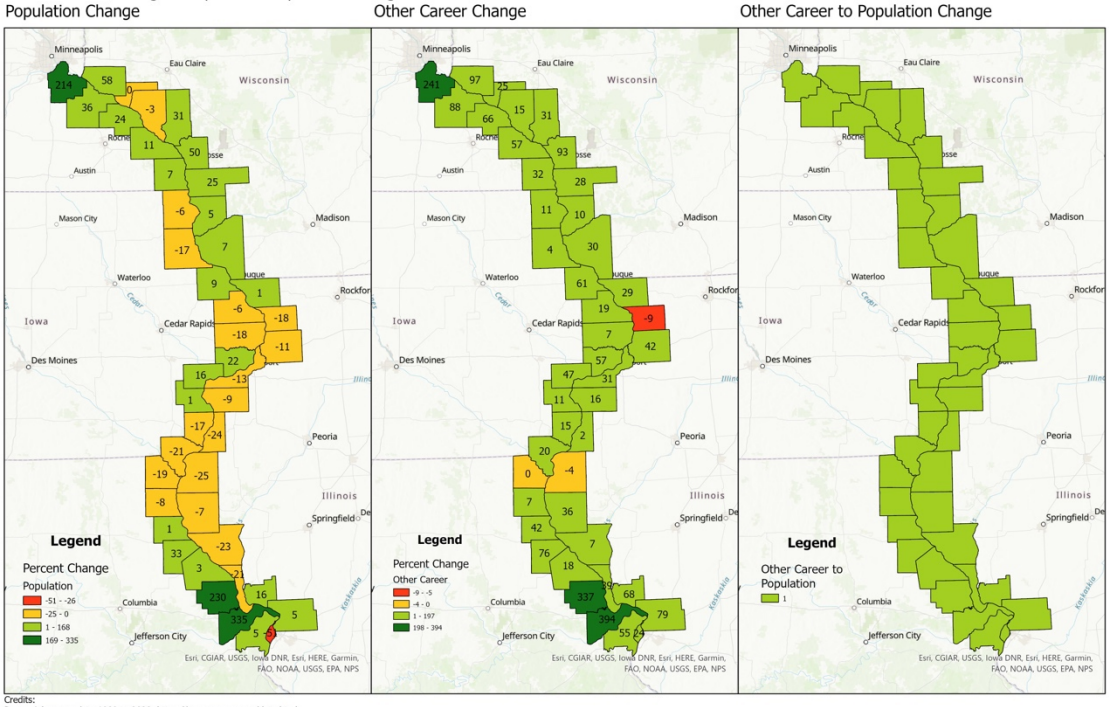
Construction Change Compared to Population Change from 1980 to 2020



Credits:
Decennial census data 1980 to 2020, https://www.census.gov/data.html
ACS 2006-2010 & 2016-2020 data, https://www.census.gov/programs-surveys/acs/data/data-tables.html
Map Projection:
NAD83(2011) / Iowa South, EPSG:6464

Other Career

Change Compared to Population Change from 1980 to 2020



Credits:
Decennial census data 1980 to 2020, https://www.census.gov/data.html
ACS 2006-2010 & 2016-2020 data, https://www.census.gov/programs-surveys/acs/data/data-tables.html
Map Projection:
NAD83(2011) / Iowa South, EPSG:6464

Income: The analysis for this section was quickly identified as flawed, as the income was clearly not adjusted for inflation. The categories either saw exaggerated declines or increases over the two decades reviewed.

Iowa Counties	Under $25,000	$25,000 to $74,999	$75,000 to $149,999	$150,000 plus
Allamakee	-38	-14	211	429
Clayton	-41	-18	286	367
Clinton	-35	-20	127	528
Des Moines	-32	-20	120	271
Dubuque	-31	-19	204	310
Jackson	-40	-25	208	791
Lee	-37	-17	119	196
Louisa	-41	-29	146	498
Muscatine	-40	-16	107	308
Scott	-32	-16	83	371

Table 8: Income category change from 2000 to 2020

Conclusion

Ideally, the findings of a study like this could be used as a basis for rationale for special economic consideration from the state or federal government. Similar economically depressed regional areas have been identified throughout the country (i.e., the rust belt) and have received extra assistance to recover. This would be of interest to economic developer, politicians, and policy makers, as well as key stakeholders in within the regional area like local governments and large employers. With further analysis to see how this region compares to its neighbors and to those other regions who have receive special attention, the argument could be made that this region needs similar support. Additionally, by comparing the performance of the states in this study and how they maybe have instituted different policies, the varying performances of the economies of these counties could indicate the efficacy of those policies.

From the perspective of the project, there were ups and downs after quite a slow start, but ultimately I am happy with the final product. I got some good experience across the three major apps we used in this course—pgAdmin, QGIS, and ArcGIS Pro—and I have the confidence to now branch out and explore geographic databases on my own. The queries in this project are not particularly sophisticated and mostly just consider rates of change over the time scale. More robust statistical analysis could be applied to get closer to the heart of the questions of this project. With more time I would incorporate these more in-depth statistical tools and broaden the comparison set of counties and states, as well as include other geographical components like how cities, universities, public schools, or any other pertinent geographic entities might change this analysis. I believe the scope is this project as a whole was a good culminating experience for the course and allowed for students to utilize the tools and skills learned in the course. I appreciate the opportunity to put these to work in this way.

Queries

Get total values of population per decade

```
select distinct(sev.name) as "County",
        sev.state as "State",
        sev.population as "1970",
        eig.population as "1980",
        nin.population as "1990",
        zer.population as "2000",
        ten.population as "2010",
        twe.population as "2020",
        sev.geom
from project.census_70 as sev
inner join project.census_80 as eig on sev.fips=eig.fips
inner join project.census_90 as nin on sev.fips=nin.fips
inner join project.census_00 as zer on sev.fips=zer.fips
inner join project.census_10 as ten on sev.fips=ten.fips
inner join project.census_20 as twe on sev.fips=twe.fips,
        project.rivers as r
where ST_Intersects(r.geom,sev.geom)
and r.name='Mississippi River'
and st_aslatlontext(st_centroid(sev.geom)) >=
        st_aslatlontext((select the_geom from world.cities where name='St. Louis'))
and st_aslatlontext(st_centroid(sev.geom)) <=
        st_aslatlontext((select the_geom from world.cities where name='Minneapolis'))
order by sev.state
```

Get percent change in population over time horizon

```
select distinct(sev.name) as "County",
        sev.state as "State",
        (twe.population-sev.population)*100/sev.population as "Population Change",
        sev.geom
from project.census_70 as sev
inner join project.census_20 as twe on sev.fips=twe.fips,
        project.rivers as r
where ST_Intersects(r.geom,sev.geom)
and r.name='Mississippi River'
and st_aslatlontext(st_centroid(sev.geom)) >=
        st_aslatlontext((select the_geom from world.cities where name='St. Louis'))
and st_aslatlontext(st_centroid(sev.geom)) <=
        st_aslatlontext((select the_geom from world.cities where name='Minneapolis'))
order by sev.state
```

Get change in each income category 2000 to 2020

```
select distinct(zer.name) as "County",
        zer.state as "State",
        (twe.under_twentyfive-zer.under_twentyfive)*100 /
                zer.under_twentyfive as "Under 25000 Change",
        (twe.twentyfive_to_seventyfive-zer.twentyfive_to_seventyfive)*100 /
                zer.twentyfive_to_seventyfive as "25000 to 75000 Change",
        (twe.seventyfive_to_onehundredfifty-zer.seventyfive_to_onehundredfifty)*100 /
                zer.seventyfive_to_onehundredfifty as "75000 to 150000 Change",
        (twe.onehundredfifty_plus-zer.onehundredfifty_plus)*100 /
                zer.onehundredfifty_plus as "150000 plus Change",
        zer.geom
from project.census_00 as zer
inner join project.census_20 as twe on zer.fips=twe.fips,
        project.rivers as r
where ST_Intersects(r.geom,zer.geom)
and r.name='Mississippi River'
and st_aslatlontext(st_centroid(zer.geom)) >=
        st_aslatlontext((select the_geom from world.cities where name='St. Louis'))
and st_aslatlontext(st_centroid(zer.geom)) <=
        st_aslatlontext((select the_geom from world.cities where name='Minneapolis'))
order by zer.state
```

Get change in age groups 1970 to 2020

```
select distinct(sev.name) as "County",
        sev.state as "State",
        (twe.under_eighteen-sev.under_eighteen)*100 /
                sev.under_eighteen as "Under Eighten Change",
        (twe.eighteen_to_fiftyfour-sev.eighteen_to_fiftyfour)*100 /
                sev.eighteen_to_fiftyfour as "Eighteen to Fiftyfour Change",
        (twe.fiftyfive_plus-sev.fiftyfive_plus)*100 /
                sev.fiftyfive_plus as "Fiftyfive plus Change",
        sev.geom
from project.census_70 as sev
inner join project.census_20 as twe on sev.fips=twe.fips,
        project.rivers as r
where ST_Intersects(r.geom,sev.geom)
and r.name='Mississippi River'
and st_aslatlontext(st_centroid(sev.geom)) >=
        st_aslatlontext((select the_geom from world.cities where name='St. Louis'))
and st_aslatlontext(st_centroid(sev.geom)) <=
        st_aslatlontext((select the_geom from world.cities where name='Minneapolis'))
order by sev.state
```

Get change in education levels 1970 to 2020

```
select distinct(sev.name) as "County",
        sev.state as "State",
        (twe.less_than_hs-sev.less_than_hs)*100 / sev.less_than_hs as "Less than HS",
        (twe.hs_grad-sev.hs_grad)*100 / sev.hs_grad as "HS Grad",
        (twe.some_college-sev.some_college)*100 / sev.some_college as "Some College",
        (twe.college_four_years_plus-sev.college_four_years_plus)*100 /
                sev.college_four_years_plus as "College Four Years Plus",
        sev.geom
from project.census_70 as sev
inner join project.census_20 as twe on sev.fips=twe.fips,
        project.rivers as r
where ST_Intersects(r.geom,sev.geom)
and r.name='Mississippi River'
and st_aslatlontext(st_centroid(sev.geom)) >=
        st_aslatlontext((select the_geom from world.cities where name='St. Louis'))
and st_aslatlontext(st_centroid(sev.geom)) <=
        st_aslatlontext((select the_geom from world.cities where name='Minneapolis'))
order by sev.state
```

Get difference between pop change and ed level change from 1970 to 2020

```
select distinct(sev.name) as "County",
        sev.state as "State",
        case when (twe.less_than_hs-sev.less_than_hs)*100 / sev.less_than_hs <
                (twe.population-sev.population)*100 /
                        sev.population then 0 else 1 end as "Less than HS Comp to Pop",
        case when (twe.hs_grad-sev.hs_grad)*100 / sev.hs_grad <
                (twe.population-sev.population)*100 /
                        sev.population then 0 else 1 end as "HS Grad Comp to Pop",
        case when (twe.some_college-sev.some_college)*100 / sev.some_college <
                (twe.population-sev.population)*100 /
                        sev.population then 0 else 1 end as "Some College Comp to Pop",
        case when (twe.college_four_years_plus-sev.college_four_years_plus)*100 /
                sev.college_four_years_plus <
                (twe.population-sev.population)*100 /
                sev.population then 0 else 1 end as "College Four Years Plus Comp to Pop",
        sev.geom
from project.census_70 as sev
inner join project.census_20 as twe on sev.fips=twe.fips,
        project.rivers as r
where ST_Intersects(r.geom,sev.geom)
and r.name='Mississippi River'
and st_aslatlontext(st_centroid(sev.geom)) >=
        st_aslatlontext((select the_geom from world.cities where name='St. Louis'))
```

```
and st_aslatlontext(st_centroid(sev.geom)) <=
        st_aslatlontext((select the_geom from world.cities where name='Minneapolis'))
order by sev.state
```

Get change in density from 1970 to 2020

```
select distinct(sev.name) as "County",
        sev.state as "State",
        sev.density as "Start Density",
        twe.density as "End Density",
        (twe.density-sev.density)*100/sev.density as "Density Change",
        sev.geom
from project.census_70 as sev
inner join project.census_20 as twe on sev.fips=twe.fips,
        project.rivers as r
where ST_Intersects(r.geom,sev.geom)
and r.name='Mississippi River'
and st_aslatlontext(st_centroid(sev.geom)) >=
        st_aslatlontext((select the_geom from world.cities where name='St. Louis'))
and st_aslatlontext(st_centroid(sev.geom)) <=
        st_aslatlontext((select the_geom from world.cities where name='Minneapolis'))
order by sev.state
```

Get change in labor participation rate and unemployment rate

```
select distinct(sev.name) as "County",
        sev.state as "State",
        sev.labor_participation_rate as "Seventy Labor Participation",
        sev.unemployment_rate as "Seventy Unemployment Rate",
        eig.labor_participation_rate as "Eighty Labor Participation",
        eig.unemployment_rate as "Eighty Unemployment Rate",
        (eig.labor_participation_rate-sev.labor_participation_rate)*100 /
                sev.labor_participation_rate as "LP Seventy to Eighty Decade Change",
        (eig.unemployment_rate-sev.unemployment_rate)*100 /
                sev.unemployment_rate as "UP Seventy to Eighty Decade Change",
        nin.labor_participation_rate as "Ninty Labor Participation",
        nin.unemployment_rate as "Ninty Unemployment Rate",
        (nin.labor_participation_rate-eig.labor_participation_rate)*100 /
                eig.labor_participation_rate as "LP Eighty to Ninty Decade Change",
        (nin.unemployment_rate-eig.unemployment_rate)*100 /
                eig.unemployment_rate as "UP Eighty to Ninty Decade Change",
        zer.labor_participation_rate as "Zero Labor Participation",
        zer.unemployment_rate as "Zero Unemployment Rate",
        (zer.labor_participation_rate-nin.labor_participation_rate)*100 /
                nin.labor_participation_rate as "LP Ninty to Zero Decade Change",
        (zer.unemployment_rate-nin.unemployment_rate)*100 /
```

```
        nin.unemployment_rate as "UP Ninty to Zero Decade Change",
    ten.labor_participation_rate as "Ten Labor Participation",
    ten.unemployment_rate as "Ten Unemployment Rate",
    (ten.labor_participation_rate-zer.labor_participation_rate)*100 /
        zer.labor_participation_rate as "LP Zero to Ten Decade Change",
    (ten.unemployment_rate-zer.unemployment_rate)*100 /
        zer.unemployment_rate as "UP Zero to Ten Decade Change",
    twe.labor_participation_rate as "Twenty Labor Participation",
    twe.unemployment_rate as "Twenty Unemployment Rate",
    (twe.labor_participation_rate-ten.labor_participation_rate)*100 /
        ten.labor_participation_rate as "LP Ten to Twenty Decade Change",
    (twe.unemployment_rate-ten.unemployment_rate)*100 /
        ten.unemployment_rate as "UP Ten to Twenty Decade Change",
    (twe.labor_participation_rate-sev.labor_participation_rate)*100 /
        sev.labor_participation_rate as "LP Seventy to Twenty Change",
    (twe.unemployment_rate-sev.unemployment_rate)*100 /
        sev.unemployment_rate as "UP Seventy to Twenty Change",
    sev.geom
from project.census_70 as sev
inner join project.census_80 as eig on sev.fips=eig.fips
inner join project.census_90 as nin on sev.fips=nin.fips
inner join project.census_00 as zer on sev.fips=zer.fips
inner join project.census_10 as ten on sev.fips=ten.fips
inner join project.census_20 as twe on sev.fips=twe.fips,
    project.rivers as r
where ST_Intersects(r.geom,sev.geom)
and r.name='Mississippi River'
and st_aslatlontext(st_centroid(sev.geom)) >=
    st_aslatlontext((select the_geom from world.cities where name='St. Louis'))
and st_aslatlontext(st_centroid(sev.geom)) <=
    st_aslatlontext((select the_geom from world.cities where name='Minneapolis'))
order by sev.state
```

Get change in households from 1970 to 2020

```
select distinct(sev.name) as "County",
    sev.state as "State",
    (twe.households-sev.households)*100 / sev.households as "Household Change",
    sev.geom
from project.census_70 as sev
inner join project.census_20 as twe on sev.fips=twe.fips,
    project.rivers as r
where ST_Intersects(r.geom,sev.geom)
and r.name='Mississippi River'
and st_aslatlontext(st_centroid(sev.geom)) >=
```

```
        st_aslatlontext((select the_geom from world.cities where name='St. Louis'))
and st_aslatlontext(st_centroid(sev.geom)) <=
        st_aslatlontext((select the_geom from world.cities where name='Minneapolis'))
order by sev.state
```

Get household change and household size change

```
select distinct(sev.name) as "County",
        sev.state as "State",
        (twe.population-sev.population)*100 / sev.population as "Population Change",
        (twe.households-sev.households)*100 / sev.households as "Households Change",
        cast(sev.population as decimal(10,2))/cast(sev.households as decimal(10,2))
                as "Seventy Household Size",
        cast(twe.population as decimal(10,2))/cast(twe.households as decimal(10,2))
                as "Twenty Household Size",
        (cast(twe.population as decimal(10,2)) /
                cast(twe.households as decimal(10,2)) -
                        cast(sev.population as decimal(10,2)) /
                        cast(sev.households as decimal(10,2)))*100 /
                        (sev.population/sev.households) as "Household Size Change",
        sev.geom
from project.census_70 as sev
inner join project.census_20 as twe on sev.fips=twe.fips,
        project.rivers as r
where ST_Intersects(r.geom,sev.geom)
and r.name='Mississippi River'
and st_aslatlontext(st_centroid(sev.geom)) >=
        st_aslatlontext((select the_geom from world.cities where name='St. Louis'))
and st_aslatlontext(st_centroid(sev.geom)) <=
        st_aslatlontext((select the_geom from world.cities where name='Minneapolis'))
order by sev.state
```

Get industry change

```
select distinct(eig.name) as "County",
        eig.state as "State",
        (twe.population-eig.population)*100/eig.population as "Population Change",
        (twe.construction-eig.construction)*100/eig.construction as "Construction Change",
        case when (twe.construction-eig.construction)*100/eig.construction <
                (twe.population-eig.population)*100/eig.population then 0 else 1 end as
"Construction to Pop",
        (twe.manufacturing-eig.manufacturing)*100/eig.manufacturing as "Manufacturing
Change",
        case when (twe.manufacturing-eig.manufacturing)*100/eig.manufacturing <
                (twe.population-eig.population)*100/eig.population then 0 else 1 end as
"Manufacturing to Pop",
```

```
	(twe.wholesale-eig.wholesale)*100/eig.wholesale as "Wholesale Change",
	case when (twe.wholesale-eig.wholesale)*100/eig.wholesale <
		(twe.population-eig.population)*100/eig.population then 0 else 1 end as
"Wholesale to Pop",
	(twe.retail-eig.retail)*100/eig.retail as "Retail Change",
	case when (twe.retail-eig.retail)*100/eig.retail <
		(twe.population-eig.population)*100/eig.population then 0 else 1 end as "Retail
to Pop",
	(twe.other_career-eig.other_career)*100/eig.other_career as "Other Career Change",
	case when (twe.other_career-eig.other_career)*100/eig.other_career <
		(twe.population-eig.population)*100/eig.population then 0 else 1 end as "Other
Career to Pop",
	eig.geom
from project.census_80 as eig
inner join project.census_20 as twe on eig.fips=twe.fips,
	project.rivers as r
where ST_Intersects(r.geom,eig.geom)
and r.name='Mississippi River'
and st_aslatlontext(st_centroid(eig.geom)) >=
	st_aslatlontext((select the_geom from world.cities where name='St. Louis'))
and st_aslatlontext(st_centroid(eig.geom)) <=
	st_aslatlontext((select the_geom from world.cities where name='Minneapolis'))
order by eig.state
```