

TUTORIAL 7.2 | HISTORICAL CENSUS DATA

Goals

- Download and use historical census data.
- Learn how to use the IPUMS platform.

Introduction

Using IPUMS, a database of census and other types of social data maintained by the University of Minnesota. IPUMS has census records dating back to 1790, in the US and in other countries. The records have been “harmonized” to correspond with contemporary shapefiles – typically the 2000 or 2010 shapefile for older records. The farther back you search, the less likely that you’ll find fine-grained data for an area, but IPUMS has census tract level data for some variables dating back decades, and even more extensive county level data

For this tutorial, we’ll be looking at historical commuter data – specifically, method of transportation to work. This data is available at the census tract level dating back to 1960. The data is available for each decade, but we’ll look at the change in 20 year increments: 1960, 80, 2000, and 2020. You’ll need a separate shapefile for each of these years, and since IPUMS data are only available for the entire country you’ll need to downsize the spreadsheets and shapefiles for each year as well. In the end you’ll have an extensive dataset of commuter change across the country, which you’ll reduce to only your city.

As with other census data, you’ll need to make sure that the GEOIDs for the spreadsheet and shapefile match and can be joined.

Step 1: Create a Free IPUMS NHGIS account.

The account and data are free for educational and research purposes.

<https://www.nhgis.org/>

Once you've created the account, login to the NHGIS website.

Step 2: Filter NHGIS tables by variable and geography.

You might not know at first which data table specifically you're looking for, or even what kinds of data are available for which years and geographical levels. I like to start by filtering by geographic level, and then by topic, and lastly by year. You'll see that each filter helpfully reduces the options for the next – for instance, once you select a geography and topic, only the years with that information available show up in the filter list.

2a Click “**Get Data**”, the green button on the NHGIS home page. Clicking “**Select Data**” at the top will also work.

The screenshot shows the IPUMS NHGIS homepage. At the top, there's a navigation bar with links for "WELCOME, MIRANDA | ACCOUNT | LOG OUT | IPUMS.ORG", "NHGIS", "GEOMARKER", and "NATIONAL HISTORICAL GIS". Below the navigation bar is a banner with five small images representing different data types: a city skyline, an aerial view of a town, autumn foliage, a person working in a field, and a historical building. Underneath the banner, there are two columns of links. The left column includes "IPUMS NHGIS", "ABOUT", "REGISTER", "DONATE TO NHGIS", "DATA", "BROWSE AND SELECT DATA", "DOWNLOAD OR REVISE MY DATA", "API", "SUPPLEMENTAL DATA", "GEOGRAPHIC CROSSWALKS", "ENVIRONMENTAL SUMMARIES", "PRIVACY-PROTECTED DEMO DATA", "SABINS SCHOOL AREAS", "ANNUAL TRACT DATA", and "DOCUMENTATION". The right column is titled "DOWNLOAD U.S. CENSUS DATA TABLES & MAPPING FILES" and contains text about the National Historical Geographic Information System (NHGIS) providing access to summary tables and time series of population, housing, agriculture, and economic data from 1790 through the present for all levels of U.S. census geography. It also features a "Read more" link. Below this text is a green button labeled "START HERE:" followed by a larger green button labeled "Get Data", which is highlighted with a red rectangular box. At the bottom of the page, there's a section titled "WHAT IS IPUMS?" with text explaining that IPUMS provides census and survey data from around the world integrated across time and space. It also mentions that IPUMS integration and documentation makes it easy to study change, conduct comparative research, and merge data.

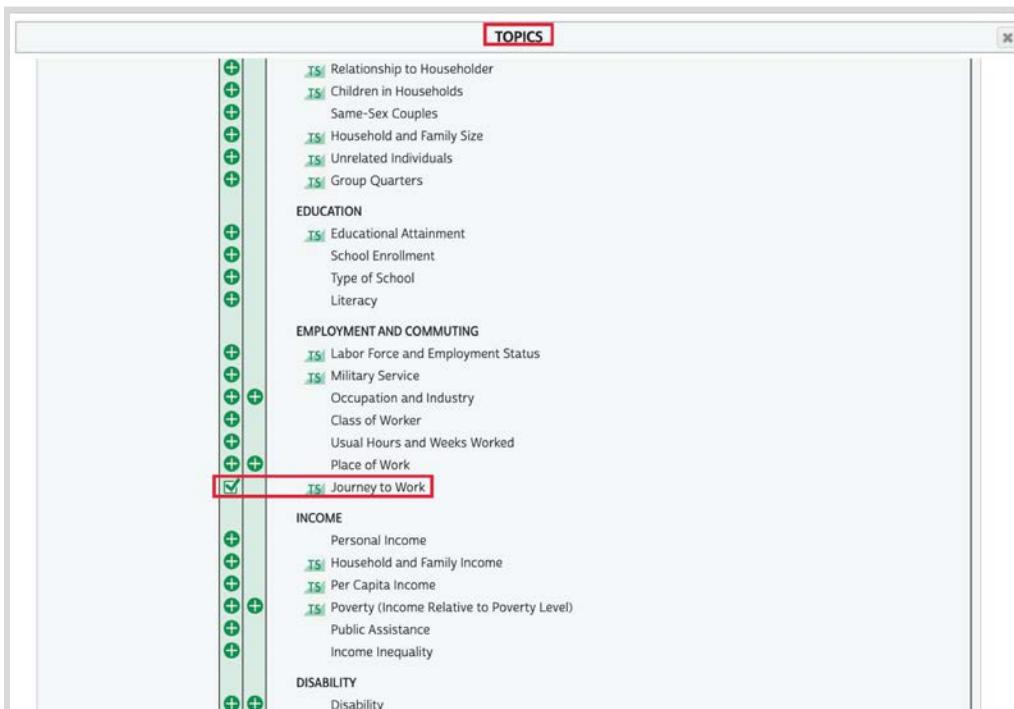
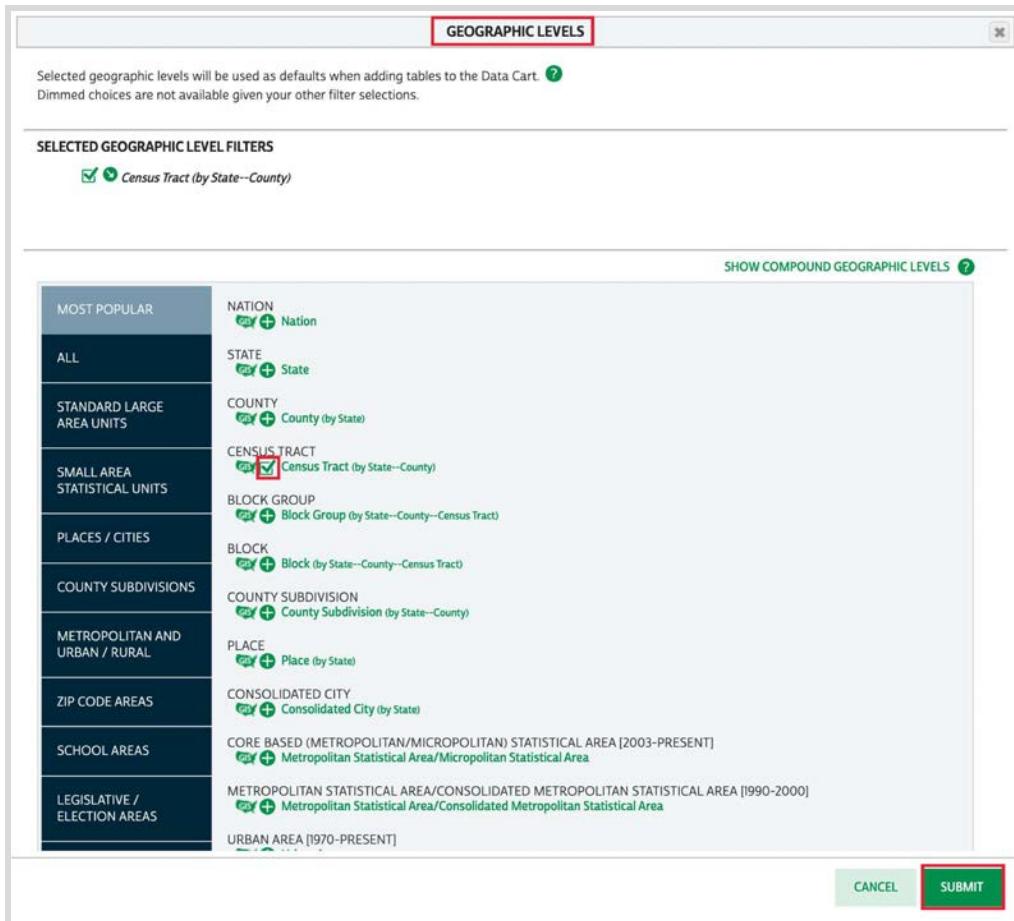
2b You will see the NHGIS search page. To show the tables we'll use, apply the following filters by clicking green plus symbol beside each filter:

Geography Level: census tract

Topic: Under “Population”, scroll down and select “Journey to Work”

Year: select 1960, 1980, 2000, and 2020

The screenshot shows the IPUMS NHGIS search page. At the top, there's a navigation bar with links for "WELCOME, MIRANDA | ACCOUNT | LOG OUT | IPUMS.ORG", "NHGIS", "GEOMARKER", and "NATIONAL HISTORICAL GIS". Below the navigation bar is a banner with five small images representing different data types: a city skyline, an aerial view of a town, autumn foliage, a person working in a field, and a historical building. Underneath the banner, there's a "FILTER + OPTIONS + REVIEW" button. To its right is a "HOW TO USE THE DATA FINDER" link. On the far right, there's a "DATA CART" section showing "0 SOURCE TABLES", "0 TIME SERIES TABLES", and "0 GIS FILES", with a "SHOW SELECTIONS" link. Below these sections is a "APPLY FILTERS" button with a question mark icon. Underneath it are four buttons: "GEOGRAPHIC LEVELS", "YEARS", "TOPICS", and "DATASETS", all of which are highlighted with red rectangular boxes. To the right of these buttons is a box containing the text "No filters selected."



YEARS

Dimmed choices are not available given your other filter selections.

DECENNIAL YEARS	NON-DECENNIAL YEARS	5-YEAR RANGES	
<input checked="" type="checkbox"/> 2020	<input type="checkbox"/> 2021	<input type="checkbox"/> 1955	<input type="checkbox"/> 1923
<input type="checkbox"/> 2010	<input type="checkbox"/> 2019	<input type="checkbox"/> 1986	<input type="checkbox"/> 1922
<input checked="" type="checkbox"/> 2000	<input type="checkbox"/> 2018	<input type="checkbox"/> 1985	<input type="checkbox"/> 1921
<input type="checkbox"/> 1990	<input type="checkbox"/> 2017	<input type="checkbox"/> 1984	<input type="checkbox"/> 1919
<input checked="" type="checkbox"/> 1980	<input type="checkbox"/> 2016	<input type="checkbox"/> 1983	<input type="checkbox"/> 1918
<input type="checkbox"/> 1970	<input type="checkbox"/> 2015	<input type="checkbox"/> 1982	<input type="checkbox"/> 1917
<input checked="" type="checkbox"/> 1960	<input type="checkbox"/> 2014	<input type="checkbox"/> 1981	<input type="checkbox"/> 1916
<input type="checkbox"/> 1950	<input type="checkbox"/> 2013	<input type="checkbox"/> 1980	<input type="checkbox"/> 1915
<input type="checkbox"/> 1940	<input type="checkbox"/> 2012	<input type="checkbox"/> 1979	<input type="checkbox"/> 1914
<input type="checkbox"/> 1930	<input type="checkbox"/> 2011	<input type="checkbox"/> 1978	<input type="checkbox"/> 1906
<input type="checkbox"/> 1920	<input type="checkbox"/> 2009	<input type="checkbox"/> 1976	<input type="checkbox"/> 1886
<input type="checkbox"/> 1910	<input type="checkbox"/> 2007	<input type="checkbox"/> 1975	<input type="checkbox"/> 1885
<input type="checkbox"/> 1900	<input type="checkbox"/> 2006	<input type="checkbox"/> 1974	<input type="checkbox"/> 1884
<input type="checkbox"/> 1890	<input type="checkbox"/> 2005	<input type="checkbox"/> 1973	<input type="checkbox"/> 1883
<input type="checkbox"/> 1880	<input type="checkbox"/> 2004	<input type="checkbox"/> 1972	<input type="checkbox"/> 1882
<input type="checkbox"/> 1870	<input type="checkbox"/> 2003	<input type="checkbox"/> 1971	<input type="checkbox"/> 1881
		<input type="checkbox"/> 1938	<input type="checkbox"/> 1880
		<input type="checkbox"/> 1879	<input type="checkbox"/> 1878
			<input type="checkbox"/> 2017-2021
			<input type="checkbox"/> 2016-2020
			<input type="checkbox"/> 2015-2019
			<input type="checkbox"/> 2014-2018
			<input type="checkbox"/> 2013-2017
			<input type="checkbox"/> 2012-2016
			<input type="checkbox"/> 2011-2015
			<input type="checkbox"/> 2010-2014
			<input type="checkbox"/> 2009-2013
			<input type="checkbox"/> 2008-2012
			<input type="checkbox"/> 2007-2011
			<input type="checkbox"/> 2006-2010
			<input type="checkbox"/> 2005-2009
			<input type="checkbox"/> 2011-2013
			<input type="checkbox"/> 2010-2012

APPLY FILTERS ? HOW TO USE THE DATA FINDER

GEOGRAPHIC LEVELS	<input checked="" type="checkbox"/>	TRACT
YEARS	<input checked="" type="checkbox"/> OR <input type="checkbox"/>	1960 or 1980 or 2000 or 2020
TOPICS	<input checked="" type="checkbox"/> INCLUDES <input type="checkbox"/>	Journey to Work
DATASETS		
RESET FILTERS		

SELECT DATA ?

86 SOURCE TABLES	5 TIME SERIES TABLES	11 GIS FILES		
PAGE 1 OF 5				
POPULARITY	TABLE NAME	UNIVERSE	CLASSIFICATIONS	YEAR - DATASET
	WT51 Working Persons 16 and Over by Means of Transportation to Work	Workers 16 Years and Over (Includes Armed Forces)	Means of Transportation to Work (B)	1960_STF1
	WT57 Employed Population by Means of Transportation to Work (from printed report)	Employed Persons	Means of Transportation to Work (B)	1960_STF1
	WT47 Means of Transportation to Work	Workers 16 Years and Over	Means of Transportation to Work (B)	1980_STF2
	WT41 Travel Time to Work	Workers 16 Years and Over Who Did Not Work at Home	Travel Time to Work (B)	1980_STF3
	WT43 Aggregate Travel Time to Work (in Minutes)	Workers 16 Years and Over Who Did Not Work at Home		1980_STF3
		Workers 16 Years and Over Using Car, Truck		

You should now see a list of around 86 source tables, 5 time series tables, and 11 GIS files returned.

Step 3: Select the tables and GIS files. Download the data.

You'll see that many source tables have similar names. We want only the data about worker commute – not travel time to work, not travel time by type of transportation. To check what kind of data a table has you can click on the Table Name to pull up a window with variables, units, available geographic levels, and some other useful information. Note: when in doubt about which tables to use, check the “popularity” column. The more popular table is usually the safer bet.

The screenshot shows a detailed view of a table from the 1960 Census. The table is titled "Working Persons 14 and Over by Means of Transportation to Work". It includes a list of variables: Railroad, Subway or elevated, Bus or streetcar, Other public means, Private auto or carpool, Walked, Worked at home (includes resident domestic), and Not reported. The measurement unit is persons. The dataset is labeled "1960_1PH" and describes the 1960 Census Population & Housing Data [Tracts: Major Cities & Surrounds]. The note states that this group provides population and housing data from the 1960 Census for several geographic levels, including census tracts where available. It also notes that updates in either the ICPSR or NHGIS data may have produced incomplete correspondence. Tract data are drawn from 2 sources (resulting in some redundancy): an electronic data file produced by the Census Bureau in 1971 and the 1960 Elizabeth Muller Bogue File (ICPSR Study 2932), which was transcribed from printed Census Reports.

3a Click the check mark beside the tables for each decade which relate to work transportation method. You'll see that the variable name changes:

Working Persons 14 and Over by Means of Transportation to Work (1960)

Means of Transportation to Work (1980 and 2000 and 2020)

Note: the tables are automatically sorted in date order, so the 2020 tables will follow several pages of 2000 results.

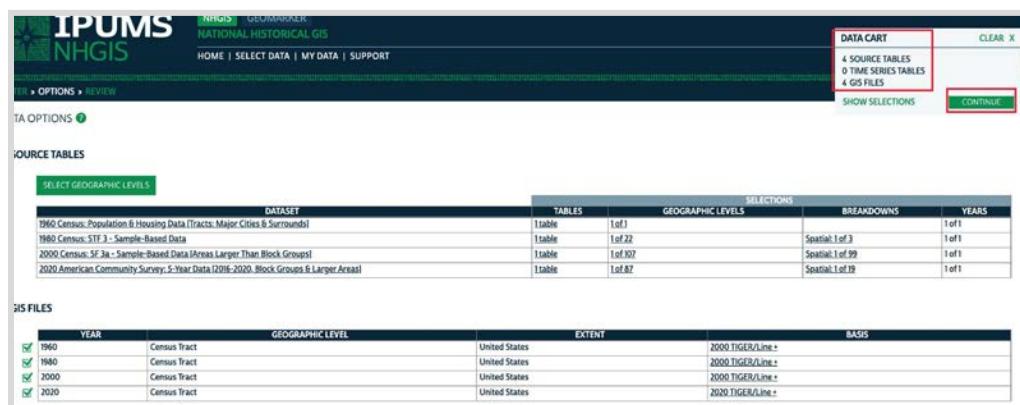
POPULARITY	TABLE NAME	UNIVERSE	CLASSIFICATIONS	YEAR + DATASET	BREAKDOWNS
<input checked="" type="checkbox"/>	WPS1. Working Persons 14 and Over by Means of Transportation to Work	Workers 14 Years and Over (Includes Armed Forces)	Means of Transportation to Work (0)	1960_1PH	
<input checked="" type="checkbox"/>	EMPLOYED. Employed Population by Means of Transportation to Work from printed report!	Employed Persons	Means of Transportation to Work (0)	1960_1PH	
<input checked="" type="checkbox"/>	MOT. Means of Transportation to Work	Workers 16 Years and Over	Means of Transportation to Work (0)	1980_SF3	Spatial
<input checked="" type="checkbox"/>	TTW. Travel Time to Work	Workers 16 Years and Over Who Did Not Work at Home	TravelTime to Work (0)	1980_SF3	Spatial
<input checked="" type="checkbox"/>	APPTRAVEL. Aggregate Travel Time to Work (in Minutes)	Workers 16 Years and Over Who Did Not Work at Home		1980_SF3	Spatial
<input checked="" type="checkbox"/>	PVO. Private Vehicle Occupancy	Workers 16 Years and Over Using Car, Truck or Van as a Means of Transportation to Work	Means of Transportation to Work (0)	1980_SF3	Spatial
<input checked="" type="checkbox"/>	MOTSF3. Means of Transportation to Work	Workers 16 Years and Over	Means of Transportation to Work (0)	2000_SF3a	Spatial
<input checked="" type="checkbox"/>	MOTSF3A. Means of Transportation to Work	Workers 16 Years and Over Using Car, Truck or Van as a Means of Transportation to Work	Means of Transportation to Work (0)	2000_SF3a	Spatial

POPULARITY	TABLE NAME	UNIVERSE	CLASSIFICATIONS	YEAR + DATASET	BREAKDOWNS
<input checked="" type="checkbox"/>	APPTRAVEL. Aggregate Travel Time to Work by Travel Time to Work by Means of Transportation to Work	Workers 16 Years and Over Who Did Not Work at Home	Means of Transportation to Work (0), TravelTime to Work (0)	2000_SF4	Race/Ethnicity, Spatial
<input checked="" type="checkbox"/>	WORKERS. Workers 16 Years and Over Who Did Not Work at Home by Time Leaving Home to Go to Work	Workers 16 Years and Over Who Did Not Work at Home	Time of Departure to Go to Work (0)	2000_SF4	Race/Ethnicity, Spatial
<input checked="" type="checkbox"/>	WORKERSF4. Workers 16 Years and Over Who Carpool to Work by Carpool Size	Workers 16 Years and Over Who Carpool to Work	Means of Transportation to Work (0)	2000_SF4	Race/Ethnicity, Spatial
<input checked="" type="checkbox"/>	TOTAL. Total Workers 16 Years and Over Who Did Not Commute to Work with a Car, Van, or Truck	Workers 16 Years and Over Commuting to Work by Other Means than Car, Truck, or Van as a Means of Transportation to Work	Means of Transportation to Work (0), TravelTime to Work (0)	2000_SF4	Race/Ethnicity, Spatial
<input checked="" type="checkbox"/>	IMPUTATION. Workers 16 Years and Over Using Car, Truck or Van as a Means of Transportation to Work by Imputation of Private Vehicle Occupancy	Workers 16 Years and Over Using Car, Truck or Van as a Means of Transportation to Work	Imputation/Allocation (0)	2000_SF4	Race/Ethnicity, Spatial
<input checked="" type="checkbox"/>	NOHOME. Workers 16 Years and Over Who Did Not Work at Home by Imputation of Time Leaving Home to Go to Work	Workers 16 Years and Over Who Did Not Work at Home	Imputation/Allocation (0)	2000_SF4	Race/Ethnicity, Spatial
<input checked="" type="checkbox"/>	NOHOMEF4. Workers 16 Years and Over Who Did Not Work at Home by Imputation of Travel Time to Work	Workers 16 Years and Over Who Did Not Work at Home	Imputation/Allocation (0)	2000_SF4	Race/Ethnicity, Spatial
<input checked="" type="checkbox"/>	MOTSF4. Means of Transportation to Work by Travel Time to Work	Workers 16 years and over who did not work from home	Means of Transportation to Work (0), TravelTime to Work (0)	2008_2020_ACSSa	Spatial
<input checked="" type="checkbox"/>	APPTRAVEL. Aggregate Travel Time to Work (in Minutes) of Workers by Travel Time to Work	Workers 16 years and over who did not work from home	TravelTime to Work (0)	2008_2020_ACSSa	Spatial
<input checked="" type="checkbox"/>	APPTRAVELF4. Aggregate Travel Time to Work (in Minutes) of Workers by Means of Transportation to Work	Workers 16 years and over who did not work from home	Means of Transportation to Work (0)	2008_2020_ACSSa	Spatial
<input checked="" type="checkbox"/>	MOTSF4. Means of Transportation to Work	Workers 16 years and over	Means of Transportation to Work (0)	2008_2020_ACSSa	Spatial
<input checked="" type="checkbox"/>	TIMEDEPARTURE. Time of Departure to Go to Work	Workers 16 years and over who did not work from home	Time of Departure to Go to Work (0)	2008_2020_ACSSa	Spatial

3b Next, go to the “GIS Files” tab. Select the files with the “2000 TIGER File” Basis. Select one for each of the 4 decades: 1960, 80, 2000, and 2020.

SELECT DATA 				
66 SOURCE TABLES		5 TIME SERIES TABLES		11 GIS FILES
POPULARITY	YEAR	GEOGRAPHIC LEVEL	EXTENT	BASIS
<input checked="" type="checkbox"/>	1960	Census Tract	United States	2000 TIGER/Line+
<input checked="" type="checkbox"/>	1960	Census Tract	United States	2008 TIGER/Line+
<input checked="" type="checkbox"/>	1980	Census Tract	United States	2000 TIGER/Line+
<input checked="" type="checkbox"/>	1980	Census Tract	United States	2008 TIGER/Line+
<input checked="" type="checkbox"/>	2000	Census Tract	United States	2000 TIGER/Line+
<input checked="" type="checkbox"/>	2000	Census Tract	United States	2008 TIGER/Line+
<input checked="" type="checkbox"/>	2000	Census Tract	United States	2009 TIGER/Line+
<input checked="" type="checkbox"/>	2000	Census Tract (Centers of Population)	United States	2000 Census Centers of Population
<input checked="" type="checkbox"/>	2020	Census Tract	United States	2020 TIGER/Line+
<input checked="" type="checkbox"/>	2020	Census Tract (Centers of Population)	United States	2020 Census Centers of Population

3c Your “Data Cart” in the upper righthand corner should show 4 Source Tables and 4 GIS Files. Click “Continue”, review the files, and click “Continue” again. On the next page, click “Submit”.

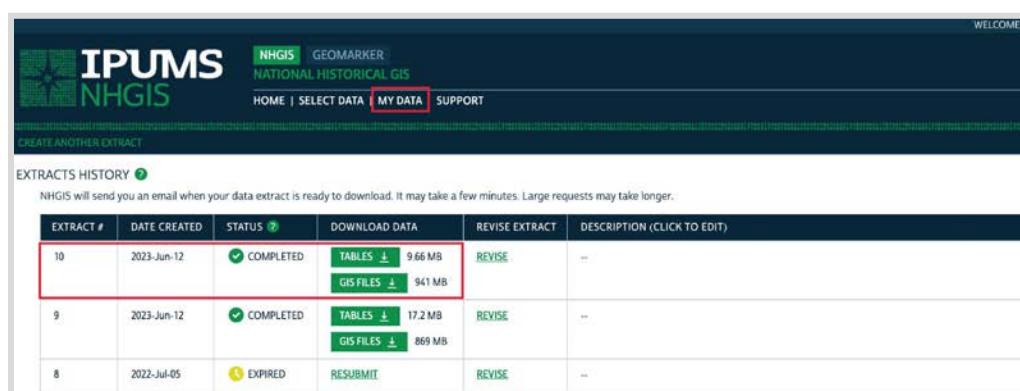


The screenshot shows the IPUMS NHGIS interface. In the top right, a "DATA CART" box displays "4 SOURCE TABLES", "4 TIME SERIES TABLES", and "4 GIS FILES". Below this, a "CONTINUE" button is highlighted with a red box. The main area shows two tables: "SOURCE TABLES" and "GIS FILES", both listing datasets for the years 1960, 1980, 2000, and 2020.

SOURCE TABLES				
SELECT GEOGRAPHIC LEVELS				
DATASET				
1960 Census: Population & Housing Data [Tracts: Major Cities & Surrounds]	Table	1st1	Spatial:1st3	1st1
1980 Census: STF 3 - Sample-Based Data	Table	1st22	Spatial:1st99	1st1
2000 Census: SF Ja - Sample-Based Data [Areas Larger than Block Groups]	Table	1st102	Spatial:1st19	1st1
2020 American Community Survey: 5-Year Data [2016-2020, Block Groups & Larger Areas]	Table	1st87		

GIS FILES				
YEAR	GEOGRAPHIC LEVEL	EXTENT	BASIS	
1960	Census tract	United States	2000 TIGER/Line+	
1980	Census tract	United States	2008 TIGER/Line+	
2000	Census tract	United States	2000 TIGER/Line+	
2020	Census tract	United States	2020 TIGER/Line+	

3d You'll be taken to the “Extracts History” page. Once the “Status” column shows “Completed”, download the tables and GIS files. Processing your request should take a few minutes.



The screenshot shows the "EXTRACTS HISTORY" page. It lists three extracts. Extract #10, created on 2023-Jun-12, is marked as "COMPLETED" with status icons. It includes download links for "TABLES" (9.66 MB) and "GIS FILES" (941 MB). Extract #9 and #8 are also listed, with Extract #8 showing an "EXPIRED" status and a "RESUBMIT" button.

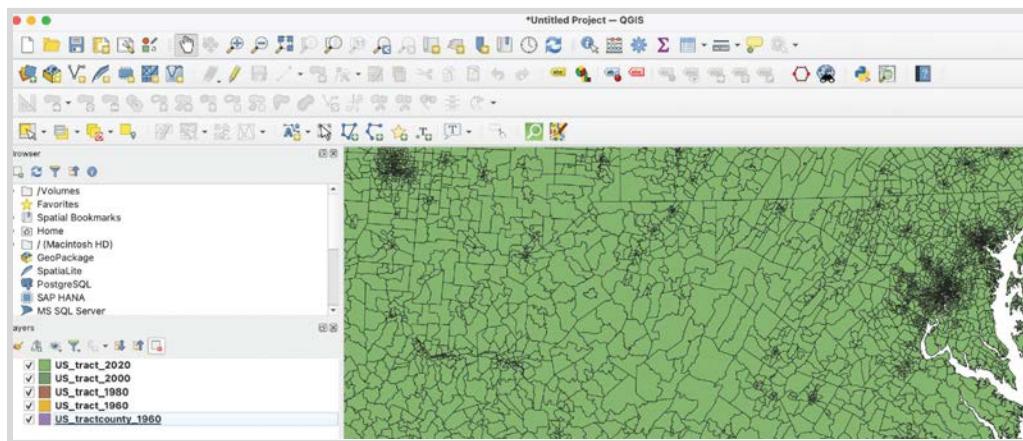
EXTRACT #	DATE CREATED	STATUS 	DOWNLOAD DATA	REVISE EXTRACT	DESCRIPTION (CLICK TO EDIT)
10	2023-Jun-12	 COMPLETED	 9.66 MB  941 MB		...
9	2023-Jun-12	 COMPLETED	 17.2 MB  869 MB		...
8	2022-Jul-05	 EXPIRED			...

Step 4: Import and join in QGIS.

4a First, open a new QGIS file. Set the CRS to “UTM 17N”.

4b Unzip and Import your TIGER shapefiles.

NOTE: the census tracts are divided into different geographical categories depending on the decade. You can import and check each file. You'll see that in 1960, Roanoke and Lynchburg were not divided into census tracts, but by 1980 they were. For each decade, import the GIS file which contains your city (tract or county).



4c Open the Attribute Table for the 1960 shapefile and scroll to the “GIS Join” columns. Open the 1960 csv file you downloaded and check that the GIS Join column and the data table’s GIS Join columns match. Sort both columns ascending to check.

	NHGIST	NHGISCTY	GISJOIN	SHAPE_AREA	SHAPE_LEN		YEAR	STATE	STATEA	COUNTY	COUNTYA	MSA	PLACE	PRETRAC
1	010	0550	G010055000001	010055000001	1328844.02...	5594.88525...	1960	Alabama	1	Etowah		55	2860	28600
2	010	0550	G010055000002	010055000002	10143051.58...	18021.17581...	1960	Alabama	1	Etowah		55	2860	28600
3	010	0550	G010055000003	010055000003	252999.29...	10765.60674...	1960	Alabama	1	Etowah		55	2860	28600
4	010	0550	G010055000004	010055000004	12378056.12...	21315.98760...	1960	Alabama	1	Etowah		55	2860	28600
5	010	0550	G010055000005	010055000005	6461961.16...	13470.81060...	1960	Alabama	1	Etowah		55	2860	28600
6	010	0550	G010055000006	010055000006	9688220.71...	25130.01289...	1960	Alabama	1	Etowah		55	2860	28600
7	010	0550	G010055000007	010055000007	1372800.45...	5104.83646...	1960	Alabama	1	Etowah		55	2860	28600
8	010	0550	G010055000008	010055000008	1863291.75...	6093.76388...	1960	Alabama	1	Etowah		55	2860	28600
9	010	0550	G010055000009	010055000009	7212531.67...	14761.35280...	1960	Alabama	1	Etowah		55	2860	28600

NOTE: if your city is included in the “tractcounty” instead of the “tract” file, you'll need to download the “county” csv for that year instead. For instance, Roanoke does not appear in the 1960 tract shapefile but does appear in the tractcounty file, so for that year I need to go back to the NHGIS website and download the “county” level geography for the 1960 “Working Persons 14 and Over by Means of Transportation to Work” table.

4d Reduce each csv table before importing it to QGIS.

First, sort the table by “State” and copy / paste the Virginia rows into a new spreadsheet. Copy / paste the header column from the first spreadsheet as well. Note: from 1980 onward you can further reduce the spreadsheet to only your county, using the following codes in the relevant county column, “CountyA”, or “CountyFP”, or “NHGISCity”:

Roanoke: 770
 Richmond: 760
 Lynchburg: 680

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
48622	G51076000705	1980	S	3	5	Virginia	51	6760	Richmond City	760	5	1035	705	
48623	G51076000706	1980	S	3	5	Virginia	51	6760	Richmond City	760	5	1035	706	
48624	G51076000707	1980	S	3	5	Virginia	51	6760	Richmond City	760	5	1035	707	
48625	G51076000708	1980	S	3	5	Virginia	51	6760	Richmond City	760	5	1035	708	
48626	G51076000709	1980	S	3	5	Virginia	51	6760	Richmond City	760	5	1035	709	
48627	G51077000001	1980	S	3	5	Virginia	51	6800	Roanoke City	770	5	1045	1	
48628	G51077000002	1980	S	3	5	Virginia	51	6800	Roanoke City	770	5	1045	2	
48629	G51077000003	1980	S	3	5	Virginia	51	6800	Roanoke City	770	5	1045	3	
48630	G51077000004	1980	S	3	5	Virginia	51	6800	Roanoke City	770	5	1045	4	
48631	G51077000005	1980	S	3	5	Virginia	51	6800	Roanoke City	770	5	1045	5	
48632	G51077000006	1980	S	3	5	Virginia	51	6800	Roanoke City	770	5	1045	6	
48633	G51077000007	1980	S	3	5	Virginia	51	6800	Roanoke City	770	5	1045	7	
48634	G51077000008	1980	S	3	5	Virginia	51	6800	Roanoke City	770	5	1045	8	
48635	G51077000009	1980	S	3	5	Virginia	51	6800	Roanoke City	770	5	1045	9	
48636	G51077000010	1980	S	3	5	Virginia	51	6800	Roanoke City	770	5	1045	10	
48637	G51077000011	1980	S	3	5	Virginia	51	6800	Roanoke City	770	5	1045	11	
48638	G51077000012	1980	S	3	5	Virginia	51	6800	Roanoke City	770	5	1045	12	
48639	G51077000013	1980	S	3	5	Virginia	51	6800	Roanoke City	770	5	1045	13	
48640	G51077000014	1980	S	3	5	Virginia	51	6800	Roanoke City	770	5	1045	14	
48641	G51077000015	1980	S	3	5	Virginia	51	6800	Roanoke City	770	5	1045	15	
48642	G51077000016	1980	S	3	5	Virginia	51	6800	Roanoke City	770	5	1045	16	
48643	G51077000017	1980	S	3	5	Virginia	51	6800	Roanoke City	770	5	1045	17	
48644	G51077000018	1980	S	3	5	Virginia	51	6800	Roanoke City	770	5	1045	18	
48645	G51077000019	1980	S	3	5	Virginia	51	6800	Roanoke City	770	5	1045	19	
48646	G51077000020	1980	S	3	5	Virginia	51	6800	Roanoke City	770	5	1045	20	
48647	G51077000021	1980	S	3	5	Virginia	51	6800	Roanoke City	770	5	1045	21	
48648	G51077000022	1980	S	3	5	Virginia	51	6800	Roanoke City	770	5	1045	22	
48649	G51077000023	1980	S	3	5	Virginia	51	6800	Roanoke City	770	5	1045	23	
48650	G51077500101	1980	S	3	5	Virginia	51	6800	Salem City	775	5	1080	101	
48651	G51077500102	1980	S	3	5	Virginia	51	6800	Salem City	775	5	1080	102	
48652	G51077500103	1980	S	3	5	Virginia	51	6800	Salem City	775	5	1080	103	
48653	G51077500104	1980	S	3	5	Virginia	51	6800	Salem City	775	5	1080	104	

Second, in the new reduced spreadsheet, delete every column except the GISJoin column and the columns with census codes (B9G001, B9G002, DHD001, etc.). For each table, refer to the accompanying “codebook” text file which contains the full name of each data column in the table.

nghis0010_ds19_1980	Table 1: Means of Transportation to Work Workers 16 Years and Over Source code: NT8E	NAME_E1	Area Name
48646	Railroad	AMKE001	Total
48647	Subway or elevated	AMKE002	Car, truck, or van
48648	Bus or streetcar	AMKE003	Car, truck, or van: Drove alone
48649	Other public means	AMKE004	Car, truck, or van: Carpool
48650	Private auto or carpool	AMKE005	Car, truck, or van: Carpool: 2-person carpool
48651	Walked	AMKE006	Car, truck, or van: Carpool: 3-person carpool
48652	Worked at home (incl.)	AMKE007	Car, truck, or van: Carpool: 4-6 person carpool
48653	Not reported	AMKE008	Car, truck, or van: Carpool: 5- or 6-person carpool
48654	All persons are granted a limited license to use the accompanying data, subject to the full accompanying data, subject to the following conditions:	AMKE009	Public transportation
48655	All persons are granted a limited license to use the accompanying data, subject to the full accompanying data, subject to the following conditions:	AMKE010	Bus
48656	All persons are granted a limited license to use the accompanying data, subject to the full accompanying data, subject to the following conditions:	AMKE011	Public transportation (excluding bus)
48657	All persons are granted a limited license to use the accompanying data, subject to the full accompanying data, subject to the following conditions:	AMKE012	Subway or elevated
48658	All persons are granted a limited license to use the accompanying data, subject to the full accompanying data, subject to the following conditions:	AMKE013	Streetcar
48659	All persons are granted a limited license to use the accompanying data, subject to the full accompanying data, subject to the following conditions:	AMKE014	Long-distance
48660	All persons are granted a limited license to use the accompanying data, subject to the full accompanying data, subject to the following conditions:	AMKE015	Light rail
48661	All persons are granted a limited license to use the accompanying data, subject to the full accompanying data, subject to the following conditions:	AMKE016	Taxicab
48662	All persons are granted a limited license to use the accompanying data, subject to the full accompanying data, subject to the following conditions:	AMKE017	Motorcycle
48663	All persons are granted a limited license to use the accompanying data, subject to the full accompanying data, subject to the following conditions:	AMKE018	Bicycle
48664	All persons are granted a limited license to use the accompanying data, subject to the full accompanying data, subject to the following conditions:	AMKE019	Walked
48665	All persons are granted a limited license to use the accompanying data, subject to the full accompanying data, subject to the following conditions:	AMKE020	Other means

Third, rename the census columns according to the codebook file. Note that, though the exact name of transportation methods changes between the decades, you should use the same name across all years: for instance, “Walk”, “Drive”, etc. If these names match between all 4 tables you’ll be able to more easily share styles between them in QGIS. In some cases, you may want to combine multiple columns using the “SUM” function (make sure not to remove columns, since you will need the full count for a “total” column).. I recommend simplifying the columns to “Walk”, “Drive”, “Public” (public transit), “Other” (including bike), “Home” (work from home), and “None” (eg no response). Note that in the codebook text file, the column descriptions indicate categories, and after colon a

subcolumn. So, for the 2020 file, “Car, truck or van: Drove alone” is a subset of “Car, truck, or van”. When totalling these columns, don’t include the subcolumns or you’ll double count.

1	GISJOIN	DHD001	DHD002	DHD003	DHD004	DHD005	DHD006
2	G510770000001	1359	423	126	29	7	8
3	G510770000002	1207	370	193	15	9	6
4	G510770000003	1649	365	100	74	5	23
						46	28
						33	27
						17	7
						17	5
						7	10
						0	14
						19	26
						5	0
						19	48
						19	19
						23	15
						32	28
						47	39
						13	22
						16	19
						34	65
						11	24
22	G510770000021	1060	145	7	13	8	9
23	G510770000022	1190	335	10	28	8	9

A	B	C	D	E	F
Table 1					
GISJOIN	Drive	Car	Carpool	Public	Walk
G510770000001	SUM	C2:D2	1359	423	126
G510770000002	1577	1207	370	193	15
G510770000003	2014	1649	365	100	74
G510770000004	2509	2097	412	110	24
G510770000005	2100	1645	455	158	70

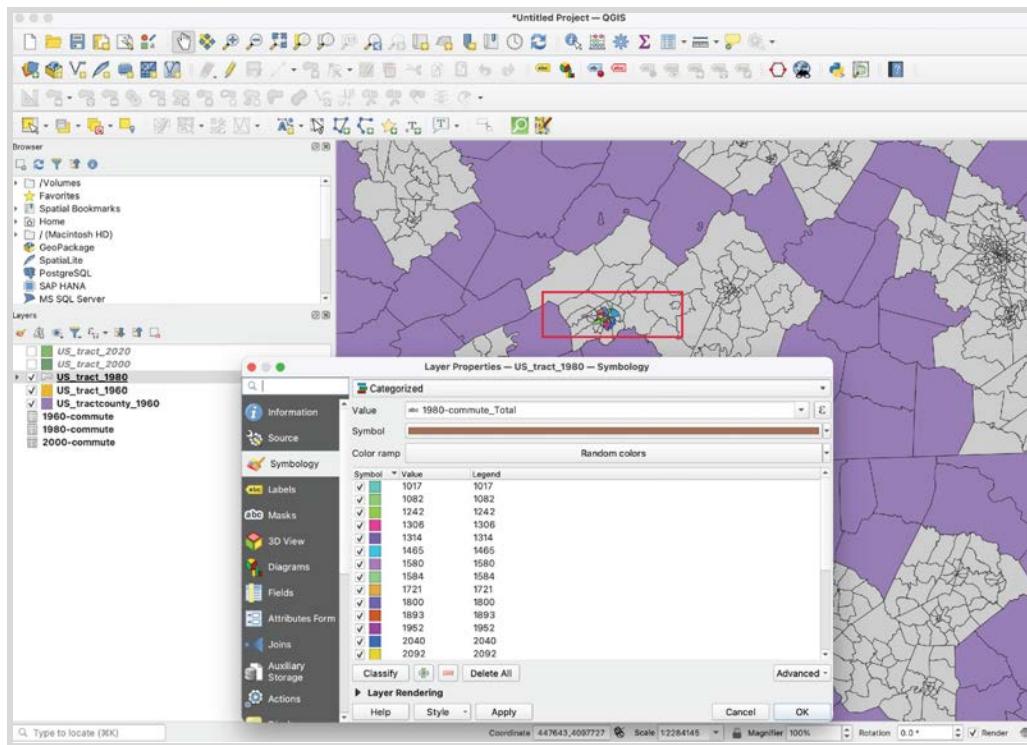
Lastly, add a column which totals the other census columns. Fill out this column with the function “=SUM(*cell range*)”, without the quotation marks. For instance, =SUM(C2:F2). Pull the function down the entire column by clicking and dragging from the bottom.

GISJOIN	Total	Drive	Public	Walk	Other	Home
G51077000000100	1632	1459	108	43	0	22
G5107700000200	1708	1526	69	36	60	17
G5107700000300	2612	2410	56	53	49	44
G5107700000400	2076	1959	31	22	46	18
G5107700000500	2178	2002	35	64	42	35

Save the new file as a csv. Repeat for the remaining 3 tables.

4d Once all the data table and shapefile GEOFID columns match for their respective years (1960 to 1960, 1980 to 1980, and so on), import the 4 data tables to QGIS.

4e Join each data table to its respective shapefile. Check the Attribute Tables of the shapefiles to make sure the new columns joined correctly and are not Null – but note that **only the tracts around your city will have data**, since you deleted everything else. You can also check by styling the data.



Step 5: Reduce the shapefiles to your city of interest.

5a Open the Attribute table of the 1960 shapefile. Click “Select by Expression” from the upper toolbar. Select “NHGISCITY” and set it equal to your city’s code with an additional 0 at the end:

Roanoke: 7700

Richmond: 7600

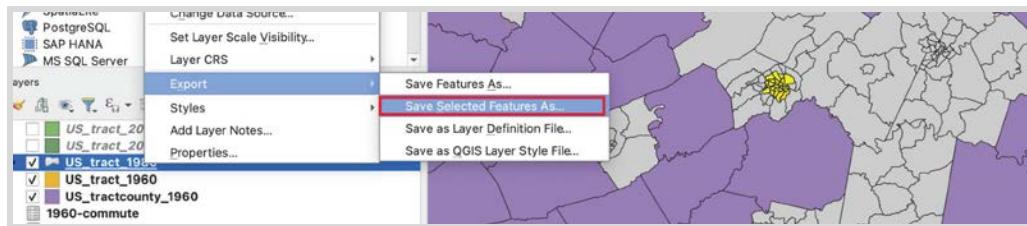
Lynchburg: 6800

US_tract_1980 — Features Total: 42917, Filtered: 42917, Selected: 0									
	NHGISST	NHGISCITY	Select features using an expression		SHAPE_AREA	SHAPE_LEN	I80-commute_Tot	I80-commute_Driv	80-commute
1	470	0010	G470001002...	47000100210	132774973.1...	60265.0224...	NULL	NULL	NULL
2	470	0010	G470001002...	47000100202	36194661.17...	26550.6686...	NULL	NULL	NULL
3	470	0010	G470001002...	47000100208	36876193.0...	36369.8706...	NULL	NULL	NULL

The screenshot shows the 'Select by Expression' dialog in ArcGIS Pro. The expression is set to `\"NHGISCITY\" = '7700'`. The expression builder highlights the field name `NHGISCITY` in red. The map view shows a single purple polygon selected.



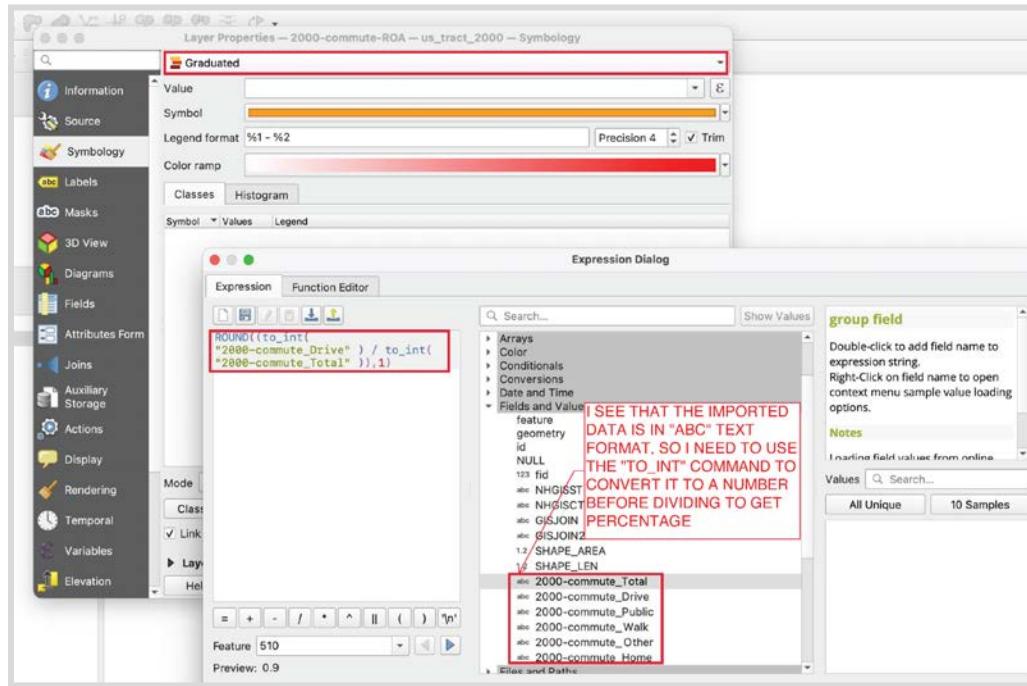
5b If you see the correct selection in your map, right click on the Shapefile layer, go to “Export”, and then “Save Selected Features As”. The new reduced shapefile will be added to your map.



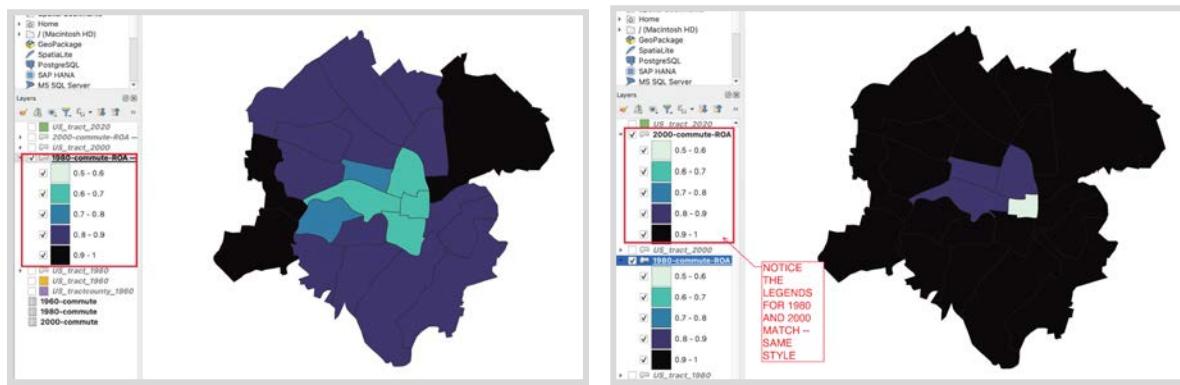
5c Repeat these steps for the other three shapefiles.

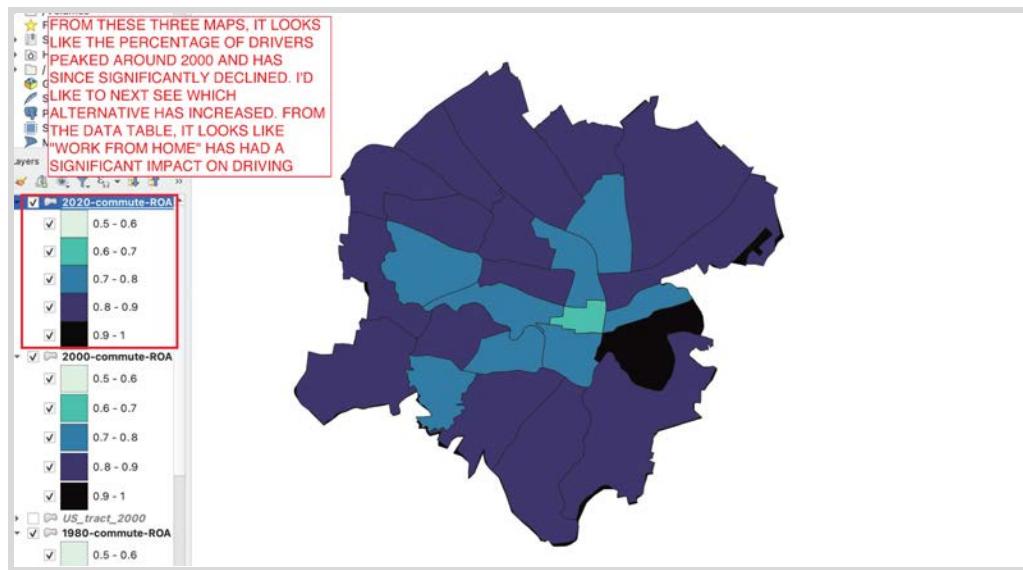
Step 6: Create a Print Layout to compare the four decades.

6a Using the “Total” and “Drive” columns, show the percentage of Car commuters for each decade. You may need to use the “to_int” function if your data is text instead of numbers.



6b Be sure to use the same color scheme for each decade. Use the “Save Style As” and “Load Style” functions.





6c In the print layout, include one map for each decade. Include some observations about trends. As always, include a title, legend, scale, north arrow, map labels, your name, and sources.