

TUTORIAL 5 | FEATURE OVERLAP ANALYSIS

Goals

- Become familiar with ACS data.
- Learn how to derive an average from census data.
- Consider the strengths and weaknesses of census data.
- Compare two different census categories geographically.
- Analyze demographic patterns through maps.

Introduction

In this tutorial, we'll be looking at a different kind of census data called the ACS (American Community Survey). After the 2000 census, the ACS began to replace the long-form decennial census. Now, every 10 years we answer only the short-form census; to supplement that data the ACS polls about 1% of the US population each year (around 3 million people). The ACS asks questions especially interesting to architects, such as age and type of housing, rental costs, occupants per room, mortgage costs, HVAC characteristics, and even building material. You can see the information ACS collects here: <https://www.census.gov/programs-surveys/acs/guidance/handbooks/general.html>

Table excerpt from Chapter 1 below:

Table 1.1. Population and Housing Data Included in American Community Survey Data Products		
Social Characteristics	Economic Characteristics	Plumbing Facilities ⁶
Ancestry	Class of Worker	Rent
Citizenship Status	Commuting (Journey to Work)	Rooms/Bedrooms
Citizen Voting-Age Population	Employment Status	Selected Monthly Owner Costs
Disability Status ¹	Food Stamps/Supplemental Nutrition Assistance Program (SNAP) ⁴	Telephone Service Available
Educational Attainment	Health Insurance Coverage ²	Tenure (Owner/Renter)
Fertility	Income and Earnings	Units in Structure
Grandparents as Caregivers	Industry and Occupation	Value of Home
Language Spoken at Home	Place of Work	Vehicles Available
Marital History ²	Poverty Status	Year Householder Moved Into Unit
Marital Status	Work Status Last Year	Year Structure Built
Migration/Residence 1 Year Ago		
Period of Military Service		
Place of Birth	Housing Characteristics	Demographic Characteristics
School Enrollment	Computer and Internet Use ⁵	Age and Sex
Undergraduate Field of Degree ³	House Heating Fuel	Group Quarters Population
Veteran Status ²	Kitchen Facilities	Hispanic or Latino Origin
Year of Entry	Occupancy/Vacancy Status	Race
	Occupants Per Room	Relationship to Householder
		Total Population

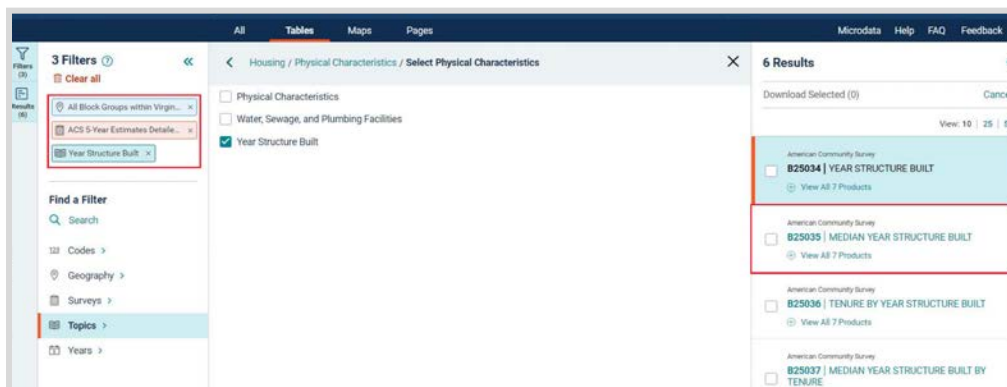
¹ Questions on Disability Status were significantly revised in the 2008 survey to cause a break in series.
² Marital History, Veterans' Service-Connected Disability Status and Ratings, and Health Insurance Coverage were added in the 2008 survey.
³ Undergraduate Field of Degree was added in the 2009 survey.
⁴ Food Stamp Benefit amount was removed in 2008.
⁵ Computer and Internet Use was added to the 2013 survey.
⁶ One of the components of Plumbing Facilities, flush toilet, and Business or Medical Office on Property questions were removed in 2016.
Source: U.S. Census Bureau.

Note: because of the small sample size, ACS data are aggregated in 5-year chunks (the 5-year estimates) for non-metropolitan areas. Because we'll be looking at urban and rural areas, and because the 5-year data are more reliable, we'll be using those. However, understand that the 5-year estimate, say, for 2019 refers to data collected from 2015-2019 – and not only 2019. Keep this in mind when interpreting and labeling the data.

Step 1: Download and process ACS census data

1a Go back to the Census Fact Finder website: <https://data.census.gov/cedsci/>

Filter for **your city's state** at the **Block Group** level. Then add a filter by survey type "**ACS, 5-year**". Lastly, filter by Topic, "Housing" and select "Physical Characteristics" and "**Median Year Built**". Select **B25035**. **Download** the file for 2019. Lastly, filter the topic by "**Race and Ethnicity**", and select Table **B02001**. Download those tables for **2019**.



1b Process your new data tables.

Like last time, we'll need the shortened GEOIDs to match the 2010 GEOIDs of your shapefile. Refer to Tutorial 4 for a reminder of how to do this.

Next, **delete the "Margin of Error" and "Name" columns**. The Margin of Error columns are important to understand (you'll notice they're huge for rural areas), but we're going to ignore them for the purposes of this map. Like last time, **rename** your columns logically (**without spaces or special characters**) in the first row and delete the second row. **Download** the cleaned file as a .csv.

	A	B	C	D	E	F	G	H	I	J	K
1	B25038_001E	B25038_001M	B25038_002E	B25038_002M	B25038_003E	B25038_003M	B25038_004E	B25038_004M	B25038_005E	B25038_005M	B25038_006E
2	EstimateTotal	Margin of Error	EstimateTotal	Margin of Error	EstimateTotal	Margin of Error	EstimateTotal	Margin of Error	EstimateTotal	Margin of Error	EstimateTotal
3	411	114	277	95	16	19	26	69	58	79	79
4	442	98	360	98	10	17	13	21	49	37	27
5	404	107	256	82	18	21	27	34	54	44	46
6	282	90	241	77	9	15	19	29	91	55	39
7	244	152	131	136	0	12	0	12	0	12	0
8	1222	250	875	227	21	38	48	44	160	99	422
9	1041	264	774	225	14	21	16	29	128	115	241
10	269	47	198	45	0	12	15	14	16	9	40
11	309	82	176	62	10	10	11	14	18	17	58
12	432	79	309	55	26	30	8	9	38	29	79

	A	B	C	D	E	F
1	GEO_ID	Total	WhiteOnly	Bl-AfAmOnly		
2	510010901001	778	716	0		
3	510010901002	699	693	0		
4	510010901003	943	881	0		
5	510010901004	479	409	12		
6	510010902001	493	43	405		
7	510010902002	3004	2638	255		
8	510010902003	2274	1289	985		
9	510010903001	557	461	84		
10	510010903002	811	500	268		
11	510010903003	1071	924	105		
12	510010904001	1313	812	501		

1c Note that to map the Median Year Built data, you'll need to replace the 1939- values with 1939, and the 2014+ values with 2014. Use the **Find and Replace** function in Google Sheets.

NOTE: make sure that your year column is formatted as a number, not as text.

	A	B	C
246	510131018022	1994	1988
247	510131018023	1994	1988
248	510131018024	2001	1996
249	510131018025	1972	1969
250	510131018031	1965	1966
251	510131018032	1950	1958
252	510131018033	1946	1948
253	510131018034	1944	1953
254	510131019001	2002	1963
255	510131019002	1939-	1949
256	510131019003	1939-	1954
257	510131020011	1995	1971
258	510131020012	1978	1978
259	510131020021	1960	1962
260	510131020031	1979	1978
261	510131020032	1949	1953
262	510131020033	1939-	1945
263	510131021001	1974	1970
264	510131021902	1952	1953
265	510131022001	1957	1955
266	510131022002	1966	1961
267	510131022003	1969	1969
268	510131022004	1969	1974
269	510131022005	1966	1966
270	510131023011	1948	1953
271	510131023012	1954	1951

TWO KINDS OF MEDIAN YEAR ENTRIES WON'T MAP CORRECTLY, BECAUSE QGIS CAN'T INTERPRET THEM AS NUMBERS: "1939-" AND "2014+". USE "FIND" TO REPLACE THESE WITH "1939" AND "2014" RESPECTIVELY

	A	B
1	GEO_ID	MedianYrBuilt
2	510010901001	1979
3	510010901002	1956
4	510010901003	1988
5	510010901004	1983
6	510010902001	1984
7	510010902002	1995
8	510010902003	1983
9	510010903001	1963
10	510010903002	1982
11	510010903003	1969
12	510010904001	1977
13	510010904002	1976
14	510010904003	1990
15	510010904004	1960
16	510010905001	1971
17	510010905002	1958
18	510010905003	1952
19	510010906001	1980
20	510010906002	1974
21	510010906003	1947
22	510010906004	1988
23	510010907001	1989
24	510010907002	1976
25	510010907003	1958
26	510010907004	1978
27	510010907005	1972
28	510010908001	1962

Find and replace

Find: 1939-

Replace with: 1939

Search: This sheet

☐ Match case

☐ Match entire cell contents

☐ Search using regular expressions [Help](#)

☐ Also search within formulas

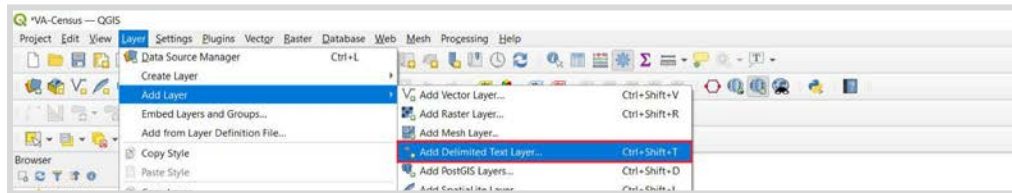
Find Replace **Replace all** Done

MAKE SURE TO SELECT "THIS SHEET"; THEN, "REPLACE ALL"

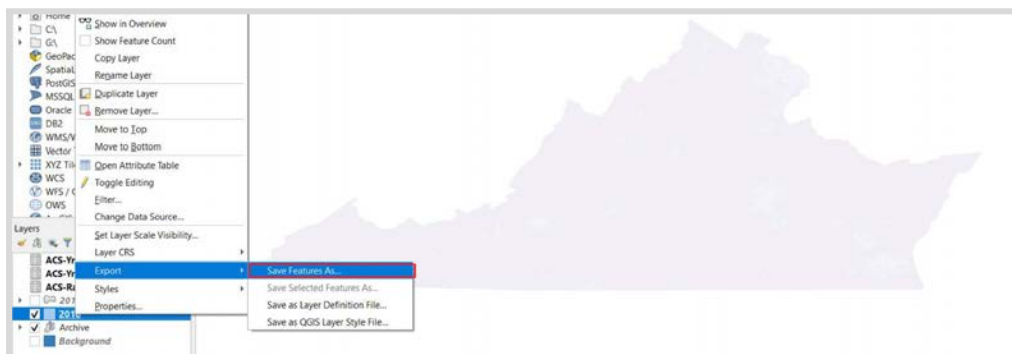
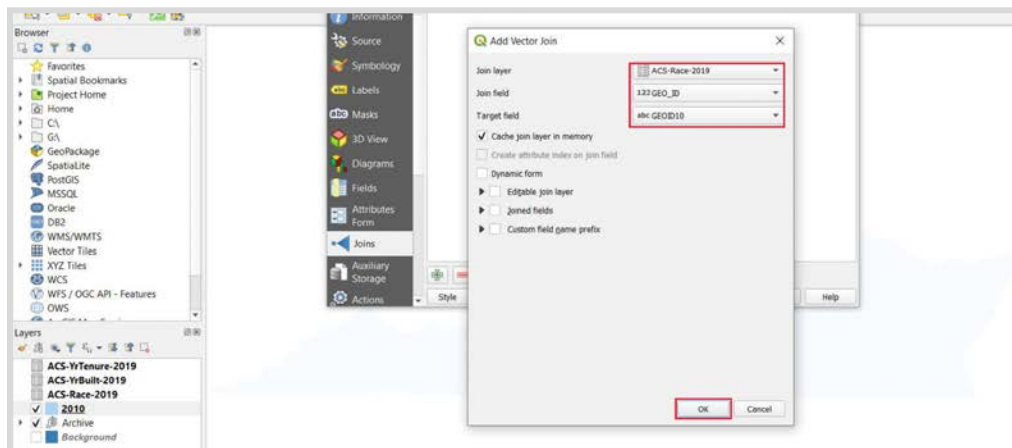
Step 2: Associate new ACS data with 2010 TIGER shapefile.

2a Add a Vector Layer for the 2010 TIGER shapefile.

2b Add a Delimited Text Layer for each ACS data table.

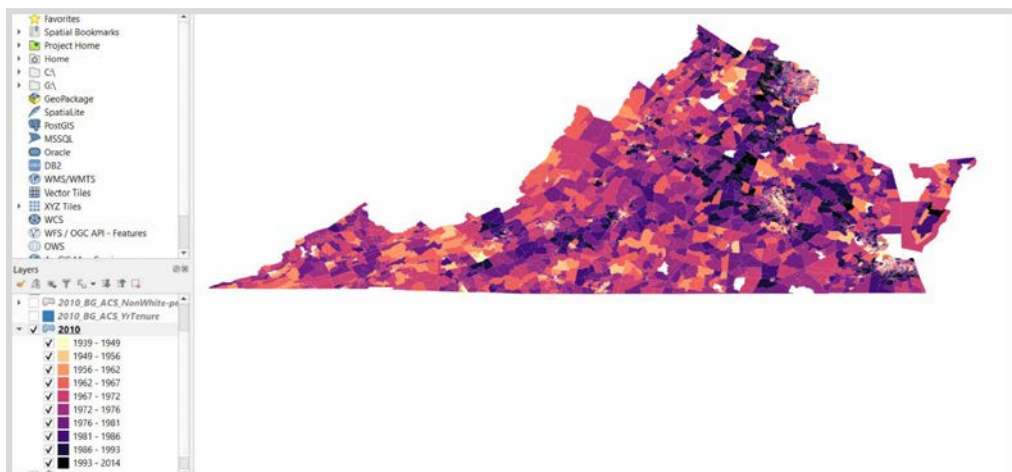
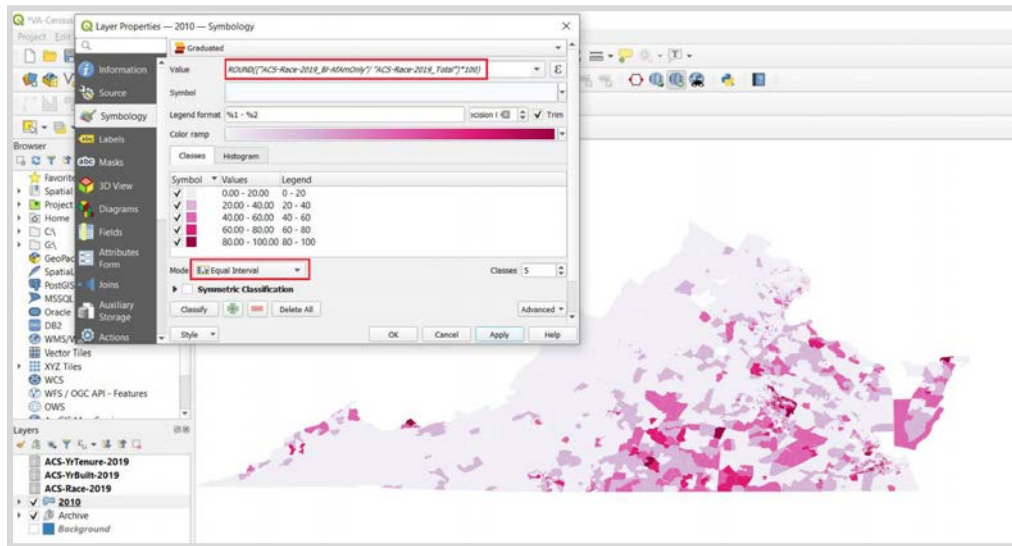


2c Join the **first .csv layer with your shapefile (double click > Join; GEOID to GEOID)**. Check that the attribute table has joined correctly, and then right click on the **layer Export > Save As...** to save the joined shapefile as a new layer.

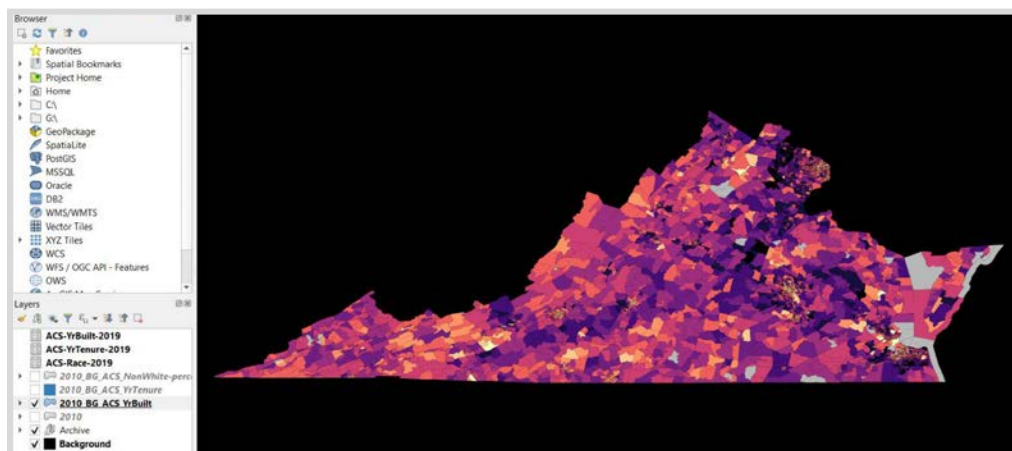


2d **Remove** the first join (red minus sign under joins). **Then, Join** the second table, and **save** again as a new layer.

Step 3: Using the Graduated style, color each layer by percentage non-white (using “Equal Interval” breaks) and median year built, respectively. Make sure line style is set to “no pen”.



3b Add a new Vector layer and draw a background polygon to offset the colors if desired.



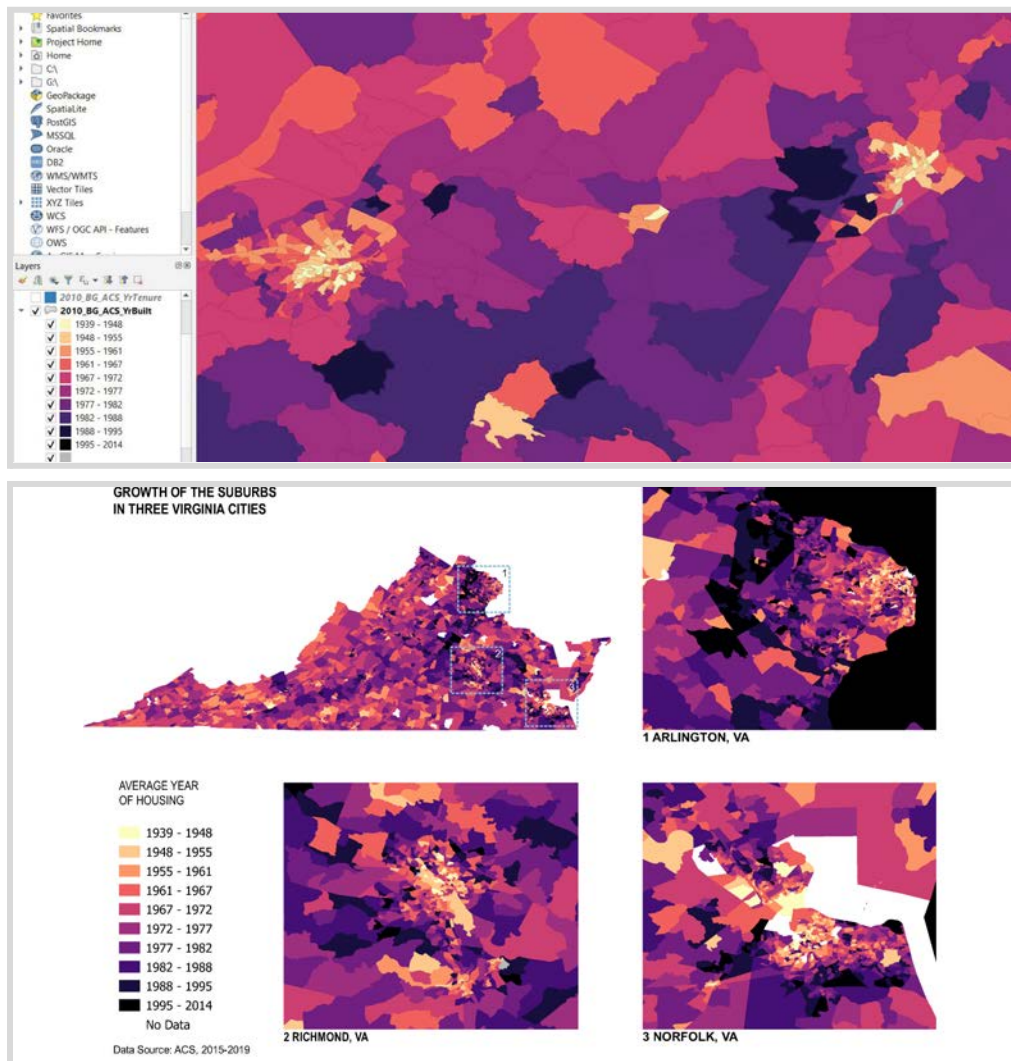
Step 4: Make a map layout showing only the median year built layer.

NOTE: If your Median Year Built variable does not work with the graduated style tool, try reformatting the field as an integer instead of a string:

1. Double click on the layer. Select "Field".
2. Click the abacus symbol (Field Calculator) at the top right.
3. Name your new field something like MedianYrBuilt-Num
4. Make sure "Integer" is selected as the field type. In the Field Calculator box, add in your Median Year Built field (double click on the field name from the right-hand list)
5. Click "Ok".
6. You should now see a new field with the years which is classified as an integer instead of a string.

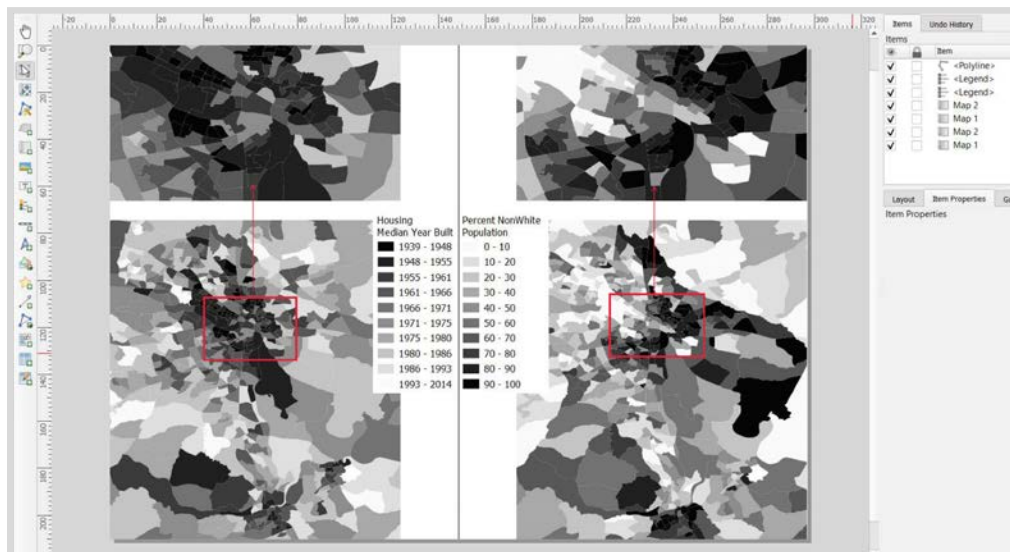
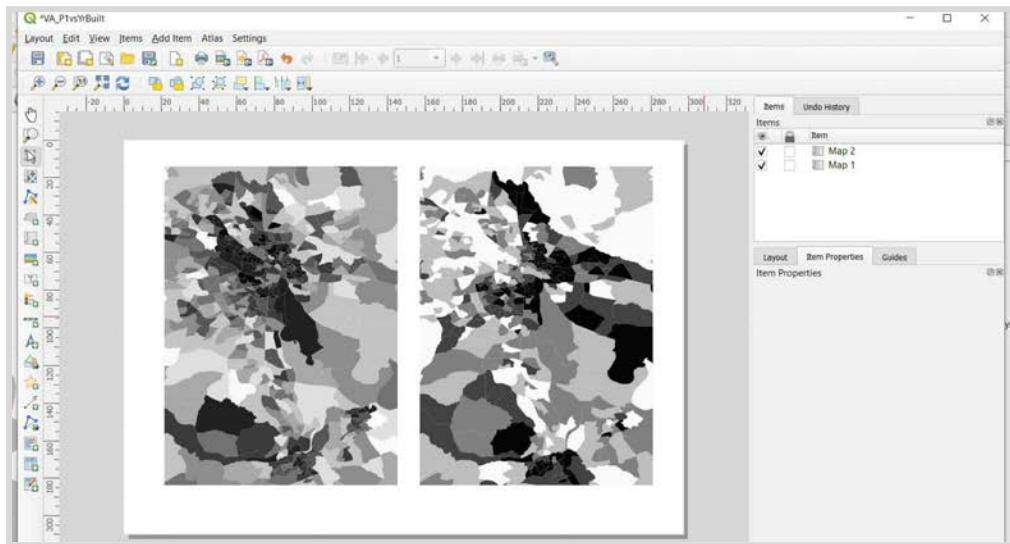
4a Choose **three** areas (including your chosen **Appalachian city**) on the map where you observe interesting patterns (eg, new suburbanization around cities), and create new maps ("Add Map") to zoom in to each area. These will function like callouts. Make sure to **lock your layers in each callout**. Make sure they're all Zoomed in to the same level (Map Properties > Scale).

4b Add a title, captions for the callouts, source, and a legend.



Step 5: Look at the correlation (negative or positive) between percentage of non-white population and average age of housing in your chosen Appalachian city.

5a Make a **second map layout** that compares these two layers. You might want to use callouts again to zoom in on specific neighborhoods, or outline areas of interest with the “Add Node Item” tool.



Step 6: Export both maps as 11x17 pdfs. (or continue to Bonus section)

- **Bonus** -

Step 7: Compare ACS racial features (table B02001) with P1 2020 racial decennial census data from Tutorial 4.

Step 8: Call out specific cities or zones to compare between the two data sets (eg zoom in on Roanoke or Richmond in both), to see the differences between decennial and ACS data.

Step 9: Export map as a pdf with legend.