Project Overview

During our Lab 1 exercise, we tested to see if the spatial distribution of assessed values per square foot in Tacoma had any correlation with the neighborhoods create in 1937 by the Home Owners' Loan Corporation (HOLC). Given what we watched in the overview video from NPR, I hope that you found it interesting that we did not find any strong connection between HOLC grades and home value. So this begs the question, what could have changed here?

This lab we plan on looking at demographic changes in Tacoma to see how populations have shifted over the last 20 years. To do this, we will be using information from the US Census to track population and race across the Decennial Census in 2000, 2010, and 2020.

Census data is tracked in several groupings of data which is covered in this video which I suggest that you watch. We are going to use Census Blocks, the smallest unit of Census Geography, for our analysis.

Our analysis will be very similar to the structure of Lab 1. We are going to assign our blocks to the associated HOLC Neighborhood spatially, aggregate the data, and calculate statistics to let us compare Neighborhood to Neighborhood. However, the point of this exercise is to show you alternatives to using operations like a spatial join, and use geographic operations with more fine control.

The first, is that we are going to use a distance calculation a given block and neighborhood based in the center of a block called the Centroid. Secondly, we are going to use a geometric operation called Union to merge the boundaries of a block and the neighborhoods, and calculate the percentage that a block overlaps with a given neighborhood.

Completing these joins, we will calculate the demographic changes between decennial census' to see if there is any correlation with the HOLC neighborhood grades. Our data will be joined to the blocks based on the geoid for the corresponding year, and we will aggregate the data to the respective neighborhoods.

We will map these changes in a Proportional Symbol map for each of the three years, so we can compare the population change over time.

Download Data

Pierce County GIS Data

- Pierce County Basemap
- Roads

IPUMS National Historic Geographic Information Systems

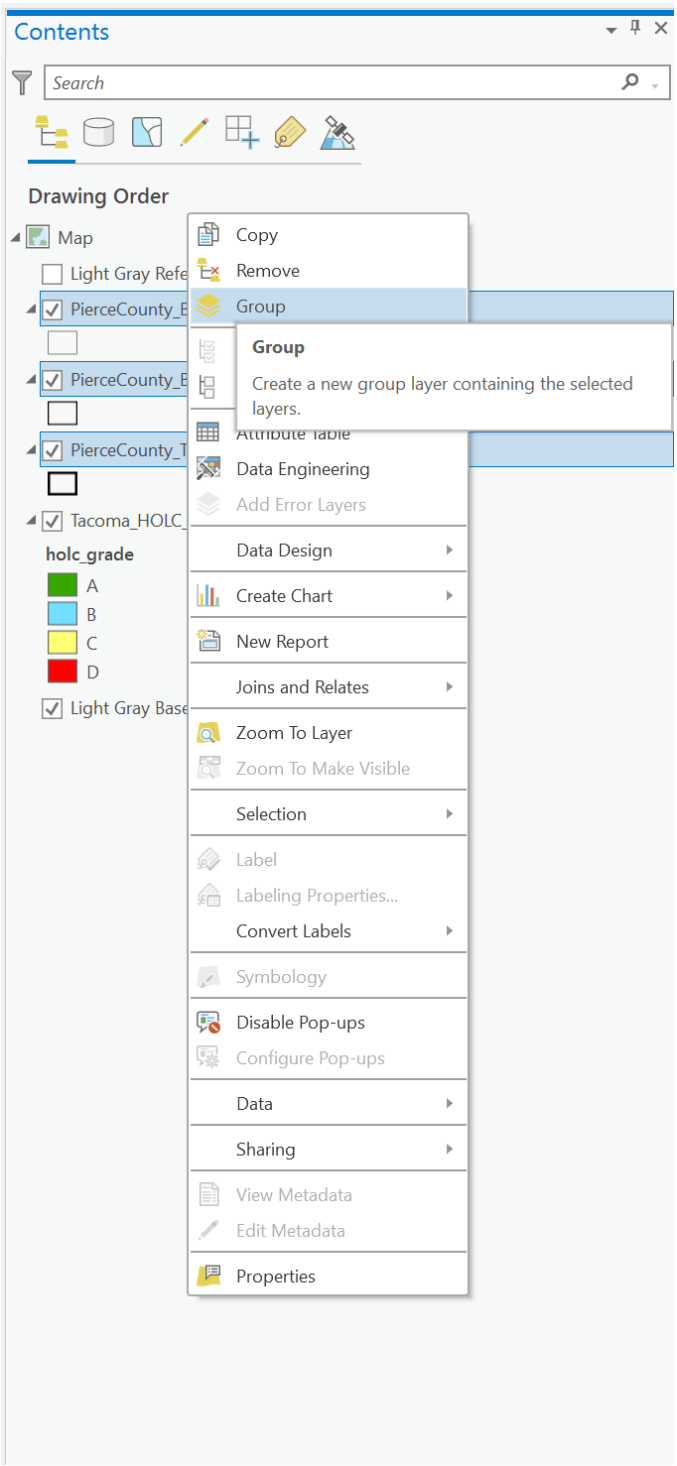
- Instructions are included below or grab data from GitHub

Census Tiger/Line

- TIGER/Line® Shapefiles Download
  - Year 2010 -> Washington State -> Pierce County
    - Blocks
    - Block Groups
    - Tracts

Map Setup

- New Project
- Import Data into Geodatabase
- Setup Working Map Document
- Import Pierce County GIS Data
  - Group Data onto Map


Contents
Search
Drawing Order
Map
Light Gray Refe
PierceCounty_E
PierceCounty_E
PierceCounty_T
Tacoma_HOLC_
holc_grade
A
B
C
D
Light Gray Base
Copy
Remove
Group
Group
Create a new group layer containing the selected layers.
Attribute Table
Data Engineering
Add Error Layers
Data Design
Create Chart
New Report
Joins and Relates
Zoom To Layer
Zoom To Make Visible
Selection
Label
Labeling Properties...
Convert Labels
Symbology
Disable Pop-ups
Configure Pop-ups
Data
Sharing
View Metadata
Edit Metadata
Properties

Analysis

Because we cannot break up Blocks, we need to be able to assign data a complete unit of a block group to a neighborhood.

Last week, we discussed in inconsistent results of a spatial join operation, so now let's spend some time actually doing the spatial operations manually.

Project Data

Our data is currently in two different Coordinate Reference Systems.

What spatial reference systems is our data in?

Project your data so that it's all in the same projection.

We can actually do this two different ways. Either Export your data in the desired Projection, or use the Projection tool in the Geoprocessing tab. We will be calculating areas in this lab, so choose an appropriate projection.

Filter to interesting Blocks

Let's filter down our blocks to those that are close to our HOLC Neighborhoods. Select Census Blocks that intersect with HOLC Neighborhoods and export our data.

Summarize the relationship

Let's explore our data a bit before we work here a bit. What is going to be difficult in our analysis here?

How many parcels are going to be easy to assign a neighborhood? How many parcels are going to cause us difficulties? What types of things do we have to be careful about when doing our spatial analysis?

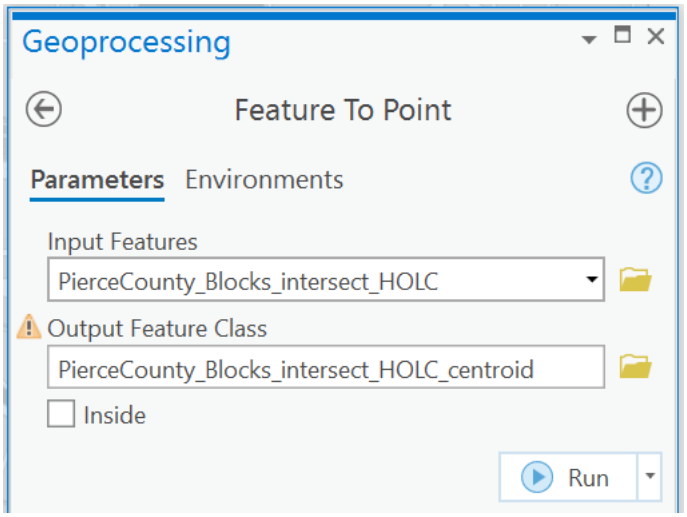
Spatial Analysis

We are going to do two different analyses to compare how our data is related. We are going to preform the nearest analysis to identify the nearest related features to a given Block, followed by some intersection analysis.

Nearest Analysis

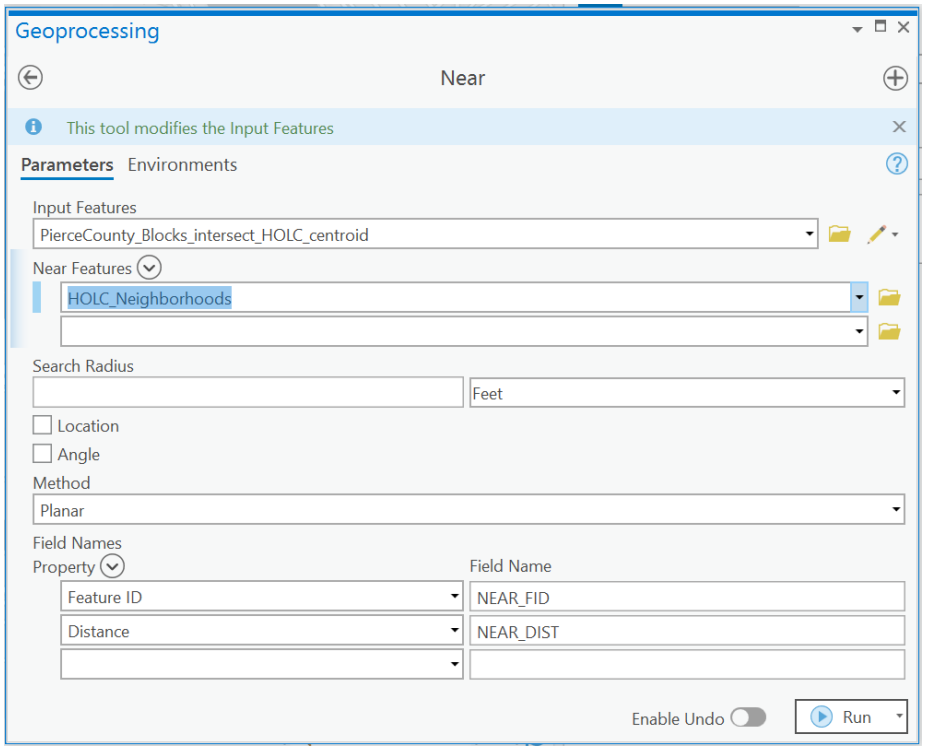
We are going to calculate the Centroid of the Blocks and see how this relates to how close the Centroid is to a given HOLC Neighborhood.

The centroids are calculated using the Feature to Point tool. A Centroid is calculated for the center for a polygon. However, this can be actually kind of difficult for complex features. Use the Help button. **Explain what the Inside checkbox in this tool does in your own words.**

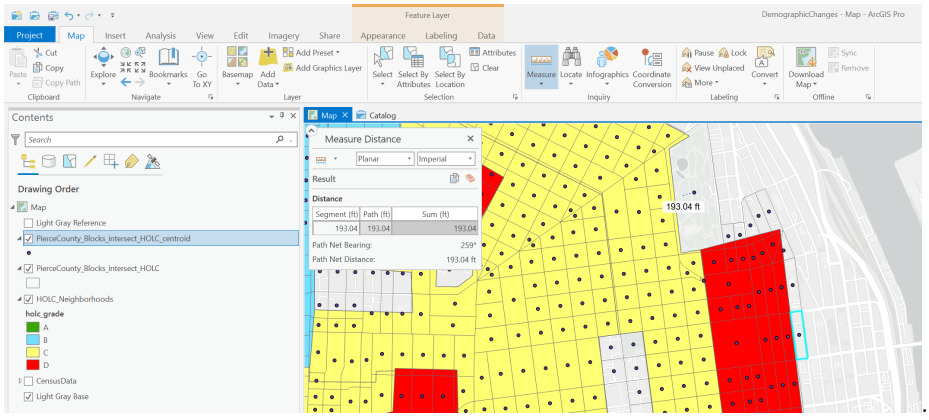


Finish running the tool and save your data to your Geodatabase and add to your map.

Now, we are going to use a Near Analysis to assign a neighborhood to each of our points. This tool will add the details about the nearest feature to a given point, which could help us bin our Blocks to our neighborhoods.



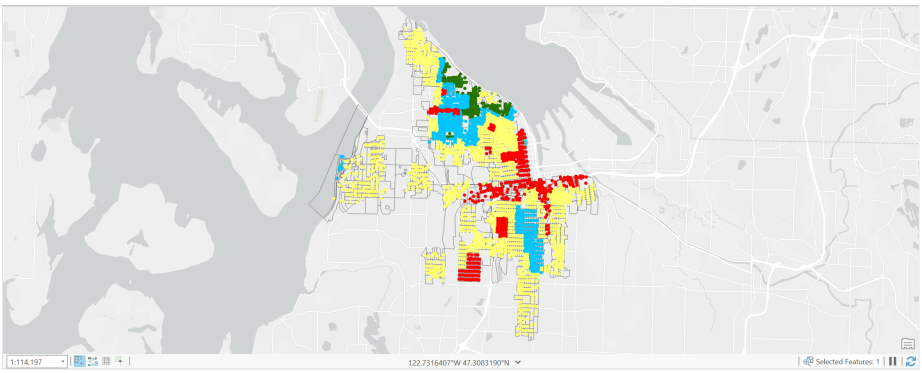
We can also limit how far we search for a neighbor in this tool, so let's explore this a bit. There is a Measure Tool in the Map tab of the ribbon. If we have done the work to set up our map correctly, we can actually measure real distances in units we know, to set up an appropriate search distance.



Explore a couple of points that fall outside our neighborhoods. Take a stab at an appropriate search distance to assign our Blocks to the correct neighborhood. Run the Near Tool with this number.

The Near tool is running a standard Nearest Neighbor Analysis, which is an example of a spatial algorithm. Give it a read here

Once that is complete, join the HOLC data back to the points data and visualize the points based on the holc_grade.



How well does this do at approximating our neighborhood boundaries?

Include a screenshot of your Blocks symbolized as points and add to your write-up. Compared to a spatial join, does this give us better flexibility to assign a block to a given location?

Topologies SideBar

Census Blocks and other census geographies use something called topologies.

This data model is pretty common and basically helps eliminate errors in your data model. Think about the neighborhoods you created, how difficult was it to make the neighborhoods match to each other perfectly? Basically, Topologies allow us to assign rules for how features are related to each other. For example, if you think about parcels, there typically isn't spaces between one parcel to another. Rather they are a cohesive 'fabric' which makes up an entire County. If you think about a parcel, if you needed to redraw the property, you are redrawing the neighboring parcels boundary too. That is, each parcel has a shared boundary, and an update to one geometry necessitates and update to the other. Otherwise, we have invalid definitions, like slivers or holes in our data, YUK!

Union

The Centroids do an okay job of assigning doing our neighborhoods. But often, we want to have that fine grain precision to control our binning.

To do so, we can use several geometric operations to be able to relate geometries. These are discussed in Chapter 10 of Learning ArcGIS Pro 2.

For our analysis, we want to be able to Union the geometries of the neighborhoods and census blocks. Then we can calculate the area of each block to see how much it intersects with a given neighborhood. Obviously, for the vast majority of our blocks, the delineation of neighborhoods and blocks are actually similar features such as a road. But in other places, a block could be in two or more neighborhoods. Becasue the way sampling works, we do not want to split a census block.

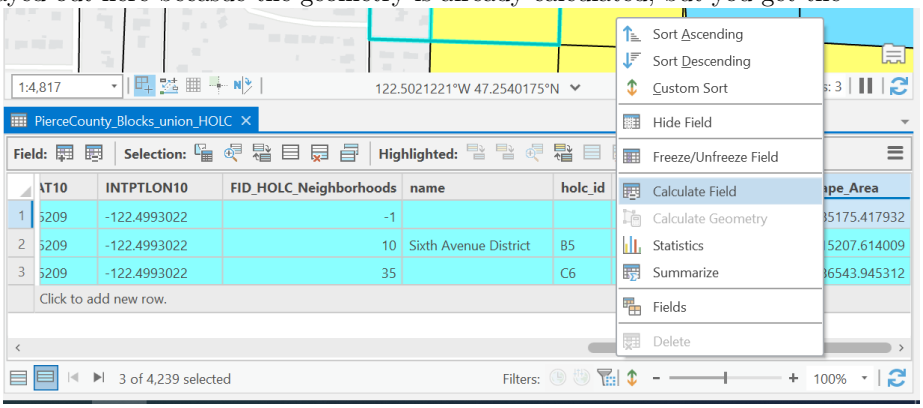
For example, consider this block (GEOID: 530530611006007). It intersects with two different neighborhoods and also is partially outside a neighborhood.

Which neighborhood would you assign this too?

The union is going to take the boundaries from our inputs and merge the boundaries, so in this case, we are left with three features based on the operation. HOLC union operation

Once completed, we also have the total area of the feature also. ArcGIS typically will run the Calculate Geometry function on any feature after Geoprocessing. However, if you ever create your own features, or edit feature geometry, you can right-click in the attribute table, and select Calculate Geometry.

It's grayed out here becasue the geometry is already calculated, but you get the point.



From here, we need to use the dissolve tool to aggregate our data back to the parcels to calculate our statistics.

Use the Dissolve Tool to

For our rules here, I am going to let you come up with your definitions of what is in a neighborhood and what is not.

For example, should a block count as being part of a neighborhood if only 5% of the feature intersects with a neighborhood?

Write out your definition for a parcel to be included in a neighborhood.

Once you have your definition for a block inclusion into a neighborhood, we have

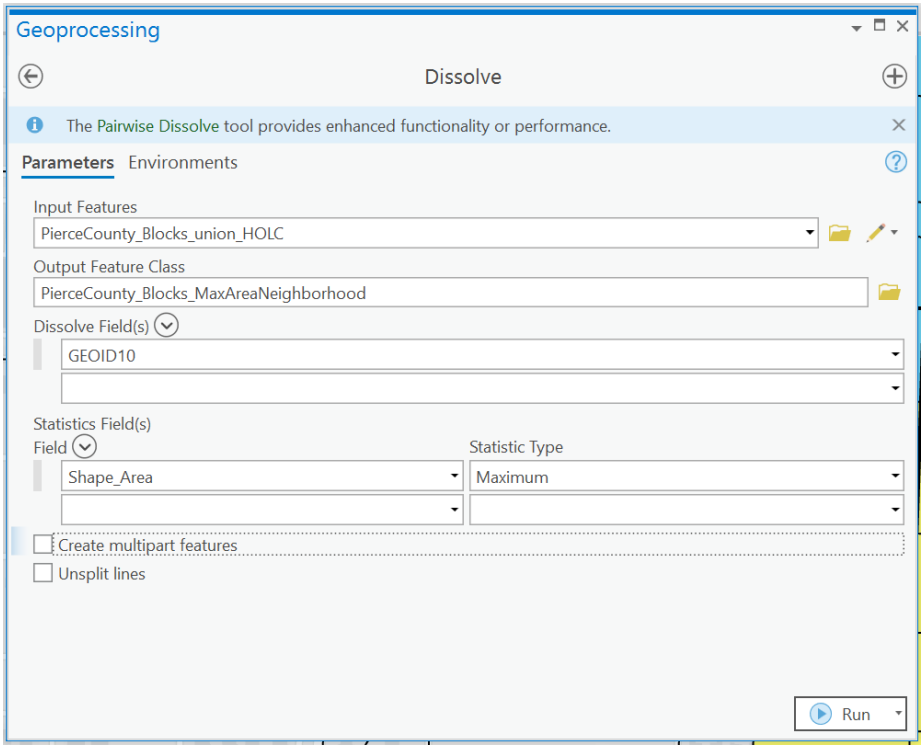
to execute a query to return the single features where our definition is true. So our example, GEOID: 530530611006007. We have three parts of the Block split by the Neghborhoods. Let's only select the part of the Block Group that has the value for the HOLC Neighborhood we want.

For me, I am going to do two different calculations here.

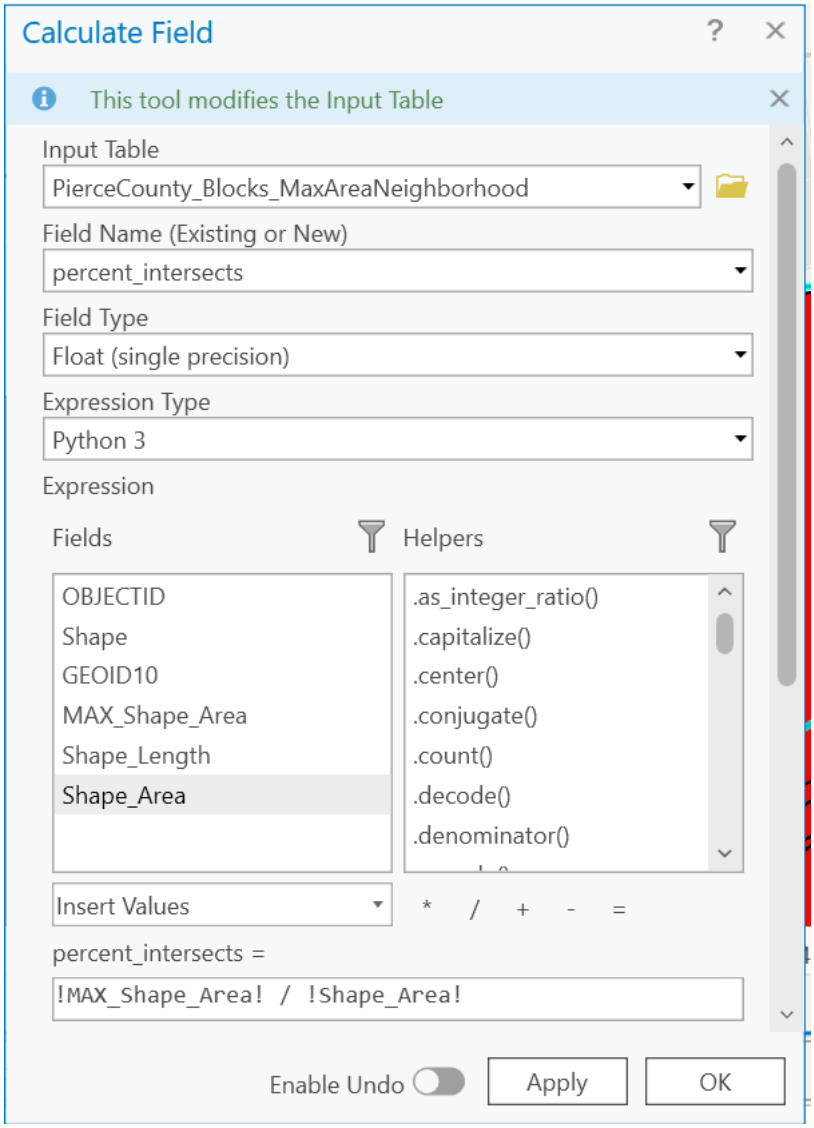
1) We want to assign a block to the nieghborhood with the most area intersecting.
2) Only for areas that intersect greater than 5% of the total area of the block group.

Make sure to use your own definition here.

We will Dissolve our Unioned Blocks layer, and calculate the Max area for a given Block.



Let's calculate a field to say if this feature matches our second condition.

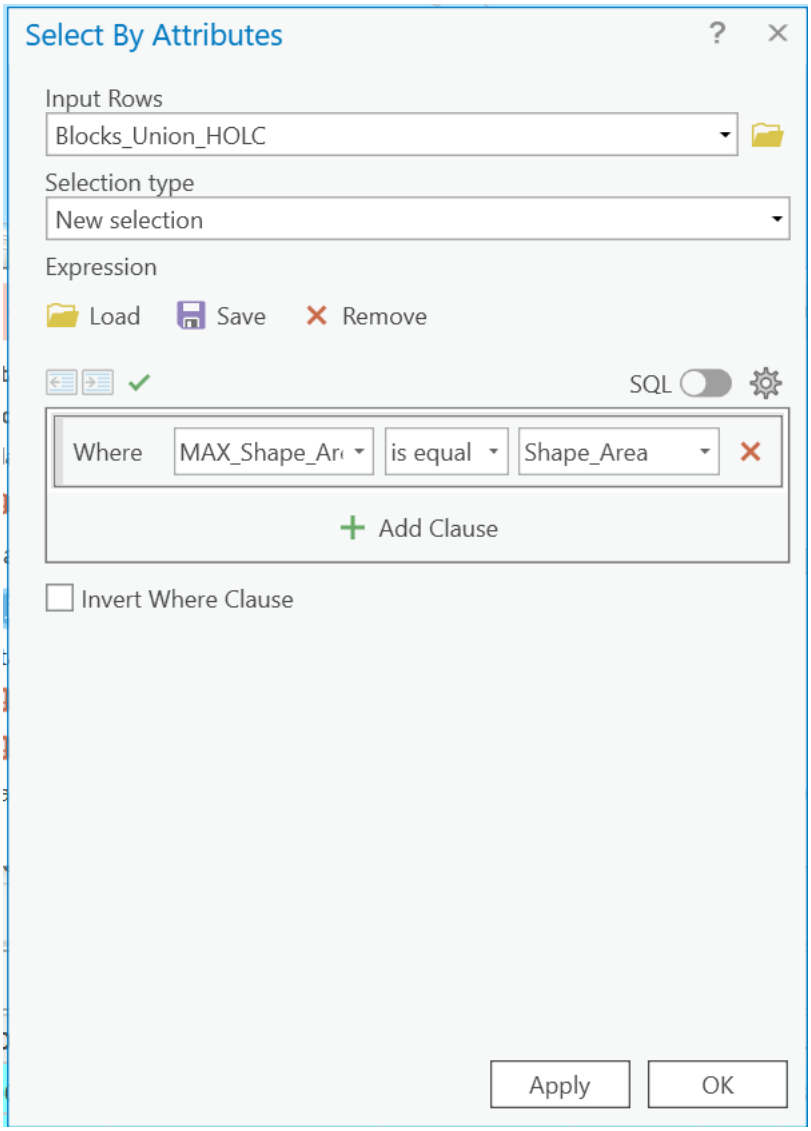


And Join the result of this feature back to our PierceCounty_Block_Union_HOLC by using the GEOID (Census Geography Identifier).

We now filter based on our conditions outlines above.

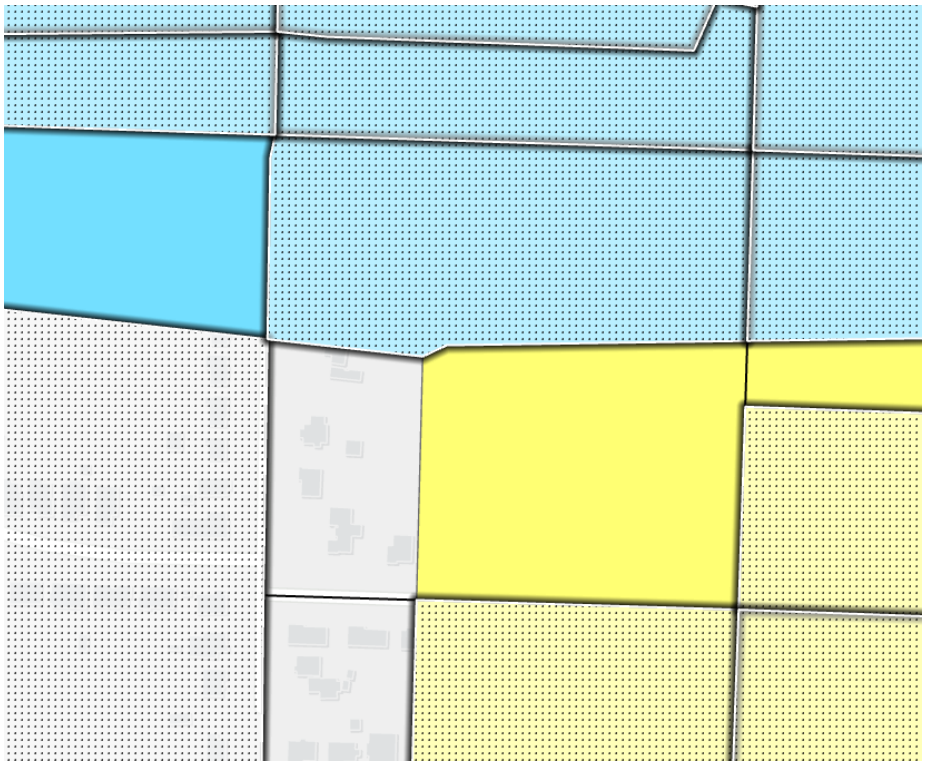
In this case, instead of querying for a certain logical condition, we need to select the features based on what's in another column of our data. So we flip to run our selection to fields where:

```
Shape_Area = Max_Shape_Area AND percent_intersects > .05
```



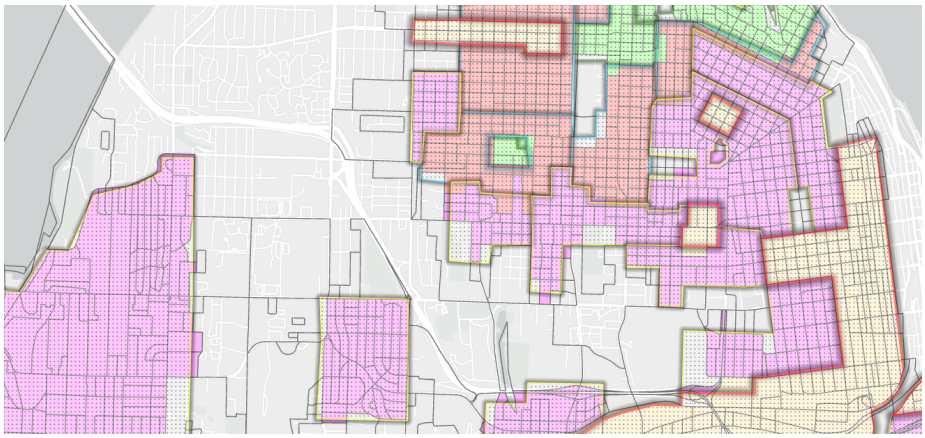
Once finished, we should only have a single part of our Blocks with the correct neighborhood information. Export this data as its own separtate layer as PierceCounty_Blocks_Assigned_HOLC.

Checking out our favorite Census Block:



Finally, these are still parts of Blocks, so we need to get the original geometry again. Join the data based on GEOID, filter the records where the joined GEOID is not NULL, Export, and clean up our data to only show the fields we care about.

Compare the results from using the Centroid data vs our geometric operations process. How much more accuracy do we get when using these methods compared to the spatial join out of the box?



So why did we do all this work?

Well, population data is an interesting data set as in we are estimating a population of people, over places of land where there may or map not be any actual people. If you watched the census video, you can see that some census blocks make up some interesting features, such as Highway Medians, or parks where most people do not live. So population is not uniform throughout a given polygon. That is one reason why we cannot break up a given census unit unless by its intermediate parts.

Getting a better population estimate would be to convert population data into a continuous surface, or a raster.

This process is called Dasymetric Mapping.

Give it a read.

Summarize what a Dasymetric Mapping is and how it would be useful for our Tacoma HOLC Neighborhoods?

Population Data

Finally, let's get into our actual Census Block Data.

As discussed, in the Census Geographies Video, Blocks change each census. So, while it would be best to do this same analysis for each the Census Blocks data from each respective year, we will rather use the Census Block Crosswalk to get the associated Blocks at a given location for each of the other years. This shows the relationship between the Census Blocks for each year. In this case we used 2010 data originally which can be related to 2000 and 2020 data respectively.

Download the relationship files to relate the following:

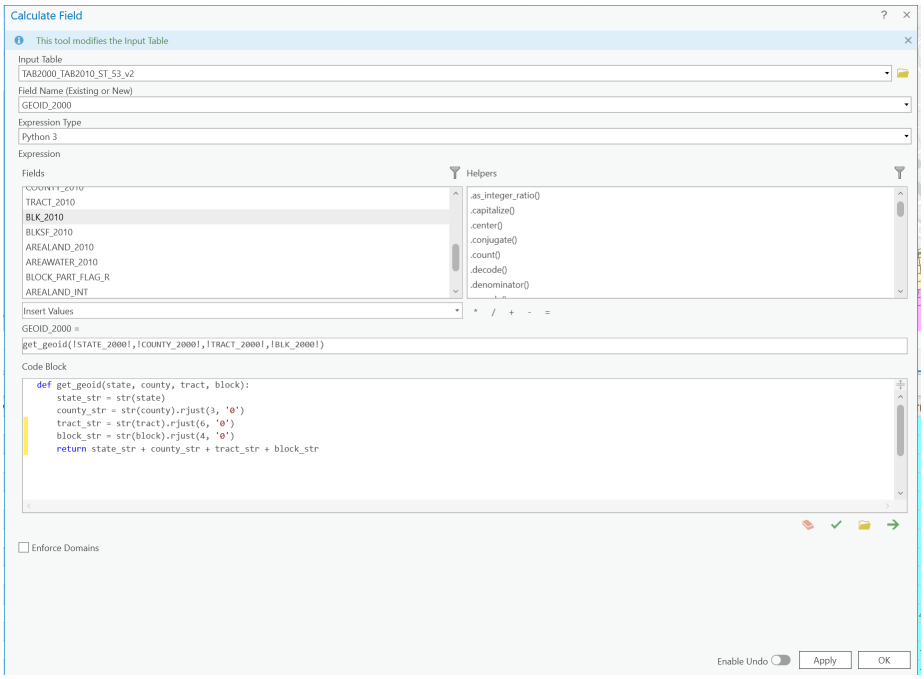
2000 Blocks -> 2010 Blocks 2010 Blocks -> 2020 Blocks

Use the following website to download the two crosswalks above. Navigate to the right year and only select the Blocks in Washington State. 2010 Census Tabulation Block to 2020 Census Tabulation Block Relationship Files

Import these tables to your geodatabase, and add to your map.

Join each of the data to your neighborhood data.

To do so, we need to calculate the GEOID for each of the year.



function call:

```
get_geoid(!STATE_2020!,!COUNTY_2020!,!TRACT_2020!,!BLK_2020!)
```

code_block:

```
def get_geoid(state, county, tract, block):
    state_str = str(state)
    county_str = str(county).rjust(3, '0')
    tract_str = str(tract).rjust(6, '0')
    block_str = str(block).rjust(4, '0')

    return state_str + county_str + tract_str + block_str
```

Once calculated, Join both tables to our Tacoma_Blocks data and export again.

Historic Population Data

Now let's work with our actual population data.

If you would like the practice and to support cool programs which make things easier for you get data, you should download that data yourself. It could be very helpful in the future.

We are going to use the IPUMS National Historical Geographic Information System. However, somehow the University email system blacklist the system to

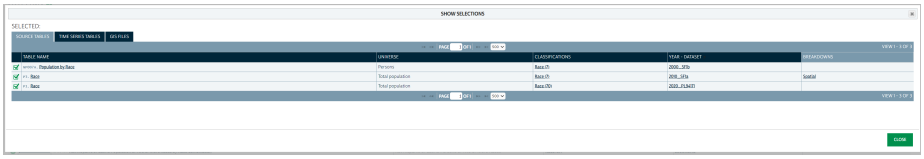
create a user account, so either use another email outside the University system, or I have provided a link to the data on GitHub.

If you are going to download the data yourself, use the filters to search for:

Dicennial Census: 2000, 2010, 2020 Geography Level Block: Topics: Race

Once these filters are completed, there will be a list of all the tables with related data, and how it can be partitioned in different ways, which is really really cool!

However, we are going to want three tables:



Once downloaded, you will see that there will be a csv for each of the tables downloaded, and a text file which shows the metadata. We will need to reference these metadata tables to complete our next steps.

Once downloaded, import these three data tables into our geodatabase. Add to your working map.

One additional thing that we need to do for this data if you look at the metadata for each of the years, is that the actual Census Questions were formatted different from year to year. We need to standardize our questions to allow us to compare the same data year to year.

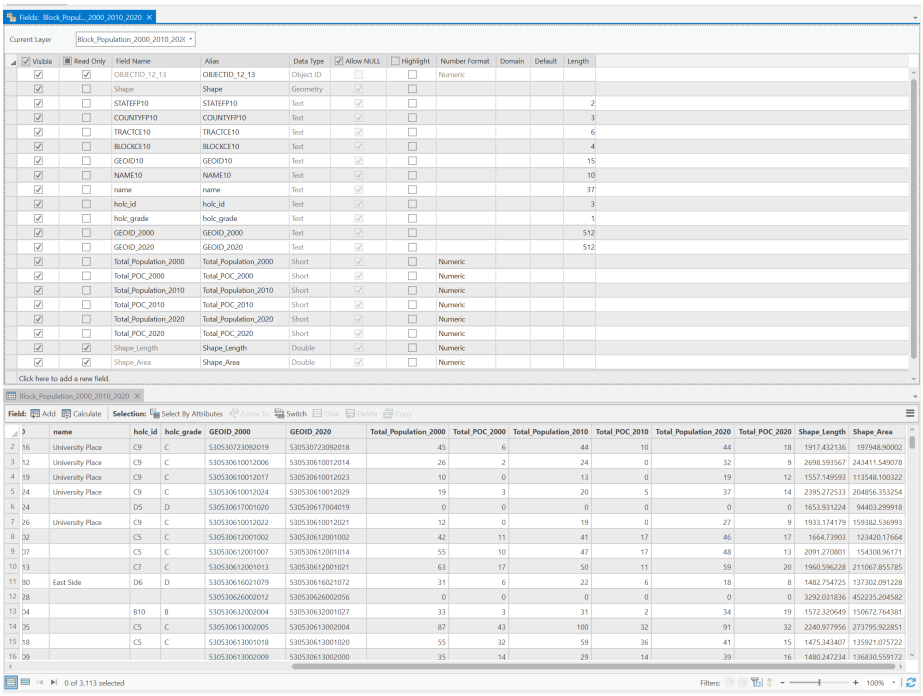
For each year, let's get the following:

- GEOID_YEAR
- Total_Population_YEAR
- Persons_Of_Color_YEAR

For each of the tables, we want to calculate these three fields. The GEOID Caclulation should be the same for each table using the function used in our crosswalk above. For Total_Population and Persons_Of_Color you will have to reference the metadata tables to identify how to calculate the correct number. Also make sure that you use the correct field types when running this calculation to save yourself work later!

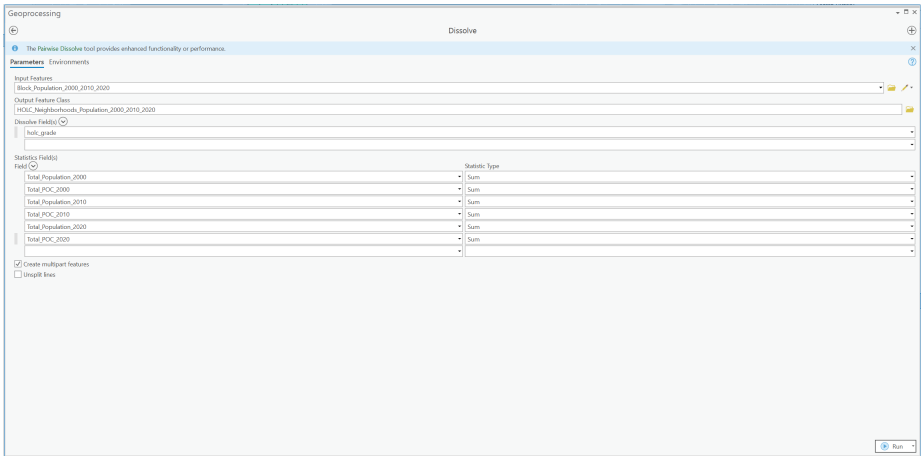
Once finished, we are going to join each of these tables to our final Blocks with our crosswalk data. Make sure to join based on the correct Census Year. Once finished, export this data so we can have one clean table.

I have the following fields:

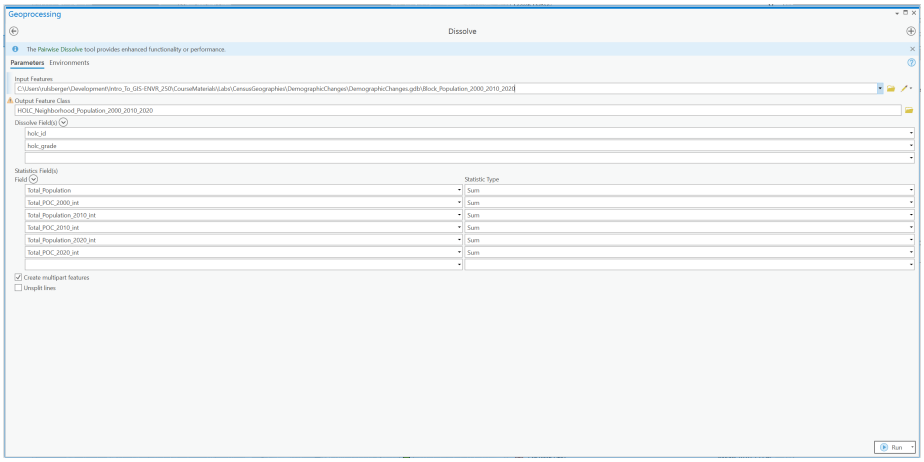


Now we finally can get somewhere with our analysis. We are going to use the Dissolve tool like we did in our last lab to identify what the total population of each of our HOLC Nieghborhoods, and summarize how they have changes over the past twenty years.

So run the Dissolve, and we are going to aggregate twice again. 1) By the Grade 2) By the HOLC_ID



And again by the HOLC_ID AND HOLC_GRADE



Now that we have our two maps, let's think about the best way to display this data.

MAP IT!!!

So now that we have this totals data, there are a lot of interesting ways we can examine it.

Totals themselves are interesting sometimes, but Census Geographies are drawn to represent groups of people. So they don't have any regard for actually how large a block is compared to each other if that makes sense. However, if we did want to actually compare census blocks, or neighborhoods in this case we could look at population density.

Additionally, we are also interested in how things change year over year, but like we have talked about, typically people are really bad at viewing totals data as it changes year over year, especially when drawn on a map. This particularly could be suited better for a different visualization like a bar chart. So we often use several visuals to help us explain our data.

For our purposes, I believe that two different statistics would be valuable for us to explore: 1) Total Persons By Year and Neighborhood 2) Percent Persons Of Color by Year and Neighborhood

For each of these we are actually communicating two different variables, 1) the HOLC Neighborhood Grade, and 2) the corresponding percentage of POCs. So we will need to use several maps, and also tables and other charts to communicate our data effectively.

For our map, we are actually going to have three maps: one for each year side by side so we can easily compare our data.

Interestingly, for this data, we have to think carefully about the range of our data, as our distributions of data actually can change per map. Instead, we

want to have the same scale over all maps, so we can compare the changes by neighborhoods and year.

To do so, we need to generate a symbology that represents the range of data across all three years of data.

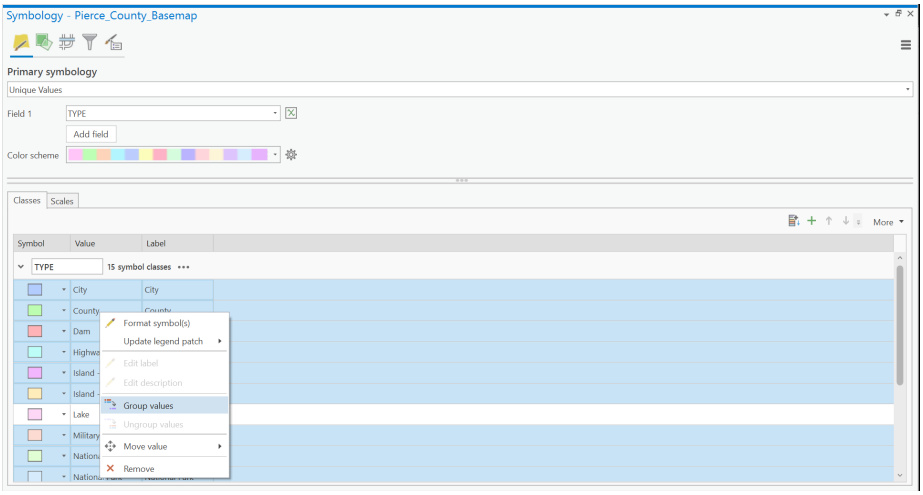
This is a very repetitive process, but it can be done quickly and efficiently if done in the right order. We are going to create a map for each year, so there will be a total of 6 maps that we will create for this lab.

First, we actually would like to get away from ESRI Basemaps. So let's recreate the light gray basemap.

Basemap

Go to the Pierce County GIS data portal, and download the Pierce County BaseMap, and Roads data. Bring into your ArcGIS Pro project, and you should have some ugly data showing.

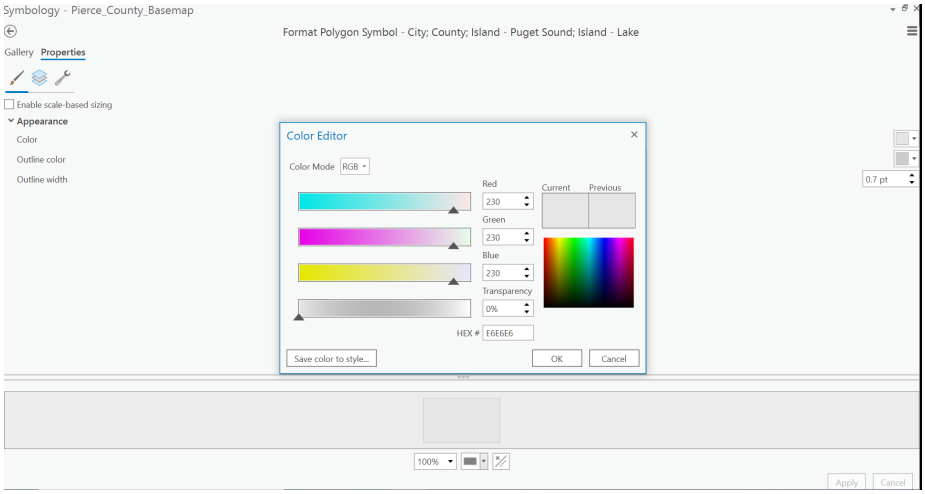
Starting with the Pierce_County_Basemap; let's symbolize this by Unique Values: Field 1 = 'TYPE'. Once doing so, we should get an ugly mess of colors. We don't actually care about all the different values for our basemap, so rather, we are going to group our values and have a very light gray represent urban boundaries, a catch all for all the other values we don't care about, and a slightly darker gray for the Puget Sound and other bodies of water. Control Click similar values, and right click -> select Group Values.



Do the same for Lake and Puget Sound so that we have two different Groups and we also want an "All Other Values" option for all the values we don't really care about.

Now for the symbols, we want to use all light grays, but still show some delineation of interesting features like city boundaries, so let's create a custom color.

Open the Symbology for our basemap and click on the patch for our urban areas.In the Properties tab, click on the color, and select Color Properties.

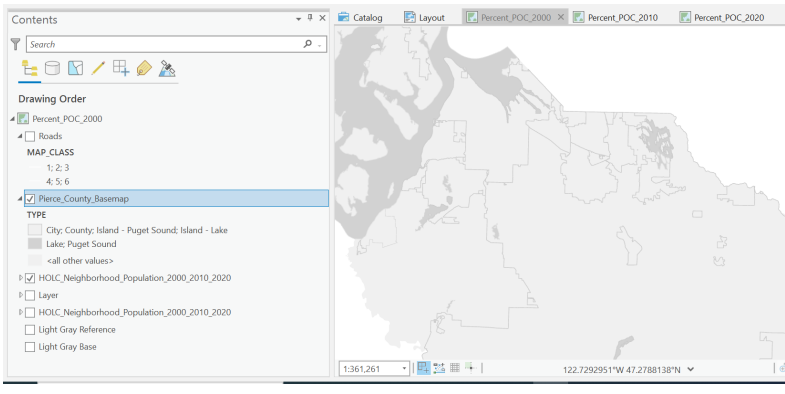


Use RGB Color Type, and select Red: 240 Green: 240 Blue 240. And assign a faint line in slightly darker gray for our urban areas.

Repeat this step for water bodies using 210 Green: 210 Blue 210.

And Finally, the same gray for the urban areas, but this time remove the line either with a width of 0, or remove the color for the "all other values" bin. This will not show any of the other details for these features that we don't want to show.

You should end up with something like this:



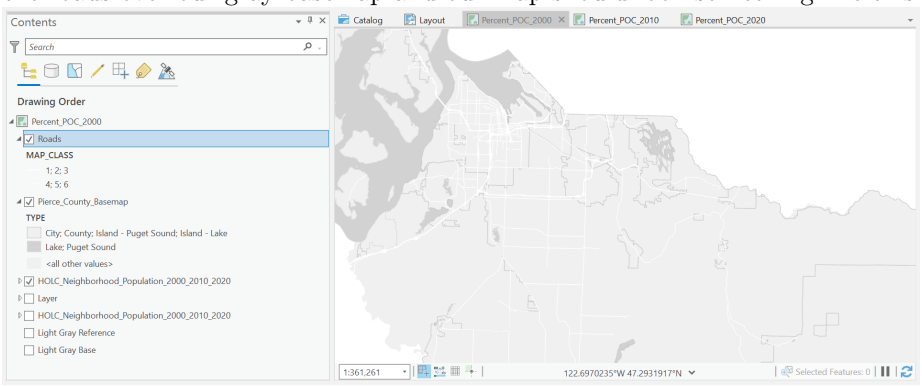
Now let's work on our roads, we want our roads to pop off our gray, so let's use a white to show the contrast.

We are going to use the same process as above, but this time, use the MAP_CLASS field. This field basically categorizes the roads in Pierce County to give them a different hierarchy for different types of roads. For example, I5 is

a common reference road and typically given a greater visual hierarchy than a normal arterial. Lets create two bins for our roads.

Like before, group the MapClass 1;2;3 into a Highway bin, and MapClass 4;5;6 into an arterial bin. Remove all other values from our symbology.

For these values, we rather are going to use size to differentiate each bin. Let's assign a Line Width of 1 to our highways, and Line Width of .7 for arterial. Order the roads over our gray basemap and our map should look something like this:



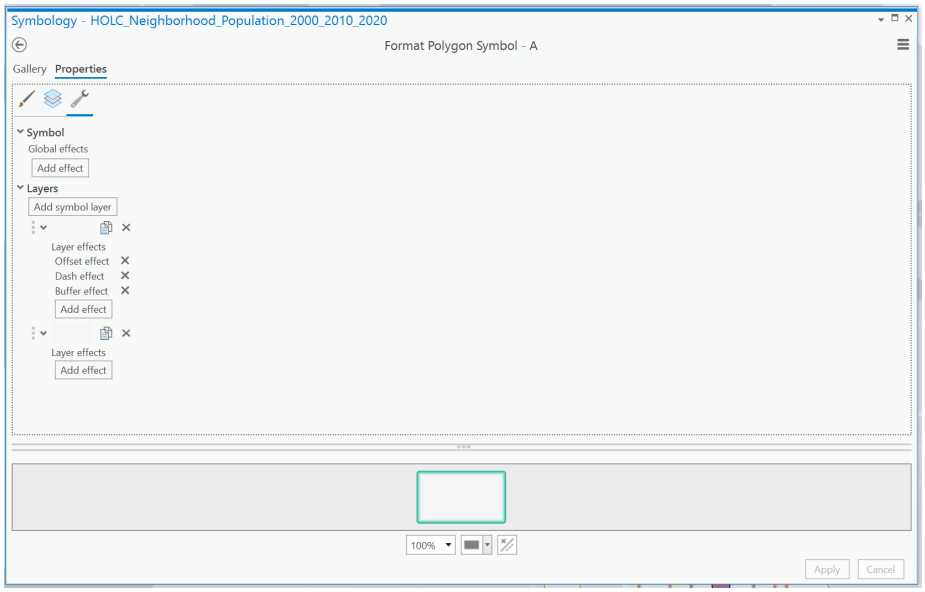
HOLC Grades

Continuing on, we now are going to add our neighborhood data. We need to add the population estimates for our grades and neighborhoods we calculated. Add these above our Pierce_County_Basemap.

We are going to create a bivariate proportional symbol map where each Proportional Symbol shows the grade, and the total population for a given neighborhood. Additionally, we want to show the actual boundaries of the neighborhoods still so we will display these in the same color as our Proportional Symbol.

Starting with Grades, lets symbolize our data using Unique Values on holc_grade.

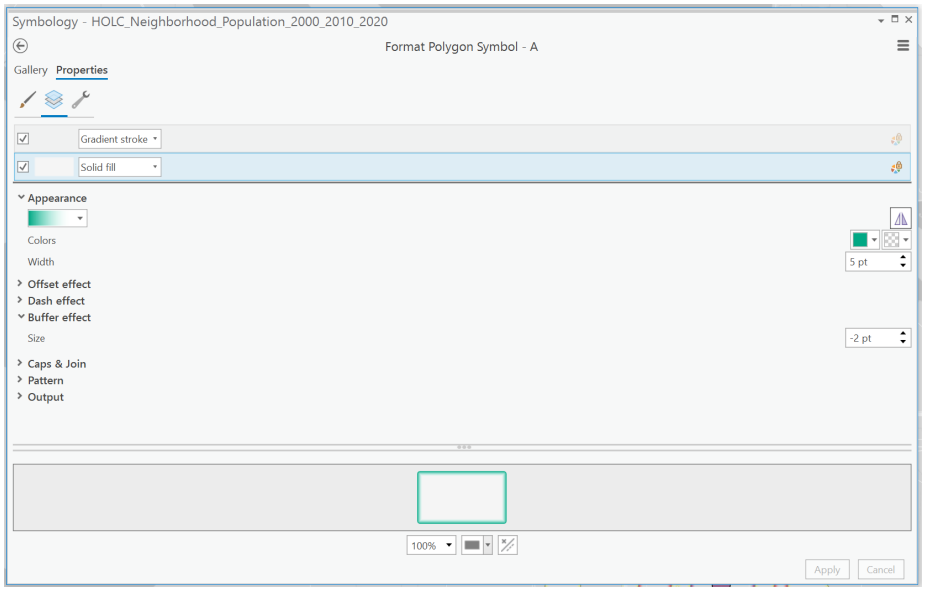
Now lets get into some fancier symbology. Click on the symbol patch for Grade A and click on the wrench (Structure). This breaks down the polygon symbol into its component parts, and lets us get into real fine control of what our symbols look like. We are going to create an outline of our grades that fades out to show our Proportional Symbol.



Click on stroke layer, and then click Add Effect to one of the stroke symbol. Add a Buffer Effect.

Now Click on the Layers Icon. Change the Stroke Type to Gradient stroke. Instead of referencing material here, please watch this again on the recording from class.

We are going to change the gradient to fade to 100% transparency and give a kinda glow effect for each neighborhood. Give the buffer a size of 5 and offset the buffer by -2 pixels. Change the colors for each grade accordingly. In the end it should look something like this:



We are going to use these layers as a base for the six maps we are going to create next.

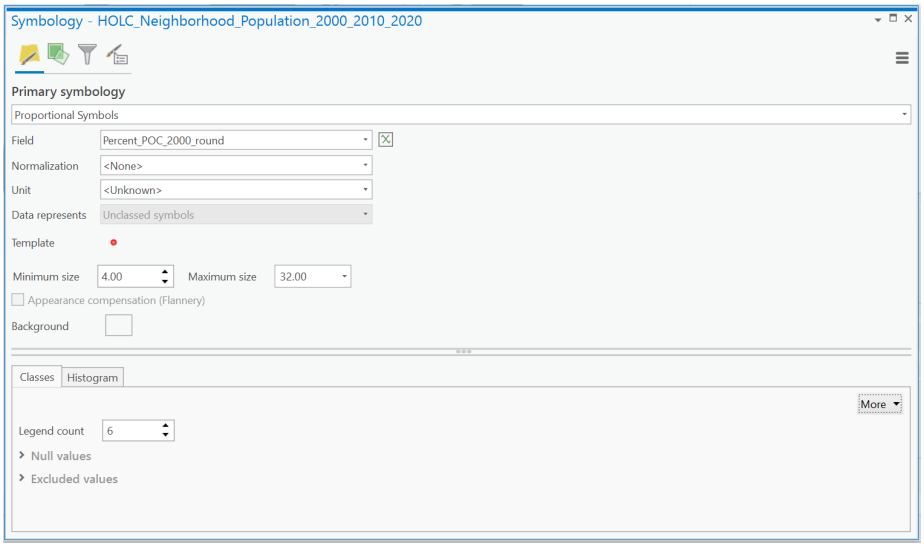
Map 1: Percent Persons Of Color By Year and Neighborhood

Now that we have our basemap, let's use some creative solutions to be show our Percent Population data.

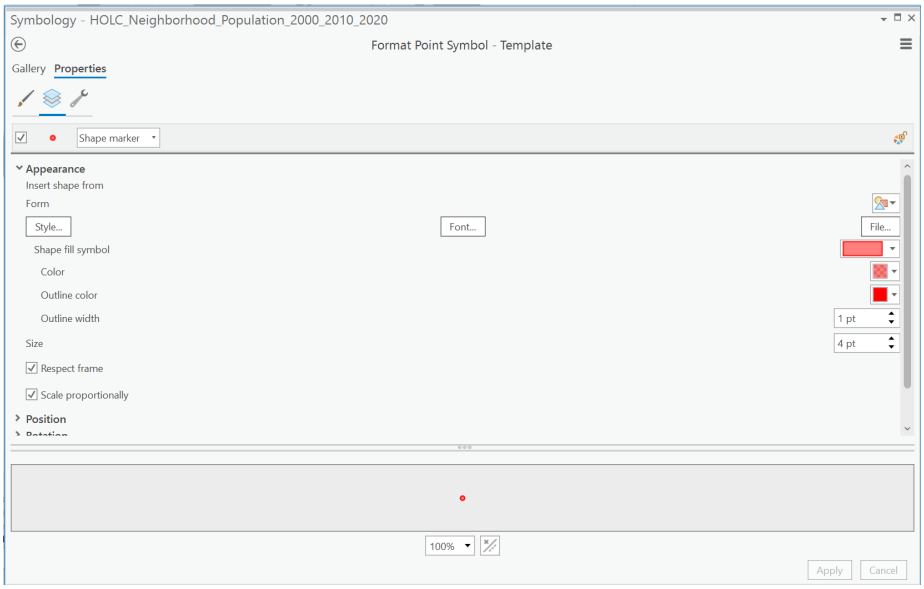
First, since you guys are now pros, calculate the Percent Persons of Color for each year so that the percent is a whole number between 1-100 (hint times by 100, and use the correct field type).

Now, let's start with the Year 2000.

Symbolize the HOLC_Nieghborhood_Population_2000_2010_2020 using Proportional Symbols. Use a minimum size of 4 and a max size of 32. Use 6 classes for this map.

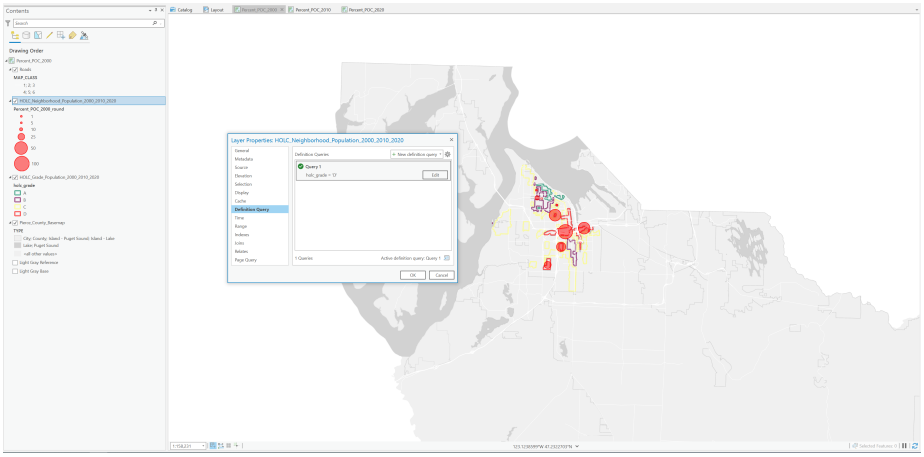


Now, click on the template, and create a symbol of a transparent circle with 50% transparency. Also give it a stroke of 1pt and the same color as the fill without any transparency.



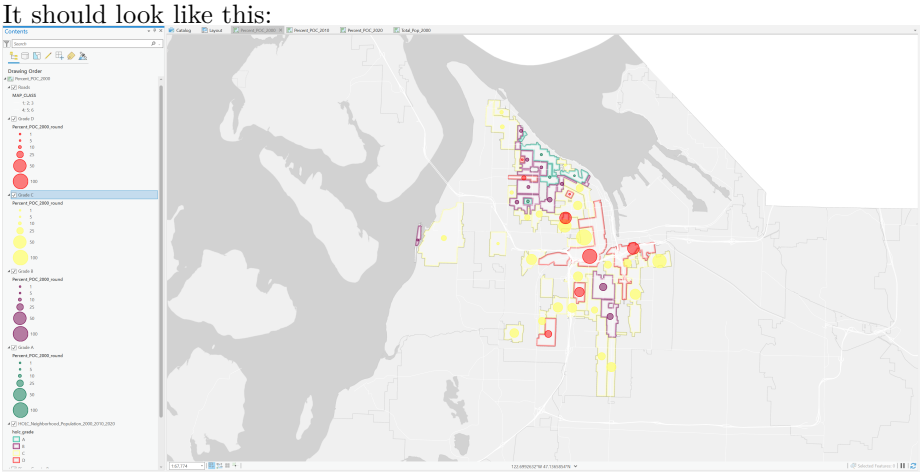
Now here is the tricky part. Since we want to show the distribution for each grade, we are going to apply a definition query to the layer to a single grade, and then copy the layer four times so we have a layer for each grade.

So apply the definition query:



And then copy the layer and paste so you have four layers. Edit the layer so that you have a definition query filtering for each grade, update the symbology so that we are using the same color for the proportional symbols as we used for the grade. Order the layers so your Grade D is on top and Descending to Grade A. We want to do this to make our Grade D's pop since they are what we are interested in examining the effect of redlining on.

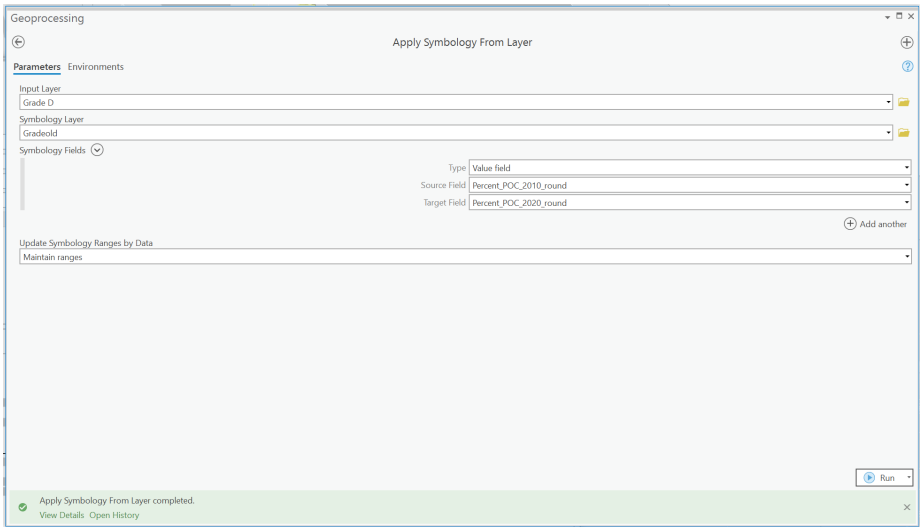
It should look like this:



Once finished, go to your catalog, and copy your map to a new map for the year 2010 and 2020. Repeat this process for each year and update the Proportion Symbols. At the end, you should have three maps that almost look identical, but are only displaying the data by different years.

One last hint, is that you potentially can use the Import Symbolgy tool to be able to apply the ranges to each new year. I will make sure to demo this in class so that you can see it as I had trouble getting the same distribution of features for the Year

2020.

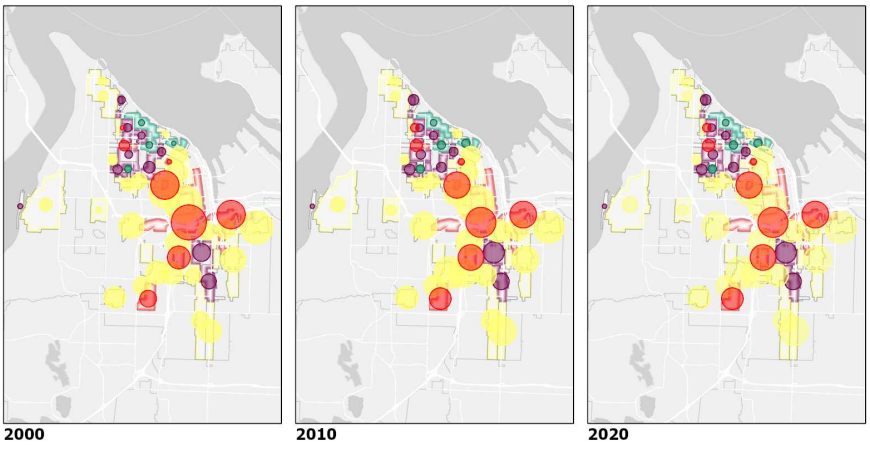


Layout Map 1: Percent Persons Of Color By Year and Neighborhood

Now that we have got our three maps, let's get them into a nice layout so we can compare each of our three years of data.

This is what we are going for:

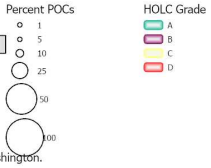
Percent Persons Of Color in HOLC Neighborhoods:
Census Years 2000, 2010, 2020



2000 **2010** **2020**

Percent Persons Of Color

HOLC Grade	2000	2010	2020
A	7	10	15
B	17	19	26
C	30	34	40
D	46	43	48



Citation: IPUMS NHGIS, University of Minnesota, www.nhgis.org. U.S. Census Bureau 2000, 2010, 2020. Pierce County Washington.

Create a new Layout in a landscape format on 8.5 x 11 template.

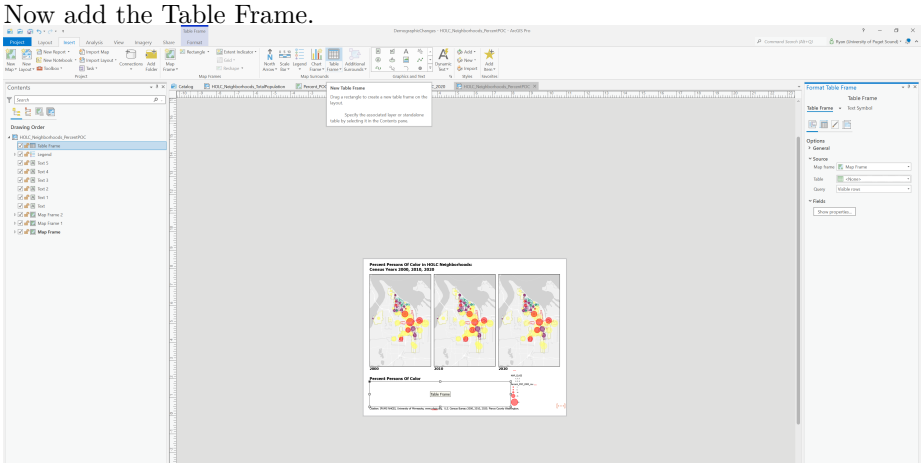
Again it might be helpful to reference the in-class recording of this data.

Add a Title at the top. Add Three Map Frames, equally spaced, each one for the corresponding year starting with 2000, 2010, and 2020. Place a label for each map.

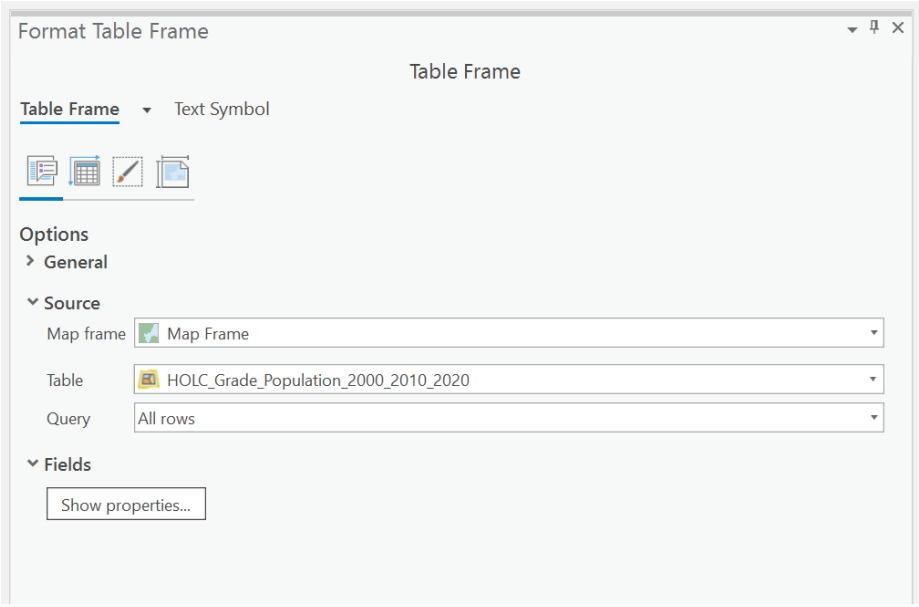
Now, we also should summarize the data for all the grades, as that is a major takeaway as well. To do so, let's use a table element.

First, go back to our Grades data, and make sure to calculate the Percent Persons of Color per Grade.

Now add the Table Frame.



In the Table Properties select any Mapframe and select the HOLC_Grades_Population_2000_2010_2020. Query all, and now this item should be added to the map and work similarly to a



Mapframe.

In the Contents pane, you can now select the columns that you want to show in the table. If you want to change the Headers, go to the HOLC_Grades_Population_2000_2010_2020 layer and change the Alias for the field in the table using the Fields Tool.

Finish off the map by adding a citation and a legend.

For the legend. We only need to know the Proportional Size and the Color for the Grades, so one trick that I did is hid a Proportion Symbol layer without the Color underneath our basemap, then I could use it without showing the color. Rather I added a transparent fill, and gave a black line so that it's easier to represent all the Proportional Symbols layers. Additionally, since we used the same colors from the symbols to the neighborhood polygons, we only need a single representation in our legend.

Whew!

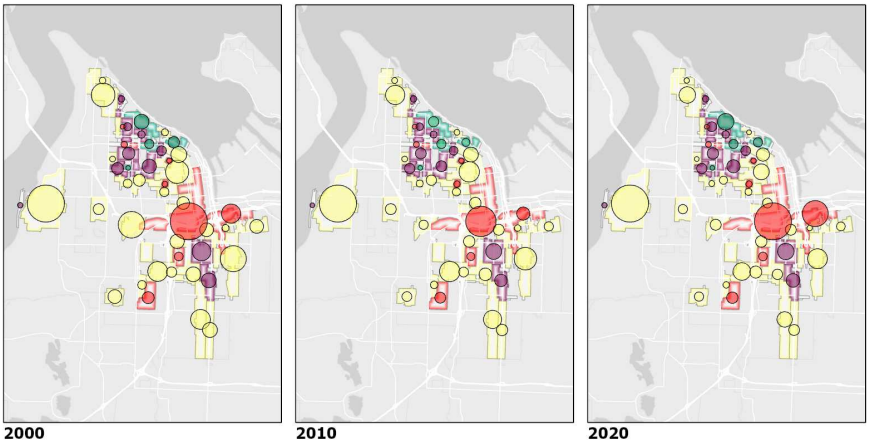
Once done, Export the map. Please upload the pdf to Canvas.

Now on to the second map.

Map 2: Total Population Data

For this map, you will repeat the processes that we did for our first map. However, no calculations are needed for this map, just show the total population. These two maps work together. As in, it's hard to get a good understanding of percentage of a population without showing a total. You map should look very similar to the Percent POC map with all the same elements in the layout.

Total Population of HOLC Neighborhoods:
Census Years 2000, 2010, 2020



Total Population

HOLC Grade	2000	2010	2020
A	7265	5303	7778
B	20960	18824	21549
C	85534	74899	92855
D	20424	18302	24685

Citation: IPUMS NHGIS, University of Minnesota, www.nhgis.org. U.S. Census Bureau 2000, 2010, 2020. Pierce County Washington.

Once done, Export the map. Please upload the pdf to Canvas.

Finally, in your writeup, write up a couple paragraph summary about our Population data and Percentages of Persons of Color in our data over the last twenty years. Are there any noticeable trends, especially as they relate to the HOLC Grade?

Deliverables:

1) WriteUp:
- **What spatial reference systems is our data in?**
- **How many parcels are going to be easy to assign a neighborhood?**
- **How many parcels are going to cause us difficulties? What types of things do we have to be careful about when doing our spatial analysis?**
- **Include a screenshot of your Blocks symbolized as points and add to your write-up. Compared to a spatial join, does this give us better flexibility to assign a block to a given location? - Which neighborhood would you assign GEOID:530530611006007 too?**
- **Compare the results from using the Centroid data vs our geometric operations process. How much more accuracy do we get when using these methods compared to the spatial join out of the box?**
- **Summarize what a Dasymetric Mapping is and how it would be useful for our Tacoma HOLC Neighborhoods?**

- **Population and Persons of Color Data Summary**

2) Percent Persons of Color Map

3) Total Population Data Map