

PNNL-32735

Better Buildings Workforce Accelerator – Technical Assistance Project

Northwest Energy Efficiency Council –
A Replicable Analysis for Building
Operator Certification Diversity, Equity,
and Inclusion Objectives

March 2022

Adrienne Rackley

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor Battelle Memorial Institute, nor any of their employees, makes **any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights.** Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or Battelle Memorial Institute. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

PACIFIC NORTHWEST NATIONAL LABORATORY
operated by
BATTELLE
for the
UNITED STATES DEPARTMENT OF ENERGY
under Contract DE-AC05-76RL01830

Printed in the United States of America

Available to DOE and DOE contractors from the
Office of Scientific and Technical Information,
P.O. Box 62, Oak Ridge, TN 37831-0062;
ph: (865) 576-8401
fax: (865) 576-5728
email: mb-reports@osti.gov

Available to the public from the National Technical Information Service
5301 Shawnee Rd., Alexandria, VA 22312
ph: (800) 553-NTIS (6847)
email: orders@ntis.gov <<https://www.ntis.gov/about>>
Online ordering: <http://www.ntis.gov>

Better Buildings Workforce Accelerator – Technical Assistance Project

Northwest Energy Efficiency Council – A Replicable Analysis for Building
Operator Certification Diversity, Equity, and Inclusion Objectives

March 2022

Adrienne Rackley

Prepared for
the U.S. Department of Energy
under Contract DE-AC05-76RL01830

Pacific Northwest National Laboratory
Richland, Washington 99354

Acronyms and Abbreviations

ACS	American Community Survey
AQI	Air Quality Index
BOC	Building Operator Certification program
CDC	Centers for Disease Control and Prevention
DEI	Diversity, Equity, and Inclusion
DOE	Department of Energy
EPA	Environmental Protection Agency
NEEC	Northwest Energy Efficiency Council
OAQPS	EPA's Office of Air Quality Planning and Standards
OAR	EPA's Office of Air and Radiation
PM2.5	Particulate Matter 2.5
PNNL	Pacific Northwest National Laboratory
TCOC	Training Certificate of Completion

Contents

Acronyms and Abbreviations	ii
1.0 Introduction	1
2.0 Analytical Approach	2
2.1 Limitations	4
3.0 Results	5
3.1 Race and Ethnicity	5
3.2 Linguistic Isolation	9
3.3 Low Income	11
3.4 Education	13
3.5 Air Quality (Ozone)	15
3.6 Social Vulnerability Index	16
3.7 Gender	18
4.0 Replicating Analysis	21
4.1 Data Preparation and Cleaning in Excel	21
4.2 Importing Data into Power BI	21
4.3 Building the Map in Power BI	28
4.3.1 Adding Reference Layers	30
4.3.2 Adding Filters	30
4.3.3 Appending New Data from Subsequent Years	31
5.0 Recommendations	35
6.0 References	36

Figures

Figure 1.	2012–2020 Participant Location by Race and Ethnicity by County, National View. Race and Ethnicity Reference Layer: ACS Race and Hispanic Origin Variables – Boundaries (Esri 2022).	6
Figure 2.	2012–2020 Participant Location by Race and Ethnicity by County, Washington State Detail. Race and Ethnicity Reference Layer: ACS Race and Hispanic Origin Variables – Boundaries (Esri 2022).	7
Figure 3.	2012–2020 Participant Location by Race and Ethnicity by Census Tract, Seattle Area Detail. Race and Ethnicity Reference Layer: ACS Race and Hispanic Origin Variables – Boundaries (Esri 2022).	8
Figure 4.	2012–2020 Participant Location by Linguistic Isolation by County, Washington State Detail. Linguistic Isolation Reference Layer: English Ability and Linguistic Isolation Variables – Boundaries (Esri 2022).	9

Figure 5.	2012–2020 Participant Location by Linguistic Isolation by Census Tract, Seattle Area Detail. Linguistic Isolation Reference Layer: English Ability and Linguistic Isolation Variables – Boundaries (Esri 2022).	10
Figure 6.	2012–2020 Participant Location by Income by County, Washington State Detail. Income Reference Layer: ACS Poverty Status Variables – Boundaries (Esri 2022).....	11
Figure 7.	2012–2020 Participant Location by Income by Census Tract, Seattle Area Detail. Income Reference Layer: ACS Poverty Status Variables – Boundaries (Esri 2022).....	12
Figure 8.	2012–2020 Participant Location by Education by County, Washington State Detail. Education Reference Layer: ACS Educational Attainment Variables – Boundaries (Esri 2022).....	13
Figure 9.	2012–2020 Participant Location by Education by Census Tract, Seattle Area Detail. Education Reference Layer: ACS Educational Attainment Variables – Boundaries (Esri 2022).....	14
Figure 10.	2012–2020 Participant Location by Air Quality (Ozone), Washington State Detail. Air Quality (Ozone) Reference Layer: AirNow – AQI Contours (Last 24 hours, Ozone only) (U.S. EPA 2020).....	15
Figure 11.	2012–2020 Participant Location by Air Quality (Ozone), Seattle Area Detail. Air Quality (ozone) Reference Layer: AirNow – AQI Contours (Last 24 hours, Ozone only) (U.S. EPA 2020).....	16
Figure 12.	2012–2020 Participant Location by Social Vulnerability Index, Washington State Detail. Social Vulnerability Reference Layer: 2018 Social Vulnerability Index (CDC 2021).....	17
Figure 13.	2012–2020 Participant Location by Social Vulnerability Index, Seattle Area Detail. Social Vulnerability Reference Layer: 2018 Social Vulnerability Index (CDC 2021).....	18
Figure 14.	2012–2020 Percentage of Participant Gender by TCOC Level (participants with missing data in the gender field are not shown).	19
Figure 15.	2012–2020 Count of Participant Gender by TCOC Year (participants with missing data in the gender field are not shown).	20

Tables

Table 1.	Desired Participant Data and Registration Form Fields	2
Table 2.	Participant Location Data	2
Table 3.	Indicators and Reference Layers	3

1.0 Introduction

The Northwest Energy Efficiency Council (NEEC) requested technical assistance through the Department of Energy (DOE) Better Buildings Workforce Accelerator program, with the objective of developing a replicable analytical approach to understanding Building Operator Certification (BOC) program participants and measuring progress toward diversity, equity, and inclusion objectives. The analysis advances NEEC's progress toward their goal of recruiting more diverse students to the BOC training and certification program. In particular, the analysis maps disadvantaged communities and participant location to assess how disadvantaged communities are represented in the student population and identify regions that may benefit from targeted outreach.

This work builds on a previous report analyzing participant proximity to pollution and environmental hazards and socioeconomic status (Tidwell 2021). The report generated insights into participant locations by manually matching participant ZIP codes to census data, data from the Environmental Protection Agency (EPA) Environmental Justice Screening & Mapping Tool, and the Washington Environmental Health Disparities Map. Manual matching of records is not a feasible approach for large datasets, and the current study sought an analysis approach that NEEC could more efficiently replicate and update annually. Recommendations for benchmarking and outreach presented by Tidwell are applicable to the current analysis.

This technical assistance develops a replicable analysis of BOC program participants based on Diversity, Equity, and Inclusion (DEI) indicators that can be applied at a national level, provides a step-by-step guide on repeating and updating the analysis, and presents recommendations for future data collection and analysis. It is broken into four sections. The analytical approach explains the data set, different options for DEI indicators, approaches to linking the disparate datasets, and selected DEI indicators. Findings presents the results maps for each selected indicator and a brief interpretation of the data. The replicating analysis section is a step-by-step how-to guide for recreating the analysis and data visualizations, including options for filtering and updating the underlying data set. Recommendations are made for updating the registration form and alternative DEI indicators.

2.0 Analytical Approach

NEEC provided registration form and Training Certificate of Completion (TCOC) records for 1,538 participants from 2010–2021. The following key DEI indicators were prioritized for analysis:

- Race and ethnicity (other than White non-Hispanic)
- Linguistic isolation
- Low income
- Environmental health risk
- Education (less than high school)
- Air quality (ozone)
- Gender.

A review of the registration form summarized in Table 1 revealed that gender was an optional field that participants could complete. This allowed direct calculation of gender participation rates. Other DEI indicators required a different approach, as the data was not collected on the registration form.

Table 1. Desired Participant Data and Registration Form Fields

Desired Participant Data	Participant Data Collected on Registration Form
Race and ethnicity	No
Linguistic isolation	No
Low income	No
Environmental health risk	No
Education	No
Air quality (ozone)	No
Gender	Yes

We explored the possibility of directly relating the remaining indicators to census data through a common key. A location field is commonly used to relate two data sets in this scenario. City, state, and ZIP code were available for participant location data (see Table 2).

Table 2. Participant Location Data

Geographic Data Point	Collected on Registration Form
Street Address	No
City	Yes
County	No
State	Yes
ZIP	Yes

The available participant data by city/ZIP is not directly relatable to census data, which is reported at block group, census tract, county, and state levels (among others). State was

deemed to be too broad a measure. Previous work with this data looked up ZIP codes using EJ Screen, Washington State Environmental Health Risk Score, and census data, requiring manual lookup for each record (Tidwell 2021). While this method provides detailed data, the time required to replicate this on an annual basis is not practical.

Rather than producing data in a table format requiring manual data entry, the selected approach uses geospatial mapping tools to plot participant ZIP code over a base map of DEI indicators. The use of a geospatial application bypasses the need to relate two tables of data on a common location field. ArcGIS Maps for Power BI was selected as the mapping platform because it didn't require NEEC to purchase additional software and allowed for the underlying data to be refreshed and offered interactivity and filtering. The participant data set was cleaned in Excel.

Participants were allowed to enter either a home location, work location, or both on the registration form; 63 percent selected work as a primary location, while 37 percent selected home as primary. The analysis used the primary location ZIP code field because home location was often blank.

ArcGIS Maps for Power BI offers publicly available reference layers through ArcGIS Living Atlas of the World. Utilizing these reference layers streamlines and simplifies the process, eliminating the need to develop custom reference layers. As shown in Table 3, public base maps were located for all but one indicator. An additional composite indicator was added to deliver DEI insights in a single visual.

Table 3. Indicators and Reference Layers

Indicator	National-Level Reference Layer
Race and ethnicity	ACS Race and Hispanic Origin Variables – Boundaries (Esri 2022)
Linguistic isolation	ACS English Ability and Linguistic Isolation Variables – Boundaries (Esri 2022)
Low income	ACS Poverty Status Variables – Boundaries (Esri 2022)
Environmental health risk	None identified, see recommendations
Education	ACS Educational Attainment Variables – Boundaries (Esri 2022)
Air quality (ozone)	AirNow – AQI Contours (Last 24 hours, Ozone only) U.S. EPA, OAR, OAQPS (U.S. EPA, OAR, OAQPS, 2020)
Combined social vulnerability reference layer	CDC Social Vulnerability Index 2018 – USA (CDC, 2021)

Participant Primary ZIP/postal code was represented as points on a map over a DEI indicator reference layer. The reference layer's level of detail varies as you zoom in and out, displaying state, county, and census tract boundaries.

Filter options were added to isolate for

- Level 1 TCOC Year
- Level 2 TCOC Year
- New Cert Year.

2.1 Limitations

While a geospatial approach bridges the data gap on individual participants, there are drawbacks. It is important to note that participant location does not mean they are members of target demographic and socioeconomic groups.

Differences in how data is tied to a location (whether data is reported at a home or work location) can cause problems interpreting the results. BOC participants were given the option to enter either a home or work location on the registration form. Most participants (63 percent) selected a work location as primary. Census data is reported by home location. A participant whose work location is primary and who lives in a different tract or county may not be represented in the work location's census records on demographic and socioeconomic indicators.

Mapping participant location by ZIP code is a rough estimate of location. Particularly where a ZIP code contains multiple census tracts, relying on a participant location point in the census tract detail view may be misleading.

A geospatial approach is useful to identify high-level trends, but it is difficult to tell what number or percentage of participants are from target areas. If a high degree of accuracy is not needed, this drawback can be overlooked.

Finally, reliance on publicly available reference layers narrows options for indicators and resulted in no reference layer identified for environmental health risk. This is an instance where a custom reference layer would be needed.

3.0 Results

The majority of participants live and/or work in Washington State and primarily the Seattle-Tacoma area, as shown in Figure 1. Within Washington State, there are clusters of participants in large and mid-sized cities (Wenatchee, Yakima, Tri-Cities, Spokane). Very few people participate in the program outside of Washington, Oregon, and Idaho. No participants are located outside of the continental United States.

Within Washington State, participants are more likely to come from counties that are predominantly White and non-Hispanic, experience lower levels of poverty than the national average, are more highly educated than the national average, and are overall in less vulnerable areas. A more nuanced view of the greater Seattle-Tacoma area at the census tract level shows fairly even distribution of participants across census tracts with varying scores. Linguistic isolation is more prevalent in the Seattle area, as well as more rural agricultural communities in central and southern Washington.

Enrollment of women is far lower than that of men. Women make up 11.82 percent of all TCOC Level 1 participants who reported gender. The percentage of female participants drops to 8.95 percent at TCOC Level 2.

The BOC program may have opportunities for expansion in more vulnerable areas south of Seattle and in rural agricultural areas in central and southern Washington. These areas have higher vulnerability scores across indicators and lower participation. Although many enrollees are from the Seattle-Tacoma area, vulnerable census tracts exist in these urban areas and bordering large cities. South Puget Sound; some areas north of Seattle, such as Northgate, Everett, and Fairmont; the Spokane Valley; and North Spokane all may benefit from targeted outreach and enrollment.

3.1 Race and Ethnicity

Participants primarily live and work in areas that are predominantly White and non-Hispanic. Smaller clusters of participants appear in mid-sized cities (in particular Yakima and the Tri-Cities area) located in or bordering areas in southern and southeastern Washington with predominantly Hispanic or Latino populations (see Figure 2). Southern and southeastern Washington are agricultural areas with higher concentrations of migrant workers. Adams County, Franklin County, and Yakima County are primarily Hispanic or Latino and are the only three counties in Washington that are not primarily non-Hispanic White.

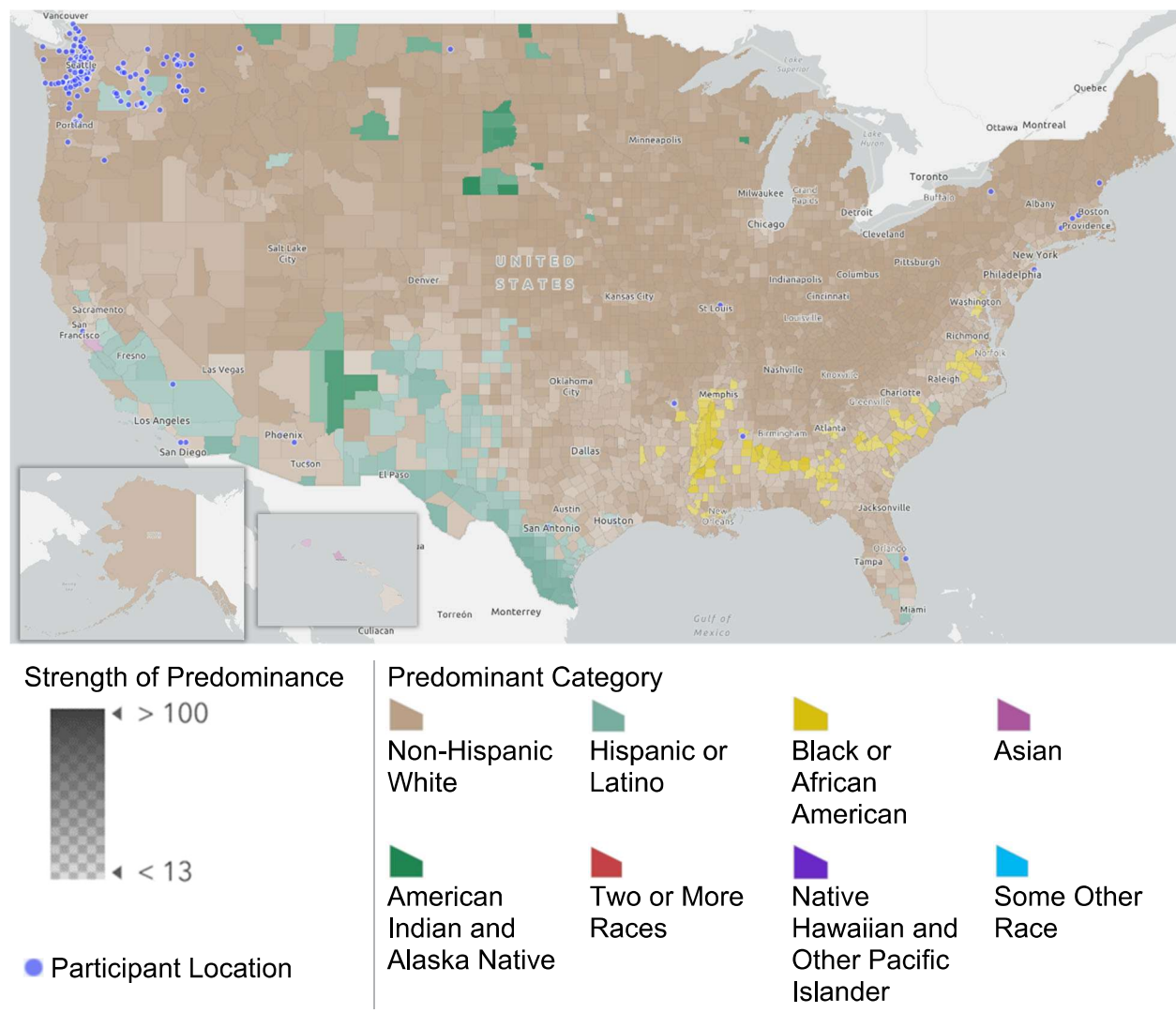


Figure 1. 2012–2020 Participant Location by Race and Ethnicity by County, National View. Race and Ethnicity Reference Layer: ACS Race and Hispanic Origin Variables – Boundaries (Esri 2022).

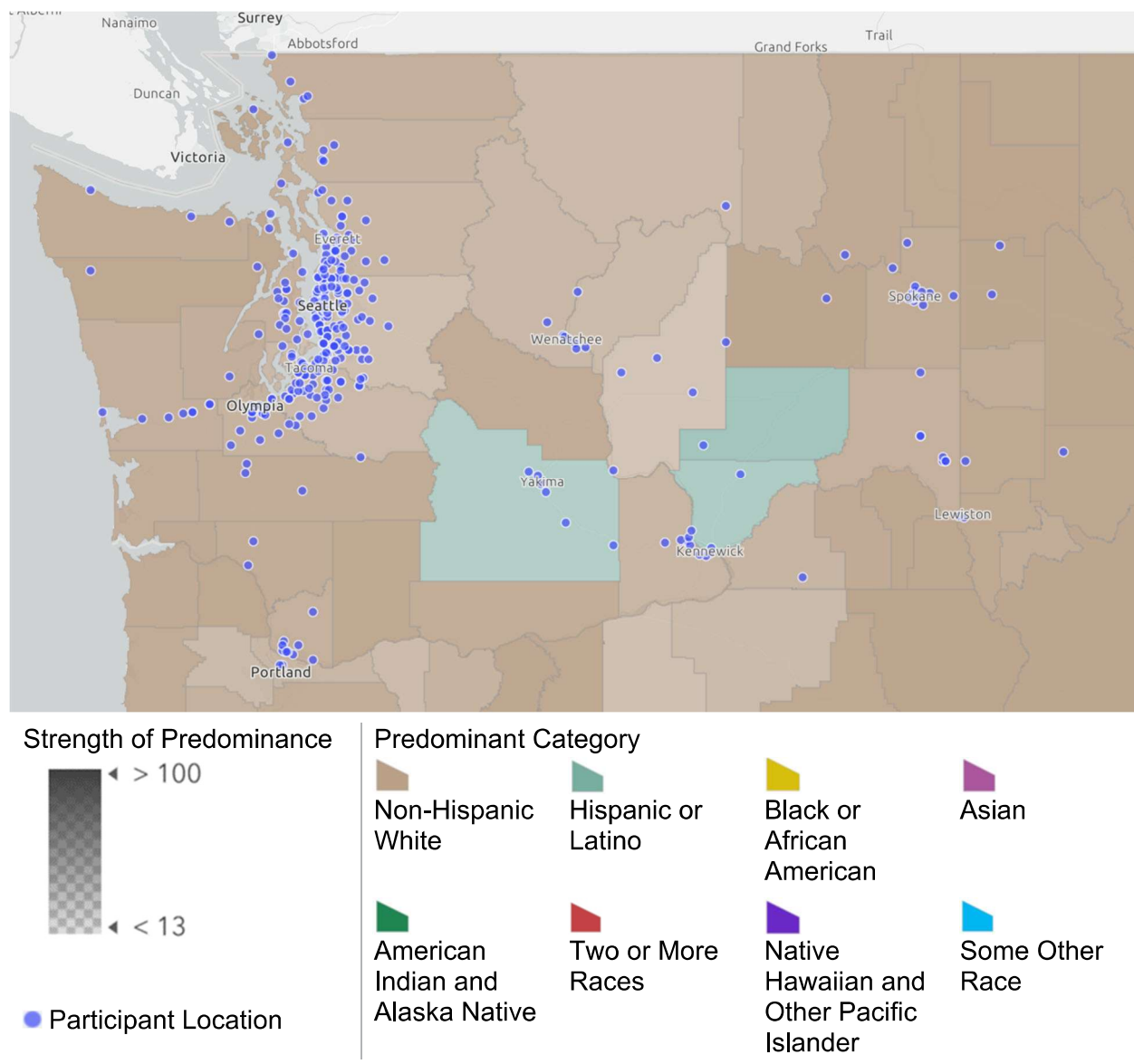


Figure 2. 2012–2020 Participant Location by Race and Ethnicity by County, Washington State Detail. Race and Ethnicity Reference Layer: ACS Race and Hispanic Origin Variables – Boundaries (Esri 2022).

The census tract-level view of the Seattle area shows a more racially and ethnically diverse view (see Figure 3). However, participants are mainly located in non-Hispanic White tracts, with few in the tracts that are shown as primarily Black or African American, Asian, or Hispanic or Latino.

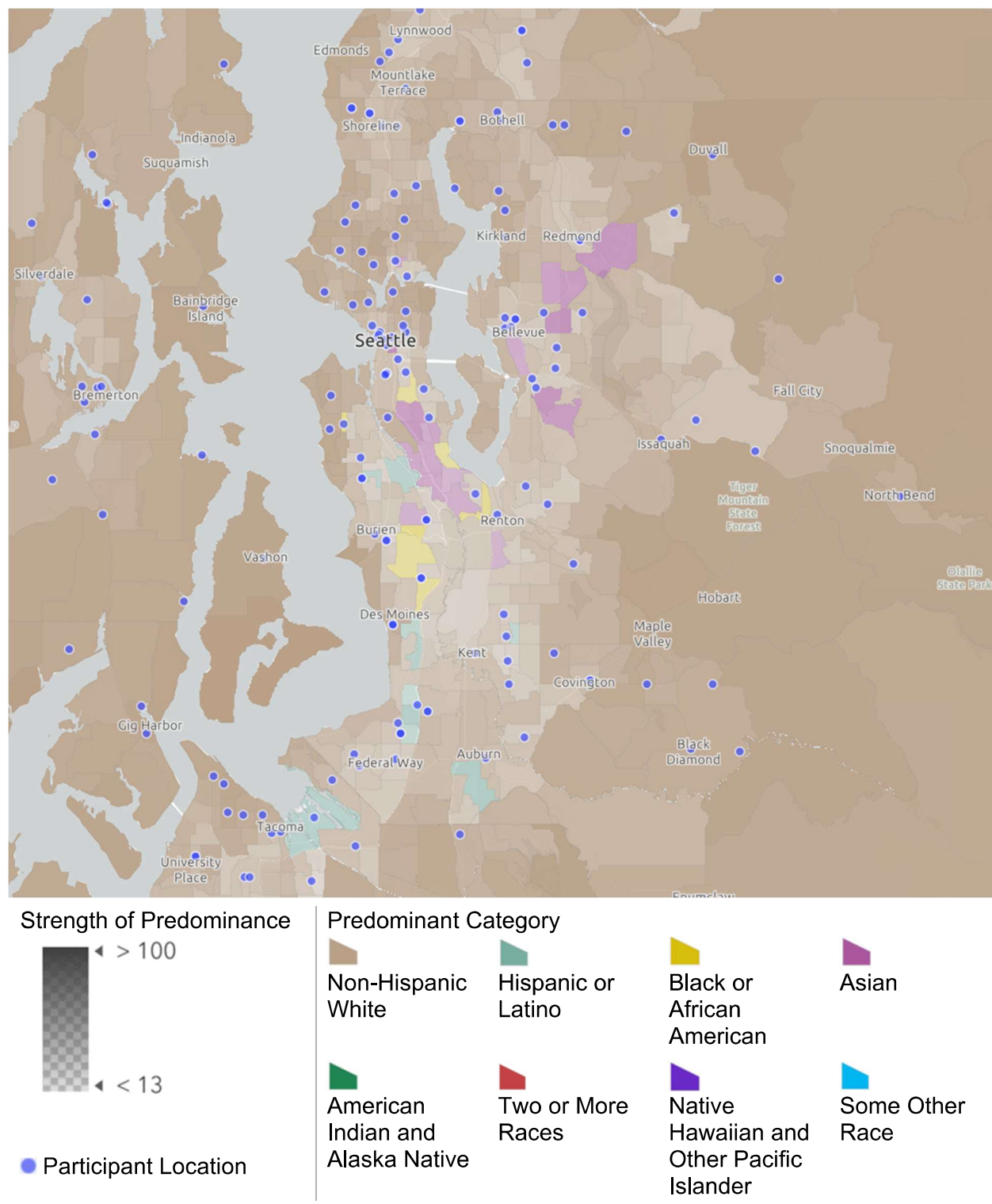


Figure 3. 2012–2020 Participant Location by Race and Ethnicity by Census Tract, Seattle Area Detail. Race and Ethnicity Reference Layer: ACS Race and Hispanic Origin Variables – Boundaries (Esri 2022).

3.2 Linguistic Isolation

Many counties in Washington rank at or above the national average for linguistic isolation. At the county level shown in Figure 4, most participants on the west side of the state are located in counties near the national average. South-central Washington counties (Adams, Grant, and Yakima) have a high percentage of adults with limited English ability.

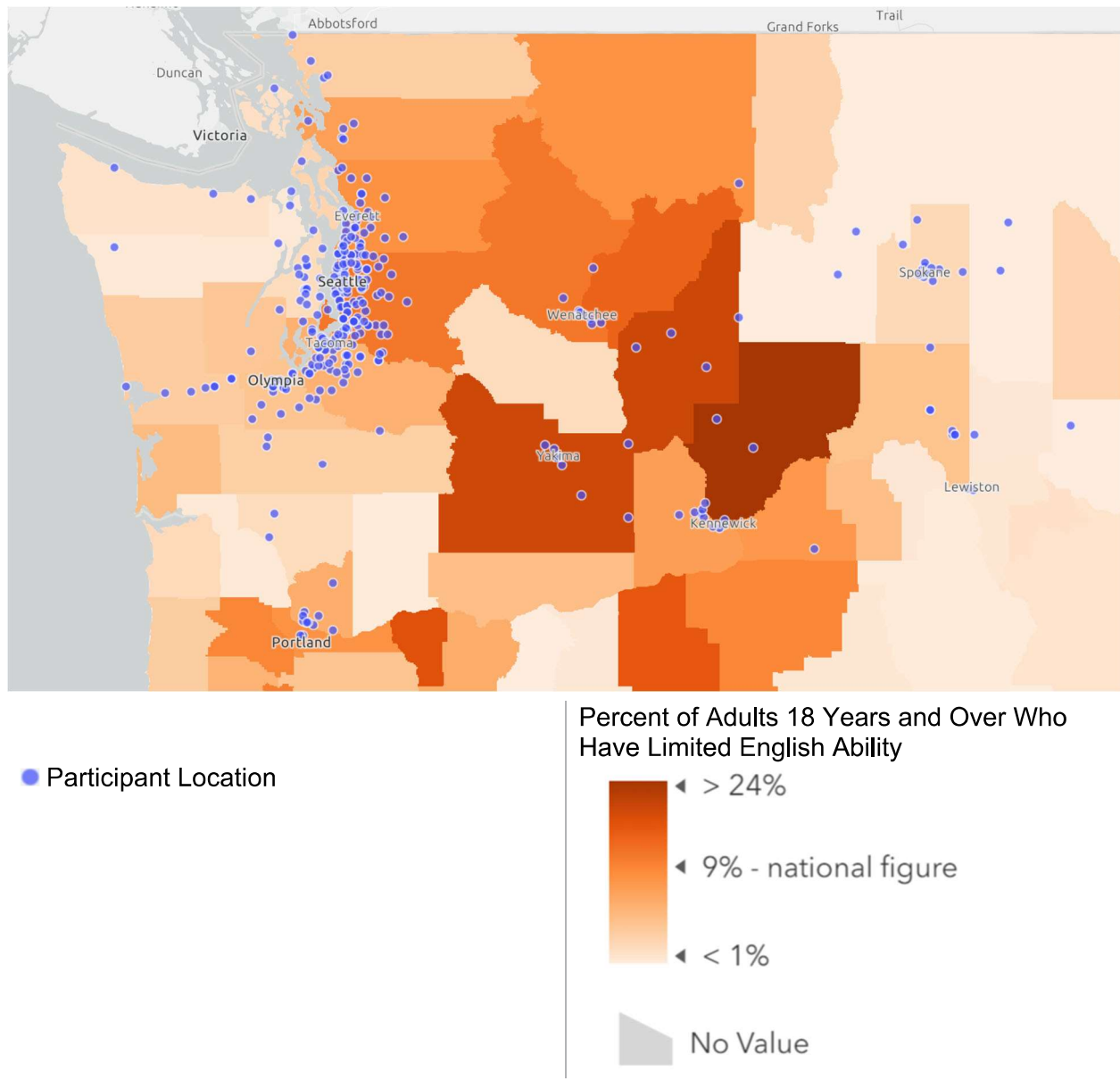


Figure 4. 2012–2020 Participant Location by Linguistic Isolation by County, Washington State Detail. Linguistic Isolation Reference Layer: English Ability and Linguistic Isolation Variables – Boundaries (Esri 2022).

The census tract-level view of the Seattle area in Figure 5 reveals that many tracts are linguistically isolated. Participants appear to be evenly distributed throughout areas with varying degrees of English language ability.

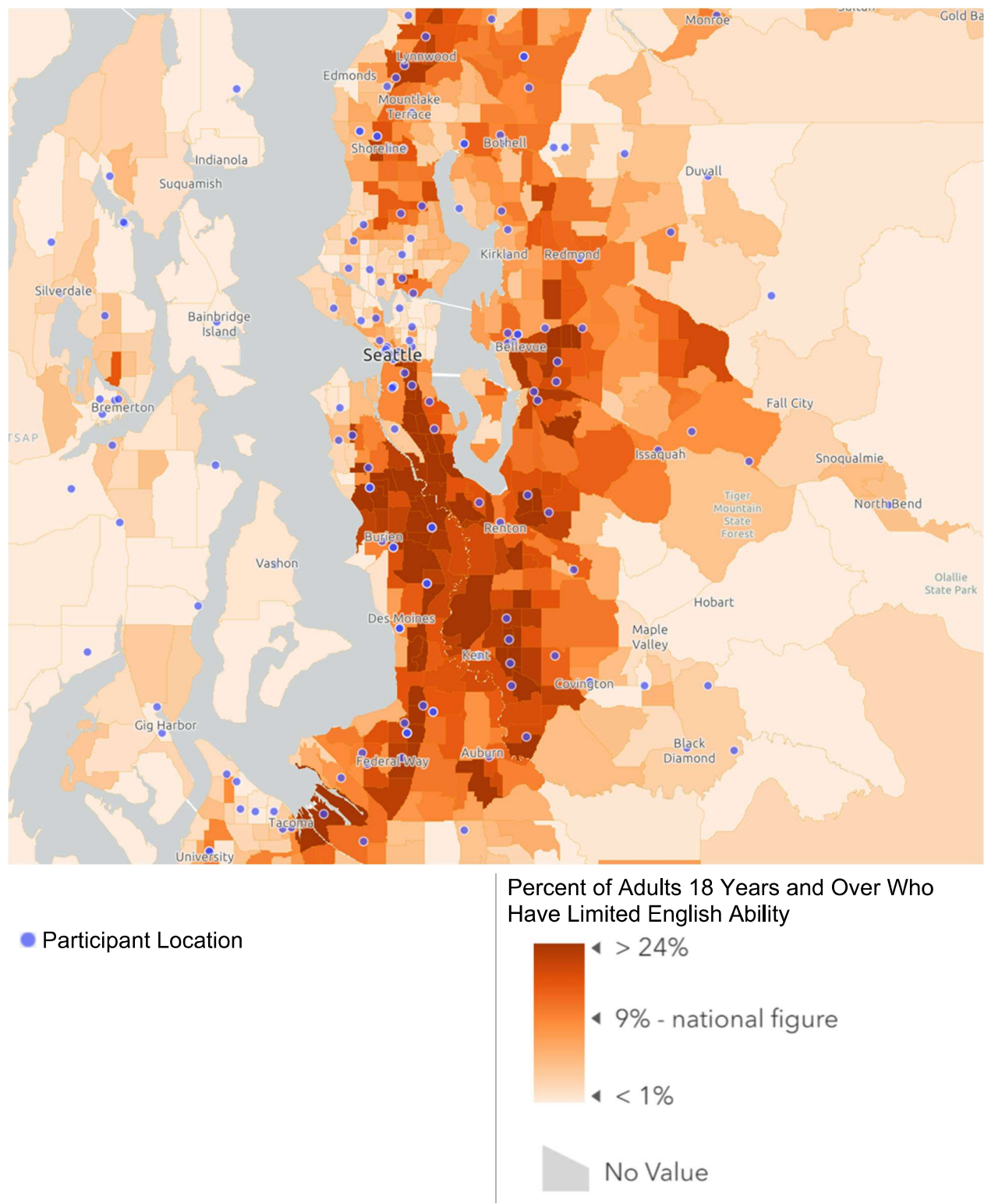


Figure 5. 2012–2020 Participant Location by Linguistic Isolation by Census Tract, Seattle Area Detail. Linguistic Isolation Reference Layer: English Ability and Linguistic Isolation Variables – Boundaries (Esri 2022).

3.3 Low Income

The Seattle metropolitan area ranks below the national average for percentage of population below poverty. As a result, at the county-level view in Figure 6, most participants are located in wealthier counties. Whitman County is the most impoverished in the state, with approximately 26 percent of the population below poverty. Neighboring Adams County (to the west) is the second-most impoverished county at 20.6 percent below poverty. Four participants are located in these two heavily agricultural counties.

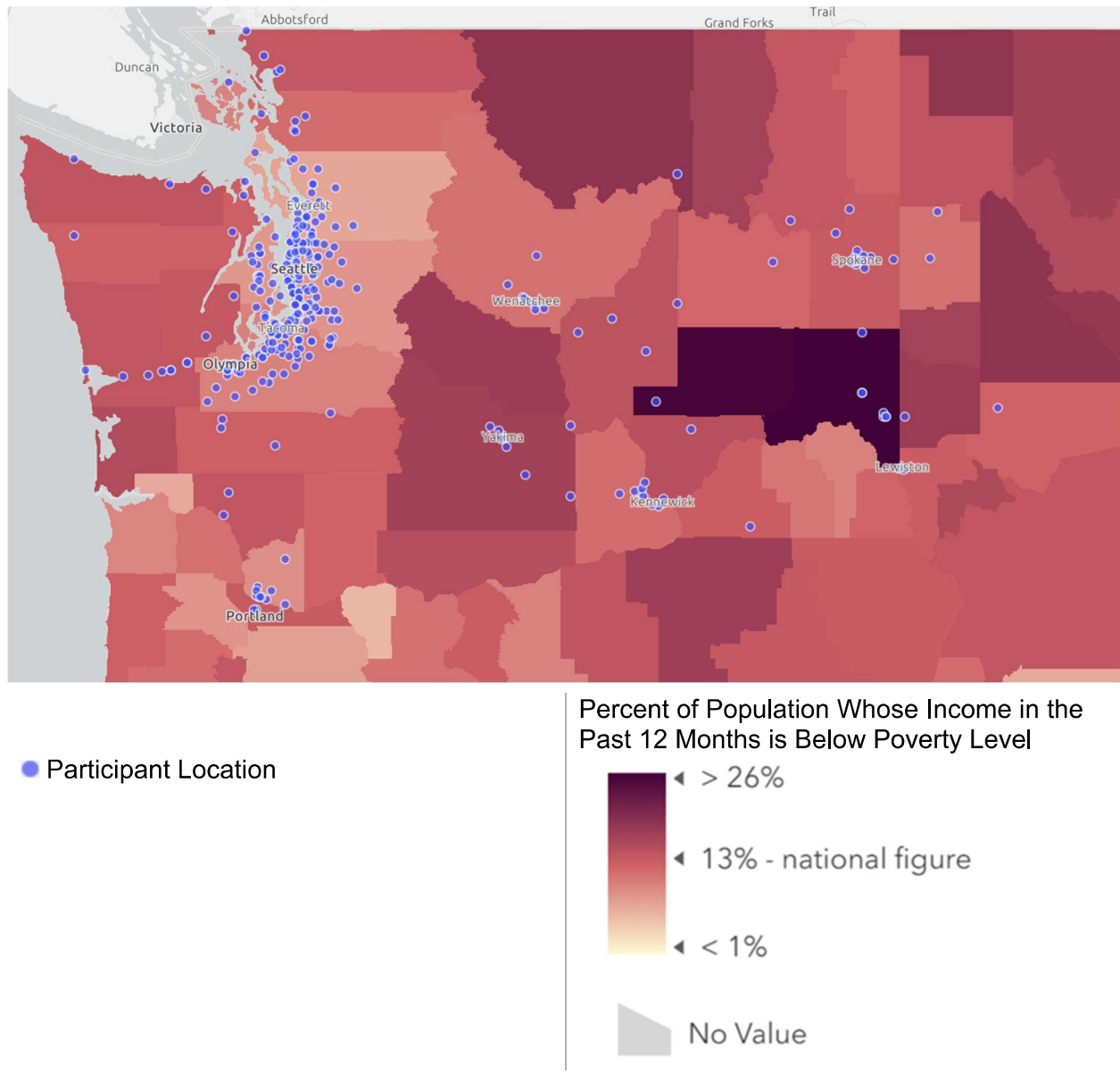


Figure 6. 2012–2020 Participant Location by Income by County, Washington State Detail. Income Reference Layer: ACS Poverty Status Variables – Boundaries (Esri 2022).

The Seattle area shows varying poverty levels scattered across census tracts (Figure 7). In general, census tracts with higher poverty rates are south of Seattle and the South Sound area. Participants do not appear to be concentrated in either high- or low-poverty areas.

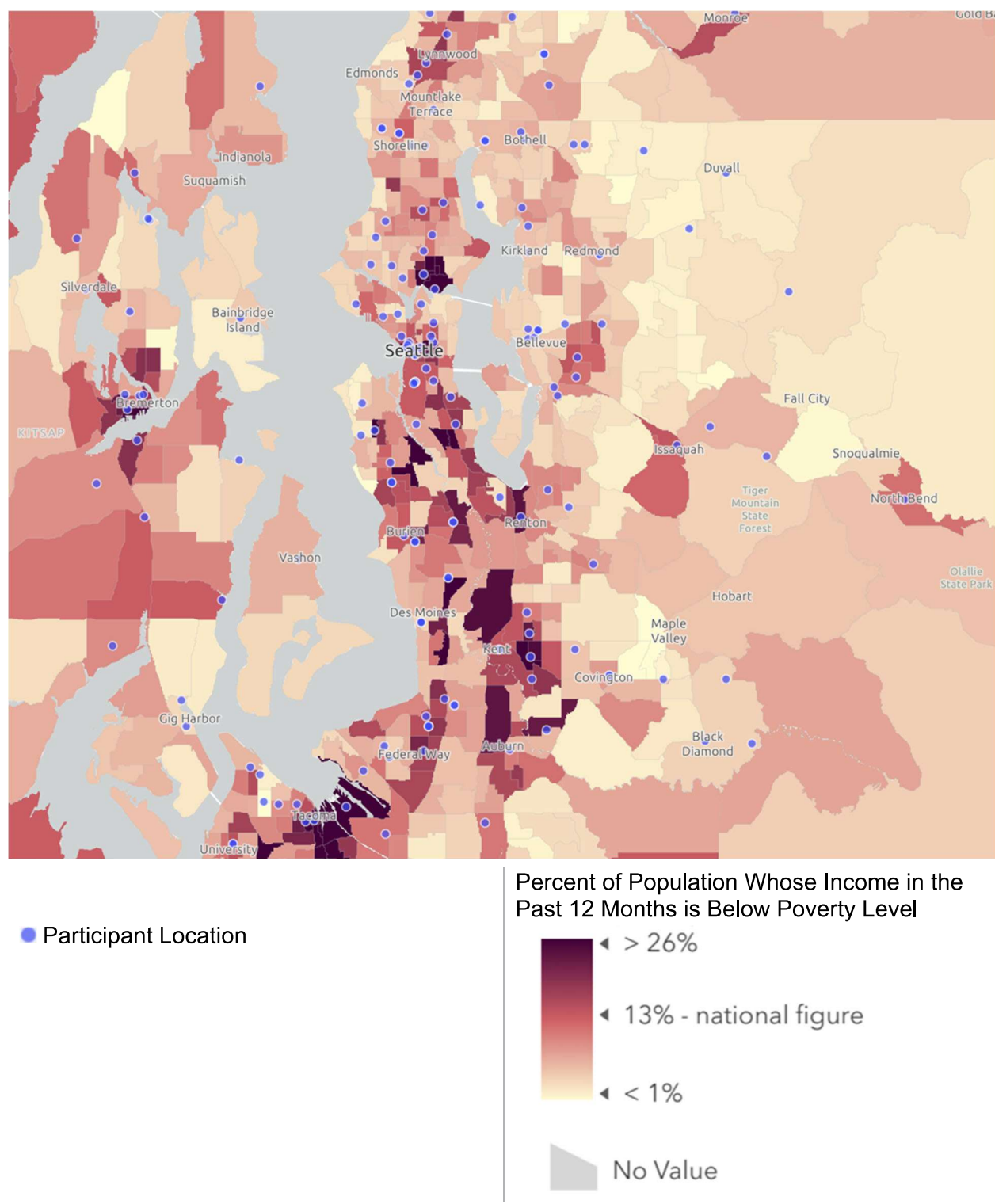


Figure 7. 2012–2020 Participant Location by Income by Census Tract, Seattle Area Detail. Income Reference Layer: ACS Poverty Status Variables – Boundaries (Esri 2022).

3.4 Education

Findings for education are similar to patterns in income and linguistic isolation. At the county level, the Seattle metropolitan area ranks below the national average for percentage of population with less than a high school education. As shown in Figure 8, most participants are located in counties that are better educated. Adams County again ranks the lowest in the state, with approximately 31.1 percent of residents without a high school diploma. Close behind are Yakima County at 25.7 percent and Grant County at 21.5 percent.

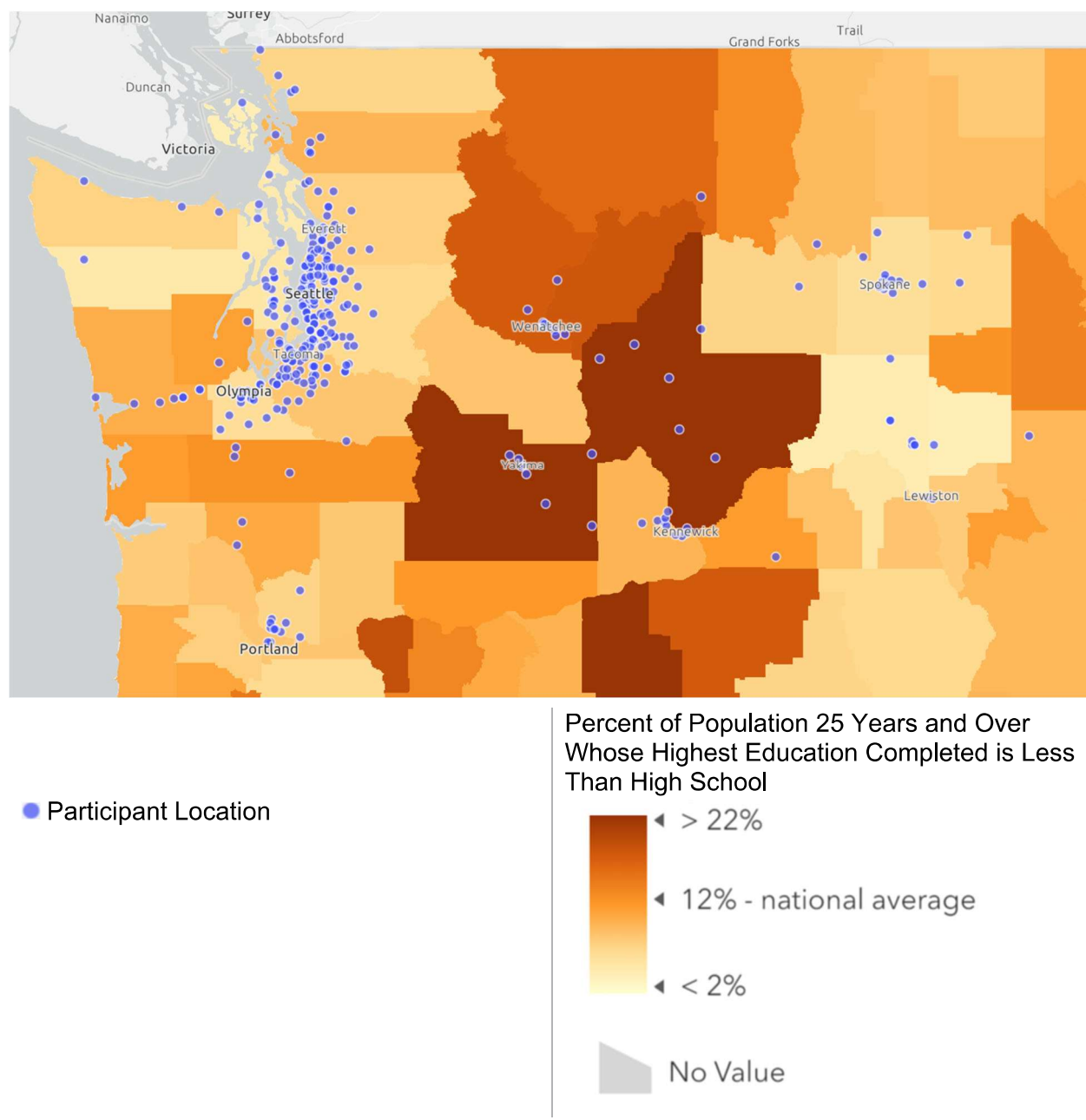


Figure 8. 2012–2020 Participant Location by Education by County, Washington State Detail. Education Reference Layer: ACS Educational Attainment Variables – Boundaries (Esri 2022).

At the Seattle area census tract level, the South Sound area and south of Seattle in general is more likely to have less than a high school diploma. Participants appear to be relatively evenly split between differently ranked areas, leaning more toward highly educated tracts.

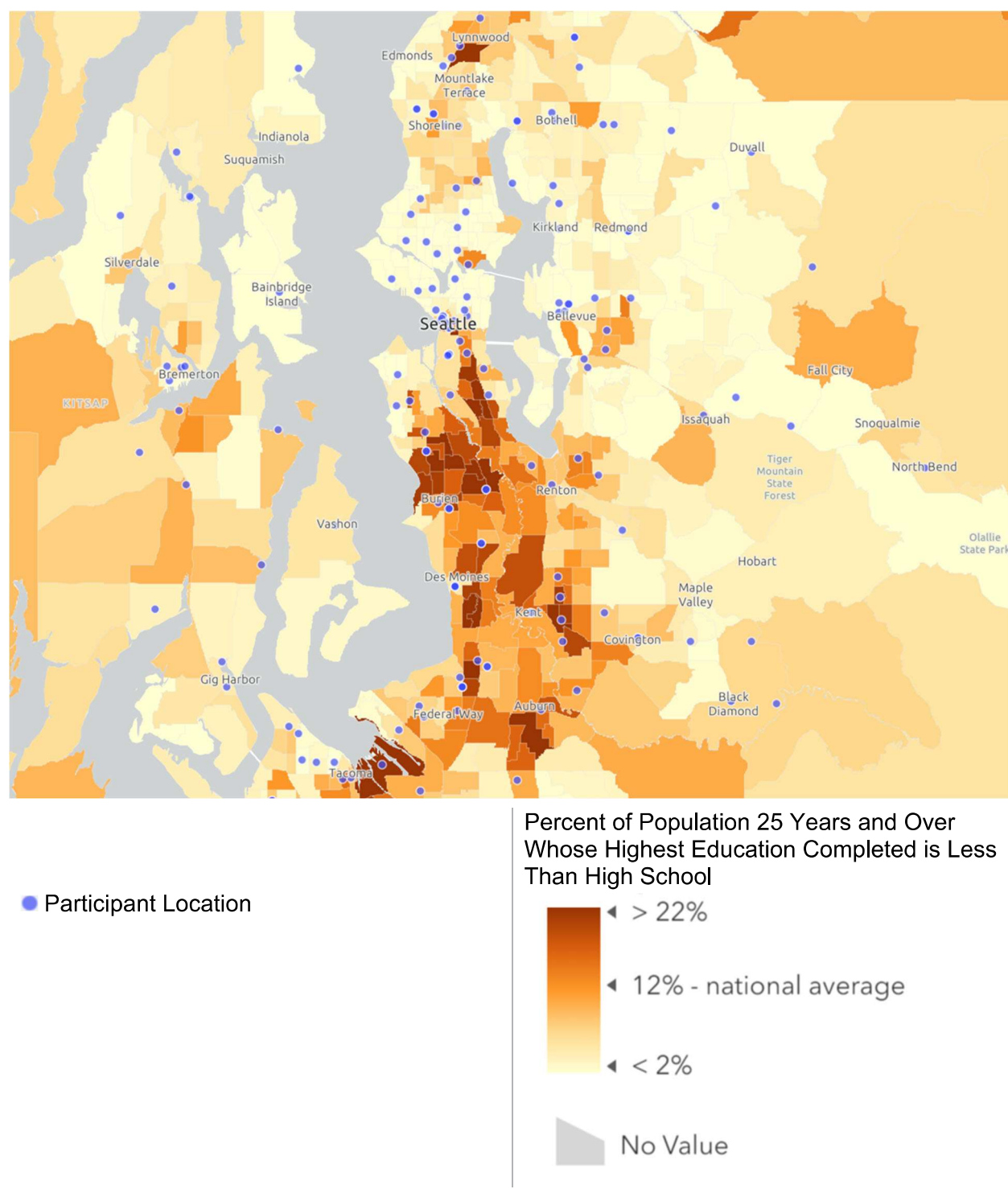


Figure 9. 2012–2020 Participant Location by Education by Census Tract, Seattle Area Detail. Education Reference Layer: ACS Educational Attainment Variables – Boundaries (Esri 2022).

3.5 Air Quality (Ozone)

The air quality reference layer has the ability to be refreshed with new data every 24 hours. At the time the images in Figure 10 and Figure 11 were captured, the air quality consistently ranked “Good” for the region.

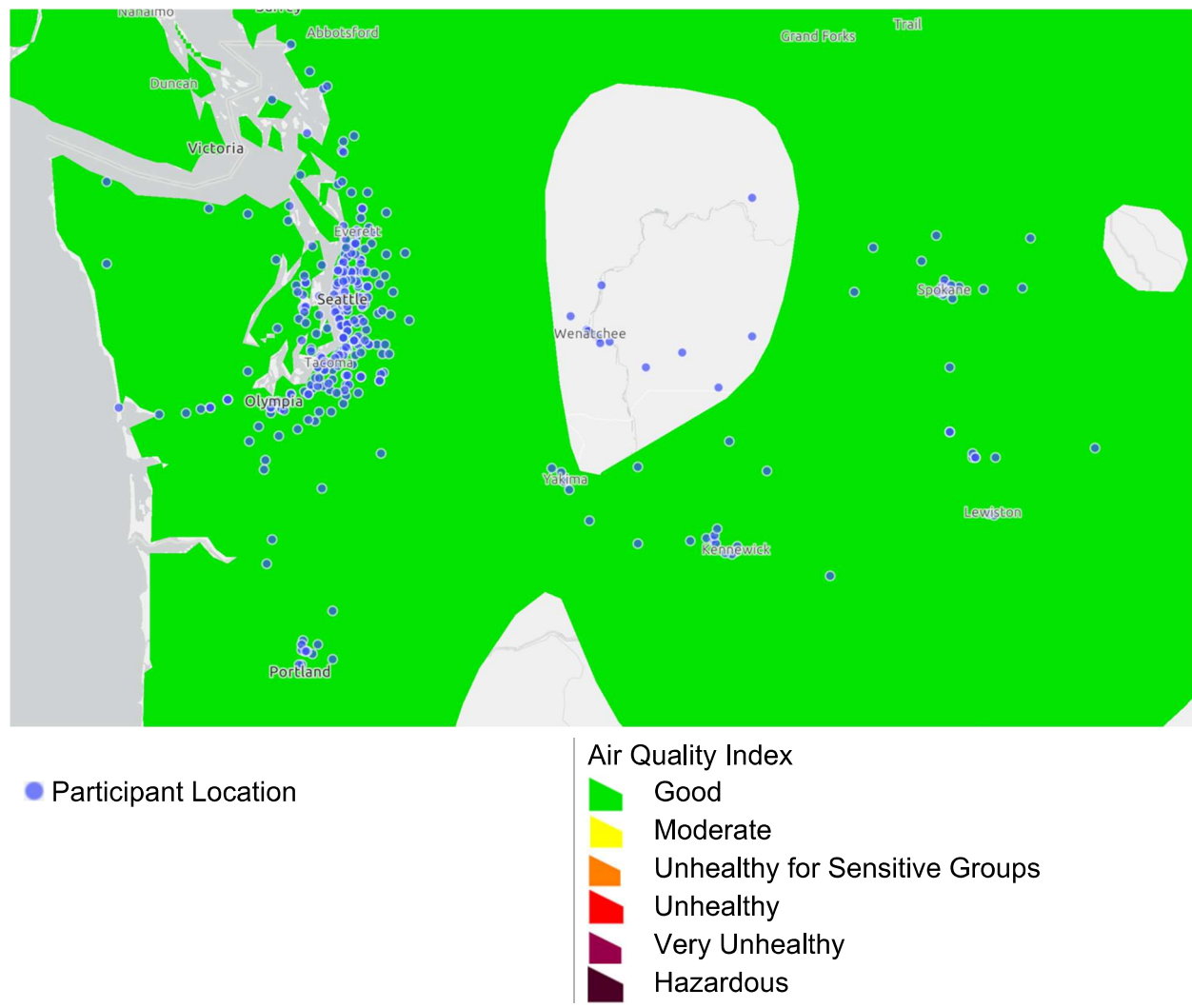


Figure 10. 2012–2020 Participant Location by Air Quality (Ozone), Washington State Detail. Air Quality (Ozone) Reference Layer: AirNow – AQI Contours (Last 24 hours, Ozone only) (U.S. EPA 2020).

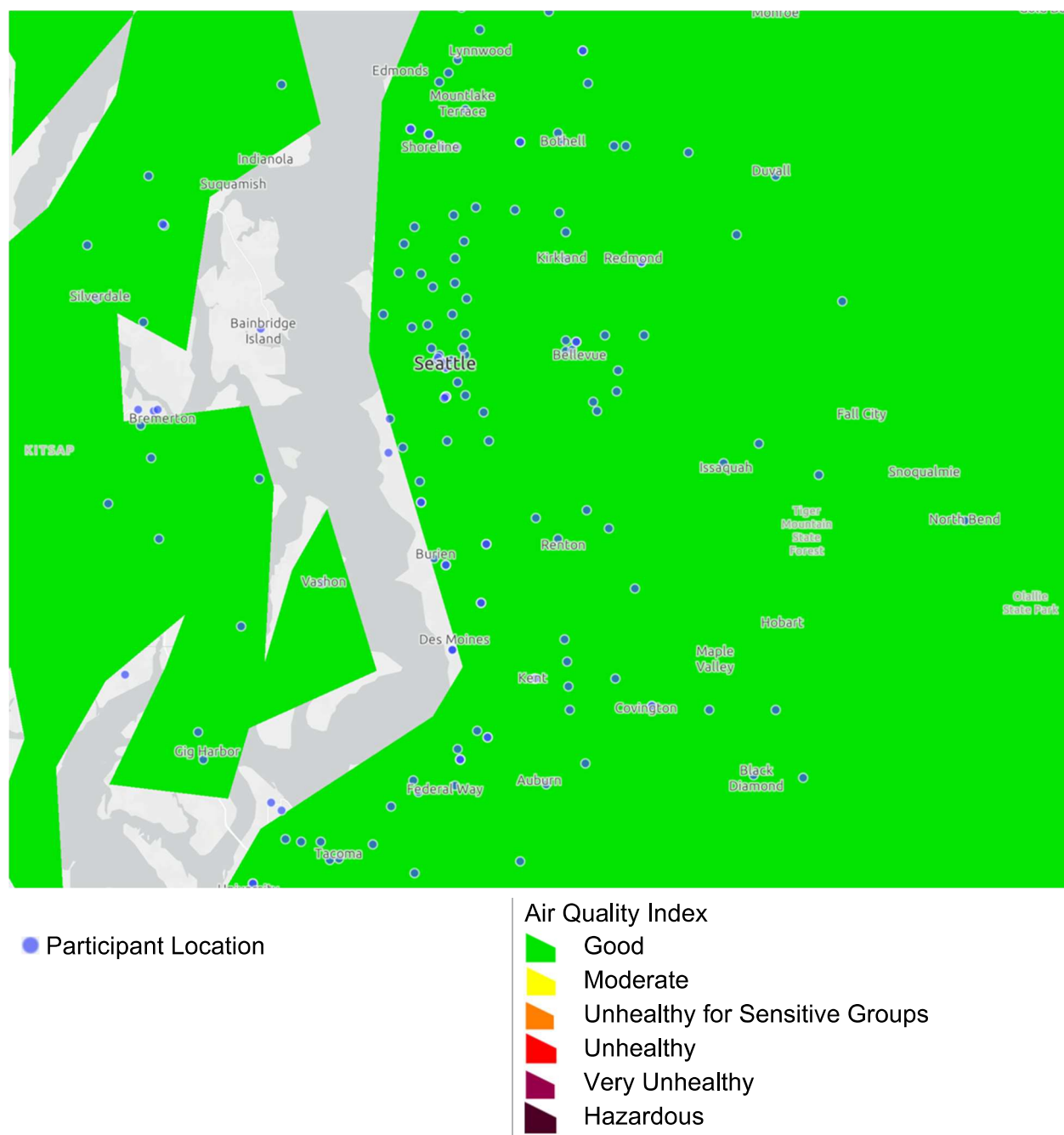


Figure 11. 2012–2020 Participant Location by Air Quality (Ozone), Seattle Area Detail. Air Quality (ozone) Reference Layer: AirNow – AQI Contours (Last 24 hours, Ozone only) (U.S. EPA 2020).

3.6 Social Vulnerability Index

A combined social vulnerability index compiles many of the indicators previously discussed in one relative percentile ranking reference layer. This allows for a more complete view of county and census tract vulnerability when time and/or space constraints require a concise analysis.

A county-level view of Washington State (see Figure 12) shows that the majority of participants are located in counties that fall in the low-moderately vulnerable range. Smaller clusters of participants are in the central and southern parts of the state near mid-sized cities (Wenatchee, Yakima, and Tri-Cities) in counties that are considered moderately to highly vulnerable. This confirms trends that were revealed with other selected indicators (see race and ethnicity, linguistic isolation, and education).

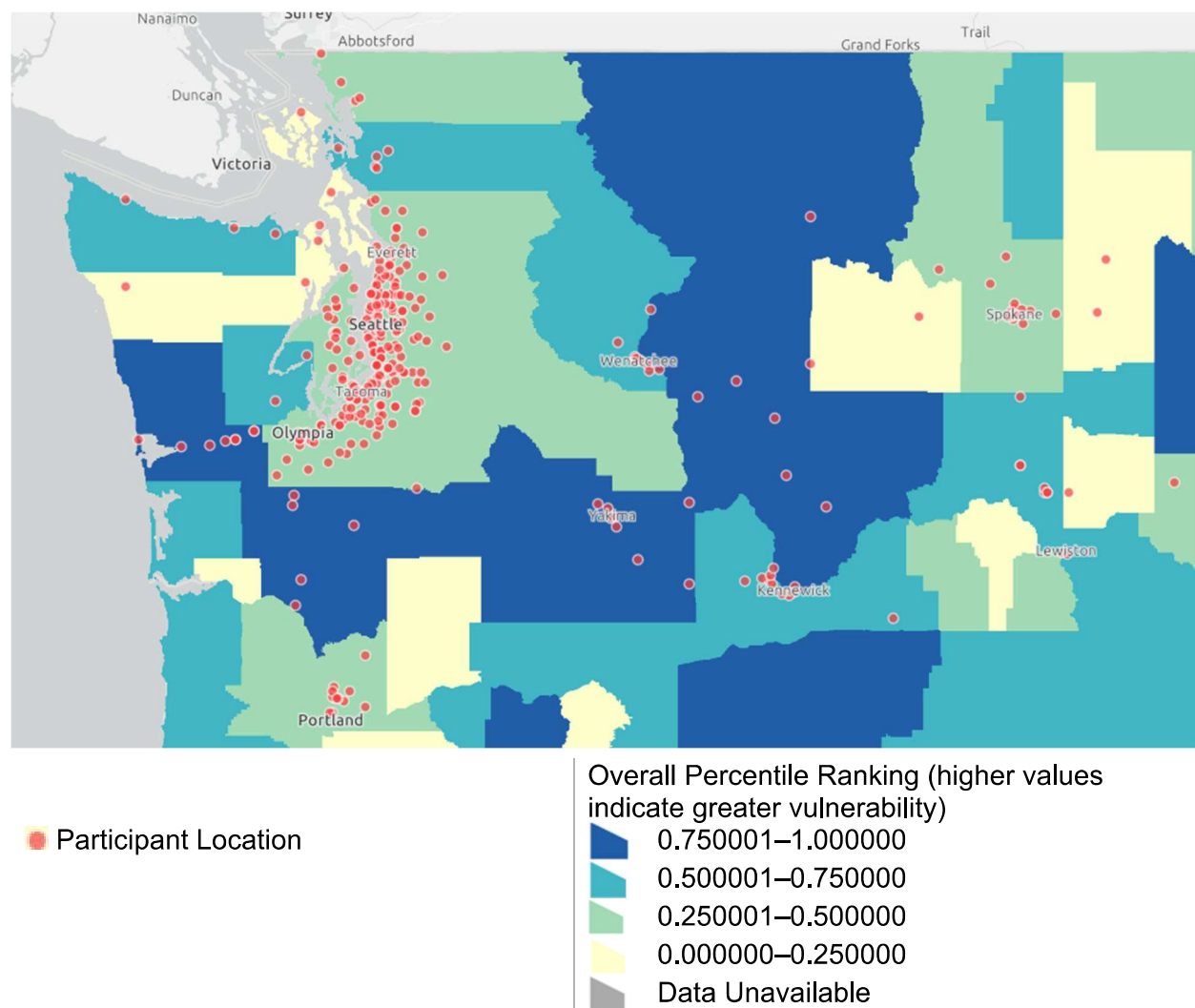


Figure 12. 2012–2020 Participant Location by Social Vulnerability Index, Washington State Detail. Social Vulnerability Reference Layer: 2018 Social Vulnerability Index (CDC 2021).

The Seattle area census tract-level view shown in Figure 13 highlights areas south of Seattle as more vulnerable. Participants appear to be distributed between more resilient census tracts and those that rank higher on the social vulnerability index.

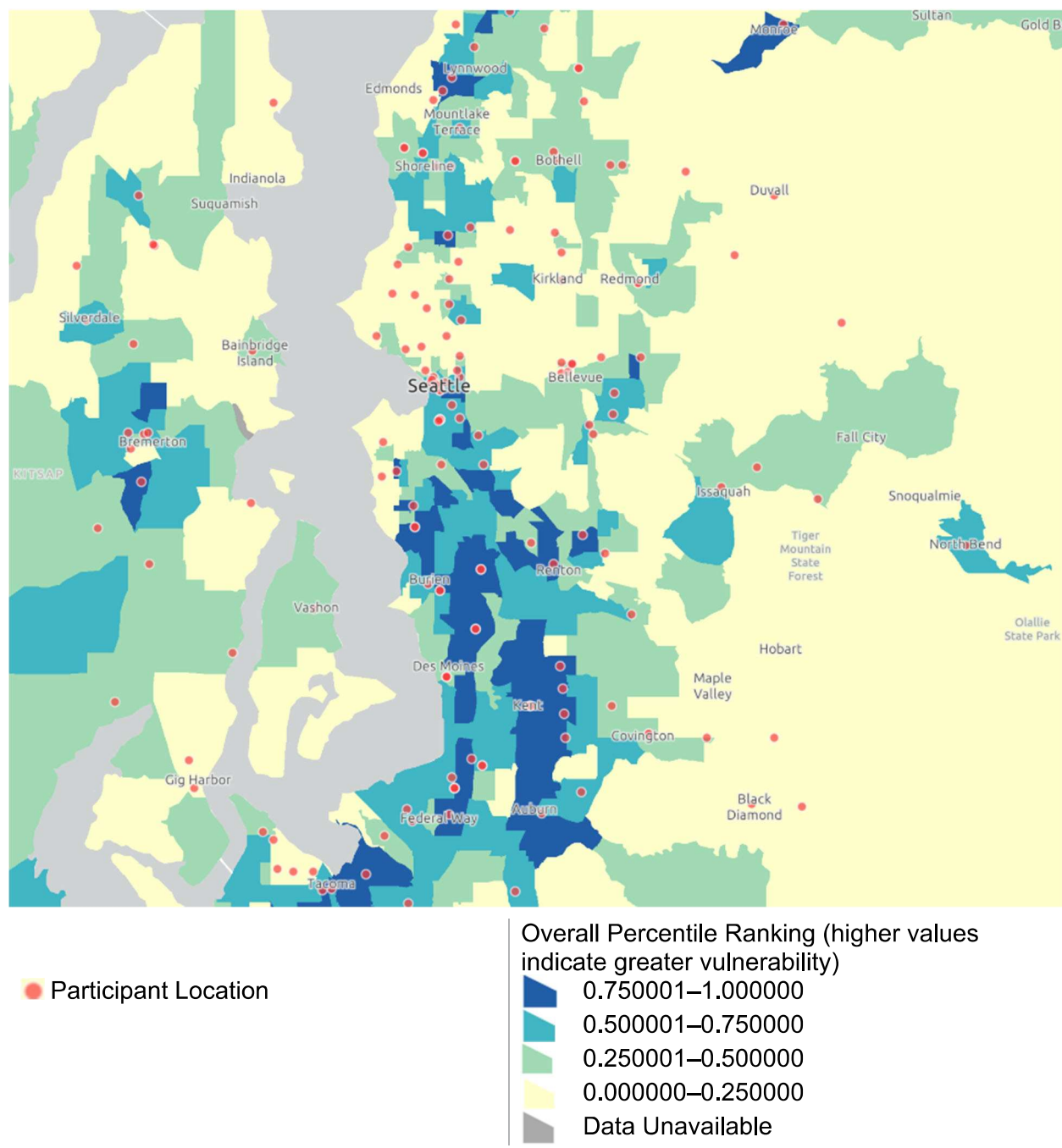


Figure 13. 2012–2020 Participant Location by Social Vulnerability Index, Seattle Area Detail. Social Vulnerability Reference Layer: 2018 Social Vulnerability Index (CDC 2021).

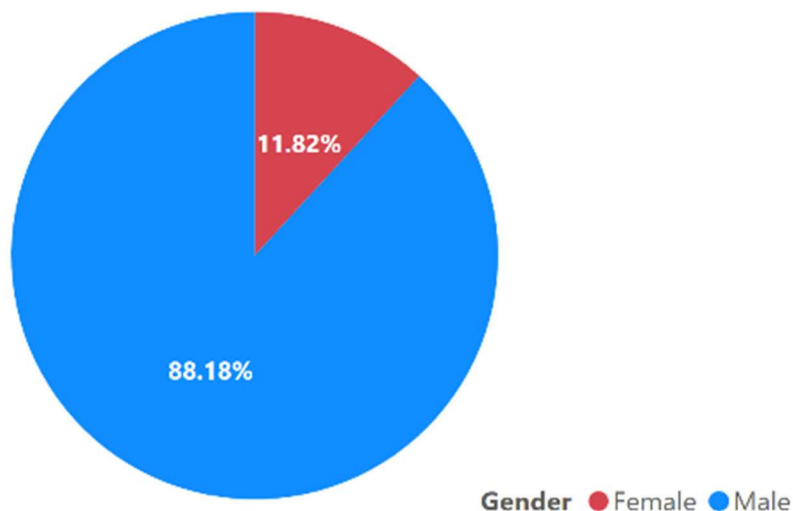
3.7 Gender

Participant gender is collected on the training and certification registration forms as an optional field. Records reflect selections of Male, Female, and blank. The data was further divided into participants who had completed Level 1 TCOC and Level 2 TCOC, to uncover patterns in certification levels between genders. A total of 1,092 participants completed the Level 1 TCOC

and 386 had completed the Level 2 TCOC. Most participants filled out the gender field, with 77.47 percent of Level 1 records and 83.94 percent of Level 2 TCOC records complete.

Over the 10-year period from 2001 to 2021, 11.82 percent of all Level 1 participants reporting gender were female. The percentage of female participants drops to 8.95 percent at Level 2, as shown in Figure 14.

Level I TCOC Year by Gender



Level II TCOC Year by Gender

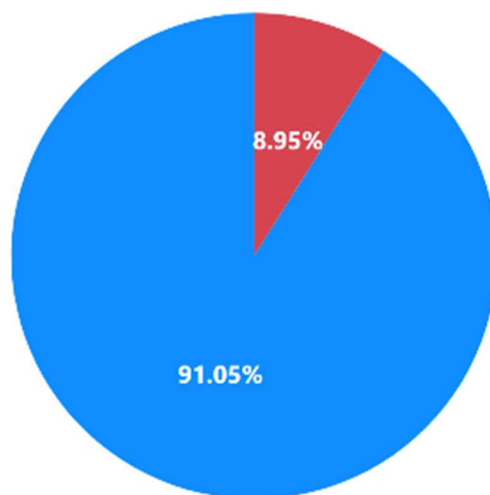


Figure 14. 2012–2020 Percentage of Participant Gender by TCOC Level (participants with missing data in the gender field are not shown).

Enrollment in Level 1 TCOC jumped sharply in 2010, followed by a jump in Level 2 enrollment in 2011. After steadily increasing to a high of 20 individuals in 2012, female participation declined to 4 in 2020. Male participation in Level 1 has been declining from a high of 103 in 2015. Female participation in Level 2 is sporadic. Participation was highest in 2015 and 2018 with 8 women. See Figure 15 below for more detail.

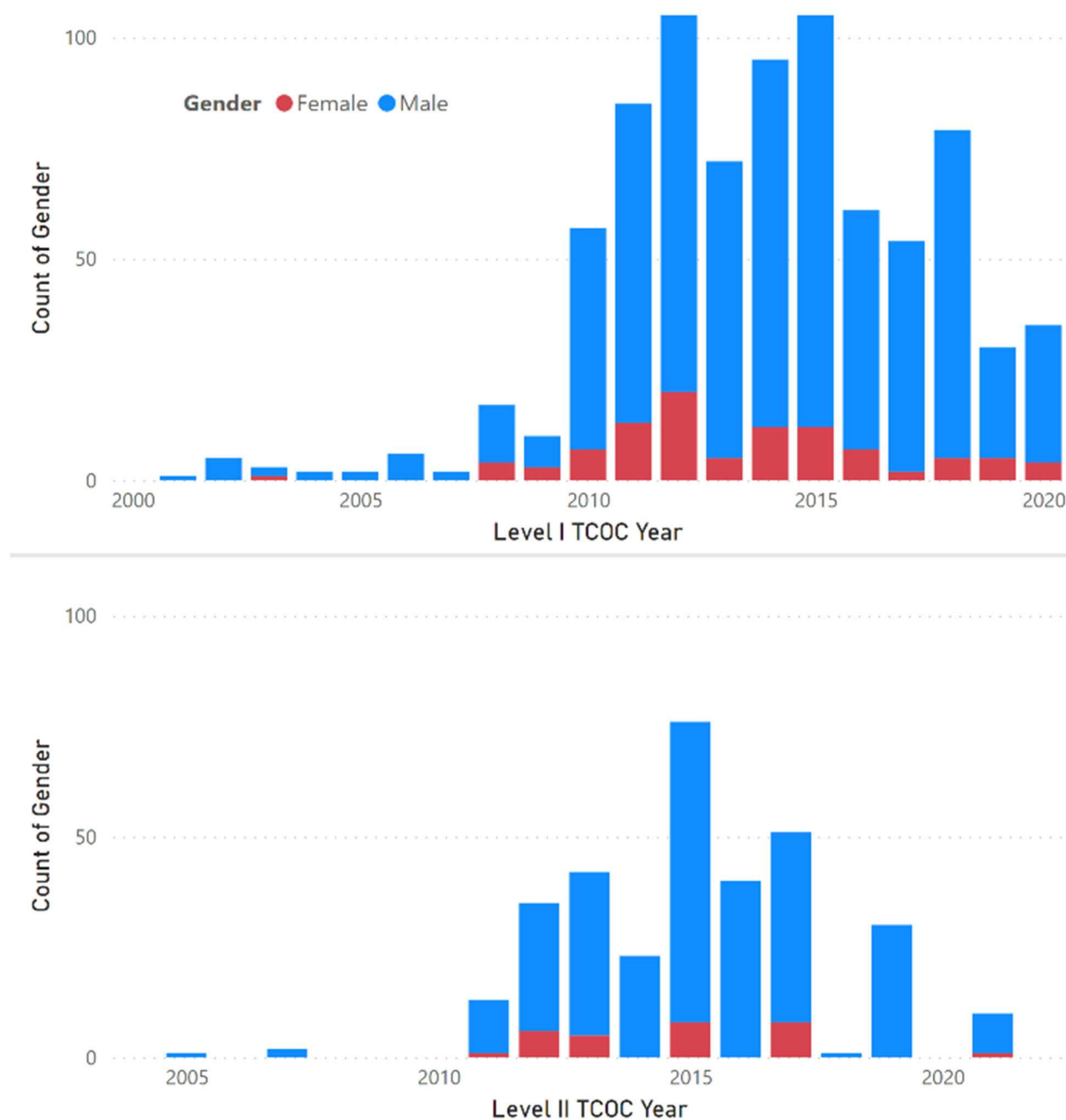


Figure 15. 2012–2020 Count of Participant Gender by TCOC Year (participants with missing data in the gender field are not shown).

4.0 Replicating Analysis

4.1 Data Preparation and Cleaning in Excel

Registration form fields:

- Primary City
- Primary State
- Primary ZIP/Postal Code
- Gender
- Level 1 TCOC Year
- New Cert Year
- Level 2 TCOC Year.

Clean data in Excel prior to importing to Power BI. Primary ZIP/Postal Code is used as the location point, so ensuring the field contains either 5- or 10-character (9-digit) ZIP codes is important for accurate point placement. Use the formula =LEN() to identify records that are not 5 or 10 characters in length and correct the records. ArcGIS Maps for Power BI can handle ZIP codes missing leading zeroes, because the country is defined. Other mapping tools may require the leading zero to accurately place the point. To preserve leading zeroes in Excel, format the field as text. If the mapping tool is unable to identify the ZIP code, the point will be unmapped.

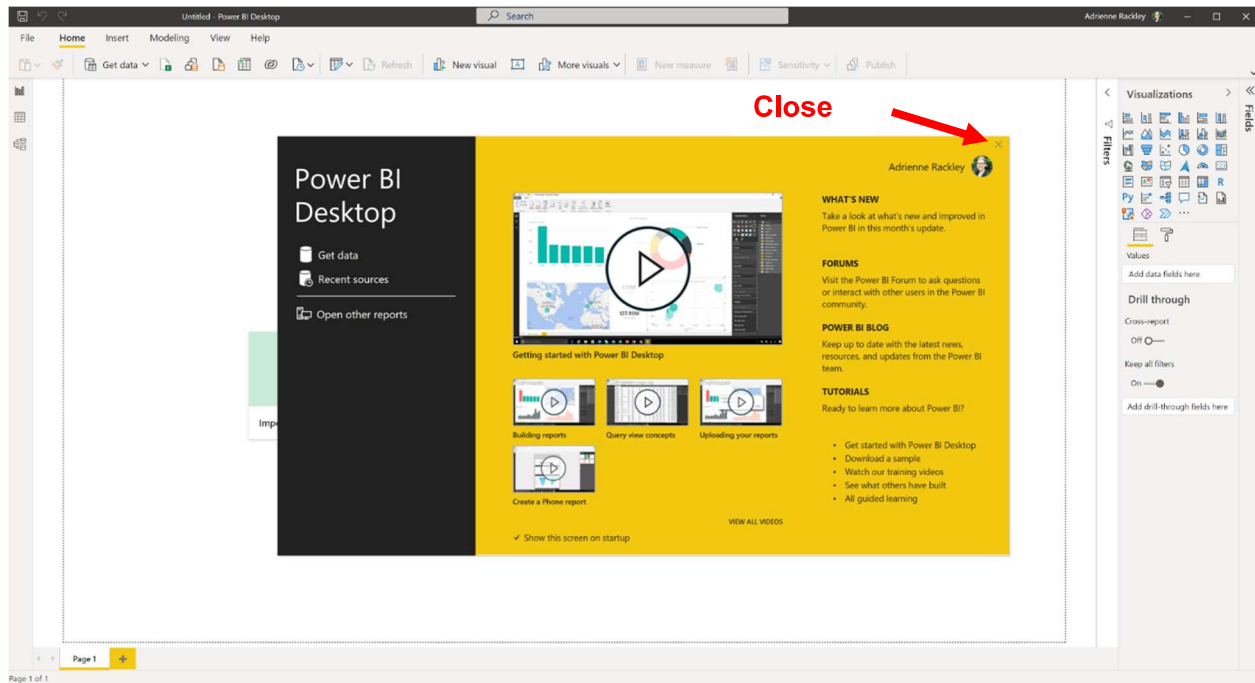
Alternatively, the Primary City field may be used as the location point. If this is desired, it is necessary to review the records for spelling errors and inconsistencies. If the ZIP code is used to locate the point, the city may still be used in the Tooltip field and would still benefit from a careful review to ensure correct spelling, punctuation, and consistent capitalization.

It is a good idea to check all other fields (city, state) for typos and spelling errors.

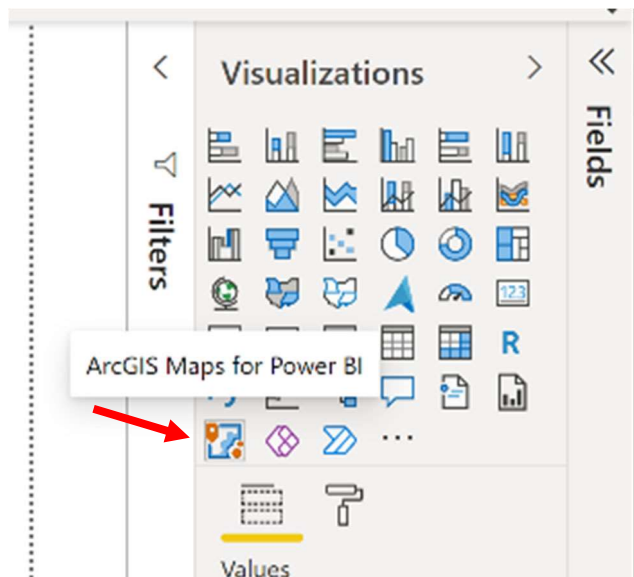
Once the data is cleaned in Excel, it can be imported into Power BI for analysis.

4.2 Importing Data into Power BI

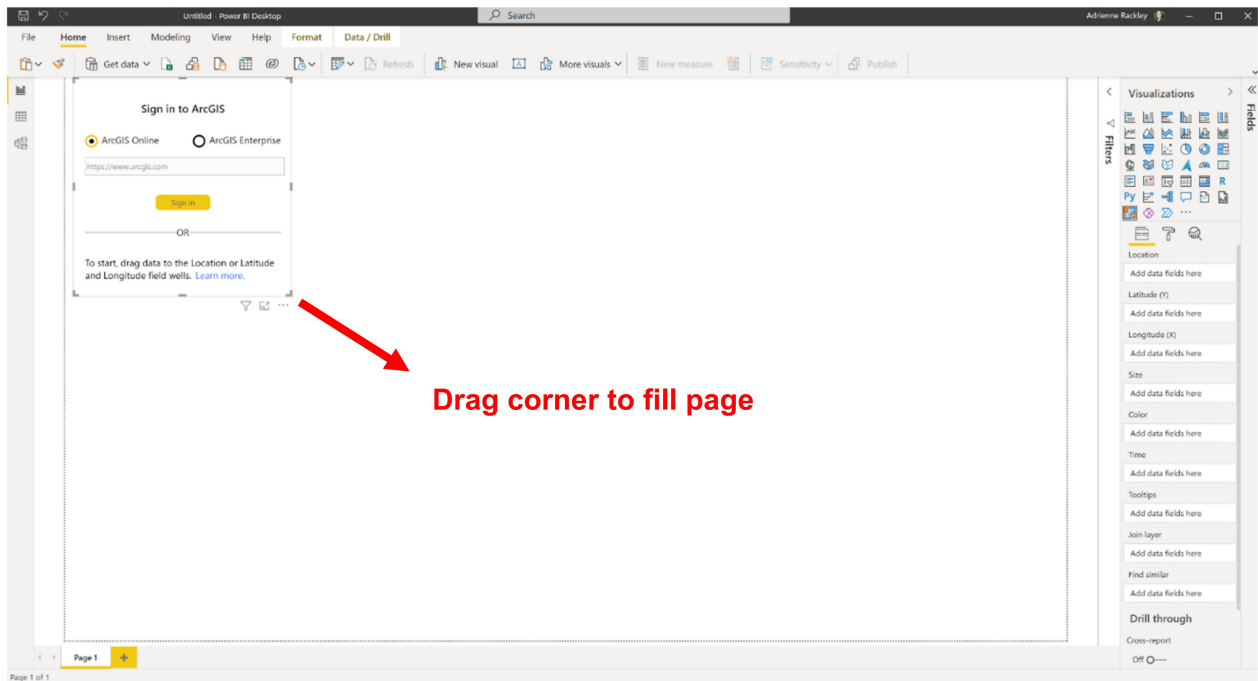
Open a new instance of Power BI Desktop. Close the start-up window. I find that ArcGIS Maps for Power BI can be temperamental if the data source is loaded first. It is less likely to throw errors if the visualization is selected first and data imported second.



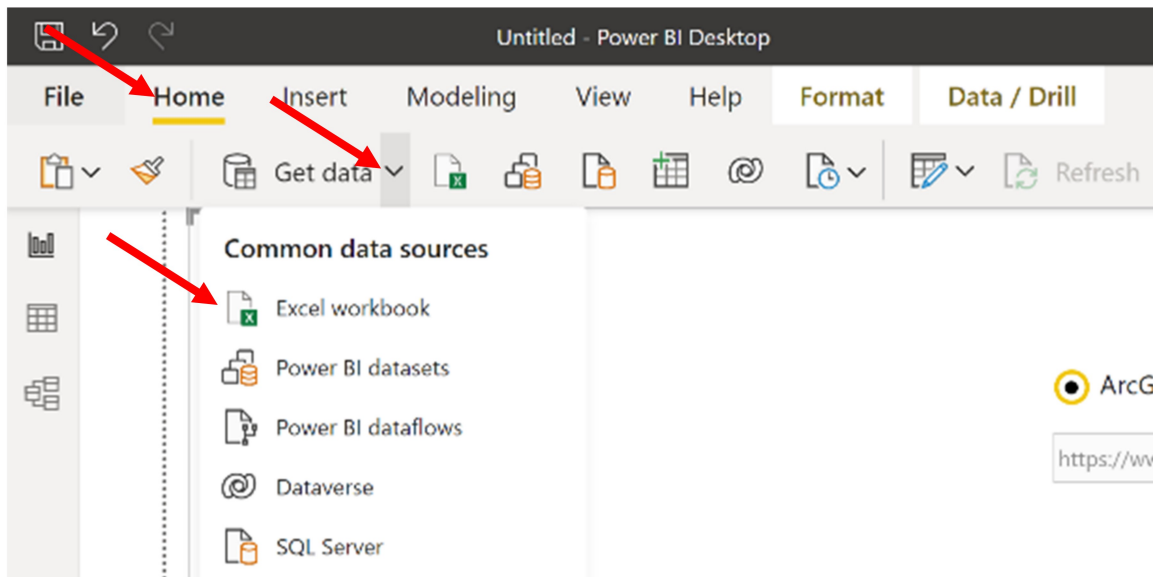
In the Visualizations Pane select ArcGIS Maps for Power BI.



Resize the visual to fill the page. Unless you have an ArcGIS Online or ArcGIS Enterprise account, do not follow prompts to sign in.



To import registration data, from the Home tab select Get data, Excel workbook.



Navigate to the Excel cleaned data file and click Open. Select the worksheet containing the data you want to use and click Transform Data.

Navigator

Display Options ▾

WA BOC Training Participants 2010-2020_v4 -...

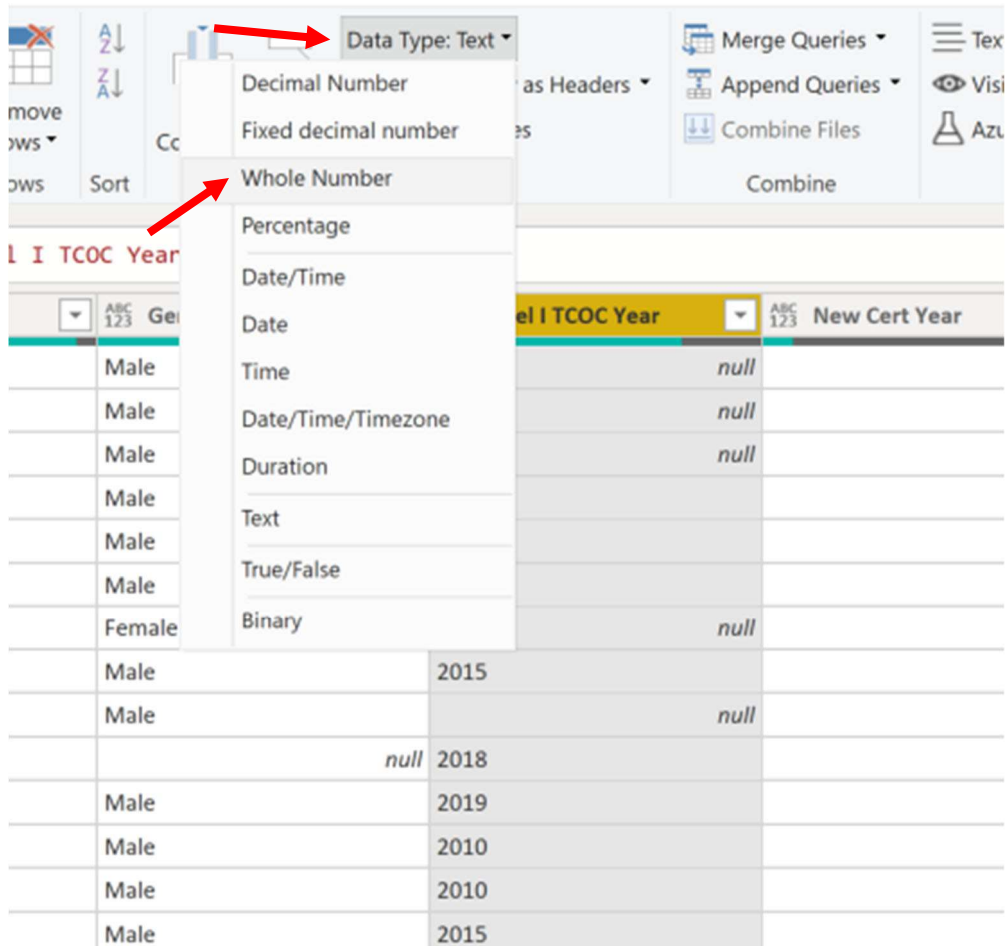
☒ Sheet1

Sheet1

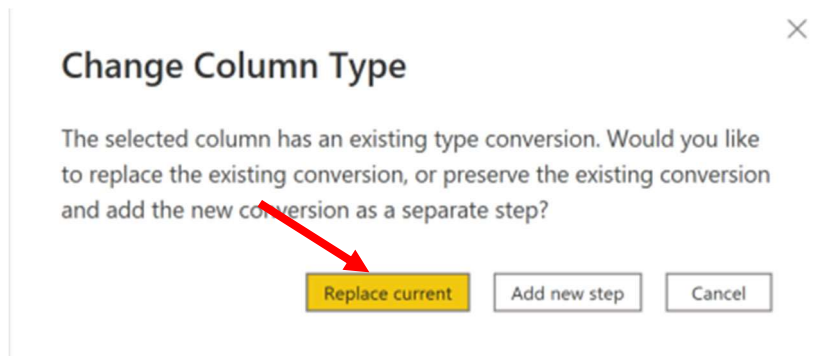
Primary City	Primary State	Primary ZIP/Postal Code	Gender	Level I
Aberdeen	Washington	98520	Male	
Aberdeen	Washington	98520	Male	
Aberdeen	Washington	98520	Male	
Alameda	California	94501	Male	
Anacortes	Washington	98221	Male	
Arlington	Washington	98223	Male	
Arlington	Washington	98223	Female	
Arlington	Washington	98223	Male	
Arlington	Washington	98223	Male	
Arlington	Washington	98223	null	
Arlington	Washington	98223	Male	
Ashford	Washington	98304	Male	
Ashford	Washington	98304	Male	
Athol	Idaho	83801	Male	
Auburn	Washington	98002	Male	
Auburn	Washington	98002	Male	
Auburn	Washington	98092	null	
Auburn	Washington	98092	null	
Auburn	Washington	98032	Male	
Auburn	Washington	98092	Male	
Auburn	Washington	98032	Male	
Auburn	Washington	98092	Male	
Auburn	Washington	98092	Male	

Load Transform Data Cancel

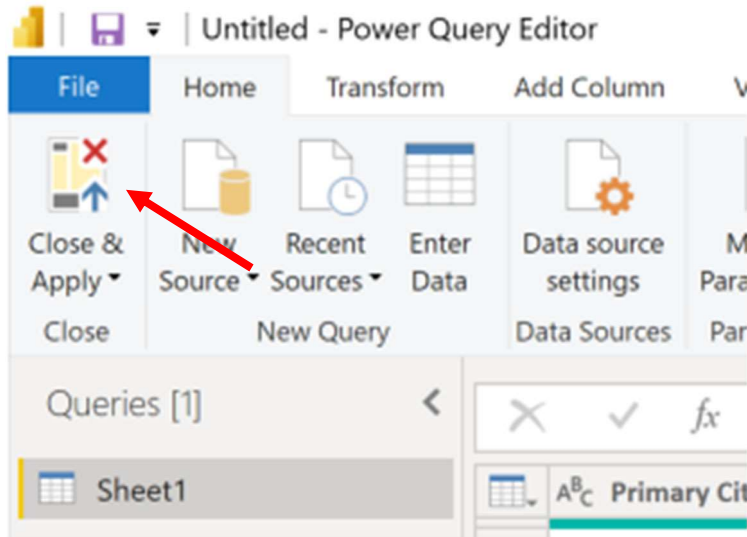
Review the data type that was automatically assigned to fields and modify where necessary. City, State, ZIP, and Gender should correctly come over as text. Year fields may come over as a mix of number and text. Year fields must be formatted as a number if they are used as a chronological axis on a chart.



When prompted with how to handle the existing type conversion, select Replace current.



You may choose to exclude unnecessary columns by clicking on the column header and then Remove Columns. Click Close & Apply.

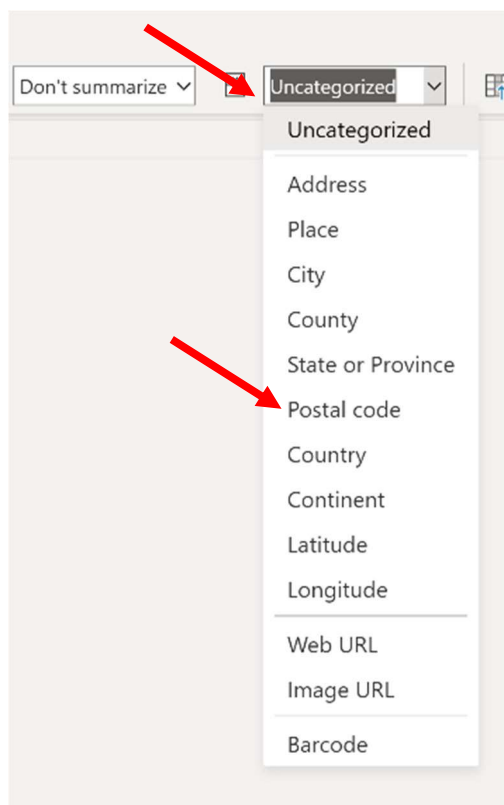


This is a good time to begin saving your work at regular intervals. Power BI does not autosave.

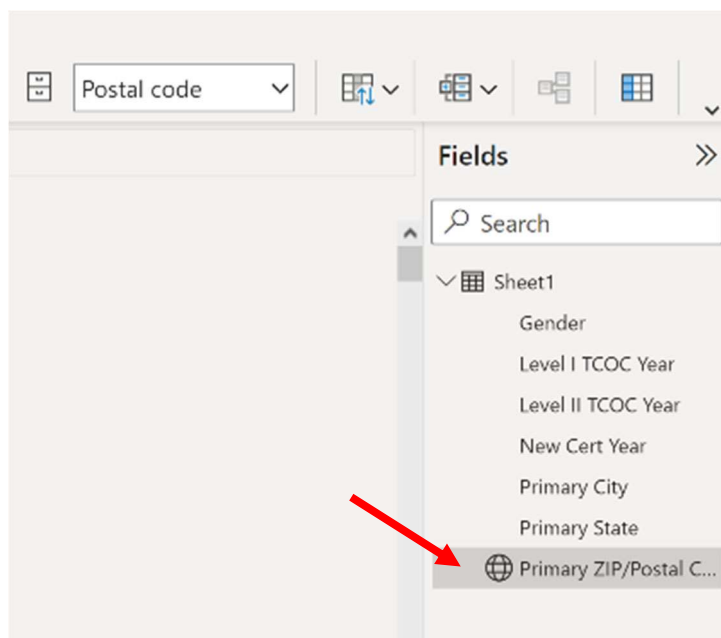
Select the Data table view on the left side to assign data categories for the geospatial fields. Click on the column header and assign the appropriate data category from the drop-down menu.

 This screenshot shows the 'Table tools' and 'Column tools' tabs in the Power Query Editor. The 'Primary ZIP/Postal Code' column header is highlighted, and a red arrow points to it. The table below contains data for Seattle, Washington, with various ZIP codes and gender information.

Primary City	Primary State	Primary ZIP/Postal Code	Gender
Seattle	Washington	98104	Female
Seattle	Washington	98103	Male
Seattle	Washington	98178	
Seattle	Washington	98121	Female
Seattle	Washington	98125	
Seattle	Washington	98104	Male
Seattle	Washington	98321	Male
Seattle	Washington	98104	Female
Seattle	Washington	98109	Male

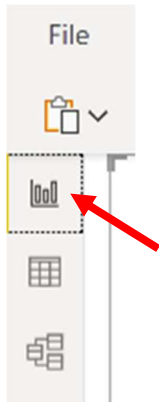


The field should now appear with a globe icon in the Fields pane to the right.



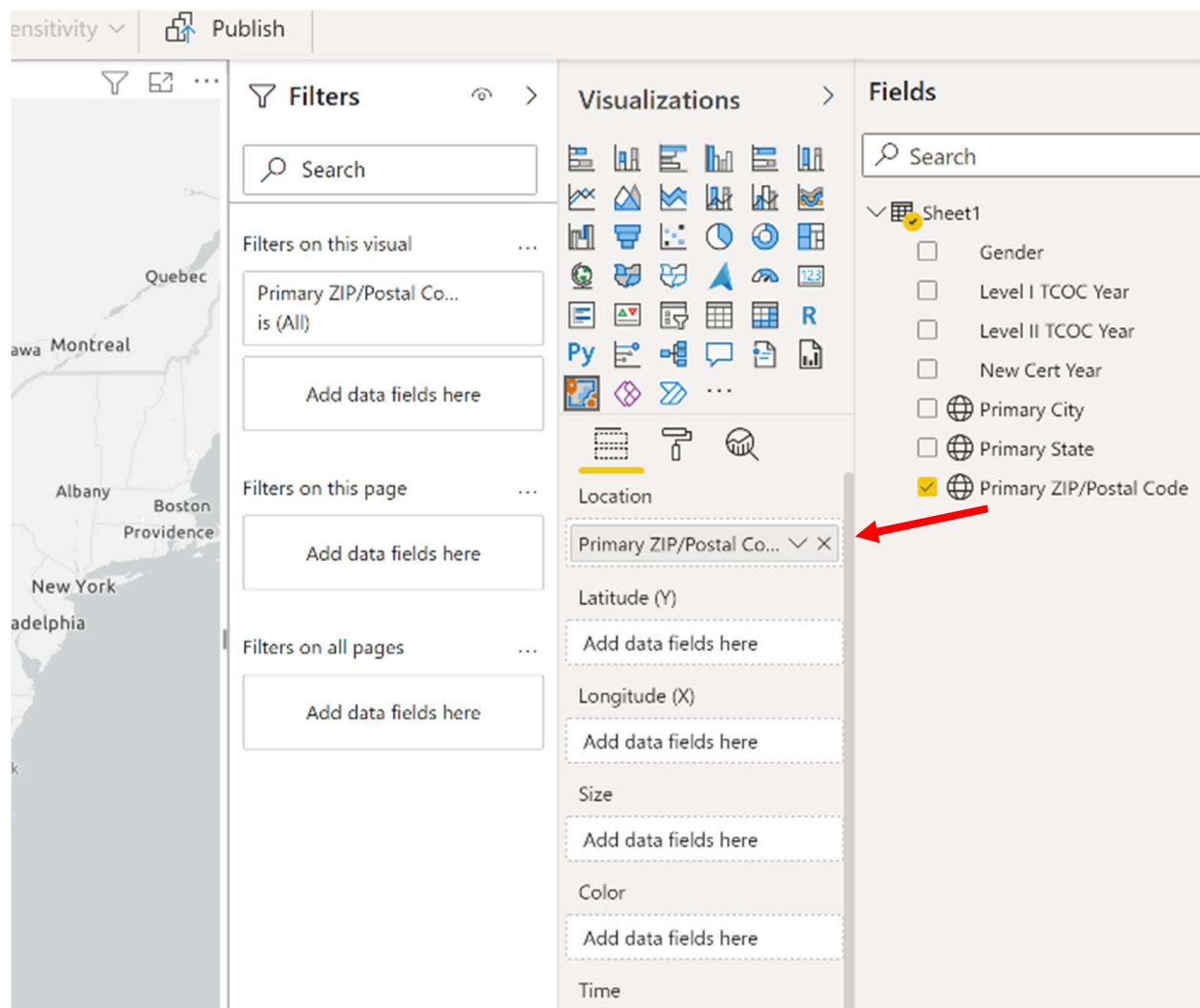
Repeat the process for all geospatial fields.

Return to the report view by clicking the bar chart icon in the upper left.

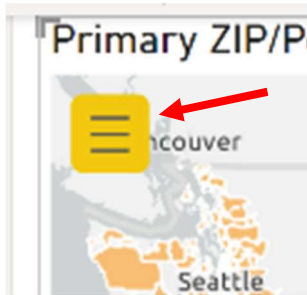


4.3 Building the Map in Power BI

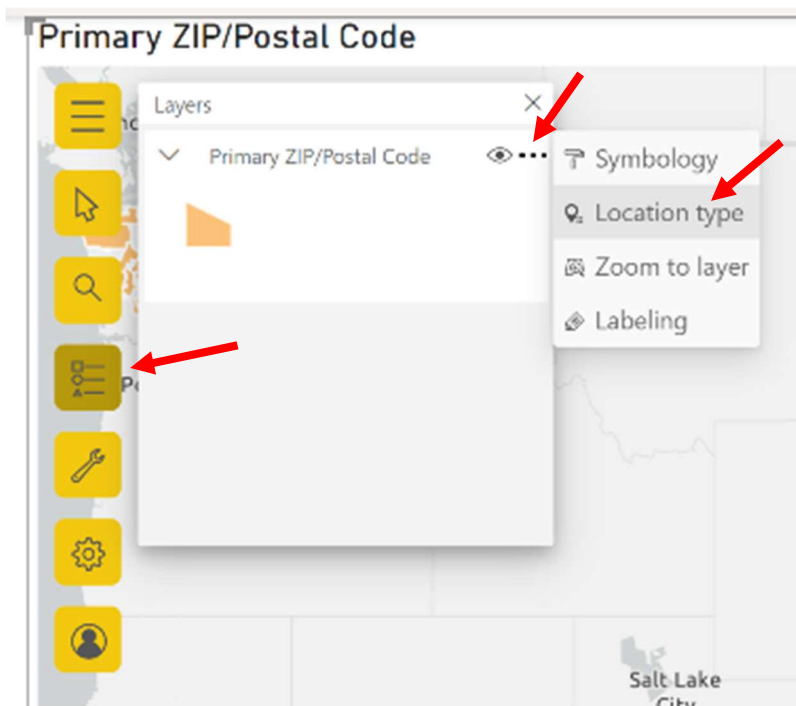
Click anywhere in the ArcGIS Maps visual to select it. Click and drag the Primary ZIP/Postal Code field from the Fields pane to the Location data field in the Visualizations Pane.



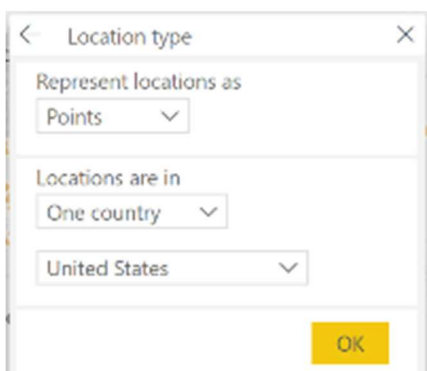
ZIP codes are automatically displayed as boundaries. To display as points, expand map tools by clicking this icon in the upper left corner:



Click to view a list of layers in the map. Click the three dots to the right of the Primary ZIP/Postal Code layer, then select Location type.



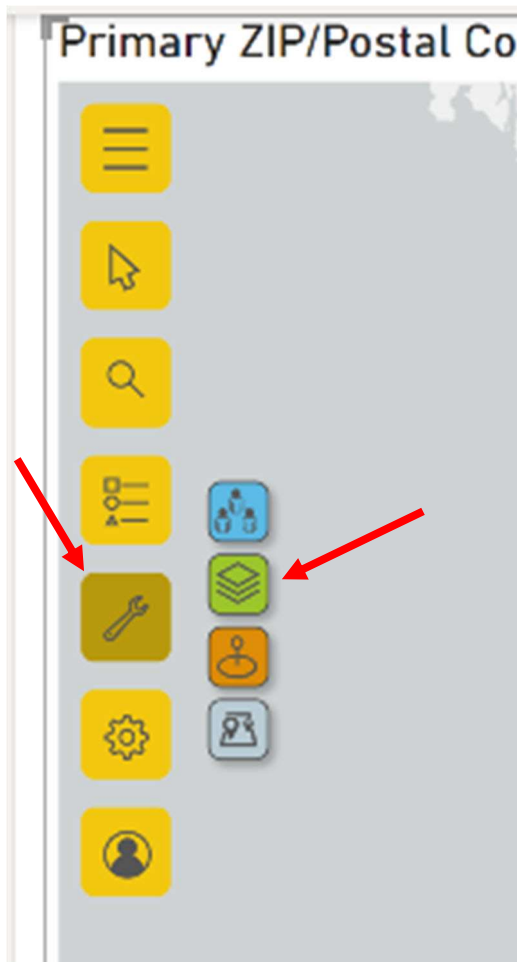
Make the following selections from the drop-down list: 1) Represent locations as Points, 2) Locations are in One country, 3) United States.



Additional formatting may be applied as desired. For example, it may be useful to display the city and state when you hover over a point on the map. To achieve this, drag the Primary City and Primary State fields into Tooltips.

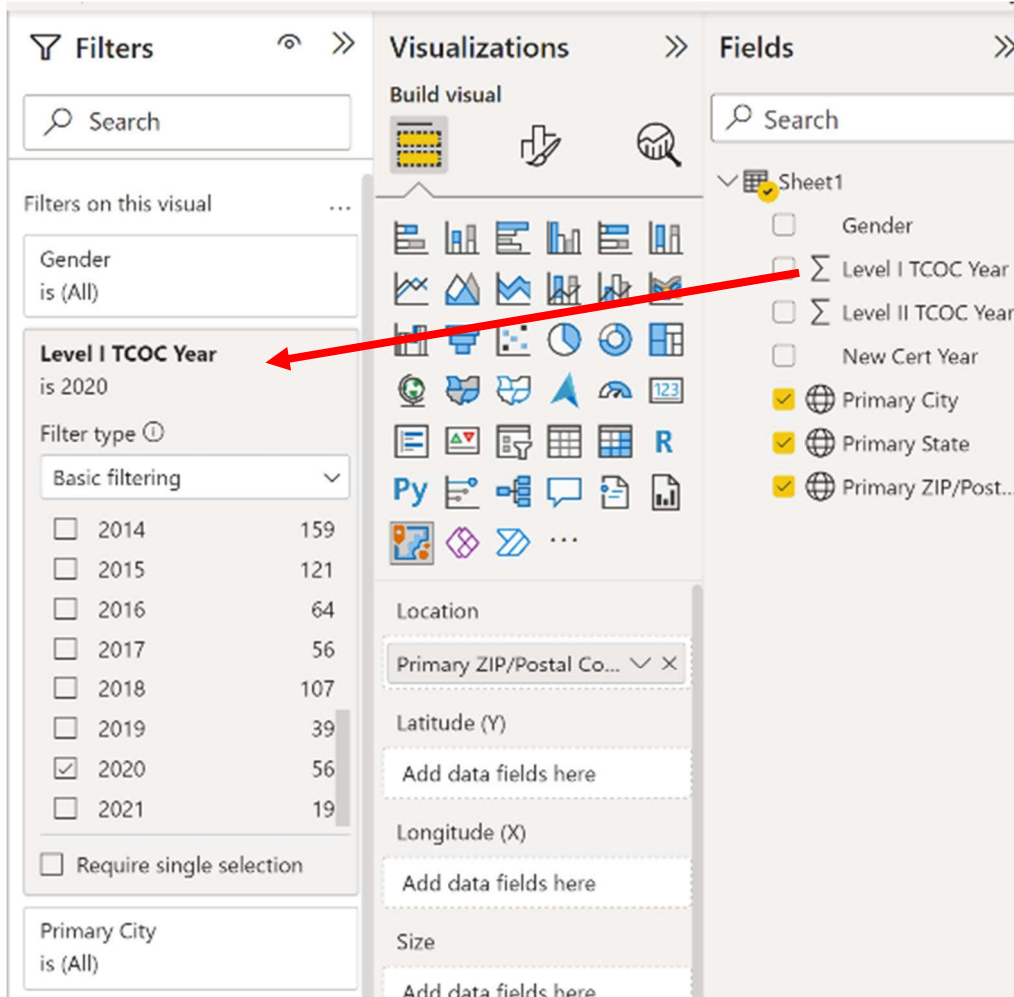
4.3.1 Adding Reference Layers

Reference layers add context to the map and begin to illustrate the socioeconomic and demographic characteristics of participant communities.



4.3.2 Adding Filters

Filters may be added to a visual, a single page, or all pages. To add a filter, click and drag a field from the Fields pane to the Filters pane.



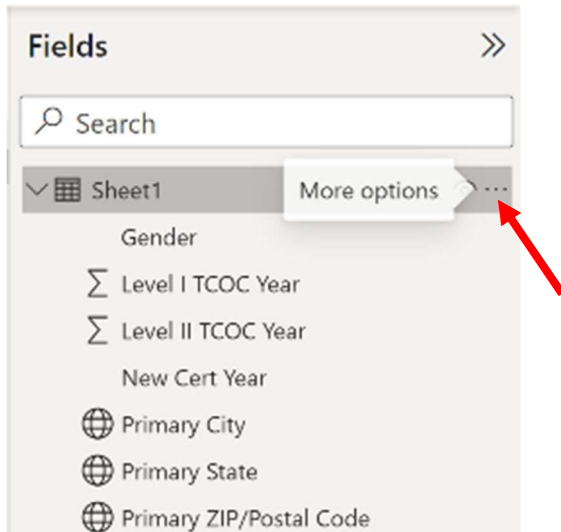
4.3.3 Appending New Data from Subsequent Years

New data may be incorporated into the visual without the need to rebuild the visualizations. Depending on how the raw data is presented, all records may be combined in a single Excel file (or database) or multiple files (or database locations). These instructions explain how to append two Excel files in Power BI to incorporate new data.

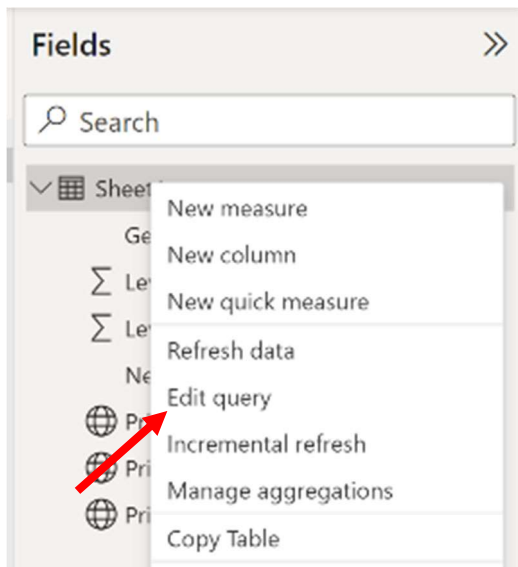
The new Excel data should be cleaned and formatted, as explained in the section “Preparation and Cleaning in Excel.” Ensure that both the new and old Excel files have the same field headers and data formats.

Create a new query in the existing Power BI file to bring in the new data. Follow the steps as outlined in 4.2 Importing Data into Power BI, omitting the first steps that deal with setting up a new file and visualization. There should now be two queries in the file.

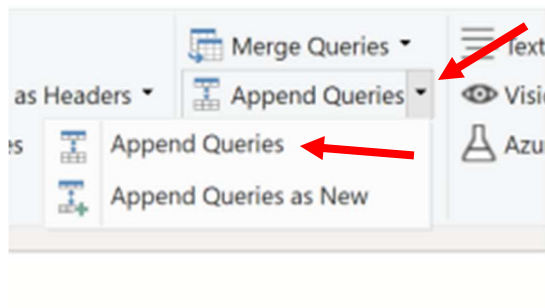
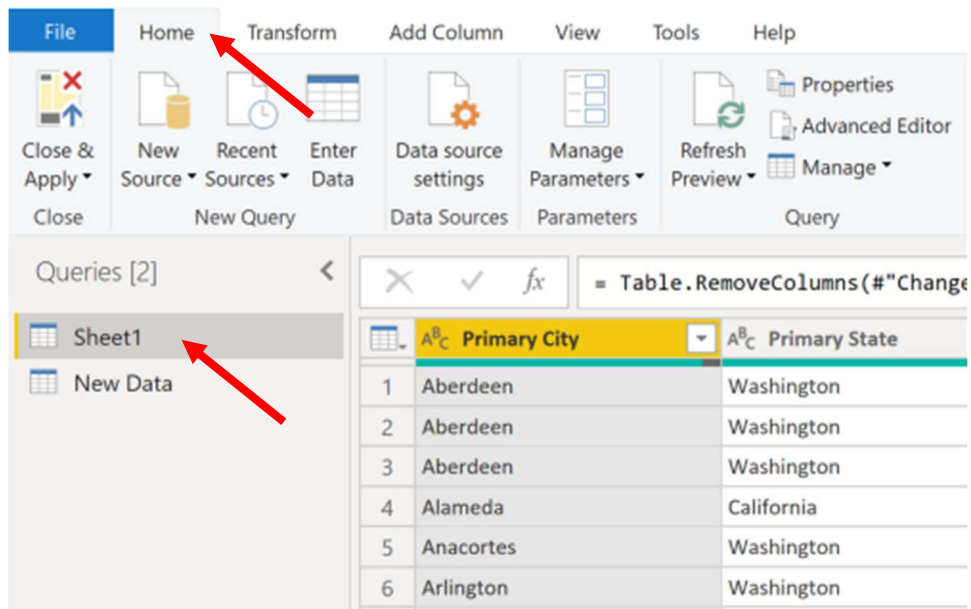
To append the new data set to the existing one in Power BI, click on the ellipsis to the right of the table name to open more options.



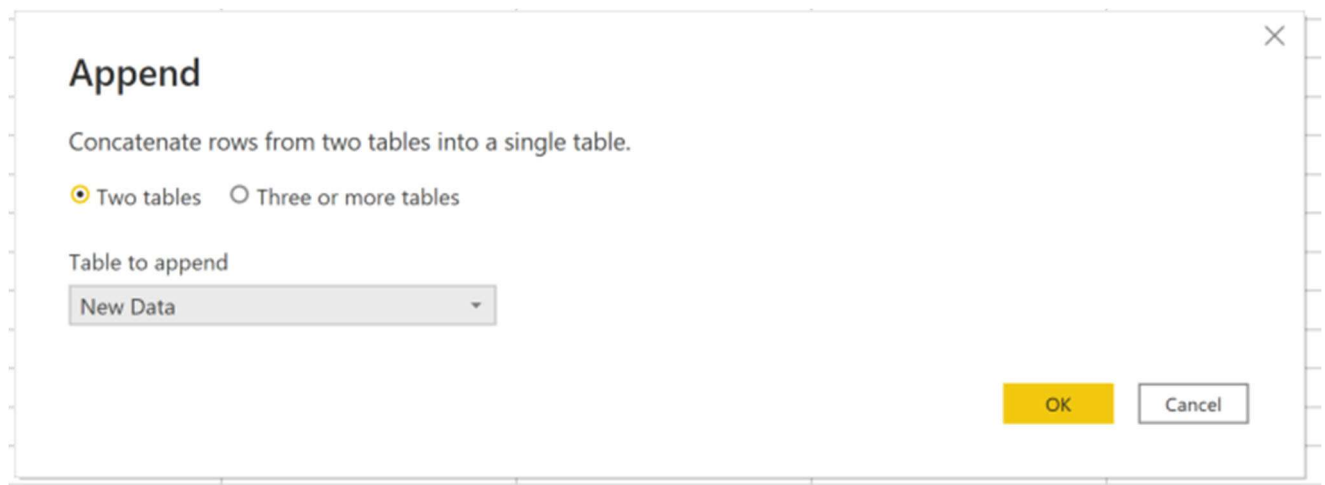
Select Edit query.



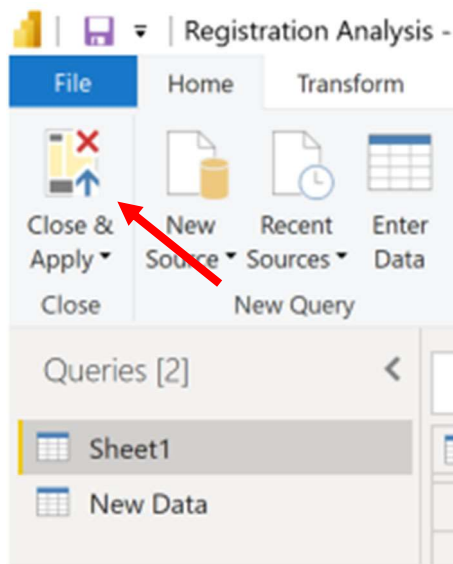
The original query (in this instance Sheet1 is the original query and New Data is the new query) should be selected in the left-hand pane. Under the home ribbon, select Append Queries.



Check two tables, select the query with your new data, and click OK.



Click Close & Apply



The new data is now added to your visuals. You may need to review your filter selections to ensure the new data is displayed.

5.0 Recommendations

No public reference layer was located for environmental health risk. This indicator may be broken out into component parts. According to the Centers for Disease Control and Prevention (CDC), the most common environmental health hazards are air and water pollution (CDC 2018). Air quality (ozone) was mapped as a reference layer. Consider including Particulate Matter 2.5 (PM2.5) with ozone using reference layer AirNow – AQI Contours (Last 24 hours, Ozone, and PM2.5 combined) U.S. EPA, OAR, OAQPS (U.S. EPA 2020). No map was located for potable water quality or risk for lead in water. Consider as an alternative proximity to polluted sites, health insurance coverage, heat islands, or disease prevalence.

A more robust statistical analysis could be conducted by modifying the registration form to collect more data about participants. This could be done with the objective of relating participant location to census data directly (collect specific location information) or gathering data about participant demographic and socioeconomic characteristics.

Suggestions for new (optional) registration form fields include fields that are commonly requested, such as

- Highest level of education completed
- Race and ethnicity
- Disability status
- County (relates to census field).

Consider how to get information that participants may be less willing to share, such as

- Income
- Home address.

To help target outreach and understand motivations and barriers to enrollment, gather more data on why people enroll by asking

- Is this credential required by your employer? If not, what do you hope to gain from this program?
 - Improve employment opportunities
 - Increase income
 - Want to learn more
 - Other (explain).

There are opportunities for further research into expanding opportunities for women and non-binary people in facilities maintenance, operation, and management. The BOC program's gender statistics could be compared with census and industry benchmarks for more context. Additionally, the gender field on the registration form could be modified for inclusivity and better data collection. Selections should include a minimum of male, female, and non-binary.

One data field, New Cert Year, includes failed attempts in text, as well as the new certification year, if an attempt was successful. Record failed attempts in a separate column.

6.0 References

Centers for Disease Control and Prevention/Agency for Toxic Substances and Disease Registry/ Geospatial Research, Analysis, and Services Program. 2021. “2018 Social Vulnerability Index.” Esri ArcGIS. Last modified October 6, 2021.

<https://www.arcgis.com/home/item.html?id=cbd68d9887574a10bc89ea4efe2b8087>

Centers for Disease Control and Prevention. 2018. “Introduction to Environmental Public Health Tracking.” Last modified January 3, 2019. <https://www.cdc.gov/nceh/tracking/tracking-intro.html>

Council on Environmental Quality. n.d. “Climate and Economic Justice Screening Tool.” Accessed March 1, 2022. <https://screeningtool.geoplatform.gov/en/#3/33.47/-97.5>

Esri. n.d. “ArcGIS Living Atlas of the World United States Edition.” Accessed March 1, 2022. <https://livingatlas.arcgis.com/en/browse/#d=2&rgnCode=US>

Esri. 2022. “ACS Race and Hispanic Origin Variables – Boundaries.” Esri ArcGIS. Credit: U.S. Census Bureau's American Community Survey (ACS) 2015-2019 5-year estimates, Table(s) B03002. Last modified March 17, 2022.

<https://www.arcgis.com/home/item.html?id=23ab8028f1784de4b0810104cd5d1c8f>

Esri. 2022. “ACS English Ability and Linguistic Isolation Variables – Boundaries.” Esri ArcGIS. Credit: U.S. Census Bureau's American Community Survey (ACS) 2016-2020 5-year estimates, Table(s) B16003, B16004. Last modified March 18, 2022.

<https://www.arcgis.com/home/item.html?id=0c4d1027de6b4d6eb896d95f1240e1aa>

Esri. 2022. “ACS Poverty Status Variables – Boundaries.” Esri ArcGIS. Credit: U.S. Census Bureau's American Community Survey (ACS) 2016-2020 5-year estimates, Table(s) B17020. Last modified March 17, 2022.

<https://www.arcgis.com/home/item.html?id=0e468b75bca545ee8dc4b039cbb5aff6>

Esri. 2022. “ACS Educational Attainment Variables – Boundaries.” Esri ArcGIS. Credit: U.S. Census Bureau's American Community Survey (ACS) 2016-2020 5-year estimates, Table(s) B15002. Last modified March 17, 2022.

<https://www.arcgis.com/home/item.html?id=84e3022a376e41feb4dd8addf25835a3>

Tidwell, A. 2021. *Better Buildings Workforce Accelerator – Technical Assistance Project: Northwest Energy Efficiency Council – Diversity, Equity, and Inclusion in Building Operator Certification Training*. Pacific Northwest National Laboratory for U.S. Department of Energy.

U.S. Census Bureau. 2022. “B03002: Hispanic or Latino Origin by Race.” 2020 American Community Survey 5-Year Data. Last modified March 17, 2022.

<https://data.census.gov/cedsci/table?q=b03002&tid=ACSDT5Y2020.B03002>

U.S. Census Bureau. 2022. “B15002: Sex by Educational Attainment for the Population 25 Years and Over.” 2020 American Community Survey 5-Year Data. Last modified March 17, 2022. <https://data.census.gov/cedsci/table?q=B15002&tid=ACSDT5Y2020.B15002>

U.S. Census Bureau. 2022. “B16003: Age by Language Spoken at Home for the Population 5 Years and Over in Limited English Speaking Households.” 2020 American Community Survey

5-Year Data. Last modified March 17, 2022.

<https://data.census.gov/cedsci/table?q=B16003&tid=ACSDT5Y2020.B16003>

U.S. Census Bureau. 2022. "B16004: Age by Language Spoken at Home by Ability to Speak English for the Population 5 Years and Over." 2020 American Community Survey 5-Year Data. Last modified March 17, 2022.

<https://data.census.gov/cedsci/table?q=B16004&tid=ACSDT5Y2020.B16004>

U.S. Census Bureau. 2022. "B17020: Poverty Status in the Past 12 Months by Age." 2020 American Community Survey 5-Year Data. Last modified March 17, 2022.

<https://data.census.gov/cedsci/table?q=B17020&tid=ACSDT5Y2020.B17020>

U.S. EPA, Office of Air and Radiation, Office of Air Quality Planning and Standards (OAQPS). 2020. "AirNow – AQI Contours (Last 24 hours, Ozone and PM2.5 combined), US EPA, OAR, OAQPS." Esri ArcGIS. Last modified November 24, 2020.

<https://www.arcgis.com/home/item.html?id=ae6bb649e4dd4b18bac4e5b0c3ba2e8a>

U.S. EPA, Office of Air and Radiation, Office of Air Quality Planning and Standards (OAQPS). 2020. "AirNow – AQI Contours (Last 24 hours, Ozone only), US EPA, OAR, OAQPS." Esri ArcGIS. Last modified August 24, 2020.

<https://www.arcgis.com/home/item.html?id=865da054330a4ac7a4c9442bcf00e01d>

Pacific Northwest National Laboratory

902 Battelle Boulevard
P.O. Box 999
Richland, WA 99354

1-888-375-PNNL (7665)

www.pnnl.gov