Geography 222

Laboratory 3

Introduction

In today's laboratory you will learn how to answer spatial questions using standard query language (SQL), and spatial data queries. You will answer questions such as: "how many bike thefts happened within Victoria's downtown core in 2017", and learn how to combine spatial questions across datasets using ArcGIS Online software. Specifically you will acquire the skills to conduct:

- Attribute Joins
- Attribute Queries
- Spatial Queries
- Raster Attribute Queries
- Summary Statistics

Attribute and spatial queries are the foundational spatial analysis tools for any GIS analyst.

An attribute query uses data in an [attribute] table to test conditions. For example: How many neighbourhoods have a population of more than 30,000 people?

A spatial query uses information about how features from two, or more, datasets are located with respect to one another. For example, how many bicycle accidents are within 500m of an intersection?

Exercise

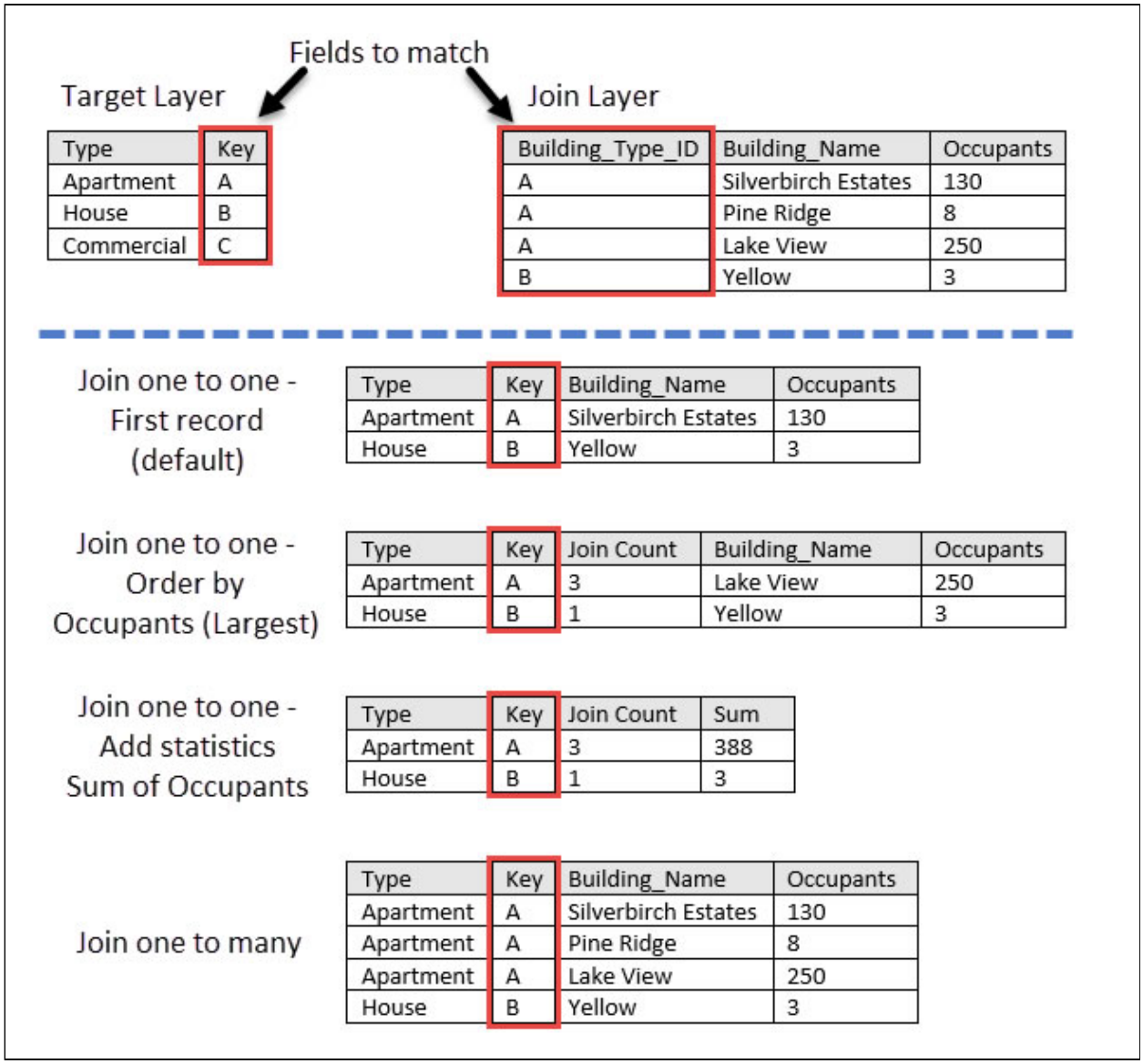
Joins

To begin any spatial query, attribute information (data) must be linked to spatial features (points, lines, or polygons). Attribute joins and spatial joins provide two ways of linking data to spatial locations.

Attribute joins, connect data in tables (e.g., an excel sheet or notepad) to spatial locations using IDs (an attribute column) which match between the table and the shapefile (e.g.,

province name). The join in this case is performed on a row by row basis, so each ID must be unique for each row/feature.

See below for a visualization:



Type	Key
Apartment	A
House	B
Commercial	C

Building_Type_ID	Building_Name	Occupants
A	Silverbirch Estates	130
A	Pine Ridge	8
A	Lake View	250
B	Yellow	3

Type	Key	Building_Name	Occupants
Apartment	A	Silverbirch Estates	130
House	B	Yellow	3

Type	Key	Join Count	Building_Name	Occupants
Apartment	A	3	Lake View	250
House	B	1	Yellow	3

Type	Key	Join Count	Sum
Apartment	A	3	388
House	B	1	3

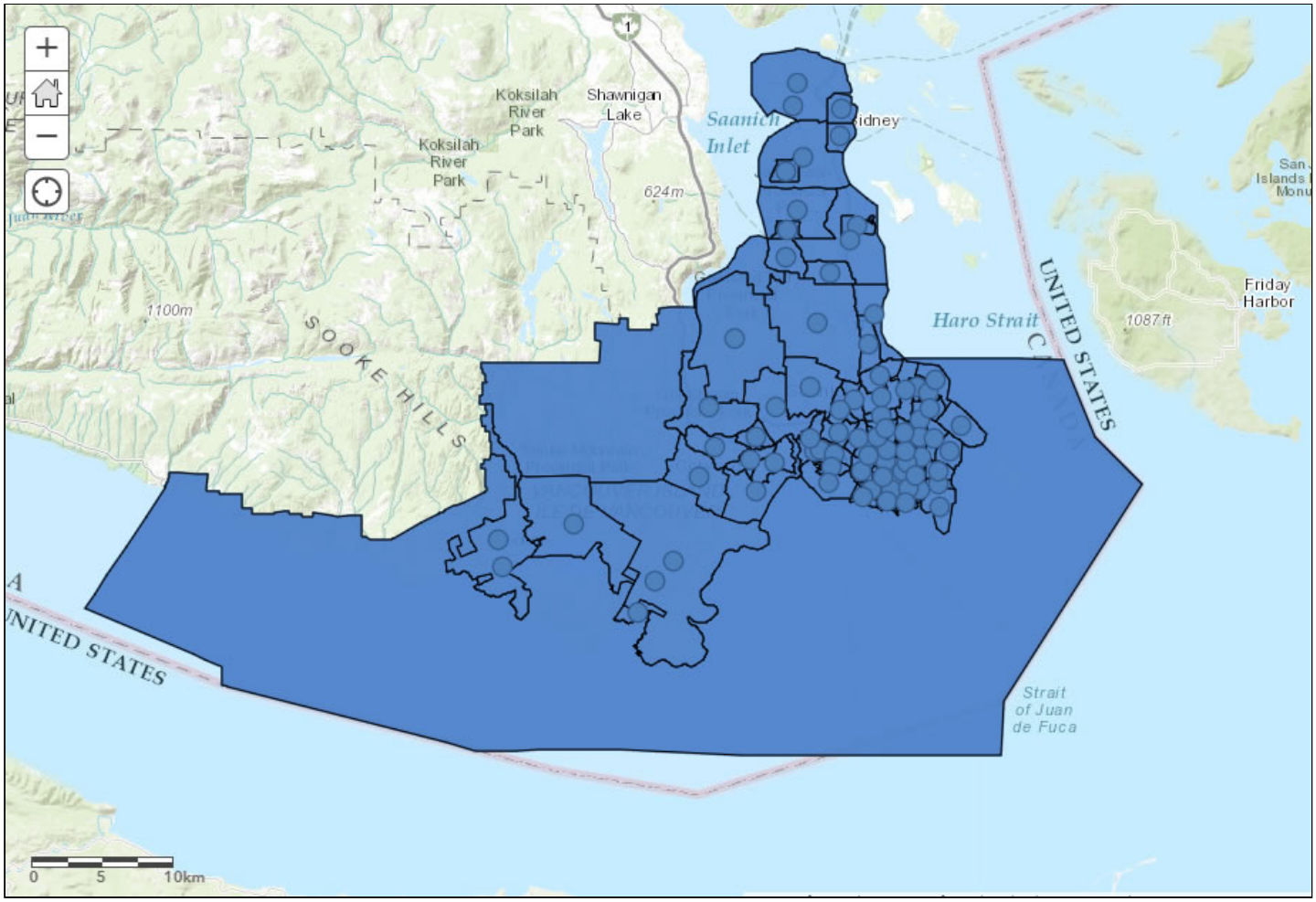
Type	Key	Building_Name	Occupants
Apartment	A	Silverbirch Estates	130
Apartment	A	Pine Ridge	8
Apartment	A	Lake View	250
House	B	Yellow	3

Image source: https://doc.arcgis.com/en/arcgis-online/analyze/join-features.htm

Spatial joins link attributes (data) between two shapefiles using spatial rules. For example, if the points that contain the attribute information are within the polygon the attributes in the point file will be joined to the polygon feature (i.e., one to one join). If multiple points were contained in the polygon, the attribute data would be summarized using statistical rules (e.g., sum, average, max, min). Summarizing attribute information into larger spatial units is called **data aggregation**.

The merge rule dictates how the spatial join will match the attributes. Since only one row will be matched with each polygon in the image below you would select the one to one method. The method will then match the first row (i.e., the only row) in the point file to the polygon. The match option was set to **Completely contains**, so that each polygon has to completely contain the point for it to be spatially joined.

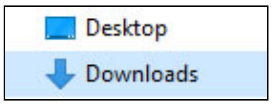
See image below:



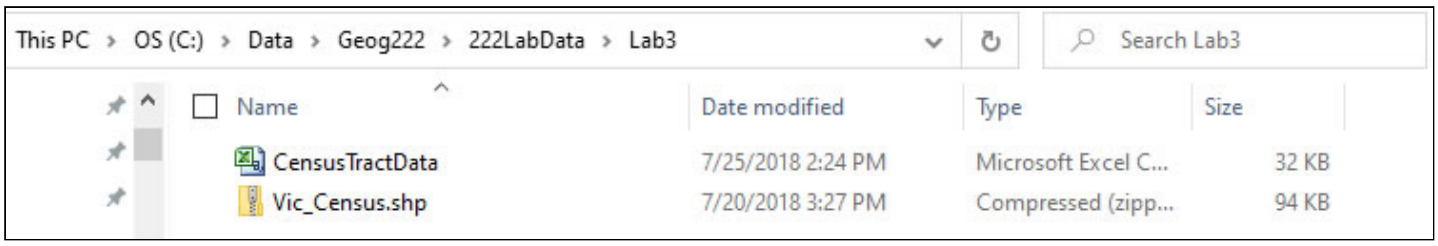
In the following example, you will link 2016 census information to the census tract boundaries using an attribute join in QGIS.

1. First, download the census boundary file (Vic_Census.shp) and the census tract data into your Lab3 folder.

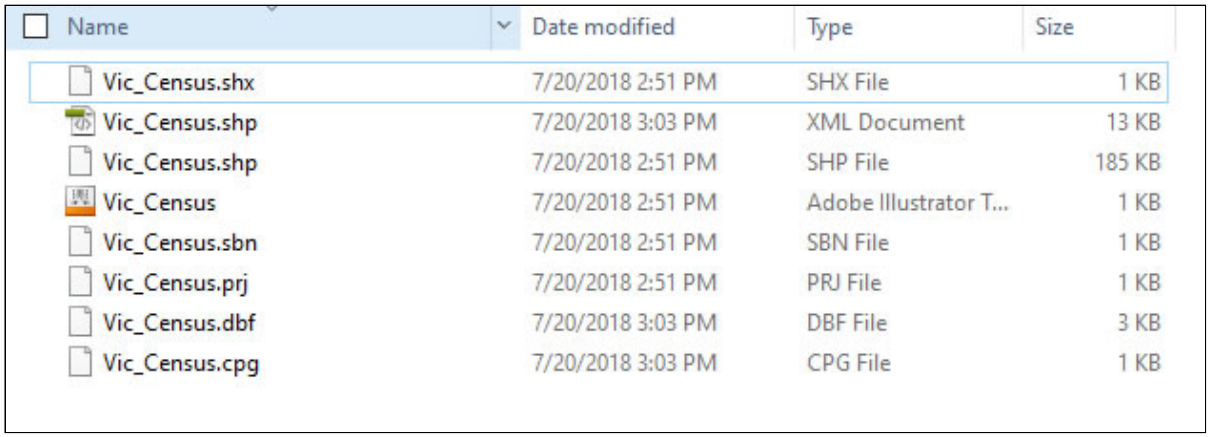
- Once you have clicked the links, and saved the files, the datasets will be available from your downloads folder



- Copy and paste the files from the downloads folder into your Lab3 folder

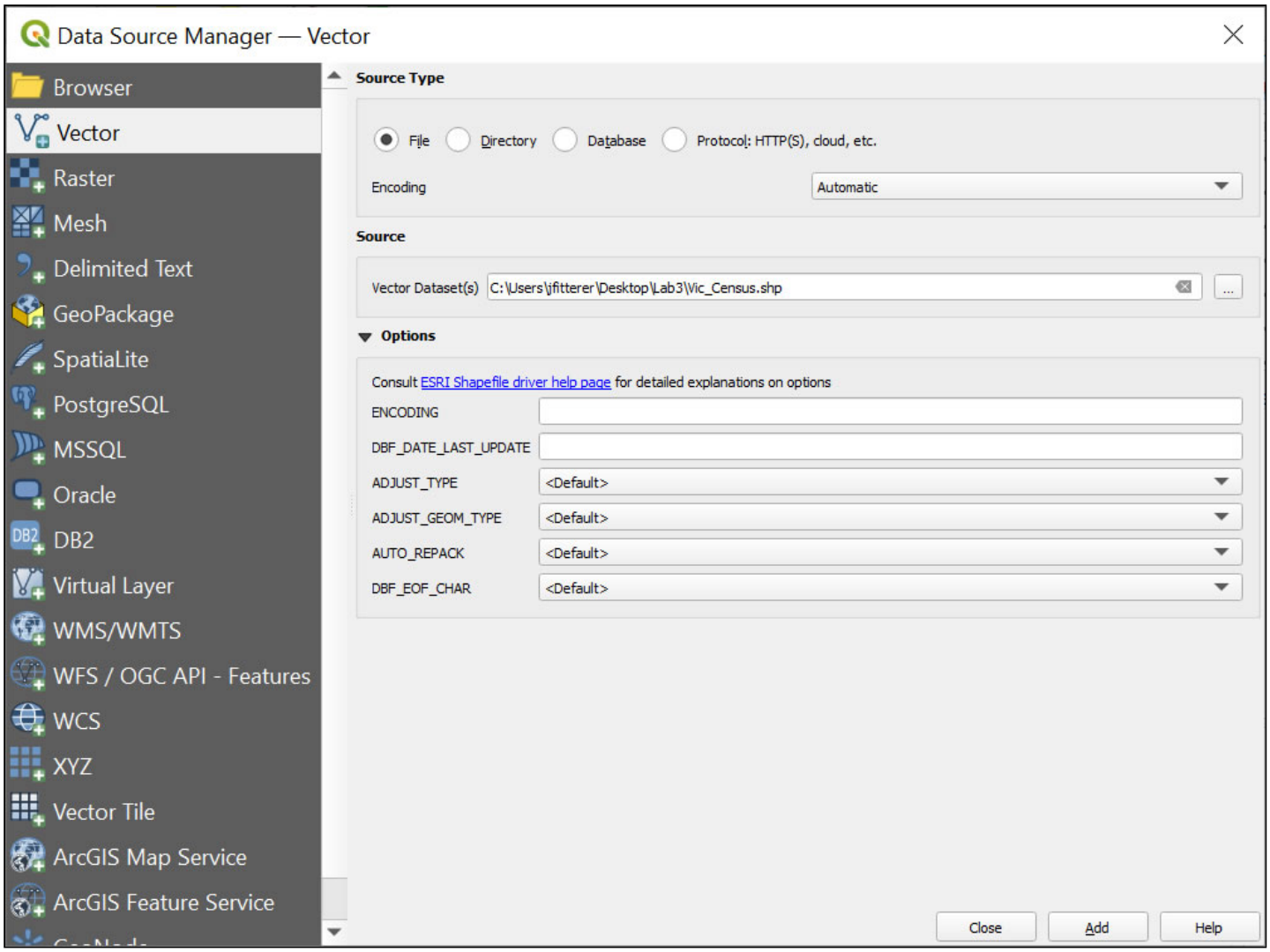


- Right click on the **Vic_Census.zip** file → and select "Extract All"

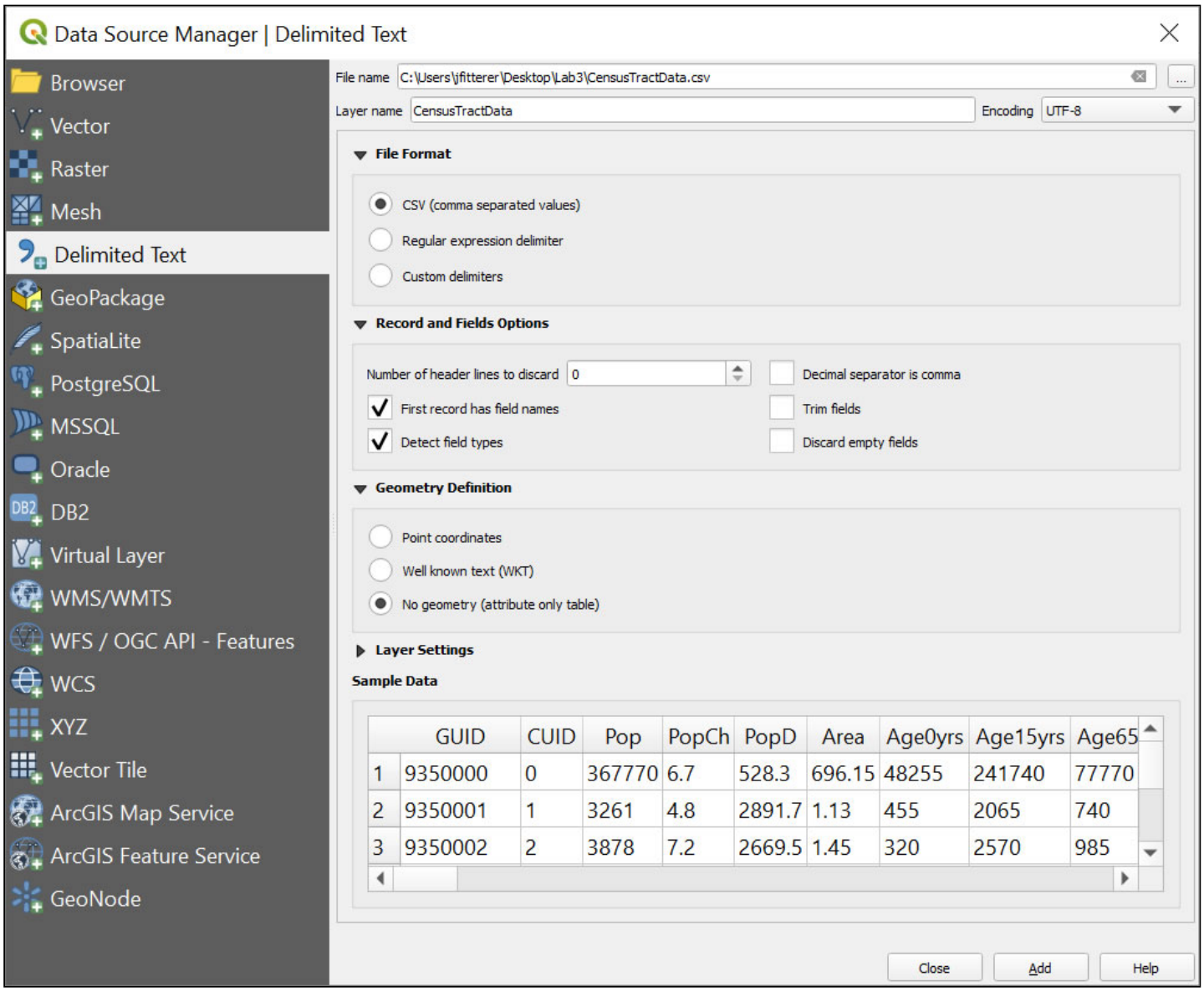


2. Open QGIS

3. Add the **Vic_Census.shp** and the **CensusTractData.csv** files into QGIS:

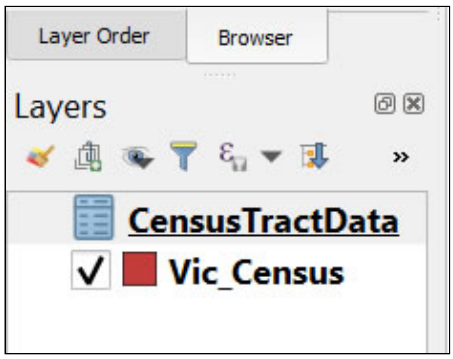


- Press Add → Close



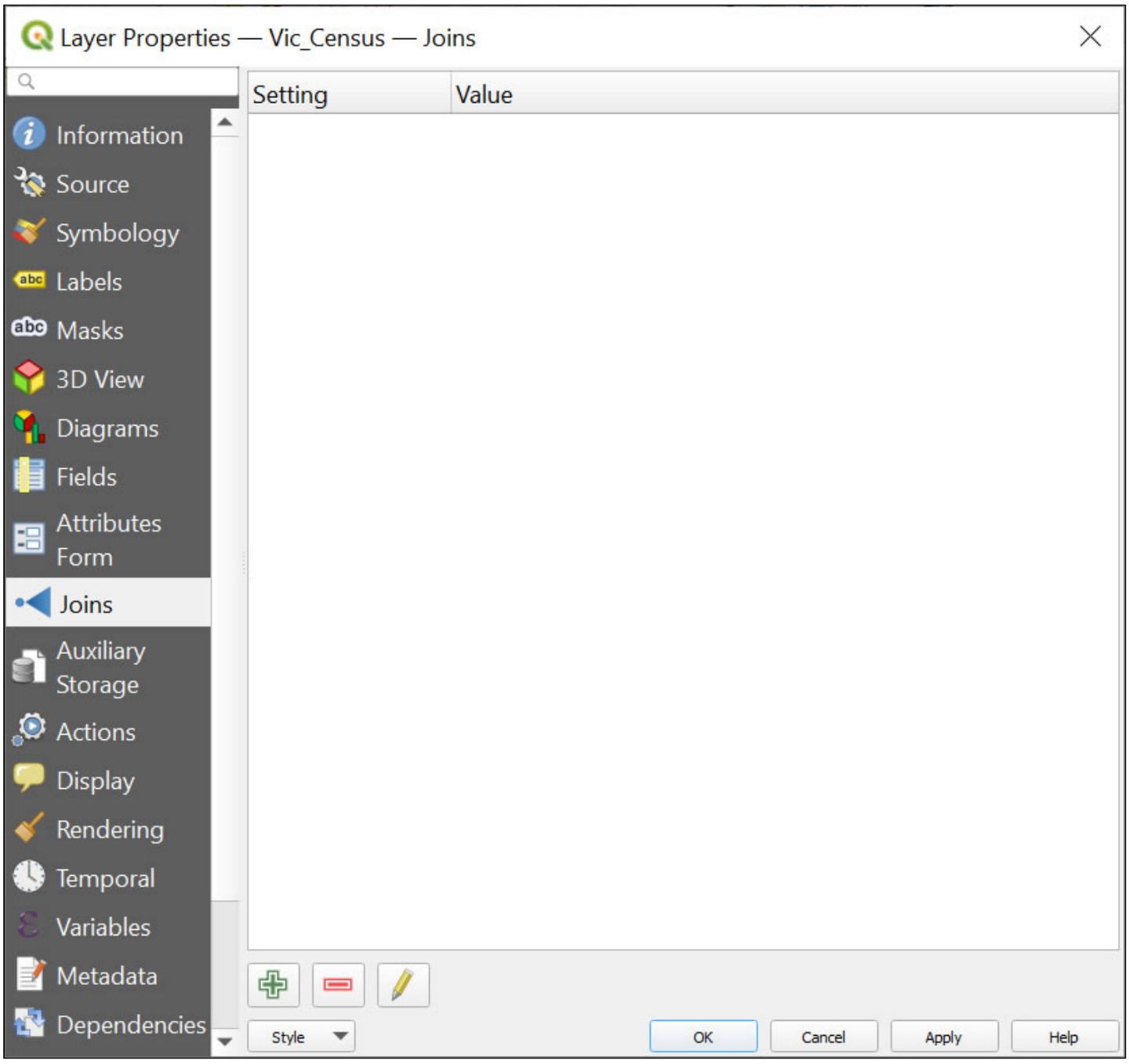
- Press Add → Close

Now you will see two layers uploaded to your contents in QGIS



The objective of the next step is to join the **Census Tract Data** table, which contains census data about each region, to the **Vic_Census** boundary. To do this, you will match the **GUID** columns in both files. The GUID is the geographic ID that is unique to each census tract (row).

4. To perform the attribute join, right click on the **Vic_Census** file → select **properties** → choose the **Join** tab



5. In the **Join** tab select the **plus** button 

Set the join options as:

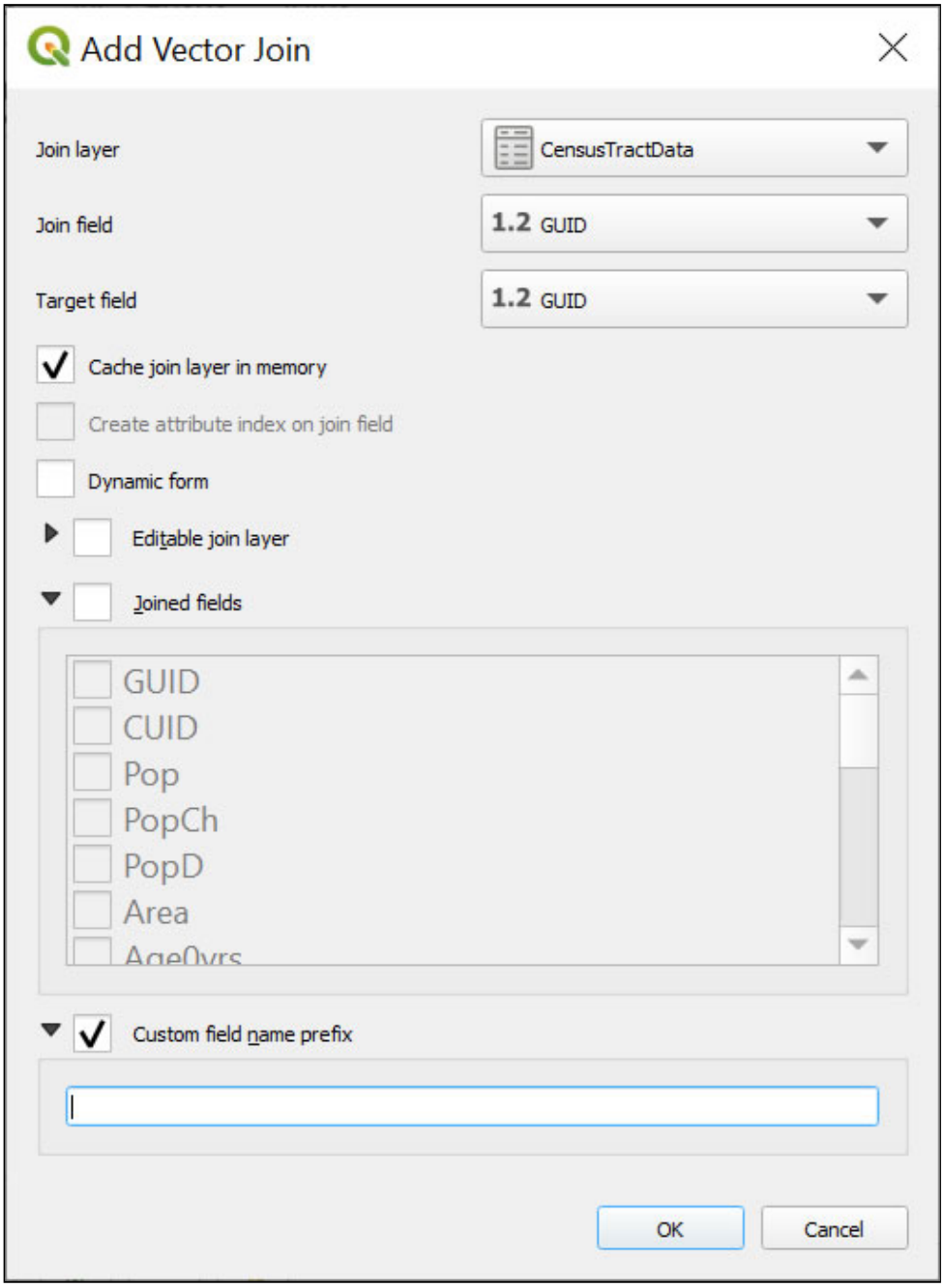
Join layer: CensusTractData

Join Field: GUID

Target Field: GUID

Check **Custom field name prefix**

- Delete the custom field name prefix

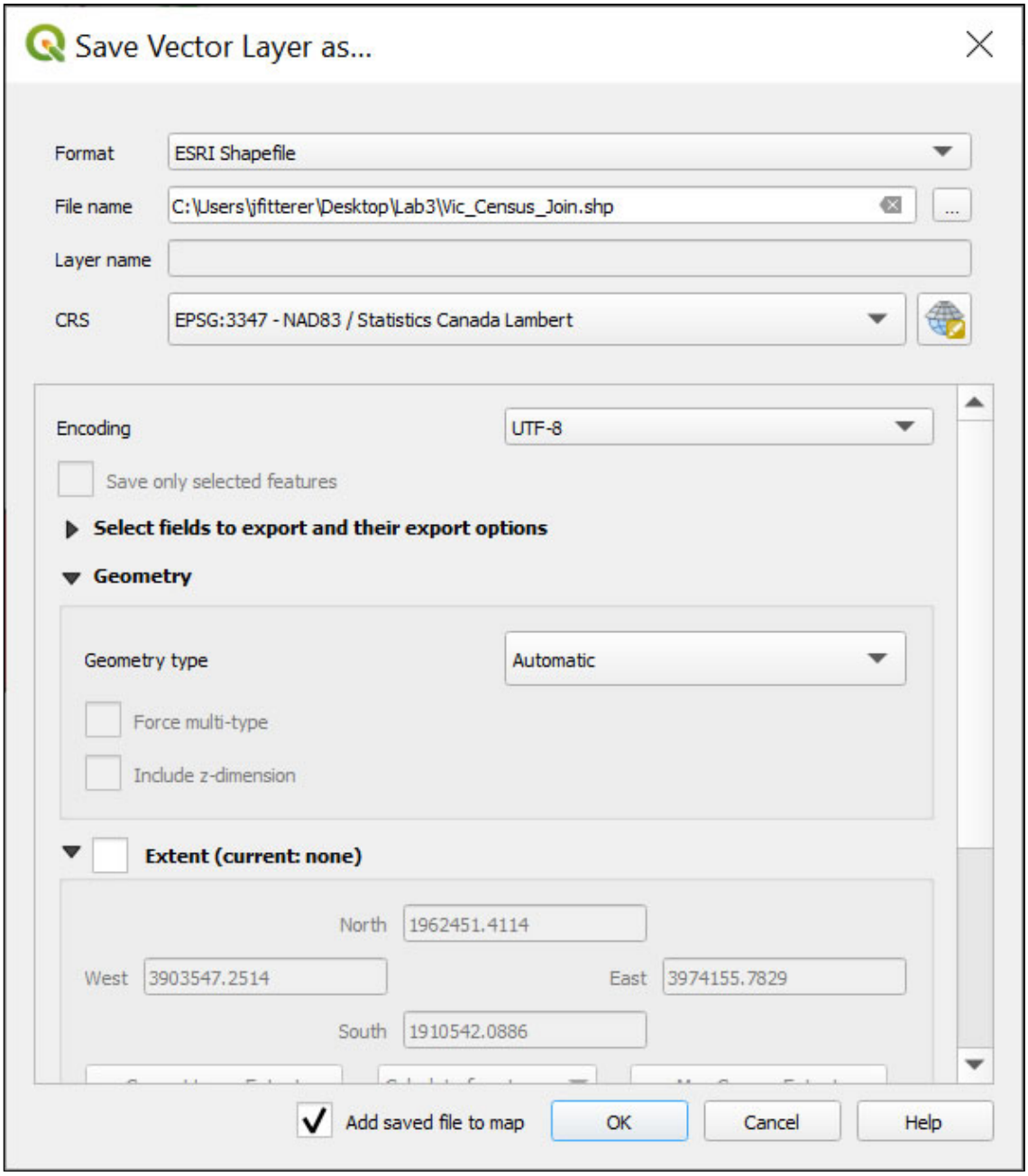


- Press **OK** → Apply → OK

Now you need to export the file in a new name to complete the join.

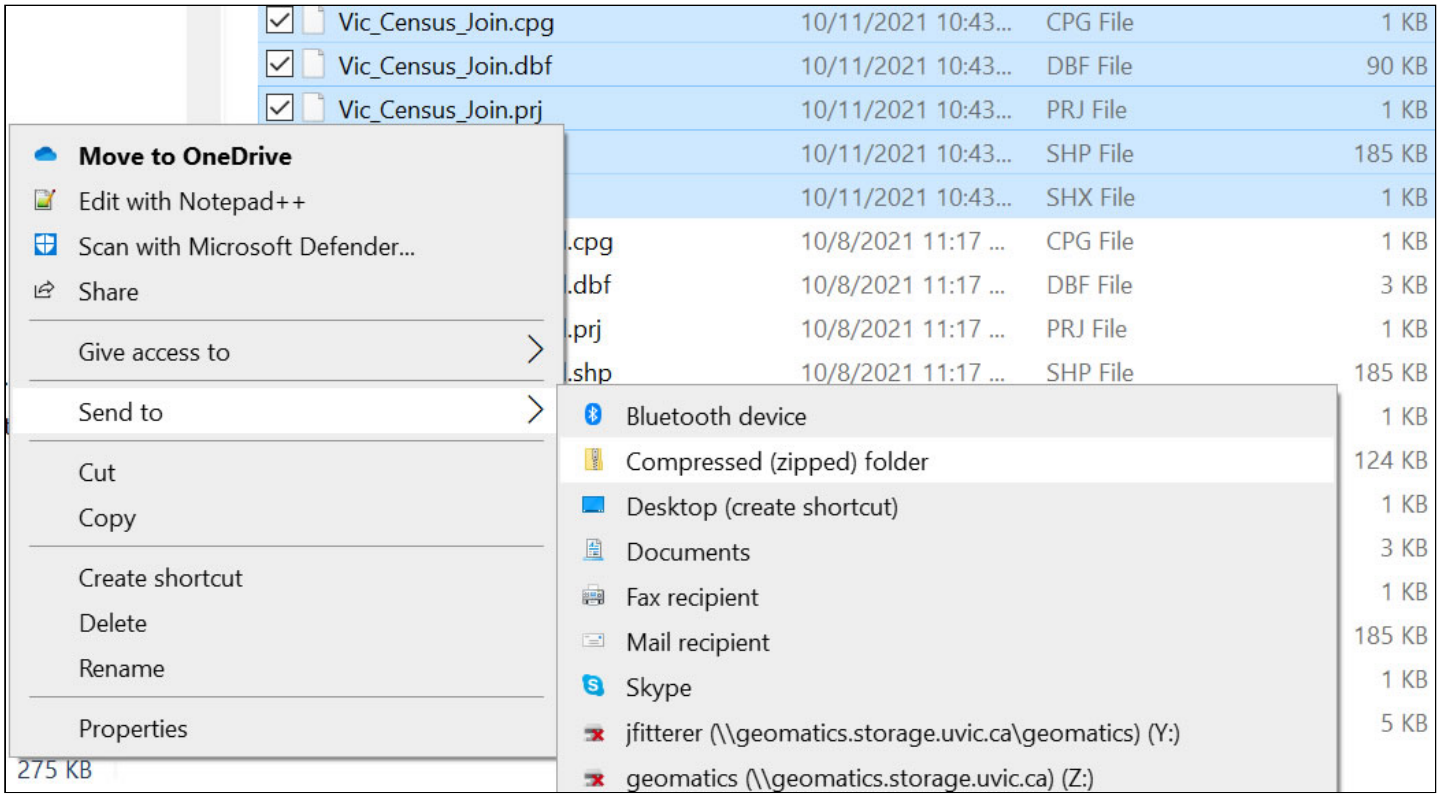
6. From the Layers menu in QGIS right click on the **Vic_Census** file → select **Export** → **Save Features As...**

- Save the features as **Vic_Census_Join** in your Lab 3 folder → press **OK**



Now you will need to zip the new shapefile in a compressed folder so you can load it into ArcGIS Online

7. Navigate to your **Lab 3** folder → select the **Vic_Census_Join** files → send to: **Compressed (Zipped) folder**



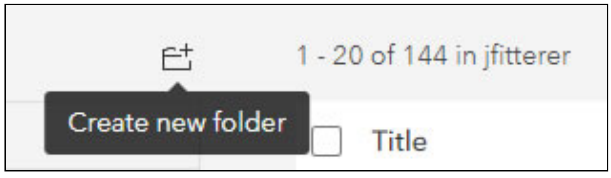
8. Now your file is ready to upload to ArcGIS online

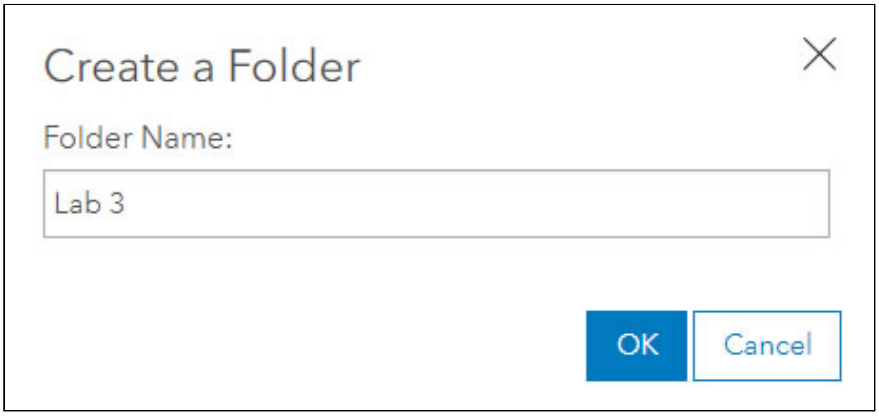
Importing Data to ArcGIS Online

1. Login to ArcGIS online with the credentials emailed to your UVic registered email. If you can't find the email search for "ArcGIS Online Notifications"

2. Once you are into ArcGIS Online, go to the **Content** → **My Content** tab

- **Create a new folder** called "Lab 3"

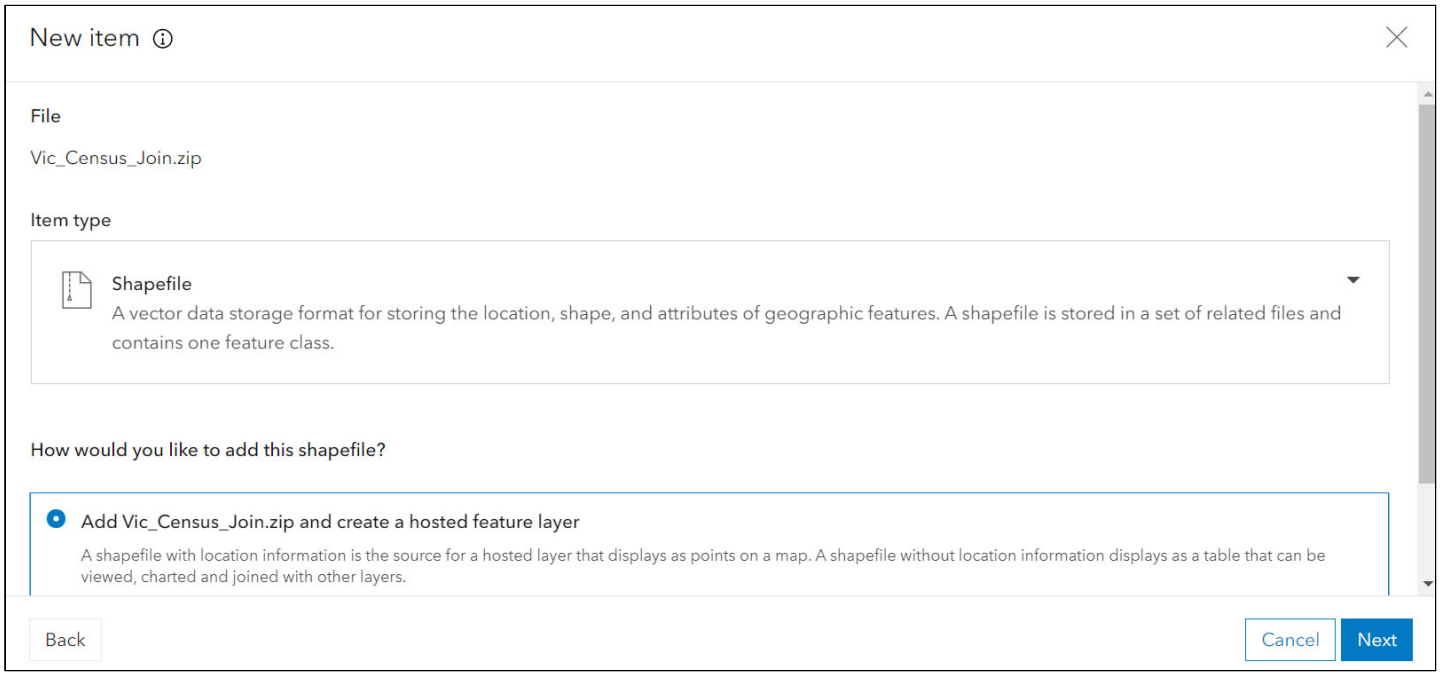




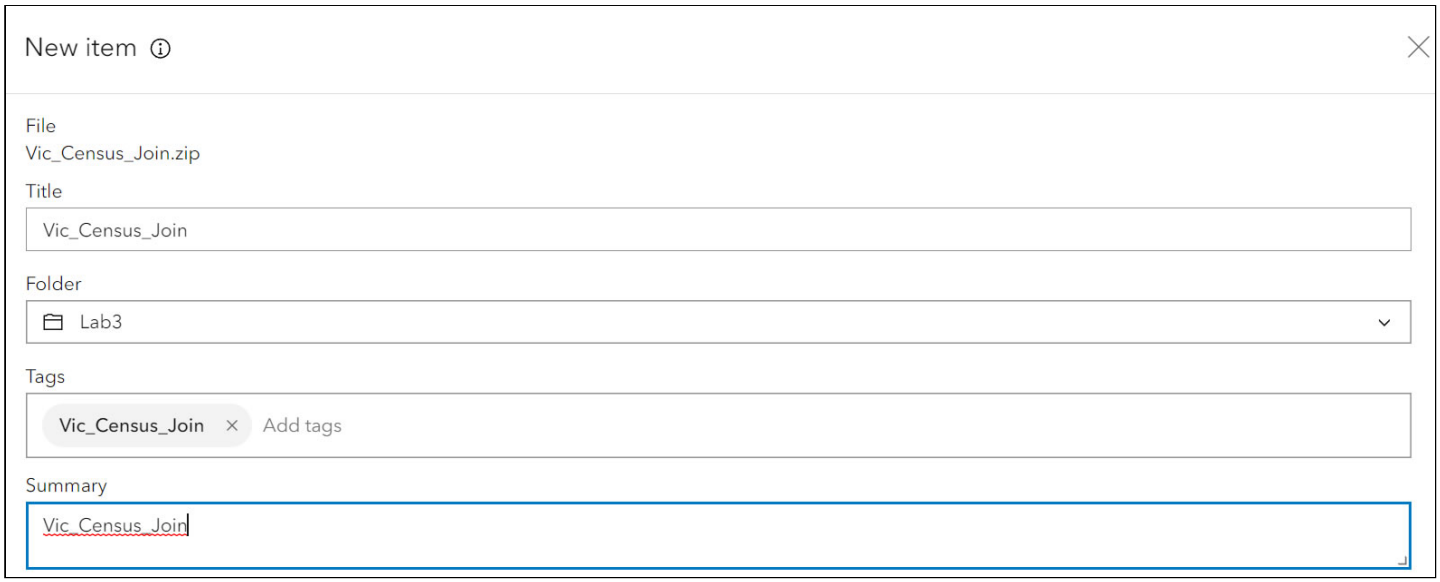
3. Once in the **Lab 3** folder → Select "**New Item**"

4. Drag and drop the **Vic_Census_Join.zip** file into the **New Item** window

- Choose to add the file as a "hosted feature layer"

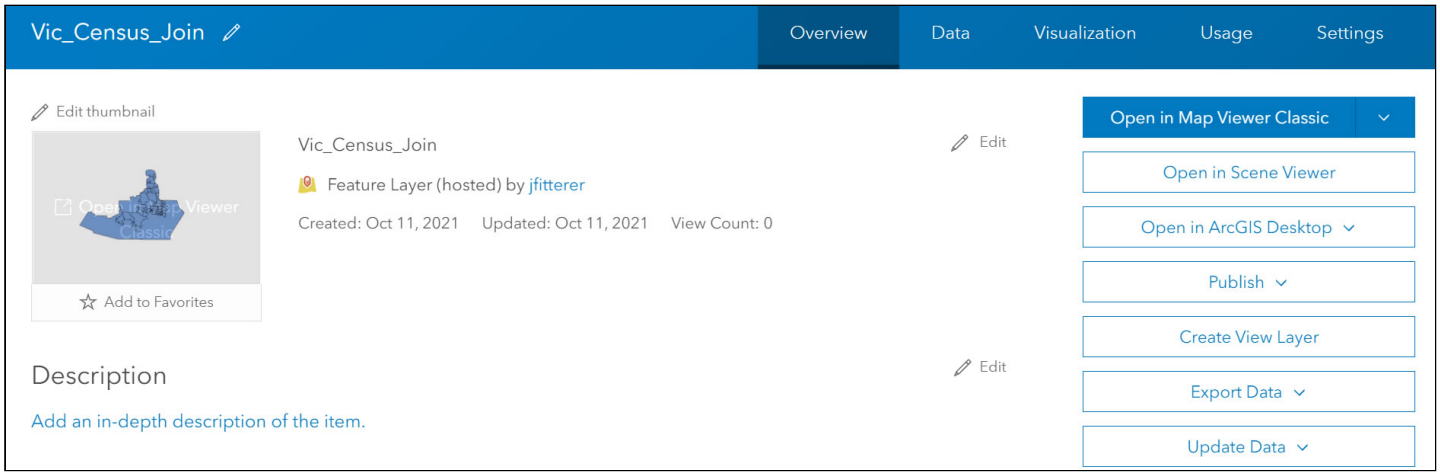


- Press "**Next**"
- Set the Tags and Summary as "**Vic_Census_Join**" **add your netlink ID to the end of the name if you get a warning about a hosted layer having the same name**



Press **Save**

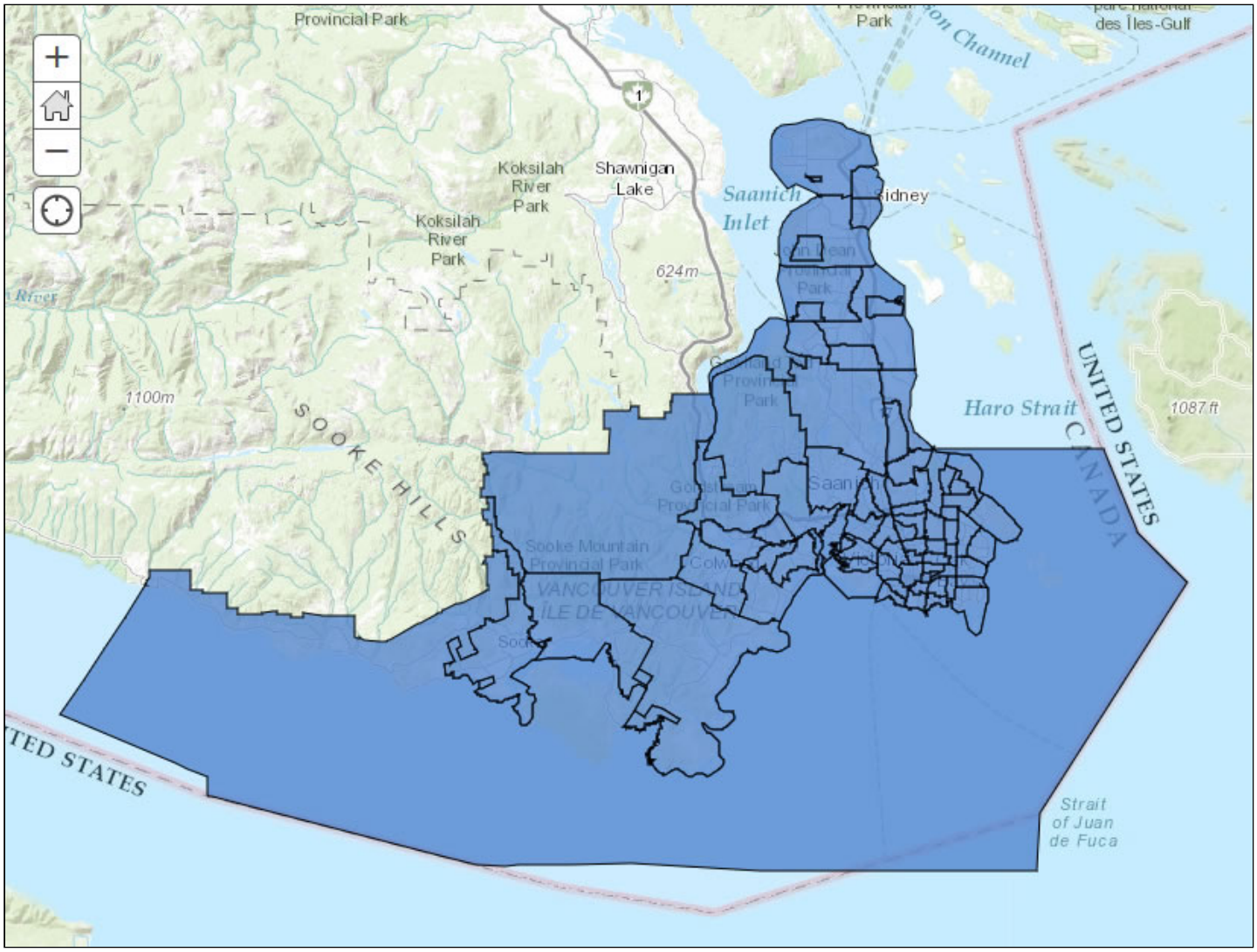
The layer is now uploaded to ArcGIS online:



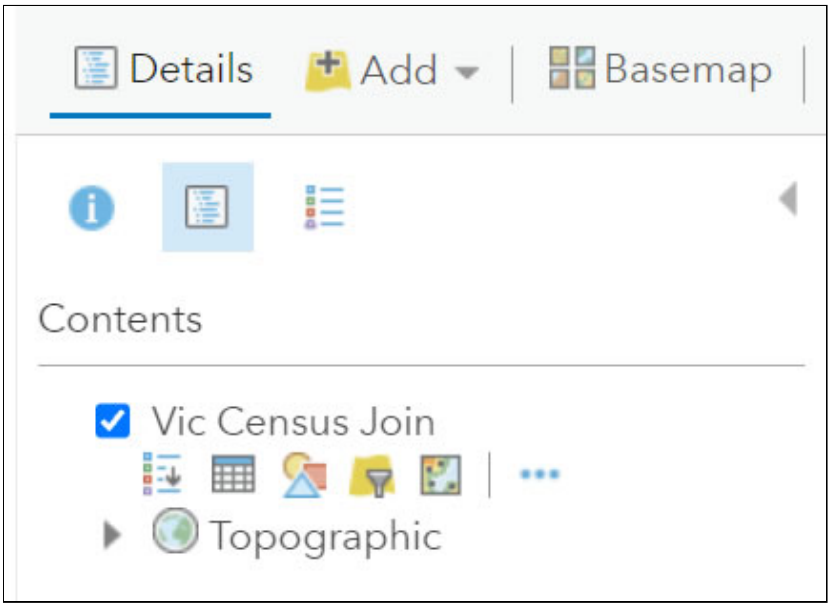
5. Select to open in **"Map Viewer Classic"**

6. Once loaded, select to **"show only locations"** for the symbology

Now you will see the census tracts on your map. Scroll to zoom in and out of the map.



8. Once the datasets are added you can select the **Details** → **Content** options to view attribute tables and set symbology



Vector Attribute Queries

Before you begin any attribute queries it's important to understand what each of the headings in the census data file represent. Visit the list here before proceeding.

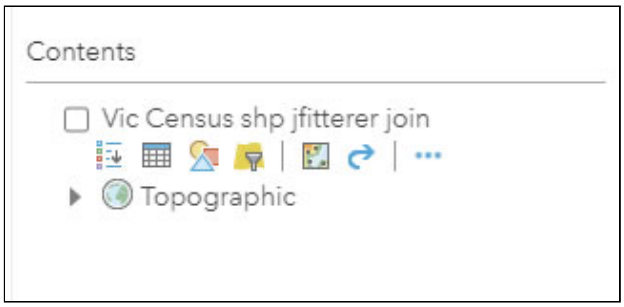
Also read the ArcGIS summary of SQL operators and example query structure.

Attribute queries answer two fundamental questions:

- What is here?
- Where is this?

A 'what is here' query is only a mouse click away from being answered.

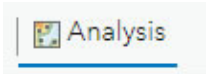
1.) In the Contents pane, have the **Vic Census Join** file showing



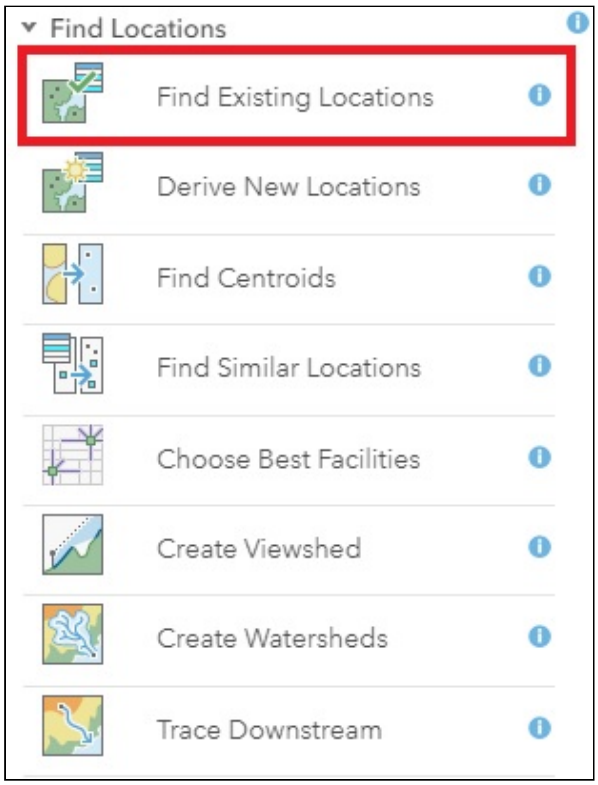
A 'where is this' query is used when you know exactly what attribute you are searching for and want to determine where it is located on the map.

For example, you may want to know how many census tracts have a population greater than 9000 people, and of these selected census tracts what the average population change was.

To answer the question above, you could open the attribute table and sort the Pop column to find out. However, this method would take too long using a large database, and you wouldn't be using the full capabilities of GIS. Instead you should use the **Find Existing Locations** tool.

1.) From the Analysis button 

→ choose **Find Locations** → select **Find Existing Locations**



Populate the tool as folows:

1. **Vic_Census_Join**

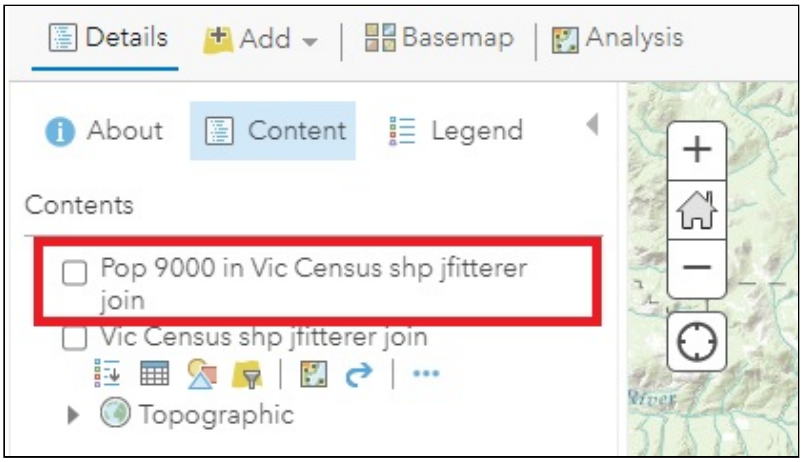
2. Add Expression

3. Results layer name: **Pop 9000 in Vic_Census_Join**

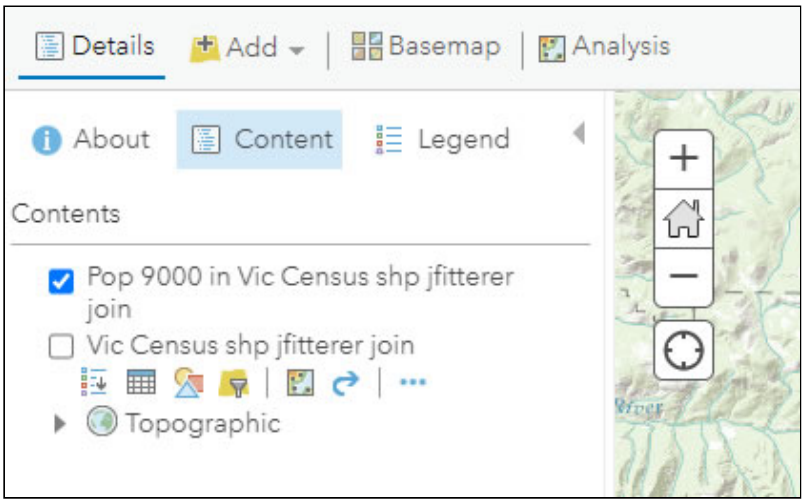
4. Uncheck "Use current map extent"

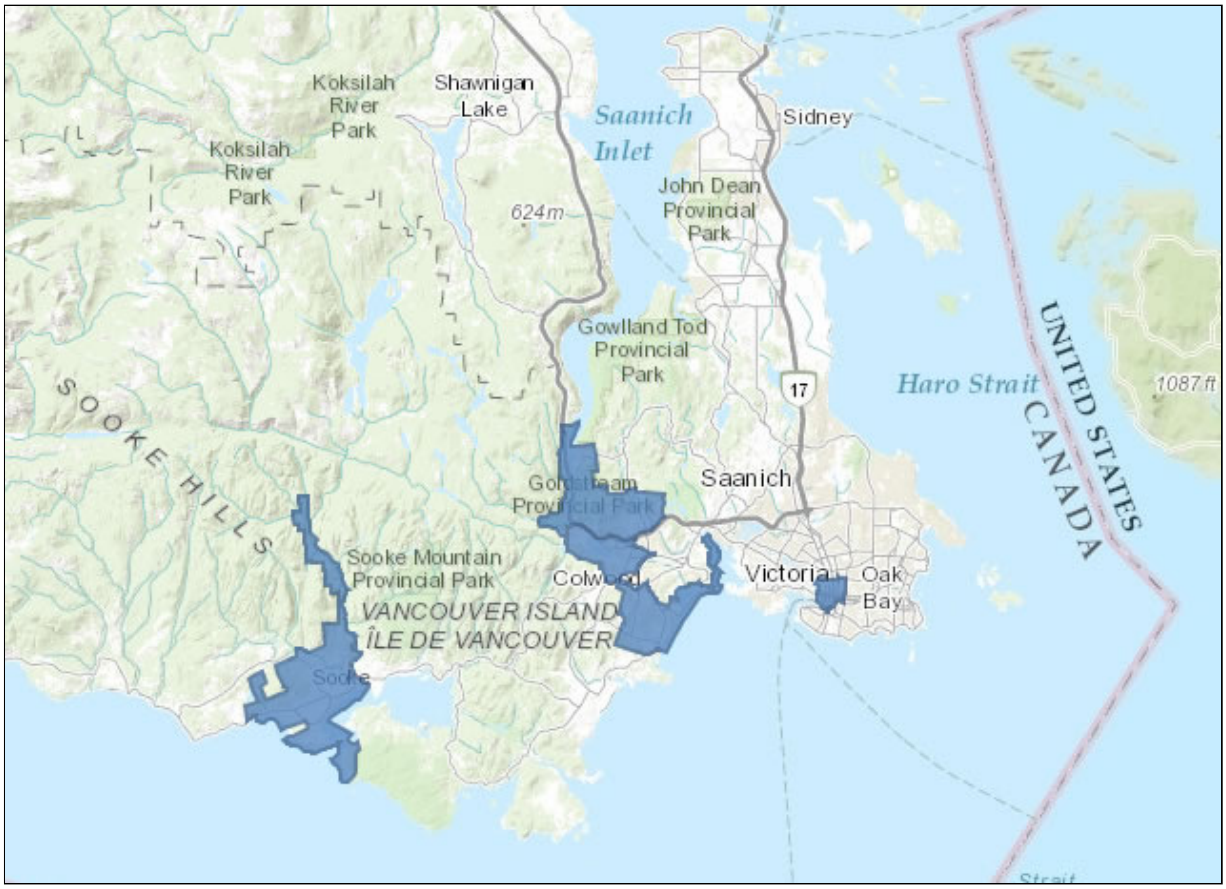
Press: **Run Analysis**

The results will be presented in a new shapefile in the table of contents



4. Uncheck the **Vic_Census_shp_netlinkID** and the **Vic_Census_shp netlinkID_join_layers** to see the query results





4.) Open the **Pop 9000 in Vic Census netlinkID join** attribute table to view the selected census tracts.

Observe that 5 of the total 78 census tracts were selected

Pop 9000 in Vic Census shp jfitterer join (Features: 5, Selected: 0)

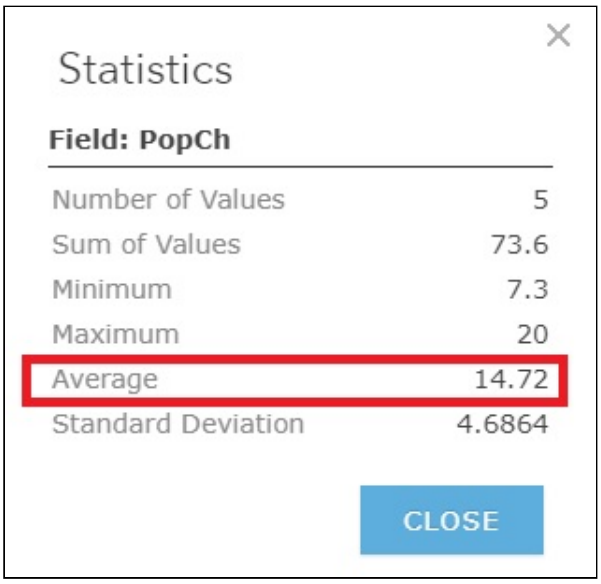
5.) If you wanted to know the average population change for these selected census tracts, you could click on the **PopCh** column and select statistics.

Pop 9000 in Vic Census shp jfitterer join (Features: 5, Selected: 0)

CTUID	GUID	CUID	Pop	PopCh	PopD	Area
9350010.00	9,350,150.03	10.00	9,207		51.50	1.86
9350151.02	9,350,150.03	151.02	9,363		65.30	6.39
9350151.03	9,350,150.03	151.03	10,030		.10	16.72
9350154.01	9,350,150.03	154.01	9,277		49.20	8.84
9350156.06	9,350,150.03	156.06	10,393	16.70	424.30	24.49

Sort Ascending
Sort Descending
Statistics
Calculate
Delete

6. The statistics results provide the average population change:

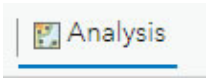


Logical Operators

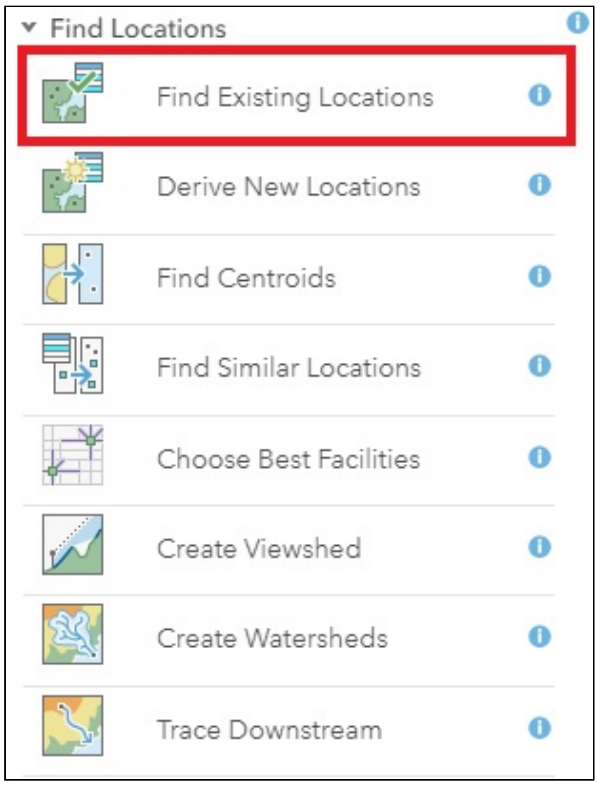
Now consider that you want to build an attribute query that will satisfy more than one clause. To do this, you would have to use a logical operator (AND, OR).

An example of a question you may ask with the AND operator would be: **Which census tracts have a population greater than 9000 people, and a land area less than 15 square kilometers?**

In this query you are asking for both clauses to be true in order to select the census tract. The census tract must have a population greater than 9000 people, **and** a land area that is less than 15 kilometers.

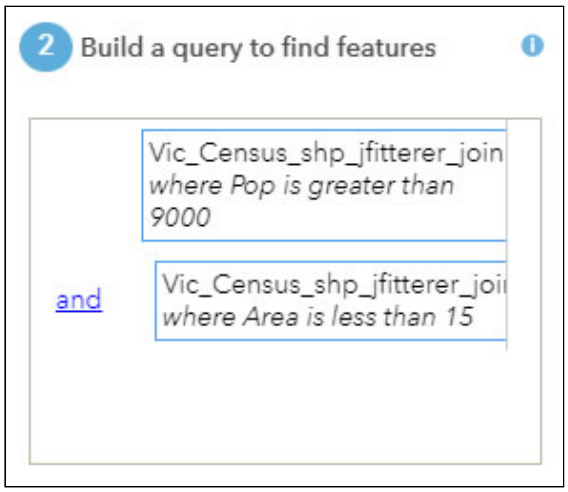
1. From the Analysis button 

→ choose **Find Locations** → select **Find Existing Locations**



Populate the tool as folows:

- **1. Vic_Census_Join**
- **2. Add Expression**



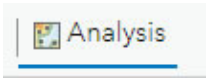
- 3. Result layer name: **Pop Land in Vic_Census_Join**
- 4. Uncheck "Use current map extent"
- Press **Run Analysis**

2. Open the **Pop Land in Vic Census shp netlinkID** attribute table to see the selected census tracts. Three of the original 78 are selected.

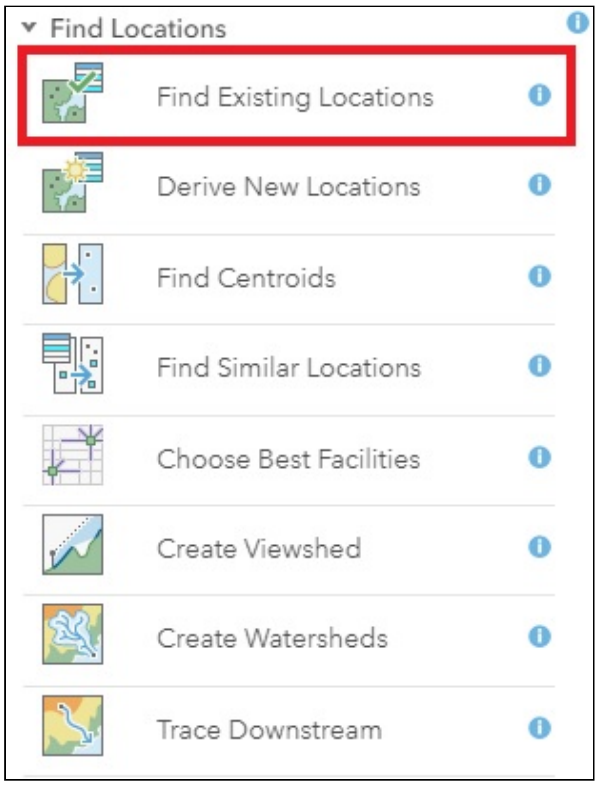
Pop Land in Vic Census shp jfitterer join (Features: 3, Selected: 0)

CTUID	GUID	CUID	Pop	PopCh	PopD	Area	Age0yrs
9350010.00	9,350,150.03	10.00	9,207	15.50	4,961.50	1.86	430
9350151.02	9,350,150.03	151.02	9,363	20.00	1,465.30	6.39	1,400
9350154.01	9,350,150.03	154.01	9,277	7.30	1,049.20	8.84	1,405

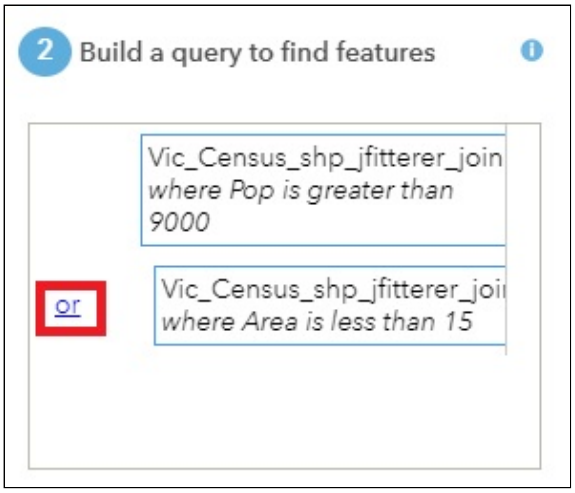
Now lets consider an OR operator question: Which census tracts have a population greater than 9000 people, or a land area less than 15 square kilometers?

3. From the Analysis button 

→ choose **Find Locations** → select **Find Existing Locations**



- 1. **Vic_Census_Join**
- 2. Build a query:



- 3. Result layer name: **Pop or Land Locations in Vic_Census_Join**
- 4. Uncheck "Use current map extent"
- Press **Run Analysis**

4. Open the **Pop or Land in Vic Census shp netlinkID** attribute table to see the selected census tracts. 70 of the original 78 are selected.

Spatial Queries

Spatial queries fall under one of two categories: adjacency or proximity.

Adjacency refers to 'having a common end point or border'.

You can perform adjacency searches on line or polygon layers **only**. Points are zero-dimensional and therefore have no way of being adjacent to each other, or to any other feature.

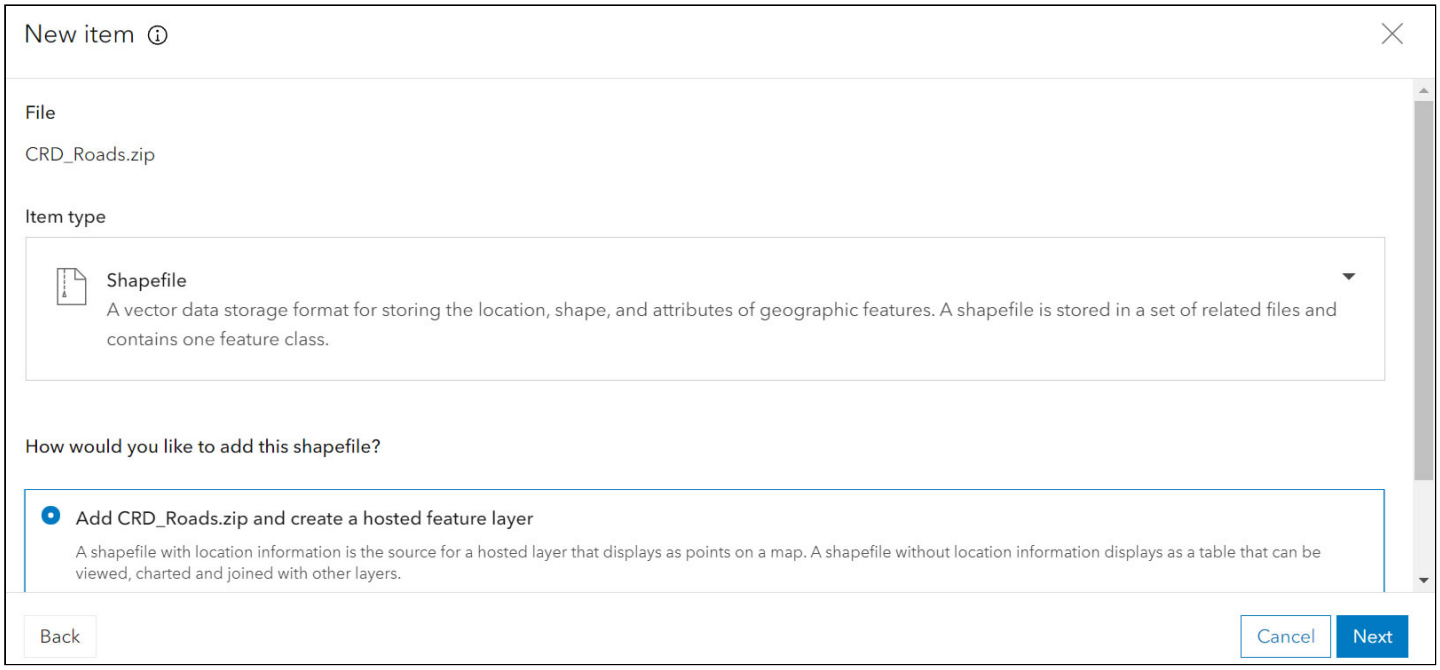
Let's start by finding out the relationship between the census tract and a road layer. **We want to find out how many Census Tracts intersect with Cook Street.**

1. Download the CRD_Roads.zip folder to your Lab 3 folder

2. In the **Content** tab in **ArcGIS Online**

3. Press the **Add Item** button

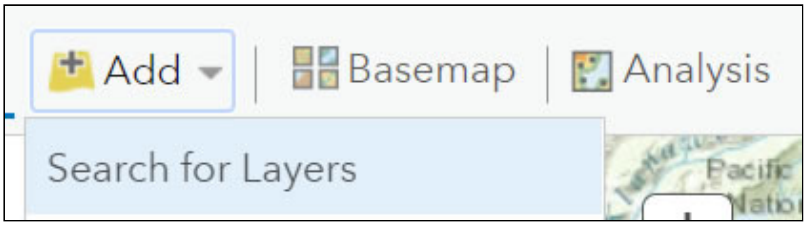
- Drag and drop the **CRD_Roads.zip** file into ArcGIS online



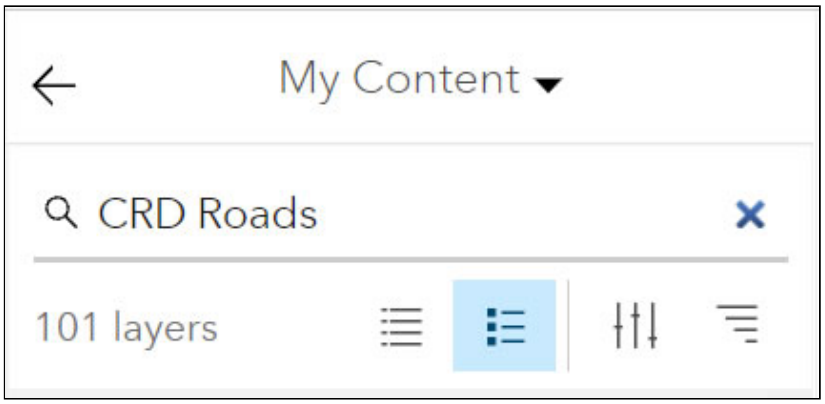
- Add the **CRD Roads.zip** as a **"hosted feature layer"**
- Set the Tags and Summary to: **CRD_Roads**
- Press **"Save"**

Now the layer is available to upload into the map viewer

4. Open in the map viewer with the **Vic_Census_Join** file uploaded
5. Press **"Add"** and then **"Search for Layers"**

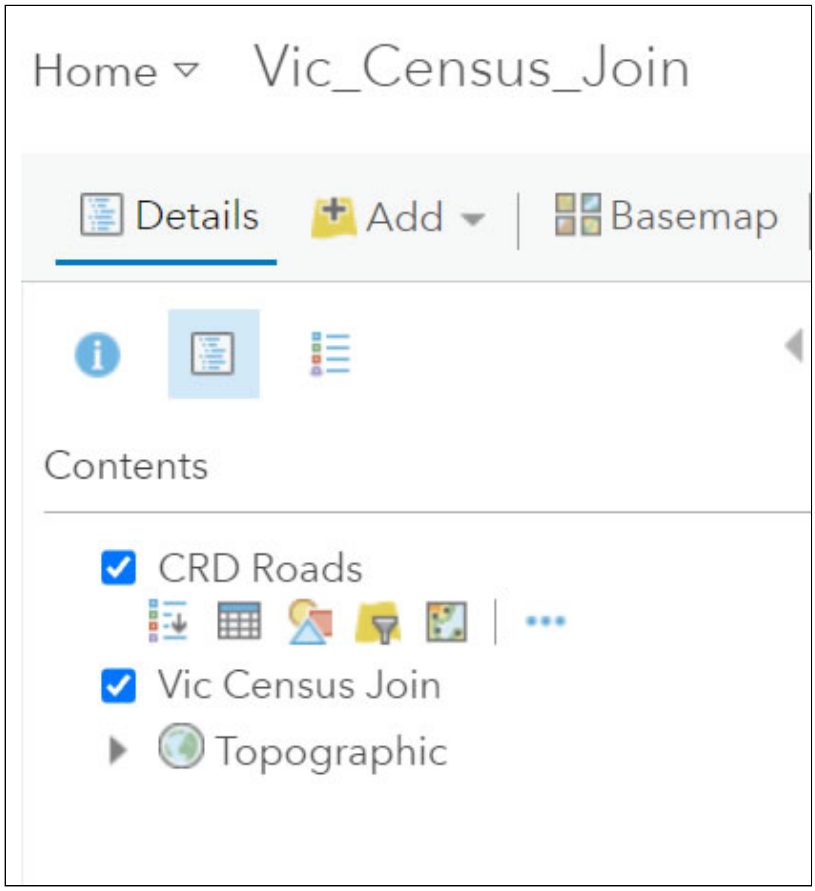


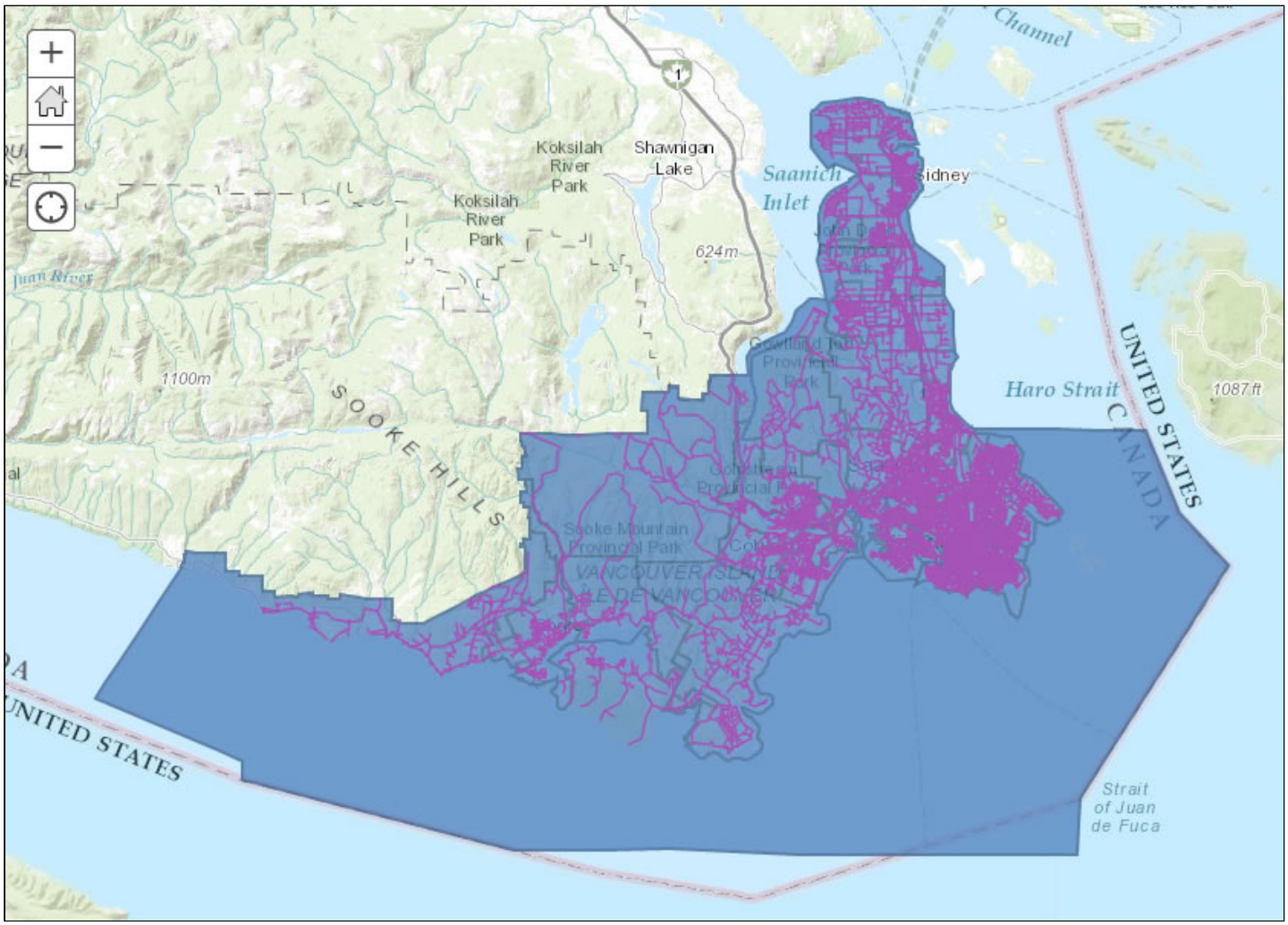
- Search for: CRD_Roads



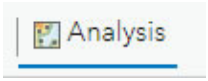
- Select the CRD_Roads layer → press "Add to Map"

The Roads layer is now uploaded to your Content pane

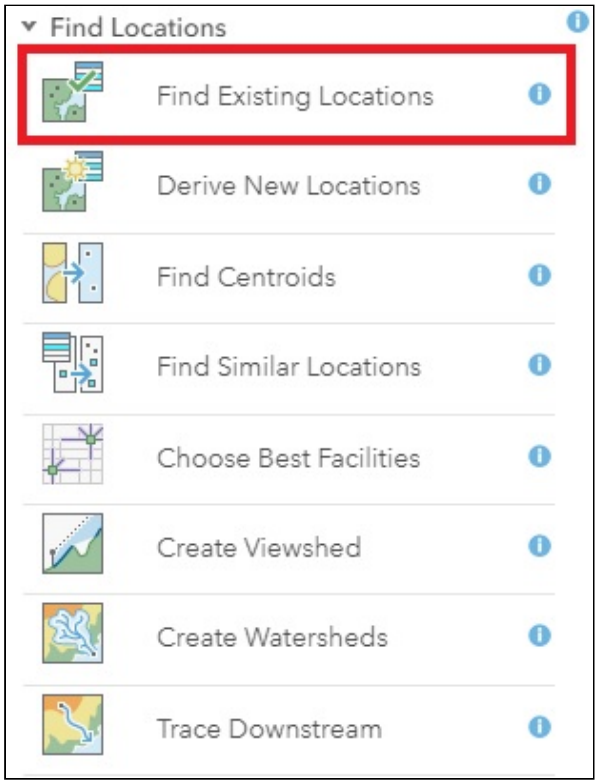




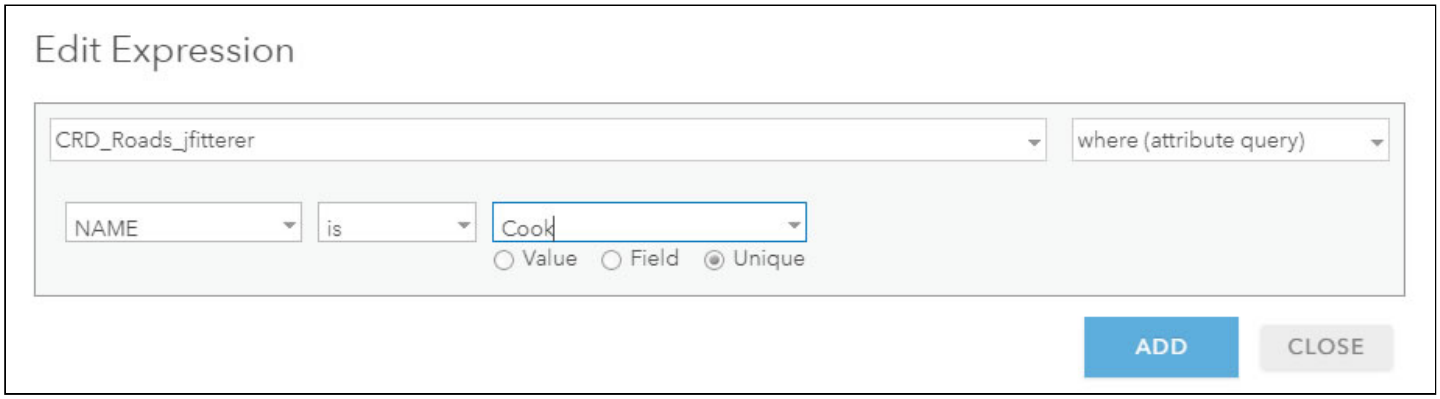
Now you will use the **Find Existing Locations** tool to select Cook St.

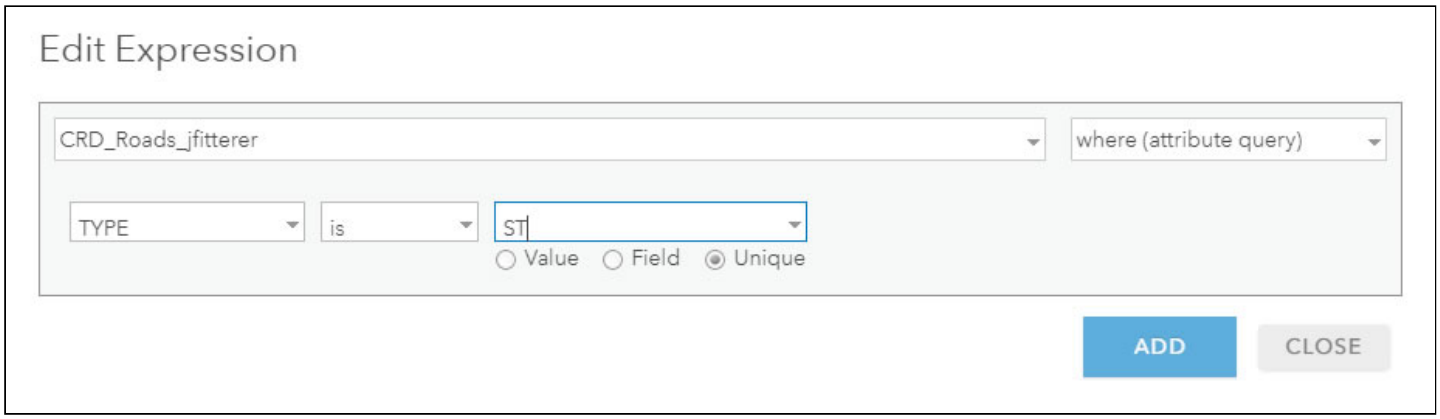
8. From the Analysis button 

→ choose **Find Locations** → select **Find Existing Locations**



- 1. **CRD_Roads**
- 2. Add an expression

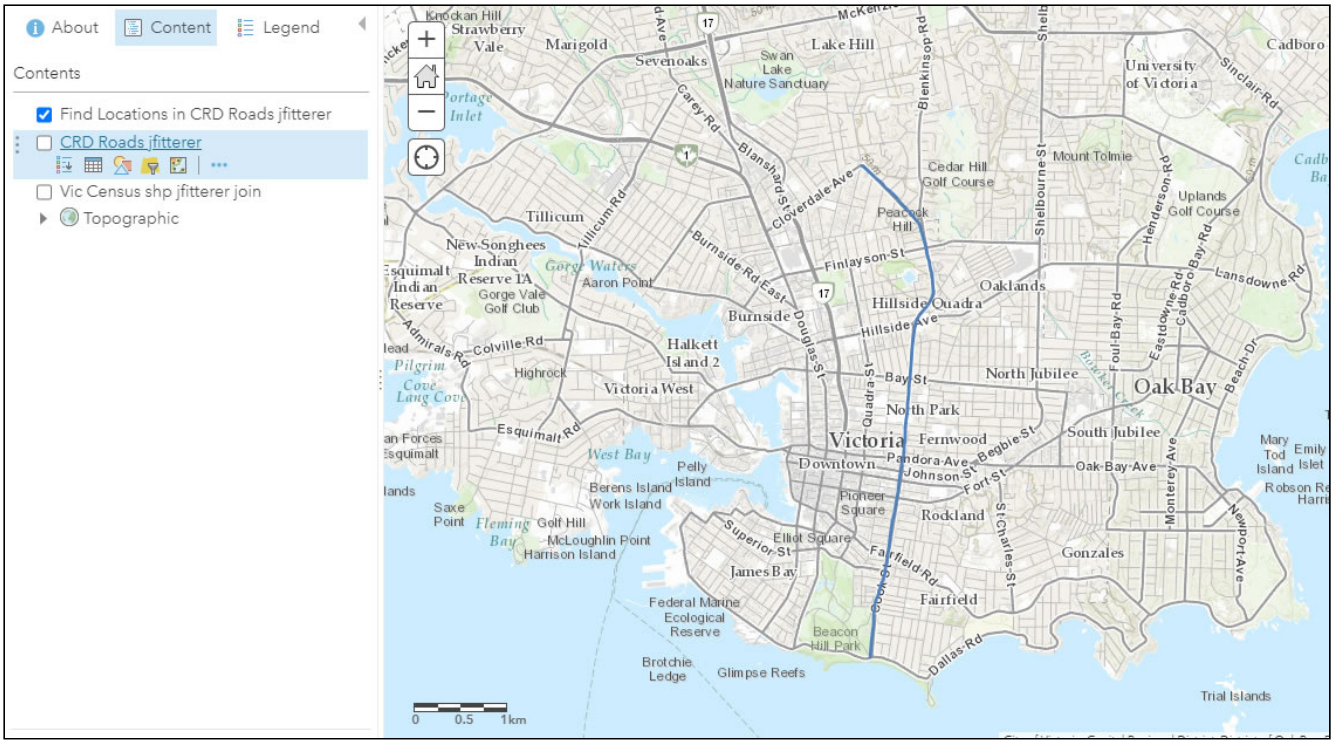




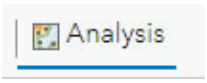
- 3. Result layer name: **Find Locations in CRD_Roads**
- 4. Uncheck "Use current map extent"

- Press **Run Analysis**

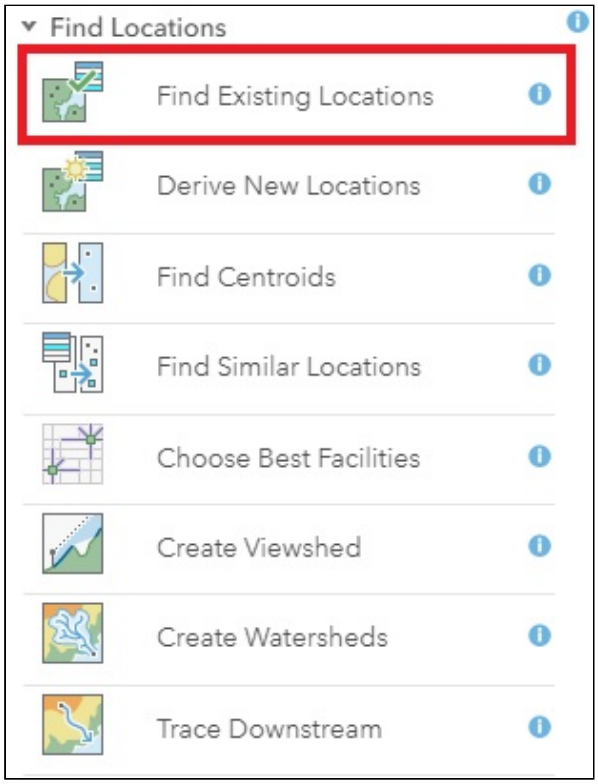
9. Uncheck the census tract layers and view the selected road



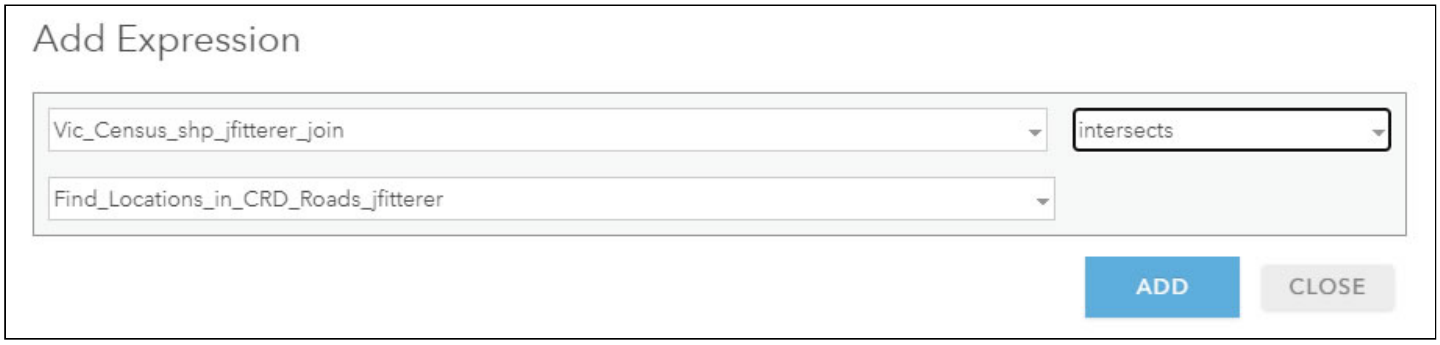
Now you are ready to conduct the spatial adjacency query to find the Census Tracts that intersect with Cook St.

10. From the Analysis button 

→ choose **Find Locations** → select **Find Existing Locations**

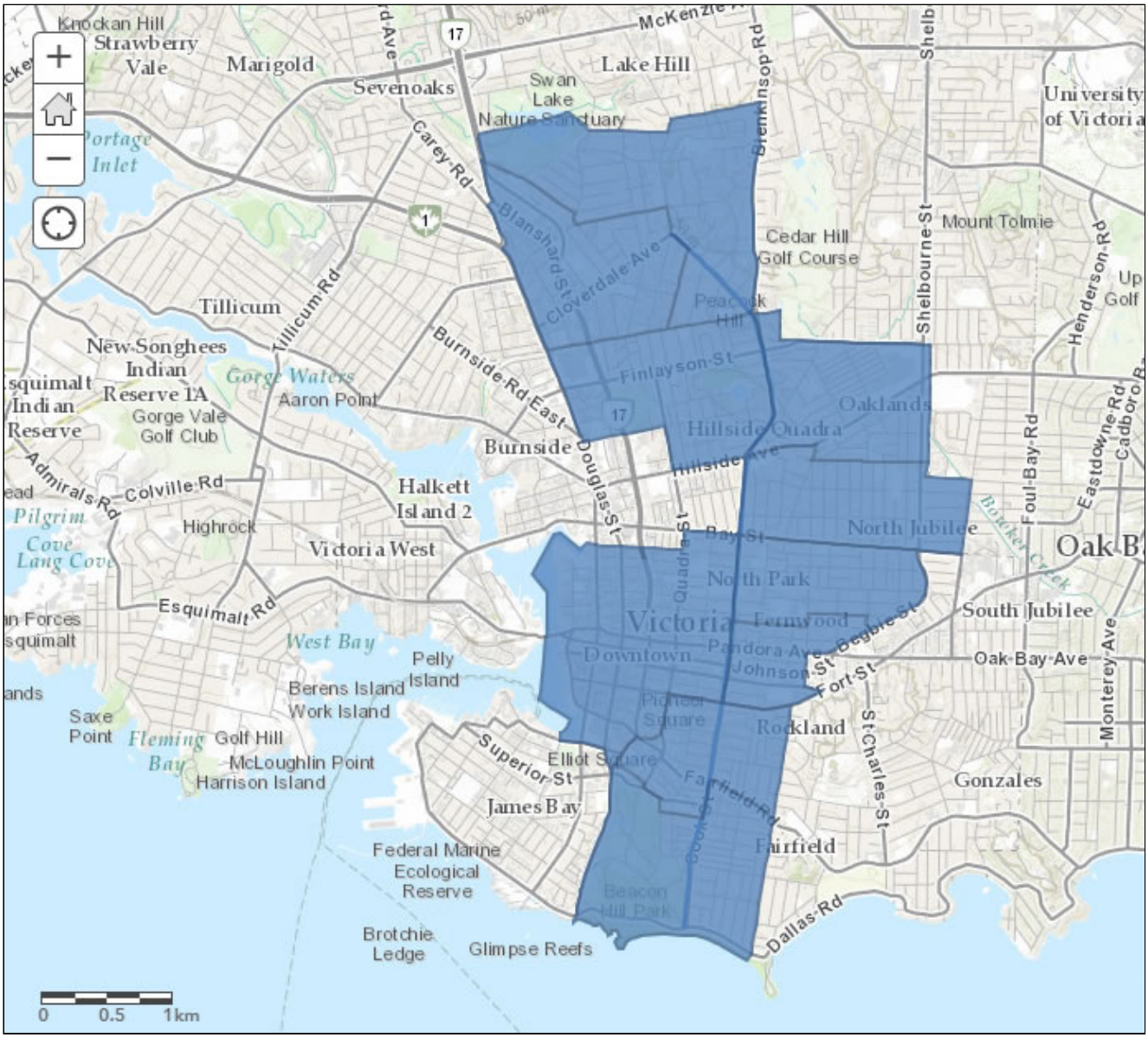


- 1. Choose the **Vic_Census_Join**
- 2. Build a query to find features:



- Press Add
- 3. Result layer name: **Intersect in Vic_Census_shp_jfitterer_join**
- 4. Uncheck "Use current map extent"
- Press **Run Analysis**

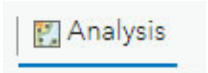
11. View the selected census tracts



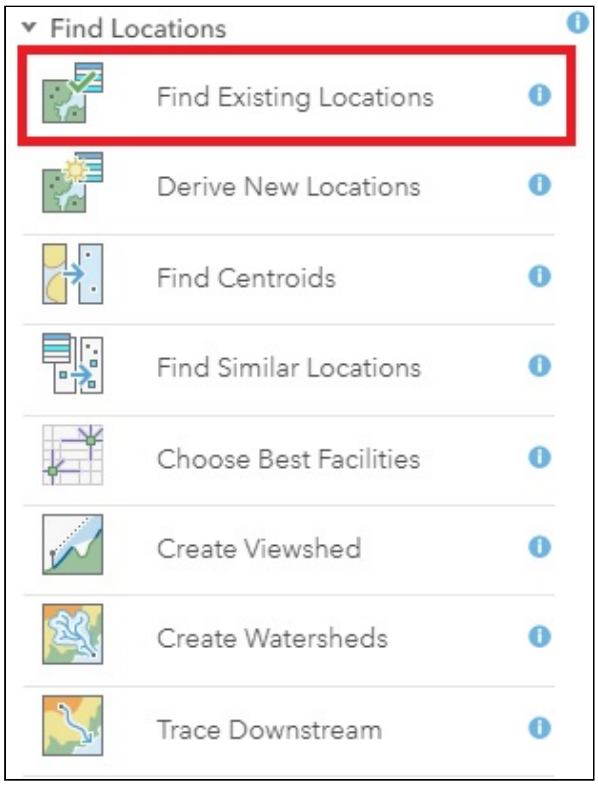
Proximity is defined as "being close or nearby". Proximity is used in GIS to find out what features are within a given distance of other features.

Suppose you wanted to find out how many Census Tracts are within 1 km of Cook Street.

You will be using the same procedures as adjacency, but you will change some of the options in the **Find Existing Locations** window.

1. From the Analysis button 

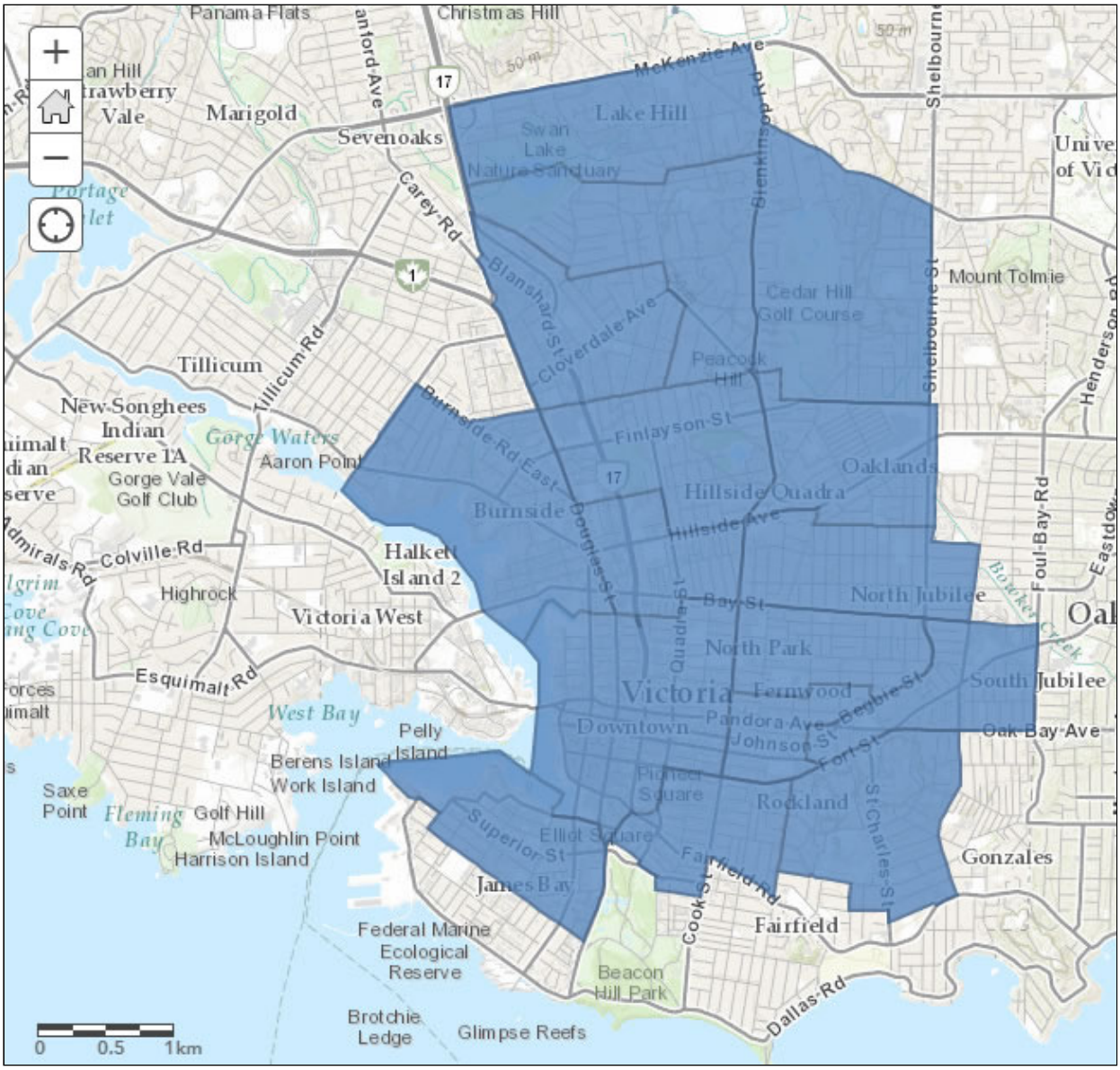
→ choose **Find Locations** → select **Find Existing Locations**



- 1. Vic_Census_Join
- 2. Add Expression

- 3. Result layer name: **1km from cook in Vic_Census_Join**
- 4. Uncheck "Use current map extent"
- 5. Press **Run Analysis**

View the results



Raster Attribute Queries

You can also perform attribute queries on raster data.

In the following examples you will examine the maximum average snow cover of the CRD region in 2006 and 2007. These data were derived from MODIS satellite imagery for ecosystem modelling. You will use QGIS to analyze the data because ArcGIS online does not have raster processing tools.

1. Open QGIS

2. Download the max_snow.zip datasets and a projected version of the Vic_Census file

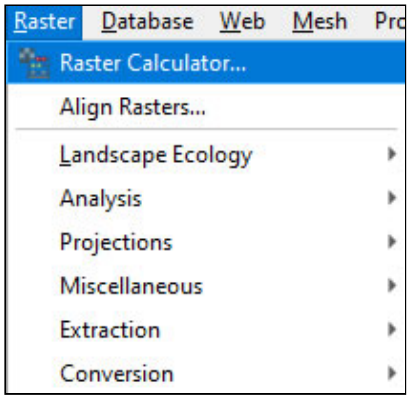
3. Unzip the **max_snow.zip** folder to your **Lab3** folder

4. In QGIS, use the **browse data button**  tab to open the **max_snow_2006** and **max_snow_2007** raster grids **(select the hdr files)**

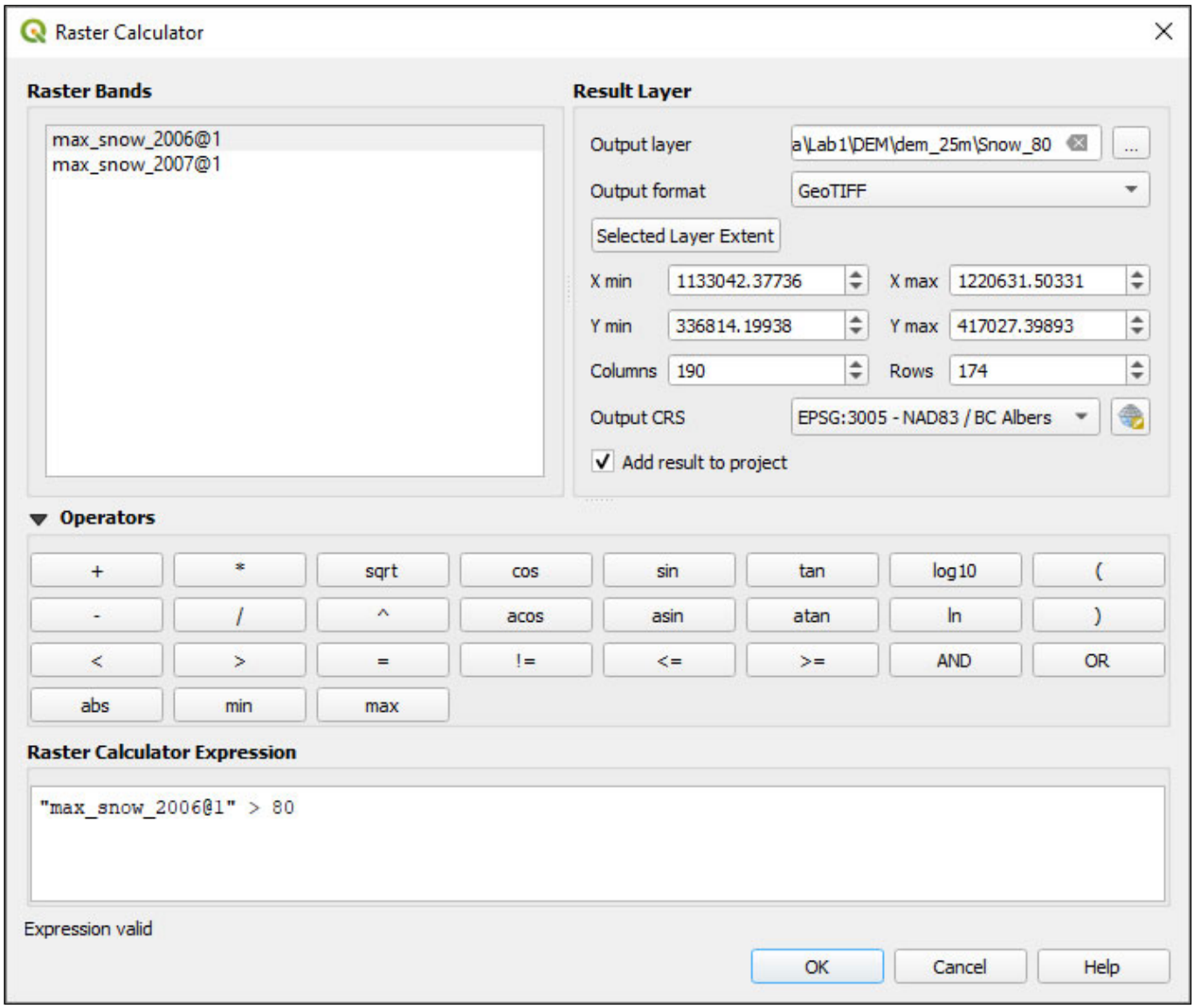
- Note that these data represent percent snow cover per 460m cell.

In the next section, you will use the raster calculator to perform an attribute query on the snow pack information. You are interested in identifying how many cells in 2006 had greater than 80% snow cover.

1. Open the **Raster** menu → **Raster Calculator**



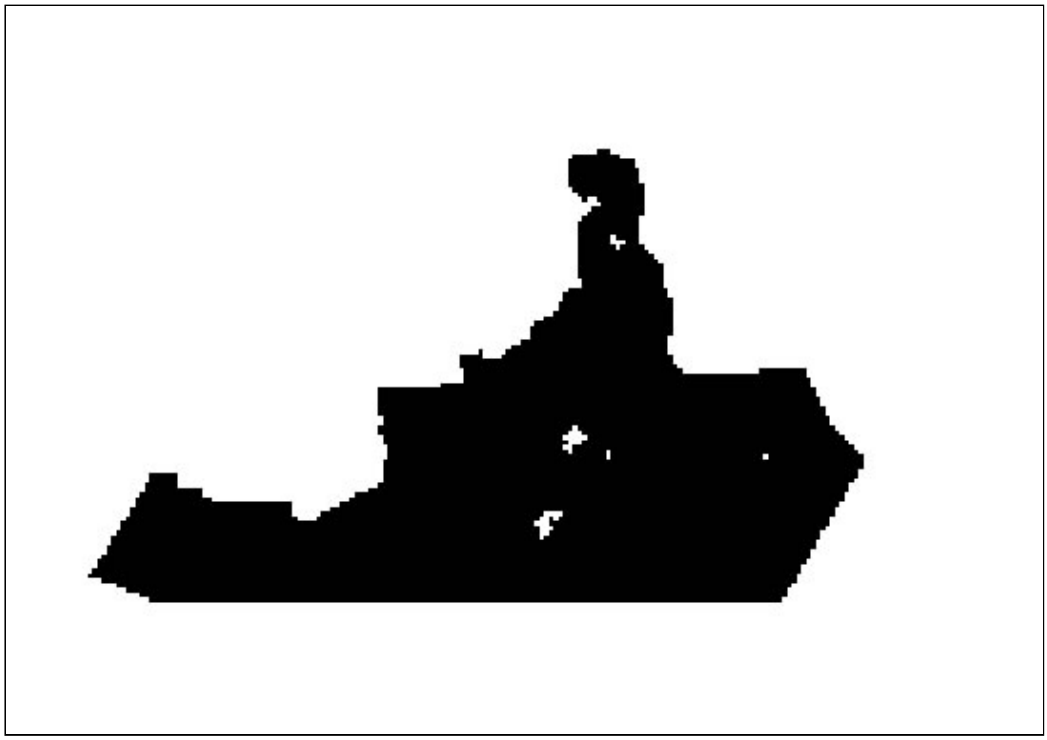
2. Set the Raster Calculator Expression as:



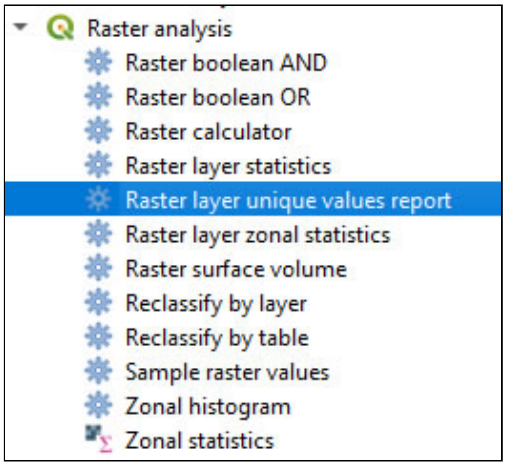
- Note that your used the **>** greater than operator
- Ensure you save the output file as **Snow_80** in your **Lab3** folder.

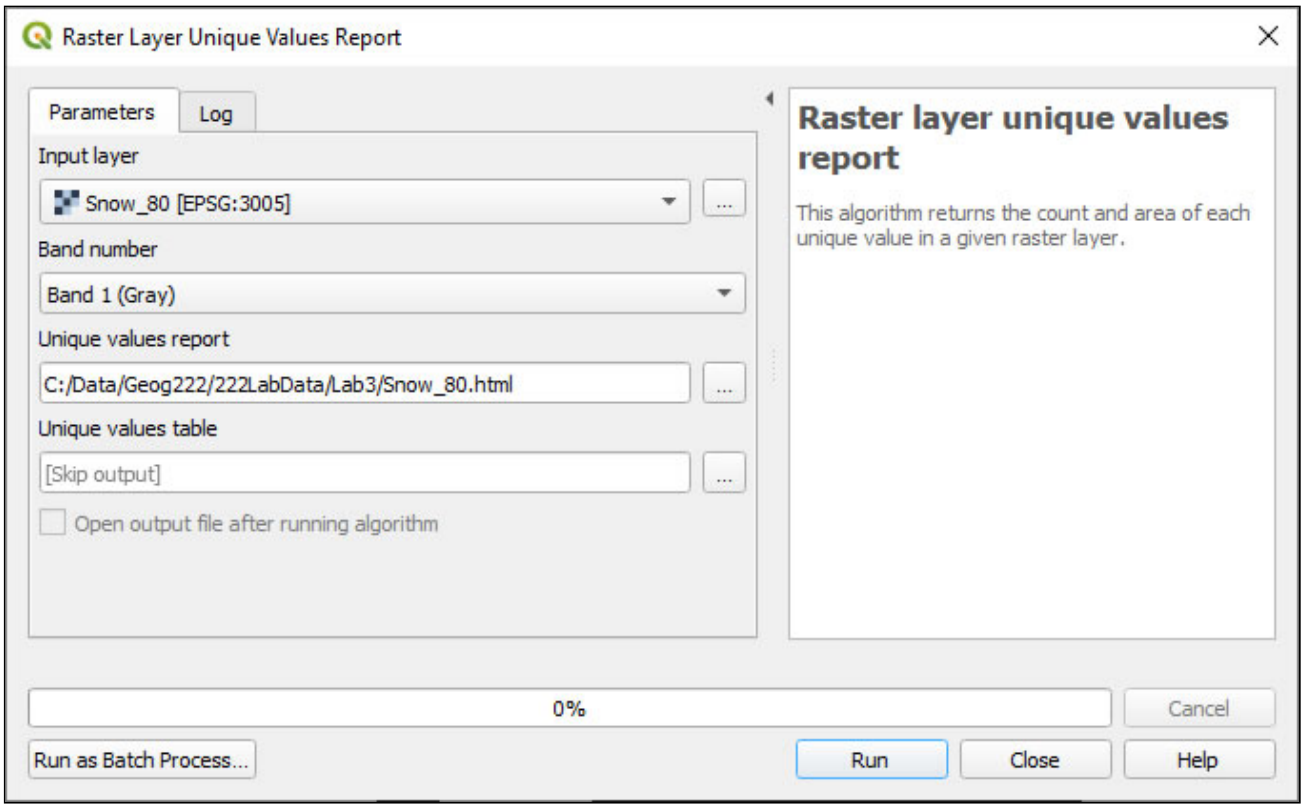
- Press **OK**

The output provides a Boolean raster (raster with values of 0 and 1). All cells with the value of 1 met your 80% snow cover criteria and all the cells with a value of 0 did not meet your criteria.



3. To identify how many cells met your attribute query, run a **Raster layer unique value report**





- Press Run

From the attribute table you can see that 6598 cells did not have a snow cover greater than 80% (VALUE 0). However, 40 cells did have a snow cover greater than 80% (VALUE 1).

Analyzed file: C:\Data\Geog222\222LabData\Lab1\DEM\dem_25m\Snow_80.tif (band 1)

Extent: 1133042.3773600000422448,336814.1993800000054762 : 1220631.5033100000582635,417027.3989300000248477

Projection: EPSG:3005 - NAD83 / BC Albers

Width in pixels: 190 (units per pixel 460.995)

Height in pixels: 174 (units per pixel 460.995)

Total pixel count: 33060

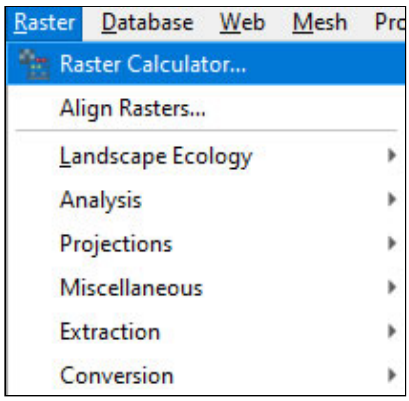
NODATA pixel count: 26422

Value	Pixel count	Area (m²)
0	6598	1402185573.027544
1	40	8500670.342695024

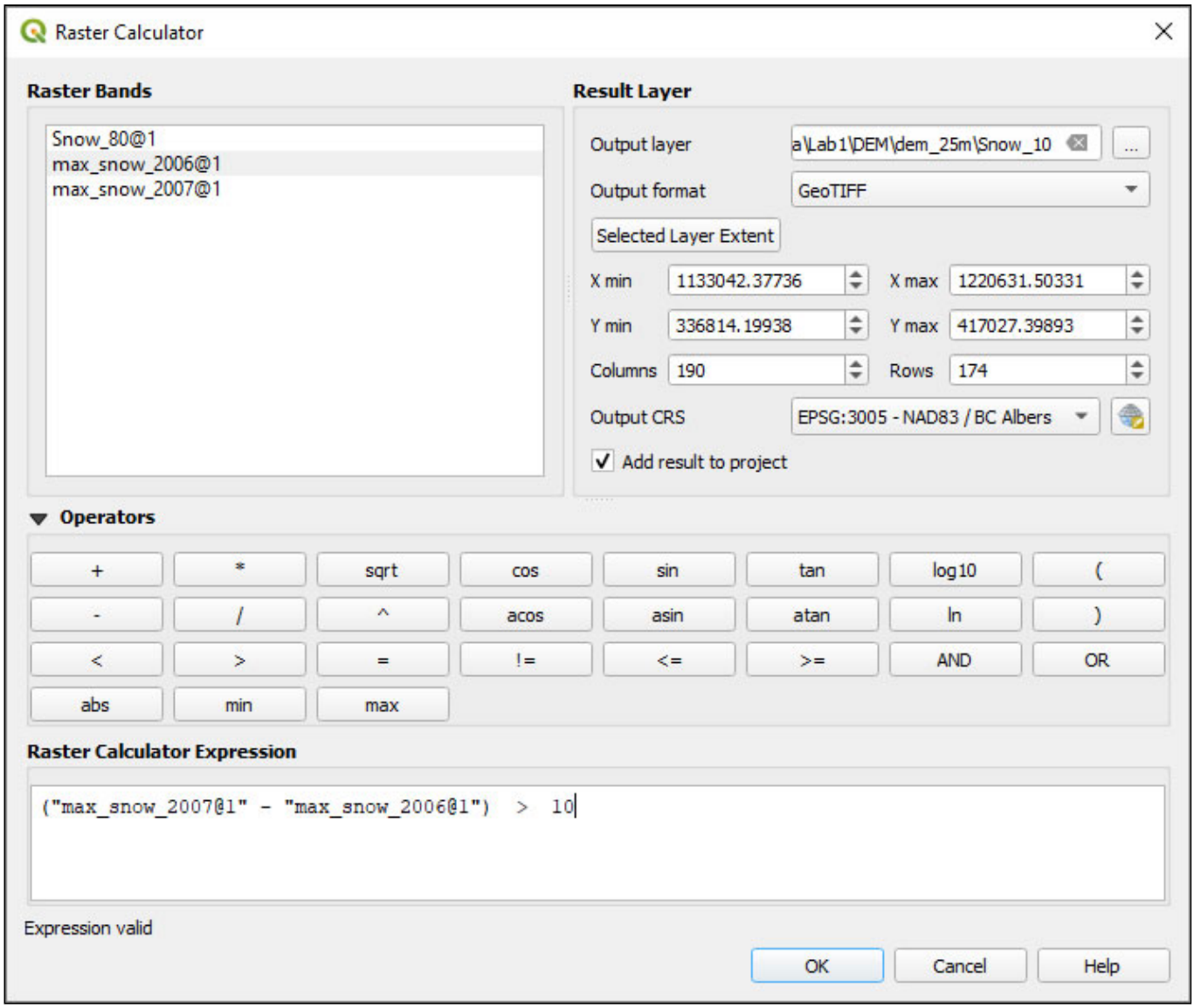
Raster Math

Consider now that you would like to identify where the change in snow cover between 2006 to 2007 exceeded 10%. That would mean that on a cell by cell basis you would like to subtract the 2006 snow cover values from the 2007 snow cover values. To do this, you will use the raster calculator.

1. Open the **Raster** menu → **Raster Calculator**

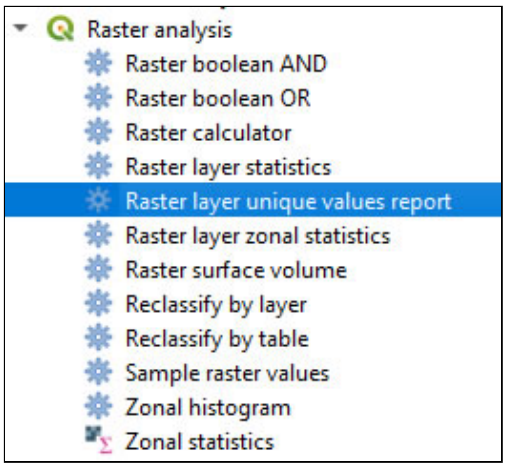


2. Set the **Raster Calculator** Expression as:

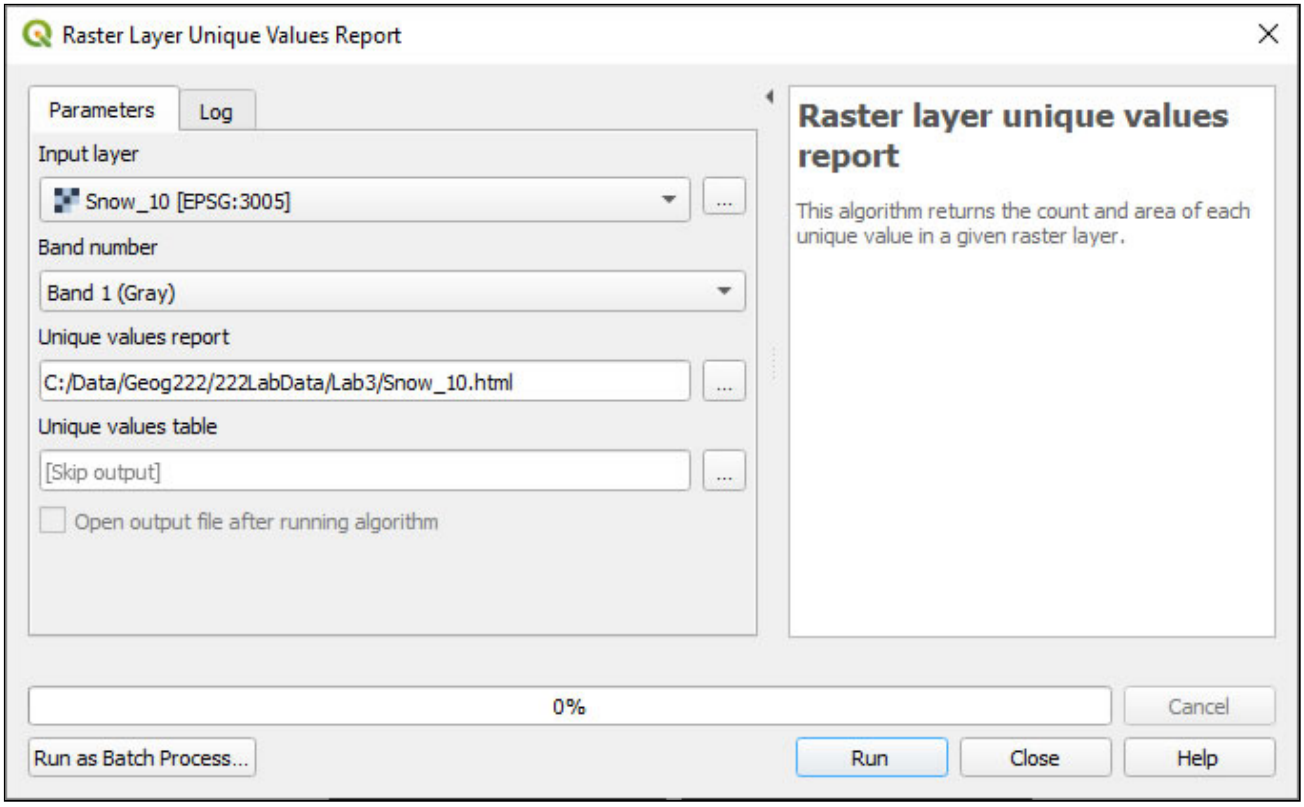


- Ensure you save the output file as **Snow_10** in your **Lab3** folder
- Press Run

3. To identify how many cells met your attribute query, run a **Raster layer unique value report**



- Set the parameters as follows:



- Press **Run**

View the report:

Analyzed file: C:\Data\Geog222\222LabData\Lab1\DEM\dem_25m\Snow_10.tif (band 1)

Extent: 1133042.3773600000422448,336814.1993800000054762 : 1220631.5033100000582635,417027.3989300000248477

Projection: EPSG:3005 - NAD83 / BC Albers

Width in pixels: 190 (units per pixel 460.995)

Height in pixels: 174 (units per pixel 460.995)

Total pixel count: 33060

NODATA pixel count: 26422

Value	Pixel count	Area (m²)
0	5826	1238122635.41353
1	812	172563607.956709

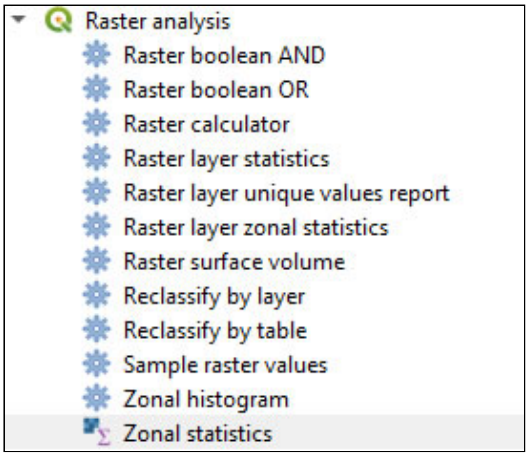
The results show that 812 cells had a change in snow cover that exceeded 10% of the 460m pixel's coverage between 2006 and 2007.

Zonal Statistics

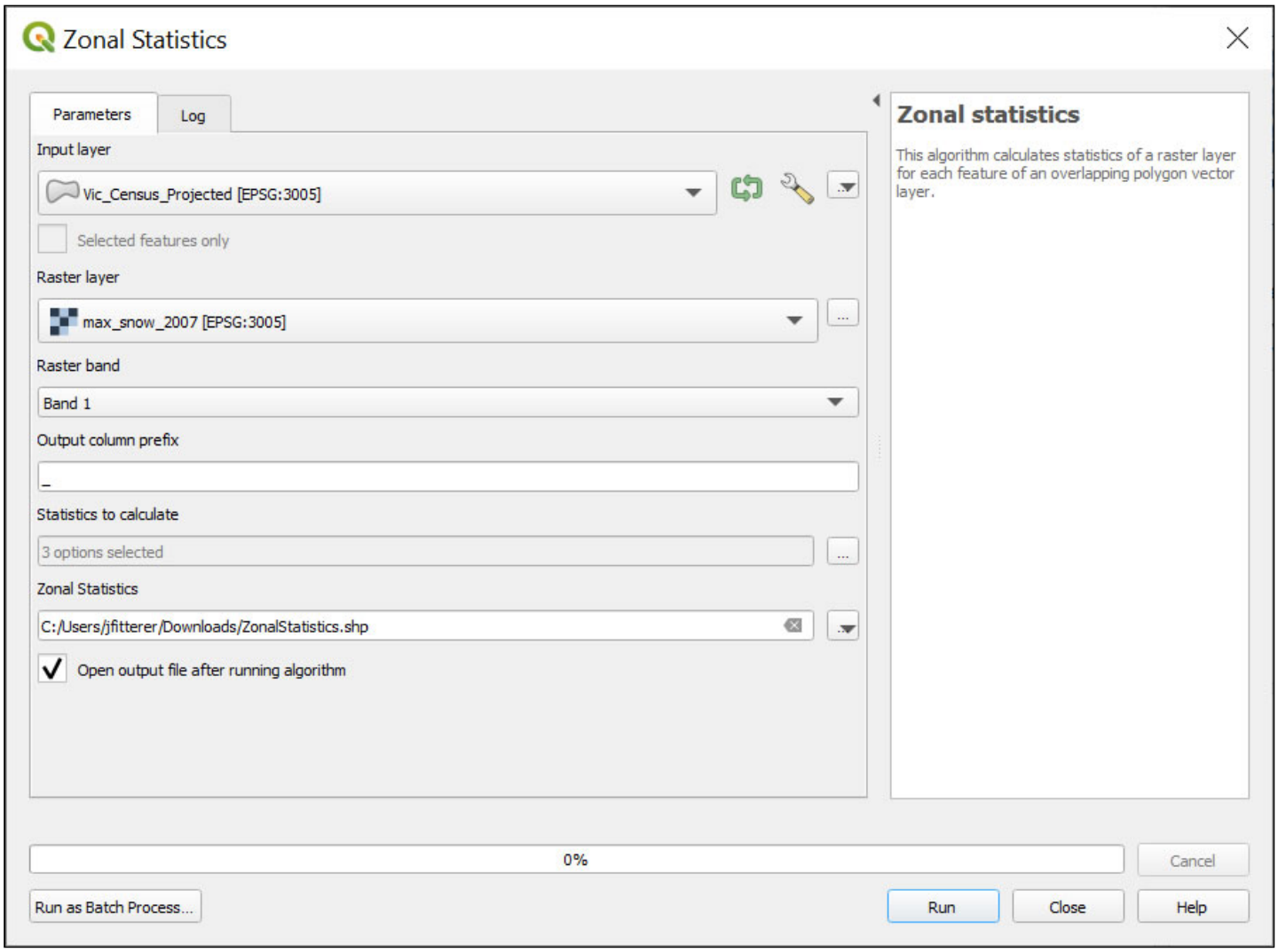
Zonal Statistics are a way to aggregate raster grid data into polygon units. In the following example you will calculate the mean 2007 snow cover per census tract unit using zonal statistics. The census tract units will form the zones, and the snow data provides the dataset to be summarized by the spatial units.

1. Open the **Vic_Census_Projected.shp** file **(You will have to unzip the folder before you will be able to add the shapefile)**

2. Open the **Raster analysis** tab → search for the **Zonal Statistics** tool

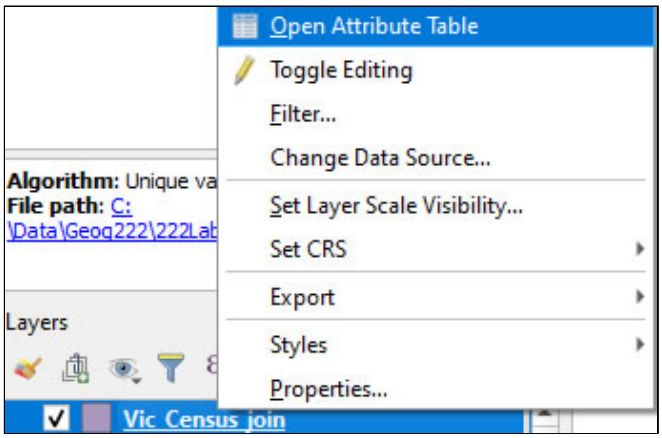


2. Set the **Zonal statistics** parameters as follows:



- Press **Run**

Open the **ZonalStatistics** attribute table



Scroll to the zonal statistics columns, and observe how the 2007 snow cover data has been summarized into the Census Tract units.

_count	_sum	_mean
4	24	6
7	13	1.857142857142...
14	65	4.642857142857...
12	45	3.75

Assignment