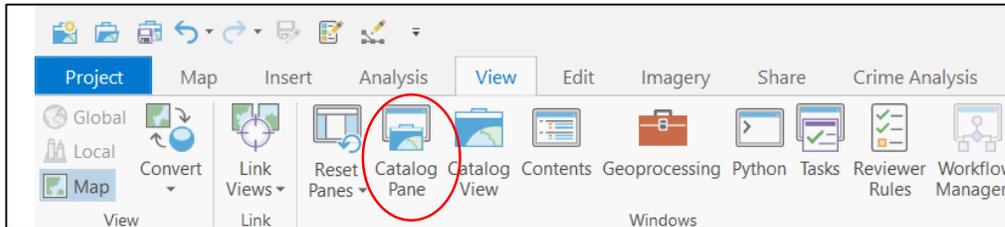


1. Cluster and Outlier Analysis

In this exercise we are going to use a 28 000+ residence units' sales information from 2010-2011 from Helsinki Metropolitan area and examine the listed square meter prices for clustering and outlier anomalies.

Open ArcGIS Pro and create a new project.

If the Catalog pane is not available on the right-side of User-Interface, you can open it from the View tab and then selecting 'Catalog Pane'.



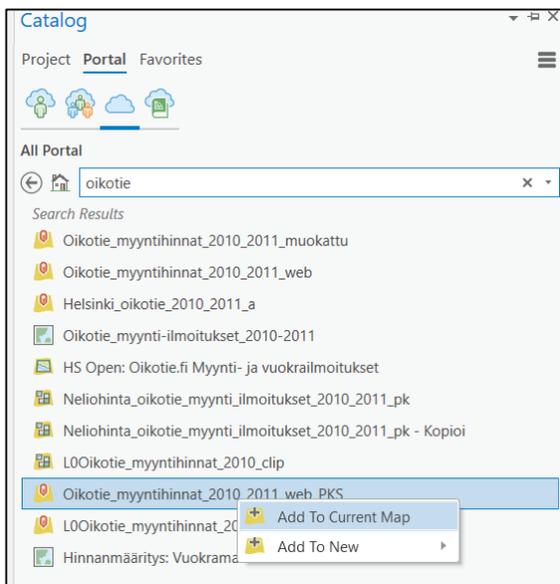
From Catalog Pane, switch the view to 'Portal' and click the cloud icon to set the search against the active and all public datasets. If you are using an ArcGIS Online organization account, then you're searching against all global public datasets. If you are using an ArcGIS Enterprise account, the search works against your organization's private datasets.

Here we assume that all students have an ArcGIS Online account.

Type 'oikotie' as a search query.

You will get a list of search results with different types of item services stored to ArcGIS Online.

Look for an item 'Oikotie_myyntihinnat_2010_2011_web_PKS' and right click on it and choose 'Add To Current Map'. This imports the dataset to Pro's map.

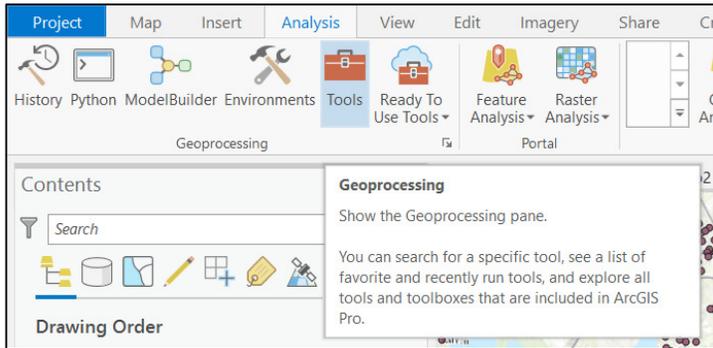


If you want, you can right click the layer in Pro's Table Of Contents (in Finnish: Sisällysluettelo) and choose 'Attribute Table' from the pop-up list to view the attribute information of features.

A quick way to check data field types is to right click the layer in 'Contents' and then choose 'Design' -> 'Fields'.

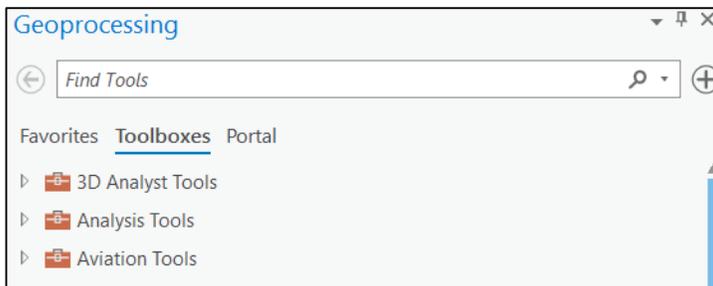
The Cluster and Outlier Analysis tool

The geoprocessing tools used in ArcGIS Pro are opened via the Geoprocessing pane. If the pane is not open already at the right-hand side of the UI, it can be opened by going to the 'Analysis' tab at the top of Pro's UI and selecting 'Tools'.



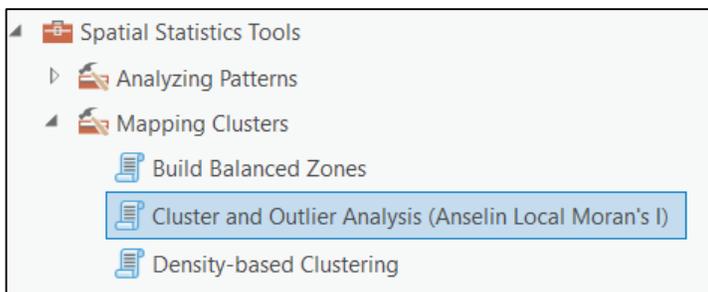
The geoprocessing pane holds three different sections: Favorites, Toolboxes and Portal.

- Favorites holds a list of often used tools
- Toolboxes holds all the ArcGIS geoprocessing frame work's tools, listed in toolboxes by categories
- Portal holds portal-based tools, which are services from either ArcGIS Online or ArcGIS Enterprise



In this exercise we want to use the 'Cluster and Outlier Analysis (Anselin Local Moran's I)' tool. The tool can be found under the Spatial Statistics toolbox and Mapping Clusters tools.

Another way to search for the tool is to type the name to the Find Tools section of the Geoprocessing Pane. Accustomed ArcGIS Pro users usually use this way of finding the tool they want.



Open the *Cluster and Outlier Analysis (Anselin Local Moran's I)* tool with left click. Fill in the tool dialogue as in picture below. Notice that when using geoprocessing tools, there are two ways of determining input datasets: the drop-down list and the folder-based input.

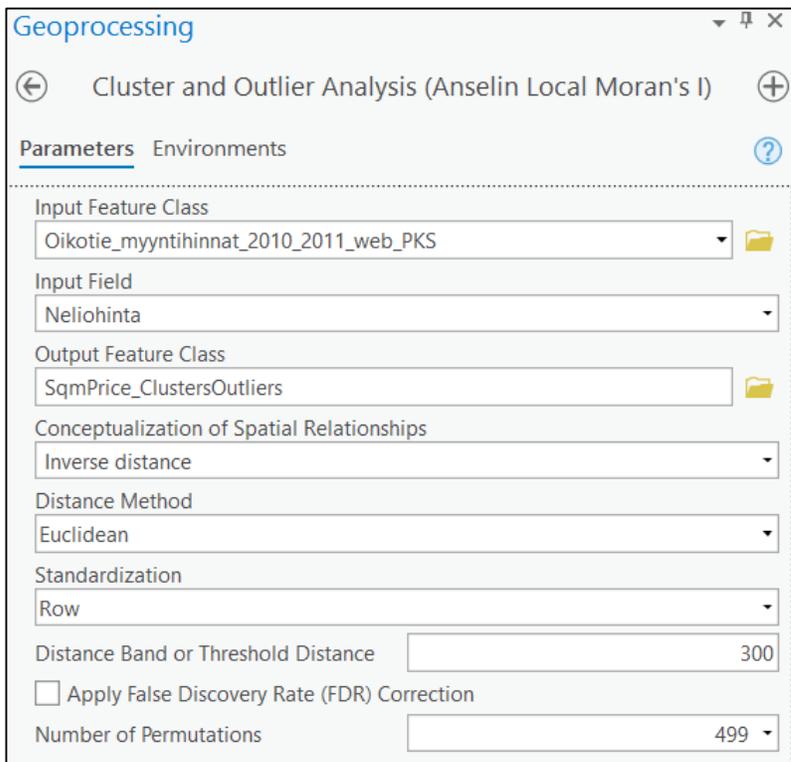
The drop-down list shows only suitable layers from the ArcGIS Pro's map you are working on. These are called in Esri language as *feature layers*. From the folder-based approach, you can determine a specific dataset from folder or geodatabase that is not necessarily in your ArcGIS Pro map. These types of inputs have different names in Esri language, but whenever the you are using a dataset from the geodatabase, it is called a *feature class*. Notice, that a dataset can be *both* if it is stored in a geodatabase and it is imported to Pro's active map.

If you click on the 'Output Feature Class' parameter, you will get to see the exact saving location of the output dataset. By default, the ArcGIS Pro uses the storing location (a geodatabase) that has been given in the Environment setting (Analysis tab).

HINT: Sometimes when we do not want to store outputs of i.e. geoprocessing tests to drive, we can determine to write the results to system memory. This can be done by replacing the folder and geodatabase part of the syntax with '*memory*'.

For example: *memory\my_output*

Where *memory* part indicates that we want to write results to memory and the *my_output* part determines the layer name. Please notice that using this in-memory writing works mainly on vector-based geoprocessing tools and not all tools might support it. Although unsupported vector tools are likely not too many. Also, some tools might require that the input is stored to drive, so memory-based inputs might not work in all cases as an input for geoprocessing tools.

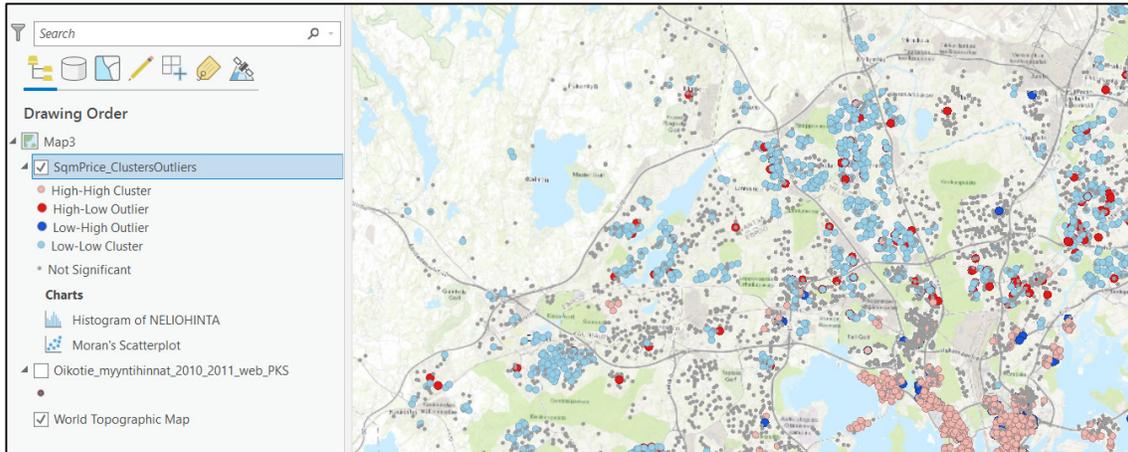


When you have the tool dialogue filled up correctly, hit the 'Run' button to begin tool processing.

Once the tool is finished, the output layer pop-ups in the Table Of Contents as a layer. Now, you can deactivate the viewing of the original sales information from the button mark, so it will not confuse the map interpretation.

Pan around the results and check especially for the outlier cases (High-Low and Low-High cases).

Also notice, the tool used outputted two charts. You can open them by double clicking them with left mouse button.



2. Hot Spot Analysis

Another tool we can use to detect clusters of high and low values is the 'Optimized Hot Spot Analysis'. This tool determines automatically the optimal clusters using Getis-Ord G_i^* statistic. The tool automatically aggregates incident data, identifies an appropriate scale of analysis, and corrects for both multiple testing and spatial dependence.

From the Geoprocessing Pane, search for 'Optimized Hot Spot' and you should receive a list of tools, where the first from top is 'Optimized Hot Spot Analysis' tool. Click on it to open tool dialogue.

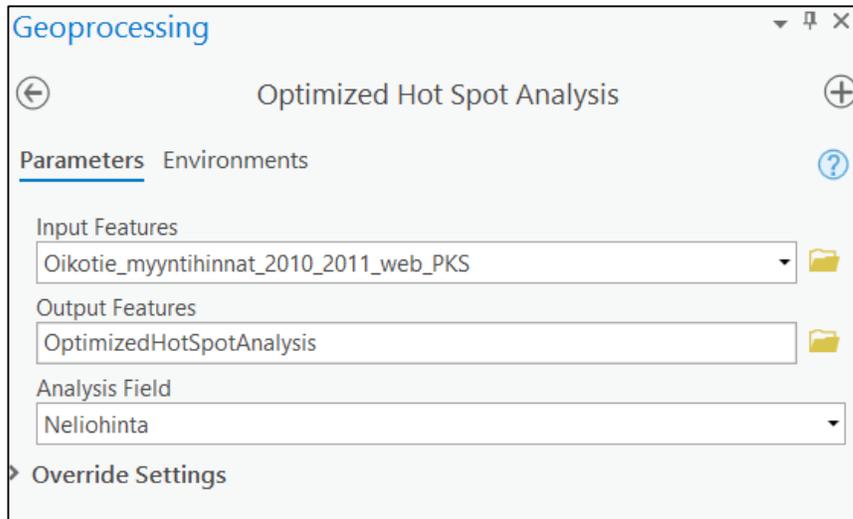
As for the Input Features, enter the 'Oikotie_myyntihinnat_2010_2011_web_PKS' layer as an input (from the previous exercise).



For 'Output Features', type: *memory\OptimizedHotSpotAnalysis*

For 'Analysis Field', select: *Neliohint*

Click 'Run' to execute the tool.



Notice: Had we not chosen any analysis field, then the analysis would be based on aggregating the features to grids or hexagons. The output would model geometrical density hot spots.

End of exercise.