

Getting Started with Mapping and Geographic Analysis in KNIME

Lingbo Liu, lingboliu@fas.harvard.edu

Tiny URL for Materials: <https://tinyurl.com/KNIMEHubGIS2026>

March 11 2026

About Me



Research Associate
Project Lead of Spatial Data Lab
Center for Geographic Analysis

Harvard University

Lecturer
Department of Government

Geospatial AI | Public Health | Urban Analytics | Visual programming

For replicable Spatial Data Science



Explainable
Geospatial Machine
Learning Model

XGeoML 0.6



Reconciled Two-Step
floating Catchment

R2SFCA1.1.3



Geospatial Analytics
Extension for KNIME

GAEK 2.0



KNIME
Google Earth Engine
Extension

KGEE 0.9- prerelease

[Google Scholar](#)

Goal

In this workshop, we will learn how to use KNIME to build a complete GIS workflow—from data input to spatial analysis and visualization.

You will be able to:

Create and manage a **KNIME workflow** (nodes, connections, branching, components)

Read and write common geospatial formats (e.g., Shapefile, GeoPackage, GeoJSON, GeoParquet)

Perform essential **geospatial preprocessing**, such as projection/CRS transformation and geometry conversion

Conduct core **spatial operations**, including buffering, overlay (clip/intersect), and spatial join

Apply **table operations** for spatial analytics (filtering, joining, aggregation, missing-value handling)

Build interactive **geospatial visualizations** and adjust map styling (color, size, classification, legends)

Package workflows into reusable mini-tools using **widgets and components** (optional)

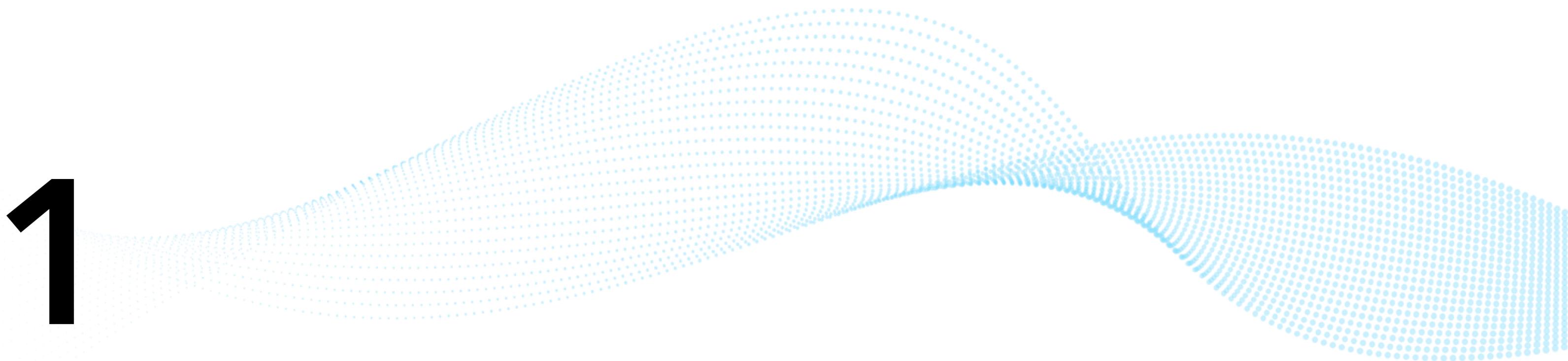
Lecture 1

Visual Programming for GIS

KNIME Analytics Platform



Center for Geographic Analysis
Institute for Quantitative Social Science
Harvard University



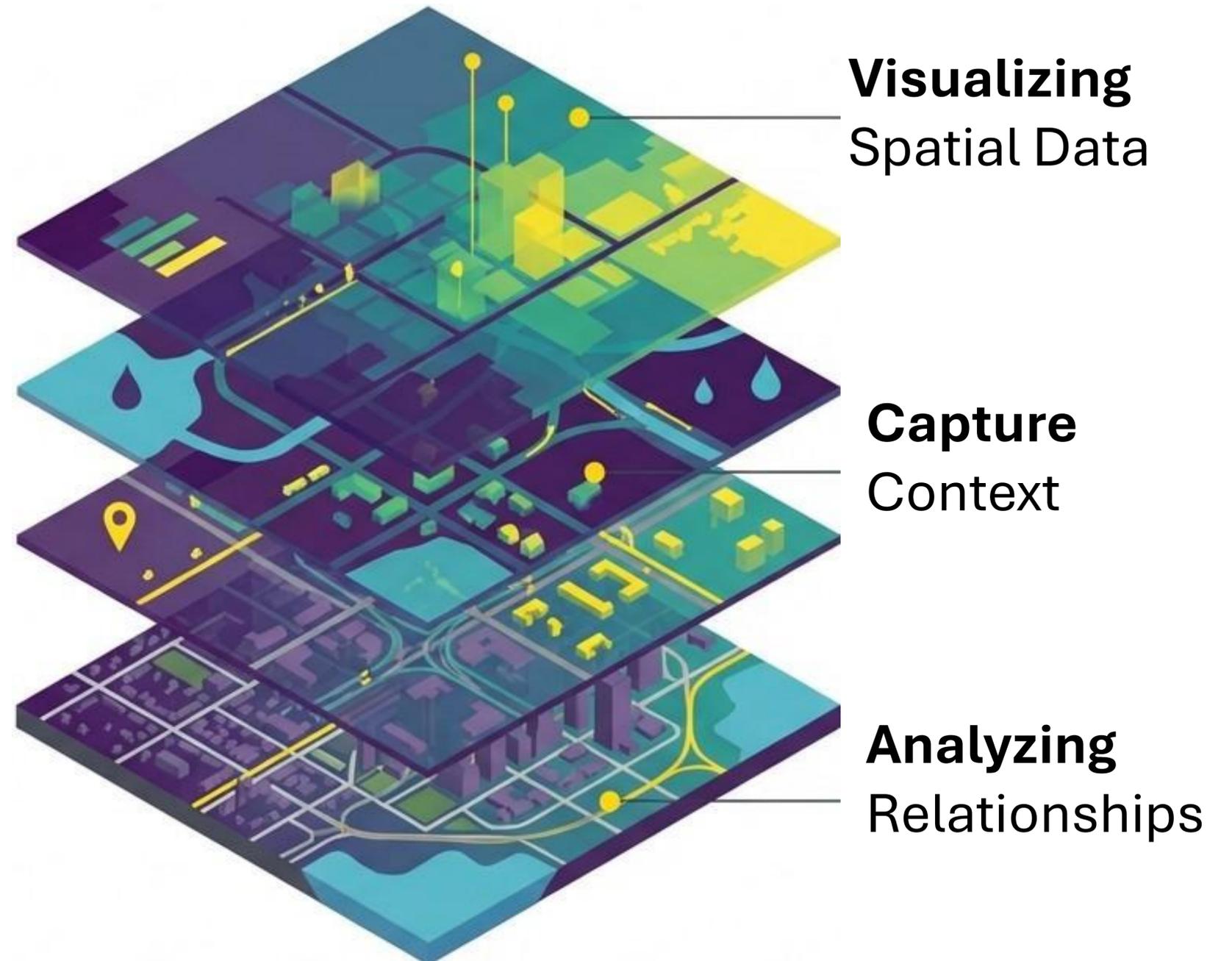
1

Where GIS meet KNIME

The Power of Location

Geographic Information Systems (GIS)

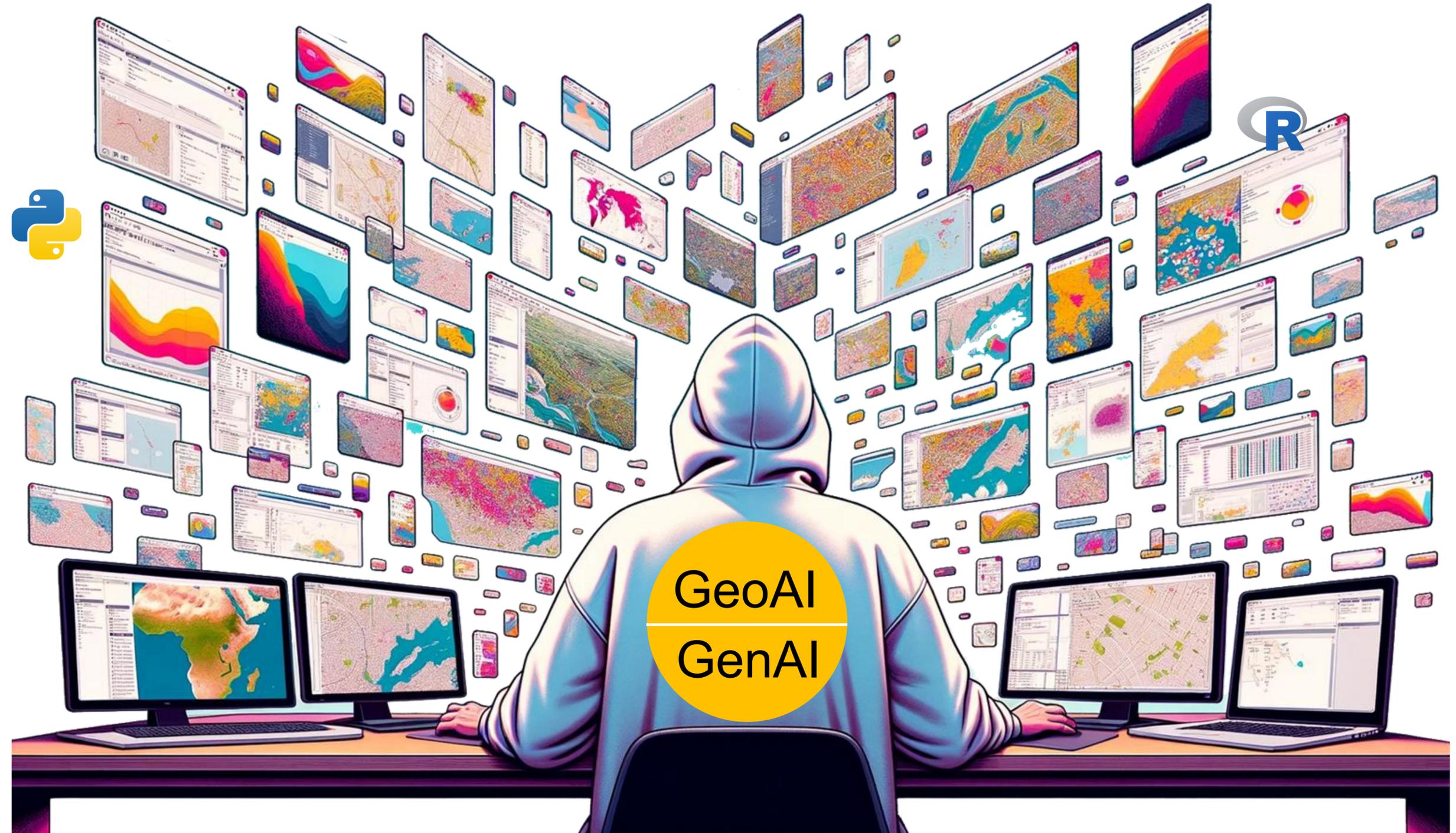
A framework to organize knowledge, revealing patterns, relationships, and processes.



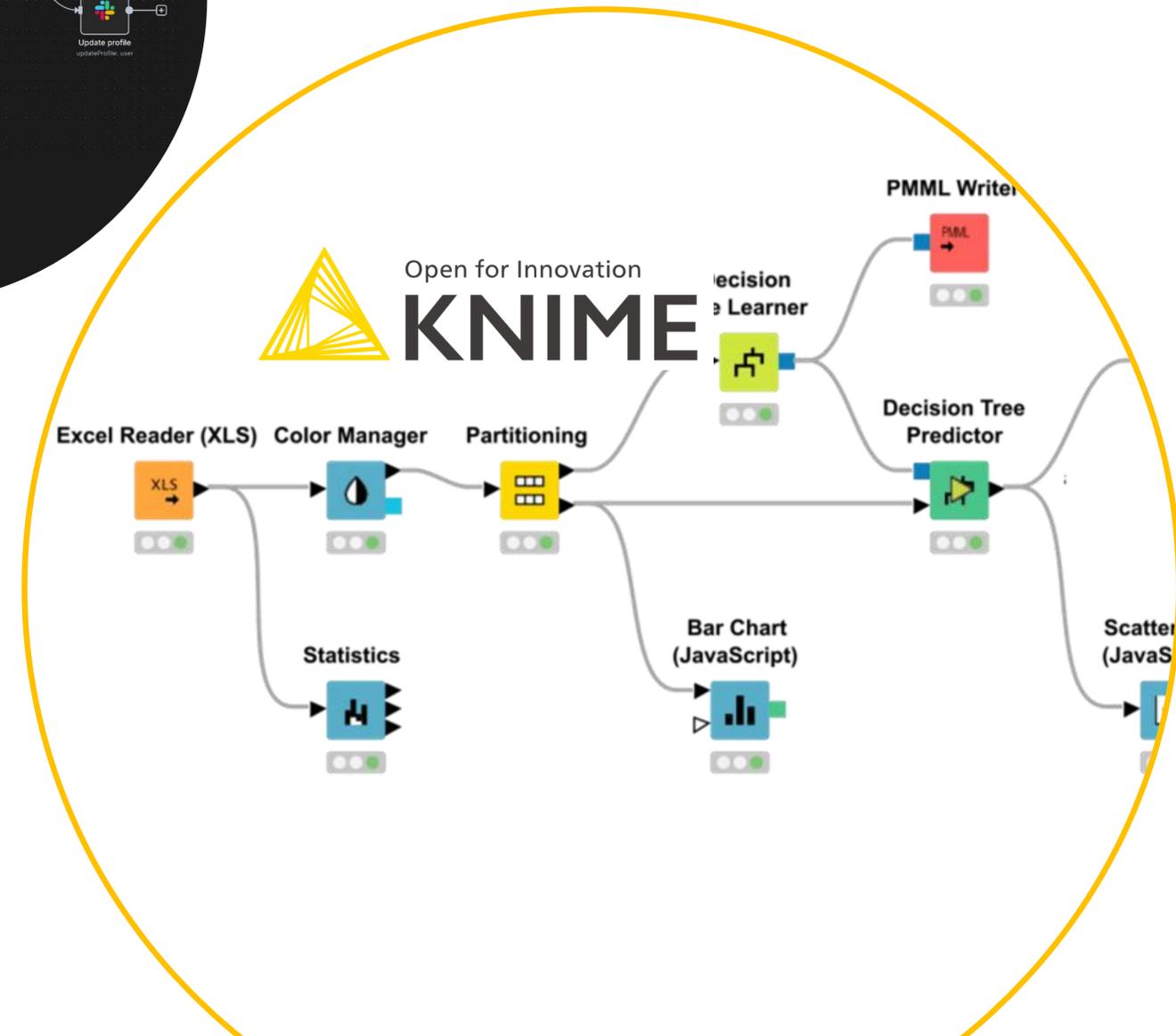
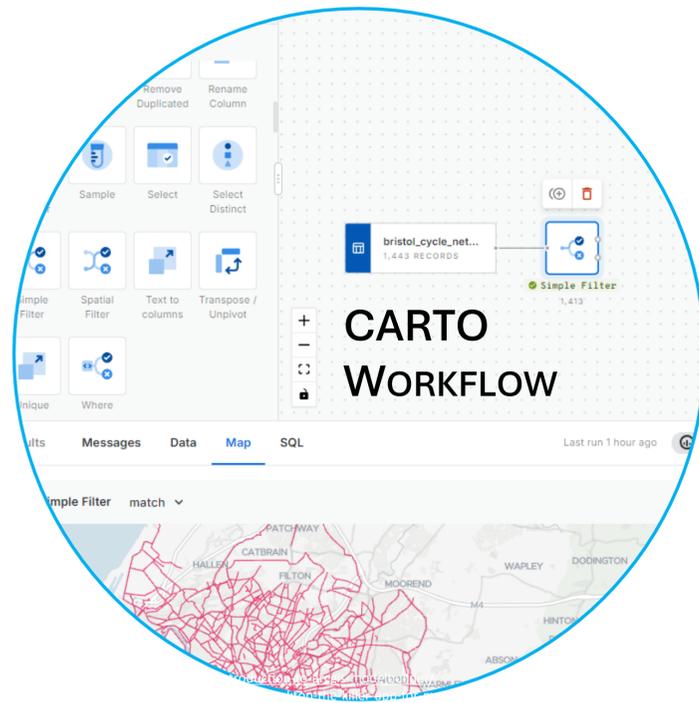
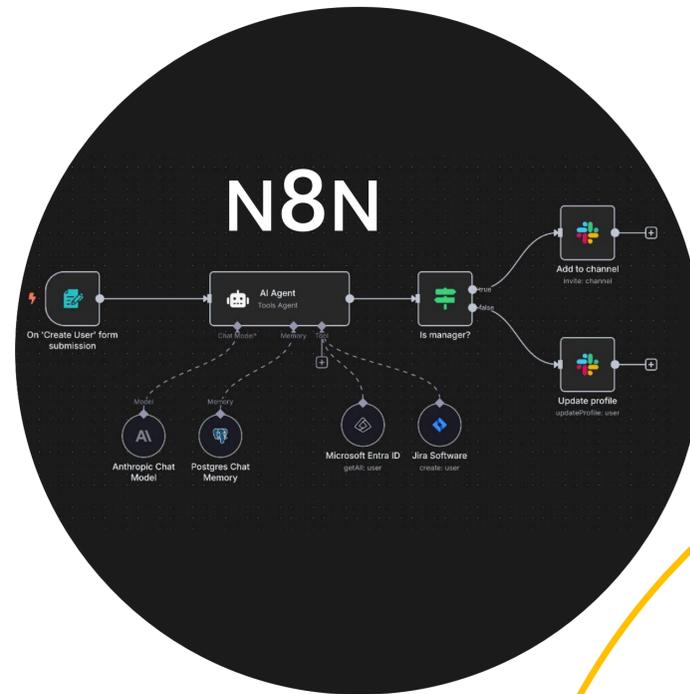
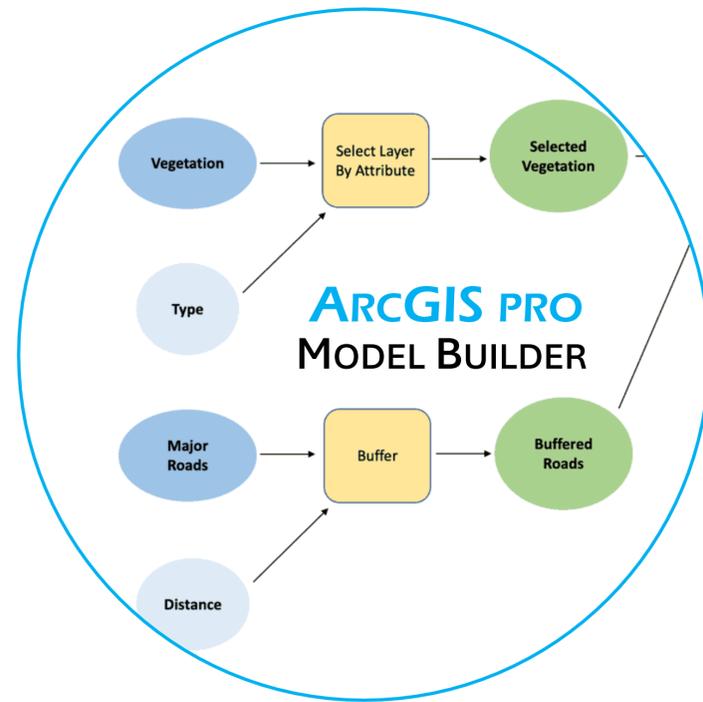
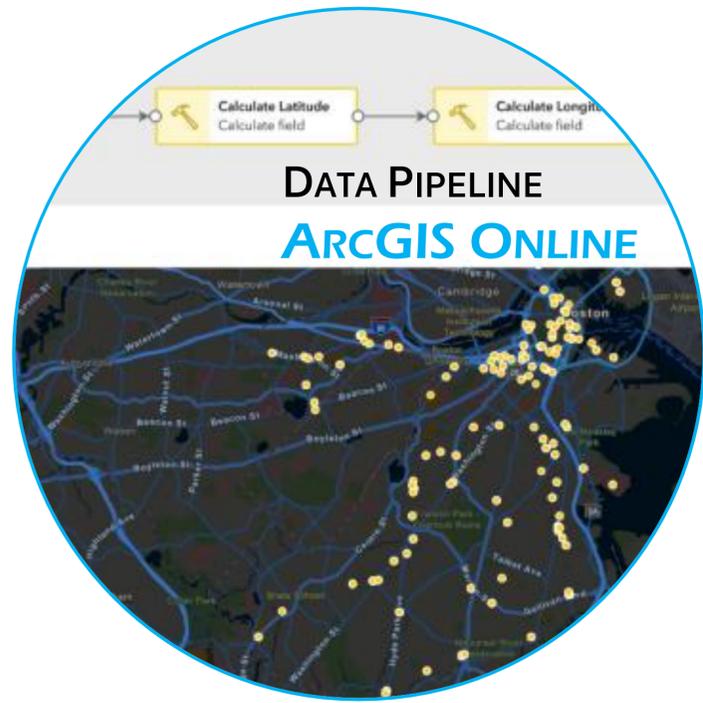
Location Challenges



Surging Learning Curve in GIS Education

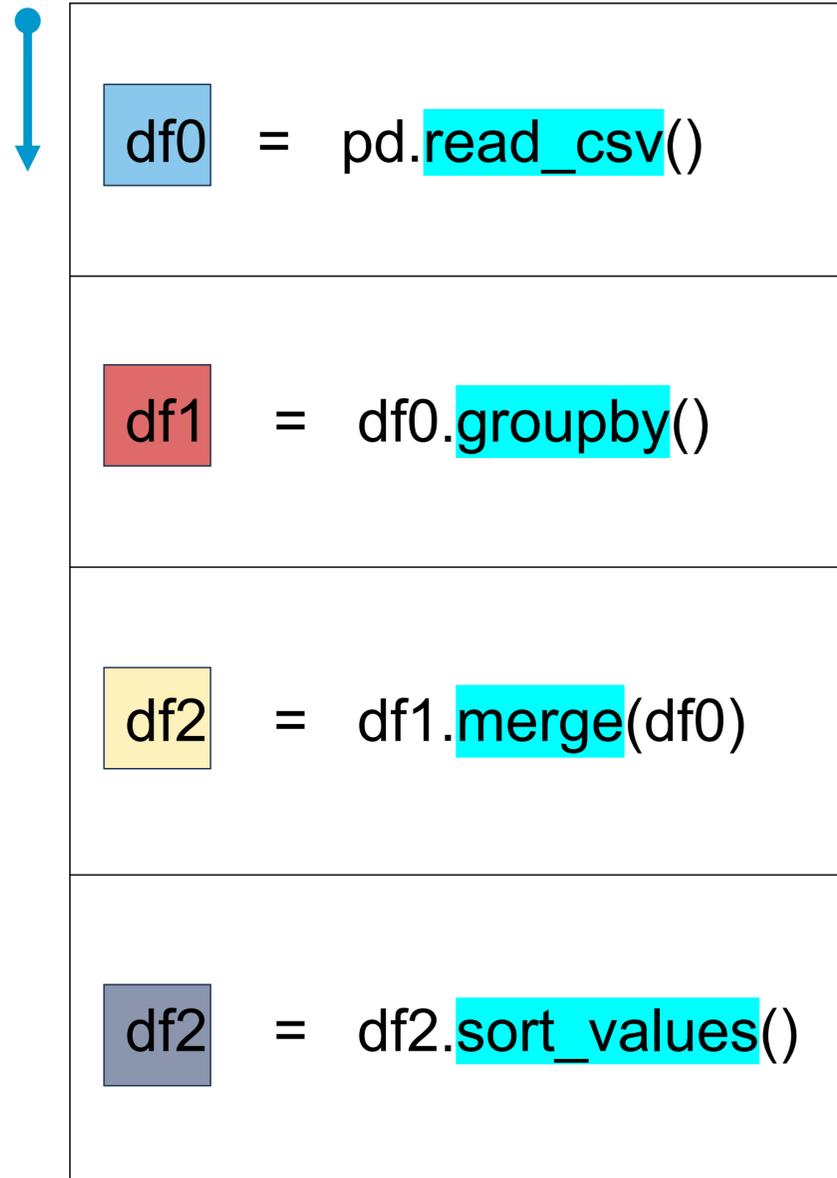


Visual Programming as a solution to GIS Learning



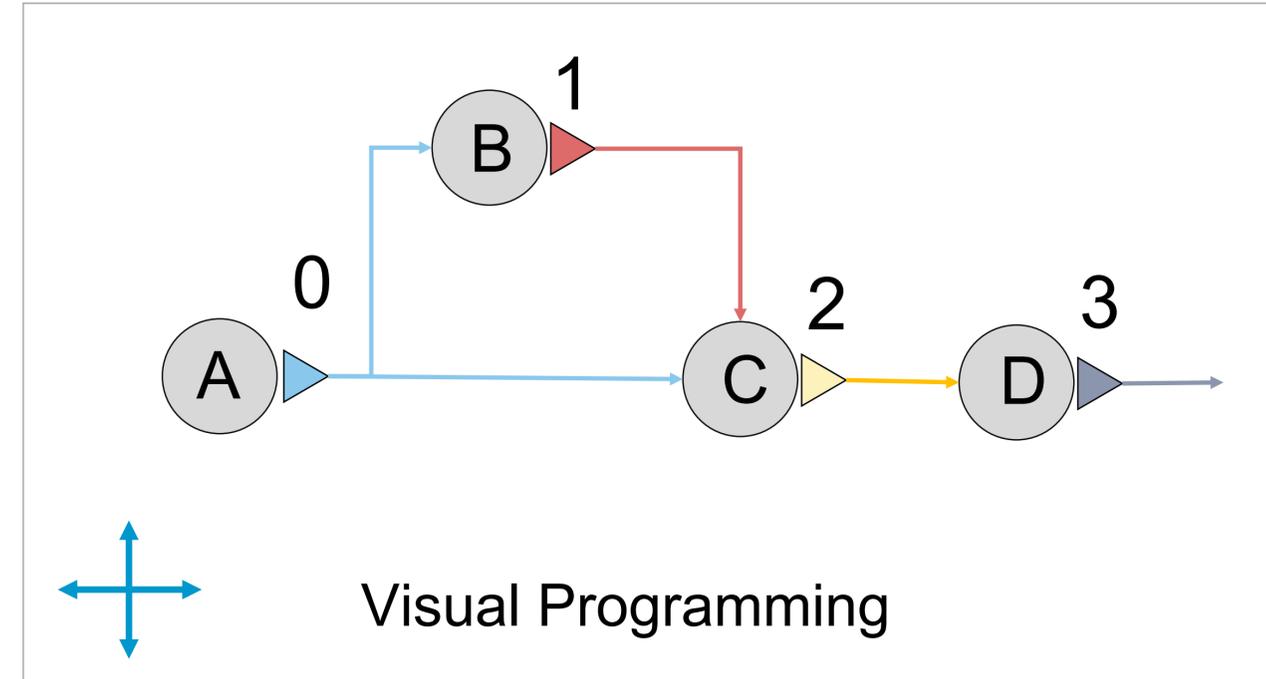
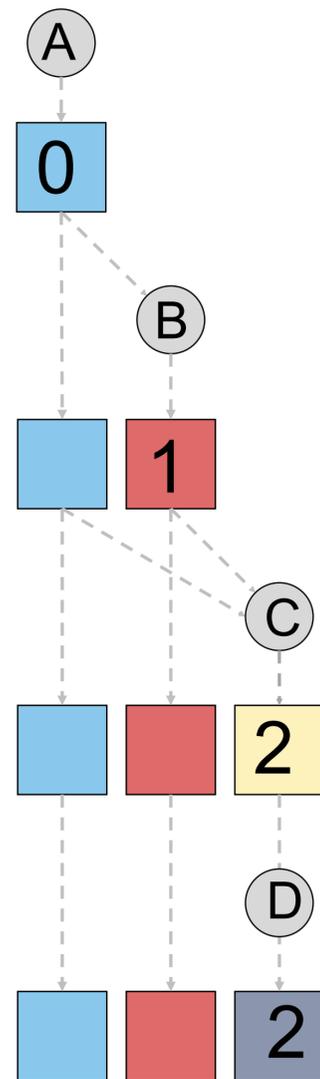
Analyst as Mind mapping : From 1D to 2D

Example



Jupyter Notebook

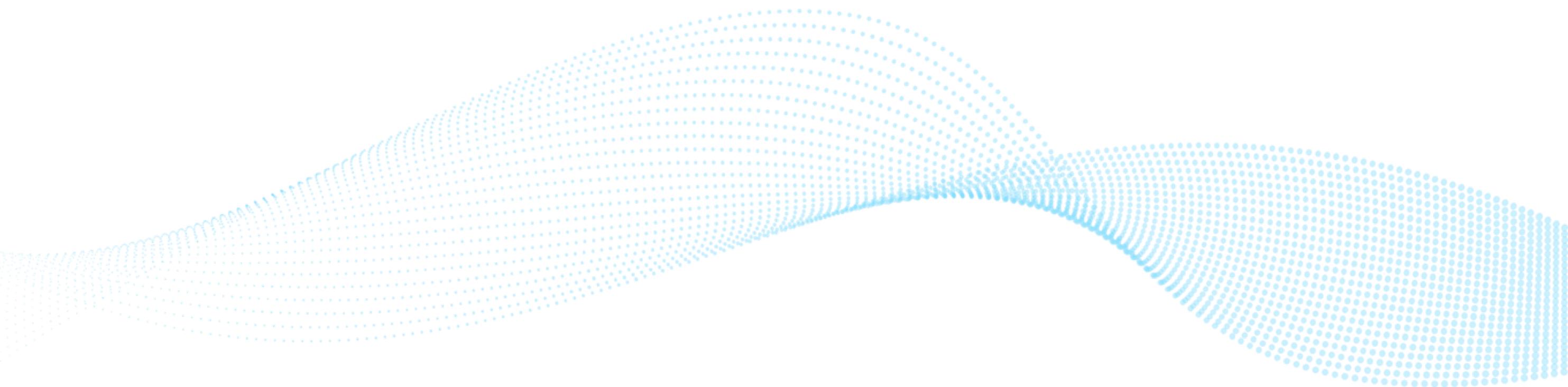
Structure



Legend

- Function
- ▶ Function Node
- Data
- Function
- - - Hidden link
- ▶ Data port
- Code line
- Data link

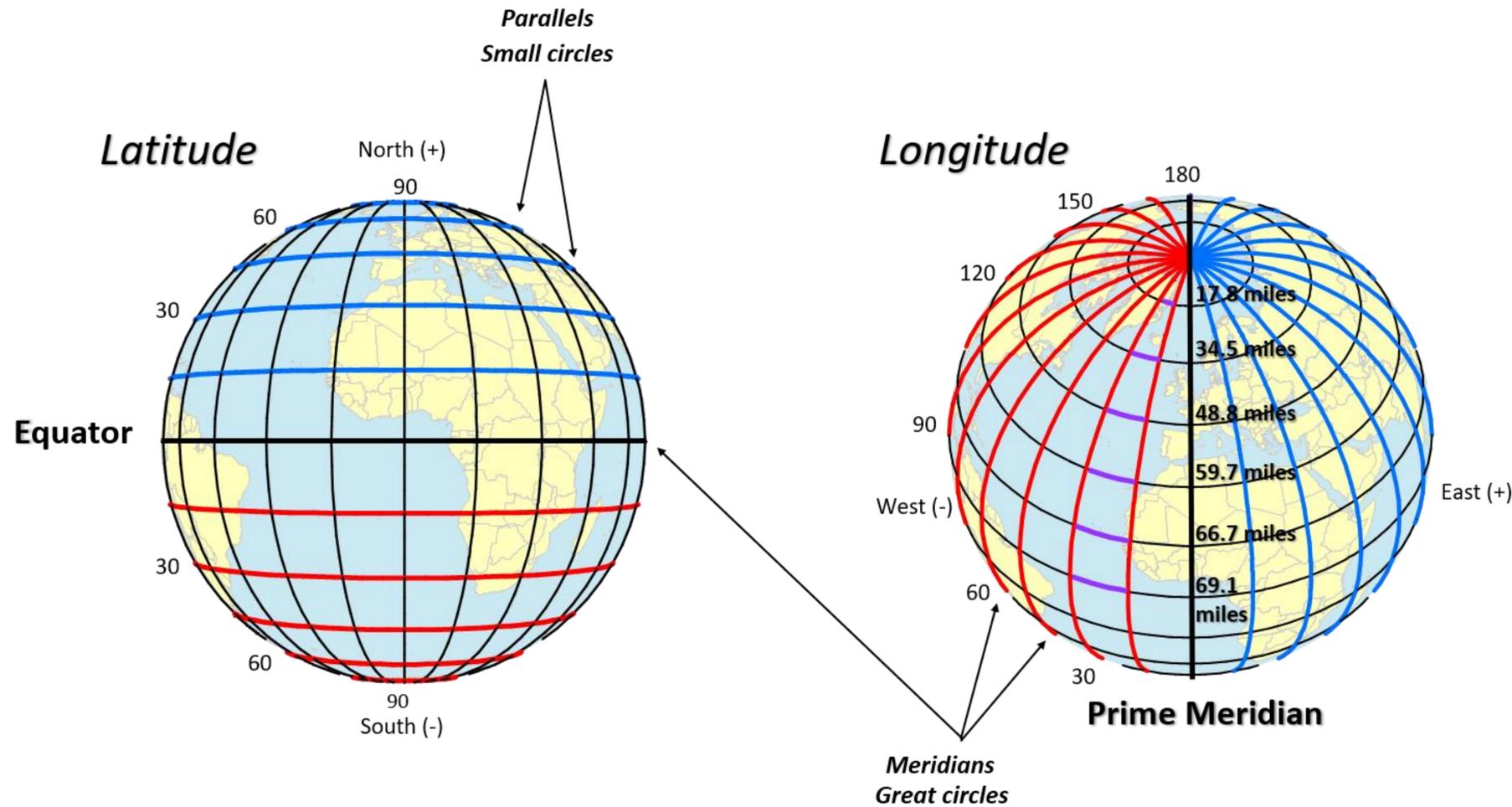
2

A decorative graphic consisting of a series of light blue dots arranged in a wavy, horizontal pattern that spans across the middle of the slide.

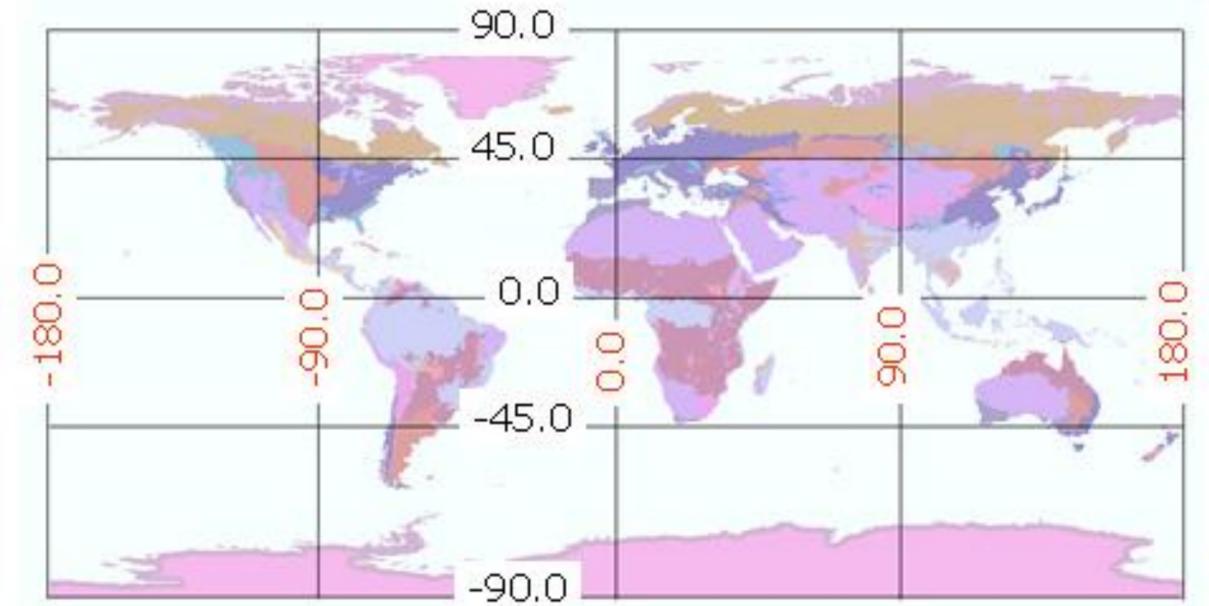
Basic GIS concepts

Geographic Coordinate Reference System(CRS)

Location is represented by longitude (-180, 180, from west to east) and latitude (-90, 90, from South to North)



- Longitude – values in **red**, range from 180 to -180
- Latitude – values in **black**, ranging from 90 to -90



WKT
POINT (Longitude, Latitude)
POINT(-71.1166 42.3744)

Projected CRS by EPSG/ESRI Code



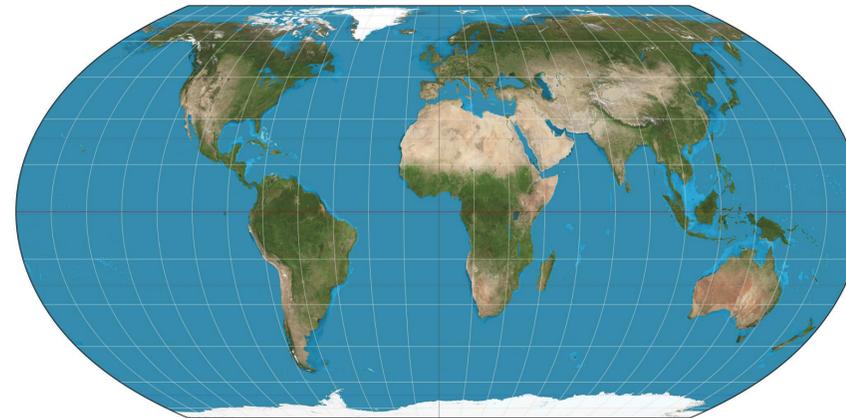
Commonly used coordinate reference systems such as **EPSG:3857 (Web Mercator)** and **EPSG:4326 (WGS 84 latitude/longitude)** are **not equal-area projections**.

EPSG:3857 (Web Mercator) significantly exaggerates area toward the poles.

EPSG:4326 (WGS 84 geographic coordinates) is mainly used for positioning and data storage, not for accurate area measurement or area-based visualization.

[Image Source](#)

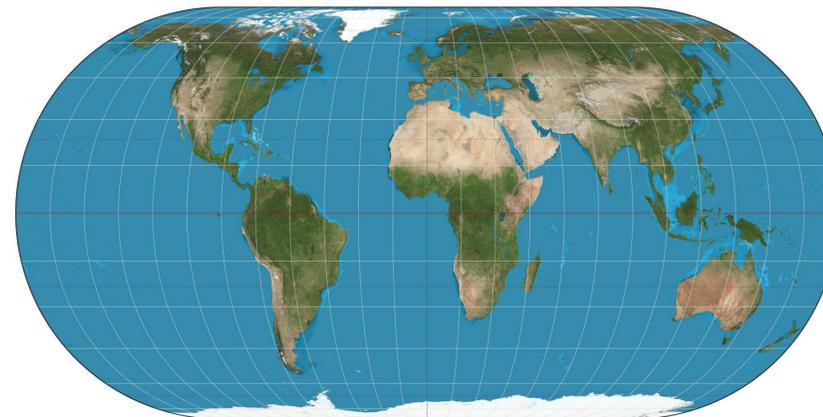
Equal Earth Projection



EPSG:8857 (WGS 84 / Equal Earth Greenwich)
ESRI:53035 (Equal Earth Greenwich)

One of the most recommended modern equal-area projections preserves area accuracy while providing a visually balanced and aesthetically pleasing

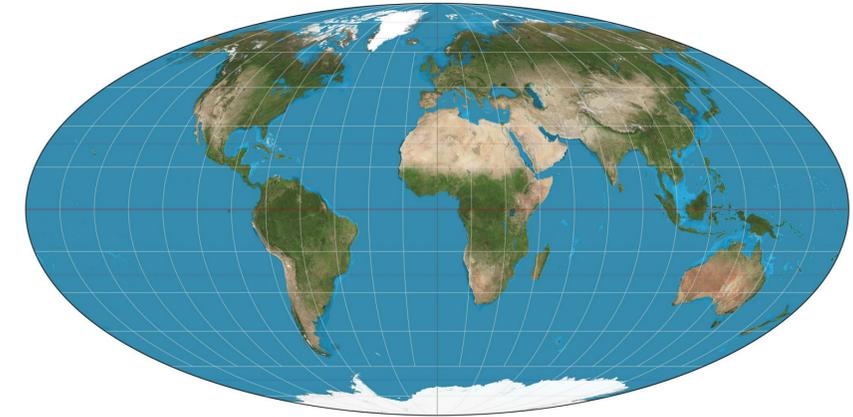
Eckert IV Projection



ESRI:54012 (World_Eckert_IV)
ESRI:53012 (spherical version)

An equal-area projection with straight poles (half the length of the equator) and curved meridians along the edges. It provides a visually smooth and balanced appearance and is commonly used in educational world maps.

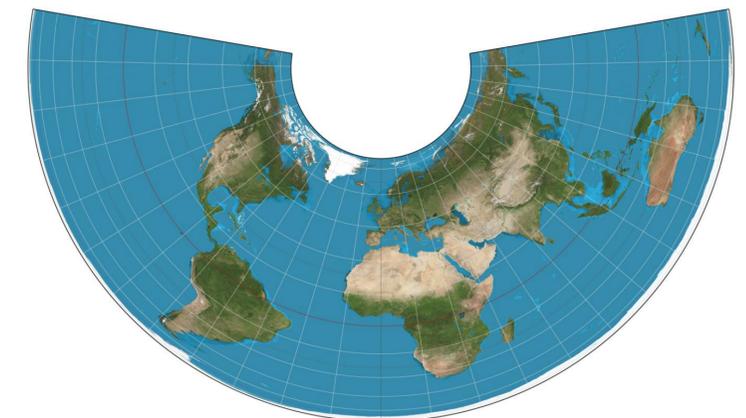
Mollweide Projection



ESRI:54009 (World_Mollweide) – most commonly used
ESRI:53009 (spherical version)

A classic pseudo-cylindrical equal-area projection elliptical in shape and is often used for global climate, environmental, or astronomical data.

Albers Equal-Area Conic Projection



ESRI:102003 (commonly used in North America; part of the Albers family)
EPSG:9822 (projection method code)

typically used for mid-latitude countries (such as the United States or China) in state/province-level area statistics, global variants also exist.

GIS Data Types-Vector in KNIME

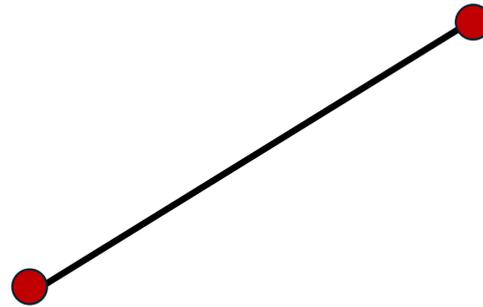
POINT



POINT(-71.116 42.374)

A pair of x and y coordinates

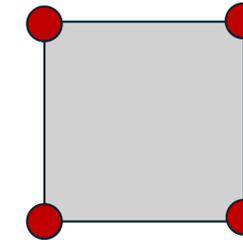
LINE



LINESTRING(-71.120 42.371, -71.112 42.371)

A series of points connected together in a specified order

POLYGON



POLYGON ((-71.120 42.371, -71.112 42.371, -71.112 42.377, -71.120 42.377, -71.120 42.371))

A line enclosed from end to end

Attribute column

Multiple geometry columns

Row0	County	Population	🌐 geometry	🌐 geometry (# 1)
Row1	A	1000	Polygon	Point
Row2	B	2000	MultiPolygon	Point
...

Geometry Type

Geo
 Geo (collection)
 Geo (line)
 Geo (multi line)
 Geo (point)
 Geo (multi point)
 Geo (polygon)
 Geo (multi polygon)

Raster Data Format

a raster consists of a matrix of pixels (or cells) organized into rows and columns (or a grid) in which each pixel contains a value representing information, such as image reflectance or temperature.

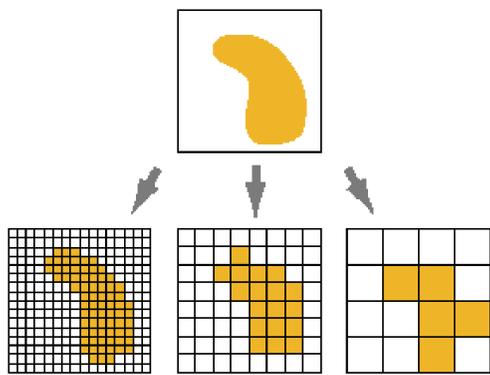
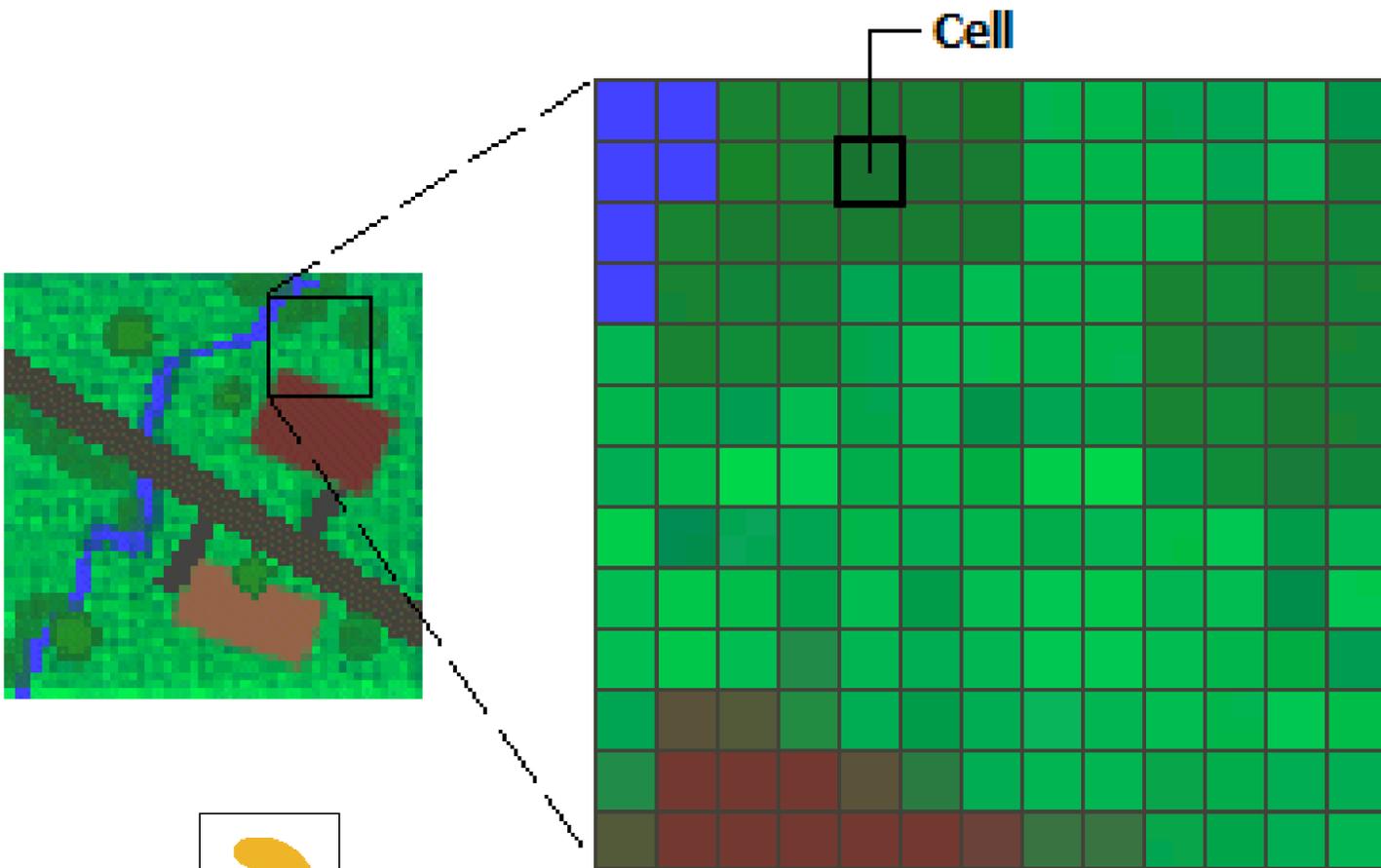
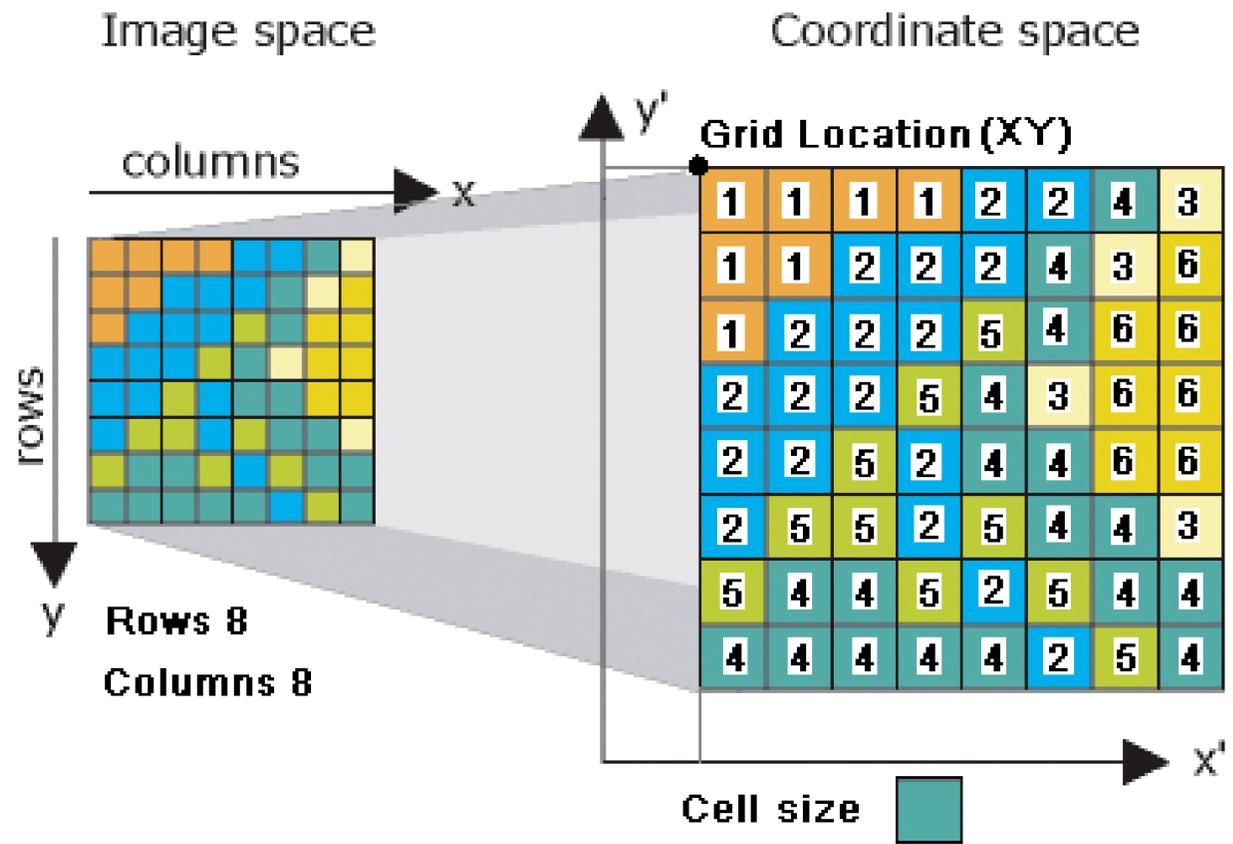


Image Source



List of cell values
[11112243112224361222546622254366225244662552544354452544444254]

GIS Data Format

Shapefile (.shp)

A shapefile is actually a collection of files, including:

.shp – geometry

.shx – shape index

.dbf – attribute table

.prj – projection information

(sometimes additional supporting files)

Important: All component files must stay together

Field name length limit (10 characters)

Zipped Shapefile

KML (.kml)

Used mainly in Google Earth

Stores geographic features and visualization styles

Often used for web-based or interactive mapping

GeoJSON (.geojson)

Text-based format using JSON structure

Widely used in web mapping

Easy to read and exchange across platforms

File Geodatabase (.gdb)

Developed by ESRI

Can store multiple layers in one folder

Supports large datasets and advanced data types

Common in ArcGIS environments

GeoPackage (.gpkg)

A modern, open-standard format based on SQLite.

Stores multiple layers in one single file

Can store vector and raster data together

Supports large datasets

No field name limitations like shapefile

Open standard (OGC)

GeoParquet (.parquet)

An extension of the Apache Parquet format designed for geospatial data. It is especially important for:

Big geospatial data

Cloud storage (AWS, GCP, Azure)

Python / Spark / GeoAI workflows

GeoTIFF (.tif)

One of the most common raster formats

Stores spatial reference information

JPEG (.jpg) with .jgw

Compressed image format

Requires a .jgw world file for spatial referencing

Smaller file size but lossy compression

PNG (.png) with .pgw

Lossless compression

Requires .pgw world file for georeferencing

.img – ERDAS Imagine format

.sid – MrSID compressed imagery

.BIL – Band Interleaved by Line

Grid – older ArcGIS raster format

netCDF – commonly used for climate and scientific multidimensional data

What is Geospatial Analysis Extension for KNIME?

A tool for spatial data analysis, manipulation, modelling and visualization

Based on visual programming platform of KNIME AP

Provides a diverse array of functionalities:

Data Input and output support for common GIS format

Spatial manipulation, calculation, conversion

Spatial network routing and Location optimization

Enriched open dataset

Spatial Modelling tools

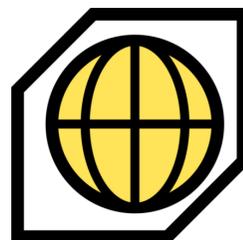
such as Geographically Weighted Regression (GWR), MGWR, SLM, SEM

Support vector data only

The KNIME GEE extension support raster data, and cloud computation via GEE



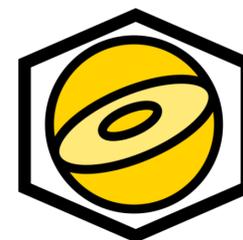
Spatial IO



Spatial Calculation



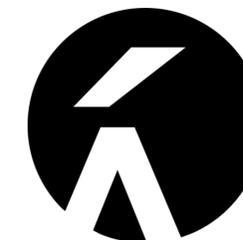
Spatial Manipulation



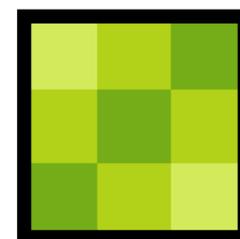
Spatial Transformation



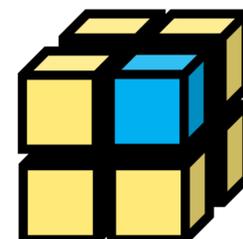
Spatial Conversion



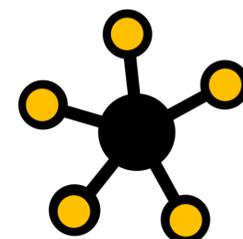
Spatial Visualization



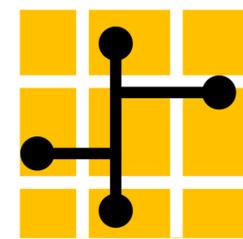
Exploratory Spatial
Data Analysis



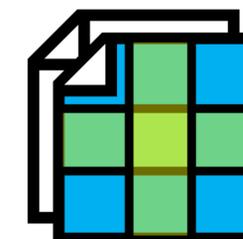
Spatial Modelling



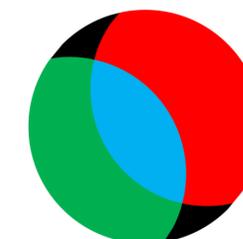
Location Analysis



Spatial Network



Open Dataset



Spatial Clustering

Hands-on 1

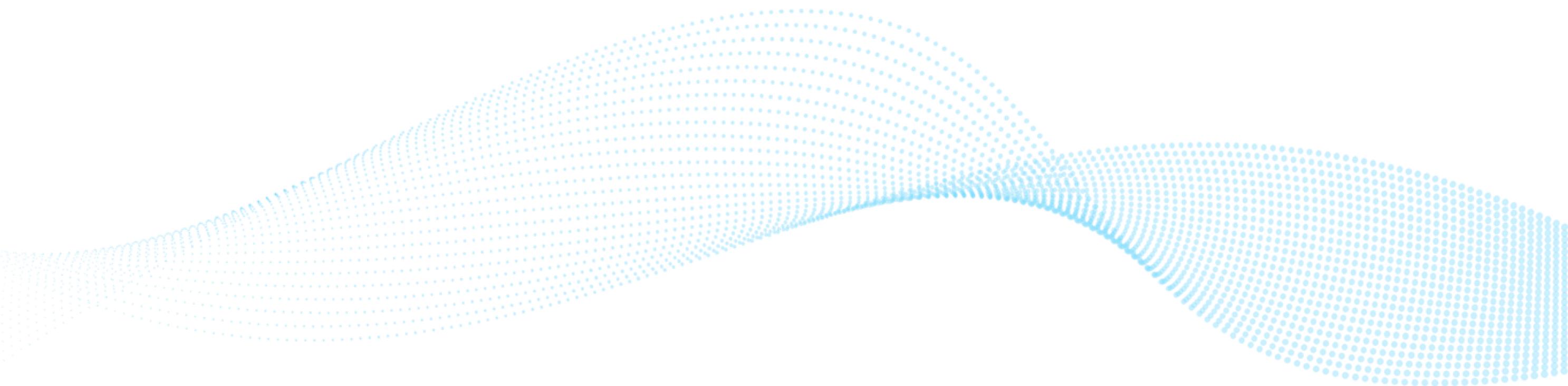
Visual Programming for GIS

KNIME Analytics Platform



Center for Geographic Analysis
Institute for Quantitative Social Science
Harvard University

3

A decorative graphic consisting of a series of light blue dots arranged in a wavy, horizontal pattern that spans across the middle of the slide.

**Under Spatial Data with Geospatial
Analytics extension for KNIME**

Hands-on

Inspect Node Monitor

Selected the executed node to inspect its outputs.

The screenshot shows the KNIME GIS interface. On the left, the 'Nodes' sidebar lists various data sources under 'Open Datasets', with 'Natural Earth Global Data' selected. The central workspace displays the 'Natural Earth Global Data' node icon. The right panel shows configuration options for the node, including 'Data category' (All scale, 10M-POV, 10M-Special), 'Scale' (1:10m, 1:50m, 1:110m), and 'All scale file type' (Country boundaries). Below the node, the 'Monitor' tab is active, showing a table of 177 rows and 169 columns. The table has columns for RowID, feature name, scalerank, and various geographical attributes. A 'Statistics' tab is also visible, and a search and filter icon is present in the top right of the table view.

#	RowID	featurecla	scalerank	LABELRA...	SOVEREI...	SOV_A3	ADM0_DIF	LEVEL	TYP
1	Row0	Admin-0 countr	1	6	Fiji	FJI	0	2	Sovereigr
2	Row1	Admin-0 countr	1	3	United Republic	TZA	0	2	Sovereigr
3	Row2	Admin-0 countr	1	7	Western Sahara	SAH	0	2	Indeterm
4	Row3	Admin-0 countr	1	2	Canada	CAN	0	2	Sovereigr

Missing values
Unique values
Min and max
Quantiles

Switch to
Statistics tab

Search and Filter

Table dimension

Column name
and type

Data Manipulation-Column filter

Keep **Natural Earth Global Data** selected, then go to: **Nodes > Manipulation > Column > Column Filter**, Double-click **Column Filter** to add the node to the workflow editor. It will automatically be linked to the first node. Alternatively, you can drag the **Column Filter** node into the editor and manually draw a connection between the two nodes. A replicable way : Choose **Regex** and input “ADMIN|POP_EST|GDP_MD|INCOME_GRP|REGION_UN|geometry”

The screenshot shows the KNIME GIS interface with the Column Filter node configuration panel open. The configuration panel is set to 'Regex' mode. The 'Includes' list contains the following columns: ADMIN, POP_EST, GDP_MD, INCOME_GRP, REGION_UN, and geometry. The 'Excludes' list contains various FCLASS_* columns. The 'Apply and Execute' button is highlighted. Below the configuration panel, the workflow editor shows a connection between the 'Natural Earth Global Data' node and the 'Column Filter' node. At the bottom, a table view shows the filtered data with 177 rows and 6 columns. The 'geometry' column is highlighted in blue.

#	RowID	ADMIN	POP_EST	GDP_MD	INCOME_GRP	REGION_UN	geometry
1	Row0	Fiji	889,953	5496	4. Lower middle inc	Oceania	MULTIPOLYGON - E
2	Row1	United Republic of 1	58,005,463	63177	5. Low income	Africa	POLYGON - EPSG:4
3	Row2	Western Sahara	603,253	907	5. Low income	Africa	POLYGON - EPSG:4
4	Row3	Canada	37,589,262	1736425	1. High income: OE	Americas	MULTIPOLYGON - E
5	Row4	United States of An	328,239,523	21433226	1. High income: OE	Americas	MULTIPOLYGON - E

Use Regex

In the configuration panel, use the <, >, <<, and >> buttons to move columns into the **Includes** box. click **Execute**.

Here is the **geometry** column Invisible in ArcGIS or QGIS

View Geospatial data Row in KNIME

In KNIME, geospatial data with a geographic coordinate system can be viewed row by row.

Double-click the 4th row in the **geometry** column (where **ADMIN = Canada**), and a floating map will pop up showing this spatial feature. You can also click the drop-down icon next to the **geometry** column name and switch to **Full Well-Known Text (WKT)** to view the textual representation of the spatial feature.



geometry ↑
Geo

```
MULTIPOLYGON (((-122.84000000000003 49.000000000000004 -122.84000000000003 49.000000000000004 -122.84000000000003 49.000000000000004)))
```

```
MULTIPOLYGON (((-48.66061601418252 -78.047000000000004 -48.66061601418252 -78.047000000000004 -48.66061601418252 -78.047000000000004)))
```

Data renderer

WKT type and CRS

Coordinate Reference...

Full Well Known Text

Multi-line String

String

Hex Dump (short)

Hex Dump (long)

Blob information

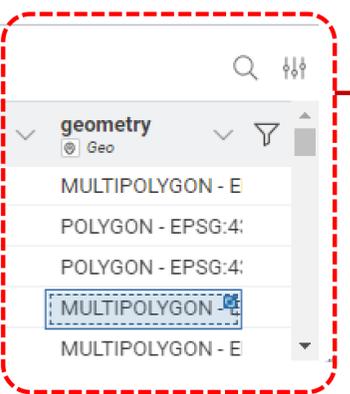
geometry ↓
Geo

1: Filtered table | Flow Variables

Rows: 177 | Columns: 6

Table | Statistics

#	RowID	ADMIN	POP_EST	GDP_MD	INCOME_GRP	REGION_UN	geometry
1	Row0	Fiji	889,953	5496	4. Lower middle inc	Oceania	MULTIPOLYGON - E
2	Row1	United Republic of T	58,005,463	63177	5. Low income	Africa	POLYGON - EPSG:4
3	Row2	Western Sahara	603,253	907	5. Low income	Africa	POLYGON - EPSG:4
4	Row3	Canada	37,589,262	1736425	1. High income: OE	Americas	MULTIPOLYGON
5	Row4	United States of Am	328,239,523	21433226	1. High income: OE	Americas	MULTIPOLYGON - E



Here is the **geometry** column Invisible in ArcGIS or QGIS

Data Manipulation-Row filter

Attribute queries are a very common task in GIS. Here, we use a **Row Filter** to extract countries in Africa.

(1)Click **Add Filter Criterion**;(2) Select **REGION_UN**, Choose **Equals**, Enter **Africa** in the *Value* box, Click **Execute**.

Then, inspect the node monitor area to view the filtered results

The screenshot displays the KNIME GIS interface. On the left, the 'Nodes' panel shows the 'Row Filter' node selected under 'Manipulation > Row > Filter'. The main workspace shows a workflow with three nodes: 'Natural Earth Global Data', 'Column Filter', and 'Row Filter'. The 'Row Filter' node is highlighted with a red dashed box. Below the workflow, the 'Node Monitor' area shows the configuration for the 'Row Filter' node. The 'Filter' section is expanded, showing 'Criterion 1' with 'Filter column' set to 'REGION_UN', 'Operator' set to 'Equals', and 'Value' set to 'Africa'. The 'Case matching' section has 'Case sensitive' selected. The 'Value' field contains 'Africa'. The 'Output' section has 'Output matching rows' selected. Below the configuration, the 'Node Monitor' area shows the results of the filter. The table has 52 rows and 6 columns. The first five rows are displayed, with the 'REGION_UN' column highlighted by a red dashed box. The values in the 'REGION_UN' column are 'Africa', 'Africa', 'Africa', 'Africa', and 'Africa'. A red arrow points from the 'REGION_UN' column in the table to the 'Only Africa remains' text box on the right.

1: Included Rows | Flow Variables

#	RowID	ADMIN	POP_EST	GDP_MD	INCOME_GRP	REGION_UN	geometry
1	Row1	United Republic of 1	58,005,463	63177	5. Low income	Africa	POLYGON - EPSG:4:
2	Row2	Western Sahara	603,253	907	5. Low income	Africa	POLYGON - EPSG:4:
3	Row11	Democratic Republi	86,790,567	50400	5. Low income	Africa	POLYGON - EPSG:4:
4	Row12	Somalia	10,192,317.3	4719	5. Low income	Africa	POLYGON - EPSG:4:
5	Row13	Kenya	52,573,973	95503	5. Low income	Africa	POLYGON - EPSG:4:

Only Africa remains

Data Manipulation-Math Formula

Nodes → **Manipulation** → **Column** → **Convert & Replace** → **Math Formula**, Add and connect it to the **Row Filter** node.

Click **Open Dialog** on the right panel of the node configuration window.

In the dialog box, enter the following formula: $\$GDP_MD\$ * 1000000 / \$POP_EST\$$. Enter **GPC** in the **Append Column** field for **GDP per capita**. Click **OK**, then click the small **Execute (triangle)** button on the node to run it.

The screenshot displays the KNIME software interface. On the left, the 'Nodes' panel shows the 'Math Formula' node selected under 'Manipulation' > 'Column' > 'Convert & Replace'. The main workspace shows a workflow: 'Natural Earth Global Data' → 'Column Filter' → 'Row Filter' → 'Math Formula'. The 'Math Formula' node configuration window is open, showing the formula $\$GDP_MD\$ * 1000000 / \$POP_EST\$$ and 'GPC' in the 'Append Column' field. The 'Column List' panel shows '\$POP_EST' and '\$GDP_MD' selected. The 'Functions' panel shows 'x / y' selected. The 'Output' panel shows a table with a new column 'GPC'.

Annotations:

- Red arrows point to the 'Open dialog' button in the node configuration window.
- Red arrows point to the 'Column List' panel, highlighting '\$POP_EST' (1) and '\$GDP_MD' (5).
- Red arrows point to the 'Functions' panel, highlighting 'x / y' (4) and 'x * y' (2).
- Red arrows point to the 'Expression' field, highlighting the formula (3).
- Red arrows point to the 'Append Column' field, highlighting 'GPC'.
- Red arrows point to the 'GPC' column in the output table.

Output Table:

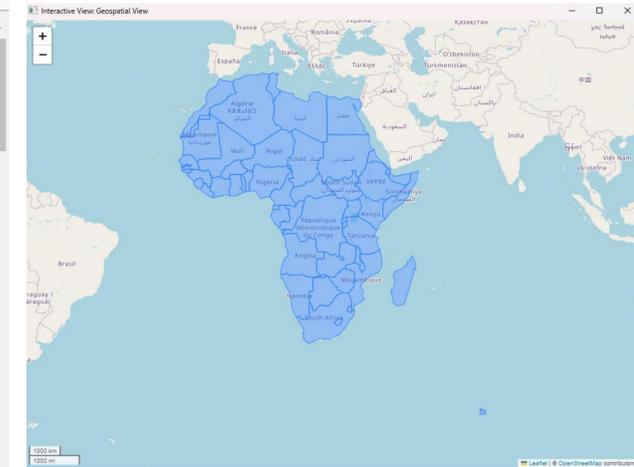
geometry	GPC
POLYGON - EPSG:31466	1,089.156
POLYGON - EPSG:31466	1,503.515
POLYGON - EPSG:31466	580.708
POLYGON - EPSG:31466	462.996
POLYGON - EPSG:31466	1,816.545

$\$GDP_MD\$$ and $\$POP_EST\$$ are column names. To avoid typing errors, you can double-click the column names in the **Column List** panel to insert them into the formula. Mathematical operators can also be inserted by double-clicking items under the **Functions** panel.

Visualization-Geospatial View

Nodes → **Community Nodes** → **Spatial Visualization** → **Geospatial View**, Add and connect it to the **Math Formula** node. Click the small **Execute (triangle)** button on the node to run it using the default settings. After execution, click **Open in New Window** to display the interactive map. You can then **pan and zoom** to explore the data dynamically.

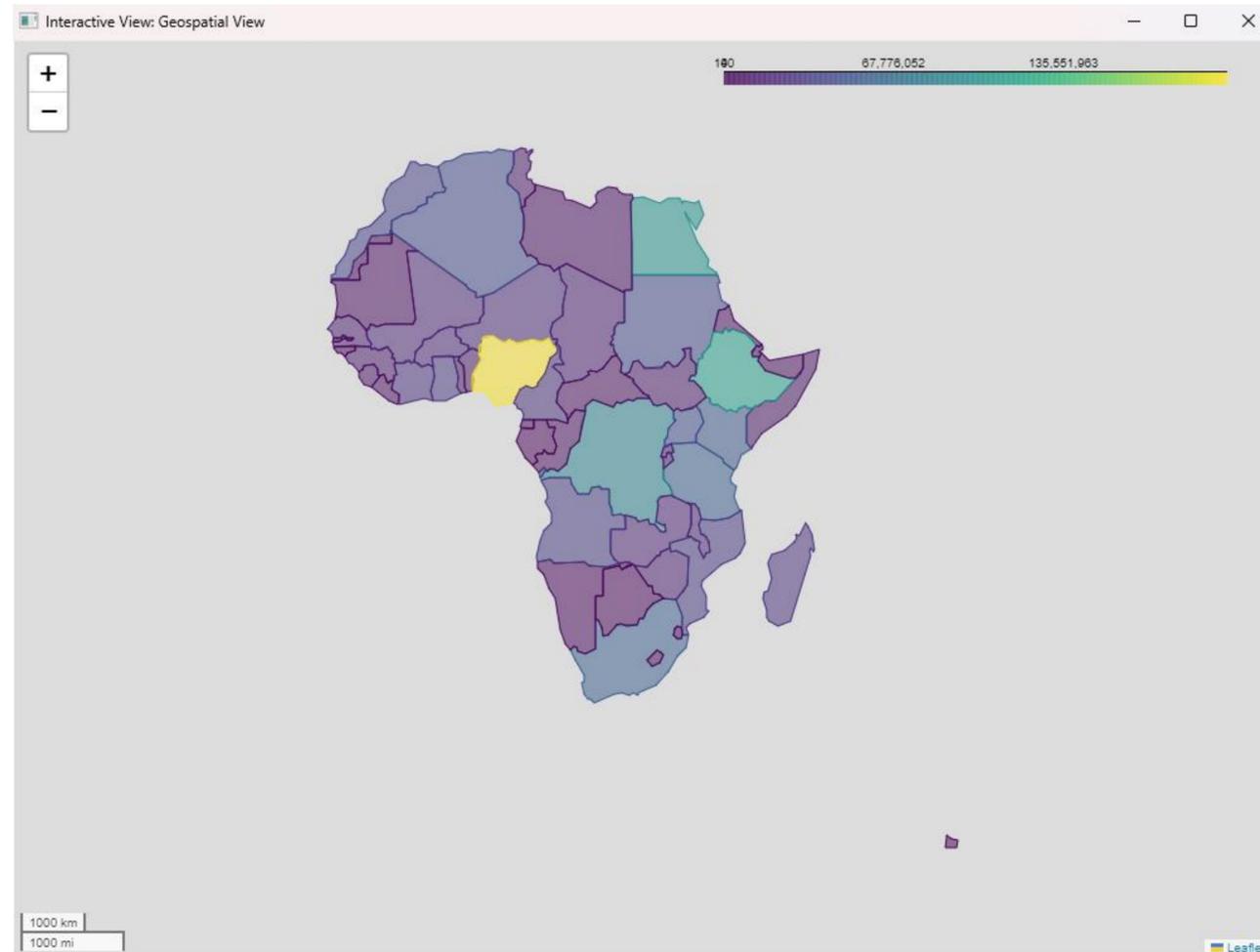
The screenshot displays the KNIME GIS interface. The top navigation bar includes 'Home', 'KNIME-GIS', 'Preferences', 'Menu', and 'pianu'. Below the navigation bar are buttons for 'Execute', 'Cancel', and 'Reset'. The main workspace shows a workflow with the following nodes: 'Natural Earth Global Data', 'Column Filter', 'Row Filter', 'Math Formula', and 'Geospatial View'. The 'Geospatial View' node is highlighted, and its configuration panel is open on the right. The configuration panel includes a 'Geometry column' dropdown set to 'geometry', a checked 'Stroke' checkbox, and a 'Marker tooltip columns' section with 'Excludes' and 'Includes' lists. The 'Excludes' list contains 'ADMIN', 'POP_EST', 'GDP_MD', 'INCOME_GRP', 'REGION_UN', and 'GPC'. The 'Includes' list is empty, with the text 'No entries in this list'. At the bottom of the configuration panel are buttons for 'Discard', 'Apply and Execute', and 'Apply'. Below the workflow is a 'View' tab with a 'Flow Variables' icon and a map showing the workflow's output. The map includes a scale bar (5000 km, 3000 mi) and a 'Leaflet | © OpenStreetMap contributors' footer. An 'Open in new window' button is visible on the map.



Visualization-Configure Geospatial View

Click >> to add all columns to the **Marker Popup Columns** (or **Tooltip Columns**) section. For **Marker Color Column**, select **POP_EST**. In the **Base Map** settings, choose **Don't show base map**. Check the option **Show Color Legend** to display the legend. Click **Execute** to rerun the node with the updated settings.

You can also experiment with other map settings



Geospatial View

Geometry column

Stroke

Marker tooltip columns

Excludes	Includes
<input type="checkbox"/> ADMIN	
<input checked="" type="checkbox"/> POP_EST	
<input checked="" type="checkbox"/> GDP_MD	
<input type="checkbox"/> INCOME_GRP	
<input type="checkbox"/> REGION_UN	
<input checked="" type="checkbox"/> GPC	

Marker popup columns

Excludes	Includes
	<input type="checkbox"/> ADMIN
	<input checked="" type="checkbox"/> POP_EST
	<input checked="" type="checkbox"/> GDP_MD
	<input type="checkbox"/> INCOME_GRP
	<input type="checkbox"/> REGION_UN
	<input checked="" type="checkbox"/> GPC

Geospatial View

Size Settings

Marker size column

Coloring Settings

Marker color column

Color map

Classify numerical marker color columns ...

Base Map Setting

Base map

Color Legend Settings

Show color legend

Change Data Stream to Visualize Global Map

Move the **Row Filter** node slightly upward to make space in the workflow. Then draw a new connection from the **Column Filter** node directly to the **Math Formula** node. This allows you to bypass the Africa-only filtering step. Next, click the **Geospatial View** node and press **Execute** to rerun the workflow and visualize the **global map** instead of only Africa. This is your first workflow with a **branch structure**.

The screenshot displays the KNIME GIS interface with a workflow and its output. The workflow consists of the following nodes: Natural Earth Global Data, Column Filter, Row Filter, Math Formula, Projection, and Geospatial View. The Column Filter node is connected to both the Row Filter and the Math Formula nodes. The Projection node is connected to the Math Formula node. The Geospatial View node is connected to the Projection node. The Geospatial View node is selected, showing a world map visualization. The map shows a color scale from 471,388,328 to 942,772,516. The map is titled "Open in new window".

The output table shows the following data:

#	RowID	ADMIN	POP_EST	GDP_MD	INCOME_GRP	REGION_UN	geometry
1	Row0	Fiji	889,953	5496	4. Lower middle inc	Oceania	MULTIPOLYGON - E
2	Row1	United Republic of 1	58,005,463	63177	5. Low income	Africa	POLYGON - EPSG: 8
3	Row2	Western Sahara	603,253	907	5. Low income	Africa	POLYGON - EPSG: 8
4	Row3	Canada	37,589,262	1736425	1. High income: OEI	Americas	MULTIPOLYGON - E
5	Row4	United States of Am	328,239,523	21433226	1. High income: OEI	Americas	MULTIPOLYGON - E

This demonstrates one of the key advantages of KNIME: it helps you clearly organize your workflow and quickly understand your data processing logic.

Reproject coordinate reference systems (CRS)

Nodes → **Community Nodes** → **Spatial Transformation** → **Projection**, Add and connect it to the **Column Filter** node.

Input **EPSG:8857** for New CRS. Click **Execute** to apply the coordinate system transformation.

You can define coordinate reference systems (CRS) using either **EPSG codes** or **ESRI codes**.

For more information about CRS definitions and codes, visit: <https://epsg.io>.

The screenshot displays the KNIME GIS interface with a workflow for reprojecting coordinate reference systems (CRS). The workflow consists of the following nodes: **Natural Earth Global Data**, **Column Filter**, **Row Filter**, **Math Formula**, and **Geospatial View**. A **Projection** node is connected to the **Column Filter** node. The **Projection** node's configuration panel is open, showing the following settings:

- Geometry column:** geometry
- New CRS:** EPSG: 8857
- Output column:** Replace (selected), Append

Below the workflow, a table view shows the output data with the following columns: #, RowID, ADMIN, POP_EST, GDP_MD, INCOME_GRP, REGION_UN, and geometry. The geometry column is highlighted with a red dashed box, and an arrow points to a callout box that says "Check the EPSG code".

#	RowID	ADMIN	POP_EST	GDP_MD	INCOME_GRP	REGION_UN	geometry
1	Row0	Fiji	889,953	5496	4. Lower middle inc	Oceania	MULTIPOLYGON - E
2	Row1	United Republic of 1	58,005,463	63177	5. Low income	Africa	POLYGON - EPSG: 8
3	Row2	Western Sahara	603,253	907	5. Low income	Africa	POLYGON - EPSG: 8
4	Row3	Canada	37,589,262	1736425	1. High income: OE	Americas	MULTIPOLYGON - E
5	Row4	United States of An	328,239,523	21433226	1. High income: OE	Americas	MULTIPOLYGON - E

Calculate Geometry Attribution

Nodes → **Community Nodes** → **Spatial Calculation** → **Area**, Add and connect it to the **Projection** node. Execute it with the default settings. The node will append a new column named **Area**, calculated in square meters (based on the projected coordinate system: **EPSG:8857**). You can return to the **Projection** node and change the projection code to another CRS to compare how the calculated area values differ under different coordinate systems.

The screenshot displays the KNIME GIS interface. The workflow consists of the following nodes: **Natural Earth Global Data**, **Column Filter**, **Row Filter**, **Math Formula**, **Geospatial View**, **Projection**, and **Area**. The **Area** node is highlighted with a blue border and the text "Add comment" below it.

The **Area** node configuration panel is open, showing the following settings:

- Geometry column: **geometry**
- Output column: Replace Append
- New column name: **area**

The output table is displayed below the workflow, showing 177 rows and 7 columns. The **area** column is highlighted with a red dashed box and a red arrow pointing to the text "area column".

#	RowID	ADMIN	POP_EST	GDP_MD	INCOME_...	REGION_UN	geometry	area
1	Row0	Fiji	889,953	5496	4. Lower middle	Oceania	MULTIPOLYGON	19,285,324,295.7
2	Row1	United Republic	58,005,463	63177	5. Low income	Africa	POLYGON - EPSG	932,728,895,075
3	Row2	Western Sahara	603,253	907	5. Low income	Africa	POLYGON - EPSG	96,735,164,427.5
4	Row3	Canada	37,589,262	1736425	1. High income:	Americas	MULTIPOLYGON	10,033,994,236.5
5	Row4	United States of	328,239,523	21433226	1. High income:	Americas	MULTIPOLYGON	9,513,248,938,61

Visualize the Projected Maps

Nodes → **Community Nodes** → **Spatial Visualization** → **Geospatial View Static**, Add and connect it to the **Area** node. Run it using the default settings. After execution, click **Open in New Window** to display the static map.

The screenshot displays the KNIME GIS interface. On the left, the 'Nodes' panel is open, showing a search bar and a tree view of node categories. Under 'Community Nodes' > 'Spatial Visualization', 'Geospatial View Static' is highlighted. The main workspace shows a workflow: 'Natural Earth Global Data' → 'Column Filter' → 'Row Filter' → 'Math Formula' → 'Geospatial View'. A second path branches from 'Column Filter' to 'Projection' → 'Area' → 'Geospatial View Static'. The 'Geospatial View Static' node is selected, and its configuration panel is open on the right. The configuration includes: 'Geometry column' set to 'geometry', 'Line width column' set to 'No value selected', 'Line width' set to '1', and 'Image output type' set to 'Image'. Below the configuration, there are 'Discard', 'Apply and Execute', and 'Apply' buttons. At the bottom, a preview window shows a world map with a blue overlay, and an 'Open in new window' button is visible.

Save Geodata File

Nodes → **Community Nodes** → **Geospatial Analytics** → **Spatial IO** → **GeoFile Writer**, Add and connect it to the **Area** node. Open the node configuration dialog and: 1) Enter an appropriate **output file path and file name**;2) Choose the desired **file format** (e.g., Shapefile, GeoPackage, GeoJSON, etc.);3) Click **Execute** to export and save the file. Click the **Save icon** in the upper-left corner to save your workflow.

The screenshot displays the KNIME GIS interface with a workflow and the configuration for the GeoFile Writer node. Annotations highlight key elements:

- Close Project**: A red box highlights the 'Close Project' button in the top toolbar.
- Save file**: A red box highlights the save icon in the top toolbar.
- Open Project**: A red box highlights the 'Explorer' icon in the left sidebar.
- Geometry column**: A red box highlights the 'geometry' dropdown in the configuration dialog.
- Icon for choosing any file in the target folder, then revise the file name**: A red box highlights the file selection icon next to the output file path.

The workflow consists of the following nodes: Natural Earth Global Data, Column Filter, Row Filter, Math Formula, Geospatial View, Projection, Area, Geospatial View Static, and GeoFile Writer. The GeoFile Writer node configuration is as follows:

Property	Value
Geometry column	geometry
Output file path	C:\Users\lil928\Downloads\WorldMap
If exists:	Fail
Output file format	GeoParquet
File compression	gzip

Flow Variables table:

Owner ID	Data Type	Variable Name	Value
	StringType	knime.workspace	C:\Users\lil928\knime-workspace

Upload Workflow to KNIME Community Hub

Close your project. Go to Space Explorer, right-click on the workflow file, choose upload, in the pop-up dialog, choose Private or Public for different workflow invisibility.

Or choose Export, to save it to your local space.

Switch between Local and Cloud

Open Project

The screenshot displays the KNIME GIS interface with a workflow diagram in the center. The workflow starts with 'Natural Earth Global Data', followed by 'Column Filter', which branches into 'Row Filter' and 'Projection'. 'Row Filter' leads to 'Math Formula', which then connects to 'Geospatial View'. 'Projection' leads to 'Area', which connects to 'Geospatial View Static' and 'GeoFile Writer'. A red dashed box highlights the 'Local space' dropdown in the top-left corner. Another red dashed box highlights the 'Explorer' icon in the left sidebar. A context menu is open over the workflow, with 'Upload' selected. An 'Upload to...' dialog box is open on the right, showing the 'KNIME Community Hub' folder selected, with sub-folders for 'piantu', 'Private', and 'Public'. A grey box with the text 'Choose workflow invisibility' is overlaid on the dialog. The dialog has 'Cancel' and 'Choose' buttons at the bottom.

Right click
Choose
Upload

Upload to...

Select an upload destination folder:

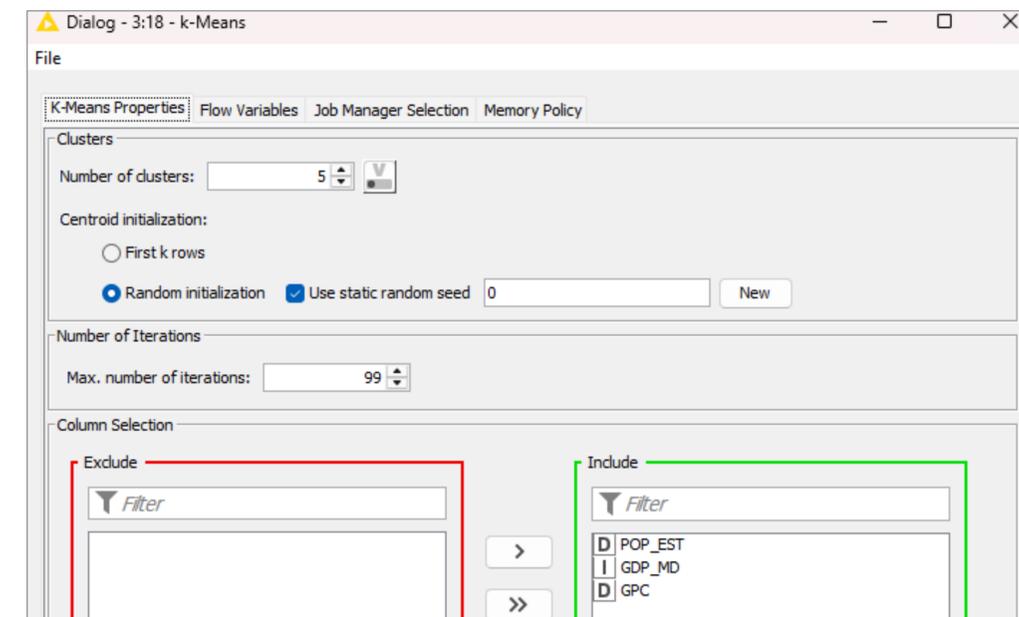
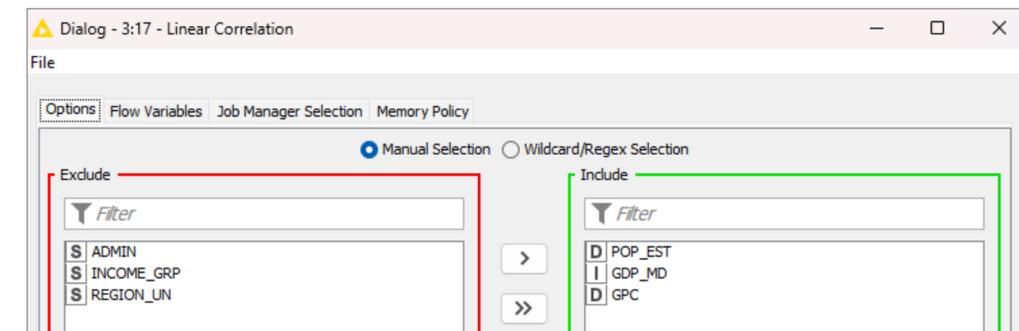
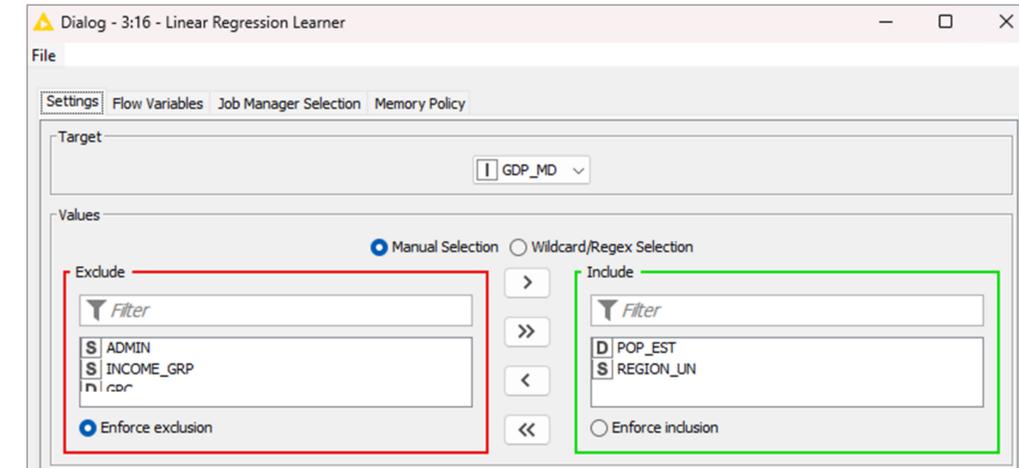
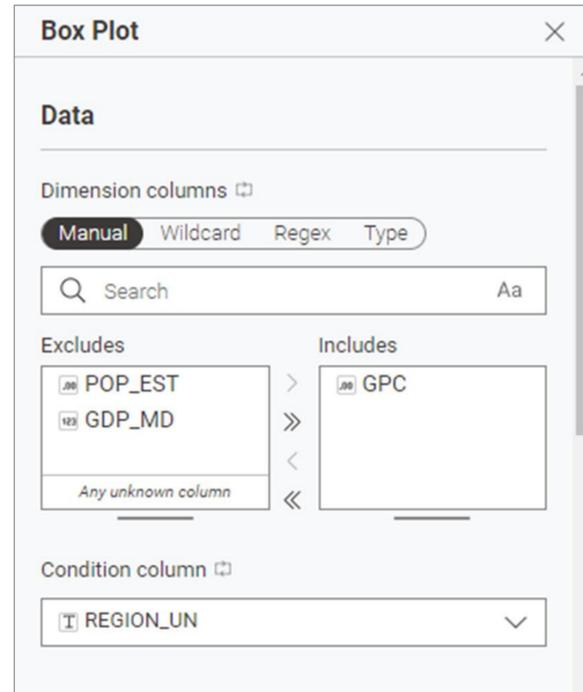
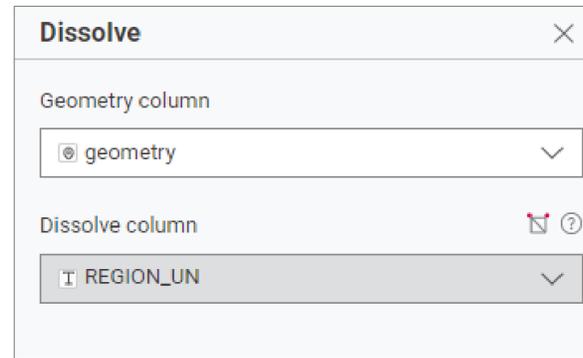
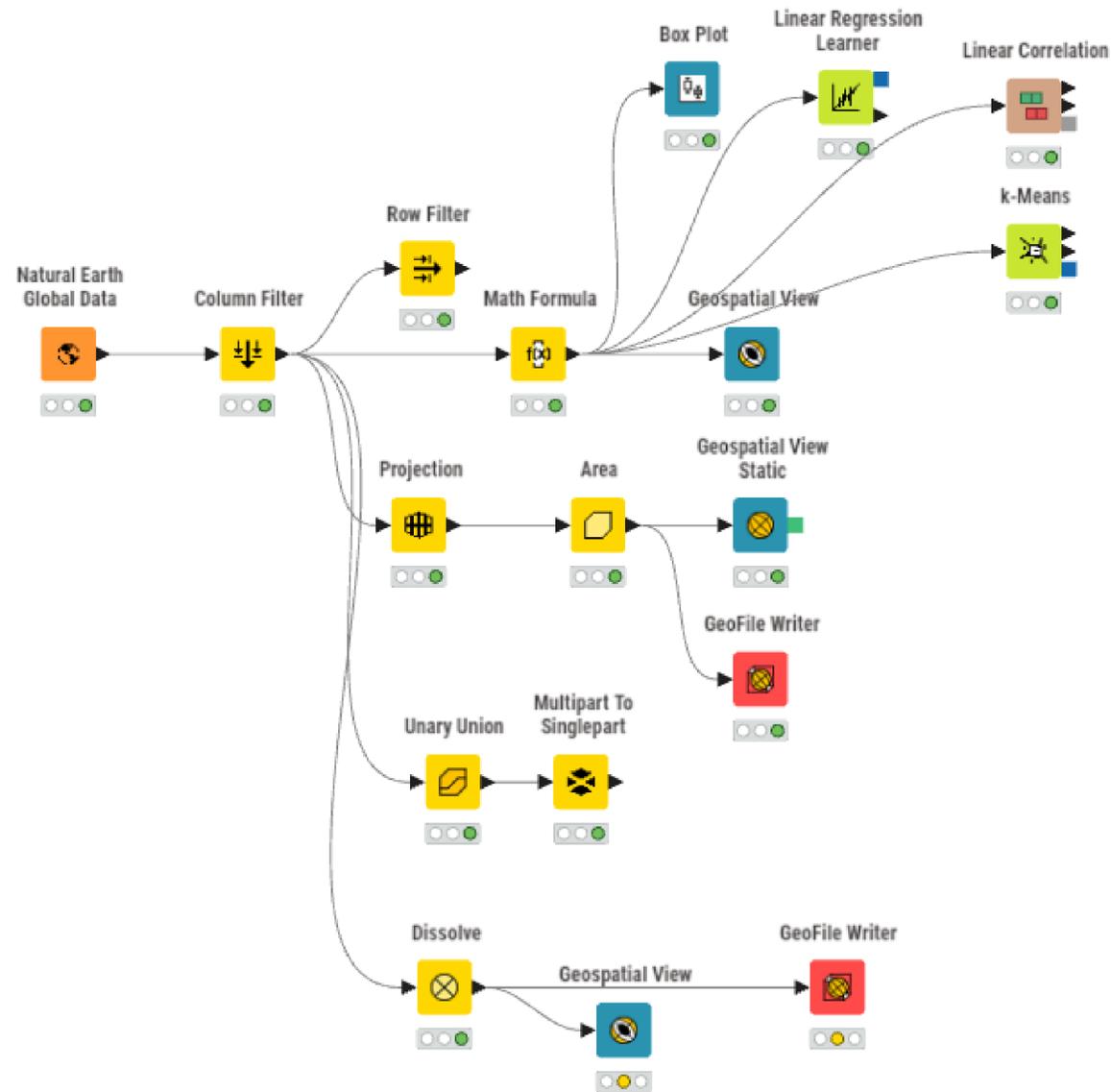
- KNIME Community Hub
 - piantu
 - Private
 - Public

Choose workflow invisibility

Cancel Choose

Want to Try other nodes?

Reopen your project, search the nodes and add them, use the same setting here...



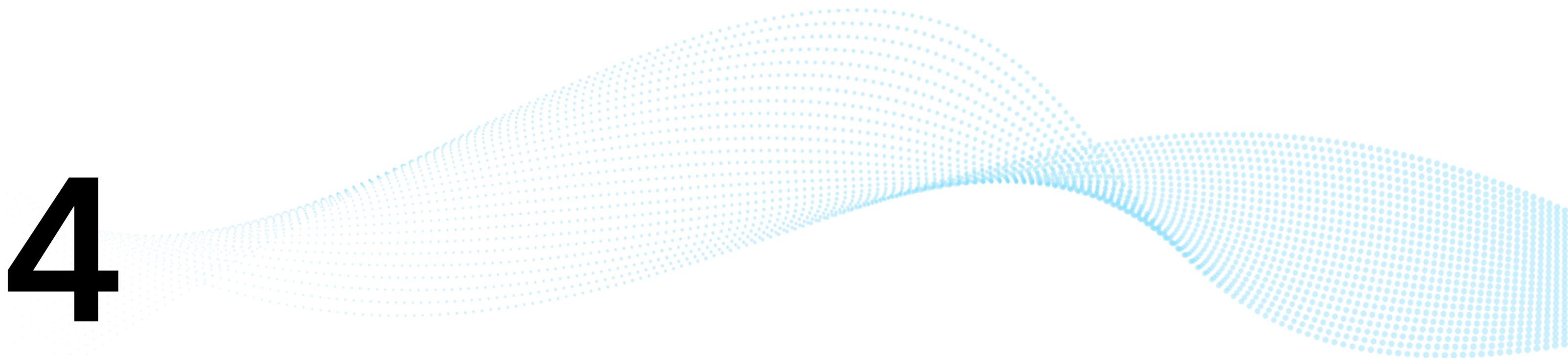
Lecture 2

Visual Programming for GIS

KNIME Analytics Platform



Center for Geographic Analysis
Institute for Quantitative Social Science
Harvard University

A decorative graphic consisting of a series of light blue dots arranged in a wavy, horizontal pattern that spans across the middle of the slide. The dots are more densely packed in some areas, creating a sense of depth and movement.

4

Common Spatial Manipulation Task

Visual Programming nodes Comparison

Most Frequently Used KNIME Core Nodes

The top commonly used **core data processing nodes** include:

Math Formula – perform column-based calculations

Joiner – combine two tables based on key columns

GroupBy – aggregate data (sum, count, mean, etc.)

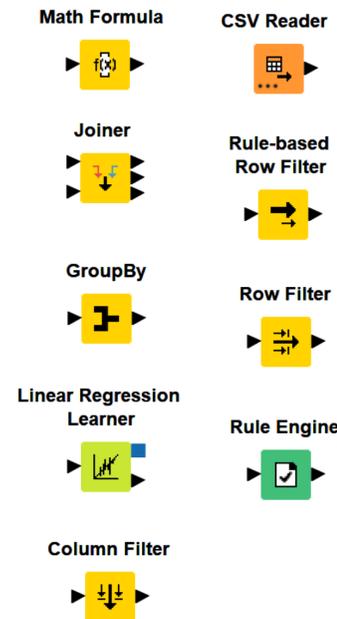
CSV Reader – import tabular data

Column Filter – select or remove columns

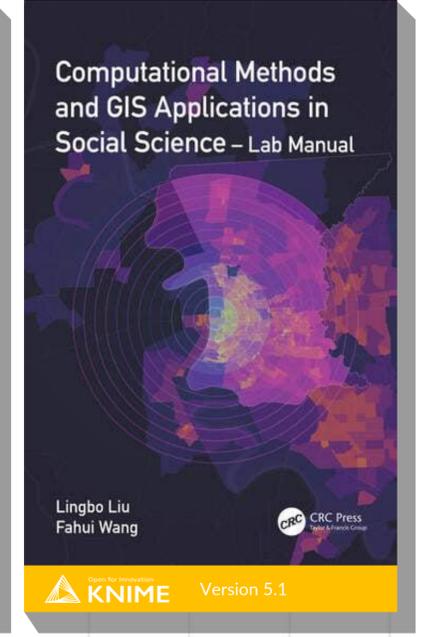
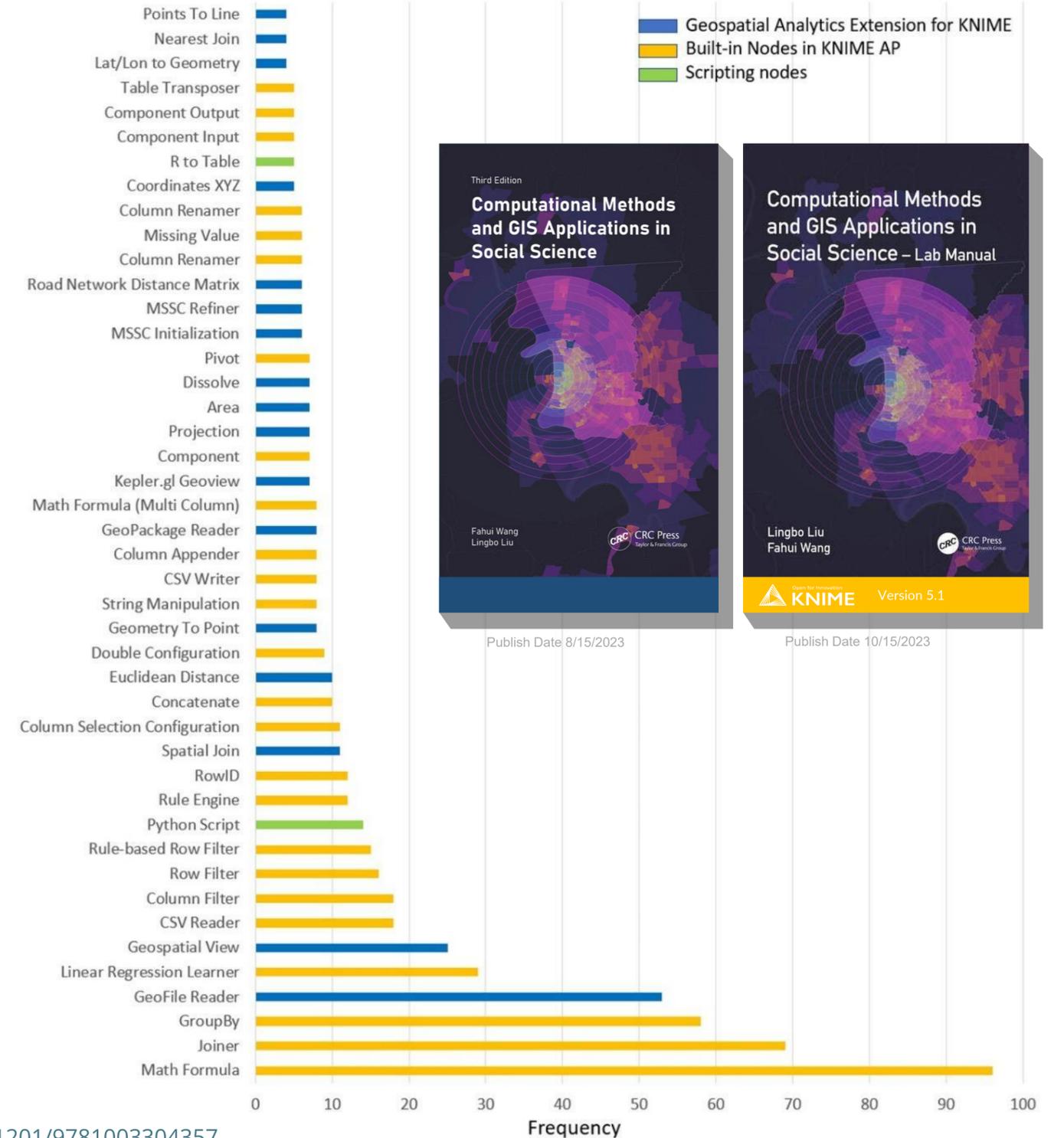
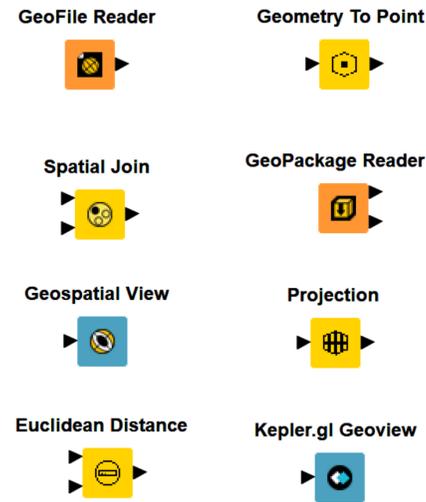
Row Filter – filter records based on conditions

These nodes form the foundation of most KNIME workflows.

KNIME Nodes



Geospatial Nodes



Lingbo Liu, Fahui Wang, Computational Methods and GIS Applications in Social Science - Lab Manual, CRC, 2023, <https://doi.org/10.1201/9781003304357>
 Fahui Wang, Lingbo Liu, Computational Methods and GIS Applications in Social Science, CRC, 2023, <https://doi.org/10.1201/9781003292302>

Visual Programming nodes Comparison

Most Frequently Used Geospatial Nodes

The top commonly used **geospatial analysis nodes** include:

GeoFile Reader – import spatial vector data

Geospatial View – visualize spatial data interactively

Spatial Join – join data based on spatial relationships

Euclidean Distance – calculate straight-line distance

Geometry to Point – convert polygons to centroids

Projection – transform coordinate reference systems (CRS)

Area – calculate polygon area

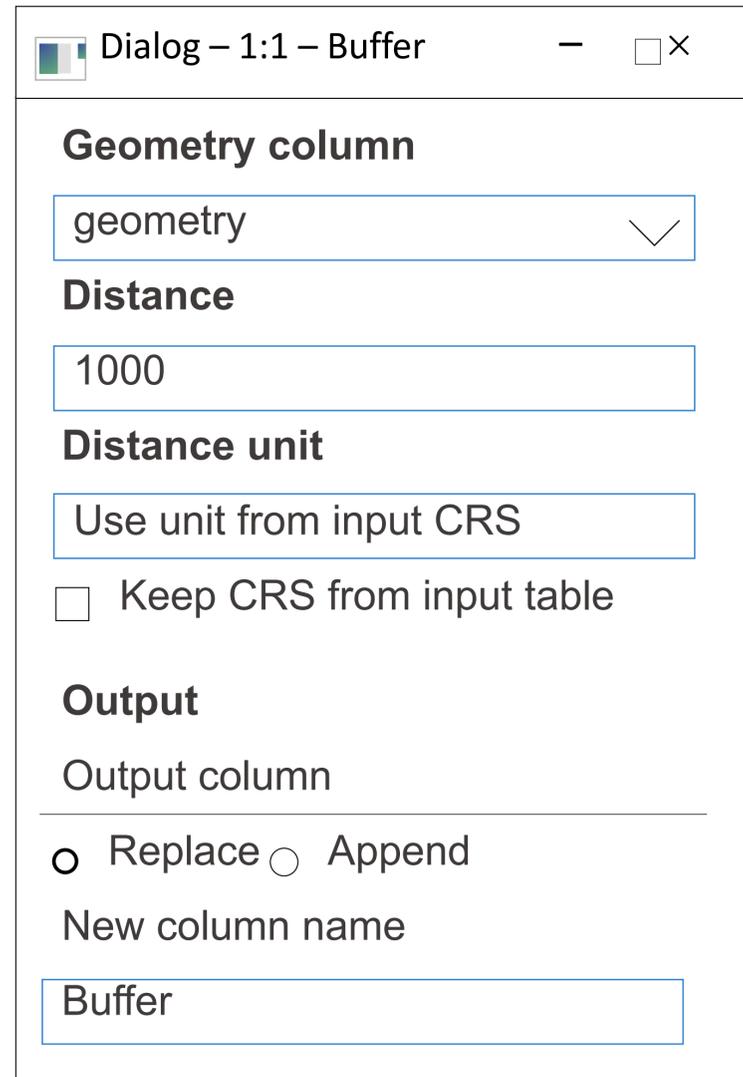
Dissolve – merge features based on attribute values

Together, these nodes support essential GIS operations such as data import, spatial transformation, measurement, aggregation, and visualization.

Spatial Manipulation – Buffer and Simplify

Buffer creates a zone around a geometry at a specified distance.

Simplify reduces the complexity of a geometry's shape while preserving its general form.



Dialog – 1:1 – Buffer

Geometry column
geometry

Distance
1000

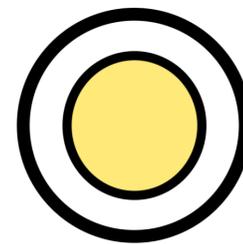
Distance unit
Use unit from input CRS

Keep CRS from input table

Output
Output column

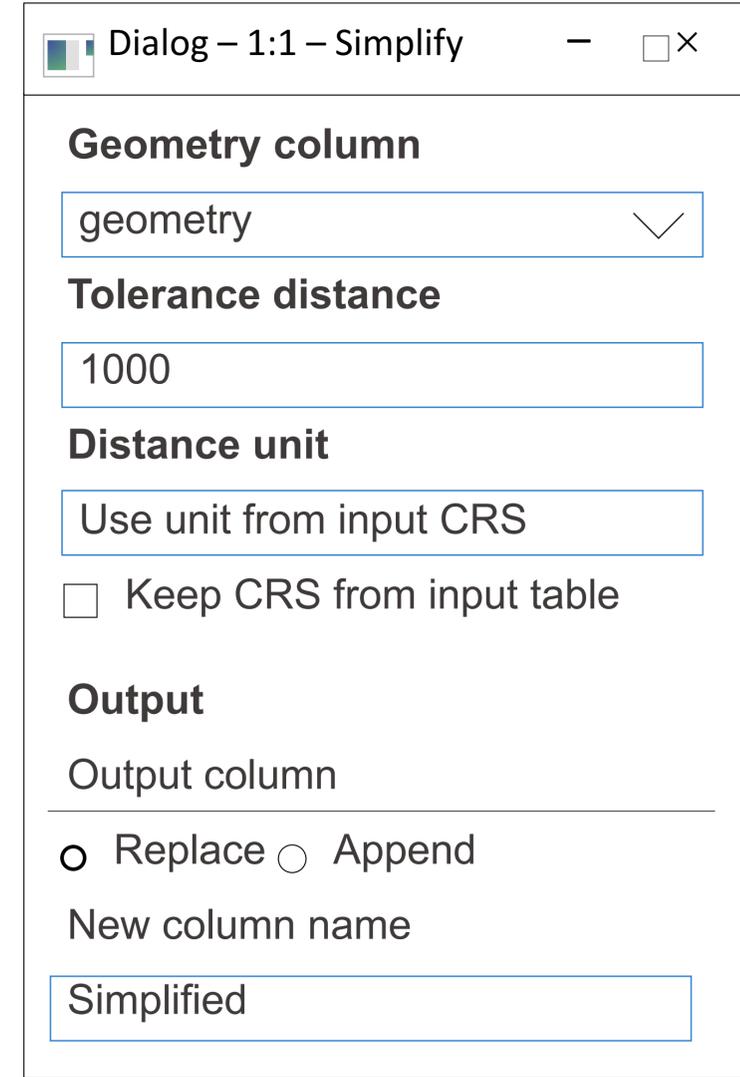
Replace Append

New column name
Buffer



Buffer

Use unit from input CRS
Meter
Kilometer
Miles
Degree



Dialog – 1:1 – Simplify

Geometry column
geometry

Tolerance distance
1000

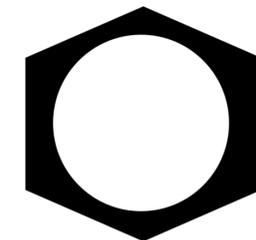
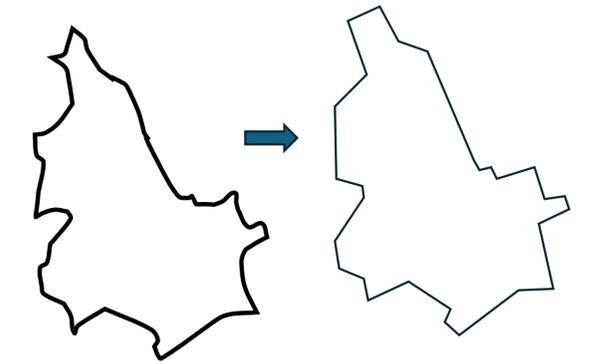
Distance unit
Use unit from input CRS

Keep CRS from input table

Output
Output column

Replace Append

New column name
Simplified



Simplify

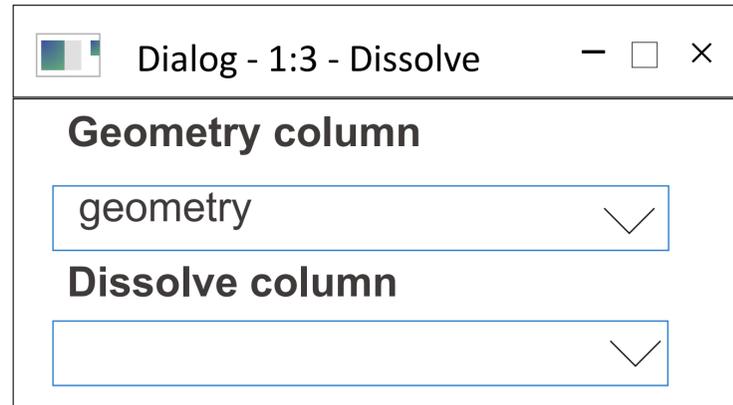
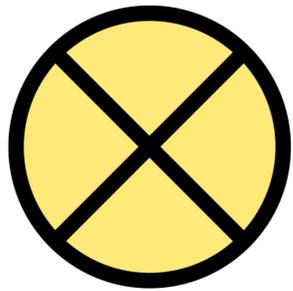
Spatial Manipulation – Dissolve and Explode

Dissolve merges adjacent or overlapping geometries into a single unified geometry, simplifying spatial boundaries.

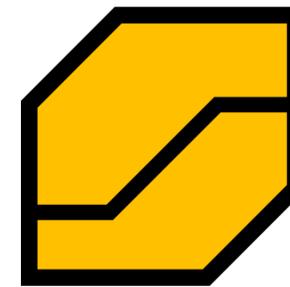
MultiPart To SinglePart splits multi-part geometries into separate single-part geometries, serving as the reverse function of Unary Union.

Unary Union merges all geometries into a single geometry row, combining spatial features into one unified shape.

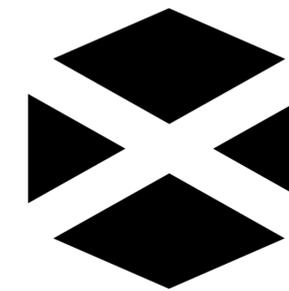
Dissolve



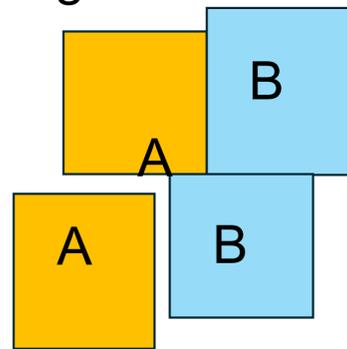
Unary Union



MultiPart To SinglePart

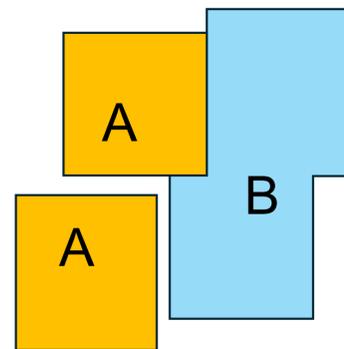


Original



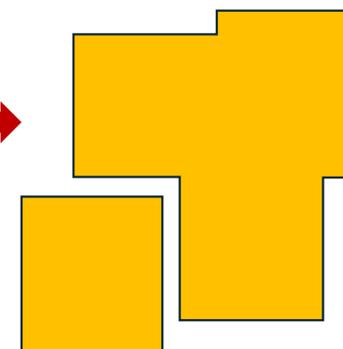
ID	Geometry
A	POLYGON
A	POLYGON
B	POLYGON
B	POLYGON

Dissolve



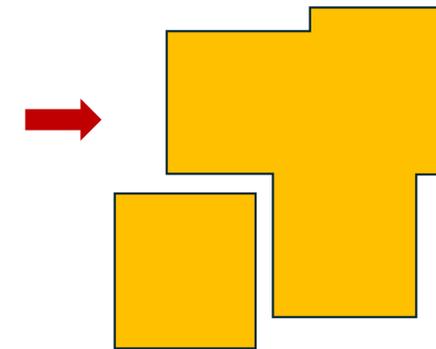
ID	Geometry
A	MULTIPOLYGON
B	POLYGON

Unary Union



Geometry
MULTIPOLYGON

MultiPart To SinglePart

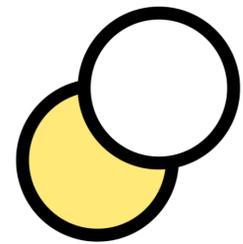


Geometry
MULTIPOLYGON
POLYGON

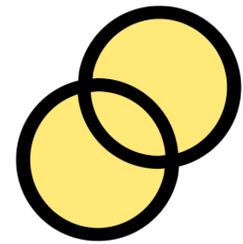
Spatial Manipulation – Overlay and Clip

Clip trims one geometry using the boundary of another, retaining only the overlapping area.

Overlay combines two geometry layers to create a new layer based on their spatial relationship, such as intersections or unions.



Clip



Overlay

Dialog – 1:1 – Clip

Target geometry column
geometry

Mask geometry column
geometry

Output

Output column

Replace Append

Dialog – 1:2 – Overlay

Left geometry column
geometry

Right geometry column
geometry

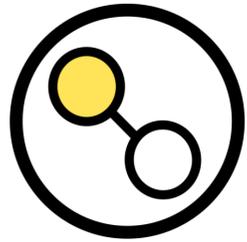
Overlay mode
intersection

Return only geometries of the same geometry type

Spatial Manipulation – Nearest Join and Spatial Join

Nearest Join links each feature to its closest neighbor in another geometry layer, adding data based on proximity.

Spatial Join merges two datasets based on their spatial relationship



Nearest Join



Spatial Join

Dialog - 1:1 - Spatial join

Top (left) geometry column
geometry

Bottom (right) geometry column
geometry

Join mode
 Inner Left Right

Match mode
Has its center in

[Show advanced settings](#)

[Show advanced settings](#)

Column selection

Top input (left table)
Manual Wildcard Regex

Search Aa

Excludes > >> < <<

Includes

Dialog -1:2 - Nearest join

Top (left) geometry column
geometry

Bottom (right) geometry column
geometry

Join mode
 Inner Left Right

Maximum distance
1000

Distance distance
Use unit from input CRS

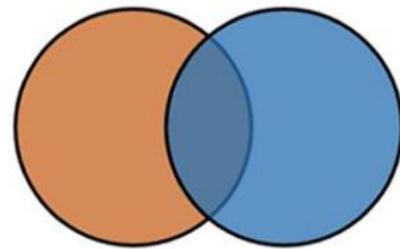
Keep CRS from left input table

[Show advanced settings](#)

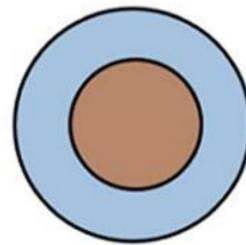
Spatial Relationship

Match Mode in Spatial Join

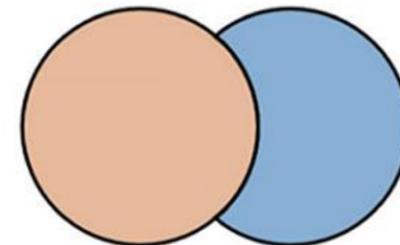
Contains
Contains the center of
Contains properly
Covers
Crosses
Has its center in
Intersects
Overlaps
Touches
Within



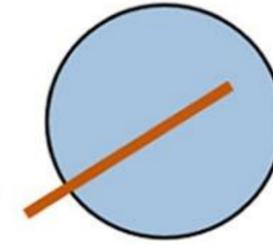
intersects



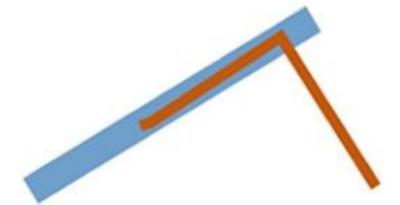
contains / within



touches



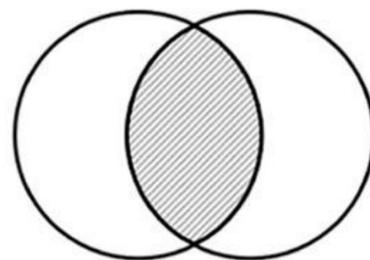
crosses



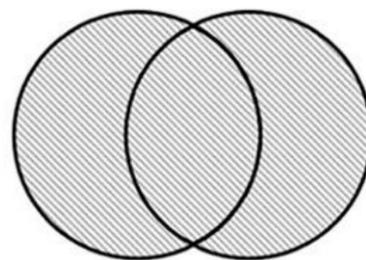
overlaps

Overlay Mode in Overlay

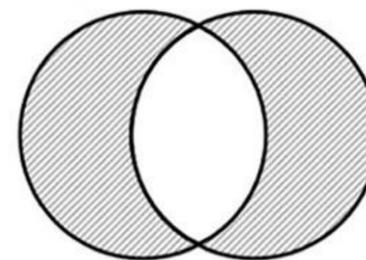
Intersection
Union
Identity
Symmetric difference
Difference



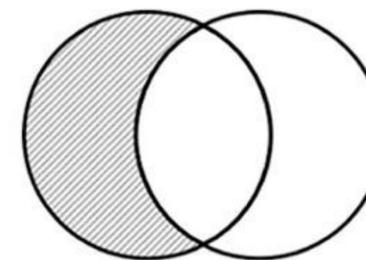
Intersection



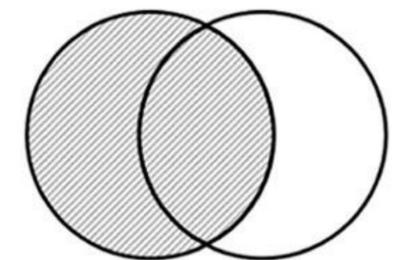
Union



Symmetrical Difference



Difference



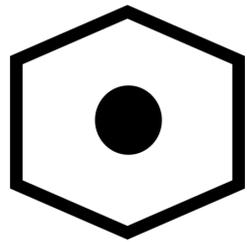
Identity

Geometry to Point, Create Grid

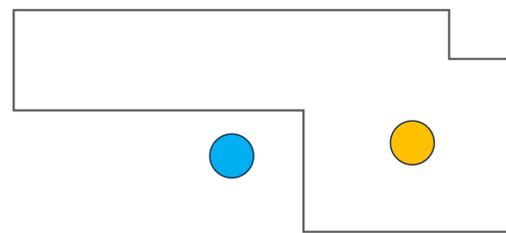
Geometry To Point converts geometries into centroids or representative points inside

Create Grid creates a rectangular grid overlay over a specified area, allowing for spatial sampling and uniform division of the space.

Create H3 Grid creates hexagonal grid inside the input polygon using H3.



Geometry To Point



-  centroid
-  representative point

Dialog - 1:1 - Geometry to Point

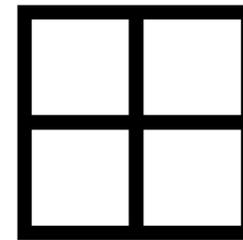
Geometry column
geometry

Point Type
 centroid representative_point

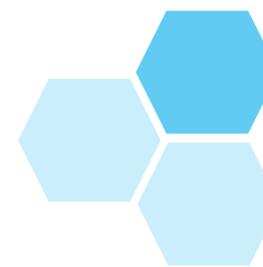
Output

Output column
 Replace Append

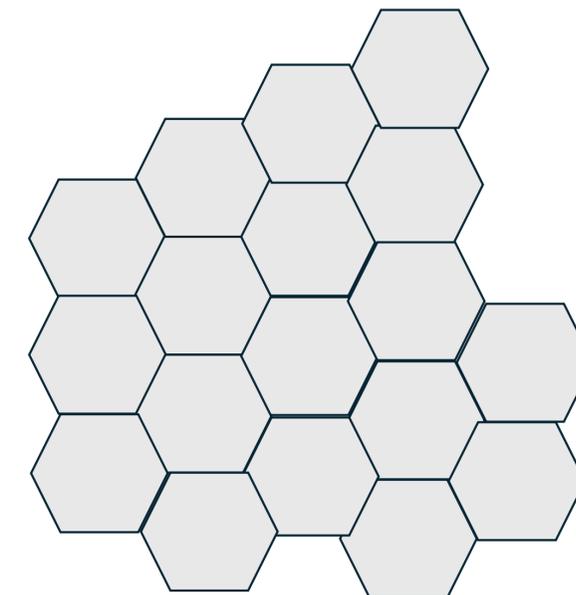
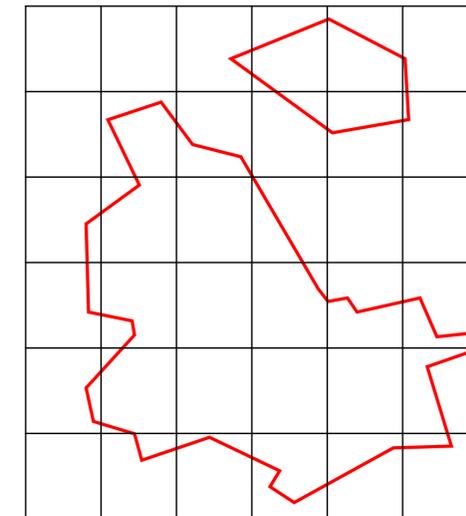
New column name
Point



Create Grid



Create H3 Grid



Geocoding

Geocoding: Converts an address into geographic coordinates.

Reverse Geocoding: Converts geographic coordinates back into a readable address.



Geocoding

1737 Cambridge ST,
Cambridge, MA 02138



Reverse
Geocoding



Latitude 42.375638
Longitude -71.1132356

Dialog - 1:1 - Reverse Geocoding

Geometry column

geometry

Geocoding Service Settings

Service provider

Nominatim

API key

Minimum delay (seconds)

1

Default timeout (seconds)

10

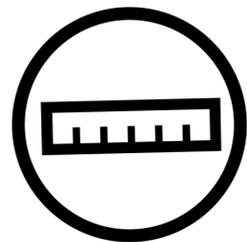
Append raw json

- arcgis: Using [ArcGIS online service](#) to do geocoding or reverse geocoding.
- azure: Using [AzureMaps geocoder based on TomTom](#) to do geocoding or reverse geocoding.
- baiduv3: Using [Baidu Map service](#) to do geocoding or reverse geocoding.
- bing: Using [Bing Map service](#) to do geocoding or reverse geocoding.
- googlev3: Using [Google Map service](#) to do geocoding or reverse geocoding.
- herev7: Using [HERE Geocoding & Search v7 API](#) to do geocoding or reverse geocoding.
- mapbox: Using [Mapbox service](#) to do geocoding or reverse geocoding.
- nominatim: Using [Nominatim service](#) to do geocoding or reverse geocoding.
- tomtom: Using [TomTom](#) to do geocoding or reverse geocoding.
- yandex: Using [Yandex service](#) to do geocoding or reverse geocoding.

Measure Distance and Routing

Euclidean Distance calculates the straight-line distance between points or features, based on Cartesian coordinates.

OSRM Distance Matrix ndoe: provides free distance calculations with Open Source Routing Machine (OSRM), but it has speed limitations.



Euclidean
Distance

Dialog - 1:1 - Euclidean Distance

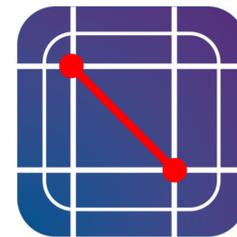
Origin geometry column
geometry

Origin ID column

Destination geometry column
geometry

Destination ID column

Distance unit
Use unit from input CRS



OSRM Distance Matrix

Dialog - 1:1 - OSRM Distance Matrix

Travel mode
 Route Travel cost Travel cost and route

Minimum delay (seconds)
1

Default timeout (seconds)
10

OSRM server
<https://router.project-osrm.org>

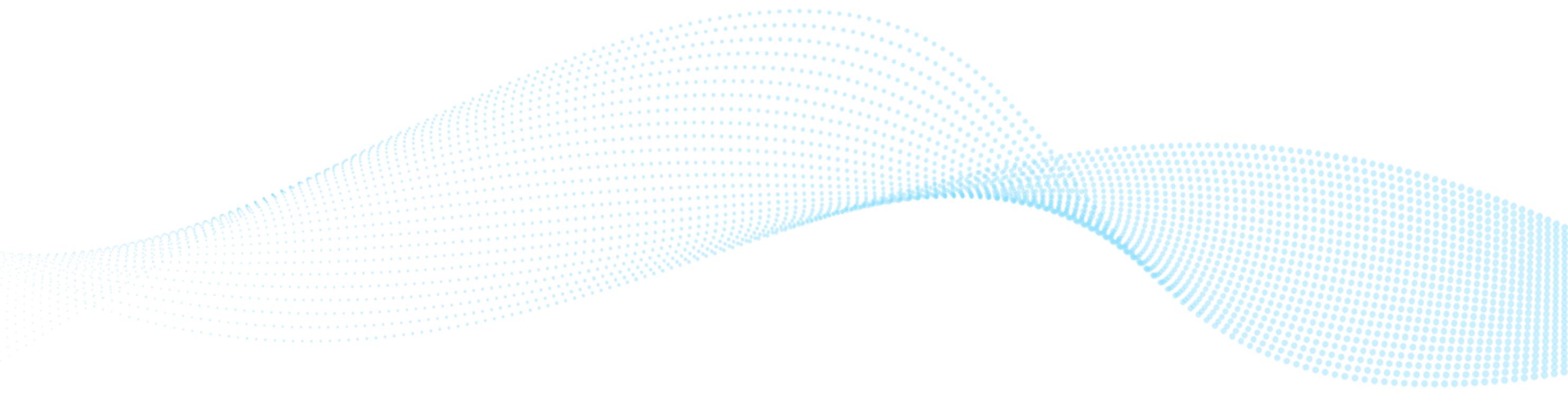
Hands-on 2

Visual Programming for GIS

KNIME Analytics Platform



Center for Geographic Analysis
Institute for Quantitative Social Science
Harvard University

A decorative graphic consisting of a series of light blue dots arranged in a wavy, horizontal pattern that spans across the middle of the page. The dots are more densely packed in some areas, creating a sense of depth and movement.

5

Find Near Facilities

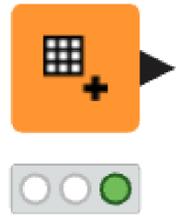


Visualize restaurants (other facilities)
within a 1,000-meter buffer
around a given location.

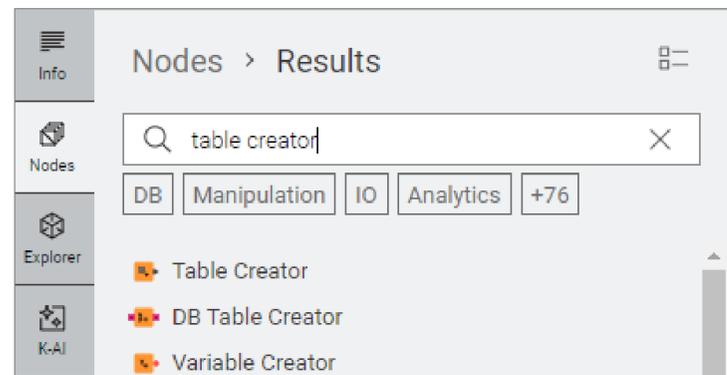
Create data like using Excel

Create a new project named **AdvancedGIS**. Within the project, add a **Table Creator** node, input “1737 Cambridge St, Cambridge, MA 02138, USA” and configure it according to the following steps.

Table Creator



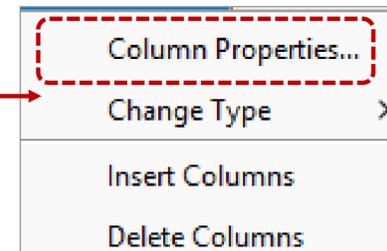
1- Search and Add **Table Creator** Node



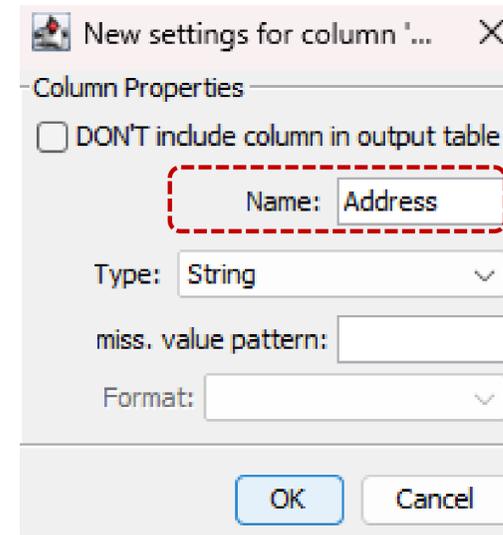
This node dialog is not supported here.

Open dialog

2 Open dialog

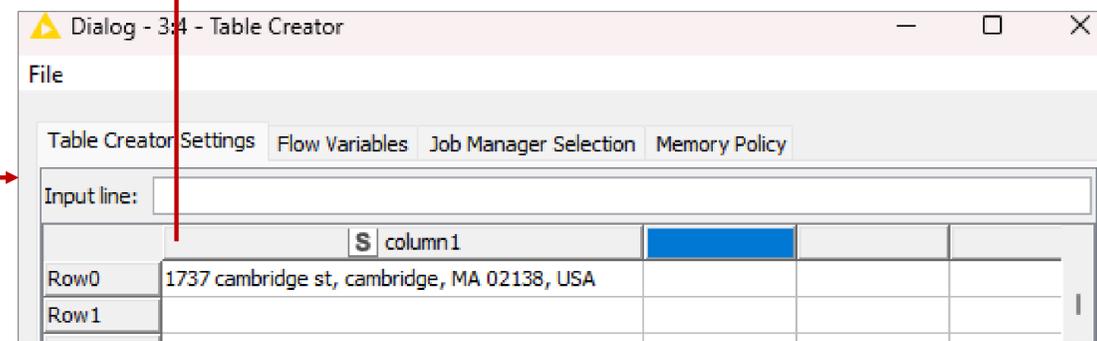


4 Right click **column 1** and choose **Column Properties**

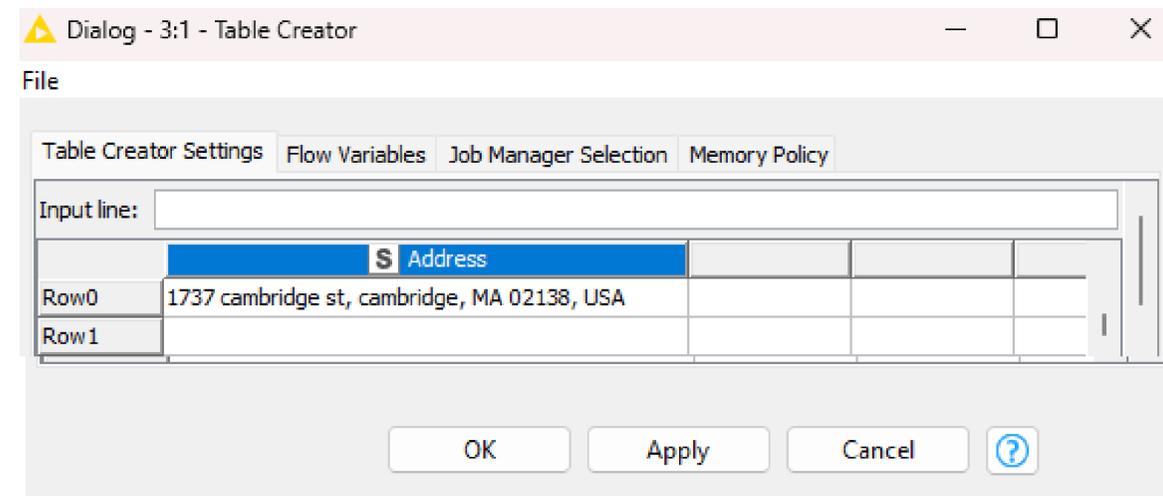


5 Change Name to **Address**

6 Click OK and run the node



3 Input text in the first empty cell

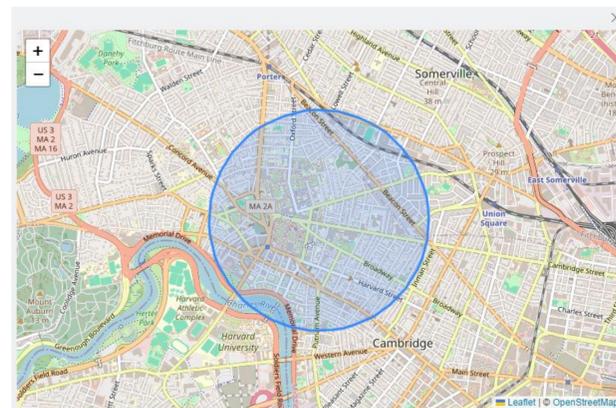
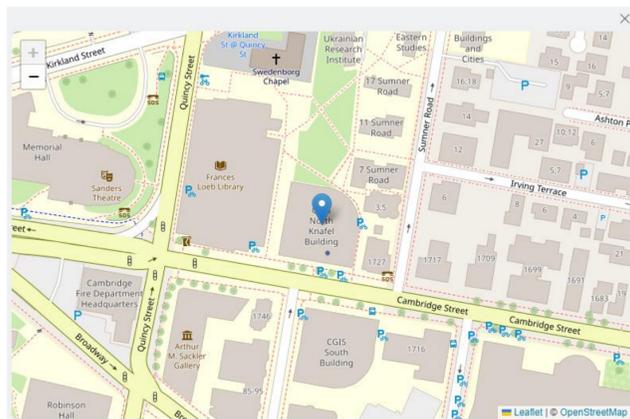
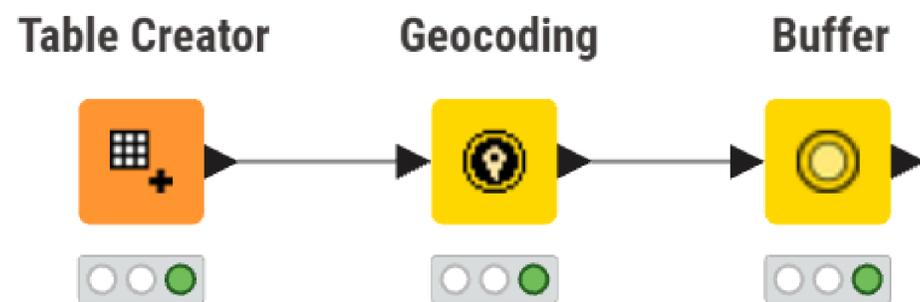


Convert Address to Spatial Point and Create Buffer Zone

Add the **Geocoding** and **Buffer** nodes to the workflow and connect them to the **Table Creator** node.

In the **Geocoding** node: Select **Address** as the address column.

In the **Buffer** node: Set **Distance** to **1000**, Choose **Meter** as the distance unit, Check **Keep CRS from input table**, Check **Replace** to keep only the generated buffer geometry (the 1,000-meter circle). Then execute the nodes to create the buffer area.



Geocoding

Address column
Address

Geocoding Service Settings

Service provider
nominatim

API key

Minimum delay (seconds)
1

Default timeout (seconds)
10

[Show advanced settings](#)

Buffer

Geometry column
geometry

Distance
1000

Distance unit
Meter

Keep CRS from input table

Output

Output column
 Replace Append

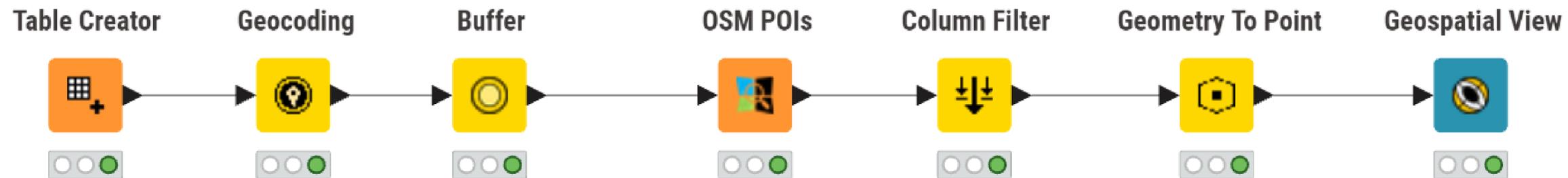
Convert Address to Spatial Point and Create Buffer Zone

Add the **OSM POIs** node to query restaurants within the buffer area using the default settings.

In the **Column Filter** node, keep **geometry**, **cuisine**, and **name** in the Includes box.

Use **Geometry to Point** to convert polygons into points by selecting **Centroid** and checking **Replace**.

Add a **Geospatial View** node to visualize the points. Set **Marker Size Column** to **None**, enter **20** for **Size Scale**, choose **cuisine** as the **Color Column**, and execute the node.



OSM POIs

Geometry column

geometry

Input place tags

amenity

Input value tags

restaurant

Column Filter

Column filter

Manual Wildcard Regex Type

Search Aa

Excludes

- air_conditioni...
- amenity
- diet:vegan
- diet:vegetarian
- opening_hours
- payment:cred...
- payment:debi...
- phone
- smoking

Includes

- geometry
- cuisine
- name

Any unknown column

Geometry To Point

Geometry column

geometry

Point Type Selection

centroid

Output

Output column

Replace Append

Geospatial view

Size Settings

Marker size column

None

Marker size scale

20

Coloring Settings

Marker color column

cuisine

Color map

viridis

Classify numerical marker color columns ...

KEY TAKEAWAYS

KNIME supports end-to-end GIS workflows: **data IO → spatial processing → spatial analysis → visualization.**

Combining **core nodes** (Filter / Join / GroupBy) with **geospatial nodes** (Projection / Buffer / Spatial Join / View) is the fastest way to build practical spatial analytics.

A clean workflow structure (branching + components) makes your analysis **reproducible, reusable, and easy to share.**

Resources

Visual Programming for GIS

KNIME Analytics Platform



Center for Geographic Analysis
Institute for Quantitative Social Science
Harvard University

KNIME Hub and GitHub Repository



Search workflows, nodes and more...

Sign in

KNIME Community Hub > Center for Geographic Analysis at Harvard University



Center for Geographic Analysis at Harvard University

Team

Spaces

Extensions

Spaces of Center for Geographic Ana... A E K P +3

Computational Methods and GIS Applications in Social Science- KNIME Lab Manual

This space contains workflows that accompany the book: Computational...

Last edited Dec 6, 2022

32 0 78

Geospatial Analytics Examples

Explorer this space for workflows and components as blueprints and building...

Last edited Nov 28, 2022

42 0 5

README Contributing MIT license

Geospatial Analytics Extension for KNIME

This repository is the home of the [Geospatial Analytics Extension](#) for [KNIME Analytics Platform](#). The extension provides a set of nodes for spatial data analysis and visualization.

The extension is developed by the [Center for Geographic Analysis](#) at [Harvard University](#) and [KNIME](#) as part of a two-year project of the [Spatiotemporal Innovation Center](#). The goal of the collaboration is to develop KNIME Analytics Platform extensions and best-practice workflows to provide a consistent and compatible platform for spatial data analysis across disciplines.

The extension is mainly based on the [GeoPandas](#) library and the [PySAL](#) library. The data types are defined in the [KNIME - Geospatial Data Types](#) extension.

<https://github.com/spatial-data-lab/knime-geospatial-extension>

<https://hub.knime.com/center-for-geographic-analysis-at-harvard-university>

<https://hub.knime.com/knime/collections/KNIME%20for%20Geospatial%20Analysis~kc8Kc8MDDqG-sqoF>

Cheat sheet, Certificate, KNIME it Challenges

Cheat Sheet: Geospatial Analytics Extension for KNIME Analytics Platform

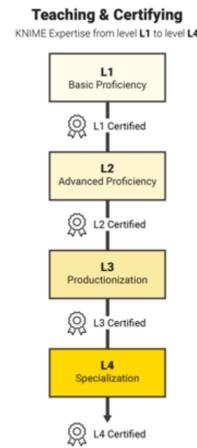
KNIME Certification Exams

- From level **L1** to level **L4**
- Online examination
- Measure your expertise with KNIME software and data skills
- Get and share your badge!



Get certified at:

knime.com/certification-program



Just KNIME It! Challenges

- The best way to keep on learning
- Weekly challenges to test your knowledge
- Easy, medium and hard challenges for any level
- Discuss the solution with the community
- Post your solution and climb the [Leaderboard](#)



Find the **challenges** at:

knime.com/just-knime-it

SPATIAL DATA I/O

GeoFile Reader: Reads single-layer spatial data from a local file path or a URL and supports various file formats such as Shapefile (.shp), zipped Shapefile (.zip), GeoJSON (.geojson), GeoPackage (.gpkg), and more.

GeoPackage Reader: Reads multi-layer spatial data from a local file path or a URL and primarily supports GeoPackage (.gpkg) and GeoDatabase (.GDB) formats but can also read zip files compressed from a GeoDatabase folder. The layer to read can be specified in the node configuration.

GeoFile Writer: Writes spatial data and its ancillary data to a local file, supporting Shapefile (.shp), GeoJSON (.geojson), and GeoParquet (.parquet) formats.

GeoPackage Writer: Writes spatial data to a new file or as a layer into an existing file and supports GeoPackage (.gpkg) and GeoDatabase (.GDB) formats.

OPEN DATASETS

OSM Boundary Map: Retrieves place boundaries from *OpenStreetMap* by geocoding the place name. The query must match the locations in the *Nominatim* database.

OSM Road Network: Retrieves a geospatial network and its associated attributes from *OpenStreetMap*, including networks for driving, cycling, walking, etc. Similarly, the **OSM POIs node** can be used to retrieve geometries and attributes for geospatial entities.

US ACS 5-Year Estimates: Retrieves American Community Survey 5-Year Data (2009-2020). The ACS covers a broad range of topics about social, economic, demographic, and housing characteristics of the U.S. population.

US2020 Census Data: Retrieves US 2020 Census Redistricting Data (Decennial Census P.L. 94-171 Redistricting Data) which provides variables such as population and household information.

US2020 TIGER Map: Retrieves the specific geospatial boundaries for one specific state of the United States. The popular TIGER/Line levels are Block group, Roads, Blocks, Tracts.

SPATIAL MANIPULATION

Buffer: Generates a buffer with the given distance for each geometric object.

Clip: Clip target geometries to the mask extent. The geometries will be clipped to the full extent of the clip object.

Spatial Join: Merge the left (top) and the right (bottom) table based on their spatial relationship of the two selected columns.

Nearest Join: Merge the left (top) and the right (bottom) table based on the distance between their geometries of the two selected columns.

Euclidean Distance: Calculates the Euclidean distance in the selected unit between each origin and destination path. To take into account the curvature of the earth, use the **Haversine Distance** node.

Voronoi (Thiessen) Polygons: Creates Voronoi (Thiessen) polygons from the input point data according to the reference boundary.

Overlay: Perform spatial overlay between two geometries based on various modes such as Union, Intersection, etc.

Dissolve: Aggregate geometries based on group id and keeps only the two columns group id and dissolved geometries.

SPATIAL TRANSFORMATION

Polygon To Line: Returns the boundaries of each polygon.

Line To MultiPoint: Returns points from the lines.

Geometry To Point: Returns centroids or representative points guaranteed to be within each geometry.

Points To Line: Generates lines from points according to group id and serial label.

Multipart To Single Part: Explodes multipart geometries into multiple single geometries.

Create Random Points: Generates random points from a uniform distribution within/along each input geometry. It can be used for polygons and lines, other geometry types are ignored.

Projection: Transforms the Coordinate Reference System (CRS) of a geometry column into a new CRS. It can be used to unify the rows with various CRS.

SPATIAL CONVERSION

Geocoding: Retrieves spatial location points based on the given addresses or vice versa with the **Reverse Geocoding** node.

WKT to Geometry: Convert the **Well-known Text (WKT)** column into a geometry column, or vice versa with the **Geometry to WKT** node.

GeoJSON to Geometry: Convert the **GeoJSON** column into a geometry column, or vice versa with the **Geometry to GeoJSON** node.

Lat/Lon to Geometry: Convert the given **latitude and longitude** column into a point geometry column, or vice versa with the **Geometry to Lat/Lon** node.

IP To Geometry: Retrieves the spatial location of the provided IP addresses by choosing the desired service provider and API key (optional).

Geometry To Metadata: Extracts metadata for each geometry within the selected column. It includes the CRS, the geometry type, and a flag indicating the presence of z-coordinates.

SPATIAL CALCULATION

Area: Calculate the geometric attributes - **Area, Length and Bounds** - for each spatial object based on the object's coordinate system.

Bounding Box: Generate new geospatial objects through calculations. **Bounding Box** and **Convex Hull** create minimum bounding geometries for each row to encapsulate the object using a box or polygon.

Unary Union: Merge all objects into a single geometry row.

SPATIAL VISUALIZATION

Geospatial View: Creates an interactive map view based on the geometric elements of the input table. It supports several modifications such as changing the base map or shape color and size.

Geospatial View Static: Creates a static visualization of the geometric elements and creates an image at its output port (supports .svg or .png files). It is useful for creating Choropleth Maps and supports various settings to customize the legend.

Kepler.gl Geoview: Creates an interactive map view using the *kepler.gl* framework. It allows changing various aspects of the view within the visualization, e.g., adding layers and filters.

Spatial Heatmap: Visualizes spatial data on an interactive heatmap using a weight column to represent the intensity at each element.

EXPLORATORY SPATIAL DATA ANALYSIS

Spatial Weights: Constructs a contiguity spatial weights matrix, considering various spatial relations such as Queen, Rook, Binary, Distance, etc., and allows the upload of a custom matrix.

Global Spatial Statistics (Spatial weight for second input port):

- Global Geary's G:** Measures overall spatial autocorrelation, examining whether values tend to cluster or disperse globally. Valuable for understanding how similar/dissimilar values are.
- Global Getis-Ord G:** Measures spatial clustering patterns on a global scale, helping to detect areas of concentration or dispersion across the study area.
- Global Moran's I:** Measures overall spatial autocorrelation within a dataset, providing a single value that summarizes the extent to which nearby data points exhibit similar or dissimilar characteristics.

Local Spatial Statistics (Spatial weight for second input port):

- Local Getis-Ord G:** Identifies statistically significant hot spots and cold spots using the Getis-Ord G* statistic based on a set of weighted features.
- Local Moran's I:** Identifies spatial clusters of features with high or low values, and spatial outliers using Moran's I statistics.

SPATIAL NETWORK

Google Distance Matrix: Uses the *Google Distance Matrix* API to create a distance matrix for the provided origins and destinations. It returns the travel distance (in meter) and duration (in minutes).

OSRM Distance Matrix: Uses the *Open Source Routing Machine (OSRM)* to create a distance matrix for the provided origins and destinations. It returns the driving travel distance and time as well as the route.

Road Network Distance Matrix: Creates a distance matrix for the provided origins and destinations using the given road network. It snaps each pair to the closest point of the network and returns the shortest path.

Road Network Isochrone Map: Calculates the *isochrone map* for the input point based on the given road network and its travel cost column. It snaps the input points to the road network and returns weighted time or distance.

SPATIAL MODELLING

OLS with Spatial Test: Performs Ordinary Least Squares (OLS) regression analysis and offers spatial statistical tests to assess model validity and spatial relationships.

2SLS with Spatial Test: Performs Spatial Two-Stage Least Squares (2SLS) analysis, providing results and diagnostics, while also performing spatial statistical tests.

Spatial Lag Model: Performs Spatial Lag Model analysis, which examines how variables are influenced by spatially neighboring observations, providing insights into spatial dependencies.

Spatial Error Model: Performs Spatial Error Model analysis, accounting for spatial autocorrelation and addressing potential biases in regression analysis.

GWR Model: Performs Geographically Weighted Regression (GWR), allowing for the exploration of spatially varying relationships in the data, considering spatial heterogeneity.

MGWR Model: Performs Multiscale Geographically Weighted Regression (MGWR) analysis, a technique that examines spatial relationships at various scales, providing insights into spatial variations.

Reference

The 1.2 release of the Geospatial Extension which now supports the experiments detailed in the first KNIME workbook for GIS: *"Computational Methods and GIS Applications in Social Science - Lab Manual"* by Lingbo Liu and Fahui Wang. For a deeper understanding and broader applications, users are encouraged to refer to the main text *"Computational Methods and GIS Applications in Social Science, 3rd Edition"* authored by Fahui Wang and Lingbo Liu. The workbooks for the Lab Manual are available in the *Geospatial Space* in the *KNIME Hub*.

Resources

- E-Books:** KNIME Advanced Luck covers advanced features & more. Practicing Data Science is a collection of data science case studies from past projects. Both available at knime.com/knimepress
- KNIME Blog:** Engaging topics, challenges, industry news, & knowledge nuggets at knime.com/blog
- Learning Courses:** Take our free online self-paced courses to learn about the different steps in a data science project (with exercises & solutions to test your knowledge) at knime.com/knime-self-paced-courses
- KNIME Community Hub:** Browse and share workflows, nodes, and components. Add ratings, or comments to other workflows at hub.knime.com
- KNIME Forum:** Join our global community & engage in conversations at forum.knime.com
- KNIME Business Hub:** For team-based collaboration, automation, management, & deployment check out KNIME Business Hub at knime.com/knime-business-hub

References

Lingbo Liu, Fahui Wang, Computational Methods and GIS Applications in Social Science - Lab Manual, CRC, 2023 , <https://doi.org/10.1201/9781003304357>

Fahui Wang, **Lingbo Liu**, Computational Methods and GIS Applications in Social Science ,CRC, 2023. <https://doi.org/10.1201/9781003292302>

Kedron, P., Li, Z., **Liu, L.** (2026). Evaluating the Impact of Open Data + FAIR Policies on Computational Reproducibility: A Systematic Analysis of IJGIS Publications, International Journal of Geographical Information Science, <https://doi.org/10.1080/13658816.2025.2603586>

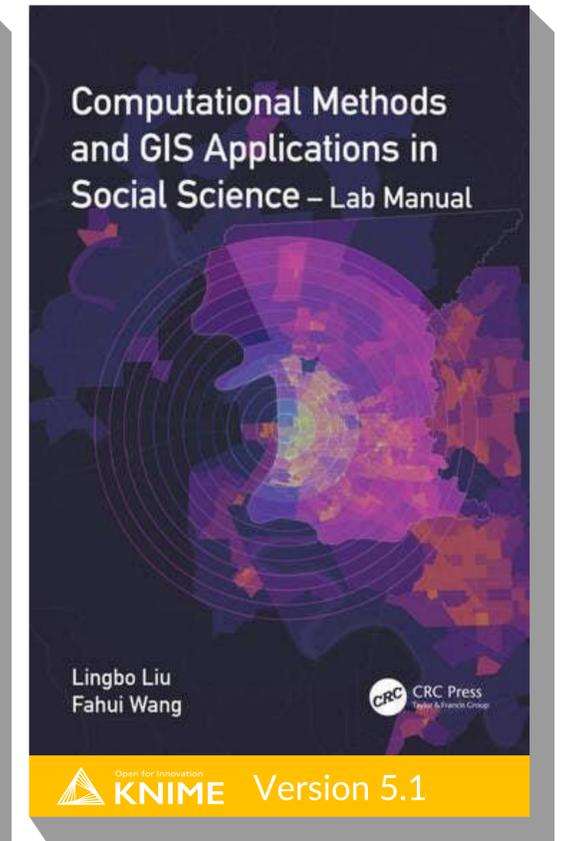
Liu, L., Guan, W. W., Wang, F., & Bao, S. (2025). Visual programming-based Geospatial Cyberinfrastructure for open-source GIS education 3.0. Cartography and Geographic Information Science, 1-13. <https://doi.org/10.1080/15230406.2025.2462342>

Liu, L.; Wang, F., Fu, X., Kötter, T., Sturm, K., Guan, W. W., & Bao, S. (2024). Elevating the RRE Framework for Geospatial Analysis with Visual Programming Platforms: An Exploration with Geospatial Analytics Extension for KNIME. International Journal of Applied Earth Observation and Geoinformation, 130, 103948. <https://doi.org/10.1016/j.jag.2024.103948>

Liu, L.; Fu, X.; Kötter, T.; Sturm, K.; Haubold, C.; Guan, W. W.; Bao, S. & Wang, F.. 2024. Geospatial Analytics Extension for KNIME. SoftwareX, 25, 101627. <https://doi.org/10.1016/j.softx.2023.101627>



Publish Date 8/15/2023



Publish Date 10/15/2023

YouTube Video

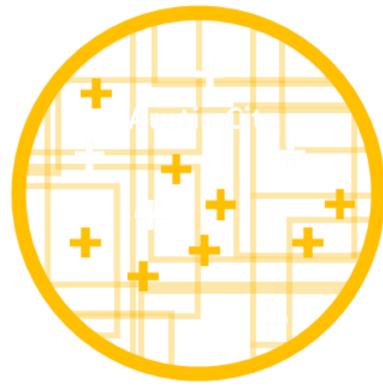
3:10



Case 1

Identification
of Nearby
Facilities

3:38



Case 2

Site Selection
by Accessibility
for Vaccination

3:36



Case 3

The Impact
Assessment of
Hurricane Ian

3:30



Case 4

Demographic Analysis
with the US and China
2020 County Census
Data

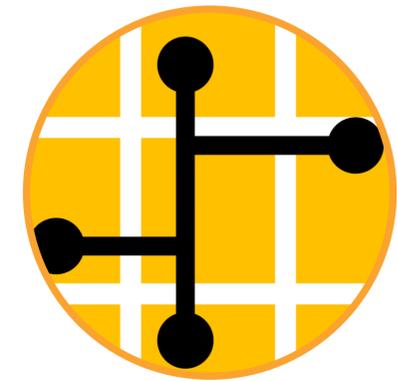
4:08



Case 5

Building an OSM
POI Heat Map
Dashboard

3:38



Case 6

New Spatial
Network Nodes in
GA 1.1

<https://www.youtube.com/playlist?list=PLnFUy1r9kH-20dWQGVKKiUAOIbPGxyBUv>

<https://www.youtube.com/@liulingbo2151>



Next Workshop on GEE



KNIME Analytics Platform

Google Earth Engine Extension

No-Code, Open Visual Programming Platform for Cloud-Based Remote Sensing

Satellite Imagery Analysis with Google Earth Engine

 **Date and Time**

April 3, 2026

01:00PM - 03:00PM EDT

 **Location**

1737 Cambridge St. Cambridge, MA. CGIS Knafel Building. Room SB-12 in the sub-basement.

Take the elevator to the sub-basement level.

Developed by Harvard Center for Geographic Analysis
Supported by NSF IUCRC Project and KNIME



Acknowledgement

