

Distance based analysis

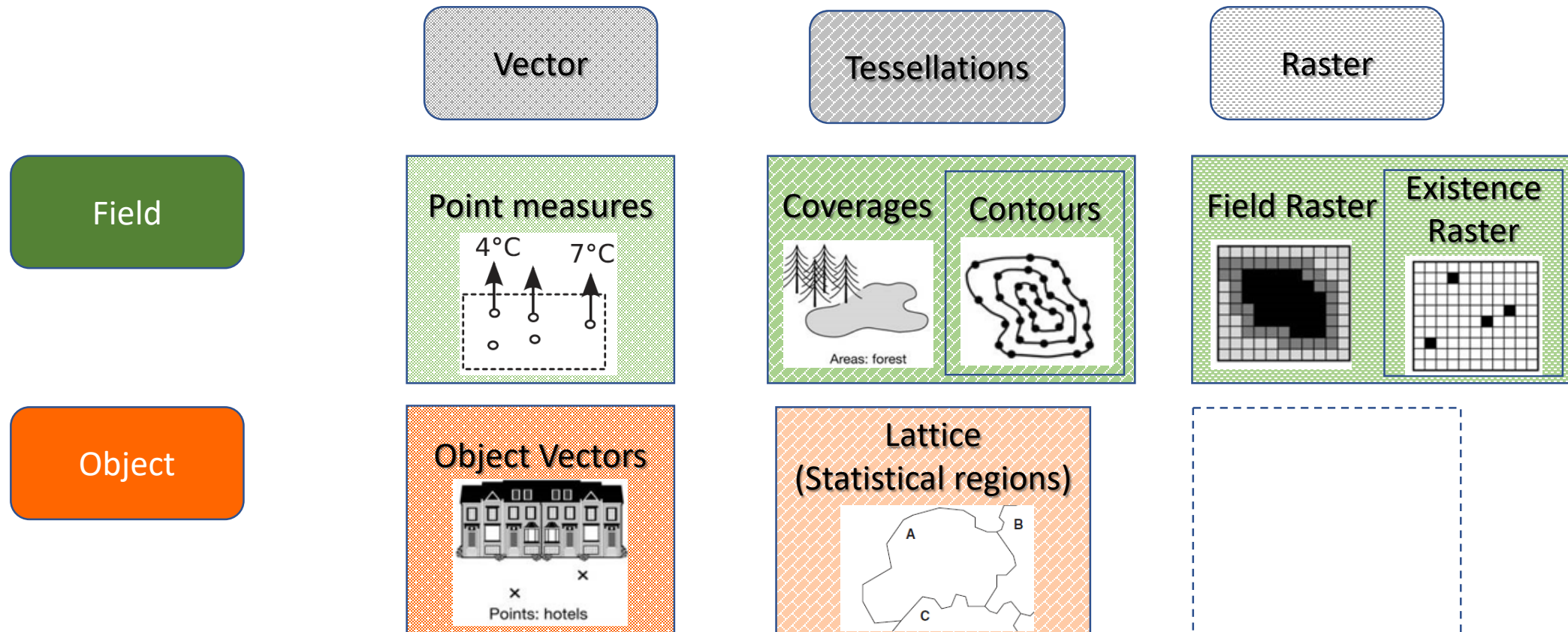
Spatial Data Analysis and Simulation modelling,
2020, Simon Scheider



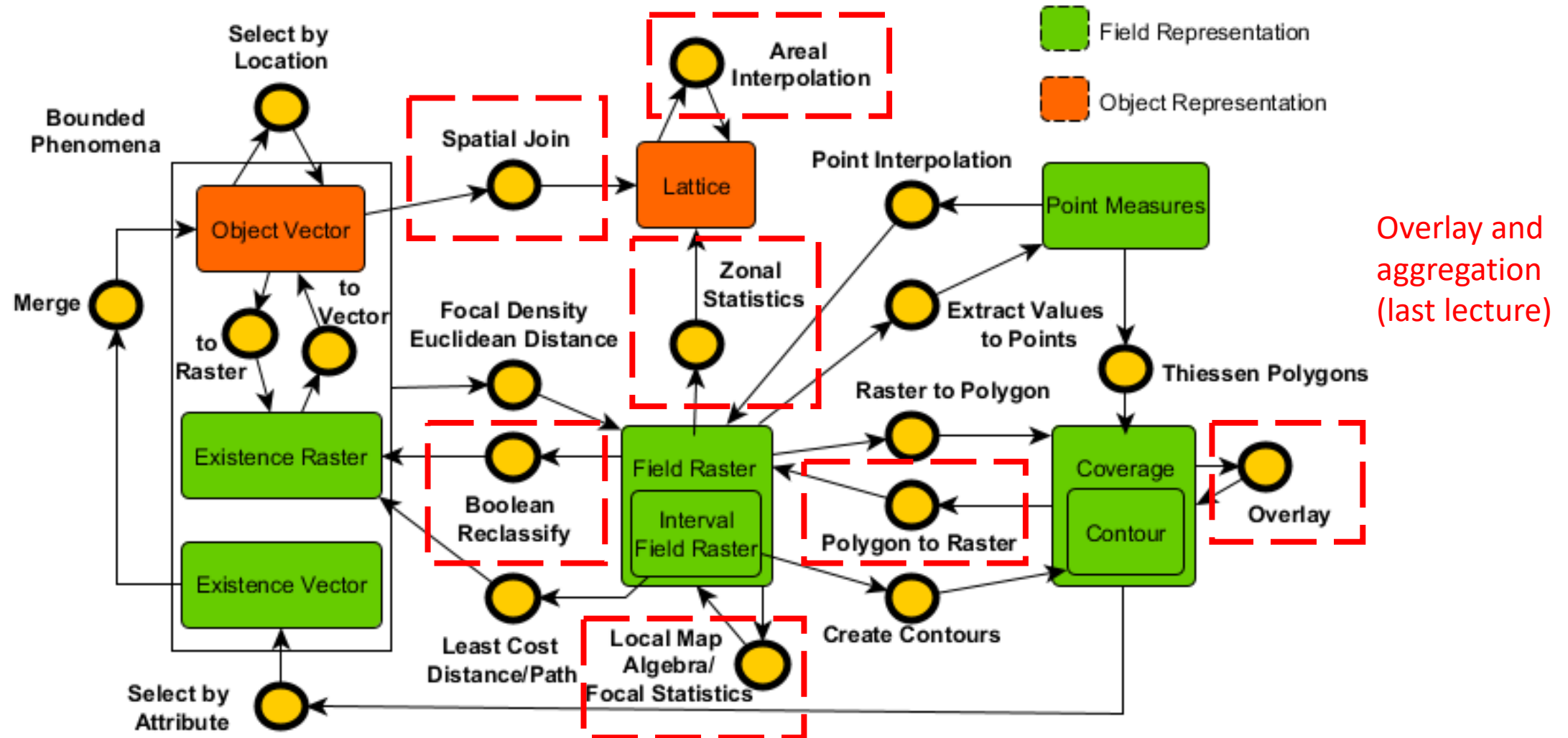
Outline

- Core concepts (recap)
- Vector Distance-Based Analysis:
 - Planar vs geodesic distances
 - Buffers
 - Nearest objects
 - Thiessen Polygons
- Raster Distance-Based Analysis:
 - Focal map algebra
 - Global map algebra (Proximity analysis)
 - Point Interpolation
- Core concept quiz

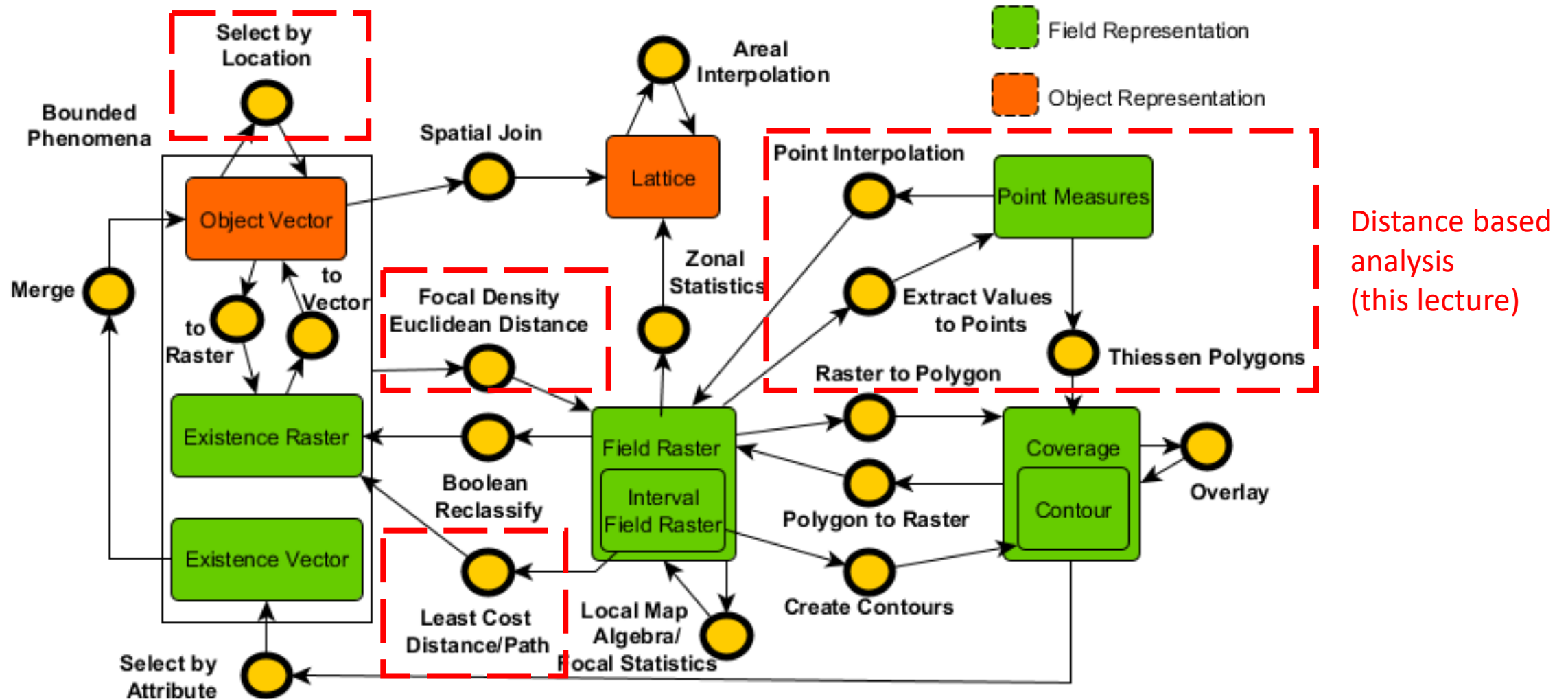
How geodata models represent core concepts



In which ways can geodata be transformed?



In which ways can geodata be transformed?

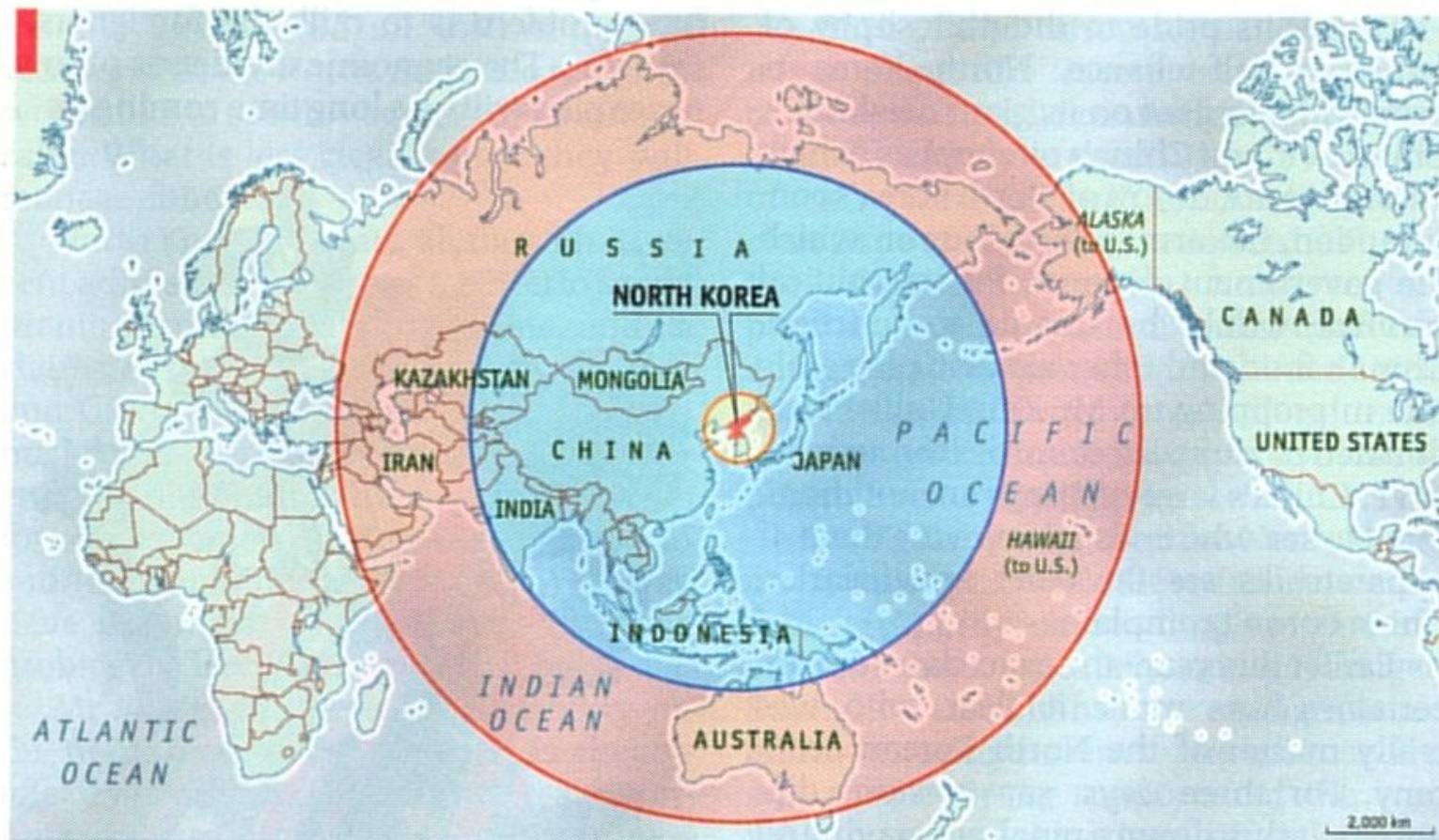


Core concept quiz




In order to make the workflows that you implement in the lab reusable on different data layers, think for a moment how they correspond to core concept transformations:

- 1) which core concepts correspond to which datasets in the lab?
- 2) which path in the computational diagram corresponds to which one of your workflows?

Vector Distance Analysis



North Korea's missile threat

Type	Maximum range	Payload	Status
Nodong	 1,300 km (810 miles)	700 kg (1,550 pounds)	Currently deployed
Taepodong-1	 Up to 10,000 km	Several hundred kg	Test failed 1998, not yet operational
Taepodong-2	 10,000-15,000 km	Several hundred kg	Not yet tested

Source: Task Force for US Korea Policy, Centre for International Policy




The Economist, May 3, 2003

correalis effect
- spin of the earth
changes range of missile

Flat-earth thinking. Thank you to those readers who pointed out that, by superimposing concentric circles on a Mercator projection, the map in our May 3rd issue greatly underestimated the potential reach of North Korea's missiles. We stand corrected.



North Korea's missile threat

Type	Maximum range	Payload	Status
Nodong	 1,300 km (810 miles)	700 kg (1,550 pounds)	Currently deployed
Taepodong-1	 Up to 10,000 km	Several hundred kg	Test failed 1998, not yet operational
Taepodong-2	 10,000-15,000 km	Several hundred kg	Not yet tested

Source: Task Force for US Korea Policy, Centre for International Policy

do we ever take this into account or is it always just the distance (airplanes etc.)

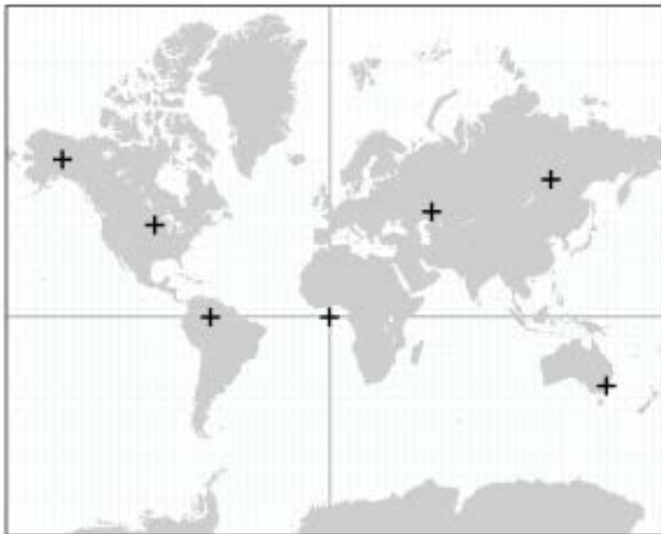
The Economist, May 17, 2003

Planar vs. Geodesic Distances

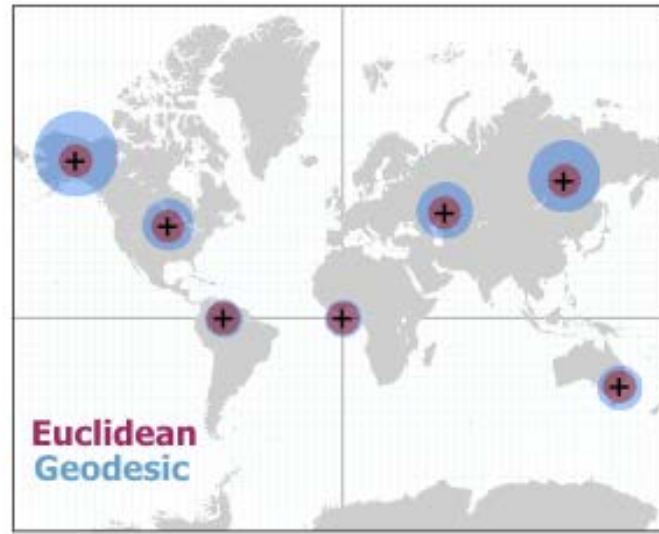
projected CRS adapts for the different distances

- Planar distances are measured in a projected coordinate system (CRS)
-> thus they are distorted on the (spherical) earth surface
- Geodesic distances are measured “on” a Spheroid model of the Earth

World - Mercator



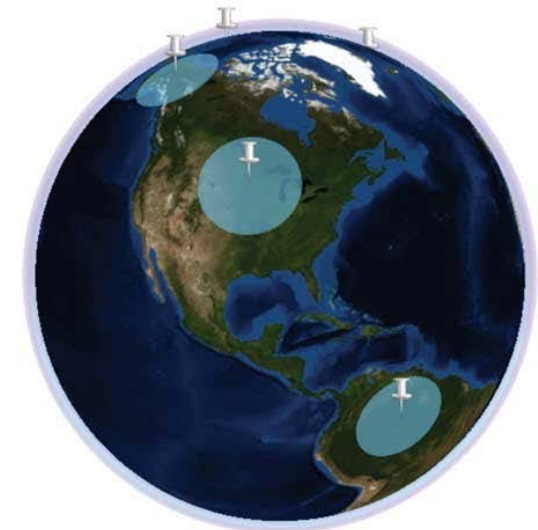
1,000 km Buffer



1,000 km Buffer



Planar



Geodesic

Proximity I: Buffering

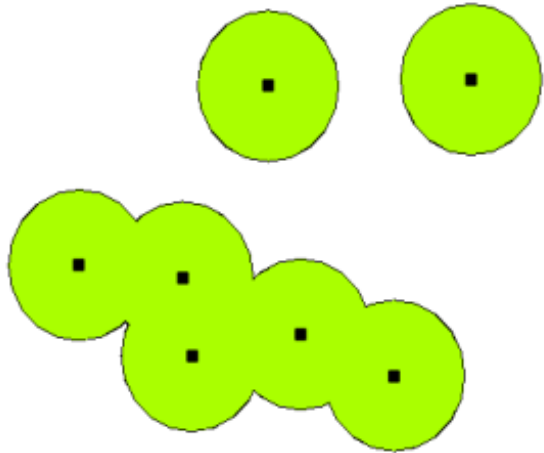
split space in area within or outside buffer

- **Buffering** usually creates two areas: one area that is **within** a specified distance to selected real world features and the other area that is **beyond**. The area that is within the specified distance is called the **buffer zone**.
- A **buffer zone** is any area that serves the purpose of keeping real world features distant from one another.
- Common types of buffer zones may be greenbelts between residential and commercial areas, border zones between countries

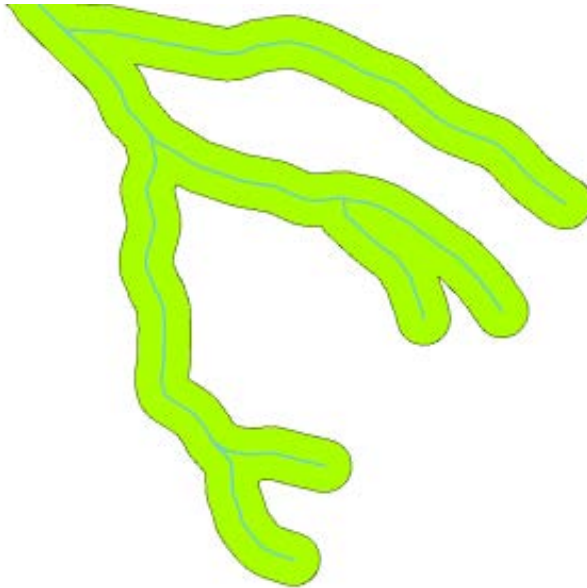


The border between the United States of America and Mexico is separated by a buffer zone. (Photo taken by SGT Jim Greenhill 2006).

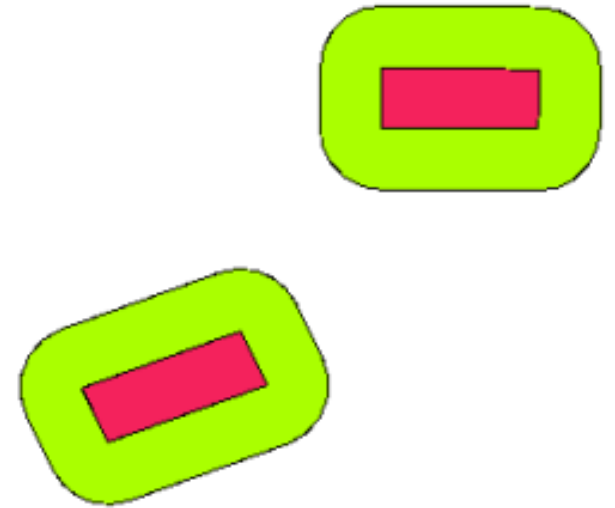
Proximity I: Buffering



A buffer zone around vector points.



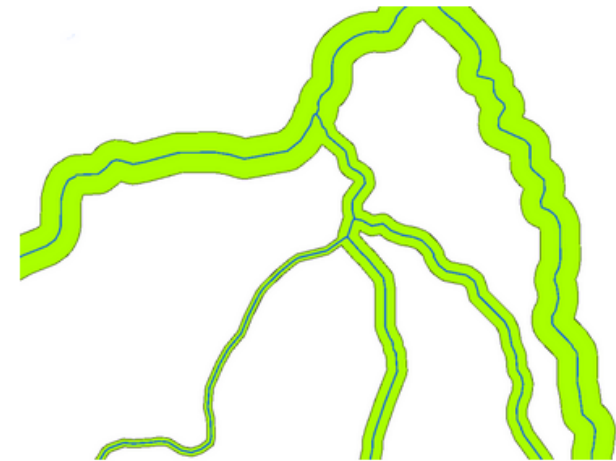
A buffer zone around vector polylines.



A buffer zone around vector polygons.

Proximity I: Buffering

- The **buffer distance** or buffer size **can vary** according to numerical values provided in the vector layer attribute table
- The numerical values have to be defined in map units according to the Coordinate Reference System (CRS) used with the data.
- For example, the width of a buffer zone along the banks of a river can vary depending on the intensity of the adjacent land use.

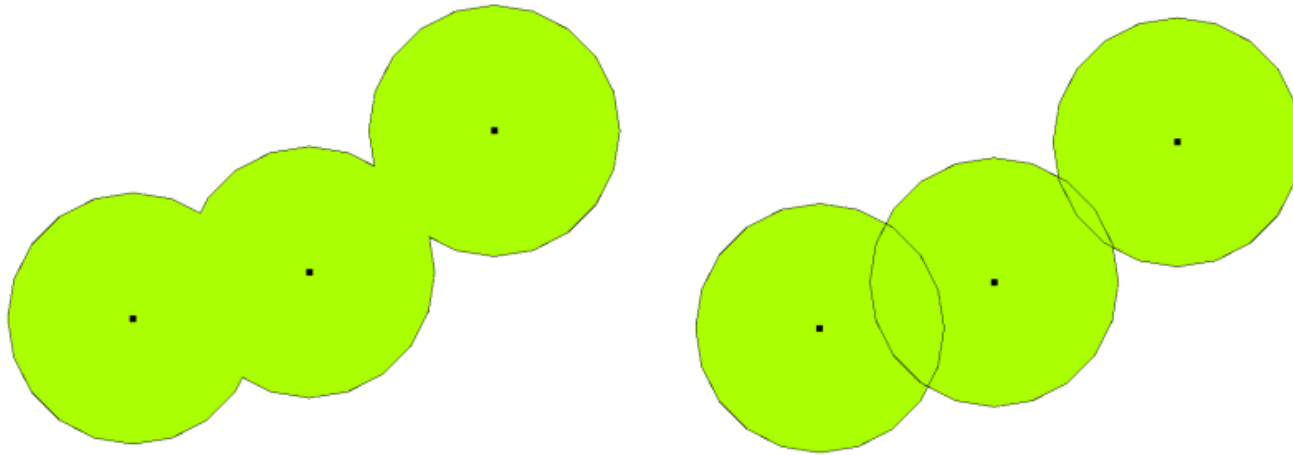


Buffering rivers with different buffer distances.

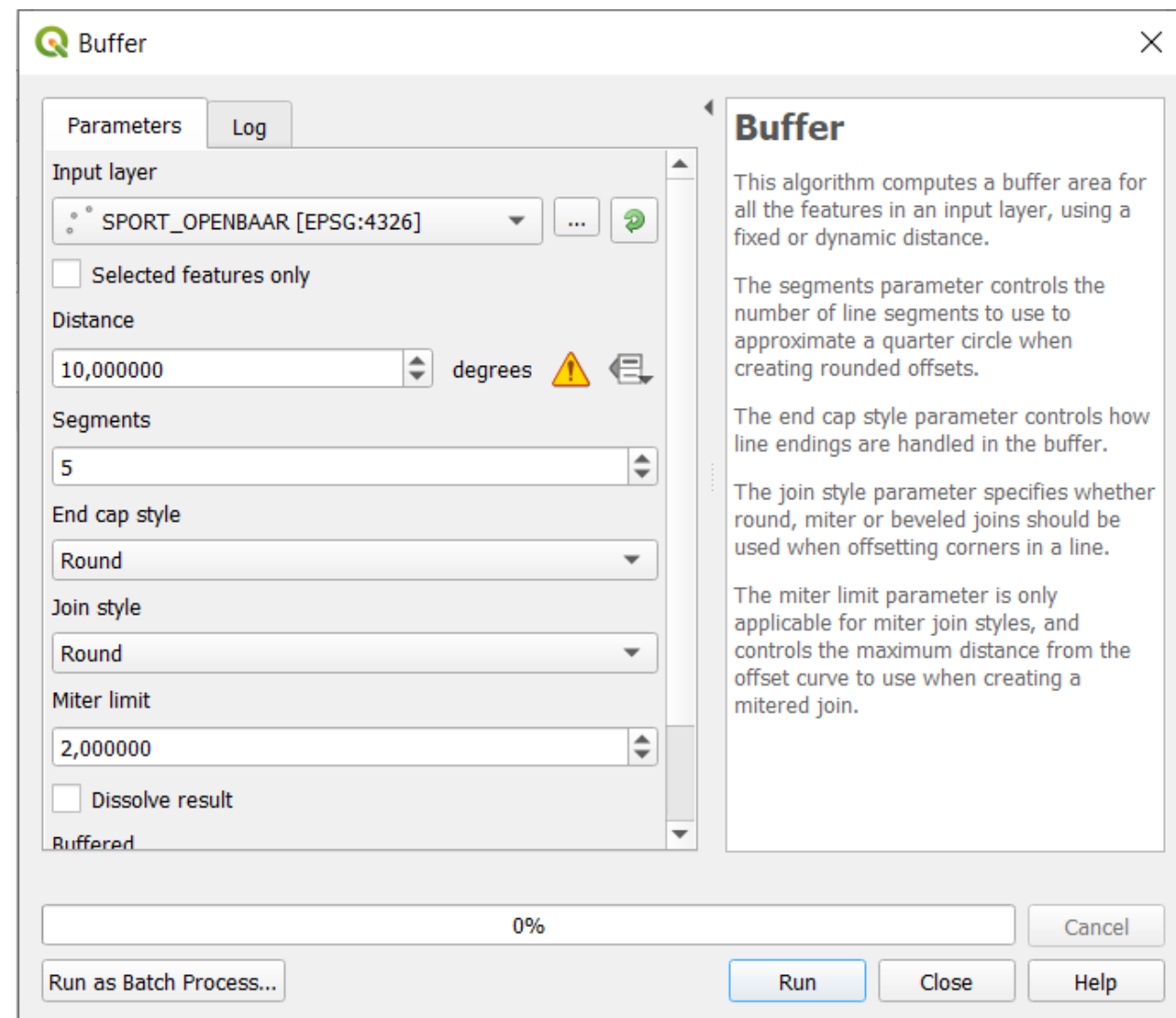
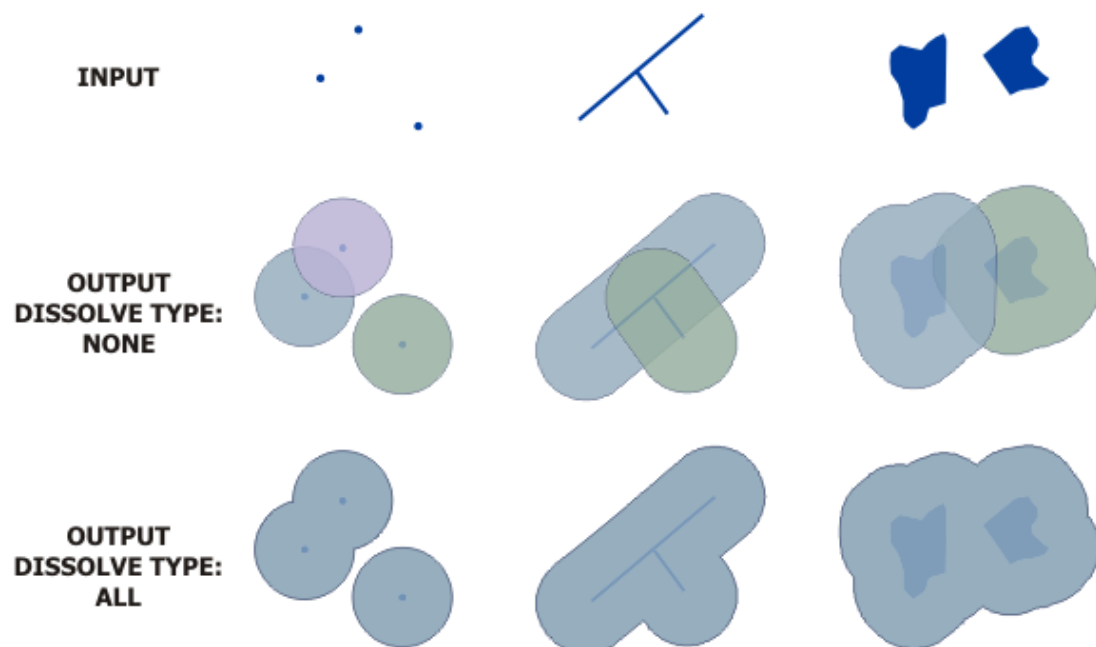
River	Adjacent land use	Buffer distance (meters)
Breede River	Intensive vegetable cultivation	100
Komati	Intensive cotton cultivation	150
Oranje	Organic farming	50
Telle river	Organic farming	50

Proximity I: Buffering

- Buffers can be **merged/dissolved** into a single geometric object to avoid overlapping areas

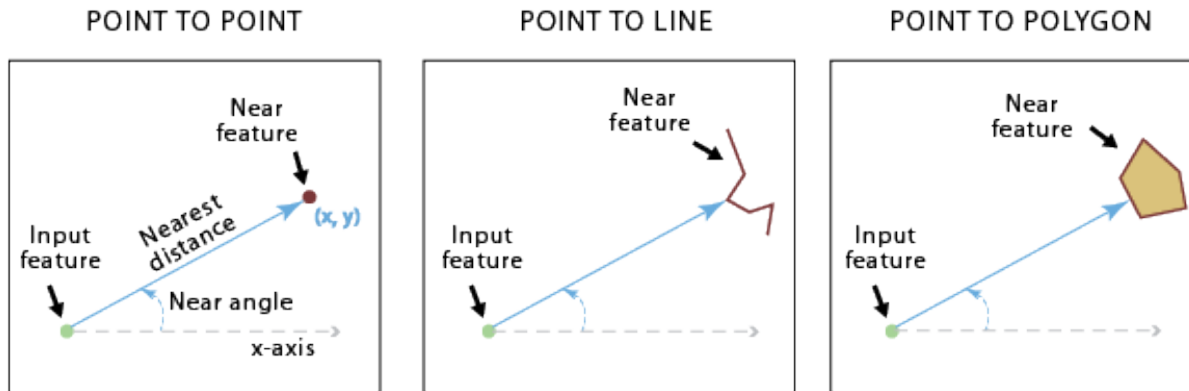


Buffer



Proximity II: Distance to nearest

- Calculates distance and additional proximity information between the source features and the **closest** feature in another layer.



The screenshot shows the 'Distance to Nearest Hub (Points)' dialog box in QGIS. The 'Parameters' tab is active, showing the following settings:

- Source points layer:** SPORT_OPENBAAR [EPSG:4326]
- Destination hubs layer:** SPORT_OPENBAAR [EPSG:4326]
- Hub layer name attribute:** abc Naam
- Measurement unit:** (empty dropdown)
- Hub distance:** [Create temporary layer]
- Open output file after running algorithm:** ☒

The 'Log' tab is also visible. The right side of the dialog contains a description of the algorithm:

Distance to nearest hub (points)

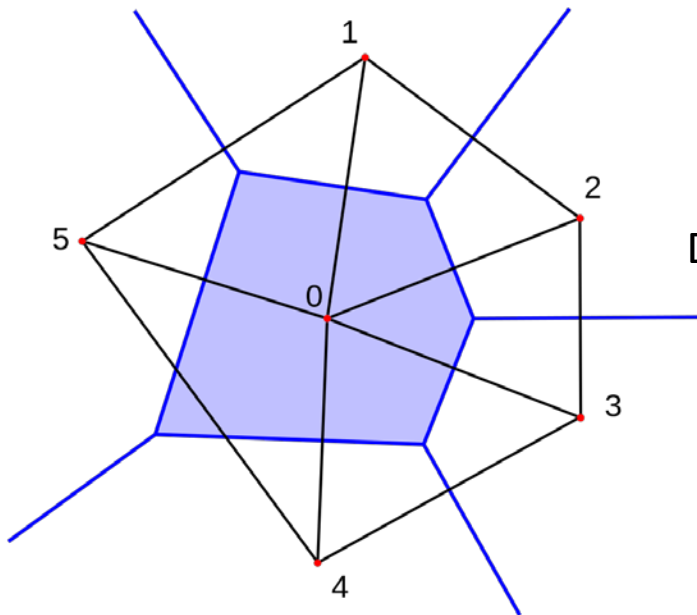
Given an origin and a destination layers, this algorithm computes the distance between origin features and their closest destination one. Distance calculations are based on the features center.

The resulting layer contains origin features center point with an additional field indicating the identifier of the nearest destination feature and the distance to it.

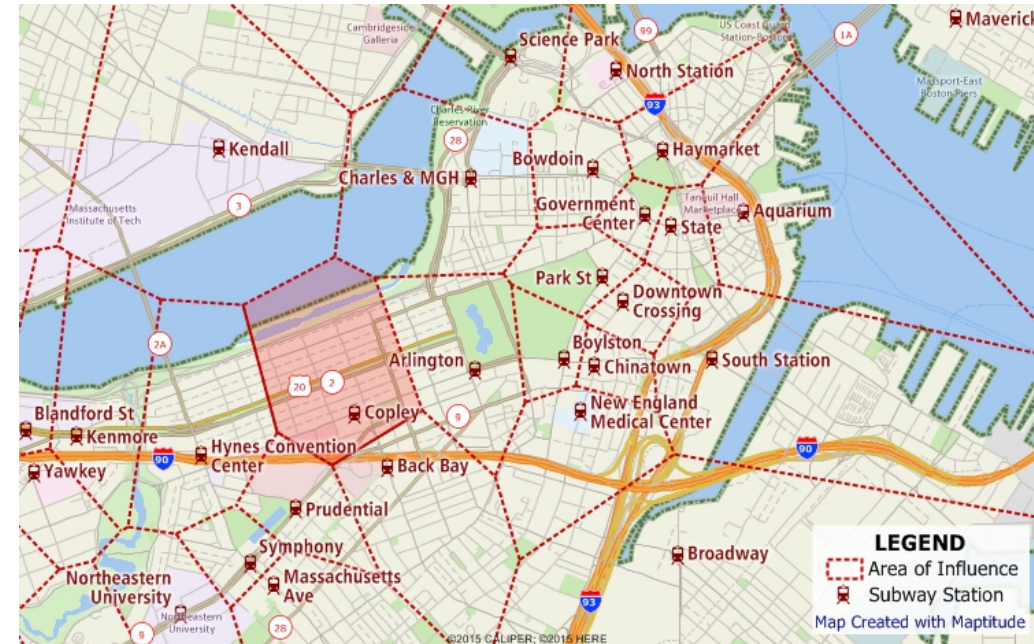
At the bottom, there is a progress bar showing 0%, and buttons for 'Run as Batch Process...', 'Run', 'Close', and 'Help'.

Proximity III: Thiessen Polygons

- Thiessen Polygons (=Voronoi Diagrams/Dirichlet tessellations) exactly bisect distances between points using a Delaunay Triangulation
- Thus all points within a polygon are closest to its centre point
- Can be used to construct catchment areas or to interpolate measures

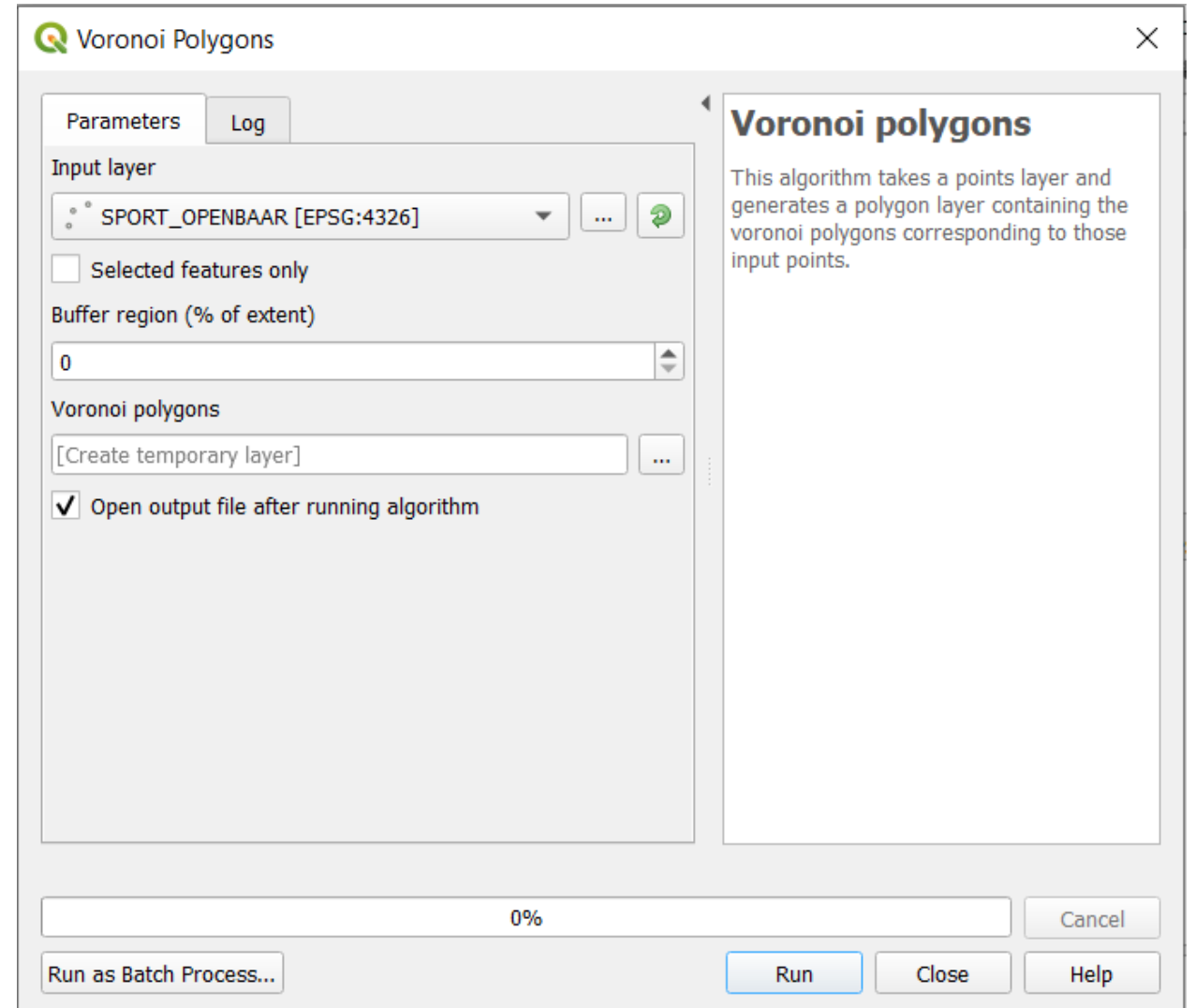
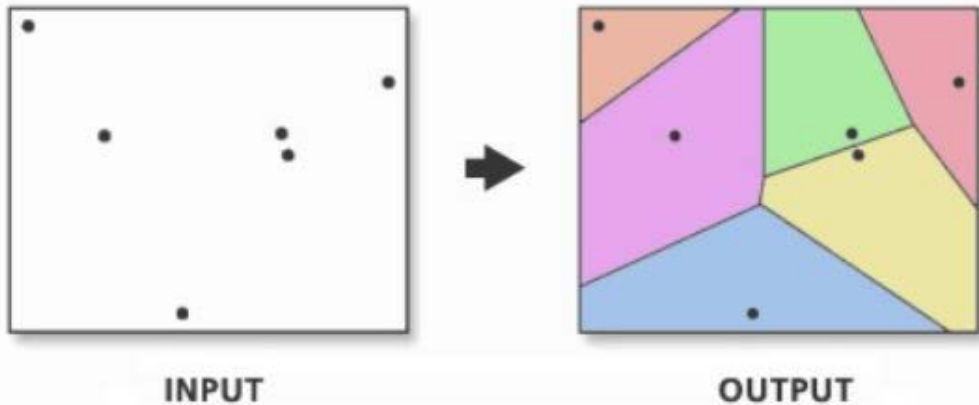


Delaunay-Triangulation



Voronoi polygons

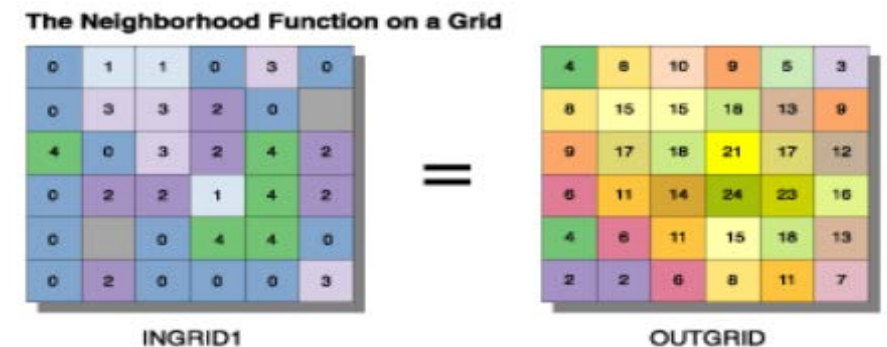
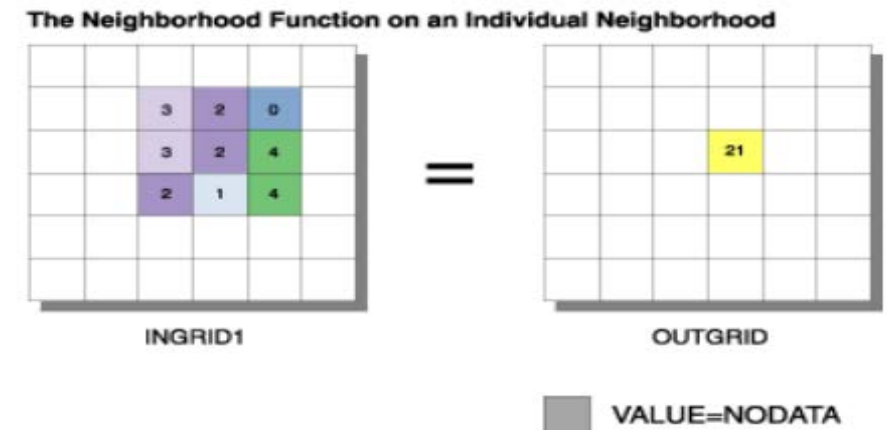
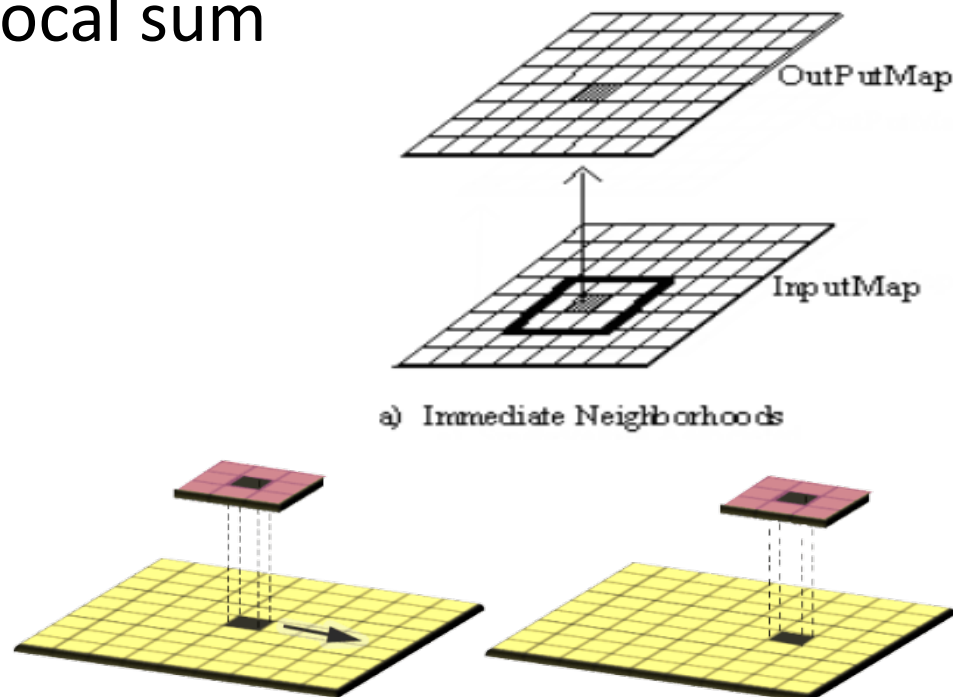
- Creates polygons from point layer



Raster Distance Analysis

Map Algebra functions: Focal

- Focal functions: on cells within a cell neighbourhood
- Quiz: What arithmetic operator is used in this focal function example?
- focal sum



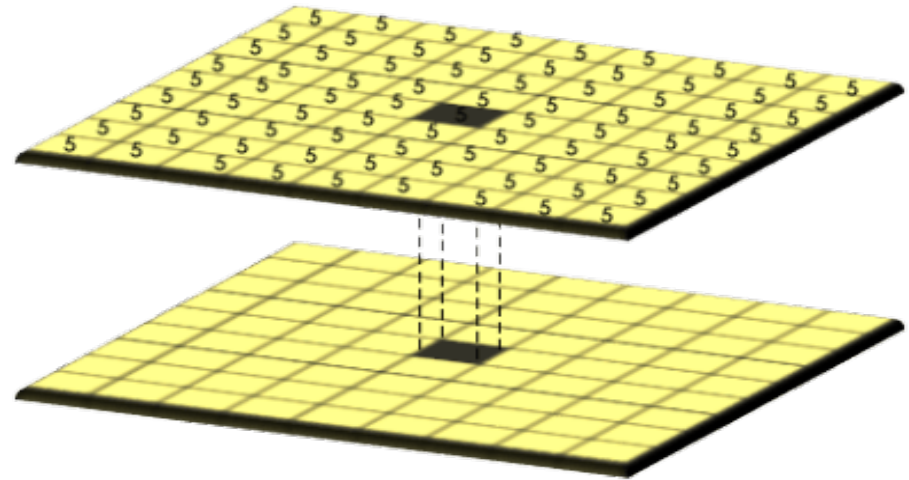
Map Algebra functions: Global

- Global functions: on all cells
- Quiz: What operator is used in this global function example?
- Euclidean distance to a source cell

1		

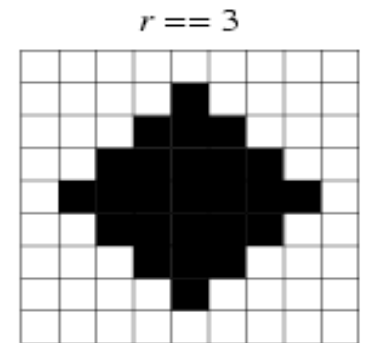
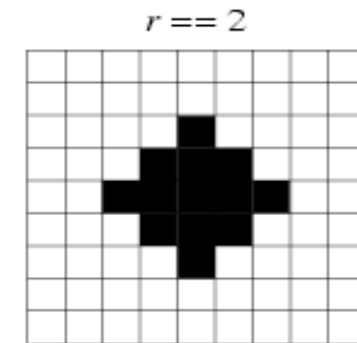
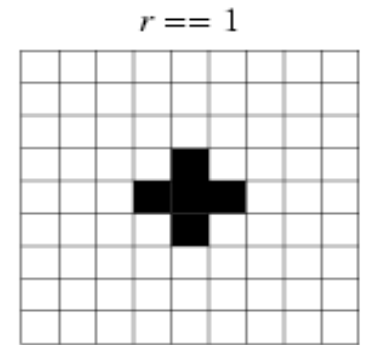
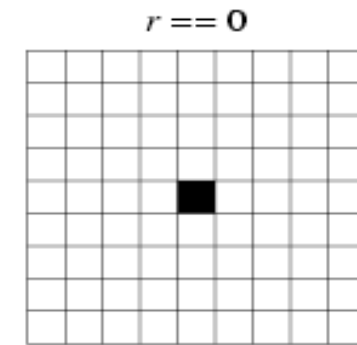
 $=$

0	1.0	2.0
1.0	1.4	2.2
2.0	2.2	2.8

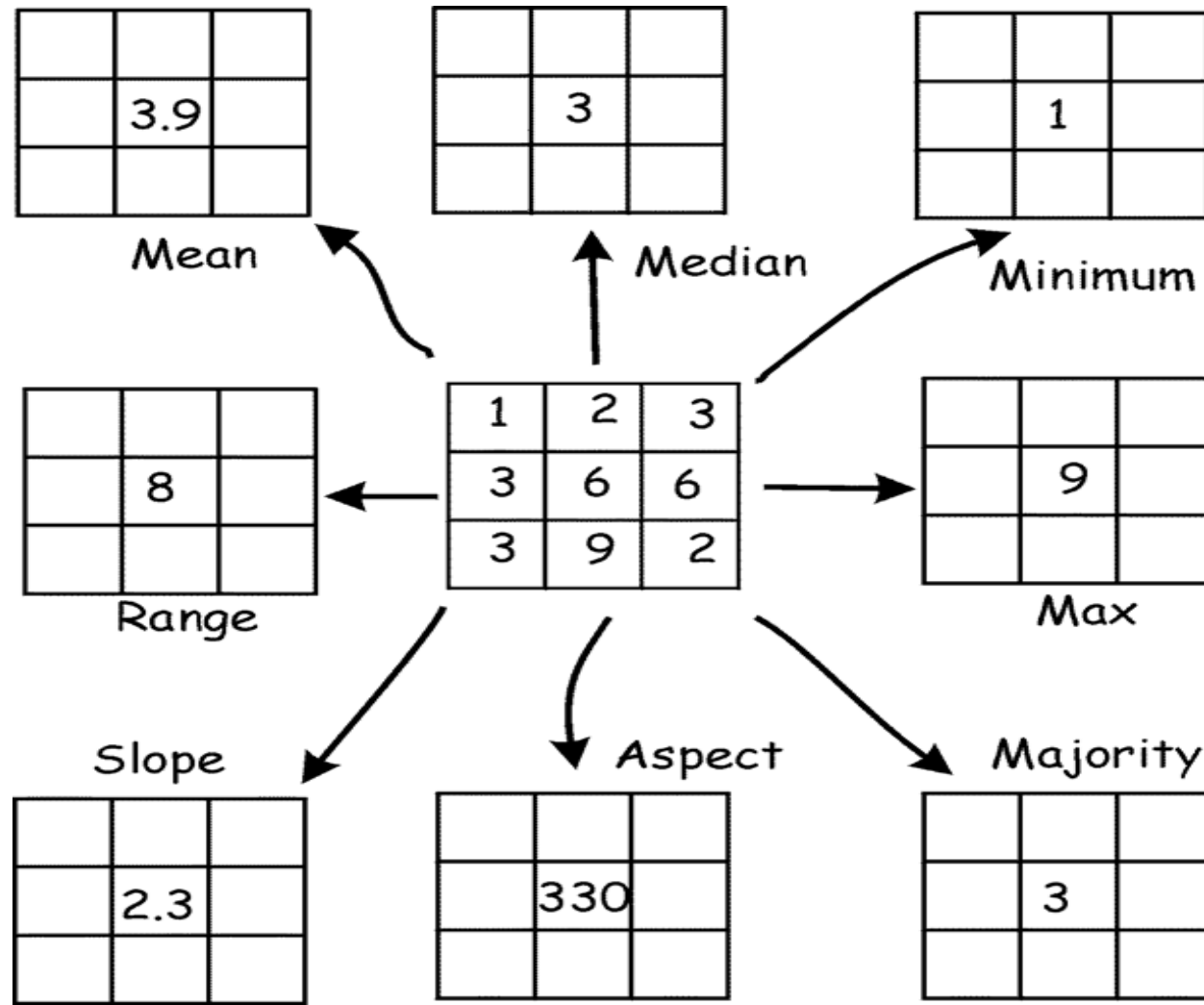


Focal functions: Von Neumann Neighborhood

- **Diamond-shaped**
- To define a set of cells surrounding a given cell
- Ranges $r = 0, 1, 2, 3$
- $N = 2 * r(r+1) + 1$
- “centered square number”

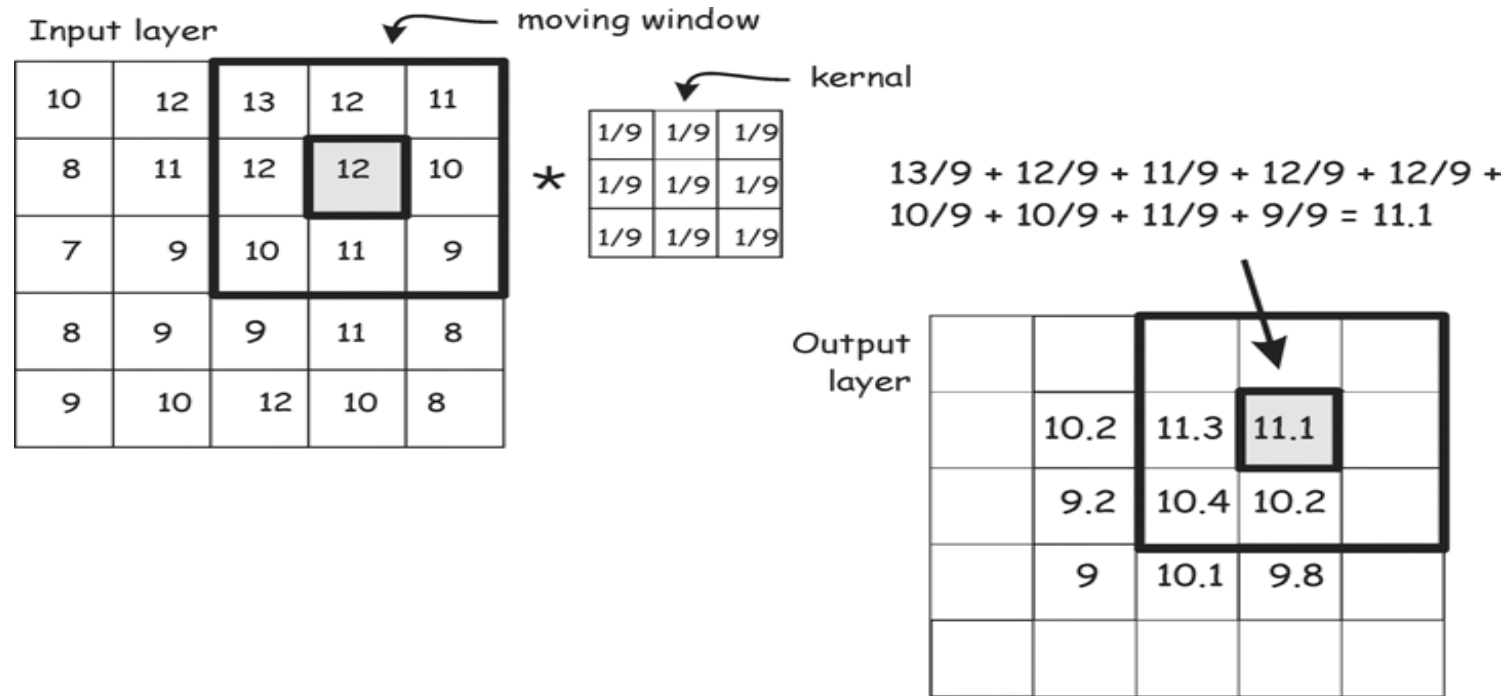


Examples of Moving Window (focal) functions



Moving Windows and Kernels

- Kernel : Set of **constants** for multiplying values within a given window
- What can you see at the **margins** of the output grid?



Edge

Detection

using Kernels

- Concentration changes; elevation changes,...

- Discovering **contrasts / differences** within the focal neighborhood



980	980	980	980	980	940	940	940	940	940
980	980	980	980	980	940	940	940	940	940
980	980	980	980	980	940	940	940	940	940
980	980	980	980	980	940	940	940	940	940
980	980	980	980	980	940	940	940	940	940
980	980	980	980	980	940	940	940	940	940
980	980	980	980	980	900	900	900	900	900
980	980	980	980	980	900	900	900	900	900
980	980	980	980	980	900	900	900	900	900
980	980	980	980	980	900	900	900	900	900

KERNELS

0	0	0
1	0	-1
0	0	0

Horizontal
difference

0	1	0
0	0	0
0	-1	0

Vertical
difference

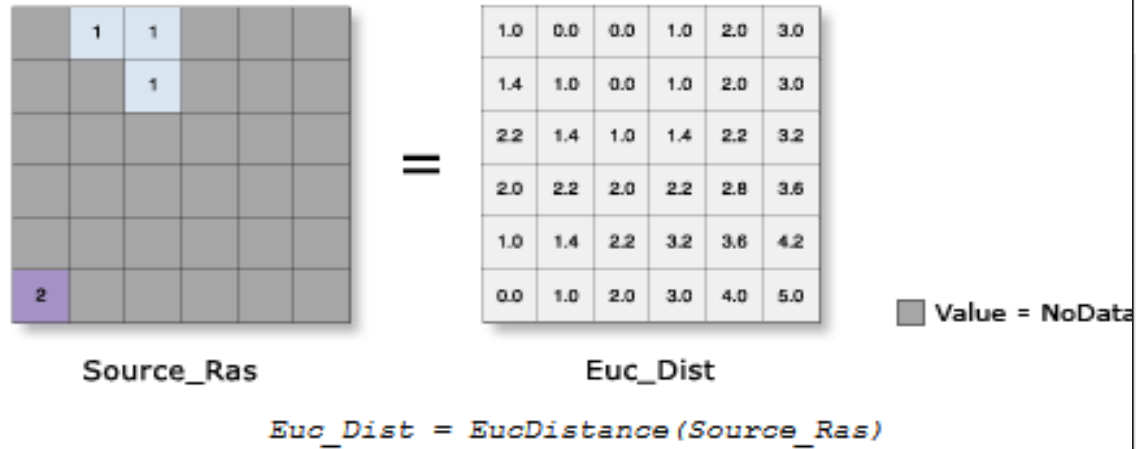
OUTPUT_A

0	0	0	0	40	40	0	0	0	0
0	0	0	0	40	40	0	0	0	0
0	0	0	0	40	40	0	0	0	0
0	0	0	0	40	40	0	0	0	0
0	0	0	0	40	40	0	0	0	0
0	0	0	0	40	40	0	0	0	0
0	0	0	0	80	80	0	0	0	0
0	0	0	0	80	80	0	0	0	0
0	0	0	0	80	80	0	0	0	0
0	0	0	0	80	80	0	0	0	0

OUTPUT_B

[illegible]

Proximity raster



Proximity (Raster Distance)

Parameters Log

Input layer
landuse [EPSG:28992]

Band number
Band 1 (Gray)

A list of pixel values in the source image to be considered target pixels [optional]
40

Distance units
Georeferenced coordinates

The maximum distance to be generated [optional]
Not set

Value to be applied to all pixels that are within the -maxdist of target pixels [optional]
Not set

Nodata value to use for the destination proximity raster [optional]
Not set

▼ Advanced parameters

0%

Run as Batch Process... Run Close Help

Focal statistics

Calculates for each input cell a neighborhood statistics.

0	1	1	0	3	0
0	3	3	2	0	
4	0	3	2	4	2
0	2	2	1	4	2
0		0	4	4	0
0	2	0	0	0	3

InRas1

=

4	8	10	9	5	3
8	15	15	18	13	9
9	17	18	21	17	12
6	11	14	24	23	16
4	6	11	15	18	13
2	2	6	8	11	7

OutRas

Value = NoData

```
OutRas = FocalStatistics(InRas1, NbrRectangle(3,3,MAP), "SUM", "")
```

r.neighbors

Description: r.neighbors

Show advanced parameters

Input raster layer: 'Rasterized' from algorithm 'Rasterize (vector to raster)'

Raster layer to select the cells which should be processed [optional]: [Not selected]

Neighborhood operation [optional]: 123 sum

Neighborhood size [optional]: 123 3

Sigma (in cells) for Gaussian filter [optional]: 123 Not set

Quantile to calculate for method=quantile [optional]: 123

Use circular neighborhood: 123 No

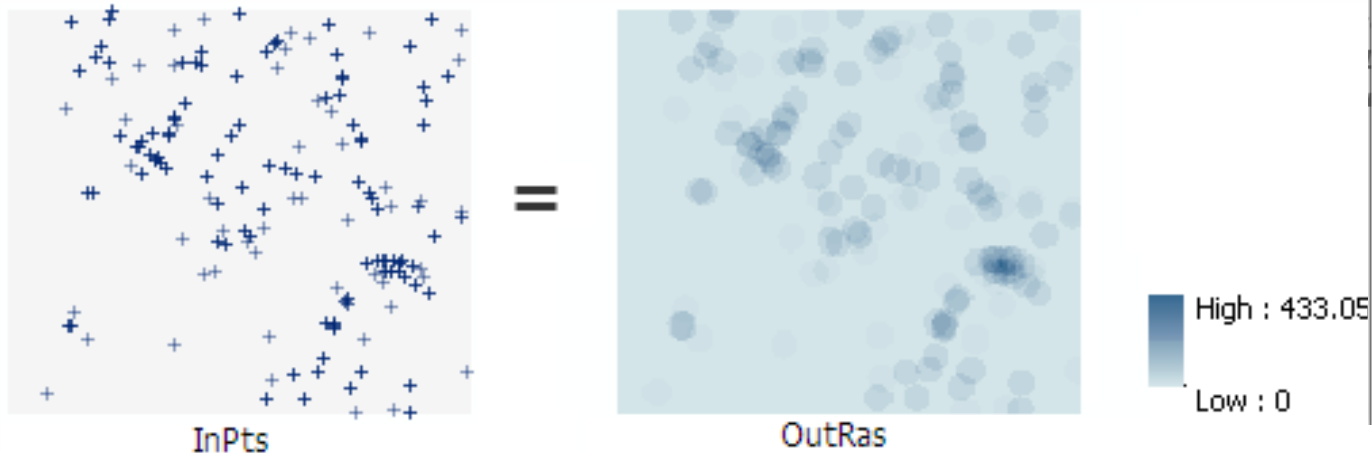
Neighbors: [Enter name if this is a final result]

Parent algorithms: 0 elements selected

OK Cancel Help

Kernel density

Calculates a magnitude-per-unit area from point features that fall within a neighborhood around each cell.



Heatmap (Kernel Density Estimation)

Parameters Log

Point layer
SPORT_OPENBAAR [EPSG:4326]

☐ Selected features only

Radius
100,000000 degrees

Output raster size
Rows 2002 Columns 2003
Pixel size X 0,100000 Pixel size Y 0,100000

Advanced parameters
Radius from field [optional]
Weight from field [optional]

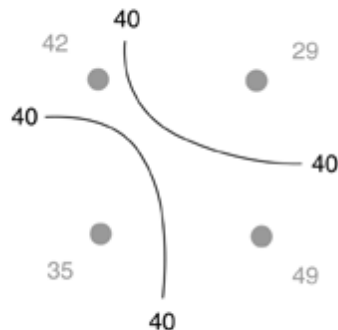
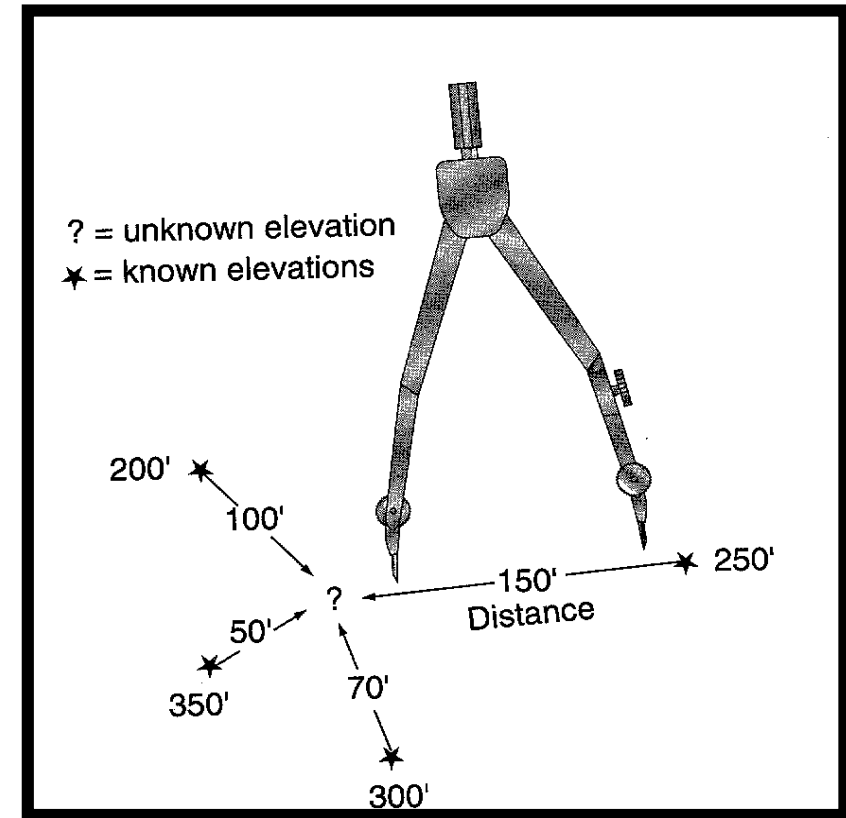
0%

Run as Batch Process... Run Close Help

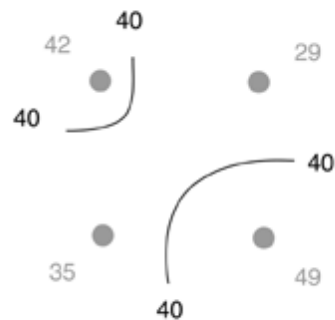
```
OutRas = PointDensity(InPts, None, 30)
```

Point interpolation

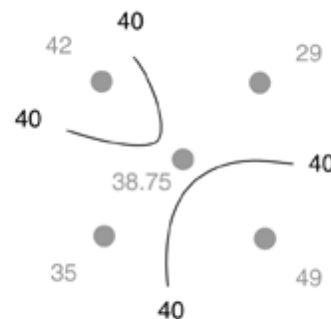
- Point sample of (field) measurements
- How to estimate the field values that were not sampled?
- How to draw isolines when linear interpolation does not work (saddle problem)?



$$(42+49)/2 = 45.5$$



$$(29+35)/2 = 32$$



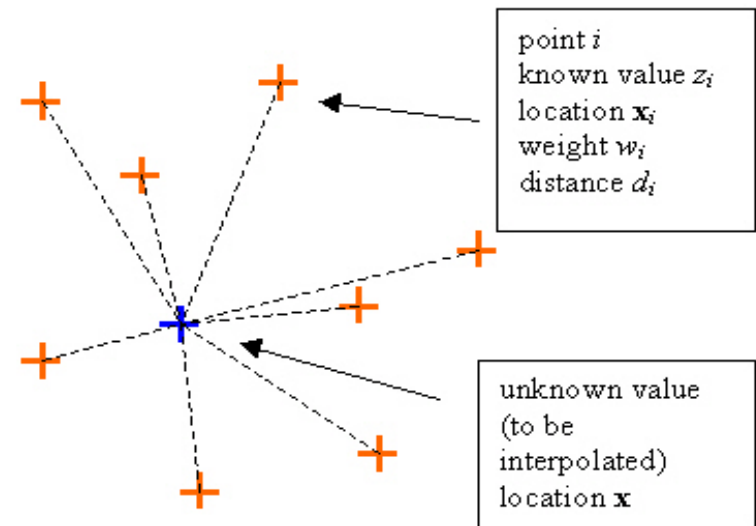
$$(45.5+32)/2 = 38.75$$

Interpolation: Inverse Distance Weighting (IDW)

Each input point has local influence that diminishes with distance

Estimates are weighted averages of values points within window R

$$Z_{i,j} = \frac{\sum_{p=1}^R z_p d_p^{-n}}{\sum_{p=1}^R d_p^{-n}}$$

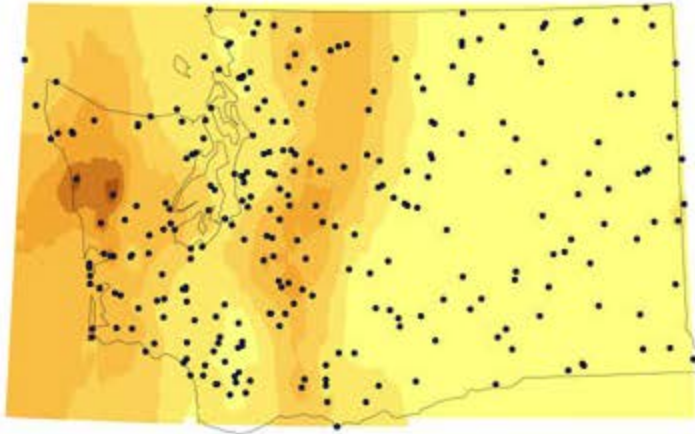


where weight is some function of distance (e.g., $w = d^{-n}$)

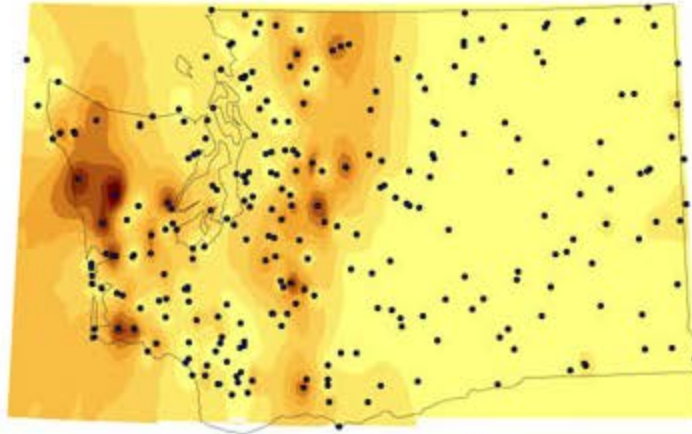
IDW parameters: exponents

increasing power = peaks in raster
the lower n the more distant points have an effect

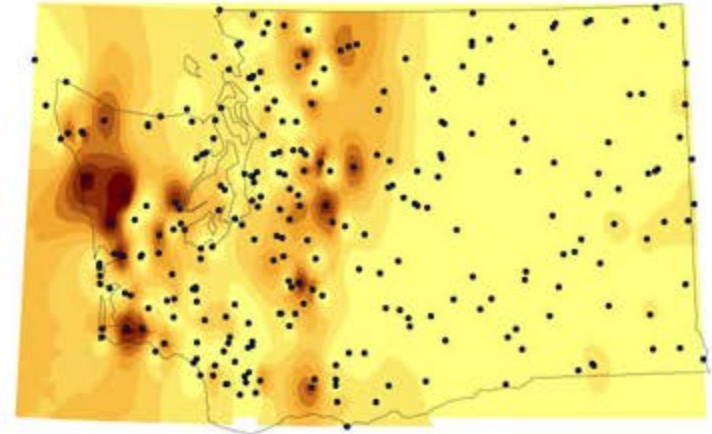
d^{-1}



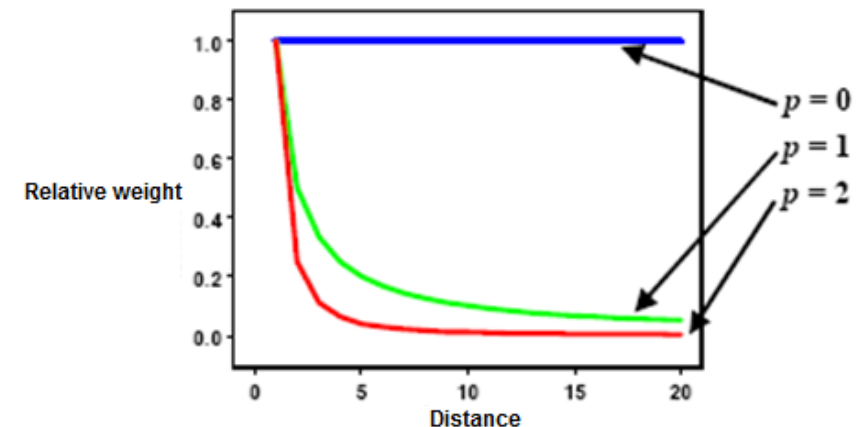
d^{-2}



d^{-3}

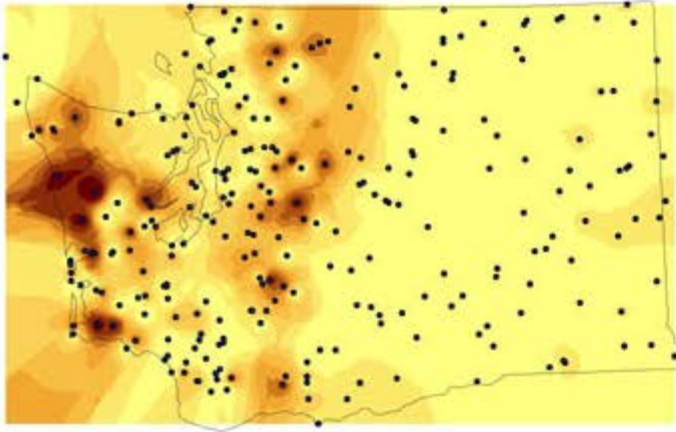


The lower n , the more distant points have a greater effect on the overall pattern, the greater n , the more local points have a greater effect

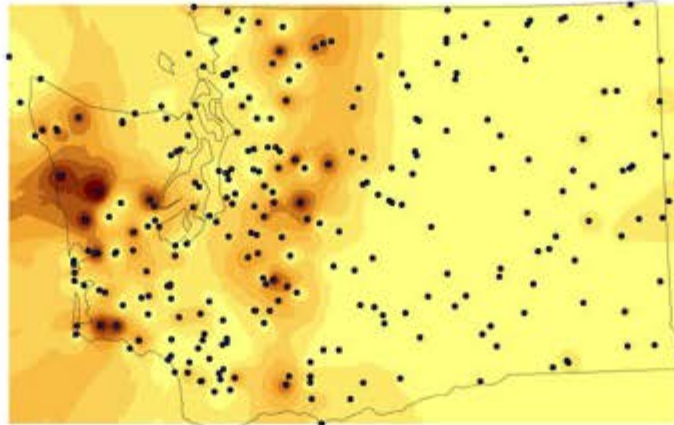


IDW parameters: search neighborhood

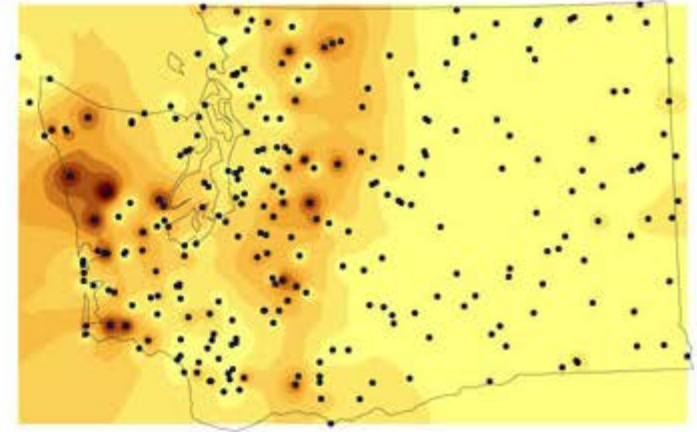
5NN



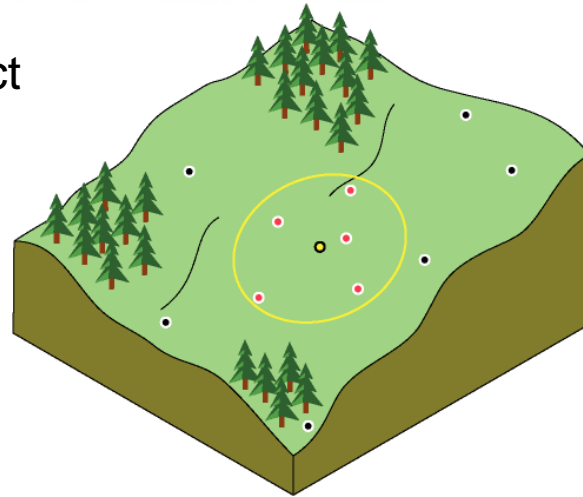
12NN



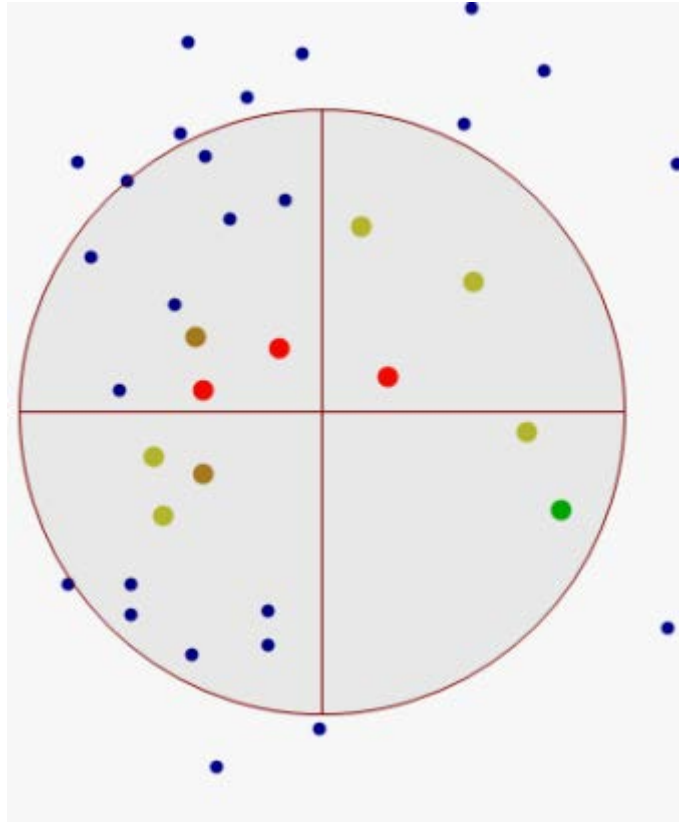
25 NN



less n = less averaging effect



IDW



Grid (IDW With Nearest Neighbor Searching) ×

Parameters Log

Point layer
SPORT_OPENBAAR [EPSG:4326] ... ↻

☐ Selected features only

Weighting power
2,000000

Smoothing
0,000000

The radius of the search circle
1,000000

Maximum number of data points to use
12

Minimum number of data points to use
0

NODATA marker to fill empty points
0.000000

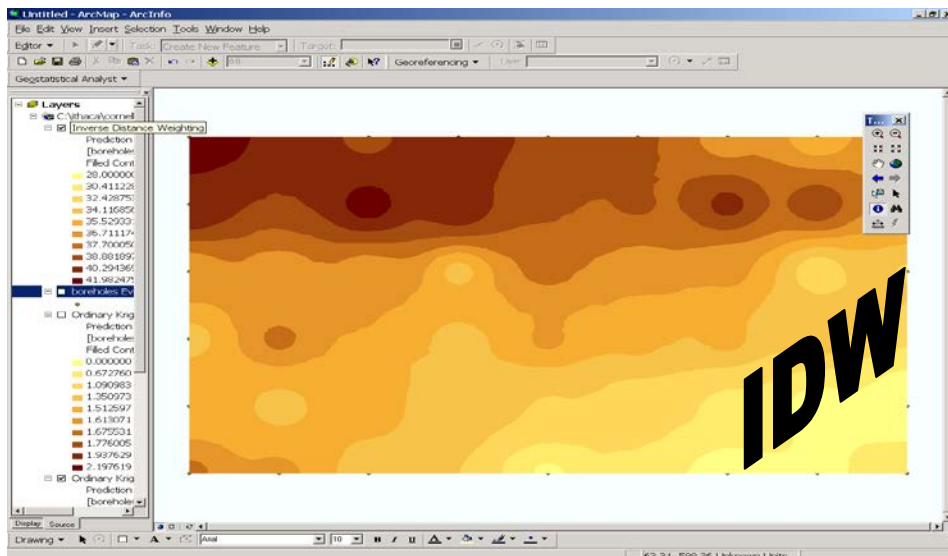
0%

Run as Batch Process... Run Close Help

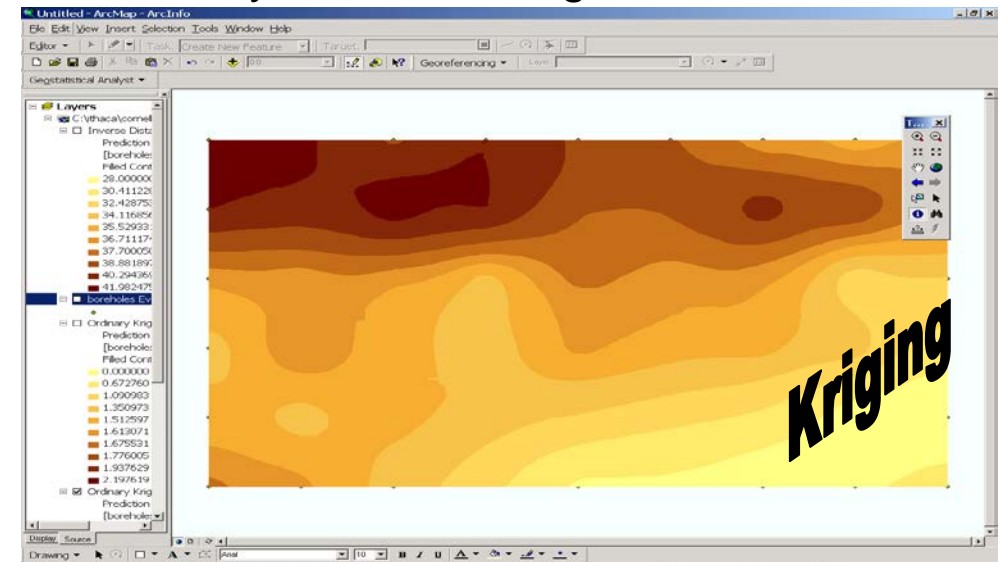
estimates distance based model

IDW vs. Kriging (Geostatistics)

- Interpolated values limited by range of the data
 - How many points should be included?
 - What about irregularly distributed points?
 - What about the map edges?
- Kriging appears to give a more “smooth” look to the data
 - Kriging avoids the “bulls eye” effect
 - Kriging gives us a standard error more smooth look to data less bulls eye effects and gives se



limited
by range



...back to the core concepts

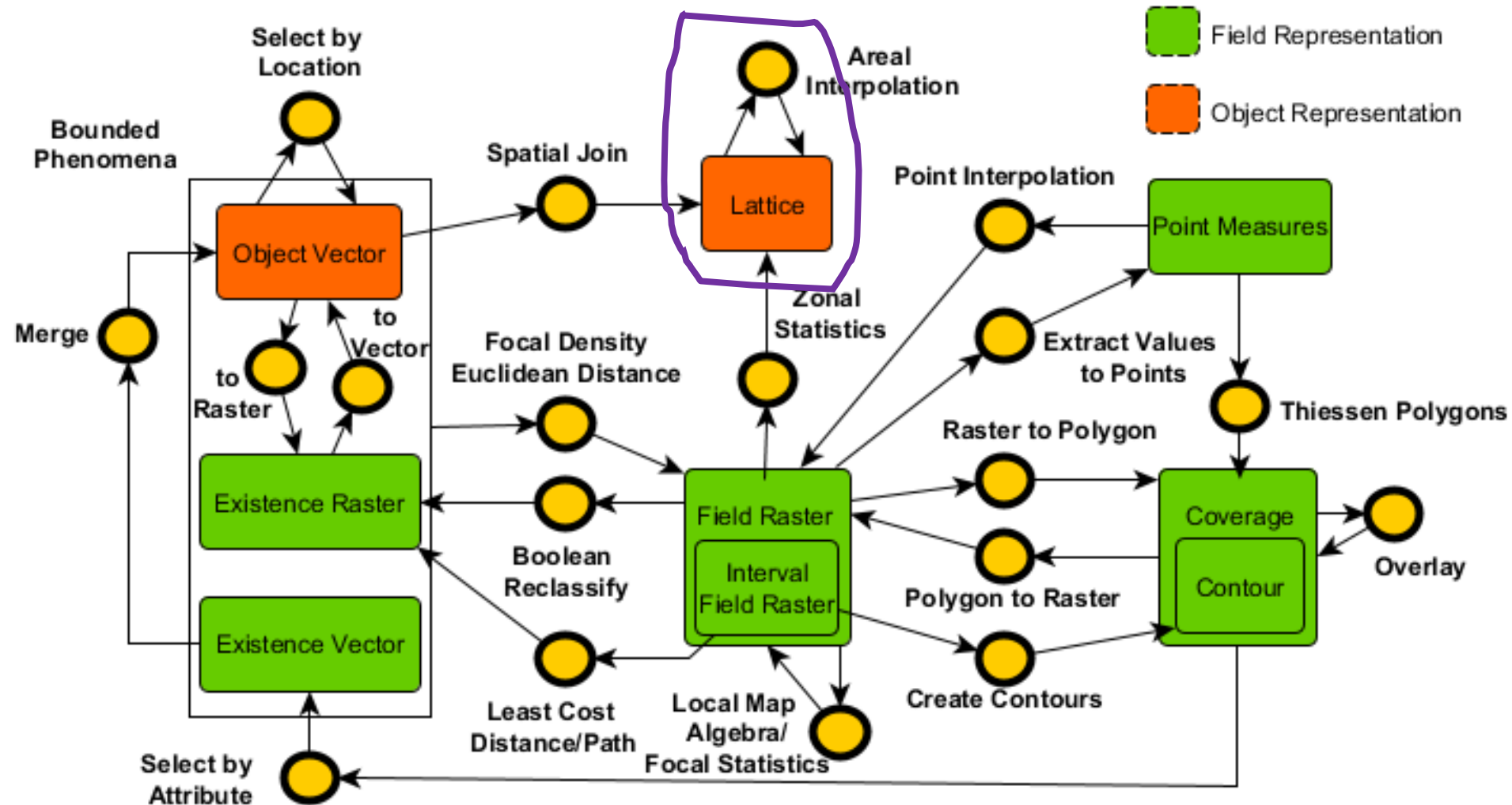
Core concept quiz 1)

Which core concepts correspond to which datasets in the practical?

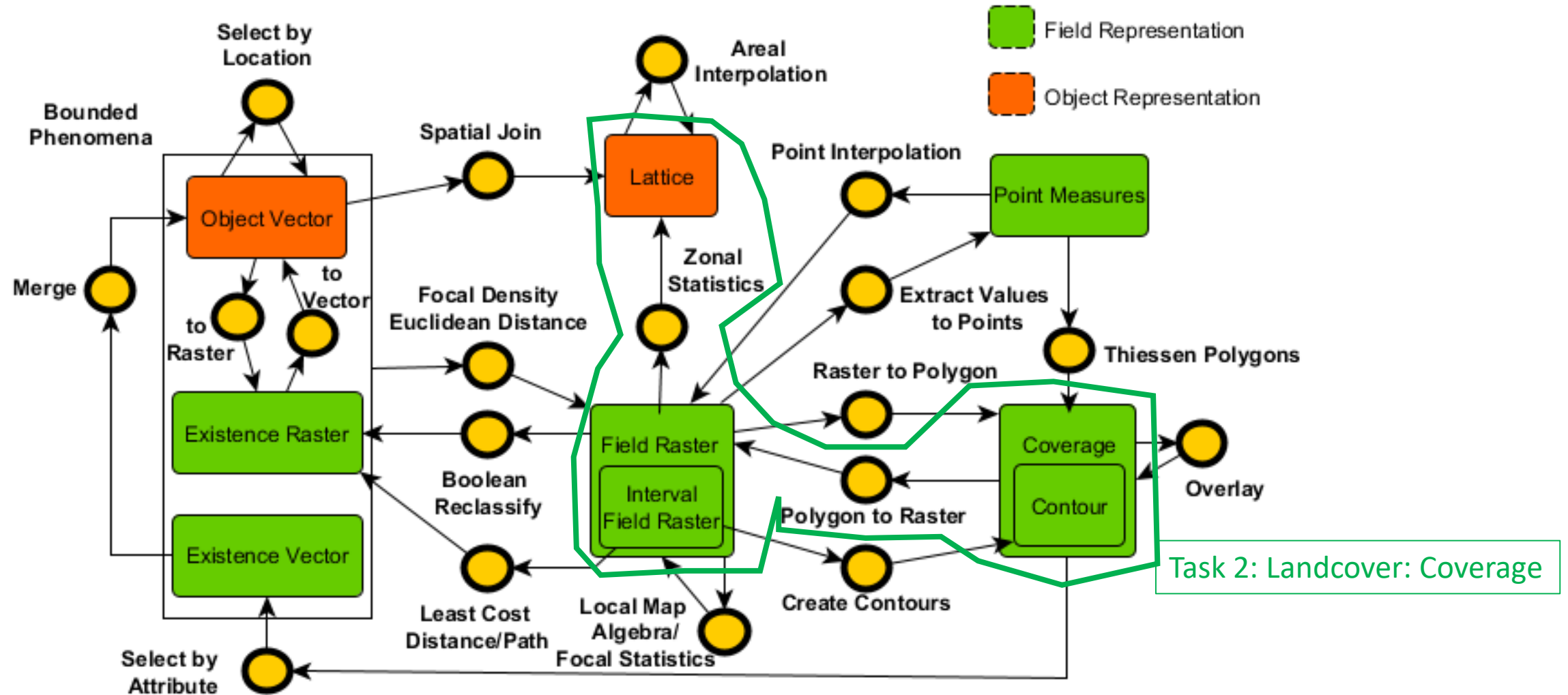
- | | |
|--------------------------|-----------------------|
| 1) CBS Buurt statistics: | Lattice |
| 2) Grondgebruik: | Coverage |
| 3) Grondgebruik: | Coverage |
| 4) Sport facilities: | Vector (Point) Object |

Core concept quiz 2)

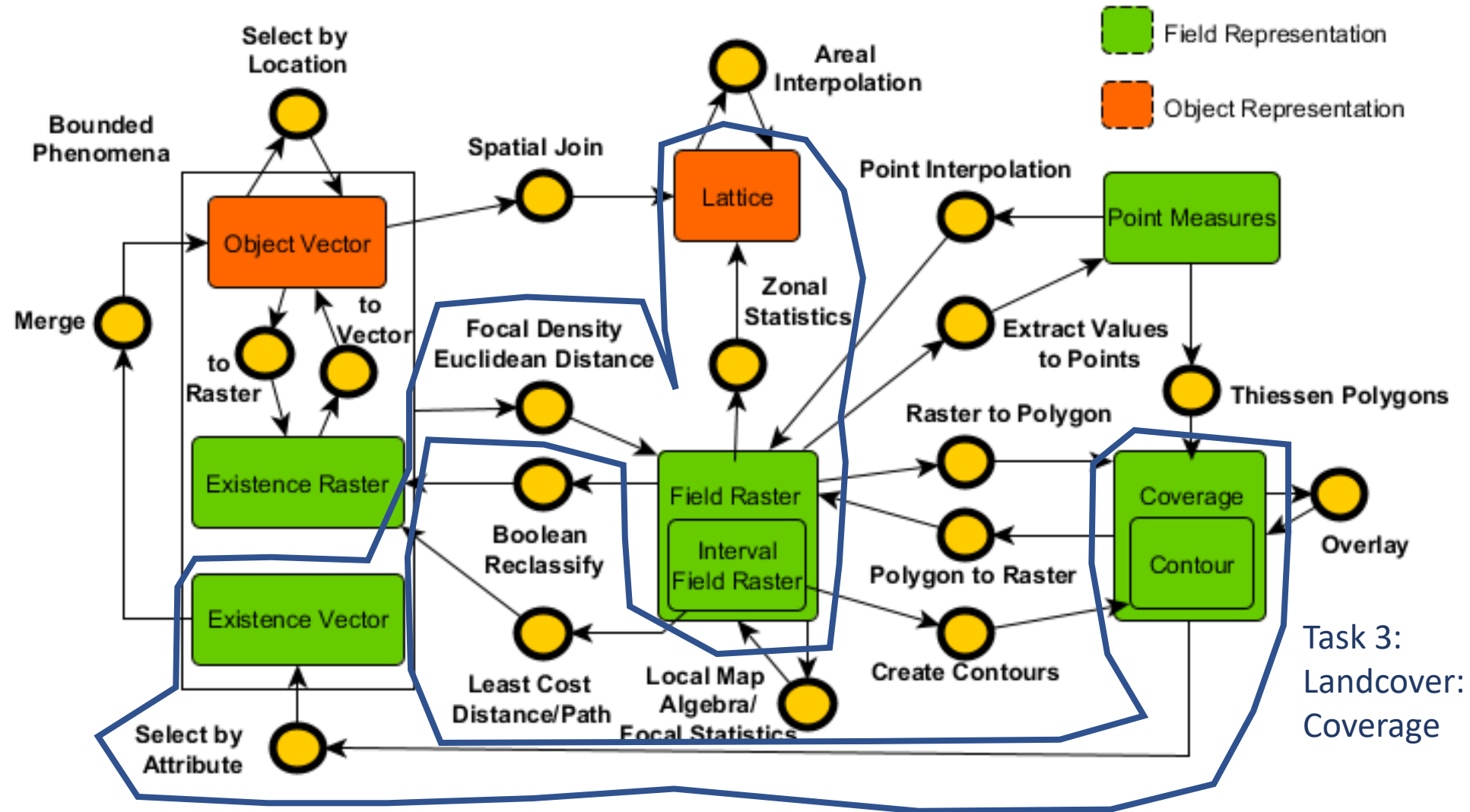
Task 1: CBS Buurt statistics: Lattice



Core concept quiz



Core concept quiz

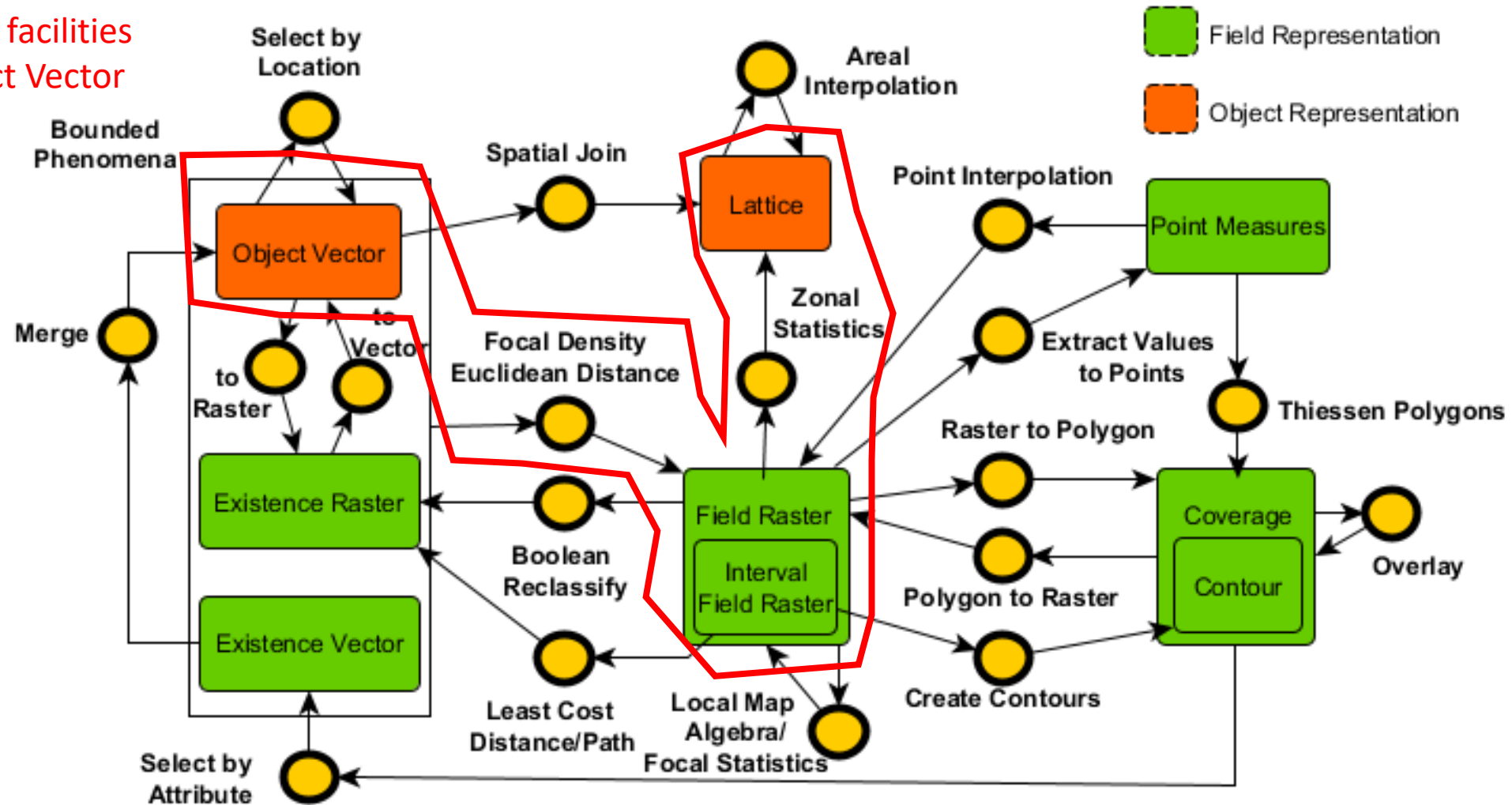


Core concept quiz

Task 4:

Sport facilities

Object Vector



Questions?
(Q&A session)

References

- Chrisman 2002: Exploring Geographic Information systems, 2nd edition, Chapter 6 “Distance Relationships” (153-167)
- Kuhn, W. (2012). Core concepts of spatial information for transdisciplinary research. International Journal of Geographical Information Science, 26(12), 2267-2276.
- Scheider, S., Meerlo, R., Kasalica, V., & Lamprecht, A. L. (2020). Ontology of core concept data types for answering geo-analytical questions. Journal of Spatial Information Science, 2020(20), 167-201 (<http://www.josis.org/index.php/josis/article/viewArticle/555>)