

# Tools and technologies for applied spatial analysis

## **Prof. Viktor Putrenko**

# Your Instructor



Dr. Viktor Putrenko is a Full Professor of American University Kyiv. He received Ph.D. from the National Academy of Science in 2008. He received Doctor of Engineering Sciences in System Analysis from National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”. Dr. Putrenko led many research projects in the field of modeling sustainable development and development of information systems to support management decisions based on spatial data. He is a developer and architect of information systems at the level of municipalities, communities, and regions. Dr. Putrenko is an expert consultant for many international projects and a member of international organizations such as IEEE, ICA, and GSDI.

# Content

1. Modern architecture of geographical information systems
2. Spatial Analysis methods
3. Case study 1. Analysis of Odessa region
4. AI in GIS
5. Case study 2. Recognition of feature
6. Visual analytic system on the GIS base
7. Case study 3. Geoportal of Lviv region

# Trend to Web GIS

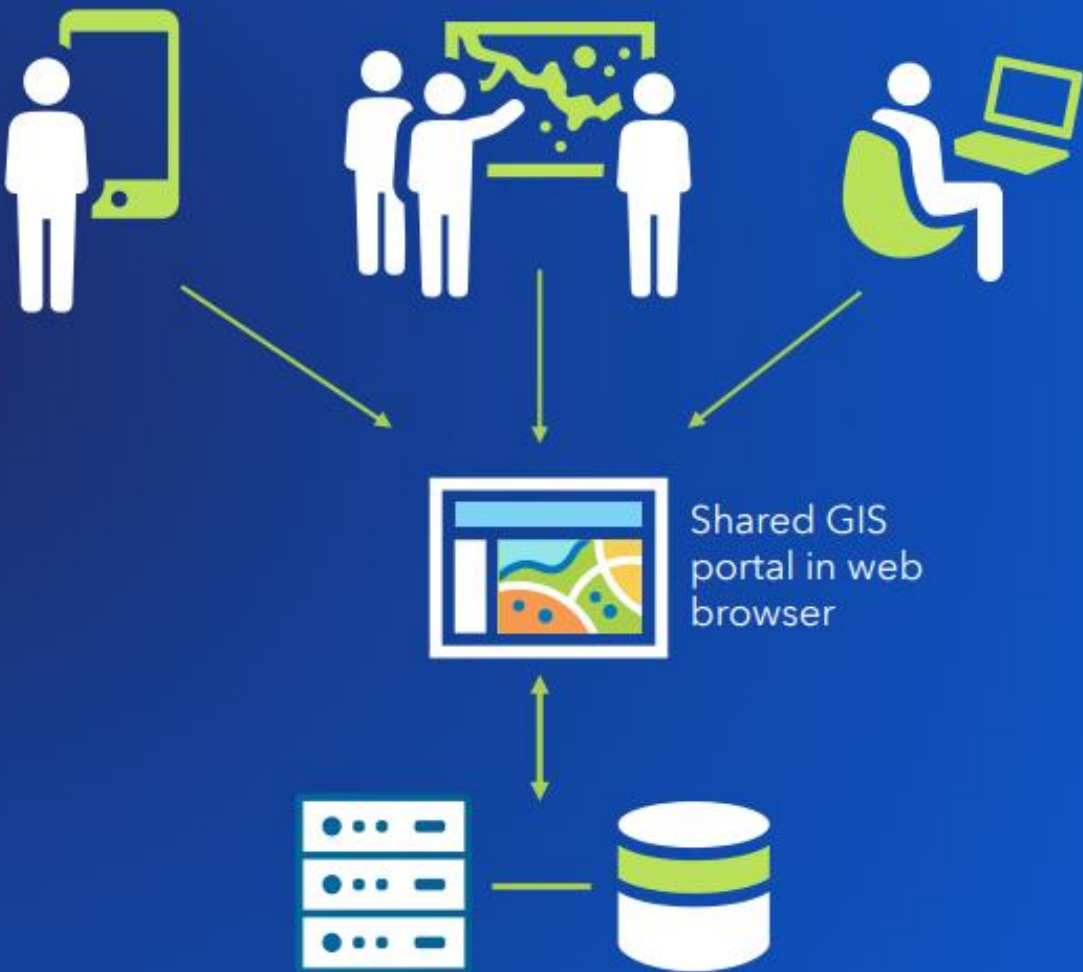
## Desktop GIS

All work is done on your local machine.



## Web GIS

Everyone's work is done by the back-end server infrastructure.



# Component of Web GIS

## Three components of Web GIS

### Front-end portal

- Share content from desktop GIS
- View and work with content in a map viewer
- Bring in GIS items from web
- Produce outputs (maps, apps, reports)



### Back-end server

- Run GIS services that power your items
- Handle all requests (zoom a map, find a location, run an analysis tool)
- Scale processes based on traffic

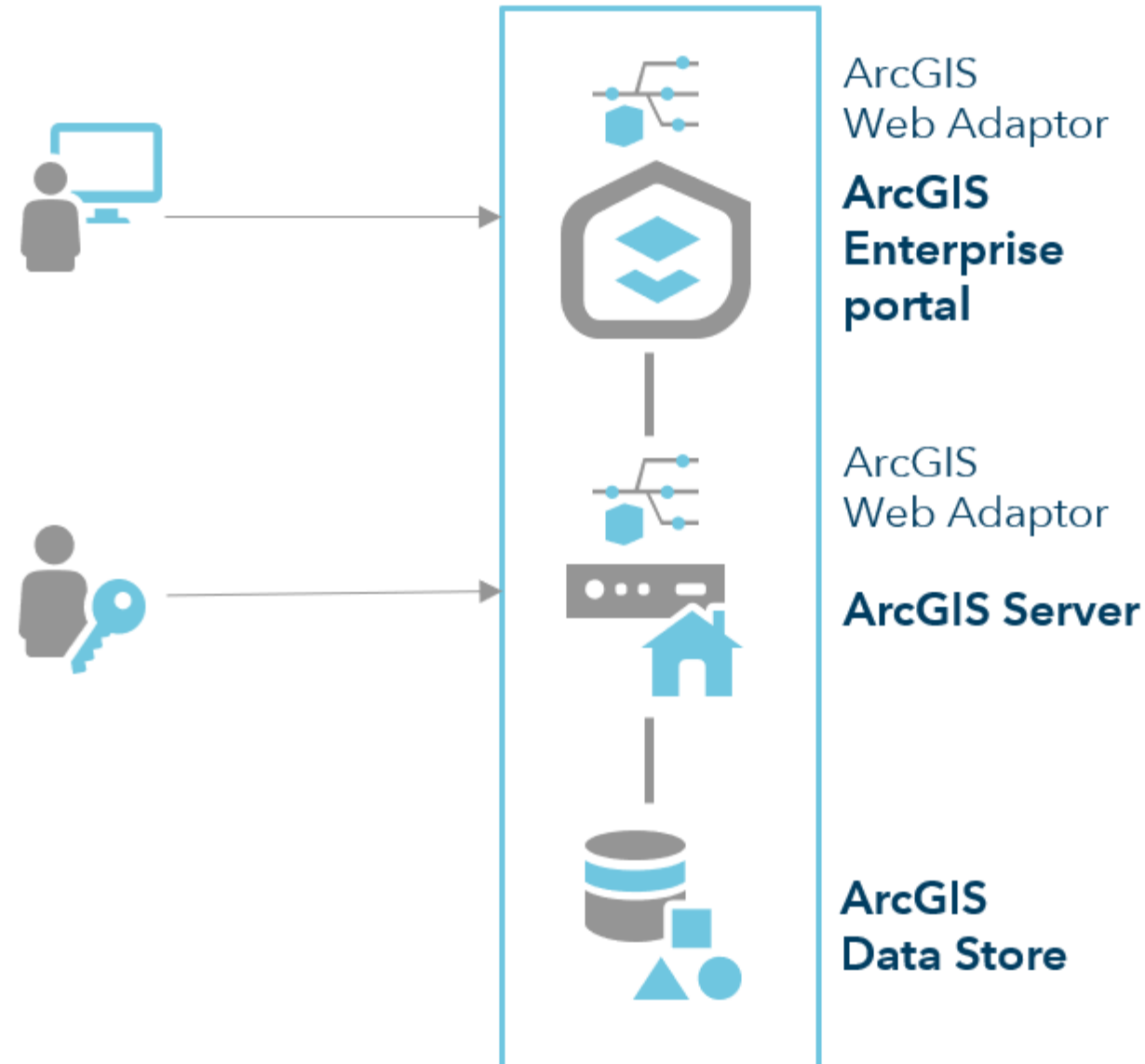


### Data store(s)

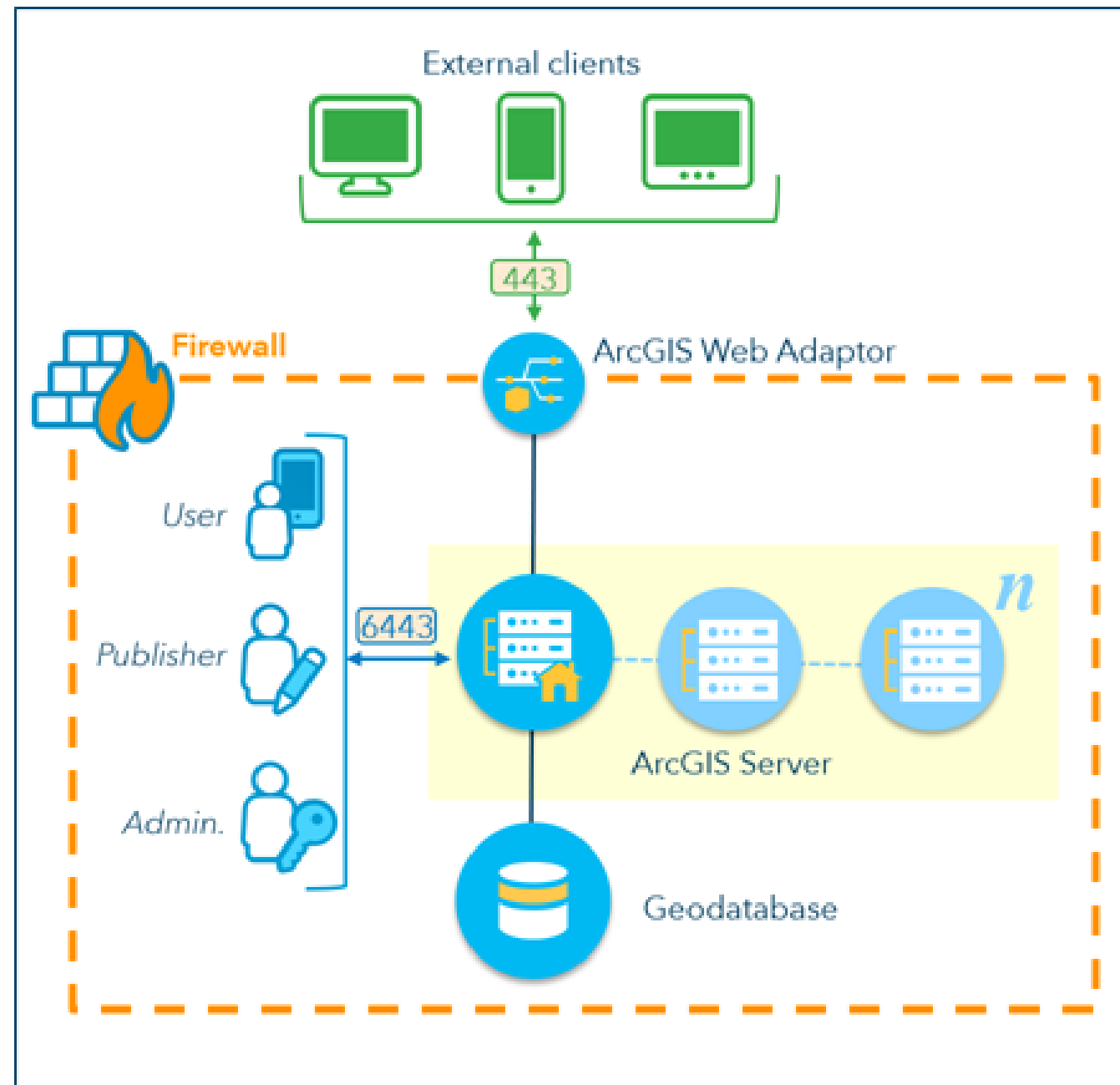
- Host the data referenced by your items
- Facilitate instant data retrieval
- (Sometimes) Allow data editing from portal



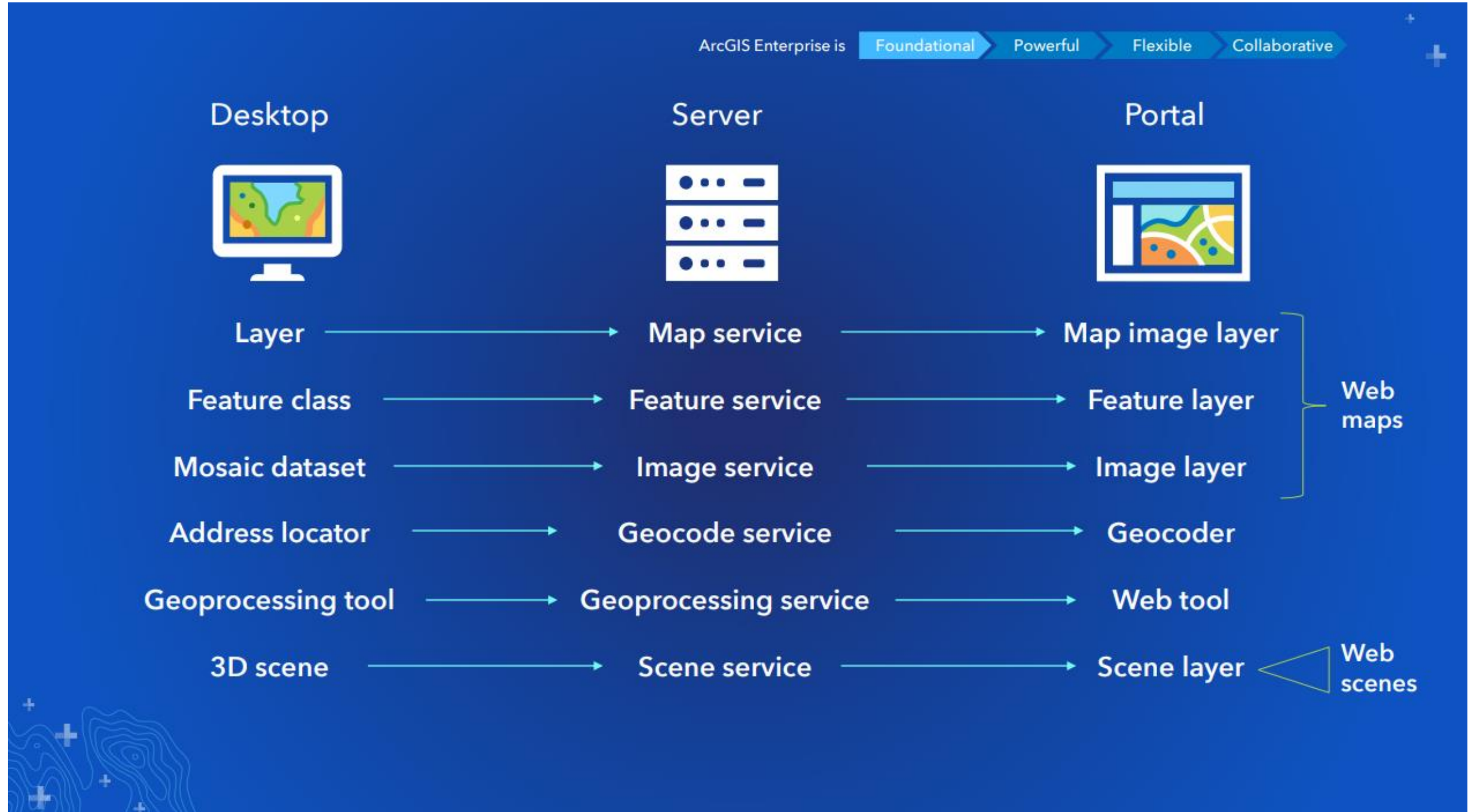
# Architecture of ArcGIS Enterprise



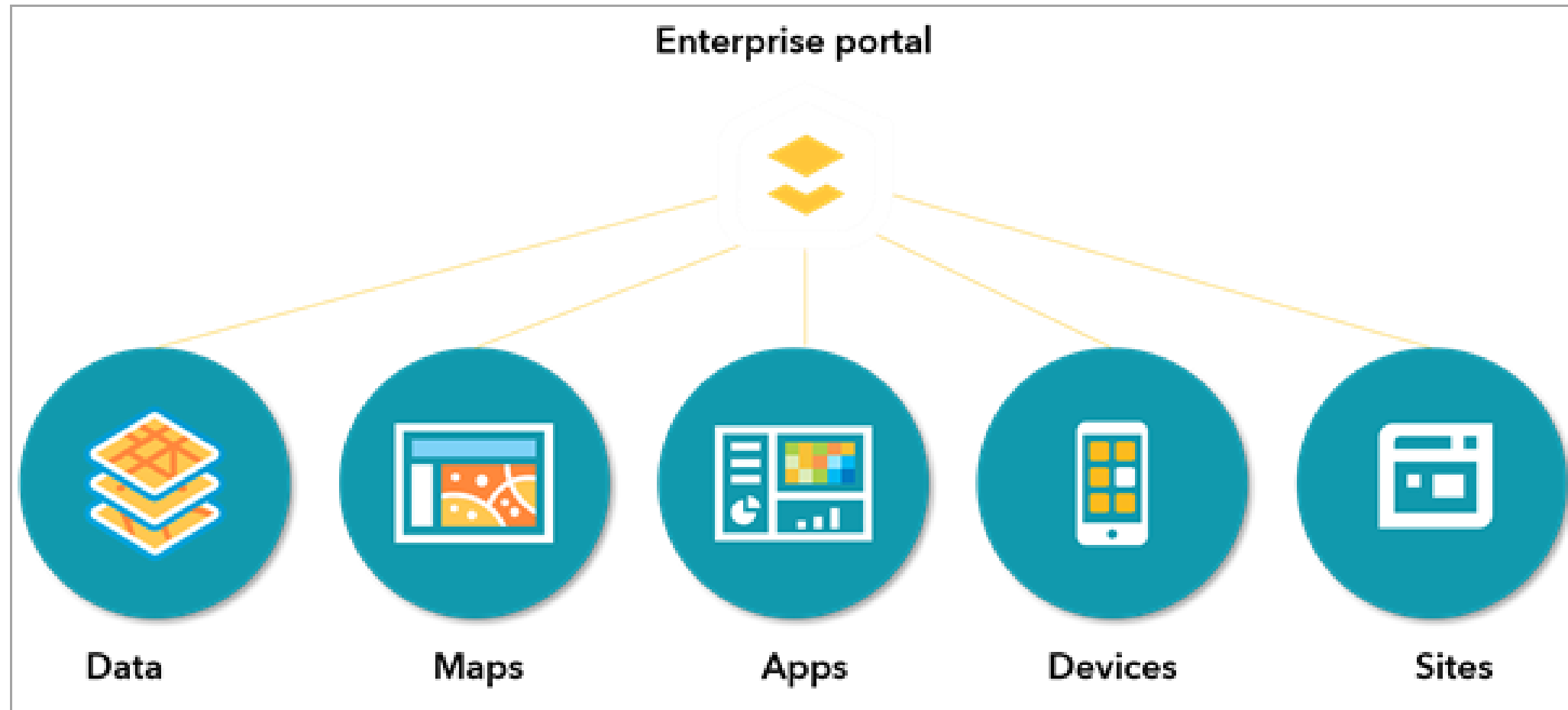
# Interaction with GIS



# WebGIS services



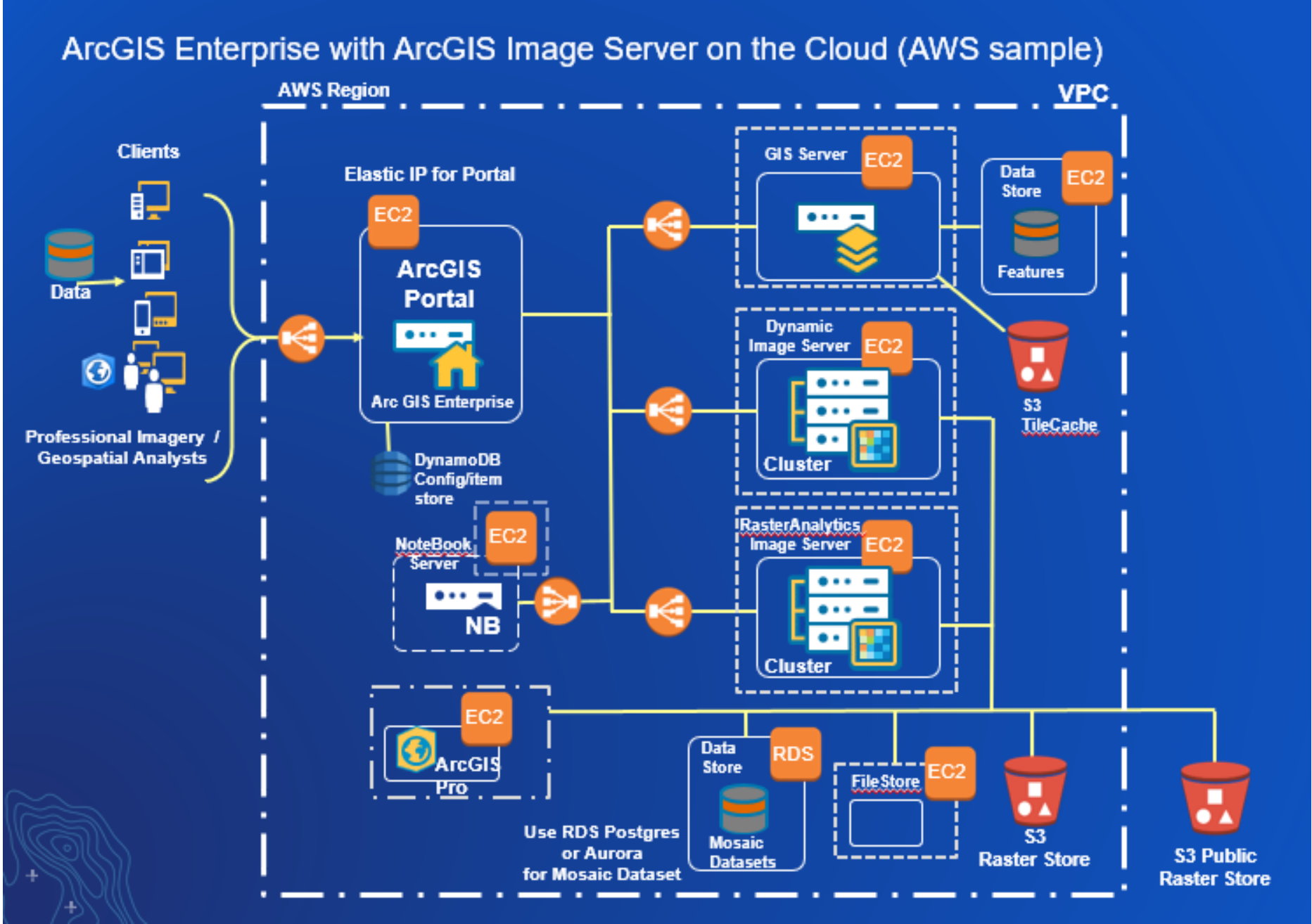
# End-user products



# Building web apps

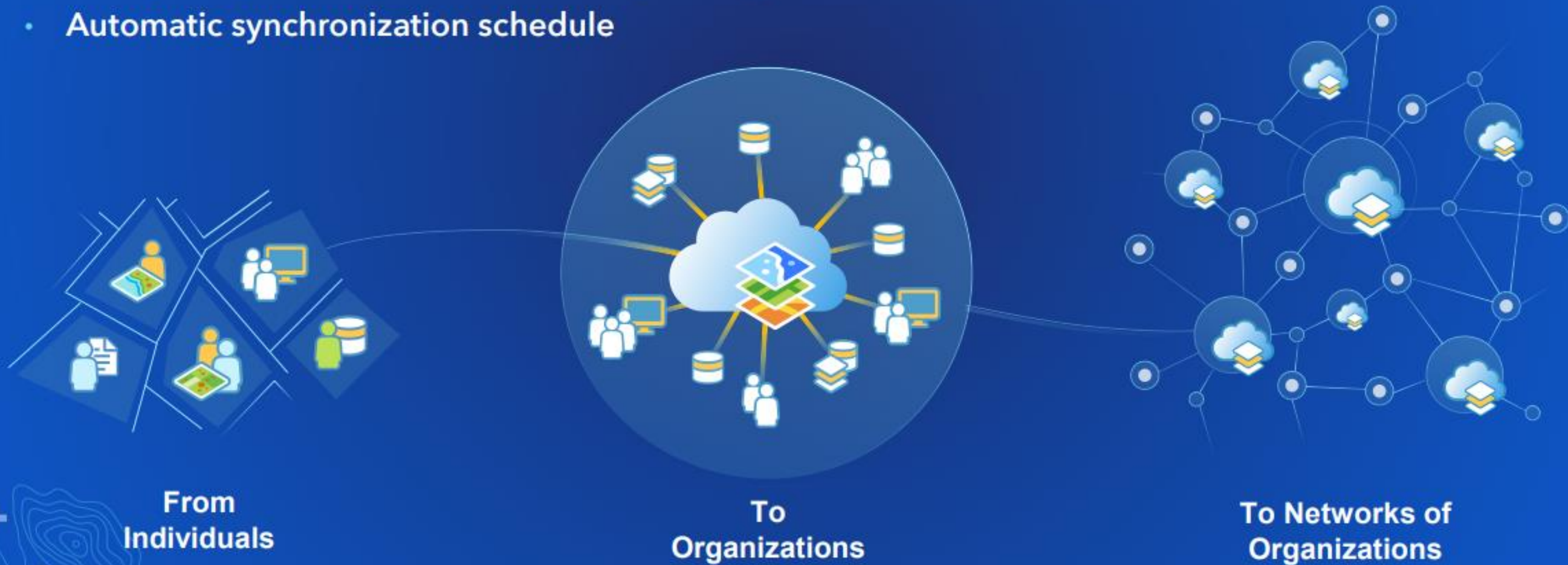


# ArcGIS Enterprise on the Cloud



# Distribute Collaboration

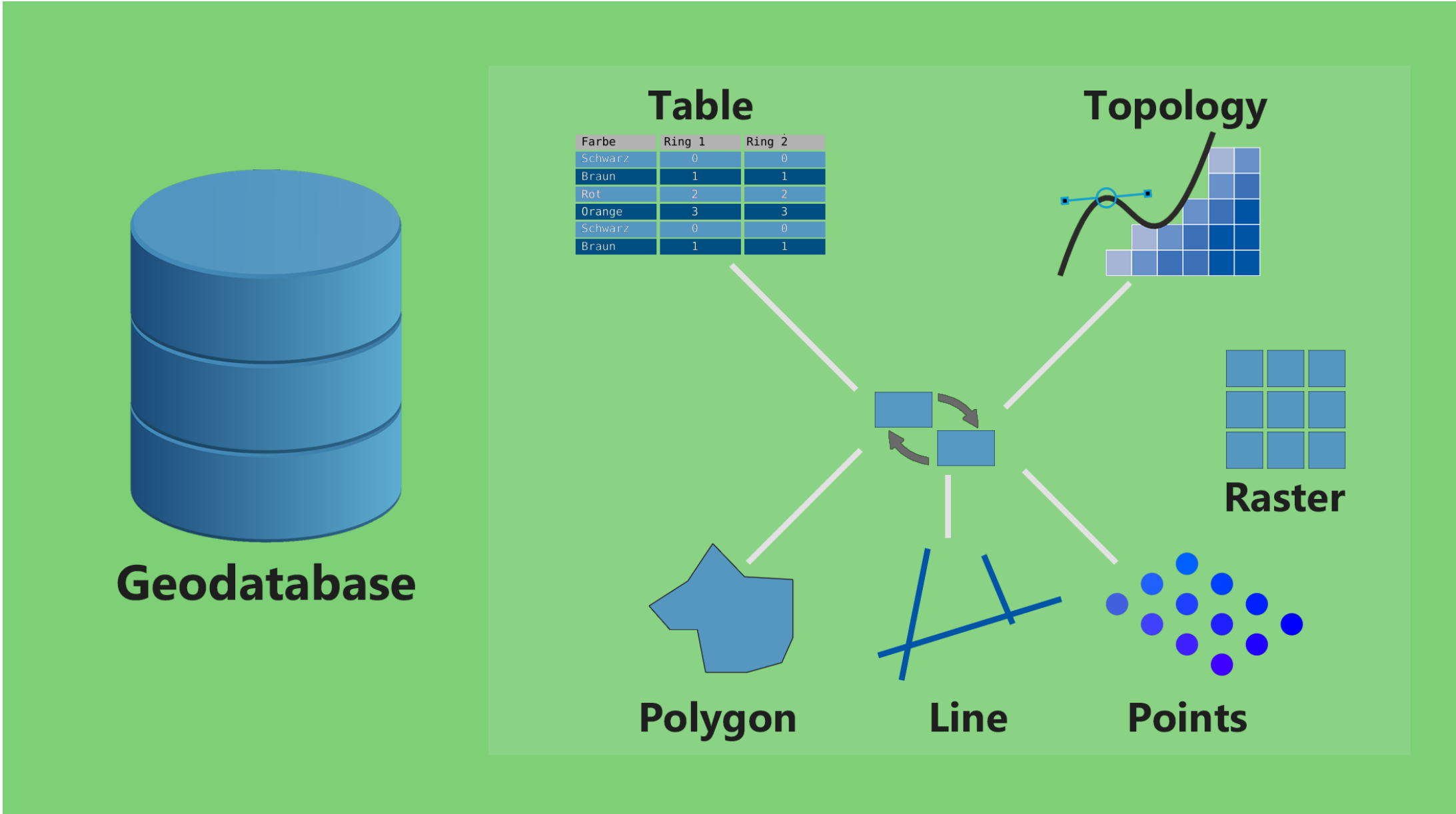
- Sharing content in a secure, trusted pipeline
- Familiar access control - group sharing model
- Each participant keeps its own security settings
- Automatic synchronization schedule



# Open source technology



# Spatial Data Science



# Spatial analysis

Map by Dr. [John Snow](#) of [London](#), showing [clusters](#) of cholera cases in the [1854 Broad Street cholera outbreak](#). This was one of the first uses of map-based spatial analysis.



# Point Pattern Analysis

Objectives of point pattern analysis are:

1. Estimate intensity of point pattern
  - Expected number of events per unit area
  - Enables visualisation of the events
2. Study structural characteristics of the spatial distribution
  - Determine if the events are random, clustered or regular
  - Indicative of a process that is underlying the occurrence of events

# Point Pattern Analysis

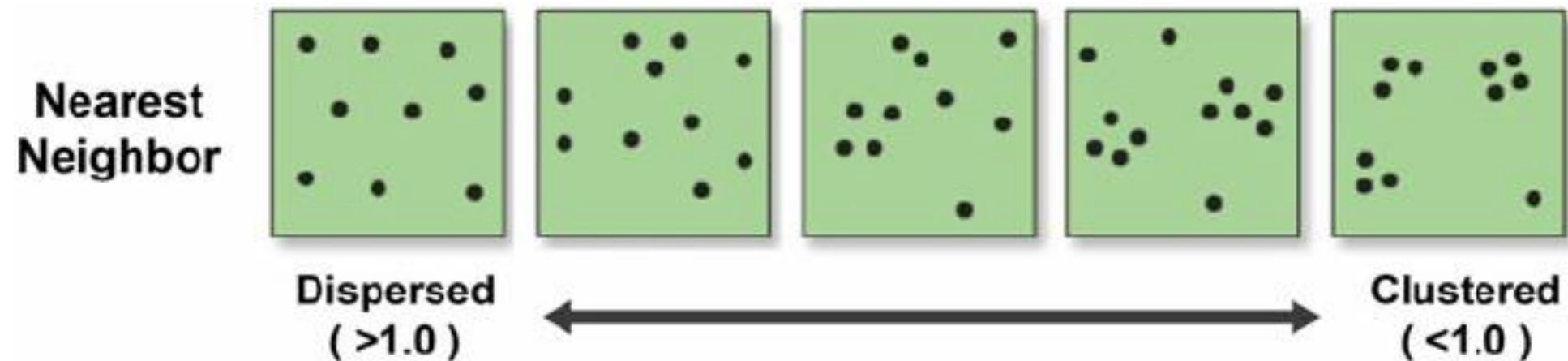
- Is the study of the spatial arrangements of points in space.
- Is the evaluation of the pattern, or distribution, of a set of points on a surface.
- It can refer to the actual spatial or temporal location of these points or also include data from point sources.

# Analysing Point Patterns

- Global and local study
- Looking for outliers & clusters
- Global – study properties for the entire dataset
  - Spatial statistic -> Avg nearest neighbour, Getis Ord,
- Local – looking at behaviour related to neighbours
  - Spatial Visual -> Voroni maps (Outliers, Smoothing, Variation)
  - Spatial statistic -> Local Moran's I (Anselin)

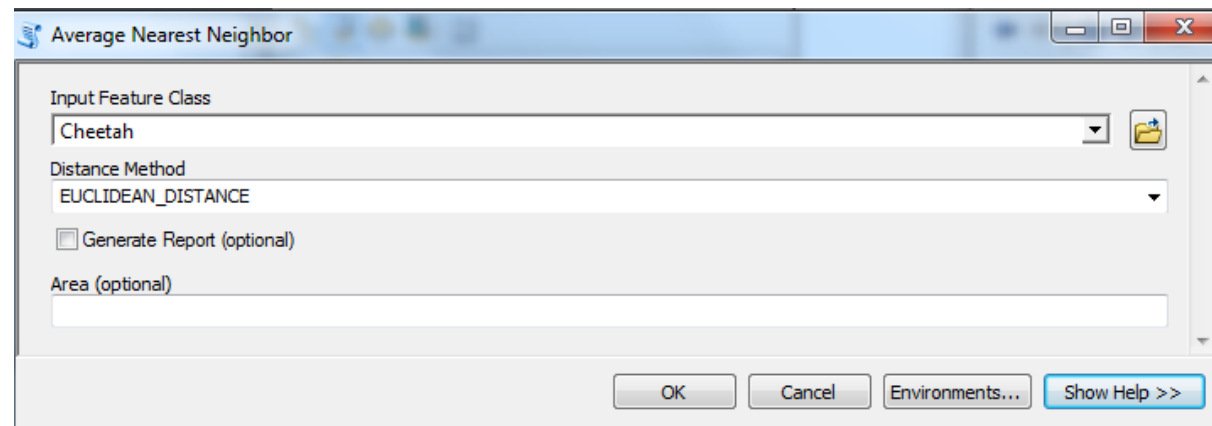
# 1. Average nearest Neighbour

Calculates a nearest neighbour index based on the average distance from each feature to its nearest neighbouring feature.



# 1. Average nearest Neighbour

## Average Nearest Neighbor Summary



Nearest Neighbor Ratio: 0.234537

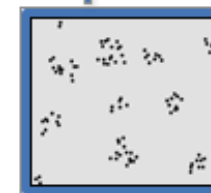
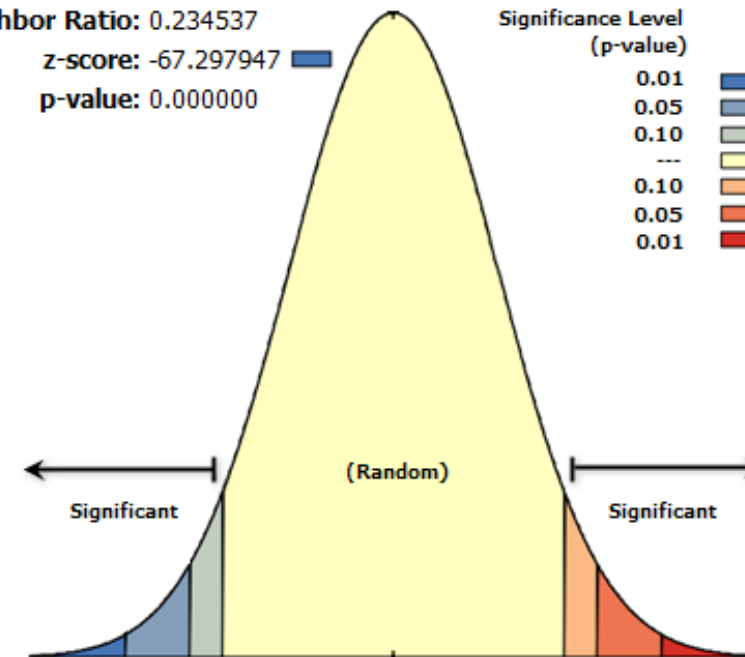
z-score: -67.297947

p-value: 0.000000

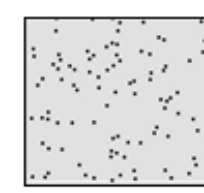
Significance Level  
(p-value)

Critical Value  
(z-score)

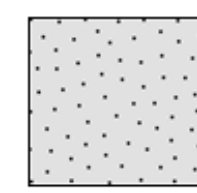
0.01	< -2.58
0.05	-2.58 - -1.96
0.10	-1.96 - -1.65
---	-1.65 - 1.65
0.10	1.65 - 1.96
0.05	1.96 - 2.58
0.01	> 2.58



Clustered



Random



Dispersed

Given the z-score of -67.30, there is a less than 1% likelihood that this clustered pattern could be the result of random chance.

# 1. Average nearest Neighbour

The Average Nearest Neighbor ratio is given as:

$$ANN = \frac{\bar{D}_O}{\bar{D}_E} \quad (1)$$

where  $\bar{D}_O$  is the observed mean distance between each feature and their nearest neighbor:

$$\bar{D}_O = \frac{\sum_{i=1}^n d_i}{n} \quad (2)$$

and  $\bar{D}_E$  is the expected mean distance for the features given a random pattern:

$$\bar{D}_E = \frac{0.5}{\sqrt{n/A}} \quad (3)$$

In the previous equations,  $d_i$  equals the distance between feature  $i$  and its nearest feature,  $n$  corresponds to the total number of features and  $A$  is the total study area.

# 1. Average nearest Neighbour

The  $z_{ANN}$ -score for the statistic is calculated as:

$$z_{ANN} = \frac{\bar{D}_O - \bar{D}_E}{SE} \quad (4)$$

where:

$$SE = \frac{0.26136}{\sqrt{n^2/A}} \quad (5)$$

where:

$$SE = \frac{0.26136}{\sqrt{n^2/A}} \quad (5)$$

## 2. Getis Ord- Hot Spot Analysis

This tool identifies statistically significant spatial clusters of high values (hot spots) and low values (cold spots).



## 2. Getis Ord- Hot Spot Analysis

- This tool works by looking at each feature within the context of neighboring features.
- A feature with a high value is interesting, but may not be a statistically significant hot spot.
- To be a statistically significant hot spot, a feature will have a high value and be surrounded by other features with high values as well.
- The local sum for a feature and its neighbors is compared proportionally to the sum of all features;
- When the local sum is much different than the expected local sum, and that difference is too large to be the result of random chance, a statistically significant Z score results.

## 2. Getis Ord- Hot Spot Analysis

The Getis-Ord local statistic is given as:

$$G_i^* = \frac{\sum_{j=1}^n w_{i,j} x_j - \bar{X} \sum_{j=1}^n w_{i,j}}{S \sqrt{\frac{n \sum_{j=1}^n w_{i,j}^2 - \left( \sum_{j=1}^n w_{i,j} \right)^2}{n-1}}} \quad (1)$$

where  $x_j$  is the attribute value for feature  $j$ ,  $w_{i,j}$  is the spatial weight between feature  $i$  and  $j$ ,  $n$  is equal to the total number of features and:

$$\bar{X} = \frac{\sum_{j=1}^n x_j}{n} \quad (2)$$

$$S = \sqrt{\frac{\sum_{j=1}^n x_j^2}{n} - (\bar{X})^2} \quad (3)$$

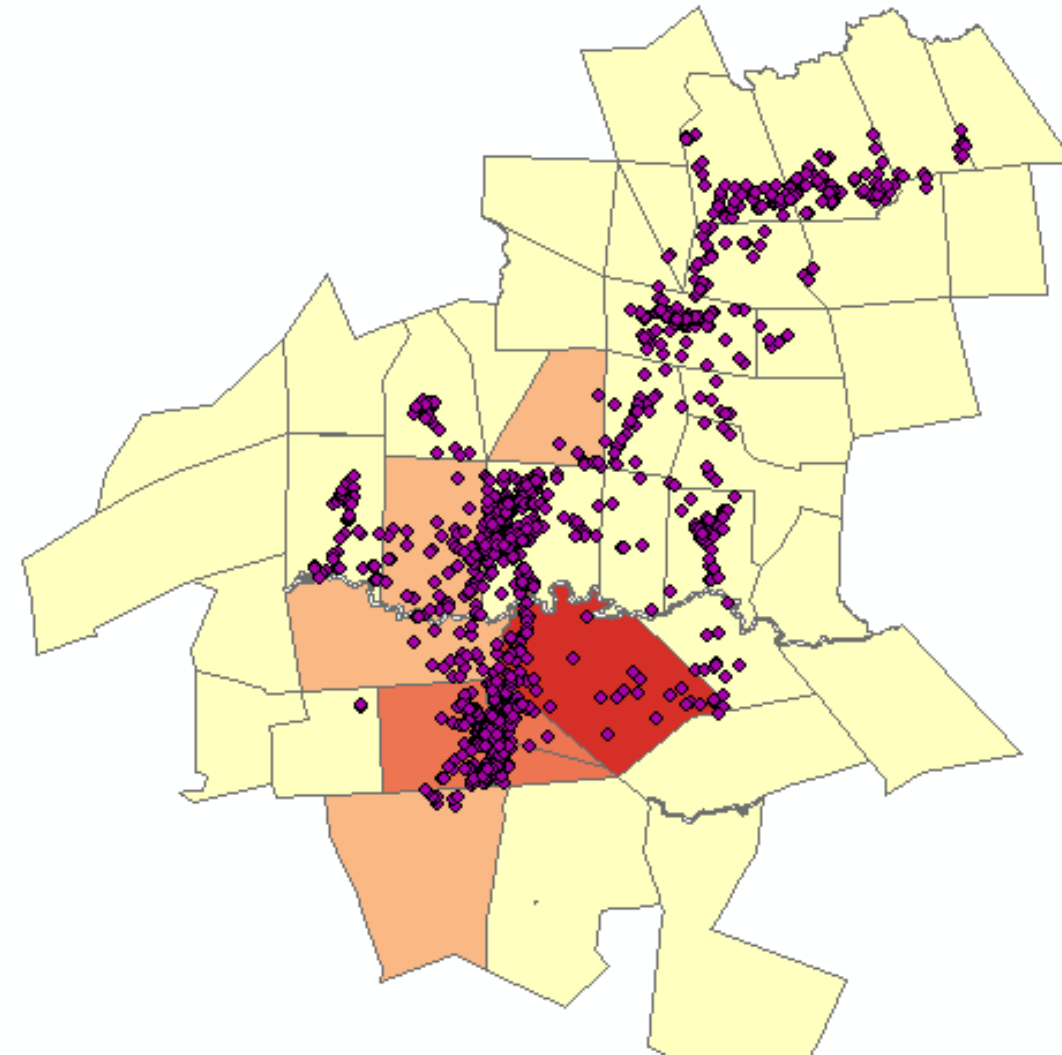
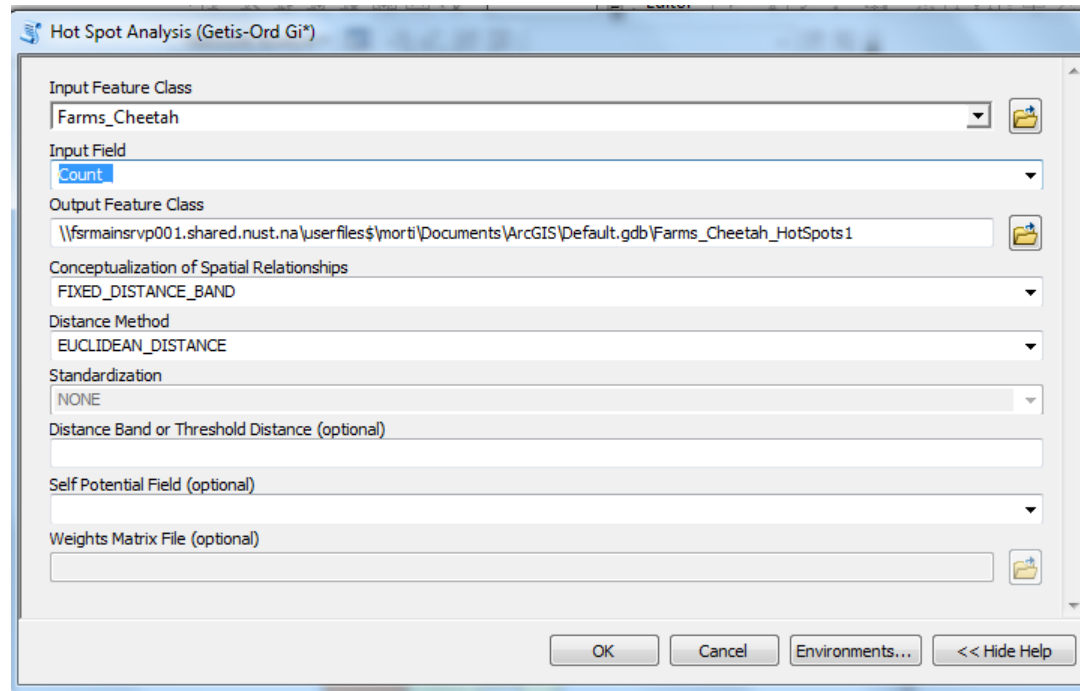
The  $G_i^*$  statistic is a z-score so no further calculations are required.

## 2. Getis Ord- Hot Spot Analysis

### **Potential Applications**

Applications can be found in crime analysis, epidemiology, voting pattern analysis, economic geography, retail analysis, traffic incident analysis, and demographics

# Getis Ord



# Voronoi maps

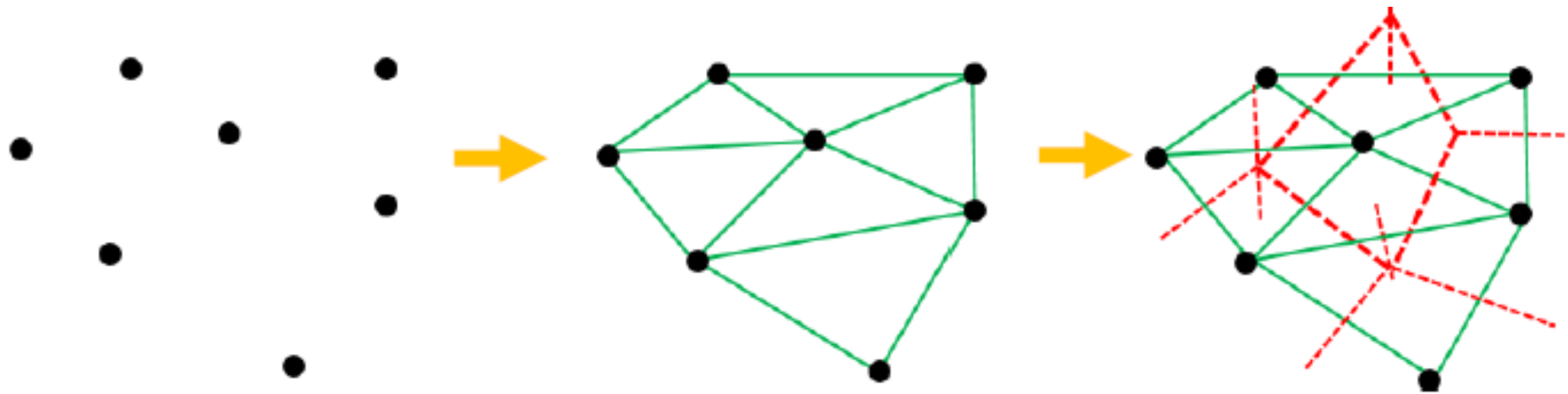
## *Thiessen (Voronoi) polygons*

Polygons are created by joining together the sample points and drawing the perpendicular bisectors of the joining segments. The intersections of these bisectors determine the sides and the number of polygons, so that the perimeters of the generated polygons are equidistant from designated points and the **area of influence**.

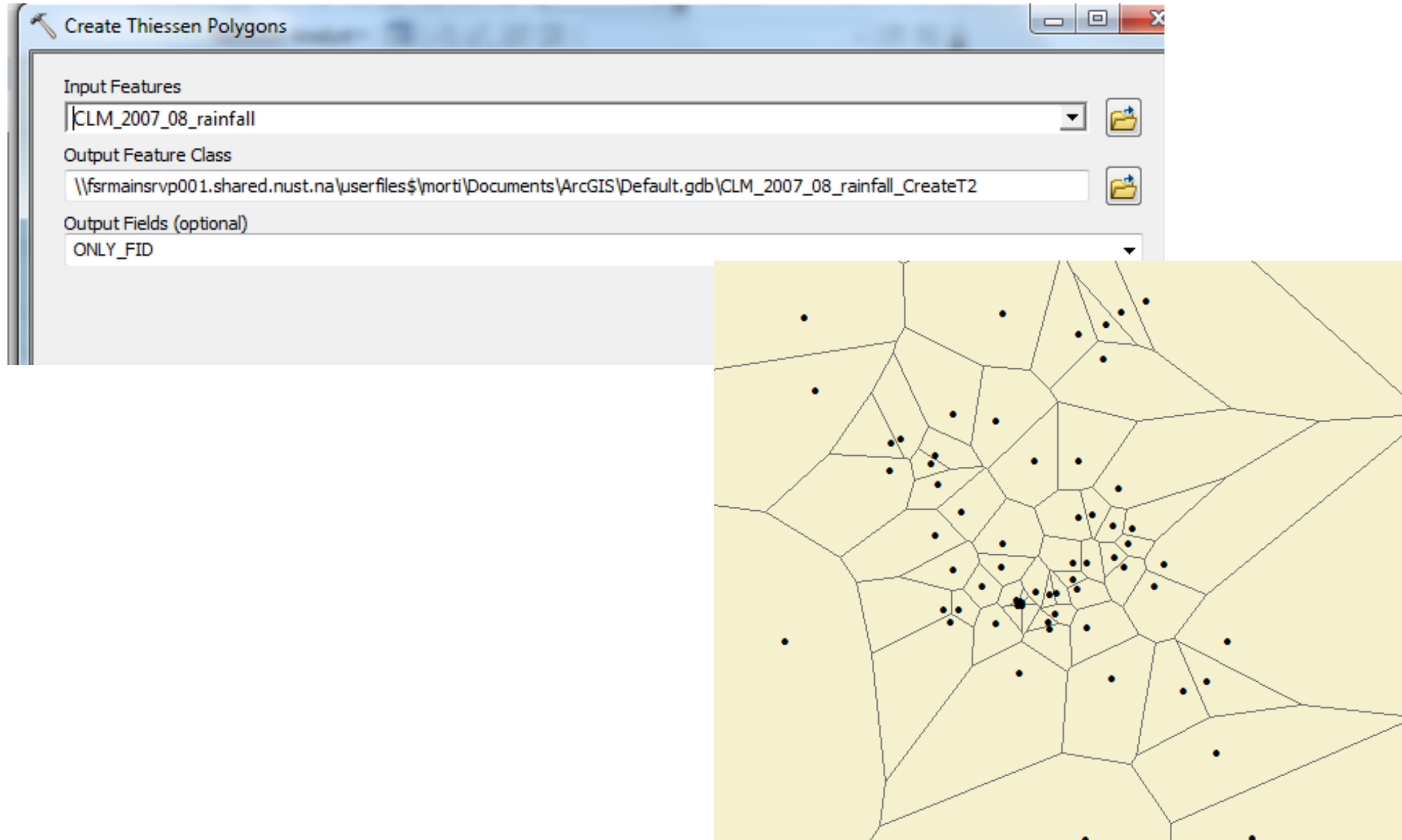
Initially *Thiessen polygons* were used for meteorological data analysis (rainfall stations), but nowadays they are also applied as GIS analytical tool in studies to determine areas of influence (hospitals, fire stations, subway stations, shopping malls, air traffic control, mobile, analysis of populations of plant species, etc.).

# Voronoi Maps

1. Link a point to its neighbouring points by straight lines.
2. Construct the perpendicular in the middle of each line.
3. The polygons defined by the intersection of this perpendicular are the *Thiessen polygons*.



# Voronoi Maps



# Local Morans Anselin

Given a set of weighted features, the Cluster and Outlier Analysis tool identifies **clusters of features with values similar in magnitude**. The tool also identifies spatial outliers. To do this, the tool calculates a Local Moran's I value, a Z score, a p-value, and a code representing the cluster type for each feature. The Z score and p-value represent the statistical significance of the computed index value.



# Local Morans Anselin

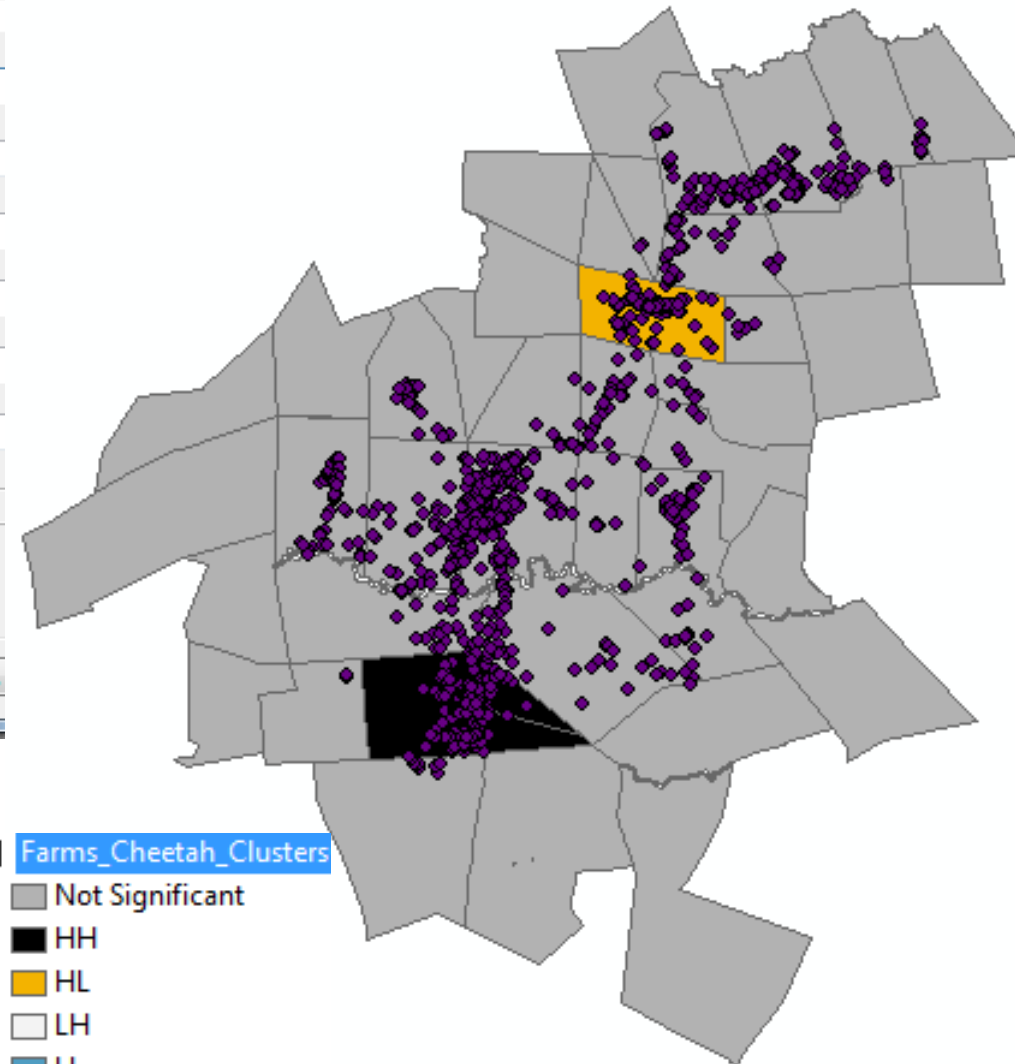
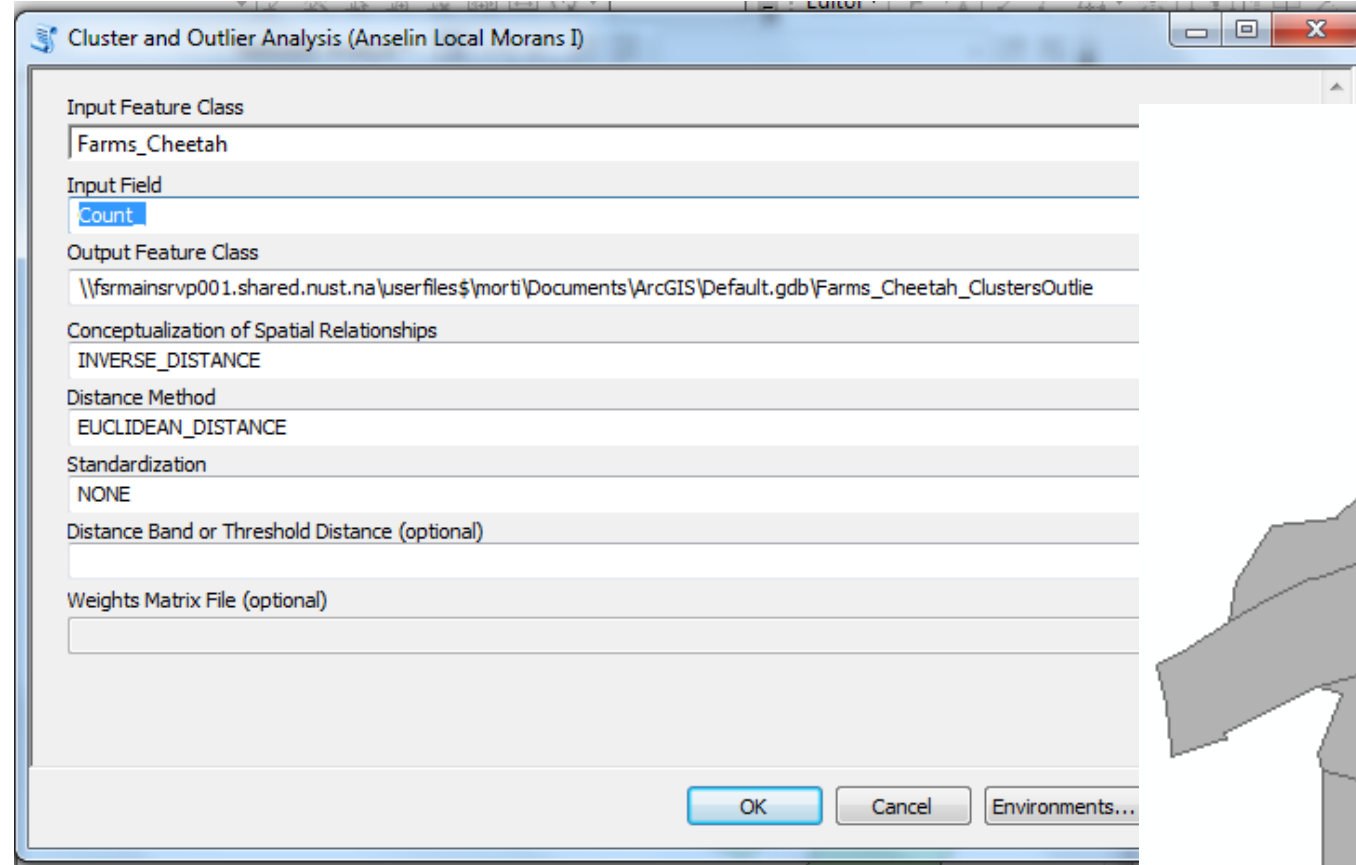
The Local Moran's  $I$  statistic of spatial association is given as:

$$I_i = \frac{x_i - \bar{X}}{S_i^2} \sum_{j=1, j \neq i}^n w_{i,j} (x_j - \bar{X}) \quad (1)$$

where  $x_i$  is an attribute for feature  $i$ ,  $\bar{X}$  is the mean of the corresponding attribute,  $w_{i,j}$  is the spatial weight between feature  $i$  and  $j$ , and:

$$S_i^2 = \frac{\sum_{j=1, j \neq i}^n (x_j - \bar{X})^2}{n - 1} \quad (2)$$

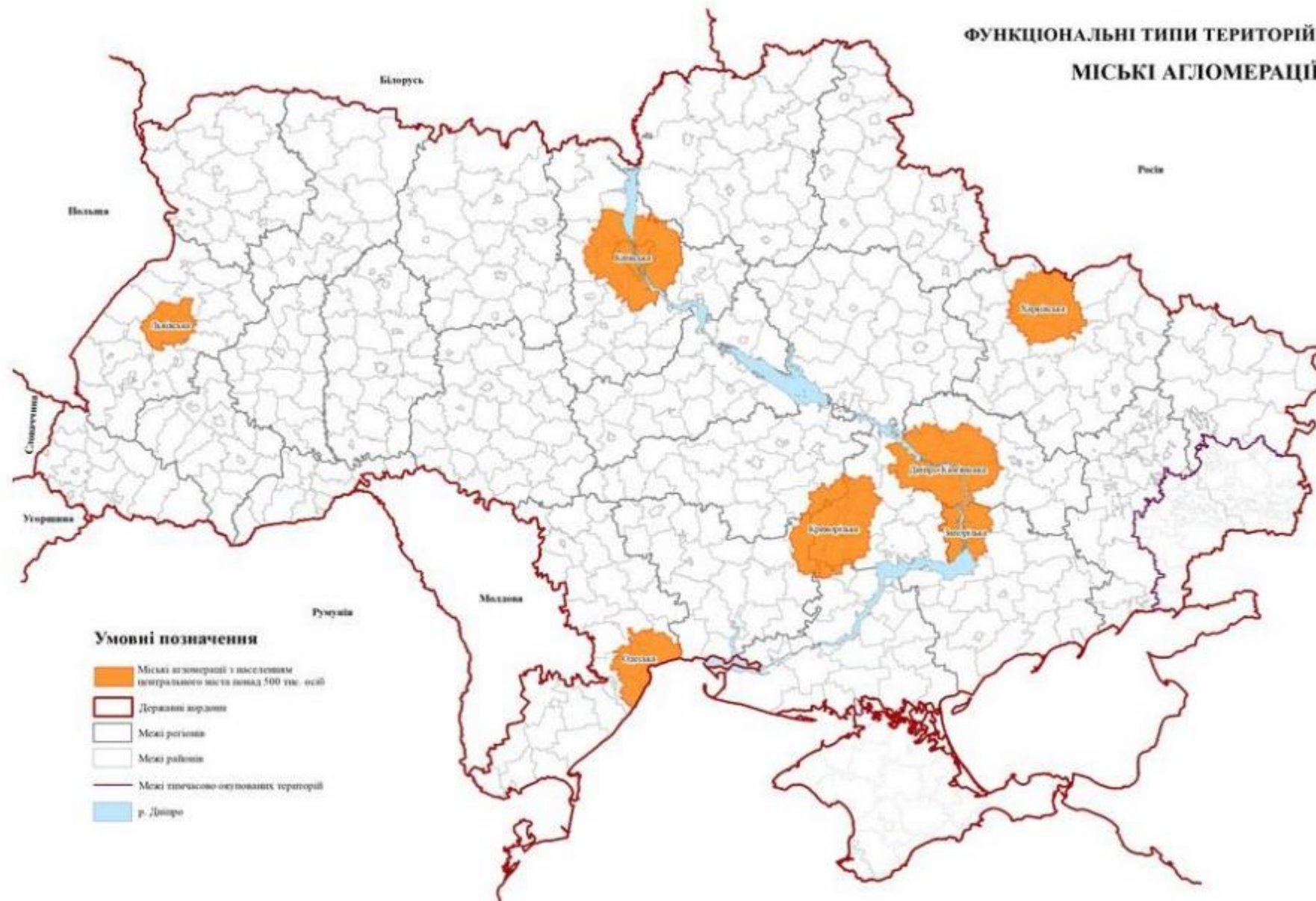
# Local Morans Anselin



# Problems

- Correlation Problem
  - Overlapping samples of  $j$ , similar local statistics.
  - Problem if statistical significance is sought.
- Small Sample Problem
  - Statistics are based on a normal distribution, which is unlikely for a small sample.
- Effects of Global Autocorrelation Problem
  - If there is significant overall global autocorrelation the local statistics will be less useful in detecting “hot spots”.

# Geospatial analysis of Odessa agglomeration



# Criteria for merging into an agglomeration

- mass labor, educational, household, cultural and recreational trips (pendulum migrations)
- 1.5-hour accessibility by transport corridors (railways, roads and rivers)
- availability of regular suburban electric trains, buses, and steamships
- locating subordinate settlements within their administrative regions except for the most closely adjacent ones the community of the airport, railway hub-terminal, dense settlement along transport corridors

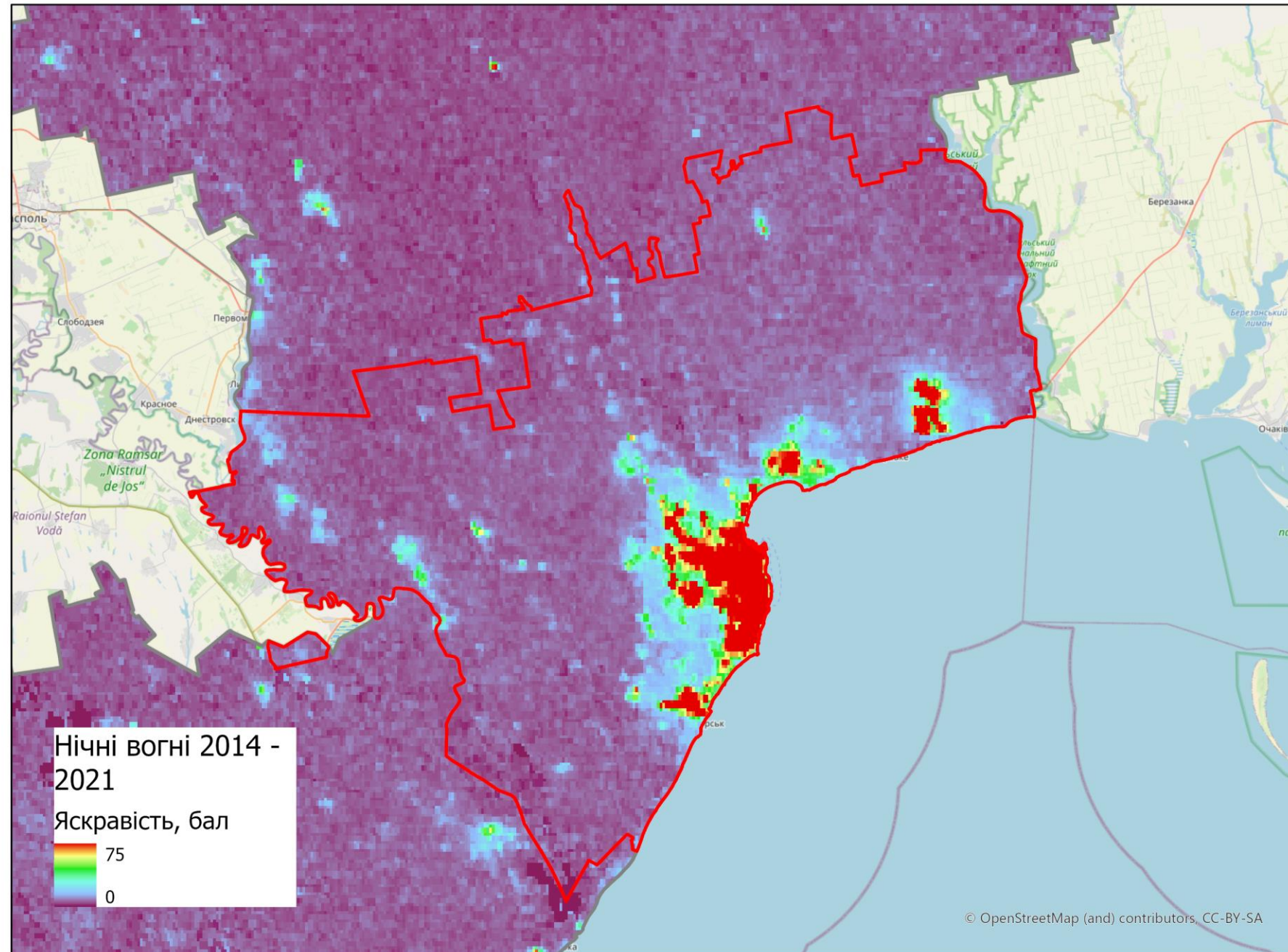
# The main factors of the formation of the Odesa agglomeration

- Historical industrial retrospective of the settlement system
- Transport and logistics support
- Seaside location
- Recreation and tourism

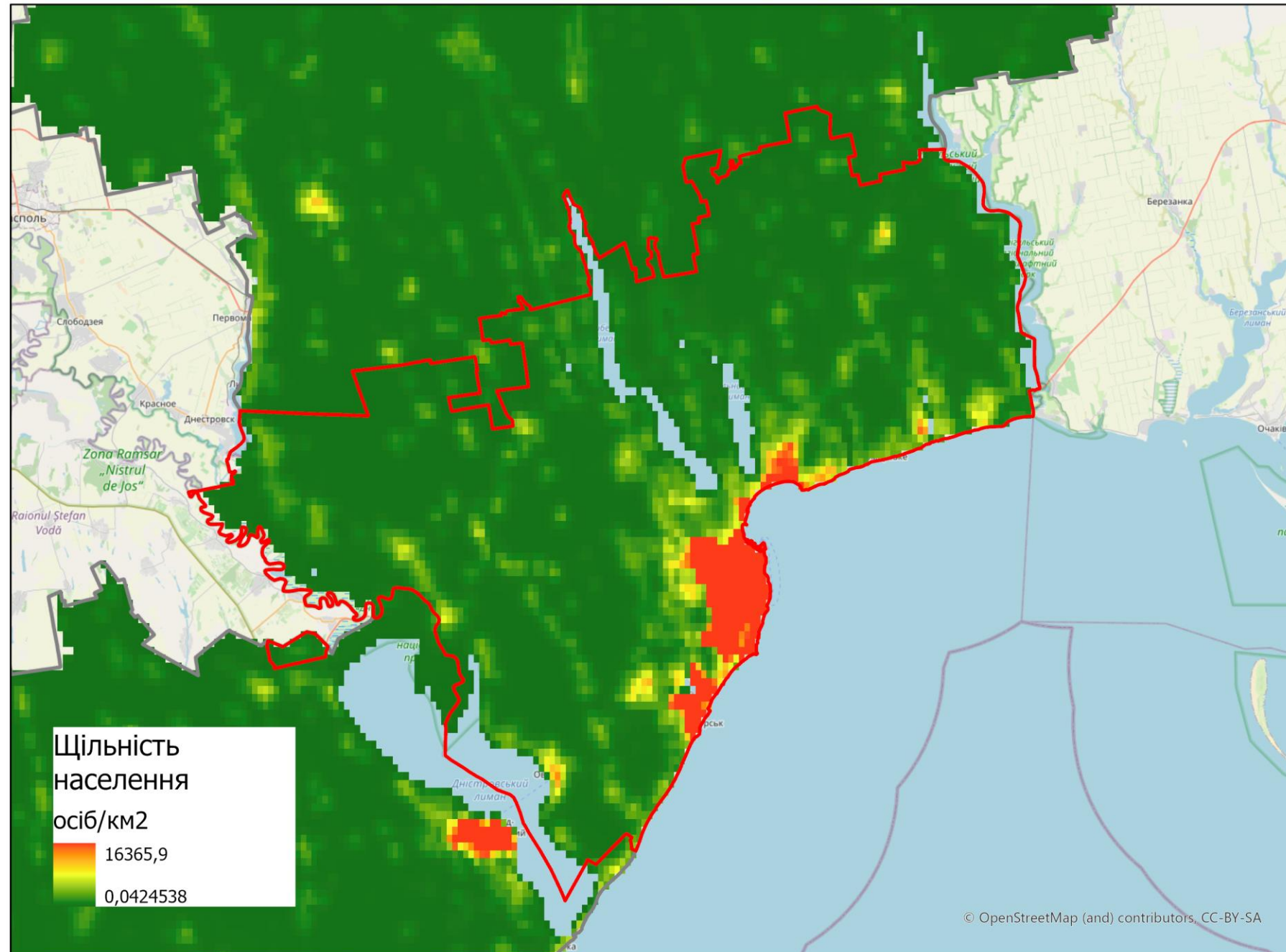
# Component evaluations of the agglomeration effect

- Density of mobile coverage of the territory
- Brightness of night lighting for the historical period 2014 – 2021
- Average value of population density, persons/km<sup>2</sup>
- Saturation of the territory with social service institutions: educational, medical, cultural institutions and provision of administrative services.
- Transport accessibility of the territory: 20, 40, 60, 80 minutes. Industrial zones

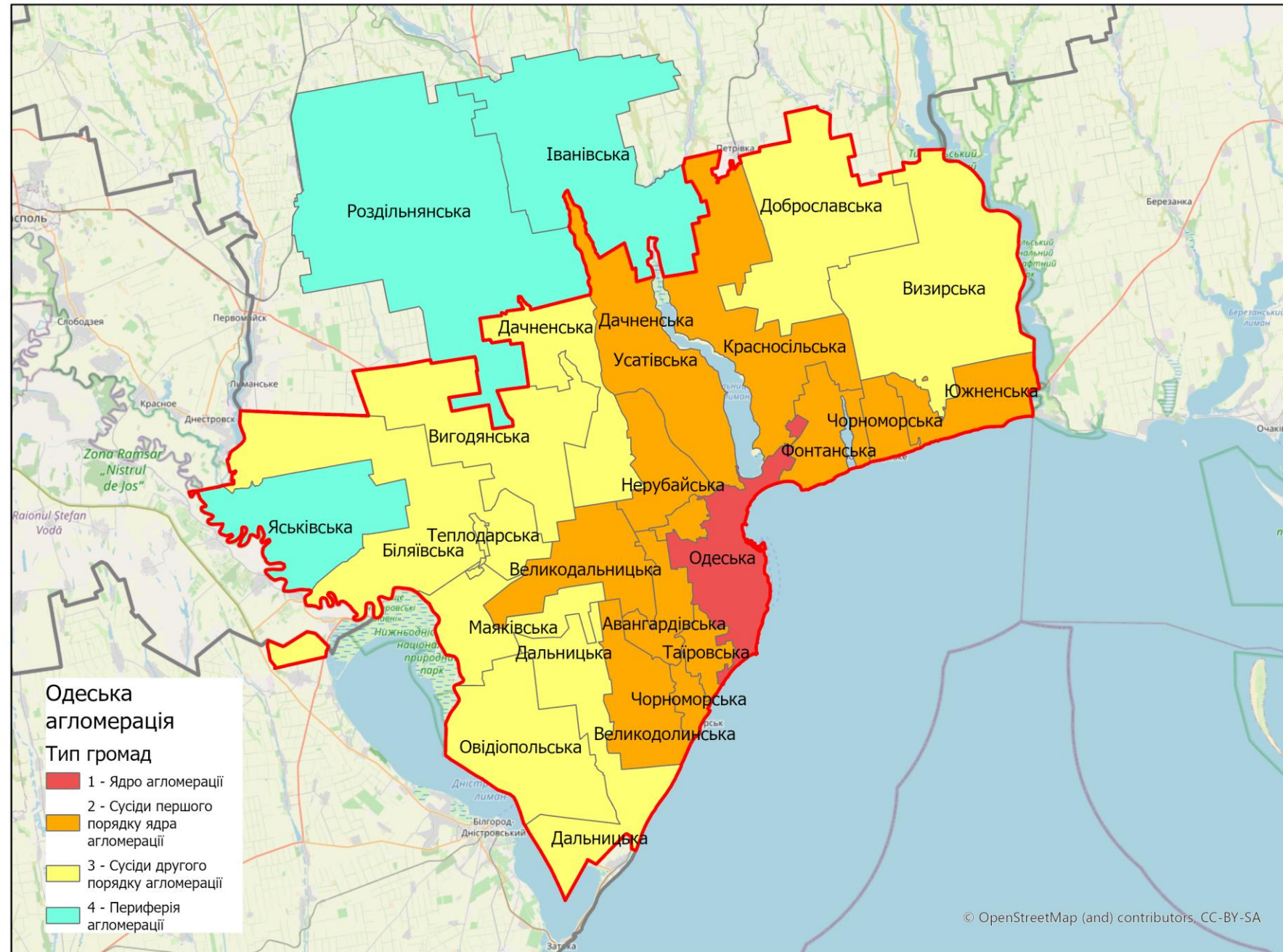
# Brightness of night lighting for the historical period 2014 - 2021



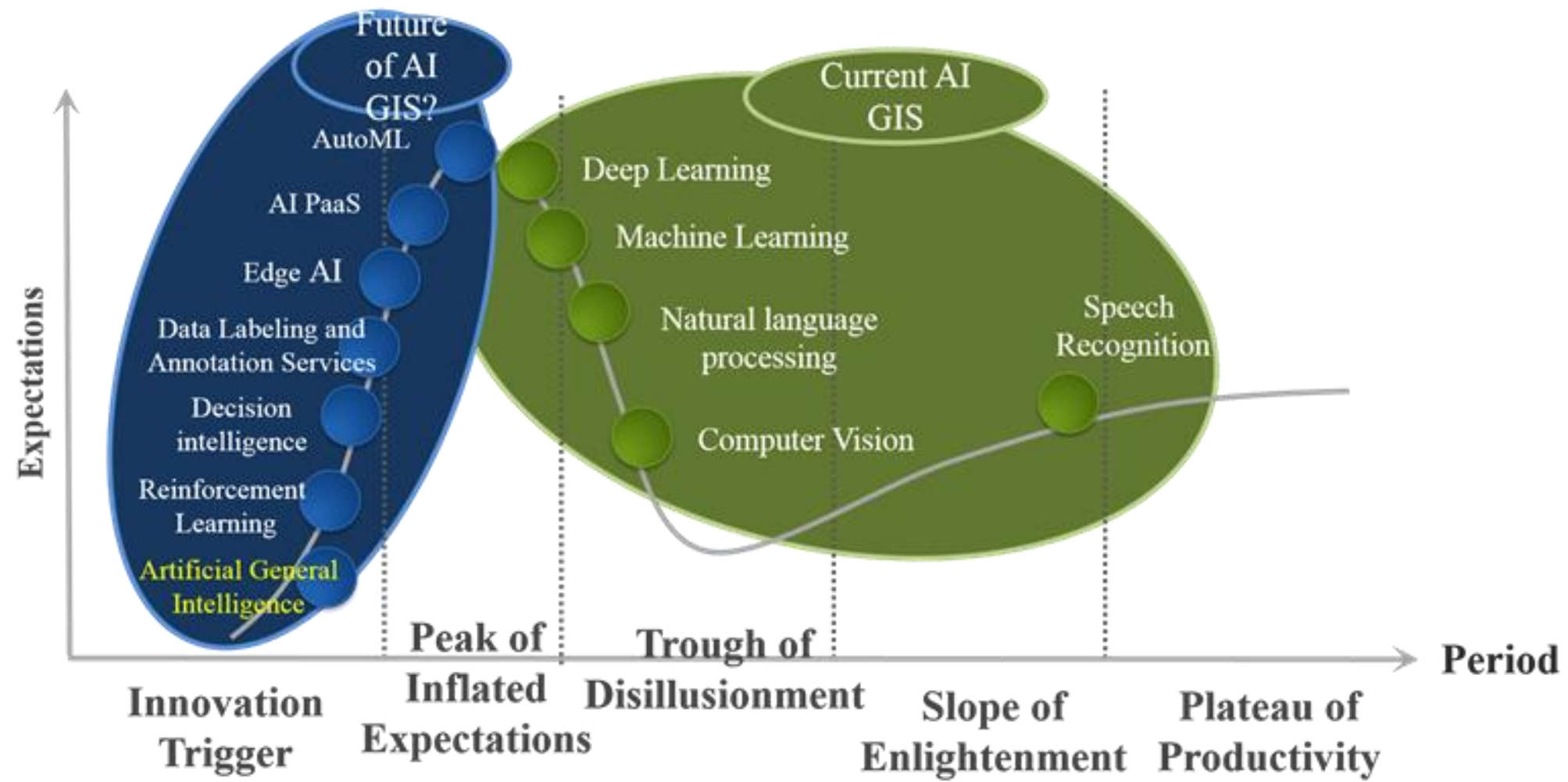
# Average value of population density, persons/km<sup>2</sup>



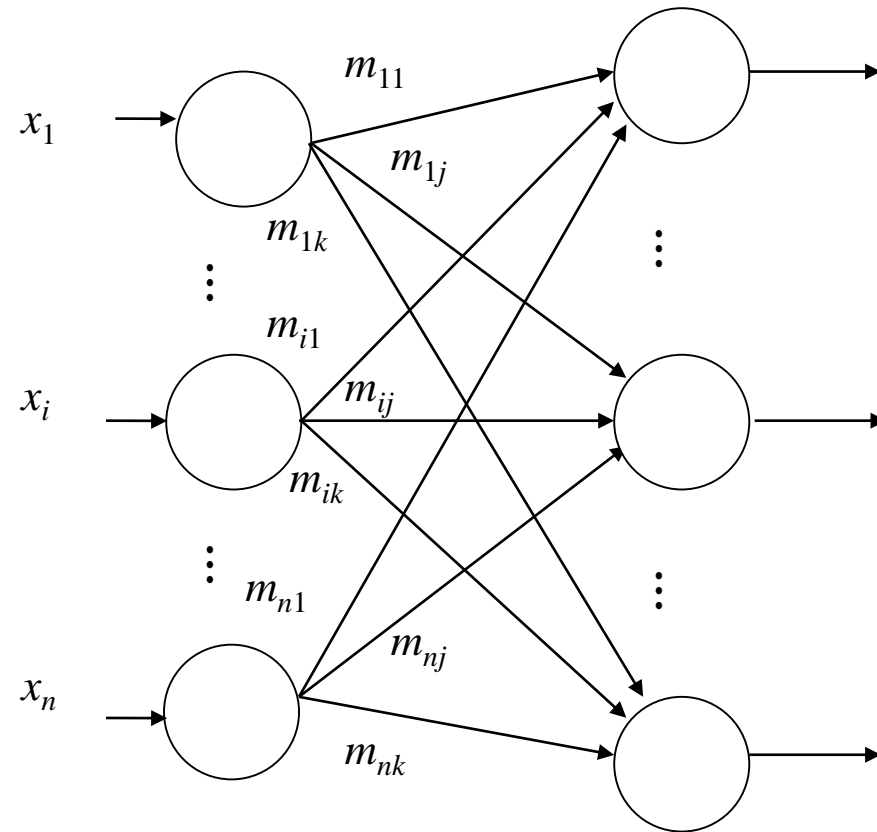
# Typification of communities of Odesa region



# AI in GIS



# Modification of SOM methods for geospatial data processing



Using Kohonen self-organization maps as neural networks for geospatial data:

$$\|x - m_i\| = \sqrt{\sum_{j=1}^n (x_j - \mu_{ji})^2}, \quad c = \arg \min_{i=1,k} \{\|x - m_i\|\},$$

$t$  - номер ітерації;

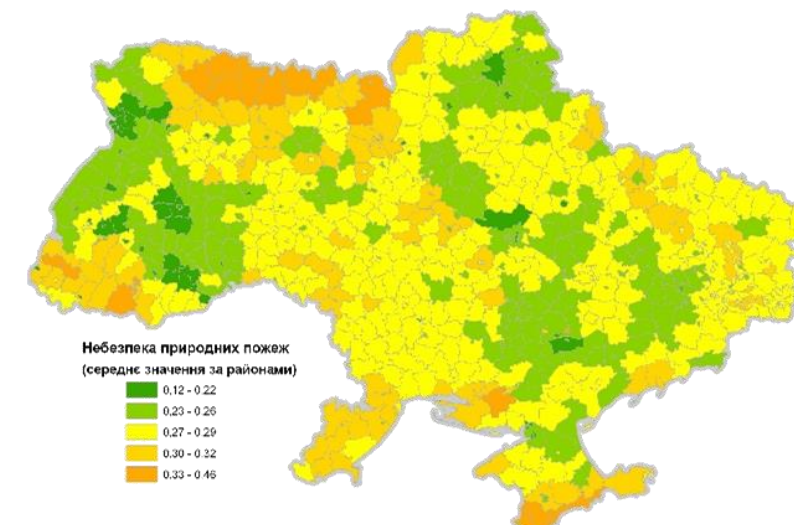
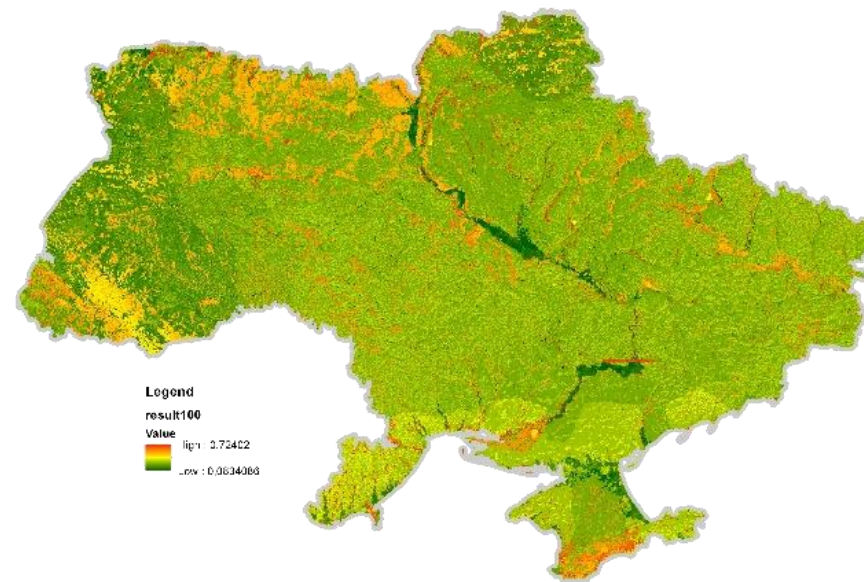
$x(t)$  - випадкове спостереження.

де  $\|x(t) - m_c(t)\| \leq \|x(t) - m_i(t)\|,$

$$h_{ci}(t) = \alpha(t) \cdot \exp\left(-\frac{\|r_c - r_i\|^2}{2\sigma^2(t)}\right),$$

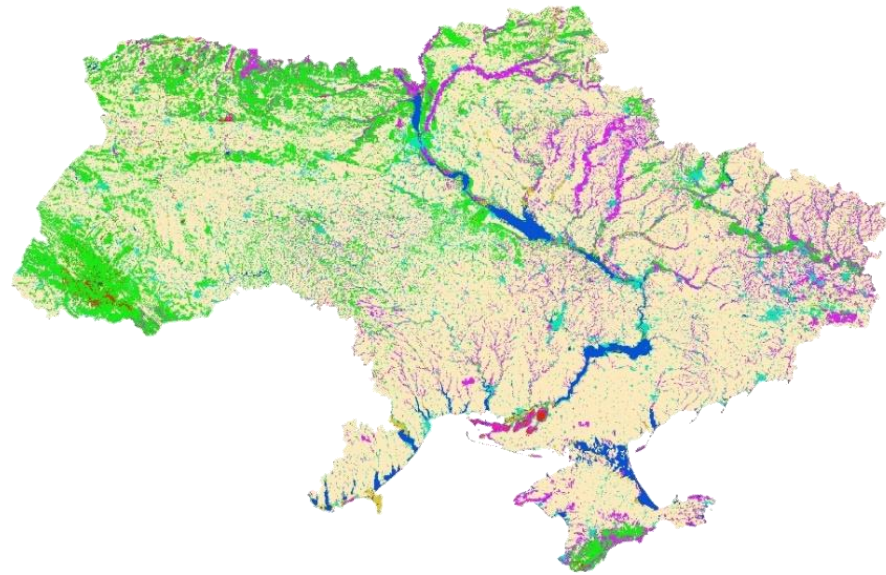
$0 < \alpha(t) < 1$  навчаючий саомножинник.

Typology according to the danger of natural fires



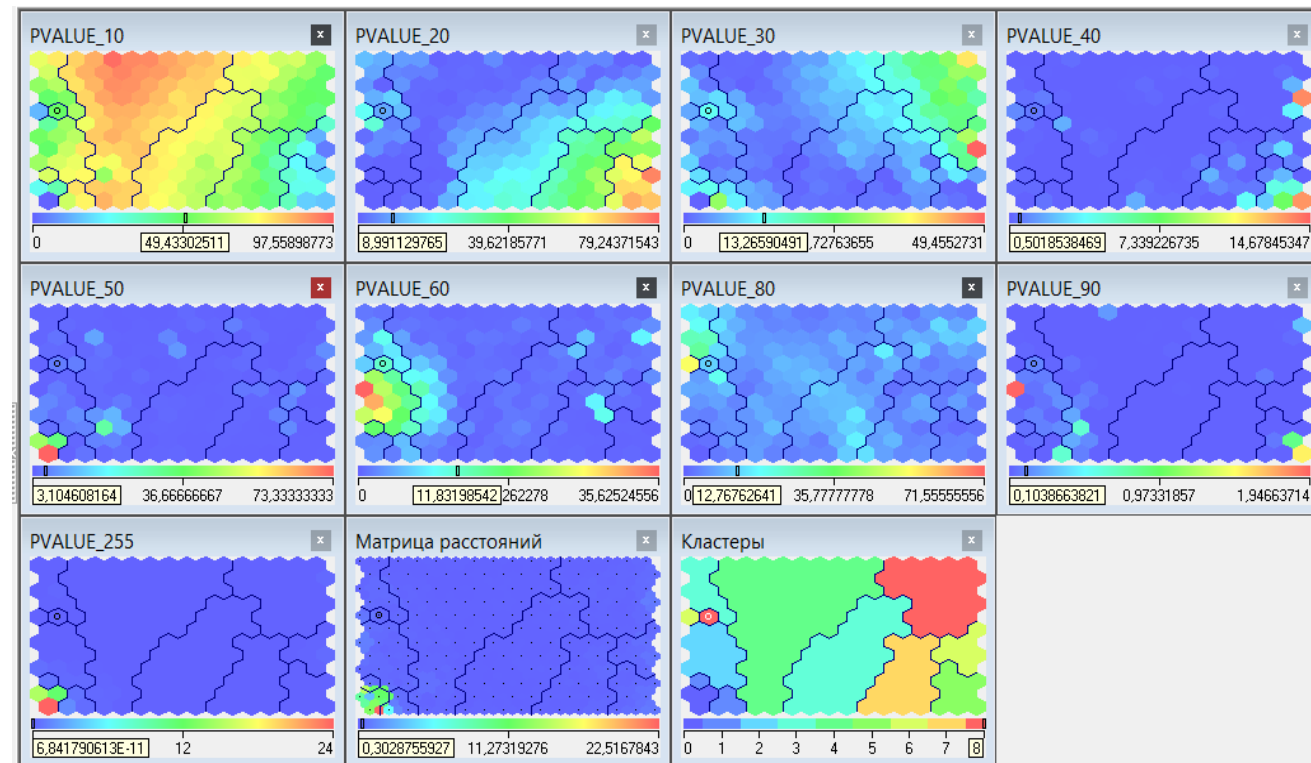
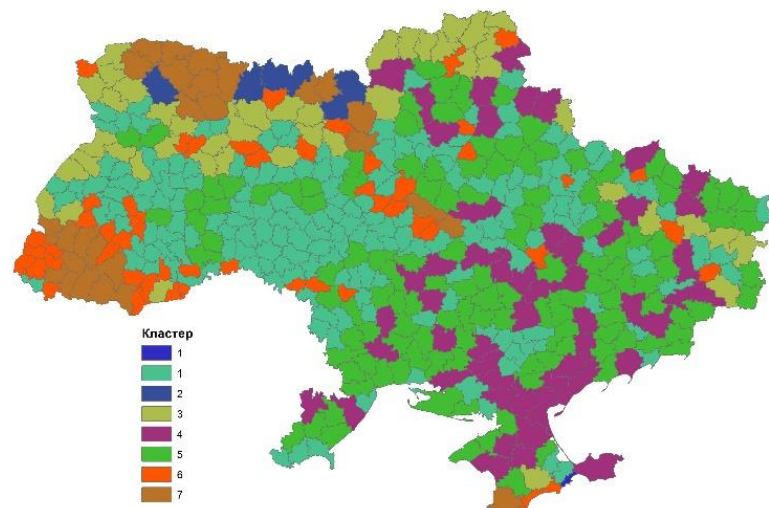
# Clustering of administrative units according to parameters of sustainable land use

Classification of land cover of Ukraine



A set	Percentages	Feature	Error
Study sample	90	445	< 0,05
Test sample	10	50	< 0,05
Total	100	495	0,05

Typology of administrative units

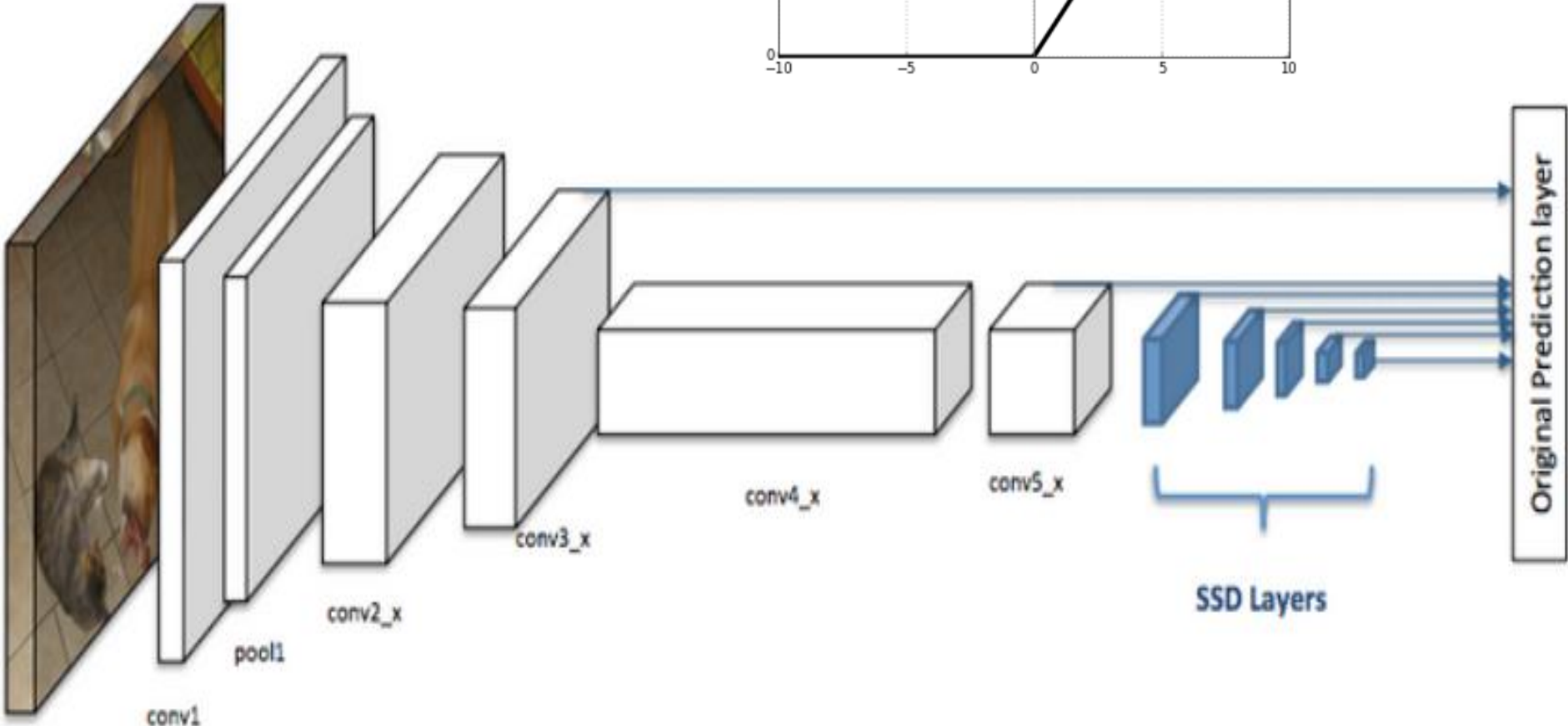
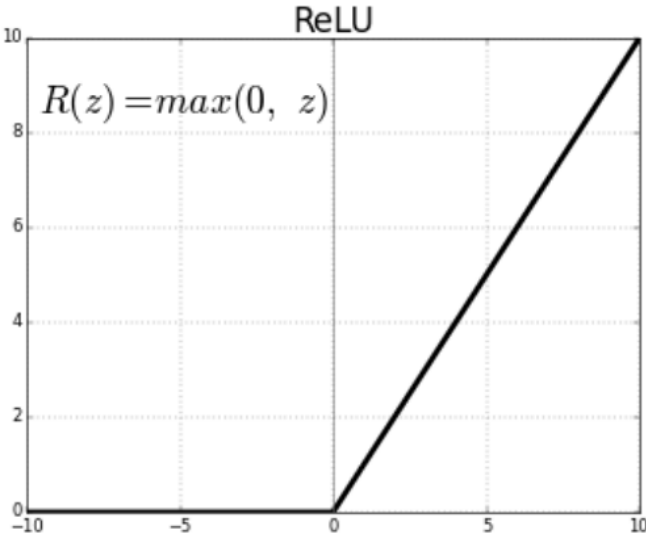


# Convolutional neural network

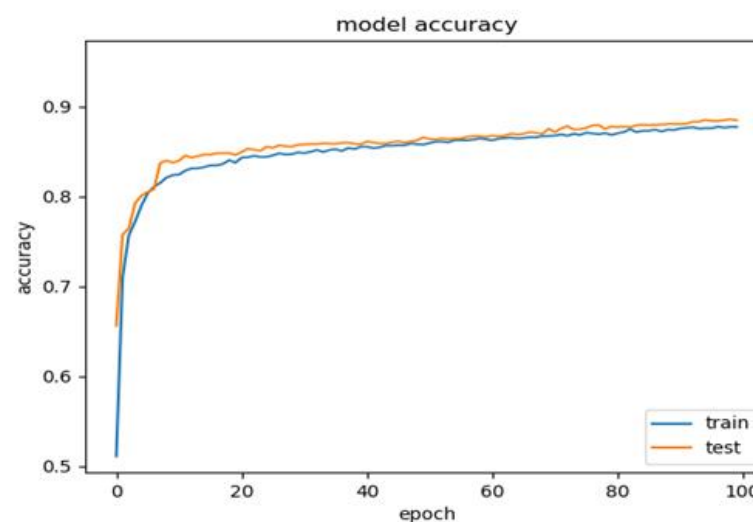
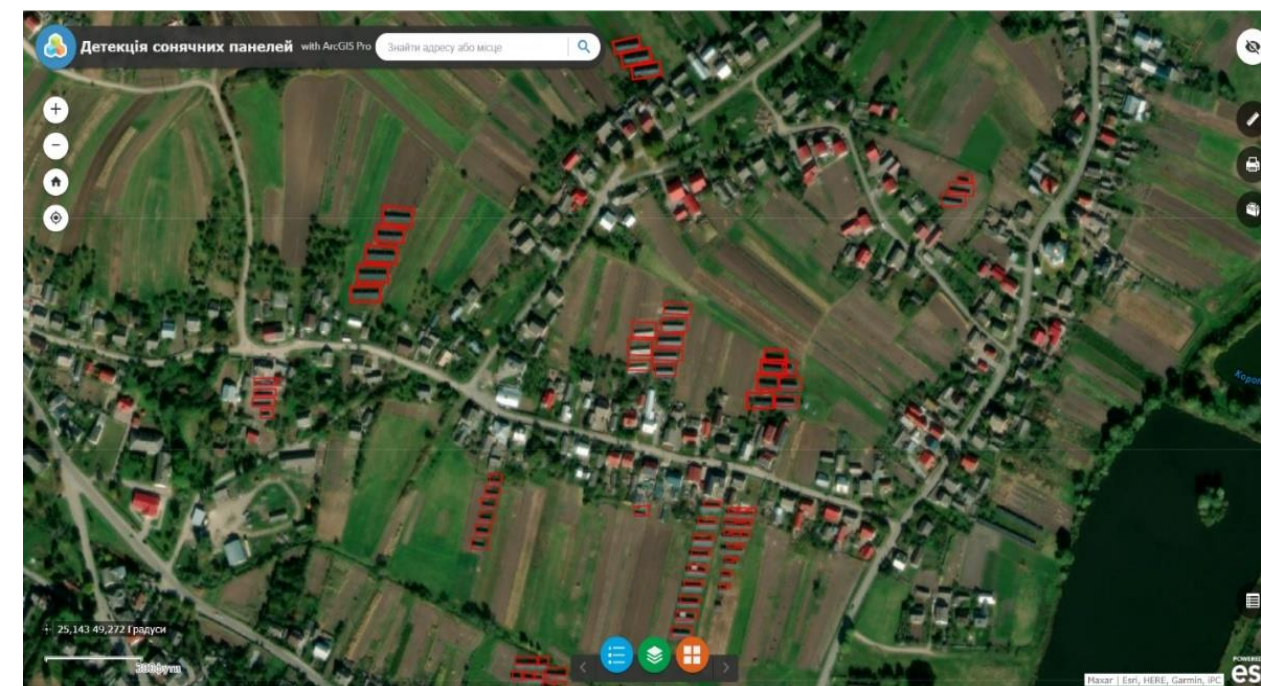
Function softmax

$$f_{ReLU} = \max(0, x)$$

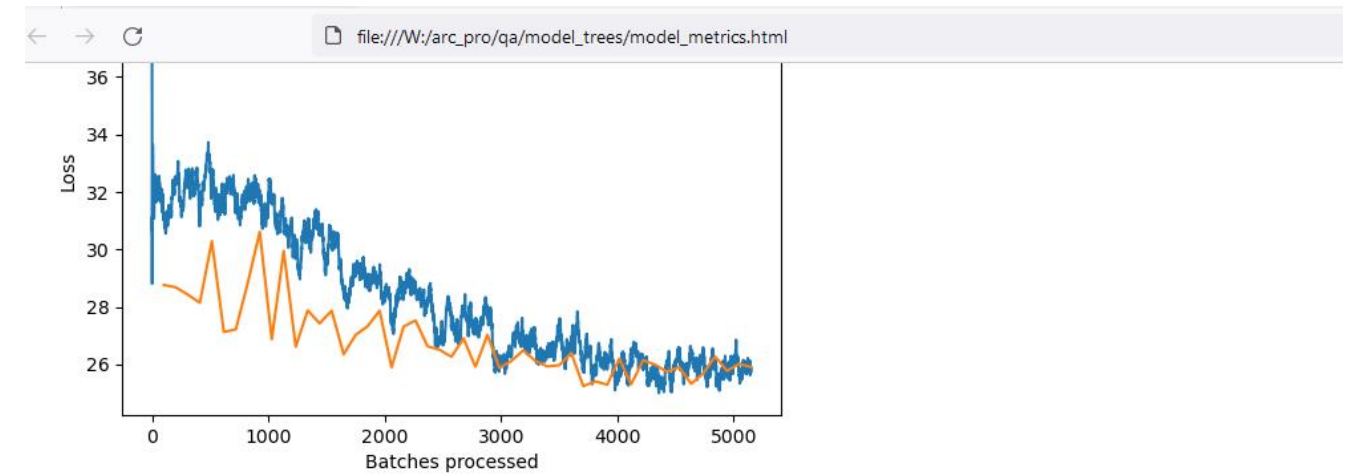
$$f_{softmax} = y_i^l = f(x_i^l) = \frac{e^{x_i^l}}{\sum_{k=0}^n e^{x_k^l}}$$



# Recognition of solar batteries



# Recognition of trees

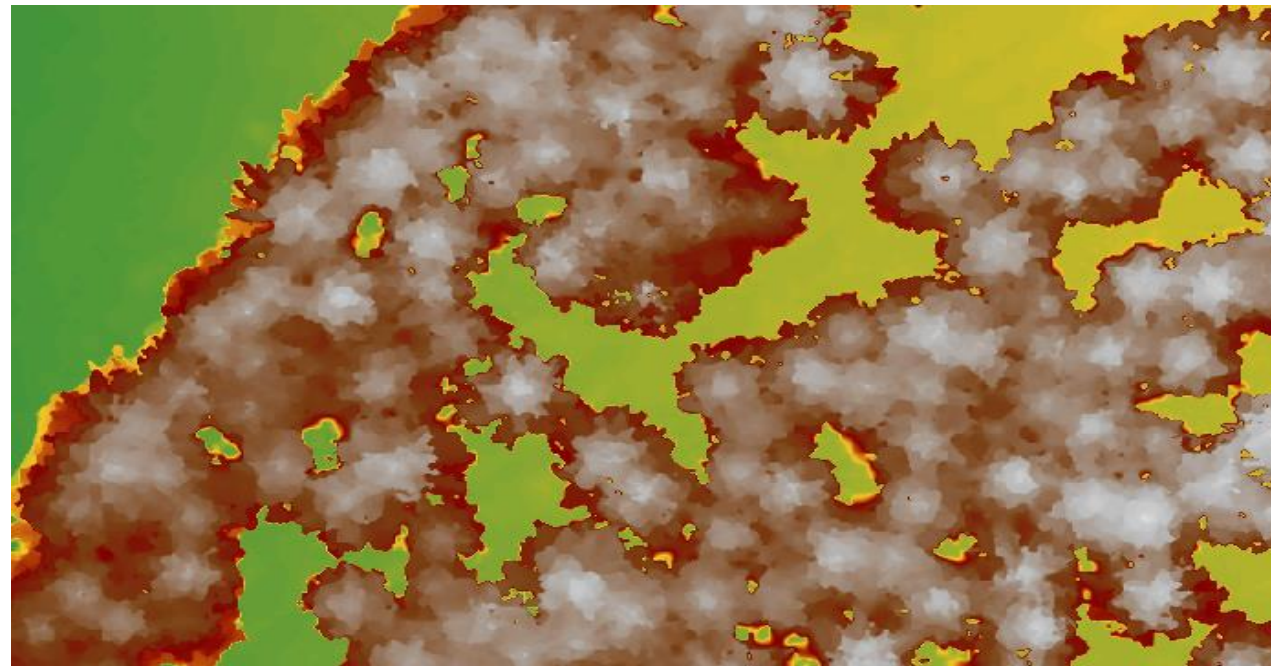
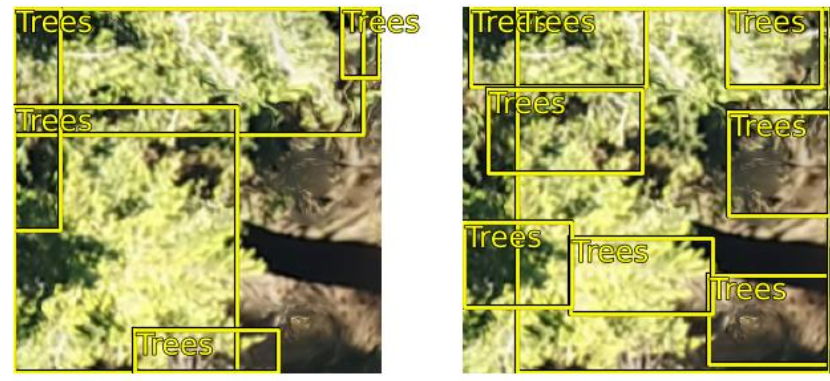


Analysis of the model

Average Precision Score: {"Trees": 0.4310948298285644}

Sample Results

Ground Truth / Predictions

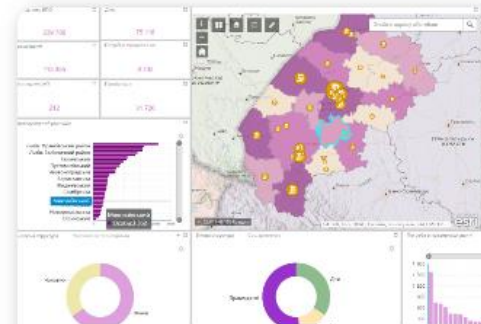


# Visual Analytic System

## Аналітичний портал Львівщини



Дашборд інвестиційної  
привабливості



Дашборд внутрішньо  
переміщених осіб

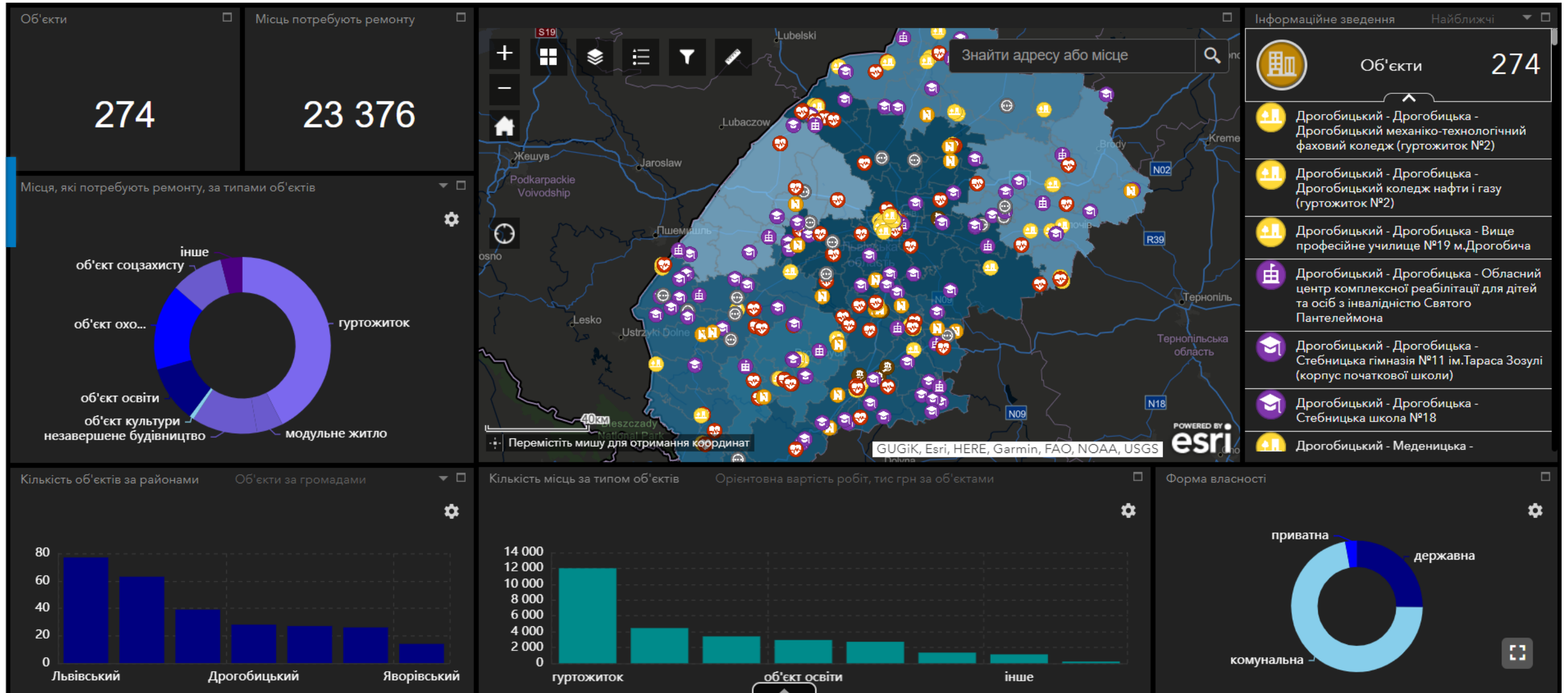


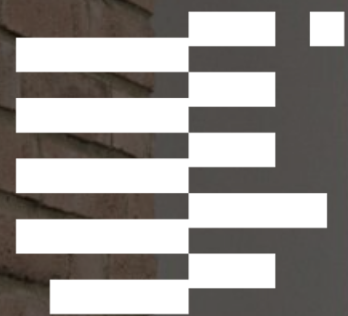
Дашборд громад  
Львівської області



Формування фондів житла  
для ВПО

# Dashboard of social services in Lviv region





American  
University  
Kyiv

Powered by  
Arizona State University

THANK YOU