technische universität
dortmund

地理信息系统概论

孟禹弛

24.10.2022

地理信息系统概论

课程大纲

1. GIS 基础
2. 空间数据
3. 空间参考坐标系
4. 基本空间分析
5. 数据可视化

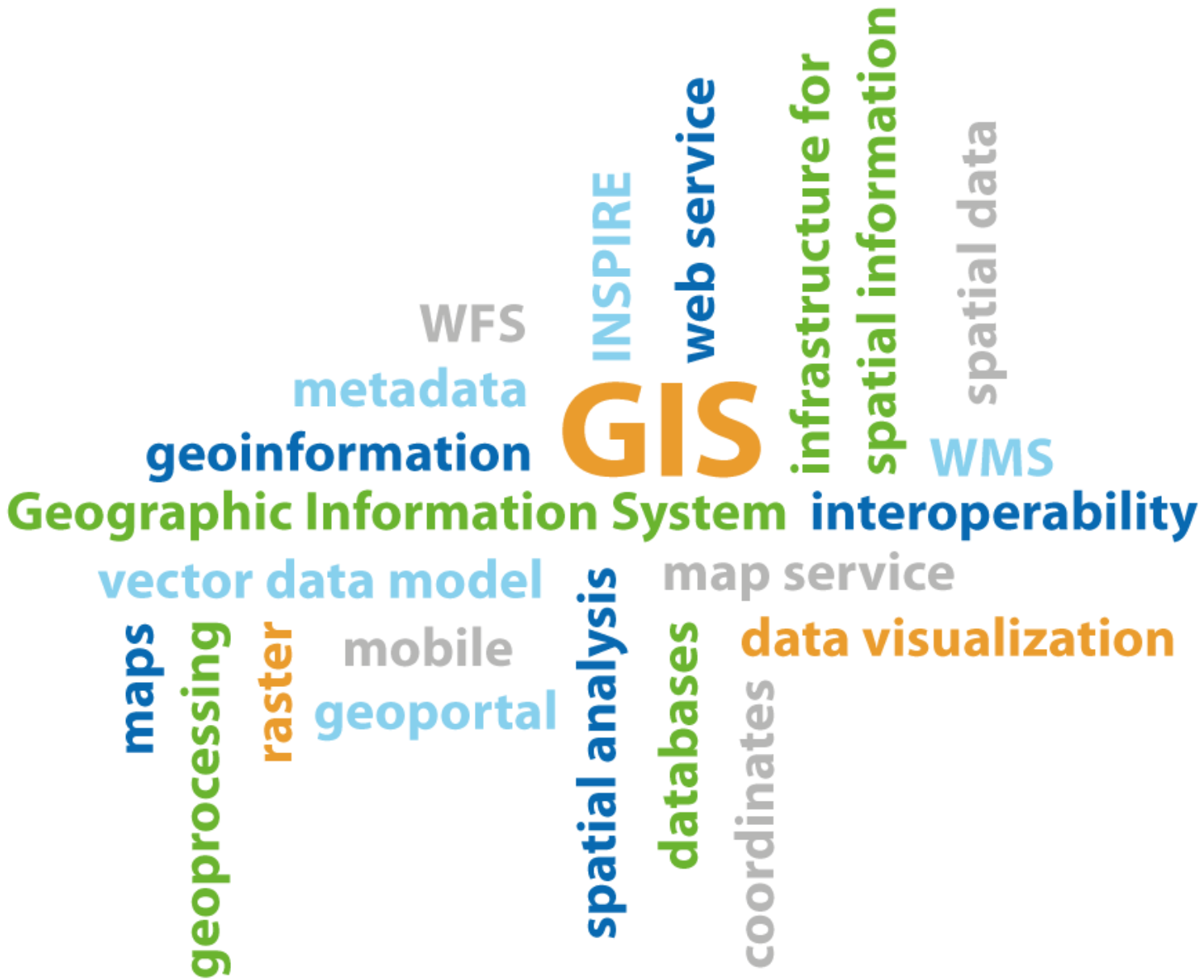
教学目的与课程重点:

1. Understand the basic concepts of Geographic Information Systems 理解地理信息系统相关基础概念、基础理论、基本原理、方法
2. Define terms related to raster and vector data models 定义矢量与栅格数据模型的基本概念
3. Compare vector and raster data models 对比矢量与栅格数据模型
4. Understand the difference between geographic and projected coordinate systems 理解地理坐标系与投影坐标系的区别
5. Select spatial objects using attribute and spatial queries 使用属性值查询空间目标
6. Perform simple analysis with geoprocessing tools 使用地理处理工具进行空间分析
7. List map elements and basic principles of map creation 列举基本地图要素与地图制作的基本原理
8. Create a thematic map using different methods of symbolization 使用不同符号化方法制作一幅主题图

1 | GIS 基础

- 1.1 什么是GIS?
- 1.2 GIS的应用

1.1 | 什么是GIS?



1. https://eo4geo.sbg.ac.at/UNEP-Grid/Introduction-to-GIS/gis.png

理解GIS

Geographic Information System

"is a computer system capable of capturing, storing, analyzing, and displaying geographically referenced information; that is, data identified according to location. Practitioners also define a GIS as including the procedures, operating personnel, and spatial data that go into the system." [1]

"is a computer-based tool for mapping and analyzing things that exist and events that happen on earth. GIS technology integrates common database operations such as query and statistical analysis with the unique visualization and geographic analysis benefits offered by maps." [2]

The spatial (geographic) part differentiates a GIS from a standard computer database.

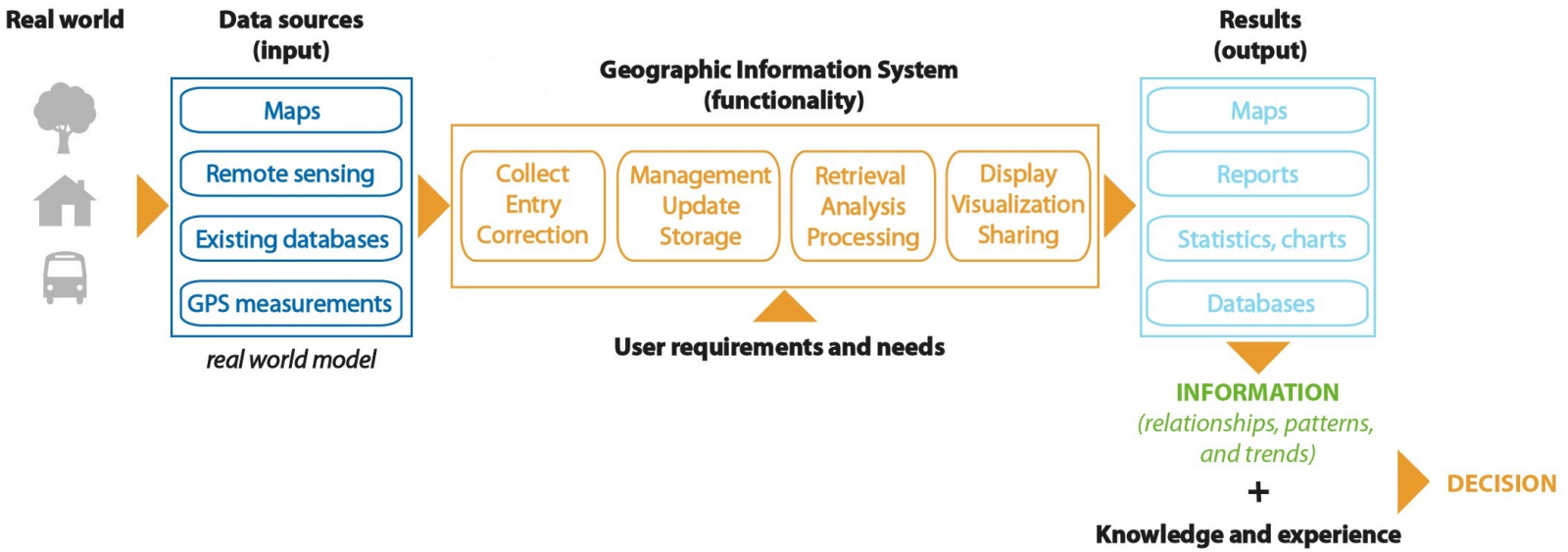
1.USGS (United States Geological Survey) 2. ESRI

GIS的组成

GIS的关键组成部分:

硬件系统, **软件系统**, **空间数据**, **应用人员**, 与 **地学模型和管理**.

GIS: main idea主要目标

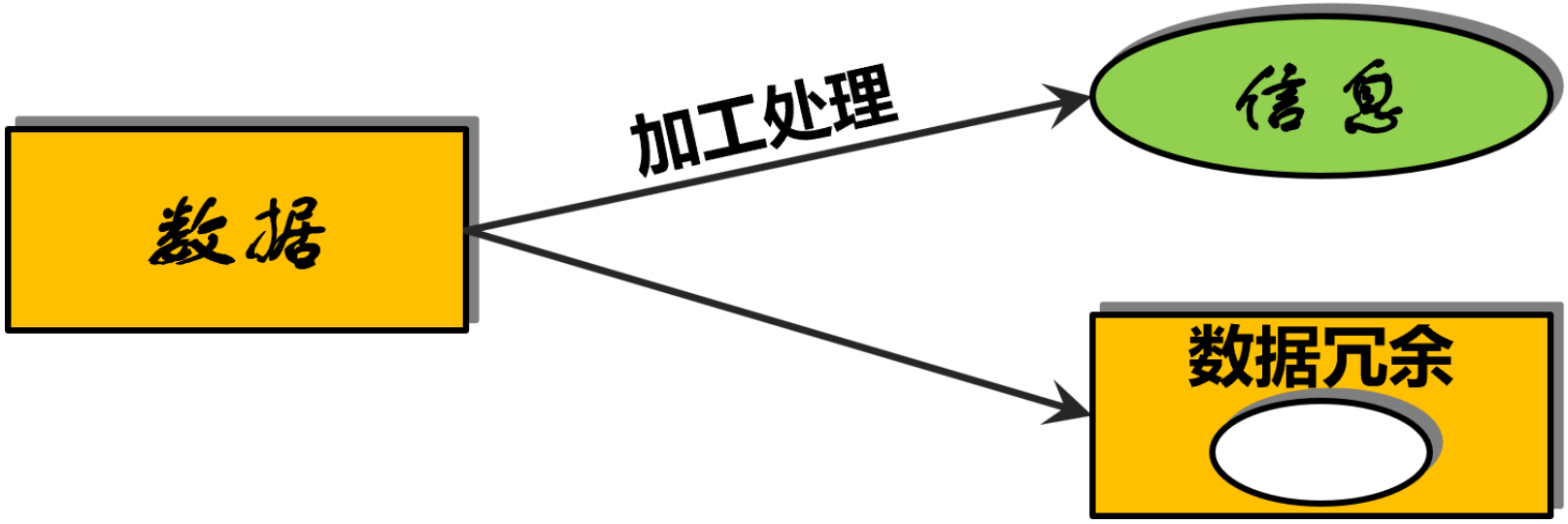


GIS 为决策者提供与 空间 相关的信息!

数据 vs. 信息

数据 是人类在认识世界和改造世界过程中，定性或定量对事物和环境描述的直接或间接原始记录，是一种未经加工的原始资料，是客观对象的表示。

信息 是用文字、数字、符号、语言、图像等介质来表示事件、事物、现象等的内容、数量或特征，从而向人们（或系统）提供关于现实世界新的事实和知识，作为生产、建设、经营、管理、分析和决策的依据。



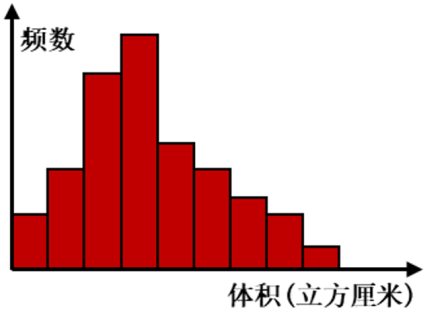
数据 vs. 信息

- **信息**来源于**数据**，是数据内涵的意义和数据内容的解释。信息是一种客观存在，而数据是客观对象的一种表示，其*本身并不是信息*。
- 数据所蕴涵的信息不会自动呈现出来，需要利用一种技术，如统计、解译、编码等对其解释，信息才能呈现出来。信息是数据的表达，数据是信息的载体。

现实世界中的对象

编号	鸡蛋体积
1	62.14
.....	.....
300	58.23

描述事物特征的数据



从数据中统计得到的信息

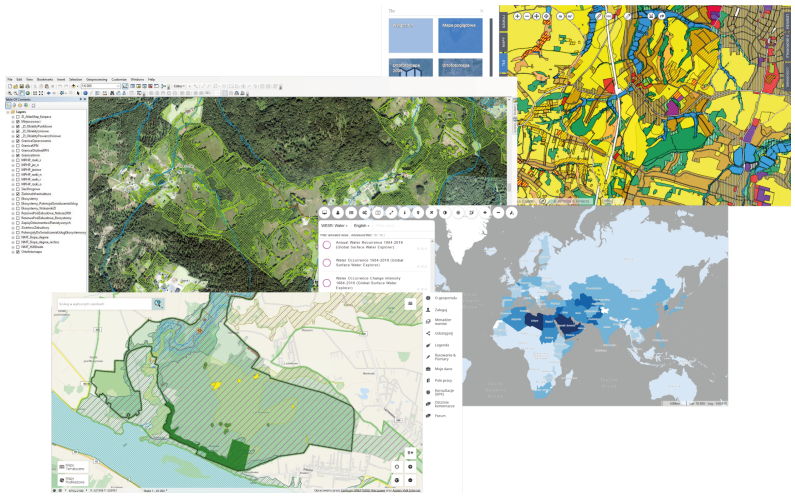
- 在GIS中，空间分析与建模是信息的主要来源。
- **空间分析** - a set of methods and tools for performing operations on spatial data in order to obtain additional information.

1.2 | GIS的应用

GIS软件与工具与:

- **Desktop GIS** (QGIS, SAGA GIS, GRASS GIS, ILWIS, IDRISI, Esri products: ArcGIS, ArcMap, ArcGlobe, GeoMedia, MapInfo, Bentley Systems: MicroStation, ENVI, ERDAS IMAGINE)
- **GIS** as a service (ArcGIS Online, Mapbox, OpenStreetMap, Google Maps, Apple Maps, Here Maps, Bing Maps)
- **Spatial database management systems** (MySQL, Oracle Spatial, Microsoft SQL Server, PostgreSQL)
- **Map servers** (Geoserver, MapServer, Mapnik)

GIS的应用领域



- urban planning
- environment protection and management
- land use monitoring
- agriculture
- transportation/logistics
- emergency management
- network infrastructure management
- tourism
- ...

GIS的优势

- Ability to view, visualize and interpret data in the form of maps, charts and reports - relationhips and trends easy to see and understand
- Improved decision making and problems solving through specific and detailed information regarding locations of features and phenomena
- Reduce costs and increase efficiency
- Improved communication between organisations or departments

2 | 空间数据

- 2.1 定义与特征
- 2.2 矢量数据模型
- 2.3 栅格数据模型
- 2.4 矢量与栅格数据模型的对比

2.1 | 定义与特征

From real world

… to GIS


Source: UNEP/GRID-Warsaw Centre, Green-Go! Carpathians,
(Local initiatives for deployment of green infrastructure
within Natura 2000 sites in the Carpathians)
ELEVATION
AIR POLLUTION
PROTECTED AREAS
LAND USE
POPULATION DENSITY
SOIL
ROADS
HYDROLOGY

空间数据 - 定义

Spatial object means an abstract representation of a real-world phenomenon related to a specific location or geographical area.

Spatial data means any data with a direct or indirect reference to a specific location or geographical area.

Spatial data set means an identifiable collection of spatial data.

DIRECTIVE 2007/2/EC of the European Parliament ad of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE)

数据模型

Data model is a set of guidelines to convert the real world (called entity) to the digitally and logically represented spatial objects consisting of the attributes and geometry.

The attributes are managed by thematic or semantic structure while the geometry is represented by geometric-topological structure.

Shunji, 1999

空间数据 - 特征

Spatial data describes **shape**, **location**, **spatial relationships** and *attributes* of features related to the Earth's surface.

GEOMETRY

! LINK !

DESCRIPTION

Trees

ID	species	age
1	oak	40
2	maple	60
3	birch	20

Roads

ID	type	length
1	highway	162
2	district road	94
3	local street	68

Land use

ID	class	area
1	water	405,08
2	agriculture	892,47
3	forest	936,91

空间数据模型

2种常用的空间数据模型:

- Vector 矢量

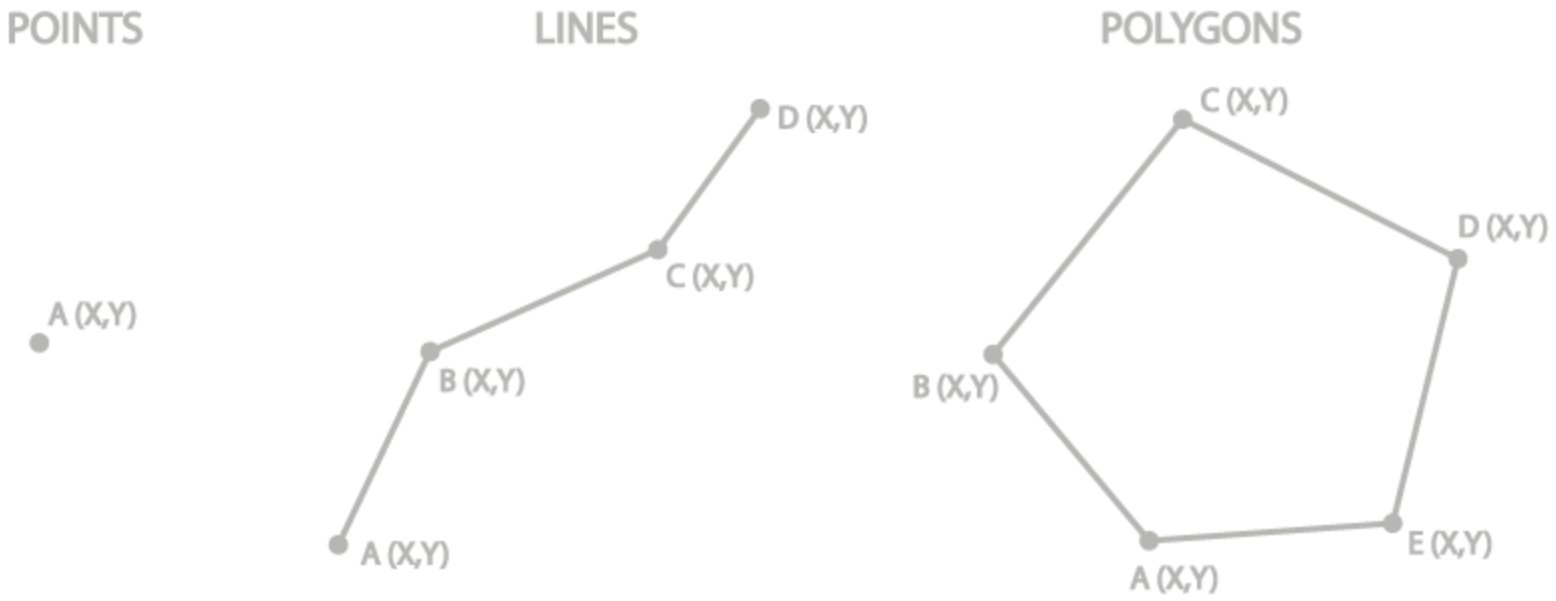
- Raster 栅格

2.2 | 矢量数据模型

Three types of geometry

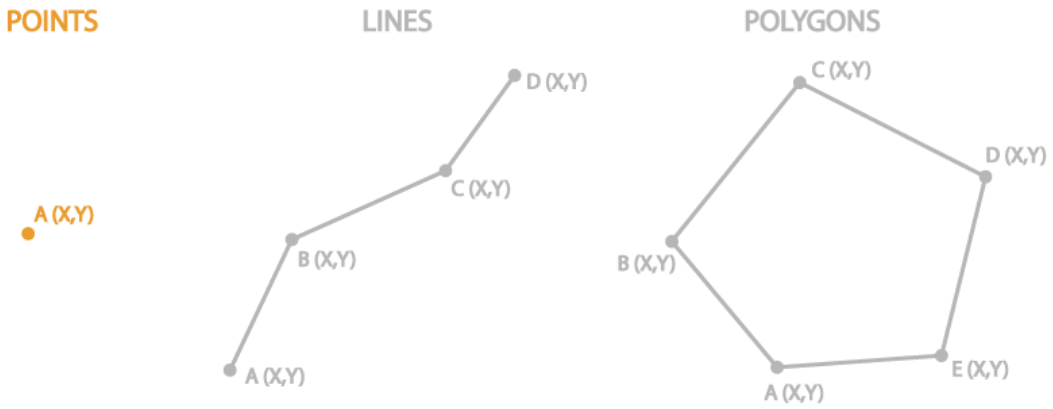
A vector data model defines discrete objects such as fire hydrants, rivers, lakes.

A vector data models divided into three basic types:



All three of these types of vector data are composed of coordinates and attributes attached to the geometry.

Geometry: points



1. 0-dimension objects

2. represented by a single pair of coordinates (X,Y)

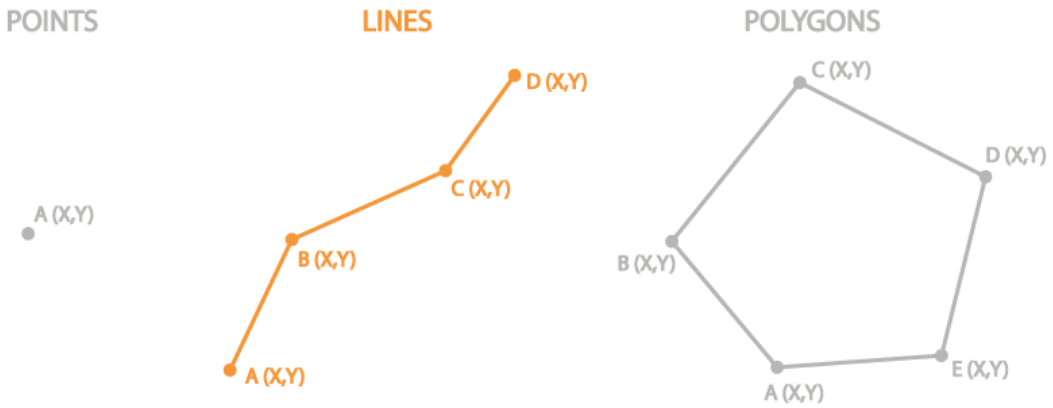
3. associated attribute information is attached to the center of the point

4. used to represent objects with no length or area (e.g. light poles, trees) or

5. used to represent a geographic feature too small to be displayed as a line or area (e.g. the location of a city on a small-scale map)

6. symbolized by a point or other sygnature (symbol) in different sizes and colors

Geometry: lines



1. 1-dimension objects

2. defined by an ordered set of two or more coordinate pairs called vertices

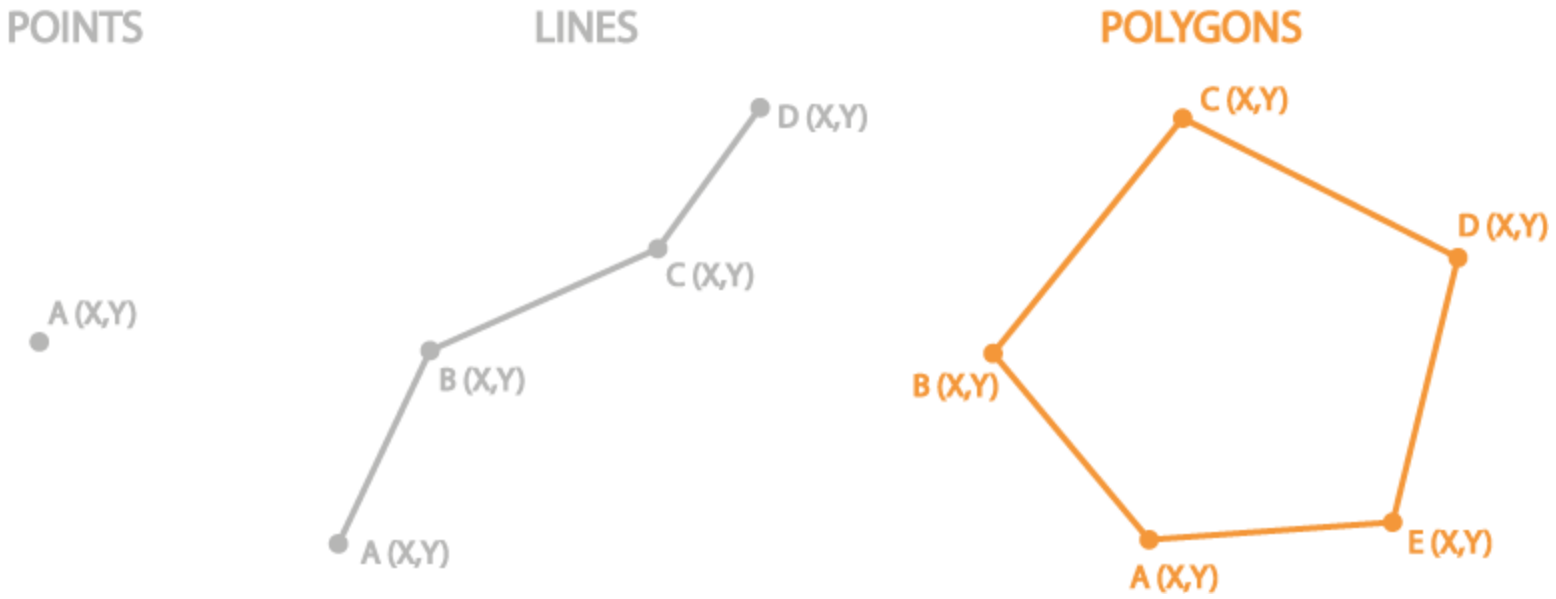
3. used to model linear features with no area (e.g. county boundary lines) or

4. used to represent the shape of geographic features too narrow to be displayed as an area at the given scale

(e.g. contours, street centrelines, streams)

5. symbolized by different types of line that have a color, width and style (solid, dashed, dotted, etc. ...)

Geometry: polygons



1. 2-dimension objects

2. composed of three or more connected lines where the start and end point have the same coordinate

3. attribute information is attached to the center of the polygon

4. used to represent areas (e.g. lakes, forests, cities)

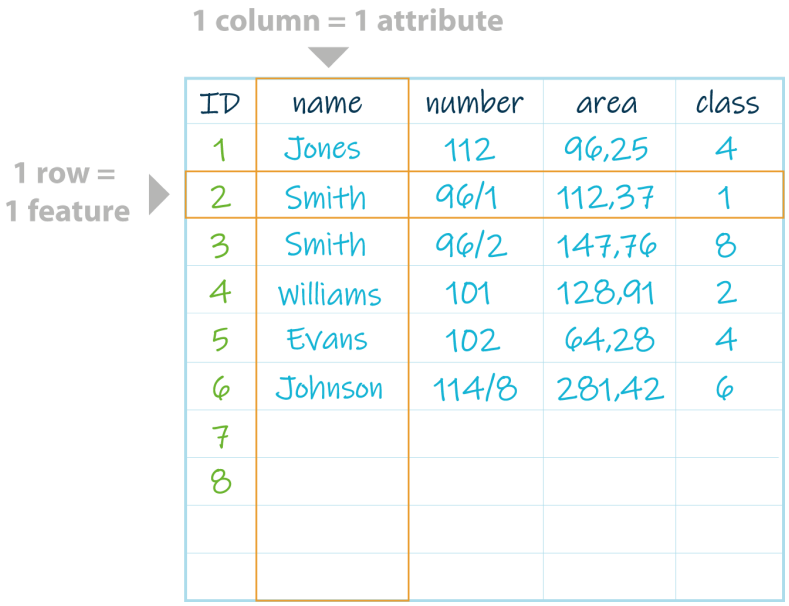
5. represent length and area, embody the idea of an inside and an outside

Attribute table

An **attribute** is a nonspatial information about a geographic feature in a GIS, usually stored in a table and linked to the feature by a unique identifier (ID).

A database or tabular file containing information about a set of geographic features, usually arranged so that:

- each row represents a feature
- each column represents one feature attribute.



ID	name	number	area	class
1	Jones	112	96,25	4
2	Smith	96/1	112,37	1
3	Smith	96/2	147,76	8
4	Williams	101	128,91	2
5	Evans	102	64,28	4
6	Johnson	114/8	281,42	6
7				
8				

The attribute values can be used to find, query,

analyze and symbolize features.

Attribute table - data types

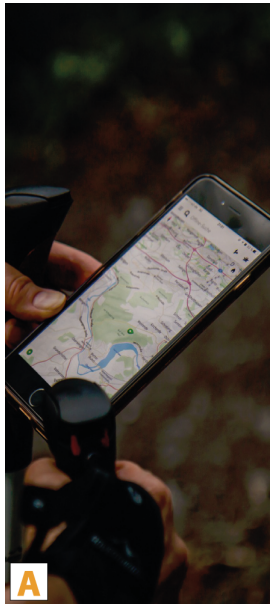
Each column in the database may contain different type of data.

ID	name	street	number	class	plan	date	time	area	length
1	Jones	Bell St.	12	4	true	2019-08-06	06:30:00	96,25	136,2
2	Smith	York St.	9	1	false	2018-01-10	08:25:15	112,37	242,0

Basic data types:

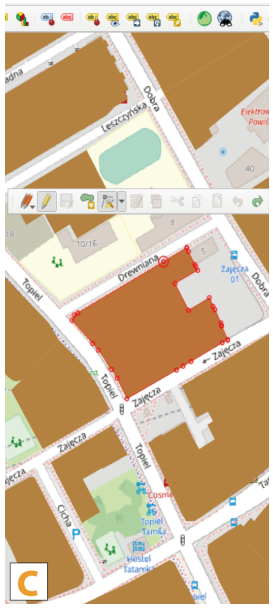
NUMERIC: INTEGER (long int, short int) - numbers, code list
NUMERIC: FLOAT (double, real) - floating-point numbers
STRING (char, varchar, text) - names and other texts
DATE/TIME (date, time, year, timestamp) - data and/or time
BOOLEAN (0/1, true/false, yes/no) - logical expression
BLOB - multimedia files

Vector data sources



L.p.	id	X	Y
1	1	589707,94	429982,80
2	1	589718,38	429961,29
3	1	589719,22	429956,83
4	1	589719,88	429957,57
5	1	589722,29	429960,28
6	1	589724,14	429962,58
7	1	589726,77	429965,83
8	1	589729,28	429969,01
9	1	589731,02	429971,37
10	1	589733,31	429974,61
11	1	589736,36	429979,07
12	1	589739,98	429982,40
13	1	589742,61	429985,18
14	1	589755,27	430001,60
15	1	589757,23	430004,71
16	1	589762,54	430014,54
17	1	589763,72	430017,03
18	1	589767,48	430026,11
19	1	589768,66	430029,47
20	1	589770,78	430036,39
21	1	589771,63	430039,14
22	1	589771,63	430039,27
23	1	589772,16	430040,88
24	1	589772,92	430042,57
25	1	589773,66	430044,78

B



(A - GPS measurements, B - list of coordinates, C - digitizing and conversion tools e.g. raster to vector, D, E - existing databases)

Vector file formats

- ESRI Shapefile - the most common geospatial file type developed by ESRI, consists of:

```
- shp (feature geometry)
- shx (shape index position)
- dbf (attribute data)
- prj (projection system metadata)
- xml (associated metadata)
```

- GML (Geography Markup Language) - XML based open standard for GIS data exchange
- KML/KMZ (Google Keyhole Markup Language) - XML based open standard for GIS data exchange
- GeoJSON (Geographic JavaScript Object Notation) - a lightweight format based on JSON, used by many open source GIS packages

2.3 | RASTER DATA MODEL

Raster data model

A raster data model defines continous data and phenomena.

Raster's are:

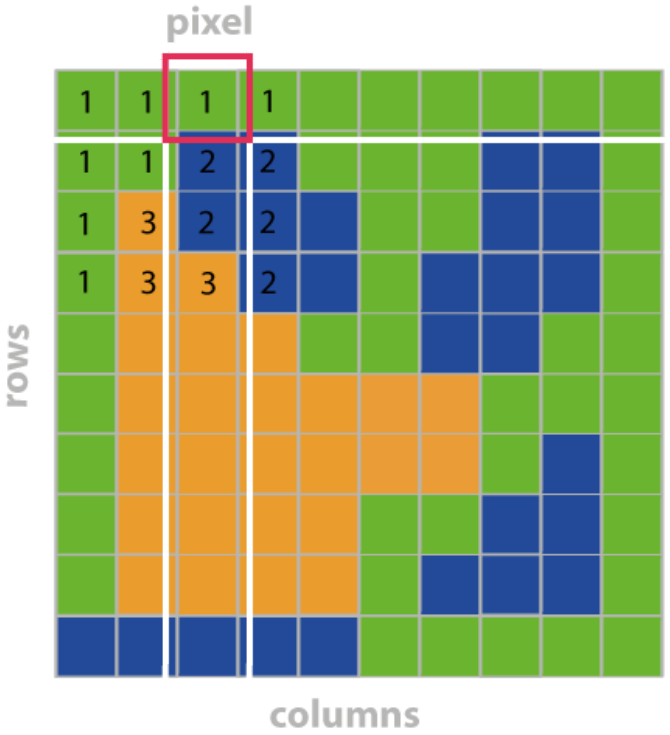
- digital aerial photographs and satellites imagery (spectral data),
- continuous data represents phenomena - e.g. temperature, precipitation, elevation, slope,
- digital pictures, scanned maps and plans.

Raster data model: geometry

A raster consists of a matrix of cells (or pixels) organized into rows and columns (or a grid) where each cell contains one value representing information such as temperature, elevation, or spectral data.

Pixel - smallest visible element of an image.

Grid - 2-D object feature that represents a single element of a continous surface.

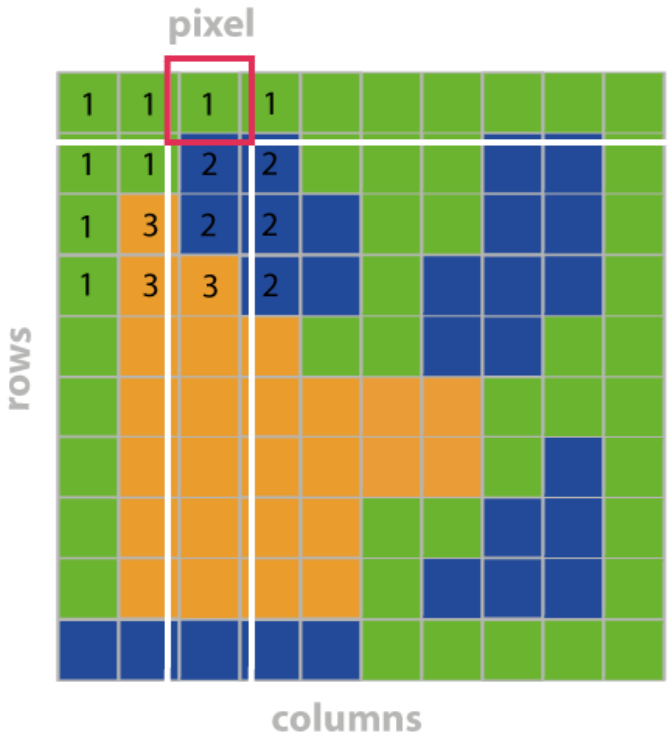


Raster data model: georeferencing

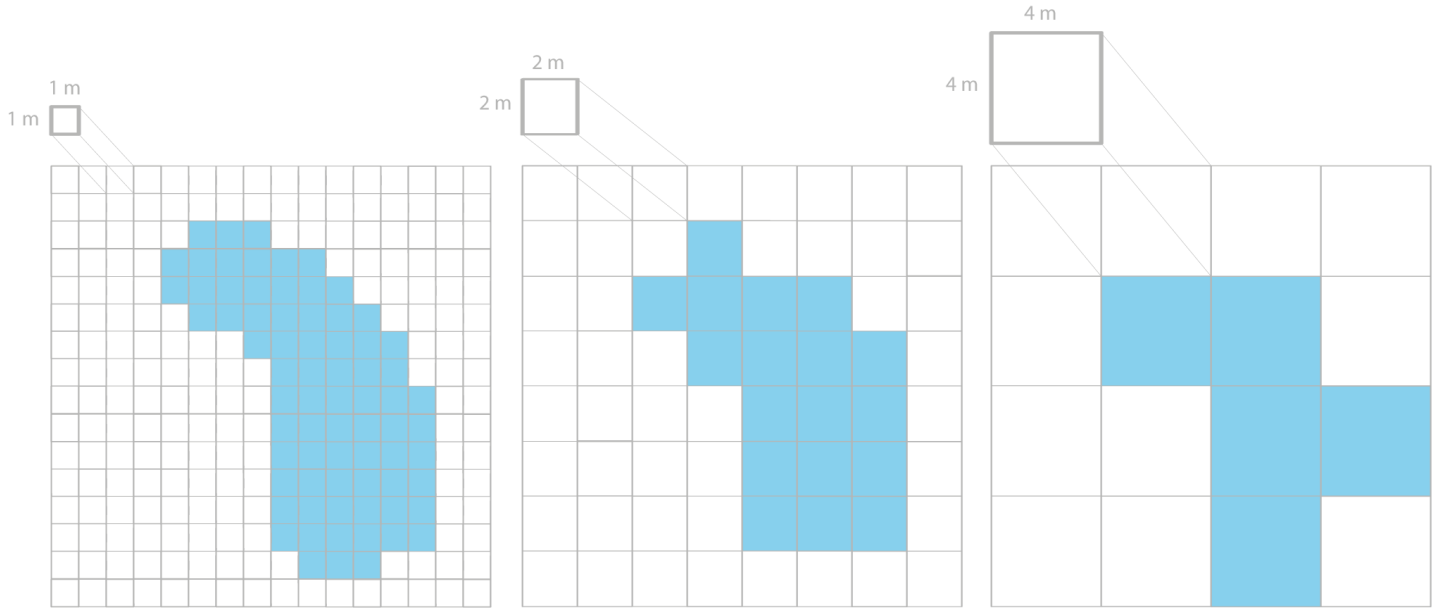
Cells are identified by their positions in the grid.

Raster data is georeferenced by:

- real world coordinates of the reference point,
- cell size in real world distance,
- using the upper-left or lower-left corner of grid as the reference point.



Spatial resolution



The same feature in images of different resolution

A **spatial resolution** refers to the dimension of the cell size representing the area covered on the ground. Higher resolution means better feature quality but it means also bigger raster file size.

Raster bands

A raster dataset contains one or more layers called bands.

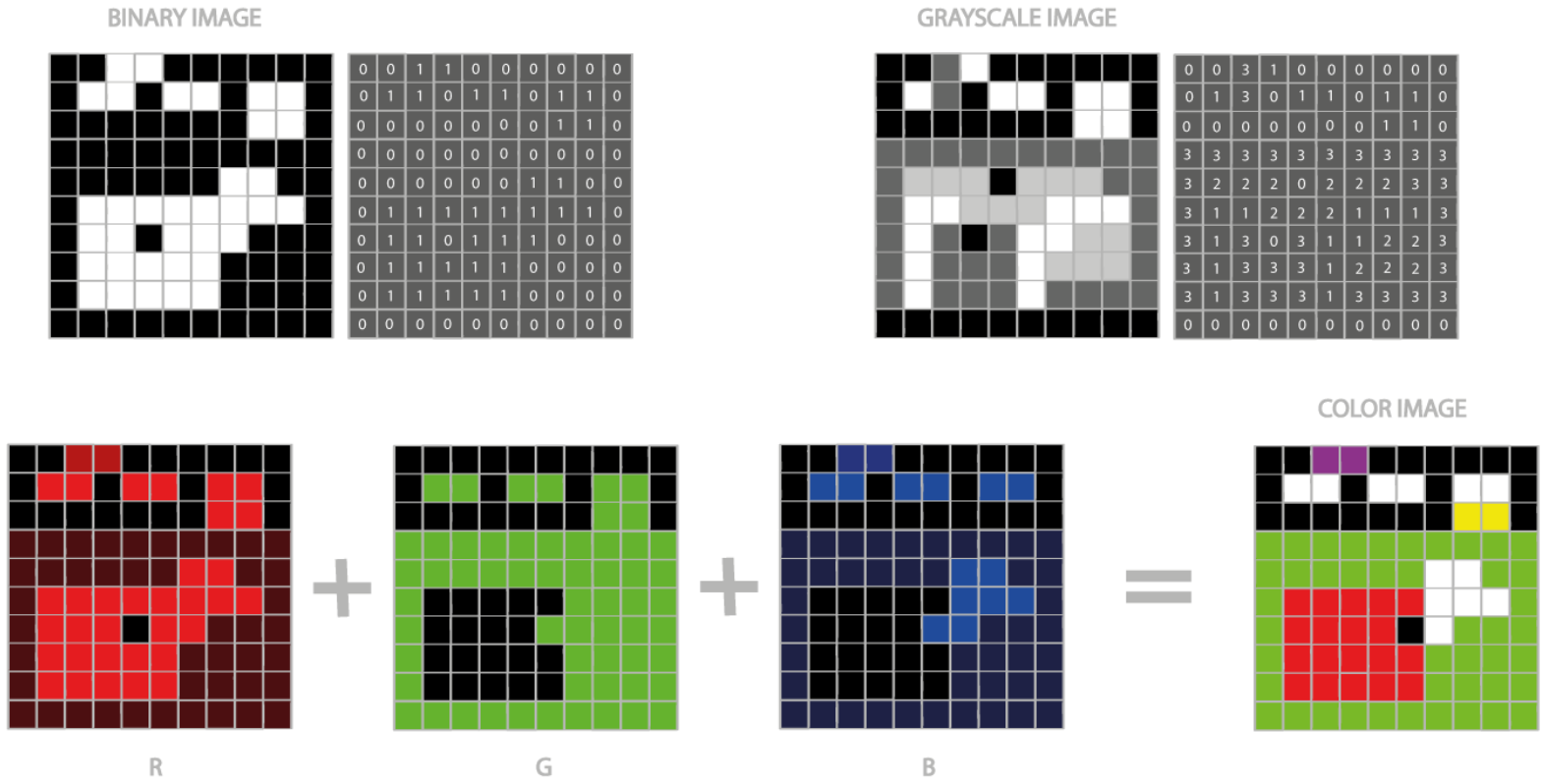
A band is represented by a single matrix of cell values.

For example, a **digital elevation model (DEM)** is a single-band raster (has one band holding elevation values) while **satellite imagery** is a multispectral image and has multiple bands.

Three main ways to display single-band raster datasets:

- binary image (each cell has a value of 0 or 1 and is often displayed using black and white),
- grayscale image (each cell has a value from 0 to another number, such as 255),
- color image (a set of values is coded to match a defined set of red, green, and blue (RGB) values).

Raster bands



Three ways to display raster dataset (binary image, grayscale image and color image)

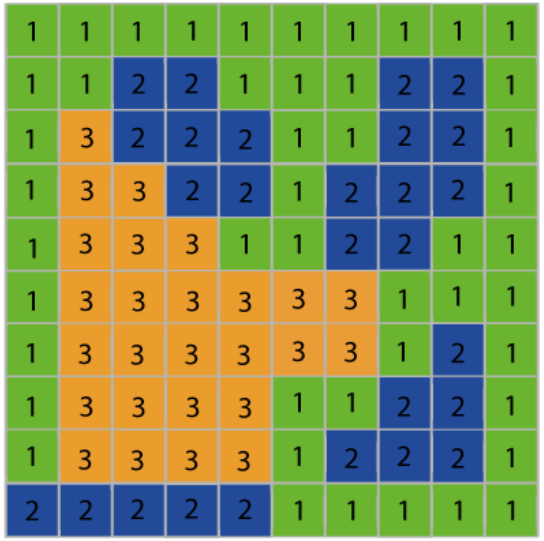
Attribute table

Raster data can also have attributes only if pixels are represented using a small set of unique integer values. Raster datasets that contain attribute tables typically have cell values that represent or define a class, group, category, or membership.

In raster datasets, each row of an attribute table corresponds to a certain zone of cells having the same value.

The attribute tables can be used to analyze datasets and symbolize raster cells.

Attribute table



Classification

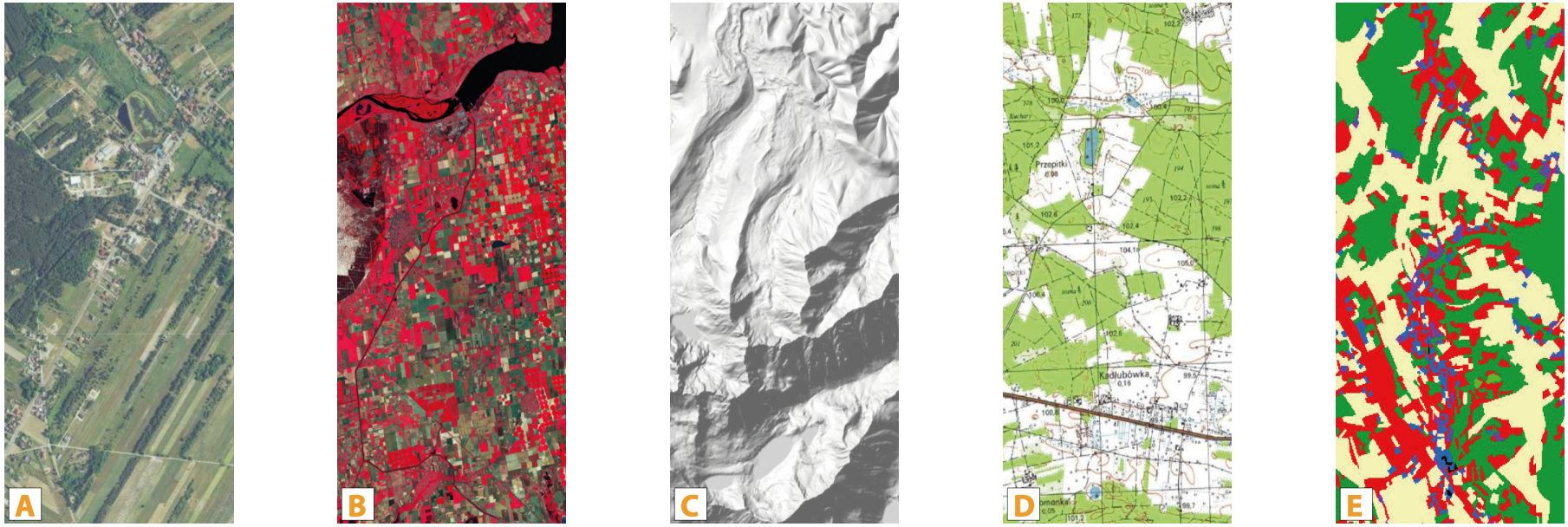
code	type
1	forest
2	water
3	agriculture

Raster attribute table

OID	value	count
0	1	47
1	2	27
2	3	26

An example of raster dataset with attribute table

Raster data sources



(A - orthophoto, B - satellite imagery, C - DEM, D - scanned maps and plans, E - conversion and analysis tools e.g. vector to raster, interpolation)

Raster file formats

- GeoTIFF - TIFF variant enriched with GIS relevant metadata, may be accompanied by other files:
  - tfw (raster geolocation)
  - xml (metadata)
  - aux (projections and other information)
  - ovr (pyramid files improves performance for raster display)
- IMG - ERDAS IMAGINE image file format
- ESRI Grid - format developed by Esri, which has two varieties: binary or ASCII

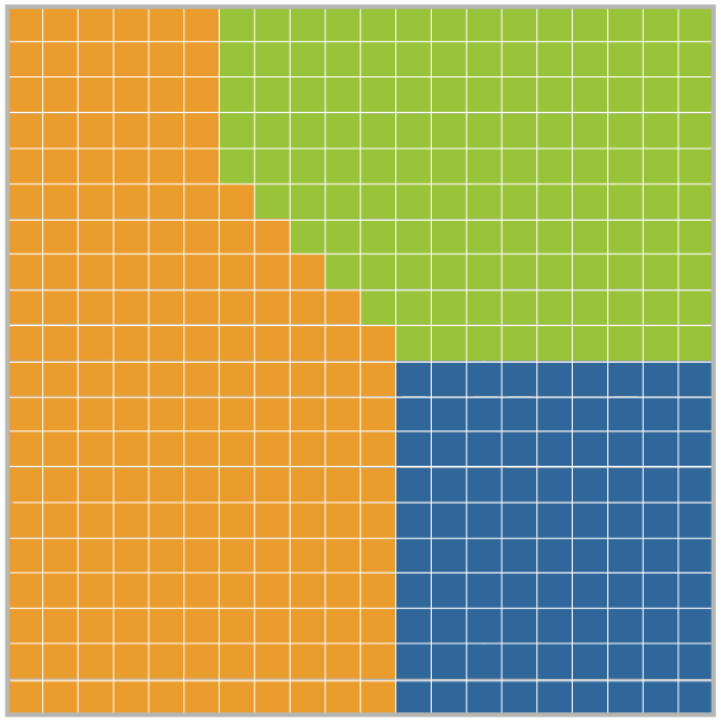
2.4 | COMPARING VECTOR AND RASTER DATA MODELS

Comparing: vector vs. raster data model

properties	vector	raster
depic	discrete features	continous data
geometry	coordinates	cells organized into a grid
attributes	attribute table (with many attributes)	cell value (only one attribute)
analysis	geoprocessing	map algebra, overlays
data structure	more complex	more simple
size	compact data structure – little storage space	greater storage needed
file formats	ESRI Shapefile, GML, KML, geoJSON, GPX	geoTIFF, IMG, grid

Which one is better?

RASTER MODEL



VECTOR MODEL

Which one is better?

RASTER MODEL

VECTOR MODEL

?

To be continued!