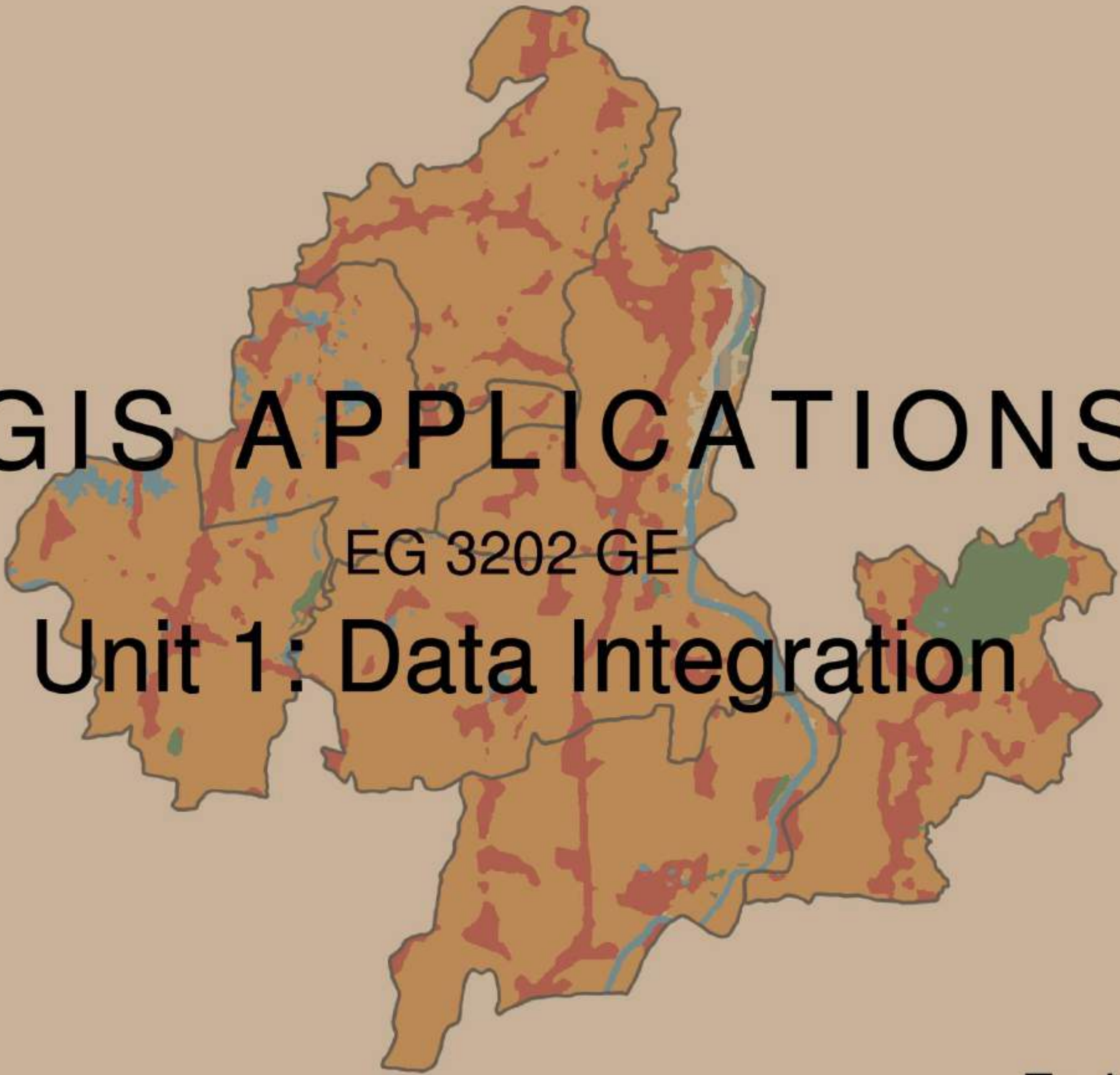


GIS APPLICATIONS

EG 3202 GE

Unit 1: Data Integration

Public Namuna Secondary School

Lecture by

Er. Keshav Raj Bhusal

Table of contents

Concept of data Integration

What is Data Integration?

- Data integration is the process of combining data from various sources into one, unified view for efficient data management, to derive meaningful insights, and gain actionable intelligence.
- Data integration is widely used in different areas of scientific circles and ordinary life as well. The process of data integration becomes significant in a variety of situations both commercial (when two commercial companies need to merge the database) and scientific (combining research results from different data sources/repositories).

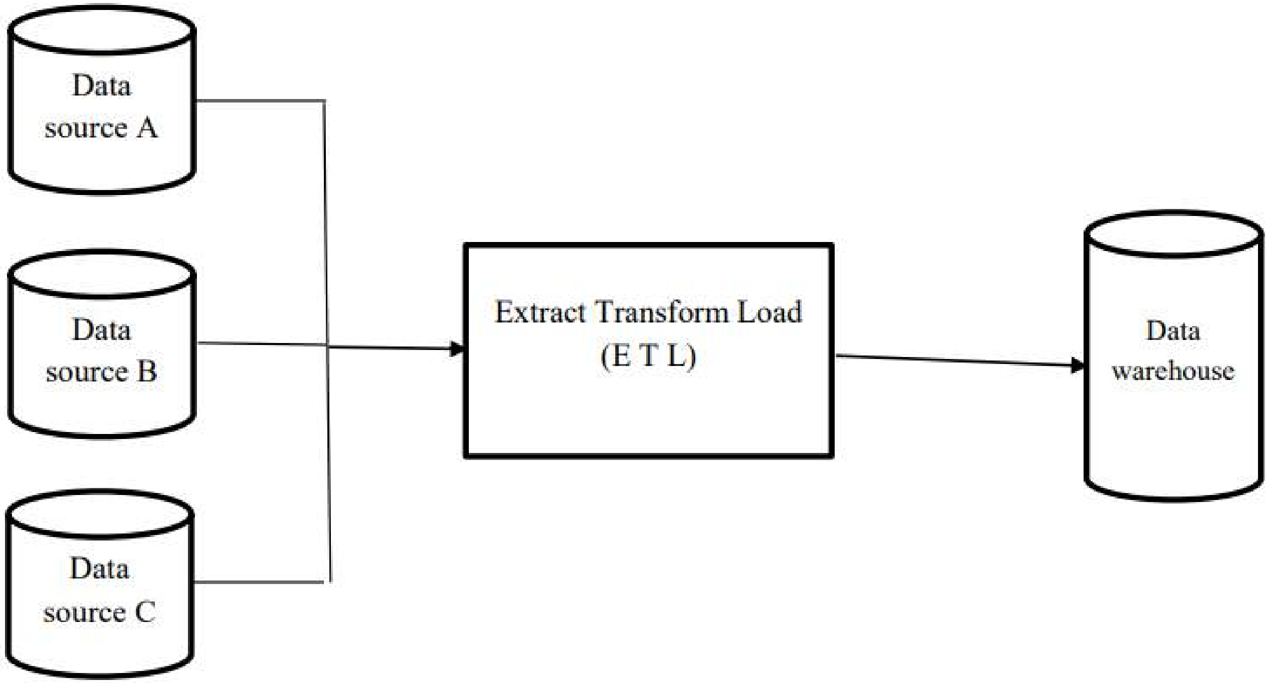


Fig: Architecture of Data integration

Concept of data Integration

Data warehouse

- A large store of data accumulated from wide range of sources.

ETL Process

- ETL process extracts information from sources databases, transforms it and then loads into the data warehouse.

Concept of data Integration

Importance of data integration

1. *Every data has its own strength:*

- Every data is designed for a reason. Each one represents information in a way no other format can, with attributes, metadata, structure and schema. For example; CAD and GIS data can be integrated to enhance a CAD drawing with specialized GIS information and attributes.

2. Take advantage of specialized application:

- Every piece of software that works with data represents, analyzes and transforms information in a specialized way. By integrating data into format accepted by the application you are giving yourself the power to open and use your data in that software.

3. Reduce data complexity:

- The information regarding the data will be available from different sources which helps in easy understanding of the data and metadata. This helps in reducing the data complexity.

Concept of data Integration

Importance of data integration

4. Increases the value of data through unified system:

- Bringing the desperate/different data set together increases the value of information.

5. Make data easily available:

- Centralizing your data makes it easy for anyone to retrieve, inspect and analyze it.

6. Make your data live:

- An integration data solution make it easy to keep information update. Eg; weather network integrates non-spatial sensor data, metadata and outputs it to spatial applications like Google earth.

Concept of data Integration

What is Spatial Data Integration?

- GIS data integration is the process of combining spatial data from multiple sources and formats to create a comprehensive, integrated dataset for analysis and decision-making.
- It is the process of combining data of different themes, content, scale or spatial extent, projections, acquisition methods, formats, schema, or even levels of uncertainty, so that they can be understood and analyzed.
- Information is passed within datasets. Thus, providing more information and insight than each type of data separately.
- The goal of spatial data integration is to facilitate the analysis, reasoning, querying, or visualization of the integrated spatial data
- Data integration is one of the main reasons why GIS software is used.

For example: layer stacking, zonal statistics, table join etc.

Concept of data Integration

What is Spatial Data Integration?

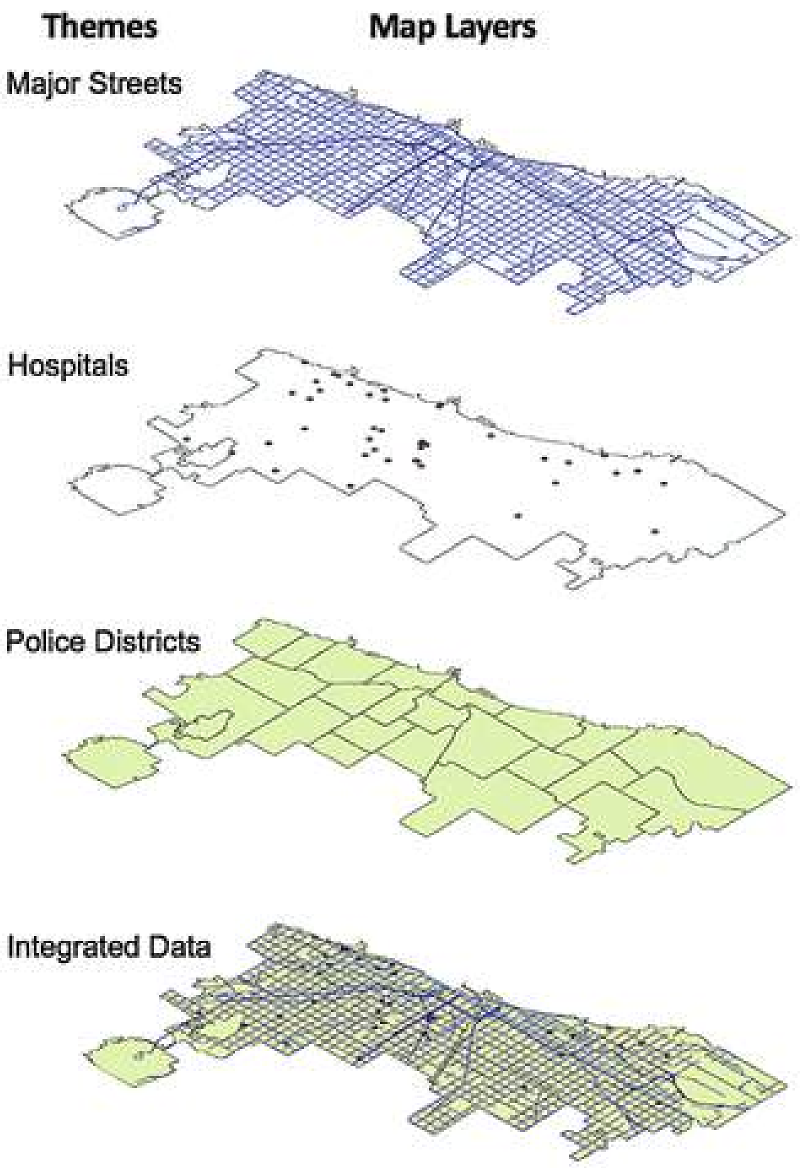


Fig: Spatial data integration of three layers or themes: major streets, hospitals, and police districts of the City of Chicago

Source: (2017) Open data portal of the city of Chicago. https://data.cityofchicago.org/. Accessed: 30 Nov 2017

Spatial data Integration

Importance of spatial data integration

Spatial data integration is important due to the following:

1. Holistic Understanding:

- Spatial data integration involves combining diverse datasets from different sources. This allows us to look at a particular geographic area or phenomenon from multiple perspectives, providing a more complete and holistic understanding.

2. Enhanced Decision-Making:

- Integrating spatial data with non-spatial data provides decision-makers with a more comprehensive view of a situation. This broader perspective enables more informed and effective decision-making.

3. Efficient Resource Management:

- In sectors like agriculture, urban planning, and natural resource management, spatial data integration optimizes resource allocation. This leads to more efficient and sustainable practices.

Importance of spatial data integration

Spatial data integration is important due to the following:

4. Infrastructure Development

- Integrating spatial data aids in infrastructure projects by assisting in site selection, environmental impact assessment, and overall project planning.

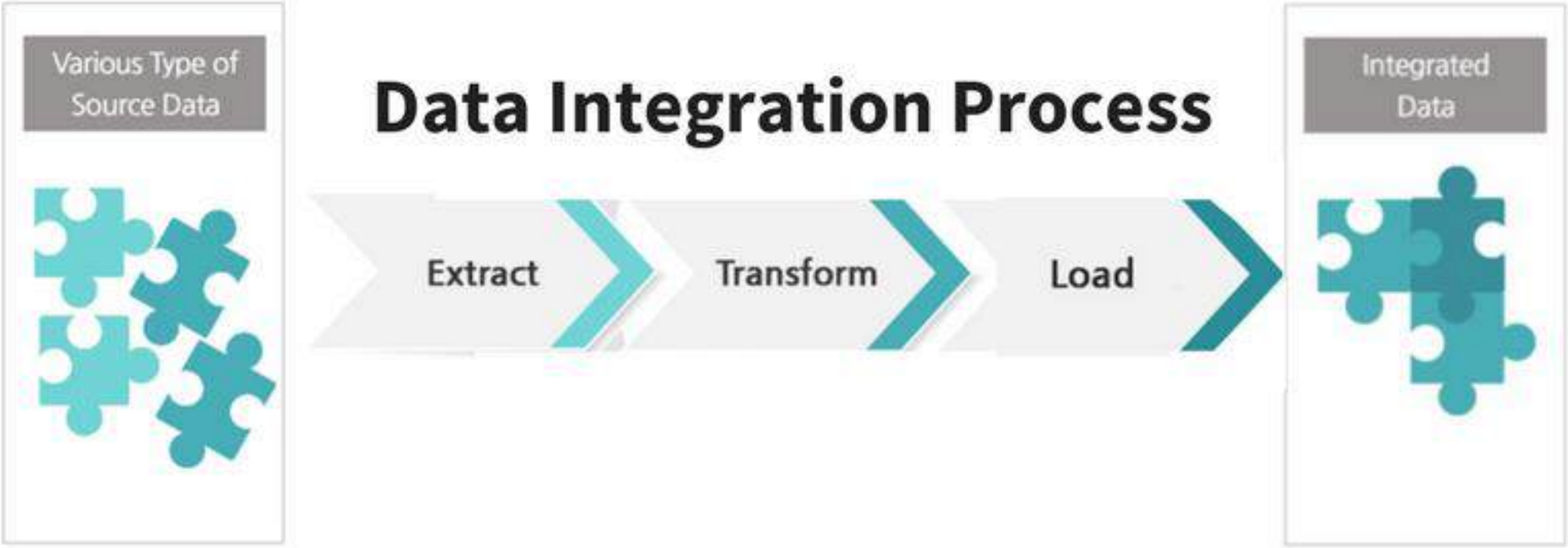
5. Interdisciplinary Collaboration:

- Spatial data integration facilitates collaboration between experts from different fields, encouraging a holistic approach to problem-solving.
- This helps to tackle the challenges faced by organizations and governments involving multiple disciplines.

Process of data integration

Data integrations are complex and technical, but at the core are three steps are known as ETL: extraction, transformation, and loading

1. ***Data extraction*** is when data is collected from all of the various data sources.
2. ***Data transformation*** is the process of transforming all of your data into a compatible format, often with the help of a universal data model that defines the datasets with common properties.
3. ***Data loading*** is when the right data is loaded into the appropriate database, ready for analysis.



Source: https://auth0.com/blog/all-you-need-to-know-about-data-management-and-integration/

Join and relate spatial and attribute data

Join

- A table join appends all the columns from one table into the other table based on the unique ID.
- A join is usually used to attach more attributes to the attribute table of a geographic layer.

Different types of relationships in join are:

1. One to one relationship
2. Many to one relationship
3. One to many relationship
4. Many to many relationship

Source: https://auth0.com/blog/all-you-need-to-know-about-data-management-and-integration/

Joining

One to one relationship

- In such type of join , you establish a one to one relationship between the layer attribute table and table containing the information you want to join.
- Combines two tables based on the value of a field that can be found in both tables (a "common field").

NepalDistrict_Old — Features Total: 75, Filtered: 75,...

	DIST_ID	DISTRICT	ZONE_NAME	REGION
1	69	Achham	Seti	Far-Western
2	51	Arghakhanchi	Lumbini	Western
3	45	Baglung	Dhaulagiri	Western
4	74	Baitadi	Mahakali	Far-Western
5	68	Bajhang	Seti	Far-Western
6	67	Bajura	Seti	Far-Western
7	57	Banke	Bheri	Mid-Western
8	33	Bara	Narayani	Central

+

total-number-of-population-and-registered-voters-per...

	District	Total number of registered voters	Total population
1	Achham	257477	127742
2	Arghakhanchi	197632	126699
3	Baglung	268613	156802
4	Baitadi	250898	124824
5	Bajhang	195159	97610
6	Bajura	134912	64950
7	Banke	491313	246078
8	Bara	687708	333678

=

NepalDistrict_Old — Features Total: 75, Filtered: 75, Selected: 0

	DIST_ID	DISTRICT	ZONE_NAME	REGION	Total number of registered voters	Total population
1	69	Achham	Seti	Far-Western	257477	127742
2	51	Arghakhanchi	Lumbini	Western	197632	126699
3	45	Baglung	Dhaulagiri	Western	268613	156802
4	74	Baitadi	Mahakali	Far-Western	250898	124824
5	68	Bajhang	Seti	Far-Western	195159	97610
6	67	Bajura	Seti	Far-Western	134912	64950
7	57	Banke	Bheri	Mid-Western	491313	246078
8	33	Bara	Narayani	Central	687708	333678

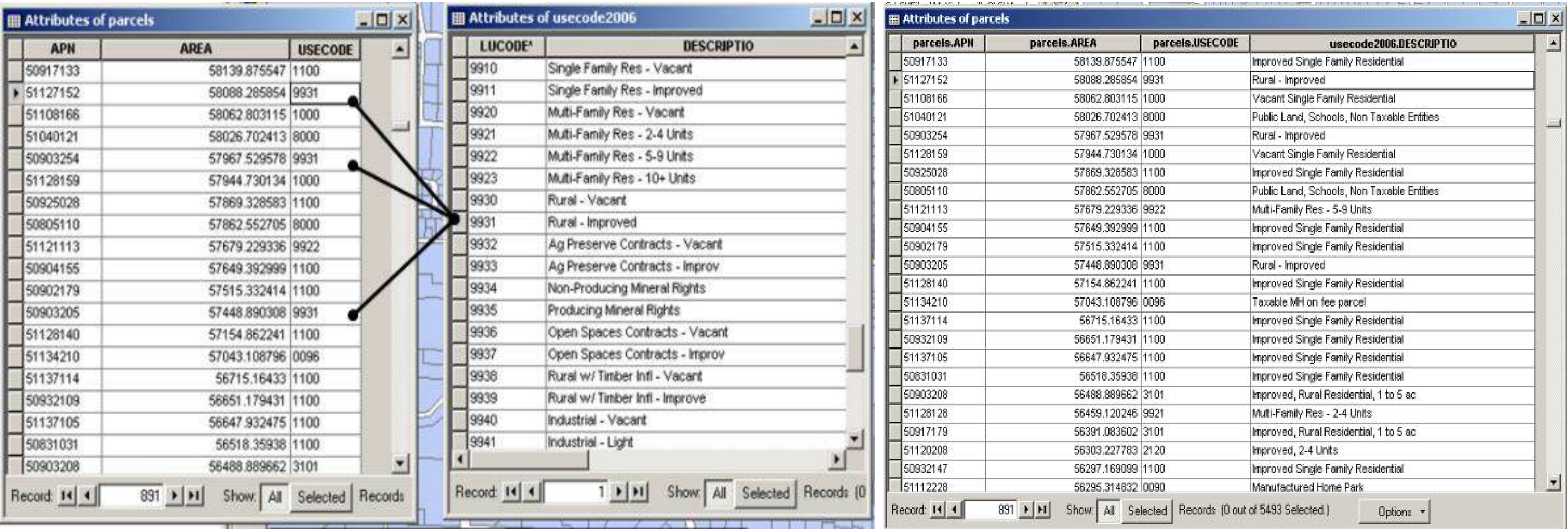
Tutorial: https://youtu.be/JUvE6KUcSdM

Fig: showing *one to one relationship* between NepalDistrict_Old and total population data.

Joining

Many to one relationship

- In such type of joining , you establish a many to one relationship between the layer attribute table and table containing the information you want to join.



Attributes of parcels

APN	AREA	USECODE
50917133	58139.875547	1100
51127152	58088.285854	9931
51108166	58062.803115	1000
51040121	58026.702413	8000
50903254	57967.529578	9931
51128159	57944.730134	1000
50925028	57869.328583	1100
50805110	57862.552705	8000
51121113	57679.229336	9922
50904155	57649.392999	1100
50902179	57515.332414	1100
50903205	57448.890308	9931
51128140	57154.862241	1100
51134210	57043.108796	0096
51137114	56715.16433	1100
50932109	56651.179431	1100
51137105	56647.932475	1100
50831031	56518.35938	1100
50903208	56488.889662	3101

Record: 891 Show: All Selected Records

Attributes of usecode2006

LUCODE	DESCRIPTIO
9910	Single Family Res - Vacant
9911	Single Family Res - Improved
9920	Multi-Family Res - Vacant
9921	Multi-Family Res - 2-4 Units
9922	Multi-Family Res - 5-9 Units
9923	Multi-Family Res - 10+ Units
9930	Rural - Vacant
9931	Rural - Improved
9932	Ag Preserve Contracts - Vacant
9933	Ag Preserve Contracts - Improv
9934	Non-Producing Mineral Rights
9935	Producing Mineral Rights
9936	Open Spaces Contracts - Vacant
9937	Open Spaces Contracts - Improv
9938	Rural w/ Timber Intl - Vacant
9939	Rural w/ Timber Intl - Improve
9940	Industrial - Vacant
9941	Industrial - Light

Record: 1 Show: All Selected Records (0

Attributes of parcels

parcels.APN	parcels.AREA	parcels.USECODE	usecode2006.DESCRIPTIO
50917133	58139.875547	1100	Improved Single Family Residential
51127152	58088.285854	9931	Rural - Improved
51108166	58062.803115	1000	Vacant Single Family Residential
51040121	58026.702413	8000	Public Land, Schools, Non Taxable Entities
50903254	57967.529578	9931	Rural - Improved
51128159	57944.730134	1000	Vacant Single Family Residential
50925028	57869.328583	1100	Improved Single Family Residential
50805110	57862.552705	8000	Public Land, Schools, Non Taxable Entities
51121113	57679.229336	9922	Multi-Family Res - 5-9 Units
50904155	57649.392999	1100	Improved Single Family Residential
50902179	57515.332414	1100	Improved Single Family Residential
50903205	57448.890308	9931	Rural - Improved
51128140	57154.862241	1100	Improved Single Family Residential
51134210	57043.108796	0096	Taxable MH on fee parcel
51137114	56715.16433	1100	Improved Single Family Residential
50932109	56651.179431	1100	Improved Single Family Residential
51137105	56647.932475	1100	Improved Single Family Residential
50831031	56518.35938	1100	Improved Single Family Residential
50903208	56488.889662	3101	Improved, Rural Residential, 1 to 5 ac
51128128	56459.120246	9921	Multi-Family Res - 2-4 Units
50917179	56391.083602	3101	Improved, Rural Residential, 1 to 5 ac
51120208	56303.227783	2120	Improved, 2-4 Units
50932147	56297.169099	1100	Improved Single Family Residential
51112228	56295.314832	0090	Manufactured Home Park

Record: 891 Show: All Selected Records (0 out of 5493 Selected.) Options

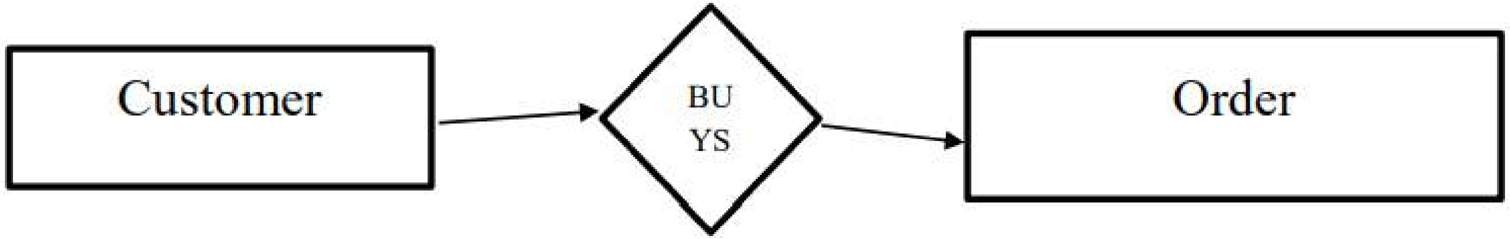
Source: *http://gsp.humboldt.edu/olm_2018/Lessons/GIS/06_2%20Vector%20Analysis%20Overlays/05_JoinsandRelates.html*

Join and relate spatial and attribute data

Joining

One to many and many to many relationship

- When using data where one to many and many to many relationship exists you should use a relationship class to establish between data sets.



Customer

Customer ID	Name	Location
1	Prajwal	KTM
2	Astro	PKR
3	Subash	DRN

Relationship table

Customer ID	Order ID	Date
1	100	1st june
1	102	2nd june
2	101	3rd june

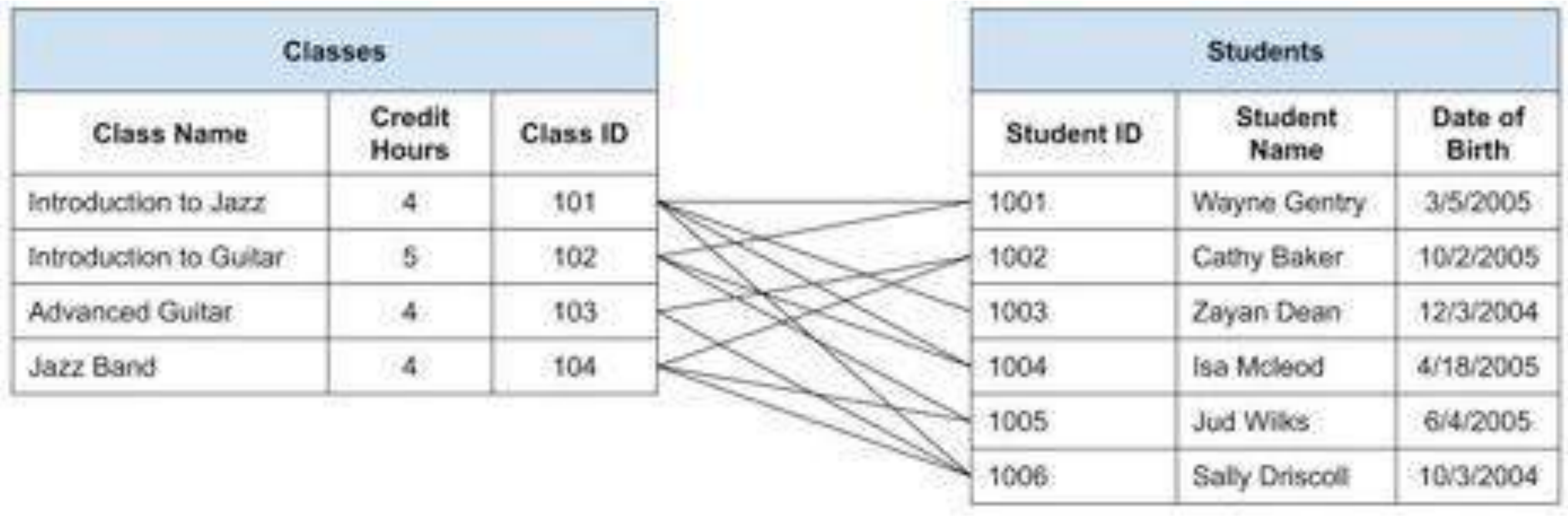
Order

Order ID	Item	Cost
100	Books	1000
101	Jeans	1500
102	Shoes	2000

Fig: showing *one to many relationship*.

Joining

One to many and many to many relationship



Classes		
Class Name	Credit Hours	Class ID
Introduction to Jazz	4	101
Introduction to Guitar	5	102
Advanced Guitar	4	103
Jazz Band	4	104

Students		
Student ID	Student Name	Date of Birth
1001	Wayne Gentry	3/5/2005
1002	Cathy Baker	10/2/2005
1003	Zayan Dean	12/3/2004
1004	Isa Mcleod	4/18/2005
1005	Jud Wilks	6/4/2005
1006	Sally Driscoll	10/3/2004

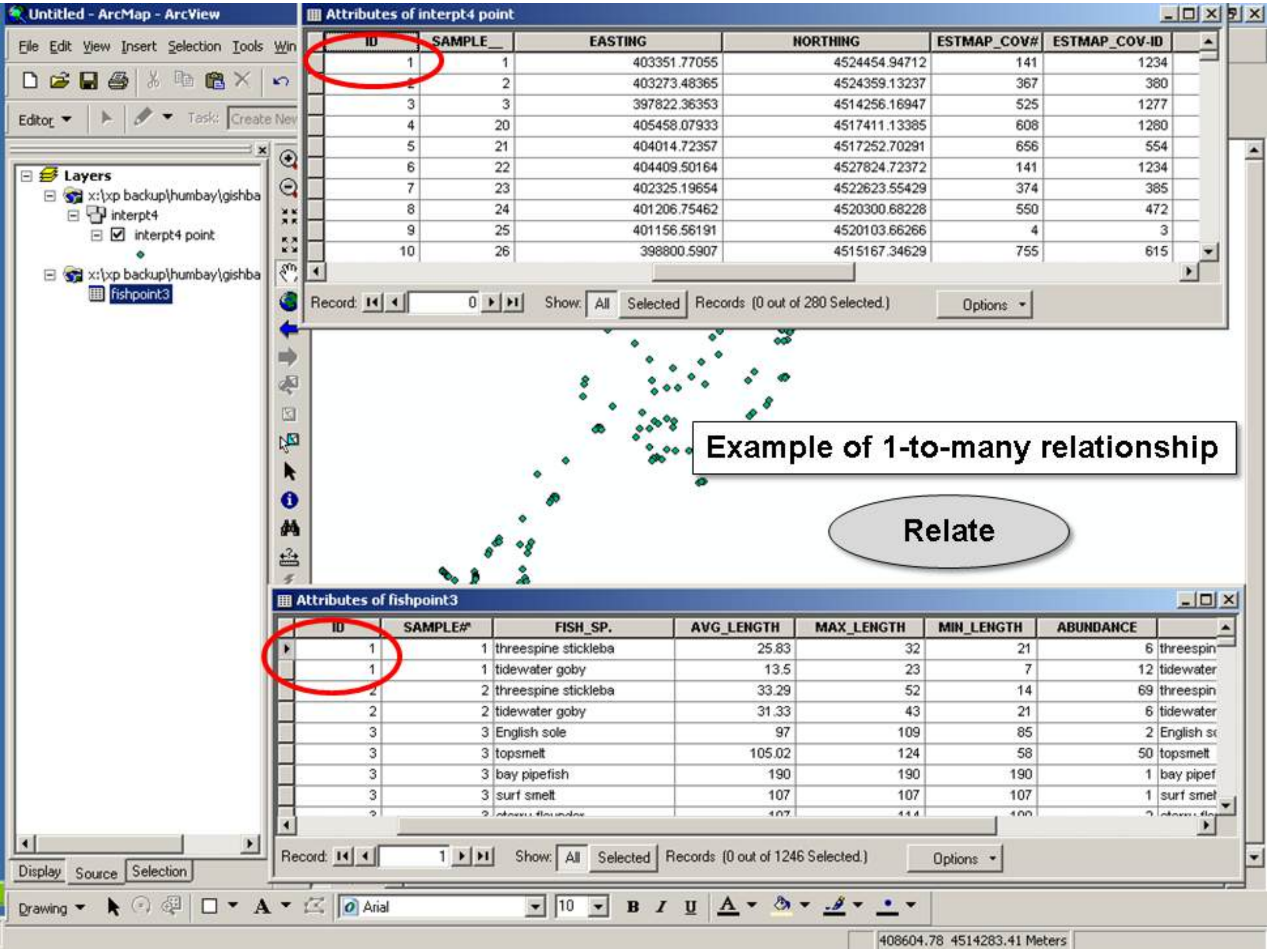
Fig: showing *Many to many relationship*

Source: https://blog.supportgroup.com/getting-started-with-relational-databases-one-to-one-and-many-to-many-relationships

Relate

- A relate operation temporarily connects two tables based on a common field but doesn't append the attributes of one to the other ; creates an entirely new table.
- Relates are used for 1-to-many relationships, or many-to-many relationships between the attribute table and an external table.

Relate



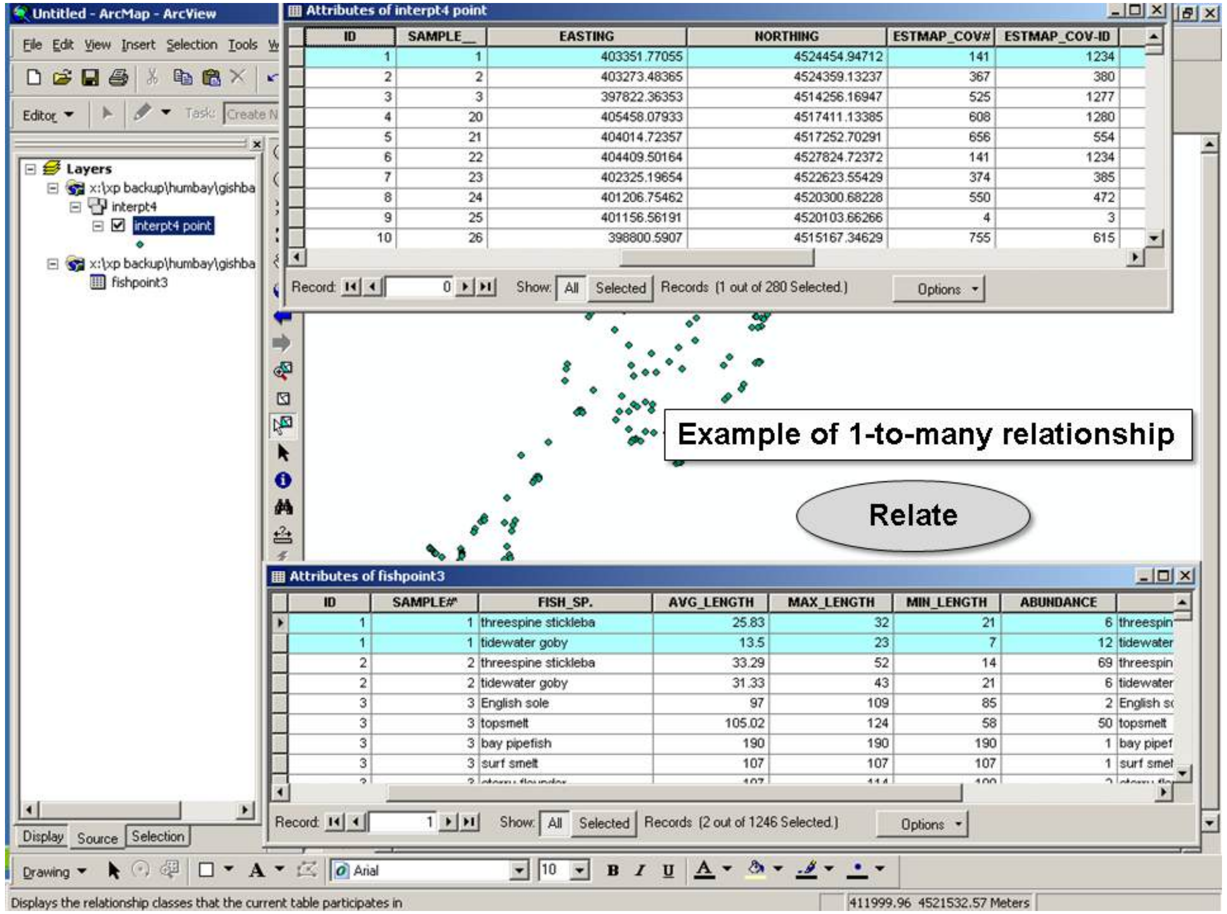


Fig: showing *one to many relationship using relate*

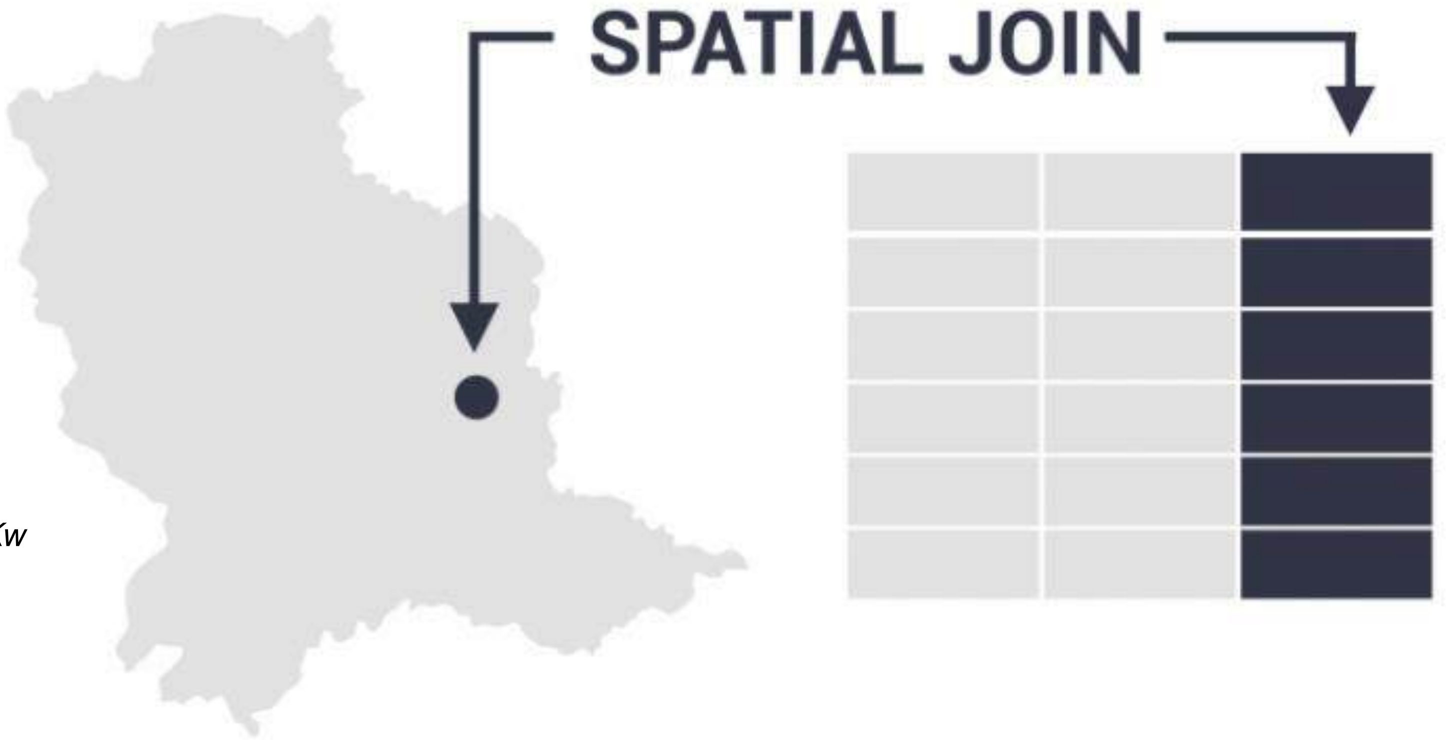
Source: https://gsp.humboldt.edu/olm/Lessons/GIS/06_2%20Vector%20Analysis%20Overlays/05_JoinsandRelates.html

Spatial Join

- Spatial join tool inserts the columns from one feature table to another based on location or proximity.

Eg. Let's say you have a set of land parcels. Each land parcel has a point inside of it. By running a spatial join, you can transfer the point table columns into the land parcel layer.

So that means that if the points have the owner's name, this field gets transferred to the land parcels.



Source: https://gisgeography.com/spatial-join/

Tutorial: https://youtu.be/pzi-eAyqUKw

Spatial Join

Types of spatial join

1. Intersect: Two features touch at any location
2. Within a distance: Two features are within a set distance
3. Completely within: The join feature is within the target feature
4. Identical: Both features match identically
5. Closest: The join feature is closest to the target feature

5

Source: https://gisgeography.com/spatial-join/

1

2

3

4

Projection and transformation of spatial data

Projection

- A projection is a mathematical technique to represent the Earth's curved surface on a flat surface.
- To represent parts of the surface of the Earth on a flat paper map or on a computer screen, the curved horizontal reference surface must be mapped onto the 2D mapping plane.
- Mapping onto a 2D mapping plane means transforming each point on the reference surface with geographic coordinates (φ, λ) to a set of Cartesian coordinates (x, y) representing positions on the map plane.
- A map projection provides a couple of distinctive advantages.

i. A map projection allows us to use two-dimensional maps, either paper or digital.

ii. A map projection allows us to work with plane coordinates rather than longitude and latitude values.

Projection

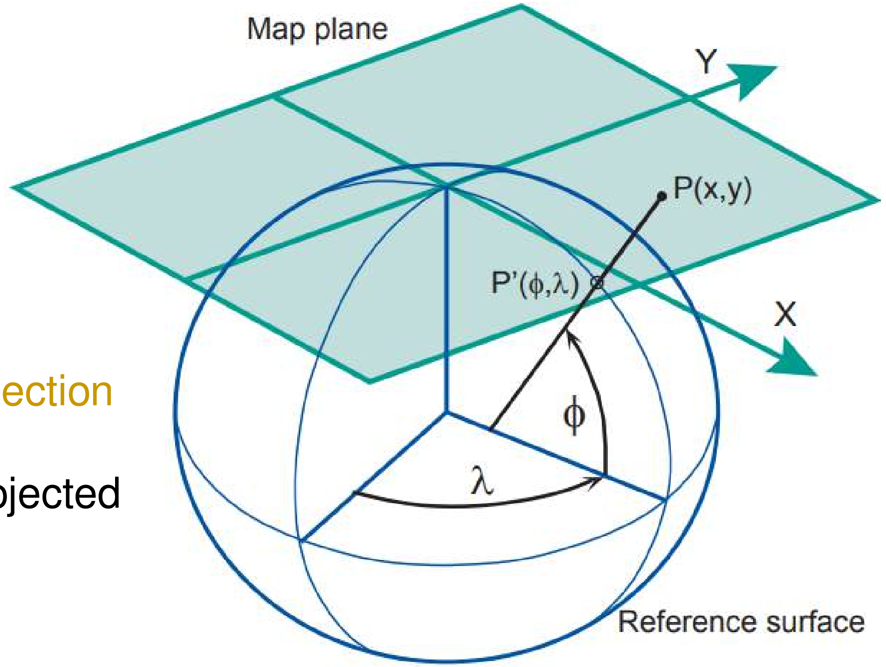


Fig: showing example of a map projection where the reference surface with geographic coordinates (φ, λ) is projected onto the 2D mapping plane with 2D Cartesian coordinates (x, y)..

Source: https://webapps.itc.utwente.nl/library www/papers_2009/general/principles gis.pdf

- A forward mapping equation transforms the geographic coordinates (φ, λ) of a point on the curved reference surface to a set of planar Cartesian coordinates (x, y), representing the position of the same point on the map plane:

$$(x, y) = f(\varphi, \lambda)$$

Transformation

Transformation

- A geographical transformation is a mathematical operation that converts the coordinates of a point in one geographic coordinate system to the coordinates of the same point in another geographic coordinate system.
- A geographic transformation always converts geographic (latitude–longitude) coordinates. Some methods convert the geographic coordinates to geocentric (X,Y,Z) coordinates, transform the X,Y,Z coordinates, and convert the new values back to geographic coordinates.
- Transformation results in either:

i. Changes in position and direction

ii. Uniform change of scale

iii. Changes in size and shape

* *Different types of transformation: search yourself*

Integration of cadastral data

Integration of cadastral data

A Cadastre is normally a parcel based, and up-to-date land information system containing a record of interests in land. For this purpose it is necessary to collect data from different source. Either primary or secondary data and integrate them centrally and distribute it when required.

Types of cadastral data integration

There are many different types of data integration followed in different parts of the world . We mainly focus on two type which is currently used in Nepalese system of cadaster.

i. Database preparation by resurveying in real field.

- In this method spatial data collection is performed by total station in the field. Along with the spatial data there will be the requirement of attribute data of all the parcels surveyed. We may collect those data primarily or we also may use the existing data secondarily. Point data is then imported in the GIS software. All the points are joined to form a parcel as in the real field and a database is prepared by adding the attribute data respectively.

Integration of cadastral data

Types of cadastral data integration

ii. Database preparation by digitizing the existing maps

- In this method, the paper maps which were previously prepared by different surveying techniques with comparatively low accuracy than the first method are scanned in the scanners to form a softcopy maps. The parcel boundaries are then digitized in a feature class within a database. Attribute data are added in each of the parcel to prepare a final data base.

There are various other methods of data integration but in context on Nepal above mentioned methods are mostly used which are also applicable in most of the parts of world.

References

Data Integration

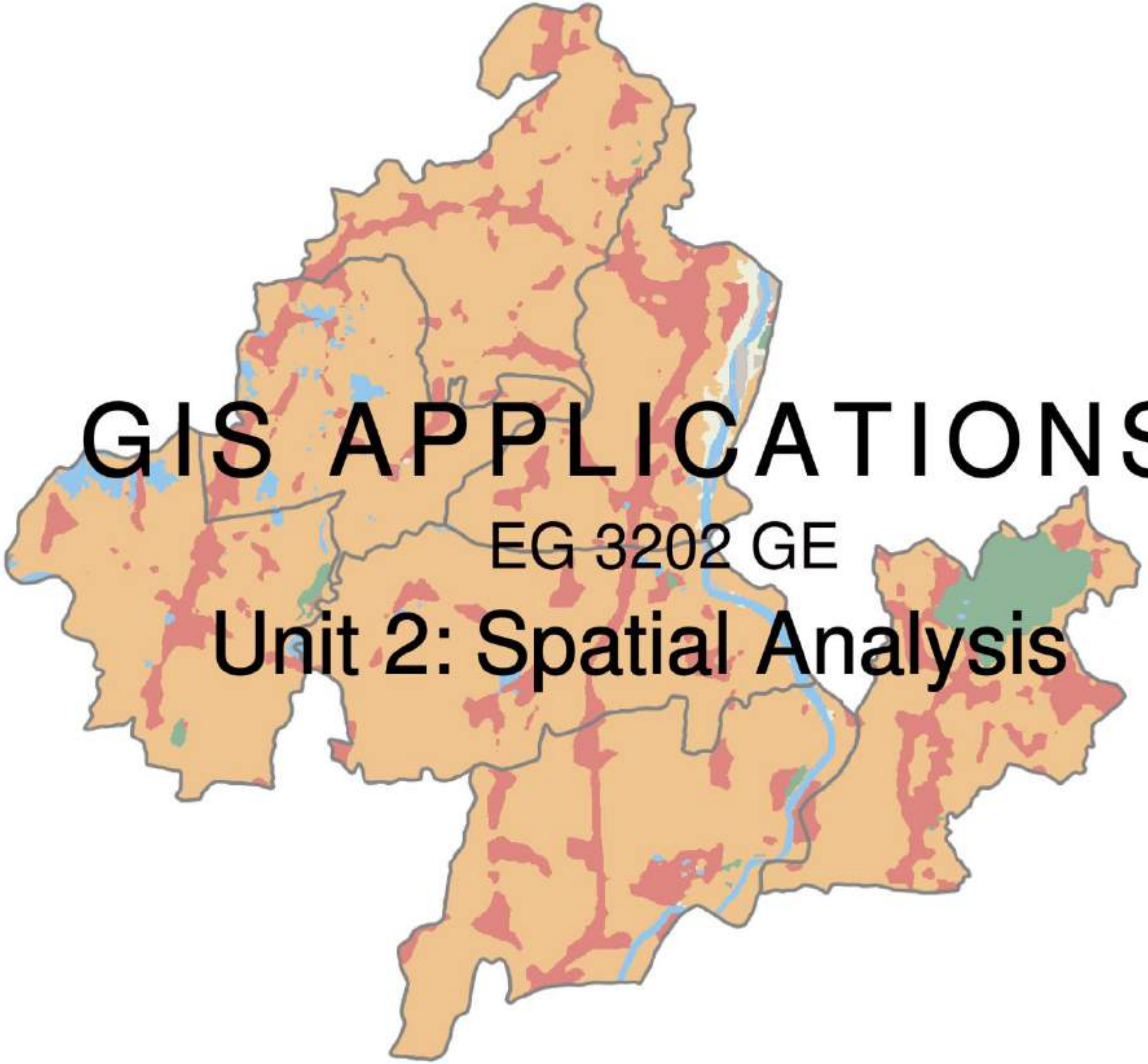
- *https://www.yorku.ca/gisweb/eats4400/lecture%20notes/lecture1.pdf*

Spatial Data Integration

- *https://link.springer.com/referencework entry/10.1007/978-3-319-63962-8_218-1#:~:text=Apart%20from%20facilitating%20complex%20geospatial,%2C%20compatibility%20with%20other%20datasets).*

- *https://auth0.com/blog/all-you-need-to-know-about-data-management-and-integration/*
- *https://webapps.itc.utwente.nl/librarywww/papers_2009/general/principlesgis.pdf*
- *http://mjcetce409.blogspot.com/2015/10/transformation-in-gis.html*
- *https://gisgeography.com/spatial-join/*
- *http://gsp.humboldt.edu/olm_2018/Lessons/GIS/06_2%20Vector%20Analysis%20Overlays/05_JoinsandRelates.html*

Suggestions:
Try to relate the definitions with the practicals.



GIS APPLICATIONS

EG 3202 GE

Unit 2: Spatial Analysis

Public Namuna Secondary School

Lecture by

Er. Keshav Raj Bhusal

Table of contents

2.1 Introduction

What is Spatial Analysis?

- It is the process of transformations, manipulations, and methods that can be applied to geographic data to add value to them, to support decisions, and to reveal patterns and anomalies that are not immediately obvious.
- Spatial analysis turns raw data into useful information.
- It is a set of methods whose results change when the locations of the objects being analyzed, or the frame used to analyze them, changes.
- Examples: hospital site suitability analysis, crime analysis etc.

2.1 Introduction

John Snow Cholera Map

Source: https://www.youtube.com/watch?v=lNjrAXGRda4&ab_channel=HarvardX

2.1 Introduction

Why spatial analysis is important?

Spatial analysis is important for the following reasons:

i. Spatial analysis helps in recognizing patterns from the distribution of features.

ii. It helps in identifying change over time. E.g landuse change over time, deforestation, urban expansion etc.

iii. It enables spatial modeling (or suitability analysis) and helps in decision making. E.g identifying best location to open a new school.

iv. It helps in prediction of certain phenomenon. E.g drought prediction

2.1 Introduction

Types of spatial analysis (Functions of spatial analysis)

On the basis of data structure

1. Vector-based Analysis

i. Data management functions
ii. Data Analysis functions
iii. Measurement function
iv. Spatial Selection
v. Vector Overlay functions

2.1 Introduction

Types of spatial analysis (Functions of spatial analysis)

On the basis of data structure

2. Raster-based Analysis

i. Spatial Interpolation methods
ii. Raster analysis including topological overlay
iii. Map calculations
iv. Spread computations
v. Classification function
vi. Reclassification

Classification of analytical GIS capabilities (Types of Spatial Analysis)

1. Classification, retrieval, measurement and generalization functions
2. Overlay functions
3. Neighborhood functions

- Search function
- Buffer zone generation
- Interpolation functions
- Topographic functions (slope angle, slope aspect, slope length, contour line)

4. Connectivity function(*these all are done on a network*)

- Network functions
- Visibility functions
- Contiguity functions

2.2 Classification and Measurement function

Coordinate and Distance measurement in vector data

- The location property of a vector feature is always stored by the GIS: a single coordinate pair for a point, or a list of pairs for a polyline or polygon boundary.
- Measuring distance between two features is another important function. If both features are points, say p and q, the computation in a Cartesian spatial reference system are given by the well-known Pythagorean distance function.

$$dist(p, q) = \sqrt{(x_p - x_q)^2 + (y_p - y_q)^2}$$

Coordinate and Distance measurement in Raster data

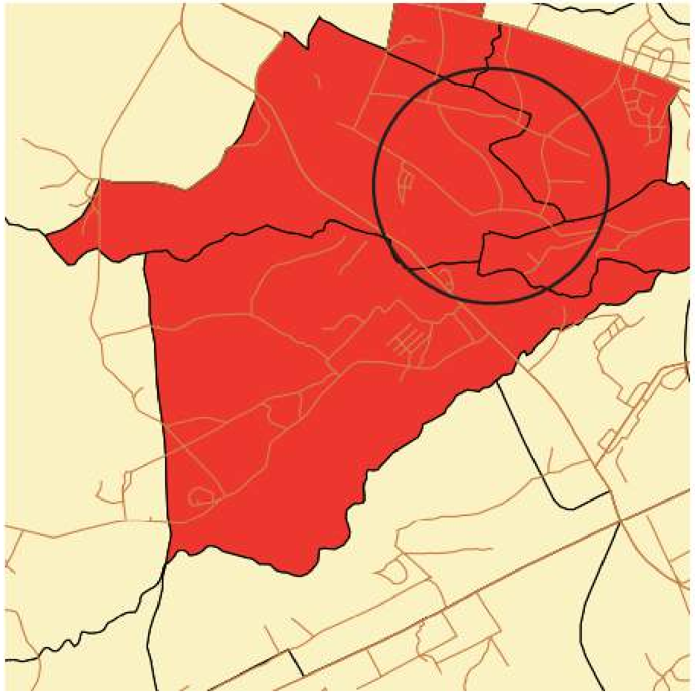
- The distance between two raster cells is the standard distance function applied to the locations of their respective mid-points, obviously taking into account the cell resolution.
- Where a raster is used to represent line features as strings of cells through the raster, the length of a line feature is computed as the sum of distances between consecutive cells.

Spatial selection queries

These queries are used to select certain features. Some of such techniques are:

1. Interactive spatial selection

- In such selection user point or draw spatial objects on screen. Such interactively defined objects are called ‘selection objects’ and can be point, line, polygon etc. GIS then select such features which are in overlap, intersect by such ‘selection objects’.



Area	Perimeter	Ward_id	Ward_nam	District	Pop88	Pop92
65420380.0000	41654.940000	1	KUNDUCHI	Kinondoni	22106	27212.00
24813620.0000	30755.620000	2	KAWE	Kinondoni	32854	40443.00
18698500.0000	26403.580000	3	MSASANI	Kinondoni	51225	63058.00
81845610.0000	49645.160000	4	UBUNGO	Kinondoni	47281	58203.00
4468546.00000	13480.130000	5	MANZESE	Kinondoni	59467	73204.00
4999599.00000	10356.850000	6	TANDALE	Kinondoni	58357	71837.00
4102218.00000	8951.096000	7	MWANANYAMALA	Kinondoni	72956	89809.00
3749840.00000	9447.420000	8	KINONDONI	Kinondoni	42301	52073.00
2087509.00000	7502.250000	9	UPANGA WEST	Ilala	9852	11428.00
2268513.00000	9028.788000	10	KIVUKONI	Ilala	5391	6254.00
1400024.00000	6883.288000	11	NDUGUMBI	Kinondoni	32548	40067.00
888966.900000	4589.110000	12	MAGOMENI	Kinondoni	16938	20851.00
1448370.00000	5651.958000	13	UPANGA EAST	Ilala	11019	12782.00
8214378.00000	14552.080000	14	MABIBO	Kinondoni	43381	53402.00
2498622.00000	7121.255000	15	MAKURUMILA	Kinondoni	54141	66648.00
1262028.00000	4885.793000	16	MZIMUNI	Kinondoni	23989	29530.00
35362240.0000	28976.090000	17	KINYEREZI	Ilala	3044	3531.00
1010613.00000	5393.771000	18	JANGIWANI	Ilala	15297	17745.00
475745.500000	3043.068000	19	KISUTU	Ilala	8399	9743.00
1754043.00000	7743.187000	20	KIGOGO	Kinondoni	21267	26180.00
29964950.0000	36964.000000	21	KIGAMBONI	Temeke	23203	27658.00
1291479.00000	5187.690000	22	MICHIKICHINI	Ilala	14852	17228.00
720322.100000	4342.732000	23	MCHAFUKOGE	Ilala	8439	9789.00
9296131.00000	16321.530000	24	TABATA	Ilala	18454	21407.00
483620.700000	3304.072000	25	KARIAKOO	Ilala	12506	14507.00
3564653.00000	9586.751000	26	BUGURUNI	Ilala	48286	56012.00
2639575.00000	6970.186000	27	ILALA	Ilala	35372	41032.00
912452.800000	4021.937000	28	GEREZANI	Ilala	7490	8688.00
6735135.00000	13579.590000	29	KURASINI	Temeke	26737	31871.00

Source: Principles of Geographic Information System - Rolf A. de By (ed.) (ITC Education Text Book Series; 1)

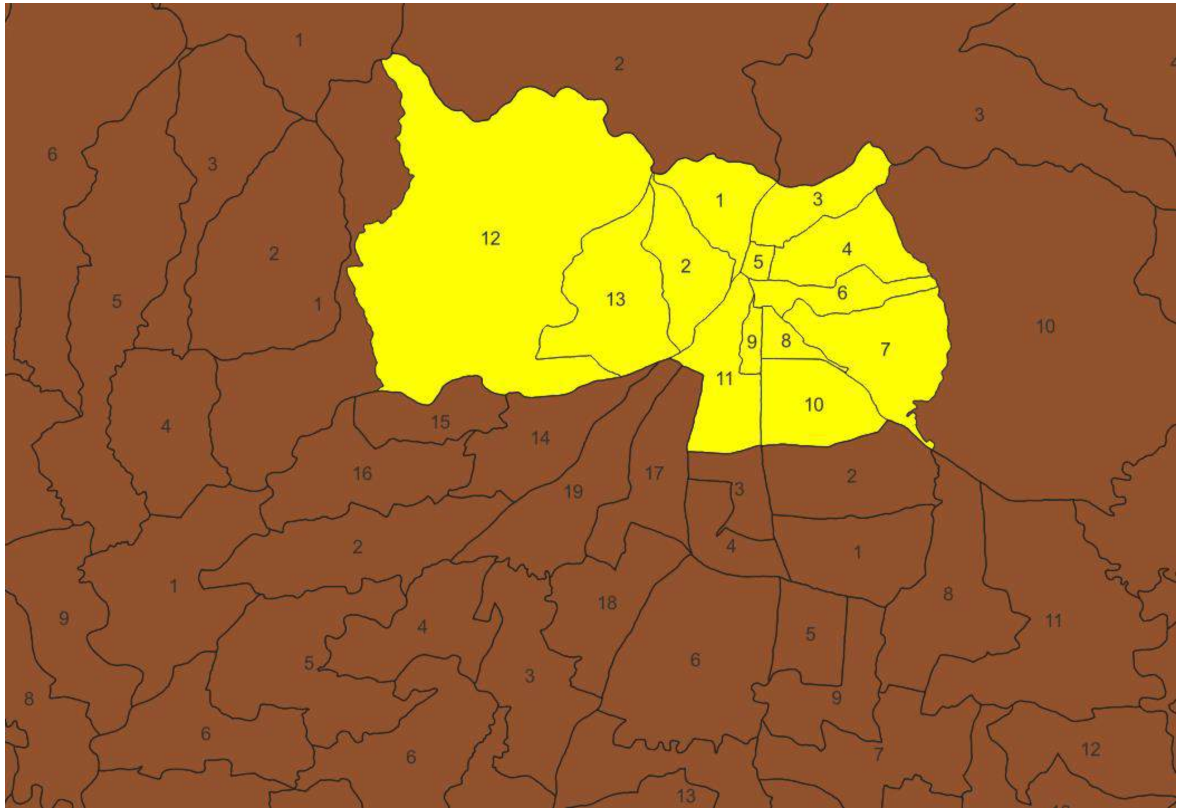
Fig: showing selection of 4 wards of a city represented by red (selected features) by selection object (circle) in [left] and attribute table selection in [right].

2.2 Classification and Measurement function

Spatial selection queries

2. Spatial selection by attribute conditions

- In such queries feature is selected using selected conditions on feature attributes. These conditions are formulated in SQL if attribute data reside in a geodatabase.



hermes_NPL_wgs_4 — Features Total: 6730, Filtered: 6730, Selected: 13

	PROVINCE	PR_NAME	DISTRICT	PALIKA	TYPE	WARD
5458	5	Province No 5	RUPANDEHI	Butwal	Upamahanagar...	1
5459	5	Province No 5	RUPANDEHI	Butwal	Upamahanagar...	2
5460	5	Province No 5	RUPANDEHI	Butwal	Upamahanagar...	3
5461	5	Province No 5	RUPANDEHI	Butwal	Upamahanagar...	4
5462	5	Province No 5	RUPANDEHI	Butwal	Upamahanagar...	5
5463	5	Province No 5	RUPANDEHI	Butwal	Upamahanagar...	6
5464	5	Province No 5	RUPANDEHI	Butwal	Upamahanagar...	7
5465	5	Province No 5	RUPANDEHI	Butwal	Upamahanagar...	8
5466	5	Province No 5	RUPANDEHI	Butwal	Upamahanagar...	9
5467	5	Province No 5	RUPANDEHI	Butwal	Upamahanagar...	10
5468	5	Province No 5	RUPANDEHI	Butwal	Upamahanagar...	11
5469	5	Province No 5	RUPANDEHI	Butwal	Upamahanagar...	12
5470	5	Province No 5	RUPANDEHI	Butwal	Upamahanagar...	13
5471	5	Province No 5	RUPANDEHI	Butwal	Upamahanagar...	14
5472	5	Province No 5	RUPANDEHI	Butwal	Upamahanagar...	15
5473	5	Province No 5	RUPANDEHI	Butwal	Upamahanagar...	16
5474	5	Province No 5	RUPANDEHI	Butwal	Upamahanagar...	17
5475	5	Province No 5	RUPANDEHI	Butwal	Upamahanagar...	18
5476	5	Province No 5	RUPANDEHI	Butwal	Upamahanagar...	19

Tutorial: https://youtu.be/GRcWB5zg5Uc

Fig: showing selection of wards of a Butwal represented by yellow (selected features) with Ward<=13 in [left] and attribute table selection in [right].
Expression used: "PALIKA" = 'Butwal' AND "WARD" <=13

Spatial selection queries

2.1 Combining attribute conditions

- In such queries feature is selected using combined conditions on feature attributes. These conditions are formulated in SQL if attribute data reside in a geodatabase.

E.g selecting Rohini and ward 3 and 5 of Omsatiya

("PALIKA" = 'Rohini') OR

("PALIKA" = 'Omsatiya' AND ("WARD" >=3 AND "WARD" <=5))

3. Selecting the features based on their distance

- Such selection can be done within a given distance from the selection objects at a given distance or even beyond the given distance.

- Which clinics are within the 2 kms of a selected school

- Which roads are within 200 mtrs of a medical clinic

4. Selecting the feature that intersect

Spatial selection using the intersect relationship between lines and polygon. E.g all buildings within that road.(*see Lab 1.5*)

2.2 Classification and Measurement function

Classification of data (manual and automatic classification of data)

- Classification is a technique of purposefully removing detail from an input data set, in the hope of revealing important patterns (of spatial distribution).
- We do so by giving a characteristic value to each element in the input set which is usually a collection of spatial features that can be raster cells or points, lines or polygons.
- In classification of vector data, there are two possible results i) nothing changes with respect to spatial extents of the original features. ii) adjacent features with the same category are merged into one bigger feature.
- For example: The pattern that we look for may be the distribution of household income in a city. Household income is called the classification parameter. If we know for each ward in the city the associated average income, we have many different values. Subsequently, we could define five different categories (or: classes) of income: ‘low’, ‘below average’, ‘average’, ‘above average’ and ‘high’, and provide value ranges for each category. If these five categories are mapped in a sensible colour scheme, this may reveal interesting information.

Classification of data (manual and automatic classification of data)

Household income range	New category value
391–2474	1
2475–6030	2
6031–8164	3
8165–11587	4
11588–21036	5

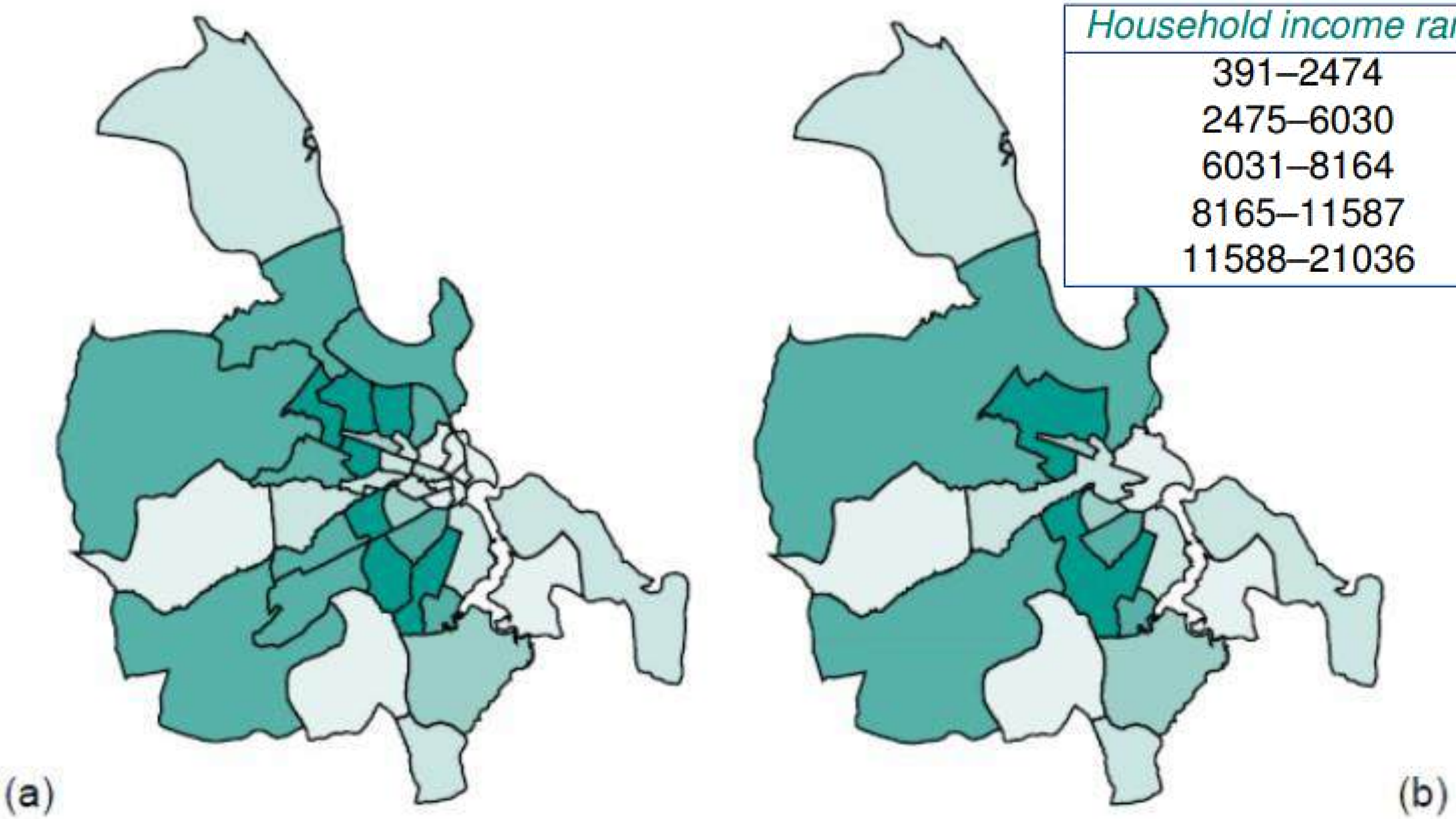


Figure 6-11: Two Classifications (a) with Original Polygons Intact (b) With Original Polygons Merged

Source: Principles of Geographic Information System - Rolf A. de By (ed.) (ITC Education Text Book Series; 1)

2.2 Classification and Measurement function

User-controlled (manual) classification of data

- In manual classification user selects the attribute that will be used as the classification parameter and defines the classification method.
- Classification method involves declaring the number of classes as well as the equivalance between the old attribute values and the new classes usually via a classification table.
- In such classification we must define the data format of the output, as a spatial data layer, which will contain the new classification attribute. The data type of this attribute is always categorical, i.e. integer or string, no matter what is the data type of the attribute(s) from which the classification was obtained.

Here classification parameter is discrete.

Code	*Old category*	*New category*
10	Planned residential	Residential
20	Industry	Commercial
30	Commercial	Commercial
40	Institutional	Public
50	Transport	Public
60	Recreational	Public
70	Non built-up	Non built-up
80	Unplanned residential	Residential

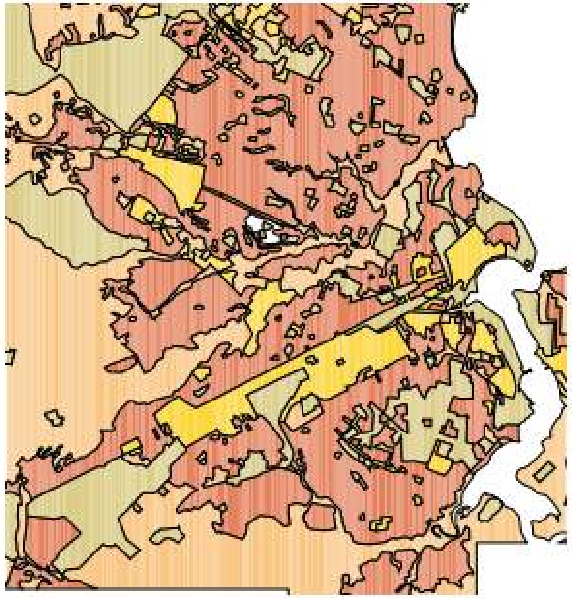


Figure 6.10: An example of a classification on a discrete parameter, namely land use unit in the city of Dar es Salaam, Tanzania. Colour scheme: Residential (brown), Commercial (yellow), Public (Olive), Non built-up (orange). Data source: Dept. of Urban & Regional Planning and Geo-information Management, ITC.

2.2 Classification and Measurement function

Automatic classification of data

- It is a classification where GIS software automatically does the classification, in which a user only specifies the number of classes in the output data set. The system automatically determines the class break points.

Two main techniques of determining break points are:

a. Equal interval technique

- In this technique the minimum and maximum values vmin and vmax of the classification parameter are determined and the (constant) interval size for each category is calculated as (vmax – vmin)/n, where n is the number of classes chosen by the user.
- This classification is useful in revealing the distribution patterns as it determines the number of features in each category.

For example, if you create 5 classes with attribute values from 0-100, the 5 classes will be 0-20, 21-40, 41-60, 61-80, and 81-100. In the example above, 100 is the maximum, and 0 is the minimum with 5 classes.

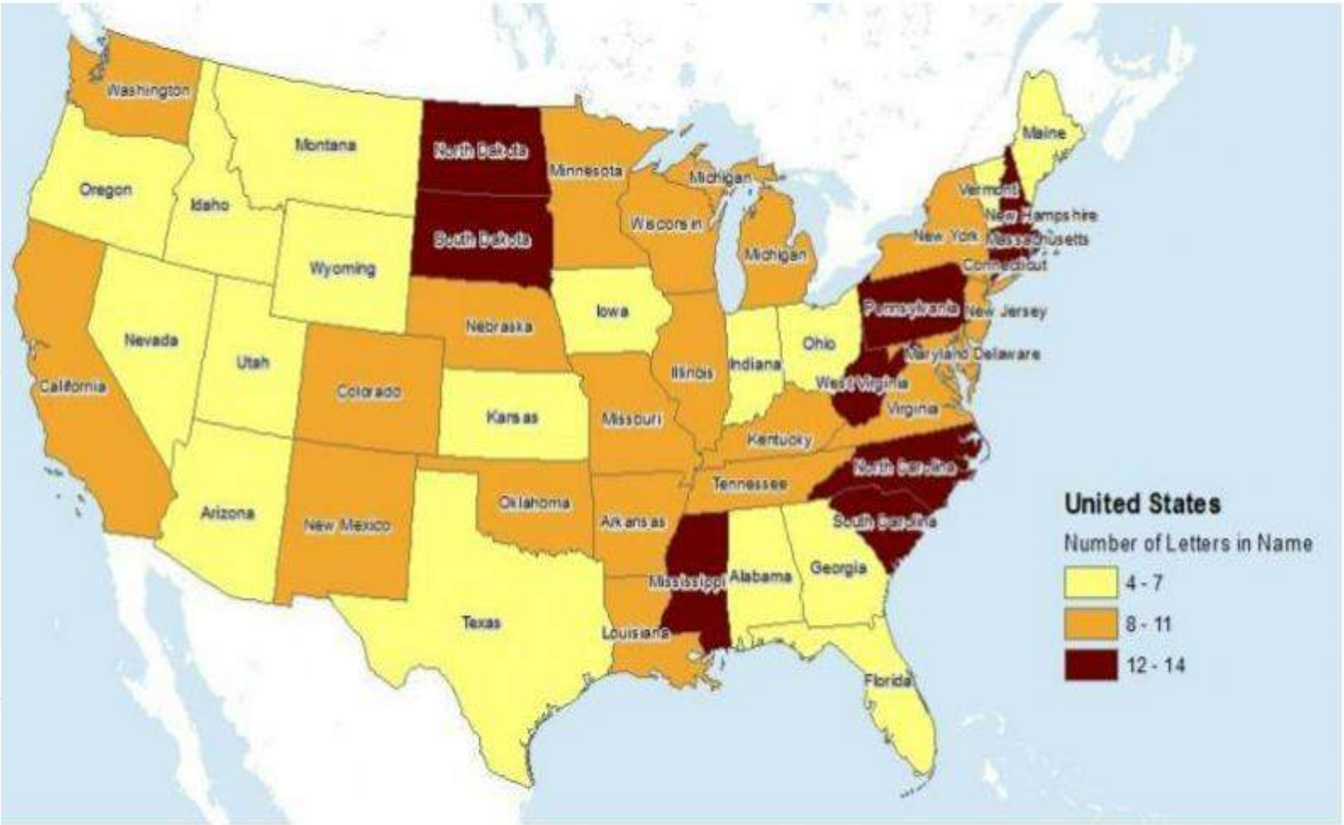
Automatic classification of data

a. Equal interval technique

Name	# of Letters	Name	# of Letters	Name	# of Letters
Iowa	4	Hawaii	6	Georgia	7
Ohio	4	Kansas	6	Indiana	7
Utah	4	Nevada	6	Montana	7
Idaho	5	Oregon	6	Vermont	7
Maine	5	Alabama	7	Wyoming	7
Texas	5	Arizona	7		
Alaska	6	Florida	7		

Name	# of Letters	Name	# of Letters	Name	# of Letters
Arkansas	8	Nebraska	8	California	10
Colorado	8	New York	8	New Jersey	10
Delaware	8	Oklahoma	8	New Mexico	10
Illinois	8	Virginia	8	Washington	10
Kentucky	8	Louisiana	9	Connecticut	11
Maryland	8	Minnesota	9	Mississippi	11
Michigan	8	Tennessee	9		
Missouri	8	Wisconsin	9		

Name	# of Letters	Name	# of Letters	Name	# of Letters
North Dakota	12	South Dakota	12	West Virginia	13
Pennsylvania	12	Massachusetts	13	North Carolina	14
Rhode Island	12	New Hampshire	13	South Carolina	14

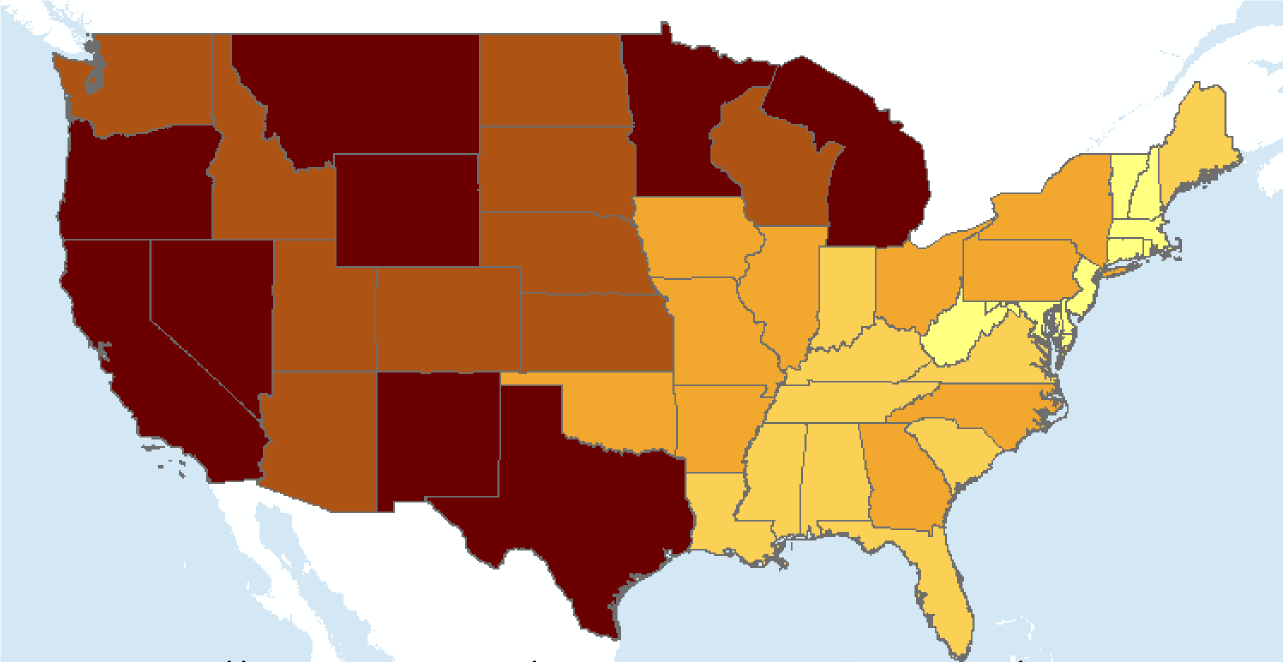


Source: https://gisgeography.com/equal-interval-classification-gis/

Automatic classification of data

b. Equal frequency technique(quantile classification)

- Quantile classification divides classes so that the total number of features in each class is approximately the same.
- In this technique the total number of features is determined first and by the required number of categories, the number of features per category is calculated.
- This technique is also known as quantile classification.



Quantile classification of US states with 5 classes; here no of total states(features) is 50.

Source: https://gisgeography.com/equal-interval-classification-gis/

2.2 Classification and Measurement function

Application of classification and measurement function

1. Useful for different local and central government agencies to keep the record of people living there and the data related to the people living there. Like, household income, etc.
2. Useful for classification of the data regarding the census of country.
3. Useful in agricultural agencies to classify the soil types.
4. Useful in cadastral survey and land administration to classify the land types.
5. Similarly there are various other sectors where the huge number of data needs to be classified.

2.3 Overlay functions and Neighborhood Analysis

Overlay functions

- It is a technique of combining two spatial data layers and producing a third from them using spatial overlay operators. For spatial overlay operators to work data layers should be georeferenced in the same system and overlap in study area.
- The principle of spatial overlay is to compare the characteristics of the same location in both data layers, and to produce a new characteristic for each location in the output data layer.
- Some of the Overlay functions are: Clip, Union, Intersect etc.

On basis of data structure there are two types of overlay functions. They are:

i. Vector Overlay Functions

ii. Raster Overlay Functions

Vector overlay operation

- It is the overlay operation which is done on the vector data using vector overlay operator.
- It can be done on point, line or polygon data layers.
- Vector overlay operators are useful, but geometrically complicated, and this sometimes results in poor operator performance.

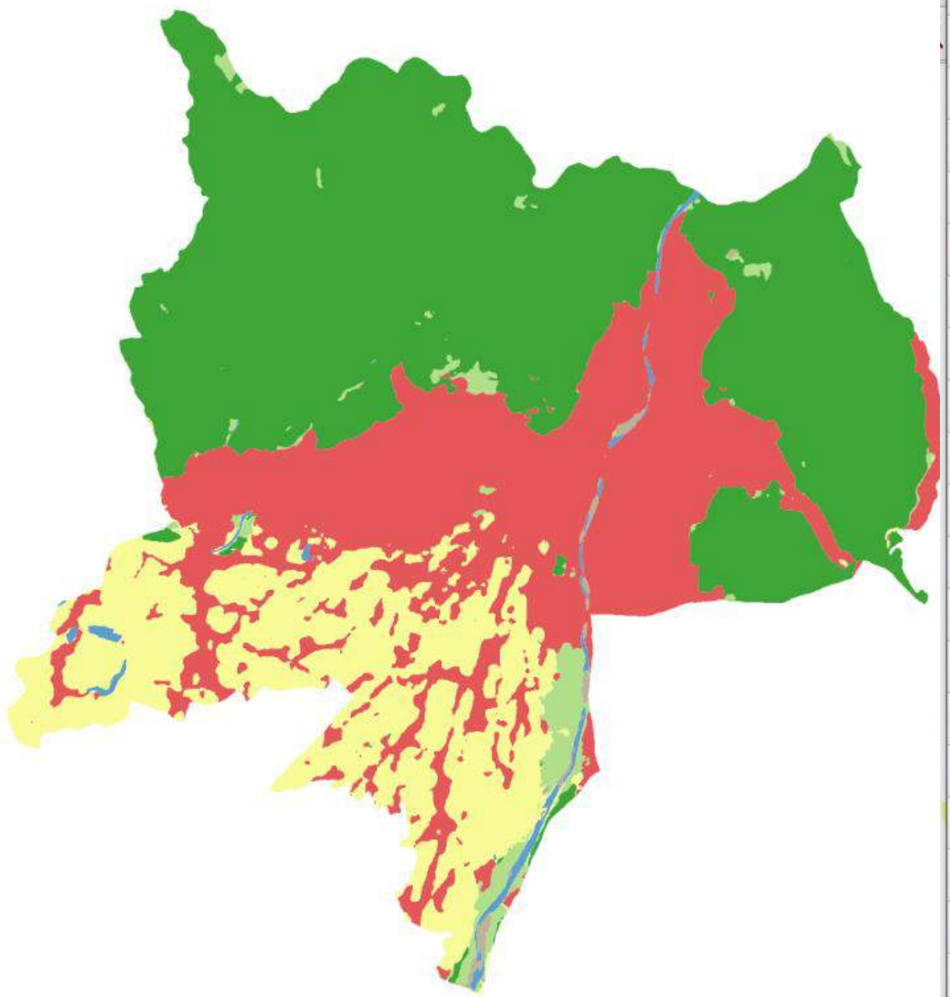
Some of the vector overlay operators are:

a. *Polygon intersection operator*

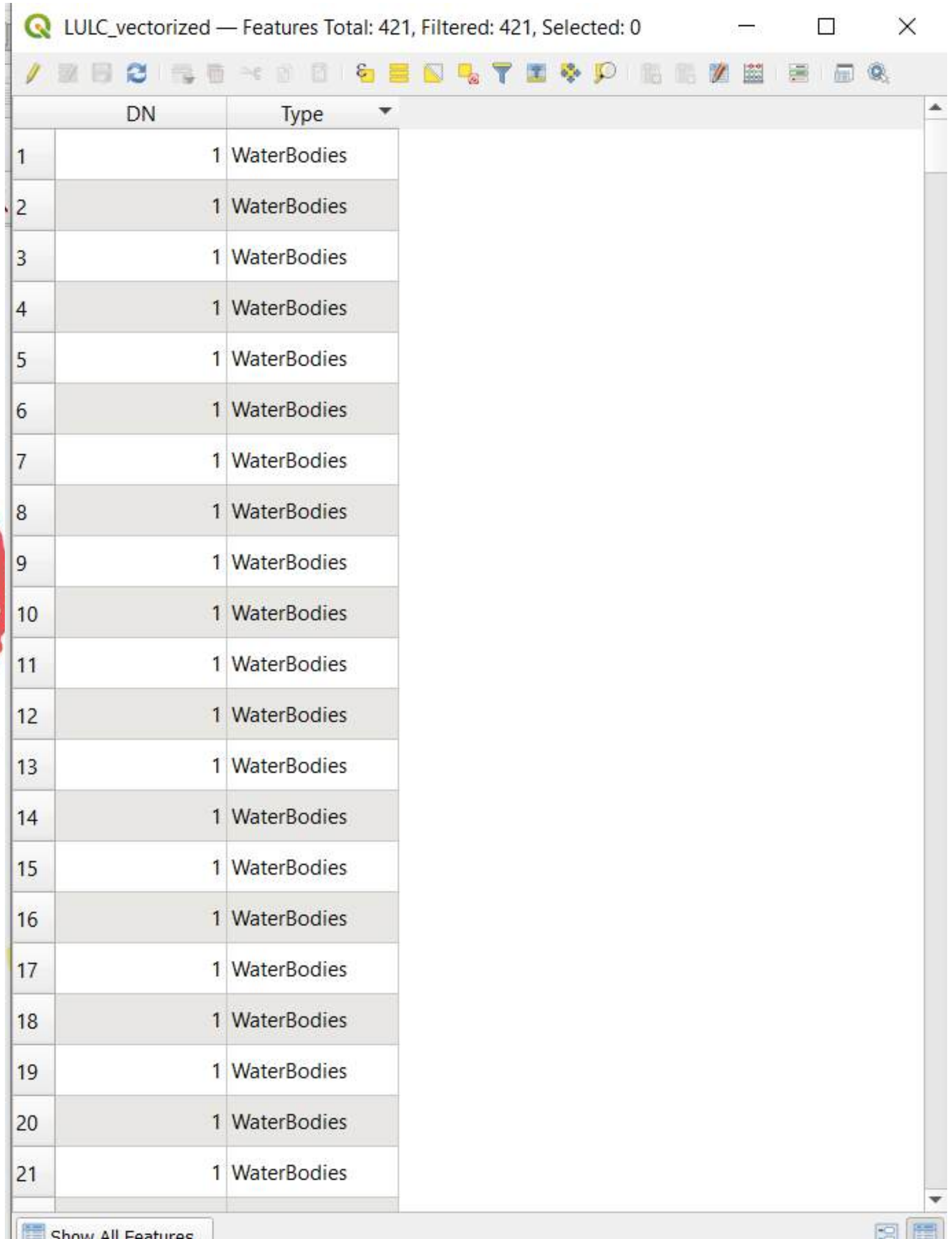
- It is the standard overlay operator for two layers of polygons.
- The result of this operator is the collection of all possible polygon intersections; the attribute table result is a join of the two input attribute tables.
- This operator is also called spatial join.

Vector overlay operation

a. Polygon intersection operator



First data layer



LULC_vectorized — Features Total: 421, Filtered: 421, Selected: 0

	DN	Type
1	1	WaterBodies
2	1	WaterBodies
3	1	WaterBodies
4	1	WaterBodies
5	1	WaterBodies
6	1	WaterBodies
7	1	WaterBodies
8	1	WaterBodies
9	1	WaterBodies
10	1	WaterBodies
11	1	WaterBodies
12	1	WaterBodies
13	1	WaterBodies
14	1	WaterBodies
15	1	WaterBodies
16	1	WaterBodies
17	1	WaterBodies
18	1	WaterBodies
19	1	WaterBodies
20	1	WaterBodies
21	1	WaterBodies

Show All Features

Source: https://youtu.be/nmM7a1vJx38

Vector overlay operation

a. *Polygon intersection operator*

Second data layer

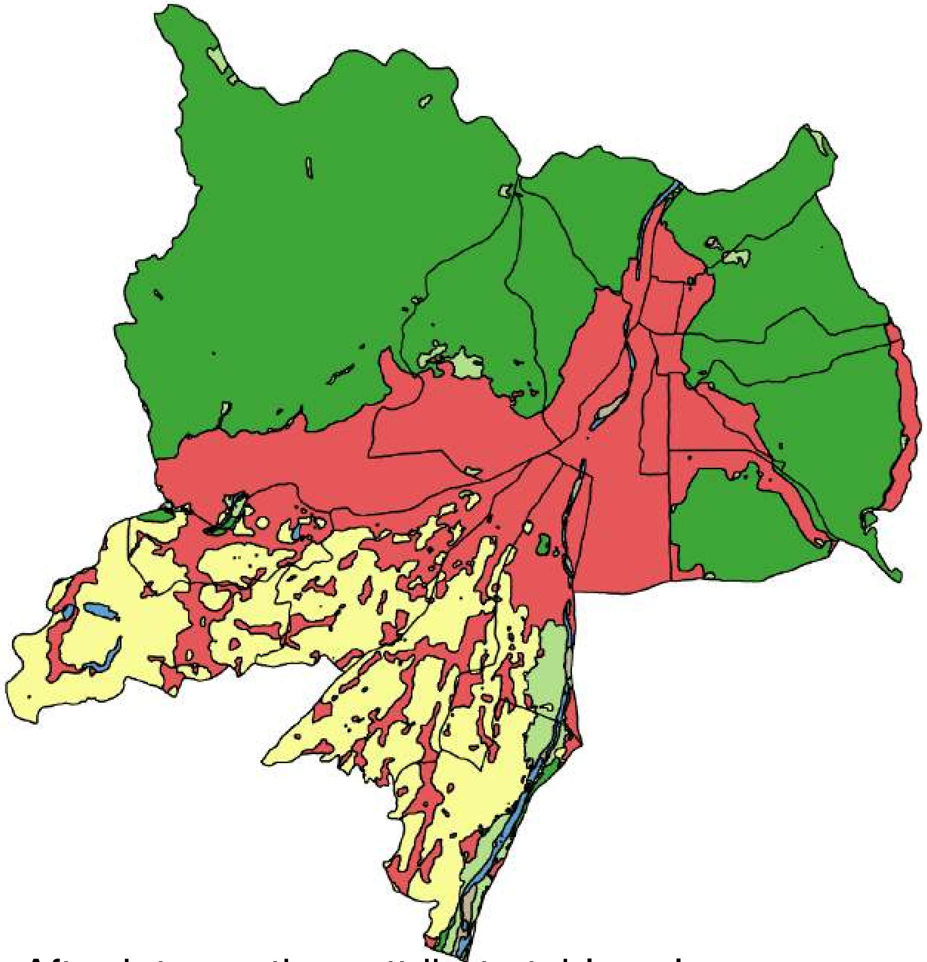
ButwalWards_44N — Features Total: 19, Filtered: 19, Selected: 0

	PROVINCE	PR_NAME	DISTRICT	PALIKA	TYPE	WARD
1	5	Province No 5	RUPANDEHI	Butwal	Upamahanagar...	1
2	5	Province No 5	RUPANDEHI	Butwal	Upamahanagar...	2
3	5	Province No 5	RUPANDEHI	Butwal	Upamahanagar...	3
4	5	Province No 5	RUPANDEHI	Butwal	Upamahanagar...	4
5	5	Province No 5	RUPANDEHI	Butwal	Upamahanagar...	5
6	5	Province No 5	RUPANDEHI	Butwal	Upamahanagar...	6
7	5	Province No 5	RUPANDEHI	Butwal	Upamahanagar...	7
8	5	Province No 5	RUPANDEHI	Butwal	Upamahanagar...	8
9	5	Province No 5	RUPANDEHI	Butwal	Upamahanagar...	9
10	5	Province No 5	RUPANDEHI	Butwal	Upamahanagar...	10
11	5	Province No 5	RUPANDEHI	Butwal	Upamahanagar...	11
12	5	Province No 5	RUPANDEHI	Butwal	Upamahanagar...	12
13	5	Province No 5	RUPANDEHI	Butwal	Upamahanagar...	13
14	5	Province No 5	RUPANDEHI	Butwal	Upamahanagar...	14
15	5	Province No 5	RUPANDEHI	Butwal	Upamahanagar...	15
16	5	Province No 5	RUPANDEHI	Butwal	Upamahanagar...	16
17	5	Province No 5	RUPANDEHI	Butwal	Upamahanagar...	17
18	5	Province No 5	RUPANDEHI	Butwal	Upamahanagar...	18
19	5	Province No 5	RUPANDEHI	Butwal	Upamahanagar...	19

Source: https://youtu.be/nmM7a1vJx38

Vector overlay operation

a. *Polygon intersection operator*



After intersection; attribute table values are joined.

LULCWards_Intersect — Features Total: 493, Filtered: 493, Selected: 0

	DN	Type	PROVINCE	PR_NAME	DISTRICT	PALIKA	TYPE_2	WARD
1	1	WaterBodies	5	Province No 5	RUPANDEHI	Butwal	Upamahanagar...	1
2	1	WaterBodies	5	Province No 5	RUPANDEHI	Butwal	Upamahanagar...	1
3	1	WaterBodies	5	Province No 5	RUPANDEHI	Butwal	Upamahanagar...	2
4	1	WaterBodies	5	Province No 5	RUPANDEHI	Butwal	Upamahanagar...	12
5	1	WaterBodies	5	Province No 5	RUPANDEHI	Butwal	Upamahanagar...	11
6	1	WaterBodies	5	Province No 5	RUPANDEHI	Butwal	Upamahanagar...	12
7	1	WaterBodies	5	Province No 5	RUPANDEHI	Butwal	Upamahanagar...	12
8	1	WaterBodies	5	Province No 5	RUPANDEHI	Butwal	Upamahanagar...	11
9	1	WaterBodies	5	Province No 5	RUPANDEHI	Butwal	Upamahanagar...	11
10	1	WaterBodies	5	Province No 5	RUPANDEHI	Butwal	Upamahanagar...	11
11	1	WaterBodies	5	Province No 5	RUPANDEHI	Butwal	Upamahanagar...	11
12	1	WaterBodies	5	Province No 5	RUPANDEHI	Butwal	Upamahanagar...	11
13	1	WaterBodies	5	Province No 5	RUPANDEHI	Butwal	Upamahanagar...	11
14	1	WaterBodies	5	Province No 5	RUPANDEHI	Butwal	Upamahanagar...	12
15	1	WaterBodies	5	Province No 5	RUPANDEHI	Butwal	Upamahanagar...	12
16	1	WaterBodies	5	Province No 5	RUPANDEHI	Butwal	Upamahanagar...	11
17	1	WaterBodies	5	Province No 5	RUPANDEHI	Butwal	Upamahanagar...	11
18	1	WaterBodies	5	Province No 5	RUPANDEHI	Butwal	Upamahanagar...	12
19	1	WaterBodies	5	Province No 5	RUPANDEHI	Butwal	Upamahanagar...	11
20	1	WaterBodies	5	Province No 5	RUPANDEHI	Butwal	Upamahanagar...	11
21	1	WaterBodies	5	Province No 5	RUPANDEHI	Butwal	Upamahanagar...	17

Vector overlay operation

b. Polygon clipping operator

- It takes a polygon data layer and restricts its spatial extent to the outer boundary of second input layer polygon.

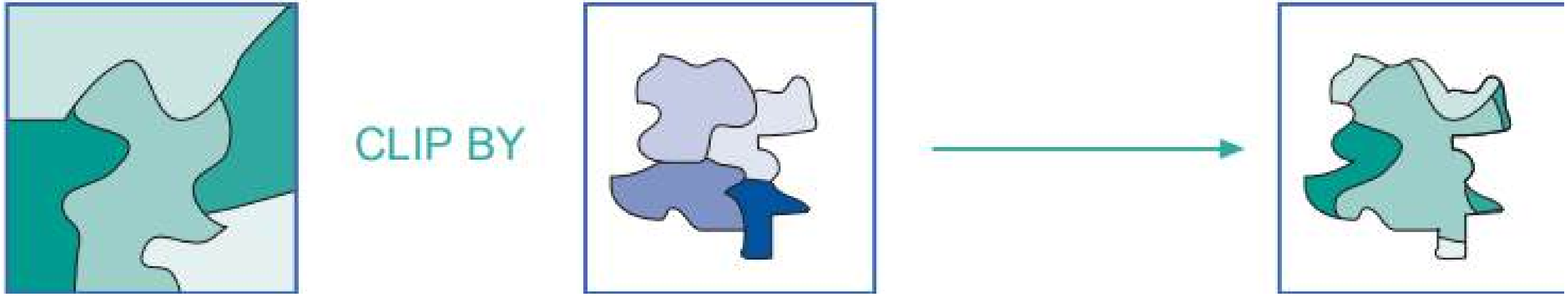


Fig: showing polygon clipping operation

Source: Principles of Geographic Information System - Rolf A. de By (ed.) (ITC Education Text Book Series; 1)

Raster overlay operation

- It is the overlay operation which is done on the raster data.
- Raster overlays do not suffer from geometrical complication, as most of them perform their computations cell by cell, and thus they are fast.
- GISs that support raster processing—as most do—usually have a language to express operations on rasters. These languages are generally referred to as map algebra or sometimes raster calculus(raster calculator in case of QGIS).

When producing a new raster a statement of the following format shall be provided:

Output raster name = Raster calculus expression

- **When the expression is evaluated,** the GIS will perform the calculation on a pixel by pixel basis, starting from the first pixel in the first row, and continuing until the last pixel in the last row.

Raster overlay operation

There is a wide range of operators and functions that can be used in raster calculus. Some of them are:

a. Arithmetic Operators

- It includes multiplication (×), division (/), subtraction (–) and addition (+).
- It also includes modulo division (MOD) and integer division (DIV). Modulo division returns the remainder of division: for instance, 10 MOD 3 will return 1 as 10–9 = 1. Similarly, 10 DIV 3 will return 3.
- More operators are goniometric: sine (sin), cosine (cos), tangent (tan), and their inverse functions asin, acos, and atan, which return radian angles as real values

2.3 Overlay functions and Neighborhood Analysis

Raster overlay operation

There is a wide range of operators and functions that can be used in raster calculus. Some of them are:

a. Arithmetic Operators

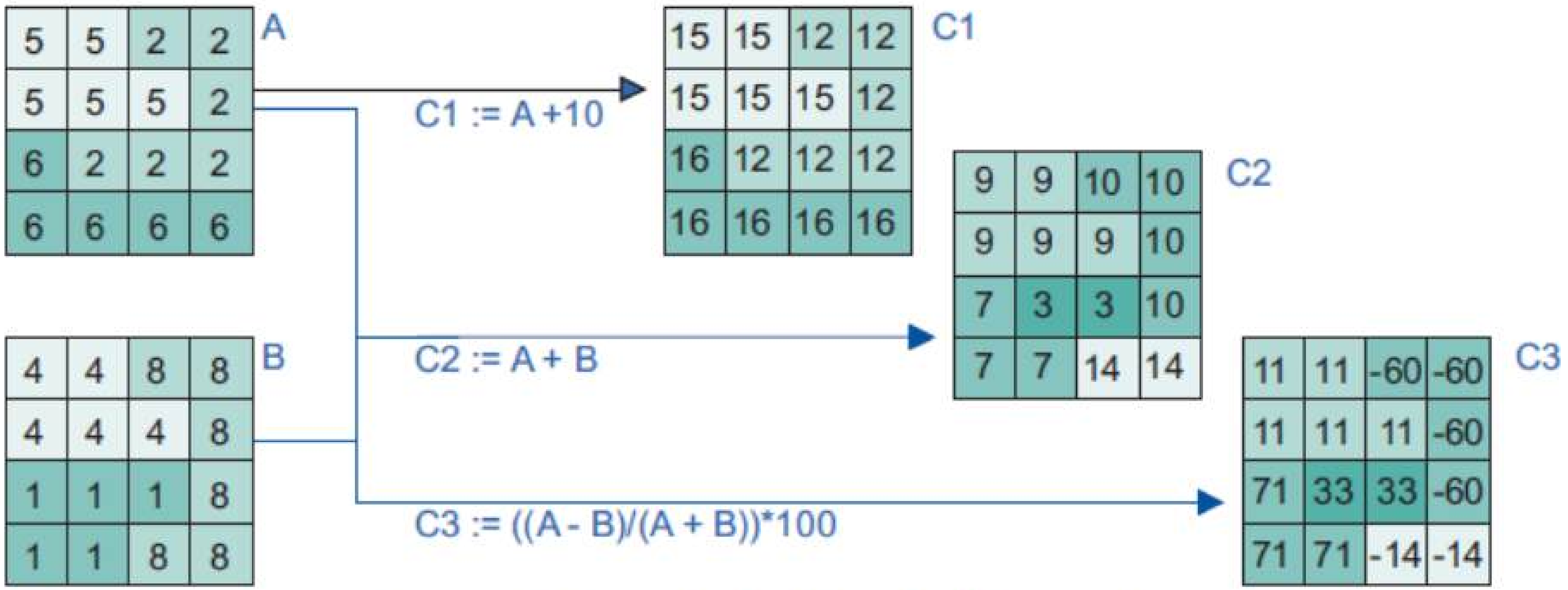


Figure 6-8: Examples of Arithmetic Operators

Source: Principles of Geographic Information System - Rolf A. de By (ed.) (ITC Education Text Book Series; 1)

2.3 Overlay functions and Neighborhood Analysis

Raster overlay operation

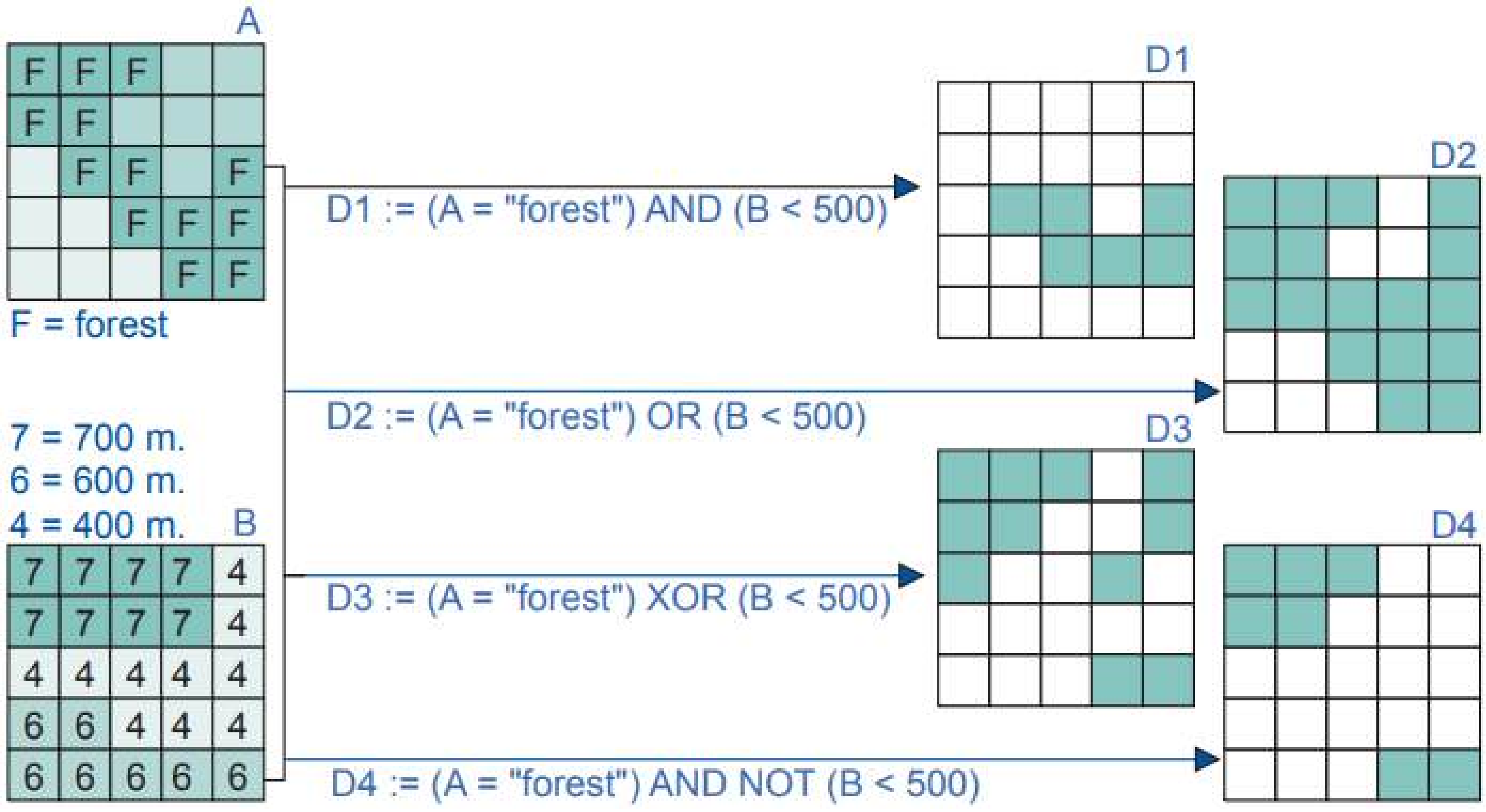
b. Comparison and Logical Operators

- Comparison operators includes (<=, =, >=, > and <>) operation.
- The comparison and logical operators produce rasters with the truth values true and false.
- Logical connectives such as AND, OR and NOT are also supported in many raster calculation. Another connective that is commonly offered in raster calculation is exclusive OR (XOR).

The expression a XOR b is true only if either a or b is true, but not both.

Raster overlay operation

b. Comparison and Logical Operators



Source: Principles of Geographic Information System - Rolf A. de By (ed.) (ITC Education Text Book Series; 1)

Examples of logical expressions in map algebra. Green cells represent true values, white cells represent false values. Here A is a classified raster for land use, and B holds elevation values.

2.3 Overlay functions and Neighborhood Analysis

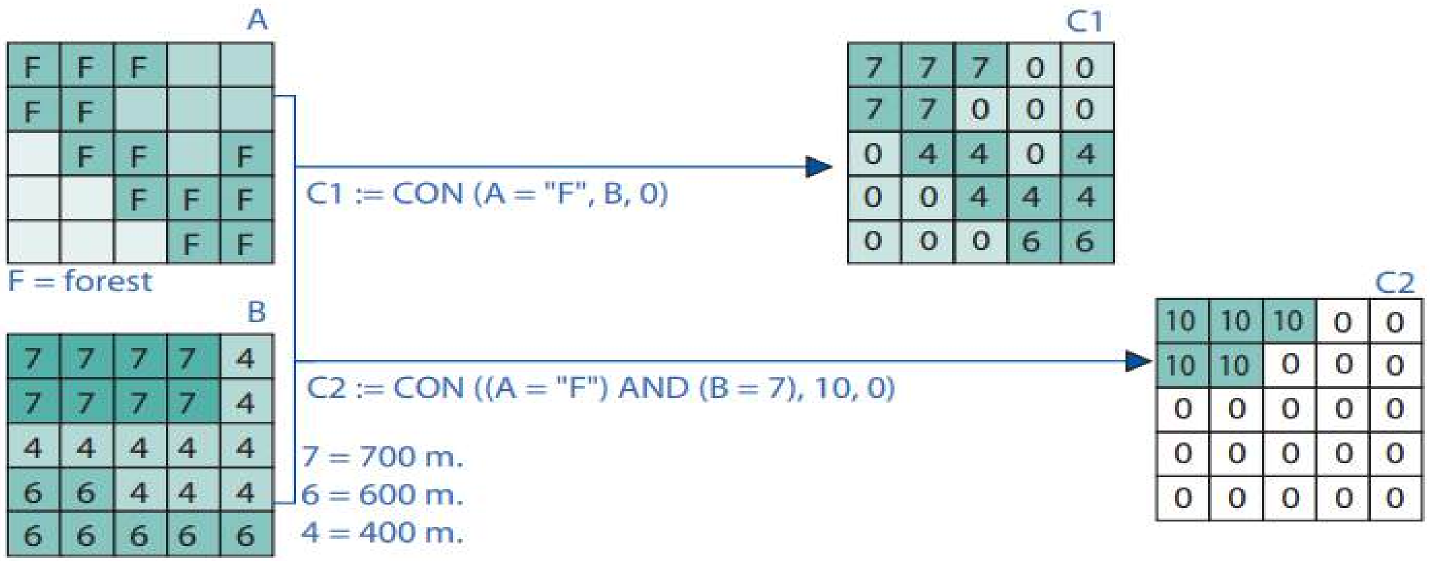
Raster overlay operation

c. Conditional Expressions

- The general format is:

Output raster = *CON (condition, then expression, else expression).*

Here, *condition* is the tested condition, *then_expression* is evaluated if condition holds, and *else_expression* is evaluated if it does not hold.



Examples of conditional expressions in map algebra. Here A is a classified raster holding land use data, and B is an elevation value raster.

Source: Principles of Geographic Information System - Rolf A. de By (ed.) (ITC Education Text Book Series; 1)

2.3 Neighborhood function (Generating buffer zone)

Neighborhood function

- Neighborhood operations evaluate the characteristics of an area surrounding a specific location

To perform neighborhood analysis, we must:

i. State which target locations are of interest to us, and define their spatial extent.

ii. Define how to determine the neighborhood for each target.

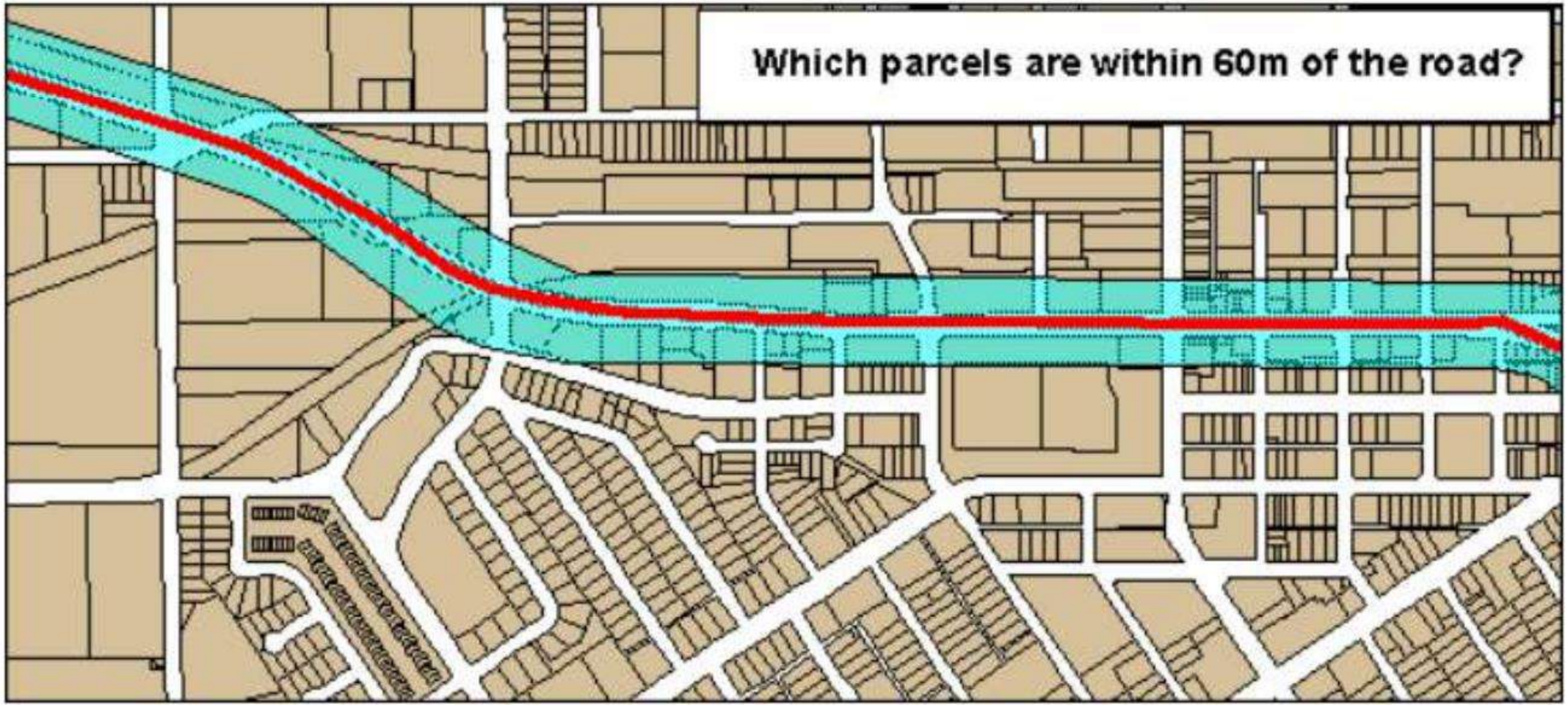
iii. Define which characteristic(s) must be computed for each neighborhood (we indicate what it is we want to discover about the phenomena that exist or occur in the neighborhood. This might simply be its spatial extent, but it might also be statistical information).

- E.g Our target might be a road. Its neighborhood could be defined as: parcels within 60 m of road, characteristic we might be looking could be spatial extent of parcels.

Some of the neighborhood function are buffer zone generation, thiessen polygon generation, flow computation etc.

2.3 Neighborhood function (Generating buffer zone)

Neighborhood function



Source: : https://planet.uwc.ac.za/nisl/GIS/Web_Page/page_3 6.htm

2.3 Neighborhood function

1. Proximity computations/analysis

• We use geometric distance to define the neighborhood of one or more target locations.

• The most common and useful technique is buffer zone generation.

1.1 Buffer zone generation

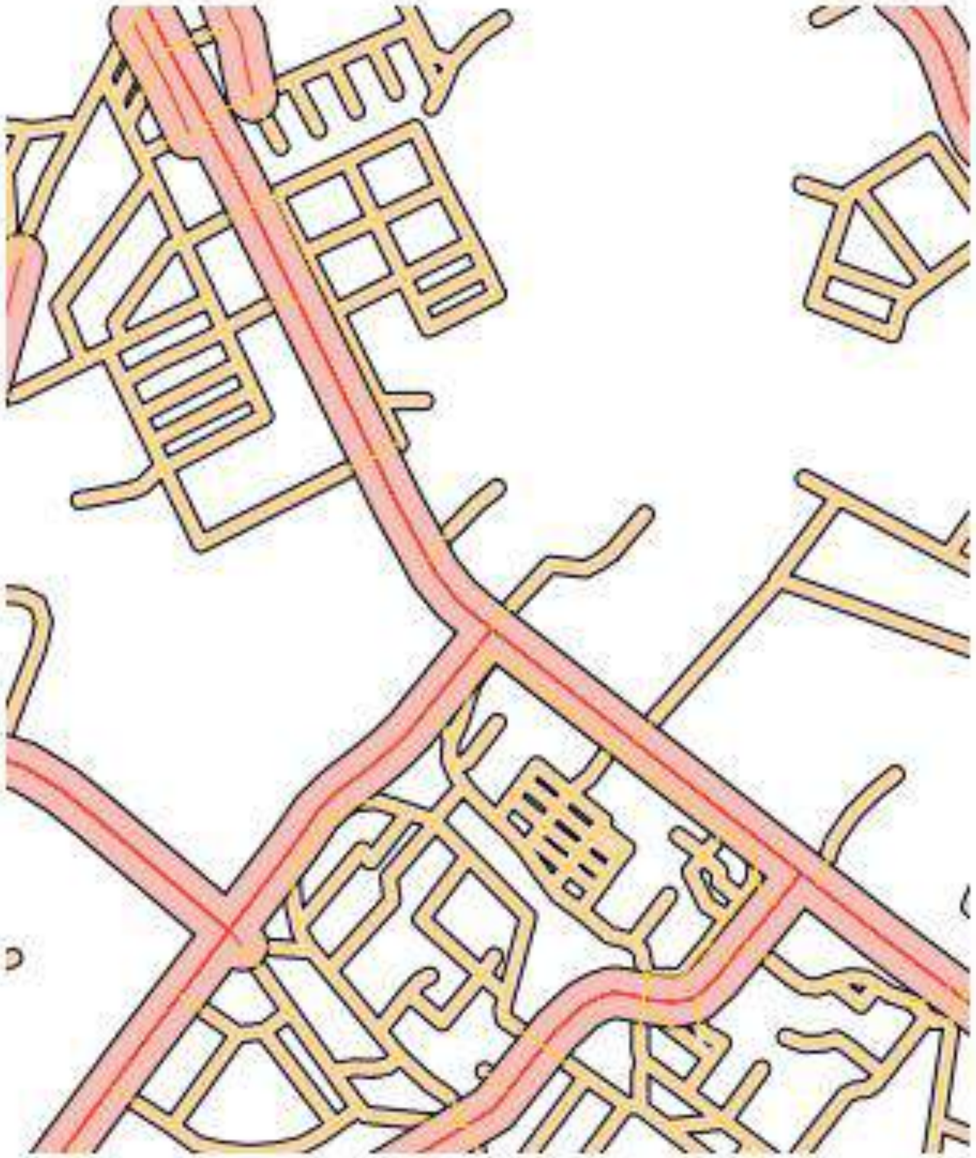
• It is done with the both vector and raster data.

• The principle of buffer zone generation is simple: we select one or more target locations, and then determine the area around them, within a certain distance.

• In vector-based buffer generation, the buffers themselves become polygon features, usually in a separate data layer, that can be used in further spatial analysis.

• Buffer generation on rasters is a fairly simple function. The target location or locations are always represented by a selection of the raster's cells, and geometric distance is defined, using cell resolution as the unit. The distance function applied is the Pythagorean distance between the cell centres.

1.1 Buffer zone generation



Buffer zone generation around main and minor roads. Different distances were applied: 25 metres for minor roads, 75 metres for main roads.

Source: https://planet.uwc.ac.za/nisl/GIS/Web_Page/page_3 6.htm

2.3 Neighborhood function

2. Flow computation

• It determine how a phenomenon spreads over the area, in principle in all directions, though with varying difficulty or resistance.

3. Raster based surface analysis

• It includes slope angle calculation, slope aspect calculation, visibility analysis, automatic catchment delineation etc.

Application examples of overlay and neighborhood function

Application examples of overlay function

1. Overlay operation can be used in site suitability analysis to find the best location to construct a new school, hospital, police station, industrial corridors etc.

2. It is also used in multicriteria decision analysis(MCDA) for overlaying multiple layers.

For example criteria of nuclear waste repository, site should be an area of suitable geology, site must not be easily accessible, site must be away from high population and site must be outside the area of conservation.

3. Some specific applications of overlay operation include habitat suitability assessment for pandas, site selection for a new business, or city expansion planning.

Application examples of neighborhood function

1. Proximity analysis is often used in urban based applications to consider areas of influence and ownership queries.
2. Proximity to roads and engineering infrastructure is typically important for development planning, tax calculations, and utility billing.
3. Raster based surface analysis are used in slope angle calculation (the calculation of the slope steepness, expressed as an angle in degrees or percentages, for any or all locations), slope aspect calculation, hillshading etc

2.4 Network analysis function

Network analysis function

- It is a kind of spatial analysis which is done on a network (a connected set of lines, representing some geographic phenomenon, typically of the railways, streams, roads, waterlines, pipelines, telecommunication lines etc).
- It analyses the movement of people, connectivity and accessibility of railway, roads, pipelines and telecommunications, flow of matter and energy and movement of goods and services.
- Network analysis can be performed on either raster or vector data layers, but they are more commonly done in the vector.

Some of the most important network analysis are:

i. Point-to-point analysis

ii. Finding Coverage

iii. Netwok allocation

iv. Origin-Destination – OD Cost Matrix

Network analysis function

Road connectivity in Delhi University north Campus

Source: http://epgp.inflibnet.ac.in/epgpdata/uploads/epgp_content/S000017GE/P001788/M027032/ET/15172071137_GIS_SEEMA_NetworkaNALYSIS.pdf

2.4 Network analysis function

i. Point-to-point analysis

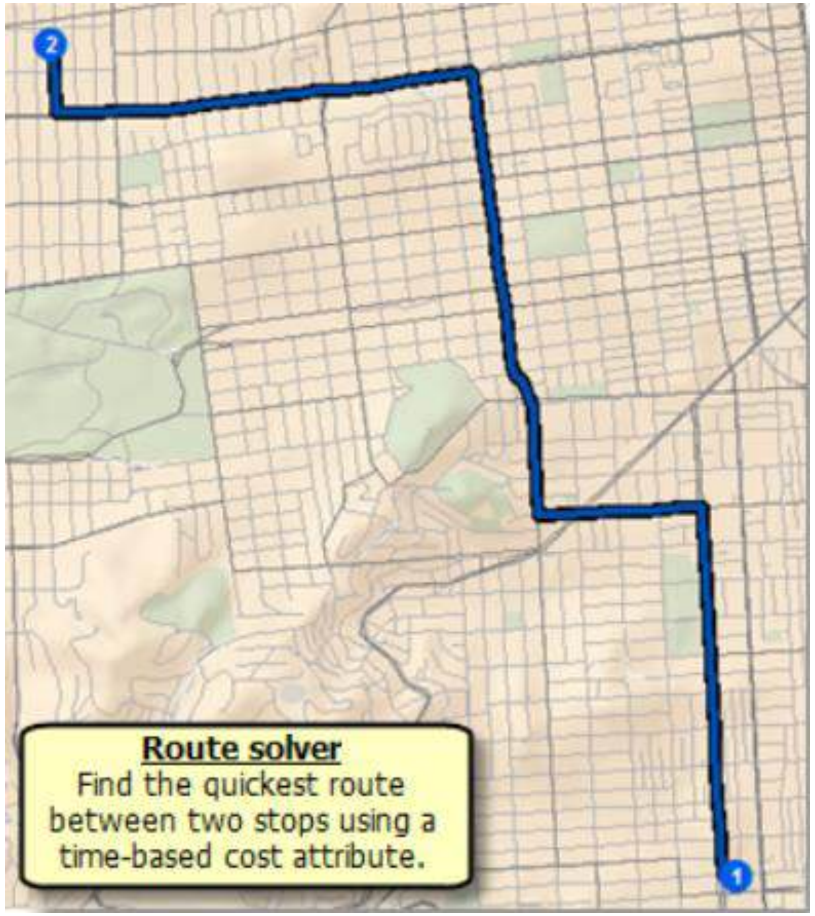
- It consists of a set of points to find the most optimal route based on specific criteria.

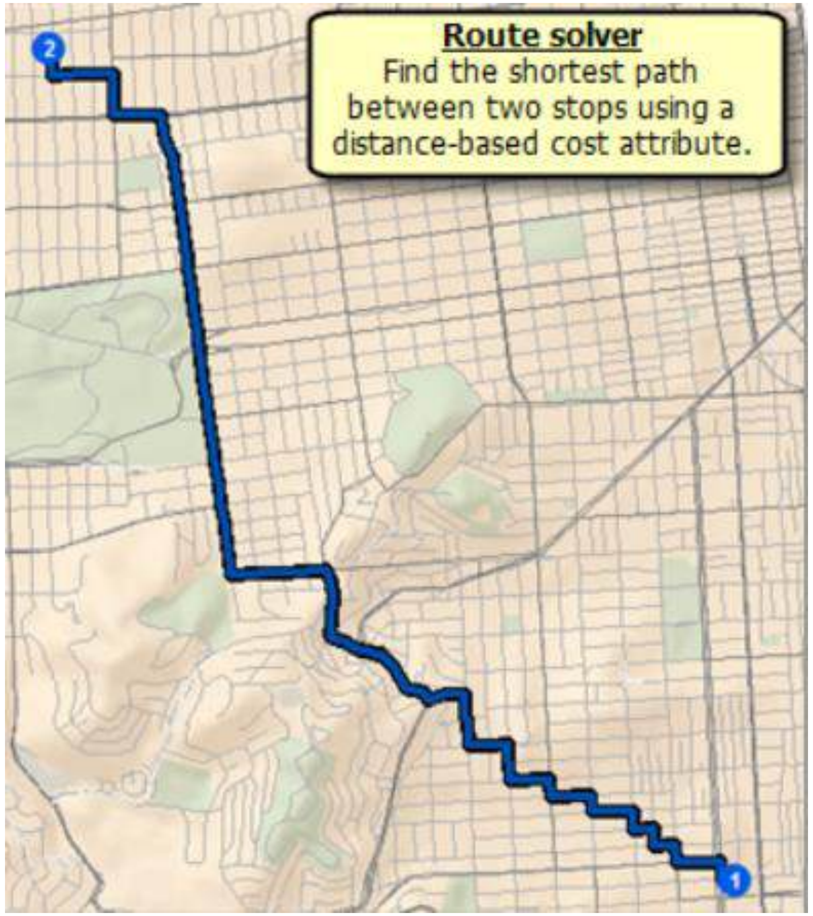
Find Nearest/Closest: Where is the closest destination? It finds the nearest destination based on a starting point with multiple potential destinations.

Shortest Distance: What's the fastest route? This analysis gathers all distances, as you travel out from one point to the other. Then, it finds the route with the least distance.

Fastest/Quickest route: Which route takes the least amount of time? This network analysis takes into account speed limits, road classification, and other costs to determine the least travel time.

i. Point-to-point analysis





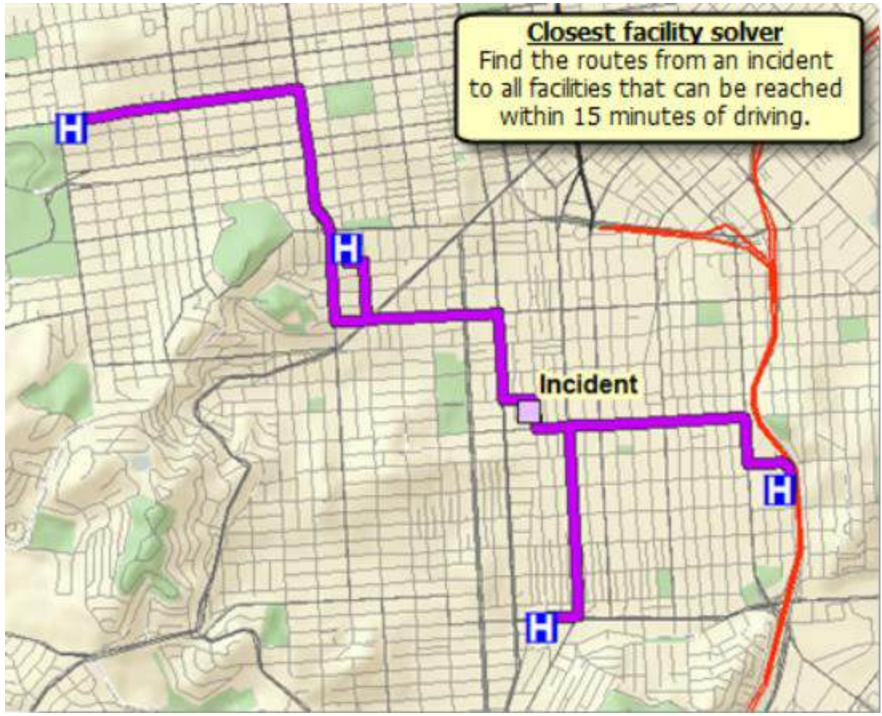


Fig 1: The quickest path is shown in blue and has a total length of 4.5 miles, which takes 8 minutes to traverse

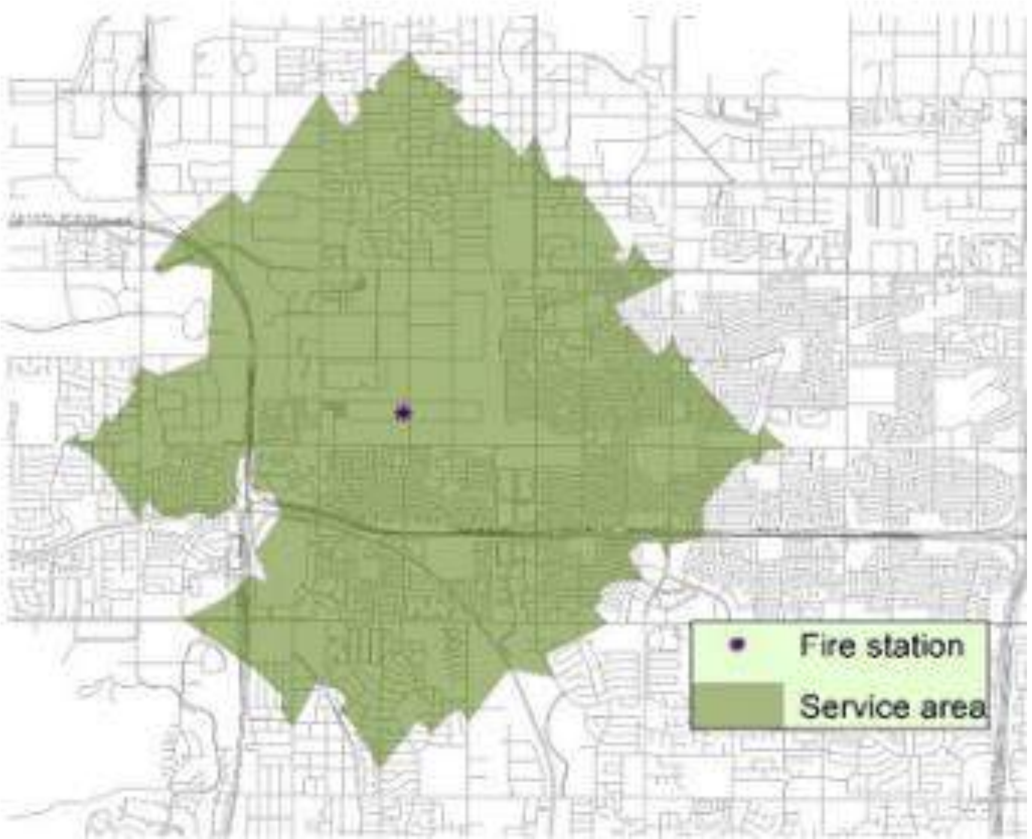
Fig 2: The shortest path is 4.4 miles, which takes 9 minutes to traverse.

Fig 3: The closest hospital(facility) within 15mins of accident spot.

Source: https://desktop.arcgis.com/en/arcmap/latest/extensions/network-analyst/types-of-network-analyses.htm

ii. Finding Coverage

- In this type of network analysis, drive-time areas correspond to the distance that can be reached within a specific amount of time.



Source: https://gistbok.ucgis.org/bok-topics/2022-quarter-02/location-and-service-area-problems

Fig: areas that can be reached within 5-minute drive distance from fire station.

iii. Network allocation

- In network allocation, we have a number of target locations that function as *resource centres*, and the problem is which part of the network to only assign to which *service centre*.

It includes:

i. ***The capacity with which a centre can produce the resources*** (whether they are medical operations, school pupil positions, or bottles of milk), and

ii. ***The consumption of the resources***, which may vary amongst lines or line segments. After all, some streets have more accidents, more children who live there, more industry in high demand of electricity.

2.4 Network analysis function

iii. Network allocation

E.g network within 2 km of school(represented by star) where the students live; demographic study data showed more student live there than the schools capacity and network was reduced to some portion and now only students in that network are admitted in school which is under schools capacity.

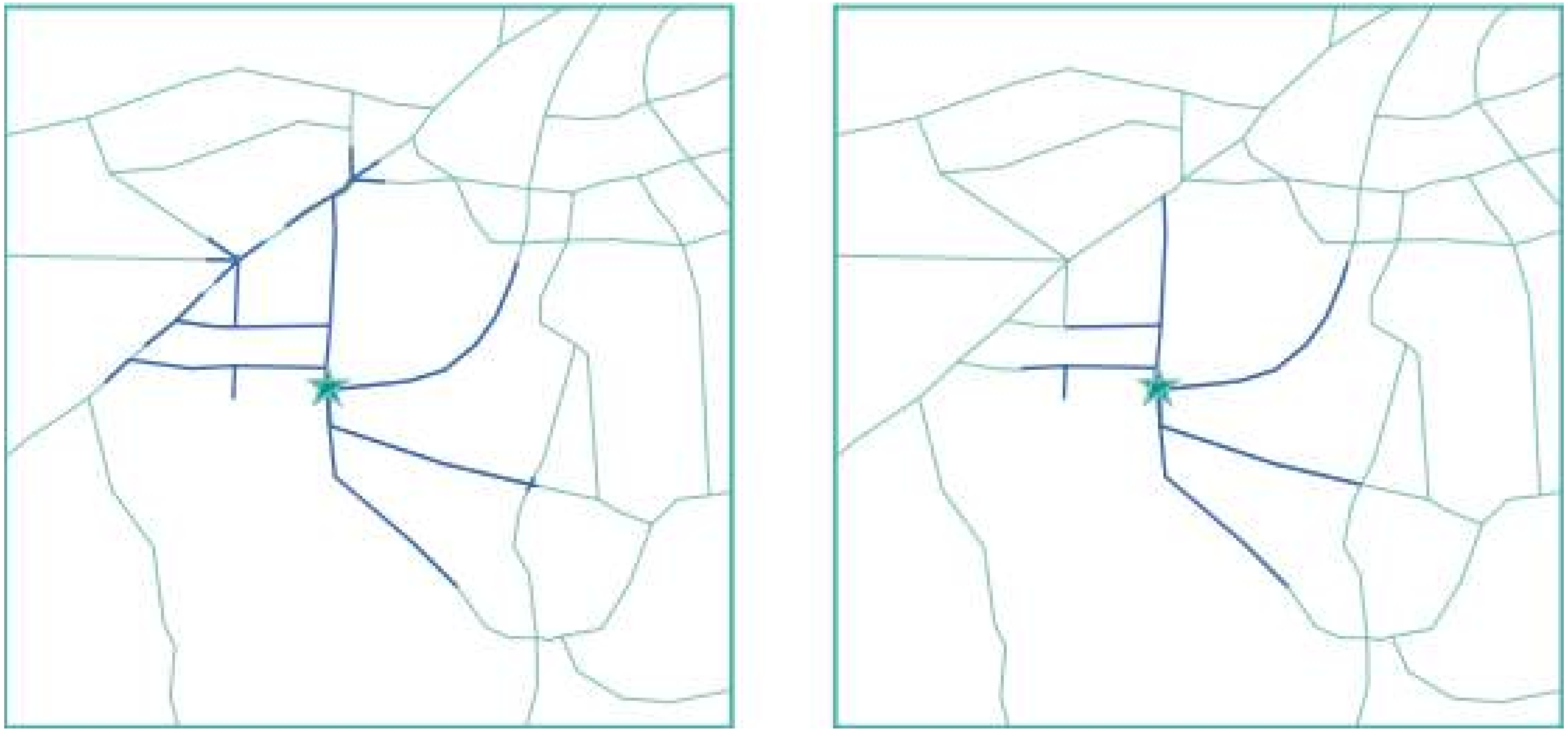
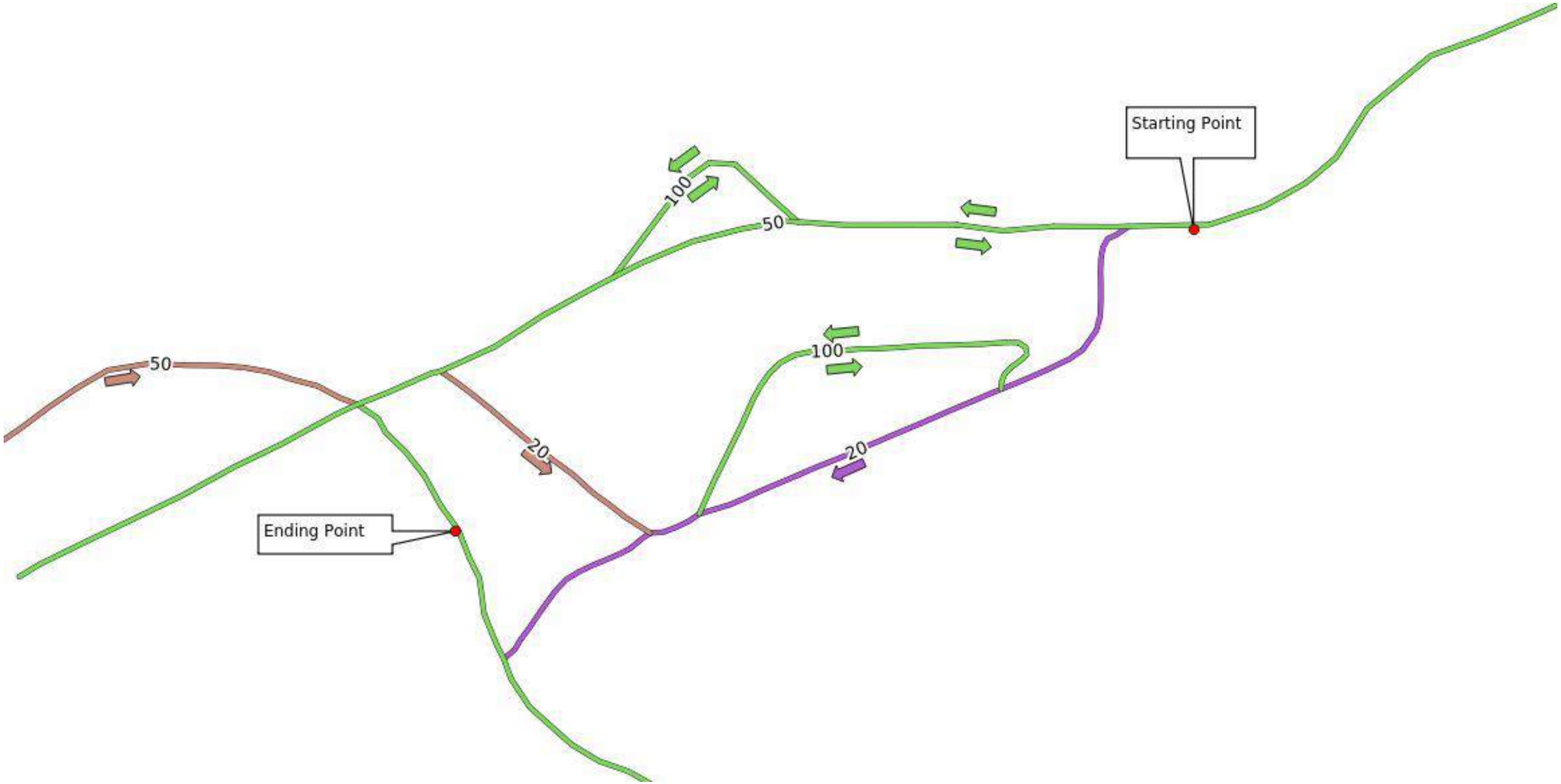


Fig: network allocation

Source: https://webapps.itc.utwente.nl/librarywww/papers_2009/general/principlesgis.pdf

Process of network analysis



Source: https://docs.qgis.org/3.4/en/_images/start_end_point.png

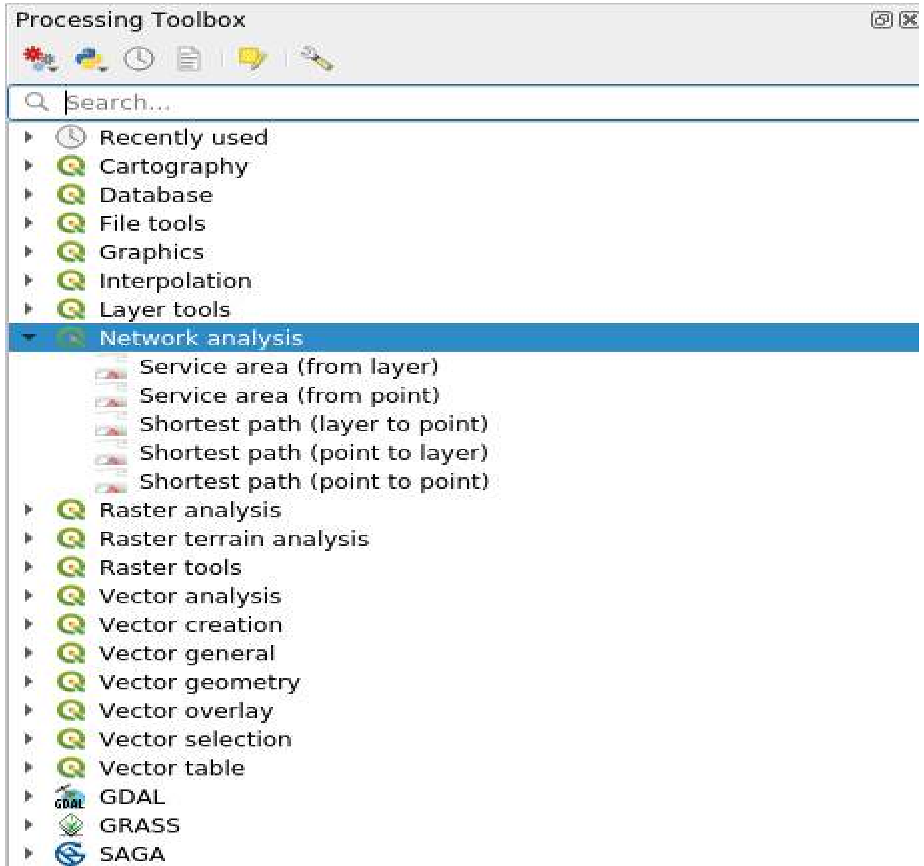
2.4 Network analysis function

Process of network analysis

The general procedure for network analysis in gis are:

1. Organizing Network Analysis toolbar:

- In any GIS system, like QGIS, first thing is you need to open the Processing Toolbar. There you need to search for ‘Network Analysis’ and select desire algorithm according to your data and purpose of analysis.

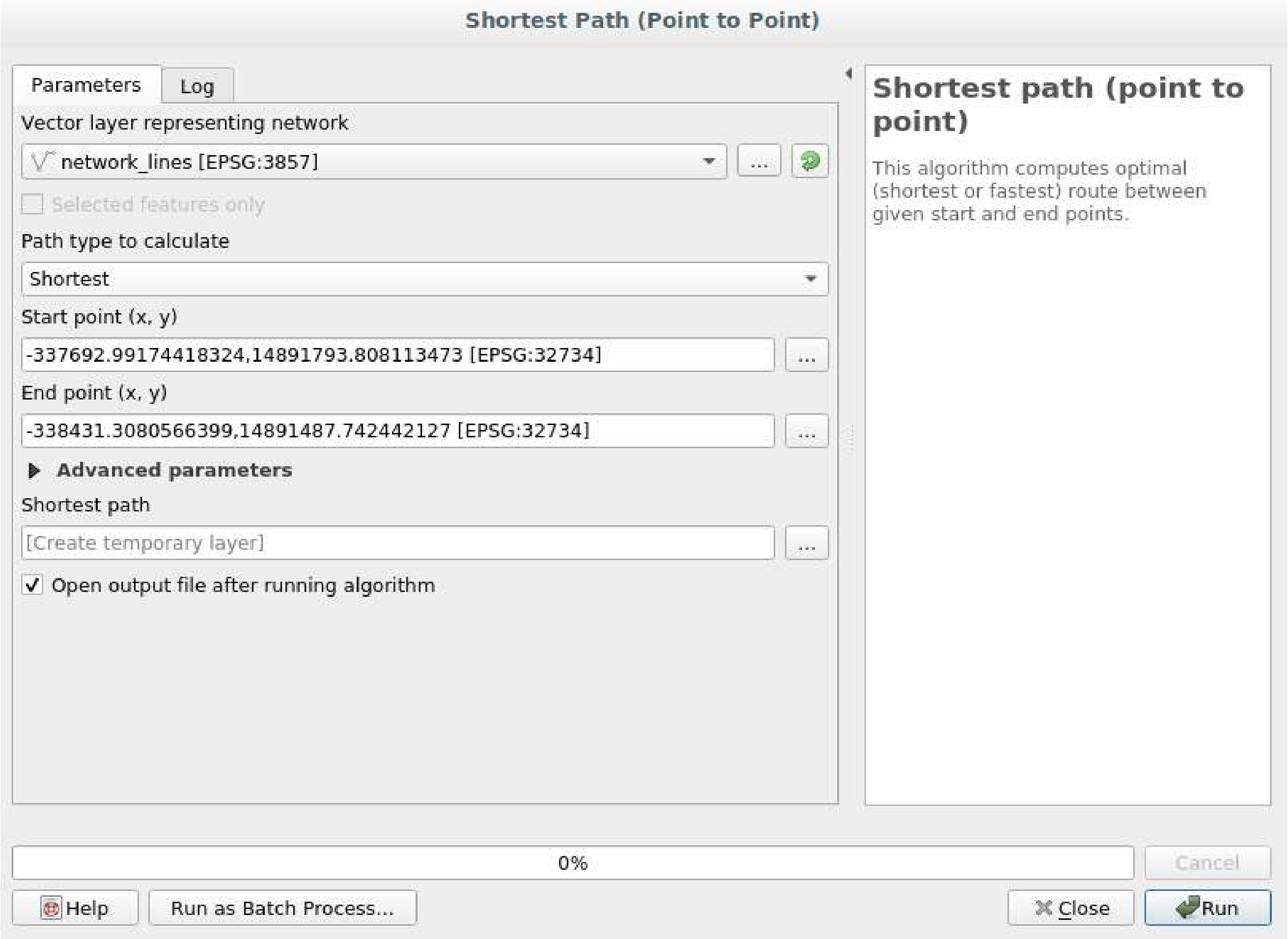


Source: https://docs.qgis.org/3.4/en/_images/start_end_point.png

Process of network analysis

2. Adding a Dataset to Your GIS Program

- Once you open the toolbar then you need to add the required dataset in parameters settings. Data may be network points or network lines.



Source: https://docs.qgis.org/3.4/en/_images/start_end_point.png

Process of network analysis

3. Selecting path type to calculate:

- In QGIS, we can choose shortest and fastest type to calculate between two points.

4. Performing the Analysis

- Once you have done all of above step you can simply click run which performs the analysis.

2.4 Network analysis function

Application of network analysis

1. Routing:

- Finding the shortest route from A to B through a road network is crucial for emergency services(like ambulance), business journeys, or simply planning routes for holiday makers touring a region.
- It is used in delivery services(food or any goods delivery) to reach clients in shortest way.

2. Resource allocation:

- The objective is to create service areas around a service centre and, if implemented successfully, allows an organization to optimise the distribution of the resources based on the capacity of each facility.

For example, enrollment of students from only within a particular network such that capacity of students can be handled by school.

Application of network analysis

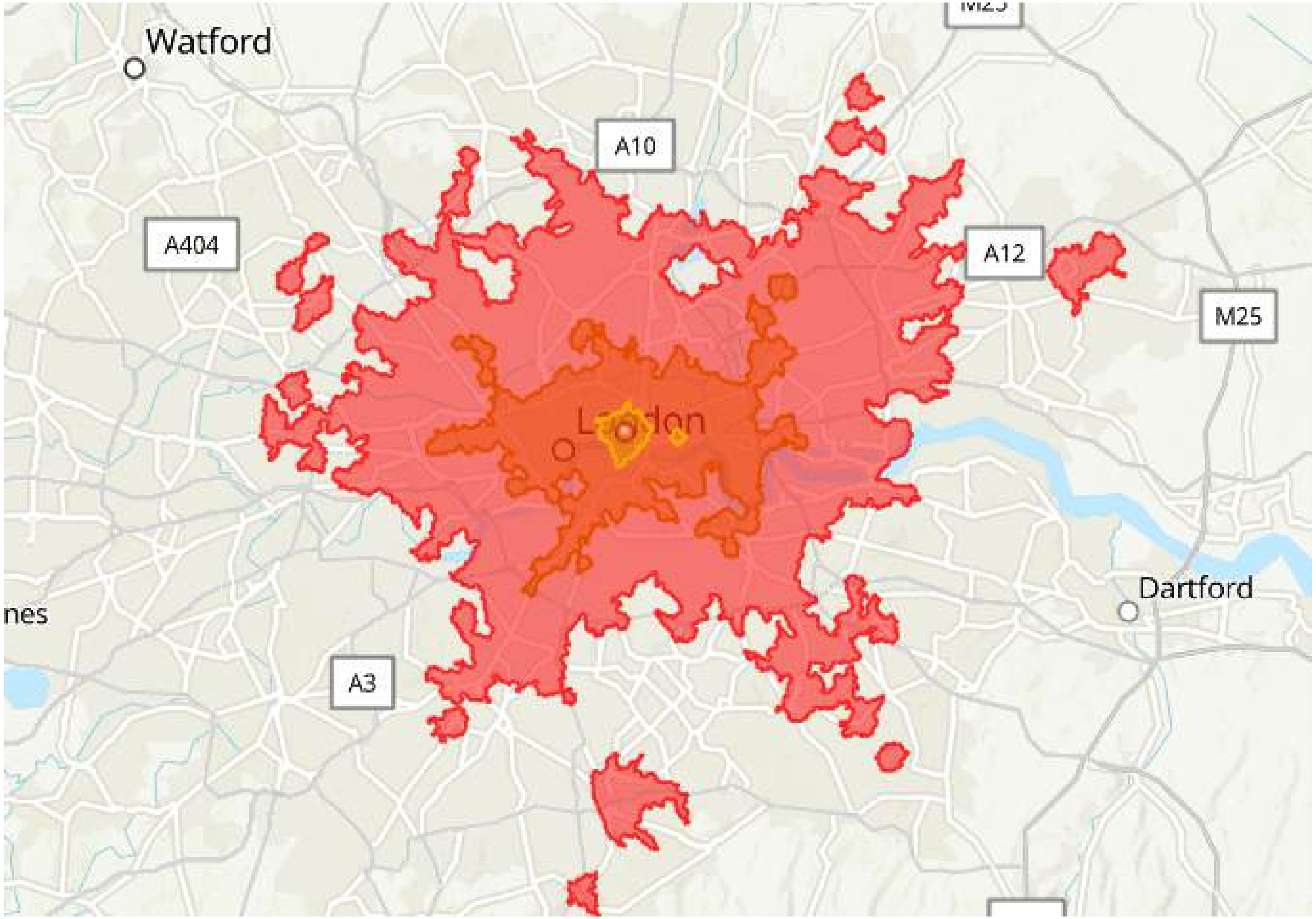
3. Isochrones (lines joining points of equal time):

- Isochrones show the area that is reachable from a starting point in a given amount of time, for a certain method of transport.
- Applications of this operation are for the establishment of cost surfaces or zones for transportation of goods and services.

For example, a furniture retailer may define zones A, B and C around the retail outlet. If you live in Zone A you can have goods delivered free of charge, if you live in Zone B, you will be charged a nominal fee, and if you live in Zone C you will be charged an reasonably high fee, because they don't want to have to send the delivery vans that far!

Application of network analysis

3. Isochrones (lines joining points of equal time):



Source: https://www.gislounge.com/how-to-create-public-transport-isochrones-in-arcgis-pro/

Fig: a public transport isochrone map for the London Bank station showing 15, 30 and 45 catchment areas:

Application of network analysis

4. The Travelling Salesman Problem (TSP):

- The travelling salesperson has to visit a number of customers then return to base. She/he needs to have details of the shortest tour possible for the sites to be visited. This is the obvious application.

3D representation of Earth surface(surface representation)

Physical surface of earth is not homogeneous and regular. Elevation varies from place to place. We need to model the elevation of earth surface for various applications.

Different methods of terrain representation:

1. Traditionally elevation of terrain can be represented or modelled in many ways like contouring, spot heights, layer tinting, hachuring, rock drawing, hybrid methods, shading etc.
2. In modern world, elevation of terrain is represented in the form of raster surface (Digital Elevation Model), TIN (Triangulated Irregular Network), DTM, DSM etc.

Methods for Generating contours

Contour can be generated through following methods:

1. Cartographic source: Topographic map can be scanned, georeferenced and digitalized to get contours and point heights. Point heights can also be used later to generated contours through various Interpolation techniques(IDW, Kriging etc.)
2. Stereo-pairs: It is traditional technique for generating contours where SPOT, Cartosat datasets etc. are used.
3. Lidar data
4. Photogrammetric methods (analytical, digital): Stereo plotters and latest drone data processing softwares can also generate contour by processing aerial images.
5. Using DEM in GIS softwares.

2.5 3D Analysis

Methods for generating 3D surface from 3D points

1. Interpolations methods

- Mathematical functions to represent the surface according to some specific methods based on the set of measured data points.

By the size of the area for interpolation, two approaches are identified. They are:

i. Area based

- In this approach the surface is constructed by using all the known points within this area and the height of any point within this area can be determined by using this constructed surface.

ii. Point based

2. Delaunay Triangulation

- Delaunay triangulation is a method for creating a mesh of non-overlapping triangles(TIN) from a set of points.

Methods for generating 3D surface from 3D points

3. Contour interpolation

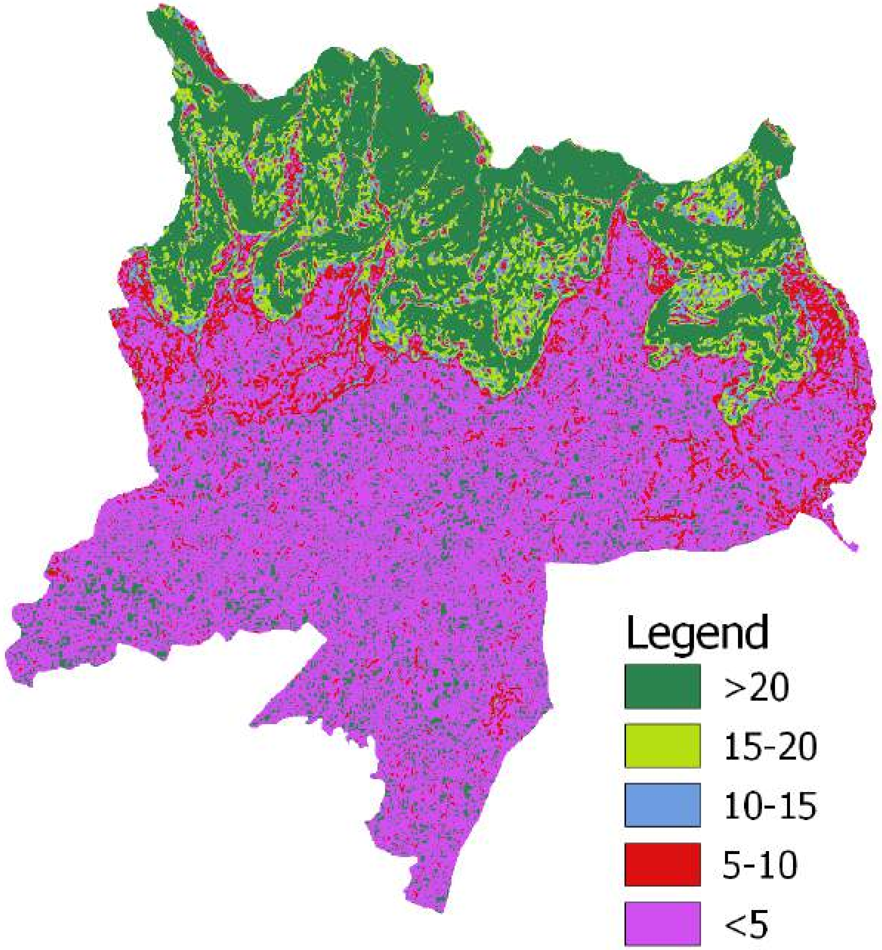
- This technique can be used to generate contour lines from given 3D points which can be later used to generate 3D surface.

4. Kriging

- It is a statistical methods used to interpolate the surface.

Slope

- It measures the rate of change of elevation of a surface. It is the first derivative of elevation.
- Slope may be expressed as percent slope or degree slope. The lower the slope value, the flatter the terrain; and the higher the slope value, the steeper the terrain.
- It is obtained from DEM in GIS.



Source: https://www.youtube.com/watch?v=wBPXsToa-9w&ab_channel=SpatialSight

Tutorial: https://youtu.be/wBPXsToa-9w?si=4sY6RaCQL4DfRj9C

Fig: showing slope of Butwal Sub-metropolitan city in degree

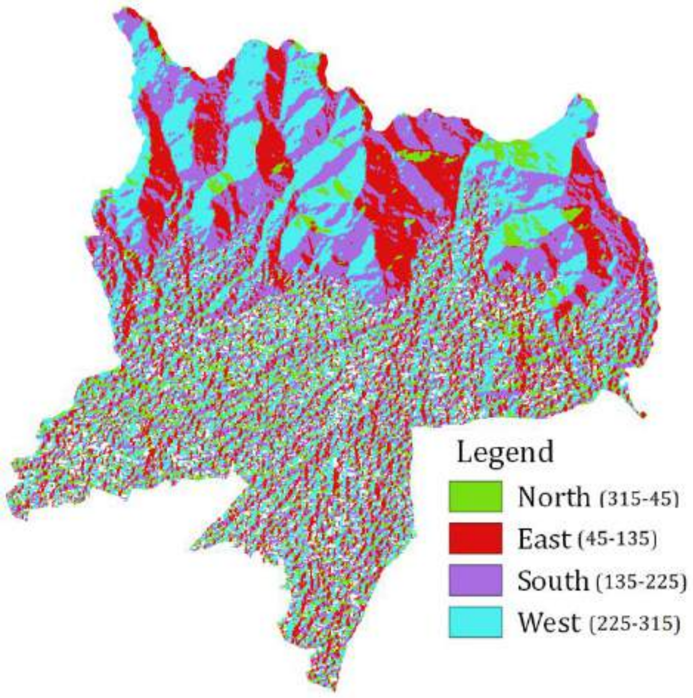
DEM Products

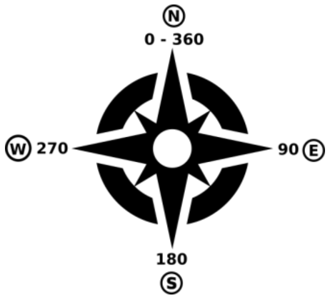
Application of slope

- Slope maps are used by architects, landscape planners, hydrologist in designing structures and predicting water flow.
- Slope is also an important input to analyses the most suitable routes across terrain for power lines, highways, and military vehicles.
- It is also used as a criteria in many Multicriteria Decision Analysis(MCDA) for suitability analysis like Landfill site selection, Landslide hazard zoning etc.

Aspect

- Aspect is the directional component of slope. An aspect of 0 means that the slope is North-facing, 90 East-facing, 180 South-facing, and 270 West-facing.
- Aspect measures are often grouped into the four principal directions or eight principal directions.
- When the terrain is flat, there is no slope. So this also means that there is no aspect.
- If no slope exists, then the cell value will be -1. These are the grey cells in the aspect map below. Where slope exists, aspect is measured clockwise starting north at 0°. It returns back as 360° north again.





Tutorial: https://youtu.be/NpYt-QW9VIM?si=UQX4yzC7dSdYuOmE

Source: https://mapadda.com/

Source: https://docs.qgis.org/3.22/en/docs/user_manual/processing_algs/qgis/rasterterrainanalysis.html?highlight=aspect

Application of aspect

i. Forestry:

- Suitability site analysis for planting trees in forest, habitat suitability analysis.

ii. Farmers grow seed according to the amount of solar radiation based on slope aspect data.

iii. Ski site selection:

For ski in snow, planners prefer north facing slope as it receives less radiation which is planned using slope aspect data.

Source: https://gisgeography.com/aspect-map/#:~:text=And%20there%20are%20some%20unique,slope%20direction%20to%20prevent%20avalanches.

DEM Products

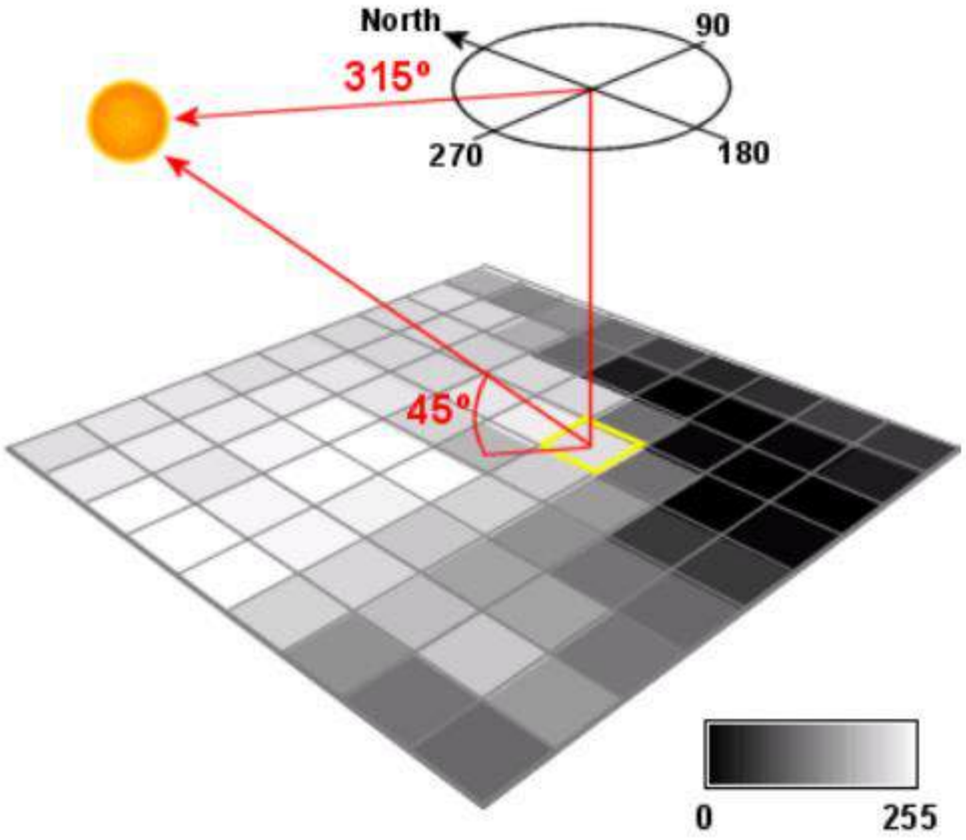
Application of aspect

iv. Specialized agriculture:

- Using aspect data farmers in Swiss Alps grow crops in south facing slopes as they are not affected by cold and dry winds which is essential for crop production.

Hillshade

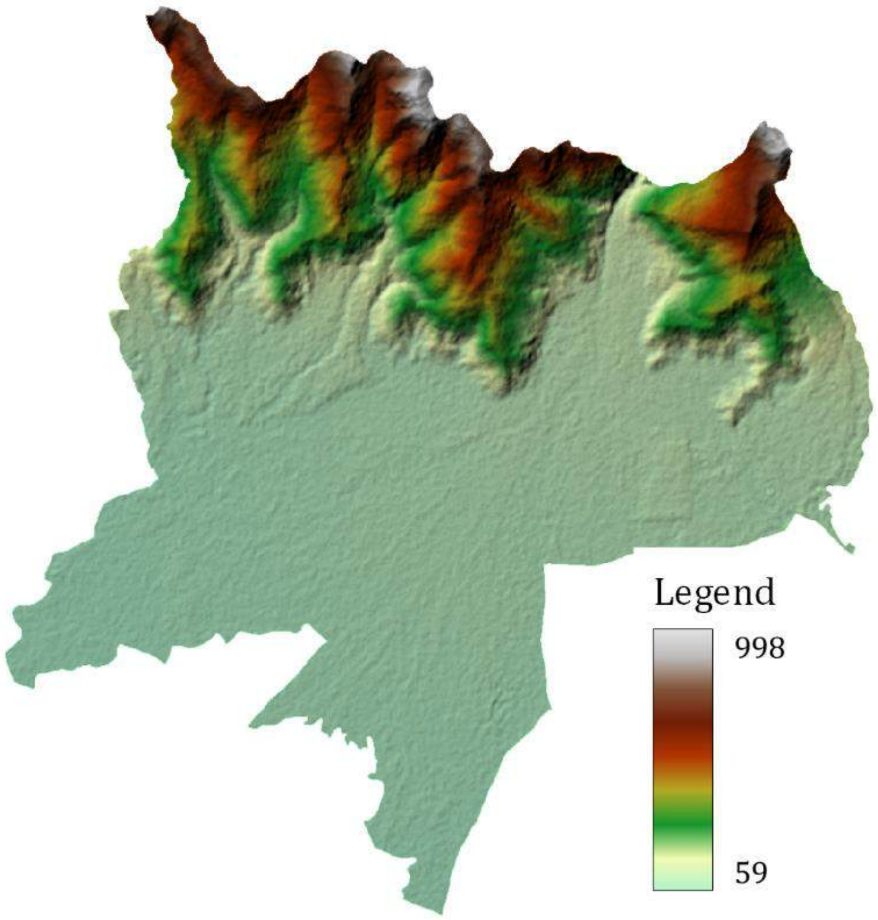
- It is a technique of hypothetical illumination of a surface by a light source on the basis of illumination value by the relative position of sun on that cell or slope and aspect on that cell.
- Hill-shading creates a 3D effect that provides a sense of visual relief for cartography and a relative measure of incident light for analysis. Hillshade is used for enhancing visualization of terrain and topography.



Source: https://gisgeography.com/aspect-map/

Application of Hillshade

i. Hillshades are often used to produce maps that are visually appealing. Used as a background, hillshades provide a relief over which you can draw raster data or vector data.



Tutorial: https://youtu.be/wBPXsToa-9w?si=9bGPEoGV9-5i4YG2

Source: https://keshavrajbhusal.com.np/

Fig: cartographic hillshade combined with elevation data showing terrain of Butwal

Surface Area

- Surface area is the total area that can be measured on the entire surface. This can only be measured if the object is a three dimensional object.
- Surface area is a fundamental parameter derived from terrain analysis in GIS for modelling the real world. Surface area can be calculated for DEM and TIN directly.

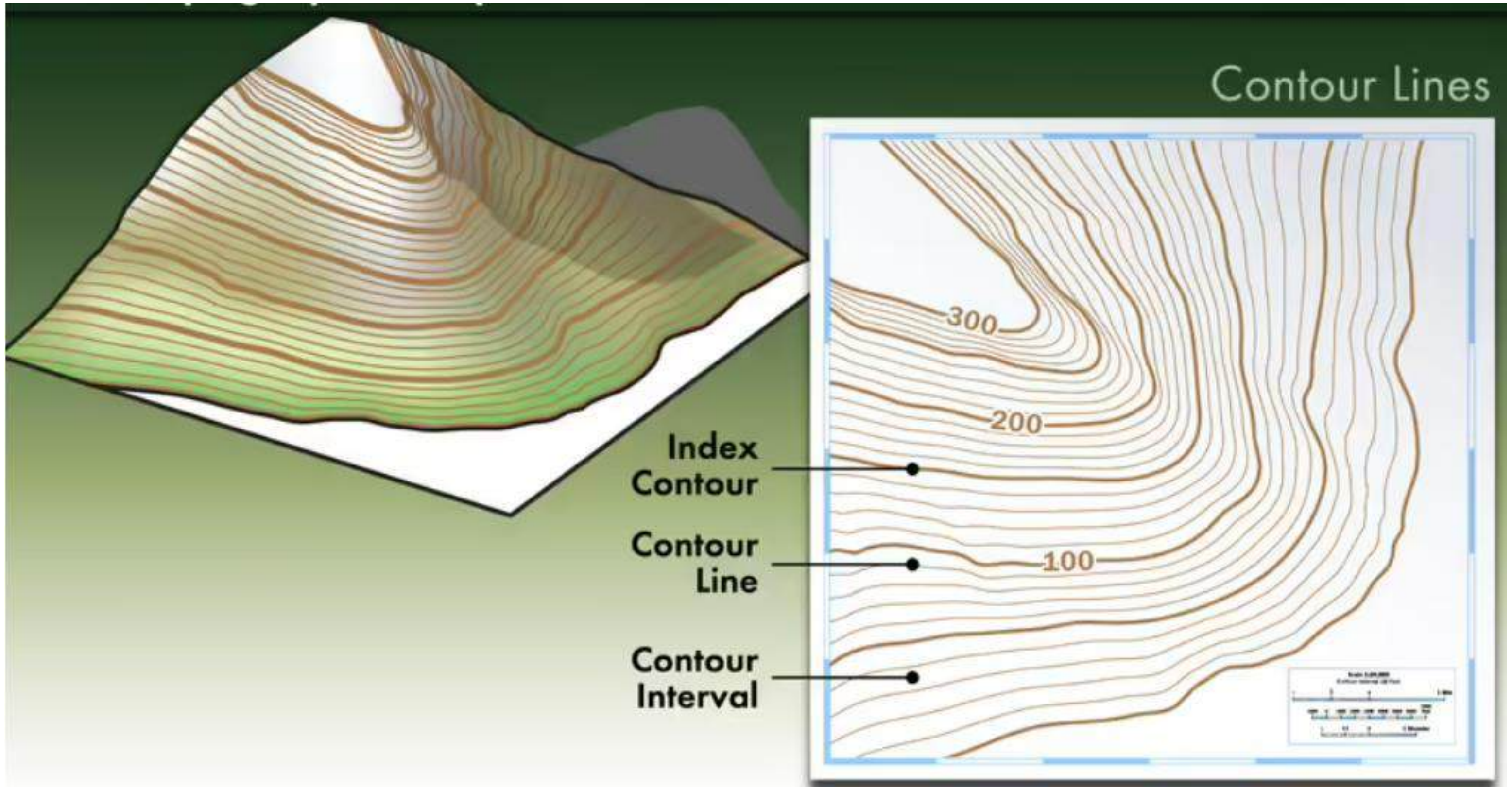
Volume

- The amount of space measured in cubic units that an object or substance occupies in measured amount that a container or other object can hold in cubic capacity.

Tutorial: Cut Fill volume analysis: https://youtu.be/JkOmhWJST8o

Contouring:-

- Contour is an imaginary line of constant elevation on the ground surface. The corresponding line on a map is called a contour line. The distribution of the line shows how elevation change across a surface.
- Contouring is the standard method of representation of terrain.
- Contour in GIS can be generated using raster surface (DEM) and vector terrain representing surface TIN.



Source: https://adventure.howstuffworks.com/outdoor-activities/hiking/how-to-read-a-topographic-map2.htm

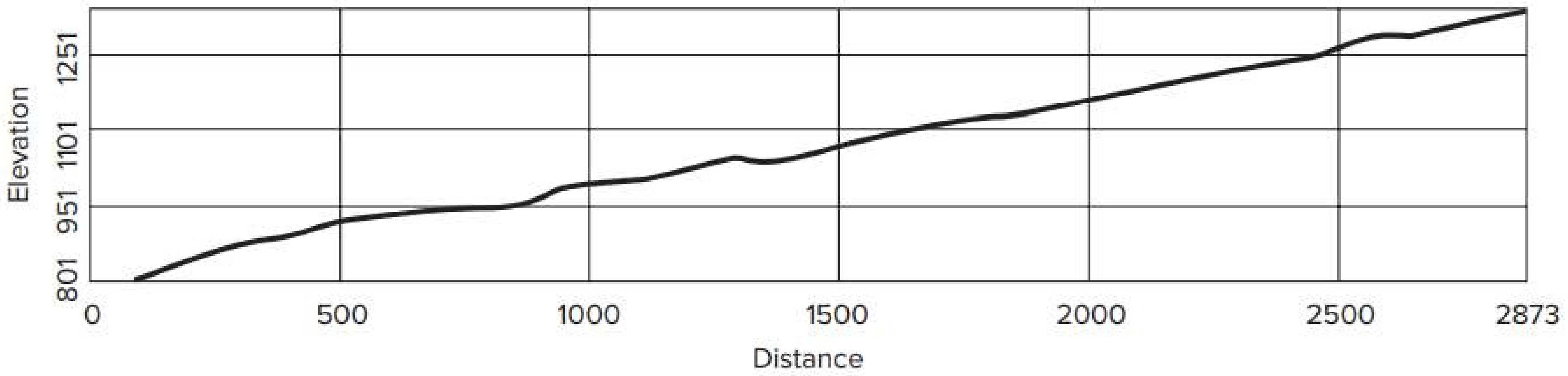
2.5 3D Analysis

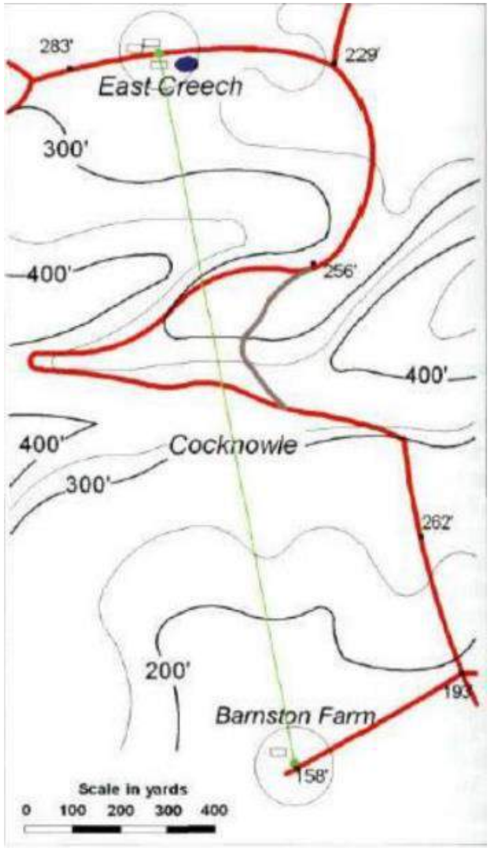
Verical profile generation

- A vertical profiles shows changes in elevation along a line such as hiking trail, road or a stream.

The manual method usually involves the following steps:

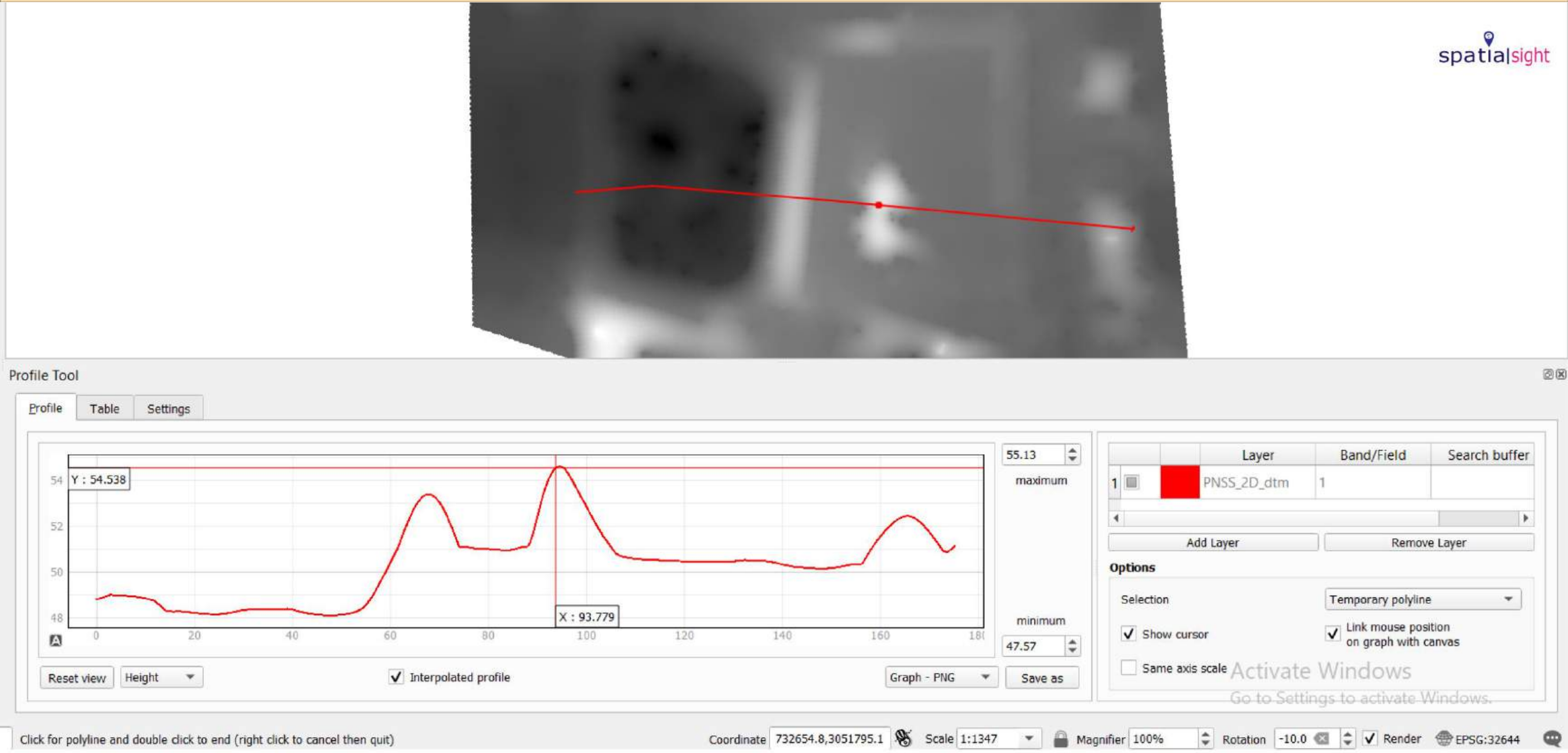
i. Draw a profile line on a contour map.
ii. Mark each intersection between a contour and the profile line and record its elevation.
iii. Raise each intersection point to a height proportional to its elevation.
iv. Plot the vertical profile by connecting the elevated points.





Source: Kang-Tsung (2019). Introduction to Geographic Information Systems. 5th Edition. McGraw Hill. ISBN: 0070658986

3D Analysis: Verical profile generation



Source: https://www.facebook.com/photo/?fbid=125886130371043&set=a.111287168497606

2.5 3D Analysis

Digital Elevation Model (DEM)

- It is the most versatile and useful representation of terrain in GIS. This is a raster representation, in which each grid cell records the elevation of the Earth's surface, and reflects a view of terrain as a field of elevation values.
- A DEM refers to bare-earth elevations, disregarding(exclusion) of the built (e.g., buildings) or natural (e.g., vegetation cover) features.
- DEM are used often in GIS and are most common basis for producing relief maps.
- The data sources for DEMs include topographic maps, aerial photographs, satellite images (from both passive and active systems), light detection and ranging (LiDAR) data, and data collected by unmanned aerial systems-based photogrammetry and terrestrial laser scanning.

210	190	170	155	140	135
204	183	165	145	125	120
200	175	160	122	110	100
208	187	165	150	126	120

Fig: DEM showing elevation value in each grid cell.

Source: https://www.researchgate.net/figure/Small-Digital-Elevation-Model-DEM-with-elevation-values-overlaid-on-each-cell-Use-a_fig1_232669735

Applications of Digital Elevation Model (DEM)

- Creation of relief map.

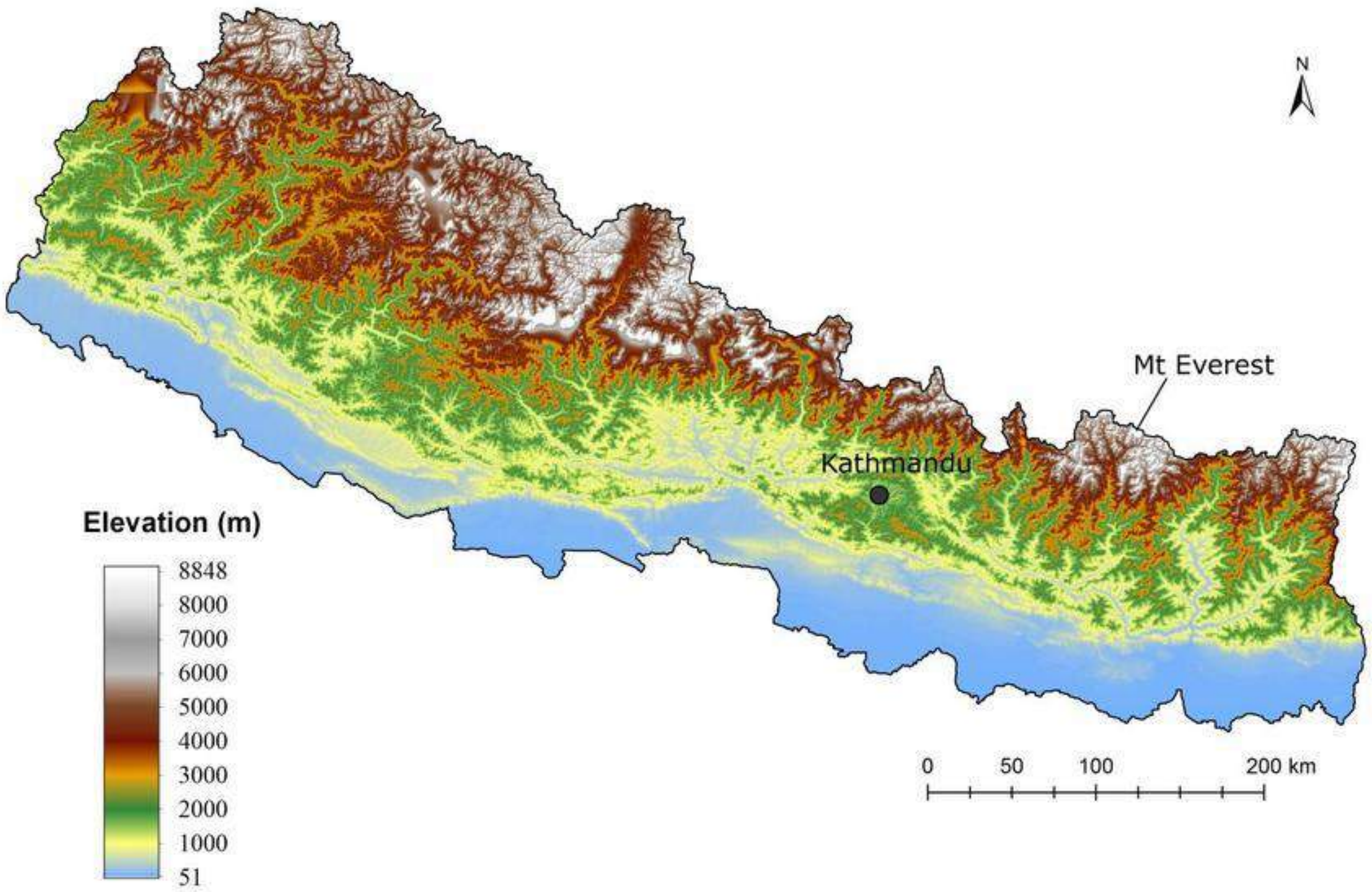


Fig: showing relief map of Nepal.

Source: https://www.researchgate.net/figure/Elevation-map-of-Nepal_fig1_319183873

2.5 3D Analysis

Applications of Digital Elevation Model (DEM)

- Providing rendering of 3D visualizations by overlay of dem with satellite imagery.
- Rectification of aerial photographs.
- TERRAIN STABILITY: Areas prone to avalanches are high slope areas with sparse vegetation. This is useful when planning a highway or residential subdivision.
- HYDROLOGIC MODELING: Hydrologists use DEMs to delineate watersheds, and calculate flow accumulation and flow direction in HEC-RAS and other GIS softwares.
- It is also used in surface analysis (creating Slope, Aspect, Hillshade, Viewshade etc.)
- Construction: Highway, railway, irrigation, (longitudinal and cross-section) survey.

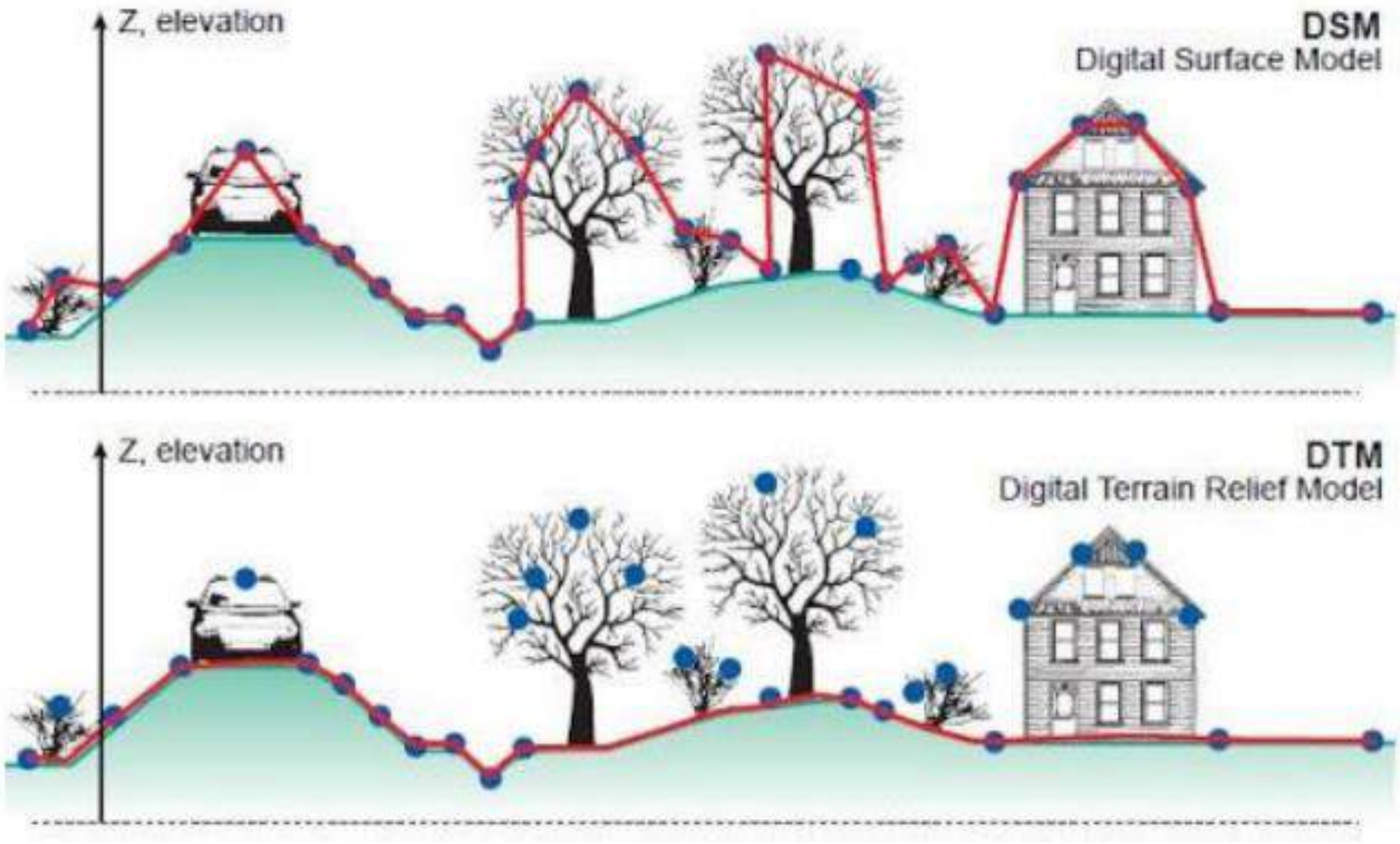
Digital Surface Model (DSM)

- A digital surface model (DSM) represents elevation, which includes the built and natural features on the Earth's surface.
- DSMs can be prepared from LiDAR data, because a LiDAR system can detect multiple return signals from objects of different elevation values.

Applications of Digital Elevation Model (DSM)

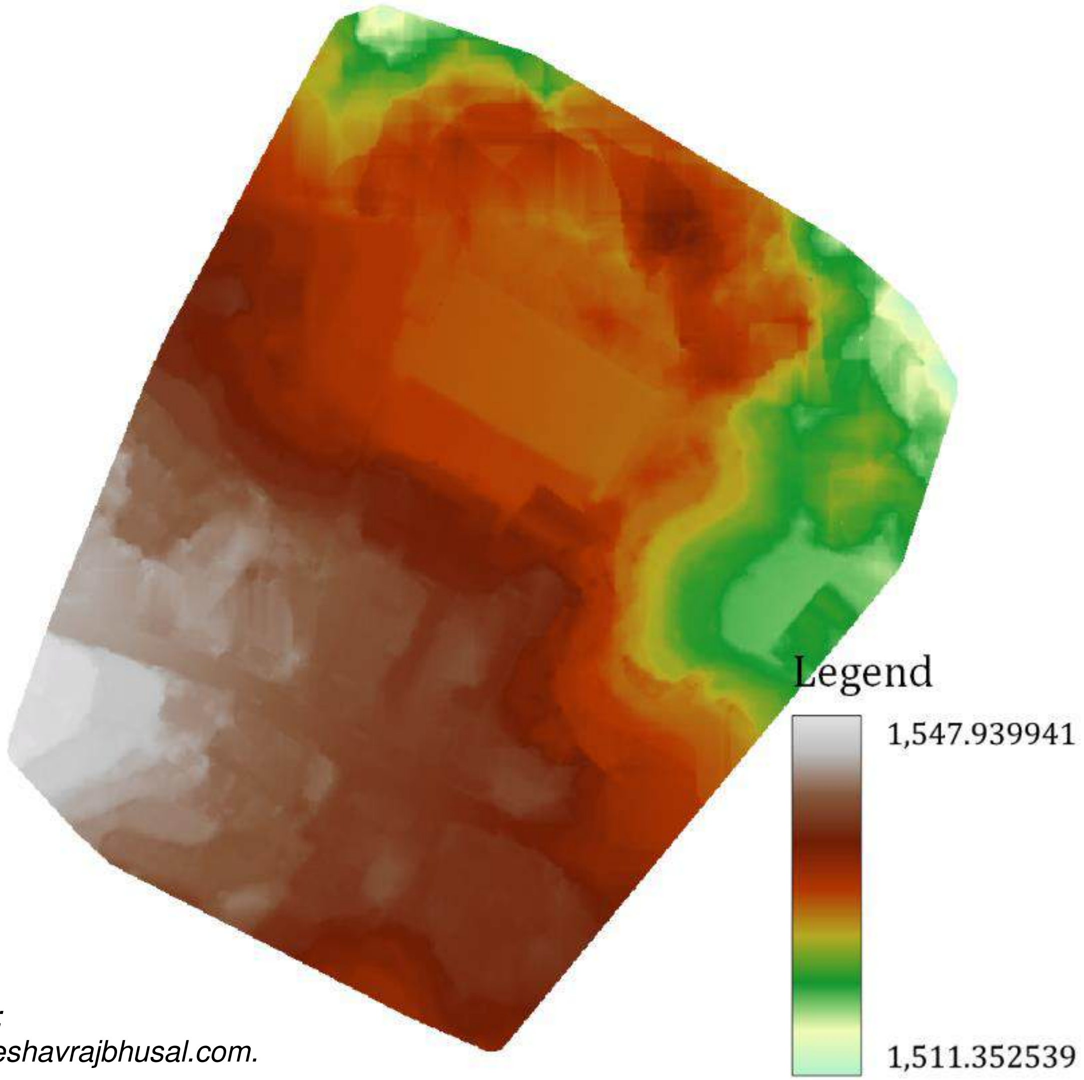
- RUNWAY APPROACH ZONE ENCROACHMENT: In aviation, DSMs can determine runway obstructions in the approach zone.
- VIEW OBSTRUCTION: Urban planners use DSM to check how a proposed building would affect the viewshed of residents and businesses.
- It is also used in flood hazard management because knowing DSM data of flood zone helps in effective planning.

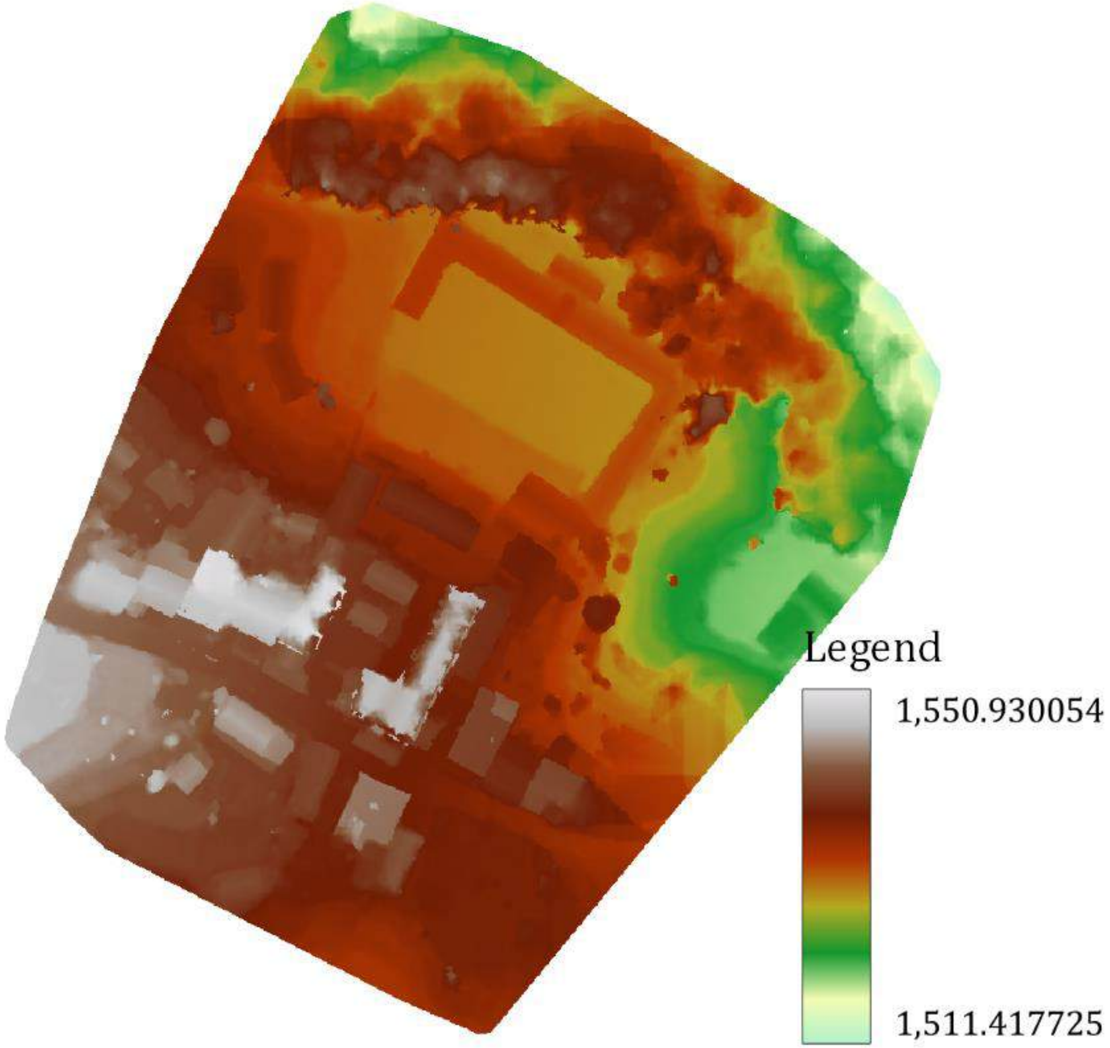
Digital Surface Model (DSM) vs DEM



2.5 3D Analysis

DEM vs DSM



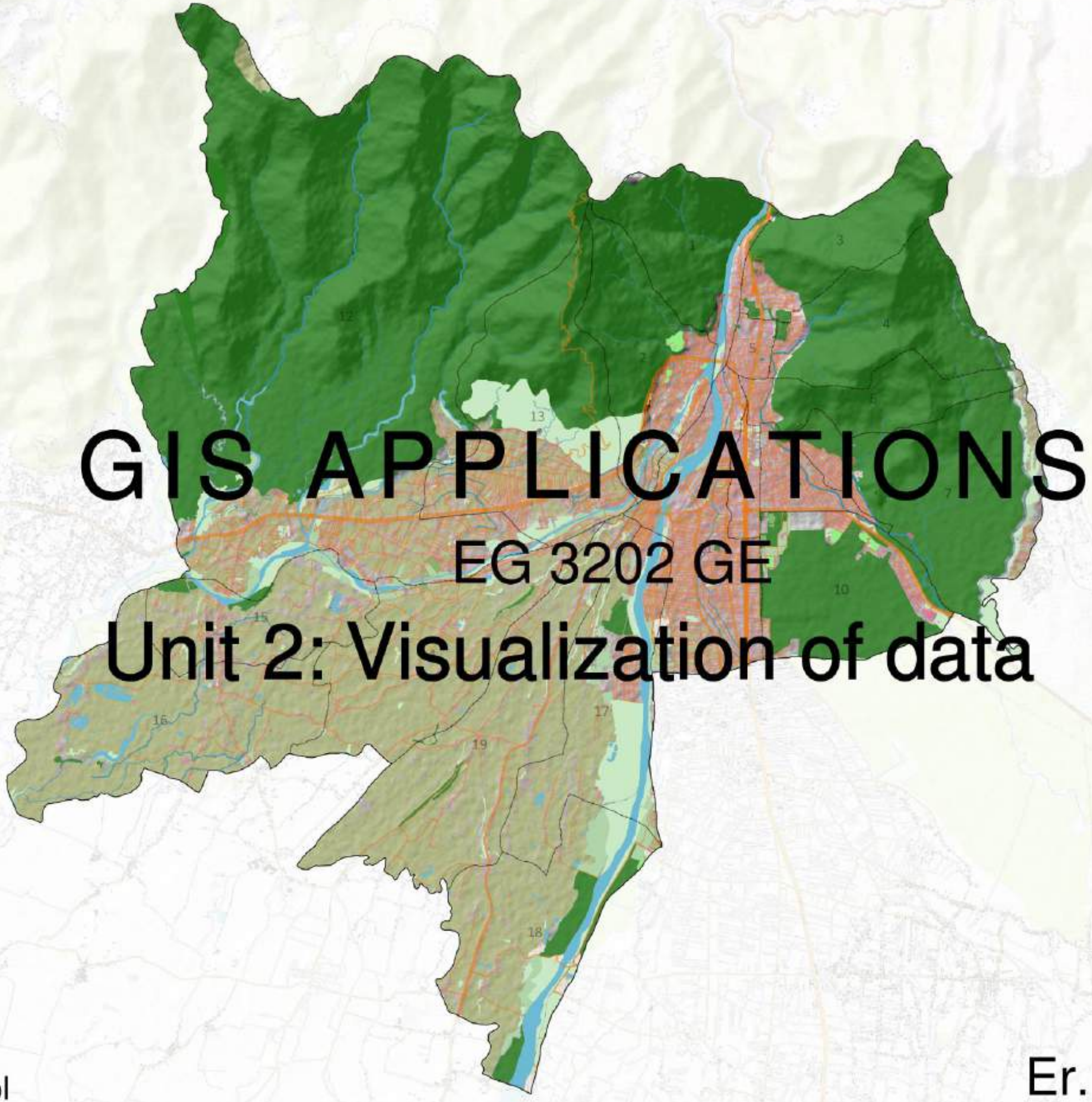


Source: http://keshavrajbhusal.com.np/

Fig: (1)DEM (2)DSM

References

- https://webapps.itc.utwente.nl/librarywww/papers_2009/general/principlesgis.pdf
- http://epgp.inflibnet.ac.in/epgpdata/uploads/epgp_content/S000017GE/P001788/M027032/ET/15172071137_GIS_SEEMA_NetworkaNALYSIS.pdf
- https://gistbok.ucgis.org/bok-topics/2022-quarter-02/location-and-service-area-problems
- Kang-Tsung (2019). Introduction to Geographic Information Systems. 5th Edition. McGraw Hill. ISBN: 0070658986
- https://gisgeography.com/aspect-map/
- https://www.youtube.com/@mapadda
- DEM vs DTM
- https://www.jouav.com/blog/elevation-models-dem-dsm-dtm.html



GIS APPLICATIONS

EG 3202 GE

Unit 2: Visualization of data

Public Namuna Secondary School

Lecture by

Er. Keshav Raj Bhusal

Table of contents

Define visualization

- Visualization is simply the representation of an object or set of information as a chart or other image. In GIS, cartographic visualization is the translation or conversion of spatial data from a database into graphics(mainly maplike products).
- During the visualization process, cartographic methods and techniques are applied.
- Visualization process can be simple or complex, while the production time can be short or long depending on visualization purpose and environment (personal computer, WWW etc.).

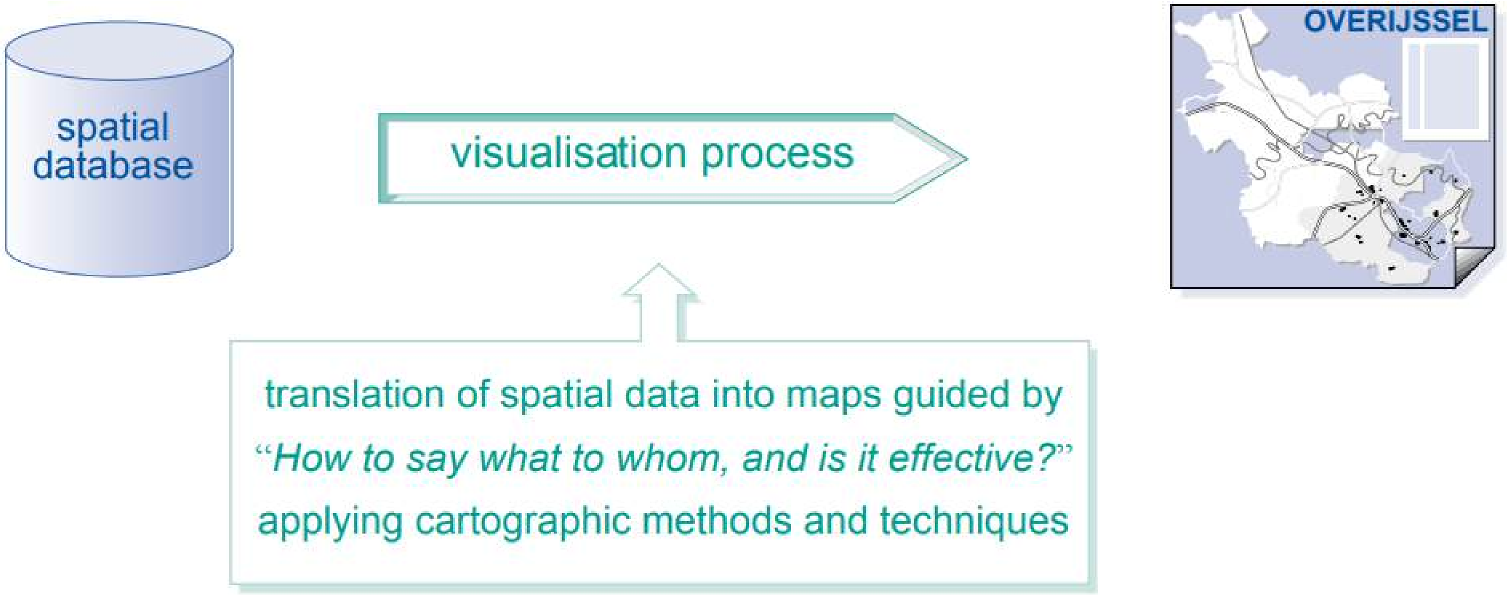


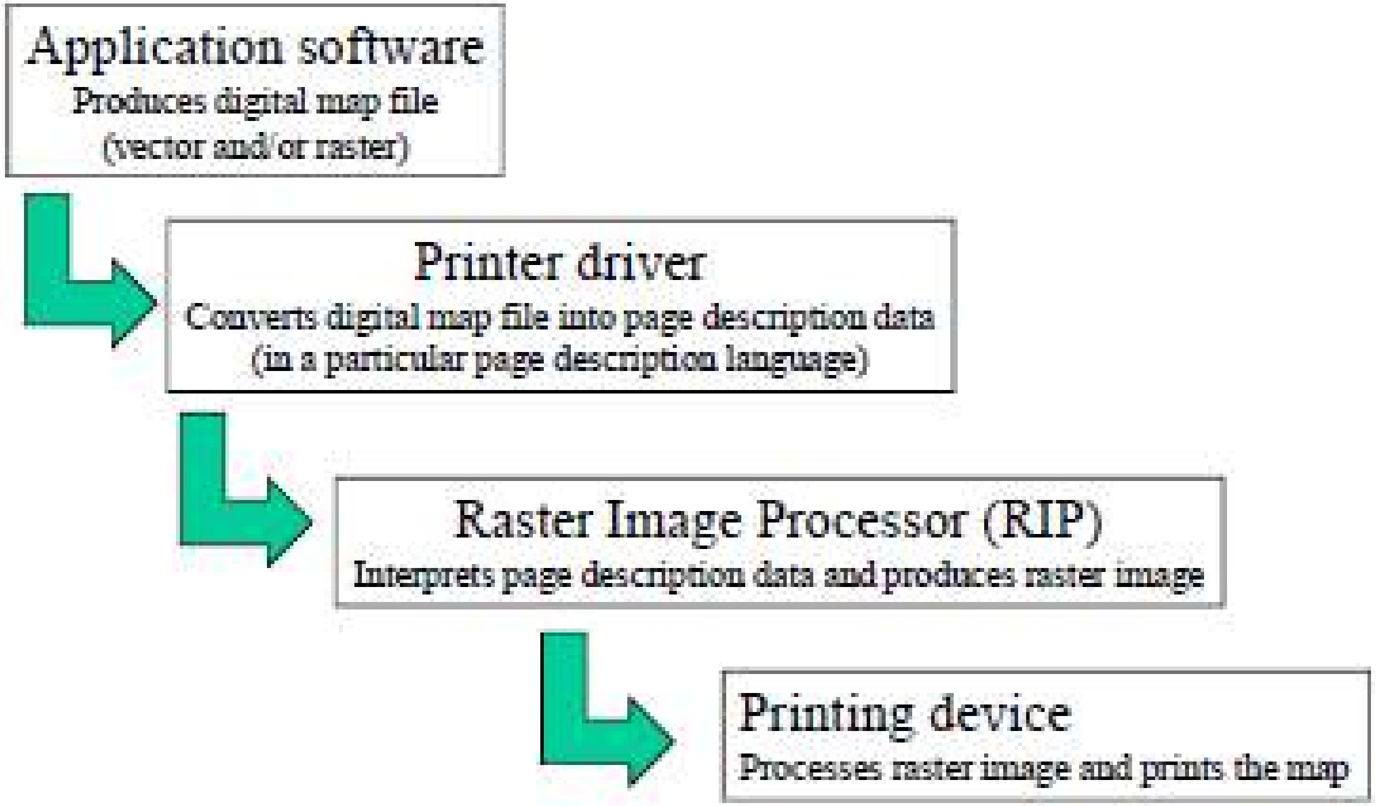
Fig: showing visualization process

Source: https://webapps.itc.utwente.nl/librarywww/papers_2009/general/principlesgis.pdf

Various visualization technique (Hard and softcopy)[Dissemination Techniques of Maps]

Hardcopy

- Maps are disseminated in physical paper by photocopying or offset printing using printers.



Source: https://webapps.itc.utwente.nl/librarywww/papers_2009/general/principlesgis.pdf

Fig: digital map printing process process

Various visualization technique [Dissemination Techniques of Maps]

Hardcopy

Advantages

- *No device needed*: You do not need a computer, mobile or other electronic device to access it.
- *Offline Access:* One of the key advantages of paper maps is that they may be consulted even when there is no signal or internet connection.
- *Are easier to understand due to the limited features*: Paper maps are easier to understand due to the limited features shown on the maps at a time.
- *Used for data comparison*: Paper maps can be used to make data comparison for historical purposes because they are static.

Various visualization technique [Dissemination Techniques of Maps]

Hardcopy

Disadvantages

- *Cannot be updated easily*: Paper maps are printed on paper and can therefore not be updated on the same paper. This means that you cannot make any changes on the map based on the changes in landforms.
- *Paper maps are limited to specific area based on the scale of the map*: Paper maps are representations that are limited on a specific area of the land and not the entire geographical location.
- *Storage requires physical space:* Storage of paper maps requires that you keep them in a dedicated place in your house or in your suitcase since it can be physically handled.
- *Extra Care is Required:* If the map becomes wet, it will become unreadable and useless.

Softcopy (Digital Map)

- It includes map displays on monitor screen and WWW (WebMap)
- WWW is the major medium for the dissemination of maps to users. E.g Google map, Google earth etc.

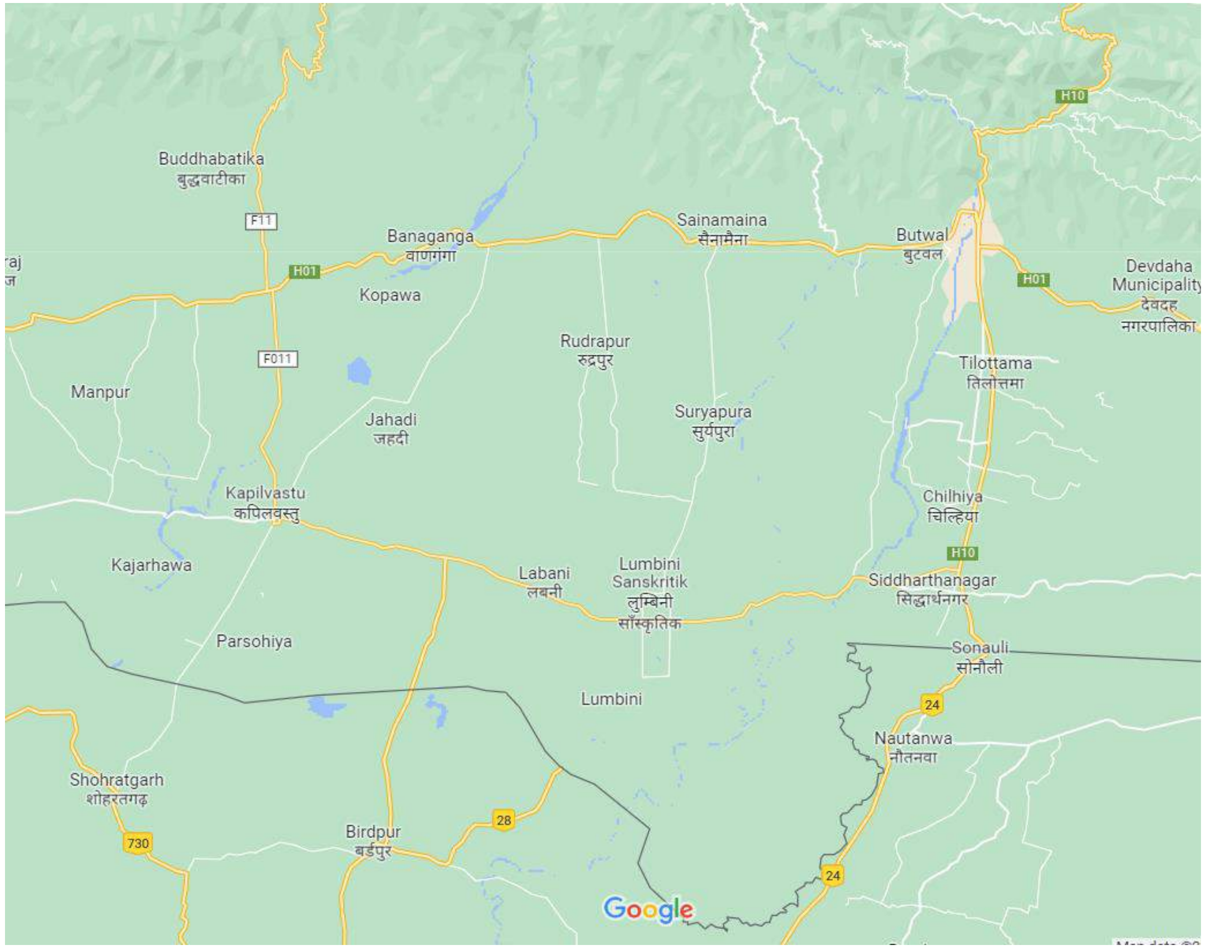


Fig: softcopy map

Source: https://www.google.com.np/maps/@27.5874118,83.2719562,10.7z?hl=en

Various visualization technique [Dissemination Techniques of Maps]

Softcopy (Digital Map)

Advantages

- *Can be easily updated*: Digital map is mostly real time representations of an area and can therefore be updated easily because the changes will be updated automatically.
- *Digital maps can represent all features at the same time*: Digital maps can be used to represent all features of a given area at the same time. They allow filtering for specific features but can also show the entire area.
- *Digital maps are not limited to any area*: Digital maps are not limited to show only a specific area based on scale. They can be widened to show the entire area.
- Panning, zooming and other interactive features.

Various visualization technique [Dissemination Techniques of Maps]

Softcopy (Digital Map)

Disadvantages

- *Device needed*: You need a computer, mobile or other electronic device to access it.
- *No offline Access:* One of the key disadvantage of softcopy maps is that they require signal or internet connection to access them. People in remote areas without internet access can’t use them.

Map Data Types

Fundamentally, maps display only two types of data: qualitative and quantitative.

Qualitative data type

- Qualitative data is nominal, meaning types, but no specific order to them.
- It shows what is found in the different locations of a map, and not how much is found.
- Map with qualitative data in geographic units should get equal attention, and none should stand out above the others.
- The categorical differences in qualitative data can be shown with symbols that vary by color hue (e.g., red, green, blue) and shape (e.g., circles, squares, triangles).

Examples Land Use land cover classification, Soil types, political boundaries relating to ward, province etc.

Map Data Types

Fundamentally, maps display only two types of data: qualitative and quantitative.

Qualitative data type

Fig: Map with delineated the boundaries of a province’s watersheds

Source: https://webapps.itc.utwente.nl/librarywww/papers_2009/general/principlesgis.pdf

Map Data Types

Fundamentally, maps display only two types of data: qualitative and quantitative.

Qualitative data type

There are mainly two types of qualitative data type. They are:

1. Nominal

- They do not have a natural order or ranking.
- This type of data is correctly represented using the visual variables like shape and hue.

2. Ordinal

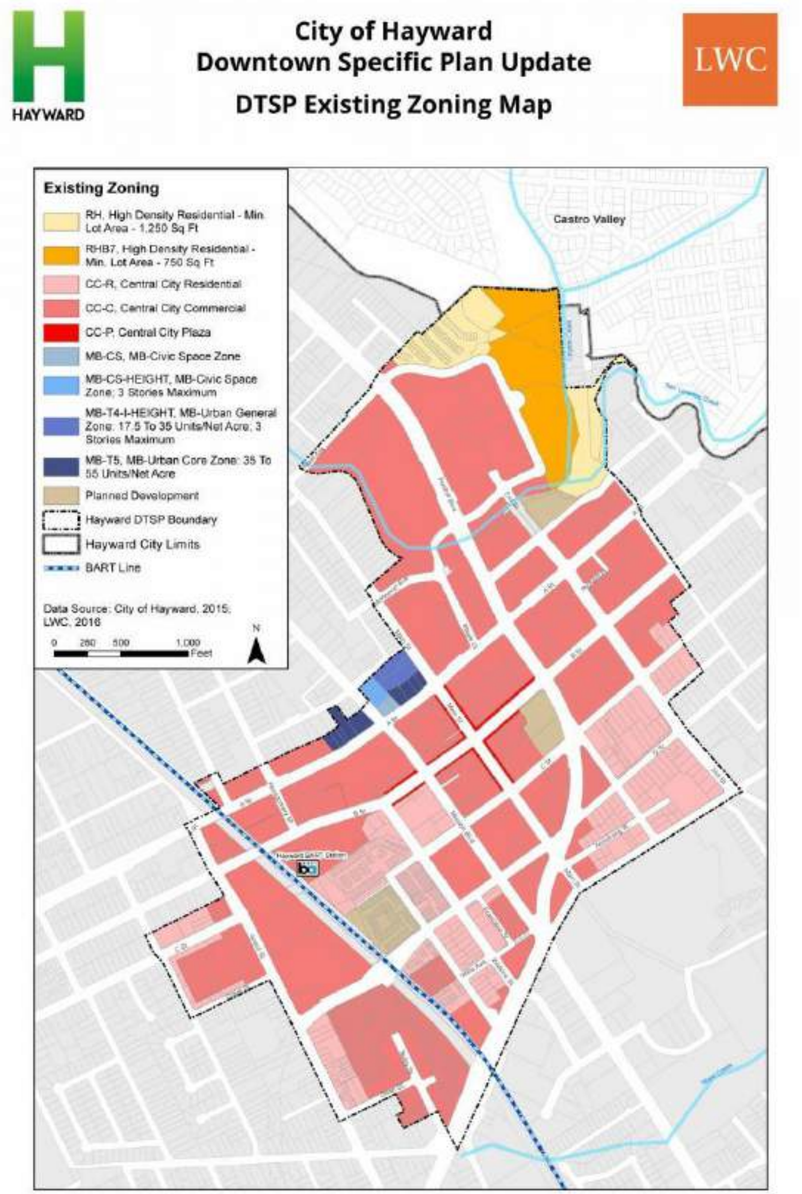


Fig: showing A zoning map using the visual variable color hue.

Source: https://www.e-education.psu.edu/geog486/node/638

Qualitative data types

Ordinal

- Ordinal data are categorical data where there is a logical inherent ordering to the categories.
- Data with a clear element of order, though not quantitatively determined (ordered).

Examples of ordinal data often seen on reference maps include political boundaries that are classified hierarchically (national, state, county, etc.) and transportation routes (primary highway, secondary highway, light-duty road, unimproved road).

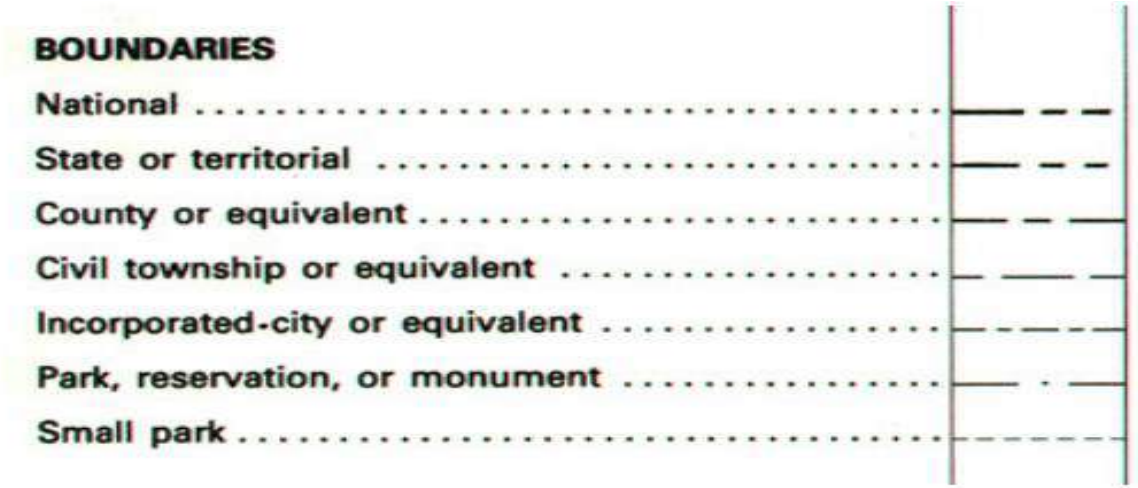


Figure 3.10.1 Attribute data measured at the ordinal level: Ranked categories of boundaries depicted on USGS topographic maps.

Credit: Steger, 1986

Source: https://webapps.itc.utwente.nl/librarywww/papers_2009/general/principlesgis.pdf

Qualitative data types

Ordinal

Examples

a. Fire danger: low, medium, high

b. Suitable sites: highly suitable, less suitable, suitable,

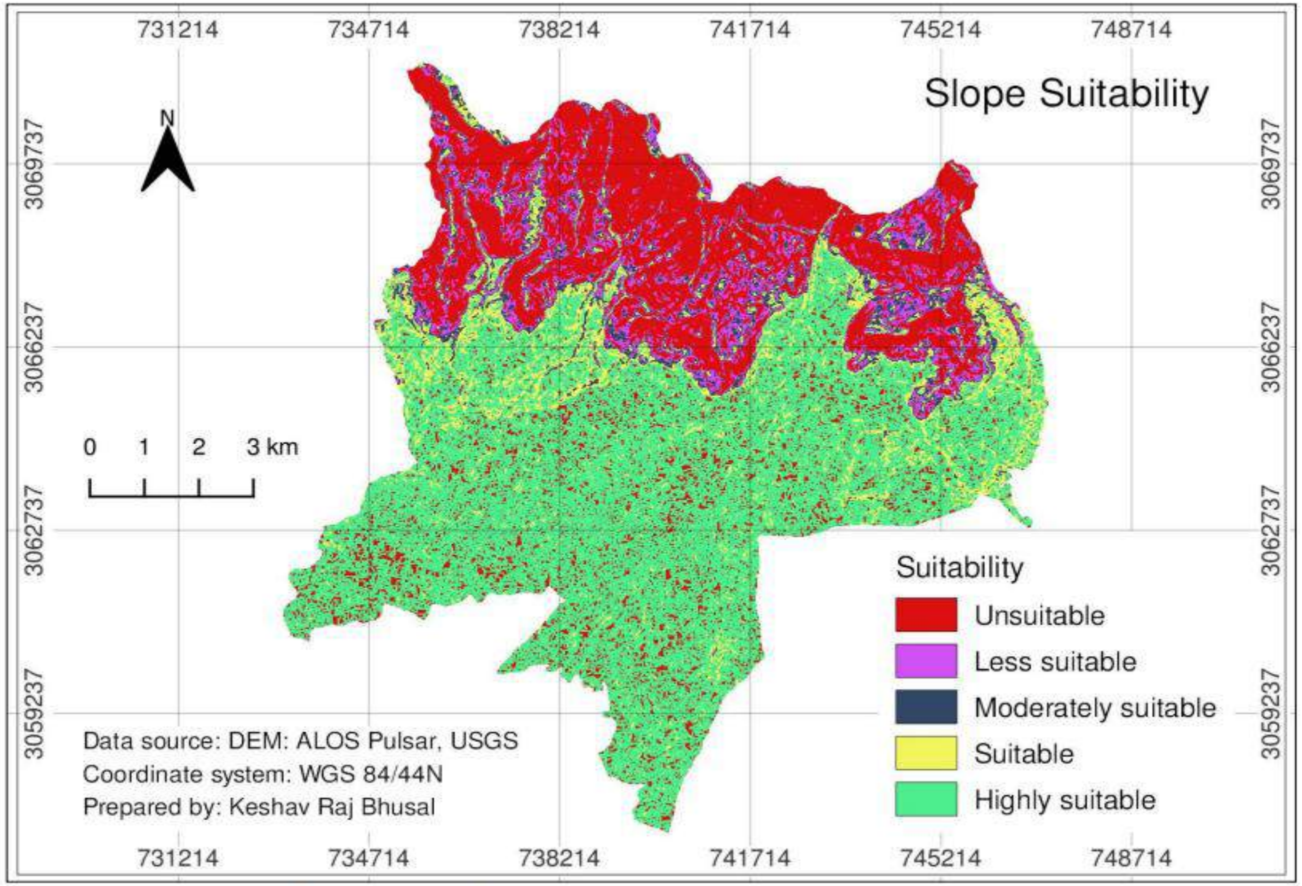


Figure 4.Showing suitable areas for landfill site according to slope criteria

Source: https://keshavrajbhusal.com.np/

Map Data Types

Quantitative data type

- Map with quantitative data in geographic units gets unequal attention, and some features stand out above the others.
- The categorical differences in quantitative data can be shown with symbols that vary in size

and color value.

Examples Dot density map, graduated color map etc.

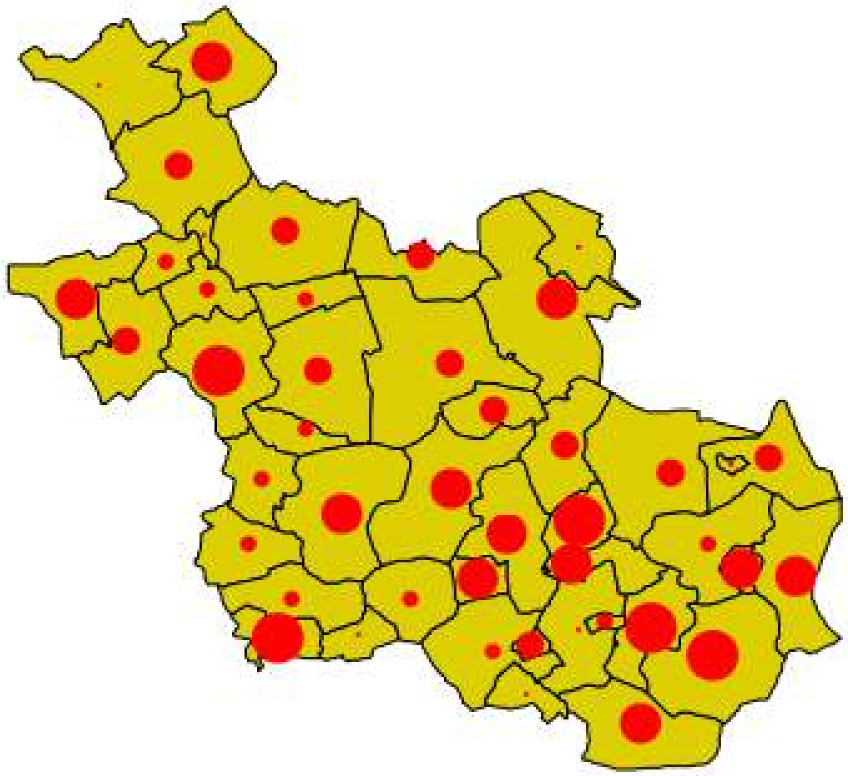


Fig: Map showing absolute quantitative data of total population in municipalities

Source: https://webapps.itc.utwente.nl/librarywww/papers_2009/general/principlesgis.pdf

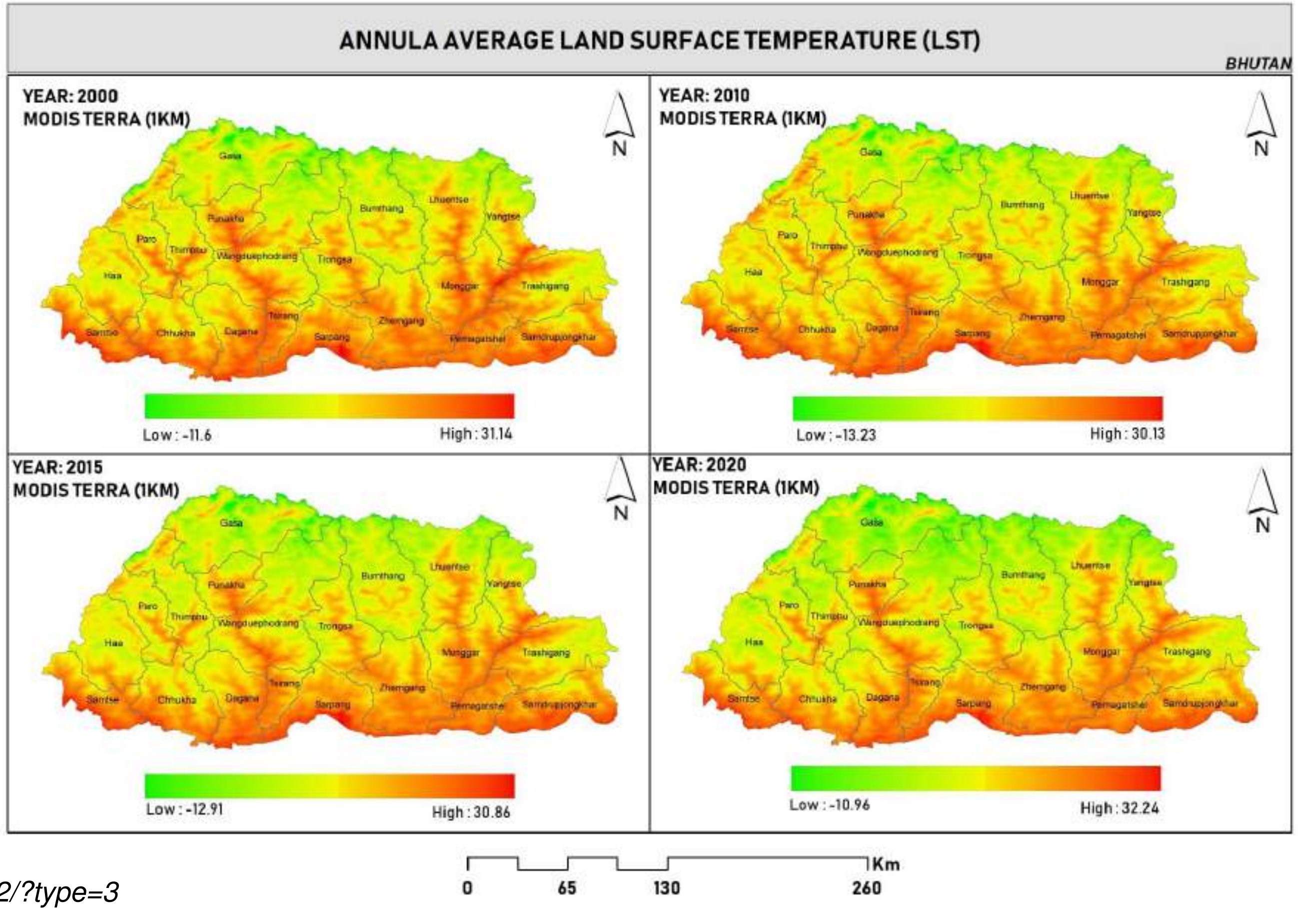
Quantitative data types

Quantitative data can be further classified as interval and ratio.

Interval

- Interval data are continuous values (not categorical) but here is no absolute zero on the scale.
- Examples include: Temperature, Ph value

etc.



Source: https://www.facebook.com/gisbhutan2019/photos/a.225288115597617/225288065597622/?type=3

Quantitative data types

Ratio

- This data type has a clear definition of 0. When the variable equals 0, there is none of that variable.
- The ratio values are calculated and expressed a series of ratios or proportional values, such as percentage, per km, per inhabitant. This kind of data is the most common.
- Examples include:, Population density etc.

Quantitative data types

Ratio

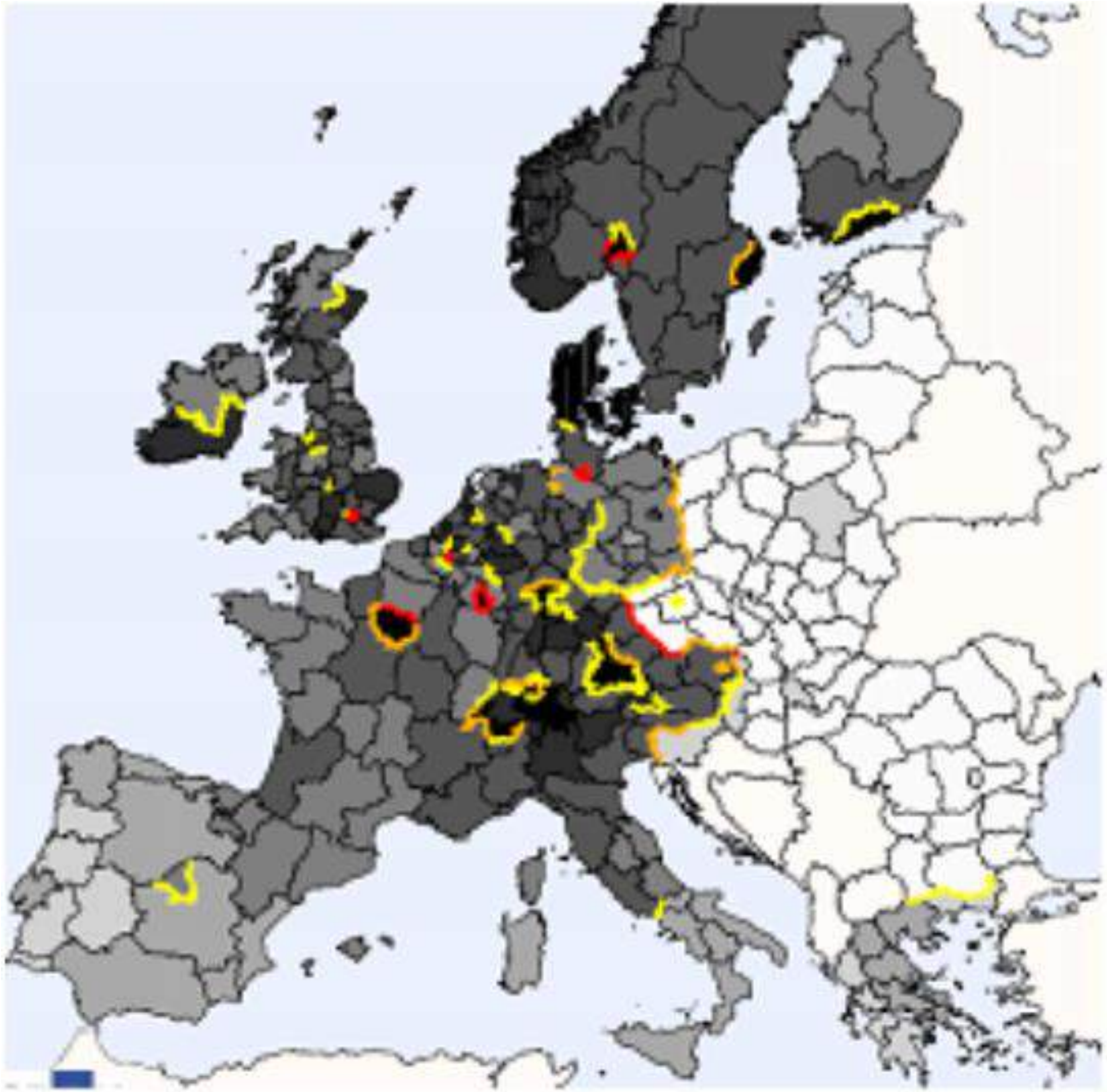


Fig: GDP per inhabitants of a city

Source: https://www.espon.eu/sites/default/files/attachments/MAPPING_GUIDE_EXTERNAL.pdf

Visual Variables

Visual variables

- Visual variables are *distinctions* that we can use to create and *differentiate* symbols on a map.
- We can make point, line, and area symbols on the map appear more or less distinctive and prominent by *altering* their shape, size, color, etc.
- The term ‘visual variable’ was introduced in 1967 by Jacques Bertin.
- The *choice* of visual variable depends on the *type of data* to be displayed.
- Qualitative Visual Variables: Hue, Orientation, Shape, Texture ; Quantitative Visual Variables: Value and Size
- The visual variables enable observers to perceive:

1. What is of equal importance(e.g through seeing same size or color used)
2. Order (e.g. the population density varies from low to high—represented by light and dark colour tints, respectively)
3. Quantities (e.g. symbols changing in size with small symbols for small amounts)
4. An instant overview of the mapped theme.

Visual variables types

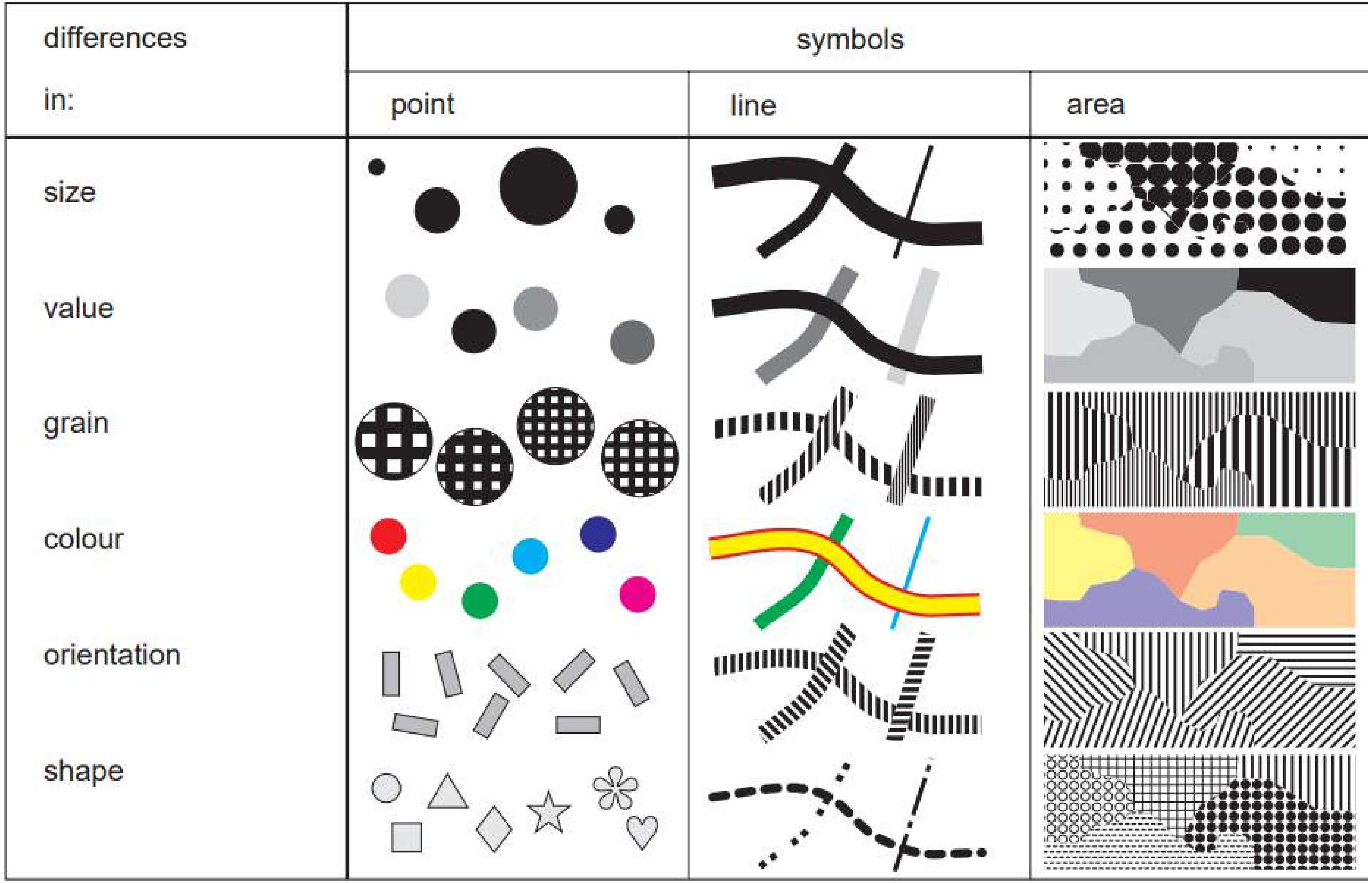


Fig: Bertin’s six visual variables illustrated

Source: https://webapps.itc.utwente.nl/librarywww/papers_2009/general/principlesgis.pdf

Visual variables types

1. Colour/Hue:

- Hue is great for identifying items as unique, or of a type of item.
- It creates a perception of groups or likeness.

2. Orientation: Orientation refers to the direction of a symbol.

3. Texture: Grain/Texture refers to the coarseness (resolution or spacing) of the graphic elements within a symbol while value remains constant.

- It refers to the pattern used to fill the body of the symbol.
- It can be to lines, using dash patterns, but it is mostly applied to areal symbols.

applied

4. Value/Saturation: is the lightness or darkness of a color, with black at the lower end and white at the higher end.

- We generally perceive darker symbols on a map as being more important, or higher in magnitude.

Visual variables types

5. Size: Size refers to the dimension of the symbol.

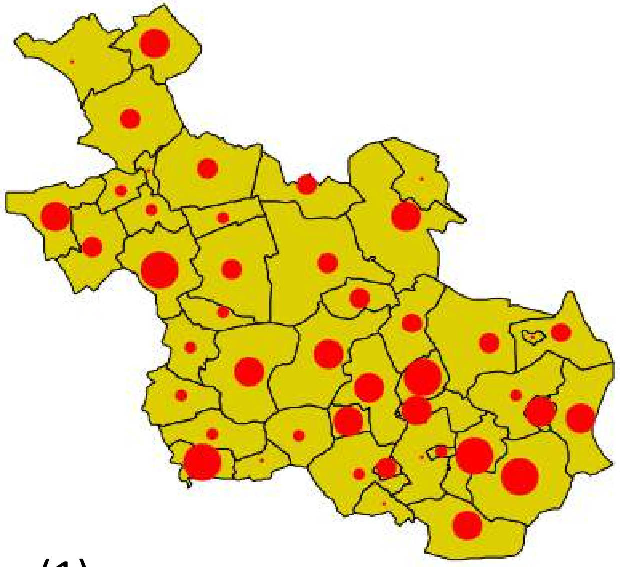
- In the case of points, it can be applied by changing the size of the symbol itself.
- In the case of lines, changing their thicknesses is the most usual way of applying this visual variable on them.
- It is not used in areal symbols, except in the case of using a texture fill, in which the size variable is applied to the texture and not to the symbol itself.

6. Shape: It is defined by the perimeter of the object.

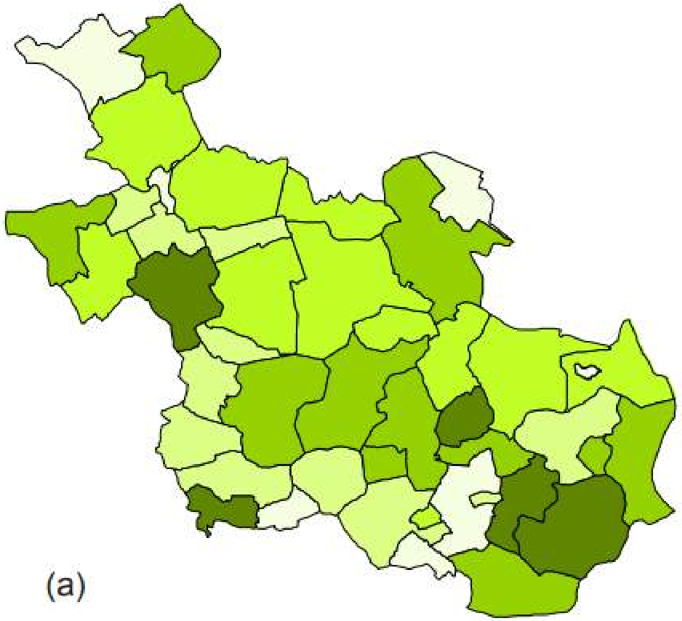
- This variable is mostly used in the case of *point* data, using a symbol of a given shape located at the exact coordinates of the point to be rendered.

Visual variables

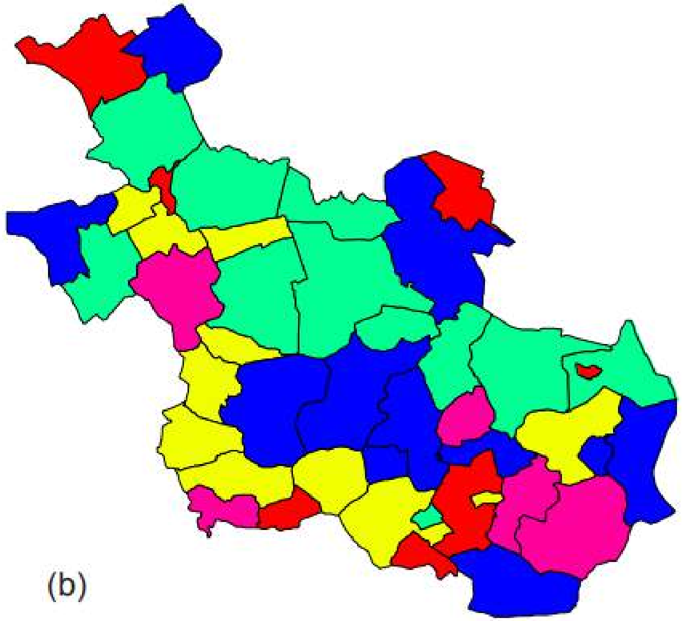
Good vs Bad use of visual variables in mapping Absolute Quantitative data



(1)



(a)



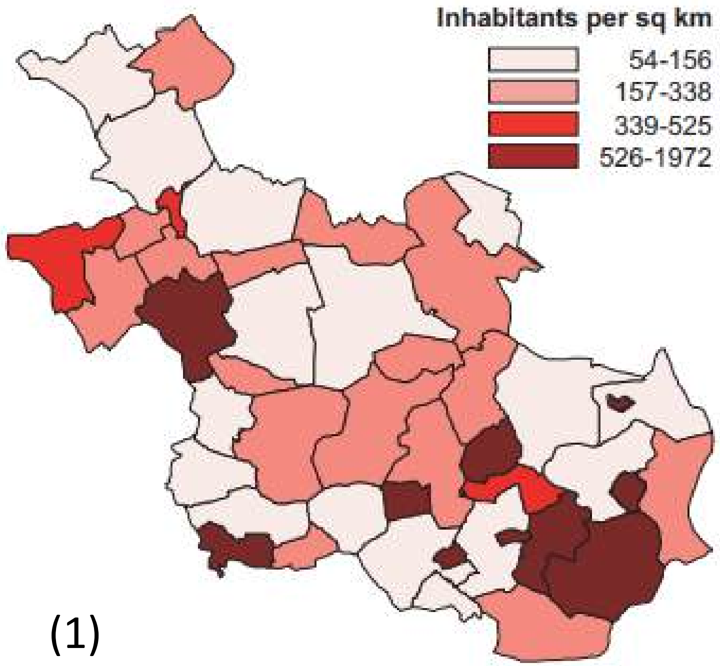
(b)

Figure 7.15: Poorly designed maps displaying absolute quantitative data: (a) wrong use of green tints for absolute population figures; (b) incorrect use of colour

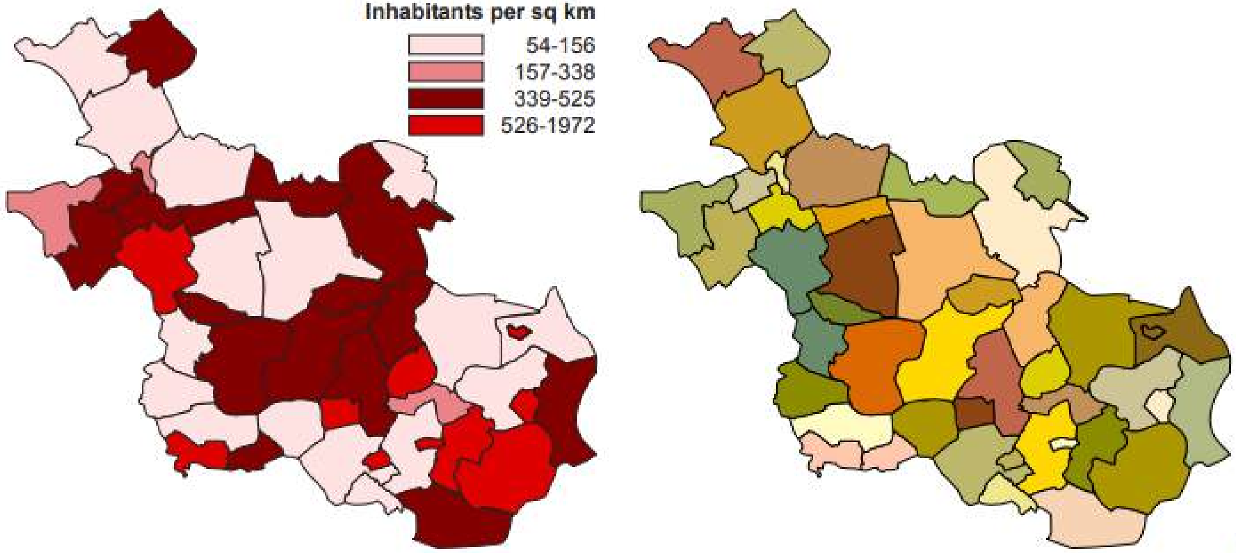
Source: https://webapps.itc.utwente.nl/librarywww/papers_2009/general/principlesgis.pdf; pg no. 474

fig(1): symbols in red color varying in size displays the data correctly; greater symbol signifies greater population and so on.
Fig (a): Viewer might think that dark green region having more area of geographic units has got more population but that might not be the case instead dark green region with smallest geographic units could have greater population
Fig(b): Due to too much use of color user can't understand where is more population.

Good vs Bad use of visual variables in mapping Relative Quantitative data



(1)



(a) (b) **Figure 7.17**: Badly designed maps representing relative quantitative data: (a) lightness values used out of sequence; (b) colour should not be used

Source: https://webapps.itc.utwente.nl/librarywww/papers_2009/general/principlesgis.pdf; pg no. 475

fig(1): Value has been used to display the density from low (light tints) to high (dark tints). This helps viewer to understand info easily.
Fig (a): Incorrect application of the visual variable value. In this map, the value tints are out of sequence. The user has to go through quite some trouble to find out high-density areas.
Fig(b): Colour has been used in combination with value. Might reader may think brown color has more population density at first. This map misguides the viewer.

Types of Maps

Types of map for visualization

1. Static map

- The digital equivalents of paper map.
- Is often *digital scan* of a paper map or an *export* from a GIS package.
- They can be shared in web and offer more than view-only functionality, they may present an interactive view to the user by offering zooming, panning, or hyperlinking to other information..
- Scale of the map cannot be changed.
- Areas off the edges of the map cannot be viewed and components of the map cannot be *toggled* on or off.

Types of map for visualization

2. Dynamic map

- Appearance changes frequently or automatically.
- Dynamic maps are about change; change in one or more of the spatial data components. On the WWW, several options to play animations are available. The so-called animated-GIF can be seen as a view-only version of a dynamic map.
- Data of dynamic maps are connected with the database and are updated in map whenever the database is updated.

Types of map for visualization

3. Interactive Maps

- Appearance changes according to user input(zoom in, zoom out, popup).
- Interaction; zooming into a particular part of the map that you are interested in, inspecting feature properties, editing existing content, panning etc.
- Most popular web map on the internet.

E.g google map, google earth etc.

Types of Maps

Base and thematic map

1. Base map

- A basemap is a layer that provides geographical context to the map and other dataset layers above it.
- The function of the basemap is to provide background detail necessary to orient the location of the map. It also makes map aesthetically appealing.
- Typical GIS data and imagery that make up the layers for a basemap: streets, parcels, boundaries (country, county, city boundaries), shaded relief of a digital elevation model, waterways, and aerial or satellite imagery.
- Without the base map layer, there would be no geographical context and the map would be useless.

Depending on the type of map, any combination of those layers can be used. For example, A map showing hiking trails could use a basemap containing a digital elevation model or topo lines that shows elevation, thus allow viewers to understand the rise and fall of a trail's path.

Base and thematic map

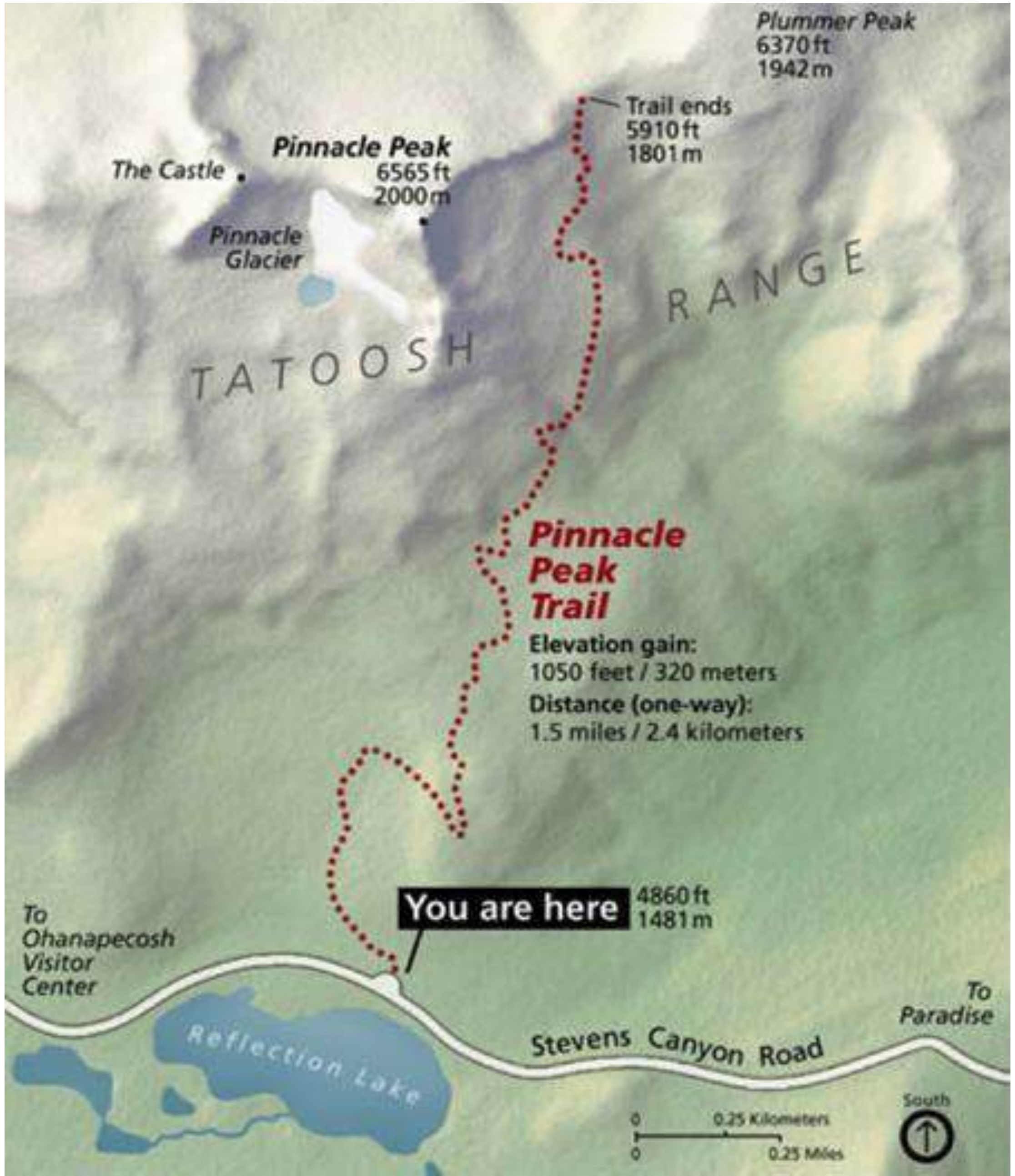


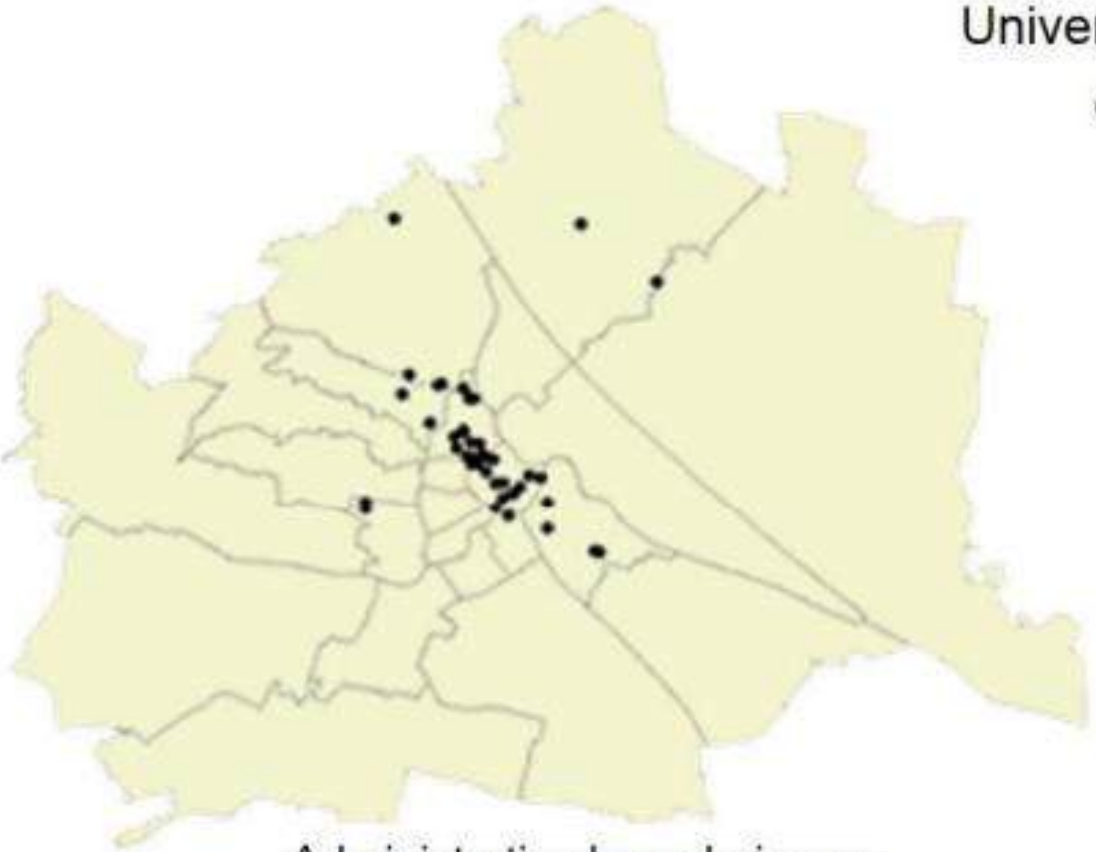
Fig: Trail map with shaded relief basemap.

Source: http://shadedrelief.com/3D_Terrain_Maps/3dterrainmapssho.html

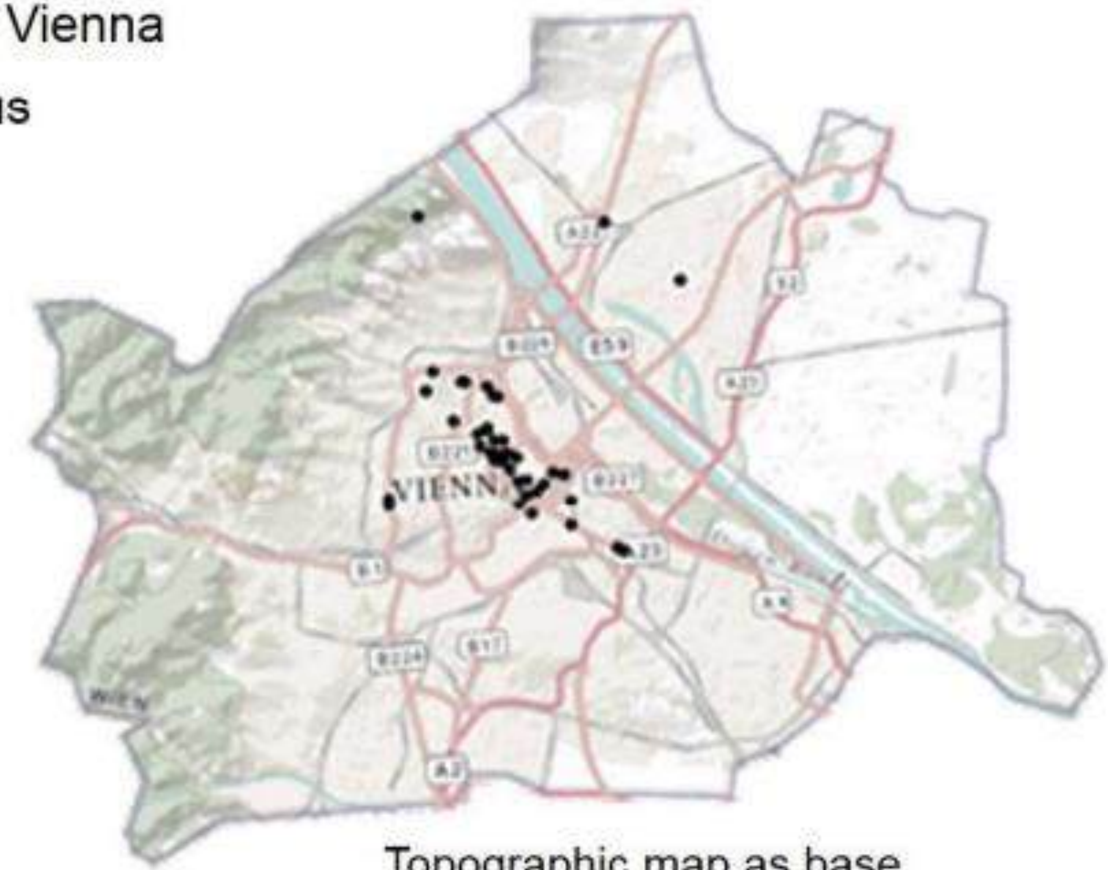
Types of Maps

Base and thematic map

University of Vienna Campus



Administrative boundaries as base map



Topographic map as base map

Different base map layers, depending on the map theme and purpose

Base and thematic map

2. Thematic map

- Maps created specifically to highlight the distribution of a particular phenomenon or theme are called thematic maps.
- Unlike topographic maps the exact location is of a secondary importance.

However users can associate values to an approximate location by using reference data such as administrative boundaries.

- A thematic map fundamentally consists of the topographic base map to provide background information and the superimposing thematic content.
- They are usually made with a single purpose in mind of revealing the spatial distribution of one or two attribute data sets.
- Thematic maps display effectively qualitative(categorical), quantitative(numerical), absolute values and relative values.

Types of Thematic map

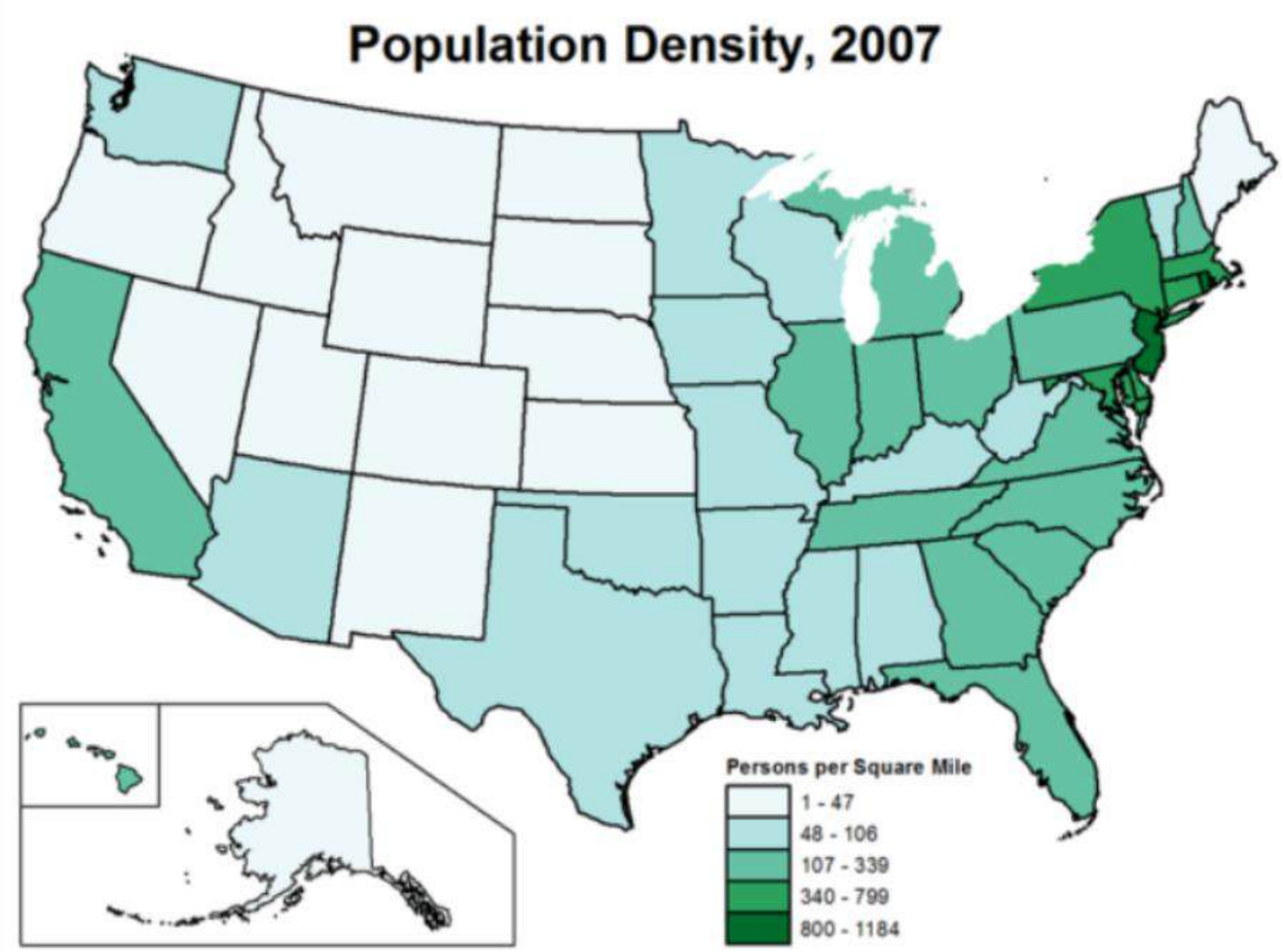
Depending on the type of the data to be displayed, the user will have to decide on the type of the thematic map.

i. Choropleth Map(Choro + pleth = area + value; Value applied to area)

- Choropleth maps depict quantities aggregated to their regions by filling the entire region(district, province etc.) with a shade or color. Typically, the quantities are grouped into “classes” (representing a range in data value) and a different fill is used to depict each class.
- The more dense an object is distributed within an area the darker the colour or the more dense the pattern should be.
- The map maker uses a type of data classification (equal interval data classification, quantile, natural breaks etc.) to produce its own unique choropleth map. Each data classification method impacts the reader differently.
- Choropleth maps are common ways to represent social and economic phenomenon.

E.g population density per square km of a district

Choropleth Map:



Source: https://alg.manifoldapp.org/read/introduction-to-cartography/section/c3c06272-8b8b-49e7-a957-da0d06550b73

Fig: Population density per square mile of county shown by varied saturation of green hue.
lighter color green represents lower population density in the darker color green represents a higher population density

Choropleth map

When to use choropleth map?

i. When your data are attached to enumeration units (e.g municipalities, provinces, wards etc.).

ii. When your data are standardized to show ratios/rates. for e.g number of people in a each municipality divided by the area of that municipality = number of people per square mile/km)

- The raw counts tell you how many people exist in each municipality, and the ratio tells you how tightly packed-in those people are.
- Standardization/Normalization allow your readers to compare places that are very different. For example, you can directly compare a large place (like Canada) with a smaller place (like Switzerland). Although Canada has more people than Switzerland, it has a far lower population density; without standardization, that fact might not be obvious.

Types of Thematic map

Choropleth Map:

Disadvantage:

- Because the data must be in ratio standardization, and relate to a defined physical area, much of the specifics of the data become lost, as the map only represents the average of what is occurring in these pre-defined physical units.
- Smaller places are easily overlooked on a choropleth map—even if they have large data values—while large countries such as Canada dominate the map no matter what color they are.

Advantage:

- They can cover large amounts of data over large geographical areas, without confusion to the reader.

Types of Thematic map

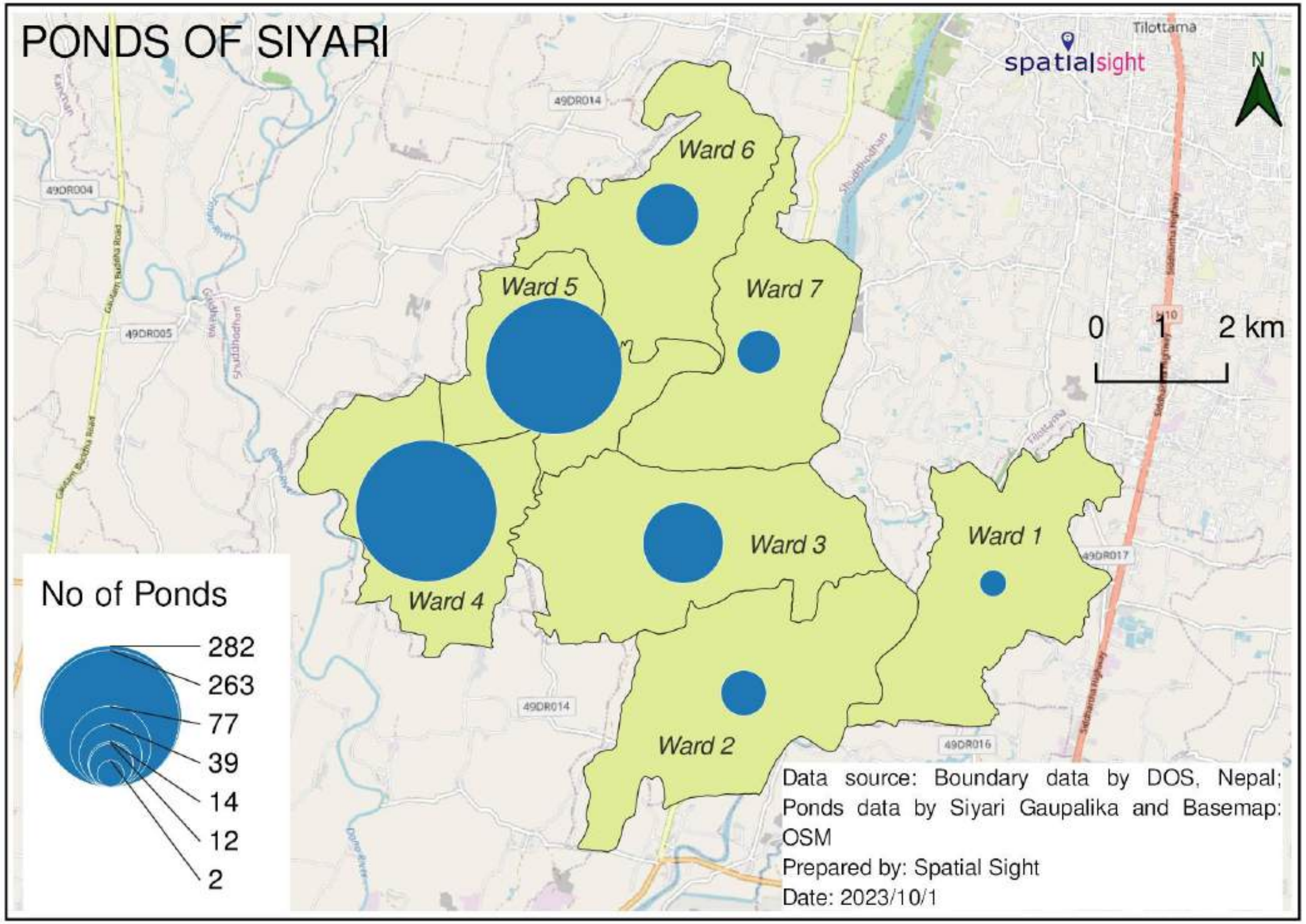
ii Proportional Symbol Map

- They visualize data using the size of a symbol (most often a circle) placed inside an enumeration unit.
- Larger symbols typically represent higher values of the variable being mapped, while smaller symbols represent lower values.
- Its symbology is unclassed. In other words, proportional symbol maps scale dots with absolute magnitude.

When to make proportional symbol map ?

- When you have absolute quantitative data for respective enumeration units. For example, population data of district, death or infected cases.

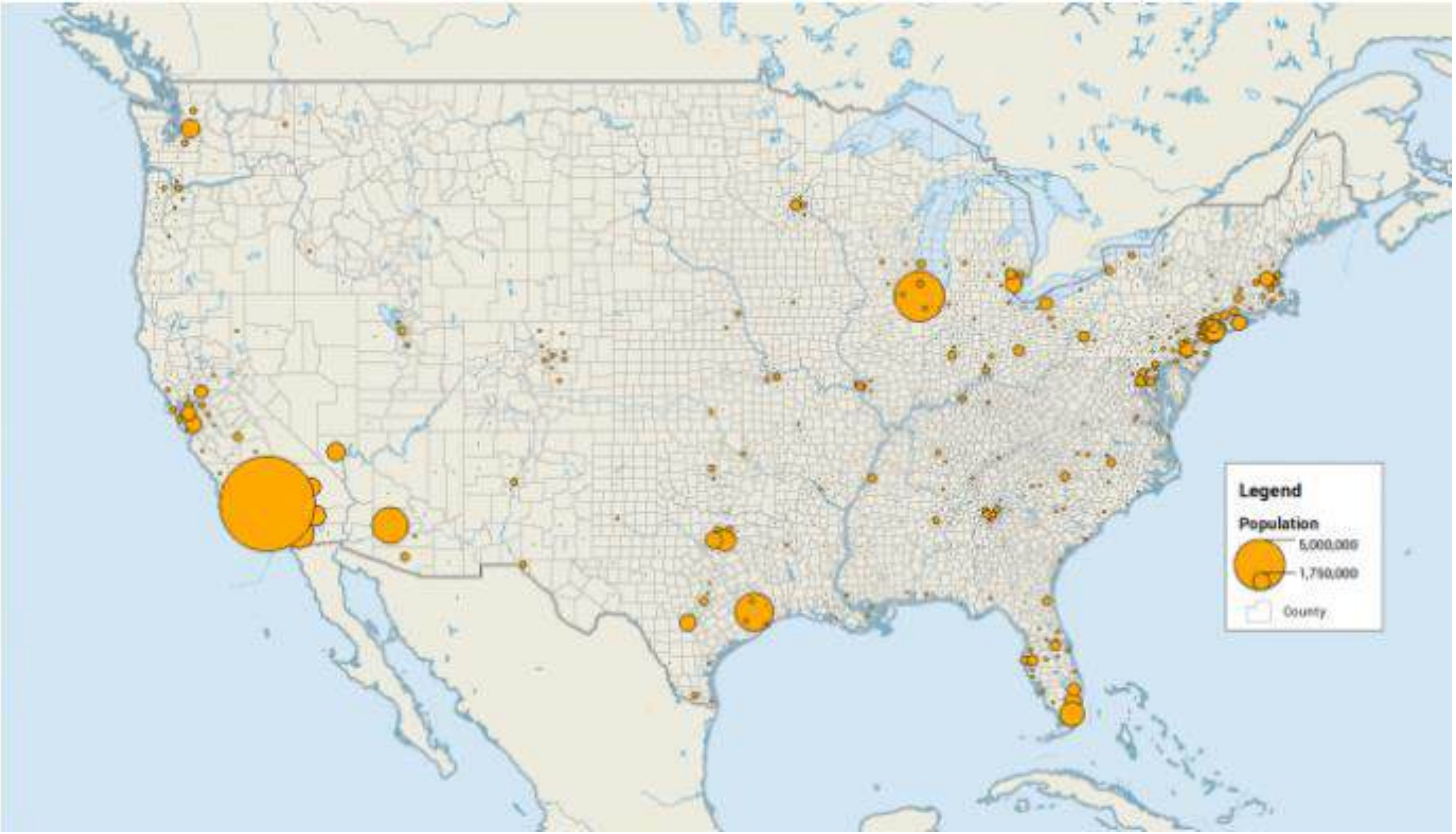
ii Proportional Symbol Map



Source: https://keshavrajbhusal.com.np/

Fig: Proportional symbol map depicting no. of ponds with circle size.

ii Proportional Symbol Map



Source: https://gisgeography.com/dot-distribution-graduated-symbols-proportional-symbol-maps/

Fig: Proportional scale map displaying the county absolute population data:

ii Proportional Symbol Map

Advantage:

i. One advantage of proportional symbol maps over dot density maps is it is generally easier for map readers to extract numbers from the map since estimating the size of a symbol is less tedious than counting many little dots.

ii. An advantage of proportional symbol maps over choropleth maps is that the size of the enumeration unit doesn’t matter. For e.g if a country with a small geographic area, such as the Netherlands, has a large data value attached to it, it will have a large symbol over it.

iii. Proportional symbol maps can use either raw data (totals, counts) or standardized data (percentages, rates, ratios).

Disadvantage:

- Displaying large data for small enumeration units symbols become congested and hard to read.

Types of Thematic map

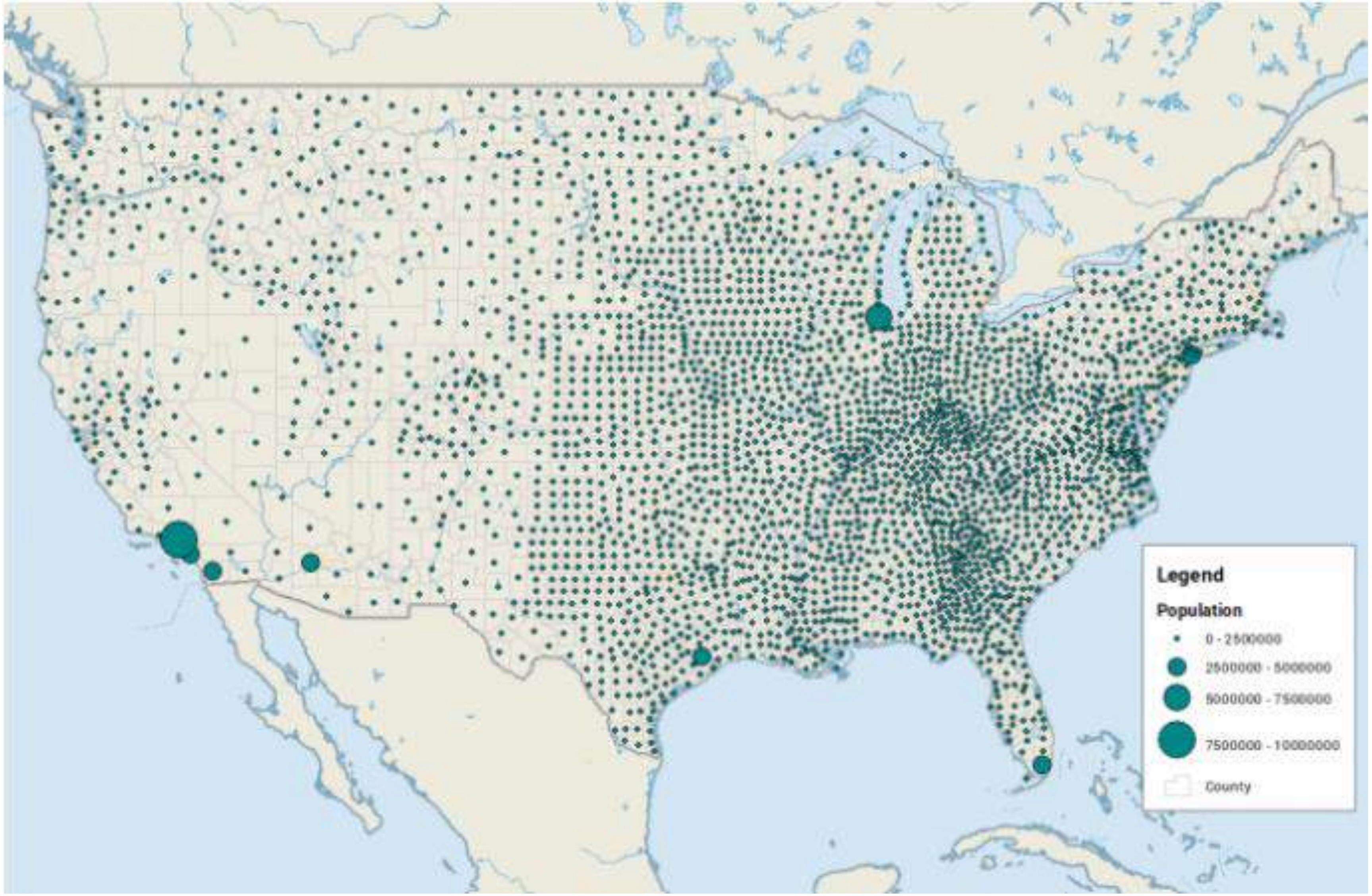
iii Gradual Symbol Map

- It also represents magnitude of data by its size.
- Its symbology is classed. It creates classes using data classification techniques like equal interval, quantile, and natural breaks.

Advantage

- An advantage of graduated symbol maps over choropleth maps is that the size of the geographic feature doesn't matter.

iii Gradual Symbol Map



Source: https://gisgeography.com/dot-distribution-graduated-symbols-proportional-symbol-maps/

Fig: Gradual symbol map displaying the county classed population data:

Types of Thematic map

iv Dot Density Map(Dot Distribution Maps)

- Dot-density maps depict a quantity for a given area by filling it in with small dots.
- Instead of larger symbols meaning “more” of something like in the previous two maps, dot distribution maps shows “more” dots.
- They are easy to understand and, at a glance, show us instantly where things clump or cluster.

For example, here is a dot distribution map for US census tracts where each dot is 50,000 people.

Advantage:

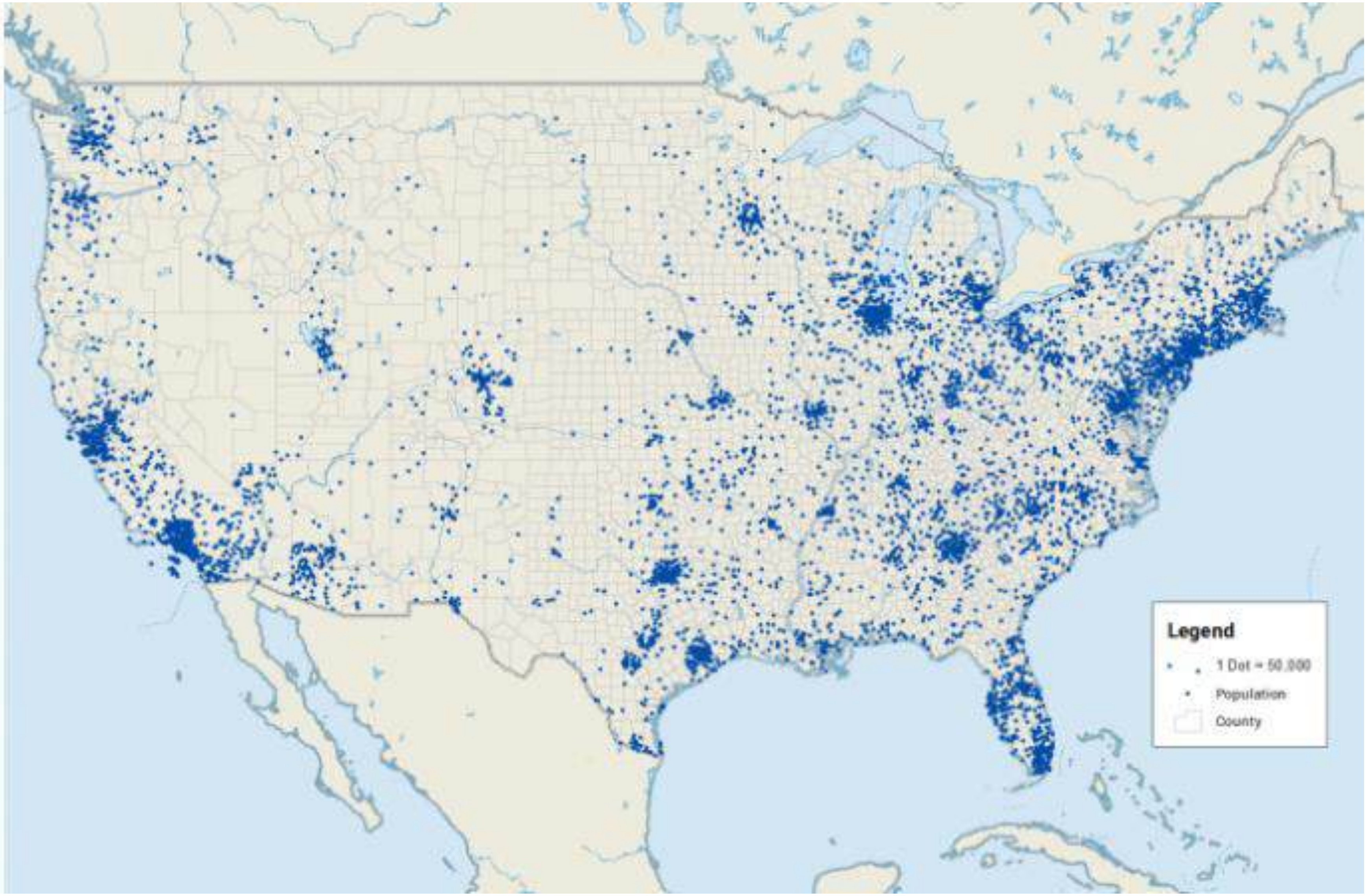
- On a dot density map you can map raw data / simple counts (e.g., number of farms) or rates and ratios (e.g., number of farms per sq kilometer).
- Your data need not be tied to enumeration units and hence some of the concerns inherent in choropleth maps can be side-stepped with dot density maps (unless, of course, your data are reported by enumeration units.

Disadvantage:

- One of the disadvantages of dot density maps is that it’s difficult to extract quantities from them. t’s tedious to count small-sized dots for a large quantity.

Types of Thematic map

iv Dot Density Map(Dot Distribution Maps)



Source: https://gisgeography.com/dot-distribution-graduated-symbols-proportional-symbol-maps/

Fig: Dot density map displaying the county population data where one dot equal to 50000:

Map design Principles

Map Design Concept and importance

- The purpose of map design is to enhance a map so that it is easy to understand and able to communicate the correct message or information.
- A well-designed map is balanced, coherent, ordered, and interesting to look at, whereas a poorly designed map is confusing and disoriented.
- Map design is both an art and a science.
- There may not be clear cut right or wrong designs for maps, but there are better or more-effective maps and worse or less effective maps.
- Map design overlaps with the field of graphic arts, and many map design principles have their origin in visual perception.

Cartographers usually study map design from the perspectives of layout and visual hierarchy.

Map Design/Cartographic Principles

There are mainly five principles of map design. They are:

1. Visual Contrast
2. Legibility
3. Figure-Ground Organization
4. Hierarchical Organization
5. Balance

Map Design Principles

1. Visual Contrast

- Contrast is a basic element in map design, important to layout as well as to visual hierarchy.
- A well-designed map with a high degree of visual contrast can result in a crisp, clean, sharp-looking map. The higher the contrast between features, the more some features will stand out (usually features that are darker or brighter).
- Contrast in color can separate the figure from the ground. Cartographers often use a warm color (e.g., orange to red) for the figure and a cool color (e.g., blue) for the ground.
- Features that have less contrast appear to belong together.
- Too much contrast can create a confusing map appearance.

Map Design Principles

1. Visual Contrast

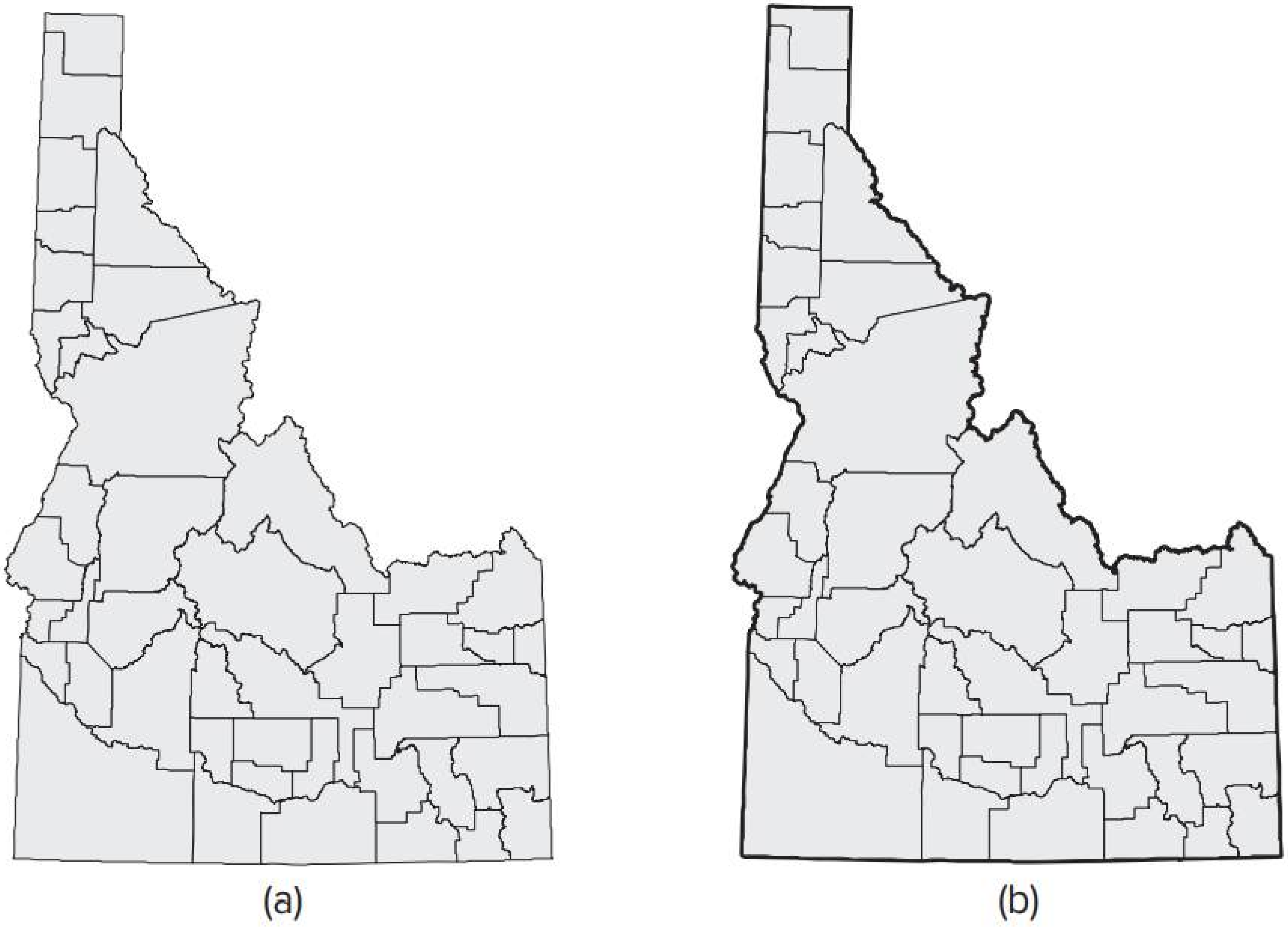
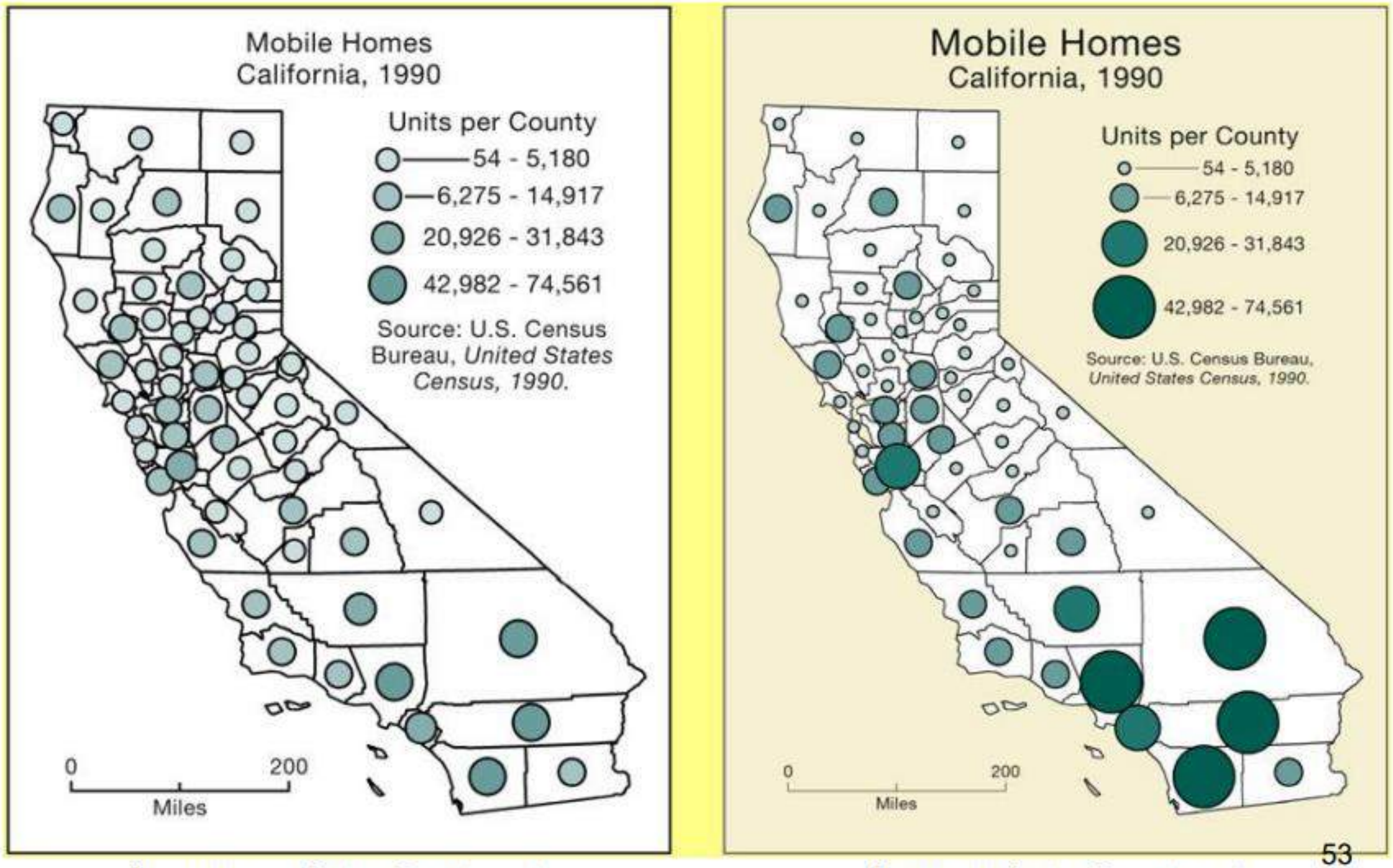


Fig: Contrast is missing in (a), whereas the line contrast makes the state outline look more important than the county boundaries in (b). Contrast in size or width can make a state outline look more important than county boundaries and larger cities look more important than smaller ones.

Map Design Principles

1. Visual Contrast



Inappropriate Contrast

Appropriate Contrast

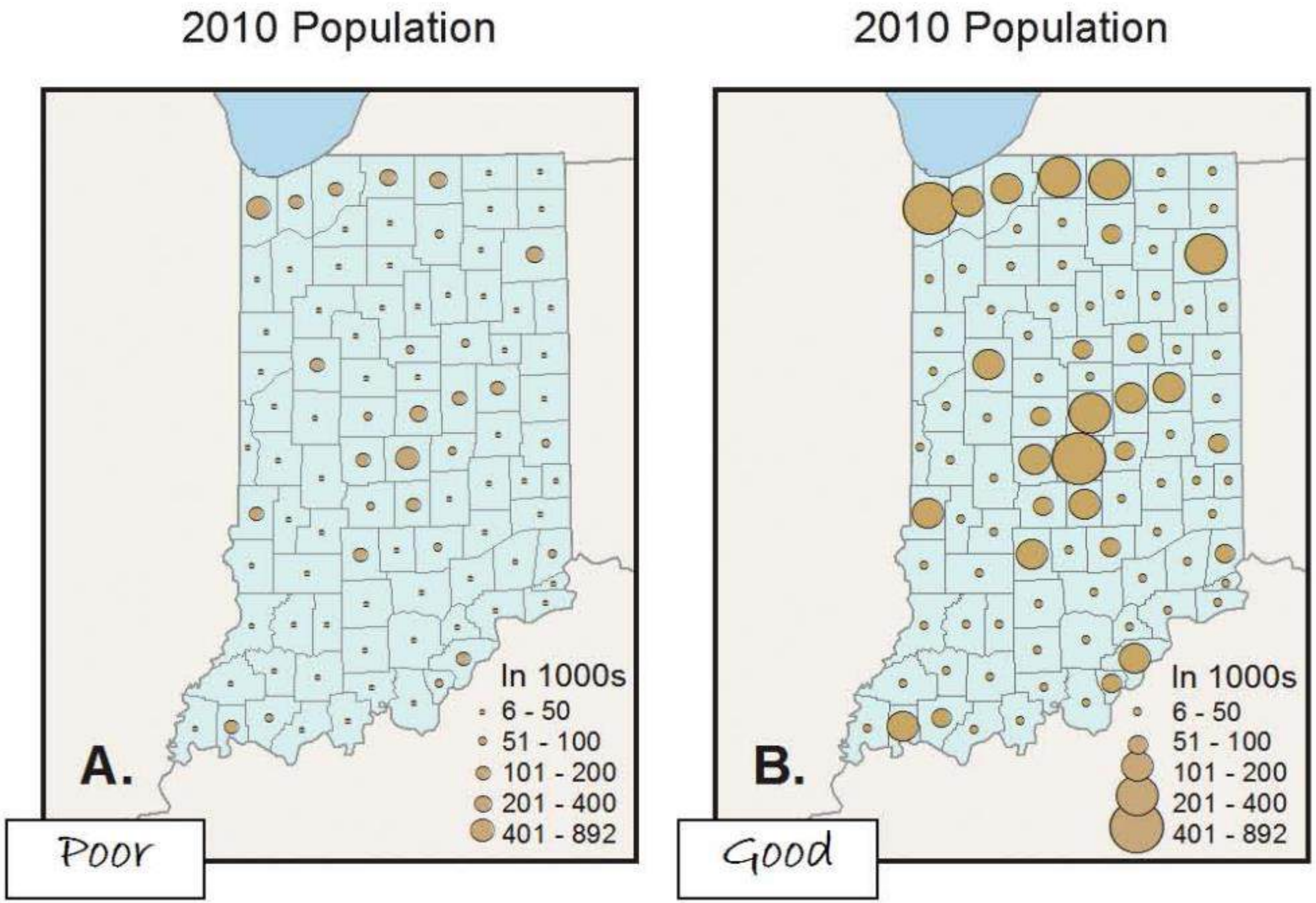
Map Design Principles

2. Legibility

- Legibility is the ability to be seen and understood.
- Legibility depends on good decision making when selecting symbols. Choosing symbols that are familiar and are appropriate sizes results in symbols that are effortlessly seen and easily understood.

Map Design Principles

2. Legibility

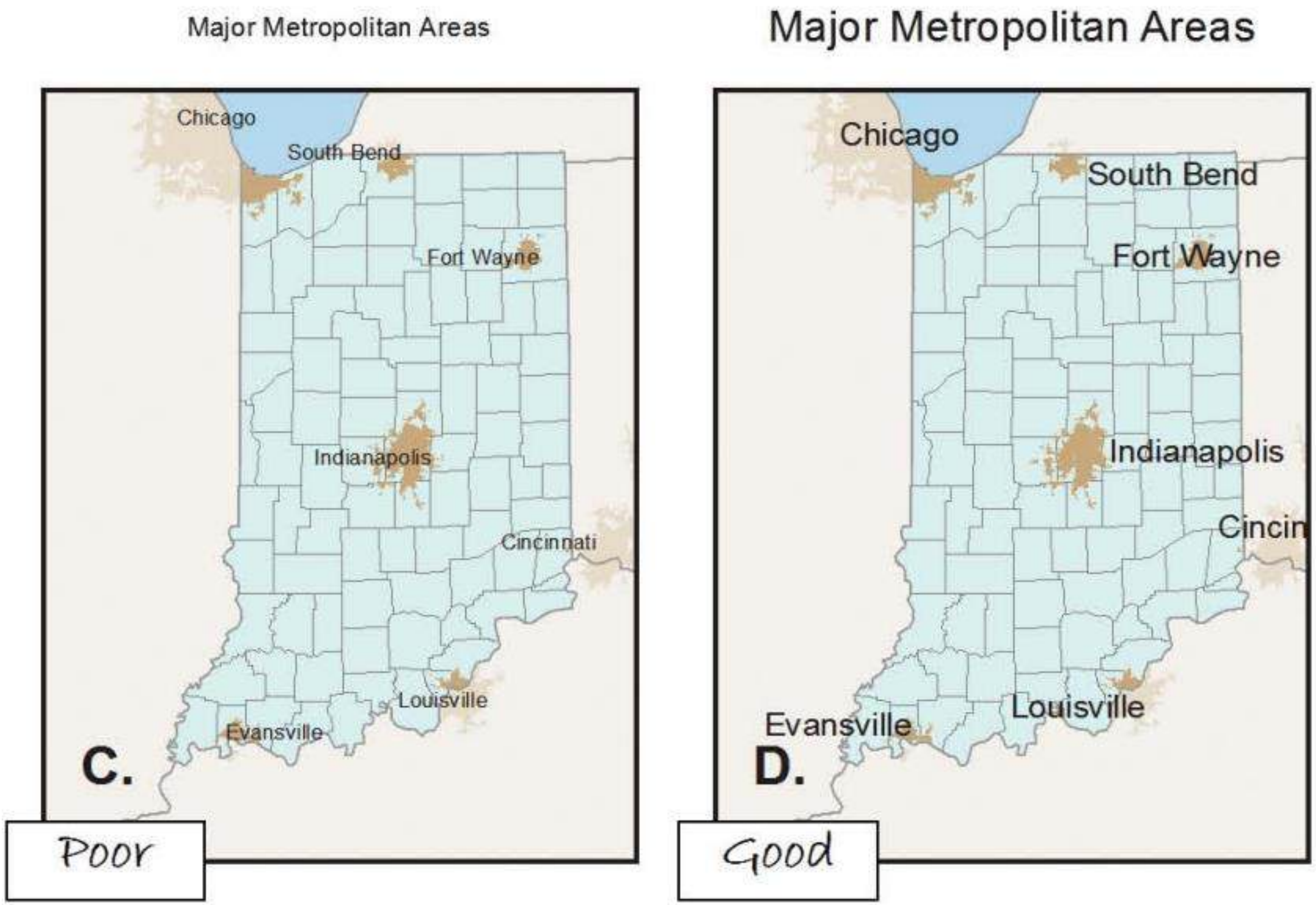


Source: https://www.esri.com/news/arcuser/0112/files/design-principles.pdf

Fig: Symbols (A) and text (C) that are too small are illegible. Appropriately sized symbols (B) and text (D) can be easily distinguished and read

Map Design Principles

2. Legibility



Source: https://www.esri.com/news/arcuser/0112/files/design-principles.pdf

Fig: Symbols (A) and text (C) that are too small are illegible. Appropriately sized symbols (B) and text (D) can be easily distinguished and read

Map Design Principles

3. Figure ground relationship

- One key to visual hierarchy is understanding figure-ground relationship.
- Some things appear prominently as foreground, and others as background.
- Color and contrast play an important role in figure-ground relationships. The less contrast in a map, the poorer the figure-ground effect, and likely the poorer the overall hierarchy.
- Your map has some background color, whether it’s a blank white page or something else. For elements that are important, use colors that are noticeably different from the background, especially in terms of lightness.

Map Design Principles

3. Figure ground relationship

Source: https://www.axismaps.com/guide/visual-hierarchy

Good use of figure ground

Fig: Color and brightness contrast let the land come forward from the ocean, and things like the river come forward from the land.

Map Design Principles

3. Figure ground relationship

Source: https://www.axismaps.com/guide/visual-hierarchy

Bad use of figure ground

Fig: Low contrasts make it harder and slower to understand map

Map Design Principles

4. Visual Hierarchy

- Visual hierarchy is used to emphasize certain important features on a map over less important features according to the map's purpose.
- An effective visual hierarchy attracts the map user's eyes to the most important aspects of the map first, and to less important aspects later.
- It results in maps that are easier to interpret, and are more attractive.
- It is the graphical representation of the intellectual hierarchy.

Map Design Principles

Map Design Principles

Intellectual hierarchy

- It is the ordering or ranking of all symbols and elements in the map according to their relative importance.
- The following is a general hierarchy for thematic maps.

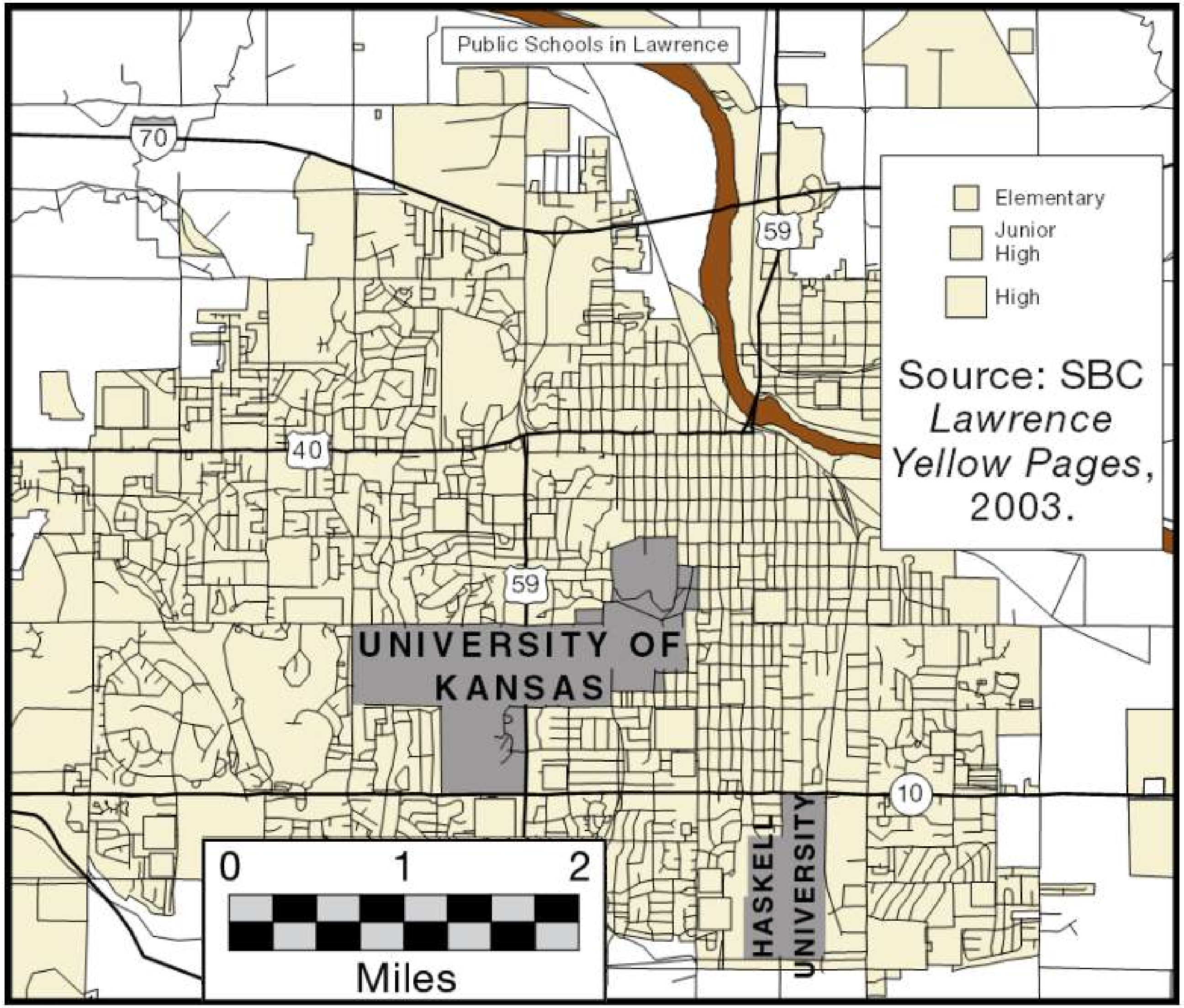
Most Important

- Thematic symbols and type labels
- Title, subtitle, and legend
- Base information (boundaries, roads, etc.)
- Scale and north arrow
- Data source and notes
- Frame and neat lines

Least Important

Map Design Principles

4. Visual Hierarchy



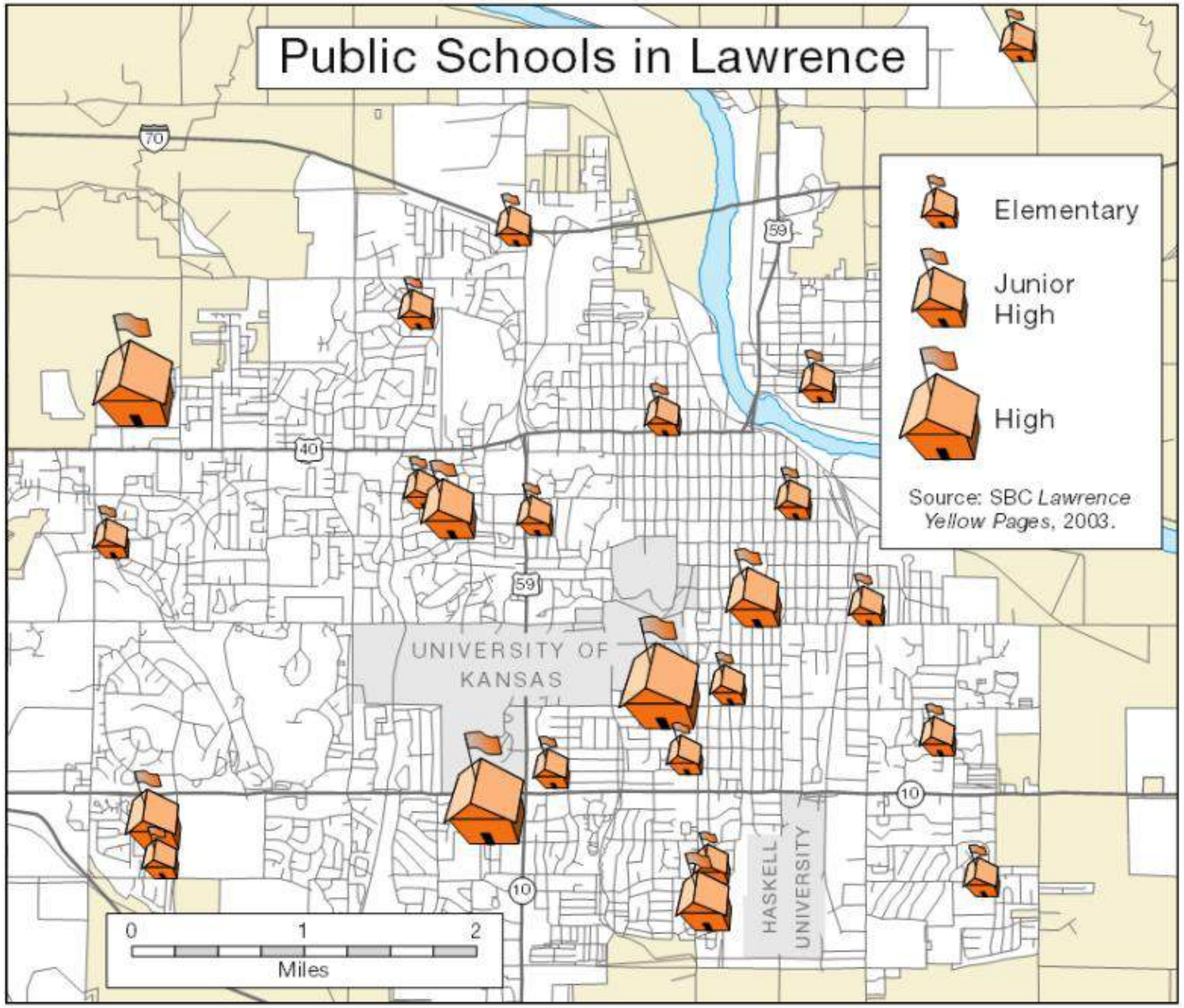
Inverted
(incorrect)
Visual Hierarchy

<u>Based on an Inverted Intellectual Hierarchy</u>

- Frame and neat lines
- Data source and notes
- Scale and north arrow
- Base information
- Title, subtitle, and legend
- Thematic symbols and type

Map Design Principles

4. Visual Hierarchy



Correct
Visual Hierarchy

<u>Based on a Correct Intellectual Hierarchy</u>

- Thematic symbols and type
- Title, subtitle, and legend
- Base information
- Scale and north arrow
- Data source and notes
- Frame and neat lines

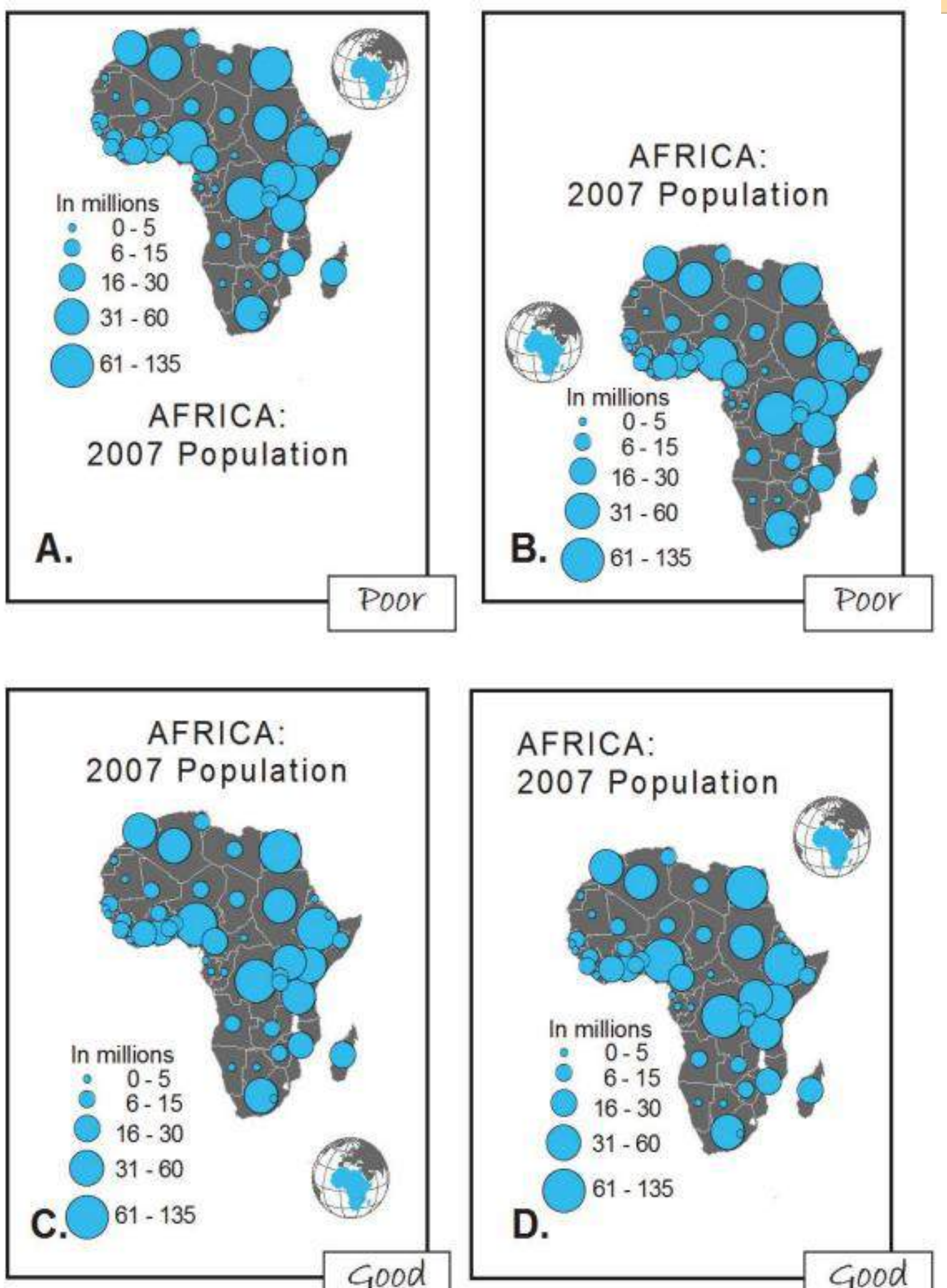
Map Design Principles

5. Balance

- Balance involves the organization of the map and other elements on the page.
- A well balanced map page results in an impression of equilibrium and harmony.
- It should not give the map reader an impression that the map “looks” heavier on the top, bottom, or side.

Map Design Principles

5. Balance



Source: https://www.esri.com/news/arcuser/0112/files/design-principles.pdf

Web map and digital map

Web map

- Web map" often implies a map that is not simply on the web, but rather one that is powered by the web.
- A web map depends on the internet. It is usually interactive.

E.g google map

Features of a webmap

i. **Dynamic scales and content**: Web maps are not static images. Different scales display varying levels of detail. For instance, zooming in may reveal information that wasn't apparent before. Real time info

ii. **Interactive:** Web maps are built to be interacted with by an end user, often in order for the user to explore a dataset and learn something.

iii. **Display real-time data updates:** Web maps are useful for geovisualizing real-time data like weather, flight info etc.

iv. **Built of** : Client side technology like html, css, javascript etc. and serverside technology like Django, Postgres and Geoserver

Web map and digital map

Web map

Elements of a webmap

i. Basemap

ii. Thematic layers

- Placed usually on top of basemap.
- Main reason why people see your map.
- Data based on particular theme(e.g population density in all districts of Nepal)

iii. Interactive elements

- Upon which user can interact.

E.g popup giving some info, charts, time slider etc.

Digital map

- A digital map is on a computer and is relatively static.
- Is often *digital scan* of a paper map or an *export* from a GIS package.

Features

i. It can be uploaded in the web and accessed through the web.

ii. It requires computer or any electronic device storage.

iii. Its scale does not change.

References

- https://webapps.itc.utwente.nl/librarywww/papers_2009/general/principlesgis.pdf
- https://civilstuff.com/digital-maps-vs-paper-maps/
- http://gis.humboldt.edu/OLM/Lessons/GIS/06%20Vector%20Analysis%20Attributes/03_AttributeTypes.html
- https://www.e-education.psu.edu/geog486/node/638
- https://alg.manifoldapp.org/read/introduction-to-cartography/section/21773ec3-7eb3-4197-b46d-8a7c2ebaa25f#:~:text=Visual%20variables%20are%20distinctions%20that,saturation%2C%20hue%2C%20and%20value.
- https://volaya.github.io/gis-book/en/Visualization.html#:~:text=From%20left%20to%20right%3A%20position,used%20to%20visualize%20geographical%20information.

Data type

https://gsp.humboldt.edu/olm/Lessons/GIS/06%20Vector%20Analysis%20Attributes/03_AttributeTypes.html#:~:text=Types%20of%20Data&text=Quantitative%20data%20is%20either%20Ordinal,data%20with%20an%20inherent%200).

https://www.espon.eu/sites/default/files/attachments/MAPPING_GUIDE_EXTERNAL.pdf

https://www.graphpad.com/support/faq/what-is-the-difference-between-ordinal-interval-and-ratio-variables-why-should-i-care/

https://geol260.academic.wlu.edu/course-notes/introduction-and-data-types/selecting-data-types/

Basemap

https://www.gislounge.com/basemaps-defined/

References

Choropleth Map

https://alg.manifoldapp.org/read/introduction-to-cartography/section/c3c06272-8b8b-49e7-a957-da0d06550b73

http://www.axismaps.com.s3-website-us-east-1.amazonaws.com/guide/data/standardizing-data/

https://www.axismaps.com/guide/choropleth

https://gis.depaul.edu/shwang/teaching/geog258/QuantitativeThematicMap.htm

Map design concept

https://alg.manifoldapp.org/read/introduction-to-cartography/section/b9662d0a-4926-4947-922d-a67a2cae0eec

Types of map

https://101gis.com/reference-map-vs-thematic-map/

GIS APPLICATIONS

EG 3202 GE

Unit 4: Data quality

Public Namuna Secondary School

Lecture by

Er. Keshav Raj Bhusal

Table of contents

4.1 Data Quality

- Define Data quality
- Data quality components
  - Spatial accuracy
  - Attribute accuracy
  - Temporal Accuracy
  - Logical Consistency
  - Lineage etc.

4.2 Accuracy Assessment

- Accuracy and precision of data
- Error propagation
- RMSE error during accuracy assessment (user accuracy and producers accuracy)

Data Quality

Data quality

- Data quality refers to the degree of excellence exhibited by the data in relation to the representation of the actual phenomena.
- It is the degree of data excellency that satisfy the given objective.
- Data quality is a pillar in any GIS implementation and application as reliable data are indispensable to allow the user obtaining meaningful results.
- It gives the answers to the questions such as:

How good are the data?

Is information available allows a user to decide if the data are suitable for our purpose?

What is the positional and attribute accuracy?

Are the data complete?

Data Quality

Data quality

Spatial Data quality can be categorized into Data completeness, Data Precision, Data accuracy and Data Consistency.

- **Data Completeness**: It is basically the measure of totality of features. A data set with minimal amount of missing features can be termed as Complete-Data.
- **Data Precision**: Precision can be termed as the degree of details that are displayed on a uniform space.
- **Data Accuracy:** This can be termed as the discrepancy between the actual attributes value and coded attribute value.
- **Data Consistency**: Data consistency can be termed as the absence of conflicts in a particular database.

Data Quality

Data quality components

i. Spatial or Positional accuracy

ii. Attribute accuracy

iii. Temporal Accuracy

iv. Logical Consistency

v. Lineage etc.

Data quality components

Spatial/Positional accuracy:

- It refers to the absolute and relative accuracy of the positions of geographic features.
- 'Absolute accuracy' is the closeness of the coordinate values in a dataset to values accepted as or being true.
- 'Relative accuracy' is the closeness of the relative positions of features to their respective relative positions accepted as or being true.

Data quality components

Attribute accuracy:

- Attribute accuracy is the assessment of the reliability of the values assigned to the features in the dataset in relation to their true real world values.
- The non spatial data linked to the location may also be inaccurate or imprecise. The inaccuracies may result due to many types of mistakes.
- Non spatial data vary greatly in precision. Precise attribute information describes the phenomenon in greater detail.

Temporal Accuracy

- Temporal accuracy refers to the update of the data or the measure of how well the database reflects the real world situation at the present or a particular point of time in the past.
- Since the positional and attribute components of the spatial data may change together or independently, it is also necessary to consider their temporal validity.

For example, For example the boundaries of the land parcel may remain fixed over a period of many years whereas their ownership attribute changes from time to time.

Data quality components

Lineage:

- Describes the history of the spatial data, including descriptions of the source material from which the data were derived, and the methods of derivation.
- It also contains the dates of the source material, and all transformations involved in producing the final digital files or map products.

Logical Consistency :

- Logical Consistency means how well the logical relationship between the items in the dataset or spatial objects in the dataset are maintained and whether or not the relationship makes sense.
- One way to test this is to check whether any given feature has contradictory attributes.

For example, suppose your database contains information about land cover, slope, and soil type. Areas with a value of "lake" for land cover should have a value of "0%" for slope. If it has "10%" or "20%." slope value then its not logically consistent.

Accuracy Assessment :

- Accuracy can be defined as the degree of closeness to which the information on map matches the real world.

Assessment is the act of for making creating comparisons forming an opinion making judgement and taking corrective actions required on something.

Thus, Accuracy assessments essentially determine the quality of the information derived from gis data.

- Accuracy assessments attempt to identify and quantify error. In these assessments, we compare map data with reference or ground truth data.

General Steps for Performing Accuracy Assessment:

1. Develop a sampling system
2. Collect reference data
3. Compare reference data to the map
4. Compute accuracy metrics

General Steps for Performing Accuracy Assessment:

1. Develop a sampling system

- Sample points are the reference data points from which the accuracy of the obtained data are determined and assessed.
- When collecting your reference data your locations should be unbiased. To achieve this you need to develop some sort of sampling scheme.
- In systematic sampling the sampling points are spaced uniformly at fixed intervals or along a grid.
- In stratified random sampling the number of random points is allocated based thematic classes.

For example if we have 10 classes and 100 sampling points, each class would be allocated 10 sampling points.

2. Collect reference data

- The reference data are collected directly from the field through the direct field data collection or from the indirect ways like high resolution satellite imageries or any other alternative existing database.

General Steps for Performing Accuracy Assessment:

3. Compare reference data to the map

- After the data has been collected and the geospatial analysis has been conducted it is time to compare the results of the two.
- This is done relatively easily by comparing the obtained map data to the reference data.

4. Compute accuracy metrics

- Data are compared and the accuracy of the map is determined.
- The accuracy of maps are quantified through various measures.

Accuracy and Precision

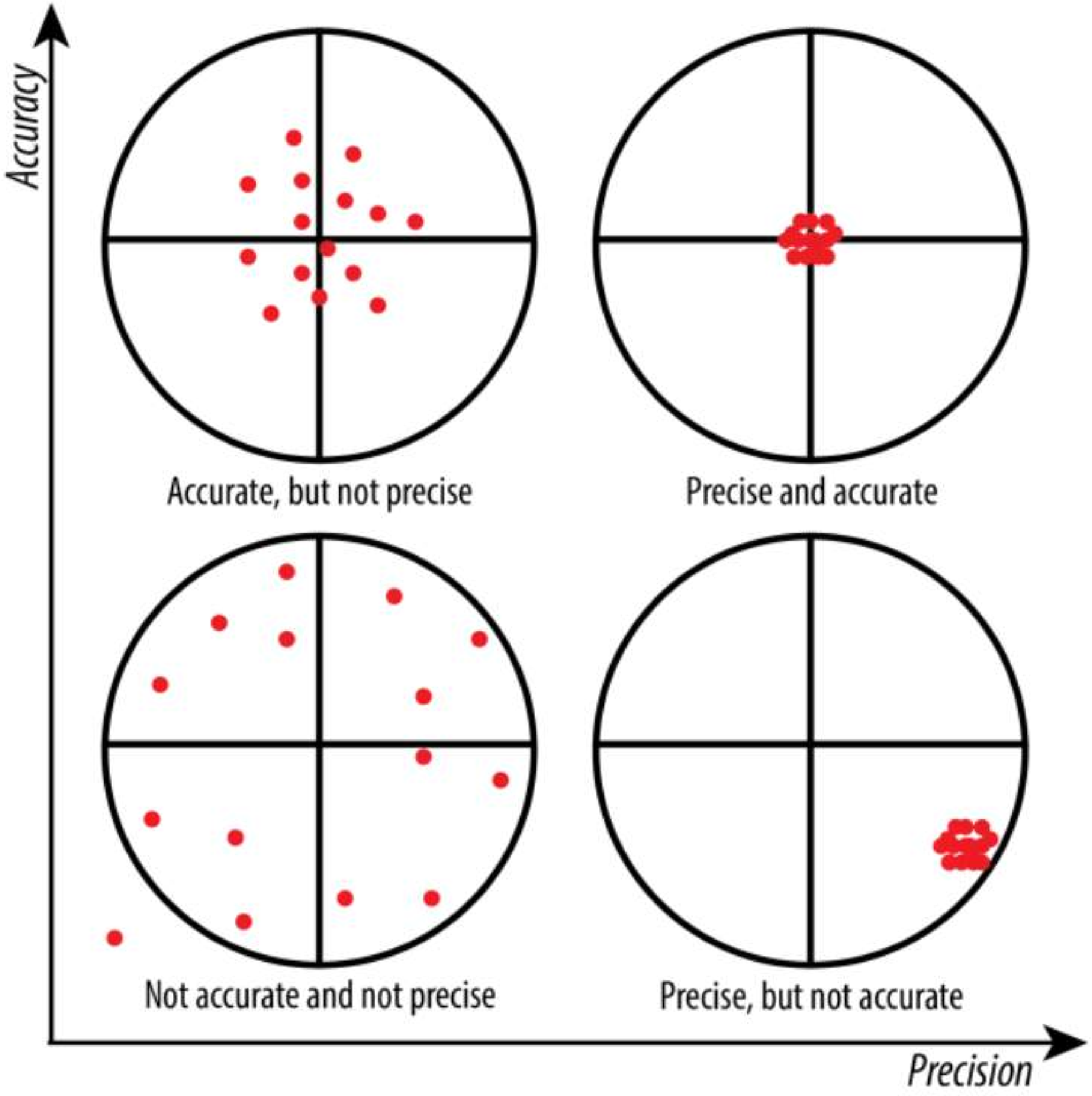
Accuracy of data

- Accuracy refers how close a measurement or observation is to the "true value".
- Accuracy in GIS is the degree to which information on a map matches real-world values.
- The level of accuracy required for particular applications varies greatly.
- Highly accurate data can be very difficult and costly to produce and compile.

Precision of data

- Precision refers how close measure values are to each other.
- Precise data no matter how carefully measured may be inaccurate.
- The level of precision required for particular applications varies greatly. Engineering projects such as road and utility construction require very precise information.
- Highly precise data can be very difficult and costly to collect.

Accuracy and Precision



Source: https://wp.stolaf.edu/it/gis-precision-accuracy/

RMSE

Root Mean Square Error

- Root Mean Square Error (RMSE) is an indicator for the accuracy.
- RMSE is the measure of the difference between the locations that are known and locations that have been interpolated or digitized.
- RMS Error is derived by squaring the difference between the known and unknown points, adding them together, dividing by the number of test points and taking the square root of that result.

$$RMSE = \sqrt{\sum_{i=1}^{n} \frac{y_{i-}\hat{y}_i}{n}}$$

where,

Σ : means “sum”

yi : is the observed value for the ith observation in the dataset

ˆi : is the actual value for the ith observation in the dataset

n : is the sample size

Accuracy Metrics

There are many different ways to look at the thematic accuracy of a classification. The error matrix allows you calculate the following accuracy metrics:

Overall Accuracy and Error

Errors of omission

Errors of commission

User’s accuracy

Producer’s accuracy

Accuracy statistics (e.g., Kappa

		Reference Data			
		Water	Forest	Urban	Total
Classified Data	Water	21	6	0	**27**
Classified Data	Forest	5	31	1	**37**
Classified Data	Urban	7	2	22	**31**
Classified Data	Total	**33**	**39**	**23**	***95***

Fig: showing error matrix show of classified data vs reference data

Source: http://gsp.humboldt.edu/olm/Courses/GSP_216/lessons/accuracy/metrics.html

Accuracy Assessment

Producers accuracy

- Producer's Accuracy is the map accuracy from the point of view of the map maker (the producer).
- This is how often are real features on the ground correctly shown on the classified map.
- It is also the number of reference sites classified accurately divided by the total number of reference sites for that class.
- Producer's Accuracy Example based on the above error matrix:

Water: Correctly classified reference sites = 21 Total # of reference sites = 33

Producer's Accuracy = 21/33 = 64%

Forest: Correctly classified reference sites = 31 Total # of reference sites = 39

Producer's Accuracy = 31/39 = 80%

Water: Correctly classified reference sites = 22 Total # of reference sites = 23

Producer's Accuracy = 22/23 =96%

Accuracy Assessment

User accuracy

- The User's Accuracy is the accuracy from the point of view of a map user, not the map maker.
- The User's accuracy essentially tells use how often the class on the map will actually be present on the ground (referred to as reliability).
- The User's Accuracy is calculated by taking the total number of correct classifications for a particular class and dividing it by the row total.
- User's Accuracy Example based on the above error matrix:

Water: Correctly classified sites = 21 Total # of classified sites = 27 User's Accuracy = 21/27 = 78%

Forest: Incorrectly classified sites = 31 Total # of classified sites = 37 User's Accuracy = 31/37 = 84%

Water: Incorrectly classified sites = 22 Total # of classified sites = 31 User's Accuracy = 22/31 = 70%

References

http://gsp.humboldt.edu/olm/Courses/GSP_216/lessons/accuracy/metrics.html

https://www.geos.ed.ac.uk/~gisteac/gis_book_abridged/files/ch12.pdf

GIS APPLICATIONS

EG 3202 GE

Unit 5: GIS Project Development

Public Namuna Secondary School

Lecture by

Er. Keshav Raj Bhusal

Table of contents

Project definition

Project

- A project is a sequence of unique, complex and connected activities having one goal or purpose that must be completed by a specific time, within budget and according to specification.
- Project is temporary and has a definite beginning and end. End is reached when the projects objectives have been achieved or because its objectives will not or cannot met.
- The project is bounded by three constraints: Time, Scope and Cost (Resources).

GIS Project

- The GIS project is the project that is done by using GIS.
- The components of GIS such as Data, Users, Hardware, Software and Methods are used for successful completion of project.

Some of the examples of GIS project are: Preparation of Topographical Map, Suitability Analysis, Network Analysis, Selection of Suitable site for waste disposal, Fixing Transmission line alignments and so on.

Stages of GIS Project Development

1. Project Conception and Initiation
2. Project Planning
3. Project Execution and Implementation
4. Project Evaluation

Stages of GIS Project Development

1. Project Conception and Initiation

- It is the start of the project and the goal of this phase is to define the project at a broad level.
- This is when you will research whether the project is feasible and if it should be undertaken or not.
- If the project is a "go" then the project charter is created, that outlines the purpose and requirements of the project.

2. Project Planning

- In this phase planning is done. Planning a GIS requires a commitment of resources and people.
- Project planning is developing a road map that everyone will follow.
- Typically begins with setting goals. Goals should be SMART (Specific Measurable Attainable Realistic and Timely).
- Scope of the project is defined and a project management plan is developed in this phase of project development.

Stages of GIS Project Development

3. Project Execution and Implementation

- Execution of the project basically involves the activity of the GIS team. This refers to the application development/customization, database creation, data encoding, data modelling and so on.
- This is the phase where all the activities are carried out in order to achieve the specific goal defined while creating the project.
- In this stage team is developed and different resources are assigned.
- Real implementation of project is done in this stage of the project.

4. Project Evaluation

- In this phase output are tested and evaluated to see if the GIS project meets its goals or not.
- This stage is done so as to control and check quality. Quality control involves agreement upon the quality standards that are applicable to project and then determining how to satisfy them.
- Quality checking involves the monitoring of the project results to determine whether they comply with quality standards.

Project Design

Project design is generally done in three stages; Conceptual Design, Logical design, Physical design

1. Conceptual Design

- Conceptual design defines the overall picture of the project.
- General work flow of the project is defined. The general steps needed to complete the project are defined in this phase.

2. Logical design

- Logical design is the detailed version of conceptual design phase.
- In this phase each and every steps and methods are defined in details. How each step is carried out is defined in this phase.

Project Design

Project design is generally done in three stages; Conceptual Design, Logical design, Physical design

3. Physical design

- Physical design is the implementation phase of the logical design phase.
- In physical design phase data models are designed that are required for the completion of the project.
- The resources required for the project are listed, possible problems that might be raised while carrying out the project is listed and possible solution is discussed.
- List of possible outcomes is also defined in this phase.
- Physical design phase generally contains following:

i. Design data models required for the project.

ii. List the resources required for the project.

iii. Discuss the possible problems.

iv. List out the expected project out comes.

References

GIS APPLICATIONS

EG 3202 GE

Unit 6: Spatial Data Infrastructure

Public Namuna Secondary School

Lecture by

Er. Keshav Raj Bhusal

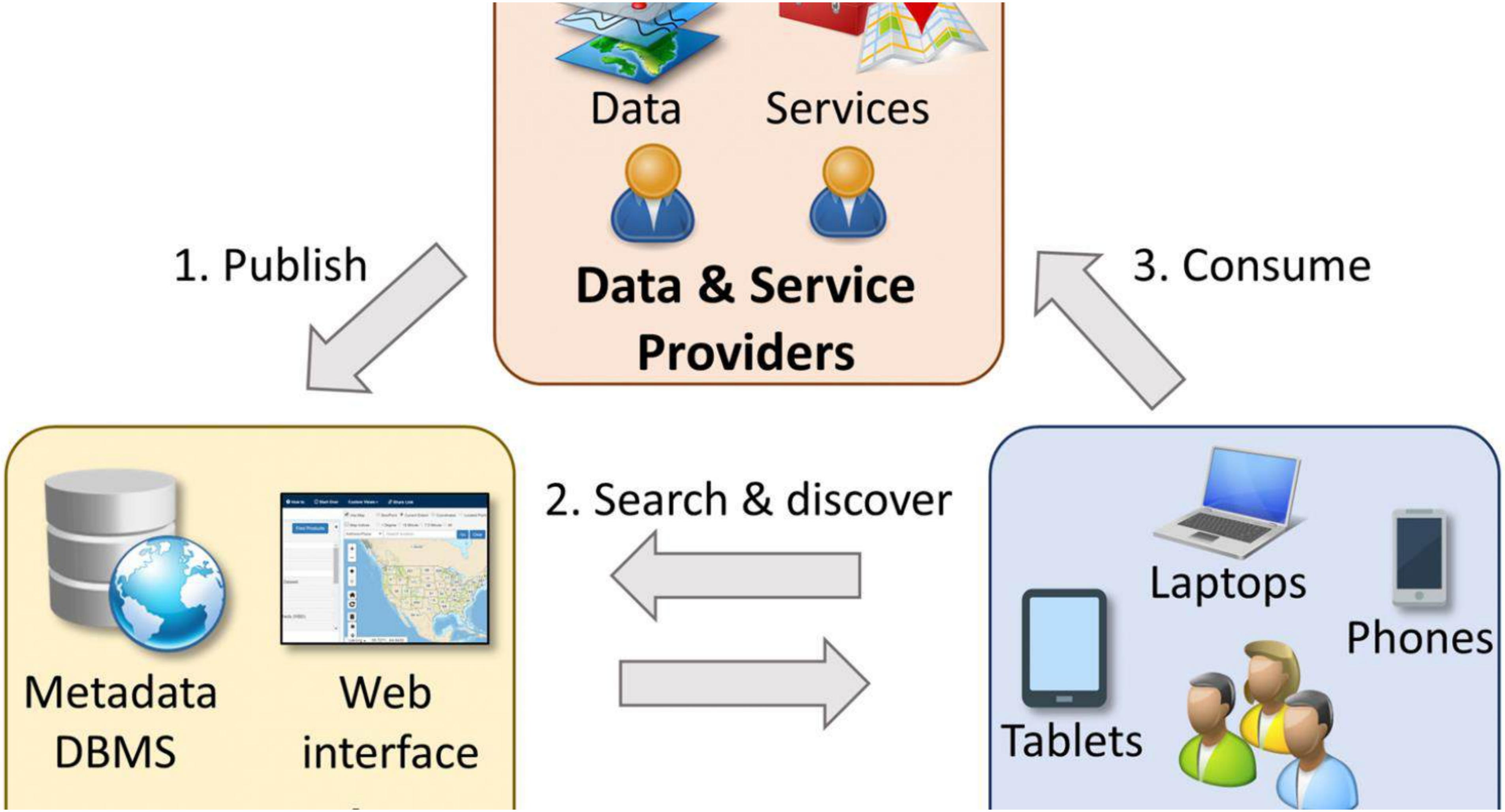
Table of contents

Concept of SDI

Spatial Data Infrastructure (SDI)

- Spatial data infrastructure (SDI) is the infrastructure that facilitates the discovery, access, management, distribution, reuse, and preservation of digital geospatial resources. These resources may include maps, data, geospatial services, and tools.
- The SDI provides a basis for spatial data discovery, evaluation, and application for users and providers within all levels of government, the commercial sector, the non-profit sector, academia and by citizens in general.
- SDI is a set of institutional, technical and economical arrangements to enhance the availability for correct, up-to-date, fit-for-purpose and integrated geoinformation, timely and at an affordable price. SDI also support decision making processes related to countries sustainable development.
- It provides a unified platform where people can go and search geospatial data, maps, services, and other digital resources.

Spatial Data Infrastructure(SDI)



Source: *https://gistbok.ucgis.org/bok-slide/spatial-data-infrastructure*

Concept of SDI

SDI Components

Basic components of SDI are data, users, policies and institutional arrangements, and technology and standards.

i. **Data:** Fundamental Datasets (digital base map, thematic, statistical, place names)

ii. **Users**: It includes people data provider as well as data consumer including different organizations and general public.

iii. **Policies**: Institutional framework (governance, data privacy & security, data sharing, cost recovery)

iv. **Technology**: It includes Access, Distribution, Storage (hardware, software, networks, databases, technical implementation plans)

- Geoportals

v. **Standards: (Data Models and Metadata)**

- Documented agreements containing technical specifications or other precise criteria to be used consistently as rules, guidelines, or definitions of characteristics, to ensure that materials, products, processes and services are fit for their purpose.

SDI Components

Geoportals

- Geoportals are probably the most visible part of SDIs, since they are the main interfaces through which people can search and find geospatial resources.
- Geoportals are typically developed using Web-based technologies and a database management system (DBMS) is used to store and manage the metadata of the geospatial resources contained in the SDI.
- A Web interface, which often contains a map, enables end users to interact with the system and to conduct searches.
- It helps users to discover the existing geospatial resources.

Some examples of such geoportals are: https://nationalgeoportal.gov.np/#/, Local Government - Data.gov etc.

Importance of SDI

i. Easy data access:

- It provides a unified platform where people can go and search geospatial data, maps, services, and other digital resources.

ii. Less costly integration:

- Integration of data from different sources can be done with reduced cost. Data can be stored in some of predefined standards and formats which ultimately helps in reducing the cost and complexity.

iii. Interoperability:

- Promotes interoperability. Data can be accessed, shared and used between two system.

iv. Reduces data redundancy:

- As multiple government agencies are sharing their data on one platform, SDI reduces data redundancy and the extra efforts in collecting duplicated geospatial data which saves time and money.

v. Reduces cost:

- SDI allows geospatial data to be collected once and reused multiple times in different applications.

Metadata

- Metadata is data about data.
- It provides information about the data in more details such as what, how, who, when about data.
- Metadata is a simple mechanism to inform others of the existence of the data sets, their purpose and scope.
- Metadata describes the content, quality, condition, limitations and restrictions and other characteristics of data.

Advantage of metadata:

i. It tell others what datasets are available. This helps prevent duplication of efforts, especially among government agencies.

ii. To provide some legal protection in the event of misuse of datasets.

iii. To inform a potential user how the data was collected, for what purpose and if any specific parameters were used. This would allow a potential user to decide whether it would be suitable for his/her purposes.

Clearinghouse

- A (spatial data) clearinghouse is a distributed network of spatial data producers, managers and users that are linked electronically together.
- Clearinghouse can be defined as a service for searching, viewing, transferring, ordering, advertising and disseminating over the internet geo-data stored at many different locations in digital format.
- Clearinghouse uses standards-based Web technology for the publication and discovery of available geospatial resources through the Geospatial Platform portal.
- Clearing housing is a distributed system of agency servers located on the Internet that contain field-level descriptions of available and planned digital spatial data, applications, and services.
- The fundamental goal of clearinghouse is to provide access to digital spatial data and related online services for data access, visualization, or order.
- A clearinghouse allows data providers to register their geographic data sets, the quality of these data and also the instructions for accessing them. Each data provider provides an electronic description of each spatial data set.
- The clearinghouse thus functions as a detailed catalogue service with support for links to spatial data and browsing capabilities.

Different levels of SDI

Global Spatial Data infrastructure (GSDI)

- Global Spatial Data Infrastructure (GSDI) refers to a framework of technologies, policies, standards, and institutional arrangements that facilitate the availability, exchange, and use of geospatial data on a global scale.
- GSDI aims to enable seamless access and sharing of spatial data among users, organizations, and nations.
- The goal of GSDI is to enable better decision-making at local, national, and global levels by providing access to accurate and up-to-date geospatial information.
- The Global Spatial Data Infrastructure Association is dedicated to international cooperation and collaboration in support of local, national and international spatial data infrastructure developments that would allow nations to better address urgent social, economic, and environmental issues.

Different levels of SDI

Global Spatial Data infrastructure (GSDI)

Key components of Global Spatial Data Infrastructure include:

1. Geospatial Data
2. Standards
3. Technologies
4. Policies
5. Institutional Arrangements:

- GSDI often requires the collaboration of multiple organizations, both governmental and non-governmental.
- Institutional arrangements involve setting up bodies or frameworks that coordinate and oversee the implementation of spatial data infrastructure.

6. Capacity Building:

- This involves training individuals and organizations in the collection, management, analysis, and application of geospatial data to ensure effective use of GSDI.

Different levels of SDI

Global Spatial Data infrastructure (GSDI)

Key components of Global Spatial Data Infrastructure include:

7. Access and Distribution:

- This can involve the development of spatial data clearinghouses, web services, and other means of data dissemination to provide open and accessible platforms for the sharing and distribution of geospatial data.

Different levels of SDI

Regional Spatial Data Infrastructure (RSDI)

- Regional Spatial Data Infrastructure (GSDI) refers to a framework of technologies, policies, standards, and institutional arrangements that facilitate the availability, exchange, and use of geospatial data across a specific geographic region.
- A Regional SDI typically involves coordination among multiple organizations, agencies, and stakeholders within a particular region.
- Some of the examples of RSDI are:

i. Asia-Pacific SDI(APSDI), European Geographic Information Infrastructure (EGII)

ii. Permanent Committee on GIS Infrastructure for Asia and the Pacific (PCGIAP)

Purpose of RSDI

- Data Sharing: It facilitates the sharing of spatial data among different entities (government agencies, private organizations, etc.) within the region.
- Decision-Making: It helps in making informed decisions based on spatial information, such as urban planning, environmental management, and disaster response.

Different levels of SDI

National SDI

- The National Spatial Data Infrastructure (NSDI) refers to a framework of policies, standards, and technologies that facilitate the development, sharing, and use of spatial data at the national level.
- Nepal also has its own national SDI in the form of national geoportal but is not in full phase.
- The national geoportal was established with the vision to enhance the public reach and easy access to the authentic Geospatial Dataset provided by the Survey Department, Nepal.

Local SDI

- The National Spatial Data Infrastructure (NSDI) refers to a framework of policies, standards, and technologies that facilitate the development, sharing, and use of spatial data at the local level.
- It is established with the aim of providing spatial data services in local level. It is managed and controlled by local organization.

SDI

Present Situation and Challenges of SDI in Nepal

- In Nepal, SDI has been developed and still a lot of work is to be done in it.
- It is used as a tool for national spatial data collection, storage, processing and dissemination.
- NSDI is the common portal to use geospatial data in effective and efficient way and link with another attribute database.
- Survey Department is the nodal agency for developing NSDI.
- NSDI helps to reduce duplication of efforts, avoid redundant expenditures and eliminate corruption practices.

Some major works done by DOS for SDI in Nepal are:

i. Development of national geoportal where all the spatial data are stored and can be downloaded by user in free or with some charge. Some datsets available are: Land use, Administrative units and other dataset are still to be added.

ii. Land Information System (LIS) is under development. Open Source Based System is also being developed.

iii. Cadastral Information System is being Transformed to Digital System for developing NSDI.

References

- SDI concept: https://informalguru.com/national-spatial-data-structure-nsdi/
- https://gistbok.ucgis.org/bok-slide/spatial-data-infrastructure
- https://eo4geocourses.github.io/KULeuven_Technical-Introduction-to-SDI/index_01_Lecture_General_Introduction_SDI.html#/01.7.0
- https://arxiv.org/ftp/arxiv/papers/1707/1707.03969.pdf